JMIR Research Protocols

Impact Factor (2023): 1.4
Volume 15 (2026) ISSN 1929-0748 Editor in Chief: Xiaomeng (Simone) Ma, PhDc, MS, BS

Contents

Protocols

The Effect of an Educational Intervention on the Knowledge and Self-Care Skills of Adult Patients With Venous Leg Ulcers: Protocol for a Randomized Controlled Trial (e84211)
Cristina Canova, Jessica Scappin, Maria Chiaruttini, Laura Tasson, Flavia Masiero, Claudia Varagnolo, Matteo Martinato. 34

Overcoming the HIV Pre-Exposure Prophylaxis Uptake Barrier With A Health Action Process Approach and Conditional Economic Incentives Among Men Who Have Sex With Men in 3 Chinese Cities: Protocol for a Randomized Controlled Trial (e82123)
Hao Lin, Mojun Ni, Chun Chang, Feng Sun, Jinghua Li, Omar Galárraga, Phoenix Mo, Wangnan Cao. 45

Implementation of a Cancer Navigation Intervention for Newly Diagnosed Survivors of Breast Cancer: Protocol for a Randomized Controlled Trial (e85820)
Melba D'Souza, Ruby Gidda, Michelle Smith, Arati Swaminadhan, Amy Strank, Ashwin Nairy. 58

Clinical Evaluation of Pediatric Pulse Oximeters in South Africa: Protocol for a Cluster Randomized Controlled Trial (e82888)
Holly Schuh, Marieke Van der Zalm, Margaret Van Niekerk, Lise-Marie Laubscher, Abenathi Mcinziba, Steve Gomas, Sunaina Kapoor, Shubhada Hooli, Lario Viljoen, Andre Gie, Bareng Nonyane, Pierre Goussard, Anneke Hesseling, Carina King, Eric McCollum. 70

Impact Assessment of the Allergy Fact Checker, a Clinical Decision Support Tool for Noninvasive Beta-Lactam Antibiotic Allergy Label Delabeling: Protocol for a Multicenter Crossover Cluster-Controlled Study (e86056)
Liesbeth Gilissen, Greet Van De Sijpe, Isabel Spriet, Rik Schrijvers. 83

Development of a Collaborative Education Program for Home Care Nursing Educators' on-the-Job Training in End-of-Life Home Care: Protocol for a Randomized Controlled Trial (e84330)
Ryousuke Yamada, Hanako Numata, Yuka Sumikawa, Kosuke Kashiwabara, Maiko Noguchi-Watanabe, Kayoko Kurita, Noriko Yamamoto-Mitani. 95

The TBI-AD/ADRD Caregiver Support Intervention (TACSI): Protocol of a Pilot Randomized Controlled Trial Evaluation of a Remote Intervention for Family Caregivers (e81125)
Samantha Ostenso, Robyn Birkeland, Elizabeth Albers, Katie Louwagie, Allison Iwan, Sherry Chesak, Joseph Gaugler. 107

Assessing the Implementation and Potential Effects of the Nishauri mHealth Intervention on HIV Care Among Men in Homa Bay County, Kenya: Protocol for a Mixed Methods Study (e85279)
Dan Omollo, Pamela Murnane, Louisa Ndunyu, Dallas Swendeman, Sheri Weiser, Collins Ouma. 129

XSL•FO
RenderX

Virtual Reality–Based Program for Pediatric Patients With Amblyopia: Protocol for a Multicenter, Randomized, Open-Label, Two-Arm Study (e85194)
Ken Nagino, Yuichi Okumura, Masakazu Hirota, Saiko Matsumura, Tadashi Matsumoto, Takashi Negishi, Akie Midorikawa-Inomata, Makiko Ui, Takao Hayashi, Yuichi Hori, Shintaro Nakao, Takenori Inomata. . . . 141

Effect of Helpers Stay Quit Online Training on Preventing Smoking Relapse and Personal Networks: Protocol for a Pragmatic Randomized Controlled Trial and Embedded Mixed Methods Personal Network Study (e82140)
Myra Muramoto, Allison Hopkins, Christopher McCarty, Alicia Allen, L Dickinson, Timothy Connolly, Janet Spradley, Jun Ying. . . . 155

Efficacy of a Virtual Reality Game on Children's Fear and Anxiety During Dental Procedures (VR-TOOTH): Protocol for a Randomized Controlled Trial (e83672)
Julien Gardner, Vallerie Markopoulos, Wenjia Wu, Gabrielle Gilbert, Daphnée Pelletier, Charlotte Fafard, Anne Gagné, Estelle Guingo, Christine Genest, Marie-Ève Asselin, Pascale Ouimet, Kate St-Arneault, Sylvie Le May. . . . 171

The Effect of Evening Technology Use on Objective Sleep in Older Adults: Protocol for a Crossover Randomized Controlled Trial (e84512)
Sarah Ghazi, Anders Behrens, Joakim Niklasson, Johan Sanmartin Berglund, Peter Anderberg. . . . 183

Youth-Centered Mobile Intervention (Next4You) to Promote Healthy Relationships and Sexual Wellness Among Adolescents in or Transitioning From Foster Care: Protocol for a Randomized Controlled Trial (e77185)
Pamela Anderson, Seow Ong, Dallas Elgin, Jennifer Laird, Darriya Starr, Sade Daniels, Jakara Rogers, Karin Coyle. . . . 196

Effectiveness and Cost-Effectiveness of a Digital Falls Prevention Program Versus Usual Care to Improve Balance, Falls Risk, and Function in Older Adults: Protocol for the Keep-On-Keep-Up Randomized Controlled Trial (e78840)
Emma Stanmore, Amelia Parchment, Bolanle Odebiyi, Peter Bower, Chloe French, Chunhu Shi, Saima Bashir, Saima Ahmed, Dawn Dowding, Jo Dumville, Roman Kislov, Alex Thompson, Dawn Skelton, Margaret Clarke, Chris Todd. . . . 210

Effectiveness of a Theory-Based Intervention in Improving Bacillus Calmette-Guérin Uptake and Preventive Practices of Childhood Tuberculosis Among Pregnant Women: Protocol for a Randomized Control Trial (e68088)
Ahmed Dirie, Nor Mohd Zulkefli, Salmiah Md Said, Ahmed Azman, Kassim Jimale, Osman Fiidow. . . . 223

Hand and Arm Bimanual Intensive Therapy Including Lower Extremities (HABIT-ILE@home) for Adults With Chronic Stroke: Protocol for a Randomized Controlled Trial (e87035)
Merlin Somville, Zélie Rosselli, Edouard Ducoffre, Carlyne Arnould, Yannick Bleyenheuft, Geoffroy Saussez. . . . 232

Evaluating the Efficacy of Repetitive Transcranial Magnetic Stimulation Combined With Auditory Integration Training for Children With Autism Spectrum Disorder: Protocol for a Randomized Sham-Controlled Trial (e80243)
Qinghong Hao, Jinying Wang, Jindi Yang, Wei Li, Sufen Hu, Zhihai Lv. . . . 248

Evaluating the Safety and Performance of the KidneYou App for Chronic Kidney Disease: Protocol for an Italian Multicenter, Randomized, Open-Label, Premarket Study (e75306)
Cesira Cafiero, Marco Gnesi, Anna Maurizi, Andreana Foresta, Massimo Beccaria, Giorgia Campilongo, Marco Fiorentino, Paola Acquaviva, Francesca Mastromauro, Marco Gorini, Loreto Gesualdo. . . . 258

Developing and Evaluating a WeChat-Based Applet Fluid Intake Reminder on Enhancing Fluid Adherence in Postoperative Patients With Urinary Calculi: Protocol for a Randomized Controlled Trial (e80214)
Xue Zhou, Tongyao Wang, Qian Chen. . . . 269

XSL•FO
RenderX

PATH Trial for Examining Yoga as a Strategy to Improve Remote-Based Weight Loss in Adults: Protocol for a Randomized Controlled Trial (e78973)
Selene Tobin, Sally Sherman, Beth Bock, Shira Dunsiger, J Thomas, Stephanie Goldstein, Jessica Unick. 284

Comparative Evaluation of AmnioGuard and Advanced Platelet-Rich Fibrin When Combined With NovaBone Putty in the Regeneration of Human Periodontal Infrabony Defects: Protocol for a Randomized Controlled Clinical Trial (e75987)
Geeta Bhandari, Priyanka Jaiswal, Shweta Bhagat, Sakshi Kotecha. 300

Preventing and Treating Insomnia Symptoms in Midlife and Older Adults (ASLEEP): Protocol for a Randomized Controlled Trial Using the PROTECT Norge Infrastructure (e81542)
Jon Aakre, Bjørn Bjorvatn, Martha Gjestsen, Ingvild Dalen, Clive Ballard, Ingelin Testad. 307

Evaluation of a Community-Based AI-Assisted Visual Impairment Screening Model for Performance, Operational Efficiency, Acceptability, Feasibility, and Costs: Protocol for a 2-Arm Pragmatic Randomized Controlled Trial (e74164)
Yibing Chen, Kai Koh, Jin Ho, Samantha Yew, Jocelyn Goh, Elaine Lum, Yih Tham. 319

Strengthening Care for Children Using a Virtual Integrated General Practitioner–Pediatrician Model of Primary Care (SUSTAIN): Protocol for a Stepped Wedge Cluster Randomized Controlled Trial (e69728)
Tammy Meyers Morris, Harriet Hiscock, Karen Wheeler, Corin Miller, Nan Hu, Shukri Hassan Shire, Carmen Crespo-Gonzalez, Susan Bullock, Natalie Taylor, Douglas Boyle, Lena Sanci, Kenny Lawson, Louisa Adams, Ken Peacock, Michael Hodgins, Annemarie Christie, Raghu Lingam.
3 2 9

Clinical and Cost-Effectiveness of Blended Cognitive Behavioral Therapy or Psychodynamic Therapy Versus Face-to-Face Psychotherapy for Depression (BLENDED Study): Protocol for a Pragmatic, Multicenter, Assessor-Blinded Randomized Controlled Noninferiority Trial (e80511)
Patrick Luyten, Danielle Speybrouck, Peter Martin, Stephan Claes, Nick Midgley, Hester van Eeren, Jan van Busschbach, Eileen Tang. 349

Assessment of Blood Glucose Measurement Using New Noninvasive Technology: Protocol and Methodology (e76558)
Eka Suradji, Satvinder Dhaliwal, Zhang Li-Feng, Entong Zhou, Amos Lim, Hui Wong, Yao Du, Fredicia Fredicia, Amanda Dianky, Puspasari Pasaribu, Ernest Masayoshi Takashi, Justin Tan, Mardi Santoso, Seng Ang. 372

Personalized Machine Learning Intervention to Improve Sleep Quality Using Wearable Technology in Healthy Middle-Aged Adults From Mexico City: Protocol for a Pilot Randomized Controlled Trial (e76415)
Rodrigo Quezada Reyes, Luis Trejo. 384

Tailored Individual Follow-Ups Versus a One-Day Group Course in Patients With Long COVID (Post–COVID-19 Condition): Protocol for a Randomized Controlled Trial (e74113)
Marte Wilson, Synne Pedersen, Nina Langeland, Rebecca Cox, Pål Aukrust, Tuva Dahl, Axel Sandvig, Maja Wilhelmsen. 406

Improving Outpatient Psychotherapy for Adults With Major Depressive and Anxiety Disorders Using Web-Based High-Frequency Monitoring and Feedback in Autosystemic Hypnotherapy: Protocol for a Two-Arm ABAB Crossed-Therapist Randomized Clinical Implementation Trial (e78166)
Günter Schiepek, Stephanie Wackernagel. 419

An Online Gamified Pain Management Program for Adults With Chronic Pain: Protocol for a Randomized Controlled Trial (e80542)
Mun Tse, Jiafan He, Tyrone Kwok, Sukki Ho. 432

Evaluation of a Tailored Multifaceted Pharmaceutical Care Intervention to Optimize Chronic Obstructive Pulmonary Disease Management: Protocol for a Cluster Randomized Controlled Trial (e82806)
Xuedi Ma, Xiaocong Li, Xinyi Li, Hui Chen, Yang Wang, Xuan Jia, Yanyan Zhao, Wei Li, Lihong Liu. 447

XSL•FO
RenderX

Impact of the Population Medicine Multimorbidity Intervention in Xishui County (POPMIX) on People at High Risk for Chronic Obstructive Pulmonary Disease Who Experience Mental Health Symptoms: Protocol for the POPMIX-MH Cluster Randomized Controlled Trial (e85853)
Yuhao Liu, Wenjin Chen, Shiyu Zhang, Yingping Wang, Zhoutao Zheng, Ke Huang, Xingyao Tang, Zhong Cao, Xunliang Tong, Lei Tang, Jinghan Zhao, Liu He, Lirui Jiao, Tianying Zhao, Yingchi Luo, Qiande Lai, Xiangqin Lyu, Qiushi Chen, Aditi Bunker, Sebastian Vollmer, Pascal Geldsetzer, Dean Jamison, Till Bärnighausen, Ting Yang, Simiao Chen, Chen Wang, POPMIX Group. 459

Effects of 5-Minute Light-Intensity Physical Activity Breaks (Sit-Cycling and Stand-Twisting) on Cognitive Function, Sleepiness, and Back Pain in Physically Inactive University Students: Protocol for a Mixed-Design Laboratory Study (e82510)
Jiameng Ma, Buan Chua, You Chin, Wei Teo, Hyunshik Kim, Seow Low, Dan Li, Yong Chia. 479

Comparison of Uncemented and Hybrid Hip Arthroplasty: Protocol for a Brazilian Randomized Controlled Trial (e79721)
Bruno Souza, Isabella Lacerda, Matheus Vasconcelos, Nathália Furtado, Julia Vieira, Iasmim Souza, Marco Magalhães, Leonã Amaral, Luiz Assad, Carlos Silva, Carolina Tsen, Valdeci Oliveira. 495

Clinical Efficacy Evaluation of Tuina Combined Medicated Oil in the Treatment of Nonspecific Low Back Pain: Protocol for a Randomized Controlled Trial (e81887)
Man Yuan, Guangxin Guo, Siyu Wang, Fei Yao, Guimao Wang, Hongxi Xu. 507

Effects of Elastic Tape Associated With Pulmonary Rehabilitation in Male Individuals With Chronic Obstructive Pulmonary Disease: Protocol for a 2-Arm, Assessor-Blinded Randomized Controlled Trial (e75029)
Thiago Pinto, Juliana dos Santos, Estéfane Reis, Cibele Marques da Silva, Fabiano de Lima, Regina Carvalho-Pinto, Celso Carvalho. 522

Kidney-Tonifying, Phlegm-Resolving, and Blood Stasis–Removing Therapy for Multiple Myeloma: Protocol for a Randomized Controlled Trial on Epigenetic and Immune Modulation (e86322)
Xiaoqi Sun, Yongming Zhou, Yujue Wang, Youya Dai, Wenwei Zhu, Hailin Chen. 533

Validating a Multibiomarker Panel for the Assessment of Quantity and Quality of Plant Foods in the Diet (PLAENTI): Protocol for a Parallel Group–Designed Randomized Controlled Trial (e77571)
Victor Schmalle, Julia Renz, Paola Ferrario, Stephanie Seifert, Ann Engelbert, Oliver Wittek, Benedikt Merz, Lorraine Brennan, Claudine Manach, Manuela Rist, Achim Bub. 546

Motor Imagery and Action Observation in Breast Cancer Survivors: Protocol for a Randomized Controlled Trial (e85469)
María Arnal-Vallés, Nelia Soto-Ruiz, Ana Bays-Moneo, Cristina García-Vivar, Paula Escalada-Hernández. 563

Effectiveness of Ayurvedic Nutritional Supplements and Yoga Protocol in Reducing the Incidence and Severity of Acute Mountain Sickness (AYAMS Study): Study Protocol for an Open-Label Randomized Controlled Trial (e81567)
Amit Rai, Sophia Jameela, Arun Yadav, Rajneesh Joshi, Pallavi Mundada, Azeem Ahmad, Bidhan Mahajon, Bhagwan Sharma, Shruti Khanduri, Babita Yadav, Arunabh Tripathi, Rakesh Rana, Bhogavalli Chandrasekhararao, Narayanam Srikanth, Raghavendra Rao, Rabinarayan Acharya.
5 7 5

The Effects of Orexin Receptor Antagonism During Early Withdrawal in Stimulant Use Disorder: Protocol for a Proof-of-Concept Study (e83842)
Heather Webber, Nikky Tran, Rita Blake, Sarah Goldman, Trevor George, Bill Bailey, Robert Suchting, Jin Yoon, Jessica Vincent, Michael Weaver, Joy Schmitz, Scott Lane. 589

Spinal Manipulative Therapy on a Healthy Population: Protocol for a Randomized Blinding Feasibility Trial (e85956)
Margaret Sliwka, Tyson Perez, Ahmed Qazi, Stephanie Sullivan. 598

XSL•FO
RenderX

Tunnel Anastomosis vs Double-Tract Jejunal Interposition Reconstruction After Proximal Gastrectomy: Protocol for a Multicenter Prospective Randomized Controlled Trial (e82712)
Qingyu Xie, Rui Peng, Chao Yue, Wei Wei, Lingli Huang, Xiaoxiao Wang, Haitian Wang, Liang Chen, Rongmin Gu, Huanqiu Chen, Xuezhi Ming, Xu Wen, Weiguo Xu, Guangli Sun, Hao Fan, Zhe Wang, Longhao Yang, Xiaohua Zhou, Xiaoyu Wu, Jin Zhou, Meng Wang, Hao Xu, Gang Li. 6 1 7

Impact of Enhanced Family Education on BMI Changes in Children and Adolescents With Overweight or Obesity: Study Protocol for a City-Wide Cluster Randomized Controlled Trial (e86508)
Rui Zhang, Ting Lu, Si Hou, Ya Qu, Qiu Chen, Meng Li, Jiao Wang, Lin Xu. 627

Effects of Sitting Baduanjin on Cancer-Related Fatigue in Patients With Advanced Cancer: Protocol for a Randomized Controlled Trial (e84925)
Qinghui Zhang, Xi Chen, Yinglong Duan, Jianfei Xie, Yongyi Chen, Yonghong Hu, Qinqin Cheng. 638

Treating Chronic Pain by Modulating Phenomenological and Psychophysiological Indices of Self-Transcendence: Single-Blind Randomized Controlled Clinical Trial Protocol (e82362)
Kennedi Childs, Carter Minnick, Geraldine Martorella, Adam Hanley. 650

Online Personalized Normative Feedback to Foster Intention to Change and Help Seeking in Young Adults With Disordered Gambling and Trading Behaviors: Protocol for a Randomized Controlled Trial (e73155)
Ainhoa Coloma-Carmona, José Carballo, Fernando Miró-Llinares, Virtudes Pérez-Jover. 668

Diffusion Tensor Imaging in the Diagnosis of Perianal Abscess: Protocol for a Single-Blind Randomized Controlled Trial (e83449)
Yuting Ma, Pingping Mei, Xiutian Guo, Yan Chen. 679

Urban Care Farming to Enhance Quality of Life Among Older Adults: Protocol for a Waitlist Randomized Trial (e78584)
Cynthia Chen, Jocelin Lam, Zhenye Shi, Katika Akksilp, Su Aw, Mary Chong, Choon Ong, Angelia Sia, Leng Thang, Xin Tham, JunXiang Pong, Kartini Omar, Joan Hung, Marek Kukumberg, Roger Ho, Elizabeth Diehl. 690

Evaluation of the Effectiveness of Injectable Platelet-Rich Fibrin as an Adjuvant to One-Stage Full-Mouth Disinfection in Patients with Stage II and III Periodontitis: Protocol for a Randomized Clinical Trial (e81032)
Mahima Kothekar, Pavan Bajaj, Sneha Dare, Shivani Thakre. 708

Immediate and Sustained Effects of Intensive Equine-Assisted Physiotherapy Based on Neuroproprioceptive “Facilitation and Inhibition” on Psychomotor Development, Clinical Functions, Quality of Life, and Molecular Biological Indicators in Children With Spinal Muscular Atrophy: Protocol for a Crossover Randomized Controlled Trial (e83266)
Katerina Marikova, Jindra Reissigova, Miloslav Vilimek, Marie Cerna, Marketa Pokorna, Kamila Rasova. 718

Efficacy of Acupuncture for Mild to Moderate Depression in Older People: Protocol for a Randomized Controlled Trial (e79327)
Qingnan Fu, Kaihui Xiao, Jie Zhang, Yang Li, Yuxian Wang, Molin Jiang, Zengqi Man, Jing Yang, Wei Lu. 740

Enhancing Upper Secondary Students’ Situational Engagement and Cognitive Prerequisites of Learning Through the Physically Active Academic Lessons Intervention: Protocol for a Mixed Methods Cluster Randomized Individual Crossover Trial (e84601)
Heidi Syväoja, Susanna Takalo, Tuomas Kukko, Nina Salmela, Harto Hakonen, Janne Kulmala, Heidi Lindfors, Hermanni Oksanen, Pekka Räsänen, Kati Mäkitalo, Tuija Tammelin. 752

Effects of Acupuncture-Combined Tuina on Patients With Knee Osteoarthritis: Protocol for a Randomized Controlled Trial (e84082)
Chendong Gu, Dong Wei, Kaiyue Zhang, Jinxi Ren, Yihui Qiao, Fengyu Zhang, Yiting Zhu, Zhixian Xu, Yinlong Cao, Shuyi Yao, Hanrui Guo, Yihang Wang, Huan Wang, Lin Wang, Quanliang Wang, Guangxin Guo. 769

XSL•FO
RenderX

Effect of Du-Moxibustion Combined With Spine-Pinching Therapy on Cognitive Frailty in Older Adults With Prefrailty: Protocol for a Randomized Controlled Trial (e82683)
Suzhen Liu, Mengxia Qi, Hua-Fang Li, Ziqiu Ye, Yu Xu, Xiangying Yang. . . . 786

Efficacy of Tuina for Myopia in Children: Protocol for a Randomized Controlled Trial (e79324)
Shouyao Zhang, Yanju Han, Lijuan Zhang, Fang Tan, Xueping Huang, Can Zhang, Zheng Yao, Lei Xiong, Xiantao Tai. . . . 801

Clinical Efficacy and Safety Assessment of Specific-Mode Electroacupuncture Stimulation Combined With Paclitaxel for Recurrent Malignant Gliomas: Study Protocol for a Single-Arm Trial (e84593)
Zhaoxing Jia, Tianxiang Jiang, Yiqing Zhang, Qianyue Chen, Zhong Di, Qi Yuan, Kecheng Qian, Lin Gan, Congcong Ma, Xianming Lin. . . . 816

Jiedu Xiaozhen Granules for Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor–Mediated Skin Toxicity: Protocol for a Randomized Controlled Trial (e79579)
Shoujiang Hao, Xiaoli Li, Shulan Hao, Xiaoying Zhang, Xiaojun Qi, Gang Jin, Fangfang Shen, Likun Liu. . . . 830

Normalization of Gastrointestinal Symptoms in Adults With Constipation With Daily Green Kiwifruit Consumption: Protocol for an Open-Label Intervention Study (e75286)
Jasjot Maggo, Hwei Ng, Simone Bayer, Catherine Wall, Caroline Hoad, Luca Marciani, Jane Mullaney, Diana Cabrera, Karl Fraser, Janine Cooney, Catrin Günther, Tania Trower, Jeffery Tang, Olivier Gasser, Amber Milan, Warren McNabb, Robin Spiller, Tamlin Conner, Chris Frampton, Meika Foster, Nicole Roy, Richard Gearry. . . . 842

Curative Effect and Autonomic Nerve Function of Patients With Primary Insomnia of Liver Depression and Spleen Deficiency Type Based on the Acupoint Selection of Meridian Theory: Protocol for a Randomized Controlled Trial (e84122)
Su Fu, Kang He, Chuanlong Zhou. . . . 861

The Effect of Virtual Reality–Based Social Cognitive Training for Autistic Adults: Protocol for STEPS (Social Cognitive Training Enhancing Pro-Functional Skills) Randomized Clinical Trial (e72854)
Johannes Andresen, Alberte Jeppesen, Anne Due, Lise Mariegaard, Rizwan Parvaiz, Carsten Hjorthøj, Amy Pinkham, Merete Nordentoft, Gasper Letnar, Jens Jepsen, Louise Glenthøj. . . . 874

The Effects of Multi-Theory Model–Based Behavior Change Intervention with Staircase Approach on Sedentary Lifestyle Among Community-Dwelling Older Adults: Study Protocol for a Randomized Controlled Trial (e81284)
Subinuer Tuerdi, HuiLing Yang, Li He, Rong Yan, YaoYi Cui, XingHui Wang, ShanShan Dong, JiaYu Yang, KeQiang Zhang, Feng Li, YueWei Li. . . . 896

Ayurvedic Therapeutic Regimen as an Add-On to Optimized Conventional Management of Parkinson Disease: Protocol for an Exploratory Randomized Controlled Trial Evaluating Clinical, Cortical Excitability, Neuroimmune, and Autonomic Function Parameters (e83336)
Umesh Chikkanna, Pramod Pal, Sophia Jameela, Rohan Mahale, Nitish Kamble, Venkataram Shivakumar, Bharath Holla, Ravi Yadav, Vikram Holla, Talakad Sathyaprabha, Kaviraja Udupa, Srinibash Sahoo, Manjunath Venkataswamy, Selva Ganapathy, Chandra Rao, Babita Yadav, Sreenithya M, Shivarama Varambally, Sulochana Bhat, Narayanam Srikanth, Kishore Ramakrishna. . . . 918

Effects of Oliceridine Versus Sufentanil on Postoperative Recovery Quality During Hysteroscopy Under Laryngeal Mask Airway Anesthesia: Protocol for a Single-Blind and Randomized Controlled Trial (e84521)
YuHang Shou, JianSheng Luo, XianJie Zhang, WenHu Zhai, Jia Han. . . . 931

Cluster Randomized Controlled Trial to Promote Physical Activity Among Low-Resourced Mothers in New York City: Protocol for the Free Time for Wellness Effectiveness Trial (e71381)
Jessica Watterson, Kate Magsamen-Conrad, Kajal Gokal, Ingrid Oakley-Girvan, Jennifer Hirsch, Min Qian, Richard Buchsbaum, Charlene Niles, Kendra Van Horn, Stephanie Moshier, Andreina Martinez, Avery Garcia Flores, Sherece Laine, Marilisa Vega, Lauren Houghton. . . . 939

XSL•FO
RenderX

Efficacy, Safety, and Economic Impact of Cytisinicline Maintenance Therapy in Patients Who Are Candidates for Smoking Cessation: Protocol for a Phase IV, Multicenter, Randomized, Open-Label, Controlled, Parallel Clinical Trial (CITISILONG Trial) (e76815)
Carlos Rabade Castedo, Ana Estany-Gestal, Carlos Jiménez Ruiz, José de Granda-Orive, Juan Riesco-Miranda, María Cristóbal Fernández, Angela Ramos-Pinedo, Jaime Signes-Costa Miñana, María Gorordo-Unzueta, Agustin Valido-Morales, Jacobo Sellarés -Torres, Eva Cabrera-César, Alejandro Frino-García, Luis Valdés Cuadrado. 953

Effectiveness and Safety of Kangjia Decoction Granules for the Treatment of Hashimoto Thyroiditis: Protocol for a Randomized, Double-Blinded, Placebo-Controlled, Multicenter Clinical Trial (e80993)
Duanchun Zhang, Dan Zhang, Xiaoxiao Qu, Huihong Cao, Yanming He. 969

Real-World Effectiveness and Noninferiority Evaluation and Comparison of Messenger RNA–Based and Protein-Based COVID-19 Vaccines: Protocol for the BEEHIVE Randomized Study With a Hybrid Effectiveness Design (e80858)
Sarang Yoon, German Ellsworth, Steph Battan-Wraith, Andrew Phillips, Rebecca Fink, Joshua Griffin, Elizabeth Rowley, Jacob McKell, Ashley Smith, Riley Campbell, Jesse Williams, Sarah Ball, Hongwei Zhao, Brandy Warren, Matthew Rousculp, Matthew Thiese. 983

Acupuncture Prescription Based on Thermal-Sensitized Acupoints for the Treatment of Knee Osteoarthritis: Protocol for a Randomized Controlled Pilot Trial (e81837)
Xue-Zhou Wang, Tao-Tao Lv, Kua-Yue Zhang, Zhuo-Ya Chen, Li-Na Qin, Wei-Juan Gang, Wei-Heng Chen, Bao-Hong Mi. 995

A Home-Based Intervention to Improve Adherence to the 24-Hour Movement Guidelines in Young Children: Protocol for a Mobile App–Based Randomized Control Trial (e75621)
Chelsea Kracht, Jerica Berge, Monique LeBlanc, Robert Newton Jr, Ryan Rhodes, Grace Bolamperti, Madigan Snodgrass, Leanne Redman.
1 0 0 4

Implementation of a Participatory Ergonomics Intervention to Reduce Musculoskeletal and Stress-Related Mental Health Risks in Australian Retail Workers: Protocol for a Randomized Controlled Trial (e84864)
Elise Condie, Victoria Weale, Katrina Lambert, Jodi Oakman. 1022

Brief Intervention for Discontinuing Inappropriate Z-Hypnotic Use Among Older Patients in Primary Care: Protocol for a Cluster Randomized Controlled Trial With a Single Crossover (e75670)
Tahreem Siddiqui, Maria Bjelkarøy, Tone Simonsen, Maria Selle, Christofer Lundqvist. 1035

Impact of the Population Medicine Multimorbidity Intervention in Xishui County (POPMIX) on People at High Risk for Chronic Obstructive Pulmonary Disease: Protocol for the POPMIX-COPD Cluster Randomized Controlled Trial (e85597)
Chen Wang, Yuhao Liu, Ke Huang, Zhoutao Zheng, Shiyu Zhang, Wenjin Chen, Xingyao Tang, Zhong Cao, Xunliang Tong, Lei Tang, Jinghan Zhao, Liu He, Lirui Jiao, Yingping Wang, Tianying Zhao, Yingchi Luo, Qiande Lai, Xiangqin Lyu, Qiushi Chen, Aditi Bunker, Sebastian Vollmer, Pascal Geldsetzer, Dean Jamison, Till Bärnighausen, Ting Yang, Simiao Chen, POPMIX Group. 1046

Effects of Allulose vs Aspartame Consumption on Postprandial Glucagon-Like Peptide-1 Profiles and Metabolic Health: Protocol for a Randomized, Crossover, Double-Blind, Placebo-Controlled Trial (e81857)
Selina Busch, Paola Ferrario, Ann-Kathrin Henk, Ann Engelbert, Oliver Wittek, Stephanie Seifert, Achim Bub, Carina Mack, Bettina Hieronimus.
1 0 6 5

Impact of Ocular Massage on Intraocular Pressure and Schlemm Canal Dimensions in Healthy Adults: Protocol for a Randomized Controlled Trial (e78864)
Mingxuan Zhang, Luning Qin, Liangzhang Tan, Yongtao Li, Qing Zhang, Shuhan Wang, Zhihui Zhang, Emmanuel Pazo, Xinjun Ren. 1080

Progressive Achilles Loading via Clinician e-Support (PACE): Protocol for a Dual-Site Randomized Controlled Trial (e90955)
Micah Wong, Mathew Frazier, Ryan Jensen, Timothy Fleagle, Barbara Van Gorp, Nicole Hall, Patrick Ten Eyck, Linder Wendt, Molly Pacha, Jessica Danielson, Mederic Hall, Jason Wilken, Ruth Chimenti. 1090

Feasibility of Vigorous Extended Reality Tele-Exergaming for Cardiometabolic Health in Youth With Mobility Disabilities: Protocol for a Case Series Study (e85246)
Byron Lai, Maggie Logan, Raven Young, Ashley Wright, Jordyn Terrell, Larsen Bright, Drew Davis, Christen Mendonca. 1103

XSL•FO
RenderX

Virtual Reality to Improve Pain Management and Mental Health in Stroke Survivors With Chronic Pain: Study Protocol for a Feasibility Randomized Controlled Trial on Virtual Reality-Acceptance and Commitment Therapy (e80611)

Sérgio Carvalho, Paulo Menezes, Catarina Duarte, David Skvarc, Ana Sousa e Silva, Ana Valentim, João Emanuel Diogo, João Sargento-Freitas, Inês Trindade, Paula Castilho, Teresa Lapa, Gerhard Andersson, Miguel Castelo-Branco. 1114

Assessing Associations Between Environmental, Sleep, and Physical Activity Factors and Metabolic Syndrome Risk: Protocol for the FEASible Study (e82034)

Tania Ramos-Santiago, Elise Huglo, Yanrong Li, Christian Corral, Congyu Wu, Kerry Kinney, Darla Castelli, Andreana Haley. 1126

Impact of Intranasal Administration of Ayurveda Medicine in Apparently Healthy Individuals on Neurophysiological Variables and Functional Connectivity Using Functional Magnetic Resonance Imaging: Protocol for an Exploratory Randomized Controlled Trial (e67132)

Devi Nair, Sandya C J, Sophia Jameela, Remya E, Srikanth Moorthy, Harilal Parasuram, Shruti Khanduri, Bhogavalli Rao, Narayanam Srikanth, Rabinarayan Acharya. 1143

An App-Based Behavioral Support Intervention Promoting Physical Activity (APPROACH) in Patients Diagnosed With Breast, Prostate, or Colorectal Cancer: Protocol for a Randomized Controlled Trial (e77096)

Fiona Kennedy, Sue Smith, Rebecca Beeken, Caroline Buck, Leanne Shearsmith, Hannah Truscott, Nicholas Counsell, Chloe Thomas, Anna Roberts, Diana Greenfield, Henry Potts, Nicholas Latimer, Ed Lowther, Lee Smith, Rosie Stevens, Amy Creaser, Jacqui Gath, Lynda Wyld, Abigail Fisher, Phillippa Lally. 1154

Enhancing Cognitive Functions in Older Adults With Mild Cognitive Impairment via Virtual Sail 3D: Protocol for a Feasibility Randomized Controlled Trial (e85089)

Federica Sancassiani, Alessandra Perra, Veronica Vacca, Roberta Montisci, Mariarita Monni, Stefano Lorrai, Giulia Cossu, Donatella Petretto, Lorenzo di Natale, Diego Primavera, Stefania Redolfi, Angelo Scuteri, Antonio Nardi, Goce Kalcev, Massimiliano Pau, Mauro Carta. 1177

A Smartphone-Based Psychological Intervention for Nonsuicidal Self-Injury (Kalmer App): Protocol for a Multicenter Double-Blind Randomized Controlled Trial (e86413)

Anna Julià, Irene Jaén, Mónica Conesa Giménez, Azucena García-Palacios, Juan Pascual, Anna Sintes Estévez, Anaís Lara, Iria Méndez, Soledad Romero, Joaquim Puntí, Joaquim Soler, Jordi Solé-Casals, Marina López-Solà, Daniel Vega. 1189

Evaluating the Effectiveness of the Motivational Interviewing–Based Wellness Coaching for Kids and Families (WC4K) Program in Pediatric Obesity Care: Protocol for a Cluster Randomized Pragmatic Trial (e78792)

Corinna Koebnick, Margo Sidell, Jessica Vallejo, Xia Li, Emerson Delacroix, Deborah Young, Poornima Kunani, Ken Resnicow. 1204

Feasibility and Acceptability of Collaborative Augmented Reality for Older Adults and Companions: Protocol for a Randomized Controlled Trial (e83864)

Nilanjan Sarkar, Judith Tate, Mahrukh Tauseef, Akshith Ullal, Cathy Maxwell, Lisa Juckett, Douglas Scharre, Alai Tan, Rendong Zhang, Zalen Ingram, Lorraine Mion. 1219

The Third Study of Infectious Intestinal Disease (IID3 Study) in the Community: Protocol for UK-Based Prospective Cohort Studies Investigating the Disease Burden (e88759)

Ben Rowland, Vanashree Sexton, Aileen Mill, Stephen Rushton, Roy Sanderson, Caroline Grundy, Simon de Lusignan, Nigel Cunliffe, Daniel Hungerford, Mark Hopkins, Saheer Gharbia, Claire Jenkins, Gauri Godbole, Roberto Vivancos, Alex Elliot, Dominic Mellor, Lesley Larkin, Rachel Chalmers, Sarah O'Brien, IID3 Consortium. 1233

A Culturally Tailored Mobile Health Intervention to Improve Quality of Life in Black Survivors With Prostate Cancer: Protocol for a Stratified Randomized Controlled Trial (e81503)

Gaurav Kumar, Parisa Ghasemi, Adam Alexander, Kathleen Dwyer, Jordan Neil, Perry Cole, Patrick Beckford, Donald Reese, Everett Montgomery, Roland Odeleye, Mark Harris, Yan Zhao, Zsolt Nagykaldi, Andrew McIntosh, Ernie Kaninjing, Mary Young, Sabrina Dickey, Daniel Morton, Opeyemi Bolajoko, Folakemi Odedina, Lourdes Planas, Michael Businelle, Darla Kendzor, Motolani Ogunsanya. 1245

XSL•FO
RenderX

Efficacy of Bacopa monnieri (Linn.) on Cognitive Function and Alterations in Blood Metabolites in Patients With Amnestic Mild Cognitive Impairment and Early Alzheimer Disease: Protocol for an Exploratory Double-Blind, Randomized, Placebo-Controlled Trial (e82891)
Abhilasha Dwivedi, Anjali Anjali, Hina Narzari, Yashwant Kumar, Hanuman Sharma, Aditi Dubey, Nilima Nilima, Roopa Rajan, Mamta Bhushan Singh, Venugopalan Vishnu, Rohit Bhatia, Gautam Sharma, Anu Gupta. . . . 1264

Effectiveness of Low-Dose Atropine Combined With Bright Light Therapy for Controlling Myopic Eye Growth in Schoolchildren: Study Protocol for a Randomized Controlled Trial (e90893)
Dennis Tse, Martin Ma, Ying Hon, Lai Ho, Wai Lam, Kendrick Shih, Xiujuan Zhang, Desmond Lam, Andy Kong, Thomas Lam, Carly Lam, Tina Lian, Ian Morgan, Chi To, Rachel Chun, Christopher Leung. . . . 1280

Effectiveness and Implementation of Adapted Physical Activity Delivery Strategies for Older Adults Living With HIV in Ivory Coast: Protocol for a Type 2 Hybrid Randomized Controlled Trial (e84677)
Nicolas Diadhiou, Nelly Assoumou, Joseph Tegbe, Herve Boa, Marie-Laura Bao, Simon Amador-Paz, Damien Vitiello, Patrick Coffie, Pierre Debeaudrap. . . . 1292

Effects of Performing Eccentric Contractions to Failure After Concentric Muscle Failure in Resistance Training Sessions: Protocol for a Within-Participant Randomized Trial (e67537)
Pedro Campos, Renan Barreto, Gabriel Fontanetti, Leonardo da Silva, Matheus Gomes, Leonardo Lima. . . . 1306

Effectiveness and Safety of a Supplement Containing a Pharmacologically Active Basidiomycete Mushroom for Chronic Fatigue and Post–COVID-19 Fatigue Syndrome: Protocol for a Randomized Controlled Trial (e82633)
Yunqing Xun, Tung Fong, Guang Chen, Yibin Feng, Linda Chan, Ning Wang. . . . 1319

Combining Noninvasive Brain Stimulation and Physiotherapy to Improve the Management of Chronic Low Back Pain in Veterans: Protocol for a Multi-Arm Randomized Controlled Trial (e78952)
Frederique Dupuis, Yannick Tousignant-Laflamme, Pascale Marier Deschênes, Philippe Fournier, Ildephonse Nduwimana, Orlane Ballot, Loris Loisel, Anne Pinard, Luc Lacombe, Pierre Langevin, Alain Gaumond, Jean-Sébastien Roy, Hugo Massé-Alarie. . . . 1329

Effectiveness of an Educational and Counseling Program (the Green Mother Project Phase 2) to Enhance Breastfeeding and Improve Mothers' Diets From an Environmental Perspective: Protocol for a Cluster Randomized Controlled Trial (e80358)
Rosa Cabedo-Ferreiro, Liudmila Liutsko, Judit Cos-Busquets, Rosa García-Sierra, Margalida Colldeforns-Vidal, Azahara Reyes-Lacalle, Pere Torán-Monserrat, M Vicente-Hernández, Miriam Gómez-Masvidal, Concepció Violán, Laura Montero-Pons, Gemma Falguera-Puig, Gemma Cazorla-Ortiz. . . . 1341

Process Evaluation of a Structured Method for Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems (SIOHPS): Protocol for a Convergent Parallel Mixed Methods Study (e89185)
Camilla Göras, Ann-Sofie Ersson, Therese Hellman, Petronella Bjurling-Sjöberg, Gunnar Bergström, Robert Sarkadi Kristiansson, Malin Lohela-Karlsson. . . . 1354

Combined Acupoint Massage and Abdominal Mirabilite Application for Accelerating Gastrointestinal Recovery in Pediatric Patients After Endoscopic Retrograde Cholangiopancreatography: Protocol for a Randomized Controlled Trial (e87961)
Xiaowei Pu, Shan Wang, Chen Cui, Libing Liang, Qunsong Shen, Wei Luan, Caiqin Wu. . . . 1370

Feasibility and Acceptability of Barbershop-Based HIV Prevention Among Heterosexual Men in Kalangala Islands, Uganda: Protocol for a Cluster Randomized Trial (HPTN 111) (e87612)
Caitlin Scoville, Brenda Gati Mirembe, Emily Voldal, Joel Maena, Juliane Etima, Jayla Harris-Wisecarver, Ivan Rukundo, Hassan Ssemere, Doreen Kemigisha, Teopista Nakyanzi, Rita Nakalega, Donaldson Conserve, Maggie Albano, Molly Dyer, Estelle Piwowar-Manning, Hans Spiegel, Clemensia Nakabiito, Andrew Mujugira, Deborah Donnell, Zubair Lukyamuzi. . . . 1382

A Digital Diabetes Prevention Program for Hispanic Adolescents (Fit24+): Protocol for a Feasibility Randomized Controlled Trial (e75331)
Erica Soltero, Marbelly Partida, Sandra Mihail, Lizette Villanueva, Callie Lopez, Salma Musaad, Teresia O'Connor, Maria Redondo, Matthew Buman, Debbe Thompson. . . . 1400

XSL•FO
RenderX

Evaluating the Role of Surgical Telementoring in the Acquisition of Surgical Skills in Laparoscopic Cholecystectomy: Protocol for a Pilot Randomized Controlled Trial (e73159)
Khayam Butt, Laurent Trichet, Gunnar Hartvigsen, Ole Sjo, Bjørn Edwin, Knut-Magne Augestad. 1412

Digital Attention Bias in Cancer Survivors Intervention for Adolescent and Young Adult Cancer Survivors: Protocol for a Pilot Randomized Controlled Trial (e82665)
Nancy Lau, Chuan Zhou, Shannon Hong, Homer Aalfs, Shannon Higgins, Isabel Badillo, Elizabeth McCauley, Tyler Ketterl, Eric Chow, Jesse Fann, Lauren Heathcote, Tonya Palermo. 1431

A Supportive Group Intervention for Caregivers to Patients Diagnosed With Glioblastoma: Protocol for the SUGRI Study (e81948)
Sara Nordentoft, Rikke Guldager, Helle Pappot, Tiit Mathiesen, Anders Tolver, Karin Piil. 1446

Evaluation of the Comparative Efficacy of an Ayurvedic Formulation (Nimba-Amalakyadi Powder) vs Metformin in the Management of Type 2 Diabetes Mellitus: Protocol for a Pilot Study (e63574)
Punam Sawarkar, Gaurav Sawarkar. 1458

Evaluating a Social Robot Reception Service in an Italian Geriatric Hospital (HOSPER Study): Protocol for a Usability Study (e81395)
Giulio Amabili, Arianna Sgolastra, Arianna Margaritini, Federico Barbarossa, Elvira Maranesi, Giacomo Cucchieri, Lidia Pascucci, Letizia Ferrara, Roberta Bevilacqua. 1466

Transcutaneous Electrical Nerve Stimulation vs Transcutaneous Electrical Nerve Stimulation With Music in Persistent Low-Back Pain: Protocol for a Pilot Feasibility Trial (e82382)
Seema Radhakrishnan, Michael Glynn, Anne-Marie Robertson, Anna Fedyczkowski, Lieu Trinh, Sabrina Naz. 1479

Adaptation of a Protocol to Evaluate Online Recruitment Strategies for Mailing HIV Self-Tests and Pilot an Intervention to Improve Linkage to HIV Prevention and Care Among Transgender Women: Protocol for a Randomized Controlled Trial (e75345)
Joanna Caldwell, Ruth Dana, Irah Lucas, Patrick Sullivan, Evelyn Olansky, Stephen Sullivan, Marissa Hannah, Travis Sanchez, Lisa Hightow-Weidman, Akshay Sharma. 1488

Therapeutic Improvement in People With Schizophrenia Undergoing tACS/CBTp (Transcranial Alternating Current Stimulation/Cognitive Behavioral Therapy for Psychosis) Associated With Usual Medication Regimen: Protocol for a Pilot, Randomized, Single-Blind Trial (e80593)
Felicia Iftene, Adriana Farcas, Simon O'Brien, Christopher Bowie, Michael Best, Oyedeji Ayonrinde, Terry Landry, Jennifer Carlson, Scott Davidson, Ellie Rodgerson, Ashley Theis. 1501

Flopropione, a Cysteine Conjugate β-Lyase 1 Inhibitor, for Prevention of Cisplatin-Induced Nephrotoxicity: Protocol for a Randomized, Open-Label, Proof-of-Concept Phase 1 and 2a Trial (e87907)
Takenao Koseki, Masashi Kondo, Hidetsugu Fujigaki, Kayoko Kikuchi, Yuko Oya, Hiroshi Kato, Tomohiro Mizuno, Naotake Tsuboi, Kenji Kawada, Yasuhiro Goto, Naozumi Hashimoto, Kazuyoshi Imaizumi, Akiko Kada, Hikaru Yabuuchi, Kuniaki Saito, Hideyuki Saya. 1519

Acupuncture Therapies for Chemotherapy-Induced Nausea and Vomiting in Patients With Breast Cancer: Protocol for a Systematic Review and Network Meta-Analysis (e86384)
Yutao Luo, Zhujun Bian, Qi Wang, Tiantian Lei, Hong Zhao. 1529

Efficacy of Nutritional Interventions in Adult Poststroke Rehabilitation: Protocol for a Systematic Review and Trial Sequential Meta-Analysis (e79478)
Chen Chen, Su-xiang Zhang, Shao-hua Chen, Yan Cao, Ji-ming Tao. 1540

Impact of Antenatal Corticosteroids on Brain Function and Underlying Mechanisms in Preclinical Studies: Protocol for a Systematic Review (e84522)
Mauritz Herselman, Tara Hislop, Sebastian McBride, Mitchell Lock, Janna Morrison, Kathryn Gatford. 1549

XSL•FO
RenderX

Determining the Impact of Nonionizing Electromagnetic Waves on Human Pregnancy and Teratogenicity: Protocol for Systematic Review and Meta-Analysis (e86479)
Desy Armalina, Neni Susilaningsih, Heri Sutanto, Sunarno Sunarno. . . . 1561

Exploring Components and Feasibility of Health Coaching Interventions for Self-Management of Type 2 Diabetes: Protocol for a Systematic Review (e71383)
Rija Mir, Shoba Poduval, Jessica Sheringham, Fiona Hamilton. . . . 1571

Efficacy and Safety of Massage Therapy for Cognitive Impairment in Patients With Autism Spectrum Disorder: Protocol for a Systematic Review and Meta-Analysis (e88284)
Yi Zhong, Sijie Dang, Xiaoqiu Wang, Xin Chia, Jing Zhang, Weijiang Li, Chongjie Yao, Zhengyu Li, Bin Xiao. . . . 1582

Impact of Physiotherapy Intervention on Pain, Quality of Life, and Function in Low Back Pain Associated With Piriformis Syndrome: Protocol for Systematic Review (e72350)
Nikita Deshmukh. . . . 1595

The Role of Metabolomics in Understanding Hypertension: Protocol for a Systematic Review and Meta-Analysis (e77536)
Mohd Mohd Nawi, Puteri Megat Kamaruddin. . . . 1602

Effects of Cash Transfer and Incentive Programs on Service Utilization and Treatment Outcomes Related to Neglected Tropical Diseases and Their Impact on Health and Nutrition in Low- and Middle-Income Countries: Protocol for a Systematic Review (e76450)
S Hasan, Radhika Dayal, Amena Nishan, Sunetra Ghatak, Puja Chakraborty, Sudeshna Maitra, Sumaiya Raisa, Nizamuddin Khan, Dinesh Mondal, Tahmeed Ahmed, Avishek Hazra. . . . 1611

Enhancing Fracture Resistance by Customizing Glass Fiber Posts in Endodontically Treated Teeth: Protocol for a Systematic Review (e76027)
Sanika Damle, Manoj Chandak, Unnati Shirbhate, Pavan Bajaj. . . . 1629

Diabetic Dyslipidemia and Its Determinants Among People With Diabetes in South Africa: Protocol for a Systematic Review and Meta-Analysis (e82716)
Mashudu Nemukula, Fentahun Adane, Arun Malaisamy, Neel Bhavesh, Olebogeng Majane, Sechene Gololo. . . . 1635

The Effectiveness of Rehabilitation Interventions for Improving Leisure Participation Following Stroke: Protocol for a Systematic Review (e71353)
Serena Alves-Stein, Natasha Lannin, Kylie Wales, Sharon Kramer, Laura Jolliffe. . . . 1643

Prevalence of Problematic Digital Media Use Among Young Adults: Protocol for a Systematic Review (e82245)
Rachelle Chanmany Pastor, Fayette Nguyen Truax, Safiye Sahin, Janet Donnelly, Jan Nick, Shanalee Tamares. . . . 1650

Determinants Associated With Pesticide Exposure in Patients With Head and Neck Cancer: Protocol for a Systematic Review (e81312)
Gustavo Sánchez-Ramírez, Manuel López Cabanillas Lomelí, Blanca González-Martínez, Ana Castro-Sánchez, Myriam De la Garza-Ramos, Guillermo Cano-Verdugo. . . . 1660

Frameworks for Guiding the Selection of Digital Data Collection Tools Used in Clinical Trials: Protocol for a Systematic Review (e78529)
Bavatharani Ananthesayanan, Kayode Fadahunsi, Huanhuan Xiong, John O'Donoghue. . . . 1668

Association and Agreement of Contact-Based Smartphone Photoplethysmography Compared With Electrocardiography: Protocol for a Systematic Review and Meta-Analysis (e84837)
James Mather, Nicholas Sculthorpe, Ethan Berry, Jacqueline Mair, Nilihan Sanal-Hayes, Lawrence Hayes. . . . 1675

XSL•FO
RenderX

Challenges Faced by Parents in Preventing Online Child Sexual Exploitation and Abuse: Protocol for a Systematic Review (e80968)
Neelam Punjani, Wenting Yan, Lisa Hartling, Amber Hussain, Farah Elgaweesh, Shannon Scott. 1684

Donor Risk Factors Affecting Graft Survival in Pediatric Kidney Transplants: Protocol for a Systematic Review and Meta-Analysis (e71620)
Cahyani Ambarsari, Jon Kim, Nur Maamor, Izzah Rosli, Nor Muhamad. 1692

Evidence for Digital Mental Health Assessment Tools in the Post–COVID-19 Era: Protocol for a Systematic Review on Diagnostic Accuracy Across Age Groups (e73492)
Kathryn Babbitt, Erin Funnell, Nayra Martin-Key, Eleanor Barker, Sabine Bahn. 1700

Digital Self-Guided Mental Health Interventions to Prevent Workplace Burnout and Enhance Psychological Wellness: Protocol for a Systematic Review (e80417)
Ehsan Etezad, John Fiset, Raghid Al Hajj. 1709

Effect of Multimodal Interventions on Pain, Stress, and Neurodevelopmental Outcomes in Neonates With Malignant and Benign Tumors: Protocol for a Systematic Review (e77969)
Sharath Hullumani, Shrutika Sharad, Aishik Ghosh, Rutuja Sawalkar, Irshad Qureshi. 1720

The Lived Experiences of Children Who Have Undergone Kidney Replacement Therapy and Their Families: Protocol for a Qualitative Systematic Review (e77725)
Kenji Takao, Maki Fujitsuka, Shingo Ueki, Tomoyuki Sakai. 1726

Quality in Use in Connected Mental Health: Protocol for a Systematic Mapping Study (e79611)
Shweta Premanandan, Sofia Ouhbi, Magdalena Stadin, Charlotte Blease, Åsa Cajander, Maria Hägglund. 1736

Effectiveness and Theoretical Foundations of mHealth Apps for Physical Activity, Healthy Eating, and Weight Loss: Protocol for a Systematic Review and Meta-Analysis (e72664)
Victoria Riccalton, Madison Milne-Ives, Rosiered Brownson-Smith, Ananya Ananthakrishnan, Cen Cong, Emily Oliver, Edward Meinert. 1746

Investigating Placebos and Controls Used in Large Language Model–Based Chatbot Intervention Trials: Protocol for a Methodological Review (e90507)
Leo Druart, Vanda Faria, Marco Annoni, John Torous, Moa Pontén, Charlotte Blease. 1756

Mathematical Prediction Models for Sentinel Node Status in Early-Stage Breast Cancer: Protocol for a Systematic Review (e82523)
Justin James, Kirti Mehta, Mohammadali Ahmadipour, Tegan Ormston, Emily Schembri, Michael Law, Shomik Sengupta, Christobel Saunders.
1 7 6 6

Longitudinal Associations Between Biomarkers and Frailty in Older Adults: Protocol for a Scoping Review (e83312)
Jo Seok, Kyusol Shim, Yihan Yang, Soomin Hong. 1776

Inside-Out Mechanism of Referred Visceral Pain: Protocol for a Systematic Review of Cutaneous Neurogenic Inflammation in Preclinical Animal Models of Visceral Diseases (e67852)
EunMee Yang, Dennis Muñoz-Vergara, Andrew Ahn, Weidong Lu, Vitaly Napadow, Alexandria Cronin, Peter Wayne. 1784

The Role of Digital Media in Chronic Disease Self-Management: Protocol for a Multimethod Study of the DISELMA Research Consortium (e77811)
Constanze Rossmann, Veronika Karnowski, Julia Metag, Juliana Raupp, Doreen Reifegerste, Claudia Riesmeyer, Nariman Sawalha, Alexandra Lux, Anna-Lena Esser, Rebecca Kammerer, Franca Singh, Natalie Rödel, Janine Brill, Elisabeth Gerling, Annemarie Wiedicke. 1792

XSL•FO
RenderX

Usability and Cultural Relevance of an mHealth App for Hispanic/Latino Individuals Living With Rheumatoid Arthritis: Protocol for a Mixed Methods Study (e88401)
Thais Alves, Ronnie Horner, Marie Montas, Melanie Cozad. 1817

Cocreation of a Digital Tool for Proactive End-of-Life Communication: Protocol for a Participatory Action Research Project (e88452)
Malin Eneslätt, Charlèss Dupont, Eva Savolainen, Petter Fjällström, Ida Goliath. 1826

Food Insecurity Determinants, Coping Strategies, and Association With Nutritional Status Among Hemodialysis Patients in Pahang, Malaysia: Protocol for a Mixed Methods Study (e84575)
Nor Mohd Shukri, Sarah Karami, Wan Mohd Abu Bakar, Roselawati Mat Ya, Norhasmah Sulaiman. 1839

Environmental Exposure and Mental Health in Hong Kong: Protocol for a GPS- and Biosensor-Based Observational Study (e84919)
Yanjia Cao, Tianyu Li, Calvin Tribby, Dorita Chang. 1853

Promoting Mental Health, Physical Activity, and Healthy and Sustainable Dietary Behavior Among Practical Education Students in the Netherlands: Protocol for a Multiphase Participatory Research Study (e84723)
Daniëlla van Uden, Annemarie Wagemakers, Margot Peeters, Marloes Kleinjan, Kirsten Verkooijen. 1867

Integrating Physical Health Dimensions in Middle-Aged and Older Adults to Predict Drug Misuse and Relapse Through a Comprehensive Needs Model: Protocol for a Mixed Methods Study (e88987)
Ema Madyaningrum, Eny Purwandary, Supriyati Supriyati, Ariani Pertiwi, Made Gautama, Nurlaela Widyarini. 1877

Understanding the Factors Explaining the Growing Use of Medical Assistance in Dying in Québec: Protocol for an Interdisciplinary Mixed Methods and Multimethods Study (e83549)
Marie-Ève Bouthillier, Isabelle Marcoux, Catherine Perron, Bruno Gagnon, David Lussier, Ghislaine Rouly, Mathieu Moreau, Michel Dorval, Sabrina Lessard, Dominique Girard, Gina Bravo, Maude Hébert, Michelle Giroux, Simon Lemyre, Alexandra Beaudin, Claude Bourque, Maryse Soulières, Maude Lévesque, Mona Gupta, Valérie Bourgeois-Guérin, David Lavoie, Bertrand Lavoie, Ariane Plaisance, Louise Bernier, Jacinthe Dupuis.
1 8 8 4

Designing a Digital Twin for the Management of Noncommunicable Diseases: Protocol for a Pilot Study and Methodology Validation (e75934)
Edgar Ross, Jason Ross, Patricia Beschler, David Guydan, Robert Jamison. 1903

Using AI Algorithms and Machine Learning in the Analysis of a Bio-Purification Method (Therapeutic Emesis, Known as “Vamana Karma”): Protocol for a Mixed Methods Study (e79875)
Pooja Rani, Sumit Kalra, Sachin Singh, Richard David, Ashutosh Gupta, Anandaraman P V. 1925

Strengthening Health Care Professionals’ Collaborative Responses to Women Experiencing Intimate Partner Violence in Pregnancy: Protocol for an Exploratory Mixed Methods Study (e86289)
Carly Jones, Belinda Lovell, Tracy Humphrey, Angela Brown. 1934

Etiologic Types and Complications of Diabetes Mellitus in Newly Diagnosed Patients at Health Institutions in Bulawayo, Zimbabwe: Protocol for a Cross-Sectional and Prospective Observational Study (e74186)
Rudo Gwini, Elopy Sibanda, Desmond Mwembe, Fraser Pirie, Ayesha Motala. 1949

Specialists’ Perceptions of Workforce Retention Strategies in the Malaysian Ministry of Health and Their Association With Job Satisfaction and Turnover Intention: Protocol for a REDCap-Based National Cross-Sectional Survey (e83377)
Siti Ahmad Subki, Ili Abdullah Sharin, Nor Zakaria, Nursyahda Zakaria, Pangie Bakit, Izzuan Adnan, Norehan Jinah, Kun Lee. 1965

Assessing the Effectiveness of an Awareness Program on Knowledge and Practice Regarding the Mothers’ Absolute Affection Program Among Postnatal Mothers in Selected Hospitals: Protocol for a Quasi-Experimental Study (e73496)
Vaishnavi Dongare, Kavita Gomase, Vaishali Taksande. 1978

XSL•FO
RenderX

Development of a Decision Aid to Support Shared Decision-Making on Cannabis Use for Arthritis: Protocol for a Multiphase Study (e76237)
Heba Aref, Yazid Al Hamarneh, Tony Kiang, Elaine Yacyshyn, Cheryl Sadowski. . . . 1983

Identifying Predictors of Gambling Episodes and Craving Using Ecological Momentary Assessment and Smartwatch-Based Physiological Measures: Protocol for a Longitudinal Observational Study (e82782)
Andreas Meyer, Esther Kim, Hae Lee, Gwanghyun Jo, Michael Schaub, Severin Haug. . . . 1997

Implementing an eHealth Model of Care for Pediatric Patients and Families at the End of Treatment for Acute Lymphoblastic Leukemia (EMERGE): Type 2 Nonrandomized Hybrid Implementation-Effectiveness Trial Study Protocol (e85901)
Maria McCarthy, Chris Williams, Michelle Tennant, Richard De Abreu Lourenco, Hannah Pring, Katie Moore, Jane Templeton, Ken Knight, Peter Downie, Stephen Hearps, Cinzia De Luca. . . . 2014

Acceptability of the Cardiff Online Cognitive Assessment for Clinical Screening of Patients With Psychosis: Protocol for a Mixed Methods Study (e84218)
Amy Lynham, Anthony Cope, Sarah Milosevic, Ian Jones, James Walters. . . . 2030

Using Implementation Science to Improve Health Care Access and Quality for People With Traumatic Brain Injury–Related Morbidity (I-HEAL): Protocol for a Translational Multiproject Program Award (e79738)
Risa Nakase-Richardson, Jolie Haun, Megan Moore, Jennifer Bogner, Jeanne Hoffman, Kristen Dams-O'Connor, Jill Coulter, Tracy Kretzmer, Aaron Martin, Marc Silva, I-HEAL Consortium. . . . 2043

Enhancing Caregivers' Quality of Life Through a Web-Based Person-Centered Solution (TechQoL4Carers): Protocol for a Mixed Methods Pilot Trial (e86602)
Betania Groba, Patricia Concheiro-Moscoso, María Miranda-Duro, Manuel Lagos-Rodríguez, Javier Pereira, Laura Nieto-Riveiro. . . . 2086

Using Alcohol and Sleep Sensors to Understand Blackout Risk in Young Adults' Natural Settings (The Lights Out Study): Protocol for an Intensive Longitudinal Pilot Study (e83980)
Veronica Richards, Ashlea Braun, Michael Sladek, Junru Zhao, Thad Leffingwell, Sydney Newell Chesebro, Julie Croff. . . . 2100

A Virtual Living Lab Platform Codeveloped for Mental Health in Youth-Onset Type 2 Diabetes (BrightSpark Care Lab): Protocol for a Mixed Methods Study (e83865)
Mandy Archibald, Brandy Wicklow, Elizabeth Sellers, Arlene Griffiths, Linda Diffey, Jonathan McGavock, Leslie Roos, Alexander Clark, Jennifer Lopez, Josephine Ho, Ernestine Ledoux, Gifty Dzorka, Shahina Parvin, Oluwatoyosi Fagbuyi, Allison Dart. . . . 2112

Traditional Chinese Medicine Syndrome Differentiation of Adult Patients With Type 2 Diabetes and Metabolic Syndrome: Protocol for a Cross-Sectional Study (e86217)
Jialing Zhang, Hoi Wong, Zhilin Lin, Shuyan Zhong, Minxia Ma, Xuejiao Wang. . . . 2127

Evaluating an AI-Enabled Mobile Mental Health Monitoring Tool Among Family Caregivers of Adults Living With Cancer: Single-Arm Feasibility and Acceptability Trial Protocol (e83276)
Chiara Acquati, Michael Aratow, Tahmida Nazreen, Arunima Bhattacharjee, Isabella Marra, Ashley Alexander. . . . 2136

Effectiveness and Cost-Effectiveness of Emergency Department–Based Violence Intervention Programs in the United Kingdom: Protocol for a Quasi-Experimental Study (e86247)
Simon Moore, Sinead Brophy, Amrita Bandyopadhyay, Annemarie Newbury, Megan Hamilton, Adele Battaglia, Trudy Lowe, David O'Reilly, David Rawlinson, Lara Snowdon, Jonathan Shepherd, Vaseekaran Sivarajasingam, Alan Watkins, Simon Walker, Shainur Premji, Sophie Borgia, Henry Yeomans. . . . 2152

Targeted Real-Time Assessment of Chronic Pain (TRAC-Pain) in Youth: Protocol for a Digital Biosignature Development Through a Prospective Observational Cohort Study (e84781)
Anna Bailes, Chi-Hung Shu, Alan Chang, Sahrish Masood, Nicole Jehl, Aliyah Davis, Jeremy Giberson, Casey Cashman, Allison Hill, Javed Gill, Ryan McGinnis, Ellen McGinnis, Nima Aghaeepour, Laura Simons. . . . 2170

XSL•FO
RenderX

Combating Dermatologic Misinformation on Social Media via a Training Program for Health Care Professionals: Protocol for a 2-Phase Pilot Mixed Methods Study (e81003)
Aliyyat Afolabi, Elijah Brown, Joanne Lyu, Pamela Ling, Andrea Rustad, Anna Davis, Beth Hoffman, Melissa Pugliano-Mauro. 2184

Impact of a Prehospital Chest Pain Alert App–Mediated Prehospital–in-Hospital Coordination Model on Treatment Delays and Clinical Outcomes in Patients With ST-Elevation Myocardial Infarction: Protocol for a 4-Year Retrospective Real-World Cohort Study (e90144)
Shuyuan Chen, Rensong Wang, Jiangwei Ma, Weichun Xiao, Changxin Song. 2193

Enhancing Continuous Medication Safety Through e-Prescription and Clinical Decision Support Systems in Outpatient Practices and Pharmacies: Protocol for a Multiperspective Study (eRIKA Study) (e87277)
Juliana Schmidt, David Lampe, Adriana Poppe, Ingo Meyer, Sara Söling, Juliane Köberlein-Neu, Daniel Grandt, Lara Düvel, Wolfgang Greiner, eRIKA Study Group. 2204

Remote Management of Patients With Diabetic Macular Edema Receiving Long-Term Intravitreal Anti–Vascular Endothelial Growth Factor Therapy (RE-SHINE Study): Protocol for a Real-World Study (e85764)
Shu Li, Yupeng Xu, Yan Xu, Xun Xu. 2219

WhatsApp-Led mHealth Intervention for Hypertension Management Among Adults Aged 40 to 59 Years in Kerala, India: Protocol for a Mixed Methods Study (e81307)
Karthika Maniyara, Prakash Kodali. 2229

Temporal Associations Between Body Checking and Eating Pathology in Adolescent Girls with Binge-Spectrum Eating Disorders: Protocol for an Ecological Momentary Assessment Study (e73447)
Jannah Moussaoui, Stephanie Manasse, Laura D'Adamo. 2244

Wellness4Students Program to Mitigate Stress, Anxiety, and Depression Among Makerere University Students in Uganda: Protocol for a Longitudinal Study (e84637)
Stephen Wandera, Olivia Nankinga, Patricia Ndugga, Christabellah Namugenyi, Claire Ashaba, Noeline Nakasujja, Rosco Kasujja, Belinda Agyapong, Vincent Agyapong. 2253

The Digital Information Environment of Lung Cancer and Lung Cancer Screening: Protocol for a Cross-Platform Social Media Content Analysis (e89479)
Lisa Carter-Bawa, Ana Vielma, Gia Nealy, Diya Vemuganti, Nidhi Patel. 2263

Development and Validation of Machine Learning Models for Predicting Early Cognitive Decline Using Home Sensor–Derived Behavioral Data: Sensors in-Home for Elder Wellbeing (SINEW) Cohort Study (e79490)
James Moon, Khalid Abdul Jabbar, Tony Tay, Laura Tay, Rathi Mahendran, Tze Ng, Shian Tan, Wilbur Koh, Weng Ying, Ah-Hwee Tan, Tih-Shih Lee, Iris Rawtaer. 2274

Longitudinal and Multimodal Recording System to Capture Real-World Patient-Clinician Conversations for AI and Encounter Research: Protocol for an Observational Study (e84688)
Misk Al Zahidy, Kerly Guevara Maldonado, Luis Vilatuna Andrango, Ana Proano, Ana Claros, Maria Lizarazo Jimenez, Esteban Gomez-Alvarez, David Toro-Tobon, Victor Montori, Oscar Ponce-Ponte, Juan Brito. 2294

Development and Formative Usability Evaluation of a Theory-Driven Progressive Web Application for Young Adult Wellness Engagement (MiCARE): Protocol for a Mixed Methods Study (e86515)
Ayesha Thanthrige, Nilmini Wickramasinghe. 2306

Real-Time and Long-Term Effects of Medical Marijuana on Older Adults: Protocol for a Prospective Cohort Study (e78900)
Kendall Robinson, Stella Seeger, Lauren Nave, Marlin Mejia, Maria Vander Meulen, Ahmed Rashid, Angela Mickle, Kimberly Sibille, Zhigang Li, Rene Przkora, Siegfried Schmidt, Margaret Lo, Robert Cook, Yan Wang. 2318

XSL•FO
RenderX

Follow-Up of Low-Acuity Patients After Redirection From a Swiss Emergency Department Using an Electronic Triage Application: Protocol for a Single-Center Prospective Cohort Study (e82195)
Flora Gobet, Anne-Laure Feral-Pierssens, Thomas Castelain, Ludovic Galofaro, Hugo Najberg, Vincent Ribordy, Youcef Guechi. 2335

Evaluation of Reduced Single-Photon Emission Computed Tomography Imaging Protocols and Software Variability in 177Lu-DOTATATE Dosimetry: Protocol for an Exploratory Observational Trial (e83248)
Takayuki Yagihashi, Kenta Miwa, Noriaki Miyaji, Satoru Sugimoto, Hideki Hayakawa, Noritoshi Kobayashi, Shoko Takano, Taro Murai. 2346

Toward Best Practice Guidelines and Curricula for Virtual Care (Telehealth) in a Cancer Center: Protocol for a Multimethod Study (e81768)
Sharon Brownie, Lauren Parkinson-Zarb, Andrew Dimech, Maria Ftanou, Patrick Broman. 2355

Surgical Outcome of Iris Claw and Scleral Tuck Intraocular Lens in Primary Cataract Surgery by Phacoemulsification in a Tertiary Care Center: Protocol for a Randomized Controlled Trial (e80943)
Mohammed Kaderi, Sachin Daigavane. 2369

The Effect of the DemensiaKITA Mobile Health App Intervention on Knowledge, Attitude, Practice, and Burden Level of Dementia Caregivers in Kuala Lumpur and Selangor, Malaysia: Protocol for a Nonrandomized Controlled Trial (e83699)
Nurulizzah Mahfar, Nik Abdullah, Dalila Roslan, Khalid Ibrahim, Ungku Ungku Mohd Zam, Noraliza Noordin Merican, Xin Chen, Mariam Mohamad. 2376

Acoustic Emission Biomarkers for the Detection and Monitoring of Early Knee Osteoarthritis: Protocol for a Prospective, Single-Center, Exploratory Study (e67032)
Louis Leuthard, Jérôme Thevenot, Tomás Teijeiro, David Atienza, Vincent Stadelmann. 2392

AI Models to Reduce Surgical Complications Through Intraoperative Video Analysis: Protocol for a Prospective Cohort Study (e62734)
António Sampaio Soares, Sophia Bano, Laura Castro, Margarida Pascoal, Ricardo Rocha, Paulo Alves, Paulo Mira, Joao Costa, Manish Chand, Danail Stoyanov, Catarina Barata. 2406

Digitally Enhanced Routine Outcome Monitoring in Italian Psychotherapy: Protocol for a Longitudinal Implementation Study (e82837)
Mattia Olive, Antonino La Tona, Vivian Ostwald, Agostino Brugnera, Gianluca Lo Coco, Angelo Compare, Dana Tzur Bitan, Cristina Masella. .
2 4 1 5

Real-Time Type 1 Diabetes Self-Management Decision-Making in Adolescents: Protocol for a Longitudinal Mixed Methods Study Using Text Messaging and Continuous Glucose Monitoring (e83218)
Melissa DeJonckheere, Samantha Chuisano, Juniar Lucien, Fouzaan Amjad, Oorvi Duvvuri, Hasan Khan, Maryam Khan, Rafee Mirza, Maya Ollivierre, Timothy Guetterman, Yu Lin, Lorraine Buis, James Aikens, Caroline Richardson, Joyce Lee. 2427

Digital Dissemination Strategies for Health Promotion Videos in Indigenous Communities in California: Protocol for a Three-Arm Comparative Study (e84674)
Lucía Abascal Miguel, Anna Epperson, Alison Comfort, Darío León, Mary Garcia, Alicia Riley, Nadia Diamond-Smith. 2443

The Development and Protocol for Testing a Co-Created Digital Intervention (Sentinel) to Improve Mental Well-Being and Help Manage and Prevent Trauma in First Responders (e72250)
Nicola Cogan, Alison Kirk, Christoph Graf. 2450

Utilization and Feasibility of a Wearable Device in Patients With Sedative Effects of Drugs: Protocol for a Prospective Observational Study for the Advanced Respiratory Monitoring Events in Drug Toxicity (ARM-ED) Study (e66961)
Lisa Dunlop, Bruce Henderson, Osian Meredith, Chris Trueman, Christopher Carlin, Robert Docking, David Lowe. 2465

XSL•FO
RenderX

Using Smartphone-Based Digital Phenotyping to Predict Relapse in Serious Mental Disorders Among Slum Residents in Dhaka, Bangladesh: Protocol for a Machine Learning Study (e79826)
Nadia Alam, Chayon Das, Neelabja Roy, Domenico Giacco, Swaran Singh, Sagar Jilka. . . . 2478

Evaluating Electroencephalogram-Based Predictive Model for Drowsiness Measurement to Reduce Accident Risk in Active Individuals: Protocol for a Preliminary Monocentric Study (e83969)
Chloé Boitard, Zoé Mazurie, Khadijeh Sadatnejad, Julien Coelho, Patricia Sagaspe, Julie Lenoir, Julien Mattei, Pierre Berthomier, Marie Brandewinder, Pierre Philip, Jean-Arthur Micoulaud Franchi, Christian Berthomier, Jacques Taillard. . . . 2495

Exploring the Cognitive and Behavioral Risks and Maintenance Factors of Hikikomori: Protocol for an Ecological Momentary Assessment Study (e81384)
Alexander MacLellan, Keisuke Takano. . . . 2513

Developing a Culturally Adapted Digital Health Application for Older Hispanic Adults With Type 2 Diabetes: Protocol for a Qualitative and Pilot Study (e76294)
Joshua Caballero, Raymond Ownby, Henry Young, Michelle McElhannon, N Bhuiya, Kenny Esho, Russ Palmer. . . . 2523

A Multilevel Network Peer Intervention Among Student Men Who Have Sex With Men Attending University: Protocol for an Implementation-Effectiveness Before-After Cohort Study (e77078)
Jingpei Xu, Zhen Dai, Yushu Qiu, Liang Wang. . . . 2535

Development of an Ethico-Legal Framework for Quality Improvement and Performance Management in Health Care: Protocol for a Qualitative Study (e82167)
Kavisha Shah, Anna Janssen, Cameron Stewart, Tim Shaw, Ian Kerridge. . . . 2550

Implementation of Digital Device–Assisted and Nurse-Led Case Management to Promote Self-Management in Adults With Noncommunicable Diseases: Protocol for a Single-Arm Intervention Study (e80083)
Raweewan Pongpoottipatchara, Kanlayawee Anonjarn, Warithorn Prawatwong, Naruemol Singha-Dong. . . . 2558

Home-Based Exercise Cardiac Telerehabilitation on Adherence and Functional Capacity for Patients After Percutaneous Coronary Intervention in Indonesia: Protocol for a Quasi-Experimental Study (e81067)
Wantiyah Wantiyah, Anggoro Hartopo, Sri Setiyarini, Budi Setianto. . . . 2575

Evaluating the Acceptability of Using Virtual Reality to Promote Physical Activity Among Latino, Latina, and Latine Adults With Cardiometabolic Risk Factors and Obesity in Underresourced Settings: Protocol for a Qualitative Focus Group Study (e80534)
Desiree Acosta, Leslie Aguilar-Hernandez, Gael Perez, Iris Guzman-Ruiz, Josyel Castellon, Jailene Cruz, Liana Gutierrez, Paulina Monteon-Garcia, Vanessa Torres, O Duru, Yelba Castellon-Lopez. . . . 2588

Cognitive Recovery After Poststroke Delirium (RECOVER): Protocol for a Longitudinal Multimodal Study on Cognitive Assessment in Patients With Stroke (e77508)
Lena Geiger, Sebastian Klütz, Christian Mychajliw, Christoph Gäbele, Andreas Jooß, Constanze Single, Katharina Feil, Ulf Ziemann, Annerose Mengel. . . . 2599

Self-Management Combined With Digital Health Interventions to Improve Dietary Behavior, Exercise Behavior, Stress Management Behavior, and Blood Pressure Among Thais With Uncontrolled Hypertension: Protocol for an Explanatory Sequential Study (e81148)
Saowaluck Sukpattanasrikul, Naruemol Singha-Dong, Yajai Sitthimongkol. . . . 2614

Just-in-Time Adaptive Intervention to Promote Walking Behavior and Reduce Stationary Time in Physically Inactive Adults: Protocol for the Walking With JITAIs Study (e79022)
Cora Firkin, Ajith Vemuri, Tanvir Rahman, Barry Bodt, Elizabeth Orsega-Smith, Keith Decker, Gregory Dominick. . . . 2627

The Effect of Digital Therapy–Based Aerobic Exercise on Children With Neuroblastoma: Protocol for a Prospective Controlled Trial (e72698)
Xuan Zhou, Shihao Huang, Lixia Wang, Haibin Guo, Sisi Chen, Tingrui Li, Qing Du. . . . 2641

XSL•FO
RenderX

Association Between Mandibular Third Molar Position and Recurrent Pericoronitis: Protocol for a Cross-Sectional Study (e72682)
Suwarna Dangore- Khasbage, Tanaya Teredesai. . . . 2652

Development of a Mobile App (MyLepto App) to Improve Knowledge, Attitude, and Practice Regarding Leptospirosis Among Wet Market Workers in Selangor, Malaysia: Protocol for a Quasi-Experimental Study (e75809)
Mas Aljunid Merican, Zaleha Md Isa, Rozita Hod, Roszita Ibrahim, Zamtira Seman, Rusdi Abd Rashid. . . . 2658

Development and Feasibility Assessment of an Intrinsic Capacity Program in Primary Care: Protocol for an Implementation Science Approach (e84257)
Qing Wang, Mimaika Ginting, Ezra Ho, Siew Goh, Jonathan Gao, Woan Tan, Yew Ding, Wei Ng, Jonathan Ng, Sing Kwek, Richard Hui, Zhen Wang, Chirk Ng, Grace Sum, The IMPACTFrail Group. . . . 2669

Contribution of Emotion Dynamics to Adolescent Psychosocial Well-Being: Protocol for a Longitudinal Study (e76333)
Carli Mastronardi, Jade Powers, Rosanne Menna, Lance Rappaport. . . . 2688

Effectiveness of a Digital Awareness App in HIV/AIDS Mitigation Among Transgender Individuals in Rawalpindi District: Protocol for a Quasi-Experimental Study (e84610)
Muhammad Farooq, Shaheer Khan, Ali Ahmed. . . . 2707

A Digital Intervention for Capturing Real-Time Health Data for Epilepsy Seizure Forecasting: Protocol for the ATMOSPHERE Study (e85993)
Lauren Thompson, Emily Nielsen, Emily Quilter, Amelia Slay, Marceli Wac, Matthew Wragg, Samuel Downes, Phil Tittensor, Leandro Junges, Liz Stuart, Rosie Charles, Peter Kissack, Yasser Qureshi, Amberly Brigden. . . . 2718

HealthData@MAD-R&I: Protocol for Design and Development of a Regional Health Data Infrastructure to Enable Secondary Use of Health Data in Research and Innovation (e82815)
Montserrat León-García, Sergio Álvarez-Pérez, Janire Gesto-Gómez, Clara Urbano-Molina, Sonia Soto-Díaz, Juan Cárdenas-Valladolid, Luis Rodríguez-Rodríguez, Antonio Díaz-Holgado, Isabel del Cura-González, Javier De La Cruz-Bertolo, Noelia García-Barrio, Juan Cruz-Bermúdez, Cristina García-Fernández, Carlos Rodríguez-Antolín, Laila García-Aldars, Elsa Moreda-Sánchez, Álvaro Roldán López, José Veganzones Alonso-Cortés, Ana Gonzalez Gonzalez, Miguel Salinero-Fort, HealthData@MAD-R&I Working Group. . . . 2733

Mobile Imaging–Based Machine Learning for Dental Caries, Sealants, and Fluorosis: Protocol for a Cross-Sectional Model Development and Validation Study (e91239)
Sang Park, Semin Kwon, Shaun Hong, Yuhyun Ji, Sreeram Nagappa, Jung Leem, Mei Lin, Eugenio Beltrán-Aguilar, Susan Griffin, Young Kim. . . . 2746

A Personalized, Home-Based, Multidisciplinary Outpatient Clinic for Managing Breathlessness in Chronic Obstructive Pulmonary Disease: Protocol for a Single-Arm, Mixed-Methods Cohort Study (e85766)
Jonathan Lee, Belinda Cochrane, Tracy Smith, Sheree Smith, Soo Foo, Luisa Garcia, Teresa Kemp. . . . 2760

Pediatric Emergency Agitation Care Enhancement: Protocol for a Prospective Mixed Methods Study (e92452)
Lillian Klein, Wendy Pomerantz, Teresa Pestian, Yin Zhang, Drew Barzman, Lisa Vaughn, Nancy Daraiseh, Ryan Siders, Holly Hanson, Bijan Ketabchi, Lynn Babcock. . . . 2772

Locus-Coeruleus Norepinephrine Functioning as a Predictor of Childhood Mental Health (LOCUS-MENTAL): Protocol for a Longitudinal Study (e89237)
Iskra Todorova, Paula Neubauer, Nico Bast. . . . 2784

Care Pathways and Patient Experiences Among Patients With Post COVID-19 Condition: Study Protocol for a Mixed-Methods Study in Germany (e91976)
Daniela Gesell, Paula Lienesch, Yanyan Shi, Ralf Strobl, Martin Tauscher, Roman Gerlach, Eva Grill, Daniela Koller. . . . 2805

XSL•FO
RenderX

Establishment, Implementation, and Impacts of the Observatory on Student Mental Health in Higher Education in Quebec, Canada: Protocol for a Mixed Methods Research Program (e83225)
Julie Lane, Benjamin Gallais, François Lauzier-Jobin, Sèverine Lanoue, Félix Guay-Dufour, Rémi Paré-Beauchemin, Rachel Guertin, Christiane Bergeron-Leclerc, Saliha Ziam, Marie-Claude Lallier Beaudoin, Martine Shareck, Véronique Boudreault, Marie-Ève Langelier, Steve Geoffrion, Nancy Beauregard, Bianca B-Lamoureux, Magaly Brodeur, Esther McSween-Cadieux, Audrey Dupuis, Jonathan Smith, Isabelle Thibault, Noémie Beauregard. . . . 2813

The Gut Bacterial Resistome in the First Two Years of Life: Protocol for a Longitudinal Observational Birth Cohort Study (e86058)
Justine Fri, Idris Njanje, Tjale Mahopo, Lufuno Mavhandu-Ramarumo, Pascal Bessong, Microbiome and Antimicrobial Resistance Collaborators. 2 8 3 2

Exploring the Role of Dairy Consumption on the Growth and Development of Canadian Children: Protocol for a Longitudinal Mixed Method Research (e87415)
Karim Karbin, Mojtaba Shafiee, Zoe Longworth, Adam Baxter-Jones, Marta Erlandson, Ginny Lane, Ling-Ling Tsao, Walter Siqueira, Hassan Vatanparast. . . . 2847

Improving Registration and Dataflows Between Pediatric Oncology Clinics and the Childhood Cancer Registry of Switzerland: Protocol for SwissPedCancer Quality Assurance Study (e87007)
Yara Shoman, Lorenz Leuenberger, Grit Sommer, Julia Bielicki, Pierluigi Brazzola, Sophia della Valle, Manuel Diezi, Daniel Drozdov, Fabienne Gumy-Pause, Ana Guerreiro Stücklin, Ursula Kuehnel, Katrin Scheinemann, Christina Schindera, Freimut Schilling, Nicolas Waespe, Ben Spycher, Luregn Schlapbach, Claudia Kuehni, Fabiën Belle, SwissPedHealth Consortium. . . . 2860

Cardiovascular Health and Financial Hardship: Protocol for a Qualitative Citizen Science Study (e89101)
Dagmar Niewold, Evy van Gestel, Latifa Abidi, Klaske Tiemstra, Myrte Thoolen, Gera Nagelhout. . . . 2876

Use of Vending Machines to Deliver Oral Rapid HIV Self-Tests to Veterans: Protocol for a Pilot Study (e84317)
Tessa Rife-Pennington, Michael Douglas, Wendy Xie, Jennifer Cocohoba. . . . 2887

Bedside Clinical Ultrasound Performed by Family Physicians in Adult Patients With Abdominal Pain in a Hospital Emergency Department: Protocol for a Pilot Quasi-Experimental Study (e82393)
Laura Carbajo Martín, Ignacio Párraga-Martínez, Luis Beltrán-Romero, Máximo Bernabeu Wittel, Northern Huelva Health Management Area Research Group. . . . 2895

Biopsychosocial Determinants, Diet Quality, Gastrointestinal Health, and Disease Activity in Adults With Rheumatoid Arthritis: Cross-Sectional Descriptive Study (e79889)
Maureen McGarrity-Yoder, Emily Cope, Heidi Wayment, Alicia Rodriguez-Pla, Tracy Crane. . . . 2911

Death Literacy and Death Competence in Undergraduate Clinical and Allied Health Education: Protocol for a Mixed Methods Study (e86867)
Sora Lee, Kerrie Noonan, Andrea Grindrod, Sumina Shrestha, Niki Read. . . . 2928

Early-Life Exposures, Neurodevelopment, and Health Outcomes: Protocol for a Birth Cohort Study (e78593)
Tor Strand, Maria Averina, Kjersti Bakken, Sudha Basnet, Yvonne Böttcher, Sandra Huber, Mari Hysing, Ingrid Kvestad, Torben Lüders, Adrian McCann, Dhiraj Pokhrel, Suman Ranjitkar, Arun Sharma, Merina Shrestha, Manjeswori Ulak, Ram Chandyo. . . . 2936

Quantitative Understanding of Advanced Novel Imaging Techniques for Fasciitis and Biosignature Yield (Quantify): Protocol for a Cross-Sectional Diagnostic Study (e87613)
Zahra Amerian, Timothy Fleagle, Utsav Tuladhar, Rachel Watson, Micah Wong, Barbara Van Gorp, Brian Smith, Michael Richards, Mederic Hall, Joe Donnelly, Jessica Danielson, Renata Leao, David Stanley, Vincent Magnotta, James Holmes, Kathleen Sluka, Ruth Chimenti. . . . 2948

Differential Factors Associated With the Presence of Persistent Symptoms in Individuals Diagnosed With Long COVID: Protocol for a Longitudinal Matched Case-Control Study (e67133)
David Lerma-Irureta, Fátima Méndez-López, Elena Navarro-Matías, Diego Lerma-Puertas, Rosa Magallón-Botaya. . . . 2963

XSL•FO
RenderX

Development of a Line of Care for the Health of People Who Engage in Chemsex: Protocol for a Multimethod Study (e84068)
Lariane Cepas, Isadora Carvalho, Alvaro de Sousa, Caíque Ribeiro, Shirley Lima, Anderson Sousa, Carlos Sorgi, Ricardo Nakamura, Sumaé Embaló, Ivone Paula, Naila Santos, Rodrigo Freitas, João Silva Netto, Talita Fernandes, Ana Fernandes. . . . 2978

Maternal Perception vs Actual Breast Milk Supply: Protocol for an Observational Cross-Sectional Study (e84776)
Bailey Bruckner, Rachael Taylor, Carmen Parata, Jillian Haszard, Barry Taylor, Samantha Bevin, Kassidy Gooding, Anne-Louise Heath, Ioanna Katiforis, Lisa Daniels. . . . 2995

Taunakitanga Takitini, Reframing Self-Management Support for All in Aotearoa New Zealand: Protocol for a Participatory Case Study Program of Research (e89658)
Leigh Hale, Bernadette Jones, Ofa Dewes, Shirley Keown, Brigit Mirfin-Veitch, Amanda Wilkinson, Donna Keen, Pauline Norris, Eileen McKinlay, Meredith Perry, Kate Diesfeld, Henrietta Trip, Tristram Ingham, Nancy Longnecker, Roberta Francis Watene. . . . 3012

Kenyan Neonatal Mortality Risk Predictor: Protocol for a User-Centered Design Evaluation (e81996)
Ronald Nyatuka, Paul Macharia, Kakhata Esther, Faith Siva, Betsy Muriithi, Md Rahman Jabin. . . . 3030

Comparative Evaluation of Effectiveness of Oxygen-Releasing Gel With Scaling and Root Planing (SRP) With That of Scaling and Root Planing Alone in the Management of Chronic Periodontitis Patients: Study Protocol for Cross-Sectional Survey (e75795)
Geeta Bhandari, Priyanka Jaiswal, Shweta Bhagat, Sakshi Kotecha. . . . 3042

Perioperative Antibiotic Prophylaxis in Cesarean Section and the Maternal Gut Microbiome: Protocol for a Remote Observational Cohort Study (e84909)
Elisabeth Feles, Frauke Mattner. . . . 3049

Children's Attitudes Toward Dementia: Protocol for a Cross-Sectional Study (e81576)
Gema Somoza-Fernández, Javier Ortiz-Alonso, José Serra-Rexach, María Vidán. . . . 3063

Evaluating the Impact of Short-Notice Accreditation Assessment on Hospitals' Patient Safety and Quality Culture: Protocol for a Mixed Methods Study (e76945)
Robyn Scanlan, Tracy Flenady, Amy-Louise Byrne, Jenni Judd. . . . 3074

Development of a Modified Intimate Wash From Yoniprakshalan Kwatha (YOSHA) With Comparative Antimicrobial and Quality Assessment: Protocol for an Experimental Study (e73239)
Prachi Jadhav, Bharat Rathi. . . . 3088

Exploring the Impact of Initiating Endocrine Therapy on Metabolic Health in Early Breast Cancer: Protocol for the Prospective Follow-Up EMETA-Study (e78589)
Lærke Nissen, Jonas Busk Holm, Signe Borgquist. . . . 3097

Community Health Worker–Integrated Model for Sickle Cell Disease Management: Protocol for a Feasibility Study (e82663)
Tilicia Mayo-Gamble, Doreen Ugwu, Tobi Oloyede, Joseph Telfair, Abdullah Kutlar. . . . 3107

Measuring Accuracy (Classification Probabilities, Positive, and Negative Predictive Values) of Executive Function Electroencephalogram Metrics in Attention-Deficit/Hyperactivity Disorder Diagnosis: Protocol for and Perspectives From the SINCRONIA Study (e79150)
Hilario Blasco-Fontecilla, Javier Sánchez-Cerezo, Irene Gómez, Georgelina Abreu-Fernández, Sandra Ortiz, Jesús Villoria, Miguel Blanco, Ana García, Julia Ballesteros, Roldán Martínez, Gerardo Gálvez, Fernando Maestú, Álvaro López-Medrano. . . . 3115

XSL•FO
RenderX

Competency and Role Development in Advanced Nursing Practice Following a Swiss Master's Cohort Through Education and Early Clinical Practice: Protocol for a Longitudinal Convergent Mixed Methods Study (e85773)
Astrid Braun, André Fringer, Heidrun Gattinger, Christian Eissler, Maya Zumstein-Shaha, Thomas Beer, Martina Roes. 3129

Multidisciplinary Treatment With Hepatic Arterial Infusion Chemotherapy, Radiotherapy, and Immunotherapy for Advanced Hepatocellular Carcinoma With Major Vascular Invasion: Prospective Registry Protocol (e82992)
Yoshiko Doi, Hiroshi Aikata, Yumi Kosaka, Takashi Nakahara, Michiyo Kodama, Masashi Hieda, Masakazu Hashimoto, Hideki Nakahara, Ippei Takahashi, Hideaki Kakizawa, Nami Mori, Keiji Tsuji, Nobuki Imano, Yuji Murakami, Saki Sueda, Tomokazu Kawaoka, Masataka Tsuge, Shiro Oka. 3142

Occupational Exposure to Resorcinol and Thyroid-Disrupting Effects: Protocol for an Exploratory Field Study in French Hairdressers (e65833)
Anca Radauceanu, Amandine Cambrai-Erb, Benedicte Adet, Marie-Isabelle Nioule, Flavien Denis, Romain Pons. 3153

Impact and Economic Evaluation of the Patient-Provider Support Agency Model Under India's National Tuberculosis Elimination Program: Protocol for a Cohort Study (e76302)
Nikita Raula, Ingeborg van der Putten, Nicole Dukers-Muijrers, Shankar Dapkekar, Priyakanta Nayak, Silvia Evers. 3169

Exploring Personal Support Worker Education, Recruitment, and Retention in Long-Term Care Homes in Ontario, Canada: Protocol for a Mixed Methods Study (e79164)
Alison Kernoghan, Danielle Just, Catherine Tong, Linda Sheiban Taucar, D'Arcy Kirkwood, Benjamin Smith, Living Classroom Team. 3183

Assessment of Programmed Cell Death Ligand-1 Expression in Oral Potentially Malignant Disorders and Tumor-Free Surgical Margins of Oral Squamous Cell Carcinoma: Protocol for a Retrospective Cross-Sectional Study (e82568)
Archana Sonone, Alka Hande, Swati Patil, Aayushi Pakhale, Preethi Sharma. 3193

Self-Management Support After Burns: Protocol for a Multicenter, Stepped-Wedge Hybrid Type II Effectiveness-Implementation Study (e86671)
Sharon Blok, Marianne Nieuwenhuis, Sonja Scholten-Jaegers, Irma Visser, Raaba Thambithurai, Monique de Craen-de Moor, Kiran Baran, Anuschka Niemeijer, Eelke Bosma, Corry van der Sluis, Sven Geelen, National Burn Care, Education & Research Group, The Netherlands. .
3 2 0 0

Molecular and Pharmacogenetic Marker Evaluation in Relation to the Toxicity and Clinical Response of Acute Lymphoblastic Leukemia Treatment in Indian Children (MPGx-INDALL): Protocol for a Prospective Observational Cohort Study (e79865)
Swetambri Sharma, Shuvadeep Ganguly, Kamali Murugadoss, Smita Kayal, Swaminathan Keerthivasagam, Jaikumar Ramamoorthy, Archna Singh, Deepam Pushpam, Jayanthi Mathaiyan, Yvonne Gloor, Frederic Baleydier, Marc Ansari, Sameer Bakhshi, Biswajit Dubashi, Chakradhara Uppugunduri. 3223

Factors Associated With Loss to Follow-Up Among People Living With HIV in a National Tertiary Care Hospital: Protocol and Baseline Analysis of a Prospective Cohort Study (e76470)
Luis Del Moral Trinidad, Jaime Andrade Villanueva, Luis Ruíz Mora, Carlos García y Nuño, Maria Perez Quintero, Brian Apodaca Escalante, Jocelyn Torres Arias, Juan Martínez Herrera, Melva Herrera Godina, Luz González Hernández. 3234

Factors Influencing the Transition From Physiological Hyperopia to Myopia in Children: Protocol for a Prospective Nested Case-Control Study (e84888)
Huaying Xu, Jianing Pu, Shimeng Bian, Xuejing Mi, Zhen Zhou, Wei Chen, Yonghong Jiao. 3244

Evaluation of the Long-Term Safety of Avacopan in Antineutrophil Cytoplasmic Antibody–Associated Vasculitis in the Real World (AvacoStar): Protocol for a Noninterventional Prospective Cohort Study (e81415)
David Jayne, Raashid Luqmani, Benjamin Terrier, Achim Obergfell, Charlotte Pollet, Marie Boff, Monica Balcells-Oliver, Bernhard Hellmich. . .
3 2 5 4

XSL•FO
RenderX

Factors That Influence Patients' Decisions About Repetitive Transcranial Magnetic Stimulation as a Treatment Option for Treatment-Resistant Depression: Protocol for a Prospective Mixed Methods Cohort Study (e82491)
Alexandra Godinho, Sanvitti Dalmia, Steven Selchen, Daniel Tziatis, Pete Wegier. 3265

Disease Progression and Outcomes in Patients With Benign Prostatic Hyperplasia: Protocol for a Multicenter Retrospective Cohort Study (e84965)
Zexuan Lv, Jiaxing Liu, Yin Lu, Xu Zhang, Fan Gao, Bingyang Guo, Hongyu Zhang, Hongzhao Li, Chao Lv. 3276

Safety of Administration of Vasopressors Through Peripheral Compared to Central Venous Catheters in a Rural Kenyan Hospital: Protocol for a Prospective Observational Cohort Study (e81794)
Grace Gao, Brian Kimani, Linette Lepore, Zahra Aghababa, Hannah Gitura, Michael Kinuthia, Jack Musau, Judy Muthoki, Alice Mwonge, Fredrick Ndibaru, Ann Ndubi, Juliah Ndungu, Kinzi Nixon, Moses Odhiambo, Winslet Okari, Benjamin Thairu, Lilian Vihenda, Rachel Wojcik, B Brotherton, Kristina Rudd. 3292

A Novel Social Network Approach to Measure Intersectional Stigma Among Latino Men Who Have Sex With Men in San Diego, California (NEXUS): Protocol for a Longitudinal Cohort Study (e72334)
Laramie Smith, Angel Algarin, Eileen Pitpitan, Heather Pines, Nicole Kelly, Carl Latkin, Aaron Gutierrez, Maryam Hussain, Francisco Soto, Ricardo Vazquez Jr, Stevie Juarez, Albert Genegaling, Rosalinda Rodriguez, Anthony Cirilo, Juan Esparza, Saul Cruz, Erick Jimenez, Ulises Reyes, Veronica Moore, Malek Guerbaoui, Beth Davenport, Jeannette Aldous, Katherine Penniga, Kenyatta Parker, Britt Skaathun. 3301

Addressing Gaps in the Hypertension and Diabetes Care Continuum in Rural Bangladesh Through Digital Technology Supported Decentralized Primary Care: Study Protocol and Baseline Results for a Hybrid Effectiveness-Implementation Trial (e71696)
Wubin Xie, Sabrina Ahmed, Ali Ahsan, Ananya Gupta, Anais Keenan, Tanmoy Sarker, Fahmida Akter, Aysha Anan, Md Hossain, Zahidul Quayyum, AHM Hussain, Robed Amin, Imran Chowdhury, Mithila Faruque, Sohel Choudhury, Ian Goon, Fred Hersch, Lora Sabin, Brian Oldenburg, John Chambers, Malay Mridha. 3323

Optimizing Antiemetic Support in Anthracycline-Based Chemotherapy for Early Breast Cancer: Protocol for a Prospective Observational Study of Four-Drug Antiemetic Therapy Including Fosnetupitant and Olanzapine (e85648)
Ayako Higuchi, Yumiko Koi, Tomoaki Eto, Yuka Maeda, Wakako Tajiri, Junji Kawasaki, Sayuri Akiyoshi, Hideki Ijichi, Yoshiaki Nakamura, Chinami Koga, Mototsugu Shimokawa, Hiroaki Shimizu, Toshihiro Matsumoto, Eriko Tokunaga. 3342

Pilot Implementation of a Post–Hypertensive Disorders of Pregnancy Education and Follow-Up Package for Health Care Providers: Protocol for a Mixed Methods Pilot Study (e81069)
Jennifer Green, Heike Roth, Ben Harris-Roxas, Kathleen Baird, Caroline Guirgis, Zarin Gundevia, Amanda Henry. 3350

Establishing Reference Intervals for Fat-Soluble Vitamins in Healthy Chinese Adults: Study Protocol for a Cross-Sectional, Multicenter Study (e80425)
Zhong Liu, Chenbing Liu, Nan Li, Chao Shen, Lihong Qiu, Di Sheng. 3360

In Vitro Comparison of the Dimensional Accuracy of Implant Impressions Using Custom Acrylic Trays and Prefabricated Self-Perforating Trays: Protocol for a Comparative Evaluation Study (e79452)
Arti Agrawal, Sharayu Nimonkar, Surekha Godbole, Rushabh Parakh, Vikram Belkhode, Namita Zilpilwar. 3372

Analysis of Associated Factors and Construction of Risk Prediction Models for Frailty in Hospitalized Older Adults Living With HIV: Protocol for a Prospective Observational Study (e84271)
Fan Li, Chaoying Xie, Fang Xiang. 3382

Stereotactic Adaptive Radiation Therapy for Borderline Resectable or Locally Advanced Pancreatic Cancer to Minimize Gastrointestinal Toxicity (ARTIA-Pancreas): Protocol for a Single-Arm Prospective Trial (e84607)
Lauren Henke, Hyun Kim, Eric Laugeman, Steven Kohlmyer, Claire McCann, Kate Pietrovito, Kenneth Russell, Jennifer Woo, Alex Price. . . .
3 3 9 3

XSL•FO
RenderX

Implementation of a Personalized Medicine Approach in Patients With Type 2 Diabetes Mellitus Receiving Multiple Daily Insulin Injections (POMA Project): Protocol for a Before-and-After Intervention Study (e85375)
Rosa Giné-Balcells, Bogdan Vlacho, Olga Vidal-Pérez, Albert Lecube, Josep Franch-Nadal, Dídac Mauricio, Àngels Molló-Iniesta, Marta Hernández, POMA Study Group. . . . 3404

Using Textural Analysis of Thermal Imaging to Predict Healing Status of Diabetes-Related Neuropathic Foot Ulcers: Protocol for a Co-Design and Longitudinal Study (e69793)
Kate Waller, Rajna Ogrin, Quoc Ngo, Barbara Polus, Samantha Hanna, Aye Tint, Richard MacIsaac, Anna Galligan, Elif Ekinci, Dinesh Kumar. 3 4 1 6

Educational Intervention to Improve Sexual Health and Quality of Life in Survivors of Breast and Gynecological Cancer: Protocol for a Mixed Methods Feasibility Study (e80567)
Nathalia Salazar Falla, María Navas Jojoa, Juan Lalinde Riviño, Auramaría Escandón Bucheli, Juan Galindo-Sánchez. . . . 3428

GORE VIABAHN VBX Balloon-Expandable Endoprostheses as a Bridging Stent With Branched and Fenestrated Endografts in the Endovascular Treatment of Aortic Aneurysms: Protocol for a Retrospective Multicenter Registry (EMBRACE Study) (e78970)
Pietro Dioni, Anders Wanhainen, Jacob Budtz-Lilly, Nuno Dias, Tab Bonny, Luca Bertoglio, EMBRACE study investigator. . . . 3439

Interventions for Students' Well-Being at the University of Helsinki (INSIGHT): Protocol and Preliminary Descriptive Results for a Quasi-Experimental Controlled Trial of a Social Identity Intervention and Two Active Comparators (e79319)
Silja Martikainen, Oona Karhunen, Iiris Mankki, Vera Gergov, Päivi Berg, Satu Venäläinen, Elisa Rissanen, Johanna Lammintakanen, Erkki Heinonen, Anu Lehtinen, Janna Pöntinen, Henna Asikainen, Nina Katajavuori, Outi Linnaranta, Milla Räisänen, Samuli Salminen, Jari Lahti. . 3 4 5 2

Investigating the Role of the Environment on Physical Activity Interventions (the InSPACE Project): Protocol for a Pooled Secondary Analysis of Randomized Controlled Trials (e83151)
Amy Youngbloom, Maya Rowland, Stephen Mooney, Kaitlin Singer, Adam Szpiro, Brian Saelens. . . . 3474

The Norwegian Adult Mental Health Registry for Quality Control in Specialized Mental Health Services: Protocol for a Nationwide Naturalistic Study (e82696)
Tore Tjora, Kent Jensen, Ingrid Dieset, Katrine Fjukstad, Per Holman, Martin Tesli, Inge Joa. . . . 3494

Utility of RECELL for Traumatic Skin Defects: Protocol for a Prospective, Single-Arm Pilot Study (e74159)
Naoki Maegawa, Hironobu Konishi, Keisuke Masuda, Akinori Okuda, Hideki Asai, Keita Miyazaki, Yasuyuki Kawai, Hidetada Fukushima. . . . 3 5 0 9

Correlation Between Prakriti (Body Constitution) and Severity of Structural Alterations in the Lungs of Patients With SARS-CoV-2: Protocol for a Retrospective Cross-Sectional Study (e63916)
Rugaved Gudadhe, Gaurav Sawarkar, Amol Deshpande. . . . 3517

Comparative Evaluation of the Antibacterial Efficacy of Metronidazole, Chlorhexidine, and Normal Saline as Laser- and Sonic-Activated Intracanal Irrigants Against Enterococcus faecalis: In Vitro Study Protocol (e76783)
Sharvari Deshmukh, Aditya Patel. . . . 3524

The Role of Iron-Overloaded Macrophages in Mesenchymal Stem Cell Senescence and Anemia in Myelodysplastic Syndromes: Protocol for an In Vitro Study (e77936)
Lu Bai, Fan Wang, Yaxian Jiang, Hongmei Ouyang. . . . 3533

Rapid Qualitative Approaches in Pandemic Research: Protocol for an Exploratory Qualitative Multimethod Study (VERDIQual) on Mpox in Italy, Nigeria, Thailand, and the United Kingdom (e77321)
Chinye Osa-Afiana, Marthe Le Prevost, Emily Nicholls, Davide Bilardi, Thomas Guadamuz, Esekwe Soje-Amadosi, Grace Adebisi, Chibueze Adirieje, Adeyosola Adetunji, Karima Yusufu, T Witzel, Worawalan Waratworawan, Nattharat Samoh, Tom May, Sarah Denford, Willian Gomes, Uzodinma Adirieje, Shema Tariq, Nadia Sam-Agudu. . . . 3541

XSL•FO
RenderX

Nutrition Education for Emerging Adults: Protocol for Program Evaluation (e81647)
Stephanie Rogus, Katelin Alfaro Hudak, Lilian Ademu, Sumathi Venkatesh, Michael Laguros, Elizabeth Racine. 3555

Comparative Evaluation of Push-Out Bond Strength of Glass Fiber Posts and Carbon Fiber Posts in Root Canals Treated With Calcium Hydroxide–Based, Resin-Based, and Bioceramic-Based Root Canal Sealers: Protocol for an In Vitro Study (e76621)
Vinus Shivlani, Manoj Chandak, Sanika Damle. 3563

Comparative Evaluation of Sonic and Ultrasonic Activation of 3 Post Space Irrigants on Push-Out Bond Strength of Fiber Posts Luted With Dual-Cure Resin Cement: Protocol for an In Vitro Study (e76907)
Apurva Wamane, Aditya Patel. 3570

Developing a Program Theory on Ventilator Weaning in Adult Intensive Care: Protocol for a Multimethods Study (e83342)
Fritz Sterr, Lydia Bauernfeind, Christian Rester, Sabine Metzing, Rebecca Palm. 3578

Establish a PrEP (Pre-Exposure Prophylaxis) Epidemiology, Modeling, and Surveillance (PREMISE) System to Analyze Trends in PrEP Uptake and the Impact of PrEP Programs and Policies: Protocol for a Natural Experiment and Modeling Study in the United States (e80911)
Aaron Siegler, Shi Koh, Wenting Huang, Eric Hall, Jeb Jones, Courtney Yarbrough, Xiao Zang, Bohdan Nosyk, Edwin Corbin-Gutierrez, Patrick Sullivan. 3591

Differentiating the COVID-19 Infection and Vaccine Experiences of Patients With Systemic, Single Organ, and Overlap Immune-Mediated Inflammatory Disease: Protocol for a Secondary Analysis for Enhancing COVID-19 Vaccine Pharmacovigilance (e68785)
Meredith Leston, José Ordóñez-Mena, Xinchun Gu, Mark Joy, F Hobbs, Lennard Y W Lee, Ioannis Parodis, Laura Andreoli, Nelly Ziade, Latika Gupta, Vikas Agarwal, Richard Conway, John Isaacs, Jeffrey Curtis, Alessia Alunno, Stefan Siebert, Ori Elkayam, Karen Schreiber, Simon de Lusignan. 3609

Evaluation of the Antimicrobial Efficacy of Cerium Oxide Nanoparticles and Chlorhexidine Gluconate Against Enterococcus faecalis in Root Canal Systems: Protocol for an In Vitro Study (e77998)
Bhuvaneshwari Khanadali, Anuja Ikhar. 3620

Effects of Time-Restricted Eating on Ketone Metabolism and Immunoregulation in Premenopausal Women: Protocol for a Pilot, Prospective, Single-Arm Dietary Intervention Study (e81063)
Geethika Thota, Natalie Macheret, Sophia Glaros, Ila Kacker, Samson Cantor, Lilian Mabundo, Noemi Malandrino, Kong Chen, Amber Courville, Sara Turner, Shanna Yang, Andrea Krenek, Thomas Recupero, Rachael Klein, Kim Han, Patrycja Puchalska, Peter Crawford, Michael Sack, Stephanie Chung. 3627

3D Comparative Evaluation of Condylar Morphology Between Chronic Areca Nut Chewers and Nonchewers: Protocol for a Case-Control Study (e84038)
Vaishnavi Tarpe, Suwarna Khasbage. 3639

Perceptions of Abortion and Sexual and Reproductive Health in Chilean Medical and Midwifery Education: Protocol for a Mixed Methods Study (e81427)
Lidia Casas-Becerra, Alejandra Ramm, Beatriz Pérez, Heidy Kaune, Mercedes Carrasco-Portiño, Adela Montero, Cecilia Bustos-Ibarra, Henry Castro-Arias, Claudio López-Labarca, Leonardo Reyes-Torres. 3647

Real-World Effectiveness and Safety of Liuwei Dihuang Pill for Menopausal Syndrome: Protocol for a Prospective, Observational, Multicenter Cohort Study (e84803)
Shuoshuo Wei, Hongyan Zhang, Mengmeng Wang, Yuhui Sun, Tao Song, Weifang Zheng, Yunmei Ye, Liran Ma, Yunshuang Yang, Tong Wei, Lan Cheng, Chunquan Sun, Xingting Ding, Shichu Zhao, Qingqiao Song, Lianxin Wang. 3661

XSL•FO
RenderX

Safety of Postpartum Herbal Medicine in Primary Care Korean Medicine Clinics: Protocol for a Prospective, Multicenter, Registry-Based Observational Study (e87543)
Anna Kim, Eunbyul Cho, Young-Eun Kim, Sumin Seo, Sungha Kim, Jae-Ahn Shin, Eunhee Lee, Mi Son. 3673

Integrating Sperm Microscopy, Environmental Exposures, and Lifestyle Factors for Male Fertility Analysis: Protocol for the Nippon Semen and Environmental Exposure Database (N-SEED) Cross-Sectional Study (e93803)
Tomoko Oguri, Kosuke Kojo, Yoshiaki Endo, Takaaki Matsuda, Haruhiko Midorikawa, Ayumi Nakazono, Kaoru Yanagida, Atsushi Ikeda, Hiroyuki Nishiyama, Isamu Ogura. 3682

Building Capacity in Institutional Operational Research in Low-Resource Settings: Protocol for an Implementation Research Study (e84284)
Sana Ali, G Murthy, Kenneth Bassett, Priya Reddy, Katie Judson, Suzanne Gilbert, Abhilash Patra, Anirudh Gudlavalleti, Debangana Mahapatra, Hira Pant, Parami Dakhwa, Sarah Jameel, Sarva Pandey, Shailaja Tetali, Sirshendu Chaudhuri, Srinivas Imandi, Suresh Kamalakannan. . . . 3 6 9 9

Sickle Cell Disease at a Tertiary Care Center in the Vidarbha Region of India: Protocol for a Clinical and Observational Study (e80483)
Tanushree Budhbaware, Jaishriram Rathored. 3710

Play and Supportive Environments Among Older Adults in the Context of Swedish Municipal Care: Protocol for the Development and Psychometric Evaluation of an Instrument to Measure Playfulness (e86373)
Albert Westergren, Anna Bergman, Sofi Knaust Lidén, Catarina Nahlén Bose, Anna Nivestam, Annika Norell, Petra Nilsson Lindström, Maria Haak. 3717

Quality of Life of Patients With Multi-Organ Autoimmune Disease and Its Relationship to Patient-Centered Care: Protocol for a Longitudinal Cohort Study (e77786)
Marianne Bayrhuber, Christian Schlett, Julia Slesaczeck, Sabrina Herbst, Gabriele Mueller, Alexandra Nieters, Bodo Grimbacher, Erik Farin-Glattacker. 3732

The Systematic Psychodynamic Multidimensional Organizational Assessment Questionnaire: Protocol for a Development and Validation Study (e78530)
Anne-Maria Müller, Sophia Sachs, Yannik Rieder, Claas Lahmann. 3741

Integration of the Integrate, Design, Assess, and Share Framework in Developing the Environmental Health Literacy Toolkit Paraben-Free & Me: Protocol for a Randomized Controlled Trial (e85474)
Graziella De Michino, Satveer Dhillon, Caroline Barakat. 3750

Developing and Validating Measures of Structural Ableism to Improve Health Outcomes for the Disability Community: Protocol for a Mixed Methods Study (e86976)
Rupa Valdez, Bonnielin Swenor, Franz Castro, Caroline Cerilli, Bailey Middleton, Kara Fitzgibbon, Ashley Shew, Varshini Varadaraj, Hari Srinivasan, Diana Cejas. 3759

Machine Learning for Estimating Cardiorespiratory Fitness in Patients With Obesity: Protocol for a Retrospective and Prospective Multicenter Cohort Study (e85069)
Jarle Berge, Vimala Nunavath, Rikke Asbjørnsen, Heidi Borgeraas, Jens Hertel, Anita Linge, Inge Groote, Trine Stensrud, Espen Gjevestad, Linda Mathisen, Matthan Caan, Martin Paulson, Eva Støa, Jan Johansen, AI and VO2max Group, Bjørn-Jostein Singstad. 3775

Assistive Robotic Arm to Support Activities of Daily Living in Individuals With Tetraplegia: Protocol for a Real-World Convergent Parallel Mixed Methods Feasibility Study (e78339)
Vera Fosbrooke, Aline Christen, Barbara Wortmann, Iris de Boer, Raphael Rätz, Julian Rösch, Gabriel Gruener, Anja Raab. 3789

Use of an AI-Based Tool (Human Experiences and Reflections Data Connector) to Improve Discovery and Reuse of Archived Qualitative Data: Protocol for an Algorithm Development and Validation Study (e82725)
Tammy Leonard, Jim Stimpson, Miguel Cano, Wenqi Shi, Ada Abaragu, Song Zhang. 3802

XSL•FO
RenderX

Life-Course Framework for Institute-Based Surveillance of Overweight, Obesity, and Metabolic Risk Factors in India: Protocol for Surveillance System Development (e86084)
Sai Challa, Biswanath Dash. . . . 3812

Standardized Methods for Evaluating Physical and Eating Behaviors: The WEALTH Cross-Sectional Study Protocol (e70186)
Grainne Hayes, Christoph Buck, Greet Cardon, Richard Cimler, Steriani Elavsky, Leopold Fezeu K, Janas Harrington, Jitka Kühnová, Jean-Michel Oppert, Luis Sigcha, Pepijn Van de Ven, Tomas Vetrovsky, Catherine Woods, Antje Hebestreit, Alan Donnelly. . . . 3826

International Clinical Practice Guidelines on Traditional Chinese Medicine Therapy for Dysmenorrhea: Development Protocol (e77423)
Li Zhang, Xiao Liang, Junlu Chen, Yuhan Du, Huilan Du, Xiaohui Wang. . . . 3843

Noninvasive Measurement of Cerebrospinal Fluid Flow in Shunted Hydrocephalus: Protocol for Scanner Calibration and Multisite Data Collection (e85918)
Jason Toliao, Matthew Borzage, Pradip Chaudhari, Joseph Ha, Isabella Friedli, Madison Gutierrez, Meghan Drastal, Peter Calabrese, Eamon Doyle, Isabel Torres, Jacob Al-Husseini, Clio González-Zacarías, John Wood, Jason Hauptman, J McComb, Stefan Blüml, Peter Chiarelli. . . 3 8 5 2

Exploring Risk Factors and Neurophysiological Mechanisms Underlying the Development of Chronic Postsurgical Pain After Thoracic Surgery: Protocol for an Observational Feasibility Study (e81042)
Mustaali Hussain, Stevie Foglia, Ameer Hamoodi, Jinhui Ma, John Agzarian, Christian Finley, Yaron Shargall, Aimee Nelson, Harsha Shanthanna. . . . 3863

Retrieving Transgender and Gender Diverse Literature: Protocol for the Development and Validation of 2 Search Hedges (e76055)
Chi Dinh, Zack Marshall, Scott Marsalis, Ryn Gagen, Preet Kang, Avery Everhart. . . . 3874

Anxiety-Free Public Dentistry for Adults With Disabilities by Using Head-Mounted Virtual Reality Technology: Protocol for a Feasibility Mixed Methods Study (e85916)
Dung Bui, Fiona McAlinden, Alice Urban, Charmine Hartel, Kristian Rotaru, Kadek Satriadi, Tanner Person, Wei Wang, Libby Callaway. . . . 3884

Cultural Adaptation of a Digital Mobile App for Bipolar Disorder (PolarUs): Protocol for a Qualitative Co-Design Study (e92600)
Leena Chau, Jill Murphy, Emma Morton, Martin Provencher, Delphine Raucher-Chene, Steven Barnes, Erin Michalak. . . . 3897

Supporting Unpaid Caregivers of Persons Living With Dementia: Protocol for a Pilot Feasibility Study to Explore Caregiver Outcomes and Impact of a Co-Designed Simulation–Based Psychoeducation Program in Virtual Reality (e87107)
Mary Chiu, Sarah Pistritto, Kristina Kokorelias, Krystina Clarke, Andrei Torres, Cole Craven, Adam Dubrowski, Bill Kapralos, Erica O'Hare, Joel Sadavoy, Adriana Shnall, Michael Smith, Nusrat Choudhury, Amer Burhan. . . . 3927

Acceptance and Commitment to Empowerment Intervention to Reduce HIV Stigma and Promote Community Resilience: Protocol for an Implementation Study (e80669)
Josephine Wong, Alan Li, Carla Hilario, Mandana Vahabi, Egbe Etowa, Isaac Luginaah, Aniela dela Cruz, Miya Narushima, Kenneth Fung. . . 3 9 4 0

Scaling Up a Diabetes Prevention Program in Geographically and Ethnoculturally Diverse Urban Regions of Canada: Protocol for a Hybrid Type 2 Implementation-Effectiveness Study (e80276)
Katie Weatherson, Jessica Bourne, Kaela Cranston, Kyra Braaten, Natalie Grieve, Joseph Kelly, Mary Jung. . . . 3953

A Context-Adapted Diabetes Prevention Program (Small Steps for Big Changes) in Australia: Protocol for a Hybrid Implementation-Effectiveness Study (e81195)
Genevieve Healy, Sjaan Gomersall, David Dunstan, Elisabeth Winkler, Katherine Heseltine, Grace Sim, Fiona Mason, Denis Giguere, Michael Tsiamis, Mary Jung, Ana Goode. . . . 3977

XSL•FO
RenderX

Determining the Impact of a Physiotherapist-Led Primary Care Model for Low Back Pain: Protocol and Analysis Plan for a Cluster Randomized Controlled Trial and Embedded Process Evaluation (e89004)
Jordan Miller, Catherine Donnelly, Chad McClintock, Kevin Varette, Yeimi Camargo, Jacquelyn Marsh, Monica Taljaard, Mir Mamun, Geneviève Bacchus, David Barber, Lynn Cooper, Simon French, Jonathan Hill, Michael Green, Joy MacDermid, Kathleen Norman, Julie Richardson, Joan Tranmer, Timothy Wideman. 4002

Updating the Good Reporting of a Mixed Methods Study (GRAMMS) Reporting Guidelines: Protocol for a Methodological Review and Modified Delphi Process (e82364)
Sarah Munce, Sergi Fàbregues, Quan Hong, Alicia O'Cathain, Timothy Guetterman, John Creswell, Cheryl Poth, Katie Dainty, Andrea Tricco, Mandy Archibald, Clementine Jarrett, Jennifer Boyle, Dorothy Luong, Frank Kiwanuka, Ahtisham Younas. 4024

Objective Assessment of Financial Decision-Making With a Simulated Online Money Management Task in Older Adults: Protocol for a Prospective Observational Study (e82488)
Preeti Sunderaraman, Madison Bouchard-Liporto, Zachary Kunicki, Edward Huey, Stephanie Cosentino. 4035

Strengthening Actions for Menstrual Health and Hygiene Interventions for Promotion of Women's Health in Nepal (SAMIP): Protocol for a Participatory Intervention Development Study Using Realist Synthesis, Human-Centered Design, Intervention Mapping, and Arts-Based Methods (e89117)
Sara Baumann, Hannah Schwarz, Annika Agarwal, Bhimsen Devkota, Elizabeth Miller, Ada Youk, Sara Parker, Sanjana Murthy, Nikita Sharma, Archana Shrestha, Madhusudan Subedi, Nirajan Khadka, Mary Hawk. 4048

Assessing Neonatal Transport, Health Outcomes, and Psychosocial Inequities Among Black, Very Preterm Infants and Their Families to Advance Health Equity: Protocol for a Mixed Methods Study (e84626)
Karen Warren, Tisha Felder, Nansi Boghossian, Robin Dawson, Robin Dail. 4067

Evaluating the Combined Effects of an Adverse Childhood Experiences–Focused Family Advocate Model and the Strengthening Families Program: Study Protocol for a Hybrid Type 1 Effectiveness-Implementation Trial in 36 New Jersey Communities (e85151)
Dallas Elgin, Phillip Graham, Kiersten Johnson, Sazid Khan, Michael Mills, Feker Wondimagegnehu, Brent Gibbons, Jason Williams, Xinyi Jiang, Andrea Zapcic, Brittany Simon. 4079

Feasibility of Internet-Based Mind-Body Training for Adults With Subjective Cognitive Decline: Protocol for a Randomized Controlled Trial (e86276)
Ruchika Prakash, Madhura Phansikar, Megan Fisher, Sarah Dove, Rebecca Andridge, Doug Scharre, Kathy Wright, Elizabeth Head, Claire O'Callaghan, Quinn Meyerson, Elizabeth Duraney, Matthew Schroeder, Niyathi Chakrapani, Elise Grant, Tyreek Jones, James Teng, Michael Esterman, Joshua Smyth, Thomas Gregoire, Rond Malhas. 4094

The National Dental Practice-Based Research Network Dental Implant Restoration Registry to Evaluate Dental Implant Outcomes in Community Practice Settings: Protocol for a Prospective Observational Study (e82795)
Joana Cunha-Cruz, Danyelle Barton, David Cochran, Phillip Crawford, Farahnaz Fahimipour, Ellen Funkhouser, Nicolaas Geurs, Gregg Gilbert, James Johnson, Ning Smith, Kaur Maninder, Sarah Startley, Kimberly Stewart, Lisa Waiwaiole, National Dental PBRN Collaborative Group. .
4 1 1 1

World Health Organization–Recommended Periodic Presumptive Treatment Versus Doxycycline Post-Exposure Prophylaxis for Sexually Transmitted Infection Control Among Men Who Have Sex With Men in Kenya: Protocol for a Randomized Controlled Trial (e81113)
Susan Graham, Fredrick Otieno, Joshua Kimani, Alé Barrientos, Olusegun Soge, Monisha Sharma, Deven Hamilton, Connie Celum, R. McClelland, Eduard Sanders. 4121

Virtual Mentored Implementation to Improve Care Transitions in Chronic Obstructive Pulmonary Disease: Protocol for a Pragmatic Implementation Study (e82043)
Cathryn Lee, Leah Traeger, Mahima Akula, April Fegley, Jenna Goldstein, Kim Erwin, Laura Damschroder, Jean Rommes, Hannah Pick, Andrew Auerbach, Peter Lindenauer, Wen Wan, Andrea Jackson Sagredo, Valerie Press. 4143

XSL•FO
RenderX

Factors Associated With Alcohol Use After Metabolic and Bariatric Surgery: Protocol for an Ecological Momentary Assessment (e87209)
Lisa Miller-Matero, Daniel Saleh, Brittany Christopher, Maha Albujuq, Alyssa Vanderziel, Erin Haley, Jordan Braciszewski, Roland Moore, Arthur Carlin, Kristina Jackson. . . . 4157

Central Nervous System Mechanisms and Treatment Response in Chronic Ocular Surface Pain: Protocol for a Cross-Sectional Observational Phenotyping Study (e84240)
Lindsey De Lott, Steven Harte, Chelsea Kaplan, David Williams, Roni Shtein, Maria Woodward, Alexander Tsodikov, Tatiana Deveney, Anat Galor, Anne Shea, Charles Schultz, Clare McKolay, Kathy Scott, Daniel Clauw. . . . 4167

Very Low–Carbohydrate Breakfast Intervention for Adults with Type 2 Diabetes and Persistent Hyperglycemia: Protocol for a Digital, Nonrandomized Pre-Post Study (e81041)
Annika Hansen, Kaitlyn Raymond, Alexander Simon, Sarah Kim, Hovig Bayandorian, Deanna Marriott, Laura Saslow. . . . 4179

Real-Time Exposure to Intersectional Minority Stressors and Alcohol Use: Protocol for an Ecological Momentary Assessment Study With Latinx and Non-Latinx Sexual Minority Youth (e87201)
Robert Rosales, Suzanne Colby, Kristina Jackson, Christina Lee, Jacob van den Berg, Madelyn Clancy, Ethan Mereish, Robert Miranda Jr. . . . 4 1 9 2

Cumulative Stress Burden and Association With DNA Methylation in Ethiopian American Immigrants: Protocol for a Community-Engaged, Biopsychosocial Study (e85971)
Lara Stone, Lili Ding, Meron Hirpa Kassa, Beakal Amsalu, Bruktawit Getnet, Hermella Eshete, Tesfaye Mersha, Lisa Vaughn. . . . 4214

Migration Effects on Cognition: Protocol for the Aging in Kerala Americans Research Study (e85493)
Kelly Cotton, Dristi Adhikari, Anne Ambrose, Emmeline Ayers, Helena Blumen, Mirnova Ceïde, VG Kumar, Mairim Melecio-Vazquez, Sanish Sathyan, Alben Sigamani, Kavita Sivaramakrishnan, Marnina Stimmel, Erica Weiss, Jie Yang, Joe Verghese. . . . 4231

Evaluation of the Positive Peers Mobile App for Supporting the Viral Suppression of Young People With HIV: Protocol for a Concurrent Mixed Methods Evaluation With Randomized Controlled Clinical Trial and Observational Cohort (e87601)
Mary Step, L Catania, Jennifer McMillen Smith, Steven Lewis, Yanis Bitar, Vinay Cheruvu, Kristen Berg, Jeffrey Hallam, Ann Avery. . . . 4247

Characterizing AIDS Drug Assistance Program Practices and Policies for Sustained Viral Suppression Using the Consolidated Framework for Implementation Research: Protocol for a Qualitative Study (e90008)
Kathleen McManus, Erin Rogers, Amber Steen, Amy Killelea, Tim Horn, Reanna Panagides, Steven Laymon, Jessica Keim-Malpass. . . . 4266

Combining Digital Cognitive Behavioral Therapy With Mindfulness Training for Binge Eating Disorder: Protocol for a Feasibility Trial (e91761)
Margaret Sala, Corey Roos, Dante Ascarrunz, Casey Stern, Jonathan Bricker, Jonathan Feldman, Cheri Levinson, Mihaela Aslan, Jennifer Hay, Hedy Kober. . . . 4277

Engaging Hospital Staff to Identify Levers for Adoption of Clinical Decision Support: Protocol for a Single-Site Case Study Using System Dynamics Group Model Building (e80848)
Nina Sperber, Lindsey Zimmerman, Lauren Caton, Tom Rust, Armando Bedoya, Sarah Haas, Lori Orlando, Ben Goldstein. . . . 4288

Synthesis and Standardization of Outcomes in Severe Malaria Treatment Trials: Protocol for the Development of a Core Outcome Set (the COSSMaT Study) (e78616)
Gideon Asamoah, Sharon Love, Diana Gibb, Kathryn Maitland, Elizabeth George. . . . 4296

Integrating Gait Analysis and AI Into Knee Osteoarthritis Care: Protocol for a 3-Phase Participatory Qualitative Study (e82860)
Owen Lindsay, Janie Astephen Wilson. . . . 4311

XSL•FO
RenderX

Application of the ADAPT Framework to Contextualize a Participatory Learning and Action Community Intervention for the Prevention and Control of Type 2 Diabetes Mellitus in Urban and Rural Settings in Afghanistan and Pakistan: Protocol for Intervention Adaptation (e71602)
Amber Tahir, Noor Sanauddin, Sara Imtiaz, Nadia Khaleeq, Mariam AbdeAli, Asima Khan, Farrukh Ahmed, Hannah Jennings, Rubia Zafar, Khalid Rehman, Abdul Shahab, Saima Afaq. . . . 4321

Chat-Based Decision Support System for the Maternal Health Journey in Assam, India: Protocol for a Mixed Methods Multiphase Implementation Study (e81873)
Ashita Munjral, Swapnil Rawat, Charlotte Warren, Jitender Nagpal, Ben Bellows, Aditi Aditi, Pompy Sridhar, Sowmya Ramesh. . . . 4334

Developing, Implementing, and Evaluating a Care Management Model for Patients With Neuromuscular Diseases: Protocol for a Participatory Mixed Methods Study (e82833)
Veronika Waldboth, Christina Schuler, Raffaella Willmann, Heidi Petry, Markus Weber, Barbara Grädel Messerli, Hannele Hediger, Iris Kramer, Gabriela Stettler-Nemecek, Martin Knoblauch, Georg Stettner. . . . 4346

Evaluating Experiences With the Newly Enacted Law on Assisted Suicide in Austria: Protocol for an Interdisciplinary Mixed Methods Study (e86740)
Klara Doppler, Tamina-Laetitia Vielgrader, Julia Fischer, Helene Mayer, Jana Hluch, Marlene Walser, Sabine Parrag, Stefan Dinges, Karl Stöger, Thomas Wochele-Thoma, Maria Kletecka-Pulker. . . . 4370

Directly Observing and Characterizing Adolescents' Self-Generated Social Media Posts: Protocol for Creation and Implementation of a Cyberethnography Informed Codebook (e84461)
Kylie Boyd, Lydia Bliss, Tingting Fan, Kayla Kern, Caitlin Carlson, Megan Moreno, Christopher Cascio, Ellen Selkie. . . . 4380

Undergraduate Interprofessional Education: Protocol for a Mixed Methods Study (e74394)
Hannah Snead, Dhwaniben Hada, Lori Eldridge. . . . 4398

Validity and Reliability of the Psychodynamic Organizational Diagnostic Instrument SyMOA: Protocol for a Mixed Methods Study (e77554)
Anne-Maria Müller, Sophia Sachs, Yannik Rieder, Claas Lahmann. . . . 4408

International Case Studies to Identify Success Factors and Contextual Conditions in the Digital Transformation of Health Care Systems and Derive Lessons for Germany: Study Protocol for a Mixed Methods Study (e80301)
Lena Kraft, Anna-Lena Brecher, Sophia Sgraja, Reinhard Busse, Volker Amelung. . . . 4417

Perinatal Health Care Among Climate Migrant Women: Protocol for a Scoping Review (e84176)
Giti Khalaj, James Plaisimond, Thierno Diallo, Lisa Merry, Steven Ouellet, Mohamed Ag Ahmed, Lara Maillet, Samira Amil, Marielle M'bangha, Erika Corona, Marie-Pierre Gagnon. . . . 4429

Evaluation Models for Undergraduate Nursing Clinical Skills: A Scoping Review Protocol (e79805)
Kateryna Metersky, Tommy Lin, Tina Lam, Areej Al-Hamad, Daria Romaniuk. . . . 4437

Factors Associated With Evolution of the Use of Medical Assistance in Dying: Protocol for a Scoping Review (e85963)
Isabelle Marcoux, David Lavoie, Amy Bergeron, Alexandra Beaudin, Simon Lemyre, Valérie Bourgeois-Guérin, Gina Bravo, Maude Hébert, Chantale Simard, Deborah Ummel, Catherine Perron, Marie-Eve Bouthillier. . . . 4444

Salutogenesis in Mexico and Latin America: Protocol For Scoping Review (e83495)
Igor Ramos Herrera, Rigoberto Cisneros Garcia. . . . 4453

School-Based Universal Mental Health Prevention Programs for Children Aged 6 to 9 Years in Low- and Middle-Income Countries: Protocol for a Scoping Review (e87105)
Cherith Langenhoven-Heber, Katherine Sorsdahl, Emily Garman, Claire van der Westhuizen. . . . 4461

XSL•FO
RenderX

Methodological Research Practices for the Adaptation of Measuring Instruments in Health Care: Protocol for a Scoping Review (e86949)
Ninive Gomes de Oliveira, Rosilene Ferreira, Andressa Bernardo Bueno, Flávia Camerini, Danielle Henrique, Cristiane Gallasch, Cintia Fassarella. . . . 4474

Methods of Analysis in Randomized Noninferiority Trials: Methodological Survey Review Protocol (e76876)
Daniel Xie, Sameer Parpia, Tyler McKechnie, Phillip Staibano. . . . 4483

Strengthening Undergraduate Medical Education for Inclusive Health Care for People With Down Syndrome and Intellectual and Developmental Disabilities in Medical Schools: Protocol for a Scoping Review (e80280)
Isabela Amaral de Almeida Bistafa, Marco Antonio Ribeiro Filho, Ana dos Santos Costa, Ana Pimenta Saad, Rauer Ferreira Franco, Janaína Aparecida de Sales Floriano, Emerson Roberto dos Santos, João de Souza Menezes, Matheus Querino da Silva, Daniele Nunes Longhi Aleixo, Eliana Fazuoli Chubaci, Andressa Karina Stefani, William Donegá Martinez, Pedro Belchior da Silveira Junior, Ana Mirtes Zecchin, Nathalia Gonçalves Cristóvão, Alex Bertolazzo Quitério, Renato Mendonça Ribeiro, Denise Cristina Móz Vaz Oliani, Antônio Hélio Oliani, Maysa Alahmar Bianchin, Luís Cesar Fava Spessoto, Rita Helú Mendonça Ribeiro, Júlio César André. . . . 4489

Public Engagement Strategies in Digital Health Ethics: Protocol for a Scoping Review (e86280)
Stella Namuganza, Mariacarla Gadebusch Bondio, Philipp Kellmeyer. . . . 4498

Clinical, Histopathological, and Molecular Prognostic Factors Associated With Survival and Disease Progression in Adult Patients With Primary and Secondary Thyroid Lymphomas: Scoping Review Protocol (e73164)
Ana Lemos-Rodriguez, Juan Lenis-Gonzalez, Oriana Arias-Valderrama, Elizabeth Arrieta, Guillermo Guzman, Andres Garcia. . . . 4507

Affective Health Messages for Adolescents: Protocol for a Scoping Review (e78927)
Siv Berg, Raul Ferrer-Conill, Kristine Stifjell, Kristin Hjorthaug Urstad, Jo Røislien, Lisbet Fjæran. . . . 4514

Understanding Responsible Development in AI-Based Clinical Prediction Models for Mortality: Protocol for a Scoping Review (e80325)
Riley Martens, Jessalyn Holodinsky, Jessica Simon, Zack Marshall. . . . 4523

Determinants of Street Food Hygiene Practices and the Effectiveness of Interventions for Vendors and Consumers in Low- and Middle-Income Countries: Protocol for a Scoping Review (e68633)
Abul Kamal, Tarique Nurul Huda, Sonjida Simi, Elizabeth Thomas, Farhana Sultana, Jesmin Sultana, Rehnuma Sarah, Soha Rafa, Om Gautam, Peter Winch, Mahbubur Rahman. . . . 4531

Inequity in Access to and Use of Digital Health Technologies in Routine Heart Failure Care: Protocol for a Scoping Review (e81949)
Nicklas Vinter, Mariam Elmegaard, Lars Køber, Morten Schou, Søren Johnsen. . . . 4542

Role of Shame and Concealment in Patient-Provider Communication Among Individuals With Low Health Literacy: Protocol for a Scoping Review (e83451)
Tsuyoshi Okuhara, Hiroko Okada, Rie Yokota. . . . 4552

Safety and Impacts of Physical Activity for Individuals Living With Hypermobility Spectrum Disorders and Hypermobile Ehlers-Danlos Syndrome: Protocol for a Scoping Review (e80394)
Donald Golden, Jonathan Low, Julia Daun, C Jones, Liz Dennett, Susan Culos-Reed, Ranita Manocha. . . . 4560

Mapping the Landscape, Knowledge Gaps, and Areas for Innovation in Brain Health and Dementia Research in Canada: Protocol for a Scoping Review of Reviews (e79020)
Juanita-Dawne Bacsu, Kiana Mero, Megan O'Connell, Megan Funk, Alixe Ménard, Myrna Norman, Sheila Blackstock, Jim Mann, Wendy Hulko, Melba D'Souza, Sarah Fraser. . . . 4572

XSL•FO
RenderX

Exploring the Well-Being, Adaptability, and Sense of Belonging of Undergraduate Nursing Students During the Transition From Simulation to Clinical Practice: Protocol for a Scoping Review (e86813)
Marjolaine Dionne Merlin, Monica McGraw, Julie Renaud, Sandrine Roy. 4578

Shared Decision-Making With a Surrogate for Life-Sustaining Treatment of Critically Ill Patients: Protocol for a Scoping Review (e83284)
Yoshiyasu Ito, Mika Moriyama, Akemi Nasu, Keiko Kamitani, Shimpei Hayashi, Satoko Ono, Yasuko Sumida, Keiko Matsumoto, Misae Ito. . . 4 5 8 5

Exploring Medical Students' Representations of Future Specialties and Parenthood: Protocol for a Scoping Review (e78133)
Sylvie Arnoux, Mia Gisselbaek, Georges Savoldelli, Nadia Bajwa. 4593

Understanding the Nurse Champion Concept and the Training Initiatives for Nurse Champions: Protocol for a Scoping Review (e80204)
Wilmer Santos, Ian Graham, Amanda Vandyk, Gillian Harvey, Janet Squires. 4601

Efficacy and Safety of Sodium-Glucose Cotransporter-2 Inhibitors in Patients With Refractory Cirrhotic Ascites and Diabetes: Scoping Review Protocol (e77587)
Tooba Laeeq, Ayesha Kamal, Arbab Kasi, Umm Shabbar Banatwala, Aniqa Baloch, Roberto Sagaribay, Vidhani Goel, Aditi Singh, Kavita Batra. 4617

Clinical Evidence Profile of Oral Chinese Patent Ethnomedicines: Protocol for a Scoping Review and Evidence Map (e77741)
Miaomiao Li, Hui Luo, Hui Zhao, Xing Liao. 4626

Mapping Algorithmic Bias in AI-Powered Electrocardiogram Interpretation Across the AI Life Cycle: Protocol for a Scoping Review (e82486)
Luqman Lawal, Christopher Paton, Mike English, Bruno Holthof, Tabitha Preston. 4639

Characteristics of Mobile Health Interventions for Repetitive Negative Thinking: Protocol for a Scoping Review (e72138)
Judith Martens, Thomas Forkmann, Inken Höller. 4648

Determinants of Equitable Data Governance for African, Caribbean, and Black Communities in Health Research in High-Income Countries: Protocol for a Scoping Review (e82403)
Josephine Etowa, Shamara Baidoobonso, Doris Kakuru, OmiSoore Dryden, Chinedu Oraka, Egbe Etowa, Peter Farrell, Sandra Mba, Akalewold Gebremeskel, Arone Fantaye. 4658

The Use and Utility of Wearable Devices for Tracking Sleep and Activity in Inpatient Mental Health Settings: Protocol for a Rapid Review (e82290)
Gillian Strudwick, Hwayeon Shin, Iman Kassam, Karishini Ramamoorthi, Benjamin Rosen, Michelle Patriquin, Justin Baker, Nicholas Neufeld. 4 6 7 0

Managing Metabolic Dysfunction–Associated Steatotic Liver Disease: Protocol for a Scoping Review of Patient Perceptions, Barriers, and Facilitators (e81404)
Sikyeong Park, Yu Park, Dahye Hong, Bada Kang. 4676

Methodology for Assessing University Members' Health and Well-Being and Their Influencing Factors: Protocol for a Rapid Scoping Review of Survey-Based Research (e83845)
Varga Zsuzsanna, László Balkányi, Anikó Gyulai, László Lippai, Andrea Rucska†, Zsuzsanna Kiss, Péter Varsányi, József Vitrai, Boróka Gács. 4 6 8 3

A Mobile Self-Assessment and Referral Platform for Family Caregivers of Individuals With Alzheimer Disease and Related Dementias: Protocol for a Pilot Randomized Controlled Trial (e90244)
Francesca Falzarano, Annabelle Greenfield, Hannah Mason, Katerina Bumbalova, Elizabeth Rojas, Aleksandar Bumbalov. 4693

XSL•FO
RenderX

A Need for Strategies to Reduce Alcohol Use After Metabolic and Bariatric Surgery: Technology-Based Intervention and Study Protocol for a Pilot Randomized Controlled Trial (e80068)
Lisa Miller-Matero, Celeste Pappas, Brittany Christopher, Roman Grossi, Alyssa Vanderziel, Nancy Barnett, Roland Moore, Sarah Bendit, Aaron Hamann, Arthur Carlin, Oliver Varban, Jordan Braciszewski. 4714

Interactive Internet-Based Motivational Interviewing Training for HIV Counseling Support Staff to Improve Health Communication in HIV Care Interactions: Protocol for Training Development and a Pilot Randomized Controlled Trial (e82241)
Iván Balán, Onna Brewer, Bryan Kutner, Rebecca Giguere. 4724

Developing a Multimodal Screening Algorithm for Mild Cognitive Impairment and Early Dementia in Home Health Care: Protocol for a Cross-Sectional Case-Control Study Using Speech Analysis, Large Language Models, and Electronic Health Records (e82731)
Maryam Zolnoori. 4740

Pilot Test of Diabetes Screening and Referral to Primary Care Within an Alaska Native Dental Clinic: Protocol for a Feasibility Study (e75619)
Kathryn Koller, Julie Beans, Diane King, Vanessa Hiratsuka, Gretchen Day, Shiela Strauss, Barbara Stillwater. 4758

Developing the Link-me+EMPHN Mental Health Model of Care to Improve General Practitioner Capacity for Mental Health Care in Australian Primary Care: Protocol for a Mixed Methods Formative Study (e79560)
Amy Coe, Jane London, Anne-Marie Martin, Aaron van Ree, Kirsty Lembke, Caroline Johnson, Bridget Bassilios, Catherine Kaylor-Hughes. .
4 7 6 9

Corrigenda and Addendas

Correction: Cumulative Stress Burden and Association With DNA Methylation in Ethiopian American Immigrants: Protocol for a Community-Engaged, Biopsychosocial Study (e95656)
Lara Stone, Lili Ding, Meron Kassa, Beakal Amsalu, Bruktawit Getnet, Hermella Eshete, Tesfaye Mersha, Lisa Vaughn. 3895

Correction: Implementation of a Personalized Medicine Approach in Patients With Type 2 Diabetes Mellitus Receiving Multiple Daily Insulin Injections (POMA Project): Protocol for a Before-and-After Intervention Study (e95267)
Rosa Giné-Balcells, Bogdan Vlacho, Olga Vidal-Pérez, Albert Lecube, Josep Franch-Nadal, Dídac Mauricio, Àngels Molló-Iniesta, Marta Hernández, POMA Study Group. 4357

Correction: Evaluation of the Comparative Efficacy of an Ayurvedic Formulation (Nimba-Amalakyadi Powder) vs Metformin in the Management of Type 2 Diabetes Mellitus: Protocol for a Pilot Study (e96344)
Punam Sawarkar, Gaurav Sawarkar. 4359

Correction: Smartphone App–Based Eating Behavior Monitoring and Feedback Intervention for Glucocorticoid-Induced Appetite Increase in Patients With Systemic Lupus Erythematosus: Protocol for a Pilot Randomized Controlled Trial (e90010)
Takashi Yamaguchi, Nobuyuki Takahashi, Ryohei Inanaga, Ryuhei So, Hiroe Kikuchi, Hisashi Noma, Hiroyuki Sasai, Kazuro Kamada, Tomohiro Sugimoto, Hiroshi Tsushima, Takanori Ichikawa, Hirofumi Miyake, Shunichi Fujita, Keisuke Ono, Yusuke Miwa, Anna Hasegawa, Naoki Suzuki, Akira Onishi, Toshihiro Matsui, Ryu Watanabe, Yasuhiro Hasegawa, Rei Ono, Takeo Isozaki, Yuichi Ishikawa, Nobuyuki Yajima, Noriaki Kurita.
4 3 6 1

Correction: Deep Learning–Based Pattern Recognition for Detecting Penile Abnormalities: Protocol for Developing a Mobile App for Circumcision Eligibility (e89351)
Irfan Wahyudi, Chandra Utomo, Samsuridjal Djauzi, Muhamad Fathurahman, Gerhard Situmorang, Arry Rodjani, Putu Raharja, Kevin Yonathan, Budi Santoso, Dwidian Khresna, Marco Raditya. 4363

XSL•FO
RenderX

Correction: Vaping, Acculturation, and Social Media Use Among Mexican American College Students: Protocol for a Mixed Methods Web-Based Cohort Study (e95651)
Bara Bataineh, C Marti, Dhiraj Murthy, David Badillo, Sherman Chow, Alexandra Loukas, Anna Wilkinson. 4365

Correction: Real-World Effectiveness and Safety of Liuwei Dihuang Pill for Menopausal Syndrome: Protocol for a Prospective, Observational, Multicenter Cohort Study (e84867)
Shuoshuo Wei, Hongyan Zhang, Mengmeng Wang, Yuhui Sun, Tao Song, Weifang Zheng, Yunmei Ye, Liran Ma, Yunshuang Yang, Tong Wei, Lan Cheng, Chunquan Sun, Xingting Ding, Shichu Zhao, Qingqiao Song, Lianxin Wang. 4367

Proposal

All-Hazards Vulnerability and Adaptation Assessment of Canadian Kidney Care Systems: Protocol for a Qualitative Study (e90059)
Saly El Wazze, Shinjini Mondal, Shaifali Sandal. 3910

Letters to the Editor

Bridging Neurofeedback and Structural Connectivity in Chronic Pain (e87420)
Alaeddin Acar, Diaa Yahya, Eray Tekirda 4785

Authors' Reply: Bridging Neurofeedback and Structural Connectivity in Chronic Pain (e89007)
Luke Bialostocki, Divya Adhia, Damith Mudiyanselage, Mark Smith, Yusuf Cakmak, Dirk De Ridder, Ramakrishnan Mani, Jerin Mathew. 4787

XSL•FO
RenderX

The Effect of an Educational Intervention on the Knowledge and Self-Care Skills of Adult Patients With Venous Leg Ulcers: Protocol for a Randomized Controlled Trial

Cristina Canova[1*], PhD; Jessica Scappin[1*], MSc; Maria Vittoria Chiaruttini[1], PhD; Laura Tasson[1], MSc; Flavia Masiero[2], BSc; Claudia Varagnolo[1], BSc; Matteo Martinato[1], PhD

[1]Department of Cardiac Thoracic Vascular Sciences and Public Health, Unit of Biostatistics, Epidemiology and Public Health, University of Padova, Via Loredan, 18, Padova, Italy

[2]Azienda ULSS 6 Euganea, Padova, Italy

[*]these authors contributed equally

Corresponding Author:
Cristina Canova, PhD
Department of Cardiac Thoracic Vascular Sciences and Public Health, Unit of Biostatistics, Epidemiology and Public Health, University of Padova, Via Loredan, 18, Padova, Italy

Abstract

Background: Venous ulcers have a prevalence of 1% in Western countries, with a tendency to become chronic and recur frequently. The lengthy periods of healing required for the treatment of these lesions result in increased costs. It is possible that patient education may enhance patients' understanding of their health condition, which could potentially lead to a reduction in recurrence rates. Nevertheless, it is uncertain whether this constitutes the optimal educational strategy.

Objective: The objective of this study is to assess the superiority of an educational intervention delivered via video compared to an informative brochure in outpatient nursing clinics, focusing on patients' knowledge of venous leg ulcers and their self-care skills.

Methods: A total of 144 patients with venous leg ulcers who attend the outpatient nursing clinics will be recruited to participate in this randomized controlled trial. Individuals with cognitive impairment, an ankle-brachial index of less than 0.90, arterial insufficiency, and arterial ulcers or ulcers of mixed origin will be excluded from the study. The intervention comprises 4 short educational videos specifically created for this study, in which patients will be provided with information on disease knowledge, medication management, compression therapy, and the importance of adequate nutrition and physical activity. The control group will be provided with the standard education offered at the outpatient clinics participating in the study, which will be disseminated through an information leaflet. Patient assessments will be performed with standardized, validated tools at baseline; at 1, 3, and 6 months postenrollment; and at 12 months via telephone follow-up. The primary outcomes are the patient's level of knowledge and their ability to self-care. The secondary outcomes are the patient's level of well-being, the success rate of the treatment, and the recurrence rates. A preliminary feasibility study was conducted to methodologically refine the protocol.

Results: A feasibility study was conducted to refine study procedures and data collection tools, confirming the acceptability of the questionnaires and the feasibility of digital data collection. The randomized controlled trial was funded in June 2025. It is currently pending ethics approval, and it is expected to start recruitment in spring 2026. Results from the main trial are expected to be available by the end of 2027.

Conclusions: A video-based educational intervention may prove to be an effective, readily accessible, and replicable method of patient education. The findings of this study will inform the most effective method of patient education for this population, ensuring the delivery of high-quality care and enhancing the quality of services.

Trial Registration: ClinicalTrials.gov NCT07031180; https://clinicaltrials.gov/ct2/show/NCT07031180

International Registered Report Identifier (IRRID): PRR1-10.2196/84211

(JMIR Res Protoc 2026;15:e84211) doi:10.2196/84211

KEYWORDS

education; venous leg ulcer; randomized controlled trial; knowledge; self-care skills

XSL•FO
RenderX

Introduction

Venous leg ulcers (VLUs) are defined as skin lesions on the lower leg or foot with a healing time of more than 6 weeks [1]. These dermatological conditions occur in the lower extremities as a consequence of elevated venous pressure, leading to vascular dilation and fluid stagnation in the affected areas. VLUs are a complication of chronic venous insufficiency, which causes damage to the vessel walls and surrounding tissues. They account for 70% of all leg ulcers [1]. In addition to venous insufficiency, several factors, such as age, reduced mobility, and obesity, have been identified as contributing to the development of VLUs [2].

The incidence of VLUs is increasing among adults, particularly among older adults who often have concomitant diseases. It affects 1% of the global population between the ages of 18 and 64 [3]. In the United States and other Western countries, the prevalence of this condition is estimated to be approximately 3% among the older adults [4-6].

In clinical practice, the treatment of VLUs typically involves prolonged healing periods and a high rate of recurrences [7,8]. A review of the literature shows that 22% of individuals who recover from VLUs will experience a recurrence within 3 months, 57% within 1 year, and up to 78% within 3 years. Furthermore, the risk of recurrence is 4.4 times higher in patients with a history of multiple VLUs [1], leading to increased treatment needs and health care costs. The diagnosis, treatment, and cure of VLUs account for up to 2% of the annual health care budget of Western societies [6].

A wound that persists for more than 1 month without following the normal healing process is classified as a chronic wound [3]. This chronicity is associated with a diminished quality of life. Patients frequently report a range of symptoms and experiences, including pain, sleep disturbances, malodor, impaired mobility, altered body image, reduced energy levels, and dissatisfaction with treatment. These symptoms and experiences can lead to a significant decline in perceived quality of life, encompassing physical, psychological, and social domains [1,9,10]. Additionally, VLUs can result in leg weakness, swelling, exudation, and recurrent infections [4].

Given the chronic and prevalent nature of these ulcers, patients or their caregivers frequently take on the responsibility of managing and providing self-care for VLUs. This involves maintaining health, monitoring the condition, and managing any issues that arise. The maintenance phase focuses on the prevention of injuries or recurrences through the reduction of risk factors. Monitoring involves observing signs and symptoms, while management encompasses the patient's response to any observed signs and symptoms [8]. A significant barrier to effective VLU self-management is the lack of adequate health information [11]. This gap in knowledge limits the ability of patients or their caregivers to maintain health, monitor the condition, and manage the wound effectively. Therefore, it is crucial to ensure that patients are fully informed about their health condition, associated risk factors, potential complications, protective measures, and available treatment options for managing symptoms.

Noninvasive treatments for VLUs include the application of local dressings to the wound and surrounding skin, as well as the use of compression therapy. The literature shows that compression therapy is a more effective method for promoting healing compared to the application of local dressings alone, resulting in a reduction of the wound surface area by 30% [4]. Furthermore, the proper application of compression therapy has been demonstrated to reduce pain, alleviate pressure, and decrease volume overload in the venous system [12], thereby improving venous and lymphatic return [4]. Given the high recurrence rates at 12 months postwound healing, it is crucial to maintain preventive measures, particularly compression therapy, in order to mitigate the risk of recurrence [13,14].

Unfortunately, a significant proportion of patients either do not adhere to or only partially adhere to compression treatment [15]. Nonadherence to recommended compression therapy is associated with a significant increase in the risk of delayed wound healing and recurrence, with estimates ranging from 2 to 20 times higher [5]. Adherence is influenced by a number of factors, including receiving conflicting information [16] or a lack of clear, objective, and easily understandable information [17]. Additionally, adherence is influenced by the individual's beliefs about the effectiveness of the treatment, their awareness and knowledge of the condition [14], and the level of patient involvement in the management of their VLU [3]. A deeper understanding of their condition empowers patients to engage in effective self-care and avoid risky behaviors [8]. Therefore, it is crucial to promote adherence to continuous compression therapy through patient education [17] to enhance knowledge and prevent recurrence.

However, the consistency of patient education across different health care settings remains variable, and there is a lack of robust scientific evidence regarding the effectiveness of interventions [14,18]. The educational interventions provided to patients with VLUs in outpatient clinics often lack clear and consistent guidelines for delivery, with the approach being dependent on the individual health care provider. In cases of high workload, a standardized tool is not always guaranteed, and education may be completely absent, as reported by the clinicians involved in this care setting.

A method for patient education described in the literature involves face-to-face sessions lasting 20 to 60 minutes, conducted at the patient's home or in the outpatient clinic, with or without written material [1]. However, given the high prevalence of VLU, it is often not feasible to carry out such lengthy sessions due to staffing limitations. Another effective educational method involves participation in a community of VLU or healed patients, where experienced VLU nurses provide education. However, transportation issues and limited availability make it difficult for all patients to attend these meetings [19].

Another potential strategy for the dissemination of information and the promotion of awareness is the use of manuals, brochures, or small books. These resources can facilitate the dissemination of knowledge and enhance the understanding of the general population, enabling patients and their families to refer back to them multiple times and support decision-making in the

XSL•FO
RenderX

management of VLU [20]. A study conducted in Italy at a vulnology clinic for ambulatory patients with VLUs demonstrated that providing an informational brochure can improve patient education, particularly in self-care capacity and disease knowledge [15]. Conversely, a study conducted by Baquerizo Nole et al [21] suggests that illustrative video recordings may be even more effective in enhancing patient knowledge compared to brochure-based education.

Indeed, educational videos are easily accessible, can be displayed in clinic waiting rooms, do not require the ability to read or comprehend written text, and also ensure that patients receive accurate and consistent information, reducing communication errors attributable to the disparate experience of health care providers or potential information distortion. In addition, video-based education could facilitate personalized care by allowing patients to request further information or specific advice from professionals. This enables nurses to prioritize delivering more relevant guidance instead of repeating information the patient may already know. These considerations make them a potentially superior method of education. However, the results of previous studies have not been statistically significant due to the limited number of enrolled patients [21].

On these premises, a project was developed to assess the superiority of an educational intervention delivered via video compared to an informational brochure in outpatient nursing clinics, focusing on patients' knowledge of VLUs and their self-care skills. The aim of this article is to present the protocol for this clinical trial.

Methods

Study Design

Study Setting and Design

The study is a randomized, controlled, single-blind trial with 2 parallel groups. The aim is to assess the effectiveness of the experimental intervention in increasing the level of knowledge and self-care skills of patients with VLUs compared to the standard of care provided in outpatient nursing clinics, where education is delivered through an informational brochure. The intervention comprises an educational video on VLU management. Patients are randomized using block randomization in order to prevent imbalance between the intervention and control groups over the course of the 12-month study period.

As a parallel project, a longitudinal prospective substudy will be dedicated to validating the Italian version of the Well-being of Wounds Inventory (WOW-I) questionnaire.

The protocol was developed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 statement [22].

Eligibility Criteria and Exclusions

Patients attending the outpatient nursing clinics involved in the study will be enrolled. To be eligible to participate in the study, patients must meet all of the criteria set out in Textbox 1. These include being aged 18 years or older, having a VLU, understanding Italian, being able to read and comprehend the questions, providing consent to participate in the study, and having access to a smart device with internet connectivity or a caregiver able to provide such a device. Individuals with cognitive impairment, an ankle-brachial index of less than 0.90, arterial insufficiency, or ulcers of arterial or mixed origin will be excluded from participation.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Aged 18 years or older
- Affected by venous leg ulcer
- Able to understand the Italian language
- Able to read and understand the questions
- Attending the nursing clinics involved in the study
- Provided consent to participate in the study
- Able to use a smart device with internet access or supported by a caregiver able to use a smart device as a video player

Exclusion criteria

- Cognitive impairment
- Ankle-brachial index less than 0.90 or arterial insufficiency
- Arterial or mixed-origin ulcers

Randomization and Blinding

After obtaining informed consent and confirming eligibility criteria, randomization will be conducted using the REDCap (Research Electronic Data Capture; Vanderbilt University) platform. Patients will be randomly assigned to 1 of 2 parallel groups in a 1:1 ratio, using block randomization to ensure balance between the 2 groups during the follow-up.

XSL•FO
RenderX

The study is single-blind. Two separate teams, blind and unblind staff, will be created and assigned different activities to be performed. Nurses belonging to the blind staff will be responsible for assessing patients and delivering a blinded QR code, thereby preventing them from knowing which intervention is assigned to each participant. The unblind staff will have access to all other information except clinical assessments and will serve as the appointed contact point for patients and their caregivers regarding any study participation–related issues, thus preventing any potential risk of unblinding. Patients' and caregivers' questions (related to study participation or clinical care) will be directed only to the unblind staff. Both the blind staff and the patients or caregivers will be informed of this logistic before the study starts and reinforced during the study conduction and follow-up.

Experimental Treatment

The experimental educational intervention consists of 4 short, professionally produced videos covering the following topics: VLU knowledge, medication management, compression therapy, and diet and physical activity.

The content of the videos was derived from a thorough literature review and further refined by 2 expert nurses specializing in wound care. The videos were recorded by an experienced videographer in collaboration with a patient volunteer, 2 nursing students, and 2 nurses with expertise in managing VLUs.

In each video, the presenter provides clear and accessible explanations, using language that is easily comprehensible to the general public. When medical terminology is used, a definition is provided. Additionally, key terminology used by the speaker is displayed on the screen, accompanied by realistic images and scenes relevant to the discussed topics, enhancing comprehension through visual representation.

The first video is approximately three and a half minutes long and provides an overview of VLUs, their anatomical location, the associated risk factors, the mechanisms of development, and the diagnostic criteria.

The second video, lasting approximately 5 minutes, provides an in-depth exploration of the importance of treating VLUs, various management and treatment approaches, and potential complications.

The third video, 7 minutes long, explains the concept of compression therapy, its benefits and contraindications, the proper technique for applying elastic stockings, and lifestyle changes, including advice on physical activity to prevent the development of ulcers.

The fourth video, which lasts about 2 minutes, presents a list of foods and nutrients that promote the healing of VLUs and help prevent their recurrence.

The videos were uploaded to a YouTube channel accessible only to individuals possessing the requisite link or QR code. This was done to prevent contamination and ensure the integrity of the study. The platform allows the researcher to determine the number of views, which serves as an indicator of adherence to the experimental treatment. To track the views of each participant, every person in the experimental group will receive the same videos but will view a differently uploaded version.

Consequently, each video link or QR code will be associated with an alphanumeric identification number assigned to a participant, a strategy that will allow researchers to link the questionnaires completed by each participant to the number of views of the individual videos. Participants will be required to retain the link or QR code for viewing the intervention, as it will differ from the links or QR codes of other participants in the study.

Video Intervention

A document containing an identification number and QR codes or links will be provided to participants who have been assigned to the intervention group. The identification code, which the patient is required to retain and keep secret, is alphanumeric and random. The code is used by the participant to complete the questionnaires at baseline and during the follow-up periods, with the intention of guaranteeing anonymity. The QR codes and links serve the purpose of enabling the patient to access the 4 experimental educational videos. Patients assigned to this group will be required to access the videos via QR codes or links and view them after completing the baseline questionnaires or within the initial 7-day enrollment period. The educational intervention can be viewed in the outpatient nursing clinic after the visit, in a waiting room, with one's own smart device or a device provided by the company, or at the participant's home.

The videos may be viewed on multiple occasions during the study period. The outpatient nurse will issue a reminder to the patient to view the intervention at least once.

Control Treatment

Similarly, the control group will receive a document containing an alphanumeric identification number and a link or QR code. The identification code will be used in the same way as it is used for the intervention group, allowing participants to complete the questionnaires at the baseline and follow-up stages. The QR code will instead direct participants to a printable educational brochure. This brochure represents the standard treatment. It was introduced following a study conducted in 2021 that demonstrated its efficacy in increasing patient compliance, self-care skills, and disease knowledge in patients with VLUs [15]. Since then, it has been routinely used to educate VLU patients in the study settings. The educational content in the brochure covers some of the topics presented in the educational videos but in a short form, such as the nature of venous ulcers, the available treatments and compression therapy, dressing management, potential risks and complications, and beneficial lifestyle changes. The brochure will contain both written text and illustrative images and will be available to participants throughout the duration of the study.

Baseline Questionnaires

Patients who meet the eligibility criteria and consent to participate in the study will be requested by the researcher to provide sociodemographic and health-related data, including age, gender, level of education, occupation, with whom they live, presence of private or family care, preexisting pathologies,

XSL•FO
RenderX

smoking habits, presence of pain, venous ulcer–related information, and medication management. The researcher will input these data and the Pressure Ulcer Scale for Healing (PUSH) Tool 3.0 measurement scale into the REDCap platform using a tablet provided by the company. To ascertain the level of knowledge about the disease, the level of self-care, and the patient's well-being, the patient will be required to complete, respectively, the validated Venous Leg Ulcer Knowledge (VLUK) measurement scale, the validated Venous Leg Ulcer Self-Efficacy Tool (VeLUSET) measurement tool, and the WOW-I scale, using the company's or their own smart device.

Follow-Up

Once the baseline questionnaires have been completed, patients will be invited to view the educational intervention assigned to them. Subsequently, at 4-week, 3-month, and 6-month intervals, participants will be required to complete the same questionnaires as at baseline. The VLUK, VeLUSET, and WOW-I scales have been employed to permit a comparison of pretreatment and posttreatment scores. Additionally, the outpatient nurse will be assigned to evaluate the patient's venous ulcer using the PUSH Tool 3.0.

The final follow-up is scheduled to occur 12 months after enrollment and will be conducted via a telephone interview. Patients will be asked whether the condition has healed and, if so, for how long. Additionally, patients will be asked whether the ulcer has recurred.

Setting and Data Collection Method

The study will be conducted in outpatient clinics located in a northern Italian region.

The data will be collected by the nurses of the outpatient clinics involved in the study. Their duties will include checking eligibility criteria, collecting sociodemographic and health data at baseline, distributing the paper containing the unique code and QR codes for accessing the educational intervention and for completing the other questionnaires, and entering the data into the PUSH Tool 3.0 at baseline and at each follow-up. The remaining questionnaires (VLUK, VeLUSET, and WOW-I) are to be completed by the participants themselves at the outset of the study and at each subsequent follow-up. Data collection forms will be accessible through the electronic case report form. Data will be collected anonymously via the REDCap platform on a smartphone or tablet provided by the company or the patient.

Monitoring and Data Management

The researcher will be responsible for the collection of data at the time of the enrollment of the patient, either directly or through adequately trained personnel. Furthermore, the researcher will provide guidance to the nurse on the correct collection of data during the follow-up period, specifically in relation to ulcer assessment and the patient's exit from the study.

It is not possible to exercise control over the questionnaires completed by patients themselves, as they are completed anonymously. During outpatient visits, the nurse will remind the patient or caregiver to view the educational intervention at least once by scanning the QR code and to complete the questionnaires during follow-ups.

The management, processing, and statistical analysis of the data will be conducted by external parties in collaboration with the sponsor. These parties will retain the data only for the duration necessary to complete the tasks assigned to them.

Outcomes Measures

Primary Outcomes

The 2 coprimary outcomes of this study are an improvement in patient knowledge about their disease (VLUK score) and improvement in self-care skills (VeLUSET score) at 4 weeks.

Secondary Outcomes

Secondary outcomes (descriptive) include (1) VLUK and (2) VeLUSET scores at 3 and 6 months, (3) the evaluation of the healing process (PUSH Tool 3.0 score), and (4) improvements in well-being (WOW-I score) at 1, 3, and 6 months.

Moreover, the number of VLU recurrences in participants and the number of patients with complete healing will be assessed at 12 months of follow-up.

The instruments used to investigate the primary outcomes and well-being are self-administered questionnaires, while the healing process will be evaluated by the nurse using the PUSH Tool 3.0. The questionnaires and tools will be administered at the beginning of the study and at subsequent follow-up points, as described below in the study timeline.

Finally, the presence of any recurrence and complete healing will be investigated at the 12-month follow-up point through a single question asked by the nurse via telephone. No specific tools will be employed. The previously described evaluations will be conducted in both the intervention and control groups (Figure 1).

XSL•FO
RenderX

Figure 1. Study flowchart. PUSH Tool 3.0: Pressure Ulcer Scale for Healing Tool 3.0; VeLUSET: Venous Leg Ulcer Self-Efficacy Tool; VLUK: Venous Leg Ulcer Knowledge; WOW-I: Well-Being of Wounds Inventory questionnaire.

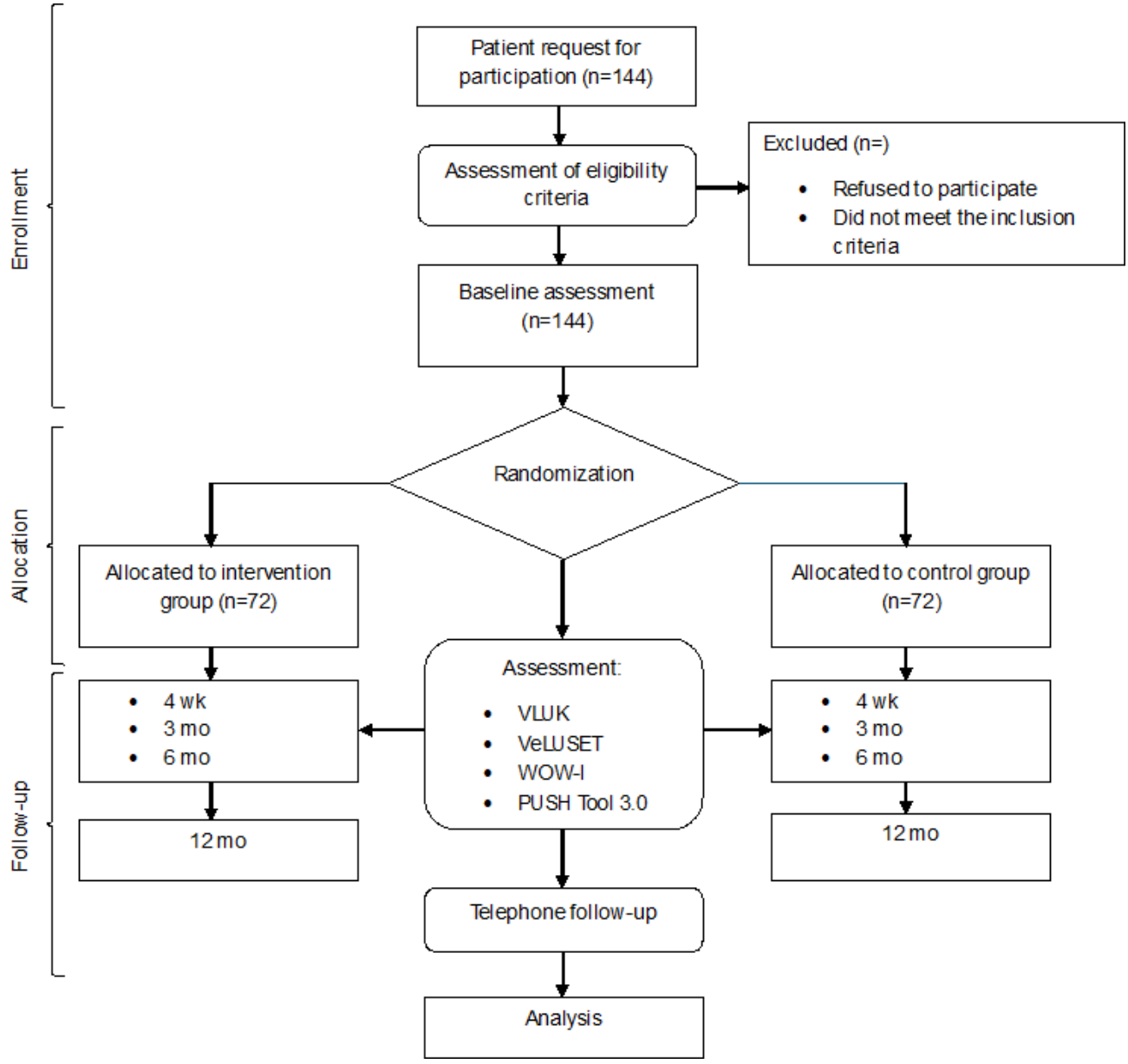

Description of Questionnaires and Tools

Venous Leg Ulcer Knowledge

Patient knowledge about the disease will be assessed using the questionnaire developed for the study "Educational Interventions in Patients with Venous Leg Ulcer Patients" (VLUK) tool, a validated instrument in the Italian context [6]. The VLUK assessment tool comprises 15 multiple-choice questions, each with one correct answer, designed to evaluate the patient's knowledge of VLU pathophysiology, treatments, and recommended lifestyle changes [15].

Venous Leg Ulcer Self-Efficacy Tool

Patients' self-care skills will be assessed using the VeLUSET, which has been validated in the Italian context in another unpublished study [23]. This tool assesses 5 dimensions of self-care in people with VLUs. The tool comprises 30 items, which can be grouped into 5 areas: general self-care (5 items), daily self-care activities (12 items), normal living (4 items), skills development (6 items), and trauma avoidance (3 items) [16,24].

Well-Being of Wounds Inventory

Patients' well-being will be evaluated using the WOW-I scale, which will be translated and validated in the Italian context.

This scale assesses the well-being of the patient with VLUs in relation to their overall quality of life, their social circumstances, and the specific experience of the wound. The scale comprises 2 subscales, the first of which includes 2 factors or categories and the second of which has 5 factors or categories, resulting in a total of 19 items, each evaluated using a Likert scale [25].

PUSH Tool 3.0

The healing process will be documented within the study using the PUSH Tool 3.0, introduced in 1997 by the National Pressure Ulcer Advisory Panel and demonstrated to be an effective

XSL•FO
RenderX

method for monitoring the healing of venous ulcers [26]. The tool was selected for its simplicity, speed, and reliability in monitoring venous ulcer healing. It comprises 3 items: lesion size (length × width), exudate quantity (none, light, moderate, and heavy), and tissue type (necrotic tissue, slough, granulation tissue, epithelial tissue, and closed). The total lesion score is obtained by summing the single-item scores, which range from 0 to 17. A reduction in the total score is indicative of progress toward the healing process.

If a patient achieves complete healing before the conclusion of the study, they will be classified as a "healed patient" and contacted again via telephone for follow-up purposes.

Validation of the Italian Version of the WOW-I Questionnaire

To psychometrically validate the Italian version of the WOW-I, patients will be invited to complete the WOW-I questionnaire, the "EuroQol-Visual Analog Scale" [27], and the well-being subscale of the "Cardiff Wound Impact Schedule" [2], at baseline and 1 week later. Internal consistency will be assessed using Cronbach α at both time points. Stability over time will be evaluated by mean differences in test-retest scores, particularly in relation to well-being levels assessed via the EuroQol-Visual Analog Scale. Convergent validity will be estimated by correlating WOW-I subscales with the Cardiff Wound Impact Schedule well-being subscale.

Study Timeline

The study will be conducted over a period of 1 year from the end of enrollment. After enrollment (t_0), the investigator will administer the WOW-I questionnaire to the patient at 7 ± 2 days (t_1) to verify the stability of the instrument, and then 3 other follow-up assessments are planned for the primary outcomes. The second follow-up evaluation (t_2) is scheduled to take place 4 weeks after the educational video is delivered, with subsequent follow-up assessments (t_3) at 3-month intervals and 1 at 6 months (t_4). If a patient is healed before the end of follow-up, they will be classified as a "healed patient."

Finally, a long-term telephone follow-up will be conducted 12 months (t_5) after enrollment (Table 1).

Table . Summary of time points, interventions, assessments, and data collected.

	Enrollment	Postallocation				Closeout
	t_0	t_1 (1 wk)	t_2 (4 wk)	t_3 (3 mo)	t_4 (6 mo)	t_5 (12 mo)
Enrollment						
Eligibility screen	✓					
Informed consent	✓					
Allocation	✓					
Interventions						
Intervention extra: video						
Intervention standard: brochure						
Assessments						
Sociodemographic factors	✓					
Health factors	✓					
PUSH Tool 3.0[a]	✓		✓	✓	✓	
VLUK[b]	✓		✓	✓	✓	
VeLUSET[c]	✓		✓	✓	✓	
WOW-I[d]	✓	✓	✓	✓	✓	
Ask if there has been healing or if there is a recurrence						✓

[a]PUSH Tool 3.0: Pressure Ulcer Scale for Healing Tool 3.0.

[b]VLUK: Venous Leg Ulcer Knowledge.

[c]VeLUSET: Venous Leg Ulcer Self-Efficacy Tool.

[d]WOW-I: Well-being of Wounds Inventory questionnaire.

XSL•FO
RenderX

Sample Size Calculation

The primary end points are the mean differences in pre-post scores (collected after 4 wk of follow-up and at baseline) of the VLUK and the VeLUSET questionnaires, which are considered coprimary hierarchical end points. This implies that the mean difference in VLUK within the treatment group must be statistically greater than the mean difference in the control group before the mean difference in VeLUSET between the control and treatment groups is evaluated. An interim analysis has been planned at the halfway point of the enrollment. The interim analysis assumes the Lan-DeMets O'Brien-Fleming approximation alpha-spending function and includes an early efficacy stop in case $Z>2.54$. The nominal P values are .006 and .048 at the interim and final analyses, respectively. Moreover, the conditional power will be calculated, and resampling will be performed in case of 50%<conditional power<80%. Otherwise, the sample size will be kept at the original size (72 patients per group), as described in the following. Particularly, a type I error of 5%, a power of 80%, and a dropout rate of 10% have been considered for the sample size estimation.

Statistical Analysis

A descriptive analysis will be conducted to summarize the baseline characteristics of the 2 randomized groups. Continuous variables will be analyzed using means and standard deviations (assuming normality) or medians and interquartile ranges. Categorical variables will be analyzed using proportions and frequencies.

Primary analysis will be conducted according to the intention-to-treat principle. The 2 coprimary outcomes, change in VLUK from baseline and VeLUSET scores at 4 weeks, will be analyzed using generalized linear models with a Gaussian family and identity link, corresponding to an analysis of covariance framework. For each outcome, the post-intervention score at 4 weeks will be used as the dependent variable, with the treatment group as the main independent variable and the corresponding baseline score included as a covariate. In the event of an imbalance between the 2 groups, multivariate generalized linear models will be employed. The coprimary outcomes will be evaluated using a hierarchical testing strategy to control the family-wise type I error rate at 5%. Specifically, the between-group difference in VLUK will be tested first; only if this comparison reaches statistical significance will the coprimary outcomes be assessed at the same significance level (ie, 5%). The secondary outcomes assessed at follow-up will be analyzed, accounting for the repeated-measures structure of the data, with time treated as a categorical variable, using appropriate parametric methods for longitudinal data. When relevant, within- and between-group comparisons over time will be performed.

Comparisons of categorical outcomes, including healing status and recurrence, will be conducted using chi-square or Fisher exact tests, as appropriate. Secondary analyses will be considered exploratory, and no formal adjustment for multiplicity will be applied beyond the hierarchical testing of the coprimary end points.

The software R version 4.4.2 (2024-10-31 UCRT) has been selected to perform both design and data analysis.

Feasibility Study

A feasibility study was conducted between September and October 2024. The purpose of the study was to evaluate the optimal timing for patient enrollment, the feasibility of data collection using REDCap and other digital tools, and the potential to recruit an adequate sample size. Initial feasibility results helped inform procedural aspects and adjustments to data collection tools.

Ethical Considerations

This study will be conducted in accordance with the ethical standards of the responsible committee on human experimentation and with the World Medical Association Declaration of Helsinki. It will also adhere to the Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine [28]. Ethical approval will be granted by the Central-North-East Area Ethics Committee of the Veneto Region. The feasibility study has been approved by the institutional review boards of the University of Padova and Azienda ULSS 6 Euganea. Informed consent will be obtained from all participants prior to their inclusion in the study. Participants will be briefed on the nature, objectives, and potential consequences of the research. To ensure privacy and confidentiality, all personal data will be processed in strict compliance with the General Data Protection Regulation (GDPR) (EU) 2016/679. Data will be deidentified (pseudonymized) at the point of collection and stored on secure, encrypted servers accessible only to the authorized research team. No identifying information (such as names or hospital IDs) will be published. Participants will not receive any financial or nonfinancial compensation for their participation in this study.

Results

In the feasibility study, 36 patients were screened, of whom 33 agreed to participate; 6 participants did not complete all questionnaire items. Initial technical issues related to the REDCap platform and the usability of digital devices were identified during the feasibility phase and were subsequently resolved.

The mean completion time was 7.67 minutes for the VLUK questionnaire and 5.12 minutes for the VeLUSET. VeLUSET scores indicated moderate-to-high confidence in general and daily self-care activities, with lower scores observed in domains related to ulcer management and trauma prevention. The feasibility results supported the acceptability of the questionnaires and informed minor procedural refinements without requiring changes to the study outcomes or measurement instruments.

The randomized controlled trial is currently pending approval by the local ethics committee. The study was funded in June 2025. Participant enrollment is expected to start in spring 2026, with an anticipated enrollment period of approximately 6 months and a total study duration of 18 months, including 12 months of follow-up.

XSL•FO
RenderX

At the time of manuscript submission, no participants had been enrolled in the randomized controlled trial beyond those included in the feasibility phase. Data analysis for the main trial had not yet been performed. The results of the randomized controlled trial are expected to be available and disseminated by the end of 2027.

Discussion

Educational interventions represent a potentially efficacious method for promoting adherence to treatment programs and empowering patients with VLUs to assume an active role in the management of their disease. It can be suggested that patient education represents an effective strategy for improving the treatment of VLUs and reducing the incidence of relapse and the recovery time required.

Unfortunately, the time available for nurses to educate patients is often limited. This is due to the considerable number of patients and the necessity for an increasing number of nurses to manage the substantial workload. Therefore, it would be useful to use standardized educational methods to support the information given to the patient during treatment time. Such methods should be reliable, readily available, and allow the patient to access the information independently.

Although some educational interventions have been documented in the literature, it remains unclear which is the most effective [18].

A study compared a leaflet intervention with a video intervention and found that the video intervention was more effective in improving patients' knowledge in this type of patient, although the results were not statistically significant [21]. Therefore, this study aims to improve the knowledge and self-care skills of patients with VLUs using a video-based educational intervention and to determine which intervention (video or brochure) is more effective.

Video interventions represent an accessible and cost-effective method of patient education, even for older adults. It is a low-cost approach that requires minimal time and resources, while providing nurses with the opportunity to educate patients in a convenient and effective manner. Moreover, it is easily available to users, thereby rendering it a highly accessible form of patient education. Additionally, an educational video provides access to accurate information from experienced professionals, even in the context of prevention or posthealing, when the patient is no longer required to attend outpatient nursing clinics for dressing.

This study has several notable strengths. Primarily, the use of videos in outpatient clinics allows nurses to prioritize personalized care. This is achieved by reducing the time spent on providing the same information to patients repeatedly and instead responding to their queries regarding content in the educational video that may be unclear to them. Moreover, as previously stated, the utilization of video technology could potentially reduce the costs associated with health care companies, thereby allowing for the reinvestment of these savings into other areas that may enhance the quality of service. Ultimately, the dissemination of educational videos is a relatively simple process, which could potentially reach individuals who are at risk of developing VLUs but have not yet sought medical attention for this condition. Consequently, the implementation of such videos could contribute to a reduction in the incidence of VLUs over time.

It is important to acknowledge that this study is not without limitations. For instance, the generalizability of the findings is restricted due to the fact that the study will involve only a single health care company. Nevertheless, should the study produce positive results, the educational intervention could be subjected to a replication study on a larger scale.

Acknowledgments

The authors thank Maryem Chaabane and Eleonora Monaco, Unit of Biostatistics, Epidemiology, and Public Health, Department of Cardiac-Thoracic-Vascular Sciences and Public Health, University of Padova, Padova, Italy, for contributing to the development of the experimental educational intervention and for their support in the development, validation, and testing of the electronic case report form for the feasibility study. Artificial intelligence was not used to generate any portion of the text in this protocol.

Funding

The recording and production of the 4 experimental educational videos were funded by the Unit of Biostatistics, Epidemiology, and Public Health of the University of Padova. The randomized controlled study was funded by the Department of Cardiac-Thoracic-Vascular Sciences and Public under the Health MART_BIRD25_01 project in June 2025.

Data Availability

The datasets generated or analyzed during the present study will be available from the corresponding author upon reasonable request.

Authors' Contributions

CC made substantial contributions to the conception and design of the work, the drafting of the work, critical revisions for important intellectual content, and the final approval of the version to be published. JS and LT contributed to drafting the work and revising it critically for important intellectual content. MVC, FM, and CV were responsible for the acquisition, analysis, and

XSL•FO
RenderX

interpretation of data for the work. MM made substantial contributions to the conception and design of the work, the drafting of the work, and critical revisions for important intellectual content.

Conflicts of Interest

None declared.

References

1. Bobbink P, Pugliese MT, Larkin P, Probst S. Nurse-led patient education for persons suffering from a venous leg ulcer in outpatient's clinics and homecare settings: a scoping review. J Tissue Viability 2020 Nov;29(4):297-309. [doi: 10.1016/j.jtv.2020.08.006] [Medline: 32907753]
2. Price P, Harding K. Cardiff Wound Impact Schedule: the development of a condition-specific questionnaire to assess health-related quality of life in patients with chronic wounds of the lower limb. Int Wound J 2004 Apr;1(1):10-17. [doi: 10.1111/j.1742-481x.2004.00007.x] [Medline: 16722893]
3. Sen CK. Human wound and its burden: updated 2020 compendium of estimates. Adv Wound Care (New Rochelle) 2021 May;10(5):281-292. [doi: 10.1089/wound.2021.0026] [Medline: 33733885]
4. Beatty A, Jones PS, Vail C, Thompson JA, Reynolds SS. Implementing venous leg ulcer education and clinical decision support: a quality improvement project. Wound Manag Prev 2022 Sep;68(9):12-18. [Medline: 36112796]
5. Bossert J, Vey JA, Piskorski L, et al. Effect of educational interventions on wound healing in patients with venous leg ulceration: a systematic review and meta-analysis. Int Wound J 2023 May;20(5):1784-1795. [doi: 10.1111/iwj.14021] [Medline: 36408782]
6. Martinato M, Ranzato C, Faggian E, et al. Knowledge assessment among subjects with chronic venous leg ulcer in outpatient setting: translation and adaptation of a tool to identify subjects at risk of poor understanding. Wound Repair Regen 2023;31(5):679-687. [doi: 10.1111/wrr.13107] [Medline: 37368793]
7. Raffetto JD, Ligi D, Maniscalco R, Khalil RA, Mannello F. Why venous leg ulcers have difficulty healing: overview on pathophysiology, clinical consequences, and treatment. J Clin Med 2020 Dec 24;10(1):29. [doi: 10.3390/jcm10010029] [Medline: 33374372]
8. Žulec M, Rotar Pavlič D, Žulec A. The effect of an educational intervention on self-care in patients with venous leg ulcers—a randomized controlled trial. Int J Environ Res Public Health 2022 Apr 12;19(8):4657. [doi: 10.3390/ijerph19084657] [Medline: 35457524]
9. Huang Y, Hu J, Xie T, et al. Effects of home-based chronic wound care training for patients and caregivers: a systematic review. Int Wound J 2023 Nov;20(9):3802-3820. [doi: 10.1111/iwj.14219] [Medline: 37277908]
10. O'Brien JA, Finlayson KJ, Kerr G, Edwards HE. Testing the effectiveness of a self-efficacy based exercise intervention for adults with venous leg ulcers: protocol of a randomised controlled trial. BMC Dermatol 2014 Oct 3;14(1):16. [doi: 10.1186/1471-5945-14-16] [Medline: 25277416]
11. Kapp S, Miller C. The experience of self-management following venous leg ulcer healing. J Clin Nurs 2015 May;24(9-10):1300-1309. [doi: 10.1111/jocn.12730] [Medline: 25422075]
12. Protz K, Dissemond J, Seifert M, et al. Education in people with venous leg ulcers based on a brochure about compression therapy: a quasi-randomised controlled trial. Int Wound J 2019 Dec;16(6):1252-1262. [doi: 10.1111/iwj.13172] [Medline: 31418532]
13. Carradice D. Superficial venous insufficiency from the infernal to the endothermal. Ann R Coll Surg Engl 2014 Jan;96(1):5-10. [doi: 10.1308/003588414X13824511650498] [Medline: 24417822]
14. Weller CD, Buchbinder R, Johnston RV. Interventions for helping people adhere to compression treatments for venous leg ulceration. Cochrane Database Syst Rev 2016 Mar 2;2016(3):CD008378. [doi: 10.1002/14651858.CD008378.pub3] [Medline: 24009097]
15. Martinato M, Comoretto RI, Azzolina D, Gregori D. Valutazione della conoscenza della patologia e capacità di autocura in pazienti con ulcere venose croniche agli arti inferiori: risultati preliminari di uno studio epidemiologico condotto per mezzo di una applicazione per smartphone e tablet [Article in Italian]. : Società Italiana di Scienze Infermieristiche; 2023 URL: https://www.sisiweb.net/wp-content/uploads/2022/10/VALUTAZIONE-DELLA-CONOSCENZA-DELLA-PATOLOGIA-E-CAPACITA-DI-AUTOCURA.pdf [accessed 2026-03-25]
16. Brown A. Evaluating the reasons underlying treatment nonadherence in VLU patients: introducing the VeLUSET Part 1 of 2. J Wound Care 2014 Jan;23(1):37-47. [doi: 10.12968/jowc.2014.23.1.37] [Medline: 24406542]
17. Domingues EAR, Kaizer UAO, Lima MHM. Effectiveness of the strategies of an orientation programme for the lifestyle and wound-healing process in patients with venous ulcer: a randomised controlled trial. Int Wound J 2018 Oct;15(5):798-806. [doi: 10.1111/iwj.12930] [Medline: 29808613]
18. Shanley E, Moore Z, Patton D, et al. Patient education for preventing recurrence of venous leg ulcers: a systematic review. J Wound Care 2020 Feb 2;29(2):79-91. [doi: 10.12968/jowc.2020.29.2.79] [Medline: 32058853]

XSL•FO
RenderX

19. Edwards H, Courtney M, Finlayson K, Lewis C, Lindsay E, Dumble J. Improved healing rates for chronic venous leg ulcers: pilot study results from a randomized controlled trial of a community nursing intervention. Int J Nurs Pract 2005 Aug;11(4):169-176. [doi: 10.1111/j.1440-172X.2005.00521.x]
20. Benevides JL, Coutinho JFV, Pascoal LC, et al. Development and validation of educational technology for venous ulcer care. Rev Esc Enferm USP 2016 Apr;50(2):309-316. [doi: 10.1590/S0080-623420160000200018] [Medline: 27384212]
21. Baquerizo Nole KL, Yim E, Van Driessche F, et al. Educational interventions in venous leg ulcer patients. Wound Repair Regen 2015;23(1):137-140. [doi: 10.1111/wrr.12247] [Medline: 25487073]
22. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
23. Martinato M, Comoretto RI, Azzolina D, Gregori D. Lo strumento di autoefficacia veluset per i pazienti con ulcera venosa delle gambe: adattamento culturale italiano e validazione psicometrica utilizzando un approccio bayesiano [Article in Italian]. University of Padua. 2021. URL: https://www.unipd-ubep.it/2023/04/20/lo-strumento-di-autoefficacia-veluset-per-i-pazienti-con-ulcera-venosa-delle-gambe-adattamento-culturale-italiano-e-validazione-psicometrica-utilizzando-un-approccio-bayesiano-dott-matteo-martinato/ [accessed 2026-03-25]
24. Brown A, Kendall S, Flanagan M, Cottee M. Encouraging patients to self-care - the preliminary development and validation of the VeLUSET©, a self-efficacy tool for venous leg ulcer patients, aged 60 years and over. Int Wound J 2014 Jun;11(3):326-334. [doi: 10.1111/iwj.12199] [Medline: 24373556]
25. Upton D, Upton P, Alexander R. Well-being in wounds inventory (WOWI): development of a valid and reliable measure of well-being in patients with wounds. J Wound Care 2016 Mar;25(3):114-120. [doi: 10.12968/jowc.2016.25.3.114] [Medline: 26947691]
26. Ratliff CR, Rodeheaver GT. Use of the PUSH tool to measure venous ulcer healing. Ostomy Wound Manage 2005 May;51(5):58-60. [Medline: 16014985]
27. Brooks R, EuroQol Group. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72. [doi: 10.1016/0168-8510(96)00822-6] [Medline: 10158943]
28. Convention for the protection of human rights and dignity of the human being with regard to the application of biology and medicine: convention on human rights and biomedicine (ETS No. 164). Oviedo, 4.IV.1997. Council of Europe. 1997. URL: https://rm.coe.int/168007cf98 [accessed 2026-04-13]

Abbreviations

PUSH Tool 3.0: Pressure Ulcer Scale for Healing Tool 3.0
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
VeLUSET: Venous Leg Ulcer Self-Efficacy Tool
VLU: venous leg ulcer
VLUK: Venous Leg Ulcer Knowledge
WOW-I: Well-being of Wounds Inventory

Edited by J Sarvestan; submitted 16.Sep.2025; peer-reviewed by YY Kristian; revised version received 29.Jan.2026; accepted 30.Jan.2026; published 21.Apr.2026.

Please cite as:
Canova C, Scappin J, Chiaruttini MV, Tasson L, Masiero F, Varagnolo C, Martinato M
The Effect of an Educational Intervention on the Knowledge and Self-Care Skills of Adult Patients With Venous Leg Ulcers: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84211
URL: https://www.researchprotocols.org/2026/1/e84211
doi:10.2196/84211

© Cristina Canova, Jessica Scappin, Maria Vittoria Chiaruttini, Laura Tasson, Flavia Masiero, Claudia Varagnolo, Matteo Martinato. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Overcoming the HIV Pre-Exposure Prophylaxis Uptake Barrier With A Health Action Process Approach and Conditional Economic Incentives Among Men Who Have Sex With Men in 3 Chinese Cities: Protocol for a Randomized Controlled Trial

Hao Lin[1], MPH; Mojun Ni[1], MPH; Chun Chang[1], PhD; Feng Sun[1], PhD; Jinghua Li[2], PhD; Omar Galárraga[3], PhD; Phoenix Kit-Han Mo[4], PhD; Wangnan Cao[1], PhD

[1]School of Public Health, Peking University Health Science Center, 38# Xueyuan Road, Haidian District, Beijing, China

[2]Department of Public Health and Medicinal Administration, Faculty of Health Sciences, University of Macau, Macau, China

[3]School of Public Health, Center for Global Public Health, Brown University, Providence, RI, United States

[4]School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Wangnan Cao, PhD
School of Public Health, Peking University Health Science Center, 38# Xueyuan Road, Haidian District, Beijing, China

Abstract

Background: Despite high pre-exposure prophylaxis (PrEP) awareness and willingness among Chinese men who have sex with men (MSM), actual uptake remains critically low. The health action process approach (HAPA) is a 2-phase theory that addresses both the motivational and volitional phases of health behavior change, making it suitable for tackling the intention-behavior gap in PrEP use. In China, where PrEP is not covered by national health insurance, financial barriers further hinder uptake. Group-based conditional economic incentives (GCEIs) represent an innovative strategy to address cost-related disincentives by leveraging peer influence.

Objective: This randomized controlled trial aims to evaluate the effectiveness of HAPA-driven interventions, both alone and combined with GCEI, to enhance PrEP initiation and adherence among high-risk, PrEP-naive MSM in China.

Methods: High-risk HIV-negative MSM naïve to PrEP were recruited via community-based organizations in 3 Chinese cities. Participants underwent individual randomization to 3 arms: (1) HAPA-based cognitive behavioral intervention featuring tailored WeChat messages and 2 structured motivational interviews targeting stage-specific barriers; (2) combined HAPA intervention with GCEI linked to peer group initiation rates, incorporating facilitated offline group activities; or (3) standard care control receiving nonpersonalized HIV prevention information. The physician-prescribed PrEP initiation rate at 3 months serves as the primary outcome. Secondary outcomes include PrEP adherence, HIV risk perception, PrEP-related stigma, and intervention acceptability. Analysis follows intention-to-treat principles using mixed-effects models.

Results: The study protocol received ethics approval in December 2024 and was funded in December 2025. As of August 2025, 3 community-based organizations in Mianyang, Zibo, and Hohhot have agreed to participate. Preliminary testing has led to refinements in the WeChat mini-program, questionnaires, and motivational interview scripts. Data collection commenced in October 2025 and is projected to be completed by August 2026. As of February 2026, baseline recruitment is complete, with 381 eligible participants enrolled and randomized. The baseline characteristics were balanced across the 3 arms. Outcome data will be collected after the intervention. Data analysis is ongoing, with full results expected to be published before December 2026.

Conclusions: This study rigorously tests a novel intervention integrating the HAPA with a GCEI model. By addressing both individual cognitive-behavioral barriers and financial disincentives within a supportive group context, this approach offers a potentially scalable strategy to bridge the PrEP uptake gap in China. Findings will provide critical evidence on combining theory-driven behavioral techniques with peer-group economic incentives for HIV prevention.

Trial Registration: ClinicalTrials.gov NCT06931106; https://clinicaltrials.gov/study/NCT06931106

International Registered Report Identifier (IRRID): PRR1-10.2196/82123

(JMIR Res Protoc 2026;15:e82123) doi:10.2196/82123

KEYWORDS

HIV; pre-exposure prophylaxis; men who have sex with men; health action process approach; economic incentive; randomized controlled trial; China

XSL•FO
RenderX

Introduction

The burden of AIDS remains substantial. It is estimated that 0.7% (95% uncertainty interval 0.6 - 0.8%) of adults aged 15 - 49 worldwide are living with HIV, with a higher prevalence of 7.5% among men who have sex with men (MSM) [1]. In China, by the end of 2023, there were 1.3 million people living with HIV or AIDS, with a cumulative death toll of 458,000 [2]. The incidence of new HIV infections among MSM in China is 5.6 per 100 person-years [3]. The risk of HIV transmission remains significant, necessitating effective measures among groups at high risk, such as MSM, to control the spread of the epidemic [4].

Pre-exposure prophylaxis (PrEP) is an efficient biological prevention method to prevent HIV infection [5]. In most high-income countries, the overall PrEP usage rate among high-risk groups ranges from 28% to 80% [6]. Currently, international recommendations for MSM include oral tenofovir disoproxil fumarate or emtricitabine and long-acting injectable-PrEP (LAI-PrEP), both of which have been proven to effectively reduce the risk of HIV infection [3,4]. China has made significant efforts to promote PrEP among high-risk groups, such as launching pilot projects and approving generic medicine [7]. However, low willingness to use oral PrEP among key populations [5], stemming from cognitive biases, leads to low PrEP adoption despite potentially high demand [8]. There are over 12 million MSM in China [9], with PrEP awareness rates ranging from 41% to 86% [10,11], and approximately 65.8% to 84.9% of MSM willing to accept PrEP [12,13]. A survey of 2388 MSM across 6 Chinese cities showed that only 13.9% had used PrEP in 2022 [14]. Factors such as low health literacy, low behavior control, distrust in PrEP efficacy, lack of knowledge about PrEP, high costs, discrimination, and stigma are major barriers to PrEP use despite high demand [15,16]. Among those who initiate PrEP, its protective efficacy depends on sufficiently high adherence [17,18]. PrEP adherence rates among MSM in China range from 29.9% to 41.1% [19,20]. For LAI-PrEP, the long-acting injectable cabotegravir was approved in China in May 2024, and its usage intention and behavior among key populations are still being explored [7].

Behavioral interventions to promote PrEP among groups at high risk need to focus on increasing initiation rates before initiation and improving adherence after initiation. Existing interventions to promote PrEP initiation among MSM include PrEP health education [21], motivational structured interviews [22], MSM peer education and guidance [23], and mobile health technology interventions [24]. In China, smartphone-based interventions and app-based PrEP initiation interventions have shown some effectiveness, with PrEP initiation rates ranging from 65% to 89% [25]. The health action process approach (HAPA) is a 2-phase theory that divides behavior into the preintentional motivational phase (focusing on promoting the generation of behavioral intention) and the postintentional volitional phase (focusing on translating intention into behavior, maintaining behavior, and managing relapse) [26]. The HAPA phase model can effectively address the high willingness–low usage gap in PrEP and is particularly suitable for explaining and measuring low willingness, low usage rates, low adherence, and low reinitiation rates after discontinuation at different PrEP stages. Chan et al [27] used motivational structured interviews to improve PrEP initiation rates, with the core of the intervention being to explore and strengthen individuals' intrinsic motivation to use PrEP and improve self-efficacy through situational simulation. Clarifying the suitability of PrEP and risk behavior is key to further effective PrEP intervention [28].

In China, where PrEP medications are not yet covered by national health insurance, financial barriers significantly hinder uptake among MSM despite high willingness. The out-of-pocket cost for oral PrEP (eg, approximately US $5.5 per on-demand dose) poses a critical obstacle, particularly for low-income individuals. In China, the monthly cost of domestically produced generic oral PrEP is approximately ¥200 - 300 (US $28 - $42); for on-demand dosing, the cost per use is about ¥40 (US $5.6). This context underscores the need for innovative strategies to mitigate cost-related disincentives. Group-based conditional economic incentives (GCEIs) offer a promising solution by leveraging behavioral economics principles. GCEI capitalizes on peer influence and shared accountability, addressing both financial and social barriers [29,30]. For MSM, whose health behaviors are often shaped by community norms, GCEI can amplify motivation through collective goal-setting (eg, group-level initiation targets) and tiered rewards (eg, US $22 for 80% - 100% group adherence). Prior studies demonstrate GCEI's efficacy in improving antiretroviral therapy adherence among youth, suggesting translatability to PrEP promotion [31-33]. By integrating financial incentives with peer dynamics, GCEI aligns with China's need for scalable, cost-sensitive interventions to bridge the intention-behavior gap in PrEP uptake.

Existing research on PrEP uptake among Chinese MSM has largely relied on cross-sectional analyses that identify barriers but fail to translate findings into effective interventions. These studies consistently highlight structural challenges—particularly financial constraints in China's self-pay PrEP model and pervasive stigma surrounding both HIV prevention and same-sex behaviors—yet proposed solutions rarely account for the cultural and logistical realities of this population. Our study addresses these limitations through an intervention model that combines theory-driven behavioral strategies with pragmatic delivery methods tailored to the Chinese context. The use of WeChat-based eHealth platforms proves particularly strategic, as it capitalizes on near-universal smartphone penetration among Chinese MSM while circumventing stigma through discreet, asynchronous engagement. This approach recognizes that target users may avoid clinic-based services due to privacy concerns, yet actively participate in virtual communities where health information is exchanged. By embedding the intervention within a familiar digital ecosystem and incorporating group-based incentives that leverage existing social networks, the design intentionally works with—rather than against—the lived experiences of Chinese MSM. The result is an implementation framework that simultaneously addresses individual behavioral determinants and systemic barriers, offering a realistic path to scaling PrEP access in stigma-affected populations.

This study aims to evaluate the effectiveness of HAPA-based interventions, alone and combined with GCEI, in improving

XSL•FO
RenderX

PrEP initiation and adherence among Chinese MSM at high risk for HIV infection. We hypothesize that both intervention arms (HAPA and HAPA+GCEI) will outperform the standard care control in promoting PrEP initiation, with the combined HAPA+GCEI arm expected to yield the greatest efficacy.

Methods

Study Design

This study uses a randomized controlled trial design. The MSM-friendly community-based organizations (CBOs) are selected, and the participants are randomly allocated to the following three arms: (1) HAPA-based cognitive intervention arm, (2) the HAPA+ conditional economic incentive combined intervention arm, and (3) control arm. The intervention will last for 3 months, during which baseline data and 2 follow-up assessments will be collected. The first follow-up will occur at the conclusion of the intervention, while the second follow-up will take place three months postintervention. We will assess PrEP uptake before and after the intervention and between different study arms. We used the SPIRIT 2025 checklist to ensure completeness of the trial protocol reporting (Checklist 1).

Study Population and Participant Recruitment

The study population is sourced from 3 MSM-friendly CBOs located in different regions. The inclusion criteria for community organizations are: (1) the organization must meet the sample size requirements for the project; (2) there must be a willingness to collaborate, making PrEP promotion a work priority; and (3) the organization should have a good partnership with local hospitals or centers for disease control and prevention to coordinate with disease control specialists and HIV specialists. Community collaborators participate in project training, assist with participant recruitment, support baseline and follow-up surveys, and help with quality control. Participants included in this study must meet the following criteria: (1) aged 18 years or older; (2) assigned male sex at birth and have had sexual intercourse with men in the past 6 months; (3) HIV negative at the start of the study; (4) no prior PrEP using experience; (5) willing to participate in the PrEP intervention and voluntarily consent to complete data collection and follow-up according to the assigned protocol; (6) classified as high risk for HIV infection in the past 6 months as suggested by the "Expert Consensus on HIV Pre-Exposure Prophylaxis in China (2023)" [7]. Participants are excluded if they meet any of the following criteria: (1) HIV positive, or unaware of their HIV status and refuse to undergo HIV testing; (2) have a health condition that makes PrEP unsuitable; and (3) have cognitive impairments or other severe illnesses that may affect the informed consent process.

The sample size calculation was based on the formula for an individually randomized controlled trial, assuming a 2-sided significance level (α) of .05 and type II error rate (β) of .10 (90% power). Based on extant literature and a 2021 multicenter PrEP survey in China indicating a median current PrEP usage rate (p_1) of 5% among MSM, the target post-intervention usage rate (p_2) was set at 20%. This yielded a minimum requirement of 100 participants per group. Accounting for an anticipated 15% attrition rate, the sample size was increased to 126 participants per group, resulting in a total target enrollment of 378 participants (126 per group) across 3 CBOs.

The randomization sequence was generated by an independent statistician using simple randomization. The sequence was then integrated into the backend of the customized WeChat mini-program. After a participant completed the baseline survey and was confirmed eligible, the system automatically executed the group assignment. Allocation concealment was strictly implemented. After a participant completed the baseline assessment and was confirmed eligible, the system automatically executed the group assignment based on the preloaded, concealed sequence. Until this point of allocation, neither the participant, the research coordinator enrolling the participant, nor the outcome assessors could access or predict the upcoming assignment. The allocation is only revealed to the system and the intervention coordinator after the participant's baseline data is irrevocably locked and enrollment is finalized. The participants are blinded to group assignment. Outcome assessors and data analysts remain blinded to group assignment throughout the trial. They have access only to deidentified participant data coded with unique study IDs, without any information linking the ID to the intervention arm.

Intervention

Tailored Health Messages

Participants receive a 3-month mobile health intervention guided by the HAPA. Tailored messages are developed using behavior change techniques based on theoretical guidance [34]. Three times weekly, the personalized health message is randomly delivered to participants via the WeChat mini-program. These messages are drawn from an information library designed based on HAPA theory, expert input, and literature review, and are tailored to the participant's current PrEP using status. For example: participants exhibiting low risk perception receive messages incorporating vivid illustrations of potential health risks associated with their PrEP adherence status, using real-life case studies and epidemiological data to enhance awareness. Participants with low self-efficacy receive messages incorporating techniques such as providing success stories of similar individuals overcoming adherence challenges and offering step-by-step guidance to bolster confidence in maintaining the PrEP regimen.

Brief PrEP Motivational Interviewing

The brief motivational interviewing sessions are conducted by experienced professionals with demonstrated expertise in both PrEP provision and motivational interviewing techniques. These 1-on-1 sessions (delivered face-to-face or online) are tailored to each participant's specific stage within the PrEP motivational cascade, such as objective identification, PrEP precontemplation, PrEP contemplation, PrEP preparation, PrEP action and initiation, and PrEP maintenance. Content is individualized to address specific barriers currently present, informed primarily by findings from the participant's most recent assessment phase. This targeted approach aims to resolve impediments to PrEP engagement through structured, theory-guided dialogue.

XSL•FO
RenderX

Conditional Economic Incentives

Participants in the combined intervention arm also receive additional peer GCEIs on top of the HAPA intervention. Incentive information is communicated immediately at the start of the intervention. The incentive amounts are determined at the end of the 3-month intervention by the researchers based on the overall completion status of each participant and their peer group according to predetermined conditions.

Participants are formed into peer groups of approximately 6 individuals. These groups participate in structured in-person sessions. Activities are designed to include: introductions and icebreakers, PrEP knowledge quizzes, partner dialogues, "Voicing My Barriers" roundtable discussions, and PrEP scenario role-playing. The objectives are to foster trust, enable deeper peer connection, enhance group belonging, strengthen HIV prevention awareness, identify and overcome barriers to PrEP use, and ultimately increase PrEP initiation rates. Six participants form a peer group, and the economic incentive amount is based on the average completion level of the group. Group-level economic incentives are stepwise reduced based on the group's average PrEP initiation rate. Participants qualify for US $22 if their group achieves an 80% - 100% rate, US $11 for a 50% - 79% rate, and US $0 for rates below 50%. PrEP initiation is operationally defined as the documented provision of a physician-issued prescription.

The WeChat mini-program automatically records intervention process evaluation indicators, including: (1) whether each participant has read the information messages; (2) the reading time for each message; and (3) each participant's judgment of whether the information was helpful.

Control

Participants receive nonpersonalized HIV prevention information, including free HIV testing and condom use, following the guidelines provided by the Chinese Center for Disease Control and Prevention and the operational standards of the community organizations, with twice-weekly information push notifications.

Outcomes

The primary outcome indicator of this study is the rate of PrEP initiation, based on formal PrEP prescriptions issued by doctors. The secondary outcome indicators include medication adherence rates after initiating PrEP (based on self-reported adherence and the medication calendar in the mini-program), intention to pay for PrEP (for 3 types of PrEP), PrEP stigma, changes in condom use, and changes in number of sexual partners before and after PrEP use.

Measurements and Follow-Ups

For the questionnaire survey, participants are informed about the study's objectives, content, and procedures, and they sign an electronic informed consent form. The following baseline assessment questionnaire is conducted via a mini-program:

1. Demographic characteristics: Includes age, ethnicity, education level, marital status, employment status, income level, and sexual orientation
2. Sexual behavior characteristics: Includes age at first sexual intercourse, types and number of sexual partners, sexual roles, condom use, HIV testing, and whether they have paid or have been paid for sex
3. PrEP usage intentions and behaviors: Includes a PrEP knowledge assessment questionnaire, awareness rates and sources of information about PrEP, willingness to use and pay for PrEP, past PrEP usage, types of PrEP, frequency of use, and willingness to use different types of PrEP
4. HAPA-based PrEP cognition scale: Includes self-efficacy for initiating, maintaining, and resuming PrEP use, positive outcome expectations for PrEP, negative outcome expectations for PrEP, perceived HIV risk, willingness to use PrEP, and action plans for initiating and maintaining PrEP use. This scale has been developed and validated among the MSM population, comprising 9 dimensions and 44 items, using a 5-point Likert scale from 1 ("strongly disagree") to 5 ("strongly agree"), with a total Cronbach α of 0.882 [35]
5. PrEP Stigma Scale: Uses a 10-item HIV PrEP Stigma Scale, which has been validated in multiple countries, eg, "using PrEP will lead to negative evaluations by others." The scale uses a 5-point Likert score from 1 ("strongly disagree") to 5 ("strongly agree"), with higher scores indicating greater PrEP stigma, and a Cronbach α of 0.880 [36].

Follow-ups are conducted at the end of the intervention and 3 months later. The follow-up questionnaire builds upon the baseline assessment, focusing on the following key areas:

1. Intervention process and satisfaction evaluation: This evaluates participants' experience and satisfaction with the mobile health intervention platform (the WeChat mini-program), including usage frequency, overall satisfaction, willingness to recommend it, and specific feedback on its ease of use, interface design, information resources, medication reminders, recording functions, and data privacy protection.
2. Risk compensation behavior assessment: It investigates changes in participants' sexual behavior patterns during the intervention to assess the possibility of risk compensation, including changes in the frequency of same-sex intercourse, number of male sexual partners, and condom use frequency compared to baseline.
3. PrEP uptake status: Based on the current PrEP status at follow-up, it further explores specific reasons for noninitiation or discontinuation, sources of medication acquisition, methods of recording medication use, episodes of treatment interruption, occurrence of adverse drug reactions, partner awareness, psychological experiences related to PrEP use, and willingness to recommend it.
4. Dynamic measurement of PrEP-related cognition: It reapplies key items from the validated HAPA-based PrEP cognition scale to assess changes in postintervention cognition, including self-efficacy for PrEP initiation and maintenance, positive and negative outcome expectations, perceived HIV risk, and action planning.
5. Reassessment of PrEP decision and payment willingness: It re-evaluates participants' perception of pricing and

XSL•FO
RenderX

willingness to use different PrEP regimens at market prices, as well as their views on potential models for cost-sharing.

6. Follow-up on PrEP perceived stigma and MSM group cohesion: It uses a brief mental health scale to assess emotional states over the past week, remeasures PrEP perceived stigma levels, and evaluates changes in perceived cohesion within the MSM community.

The sexually transmitted infections testing data (HIV, syphilis, hepatitis B, and hepatitis C) from participants at baseline and during 2 follow-up visits are collected. All tests are performed using standardized diagnostic kits (Wondfo Biotech) according to the manufacturer's protocols. Serological results are recorded as binary outcomes (positive or negative). For participants who initiated PrEP during the study, adverse drug reaction evaluations are integrated throughout the PrEP medical service process. A formal safety stopping rule is prespecified. The trial will be halted immediately if the independent Data and Safety Monitoring Board identifies clear, serious, and related adverse events attributable to the study procedures that pose unacceptable risks to participants, or if new, definitive external evidence emerges that renders the trial unethical or obsolete.

Preliminary Testing and Modifications

Preliminary testing with key stakeholders enabled critical refinements to study instruments and procedures, including substantial enhancements to the WeChat mini-program's navigation flow to facilitate access to PrEP information and intervention components, along with added privacy protection features such as local data storage and optional fingerprint authentication, while message delivery timing was optimized based on observed user activity patterns. Questionnaire modifications involved clarifying wording in the PrEP Stigma Scale to better differentiate between perceived and internalized stigma, and supplementing HAPA-based scales with concrete examples to improve comprehension of abstract constructs like self-efficacy.

The motivational interview protocol is finalized after cultural appropriateness testing, resulting in standardized prompts for exploring financial and structural barriers specific to Chinese MSM populations, as well as incorporating guidance for discussing newly approved LAI-PrEP options, collectively strengthening the intervention's relevance and implementation fidelity for the target population.

Ethical Considerations

The study protocol was approved by the Ethics Commission of Peking University Health Science Center (IRB00001052-24167) on December 22, 2024. Prior to study initiation, all participants are fully informed about the study objectives, procedures, potential risks, and their rights, and provide written informed consent. Participants receive a compensation of US $2.8 (approximately CNY 20) for completing each questionnaire at baseline and follow-up assessments. The study does not provide free PrEP medication; participants are required to purchase medications at market prices. To protect participant privacy and confidentiality, several measures have been implemented: data collection uses a customized WeChat mini-program featuring local data storage and optional fingerprint authentication for enhanced privacy; all questionnaires are administered anonymously and linked only via a unique study ID; and sexually transmitted infections test results (HIV, syphilis, hepatitis B, and hepatitis C) are recorded solely as binary (positive or negative) outcomes without any personally identifiable information.

Statistical Analysis

Descriptive Analysis of Baseline Data and Intergroup Comparisons

For numeric variables such as PrEP usage intentions and HAPA component scores, the mean and SD (or median and IQR) will be reported based on the data distribution. Mann-Whitney tests or *t* tests will be used to compare the baseline data balance between the intervention and control groups according to the data distribution type. Categorical variables such as PrEP initiation and adherence will be described using rates, and chi-square tests or Mann-Whitney tests will be used to compare the baseline data balance between the intervention and control groups.

Evaluation of Intervention Effectiveness

This study will use mixed-effects models to evaluate the effectiveness of primary and secondary outcome indicators. Different evaluation models will be selected based on the type of outcome indicators. Individual participants will be the unit of analysis. For the evaluation of categorical variable effects, a generalized mixed linear model will be used, reporting the adjusted odds ratio. For continuous variable effects, a mixed linear model will be used, reporting the adjusted mean differences between groups. The effectiveness evaluation will adhere to the intention-to-treat principle, including all randomized participants in the analysis. Stratified analysis will be conducted based on participants' age, education level, income level, sample source, and city characteristics to explore evidence of effects related to these various characteristics. Statistical analyses will be performed using R (R Foundation for Statistical Computing) and SPSS software (IBM Corp), with a P value<.05 considered statistically significant. The study assumes that missing values are randomly distributed, and missing data will be handled using maximum likelihood estimation. Sensitivity analyses will be conducted after imputing missing values using multiple imputation methods.

Results

The trial is a 3-arm, parallel-group, randomized controlled trial conducted in collaboration with 3 CBOs located in Mianyang (Sichuan Province), Zibo (Shandong Province), and Hohhot (Inner Mongolia), China. The study protocol was funded in December 2025. Key operational phases, from recruitment to analysis, with specific dates and participant flow, are presented in Figure 1.

XSL•FO
RenderX

Figure 1. Participant flowchart (status as of February 2026: baseline assessment and randomization completed; interventions are ongoing). GCEI: group-based conditional economic incentive; HAPA: health action process approach; ITT: intention-to-treat.

XSL•FO
RenderX

Participant Recruitment

Participant recruitment and screening were conducted from October to December 2025. During this period, 743 eligible MSM from the 3 CBOs were initially identified and assessed. Of these, 362 individuals were excluded for the following reasons: failure to meet predefined inclusion criteria (n=306), declining to participate (n=52), or other logistical reasons (n=4). Consequently, 381 participants provided written informed consent and were formally enrolled in the study. These 381 participants were randomly allocated into 1 of 3 study arms using a computer-generated sequence: HAPA-based cognitive behavioral intervention group (n=128), HAPA plus GCEI intervention group (n=126), and control group (n=127).

Participant Baseline Characteristics

As shown in Table 1, the baseline sociodemographic characteristics of the participants were well-balanced across the 3 study groups. Specifically, the average ages across the 3 groups were similar, with the HAPA cognitive intervention group averaging 34.5 (SD 9.9) years, the HAPA combined conditional economic incentive intervention group averaging 35.4 (SD 10.6) years, and the control group averaging 34.3 (SD 10.6) years ($F_{2,378}$=0.411; P=.66). The geographical distribution of participants among the 3 recruitment cities (Mianyang, Zibo, Hohhot) was also balanced and consistent across all groups (χ^2_4=0.205; P=.96).

XSL•FO
RenderX

Table . Baseline sociodemographic characteristics of participants (N=381).

Characteristic	Group			*F*/chi-square (*df*)[a]	*P* value
	HAPA[b] cognitive intervention (n=128)	HAPA+CEI[c] intervention (n=126)	Control (n=127)		
Age (y), mean (SD)	34.5 (9.9)	35.4 (10.6)	34.3 (10.6)	0.411 (2, 378)	.66
City, n (%)					
Mianyang	42 (32.8)	43 (34.1)	41 (32.3)	0.205 (4)	.96
Zibo	44 (34.4)	41 (32.5)	42 (33.1)		
Hohhot	42 (32.8)	42 (33.3)	44 (34.6)		
Current relationship, n (%)[d]					
Single	51 (39.8)	65 (51.6)	55 (43.3)	3.731 (2)	.16
Has a boyfriend	54 (42.2)	41 (32.5)	52 (40.9)	2.943 (2)	.23
Heterosexual marriage or has a girlfriend	17 (13.3)	12 (9.5)	18 (14.2)	1.424 (2)	.49
Divorced or widowed	13 (10.2)	11 (8.7)	8 (6.3)	1.260 (2)	.53
Education level, n (%)[e]					
Junior high school or below	17 (13.3)	11 (8.7)	9 (7.1)	6.156 (6)	.41
High school or associate degree	50 (39.1)	51 (40.5)	44 (34.6)		
Bachelor’s degree	38 (29.7)	43 (34.1)	48 (37.8)		
Master’s degree or above	15 (11.7)	16 (12.7)	11 (8.7)		
Employment status, n (%)					
Full-time	94 (73.4)	96 (76.2)	98 (77.2)	1.548 (6)	.96
Unemployed or retired	15 (11.7)	15 (11.9)	13 (10.3)		
Part-time	11 (8.6)	7 (5.6)	10 (7.9)		
Student	8 (6.3)	8 (6.3)	6 (4.7)		
Personal monthly income (CNY[f]), n (%)					
≤3000	28 (21.9)	35 (27.8)	23 (18.1)	4.141 (6)	.66
3001 - 6000	55 (43.0)	52 (41.3)	61 (48.0)		
6001 - 9000	31 (24.2)	28 (22.2)	28 (22.0)		
>9000	14 (10.9)	11 (8.7)	15 (11.8)		
Sexual orientation, n (%)					
Homosexual	101 (78.9)	101 (80.2)	97 (76.4)	2.347 (4)	.67
Bisexual or heterosexual	24 (18.8)	19 (15.1)	23 (18.1)		
Uncertain	3 (2.3)	6 (4.8)	7 (5.5)		

[a]An *F* test was conducted for Age. All other categories were analyzed with a chi-square test.

[b]HAPA: health action process approach.

[c]CEI: conditional economic incentives.

[d]Multiple selection allowed; percentages do not sum to 100%.

[e]Due to technical issues with the data collection system, some data on education level are missing.

[f]1 CNY=US $0.145.

XSL•FO
RenderX

Furthermore, analyses of various personal and social characteristics revealed no significant group differences. This included educational attainment (χ^2_6=6.156, *P*=.41), current relationship status (*P*>.05), employment status (χ^2_6=1.548, *P*=.96), level of personal monthly income (χ^2_6=4.141, *P*=.66), and sexual orientation distribution (χ^2_4=2.347, *P*=.67). In terms of these characteristics, the descriptive profile shows that the majority of participants were either single or in a relationship, were primarily engaged in full-time employment, and reported a personal monthly income concentrated in the range of 3001 to 6000 CNY (US $436.02-$871.76). The predominant self-identified sexual orientation among participants was homosexual.

Intervention Delivery and Follow-Up Status

Following the baseline assessment, the 3-month active intervention period commenced in January 2026. The intervention delivery, including tailored WeChat messaging, scheduled motivational interviews, and structured in-person group sessions for the HAPA+GCEI arm, is scheduled for completion in April 2026.

Data Collection and Analysis Timeline

Data collection for the primary outcome and secondary outcomes is scheduled to conclude with the final follow-up assessment in August 2026, with full results expected to be published before December 2026. Database lock and statistical analysis are planned for December 2026.

Discussion

This protocol outlines a study aimed at enhancing the uptake of HIV PrEP among Chinese MSM through a multifaceted approach, integrating the HAPA and conditional economic incentives. This research addresses the critical public health challenge of low PrEP use among high-risk populations, particularly MSM, who are disproportionately affected by HIV.

The HAPA model, a robust theoretical framework, provides a comprehensive strategy for understanding and promoting health behavior change. By focusing on both motivational and volitional phases, the HAPA model can effectively address the barriers to PrEP initiation and adherence [26]. This approach aligns with international research advocating for theory-driven, stage-matched interventions to bridge the intention-behavior gap, such as motivational interviewing for PrEP uptake among MSM in the United States [37]. However, many existing behavioral interventions in high-income settings do not adequately address the substantial financial barriers that are paramount in contexts like China, where PrEP is not covered by national health insurance [38]. This study directly responds to this gap by integrating HAPA with a pragmatic financial component.

Additionally, the incorporation of economic incentives represents an innovative strategy that leverages the power of financial rewards to encourage health-promoting behaviors. While economic incentives have shown promise in improving antiretroviral therapy adherence [39] and in pilot PrEP studies [33] elsewhere, our GCEI model innovates by applying a group-contingent structure within a Chinese MSM context. Economic incentives can reduce the financial burden associated with accessing PrEP, thereby overcoming one of the significant barriers to its use in the Chinese context [40]. By analyzing objective measures of PrEP initiation and adherence, this study will contribute valuable insights into the effectiveness of integrating behavioral interventions with economic incentives in promoting HIV prevention strategies, which is of global interest [41].

Several practical and operational considerations are crucial to ensuring the trial's success. To maintain participant engagement and facilitate follow-up, providing clear and accessible information about the study and PrEP is essential. Regular reminders through various communication channels will be used to enhance retention rates. Furthermore, standard training for all staff involved in participant recruitment and data collection will be implemented to ensure high-quality and consistent data throughout the study.

However, this study does have potential limitations. First, the generalizability of the findings may be affected by the implementation through selected CBOs in specific urban Chinese settings, which may limit applicability across different cultural and socioeconomic contexts. The 3-month intervention period may not capture long-term behavioral patterns, suggesting that future studies should incorporate extended follow-up assessments. Additionally, the effectiveness of economic incentives could be influenced by external macroeconomic factors and individual attitudes toward financial rewards, though socioeconomic status will be included as a moderating variable in the analysis.

This study demonstrates several key strengths. An innovative intervention model has been designed by integrating eHealth technologies with established CBO partnerships. The WeChat-based platform is implemented in collaboration with 3 well-established MSM-friendly CBOs (each with 5 - 10 y of documented HIV prevention experience), creating a potentially replicable implementation framework. The study design includes comprehensive pilot testing of behavioral change techniques prior to full implementation, with systematic application of HAPA theory across all intervention components. Both individual-level and group-level mechanisms have been carefully operationalized to allow for thorough evaluation. The hybrid intervention approach (combining digital and in-person components) addresses several implementation challenges. Existing social networks will be leveraged through CBO infrastructures, while privacy concerns will be mitigated through discreet eHealth communication channels. The protocol includes specific strategies for cost-effective scaling in resource-limited settings. These design features may provide valuable insights for bridging the intention-behavior gap in PrEP uptake among key populations.

In conclusion, this study outlines a comprehensive research protocol aimed at addressing the pressing challenge of low PrEP uptake among Chinese MSM. By integrating behavioral interventions based on HAPA with conditional economic incentives, we aim to identify effective strategies to enhance PrEP initiation and adherence in this vulnerable population.

XSL•FO
RenderX

The outcomes have the potential to inform both public health practice in China and similar global settings by demonstrating how individual-level cognitive support and group-level financial strategies can be synergistically deployed within community-based and digital platforms to enhance HIV prevention.

Acknowledgments

The authors sincerely thank the staff of the community-based organizations in Mianyang, Zibo, and Hohhot for their invaluable collaboration and support throughout this study. They are also deeply grateful to all the participants for their time and contribution. During the preparation of this manuscript, the authors used ChatGPT (GPT-4o) to assist with proofreading and editing for grammar and clarity. The authors reviewed and edited the artificial intelligence–generated content as needed and take full responsibility for the content of the publication.

Funding

The study was supported by the National Natural Science Foundation of China (82504531) and the Start-up Research Grant of the University of Macau (SRG2025-00031-FHS).

Authors' Contributions

Conceptualization: WC, JL
Data curation: HL
Formal analysis: WC, HL
Funding acquisition: WC
Investigation: HL, MN
Methodology: WC, HL
Project administration: WC
Resources: CC, WC
Software: HL
Supervision: WC
Validation: CC, WC
Visualization: HL
Writing—original draft: HL, WC
Writing—review and editing: PKHM, FS, OG
All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[DOCX File, 34 KB - resprot_v15i1e82123_app1.docx]

References

1. Global AIDS Strategy 2021–2026: End inequalities. End AIDS. : Joint United Nations Programme on HIV/AIDS (UNAIDS); 2021 URL: https://www.unaids.org/sites/default/files/media_asset/global-AIDS-strategy-2021-2026_en.pdf [accessed 2026-04-02]
2. National Center for AIDS/STD Control and Prevention CC. National AIDS and STI epidemic in December 2023 [Article in Chinese]. Chin J AIDS STD 2024;30(3):225. [doi: 10.13419/j.cnki.aids.2024.03.01]
3. Xu JJ, Huang XJ, Liu XC, et al. Expert consensus on pre-exposure prophylaxis for HIV in China [Article in Chinese]. Chin J AIDS STD 2020:1265-1271. [doi: 10.13419/j.cnki.aids.2020.11.34]
4. Xu JJ, Han MJ, Jiang YJ, et al. Prevention and control of HIV/AIDS in China: lessons from the past three decades. Chin Med J (Engl) 2021 Nov 10;134(23):2799-2809. [doi: 10.1097/CM9.0000000000001842] [Medline: 34759226]
5. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. : World Health Organization; 2015 URL: https://iris.who.int/server/api/core/bitstreams/90551903-a89b-49c7-abb5-47ec8ae3f23c/content [accessed 2026-04-02]
6. Buffel V, Reyniers T, Masquillier C, et al. Awareness of, willingness to take PrEP and its actual use among Belgian MSM at high risk of HIV infection: secondary analysis of the Belgian European MSM Internet Survey. AIDS Behav 2022 Jun;26(6):1793-1807. [doi: 10.1007/s10461-021-03526-z] [Medline: 34817766]

XSL•FO
RenderX

7. Zhang L, Wang H. Chinese expert consensus on HIV pre-exposure prophylaxis (2023 edition) [Article in Chinese]. Chin J AIDS STD 2023;29(9):954-961. [doi: 10.13419/j.cnki.aids.2023.09.02]
8. Liu Y, Xian Y, Liu X, et al. Significant insights from a national survey in China: PrEP awareness, willingness, uptake, and adherence among YMSM students. BMC Public Health 2024 Apr 11;24(1):1009. [doi: 10.1186/s12889-024-18512-y] [Medline: 38605341]
9. Yang Z, Wei L, Xie W, et al. Estimating changes in population size and behavioral characteristics in men who have sex with men between 2014 and 2019: longitudinal study. JMIR Public Health Surveill 2022 Aug 16;8(8):e34150. [doi: 10.2196/34150] [Medline: 35972779]
10. Dai MN, Xi Y, Yin WW, et al. Meta analysis on acceptance rate of HIV pre-exposure prophylaxis among men who have sex with men in China [Article in Chinese]. Zhonghua Yu Fang Yi Xue Za Zhi 2022 Feb 6;56(2):197-202. [doi: 10.3760/cma.j.cn112150-20210611-00570] [Medline: 35184450]
11. Pan L, Xue H, Yu F, Shan D, Zhang DP, Wang JJ. Status and associated factors of pre-exposure prophylaxis use among men who have sex with men in 24 cities in China [Article in Chinese]. Zhonghua Liu Xing Bing Xue Za Zhi 2023 Jun 10;44(6):905-911. [doi: 10.3760/cma.j.cn112338-20220831-00749] [Medline: 37380411]
12. Han J, Bouey JZ, Wang L, et al. PrEP uptake preferences among men who have sex with men in China: results from a national internet survey. J Int AIDS Soc 2019 Feb;22(2):e25242. [doi: 10.1002/jia2.25242] [Medline: 30724470]
13. Peng L, Cao W, Gu J, et al. Willingness to use and adhere to HIV pre-exposure prophylaxis (PrEP) among men who have sex with men (MSM) in China. Int J Environ Res Public Health 2019 Jul 23;16(14):2620. [doi: 10.3390/ijerph16142620] [Medline: 31340482]
14. Xiao L, Tian ML, Zhu YR, et al. The utilization of pre-exposure prophylaxis and related factors among MSM in six cities in China [Article in Chinese]. Chinese J AIDS STD 2023:650-654. [doi: 10.13419/j.cnki.aids.2023.06.07]
15. Voglino G, Gualano MR, Rousset S, et al. Knowledge, attitudes and practices regarding pre-exposure prophylaxis (PrEP) in a sample of Italian men who have sex with men (MSM). Int J Environ Res Public Health 2021 Apr 29;18(9):4772. [doi: 10.3390/ijerph18094772] [Medline: 33947130]
16. Li C, Xiong Y, Liu H, et al. Multifaceted barriers to rapid roll-out of HIV pre-exposure prophylaxis in China: a qualitative study among men who have sex with men. Int J Behav Med 2024 Apr;31(2):252-262. [doi: 10.1007/s12529-023-10177-y] [Medline: 37156941]
17. Fonner VA, Dalglish SL, Kennedy CE, et al. Effectiveness and safety of oral HIV preexposure prophylaxis for all populations. AIDS 2016 Jul 31;30(12):1973-1983. [doi: 10.1097/QAD.0000000000001145] [Medline: 27149090]
18. Martinez JE, Jonas KJ. Pre-exposure prophylaxis sorting among men who have sex with men. AIDS Care 2019 Mar;31(3):388-396. [doi: 10.1080/09540121.2018.1533229] [Medline: 30301371]
19. Mou YC, Xu JY, Ma YL, Zhang JY. Cohort retention in HlV pre-exposure prophylaxis clinical trial among men who have sex with men [Article in Chinese]. Mod Prev Med 2017:2886-2889.
20. Pan H, Li LF, Lin B, et al. Predicting pre-exposure prophylaxis adherence in men who have sex with men in Western China based on nomogram [Article in Chinese]. Chin J Health Stat 2021:528-532.
21. McMahan VM, Martin A, Garske L, et al. Development of a targeted educational intervention to increase pre-exposure prophylaxis uptake among cisgender men and transgender individuals who have sex with men and use methamphetamine in Seattle (WA, USA). Sex Health 2019 Apr;16(2):139-147. [doi: 10.1071/SH18060] [Medline: 30739638]
22. Moitra E, van den Berg JJ, Sowemimo-Coker G, Chau S, Nunn A, Chan PA. Open pilot trial of a brief motivational interviewing-based HIV pre-exposure prophylaxis intervention for men who have sex with men: preliminary effects, and evidence of feasibility and acceptability. AIDS Care 2020 Mar;32(3):406-410. [doi: 10.1080/09540121.2019.1622644] [Medline: 31130000]
23. Young LE, Schumm P, Alon L, et al. PrEP Chicago: a randomized controlled peer change agent intervention to promote the adoption of pre-exposure prophylaxis for HIV prevention among young Black men who have sex with men. Clin Trials 2018 Feb;15(1):44-52. [doi: 10.1177/1740774517730012] [Medline: 28862483]
24. Rouffiac AE, Whiteley L, Brown L, et al. A mobile intervention to improve uptake of pre-exposure prophylaxis for southern Black men who have sex with men: protocol for intervention development and pilot randomized controlled trial. JMIR Res Protoc 2020 Feb 20;9(2):e15781. [doi: 10.2196/15781] [Medline: 32130196]
25. Wang X, Xu J, Wu Z. A pilot program of pre-exposure and post-exposure prophylaxis promotion among men who have sex with men - 7 study sites, China, 2018-2019. China CDC Wkly 2020 Nov 27;2(48):917-919. [doi: 10.46234/ccdcw2020.250] [Medline: 34594800]
26. Schwarzer R. Self-regulatory processes in the adoption and maintenance of health behaviors. J Health Psychol 1999 Mar;4(2):115-127. [doi: 10.1177/135910539900400208] [Medline: 22021474]
27. Chan PA, Nunn A, van den Berg JJ, et al. A randomized trial of a brief behavioral intervention for PrEP uptake among men who have sex with men at increased risk for HIV infection. J Acquir Immune Defic Syndr 2021 Jul 1;87(3):937-943. [doi: 10.1097/QAI.0000000000002671] [Medline: 33734099]
28. Kesler MA, Kaul R, Myers T, et al. Perceived HIV risk, actual sexual HIV risk and willingness to take pre-exposure prophylaxis among men who have sex with men in Toronto, Canada. AIDS Care 2016 Nov;28(11):1378-1385. [doi: 10.1080/09540121.2016.1178703] [Medline: 27136725]

XSL•FO
RenderX

29. Williams DM, Bohlen LC, Dunsiger S, et al. Testing an insurance-based monetary incentive program for exercise: RCT design and rationale. Contemp Clin Trials 2023 Dec;135:107382. [doi: 10.1016/j.cct.2023.107382] [Medline: 37935304]
30. Boonmanunt S, Pattanaprateep O, Ongphiphadhanakul B, et al. Evaluation of the effectiveness of behavioral economic incentive programs for goal achievement on healthy diet, weight control and physical activity: a systematic review and network meta-analysis. Ann Behav Med 2023 Apr 12;57(4):277-287. [doi: 10.1093/abm/kaac066] [Medline: 36367428]
31. Galárraga O, Wilson-Barthes M, Chivardi C, et al. Incentivizing adherence to pre-exposure prophylaxis for HIV prevention: a randomized pilot trial among male sex workers in Mexico. Eur J Health Econ 2025 Mar;26(2):299-311. [doi: 10.1007/s10198-024-01705-y] [Medline: 39002005]
32. LeGrand S, Knudtson K, Benkeser D, et al. Testing the efficacy of a social networking gamification app to improve pre-exposure prophylaxis adherence (P3: Prepared, Protected, emPowered): protocol for a randomized controlled trial. JMIR Res Protoc 2018 Dec 18;7(12):e10448. [doi: 10.2196/10448] [Medline: 30563818]
33. Lee SS, Kwan TH, Wong NS, et al. Piloting a partially self-financed mode of human immunodeficiency virus pre-exposure prophylaxis delivery for men who have sex with men in Hong Kong. Hong Kong Med J 2019 Oct;25(5):382-391. [doi: 10.12809/hkmj198030] [Medline: 31619578]
34. Carey RN, Connell LE, Johnston M, et al. Behavior change techniques and their mechanisms of action: a synthesis of links described in published intervention literature. Ann Behav Med 2019 Jul 17;53(8):693-707. [doi: 10.1093/abm/kay078] [Medline: 30304386]
35. Lin H, Guo P, Chang C, et al. Development and preliminary validation of an HIV pre-exposure prophylaxis stigma scale among men who have sex with men in China [Article in Chinese]. Chin J AIDS STD 2024;30(5):487-492. [doi: 10.13419/j.cnki.aids.2024.05.07]
36. Siegler AJ, Wiatrek S, Mouhanna F, et al. Validation of the HIV pre-exposure prophylaxis stigma scale: performance of Likert and semantic differential scale versions. AIDS Behav 2020 Sep;24(9):2637-2649. [doi: 10.1007/s10461-020-02820-6] [Medline: 32157490]
37. Dangerfield DT 2nd, Davis G, Pandian V, Anderson JN. Using motivational interviewing to increase HIV PrEP initiation and adherence: a scoping review. Prev Sci 2023 Oct;24(7):1365-1375. [doi: 10.1007/s11121-023-01554-w] [Medline: 37249729]
38. Hong C, Wang L, Wang Y, et al. Barriers and facilitators to engagement with the PrEP care continuum among sexual minority men in China: a systematic review. Sex Health 2025 Aug;22:SH25021. [doi: 10.1071/SH25021] [Medline: 40977211]
39. DeAtley T, Harrison A, Mtukushe B, et al. Conditional economic incentives for HIV treatment adherence: aligning adolescent developmental hallmarks with behavioral economic theory to improve HIV treatment adherence. AIDS Patient Care STDS 2022 Jul;36(7):272-277. [doi: 10.1089/apc.2022.0060] [Medline: 35797651]
40. Galárraga O, Sosa-Rubí SG. Conditional economic incentives to improve HIV prevention and treatment in low-income and middle-income countries. Lancet HIV 2019 Oct;6(10):e705-e714. [doi: 10.1016/S2352-3018(19)30233-4] [Medline: 31578955]
41. Galárraga O, Linnemayr S, McCoy SI, Thirumurthy H, Gordon C, Vorkoper S. We must invest in behavioural economics for the HIV response. Nat Hum Behav 2023 Aug;7(8):1241-1244. [doi: 10.1038/s41562-023-01664-z] [Medline: 37463988]

Abbreviations

CBO: community-based organization
GCEI: group-based conditional economic incentive
HAPA: health action process approach
LAI-PrEP: long-acting injectable-pre-exposure prophylaxis
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by A Schwartz; submitted 11.Aug.2025; peer-reviewed by J Xu; revised version received 10.Mar.2026; accepted 10.Mar.2026; published 20.Apr.2026.

Please cite as:
Lin H, Ni M, Chang C, Sun F, Li J, Galárraga O, Mo PKH, Cao W
Overcoming the HIV Pre-Exposure Prophylaxis Uptake Barrier With A Health Action Process Approach and Conditional Economic Incentives Among Men Who Have Sex With Men in 3 Chinese Cities: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e82123
URL: https://www.researchprotocols.org/2026/1/e82123
doi:10.2196/82123

XSL•FO
RenderX

© Hao Lin, Mojun Ni, Chun Chang, Feng Sun, Jinghua Li, Omar Galárraga, Phoenix Kit-Han Mo, Wangnan Cao. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Implementation of a Cancer Navigation Intervention for Newly Diagnosed Survivors of Breast Cancer: Protocol for a Randomized Controlled Trial

Melba Sheila D'Souza[1], MScN, MPhilN, RN, PhD, CCNE; Ruby Gidda[2,3], BSN, MEd, RN; Michelle Smith[4,5], AA, MSc, Cert III VA; Arati Swaminadhan[6], MSc; Amy Jean Strank[7], BScN, RN; Ashwin Nairy[8]

[1]Nursing and Population Health, School of Nursing, Thompson Rivers University, 805 TRU Way Kamloops, Kamloops, BC, Canada

[2]School of Nursing, University of British Columbia, Abbotsford, BC, Canada

[3]Clinical Capacity Optimization, BC Cancer Agency, Abbotsford, BC, Canada

[4]Faculty of Health and Social Development, UBC School of Health and Exercise Sciences, University of British Columbia, Kelowna, BC, Canada

[5]Learning Health Systems, Research Department, Interior Health, Kelowna, BC, Canada

[6]The University of British Columbia, UBC Experimental Medicine, University of British Columbia, Vancouver, BC, Canada

[7]Home Health, Interior Health, Kamloops, BC, Canada

[8]The University of British Columbia, Faculty of Science, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Melba Sheila D'Souza, MScN, MPhilN, RN, PhD, CCNE
Nursing and Population Health, School of Nursing, Thompson Rivers University, 805 TRU Way Kamloops, Kamloops, BC, Canada

Abstract

Background: Cancer inequities among vulnerable populations in rural areas remain a public health challenge in Canada. Rural populations are defined as vulnerable due to geographic isolation, limited access to specialized oncology care, and socioeconomic barriers such as transportation and financial toxicity. Professional navigation offers a potential solution to bridge these gaps, yet there is a lack of evidence on the barriers to and facilitators of its adoption in breast cancer survivorship.

Objective: The objective of this study is to evaluate the effectiveness of a cancer navigation intervention using professional navigators compared to the standard of care (medical care) in improving the quality of life and functional outcomes of newly diagnosed survivors of breast cancer in interior British Columbia.

Methods: A single-center, parallel-group, open-cohort randomized controlled trial is being conducted over 3 years. Ethics approval was obtained for the study. Participants who provide informed consent are randomized into 2 groups: the intervention group receives the cancer navigation intervention and the control group receives the standard of care (the usual medical care offered by health care practitioners). The baseline study time point spanned January to March 2025, the first follow-up spanned April to June 2025 at 3 months after enrollment, and the second follow-up spanned July to September 2025 at the end of 6 months after enrollment. The cancer navigation intervention comprises direct psychosocial and educational webinars, coordinated telephone support services, and community-based cancer care resources. Professional navigators are qualified registered nurses who facilitate information and connect participants with available supportive resources, services, and programs. The main outcomes are financial distress, quality of life, and satisfaction with navigation and interpersonal relationships. The Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy, Functional Assessment of Cancer Therapy–Breast, Breast Cancer Navigation Survey, Participant Satisfaction With Navigation Scale, Satisfaction With Interpersonal Relationships Survey, and Breast Cancer Navigation Interview are used in the study. Steps are being taken to ensure the trustworthiness of the qualitative data. With a 5% level of significance (2 tailed) and 90% power, the sample size was calculated as 108.

Results: Data collection took place from January 2 to September 30, 2025. A total of 164 participants were recruited.

Conclusions: This study aims to demonstrate effectiveness and satisfaction with professional navigation and knowledge translation for future implementation of a cancer navigation intervention in British Columbia.

Trial Registration: International Standard Randomised Controlled Trial Number (ISRCTN) ISRCTN10220433; https://www.isrctn.com/ISRCTN10220433

International Registered Report Identifier (IRRID): DERR1-10.2196/85820

(JMIR Res Protoc 2026;15:e85820) doi:10.2196/85820

XSL•FO
RenderX

KEYWORDS

oncology; cancer; cancer navigation; early survivorship; supportive care; digital health; rural communities; patient engagement; randomized controlled trial; mixed methods

Introduction

Background

Cancer is Canada's leading cause of mortality, affecting not only individuals but also the health care system [1]. According to the Canadian Cancer Society, 2 in 5 Canadians are expected to be diagnosed with cancer in their lifetime; approximately 1 in 4 Canadians was expected to die of the disease in 2024 [2]. In 2022, a total of 2.3 million women worldwide were diagnosed with breast cancer, resulting in 670,000 deaths [3]. As of 2024, breast cancer is still the most commonly diagnosed cancer among Canadian women [2]. The Canadian Cancer Society estimates that 30,500 women will receive a breast cancer diagnosis in 2024, accounting for 25% of all new cancer cases in women. Approximately 5500 women are expected to die from the disease in 2024, which represented 13% of all cancer-related deaths among Canadian women in 2024 [2]. Early diagnoses, prompt referrals, and timely access to appropriate treatments are associated with better reported outcomes.

Breast cancer accounts for 28% of all cancer diagnoses for women and is the most common cancer in this population [4]. Ensuring equitable access to cancer care, whereby individuals receive care based on their needs, is crucial for timely treatment and the effective navigation of complex care pathways [5]. A cancer diagnosis poses challenges, and patient navigation programs assist individuals and their families in managing the complexities of survivorship. Cancer treatment involves complex services and diverse health care professionals in various units, which challenges equitable care and emphasizes the need to identify where disparities occur along the care continuum [5]. Socioeconomic factors, such as the lack of insurance and transportation as well as financial concerns, affect the access to and use of cancer care services [6]. While survival rates have improved, the transition from active treatment to survivorship care is often fragmented.

Patient navigation programs have been shown to address physical, emotional, and financial challenges through interventions such as psychosocial counseling, educational webinars, and coordinated access to community-based resources [7]. Access to and the use of health care services remains crucial for individuals following cancer treatment and throughout posttreatment care [8]. There are variations in access to cancer care after treatment and supportive care in rural areas [9]. Patient navigation programs play a crucial role in enhancing reported outcomes and addressing treatment inequities stemming from nonmedical factors [10]. Fewer interventions address psychosocial well-being and the navigation of health care services for adults diagnosed with breast cancer [11]. Hence, it is important to target reported outcomes, increase access to supportive care, and increase understanding of supportive care [10].

Approximately 17% of the population resides in British Columbia [12]. In interior regions, rural isolation, limited health care resources, and inadequate access to specialized care often lead to poorer reported outcomes [13]. Black individuals, Asian individuals, and those from other minority ethnic groups, as well as those from low-income and remote areas, likely face delayed cancer treatment [14]. Early findings from one study show that 24% of participants reported not receiving adequate emotional support during their breast cancer treatment journey and 50% felt adequately supported at diagnosis [15]. A total of 24% of survivors report insufficient emotional support.

Several studies have demonstrated positive associations between patient navigation and ameliorated distress, quality of life (QOL), and anxiety [16-19]. Patients aged 65 years and younger and those residing in rural areas reported significantly lower distress scores at discharge, supporting the value of incorporating oncology nurse navigators into care pathways [20]. People from socioeconomically disadvantaged groups in rural settings experience unique challenges in accessing care, which can impact reported outcomes [21]. The lack of continuity often leads to a limited awareness of care. Moreover, in another study, 24% of women reported insufficient emotional support during their breast cancer journey, whereas only 50% felt adequately supported at diagnosis [22]. In one study, 20% of women expressed a desire for more emotional support [6]. In another study, engagement in a patient navigation program led to improvements in distress and QOL over 18 months among participants, positively influencing patients' well-being [23]. Survivors of cancer often face psychosocial issues that can harm QOL and may decrease adherence to follow-up care [24]. Health care providers should assess psychosocial and mental health distress and provide appropriate referrals to counseling or support groups. Several factors impact patients living with cancer, such as difficulty navigating services, a lack of follow-up, and inadequate care coordination [25]. Addressing the complex needs of survivors of cancer requires an integrated approach that encompasses supportive and survivorship care planning, psychosocial support, and effective communication among health care providers [26].

Patient navigators serve as liaisons within the health care system, performing core functions, such as providing patient education, offering emotional support, coordinating logistics, and facilitating access to resources [27]. Patient navigators assist patients in understanding their care plans while addressing practical challenges such as transportation, scheduling, and financial concerns [28]. A systematic review found that patient navigator interventions decreased delays in care among survivors of cancer [29]. Improving accessibility and health literacy requires tailored interventions and consistent follow-up, informed by an assessment of the needs of immigrant and refugee women [30]. Studies have demonstrated that, with proper implementation and evaluation, patient-reported outcomes, satisfaction, and self-efficacy in underserved communities can be improved [31,32]. Despite the continued

XSL•FO
RenderX

lack of culturally and linguistically appropriate content for equity-deserving populations, the Functional Assessment of Cancer Therapy–Breast (FACT-B) remains a vital tool to measure QOL across four domains: physical, social, emotional, and functional well-being.

Aim

The aim of this study is to evaluate the effectiveness of a cancer navigation intervention compared to the standard of care (SOC) in improving QOL and functional outcomes for newly diagnosed survivors of breast cancer in interior British Columbia.

Methods

Research Design

This study uses implementation science through an open-cohort randomized controlled trial (RCT) involving an intervention and a control group. Implementation science is defined as the promotion of the systematic uptake of research findings and other evidence-based practices into routine practice to improve the quality and effectiveness of health services [33,34]. An RCT is a study in which participants are randomly allocated to receive a new treatment (experimental group), a standard treatment (comparison group), or no treatment at all (placebo group) [35]. This RCT is a single-center, parallel-group study with stratified block randomization that follows the CONSORT (Consolidated Standards of Reporting Trials) guidelines (Checklist 1) and assesses a cancer navigation intervention provided by professional navigators.

Participants and Setting

Study participants were recruited from Kamloops, Revelstoke, Merritt, 100 Mile House, Kelowna, Vernon, Penticton, and Okanagan, all situated in interior British Columbia.

Eligibility Criteria

The inclusion criteria are adults older than 18 years who can read, understand, and write in English or their preferred primary language; are newly diagnosed with and being treated for breast cancer; and are willing to participate in the study.

The exclusion criteria are adults with advanced cancer, secondary cancer, or concurrent malignancies; neonates, infants, or children younger than 19 years; and self-reported severe, undiagnosed, or untreated neurological, psychiatric, and cognitive problems impairing understanding and provision of informed consent.

Recruitment and Data Collection

Recruitment for the study took place with permission from local groups, institutions, and establishments. The recruitment call was distributed through printed posters, social media, community networks, and online advertisements. During the initial recruitment conversation, the research assistant described the study project, informing participants of what is expected, reading the study information sheet, and providing clarification. Participants were screened for eligibility and randomized into the study. Participants who met the criteria were enrolled via a web link or QR code.

Participants who provided informed consent were randomized (1:1) into one of the two study groups: the intervention group received the cancer navigation intervention and the control group received the SOC (the usual medical care offered by health care practitioners; Figure 1). The computer-based randomization balanced random assignments to the two groups in blocks of 5 participants. The baseline study time point spanned January 2025 to March 2025, the 3-month follow-up spanned April 2025 to June 2025, and the 6-month follow-up spanned July 2025 to September 2025.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.

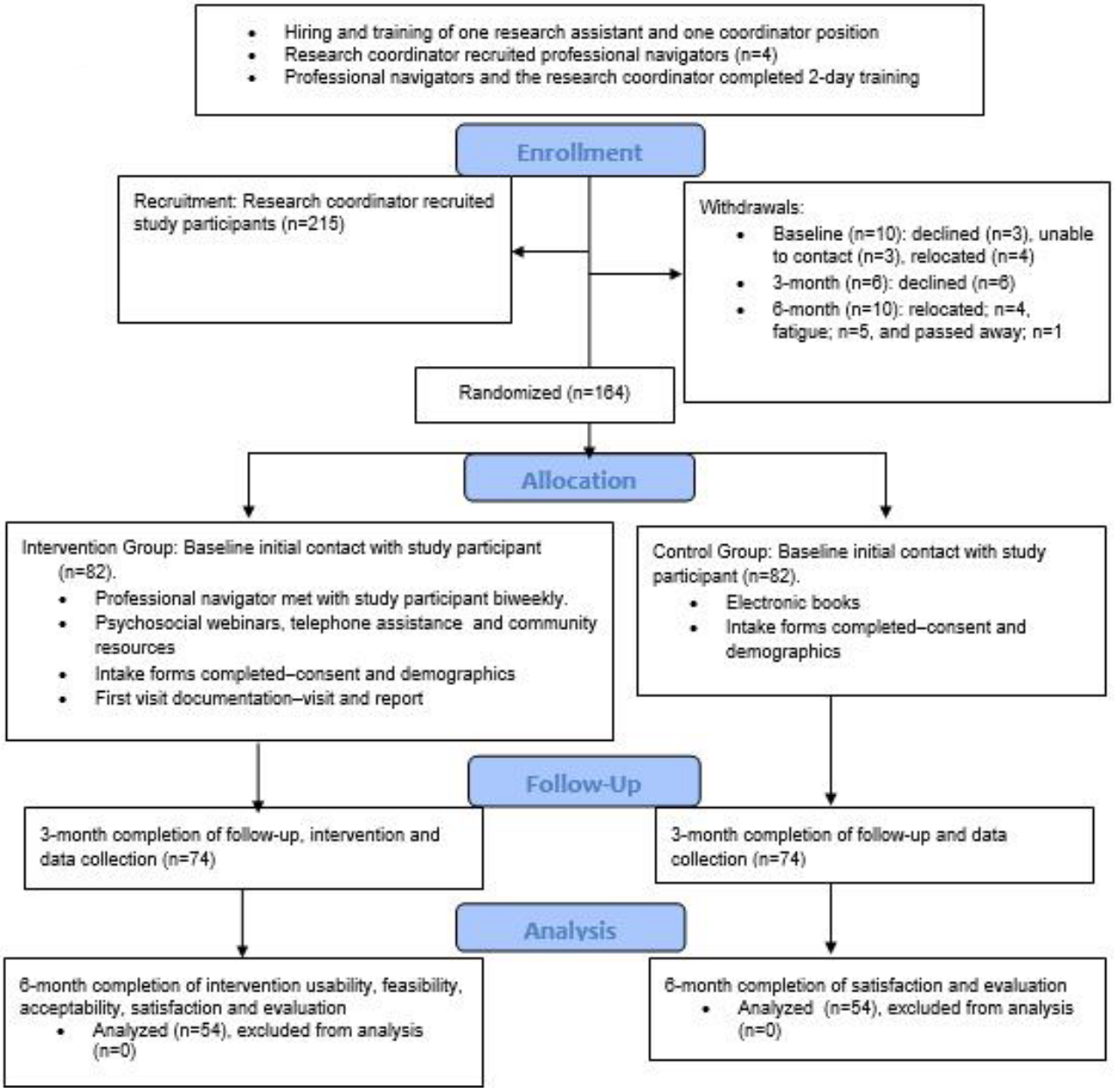

Intervention and Control Groups

The cancer navigation intervention comprises direct psychosocial and educational webinars, coordinated telephone support services, and community-based cancer care resources offered by professional navigators (Table 1). Professional navigators are qualified registered nurses trained in professional navigation to provide information, connect patients with available supportive services, and help identify information for survivorship. Participants engage in navigator follow-ups, ensuring continuous access to professional guidance.

XSL•FO
RenderX

Table . Cancer Navigation Reported Outcomes research protocol.

Activity	Protocol	Expectation
1. Recruit eligible participants	• Advertisement for participant recruitment: recruit participants by advertisement and poster announcements. • When: weekly	• Review eligibility criteria to yield eligibility determination. • When: at the time of initial participant contact
2. Initial contact by the research team	• Cover letter to participants and informed consent form and code of conduct: meet with participants via phone for study information, an introduction, and the baseline survey. • When: within 1 week of recruitment or scheduling of the initial appointment	• Speak to participants by phone for informed consent and the baseline survey: sociodemographic survey (5 minutes), COST-FACIT[a] (5 minutes), and FACT-B[b] (10 minutes) • When: within 1 week of initial participant recruitment appointment
3. Communication with the navigator	• Participant permission script: communicate with the navigator by email and phone at regular intervals regarding the participant's enrollment and check-ins. • When: within 1 week following randomization	• Communicate with the navigator by email and phone at scheduled appointments • When: within 1 week following randomization
4. Tracking participants over time	• Coordinated telephone support service: meet with participants online based on the schedule. • When: within 1 week of randomization	• When: (1) at initial intake, (2) within 1 month of randomization, and (3) at the end of each month until completion • Speak to a participant by phone at least 6 times during navigation, a total of 12 sessions.
5. Number of contact attempts to reach the participant	• Attempt to reach participants by the preferred contact method, usually by phone at least 3 times, including one in the "evening" hours, followed by 1 follow-up letter from the research team. • When: once daily (1 attempt/day for 3 different days)	• Attempt to reach participant by preferred contact method, usually by phone at least 3 times on 3 different days • When: once daily (one attempt/day for 3 different days)
6. Missed appointment follow-up	• Any missed appointments should result in a navigator phone call and completion of the navigation activity. If the participant has a history of missed appointments, the navigator should provide a reminder call the day before every scheduled appointment. • When: within 48 hours of missed appointment	• After the first missed appointment, the navigator should call to determine the cause and complete an assessment if determined necessary. • When: within 1 week of missed appointment
7. Participants lost to follow-up	• Message participants, and reach out to the research team. • When: after 3 contact attempts, being unable to reach the participant, and not showing up for scheduled appointments	• Call the participant. • When: after 3 contact attempts, being unable to reach the participant, and not showing up for the scheduled appointments
8. Completing the navigation	• Complete direct psychosocial and educational webinars: interactive teaching-learning and interactive webinars • When: end of 3 months and at the end of 6 months	• When: within 1 month of randomization and at the end of 3 months until completion • Interaction with the participant online at least two times during navigation, a total of 4 sessions. Interactive webinars with the participant online at least 2 times, a total of 4 sessions.
9. Follow-up	• Schedule of enrollment, trial, and metrics: the navigator should ask participants at each follow-up encounter if their identified needs have been met and documented.	• Follow-up by phone • When: 1 week after assessment

[a]COST-FACIT: Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy.

XSL•FO
RenderX

[b]FACT-B: Functional Assessment of Cancer Therapy–Breast.

Participants were randomized to either the cancer navigation intervention or the SOC (the control group). The cancer navigation intervention took place biweekly over 3 months and was facilitated by a professional navigator. The intervention follows tailored approaches to care based on full assessments of participants' needs and preferences. The intervention includes a full assessment, information and emotional support, and care planning specific to breast cancer survivorship. Professional navigators interact with participants in person during the intervention and follow up with them to provide psychoeducational webinars, telephone helpline support, and cancer care resources (Table 1). The intervention comprises 15 minutes of telephone calls biweekly for 3 months and 6 months, 45 minutes of interactive teaching-learning, a 30-minute webinar, and a 50-minute participant interview, totaling 5 hours (Table 2).

Table . Schedule of enrollment, trial, and metrics (outcomes) of the intervention and control groups.

	Study period and time point			
	Enrollment (T1), n	Postallocation (T1 at baseline; 0 weeks)	Follow-up (T2 at the end of 3 months; 12 weeks)	Evaluation (T3 at the end of 6 months; 12 weeks)
Enrollment, eligibility screening, informed consent, and randomization allocation	164	—[a]	—	—
Cancer navigation intervention: intervention group	82	—	—	—
Telephone calls	82	6 sessions	6 sessions	—
Interactive teaching-learning	82	2 sessions	2 sessions	—
Interactive webinar	82	2 sessions	2 sessions	—
Standard of care: control group	82	No sessions	No sessions	—
Interview and surveys				
Sociodemographic survey	164	5 minutes	5 minutes	5 minutes
Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy	164	10 minutes	—	10 minutes
The Functional Assessment of Cancer Therapy–Breast	164	5 minutes	—	5 minutes
Breast Cancer Navigation Survey	164	10 minutes,	—	10 minutes
Participant Satisfaction With Navigation Survey	164	10 minutes	—	10 minutes
Satisfaction With Interpersonal Relationships Survey	164	10 minutes	—	10 minutes
Breast Cancer Navigation Interview	82	—	—	45 - 50 minutes

[a]Not applicable.

Data Measurement Instruments and Outcomes

The main outcomes are financial distress; QOL, such as physical, social, emotional, and functional well-being; and satisfaction with navigation and interpersonal relationships, measured using the instruments outlined in this section.

A sociodemographic survey is used to collect the demographics of the participants.

The COST-FACIT was developed by de Souza et al [36] in conjunction with the University of Chicago and is a patient-reported outcome measure that describes the financial distress experienced by patients with cancer.

The FACT-B is a 37-item instrument designed to measure 5 domains of health-related QOL in patients with breast cancer: physical, social, emotional, and functional well-being and a breast cancer subscale. The FACT-B was developed with an emphasis on patients' values and brevity [37].

The Breast Cancer Navigation Survey (BCNS) comprises 10 items on a 5-point rating scale from "very dissatisfied" to "very satisfied" and 3 items on a 3-point rating scale from "very

XSL•FO
RenderX

dissatisfied" to "very satisfied" [38]. This survey takes 10 minutes to complete (Multimedia Appendix 1).

The Participant Satisfaction With Navigation Scale (PSNS) has 26 items on a 5-point rating scale, from "very satisfied" to "do not know" or "refuse to answer," was developed by a multidisciplinary team of researchers and clinicians from the National Cancer Institute (NCI)–sponsored Patient Navigation Research Program, led by Jean-Pierre [39]. This survey takes 10 minutes to complete.

The Satisfaction With Interpersonal Relationships Survey comprises 10 items on a 5-point rating scale from "very dissatisfied" to "very satisfied" and 3 items on a 3-point rating scale from "very dissatisfied" to "very satisfied" [40]. This survey takes 10 minutes to complete.

The Breast Cancer Navigation Interview includes 10 questions on perspectives on diagnosis, treatment, and navigation (Multimedia Appendix 2).

Internal consistency reliability for the COST-FACIT for participants diagnosed with breast cancer using Cronbach α for the 2-factor structure was 0.76 to 0.82. A pretest with 30 female patients with breast cancer found a mean COST-FACIT score of 21.23 (SD 9.85). Construct validity for the COST-FACIT score correlated negatively with higher psychological distress and positively with better QOL and lower out-of-pocket costs.

Internal consistency reliability for the Functional Assessment of Cancer Therapy–Breast (FACT-B) total score was Cronbach α >0.90, with subscale Cronbach α values for the physical, social, emotional, and functional domains being between 0.71 and 0.83 for participants diagnosed with breast cancer. The tool exhibited concurrent validity with the EORTC QLQ-C30 and high sensitivity to clinical changes, identifying significant shifts in performance status and disease stage via a 6- to 7-point score variance.

Internal consistency reliability for the BCNS using the Cronbach α was 0.70 to 0.78 when pretesting with participants diagnosed with breast cancer. Face and content validity were established through an expert panel consisting of 5 members: one oncologist, one physician, one nurse, one patient, and one caregiver.

Internal consistency reliability for the PSNS using the Cronbach α was 0.81 when pretesting among participants diagnosed with breast cancer. Predictive validity showed that higher satisfaction scores correlated with lower levels of decisional conflict and higher adherence to follow-up treatments. The mean satisfaction score was 3.7 out of 5.0.

Internal consistency reliability for the Satisfaction With Interpersonal Relationships Survey ranged from α=0.75 to α=0.84 among participants with breast cancer. The tool exhibited discriminant validity by differentiating between technical competence and interpersonal attributes. Satisfaction regarding trust was associated with emotional well-being subscales.

Trustworthiness

To ensure the trustworthiness of the qualitative findings, the researchers followed the criteria of credibility, dependability, confirmability, and transferability. Credibility and confirmability were preserved through reflexivity and bracketing; researchers used reflexive journaling to critically examine and set aside preconceived notions, ensuring the findings emerged directly from the lived experiences of participants. To establish dependability, an audit trail was carefully maintained to document and track all analytical decisions. These strategies enhanced the transferability of the work, providing a rigorous framework for the study.

Sample Size and Power

The study is sufficiently powered to detect a clinically meaningful difference while accounting for the longitudinal design. For effect size, the power assumption calculation was based on a minimum clinically significant difference of 7 points on the FACT-B [41]. An estimated SD of 16.93 was used, derived from existing literature on adult survivors of cancer. An attrition rate of 25% was estimated based on the potential for dropout during the 6-month follow-up schedule. While the study is a single-center RCT, the randomization is balanced using blocks of 5 to mitigate potential clustering effects and ensure equal distribution across the 2 arms. With a 5% level of significance (2 tailed) and 90% power, the required sample size was calculated as 108 participants. The total recruitment target is set to 150 participants (75 in the intervention group and 75 in the control group).

Ethical Considerations

The study received an approval certificate from the BC Cancer Agency, Interior Health Authority, and Thompson Rivers University Research Ethics Board (H22-03105/27-04-2023/27-04-2024). Research team members completed the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans, a course on research ethics. Participants will receive a gift voucher for their time and efforts at baseline and after 3 months.

Quality, Safety, and Monitoring

The research assistants developed a master sheet, which includes the study location and site; participant codes (study ID); follow-up tracking; the participant's name and contact information (email, phone number, and address); the randomization arm; date and time and modules completed; and whether the intervention is in progress or incomplete. Research assistants reminded participants via email and phone to review intervention module information. Discussions among the research team addressed any adverse effects that could result from study participation, the validity and integrity of the data, the enrollment rate relative to expectations, participant characteristics, retention, adherence to the protocol (potential or real deviations), and data completeness. We have established a data and safety monitoring plan to ensure participants' safety. These activities are reviewed by screening results and other available data at weekly research meetings among the principal investigator, the research team, and other relevant research team members; in a quarterly review of data safety and enrollment

XSL•FO
RenderX

by the research team, including the principal investigator, research assistants, and relevant research team members; and in an annual review by the ethics board.

Data Analysis

All quantitative data will be analyzed using the intention-to-treat principle. Participants who engaged with the study from October 2025 to December 2025 will be included in the analysis.

The FACT-B is the primary measure used to evaluate the effectiveness of professional navigation on health-related QOL. Changes in FACT-B total scores and the scores on its 5 subscales (physical, social, emotional, and functional well-being and the breast cancer subscale) from baseline (T1) to 3 months (T2) and 6 months (T3) will be compared between the cancer navigation intervention and SOC control groups. A repeated-measure ANOVA will be used to determine the interaction between time and the treatment group. A change of 6 to 7 points on the FACT-B is considered clinically significant for this population.

Financial toxicity is an indicator of socioeconomic distress. Financial burden will be quantified using the COST-FACIT instrument. Descriptive statistics will be generated. To determine the impact of financial toxicity on health-related QOL, ANOVAs will be conducted to assess the associations between COST-FACIT scores and health-related QOL outcomes.

The BCNS analyzes participant satisfaction across 13 items to evaluate the user-friendliness and clarity of the program objectives.

The PSNS uses 26 items to measure overall satisfaction with the navigation process. Higher scores will be correlated with secondary outcomes such as treatment adherence using predictive validity testing.

The Satisfaction With Interpersonal Relationships Survey evaluates soft skills and trust in the professional navigator. Discriminant validity will be assessed by distinguishing between technical competence and interpersonal satisfaction scores.

Qualitative data from the Breast Cancer Navigation Interview will be analyzed using the thematic analysis approach by Braun and Clarke [42] to identify barriers to and facilitators of the adoption of professional navigation. These qualitative findings will be synthesized with the quantitative results to explain the how and why behind the intervention and effectiveness data. Two independent researchers will read the transcripts multiple times to achieve immersion. Participant quotations will be analyzed in terms of the content; what was being discussed; context; language use, such as metaphors, symbols, repetitions, and pauses; and initial interpretative comments by the participants [43]. Researchers will look at associations among emerging themes and group them according to conceptual similarities; track them; log them; and then extract major themes, subthemes, and relevant excerpts from the interview transcripts. Codes will be generated to capture emerging concepts and grouped into clusters based on conceptual similarities to categorize themes.

Results

Study Timeline

Data collection took place from January 2 to September 30, 2025. Funding was secured in June 2022; the study has progressed through data validation and cleaning; publication of the results is anticipated in December 2026. A total of 164 participants were recruited at baseline and randomized into either the intervention group (n=82) or the control group (n=82), as shown in Figure 1. At the end of the 6-month follow-up, 108 participants had completed the study (54 in the intervention group and 54 in the control group), resulting in a response rate of 65.85%.

Sociodemographic Characteristics

In the intervention group (n=54), the largest age group was 40-49 years (17/54, 31.48%), whereas the control group (n=54) was slightly older, with the largest age group being 50-59 years (24/54, 44.44%). Educational attainment was comparable across both groups, with individuals who had completed college, associate, or technical levels comprising 33.33% (18/54) of the intervention group and 37.04% (20/54) in the control group.

Differences were observed in employment and residency. The control group reported a higher employment rate (25/54, 46.3%) than the intervention group (16/54, 29.63%). A higher percentage of the intervention group resided in rural areas (44/54, 81.48%) compared to the control group (30/54, 55.56%).

Functional Assessment of Cancer Therapy–Breast

At the 6-month follow-up, participants in the intervention group demonstrated higher mean scores across all FACT-B domains than the control group. The intervention group reported higher mean scores (mean 25.6, SD 5.2) on the breast cancer subscale than the control group (mean 24.5, SD 7.1; $P<.01$), indicating improved QOL outcomes.

Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy

Age group ($P=.05$), education level ($P<.05$), employment status ($P=.01$), and place of residence ($P<.03$) were significantly associated with COST-FACIT scores.

Discussion

The study seeks to identify how navigation can enhance supportive and survivorship care and anticipates improving QOL. This study illustrates findings that emphasize the necessity of addressing the transition to survivorship through dedicated psychosocial support. The trial is recruiting participants from rural communities, yet it remains a single-center study, which may limit the generalizability of the findings to other provinces. The effectiveness of the study components may be hindered by varying levels of health literacy. The inclusion criteria allow participants who speak English or their preferred primary language, but the availability of culturally and linguistically tailored services remains a challenge for equity-deserving groups. A longitudinal design over 6 months introduces the

XSL•FO
RenderX

possibility of attrition, which could impact the analysis even with the use of the intention-to-treat principle.

This RCT plans to establish an intervention framework for evaluating the effectiveness of a cancer navigation intervention in improving QOL and functional outcomes for newly diagnosed survivors of breast cancer. The study aims to demonstrate how professional navigation can bridge the gap between medical care and the psychosocial needs of rural and vulnerable populations. This study seeks to provide a contextual, evidence-informed model for professional navigation to enhance patient satisfaction. The study will demonstrate that professional navigators enhance trust and communication, reinforcing the value of person-centered care. Data will be disseminated through community engagement forums that will aim to influence health policy and evidence synthesis regarding breast cancer survivorship.

Acknowledgments

This study acknowledges the partnerships and collaborations with BC Cancer, Interior Health, Royal Inland Hospital Foundation, Rae Fawcett Breast Health Clinic, Thompson Region Division of Family Practice, Beyond Breast Cancer, Kamloops Immigrant Services, Kamloops Massage Matters, Kamloops Breast Cancer Support Group, The Spirit Warriors Society, InspireHealth Supportive Cancer Care, and Sensational Survivors Kamloops.

Funding

This work was funded by Breast Cancer Canada (2022-005-R/23-06-2022-22-06-2026).

Data Availability

The datasets generated during the study will be available from the corresponding author on reasonable request. To protect participant privacy, data will be coded and de-identified in accordance with the institutional review board requirements before sharing.

Authors' Contributions

MSD, RG, MS, A Swaminadh, A Strank, and AN made substantial contributions to the conception and design of the work, methodology, software, validation, funding acquisition, data curation, analysis, interpretation of data for the work. MSD, RG, MS, A Swaminadh, A Strankand AN drafted the original preparation, edited and reviewed it critically for important intellectual content; and gave final approval of the version to be published. MSD, RG, MS, A Swaminadh, A Strank and AN agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Breast Cancer Navigation Survey and Participant Satisfaction With Navigation Scale.
[PDF File, 270 KB - resprot_v15i1e85820_app1.pdf]

Multimedia Appendix 2
Breast Cancer Navigation Interview.
[PDF File, 191 KB - resprot_v15i1e85820_app2.pdf]

Checklist 1
CONSORT checklist.
[PDF File, 172 KB - resprot_v15i1e85820_app3.pdf]

References

1. Brenner DR, Poirier A, Woods RR, et al. Projected estimates of cancer in Canada in 2022. CMAJ 2022 May 2;194(17):E601-E607. [doi: 10.1503/cmaj.212097] [Medline: 35500919]
2. Canadian cancer statistics. Canadian Cancer Society. URL: https://cancer.ca/en/research/cancer-statistics/canadian-cancer-statistics [accessed 2026-4-10]
3. Breast cancer. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/breast-cancer [accessed 2026-4-10]
4. Breast. BC Cancer. URL: http://www.bccancer.bc.ca/health-professionals/clinical-resources/cancer-management-manual/breast/breast [accessed 2026-4-10]

XSL•FO
RenderX

5. Dixit N, Rugo H, Burke NJ. Navigating a path to equity in cancer care: the role of patient navigation. Am Soc Clin Oncol Educ Book 2021 Mar;41:1-8. [doi: 10.1200/EDBK_100026] [Medline: 33830828]
6. D'Souza MS, Apelegan T, Nairy AA. Experiences and meaningfulness of breast cancer survivorship care in improving the quality of life of immigrant women: a qualitative systematic review. Can Oncol Nurs J 2024;34(3):304-356. [doi: 10.5737/23688076343304] [Medline: 39502095]
7. What is breast cancer? Canadian Breast Cancer Network. URL: https://www.cbcn.ca/en/what-is-breast-cancer [accessed 2026-4-10]
8. Bourgeois A, Horrill TC, Mollison A, Lambert LK, Stajduhar KI. Barriers to cancer treatment and care for people experiencing structural vulnerability: a secondary analysis of ethnographic data. Int J Equity Health 2023 Mar 30;22(1):58. [doi: 10.1186/s12939-023-01860-3] [Medline: 36998035]
9. Ozcan BB, Dogan BE, Mootz AR, et al. Breast cancer disparity and outcomes in underserved women. Radiographics 2024 Jan;44(1):e230090. [doi: 10.1148/rg.230090] [Medline: 38127658]
10. Howard EC, Murray Horwitz ME, Gunn CM, et al. Burden and trajectory of social needs after breast cancer diagnosis at a safety-net hospital. Breast Cancer Res Treat 2024 Oct;207(3):579-585. [doi: 10.1007/s10549-024-07389-5] [Medline: 38851661]
11. D'Souza MS, Bacsu JD, Sharma A, Nairy A. Culturally sensitive approaches in psychosocial interventions to enhance well-being of immigrant adults diagnosed with breast cancer: a systematic review. Int J Environ Res Public Health 2025 Feb 25;22(3):335. [doi: 10.3390/ijerph22030335] [Medline: 40238326]
12. Population growth in Canada's rural areas, 2016 to 2021. Government of Canada. 2022 Feb 9. URL: https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-x/2021002/98-200-x2021002-eng.cfm [accessed 2026-04-10]
13. Kornelsen J, Carthew C, Míguez K, et al. Rural citizen-patient priorities for healthcare in British Columbia, Canada: findings from a mixed methods study. BMC Health Serv Res 2021 Sep 18;21(1):987. [doi: 10.1186/s12913-021-06933-z] [Medline: 34537027]
14. Giaquinto AN, Sung H, Newman LA, et al. Breast cancer statistics 2024. CA Cancer J Clin 2024;74(6):477-495. [doi: 10.3322/caac.21863] [Medline: 39352042]
15. D'Souza MS, Latif E, McCarthy A, Karkada SN. Experiences and perspectives of ethnocultural breast cancer survivors in the interior region of British Columbia: a descriptive cross-sectional approach. Clin Epidemiol Glob Health 2022;16:101095. [doi: 10.1016/j.cegh.2022.101095]
16. Chan RJ, Milch VE, Crawford-Williams F, et al. Patient navigation across the cancer care continuum: an overview of systematic reviews and emerging literature. CA Cancer J Clin 2023;73(6):565-589. [doi: 10.3322/caac.21788] [Medline: 37358040]
17. Gonzalez-Ling A, Galindo Vázquez O, Espinoza Bello M, et al. Quality of life, anxiety, depression, and distress in patients with advanced and metastatic lung cancer. Palliat Support Care 2023 Aug;21(4):608-615. [doi: 10.1017/S147895152200116X]
18. Papadopoulou A, Govina O, Tsatsou I, et al. Quality of life, distress, anxiety and depression of ambulatory cancer patients receiving chemotherapy. Med Pharm Rep 2022 Oct;95(4):418-429. [doi: 10.15386/mpr-2458] [Medline: 36506609]
19. Qaderi SM, Swartjes H, Vromen H, Bremers AJ, Custers JA, de Wilt JH. Acceptability, quality of life and cost overview of a remote follow-up plan for patients with colorectal cancer. Eur J Surg Oncol 2021 Jul;47(7):1637-1644. [doi: 10.1016/j.ejso.2020.12.018] [Medline: 33423826]
20. Tulk J, Wurz A, Hou SH, et al. Rural-urban differences in distress, quality of life, and social support among Canadian young adult cancer survivors: a Young Adults with Cancer in Their Prime (YACPRIME) study. J Rural Health 2024 Jan;40(1):121-127. [doi: 10.1111/jrh.12774] [Medline: 37355833]
21. D'Souza MS, Fehr FC, Smith ML, Marshall MC. Mediators of psychosocial well-being for immigrant women living with breast cancer in Canada: a critical ethnography. J Psychosoc Oncol Res Pract 2023;5(4):119. [doi: 10.1097/OR9.0000000000000119]
22. Jenstad LM, Howe T, Breau G, et al. Communication between healthcare providers and communicatively-vulnerable patients with associated health outcomes: a scoping review of knowledge syntheses. Patient Educ Couns 2024 Feb;119:108040. [doi: 10.1016/j.pec.2023.108040] [Medline: 37951163]
23. Bell JG, Shaffer LE, Schott VA, et al. Patient navigation effect on cancer patients' quality of life and distress. J Oncol Navig Surviv 2020;11(10) [FREE Full text]
24. Ramirez AG, Choi BY, Munoz E, et al. Assessing the effect of patient navigator assistance for psychosocial support services on health-related quality of life in a randomized clinical trial in Latino breast, prostate, and colorectal cancer survivors. Cancer 2020 Mar 1;126(5):1112-1123. [doi: 10.1002/cncr.32626] [Medline: 31743436]
25. D'Souza MS, O'Mahony J, Karkada SN. Effectiveness and meaningfulness of breast cancer survivorship and peer support for improving the quality of life of immigrant women: a mixed methods systematic review protocol. Clin Epidemiol Glob Health 2021;10:100678. [doi: 10.1016/j.cegh.2020.100678]
26. Oppong BA, Rumano RP, Paskett ED. Expanding the use of patient navigation: health coaching-based navigation as a novel approach to addressing deficits in breast cancer survivorship support. Breast Cancer Res Treat 2024 May;205(1):1-3. [doi: 10.1007/s10549-023-07213-6] [Medline: 38273216]

XSL•FO
RenderX

27. Miller SD, Urquhart R, Kephart G, Asada Y, Younis T. Nurse navigators' views on patient and system factors associated with navigation needs among women with breast cancer. Curr Oncol 2021 Jun 5;28(3):2107-2114. [doi: 10.3390/curroncol28030195] [Medline: 34198915]
28. Budde H, Williams GA, Scarpetti G, Kroezen M, Maier CB. What are patient navigators and how can they improve integration of care? : European Observatory on Health Systems and Policies; 2022 URL: https://iris.who.int/server/api/core/bitstreams/e089c8d2-26a4-48ac-a63f-cd1354fa8fde/content [accessed 2026-04-10]
29. Chen M, Wu VS, Falk D, Cheatham C, Cullen J, Hoehn R. Patient navigation in cancer treatment: a systematic review. Curr Oncol Rep 2024 May;26(5):504-537. [doi: 10.1007/s11912-024-01514-9] [Medline: 38581470]
30. Racine L, D'Souza MS, Tinampay C. Effectiveness of breast cancer screening interventions in improving screening rates and preventive activities in Muslim refugee and immigrant women: a systematic review and meta-analysis. J Nurs Scholarsh 2023 Jan;55(1):329-344. [doi: 10.1111/jnu.12818] [Medline: 36199240]
31. Rahimipour Anaraki N, Mukhopadhyay M, Wilson M, Karaivanov Y, Asghari S. Virtual healthcare in rural and remote settings: a qualitative study of Canadian rural family physicians' experiences during the COVID-19 pandemic. Int J Environ Res Public Health 2022 Oct 17;19(20):13397. [doi: 10.3390/ijerph192013397] [Medline: 36293990]
32. Aapro M, Bossi P, Dasari A, et al. Digital health for optimal supportive care in oncology: benefits, limits, and future perspectives. Support Care Cancer 2020 Oct;28(10):4589-4612. [doi: 10.1007/s00520-020-05539-1] [Medline: 32533435]
33. Eccles MP, Mittman BS. Welcome to implementation science. Implement Sci 2006;1:1. [doi: 10.1186/1748-5908-1-1]
34. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? Psychiatry Res 2020 Jan;283:112376. [doi: 10.1016/j.psychres.2019.04.025] [Medline: 31036287]
35. Braga LH, Farrokhyar F, Dönmez M, et al. Randomized controlled trials - the what, when, how and why. J Pediatr Urol 2025 Apr;21(2):397-404. [doi: 10.1016/j.jpurol.2024.11.021] [Medline: 39701869]
36. de Souza JA, Yap BJ, Wroblewski K, et al. Measuring financial toxicity as a clinically relevant patient-reported outcome: the validation of the COmprehensive Score for financial Toxicity (COST). Cancer 2017 Feb 1;123(3):476-484. [doi: 10.1002/cncr.30369] [Medline: 27716900]
37. Brady MJ, Cella DF, Mo F, et al. Reliability and validity of the Functional Assessment of Cancer Therapy-Breast quality-of-life instrument. J Clin Oncol 1997 Mar;15(3):974-986. [doi: 10.1200/JCO.1997.15.3.974] [Medline: 9060536]
38. Skrutkowski M, Saucier A, Meterissian S. The breast cancer patient navigation kit: development and user feedback. J Cancer Educ 2011 Dec;26(4):782-792. [doi: 10.1007/s13187-011-0216-0] [Medline: 21461955]
39. Carle AC, Jean-Pierre P, Winters P, et al. Psychometric evaluation of the patient satisfaction with logistical aspects of navigation (PSN-L) scale using item response theory. Med Care 2014 Apr;52(4):354-361. [doi: 10.1097/MLR.0000000000000089] [Medline: 24848207]
40. Jean-Pierre P, Shao C, Cheng Y, Wells KJ, Paskett E, Fiscella K. Patient Satisfaction with Navigator Interpersonal Relationship (PSN-I): item-level psychometrics using IRT analysis. Support Care Cancer 2020 Feb;28(2):541-550. [doi: 10.1007/s00520-019-04833-x] [Medline: 31076896]
41. Giner-Sorolla R, Montoya AK, Reifman A, et al. Power to detect what? Considerations for planning and evaluating sample size. Pers Soc Psychol Rev 2024 Aug;28(3):276-301. [doi: 10.1177/10888683241228328] [Medline: 38345247]
42. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2022;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
43. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

Abbreviations

BCNS: Breast Cancer Navigation Survey
CONSORT: Consolidated Standards of Reporting Trials
COST-FACIT: Comprehensive Score for Financial Toxicity–Functional Assessment of Chronic Illness Therapy
FACT-B: Functional Assessment of Cancer Therapy–Breast
PSNS: Participant Satisfaction With Navigation Scale
QOL: Quality of life
RCT: Randomized controlled trial
SOC: Standard of care

XSL•FO
RenderX

Edited by A Schwartz; submitted 13.Oct.2025; peer-reviewed by D Chong, M Lemonde; revised version received 16.Mar.2026; accepted 16.Mar.2026; published 20.Apr.2026.

Please cite as:
D'Souza MS, Gidda R, Smith M, Swaminadhan A, Strank AJ, Nairy A
Implementation of a Cancer Navigation Intervention for Newly Diagnosed Survivors of Breast Cancer: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e85820
URL: https://www.researchprotocols.org/2026/1/e85820
doi:10.2196/85820

© Melba Sheila D'Souza, Ruby Gidda, Michelle Smith, Arati Swaminadhan, Amy Jean Strank, Ashwin Nairy. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical Evaluation of Pediatric Pulse Oximeters in South Africa: Protocol for a Cluster Randomized Controlled Trial

Holly B Schuh[1,2], PhD; Marieke Van der Zalm[3], PhD; Margaret Van Niekerk[3], MNutrition; Lise-Marie Laubscher[3], MBChB; Abenathi Mcinziba[3], MPhil; Steve Gomas[4]; Sunaina Kapoor[2], MD; Shubhada Hooli[2,5], MD; Lario Viljoen[3], PhD; Andre Gie[6], PhD; Bareng A S Nonyane[7], PhD; Pierre Goussard[6], PhD; Anneke C Hesseling[3], PhD; Carina King[8], PhD; Eric D McCollum[2,3], MD

[1]Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

[2]Global Program in Pediatric Respiratory Sciences, Eudowood Division of Pediatric Respiratory Sciences, Department of Pediatrics, Johns Hopkins School of Medicine, Rubenstein Building, 200 North Wolfe Street, Baltimore, MD, United States

[3]Desmond Tutu TB Centre, Department of Paediatrics and Child Health, Stellenbosch University, Clinical Building, Francie Van Zijl Dr., Parow Valley, Cape Town, South Africa

[4]Springer Design, Danville, CA, United States

[5]Division of Pediatric Emergency Medicine, Department of Pediatrics, Baylor College of Medicine, Houston, TX, United States

[6]Department of Paediatrics and Child Health, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

[7]Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

[8]Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Eric D McCollum, MD
Global Program in Pediatric Respiratory Sciences, Eudowood Division of Pediatric Respiratory Sciences, Department of Pediatrics, Johns Hopkins School of Medicine, Rubenstein Building, 200 North Wolfe Street, Baltimore, MD, United States

Abstract

Background: The burden of children with lower respiratory infections and low blood oxygen levels (hypoxemia) is high, and outcomes are poor in low- and middle-income countries (LMICs). Pulse oximeters noninvasively measure the capillary oxyhemoglobin saturation (SpO_2) to identify hypoxemia, but high-quality devices designed for the unique needs of children are rarely available in primary health care clinics (PHCs) in LMICs, where children initially access care.

Objective: This study aims to evaluate whether 2 pediatric pulse oximeters co-designed with health care workers (HCWs) in LMICs improve the correct SpO_2 management of children in PHCs compared to a standard pulse oximeter.

Methods: We are conducting a pragmatic 3-arm cluster randomized controlled trial in the Eastern Khayelitsha, Northern, and Tygerberg areas of Cape Town, South Africa, over 18 months between 2024 and 2026. We plan to enroll 1200 children aged younger than 2 years with an acute respiratory infection from 18 PHCs randomized to implement one of 3 pulse oximeters, either 1 standard-of-care device or 2 intervention devices. HCWs in selected PHCs will administer the intervention. Our primary outcome will be "correct SpO_2 management," an intermediate clinical end point between device implementation and hypoxemia outcome, defined by the following three elements necessary to reduce inappropriately treated hypoxemia: (1) device adoption—HCW use of the device as evidenced by a HCW-documented SpO_2 and pulse rate; (2) quality SpO_2 measurement—SpO_2 confirmed by reference device measurement within 2% SpO_2 above or below the HCW-measured SpO_2; and (3) correct SpO_2 decision-making—an appropriate referral recommendation by the HCW. A concurrent mixed methods process evaluation will explore how, why, for whom, and to what extent these devices impact the clinical management of hypoxemic children. The primary analysis will be intention-to-treat. For all primary and secondary outcomes, we will conduct pairwise comparisons between the 2 intervention arms and the control arm.

Results: Data collection commenced in 2024, and results are expected from 2026 to 2027. As of December 2025, enrollment has been completed in 12 clinics.

Conclusions: While there are notable challenges inherent in designing a trial to evaluate whether pulse oximeters improve HCW SpO_2 management of children at PHCs, our protocol development process attempted to address all potential limitations and sources of bias to maximize the trial's future impact.

Trial Registration: ClinicalTrials.gov NCT05914324; https://clinicaltrials.gov/study/NCT05914324

International Registered Report Identifier (IRRID): PRR1-10.2196/82888

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e82888) doi:10.2196/82888

KEYWORDS

respiratory tract infections; pneumonia; hypoxemia; child; infant; pulse oximetry; health facilities; low- and middle-income countries; clinical decision-making

Introduction

Lower respiratory infections, including pneumonia, remain the leading infectious cause of death globally for children aged younger than 5 years [1,2]. Most lower respiratory infection–related deaths occur in low- and middle-income countries (LMICs), with inequitable distribution both between and within countries [1,3,4]. Key quality-of-care implementation gaps have hampered the effectiveness of the World Health Organization's (WHO) Integrated Management of Childhood Illnesses (IMCI) guidelines [5], which are commonly used for pediatric pneumonia case management in LMICs. Interventions to improve the identification of severely ill children within the IMCI approach could improve outcomes, with routine use of pulse oximetry, to noninvasively measure blood oxygen levels via the capillary oxyhemoglobin saturation (SpO_2), providing one such opportunity [6,7].

Hypoxemia, defined as a low SpO_2, is both common in children with pneumonia, estimated at 28% within the African context and 23% in outpatient settings [8], and is associated with increased mortality [9]. Therefore, effectively identifying these children early in their care-seeking pathway is fundamental to reducing mortality in high-burden, low-resource contexts. While IMCI recommends SpO_2 measurements for children with suspected pneumonia, pulse oximetry devices for measuring SpO_2 are not widely implemented in primary health care clinics (PHCs) in LMICs, where most children first access care [6]. Referral decisions are therefore largely based on subjective clinical danger signs, which do not accurately identify hypoxemia [10,11]. Routine care is ineffective at identifying treatment failure in nonsevere pneumonia patients sent home with first-line antibiotics, and little evidence exists on its burden [12-14]. Estimating the prevalence of treatment failure among children with pneumonia at PHCs would fill a key evidence gap on outcomes and resources needed in this setting.

A challenge in the implementation of pediatric pulse oximetry is that devices are largely designed for surgical and inpatient continuous monitoring in high-resource and adult populations. The use case for outpatient pediatric patients seen in low-resource settings differs considerably from adults—ease of placement, robustness, battery life, motion artifact and blood perfusion variability, and smaller surface areas for measurement are important and can pose greater difficulty [15]. To date, there have been a limited number of efforts to design pulse oximeter devices specifically for this target group, meaning that many devices currently used in pediatric outpatients are not fit for purpose. Devices that have been promoted for use on infants and children are either too costly for scale-up in LMICs, as they are dependent on expensive and often fragile specialized probes, or their performance has not been validated in the pediatric population, leading to performance issues when implemented in real-world contexts.

The Phefumla (meaning "breathe" in isiXhosa) trial will evaluate the clinical impact of 2 novel pediatric pulse oximeters, both co-designed with HCWs and specifically for children presenting to high-burden and low-resource outpatient settings, on the quality of care for children with acute respiratory infections (ARIs) in Cape Town, South Africa. Both devices were designed to address known issues with pediatric oximetry measurements related to poor probe fit onto the digits of infants and children and related performance issues, but with different approaches. Specifically, one device is an investigational probe-less reflectance oximeter designed to obtain measurements on digits, hands, feet, and the forehead [16]. The second device is a regulatory-approved transmissive oximeter clip probe uniquely designed to fit on the big toe and across the foot of infants and younger children and on the fingers of older children [17]. In transmissive oximetry, light passes through tissue to a detector located on the opposite side of the body, while in reflectance oximetry, light is emitted and detected on the same side of the body [18]. This trial aims to understand not just the impact of these pediatric pulse oximeters on clinical outcomes but also how these devices perform and are implemented within routine care, when compared to standard pulse oximeter devices in PHCs in the Western Cape of South Africa.

Methods

Study Design

The Phefumla trial is a pragmatic 3-arm cluster randomized controlled trial consistent with elements of a type 1 hybrid effectiveness-implementation design, given that the primary end point ("correct SpO_2 management") is a composite intermediate clinical outcome (see the Trial Outcomes section) [19]. The trial is conducted in the Eastern Khayelitsha, Northern, and Tygerberg areas of Cape Town, South Africa, over an 18-month period between 2024 and 2026 (Figure 1). A total of 18 clusters, defined as PHCs, will be randomized to the Phefumla device [16] (prototype version 1.1; Springer Designs), the LB-01 device [17] (Acare), or standard care, with the primary outcome being "correct SpO_2 management" among children aged younger than 2 years with ARI (Textbox 1). Implementation outcomes will be defined according to the "reach, effectiveness, adoption, implementation, and maintenance" (RE-AIM) framework [19,20] and will be used to explore how, why, for whom, and to what extent these novel pediatric pulse oximetry devices impact the clinical management of hypoxemic children.

XSL•FO
RenderX

Figure 1. Study design. SpO_2: capillary oxyhemoglobin saturation.

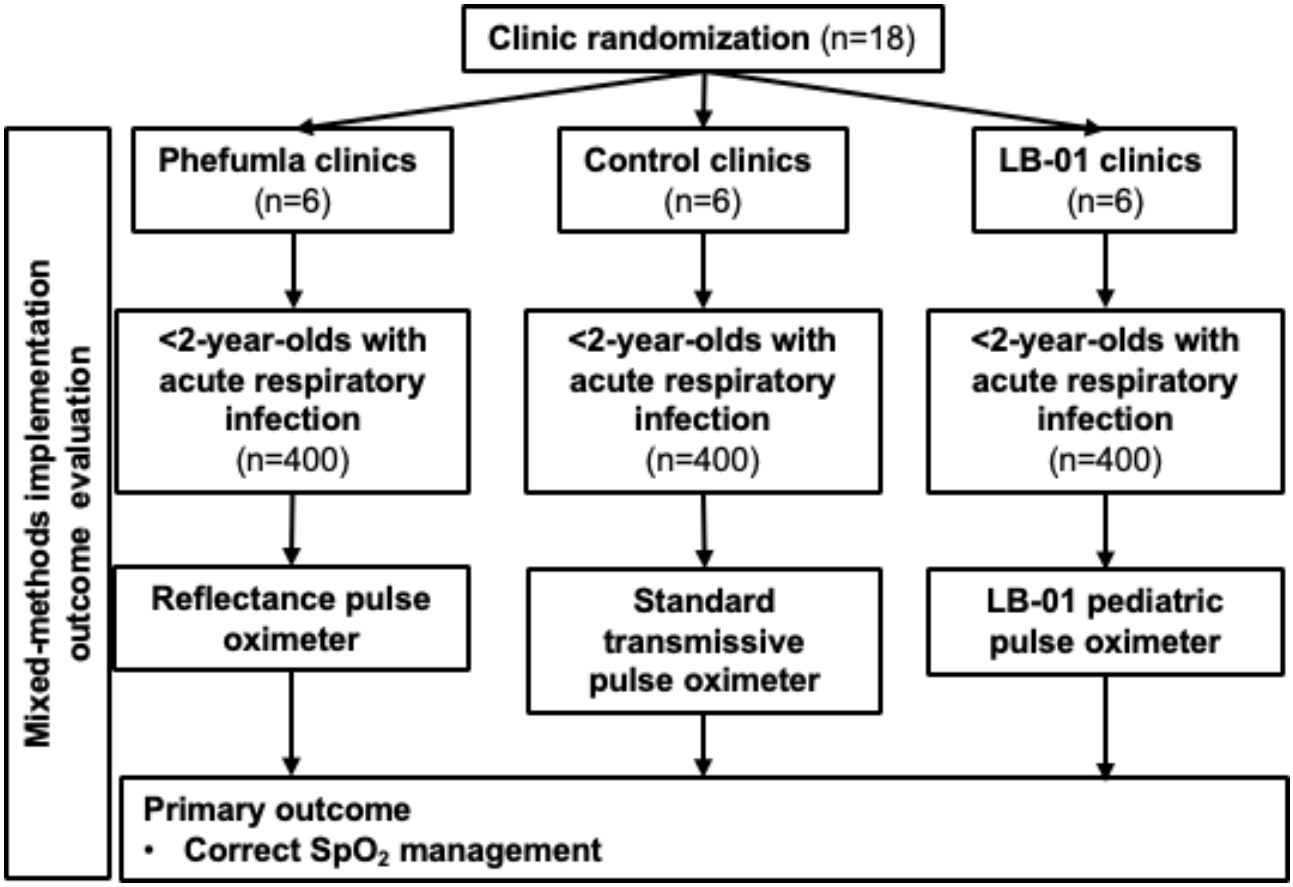

Textbox 1. Primary trial research question and hypotheses.

Research question: Do pediatric pulse oximeters designed for low-resource contexts improve capillary oxyhemoglobin saturation management in primary health care clinics (PHCs) compared to standard care?

Hypothesis 1: Health care workers (HCWs) using the LB-01 oximeter will correctly manage a higher proportion of children aged 0 to <24 months at PHCs with acute respiratory infection (ARI) than with the current standard pulse oximeter.

Hypothesis 2: HCWs using the prototype Phefumla oximeter will correctly manage a higher proportion of children aged 0 to <24 months at PHCs with ARI than with the current standard pulse oximeter.

The evaluation was designed in line with the RE-AIM framework [20]. Effectiveness will be assessed among children aged younger than 2 years through a prospective cohort of children presenting to study clinics. Reach, adoption, and implementation will be assessed through the mixed method process evaluation, involving facility audits, qualitative observations, and health care worker skills assessments. Maintenance of HCW pulse oximeter practice will be assessed through a follow-up facility visit 6 months after patient recruitment is completed.

We will also nest a descriptive prospective cohort study throughout the trial period to evaluate clinical outcomes for recruited children, pooling 3 trial arms and following them for 2 weeks after enrollment. The primary outcome of this nested substudy will be 2-week postenrollment treatment failure.

Setting

The trial will sample PHCs from the Eastern Khayelitsha (6 PHCs), Tygerberg (6 PHCs), and Northern (6 PHCs) health subdistricts. Khayelitsha has an official population of 450,000 and is 90.5% Black African [21], although researchers and organizations working there estimate the population to be closer to 1 million [22]. Khayelitsha has a young population, with more than 40% of its residents aged younger than 19 years, residing in informal housing with high caregiver unemployment and limited access to running water. HIV and tuberculosis prevalence are high, with maternal HIV prevalence being 29.5%—the highest in the Western Cape Province— and the tuberculosis incidence of 1389 per 100,000 population exceeded the national average of 834 per 100,000 in 2017 [23]. The National Department of Health of South Africa uses IMCI guidelines for the management of children at PHCs.

PHC Eligibility and Recruitment

Eligible PHCs are defined as public facilities with no known planned closures during the study period, that provide health care to children aged 0 to less than 24 months, and are based in the Eastern Khayelitsha, Tygerberg, or Northern health subdistricts of the Western Cape. A list of eligible facilities will be obtained from the Western Cape Department of Health and City of Cape Town, and this will act as the sampling frame for

XSL•FO
RenderX

clusters. Through discussion with Western Cape Department of Health and the City of Cape Town officials, we purposively selected the 18 study PHCs. The intervention will be delivered at a facility level, with all HCWs of any cadre currently employed (either on a temporary or permanent basis) who provide care for children aged younger than 2 years being eligible for participation. Activities will occur through sequential site enrollment (see also Randomization and Blinding and Impact Evaluation Procedures sections).

Intervention and Control

We will be comparing the Phefumla (arm 1) and LB-01 (arm 2) devices to pulse oximeter devices currently used in routine care (arm 3 control). The intervention devices were both developed through a human-centered design approach with HCWs working in LMIC contexts and are not currently routinely used in the trial setting. The transmissive (ie, measuring SpO_2 by shining light *through* tissue to the opposite side of the body) LB-01 device is Conformité Européenne (CE) marked and commercially available [17]. The reflectance (ie, measuring SpO_2 by detecting light *reflected* from vessels below the skin on the same side of the body). Phefumla device is an investigational device that passed International Organization for Standardization (ISO) 80601-2-61:2017 regulatory testing for accuracy and safety and is eligible for CE marking and US Food and Drug Administration clearance but cannot be used for clinical decision-making during routine care [16]. The reference device (Rad-G; Masimo), a transmissive oximeter device, will be used across all 3 trial arms to ensure participant safety and comply with nonsignificant risk study standards. A summary of the technical specifications of the trial devices is provided in Table 1.

Table . Technical specifications of trial pulse oximeters.

Characteristic	Arm 1	Arm 2	Arm 3[a]	Reference
Device	Phefumla pulse oximeter	AH-MX LB-01 pulse oximeter	CONTEC CMS5100 patient monitor	Masimo Rad-G pulse oximeter
Oximeter type	Handheld	Handheld	Desktop	Handheld
Regulatory markings	ISO[b] 80601-2-61:2017 (passed), investigational device	ISO 80601-2-61:2017; CE[c] IIb classification	ISO 80601-2-61:2017; US FDA[d] clearance	ISO 80601-2-61:2017; US FDA clearance, CE IIb classification
Patient population	Neonate, pediatric, and adult	Neonate, pediatric, and adult	Neonate, pediatric, and adult	Neonate and pediatric
Recommended measurement location	Skin surface	Big toe, finger, or thumb	Big toe, index finger, or thumb	Big toe, index finger, or thumb

[a]If a control facility has a different device, the health care workers will continue to use their standard device. If a control facility has no functional oximeter device, the CONTEC pulse oximeter will be provided.

[b]ISO: International Organization for Standardization.

[c]CE: Conformité Européenne.

[d]US FDA: US Food and Drug Administration.

All facility staff who routinely provide care for children aged younger than 5 years, regardless of cadre and trial arm, will be trained in pulse oximetry. Control facilities will use the Contec Medical Systems (CONTEC) pulse oximeter (Table 1). If a control facility has a different device, the HCWs will continue to use their standard device. If a control facility has no functional oximeter device, the CONTEC pulse oximeter will be provided. All intervention components are summarized in Table 2.

Table . Summary of intervention components across the 3 study arms.

	Arm 1	Arm 2	Arm 3
Pulse oximeter provision[a]	Phefumla device	LB-01	Existing device in facility[a]
Baseline training and pilot by study physician	Days 1-2 or 3 of enrollment period	Days 1-2 or 3 of enrollment period	Days 1-2 or 3 of enrollment period
Intervention enrollment period	Days 2 or 3 of enrollment period until 4-6 wk enrollment with a target sample size of 66-67 children	Days 2 or 3 of enrollment period until 4-6 wk enrollment with a target sample size of 66-67 children	Days 2 or 3 of enrollment period until 4-6 wk enrollment with a target sample size of 66-67 children
Direct observation(s) by study physician	After at least 3 wk of enrollment	After at least 3 wk of enrollment	After at least 3 wk of enrollment

[a]If the clinic does not have a functional oximeter, they are provided with a CONTEC device.

Participants

Our primary target population is children aged 0 to less than 24 months. Eligibility criteria for trial recruitment include children aged 0 to less 24 months, routinely presenting to care with HCW-observed and/or caregiver-reported history of either cough and/or difficult breathing (ie, ARI), being a resident of the Western Cape, having a primary caregiver who agrees to provide contact details for follow-up and is able and willing to provide written informed consent. Children with any of the following

XSL•FO
RenderX

WHO-defined emergency signs will be excluded: obstructed or absent breathing, severe respiratory distress (ie, grunting), signs of uncompensated shock (capillary refill time >3 s, elevated heart rate with weak pulse, and low or unmeasurable systolic blood pressure), coma, active convulsions, or diarrhea with severe dehydration (altered mental status, sunken eyes, very slow return after skin pinch, or any 2 of these).

If there is no viable site for measurement (eg, the participant is missing digits or has severe burns), this will be classified as an "unsuccessful measurement." The inclusion criteria for ARI were chosen to reflect the WHO IMCI guidelines [5], as these signs are the entry point into an HCW conducting a full respiratory examination.

Randomization and Blinding

Clinics will be randomized in a 1:1:1 ratio to intervention arm 1 (Phefumla device: 6/18, 33%), intervention arm 2 (LB-01 device: 6/18, 33%), and control (6/18, 33%). The 18 study clinics will be equally distributed across the 3 subdistricts (Khayelitsha Eastern, Northern, and Tygerberg) and were randomized using stratified randomization to give 2 clinics in each arm in each subdistrict. Using the same random number, these facilities were then put into blocks, with 1 clinic from each arm in each block (ie, 6 blocks of 3 clinics each). Within these blocks, a new random number was generated to assign the clinic recruitment order. Randomization will be done using Stata SE version 17 (StataCorp) [24].

Owing to the nature of the intervention, the only group masked to trial arm allocation is the trial statistician (BAN), who will conduct the primary trial analysis. All investigators will also be masked during the analysis phase of the trial. The key listing of PHCs by trial arm will be held by the trial epidemiologist (CK) and shared after the primary analysis results have been presented to a study advisory group comprising health professionals and researchers responsible for providing independent advising to the project.

PHC staff, caregivers, participants, and study staff at participating facilities cannot be blinded to the type of pulse oximeter. Caregivers at the PHC will be informed that there is an ongoing study. HCWs and eligible participants will be provided with information, but they will not know the trial arm of the PHC. HCWs and facility managers will be provided with information on the study design, research questions, procedures, and randomization. However, these individuals will not be informed of the primary outcome definition, nor which device is the control device.

Impact Evaluation Procedures

Two study teams comprised 1 research nurse and 1 research counselor who will be supervised by 1 study physician and will conduct child and caregiver recruitment and follow-up. Study staff will be trained on the study protocol, data collection tools, ethics and informed consent, and SpO_2 measurements. This will be followed by observed mock interviews, with a formal assessment of SpO_2 measurement and interpretation skills. Study clinics will be enrolled sequentially, with participant recruitment occurring in 1 to 2 clinics simultaneously. At each clinic, participation will entail HCW training for the first 1 to 2 days of the enrollment period, followed by 1 to 2 days of piloting. At HCW trainings, study staff will review IMCI guidelines that include guidance on pulse oximetry use and demonstrate and train HCWs on the use of the assigned pulse oximeter devices. Study staff at control facilities will receive the same training but with a focus on the commercially available device available at the facility. Study staff will recruit 4 to 5 d/wk during clinic operating hours, for a 4 - to 6-week period, depending on patient volume.

Study staff will follow a standard operating procedure for recruitment. Children will be screened, identified as eligible, managed routinely by clinic staff, selected to participate, and then enrolled by study staff. Informed consent will be provided in the caregiver's preferred language of isiXhosa, English, or Afrikaans by trained study staff. After a child is triaged and completes their SpO_2 evaluation by the HCW, study staff will collect 2 SpO_2 measurements. The first SpO_2 measurement will be done using the reference device, followed by the study arm device, with both initiated within 10 minutes of the routine HCW assessment. Following the final clinical assessment by the registered nurse or attending HCW, additional sociodemographic, vaccination, and contact information will be collected, and a clinical examination will be performed by study staff. If the reference SpO_2 measurement clinically differs from the HCW's measurement and may impact onward care, then it will be shared with the HCW to support their clinical decision-making (eg, the reference device and HCW measurement differ by being either above or below the 90% threshold used for oxygen initiation and onward referral). If an SpO_2 reference measurement cannot be initiated within 10 minutes of the HCW measurement, the study data collector will identify this as a failed measurement, although all children will still have a reference SpO_2 measurement obtained. Participant flow through the study will depend on patient flow in the PHC, but in all cases, a valid study measurement must be initiated within 10 minutes of the HCW measurement. In control clinics that are not using the CONTEC device, study staff will still use the CONTEC device for reference measurements. To optimize both participant safety and the timing of study SpO_2 measurements, a delayed consent waiver is available for study staff to use in PHCs that more systematically obtain SpO_2 measures during patient triage. Finally, trial pulse oximeters will be removed after each clinic's implementation period, and clinics will be provided with a reference device for continued use for patient care.

Trial Outcomes

The primary trial outcome is "correct SpO_2 management" [15], a composite intermediate clinical end point between correct pulse oximeter use and hypoxemia outcomes, and is defined by the three elements necessary to reduce inappropriately treated hypoxemia: (1) device adoption—HCW use of the device as evidenced by an HCW-documented SpO_2 and heart rate measured in room air (ie, off supplemental oxygen); (2) quality SpO_2 measurement—a study staff–measured SpO_2 with the reference device that is within a 2% SpO_2 range above or below the HCW-documented SpO_2; and (3) correct SpO_2

XSL•FO
RenderX

decision-making—an appropriate referral recommendation provided by the HCW according to the WHO-defined hypoxemia definition (SpO_2 <90%). A ±2% SpO_2 range was selected, as it is comparable to the standard acceptable difference between the reference arterial blood gas oxyhemoglobin saturation and SpO_2 measurements used in pulse oximeter development and testing and is the performance range stipulated in pulse oximeter manufacturer product specifications. A ±2% SpO_2 range also applies to individuals with darker skin tones more common in LMIC settings [25] (see Table 3 for key terms). The secondary outcomes are correct SpO_2 management according to varied definitions, quality measurement frequency, referral acceptance, and oxygen treatment.

Table . Key terms.

Term	Definition
ARI[a]	Reported or observed cough or difficulty breathing.
Pneumonia	ARI and fast breathing for age, chest wall indrawing, a general danger sign, or hypoxemia [5].
WHO[b]-defined hypoxemia	SpO_2[c] <90% while breathing room air [5].
Study-defined hypoxemia	SpO_2 <94% while breathing room air [26].
Moderate hypoxemia	SpO_2 90%-93% while breathing room air.
Biologically plausible	SpO_2 measurements achieving a stable plethysmography waveform with stable SpO_2 (±1%) for 3-5 s and a heart rate between 1st and 99th centile for age [27].
Appropriate referral	Children with a SpO_2 <90% were counseled to go to the hospital, given a prereferral dose of antibiotic according to IMCI[d], and received supplemental oxygen (if available). Children with an SpO_2 >90% may be referred due to other clinical signs—this will not be considered inappropriate.
Difficult breathing	Any abnormal breathing pattern reported by caregiver or observed by a HCW[e] or study staff.
Fast breathing for age	Respiratory rate >60 breaths per minute if 0 to <2 mo of age, >50 breaths if 2 to <12 mo of age, or >40 breaths if 12 to <24 mo of age [5].
Chest wall indrawing	Bilateral inward pulling of the anterior chest wall subcostal tissue and lower ribs during inspiration [5].
General danger signs (2- to <24-mo-olds)	Vomiting everything, lethargy or unconscious, unable to drink or breastfeed, convulsions, or stridor at rest [5].
Neonatal danger signs (0- to <2-m-olds)	Unable to feed well, not moving at all or moves only when stimulated, severe chest indrawing, or grunting.
Unsuccessful SpO_2 measurement	Study staff measurements initiated within 10 min of the HCW measurement using the device assigned to arm and reference device both indicate performance is not within ±2%.

[a]ARI: acute respiratory infection.

[b]WHO: World Health Organization.

[c]SpO_2: capillary oxyhemoglobin saturation.

[d]IMCI: Integrated Management of Childhood Illnesses.

[e]HCW: health care worker.

Power and Sample Size

The trial is powered to detect a difference in the primary outcome of "correct SpO_2 management" in a pairwise analysis between each intervention arm and the control arm. We estimate that 55% of children within control clinics will meet the definition for correct SpO_2 management [28]. Using the following parameters, we will be able to detect a 22% increase in the primary outcome: 18 months recruitment, 6 clusters per arm, 60 children per cluster (targeting 67 and allowing for ~10% attrition, ie, child enrolled but post-HCW visit examination incomplete), 2.5% alpha after applying Bonferroni correction [29], greater than 80% power, intraclass correlation coefficient of 0.05, and a coefficient of variation of cluster size of 0.2. This will give us a sample of 360 children per arm (N=1080). We will aim to recruit an average of 3 to 4 children each day from each facility, yielding approximately 1200 children.

Statistical Analysis

The primary analysis will be intention-to-treat comparing correct SpO_2 management between each intervention arm and the control arm, using a mixed-effects logistic regression model. We will assess the variability in cluster sizes and whether these are associated with the outcome and, if so, will use independence

XSL•FO
RenderX

estimating equations with a logit link. As we anticipate a low level of missing data, our primary analysis will be a complete case analysis. The detailed statistical analysis plan is provided in Multimedia Appendix 1.

Mixed Method Process Evaluation

The mixed method process evaluation will be conducted throughout the trial enrollment period, with quantitative data collected from all facilities, and ethnographic observations conducted in a subsample of 6 facilities (2 per arm). The process evaluation will aim to explore the contexts that shape HCW willingness to use pulse oximetry, how facility structures and patient pathways influence pulse oximetry uptake, HCW perceptions and preference for different devices and their usability, and the association of pulse oximetry use with referral acceptance. Process indicators organized according to the RE-AIM framework [20] are summarized in Table 4. Information on the wider context in which the intervention is being run will be documented through study team meeting notes, including health system disruptions (eg, staff strikes) and political actions (eg, the launch of new policies and elections).

XSL•FO
RenderX

Table . Summary process evaluation indicators and data sources.

RE-AIM[a] domain	Definition	Indicator
Reach	The absolute number, proportion, and representativeness of children who are willing to participate in the study	• Consent rate (data source: recruitment log; perspective: caregivers) • Comparison of participants to nonparticipants (data source: recruitment log; perspective: caregivers) • Socioeconomic profile of participants (data source: enrollment survey; perspective: caregivers) • SpO_2[b] measurement acceptance (data source: enrollment survey; perspective: caregivers) • Referral acceptance (data source: enrollment survey; perspective: caregivers)
Adoption	The absolute number, proportion, and representativeness of facilities and HCWs[c] who are willing to participate in the study	• Characteristics of PHCs[d] (data source: baseline facility audit; perspective: PHCs) • Availability of functional pediatric pulse oximeters at baseline (data source: baseline facility audit; perspective: PHCs) • Characteristics of HCWs who participate (data source: baseline HCW survey; perspective: HCWs)
Implementation	At the facility level, fidelity to the intervention protocol, including consistency of delivery	• HCW knowledge of pulse oximetry (data source: baseline HCW survey; perspective: HCWs) • HCW proficiency in pulse oximetry (data source: HCW skills test; perspective: HCWs) • Time to successful SpO_2 measurement (data source: HCW skills test; perspective: HCWs) • Perception of HCWs about the pulse oximeter (data source: HCW FDG[e]; perspective: HCWs) • Perception of the HCWs about the mode of pulse oximeter implementation (data source: HCW FDG[e]; perspective: HCWs) • Adaptation to the pulse oximeter implementation approach (data source: baseline facility audit and observations; perspective: PHCs)
Maintenance	The extent to which the use of pulse oximetry becomes institutionalized, measured 6 mo after intervention completion	• HCW knowledge after 6 mo (data source: follow-up HCW survey; perspective: HCWs) • HCW proficiency after 6 mo (data source: Follow-up HCW skills test; perspective: HCWs) • Availability of functional pediatric pulse oximeters after 6 mo (data source: check-in logs; perspective: PHCs)

[a]RE-AIM: reach, effectiveness, adoption, implementation, and maintenance.

[b]SpO_2: peripheral blood oxygen saturation.

[c]HCW: health care worker.

[d]PHC: primary health care clinic.

[e]FDG: focus group discussion.

Each PHC will have a baseline facility audit conducted after the PHC has provided consent and before any other trial intervention activities take place. The audit will collect data on clinical services, infrastructure, and human resources and will include inspections of oxygen and pulse oximetry equipment. A subset of 6 purposefully selected PHC will have 1 to 2 days

XSL•FO
RenderX

of ethnographic observations, following a structured observation checklist, to explore pulse oximetry practice and patient pathways. For this subset of facilities, observations will be repeated during weeks 1 and 4 of trial recruitment and then again at the 6-month visit.

HCW surveys will be conducted with all HCWs before they take part in the baseline training to assess demographics and work history, knowledge of the IMCI guidelines and pulse oximetry, and training history. This survey will be repeated at the end of the clinic recruitment period (between 4 and 6 wk later), and again at the 6-month visit. HCWs will also be asked to do a pulse oximetry skills test at these 2 time points, during which they will conduct 2 measurements on different children, while being observed by study staff [17].

After the end of trial recruitment across all facilities, a discussion with HCWs from the study clinics will be held. We will invite 1 HCW from each of the study facilities to attend a meeting where we will present the key trial findings and facilitate a discussion of the results. We will aim for a mix of HCWs to ensure a range of trial perspectives and experiences.

Embedded Prospective Cohort

The embedded prospective cohort comprised the same children enrolled in the trial. The objectives of the prospective cohort analyses are to determine the outpatient burden of treatment failure and hypoxemia in PHCs and to describe risks for treatment failure, hypoxemia, and mortality among children presenting with ARIs. Treatment failure will be defined as the proportion of enrolled children who reported that they were still sick or whose caregiver reported any of the following signs: cough at 14 days, difficulty breathing at 14 days, change in antibiotic treatment, readmission to the clinic or hospital, or death any time at or before 14 days after study enrollment, as confirmed via phone or by home visit.

Ethical Considerations

Our protocol has been approved by the Human Research Ethics Committee of Stellenbosch University in South Africa (M23/08/024) and the Johns Hopkins School of Medicine Institutional Review Board (IRB00444460) for ethical research conduct and the protection of human subjects. Study nurses and coordinators will obtain informed consent in a private area of the clinic separate from the public waiting area. HCW and mother and child participation are voluntary, and participants will be informed of their ability to opt out. Participants will be compensated as required by the South African Health Products Regulatory Authority for time, inconvenience, and expense. Data collected will be manually entered into a secure, password-protected, encrypted tablet and will be downloaded daily with the REDcap (Research Electronic Data Capture) data management system hosted on a dedicated and secure server at Stellenbosch University. All audio recordings and transcriptions will be held on password-protected and encrypted study laptops, and access to data on REDcap will be restricted to only key personnel. Linking files for personal identifiers will only be accessible to the study principal investigator and key personnel. We will disseminate findings at international research conferences and in peer-reviewed journals.

Results

The study was funded in 2023, with data collection commencing in 2024. As of December 2025, we have completed enrollment in 12 clinics. Trial results are expected in 2026 and will be disseminated through academic publications, conferences, and a stakeholder meeting with HCWs from the trial PHCs.

Discussion

Anticipated Findings

In this trial, we will evaluate the correct SpO_2 management of two pediatric pulse oximeters: the CE-marked LB-01 device and the first prototype of the investigational Phefumla device, both of which have undergone extensive development and validation using an end user–informed, participatory, human-centered oximeter design process with HCWs in Malawi, Bangladesh, the United Kingdom, and South Africa [16]. Both devices were designed to address common barriers that limit the use of pulse oximetry for young children in low-resource settings, including lack of devices designed for pediatric patients, challenges with children moving during measurement, prohibitive device cost, unavailability of devices at facilities, and lack of health care provider awareness, training, and supervision [10,15,16,30-34]. During the development phase of both devices, we learned that important features of durability, battery life, and usability with children are essential to providers in low-resource contexts. We intend to compare correct SpO_2 management—a composite intermediate clinical end point between correct pulse oximeter use and hypoxemia outcome—between these devices and other standard-of-care devices on the market to understand whether their design features favorably support effective implementation in real-world settings and what improvements can be incorporated into future versions to address any identified performance issues.

Strengths and Limitations

Our study design and implementation have challenges. First, there is difficulty in not disrupting clinical care while still generating meaningful comparison measurements to evaluate real-world clinical decision-making. To lessen any disruption, we include a delayed consent option to permit study staff to obtain SpO_2 measures immediately before HCWs in clinics that more rapidly do pulse oximeter testing during patient triage at the PHC. Second, the Phefumla device cannot be used for clinical decision-making in this study as it is an investigational device. This project did not have the implementation time frame and resources to seek the regulatory approvals required to conduct a significant risk study. To address this, we pair reference and HCW device measurements to augment participant safety. Third, to address any bias on our primary outcome of "correct SpO_2 management" that may be introduced by the presence of study staff at clinics (Hawthorne effect), we conduct study device measurements in a separate setting to avoid observation of our staff measurements. Fourth, as SpO_2 is not a static measurement, there is the possibility of measurement bias from sequential HCW and study staff SpO_2 measurements. To mitigate this potential issue, we set a 10-minute time limit

XSL•FO
RenderX

for when the study measurements must be initiated following HCW SpO_2 measurement. While preparatory work informed our decision to include the CONTEC control device, we cater our training to allow control clinics to continue to use whatever device is available at the PHC, even if different. We prioritized pragmatism and the implementation focus of our study and expect minimal effect on device measurement differences. Collectively, these necessary but challenging aspects of our protocol require patient flow maps specific to each participating clinic to maximize opportunities for recruitment, consenting, and study measurement procedures without disrupting patient services. We expect that findings from our study will be generalizable to low-resource settings globally as well as areas of high disease burden.

In sum, although the importance of SpO_2 monitoring is well accepted, with calls to make SpO_2 a vital sign [35], translating this into practice has proven challenging. Research gaps remain around optimizing implementation and sustaining best practice, with 7 of the top 20 pediatric pneumonia research priorities [36] centered around pulse oximetry and oxygen adoption and health system impacts. This trial will fill some of these gaps in evidence.

Acknowledgments

The authors are grateful to Dr. Anneke Hesseling, Dr. Natasha O'Connell, and Dr. Andrew Redfern for their assistance in designing the study for the South African context. The authors appreciate the contributions of Mr. Mike Bernstein in the development of the Phefumla prototype device.

Funding

This work was supported by the Fogarty International Center of the National Institutes of Health (grant 1R21TW012212-01) and Thrasher Research Fund (award number 01724). MVdZ is supported by a career development grant from the European Union EDCTP2 program (TMA2019SFP-2836 tuberculosis lung-FACT2) and the Fogarty International Center of the National Institutes of Health (grant K43TW011028). The funders were not involved in the study design or the writing of the manuscript, and the content is solely the responsibility of the 22 authors and does not necessarily represent the official views of the National Institutes of Health or Thrasher Research Fund.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: CK, EDM
Methodology: HBS, BASN, CK, EDM
Data collection: LML, AM, EDM
Writing – original draft preparation: HBS, CK, EDM
Writing – review & editing: HBS, MVdZ, MVN, LML, AM, SG, SK, SH, LV, AG, BASN, PG, ACH, CK, EDM
Supervision: MVdZ, MVN, LML, ACH, CK, EDM
Project administration: MVdZ, MVN, CK, EDM
Funding acquisition: EDM
All authors have read and agreed to the published version of the manuscript

Conflicts of Interest

None declared.

Multimedia Appendix 1
Statistical analysis plan.
[DOCX File, 180 KB - resprot_v15i1e82888_app1.docx]

References

1. Troeger CE, Khalil IA, Blacker BF. Quantifying risks and interventions that have affected the burden of lower respiratory infections among children younger than 5 years: an analysis for the Global Burden of Disease Study 2017. Lancet Infect Dis 2020 Jan;20(1):60-79. [doi: 10.1016/S1473-3099(19)30410-4] [Medline: 31678026]
2. Villavicencio F, Perin J, Eilerts-Spinelli H, et al. Global, regional, and national causes of death in children and adolescents younger than 20 years: an open data portal with estimates for 2000-21. Lancet Glob Health 2024 Jan;12(1):e16-e17. [doi: 10.1016/S2214-109X(23)00496-5] [Medline: 37898143]

XSL•FO
RenderX

3. GBD 2021 Causes of Death Collaborators. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):2100-2132. [doi: 10.1016/S0140-6736(24)00367-2] [Medline: 38582094]
4. GBD 2021 Lower Respiratory Infections and Antimicrobial Resistance Collaborators. Global, regional, and national incidence and mortality burden of non-COVID-19 lower respiratory infections and aetiologies, 1990-2021: a systematic analysis from the Global Burden of Disease Study 2021. Lancet Infect Dis 2024 Sep;24(9):974-1002. [doi: 10.1016/S1473-3099(24)00176-2] [Medline: 38636536]
5. Integrated management of childhood illness: chart booklet. : World Health Organization; 2014 URL: https://cdn.who.int/media/docs/default-source/mca-documents/9789241506823_chartbook_eng.pdf?sfvrsn=cccea3a6_1&download=true [accessed 2026-01-28]
6. Graham HR, King C, Rahman AE, et al. Reducing global inequities in medical oxygen access: the Lancet Global Health Commission on medical oxygen security. Lancet Glob Health 2025 Mar;13(3):e528-e584. [doi: 10.1016/S2214-109X(24)00496-0] [Medline: 39978385]
7. Floyd J, Wu L, Hay Burgess D, Izadnegahdar R, Mukanga D, Ghani AC. Evaluating the impact of pulse oximetry on childhood pneumonia mortality in resource-poor settings. Nature New Biol 2015 Dec 3;528(7580):S53-S59. [doi: 10.1038/nature16043] [Medline: 26633766]
8. Rahman AE, Hossain AT, Nair H, et al. Prevalence of hypoxaemia in children with pneumonia in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Glob Health 2022 Mar;10(3):e348-e359. [doi: 10.1016/S2214-109X(21)00586-6] [Medline: 35180418]
9. Lazzerini M, Sonego M, Pellegrin MC. Hypoxaemia as a mortality risk factor in acute lower respiratory infections in children in low and middle-income countries: systematic review and meta-analysis. PLoS One 2015;10(9):e0136166. [doi: 10.1371/journal.pone.0136166] [Medline: 26372640]
10. McCollum ED, Ahmed S, Roy AD, et al. Risk and accuracy of outpatient-identified hypoxaemia for death among suspected child pneumonia cases in rural Bangladesh: a multifacility prospective cohort study. Lancet Respir Med 2023 Sep;11(9):769-781. [doi: 10.1016/S2213-2600(23)00098-X] [Medline: 37037207]
11. McCollum ED, King C, Deula R, et al. Pulse oximetry for children with pneumonia treated as outpatients in rural Malawi. Bull World Health Organ 2016 Dec 1;94(12):893-902. [doi: 10.2471/BLT.16.173401] [Medline: 27994282]
12. Fox MP, Thea DM, Sadruddin S, et al. Low rates of treatment failure in children aged 2-59 months treated for severe pneumonia: a multisite pooled analysis. Clin Infect Dis 2013 Apr;56(7):978-987. [doi: 10.1093/cid/cis1201] [Medline: 23264361]
13. Bhutta ZA. Dealing with childhood pneumonia in developing countries: how can we make a difference? Arch Dis Child 2007 Apr;92(4):286-288. [doi: 10.1136/adc.2006.111849] [Medline: 17376934]
14. McCollum ED, King C, Hollowell R, et al. Predictors of treatment failure for non-severe childhood pneumonia in developing countries--systematic literature review and expert survey--the first step towards a community focused mHealth risk-assessment tool? BMC Pediatr 2015 Jul 9;15:74. [doi: 10.1186/s12887-015-0392-x] [Medline: 26156710]
15. King C, Boyd N, Walker I, et al. Opportunities and barriers in paediatric pulse oximetry for pneumonia in low-resource clinical settings: a qualitative evaluation from Malawi and Bangladesh. BMJ Open 2018 Jan 30;8(1):e019177. [doi: 10.1136/bmjopen-2017-019177] [Medline: 29382679]
16. Ilhan EI, Jola LN, van der Zalm MM, et al. Designing a smartphone-based pulse oximeter for children in South Africa (Phefumla Project): qualitative analysis of human-centered design workshops with health care workers. JMIR Hum Factors 2024 May 30;11:e54983. [doi: 10.2196/54983] [Medline: 38825834]
17. Boyd N, King C, Walker IA, et al. Usability testing of a reusable pulse oximeter probe developed for health-care workers caring for children < 5 years old in low-resource settings. Am J Trop Med Hyg 2018 Oct;99(4):1096-1104. [doi: 10.4269/ajtmh.18-0016] [Medline: 30141389]
18. Lee H, Ko H, Lee J. Reflectance pulse oximetry: practical issues and limitations. ICT Express 2016 Dec;2(4):195-198. [doi: 10.1016/j.icte.2016.10.004]
19. Landes SJ, McBain SA, Curran GM. An introduction to effectiveness-implementation hybrid designs. Psychiatry Res 2019 Oct;280:112513. [doi: 10.1016/j.psychres.2019.112513] [Medline: 31434011]
20. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
21. Mumm R, Diaz-Monsalve S, Hänselmann E, et al. Exploring urban health in Cape Town, South Africa: an interdisciplinary analysis of secondary data. Pathog Glob Health 2017 Feb;111(1):7-22. [doi: 10.1080/20477724.2016.1275463] [Medline: 28093045]
22. Living in Khayelitsha in South Africa. : Peace centre; 2019 URL: https://www.peacecentre.org.za/wp-content/uploads/2019/08/Living-in-Khayelitsha-in-South-Africa.pdf [accessed 2026-01-28]
23. Richards D, Hunter L, Forey K, Myers C, Christensen E, Cain S, et al. Demographics and predictors of mortality in children undergoing resuscitation at Khayelitsha Hospital, Western Cape, South Africa. S Afr J Child Health 2018;12(3):127-131. [doi: 10.7196/sajch.2018.v12i3.1604]
24. Stata: Release 17. Statistical Software. StataCorp. 2021. URL: https://www.stata.com/manuals17/i.pdf [accessed 2025-09-01]

XSL•FO
RenderX

25. Shi C, Goodall M, Dumville J, et al. The accuracy of pulse oximetry in measuring oxygen saturation by levels of skin pigmentation: a systematic review and meta-analysis. BMC Med 2022 Aug 16;20(1):267. [doi: 10.1186/s12916-022-02452-8] [Medline: 35971142]
26. McCollum ED, King C, Ahmed S, et al. Defining hypoxaemia from pulse oximeter measurements of oxygen saturation in well children at low altitude in Bangladesh: an observational study. BMJ Open Respir Res 2021 Nov;8(1):e001023. [doi: 10.1136/bmjresp-2021-001023] [Medline: 34728475]
27. King C, Mvalo T, Sessions K, et al. Performance of a novel reusable pediatric pulse oximeter probe. Pediatr Pulmonol 2019 Jul;54(7):1052-1059. [doi: 10.1002/ppul.24295] [Medline: 30912314]
28. Bjornstad E, Preidis GA, Lufesi N, et al. Determining the quality of IMCI pneumonia care in Malawian children. Paediatr Int Child Health 2014 Feb;34(1):29-36. [doi: 10.1179/2046905513Y.0000000070] [Medline: 24091151]
29. Vickerstaff V, Omar RZ, Ambler G. Methods to adjust for multiple comparisons in the analysis and sample size calculation of randomised controlled trials with multiple primary outcomes. BMC Med Res Methodol 2019 Jun 21;19(1):129. [doi: 10.1186/s12874-019-0754-4] [Medline: 31226934]
30. Enoch AJ, English M, Clinical Information Network, McGivern G, Shepperd S. Variability in the use of pulse oximeters with children in Kenyan hospitals: a mixed-methods analysis. PLoS Med 2019 Dec;16(12):e1002987. [doi: 10.1371/journal.pmed.1002987] [Medline: 31891572]
31. Graham HR, Bakare AA, Gray A, Ayede AI, Qazi S, McPake B. Correction: Adoption of paediatric and neonatal pulse oximetry by 12 hospitals in Nigeria:a mixed-methods realist evaluation. BMJ Glob Health 2018 Dec;3(6):e000812corr1. [doi: 10.1136/bmjgh-2018-000812corr1]
32. Sheikh M, Ahmad H, Ibrahim R, Nisar I, Jehan F. Pulse oximetry: why oxygen saturation is still not a part of standard pediatric guidelines in low-and-middle-income countries (LMICs). Pneumonia (Nathan) 2023 Feb 5;15(1):3. [doi: 10.1186/s41479-023-00108-6] [Medline: 36739442]
33. Zhang J, Tüshaus L, Nuño Martínez N, et al. Data integrity-based methodology and checklist for identifying implementation risks of physiological sensing in mobile health projects: quantitative and qualitative analysis. JMIR mHealth uHealth 2018 Dec 14;6(12):e11896. [doi: 10.2196/11896] [Medline: 30552079]
34. Sarin E, Kumar A, Alwadhi V, Saboth P, Kumar H. Experiences with use of a pulse oximeter multimodal device in outpatient management of children with acute respiratory infection during Covid pandemic. J Family Med Prim Care 2021 Feb;10(2):631-635. [doi: 10.4103/jfmpc.jfmpc_1410_20] [Medline: 34041052]
35. Zar HJ, McCollum ED. Pulse oximetry to detect paediatric hypoxaemia-the fifth vital sign. Lancet Glob Health 2023 Nov;11(11):e1684-e1685. [doi: 10.1016/S2214-109X(23)00380-7] [Medline: 37793415]
36. King C, Baker K, Richardson S, et al. Paediatric pneumonia research priorities in the context of COVID-19: an eDelphi study. J Glob Health 2022;12:09001. [doi: 10.7189/jogh.12.09001] [Medline: 35265333]

Abbreviations

ARI: acute respiratory infection
CE: Conformité Européenne
CONTEC: Contec Medical Systems
HCW: health care worker
IMCI: Integrated Management of Childhood Illnesses
ISO: International Organization for Standardization
LMIC: low- and middle-income country
PHC: primary health care clinic
REDCap: Research Electronic Data Capture
SpO2: oxyhemoglobin saturation
WHO: World Health Organization

Edited by A Mavragani; submitted 24.Aug.2025; peer-reviewed by H Ruppel, H Langet; revised version received 11.Jan.2026; accepted 19.Jan.2026; published 30.Mar.2026.

Please cite as:
Schuh HB, Van der Zalm M, Van Niekerk M, Laubscher LM, Mcinziba A, Gomas S, Kapoor S, Hooli S, Viljoen L, Gie A, Nonyane BAS, Goussard P, Hesseling AC, King C, McCollum ED
Clinical Evaluation of Pediatric Pulse Oximeters in South Africa: Protocol for a Cluster Randomized Controlled Trial
JMIR Res Protoc 2026;15:e82888
URL: https://www.researchprotocols.org/2026/1/e82888
doi:10.2196/82888

XSL•FO
RenderX

© Holly B Schuh, Marieke Van der Zalm, Margaret Van Niekerk, Lise-Marie Laubscher, Abenathi Mcinziba, Steve Gomas, Sunaina Kapoor, Shubhada Hooli, Lario Viljoen, Andre Gie, Bareng A S Nonyane, Pierre Goussard, Anneke C Hesseling, Carina King, Eric D McCollum. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact Assessment of the Allergy Fact Checker, a Clinical Decision Support Tool for Noninvasive Beta-Lactam Antibiotic Allergy Label Delabeling: Protocol for a Multicenter Crossover Cluster-Controlled Study

Liesbeth Gilissen[1], PhD; Greet Van De Sijpe[2], PharmD; Isabel Spriet[2*], PhD; Rik Schrijvers[1*], MD, PhD

[1]Allergy and Clinical Immunology Research Group, Department of Microbiology, Immunology and Transplantation, KU Leuven, Herestraat 49, Leuven, Flanders, Belgium

[2]Department of Pharmaceutical and Pharmacological Sciences, Clinical Pharmacology and Pharmacotherapy, KU Leuven, Leuven, Belgium

[*]these authors contributed equally

Corresponding Author:
Rik Schrijvers, MD, PhD
Allergy and Clinical Immunology Research Group, Department of Microbiology, Immunology and Transplantation, KU Leuven, Herestraat 49, Leuven, Flanders, Belgium

Abstract

Background: Beta-lactam allergy labels (BLALs), especially penicillin allergy labels, are frequently recorded in hospitalized patients and are associated with increased use of broad-spectrum and second-line antibiotics. Most BLALs are incorrect, but current allergy workups require invasive testing and specialized resources. We recently developed a strictly noninvasive, electronic patient record–embedded clinical decision support tool, the Allergy Fact Checker (AFC), which proactively identifies potentially incorrect BLALs by detecting uneventful re-exposures to the culprit or other beta-lactams since introduction of the BLAL.

Objective: This study aims to evaluate the clinical, antimicrobial, and economic impact of the AFC in hospitalized adults with BLALs compared to the standard of care (no AFC).

Methods: We are conducting a multicenter, open-label, crossover cluster-controlled study in 9 hospitals in Flanders, Belgium. All hospitalized adults with a BLAL are eligible, excluding patients in palliative care, discharged within 24 hours, or previously enrolled. Each hospital will alternate between intervention (use of the AFC) and control (standard practice) phases, separated by washout periods. The primary endpoint is cumulative guideline-concordant prescribing of first-line and/or narrow-spectrum beta-lactams following local antibiotic treatment guidelines, across predefined care windows up to day 100, expressed as a weighted per-patient proportion. Secondary endpoints include delabeling or refinement rate, beta-lactam tolerance, antibiotic switching, hospital length of stay, in-hospital and 3-month mortality, intensive care unit admission, readmission, multidrug-resistant organism colonization or infection, and costs. Based on prior data, we calculated that a total of 3285 participants are required to achieve 80% power to detect superiority of the intervention (4.6% vs 9.9% appropriate prescribing). Recruitment started in March 2025 and is ongoing.

Results: The primary outcome will be analyzed using hierarchical mixed-effect models accounting for hospital-level clustering and period effects. Data collection will continue until 200 days after the last patient is discharged. This trial is expected to conclude in 2026. Ethical approval was obtained from the institutional review board in November 2024. Recruitment started in March 2025 and is ongoing. As of manuscript submission, more than 3000 participants have been enrolled across the participating hospitals. According to the prespecified protocol, an interim assessment of nuisance parameters will be performed in the following month to evaluate the initial design assumptions. Data collection is expected to continue until the end of 2026. The main study results are anticipated to be reported in 2027.

Conclusions: This is the first multicenter European study evaluating a strictly noninvasive BLAL delabeling approach. If successful, this AFC tool could improve antimicrobial stewardship, reduce costs, and provide a scalable model for centralized allergy label management.

Trial Registration: ClinicalTrials.gov NCT06771440; https://clinicaltrials.gov/study/NCT06771440

International Registered Report Identifier (IRRID): DERR1-10.2196/86056

(JMIR Res Protoc 2026;15:e86056) doi:10.2196/86056

XSL•FO
RenderX

KEYWORDS

antibiotic allergy label; antimicrobial stewardship; beta-lactams; clinical decision support; delabeling; drug allergy; electronic patient records; noninvasive delabeling; medical records; penicillin

Introduction

Drug allergy labels are an important tool to improve patient safety. They aim to reduce potential harmful re-exposure by alerting health care providers (HCP) of a prior allergic reaction. However, labeling is often insufficiently accurate, and an incorrect label may even cause harm through unnecessarily restraining patients from a first-choice treatment. Antibiotic allergy labels are the most frequently reported drug allergy labels [1]. Especially the prevalence of beta-lactams, which are first-choice treatments in most bacterial infections because of their spectrum, safety, low cost, and availability, is high. In US studies, the prevalence of beta-lactam allergy labels (BLALs) reported in electronic patient records (EPR) ranges between 9% and 16% [2-6], whereas European studies report a significantly lower prevalence of 1% to 5% [7-10]. In our center, covering over 1 million patients over an 8-year period, the observed BLAL prevalence was 2%, with rates varying from 1% in outpatient settings to 6% among hospitalized patients [11].

Most BLALs are incorrect. Upon a complete drug allergy workup, which includes a thorough clinical history, invasive skin and drug provocation testing, the majority of BLAL, especially penicillin, can be invalidated ranging from 95% - 99% of patients with a penicillin allergy label in United States and Australian studies [1-9], to 77% - 91% in European Union (EU) studies [10,11].

The impact of BLAL, irrespective of the correctness of the label, has been studied primarily in non-EU settings and mostly for penicillin. Penicillin allergy labels were associated with increased use of second-line and broad-spectrum antibiotics, a delay in time-to-first-dose upon admission [1,6,12], and increased odds for infection with antibiotic-resistant organisms [13-15]. Moreover, an increased length of hospital stay (LOS), mortality, intensive care unit (ICU) admission rate [5,16,17], and a higher economic cost [18-20] were found. Recent work from our group suggested that BLALs have limited impact on individual clinical outcomes (ie, no increased mortality, ICU admission rate, or increased LOS), although their effect on antibiotic use remains in line with findings from non-EU studies [21].

These concerns led to the emergence of invasive delabeling protocols for penicillin allergy labels in the United States and Australia [22,23]. Current approaches use risk stratification based on clinical history and direct challenges (ie, in-hospital, but without prior skin tests) with penicillins or alternative beta-lactams in identified low-risk patients [24,25]. The efforts demonstrate to be cost-effective [22] and impact antibiotic use, shifting more to first-line beta-lactams and narrow-spectrum penicillins [26]. However, a relevant fraction of readministrations was accompanied by a repeat hypersensitivity reaction (3.8%), of which a minority (<1%) were severe [26]. This illustrates the delicate balance between indirect harm through incorrect BLAL and repeat iatrogenicity through invasive delabeling protocols themselves.

An often forgotten step in the delabeling process that encompasses the lowest risk for repeat iatrogeny is noninvasive delabeling. Here, without skin or drug provocation testing, but solely by using patient questionnaires and information from tolerated antibiotic exposures, discordances in the EPR are exploited to the advantage of the patient [9,27-29]. Our group prospectively evaluated in an EU context a noninvasive delabeling protocol in adult internal medicine inpatients with a BLAL [30], combining (1) an allergy-oriented clinical history, (2) an EPR search to identify incorrect BLAL and/or uneventful re-exposures to beta-lactams, and (3) a systematic primary care physician and community pharmacist contact. Using this approach, 38% of BLALs could be delabeled, and another 27% could be refined, largely driven by EPR findings. In the next step, we optimized the EPR search by developing a semiautomated clinical decision support (CDS) tool, designed to identify those patients with documented re-exposure without allergic reaction. A proof-of-concept study, limited to penicillin allergy labels, demonstrated the feasibility and efficacy of this novel CDS tool, identifying nearly 10% of alleged penicillin-allergic patients and showing approximately sevenfold higher odds of delabeling after its introduction [31]. Importantly, the impact on prescription behavior and major clinical outcome parameters is unknown and would be the basis for broad roll-out. Therefore, we propose to perform the first multicenter study on noninvasive BLAL delabeling, using an EPR-embedded, semiautomatic, intelligent “Allergy Fact Checker (AFC)” CDS tool.

Methods

Aim

This study aims to evaluate the clinical and economic benefits of a noninvasive CDS tool, critically assessing BLALs the AFC, while following local antibiotic treatment guidelines. Specifically, we will assess its impact on antibiotic use, as well as clinical, antimicrobial, and economic end points in hospitalized adults with BLAL. We hypothesize that a relevant fraction of BLAL can be rationalized by the AFC, thereby improving antibiotic use and costs without impacting morbidity or mortality.

Design

A multicenter, open-label, crossover cluster-controlled trial will be conducted to compare the use of our AFC on BLAL (intervention, I) and subsequent antibiotic prescription with current practice (control, C, ie, no AFC). All hospitalized adults with one or more BLAL will be included, except those previously enrolled in this study, discharged within 24 hours, or in palliative care. The primary endpoint is cumulative guideline-concordant prescribing of first-line and/or narrow-spectrum beta-lactams (based on the IGGI

XSL•FO
RenderX

[Infectiologiegids/Guide d'Infectiologie] guidelines of the Belgian Society for Infectiology and Clinical Microbiology) across predefined care windows up to day 100, expressed as a weighted per-patient proportion. Within each predefined assessment window, prescribing is scored in an any or ever binary manner, and a cumulative weighted score per patient is calculated across applicable windows. Secondary endpoints include proportion of delabeled and refined BLALs; beta-lactam antibiotic tolerance; need to switch antibiotics; in-hospital and 3-month posthospitalization mortality; LOS; 3-month readmission rate; ICU admission rate; colonization and/or infection with methicillin-resistant *Staphylococcus aureus*, vancomycin-resistant *Enterococcus*, *Clostridium difficile* (*C. diff*); antibiotic and hospitalization cost, time, and resources needed for the AFC; and physician appraisal of the approach.

We chose a mono-faceted intervention without an awareness campaign or educational component (eg, e-learning for physicians on BLAL labeling and delabeling) to evaluate the impact of the AFC itself, enhancing its extrapolation afterward. To correct for interhospital variability, we choose a clustered crossover design. The crossover design also accounts for seasonal variability, for example, in overall antibiotic use.

Allocation is performed on the hospital level (participating sites) before the start of this study to prevent contamination of effect between the intervention and control phases (ie, so that physicians will not receive information through the AFC while having to use conventional means for different patients at the same time). The allocation is nonrandomized and operationally predefined based on the availability of a single trained research team responsible for on-site implementation across hospitals. Inclusion periods are sufficiently long and distributed over calendar time to account for seasonal variation. A crossover cluster-controlled design was used to reduce potential confounding related to temporal and organizational factors in clinical practice, such as variation in staffing and workload among treating physicians. Residual temporal and cluster-level confounding will be further addressed in the statistical analysis by including study period and hospital as model terms.

During separate intervention phases, hospital physicians will receive the intervention, that is, BLAL CDS. Each of the crossover periods will be 3 weeks long. However, a washout period of 3 weeks is also taken into account to allow the patient population to be sufficiently "new" before the cluster transitions to the next condition. We will use the sequences I-x-C-x, x-I-x-C, C-x-I-x, and x-C-x-I (where I=intervention, ie, use of the AFC, and C=control, ie, standard of care [SOC], no use of the AFC, and x=washout period), and each cluster will be allocated to one of these sequences (Figure 1). Information about the assigned sequence is available to all people and groups involved in the research (ie, unblinded, open study).

Figure 1. Schematic representation of this study's design.

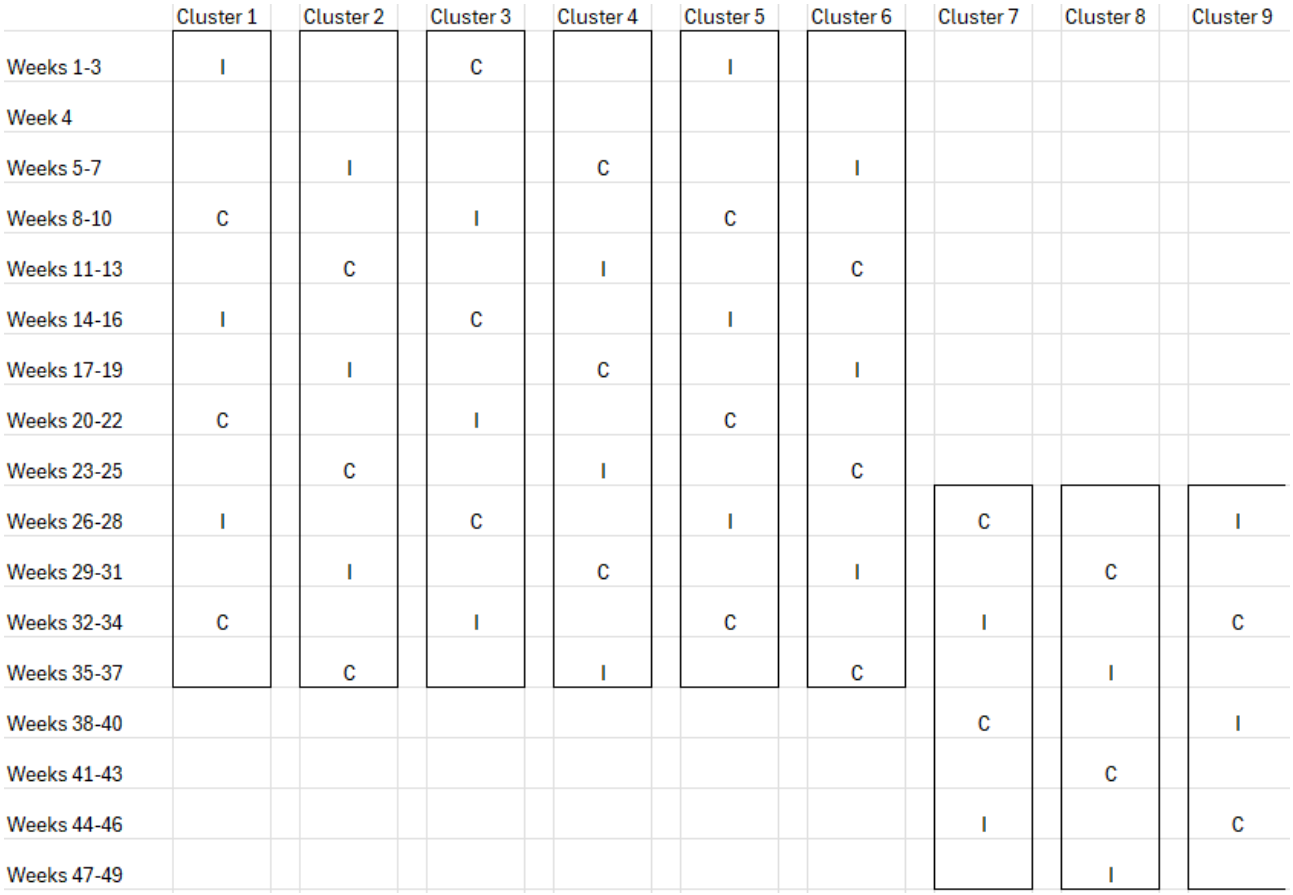

XSL•FO
RenderX

Setting

Patients will be recruited in 9 participating centers across Flanders, that is, the University Hospitals of Leuven (UZ [Universitair Ziekenhuis] Leuven), and the general hospitals of Pelt (Noorderhart Mariaziekenhuis), Mol (Heilig Hartziekenhuis), Herentals (AZ [Algemeen Ziekenhuis (general hospital)] Herentals), Turnhout (AZ Turnhout), Izegem (Sint-Jozefskliniek), Halle (AZ Sint-Maria Halle), Tienen (RZ Heilig Hart Tienen), and Malle (AZ Voorkempen). The expected recruitment period, from First Patient In to Last Patient In, is 11 months, starting March 10, 2025. The total inclusion period from First Patient In to Last Patient Out is expected to take approximately 18 months. The total period of intervention plus follow-up for every single patient depends on the LOS, that is, until 200 days after discharge from the hospital, to collect the secondary end point parameters: in-hospital mortality, 3-month posthospitalization mortality, LOS, and 3-month readmission rate. Patients who are still in the hospital when the end of the recruitment period is reached (estimated February 2026, when the needed sample size was reached) will be excluded. For comparison, in a recent retrospective study in UZ Leuven, involving 21,999 patients with pneumonia, pyelonephritis, infections associated with heart, kidney, liver, or lung transplantation, appendectomy, coronary artery bypass grafting, and total knee or hip replacement, the average LOS in UZ Leuven for patients with BLAL was 21 days (SD 39) [21].

Study Procedures

Study procedures and their timing are summarized in this study's flowchart (Table S1 in Multimedia Appendix 1). Figure 2 shows a schematic representation of the different steps of the intervention vs SOC. All study-specific procedures will be carried out by the team of the coordinating investigator at the different participating sites to minimize time consumption from principal investigators. This allergy-trained central study team consists of an allergy specialist (RS), an internal medicine specialist in training for allergy and clinical immunology and dedicated to this study (Dries Wets, MD), a biomedical scientist acting as a data manager (Nghia Nguyen, MSc), and a pharmacist-PhD trained in allergy (LG). This study's team (LG, RS, Dries Wets and Nghia Nguyen) evaluates the cases that are flagged and ensures communication with the treating physician (Figure S1 in Multimedia Appendix 1).

Figure 2. Schematic representation of the intervention vs standard of care. (1) Report baseline measurements eCRF; (2) automated search in the hospital EPR for re-exposure; (3) study team members receive AFC results on working list; (4) inclusion of patient in either standard of care or intervention (depending on scheme in Figure 1); (5) EPR check to identify BLAL for delabeling, refinement, or no action; (6) action is communicated to treating hospital physician (structured notes); (7) and (8) treating hospital physician can perform bedside verification and/or contact primary care HCP; (9) formal report of the results in the EPR; (10) letter to the primary care HCP; and (11) outcome parameters are followed up and scored in eCRF. AFC: Allergy Fact Checker; BLAL: beta-lactam allergy label; eCRF: electronic case report form; EPR: electronic patient record; GP: general practitioner; HCP: health care provider; REDCap: Research Electronic Data Capture; Vanderbilt University.

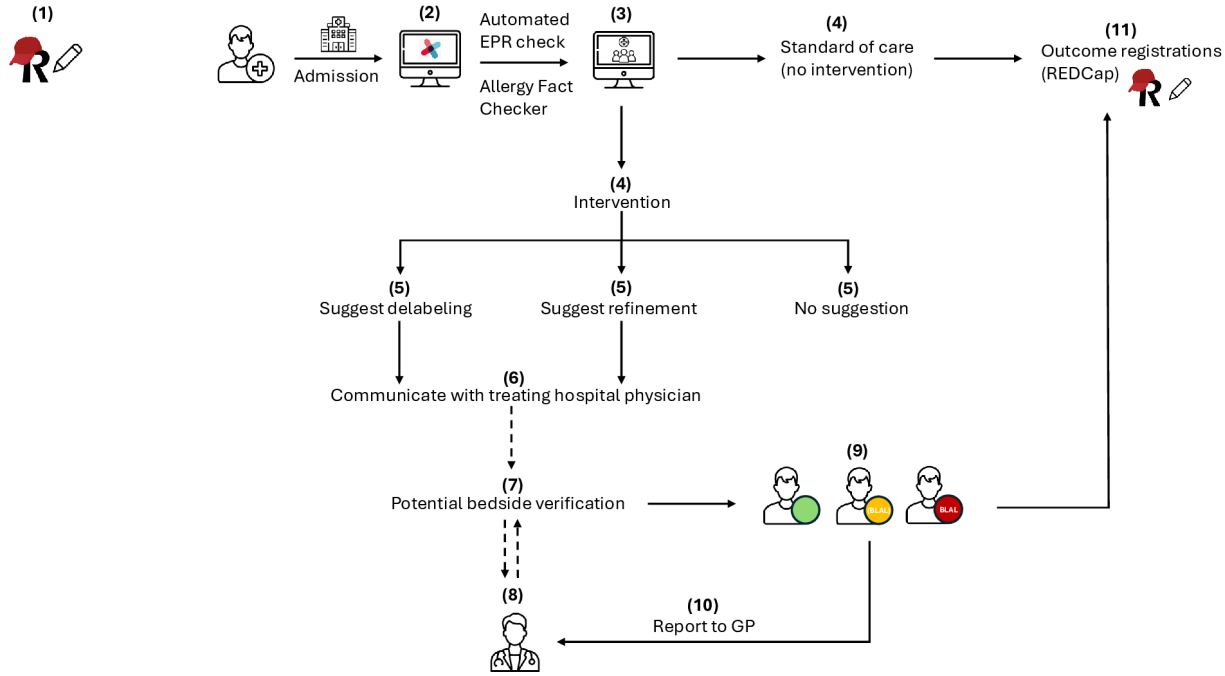

Patient Inclusion

Potential participants will be identified and recruited during hospitalizations in one of the participating sites. An EPR-embedded clinical rule identifies in near-real time all hospitalized patients with a BLAL in the allergy module. After checking inclusion and exclusion criteria, all eligible patients will be included in this study.

Ethical Considerations

This clinical study was initiated after evaluation and approval by the Research Ethics Committee UZ/KU Leuven (S68439, protocol version 5, July 2, 2025), and after consultation with

XSL•FO
RenderX

the ethics committees (EC) of each participating center in Belgium (Commissie voor Medische Ethiek [CME] Noorderhart [S504409]; EC St-Jozefskliniek Izegem [S504492]; CME Heilig Hart Ziekenhuis Mol [S504443]; CME AZ Herentals [S504441]; CME AZ Turnhout [S504448]; CME AZ Sint-Maria, CME RZ Tienen and CME AZ Voorkempen [approval granted; S-numbers pending]).

This study was classified as a noninterventional quality improvement project. The project does not introduce a new clinical treatment or experimental procedure but consists of systematically identifying discordances in existing electronic health records (ie, documented BLALs despite proven prior tolerance) and communicating this information to the treating physician through a decision-support process. The final clinical decision always remains with the treating physician and follows existing international and national guidelines. No additional diagnostic or monitoring procedures are imposed, and no automatic delabeling is performed by this study's team. Therefore, this study does not mandate changes in patient management but supports standard-of-care decision-making. Under Article 2, §8 of the Belgian Law of 7 May 2004 [31], this project qualifies as a noninterventional study, as medications are prescribed according to routine clinical practice, no additional diagnostic or monitoring procedures are imposed, and epidemiological methods are used for analysis. Moreover, the risks for participants are minimal and proportionate to the anticipated benefits. However, for methodological clarity, the term "intervention" is used in this protocol but refers solely to an informational decision-support process and not to a patient-level clinical intervention.

Therefore, individual informed consent is not required. Instead, a structured patient notification and opt-out procedure is implemented in accordance with institutional policies[32].

To protect participant privacy, all study data are recorded in a secure electronic data capture system (REDCap [Research Electronic Data Capture]). Only authorized study personnel have access to identifiable data. For analysis, data are pseudonymized and stored on secure institutional servers in accordance with applicable data protection regulations, including the General Data Protection Regulation.

Participants do not receive any financial compensation or incentives for participation in this study.

Informed Consent

Individual informed consent is not required, as this study was classified as a noninterventional quality improvement project.

Moreover, obtaining individual informed consent was considered inappropriate for several reasons: it could introduce bias by alerting physicians and/or patients of their label, potentially leading to behavioral changes compromising study validity, particularly in the control group; this study poses minimal risk, with potential disadvantages proportional to the expected benefits; the intervention targets physician prescribing behavior, not individual patients; a complete dataset is essential, excluding nonconsenting patients would limit validity and representativeness.

Instead of traditional informed consent, a structured patient notification and opt-out procedure is applied. Participants are informed through a study information letter made available via the EPR user application (Mynexuzhealth) and via the publicly accessible study webpage [33]. This study is also publicly announced through posters displayed in participating hospitals (Figure S2 in Multimedia Appendix 1). Patients are explicitly informed that they can withdraw at any time without consequences for their treatment by contacting this study's team, as indicated in the information letter. If the dataset is not yet locked, previously collected data will be removed, and no further data will be collected. Any proposed label adjustments are actively communicated to both patients and their HCPs through the EPR (cf. Infra, communication of the AFC results, intervention group, if applicable).

These procedures were reviewed in depth and formally approved by the central ethics committee as well as by the local ethics committees of all participating sites.

Inclusion Criteria

The inclusion criteria are hospitalized patients and BLAL (for one or more beta-lactams).

Exclusion Criteria

The exclusion criteria are being aged younger than 18 years, previous enrollment in this study, patients receiving palliative care, and patients expected to be discharged within 24 hours.

Baseline Measurements

For each patient, baseline data will be recorded in the electronic case report form using the REDCap (cf. Infra). The collected information will encompass sociodemographic details, clinical data, and specific information on the BLALs as registered in the EPR. Sociodemographic information will include age, gender, and contact details of both the general physician and the community pharmacist. Clinical information will capture the reason for admission and, if applicable, the condition requiring antibiotic treatment, categorized according to the *ICD* (*International Classification of Diseases*) codes. Additionally, any antibiotics (Anatomical Therapeutic Chemical codes J) already administered during the current hospitalization will be documented. To assess comorbidity burden, the presence of specific conditions contributing to the Charlson Comorbidity Index will be recorded, including myocardial infarction, congestive heart failure, peripheral vascular disease, cerebrovascular disease, dementia, chronic pulmonary disease, rheumatologic disease, peptic ulcer disease, liver disease, diabetes, hemiplegia or paraplegia, renal disease, malignancy (localized or metastatic), leukemia, lymphoma, and acquired immunodeficiency syndrome. Information on BLALs will be extracted from the allergy registration module within the EPR. This will include the registration date of the BLAL, the specific drug involved (Anatomical Therapeutic Chemical code), the onset date of the allergic reaction, reported symptoms, and the source of information. Additionally, the probability of label accuracy, as indicated by the reporter, will be documented.

XSL•FO
RenderX

AFC CDS Tool (Intervention Group)

A clinical rule embedded in the EPR checks in near-time for re-exposure to the culprit antibiotic or class member, from the date of BLAL registration to the date of patient inclusion (the AFC). The AFC is incorporated into the hospital information system as an "if-then" algorithm, using real-time structured data available in the EPR. The AFC incorporates information from the electronic prescribing module in the EPR and uses an integrated natural language processing tool to identify re-exposures since the introduction of the BLAL in the various EPR records. Additionally, it lists data on coadministration of antihistamines, glucocorticoids, or epinephrine during the re-exposure period, up to 24 hours after the last antibiotic dose. These medications are considered proxies for possible allergic reactions and help inform the subsequent review process. The results of the screening, that is, the clinical rule alerts, are compiled on a structured worklist in the hospital information system for secondary review by a specified trained person, that is, a member of the study team in the current study, before alerting the treating physician [24]. Validated flowcharts are used to support the review, ensuring a standardized and systematic approach. Schematic algorithms are used to guide the actual delabeling or refinement advice based on the allergy label, subsequent exposure, and consequences linked to this exposure (Figure S3 in Multimedia Appendix 1).

Consulting primary care data could not yet be automated due to the lack of interoperability between different EPRs. In order to retain a focus on the impact of the AFC alone, we did not systematically obtain information from the primary care HCP (physician and/or pharmacist) on beta-lactam re-exposure. In our pilot study [30], this step provided only a minor contribution.

Based on the collected information, BLALs are allocated to 3 groups: to delabel, to refine, or to leave as such. An incorrect BLAL is assigned when patients reported a clearly nonallergic adverse drug reaction, when a negative rechallenge or uneventful re-exposure was identified since registration of the BLAL; a correct BLAL is assigned when an allergic reaction was observed in the hospital or when positive skin and/or provocation testing was found; and an undetermined BLAL is assigned when there is no proof of correctness or incorrectness of the BLAL. Allocation is performed by a member of this study's team, and if needed, discussed and validated with the principal investigator.

BLAL Adjustment (Intervention Group, if Applicable)

In the case of a correct or undetermined BLAL, no action will be taken. In case of an incorrect BLAL, a suggestion will be sent to the treating physicians in the EPR ("follow-up note"), to delabel (re-exposure to the culprit found) or to refine (harmless exposure to nonculprit beta-lactams, excluding monobactams or carbapenems found) the BLAL accordingly. The note will include a summary of the algorithm's findings, a suggestion for review of the label in consultation with the patient, a reference to the in-house delabeling procedure (written and video), and contact details for questions and additional information. If the BLAL is not changed in the following 24 - 72 hours, the treating physician will be contacted by phone to discuss potential delabeling; reasons to discard the advice will be noted in the electronic case report form (as part of the process evaluation) and in the formal study report (cf. Infra). A distinction is made between definite tolerated re-exposure and presumed tolerance. The former is scored in case of documented administration through the electronic prescribing module (in which prescription and registered administration are required), in the EPR, and/or when bedside uneventful re-exposure could be orally confirmed by the treating hospital physician. The latter is scored when free-text notes or ambulatory prescriptions of antibiotic delivery are identified in the EPR, in the absence of EPR-registered bedside verification by the treating hospital physician.

Communication of the AFC Results (Intervention Group, if Applicable)

In case of an incorrect BLAL, this study's member will make a formal report of the AFC result in the EPR (Figure S4 in Multimedia Appendix 1), available for hospital physicians and, after validation of the coordinating investigator, also for the patient in the Mynexuzhealth app. This report is also sent to the primary care HCP via e-letter through the eHealth platform. The report will include a summary of our findings and, if applicable, label adjustment. Patients are referred to their primary care HCP, this study's information letter for participants, and this study's webpage with contact details of this study's team for further questions and/or disagreement with the label adjustment.

Follow-Up or Outcome Measurement

To evaluate the effectiveness and safety of the intervention, information on antibiotic prescriptions, BLAL details, safety indicators, clinical outcomes, microbiological end points, and economic factors will be extracted from the EPR at days 30, 100, and 200. Antibiotic prescription will be monitored throughout the follow-up period. BLAL details will include the outcome of the assessment, along with any newly recorded or relabeled BLALs. The primary outcome is assessed per patient at the predefined time points. Inconclusive outcomes will be single-blinded evaluated by 2 independent infectious disease specialists who, upon disagreement, discuss the outcome with a third senior infectious disease specialist for a consensus. In the absence of doubt regarding the use (or not) of a first-line regimen, the scores are scored (unblinded) by the central study team (n=2). Safety parameters will be evaluated based on antibiotic tolerance and the need to switch antibiotics during treatment. Clinical outcomes will be assessed by measuring the number of days until in-hospital mortality, mortality at 3 months posthospitalization, LOS, and ICU admission rate. Microbiological end points will include colonization and/or infection of methicillin-resistant *Staphylococcus aureus*, vancomycin-resistant *Enterococcus*, and *C. diff.* economic end points, focusing on the societal perspective, will capture the cost of antibiotic treatment and hospitalization, as well as the time and human resources required. However, detailed economic analysis will be the subject of a separate protocol. No patients will be lost to follow-up within participating hospitals, as this study relies exclusively on back-office EPR data extraction. Outcomes are captured through the shared EPR platform "Clinical Workstation" (CWS, Nexuzhealth), a multihospital

XSL•FO
RenderX

EPR network used by several tertiary and regional Belgian hospitals. However, antibiotic exposure and readmissions are only captured for hospitals using this EPR network; outpatient antibiotic use and care in nonnetwork hospitals are not captured. Moreover, survival is centrally registered and automatically linked to the patient's Belgian social security number, allowing us to reliably follow up on this parameter irrespective of subsequent hospitalizations.

Missing Data

Baseline variables used by the AFC tool are routinely collected in the electronic health record and are therefore expected to be minimally missing. For follow-up outcomes, mortality is centrally collected (and therefore not expected to yield missing data). However, intercenter variability in data capture may occur. Therefore, for time-to-event outcomes (survival), patients will be censored at the last known follow-up in accordance with standard survival analysis methods. For binary and continuous secondary outcomes (eg, antimicrobial resistance and antibiotic use), we will assume data are missing at random and apply multilevel multiple imputation, using Rubin rules. For the window-based prescribing outcomes, windows without in-hospital systemic antibiotic exposure are considered not applicable ("no opportunity") rather than missing and are excluded from the denominator of the cumulative primary end point. Patients without any applicable windows (no in-hospital antibiotic exposure in any assessment window up to day 100) will be excluded from the primary end point analysis, as the cumulative guideline-concordance proportion is not defined for these patients. The number and proportion of such patients will be reported by study phase. The proportion and patterns of missing data will be reported accordingly, and sensitivity analyses will be conducted using complete-case analyses and worst-case assumptions to assess the robustness of findings.

Sample Size Determination

The main aim of our study is to establish the assumed superiority of using the AFC over SOC in increasing the proportion of patients with first-line antibiotics prescribed. Therefore, we look at the proportion of patients receiving first-line or narrow-spectrum beta-lactam antibiotics within the total number of patients receiving antibiotics. The crossover design accounts for seasonal as well as intra- and interhospital differences in overall antibiotic use. Nevertheless, overall antibiotic use will be monitored at the participating sites during this trial's period.

The following assumptions were applied.

For the control arm, in a recent retrospective study in UZ Leuven, involving 21,999 patients with pneumonia, pyelonephritis, infections associated with heart, kidney, liver, or lung transplantation; appendectomy; coronary artery bypass grafting; and total knee or hip replacement, we investigated the difference in antibiotic use between patients with and without BLAL [21]. It showed that 10.5% of patients with a BLAL received amoxicillin despite their label. Based on the latter, we estimate that 10% of the patients in the control group already receive first-line antibiotics. We add a 5% increase due to indirect learning of our protocol because we cannot exclude that treating physicians learn to use discordances in the medical file for delabeling or refinement of the allergy labels, independent of our study and AFC tool. Considering around 30% of hospitalized patients receive antibiotics during their stay [34], this brings the percentage down to 4.6% in the control group.

For the intervention arm, we previously showed that our CDS approach has the potential to delabel 38% of BLALs [30]. We add a margin of 5% for physicians or patients who refuse to delabel or be delabeled [9]. Considering around 30% of hospitalized patients will receive antibiotics during their stay, this results in an overall estimate for the intervention group of 9.9%.

For the intra- and intercluster correlation (ICC), literature indicates a moderate ICC of 0.059 for antibiotic prescription behavior [35,36]. This estimate seems appropriate as moderate differences in antibiotic prescription behavior are expected within and between different physicians at participating sites, which mostly follow the IGGI guidelines, however, with some internal adjustments. Indeed, the most recent point prevalence survey of antimicrobial use and health care–associated infections in Belgian acute care hospitals (results of the global point prevalence survey and European Center for Disease Prevention and Control point prevalence survey) reports 77% compliance with IGGI guidelines [37]. Additionally, 3 participating sites are part of the same Hospital Outbreak Support Team network, that is, Heilig Hartziekenhuis Mol, AZ Turnhout, and AZ Herentals in the Hospital Outbreak Support Team Network Kempen, which has the specific goal to align antimicrobial stewardship. All centers use the same EPR (CWS). The major difference would be expected between UZ Leuven, as an academic hospital, and the other participating sites, as regional hospitals. The AFC is currently only available in the CWS EPR system, and therefore, we were unable to include other academic centers as well. A sensitivity analysis of academic vs regional hospitals will be performed.

Considering the number of hospitalizations per year in each center and a prevalence of 6% of BLALs [8], we expect an inclusion rate of 3285 participants.

Sample size and power calculations were carried out according to the methods described by Hemming et al [38], using the Shiny CRT Calculator [39]. Based on the design given in Figure 1, assuming a discrete time decay (allowing correlations between periods to decay with each period), a within-period ICC of 0.059 and an autocorrelation (ratio of between-period and within-period ICC) of 0.6, this study will have 80% power to detect a statistically significant difference between control and intervention using a 2-sided significance level of 5%, assuming 4.6% and 9.9% of patients to have the primary end point in the control and intervention group, respectively. This sample size calculation is based on patient-level appropriate prescribing proportions and is considered a conservative approximation for the cumulative window-based primary end point.

In case key assumptions differ importantly from those used in the initial calculations, we will implement a preplanned sensitivity and adaptation strategy focused on nuisance parameters rather than treatment effect estimates. Once the initial target number of inclusions has been reached, we will

XSL•FO
RenderX

conduct a blinded (or arm-masked) internal review of (1) overall primary end point incidence, (2) observed within-period ICC and between-period correlation structure, and (3) completeness of follow-up and outcome ascertainment across centers and periods. In parallel, we will run sensitivity power analyses across plausible ranges of ICC or autocorrelation to quantify the robustness of power under departures from assumptions. If this review indicates that statistical power is likely to be compromised (eg, lower-than-expected event rate or higher-than-expected clustering), we will adapt by extending recruitment and/or adding additional crossover periods and/or centers, while keeping the intervention allocation scheme and primary outcome definition unchanged, depending on the available human resources to perform this. Any such modifications will be documented as protocol amendments and reflected in an updated statistical analysis plan. Importantly, to avoid operational recruitment gaps and preserve feasibility, recruitment will continue while this internal review is performed, and the adaptation decision will be based on updated nuisance-parameter estimates rather than interim efficacy signals.

Statistical Analysis

The primary endpoint is cumulative guideline-concordant use of first-line and/or narrow-spectrum beta-lactam antibiotics according to IGGI-based local guidelines, assessed across predefined assessment windows up to day 100 (days 1, 30, and 100). Within each window, guideline concordance is scored at the patient level in an any or ever manner as 1 if at least one in-hospital systemic antibiotic course is guideline-concordant and 0 if otherwise. If multiple courses occur within a window, the window is considered concordant if any course meets guideline criteria. This any or ever approach was chosen in line with the quality-improvement objective of the AFC intervention, capturing whether appropriate first-line therapy was achieved within a care window while limiting instability due to multiple within-episode treatment changes. Windows without in-hospital systemic antibiotic administration are considered not applicable (“no opportunity”) and are excluded from the denominator.

For each patient, a cumulative weighted guideline-concordance score is calculated at day 100 as the proportion of applicable windows with guideline-concordant therapy (s/n). The primary comparison between intervention (AFC) and control will use a binomial generalized linear mixed-effects model with logit link, modeling s out of n, including fixed effects for study phase and period, and a random intercept for hospital to account for clustering in the crossover design. A prespecified adjusted model will additionally include age, sex, and the Charlson comorbidity index. Patients without any applicable windows will be excluded from this analysis and reported separately. The primary estimand targets the period during which the AFC is expected to plausibly influence prescribing through contemporaneous decision support and short-term label adjustment.

Secondary endpoints include guideline-concordant beta-lactam use at each individual assessment window (including day 200), delabeling or refinement rate, beta-lactam tolerance, antibiotic switching, ICU admission, readmission, resistant organism outcomes, LOS, mortality, and cost outcomes. Outcomes at day 200 are analyzed as secondary durability measures, as longer-term prescribing may be influenced by label reappearance or documentation changes unrelated to the index intervention. Secondary endpoints will be analyzed using hierarchical mixed-effects models appropriate to outcome type (binary, continuous, or time-to-event), with fixed effects for study phase and period and clustering by hospital. Process evaluation outcomes will be analyzed descriptively.

Predefined subgroup analyses will assess potential effect modification by including phase-by-subgroup interaction terms in the primary model. Subgroups of interest are BLAL type (penicillin vs other labels), reported reaction severity category, time since label registration, infection type (prophylaxis vs therapeutic indication), ICU admission status, and hospital type (academic vs regional). Subgroup analyses will focus on the primary endpoint and are considered exploratory and hypothesis-generating.

Moreover, because the AFC is expected to influence prescribing primarily through label delabeling or refinement, we will perform an exploratory implementation analysis within the AFC phase. Patients will be classified as with implementation if the BLAL was delabeled or refined after the AFC recommendation, and as not with implementation if otherwise. The cumulative primary endpoint (s/n up to day 100) will be summarized by implementation status; any between-group comparisons will be considered descriptive and interpreted cautiously, as implementation is a postallocation process measure and may be confounded.

No formal adjustment for multiplicity will be applied. Secondary and subgroup analyses will be interpreted with caution.

Sensitivity analyses will include evaluation of alternative covariance structures for within-cluster and within-period correlation (eg, exchangeable or autoregressive), with model selection guided by goodness-of-fit criteria, and estimation using generalized estimating equations with robust SEs to obtain marginal effect estimates and assess robustness to modeling assumptions. As a robustness check, we will report the proportion of patients with any recorded in-hospital follow-up within the shared EPR network (CWS) by study phase and period.

A detailed statistical analysis plan and data management plan will be finalized before database lock.

Quality Assurance

In order to ensure the same quality and safety standards in patient care for clinical research as commonly applied by the research team in its regular activities, the research team shall comply with the following obligations: (1) the research team will use trained and qualified employees or contractors to manage and coordinate this study; (2) the sponsor will ensure that this (multicenter) study’s reporting is reliable and valid, statistically accurate, ethical, and unbiased; (3) the research team will not grant incentives, other than standard compensations and reimbursement of costs, to this study’s participants or to participating site’s staff that would compromise the integrity of the research; (4) the research team is responsible for monitoring and evaluating the quality, safety,

XSL•FO
RenderX

and ethics of this study and will respect the participating site's policies and processes when performing such monitoring and evaluation activities; and (5) the research team will protect the privacy and confidentiality of this study participants in accordance with all applicable laws.

A dedicated postdoctoral researcher and pharmacist (LG) will be in charge of overall trial (quality) management, and monthly coordinating investigator study team meetings will be organized.

Moreover, a dedicated advisory board is installed for this project, consisting of all principal investigators, other experts, and additional stakeholders for further valorization of our approach. This advisory board will be informed and consulted about the project progress at least every 6 months (ie, 6 times in total). A dedicated Teams (Microsoft Corp) environment will be created to ensure efficient communication during and in between the semestral board meetings. The overall role of the advisory board is to ensure the project goals, milestones, and deliverables are reached in time, to ensure the technicalities of the tool match the needs and criteria for the respective stakeholder group, and to promote usage of the tool postproject in the relevant stakeholder groups.

Results

Recruitment for this multicenter crossover cluster-controlled study is ongoing across hospitals in Flanders, Belgium. Recruitment started in March 2025. As of March 2026, more than 3000 participants have been enrolled across 6 participating hospitals, which approximates the initially estimated sample size required for the planned power calculation.

According to the prespecified protocol, an interim assessment of nuisance parameters will be conducted in the following month to evaluate the initial design assumptions. Following this assessment, recruitment will continue with the participation of 3 additional hospitals, with data collection expected to continue until late 2026.

The final study database is anticipated to be locked in late 2026, after completion of follow-up. The prespecified statistical analyses will then be performed. The main study results are expected to be reported in 2027.

Discussion

Noninvasive delabeling is an often-forgotten step in the process of correct BLAL labeling and delabeling. Previous work, including ours, has shown its potential. However, its impact on clinical end points is unclear. Here, a novel CDS tool, the AFC, embedded in the EPR and driven by an expert-elaborated algorithm, is expected to be a key tool for (proactive) noninvasive antibiotic allergy delabeling or refinement. This study aims to evaluate the clinical, antimicrobial, and economic impact of implementing the AFC for hospitalized adults with BLALs compared to SOC (no AFC). The AFC holds the potential to also optimize other (nonantibiotic, eg, iodinated contrast media and others) allergy labels. Moreover, it would offer additional arguments to support health data sharing and the development of an overarching patient record system, which would stimulate interdisciplinary communication. Additional valorization steps, such as leveraging our tool to nonhospital EPRs, would be relevant as improving BLAL evidently transcends the boundaries of primary, secondary, and tertiary care facilities. The heterogeneous primary care EPR antibiotic allergy labels are not yet linked to the hospital EPR in our country. Nevertheless, we believe the future integration and sharing of extramural medical data holds great promise and should remain a next step. A centralized registration module (in eHealth Belgium or through a European Allergy Module in the European Health Data Space) would provide a solution, but requires further optimization. However, additional elements need to be taken into account. First, the primary endpoint used a cumulative, window-based guideline-concordance measure with any or ever scoring up to day 100. This pragmatic definition aligns with the quality-improvement objective of the AFC intervention and captures whether appropriate first-line therapy was achieved within each care window, but it simplifies complex treatment courses and may overestimate concordance when multiple regimens occur within a window. Second, follow-up antibiotic use and readmissions are ascertained from in-hospital records across hospitals using the shared EPR platform (CWS), which includes several large Belgian hospitals. Care delivered outside this network and outpatient antibiotic use are not captured, which may lead to underestimation of total exposure and guideline concordance. If nondifferential between study phases, this would likely bias effect estimates toward the null; the degree of in-network follow-up capture will therefore be reported. Third, exposure to the AFC may directly or indirectly induce a learning effect, potentially reducing the observed difference between the intervention and control groups. To avoid overestimation of the AFC effect, we conservatively corrected for an estimated 5% increase in non-AFC–induced delabeling in the control group. Importantly, a learning effect is likely to occur in both groups over time; therefore, this correction should be regarded as an approximation allowing a more accurate estimation of the true AFC-specific impact. Residual confounding due to non-randomized allocation and temporal variation in clinical context cannot be fully excluded, although this risk was mitigated through the crossover cluster design and adjustment for period and hospital effects in the statistical analysis. In addition, truly allergic patients may lack documented allergy labels, and information on drugs given outside of the registered hospital context may be incorrect. Finally, the scalability of our AFC tool will depend on its ability to be integrated into other EPR software platforms. The AFC tool is based on the principle of discordant patient information. Exploiting these discrepancies to the benefit of the patients would allow for improved EPRs and is in line with the goal of a unified EPR.

XSL•FO
RenderX

Acknowledgments

We thank Sofie Heeckhout, PharmD, Dries Grymonprez, MSc, Lisa Santy, MD, Eef Verhesen, PharmD, Charline Dilen, MSc, Dana Van Kerkhoven, MD, Erwin Ho, MD, Peter Verbeeck, MD, Geanna Peeters, MSc, Frederik Staels, MD, Anneleen Peeters, MD, and Clara Ceyssens, MD, for their support with the implementation of this protocol within the different centers. We thank Nghia Nguyen, MSc, and Dries Wets, MD, for writing assistance.

Funding

This research is supported by the Research Foundation Flanders (Fonds Wetenschappelijk Onderzoek [FWO]; project T003023N) and KU Leuven (C3/23/028). RS holds a senior clinical investigator fellowship from the FWO (1805518N, 1805523N). LG is a FWO junior postdoctoral fellow (12A6Q24N). The funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of this paper.

Authors' Contributions

RS and IS conceptualized this study, and LG drafted the protocol and paper. GVDS contributed to the methodological design of this study and critically revised this paper. All authors read and approved this final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supporting figures and table.
[DOCX File, 2155 KB - resprot_v15i1e86056_app1.docx]

References

1. Blumenthal KG, Peter JG, Trubiano JA, Phillips EJ. Antibiotic allergy. Lancet 2019 Jan 12;393(10167):183-198. [doi: 10.1016/S0140-6736(18)32218-9] [Medline: 30558872]
2. Moran R, Devchand M, Smibert O, Trubiano JA. Antibiotic allergy labels in hospitalized and critically ill adults: a review of current impacts of inaccurate labelling. Br J Clin Pharmacol 2019 Mar;85(3):492-500. [doi: 10.1111/bcp.13830] [Medline: 30521088]
3. Macy E, Poon K-Y T. Self-reported antibiotic allergy incidence and prevalence: age and sex effects. Am J Med 2009 Aug;122(8):778. [doi: 10.1016/j.amjmed.2009.01.034]
4. Zhou L, Dhopeshwarkar N, Blumenthal KG, et al. Drug allergies documented in electronic health records of a large healthcare system. Allergy 2016 Sep;71(9):1305-1313. [doi: 10.1111/all.12881] [Medline: 26970431]
5. Charneski L, Deshpande G, Smith SW. Impact of an antimicrobial allergy label in the medical record on clinical outcomes in hospitalized patients. Pharmacotherapy 2011 Aug;31(8):742-747. [doi: 10.1592/phco.31.8.742] [Medline: 21923600]
6. Sousa-Pinto B, Fonseca JA, Gomes ER. Frequency of self-reported drug allergy: a systematic review and meta-analysis with meta-regression. Ann Allergy Asthma Immunol 2017 Oct;119(4):362-373. [doi: 10.1016/j.anai.2017.07.009] [Medline: 28779998]
7. van Dijk SM, Gardarsdottir H, Wassenberg MWM, Oosterheert JJ, de Groot MCH, Rockmann H. The high impact of penicillin allergy registration in hospitalized patients. J Allergy Clin Immunol Pract 2016;4(5):926-931. [doi: 10.1016/j.jaip.2016.03.009] [Medline: 27131826]
8. Gilissen L, Spriet I, Gilis K, Peetermans WE, Schrijvers R. Prevalence of antibiotic allergy labels in a tertiary referral center in Belgium. J Allergy Clin Immunol Pract 2021 Jun;9(6):2415-2425. [doi: 10.1016/j.jaip.2021.01.047] [Medline: 33607341]
9. Chua KYL, Vogrin S, Bury S, et al. The penicillin allergy delabeling program: a multicenter whole-of-hospital health services intervention and comparative effectiveness study. Clin Infect Dis 2021 Aug 2;73(3):487-496. [doi: 10.1093/cid/ciaa653]
10. Chiriac AM, Wang Y, Schrijvers R, et al. Designing predictive models for beta-lactam allergy using the drug allergy and hypersensitivity database. J Allergy Clin Immunol Pract 2018;6(1):139-148. [doi: 10.1016/j.jaip.2017.04.045] [Medline: 28739365]
11. Van Gasse AL, Oulkadi R, Mousati Z, et al. Prevalence of self-reported and confirmed penicillin allergy in a Belgian outpatient population. Allergy 2020 Aug;75(8):2111-2115. [doi: 10.1111/all.14292] [Medline: 32243598]
12. Conway EL, Lin K, Sellick JA, et al. Impact of penicillin allergy on time to first dose of antimicrobial therapy and clinical outcomes. Clin Ther 2017 Nov;39(11):2276-2283. [doi: 10.1016/j.clinthera.2017.09.012] [Medline: 29032850]
13. Lam W, Staicu ML, Conn KM, Ramsey AC. Is there a higher prevalence of methicillin-resistant Staphylococcus aureus (MRSA) or vancomycin-resistant Enterococcus (VRE) colonization in patients with antibiotic allergy labels? Am J Infect Control 2020 Jun;48(6):663-667. [doi: 10.1016/j.ajic.2019.09.021] [Medline: 31735590]

XSL•FO
RenderX

14. Macy E, Contreras R. Health care use and serious infection prevalence associated with penicillin “allergy” in hospitalized patients: a cohort study. J Allergy Clin Immunol 2014 Mar;133(3):790-796. [doi: 10.1016/j.jaci.2013.09.021] [Medline: 24188976]
15. Reddy V, Baman NS, Whitener C, Ishmael FT. Drug resistant infections with methicillin-resistant Staphylococcus aureus, clostridium difficile, and vancomycin resistant Enterococcus are associated with a higher prevalence of penicillin allergy. J Allergy Clin Immunol 2013 Feb;131(2):AB170. [doi: 10.1016/j.jaci.2012.12.1269] [Medline: 23916153]
16. Blumenthal KG, Lu N, Zhang Y, Walensky RP, Choi HK. Recorded penicillin allergy and risk of mortality: a population-based matched cohort study. J Gen Intern Med 2019 Sep;34(9):1685-1687. [doi: 10.1007/s11606-019-04991-y] [Medline: 31011962]
17. Wu JC, Langford BJ, Schwartz KL, et al. Potential negative effects of antimicrobial allergy labelling on patient care: a systematic review. Can J Hosp Pharm 2018;71(1):29-35. [Medline: 29531395]
18. Picard M, Bégin P, Bouchard H, et al. Treatment of patients with a history of penicillin allergy in a large tertiary-care academic hospital. J Allergy Clin Immunol Pract 2013;1(3):252-257. [doi: 10.1016/j.jaip.2013.01.006] [Medline: 24565481]
19. Sade K, Holtzer I, Levo Y, Kivity S. The economic burden of antibiotic treatment of penicillin-allergic patients in internal medicine wards of a general tertiary care hospital. Clin Exp Allergy 2003 Apr;33(4):501-506. [doi: 10.1046/j.1365-2222.2003.01638.x] [Medline: 12680867]
20. Mattingly TJ, Fulton A, Lumish RA, et al. The cost of self-reported penicillin allergy: a systematic review. J Allergy Clin Immunol Pract 2018;6(5):1649-1654. [doi: 10.1016/j.jaip.2017.12.033] [Medline: 29355644]
21. Gilissen L, Van De Sijpe G, Laenen A, et al. Clinical outcome associated with beta‐lactam allergy labels in hospitalized patients in Belgium. Clinical and Translational Allergy (forthcoming) . [doi: 10.1002/clt2.70166]
22. Blumenthal KG, Shenoy ES, Wolfson AR, et al. Addressing inpatient beta-lactam allergies: a multihospital implementation. J Allergy Clin Immunol Pract 2017;5(3):616-625. [doi: 10.1016/j.jaip.2017.02.019] [Medline: 28483315]
23. Trubiano JA, Vogrin S, Chua KYL, et al. Development and validation of a penicillin allergy clinical decision rule. JAMA Intern Med 2020 May 1;180(5):745. [doi: 10.1001/jamainternmed.2020.0403]
24. Romano A, Atanaskovic-Markovic M, Barbaud A, et al. Towards a more precise diagnosis of hypersensitivity to beta-lactams - an EAACI position paper. Allergy 2020 Jun;75(6):1300-1315. [doi: 10.1111/all.14122] [Medline: 31749148]
25. Khan DA, Banerji A, Blumenthal KG, et al. Drug allergy: a 2022 practice parameter update. J Allergy Clin Immunol 2022 Dec;150(6):1333-1393. [doi: 10.1016/j.jaci.2022.08.028] [Medline: 36122788]
26. Wolfson AR, Huebner EM, Blumenthal KG. Acute care beta-lactam allergy pathways: approaches and outcomes. Ann Allergy Asthma Immunol 2019 Jul;123(1):16-34. [doi: 10.1016/j.anai.2019.04.009] [Medline: 31009700]
27. du Plessis T, Walls G, Jordan A, Holland DJ. Implementation of a pharmacist-led penicillin allergy de-labelling service in a public hospital. J Antimicrob Chemother 2019 May 1;74(5):1438-1446. [doi: 10.1093/jac/dky575] [Medline: 30753497]
28. Torda A, Chan V. Antibiotic allergy labels-the impact of taking a clinical history. Int J Clin Pract 2018 Mar;72(3):e13058. [doi: 10.1111/ijcp.13058] [Medline: 29341373]
29. Devchand M, Kirkpatrick CMJ, Stevenson W, et al. Evaluation of a pharmacist-led penicillin allergy de-labelling ward round: a novel antimicrobial stewardship intervention. J Antimicrob Chemother 2019 Jun 1;74(6):1725-1730. [doi: 10.1093/jac/dkz082] [Medline: 30869124]
30. Van De Sijpe G, Gilissen L, Vandebotermet M, Peetermans WE, Spriet I, Schrijvers R. Non-invasive delabeling and refining of beta-lactam allergy labels in inpatients to optimize antimicrobial stewardship. Allergy 2022 Oct;77(10):3157-3159. [doi: 10.1111/all.15453] [Medline: 35876226]
31. Van De Sijpe G, Gilissen L, Wets D, et al. Evaluation of the Allergy Fact Checker, a clinical decision support system for non-invasive beta-lactam delabeling: a mixed-methods study. NPJ Digit Med 2025 Nov 17;8(1):662. [doi: 10.1038/s41746-025-02030-1] [Medline: 41249542]
32. Wet inzake experimenten op de menselijke persoon (law on experiments on the human person). . 2004 May 7.
33. STAAL studie [article in dutch]. KU Leuven. URL: https://gbiomed.kuleuven.be/english/research/50000698/50000631/staal-studie/staalstudie [accessed 2026-03-07]
34. Catteau. Antimicrobial Use in Belgian Acute Care Hospitals: Results of the 2022 ECDC Point Prevalence Survey sciensano.be.: SHEA Spring; 2024. URL: https://www.sciensano.be/sites/default/files/shea_poster_pps_lc.pdf [accessed 2026-03-07]
35. Ferreira-González I, Carrillo X, Martín V, et al. Variabilidad interhospitalaria en la prescripción tras un síndrome coronario agudo: hallazgos del estudio ACDC. Rev Esp Cardiol 2016 Feb;69(2):117-124. [doi: 10.1016/j.recesp.2015.04.018]
36. O’Brien K, Bellis TW, Kelson M, Hood K, Butler CC, Edwards A. Clinical predictors of antibiotic prescribing for acutely ill children in primary care: an observational study. Br J Gen Pract 2015 Sep;65(638):e585-e592. [doi: 10.3399/bjgp15X686497] [Medline: 26324495]
37. Vandael E, Latour K, Goossens H, et al. Point prevalence survey of antimicrobial use and healthcare-associated infections in Belgian acute care hospitals: results of the Global-PPS and ECDC-PPS 2017. Antimicrob Resist Infect Control 2020 Dec;9(1):13. [doi: 10.1186/s13756-019-0663-7]

XSL•FO
RenderX

38. Hemming K, Kasza J, Hooper R, Forbes A, Taljaard M. A tutorial on sample size calculation for multiple-period cluster randomized parallel, cross-over and stepped-wedge trials using the Shiny CRT Calculator. Int J Epidemiol 2020 Jun 1;49(3):979-995. [doi: 10.1093/ije/dyz237] [Medline: 32087011]
39. The Shiny CRT Calculator: Power and Sample size for Cluster Randomised Trials. URL: https://clusterrcts.shinyapps.io/rshinyapp/ [accessed 2026-03-07]

Abbreviations

AFC: Allergy Fact Checker
AZ: Algemeen Ziekenhuis (general hospital)
BLAL: beta-lactam allergy label
CDS: clinical decision support
CME: Commissie voor Medische Ethiek
CWS: Clinical Workstation
EPR: electronic patient record
EU: European Union
HCP: health care provider
ICC: intra- and intercluster correlation
***ICD*:** *International Classification of Diseases*
ICU: intensive care unit
IGGI: Infectiologiegids/Guide d'Infectiologie of the Belgian Society for Infectiology and Clinical Microbiology
LOS: length of hospital stay
REDCap: Research Electronic Data Capture
SOC: standard of care
UZ: Universitair Ziekenhuis (university hospital)

Edited by J Sarvestan; submitted 17.Oct.2025; peer-reviewed by AP Douglas, MB Alnaes; revised version received 09.Feb.2026; accepted 19.Feb.2026; published 27.Mar.2026.

Please cite as:
Gilissen L, Van De Sijpe G, Spriet I, Schrijvers R
Impact Assessment of the Allergy Fact Checker, a Clinical Decision Support Tool for Noninvasive Beta-Lactam Antibiotic Allergy Label Delabeling: Protocol for a Multicenter Crossover Cluster-Controlled Study
JMIR Res Protoc 2026;15:e86056
URL: https://www.researchprotocols.org/2026/1/e86056
doi:10.2196/86056

© Liesbeth Gilissen, Greet Van De Sijpe, Isabel Spriet, Rik Schrijvers. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of a Collaborative Education Program for Home Care Nursing Educators' on-the-Job Training in End-of-Life Home Care: Protocol for a Randomized Controlled Trial

Ryousuke Yamada[1], MHS; Hanako Numata[1], PhD; Yuka Sumikawa[1], PhD; Kosuke Kashiwabara[2], PhD; Maiko Noguchi-Watanabe[1], PhD; Kayoko Kurita[3], PhD; Noriko Yamamoto-Mitani[1], PhD

[1]Department of Gerontological Home Care and Long-Term Care Nursing, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo, Japan

[2]Data Science Office, Clinical Research Promotion Center, The University of Tokyo Hospital, Tokyo, Japan

[3]Center for Research and Development in Higher Education, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Hanako Numata, PhD
Department of Gerontological Home Care and Long-Term Care Nursing, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-Ku, Tokyo, Japan

Abstract

Background: Despite the growing demand for end-of-life home care, nursing educators responsible for on-the-job training face substantial barriers, such as time and geographical constraints, which limit their access to professional development. Collaborative online training programs offer a potential solution to these challenges.

Objective: This study aims to develop and evaluate a collaborative online training program designed to improve the educational ability of home care nursing educators in delivering effective on-the-job training for home-based end-of-life care.

Methods: This parallel group, waitlist-controlled trial recruited home care nursing managers or educators from agencies with at least one home death client in the past year. We will provide this program on the internet. The intervention will comprise three online sessions, each lasting 60 minutes and held monthly, involving one lecture and two interactive case studies related to on-the-job training in end-of-life home care. This will be followed by 8 weeks of weekly educational messages via LINE, a mobile messaging app widely used in Japan. The control group will serve as a waitlist and will receive the same intervention after the control period. The primary outcomes are participants' attitudes, measured using the End-of-Life Nursing Education Questionnaire. Using a centralized computer-generated randomization system, 125 home care nursing managers and educators will be recruited and randomly assigned to either the intervention or control group at a ratio of 1:1. Stratification will be based on their agency's annual home death count between November 29, 2023, and November 29, 2024. To minimize performance bias, the intervention content has been standardized before the intervention to ensure fidelity.

Results: This study was approved by the Ethics Committee of the Graduate School of Medicine, The University of Tokyo (2024359NI-(2)) on November 15, 2024. Participant recruitment began on November 29, 2024, and closed on December 25, 2024, with 120 participants enrolled. As of February 15, 2026, the following number of surveys were completed at each time point: 106 at T0 (the intervention group:the control group; 52:54); 103 at T1 (51:52); 102 at T2 (50:52); 40 at T3 and 43 at T4 in the intervention group; and 47 at T5, 44 at T6, and 41 at T7 in the control group. Quantitative data collection was completed by January 2026, and qualitative interviews will be completed by March 2026. Primary quantitative results and integrated results are expected to be submitted for publication in 2026-2027 and 2027-2028, respectively.

Conclusions: This protocol can address geographical and time constraints by providing flexible, online educational resources for home care nursing educators. The anticipated outcomes include enhanced on-the-job training ability among nursing educators for home-based end-of-life care. This study's findings could support broader implementation of scalable training strategies in home care settings.

Trial Registration: The University Hospital Medical Information Network UMIN000056165; https://center6.umin.ac.jp/cgi-open-bin/ctr/ctr_view.cgi?recptno=R000063254

International Registered Report Identifier (IRRID): DERR1-10.2196/84330

(JMIR Res Protoc 2026;15:e84330) doi:10.2196/84330

KEYWORDS

continuing nursing education; end-of-life care; home health nursing; terminal care; on-the-job training

XSL•FO
RenderX

Introduction

With the growing aging population worldwide, the demand for home health care services has increased, as more individuals choose to receive medical care at home while managing chronic illnesses and disabilities [1]. Further, the proportion of individuals desiring to spend their final moments at home has risen, while actual home deaths have also increased. However, providing end-of-life (EOL) care presents several challenges, including physical and psychological distress among clients and families [2].

Home care nursing plays an important role in enabling EOL care at home. Previous studies have identified key factors facilitating home deaths, including medical conditions, personal preferences, and environmental support [3]. Among these, environmental support—including the availability of skilled home care nurses—has been essential for ensuring EOL home care. However, despite an increasing number of home care nursing agencies, a lack of systematic training opportunities for cultivating skilled home care nurses—primarily owing to limited time, funding, and human resources—has been reported as a significant barrier [4]. Therefore, establishing structured training systems to educate skilled nursing staff is crucial.

One of the key factors for enabling EOL home care is the availability of systematic training opportunities within the home care nursing agencies. Training opportunities for home care nurses at their agencies have been shown to contribute to improved staff retention and ability [5,6]. Several external training programs about EOL care, such as the End-of-Life Nursing Education Consortium [7] and the Program of End-of-Life Care for Home Visiting Nurses Training [8], provide foundational EOL care knowledge. However, as EOL home care requires the practical application of this knowledge within individual clients [9], such foundational knowledge is insufficient. On-the-job training (OJT) is particularly important in home care nursing agencies [10] because it can be provided in a tailored manner to suit the individual needs of each staff member and the client's context. In addition, among these constraints on time, funding, and human resources that challenge systematic training, OJT offers a practical option by being integrated into daily work. Nevertheless, home care nursing managers and staff have reported unsatisfactory preparation for delivering effective EOL home care [11].

Home care nursing managers and educators face challenges in attending external training programs owing to geographic and time constraints. Particularly, rural home care nursing agencies have limited access to such programs [12]. Furthermore, almost all home care nursing agencies have only one designated home care nursing manager. This single manager is responsible for multiple roles, including that of an educator [13]. This is why they often lack the time and scheduling flexibility to participate in lengthy in-person training courses, which require several hours to multiple days. Although some home care nursing agencies in Japan have staff designated as educators [14], most employ only one person in this role. Therefore, both managers and educators often face similar barriers in accessing training opportunities. Based on these limitations, a practical and accessible program that specifically supports educators in delivering OJT in EOL home care is needed.

To address these challenges, we will develop and evaluate an online training program designed specifically for home care nursing managers and educators to improve their educational abilities in conducting OJT for EOL home care. Given the geographic and time constraints of these managers and educators, we will develop a program that is entirely online, using a private messaging app for interaction to provide a short, flexible, and accessible learning format.

This study will evaluate the effectiveness of this training program through a randomized controlled trial (RCT). The primary outcome will be fostering changes in participants' attitudes as educators. Secondary outcomes will include improvements in knowledge and educational practices related to OJT in EOL home care. This study aims to establish a scalable evidence-based training model for widespread implementation in home care nursing agencies, thereby enhancing the quality of EOL home care.

Methods

Ethical Considerations

This study was approved by the Ethics Committee of the School of Medicine, The University of Tokyo (2024359NI-(2)). We will use an informed consent form to explain the study's purpose, voluntary participation, and privacy protection to all participants before enrollment. This is a minimal risk study. If participants experience distress or feel unwell, they can choose to pause or discontinue participation at any time. The principal investigator will serve as the contact point for any concerns. We will use an informed consent form to explain the study's purpose, voluntary participation, and privacy protection to the participants. After obtaining the consent of all participants, the study procedures will be conducted in accordance with the ethical guidelines of The Medical and Health Research Involving Human Participants in Japan [15] and the 1964 Declaration of Helsinki. Regarding incentives, participants will receive ¥5000 upon completion of all required surveys through T4 (course A) or T7 (course B; an exchange rate of ¥159.49 to US $1 is applicable). An additional ¥1000 will be provided to participants who contribute real-world educational cases for the case study sessions and to those who participate in interviews. Participants and staff members who participate in interviews will receive ¥1000 immediately after interview completion. No payments will be made to participants who withdraw before completing the required procedures. No additional compensation for harm is planned.

Design

This study will be an RCT whose design, implementation, and reporting adhere to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [16]. This RCT has been registered with the UMIN Clinical Trials Registry (UMIN000056165).

XSL•FO
RenderX

Setting

The program will be conducted via Zoom (Zoom Communications), a video communication platform. The sessions will include lectures, case studies, and interactive communications to enhance participants' learning and engagement. LINE, a freeware app for instant messaging and social networking developed by LY Corporation, will be used to distribute weekly educational messages and facilitate communication among participants throughout the intervention period.

Eligibility Criteria for Participants

The eligibility criteria for participants are as follows: home care nursing managers or educators (staff primarily responsible for education in their agency) working at an agency with at least one home death client in the previous year; access to Zoom using a computer; the ability to attend all scheduled training program sessions; and fluency in Japanese, which is defined as the capacity to read and understand written materials in Japanese.

Intervention Program

We developed an educational program for home care nursing managers and educators to improve their educational ability in OJT in EOL home care. The intervention group will receive this program after the baseline survey (T0) on a specific day. The control group will constitute a waitlist lasting 4 months after the baseline survey, and the same program will be delivered after the third survey (T2).

Program Theory of This Project

This section outlines the theoretical framework that guides the design and structure of our educational program. This project is primarily based on the theory of transfer of learning [17], which explains how knowledge is applied in the workplace post training. To apply this theory, our program's design focuses on three key domains: (1) trainee characteristics, (2) training design, and (3) work environment. The conceptual framework illustrating how these intervention components are designed to influence outcomes is presented in Figure 1.

Figure 1. Conceptual framework of the training program.

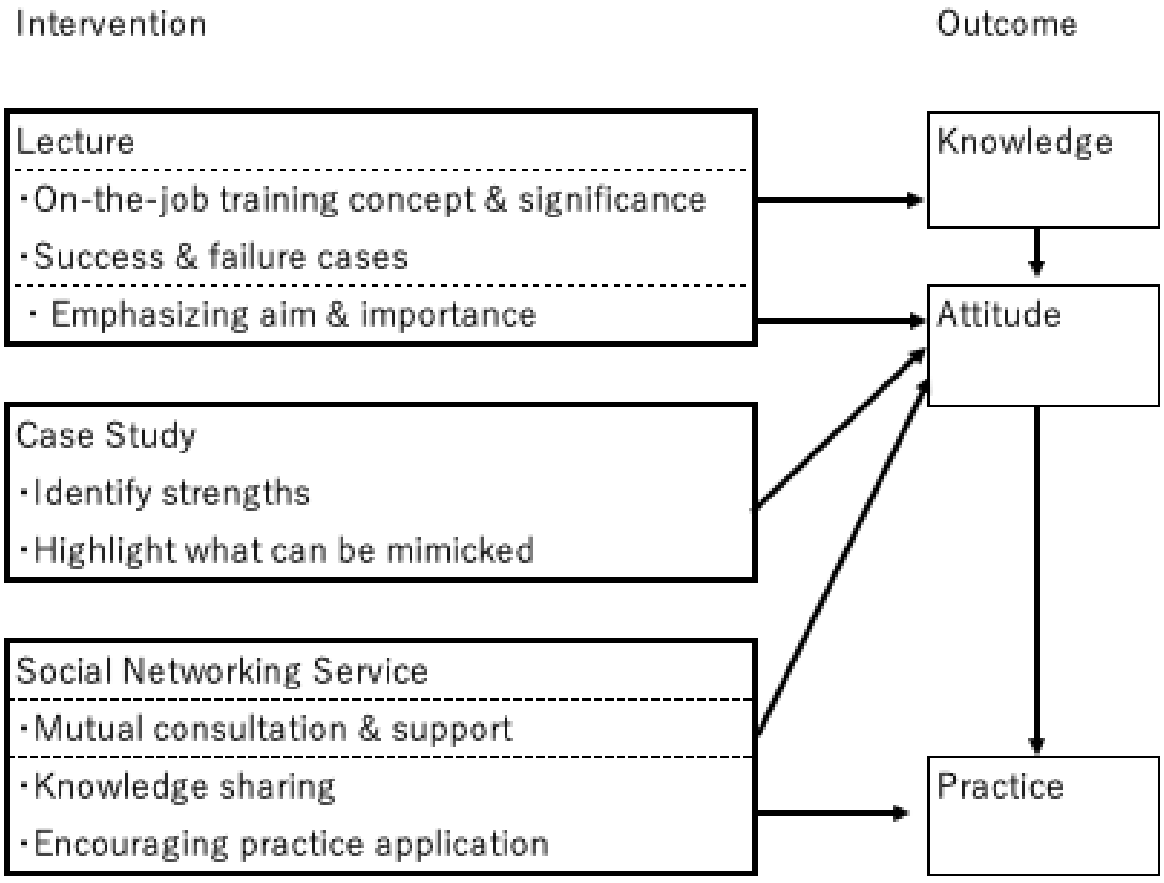

First, to address trainee characteristics, the model emphasizes the importance of high motivation, confidence, and positive outcome expectations to ensure effective training. To achieve this, we emphasize the aim and significance of the program based on the ARCS (attention, relevance, confidence, and satisfaction) model [18]. For example, at the beginning of the content, we present the objectives and goals that directly relate to real-world challenges faced by participants, thereby capturing their attention and establishing relevance. Second, regarding training designs, this program is grounded in Merrill's First Principles of Instruction [19] to ensure the learning is practical and effective. This program activates the participants' prior experience and facilitates the learning of new information in the lecture session. Subsequently, participants engage in

XSL•FO
RenderX

application and integration in the following case study session, where they use both their experiential and new knowledge to solve real-world problems. Third, to support the work environments and promote long-term maintenance of skills, we provide ongoing reinforcement. Based on previous research showing that social networking service–based messaging was an effective tool for prompting action [20], we use LINE for this purpose. Through this platform, we send weekly messages to reinforce key concepts from the sessions, prompt the application of skills, and foster a peer support community.

Program Content and Delivery

The program consists of three sessions: one lecture and two case studies, each lasting 60 minutes and conducted once every third Sunday of the month. This structure is based on the quality improvement collaborative approach, which promotes learning through periodic collaborative meetings supported by intersession communication [21]. This program is composed of three main types of sessions: a lecture, a case study, and weekly messages via LINE, which are designed to build on each other. The detailed objectives, goals, and content for each component are outlined in Textbox 1. The overall objective is for learners to be able to conduct effective OJT with adequate knowledge and attitude in EOL home care. A graphic syllabus provides participants with a clear, visual road map of this learning process (Figure 2).

Textbox 1. Table of contents.

Course objective

Learners can conduct effective on-the-job training with adequate knowledge and attitude for staff providing end-of-life home care.

Course goals

1. Learners will state adequate knowledge about on-the-job training in end-of-life home care.
2. Learners will describe practical educational methods that can be implemented in practice.
3. Learners will build confidence and motivation to conduct on-the-job training in end-of-life home care.
4. Learners will clarify their uncertainties regarding educational methods.

Program

Lecture objective

Learners can acquire educational knowledge and attitudes to conduct on-the-job training in end-of-life home care.

Lecture goals

1. Learners will identify challenges staff encounter during end-of-life home care.
2. Learners will describe the educational support required to address these challenges.

Lecture contents

1. The current situations and challenges faced by staff providing end-of-life home care.
2. Concrete methods for conducting effective on-the-job training in end-of-life home care.

Case study objective

Learners can understand practical educational methods in on-the-job training for staff providing end-of-life home care.

Case study goals

1. Learners will explain practical educational methods for on-the-job training in end-of-life home care.
2. Learners will gain confidence and motivation to implement on-the-job training in their practice.
3. Learners will be able to apply the acquired methods to conduct effective on-the-job training starting immediately.

Case study contents

1. Learners will understand one real-world case shared by a representative.
2. Group discussions will focus on best practices for conducting on-the-job training in end-of-life home care, based on a case scenario.
3. Learners will connect the knowledge from a lecture to a case scenario, fostering reflective and collaborative learning.

XSL•FO
RenderX

Figure 2. Graphic syllabus of this program.

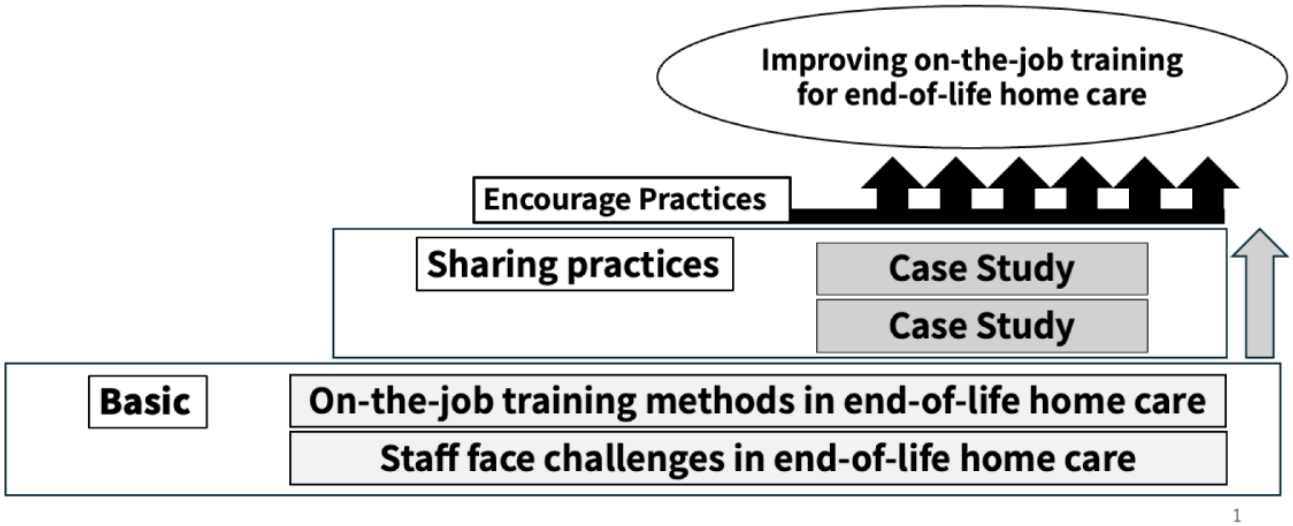

Session 1 is a lecture on foundational principles. We introduce educational methods for staff involved in EOL care, structured within a framework that is grounded in previous research findings and educational theory. Moreover, we highlight the common challenges that staff encounter in providing EOL care. The session concludes with a short group activity based on a real case designed to encourage practical reflection and discussion. Sessions 2 and 3 depict case studies. Providing tailored OJT in EOL home care requires applying practical knowledge to various situations; for this reason, these sessions focus on practical application. Discussing these real-world cases helps participants learn from diverse experiences and acquire the practical knowledge necessary for their practice. To ensure relevance, a real-world case based on their actual educational experiences is gathered beforehand from volunteering participants through an online interview. The case details focus on the educational interactions between the educator, staff, and the client. After gathering a case, we summarize it in a table to facilitate discussion. The content is anonymized, and case details not essential for educational discussion are omitted. The case materials include no identifiable client information; accordingly, client consent was not required, as confirmed by the ethics committee. The anonymized cases are reviewed by five nursing researchers, including an expert in home care nursing and case study methodology, to ensure educational relevance and validity. In the sessions, participants analyze the case scenario, discuss best practices as educators, and connect theoretical knowledge to the practical situation, fostering reflective and collaborative learning. The goal is for participants to improve attitudes and be able to apply the acquired methods in their own practice.

To provide continuous support and promote the application of learning, a LINE group is used between sessions. The group, created immediately after the lecture, serves two main functions. The primary function is to prompt action and reinforce key concepts through regular messages sent by the researchers. These messages start after the first case study and continue weekly at 6 PM on Fridays for 8 weeks. We post brief messages weekly to reinforce key learning points from the program and encourage reflection and peer discussion about OJT for EOL home care. The secondary function is to provide a community for peer support. This platform allows participants to share experiences and mitigate the professional isolation common among educators. To ensure sustainable support, the community remains accessible for participants' use even after the study concludes. Participation in the LINE group is voluntary. Participants may opt out of the online messaging component at any time by leaving the LINE group, and this will not affect their participation in the study or completion of the surveys. Participants may interact by reading messages, posting text or images, replying to posts, and using reactions; passive participation (reading only) is also permitted. If inappropriate content is posted, the research team will request deletion. Messaging data used for research will be deidentified before analysis.

Measures

All measurements will be taken at the individual level using a self-administered survey. T0 is the baseline survey before the program begins. T1 is administered immediately after the lecture (session 1), and T2 immediately after completion of the full program. For the early intervention group, follow-up surveys are administered 3 and 6 months after T2 (T3 and T4, respectively). For the waitlist control group, surveys are administered at the same schedule during the waiting period (T0 to T2) and then immediately after they complete the program and at 3 and 6 months post program (T5, T6, and T7, respectively). Depending on group assignment (early intervention group: T0, T1, T2, T3, T4; waitlist control group: T0, T1, T2, T5, T6, T7), all variables, except covariates, will be measured at 5 or 6 time points. Unless otherwise indicated, the survey items use a 5-point Likert scale, wherein one represents "strongly disagree" and five represents "strongly agree."

Primary Outcome

The primary outcome is participants' attitudes as educators in EOL care, assessed using the End-of-Life Nursing Education Questionnaire (ELNEQ). The ELNEQ consists of five domains: confidence in teaching, motivation for teaching, preparation to provide teaching, preparation to lead initiatives in EOL care, and expected influences on participants. For this study, we selected three domains as the primary focus: confidence in

XSL•FO
RenderX

teaching, motivation for teaching, and expected influences on participants. These domains were chosen because attitudinal changes are key predictors of behavioral change in practice [22]. Higher scores on the ELNEQ indicate more positive attitudes toward EOL care education. The ELNEQ consists of 12 items evaluated on a 5-point Likert scale, where one represents "strongly disagree" and five represents "strongly agree." Sample items include "I am satisfied with my ability to teach EOL care" and "I am motivated to teach EOL care." The reliability of these domains was demonstrated in the original study [23], with Cronbach α values of 0.86 (confidence in teaching), 0.94 (motivation for teaching), and 0.93 (expected influences on participants). The effect sizes of Cohen *d* in the development study of confidence in teaching, motivation for teaching, and expected influences on participants are 0.75, 0.13, and 0.58, respectively [23]. To calculate the sample size, given the feasibility of the study and the expectation that participants will be highly motivated by education, an effect size of 0.58 was used. Finally, the ELNEQ's psychometric properties, including face validity, content validity, construct validity, and reliability, have been established in a prior study, supporting its suitability for this research. In this study, the primary end point is the change in this outcome from baseline (T0) to post intervention (T2).

Secondary Outcomes

The first secondary outcome is the participants' knowledge about the roles and responsibilities of educators in EOL home care. This knowledge encompasses understanding required care practices, the systems and roles of educators, and OJT methods for staff providing EOL home care. A scale was developed based on Shulman's pedagogical content knowledge theory [24]. The theory includes six categories: content knowledge; curriculum knowledge; pedagogical content knowledge; knowledge of learners and their characteristics; knowledge of educational contexts; and knowledge of educational ends, purposes, and values and their philosophical and historical grounds. These categories were adapted to fit the context of EOL home care nursing. The scale consists of six items, evaluated on a 5-point Likert scale, wherein one represents "strongly disagree" and five represents "strongly agree." Sample items include "I have adequate knowledge about required care practices at end-of-life home care as an educator" and "I can clearly describe the system and role of educator for OJT at end-of-life home care." A higher score implies a higher level of knowledge about educators in EOL home care.

The second secondary outcome is the attitude of educators in EOL home care. This scale was developed based on Bandura's theory and previous research [22,25], and assesses confidence, motivation, and outcome expectations. It consists of three items, measured on a 5-point Likert scale, wherein one represents "strongly disagree" and five represents "strongly agree." It includes three items: "I am confident in my ability to teach staff about on-the-job training in end-of-life home care"; "I will actively teach staff who provide end-of-life home care"; and "I believe my education can positively influence staff, enabling them to gain confidence in providing end-of-life home care." A higher score implies a more positive attitude among educators in EOL home care.

The third secondary outcome is the practice of educators in EOL home care. Key components of practice were extracted from a previous study and converted into a measurable scale [26], which includes five items, evaluated on a 5-point Likert scale, wherein one represents "never" and five represents "always." It includes five items, such as "I confirm whether staff have sufficient readiness to provide end-of-life home care" and "I assign clients to staff based on their experience and skill level." A higher score implies more frequent educational practice in EOL home care.

Finally, we will evaluate general educational practice based on Oiwa et al's [27] research. From this study's four domains, we selected factor 3 (attitude and behavioral skills as an educator). This scale includes five items, such as "I create an open environment where learners feel comfortable consulting me" and "I interact with learners in a respectful manner." A higher score implies better overall practice as an educator in a general hospital context.

Covariate

Demographic attributes, including gender, age, and participants' educational background, will be collected in EOL care. Additionally, workplace-related information will be gathered, such as agency size, the number of home deaths managed annually, and whether the agency receives additional payment, such as 24-hour home care nursing support or high-function home care nursing. In Japan, high-function home care nursing is one of the additional payments that can be obtained by agencies that meet numerous requirements, such as the number of deaths in their homes, pediatric clients, and so on.

Participants and Procedure

Recruitment and Randomization

Figure 3 illustrates the flow of participant recruitment and selection in this study. We aim to recruit 125 home care nursing managers and educators. Recruitment will be carried out through online platforms such as NsPace (a community space for home care nurses managed by Teijin, Japan) and our university's laboratory website. Further, snowball sampling will be used by asking participants to share study information with their colleagues who meet the inclusion criteria. Information on this study will include the objectives, procedures, and registration methods.

XSL•FO
RenderX

Figure 3. The study protocol flowchart shows participant recruitment, allocation, and intervention timeline. T0 represents the baseline survey; T1 is administered immediately after the lecture (session 1); T2 is administered immediately after completion of the full program. Follow-up surveys are conducted at 3 and 6 months after T2 for the intervention group (T3 and T4, respectively) and at corresponding intervals after the waitlist control group completes the program (T5 to T7).

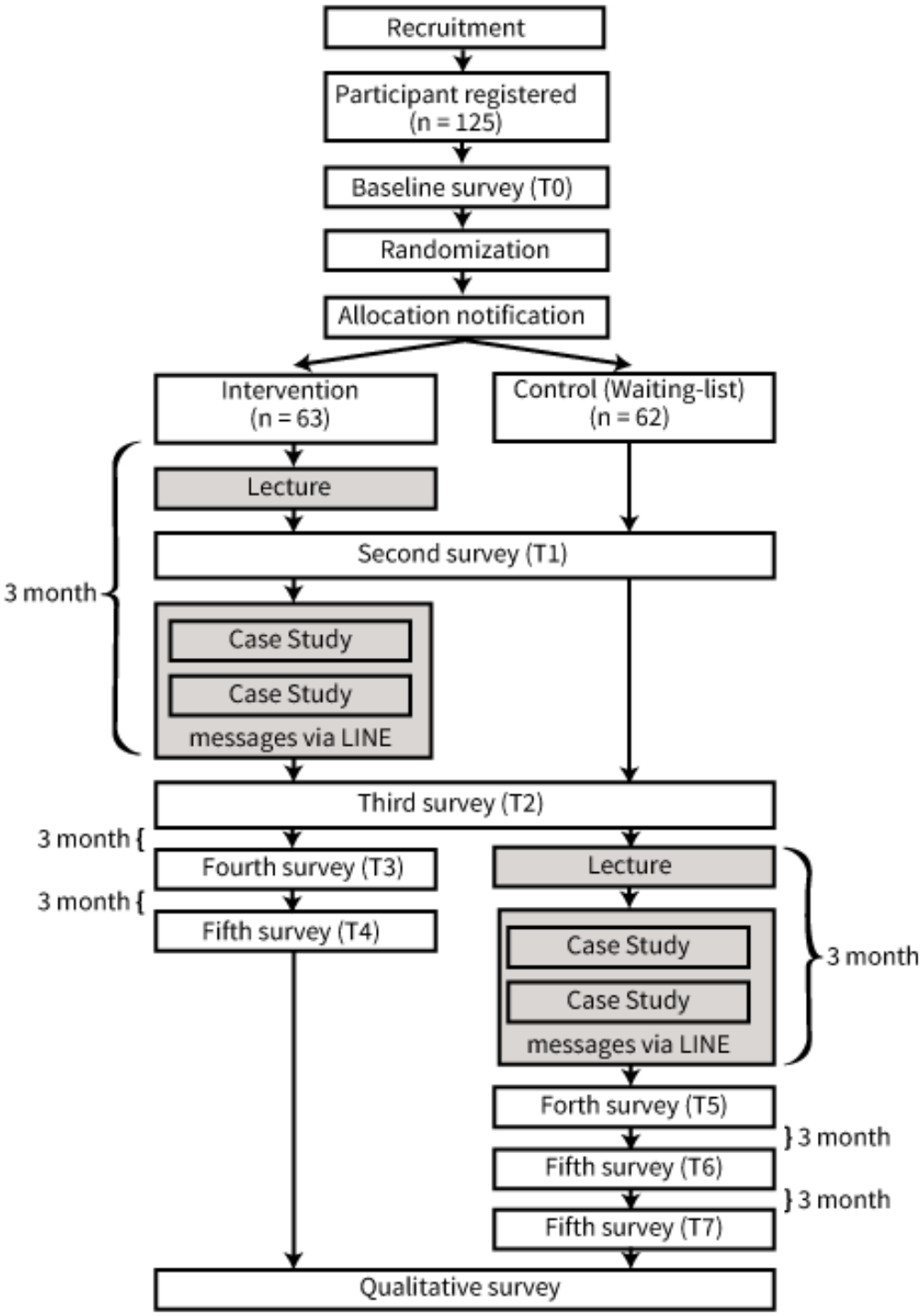

To allocate participants to the intervention and control groups at a 1:1 ratio, restricted randomization with stratification will be performed. Participants will be stratified into three groups based on the number of home deaths managed by their agency between November 29, 2023, and November 29, 2024, as reported at T0: fewer than 5, 5 to 10, and more than 10, following the criteria of previous studies [28]. Participants will be randomized to either course A (earlier program) or course B (later program) using the UMIN Internet Data and Information System for Clinical and Epidemiological Research (Cloud version). The allocation sequence is automatically generated by a centralized system using permuted block randomization with variable block sizes, ensuring appropriate allocation concealment. The lead investigator will assess participant

XSL•FO
RenderX

eligibility, enroll them, and perform the allocation within the system. Allocation concealment is strictly maintained; the assigned group is only revealed by the centralized system after the participant is formally registered and assigned. As this is a waitlist-controlled trial, participants will be informed of their program schedule 14 days before the program begins; thus, this study is conducted in an open-label format without participant blinding. To minimize the risk of performance bias, the intervention will be delivered with fidelity according to preplanned content. In addition, to reduce the risk of reporting and expectation biases in the intervention group, we will use validated and standardized scales for our primary outcomes to ensure measurement consistency.

Baseline data (T0) will be collected immediately after registration. Two weeks before the program begins, participants will be informed of their group schedule (course A or course B). Course A participants will attend the first program session, which consists of a lecture delivered via Zoom. Immediately after the lecture, a follow-up survey (T1) will be administered to all participants. Following the lecture, course A participants will be added to a LINE group and registered with our official LINE account. One month later, a case study session will be held. Additionally, weekly educational messages will be sent every Friday at 6 PM for 8 weeks (a total of 8 messages). After completing all interventions, the third survey (T2) will be conducted immediately after the intervention period. Subsequently, the fourth survey (T3) will be conducted 3 months after T2, while the fifth survey (T4) will be conducted 3 months after T3. Course B participants will continue their routine clinical practice from T0 to T2 and will then participate monthly in the program (the lecture and two case studies). Surveys T5, T6, and T7 will be conducted at corresponding intervals after course B completes the program.

Following quantitative evaluations, a qualitative assessment will be conducted. At T4 (course A) or T7 (course B), participants will be asked whether they can join an interview alongside their staff. We plan to interview approximately 20 educators and 20 staff members. Participants will be recruited by convenience sampling (volunteers). For participants, questions will focus on changes in OJT after the training program and perceived variations in staff performance. The questions from the staff will address the education received from the educators. Interviews will be conducted online via Zoom, last approximately 60 minutes, and will be audio-recorded. To protect confidentiality and encourage feedback, interviews for educators and staff will be conducted in entirely separate sessions. Transcripts will be prepared by the research team and anonymized after transcription. Data will be analyzed using thematic analysis. Initial coding will be conducted by the lead investigator, and emerging codes and themes will be iteratively reviewed and discussed by the research team to enhance analytic rigor and credibility.

Sample Size

The required sample size was calculated to detect participants' change in the primary outcome, specifically participants' attitudes as educators. The required sample size (n=96) was calculated using G*Power (version 3.1; Universität Düsseldorf), assuming a power of 80%, an α level of .05, and an effect size *d* of 0.58. To account for a projected dropout rate of 25%, we determined a final target sample size total of 125 participants.

Data Management and Analysis

All quantitative data will be collected using SurveyMonkey. Data will be downloaded by the lead investigator and stored on a password-protected file within a locked cabinet at the research institution. Identifiable source data (eg, data linkage key files and unencrypted recordings) will be accessible only to the lead investigator; deidentified, analysis-ready datasets and anonymized transcripts will be shared with the research team only after all personal identifiers have been removed. Related to the qualitative data governance, to strictly maintain participant privacy, interview recordings and transcripts for educators and staff will be stored in separate secure digital folders. Regarding digital platform security, Zoom sessions (lecture and case studies) will be recorded and chat logs saved for fidelity assessment; recordings will also be shared with participants who were absent from a session. Zoom data will be stored in the Zoom cloud, with access limited to the lead investigator. LINE group posts (text, images, and stickers) will be analyzed after anonymization; engagement will be summarized as counts of reads and reactions, and will be collected by manual counting rather than export. All study data will be retained for 10 years. As this is a waitlist-controlled trial evaluating an educational program with minimal risk to participants, no formal data monitoring committee was established. The research team, led by the principal investigator, will oversee the trial progress, data integrity, and participant safety. No interim analyses are planned, and no stopping guidelines have been established, given that the intervention carries minimal risk to participants. Any important modifications to the protocol that may impact the conduct of the study, including changes to study objectives, study design, sample sizes, or eligibility criteria, will be reported to the Ethics Committee of the Graduate School of Medicine, The University of Tokyo, and updated in the UMIN Clinical Trials Registry. Statistical analyses will be performed using RStudio (version 2022.12.0+353; Posit PBC). All *P* values will be 2-sided, and a threshold of *P* less than .05 will be considered statistically significant. We will perform an intention-to-treat analysis, including all participants who were randomized as the primary analysis. The primary effect will be estimated using analysis of covariance (ANCOVA) with the outcome at T2 as the dependent variable and group and baseline (T0) as covariates. Missing outcome data will be handled using a complete case approach.

Estimation of a Lecture Effect

As part of the process evaluation, we will evaluate the effect of lectures on the primary and secondary outcomes of T0 (baseline) and T1 (post lecture). Primary and secondary outcomes, which will be continuously distributed and measured at T0 and T1, will be analyzed using ANCOVA; a linear model will be used where each outcome at T1 is the dependent variable and T0 and group are the covariates. If the assumption of homogeneity of regression slopes is violated, an interaction term between the baseline values and group assignment will be tested.

XSL•FO
RenderX

Estimation of Overall Intervention Effect

As this study's primary analysis, we will evaluate the overall intervention effect on the primary and secondary outcomes of T0 (baseline) and T2 (post intervention). Primary and secondary outcomes will be analyzed with ANCOVA using a linear model where each outcome at T2 is the dependent variable and T0 and group are the covariates. If the assumption of homogeneity of regression slopes is violated, an interaction term between the baseline values and group assignment will be tested.

Estimating the Duration of the Overall Intervention Effect

To evaluate the duration of the treatment effect, primary and secondary outcomes will be assessed at T0 and at three postintervention time points for both groups: T2, T3, and T4 for the intervention group, and T5, T6, and T7 for the control group, representing immediately after, 3 months after, and 6 months after the intervention, respectively. We will use repeated measures ANOVA, with time as the within-subject factor and T0 or T2 (baseline) as a covariate. To determine whether the intervention effect persists, diminishes, or increases over time, mean values will be compared between adjacent time points (eg, T2 to T3 and T3 to T4 for the intervention group, and T5 to T6 and T6 to T7 for the control group).

Results

The study protocol was approved by the Ethics Committee of the School of Medicine, The University of Tokyo (2024359NI-(2)) on November 15, 2024. Participant recruitment started on November 29, 2024, and closed on December 25, 2024, with 120 participants enrolled. As of February 15, 2026, the following number of surveys have been completed: 106 at T0 (course A:course B; 52:52); 103 at T1 (52:51); 102 at T2 (50:52); 40 at T3 and 43 at T4 for course A; and 43 at T5, 38 at T6, and 41 at T7 for course B. Quantitative data collection was completed by January 2026, and qualitative interviews will be completed by March 2026. These data demonstrate the feasibility of recruitment and follow-up within the planned timeline. Primary quantitative results and integrated findings are expected to be submitted for publication in 2026-2027 and 2027 - 2028, respectively.

Discussion

This program aims to enhance the ability of home care nursing educators in providing OJT in EOL home care through multiple approaches. Further, we will investigate the staff experience with educational practices conducted by their educator. This study's strength is that it presents a strategy for the dissemination of a required course for home care nursing educators while addressing time and geographic constraints. First, this program mitigates the time constraints of training by conducting a 60-minute session once a month on a Sunday. Moreover, it may be possible for home care nursing educators to take the course in their spare time because home care nursing agencies tend to be closed on Sundays. Second, based on an online training program, we explain how to address geographically challenging situations of home care nursing agencies. Home care nursing agencies in rural areas tend to find it difficult to use off-the-job training. Therefore, providing an online training program may resolve this situation. Additionally, this program will be feasible to deliver and evaluate within this trial. In Japan, there are 15,055 designated home care nursing agencies as of August 2024 [29]. Because home care nursing managers or educators often cover OJT within their agencies, the potential target population could be on the order of one educator per agency (approximately 15,000 nationwide), although this is an estimate. If this trial demonstrates feasibility and effectiveness, the program may support broader dissemination in the future.

In this study, rewards will be provided to reduce the burden on participants. This additional payment may affect changes in outcomes; however, we minimized this risk by providing payment after the survey responses. In addition, all outcome measures in this study are self-reported. Another obstacle is represented by the difficulty in excluding the risk of social desirability, reporting, and expectation biases in the intervention group. To partially address this limitation, we will use validated and standardized scales for our primary outcomes to ensure measurement consistency. In addition, we will conduct interviews with a subset of staff members who received training from the participants 6 months after this program. These staff interviews are intended to supplement the self-reported data by providing external perspectives on behavioral changes.

This RCT will be the first experimental study regarding a training program for educators in EOL home care and will explain a dissemination strategy. If successful, this program can become an evidence-based model that can be used as a framework for educators in other care settings. The diffusion of training programs in the workplace will encourage staff who provide EOL home care to carry out their duties with sufficient confidence in their workplace.

Acknowledgments

The authors would like to thank the participants for their time and contribution to this study.

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT (Generative AI Delegation Taxonomy; 2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing. The GAI tool used was Grammarly. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the outcomes. This declaration was made by RY.

XSL•FO
RenderX

Funding

This study was supported by the Japan Society for the Promotion of Science (23K19798) in September 2023, JST SPRING (JPMJSP2108) in April 2023, and the Yamaji Fumiko Nursing Research Fund in September 2024.

Data Availability

The datasets to support these findings are not publicly available owing to the lack of participant consent. They are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: RY, HN, YS, MN-W, K Kurita, NY-M
Methodology: RY, HN, YS, K Kashiwabara, MN-W, K Kurita, NY-M
Formal analysis: RY, HN, K Kashiwabara
Investigation: RY, HN
Funding acquisition: RY, HN
Project administration: RY
Supervision: NY-M
Writing – original draft: RY, HN, NY-M
Writing – review & editing: RY, HN, YS, K Kashiwabara, MN-W, K Kurita, NY-M
All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. UN, Population Division. World population ageing, 2019: highlights. : United Nations; 2020 URL: https://digitallibrary.un.org/record/3846855?v=pdf [accessed 2026-03-25]
2. Lynn J, Adamson DM. Living well at the end of life: adapting health care to serious chronic illness in old age. : RAND; 2003. [doi: 10.7249/WP137]
3. Costa V, Earle CC, Esplen MJ, et al. The determinants of home and nursing home death: a systematic review and meta-analysis. BMC Palliat Care 2016 Jan 20;15:8. [doi: 10.1186/s12904-016-0077-8] [Medline: 26791258]
4. Brant JM, Fink RM, Thompson C, et al. Global survey of the roles, satisfaction, and barriers of home health care nurses on the provision of palliative care. J Palliat Med 2019 Aug;22(8):945-960. [doi: 10.1089/jpm.2018.0566] [Medline: 31380727]
5. Morioka N, Okubo S, Yumoto Y, Ogata Y. Training opportunities and the increase in the number of nurses in home-visit nursing agencies in Japan: a panel data analysis. BMC Health Serv Res 2019 Jun 20;19(1):398. [doi: 10.1186/s12913-019-4225-8] [Medline: 31221149]
6. Yamada R, Igarashi A, Kashiwabara K, et al. Association between education for home care nurses and their practice of comprehensive client assessment: a cross‐sectional study. Geriatr Gerontol Int 2025 Apr;25(4):572-582. [doi: 10.1111/ggi.70024] [Medline: 40071622]
7. Sherman DW, Matzo ML, Rogers S, McLaughlin M, Virani R. Achieving quality care at the end of life: a focus of the End-of-Life Nursing Education Consortium (ELNEC) curriculum. J Prof Nurs 2002;18(5):255-262. [doi: 10.1053/jpnu.2002.129229] [Medline: 12434318]
8. Hamatani M, Hirahara Y, Onuma E, et al. A randomized controlled trial of the effectiveness of the Program of End-Of-Life Care for Home Visiting Nurses Training (PENUT). J Jpn Acad Nurs Sci 2024;44:218-227. [doi: 10.5630/jans.44.218]
9. Furuse M. Difficulties perceived by visiting nurses who care for terminal cancer patients [Article in Japanese]. J Jpn Soc Cancer Nurs 2013;27(1):61-66. [doi: 10.18906/jjscn.2013-27-1-61]
10. Yamada R, Noguchi-Watanabe M, Numata H, Yamamoto-Mitani N. On-the job training programs for newly graduated and novice home visit nurses in Japan: literature review [Article in Japanese]. J Jpn Acad Nurs Home Care 2022;11:52-61.
11. Sørstrøm AK, Ludvigsen MS, Kymre IG. Facilitating planned home death: a qualitative study on home care nurses' experiences of enablers and barriers. J Adv Nurs 2025 Jan;81(1):340-352. [doi: 10.1111/jan.16171] [Medline: 38515196]
12. Pavloff M, Labrecque ME. Continuing education for home care nurses: an integrative literature review. Home Healthc Now 2021;39(6):310-319. [doi: 10.1097/NHH.0000000000001005] [Medline: 34738966]
13. Rosenfeld P, Pyc LS, Rosati RJ, Marren JM. Developing a competency tool for home health care nurse managers. Home Health Care Manag Pract 2012 Feb;24(1):5-12. [doi: 10.1177/1084822311416499]
14. Matsubara M, Masaki N, Nagao M, et al. Nurse educators' perceptions of the practical competency required by new visiting nurses for home-care visits without assistance and factors influencing their perception formation [Article in Japanese]. Japanese Red Cross Hiroshima Coll Nurs 2019;19:13-22 [FREE Full text]

XSL•FO
RenderX

15. Ethical guidelines for medical and biological research involving human subjects. Ministry of Education, Culture, Sports, Science and Technology. 2021 Mar 23. URL: https://www.mext.go.jp/content/20250325-mxt_life-000035486-01.pdf [accessed 2026-03-26]
16. Hróbjartsson A, Boutron I, Hopewell S, et al. SPIRIT 2025 explanation and elaboration: updated guideline for protocols of randomised trials. BMJ 2025 Apr 28;389(389):e081660. [doi: 10.1136/bmj-2024-081660] [Medline: 40294956]
17. Baldwin TT, Ford JK, Blume BD. Transfer of training 1988-2008: an updated review and agenda for future research. In: Hodgkinson GP, Ford JK, editors. International Review of Industrial and Organizational Psychology: John Wiley & Sons, Ltd; 2012:41-70. [doi: 10.1002/9780470745267]
18. Keller JM. Motivational Design for Learning and Performance: The ARCS Model Approach: Springer; 2010.
19. Merrill MD. First principles of instruction. Educ Technol Res Dev 2002 Sep;50(3):43-59. [doi: 10.1007/BF02505024]
20. Suffoletto B. Deceptively simple yet profoundly impactful: text messaging interventions to support health. J Med Internet Res 2024 Aug 27;26(1):e58726. [doi: 10.2196/58726] [Medline: 39190427]
21. Wells S, Tamir O, Gray J, Naidoo D, Bekhit M, Goldmann D. Are quality improvement collaboratives effective? A systematic review. BMJ Qual Saf 2018 Mar;27(3):226-240. [doi: 10.1136/bmjqs-2017-006926] [Medline: 29055899]
22. Bandura A, National Institute of Mental Health. Social Foundations of Thought and Action: A Social Cognitive Theory: Prentice-Hall, Inc; 1986.
23. Takenouchi S, Miyashita M, Tamura K, Kizawa Y, Kosugi S. Evaluation of the End-of-Life Nursing Education Consortium-Japan Faculty Development Program: validity and reliability of the “End-of-Life Nursing Education Questionnaire”. J Hosp Palliat Nurs 2011 Jan;13(6):376-377. [doi: 10.1097/NJH.0b013e31822da9d7]
24. Shulman LS. Those who understand: knowledge growth in teaching. Educ Res 1986 Feb;15(2):4-14. [doi: 10.3102/0013189X015002004]
25. Sørstrøm AK, Ludvigsen MS, Kymre IG. Home care nurses facilitating planned home deaths. A focused ethnography. BMC Palliat Care 2023 Nov 9;22(1):175. [doi: 10.1186/s12904-023-01303-4] [Medline: 37940911]
26. Numata H. On-the-job training to enable support for end-of-life care at home: a qualitative study of the experiences of novice home care nurses and their mentors [Article in Japanese] [PhD Thesis]. : The University of Tokyo; 2023 URL: https://repository.dl.itc.u-tokyo.ac.jp/records/2011829 [accessed 2025-04-14]
27. Oiwa M, Ohara A, Ouchi C, et al. Development of clinical nursing educators employing an education and teaching skills evaluation index at a hospital organization [Article in Japanese]. Bull Health Sci Kobe 2019;35:35-48.
28. Jona Y, Tanahashi S. Contents and methods of educating home care nurses in home care nursing agencies [Article in Japanese]. J Jpn Nurs Soc 2020;50:71-74 [FREE Full text]
29. Visiting nursing system in Japan: the present and future of visiting nursing. Japan Visiting Nursing Foundation. 2025. URL: https://jvnf-official.g.kuroco-mng.app/files/user/houmon/present_future_visiting_nursing/2025_EN-memo.pdf [accessed 2026-03-13]

Abbreviations

ANCOVA: analysis of covariance
ARCS: attention, relevance, confidence, and satisfaction
ELNEQ: End-of-Life Nursing Education Questionnaire
EOL: end-of-life
OJT: on-the-job training
RCT: randomized controlled trial
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by J Sarvestan; submitted 17.Sep.2025; peer-reviewed by EA Parker, J Steinmiller; revised version received 10.Mar.2026; accepted 12.Mar.2026; published 27.Mar.2026.

Please cite as:
Yamada R, Numata H, Sumikawa Y, Kashiwabara K, Noguchi-Watanabe M, Kurita K, Yamamoto-Mitani N
Development of a Collaborative Education Program for Home Care Nursing Educators' on-the-Job Training in End-of-Life Home Care: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84330
URL: https://www.researchprotocols.org/2026/1/e84330
doi:10.2196/84330

© Ryousuke Yamada, Hanako Numata, Yuka Sumikawa, Kosuke Kashiwabara, Maiko Noguchi-Watanabe, Kayoko Kurita, Noriko Yamamoto-Mitani. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026.

XSL•FO
RenderX

This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The TBI-AD/ADRD Caregiver Support Intervention (TACSI): Protocol of a Pilot Randomized Controlled Trial Evaluation of a Remote Intervention for Family Caregivers

Samantha Ostenso[1], BA; Robyn W Birkeland[1], PhD; Elizabeth Albers[1], MPH; Katie W Louwagie[1], DNP; Allison Iwan[1], BS; Sherry S Chesak[2], MS, PhD, RN; Joseph E Gaugler[1], PhD

[1]Families and Long Term Care, School of Public Health, University of Minnesota, 420 Delaware St SE, D351 Mayo, Minneapolis, MN, United States
[2]School of Nursing, University of Minnesota, Minneapolis, MN, United States

Corresponding Author:
Samantha Ostenso, BA
Families and Long Term Care, School of Public Health, University of Minnesota, 420 Delaware St SE, D351 Mayo, Minneapolis, MN, United States

Abstract

Background: Across the United States, there are millions of informal (ie, unpaid) caregivers helping individuals with Alzheimer Disease/Alzheimer Disease and Related Dementias (AD/ADRD), traumatic brain injury (TBI), or both. TBI is a risk factor for developing, and often co-occurs with, AD/ADRD. In the next decade, more informal caregivers will have to navigate the complexities of the dual diagnosis of TBI and AD/ADRD. Currently, there is a paucity of interventions for caregivers dealing with this dual diagnosis. Our team designed the TBI-AD/ADRD caregiver support intervention (TACSI) as a support program to meet the needs of those providing care to individuals with dual diagnoses of TBI and AD/ADRD.

Objective: This pilot randomized controlled trial (RCT) evaluates the preliminary efficacy and successful implementation of TACSI, a 6-session semistructured psychosocial and psychoeducational program, over a 6-month period.

Methods: This pilot RCT uses an embedded mixed methods randomized controlled evaluation. Caregivers of individuals with a dual diagnosis of AD/ADRD and a history of TBI are randomly assigned to the TACSI (n=41) or a usual care control condition (n=39). Primary outcomes include perceived burden, caregiver relationship satisfaction, and caregiving mastery. Secondary outcomes include caregiver self-efficacy, well-being, and personal resources. Qualitative elements (up to 20 semistructured interviews at the conclusion of the 6-month evaluation and an open-response item at follow-up surveys) allow for better interpretation of quantitative results and understanding of TACSI's mechanisms of benefit.

Results: Recruitment for the study was completed in October 2024. Survey collection and intervention were completed in June 2025. Data analysis will follow.

Conclusions: Having a dual diagnosis of AD/ADRD and a history of TBI is relatively common. However, there are currently no caregiver support interventions specific to this dual diagnosis. It is important to support these caregivers as they encounter the unique challenges of managing TBI and AD/ADRD. This pilot RCT will help determine if the TACSI program can help to increase caregiving skills and strategies and provide support to improve mood and reduce stress for caregivers of individuals with AD/ADRD and a history of TBI, and also determine if it is feasible to proceed to a full-scale RCT.

Trial Registration: Clinical Trials NCT05465109; https://clinicaltrials.gov/study/NCT05465109

International Registered Report Identifier (IRRID): DERR1-10.2196/81125

(JMIR Res Protoc 2026;15:e81125) doi:10.2196/81125

KEYWORDS

traumatic brain injury; dementia; caregiver; mixed methods; support

Introduction

In the United States in 2025, an estimated 7.2 million adults 65 years and older have an Alzheimer Disease/Alzheimer Disease and Related Dementias (AD/ADRD) diagnosis, while each year approximately 2.8 million Americans are diagnosed with a traumatic brain injury (TBI) [1,2]. Each of these conditions can significantly impact one's well-being and ability to complete activities of daily living (ADL) and function independently. Informal (unpaid) caregivers, typically family, often step forward to provide support to these individuals. Over 11 million adults in the United States provide informal care for someone with AD/ADRD [1]. Additionally, about 2.5 million adults serve

XSL•FO
RenderX

as informal caregivers for someone who has experienced a TBI [3].

Research has found that individuals with a history of TBI are more likely to develop AD/ADRD and other neurological disorders than those without TBI [4-7]. In fact, a cohort study following a proportion of the general population for 5 years found that participants with TBI had 1.68 times the risk of developing dementia compared with those without TBI [8]. Given the increased risk of developing AD/ADRD from TBI, it is likely that many caregivers provide support to an individual with a dual diagnosis of AD/ADRD and a history of TBI.

On their own, the care demands associated with AD/ADRD and TBI have negative mental, social, financial, and physical health impacts on caregivers as evidenced in decades of research [9-16]. Together, these conditions can be even more complex and challenging, creating new and unique needs for the caregiver and care recipient. For example, individuals with post-TBI dementia have an increased risk of co-occurring mental health conditions such as mixed affective disorders, bipolar disorder, and post-traumatic stress disorder [17]. They may also experience more agitation, irritability, disinhibition, and depressive symptoms than people with AD/ADRD alone [17]. Individuals with post-TBI dementia also experience more falls and mobility issues than those solely diagnosed with dementia [17] and often require more medications, medical visits, and hospital days than individuals with solely TBI [17]. Unsurprisingly, given these additional stressors, informal caregivers for veterans with the dual diagnoses of TBI and AD/ADRD reported experiencing symptoms of anxiety and depression and felt severely burdened by their duties [17].

Clearly, these unique dual-diagnosis caregivers need support. There are many evidence-based interventions designed to support dementia caregivers, including the Resources for Enhancing Alzheimer's Caregiver Health (REACH) intervention, a structured multicomponent program tailored to caregiver needs that has been shown to increase the quality of life of ethnically diverse dementia caregivers [18,19]. Other multicomponent and psychoeducational interventions for dementia caregivers have also been found to reduce caregiver depressive symptoms and sense of burden and improve their well-being, confidence, and social support [20]. Fewer interventions exist and have been evaluated for their efficacy for TBI caregivers [21-24]. The limited research on these interventions for TBI caregivers suggests that the inclusion of education, skill-building, social support, coping strategies, and resilience holds potential for increasing quality of life and caregiver mastery [22,23,25-27].

Despite these interventions for these individual diagnoses, there is a complete lack of support interventions designed specifically for caregivers of individuals with a dual diagnosis of AD/ADRD and TBI. A 2020 randomized controlled trial (RCT) protocol paper described an intervention at the primary care level for Veteran caregiver dyads with a care recipient who had either TBI or dementia, but not both [28]. To our knowledge, there is only one other intervention aimed at supporting TBI-AD/ADRD dual diagnosis caregivers, and it is also being developed. Researchers in the Veterans Administration are currently evaluating this intervention, REACH Hope (a modification of the REACH program), for AD/ADRD and TBI dual diagnosis caregivers.

We designed the TBI-AD/ADRD caregiver support intervention (TACSI) as an intervention to meet the unique needs of those providing care to individuals with dual diagnoses of a history of TBI and AD/ADRD. TACSI is a psychoeducational and psychosocial support intervention designed to help informal caregivers build caregiving skills and strategies while providing education and support. TACSI is a semistructured program, tailored to meet the needs of individual caregivers as well as the family. The intervention offers one-on-one coaching designed to increase caregiver understanding of TBI-AD/ADRD; provides strategies to improve engagement and communication with the care recipient, family, and health care team; and develops skills to better manage caregiver stress and increase external support.

A 3-month preliminary pilot study established the feasibility, acceptability, and appropriateness of TACSI among 15 caregivers of individuals with TBI and AD/ADRD [29]. Following minor revisions to the program based on the preliminary pilot, the purpose of this study is to deploy a pilot RCT design with 80 caregivers to evaluate the preliminary efficacy and successful implementation of TACSI over a 6-month period.

Methods

Overview

Coinvestigators from Mayo Clinic and the Minneapolis Veterans Affairs Health Care System (MVAHCS) collaborated with the University of Minnesota (UMN) team in designing the TACSI protocol.

The UMN maintains oversight of all study procedures, data collection, data management, and dissemination. In addition to ongoing review of protocol modifications, the UMN Institutional Review Board requires an annual continuing review process, with approval contingent upon review and approval of a study submission regarding participant accrual, withdrawal, and other study-related issues. An independent safety monitor provided additional study oversight, reviewing the annual data and safety monitoring report. Quarterly reports, including enrollment and safety updates, were submitted to the Department of Defense. This protocol was reported following the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) reporting guidelines (Checklist 1) [30].

Ethical Considerations

The UMN Institutional Review Board approved the study protocol for the UMN and Mayo Clinic (STUDY00014886); current version 7, June 4, 2025. The MVAHCS obtained site-specific approval (1641329). All study procedures (consent, data collection, and intervention delivery) were conducted remotely with participating caregivers residing in the United States. Based on the caregiver's preference, consent was completed either electronically via the secure survey platform, Qualtrics, via telephone, or mailed with telephone consent.

XSL•FO
RenderX

All electronic data (eg, administrative tracking logs, survey data, and coaches' notes) from potential and enrolled participants were maintained in secure environments (ie, Qualtrics, secure project folder, and Box). Participant data and their identifying information were kept in separate files. Links to identifiable data will be destroyed 2‐3 years after data collection is complete. All data on the team's computers was encrypted and protected by strong passwords only accessible to the UMN research team. The team used Duo 2-step verification to access VPN/Secure servers. Any hard copy forms were stored in a file cabinet in an office that was locked at all times. Interview audio recordings were deleted from secure recording devices after they were uploaded to Box and a secure server. Transcripts were created by a HIPAA (Health Insurance Portability and Accountability Act)-compliant organization and deidentified.

Upon completion of the study, caregivers were compensated for their participation. Caregivers received US $25 for each completed survey (T0, T1, and T2), for a total of up to US $75. An additional US $25 was provided to those caregivers in the treatment group who completed a final semistructured interview. Separate compensation for completion of the treatment receipt checklist (TRC) was not provided.

Aims

Prior to initiating this pilot study of TACSI, the study team developed and implemented TACSI for 15 family members of persons with a dual diagnosis of TBI & AD/ADRD to assess the feasibility, acceptability, and appropriateness of TACSI over a 3-month period (phase I). The results of the pilot study supported the feasibility, acceptability, and appropriateness of TACSI [29].

In this study (phase II), the specific aims were to use a pilot RCT design with 80 caregivers to evaluate the implementation and preliminary efficacy of a refined version of TACSI over a 6-month period. Specifically, we sought to evaluate the impact of TACSI on primary subjective stressors, including caregiver perceived burden, relationship satisfaction, ideology, and mastery (primary outcomes), primary objective stressors (assistance with ADL/instrumental activities of daily living), and well-being, resource usage, and coping strategies.

Design

This embedded experimental mixed methods study used a pilot RCT design to determine the feasibility and performance of TACSI with 80 caregivers of persons with TBI-ADRD. The mixed methods approach integrated elements of quantitative and qualitative data collection at multiple time points throughout the study for a more robust analysis, enriching our understanding of the caregiver experience in comparison to solely collecting qualitative or quantitative data (Figure 1) [31,32].

Inclusion of qualitative data provides pertinent context for interpretation of quantitative results and rich insight into the intervention's mechanisms of benefit [33,34]. The mixed methods pilot RCT design allowed for the evaluation of appropriateness and sufficiency of the recruitment, screening, randomization, treatment receipt, and retention processes (all appropriate pilot RCT goals over a 6-month follow-up period) [35].

XSL•FO
RenderX

Figure 1. Mixed methods design. TACSI: TBI (traumatic brain injury)-AD/ADRD (Alzheimer disease and related dementias) caregiver support intervention.

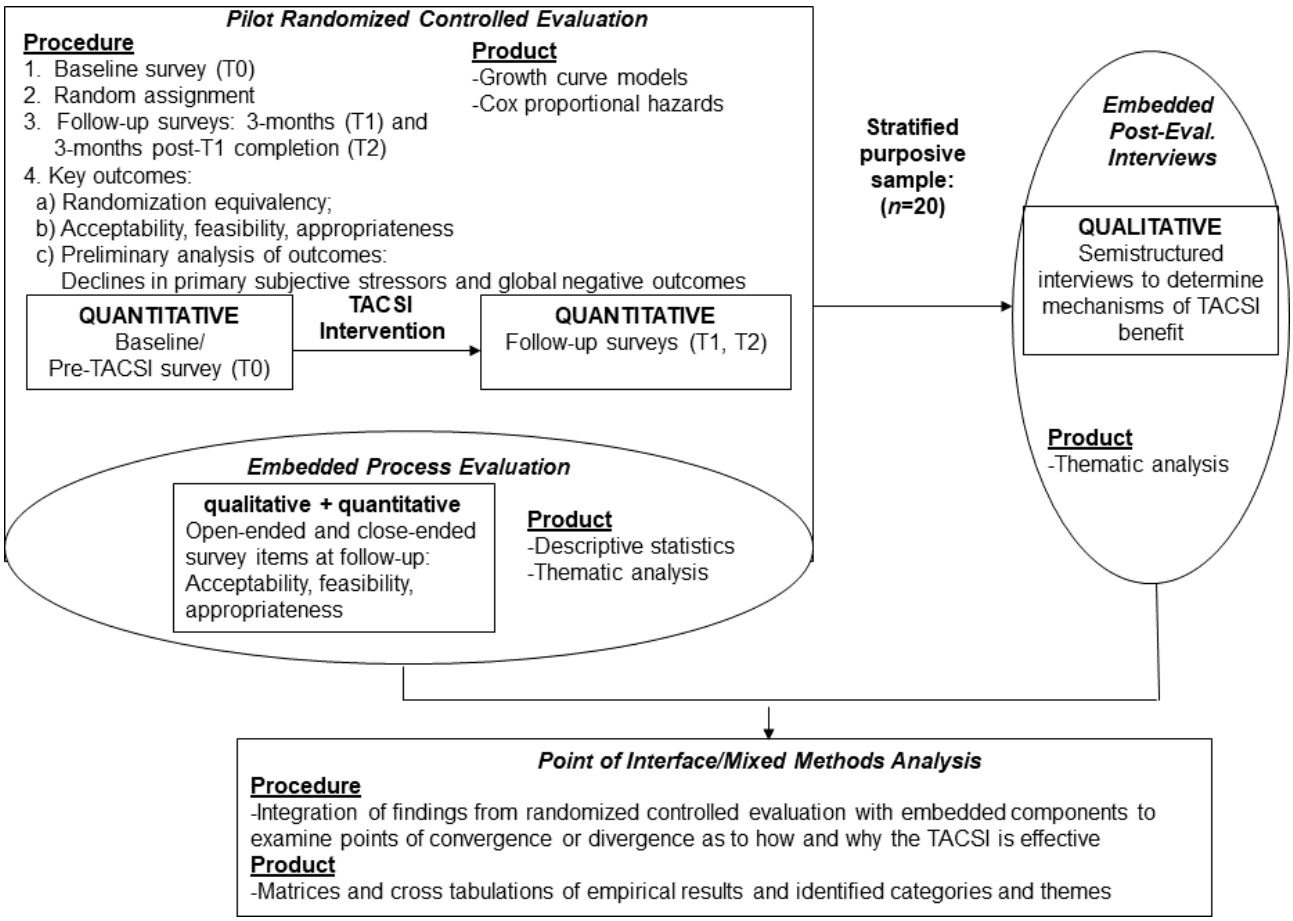

Setting

Study survey administration was offered via web or mail to be completed in hard copy. Two trained coaches from the UMN delivered TACSI to caregivers in the treatment group via secure video conferencing (ie, Zoom) or telephone. Finally, caregivers selected to participate in a semistructured interview could complete the interview by videoconference or telephone.

Participant Eligibility

Eligible caregivers were those who (1) cared for someone with a dual diagnosis of progressive cognitive decline (eg, AD/ADRD) and history of a TBI, (2) provided the most help to the care recipient (or shared the role equally), (3) were willing to participate in the TACSI evaluation, (4) spoke English, (5) were 21 y or older, (6) were not currently participating in other 1:1 psychosocial consultation specific to caregiving, and (7) resided in the United States. Caregivers were excluded if they endorsed a new or worsening mental health condition for which they were not receiving ongoing treatment. If using a psychotropic medication, caregivers needed to have been on a stable dose for at least 3 months prior to study enrollment.

Processes, Interventions, and Procedures

Recruitment

Based on the target population of those caring for patients with a dual diagnosis of TBI and AD/ADRD, the MVAHCS and Mayo Clinic generated internal data requests for patient records using *ICD* (*International Classification of Diseases*) codes. Specific *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) codes included G30, G30.9–Alz Disease, F03-Unspecified dementia, S06.2–Diffuse brain injury, and S06.3–Focal brain injury. Potential participants were listed as caregivers in the patient's medical records.

The 2 health care systems (HCSs) were responsible for making initial contact to describe the study. The UMN study team met weekly with staff from Mayo Clinic and MVAHCS throughout study recruitment. Specifics of the HCS contact protocol slightly varied by site. Figure 2 shows the general flow of contacts at both sites. Patients and caregivers affiliated with Mayo Clinic received initial contacts via email, mail, or both, and follow-up telephone calls from the HCS for those contacts who did not respond. Contacts could opt out of study-related communication (via telephone or web-based form) or indicate study interest by completing a web-based study contact form, calling the UMN team study telephone or HCS, or returning a mailed permission form to the HCS using a pre-paid envelope. For those providing permission, Mayo Clinic shared contact information with the UMN study team.

Patients and caregivers affiliated with the MVAHCS received contacts similarly. The MVAHCS required a letter or encrypted email to be sent prior to study-related follow-up telephone calls by the HCS. Contacts could opt out of further study-related communication. Since the MVAHCS could not share contact information for interested caregivers directly with the UMN

XSL•FO
RenderX

study team, the MVAHCS protocol required interested contacts to alert the study team directly via a web-based form or a call (by the caregiver or by transferring their call to the study line). To facilitate recruitment, staff at either HCS could provide a phone transfer to the UMN study telephone number for interested contacts to discuss screening with the research team.

Figure 2. HCS recruitment diagram. HCS: health care system. *Applies to Mayo Sites Only; ** Applies to VA (Veterans Affairs) site only.

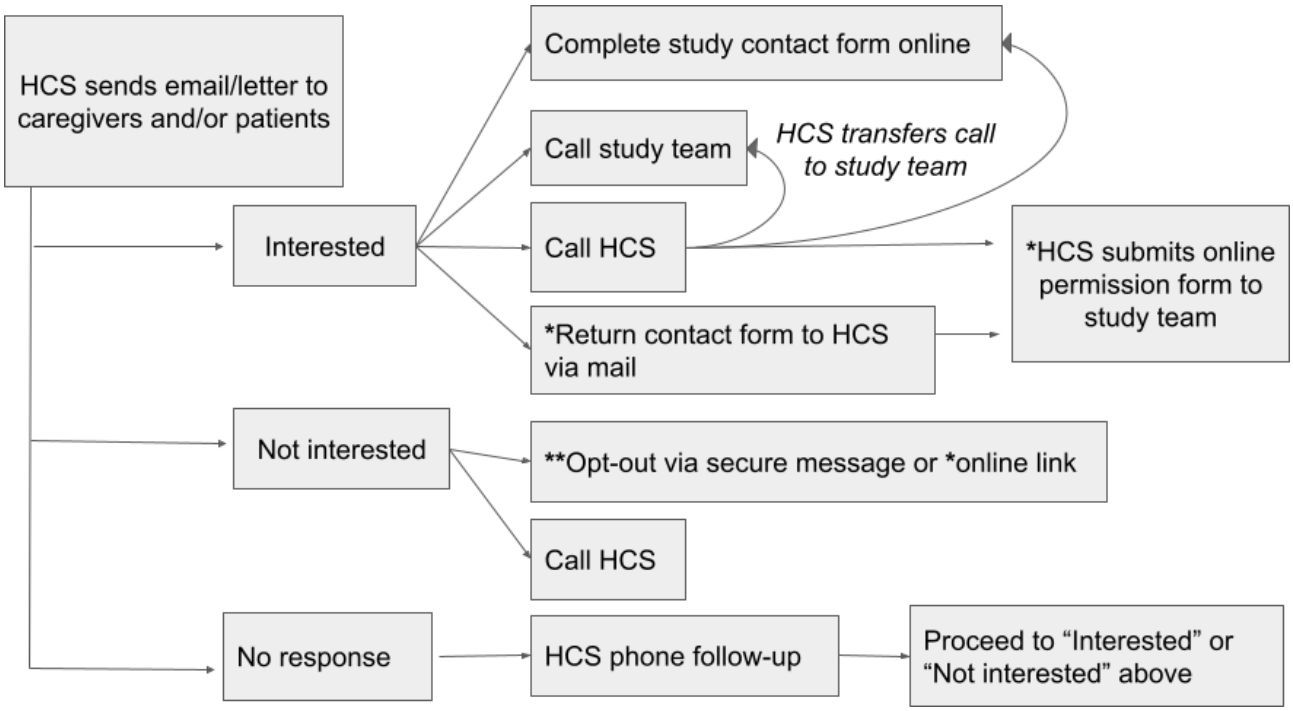

The majority of participants were recruited through the MVAHCS or Mayo Clinic. In addition to HCS recruitment solely based on diagnosis codes, caregivers of someone with a dual diagnosis of TBI and AD/ADRD could express interest in participating in this research study when recruited by MVAHCS or through study postings on ClinicalTrials.gov and StudyFinder. Interested caregivers could call or email the study team directly or complete a web-based permission to contact form to obtain more information about the study.

Each potential participant completed a telephone-based screening procedure to determine eligibility. Only caregivers who were not referred from the MVAHCS or Mayo Clinic were required to answer eligibility questions related to their care recipient's diagnoses to ensure the presence of a dual diagnosis of TBI-AD/ADRD. If eligible, the staff member described the study activities and proceeded to the consent procedures.

Enrollment

Study enrollment began in June 2023. Target enrollment (n=80) was reached in November 2024. Final data collection was completed in June 2025 (Figure 3). Participants were administered consent procedures via their preferred modality: (1) a coordinator read a consent script and obtained caregiver consent verbally via telephone; (2) the caregiver reviewed a web-based consent form and consented via an electronic signature; or (3) the caregiver reviewed a mailed hard copy of the consent form prior to providing telephone consent. Interested caregivers were encouraged to ask questions prior to providing consent for study participation; participants were sent a copy of the completed consent form for their records.

Following consenting procedures, participants were administered a baseline survey; baseline surveys included demographic items for the participating caregiver and the person they care for, as well as a battery of quantitative survey measures. Table 1 displays the SPIRIT timeline [30].

XSL•FO
RenderX

Figure 3. TACSI study flowchart. SSI: semistructured interviews; TACSI: TBI (traumatic brain injury)-AD/ADRD (Alzheimer disease and related dementias) caregiver support intervention.

XSL•FO
RenderX

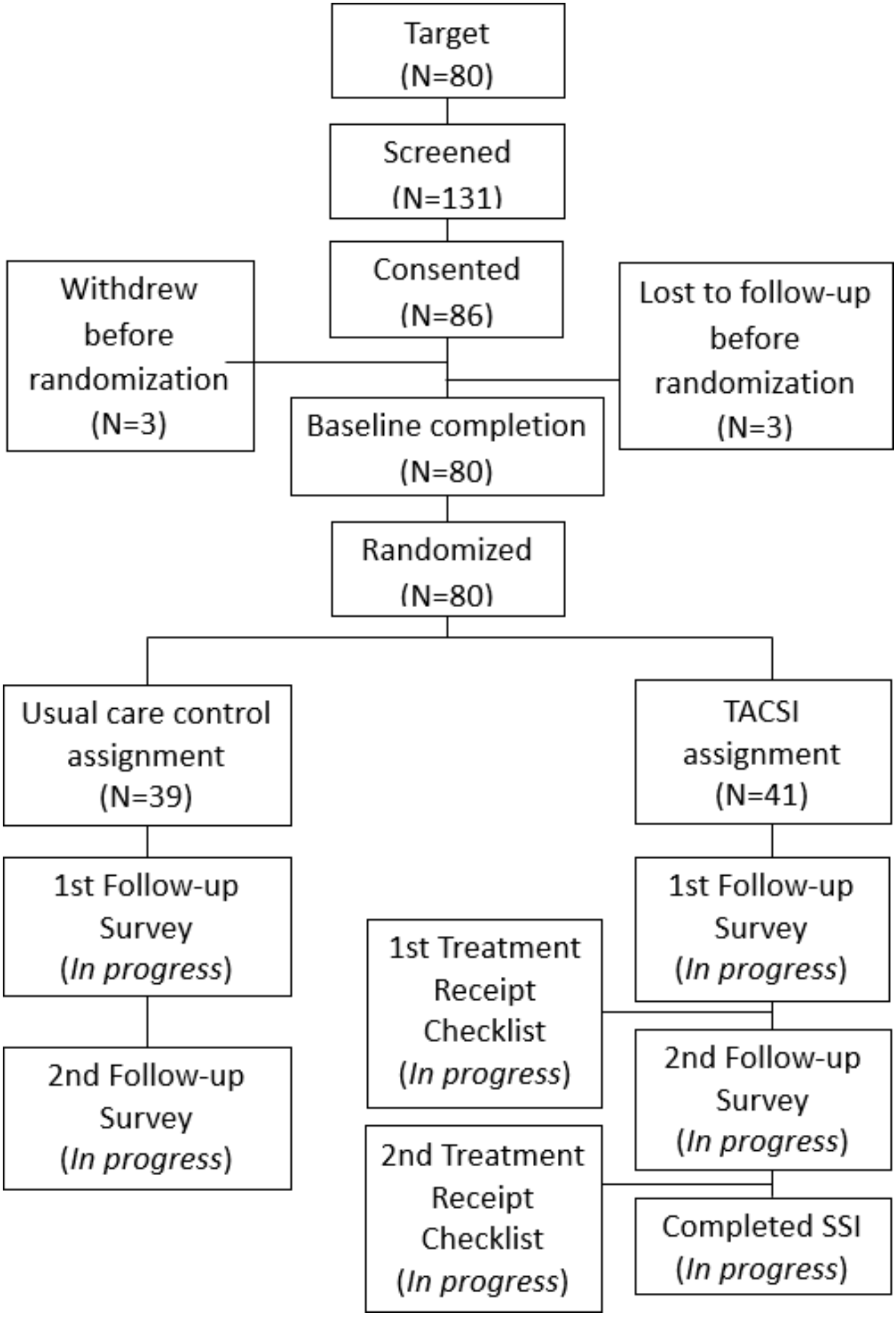

XSL•FO
RenderX

Table . SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) timeline of enrollment, intervention, and measures.

Activity or assessment	Prestudy	Baseline survey (t_0)	Randomization	Follow-up survey (t_1) 3-months post-t_0 completion	Follow-up survey (t_2) 3-months post-t_1 completion	End of study
Eligibility screen	✓					
Informed Consent	✓					
Outcome measures[a]						
Modified Caregiver Appraisal Scale [36] measuring Perceived burden, Caregiver relationship satisfaction, Caregiving ideology, and Caregiving mastery		✓		✓[g]	✓[g]	
Psychological Distress: The Center for Epidemiological Studies-Depression Scale [37]		✓		✓	✓	
TBI-CareQOL[c] Feelings of Loss – Self – Short Form 6a [38]		✓		✓[g]	✓[g]	
TBI-CareQOL Feelings of Loss – Person with TBI – Short Form 6a [38]		✓		✓[g]	✓[g]	
TBI-CareQOL Feeling Trapped – Short Form 6a [38]		✓		✓[g]	✓[g]	
TBI-CareQOL Caregiver-Specific Anxiety – Short Form 6a [38]		✓		✓[g]	✓[g]	
TBI-CareQOL Caregiver Strain – Short Form 6a [38]		✓		✓[g]	✓[g]	
TBI-CareQOL Caregiver Vigilance – Short Form 6a [38]		✓		✓[g]	✓[g]	
TBI-CareQOL Emotional Suppression – Short Form 6a [38]		✓		✓[g]	✓[g]	
TBI-CareQOL Family Disruption – Short Form 3a [38]		✓	✓	✓[g] Omit if bereaved	✓[g] Omit if bereaved	
Health Services Service Use		✓		✓	✓[d] Omit if previously completed post-bereavement	
Community-Based Services Use [39]		✓		✓	✓[d] Omit if previously completed post-bereavement	
COPE [40]		✓		✓	✓	

XSL•FO
RenderX

Activity or assessment	Prestudy	Baseline survey (t_0)	Randomization	Follow-up survey (t_1) 3-months post-t_0 completion	Follow-up survey (t_2) 3-months post-t_1 completion	End of study
Interpersonal Support Evaluation List [41]		✓		✓	✓	
Caregiver Self-Efficacy Scale [42]		✓		✓[g]	✓[g]	
Activities of Daily Living/Instrumental Activities of Daily Living (ADL/IADL) [43,44]		✓		✓	✓[d]	
Revised-Memory and Behavior Checklist [45]		✓		✓[g]	✓[g]	
Intervention						
Caregiver support intervention			✓			
Usual care control assignment			✓			
Treatment Receipt Checklist				Treatment group only	Treatment group only	
Semistructured Interview[e]						Select treatment group participants

[a]More information about the outcome measures can be found in the Multimedia Appendix 1 [34,35,46-53].

[b]Omit if bereaved.

[c]TBI-CareQOL: Traumatic Brain Injury Caregiver Quality of Life.

[d]Omit if previously completed post-bereavement.

[e]Interview questions can be found in the Multimedia Appendix 2.

Randomization Procedures

After baseline survey completion, participating caregivers were randomly assigned on a 1:1 basis to participate in either (1) the TACSI program treatment group or (2) the usual care control group (the latter continued routine care through their HCS). The random assignment schedule was created digitally by a staff member involved in the study solely for this task, via Research Randomizer. Participants had equal chances of being assigned to either the treatment or control group. The assignment schedule was input into a series of sequenced envelopes to be opened in order at the time of randomization (following a participant's baseline survey completion).

The study coordinator informed participants about the results of their randomization to either the TACSI treatment group or the usual care control group via phone, email, or both. Coordinators also alerted the TACSI study coaches, who divided the participant caseload based on workload and availability. Those in the control group were given the opportunity to request resources or referrals during their randomization notification and as a part of the survey completion thank you procedures in order to provide support and bolster retention.

Intervention

Coaching Sessions

Overview

After randomization, the assigned TACSI coach contacted the caregiver assigned to the treatment group to offer introductory information, including a description of the semistructured program (eg, design, length, and frequency of sessions) and to determine caregiver preferences for participation (eg, modality of delivery).

The TACSI program is a 6-session semistructured psychosocial and psychoeducational program, developed to increase caregiving skills and strategies and provide support to improve mood, reduce stress, and increase caregiver self-efficacy. TACSI provides individual coaching and ad hoc support to caregivers. Intervention sessions are conducted on a weekly basis, as schedules permit. Sessions typically last from 45 minutes to an hour and a half, depending on discussion needs and participant preferences. Sessions are held remotely via telephone or secure video conferencing (ie, Zoom) based on participant requests. Table 2 includes descriptions of TACSI's mode of delivery, frequency, and other delivery characteristics adapted from the intervention taxonomy of Schulz et al [46].

XSL•FO
RenderX

The initial TACSI session includes an informal interview to build rapport and allow the coach to learn about the caregiver and care recipient, identifying critical stressors and priorities for the intervention. All subsequent sessions include a brief check-in for caregivers to share updates of their previous week and ascertain the most relevant focus of the session, followed by a discussion of pertinent module content. The conclusion of each session may include discussion of self-care plans for the upcoming week, determining goals or homework, and scheduling the next session, if applicable. The order of delivery of TACSI modules 2 through 5 is intentionally flexible to respond to the priorities and needs of the participant. The sixth session serves to discuss any unexplored intervention content, review previously presented material, and address any remaining caregiver questions.

A TACSI intervention manual provided an overarching structure and guide to facilitate the coach's intervention delivery. The semistructured nature of the program allows coaches the flexibility to tailor the intervention based on the caregiver's needs and priorities. All topics in the manual are offered during the intervention, but the depth of exploration, if any, is determined by the caregiver. Further, tailoring and personalization of the session content and focus are emphasized by the caregiver guiding the selection of topics to discuss and deciding when to discuss them during the intervention.

XSL•FO
RenderX

Table . TBI-AD/ADRD[a] caregiver support intervention (TACSI) delivery characteristics.

Dimension	Options
Method of caregiver/coach contact	• Telephone • Secure video conferencing (ie, Zoom) • Email • Mail
Intervention materials	• Coaches: • Treatment manual • Computer with word processing for session notes • Internet access (for secure video conferencing, ie, Zoom; resource identification/use) • Printer to share resources if needed • Caregivers: • Computer, phone, or both to access remote intervention

Delivery location	• Remote delivery via:
	  • Telephone
	  • Secure video conference calls

Schedule: duration and intensity	• Session types:
	  • Six weekly TACSI sessions (approximately 60 minutes each)
	  • Ad hoc consultation as needed (approximately 45 minutes)
	  • Participants could voluntarily withdraw from intervention at any time.
	• Time needed for documentation (approximately 30 minutes post-session):
	  • Session dates
	  • Session length
	  • Number of ad hoc sessions
	  • Length of ad hoc sessions
	  • Session notes

Scripting/tailoring of intervention	• Semistructured intervention based on treatment manual • General guidelines and some language provided; elaboration and discussion encouraged • Topics specified; coach tailoring of which topics to discuss and when, based on caregiver priorities and needs • Personalization emphasized based on caregiver and care recipient background, experiences, and abilities
Sensitivity to participant characteristics	• Personalization of session content/topics (see Scripting/Tailoring of Intervention) • Language preferences, literacy, visual supplements/augmented communication have not been incorporated or requested by participants

Interventionist qualifications, training, and experience	• Bachelor's degree or higher
	• Experience working with individuals with dementia and their caregivers
	• Training relevant to understanding the treatment manual and ongoing support for accurately implementing the intervention:
	  • Thorough step-by-step review of the treatment manual
	  • Regular coaching meetings (problem-solve questions and concerns, share available resources, etc)
	  • Role play intervention material (Consider having trainees shadow/listen to sessions or mentors observe and provide feedback on trainee's sessions)
	• Competence in remote intervention delivery
	• Ability to build and maintain rapport

XSL•FO
RenderX

Dimension	Options

Adaptability: extent to which intervention can be modified	• What:
	  • Ad-hoc sessions
	  • Session length
	  • Caregiver location
	  • Mode of delivery (telephone vs secure video conference)
	  • Session Content
	• When:
	  • Any time while enrolled in study
	• Why:
	  • Participant request
	  • Participant availability
	  • Participant preference
	  • Coach judgment

Treatment implementation: • Delivery: documentation of interventionist compliance to intended treatment and modifications • Receipt: extent to which processes are implemented by participant and goals are met • Enactment: extent to which knowledge and skills acquired during treatment are applied in real-world settings outside of treatment	Documentation: • Number and duration of sessions • Content delivered Participant Data Collection: • Treatment receipt checklist survey responses • Participant self-report during sessions • Coach appraisal of participant understanding, acquired skills, caregiving self-efficacy, engagement, and use of supports (recorded in coaching notes) • Semistructured interviews conducted with purposively sampled primary caregivers at conclusion of participation
Content and goals	
Treatment content strategies: specific strategies aimed at improving outcomes	• Psychoeducation/didactic instruction • Personalized resource identification and provision • Facilitation of family and social support and involvement • Effective communication and conflict resolution skills, including advocacy • Skill-building techniques and problem-solving techniques • Stress management techniques and relaxation exercises • Processing and emotional support, including for guilt and grief • Active coping strategies
Mechanisms of action: key processes, goals, or mediators of desired treatment outcomes	• Improved knowledge of dementia • Improved engagement with care recipient, family, and health care team • Increased awareness of resources and supports • Improved social and family support • Improved communication and conflict resolution skills • Increased problem-solving skills • Increased prioritization and engagement in self-care and stress management • Increased caregiving self-efficacy • Increased caregiving competence

[a]TBI: traumatic brain injury; AD/ADRD: Alzheimer disease and related dementias.

In each session, coaches aim to establish and maintain a therapeutic rapport with the caregiver while providing a safe, neutral third-party environment to identify, explore, and discuss their stressors. Coaches engage in thoughtful conversation, providing psychoeducation and resources to support each caregiver. They provide individualized constructive guidance and feedback to help caregivers address their unique needs, respond to challenging situations, and use newly learned skills. Coaches also have the flexibility to tailor the length of sessions and to include topics outside of the manual to meet the participant's needs. After each meeting, coaches record session notes, detailing key caregiving issues, relevant family dynamics, the caregiver's reaction to the session, and anticipated content for the subsequent session.

TACSI is a multicomponent program combining psychosocial support with psychoeducational and skills training. Multicomponent interventions have been associated with improving caregiving outcomes [47,48]. TACSI was developed by one of the study coaches (RB) by modifying a well-received and tested intervention, the residential care transition module, created by the UMN team for a previous dementia caregiver

XSL•FO
RenderX

support intervention [49-52]. Moreover, the residential care transition module was modeled after the New York University Caregiver Intervention, which has been found to improve dementia caregivers' well-being, depressive symptoms, and their appraisal of behavioral issues in their care recipients [53-57]. These interventions were grounded in the Stress Process Model [58].

Modifications for TACSI included adding TBI-specific content, such as education on how a TBI may affect the brain and behavior, managing TBI-related concerns, and resources for caring for individuals with TBI. The TACSI program also featured an increased number of approaches for managing negative or anxious thoughts and additional new topics such as how to work effectively with the care recipient's health care team, how to ask for and accept help, and how to navigate changes in partner intimacy. Additionally, coinvestigator JF, an expert on TBI caregiving, reviewed and edited the intervention manual, providing feedback and recommended resources. Table 3 provides an overview of the TACSI program content.

Table . TBI-AD/ADRD[a] caregiver support intervention program content.

Session	Content
Session 1:	• Module 1 (caregiver interview): to explore the caregiver's journey (eg, individuals involved, roles, stressors, strengths/weaknesses of care, etc) to guide the intervention.
Modules 2-5:	• Module 2 (increasing understanding of dementia and traumatic brain injury and building awareness of supports and services): (1) to provide psychoeducation TBO and dementia, as well as impact on daily life (brain, cognition, personality, and behavior), (2) to discuss local/national supports tailored to individual participants' needs, and (3) to introduce/set SMART[b] goals. • Module 3 (supporting emotional health and well-being): (1) to explore experiences of guilt and grief and (2) to increase the understanding of stress and coping strategies. • Module 4 (encouraging realistic expectations and adjusting to lifestyle changes): (1) to explore roles and expectations and (2) to share activity ideas. • Module 5 (effectively working with a health care team/enhancing communication): (1) to review tips for navigating and partnering with the health care team and (2) to discuss conflict resolution.
Session 6:	• Module 6 (review of intervention topics and discussion of wrap-up questions): (1) to wrap up/highlight any intervention content and (2) to review wrap-up questions assessing caregiver adjustment, strengths, and challenges.

[a]TBI: traumatic brain injury; AD/ADRD: Alzheimer disease and related dementias.

[b]SMART: Specific, Measurable, Achievable, Relevant, Time-Bound.

Intervention sessions focus on exploring the participant's caregiving experience, including the supports and strategies they currently use, while determining and subsequently sharing skills, approaches, and new information they likely will benefit from acquiring. A key objective for each session is to build the caregiver's resources in order to alleviate primary subjective stressors, increase caregiver confidence, and enhance awareness and use of informal and formal supports and services. Intervention content and coaching strategies were purposefully selected to positively impact primary (caregiver perceived burden, relationship satisfaction, ideology, and mastery) and secondary outcomes (table of measures is provided in Multimedia Appendix 1 [34,35,46-53]). The following 4 domains serve as the foundation for achieving the 6 objectives and affiliated key activities as detailed in the intervention map (Figure 4).

XSL•FO
RenderX

Figure 4. TBI (traumatic brain injury)-AD/ADRD (Alzheimer disease and related dementias) caregiver support intervention (TACSI) map. AD: Alzheimer disease.

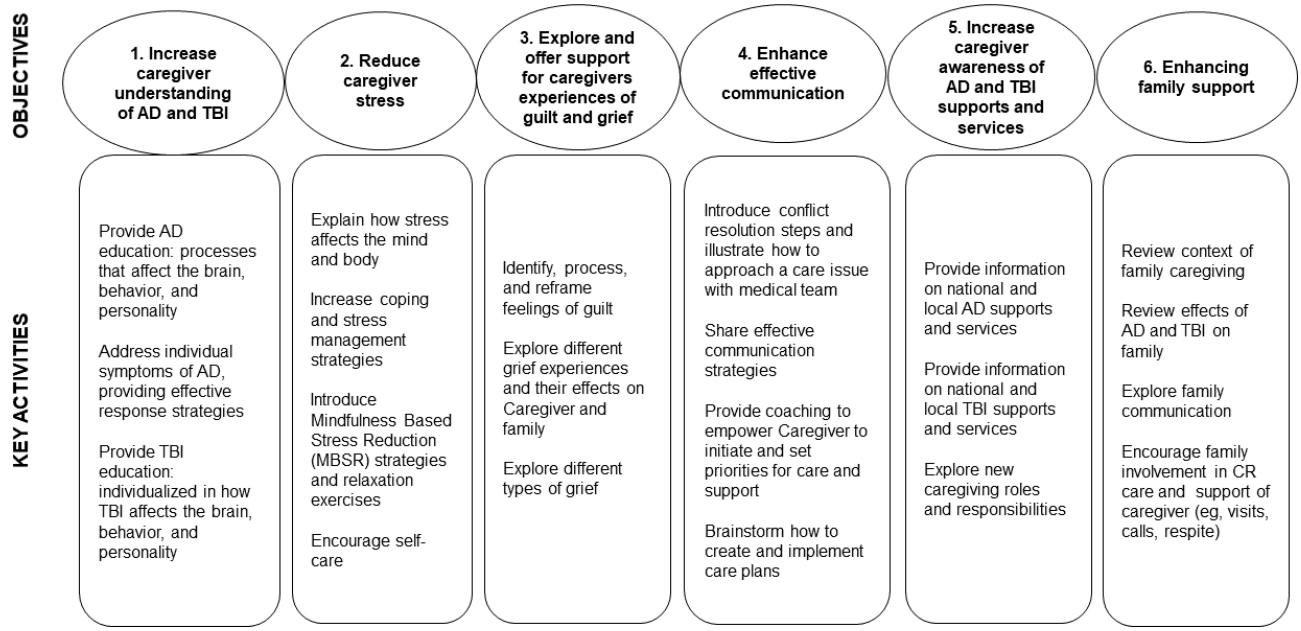

Dementia and TBI Education (Addressing Objective 1)

To increase understanding of AD/ADRD and TBI, coaches share how TBI and dementia affect the brain and how that may impact behavior, personality, and cognition. This education provides an explanation and foundation for the caregiver to understand why the relative may be functioning as they are and how they may be impacted in the future. Learning that their care recipient's behavior is not an intentional choice, but a result of the injury or disease process, can help reduce the annoyance some caregivers may experience. It also helps them to have appropriate expectations and to make plans for the future. In addition, the intervention shares approaches that help to prevent and better manage frustrating behavior (eg, aggression, wandering, hygiene avoidance). Education also focuses on processing and addressing uncomfortable role changes that the caregiver may be experiencing, including adjusting to new responsibilities and changes in spousal intimacy.

Promoting Stress Management and Well-Being (Addressing Objectives 2 and 3)

The intervention aimed to help improve mood and manage stressors. Exploration of the caregiver's experiences with guilt and grief sought to help identify sources of negativity and subsequently, process, validate, normalize, and reframe those negative emotions. Stressors were identified, including those related to caregiving, as well as outside sources of stress. Stress management techniques, including relaxation exercises and problem-solving strategies, such as breaking a larger problem into smaller tasks and setting specific, measurable, achievable, relevant, and time-bound goals, were explored. The caregiver's experience with loneliness and social isolation was explored, along with the importance of self-care and social engagement.

Enhancing Communication (Addressing Objective 4)

Coaches focused on providing a range of communication skills and strategies for caregivers to use with their care recipients, families, health care teams, and others. Through psychoeducation on the tenets of effective communication and the sharing of conflict resolution strategies, the intervention aims to help caregivers convey their perspective, gain understanding and support from others, and advocate for their care recipient. Coaches also offer approaches on how to best engage with a health care team, adapted from the Health Resources and Services Administration's "Working with a Healthcare Team" module [59]. Additionally, coaches share strategies to improve communication and engagement with the care recipient.

Building Awareness of Supports (Addressing Objectives 5 and 6)

Caregivers learn about the local and national supports and services that can meet their unique needs. Provision of these personalized local and national formal supports is complemented by the identification of sources of informal support, such as family, friends, and neighbors. Once these informal sources of support are identified, coaches process caregiver reservations about asking for assistance, provide guidance, and encourage caregivers to ask for and accept offers of help and respite.

Ad Hoc Sessions

Ad hoc sessions, delivered via telephone, secure web conferencing, or email, were available at any time between sessions and following intervention completion. Caregivers could schedule a standing ad hoc session or request impromptu ad hoc coaching. Ad hoc sessions and email consultations were unstructured and designed to address any caregiver concerns in a timely manner.

Interventionist Characteristics

The program is designed to be delivered by a qualified coach based on education or training. Recommended qualifications include: (1) a Bachelor's degree or higher in a health or wellness-related field; (2) relevant experience working with people with memory loss or their caregivers; (3) experience or proficiency in remote intervention delivery, particularly having

XSL•FO
RenderX

the ability to establish rapport without being in person; and (4) completion of TACSI training based on the intervention manual. See Table 2 for additional interventionist characteristics [46].

TACSI coaches were 2 doctorally educated staff (DNP and PhD in clinical psychology). Both have personal and professional experience with individuals with dementia and their caregivers, as well as experience with remote intervention delivery. The coaches worked together previously on ADRD caregiving studies. Similar procedures were used in this study as in the previous intervention studies.

The coach-developed (RB) intervention includes a manual that specifies the procedures for delivering the intervention and details module content. The manual serves as a guide on intervention structure, delivery processes, and session content. Coach training consisted of a thorough step-by-step review of the treatment manual, including discussion of appropriate question prompts, how to start topic discussions, and how to tailor the content to the individual participant. Coaches meet regularly to discuss the intervention, problem-solve any concerns, and share information and newly discovered resources. Coinvestigator JF, a MVAHCS staff psychologist, was present at weekly meetings throughout study recruitment and was available to confer with coaches on any issues as needed.

Data Collection and Measures

Overview

Data collection for both treatment and control groups was conducted via surveys. Additional data were collected from treatment group caregivers via TRCs and semistructured interviews.

Surveys

Caregivers received a baseline survey collecting information on sociodemographics, context of care, and outcome measures. The baseline survey contained sociodemographic measures for both the caregiver and care recipient (date of birth, gender, racial background, marital status, education level, and household income). Context of care variables, such as geographic location, caregiver work status, health conditions, and relationship to care recipient were assessed. Additional variables included year and severity/type of TBI and AD/ADRD diagnoses with concurrent symptoms, care recipient Veteran and Medicaid status, and living arrangement.

Two follow-up surveys were administered by a research assistant blinded to participant group assignment from randomization through the end of analysis. The first follow-up (T1) was administered at 3 months post completion of the baseline survey; the second follow-up was administered 3 months after completion of the 3-month survey (T2). If a care recipient passed away during the study, modified bereavement surveys were administered at follow-up.

Three options for completing surveys were offered: online via Qualtrics, a mailed survey, or over the phone with study staff. The specific measures selected for assessment have strong psychometric properties; Table 1 displays outcome measures administered by time point and survey type. Primary outcome measures assessed perceived burden, caregiver relationship satisfaction, caregiving ideology, and caregiving mastery using the Modified Caregiver Appraisal Scale [36]. Secondary outcome measures evaluated caregiver well-being via the Center for Epidemiological Studies-Depression Scale [37] and multiple TBI caregiver quality of life [38] subscales assessing feelings of loss, feeling trapped, caregiver anxiety, strain, vigilance, emotional suppression, and family disruption. Secondary outcome measures also evaluated caregiver resource usage, including hospital and residential care service use and community-based services such as housekeeping [39], caregiver personal resources (coping style [40], interpersonal supports [41], and self-efficacy [42]), ADLs/instrumental activities of daily living [43,44], and care recipient behaviors [45].

Adverse events were defined prior to the onset of the study, and standardized procedures were determined to address these events, if encountered. At the 3 and 6 month follow-up timepoints, participants were asked: “In the last three months, since your last survey, was there any new or worsening health problem that caused you to be unable to perform your daily routine (ie, not go to work or volunteer), seek medical care (ie, go see your doctor, go to the ER or hospital), or take a new medication?” Those medical events were documented as adverse events. In addition, during each intervention session, coaches inquired into and addressed participant health and other concerns, as applicable.

Staff provided up to 4 reminders to complete each survey via phone or email. If a follow-up survey was not completed, subsequent surveys were still sent unless the participant withdrew from the study. Staff provided a thank-you phone call or email after completion of each survey.

Treatment Receipt Checklist

Within 2 weeks of the completion of their follow-up surveys (T1, T2), treatment group caregivers were sent an additional survey by an unblinded staff member to evaluate treatment receipt via a brief TRC, as well as assess the acceptability, appropriateness, and feasibility of the intervention. Unblinded staff provided reminders to complete the TRCs; thank you emails or calls were completed following survey completion.

The TRC is a 13-item survey evaluating participant receipt of various components of the intervention using 12 Likert scale items (1=strongly disagree to 5=strongly agree, with an option for not applicable/does not apply to me). An open-ended question inquiries about how TACSI was or was not helpful (Wilson et al [29] provides TRC items).

Acceptability, appropriateness, and feasibility of the intervention were evaluated using the acceptability of intervention measure (AIM), the intervention appropriateness measure (IAM), and the feasibility of intervention measure (FIM) [60]. Each measure consists of 4 items using a 5-point Likert scale (1=completely disagree; 5=completely agree). These measures demonstrate strong validity, high internal consistency (AIM, α=.85; IAM, α=.91; and FIM, α=.89), and high test-retest reliability (AIM=0.80, IAM=0.73, FIM=0.88).

XSL•FO
RenderX

Bereavement or Residential Placement

If the care recipient passed away during the study, the participating caregiver was offered condolences by the staff. Pending their willingness to continue the study, these caregivers were administered modified bereavement surveys at follow-up. If the caregiver was in the treatment group, they no longer participated in standard TACSI sessions. However, the interventionist offered optional ad hoc sessions focused on the participant's needs, integrating TACSI content only if relevant. Depending on the number of study time points remaining, either one or two bereavement surveys would be administered, both omitting measures no longer applicable, as shown in Table 1.

In the case of a care recipient's nursing home admission, the caregiver could continue study involvement (TACSI intervention and routine follow-up surveys) as usual.

Semistructured Interviews

Twenty participants in the treatment group were purposely selected to participate in an audio-recorded semistructured interview to provide qualitative data regarding the TACSI program. Staff selected participants who rated TACSI with higher and lower average scores on the TRC and measures assessing the intervention's appropriateness, acceptability, and feasibility (see above). Additionally, as the sample allowed, the study team made an effort to select participants with varied characteristics: relationship to care recipient (partner, adult child), severity of condition (TBI-AD/ADRD), and racial/ethnic identity. Participants could complete the semistructured interview via phone or secure video conferencing (Zoom). Multimedia Appendix 2 contains the interview guide.

Coach Documentation

Throughout each caregiver's time in the study, coaches documented any participant contact for treatment and control participants: dates, method of contact, length of contact, and general topics of discussion. Session notes were also completed by coaches, documenting key issues, family dynamics, participant response or engagement with session content, and plans for the next session's content.

Analysis

Quantitative analysis will include descriptive statistics, such as frequency tables and means, to describe sample characteristics and TBI-AD/ADRD caregiver outcomes. Enrollment and recruitment data, such as the number of contacts or referrals by site, reasons for exclusion, and retention, will be summarized. Distribution/normality and missing data will also be examined.

The number of TBI-AD/ADRD caregivers enrolled to address the study hypotheses (n=80) was determined using power analysis procedures that relied on Wänström's Monte Carlo simulations [61]. A final sample of 72 TBI-AD/ADRD caregivers was found sufficient for detecting a medium effect size of 0.5 at a statistical power level of 0.8, while adjusting for 10% loss to follow-up and for multiple potential statistical differences between treatment and control groups at baseline. In order to evaluate whether a full-scale randomized controlled design is feasible and appropriate, we will conduct paired *t* tests to determine whether the TACSI treatment and control groups are statistically ($P>.05$) equivalent across all baseline outcomes and TBI-AD/ADRD caregiver covariates. Multilevel regression analyses of outcome and growth curve modeling will be used to compare rates of change between the treatment and control groups in primary subjective stressors and other caregiver outcomes. The primary independent variable will indicate random assignment into the TACSI treatment or control group. If any sample characteristics or outcomes vary significantly at baseline between the treatment and control groups, the baseline values will be included as covariates. In order for the intercept effect to be a main effect and to have a treatment*time interaction effect, the baseline value of each outcome will be a covariate, and time will be centered at 3 months after the baseline survey. We will explore whether treatment participants experienced statistically significant ($P<.05$) declines in primary subjective stress, increased well-being, and other changes in TBI-AD/ADRD caregiver outcomes when compared with controls over the study period. Further, Cox proportional hazard models will be run with the events including admission to a nursing home, assisted living, similar residential living situation, and emergency department and overnight hospital stays.

We will calculate means and create frequency tables to evaluate AIM, IAM, and FIM measures and TRC at T1 and T2. We will determine if 80% of treatment caregivers found the intervention appropriate, feasible, and acceptable at the 2 follow-up timepoints (T1 and T2).

Post hoc analyses will also be conducted. Specifically, to examine the effects of TACSI use on changes in primary subjective stress and caregiver well-being, intervention session data regarding variations in use of TACSI, such as frequency and duration of counseling sessions, will also be included in the growth curve models.

The qualitative analysis will use reflexive thematic content analytic techniques as described by Morse, Moser, and others [62-66]. The semistructured interviews and open-ended comments from the TRCs will comprise the qualitative data. The research staff will read the qualitative data to familiarize themselves with the data and make changes to the Phase I codebook accordingly. Staff will review and discuss coding assumptions before making any revisions to the codebook. This qualitative analysis will help determine why TACSI did or did not meet the needs of TBI-AD/ADRD caregivers.

The mixed methods analysis aims to provide empirical context for the participants' perceptions of TACSI. A sequential mixed methods sampling approach was used in selecting treatment participants for the semistructured interviews. This will enable research staff to better understand participants' perceptions of TACSI's usefulness while also linking qualitatively derived themes with the outcome data. Specifically, the empirical outcome data will be sorted by qualitative themes that appeared across the 2 subgroups of the semistructured interview sample (those who rated TACSI as highly acceptable, feasible, and useful in the TRCs vs those who rated TACSI poorly on the same measures). This mixed methods approach will allow us to determine if certain themes appear more often for one subgroup than the other subgroup, or if the subgroups use TACSI recommendations differently.

XSL•FO
RenderX

Data Sharing

To facilitate data preservation and sharing following project completion, the study protocol includes a plan to deposit quantitative survey study data into a centralized repository. Pending regulatory recommendations and availability at the time of post-study submission, the team has focused data preparation efforts on procedures developed by the Federal Interagency Traumatic Brain Injury Research Informatics System, developed by the National Institute of Health in collaboration with the Department of Defense. Alternatively, the Data Repository for University of Minnesota (a data repository located at UMN) or a similar resource will be used.

Dissemination Plan

Results of this study will be shared with participants via a one-page summary and posted to the team's website. The team will publish the results in academic journals and potentially present at gerontological conferences and community sites, including at TBI and dementia-focused organizations.

Results

Figure 3 depicts participants' flow through the study. All participants have been recruited. Enrollment of 80 participants began in June 2023. Survey data collection completed in late June 2025. Figure 3 shows the general flow of participants throughout the study.

Discussion

Within the dementia care intervention research literature, care recipient condition (eg, "Alzheimer's disease or a related dementia") is often conceptualized as a solitary challenge influencing care recipient symptoms, care demands, and ultimately, caregiver distress and well-being. However, persons living with ADRD often experience other serious conditions either concurrently (eg, "mixed" dementia) or in instances where another condition's onset may have predated the onset of ADRD, such as TBI [1,67,68]. With the potential complexities for caregivers in managing symptoms and care demands in these situations, we developed TACSI as a potential support resource for family members and other unpaid individuals to assist in navigating the TBI/ADRD care context.

With the exception of concurrent efforts to adapt the evidence-based REACH intervention in the Veterans Administration HCS, we know of no other caregiver intervention model that is tailored to the needs of caregivers of individuals with TBI/ADRD. Given how TBI is a consistent risk factor for ADRD (see above), creating a psychosocial and psychoeducational program that targets increased understanding, stress reduction, guilt and grief management, effective communication, and awareness of support locally and within the family may help to address an important gap in care for people living with TBI/ADRD and those who support them.

This pilot randomized controlled evaluation, along with the mixed methods experimental design, will yield important insights as to the readiness of TACSI to proceed to a larger-scale randomized controlled evaluation to demonstrate efficacy, as well as how and why the intervention is beneficial to TBI/ADRD caregivers [30,34]. The latter is particularly crucial in the developmental life cycle of a behavioral intervention, as the use of qualitative data to identify mechanisms of action/benefit may help to guide their measurement and testing in subsequent efficacy trials and help to establish the essential aspects of an intervention that must be maintained as it progresses from efficacy to later effectiveness testing and, ultimately, successful dissemination/implementation.

Across chronic disease contexts, and particularly in TBI and ADRD, the ramifications of the disease trajectory extend far beyond the person living with one or both of these conditions. To achieve optimal care for persons with TBI/ADRD, a socioecological perspective is warranted, where medical and psychosocial support target both the person living with these conditions and those involved in their care, most often family members. The TACSI program adopts such a perspective in its focus, intervention targets, and outcomes. The use of a mixed methods experimental intervention design will further facilitate the: 1) piloting of the RCT design to establish its feasibility; and 2) identification of how and why the intervention yields benefits for caregivers of individuals with TBI/ADRD, establishing a strong scientific foundation from which to proceed to a larger scale RCT design to demonstrate efficacy.

When considering the design, testing, and (it is hoped) successful dissemination and implementation of TACSI, several key issues are essential to note. One aspect of TACSI that helps to facilitate its translation and dissemination is its remote/telehealth delivery. A potential implementation barrier will be to scale the program so that interventionists can deliver the protocol and not require advanced professional training and credentials (as the current TACSI interventionists hold). Moreover, if the latter translational objective is achieved, the development of an asynchronous training protocol so that interventionists can easily obtain the skills and certification needed to deliver TACSI will require development and refinement in order for the intervention to achieve its full implementation/dissemination potential. Another key aspect of our design that will facilitate subsequent efficacy and effectiveness testing is the use of mixed methods to identify mechanisms of action (in addition to the ones already hypothesized as part of the TACSI via the Stress Process Model; please see above) that will underscore the essential elements of the intervention that should be maintained if/when the intervention is adapted for effectiveness testing and dissemination/implementation in heterogeneous clinical and community settings.

The prevalence of mixed dementia, where persons diagnosed with Alzheimer disease are found to have multiple causes of dementia (eg, Alzheimer disease and TBI), is present in half or more of individuals living with ADRD [1,67,68]. Given the long-established challenges of dementia care for unpaid family members, friends, and others, there is a pressing need to develop support programs and resources for caregivers of persons with potentially concurrent causes of dementia. It is hoped that TACSI can fill this gap for caregivers of persons with TBI/ADRD.

XSL•FO
RenderX

Funding

This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Peer-Reviewed Alzheimer Research Program (award W81XWH2110726).

Disclaimer

The views, opinions, and findings contained in this paper are those of the authors and should not be construed as an official Department of Defense, US Department of Veterans Affairs, or other federal agency position, policy, or decision unless so designated by other official documentation.

Data Availability

Pending regulatory recommendations and availability at the time of poststudy submission, the team has focused data preparation efforts on procedures developed by the Federal Interagency Traumatic Brain Injury Research (FITBIR) Informatics System, developed by the National Institute of Health (NIH) in collaboration with the Department of Defense. Alternatively, the Data Repository for the University of Minnesota (DRUM; a data repository located at UMN) or a similar resource will be used.

Authors' Contributions

JG conceived and planned the study. RB developed the intervention. KL and RB delivered the intervention. SO, AI, EA, RB, and KL administer the study, and AI conducts semistructured interviews. EA is responsible for data analysis and interpretation of the results. SO took the lead in writing the manuscript. All authors (SO, AI, EA, RB, KL, JG, and SC) contributed to writing the manuscript. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey measures.
[DOCX File, 20 KB - resprot_v15i1e81125_app1.docx]

Multimedia Appendix 2
Semistructured interview guide.
[DOCX File, 16 KB - resprot_v15i1e81125_app2.docx]

Checklist 1
SPIRIT 2025 checklist.
[PDF File, 102 KB - resprot_v15i1e81125_app3.pdf]

References

1. 2025 Alzheimer's disease facts and figures. Alzheimers Dement 2025 Apr;21(4):e70235. [doi: 10.1002/alz.70235]
2. Brain injury association of America. brain injury community pushes for recognition of brain injury as a chronic health condition. Brain Injury Association of America. 2024 Mar 29. URL: https://biausa.org/public-affairs/public-awareness/news/brain-injury-community-pushes-for-recognition-of-brain-injury-as-a-chronic-health-condition [accessed 2025-05-27]
3. Malec JF, Van Houtven CH, Tanielian T, Atizado A, Dorn MC. Impact of TBI on caregivers of veterans with TBI: burden and interventions. Brain Inj 2017;31(9):1235-1245. [doi: 10.1080/02699052.2016.1274778] [Medline: 28981343]
4. Elder GA, Ehrlich ME, Gandy S. Relationship of traumatic brain injury to chronic mental health problems and dementia in military veterans. Neurosci Lett 2019 Aug 10;707:134294. [doi: 10.1016/j.neulet.2019.134294] [Medline: 31141716]
5. Schaffert J, LoBue C, White CL, et al. Traumatic brain injury history is associated with an earlier age of dementia onset in autopsy-confirmed Alzheimer's disease. Neuropsychology 2018 May;32(4):410-416. [doi: 10.1037/neu0000423] [Medline: 29389151]
6. Mendez MF. What is the relationship of traumatic brain injury to dementia? J Alzheimers Dis 2017;57(3):667-681. [doi: 10.3233/JAD-161002] [Medline: 28269777]
7. Barnes DE, Kaup A, Kirby KA, Byers AL, Diaz-Arrastia R, Yaffe K. Traumatic brain injury and risk of dementia in older veterans. Neurology (ECronicon) 2014 Jul 22;83(4):312-319. [doi: 10.1212/WNL.0000000000000616] [Medline: 24966406]
8. Wang HK, Lin SH, Sung PS, et al. Population based study on patients with traumatic brain injury suggests increased risk of dementia. J Neurol Neurosurg Psychiatry 2012 Nov;83(11):1080-1085. [doi: 10.1136/jnnp-2012-302633] [Medline: 22842203]

XSL•FO
RenderX

9. Brodaty H, Woodward M, Boundy K, Ames D, Balshaw R, PRIME Study Group. Prevalence and predictors of burden in caregivers of people with dementia. Am J Geriatr Psychiatry 2014 Aug;22(8):756-765. [doi: 10.1016/j.jagp.2013.05.004] [Medline: 24012226]
10. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
11. Schulz R, Beach SR, Czaja SJ, Martire LM, Monin JK. Family caregiving for older adults. Annu Rev Psychol 2020 Jan 4;71(1):635-659. [doi: 10.1146/annurev-psych-010419-050754]
12. Schulz R, Sherwood PR. Physical and mental health effects of family caregiving. Am J Nurs 2008 Sep;108(9 Suppl):23-27. [doi: 10.1097/01.NAJ.0000336406.45248.4c] [Medline: 18797217]
13. Eden J. Committee on family caregiving for older adults, board on health care services, health and medicine division, national academies of sciences, engineering, and medicine. In: Schulz R, editor. Families Caring for an Aging America: National Academies Press (US); 2016. URL: http://www.ncbi.nlm.nih.gov/books/NBK396401 [accessed 2026-03-04] [Medline: 27905704]
14. Mitchell LA, Hirdes J, Poss JW, Slegers-Boyd C, Caldarelli H, Martin L. Informal caregivers of clients with neurological conditions: profiles, patterns and risk factors for distress from a home care prevalence study. BMC Health Serv Res 2015 Aug 28;15(1):350. [doi: 10.1186/s12913-015-1010-1] [Medline: 26315104]
15. Kjeldgaard A, Soendergaard PL, Wolffbrandt MM, Norup A. Predictors of caregiver burden in caregivers of individuals with traumatic or non-traumatic brain injury: a scoping review. NeuroRehabilitation 2023;52(1):9-28. [doi: 10.3233/NRE-220134] [Medline: 36617762]
16. Griffin JM, Lee MK, Bangerter LR, et al. Burden and mental health among caregivers of veterans with traumatic brain injury/polytrauma. Am J Orthopsychiatry 2017(2):139-148. [doi: 10.1037/ort00002074]
17. Nichols LO, Martindale-Adams J, Seel RT, Zuber JK, Perrin PB. Demographics, clinical characteristics, and well-being of veterans with TBI and dementia and their caregivers. Geriatrics (Basel) 2024 Oct 8;9(5):130. [doi: 10.3390/geriatrics9050130] [Medline: 39451862]
18. Belle SH, Burgio L, Burns R, et al. Enhancing the quality of life of dementia caregivers from different ethnic or racial groups: a randomized, controlled trial. Ann Intern Med 2006 Nov 21;145(10):727-738. [doi: 10.7326/0003-4819-145-10-200611210-00005] [Medline: 17116917]
19. Cheng ST, Au A, Losada A, Thompson LW, Gallagher-Thompson D. Psychological interventions for dementia caregivers: what we have achieved, what we have learned. Curr Psychiatry Rep 2019 Jun 6;21(7):59. [doi: 10.1007/s11920-019-1045-9] [Medline: 31172302]
20. Goy E, Kansagara D, Freeman M. A Systematic Evidence Review of Interventions for Non-Professional Caregivers of Individuals with Dementia: Department of Veterans Affairs (US); 2010. URL: https://www.ncbi.nlm.nih.gov/books/NBK49194/ [accessed 2026-03-04] [Medline: 21155197]
21. Kreitzer N, Kurowski BG, Bakas T. Systematic review of caregiver and dyad interventions after adult traumatic brain injury. Arch Phys Med Rehabil 2018 Nov;99(11):2342-2354. [doi: 10.1016/j.apmr.2018.04.016] [Medline: 29752909]
22. Backhaus SL, Ibarra SL, Klyce D, Trexler LE, Malec JF. Brain injury coping skills group: a preventative intervention for patients with brain injury and their caregivers. Arch Phys Med Rehabil 2010 Jun;91(6):840-848. [doi: 10.1016/j.apmr.2010.03.015]
23. Kreutzer JS, Marwitz JH, Sima AP, Godwin EE. Efficacy of the brain injury family intervention: impact on family members. J Head Trauma Rehabil 2015;30(4):249-260. [doi: 10.1097/HTR.0000000000000144] [Medline: 26147315]
24. Shepherd-Banigan ME. Interventions to Support Caregivers or Families of Patients with TBI, PTSD, or Polytrauma: A Systematic Review: Department of Veterans Affairs (US); 2018. URL: http://www.ncbi.nlm.nih.gov/books/NBK513786 [accessed 2026-03-04] [Medline: 30063308]
25. Hanks RA, Rapport LJ, Vangel S. Caregiving appraisal after traumatic brain injury: the effects of functional status, coping style, social support and family functioning. NeuroRehabilitation 2007;22(1):43-52. [Medline: 17379947]
26. Moriarty H, Winter L, Robinson K, et al. A randomized controlled trial to evaluate the veterans' in-home program for military veterans with traumatic brain injury and their families: report on impact for family members. PM R 2016 Jun;8(6):495-509. [doi: 10.1016/j.pmrj.2015.10.008] [Medline: 26514790]
27. Damianakis T, Tough A, Marziali E, Dawson DR. Therapy online: a web-based video support group for family caregivers of survivors with traumatic brain injury. J Head Trauma Rehabil 2016;31(4):E12-E20. [doi: 10.1097/HTR.0000000000000178] [Medline: 26291634]
28. Carnahan JL, Judge KS, Daggy JK, et al. Supporting caregivers of veterans with Alzheimer's disease and traumatic brain injury: study protocol for a randomized controlled trial. Trials 2020 Apr 19;21(1):340. [doi: 10.1186/s13063-020-4199-1] [Medline: 32306982]
29. Wilson M, Birkeland RW, Albers E, et al. The traumatic brain injury-Alzheimer's disease and Alzheimer's disease-related dementia caregiver support intervention: a mixed methods evaluation of program feasibility, acceptability, and utility. Innov Aging 2025;9(6):igaf057. [doi: 10.1093/geroni/igaf057] [Medline: 40709322]
30. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ 2025 Apr 28;389:e081477. [doi: 10.1136/bmj-2024-081477] [Medline: 40294953]

XSL•FO
RenderX

31. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research: SAGE; 2011.
32. Hanson WE, Creswell JW, Clark VLP, Petska KS, Creswell JD. Mixed methods research designs in counseling psychology. J Couns Psychol 2005;52(2):224-235. [doi: 10.1037/0022-0167.52.2.224]
33. Creswell JW, Fetters MD, Plano Clark VL, Morales A. Mixed methods intervention trials. In: Andrew S, Halcomb EJ, editors. Mixed Methods Research for Nursing and the Health Sciences: Wiley-Blackwell; 2009:159-180. [doi: 10.1002/9781444316490]
34. Gallo JJ, Lee SY. Mixed methods in behavioral intervention research. In: Behavioral Intervention Research: Designing, Evaluating, and Implementing: Springer; 2016:195-212.
35. Abbott JH. The distinction between randomized clinical trials (RCTs) and preliminary feasibility and pilot studies: what they are and are not. J Orthop Sports Phys Ther 2014 Aug;44(8):555-558. [doi: 10.2519/jospt.2014.0110] [Medline: 25082389]
36. Struchen MA, Atchison TB, Roebuck TM, Caroselli JS, Sander AM. A multidimensional measure of caregiving appraisal: validation of the caregiver appraisal scale in traumatic brain injury. J Head Trauma Rehabil 2002 Apr;17(2):132-154. [doi: 10.1097/00001199-200204000-00005] [Medline: 11909511]
37. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1:385-401. [doi: 10.1177/014662167700100306] [Medline: 23469262]
38. Carlozzi NE, Kallen MA, Hanks R, et al. The TBI-CareQOL measurement system: development and preliminary validation of health-related quality of life measures for caregivers of civilians and service members/veterans with traumatic brain injury. Arch Phys Med Rehabil 2019 Apr;100(4S):S1-S12. [doi: 10.1016/j.apmr.2018.08.175] [Medline: 30195987]
39. Amjad H, Wong SK, Roth DL, et al. Health services utilization in older adults with dementia receiving care coordination: the MIND at home trial. Health Serv Res 2018 Feb;53(1):556-579. [doi: 10.1111/1475-6773.12647] [Medline: 28083879]
40. Carver CS, Scheier MF, Weintraub JK. Assessing coping strategies: a theoretically based approach. J Pers Soc Psychol 1989;56(2):267-283. [doi: 10.1037/0022-3514.56.2.267]
41. Cohen S, Mermelstein R, Kamarck T, Hoberman HM. Measuring the functional components of social support. In: Sarason IG, Sarason BR, editors. Research and Applications NATO ASI Series: Springer Netherlands; 1985:73-94. [doi: 10.1007/978-94-009-5115-0_5]
42. Ritter PL, Sheth K, Stewart AL, Gallagher-Thompson D, Lorig K. Development and evaluation of the eight-item caregiver self-efficacy scale (CSES-8). Gerontologist 2022 Mar 28;62(3):e140-e149. [doi: 10.1093/geront/gnaa174] [Medline: 33146727]
43. Graf C. The Lawton instrumental activities of daily living scale. Am J Nurs 2008;108(4):52-62. [doi: 10.1097/01.NAJ.0000314810.46029.74]
44. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [doi: 10.1093/geront/9.3_Part_1.179] [Medline: 5349366]
45. Teri L, Truax P, Logsdon R, Uomoto J, Zarit S, Vitaliano PP. Assessment of behavioral problems in dementia: the revised memory and behavior problems checklist. Psychol Aging 1992;7(4):622-631. [doi: 10.1037/0882-7974.7.4.622]
46. Schulz R, Czaja SJ, McKay JR, Ory MG, Belle SH. Intervention taxonomy (ITAX): describing essential features of interventions. Am J Health Behav 2010;34(6):811-821. [doi: 10.5993/ajhb.34.6.15] [Medline: 20604704]
47. Schulz R, O'Brien A, Czaja S, et al. Dementia caregiver intervention research: in search of clinical significance. Gerontologist 2002 Oct;42(5):589-602. [doi: 10.1093/geront/42.5.589] [Medline: 12351794]
48. Sörensen S, Pinquart M, Duberstein P. How effective are interventions with caregivers? an updated meta-analysis. Gerontologist 2002 Jun;42(3):356-372. [doi: 10.1093/geront/42.3.356] [Medline: 12040138]
49. Gaugler JE, Reese M, Sauld J. A pilot evaluation of psychosocial support for family caregivers of relatives with dementia in long-term care: the residential care transition module. Res Gerontol Nurs 2015;8(4):161-172. [doi: 10.3928/19404921-20150304-01] [Medline: 25751083]
50. Albers EA, Birkeland RW, Louwagie KW, et al. A qualitative analysis of mechanisms of benefit in the residential care transition module: a telehealth intervention for caregivers of relatives with dementia living in residential long-term care. Inquiry 2023;60:469580231217981. [doi: 10.1177/00469580231217981] [Medline: 38142369]
51. Gaugler JE, Statz TL, Birkeland RW, et al. The ResidentialCare Transition Module: a single-blinded randomized controlled evaluation of a telehealth support intervention for family caregivers of persons with dementia living in residential long-term care. BMC Geriatr 2020 Apr 15;20(1):133. [doi: 10.1186/s12877-020-01542-7] [Medline: 32293314]
52. Gaugler JE, Birkeland RW, Albers EA, et al. Efficacy of the residential care transition module: a telehealth intervention for dementia family caregivers of relatives living in residential long-term care settings. Psychol Aging 2024 Aug;39(5):565-577. [doi: 10.1037/pag0000820] [Medline: 38753405]
53. Mittelman MS, Bartels SJ. Translating research into practice: case study of a community-based dementia caregiver intervention. Health Aff (Millwood) 2014 Apr;33(4):587-595. [doi: 10.1377/hlthaff.2013.1334]
54. Gaugler JE, Reese M, Mittelman MS. Process evaluation of the NYU caregiver intervention-adult child. Gerontologist 2018 Mar 19;58(2):e107-e117. [doi: 10.1093/geront/gnx048] [Medline: 29562359]
55. Mittelman MS, Roth DL, Coon DW, Haley WE. Sustained benefit of supportive intervention for depressive symptoms in caregivers of patients with Alzheimer's disease. AJP 2004 May 1;161(5):850-856. [doi: 10.1176/appi.ajp.161.5.850]

XSL•FO
RenderX

56. Mittelman MS, Roth DL, Haley WE, Zarit SH. Effects of a caregiver intervention on negative caregiver appraisals of behavior problems in patients with Alzheimer's disease: results of a randomized trial. J Gerontol B Psychol Sci Soc Sci 2004 Jan;59(1):27-34. [doi: 10.1093/geronb/59.1.p27] [Medline: 14722336]
57. Mittelman MS, Haley WE, Clay OJ, Roth DL. Improving caregiver well-being delays nursing home placement of patients with Alzheimer disease. Neurology (ECronicon) 2006 Nov 14;67(9):1592-1599. [doi: 10.1212/01.wnl.0000242727.81172.91] [Medline: 17101889]
58. Pearlin LI, Lieberman MA, Menaghan EG, Mullan JT. The stress process. J Health Soc Behav 1981 Dec;22(4):337-356. [doi: 10.2307/2136676] [Medline: 7320473]
59. Train Health Care Workers About Dementia. Health Resources & Services Administration. 2017 Dec. URL: https://bhw.hrsa.gov/alzheimers-dementia-training [accessed 2025-05-27]
60. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implementation Sci 2017 Dec;12(1):108. [doi: 10.1186/s13012-017-0635-3]
61. Wänström L. Sample sizes for two-group second-order latent growth curve models. Multivariate Behav Res 2009 Sep 30;44(5):588-619. [doi: 10.1080/00273170903202589]
62. Morse JM, Field PA. Qualitative Research Methods for Health Professionals: SAGE Publications; 1995:0803973276.
63. Strauss A, Corbin J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory: SAGE Publications; 1998.
64. Morse JM. Confusing categories and themes. Qual Health Res 2008 Jun;18(6):727-728. [doi: 10.1177/1049732308314930]
65. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
66. Williams M, Moser T. The art of coding and thematic exploration in qualitative research. Int Manag Rev 2019;15(1):45-55 [FREE Full text]
67. Kapasi A, DeCarli C, Schneider JA. Impact of multiple pathologies on the threshold for clinically overt dementia. Acta Neuropathol 2017 Aug;134(2):171-186. [doi: 10.1007/s00401-017-1717-7] [Medline: 28488154]
68. Brenowitz WD, Hubbard RA, Keene CD, et al. Mixed neuropathologies and estimated rates of clinical progression in a large autopsy sample. Alzheimers Dement 2017 Jun;13(6):654-662. [doi: 10.1016/j.jalz.2016.09.015] [Medline: 27870939]

Abbreviations

AD/ADRD: Alzheimer Disease/Alzheimer Disease and Related Dementias
ADL: activities of daily living
AIM: acceptability of intervention measure
FIM: feasibility of intervention measure
HCS: health care system
HIPAA: Health Insurance Portability and Accountability Act
IAM: intervention appropriateness measure
***ICD*:** *International Classification of Diseases*
***ICD-10*:** *International Statistical Classification of Diseases, Tenth Revision*
MVAHCS: Minneapolis Veterans Affairs Health Care System
RCT: randomized controlled trial
REACH: Resources for Enhancing Alzheimer's Caregiver Health
SPIRIT : Standard Protocol Items: Recommendations for Interventional Trials
TACSI: TBI-AD/ADRD Caregiver Support Intervention
TBI: traumatic brain injury
TRC: treatment receipt checklist
UMN: University of Minnesota

Edited by E Andrikopoulou; submitted 22.Jul.2025; peer-reviewed by H Feng, T Bashford; accepted 26.Jan.2026; published 17.Mar.2026.

Please cite as:
Ostenso S, Birkeland RW, Albers E, Louwagie KW, Iwan A, Chesak SS, Gaugler JE
The TBI-AD/ADRD Caregiver Support Intervention (TACSI): Protocol of a Pilot Randomized Controlled Trial Evaluation of a Remote Intervention for Family Caregivers
JMIR Res Protoc 2026;15:e81125
URL: https://www.researchprotocols.org/2026/1/e81125
doi:10.2196/81125

XSL•FO
RenderX

© Samantha Ostenso, Robyn W Birkeland, Elizabeth Albers, Katie W Louwagie, Allison Iwan, Sherry S Chesak, Joseph E Gaugler. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Implementation and Potential Effects of the Nishauri mHealth Intervention on HIV Care Among Men in Homa Bay County, Kenya: Protocol for a Mixed Methods Study

Dan Omollo[1,2], BSc, MPH; Pamela Murnane[3], PhD; Louisa Ndunyu[2], PhD; Dallas Swendeman[4], PhD; Sheri D Weiser[5], PhD; Collins Ouma[6], PhD

[1]Department of Epidemiology and Biostatistics, Global Health Institute, University of California, San Francisco, 590 Minnesota, San Fransisco, CA, United States

[2]Department of Public Health, School of Public Health and Community Development, Maseno University, Kisumu, Kenya

[3]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

[4]Semel Institute for Neuroscience and Human Behavior, University of California, Los Angeles, Los Angeles, CA, United States

[5]Department of Medicine, University of California, San Francisco, San Fransisco, CA, United States

[6]School of Public Health and Community Development, Maseno University, Kisumu, Kenya

Corresponding Author:
Dan Omollo, BSc, MPH
Department of Epidemiology and Biostatistics, Global Health Institute, University of California, San Francisco, 590 Minnesota, San Fransisco, CA, United States

Abstract

Background: About 2.4% of Kenyan people (approximately 1.3 million people) are living with HIV. Despite advances in antiretroviral therapy, men continue to experience disproportionately poor engagement in HIV care due to entrenched masculine norms, stigma, and lack of tailored interventions. Mobile health (mHealth) platforms offer a promising strategy to improve care engagement, but evidence on its implementation and impact among men living with HIV is limited.

Objective: This study aims to assess the implementation and potential effects of the Nishauri mHealth intervention on HIV care and treatment outcomes among men in Western Kenya. Specifically, it seeks to (1) analyze its effects on HIV care engagement and treatment outcomes, (2) explore the role of masculine identity in modifying acceptability and uptake, and (3) identify barriers and facilitators of adoption, use, and sustainment.

Methods: We will use a mixed methods design combining a stepped-wedge cluster approach and a pre- and postimplementation assessment across 4 health facilities in Homa Bay County, Kenya. Approximately 347 men receiving HIV treatment who own a smartphone will be enrolled. The stepped-wedge design will sequentially introduce the intervention across the 4 facilities in 2-month intervals following baseline assessments, allowing each site to serve as its own control. Surveys will collect data on sociodemographics, masculinity, intervention acceptability and uptake, and HIV clinical outcomes using validated measures. Intervention effects on pre- and postbinary outcomes will be assessed using the McNemar test, while generalized estimating equations (α=.05; β=.2; 95% CI) will account for clustering and repeated measures in the stepped-wedge analysis. Focus group discussions (n=5-6) will be conducted with men living with HIV, health care providers, and app developers to explore barriers and facilitators of implementation and adoption. Focus group discussions will be audio-recorded, transcribed, coded, and analyzed thematically.

Results: This study received institutional review board approval in July 2025 and was registered on ClinicalTrials.gov in August 2025. Recruitment began in September 2025 and concluded in November 2025. A total of 307 men living with HIV were recruited across the 4 clinics for the pre- and postquantitative assessment. Preliminary findings will describe implementation outcomes and early effects on HIV care engagement.

Conclusions: This trial will use a stepped-wedge design to evaluate the implementation and effects of the Nishauri mHealth intervention on antiretroviral therapy adherence and clinic attendance among men in Homa Bay County. By examining both clinical outcomes and the influence of masculine norms on intervention uptake, it will provide robust evidence on the effectiveness of mHealth strategies tailored for men in low-resource, high–HIV-burden settings. Findings will inform the design, scalability, and optimization of similar interventions by identifying key implementation barriers and facilitators.

Trial Registration: ClinicalTrials.gov NCT07116538; https://clinicaltrials.gov/study/NCT07116538

International Registered Report Identifier (IRRID): DERR1-10.2196/85279

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e85279) doi:10.2196/85279

KEYWORDS

mHealth; mobile health; HIV care; men living with HIV; implementation; masculinity; Kenya; Nishauri

Introduction

Globally, men are 27% less likely than women to seek HIV testing and often present to care at later stages of illness [1]. In sub-Saharan Africa, men's participation in HIV care is further constrained by cultural expectations that prioritize economic provision, requiring men to focus on work and income generation over health-seeking. Additionally, norms around emotional suppression discourage men from expressing vulnerability or seeking support, including attending health services [2,3]. Recent studies highlight that men who conform to traditional masculine norms are less likely to attend routine HIV care and more likely to disengage from treatment [4-6]. Despite advances in antiretroviral therapy (ART) and HIV care services, men experience persistent barriers to accessing and engaging with care. Masculine norms around strength, stoicism, and self-reliance often discourage men from seeking health services, including HIV testing and treatment [4], contributing to lower engagement and adherence [7,8]. The lack of men-targeted interventions compounds these challenges, making it critical to develop and evaluate strategies that effectively engage men in care to achieve the Joint United Nations Programme on HIV/AIDS 95-95-95 targets by 2030.

Mobile health (mHealth) interventions have emerged as promising tools for improving health care delivery and patient engagement, particularly in resource-constrained settings [8-10]. By leveraging mobile devices to deliver health-related information and services, mHealth strategies help bridge critical gaps in access, communication, and continuity of care. Evidence from low- and middle-income countries indicates that mHealth interventions, such as SMS reminders and mobile apps, can improve ART adherence, retention in care, and health literacy [4,7] among people living with HIV. Studies globally and in sub-Saharan Africa have shown the potential of mHealth interventions to enhance clinical outcomes, reduce loss to follow-up, and strengthen engagement in care [4,7,10-13]. For instance, studies in Uganda and South Africa show that mobile reminders and SMS-based counseling significantly improve adherence and retention among individuals with HIV [14-16]. These tools facilitate behavior change through personalized messaging, medication reminders, and educational content. Critically, culturally adapted mHealth interventions targeted for men are more effective, as they address stigma and align health-seeking behaviors with socially accepted masculine norms. For example, SMS reminders framed around men's responsibilities or apps providing discreet, personalized health information have been shown to improve ART adherence and clinic attendance among men in sub-Saharan Africa [14-17]. This demonstrates how aligning messages and delivery with socially accepted masculine norms can reduce stigma and enhance engagement. However, there remains limited evidence on the implementation outcomes, service outcomes, and long-term sustainability of mHealth interventions specifically tailored for men in HIV care in Kenya, a country that continues to face significant challenges in the fight against HIV/AIDS, with an estimated 2.4% of the population being HIV-seropositive [18].

In Kenya, the Nishauri (client-centered mHealth digital) platform was developed to enhance HIV care management by supporting appointment reminders, ART adherence, patient-provider communication, and educational messaging. Pilot implementations of Nishauri and its predecessor, Ushauri, have shown promise in improving engagement and retention, though systematic evidence on its adoption and sustained use, particularly among men, remains limited [19,20]. While mHealth technologies such as Nishauri have shown promise in improving HIV care outcomes [21,22], their adoption and sustained use in Kenya remain at an early stage and face significant challenges. Structural barriers, including digital literacy gaps, inconsistent device ownership, limited infrastructure, and concerns around data privacy and confidentiality, continue to hinder equitable uptake [23,24]. Mobile phone ownership and digital access also vary by gender, education level, and socioeconomic status [25], with men in rural, high-prevalence settings like Homa Bay County often facing compounded disadvantages. Beyond structural issues, many mHealth interventions are not adequately aligned with the gendered and sociocultural realities that shape care-seeking behaviors. Programs often lack meaningful provider involvement and fail to tailor delivery to address male-specific barriers or local norms [26]. Although designed to be discreet and convenient, mHealth tools may fall short if they do not resonate with prevailing masculine identities or reframe care-seeking in ways that affirm strength and responsibility [17]. In high-burden regions such as Homa Bay, where HIV prevalence is 16.2% [18,27], men frequently disengage from care due to stigma and perceived threats to masculinity [28].

In parts of Homa Bay and other counties where Nishauri has been implemented on a pilot basis for over 6 months, its uptake and use remain low and highly variable across facilities, providers, and patient populations. This variability hints at challenges with its implementation and effectiveness among men, a population that faces unique sociocultural barriers to HIV care. Existing research often overlooks the role of gender norms and contextual factors in influencing men's engagement with digital health interventions. Addressing this gap is essential for designing scalable, culturally relevant mHealth solutions that promote sustained engagement and improved clinical outcomes among men. This study will be guided by an integrated conceptual framework combining Connell's theory of hegemonic masculinity [29] and the technology acceptance model (TAM) [30,31] to underpin the design and evaluation of Nishauri, predicting how individual factors (eg, perceived usefulness, usability, and masculinity norms) and system-level factors (eg, workflow integration and provider support) influence adoption, engagement, and sustainment of the intervention.

XSL•FO
RenderX

This study aims to assess the implementation and potential effects of the Nishauri mHealth intervention on HIV care and treatment outcomes among men aged 18 - 55 years in Homa Bay County, Western Kenya. Specifically, it seeks to (1) analyze its potential effects on HIV care and treatment outcomes, (2) explore the role of masculine identity in modifying its acceptability and uptake, and (3) identify barriers and facilitators of its adoption, use, and sustainment.

Methods

Study Design

This study will use an explanatory sequential mixed methods design, combining quantitative and qualitative approaches to understand both implementation outcomes and contextual factors influencing intervention uptake [32]. The quantitative component will use a stepped-wedge approach across 4 health facilities (n=1899) [33,34] analyzed using generalized linear mixed models to account for clustering and repeated measures [34,35] with pre- and postassessment of implementation outcomes [36,37] (n=347; Figure 1). The qualitative component will involve thematic analysis of focus group discussions (FGDs; 5 - 6 groups of ~6 key stakeholders) across 4 health facilities in Homa Bay County (Table 1). This analysis will combine deductive coding guided by Proctor's implementation outcomes framework with inductive coding to capture emergent themes [36]. This integrated approach ensures rigorous evaluation of the Nishauri mHealth intervention across multiple levels and stakeholder perspectives.

Figure 1. Study Schema. FGD: focus group discussion; HCP: healthcare providers.

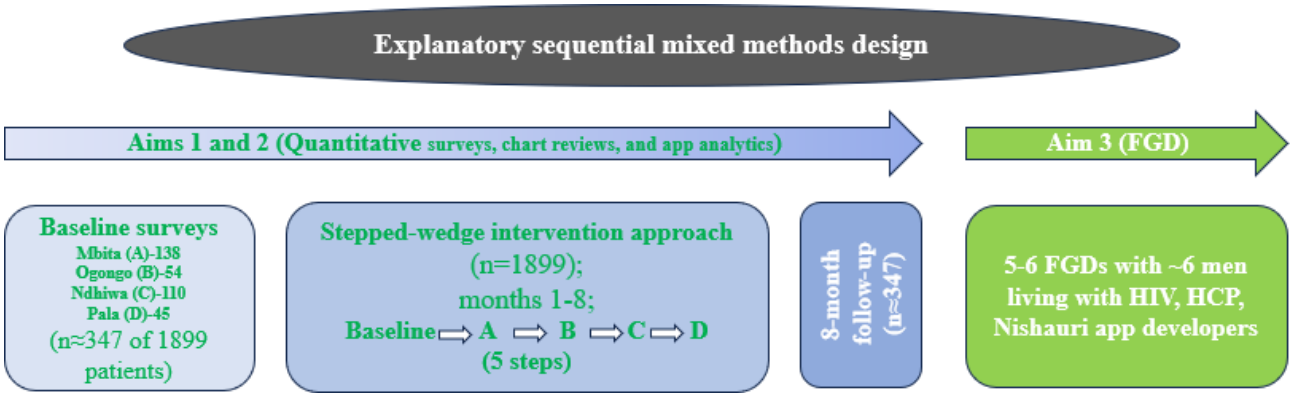

Table . Focus group discussion (FGD) sampling plan or rationale.

Group	FGDs, n	Participants per FGD, n	Total participants, n	Plan or rationale
Men living with HIV	2 - 3	6 - 8	12 - 24	• Stratified by age, missed visits, antiretroviral therapy adherence, and intervention adoption.
Health care providers	2	6 - 8	12 - 16	• Grouped by facility or role (eg, nurse vs adherence counselor). • County or subcounty officials to provide valuable insight into policy, scale-up, and sustainability.
App developers	1	4 - 6	4 - 6	• Being a homogeneous group, 1 FGD session should suffice.

After baseline data collection is completed, a standardized readiness assessment will be conducted across the 4 facilities to determine the procedure of the phased intervention rollout. We will leverage a partnership with the Palladium Group (the app developers) to access backend use data from the Nishauri app, providing detailed insights into how different app functions are used in real-world settings. This mixed methods design will enable both within- and between-facility comparisons, offering a nuanced understanding of how implementation processes and gendered experiences shape the intervention's effectiveness among men. This study protocol was developed and reported in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 guidelines (Checklist 1).

Study Setting

We will conduct this study in 4 subcounty HIV Comprehensive Care Clinics (CCC) in Homa Bay County, Kenya—Mbita,

XSL•FO
RenderX

Ogongo, Ndhiwa, and Pala. Homa Bay County was chosen for its highest HIV prevalence, 16.2% [18] in Kenya, making it suitable to assess the implementation process and outcomes of an mHealth intervention for enhancing HIV care. The county covers an area of about 3154 km^2 [38,39]. As of the 2019 Kenya National Census, it had an estimated population of approximately 1.1 million people [39]. There are approximately 160 health facilities providing HIV comprehensive care services across the county. ART uptake is at 96%, while viral load suppression (VLS) prevalence is at 83.8% among people living with HIV, with significant disparities between men and women [18,27,40].

The 4 health facilities have been selected based on feasibility factors such as patient volume and mobile network coverage, representativeness of diverse service delivery contexts, and requirements of the stepped-wedge design to ensure adequate clusters while minimizing contamination. They serve a total of 1899 men living with HIV aged 18 - 55 years: Mbita (n=758), Ogongo (n=294), Ndhiwa (n=598), and Pala (n=249). The average distance between each neighboring facility is 25 km. Mbita and Ndhiwa subcounty hospitals are situated in an urban setup, while Ogongo and Pala subcounty health centers are located in relatively rural setups.

About Nishauri mHealth Intervention

The Nishauri mHealth intervention is a client-centered digital platform developed by Palladium, in collaboration with Kenya's Ministry of Health, and funded by the President's Emergency Plan for AIDS Relief through the Centers for Disease Control and Prevention, to enhance HIV care and treatment outcomes. Nishauri supports key aspects of HIV care management, including appointment scheduling, ART referrals, and ongoing patient engagement. Its core functionalities include automated appointment and medication adherence reminders. It also enables tailored health education messages, 2-way communication between patients and health care providers, and behavior-change communication strategies to support retention in care. Patients with HIV using Nishauri can access their treatment engagement information and schedule or reschedule appointments through the app. A screenshot of the Nishauri app, captured directly from the live platform, is presented in Figure 2 for illustrative purposes.

Figure 2. Screenshot of the Nishauri mobile health app showing the home and dashboard interfaces.

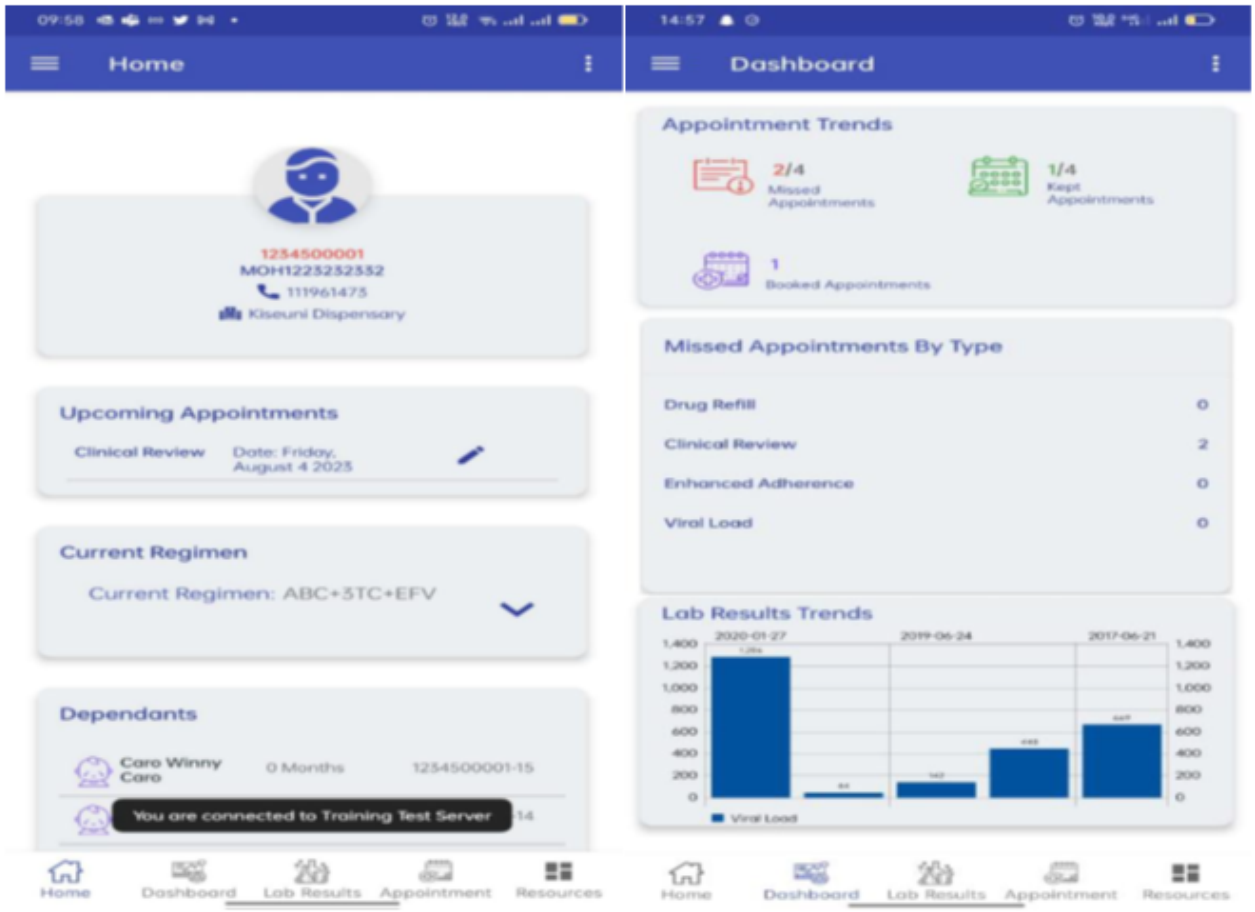

Nishauri is the Ministry of Health–endorsed scale-up of Kenya's earlier Ushauri system, which was piloted to support appointment reminders, treatment adherence messaging, and defaulter tracing through SMS, web, and mobile interfaces [19]. Building on the functionality and adoption of Ushauri, Nishauri is a client-centered mobile app that integrates in real time with the Kenya electronic medical record system and the Ushauri platform to support appointment scheduling, adherence reminders, 2-way communication, and behavior-change messaging for people living with HIV [19,20].

XSL•FO
RenderX

Conceptual Framework and Theoretical Grounding

This study is guided by an integrated conceptual framework (Figure 3) drawing on Connell's theory of hegemonic masculinity [29,41] and the TAM. Hegemonic masculinity provides a lens for understanding how dominant gender norms, such as expectations of strength, self-reliance, and economic provision, influence men's engagement with HIV care and perceptions of digital health interventions. TAM informs the assessment of how perceived usefulness and ease of use of the Nishauri mHealth platform shape adoption and sustained use. These frameworks are integrated to examine how masculine norms modify technology acceptance and implementation outcomes, including adoption, acceptability, and engagement, ultimately influencing HIV care outcomes among men.

Figure 3. Study conceptual framework. ART: antiretroviral therapy; mHealth: mobile health; VLS: viral load suppression.

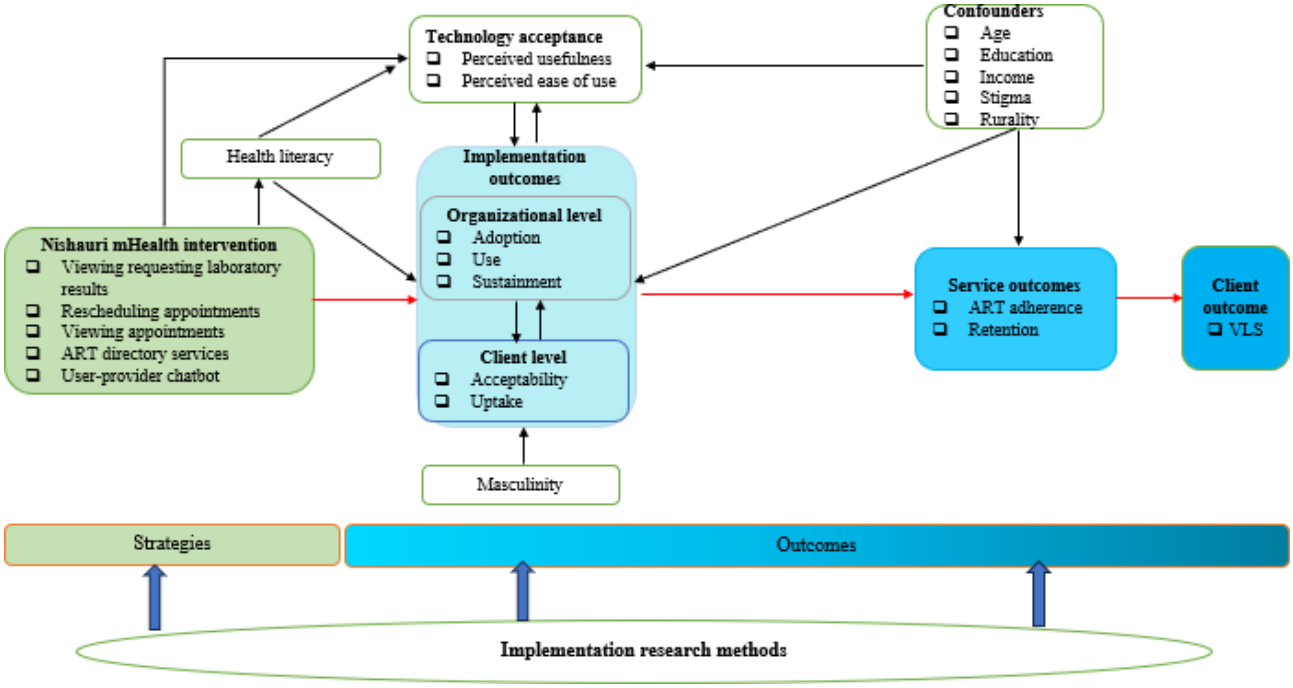

Stepped-Wedge Intervention Approach

Strategy

A stepped-wedge design was chosen to allow all facilities to receive the intervention, accommodate phased training and app deployment, minimize contamination between sites, and enable within-facility comparisons over time while controlling for secular trends. The 4 CCCs will undergo a standardized readiness assessment, and phased implementation of the intervention will proceed in sequence according to facility readiness, with sites transitioning from standard care to the mHealth intervention arm over an 8-month period. The phased rollout will occur at 2-month intervals across the 4 participating facilities, resulting in 5 study periods (1 baseline and 4 intervention steps) over approximately 8 months. The 2-month interval was selected to align with routine HIV clinic visit schedules in Kenya [18], allow sufficient time for provider training and participant onboarding, and permit stabilization of app engagement before subsequent facility transitions. This interval is consistent with stepped-wedge evaluations of health service and digital interventions [42] and supports the detection of short-term changes in missed clinic visits and adherence behaviors while accounting for secular trends.

We acknowledge that the 2-month step length is relatively short and may limit detection of longer-term intervention effects within each wedge; however, this interval was selected for programmatic feasibility and alignment with routine HIV care cycles, and outcome measures were therefore chosen to capture proximal, short-term indicators of engagement and retention.

After baseline, all men who access HIV care from each facility will be trained on how to install and use the Nishauri app by adherence counselors or peer educators already trained by Palladium Group. This will happen during clinic education sessions as well as community outreaches organized through community ART (support) groups. Peer educators, through the community support groups, will provide follow-up support to ensure that all eligible men are reached. Contamination will be minimized by staggering rollout, restricting app activation to intervention sites, and monitoring cross-facility clinic attendance. We will track each participant's primary facility at enrollment and confirm the facility where ART is routinely dispensed. Study inclusion will be restricted to men who are primarily engaged at 1 of the 4 participating facilities during the intervention period. Any participant who routinely receives care at multiple facilities will be carefully documented, and their data will be linked to the facility where they are primarily enrolled. These procedures are consistent with both national treatment guidelines and best practices for stepped-wedge designs in HIV implementation research.

Due to its facility-level implementation, individual-level exclusion criteria will not apply, as specified for the pre- and postassessment, which involves individual-level measurements and participant follow-up.

XSL•FO
RenderX

Outcomes

The primary clinical outcome will be missed clinic appointments, while secondary outcomes will include ART adherence (30-day percent visual analog), pharmacy refill consistency (medication possession ratio), and VLS—plasma HIV RNA <200 copies per milliliter. CCC data will be retrieved from facility electronic medical records and patient files, focusing on routinely collected HIV care indicators.

Implementation Outcomes

Guided by Proctor's implementation outcomes framework, this study seeks to evaluate multiple implementation outcomes of the Nishauri mHealth intervention using a mixed methods approach. The primary implementation outcome will be adoption. Adoption will be measured as completion of app installation and account activation within 60 days of introduction. Acceptability (the perceived satisfaction and appropriateness of the intervention among users and providers) and use will be assessed through validated survey instruments, including the System Usability Scale and intervention acceptability measures, alongside qualitative data from FGDs with users and providers. Early indicators of sustainment are assessed through continued engagement metrics over the study period and stakeholder perspectives on long-term integration and scalability, acknowledging that full sustainment extends beyond the study time frame.

Stepped-Wedge Analysis

We will analyze intervention effects using intention-to-treat principles and generalized linear mixed models appropriate to each outcome. Models will include a fixed effect for time period (step) to account for secular trends, a fixed indicator for intervention exposure during each step (within-cluster), and random intercepts for clusters (facilities) to account for within-cluster correlation. Where appropriate, we will include random slopes for time if supported by the data. Covariates prespecified for adjustment include patient-level baseline characteristics (age, gender, income, education, and stigma) and facility-level covariates (facility size and urban or rural). We will use robust SEs (cluster-robust SEs) to ensure valid inference with a limited number of clusters. Additional sensitivity analyses will include per-protocol analyses and models testing for time-by-intervention interaction to explore whether effect sizes change across steps. Primary inference will be 2-sided with α=.05. Analyses will be performed in Stata (StataCorp LLC; xtmelogit/xtmixed), and code will be archived with the study repository for reproducibility.

Power Estimation for Stepped Wedge Analysis

Sample size and power were estimated for the stepped-wedge cluster design using simulation under a Hussey-Hughes mixed effects logistic model. With facility totals of 1899—Mbita (n=758), Ogongo (n=294), Ndhiwa (n=598), and Pala (n=249), assuming baseline missed-appointment proportion (p0) of 20% (0.20) and 5 periods, the study has ~80% power to detect an intervention effect (missed visits) corresponding to an odds ratio of 0.75, assuming intracluster correlation coefficient=0.02 (α=.05). Sensitivity analyses indicated that power remained ≥70% for odds ratios of 0.80 and intracluster correlation coefficient values up to 0.05 and exceeded 90% for odds ratios of ≤0.70, supporting the robustness of the proposed sample size.

Pre- and Postassessment of Implementation Outcomes

Participants and Recruitment

The primary study population will comprise men who initiated HIV care and treatment in 2016 or later, following the updated national HIV treatment guidelines. Eligible participants must be between 18 and 55 years of age, own a smartphone or tablet, and be able and willing to provide written informed consent (Multimedia Appendix 1). Men with severe comorbidities or chronic coinfections (eg, cancer or hypertension) that may affect the adoption or use of mHealth interventions will be excluded. Only men who are primarily enrolled at 1 of the 4 participating facilities during the study period will be included. Participant recruitment in each CCC will proceed according to probability proportional to size sampling until the required sample size of 347 participants is reached. Recruitment will occur during routine clinic visits in each facility before the stepped-wedge rollout of Nishauri. Trained study staff or peer educators will introduce the study to clients, provide study information, and collect contact information from those expressing interest. Screening for eligibility will take place in private, either in a clinic or at a location convenient and confidential for the participant. Those who are eligible and interested will then undergo written informed consent.

Quantitative Data Collection and Storage

Quantitative data will be collected at baseline, at 8-month follow-up, and via the Nishauri app data. At baseline, potential participants will be screened for eligibility. Eligible participants will provide written informed consent and locator information before undergoing a short baseline survey lasting approximately 40 minutes. The survey includes validated and pilot-tested modules on sociodemographics, technology acceptance, HIV-related stigma, health literacy, masculinity norms, the System Usability Scale, and intervention acceptability. The survey will be administered electronically using tablets with the REDCap (Research Electronic Data Capture) platform, hosted by the San Francisco Coordinating Center. Data will be collected by trained research assistants (RAs), who were locally recruited and had prior experience in health or social science research. RAs have backgrounds in public health or related disciplines; are fluent in English, Kiswahili, and Dholuo; and will receive study-specific training on the study protocol, ethical conduct of research, informed consent, and standardized data collection procedures prior to data collection. RAs will perform daily quality checks on all forms to identify and resolve errors prior to secure upload to REDCap servers. These servers are firewall-protected, regularly backed up, and housed in a secure environment.

Throughout implementation at each site, app-based metrics will be assessed to capture implementation outcomes related to acceptability and uptake of the Nishauri intervention. Facility-based data will be collected, including who gets the intervention, the date of training, and the date of app installation. App analytics will be used to capture patterns of engagement and uptake among participants. This includes the number of

XSL•FO
RenderX

log-ins per user, frequency of use over time, cumulative duration of app sessions, and responsiveness to reminders and notifications (whether SMS or in-app prompts were opened or acted upon). Aggregate metrics such as average log-ins per month, median engagement time, and drop-off points will be analyzed to assess acceptability and sustained use across the study period.

All data will be transferred securely to the San Francisco Coordinating Center network via a secure File Transfer Protocol site under a formal data transfer agreement between the University of California, San Francisco, and the local research center. Data will be stored on encrypted, password-protected servers with limited access, using study ID numbers in place of participant names. Physical documents will be stored securely onsite, and all sensitive data will be separated from identifying information to ensure confidentiality.

Quantitative Data Analysis

Descriptive statistics will be used to summarize demographic variables, intervention adoption, and predictors of adoption (health literacy, technology acceptance, and masculinity). A participant will be considered to have adopted Nishauri if he has completed device installation and created or activated an account within 60 days of being trained or introduced to the app. Participants who occasionally attend multiple facilities will be documented, and their data will be analyzed based on their primary facility of care. This aligns with national HIV care guidelines, which encourage patients to register at a single facility for longitudinal HIV treatment and monitoring.

Power Calculation for the Pre- and Postassessment

With 347 participants, an anticipated 10% attrition rate (effective n≈312), and a primary binary uptake outcome, the study will have adequate power to examine associations between key predictors (health literacy, technology acceptance, and masculinity) and adoption of Nishauri. Assuming an uptake rate of ~70% (adopters: n≈243; nonadopters: n≈104) as seen in other studies [22,23], we will have >80% power (α=.05) to detect moderate associations corresponding to odds ratios of roughly 1.7 - 2.0 or higher, depending on predictor prevalence. Sensitivity analyses showed that with effective sample sizes ranging from 280 to 330 participants, the study would retain ≥80% power to detect odds ratios between 1.7 and 2.0, depending on predictor prevalence.

Retention Strategy

Participants will receive automated appointment reminders via the Nishauri platform and additional study reminders 1 week prior to their scheduled 8-month follow-up visit. For missed visits, trained study staff will attempt to contact participants by phone call or SMS text message to reschedule as soon as possible. If participants cannot be reached remotely, study staff will arrange an in-person follow-up at a safe and confidential location. Participants who relocate outside the study area will be contacted to facilitate follow-up data collection at an appropriate alternative site. These strategies are designed to minimize loss to follow-up and support sustained engagement with both the intervention and routine HIV care.

Focus Group Discussions

Qualitative Data Collection

To complement the quantitative findings, 5 to 6 FGDs with approximately 6 - 8 participants per group will be conducted in person by trained and qualified RAs and the principal investigator (DO). The FGDs will comprise 2 - 3 with men living with HIV, 2 with health care providers (nurses, adherence counselors, or peer educators), and 1 with app developers. FGD sampling among men living with HIV will be stratified by factors such as age, ART adherence, and proportion of missed clinic visits (Table 1). We will invite 10 participants to each FGD and expect at least 6 to participate in each group. This sample size is expected to achieve thematic saturation in this study context, based on qualitative research standards [43] and was chosen based on guidelines, suggesting that 4 - 6 FGDs are sufficient to reach thematic saturation in studies of similar scope and target populations [43-45]. This sampling strategy balances the need for rich qualitative data with feasibility considerations, allowing for in-depth exploration of implementation experiences across stakeholder groups. FGDs will continue until no new themes emerge, ensuring comprehensive coverage of barriers, facilitators, and contextual factors influencing the adoption, use, and sustainment of the Nishauri mHealth intervention.

To preserve the integrity of the data, most FGD participants will be different from those in the quantitative survey arm to reduce contamination, social desirability bias, and respondent fatigue. However, a limited number of individuals who participated in the pre- and poststudy may be included in FGDs to help explain emerging patterns in the data, as this study uses an explanatory mixed methods approach. Overlap will be explicitly tracked, and all analyses will account for repeated participation when interpreting themes, ensuring that qualitative findings complement, rather than double-count, quantitative data. Health care providers and app developers will also be purposively sampled to ensure representation of different implementation perspectives across the system.

The sampling plan and rationale for the FGDs is shown in Table 1. The FGD guides (Multimedia Appendix 2) have been organized thematically and include open-ended questions and follow-up probes to facilitate in-depth exploration of participants' experiences with the intervention. We will explore contextual factors, including masculinity, stigma, and social norms, that influence the adoption, use, and sustainment of the Nishauri mHealth intervention. The FGDs will be audio-recorded using digital recorders and conducted in private settings to ensure confidentiality and adherence to ethical standards.

Qualitative Data Analysis

Audio recordings will be transcribed verbatim and translated into English and analyzed using Dedoose. We will use a thematic analysis approach combining deductive coding guided by Proctor's implementation outcomes framework [36,37] with inductive coding to capture emergent themes not covered by the framework. Descriptive statistics will summarize FGD participant demographics. Narrative summaries, tables,

XSL•FO
RenderX

diagrams, and cross-case comparisons will be used to deepen interpretation. Triangulation with app analytics (eg, proportion of missed visits and use metrics) and member checking will enhance credibility. This study uses an explanatory sequential mixed methods design, using quantitative data to assess intervention uptake and outcomes, followed by qualitative inquiry to explain how factors like masculine norms, stigma, and system-level challenges influence these patterns. Qualitative findings will be integrated with application data from the Nishauri dashboard to uncover underlying mechanisms and contextual factors shaping implementation outcomes. Standard operating procedures will guide FGD facilitation, qualitative data handling, and analysis.

Dissemination Plan

Study findings will be shared with participants, health care providers, Homa Bay County Ministry of Health officials, and other key stakeholders. Results will be disseminated through presentations at local and international conferences focused on HIV prevention, treatment, and behavioral medicine, as well as through publication in peer-reviewed journals. The study team has established an authorship agreement outlining contributions to primary and secondary publications. Participants' identities will remain confidential and will not appear in any reports or publications. Direct feedback will also be provided to research participants and community partners to ensure inclusive and transparent communication of the findings.

Ethical Considerations

This study will be conducted in accordance with the ethical principles of the Belmont Report—respect for persons, beneficence, and justice. Ethics approval was obtained from the Maseno University Scientific Ethical Review Committee (MSU/DRPI/MUSERC/01502/25), which serves as the primary institutional review board. All study materials, including the study protocol, data collection tools, and any protocol amendments, will be submitted to the institutional review board for review and approval prior to implementation. In addition, administrative and institutional approvals were obtained from the National Commission on Science, Technology and Innovation, Homa Bay County—Department of Health, the participating health facilities (Mbita, Ndhiwa, Ogongo, and Pala subcounty hospitals), and the University of California, San Francisco Department of Epidemiology and Biostatistics. All study staff will complete training in human subjects' protection and Good Clinical Practice prior to study initiation. Written informed consent will be obtained by trained RAs before any study procedures are conducted. Consent forms will be available in English, Dholuo, and Swahili, and participants will choose their preferred language. Literate participants will read and sign the form; those with limited literacy will have the form read aloud in the presence of an impartial witness who will document the consent accordingly. All participants will receive a copy of the signed consent form, or if deemed unsafe to keep, may store it at the clinic and receive a contact slip for future reference. Participants will be reimbursed Ksh 300 (US $2.32) for their time. Strict confidentiality measures will be observed, given the sensitivity of HIV-related data. Personal identifiers will be anonymized or pseudonymized, and all data will be stored in secure, encrypted systems accessible only to authorized, trained personnel. Participants will be informed of confidentiality protections during the consent process. Hard copies of study documents will be stored securely for 5 years before being shredded; electronic and audio files will be permanently deleted at that time.

Results

The study obtained institutional review board approval on July 28, 2025. Baseline data collection began on September 8, 2025, and has been completed in 3 health facilities and is anticipated to end by January 31, 2026. A total of 307 men living with HIV were recruited across the 4 clinics for the pre- and postquantitative assessment. Follow-up and app use monitoring are ongoing in the first facility. Follow-up data collection is expected to be completed in August 2026. Analysis is projected to be concluded in the third quarter of 2026, and the results will be submitted for publication at the end of 2026. Findings are anticipated to inform gender-responsive adaptations of digital health strategies and guide policy recommendations for scaling mHealth interventions for men living with HIV in Kenya.

Discussion

Principal Findings

This study is guided by a conceptual framework integrating Proctor's implementation outcomes framework, TAM, and the masculinity model (Figure 3). It is designed to generate evidence on the implementation and impact of the Nishauri mHealth intervention on HIV care and treatment among men living with HIV in Homa Bay County, Kenya. With this model guiding data collection and analysis, we anticipate findings that will address critical gaps in understanding how mHealth tools perform in resource-constrained, gendered contexts.

By identifying both individual- and system-level factors that influence the adoption, use, and sustainment of the Nishauri platform, this study will generate valuable insights into the dynamic interplay between client- and organizational-level implementation outcomes. Barriers such as limited digital literacy, stigma, poor network coverage, and workflow integration challenges, as well as facilitators like ease of use, perceived usefulness, provider support, and trust in the health system, will be explored in depth. These findings will inform the development of gender-responsive, culturally grounded adaptations to Nishauri as well as other digital health strategies aimed at improving long-term health outcomes for men living with HIV in Kenya and advancing national efforts to close gender gaps in HIV care.

We hypothesize that there will be measurable improvements in HIV service outcomes, particularly ART adherence and clinic retention, among men exposed to the intervention, compared to control periods within the stepped-wedge design. These outcomes will be interpreted in light of confounding variables, such as age, education, and stigma, and supported by app analytics and chart review data. Furthermore, the client outcome of VLS will offer a critical downstream indicator of the intervention's potential clinical impact.

XSL•FO
RenderX

This study will contribute critical insights into how masculine norms shape the implementation and impact of digital health tools for HIV care. By unpacking how behavioral constructs such as autonomy, stoicism, and fear of vulnerability influence engagement with mHealth interventions, the findings will inform the development of more gender-responsive and contextually grounded digital strategies. These insights have the potential to guide national policies and programs aiming to close gender gaps in HIV care by tailoring interventions to the unique needs and experiences of men.

By integrating implementation science and gender theory within a mixed methods design, this study will generate a nuanced understanding of individual- and system-level factors influencing the success of mHealth tools among men in HIV care. While the findings may be context-specific to Homa Bay County, they offer valuable insights to inform the design and scale-up of culturally relevant, gender-sensitive digital health strategies in Kenya and similar low-resource settings. Future research should build on these results to codevelop scalable interventions and policy frameworks that support sustainable integration into routine HIV care for men.

Strengths

This study has several strengths, including its use of a stepped-wedge design within real-world public health settings, enhancing both rigor and relevance. By focusing on men, an underserved population in HIV care, and guided by Proctor's implementation outcomes framework, the study offers a structured analysis of adoption, use, and sustainability. The explanatory sequential mixed methods approach, combined with mHealth app data, allows for deep insights into both outcomes and underlying mechanisms. Additionally, the study's gender-informed lens and alignment with national HIV priorities position it to generate policy-relevant evidence for more effective and equitable digital health interventions.

Limitations

While this study uses a stepped-wedge cluster randomized design, several limitations should be noted. First, the number of clusters (health facilities) is small, which may limit statistical power and the ability to control for facility-level confounding. The small number of clusters limits the precision with which facility-level confounders can be modeled and reduces degrees of freedom for estimating random effects. As a result, adjustment for cluster-level confounding is restricted to a limited set of prespecified facility characteristics (eg, size and urban-rural setting), and residual confounding at the facility level cannot be entirely ruled out. To mitigate this limitation, analyses will include fixed effects for time, intention-to-treat estimation, and triangulation with qualitative data to support interpretation of observed effects. Second, reliance on self-reported survey measures for some implementation outcomes (eg, acceptability and adherence) and masculinity factors may introduce recall or social desirability bias. Third, inclusion criteria requiring smartphone or tablet ownership may exclude more men from resource-limited countries, potentially limiting the generalizability of findings. Finally, variations in how the intervention is implemented across facilities—due to staffing, infrastructure, or engagement differences—may influence observed effects. Despite these limitations, the study's mixed methods approach and triangulation of data sources will enhance the robustness and contextual relevance of the findings.

Conclusions

The Nishauri mHealth intervention was developed to address persistent gaps in HIV care engagement and outcomes among people living with HIV in Kenya, a group historically underserved by conventional HIV programming. While mHealth solutions have shown promise in improving HIV-related outcomes, few have been systematically implemented or evaluated with a gender lens in high-burden settings like Homa Bay County. This study will evaluate the real-world implementation and effects of Nishauri using a mixed methods, stepped-wedge design—generating evidence on both process and clinical outcomes while integrating perspectives from users, providers, and developers. If successful, Nishauri could offer a scalable, gender-responsive mHealth model that enhances retention, adherence, and viral suppression among men. The findings will inform future efforts to strengthen digital health strategies and promote more equitable HIV care for men in Kenya and similar resource-limited settings.

Acknowledgments

The authors would like to thank in advance the study participants, adherence counselors, and peer educators for supporting the recruitment planning and process. The authors also acknowledge the support of the County Department of Health for granting them local approval and fruitful inception and their research assistants for their contributions in data collection. This work has been supported in part by mentorship and guidance from the Sustainable Development for Improving HIV Health in Kenya program mentors from University of California, San Francisco, and Maseno University. The authors are also grateful to the developers of the Nishauri mobile health platform—the Palladium Group—for their technical support throughout the study planning process.

Funding

This study was made possible through the support of the Fogarty International Center of the National Institutes of Health (D43TW011306). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

XSL•FO
RenderX

Data Availability

Deidentified data are available upon request from Maseno University and Sustainable Development for Improving HIV Health repositories.

Authors' Contributions

SDW and LN were responsible for funding acquisition. DO and PM were responsible for the conceptualization of the study, data curation, and methodology. DO wrote the original draft of the manuscript and was responsible for project administration. All authors assisted in the writing, reviewing, and editing of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Informed consent form.
[PDF File, 175 KB - resprot_v15i1e85279_app1.pdf]

Multimedia Appendix 2
Focus group discussion guide.
[PDF File, 247 KB - resprot_v15i1e85279_app2.pdf]

Checklist 1
SPIRIT checklist.
[PDF File, 128 KB - resprot_v15i1e85279_app3.pdf]

References

1. Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases: interim guidance. World Health Organization. 2020. URL: https://www.who.int/publications/i/item/10665-331501 [accessed 2025-03-07]
2. Pulerwitz J, Gottert A, Kahn K, et al. Gender norms and HIV testing/treatment uptake: evidence from a large population-based sample in South Africa. AIDS Behav 2019 Sep;23(Suppl 2):162-171. [doi: 10.1007/s10461-019-02603-8] [Medline: 31359218]
3. Chikovore J, Gillespie N, McGrath N, Orne-Gliemann J, Zuma T, ANRS 12249 TasP Study Group. Men, masculinity, and engagement with treatment as prevention in KwaZulu-Natal, South Africa. AIDS Care 2016;28(Suppl 3):74-82. [doi: 10.1080/09540121.2016.1178953] [Medline: 27421054]
4. Colvin CJ. Strategies for engaging men in HIV services. Lancet HIV 2019 Mar;6(3):e191-e200. [doi: 10.1016/S2352-3018(19)30032-3] [Medline: 30777726]
5. Inman EM, Zharima C, Hammock A, Violari A, Kidman R. The role of masculinity in HIV disclosure among adolescent South African boys living with HIV. AIDS Care 2025 Oct;37(10):1785-1794. [doi: 10.1080/09540121.2025.2547941] [Medline: 40838871]
6. Young AM, Moji M, Duby Z, et al. The “Ideal Man”: how gender norms and expectations shape South African men’s masculinity, sexual identities, and well-being. J Men’s Stud 2025 Jun;33(2):417-440. [doi: 10.1177/10608265241303734]
7. Okal J, Lango D, Matheka J, et al. “It is always better for a man to know his HIV status”—a qualitative study exploring the context, barriers and facilitators of HIV testing among men in Nairobi, Kenya. PLoS ONE 2020;15(4):e0231645. [doi: 10.1371/journal.pone.0231645] [Medline: 32294124]
8. Whiteley LB, Olsen EM, Haubrick KK, Odoom E, Tarantino N, Brown LK. A review of interventions to enhance HIV medication adherence. Curr HIV/AIDS Rep 2021 Oct;18(5):443-457. [doi: 10.1007/s11904-021-00568-9] [Medline: 34152554]
9. McCool J, Dobson R, Whittaker R, Paton C. Mobile health (mHealth) in low- and middle-income countries. Annu Rev Public Health 2022 Apr 5;43:525-539. [doi: 10.1146/annurev-publhealth-052620-093850] [Medline: 34648368]
10. Sharma S, Gergen Barnett K, Maypole JJ, Grochow Mishuris R. Evaluation of mHealth apps for diverse, low-income patient populations: framework development and application study. JMIR Form Res 2022 Feb 11;6(2):e29922. [doi: 10.2196/29922] [Medline: 35147502]
11. Esmaeili ED, Azizi H, Dastgiri S, Kalankesh LR. Does telehealth affect the adherence to ART among patients with HIV? A systematic review and meta-analysis. BMC Infect Dis 2023 Mar 17;23(1):169. [doi: 10.1186/s12879-023-08119-w] [Medline: 36932376]
12. Nelson KM, Perry NS, Horvath KJ, Smith LR. A systematic review of mHealth interventions for HIV prevention and treatment among gay, bisexual, and other men who have sex with men. Transl Behav Med 2020 Oct 12;10(5):1211-1220. [doi: 10.1093/tbm/ibaa007] [Medline: 33044531]

XSL•FO
RenderX

13. Ameli V, Wong G, Barlow J, et al. How mobile health can change the contexts of living with HIV and engaging with treatment and care in Iran: a realist-informed qualitative study. Qual Health Res 2025 Feb;35(2):156-173. [doi: 10.1177/10497323241256865] [Medline: 39106351]
14. Buh AW. Non-Adherence to HIV Treatment Among Patients in Cameroon: Prevalence, Predictors And Effective Strategies Improving Treatment Adherence [Doctoral Dissertation]. Université d'Ottawa/University of Ottawa; 2024. : Université d'Ottawa/University of Ottawa; 2024.
15. Williams LD, van Heerden A, Friedman SR, Chibi B, Rodriguez WA, Memela P. Changes in stigma and social support among participants in a randomized trial of a novel expanded social network-based HIV testing intervention in KwaZulu-Natal, South Africa. AIDS Behav 2024 Aug;28(8):2619-2629. [doi: 10.1007/s10461-024-04379-y] [Medline: 38833064]
16. Sam-Agudu NA, Adirieje C, Agwu AL, Rakhmanina N. Barriers and facilitators to equitable implementation of long-acting ART for adolescents and youth with HIV in low- and middle-income settings. Pan Afr Med J 2024;49:53. [doi: 10.11604/pamj.2024.49.53.45322] [Medline: 39911367]
17. Sileo KM, Fielding-Miller R, Dworkin SL, Fleming PJ. What role do masculine norms play in men's HIV testing in sub-Saharan Africa?: A scoping review. AIDS Behav 2018 Aug;22(8):2468-2479. [doi: 10.1007/s10461-018-2160-z] [Medline: 29777420]
18. Preliminary KENPHIA 2018 report. : National AIDS and STI Control Programme (NASCOP); 2020 URL: https://phia.icap.columbia.edu/wp-content/uploads/2020/04/KENPHIA-2018_Preliminary-Report_final-web.pdf [accessed 2026-03-07]
19. Mudogo CM, Mulwa A, Kyalo DN. Increasing management efficiencies and effectiveness of HIV/AIDS projects using digital health interventions in Kenya: a qualitative inquiry. J Prof Bus Rev 2024;9(12):e04890. [doi: 10.26668/businessreview/2024.v9i12.4890]
20. Mwangi C, Mukanya C, Maghanga C. Assessing the interoperability of mLab and Ushauri mHealth systems to enhance care for HIV/AIDS patients in Kenya. JIPIT 2022;2(1):83-116. [doi: 10.52907/jipit.v2i1.209]
21. El Joueidi S, Bardosh K, Musoke R, et al. Evaluation of the implementation process of the mobile health platform "WelTel" in six sites in East Africa and Canada using the modified consolidated framework for implementation research (mCFIR). BMC Med Inform Decis Mak 2021 Oct 26;21(1):293. [doi: 10.1186/s12911-021-01644-1] [Medline: 34702229]
22. Ojo JO, Ipinnimo TM, Afolayan AC, et al. Factors associated with willingness to use mHealth interventions for medication adherence among people living with HIV attending a tertiary hospital in sub-Saharan Africa. In: Jahun I, editor. PLoS ONE 2024;19(8):e0309119. [doi: 10.1371/journal.pone.0309119] [Medline: 39146337]
23. Gagnon MP, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
24. Kruse CS, Pacheco GJ, Rosenthal N, Kopp CJ, Omorotionmwan O, Cruz JE. Leveraging mHealth for the treatment and management of PLHIV. Risk Manag Healthc Policy 2023;16:677-697. [doi: 10.2147/RMHP.S403946] [Medline: 37077534]
25. Jennings Mayo-Wilson L, Kang BA, Mathai M, Mak'anyengo MO, Ssewamala FM. Mobile phone access, willingness, and usage for HIV-related services among young adults living in informal urban settlements in Kenya: a cross-sectional analysis. Int J Med Inform 2022 May;161:104728. [doi: 10.1016/j.ijmedinf.2022.104728] [Medline: 35228007]
26. Olang APBO. mHealth Utilization and Access to Treatment by Teenagers Living With HIV/AIDS in Island Communities of Lake Victoria, Kenya [Doctoral Dissertation]. : Jomo Kenyatta University of Agriculture and Technology; 2024 URL: https://ir.jkuat.ac.ke/bitstream/handle/123456789/6359/Olang%E2%80%99%2C%20A.%20P.%20B.%20O.%20PhD%20Thesis%202024.pdf?isAllowed=y&sequence=1 [accessed 2026-03-07]
27. Banadakoppa Manjappa R, Bhattacharjee P, Shaw SYN, et al. A sub-national HIV epidemic appraisal in Kenya: a new approach for identifying priority geographies, populations and programmes for optimizing coverage for HIV prevention. J Int AIDS Soc 2024 Jul;27(Suppl 2):e26245. [doi: 10.1002/jia2.26245] [Medline: 38982894]
28. Nagata JM, Magerenge RO, Young SL, Oguta JO, Weiser SD, Cohen CR. Social determinants, lived experiences, and consequences of household food insecurity among persons living with HIV/AIDS on the shore of Lake Victoria, Kenya. AIDS Care 2012;24(6):728-736. [doi: 10.1080/09540121.2011.630358] [Medline: 22150119]
29. Reddy R, Sharma AK, Jha M. Hegemonic masculinity or masculine domination: toward a comprehensive social theory of gender. IJSSP 2019;39:296-310. [doi: 10.1108/IJSSP-08-2018-0133]
30. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
31. Amadu L, Syed Muhammad S, Mohammed AS, Owusu G, Lukman S. Using technology acceptance model to measure the ese of social media for collaborative learning in Ghana. J Technol Sci Educ 2018;8(4):321. [doi: 10.3926/jotse.383]
32. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
33. Hemming K, Eldridge S, Forbes G, Weijer C, Taljaard M. How to design efficient cluster randomised trials. BMJ 2017 Jul 14;358:j3064. [doi: 10.1136/bmj.j3064] [Medline: 28710062]
34. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. Contemp Clin Trials 2007 Feb;28(2):182-191. [doi: 10.1016/j.cct.2006.05.007] [Medline: 16829207]

XSL•FO
RenderX

35. Twisk J, Bosman L, Hoekstra T, Rijnhart J, Welten M, Heymans M. Different ways to estimate treatment effects in randomised controlled trials. Contemp Clin Trials Commun 2018 Jun;10:80-85. [doi: 10.1016/j.conctc.2018.03.008] [Medline: 29696162]
36. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
37. Proctor EK, Bunger AC, Lengnick-Hall R, et al. Ten years of implementation outcomes research: a scoping review. Implementation Sci 2023;18(1):31. [doi: 10.1186/s13012-023-01286-z]
38. Agala CB, Xiong K, Thomas JC, Powell R. Referral network analysis for improved HIV care in Homa Bay County, Kenya. : MEASURE Evaluation; 2016 URL: https://www.measureevaluation.org/resources/publications/wp-16-159/at_download/document [accessed 2026-03-07]
39. Wellbeing in Kenya: special focus on inequalities in wellbeing in Kenya. : Kenya National Bureau of Statistics (KNBS); 2024 URL: https://www.knbs.or.ke/wp-content/uploads/2024/05/Inequalities-in-wellbieng-in-Kenya-Report-.pdf [accessed 2026-03-07]
40. Masaba R, Woelk G, Siamba S, et al. Antiretroviral treatment failure and associated factors among people living with HIV on therapy in Homa Bay, Kenya: a retrospective study. In: Shin SS, editor. PLOS Glob Public Health 2023;3(3):e0001007. [doi: 10.1371/journal.pgph.0001007] [Medline: 36962996]
41. Mathenjwa T, Adeagbo O, Zuma T, et al. Development and acceptability of a tablet-based app to support men to link to HIV care: mixed methods approach. JMIR Mhealth Uhealth 2020 Nov 24;8(11):e17549. [doi: 10.2196/17549] [Medline: 33231558]
42. Cooper V, Clatworthy J, Whetham J, Consortium E. mHealth interventions to support self-management in HIV: a systematic review. TOAIDJ 2017 Nov 21;11(1):119-132. [doi: 10.2174/1874613601711010119]
43. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017 Mar;27(4):591-608. [doi: 10.1177/1049732316665344] [Medline: 27670770]
44. Morse WC, Lowery DR, Steury T. Exploring saturation of themes and spatial locations in qualitative public participation geographic information systems research. Soc Nat Resour 2014 May;27(5):557-571. [doi: 10.1080/08941920.2014.888791]
45. Crabtree BF, Miller WL. Doing Qualitative Research, 3rd edition: Sage Publications, Inc; 2023.

Abbreviations

ART: antiretroviral therapy
CCC: Comprehensive Care Clinic
FGD: focus group discussion
mHealth: mobile health
RA: research assistant
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TAM: technology acceptance model
VLS: viral load suppression

Edited by A Schwartz; submitted 05.Oct.2025; peer-reviewed by CK Nalwadda, DI Mangale; accepted 30.Jan.2026; published 24.Mar.2026.

Please cite as:
Omollo D, Murnane P, Ndunyu L, Swendeman D, Weiser SD, Ouma C
Assessing the Implementation and Potential Effects of the Nishauri mHealth Intervention on HIV Care Among Men in Homa Bay County, Kenya: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e85279
URL: https://www.researchprotocols.org/2026/1/e85279
doi:10.2196/85279

© Dan Omollo, Pamela Murnane, Louisa Ndunyu, Dallas Swendeman, Sheri D Weiser, Collins Ouma. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Virtual Reality–Based Program for Pediatric Patients With Amblyopia: Protocol for a Multicenter, Randomized, Open-Label, Two-Arm Study

Ken Nagino[1,2,3,4], PhD; Yuichi Okumura[2,3,4], MD, PhD; Masakazu Hirota[3,5,6,7], PhD; Saiko Matsumura[8], MD, PhD; Tadashi Matsumoto[8], MD, PhD; Takashi Negishi[4], MD, PhD; Akie Midorikawa-Inomata[1,2,3], RN, MPH, PhD; Makiko Ui[9], MD; Takao Hayashi[5,6], MD, PhD; Yuichi Hori[8], MD, PhD; Shintaro Nakao[4], MD, PhD; Takenori Inomata[1,2,3,4], MD, PhD, MBA

[1]Department of Hospital Administration, Juntendo University Graduate School of Medicine, Tokyo, Japan

[2]Department of Telemedicine and Mobile Health, Juntendo University Graduate School of Medicine, Tokyo, Japan

[3]InnoJin, Inc, Tokyo, Japan

[4]Department of Ophthalmology, Juntendo University Graduate School of Medicine, 2-2-1 Hongo, Bunkyo-ku, Japan

[5]Department of Orthoptics, Faculty of Medical Technology, Teikyo University, Tokyo, Japan

[6]Department of Ophthalmology, School of Medicine, Teikyo University, Tokyo, Japan

[7]Graduate Degree Program of Health Data Science, Teikyo University, Tokyo, Japan

[8]Department of Ophthalmology, Toho University Faculty of Medicine, Tokyo, Japan

[9]CS Eye Clinic, Tokyo, Japan

Corresponding Author:
Takenori Inomata, MD, PhD, MBA
Department of Hospital Administration, Juntendo University Graduate School of Medicine, Tokyo, Japan

Abstract

Background: Conventional amblyopia treatment involves the occlusion of the fellow eye using an eye patch. However, this approach imposes a substantial psychological and physical burden on pediatric patients with amblyopia, resulting in low adherence and suboptimal visual outcomes. As responsiveness to therapy declines beyond early childhood, treatments that can enhance adherence and improve efficacy are needed. We developed a virtual reality dichoptic training app (VR app) that integrates gamification and hand-eye coordination training in pediatric patients with amblyopia.

Objective: This study investigated the effect of a VR app on visual acuity improvement in pediatric patients with amblyopia compared with conventional occlusion therapy using an eye patch.

Methods: This is a multicenter, open-label, prospective, randomized controlled trial that will be conducted at 3 centers in Tokyo, Japan. Pediatric patients aged 3-10 years with anisometropic, strabismic, or refractive amblyopia will be enrolled. Participants will be assigned 1:1 to a VR app group or a control group. The VR app group will use the VR app at home for 1 hour per day for 6 months. The control group will undergo conventional occlusion therapy with an eye patch for 6 months. Each group will comprise 15 participants (30 participants in total). Ophthalmologic examinations will be performed at baseline and at weeks 4, 8, 12, 16, 20, and 24. The primary endpoint will be the change in the best-corrected visual acuity in the amblyopic eye from baseline to week 12. Secondary endpoints will include changes in best-corrected visual acuity, stereopsis, and ocular deviation through week 24. We will also assess treatment adherence, defined as the ratio of the cumulative actual treatment time to the cumulative prescribed time; record adverse events, and evaluate usability in the VR app group. Longitudinal changes in outcome measures will be analyzed using a mixed-effects model.

Results: Participant enrollment will start from January 1, 2026, to September 30, 2026. The data analysis will begin on October 1, 2026, and the results will be reported by March 31, 2027.

Conclusions: This study will clarify the effectiveness of the newly developed VR app in improving treatment outcomes and adherence among pediatric patients with amblyopia. By addressing the limitations of conventional occlusion therapy and providing a potentially more efficacious and acceptable treatment option, the VR app may enhance clinical outcomes and represent a paradigm shift in the treatment of pediatric amblyopia.

International Registered Report Identifier (IRRID): PRR1-10.2196/85194

(JMIR Res Protoc 2026;15:e85194) doi:10.2196/85194

XSL•FO
RenderX

KEYWORDS

amblyopia; dichoptic training; pediatric; randomized controlled trial; treatment adherence; virtual reality; digital therapeutics; software as a medical device; telemedicine; gamification; hand-eye coordination

Introduction

Amblyopia is primarily characterized by reduced visual acuity in one or both eyes, with a prevalence of approximately 1% to 5% among children worldwide and an estimated prevalence of 0.1% to 0.2% in Japan [1-5]. Pediatric amblyopia is most commonly attributable to refractive errors, anisometropia, strabismus, and visual deprivation [6]. The treatment of amblyopia involves the use of fully corrective glasses in combination with occlusion therapy using an eye patch to train the amblyopic eye. However, the sensitive period of visual development extends only until approximately the age of 7 years [7], after which the effectiveness of treatment becomes significantly limited. Accordingly, amblyopia necessitates public health strategies aimed at early detection, the establishment of effective therapeutic approaches, and efficient management [8].

In Japan, under the Maternal and Child Health Law, all 3-year-old children undergo ophthalmologic screening, including testing for amblyopia, as part of their routine health examination [9]. The screening process is implemented in successive stages. Initially, parents conduct vision testing at home, and the results are submitted to the health examination site. At the venue, pediatricians, in collaboration with public health nurses and other nonophthalmic professionals, review the submitted data and, when indicated, perform supplementary preliminary assessments. If abnormalities are suspected, the child is subsequently referred to an ophthalmology clinic, where ophthalmologists conduct comprehensive examinations [9]. Nevertheless, reports indicate that home-based vision screening overlooks approximately 24.2% of pediatric patients with amblyopia [9]. According to a report assessing vision screening at 30-35 months of age in the United Kingdom, orthoptist-administered screening achieved sensitivities and specificities approaching 100%. Contrastingly, screening performed by healthy visitors yields a sensitivity of only approximately 50%, with approximately half of the affected cases remaining undetected [10]. Accordingly, evidence suggests that certain cases of amblyopia go unrecognized when comprehensive examinations at medical institutions are not undertaken, thereby precluding the timely initiation of appropriate therapy [10].

In occlusion therapy utilizing an eye patch, whereby the healthy eye is covered to promote the use of the amblyopic eye, the duration of daily occlusion must be extended with advancing patient age, thereby imposing an increasing psychological and physical burden on the child [11,12]. Excessive burden associated with amblyopia treatment may result in reduced adherence, manifested by noncompliance with the occlusion of the healthy eye by the child or premature discontinuation of therapy by parents of their own judgment, thereby attenuating the overall effectiveness of amblyopia management [13-15]. A study assessing adherence to occlusion therapy using an eye patch reported that patients complied with the patch wear for approximately half of the prescribed occlusion time [16]. Adherence to amblyopia therapy has been demonstrated to exhibit a dose-response relationship with therapeutic efficacy [17]. Accordingly, the development of alternative modalities to occlusion therapy that lessen the psychological and physical burden, thereby mitigating the decline in adherence, is warranted.

To overcome the challenges associated with pediatric amblyopia therapy, we developed a clinical investigational device, the virtual reality (VR) app, which utilizes VR as a programmatic medical device for amblyopia training in children (Figure 1A) [18]. The VR-based Software as Medical Device (SaMD) for pediatric amblyopia incorporates functionality to partially render the visual input to the healthy eye transparent and includes a gaming module operated through hand-tracking controllers [18]. The VR app enables amblyopia training through interactive gameplay, wherein part of the image projected onto a healthy eye is rendered transparent to facilitate the engagement of the amblyopic eye during the task [18]. An illustrative example of a VR-based SaMD for pediatric amblyopia training is the Luminopia One (Luminopia, Inc.), which was granted authorization by the U.S. Food and Drug Administration [19]. The therapeutic effect of Luminopia One is achieved by attenuating the contrast of visual stimuli presented to the healthy eye, which in turn promotes the use of the amblyopic eye and enhances the efficacy of amblyopia training [19]. The principal difference between Luminopia One and the VR app is that Luminopia One delivers only contrast-modified visual content, whereas the VR app integrates interactive games requiring active physical engagement via hand-tracking controllers, such as kendama (Figure 1A), tennis (Figure 1B), and table tennis (Figure 1C) [18,19]. In contrast to the amblyopia training implemented in Luminopia One, which is limited to passive viewing of visual content with the amblyopic eye, the VR app incorporates VR-based games that require active hand-eye coordination, potentially conferring greater therapeutic efficacy in amblyopia treatment [20]. Thus, the VR app–based amblyopia training may confer greater therapeutic efficacy than conventional approaches, owing to its distinctive advantages: alleviation of psychological and physical burden by obviating the use of an eye patch, improvement of adherence through gamification, and provision of effective training that incorporates hand-eye coordination [17,18,20,21].

This study aims to conduct a multicenter clinical trial to assess the therapeutic effectiveness of a VR app in pediatric amblyopia. The use of the VR app is expected to enable amblyopia training with higher adherence and enhanced efficacy compared to conventional occlusion therapy, with the potential to mitigate the societal burden attributable to pediatric amblyopia.

XSL•FO
RenderX

Figure 1. The virtual reality (VR) app for pediatric patients with amblyopia. (A) VR head–mounted display and a screenshot of the Kendama game projected to each eye. In the fellow eye, the ball is displayed as semitransparent. (B) Screenshot of the tennis game projected to the fellow eye. (C) Screenshot of the table tennis game projected to the fellow eye.

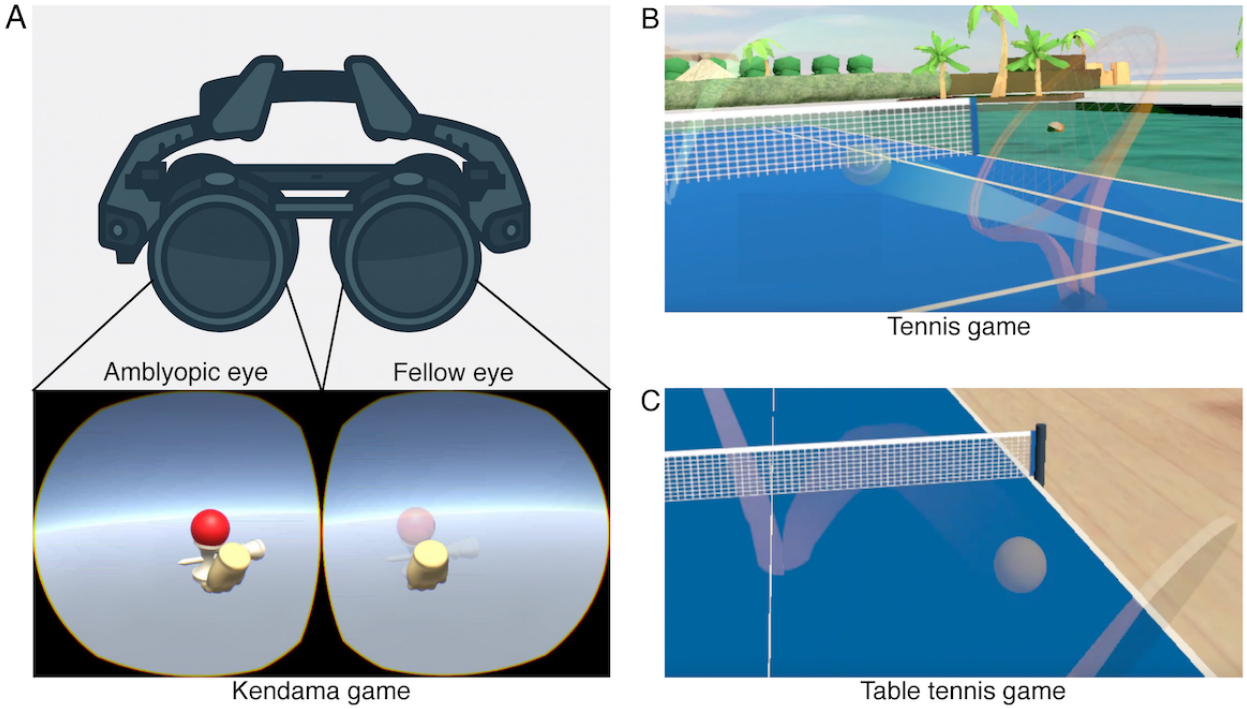

Methods

The VR Dichoptic Training App for Pediatric Patients With Amblyopia

A VR app for pediatric patients with amblyopia was developed by InnoJin Inc., ImaCreate Co. Ltd., and Sumitomo Corporation. The VR app has not yet been approved as an SaMD in Japan as of August 31, 2025. After being approved for SaMD, pediatric patients with amblyopia will be able to download and use the VR app for pediatric amblyopia on a VR device after receiving a prescription from their physicians.

The VR app for pediatric patients with amblyopia uses a gamification approach to train the amblyopic eye in pediatric patients. Currently, the VR app for pediatric patients with amblyopia provides kendama, tennis, and table tennis games (Figure 1A-C, respectively). Kendama is a children's activity that uses the traditional Japanese skill toy called "kendama" [22], which consists of a handle (called "ken," meaning "sword") and a ball (called "tama" or "dama") connected by a string. The basic way to play is to catch the thrown ball with 1 of the 3 cups on the handle, or to "spike" the ball by inserting the tip of the handle into a hole on the ball. Countless combinations and styles exist to catch the ball, and players often combine techniques in unique ways to perform tricks. In the VR app, players will wear the VR head–mounted display and swing the controllers, allowing them to move the kendama handle and catch the ball in a virtual world. The "tennis" and "table tennis" games will both be sports in which players hit a ball back and forth with a racket by swinging the controllers to move a racket and hit the ball in the virtual world. This study will use the MeganeX Superlight 8K (Shiftall Inc.) as the head-mounted display.

The VR app will provide amblyopia training within the games. By ensuring that the ball image projected to the healthy eye is semitransparent or transparent, the program replicates the occlusion of the healthy eye and encourages the use of the amblyopic eye while playing the game. Additionally, the VR app for pediatric patients with amblyopia will incorporate hand-eye coordination training using hand-tracking controllers, which may promote the greater development of visual pathways compared with conventional approaches that require only image viewing without physical movement [20].

Study Setting

This multicenter, open-label, randomized controlled trial is scheduled to be conducted from January 1, 2026, to March 31, 2027. Patients will be enrolled at 3 medical institutions: Toho University Omori Medical Center (Tokyo, Japan), Teikyo University Hospital (Tokyo, Japan), and the CS Eye Clinic (Tokyo, Japan).

Ethical Considerations

This study has been approved by the Teikyo University Certified Review Board, Tokyo, Japan (protocol version 1.0; August 14, 2025; approval number: C0130-6). The parents of the participating children, acting as their representatives, will provide written informed consent. Informed assent will also be obtained whenever possible from children judged to have at least an elementary school level of understanding. All individuals involved in this study will make every effort to protect the personal information and privacy of the participating patients, and patient data will be anonymized to ensure that no individual can be identified. Participants who complete all study procedures through the final visit at week 24 will be provided with a voucher valued at 20,000 yen (approximately US $128,

XSL•FO
RenderX

as of January 5, 2026). Participants who discontinue participation prior to the final visit, after completing at least 12 weeks but fewer than 24 weeks of study participation, will be provided with a voucher valued at 10,000 yen (approximately US $64, as of January 5, 2026).

Inclusion and Exclusion Criteria

Outpatients who visit the collaborating study institutions will be recruited and must meet the following inclusion criteria: between 3 and 10 years old at time of enrollment; diagnosis of anisometropic amblyopia, strabismic amblyopia, or refractive amblyopia; best-corrected visual acuity (BCVA) of 0.3 logMAR or worse in the amblyopic eye at screening; and have a BCVA in the fellow eye of 0.2 logMAR or better at screening [19]. The inclusion age range was chosen because amblyopia treatment responsiveness declines with age, and improvement is less likely in older children. In this study, limiting enrollment to younger children was considered ethical because the VR-based intervention's effectiveness remains uncertain.

In this study, amblyopia is confirmed before enrollment, and since anisometropic and refractive amblyopia may occur even with small interocular acuity differences, we did not require a 2-line difference as eligibility criterion, though commonly referenced as a diagnostic indicator of amblyopia [23]. The exclusion criteria will be as follows: inability to obtain consent from the legal guardian; where applicable, inability to obtain assent from the patient; history of intraocular surgery or refractive surgery; history of photosensitive epilepsy; inability to comply with the study protocol due to psychological, familial, or geographic reasons; or inability to perform hand-eye coordination tasks, as determined by physical or mental conditions or by observation during a trial use of the study device, indicating a clear inability to establish hand-eye coordination.

Patient Withdrawal

Patients will be withdrawn from the study if, due to the occurrence of adverse events, it becomes difficult to continue the research; if patients or their parents or guardians request to end their participation; if the research study is discontinued; or if the principal investigator and subinvestigators judge it appropriate to discontinue the research.

Study Procedures

Figure 2 illustrates the study enrollment procedure. After obtaining informed consent, pediatric patients will be randomly assigned to either a VR app group, which will undergo amblyopia treatment using the VR app for pediatric patients with amblyopia combined with full refractive correction glasses, or a control group, which will undergo amblyopia treatment based on occlusion treatment with an eye patch combined with full refractive correction glasses [24]. Patients assigned to the VR app group will use the VR app for amblyopia at home for 1 hour per day for 6 months, in combination with full refractive correction glasses. Patients assigned to the control group will follow occlusion therapy with an eye patch for 6 months, also wearing full refractive correction glasses, and the duration of eye patch use will be determined by the prescribing physician. Both groups underwent ophthalmological examinations on the first day of the trial and at weeks 4, 8, 12, 16, 20, and 24. Ophthalmological examinations will include measurements of visual acuity, intraocular pressure, strabismus angle, stereopsis, refractive error, and corneal curvature. Only the VR app group will perform a digital stereopsis test using the VR app once per week and complete a usability questionnaire on the VR app at weeks 2, 12, and 24. The week 2 visit conducted only for the VR app group is intended to confirm that the study participants, particularly those unfamiliar with VR devices, are using the VR app correctly and facilitate their continued appropriate use. Only the control group will record their daily duration of eye patch use in a diary. Study participation will end for both groups after the 24-week visit.

XSL•FO
RenderX

Figure 2. Study flowchart. VR: virtual reality.

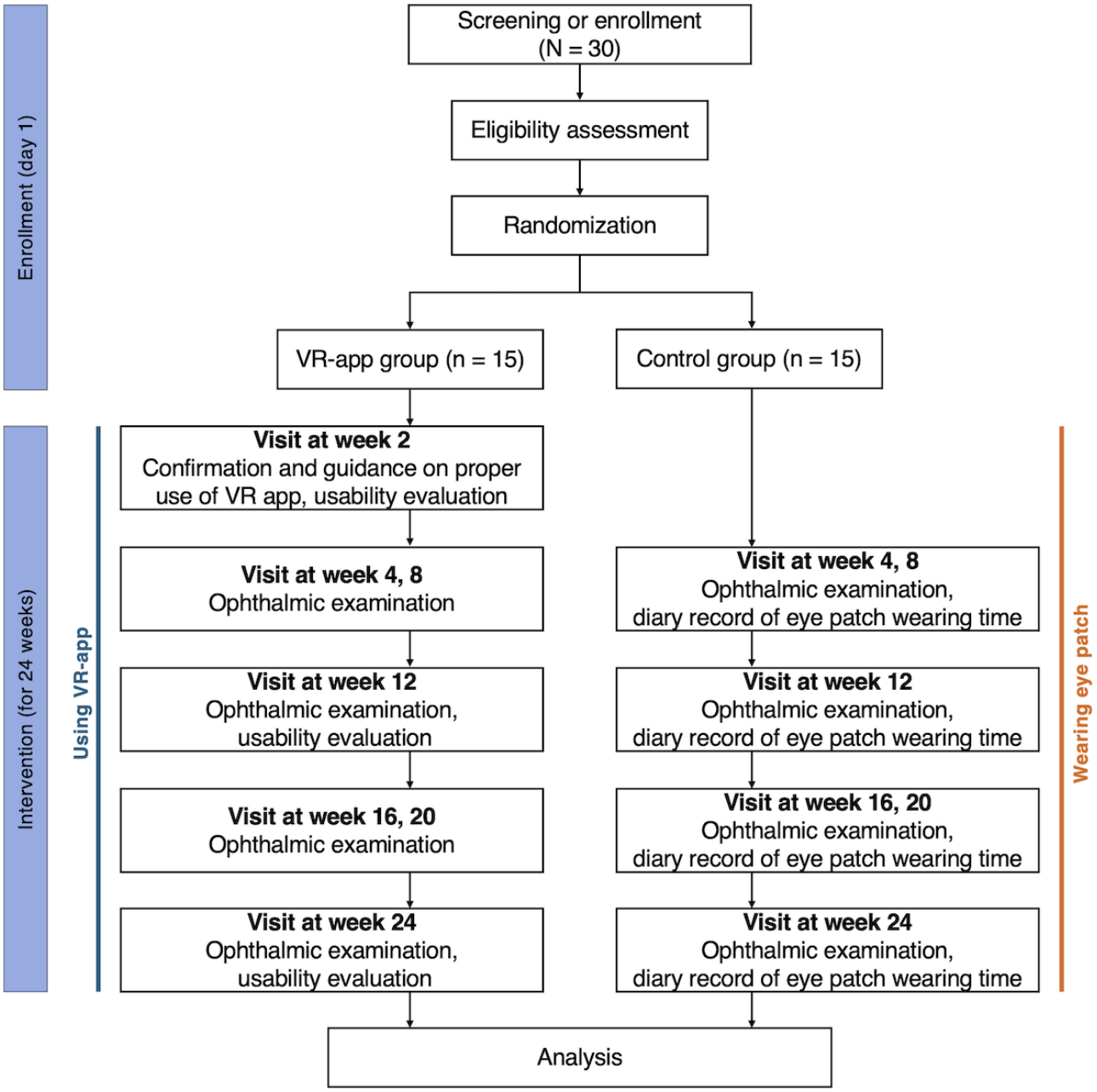

Data Collection and Outcome Assessments

The data collection items are listed in Textbox 1. Data on patient characteristics, ophthalmological examination results, VR app malfunctions, and adverse events will be collected from both the VR app and control groups. The daily usage time of the VR app, stereopsis test results obtained via the VR app, and usability evaluation results were collected only for the VR app group. Daily eye patch usage time will be collected only in the control group.

Visual acuity will be measured by BCVA using the Landolt-C chart at a distance [25]. Intraocular pressure will be measured using a noncontact tonometer [26]. The strabismus angle will be assessed using the alternate prism cover test [27]. Stereopsis will be evaluated using the Titmus stereo test [28]. The refractive error and corneal curvature radius will be measured using an autorefractor keratometer [29]. Daily VR app usage time will be calculated from the VR app usage logs, and daily eye patch wearing time will be reported by the patient's parent in a diary.

XSL•FO
RenderX

Textbox 1. Data collection.

Patient characteristics

- Age, sex, medical history (including anisometropic amblyopia, strabismic amblyopia, and refractive amblyopia), and duration of treatment for amblyopia

Ophthalmological examinations

- Visual acuity, intraocular pressure, angle of strabismus, stereopsis test, refractive error, corneal curvature

VR app group only

- Daily virtual reality (VR) app usage time
- VR app–based stereopsis test
- Usability evaluation

Control group only

- Daily eye patch usage time

Other outcomes

- VR app failure
- Adverse events

VR App–Based Stereopsis Test

In the VR app group, a stereopsis test will be conducted using the VR app (Figure 3). In this test, 4 spheres are projected to both eyes, with 1 sphere placed directly in front of each eye. One of the 4 spheres will be moved forward along a path perpendicular to the viewing direction, and the participants will be asked to identify which sphere moved. When the sphere moves closer, its size decreases, whereas when it moves farther away, its size increases. This feature prevents the detection of sphere movement based solely on changes in size and ensures that the movement cannot be recognized by a single eye on the side of the moving sphere. This stereotest is a newly developed measure that has not yet undergone formal validation. The results of the VR app–based stereopsis test will be compared with those of the Titmus stereo test [28].

XSL•FO
RenderX

Figure 3. Virtual reality (VR) app–based stereopsis test. One of 4 spheres is moved forward along a path perpendicular to the viewing direction, and participants are asked to identify the moving sphere. To prevent detection of movement based solely on changes in size, the sphere is rendered such that its size decreases when approaching the viewer and increases when receding.

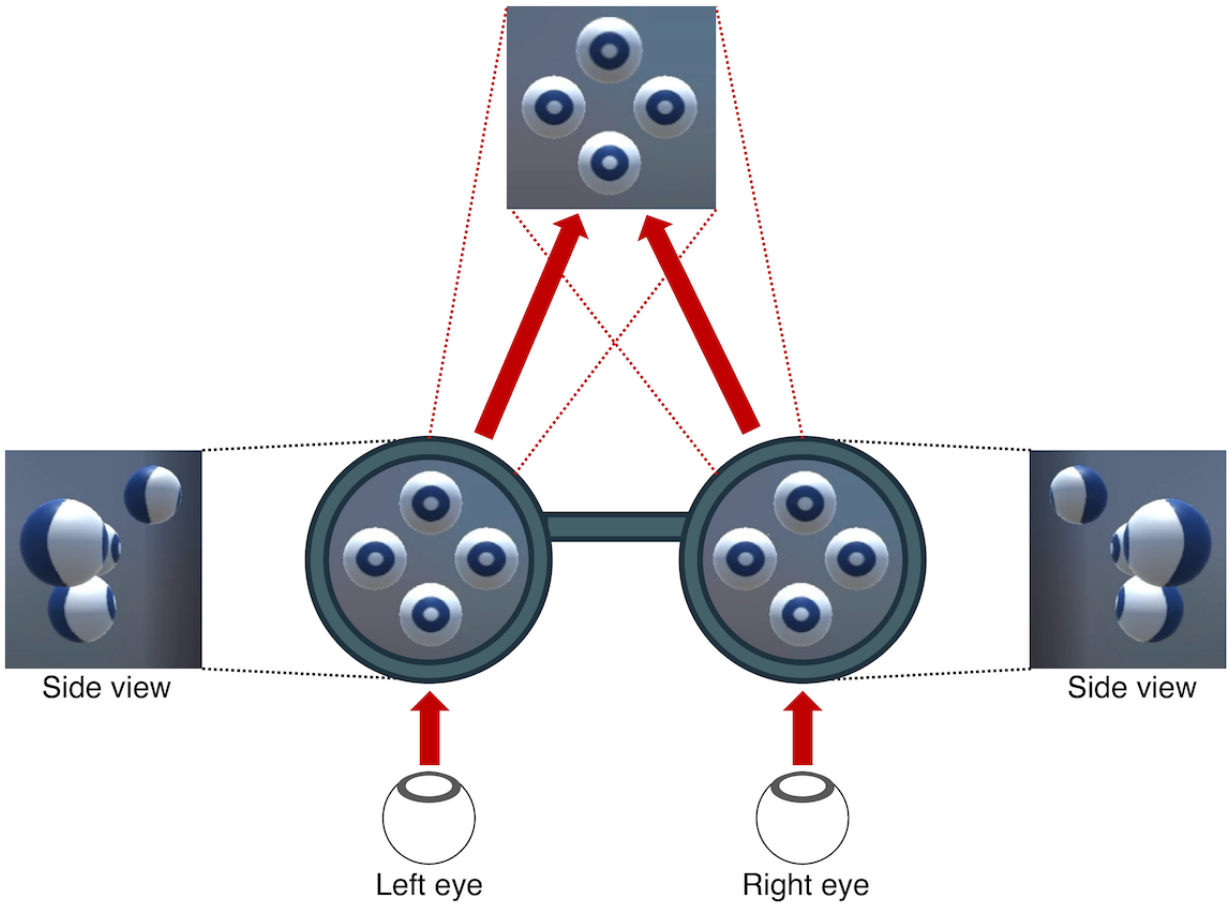

Usability Evaluation

The parents and guardians of patients will respond to a usability questionnaire consisting of 15 fixed-choice statements regarding the VR app. This questionnaire was custom-designed for this study by selecting items from previously published questionnaires that were applicable to smartphone-based apps [30-32] and by adding original items relevant to the present VR app. Therefore, it has not been previously validated and was used for exploratory or descriptive purposes.

The questionnaire will cover 3 domains: game display (8 items), instruction manual (3 items), and overall impression of the parents (4 items), as shown in Textbox 2. In the game display domain, items will evaluate the clarity of background and text colors, the legibility of on-screen text size, the comprehensibility of on-screen text, the visibility of the size of displayed buttons, the comprehensibility of button labels or icons, the recognizability of button locations, the simplicity of operation, and the appropriateness of sound volume. In the instruction manual domain, items will examine the legibility of the manual text, the clarity of the manual content, and the clarity of guidance on operating the VR app. In the overall impression of the parents' domain, items will assess children's speed and ease of learning to operate the VR app, children's ability to use it independently soon after instruction, children's willingness to continue treatment using the VR app, and parents' willingness for their child to continue the treatment using the VR app. Each question will be rated using a 5-point Likert scale (1, strongly disagree; 2, disagree; 3, neutral; 4, agree; and 5, strongly agree) [32,33].

XSL•FO
RenderX

Textbox 2. Usability evaluation questionnaire.

Game display domain

1. The colors of the background and text were easy to see.
2. The size of the displayed text was easy to read.
3. The meaning of the displayed text was easy to understand.
4. The size of the displayed buttons was easy to see.
5. The meaning of the displayed labels or icons of buttons was easy to understand.
6. The location of the displayed buttons was easy to recognize.
7. The operation was simple.
8. The sound volume was appropriate.

Instruction manual

1. The text size in the manual was easy to read.
2. The content of the manual was easy to understand.
3. The manual clearly explained how to operate the VR app.

Overall impression of the parents

1. The child was able to learn how to operate the VR app quickly.
2. The child was able to use the VR app independently soon after learning.
3. The child would like to continue treatment using the VR app.
4. The parent would like the child to continue treatment using the VR app.

Adherence Evaluation

Treatment adherence will be evaluated for VR app usage and eye patch wearing, defined as the ratio of the actual treatment time to the target treatment time. In the VR app group, the target daily treatment time will be 1 hour of VR app usage per day, and the ratio will be calculated by dividing the actual VR app usage time by the total target treatment time (number of participation days in the study multiplied by 1 h). In the control group, the sum of the daily eye patch wearing times prescribed by the physician will constitute the target treatment time. This ratio will be calculated by dividing the actual eye patch wearing time by the total target treatment time throughout the study period.

Primary Outcome

The primary outcome measure will be that the improvement in BCVA in the amblyopic eye from enrollment to 12 weeks in the VR app group is noninferior to the improvement in BCVA in the amblyopic eye from enrollment to 12 weeks in the control group. This outcome measure will be designed to confirm that the amblyopia treatment effect of the VR-based program is not inferior to that of conventional eye patch therapy. In addition, since previous studies investigating amblyopia treatment using a VR headset have shown a significant treatment effect following 12 weeks of use, comparing BCVA at 12 weeks will enable an efficacy assessment of the VR-based program for amblyopia under standard amblyopia treatment conditions, and thus, the 12-week time point will be adopted [19,34].

Secondary Outcomes

Secondary outcomes are as follows:

1. Comparison of BCVA in the amblyopic eye at baseline and at weeks 4, 8, 12, 16, 20, and 24 between and within the control group (eye patch for the sound eye) and the VR app group
2. Comparison of stereopsis in the amblyopic eye at baseline and at weeks 4, 8, 12, 16, 20, and 24 between and within the control and VR app groups
3. In the VR app group, stereopsis in the amblyopic eye measured using the VR-based program for pediatric amblyopia will be compared with stereopsis measured using the Titmus stereo test at baseline and at weeks 4, 8, 12, 16, 20, and 24 during clinic visits
4. Comparison of the ratio of total wearing time to target wearing time (treatment adherence) at weeks 4, 8, 12, 16, 20, and 24 between the control and VR app groups
5. Comparison of the angle of strabismus in the amblyopic eye at baseline and at weeks 4, 8, 12, 16, 20, and 24 between the control and VR app groups
6. In the VR app group, the association between the VR gameplay status (including game type, play date/time, play duration, and game settings) and BCVA in the amblyopic eye, stereopsis, the ratio of total wearing time to the target wearing time, and ocular deviation will be evaluated
7. Malfunctions of the VR-based program for pediatric amblyopia
8. Adverse events
9. Usability evaluation (VR app group only)

XSL•FO
RenderX

To evaluate the effects of amblyopia treatment in the control and VR app groups, BCVA in the amblyopic eye will be assessed as secondary outcome 1 and stereopsis in the amblyopic eye as secondary outcome 2. To assess the reliability and validity of stereopsis measurement using the VR app, secondary outcome 3 will involve comparing stereopsis in the amblyopic eye measured using the VR app for pediatric amblyopia with that obtained during clinical visits using the Titmus stereo test. To evaluate treatment adherence in the control and VR app groups, secondary outcome 4 will involve assessing the ratio of the total wearing time to the target wearing time. To evaluate the safety of amblyopia treatment using the VR app, secondary outcome 5 will involve comparing the angle of strabismus in the amblyopic eye between the control and the VR app groups. To investigate how VR gameplay status correlates with efficacy and safety, secondary outcome 6 will involve examining the VR gameplay status in the VR app group (including game type, play date or time, play duration, and game settings) in relation to BCVA, stereopsis, treatment adherence, and ocular deviation. Secondary outcome 7 will evaluate malfunctions of the VR app for pediatric amblyopia, secondary outcome 8 will monitor adverse events, and secondary outcome 9 will assess usability in the VR app group. Adverse events will be monitored throughout the trial. Based on our safety evaluation of the VR app in adults, where no visually induced motion sickness (VIMS) or vision-related adverse events occurred [18], the intervention is not expected to pose major safety concerns. However, careful monitoring remains essential in pediatric populations. Possible adverse events include VIMS, asthenopia, diplopia, transient blurring, and headaches. All adverse events reported by participants or observed during clinic visits will be documented and evaluated for severity and relationship with the intervention. Among the secondary outcomes, those related to BCVA, stereopsis, ocular deviation, and treatment adherence (secondary outcomes 1 - 4) have been defined as key secondary end points. Since this is an exploratory clinical study, all secondary outcomes other than the primary outcome will be analyzed and interpreted in an exploratory (hypothesis-generating) manner.

Sample Size and Statistical Analyses

The primary outcome of this study will be to determine whether the improvement in BCVA from enrollment to 12 weeks in the VR app group is noninferior to the improvement in BCVA from enrollment to 12 weeks in the control group. Based on the assumption that the difference in mean BCVA at 12 weeks between the control and VR app groups will be 0.07 LogMAR [35], the common standard deviation will be 0.14 [36] LogMAR, the noninferiority margin will be 0.10 LogMAR [37], the significance level will be 2.5%, and statistical power will be 90%, a sample size of 24 participants (12 per group) is required [19]. Anticipating a dropout rate of approximately 20%, the target sample size has been set to 30 participants (15 per group).

To assess whether the VR-based program for amblyopia is noninferior to eye patch therapy, the 95% CI for the change in BCVA between 0 and 12 weeks will be calculated for both the control and VR app groups. If the lower bound of the 95% CI for the change in BCVA in the VR app group is greater than that in the control group, the VR-based program for amblyopia will be considered noninferior to eye patch therapy, and the primary outcome criterion will be met.

For between-group comparisons of BCVA, stereopsis, ocular deviation, and the ratio of total treatment time to the target treatment time, 2-tailed *t* tests or repeated-measures mixed-effects models will be used. Within-group comparisons will be conducted with 2-tailed paired *t* tests, repeated-measures analysis of variance, or repeated-measures mixed-effects models. Reliability and validity evaluations of stereopsis measurements using the VR app compared with the Titmus stereo test administered at clinical visits will be performed using 2-tailed paired *t* tests, Cronbach α coefficients, or intraclass correlation coefficients. The associations between VR gameplay status, ophthalmological examination results, and treatment adherence will be examined using multiple regression analyses or mixed-effects models adjusted for confounding factors. The number of adverse events in the control and VR app groups will be compared using the chi-square test. If continuous variables deviate substantially from a normal distribution, more appropriate methods such as the Mann-Whitney *U* test will be applied. A 2-sided significance level of 5% ($P<.05$) and 95% CI will be used for all analyses. We will not apply multiplicity adjustments across secondary outcomes that represent different endpoint domains. However, when repeated measurements of the same outcome are compared at multiple time points (eg, longitudinal analyses of BCVA), appropriate multiple-comparison adjustments (eg, Bonferroni correction or equivalent family-wise error control) will be applied within that end point family. Primary and secondary analyses will follow intention-to-treat, including all data collected before withdrawal. A per-protocol analysis will be performed as supplementary to assess the robustness of the findings.

Data Management

Data management will be performed by an independent external organization (Micron, Inc.) in accordance with a predetermined data management plan. Data will be collected and linked at multiple time points for each participant. Study investigators, including sponsor-affiliated authors, will not access study data before database lock. After lock, the final dataset will be transferred to an independent statistician within the external organization who has no financial or organizational relationship with the developer and will conduct all analyses. Because the external organization lacks medical expertise in pediatric amblyopia, investigators will help interpret the independently analyzed results and prepare the manuscript. The investigator group will make decisions about reporting and publishing study findings. However, once locked, the dataset cannot be modified; sponsor-affiliated authors cannot alter study data or analysis outputs. The final study report, including results and interpretation, will undergo review by the institutional ethics committee, and the manuscript will undergo independent peer review. These external oversight layers enhance transparency and safeguard research objectivity.

Results

This study will begin on January 1, 2026, at 3 medical institutions in Japan. Patient enrollment will run from January

XSL•FO
RenderX

1, 2026, to September 30, 2026. The data analysis will begin on October 1, 2026, and the results will be reported by March 31, 2027.

Discussion

Pediatric amblyopia can cause an irreversible reduction in visual acuity and has a profound impact on the quality of life. The conventional therapeutic approach, comprising full refractive correction and occlusion of the healthy eye, has established efficacy. However, the substantial psychological and physical burden imposed on affected children predisposes them to poor adherence and compromises treatment effectiveness, thereby undermining the overall outcomes of amblyopia therapy. Therefore, therapeutic modalities that minimize the burden on pediatric patients and concurrently enhance patient adherence and therapeutic efficacy in the treatment of amblyopia are warranted. This protocol paper outlines a multicenter clinical trial designed to evaluate the therapeutic efficacy of a VR app for pediatric amblyopia. The implementation of VR app–based amblyopia training has the potential to improve treatment outcomes by enhancing adherence through gamification and augmenting therapeutic efficacy via hand-eye coordination.

VR-based digital therapeutics are increasingly recognized as promising therapeutic modalities for the management of pediatric amblyopia. In contrast to traditional treatment modalities, VR systems can independently modulate the visual input to each eye, thereby facilitating binocular training, which is regarded as critical in amblyopia therapy [38]. Furthermore, game-based amblyopia training that leverages gamification strategies to engage children's interests is expected to enhance treatment adherence [39]. The objective of this study is to evaluate the therapeutic efficacy of a VR app for pediatric patients with amblyopia, utilizing the distinctive features of VR, in comparison with conventional occlusion therapy using an eye patch on the healthy eye. The VR app delivers game-based tasks that incorporate active bodily movements via hand-tracking controllers. The execution of these tasks necessitates hand-eye coordination. The combination of hand-eye coordination exercises with occlusion therapy in the healthy eye has the potential to augment visual recovery in the amblyopic eye through both direct visual stimulation and facilitation of neuroplasticity [40]. Unlike eye patch occlusion, which entirely blocks visual input to a healthy eye, this method permits simultaneous binocular training, offering the potential for greater visual improvement than conventional amblyopia treatments [41]. The demonstration of the efficacy of VR-based amblyopia training with a VR app in pediatric patients with amblyopia could introduce a novel therapeutic option that integrates the established benefits of occlusion therapy with improved adherence through gamification, promotion of neuroplasticity and visual recovery via hand-eye coordination, and additional visual gains through binocular training. Collectively, these features may drive a paradigm shift in amblyopia treatment.

Several studies have reported the efficacy of VR-based amblyopia training. One study incorporating the VR-based therapeutic device Luminopia One (Luminopia, Inc.) used a method wherein the contrast of streaming content presented to the healthy eye was reduced to encourage viewing with the amblyopic eye. Over a 12-week treatment period, significant improvements in visual acuity and stereopsis were observed, accompanied by high treatment adherence [19,42,43]. As an early example of a VR system, the Interactive Binocular Treatment system, a VR gaming platform employing a stereoscopic display and personal computer, was reported as early as 2006. Using this system, patients underwent treatment once or twice a week for approximately 20 min per session. In a cohort with a mean total treatment duration of 4.4 hours, an average improvement of 10 letters in logMAR visual acuity was achieved [44]. Additional VR-based approaches have been described, such as a photo-hunt game that requires the detection of inconsistencies between images presented to the 2 eyes [45] and NEIVATECH, a serious game–based system intended for binocular training in amblyopia that employs dichoptic presentation of images with differential contrast in a VR environment to implement perceptual learning tasks [41]. Each of these gaming systems seeks to achieve amblyopia training by projecting distinct visual stimuli onto the amblyopic and healthy eyes, thereby enhancing adherence through gamification. Contrastingly, the VR app for pediatric patients with amblyopia used in this study is intended to deliver, in addition to visual stimulation–based amblyopia training and binocular training akin to prior VR-based interventions, the promotion of neuroplasticity via tasks requiring hand-eye coordination. Accordingly, the VR-based SaMD for pediatric patients with amblyopia used in this study may yield superior therapeutic outcomes compared with conventional VR-based amblyopia training systems.

This study has certain limitations. First, despite the explicit inclusion criteria specifying participants with anisometropic, strabismic, or refractive amblyopia and a BCVA of ≤0.5 logMAR in the amblyopic eye, heterogeneity in disease severity, underlying etiology, and associated comorbidities may influence therapeutic outcomes. Second, interindividual differences in age, prior exposure to VR devices, and degree of familiarity may have affected both adherence and therapeutic response, especially during the initial phase of the study. To minimize the influence of such heterogeneity, participants allocated to the VR app group underwent an interim visit at week 2—prior to the first ophthalmologic assessment at week 4—during which VR app usage was reviewed to facilitate the proper continuation of treatment, even among participants with limited prior experience using VR systems. Third, the extended use of the VR app carries the potential risk of adverse effects, including VIMS and asthenopia. Our prior safety evaluation of the VR app in adults revealed no VIMS- or vision-related adverse events [18]. However, data on the long-term use of VR apps in pediatric populations remain limited, and the occurrence of adverse events could potentially affect treatment adherence and diminish the therapeutic efficacy of amblyopia management. Fourth, while amblyopia therapy in clinical settings is generally expected to extend beyond the 6-month duration of this study, this protocol does not permit the evaluation of long-term efficacy, posttreatment sustainability of therapeutic benefits, or the risk of amblyopia recurrence. Upon the completion of this study, an extension study is being conducted to allow interested participants to continue treatment with the VR app. The

XSL•FO
RenderX

implementation of such a study with an extended follow-up would facilitate the assessment of long-term therapeutic efficacy and safety, persistence of treatment effects, and the likelihood of amblyopia recurrence. Fifth, the VR app–based stereopsis test used here was newly developed without formal validation. From a visual psychophysics perspective, this task uses a 4-alternative forced-choice design with size-constancy adjustments to reduce monocular cues. However, the extent of monocular cue elimination remains unverified. The floor and ceiling disparity levels of this stereotest are undetermined. Therefore, the stereopsis outcomes should be interpreted as preliminary. We included an in-headset stereopsis measure as it may provide a practical way to assess stereopsis under intervention conditions and enable monitoring between clinic visits, with potential utility for future clinical implementation and remote assessment in telemedicine.

In conclusion, a VR app may overcome the limitations of conventional amblyopia training by offering a highly effective therapeutic alternative, thereby facilitating a paradigm shift in amblyopia management. Future confirmatory clinical trials are required to delineate the efficacy and limitations of the VR app and to determine its potential integration into standard treatment strategies for pediatric amblyopia.

Acknowledgments

The authors thank the collaborators of the Toho University Omori Medical Center, Teikyo University Hospital, and the CS Eye Clinic for their research support. No generative artificial intelligence or artificial intelligence–assisted technologies were used to write this manuscript.

Funding

This work was supported by the Japan Agency for Medical Research and Development (grant numbers 22he2722008j0001 [TI] and 24he1022014h0001 [TI]), JST FOREST program (JPMJFR234I [TI]), Japan Society for the Promotion of Science, Grants-in-Aid for Scientific Research (grant numbers: 25K20553 [KN], 25K20211 [YO], and 25K12857 [TI]), and Tateishi Science and Technology Foundation Research Grant (A) 2024 (grant number 2241005 [TI]). The funders provided financial support and may have set general programmatic requirements consistent with the grant terms; however, they had no role in data collection, data management, statistical analysis, interpretation of the results, manuscript preparation, or decision to publish. InnoJin, Inc. provided financial support and was involved in study planning, conduct, manuscript preparation, and publication. To safeguard objectivity, investigators (including sponsor-affiliated authors) did not have access to study data prior to the database lock; after the database lock, the final locked dataset was analyzed by an independent statistician at an external organization. Once locked, the dataset cannot be modified.

Conflicts of Interest

TI, YO, and AM-I were the owners of InnoJin, Inc., for developing VR–based SaMD for pediatric patients with amblyopia. KN and MH are employees of InnoJin, Inc. TI reported receiving grants from Kandenko, Co, Ltd, outside the submitted work, as well as personal fees from Santen Pharmaceutical Co., Ltd. and InnoJin, Inc. YO, MH, and AM-I reported receiving personal fees from InnoJin, Inc. The remaining authors declare no conflicts of interest.

References

1. Attebo K, Mitchell P, Cumming R, Smith W, Jolly N, Sparkes R. Prevalence and causes of amblyopia in an adult population. Ophthalmology 1998 Jan;105(1):154-159. [doi: 10.1016/s0161-6420(98)91862-0] [Medline: 9442792]
2. Faghihi M, Hashemi H, Nabovati P, et al. The prevalence of amblyopia and its determinants in a population-based study. Strabismus 2017 Dec;25(4):176-183. [doi: 10.1080/09273972.2017.1391849] [Medline: 29144181]
3. Wu C, Hunter DG. Amblyopia: diagnostic and therapeutic options. Am J Ophthalmol 2006 Jan;141(1):175-184. [doi: 10.1016/j.ajo.2005.07.060] [Medline: 16386994]
4. Matsuo T, Matsuo C. Comparison of prevalence rates of strabismus and amblyopia in Japanese elementary school children between the years 2003 and 2005. Acta Med Okayama 2007 Dec;61(6):329-334. [doi: 10.18926/AMO/32877] [Medline: 18183077]
5. Matsuo T, Matsuo C. The prevalence of strabismus and amblyopia in Japanese elementary school children. Ophthalmic Epidemiol 2005 Feb;12(1):31-36. [doi: 10.1080/09286580490907805] [Medline: 15848918]
6. Ganekal S, Jhanji V, Liang Y, Dorairaj S. Prevalence and etiology of amblyopia in Southern India: results from screening of school children aged 5-15 years. Ophthalmic Epidemiol 2013 Aug;20(4):228-231. [doi: 10.3109/09286586.2013.809772] [Medline: 23865603]
7. Kuman IG, Fedorov SN, Novikova LA. The sensitive period in the development of the human visual system. Sov Psychol 1984 Apr;22(3):35-46. [doi: 10.2753/RPO1061-0405220335]
8. Hu B, Liu Z, Zhao J, et al. The global prevalence of amblyopia in children: a systematic review and meta-analysis. Front Pediatr 2022;10(819998):819998. [doi: 10.3389/fped.2022.819998] [Medline: 35601430]

XSL•FO
RenderX

9. Tamura S, Yoshitake S, Okano M, Uchikawa Y, Onitsuka S, Kawano R. Benefits of consulting a certified orthoptist in a pediatric vision screening program. J AAPOS 2020 Dec;24(6):371-374. [doi: 10.1016/j.jaapos.2020.08.008] [Medline: 33246111]
10. Jarvis SN, Tamhne RC, Thompson L, Francis PM, Anderson J, Colver AF. Preschool vision screening. Arch Dis Child 1991 Mar;66(3):288-294. [doi: 10.1136/adc.66.3.288] [Medline: 2025002]
11. Sen S, Singh P, Saxena R. Management of amblyopia in pediatric patients: current insights. Eye (Lond) 2022 Jan;36(1):44-56. [doi: 10.1038/s41433-021-01669-w] [Medline: 34234293]
12. Erdem E, Çınar GY, Somer D, Demir N, Burcu A, Örnek F. Eye patching as a treatment for amblyopia in children aged 10-16 years. Jpn J Ophthalmol 2011 Jul;55(4):389-395. [doi: 10.1007/s10384-011-0029-z] [Medline: 21647569]
13. Levi DM. Rethinking amblyopia 2020. Vision Res 2020 Nov;176:118-129. [doi: 10.1016/j.visres.2020.07.014] [Medline: 32866759]
14. Maconachie GDE, Farooq S, Bush G, Kempton J, Proudlock FA, Gottlob I. Association between adherence to glasses wearing during amblyopia treatment and improvement in visual acuity. JAMA Ophthalmol 2016 Dec 1;134(12):1347-1353. [doi: 10.1001/jamaophthalmol.2016.3793] [Medline: 27737444]
15. Wang J. Compliance and patching and atropine amblyopia treatments. Vision Res 2015 Sep;114:31-40. [doi: 10.1016/j.visres.2015.02.012] [Medline: 25743080]
16. Le T, Örge F. Treatment compliance in amblyopia: a mini-review and description of a novel online platform for compliance tracking. Surv Ophthalmol 2022;67(6):1685-1697. [doi: 10.1016/j.survophthal.2022.08.003] [Medline: 35970235]
17. Stewart CE, Moseley MJ, Stephens DA, Fielder AR. Treatment dose-response in amblyopia therapy: the Monitored Occlusion Treatment of Amblyopia Study (MOTAS). Invest Ophthalmol Vis Sci 2004 Sep;45(9):3048-3054. [doi: 10.1167/iovs.04-0250] [Medline: 15326120]
18. Hirota M, Okumura Y, Nagino K, et al. Safety evaluation in healthy adults of motion-based virtual reality dichoptic training for pediatric patients with amblyopia: prospective intervention study. JMIR Form Res 2025 Jun 17;9:e69801. [doi: 10.2196/69801] [Medline: 40526901]
19. Xiao S, Angjeli E, Wu HC, et al. Randomized controlled trial of a dichoptic digital therapeutic for amblyopia. Ophthalmology 2022 Jan;129(1):77-85. [doi: 10.1016/j.ophtha.2021.09.001] [Medline: 34534556]
20. Shah S, Shah M, Purohit A, Khanguda V. Perceptual learning with hand–eye coordination as an effective tool for managing amblyopia: a prospective study. Indian J Ophthalmol 2022 Oct;70(10):3630-3633. [doi: 10.4103/ijo.IJO_852_22] [Medline: 36190061]
21. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3. [doi: 10.2196/games.3139] [Medline: 25658754]
22. Watanabe T, Hirowatari H, Tokura Y, Takeda K, Hanaoka H. A 30-day kendama program for community-dwelling elderly: effects on participants' physical and cognitive functions and the practicality of the program. Jpn J Compr Rehabil Sci 2024;15:34-41. [doi: 10.11336/jjcrs.15.34] [Medline: 39473971]
23. Bui Quoc E, Kulp MT, Burns JG, Thompson B. Amblyopia: a review of unmet needs, current treatment options, and emerging therapies. Surv Ophthalmol 2023;68(3):507-525. [doi: 10.1016/j.survophthal.2023.01.001] [Medline: 36681277]
24. Kaur S, Sharda S, Aggarwal H, Dadeya S. Comprehensive review of amblyopia: types and management. Indian J Ophthalmol 2023 Jul;71(7):2677-2686. [doi: 10.4103/IJO.IJO_338_23] [Medline: 37417105]
25. Chaikitmongkol V, Nanegrungsunk O, Patikulsila D, Ruamviboonsuk P, Bressler NM. Repeatability and agreement of visual acuity using the ETDRS number chart, Landolt C chart, or ETDRS alphabet chart in eyes with or without sight-threatening diseases. JAMA Ophthalmol 2018 Mar 1;136(3):286-290. [doi: 10.1001/jamaophthalmol.2017.6290] [Medline: 29346499]
26. Shields MB. The non-contact tonometer. Its value and limitations. Surv Ophthalmol 1980;24(4):211-219. [doi: 10.1016/0039-6257(80)90042-9] [Medline: 6987761]
27. Mori DM, Kuchhangi A, Tame J, et al. Evaluation of a novel virtual reality simulated alternate cover test to assess strabismus: a prospective, masked study. Am J Ophthalmol 2025 Jan;269:266-272. [doi: 10.1016/j.ajo.2024.08.042] [Medline: 39245131]
28. Ancona C, Stoppani M, Odazio V, La Spina C, Corradetti G, Bandello F. Stereo tests as a screening tool for strabismus: which is the best choice? Clin Ophthalmol 2014;8:2221-2227. [doi: 10.2147/OPTH.S67488] [Medline: 25419114]
29. Lebow KA, Campbell CE. A comparison of a traditional and wavefront autorefraction. Optom Vis Sci 2014 Oct;91(10):1191-1198. [doi: 10.1097/OPX.0000000000000378] [Medline: 25198541]
30. Kim JH, Kwon SS, Shim SR, et al. Validation and reliability of a smartphone application for the International Prostate Symptom Score questionnaire: a randomized repeated measures crossover study. J Med Internet Res 2014 Feb 10;16(2):e38. [doi: 10.2196/jmir.3042] [Medline: 24513507]
31. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Using smartphone technology to monitor physical activity in the 10,000 Steps program: a matched case-control trial. J Med Internet Res 2012 Apr 20;14(2):e55. [doi: 10.2196/jmir.1950] [Medline: 22522112]
32. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation in Industry: Taylor & Francis; 1996:189-194 URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke [accessed 2025-12-31]

XSL•FO
RenderX

33. Nagino K, Okumura Y, Yamaguchi M, et al. Diagnostic ability of a smartphone app for dry eye disease: protocol for a multicenter, open-label, prospective, and cross-sectional study. JMIR Res Protoc 2023 Mar 13;12:e45218. [doi: 10.2196/45218] [Medline: 36912872]
34. Xiao S, Gaier ED, Mazow ML, et al. Improved adherence and treatment outcomes with an engaging, personalized digital therapeutic in amblyopia. Sci Rep 2020 May 20;10(1):8328. [doi: 10.1038/s41598-020-65234-3] [Medline: 32433490]
35. Shao W, Niu Y, Wang S, et al. Effects of virtual reality on the treatment of amblyopia in children: a systematic review and meta-analysis. J Pediatr Nurs 2023;72:106-112. [doi: 10.1016/j.pedn.2023.07.014] [Medline: 37494854]
36. Chen CW, Zhu Q, Duan YB, Yao JY. Comparison between binocular therapy and patching for treatment of amblyopia: a meta-analysis of randomised controlled trials. BMJ Open Ophthalmol 2021;6(1):e000625. [doi: 10.1136/bmjophth-2020-000625] [Medline: 33718612]
37. Wygnanski-Jaffe T, Kushner BJ, Moshkovitz A, Belkin M, Yehezkel O, CureSight Pivotal Trial Group. An eye-tracking-based dichoptic home treatment for amblyopia: a multicenter randomized clinical trial. Ophthalmology 2023 Mar;130(3):274-285. [doi: 10.1016/j.ophtha.2022.10.020] [Medline: 36306974]
38. Li SL, Reynaud A, Hess RF, et al. Dichoptic movie viewing treats childhood amblyopia. J AAPOS 2015 Oct;19(5):401-405. [doi: 10.1016/j.jaapos.2015.08.003] [Medline: 26486019]
39. Waddingham P, Cobb SV, Eastgate R, Gregson R. Virtual reality for interactive binocular treatment of amblyopia. Int J Disabil Hum Dev 2006;5(2):155-162. [doi: 10.1515/IJDHD.2006.5.2.155]
40. Coco-Martin MB, Piñero DP, Leal-Vega L, et al. The potential of virtual reality for inducing neuroplasticity in children with amblyopia. J Ophthalmol 2020;2020:7067846. [doi: 10.1155/2020/7067846] [Medline: 32676202]
41. Leal Vega L, Piñero DP, Hernández Rodríguez CJ, et al. Study protocol for a randomized controlled trial of the NEIVATECH virtual reality system to improve visual function in children with anisometropic amblyopia. BMC Ophthalmol 2022 Jun 7;22(1):253. [doi: 10.1186/s12886-022-02466-z] [Medline: 35672688]
42. Xiao S, Gaier ED, Wu HC, et al. Digital therapeutic improves visual acuity and encourages high adherence in amblyopic children in open-label pilot study. J AAPOS 2021 Apr;25(2):87.e1-87.e6. [doi: 10.1016/j.jaapos.2020.11.022] [Medline: 33905837]
43. Miller AM, Bohra LI, Morse CL, et al. Novel digital therapeutic improves visual acuity and encourages high adherence in amblyopic children. J AAPOS 2019 Aug;23(4):e10. [doi: 10.1016/j.jaapos.2019.08.028]
44. Eastgate RM, Griffiths GD, Waddingham PE, et al. Modified virtual reality technology for treatment of amblyopia. Eye (Lond) 2006 Mar;20(3):370-374. [doi: 10.1038/sj.eye.6701882] [Medline: 15832182]
45. Vichitvejpaisal P, Chotined T. Virtual reality photo hunt game application for amblyopia patient. 2017 Presented at: Proceedings of the 21st International Computer Science and Engineering Conference (ICSEC); Nov 15-18, 2017. [doi: 10.1109/ICSEC.2017.8443805]

Abbreviations

BCVA: best-corrected visual acuity
SaMD: software as a medical device
VIMS: visually induced motion sickness
VR: virtual reality

Edited by J Sarvestan; submitted 03.Oct.2025; peer-reviewed by M Piano; revised version received 13.Dec.2025; accepted 18.Dec.2025; published 21.Jan.2026.

Please cite as:
Nagino K, Okumura Y, Hirota M, Matsumura S, Matsumoto T, Negishi T, Midorikawa-Inomata A, Ui M, Hayashi T, Hori Y, Nakao S, Inomata T
Virtual Reality–Based Program for Pediatric Patients With Amblyopia: Protocol for a Multicenter, Randomized, Open-Label, Two-Arm Study
JMIR Res Protoc 2026;15:e85194
URL: https://www.researchprotocols.org/2026/1/e85194
doi:10.2196/85194

© Ken Nagino, Yuichi Okumura, Masakazu Hirota, Saiko Matsumura, Tadashi Matsumoto, Takashi Negishi, Akie Midorikawa-Inomata, Makiko Ui, Takao Hayashi, Yuichi Hori, Shintaro Nakao, Takenori Inomata. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly

XSL•FO
RenderX

cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effect of Helpers Stay Quit Online Training on Preventing Smoking Relapse and Personal Networks: Protocol for a Pragmatic Randomized Controlled Trial and Embedded Mixed Methods Personal Network Study

Myra Muramoto[1], MD, MPH; Allison Hopkins[2], PhD; Christopher McCarty[3], PhD; Alicia Allen[4], MPH, PhD; L Miriam Dickinson[1], PhD; Timothy Connolly[1], RN, MSN; Janet Spradley[1], MSW; Jun Ying[1], PhD

[1]Department of Family Medicine, School of Medicine, University of Colorado, Rm 3613A, 12631 E. 17th Ave, Aurora, CO, United States

[2]Department of Anthropology, Texas A&M University, College Station, TX, United States

[3]Department of Anthropology, College of Liberal Arts and Sciences, University of Florida, Gainesville, FL, United States

[4]Department of Pediatrics, National Center for Opioid Research and Clinical Efficacy, Arkansas Children's Research Institute, University of Arkansas for Medical Sciences, Little Rock, AR, United States

Corresponding Author:
Myra Muramoto, MD, MPH
Department of Family Medicine, School of Medicine, University of Colorado, Rm 3613A, 12631 E. 17th Ave, Aurora, CO, United States

Abstract

Background: Despite major gains in smoking cessation treatment, over half of those who recently quit will relapse within one year. Two systematic reviews of relapse prevention studies reached differing conclusions on the effectiveness of behavioral interventions. Existing relapse prevention evidence is limited by study designs, methodology, and conceptual approaches to behavioral interventions. Personal networks exert powerful effects on initiating and maintaining smoking behavior and can facilitate maintaining abstinence or trigger relapse. To date, relapse prevention interventions have focused on those who are newly abstinent ("abstainers") and have not attempted to influence the abstainer's personal network. The Helpers Stay Quit (Helpers Stay Quit) online training is a conceptually novel "help others" intervention to increase abstainers' public identification as a nonsmoker and their ability to influence those in their personal network to also quit smoking—thereby creating a personal network social environment supportive of long-term abstinence.

Objective: This study is a 2-arm, pragmatic randomized controlled trial (pRCT) testing the hypothesis that quitline abstainers receiving the Helpers Stay Quit intervention will have higher 30-day and 7-day point prevalence of abstinence than those receiving quitline follow-up usual care, and that outcomes may be mediated by characteristics of abstainers' personal networks. The embedded mixed methods study examines the effects of the Helpers Stay Quit intervention on the abstainers' personal network interactions related to smoking and smoking cessation.

Methods: The study design is a 2-group pRCT (N=940) comparing the Helpers Stay Quit online training intervention with a quitline usual care condition. Baseline, 3-, 6-, and 12-month surveys collect data on cognitive and emotional factors potentially influencing relapse. Text messages survey tobacco use status and participants' use of the Helpers Stay Quit training content. The composition and structure of participants' personal networks are assessed at baseline and 12 months. We interviewed 60 participants (both relapsed and abstinent) at differing intervals in the last 6 months of study participation to qualitatively assess personal network influences on relapse or abstinence.

Results: A total of 9 state quitlines are participating by referring potentially eligible clients for screening and potential enrollment. Recruitment began in December 2022. Enrollment of 940 participants was completed in September 2025. When the manuscript was submitted, as of August 31, 2025, 337 participants (65%) had completed the study with 12-month follow-up surveys.

Conclusions: This pRCT tests whether exposure to the Helpers Stay Quit intervention decreases relapse rates of newly abstinent smokers enrolled in state quitline coaching treatment. The embedded personal network mixed methods study is designed to examine the characteristics of a newly abstinent smoker's personal network that may influence relapse and the potential spread of cessation-related information and behaviors through their personal network. The study's design and measures will provide insights into the influences of personal networks and the cognitive and emotional factors impacting the likelihood of relapse or maintaining abstinence in those who are newly abstinent.

Trial Registration: ClinicalTrials.gov NCT05641974; https://clinicaltrials.gov/study/NCT05641974

International Registered Report Identifier (IRRID): DERR1-10.2196/82140

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e82140) doi:10.2196/82140

KEYWORDS

smoking cessation; tobacco cessation; relapse prevention; online training; cessation training; brief intervention; behavioral intervention; mHealth; social network; randomized controlled trial; pragmatic trial; mixed methods

Introduction

Background

Effective tobacco dependence treatments are more numerous and accessible than ever before [1,2]. However, tobacco use remains the leading preventable cause of premature morbidity and mortality in the United States. In addition, cessation treatments are severely underused [3-5], due to both knowledge and attitudinal barriers (eg, low awareness of and misinformation about cessation medications' effectiveness, safety and proper use, and negative attitudes about seeking counseling or other behavioral support for quitting) [6-9]. Despite major gains in the use of smoking cessation treatment, over half of those who have recently quit will relapse within the first year. In an analysis of prospective data from a large representative sample of the general population of the United States, the risk of relapse is more than 50% [10] for abstainers with less than 12 months of abstinence. Identifying effective relapse prevention strategies will enable more individuals to maintain their abstinence beyond a year and, consequently, prevent tobacco-related morbidity and mortality.

Published research on smoking relapse prevention has limitations in methodological design and intervention approaches. Two systematic reviews of relapse prevention interventions reached differing conclusions regarding the efficacy of behavioral interventions [11,12]. Agboola et al [11] concluded minimal self-help intervention was effective. Livingstone-Banks et al [13] concluded with only "moderate certainty" that behavioral interventions were not effective, with the caveat that "further research is likely to have an important impact on our confidence," [13] and a recommendation that future relapse prevention research examine alternative behavioral approaches [13]. A critique of extant research cited methodological weaknesses and limited conceptual approaches to treatment [13]. Different approaches to both studies of relapse prevention and to the interventions themselves are needed to advance the long-term understanding and outcomes of smoking relapse prevention. Livingstone-Banks et al recommended randomization of smokers who were already abstinent ("abstainers") as the strongest design for relapse prevention studies [13]. The protocol for a pragmatic randomized control trial (pRCT) we present here addresses both methodological and treatment approach limitations of extant research.

Conceptual Basis for the Helpers Stay Quit Intervention

Conceptualizing relapse as a dynamic process is the organizing framework [14] guiding the overall Helpers Stay Quit intervention and our evaluation approach. Within this working model, relapse susceptibility or "Relapse Proneness" [15] shifts as a function of physiological, environmental, and cognitive components, each independently contributing to relapse risk at various points in the continuum from quit attempt to long-term abstinence. Relapse proneness results from the interplay of 3 forces—2 of which are cornerstones of relapse research: physical withdrawal symptoms and triggers. The third, more speculative force—"cessation fatigue"—is a latent construct operationalized by declines in motivation, expectancy for success, coping attempts, self-efficacy, and self-control resources [14]. It has been suggested that exploration of "cessation fatigue" may illuminate new possible mechanisms related to relapse [15-17]. To address the dynamic and complex nature of the relapse process and to guide the approach for the proposed pRCT, the conceptual model for the intervention and study uses an integration of interpersonal [18,19], social network [20], social support [21], behavioral theories and concepts and stress and coping theories [22], and Helper Therapy Principle [23] to examine the interrelatedness, time sensitivity, and potential feedback loops related to relapse outcomes [16]. Helpers Stay Quit is novel in its integration of the above theories into an intervention model that is presented as helping others and teaches trainees how to offer a "helping conversation" to help others quit.

Personal Networks and Smoking Behavior

Personal networks exert powerful influences on smoking and quitting behavior [24]. Efforts to increase the use of cessation aids have focused almost exclusively on the smoker [25,26]. Personal networks, wherein abstainers ("egos") are in direct contact with members of their personal network ("alters"), warrant exploration as an avenue for relapse intervention. Personal networks can provide positive feedback and support for abstinence. Importantly, personal networks can also be sources of powerful cues to smoke and ready access to cigarettes—directly impacting relapse risk [27-31]. Personal networks may be an important part of the relapse process, but little is known about how relationships and exchanges of knowledge and information between people in the network feed back into the cessation context. Cessation behavior transmission can occur within three degrees of separation from the original quitter [32], suggesting network-based interventions may be beneficial for initiating quit attempts beyond the abstainer's personal network. In previous work, 89% of general community member participants (n=906) were motivated to enroll in our study to learn tobacco cessation skills to help a friend or family member quit [33]. Furthermore, the majority of participants (up to 80% - 86%) [34,35] reported taking action to help others quit, including offering information and referrals [35,36] primarily to persons in their personal network [35,36].

Preliminary Evidence Supporting Helpers Stay Quit for Relapse Prevention

We conducted a single-group, observational, pilot feasibility study of Helpers Stay Quit, with an embedded personal network study (n=104), to assess design, methods, and procedures in

XSL•FO
RenderX

preparation for the current pRCT and personal network study protocol presented in this paper. Participants were cigarette smokers enrolled in Arizona's state quitline coaching service contacted for routine follow-up approximately 14 - 30 days after their quit date and reporting abstinence from smoking (not even a puff) for at least 14 days. Eligible participants were adults, able to speak and read English, had access to a computer to take the online Helpers Stay Quit training, had a mobile phone to receive and send text messages, agreed to participate in two personal network interviews, and consented to share their quitline client data with the research team. Participants completed online questionnaires at baseline, 3 months, and 6 months, querying smoking status and other behavioral measures related to smoking and cessation. Monthly text message surveys queried smoking status and offering of helping conversations. Personal network telephone interviews were conducted at baseline and 6 months.

We used propensity score matching to compare the quitline clientele with the feasibility study sample at 6 - 7 months. Matches were drawn from deidentified data from the quitline's complete participant pool for the 2-year period corresponding to the study in a ratio of 4 to 1. Variables for matching included age (>50 years vs 18 - 50 years), sex (male vs female—no nonbinary participants were enrolled), race and ethnicity (non-Hispanic White, Hispanic, and non-White non-Hispanic), Fagerström score (>5 vs 1 - 6), and presence of a chronic condition (any vs none). Study results showed Helpers Stay Quit participants reporting higher 30-day abstinence than nonparticipants (86% vs 36%; difference 49%, 95% CI 40%-59%; *P*<.001) [37].

Participants in the longitudinal embedded personal network study (egos) completed a personal network telephone interview in which they named members of their personal network with whom they had at least monthly contact (alters). With each monthly text message survey, egos who reported offering a helping conversation were given a link to an online survey that could be passed to an alter who could respond anonymously. The alter survey queried smoking status, content of the helping conversation, and the alter's reaction to the helping conversation. Results showed that a participant's behavior to help others quit extended to their personal networks, where 89% of abstainers offered helping conversations, and 86% abstained from smoking for the 6-month study duration (n=64) [38]. Abstaining individuals significantly increased the number of nonfamily members (mean difference [MD] 1.41; *P*<.001) and smokers in their networks (MD 0.77; *P*=.006) at follow-up. This change provided participants with more opportunities to carry out helping conversations. After receiving a helping conversation, alters reported positive changes in cessation-related behavior. These findings further support the potential of personal network behavioral interventions, like Helpers Stay Quit, to facilitate smoking relapse prevention.

Research Objectives

This study is designed to assess the effect of Helpers Stay Quit training on the proportion and duration of participants maintaining abstinence or relapsing over time and on participants' personal network structure and interactions related to smoking and smoking cessation. Primary and secondary research questions are as described in study-specific aims (Textbox 1).

Textbox 1. The Helpers Stay Quit study-specific aims.

- Aim 1: Using a 2-group, pragmatic randomized, controlled trial (pRCT; N=940) in which abstinent smokers (2-6 weeks abstinent) are randomized to the Helpers Stay Quit intervention or quitline follow-up usual care, to evaluate the effectiveness of Helpers Stay Quit on:
  - Smoking status (30-day and 7-day point prevalence abstinence) at 6- and 12-month intervals (primary outcome).
  - The number, timing, and duration of participants' lapses and relapses to smoking (secondary outcome).
  - Rates of participants' self-reported delivery of helping conversations to other tobacco users (secondary outcome).
- Aim 2: embed a mixed methods social network study into the pRCT to examine the Helpers Stay Quit effects on:
  - Composition and structure of participants' personal networks.
  - Characteristics of participants' personal networks over the course of the study (eg, proportion of smokers in personal network).
  - Spread of smoking cessation-related information, assistance, and behaviors from participants to others in their personal networks.
  - Recipients of helping conversations report about smoking-related interactions and information exchange with participants and associated cessation-related behavior change.
- Aim 3: using personal network metrics derived from the personal network study (eg, size, composition, and structure), conduct mediational analyses of personal network effects on the primary and secondary study outcomes in Aim 1.

Methods

Study Design

This is a 2-group, parallel pRCT with embedded personal network study comparing: (1) Intervention: Helpers Stay Quit training, and (2) control: quitline postabstinence follow-up usual care. The primary outcome is smoking status at 6 and 12 months. The purpose of the pRCT is to assess if a novel behavioral intervention (Helpers Stay Quit) overall has any effect on relapse in newly abstinent smokers following standard quitline treatment compared to quitline usual care, in real-world conditions of serving state quitline client populations. Relapse and the number of helping conversations will be assessed using a text survey weekly for the first 6 months and biweekly for months 7 - 12. The prestudy quitline treatment, quitline referral, study

XSL•FO
RenderX

recruitment, screening, consent, and prerandomization and run-in period are depicted in Figure 1. The postrandomization study intervention exposure, assessment intervals, and qualitative interviews are depicted in Figure 2.

Figure 1. Timeline for recruitment, screening, consent, and prerandomization run-in for the helpers stay quit pragmatic randomized controlled trial and embedded personal network assessment (PNA).

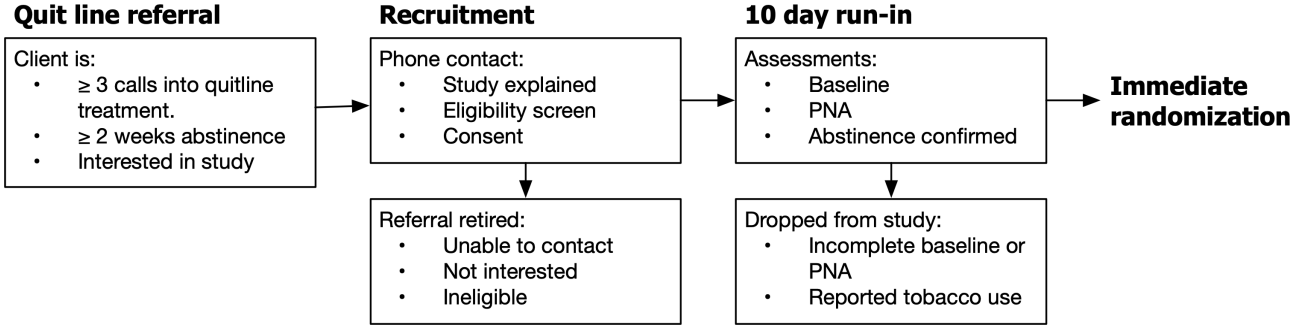

Figure 2. Postrandomization study timeline and data collection schedule for the Helpers Stay Quit (HSQ) pragmatic randomized controlled trial.

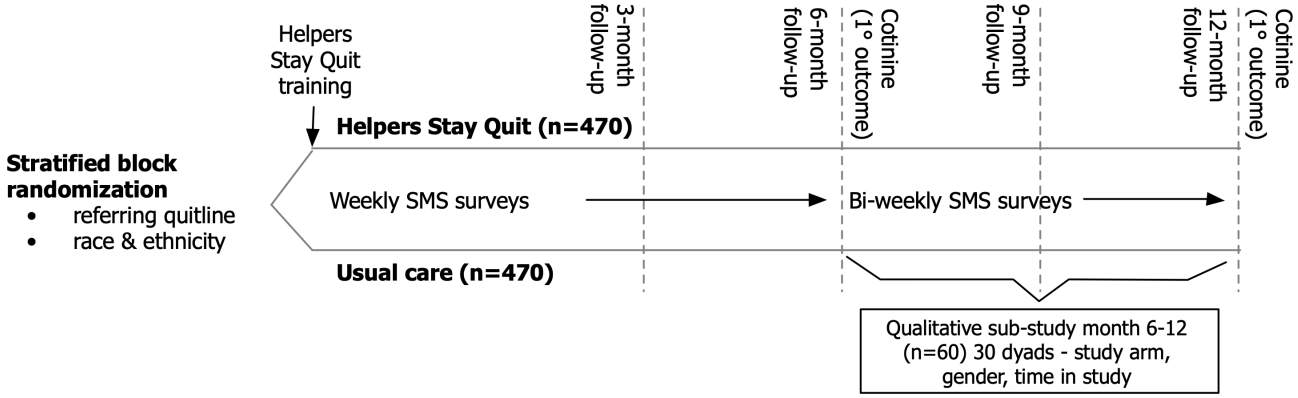

Ethical Considerations

The Helpers Stay Quit study was approved by all participating institutions' institutional review boards (IRBs), including the Combined Multiple Institution Review Board of the University of Colorado (21 - 4624), Arizona Department of Health Services (HSRB # 23 - 0016), Kentucky Cabinet for Health and Family Services IRB (CHFS-IRB-DPH-FY23-35), Massachusetts Department of Public Health (MDPH IRB 00000701), Ohio Department of Health Human Subjects IRB (2023 - 21), Pennsylvania Department of Health IRB (2022 - 065), and Utah Department of Health and Human Services IRB (686).

The IRBs for Texas A&M University, University of Florida, Kansas Department of Health and Environment, Michigan Department of Health and Human Services, and the Vermont Agency of Human Services each determined that the Helpers Stay Quit study could be implemented without additional IRB review.

Informed consent was conducted by telephone; participants could see the consent form on a screen, could electronically sign the form, and were emailed a copy of their signed consent. As part of the informed consent process, participants were read paragraphs from the consent form as shown in Multimedia Appendix 1.

Study participants' confidentiality is protected by the following measures: all study participants are assigned a numerical study ID. The linking file that contains both study ID and participant-identifying information (name, address, telephone number, and email) is stored on a secure server in a physically secured server room. The linking file is only accessible by research staff employed by the University of Colorado. Data files shared with co-investigators, statisticians, and analysts are deidentified.

Participants completing every protocol activity and qualifying for every incentive could receive a total of US $419. A breakdown of the study protocol items and their compensation rates is as follows: baseline (US $20), social network assessment 1 (US $25), training (US $40), incentive for completing training within two weeks (US $10), 3-month survey (US $20), 6-month survey (US $20), 9-month survey (US $20), 12-month survey (US $20), bonus for completing all 4 surveys (US $5), social network assessment 2 (US $25), qualitative personal network interview (US $30), weekly text survey (24 total; US $4 each), biweekly text survey (12 total; US $4 each), 6-month follow-up specimen completion incentive (US $15), 12-month follow-up specimen completion incentive (US $15), and maximum completion bonus for completing all study activities (US $10).

Participant Recruitment, Eligibility, Enrollment, and Randomization

Recruitment

Participants are referred from state quitlines based on a prescreening algorithm programmed into the coaching platform that identifies potentially eligible quitline callers self-reporting abstinence during their third, fourth, or fifth coaching call, and interested in being referred to the study. Upon referral from the

XSL•FO
RenderX

quitlines, study staff contact participants via telephone to screen for initial study eligibility, provide the participants with a detailed explanation of the study's purpose and procedures, and obtain electronic informed consent from those who are eligible and wish to proceed to the prerandomization run-in period.

Eligibility Criteria

The eligibility criteria are shown in Textbox 2.

Textbox 2. Eligibility criteria.

Inclusion criteria:

- Participants are aged 18 years and older.
- Abstinent from smoking for between 14 and 60 days per self-report.
- Primary tobacco use is cigarettes.
- Has access to the internet via computer or mobile device.
- Self-described proficiency with English.
- Willing and able to send and receive weekly text messages using personal mobile phone.
- Will allow the quitline to share its client data with the research team.
- Willing to complete online surveys at baseline, 3, 6, 9, and 12 months.
- Willing to self-collect dried blood spot for biological confirmation of smoking status and send back to the research team.
- If assigned to the Helpers Stay Quit condition, willing to complete the training within 14 days.
- If selected, willing to participate in a qualitative interview.
- Willing to forego any other training for tobacco cessation intervention or support (ie, to become a cessation counselor or facilitator or support person, eg, "quit buddy") for the duration of their study enrollment.

Exclusion criteria:

- Any previous exposure to Helpers Stay Quit or other cessation training in the previous 2 years.
- Relapse within 10 days (run-in phase) before enrollment and randomization.
- Participants are not precluded from continuing their personal tobacco cessation activities with quitline or other services or aids (eg, support groups, medication, etc).
- Concomitant use of cessation aids is tracked while participants are enrolled in the study via online surveys (baseline, 3, 6, 9, and 12 months, and weekly or biweekly text surveys).

Prerandomization Run-in/Baseline Measures

A major strength of the Helpers Stay Quit study design is the clear separation of quitline cessation treatment from the Helpers relapse prevention intervention. This separation is maintained by requiring that all participants attest to their continued abstinence immediately before randomization. After initial study eligibility screening, consented participants have up to 10 days to complete study run-in baseline measures, which include a survey of cognitive and emotional measures (Table 1) and a personal network assessment (PNA). Those who fail to complete run-in baseline procedures are not randomized into the study. Immediately before randomization, participants attest to their continued abstinence (7-day point prevalence; Figure 1).

XSL•FO
RenderX

Table . Data collection schedule for the Helpers Stay Quit study pragmatic randomized controlled trial.

Data collected	Baseline	Training pre-post	Weeks 1 - 24	3-month follow-up	6-month follow-up	Weeks 25 - 52	9-month follow-up	12-month follow-up	Dyad qualitative interviews, months 5 - 12
Tobacco use or abstinence (1° outcome)	✓		✓	✓	✓	✓	✓	✓	
Cotinine					✓			✓	
Helping conversations (2° outcome)			✓	✓	✓	✓	✓	✓	
Helping conversation content				✓	✓		✓	✓	
Personal network assessment	✓							✓	
Cessation knowledge		✓							
Cessation fatigue	✓			✓	✓		✓	✓	
Abstinence Related Motivational Engagement-Short Form	✓			✓	✓		✓	✓	
Smoker or quitter identity	✓			✓	✓		✓	✓	
Smoker group identity	✓			✓	✓		✓	✓	
Residual attraction or vulnerability to smoking	✓			✓	✓		✓	✓	
Situational smoking abstinence self-efficacy	✓			✓	✓		✓	✓	
Proactive Coping Inventory	✓			✓	✓		✓	✓	
Patient Health Questionnaire-4	✓			✓	✓		✓	✓	
Qualitative aspects of network change, continuity									✓

Randomization

Participants are randomized using stratified block randomization to ensure balance between study arms within key subgroups. Randomization is stratified by quitline and race and ethnicity (categorized as non-Hispanic White, Hispanic, non-Hispanic Black, and other). Within each stratum, participants are assigned to intervention or control groups in random block sizes of 4 and 6. The use of variable block sizes minimizes the predictability of treatment allocation while maintaining approximate balance

XSL•FO
RenderX

in group assignments throughout the enrollment process. Randomization is implemented using a computer-generated randomization schedule prepared by the study biostatistician and concealed from study personnel involved in participant recruitment and assessment.

Intervention and Comparison Conditions

Helpers Stay Quit Intervention

The Helpers Stay Quit intervention is a web-based tobacco cessation brief intervention training designed specifically for lay community members interested in helping someone in their lives quit tobacco. The multimedia, interactive, web-based training emphasizes a tobacco user–centered, nonconfrontational approach to encouraging others to quit. Helpers Stay Quit trainees (ie, "Helpers") learn how to offer a 4-step "helping conversation" to encourage quitting and use of evidence-based cessation aids (eg, cessation medications, and quitlines), and to offer referrals to quitlines and reliable information sources (eg, pharmacists, National Cancer Institute's SmokeFree website). The 4 steps of a helping conversation parallel the 4 core training modules of Helpers Stay Quit (awareness, understanding, helping, and relating), incorporating key components of the Public Health Service guideline's recommended 5As [39].

A key learning objective of Helpers Stay Quit training is for trainees to learn how to manage their own expectations for the process and outcome of a helping conversation. For example, they cannot "make" anyone quit, but they can offer nonjudgmental encouragement and support and information about effective cessation aids. Helpers are taught to avoid nagging, confrontation, or pushing anyone to quit. The helping conversation focuses on encouraging behavior change that is aligned with the tobacco user's current willingness and readiness to take any action toward quitting. Helpers Stay Quit includes content on evidence-based cessation aids and sources of reliable information about quitting. The training includes video role-plays of helping conversations and testimonials of current and former tobacco users and "concerned others" to illustrate core concepts. Learning activities prompt personal reflection and require application of new knowledge through scenario-based quizzes and decision-testing (eg, tobacco use, benefits of quitting) and skills (eg, recognition of appropriate use of communication techniques, identifying optimal or suboptimal responses to helping conversation interactions). Participants also have access to downloadable "Helpers tools," which are simple brochures on topics such as "Thinking of Quitting" and "Benefits of Quitting." The training takes approximately 2.5 hours to complete and includes a pre- and posttest and a downloadable certificate for successful training completion.

Quitline Usual Care Comparison

Quitline posttreatment usual care for our participating quitlines is to contact clients for follow-up assessment of abstinence at 7 months after enrollment in services, regardless of the number of calls completed. If the client has relapsed, the quitline attempts to reengage the client in cessation services. The participating quitlines each have a website that accompanies and complements their telephonic services. The websites contain information about tobacco use, tobacco cessation, descriptions of telephonic quit coaching services, and how to access quit coaching and cessation medications offered by the quitline. None of the quitline websites offer information specific to how to help others quit tobacco.

Measures and Data Collection for Pragmatic Randomized Controlled Trial (pRCT)

Quitline Client Data

Participants consent for their quitline data to be shared with the study team including demographics (age, gender, and race and ethnicity), education, income level, presence or absence of behavioral health or chronic disease conditions, zip code, detailed tobacco use history and current use, time to first cigarette, use of electronic nicotine delivery systems, cessation medications provided and used, and number of quitline coaching calls completed.

Relapse Status (Primary Outcome) and Helping Conversations (Secondary Outcome)

Smoking or other tobacco use, electronic nicotine delivery device use, use of cessation aids, and helping conversations are assessed by online surveys at baseline, 3, 6, and 12 months, and by weekly text message surveys for months 1 - 6, then biweekly text surveys for months 7 - 12. Relapse is defined as (1) any smoking on 7 consecutive days or (2) smoking at least once each week over 2 consecutive weeks. Multiple relapses are defined as relapsing, reporting being quit (7-day point prevalence), and then subsequently reporting smoking again. Offers of helping conversations are assessed by baseline survey and weekly text message surveys (months 1 - 6), then biweekly text surveys (months 7 - 12). A 5-point Likert scale queries confidence in helping conversations, for example, talking to others about quitting smoking. Content of typical helping conversation (eg, setting quit date, quit plan, cessation medication, quitline referrals, and avoiding relapse) is queried via a yes or no checklist at baseline, 3, 6, 9, and 12 months (Table 1).

Cognitive or Emotional Measures

A dynamic model of smoking relapse is the organizing framework guiding the overall intervention and evaluation approach [14]. Selected measures evaluate possible factors in "cessation fatigue," a latent construct proposed to contribute to "relapse proneness" and operationalized by declines in motivation, expectancy for success, coping attempts, self-efficacy, and self-control resources [14]. To explore the concept of "cessation fatigue" [16], the following patient-reported outcome measures are collected at baseline, 3, 6, 9, and 12 months post study enrollment: cessation fatigue [17], Abstinence Related Motivational Engagement – Short Form [40], smoker versus quitter identity [41], smoker group identity [41], residual attraction and vulnerability to smoking [42], situational smoking abstinence self-efficacy [43], Proactive Coping Inventory [44,45]; and Patient Health Questionnaire-4 for anxiety and depression [46-48] (see Table 1).

XSL•FO
RenderX

Biochemical Confirmation of Abstinence

Participants who self-report abstinence on 6- and 12-month surveys are mailed a kit for self-collection of the dried blood spot. Participants are trained on self-collection of dried blood spots using an online video and instructional booklet [49]. Dried blood spot samples will be analyzed for cotinine to confirm self-report, following previously validated approaches [50] (Table 1).

Measures and Data Collection for Embedded Mixed Methods Personal Network Study

Overview

Personal network recruitment, screening, and inclusion and exclusion criteria are the same as for the main study. All main study participants (egos) complete two online personal network assessments, one at baseline and the other at 12 months post baseline. Personal network assessment questions address network composition and structure and are administered using REDCap (Research Electronic Data Capture; Vanderbilt University) [51-54]. Each participant (ego) is asked to name 25 individuals (alters) with whom they have had any kind of contact over the past six months [55,56]. Egos are then asked questions about each alter to capture the compositional variables, such as the proportion of individuals who smoke in their personal network and the social support and influence for abstinence. Personal network structure is defined by asking the participant (ego) whether each unique pair of alters talks to each other independently of ego [52].

Alter Surveys

We use respondent-driven sampling [57-62] to survey smoking alters who receive a helping conversation from the ego to measure the spread of smoking cessation information and behaviors. With each weekly (months 1‐6) or biweekly (months 7‐12) period, egos also receive a text message link to an anonymous online survey (alter survey) that can be passed on to alters for completion. Alter surveys query smoking status, reaction to the helping conversation, context and content of the helping conversation, and if the alter has shared helping conversation information with others in their own personal network.

Personal Network Qualitative Interviews

A subsample (n=60) was selected for qualitative interviews about relapse or abstinence and the qualitative aspects of personal network continuity or change [63,64], interactions with members of personal network, exchange of information, assistance, or goods related to smoking cessation. Participants who had completed at least 5 months of the study were selected in 30 dyads (a dyad is one relapser paired with a like abstainer or maintainer from the same study arm, matched on gender). The 15 relapser or maintainer dyads from each arm are matched on the approximate number of weeks participating in the study. Network visualizations on connectedness, alter centrality, closeness, interaction, demographics, and smoking status are generated using the R software (R Foundation for Statistical Computing) package *ggraph* and shown to participants to facilitate discussion in qualitative interviews.

Sample Size, Power Estimates, and Analysis

Power and Sample Size for Pragmatic Randomized Controlled Trial (pRCT)

A sample size of 940 participants would yield 80% power to detect a difference in smoking cessation rates between arms of 10%, assuming a conservative control arm rate of 0.25 [65], 30% dropout, and a noncontinuity corrected χ^2 test. This difference is smaller than that found in the preceding pilot feasibility study (difference=46%); this effect size was chosen as a clinically significant difference with public health implications and is feasible, given the pilot results. This sample size also yields 90% power to detect standardized effect sizes for continuous secondary outcomes of Cohen d=0.26, a small-to-medium effect size, and considered to be a good estimate for a minimum important difference for patient-reported outcomes.

Personal Network Study Sample Size

To measure change in personal networks at two times, 940 participants (470 in each arm) will be sufficient. Although there are few longitudinal personal network studies among those with substance use disorders or those who have recently quit using as models [55], studies using dynamic network analyses to estimate the impact of life events on an individual's social context typically include between 33‐250 participants [66-72]. Informed by the previous feasibility study, the present personal network study features an increased sample size, a lengthened interval between the first and second personal network assessment to 1 year, and broadened the personal network name generator question to capture both strong ties and weak ties. The qualitative sample (n=60) is theory-driven; sufficient for coverage of basic consensus and range of variability of critical themes [73].

Analytic Plan

General Quantitative Approaches

For this pRCT, descriptive statistics (mean, SD, and frequencies) will be computed for baseline participant characteristics. Participant-level covariates will be screened in bivariate analyses and included in multivariate analysis if they are related to the outcome at P<.2, are associated with dropout, or used in the randomization procedure. Before beginning the analyses described below to address study hypotheses, we will compare baseline characteristics of participants who drop out to participants with complete data to assess the missingness mechanism (missing completely at random [MCAR], missing at random [MAR], and missing not at random [MNAR]) and generalizability of our results [74-76]. If missingness is ignorable (MCAR and MAR), we will use likelihood-based methods that use all available data, adjusting for covariates that are associated with dropout. If missingness is nonignorable, we will use pattern mixture models. Sensitivity analyses with multiple imputation will be undertaken if missing data rates for the primary outcome are greater than 10% at any follow-up time used in the analysis.

XSL•FO
RenderX

Aim 1: Abstinence and Relapse

Abstinence rates at 6- and 12-months postrandomization and biochemically confirmed abstinence at 6- and 12-months postrandomization will be compared between arms using multivariable logistic regression for each outcome (6- or 12-months). Independent variables will include intervention status, strata (race and ethnicity), quitline, and covariates (as described above). We will compare time to first relapse using Cox models and Kaplan-Meier curves. The number of relapses and helping conversations will be compared between study arms using Poisson regression, with the zero-inflated Poisson or negative binomial model used if appropriate. We will carry out secondary analyses using general linear mixed models to examine change in cognitive and emotional measures of participant-reported outcomes (described above in the General Quantitative Approaches section) for participants randomized to receive Helpers Stay Quit compared to participants randomized to the usual care arm.

Aim 2: Personal Network Study

We will use quantitative personal network analysis to address Aims 2a, b, and c. Personal network measures consist of compositional variables, which describe who is in the network, and structural variables, which measure how network alters are arranged around the ego. Personal network measures will be obtained at two points, baseline and 12-month follow-up. Compositional variables for each alter include: demographic data (ie, sex and age); relationship to participant (ie, relationship, how close, frequency of interaction, and smoking- and quitting-related support); and smoking-related information (ie, use of tobacco products, previous and current smoking status, and quit attempts). Structural variables will be calculated in R package *egor* based on alter pairs that very likely interact and include: density (ie, the extent to which all possible connections between alters are present), alter centrality measures (ie, to identify the most central individuals in the network), network centralization (ie, the extent to which a network is dominated by one or a few alters), and measures of subgroups (ie, to identify cohesive or disconnected subgroups). Additionally, some combined compositional and structural variables will be calculated with smoking status and alter centrality measures.

Descriptive statistics will be run on all compositional and structural personal network variables at both time points. We will assess using analysis of variance differences in network measures between the Helpers Stay Quit (vs usual care completers, completers who carried out helping conversations (vs no helping conversations), and completers who relapsed (vs abstained). Paired samples *t* tests will be used to compare network measures between the two different points.

Qualitative analysis will be used to address Aims 2a, b, and c, along with the quantitative personal network analysis. Participants were shown a series of network visualizations to facilitate discussion during the interview. Interviews are transcribed, using a verbatim transcription protocol, and analyzed using a combination of a priori (theory-based) coding combined with emergent codes to produce themes, which are used to assign higher-level meaning to the data [77-79]. Manual coding using ATLAS.ti (Scientific Software Development GmbH) will be compared to AI coding using natural language processing in R. Both descriptive and interpretive codes will be used in this analysis, with development of additional codes for emergent themes and issues [79].

Aim 3: Mediational Analysis

Using data from Aims 1 and 2, we will use multiple mediation models to investigate mechanisms of Helpers Stay Quit. Specifically, we will estimate the direct effects of the intervention on abstinence, as well as the indirect effects that are mediated through the personal networks of participants and longitudinal change in scores collected through surveys. Potential candidates for mediation include the number of helping conversations, change in personal network variables (smoking status, smoking status and degree centrality, smoking status and sex, density, and degree centralization), and change in patient-reported outcome variables from the cognitive and emotional measures. To narrow down the list of potential mediators from the conceptual framework, we will first examine relationships between the intervention arm and the mediator, as well as the mediator and the outcome variable. Variables associated with both the intervention and abstinence, with a *P* value of .1, will be included in this analysis as potential mediators. Variables associated with abstinence but not treatment arm at the 0.1 alpha level will be included as a covariate but not explored further as a mediator. For those who meet the initial criteria for a mediator (associated with both intervention arm and outcome), we will examine direct and indirect effects using causal mediation modeling [80].

Results

This trial was funded in 2021 and is currently ongoing. Initial study start-up was delayed when the first author, MM, moved to the University of Colorado. Recruitment began in December 2022, and we completed enrollment and randomization of 940 participants on September 15, 2025. As of August 31, 2025, when the paper was submitted, 337 participants (65%) had completed the study with the 12-month follow-up surveys.

Discussion

Principal Findings

We describe the protocol for a rigorous evaluation of a conceptually novel behavioral intervention for smoking relapse prevention with a study design and intervention addressing methodologic weaknesses and limited conceptual approaches to treatment in extant research [81].

Treatment approaches described in the literature have focused on skills-based interventions that encouraged identifying and resolving tempting situations, and minimal interventions using single sessions or written materials [81]. All behavioral interventions had abstainers apply skills to themselves [81]. Methodological limitations included small sample size, lack of clear separation between treatment and relapse prevention interventions, or insufficient detail to permit survival analyses [81]. Randomization of smokers who were already abstinent is recommended as the strongest design [81]. This ongoing pRCT addresses both methodological and treatment approach

XSL•FO
RenderX

limitations of extant research, with a clear separation of cessation treatment from relapse prevention by randomly assigning smokers who have quit for at least 14 days. Helpers Stay Quit's integration of behavioral theories and conceptual models is a conceptually novel approach to behavioral intervention, not previously studied.

Helpers Stay Quit's "help others" presentation encourages abstainers to help other smokers to quit. Helpers Stay Quit integrates multiple behavioral theories into a novel intervention model that is presented as "help others." Helpers Stay Quit teaches knowledge and skills to help others to quit tobacco, and thus could potentially increase an abstainer's ability to influence their personal network in ways supportive of prolonged abstinence, for example, transmitting cessation behavior to smokers in the personal network [32,82]. This approach could potentially reinforce an abstainer's identity as a nonsmoker and help shift the personal network social environment toward nonsmoking. Smoking-related identities and smoking behavior are reciprocally interrelated [41,83]. By offering helping conversations to others in their personal network, an abstainer publicly declares and reinforces a "quitter" or "non-smoker" social identity (vs smoker identity) within their personal network and to the abstainer themselves, creating a social expectation that favors abstinence and identification as a nonsmoker over time [84]. Engaging in prosocial helping behavior may be an alternate or additional proactive coping strategy for abstainers to deal with stressors potentially impacting relapse [85-87].

Research on the effect of social networks on smoking cessation and relapse among adults in the general population is limited. Most current studies of adult smoking and social networks focus on large online communities of tobacco users engaged in social media related to smoking and smoking cessation, for example, QuitNet [32,88-91], Become an EX [92,93], and Facebook pages with a smoking-cessation theme [94]. Much of the other research on smoking and social networks addresses adolescent or young adult smokers and smoking initiation [95] or pregnant and postpartum women [81]. Stone and colleagues [96] identified two different network approaches used in studies of personal networks and recovery from substance use, and by extension, smoking cessation: network predicting outcome and treatment predicting network change. The network predicting outcome studies are more prevalent and focus on the changes in the smokers' personal network after cessation in a general population [90], during pregnancy [27], or among adults with serious mental illness [97]. Treatment-predicting network change studies are less common and include Helpers Stay Quit. Helpers Stay Quit is an intervention ("treatment") intended to facilitate ("predict") network change through helping conversations. There are several benefits to this approach. First, it allows participants to maintain their network members, which was identified as a factor in Nguyen and colleagues' [27] study on why women did not stay quit after they gave birth. It also potentially reduces the participants' contact with people who smoke by assisting their network members in quitting smoking, which was an important factor for those who stayed quit in Bray and colleagues' study [90]. Finally, it leverages the valuable resource that people who quit smoking become for assisting other network members in quitting [97].

Strengths and Limitations

The Helpers Stay Quit study's strengths include a methodologically rigorous design, a pRCT with "real world" participants recruited from multiple state quitlines, and the embedded mixed methods personal network study. The study's innovation, described below, is another key strength.

Helpers Stay Quit integrates multiple behavioral theories into an intervention model that is innovative both in its theoretical approach and its "help others" presentation. The beneficial effects of helping others on the helpers themselves [23] have been demonstrated in diverse other groups well beyond mental illness and addiction treatment. Yet, the impact of helping others quit on the newly abstinent smokers themselves has not been previously studied. Also novel is the integration of behavioral theories from interpersonal [18,19], social network [20], and social support [21], stress and coping theories [22], and Helper Therapy Principle [23] into a multifaceted intervention addressing relapse proneness within a framework of relapse as a complex and dynamic process. This is one of the first attempts to include the structural and compositional properties of a personal network as a covariate in smoking cessation.

Activating abstainers to offer helping conversations within their personal network could create new avenues of exploration for relapse prevention. This study examines the effect of the Helpers Stay Quit intervention on the structure and composition of abstainers' personal networks. By measuring personal network changes over time, we can assess if abstainers could potentially reduce their risk of relapse by decreasing the number of smokers in their personal network, thereby decreasing exposure to smoking cues, access to cigarettes from other smokers, potentially shifting their personal network's normative behaviors toward nonsmoking, and overall creating a social environment that is more conducive to maintaining abstinence [16]. Abstainers may also be able to make changes in their relationships that affect the position of smokers within their networks, such as making smokers less central.

The study's limitations include the following: the Helpers Stay Quit intervention is currently only available in English, so participants must be proficient in English. The state quitline prescreening and referral algorithm only identifies clients who are abstinent and have completed 3 or more coaching calls, so we will be unable to assess intervention impact on clients who could be abstinent after fewer than 3 calls. The Helpers Stay Quit intervention is an integration of concepts from interpersonal [18,19], social network [20] and social support [21], stress and coping theories [22], and Helper Therapy Principle [23] into a multifaceted intervention model collectively presented to participants as "helping others." This first trial after our feasibility trial is designed to assess whether Helpers Stay Quit overall has any effect on relapse, not which of the theoretical concepts or behavioral mechanisms has an effect. Thus, we will be unable to discern if there is a differential effect from a particular intervention component. If Helpers Stay Quit overall has a favorable effect on relapse, future studies with different designs and comparator conditions could examine which aspects of the model may be most salient. National Jewish Health operates the quitline for all the participating states, so we will

XSL•FO
RenderX

not know if results from the Helpers Stay Quit study would also apply to other states with different quitline vendors. As with any study with inclusion and exclusion criteria, and participant self-selection to enroll, the study sample may differ from the general population of quitline clients who did not qualify or chose not to enroll. The study is recruiting from clients enrolled in state quitline services, and the majority of quitline clients are women. Since the majority of smokers attempting to quit do not engage with quitlines, study findings may not generalize to the general population of smokers.

Conclusions and Future Work

In summary, this study aims to address current limitations in smoking relapse prevention research using the recommended strongest study design and a conceptually novel approach and presentation of an online behavioral intervention. Should Helpers Stay Quit demonstrate a beneficial effect on smoking relapse, the low marginal cost and scalability of a web-based intervention, and the ease of integration into existing cessation services increase the potential reach. Even small effects can be meaningful on a population level if the intervention can be delivered efficiently at a large scale and low cost. Follow-up papers will report on trial results following the analytic plans described in this paper. Findings from this trial will help advance scientific knowledge on smoking relapse prevention from different perspectives and domains of possible inquiry.

Acknowledgments

The authors are grateful for the recruitment partnership of National Jewish Health, and the study participation of Arizona Department of Health Services, Kansas Department of Health and Environment, Kentucky Cabinet for Health and Family Services, Massachusetts Department of Public Health, Michigan Department of Health and Human Services, Ohio Department of Health Human Subjects, Pennsylvania Department of Health, Utah Department of Health and Human Services, and the Vermont Agency of Human Services. We are especially grateful to our research participants, the students and research assistants on the Helpers Stay Quit Study team, and Elizabeth Staton for her editorial assistance.

Funding

The Helpers Stay Quit study is funded by the National Cancer Institute (grant R01CA248658) for a 6-year period, beginning in March 2021. The National Cancer Institute had no involvement in the study design, data collection, analysis, interpretation, or writing of this manuscript. The study was originally recommended for a total of US $1,859,408 in funding before subsequent funding cuts.

Authors' Contributions

Conceptualization: MM, AH, CM, AA, TC
Methodology: MM, AH, CM, AA, TC
Funding acquisition: MM, AH, CM, AA, TC
Resources: MM, TC
Investigation: MM, AH, CM, TC, JS
Project administration: MM, TC, JS
Supervision: MM, TC, JS
Data curation: LMD, TC, JY
Formal analysis: LMD, JY
Writing – original draft: MM, AH, CM, AA, LMD, TC, JS, JY

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consent form paragraphs read to participants as part of the informed consent process.
[DOCX File, 14 KB - resprot_v15i1e82140_app1.docx]

Checklist 1
SPIRIT checklist.
[DOCX File, 34 KB - resprot_v15i1e82140_app2.docx]

References

1. Tobacco Use and Dependence Guideline Panel. Treating tobacco use and dependence: 2008 update. : US Department of Health and Human Services; 2008.

XSL•FO
RenderX

2. Zhu SH, Lee M, Zhuang YL, Gamst A, Wolfson T. Interventions to increase smoking cessation at the population level: how much progress has been made in the last two decades? Tob Control 2012 Mar;21(2):110-118. [doi: 10.1136/tobaccocontrol-2011-050371]
3. Edwards SA, Bondy SJ, Callaghan RC, Mann RE. Prevalence of unassisted quit attempts in population-based studies: a systematic review of the literature. Addict Behav 2014 Mar;39(3):512-519. [doi: 10.1016/j.addbeh.2013.10.036] [Medline: 24333037]
4. Smith AL, Chapman S, Dunlop SM. What do we know about unassisted smoking cessation in Australia? A systematic review, 2005-2012. Tob Control 2015 Jan;24(1):18-27. [doi: 10.1136/tobaccocontrol-2013-051019] [Medline: 24026163]
5. United States Public Health Service Office of the Surgeon General, National Center for Chronic Disease Prevention and Health Promotion (US) Office on Smoking and Health. Smoking Cessation: A Report of the Surgeon General: US Department of Health and Human Services; 2020. [Medline: 32255575]
6. Smith AL, Carter SM, Chapman S, Dunlop SM, Freeman B. Why do smokers try to quit without medication or counselling? A qualitative study with ex-smokers. BMJ Open 2015 Apr 30;5(4):e007301. [doi: 10.1136/bmjopen-2014-007301] [Medline: 25933811]
7. Etter JF, Perneger TV. Attitudes toward nicotine replacement therapy in smokers and ex-smokers in the general public. Clin Pharmacol Ther 2001 Mar;69(3):175-183. [doi: 10.1067/mcp.2001.113722] [Medline: 11240982]
8. Gross B, Brose L, Schumann A, et al. Reasons for not using smoking cessation aids. BMC Public Health 2008 Apr 22;8(1):129. [doi: 10.1186/1471-2458-8-129] [Medline: 18430206]
9. Vogt F, Hall S, Marteau TM. Examining why smokers do not want behavioral support with stopping smoking. Patient Educ Couns 2010 May;79(2):160-166. [doi: 10.1016/j.pec.2009.10.007] [Medline: 19932580]
10. García-Rodríguez O, Secades-Villa R, Flórez-Salamanca L, Okuda M, Liu SM, Blanco C. Probability and predictors of relapse to smoking: results of the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). Drug Alcohol Depend 2013 Oct 1;132(3):479-485. [doi: 10.1016/j.drugalcdep.2013.03.008] [Medline: 23570817]
11. Agboola S, McNeill A, Coleman T, Leonardi Bee J. A systematic review of the effectiveness of smoking relapse prevention interventions for abstinent smokers: relapse prevention interventions for abstinent smokers. Addiction 2010 May 28;105(8):1362-1380. [doi: 10.1111/j.1360-0443.2010.02996.x]
12. Coleman T, Agboola S, Leonardi-Bee J, Taylor M, McEwen A, McNeill A. Relapse prevention in UK Stop Smoking Services: current practice, systematic reviews of effectiveness and cost-effectiveness analysis. Health Technol Assess 2010 Oct;14(49):1-152. [doi: 10.3310/hta14490] [Medline: 21040645]
13. Livingstone-Banks J, Norris E, Hartmann-Boyce J, et al. Relapse prevention interventions for smoking cessation. Cochrane Database of Systematic Reviews 2019 Feb 13. [doi: 10.1002/14651858.CD003999.pub5]
14. Piasecki TM, Fiore MC, McCarthy DE, Baker TB. Have we lost our way? The need for dynamic formulations of smoking relapse proneness. Addiction 2002 Sep;97(9):1093-1108. [doi: 10.1046/j.1360-0443.2002.00216.x] [Medline: 12199822]
15. Shiffman S, Shumaker SA, Abrams DB, Cohen S, et al. Task Force 2: Models of smoking relapse. Health Psychol 1986;5(Suppl):13-27. [doi: 10.1037/0278-6133.5.Suppl.13]
16. Brandon TH, Vidrine JI, Litvin EB. Relapse and relapse prevention. Annu Rev Clin Psychol 2007;3(1):257-284. [doi: 10.1146/annurev.clinpsy.3.022806.091455] [Medline: 17716056]
17. Heckman BW, Cummings KM, Stoltman JJK, et al. Longer duration of smoking abstinence is associated with waning cessation fatigue. Behav Res Ther 2019 Apr;115:12-18. [doi: 10.1016/j.brat.2018.11.011] [Medline: 30509485]
18. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: Englewood Cliffs; 1986.
19. Bandura A. Social cognitive theory of self-regulation. Organ Behav Hum Decis Process 1991 Dec;50(2):248-287. [doi: 10.1016/0749-5978(91)90022-L]
20. Lawn SJ, Pols RG, Barber JG. Smoking and quitting: a qualitative study with community-living psychiatric clients. Soc Sci Med 2002 Jan;54(1):93-104. [doi: 10.1016/s0277-9536(01)00008-9] [Medline: 11820684]
21. Heaney CA, Israel BA. Chapter 9: social networks and social support. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research and Pract, 4th edition: John Wiley & Sons, Inc; 2008:14-21.
22. Biggs A, Brough P, Drummond S. Lazarus and Folkman's psychological stress and coping theory. In: The Handbook of Stress and Health: A Guide to Research and Practice: John Wiley & Sons, Ltd; 2017:349-364. [doi: 10.1002/9781118993811]
23. Riessman F. The "Helper" therapy principle. Social Work 1965 Apr 1;10(2):27-32. [doi: 10.1093/sw/10.2.27]
24. Baha M, Le Faou AL. Smokers' reasons for quitting in an anti-smoking social context. Public Health 2010 Apr;124(4):225-231. [doi: 10.1016/j.puhe.2010.02.011] [Medline: 20371089]
25. Baker TB, Fiore MC. Treating more smokers, more of the time, more successfully. Addiction 2015 Mar;110(3):388-389. [doi: 10.1111/add.12801]
26. Mermelstein R, Cohen S, Lichtenstein E, Baer JS, Kamarck T. Social support and smoking cessation and maintenance. J Consult Clin Psychol 1986 Aug;54(4):447-453. [doi: 10.1037/0022-006X.54.4.447]
27. Nguyen SN, Von Kohorn I, Schulman-Green D, Colson ER. The importance of social networks on smoking: perspectives of women who quit smoking during pregnancy. Matern Child Health J 2012 Aug;16(6):1312-1318. [doi: 10.1007/s10995-011-0896-4] [Medline: 21989676]

XSL•FO
RenderX

28. Powell J, Dawkins L, West R, Powell J, Pickering A. Relapse to smoking during unaided cessation: clinical, cognitive and motivational predictors. Psychopharmacology (Berl) 2010 Dec;212(4):537-549. [doi: 10.1007/s00213-010-1975-8] [Medline: 20703450]
29. Cook MR, Gerkovich MM, O'Connell KA, Potocky M. Reversal theory constructs and cigarette availability predict lapse early in smoking cessation. Res Nurs Health 1995 Jun;18(3):217-224. [doi: 10.1002/nur.4770180305] [Medline: 7754092]
30. O'Connell KA, Shiffman S, Decarlo LT. Does extinction of responses to cigarette cues occur during smoking cessation? Addiction 2011 Feb;106(2):410-417. [doi: 10.1111/j.1360-0443.2010.03172.x] [Medline: 21054617]
31. Flatt JD, Agimi Y, Albert SM. Homophily and health behavior in social networks of older adults. Fam Community Health 2012;35(4):312-321. [doi: 10.1097/FCH.0b013e3182666650] [Medline: 22929377]
32. Christakis NA, Fowler JH. The collective dynamics of smoking in a large social network. N Engl J Med 2008 May 22;358(21):2249-2258. [doi: 10.1056/NEJMsa0706154] [Medline: 18499567]
33. Campbell J, Mays M, Yuan N, Muramoto MLML. Who are health influencers? Characterizing a sample of tobacco cessation interveners. Am J Health Behav 2007;31(2):181-192. [doi: 10.5555/ajhb.2007.31.2.181] [Medline: 17269908]
34. Muramoto ML, Hall JR, Nichter M, et al. Activating lay health influencers to promote tobacco cessation. Am J Hlth Behav 2014 May 1;38(3):392-403. [doi: 10.5993/AJHB.38.3.8]
35. Muramoto ML, Wassum K, Connolly T, Matthews E, Floden L. Helpers program: a pilot test of brief tobacco intervention training in three corporations. Am J Prev Med 2010 Mar;38(3 Suppl):S319-S326. [doi: 10.1016/j.amepre.2009.12.009] [Medline: 20176303]
36. Castañeda H, Nichter M, Nichter M, Muramoto M. Enabling and sustaining the activities of lay health influencers: lessons from a community-based tobacco cessation intervention study. Health Promot Pract 2010 Jul;11(4):483-492. [doi: 10.1177/1524839908318288] [Medline: 18540006]
37. Muramoto ML, Hopkins A, Bell M, Allen A, Nair U, Connolly TE. Results of a feasibility study of helpers stay quit training for smoking relapse prevention. Nicotine Tob Res 2021 Mar 19;23(4):711-715. [doi: 10.1093/ntr/ntaa176] [Medline: 32966558]
38. Hopkins AL, Muramoto ML, McCarty C, et al. Personal network changes among smoking relapse prevention intervention participants. Am J Health Behav 2023 Nov 11;47(5):958-964. [doi: 10.5993/AJHB.47.5.8]
39. Fiore MC, Jaen CR, Baker TB, et al. Treating Tobacco Use and Dependence: 2008 Update: U.S. Department of Health and Human Services; 2008.
40. Simmons VN, Heckman BW, Ditre JW, Brandon TH. A measure of smoking abstinence-related motivational engagement: Development and initial validation. Nicotine & Tobacco Research 2010 Apr 1;12(4):432-437. [doi: 10.1093/ntr/ntq020]
41. Meijer E, Van den Putte B, Gebhardt WA, et al. A longitudinal study into the reciprocal effects of identities and smoking behaviour: Findings from the ITC Netherlands Survey. Soc Sci Med 2018 Mar;200:249-257. [doi: 10.1016/j.socscimed.2017.12.006] [Medline: 29321102]
42. Vangeli E, Stapleton J, West R. Residual attraction to smoking and smoker identity following smoking cessation. Nicotine & Tobacco Research 2010 Aug 1;12(8):865-869. [doi: 10.1093/ntr/ntq104]
43. Velicer WF, Diclemente CC, Rossi JS, Prochaska JO. Relapse situations and self-efficacy: an integrative model. Addict Behav 1990;15(3):271-283. [doi: 10.1016/0306-4603(90)90070-e] [Medline: 2378287]
44. Roesch SC, Aldridge AA, Huff TLP, Langner K, Villodas F, Bradshaw K. On the dimensionality of the Proactive Coping Inventory: 7, 5, 3 factors? Anxiety Stress Coping 2009 May;22(3):327-339. [doi: 10.1080/10615800802082304] [Medline: 19253071]
45. Greenglass E. Chapter 3. proactive coping. In: Frydenberg E, editor. Coping Meet Goals Vis Chall: Oxford University Press; 2002:37-62. [doi: 10.1093/med:psych/9780198508144.003.0003]
46. Kroenke K, Spitzer RL, Williams JBW, Löwe B, Lowe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
47. Mills SD, Fox RS, Pan TM, Malcarne VL, Roesch SC, Sadler GR. Psychometric evaluation of the Patient Health Questionnaire–4 in Hispanic Americans. Hisp J Behav Sci 2015 Nov;37(4):560-571. [doi: 10.1177/0739986315608126] [Medline: 29097834]
48. Löwe B, Wahl I, Rose M, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. J Affect Disord 2010 Apr;122(1-2):86-95. [doi: 10.1016/j.jad.2009.06.019] [Medline: 19616305]
49. Allen AM, Lundeen K, Murphy SE, Spector L, Harlow BL. Web-delivered multimedia training materials for the self-collection of dried blood spots: a formative project. JMIR Form Res 2018 Nov 5;2(2):e11025. [doi: 10.2196/11025] [Medline: 30684406]
50. Murphy SE, Wickham KM, Lindgren BR, Spector LG, Joseph A. Cotinine and trans 3'-hydroxycotinine in dried blood spots as biomarkers of tobacco exposure and nicotine metabolism. J Expo Sci Environ Epidemiol 2013;23(5):513-518. [doi: 10.1038/jes.2013.7] [Medline: 23443235]
51. McCarty C. Measuring structure in personal networks. J Soc Struct 2002 [FREE Full text]
52. McCarty C, Killworth PD, Rennell J. Impact of methods for reducing respondent burden on personal network structural measures. Soc Networks 2007 May;29(2):300-315. [doi: 10.1016/j.socnet.2006.12.005]

XSL•FO
RenderX

53. McCarty C, Molina JL, Aguilar C, Rota L. A comparison of social network mapping and personal network visualization. Field methods 2007 May;19(2):145-162. [doi: 10.1177/1525822X06298592]
54. EgoNet. SourceForge. 2013. URL: http://sourceforge.net/projects/egonet/ [accessed 2025-12-12]
55. Min MO, Tracy EM, Kim H, et al. Changes in personal networks of women in residential and outpatient substance abuse treatment. J Subst Abuse Treat 2013 Oct;45(4):325-334. [doi: 10.1016/j.jsat.2013.04.006] [Medline: 23755971]
56. Tracy EM, Kim H, Brown S, Min MO, Jun MK, McCarty C. Substance abuse treatment stage and personal networks of women in substance abuse treatment. J Soc Social Work Res 2012 Mar 23;3(2):65-79. [doi: 10.5243/jsswr.2012.5] [Medline: 22639705]
57. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. Soc Probl 1997 May;44(2):174-199. [doi: 10.1525/sp.1997.44.2.03x0221m]
58. Heckathorn DD. 6. Extensions of respondent-driven sampling: analyzing continuous variables and controlling for differential recruitment. Sociol Methodol 2007 Aug;37(1):151-208. [doi: 10.1111/j.1467-9531.2007.00188.x]
59. Salganik MJ, Heckathorn DD. 5. Sampling and estimation in hidden populations using respondent-driven sampling. Sociol Methodol 2004 Dec;34(1):193-240. [doi: 10.1111/j.0081-1750.2004.00152.x]
60. Volz E, Heckathorn DD. Probability based estimation theory for respondent driven sampling. J Off Stat 2008;24(1):79-97 [FREE Full text]
61. Wejnert C. An empirical test of respondent-driven sampling: point estimates, variance, degree measures, and out-of-equilibrium data. Sociol Methodol 2009 Aug 1;39(1):73-116. [doi: 10.1111/j.1467-9531.2009.01216.x] [Medline: 20161130]
62. Wejnert C, Heckathorn DD. Web-based network sampling. Sociol Methods Res 2008 Aug;37(1):105-134. [doi: 10.1177/0049124108318333]
63. Miller G, Dingwall R. Network analysis and qualitative research: a method of contextualization. In: Context and Method in Qualitative Research: SAGE Publications; 2016:120-138. [doi: 10.4135/9781849208758.n9]
64. Edwards G. Mixed-method approaches to social network analysis. NCRM 2010 Jan [FREE Full text]
65. Nohlert E, Öhrvik J, Helgason Á. Non-responders in a quitline evaluation are more likely to be smokers – a drop-out and long-term follow-up study of the Swedish National Tobacco Quitline. Tob Induced Dis 2016 Dec;14(1):5. [doi: 10.1186/s12971-016-0070-2]
66. Degenne A, Lebeaux MO. The dynamics of personal networks at the time of entry into adult life. Soc Networks 2005 Oct;27(4):337-358. [doi: 10.1016/j.socnet.2004.11.002]
67. Morgan DL, Neal MB, Carder P. The stability of core and peripheral networks over time. Soc Networks 1997 Jan;19(1):9-25. [doi: 10.1016/S0378-8733(96)00288-2]
68. Suitor J, Keeton S. Once a friend, always a friend? Effects of homophily on women's support networks across a decade. Soc Networks 1997 Jan;19(1):51-62. [doi: 10.1016/S0378-8733(96)00290-0]
69. Terhell EL, Broese van Groenou MI, van Tilburg T. Network contact changes in early and later postseparation years. Soc Networks 2007 Jan;29(1):11-24. [doi: 10.1016/j.socnet.2005.11.006]
70. van Tilburg T. Support networks before and after retirement. J Soc Pers Relat 1992 Aug;9(3):433-445. [doi: 10.1177/0265407592093006]
71. Wellman B, Wong RL, Tindall D, Nazer N. A decade of network change: Turnover, persistence and stability in personal communities. Soc Networks 1997 Jan;19(1):27-50. [doi: 10.1016/S0378-8733(96)00289-4]
72. Lubbers MJ, Molina JL, Lerner J, Brandes U, Ávila J, McCarty C. Longitudinal analysis of personal networks. The case of Argentinean migrants in Spain. Soc Networks 2010 Jan;32(1):91-104. [doi: 10.1016/j.socnet.2009.05.001]
73. Trotter RT. Qualitative research sample design and sample size: resolving and unresolved issues and inferential imperatives. Prev Med 2012 Nov;55(5):398-400. [doi: 10.1016/j.ypmed.2012.07.003] [Medline: 22800684]
74. Dempster AP, Laird NM, Rubin DB. Maximum likelihood estimation from incomplete data via the EM algorithm. J R Stat Soc Ser B 1977;39(1):1-22. [doi: 10.1111/j.2517-6161.1977.tb01600.x]
75. Little RJA, Rubin DB. Statistical Analysis with Missing Data, 2nd edition: Wiley; 2010.
76. Diggle P, Kenward MG. Informative drop-out in longitudinal data analysis. Appl Stat 1994;43(1):49. [doi: 10.2307/2986113]
77. LeCompte MD, Schensul JJ. Ethnographer's Toolkit: Rowman & Littlefield; 1999.
78. Bernard HR, Ryan GW. Analyzing Qualitative Data: Systemic Approaches: Sage Publications; 2010.
79. Miles MB, Huberman AM. Qualitative Data Analysis: An Expanded Sourcebook: Sage Publications; 1994.
80. Yung YF, Lamm M, Zhang W. Causal mediation analysis with the CAUSALMED procedure. : SAS Institute; 2018 URL: https://support.sas.com/resources/papers/proceedings18/1991-2018.pdf [accessed 2025-12-31]
81. Livingstone-Banks J, Norris E, Hartmann-Boyce J, West R, Jarvis M, Hajek P. Relapse prevention interventions for smoking cessation. Cochrane Database Syst Rev 2019 Feb 13;2(2):CD003999. [doi: 10.1002/14651858.CD003999.pub5] [Medline: 30758045]
82. Christakis NA, Fowler JH. Social contagion theory: examining dynamic social networks and human behavior. Stat Med 2013 Feb 20;32(4):556-577. [doi: 10.1002/sim.5408]
83. Mermelstein R, Cohen S, Lichtenstein E, Baer JS, Kamarck T. Social support and smoking cessation and maintenance. J Consult Clin Psychol 1986;54(4):447-453. [doi: 10.1037/0022-006X.54.4.447]

XSL•FO
RenderX

84. Meijer E, Gebhardt WA, van Laar C, van den Putte B, Evers AWM. Strengthening quitter self-identity: An experimental study. Psychology & Health 2018 Oct 3;33(10):1229-1250. [doi: 10.1080/08870446.2018.1478976]
85. Greenglass ER. Chapter 3: proactive coping. In: Frydenberg E, editor. Beyond Coping: Meeting Goals, Visions, and Challenges: Oxford University Press; 2002:37-62. [doi: 10.1093/med:psych/9780198508144.003.0003]
86. Raposa EB, Laws HB, Ansell EB. Prosocial Behavior Mitigates the Negative Effects of Stress in Everyday Life. Clin Psychol Sci 2016 Jul;4(4):691-698. [doi: 10.1177/2167702615611073] [Medline: 27500075]
87. Riessman F. The “Helper” therapy principle. Soc Work 1965 Apr;10(2):27-32. [doi: 10.1093/sw/10.2.27]
88. Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. Am J Public Health 2010 Jul;100(7):1282-1289. [doi: 10.2105/AJPH.2009.165449] [Medline: 20466971]
89. Graham AL, Zhao K, Papandonatos GD, et al. A prospective examination of online social network dynamics and smoking cessation. In: Amblard F, editor. PLOS ONE 2017;12(8):e0183655. [doi: 10.1371/journal.pone.0183655] [Medline: 28832621]
90. Bray BC, Smith RA, Piper ME, Roberts LJ, Baker TB. Transitions in smokers’ social networks after quit attempts: a latent transition analysis. Nicotine Tob Res 2016 Dec;18(12):2243-2251. [doi: 10.1093/ntr/ntw173] [Medline: 27613925]
91. Burgess-Hull AJ, Roberts LJ, Piper ME, Baker TB. The social networks of smokers attempting to quit: An empirically derived and validated classification. Psychol Addict Behav 2018 Feb;32(1):64-75. [doi: 10.1037/adb0000336] [Medline: 29251951]
92. Vallone DM, Duke JC, Cullen J, McCausland KL, Allen JA. Evaluation of EX: a national mass media smoking cessation campaign. Am J Public Health 2011 Feb;101(2):302-309. [doi: 10.2105/AJPH.2009.190454] [Medline: 21164094]
93. Zhao K, Wang X, Cha S, et al. A multirelational social network analysis of an online health community for smoking cessation. J Med Internet Res 2016 Aug 25;18(8):e233. [doi: 10.2196/jmir.5985] [Medline: 27562640]
94. Cole-Lewis H, Perotte A, Galica K, et al. Social network behavior and engagement within a smoking cessation Facebook page. J Med Internet Res 2016 Aug 2;18(8):e205. [doi: 10.2196/jmir.5574] [Medline: 27485315]
95. Lakon CM, Hipp JR, Timberlake DS. The social context of adolescent smoking: a systems perspective. Am J Public Health 2010 Jul;100(7):1218-1228. [doi: 10.2105/AJPH.2009.167973] [Medline: 20466966]
96. Stone A, Jason LA, Light JM, Stevens EB. The role of ego networks in studies of substance use disorder recovery. Alcohol Treat Q 2016;34(3):315-328. [doi: 10.1080/07347324.2016.1182818] [Medline: 27594761]
97. Aschbrenner KA, Bobak C, Schneider EJ, Naslund JA, Brunette MF, O’Malley AJ. Egocentric social networks and smoking among adults with serious mental illness. Transl Behav Med 2018 Jul 17;8(4):531-539. [doi: 10.1093/tbm/ibx014] [Medline: 30016519]

Abbreviations

IRB: institutional review board
MAR: missing at random
MCAR: missing completely at random
MD: mean difference
MNAR: missing not at random
PNA: personal network assessment
pRCT: pragmatic randomized controlled trial

Edited by J Sarvestan; submitted 11.Aug.2025; peer-reviewed by C Cottrell-Daniels; revised version received 07.Nov.2025; accepted 19.Nov.2025; published 27.Jan.2026.

Please cite as:
Muramoto M, Hopkins A, McCarty C, Allen A, Dickinson LM, Connolly T, Spradley J, Ying J
Effect of Helpers Stay Quit Online Training on Preventing Smoking Relapse and Personal Networks: Protocol for a Pragmatic Randomized Controlled Trial and Embedded Mixed Methods Personal Network Study
JMIR Res Protoc 2026;15:e82140
URL: https://www.researchprotocols.org/2026/1/e82140
doi:10.2196/82140

© Myra Muramoto, Allison Hopkins, Christopher McCarty, Alicia Allen, L Miriam Dickinson, Timothy Connolly, Janet Spradley, Jun Ying. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR

XSL•FO
RenderX

Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy of a Virtual Reality Game on Children's Fear and Anxiety During Dental Procedures (VR-TOOTH): Protocol for a Randomized Controlled Trial

Julien Gardner[1,2]; Vallerie Markopoulos[2,3,4], DDS; Wenjia Wu[3], DDS, MSc; Gabrielle Gilbert[5]; Daphnée Pelletier[2,5]; Charlotte Fafard[1,2], BSc; Anne Gagné[3]; Estelle Guingo[2,6], PhD; Christine Genest[7,8], RN, PhD; Marie-Ève Asselin[3,4], DDS, MSc; Pascale Ouimet[2,8], BScN; Kate St-Arneault[2,8], RN, PhD(c); Sylvie Le May[2,3,8], RN, PhD, FCAN

[1]Faculty of Medicine, Université de Montréal, Montreal, QC, Canada

[2]Research Center, Centre Hospitalier Universitaire Sainte-Justine, Montreal, QC, Canada

[3]Faculty of Dental Medicine, Université de Montréal, Montreal, QC, Canada

[4]Department of Dental Medicine, Centre Hospitalier Universitaire Sainte-Justine, Montréal, QC, Canada

[5]Faculty of Arts and Sciences, Univeristé de Montréal, Montreal, QC, Canada

[6]Université du Québec en Abitibi-Témiscamingue, Abitibi-Témiscamingue, QC, Canada

[7]Trauma Studies Centre, Institut Universitaire en Santé Mentale de Montréal, Montréal, QC, Canada

[8]Faculty of Nursing, Université de Montréal, Pavillon Marguerite-d'Youville, 2375 chemin de la Côte-Sainte-Catherine, Montreal, QC, Canada

Corresponding Author:
Sylvie Le May, RN, PhD, FCAN
Research Center, Centre Hospitalier Universitaire Sainte-Justine, Montreal, QC, Canada

Abstract

Background: Dental fear and anxiety (DFA) affects approximately a quarter of children and adolescents. It significantly contributes to pediatric patients avoiding dental care later in adulthood. Lack of cooperation due to DFA can create a stressful environment, often forcing dentists to end appointments prematurely and consider alternative pharmacological treatments. The use of virtual reality (VR) during dental procedures, offering an immersive sensory experience, may serve as an additional nonpharmacologic tool to better manage DFA in children with special health care needs (SHCN) undergoing dental treatment.

Objective: This study aims to assess the effectiveness of VR immersion in reducing anxiety and pain among pediatric patients with SHCN undergoing dental procedures. The study also seeks to understand the satisfaction of parents and health care providers with the use of VR during dental appointments.

Methods: This randomized controlled trial follows a parallel design with two groups: a control group receiving standard care and an experimental group using VR. A sample size of 400 participants was calculated. Participants will be randomly assigned equally to each group. Recruitment will take place at the dental clinic of the Centre Hospitalier Universitaire Sainte-Justine, a tertiary- and quaternary-care center that primarily serves pediatric patients with SHCN. The two primary outcomes will include both observed and objective biomarker-based measures of anxiety. DFA will be evaluated using the Venham Anxiety Rating Scale as well as changes in mean levels of salivary alpha-amylase. Sociodemographic characteristics, parents' and health care professionals' satisfaction levels, participants' pain intensity and behavior during the procedure, changes in heart rate, occurrence of side effects, procedure duration, and any deviations from normal procedural length will also be collected. Descriptive and comparative statistics will be conducted for demographic and clinical comparisons and will be used to present sociodemographic and clinical data, parents' and health care professionals' satisfaction levels, child satisfaction with the game, and procedural time.

Results: This study will be conducted from November 2023 to December 2025. As of November 2025, 300 participants have been recruited. Results are expected to be available in June 2026.

Conclusions: We believe that the results of this study will confirm the efficacy of VR in reducing DFA in children with SHCN, providing an additional nonpharmacological alternative for better managing this condition in pediatric hospital settings.

Trial Registration: Clinicaltrials.gov NCT06182462; https://clinicaltrials.gov/study/NCT06182462

International Registered Report Identifier (IRRID): DERR1-10.2196/83672

(JMIR Res Protoc 2026;15:e83672) doi:10.2196/83672

KEYWORDS

virtual reality; dentistry; paediatrics; pain; special health care needs; anxiety

XSL•FO
RenderX

Introduction

Background

Dental fear and anxiety (DFA) is defined as a patient's psychological state of worry, anxiety, or even dread that is specific to a dental procedure [1]. Although different, dental fear and dental anxiety, which respectively represent an emotional response and a heightened state of apprehension to potential distress from dental treatment, are challenging to distinguish in clinical practice [2,3]. Consequently, the term "dental fear and anxiety" is frequently used to describe these intense negative emotional states associated with dental procedures. Children's DFA disorder refers to the clinical manifestations of dental phobia or anxiety in children, including reduced cooperation, lessened compliance, and even resistance to treatment [4,5]. DFA are conditions that affect approximately 13.3% to 29.3% of children and adolescents and are a significant reason why patients avoid dental care [3,6]. In children, DFA is also associated with lower oral health-related quality of life [7]. Although the etiology of DFA is multifactorial, often stemming from both exogenous and endogenous sources, a past traumatic dental experience is the most predictive factor for DFA [8,9]. A study by Ten Berge et al [10] states that the majority of DFAs experienced by adults originates from poor dental experiences during childhood. These findings highlight the importance of making each dental visit a positive experience for pediatric patients. Short-term distress during appointments, if not managed appropriately, can accumulate into poor dental experiences and consequently, reinforce DFA into adulthood [11]. The long-term effects of these negative childhood dental experiences can lead patients to avoid seeking proper dental care in the future [12].

Pediatric dental patients with special health care needs (SHCN), such as patients with cancer, neurodivergent patients, or those with chronic conditions, including cardiovascular or pulmonary conditions or neurological afflictions, require extra time and special considerations during treatment due to medical, physical, cognitive, or developmental conditions [13]. This population includes children with behavior disorders (eg, autism spectrum, anxiety, and attention-deficit/hyperactivity disorder), congenital conditions (eg, trisomy 21 or congenital heart disease), developmental disorders (eg, cerebral palsy), systemic illnesses (eg, childhood cancer or sickle cell disease), or cognitive disabilities (eg, intellectual disability) [13]. Children with SHCN face more barriers to dental care than the overall population [13], including external factors like transportation, cost, inadequate dental facilities, as well as internal factors such as fear and poor tolerance [13,14]. Many children with SHCN are cared for by dentists in community settings or, more often, in hospital dental clinics [13]. They tend to experience a higher level of DFA, which can make dental visits more challenging [13]. Providing comprehensive care and ensuring each dental visit is positive for pediatric patients with SHCN is crucial in promoting good oral health habits and improving their oral health as they transition into adulthood [13]. Understanding and assessing DFA in children is key to delivering successful dental care with high satisfaction in this age group. Among the vast array of assessment methods available, self-report, parental proxy, observation-based, and physiological assessments are the four major types used to evaluate DFA in children [15]. Recently, salivary alpha-amylase, an enzyme linked to both adrenaline and noradrenaline, has also been used as a marker for autonomic nervous system activity and stress [16].

Factors in the dental setting that trigger DFA include loud dental instrument sounds, the presence of unfamiliar people examining the oral cavity, and fear of pain [17]. Of these, the biggest trigger of DFA is the anticipation and use of local anesthesia injections [9,17]. Although necessary for effective pain control during certain procedures, local anesthesia injections are understandably uncomfortable [9]. For children, the initial injection combined with the feeling of numbness during the procedure can be especially distressing [9]. Pharmacological agents, combined with light to moderate sedation or general anesthesia, can also be considered for noncooperative patients, who are often time-consuming and more costly and carry higher health risks [3]. In current research, audiovisual distractions such as tablets and TV screens have been used as additional distraction techniques beyond traditional tell-show-do methods, yielding overall positive results [18]. However, these techniques lack interactivity, resulting in less immersive distraction environments for children [19]. The lack of cooperation due to DFA often forces dentists treating pediatric patients to end appointments prematurely, sometimes without completing the planned procedures. Treating anxious and fearful patients can create a stressful environment for the clinician and the dental team [3]. Particularly when treating children with SHCN, extra time and tools are needed to ensure comfortable dental care. Depending on the child's diagnosis, some may experience hypersensitivity to external stimuli such as loud noises, aversion to specific tastes, or difficulty straying from usual daily routines [20]. Furthermore, Pagano et al [20] showed that augmented reality was well suited for patients with autism spectrum disorder as preparation for their dental visits.

Additionally, a systematic review by Cunningham et al [21] concluded that virtual reality (VR) is a promising tool in dentistry, especially for children with autism spectrum disorder or other special health care needs. VR is defined as an artificial environment experienced through sensory stimuli [22]. It is a modern tool that immerses patients in a "game" or "virtual world." Commonly used in medicine to distract patients during unpleasant procedures such as vaccination, cast removal, and short bedside interventions, VR has been proven effective at decreasing procedural anxiety and creating a more positive patient experience [23]. Limited existing research on VR in dentistry shows positive results for anxiety management during dental procedures. A recent clinical trial by Alshatrat et al [19] found that VR was effective in reducing anxiety in young children during dental procedures.

Additionally, a previous study by Ram et al [24] reported high satisfaction among both parents and clinicians using audiovisual glasses to guide children during dental treatments. However, research on clinical VR applications in pediatric dentistry remains limited, especially within populations with SHCN. VR offers potential as an additional nonpharmacologic tool for pediatric dentists. A clinical study assessing VR's effectiveness during dental appointments in pediatric patients with SHCN

XSL•FO
RenderX

could lead to better control of DFA and possibly facilitate easier procedures.

The Study

Aims

The study aims to evaluate the efficacy of VR immersion as a tool in reducing anxiety in pediatric patients with SHCN undergoing dental procedures. The study also aims to gain insight into parents' and health care providers' satisfaction levels with the use of VR by children during dental treatments.

Hypothesis

We hypothesize that VR immersion will more effectively reduce anxiety in pediatric patients undergoing dental procedures than the standard intervention of watching a cartoon on a muted mounted TV set or tablet.

Objectives

Primary Outcomes

The primary outcomes of this study are to measure both observed and objective biomarker-based indicators associated with DFA. The observed indicator will be evaluated using the Venham Anxiety Rating Scale (VARS) [25] (Table 1) and will be assessed by examining changes in mean levels of salivary alpha-amylase (SAA) between baseline (T0) and the middle of the intervention (T1) [26].

Table . The Venham Anxiety Rating Scale.

Rating	Definition
0	Relaxed, smiling, willing, and able to converse.
1	Uneasy and concerned. During a stressful procedure, the child may protest briefly and quietly to indicate discomfort. Hands remain down or partially raised to signal discomfort. Child willing and able to interpret experience as requested. Tense facial expression, may have tears in eyes.
2	Child appears scared. Tone of voice and questions and answers reflect anxiety. During a stressful procedure, verbal protest, crying or quiet crying, hands tense and raised (not interfering much—may touch the dentist's hand or instrument, but not pull at it). The child interprets the situation with reasonable accuracy and continues to work to cope with his or her anxiety.
3	Shows reluctance to enter the situation, difficulty in correctly assessing situational threat. Pronounced verbal protest and crying. Uses hands to try to stop procedure. Protest out of proportion to threat. Copes with the situation with great reluctance.
4	Anxiety interferes with the ability to assess the situation. General crying not related to treatment. More prominent body movements. The child can be reached through verbal communication, and eventually, with reluctance and great effort, he or she begins the work of coping with the threat.
5	The child is out of contact with the reality of the threat. General loud crying, unable to listen to verbal communication, and makes no effort to cope with the threat. Actively involved in escape behavior. Physical restraint required.

Secondary Outcomes

The secondary outcomes of this study are to compare the following indicators between the VR distraction and the standard intervention groups:

1. Parents' (Multimedia Appendix 1) and health care professionals' (Multimedia Appendix 2) satisfaction levels
2. Child satisfaction with the use of VR (Multimedia Appendix 1).
3. Physiological parameters such as heart rate and oxygen saturation
4. Behavior during the procedure (Table 2)
5. Occurrence of side effects
6. Length of the dental procedure
7. Child's pain intensity (Multimedia Appendix 3)

XSL•FO
RenderX

Table . The Venham Behavioral Rating Scale.

Rating	Definition
0	Total cooperation, best possible working conditions, and no crying or physical protest.
1	Mild or soft verbal protest, crying or quiet crying as a signal of discomfort, but does not obstruct progress. Appropriate behavior for the procedure, that is, slight start at injection, “ow” during drilling if hurting, and so on.
2	Protest more prominent. Both crying and hand signals. May move the head around making it hard to administer treatment. Protest is more distracting and troublesome. However, the child still complies with the request to cooperate.
3	Protest presents a real problem to the dentist. Complies with demands reluctantly, requiring extra effort by dentist. Body movement.
4	Protest disrupts procedure and requires that all of the dentist’s attention be directed toward the child’s behavior. Compliance was eventually achieved after considerable effort by the dentist, but without much actual physical restraint (may require holding the child’s hands or the like to start). More prominent body movement.
5	General protest, no compliance or cooperation. Physical restraint is required.

Methods

Design

This randomized controlled trial (RCT) will follow a parallel design with two groups: a control group (watching cartoons on a muted wall-mounted TV or a tablet) and an experimental group (VR intervention during the dental procedure).

Sample and Setting

In mid-2023, we conducted a pilot study to evaluate feasibility and identify recruitment challenges before moving forward with the full RCT [27]. Between June 6, 2023, and July 20, 2023, 25 patients were recruited over 44 days at a single research site. Among these 25 patients, the mean VARS and Venham Behavioral Rating Scale scores before the procedure were 0.08 (SD = 0.28) for both scales, and during the procedure, the mean scores were 0.625 (SD = 1.35) and 0.667 (SD = 1.37), respectively. The variability in scores was higher at the time point with the higher average scores; therefore, we assumed a test with unequal variances. For the SAA outcome, data from a separate RCT conducted by our group were used [28]. This study measured changes in SAA before and during magnetic resonance imaging in children assigned to either a VR or a control condition. In that RCT, the mean difference in log-transformed SAA was 0.171 (SD = 1.05) in the control group and −0.096 (SD = 0.495) in the VR group. Because the expected difference in log SAA was smaller, this outcome guided our sample size calculations.

Therefore, to achieve 80% power to detect a mean difference of −0.096 (SD = 0.495) in log SAA using a 2-sided, 2-sample *t* test with unequal variances and a type-I error (α) of .05, a sample size of 364 participants (n=182 per group) was needed. Given the two primary outcomes, a Bonferroni correction was applied, reducing the α to .025. Based on an estimated 10% attrition rate from our pilot study, the final sample size was increased to 400 participants. Patients are currently being recruited for this study.

Recruitment is taking place at the dental clinic of the Centre Hospitalier Universitaire Sainte-Justine, a pediatric hospital in Montréal, Canada. This clinic mainly serves patients with SHCN, such as craniofacial abnormalities, autism spectrum disorder, or children battling cancer. Pediatric patients with SHCN make up about 80% of the clinic’s total clientele, while the remaining patients are otherwise healthy with dental trauma or other emergencies. The clinic sees roughly 5000 visits annually, including around 4000 patients with SHCN. Typically, about 600 (15%) of these patients have a severe diagnosis and are known to be completely uncooperative. Therefore, we estimate that approximately 3400 patients would be eligible for recruitment. Based on our pilot recruitment rate, we anticipate recruiting roughly 3 participants per week during September through April and 9 participants per week during May through August. Date collection is currently ongoing, and we expect to reach our target enrollment of 400 patients by the end of December 2025.

Participants will be identified by the clinic’s resident dentist using the scheduling system for upcoming appointments involving specific painful dental procedures, such as examinations, prophylaxis, extractions, and restorations. The research assistants will then contact the parents or legal guardians to provide information about the study and obtain their approval in advance. An independent individual, separate from the clinical dental team, will review the consent with both participants and parents. The information and consent form will be signed by a parent or legal guardian on the day of the visit. Whenever possible, patient assent will be sought.

Inclusion Criteria

Children participating in this study must meet the following criteria: (1) be aged 6 to 17 years; (2) have received the dentist’s recommendation to participate; (3) need to undergo any dental

XSL•FO
RenderX

procedure; and (4) be accompanied by a parent or legal guardian who can read, write, and understand either French or English.

Exclusion Criteria

Participants will be excluded from this study if they (1) are diagnosed with epilepsy; (2) have paralysis or paresis of the hand; (3) have any diagnosed eye disease or problem; or (4) have any other conditions preventing them from using VR (eg, epidermolysis bullosa). Patients with a strong history of motion sickness will not be excluded from this study, but this information will be documented, and additional monitoring will be provided. The use of medication (opioid and nonopioid analgesics, antiemetics, anxiolytics, or any other drugs) prior to the procedure within the last 4 hours will not exclude participants from this study, but this information will be collected before the procedure and included as part of the sociodemographic and clinical information form. Patients who have received previous dental procedures at the recruiting setting will not be excluded.

Randomization and Allocation

Randomization will be performed using the electronic REDCap (Research Electronic Data Capture; Vanderbilt University) system. Allocation to either intervention will be randomized by an independent biostatistician from the Unité de recherche clinique appliquée (Applied Clinical Research Unit). To equalize participants across both arms, permuted block randomization with randomly selected block sizes design will be used to randomize participants to their intervention [29]. Access to the randomization list will be granted only to the biostatistician, and allocation will be concealed in REDCap to reduce selection bias.

Interventions

Control Intervention

Participants assigned to the control intervention will only receive the care-as-usual approach. This includes viewing cartoons on a muted, mounted television already set by the clinic, watching a video of their choice on a tablet on YouTube, and the use of pharmaceutical treatment during the procedure, such as injected local anesthesia, if indicated by the procedure. In the event of noncooperation during the appointment, any necessary procedural retakes or rescheduling of appointments as determined by the dentist will be documented by the research assistant. One parent will be permitted to be present in the room during the procedure as part of the clinic's usual protocol, and their presence will be recorded. Children in the control intervention will be offered the option to try the VR game after the study period, if they choose to do so.

Experimental Intervention

Participants assigned to the experimental intervention will engage with a VR game called "Dream Dental," tailored for a pediatric population. The simplified no-success game, developed by Paperplane Therapeutics [30], is enjoyable regardless of a child's video game experience, and its point-and-shoot arcade style makes it easy to understand (Figure 1). It was designed specifically for this study, with the horizontal position required for dental procedures in mind (Figure 1). The VR video game aims to reduce anxiety in children aged 6 to 17 years by means of immersive distraction.

Figure 1. Patient's position during a dental procedure and virtual reality accessory [30]. The images or screen captures were reproduced with permission from Paperplane Therapeutics.





We will use a lightweight head-mounted device (HMD; Pico Neo 4) with built-in eye tracking [31]. The Pico Neo 4 is a standalone VR headset featuring a 4K Super-Vision Display with pancake lenses for a clearer, broader field of view. Its key features include a balanced design with a rear-mounted battery, a high-resolution display (2160×2160 per eye), a 90 Hz refresh rate, and full-color pass-through via its 16 MP RGB camera. It also uses 6 degrees of freedom inside-out tracking and offers both wireless and wired PC VR connectivity. This headset was specifically designed to be lightweight and smaller to better fit children's smaller heads. Children can play throughout the entire dental procedure using a clicker in one hand (Figure 1); eye-tracking sensors will assist navigation through the game, but will not be used for data collection. The VR game "Dream Dental" features a rail mechanism (predefined path or sequence of scenes to reduce body and head movement) that guides children through an interactive environment filled with entertaining characters. The game is designed for dental procedures (Figure 2) and incorporates innovative eye-tracking technology integrated into the headset. This technology simplifies navigation by enabling users to target objects with their eyes instead of relying on head movements. The goal is to make the game easier to use during dental procedures, where head movements need to be restricted. These features are also intended to reduce virtual motion sickness. The VR HMD allows children to view the game they are playing in real time while

XSL•FO
RenderX

blocking their view of the procedure, and it was determined not to hinder dental procedures [27].

Figure 2. Dream Dental game environment [30].



Pharmaceutical treatment during the procedure, such as injected local anesthesia, will be used as indicated. In the event of noncooperation during the appointment, any procedural retake or rescheduled appointments will be noted by the research assistant or resident dentist. One parent will be allowed in the room during the dental procedure, and their presence will be recorded.

Study Time Periods

The study time periods are as follows and are detailed in Table 3. T0 corresponds to the time before the start of the intervention, T1 represents the time during or at the midpoint during the procedure, and T2 corresponds to the period 10 minutes after the intervention.

Table . Study time periods and outcomes.

Study time periods	Outcomes									
	Patients' SDC[a]	VARS[b] anxiety	SAA[c]	Parents' and health care professionals' satisfaction	Patients' satisfaction	Patients' physiological parameters	VBRS[d] behavior	Side effects	Length of procedure	FLACC-R[e]
T0[f]	✓	✓	✓			✓	✓	✓		✓
T1[g]		✓				✓	✓			✓
T2[h]			✓	✓	✓	✓		✓	✓	

[a]SDC: sociodemographic characteristics.

[b]VARS: Venham Anxiety Rating Scale.

[c]SAA: salivary alpha-amylase.

[d]VBRS: Venham Behavioral Rating Scale.

[e]FLACC-R: Face, Legs, Activity, Cry, and Consolability—Revised.

[f]T0: before the procedure.

[g]T1: midpoint during the procedure.

[h]T2: 10 minutes after the procedure.

Measures and Outcomes

Sociodemographic and Clinical Characteristics

Sociodemographic characteristics will be collected in the waiting room, after consent and prior to the intervention, by the parent or legal guardian present, including age, sex, ethnicity, and the procedure needed. Other clinical information will also include the participant's clinical diagnosis and any medication taken within the last 4 hours (name, class, and posology) that could affect the study's outcome. Participants will also be asked if they have taken any medication (name) for attention or hyperactivity to assess its potential impact on their ability to be distracted by VR.

Primary Outcomes

The two primary outcomes of this study are to measure both observed and objective biomarker-based indicators associated with DFA.

Measures of Primary Outcomes

The first primary outcome will evaluate DFA using the VARS (Table 1) [25]. The VARS is an observation-based assessment by proxy for DFA and is among the most frequent anxiety scoring instruments for DFA [25]. The scale has been used as an anxiety rating scale in other studies that evaluated the efficacy of VR distraction in the management of DFA [32,33]. The anxiety scale consists of one item, with six defined anxiety levels that range from 0, representing the child being relaxed, smiling, and willing to cooperate, to 5, representing the child being out of contact with the reality of the threat. The highest score represents the highest anxiety level [25]. A high degree

XSL•FO
RenderX

of reliability was observed for this scale, even by untrained observers [25,34]. A research assistant will indicate a score from 0 to 5 on the anxiety scale at T0 and T1 (Table 3).

The second primary outcome will assess the mean difference in SAA levels between T0 and T2, 10 minutes after the dental procedure. This enzyme is linked to both adrenaline and noradrenaline and is a validated and reliable specific biomarker of autonomic nervous system activity. It has been proposed as an alternative biomarker for anxiety, over salivary cortisol, as it reaches peak levels within 10 minutes after a stressful procedure compared to cortisol, which requires a 30-minute delay [16,35]. Enzyme sampling has been performed in previous studies to evaluate DFA in children undergoing dental procedures [35-37]. This enzyme is considered an accurate and noninvasive procedure for assessing anxiety-related events such as dental examination, cleaning, restoration, or extraction in children [26,35]. Alpha-amylase sampling has been performed in previous studies to evaluate DFA for children undergoing dental procedures [35,38]. Salivary samples will be collected using a sterile dental cotton swab from under the tongue and stored immediately within a sterile vial in a research-specific on-site freezer at −80 degrees Celsius. The research assistant will collect the samples and take them to the laboratory. Results will be entered in REDCap for analysis purposes after the study.

Secondary Outcomes

The secondary outcomes include (1) parents' and health care professionals' satisfaction in both groups and children's satisfaction in the VR intervention; (2) changes in physiological parameters, such as heart rate and oxygen saturation during and after the intervention compared to baseline; (3) behavior during the procedure; (4) occurrence of side effects (nausea, dizziness, and vomiting); (5) length of the procedure; and (6) the effectiveness of VR in reducing pain (Table 3).

Measures of Secondary Outcomes

The secondary outcomes will be measured as follows:

1. Parents' or legal guardians' satisfaction with the intervention will be assessed using a Numerical Rating Scale (0 - 10, where 0 is very dissatisfied and 10 is very satisfied) and according to the recommended question by the Pediatric Initiatives on Methods, Measurement and Pain Assessment in Clinical Trials: "Considering anxiety relief, side effects, and emotional recovery, how satisfied were you with the intervention used to manage anxiety experienced by your child?" [39] (Multimedia Appendix 1).
2. Health care professionals' satisfaction with the intervention will be assessed using a 6-question satisfaction questionnaire, with a range of scores from 1 to 4 (1="Totally disagree," 2="Disagree," 3="Agree," and 4="Totally agree"; Multimedia Appendix 2).
3. Child satisfaction with VR will be assessed using a 0 - 10 numerical rating scale (where 0 is very dissatisfied and 10 is very satisfied; Multimedia Appendix 1).
4. Physiological parameters (heart rate and oxygen saturation) will be measured using the COVIDIEN Nellcor pulse oxygen saturation meter, an approved and validated device.
5. The patient's behavior during the procedure will be assessed by proxy via a research assistant using the Venham Behavioral Rating Scale at T0 and T1 (Table 3) [25,40]. This 6-point scale assesses children's behavior and provides information about patients with disruptive behavior. The scale points are based on objective, and readily observable behaviors and classify children's behavior into 6 categories ranging from 0 (total cooperation) to 5 (general protest, no compliance, or cooperation). This scale has demonstrated good interrater reliability and has proven to be a valuable instrument for assessing children's behavioral responses to dental procedures [40].
6. Side effects will be collected in children in both groups from arrival on site until discharge, using a checklist of side effects commonly associated with VR, namely nausea, dizziness, and vomiting. Any of these side effects will be recorded along with the time they will occur.
7. The length of the procedure in minutes will be measured and recorded for each participant and compared between the two groups.
8. Pain will be assessed using the Face, Legs, Activity, Cry, and Consolability—Revised scale (FLACC-R) [41] (Multimedia Appendix 3). This questionnaire will serve to assess VR efficacy in reducing pain by comparing the mean pain level of both groups (experimental and control intervention) during the dental procedure (T1).

The FLACC-R scale (Multimedia Appendix 3) is a tool developed by Malviya et al [41] to assess pain in infants and young children who are unable to communicate their pain verbally. Health care professionals use this scale to evaluate pain in children who may not be able to express their discomfort in words. The FLACC-R scale consists of five categories, each rated on a scale of 0 to 2, with 0 indicating no pain and 2 indicating severe pain. The categories are as follows:

1. *Face* (F): this category assesses the child's facial expressions. It looks at whether the child is relaxed, smiling, or showing pain-related facial expressions, such as frowning, grimacing, or jaw clenching. A score of 0 indicates a relaxed face, while a score of 2 indicates a highly distressed or painful expression.
2. *Legs* (L): this category considers whether the child's legs are relaxed or whether there is increased movement, kicking, or tensing of the leg muscles. A score of 0 implies relaxed legs, while a score of 2 indicates significant leg movement due to pain.
3. *Activity* (A): this category looks at whether the child is calm and content or if they are restless, agitated, or unable to stay still. A score of 0 indicates calmness, while a score of 2 indicates high activity and restlessness due to pain.
4. *Cry* (C): this category assesses the quality of the child's cry, considering whether the child is not crying, whimpering, or having a high-pitched, continuous cry. A score of 0 means no crying, while a score of 2 implies a high-pitched, continuous cry.
5. *Consolability* (C): the *consolability* considers whether the child can be comforted easily or if they remain inconsolable despite efforts to soothe them. A score of 0 suggests easy

XSL•FO
RenderX

consolability, while a score of 2 indicates that the child cannot be consoled.

The total score on the FLACC scale is obtained by summing the scores across the 5 categories and ranges from 0 (no pain) to 10 (severe pain). The FLACC-R score will be evaluated at T0 (baseline) and at 10 minutes after the start of the procedure or at midprocedure for shorter dental treatments (T1).

Study Proceedings Including Data Collection

Participants will be screened by a resident dentist or a research assistant via the appointment scheduling system for any dental procedure. Since DFA can arise from multiple procedures, such as simple cleaning to teeth extraction. A research assistant will review consent with participants and parents. Written consent by the parents or legal guardian—including the child's assent (if possible)—will be obtained on arrival at the clinic, the day of the procedure, while completing the preprocedural questionnaire. After verifying that the participant meets all inclusion and exclusion criteria, the research assistant will log into REDCap only 10 minutes before the start of the intervention to minimize the risk of bias toward the intervention, obtain the group allocation, and then announce it to the participant. Because of the nature of VR, no blinding will be possible to staff, participants, parents, or legal guardians. Baseline data collection prior to the dental procedure, including a sociodemographic questionnaire, recording of physiological parameters, and SAA sampling, will take approximately 15 minutes to complete. The VR headset will be adjusted to the child's size, and approximately 5 minutes will be allotted for the child to familiarize themselves with the room, equipment, and the game (VR or cartoon on TV or tablet) before the start of the procedure. The intervention will last the entire dental procedure, and its duration will be documented for each patient. DFA assessment by proxy using the VARS will be performed on site by the research assistant present. As per the clinic's protocol, if a child becomes restless and uncooperative, they will be held by parents for the remainder of the procedure if it cannot be safely stopped at this time or if the procedure is considered an emergency. If the procedure can be safely stopped, rescheduling will be discussed with the parents, including the possible need for sedation. The involvement of parents will be recorded, as will any other distractions used during treatment, such as stickers, teddy bears, music, and others. The occurence of side effects and medications used will be recorded. The former will be transmitted to the ethics committee if they are serious or life-threatening. As VR has the potential to cause nausea, dizziness, and motion sickness, if participants experience any of these, they will be managed in accordance with the clinic's protocols.

If participants experience severe side effects, the treatment will be safely discontinued. All instruments in the mouth will be removed. The VR device will be removed from the participant. The dental chair will be adjusted to a comfortable position, and the patient will rest until they feel comfortable and recovered. The option to continue the treatment (without the VR headset) or to postpone it to another day will be discussed. Postprocedure SAA sampling and completion of questionnaires on parents' and health care professionals' satisfaction levels will take approximately 10 minutes. Equipment used by children will be disinfected and reset before the next participant.

Statistical Analysis

The primary analyses for this study will follow an intent-to-treat approach, using randomization status as the independent variable, and the mean difference in anxiety (using the anxiety subscale of the Venham) during the dental procedure (T1) and the SAA mean difference between T0 and T2 as the primary outcomes. As there are two primary outcomes (anxiety subscale and SAA), a Bonferroni correction will be applied, and a significance level of P=.025 will be used to determine statistical significance of primary outcomes. An ANCOVA, adjusting for unequal variances and adjusted for the time of the procedure, age, and baseline (T0) anxiety or behavioral score, will be used to assess the mean difference in Venham anxiety score during the dental procedure (T1) between the experimental and the control groups. An ANCOVA, adjusted for time of procedure, neurodevelopmental diagnosis, age, and baseline SAA level (T0), will be used to assess the mean difference in SAA level between groups 10 minutes after the procedure (T2).

For the secondary outcomes analyses, parametric tests adjusted for procedural type, age, sex, and parental presence will be used. Comparisons of dichotomous variables, including the necessity of procedural retakes, the use of other nonpharmacological interventions, and the occurrence of side effects, will be assessed using the Cochran-Mantel-Haenszel test. We will also perform subgroup analyses by age group (6 - 12 y vs 13 - 17 y) and sex and gender (boys, girls, or other) using ANCOVA. Any side effects will be reported using the Medical Dictionary for Regulatory Activities terminology [42], and their proportions will be compared between groups.

Ethical Considerations

A research assistant will review consent with participants and parents. The information and consent form will be signed by one of the parents on the day of the visit. At enrollment, children are assigned a code used on all data collection forms to protect their confidentiality. Patient identifying information is kept separate from the case reports and is linked to study enrollment by a study ID number. The consent form also identifies all the information that may be collected during this study. All data collected on paper from the day of the intervention will subsequently be transferred into REDCap. No recording of the child's screen will be collected, nor will their game inputs. As per usual, all information, including links between patient identifiers and study ID, and paper copies of forms and questionnaires, will be stored and double-locked in a cabinet for 7 years (starting from the end of the trial) at the principal investigator's office at the research center. This trial was approved by the research and ethics board of the Centre Hospitalier Universitaire Sainte-Justine (#2024 - 5983). The participants received no compensation.

Results

This study will be conducted from November 2023 to December 2025. As of November 2025, 300 participants have been recruited. Results are expected to be available in June 2026.

XSL•FO
RenderX

The study findings will be disseminated through peer-reviewed publications and presentations at major dental conferences.

Discussion

Anticipated Findings

The literature reports that some VR games can induce cybersickness, such as nausea or dizziness associated with computer games, especially in younger individuals with motion sickness susceptibility, or in those who have experienced recent headaches [43]. However, the game used in this study was specifically designed for pediatric patients with reduced speed and eye tracking to minimize these side effects. If the participant feels any symptoms, the game will be stopped and the HMD will be removed. Should the participant experience an injury of any kind following the administration of the study intervention or following any other procedure related to this research, they will receive the appropriate care and services required by their condition.

Conclusions

The results of the study could be used to better understand how VR distraction reduces dental anxiety and pain in children with SHCN, as well as the satisfaction of parents, patients, and health care professionals with its use during dental procedures. This study may also lead to more comfortable and less traumatic dental care for pediatric patients. In turn, VR can be a valuable tool for managing anxiety in children, which can be a disturbing and anxiety-inducing experience for both parents and health care professionals during dental procedures.

Funding

This study was funded by the MayDay Funds Project Grant.

Data Availability

Once the study is completed, data supporting the findings will be available by contacting the corresponding author.

Authors' Contributions

Study design and conception: SLM, MEA, WW, CG.
Data supervision and acquisition: SLM, MEA, WW, JG, GG, DP, CF, AG, PO, VM, KS-A.
Data analysis: SLM, JG, VM.
Manuscript writing: SLM, JG, WW, VM.
Manuscript revision: SLM, JG.
All authors have approved the final manuscript version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Parental and participant satisfaction questionnaire.
[DOCX File, 52 KB - resprot_v15i1e83672_app1.docx]

Multimedia Appendix 2
Health care professional satisfaction questionnaire.
[DOCX File, 41 KB - resprot_v15i1e83672_app2.docx]

Multimedia Appendix 3
Face, Legs, Activity, Cry, and Consolability—Revised scale scoring.
[DOCX File, 16 KB - resprot_v15i1e83672_app3.docx]

Checklist 1
SPIRIT Checklist 2025.
[DOCX File, 35 KB - resprot_v15i1e83672_app4.docx]

References

1. Lee DW, Kim JG, Yang YM. The influence of parenting style on child behavior and dental anxiety. Pediatr Dent 2018 Sep 15;40(5):327-333. [Medline: 30355427]
2. Armfield JM. How do we measure dental fear and what are we measuring anyway? Oral Health Prev Dent 2010;8(2):107-115. [Medline: 20589243]

XSL•FO
RenderX

3. Cianetti S, Lombardo G, Lupatelli E, et al. Dental fear/anxiety among children and adolescents: a systematic review. Eur J Paediatr Dent 2017 Jun;18(2):121-130. [doi: 10.23804/ejpd.2017.18.02.07] [Medline: 28598183]
4. Klingberg G, Broberg AG. Dental fear/anxiety and dental behaviour management problems in children and adolescents: a review of prevalence and concomitant psychological factors. Int J Paediatr Dent 2007 Nov;17(6):391-406. [doi: 10.1111/j.1365-263X.2007.00872.x] [Medline: 17935593]
5. Seligman LD, Hovey JD, Chacon K, Ollendick TH. Dental anxiety: an understudied problem in youth. Clin Psychol Rev 2017 Jul;55:25-40. [doi: 10.1016/j.cpr.2017.04.004] [Medline: 28478271]
6. Murad MH, Ingle NA, Assery MK. Evaluating factors associated with fear and anxiety to dental treatment: a systematic review. J Family Med Prim Care 2020 Sep;9(9):4530-4535. [doi: 10.4103/jfmpc.jfmpc_607_20] [Medline: 33209758]
7. Luoto A, Lahti S, Nevanperä T, Tolvanen M, Locker D. Oral-health-related quality of life among children with and without dental fear. Int J Paediatr Dent 2009 Mar;19(2):115-120. [doi: 10.1111/j.1365-263X.2008.00943.x] [Medline: 19250394]
8. Beaton L, Freeman R, Humphris G. Why are people afraid of the dentist? Observations and explanations. Med Princ Pract 2014;23(4):295-301. [doi: 10.1159/000357223] [Medline: 24356305]
9. Humphris G, King K. The prevalence of dental anxiety across previous distressing experiences. J Anxiety Disord 2011 Mar;25(2):232-236. [doi: 10.1016/j.janxdis.2010.09.007] [Medline: 20952156]
10. Ten Berge M, Veerkamp JSJ, Hoogstraten J. The etiology of childhood dental fear: the role of dental and conditioning experiences. J Anxiety Disord 2002;16(3):321-329. [doi: 10.1016/s0887-6185(02)00103-2] [Medline: 12214817]
11. Shim YS, Kim AH, Jeon EY, An SY. Dental fear & anxiety and dental pain in children and adolescents: a systemic review. J Dent Anesth Pain Med 2015 Jun;15(2):53-61. [doi: 10.17245/jdapm.2015.15.2.53] [Medline: 28879259]
12. Oliveira MA, Vale MP, Bendo CB, Paiva SM, Serra-Negra JM. Influence of negative dental experiences in childhood on the development of dental fear in adulthood: a case-control study. J Oral Rehabil 2017 Jun;44(6):434-441. [doi: 10.1111/joor.12513] [Medline: 28386938]
13. Management of dental patients with special health care needs. : American Academy of Pediatric Dentistry (AAPD); 2025 URL: https://www.aapd.org/globalassets/media/policies_guidelines/bp_shcn.pdf?utm_source=chatgpt.com [accessed 2026-01-08]
14. Koneru A, Sigal MJ. Access to dental care for persons with developmental disabilities in Ontario. J Can Dent Assoc 2009 Mar;75(2):121. [Medline: 19267962]
15. Yon MJY, Chen KJ, Gao SS, Duangthip D, Lo ECM, Chu CH. An introduction to assessing dental fear and anxiety in children. Healthcare (Basel) 2020 Apr 4;8(2):86. [doi: 10.3390/healthcare8020086] [Medline: 32260395]
16. Vacaru RP, Didilescu AC, Sfeatcu R, et al. The effect of dental treatments in caries management on stress and salivary protein levels. J Clin Med 2022 Jul 27;11(15):4350. [doi: 10.3390/jcm11154350] [Medline: 35893440]
17. Morgan AG, Rodd HD, Porritt JM, et al. Children's experiences of dental anxiety. Int J Paediatr Dent 2017 Mar;27(2):87-97. [doi: 10.1111/ipd.12238]
18. Al-Khotani A, Bello LA, Christidis N. Effects of audiovisual distraction on children's behaviour during dental treatment: a randomized controlled clinical trial. Acta Odontol Scand 2016 Aug;74(6):494-501. [doi: 10.1080/00016357.2016.1206211] [Medline: 27409593]
19. Alshatrat SM, Sabarini JM, Hammouri HM, Al-Bakri IA, Al-Omari WM. Effect of immersive virtual reality on pain in different dental procedures in children: a pilot study. Int J Paediatr Dent 2022 Mar;32(2):264-272. [doi: 10.1111/ipd.12851] [Medline: 34133809]
20. Pagano S, Lombardo G, Coniglio M, et al. Autism spectrum disorder and paediatric dentistry: a narrative overview of intervention strategy and introduction of an innovative technological intervention method. Eur J Paediatr Dent 2022 Mar;23(1):54-60. [doi: 10.23804/ejpd.2022.23.01.10] [Medline: 35274543]
21. Cunningham A, McPolin O, Fallis R, Coyle C, Best P, McKenna G. A systematic review of the use of virtual reality or dental smartphone applications as interventions for management of paediatric dental anxiety. BMC Oral Health 2021 May 7;21(1):244. [doi: 10.1186/s12903-021-01602-3] [Medline: 33962624]
22. Virtual reality definition & meaning. Merriam-Webster. 2023. URL: https://www.merriamwebster.com/dictionary/virtual%20reality [accessed 2023-04-09]
23. Eijlers R, Utens E, Staals LM, et al. Systematic review and meta-analysis of virtual reality in pediatrics: effects on pain and anxiety. Anesth Analg 2019 Nov;129(5):1344-1353. [doi: 10.1213/ANE.0000000000004165] [Medline: 31136330]
24. Ram D, Shapira J, Holan G, Magora F, Cohen S, Davidovich E. Audiovisual video eyeglass distraction during dental treatment in children. Quintessence Int 2010 Sep;41(8):673-679. [Medline: 20657857]
25. Venham LL. The effect of mother's presence of child's response to dental treatment. ASDC J Dent Child 1979;46(3):219-225. [Medline: 285948]
26. Chaturvedi Y, Chaturvedy S, Marwah N, Chaturvedi S, Agarwal S, Agarwal N. Salivary cortisol and alpha-amylase-biomarkers of stress in children undergoing extraction: an in vivo study. Int J Clin Pediatr Dent 2018;11(3):214-218. [doi: 10.5005/jp-journals-10005-1514] [Medline: 30131644]
27. Wu W, Le May S, Hung N, et al. Effects of a virtual reality game on children's anxiety during dental procedures (VR-TOOTH): protocol for a pilot randomized controlled trial. JMIR Res Protoc 2023 Nov 10;12:e49956. [doi: 10.2196/49956] [Medline: 37948113]

XSL•FO
RenderX

28. Le May S, Tsimicalis A, Noel M, et al. Immersive virtual reality vs. non-immersive distraction for pain management of children during bone pins and sutures removal: a randomized clinical trial protocol. J Adv Nurs 2021 Jan;77(1):439-447. [doi: 10.1111/jan.14607] [Medline: 33098330]
29. Efird J. Blocked randomization with randomly selected block sizes. Int J Environ Res Public Health 2011 Jan;8(1):15-20. [doi: 10.3390/ijerph8010015] [Medline: 21318011]
30. Dream Dental: therapeutic virtual reality for dental procedures. Paperplane Therapeutics. 2025. URL: https://paperplanetherapeutics.com/en/produits/dream-dental [accessed 2026-01-08]
31. PICO. picoxr. 2025. URL: https://www.picoxr.com/global [accessed 2025-12-18]
32. Sahebalam R, Rafieinezhad R, Boskabad M. Comparison of the efficacy of Jilo Animation Approach versus conventional Tell-Show-Do (TSD) technique on cooperation and anxiety levels of children during dental practice: a randomized controlled clinical trials. J Dent (Shiraz) 2020 Dec;21(4):284-291. [doi: 10.30476/dentjods.2020.81897.1001] [Medline: 33344678]
33. Nunna M, Dasaraju RK, Kamatham R, Mallineni SK, Nuvvula S. Comparative evaluation of virtual reality distraction and counter-stimulation on dental anxiety and pain perception in children. J Dent Anesth Pain Med 2019 Oct;19(5):277-288. [doi: 10.17245/jdapm.2019.19.5.277] [Medline: 31723668]
34. Narayan VK, Samuel SR. Appropriateness of various behavior rating scales used in pediatric dentistry: a review. JGOH 2020;2(2):112-117. [doi: 10.25259/JGOH_64_2019] [Medline: 31429749]
35. Furlan NF, Gavião MBD, Barbosa TS, Nicolau J, Castelo PM. Salivary cortisol, alpha-amylase and heart rate variation in response to dental treatment in children. J Clin Pediatr Dent 2012;37(1):83-87. [doi: 10.17796/jcpd.37.1.n32m21n08417v363] [Medline: 23342572]
36. Alaki SM, Safi A, Ouda S, Nadhreen A. Comparing dental stress in new child patients and returning patients using salivary cortisol, immunoglobulin-A and alpha-amylase. J Clin Pediatr Dent 2017;41(6):462-466. [doi: 10.17796/1053-4628-41.6.8] [Medline: 28937902]
37. Wu W, Asselin ME, Hung N, et al. Effects of a virtual reality game on children's fear and anxiety during dental procedures (VR-TOOTH): a pilot randomized controlled trial. Journal of Medical Extended Reality 2024 Mar 1;1(1):137-146. [doi: 10.1089/jmxr.2024.0021]
38. AlMaummar M, AlThabit HO, Pani S. The impact of dental treatment and age on salivary cortisol and alpha-amylase levels of patients with varying degrees of dental anxiety. BMC Oral Health 2019 Sep 6;19(1):211. [doi: 10.1186/s12903-019-0901-7] [Medline: 31492133]
39. McGrath PJ, Walco GA, Turk DC, et al. Core outcome domains and measures for pediatric acute and chronic/recurrent pain clinical trials: PedIMMPACT recommendations. J Pain 2008 Sep;9(9):771-783. [doi: 10.1016/j.jpain.2008.04.007] [Medline: 18562251]
40. Venham LL, Gaulin-Kremer E, Munster E, Bengston-Audia D, Cohan J. Interval rating scales for children's dental anxiety and uncooperative behavior. Pediatr Dent 1980 Sep;2(3):195-202. [Medline: 6938934]
41. Malviya S, Voepel-Lewis T, Burke C, Merkel S, Tait AR. The revised FLACC observational pain tool: improved reliability and validity for pain assessment in children with cognitive impairment. Paediatr Anaesth 2006 Mar;16(3):258-265. [doi: 10.1111/j.1460-9592.2005.01773.x] [Medline: 16490089]
42. MedDRA version 28.0. MedDRA - Medical Dictionary for Regulatory Activities. 2025. URL: https://www.meddra.org [accessed 2025-12-08]
43. Garrido LE, Frías-Hiciano M, Moreno-Jiménez M, et al. Focusing on cybersickness: pervasiveness, latent trajectories, susceptibility, and effects on the virtual reality experience. Virtual Real 2022;26(4):1347-1371. [doi: 10.1007/s10055-022-00636-4] [Medline: 35250349]

Abbreviations

DFA: dental fear and anxiety
FLACC-R: Face, Legs, Activity, Cry, and Consolability—Revised
HMD: head-mounted device
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SAA: salivary alpha-amylase
SHCN: special health care needs
VARS: Venham Anxiety Rating Scale
VR: virtual reality

XSL•FO
RenderX

Edited by GRML Rosa; submitted 09.Sep.2025; peer-reviewed by AM Meldrum, K Horvath; accepted 24.Dec.2025; published 29.Jan.2026.

Please cite as:
Gardner J, Markopoulos V, Wu W, Gilbert G, Pelletier D, Fafard C, Gagné A, Guingo E, Genest C, Asselin MÈ, Ouimet P, St-Arneault K, Le May S
Efficacy of a Virtual Reality Game on Children's Fear and Anxiety During Dental Procedures (VR-TOOTH): Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e83672
URL: https://www.researchprotocols.org/2026/1/e83672
doi:10.2196/83672

© Julien Gardner, Vallerie Markopoulos, Wenjia Wu, Gabrielle Gilbert, Daphnée Pelletier, Charlotte Fafard, Anne Gagné, Estelle Guingo, Christine Genest, Marie-Ève Asselin, Pascale Ouimet, Kate St-Arneault, Sylvie Le May. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 29.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effect of Evening Technology Use on Objective Sleep in Older Adults: Protocol for a Crossover Randomized Controlled Trial

Sarah Nauman Ghazi, MD, Lic; Anders Behrens, MD, PhD; Joakim Niklasson, PhD; Johan Sanmartin Berglund, MD, PhD; Peter Anderberg, PhD

Department of Health, Blekinge Institute of Technology, Karlskrona, Sweden

Corresponding Author:
Sarah Nauman Ghazi, MD, Lic
Department of Health, Blekinge Institute of Technology, Karlskrona, Sweden

Abstract

Background: Evening technology use (ETU) has been associated with sleep disturbances, often attributed to blue light exposure and cognitive arousal. However, most of the existing evidence focuses on younger populations and relies primarily on subjective measures. As older adults increasingly engage with both passive and active technology use, it is important to investigate how ETU impacts objective sleep. Currently, there is also a limited understanding of how particular evening digital activities, especially active versus passive engagement, affect objective sleep in older adults.

Objective: This study aims to investigate the impact of exposure to ETU on both objective and subjective sleep outcomes in older adults.

Methods: This is a randomized crossover trial involving approximately 55 adults aged 60‐75 years from the ongoing Swedish National Study on Aging and Care – Blekinge. Each participant will undergo 3 one-week intervention periods: active ETU, passive ETU, and a nondigital activity (book reading), with one-week washout periods in between. The order of interventions will be randomized. Sleep will be assessed using a home-based electroencephalography device (MUSE headband) and daily self-reports. Primary outcomes are sleep onset latency and wake after sleep onset. Secondary outcomes include objective measures such as total sleep time, sleep efficiency, and time spent in REM, deep, and light sleep, subjective sleep quality, adherence, and perception of the intervention and comfort of using the objective measurement tool, that is, the electroencephalography headband. Linear mixed-effects models (with fixed effects for condition and period and a random participant intercept) were used to analyze crossover effects on sleep outcomes.

Results: Participant recruitment and data collection began in the fall of 2025 and will continue through summer 2026 or until the target sample size is reached. Data collection is scheduled to be completed by spring 2027. Results will include participant flow, baseline characteristics, adherence data, and comparative analyses of the 3 intervention conditions. Within-subject statistical models will be used to evaluate differences in sleep outcomes and investigate the associations between ETU and sleep quality.

Conclusions: This crossover study will clarify how active and passive ETU, compared with a nondigital activity, relate to objective sleep in older adults. Findings will inform simple, practical recommendations for technology use before bed in late life.

Trial Registration: ClinicalTrials.gov NCT07001514; https://clinicaltrials.gov/study/NCT07001514

International Registered Report Identifier (IRRID): DERR1-10.2196/84512

(JMIR Res Protoc 2026;15:e84512) doi:10.2196/84512

KEYWORDS

sleep; technology use; older adults; cross-over randomized controlled trial; gerontechnology; eHealth.

Introduction

The digital society has made it easy for its citizens to use technology for multiple purposes, from everyday tasks such as banking to entertainment, relaxation, and promoting health. Recent data indicate that about 90% of Swedes aged 60 years and older use the internet, with approximately 75% using the internet daily, showcasing digital engagement in later life [1], raising important questions about how this use may influence restorative behaviors such as sleep. In particular, evening technology use (ETU) before bed often becomes a routine; therefore, it is important to clarify its effects on sleep in order to protect sleep health among older adults.

Sleep is an important determinant of health and well-being, particularly in older adults, who may experience changes in sleep architecture and quality due to aging and lifestyle factors [2]. Even factors such as technology use may shape these sleep patterns. The increasing use of technology within the general

XSL•FO
RenderX

population and among older adults [1], especially ETU and its potential effects on sleep, has raised concerns among health care professionals and researchers. Traditionally, these concerns have focused on the impact of bright light or blue light on melatonin release, which can disrupt the circadian rhythm. However, this unidirectional view has been refined by a bidirectional model that recognizes that technology can disrupt sleep and vice versa through multiple mechanisms. These include sleep time displacement, the arousal hypothesis, and the bright light hypothesis. Consequently, poor sleep can also lead to increased technology use, as it fills the time and provides cognitive distraction [3]. According to the arousal hypothesis proposed by Bauducco et al [3], using engaging or active digital content close to bedtime (eg, video games or social media exchanges) increases a person's physiological and psychological arousal by increasing heart rate, alertness, or stress levels before bed. This, in turn, makes it harder to fall asleep quickly or stay asleep.

Existing research also highlights the influence of digital device use on sleep through mechanisms such as blue light exposure [4], cognitive stimulation [5], and emotional arousal [6,7]. Much of the current research on technology and sleep focuses on children, adolescents, and younger adults [8-10]. Far fewer studies have investigated whether and how these mechanisms generalize to older adults.

Older adults use technology for social interaction [11], entertainment [12], cognitive exercises [13], and overall well-being [14]. Older adults between the ages of 60 and 75 years are often overlooked in current research on sleep. This age group is old enough to experience age-related changes in sleep patterns, yet young enough to minimize the effects of frailty, multiple health conditions, and the use of multiple medications, which could skew the results. Additionally, their experiences with technology and sleep may differ from those of younger populations [15].

Additionally, little is known about how the type of ETU, whether passive (such as watching a documentary) or active (such as playing a video game), affects sleep compared with traditional, nondigital presleep habits. There has been some research that suggests that watching nature videos can promote relaxation, reduce stress, and have positive effects on cognitive and emotional health [10], which are factors that could positively influence sleep quality. Furthermore, it has also been shown that engaging in complex cognitive tasks before sleep can significantly improve both objective and subjective sleep quality, particularly for individuals who typically struggle with sleep issues [11]. On the other hand, video gaming before bed is associated with delayed sleep onset in adolescents [7]. Conversely, some research suggests that casual gaming may not significantly disrupt sleep, indicating that the type and intensity of gaming matter [8,16]. Moreover, to measure sleep objectively is important as it provides reliable, instrument-based evidence of sleep health that complements subjective reports and reduces bias [17,18]. Therefore, the aim of the study is to investigate the impact of exposure to ETU on both objective and subjective sleep outcomes in older adults.

Methods

Objectives

Study Design

This study adopts a crossover randomized controlled trial (RCT) to examine how ETU influences objective and subjective sleep in older adults aged 60 to 75 years. Three conditions will be compared: (1) passive ETU, (2) active ETU, and (3) nondigital activity.

After completing an initial baseline assessment via survey, participants will undertake a 5-week schedule composed of 3 separate, one-week intervention periods. At baseline, each participant is randomized to one of six possible sequences of the 3 conditions (passive ETU, active ETU, nondigital; 3!=6). Participants then complete 3 one-week intervention periods in their assigned order, separated by 2 one-week washouts (total duration: 5 wk). We chose a 7-day washout period because systematic reviews of digital detox interventions often recommend this duration for detoxification [19].

During each intervention week, participants will maintain a daily evening log to record subjective sleep quality, ETU, and the comfort of using a wearable headband. Objective sleep outcomes will be assessed using electroencephalography (EEG) collected via a headband (Muse-S Athena, InteraXon, Toronto, ON, Canada), while self-reported logs capture subjective experiences. This crossover design allows for within-subject comparisons, thereby reducing variability, controlling for individual differences in sleep patterns, and confounding factors, while enhancing statistical power.

Population and Recruitment

Participants in this study will be recruited from the Swedish National Study on Aging and Care – Blekinge (SNAC-B) cohort. SNAC-B is part of a large, national, population-based research initiative in Sweden focused on individuals aged 60 years and older [20]. The study is conducted at the Research and Development Clinic affiliated with the Department of Health at Blekinge Institute of Technology (BTH).

The SNAC-B population offers a well-characterized cohort of older adults, increasing the likelihood of identifying individuals who meet the inclusion criteria for this study. Eligible participants will be directly contacted by the research team, ensuring both transparency and voluntary participation. Detailed study information will be provided before obtaining informed consent.

Participants eligible for inclusion in the study will be aged between 60 and 75 years and must be regular users of digital devices such as smartphones or tablets. They should be willing and able to use an EEG headband for sleep monitoring throughout the intervention period. Additionally, participants must be able to take part in either the RCT, follow-up interviews, or both, depending on study needs and personal preference.

Individuals will be excluded from participation if they have a diagnosed sleep disorder or severe cognitive impairment. Use of medications that significantly influence sleep, such as

XSL•FO
RenderX

melatonin or benzodiazepines, will also serve as an exclusion criterion.

Eligible participants aged 60 - 75 years will be identified from the SNAC-B cohort. A random number will be generated using Excel to generate a random order of potential participants, who will be contacted sequentially until the target sample size of 55 is reached. One investigator (SNG) generated the randomization list for the 6 possible condition sequences using computer-generated random numbers in Excel. Random numbers, not participant IDs, determined assignment. Sequence codes were kept in a secure file on the university's server and accessed only by the authors of this study.

Before analysis, a coauthor not involved in analysis (AB) will replace all participant IDs with pseudonymous IDs; the ID key is stored separately. AB will also remap the 3 conditions to neutral labels (eg, X/Y/Z). The primary analysts (SNG and JN) will conduct the primary analyses on masked labels and pseudonymous IDs. Unmasking will occur only after the analysis code is finalized and the results are locked.

The sample size is determined through performing a priori power analysis, conducted using effect size estimates from Leger et al [21] on evening screen use and sleep in older adults. Details of power analysis are shown in Multimedia Appendix 1. Recruitment will proceed in waves of 10 participants for logistical feasibility. All enrolled participants will complete all 3 intervention conditions in a randomized order.

Intervention

Overview

This study uses a crossover RCT design in which participants progress through 3 different intervention conditions: passive ETU, active, and nondigital activity, in a randomized order. Each intervention phase lasts for one week, followed by a one-week washout period to reduce carryover effects, resulting in a total study duration of 5 weeks (Figure 1).

Figure 1. Schematic of the randomized crossover trial design. Participants complete three 1-week conditions in randomized order, separated by 1-week washout periods, following baseline investigation.

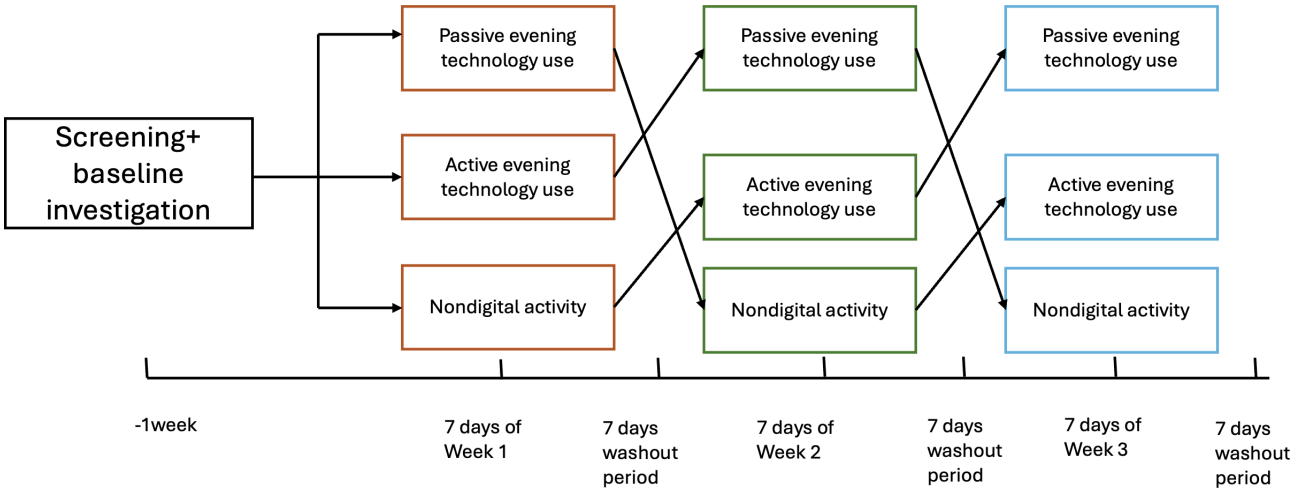

Baseline Assessment

Before the first intervention phase, participants complete a baseline survey (in Swedish language) assessing sociodemographic characteristics, subjective sleep health, chronotype, lifestyle factors, sleeping habits, general health status, and technology use (Multimedia Appendix 2).

Main Intervention

Participants will engage with the assigned evening activity for 30 - 60 minutes within the final 60 minutes before intended bedtime. To minimize variability in the interventions, the digital activities are standardized. To reduce the likelihood of accidentally falling asleep, participants should complete the evening activity seated (not lying in bed) and move to bed only when ready for lights-out. If they fall asleep during the activity before starting the MUSE recording, they should start the recording as soon as they notice and select "fell asleep before recording" in the morning survey. For all technology sessions, participants will set the device display to medium brightness (≈50% of the slider) at the start of the activity and maintain this setting throughout. Auto/adaptive brightness should be turned off to prevent fluctuations. The activity may continue until the participant is ready to sleep (lights-out). Immediately before lying down to sleep, participants will put on the MUSE headband, start the recording, then switch off the lights and attempt to sleep. No additional screen use is permitted after recording starts. Compliance will be checked through self-reporting in the daily log.

Passive Evening Technology Use

Participants watch a curated selection of soothing documentaries, those narrated by Sir David Attenborough or the secrets of the forest on SVT Play, for 30 - 60 minutes each evening, one hour before bedtime, using smartphones or tablets.

Active Evening Technology Use

Participants engage with the interactive word-based game Ruzzle for 30 - 60 minutes each evening, one hour before bedtime [2], using smartphones or tablets. It is to provide mental stimulation and engage participants actively.

XSL•FO
RenderX

Nondigital Activity

Participants read a nonfiction book of their choice for the same duration, avoiding all digital screens for one hour before bedtime.

Daily Logs

Throughout each intervention week, participants maintain a daily log (in Swedish) of their intervention adherence, morning assessments of subjective sleep quality, and the feasibility or comfort of using the EEG headband. The daily survey also includes an open-ended item asking whether anything unusual since the prior evening could have affected sleep, with examples written such as more caffeine than usual (especially after ≈4 PM) and alcohol (Multimedia Appendix 2).

Measures

The baseline and daily log surveys, along with the sociodemographic and subjective measures, are attached in Multimedia Appendices 2 and 3.

Sociodemographic and General Measures

The included variables were age, gender, living status, economic situation, employment situation, education, health status, sleep health scores, and the reduced Morningness–Eveningness Questionnaire (for chronotype) [22,23] (Multimedia Appendices 2 and 3 contain detailed questions).

Primary Outcome Measures

Overview

The primary outcomes of this study focus on changes in objective sleep onset latency (SOL) and wake after sleep onset (WASO) assessed using objective data collected via the MUSE EEG headband. This wearable device records brainwave activity throughout the night, enabling detailed analysis of various sleep parameters. The primary sleep parameters are provided below.

Sleep Onset Latency (SOL)

This refers to the duration of time it takes for a participant to transition from full wakefulness to sleep. Sleep latency will be automatically calculated by the MUSE EEG headband as the time from the start of recording to the first detected epoch of NREM sleep.

Wake After Sleep Onset (WASO)

WASO measures the cumulative amount of time (in minutes) a participant spends awake after initially falling asleep until final morning awakening. Elevated WASO values may indicate fragmented or disrupted sleep, which can negatively affect sleep quality and daytime functioning. It is automatically computed by summing detected wake periods within the sleep period time by the MUSE device algorithm.

Secondary Outcomes Measures

Overview

In addition to the objective measures of sleep, the study will assess several secondary outcomes that capture participants' subjective experiences, intervention adherence, and device usability.

Total Sleep Time (TST)

Total sleep time (TST) is the total amount of time spent asleep during the night, excluding any periods of wakefulness after sleep onset. It reflects the overall quantity of sleep and is a core indicator of sleep sufficiency. Derived directly from the MUSE's sleep stage classification, summing up the total minutes spent in light, deep, and REM sleep stages.

Sleep Efficiency (SE)

Sleep efficiency (SE) is the percentage of time in bed actually spent sleeping. It is calculated by dividing TST by the total time spent in bed and multiplying by 100. Higher SE suggests less wakefulness during the night and better overall sleep quality. SE is calculated automatically by the MUSE device algorithm.

Sleep Stages

The EEG data will be analyzed to determine the time spent in different sleep stages: REM (rapid eye movement) sleep, deep sleep (slow-wave sleep), and light sleep. These are expressed in percentages.

Subjective Sleep Quality

Each morning, participants will complete a low-burden Single-Item Sleep Quality Scale, a validated tool designed to assess overall perceived sleep quality [24]. We chose this brief item to minimize burden and support daily completion in older adults. This brief measure captures participants' evaluations of how well they slept the previous night, ranging from 0 (terrible) to 10 (excellent). Subjective ratings are analyzed alongside objective EEG outcomes as complementary indices of sleep and are tested across passive ETU, active ETU, and nondigital activity.

Intervention Adherence and Perception

Participants will be asked whether they were able to complete the assigned evening activity the previous night. If the answer is "yes," they will also report how long they engaged in the intervention (in minutes). This self-report measure will help evaluate the feasibility and real-world compliance with each intervention condition.

To assess how participants perceived each evening activity, they will complete a brief daily rating scale immediately after the intervention. The scale includes 5 items rated from 0 (not at all) to 10 (extremely), adapted from established tools: the Visual Analog Mood Scales for calmness [25], the Positive and Negative Affect Schedule for enjoyment and stress [26], and the User Engagement Scale for engagement and perceived meaningfulness [27]. This approach allows for low-burden, repeated assessment of affective and cognitive responses relevant to sleep-related outcomes in older adults. The survey is included in Multimedia Appendices 2 and 3. The internal consistency of the items (Cronbach α), once the data is collected, will be reported in the main study.

Comfort Level of EEG Headband

To assess usability and participant burden, participants will rate the comfort of wearing the EEG headband during the night using a 5-point Likert scale. The scale ranges from 1 (very uncomfortable) to 5 (very comfortable). This data will inform

XSL•FO
RenderX

the acceptability of the MUSE EEG headband as a sleep monitoring tool and identify any potential discomfort that may influence sleep or compliance.

Statistical Analysis Plan

Analyses will be conducted in SPSS Statistics (version 29.0.x; IBM Corp, Armonk, NY); the exact patch (eg, 29.0.2) will be reported at analysis. Descriptive statistics will be used to summarize demographic information, baseline sleep characteristics, chronotype prevalence, overall intervention adherence (percentage of nights completed; minutes engaged), perception ratings (calmness, stress, engagement, enjoyment, meaningfulness on 0 - 10), and EEG headband comfort (1 - 5 Likert), presented as mean (SD) or median (IQR), as appropriate.

Primary and secondary sleep outcomes will be derived from MUSE EEG auto-scoring using the device's standard algorithm. This study does not include manual rescoring or a new polysomnography comparison; instead, we rely on published validation studies demonstrating agreement with polysomnography/manual scoring [28,29]. All analyses will use participant-level week means (nightly values averaged within each intervention week/condition). Before analysis, we will check data quality. A night is classified as VALID if the device records ≥4 hours within the main sleep window and yields plausible values (eg, TST 2 - 12 h, SOL 0 - 180 min, WASO 0 - 240 min); otherwise, it is INVALID. All the valid nights will be included in the analysis. During the analysis, the data analyst will be blinded to intervention allocation. Unblinding will occur only after the analysis code is finalized, outputs are generated, and the primary results are locked. Participant-level week means (nightly values averaged within each intervention week/condition) will be calculated for descriptive purposes.

The coprimary outcomes are SOL and WASO derived from the MUSE EEG. The primary objective compares combined ETU versus nondigital activity, where ETU is defined by collapsing the 2 technology conditions (passive and active) into a single ETU category. This will be tested as a prespecified linear contrast of the condition marginal means from a linear mixed-effects model (LMM). For each primary outcome (SOL and WASO), we will fit an LMM with fixed effects for condition and period and a participant-specific random intercept. The LMM analyses will use all available valid nights. In addition to the prespecified combined-ETU contrast, we will estimate all pairwise condition contrasts (active vs passive, active vs nondigital, passive vs nondigital) using the Tukey adjustment. Model-based estimates, 95% CIs, and *P* values will be reported, together with a standardized within-participant effect size. Multiplicity control across the 2 coprimary endpoints (SOL and WASO) will be performed using the Holm procedure with a family-wise α of .05 (2-sided). If at least one coprimary endpoint is statistically significant after the Holm adjustment, we will proceed to confirmatory testing of the prespecified key secondary outcomes; otherwise, these secondary outcomes will be reported as exploratory. Model assumptions will be checked; if needed, outcomes will be transformed or analyzed using an appropriate generalized mixed model.

Key secondary objective outcomes are TST, SE, and sleep stages (REM%, Deep% [SWS], Light%). These outcomes will be analyzed using the same LMM framework (fixed effects for condition and period; random participant intercept), with Tukey adjustment for pairwise condition comparisons. To control for multiplicity across the key secondary outcomes, *P* values will be adjusted using Holm's procedure (2-sided α=.05), applied only if the gatekeeping criterion on the coprimary outcomes is met. All remaining outcomes (eg, subjective sleep, adherence, and comfort) will be treated as exploratory and interpreted primarily via effect estimates and 95% CIs.

We will briefly explore how subjective sleep aligns with objective measures using simple descriptive statistics (mean, SD and median, IQR). The full analysis code will be included as a supplementary file (including SPSS syntax and output) when the main study is reported. The statistical plan for each research question is presented in Table 1.

XSL•FO
RenderX

Table . Research questions with hypotheses, planned analysis, and expected endpoints.

Objectives	Research question	Hypothesis (H_1)	Null hypothesis (H_0)	Endpoints and analysis
Primary objective	Does evening technology use (ETU) affect primary objective sleep outcomes (WASO[a] and SOL[b]) compared with engaging in nondigital activity like reading physical books?	Older adults exposed to ETU before bedtime will show significant impairments in primary objective sleep outcomes than those exposed to nondigital activity.	Older adults exposed to ETU before bedtime will show no significant impairments in primary objective sleep outcomes than those exposed to nondigital activity.	• Primary outcomes: SOL, WASO (MUSE EEG)[c]. Linear mixed-effects model with fixed effects for condition and period and a random participant intercept. Multiplicity across SOL and WASO controlled using Holm (α=.05, 2-sided). Pairwise condition comparisons reported with Tukey adjustment. • Report estimated marginal mean 95% CIs, *P* values, and Cohen *d* (standardized within-participant effect sizes).
Key secondary	Does passive ETU[d] lead to better objective sleep outcomes than active ETU?	Passive ETU will show significant differences in sleep outcomes than active ETU.	Passive ETU will lead to no difference in sleep outcomes than active ETU.	• SOL, WASO, SE[e], TST[f], Sleep Stages (REM%[g], Deep%, Light%). LMM[h] with fixed effects for condition and period and a random participant intercept. Confirmatory testing follows a hierarchical (gatekeeping) strategy: these outcomes are formally tested only if ≥1 coprimary endpoint is significant after Holm adjustment. Multiplicity across key secondary outcomes controlled using Holm (α=.05). Pairwise condition comparisons use Tukey adjustment.

XSL•FO
RenderX

Objectives	Research question	Hypothesis (H_1)	Null hypothesis (H_0)	Endpoints and analysis
Key secondary	Does passive ETU lead to better objective sleep outcomes than nondigital activity[i]?	Passive ETU will show significant differences in sleep outcomes than nondigital activity.	Passive ETU will show no significant differences in sleep outcomes than nondigital activity.	• SOL, WASO, SE, TST, Sleep Stages (REM%, Deep%, Light%). LMM with fixed effects for condition and period and a random participant intercept. Confirmatory testing follows a hierarchical (gatekeeping) strategy: these outcomes are formally tested only if ≥1 coprimary endpoint is significant after Holm adjustment. Multiplicity across key secondary outcomes controlled using Holm (α=.05). Pairwise condition comparisons use Tukey adjustment.
Key secondary	Does active ETU[j] lead to better objective sleep outcomes than nondigital activity?	Active ETU will show significant differences in sleep outcomes than nondigital activity.	Active ETU will show no significant differences in sleep outcomes than nondigital activity.	• SOL, WASO, SE, TST, Sleep Stages (REM%, Deep%, Light%). LMM with fixed effects for condition and period and a random participant intercept. Confirmatory testing follows a hierarchical (gatekeeping) strategy: these outcomes are formally tested only if ≥1 coprimary endpoint is significant after Holm adjustment. Multiplicity across key secondary outcomes controlled using Holm (α=.05). Pairwise condition comparisons use Tukey adjustment.
Key secondary	Are there differences in subjective sleep quality across passive ETU, active ETU, and nondigital activity?	Subjective sleep quality will differ across conditions.	Subjective sleep quality does not differ across conditions.	• Single-Item Sleep Quality (0 - 10). Exploratory Analysis using LMM with fixed effects for condition and period (and sequence if included) and a random intercept for participant. Report model-based estimated marginal means for each condition and pairwise contrasts with Tukey-adjusted 95% CIs (*P* values reported descriptively)
Secondary			Dropout and nonadherence rates are equal across all 3 conditions.	

XSL•FO
RenderX

Objectives	Research question	Hypothesis (H_1)	Null hypothesis (H_0)	Endpoints and analysis
	What do participants' adherence to the assigned evening activity and their perceived experience of the activity look like over the study?	Dropout and nonadherence rates will be higher in the active ETU condition compared with passive and nondigital activity.		• Adherence: % of nights with ≥30 min activity; mean engagement time. Exploratory Analysis using LMM with fixed effects for condition and period (and sequence if included) and a random intercept for participant.
Implementation	What is the overall comfort/usability of the MUSE EEG headband during sleep?	(Exploratory, no formal hypothesis)	Not applicable	• Comfort rating (1 - 5 Likert). Exploratory/descriptive only. Report mean (SD), median (IQR), % nights rated ≥4.

[a]WASO: wake after sleep onset.

[b]SOL: sleep onset latency.

[c]EEG: electroencephalography.

[d]Passive ETU: passive evening technology use (watching a documentary on SVT PLAY app).

[e]SE: sleep efficiency.

[f]TST: total sleep time.

[g]REM: rapid eye movement.

[h]LMM: linear mixed-effects model.

[i]Nondigital activity (reading a book).

[j]Active ETU: active evening technology use (playing a video game called Ruzzle).

Ethical Considerations

The study will adhere to ethical standards as outlined by the Swedish research ethics framework. Ethical approval was obtained from the Regional Ethical Review Board prior to study initiation (Dnr 2025-02006-01). All participants will receive comprehensive information about the study's aims, procedures, potential risks, and their rights before providing written informed consent.

Participant confidentiality will be maintained through strict data management procedures. All data will be stored securely, ensuring that individual identities cannot be traced. All study data are stored on an encrypted, firewall-protected BTH server. Records are pseudonymized; identifiers are not stored with research data. Participation in the study is entirely voluntary, and participants may withdraw at any time without consequence. Procedures comply with GDPR, including participants' rights to access, rectification, and deletion.

Results

Participant recruitment and data collection commenced in the Fall of 2025. Recruitment is ongoing and will continue through Summer 2026 or until the target sample size is reached. The completion of data collection is projected for Spring 2027. We hypothesize that older adults exposed to ETU before bedtime will show significant differences in primary objective sleep outcomes compared with those exposed to nondigital activity. We also hypothesize that Passive and Active technology use will show significant differences in sleep outcomes among themselves and versus nondigital activity.

In the final trial report, the following data will be presented.

1. Number of participants screened, enrolled, randomized, and completed each intervention phase
2. Reasons for dropouts, exclusions, and nonadherence during any of the 3 conditions
3. Baseline characteristics of all the variables measured
4. Results of the statistical analysis based on the research questions of the primary and secondary outcome measures

Discussion

Principal Findings

This study aims to evaluate how ETU and types of evening activities, specifically Active technology use, Passive technology use, and nondigital activity, affect both objective sleep parameters and subjective sleep quality in older adults. In this 3-period crossover study of older adults, we hypothesize that, compared with a nondigital presleep activity, both passive and active ETU will be associated with longer SOL and greater WASO on EEG. We further expect active technology use to show the largest effects. As a secondary expectation, subjective sleep quality ratings will be lower on technology-use nights and will broadly follow the objective patterns.

A previous study revealed in OA that screen use before bed is positively associated with subjective sleep health [15], contrary to the conventional opinion that screen use has a negative

XSL•FO
RenderX

relationship in the younger population [30-32]. This variability is an empirical motivation to test whether effects differ by activity type. Our within-person crossover study compares active and passive ETU to a nondigital activity, allowing us to determine if any effects are specific to the activity without inferring mechanisms. Evening activities can affect presleep arousal through 2 main routes: light exposure from screens (blue-enriched light) [3,33] and cognitive/emotional engagement [5]. The study expects that active ETU (interactive gaming) will produce the highest engagement/arousal and screen exposure; passive ETU (neutral, low-arousal videos) will produce moderate-light engagement; and nondigital reading will be lowest on both. Accordingly, we hypothesize that active ETU has a greater impact on disrupting objective sleep and lower subjective sleep quality compared with passive ETU and nondigital technology nights.

Older adults are an important population for studying the underlying mechanisms related to sleep. Age-related changes in circadian and homeostatic regulation often lead to earlier sleep onset, reduced amplitude, and more fragmented sleep [34]. Consequently, factors like presleep arousal through the use of technology become especially significant. Additionally, changes in ocular and retinal function can affect spectral sensitivity, meaning the way older adults perceive light may differ from that of younger individuals [35].

Digital usage patterns among older adults also vary, typically involving more passive activities, such as viewing content, communicating, and playing simple games. These behaviors influence cognitive and emotional engagement [36]. Therefore, it is important, both demographically and conceptually, to investigate how different types of evening activities, such as those involving screen light and engaging content, affect sleep quality in older adults.

Unlike much of the existing research, which focuses primarily on children and adolescents [32,37], this study examines an underexplored demographic: older adults who are increasingly active users of digital technology and often have health issues linked to poor sleep.

Methodological Consideration

This study will use a randomized crossover design. This within-subject approach limits confounding, minimizes the influence of inter-individual variability, and increases statistical power, which is especially important given the moderate sample size [3,4]. Objective sleep measures will be captured via a wearable EEG headband (MUSE), allowing for naturalistic, at-home sleep tracking that is both minimally intrusive and ecologically valid as it supports generalizability to real-world settings [29]. Additionally, participants will complete daily self-report ratings assessing their subjective sleep quality and mood. Single-item sleep quality measures have limitations because they lack the detail needed to capture the multi-dimensional nature of sleep problems, leading to issues with sensitivity, content validity, and reliability. However, the item is low-burden and has good construct validity [24]. The inclusion of a brief affective experience scale, adapted from validated instruments such as the Visual Analog Mood Scales [25], Positive and Negative Affect Schedule [26], and the User Engagement Scale [27], enables examination of psychological and experiential factors that may mediate the relationship between ETU type and sleep outcomes. Combining objective EEG data with subjective daily logs will provide a holistic, multidimensional understanding of sleep quality [18].

In this randomized crossover trial, participants will receive structured instructions and technical support to guide them in using the EEG headband and completing the assigned activities. These elements are designed to support compliance, usability, and data completeness. In this study design, there is a bias due to the carryover effect (residual effects of one condition influencing the next) and period/sequence effects (time- or order-related differences) [38]. These biases will be mitigated with a one-week washout period and by randomizing condition order, analyzing participant-level week means (which dampens night-to-night variability). This will also help to reduce weekend bias, as weekend sleep might differ from weekday sleep.

It is important to consider certain limitations of the study design and the expected results. Randomized crossover trials might increase the risk of participant dropout because, for a participant's data to be included in the analysis, they must complete all intervention periods. To reduce attrition and support adherence, participants receive direct contact details (phone and email) for the lead author (SNG) and the university research clinic. They are also welcome to visit the clinic if they wish to discuss any issues. Recruitment and the initial information meeting are scheduled flexibly with 4 alternative time slots. Adherence is monitored via the daily log. If participants drop out before finishing all interventions, their data may be unusable, reducing the effective sample size and potentially introducing bias [39]. Although the crossover design helps control for inter-individual differences, subjective perceptions of each intervention may still influence self-reported sleep outcomes (eg, enjoyment or bias toward novelty). Nonadherence and usability challenges with the EEG device may affect the implementation of the intervention. Some participants may not follow the assigned activity consistently or may experience discomfort while using the MUSE EEG headband, which can impact their sleep. Technical issues or limited digital literacy could further reduce compliance or introduce usage bias [40]. Therefore, we will report the number of valid nights per condition and provide an attrition/participant flow summary in our results manuscript, allowing readers to assess compliance. These risks will be mitigated through initial training, daily logs, and reminders, but complete control is not guaranteed. Regarding the results, the findings will reflect the experiences of relatively healthy, digitally literate older adults and may not generalize to older populations with cognitive impairments or severe sleep disorders. Furthermore, the use of EEG headbands at home relies on participant compliance and technical reliability, which may affect data completeness.

Future Research

This trial will be a first step toward clarifying the relationship between passive and active ETU and sleep in older adults. Findings should be interpreted with consideration for the short observation period and the emphasis on users experienced with digital technology. Future research could extend the duration

XSL•FO
RenderX

of the intervention, include follow-ups, vary the "dose" and timing (eg, 15, 30, or 60 minutes, or the last hour before bed), and test device features such as blue light filters, night mode, and different types or paces of content. Implementation studies in clinics and community programs can assess the practicality, adherence, and effects of brief guidance on evening technology habits.

Dissemination Plan

The findings from this trial will be disseminated through a peer-reviewed journal article, presentations at academic conferences, and promoted using social media.

Conclusions

This protocol describes a crossover trial to evaluate the relationship between objective and subjective sleep and various types of evening activities, such as active technology use, passive technology use, and nondigital activities. If the hypothesized patterns are observed, the results could provide practical guidance on the use of presleep technology in late life. The findings may be useful in improving digital wellness programs for older adults, providing guidance on the timing, duration, and type of evening use.

Acknowledgments

The authors thank the Swedish National Graduate School on Aging and Health (SWEAH) for supporting the learning process.

Funding

This research was funded by the Blekinge Institute of Technology.

Authors' Contributions

SNG conceptualized the study and developed the research design in collaboration with AB, JN, and PA. SNG led the development of the study protocol, drafted the original manuscript, and designed the data collection instruments. AB, JN, PA, and JSB contributed to the methodology and protocol refinement. All authors reviewed and contributed to revisions of the manuscript and approved the final version for submission. SNG will coordinate the study and oversee data collection. AB, JN, PA, and JSB provided supervision and guidance throughout the development of the protocol and will contribute to the interpretation of results and dissemination.

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision:
- Proofreading and editing
The GAI tool used was: Grammarly 1.146.3.0. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by: Sarah Nauman Ghazi

Conflicts of Interest

None declared.

Multimedia Appendix 1
Power analysis.
[DOCX File, 34 KB - resprot_v15i1e84512_app1.docx]

Multimedia Appendix 2
Baseline questionnaire–English.
[DOCX File, 46 KB - resprot_v15i1e84512_app2.docx]

Multimedia Appendix 3
Baseline questionnaire–Swedish.
[DOCX File, 1759 KB - resprot_v15i1e84512_app3.docx]

References

1. Internetanvändning [Internet use]. Svenskarna och Internet [Swedes and the Internet]. 2024. URL: https://svenskarnaochinternet.se/rapporter/svenskarna-och-internet-2024/internetanvandning/ [accessed 2025-01-15]
2. Ancoli-Israel S. Sleep and its disorders in aging populations. Sleep Med 2009 Sep;10 Suppl 1:S7-11. [doi: 10.1016/j.sleep.2009.07.004] [Medline: 19647483]

XSL•FO
RenderX

3. Bauducco S, Pillion M, Bartel K, Reynolds C, Kahn M, Gradisar M. A bidirectional model of sleep and technology use: a theoretical review of how much, for whom, and which mechanisms. Sleep Med Rev 2024 Aug;76:101933. [doi: 10.1016/j.smrv.2024.101933] [Medline: 38657359]
4. Heo JY, Kim K, Fava M, et al. Effects of smartphone use with and without blue light at night in healthy adults: a randomized, double-blind, cross-over, placebo-controlled comparison. J Psychiatr Res 2017 Apr;87:61-70. [doi: 10.1016/j.jpsychires.2016.12.010] [Medline: 28017916]
5. Alkozei A, Dailey NS, Bajaj S, et al. 0070 the effects of acute blue wavelength light exposure on functional brain connectivity and mood. Sleep 2020 May 27;43(Supplement_1):A28-A29. [doi: 10.1093/sleep/zsaa056.068]
6. Han X, Zhou E, Liu D. Electronic media use and sleep quality: updated systematic review and meta-analysis. J Med Internet Res 2024 Apr 23;26:e48356. [doi: 10.2196/48356] [Medline: 38533835]
7. Weaver E, Gradisar M, Dohnt H, Lovato N, Douglas P. The effect of presleep video-game playing on adolescent sleep. J Clin Sleep Med 2010 Apr 15;6(2):184-189. [doi: 10.5664/jcsm.27769] [Medline: 20411697]
8. De Rosa O, Baker FC, Barresi G, Conte F, Ficca G, de Zambotti M. Video gaming and sleep in adults: a systematic review. Sleep Med 2024 Dec;124:91-105. [doi: 10.1016/j.sleep.2024.09.015] [Medline: 39284267]
9. Potapova R, Potapov V, Lebedeva N, Karimova E, Bobrov N. Media content vs nature stimuli influence on human brain activity. In: Speech and Computer SPECOM 2021 Lecture Notes in Computer Science, Vol 12997: Springer, Cham; 2021, Vol. 12997:528-539. [doi: 10.1007/978-3-030-87802-3_48]
10. Grassini S, Segurini GV, Koivisto M. Watching nature videos promotes physiological restoration: evidence from the modulation of alpha waves in electroencephalography. Front Psychol 2022;13:871143. [doi: 10.3389/fpsyg.2022.871143] [Medline: 35747675]
11. Cerasuolo M, Conte F, Cellini N, et al. The effect of complex cognitive training on subsequent night sleep. J Sleep Res 2020 Dec;29(6):e12929. [doi: 10.1111/jsr.12929] [Medline: 31651070]
12. Kuoppamäki SM, Taipale S, Wilska TA. The use of mobile technology for online shopping and entertainment among older adults in Finland. Telematics and Informatics 2017 Jul;34(4):110-117. [doi: 10.1016/j.tele.2017.01.005]
13. Kueider AM, Parisi JM, Gross AL, Rebok GW. Computerized cognitive training with older adults: a systematic review. In: Brucki S, editor. PLoS One 2012;7(7):e40588. [doi: 10.1371/journal.pone.0040588] [Medline: 22792378]
14. Sen K, Prybutok G, Prybutok V. The use of digital technology for social wellbeing reduces social isolation in older adults: a systematic review. SSM Popul Health 2022 Mar;17:101020. [doi: 10.1016/j.ssmph.2021.101020] [Medline: 35024424]
15. Ghazi SN, Behrens A, Berglund JS, Berner J, Anderberg P. Examining sleep health and its associations with technology use among older adults in Sweden: insights from a population-based study. BMC Public Health 2025 Aug 22;25(1):2896. [doi: 10.1186/s12889-025-23894-8] [Medline: 40847298]
16. De Rosa O, Conte F, D'Onofrio P, et al. Habitual videogame playing does not compromise subjective sleep quality and is associated with improved daytime functioning. Brain Sci 2023 Feb 7;13(2):279. [doi: 10.3390/brainsci13020279] [Medline: 36831822]
17. Pierson-Bartel R, Ujma PP. Objective sleep quality predicts subjective sleep ratings. Sci Rep 2024 Mar 11;14(1):5943. [doi: 10.1038/s41598-024-56668-0] [Medline: 38467694]
18. Landry GJ, Best JR, Liu-Ambrose T. Measuring sleep quality in older adults: a comparison using subjective and objective methods. Front Aging Neurosci 2015;7(SEP):166. [doi: 10.3389/fnagi.2015.00166] [Medline: 26441633]
19. Radtke T, Apel T, Schenkel K, Keller J, von Lindern E. Digital detox: an effective solution in the smartphone era? a systematic literature review. Mobile Media & Communication 2022 May;10(2):190-215. [doi: 10.1177/20501579211028647]
20. Lagergren M, Fratiglioni L, Hallberg IR, et al. A longitudinal study integrating population, care and social services data. The Swedish National study on Aging and Care (SNAC). Aging Clin Exp Res 2004 Apr;16(2):158-168. [doi: 10.1007/BF03324546] [Medline: 15195992]
21. Leger KA, Kim YK, Zhang S, Gao S, Fingerman KL. Daily electronic media use and sleep in late life. In: Sneed RS, editor. J Gerontol B Psychol Sci Soc Sci 2024 Aug 1;79(8):gbae095. [doi: 10.1093/geronb/gbae095] [Medline: 38845388]
22. Adan A, Almirall H. Horne & Östberg morningness-eveningness questionnaire: a reduced scale. Pers Individ Dif 1991;12(3):241-253. [doi: 10.1016/0191-8869(91)90110-W]
23. Danielsson K, Sakarya A, Jansson-Fröjmark M. The reduced morningness-eveningness questionnaire: psychometric properties and related factors in a young Swedish population. Chronobiol Int 2019 Apr;36(4):530-540. [doi: 10.1080/07420528.2018.1564322] [Medline: 30614272]
24. Snyder E, Cai B, DeMuro C, Morrison MF, Ball W. A new single-item sleep quality scale: results of psychometric evaluation in patients with chronic primary insomnia and depression. J Clin Sleep Med 2018 Nov 15;14(11):1849-1857. [doi: 10.5664/jcsm.7478] [Medline: 30373688]
25. Stern RA, Arruda JE, Hooper CR, Wolfner GD, Morey CE. Visual analogue mood scales to measure internal mood state in neurologically impaired patients: description and initial validity evidence. Aphasiology 1997 Jan;11(1):59-71. [doi: 10.1080/02687039708248455]
26. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037/0022-3514.54.6.1063] [Medline: 3397865]

XSL•FO
RenderX

27. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. Int J Hum Comput Stud 2018 Apr;112:28-39. [doi: 10.1016/j.ijhcs.2018.01.004]
28. Jaoude MA, Ravi A, Niu J, Banville H, Torres NF, Aimone C. Automated sleep staging on wearable EEG enables sleep analysis at scale. : IEEE Presented at: 2023 11th International IEEE/EMBS Conference on Neural Engineering (NER); Apr 24-27, 2023; Baltimore, MD, USA p. 1-4. [doi: 10.1109/NER52421.2023.10123829]
29. Lanthier M, Foti MC, Higginson C, et al. 1119 Validation of a portable sleep electroencephalography device in good sleepers and people with sleep apnea. Sleep 2024 Apr 20;47(Supplement_1):A480. [doi: 10.1093/sleep/zsae067.01119]
30. Bozkurt A, Demirdöğen EY, Akıncı MA. The association between bedtime procrastination, sleep quality, and problematic smartphone use in adolescents: a mediation analysis. Eurasian J Med 2024 Feb;56(1):69-75. [doi: 10.5152/eurasianjmed.2024.23379] [Medline: 39128061]
31. Bruni O, Sette S, Fontanesi L, Baiocco R, Laghi F, Baumgartner E. Technology use and sleep quality in preadolescence and adolescence. J Clin Sleep Med 2015 Dec 15;11(12):1433-1441. [doi: 10.5664/jcsm.5282] [Medline: 26235161]
32. Brautsch LAS, Lund L, Andersen MM, Jennum PJ, Folker AP, Andersen S. Digital media use and sleep in late adolescence and young adulthood: a systematic review. Sleep Med Rev 2023 Apr;68:101742. [doi: 10.1016/j.smrv.2022.101742] [Medline: 36638702]
33. Killgore WD, Dailey NS, Raikes AC, et al. 0079 blue light exposure enhances neural efficiency of the task positive network during a cognitive interference task. Sleep 2020 May 27;43(Supplement_1):A32. [doi: 10.1093/sleep/zsaa056.077]
34. Cajochen C, Münch M, Knoblauch V, Blatter K, Wirz-Justice A. Age-related changes in the circadian and homeostatic regulation of human sleep. Chronobiol Int 2006;23(1-2):461-474. [doi: 10.1080/07420520500545813] [Medline: 16687319]
35. Najjar RP, Prayag AS, Gronfier C. Melatonin suppression by light involves different retinal photoreceptors in young and older adults. J Pineal Res 2024 Jan;76(1):e12930. [doi: 10.1111/jpi.12930] [Medline: 38241677]
36. Shaleha R, Roque N. From screens to cognition: a scoping review of the impact of screen time on cognitive function in midlife and older adults. Digit Health 2025;11:20552076251343989. [doi: 10.1177/20552076251343989] [Medline: 40656848]
37. Fuller C, Lehman E, Hicks S, Novick MB. Bedtime use of technology and associated sleep problems in children. Glob Pediatr Health 2017;4:2333794X17736972. [doi: 10.1177/2333794X17736972] [Medline: 29119131]
38. Laycock PJ, Jones B, Kenward MG. Design and analysis of cross-over trials. Biometrics 1991 Jun;47(2):787. [doi: 10.2307/2532176]
39. Reda DJ. Cross-over trials. In: Clinical Trials Design in Operative and Non Operative Invasive Procedures: Springer, Cham; 2017:79-86. [doi: 10.1007/978-3-319-53877-8_9]
40. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11. [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

Abbreviations

EEG: electroencephalography
ETU: evening technology use
LMM: linear mixed-effects model
RCT: randomized controlled trial
REM: rapid eye movement
SE: sleep efficiency
SNAC-B: Swedish National Study on Aging and Care – Blekinge
SOL: sleep onset latency
TST: total sleep time
WASO: wake after sleep onset

Edited by J Sarvestan, S Basha; submitted 20.Sep.2025; peer-reviewed by A Hu, R Shaleha; accepted 24.Dec.2025; published 30.Jan.2026.

Please cite as:
Ghazi SN, Behrens A, Niklasson J, Sanmartin Berglund J, Anderberg P
The Effect of Evening Technology Use on Objective Sleep in Older Adults: Protocol for a Crossover Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84512
URL: https://www.researchprotocols.org/2026/1/e84512
doi:10.2196/84512

XSL•FO
RenderX

© Sarah Nauman Ghazi, Anders Behrens, Joakim Niklasson, Johan Sanmartin Berglund, Peter Anderberg. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Youth-Centered Mobile Intervention (Next4You) to Promote Healthy Relationships and Sexual Wellness Among Adolescents in or Transitioning From Foster Care: Protocol for a Randomized Controlled Trial

Pamela M Anderson[1*], PhD; Seow Ling Ong[1*], MSW; Dallas Elgin[2*], PhD; Jennifer Laird[2*], PhD; Darriya Starr[2*], MS; Sade Daniels[3*], MSW; Jakara Rogers[1*], AA; Karin K Coyle[1*], PhD

[1]Education Training and Research, 500 Westridge Drive, Suite 102, Watsonville, CA, United States

[2]RTI International, Research Triangle Park, NC, United States

[3]Dae Dreams Affirmed, LLC, Douglasville, GA, United States

[*]all authors contributed equally

Corresponding Author:
Pamela M Anderson, PhD
Education Training and Research, 500 Westridge Drive, Suite 102, Watsonville, CA, United States

Abstract

Background: National birth rates among adolescents have consistently decreased since 1991, yet substantial disparities remain, particularly among youth in foster care, who experience higher risks of unintended pregnancies and sexually transmitted infections. Few sexual health programs rarely address the specific needs of foster youth or incorporate youth perspectives into their design, development, and implementation.

Objective: This study aims to refine, pilot-test, and evaluate Next4You, a fully mobile, youth-centered sexual health education platform tailored specifically for adolescents in or transitioning from foster care. The program intends to reduce sexual risk behaviors and related psychosocial outcomes by promoting healthy relationships, sexual wellness, contraception use, communication skills, self-respect, and wellness education.

Methods: Next4You is a fully mobile, self-paced, 4-week intervention evaluated using an individual-level randomized controlled trial involving a target sample of 500 youths aged 16 - 19 years with current or previous foster care experience in California. Participants are randomly assigned to either the intervention group, which receives access to Next4You modules, or the comparison group, which accesses general health materials through a similar web platform. Both groups complete baseline, 3-month, and 9-month follow-up surveys. Immediately following the 9-month follow-up survey, participants in each group receive access to the platform they were not initially assigned. Data collection assesses contraceptive behaviors, sexual communication, consent self-efficacy, and knowledge related to health rights and financial literacy. Intervention engagement is tracked through platform analytics, and qualitative interviews supplement data collection. Data analysis will adhere to intent-to-treat principles, using multilevel regression models to assess impacts.

Results: Funding was awarded in 2021, with institutional review board approval in October 2022. Intervention development lasted approximately 7 months, engaging foster youth in a co-design process to ensure relevance and cultural competence. Study recruitment began in September 2023, continuing until May 2025. Final data collection is anticipated by March 2026, followed by data analysis.

Conclusions: Next4You presents an innovative approach to sexual health education by addressing critical gaps in sexual health education for foster youth through mobile technology. If effective, Next4You could provide another evidence-based option for promoting sexual health among foster youth, guiding policymakers and practitioners in adopting similar trauma-informed and youth-centered interventions.

Trial Registration: ClinicalTrials.gov NCT06127277; https://clinicaltrials.gov/study/NCT06127277

International Registered Report Identifier (IRRID): DERR1-10.2196/77185

(JMIR Res Protoc 2026;15:e77185) doi:10.2196/77185

KEYWORDS

adolescents; sexual health; technology; randomized controlled trial; mobile app

XSL•FO
RenderX

Introduction

Background and Rationale

National birth rates for adolescent women have continuously declined since 1991, falling to 5.6 per 1000 for 15‐ to 17-year-olds (approximately an 85% decrease) and to 25.8 per 1000 for 18‐ to 19-year-olds (approximately a 73% decrease) by 2022 [1]. This improvement has been attributed to a variety of factors, including increased use of contraception facilitated by improved access to clinical services, publicly funded adolescent pregnancy prevention programs, and comprehensive sexual health education [2,3]. Despite this progress, substantial disparities in birth rates persist by race, ethnicity, socioeconomic status, geographic location, and some populations [1]. In particular, foster youth are at a higher risk for unplanned pregnancy and sexual risk behaviors that lead to sexually transmitted infections (STIs). One in three young women in foster care becomes pregnant at least once by age 17‐18 years, with more than half experiencing a pregnancy by age 19 years, twice the rate of their peers not in foster care [4-6]. Young men in foster care are also at high risk for getting someone pregnant by age 19 years [4]. Youth in foster care are more likely than their peers to have had sexual intercourse, to have engaged in transactional sex, and to not have used birth control at last sex [6]. While 70% of pregnant young women in foster care did not plan their pregnancy, only a quarter of them were using birth control when they became pregnant [6], demonstrating a clear gap in prevention services reaching vulnerable youth [7-9]. Youth in foster care may be using birth control inconsistently or incorrectly because they experience instability in access to sexual health information and services, such as changes in caseworkers and living arrangements, and a lack of coordination among agencies providing care [4,5,10]. Additionally, many vulnerable youth have concerns about provider attitudes, privacy, and stigma when they access services [8]. LGBTQ youth, who are disproportionately represented in foster care and have an increased risk for pregnancy and STIs, face additional challenges in accessing health services and education that are inclusive and represent LGBTQ youth's experiences, relationships, and identities [8,11,12].

For youth in foster care, there are few sex education programs designed specifically for them, and none that are evidence-based and available through the teen pregnancy prevention evidence review use a fully technology-based approach [13]. *Power Through Choices* was the first sex education program designed exclusively for foster youth to be evaluated by a randomized controlled trial [14]. This 10-session program used interactive methods to support youth in developing critical knowledge and sexual decision-making skills. Youth participating in the program were less likely to engage in sexual intercourse without birth control in the past 3 months at 6-month follow-up and had significantly lower odds of ever being or getting someone pregnant at 12-month follow-up than control youth. This program demonstrates that sexual health education can be implemented effectively with foster youth in nonclassroom settings. However, similar to other traditional sexual health education programs, there are limitations to a lengthy, multisession program among a population where high instability is associated with increased risk of pregnancy [5]. And when vulnerable youth cannot access information from trusted sources such as educators and providers, they may turn to the internet [12]. In a nationally representative survey, more than 80% of adolescents reported that they have sought health advice on the internet, with 34% modifying their behavior based on what they found [15]. Mobile technology is almost universally accessible—95% of adolescents have access to a smartphone regardless of gender, race and ethnicity, and socioeconomic status, and the majority (89%) are accessing the internet at least several times a day [16]. Yet, to our knowledge, there has been only one fully mobile sexual health intervention for young people that has been rigorously evaluated, and none that have been designed and evaluated specifically for youth in foster care [13,17].

Engaging youth in the design of interventions that impact their health and well-being is believed to support young people's development [18], strengthen programs and services [19], and lead to social, political, and economic benefits [20]. Engagement in these processes is also what young people want [19]. Yet, sexual health interventions are rarely designed in collaboration with young people, especially those from systemically impacted communities such as youth in foster care [21]. Additionally, many existing sexual health interventions that meaningfully involve youth lack rigorous evidence on their ability to engage youth and impact health outcomes [21].

Young people, including those in foster care, routinely identify relationships and relationship health as important areas of focus [22]. Romantic relationships play a significant role in adolescent development [23]. By middle adolescence, most young people have been involved in at least one romantic relationship [24], a context in which most sexual interactions occur [25]. Adolescence is a period where young people begin to develop standards and expectations for relationships and build their communication skills. For youth in foster care, these tasks may be more difficult if they have experienced contexts where neglect, aggression, and violence are modeled [26]. Thus, youth in foster care are at increased risk for experiencing intimate partner violence and, because of prior trauma, may be more vulnerable to negative health impacts from intimate partner violence [27]. Yet, developmental psychologists emphasize that relationships developed during adolescence represent a new period where past models of relationships may be reshaped [23], an incredible potential to change unhealthy models and carry positive ones into adulthood, underscoring the potential impact of this content on adolescents' well-being.

Study Objectives and Research Questions

The primary study objectives are as follows: (1) refine and update the Next4You mobile responsive platform using trauma-informed and healing-centered principles, centering the voices of youth in foster care and engaging them directly in development; (2) pilot test the updated intervention and planned evaluation activities with youth representing the priority population and finalize based on pilot data and insights from youth and program partners; (3) assess the effectiveness of the new intervention using a randomized controlled trial; (4) develop an intervention manual and implementation package (eg, logic

XSL•FO
RenderX

model, core components, adaptation guidelines, and fidelity tool) to support replications; and (5) collect other implementation data and document lessons learned throughout the study.

Textbox 1 summarizes the primary and secondary research questions for the study.

Textbox 1. Primary and Secondary Research Questions

Primary research question

Three months following the intervention:

1. What is the effect of the 4-week Next4You intervention on the prevalence of choosing abstinence or using condoms among participants who had vaginal or anal sex in the past 3 months?

Secondary research questions

Nine months following the intervention:

1. What is the effect of the 4-week Next4You intervention on the prevalence of choosing abstinence or using condoms among youth in foster care who had vaginal or anal sex in the past 3 months?

Three and nine months following the intervention period:

2. What is the effect of the 4-week Next4You intervention on rates of condomless vaginal or anal sex?

3. What is the effect of the 4-week Next4You intervention on sexual consent self-efficacy?

4. What is the effect of the 4-week Next4You intervention on sexual communication self-efficacy?

5. What is the effect of the 4-week Next4You intervention on condom use attitudes?

6. What is the effect of the 4-week Next4You intervention on contraceptive use attitudes and beliefs?

7. What is the effect of the 4-week Next4You intervention on experiences with equity and power balance in relationships?

8. What is the effect of the 4-week Next4You intervention on knowledge of health care rights and location of services?

9. What is the effect of the 4-week Next4You intervention on contraceptive use (specifically experience with condom errors)?

10. What is the effect of the 4-week Next4You intervention on knowledge of educational rights and programs?

11. What is the effect of the 4-week Next4You intervention on financial literacy knowledge?

Methods

Study Design Overview

Next4You will be tested using an individual-level (randomized controlled trial) involving up to 500 young people ages 16 - 19 years with former or current experience in the California foster care system. Half will be assigned to an intervention group and the other half to a comparison group. The study survey will be collected at three time points: baseline (ie, after assent and consent and before randomization) and 3 and 9 months after a 4-week intervention period (applied to both the intervention and comparison groups). The evaluation of Next4You was designed to meet the Institute of Education Sciences' *What Works Clearinghouse* (WWC) design standards without reservations [28]. Figure 1 presents the CONSORT (Consolidated Standards of Reporting Trials) flow diagram illustrating participant enrollment, allocation, follow-up, and analysis for the study.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram for the Next4You study.

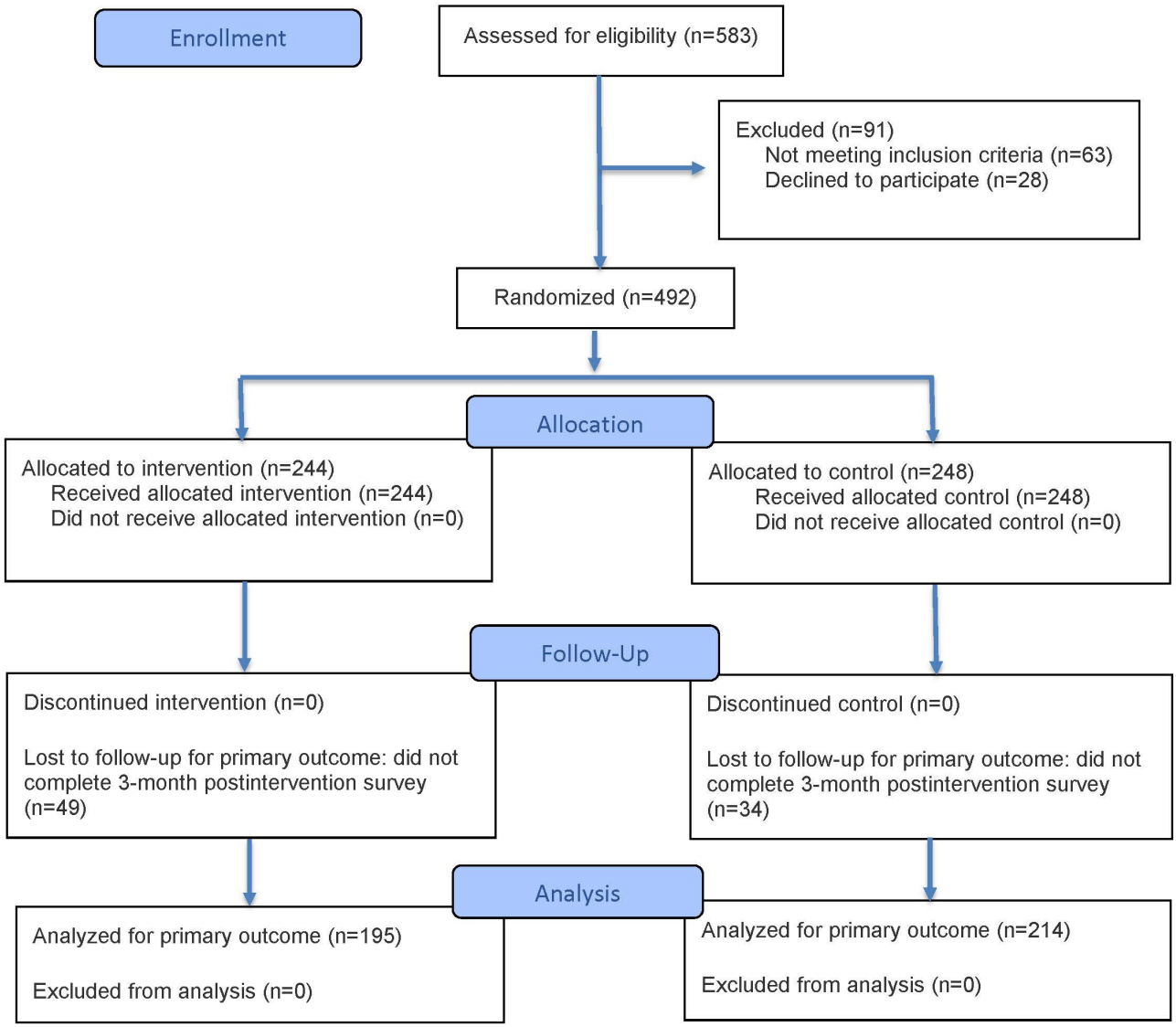

Note: This diagram reflects the primary endpoint of the study: 3-months postintervention. The secondary endpoint, 9-months postintervention, will be reflected in CONSORT diagrams included in subsequent papers.

Study Setting

The study will be completed digitally, among participants who have been referred to the study through partner organizations in 24 counties throughout California.

Eligibility Criteria

Youth must meet the following criteria to be eligible for the study: (1) age 16 to 19 years at time of study enrollment; (2) currently in foster care or formerly in foster care (and were in foster care on or after their 13th birthday); (3) currently residing in California; (4) proficient in English; and (5) have access to a tablet, laptop, desktop computer, or phone that is Wi-Fi- or cellular-enabled. They also had to be free of severe mental health issues and have the cognitive ability to meaningfully engage with the web platform as determined by the recruitment ambassadors.

Intervention and Comparison Protocols

Intervention Condition

Participants randomized to the active intervention condition will receive access to Next4You, a fully mobile intervention designed to reduce rates of unintended pregnancy and STIs among young people currently or recently in the California foster care system. The platform was designed in partnership with foster care youth and with input from community-based organizations serving this population. The responsive, web-based platform can be accessed by young people at any time from an internet-enabled mobile device.

XSL•FO
RenderX

The intervention features 6 content modules (Table 1): Sexual Wellness, Contraception, Relationships, Respecting Self and Partners, Communication, and Building Wealth and Wellness (eg, financial literacy, educational rights, and mental health), with a total instructional time of approximately 3 to 4 hours. Each module includes 5 - 12 microlessons, designed as brief, informal snippets of content that are 5 minutes or less; focused on a single learning concept; and are delivered through diverse approaches such as games, videos, quizzes, and others [29]. One advantage of this method includes the repeatability of the content in time, place, and format preferred by most young people [30]. In addition to microlearning, the site includes gamification features, a youth-designed retention strategy to increase motivation to engage youth in the content.

Table . Intervention modules.

Module	Microlessons, n	Brief description of module content
Sexual Wellness	9	Defining sexual wellness; sexual identity and orientation; sexually transmitted infections and ways to reduce risk; using and accessing health care to maintain sexual wellness; sexual boundaries; sexual and reproductive health rights; sexual anatomy and tips for increasing pleasure.
Relationships	10	Elements of a safe and secure relationship; factors that suggest it may be time to end a relationship; tips on navigating a breakup; navigating loss and feelings of rejection after a relationship ends; intimate partner violence; love bombing; strengthening and building connection in relationships; exploring factors that make young people with lived experience in foster care more susceptible to trafficking; exploring power and control behaviors that may show up in relationships.
Communication	7	"I" versus "you" statements; preparing for having difficult conversations; strategies for managing emotions that may surface during difficult conversations; strategies for de-escalating conversations; strategies for active listening to improve communication skills.
Respecting Self and Partners	7	Self-love and self-worth; affirmative consent; what respecting a partner includes; personal boundaries; strategies for navigating different personal boundaries; gender and gender identity.
Contraception	5	Condom use and steps; common methods of birth control; withdrawal from a harm reduction perspective; contraception rights; options if methods of contraception fail.
Building Wealth and Wellness	12	Passion and purpose; educational rights; college transition and support services; financial rights; coping skills; financial aid options; trauma—building resilience and navigating past trauma; mental health; budgeting; body image.

The Next4You platform includes a total of 50 microlessons, and participants earn points for each completed lesson. To motivate early engagement with the website, participants receive a US $25 e-gift card upon completing 5 lessons. Subsequent points are banked, and the remaining incentive amount is distributed at the end of the 4 weeks, based on the total amount of points earned. Participants may earn up to US $100 in e-gift cards for completing all 50 microlessons. In addition to lesson completion, participants can earn up to US $25 in bonus incentives through all 4 achievement badges: Start Up, On Fire, My Faves, and Consistency. In total, participants may earn a maximum of US $125 in e-gift cards over the 4-week program period.

The online platform will collect paradata in the form of timestamps and counts of user sessions for all actions performed by each user, thereby allowing for calculation of the time spent engaging with the intervention (including total time spent on each microlesson, modules, and engaging with various activities and resources), frequency of logins, and the amount of completed content. Users will be automatically logged out after approximately 10 minutes of inactivity to protect privacy and reduce the likelihood of inflated duration data. The evaluation will use the duration data to conduct exploratory analyses examining differing levels of exposure and the relation to study outcomes.

Intervention Conceptual and Development Frameworks

The development of the Next4You intervention draws upon other pertinent knowledge bases to center youth voice through a youth participatory process and draws on the latest research

XSL•FO
RenderX

in positive youth development and microlearning to strengthen the intervention's content and strategies (see the screenshots in Multimedia Appendix 1). Positive youth development is a strength-based approach used to promote adolescents' prosocial competencies and build skills related to their positive health and well-being [31]. The adolescent development literature guided Next4You's relationship development content and helped ensure it is age appropriate [23]. The intervention also draws on dual process theories to address the social-emotional and cognitive influences on sexual decision-making [32]. Finally, the intervention draws from social cognitive theory both in terms of key constructs in skill acquisition, such as building self-efficacy or confidence in one's ability to perform a skill, and in shaping the process of learning, such as through using observational learning or modeling during instruction [33].

Intervention Co-Design Process

Engaging youth with lived experience in the foster care system as subject matter experts was an essential first step in developing the Next4You platform and its content. This approach provided a unique opportunity to deliver accurate, inclusive, culturally competent, and on-demand information, resources, and tools tailored to the needs of transition-age youth. To refine and enhance the platform before piloting, the team used youth participatory methods to engage these subject matter experts. These methods incorporate positive youth development principles, emphasizing the importance of engaging young people as partners by respecting their knowledge, insights, expertise, and leadership [34]. Recognizing that youth in foster care often experience trauma, we used trauma-informed and healing-centered principles when engaging co-designers, including safety, equity, empowerment, and resilience. This ensured that the co-design process was inclusive, culturally relevant, and supportive, even when difficult topics arose.

The co-design process involved 9 participants, ages 19 - 22 years, from California, who engaged individually and as part of a group over 7 weeks through weekly 90-minute meetings. The facilitation approach was intentionally structured yet flexible. The sequence included 2 sessions on discovery (eg, community building), 2 on defining (research and product design and function), and 3 on designing content for intervention modules, concluding with a design and testing session. Slide decks structured each meeting, but sessions were highly interactive, incorporating icebreakers, small-group discussions, and digital tools (eg, Mentimeter or Kahoot). Participants were invited to be active cocreators rather than passive reviewers; facilitators used semistructured prompts such as "What keeps users coming back?" or "What does wellness mean to you?" to elicit ideas that shaped both content and user experience. Through this process, participants influenced the platform's esthetics, navigation, and functionality, including customizable color schemes, authentic imagery, gamified features, and microlesson delivery. They identified priority content areas and co-developed stories, scenarios, and audio narration reflecting real experiences. The team used mockups and mood boards to align design decisions with participant input and maintained engagement between sessions through asynchronous channels (eg, GroupMe). Following co-design, pilot testing of 18 lessons yielded highly positive feedback. Participants recommended adding more interactive activities, videos, and visual animations, as well as expanding life-skills content such as improving communication with adults in positions of authority including teachers and supervisors. These refinements strengthened both usability and relevance for youth transitioning from foster care.

By centering youth as co-designers, the Next4You project produced a platform that authentically reflects their experiences and priorities, demonstrating the power of participatory design to advance youth-driven innovation in adolescent health.

Comparison Intervention

Youth randomized to the comparison condition will have access to a mobile-responsive website featuring digital resources focused on nutrition, sleep, stress and anxiety, and exercise. Consistent with the intervention group, participants in the comparison group will have unlimited access to the general health digital resources for the initial 4-week period. Throughout the evaluation, youth assigned to the control condition will also continue to receive the standard resources and services typically available to them through their involvement in the foster care system. The comparison condition was selected to establish an effective treatment contrast [35] by isolating the difference between the treatment received by youth with access to the Next4You platform and the treatment received by youth who instead received access to digital resources focusing on nutrition, sleep, stress and anxiety, and exercise. After the initial 4-week intervention phase, comparison group participants will retain access to the comparison website. Upon completion of the 9-month evaluation period, these participants will additionally receive access to the Next4You intervention website for the remainder of the study, concluding on September 29, 2026.

Sample Size and Power Analyses

Our proposed sample of 500 young people (with a 1:1 allocation to each condition) would result in the Next4You evaluation being powered to identify effects in the range of small to "substantively important" (ie, 0.25 SDs, in accordance with the WWC standards) [36]. The power analysis estimates were obtained and validated via Optimal Design Plus Empirical Evidence software and the HHS Office of Adolescent Health's Teen Pregnancy Prevention MDI and MDES calculator. The estimates were developed based on the following parameters in accordance with the funder evaluation standards: simple randomization at the individual level; 25% attrition rate; type I error rate of 5%; type II error rate of 20% (ie, power of 0.80); 2-sided hypothesis test; 50% probability of assignment to the treatment group; outcome prevalence of 50%; 25% of the outcome variance is explained by covariates. The impact model will include the baseline measure of the outcome and age or grade level, sex, and race and ethnicity as covariates consistent with funder requirements. The study team will use a combination of data- and theory-driven approaches to select additional covariates for inclusion in the model to further improve the precision of the impact estimates.

Recruitment

Youth will be recruited from community and school programs serving young people in or transitioning from foster care. Recruitment and enrollment will occur on a rolling basis during

XSL•FO
RenderX

years 2 and 3 and the first quarter of year 4 across 24 California counties, including postsecondary on-campus support programs such as Guardian Scholars and NextUp. Each collaborating site designates a recruitment ambassador, typically an adult supporter who works directly with foster youth. The primary criterion for ambassadors is that they are staff members who meet individually and have direct contact with youth who meet study eligibility criteria. The study team trains all recruitment ambassadors and provides them with a standardized protocol that includes approved materials (recruitment flyers, brief videos explaining the study, a frequently asked questions document, and recruitment scripts). Upon completion of the training, ambassadors promote the study at their site, share the opportunity with youth who come in for direct services, and incorporate the study into the site's normal outreach mechanisms (eg, newsletters, direct mailing, and virtual and hard copy posters with a QR code). Recruitment ambassadors are trained to only present the study opportunity to youth who meet the study criteria based on their knowledge of the young people. Youth enroll in the study through an enrollment portal at partner sites (with ambassadors supporting youth through the process), which involves completing a screener and reviewing and signing the study consent form. Youth are also asked to input a recruitment code (which is unique to each recruitment ambassador), which allows the study team to track recruitment and confirm eligibility.

Consent and Assent Process

We received a waiver of parental permission from the ETR institutional review board for youth ages 16 - 17 years, who will complete an assent form written at an eighth-grade reading level, while youth ages 18 - 19 years will complete a consent form at the same readability level. Both forms are collected online to align with the fully mobile nature of the intervention. The assent and consent forms include 3 multiple-choice questions to assess young people's understanding of what they are being asked to do. If youth mark a question incorrectly, a clarifying message will be provided to ensure they have the information needed to make their decision. At the end of the consent form, youth will be asked to indicate their decision of whether they want to take part or not or if they would like to talk with a study team member to get more information before deciding. Youth who decline participation will be thanked for their consideration and will not be contacted again. Youth who say they would like to talk with a study team member will be contacted by a study team member within 2 business days to provide clarifying information. These youth would then be allowed to make their decision online.

Randomization

Randomization takes place after young people submit their baseline survey using the randomizing feature in Qualtrics with a 50 - 50 allocation fraction. More specifically, 50% of participants will be allocated to the intervention group and 50% of participants will be allocated to the comparison group. The study evaluator oversees and manages the randomization protocol. Once randomized, youth are automatically re-directed to the Next4You log in page, based on their assigned conditions along with steps for logging in and establishing their accounts.

Data Collection Methods

Baseline Survey

The evaluation of Next4You will rely on a confidential youth self-report survey assessing knowledge, attitudes, beliefs, and behaviors related to sexual health through web-based surveys administered via the Qualtrics platform. Young people receive a unique link via email to the survey immediately after affirming their assent to study participation. The study team will follow up with youth who do not complete the baseline survey by sending up to 3 texts and 2 additional emails over the course of approximately 3 weeks, each with the unique link to their baseline survey, and an offer to answer questions or support in completing the survey. The study team will then engage the recruitment ambassador for assistance in reaching out to the youth and encouraging them to complete the survey.

Participants will be randomized after the baseline survey and are automatically re-directed at the end of the survey to access their assigned intervention condition. The baseline survey is estimated to take approximately 25 minutes to complete. Participants receive a gift card worth US $40 for filling out the baseline survey.

Preparing for Follow-Up Surveys

One month prior to sending each follow-up survey, youths are asked to provide the study team with their latest contact information. Each youth is contacted simultaneously at their last known phone number and email address—one of which has usually not changed in the 3- and 6-month periods since their last follow-up survey. This outreach serves to both remind youth that a follow-up survey will be sent to them in the next few weeks and maintain the most recent, preferred contact information to reach the youth when their next survey is sent. Doing so 1 month in advance gives youth time to reply or receive up to 3 reminders if needed. This confirmation of contact information request is associated with a US $10 e-gift card when the youth replies to the study team with their current contact information.

Follow-Up Surveys (3 and 9 Months)

Both follow-up surveys are administered online using procedures like those used at baseline. The study team will use a multimethod outreach process to contact youth who do not respond to the surveys. In the first phase, the study team will send up to 3 simultaneous texts and emails to encourage nonrespondents to complete the survey. During the second phase, youth who have not responded will receive up to 2 calls (with voicemails) from the study team during the following week. If youth have not completed the survey during the second phase, the study team will wait 1 week before implementing the third phase of outreach, which involves engaging with the study recruitment ambassadors for assistance in reaching out to the youth and encouraging them to complete the survey. At all outreach attempts by the study team, participants are provided with the study team's email and phone for any questions or need for assistance in completing the survey.

XSL•FO
RenderX

Implementation Evaluation Data Collection

Given the intervention is fully mobile, our implementation evaluation data collection focuses on four areas related to fidelity that have been found to influence dose and treatment effects in mobile interventions [37]: (1) amount (count of lessons completed), (2) frequency (counts of user sessions and patterns of use), (3) duration (amount of time spent in each module and lesson), and (4) depth (total content and engagement as a percentage of all available content). These data will be collected from the web platform directly. We will also conduct selected postintervention interviews to gather qualitative data on user experiences during the study.

Retention

The study team will implement a comprehensive participant tracking and retention strategy designed to maximize response rates and minimize attrition across all follow-up periods. This approach includes allocating sufficient resources for incentives, dedicating staff time to personalize follow-up, and conducting multiple outreach attempts to nonrespondents. Consistent with intent-to-treat (ITT) principles, all youth enrolled in the evaluation will be tracked and will receive retention-focused outreach, regardless of level of engagement or participation.

Participants will receive e-gift card incentives for completing evaluation surveys, verifying or updating contact information, and completing intervention activities. In total, youth may earn up to US $315 for completing all activities. Incentive amounts include US $40 for the baseline survey, US $60 each for the 3- and 9-month follow-up surveys, US $10 per response (up to 2 times) for confirming or updating contact information, and up to US $125 for completing intervention activities. Incentive values were informed by a consultant with lived and professional experience in the foster care system to ensure that compensation reflects and honors participants' time and expertise.

Retention in the intervention group platform will include gamification features such as badges and a point-based reward system that allows for points to be converted into electronic gift cards. Additionally, push factor text and email notifications will be used to promote engagement. For example, if there has been no activity for at least 3 days, an automatic notification will go out to participants inviting them to log onto the platform and continue earning points through the completion of lessons. Youth assigned to the comparison group platform will receive equivalent total incentives for completing a minimum of 12 online pamphlets [38-49] on promoting healthy eating and nutrition, sleep, exercise, and reducing stress and anxiety. The total amount of incentives that youth may receive is the same for both study groups.

Outcomes

Primary Outcome

The primary outcome of this study is the impact of the 4-week Next4You intervention on contraceptive use among youth in foster care. Specifically, this study will examine the prevalence of participants who report either abstaining from sex or using c condoms during vaginal or anal sex in the past 3 months, compared with participants provided access to a mobile-responsive website containing only general wellness resources (sleep, nutrition, stress and anxiety, and exercise). This outcome will be assessed using data from the 3-month postintervention follow-up survey.

Secondary Outcomes

The secondary outcomes will examine the impact of the 4-week Next4You intervention, compared with a mobile-responsive website on wellness topics, on a range of sexual and reproductive health and well-being outcomes among youth in foster care. Specifically, analyses will assess effects on condom use behaviors, condomless sex, and condom errors; self-efficacy related to sexual consent and communication; attitudes toward condom and contraceptive use; power balance and experiences within sexual relationships; and knowledge of health care rights, educational rights, service locations, and financial literacy. Outcomes will be assessed at 3- and 9-month postintervention follow-ups.

Multimedia Appendix 2 [50-57] highlights our primary and secondary measures, with sample survey items.

Statistical Methods

Overview of Analytical Strategy

Our analysis plan follows a series of analytic steps, including data cleaning, analysis of baseline equivalence and attrition, handling of missing data, descriptive analysis, impact analyses of confirmatory research questions, and exploratory analyses. Throughout the analysis phase, we will use a significance level of 0.05 and employ an ITT framework, whereby young people who were randomly assigned to a condition are considered to be in that condition regardless of engagement level or intervention dosage. All analyses will be carried out using Stata. The results of the power analyses suggest a minimum detectable impact between 0.11 and 0.13, or 11 to 13 percentage points, for the dichotomous outcomes. These figures correspond to a minimum detectable effect size of 0.23 to 0.25 SD, or a reduction in continuous outcome from 50% to 37.5%, assuming a baseline prevalence of 50%.

Baseline Equivalence

Prior to estimating any outcome models, the study team will follow WWC group design standards related to baseline equivalence, attrition, and missing data [58]. The team will first test for differences between treatment conditions on baseline measures of all outcomes and demographic factors. Following WWC standards, effect sizes will be computed using Hedge *g* and any differences above 0.05 noted. Because of random assignment, group differences on baseline or demographic factors are unlikely, but baseline measures for all outcomes will be included as a covariate in all planned impact analyses, and any significant differences in demographic factors (ie, effect sizes between 0.05 and 0.25 SD) will be added as additional covariates.

Attrition

The team will then calculate overall and differential attrition rates. Overall youth attrition rates will consist of the number of units without observed outcome data in the analysis against the number of units randomized. Differential attrition will be

XSL•FO
RenderX

obtained by contrasting the attrition rates between treatment and control groups. Overall youth attrition and differential attrition between the treatment and control groups will be carefully monitored to determine the levels of potential bias.

Handling Missing Data

Missing data for all youth who agree to participate in data collection will be addressed with regression imputation using Stata, with special attention paid to WWC recommendations (eg, separately imputing data from treatment and control conditions). Regression imputation is the preferred approach for this study as it can be used to include participants with missing baseline data and missing outcome data (whereas nonresponse weighting, another imputation method, can only be used for the former). The team will use imputation regression models that (1) are conducted separately by condition or include an indicator variable for condition, (2) include all covariates used for adjustment in the impact model, and (3) include the outcome when imputing missing baseline data.

Descriptive and Impact Analyses

The study team will use simple averages and frequency distributions to describe the prevalence of each outcome for youth overall and for the treatment group, by selected demographic characteristics and by usage estimates, to provide an initial understanding of changes in the primary and secondary outcomes before estimating impact estimates using multivariate models. The impact of Next4You will be estimated under the ITT framework by comparing outcomes for individuals assigned to Next4You and those assigned to the control condition. The sample will include all youth who participated in the baseline survey and completed the 3-month follow-up, 9-month follow-up, or both. The study team will model all outcomes with a linear-based 2-level multilevel regression model with youth (level 1) nested within level 2 (sites) to account for the nested structure of the data, adjust for correlated error terms, and avoid underestimating standard errors and increasing the probability of type I errors. The following 2-level random intercept multilevel model will estimate the impacts of Next4You on the primary outcome with separate models estimated for the 3- and 9-month follow-ups:

$$yij = \beta o + \beta 1 Treat1ij + \beta 2 Baseline2ij + \beta 3 X3ij \ldots + \beta pXpij + uj + eij$$

where yij refers to the expected probability of exhibiting the corresponding outcome for individual *i* in site *j*; β0 is the overall mean of the outcome across all groups; β1Treat1ij is an indicator equal to one for participants assigned to the treatment group and zero for those assigned to the control group; β2Baseline2ij is the baseline measure of the outcome; β3X3ij+…+βpXpij represents a vector of individual-level baseline covariates (such as sex, race and ethnicity, and age); μj is a site-level random effect; and εij is an individual-level error term. The estimation of the effects of the primary outcome and other dichotomous outcomes via multilevel linear probability models provides a convenient, alternative approach to logit models for interpreting dichotomous outcomes [59]. In particular, linear probability models provide model results on a probability scale that is easier to interpret, whereas logit-based versions of the models express the findings in terms of log odds that are more challenging to interpret.

The use of a 2-level random intercept multilevel model is contingent upon the results of a likelihood-ratio test comparing the fit of the 2-level model against a single-level linear probability model. The null hypothesis for the likelihood-ratio test is that there is no significant difference between the two models. A *P* value of <.05 for the likelihood ratio test would result in the rejection of the null hypothesis, as there was a statistically significant difference between the models with the 2-level model fitting significantly better than the single-level linear probability model. In the event that the null hypothesis could not be rejected, the study team will use a single-level model.

The team will adjust *P* values to account for multiple hypothesis testing within each primary and secondary outcome domain using the Benjamini-Hochberg procedure [60]. Finally, the study team will conduct sensitivity analyses to test the robustness of the findings to alternative assumptions or specification of the multilevel models. Sensitivity analyses may involve comparing findings between multilevel logistic regression and multilevel linear probability models, assessing the sensitivity of approaches for addressing missing data, and assessing the impact of sibling participation on outcomes.

Implementation Data Analyses

User analytics captured directly on the platform will be analyzed to monitor participant engagement with both the intervention and comparison group platforms, exploring user profiles and how engagement may relate to behavioral outcomes in our analyses. Specifically, the study team will use these data to conduct exploratory analyses of dosage and treatment effects as moderators to assess how, if at all, the level of exposure to the platform influences the strength of the observed intervention effect. These data will also be within sensitivity analyses where a "treatment on the treated" analysis framework is used to examine the robustness of the impact estimates under the ITT framework. Various definitions of intervention uptake will be examined, including amount, frequency, duration, and depth among participants who ultimately received the condition to which they were assigned (ie, "compliers"). The effect of treatment receipt for the subgroup of compliers, or the local average treatment effect, will be estimated using instrumental variable methods with a two-stage least squares estimator. In the two-stage least squares, the first stage will involve using the random assignment indicator to predict intervention participation, and the second stage will use the predicted values of intervention participation from the first stage intervention to estimate intervention effectiveness.

Ethical Considerations

This study was funded in 2021 through the Family and Youth Services Bureau as part of the Personal Responsibility Education Innovative Strategies program as an award to ETR. This project was approved by the ETR institutional review board in October 2022. The study is registered at ClinicalTrials.gov (NCT06127277).

All participants provided informed consent or assent prior to participation. Youth aged 16-17 years provided assent under a waiver of parental permission, and youth aged 18-19 years

XSL•FO
RenderX

provided informed consent. Consent and assent procedures were conducted electronically and designed to ensure participants' understanding of study procedures, risks, and benefits, consistent with the fully mobile nature of the intervention. Participants were compensated for their time through electronic gift cards for survey completion, confirmation of contact information, and engagement with the intervention activities. Incentive amounts were structured to reflect participants' time and effort; full details of compensation and retention strategies are described in the Retention section. All consent forms and survey data will be stored securely within the Qualtrics platform, which employs firewalls, encryption, and independent security audits to protect participant information. Data will be accessible only to authorized project staff using password-protected, two-factor-authenticated devices. Personally identifiable information, such as names or email addresses, will be stored separately from survey responses, and each participant will be assigned a unique study ID used in all analyses. No identifiable information will appear in reports or publications.

Results

The study was funded in 2021, with ETR institutional review board approval in October 2022. The Next4You intervention was built over a period of 6 - 8 months. Study recruitment launched in September 2023 and completed in May 2025, with final follow-up surveys expected to be completed by March 2026, at which time we will initiate final study analyses. The study ends in September 2026.

Discussion

Strengths and Limitations

Adolescents in foster care experience disproportionately higher rates of teen pregnancy and STIs compared to their peers [61]. Despite these disparities, there are few evidence-based sexual health programs tailored for youth in foster care [62], a critical gap given their sexual health needs and the disparities they are navigating [61]. Further, existing evidence-based programs for these young people, like *Power Through Choices* [14] and *Making Proud Choices* (adapted for systems involved youth) [63], are delivered via group-based instruction, which may limit young people's ability to access content when they need it. A recent meta-analysis focused on digital sexual health promotion interventions found that most of the studies reviewed had a positive impact on cognitive perceptions and sexual behavior outcomes, highlighting the potential for digital interventions [64].

The Next4You study addresses current gaps in sexual health promotion for young people in foster care by testing a fully mobile sexual health program created for foster youth and informed by their lived experiences, allowing young people to access content they identified as key and return to it when they want or need it. Our co-design process actively engaged young people, ensuring the intervention was relevant, authentic, and tailored to their specific needs. A "nothing about us without us" approach not only makes the content more engaging and relatable for participants, but also represents an innovative, trauma-informed methodology that is rarely implemented in adolescent sexual health interventions [65]. The content identified by Next4You co-designers aligns with the findings reported by Aparicio et al [22], with microlessons focused on content from all 3 core theoretical categories identified by foster youth and foster care staff they interviewed (reproductive system health, relationship health, and mental health).

To illustrate, a core focus of the Next4You intervention is its concentration on fostering strong relationships—examining their characteristics, exploring power and control and how that may show up in relationships, and learning and practicing communication skills to enhance communication without abuse, among other topics. This content was deemed as most relevant by the co-designers and is taking a more prominent role in sexual health programs now than in the past, which aligns with adolescent developmental needs [66]. Nonetheless, most existing evidence-based sexual health programs include limited coverage on healthy relationships, such as single lessons or activities on healthy and unhealthy elements of relationships. This is insufficient to help youth build healthy relationships, end unhealthy ones, and learn skills to navigate and communicate personal sexual boundaries and negotiate contraceptive use with a partner. *Love Notes* [67], one of the few healthy relationship education curricula for youth, highlights the potential of a more focused approach; however, there remains a significant gap in the current evidence base, particularly for young people in foster care settings. The Next4You study will address this gap and examine the impact of a relationship-based approach for youth with experience in foster care.

Next4You was designed using a fully mobile, self-paced delivery format because digital health interventions can offer important advantages for adolescents, such as greater accessibility, privacy, and personalization of content [64]. Next4You will allow young people to engage with the content at their own pace, on their own time—a flexibility that is especially valuable given the myriad demands on young people's time. The gamified, interactive system is expected to sustain youths' interest and facilitate learning in a youth-friendly way. The mobile format of Next4You is anticipated to lower many traditional barriers to implementation, complementing services provided by existing organizations and educational programs providing direct services and support to foster youth.

Despite these strengths, there are important limitations to consider. A self-paced digital intervention relies on adolescents' motivation and engagement. Foster youth juggling the stresses of placement changes, trauma, and daily life may not consistently complete online modules without encouragement, despite the incentives that we provide for participation. The absence of a live facilitator means there is no built-in adult support to answer questions in real time or to prompt participation if a young person loses interest. Fully mobile programs also require access to devices and Wi-Fi that could create barriers for some, though most young people have phones [16], and there are programs that provide support for home internet, phones, and data plans (eg, California LifeLine Foster Youth Program and FCC LifeLine Program). It is also possible that some sensitive topics could elicit emotional reactions, given participants' trauma histories. In anticipation of this, we have included resources for follow-up and are using a recruitment

XSL•FO
RenderX

model that is anchored at service organizations so that young people have options for follow-up support. This is an area of focus in exploring participants' reactions to the program. Further, Next4You was developed expressly for youth involved in the California foster care system; therefore, findings and program components may not be directly generalizable to foster youth in other states. Further, although this study was sufficiently powered to detect a small effect (approx. 0.25 SDs, considered substantively important per WWC criteria) [58], it may have limited power to detect smaller yet potentially meaningful intervention effects. More broadly, while digital interventions show promise in improving adolescents' sexual health knowledge and attitudes, the evidence base for fully self-paced programs is still emerging, underscoring the value of studies like this one. In summary, the innovative, youth-driven design and scalable digital platform of Next4You are major strengths, but these strategies are not without potential challenges.

Conclusions

The Next4You intervention represents an innovative, youth-centered approach to addressing sexual health disparities among foster youth. If proven effective, Next4You could provide an accessible and scalable solution to address a critical gap in sexual health education for this group of young people. Policymakers and practitioners could leverage this approach to normalize sexual health conversations, reduce stigma, and ultimately empower foster youth to make informed, healthy decisions.

Acknowledgments

The authors wish to thank Alexis Obinna, Julio Romero, MPH, and the staff at Fresno EOC, and all of the recruitment site ambassadors and youth who have participated in this study.

Funding

This publication was made possible by grant 90AP2694-04-00 from the Department of Health and Human Services, Administration for Children and Families, Family and Youth Services Bureau, Personal Responsibility Education Innovative Strategies. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Department of Health and Human Services, Administration for Children and Families. The study methods were reviewed and approved by the funder and technical advisors at Abt Associates before the initiation of recruitment.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to confidentiality and data protection considerations but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshots of the intervention.
[PDF File, 242 KB - resprot_v15i1e77185_app1.pdf]

Multimedia Appendix 2
Study outcome measures.
[DOCX File, 32 KB - resprot_v15i1e77185_app2.docx]

References

1. Osterman MJK, Hamilton BE, Martin JA, Driscoll AK, Valenzuela CP. Births: final data for 2022. Natl Vital Stat Rep 2024 Apr;73(2):1-56. [doi: 10.15620/cdc:145588] [Medline: 38625869]
2. Boonstra HD. Winning campaign: California's concerted effort to reduce its teen pregnancy rate. Guttmacher Policy Rev 2010;13(2) [FREE Full text]
3. Santelli JS, Lindberg LD, Finer LB, Singh S. Explaining recent declines in adolescent pregnancy in the United States: the contribution of abstinence and improved contraceptive use. Am J Public Health 2007 Jan;97(1):150-156. [doi: 10.2105/AJPH.2006.089169] [Medline: 17138906]
4. Oshima KMM, Narendorf SC, McMillen JC. Pregnancy risk among older youth transitioning out of foster care. Child Youth Serv Rev 2013 Oct;35(10):1760-1765. [doi: 10.1016/j.childyouth.2013.08.001] [Medline: 24489422]
5. Dworsky A, Courtney ME. The risk of teenage pregnancy among transitioning foster youth: implications for extending state care beyond age 18. Child Youth Serv Rev 2010 Oct;32(10):1351-1356. [doi: 10.1016/j.childyouth.2010.06.002]

XSL•FO
RenderX

6. Courtney M, Dworsky A, Cusick G, Havlicek J, Perez A, Keller T. Midwest Evaluation of the Adult Functioning of Former Foster Youth: Outcomes at Age 21: Chapin Hall Center for Children; 2007. URL: https://pdxscholar.library.pdx.edu/cgi/viewcontent.cgi?article=1059&context=socwork_fac [accessed 2025-12-04]
7. Tyler KA, Whitbeck LB, Chen X, Johnson K. Sexual health of homeless youth: prevalence and correlates of sexually transmissible infections. Sex Health 2007 Mar;4(1):57-61. [doi: 10.1071/sh06045] [Medline: 17382040]
8. Khalili J, Leung LB, Diamant AL. Finding the perfect doctor: identifying lesbian, gay, bisexual, and transgender-competent physicians. Am J Public Health 2015 Jun;105(6):1114-1119. [doi: 10.2105/AJPH.2014.302448] [Medline: 25880937]
9. Decker MJ, Berglas NF, Brindis CD. A call to action: developing and strengthening new strategies to promote adolescent sexual health. Societies (Basel) 2015;5(4):686-712. [doi: 10.3390/soc5040686]
10. Boonstra HD. Teen pregnancy among young women in foster care: a primer. Guttmacher Policy Rev 2011;14(2) [FREE Full text]
11. Wilson BDM, Cooper K, Kastanis A, Nezhad S. Sexual and gender minority youth in foster care: assessing disproportionality and disparities in Los Angeles. : The Williams Institute, UCLA School of Law; 2014 Aug URL: https://williamsinstitute.law.ucla.edu/wp-content/uploads/SGM-Youth-in-Foster-Care-Aug-2014.pdf [accessed 2025-12-04]
12. Gowen LK, Winges-Yanez N. Lesbian, gay, bisexual, transgender, queer, and questioning youths' perspectives of inclusive school-based sexuality education. J Sex Res 2014;51(7):788-800. [doi: 10.1080/00224499.2013.806648] [Medline: 24003908]
13. Widman L, Nesi J, Kamke K, Choukas-Bradley S, Stewart JL. Technology-based interventions to reduce sexually transmitted infections and unintended pregnancy among youth. J Adolesc Health 2018 Jun;62(6):651-660. [doi: 10.1016/j.jadohealth.2018.02.007] [Medline: 29784112]
14. Oman RF, Vesely SK, Green J, Clements-Nolle K, Lu M. Adolescent pregnancy prevention among youths living in group care homes: a cluster randomized controlled trial. Am J Public Health 2018 Feb;108(S1):S38-S44. [doi: 10.2105/AJPH.2017.304126] [Medline: 29443557]
15. Wartella E, Rideout V, Montague H, Beaudoin-Ryan L, Lauricella A. Teens, health and technology: a national survey.. MaC 2016;4(3):13-23. [doi: 10.17645/mac.v4i3.515]
16. Anderson M, Jiang J. Teens, social media and technology 2018. Pew Research Center. 2018 May 31. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/ [accessed 2025-12-04]
17. Manlove J, Cook E, Whitfield B, Johnson M, Martínez-García G, Garrido M. Short-term impacts of pulse: an app-based teen pregnancy prevention program for Black and Latinx women. J Adolesc Health 2020 Feb;66(2):224-232. [doi: 10.1016/j.jadohealth.2019.08.017] [Medline: 31690536]
18. Zeldin S, Petrokubi J, Collura J, Camino L, Skolaski J. Strengthening communities through youth participation: lessons learned from the ACT for Youth Initiative. : ACT for Youth Center of Excellence; 2009 Mar URL: https://ecommons.cornell.edu/server/api/core/bitstreams/8805028b-6d4a-41e2-99cf-bfdb008c0070/content [accessed 2025-12-04]
19. Patton GC, Sawyer SM, Santelli JS, et al. Our future: a Lancet commission on adolescent health and wellbeing. Lancet 2016 Jun 11;387(10036):2423-2478. [doi: 10.1016/S0140-6736(16)00579-1] [Medline: 27174304]
20. Adolescent and youth engagement strategic framework. UNICEF. 2017. URL: https://childhub.org/sites/default/files/library/attachments/adolescent_and_youth_engagement_en.pdf [accessed 2025-12-04]
21. Peterson A, Drake P, Tat S, Silver G, Bell H, Guinosso S. Youth engagement in sexual health programs and services: findings from the youth engagement network's environmental scan. : ETR; 2020 Dec URL: https://pages.etr.org/hubfs/YEN/ETR_YEN%20Environmental%20Scan%20Full%20Report_2020.pdf [accessed 2025-12-04]
22. Aparicio EM, Kachingwe ON, Salerno JP, Geddings-Hayes M, Boekeloo BO. Addressing sexual health among youth in foster care group homes: a community-engaged grounded theory study. Sex Res Soc Policy 2021 Dec;18(4):1136-1147. [doi: 10.1007/s13178-020-00519-w]
23. Furman W. The romantic relationships of youth. In: Bukowski W, Laursen B, Rubin KH, editors. Handbook of Peer Interactions, Relationships, and Groups. 2nd ed: Guilford Press; 2018:410-428.
24. Carver K, Joyner K, Udry JR. National estimates of adolescent romantic relationships. In: Florsheim P, editor. Adolescent Romantic Relations and Sexual Behavior: Psychology Press; 2003:37-70 URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9781410607782-3/national-estimates-adolescent-romantic-relationships-karen-carver-kara-joyner-richard-udry
25. Manning WD, Giordano PC, Longmore MA. Hooking up: the relationship contexts of "nonrelationship" sex. J Adolesc Res 2006;21(5):459-483. [doi: 10.1177/0743558406291692]
26. Forenza B, Bermea A, Rogers B. Ideals and reality: perceptions of healthy and unhealthy relationships among foster youth. Child Adolesc Soc Work J 2018 Jun;35:221-230. [doi: 10.1007/s10560-017-0523-3]
27. Wekerle C, Leung E, Wall AM, et al. The contribution of childhood emotional abuse to teen dating violence among child protective services-involved youth. Child Abuse Negl 2009 Jan;33(1):45-58. [doi: 10.1016/j.chiabu.2008.12.006] [Medline: 19167066]
28. What Works Clearinghouse Standards Handbook, Version 4.1. : U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance; 2020 URL: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/WWC-Standards-Handbook-v4-1-508.pdf [accessed 2025-12-04]
29. Hug T, Lindner M, Bruck PA, editors. Micromedia & e-Learning 20: Gaining the Big Picture: Proceedings of Microlearning Conference 2006: Innsbruck University Press; 2006.

XSL•FO
RenderX

30. Zarshenas L, Saranjam E, Mehrabi M, Setoodeh G. Microlearning and gamification in anxiety management among girl adolescents in Iran: an interventional study. Pak J Med Health Sci 2020;14(6):689-694 [FREE Full text]
31. Damon W. What is positive youth development? Ann Am Acad Pol Soc Sci 2004 Jan;591(1):13-24. [doi: 10.1177/0002716203260092]
32. Reyna VF, Brainerd CJ. Dual processes in decision making and developmental neuroscience: a fuzzy-trace model. Dev Rev 2011 Sep;31(2-3):180-206. [doi: 10.1016/j.dr.2011.07.004] [Medline: 22096268]
33. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: Prentice Hall; 1986.
34. Youth-centred digital health interventions: a framework for planning, developing and implementing solutions with and for young people. World Health Organization. 2021 Jul 19. URL: https://www.who.int/publications/i/item/9789240011717 [accessed 2025-12-04]
35. Hamilton G, Scrivener S. The central role of implementation research in understanding treatment contrast. Implementation Research Incubator (Blog). 2018 Jul. URL: https://www.mdrc.org/publication/central-role-implementation-research-understanding-treatment-contrast [accessed 2025-12-04]
36. What Works Clearinghouse Procedures Handbook, Version 4.0. : U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance; 2017 URL: https://ies.ed.gov/ncee/wwc/docs/referenceresources/wwc_procedures_handbook_v4.pdf [accessed 2025-12-04]
37. Hightow-Weidman LB, Bauermeister JA. Engagement in mHealth behavioral interventions for HIV prevention and care: making sense of the metrics. Mhealth 2020;6:7. [doi: 10.21037/mhealth.2019.10.01] [Medline: 32190618]
38. Losing weight: healthy eating/being active [Pamphlet]. ETR. 2008. URL: http://www.etr.org/store/product/losing-weight-healthy-eatingbeing-active/ [accessed 2025-12-04]
39. MyPlate on a budget [Pamphlet]. ETR. 2012. URL: https://web.archive.org/web/20250916043416/https://www.etr.org/store/product/myplate-on-a-budget/ [accessed 2025-12-4]
40. Can you eat healthy? Try it out! [Pamphlet]. ETR. 2012. URL: https://web.archive.org/web/20231202212131/http://www.etr.org/store/product/can-you-eat-healthy-try-it-out/ [accessed 2025-12-04]
41. How to read a food label [Pamphlet]. ETR. 2011. URL: https://web.archive.org/web/20250421050025/http://www.etr.org/store/product/how-to-read-a-food-label/ [accessed 2025-12-04]
42. Fitness facts: get moving for better health. ETR. 2015. URL: http://www.etr.org/store/product/fitness-facts-get-moving-for-better-health/ [accessed 2025-12-04]
43. Move more: feel great. ETR. 2010. URL: http://www.etr.org/store/product/move-more-feel-great/ [accessed 2025-12-04]
44. Getting fit with no time and no money [Pamphlet]. ETR. 1999. URL: http://www.etr.org/store/product/getting-fit-with-no-time-and-no-money/ [accessed 2025-12-4]
45. 101 stress busters. ETR. 1997. URL: http://www.etr.org/store/product/101-stress-busters/ [accessed 2025-12-4]
46. 5 ways to stop stress. ETR. 2003. URL: http://www.etr.org/store/product/5-ways-to-stop-stress/ [accessed 2025-12-4]
47. Anxiety: help on the way. ETR. 2005. URL: https://www.etr.org/store/product/anxiety-help-on-the-way/ [accessed 2025-12-4]
48. 5 smart steps to less stress. ETR. 1997. URL: http://www.etr.org/store/product/5-smart-steps-to-less-stress/ [accessed 2025-12-4]
49. Getting what you want from sleep. ETR. 1997. URL: https://www.etr.org/store/product/getting-what-you-want-from-sleep/ [accessed 2025-12-4]
50. Anderson P, Coyle K, Guinosso S, et al. Promoting adolescent healthy relationships (The About Us Program): protocol for a randomized clinical trial. JMIR Res Protoc 2021 Sep 1;10(9):e30499. [doi: 10.2196/30499] [Medline: 34468330]
51. Humphreys TP, Brousseau MM. The sexual consent scale-revised: development, reliability, and preliminary validity. J Sex Res 2010 Sep;47(5):420-428. [doi: 10.1080/00224490903151358] [Medline: 19685367]
52. Quinn-Nilas C, Milhausen RR, Breuer R, et al. Validation of the Sexual Communication Self-Efficacy Scale. Health Educ Behav 2016 Apr;43(2):165-171. [doi: 10.1177/1090198115598986] [Medline: 26286296]
53. Helweg-Larsen M, Collins BE. The UCLA Multidimensional Condom Attitudes Scale: documenting the complex determinants of condom use in college students. Health Psychol 1994;13(3):224-237. [doi: 10.1037/0278-6133.13.3.224] [Medline: 8055858]
54. Kyes BK. Contraceptive attitude scale. In: Davis CM, Yarber WL, Bauserman R, Schreer G, Davis SL, editors. Handbook of Sexuality-Related Measures: Sage Publications; 1998:164-165.
55. Pulerwitz J, Gortmaker SL, DeJong W. Measuring sexual relationship power in HIV/STD research. Sex Roles 2000 Apr;42(7-8):637-660. [doi: 10.1023/A:1007051506972]
56. Coyle K, Anderson P, Laris BA, Barrett M, Unti T, Baumler E. A group randomized trial evaluating high school FLASH, a comprehensive sexual health curriculum. J Adolesc Health 2021 Apr;68(4):686-695. [doi: 10.1016/j.jadohealth.2020.12.005] [Medline: 33583683]
57. Danes SM, Haberman HR. Teen financial knowledge, self-efficacy, and behavior: a gendered view. Financ Couns Plan 2007;18(2):48-60 [FREE Full text]
58. What Works Clearinghouse Procedures and Standards Handbook, Version 5.0. : U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance; 2022 URL: https://ies.ed.gov/ncee/wwc/Docs/referenceresources/Final_WWC-HandbookVer5.0-0-508.pdf [accessed 2025-12-04]

XSL•FO
RenderX

59. Wooldridge JM. Introductory Econometrics: A Modern Approach: Cengage Learning; 2003.
60. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
61. Ball B, Hoefer S, Faulkner M, et al. Innovation in sexuality and relationship education in child welfare: shifting toward a focus on ongoing conversations, connection, and consent. Prev Sci 2023 Dec;24(Suppl 2):163-173. [doi: 10.1007/s11121-022-01476-z] [Medline: 36630023]
62. Oman RF, Vesely SK, Green J, Fluhr J, Williams J. Short-term impact of a teen pregnancy-prevention intervention implemented in group homes. J Adolesc Health 2016 Nov;59(5):584-591. [doi: 10.1016/j.jadohealth.2016.07.002] [Medline: 27544458]
63. Combs KM, Aparicio EM, Prince DM, Grinnell-Davis C, Marra L, Faulkner M. Evidence-based sexual health programs for youth involved with juvenile justice and child welfare systems: outcomes across settings. Child Youth Serv Rev 2019 May;100:64-69. [doi: 10.1016/j.childyouth.2019.02.032]
64. Sewak A, Yousef M, Deshpande S, Seydel T, Hashemi N. The effectiveness of digital sexual health interventions for young adults: a systematic literature review (2010-2020). Health Promot Int 2023 Feb 1;38(1):daac104. [doi: 10.1093/heapro/daac104] [Medline: 36757346]
65. Checkoway B. What is youth participation? Child Youth Serv Rev 2011 Feb;33(2):340-345. [doi: 10.1016/j.childyouth.2010.09.017]
66. Goldfarb ES, Lieberman LD. Three decades of research: the case for comprehensive sex education. J Adolesc Health 2021 Jan;68(1):13-27. [doi: 10.1016/j.jadohealth.2020.07.036] [Medline: 33059958]
67. Barbee AP, Cunningham MR, van Zyl MA, Antle BF, Langley CN. Impact of two adolescent pregnancy prevention interventions on risky sexual behavior: a three-arm cluster randomized control trial. Am J Public Health 2016 Sep;106(S1):S85-S90. [doi: 10.2105/AJPH.2016.303429] [Medline: 27689500]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
ITT: intent-to-treat
STI: sexually transmitted infection
WWC: What Works Clearinghouse

Edited by A Schwartz; submitted 08.May.2025; peer-reviewed by J Noel-MacDonell; accepted 12.Nov.2025; published 03.Feb.2026.

Please cite as:
Anderson PM, Ong SL, Elgin D, Laird J, Starr D, Daniels S, Rogers J, Coyle KK
Youth-Centered Mobile Intervention (Next4You) to Promote Healthy Relationships and Sexual Wellness Among Adolescents in or Transitioning From Foster Care: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e77185
URL: https://www.researchprotocols.org/2026/1/e77185
doi:10.2196/77185

© Pamela M Anderson, Seow Ling Ong, Dallas Elgin, Jennifer Laird, Darriya Starr, Sade Daniels, Jakara Rogers, Karin K Coyle. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 3.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness and Cost-Effectiveness of a Digital Falls Prevention Program Versus Usual Care to Improve Balance, Falls Risk, and Function in Older Adults: Protocol for the Keep-On-Keep-Up Randomized Controlled Trial

Emma Stanmore[1,2,3], PhD; Amelia Parchment[2,3], PhD; Bolanle Odebiyi[3], MSc; Peter Bower[2,3], PhD; Chloe French[3], PhD; Chunhu Shi[2,3], PhD; Saima Bashir[1,2], PhD; Saima Ahmed[2,3], PhD; Dawn Dowding[2,3], PhD; Jo Dumville[2,3], PhD; Roman Kislov[2,3,4], PhD; Alex Thompson[3], PhD; Dawn A Skelton[5], PhD; Margaret Clarke[3]; Chris Todd[1,2,3], PhD

[1]Manchester Academic Health Science Centre, Manchester, United Kingdom

[2]NIHR Applied Research Collaboration Greater Manchester, Manchester, United Kingdom

[3]School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Jean McFarlane Building, Oxford Road, Manchester, United Kingdom

[4]Faculty of Business and Law, Manchester Metropolitan University, Manchester, United Kingdom

[5]Glasgow Caledonian University, School of Health and Life Sciences, Research Centre for Health (ReaCH), Glasgow, United Kingdom

Corresponding Author:
Emma Stanmore, PhD
Manchester Academic Health Science Centre, Manchester, United Kingdom

Abstract

Background: Falls are the primary cause of fatal and nonfatal accidental injuries in older adults. The World Falls Prevention Guidelines recommend balance-challenging, functional exercise programs as a key strategy for falls prevention, but access, uptake, and adherence to these programs in community settings remain suboptimal. Keep-On-Keep-Up (KOKU), a digital, National Health Service–approved program, was codeveloped with older adults and therapists to provide progressive, evidence-based exercises and to raise awareness of falls prevention strategies.

Objective: This trial aims to investigate the effectiveness and cost-effectiveness of the KOKU digital strength and balance program for improving balance, enhancing physical function, and reducing falls risk among community-dwelling older adults.

Methods: This is a 2-arm, parallel-group randomized controlled trial. A total of 196 community-dwelling older adults 60 years and older will be randomized to either the intervention group, comprising a digital strength and balance program (KOKU) alongside standard care (strength and balance exercise advice and a falls prevention leaflet), or to a control group, receiving standard care only. Participants receiving the intervention will be asked to exercise 3 times per week following the tailored and progressive program. Randomization will take place after recruitment and baseline data collection. The trial's primary outcome measure is balance function (Berg Balance Score) at 12 weeks post randomization. Secondary trial outcomes include lower limb strength, health care utilization and health-related quality of life, self-reported concerns about falling, self-reported physical activity, falls risk, pain, mood, fatigue, self-reported falls, and acceptability and usability of the KOKU program. Intention-to-treat analysis and a cost-effectiveness analysis will be employed for trial data analysis. Qualitative interviews and focus groups will be undertaken with around 10 care providers and 13 participants to further understand views of the intervention and trial processes.

Results: This study began recruitment in July 2024 and concluded in March 2025, recruiting a total of 202 participants (102 intervention group and 100 control group). Following protocol publication, data compilation and analysis will be conducted, with results anticipated to be published in 2027.

Conclusions: This trial will provide important evidence on whether a digital strength and balance program can improve balance and related outcomes in older adults compared with usual care.

Trial Registration: ClinicalTrials.gov NCT06687135; https://clinicaltrials.gov/study/NCT06687135

International Registered Report Identifier (IRRID): DERR1-10.2196/78840

(JMIR Res Protoc 2026;15:e78840) doi:10.2196/78840

KEYWORDS

aged; accidental falls/prevention; postural balance; randomized controlled trial; digital health intervention; mobile apps

XSL•FO
RenderX

Introduction

With population aging, supporting older adults to live longer, healthier, independent lives is a major societal challenge [1]. Falls are a common problem for older adults, with risk increasing exponentially with age, leading to growing health and social care costs [2,3]. These include ambulance callouts, fractures, hospital admissions, premature admission to care homes, and increased dependency. The Office for Health Improvement and Disparities estimates that in England, the total annual cost of falls-related fragility fractures was £4.4 billion (~US $6 billion) [4]. Systematic reviews of randomized controlled trial (RCT) evidence conclude that there is good evidence that strength and balance exercises are the most effective approach to improve function and reduce falls [5,6]. Effectiveness increases if the programs are challenging, of sufficient dose, progressive, and maintained [7].

In the United Kingdom, strength and balance training programs are usually delivered in community centers by qualified professionals. These programs can be effective but may be inaccessible to those who are homebound, without transport, and unavailable in some locations [8].

A growing evidence base suggests that digital interventions may be a feasible tool for the delivery of falls prevention and exercise programs among older adults [9-11]. Moreover, digital formats can incorporate health behavior change techniques such as goal setting, progress tracking, and feedback on individual progress and can be more user-friendly and appealing, which may improve engagement and adherence [12]. In comparison with face-to-face alternatives, digital platforms are relatively inexpensive, scalable, and can be safely used in the user's own home at their convenience [12,13].

Keep-On-Keep-Up (KOKU) is a UK National Health Service (NHS)-approved digital health intervention for falls prevention [14]. KOKU can be used independently by older adults to instruct them how to undertake progressive strength and balance exercises based on the Otago Exercise Program and the Falls Management Exercise Program (FaME) [15,16]. KOKU also has 6 health literacy games to raise awareness of home hazards and ways to improve bone health, brain health, nutrition, and hydration. KOKU's development was guided by the Healthtech Innovation Cycle [17], the EAST Behavioral Insights Framework [18], and the Medical Research Council guidance for designing and evaluating complex interventions [19]. KOKU has been iteratively co-developed with therapists and older adults, informed by international guidelines for falls prevention [7], a meta-analysis of exercises for falls prevention [6], health behavior change techniques [20], and consensus studies with outcome and reporting recommendations for falls prevention strategies [21].

This is one of the first digital evidence-informed strength and balance programs with both progressive exercises and gamified education for older people that acts as an intervention to prevent falls and functional decline. Feedback from users and stakeholders in usability and acceptability studies in the United Kingdom [22] and the United States [9] has been positive.

In this paper, we describe an RCT to evaluate the effectiveness and cost-effectiveness of a digital exercise program, KOKU, for improving balance and function and reducing falls risk in community-dwelling older adults aged 60 years and above.

Methods

Public and Community Involvement and Engagement

Patient and public involvement was used to inform the design of the first iteration of the KOKU program and has continued to be important for ongoing development and to inform the design of the trial. Patient and public involvement feedback has led to the development of KOKU in both iOS and Android formats to be more digitally inclusive and informed the study methods such as how to recruit participants through care providers. The study is supported by the NIHR Applied Research Collaboration Greater Manchester (ARC-GM), which has a public and community involvement and engagement lead who has convened 2 meetings during the preparation of the study methods to ensure that a diverse group of older adults' views were incorporated into the research processes. During the design phase of this study, the research team visited community organizations supporting social care (ie, not health service providers) to explain the study and gain an insight into the organization's daily activities and operations to inform study processes.

Trial Design

The 2-arm parallel-group RCT will compare the effectiveness of a digital strength and balance program (KOKU) plus standard care (strength and balance exercise advice and a falls prevention leaflet) against standard care alone. The study will test the effectiveness and cost-effectiveness of the KOKU program, and the primary outcome measure will be balance, as measured by the Berg Balance Scale [23].

This study is funded by the NHS England CareTech Program, which has a remit to identify care technology solutions and provide support to gain robust evidence that could have potential for wider rollout within the care sector. The trial opened to recruitment in July 2024 and closed in March 2025, with findings expected to be published in 2027. This protocol paper follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [24] (Checklist 1) and the TIDieR (Template for Intervention Description and Replication) checklist for the description of the intervention [25] (Checklist 2).

Participants

Participants will be recruited from a variety of organizations and social care providers based in community settings in Greater Manchester and Lancashire areas in the Northwest of England. These include adult social care teams, housing providers, and charity organizations that support older adults to remain as independent as possible in their own homes. The community organization providers will be recruited via the Greater Manchester local authorities' adult social care managers, who will then act as gatekeepers to recruit older adults. All eligible research participants will receive a participant information sheet and will be asked to complete an informed consent form by the

XSL•FO
RenderX

research team, attesting to their voluntary participation in the study.

Participants will include community-dwelling older adults aged 60 years and over, referred via community care providers in Greater Manchester and Lancashire, and eligible to take part in the study if they have the following characteristics:

- are willing and able to give informed consent
- speak English
- are able to see and safely use the tablet-based KOKU program
- can read instructions with or without glasses, as assessed by the trained research staff
- are able to mobilize indoors without the help of another person, with or without a walking aid (self-reported)

The exclusion criteria for the older adult participants are as follows:

- those who are currently using digital technologies to exercise
- currently taking part in an exercise program or undertaking physiotherapy
- those who are unable to comprehend the study procedures
- those who have medical contraindications to exercise due to unstable health or severe conditions (eg, acute illness, severe congestive cardiac failure, uncontrolled hypertension, recent fracture or surgery; myocardial infarction or stroke in past 6 mo; severe cognitive impairment; orthopedic surgery in last 6 mo or on a waiting list to have orthopedic surgery; wheelchair users; severe auditory or visual impairment; or other uncontrolled medical conditions likely to compromise the ability to exercise)

We also aim to recruit care provider staff (n=10) and a purposive sample of the participating older adults (n=13) to participate in a qualitative interview or focus group at the end of the intervention period to provide insights and feedback on the usability of the KOKU program and the trial processes.

Recruitment, randomization, data analysis, and interpretation will be undertaken by the Applied Research Collaboration-Greater Manchester (led by PB), independent from KOKU Health.

Randomization and Blinding

Randomization will be performed using Research Electronic Data Capture (REDCap V 14.0.31 [26]) randomization module to allocate older adult participants in a 1:1 ratio to either the intervention or control group. Random permuted blocks of various sizes will be used to produce similar group sizes and preserve allocation concealment. The allocation sequence will be computer-generated and uploaded into the REDCap system by a researcher who is not involved in enrolling or assigning participants to groups.

Responsibility for enrolling participants into the trial lies with the principal investigator and staff at the sites. Allocation concealment will be ensured, as the system will not release the allocation until the participant has been randomized by a researcher not involved in recruitment, after baseline assessments have been taken. Given the nature of the trial, researchers completing data collection and participants will not be blinded. The researcher undertaking data analyses will be blinded to the intervention and control groups to minimize bias.

Intervention

Participants allocated to the intervention group will receive an iPad for the intervention period and will be recommended to train with KOKU for 20 to 30 minutes, 3 times per week for 12 weeks. The researcher will explain and also provide written instructions on how to use the iPad and the KOKU program, including how to switch on and charge the device. Upon first opening the KOKU program, participants are asked to complete a short onboarding questionnaire (eg, age and gender) and a walkthrough featuring written, audio, and animation instructions. The researchers will give their contact details so additional digital literacy support can be provided in person or via phone calls, if required.

The digital program includes 26 strength and balance exercises that progress from seated to standing and walking exercises. As participants complete the program, the exercises progress according to responses to questions such as “Do you feel out of breath?“ or “Could you perform the exercises easily?” Each exercise has safety information and an animated coach who demonstrates correct movements with audio and written information. The exercises include heel raises, sit-to-stands, side leg raises (abductors), side, front, and back lunges, heel-to-toe walking, sideways walking, and walking backward. If participants forget to complete their weekly exercises, they are sent an app notification after 1 week to remind them to continue.

After each exercise, participants can see how many training sessions have been completed on a progress screen embedded within the KOKU program and how many sessions remain before the 12-week intervention is completed.

The KOKU program is informed by behavior change theory and techniques such as gamification to improve access and uptake of strength and balance exercises. In addition to the exercises, there are 6 complementary health literacy games to help users learn about how to promote bone health, nutrition, hydration, and safety in the home. The bone health game, for example, asks participants to choose items from a fridge that are known to deplete or enhance calcium, vitamin D, and associated nutrients or minerals. The animated character’s X-ray bones grow or shrink, and a bone barometer increases or decreases. Evidence-based messages in lay language are presented after each selection to explain why a choice of food or drink is beneficial or not for improving bone health. The embedded health literacy games were developed with international experts and are popular with users due to the gamified elements that increase engagement and enjoyment. A user instruction video provides a short demonstration of how to download and use KOKU [27].

Participants in the intervention group will also receive home exercise leaflets based on FaME [15] and Otago Exercise Program [16], and falls prevention advice alongside the KOKU digital program.

XSL•FO
RenderX

Participants in the control group will receive the same leaflets without KOKU and will be asked to undertake the exercises listed in the leaflet 3 times a week in their leisure time.

Outcomes

At baseline, a bespoke questionnaire will capture past medical history (including fractures), medication, comorbidities, self-reported level of vision, and demographic data (age, gender, ethnicity, and socioeconomic status). Outcome data will be collected by trained researchers from all participants at baseline, 6 weeks, and at the end of the 12-week period (Table 1).

XSL•FO
RenderX

Table . KOKU[a] trial schedule of enrollment, interventions, and assessments.

Timepoint	Prescreening	Enrolment/allocation	Study period postallocation			
			Baseline	Week 6	Week 12	Week 24
Recruitment—research team						
Identify potential care providers for study	✓					
Familiarize care providers with research project	✓					
Recruitment—care providers						
Identify potential study participants	✓					
Patient enrollment						
Eligibility screen	✓	✓				
Informed consent		✓				
Training to use KOKU		✓				
Interview/focus group					✓	
Care provider staff enrollment						
Eligibility screen		✓				
Informed consent		✓				
Training to use KOKU		✓				
Interview/focus group					✓	
Allocation—Interventions						
KOKU			✓	✓	✓	
OEP[b] leaflet and falls prevention leaflet			✓	✓	✓	
Assessments						
BBS[c]			✓	✓	✓	
5StS[d]			✓	✓	✓	
Health questionnaires[e]				✓	✓	
Adverse events			✓	✓	✓	
Usability/acceptability questionnaires[f]					✓	
Daily falls calendar follow-up			✓	✓	✓	✓

[a]KOKU: Keep-On-Keep-Up.

[b]OEP: Otago Exercise Programme.

[c]BBS: Berg Balance Scale.

[d]5StS: Five Times Sit-to-Stand chair test.

XSL•FO
RenderX

[e]Health questionnaires include 5-item Geriatric Depression Scale, Physical Activity Scale for the Elderly, Short Falls Efficacy Scale-International, Assessment of Falls Risk tool, and EQ-5D-5L.

[f]Usability/acceptability questionnaires include the System Usability Scale, the Short Version of the User Experience Questionnaire, and the theoretical framework for acceptability.

The primary outcome measure is balance function, as measured by the validated Berg Balance Scale (BBS) [23]. The test involves 14 assessments of static and dynamic balance and is useful in predicting the risk of falls and indicating an individual's ability to mobilize and carry out activities of daily living.

Secondary outcomes measured at 6 and 12 weeks will be measured by the following assessments:

1. The Five Times Sit-to-Stand (5-STS) chair test [28] will be completed to assess lower limb strength. For this test, the participant is asked to stand up and sit down as quickly as possible 5 times from a standard-height chair (43.2 cm) with their arms folded. The researcher records the time taken to complete the test.
2. Health-related quality of life will be measured using the EQ-5D-5L [29], which considers 5 dimensions: mobility, self-care, usual activities, pain or discomfort, and anxiety. Each dimension is rated as one of the following 5 options: no problems, slight problems, moderate problems, severe problems, and extreme problems.
3. Mood will be assessed using the 5-item Geriatric Depression Scale. This screening tool for depression in older adults is validated to be as effective as the 15-item Geriatric Depression Scale in a wide range of settings [30].
4. The level of physical activity will be measured using the Physical Activity Scale for the Elderly questionnaire [31]. This is a validated and easy-to-administer questionnaire designed to assess physical activity in those 65 years and older.
5. Concerns about falling will be measured using the Short Falls Efficacy Scale—International [32]. The Short Falls Efficacy Scale—International is a validated and reliable 7-item tool that measures concerns about falling in relation to a range of everyday activities and is available in a number of languages [32,33].
6. The risk of falling will be assessed using the assessment of falls risk tool. This validated measure [34] has 5 items, which include a history of any fall in the previous year, 4 or more prescribed medications, diagnosis of stroke or Parkinson disease, reported problems with balance, and inability to rise from a chair without using arms. A visual analog score for pain and fatigue will be included, as these are all strong indicators of risk of falling.
7. Secondary outcomes will also consider the usability and acceptability of KOKU. At the 12-week assessment, participants in the intervention group will complete questionnaires focusing on usability, acceptability, and motivation to use KOKU. The System Usability Scale will assess the usability of KOKU from the users' perspective [35]. The theoretical framework for acceptability will assess perceived acceptability of the KOKU system [36]. The short version of the User Experience Questionnaire will assess the usability of the KOKU system [37].
8. Falls will be assessed using daily self-report fall calendars for 24 weeks, with follow-up phone calls from a researcher to find out more details of any falls reported and any use of health care resources. This will include details on how the fall occurred, if a visit to a general practitioner (GP), hospital, or emergency department was required, and if the fall resulted in a fracture, swelling, or other self-reported injury. Fall severity will be coded as no injury, moderate injury, or serious injury based on consultation between 2 authors (CF, ES), including a registered nurse and without discussion of group allocation. Falls and all injuries will be recorded during the 6-month period, with falls defined as "an unexpected event in which participants come to rest on the ground, floor, or other lower level" [21].
9. Health care resource use: Resource use data will be collected during follow-up to measure health service utilization related to falls. This includes ambulance attendances, accident and emergency visits, general practitioner consultations, and hospital admissions (public or private). This will be analyzed descriptively as a secondary outcome and will inform the cost-effectiveness analysis.

Data Management and Quality Assurance

Data generated will be accessible only to the research team. All data will be collected and managed using the secure, web-based software platform REDCap [26,38], hosted at The University of Manchester. The data collection forms in REDCap will use functions that check for mandatory information, data ranges, and alerts whenever data violate specific limits [38]. Paper documents, such as written consent forms, will be stored in a locked cabinet in a secure area.

Digital copies of data will be stored securely using the University of Manchester's Research Data Storage (RDS) service while pending, during, and following analysis. Any physical copies of the data will be stored in a locked filing cabinet within a locked office building on the university campus.

Study data and materials may be looked at by individuals from the University of Manchester, regulatory authorities, or the NHS Trust for monitoring and auditing purposes, and this may well include access to personal information. A member of the participant's care team trained for this study may also gain access to these data where essential for the identification of an individual participant.

A unique alphanumeric identification code, and subsequently a pseudonym, will be assigned for each participant. Data entry will be checked for completeness intermittently throughout the study by a researcher not involved in data collection (CS), and a trial monitoring team, independent from the sponsor and funder, will monitor progress and data management procedures every month. Direct quotations from interviews or focus groups may be used, but these will be anonymized prior to publication

XSL•FO
RenderX

and in interview transcripts. Participants will be given pseudonyms in transcripts and publications.

Sample Size

Parameter estimates are based on a previous study of a digital intervention for adults in community settings with BBS as the primary outcome measure at 12 weeks [12]. Based on data from this previous study, we estimate the SD to be 14. Correlation between baseline and follow-up was 0.81 in the previous trial, but we will use a conservative value of 0.7. To achieve 90% power at a 5% significance level, with inflation for 13% attrition, we estimate that a sample size of 196 (98 per arm) is required to demonstrate the minimal clinically important change needed to reduce falls risk and improve functional balance of a difference in means of 5 points on the BBS at 12 weeks [39]. A minimal clinically important change of 5 points in the BBS was chosen according to baseline BBS scores of 35 to 44 observed in a previous study of older adults in the same region where the trial will be conducted [12].

Adverse Events and Risk

The main concern regarding adverse events and/or risk is whether or not the participant can come to any harm from using the KOKU digital program. No disproportionate adverse events are expected, as all of the exercises within the KOKU digital program have been well researched with older people and mimic those of a routine physiotherapy session. Participants will be trained on how to use KOKU safely and will be requested to inform the Principal Investigator in the event of any injury, after seeking appropriate medical assistance. Any participant suspected of having an adverse reaction will be immediately withdrawn from the study, with the cause of the event thoroughly investigated. The Principal Investigator will determine seriousness and causality in conjunction with any treating medical practitioners. A serious adverse event that is deemed directly related to, or suspected to be related to, the trial intervention will be reported to the ethics committee. All trial staff in contact with participants will be responsible for recording adverse events, and the adverse event procedure will be followed.

Statistical Analysis

All statistical analyses will be performed using Stata (StataCorp) or SPSS (version 28, IBM Corp; 2021).

Baseline demographic and clinical characteristics will be summarized by trial arm using means and SD, or median and interquartile ranges for continuous variables, as appropriate, and frequencies with percentages for categorical variables.

The primary analysis will be conducted following the intention-to-treat (ITT) principle, including all participants providing evaluable outcome data irrespective of their engagement with the KOKU program or control intervention.

BBS will be analyzed using a mixed model for repeated measures (MMRM). The model will include the fixed effects of trial arm (ie, KOKU and control), time point (ie, 6 and 12 wk), and the fixed covariate of baseline scores, as well as the interactions of all fixed terms by time point, ensuring the effects are permitted to vary for each of the 2 time points. An unstructured covariance matrix will be used, and the model will be fitted using restricted maximum likelihood (RML) estimation. The primary comparison of interest is between group differences at 12 weeks. Results will be presented as regression coefficients with 95% CIs. Continuous secondary outcomes will be analyzed using a similar framework to the primary outcome.

The rate of falls will be analyzed using a parametric model for count data (eg, Poisson or negative binomial), adjusting for baseline values. The appropriate model will be chosen depending on the distribution of this variable. The incident rate ratio will be presented with a 95% CI.

All the models will be adjusted for baseline values. Thus, to minimize data loss, simple mean imputation (across the randomized groups) of the corresponding outcome data will be used. MMRM models are likelihood-based and therefore consistent with the missing-at-random assumption. Sensitivity analyses to assess the robustness of the MAR assumption will be performed for the primary outcome.

Quantitative data will be managed using REDCap [26], a secure, web-based software platform designed to support data capture for research studies, providing (1) an intuitive interface for validated data capture; (2) audit trails for tracking data manipulation and export procedures; (3) automated export procedures for seamless data downloads to common statistical packages; and (4) procedures for data integration and interoperability with external sources.

Economic Analysis

The primary objective of the economic analysis is to assess the incremental cost-effectiveness of the KOKU program compared with usual care. A cost-effectiveness analysis (CEA) will be conducted alongside the trial from the NHS perspective, excluding patient-incurred costs. The base-case CEA will adopt a 12-week time horizon, consistent with the trial follow-up period.

Health care resource–use data collected during the telephone follow-ups of participants who report a fall will be combined with relevant unit cost data for the financial year 2023 to 2024 to calculate the total costs incurred over the study period. Unit costs will be obtained from available UK health care reference sources. For primary care, the latest Personal Social Services Research Unit reference costs will be used, while for secondary care, costs will be sourced from the latest national reference cost schedule.

Quality-adjusted life years (QALYs) will be calculated based on EQ-5D-5L utility values using the area-under-the-curve method, assuming linear extrapolation of utility between time points. EQ-5D-5L will be scored using the appropriate tariff as recommended by the National Institute for Health and Care Excellence at the time of data collection. Multiple imputation techniques will be used to handle missing data in the main analysis, with the number of imputations determined by the extent of missing data. Regression analysis will estimate incremental costs and QALYs between the intervention and control groups, adjusting for baseline covariates including sex, age, ethnicity, marital status, socioeconomic status, and baseline utility score. Model selection (such as ordinary least

XSL•FO
RenderX

squares/general linear models) will be based on goodness-of-fit statistics and diagnostic tests for model specification. Incremental net-monetary benefit will be calculated using a willingness-to-pay threshold of £20,000 (~$26,900) per QALY. Cost-effectiveness acceptability curves will be produced to show the probability of cost-effectiveness across a range of thresholds.

Scenario analyses will explore the effects of various assumptions, including variations in the implementation process, the costs associated with training, and the number of imputations. Additionally, exploratory analyses will extend the trial findings and time horizon to a longer-term time horizon (2 y), providing insights into potential long-term outcomes. Both costs and QALYs will be discounted at a rate of 3.5% per annum, in accordance with National Institute for Health and Care Excellence guidelines. Sensitivity analysis will be conducted to assess the impact of assumptions related to the duration and sustainability of the intervention's effects.

In addition to the within-trial CEA, we will undertake an exploratory budget impact and social value assessment. The impact of the intervention will be scaled to a larger, more representative population to illustrate its potential budgetary impact and the cash/noncash releasing benefits of its implementation to commissioners and other stakeholders. Return on investment and net present social/public value (NPSV) will be calculated using the methodologies outlined in the Treasury Green Book. These approaches are designed to support strategic decision-making and will complement the CEA by providing a broader value-for-money perspective for commissioners and stakeholders.

Qualitative Analysis

All interviews and focus groups will be digitally audio recorded and transcribed verbatim. An inductive approach will be taken to identify patterns within the data. Reflexive thematic analysis will be used to analyze the qualitative data, following the steps outlined by Braun and Clarke [40]: familiarization with the data, generating codes, constructing prototype themes, revising themes, defining themes, and producing the report. This is an iterative process that involves returning to the data several times, revising codes and themes, and discussing the themes with the research team.

Integrating data from different stakeholders and the researcher aids the triangulation process, which increases the comprehensiveness and trustworthiness of the final analysis.

Ethical Considerations

The study will be conducted in accordance with the UK Policy Framework for Health and Social Care Research and other applicable guidance. Ethical approval has been obtained from the University of Manchester Research Ethics Committees (Ref: 2024-18620-35607) prior to commencing participant recruitment. All participants will be informed about the study and will provide written informed consent before enrollment. Participation is entirely voluntary. Participants may decline to take part or withdraw from the study at any time, without providing a reason by informing a member of the research team. All data on research participants will be pseudonymized and kept strictly confidential. It will not be possible to identify participants in any publications. The principal investigator will be responsible for preserving and maintaining all data collected during and after the study. Any deviations from the protocol will be put through the appropriate ethical amendment process. The study will be conducted in full conformance with principles of the "Declaration of Helsinki," Good Clinical Practice, and within the laws and regulations of the United Kingdom and the European Union.

Results

Recruitment for this study began in July 2024 and concluded in March 2025, with a total of 202 participants enrolled. A total of 102 participants were randomized into the intervention group (n=102), and 100 were randomized into the control group (n=100). Following protocol publication, data compilation and analysis will be conducted, with results anticipated to be published in 2027. The flow of participants through the trial was recorded in compliance with the CONSORT (Consolidated Standards of Reporting Trials) statement (Figure 1).

XSL•FO
RenderX

Figure 1. Flow diagram of study design. KOKU: Keep-On-Keep-Up.

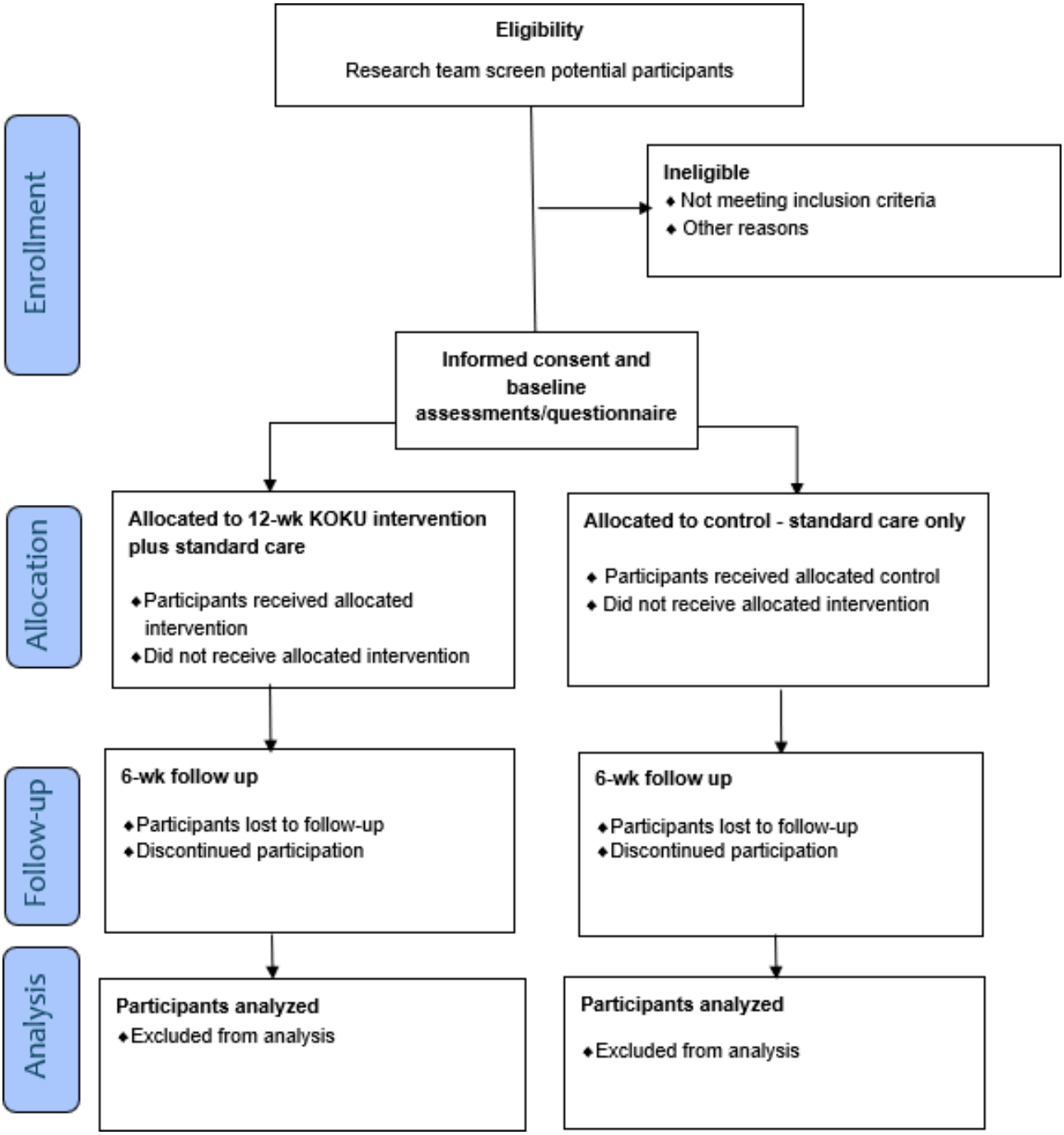

Discussion

Anticipated Findings and Implications

With the growing need for self-manageable, accessible, and scalable interventions to improve function and reduce fall risk, digital health technologies present a promising option for older adults. However, evidence on their clinical effectiveness and cost-effectiveness remains limited. This trial aims to address this gap by evaluating the effectiveness of the KOKU digital program compared with usual care in improving balance, as measured by the BBS [23]. We hypothesize that participants using the KOKU digital program will show better improvements in BBS than those in the control group at 12 weeks.

In addition to balance improvements, this study will also assess other important outcomes such as concerns about falling, falls risk, and health-related quality of life (EQ-5D-5L). Concern about falling is a major factor contributing to reduced physical activity, deconditioning, and increased fall risk. Improvements in confidence may enhance overall well-being and engagement in daily activities. Furthermore, by conducting a cost-effectiveness analysis, this study will provide useful insights into whether digital interventions such as KOKU offer good value for money.

Digital programs such as KOKU may add to and enhance current service provision by offering a personalized, engaging, and home-based exercise option. If found to be effective, this program could provide an alternative or supplement to in-person

XSL•FO
RenderX

programs, which may be particularly helpful for people in remote locations or with limited access to structured exercise classes.

We hope to recruit a diverse sample and have considered a variety of inequalities, including digital inequity, to reduce disparities in access, understanding, and use of KOKU. For example, KOKU was developed with older adults, including those with visual or hearing impairments as well as a range of comorbidities. In addition, we will actively recruit older adults from deprived, socially isolated communities who are digitally marginalized, and participants with impairments, comorbidities, and underserved diverse populations. To support these individuals, we can loan iPads to older adults for the duration of the study, support digital skills specifically on how to use the KOKU program, and continue to work with digital inclusion teams to extend access.

Strengths and Limitations

This study has several strengths. Firstly, it employs an RCT design, minimizing bias and allowing for robust comparisons between the intervention and control groups. Secondly, by using the BBS as the primary outcome measure, the study employs a well-validated and clinically relevant tool to assess balance changes. Balance is widely used as a proxy for falls risk due to its being a key risk factor for falls. As a modifiable risk factor, physical interventions that target balance can directly reduce falls risk. Many falls prevention guidelines recommend balance training with improvements in balance considered a marker of intervention success [7]. Thirdly, the trial incorporates an economic evaluation, which is important for informing policymakers and health care providers about the cost-effectiveness of digital interventions, particularly those tested in real-world conditions. Additionally, by utilizing an intention-to-treat analysis, the study findings will reflect real-life effectiveness in practice.

Despite its strengths, the study has several limitations. In exercise studies, it is very difficult to blind participants or assessors, which can lead to bias. Furthermore, although unlikely, it is possible that participants in the intervention group could share information so that those in the control group have access to the intervention given that some participants were recruited from the same care providers. In addition, larger, fully powered studies are also needed to investigate the clinical and cost-effectiveness of the KOKU intervention with a primary outcome of fall rates.

Conclusions

This trial will provide important evidence on whether a digital strength and balance program can improve balance and related outcomes in older adults compared to usual care. The findings may have important implications for promoting scalable, cost-effective digital solutions to augment current existing service provision and support healthy aging in the community.

Acknowledgments

The research team acknowledges the health and social care professionals and patients across Greater Manchester involved in the development and participation of this intervention and study. Generative artificial intelligence was not used in the preparation or revision of this manuscript.

Funding

This study is funded by National Health Service (NHS) England (Ref: 31088) and supported by the National Institute for Health and Care Research, Applied Research Collaboration-Greater Manchester (NIHR206119, NIHR200174). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The views expressed are those of the authors and not necessarily those of the NHS, the National Institute for Health and Care Research, the Department of Health and Social Care, or its partner organizations.

Data Availability

The research dataset will be stored in an encrypted and secure location to be accessed for secondary analyses or at the request of the participants involved or the research ethics committee.

Authors' Contributions

ES secured funding for the study. ES, CT, PB, and DD contributed to designing the study protocol. AP, BO, MC, and CF are responsible for undertaking participant recruitment and data collection. CS and SB contributingcontributed to data analysis. ES, SA, BO, and CF drafted the manuscript with contributions from all authors. All reviewed and edited the manuscript. Generative AI was not used in the preparation or reviewing of this manuscript.

Conflicts of Interest

ES is the Director of Keep-On-Keep-Up Health, a UK-based nonprofit organization. DAS is a Director of Later Life Training, a not-for-profit training company that provides training in the delivery of the Otago Exercise Program and Falls Management Exercise Program in the United Kingdom, Europe, and Singapore. All other authors declare no conflicts of interest.

Checklist 1

XSL•FO
RenderX

SPIRIT checklist.
[DOC File, 124 KB - resprot_v15i1e78840_app1.doc]

Checklist 2
TIDieR checklist.
[PDF File, 152 KB - resprot_v15i1e78840_app2.pdf]

References

1. Step safely: strategies for preventing and managing falls across the life-course. : World Health Organization; 2021 URL: https://iris.who.int/server/api/core/bitstreams/e7aabc9b-6978-43eb-a6b7-5172d11ce379/content [accessed 2026-01-03]
2. Rubenstein LZ. Falls in older people: epidemiology, risk factors and strategies for prevention. Age Ageing 2006 Sep;35(Suppl 2):ii37-ii41. [doi: 10.1093/ageing/afl084] [Medline: 16926202]
3. Murray CJL, Vos T, Lozano R, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2197-2223. [doi: 10.1016/S0140-6736(12)61689-4] [Medline: 23245608]
4. Health disparities and health inequalities: applying all our health. GOV.UK. 2022 Oct 11. URL: https://www.gov.uk/government/publications/health-disparities-and-health-inequalities-applying-all-our-health/health-disparities-and-health-inequalities-applying-all-our-health [accessed 2026-01-03]
5. Sherrington C, Fairhall N, Wallbank G, et al. Exercise for preventing falls in older people living in the community: an abridged Cochrane systematic review. Br J Sports Med 2020 Aug;54(15):885-891. [doi: 10.1136/bjsports-2019-101512] [Medline: 31792067]
6. Sherrington C, Fairhall NJ, Wallbank GK, et al. Exercise for preventing falls in older people living in the community. Cochrane Database Syst Rev 2019 Jan 31;1(1):CD012424. [doi: 10.1002/14651858.CD012424.pub2] [Medline: 30703272]
7. Montero-Odasso M, van der Velde N, Martin FC, et al. World guidelines for falls prevention and management for older adults: a global initiative. Age Ageing 2022 Sep 2;51(9):afac205. [doi: 10.1093/ageing/afac205] [Medline: 36178003]
8. Hawley-Hague H, Ventre J, Quigley C, Skelton DA, Todd C. Understanding the delivery of the Falls Management Exercise Programme (FaME) across the U.K. J Frailty Sarcopenia Falls 2024 Jun;9(2):96-121. [doi: 10.22540/JFSF-09-096] [Medline: 38835620]
9. Choi NG, Stanmore E, Caamano J, Vences K, Gell NM. A feasibility study of multi-component fall prevention for homebound older adults facilitated by lay coaches and using a tablet-based, gamified exercise application. J Appl Gerontol 2021 Nov;40(11):1483-1491. [doi: 10.1177/0733464821991024] [Medline: 33541199]
10. Delbaere K, Valenzuela T, Lord SR, et al. E-health StandingTall balance exercise for fall prevention in older people: results of a two year randomised controlled trial. BMJ 2021 Apr 6;373:n740. [doi: 10.1136/bmj.n740] [Medline: 33824131]
11. Pettersson B, Lundell S, Lundin-Olsson L, Sandlund M. "Maintaining balance in life'-exploring older adults" long-term engagement in self-managed digital fall prevention exercise. Eur Rev Aging Phys Act 2023 Jul 18;20(1):12. [doi: 10.1186/s11556-023-00322-7] [Medline: 37464299]
12. Stanmore EK, Mavroeidi A, de Jong LD, et al. The effectiveness and cost-effectiveness of strength and balance Exergames to reduce falls risk for people aged 55 years and older in UK assisted living facilities: a multi-centre, cluster randomised controlled trial. BMC Med 2019 Feb 28;17(1):49. [doi: 10.1186/s12916-019-1278-9] [Medline: 30813926]
13. Meekes W, Stanmore EK. Motivational determinants of Exergame participation for older people in assisted living facilities: mixed-methods study. J Med Internet Res 2017 Jul 6;19(7):e238. [doi: 10.2196/jmir.6841] [Medline: 28684385]
14. KOKU Health. 2021. URL: https://kokuhealth.com [accessed 2026-01-03]
15. Skelton D, Dinan S, Campbell M, Rutherford O. Tailored group exercise (Falls Management Exercise -- FaME) reduces falls in community-dwelling older frequent fallers (an RCT). Age Ageing 2005 Nov;34(6):636-639. [doi: 10.1093/ageing/afi174] [Medline: 16267192]
16. Campbell AJ, Robertson MC, Gardner MM, Norton RN, Tilyard MW, Buchner DM. Randomised controlled trial of a general practice programme of home based exercise to prevent falls in elderly women. BMJ 1997 Oct 25;315(7115):1065-1069. [doi: 10.1136/bmj.315.7115.1065] [Medline: 9366737]
17. Navigating the healthcare innovation cycle. : Consortia for Improving Medicine with Innovation & Technology (CIMIT); 2018 URL: https://cimti.cat/wp-content/uploads/2018/04/Navigating-the-Healthcare-Innovation-Cycle_EU.pdf [accessed 2026-01-03]
18. EAST: four simple ways to apply behavioural insights. : Behavioural Insights Ltd; 2017 URL: https://transparencyschool.org/wp-content/uploads/EAST_ExecSummary.pdf [accessed 2026-01-03]
19. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
20. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies

XSL•FO
RenderX

involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015 Nov;19(99):1-188. [doi: 10.3310/hta19990] [Medline: 26616119]
21. Lamb SE, Jørstad-Stein EC, Hauer K, Becker C, Prevention of Falls Network Europe and Outcomes Consensus Group. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. J Am Geriatr Soc 2005 Sep;53(9):1618-1622. [doi: 10.1111/j.1532-5415.2005.53455.x] [Medline: 16137297]
22. Stanmore EK, Oldham J, Skelton DA, O'Neill T, Pilling M, Todd C. Fear-of-falling and associated risk factors in persons with rheumatoid arthritis: a 1 year prospective study. BMC Musculoskelet Disord 2021 Mar 10;22(1):260. [doi: 10.1186/s12891-021-04068-0] [Medline: 33691660]
23. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83(Suppl 2):S7-S11. [Medline: 1468055]
24. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
25. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 7;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
26. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
27. KOKU (Keep On Keep Up). App Store. 2023. URL: https://apps.apple.com/us/app/koku-keep-on-keep-up/id1669009161 [accessed 2026-01-03]
28. Whitney SL, Wrisley DM, Marchetti GF, Gee MA, Redfern MS, Furman JM. Clinical measurement of sit-to-stand performance in people with balance disorders: validity of data for the Five-Times-Sit-to-Stand Test. Phys Ther 2005 Oct;85(10):1034-1045. [doi: 10.1093/ptj/85.10.1034] [Medline: 16180952]
29. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
30. Hoyl MT, Alessi CA, Harker JO, et al. Development and testing of a five-item version of the Geriatric Depression Scale. J Am Geriatr Soc 1999 Jul;47(7):873-878. [doi: 10.1111/j.1532-5415.1999.tb03848.x] [Medline: 10404935]
31. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. J Clin Epidemiol 1993 Feb;46(2):153-162. [doi: 10.1016/0895-4356(93)90053-4] [Medline: 8437031]
32. Kempen G, Yardley L, van Haastregt JCM, et al. The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. Age Ageing 2008 Jan;37(1):45-50. [doi: 10.1093/ageing/afm157] [Medline: 18032400]
33. McGarrigle L, Yang Y, Lasrado R, Gittins M, Todd C. A systematic review and meta-analysis of the measurement properties of concerns-about-falling instruments in older people and people at increased risk of falls. Age Ageing 2023 May 1;52(5):afad055. [doi: 10.1093/ageing/afad055] [Medline: 37211363]
34. Nandy S, Parsons S, Cryer C, et al. Development and preliminary examination of the predictive validity of the Falls Risk Assessment Tool (FRAT) for use in primary care. J Public Health (Oxf) 2004 Jun;26(2):138-143. [doi: 10.1093/pubmed/fdh132] [Medline: 15284315]
35. Brooke J. SUS: a "quick and dirty" usability scale. In: Usability Evaluation in Industry 1996:189-194 URL: https://digital.ahrq.gov/sites/default/files/docs/survey/systemusabilityscale%2528sus%2529_comp%255B1%255D.pdf [accessed 2026-01-03]
36. Sekhon M, Cartwright M, Francis JJ. Acceptability of health care interventions: a theoretical framework and proposed research agenda. Br J Health Psychol 2018 Sep;23(3):519-531. [doi: 10.1111/bjhp.12295] [Medline: 29453791]
37. Alberola C, Walter G, Brau H. Creation of a short version of the User Experience Questionnaire UEQ. i-com 2018 Apr 25;17(1):57-64. [doi: 10.1515/icom-2017-0032]
38. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
39. Donoghue D, Physiotherapy Research and Older People (PROP) group, Stokes EK. How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. J Rehabil Med 2009 Apr;41(5):343-346. [doi: 10.2340/16501977-0337] [Medline: 19363567]
40. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]

Abbreviations

5-STS: five times sit-to-stand
ARC-GM: Applied Research Collaboration–Greater Manchester
BBS: Berg Balance Scale
CEA: cost-effectiveness analysis
CONSORT: Consolidated Standards of Reporting Trials
KOKU: Keep-On-Keep-Up

XSL•FO
RenderX

NHS: National Health Service
QALY: quality-adjusted life years
RCT: randomized controlled trial
REDCap : Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TIDieR: Template for Intervention Description and Replication

Edited by J Sarvestan; submitted 11.Jun.2025; peer-reviewed by A Adepoju, Y Wang, Z Davey; revised version received 26.Nov.2025; accepted 27.Nov.2025; published 05.Feb.2026.

Please cite as:
Stanmore E, Parchment A, Odebiyi B, Bower P, French C, Shi C, Bashir S, Ahmed S, Dowding D, Dumville J, Kislov R, Thompson A, Skelton DA, Clarke M, Todd C
Effectiveness and Cost-Effectiveness of a Digital Falls Prevention Program Versus Usual Care to Improve Balance, Falls Risk, and Function in Older Adults: Protocol for the Keep-On-Keep-Up Randomized Controlled Trial
JMIR Res Protoc 2026;15:e78840
URL: https://www.researchprotocols.org/2026/1/e78840
doi:10.2196/78840

© Emma Stanmore, Amelia Parchment, Bolanle Odebiyi, Peter Bower, Chloe French, Chunhu Shi, Saima Bashir, Saima Ahmed, Dawn Dowding, Jo Dumville, Roman Kislov, Alex Thompson, Dawn A Skelton, Margaret Clarke, Chris Todd. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of a Theory-Based Intervention in Improving Bacillus Calmette-Guérin Uptake and Preventive Practices of Childhood Tuberculosis Among Pregnant Women: Protocol for a Randomized Control Trial

Ahmed Mohamed Dirie[1*], MPhil; Nor Afiah Mohd Zulkefli[1*], PhD; Salmiah Md Said[1*], MCommMed; Ahmed Zaid Fatah Azman[1*], MB ChB; Kassim Abdi Jimale[2], MPhil; Osman Abubakar Fiidow[2*], PhD

[1]Department of Community Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

[2]Faculty of Medicine and Health Sciences, Jamhuriya University of Science and Technology, Mogadishu, Somalia

[*]these authors contributed equally

Corresponding Author:
Nor Afiah Mohd Zulkefli, PhD
Department of Community Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

Abstract

Background: Tuberculosis (TB) in children is one of the most significant public health crises in Somalia. This issue is aggravated by the fact that only 36.7% of children aged 12 to 23 months receive the Bacillus Calmette–Guérin (BCG) vaccine in Somalia, which helps prevent TB-disseminated diseases. Among the major factors that contribute to poor BCG uptake and TB prevention practices are the lack of maternal knowledge, negative attitude, and poor self-efficacy toward the BCG vaccine and TB prevention practices. As such, pregnant women play a vital role in ensuring timely, routine BCG vaccination for their newborns and adherence to TB prevention practices.

Objective: This study aims to develop, implement, and evaluate the impact of health education intervention using the information-motivation-behavioral (IMB) skills model to improve knowledge, attitudes, and self-efficacy regarding BCG uptake and TB prevention among pregnant women in Banadir Hospital.

Methods: This single-blind randomized controlled trial enrolled a sample of 370 pregnant women recruited at Banadir Hospital in Mogadishu, Somalia. Eligible participants will be randomized to an intervention group receiving an IMB-based health education and to a waiting list control group, in a 1:1 ratio. Outcome assessments will be conducted at baseline, 2-month follow-up, and 4-month follow-up. The primary outcomes are BCG vaccine uptake and TB prevention practices. Secondary outcomes include knowledge, attitudes, and self-efficacy related to the BCG vaccine and TB prevention practices. The IMB-based health education intervention program consists of 6 sessions, with 1 group per session, and with each session containing 30 participants. The effects of the intervention will be assessed by handing out the same self-administered questionnaires at baseline, 2-month postintervention, and 4-month postintervention.

Results: A total of 370 pregnant women were recruited at baseline in November 2021, with 185 assigned to the intervention group and 185 assigned to the control group. In January 2022, the 185 pregnant women enrolled in the intervention program and completed the sessions by February 2022. Data collection for the 2-month and 4-month postintervention assessments was completed in June 2022. The findings of this study will be reported by the beginning of 2026.

Conclusions: The developed health education intervention module in this study has the potential to be adopted and included as part of routine antenatal care services. Its implementation could effectively raise awareness among pregnant women in Somalia regarding the importance of BCG vaccination and TB prevention practices, ultimately mitigating childhood mortality rates associated with TB-disseminated diseases in the country.

Trial Registration: Pan African Clinical Trials Registry PACTR202105554130782; https://pactr.samrc.ac.za/TrialDisplay.aspx?TrialID=15863

International Registered Report Identifier (IRRID): DERR1-10.2196/68088

(JMIR Res Protoc 2026;15:e68088) doi:10.2196/68088

XSL•FO
RenderX

KEYWORDS

childhood tuberculosis; prevention practices; Bacillus Calmette–Guérin; BCG; randomized controlled trial; information-motivation-behavioral (IMB) skills model

Introduction

Tuberculosis (TB) remains a prevalent health issue worldwide, with over 10 million new cases annually [1]. TB cases are also reported in children younger than 15 years and are referred to as childhood TB. Moreover, children younger than 5 years are particularly vulnerable to developing active TB disease [2]. In 2011, the World Health Organization reported its first estimate of annual global childhood TB, which amounted to 490,000 cases [3]. In 2022, approximately 1.25 million children and young adolescents were diagnosed with TB, with 47% of these cases occurring in children younger than 5 years. This age group represents 12% of the overall TB incidence of 10.6 million [4]. However, approximately 51% of these cases were either undiagnosed or not reported to national TB programs. The gap is even more pronounced among children younger than 5 years, with 58% of estimated cases being missed, compared to 30% for individuals aged 15 years and older [3].

The risk factors contributing to the development of TB disease among children have received increasing attention over the years. Contact with known adult TB cases is a significant risk factor for both children and adults [5,6]. A study conducted in Guinea-Bissau found that children younger than five years living in the same household as an adult with TB faced a markedly higher risk of death, with the risk being up to 8 times higher for those whose mothers were affected by the disease [7]. Other important risk factors include lack of BCG vaccination, severe malnutrition, and knowledge and attitude toward TB [8-10].

Somalia is classified as one of the least developed countries and faces a significant burden of TB, particularly multidrug-resistant TB [11]. In 2023, the TB incidence in Somalia was reported at 246 cases per 100,000 population [12]. Despite the heightened vulnerability of children due to malnutrition and limited access to health care, only about 56.4% of TB cases are detected, resulting in many children going without the necessary treatment [13]. Furthermore, according to the Somali National Strategic Plan 2020 - 2024 for Tuberculosis, 22% of TB cases occur in children younger than 15 years [14].

Somalia, as a TB-endemic country, faces significant challenges, particularly low BCG vaccine uptake, limited knowledge of childhood TB, and negative attitudes toward the disease. According to the Somali Health Demographic Survey conducted in 2020, only 36.7% of children aged 12 to 23 months received the BCG vaccine, marking one of the lowest coverage rates in Africa [15]. Similarly, community-based studies in Somalia revealed that only 4.7% of participants correctly identified bacteria as the cause of TB, while just 44.7% of participants were aware of the importance of covering their nose and mouth when coughing and sneezing to prevent TB transmission [16,17]

Previous intervention studies have demonstrated the positive impact of health education on improving BCG uptake and TB prevention practices [18,19]. Similarly, health theories are important for understanding health behaviors, designing effective interventions, and evaluating program success. This study suggests using the information-motivation-behavioral (IMB) skills model construct over theories such as the health belief model and the theory of planned behavior, which predominantly center on perceptions and beliefs concerning health threats but may underestimate the crucial motivational aspect in behavior change. The IMB skills model emphasizes 3 essential constructs for promoting behavior change: providing relevant information regarding the desired behavior (which alone is insufficient for behavior change), fostering appropriate motivation to engage in the desired behaviors, and ensuring that individuals possess the necessary skills and competencies to effectively perform the desired behaviors [20].

However, no studies were identified during the literature search that assessed health education interventions specifically related to a model concerning BCG vaccination and TB prevention in Somalia. Therefore, this study proposes to develop, implement, and assess the effects of an IMB-based health education intervention on BCG uptake and TB prevention practices among pregnant women attending the antenatal care (ANC) clinic at Banadir Maternal and Child Hospital, Mogadishu, Somalia.

Methods

Study Location

This study selected Banadir Maternal and Child Hospital, which is a public referral hospital, as the study location. This hospital is located in Mogadishu, the capital city of Somalia. Also known as Xamar, it is the most populous city in Somalia, with an estimated 2 million residents [21], and stretches along the southern coast of Somalia on the Horn of Africa. This particular hospital was chosen due to its strategic geographical location at the center of the city and its considerable size and high patient load, making it one of the largest hospitals in the area. The hospital comprises 2 main departments: maternity and pediatrics. Under the Maternity Department, the hospital operates an ANC clinic from Saturday to Thursday, providing support to expectant mothers throughout their pregnancy. Other essential services that are available at this hospital include the Diarrhea Treatment Center, Blood Transfusion Unit, Emergency Unit, TB Center, and Research and Training Center.

Study Design

A single-blind randomized controlled trial will be conducted in this study, with only the participants blinded to their assigned group. The participants will be randomly categorized into the intervention group and control group. The study protocol will be reported in accordance with the Standard Protocol Items: Recommendations for Intervention Trial (SPIRIT) 2013 Guideline [22] (Figure S1, Multimedia Appendix 1).

Study Duration

Planning of the study protocol began in January 2021, and data collection was completed in June 2022.

XSL•FO
RenderX

Study Population and Study Setting

The study population consists of pregnant women attending the Banadir Maternal and Child Hospital, Mogadishu, for ANC visits. The inclusion criteria include mothers with a gestational age range of 25 to 33 weeks, plans for postnatal follow-up at the hospital, and residence in Mogadishu, including internally displaced persons. The exclusion criteria include mothers who visited the hospital for other health conditions, those of non-Somali nationality, and those who had previous health education training related to immunization and TB prevention practices.

Sample Size

The 2-sample mean formula was used to determine the sample size for this study. This formula considers the mean and SD of TB preventive practices from a previous health education program conducted by John [19], where a SD of 4.40, an intervention group mean of 20.71, and a control group mean of 18.45 were used in the calculation. To ensure adequate statistical power, several adjustments were made for the estimated eligibility (90%) and attrition rate (30%), with a 95% significance level and 80% desired power. The anticipated high attrition rate is primarily due to the impact of the COVID-19 pandemic, which created significant barriers to participation in the study. Based on this formulation, the total sample size for this study was 370 participants, with an equal number of participants in both the intervention and control groups.

Recruitment

Potential participants were recruited over a 4-week period from the ANC clinic at Banadir Maternal and Child Hospital. To ensure an organized and systematic selection, their names were sorted into the ANC electronic register in chronological order. After obtaining informed consent from all participants who agreed to take part in the study, their eligibility was assessed by applying the inclusion and exclusion criteria checklist. The recruitment process is summarized in Figure S2 (Multimedia Appendix 1).

Randomization, Allocation Concealment, and Blinding

Randomization was performed at the individual level, with pregnant women as the sampling unit. Ensuring an equal allocation ratio (1:1), participants were randomly assigned into the intervention group, which received the IMB-based health education intervention, or the control group, which was placed on a waiting list. The randomization process was conducted using computer-generated random numbers using a web-based calculator (GraphPad; GraphPad Software) by hospital staff who were independent of all other research processes. The generated sequences were placed in sealed opaque envelopes. Subsequently, another hospital staff member, who was also independent of the research team, was responsible for opening the envelopes sequentially and assigning participants to either the intervention or control group.

Intervention

Intervention Development and Validation

The intervention modules were developed based on relevant sources from the World Health Organization, the Ministry of Health of Somalia [23], and publications and guidelines from previous studies [24]. As presented in Table 1, the IMB components, namely information, motivation, and behavior skills, were used to develop the intervention program. Each module's content was examined by a panel of experts consisting of 6 public health specialists and pediatricians from Universiti Putra Malaysia, the Ministry of Health of Somalia, and Bandir Hospital. To ensure effective delivery of the module content to participants, 2 female intern students provided a single, intensive training session lasting approximately 3 hours. To further prepare the interns, a follow-up session was scheduled a day before the intervention sessions began. Both interns had backgrounds in medicine and have demonstrated competency in Somali communication, which enhanced their ability to engage effectively with participants.

XSL•FO
RenderX

Table . Summary of intervention program content and delivery.

Theory construct module	Health education content	Delivery strategy	Estimated duration
Module 1: information	Basic knowledge of childhood TB[a], including cause, transmission, signs and symptoms, severity, and complications.	Lecture using PowerPoint presentation, question-and-answer sessions.	30 min
Module 2: information	Benefits, schedule, and side effects of the BCG[b] vaccine; respiratory hygiene (cough etiquette); and environmental control measures.	Lectures using PowerPoint presentations, question-and-answer sessions.	30 min
Module 3: motivation	Importance of the BCG vaccine, respiratory hygiene (cough etiquette), and environmental control measures; religious and cultural leaders addressing the benefits, safety, and halal status of the BCG vaccine, and normalizing face mask use.	Interactive discussion, brainstorming sessions, video presentations, and lectures.	1 h
Module 4: behaviorSkills	Self-efficacy evaluation related to receiving the BCG vaccine and childhood TB prevention practices; problem-solving focused on challenges and solutions; reinforcement through a 1-month follow-up call for mothers in the intervention group to reinforce key intervention points and prioritize the administration of the BCG vaccine for their newborns.	Role play (video), interactive session, brainstorming exercise, and a 1-month follow-up call to reinforce.	1 h and 30 min

[a]TB: tuberculosis.

[b]BCG: Bacillus Calmette–Guérin.

Pretesting of the questionnaire was conducted among 36 pregnant women in Banadir Hospital who were not part of the study to assess the reliability of the questionnaire. Reliability was carried out using Cronbach α to assess for internal consistency. The Cronbach α coefficients for knowledge, attitudes, and prevention practices were 0.774, 0.829, and 0.689, respectively. The results were within the acceptable range of 0.60 to 0.95 [25].

Intervention Structure and Delivery

The intervention consists of 4 modules. The first module utilizes a combination of lectures, PowerPoint presentations, and a question-and-answer session to provide a foundational understanding of childhood TB. The duration of this module is estimated to be 30 minutes. The second module concentrates on understanding the BCG vaccine and TB prevention practices. The delivery method is similar to that of the first module and takes approximately 30 minutes. The third module explores participants' attitudes toward the BCG vaccine and TB prevention practices. Various strategies are used, including video presentations featuring traditional and religious leaders, interactive sessions, and brainstorming exercises. This module is estimated to take 1 hour to be completed. The fourth and final module aims to enhance participants' self-efficacy regarding the BCG vaccine and TB prevention practices, as well as equip participants with useful skills to make informed decisions and overcome potential challenges. This module incorporates a diverse range of strategies, including role-play videos, interactive sessions, and brainstorming exercises. The duration of this module is estimated to last 1 hour and 30 minutes.

Participants in the intervention group were divided into 6 small groups for health education sessions. There were 6 intervention sessions, with 1 group per session; each session accommodated 30 participants, except for the first session, which had 35 participants. The language of delivery was Somali, and participants received US $2.00 for transportation. The delivery of the modules was closely monitored and supervised by the researchers, who provided necessary feedback to the facilitators to ensure strict adherence to the protocol without deviation. After the completion of the sessions, mothers participating in the intervention received a follow-up call 1 month later to reinforce the key points discussed and to highlight the importance of prioritizing the administration of the BCG vaccine for their newborns.

Control Group

The control group was assigned to a waiting list status throughout the intervention program. Participants received their usual care during their visits to the ANC clinic and had access to general health education content through television programming. The hospital does not provide personalized health education for antenatal mothers. To avoid potential contamination or cross-group influence, participants in both groups were given different follow-up dates. Additionally, participants were clearly instructed on the importance of not discussing the intervention sessions with others.

XSL•FO
RenderX

Data Collection

Data were collected using paper questionnaires at 3 distinct time intervals: baseline, 2-months postintervention, and 4-months postintervention. Participants were asked to complete a consent form before participating and were informed that their information would remain confidential and used solely for research purposes. The recorded data were securely stored in protected files, and individual responses were not reported publicly; instead, findings were presented in aggregate form. The random assignment took place on the same day that participants were selected for the study, after the collection of baseline data. To facilitate the data collection process, each participant was given a study follow-up card and was requested to carry them during each follow-up visit to the hospital.

Quality Control

Several steps were taken to ensure the validity and reliability of the questionnaire. First, content and face validity were evaluated to ensure that the wording, ordering, and language were appropriate and well understood. The analysis categorized items as good, average, or poor; language clarity was assessed as clear, average, or confusing; and the appropriateness of construct measurement was classified as good, average, or poor. Feedback from experts and participants led to important modifications, including rewording several items for clarity and removing 3 items deemed irrelevant to the study objectives.

Additionally, factor analysis was conducted using SPSS software (version 27.0; IBM Corp) on a separate sample set of 134 participants who were not part of the main study but were recruited from the same hospital to assess sampling adequacy. Cohen's κ and intraclass correlation coefficient values were also utilized to evaluate the reliability of the questionnaire. All values exceeded 0.60, which is considered acceptable [25].

Outcome Assessment

Table 2 presents the primary and secondary outcome assessments, including the timing of each measurement.

XSL•FO
RenderX

Table . Outcome measurements.

Outcome measure	Definition	Time of measurement		
		T0 (baseline)	T1 (2-month follow-up)	T2 (4-month follow-up)
Primary outcomes				
BCG[a] uptake	Refers to receiving 1 dose of the BCG vaccine at birth. Infants who receive the vaccine after 6 weeks will be considered delayed.	No	Yes	Yes
TB[b] preventive practices	Refers to the change in total practice scores of actions taken by participants to reduce the risk of TB infection and transmission, including respiratory hygiene (cough etiquette) and environmental control, between baseline, 2-month follow-up, and 4-month follow-up.	Yes	Yes	Yes
Secondary outcomes				
Knowledge	Refers to the total score change of participants' knowledge regarding BCG vaccine benefit, schedule, and side effect, as well as knowledge of preventive practices (respiratory hygiene/cough etiquette and environmental control) between baseline, 2-month follow-up, and 4-month follow-up.	Yes	Yes	Yes
Attitude	Refers to the change in total attitude scores toward the BCG vaccine and preventive practices between baseline, 2-month follow-up, and 4-month follow-up.	Yes	Yes	Yes
Self-efficacy	Refers to the change in total scores reflecting participants' confidence in their ability to obtain and receive the BCG vaccine, as well as their confidence in engaging in TB prevention practices between baseline, 2-month follow-up, and 4-month follow-up.	Yes	Yes	Yes

[a]BCG: Bacillus Calmette–Guérin.

[b]TB: tuberculosis.

Data Analysis

Collected data were entered into SPSS software (version 27.0; IBM Corp) before proceeding with data cleaning procedures. Both statistical and graphical normality tests were applied to continuous variables, such as age, knowledge, attitude, and self-efficacy. At baseline, the comparison of variables between groups used appropriate statistical measures based on the distribution of data. For normally distributed data, results are reported as mean (SD), and comparisons were conducted using an independent *t* test. For non-normally distributed data, results were reported as median (IQR), and comparisons were conducted using the chi-square test and the Mann-Whitney *U* test.

To evaluate the main effect of the IMB-based intervention modules on the response variables over time, as well as the group-time interactions, repeated-measures ANOVA was used for normally distributed data. For non-normally distributed data, generalized linear mixed models were applied to account for individual-level variability among participants, or generalized

XSL•FO
RenderX

estimating equations were used when focusing on overall population-level trends. Intention-to-treat analyses were conducted separately for the intervention and control groups, and missing data were addressed using multiple imputations. Results were reported as differences between groups with 95% CIs. Statistical significance in this study was set at α=.05.

Ethical Considerations

The trial protocol was approved by the Ethics Committee for Research Involving Human Subjects at the Universiti Putra Malaysia (UPM) (UPM/TNCPI/RMC/JKEUPM/1.4.18.2), the Ethical Review Committee of the Ministry of Health of Somalia (MOHHS/DGO/1166/Feb2024), and Banadir Hospital (BH/30594/4/22). The study was registered on May 24, 2022, with the Pan African Clinical Trials Registry and assigned a unique identification number PACTR202105554130782.

A participant information sheet in the Somali language was provided in written form by trained enumerators to each participant, from whom informed consent was obtained. Participant information and identities remained confidential and used solely for the purpose of this research. Participation in the study was voluntary, and participants had the right to withdraw at any point during the intervention. Participants received compensation of USD 2.00 per follow-up visit to cover transportation costs.

Any significant protocol modifications, such as alterations to eligibility criteria, outcomes, or analysis methods, will be promptly communicated to both the Pan African Clinical Trials Registry and the Jawatankuasa Etika Universiti Putra Malaysia.

Results

Recruitment began on November 25, 2021, and by December 21, 2021, a total of 370 participants had been enrolled in the trial. The intervention sessions were completed on February 4, 2022. Follow-up data collection was completed on June 14, 2022. Results are expected to be published by the end of 2025.

Discussion

Expected Findings and Implications

Health theories play a crucial role in public health by providing frameworks for understanding health behaviors, designing effective interventions, and evaluating improvements of outcomes. This study will use IMB-based theory constructs to improve knowledge, attitudes, and self-efficacy among antenatal mothers, ultimately aiming to improve BCG uptake and childhood TB prevention practices. However, previous studies in Somalia have primarily been observational in design and have lacked a theoretical framework, which limits an inclusive understanding of the factors influencing behavior change. Therefore, investigating this gap is essential for improving outcomes for mothers and children.

This study is the first to examine the effectiveness of an IMB-based health education intervention in improving BCG uptake and childhood TB prevention practices by comparing intervention and control groups. In particular, the findings of this study may highlight the potential for improving maternal knowledge, attitudes, and self-efficacy regarding BCG vaccination and childhood TB prevention practices in Mogadishu, Somalia. Similarly, it is expected that pregnant women in the intervention group will demonstrate higher TB prevention practices and increased BCG uptake for their children compared with those in the control group. Previous intervention studies supported the effectiveness of health education interventions in improving BCG uptake and TB prevention practices [26-28]. Finally, the findings are anticipated to contribute to improving TB prevention efforts in Somalia, where low BCG coverage, poor maternal knowledge, and fragile health systems continue to pose major public health challenges.

Strengths and Limitations

This is the first IMB-based health education intervention designed to improve BCG uptake and childhood TB prevention practices. Additionally, its randomized controlled design, single-blind allocation, theoretical framework, and use of validated tools enhance internal validity and reduce the risk of bias. Furthermore, conducting the trial at Banadir Hospital, a referral facility for maternal and child health, allows for the recruitment of a diverse population of pregnant women, including internally displaced persons, thereby enhancing the generalizability of outcomes.

Nevertheless, this study has several limitations. Despite blinding, cross-group information sharing among participants cannot be entirely ruled out, which may influence the findings. Additionally, outcomes are largely based on self-reported measures, introducing the potential for reporting bias. The inability to blind outcome assessors may also introduce bias in the evaluation of findings. Furthermore, because the trial is facility-based, the findings may not be fully generalizable to women who lack access to health care. Finally, external factors, such as sociopolitical instability or disruptions in the health system, may independently affect BCG uptake regardless of the intervention.

Conclusion

This study is a single-blinded randomized controlled trial involving 370 mothers at Banadir Hospital in Mogadishu, Somalia. Participants were randomized into either an intervention group receiving the IMB-based health education intervention or a waitlist control group. It is anticipated that pregnant women in the intervention program will show significant improvements in the primary outcomes of BCG vaccine uptake and TB prevention practices. Additionally, secondary outcomes, including knowledge, attitudes, and self-efficacy related to childhood TB prevention and the BCG vaccine, are expected to improve compared with those in the control group.

The effectiveness of the IMB-based health education modules developed in this intervention may have the potential to be adopted in routine health education programs for pregnant women during their ANC visits in Mogadishu, Somalia.

XSL•FO
RenderX

Acknowledgments

We would like to express our gratitude to all contributors for their invaluable support and insights throughout the development of this research protocol.

Funding

None of the authors received any form of funding for this work.

Data Availability

The data of this study are available upon request from the corresponding author, and it is subject to approval from the Universiti Putra Malaysia.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Figures S1-S2
[DOCX File, 32 KB - resprot_v15i1e68088_app1.docx]

References

1. Chakaya J, Khan M, Ntoumi F, et al. Global Tuberculosis Report 2020 - reflections on the global TB burden, treatment and prevention efforts. Int J Infect Dis 2021 Dec;113 Suppl 1(Suppl 1):S7-S12. [doi: 10.1016/j.ijid.2021.02.107] [Medline: 33716195]
2. Basu Roy R, Whittaker E, Seddon JA, Kampmann B. Tuberculosis susceptibility and protection in children. Lancet Infect Dis 2019 Mar;19(3):e96-e108. [doi: 10.1016/S1473-3099(18)30157-9] [Medline: 30322790]
3. Global tuberculosis control: WHO report 2011. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/44728/9789241564380_eng.pdf?sequence=1 [accessed 2026-01-30]
4. Verkuijl S, Bastard M, Brands A, et al. Global reporting on TB in children and adolescents: how far have we come and what remains to be done? IJTLD Open 2024 Jan;1(1):3-6. [doi: 10.5588/ijtldopen.23.0529] [Medline: 38919413]
5. Gessner BD, Weiss NS, Nolan CM. Risk factors for pediatric tuberculosis infection and disease after household exposure to adult index cases in Alaska. J Pediatr 1998 Mar;132(3 Pt 1):509-513. [doi: 10.1016/s0022-3476(98)70029-0] [Medline: 9544910]
6. Russell FM, Carapetis JR, Tikoduadua L, et al. Invasive pneumococcal disease in Fiji: clinical syndromes, epidemiology, and the potential impact of pneumococcal conjugate vaccine. Pediatr Infect Dis J 2010 Sep;29(9):870-872. [doi: 10.1097/INF.0b013e3181ec7ae2] [Medline: 20622710]
7. Gomes VF, Andersen A, Wejse C, et al. Impact of tuberculosis exposure at home on mortality in children under 5 years of age in Guinea-Bissau. Thorax 2011 Feb;66(2):163-167. [doi: 10.1136/thx.2010.141309] [Medline: 21148136]
8. Soysal A, Millington KA, Bakir M, et al. Effect of BCG vaccination on risk of mycobacterium tuberculosis infection in children with household tuberculosis contact: a prospective community-based study. The Lancet 2005 Oct;366(9495):1443-1451. [doi: 10.1016/S0140-6736(05)67534-4]
9. Siddalingaiah N, Chawla K, Nagaraja SB, Hazra D. Risk factors for the development of tuberculosis among the pediatric population: a systematic review and meta-analysis. Eur J Pediatr 2023 Jul;182(7):3007-3019. [doi: 10.1007/s00431-023-04988-0] [Medline: 37130994]
10. Singh M, Mynak ML, Kumar L, Mathew JL, Jindal SK. Prevalence and risk factors for transmission of infection among children in household contact with adults having pulmonary tuberculosis. Arch Dis Child 2005 Jun;90(6):624-628. [doi: 10.1136/adc.2003.044255] [Medline: 15908630]
11. Asuke S, Isah HO, Jimoh AO, Achema T. Predictors of tuberculosis knowledge among mothers of under-fives, seen at Bingham University Teaching Hospital, Jos Nigeria. J Infect Dev Ctries 2022 Apr 30;16(4):691-697. [doi: 10.3855/jidc.13845] [Medline: 35544632]
12. Ali MK, Karanja S, Karama M. Factors associated with tuberculosis treatment outcomes among tuberculosis patients attending tuberculosis treatment centres in 2016-2017 in Mogadishu, Somalia. Pan Afr Med J 2017;28:197. [doi: 10.11604/pamj.2017.28.197.13439] [Medline: 29610635]
13. The national strategic plan for tuberculosis control 2024-2026 national leprosy and tuberculosis control program. : Somalia Ministry of Health; 2023 URL: https://moh.pl.so/wp-content/uploads/2024/08/About-Nationa-TB-Program.pdf
14. Somali national tuberculosis strategic plan 2020–2024. : Somalia Ministry of Health; 2020 URL: https://moh.pl.so/wp-content/uploads/2024/08/About-Nationa-TB-Program.pdf
15. Somalia National Bureau of Statistics, Federal Government of Somalia The Somali Health and Demographic Survey 2020.: Mogadishu: SNBS; 2020. URL: https://nbs.gov.so/ [accessed 2026-01-30]

XSL•FO
RenderX

16. Yusuf-Isleged M. Community knowledge, attitude and practices towards tuberculosis among household heads in Mogadishu, Somalia. Integr J Res Arts Humanities 2022 Sep 23;2(5):74-84. [doi: 10.55544/ijrah.2.5.13]
17. Nur AY. Factors influencing delay in seeking tuberculosis treatment in Belet-Weyne District, Somalia. : University of the Western Cape; 2008 URL: https://api.semanticscholar.org/CorpusID:55262575
18. Oche MO. An assessment of the impact of health education on maternal knowledge and practice of childhood immunization in Kware, Sokoto State. J Public Health Epidemiol 2011;3(10):440-447 [FREE Full text]
19. John S. Effect of health education program on knowledge, attitude, and practice towards tuberculosis among nomads in Adamawa State, Nigeria. Texila Int J Public Health 2024;12(3). [doi: 10.21522/TIJPH.2013.12.03.Art067]
20. Fisher JD, Fisher WA, Williams SS, Malloy TE. Empirical tests of an information-motivation-behavioral skills model of AIDS-preventive behavior with gay men and heterosexual university students. Health Psychol 1994;13(3):238-250. [doi: 10.1037//0278-6133.13.3.238]
21. Elmi AA, Ali FN, August PR. Mogadishu: city report. 2024. URL: https://www.african-cities.org/wp-content/uploads/2024/08/ACRC_Working-Paper-20_August-2024.pdf [accessed 2026-01-30]
22. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346:e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
23. UNICEF. Knowledge, attitudes and practices in relation to immunization of children in Serbia.: UNICEF; 2018. URL: https://www.unicef.org/serbia/media/9146/file/Knowledge,%20attitudes%20and%20practices.pdf [accessed 2026-01-30]
24. Advocacy, communication and social mobilization for TB control: a guide to developing knowledge, attitude and practice surveys. : World Health Organization: Stop TB Partnership; 2008 URL: https://www.who.int/publications/i/item/9789241596176 [accessed 2026-01-31]
25. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282. [Medline: 23092060]
26. Bisallah CI, Rampal L, Lye MS, et al. Effectiveness of health education intervention in improving knowledge, attitude, and practices regarding Tuberculosis among HIV patients in General Hospital Minna, Nigeria - a randomized control trial. PLoS ONE 2018;13(2):e0192276. [doi: 10.1371/journal.pone.0192276] [Medline: 29470530]
27. Mohammadi K, Tavafian SS, Ghofranipoor F, Amin-Shokravip F. Health education program and tuberculosis preventive behaviors. Zahedan J Res Med Sci 2012 [FREE Full text]
28. Balogun M, Sekoni A, Meloni ST, et al. Trained community volunteers improve tuberculosis knowledge and attitudes among adults in a periurban community in southwest Nigeria. Am J Trop Med Hyg 2015 Mar;92(3):625-632. [doi: 10.4269/ajtmh.14-0527] [Medline: 25510722]

Abbreviations

ANC: antenatal
BCG: Bacillus Calmette-Guérin
IMB: information-motivation-behavior
RCT: randomized controlled trial
TB: tuberculosis

Edited by J Sarvestan; submitted 29.Dec.2024; peer-reviewed by AB Fisker, JL Karim; revised version received 08.Sep.2025; accepted 15.Sep.2025; published 09.Feb.2026.

Please cite as:
Dirie AM, Mohd Zulkefli NA, Md Said S, Azman AZF, Jimale KA, Fiidow OA
Effectiveness of a Theory-Based Intervention in Improving Bacillus Calmette-Guérin Uptake and Preventive Practices of Childhood Tuberculosis Among Pregnant Women: Protocol for a Randomized Control Trial
JMIR Res Protoc 2026;15:e68088
URL: https://www.researchprotocols.org/2026/1/e68088
doi:10.2196/68088

© Ahmed Mohamed Dirie, Nor Afiah Mohd Zulkefli, Salmiah Md. Said, Ahmed Zaid Fatah Azman, Kassim Abdi Jimale, Osman Abubakar Fiidow. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Hand and Arm Bimanual Intensive Therapy Including Lower Extremities (HABIT-ILE@home) for Adults With Chronic Stroke: Protocol for a Randomized Controlled Trial

Merlin Somville[1]; Zélie Rosselli[1]; Edouard Ducoffre[1,2]; Carlyne Arnould[2], PhD; Yannick Bleyenheuft[1*], Prof Dr; Geoffroy Saussez[1,2*], PhD

[1]Institute of NeuroSciences (IoNS), UCLouvain, Avenue Emmanuel Mounier 83, Bruxelles, Belgium

[2]CeREF-Santé, Haute École Louvain en Hainaut, Montignies-sur-Sambre, Belgium

[*]these authors contributed equally

Corresponding Author:
Yannick Bleyenheuft, Prof Dr
Institute of NeuroSciences (IoNS), UCLouvain, Avenue Emmanuel Mounier 83, Bruxelles, Belgium

Abstract

Background: Intensive therapies based on motor skill learning have been widely used in stroke rehabilitation for improving upper extremity abilities, demonstrating significant improvements in arm function and daily life activities. Based on the same therapeutic principles of motor skill learning, hand and arm bimanual intensive therapy including lower extremities (HABIT-ILE) was developed focusing on bimanual coordination and constant concomitant stimulation of trunk control and lower extremities. However, the implementation of such high-dosage interventions in stroke rehabilitation might face barriers due to limited accessibility and high resource requirements. Delivering HABIT-ILE therapy at home (HABIT-ILE@home) via telerehabilitation may reduce logistical barriers while maintaining efficacy. In addition, the added value of a 9-week specific follow-up program will be tested after high dosage interventions.

Objective: The first randomized controlled trial (RCT1) aims to evaluate the noninferiority of a high-dosage HABIT-ILE@home program compared to its on-site counterpart. The second randomized controlled trial (RCT2) aims to test the superiority of a 9-week specific HABIT-ILE@home follow-up vs a nonspecific home program.

Methods: A total of 48 adults with chronic stroke will be randomized to either HABIT-ILE@home or HABIT-ILE on-site group (RCT1, 65 hours over 2 weeks). HABIT-ILE@home will follow the same principles as HABIT-ILE on-site but will be delivered by caregivers with remote supervision by trained therapists and the use of a dedicated telerehabilitation device to facilitate intervention delivery and remote monitoring (ie, REAtouch Lite; Axinesis). All participants will then participate in a HABIT-ILE@home follow-up program or nonspecific follow-up (RCT2, 45 hours over 9 weeks). Primary outcomes will be the change in Fugl-Meyer Assessment (FMA), while secondary outcomes include feasibility and adherence questionnaires, upper and lower extremity motor function assessments, daily activities, and quality of life questionnaires. Assessments will be performed before (T0) and after (T1) the 2 weeks of high dosage intervention, followed by an assessment after the 9-week follow-up (T2).

Results: Recruitment for the trial started in March 2023 and ended in March 2025, and data collection has been completed for this study. Data analysis is planned to start early 2026; we expect to submit the results for publication in spring 2026.

Conclusion: This study will provide evidence on the feasibility and efficacy of delivering HABIT-ILE through a home-based telerehabilitation model for adults with chronic stroke. Demonstrating noninferiority of HABIT-ILE@home compared to on-site therapy would support wider accessibility to intensive rehabilitation while reducing logistical and human resource constraints. Additionally, showing the added benefit of a structured follow-up could emphasize the importance of continuity of care to sustain and enhance motor recovery after intensive interventions.

Trial Registration: ClinicalTrials.gov NCT05727111; https://clinicaltrials.gov/study/NCT05727111

International Registered Report Identifier (IRRID): DERR1-10.2196/87035

(JMIR Res Protoc 2026;15:e87035) doi:10.2196/87035

KEYWORDS

chronic stroke; intensive rehabilitation; motor function; motor skill learning; bimanual therapy; HABIT-ILE; home-based training; telerehabilitation; caregiver involvement; randomized controlled trial; hand and arm bimanual intensive therapy including lower extremities

XSL•FO
RenderX

Introduction

Background

The global lifetime risk of stroke in the adult population is 24.9% [1]. Stroke is the second leading cause of death and the primary etiology of motor disability in adults worldwide [2]. After a stroke, patients experience variable acute symptoms depending on the type, location, and size of the brain lesions: hemiparesis, which often causes partial or complete inability to use one upper limb, spasticity, loss of independent walking, chronic pain, alterations of sensations, and visuospatial neglect [3,4]. It is important to note that some of these symptoms may become chronic. The long-term consequences of the pathology, particularly the motor impairments, can lead to difficulties in daily activities and a lower quality of life [5,6].

Rehabilitation plays an important role in reducing the impact of symptoms and in improving the quality of life of stroke patients. A wide variety of rehabilitation interventions are available, including strength or manual dexterity training, neuromuscular stimulation, balance exercises, walking and task training, and stretching of the affected side [7]. The most effective interventions identified in the literature are those based on the principles of motor skill learning, such as constraint-induced movement therapy (CIMT), hand-arm bimanual intensive training, bilateral arm training, and high dose repetitive task practice [8-11]. These therapies have demonstrated significant improvement in upper extremity functions and activities of daily living [12-15]. More recently, hand and arm bimanual intensive therapy including lower extremities (HABIT-ILE) has been developed with a focus on bimanual coordination training with stimulation of postural control and lower extremities [16]. In adults with stroke, the implementation of a HABIT-ILE training is under investigation through a randomized controlled trial with expectations of similar improvements [17]. Although no results on the efficacy of HABIT-ILE in stroke patients are currently available, improvements similar to those observed in upper extremities with other intensive therapies based on motor skill learning are expected.

HABIT-ILE is usually organized through a 2-week on-site camp. Although this high therapeutic dosage is one of the ingredients necessary for its efficacy, it can also be considered an obstacle to the implementation of HABIT-ILE in clinical practice. As the number of locations offering HABIT-ILE therapy is currently limited, patients and their families often have to travel long distances to reach the therapy site and find accommodation for the duration of the treatment, and those unable to meet these burdens will not have access to therapy. The organization of HABIT-ILE sessions on-site is also challenging, as it requires specialized facilities, extensive equipment, and a large number of trained therapists (ie, one per participant), in a context where there is a lack of therapists qualified to administer this type of therapy. Additionally, public health crises such as the COVID-19 pandemic have demonstrated how vulnerable such therapy models can be, as they can be significantly disrupted by global health emergencies. These constraints are not specific to HABIT-ILE but are shared by other intensive motor skill learning-based interventions and contribute to limiting their dissemination in routine practice [18,19]. To overcome these potential barriers and improve the accessibility of HABIT-ILE, we propose a telerehabilitation modality.

Besides intensive interventions, access to outpatient therapies after stroke is sometimes limited, especially in rural areas [20]. Although some motor skills acquired during intensive motor skill learning therapies are maintained in the long term [21], some functional goals or newly acquired motor abilities might be difficult to transfer to daily life directly after a camp. The long-term and additional effects of such intensive interventions could be enhanced with a specific follow-up intervention focusing on the potentiation of the newly acquired abilities and their transfer in daily life. Based on this assumption, we aim to investigate the efficacy of a follow-up program of HABIT-ILE therapy at home (HABIT-ILE@home) in adults with stroke after the completion of intensive HABIT-ILE intervention (either on-site or at home).

Aims and Hypotheses

This protocol comprises 2 randomized controlled trials (RCTs), hereafter referred to as RCT1 and RCT2, conducted in the same cohort of 48 adults with chronic stroke. The first trial (RCT1) aims to test whether a 2-week high-dosage HABIT-ILE telerehabilitation program (HABIT-ILE@home) is noninferior to a 2-week high-dosage on-site HABIT-ILE camp in improving motor function. We hypothesize that both interventions will lead to significant improvements in motor function, activities, and participation, and that the HABIT-ILE@home group will be noninferior to the on-site group when comparing changes from baseline to postintervention (T1-T0). We further hypothesize that HABIT-ILE@home will demonstrate higher self-perceived burden for caregivers but will show acceptable feasibility, adherence, and perceived efficacy.

The second trial (RCT2) aims to evaluate whether a structured 9-week specific HABIT-ILE@home follow-up program is more effective than a dose-matched nonspecific home-based follow-up in maintaining and further enhancing the improvements obtained after the intensive phase. We hypothesize that, over the follow-up period (T2-T1), participants allocated to the specific HABIT-ILE@home follow-up will show superior retention and/or additional improvements in motor function, activity performance, and social participation compared with those receiving the nonspecific follow-up. We further hypothesize that HABIT-ILE@home follow-up will demonstrate higher self-perceived burden for caregivers but will show acceptable feasibility, adherence, and perceived efficacy. All outcomes will be assessed at baseline (T0), immediately after the intensive 2-week intervention (T1), and after the 9-week follow-up period (T2).

Methods

Ethical Considerations

The study was approved by the Hospital-Faculty Ethics Committee of Saint-Luc – UCLouvain, Belgium (clinical trial number: B4032022000142) and was registered on ClinicalTrials.gov (NCT05727111) on February 2, 2023.

XSL•FO
RenderX

Participants were informed by means of a comprehensive information document and signed an informed consent form. All data collected for this study will be treated anonymously. Compensation for research-related harm, if applicable, will be covered according to the institutional insurance arrangements in place for this study. Written consent for the use of data, including images and videos, was obtained from the participants, or from their parents or legal guardians in the case of minors. The data will be anonymized using allocation codes for analysis after both RCTs.

Study Design

This study consists of 2 sequential, 2-arm, parallel-group pragmatic, single-blind RCTs conducted within the same cohort of adults with chronic stroke. RCT1 is a noninferiority trial comparing a 2-week high-dosage HABIT-ILE@home program with a dose-matched on-site HABIT-ILE camp (allocation ratio 1:1). RCT2 is a superiority trial comparing a 9-week specific HABIT-ILE@home follow-up to a dose-matched non-specific home-based follow-up. Randomization will be performed using pairwise randomization, and outcome assessors are blinded to group allocation. All evaluations will occur at baseline (T0), postintensive intervention (T1), and after the follow-up period (T2).

Patient and Public Involvement

Before this RCT, a pilot HABIT-ILE@home study was conducted in which participants and caregivers completed questionnaires on feasibility, acceptability, and burden; their feedback was used to modify the current study protocol and enhance several components of the HABIT-ILE@home intervention. In the present trial, participants' perception is gathered through completing satisfaction and burden questionnaires at the end of each intervention phase.

Participants and Setting

A total of 48 participants with chronic stroke will be recruited from Belgium and surrounding countries through mailing, spontaneous applications, social networks, physicians, physiotherapists, occupational therapists, and hospitals. Participants will be considered eligible if they meet the following criteria: time since stroke of more than 6 months, ages 18 years and older at inclusion, ability to interact, and understand simple instructions for completing assessments and therapy, ability to voluntarily initiate minimal movement of the affected upper extremity (eg, observable slight minimal movement in either shoulder, elbow, wrist, or fingers). A caregiver must also be available for 6.5 hours per day over 2 weeks for the needs of the high-dosage HABIT-ILE@home intervention (RCT1) and for the duration of follow-up programs (RCT2). Participants will be ineligible if one of the following criteria is present: botulinum toxin injections, orthopedic surgery, or another intensive therapy during the 3 months before the first assessment (T0) or during the study period. Additionally, participants will not be included if they are pregnant or present uncontrolled seizures or other uncontrolled health issues such as heart or renal failure. Cognitive function will be assessed at baseline using the Montreal Cognitive Assessment. This measure will only serve as a descriptor, not for eligibility or outcomes. All assessments will be conducted on the UCLouvain premises in Brussels or Louvain-la-Neuve. This study will be reported following CONSORT guidelines (Figure 1).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards for Reporting Trials) flowchart. ACTIVLIM: Activity Limitations Questionnaire; 6MWT: 6-minute walk test; BBT: box and block test; CaSM: Confidence After Stroke Measure; COPM: Canadian Occupational Performance Measure; FMA: Fugl-Meyer Assessment; HADS: Hospital Anxiety and Depression Scale; Mini-BESTest: Mini Balance Evaluation Systems Test; MoCA: Montreal Cognitive Assessment; mRS: modified Rankin scale; PILS-Stroke: Participation In Life Situations - Stroke; SIS: Stroke Impact Scale; TUG: timed up and go test; WMFT: Wolf Motor Function Test.

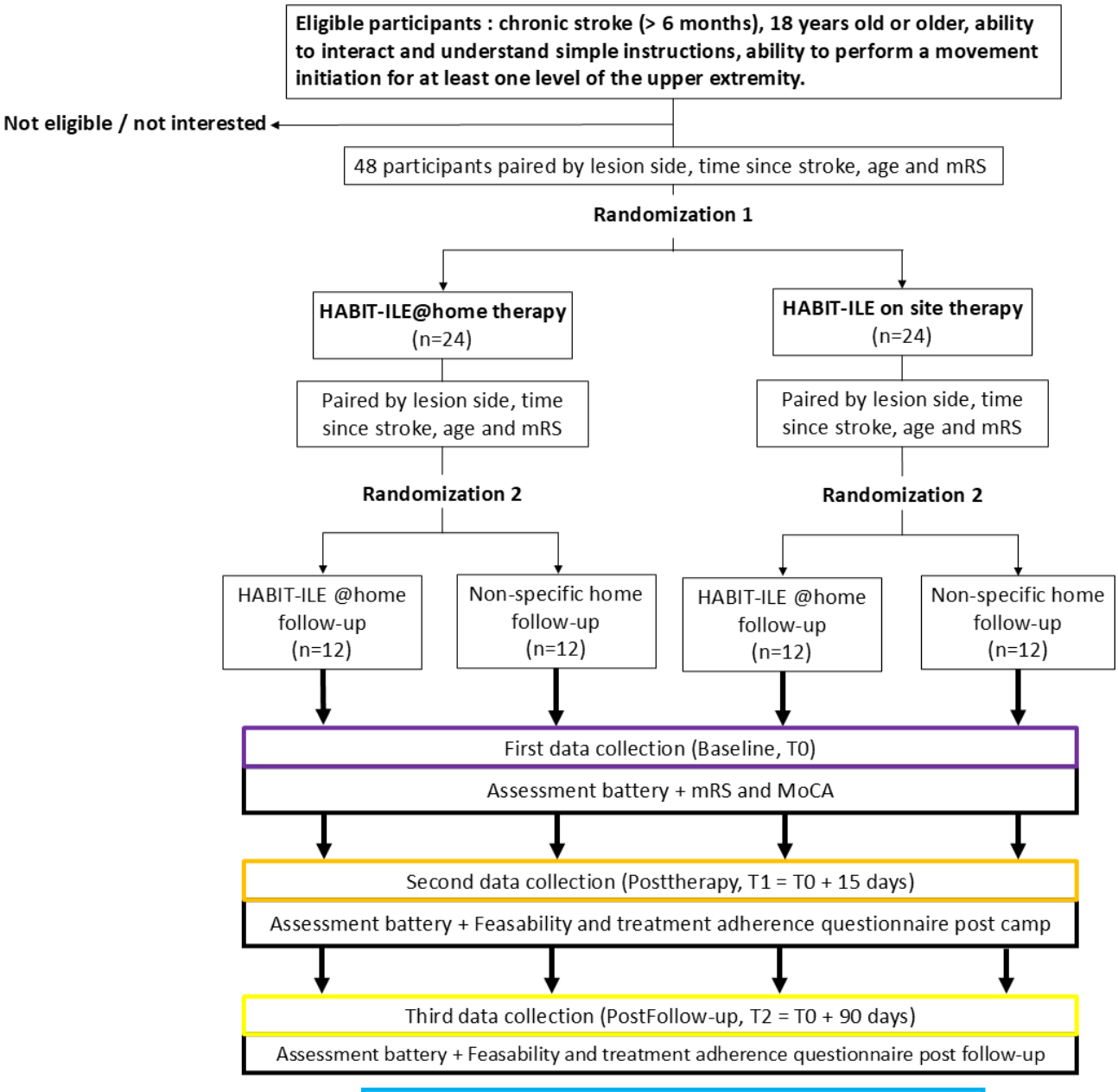

Assessment battery			
Body functions and anatomical structures	**FMA (Main Outcome)** BBT miniBESTest		TUG HADS CaSM
Activities	**Mobility**	**Self-care**	
	WMFT 6MWT	COPM ABILHAND-Stroke	ACTIVLIM-Stroke ABILOCO-Stroke
Participation	PILS-Stroke SIS	SATIS-Stroke	

Randomization Process

For RCT1, randomization will be performed using pairwise randomization. Participants will first be paired according to lesion side, time since stroke, age, and modified Rankin Scale (mRS) level. Participants in each pair will then be randomized into one of the two groups using an online external blind electronic randomization system [22], which will provide the

XSL•FO
RenderX

assignment only after pairing is completed. This procedure will ensure allocation concealment, as investigators will not be able to predict or influence the upcoming allocation. For RCT2, the same procedure will be repeated within each RCT1 allocation arm, such that participants previously assigned to HABIT-ILE@home or HABIT-ILE on-site will be again matched and randomized in the HABIT-ILE@home follow-up subgroup or the nonspecific follow-up subgroup.

Sample Size

For RCT1, assessing the noninferiority of HABIT-ILE@home compared to HABIT-ILE on-site, the sample size was calculated using PASS (version 14.0.15; NCSS) for a noninferiority 2-tailed *t* test of the difference between two means. The primary outcome is the change in the upper extremity subscale of the Fugl-Meyer Assessment (FMA-UE) score from baseline to postintervention (T1-T0), with higher scores indicating better motor abilities. We defined the noninferiority margin (NIM) using a method that ensures both clinical and statistical relevance, rather than only relying on the minimal clinically important difference. We chose the NIM as 50% of the difference between the changes observed with the gold-standard treatment and an interventional control group. This method was chosen to provide a more stringent criterion for noninferiority, ensuring that any treatment deemed non-inferior demonstrates a real benefit for stroke patients. As no previous studies have investigated such changes after a HABIT-ILE intervention, the sample size calculation was based on the mean change observed after a CIMT intervention (+5.6 points at FMA-UE; SD 3.1) [13]. The non-inferiority hypothesis was defined as follows: the HABIT-ILE@home intervention will be considered noninferior if the improvement produced is no more than 2.8 points lower than that obtained with the on-site HABIT-ILE intervention on the FMA-UE. An α level of 0.05 and a 1-β level of 0.8 were used for calculation. Under these assumptions, sample size calculation defined a number of 17 participants per group, resulting in a total of 34 participants.

For RCT2, which assesses the superiority of the specific HABIT-ILE@home follow-up compared with the nonspecific follow-up, the sample size calculation was also based on the mean changes in the FMA-UE observed in the study of Bonifer et al [13] by adjusting the expected improvements to account for the lower number of therapy hours in RCT2. Scaled proportionally, a 45-hour program corresponds to an expected improvement of 2.8 points. Because both follow-up arms constitute active interventions, the superiority margin was set at 1.4 points (50% of the expected 45-hour gain), reflecting a meaningful difference between two active interventions. Consequently, 17 participants are needed per group, resulting in a total of 34 participants, with a α level of .05 and a 1-β level of .8. To ensure that the required number of participants complete both phases, we plan to recruit 48 individuals, corresponding to an anticipated attrition rate of about 30%. This conservative estimate was based on a previous study with a similar population and reflects the two-phase design and the substantial differences between intervention modalities, with participants enrolled before being informed of their allocation [17].

Blinding Procedure

The main outcome (FMA) will be video-recorded and scored by experienced external raters, who are blind to group allocation and timing of assessment. Regarding secondary outcomes, the videos of the Wolf Motor Function Test (WMFT) and Mini Balance Evaluation Systems Test (Mini-BESTest) will also be scored blindly by external raters. As mentioned, the data will be anonymized using allocation codes for analysis after the end of both RCTs.

Study Interventions

RCT 1: HABIT-ILE@home vs HABIT-ILE on-Site Camp (2 Weeks)

Overview

During the first RCT, all participants will complete 2 weeks of HABIT-ILE therapy either at home or in an on-site camp. All the participants will benefit from 65 hours of intervention over 2 weeks (6.5 hours per day, 5 days per week).

HABIT-ILE is an intensive therapy that involves both hands and systematically engages the trunk and/or lower extremities in various functional activities [16,17]. It is based on the key therapeutic principles of motor skill learning [23,24]. HABIT-ILE is a goal-directed intervention, focusing on self-determined goals (maximum 5 goals/participant). It involves massed practice, which includes a high dosage of therapy, repetition of movements (trial-error training), and a high motor engagement time (ie, the percentage of time during which participants practice active movements). Therefore, the stimulation of voluntary motor control is achieved through hands-off practice. Throughout the intervention, the difficulty of the proposed structured activities is gradually increased based on motor, environmental, or task components. Participants' motivation is supported, notably through playful activities based on the patients' interests. For more details about the HABIT-ILE therapy, see Bleyenheuft and Gordon [16] and Ebner-Karestinos et al [17].

HABIT-ILE On-Site Camp (Control Group)

HABIT-ILE on-site intervention will be held in a rehabilitation center in Belgium. All participants will be in the same room. Each participant will benefit from an individualized intervention with one or two dedicated interventionists (occupational therapists, physiotherapists, or students of occupational therapy or physiotherapy) during the whole process. At the end of each day, a group activity will be performed with all the participants.

The therapeutic plan of each participant will be established, supervised, and adapted day after day by a team of supervisors (occupational therapist and/or physiotherapist specifically trained for the supervision of a HABIT-ILE intervention). The supervisors will coach the interventionists throughout the day to ensure the application of HABIT-ILE and motor skill learning principles. In addition, at the end of each treatment day, the supervision team will discuss the session and coach the interventionists to adapt the therapeutic plan and define the type of activities for the next day, in order to ensure goal-directed motor learning throughout the intervention.

XSL•FO
RenderX

HABIT-ILE@home (Treatment Group)

During the HABIT-ILE@home intervention, all participants will perform the intervention at home, accompanied by a non-therapist caregiver (patient's relative, partner, friends, etc). The caregiver will be coached by a supervision team through remote supervision to monitor the delivery of the intervention. The supervision team will include 1 to 2 supervisors and trained therapists, with each trained therapist coaching up to 4 patients during the intervention.

The caregiver will be involved throughout the entire therapeutic process to observe movement parameters, identify compensatory movements, and adapt the therapeutic environment in line with the "hands-off" and "increasing difficulty" principles (eg, by modifying the type of manipulated objects, adjusting the spatiotemporal requirements of the task, the role of the hands, and the standing or sitting environment). The caregiver will also provide positive feedback, motivate the patient, and ensure safety.

To fulfill their role successfully, caregivers will receive support from the supervision team in various ways. Before the therapy, caregivers will receive an adapted HABIT-ILE@home training session (one hour) as well as a "caregiver manual" to explain: (1) how to apply HABIT-ILE intervention and motor skill learning therapeutic principles, (2) their role as a caregiver, and (3) the rehabilitation equipment loaned to the participants. Two extra training hours will be provided the week before the therapy to familiarize participants with the telerehabilitation system used throughout the intervention. Moreover, the supervision team will generate and provide a customized treatment schedule to caregivers each evening for the next therapeutic day. Each program will include a description of the trained movements and skills as well as information about the proposed task-oriented activities or games (type of objects, type of grasping, orientation of objects, types of games, and duration).

Remote supervision will be scheduled for one hour each day within the 6.5 hours of therapy (Figure 2). The supervision sessions will be divided into two 30-minute sessions, one at the beginning of each therapy day, and one at the end. The first 30-minute session will be used to explain the daily therapeutic plan, demonstrate how to set up and carry out goal-oriented activities (instructions regarding the trained movements, the way to present manipulated objects, ensure safety, the type of instructions, and feedback needed) and adapt the therapeutic environment. Note that this initial session will last 1 hour on the first training day. The second 30-minute end session will be used to address participants' questions, discuss any difficulties encountered, assess the participants' progress, test new activities to adjust the difficulty level, and plan the therapeutic program for the following day.

Figure 2. Course of a HABIT-ILE@home therapy day: A day of HABIT-ILE@home will begin with 30 minutes of remote supervision via video conferencing software, then the participants and their caregivers will be left for 5 hours and 30 minutes in autonomy. Finally, the day will end with another 30 minutes of remote supervision. At the end of the day, the supervisor will send the treatment schedule for the next day.

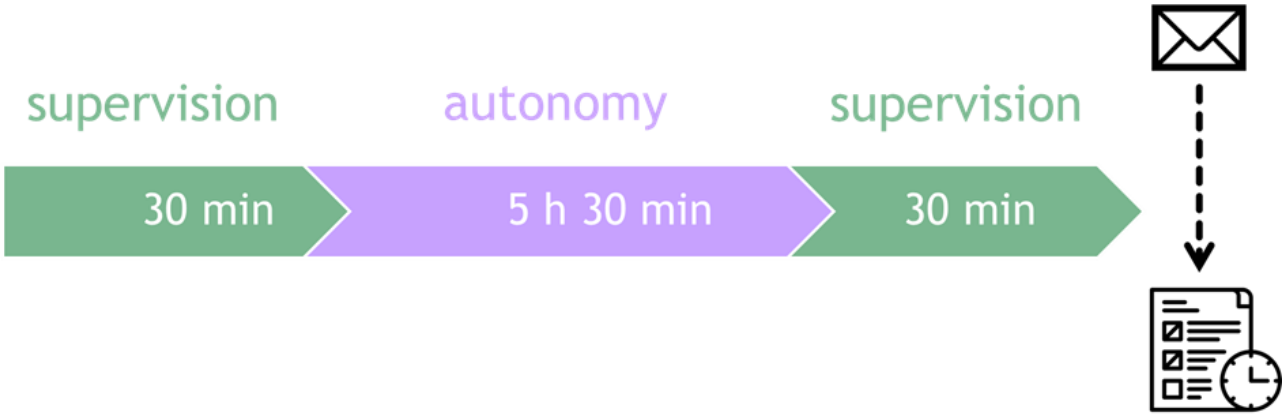

The implementation of HABIT-ILE therapy at home will be facilitated using a telerehabilitation device, the REAtouch Lite (Axinesis). REAtouch Lite is a smaller and lighter version of the REAtouch device, which was previously used in HABIT-ILE on-site camps with a pediatric population [25]. The REAtouch device is an interactive interface that interacts with customizable tangible objects through a range of virtual games and activities. It offers an environment designed to assist clinicians or caregivers in motivating the patient throughout the intervention and facilitate the decision-making process for implementing motor skill learning principles [25]. Note that, depending on the patients' therapeutic needs, goals, and the course of the intervention, activities will be performed with or without direct use of the REAtouch Lite. In addition to being a smaller and more transportable version of the initial device, the REAtouch Lite is equipped with a telecommunication module designed for remote communication and rehabilitation, including a smartphone for directly interacting with the patients and caregivers during supervision sessions. The use of a smartphone will enable supervision sessions both inside and outside the patient's home, and the caregiver will be able to take photos and videos and send them easily to the supervision team. The smartphone will be supplied with three phone holders to easily adjust camera angles during sessions; position 1: in front of the REAtouch Lite, providing a precise view of the patient. Position 2: positioned to the side of the patient, focusing on the lower limbs. Position 3: positioned for a wider perspective, allowing observation of movements within a 3 meter range.

Moreover, in the context of the HABIT-ILE@home program, the REAtouch Lite device will be supplied with a table with adjustable height to ensure adaptability based on postural control

XSL•FO
RenderX

and lower extremities involvement. The therapeutic environment will be enhanced by the addition of an adjustable bench, an inflatable ball, and/or a standing balance board, depending on the participants' motor abilities and functional goals. Furthermore, a box containing therapeutically useful objects and materials for adaptations (such as Velcro straps, tape, etc) will be delivered alongside the REAtouch Lite device (Figure 3).

Figure 3. Therapeutic environment delivered for HABIT-ILE@home, including: (A) an illustration of delivered therapeutic environment: phone holders, the Reatouch Lite, gym ball and/or adjustable bench, an adjustable table; (**B**) objects of various shapes and sizes to elicit different types of grasping; (**C**) adaptation material, such as velcro straps, tape, nonslip mat, etc, along with objects delivered.

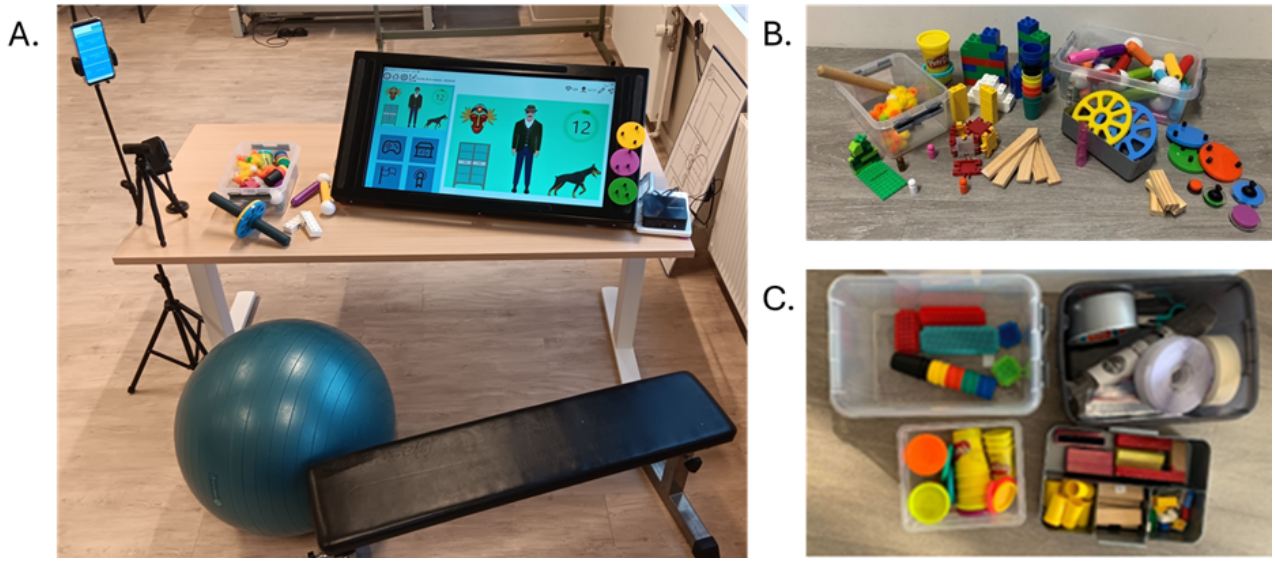

RCT 2: HABIT-ILE@home vs Nonspecific Follow-Up Interventions (9 Weeks)

Overview

During RCT2, all participants will take part in a low- to moderate-dosage home follow-up intervention, either with a specific HABIT-ILE@home program or with a nonspecific follow-up program managed by patients' usual therapists. In both groups, the follow-up program will last 9 weeks, with a recommendation of 5 hours of therapy per week (in addition to usual care) performed at home (45 hours in total).

Specific HABIT-ILE@home Follow-Up (Treatment Group)

For the specific HABIT-ILE@home follow-up, the treatment modalities will be the same as for the 2-week HABIT-ILE@home intervention described previously, except for the therapeutic and supervision dosage. A 1-hour session will be planned each week, as part of the 5 recommended hours. Like the previously mentioned daily program, a detailed therapeutic schedule will be sent and adapted every 1 or 2 weeks, depending on the participants' needs.

Nonspecific Home Follow-Up (Control Group)

For the nonspecific home follow-up, participants' usual therapist will be contacted and asked to develop a 5-hour weekly program based on their usual practice and treatment plan for 9 weeks. If contact could not be established with the participants' usual therapists, documentation with instructions for self-rehabilitation exercises will be sent to the patients and their caregivers, including stretching of the affected side, muscle strengthening, and task-oriented exercises as described in an online booklet [26].

Criteria for Discontinuing or Modifying Allocated Interventions

The intervention may be discontinued at the participant's request or if safety/tolerability concerns arise. All reasons will be documented, and follow-up assessments will be completed whenever possible.

Data Collection and Data Management

After each assessment session, data will be entered into the Research Electronic Data Capture (REDCap) system hosted at UCLouvain. A detailed codebook will be created to document each variable, including its name, possible values, data labels, and units.

Data Monitoring Committee

At UCLouvain, there is no data monitoring committee for nondrug interventional studies. However, the Hospital-Faculty Ethics Committee of Saint-Luc, composed of a majority of doctors (from UCLouvain or the Cliniques Universitaires de Saint-Luc) but also including a general practitioner, nurses, pharmacists, lawyers, ethicists, a methodologist, a psychologist, scientific collaborators, patients' representatives, and healthy volunteers' representatives, ensures that participants in the clinical trials are not exposed to undue risks. The Ethics Committee reviews the study protocol and all relevant information about the experimental treatment to ensure research safety. Additionally, it verifies that the information provided to participants is clear and accessible, enabling fully informed consent. Importantly, the Committee operates independently from the study sponsor, with no members affiliated with the research project.

XSL•FO
RenderX

Outcomes

Primary Outcomes

The primary outcome measure is change in the FMA. The FMA is a standardized assessment tool initially designed to measure motor function, sensation (light touch and position sense), balance, joint range of motion, and joint pain in the upper and lower extremities of poststroke patients [27]. More specifically, the motor function subsection assesses reflex activity, volitional movement, and coordination/speed. Only the total motor score (upper and lower extremities) will be used in this study. The score varies from 0 to 100 points, with 66 points for upper extremities and 34 for lower extremities. Higher values represent better motor function. The FMA has excellent reliability [28], and its validity [29] and responsiveness [30] have been demonstrated. The training procedure proposed to enhance standardization of the assessment [31] will be completed by both examiners and blind assessors. For RCT1, the primary outcome for comparison will be the change from T0 to T1, and for RCT2, the primary outcome for comparison will be the change from T1 to T2.

Secondary Outcomes

The same secondary outcome battery will be administered at T0, T1, and T2 for both RCT1 and RCT2. For RCT1, secondary outcomes for comparison will be the change from T0 to T1, and for RCT2 the secondary outcomes will be the change from T1 to T2.

Body Structures and Functions

Anxiety and depression will be measured by the Hospital Anxiety and Depression Scale (HADS) [32]. The HADS is a 14-item self-perceived questionnaire (7 items for the anxiety subscale and 7 items for the depression subscale) asking participants to estimate their mood in the past week [33]. Each item is rated on a 4-level scale, leading to a total score ranging from 0 to 21 points for each subscale used. Higher scores indicate higher anxiety and depression. HADS has shown an excellent internal consistency in the stroke population [34].

Gross unimanual dexterity will be assessed by the Box and Block Test, which measures the number of blocks transported from one compartment of a box to another in a minute [35]. This tool shows excellent test-retest reliability and is sensitive to detect changes in stroke patients [36,37].

Activities

Upper extremity motor capacity will be measured by the 17-item version of the WMFT [38]. A total of 15 items must be performed as quickly as possible, with a 120-second time limit. A total of 6 items assess joint-segment movements while 9 items evaluate integrative functional movements. Each of these 15 items is rated on a 6-level scale, leading to a total score ranging from 0 to 75 points. Higher scores represent higher upper extremity motor capacity. The remaining 2 items involve measures of strength. The WMFT is considered a reliable and valid measurement tool [38].

Manual activity performance, defined as the ability to manage daily activities requiring the use of the upper extremities (whatever the strategies involved), will be assessed using the self-administered Abilities of the Hand (ABILHAND-Stroke) questionnaire. Patients will be asked to report their ease or difficulty in performing 23 bimanual activities. As a Rasch-built questionnaire, the measures which are linear correspond to the percentage of the full range scale expressed in logits, from 0% (low manual performance) to 100% (high manual performance). This questionnaire shows a high reliability and a good validity [39].

Balance will be measured by the Mini-BESTest, which is a shortened version of the Balance Evaluation Systems Test. A total of 4 main domains of balance (anticipatory postural adjustments, reactive postural control, sensory orientation, and dynamic gait) are evaluated through 14 items. Each item is rated on a 3-level scale, leading to a total score ranging from 0 to 28 points. Higher values indicate better balance. This test is considered reliable and valid [40].

Lower extremity capacity, including mobility, balance, and walking skills, will be assessed by the timed up and go test. From a seated position, the patient will be asked to stand up, walk 3 meters, then turn around, walk 3 meters, and sit down. The score (in seconds) is the time taken to complete the functional task. The timed up and go test is a reliable and valid test for stroke patients [41].

Walking endurance will be evaluated by the 6-minute walk test [42]. The patients will have to cover the greatest possible distance in 6 minutes in a 30 m hallway, without taking a break. The 6-minute walk test is a good measure of endurance, giving an idea of daily life performance. This test has excellent validity and reliability [41].

Locomotor performance, defined as the patients' ability to move about effectively in their environment, will be measured by the ABILOCO-Stroke questionnaire. Patients will be asked whether they are able to perform 13 locomotor activities on a 2-level scale (impossible or possible). ABILOCO-Stroke shows good psychometric properties [43]. As a Rasch-built questionnaire, the measures which are linear correspond to the percentage of the full range scale expressed in logits, from 0% (low locomotor performance) to 100% (high locomotor performance).

Global activity performance, defined as the ability to manage daily activities requiring the use of the upper and/or lower extremities (whatever the strategies involved), will be assessed by the Activity Limitations Questionnaire (ACTIVLIM-Stroke) questionnaire. Patients will be asked to report their ease or difficulty in performing 20 daily activities. As a Rasch-built questionnaire, the measures which are linear correspond to the percentage of the full range scale expressed in logits, from 0% (low global performance) to 100% (high global performance). The questionnaire shows a high validity [44].

The self-perceived patient's performance in functional goals and their satisfaction with achieved performance will be evaluated through the Canadian Occupational Performance Measure. Patients will be asked to estimate their performance and their satisfaction on a 10-level scale, from 1 (low performance/satisfaction) to 10 (high performance/satisfaction). The Canadian Occupational Performance Measure is a reliable, valid, and responsive outcome measure [45].

XSL•FO
RenderX

Global disability level will be classified using the mRS. This scale classifies stroke patients into 7 levels, from 0 (no symptoms) to 5 (severe disability: confined to bed, incontinent, and requiring constant nursing care and attention) and 6 (death) [46]. The scale has shown excellent test-retest reliability [47].

Participation

Satisfaction in activities and social participation will be assessed by the SATIS-Stroke questionnaire. Patients will be asked to rate their satisfaction in the achievement of 36 activities or social life situations on a 4-level scale. As a Rasch-built questionnaire, the measures which are linear correspond to the percentage of the full range scale expressed in logits, from 0% (low satisfaction) to 100% (high satisfaction). Excellent test-retest reliability has been shown [48].

Participation, defined as the patient's involvement in life situations (chosen because they make sense to them) in which they interact with others or which are linked to social roles, will be measured by the Participation In Life Situations – Stroke questionnaire (PILS-Stroke). This tool, which is still under development, will generate linear measures (in percentage of the full range scale expressed in logits), from 0% (low participation) to 100% (high participation).

Personal Factors

Confidence after a stroke will be evaluated using the 27-item Confidence After Stroke Measure. This self-administered questionnaire consists of 9 items of self-confidence (ie, feeling confident in one's abilities, qualities, and judgment), 8 items of positive attitudes (ie, being optimistic about situations, interactions, and oneself), and 10 items of social confidence (ie, feeling comfortable in social situations and in social acceptance). Each item is rated on a 4-level scale, leading to a total score ranging from 27 to 108 points. The higher the values, the higher the confidence. The Confidence After Stroke Measure has shown good psychometric properties [49].

Health-Related Quality of Life

Health-related quality of life after stroke will be evaluated using the 59-item Stroke Impact Scale (SIS) [50]. This self-reported questionnaire covers 8 main health-related quality of life domains: strength (4 items), hand function (5 items), (instrumental) activities of daily living (10 items), mobility (9 items), communication (7 items), emotion (9 items), memory/thinking (7 items), and participation/role function (8 items). Patients will be asked to estimate the difficulty they experience in completing each item on a 5-level scale. The SIS shows adequate to excellent reliability [51] and excellent validity [52]. The proxy-patient version of the SIS will also be completed by the caregiver [53]. For each domain, an ordinal score is calculated with the following formula: Transformed Scale = [(Actual raw score - lowest possible raw score) ÷ Possible raw score range] x 100. A composite physical subscale can be calculated by combining 4 of the physical domains.

Feasibility and Adherence to Treatment (Satisfaction Questionnaire)

A satisfaction questionnaire was developed to assess the feasibility of both the 2-week HABIT-ILE@home intervention and the 9-week HABIT-ILE@home follow-up, and to compare them with the HABIT-ILE on-site camp and nonspecific follow-up, respectively. Patients' and caregivers/interventionists' adherence to treatment will also be recorded through the satisfaction questionnaire. The development of this questionnaire was informed by the HABIT-ILE@home pilot study, in which the feasibility, acceptability, and perceived burden of the telerehabilitation approach were assessed using instruments derived from validated scales [54]. Specifically, the questionnaire integrates dimensions adapted from the System Usability Scale, the User Experience Questionnaire, AttrakDiff2, the Dimensions of Mastery Questionnaire, the Intrinsic Motivation Inventory, the Exercise Adherence Rating Scale, and the Exercise Therapy Burden Questionnaire. Based on this, the questionnaire used in the present trial assesses the following dimensions: REAtouch Lite interactive system features, caregivers' ability to follow the therapy schedule, and to apply motor skill learning principles, caregivers' compliance, patients' motivation and involvement, patients' compliance (perceived usefulness of the therapy, enjoyment, and burden of the therapy), and preferences regarding therapy modalities (remote, in-person, or hybrid). The same questionnaire structure was adapted for the on-site modality, ensuring comparability (eg, removal of virtual-related sections, rewording of items). The satisfaction questionnaire will be filled in by patients, caregivers, and/or interventionists (only for on-site at T1), at the end of each intervention phase (T1 and T2).

Documentation of the Therapy Content

The content of the 2-week intensive therapy will be documented using multiple methods (RCT1). First, inertial measurement units, specifically the Xsens Dot (Xsens Technologies BV), will be used to measure and document motor activity (quantity and type of movements) during both on-site camp and at-home HABIT-ILE modalities. Each participant will be equipped with three wearable sensors during the 6.5 hours of daily high-dosage therapy: two placed dorsally on the wrists and one on the front of the right thigh. Activity count (based on a predefined threshold for movement) and asymmetry ratio (comparing right and left upper extremities activity) will be calculated and should make it possible to differentiate between uni- and bimanual tasks [55]. Accelerometer data should also make it possible to evaluate participants' positions to determine whether they are predominantly sitting or standing during therapy sessions, as well as to quantify the percentage of time spent walking. However, inertial measurement unit–based classification remains approximate, and specific tasks cannot be distinguished with certainty. During the on-site modality, interventionists will complete a structured timesheet detailing each activity in terms of duration and type of motor activity performed throughout the day. For the home modality, therapy content will be partly documented based on the prescribed daily program and data extracted from the REAtouch Lite system (session duration and touch count on the screen). Additionally, at the end of each day, participants will report the duration of therapy they were able to complete.

For the follow-up interventions (RCT2), participants will maintain a weekly log over a 9-week period, briefly describing the type and duration of activities performed each day. Monitoring the completion rates of therapy sessions and

XSL•FO
RenderX

comparing the hours of therapy actually performed with those planned will enable us to judge the feasibility of setting up such follow-ups.

Statistical Analysis

Overview

Statistical analysis will be performed after the last measurements at the follow-up assessment using RStudio software (version 2023.12.15+402; Posit). Baseline characteristics will be summarized with means and SDs for normally distributed variables, medians and IQRs for nonnormal variables. Baseline group equivalence will be explored descriptively using independent *t* tests or nonparametric equivalents.

RCT 1: HABIT-ILE@home vs HABIT-ILE on-Site Camp (2 Weeks)

Non-inferiority statistics will be used to compare the efficacy of the HABIT-ILE@home modality with the HABIT-ILE on-site camp modality. First, within each intervention arm, changes from baseline (T0) to postintervention (T1) will be evaluated for all outcomes. Depending on data normality, paired samples *t* tests or nonparametric equivalents will be applied. Then, the 95% CI of the difference between groups will be computed by comparing group differences in change scores using independent *t* tests or nonparametric equivalents. Noninferiority will be concluded if the lower bound of this interval is above the prespecified NIM. For all outcomes, NIM will be defined as 50% of the difference between the changes observed with a gold standard (ie, CIMT) and a control group [13].

RCT 2: HABIT-ILE@home vs Nonspecific Follow-Up Interventions (9 Weeks)

Superiority statistical analyses will be used to examine changes during the 9-week follow-up period (T2-T1). Depending on the distribution of the data, the analyses will be conducted using either analysis of covariance (ANCOVA; for normally distributed data) or nonparametric equivalents (eg, rank-based ANCOVA) for nonnormally distributed data. Covariates such as initial T1 scores, mRS, and/or type of initial intervention will be included in the model to adjust comparisons. In case of analyses including all three time points, linear mixed-effects models will be used, with baseline score and group allocation entered as fixed effects.

Feasibility and Adherence

For questionnaire items shared by both groups, between-group comparisons will use Wilcoxon rank-sum tests. Items specific to the home-based condition will be summarized descriptively.

Missing Data

All analyses will follow an intention-to-treat approach. Missing data will be addressed using multiple imputation. All variables and auxiliary variables predictive of missingness will be included in the imputation model.

Clinical Significance

The proportion of participants exceeding established minimal clinically important differences on each outcome will be reported for both groups.

Results

Recruitment started in March 2023 and ended in March 2025. All participants have been recruited and have undergone interventions in RCT1 and RCT2. Data analysis is expected to start in spring 2026.

Discussion

Principal Findings

Previous studies have shown that home-based telerehabilitation may be as effective as in-person therapies [56] with good treatment adherence [57]. Specifically, intensive interventions based on motor skill learning principles such as CIMT have shown similar effectiveness at home (both with or without a rehabilitation device) compared to in-person CIMT [19] or standard therapy [58]. Based on these previous results, we anticipate similar improvements when comparing HABIT-ILE@home and HABIT-ILE on-site camp within the first RCT.

However, there may be some challenges for participants to engage in such an intensive home intervention (short duration, high dosage), or even for a home follow-up intervention (longer duration, low to moderate dosage). The higher dosage of the therapy (65 hours) compared to previous studies may be more difficult for participants and caregivers to manage in terms of tiredness and motivation. Nevertheless, intrinsic motivation developed by participants based on the self-determined functional goals and the perceived usefulness of the therapy might help to maintain their treatment compliance. Indeed, a HABIT-ILE@home pilot study conducted in 3 adults with chronic stroke showed that the perspective of improving their motor control capacities, reaching their goals, and gaining autonomy was their main source of motivation, more than the playful nature of the therapy [54]. The daily program sent each day by the supervision team will aim to target the functional goals considered as important by the participant, thus supporting their intrinsic motivation.

In addition to the patients' and caregivers' motivation and adherence to treatment, a second aspect of the home-based protocol that needs to be documented is the burden of treatment. Indeed, caring for patients with chronic stroke may impose a significant burden on caregivers and could lead to heightened levels of anxiety and depression [59]. This affected mental health in caregivers can result in slower progress and poorer overall outcomes for the patient [60]. Moreover, the emotional state of caregivers and stroke patients is interconnected. Patients whose caregivers report greater burdens often exhibit higher levels of anxiety themselves, which could also impact patient improvements [59,61].

Introducing a new rehabilitation approach that involves caregivers may initially increase their burden, as they must understand and adapt to new responsibilities associated with the therapeutic process [61,62]. However, over time, as caregivers gain competence in managing their relative's health condition, they may experience reduced anxiety and greater satisfaction in their caregiving role [63,64]. This shift highlights

XSL•FO
RenderX

the potential of involving the caregivers in rehabilitation to not only enhance the stroke patients' recovery but also improve the mental health of the caregivers. Involving caregivers in rehabilitation can lead to mutual benefits for themselves and stroke patients, ultimately promoting better health outcomes. However, this can only be achieved if caregivers feel sufficiently supported by therapists by receiving adequate clinical supervision [65]. As planned in the present intervention, previous studies have observed a positive effect of a dedicated initial training to reduce the caregivers' burden and enhance perceived competence in managing therapeutic tasks, resulting in less stress and a greater engagement in rehabilitation [66].

Based on the current recommendations for motor and functional stroke rehabilitation, we hypothesize that the specific HABIT-ILE@home follow-up will result in better retention of changes following 2-week high-dosage therapies, and even better additional improvements than the nonspecific follow-up intervention (RCT 2) [67,68]. The HABIT-ILE follow-up program follows these recommendations by using a patient-centered approach that focuses on setting individualized goals to shape the therapy. It encourages self-initiated movements through goal-oriented training, promoting motor learning through trial and error. The program also incorporates shaping techniques, where activities are adapted to the patient's current abilities and gradually become more challenging as they progress. This is further supported by intensive practice sessions designed to maximize the number of repetitions. However, adding 5 hours of home therapy per week to the participants' usual schedule may be a challenge. Multiple determinants may cause nonadherence to a home-based program: psychological factors, lack of social support, time constraints, and complexity of activities proposed [69,70]. To avoid nonadherence to the prescribed exercise regimen during HABIT-ILE@home follow-up, we plan to set up a weekly schedule of exercises tailored to individual goals of each participant along with gamified activities in addition to the 1-hour supervised session each week [71].

The participants' adherence to the 2-week HABIT-ILE@home intervention and the 9-week HABIT-ILE@home follow-up, as well as the burden and feasibility of these interventions, will be documented through a satisfaction questionnaire. This should help us highlight some of the pros and cons of such interventions. Another concern that will be investigated is the possibility of adhering to the HABIT-ILE therapeutic principles during the whole therapeutic time in the home-based setting. This will be partially inferred from the accelerometers, REAtouch Lite, and logbook data documenting the content of the therapy. Even if we cannot be sure of the correct application of all the HABIT-ILE principles, we assume that caregivers' and participants' training before therapy, provision of a daily detailed treatment schedule, and adequate remote supervision will help to address this concern.

This trial is expected to determine whether HABIT-ILE@home can produce improvements comparable to those obtained during the on-site intensive program across motor, activity, and participation-related outcomes, while substantially reducing logistical and accessibility barriers. Demonstrating noninferiority would support the scalability of high-dose bimanual motor-learning interventions in community settings. The follow-up phase will also provide important insights into the role of structured, task-specific home programs in maintaining or enhancing gains achieved during intensive interventions. If the specific HABIT-ILE@home follow-up proves superior to the nonspecific program, this would highlight the importance of targeted postintensive rehabilitation to promote retention and transfer of improvements into daily life.

Conclusion

This study will provide evidence on the feasibility and efficacy of delivering HABIT-ILE through a home-based telerehabilitation modality for adults with chronic stroke. Demonstrating noninferiority of HABIT-ILE@home compared to on-site therapy would support wider accessibility to intensive rehabilitation while reducing logistical and human resource constraints. Additionally, showing the added benefit of a structured follow-up could emphasize the importance of continuity of care to sustain and enhance motor recovery after intensive interventions.

Acknowledgments

The generative artificial intelligence tool, ChatGPT-5.1, was used for refinement, correction, editing, or formatting the manuscript to improve clarity of language.

Funding

This project is funded by BioWIN grant of the Région Wallonne (31/8526). The funding body has reviewed this protocol during the selection process, but after this initial point, BioWIN has not participated in the design study and will not participate in data collection, analysis, interpretation of the data, or in writing the manuscript.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study. Upon completion of the study, data will be stored in the UCLouvain Dataverse repository; access to selected data will be provided upon reasonable request to the corresponding author and after signing a data sharing agreement.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: MS, ZR, ED, CA, YB, GS
Methodology: MS, ZR, ED, CA, YB, GS
Funding acquisition: CA, YB, GS
Project administration: YB, GS
Supervision: YB, GS
Investigation: MS, ZR, ED, YB, GS
Data curation: MS, ZR, ED
Formal analysis: MS, CA, YB, GS
Ethics/regulatory approvals: YB
Data monitoring: GS
Writing – original draft: MS
Writing – Review & Editing: MS, CA, YB, GS.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[DOCX File, 35 KB - resprot_v15i1e87035_app1.docx]

References

1. GBD 2016 Lifetime Risk of Stroke Collaborators, Feigin VL, Nguyen G, et al. Global, regional, and country-specific lifetime risks of stroke, 1990 and 2016. N Engl J Med 2018 Dec 20;379(25):2429-2437. [doi: 10.1056/NEJMoa1804492] [Medline: 30575491]
2. Bonita R, Mendis S, Truelsen T, et al. The global stroke initiative. Lancet Neurol 2004 Jul;3(7):391-393. [doi: 10.1016/S1474-4422(04)00800-2] [Medline: 15207791]
3. Lee KE, Choi M, Jeoung B. Effectiveness of rehabilitation exercise in improving physical function of stroke patients: a systematic review. Int J Environ Res Public Health 2022 Oct 5;19(19):12739. [doi: 10.3390/ijerph191912739] [Medline: 36232038]
4. Langhorne P, Coupar F, Pollock A. Motor recovery after stroke: a systematic review. Lancet Neurol 2009 Aug;8(8):741-754. [doi: 10.1016/S1474-4422(09)70150-4] [Medline: 19608100]
5. Winstein CJ, Stein J, Arena R, et al. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2016 Jun;47(6):e98-e169. [doi: 10.1161/STR.0000000000000098] [Medline: 27145936]
6. Cauraugh JH, Kang N. Bimanual movements and chronic stroke rehabilitation: looking back and looking forward. Appl Sci (Basel) 2021;11(22):10858. [doi: 10.3390/app112210858]
7. Stinear CM, Lang CE, Zeiler S, et al. Advances and challenges in stroke rehabilitation. Lancet Neurol 2020 Apr;19(4):348-360. [doi: 10.1016/S1474-4422(19)30415-6] [Medline: 32004440]
8. Veerbeek JM, van Wegen E, van Peppen R, et al. What is the evidence for physical therapy poststroke? A systematic review and meta-analysis. PLoS ONE 2014;9(2):e87987. [doi: 10.1371/journal.pone.0087987] [Medline: 24505342]
9. Daly JJ, McCabe JP, Holcomb J, et al. Long-dose intensive therapy is necessary for strong, clinically significant, upper limb functional gains and retained gains in severe/moderate chronic stroke. Neurorehabil Neural Repair 2019 Jul;33(7):523-537. [doi: 10.1177/1545968319846120] [Medline: 31131743]
10. Hatem SM, Saussez G, Della Faille M, et al. Rehabilitation of motor function after stroke: a multiple systematic review focused on techniques to stimulate upper extremity recovery. Front Hum Neurosci 2016;10(442):442. [doi: 10.3389/fnhum.2016.00442] [Medline: 27679565]
11. Meng G, Meng X, Tan Y, et al. Short-term efficacy of hand-arm bimanual intensive training on upper arm function in acute stroke patients: a randomized controlled trial. Front Neurol 2017;8(726):726. [doi: 10.3389/fneur.2017.00726] [Medline: 29403422]
12. Wolf SL, Winstein CJ, Miller JP, et al. Effect of constraint-induced movement therapy on upper extremity function 3 to 9 months after stroke: the EXCITE randomized clinical trial. JAMA 2006 Nov 1;296(17):2095-2104. [doi: 10.1001/jama.296.17.2095] [Medline: 17077374]
13. Bonifer NM, Anderson KM, Arciniegas DB. Constraint-induced movement therapy after stroke: efficacy for patients with minimal upper-extremity motor ability. Arch Phys Med Rehabil 2005 Sep;86(9):1867-1873. [doi: 10.1016/j.apmr.2005.04.002] [Medline: 16181956]

XSL•FO
RenderX

14. Stevenson T, Thalman L, Christie H, Poluha W. Constraint-induced movement therapy compared to dose-matched interventions for upper-limb dysfunction in adult survivors of stroke: a systematic review with meta-analysis. Physiother Can 2012;64(4):397-413. [doi: 10.3138/ptc.2011-24] [Medline: 23997396]
15. Pollock A, Farmer SE, Brady MC, et al. Interventions for improving upper limb function after stroke. Cochrane Database Syst Rev 2014 Nov 12;2014(11):CD010820. [doi: 10.1002/14651858.CD010820.pub2] [Medline: 25387001]
16. Bleyenheuft Y, Gordon AM. Hand-arm bimanual intensive therapy including lower extremities (HABIT-ILE) for children with cerebral palsy. Phys Occup Ther Pediatr 2014 Nov;34(4):390-403. [doi: 10.3109/01942638.2014.932884] [Medline: 25271469]
17. Ebner-Karestinos D, Gathy E, Carton de Tournai A, et al. Hand-Arm Bimanual Intensive Therapy Including Lower Extremities (HABIT-ILE) in adults with chronic stroke: protocol of a randomised controlled trial. BMJ Open 2023 Apr 13;13(4):e070642. [doi: 10.1136/bmjopen-2022-070642] [Medline: 37055214]
18. Claflin ES, Krishnan C, Khot SP. Emerging treatments for motor rehabilitation after stroke. Neurohospitalist 2015 Apr;5(2):77-88. [doi: 10.1177/1941874414561023] [Medline: 25829989]
19. Uswatte G, Taub E, Lum P, et al. Tele-rehabilitation of upper-extremity hemiparesis after stroke: Proof-of-concept randomized controlled trial of in-home Constraint-Induced Movement therapy. Restor Neurol Neurosci 2021;39(4):303-318. [doi: 10.3233/RNN-201100] [Medline: 34459426]
20. Morrow C, Woodbury M, Simpson AN, Almallouhi E, Simpson KN. Differences in rehabilitation evaluation access for rural and socially disadvantaged stroke survivors. Top Stroke Rehabil 2024 Sep;31(6):625-631. [doi: 10.1080/10749357.2024.2312638] [Medline: 38369788]
21. Kwakkel G, Veerbeek JM, van Wegen EEH, Wolf SL. Constraint-induced movement therapy after stroke. Lancet Neurol 2015 Feb;14(2):224-234. [doi: 10.1016/S1474-4422(14)70160-7]
22. Research Randomizer. URL: https://www.randomizer.org/ [accessed 2026-01-23]
23. Leech KA, Roemmich RT, Gordon J, Reisman DS, Cherry-Allen KM. Updates in motor learning: implications for physical therapist practice and education. Phys Ther 2022 Jan 1;102(1):pzab250. [doi: 10.1093/ptj/pzab250] [Medline: 34718787]
24. Spampinato D, Celnik P. Multiple motor learning processes in humans: defining their neurophysiological bases. Neuroscientist 2021 Jun;27(3):246-267. [doi: 10.1177/1073858420939552] [Medline: 32713291]
25. Saussez G, Bailly R, Araneda R, et al. Efficacy of integrating a semi-immersive virtual device in the HABIT-ILE intervention for children with unilateral cerebral palsy: a non-inferiority randomized controlled trial. J Neuroeng Rehabil 2023 Jul 29;20(1):98. [doi: 10.1186/s12984-023-01218-4] [Medline: 37516873]
26. Bonnyaud C, Roche N. Livret d'auto-rééducation pour les personnes hémiplégiques [Report in French]. : Allergan; 2017 URL: https://www.handicap.org/wp-content/uploads/2020/04/LIVRE-FICHES_GARCHES_11_01-2017-A4-BAT.pdf [accessed 2026-01-23]
27. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. JRM 1975;7(1):13-31. [doi: 10.2340/1650197771331]
28. Lin JH, Hsueh IP, Sheu CF, Hsieh CL. Psychometric properties of the sensory scale of the Fugl-Meyer Assessment in stroke patients. Clin Rehabil 2004 Jun;18(4):391-397. [doi: 10.1191/0269215504cr737oa] [Medline: 15180122]
29. Woodbury ML, Velozo CA, Richards LG, Duncan PW, Studenski S, Lai SM. Longitudinal stability of the Fugl-Meyer Assessment of the upper extremity. Arch Phys Med Rehabil 2008 Aug;89(8):1563-1569. [doi: 10.1016/j.apmr.2007.12.041] [Medline: 18674991]
30. Rabadi MH, Rabadi FM. Comparison of the action research arm test and the Fugl-Meyer assessment as measures of upper-extremity motor weakness after stroke. Arch Phys Med Rehabil 2006 Jul;87(7):962-966. [doi: 10.1016/j.apmr.2006.02.036] [Medline: 16813784]
31. See J, Dodakian L, Chou C, et al. A standardized approach to the Fugl-Meyer assessment and its implications for clinical trials. Neurorehabil Neural Repair 2013 Oct;27(8):732-741. [doi: 10.1177/1545968313491000] [Medline: 23774125]
32. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
33. Dunbar M, Ford G, Hunt K, Der G. A confirmatory factor analysis of the Hospital Anxiety and Depression scale: comparing empirically and theoretically derived structures. Br J Clin Psychol 2000 Mar;39(1):79-94. [doi: 10.1348/014466500163121] [Medline: 10789030]
34. Aben I, Verhey F, Lousberg R, Lodder J, Honig A. Validity of the beck depression inventory, hospital anxiety and depression scale, SCL-90, and hamilton depression rating scale as screening instruments for depression in stroke patients. Psychosomatics 2002;43(5):386-393. [doi: 10.1176/appi.psy.43.5.386] [Medline: 12297607]
35. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block Test of manual dexterity. Am J Occup Ther 1985 Jun;39(6):386-391. [doi: 10.5014/ajot.39.6.386] [Medline: 3160243]
36. Lin KC, Chuang LL, Wu CY, Hsieh YW, Chang WY. Responsiveness and validity of three dexterous function measures in stroke rehabilitation. J Rehabil Res Dev 2010;47(6):563-571. [doi: 10.1682/jrrd.2009.09.0155] [Medline: 20848369]
37. Platz T, Pinkowski C, van Wijck F, Kim IH, di Bella P, Johnson G. Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study. Clin Rehabil 2005 Jun;19(4):404-411. [doi: 10.1191/0269215505cr832oa]

XSL•FO
RenderX

38. Morris DM, Uswatte G, Crago JE, Cook EW III, Taub E. The reliability of the Wolf Motor Function Test for assessing upper extremity function after stroke. Arch Phys Med Rehabil 2001 Jun;82(6):750-755. [doi: 10.1053/apmr.2001.23183]
39. Penta M, Tesio L, Arnould C, Zancan A, Thonnard JL. The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. Stroke 2001 Jul;32(7):1627-1634. [doi: 10.1161/01.str.32.7.1627] [Medline: 11441211]
40. Tsang CSL, Liao LR, Chung RCK, Pang MYC. Psychometric properties of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in community-dwelling individuals with chronic stroke. Phys Ther 2013 Aug;93(8):1102-1115. [doi: 10.2522/ptj.20120454] [Medline: 23559522]
41. Flansbjer UB, Holmbäck AM, Downham D, Patten C, Lexell J. Reliability of gait performance tests in men and women with hemiparesis after stroke. J Rehabil Med 2005 Mar;37(2):75-82. [doi: 10.1080/16501970410017215] [Medline: 15788341]
42. Kosak M, Smith T. Comparison of the 2-, 6-, and 12-minute walk tests in patients with stroke. J Rehabil Res Dev 2005;42(1):103-107. [doi: 10.1682/jrrd.2003.11.0171] [Medline: 15742254]
43. Caty GD, Arnould C, Stoquart GG, Thonnard JL, Lejeune TM. ABILOCO: a Rasch-built 13-item questionnaire to assess locomotion ability in stroke patients. Arch Phys Med Rehabil 2008 Feb;89(2):284-290. [doi: 10.1016/j.apmr.2007.08.155] [Medline: 18226652]
44. Batcho CS, Tennant A, Thonnard JL. ACTIVLIM-Stroke: a crosscultural Rasch-built scale of activity limitations in patients with stroke. Stroke 2012 Mar;43(3):815-823. [doi: 10.1161/STROKEAHA.111.638965] [Medline: 22223234]
45. Carswell A, McColl MA, Baptiste S, et al. The Canadian Occupational Performance Measure: a research and clinical literature review. Can J Occup Ther 2004 Oct;71(4):210-222. [doi: 10.1177/000841740407100406] [Medline: 15586853]
46. RANKIN J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. Scott Med J 1957 May;2(5):200-215. [doi: 10.1177/003693305700200504] [Medline: 13432835]
47. De Haan R, Horn J, Limburg M, et al. A comparison of five stroke scales with measures of disability, handicap, and quality of life. Stroke 1993 Aug;24(8):1178-1181. [doi: 10.1161/01.str.24.8.1178] [Medline: 8342193]
48. Bouffioulx E, Arnould C, Thonnard JL. SATIS-Stroke: A satisfaction measure of activities and participation in the actual environment experienced by patients with chronic stroke. J Rehabil Med 2008 Nov;40(10):836-843. [doi: 10.2340/16501977-0272] [Medline: 19242621]
49. Horne JC, Lincoln NB, Logan PA. Measurement of confidence: the development and psychometric evaluation of a stroke-specific, measure of confidence. Clin Rehabil 2017 Nov;31(11):1529-1537. [doi: 10.1177/0269215517705424] [Medline: 28466669]
50. Duncan PW, Bode RK, Min Lai S, Perera S, Glycine Antagonist in Neuroprotection Americans Investigators. Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. Arch Phys Med Rehabil 2003 Jul;84(7):950-963. [doi: 10.1016/s0003-9993(03)00035-2] [Medline: 12881816]
51. Edwards B, O'Connell B. Internal consistency and validity of the Stroke Impact Scale 2.0 (SIS 2.0) and SIS-16 in an Australian sample. Qual Life Res 2003 Dec;12(8):1127-1135. [doi: 10.1023/a:1026109920478] [Medline: 14651430]
52. Vellone E, Savini S, Fida R, et al. Psychometric evaluation of the Stroke Impact Scale 3.0. J Cardiovasc Nurs 2015;30(3):229-241. [doi: 10.1097/JCN.0000000000000145] [Medline: 24695074]
53. Duncan PW, Lai SM, Tyler D, et al. Evaluation of proxy responses to the Stroke Impact Scale. Stroke 2002 Nov;33(11):2593-2599. [doi: 10.1161/01.str.0000034395.06874.3e] [Medline: 12411648]
54. Ducoffre E, Arnould C, Somville M, et al. Feasibility of HABIT-ILE@home in children with cerebral palsy and adults with chronic stroke: a pilot study. PLOS Digit Health 2025 May;4(5):e0000850. [doi: 10.1371/journal.pdig.0000850] [Medline: 40338898]
55. Coker-Bolt P, Downey RJ, Connolly J. Exploring the feasibility and use of accelerometers before, during, and after a camp-based CIMT program for children with cerebral palsy. J Pediatr Rehabil Med 2017;10(1):27-36. [doi: 10.3233/PRM-170408] [Medline: 28339408]
56. Sarfo FS, Ulasavets U, Opare-Sem OK, et al. Tele-rehabilitation after stroke: an updated systematic review of the literature. J Stroke Cerebrovasc Dis 2018 Sep;27(9):2306-2318. [doi: 10.1016/j.jstrokecerebrovasdis.2018.05.013] [Medline: 29880211]
57. Cramer SC, Dodakian L, Le V, et al. Efficacy of home-based telerehabilitation vs in-clinic therapy for adults after stroke: a randomized clinical trial. JAMA Neurol 2019 Sep 1;76(9):1079-1087. [doi: 10.1001/jamaneurol.2019.1604] [Medline: 31233135]
58. Barzel A, Ketels G, Stark A, et al. Home-based constraint-induced movement therapy for patients with upper limb dysfunction after stroke (HOMECIMT): a cluster-randomised, controlled trial. Lancet Neurol 2015 Sep;14(9):893-902. [doi: 10.1016/S1474-4422(15)00147-7] [Medline: 26231624]
59. Tziaka E, Tsiakiri A, Vlotinou P, et al. A holistic approach to expressing the burden of caregivers for stroke survivors: a systematic review. Healthcare (Basel) 2024 Feb 29;12(5):565. [doi: 10.3390/healthcare12050565] [Medline: 38470676]
60. Zhao J, Zeng Z, Yu J, et al. Effect of main family caregiver's anxiety and depression on mortality of patients with moderate-severe stroke. Sci Rep 2021;11(1):2747. [doi: 10.1038/s41598-021-81596-8]

XSL•FO
RenderX

61. Maggio MG, Corallo F, De Francesco M, et al. Understanding the family burden and caregiver role in stroke rehabilitation: insights from a retrospective study. Neurol Sci 2024 Nov;45(11):5347-5353. [doi: 10.1007/s10072-024-07668-5] [Medline: 38958795]
62. Ashghali Farahani M, Najafi Ghezeljeh T, Haghani S, et al. The effect of a supportive home care program on caregiver burden with stroke patients in Iran: an experimental study. BMC Health Serv Res 2021 Apr 15;21(1):346. [doi: 10.1186/s12913-021-06340-4] [Medline: 33858400]
63. Otaka Y, Kitamura S, Suzuki M, et al. Effects of rehabilitation program focused on improving real-life daily activities of patients with mild cognitive impairments or dementia and their caregivers. J Rehabil Med Clin Commun 2023;6:12293. [doi: 10.2340/jrmcc.v6.12293] [Medline: 37829668]
64. Ugur HG, Erci B. The effect of home care for stroke patients and education of caregivers on the caregiver burden and quality of life. Acta Clin Croat 2019 Jun;58(2):321-332. [doi: 10.20471/acc.2019.58.02.16] [Medline: 31819329]
65. Essery R, Geraghty AWA, Kirby S, et al. Predictors of adherence to home-based physical therapies: a systematic review. Disabil Rehabil 2017 Mar;39(6):519-534. [doi: 10.3109/09638288.2016.1153160] [Medline: 27097761]
66. Blanton S, Scheibe DC, Rutledge AH, et al. Family-centered care during constraint-induced therapy after chronic stroke: a feasibility study. Rehabil Nurs 2019;44(6):349-357. [doi: 10.1097/rnj.0000000000000197] [Medline: 31688561]
67. Bhalla A, Clark L, Fisher R, et al. The new national clinical guideline for stroke: an opportunity to transform stroke care. Clin Med (Lond) 2024 Mar;24(2):100025. [doi: 10.1016/j.clinme.2024.100025] [Medline: 38387534]
68. O'Flaherty D, Ali K. Recommendations for upper limb motor recovery: an overview of the UK and European Rehabilitation after Stroke Guidelines (2023). Healthcare (Basel) 2024 Jul 18;12(14):1433. [doi: 10.3390/healthcare12141433] [Medline: 39057576]
69. Jack K, McLean SM, Moffett JK, et al. Barriers to treatment adherence in physiotherapy outpatient clinics: a systematic review. Man Ther 2010 Jun;15(3):220-228. [doi: 10.1016/j.math.2009.12.004] [Medline: 20163979]
70. Peek K, Carey M, Mackenzie L, et al. Patient-perceived barriers and enablers to adherence to physiotherapist prescribed self-management strategies. NZJP 2018 Nov;46(3):105-112. [doi: 10.15619/NZJP/46.3.03]
71. Argent R, Daly A, Caulfield B. Patient involvement with home-based exercise programs: can connected health interventions influence adherence? JMIR Mhealth Uhealth 2018 Mar 1;6(3):e47. [doi: 10.2196/mhealth.8518] [Medline: 29496655]

Abbreviations

ACTIVLIM: Activity Limitations Questionnaire
ANCOVA: analysis of covariance
CIMT: constraint-induced movement therapy
FMA: Fugl-Meyer Assessment
FMA-UE: upper extremity subscale of the Fugl-Meyer Assessment
HABIT-ILE: hand and arm bimanual intensive therapy including lower extremities
HADS: Hospital Anxiety and Depression Scale
Mini-BESTest: Mini Balance Evaluation Systems Test
mRS: modified Rankin Scale
NIM: noninferiority margin
PILS-Stroke: Participation In Life Situations – Stroke questionnaire
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SIS: Stroke Impact Scale
WMFT: Wolf Motor Function Test

Edited by J Sarvestan; submitted 04.Nov.2025; peer-reviewed by ES Sogbossi, F Nindorera; revised version received 12.Dec.2025; accepted 15.Dec.2025; published 10.Feb.2026.

Please cite as:
Somville M, Rosselli Z, Ducoffre E, Arnould C, Bleyenheuft Y, Saussez G
Hand and Arm Bimanual Intensive Therapy Including Lower Extremities (HABIT-ILE@home) for Adults With Chronic Stroke: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e87035
URL: https://www.researchprotocols.org/2026/1/e87035
doi:10.2196/87035

XSL•FO
RenderX

© Merlin Somville, Zélie Rosselli, Edouard Ducoffre, Carlyne Arnould, Yannick Bleyenheuft, Geoffroy Saussez. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Efficacy of Repetitive Transcranial Magnetic Stimulation Combined With Auditory Integration Training for Children With Autism Spectrum Disorder: Protocol for a Randomized Sham-Controlled Trial

Qinghong Hao*, MMed; Jinying Wang*, MMed; Jindi Yang*, BSc; Wei Li, MMed; Sufen Hu, BSc; Zhihai Lv, MMed

Department of Child Rehabilitation, Longgang District Maternity & Child Healthcare Hospital of Shenzhen City (Longgang Maternity and Child Institute of Shantou University Medical College), No. 6, Ailong Road, Longgang District, Shenzhen, China

*these authors contributed equally

Corresponding Author:
Zhihai Lv, MMed
Department of Child Rehabilitation, Longgang District Maternity & Child Healthcare Hospital of Shenzhen City (Longgang Maternity and Child Institute of Shantou University Medical College), No. 6, Ailong Road, Longgang District, Shenzhen, China

Abstract

Background: Autism spectrum disorder (ASD) represents a significant public health challenge characterized by persistent social communication deficits and restricted, repetitive patterns of behavior. Current interventions show limited efficacy, particularly for core symptoms. Repetitive transcranial magnetic stimulation (rTMS) and auditory integration training (AIT) have independently demonstrated promise in addressing neurophysiological abnormalities associated with ASD.

Objective: This study aims to evaluate the clinical efficacy of a combined rTMS and AIT intervention compared to rTMS alone and sham stimulation in children with ASD.

Methods: This is a randomized, sham-controlled trial that will recruit 80 children aged 2 to 6 years with a confirmed ASD diagnosis. The randomization of the first 8 participants in this study used a 1:1:1 ratio. To more effectively test the core hypothesis (ie, the efficacy of the combined intervention), greater statistical power will be concentrated on the intervention group (rTMS+AIT), and the randomization ratio was ultimately adjusted to 2:1:1—specifically, (1) rTMS combined with AIT (n=40), (2) rTMS alone (n=20), or (3) sham rTMS (n=20). Primary outcome measures include the Autism Behavior Checklist and Childhood Autism Rating Scale. Secondary outcomes are the Strengths and Difficulties Questionnaire and Repetitive Behavior Scale–Revised. Assessments will be conducted at baseline, an interim time point, and immediately after the intervention. Data will be analyzed using SPSS (version 25.0; IBM Corp).

Results: This study has received funding, with data collection commencing in April 2024. Due to the small initial sample size of 8 participants (5 male and 3 female), no formal statistical comparisons of baseline characteristics between groups have been conducted at this time. It is anticipated that the rTMS combined with AIT intervention will exhibit superior efficacy compared to rTMS only.

Conclusions: This will be the first sham-controlled trial to systematically investigate the potential synergistic effects of a combined rTMS and AIT intervention in children with ASD. The results will provide valuable insights into the neurotherapeutic potential of this combined approach and contribute to the development of evidence-based interventions for ASD.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400082706; https://www.chictr.org.cn/showproj.html?proj=225665

International Registered Report Identifier (IRRID): DERR1-10.2196/80243

(JMIR Res Protoc 2026;15:e80243) doi:10.2196/80243

KEYWORDS

autism spectrum disorder; repetitive transcranial magnetic stimulation; auditory integration training; sham stimulation; randomized controlled trial

Introduction

According to the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, autism spectrum disorder (ASD) is a complex neurodevelopmental condition characterized by impaired social interaction and communication as well as restricted and repetitive patterns of behavior [1,2]. Individuals with ASD may also exhibit abnormal responses to sensory stimuli, unusual emotional behaviors, difficulties with speech and language comprehension, and challenges with self-care and

XSL•FO
RenderX

social adjustment [3]. The prevalence of ASD is a significant public health concern. According to the World Health Organization, approximately 1 in 100 children worldwide is diagnosed with ASD [4]. Data released by the US Centers for Disease Control and Prevention in 2023 showed a sharp increase in the prevalence of autism among children in the United States over the past 20 years [5,6]. Epidemiological survey data from a large sample population in China also indicate a prevalence rate of 0.7% for ASD [7]. Given the growing prevalence of ASD, early assessment, diagnosis, and intervention are crucial. However, the pathogenesis of ASD remains unclear, and there is a lack of specific drug treatments. It is necessary to actively explore effective clinical intervention methods to address this pressing public health challenge.

In recent years, noninvasive brain stimulation, particularly repetitive transcranial magnetic stimulation (rTMS), has been widely applied in the treatment of ASD and has gained recognition among researchers [8,9]. Transcranial magnetic stimulation (TMS) is believed to induce lasting neuroplastic changes in the brain, thereby potentially normalizing social and cognitive functioning in individuals with ASD by stabilizing abnormal neuroplasticity [10]. A meta-analysis by Smith et al [11] of 12 studies demonstrated that TMS improved social withdrawal, repetitive behaviors, and executive function in adolescents with ASD who had an IQ greater than 65, highlighting the significant therapeutic potential of TMS for ASD. Current TMS interventions for ASD primarily target the dorsolateral prefrontal cortex (DLPFC), medial prefrontal cortex, and temporo-parietal junction [12,13]. The DLPFC is involved in multiple higher-order cognitive functions, including emotional regulation, decision-making, attentional control, and adaptive adjustment. In ASD, the core symptoms of repetitive, stereotyped behaviors and social impairments are closely associated with DLPFC dysfunction. Studies using low-frequency rTMS targeting the bilateral DLPFC have demonstrated significant reductions in Repetitive Behavior Scale–Revised (RBS-R) scores following intervention [14,15]. Another study using low-frequency rTMS on the bilateral DLPFC in individuals with ASD showed significant improvements in social behavior after 18 treatment sessions. Similarly, Masuda et al [16] and Casanova et al [17] have indicated that low-frequency rTMS may exert therapeutic effects on social interaction and repetitive behaviors in ASD. In summary, DLPFC dysfunction represents a key neural mechanism underlying the core symptoms of ASD, with particularly pronounced effects on social skills and behavioral regulation. Given its favorable safety profile and emerging evidence of efficacy, rTMS offers a promising intervention for addressing the challenges faced by individuals with ASD.

Studies have shown that most children with ASD experience sensory integration dysfunction [18]. Moreover, the core symptoms of ASD are closely related to abnormal sensory processing, with auditory processing abnormalities being a primary concern [19]. Auditory integration training (AIT), a novel audio therapy initially proposed by French physician Guy Bérard [20], filters out hypersensitive frequencies to reduce auditory cortex sensitivity to sound stimuli, thereby significantly improving auditory attention in children with ASD [21]. Studies indicate that AIT effectively improves language, social interaction, communication, and behavioral skills in children with ASD who also have auditory attention deficits. International studies have explored the physiological effects of AIT on plasma glial cell–derived neurotrophic factor levels. Results demonstrate that AIT significantly elevates plasma glial cell–derived neurotrophic factor levels, thereby improving abnormal sensory processing in affected children [22]. Multiple studies support the efficacy of this method, indicating that when children with ASD can more accurately perceive and process auditory information, their comprehension and expression abilities improve, leading to enhanced functioning in social skills, communication, and related domains [18,23]. However, AIT is considered limited in its capacity to fundamentally activate and repair the nervous system, which may hinder sustained functional recovery [24]. Given this limitation, current research predominantly combines AIT with traditional behavioral training to evaluate treatment outcomes [25,26]. Furthermore, previous studies have often lacked appropriate control groups, necessitating additional research to validate AIT's efficacy. Nevertheless, AIT can be easily implemented in laboratory or clinic settings, facilitating researchers' collection of physiological data for objective assessment of treatment outcomes.

Although preliminary findings from these individual interventions are encouraging, the evidence base for their combined application remains limited and is subject to methodological constraints. To our knowledge, the few existing studies on combined rTMS and AIT lack adequate control conditions and are susceptible to placebo effects and observer bias [27,28]. A critical and pervasive limitation in recent studies of noninvasive brain stimulation interventions for ASD is the absence of high-quality sham-controlled trials [29,30]. This significant methodological gap prevents definitive conclusions about the true synergistic efficacy of these combined interventions. Consequently, there is an urgent need for rigorous, well-controlled research evaluating the synergistic effects of these 2 treatment modalities to provide stronger evidence-based support for ASD interventions.

The primary objective of this study is to investigate whether the combination of rTMS and AIT produces superior effects on brain activity and behavior outcomes in children with ASD compared to sham-controlled conditions. We hypothesize that the combined intervention will yield synergistic benefits, and the findings from this study are expected to provide a scientific foundation for optimizing clinical interventions for ASD.

Methods

Design and Setting

This is a randomized, parallel-group, sham-controlled trial conducted at the Shenzhen Longgang District Maternal and Child Health Hospital. The flowchart of the experimental design is shown in Figure 1, and the schedule of participant assessments is shown in Table 1. Specifically, participants are randomly assigned in a 2:1:1 ratio to 1 of 3 intervention groups: rTMS combined with AIT, rTMS alone, or sham stimulation. Sham stimulation is used via a "sham coil" with an identical

XSL•FO
RenderX

appearance to that of the active TMS coil. Participants, outcome assessors, treating therapists, data analysts, and other study personnel are kept blind to group allocation throughout the trial.

Figure 1. Flowchart of the trial. ABC: Autism Behavior Checklist; AIT: auditory integration training; ASD: autism spectrum disorder; CARS: Childhood Autism Rating Scale; RBS-R: Repetitive Behavior Scale–Revised; rTMS: repetitive transcranial magnetic stimulation; SDQ: Strengths and Difficulties Questionnaire.

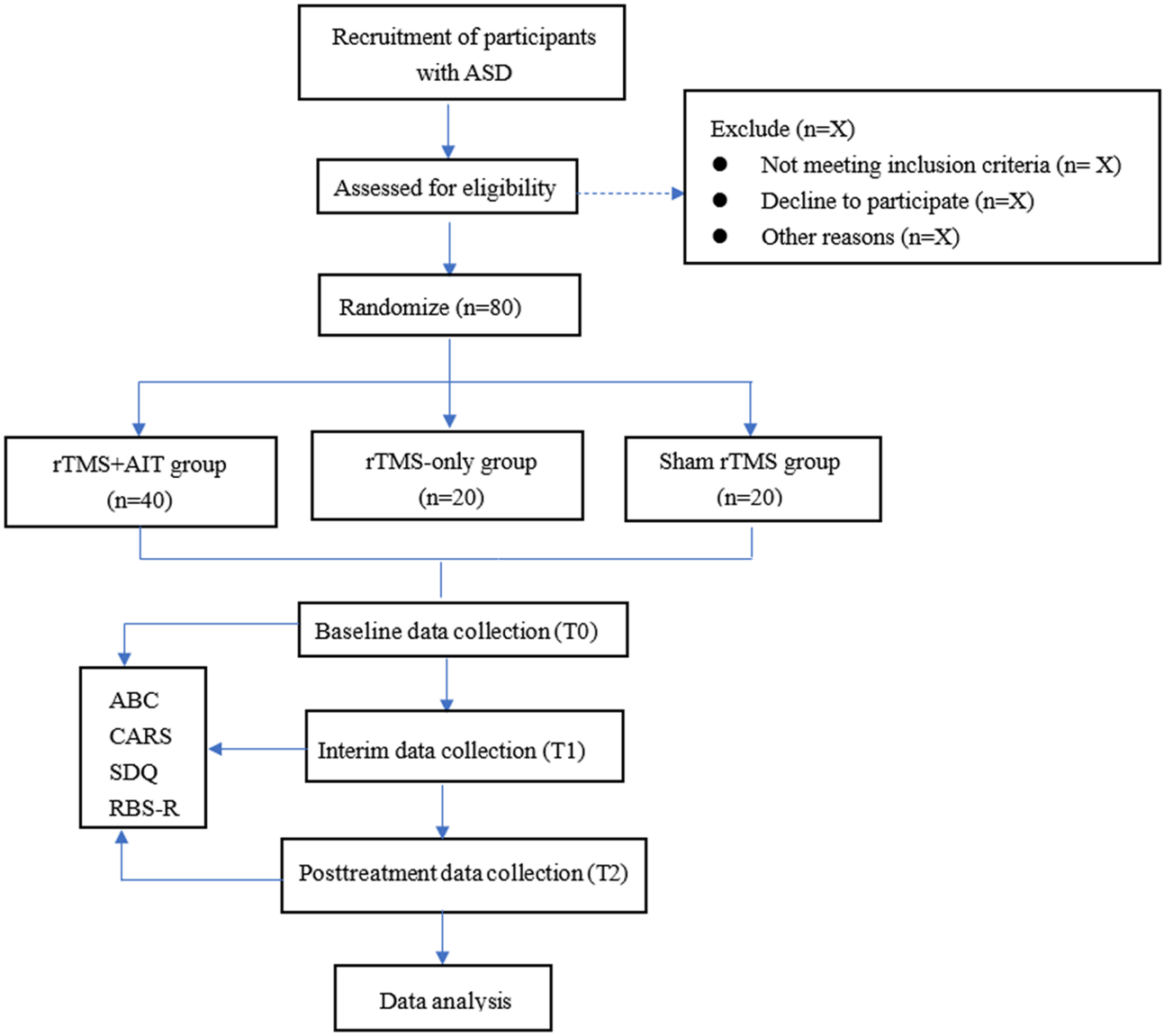

XSL•FO
RenderX

Table . Schedule of the participant assessments.

	Study period									
	Baseline (week –1)	Allocation (week 0; T0)	Treatment phase—left DLPFC[a]			Interim time point (T1)	Treatment phase—right DLPFC			Posttreatment time point (week 13; T2)
			Weeks 1-2	Weeks 3-4	Weeks 5-6		Weeks 7-8	Weeks 9-10	Weeks 11-12	
Enrollment										
Eligibility screen	✓									
Informed consent	✓									
Randomization		✓								
Interventions										
rTMS[b]+AIT[c] group			✓	✓	✓		✓	✓	✓	
rTMS group			✓	✓	✓		✓	✓	✓	
Sham rTMS group			✓	✓	✓		✓	✓	✓	
Assessment										
ABC[d]		✓				✓				✓
CARS[e]		✓				✓				✓
SDQ[f]		✓				✓				✓
RBS-R[g]		✓				✓				✓

[a]DLPFC: dorsolateral prefrontal cortex.

[b]rTMS: repetitive transcranial magnetic stimulation.

[c]AIT: auditory integration training.

[d]ABC: Autism Behavior Checklist.

[e]CARS: Childhood Autism Rating Scale.

[f]SDQ: Strengths and Difficulties Questionnaire.

[g]RBS-R: Repetitive Behavior Scale–Revised.

Participants

We are recruiting participants from inpatient and outpatient populations of the Children's Rehabilitation Department of Shenzhen Longgang District Maternal and Child Health Hospital. Recruitment is being conducted through structured screening procedures. During recruitment, gender is not a restriction. If the proportion is unbalanced, the sample size will need to be increased.

Eligibility Criteria

Inclusion Criteria

Participants who meet the following criteria are included:

- Fulfillment of the diagnostic criteria for ASD according to the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* [31]
- Age of 2 to 6 years (this age range represents the period of greatest neural plasticity in the developing brain) [32]
- No significant abnormalities on routine electroencephalography (EEG) examination
- Normal olfactory, visual, and auditory function, with no other physical or psychiatric comorbidities
- Legal guardian being informed of the study procedures and providing written informed consent

Exclusion Criteria

Participants who meet any of the following criteria are excluded:

- Significant abnormalities on cranial magnetic resonance imaging or severe cardiopulmonary insufficiency
- Children with schizophrenia, epilepsy, deafness, metabolic diseases, or other mental conditions

XSL•FO
RenderX

- Use of psychotropic medications or nutritional supplements for symptom management within the previous 3 months
- Hearing impairments or abnormalities
- Presence of intracranial or peristimulation metallic foreign bodies, cardiac pacemakers, or other implanted electronic devices
- Previous receipt of ASD-specific interventions or comprehensive rehabilitation training

Sample Size

Sample size calculation was based on one-way ANOVA comparing the primary outcome measure (Autism Behavior Checklist [ABC] scores) among the 3 groups. The calculation was performed using the PASS software (version 15; NCSS, LLC) with the following parameters: significance level of α=.05 (2 sided), statistical power of $1 - \beta = 0.90$, and number of groups of k=3. On the basis of previous similar studies [33,34], the within-group SD is set at σ=6.5, yielding a medium to large effect size (Cohen *f*). The calculation indicated that a minimum sample size of 15 participants per group was required. Accounting for an anticipated dropout rate of approximately 20%, a total of 20 participants will be recruited for each group. The following standard ANOVA sample size formula was applied: $n = \psi^2 \times (\sigma^2) / (k \times d^2)$.

Randomization and Allocation Concealment

A researcher (WL) who is not involved in participant recruitment or treatment uses SPSS (version 25.0; IBM Corp) to randomly allocate all eligible participants to 3 groups. The sequence number and group assignment are placed in consecutively numbered, opaque, sealed envelopes. When a participant is enrolled and found to meet the inclusion criteria, the next available envelope is opened to reveal the group allocation, which is then implemented. To ensure balance in terms of important prognostic factors such as age and gender across the intervention groups after randomization, a stratified randomization approach has been adopted. All other researchers involved in this study are blinded to the treatment allocations.

Intervention Protocol

All participants receive conventional comprehensive rehabilitation treatment, including speech and cognitive training, social skill training, occupational therapy, sensory integration training, and music therapy [35,36], for 30 minutes per session once daily, 5 times per week, for 12 weeks in total.

rTMS-Only Group

The rTMS intervention is administered using a CCY-1 TMS device (Yiruide Medical Equipment New Technology Co., Ltd) with a circular coil. Before the first treatment session, the resting motor threshold (RMT) is determined by identifying the cortical representation of the contralateral abductor pollicis brevis muscle within the primary motor cortex (M1). The optimal stimulation site is defined as the location that consistently produces the maximum amplitude (>50 µV) and shortest latency motor-evoked potential in the relaxed abductor pollicis brevis muscle. RMT is defined as the minimum stimulation intensity required to induce motor-evoked potentials of ≥50 µV in at least 5 of 10 consecutive trials [37,38].

During treatment sessions, participants are seated and wear a positioning cap. The bilateral DLPFC is targeted sequentially: the left DLPFC during weeks 1 to 6 and the right DLPFC during weeks 7 to 12 [2]. This sequencing is based on the hypothesis of hemispheric excitation-inhibition imbalance in ASD [39]. The frequency is set to 1 Hz, and the stimulation intensity is set to 80% the RMT with 400 pulses each time at a 20-second interval every 10 pulses 5 times a week for a total of 12 weeks as a course of treatment. Each session lasts approximately 20 minutes. Following rTMS, participants listen to unmodulated natural environmental sounds or music through headphones for an additional 20 minutes to match the total treatment duration of the combined intervention.

rTMS Combined With AIT Group

In addition to the standard rehabilitation treatment, participants in this group receive a combination of rTMS and AIT. For the first 6 weeks, the left DLPFC is targeted, followed by right DLPFC stimulation in the subsequent 6 weeks. The 2 treatments are administered simultaneously, with each session consisting of 20 minutes of rTMS followed by 20 minutes of AIT for a total session duration of 40 minutes. The rTMS protocol matches the parameters used in the rTMS-only group.

The AIT intervention is delivered using the Elite RT510 AIT instrument (Shenzhen Elite Medical Equipment Co, Ltd), a digital integrated therapy device. Participants wear headphones during the AIT sessions. The operator selects the appropriate audio filtration settings based on the participant's age (distinguishing between those over and under 4 years of age) and then plays a series of filtered and modulated music CDs according to the established AIT protocol [40]. The treatment process involves precise control of sound intensity ("sound level") and attenuation values following a fixed playback sequence. All equipment parameters are reset after each session.

Sham rTMS Group

Participants in this group receive conventional comprehensive rehabilitation treatment as well as sham rTMS stimulation using a "sham coil" with the same appearance as that of the active TMS coil. The sham stimulation parameters, including sequence, sound, and frequency, match those of the rTMS-only group. Additionally, during the 20 minutes following the sham rTMS, participants listen to unmodulated and unfiltered natural environmental sounds or music, similar to the rTMS-only group.

Outcome Measures

Primary Outcome Measures

ABC Measure

This 57-item scale [41] completed by parents is suitable for screening and assisting in the diagnosis of ASD in individuals aged 8 months to 28 years. It evaluates behavioral symptoms across 5 factors: sensory abilities, communication abilities, motor skills, language, and self-care. Each item is scored on a 4-point scale. A total score of ≥53 indicates a screening cutoff for ASD, whereas a total score of ≥67 indicates a positive diagnosis of ASD.

XSL•FO
RenderX

Childhood Autism Rating Scale

This 15-item scale [42] is used to assess symptom severity in children with ASD aged ≥2 years. Qualified personnel will conduct the assessments, with each item scored from 1 to 4. A total score of 30 to 36 with fewer than 5 items with a score below 3 indicates mild to moderate ASD; a total score of >36 with at least 5 items with a score above 3 indicates severe ASD.

Secondary Outcome Measures

Strengths and Difficulties Questionnaire

Developed by Dr Robert Goodman [43], this 25-item questionnaire is used to assess emotional and behavioral problems in individuals aged 2 to 17 years, including those with ASD. The Strengths and Difficulties Questionnaire (SDQ) evaluates 4 areas of difficulty: emotional symptoms, conduct problems, hyperactivity, and peer relationship problems, as well as prosocial behavior. Parents rate their children's emotional and behavioral performance over the previous 6 months on a 3-point Likert scale (0="not true," 1="somewhat true," and 2="certainly true"). Higher scores on the difficulty scales indicate more severe problems, whereas higher scores on the prosocial scale indicate better prosocial behavior. The Cronbach α coefficient is 0.78, demonstrating good reliability and validity.

RBS-R Measure

This 43-item scale [44] comprises 6 core components: stereotyped behavior, self-injurious behavior, compulsive behavior, ritualistic behavior, fixated behavior, and restricted behavior. Participants' parents complete the RBS-R, with higher total scores indicating more severe stereotyped behaviors. The Cronbach α coefficient is 0.85. The Chinese version of this scale has shown good reliability and validity, making it an effective tool for evaluating treatment effects on repetitive behaviors in children with ASD.

Data Management

Clinical data are collected using a case report form (CRF) provided exclusively by the research team. The CRF serves as the original record and will not be altered. One member of the research team (QH) is responsible for recording and maintaining the CRFs. Upon completion of the trial, 2 researchers (JW and JY) will independently enter the CRF data into the computer for statistical analysis.

Ethical Considerations

This study has been reviewed and approved by the ethics committee of Longgang Maternal and Child Health Hospital of Shenzhen, China (KYXM-2023-032), in accordance with the Declaration of Helsinki. Parents or legal guardians of child patients who meet the inclusion criteria are informed of the study's purpose, specific procedures, and potential risks and possible benefits of participation. Those who choose to participate then provide written informed consent. Participants' personal data will be kept confidential throughout the study and accessible only to the research team. There will be no compensation provided to participants, and no identifiable images of individual participants will be included in the manuscript or supplementary materials.

Statistical Analysis

Statistical analyses will be conducted using SPSS. All participants, including the first 8 randomized under the original scheme, will be analyzed according to the intention-to-treat principle. At baseline, continuous variables among the 3 groups will be compared using the Kruskal-Wallis test, whereas categorical variables will be compared using the chi-square test to assess the balance of demographic and clinical characteristics (including age, gender, and baseline scores on all outcome measures) across the randomized groups.

Data normality will be assessed, and appropriate statistical tests will be selected accordingly. For normally distributed measurement data, between-group comparisons will be conducted using one-way ANOVA, and within-group pretest-posttest treatment comparisons will be made using paired-sample 2-tailed *t* tests presented as means and SDs. For nonnormally distributed data, nonparametric tests will be used, with results reported as medians and IQRs. $P<.05$ will be considered statistically significant.

In addition to the primary end-point comparisons at study completion, an exploratory interim analysis of outcomes at 6 weeks will be performed to preliminarily assess the effects of left DLPFC stimulation.

Protocol Amendment

This study randomized the first 8 participants in a 1:1:1 ratio. Subsequently, to more effectively test the core hypothesis (ie, the efficacy of the combined intervention), greater statistical power will be concentrated on the intervention group (rTMS+AIT), and the randomization ratio was ultimately adjusted to 2:1:1.

Results

Overview

This study commenced data collection in April 2024. It is anticipated that the rTMS combined with AIT intervention will exhibit superior efficacy compared to rTMS only.

Currently, 8 participants (5 male and 3 female) have been recruited and randomized. Due to the limited sample size, this study has not yet conducted statistical comparisons of baseline characteristics among the 3 groups. The pooled baseline characteristics of the enrolled participants are as follows: mean age of 4.12 years, mean total ABC score of 53.38, and mean total Childhood Autism Rating Scale (CARS) score of 35.25. Recruitment and follow-up of all participants are ongoing.

Study Timeline and Status

The detailed time points and status of this study are shown in Table 2.

XSL•FO
RenderX

Table . Study timeline and status notes.

Period	Research activities	Status
January 2024-March 2024	Final protocol development and ethics approval	Completed
April 2024	Recruitment and randomization commenced	From the first participant onward, we consistently used a gender- and age-stratified randomization method.
May 2024-September 2025	Data collection and intervention; protocol refinement	Ongoing. All participants have received interventions per the established protocol. Interim assessments (6 weeks) and baseline evaluations as predefined assessment points have been conducted since study initiation. The first 8 participants were randomized at the initial 1:1:1 ratio. To enhance statistical power for effect size and variability in the combined intervention group (rTMS[a]+AIT[b]), the ratio was ultimately adjusted to 2:1:1.
October 2025-March 2026	Data collection and intervention; recruitment completion	Ongoing; strictly following protocol
March 2026-June 2026	End point assessment; data collation and analysis	Future work

[a]rTMS: repetitive transcranial magnetic stimulation.

[b]AIT: auditory integration training.

Discussion

We hypothesize that, compared to the sham stimulation control group, participants with ASD receiving the active rTMS combined with AIT intervention will exhibit significant improvements in core symptoms such as social difficulties and stereotyped behaviors as measured using standardized scales such as the ABC, CARS, RBS-R, and SDQ. This study aims to provide the latest evidence-based support for the efficacy of this combined intervention approach.

Although preliminary studies have reported positive effects of rTMS combined with AIT in ASD, most have lacked blinding and placebo controls. By implementing a double-blind, sham-controlled design, this trial will address that key methodological limitation. If the results align with our hypotheses, they will provide the first clear demonstration of the specific therapeutic benefits of this combined intervention, thereby advancing the evidence level in this field from the exploratory stage to preliminary validation.

The core strength of this study lies in its rigorous methodological design. As a prospective randomized controlled trial, it uses a double-blind, sham stimulus–controlled approach, which minimizes the potential for expectation and observer biases during the evaluation process, thereby strengthening the causal inferences that can be drawn from the observed efficacy differences. Additionally, the study uses recognized, standardized assessment tools, including the ABC, CARS, SDQ, and RBS-R, to provide a comprehensive and objective evaluation of the core symptom domains in ASD. This assessment combination ensures good reliability and clinical relevance of the study results. Finally, the rTMS and AIT interventions are noninvasive, well tolerated, and easy to implement in clinical settings. These features support participant compliance and trial feasibility and, if the interventions prove effective, will facilitate wider clinical translation as a safe, scalable treatment option for children with ASD.

This study also has some important limitations. First, while rigorous behavioral assessments are used, neurophysiological indicators (eg, EEG and functional near-infrared spectroscopy) are not included. Therefore, this study cannot directly reveal the underlying neural mechanisms of the combined intervention. Future studies should verify these mechanisms in conjunction with neuroimaging techniques. Second, the primary objective is to preliminarily evaluate whether adding AIT to rTMS could yield superior clinical efficacy. As a result, a separate control group receiving AIT alone was not established. Subsequent studies will need to expand the sample size to explore this question. Additionally, the follow-up period after the intervention may be insufficient to assess the long-term durability of the effects, so our conclusions will be limited to discussing the short-term efficacy.

On the basis of the expected results and limitations of this study, future research should focus on the following directions. First, at the mechanistic level, follow-up studies should integrate neuroimaging techniques such as EEG, functional near-infrared spectroscopy, or functional magnetic resonance imaging to directly observe the effects of the combined intervention on brain network connectivity and neural activity. Second, at the level of efficacy optimization, trials with larger sample sizes, longer follow-up periods (eg, 6 months or 1 year), and inclusion of an AIT-only group will be needed to verify the durability of the intervention's benefits and clarify the individual vs additive benefits of rTMS and AIT.

XSL•FO
RenderX

Acknowledgments

The authors acknowledge the help and contributions of the experts, therapists, and researchers who helped design this study. They thank all patients for participating in this study.

Funding

This research was supported by Shenzhen Longgang District Science and Technology Innovation Special Funds (LGWJ2022-44, LGWJ2022-52, and LGWJ2022-51), Shenzhen Longgang District Science and Technology Innovation Special Fund Medical and Health Technology Research Support Project (LGKCYLWS2024-17) and the 2023 Annual Research Project Plan of the Chinese Rehabilitation Medical Association (KFKT-2023-034).

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Authors' Contributions

ZL and QH are responsible for this study. QH, JW, and JY conceived and designed the study. QH, JW, and JY drafted the manuscript of this trial protocol. WL and SH participated in recruitment and data collection. JW and ZL were responsible for treatment. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT 2025 checklist.
[DOCX File, 31 KB - resprot_v15i1e80243_app1.docx]

References

1. Gnanavel S, Robert RS. Diagnostic and statistical manual of mental disorders, fifth edition, and the impact of events scale-revised. Chest 2013 Dec;144(6):1974. [doi: 10.1378/chest.13-1691] [Medline: 24297138]
2. Kang JN, Song JJ, Casanova MF, Sokhadze EM, Li XL. Effects of repetitive transcranial magnetic stimulation on children with low-function autism. CNS Neurosci Ther 2019 Nov;25(11):1254-1261. [doi: 10.1111/cns.13150] [Medline: 31228356]
3. Dandan L, Yuan L, Min Y, Zewei H. Research trends and hotspots in emotional interventionfor children with autism in China in the past 20 years: visual analysis based on CiteSpace. Psychological Monthly 2025;20(1) [FREE Full text] [doi: 10.19738/j.cnki.psy.205.01.004]
4. Zeidan J, Fombonne E, Scorah J, et al. Global prevalence of autism: a systematic review update. Autism Res 2022 May;15(5):778-790. [doi: 10.1002/aur.2696] [Medline: 35238171]
5. Goodwin A, Matthews NL, Smith CJ. Parent-reported early symptoms of autism spectrum disorder in children without intellectual disability who were diagnosed at school age. Autism 2019 Apr;23(3):770-782. [doi: 10.1177/1362361318777243] [Medline: 29852752]
6. Samadi SA. Overview of services for autism spectrum disorders (ASD) in low- and middle-income countries (LMICs) and among immigrants and minority groups in high-income countries (HICs). Brain Sci 2022 Dec 8;12(12):1682. [doi: 10.3390/brainsci12121682] [Medline: 36552142]
7. Zhou H, Xu X, Yan W, et al. Prevalence of autism spectrum disorder in China: a nationwide multi-center population-based study among children aged 6 to 12 years. Neurosci Bull 2020 Sep;36(9):961-971. [doi: 10.1007/s12264-020-00530-6] [Medline: 32607739]
8. Cole EJ, Enticott PG, Oberman LM, et al. The potential of repetitive transcranial magnetic stimulation for autism spectrum disorder: a consensus statement. Biol Psychiatry 2019 Feb 15;85(4):e21-e22. [doi: 10.1016/j.biopsych.2018.06.003] [Medline: 30103951]
9. McPartland JC, Bernier RA, Jeste SS, et al. The Autism Biomarkers Consortium for Clinical Trials (ABC-CT): scientific context, study design, and progress toward biomarker qualification. Front Integr Neurosci 2020;14:16. [doi: 10.3389/fnint.2020.00016] [Medline: 32346363]
10. Desarkar P, Rajji TK, Ameis SH, Daskalakis ZJ. Assessing and stabilizing aberrant neuroplasticity in autism spectrum disorder: the potential role of transcranial magnetic stimulation. Front Psychiatry 2015;6:124. [doi: 10.3389/fpsyt.2015.00124] [Medline: 26441685]
11. Smith JR, DiSalvo M, Green A, et al. Treatment response of transcranial magnetic stimulation in intellectually capable youth and young adults with autism spectrum disorder: a systematic review and meta-analysis. Neuropsychol Rev 2023 Dec;33(4):834-855. [doi: 10.1007/s11065-022-09564-1] [Medline: 36161554]

XSL•FO
RenderX

12. Jin J, Wang XQ, Yang X, et al. Abnormal individualized peak functional connectivity toward potential repetitive transcranial magnetic stimulation treatment of autism spectrum disorder. Hum Brain Mapp 2023 Nov;44(16):5450-5459. [doi: 10.1002/hbm.26455] [Medline: 37694907]
13. Enticott PG, Barlow K, Guastella AJ, et al. Repetitive transcranial magnetic stimulation (rTMS) in autism spectrum disorder: protocol for a multicentre randomised controlled clinical trial. BMJ Open 2021 Jul 7;11(7):e046830. [doi: 10.1136/bmjopen-2020-046830] [Medline: 34233985]
14. Sokhadze EM, Lamina EV, Casanova EL, et al. Exploratory study of rTMS neuromodulation effects on electrocortical functional measures of performance in an oddball test and behavioral symptoms in autism. Front Syst Neurosci 2018;12:20. [doi: 10.3389/fnsys.2018.00020] [Medline: 29892214]
15. Casanova MF, Sokhadze EM, Casanova EL, Li X. Transcranial magnetic stimulation in autism spectrum disorders: neuropathological underpinnings and clinical correlations. Semin Pediatr Neurol 2020 Oct;35:100832. [doi: 10.1016/j.spen.2020.100832] [Medline: 32892959]
16. Masuda F, Nakajima S, Miyazaki T, et al. Clinical effectiveness of repetitive transcranial magnetic stimulation treatment in children and adolescents with neurodevelopmental disorders: a systematic review. Autism 2019 Oct;23(7):1614-1629. [doi: 10.1177/1362361318822502] [Medline: 30663323]
17. Casanova MF, Sokhadze EM, Casanova EL, et al. Translational neuroscience in autism: from neuropathology to transcranial magnetic stimulation therapies. Psychiatr Clin North Am 2020 Jun;43(2):229-248. [doi: 10.1016/j.psc.2020.02.004] [Medline: 32439019]
18. Li Y, Huan L, Han-Yu W, He L, Jia-Chen W, Hong-Chao Z. Clinical efficacy of rTMS combined with sensory integration training on children with autism spectrum disorder. Hainan Med J 2023;34(1):49-53 [FREE Full text]
19. Siemann JK, Veenstra-VanderWeele J, Wallace MT. Approaches to understanding multisensory dysfunction in autism spectrum disorder. Autism Res 2020 Sep;13(9):1430-1449. [doi: 10.1002/aur.2375] [Medline: 32869933]
20. Turri S, Rizvi M, Rabini G, Melonio A, Gennari R, Pavani F. Orienting auditory attention through vision: the impact of monaural listening. Multisens Res 2021 Aug 11;35(1):1-28. [doi: 10.1163/22134808-bja10059] [Medline: 34384046]
21. Huang Q, Zhou K, Liu Z, Yang J. Effect of auditory integration training on children with autism spectrum disorder and auditory attention disorder. Rehabil Med 2022 Aug 1;32(4):314-319. [doi: 10.3724/SP.J.1329.2022.04005]
22. Al-Ayadhi L, El-Ansary A, Bjørklund G, Chirumbolo S, Mostafa GA. Impact of auditory integration therapy (AIT) on the plasma levels of human glial cell line-derived neurotrophic factor (GDNF) in autism spectrum disorder. J Mol Neurosci 2019 Aug;68(4):688-695. [doi: 10.1007/s12031-019-01332-w] [Medline: 31073917]
23. Jianqing F, Qingzheng J, Lichao X. Effects of repetitive transcranial magnetic stimulation combined with multi-sensory training on perception ability, psychological impairment, and social ability in children with autism spectrum disorder. J Pathol Clin Res 2023;43(3):502-507 [FREE Full text]
24. Minghe Z. Application effect of rTMS combined with auditory integration training in children with autism spectrum disorder. Chin J Med Innov 2024;21(6):133-137 [FREE Full text]
25. Juan C, Yang Z, Yi-xin C. Effect of auditory integration training and applied behavior analysis on autism spectrum disorder. J Clin Psychiatry 2021;31(5):370-373 [FREE Full text]
26. Xiaoli H, Chunyun W, Jianhui X. Intervention study of auditory integration trainer combined with structured instruction in behavior management of children with autism. Heilongjiang Med J 2023;47(13):1580-1583 [FREE Full text]
27. Ruxia Q, Chuizhou Z, Zhen C, Dongwei X. Efficacy evaluation of high-frequency repetitive transcranial magnetic stimulation combined with auditory integration training for children with autism spectrum disorder. Chronic Pathematol J 2024;25(6):909-913 [FREE Full text]
28. Yu-zhen Z. Effects of repeated transcranial magnetic stimulation combined with auditory integration on childhood autism rating scale score and ABC score of children with autism. Chin Med Devices Inform 2025;31(5):120-122. [doi: 10.15971/j.cnki.cmdi.2025.05.050]
29. Oberman LM, Francis SM, Lisanby SH. The use of noninvasive brain stimulation techniques in autism spectrum disorder. Autism Res 2024 Jan;17(1):17-26. [doi: 10.1002/aur.3041] [Medline: 37873560]
30. López-Rodríguez S, Coelho DR, Renet C, et al. Noninvasive brain stimulation for neurodevelopmental disorders: a systematic review. J Neuropsychiatry Clin Neurosci 2025;37(4):297-312. [doi: 10.1176/appi.neuropsych.20240127] [Medline: 40350963]
31. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Association; 2015. [doi: 10.1176/appi.books.9780890425596]
32. Petanjek Z, Sedmak D, Džaja D, Hladnik A, Rašin MR, Jovanov-Milosevic N. The protracted maturation of associative layer IIIC pyramidal neurons in the human prefrontal cortex during childhood: a major role in cognitive development and selective alteration in autism. Front Psychiatry 2019;10:122. [doi: 10.3389/fpsyt.2019.00122] [Medline: 30923504]
33. Mengqing L, zhimei J. rTMS and neurofeedback treatment on stereotyped behavior of children with autism spectrum disorder. Chin J Rehabil 2018;33(2):114-117 [FREE Full text]
34. Yan L, Yuanyue Z, Guoling Z, Kaijing D, Shaohua W, Chuang X. A meta-analysis of transcranial magnetic stimulation combined with rehabilitation training in the treatment of loneliness spectrum disorder. J Mod Pract Med 2021;33(6):726-729 [FREE Full text]

XSL•FO
RenderX

35. Jiayi Z, Qing S, Haibei L, Xiaoxia L. Application effects of repetitive transcranial magnetic stimulation combined with auditory integration training in children with autism. Med J Chin People Health 2024;36(21):80-83 [FREE Full text]
36. Xiaoxiao Z, Baoyuan D, Kuo W. Effects of auditory integration training combined with repetitive transcranial magnetic stimulation on behavioral symptoms in children with autism spectrum disorder. Chin J Rehabil 2024;39(10):599-603 [FREE Full text]
37. Mills KR, Nithi KA. Corticomotor threshold to magnetic stimulation: normal values and repeatability. Muscle Nerve 1997;20(5):570-576. [doi: 10.1002/(sici)1097-4598(199705)20:5<570::aid-mus5>3.0.co;2-6] [Medline: 9140363]
38. Yi X, Da L, Liwen T, Weiqing L, Chuan S, Jianming F, et al. Chinese experts consensus on repetitive transcranial magnetic stimulation. Transl Med J 2018;7(1):4-9 [FREE Full text]
39. Yuan LX, Wang XK, Yang C, et al. A systematic review of transcranial magnetic stimulation treatment for autism spectrum disorder. Heliyon 2024 Jun 15;10(11):e32251. [doi: 10.1016/j.heliyon.2024.e32251] [Medline: 38933955]
40. Lanya L. Effects of auditory integration training on behavior and cognition of children with mental retardation. Med Public Health 2016:47 [FREE Full text]
41. Jianping L, Zhiwei Y, Mingyao S, Linyan S. Reliability, validity analysis of the childhood autism rating scale. China J Mod Med 2004;014(13):119-121 [FREE Full text]
42. Leekam S, Tandos J, McConachie H, et al. Repetitive behaviours in typically developing 2-year-olds. J Child Psychol Psychiatry 2007 Nov;48(11):1131-1138. [doi: 10.1111/j.1469-7610.2007.01778.x] [Medline: 17995489]
43. Jianhua K, Yasong D, Liming X. Reliability and validity of “children strengths and difficulties questionnaire” in Shanghai norm. Shanghai Arch Psychiatry 2005:25-28 [FREE Full text]
44. Jingya L, Zhimei J, Guixia C, Lanmin G, Junli Z. Reliability and validity of the Chinese version of the Repetitive Behavior Scale-Revised (RBS-R) in children with autism. Chin Pediatric Integrat Trad West Med 2013;5(3):4 [FREE Full text]

Abbreviations

ABC: Autism Behavior Checklist
AIT: auditory integration training
ASD: autism spectrum disorder
CARS: Childhood Autism Rating Scale
CRF: case report form
DLPFC: dorsolateral prefrontal cortex
EEG: electroencephalography
RBS-R: Repetitive Behavior Scale–Revised
RMT: resting motor threshold
rTMS: repetitive transcranial magnetic stimulation
SDQ: Strengths and Difficulties Questionnaire
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TMS: transcranial magnetic stimulation

Edited by A Schwartz; submitted 07.Jul.2025; peer-reviewed by R Vani Lakshmi, S Rampersad; accepted 30.Dec.2025; published 11.Feb.2026.

Please cite as:
Hao Q, Wang J, Yang J, Li W, Hu S, Lv Z
Evaluating the Efficacy of Repetitive Transcranial Magnetic Stimulation Combined With Auditory Integration Training for Children With Autism Spectrum Disorder: Protocol for a Randomized Sham-Controlled Trial
JMIR Res Protoc 2026;15:e80243
URL: https://www.researchprotocols.org/2026/1/e80243
doi:10.2196/80243

© Qinghong Hao, Jinying Wang, Jindi Yang, Wei Li, Sufen Hu, Zhihai Lv. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Safety and Performance of the KidneYou App for Chronic Kidney Disease: Protocol for an Italian Multicenter, Randomized, Open-Label, Premarket Study

Cesira Cafiero[1], PhD; Marco Gnesi[2], PhD; Anna Rita Maurizi[3], MD, PhD; Andreana Foresta[4], MSc; Massimo Beccaria[4], MSc; Giorgia Campilongo[5], MD; Marco Fiorentino[5], MD, PhD; Paola Maria Acquaviva[1], MSc; Francesca Mastromauro[6], MSc; Marco Gorini[7], MSc, EMBA; Loreto Gesualdo[5], MD

[1]Nephrology, Dialysis and Transplantation Unit of the University Hospital Consortium, Bari, Italy

[2]Medical Evidence, Biopharmaceuticals Medical, AstraZeneca, Milan, Italy

[3]Medical Affairs, Biopharmaceuticals Medical, AstraZeneca, Milan, Italy

[4]Advice Pharma Group Srl, Milan, Italy

[5]Department of Precision and Regenerative Medicine and Ionian Area (DiMePre-J), Nephrology Dialysis and Transplantation Unit, University of Bari Aldo Moro, Piazza Umberto I, Bari, BA, Italy

[6]Medical Evidence EUCAN, Biopharmaceuticals Medical, AstraZeneca, Milan, Italy

[7]Innovation Unit & Business Excellence, AstraZeneca, Milan, Italy

Corresponding Author:
Loreto Gesualdo, MD
Department of Precision and Regenerative Medicine and Ionian Area (DiMePre-J), Nephrology Dialysis and Transplantation Unit, University of Bari Aldo Moro, Piazza Umberto I, Bari, BA, Italy

Abstract

Background: Chronic kidney disease (CKD) is characterized by long-term structural or functional kidney abnormalities, often progressing over decades and potentially leading to kidney failure, treatable only by dialysis or transplantation. Nutritional programs (NPs), physical activity (PA) programs, and mindfulness programs (MPs) can play a key role in conservative CKD management, aiming to slow the progression of symptoms, decrease drug load, reduce stress, and delay dialysis. KidneYou (Advice Pharma Group Srl) integrates nutrition, exercise, and MPs to improve the health of patients with CKD by promoting lifestyle changes and stress reduction. As an innovative medical tool, KidneYou is designed to address an unmet need by providing a nonpharmacological approach for better managing CKD, empowering patients to manage their condition more effectively and sustain a healthier lifestyle, which could ultimately lead to improved disease outcomes.

Objective: This study aims to evaluate the efficacy and safety of KidneYou, a digital medical device that delivers personalized NPs, PA programs, and MPs to patients with CKD, in comparison with the standard of care.

Methods: This is a multicenter, open-label, randomized, parallel-arm trial aimed at evaluating health improvement in patients with CKD exposed to a nonpharmacological treatment involving an NP, PA program, and MP delivered to the patient by digital technology (investigational arm) or a standard approach (paper diary and control arm) for 3 months. The primary aim of the study is to evaluate the efficacy of KidneYou, defined by the achievement of at least 1 of the 3 conditions described in the composite primary end point: a reduction of at least 10% of azoturia (gram per 24 h), an increase of at least 15% of distance (meters in the 6-min walk test), or a decrease of at least 10% of perceived stress in CKD KidneYou app users compared to CKD KidneYou app nonusers after 3 months.

Results: Data will be analyzed and presented in accordance with international CONSORT (Consolidated Standards of Reporting Trials) guidelines. Recruitment began in July 2022 and stopped at all sites in April 2024. Data analysis is currently ongoing. The results are expected to be published in early 2026.

Conclusions: The results of this trial will guide further studies on the impact of digital devices to help patients with CKD in improving their condition more effectively and sustain a healthier lifestyle, which could ultimately lead to improved disease outcomes.

Trial Registration: ClinicalTrials.gov NCT05286632; https://clinicaltrials.gov/study/NCT05286632

International Registered Report Identifier (IRRID): DERR1-10.2196/75306

(JMIR Res Protoc 2026;15:e75306) doi:10.2196/75306

XSL•FO
RenderX

KEYWORDS

eHealth; digital therapeutics; mHealth; mobile health; lifestyle intervention.; chronic kidney disease

Introduction

Chronic kidney disease (CKD) is a progressive condition defined by structural or functional abnormalities of the kidneys with health implications, persisting for over 3 months [1]. It is classified based on cause, glomerular filtration rate (GFR) category, and albuminuria category and includes various disorders that affect kidney structure and function, with diverse clinical presentations depending on the cause, severity, and progression rate. While some forms of CKD progress rapidly, most evolve over decades. Kidney failure, the most serious outcome, can only be treated with dialysis or transplantation [1]. Moreover, CKD is linked to an increased risk of acute kidney injury, which can worsen outcomes and accelerate CKD progression. Early identification and referral for specialized care can improve both clinical and economic outcomes [1].

Adherence to lifestyle interventions in patients with CKD presents persistent challenges, as patients often contend with physical, emotional, and cognitive barriers that impede consistent engagement with recommended dietary, exercise, and stress management routines. Digital health tools, particularly mobile apps, represent a promising solution to improve adherence by offering tailored, accessible support. With smartphone use increasingly widespread, these platforms can provide patients with CKD with real-time, personalized guidance, enhancing both engagement and self-management.

The nutritional program (NP) is an important component of the conservative management of patients with CKD, which precedes and integrates the pharmacological therapies. Generally, the objectives of NP are to maintain an optimal nutritional state, prevent or correct the signs, symptoms, and complications related to chronic renal insufficiency, and delay the start of dialysis. In addition, NP enables a reduction in the drug load (and related side effects and interactions) and can allow for safe and effective use of lower doses of dialysis, even when the GFR continues to diminish [2]. The aim of NP is to provide patients with CKD with an adequate caloric intake, control the sodium and potassium intake, and reduce phosphate intake. Beyond the quantitative aspects, dietary support also calls for modification in the quality of food, favoring foods of plant origin, which induce favorable effects on phosphate metabolism and acid-base balance, with better control of arterial pressure and renal hemodynamics [2,3]. Physical exercise in patients with CKD is another nonpharmacological tool aimed at improving health in this patient population. Despite large scientific evidence produced over the years about exercise tolerance and training among patients on dialysis, individuals with nondialysis-dependent CKD have been relatively understudied, perhaps because of the heterogeneity of the CKD population [4,5]. The clinical evidence produced for individuals with CKD is limited in both quality and quantity. Data from uncontrolled studies and small randomized controlled trials showed that exercise training results in improved physical performance and functioning. Of note, although the current literature does not allow for drawing definitive conclusions about the role of exercise training on the progression of CKD, no study has reported worsening of kidney function as a result of exercise training. Exercise appears to be safe in patients with CKD if begun at moderate intensity and increased gradually. The evidence suggests that the risk of remaining inactive is greater [6].

The mindfulness program (MP) aims to reduce stress perceived by patients with CKD and to improve their coping capacities in response to the psychological burden caused by living with an active chronic disease, the complexity of ongoing therapies, including concomitant ones, and the lifestyle changes required by dietary and exercise programs. Scientific evidence has shown that, when a stress-coping program based on mindfulness meditation is applied to stressed individuals, a significant reduction in perceived stress, anxiety, and depression is observed [7-9].

While the benefits of nutritional counseling, physical activity (PA) promotion, and behavioral support have been extensively demonstrated as stand-alone interventions, growing evidence suggests that integrating them in multicomponent lifestyle programs, particularly when supported by digital tools, may enhance patient engagement and lead to more sustained health improvements over time [10]. Despite the growing availability of digital tools, few interventions specifically designed for patients with CKD combine nutritional guidance, PA promotion, and stress management into a single, integrated platform [11].

The KidneYou app (Advice Pharma Group Srl) is an investigational medical device intended to administer these 3 nonpharmacological programs simultaneously to patients with CKD, to monitor patients' performance and adherence during the study, and to improve their health status.

The aim of this study is to determine whether the investigational medical device KidneYou, randomly assigned to patients with CKD (G3b or G4 categories; A1 or A2 or A3 categories), can improve the CKD-related health conditions by positive changes in the patient's lifestyle (ie, healthy eating and PA participation) and by means of cumulative stress reduction (ie, MP) originating from the underlying chronic disease and from diagnosis and treatment interventions necessary to manage it.

Methods

Study Intervention

KidneYou, an app for mobile devices developed by Advice Pharma Group Srl, on its electronic data capture (EDC) ICE technology, is an investigational medical device designed to deliver digital therapy (DT) to patients. DT, as part of digital health, uses high-quality digital technologies to stimulate lifestyle changes in patients.

A multidisciplinary board, composed of clinicians, nutritionists, psychiatrists, and graphic designers, has been established. The team worked collaboratively to map out the patient journey for individuals with CKD, identifying key stages within the disease

XSL•FO
RenderX

management process. From this journey, critical touchpoints and challenges in CKD management were pointed out, emphasizing factors that play a pivotal role in improving patients' physical and mental well-being.

Each specialist contributed to defining the core domains of the app. Nutritionists were involved in personalized NPs based on patients' characteristics and needs, whereas psychiatrists contributed to the design of mindfulness practices tailored to address mental health challenges in patients with CKD. Physiatrists played a crucial role in developing the PA components, ensuring that exercise regimens were safe, effective, and adaptable to individuals with CKD.

Using these multidisciplinary inputs, a computational framework that generates a personalized algorithm was developed, incorporating customized nutrition, PA, and mindfulness strategies, all tailored to individual needs based on their specific health parameters.

To further enhance patient engagement, gamification elements were incorporated into the app design. These features were carefully crafted to encourage consistent app usage and promote a more interactive and enjoyable experience. By turning health management into an engaging activity, the app aims to foster long-term adherence to the prescribed treatment plan.

The overall design of the app reflects a patient-centered approach, combining clinical expertise from multiple fields with innovative digital solutions. This ensures that patients with CKD receive a comprehensive, holistic management experience that is both scientifically grounded and tailored to their specific needs.

KidneYou has been designed to improve the commitment of patients in following clinical directives by working on three therapeutic areas: NPs, PAs, and MPs. All these programs are administered by means of DT for patients included in the intervention group. Images of KidneYou are presented in Figure 1.

The NP assigned by the app is based on gender, BMI, and estimated GFR (eGFR; values both measured at baseline) and is a customized low-protein dietary plan. The plan includes 5 daily meals, which vary every day of the week to maintain the proposed meals attractive for the patient and increase their compliance with the dietary plan. In addition, the NP proposes to each patient further daily alternatives to the first and second courses and side dishes to further augment food variability, reduce as much as possible dietetic monotony, and minimize abandonment or severe deviation from the assigned diet.

The PA assigned by the DT is based on an individual's background fitness level and patient with CKD history, assessed at baseline. The exercise program is customized in terms of recommended type of PA, minutes of exercise per day, number of days per week, and level of intensity. At the end of each daily exercise, the patient selects the real number of minutes they spent exercising and rates the PA as too light, very good, or too heavy.

The MP aims to help patients follow stress-reducing activities; it consists of multimedia content (audio files) with different mindfulness meditation paths, allowing the patient to be guided by the therapist's voice through different forms of exercise (eg, breath control, body scan, and virtual sound environments), all aimed at containing stress.

XSL•FO
RenderX

Figure 1. Login page of KidneYou (Advice Pharma Group Srl).

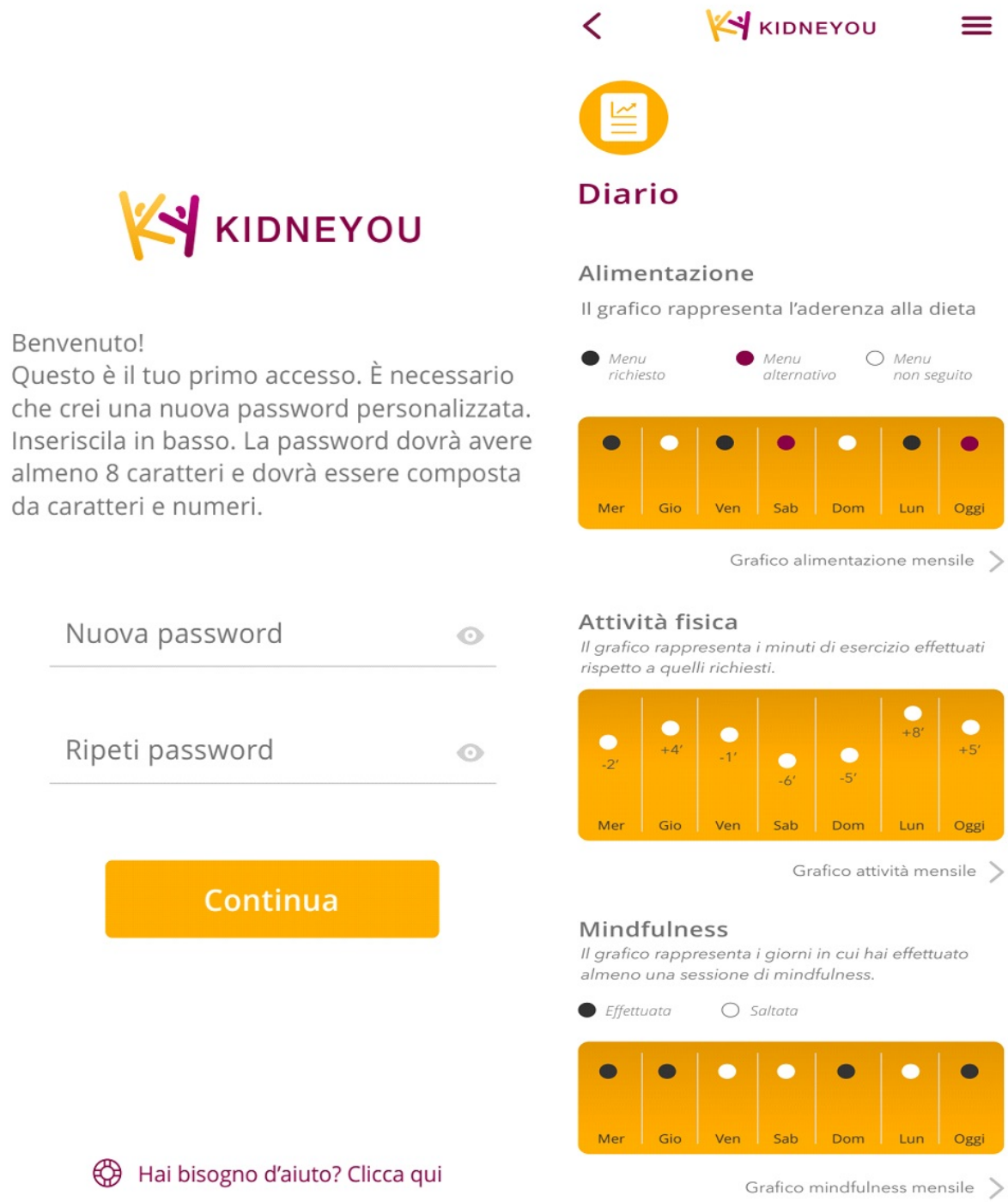

Study Design

This is a multicenter, randomized, 2-arm, comparative, parallel, open-label, premarket study on individuals affected by CKD belonging to GFR categories G3b or G4, and with persistent albuminuria categories A1, A2, or A3 (A3 limited to >300 mg/g [>30 mg/mmol] and <3000 mg/g [<300 mg/mmol]).

This study is a premarket clinical investigation conducted in accordance with European Union Medical Device Regulation n.2017/745 requirements, with the goal of generating clinical evidence necessary for Conformité Européenne marking of the investigational medical device.

Enrolled participants are randomly allocated in a ratio of 1:1 to one of the two treatment arms using a randomization list that is computer generated by the trial statistician and provided through a web-based approach (via interactive response technology or randomization and trial supply management): group A (KidneYou users; investigational arm) or group B (KidneYou nonusers; control arm). A randomization list with random permuted blocks of different sizes (2 or 4 or 6 or 8) has

XSL•FO
RenderX

been prepared and included in the electronic case report form (eCRF).

Ethical Considerations

The study is conducted under the clinical investigation plan and international guidelines, including the most recent versions of the Declaration of Helsinki, the International Conference on Harmonization guidelines for Good Clinical Practice, European Union Medical Device Regulation 2017/745, and International Organization for Standardization (ISO) 14155, as well as all applicable local laws and regulations. The study was approved by the Territorial Ethics Committee (Comitato Etico Territoriale Regione Puglia of the Azienda Ospedaliero-Universitaria Consorziale Policlinico Di Bari, Institutional Ethics Committee protocol number 0036279, April 19, 2022) and by the Italian Ministry of Health (protocol number 2061; October 19, 2022). Informed consent is collected from all the participants involved in the study. No financial compensation was provided to the participants. All participant data are collected in a pseudonymized form and treated confidentially in compliance with applicable privacy laws, including the General Data Protection Regulation. Data are stored securely and used solely for research purposes.

The study team disseminates the trial results by making a summary publicly available on ClinicalTrials.gov, presenting findings at national and international conferences, and publishing them in peer-reviewed journals.

Treatment Arms

In the intervention group (group A), the NP and PA program are administered through KidneYou, supplemented by guided stress-reduction activities.

In the control group (group B), the NP and PA program are administered by means of a paper diary. No indications are given to patients in this group regarding stress reduction activities.

Duration of Study

For each patient, the treatment lasts for 3 consecutive months (12 wk), during which lifestyle changes are implemented with or without KidneYou. At the end of this period, the lifestyle program ends, followed by a follow-up visit 1 month later (at 16 wk).

Study Setting

The trial involves 12 Italian clinical centers, all specialized in the treatment of renal diseases, distributed across different Italian regions, and provides a comprehensive representation of the country's diverse lifestyles and dietary habits. A detailed list of the participating sites and their respective staff can be found in Multimedia Appendix 1.

Study Population

The inclusion and exclusion criteria are presented in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Age ≥18 years at the time of signing informed consent.
- Participants with chronic kidney disease classified as glomerular filtration rate (GFR) category G3b (30-44 mL/min/1.73 m²) or G4 (15-29 mL/min/1.73 m²); persistent albuminuria category A1 (<30 mg/g), A2 (30-300 mg/g), or A3 (>300 mg/g and <3000 mg/g).
- Presence of at least 1 causal etiology: diabetes, arterial hypertension, chronic glomerulonephritis, and cystic kidney diseases (eg, polycystic kidney, nephronophthisis, cortical or tubular glomerular cysts, renal medulla cystic diseases, and tuberous sclerosis).
- Abnormalities of kidney function (GFR and albuminuria) and structure (causal etiologies) present for >3 months with health implications.
- Perceived Stress Scale total score >12.
- Sex: male and female.
- Capability of giving informed consent, including compliance with requirements and restrictions.
- Ownership of a mobile phone, willingness to use mobile apps, and ability to download KidneYou.
- Native speakers of Italian or foreign individuals with a full understanding of Italian (all programs provided in Italian).

Exclusion criteria

- Medical conditions: (1) acute and chronic joint disease; (2) acute or chronic muscle diseases; (3) history or current depression; (4) sleep disturbance; (5) suicidal ideation or any mental disorder; (6) refusal or inability to follow dietary rules; (7) chewing disorder; (8) lack of motivation to follow nutritional program, physical activity, and psychological programs; (9) positive hepatitis C antibody, hepatitis B virus surface antigen or core antibody, Child-Pugh class C, and HIV positive; (10) drug or alcohol abuse; (11) QT-interval prolongation or congenital long QT-interval syndrome and history of arrhythmia; (12) cancer; and (13) autoimmune diseases, Addison disease, and amyloidosis.
- Prior or concomitant therapies and other exclusions: (1) any change, termination, addition, or replacement of medication in the current treatment strategy; (2) participation in another clinical study in the last 6 months; (3) involvement in the planning or conduct of the study; (4) investigator judgment that participation is not advisable; (5) previous enrollment or randomization in this study; (6) currently pregnant or breastfeeding (women only); and (7) any other condition making participation undesirable in the investigator's opinion.

XSL•FO
RenderX

Main Objectives and End Points

Primary Objective

The primary objective is (1) to improve the quality and quantity of dietary intake in patients using KidneYou compared to patients not provided with the app after 3 months of a customized NP, (2) to increase energy expenditure in patients using KidneYou compared to patients not provided with KidneYou after 3 months of a customized PA program, or (3) to relieve the stress related to the underlying CKD condition in patients using KidneYou compared to patients not using KidneYou after 3 months of study.

Secondary Objectives

The secondary objectives are as follows:

1. To help patients achieve 2 or more of the aforementioned objectives
2. To describe each primary objective achieved in each randomized group
3. To stratify the primary objective results by G3b and G4 categories of CKD classification
4. To stratify the primary objective results by A1, A2, and A3 (A3 limited to >300 mg/g [>30 mg/mmol] and <3000 mg/g [<300 mg/mmol]) category of CKD classification
5. To assess the degree of azoturia reduction (gram per 24 h) in each treatment group during the study
6. To assess differences in azoturia reduction (gram per 24 h) between the 2 randomized groups during the study
7. To assess the degree of physical function improvements in each treatment group during the 6-minute walk test (6-MWT)
8. To assess differences in physical function improvements between the 2 randomized groups during the study
9. To assess the degree of PA increase (min) in each treatment group during the study
10. To assess the degree of stress reduction (Perceived Stress Scale [PSS] of Cohen [12]) in each treatment group during the study
11. To assess the differences in stress reduction (PSS of Cohen) between the 2 randomized groups during the study
12. To detect the acceptability level among KidneYou users.

Safety Objective

The safety objective is to describe the safety profile in the 2 treatment groups.

Primary End Point

The primary end point is as follows: (1) to evaluate the achievement of a mean reduction of at least 10% of azoturia (gram per 24 h) in CKD KidneYou users exposed to a 3-month NP compared to CKD KidneYou nonusers exposed to the same 3-month NP, (2) to evaluate the achievement of a mean increase of at least 15% of the distance (meters in the 6-MWT) in CKD KidneYou users exposed to a 3-month PA program compared to CKD KidneYou nonusers exposed to the same 3-month PA program, or (3) to evaluate the achievement of a mean decrease of at least 10% of perceived stress (according to the PSS of Cohen) in CKD KidneYou users exposed to a 3-month MP compared to CKD KidneYou nonusers and not exposed to the same MP.

Secondary End Points

The secondary end points are as follows:

1. To evaluate the achievement of 2 or more of the aforementioned components of the primary end point
2. To assess the achievement of every single component of the primary end point in each randomized group
3. To stratify the within-group changes and the between-group differences in changes of primary end point results by the eGFR category of patients with CKD (eg, G3b and G4)
4. To stratify the within-group changes and the between-group differences in changes of primary end point results by the albuminuria category of patients with CKD (eg, A1, A2, and A3)
5. To assess the within-group changes of mean values of azoturia (gram per 24 h) in each treatment group during the study visits (ie, V0, V2, and V3)
6. To assess the between-group differences in changes of mean values of azoturia (gram per 24 h) during the study visits (ie, V0, V2, and V3)
7. To assess the within-group changes of mean values of heart rate (HR), blood pressure (mm Hg), and O_2 saturation in each treatment group during the performance of the 6-MWTs
8. To assess the between-group differences in changes in mean values of HR, blood pressure (mm Hg), and O_2 saturation during the performance of the 6-MWTs
9. To assess the within-group changes in mean values of PA (min) self-reported by patients in each treatment group during the study visits (ie, V0, V2, and V3)
10. To assess the within-group changes of mean values of stress reduction (according to the PSS of Cohen) in each treatment group during the study visits (ie, V0, V2, and V3)
11. To assess the between-group differences in changes of mean values of stress reduction (according to the PSS of Cohen) during the study visits (ie, V0, V2, and V3)
12. To assess patient satisfaction with the KidneYou App by using the Net Promoter Score system. With a score ≥7, the participant will be considered a promoter of the app

Safety End Point

The safety end point is to assess the rate of adverse events (AEs), adverse device effects, and abnormality of laboratory parameters between the 2 treatment groups across the study (ie, V1, V2, and V3).

Tertiary or Exploratory End Points

The tertiary or exploratory end points are as follows:

1. To assess the between-group differences in changes of mean values of PA (min) self-reported by patients during the study visits (ie, V0, V2, and V3)
2. To assess the within-group changes and the between-group differences in changes of urinary levels of phosphorus, calcium, potassium, sodium, chlorine, bicarbonate, albumin, creatinine, glucose, and uric acid by urine samples during the study visits (ie, V0, V2, and V3)

XSL•FO
RenderX

3. To assess the within-group changes and the between-group differences in changes of urinary levels of other compounds (ie, density, pH, proteins, ketones, hemoglobin, bilirubin, nitrites, white cells, red blood cells, casts, crystals, and epithelial cells) by urine samples during the study visits (ie, V0, V2, and V3)
4. To assess the within-group changes and the between-group differences of changes in blood levels of electrolytes (potassium, sodium, chloride, and bicarbonate), minerals (phosphorus and calcium), proteins (albumin), metabolic products (nitrogen and creatinine), sugar (glucose), and uric acid in the blood samples during the study visits (ie, V0, V2, and V3)
5. To assess the within-group changes and the between-group differences of changes of vital signs (systolic blood pressure or diastolic blood pressure, HR, and respiratory rate) during the study visits (ie, V0, V1, V2, and V3)
6. To assess the between-group differences of changes in mean values of impedancemetry (lean mass and fat mass) during the study visits (ie, V0, V2, and V3)
7. To assess the within-group changes and the between-group differences of changes of signs and symptoms characterizing CKD during the study visits (ie, V0, V1, V2, and V3)
8. To assess the within-group compliance (adherence) by measuring the overall time (min) of KidneYou use during the 3 months of the study
9. To assess a possible correlation between levels of compliance with KidneYou use and the rate of achievement of the primary end point
10. To assess a possible correlation between levels of compliance (min) with KidneYou and the customer experience rate according to Net Promoter Score

Assessments

Each participant undergoes 4 visits and is examined according to the following schedule: baseline (V0), 6-week visit (V1), 12-week visit (V2), and 4-week follow-up visit (V3). During the baseline visit, the patient signs the informed consent form and is evaluated for inclusion and exclusion criteria. Once enrollment is confirmed, the patient is randomized. The PSS questionnaire is administered to all enrolled patients, regardless of the randomization group. At V0, V1, V2, and V3, a complete physical examination is performed; vital signs (systolic blood pressure, diastolic blood pressure, HR, and respiratory rate), ongoing medications for CKD, and concomitant medications are collected.

At V0, V2, and V3, laboratory tests are also performed and include hematology (complete blood count, glycemia, proteins [albumin], and creatinine), minerals (phosphorus and calcium), electrolytes (sodium, potassium, chloride, and bicarbonate), azotemia, and uric acid; lipid profile, that is, total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and triglycerides; and urine tests (24 h urinary excretion collection), that is, urea nitrogen, uric acid, minerals (phosphorus and calcium), electrolytes (potassium, sodium, chloride, and bicarbonate), albumin, creatinine, glucose, density, pH, proteins, ketones, hemoglobin, bilirubin, nitrites, white cells, red blood cells, casts, crystals, epithelial cells, and serum pregnancy test, only for women of childbearing potential.

All patients undergo the 6-MWT and bioimpedance test. Safety, including the appearance of adverse device effects and abnormalities in physical and laboratory parameters, is assessed throughout the entire study period.

Data Collection and Management

The KidneYou software consists of 3 different components.

The mobile app enables the patient to receive dietary indications, exercise categories, and mindfulness content and to enter information related to the activities performed. It also allows the patient to provide feedback through questions and evaluations on the activities performed.

The ICE Platform, which can be accessed through a dedicated control panel after authentication and authorization, allows the physician to enter patient data into the platform. The platform also displays data on patient adherence to recommended activities.

The backend is the part of the software dedicated to exchanging communications between the mobile app and the ICE platform. It receives patient input from the app, including self-reported adherence to activities, which is sent to the platform for visualization.

The KidneYou database has been developed using EDC technology (ICE, Advice Pharma), compliant with the regulations for clinical data management to facilitate the study data analysis.

Data are collected for each participant in eCRFs. The eCRF is designed by Advice Pharma Group using the ICE software.

ICE (version 2.6.2) is the EDC solution provided by Advice Pharma. This system is a fully web-based Integrated Development Environment designed to generate customizable eCRFs.

The EDC system consists of the database application and integrated tools for query management and audit trail. Data are reviewed and checked for completeness and validity before, during, and after data entry. The data cleaning process includes different levels of control: source data verification and quality checks at the data entry level, automatic controls on the ICE database, and database data review performed by the data manager, with the production of manual eQueries.

Entered data are subject to visual and electronic validation by data management and monitoring staff. Any errors, inconsistencies, and unexpected values in the data are noted and queried with the site staff for resolution.

The data manager reviews data for each patient upon study withdrawal, after all data are entered, and after source data verification is completed. Data cleaning continues until all discrepancies are resolved. Data changes resulting from queries are made by the investigator directly in the clinical database.

Concomitant medications are coded using the latest available version of the *World Health Organization Drug Dictionary Enhanced* (WHO DDE). The data manager performs the coding centrally using the dedicated ICE coding tool, matching the

XSL•FO
RenderX

term entered by the investigator with the appropriate term in the relevant dictionary.

Medical coding is performed at the end of the study before database lock.

Statistics

Sample Size Calculation

This study is designed to test the null hypothesis that the percentage *success* (defined by the achievement of at least one of the 3 conditions described in the primary end point, with no hierarchical structure among them) in KidneYou nonusers is equal to the percentage "success" in KidneYou users. The alternative hypothesis is that the percentage "success" in KidneYou nonusers is different from the percentage "success" in KidneYou users.

The sample size is estimated to assess the study hypothesis on the primary end point.

The sample size calculation is based on the combination of the expected proportion of success in the control group and the expected delta (Δ) of success in the experimental group.

With a significance level α equal to .05 and a power equal to 80%, a sample size of 190 participants (95 per group) is required to detect as significant, if any, a success delta in favor of the experimental group equal to 20% (under the hypothesis of success rate of 40% in the control group and success rate of 60% in the experimental group). Considering a dropout rate of approximately 10%, a total of 210 patients (105 per group) will have to be enrolled. A competitive enrollment approach across the centers involved in the study will be used to reach the sample size efficiently.

Statistical Analysis

All the analyses on primary, secondary, and tertiary or exploratory end points will be carried out in the intention-to-treat population—as main analyses—and in per protocol population, as sensitivity analyses. Safety end points will be described only for the per protocol population.

Data will be analyzed and presented in accordance with the international CONSORT (Consolidated Standards of Reporting Trials) guidelines and analyzed according to the intention-to-treat principle.

Data analyses will be conducted using SAS (version 9.4; SAS Institute) statistical software.

Primary Analysis on Primary End Point

The primary efficacy end point is the percentage of achievement of a mean reduction of at least 10% of azoturia (grams per 24 h) or the achievement of a mean increase of at least 15% of distance (meters in the 6-MWT) or the achievement of a mean decrease of at least 10% of perceived stress (according to the PSS of Cohen) in KidneYou users exposed to a 3-month NP, PA program, and MP compared to KidneYou nonusers and not exposed to the same 3-month programs.

The observed percentages (in KidneYou users vs KidneYou nonusers) will be tested using a chi-square test or the Fisher exact test, as appropriate.

The secondary analyses on the primary end point include (1) achievement of each single component of the primary end point in each randomized group, (2) achievement of 2 or more of the aforementioned components of the primary end point, and (3) assessment of between-group differences in primary end point results by estimated eGFR category of patients with CKD (eg, G3b and G4) and by albuminuria category of patients with CKD (eg, A1, A2, and A3).

Secondary End Points

The secondary end points include the evaluation of many parameters (both continuous and categorical) through comparisons both between groups and within groups.

Between-group differences will be investigated with 2-tailed unpaired *t* tests (or the corresponding nonparametric Wilcoxon-Mann-Whitney tests, as appropriate) in case of a continuous variable and with chi-square test or the Fisher exact test in case of a categorical variable.

Within-group differences will be investigated with paired *t* tests (or the corresponding nonparametric Wilcoxon-Mann-Whitney tests, as appropriate) in case of a continuous variable, with the McNemar test in cases of a categorical one.

If appropriate, the percentage change with respect to the baseline value will be calculated for each interesting time point.

Global approaches, such as mixed linear models for repeated measures, will also be considered.

Tertiary or Exploratory End Points

Tertiary or exploratory end points that involve a comparison between and within groups will be analyzed as described in the "Secondary End Points" section.

To assess possible correlation between levels of compliance to KidneYou use and other parameters of interest, the most appropriate test will be used. In particular, in case of continuous or numerical parameters, the Pearson correlation coefficient (or Spearman, as appropriate) will be computed; in case of discrete or dichotomous parameters, the association will be assessed with 2-tailed unpaired *t* tests or the corresponding nonparametric Wilcoxon-Mann-Whitney tests, as appropriate.

Safety

Medical AEs, medical device AEs, and laboratory parameter abnormalities will be collected during the study.

Medical AEs will be coded to a preferred term and primary system organ class, using the current version of the *Medical Dictionary for Regulatory Activities* (MedDRA; International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) dictionary. AEs will be tabulated as preferred term, system organ class, and frequency for the safety population.

XSL•FO
RenderX

The statistical difference in the rate of AEs, device AEs, and laboratory parameter abnormalities will be assessed through the chi-square test or the Fisher exact test.

Results

Data will be analyzed and presented following the guidelines outlined in the international CONSORT to ensure the integrity and transparency of the results. The primary and secondary end points will be rigorously examined to assess the effectiveness of the KidneYou app in promoting adherence to nonpharmacological interventions (such as NP, PA program, and MPs) and improving health outcomes in patients with CKD.

Recruitment began in July 2022 and was completed in April 2024 across 12 Italian nephrology centers with a total number of 210 participants enrolled. Data analysis is currently ongoing. The results are expected to be published by early 2026.

Discussion

Anticipated Findings

This study is designed to evaluate the anticipated performance of KidneYou, a digital therapeutic tool delivering personalized NP, PA program, and MPs, which can improve adherence to lifestyle changes and health outcomes in patients with CKD, compared to the standard approach.

If effective, the KidneYou model could guide future digital health solutions for CKD and other chronic diseases, offering a scalable, nonpharmacological approach. These solutions may complement medical treatments, reduce costs, and enhance patient adherence to lifestyle changes. By providing an accessible, patient-centered tool, this approach could improve treatment outcomes and help patients with CKD manage their condition more effectively.

Comparison With Prior Work

Previous digital interventions for patients with CKD have largely focused on single domains such as nutritional guidance or physical activity, with limited evidence on integration across multiple lifestyle components. To the best of our knowledge, no existing digital solution for CKD has combined all 3 nonpharmacological domains in a single, personalized platform. This trial is designed to fill that gap.

Strengths and Limitations

The strengths of this study include its multicenter design, the use of a composite primary end point, and the personalization of interventions to individual patient needs. Limitations include the relatively short intervention period, the absence of a prespecified plan to disentangle the independent contribution of engagement features to adherence or outcomes, and the exclusion of patients with mental health conditions, such as depression or sleep disturbances, which are relatively common among individuals with CKD who could also benefit from mindfulness-based interventions. Nevertheless, this approach was necessary to standardize the study population and minimize potential confounding factors in this interventional trial. We also acknowledge that restricting participation to Italian speakers and tailoring the content to the Italian context may limit the international applicability of the findings. However, this design was necessary because the KidneYou App, including all its content, was specifically developed for Italian patients and is culturally, linguistically, and in terms of dietary content, tailored to Italy. Future studies may explore the adaptation of the app for other countries and languages to assess generalizability.

Future Directions

If the intervention proves to be effective, future research should investigate long-term outcomes and improve mechanisms of engagement. Studies could also explore optimization of individual program components and adaptation of the model for populations with other chronic diseases.

Dissemination

The results of this study will be made publicly available on ClinicalTrials.gov and disseminated through peer-reviewed publications and presentations at national and international conferences, ensuring accessibility to both the scientific community and clinical stakeholders.

Monitoring

The study is monitored in accordance with Good Clinical Practice guidelines and the study-specific monitoring plan. On-site monitoring visits are conducted to ensure protocol adherence, data accuracy, and participant safety. Source data verification is performed as per the monitoring plan. AEs and serious adverse events are assessed and reported in compliance with regulatory requirements.

Funding

KidneYou is a co-design project of AdvicePharma and AstraZeneca SpA, funded by AstraZeneca. The legal manufacturer of the Investigational Medical Device (Software as Medical Device) and the sponsor of the clinical investigation is AdvicePharma. The sites involved in the study are listed in Multimedia Appendix 1. KidneYou is a registered European Union (EU) trademark (018432826) at the EU Intellectual Property Office.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to participant privacy but are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: M Gorini, M Gnesi, ARM, MB, FM, MF, LG

XSL•FO
RenderX

Writing – original draft: CC, M Gnesi, AF, GC, MF, PMA
Writing – review & editing: M Gorini, M Gnesi, ARM, MB, FM, LG, CC, AF, GC, MF, PMA

Conflicts of Interest

M Gorini, M Gnesi, FM, and ARM are employees of AstraZeneca SpA, which funded the study and co-designed the KidneYou app.
AF and MB are employees of AdvicePharma, the legal manufacturer of the KidneYou App and study sponsor.
LG, CC, MF, and PMA declare no commercial or financial conflicts of interest related to this study.

Multimedia Appendix 1
List of study sites participating in the trial and corresponding principal investigators.
[DOCX File, 18 KB - resprot_v15i1e75306_app1.docx]

References

1. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int 2024 Apr;105(4S):S117-S314. [doi: 10.1016/j.kint.2023.10.018] [Medline: 38490803]
2. Rhee CM, Wang AYM, Biruete A, et al. Nutritional and dietary management of chronic kidney disease under conservative and preservative kidney care without dialysis. J Ren Nutr 2023 Nov;33(6S):S56-S66. [doi: 10.1053/j.jrn.2023.06.010] [Medline: 37394104]
3. Kaimori JY, Sakaguchi Y, Oka T, Isaka Y. Plant-dominant low-protein diets: a promising dietary strategy for mitigating disease progression in people with chronic kidney disease-a comprehensive review. Nutrients 2025 Feb 11;17(4):643. [doi: 10.3390/nu17040643] [Medline: 40004970]
4. Zhang L, Wang Y, Xiong L, Luo Y, Huang Z, Yi B. Exercise therapy improves eGFR, and reduces blood pressure and BMI in non-dialysis CKD patients: evidence from a meta-analysis. BMC Nephrol 2019 Oct 29;20(1):398. [doi: 10.1186/s12882-019-1586-5] [Medline: 31664943]
5. Davies M, Sandoo A, Macdonald J. The role of exercise training in delaying kidney function decline in non-dialysis-dependent chronic kidney disease. Kidney Dial 2022;2(2):262-286. [doi: 10.3390/kidneydial2020026]
6. Battaglia Y, Baciga F, Bulighin F, et al. Physical activity and exercise in chronic kidney disease: consensus statements from the physical exercise working group of the Italian society of nephrology. J Nephrol 2024 Sep;37(7):1735-1765. [doi: 10.1007/s40620-024-02049-9] [Medline: 39269600]
7. Kang YS, Choi SY, Ryu E. The effectiveness of a stress coping program based on mindfulness meditation on the stress, anxiety, and depression experienced by nursing students in Korea. Nurse Educ Today 2009 Jul;29(5):538-543. [doi: 10.1016/j.nedt.2008.12.003] [Medline: 19141364]
8. Huberty J, Green J, Glissmann C, Larkey L, Puzia M, Lee C. Efficacy of the mindfulness meditation mobile app “calm” to reduce stress among college students: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14273. [doi: 10.2196/14273] [Medline: 31237569]
9. Li X, Gao W, Yu J, et al. Mental health and quality of life in chronic kidney disease patients with mild-to-moderate depression: a retrospective cohort study of mindfulness-based stress reduction therapy. Actas Esp Psiquiatr 2024 Apr;52(2):138-148. [doi: 10.62641/aep.v52i2.1600] [Medline: 38622007]
10. Curtis F, Burton JO, Butt A, et al. Lifestyle interventions delivered by eHealth in chronic kidney disease: a scoping review. PLoS ONE 2024;19(1):e0297107. [doi: 10.1371/journal.pone.0297107] [Medline: 38266006]
11. Letton ME, Tr n TB, Flower S, et al. Digital physical activity and exercise interventions for people living with chronic kidney disease: a systematic review of health outcomes and feasibility. J Med Syst 2024 Jul 1;48(1):63. [doi: 10.1007/s10916-024-02081-z] [Medline: 38951385]
12. Cohen S. Perceived stress in a probability sample of the United States. In: Spacapan S, Oskamp S, editors. The Social Psychology of Health: Sage Publications, Inc; 1988:31-67.

Abbreviations

6-MWT: 6-minute walk test
AE: adverse event
CKD: chronic kidney disease
DT: digital therapy
eCRF: electronic case report form
EDC: electronic data capture
eGFR: estimated glomerular filtration rate
GFR: glomerular filtration rate
HR: heart rate

XSL•FO
RenderX

ISO: International Organization for Standardization
MedDRA: *Medical Dictionary for Regulatory Activities*
MP: mindfulness program
NP: nutritional program
PA: physical activity
PSS: Perceived Stress Scale
WHO DDE: *World Health Organization Drug Dictionary Enhanced*

Edited by J Sarvestan; submitted 01.Apr.2025; peer-reviewed by IA Eide, N Maeda; accepted 05.Nov.2025; published 19.Feb.2026.

Please cite as:
Cafiero C, Gnesi M, Maurizi AR, Foresta A, Beccaria M, Campilongo G, Fiorentino M, Acquaviva PM, Mastromauro F, Gorini M, Gesualdo L
Evaluating the Safety and Performance of the KidneYou App for Chronic Kidney Disease: Protocol for an Italian Multicenter, Randomized, Open-Label, Premarket Study
JMIR Res Protoc 2026;15:e75306
URL: https://www.researchprotocols.org/2026/1/e75306
doi:10.2196/75306

© Cesira Cafiero, Marco Gnesi, Anna Rita Maurizi, Andreana Foresta, Massimo Beccaria, Giorgia Campilongo, Marco Fiorentino, Paola Maria Acquaviva, Francesca Mastromauro, Marco Gorini, Loreto Gesualdo. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developing and Evaluating a WeChat-Based Applet Fluid Intake Reminder on Enhancing Fluid Adherence in Postoperative Patients With Urinary Calculi: Protocol for a Randomized Controlled Trial

Xue Zhou[1], MSc; Tongyao Wang[1], PhD; Qian Chen[2], MSc

[1]School of Nursing, Li Ka Shing Faculty of Medicine, University of Hong Kong, Room 532, Academic Building at No 3, Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

[2]Department of Urology, Shenzhen Qianhai Shekou Free Trade Zone Hospital, Shenzhen, China

Corresponding Author:
Tongyao Wang, PhD
School of Nursing, Li Ka Shing Faculty of Medicine, University of Hong Kong, Room 532, Academic Building at No 3, Sassoon Road, Pokfulam, Hong Kong, China (Hong Kong)

Abstract

Background: Urinary calculi (UC), affecting 1% - 13% globally, pose a significant health burden due to high recurrence rates (up to 50% within 10 years) and substantial health care costs. Adequate fluid intake is a cornerstone of prevention; yet, its adherence remains poor due to forgetfulness, lifestyle barriers, and limited patient education. Existing mobile health interventions for UC prevention often lack medical oversight and clinical validation. WeChat-based digital therapeutic intervention may have a positive effect on fluid adherence in this patient group.

Objective: Our objective is to develop a WeChat applet to improve hydration behavior and reduce stone recurrence among postoperative patients with UC.

Methods: This is an open-label, 2-arm, parallel-group randomized controlled trial. We will recruit 148 participants from China's tertiary hospital and randomly allocate them in a ratio of 1:1 to the intervention or control group. The intervention group received standard postoperative care supplemented by the WeChat-Based Applet Fluid Intake Reminder (WAFIR), which delivers personalized fluid intake reminders, urine color monitoring, 24-hour fluid intake and urine output tracking, and evidence-based educational content on hydration and urolithiasis management. The control group receives standard care of general discharge instructions from nurses. The primary outcome is the fluid adherence, measured by 24-hour fluid intake and urine volume; secondary outcome measures are the Wisconsin Stone Quality of Life Questionnaire, Patient Health Questionnaire-9, Electronic Health Literacy Scale, physical activity (International Physical Activity Questionnaire—Short Form), and recurrence rate of UC. Outcomes are measured before intervention (T0) and after a 1-month (T1) and 3-month (T2) follow-up period. Intention-to-treat analysis, 2-tailed *t* tests, and repeated measures ANOVA will be used to compare outcomes; statistical significance is set at a $P<.05$ significance threshold. The study was approved by the ethics review board in December 2024.

Results: The development of WAFIR, conducted in collaboration with stakeholders, was finalized in February 2025. Recruitment commenced on March 1, 2025; data collection was completed in September 2025, and data analysis was analyzed in December 2025. Dissemination of findings is planned through conferences and publications in 2026.

Conclusions: This research evaluates the effectiveness of a nurse-led, evidence-based digital therapeutic intervention, WAFIR, in overcoming fluid adherence barriers among postoperative patients following urolithiasis surgery, aiming to increase daily fluid intake and urine output, reduce recurrence rates, enhance quality of life, and generate empirical evidence for its application in urology care, thereby optimizing postoperative management within clinical settings.

Trial Registration: ClinicalTrials.gov NCT06990672; https://clinicaltrials.gov/study/NCT06990672

International Registered Report Identifier (IRRID): DERR1-10.2196/80214

(JMIR Res Protoc 2026;15:e80214) doi:10.2196/80214

KEYWORDS

fluid adherence; urinary calculi; digital therapeutic; WeChat applet; fluid reminder

XSL•FO
RenderX

Introduction

Urolithiasis, or urinary calculi (UC), is emerging as a significant global health crisis, affecting individuals of all ages and genders. The prevalence of UC has been increasing worldwide over the past decades [1]. Estimated incidence rates are 7% - 13% in North America, 5% - 9% in Europe, and 1% - 5% in Asia [2]. In low-income countries, up to 25% of the population can be affected, often with fatal outcomes due to inadequate urological care [3]. In China, the prevalence of nephrolithiasis is around 5.8%, exceeding 10% in the southern region [4]. Meanwhile, the high recurrence rate of UC poses a significant challenge, with approximately 10%, 30%, and 50% of individuals experiencing recurrence at 1 year, 5 years, and 10 years, respectively [5-7]. The financial burden of UC is substantial, with yearly medical expenses estimated at 10 billion dollars in the United States and US $695.51 million in Europe [8,9]. Predictive models indicate that the high prevalence and recurrence rate will lead to a 25% increase in health care costs over the next 50 years [10]. This makes UC one of the most expensive urological diseases, highlighting the urgent need for better understanding and management of this condition.

The occurrence of UC is influenced by a variety of environmental, climatic, and dietary factors. Environmental and climatic conditions, particularly in hot climates, can contribute significantly to water loss and water quality issues. High temperatures often lead to increased perspiration, which in turn reduces overall water levels in the body. This dehydration can concentrate the urine, increasing the likelihood of stone formation. Additionally, water quality, specifically the levels of sodium and other minerals, plays a crucial role in this process [11,12]. Dietary habits are another critical factor in the development of urinary stones. Insufficient water intake is one of the primary dietary risk factors, as it leads to more concentrated urine, which facilitates the crystallization process. Poor dietary habits, such as the prolonged consumption of oxalate-rich foods—like spinach, nuts, and chocolates—significantly elevate the propensity for developing calcium oxalate stones [13]. High-sugar and high-fat diets further exacerbate this risk by influencing urinary elimination processes, making the urine more conducive to stone formation [14]. These risk factors highlight the important role of nurses in implementing targeted hydration protocols and dietary interventions to mitigate stone recurrence.

The European Association of Urology and the American Urological Association both strongly emphasize the importance of increasing fluid intake to prevent urinary stone formation. Research indicates that increasing fluid consumption stands out as a simple and economically feasible way to lower the risk of stone formation [15]. Then, McCauley et al [16] discovered that patients tend to be more open to increasing fluid intake than adopting other dietary or pharmaceutical therapies. Yet, Khambti et al [17] identified a challenge, noting that only about half of patients with urinary stones adhere to hydration recommendations, as determined through 24-hour urine output tests. Building on this, further research elucidated the underlying causes: at first, patients might not know or remember to drink enough; then, some dislike the taste of water or do not feel thirsty; and finally, the need for frequent voiding or workplace factors gets in the way [16]. Other research backs this up, showing that forgetfulness or not realizing their need for hydration plays a big role in poor fluid habits [18,19]. Therefore, enhancing adherence to fluid intake recommendations offers substantial benefits, including reduced health care expenditure through lower incidence of urinary stones and associated costs, improved public health outcomes, and an elevated quality of life (QoL) by decreasing the frequency and severity of symptoms.

Digital therapeutic interventions (DTx) deliver evidence-based, software-driven solutions via mobile apps, wearables, and web platforms to manage conditions. Unlike wellness tools, DTx is clinically validated, often regulated, and designed for measurable health outcomes [20]. By 2026, the global mobile app market was expansive, with approximately 4 million apps available on the Apple App Store and Google Play, reflecting the vast potential of digital health tools [21]. However, conventional mobile apps face significant barriers to widespread adoption, such as complex development processes, protracted timelines, and substantial financial investment, which often limit their scalability and accessibility in clinical settings. In contrast, WeChat applets, which require no installation and are integrated into China's predominant social media platform, facilitate seamless, on-demand access, effectively addressing technical barriers and offering distinct advantages by enhancing user convenience and minimizing operational challenges. With over 1.4 billion monthly active users, WeChat has a vast reach and engagement, making it a widely used communication tool in China. Using WeChat's existing infrastructure, the applets simplify the process by reducing the need for extensive coding and platform-specific adaptation and allow individuals to access them instantly. The significance of WeChat in health care has been amplified by its rapid adoption during the COVID-19 pandemic, which catalyzed its integration into various medical contexts, from telemedicine to patient education, markedly improving care quality [22,23]. Notably, the feasibility of using WeChat to enhance adherence to prevention strategies has been demonstrated among multiple chronic diseases [24]. WeChat-based DTx overcomes conventional mobile app barriers and drives health care innovation by enhancing patient adherence and care quality across diverse medical contexts, including chronic disease management and preventive strategies.

Previous studies have highlighted the limitations in the current landscape of DTx for fluid management and urinary stone prevention. Stevens et al [25] noted that many apps lack the support of medical professionals, leading to the dissemination of potentially false information and inaccuracies that could negatively impact patient health. This underscores the importance of involving medical professionals in the development and content review of these apps to ensure their reliability and clinical validity. The advent of digital technologies, such as wearable devices and intelligent water bottles, has introduced new methods for monitoring water intake and preventing urinary stones. However, the effectiveness of these devices is highly dependent on the practicality of their use. For instance, patients may not always have access to smart bottles in all settings, which can lead to incomplete fluid intake

XSL•FO
RenderX

records. Karagöz and Sarıca [26] further pointed out that the high cost and inconvenience associated with these devices may deter long-term patient compliance, highlighting the need for more cost-effective and user-friendly solutions for fluid intake notification. A review of the literature reveals that most published studies are small-scale, single-arm pilot interventions lacking robust controlled trials designed to evaluate long-term effectiveness and adherence. Moreover, Philip-McKenzie et al [27] analyzed 51 apps from the Apple and Google Play Store and found that none of the apps measured urine volume and color, and only 1 provided medical advice or educational content regarding the consequences of poor hydration. This reveals a substantial gap in the provision of comprehensive medical advice and education within these apps to improve patients' recognition of disease severity. Furthermore, Winoker et al [28] highlighted the scarcity of evidence supporting the benefits of mobile health (mHealth) interventions in urological care and emphasized the need for more rigorous research to validate the effectiveness of these technologies. Karagöz and Sarıca [26] also recommended that future research should focus on the practicality, security, efficacy, and overall benefits of these apps in preventing stone recurrence.

This study aims to investigate the efficacy of the WeChat-Based Applet Fluid Intake Reminder (WAFIR) in improving fluid intake among postoperative patients with UC compared to standard care, with the ultimate goal of reducing stone recurrence and enhancing patient well-being. The specific objectives are to quantify changes in daily fluid intake and urine output, evaluate its effect on the QoL, and determine its influence on readmission rates, recurrence rates, and eHealth literacy over the follow-up period. We hypothesize that patients receiving WAFIR exhibit statistically significant differences in 24-hour urine output (mL), fluid intake (mL/24 hours), QoL (Wisconsin Stone Quality of Life Questionnaire [WISQOL] score), eHealth literacy (Electronic Health Literacy Scale [eHEALS] score), or urinary stone recurrence rates (3-month follow-up) compared to those receiving usual care.

Methods

Research Design

We will conduct a 3-month, 2-arm randomized controlled trial (RCT) involving participants assigned to either a WAFIR intervention group or standard care group. Participants will be enrolled in batches, using block randomization with blocks of 10, and allocated in a 1:1 ratio, with a total sample size of 148 participants determined. An independent researcher (QC) will generate the randomization list using the sealed envelope method. A detailed research technology roadmap is displayed in Figure 1. Key aspects of the study design can be found in Figure 2.

XSL•FO
RenderX

Figure 1. RCT CONSORT flowchart. CONSORT: Consolidated Standards of Reporting Trials; eHEALS: Electronic Health Literacy Scale; IPAQ-SF: International Physical Activity Questionnaire—Short Form; PHQ-9: Patient Health Questionnaire-9; RCT: randomized controlled trial; WISQOL: Wisconsin Stone Quality of Life Questionnaire.

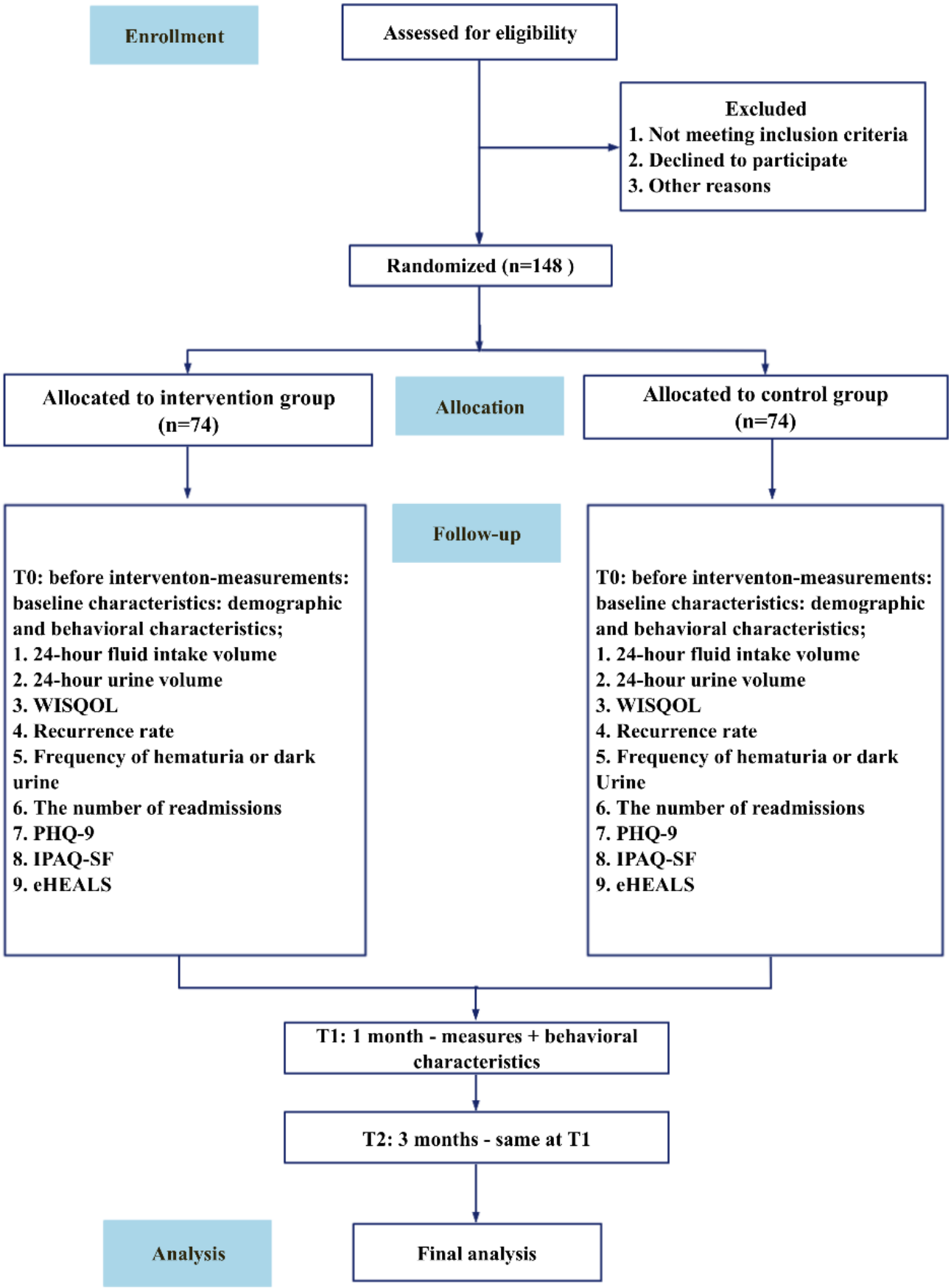

XSL•FO
RenderX

Figure 2. Study design table. MARS: Mobile App Rating Scale; WAFIR: WeChat-Based Applet Fluid Intake Reminder.

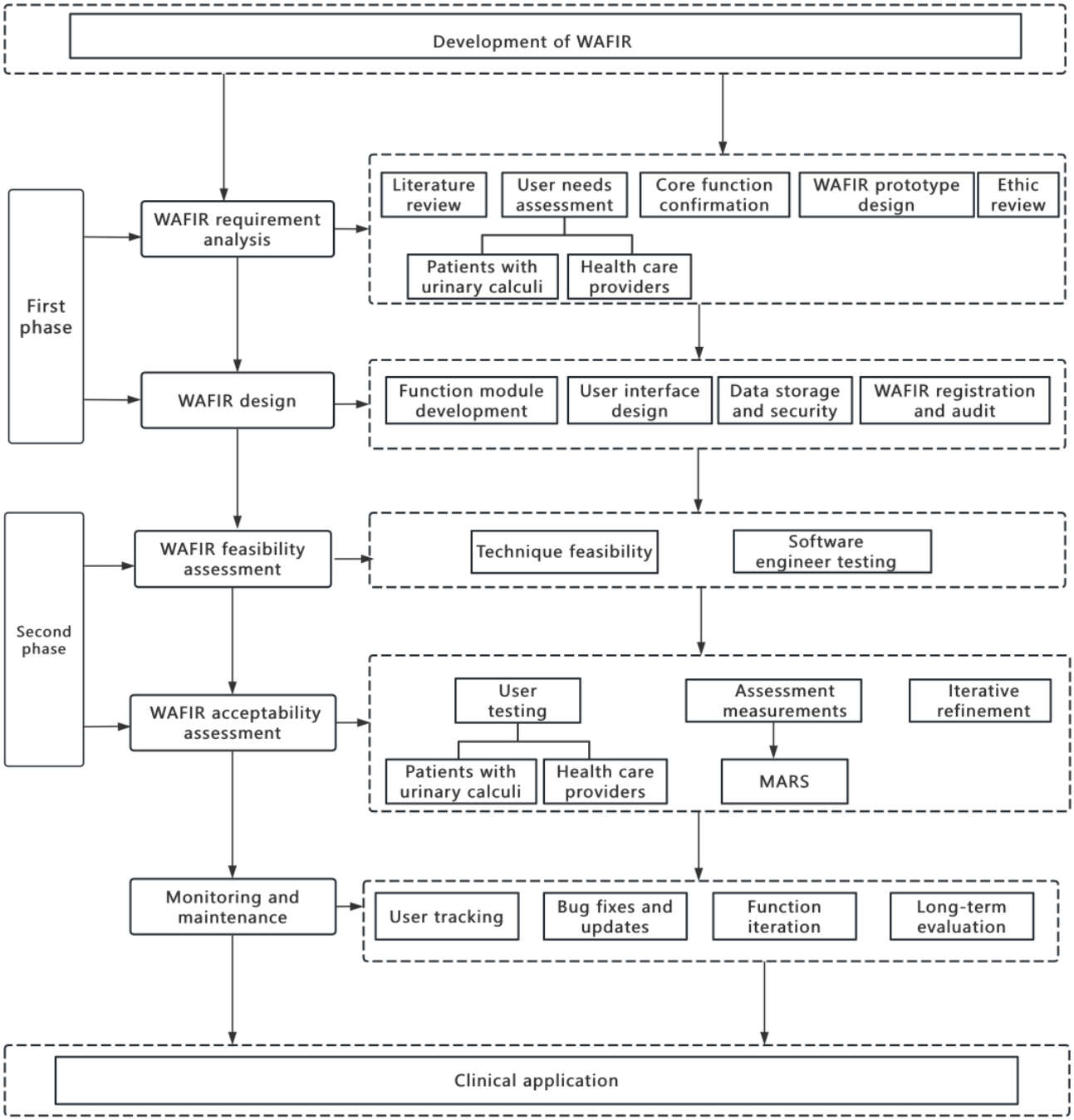

Characteristics of Participants

The study is open to individuals aged 18 years or older, diagnosed with UC and admitted to Shekou Free Trade Zone Hospital (a tertiary general hospital) for urinary stone surgery, to participate. Eligible patients will be enrolled prior to discharge, either by a nephrologist or a urologist (Textbox 1).

XSL•FO
RenderX

Textbox 1. Inclusion, exclusion, and withdrawal criteria for participants' recruitment.

Inclusion criteria

- With a history of urinary calculi and having received urolithiasis surgery before.
- 24-hour fluid intake <2500 mL.
- Age ≥18 years.
- Able to use WeChat daily.
- Able to read and write informed consent.

Exclusion criteria

- Participants with obstructive uropathy, chronic urosepsis, renal failure, and renal tubular acidosis.
- Participants with congestive heart failure, psychiatric conditions, pregnancy, and primary hyperparathyroidism.
- Participants experienced complicated urolithiasis surgery (surgical procedures for urolithiasis that involve complex interventions, often requiring multiple operations during the follow-up period due to factors such as large stone burden, anatomical abnormalities, or postoperative complications, eg, percutaneous nephrolithotomy with residual stones or ureteroscopy requiring reintervention).
- History of recurrent urinary tract infection (>3 per year).
- Medication use increases stone risk.
- Participants with physical disabilities.

Withdrawal criteria

- Participants will be deemed to have withdrawn from the study upon the voluntary withdrawal of their consent or if they fail to return a postintervention assessment within 30 days after the end of the intervention phase. All instances of withdrawal and the discontinuation of any participant's involvement in the trial are meticulously documented.

Intervention

Standard Care

Participants will receive standard dietary and regular recommendations, including an education profile and counseling to achieve a water intake and urine output ≥2500 mL, along with verbal health education counseling regarding urological calculi prevention during hospitalization and guidance on adequate hydration at discharge. A fixed-capacity water bottle will be provided. Phone calls will be given at each follow-up phase to reinforce the educational content and collect primary and secondary data through questionnaires.

WAFIR Intervention

Participants will receive the standard dietary counseling and an educational handout. The WAFIR will be introduced to the intervention group to enhance adherence to fluid intake and facilitate self-monitoring for the prevention of UC. The WAFIR will include features such as fluid intake reminders, fluid and urine value recording, urine color tracking, and health education, as well as an interactive group for communication among postoperative patients with urinary stones. The WAFIR will deliver personalized daily notifications to remind participants of increasing fluid intake exceeding 2500 mL. Additional educational content highlighting the benefits of adherence will be provided to the experimental group through the applet as well. Phone calls will be given at each follow-up phase to collect primary and secondary data via the questionnaires.

To address the limitations observed in existing mobile apps, we are developing an innovative WAFIR aimed at improving participants' adherence to fluid intake and reducing the recurrence of UC. This intervention will incorporate multiple functionalities to enhance participants' awareness and control over their fluid consumption. The WAFIR will consist of 4 main pages:

- Homepage: The primary interface of the WAFIR provides a user-centric platform for monitoring and managing daily hydration objectives. It enables patients to set personalized fluid consumption targets, log daily fluid intake, record urine output volumes, and visually assess urine color to determine hydration status.
- Statistical Analysis page: This offers a comprehensive overview of patients' hydration and urination patterns over multiple time frames. It presents historical data and trends in fluid intake and urine output, displayed on a daily, monthly, and annual basis. Through graphical visualizations, this module empowers patients and health care providers to monitor adherence trends, identify patterns, and adjust intervention strategies to optimize long-term compliance with fluid intake recommendations.
- Health Education Knowledge page: The evidence-based materials tailored to kidney stone prevention are provided, which include detailed information on various types of UC and targeted prevention strategies. This is delivered through accessible formats like articles.
- Function page: The applet will send notification messages at regular intervals, customized based on the patient's daily habits. The notification module delivers customized reminders designed to promote consistent fluid intake. By sending timely and context-relevant messages, the system encourages adherence to prescribed hydration regimens,

XSL•FO
RenderX

addressing barriers such as forgetfulness or situational constraints.

The development of the WAFIR will begin with its registration on the WeChat Official Accounts platform, initiating a structured 6-phase development process as depicted in Figure 3. The first phase, WAFIR requirement analysis, will involve a multidisciplinary team—including urological nurses, doctors, nursing education specialists, nutritionists, and postoperative patients with UC—to conduct needs assessments, define functional specifications (eg, fluid intake reminders and urine output tracking), and develop evidence-based content through stakeholders' feedback. The second phase will be the WAFIR design phase, where a user-friendly interface and function module will be created in collaboration with software developers and health care providers. The third phase of the WAFIR feasibility assessment will evaluate the prototype's technical performance and clinical utility through testing by the development team and clinical professionals. In the fourth phase, the WAFIR acceptability assessment will engage a cohort of 19 participants—comprising 3 urological nurses, 3 urological doctors, 2 nursing education nurses, 1 nutritionist, and 10 postoperative patients—to assess usability and collect feedback using validated tools such as the Mobile App Rating Scale [29] and the mHealth Evidence Reporting and Assessment checklist [30], ensuring that the WAFIR meets patient needs prior to the RCT. The fifth phase, monitoring and maintenance, will focus on ongoing improvements, incorporating user data and feedback to refine content and functionality. Finally, the clinical application phase will implement the fully validated WAFIR in real-world clinical settings for the RCT, aiming to enhance health management and prevention outcomes.

XSL•FO
RenderX

Figure 3. Development of WAFIR flowchart. eHEALS: Electronic Health Literacy Scale; IPAQ-SF: International Physical Activity Questionnaire—Short Form; PHQ-9: Patient Health Questionnaire-9; RCT: randomized controlled trial; WISQOL: Wisconsin Stone Quality of Life Questionnaire.

	Study period					
	Enrollment	Allocation	Postallocation			Close-out
Time point(month)	-t1	0	*TO(0)*	*T1(1)*	*T2(3)*	7
Enrollment:						
Eligibility screen	√					
Informed consent	√					
Allocation		√				
Interventions:						
WeChat applet intervention			←	—	→	
Standard intervention			←	—	→	
Assessments:						
Demographic variables, use of digital devices, health status, dietary habits, social-psychological condition	√	√				
24-h fluid intake, 24-h urine volume, WISQOL, PHQ-9, IPAQ-SF, eHEALS, number of readmissions, outpatients visits, recurrence rate, frequency of hematuresis			←	—	→	
Dissemination:						
Data analysis						√
Writing report						√

Intervention Nurses

The principal investigator (XZ), a clinical nurse on the unit, will train the nursing team on the use of the WAFIR app. Nurses will be involved in securing informed consent from patients by interpreting the purpose, procedures, and anticipated benefits of the study and instructing patients how to use the WAFIR to log their fluid intake. Additionally, they will collect baseline and follow-up data through questionnaires at hospitalization and at 1-month and 3-month intervals. Physicians will evaluate and screen admitted patients to determine whether they meet the inclusion criteria for the research.

Intervention Mechanism

The mechanism of the WAFIR provides a robust and constructive framework by providing a feedback loop, which is depicted in Figure 4, to enhance the adherence awareness of fluid intake for patients with UC, which is applied following these processes: goal setting to define the objective and specify the desired outcome: clearly define the water intake goals for patients, such as the amount of drink to be consumed at designated times; begin by establishing long-term and short-term goals. The long-term goal could be to prevent further formation of UC, while a short-term goal could be to ensure a specific amount of water intake per day. Action planning outlines a detailed process on how to achieve the set goals. This could involve developing a WeChat applet monitoring system to track each patient's water intake, deliver hydration reminders,

XSL•FO
RenderX

formulate strategies to overcome barriers to hydration, provide educational documents, facilitate peer support from the community, and offer access to professional guidance. Essential training is also provided to ensure that patients are adept at using the WAFIR to manage and monitor their water intake. Self-monitoring of behavior: encourage patients to monitor their behavior, for instance, recording their daily water intake, the color of their urine, and the physical activity level they perform daily.

Figure 4. The mechanism of WAFIR. WAFIR: WeChat-Based Applet Fluid Intake Reminder.

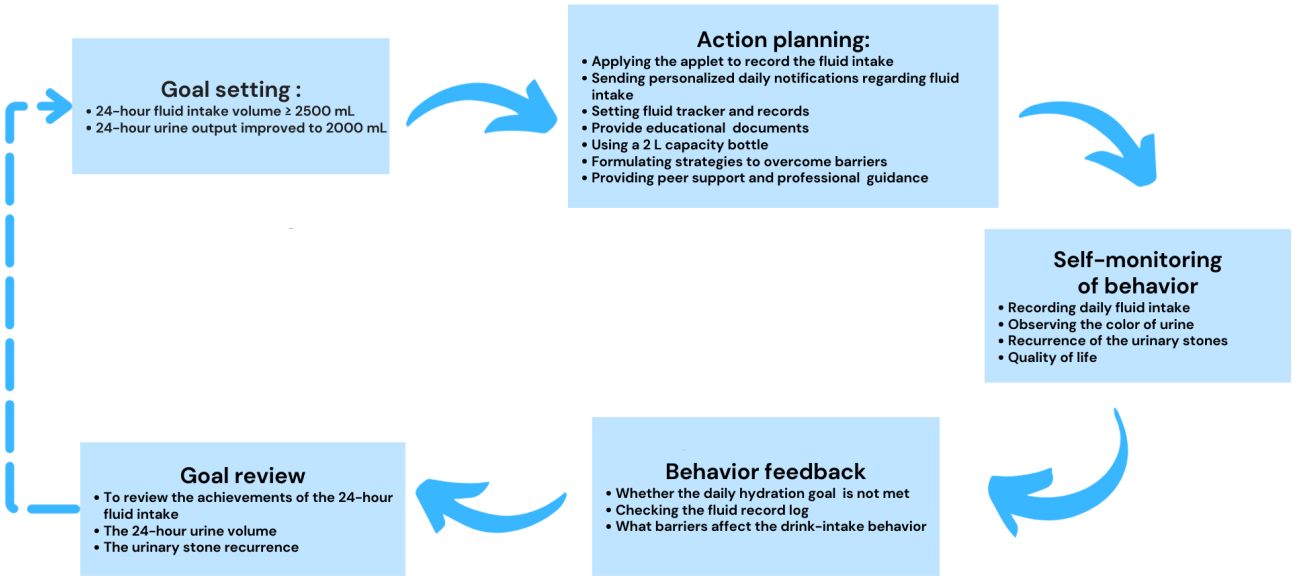

Concurrently, educating patients on the importance of proper hydration and its significant impact on health increases their intrinsic motivation. Nurses assist patients with behavioral skills development by setting reminders, recording water intake, and developing strategies to overcome barriers to hydration. Behavior feedback: the WAFIR incorporates feedback mechanisms to provide real-time feedback to patients. This is using notifications, alerts, or visual indicators to signal when patients' hydration goals are not met. Data collected through the WAFIR are analyzed to recognize patterns of compliance and noncompliance, which help regulate behavior and reduce disparities between current and desired behaviors. This fosters self-monitoring and self-regulation through the WAFIR to enhance patients' sense of self-efficacy for their drink intake. Goal review: regularly review the goals in light of the monitored behavior and feedback, which may involve adjusting the frequency of reminders, fluid intake goals, physical activity levels, dietary habits, or health literacy to align with patient behavior and feedback. This approach is supported by Partridge et al [31], who demonstrated that mHealth interventions using techniques such as goal setting, self-monitoring of behavior, and feedback are more effective in promoting behavior modification than apps lacking these features. Similarly, a systematic review by Middelweerd et al [32] evaluated 41 apps in iTunes and 23 apps in Google Play, identifying goal setting, self-monitoring, and feedback on performance as common approaches to promote intention to behavioral alteration.

Theory-Based Technique

Control theory examines how individuals monitor and modulate their behaviors to attain specific objectives. It underscores the importance of self-regulation, including behavior self-assessment, the establishment of targets, and the use of feedback mechanisms [33]. Within the domain of health behaviors, control theory delineates the strategic adjustment of behaviors to fulfill health-related goals. Meanwhile, control theory, which comprises an input function, a comparator with reference values, an output function, an effect on the environment or others, and disturbance, is majorly applied to bridge the intention-behavior gap in artificial intelligence (AI) for physical action change. The detailed diagram described in Figure 5 closely aligns with the WAFIR mechanism. The goal-setting and action-planning stages correspond to the input function and comparator phases; the self-monitoring of behavior and behavior feedback phases can be reflected in the output function; and the goal review stages can be found in the effects phase of the control theory, with disturbances identified accordingly.

XSL•FO
RenderX

Figure 5. Control theory diagram.

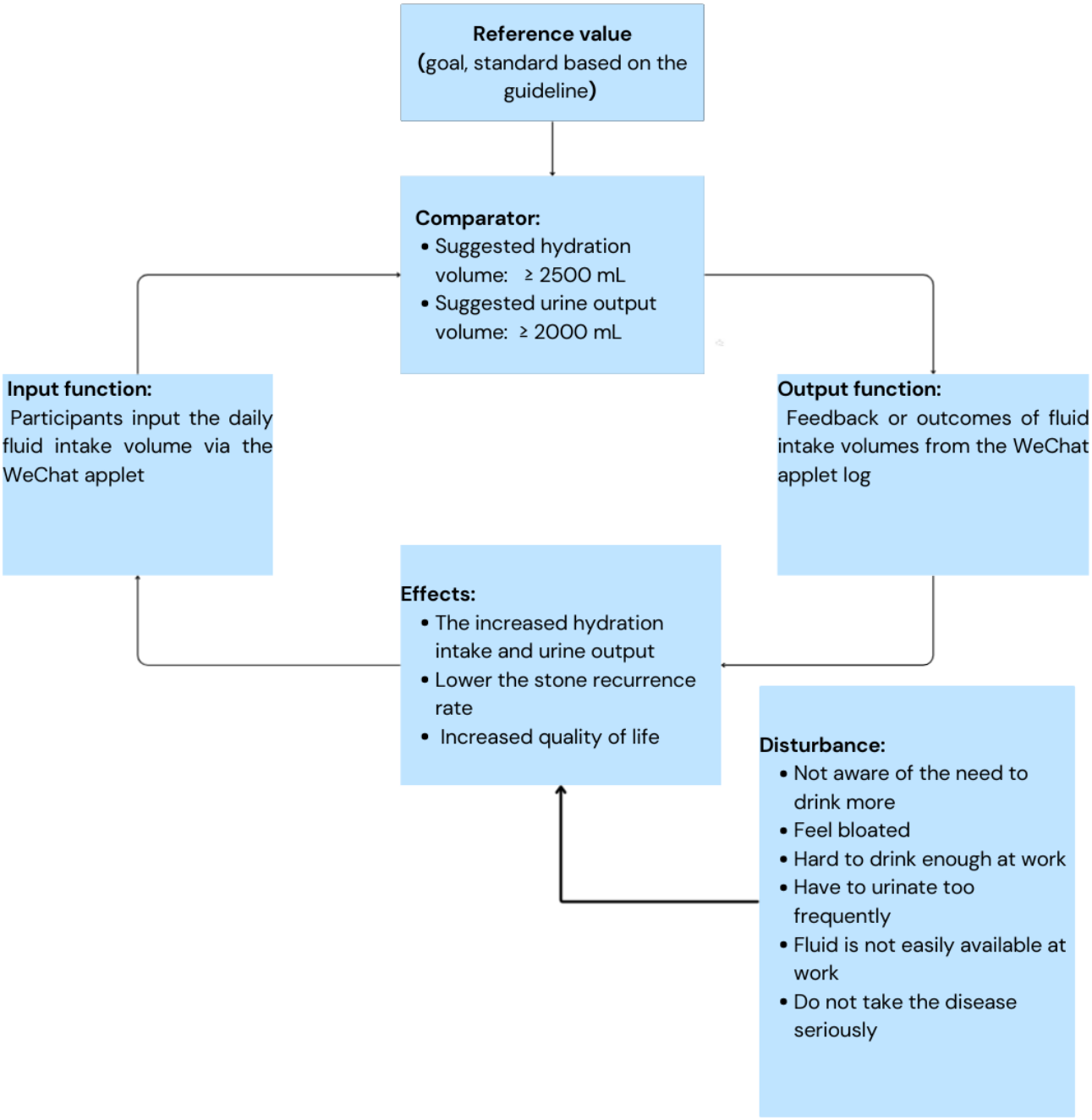

Ethical Considerations

This study received ethics approval from the University of Hong Kong/Hospital Authority Hong Kong West Cluster Ethics Review Board (UW24-754) and the Qianhai Shekou Free Trade Zone Hospital Ethics Committee (2024-055(Shen)-01) in December 2024. Written informed consent will be obtained from all participants, ensuring their right to withdraw, per the Declaration of Helsinki. Participant data will be deidentified, stored securely, and accessed only by authorized personnel to ensure privacy and confidentiality. The trial does not involve physical harm. No compensation will be provided, with voluntary participation incentivized by potential health benefits. No identifiable images or information will be included in the manuscript or supplementary materials. All procedures adhere to institutional and international ethical standards.

Trial Registration

Our trial was retrospectively registered on ClinicalTrials.gov (NCT06990672) on May 24, 2025, after the initiation of recruitment. This delay was unintentional and resulted from unforeseen administrative challenges, including delays in securing institutional registration and coordinating with the tertiary hospital's research office in Mainland China, which were exacerbated by the rapid progression of the study timeline following the development of the WAFIR applet in February 2025. Recruitment commenced on March 1, 2025, and was completed by June 27, 2025, prior to registration, due to these logistical constraints. We affirm that this retrospective registration does not reflect an intent to bias reporting, as the study protocol, including design, outcomes, and intervention details, was finalized and documented internally before recruitment began, adhering to the original research plan.

XSL•FO
RenderX

Outcome Measures

Primary Outcomes

24-Hour Urine Volume

To initiate the 24-hour urine collection, upon waking in the morning, void into a graduated container and document the precise start time (eg, 8 AM). On the subsequent morning, near the 24-hour end point (eg, 8 AM), we record the urine volume once more. Throughout this period, we use platforms such as the WeChat applet or alternative recording methods to track and document the cumulative urine output.

24-Hour Fluid Intake Volume

This measure is to calculate the participants' fluid intake for a full 24-hour period. This comprehensive record should span a full 24-hour period, ideally starting and ending at the same time of day to standardize the assessment. Participants will be instructed to record the fluid consumed using either the WAFIR or another recording method, including water, tea, coffee, Chinese herbal tea, and other eligible liquids. The data on daily fluid consumption volume can be summarized automatically once patients log the data in the WAFIR in the intervention group. However, the participants must manually report the fluid intake volume in the control group. This is the direct parameter that reflects the adherence to fluid intake in the WAFIR.

Secondary Outcomes

Wisconsin Stone Quality of Life Questionnaire

Penniston et al [34] developed the WISQOL as a disease-specific measure to assess the QoL of patients with UC. It encompasses 7 dimensions: vitality and fatigue, sleep, work and family, nutrition (diet) and pharmacotherapy, physical conditions, concerns over intimacy and traveling, and emotional status in general [34]. The scale consists of 28 items, each rated on a 5-point scale from 1=very true to 5=not at all true, yielding a total score range between 28 and 140. A higher score indicates a better QoL. This scale is widely used in multicenter studies in various countries, such as the United States and Canada, for its well-confirmed reliability and validity. Cronbach α coefficient was 0.948, and the reliability was 0.907 in the Chinese version.

Patient Health Questionnaire-9

This assesses participants' mental health status after receiving different interventions, which can help monitor the health status and psychological well-being of patients [35]. Nine items with a cumulative score on the Patient Health Questionnaire-9 (PHQ-9) range from 0 to 27. Cronbach α coefficient was 0.839, and the reliability was 0.846 in the Chinese version.

International Physical Activity Questionnaire—Short Form

This assesses the participants' physical activity levels and the frequency and duration of the workout [36], which can affect the fluid requirements and the risk of UC formation. The questionnaire consists of 7 questions to assess physical activity level over the past week. Activity is classified as low if the total score is <600 metabolic equivalent of task-minutes per week and classified as high if it is >3000 metabolic equivalent of task-minutes per week. The intraclass correlation coefficients for all physical activity categories within the International Physical Activity Questionnaire were above 0.7 in the Chinese version.

Electronic Health Literacy Scale

This scale is designed by Norman and Skinner [37] to measure an individual's self-perceived skills in using information technology for seeking, searching, understanding, and appraising health information from electronic sources. Including 8 items with a 5-point Likert scale, the total score ranges from 8 to 40. The Cronbach α coefficient of the eHEALS scale is 0.963. The retest reliability coefficient r was 0.984 (P<.001) in the Chinese version.

The Number of Readmissions, Outpatient Visits, and Emergency Department Visits

This assesses the frequency of hospital readmissions related to UC, including the timing of such readmissions due to stone recurrence or related symptoms (eg, hematuria, pain, and infection) and the number of outpatient visits in terms of UC management. The frequency of emergency department visits is due to UC. Any medications might affect fluid intake or urinary output or improve the symptoms that triggered postoperatively.

Recurrence Rate

The rate at which new UC forms during the 3-month follow-up.

Frequency of Hematuria or Dark Urine

The frequency of the presence of blood in urine or dark urine, which can be a symptom of a new urinary stone or a signal of less fluid intake.

Sample Size

An adequate sample size and statistical power ensure that the findings can be reliably generalized to a broader population. The required sample size for hypothesis testing is determined based on the expected effect size [38]. In this study, the sample size will be determined by a statistical tool, G*Power software (Heinrich Heine University Düsseldorf) [39], with an α of .05, a power of 0.80, and an allocation ratio of 1:1. The sample size for this study was calculated based on a comparable RCT that used an eHealth device (an intelligent water bottle) to promote increased fluid intake [18], of which the calculated Cohen d was 0.61, and the effect size is considered significant since Cohen d is over 0.5. Based on the calculations, a sample size of 92 participants is estimated. Accounting for an attrition rate of 38%, as observed in this research, a total of 148 participants are anticipated to be recruited.

Randomization, Blinding, and Concealed Allocation

An RCT will be conducted to recruit participants in a single site, the Shekou Free Trade Zone Hospital. Randomization may lead to baseline characteristic imbalances between intervention groups [40]. In small-scale trials, stratification might be beneficial, as it can prevent significant imbalances in factors predicting outcomes [41]. Stratified randomization will be applied to the eligible participants to ensure an equal chance to be allocated to the research field [42]. Balanced randomization aims to create unbiased groups for comparison and also seeks

XSL•FO
RenderX

to maintain groups of roughly equivalent size throughout the trial [43]. In this research, convenience sampling will be applied, with block randomization applied using a block size of 10. For every group of 10 participants enrolled in sequence, 5 will be allocated to the intervention group and the remaining 5 to the control group, which will be done by the statistician and urologists who are not engaged in conducting the research interventions.

Data Collection

The study will be conducted on postoperative patients with UC. It will span 3 periods: T0 (before the intervention), T1 (at the end of 1 month), and T2 (at the end of 3 months). The data collected at each time point will allow for evaluating the intervention's effectiveness and identifying areas for improvement. Reminders will be sent prior to each follow-up visit using the participant's preferred method of communication, with a thank-you message after the visit. For any participant who discontinues the intervention prematurely, we will record the reason for their discontinuation and collect any serious adverse events that occurred up to the point of discontinuation.

Methods of Analysis

The study outcomes will be reported to adhere to the guidelines outlined in the CONSORT (Consolidated Standards of Reporting Trials) 2010 statements [44]. Statistical tests such as a 2-tailed *t* test and chi-square test will be used for continuous and categorical variables within 2 groups. The principal analysis will compare the fluid intake volume between the intervention and control groups using the 2-tailed *t* test. Secondary outcomes such as WISQOL, PHQ-9, and eHEALS will use repeated measures of ANOVA since they are measured over time and between 2 groups. If necessary, subgroup analyses will be conducted to explore whether the effect of the intervention varies across different demographic or clinical subgroups. Intention-to-treat analysis will be used to assess the impact of the treatment and its generalizability as well as the missing data. A planned interim analysis will be performed to evaluate both efficacy (whether the intervention shows sufficient benefit to potentially stop the trial early) and futility (whether the intervention is unlikely to demonstrate benefit if the trial continues) after 50% (74/148) of the target sample size have completed the primary end-point assessment. A significance level of $P<.05$ will be used to determine statistical significance, and all data analysis procedures will be conducted using the SPSS (version 26.0; IBM Corp) software package.

Access to Data

The researcher is granted access to the complete trial dataset, with a contractual agreement established with the statistician to manage the data confidentially in accordance with the research protocol.

Results

This research has achieved notable advancements in assessing the efficacy of a nurse-led digital health intervention, WAFIR. The development of WAFIR, conducted in collaboration with a software developer, was finalized in February 2025, establishing a technological platform for the study. Recruitment, a critical initial stage, was fully completed, with all 148 participants enrolled from a tertiary hospital in Mainland China between March 1, 2025, and June 27, 2025, exclusively during the predischarge hospitalization period. Baseline data collection (T0) was completed by June 30, 2025, providing a solid baseline for longitudinal analysis. Follow-up assessments at the 1-month mark (T1) were finalized on July 31, 2025, while the 3-month follow-up (T2) is set to conclude in October 2025, aligning with the present date. Subsequent data analysis is scheduled for completion by December 2025, with dissemination of results planned through academic conferences and peer-reviewed publications throughout 2026.

Discussion

This study represents, to our knowledge, one of the first RCTs to assess a nurse-led, WeChat-based DTx specifically tailored for UC management, addressing a critical gap in mHealth interventions that often lack medical oversight or clinical validation. The primary outcomes of this study are to evaluate the effectiveness of the WAFIR in improving fluid intake and urine output among postoperative patients with UC. We hypothesize that the WAFIR intervention, combining fluid intake reminders, urine color monitoring, fluid and urine tracking, and evidence-based educational content, will significantly enhance adherence to hydration recommendations compared to standard care, potentially lowering recurrence rates and improving patient-reported outcomes such as QoL (via WISQOL), eHealth literacy (via eHEALS), and mental health (via PHQ-9).

This study used face-to-face and telephone interviews to collect data. However, interviews are subject to limitations, including the absence of visual cues (particularly in telephone interviews), reduced ability to observe contextual details, and potential difficulties in eliciting sensitive information, which may affect data quality [45]. The primary outcome, 24-hour urine volume, is required to be collected. However, previous studies indicate challenges in obtaining precise 24-hour urine samples due to poor patient compliance, which may result in incomplete or abandoned collections stemming from the complex or inconvenient nature of the process, as well as deviations from the 24-hour collection period. To enhance the measurement accuracy of 24-hour urine collection, it is crucial to improve patient education on the importance of urine collection; simplify the collection process with clear guidelines, flowcharts, and user-friendly containers; and provide necessary tools like graduated urinals or collection bags. Regular monitoring and feedback sessions are essential to address any issues promptly. Additionally, incentivizing participation with rewards can boost patient motivation, ensuring higher compliance and accurate urine volume measurements within the specified time frame.

A key strength of this study is its rigorous RCT design, with clear inclusion criteria, randomization, and intention-to-treat analysis, ensuring robust evidence generation. However, there are several limitations to this protocol that should be acknowledged. First, patients were recruited from a single site. This limitation may affect the generalizability of the findings, as the sample may not be representative of the broader

XSL•FO
RenderX

population. Second, the current WAFIR is equipped with a fluid record function, evidence-based educational content delivery, automated reminders and notifications, and data collection and reporting. AI technologies, such as AI-driven predictive models, are not equipped within the WAFIR, as the integration could enhance its effectiveness by providing tailored health education, motivational messages, and personalized recommendations based on individual patient data and behavior patterns.

The dissemination plan includes presenting research findings at international conferences and forums and publishing in peer-reviewed journals in 2026. If proven effective, WAFIR could establish a pathway for developing other nurse-led digital interventions, offering a low-cost and scalable approach to enhance patient outcomes in chronic disease management. Furthermore, the findings will guide app developers in creating clinically validated mHealth tools and assist clinicians in integrating such interventions into standard urolithiasis care, thereby potentially transforming postoperative management practices.

Acknowledgments

The authors would like to thank He Ruoying for her valuable assistance with the sample size calculations and certain statistical aspects of the research. The authors disclose that generative artificial intelligence (specifically, ChatGPT or DeepSeek or QuillBot) was used during the preparation of this manuscript. The tool was used solely for the purpose of improving the language and readability of the text.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Data Availability

The datasets generated and/or analyzed will not be publicly available due to privacy and ethical restrictions. However, requests for data access will be considered on an individual basis by contacting the corresponding author.

Authors' Contributions

XZ and TW conceptualized the research design and developed the WeChat-based digital therapeutic intervention, and TW led the protocol development, statistical planning, and the ethical considerations for the randomized controlled trial. XZ drafted the initial manuscript. QC coordinated participant recruitment strategies, ensuring the effective enrollment of patients with urinary calculi. All authors (XZ, TW, and QC) critically reviewed, drafted, and approved the final manuscript and agree to be accountable for all aspects of the work, ensuring its accuracy and integrity.

Conflicts of Interest

None declared.

References

1. Shafi H, Moazzami B, Pourghasem M, Kasaeian A. An overview of treatment options for urinary stones. Caspian J Intern Med 2016;7(1):1-6. [Medline: 26958325]
2. Sorokin I, Mamoulakis C, Miyazawa K, Rodgers A, Talati J, Lotan Y. Epidemiology of stone disease across the world. World J Urol 2017 Sep;35(9):1301-1320. [doi: 10.1007/s00345-017-2008-6] [Medline: 28213860]
3. Romero V, Akpinar H, Assimos DG. Kidney stones: a global picture of prevalence, incidence, and associated risk factors. Rev Urol 2010;12(2-3):e86-e96. [Medline: 20811557]
4. Zeng G, Mai Z, Xia S, et al. Prevalence of kidney stones in China: an ultrasonography based cross-sectional study. BJU Int 2017 Jul;120(1):109-116. [doi: 10.1111/bju.13828] [Medline: 28236332]
5. Johnson CM, Wilson DM, O'Fallon WM, Malek RS, Kurland LT. Renal stone epidemiology: a 25-year study in Rochester, Minnesota. Kidney Int 1979 Nov;16(5):624-631. [doi: 10.1038/ki.1979.173] [Medline: 548606]
6. Pearle MS, Calhoun EA, Curhan GC, Urologic Diseases of America Project. Urologic diseases in America project: urolithiasis. J Urol 2005 Mar;173(3):848-857. [doi: 10.1097/01.ju.0000152082.14384.d7] [Medline: 15711292]
7. Moftakhar L, Jafari F, Ghoddusi Johari M, Rezaeianzadeh R, Hosseini SV, Rezaianzadeh A. Prevalence and risk factors of kidney stone disease in population aged 40-70 years old in Kharameh cohort study: a cross-sectional population-based study in southern Iran. BMC Urol 2022 Dec 19;22(1):205. [doi: 10.1186/s12894-022-01161-x] [Medline: 36536352]
8. Becerra AZ, Khusid JA, Sturgis MR, et al. Contemporary assessment of the economic burden of upper urinary tract stone disease in the United States: analysis of one-year health care costs, 2011-2018. J Endourol 2022 Apr;36(4):429-438. [doi: 10.1089/end.2021.0485] [Medline: 34693752]
9. Lotan Y, Buendia Jiménez I, Lenoir - Wijnkoop I, et al. Primary prevention of nephrolithiasis is cost - effective for a national healthcare system. BJU Int 2012 Dec;110(11c):E1060-E1067. [doi: 10.1111/j.1464-410X.2012.11212.x]

XSL•FO
RenderX

10. Fakheri RJ, Goldfarb DS. Ambient temperature as a contributor to kidney stone formation: implications of global warming. Kidney Int 2011 Jun;79(11):1178-1185. [doi: 10.1038/ki.2011.76] [Medline: 21451456]
11. Mitra P, Pal DK, Das M. Does quality of drinking water matter in kidney stone disease: a study in West Bengal, India. Investig Clin Urol 2018 May;59(3):158-165. [doi: 10.4111/icu.2018.59.3.158] [Medline: 29744472]
12. Mathiyalagen P, Neelakantan A, Balusamy K, Vasudevan K, Cherian J, Sunderamurthy B. A case-control study on environmental and biological risk factors for renal calculi persisting in a coastal Union Territory, India. J Family Med Prim Care 2017;6(1):126-130. [doi: 10.4103/2249-4863.214981] [Medline: 29026764]
13. Ferraro PM, Bargagli M, Trinchieri A, Gambaro G. Risk of kidney stones: influence of dietary factors, dietary patterns, and vegetarian-vegan diets. Nutrients 2020 Mar 15;12(3):779. [doi: 10.3390/nu12030779] [Medline: 32183500]
14. Cheraghian B, Meysam A, Hashemi SJ, et al. Kidney stones and dietary intake in adults: a population-based study in southwest Iran. BMC Public Health 2024 Apr 4;24(1):955. [doi: 10.1186/s12889-024-18393-1] [Medline: 38575950]
15. Streeper NM, Dubnansky A, Sanders AB, Lehman K, Thomaz E, Conroy DE. Improving fluid intake behavior among patients with kidney stones: understanding patients' experiences and acceptability of digital health technology. Urology 2019 Nov;133:57-66. [doi: 10.1016/j.urology.2019.05.056] [Medline: 31374289]
16. McCauley LR, Dyer AJ, Stern K, Hicks T, Nguyen MM. Factors influencing fluid intake behavior among kidney stone formers. J Urol 2012 Apr;187(4):1282-1286. [doi: 10.1016/j.juro.2011.11.111] [Medline: 22341296]
17. Khambati A, Matulewicz RS, Perry KT, Nadler RB. Factors associated with compliance to increased fluid intake and urine volume following dietary counseling in first-time kidney stone patients. J Endourol 2017 Jun;31(6):605-610. [doi: 10.1089/end.2016.0836] [Medline: 28318298]
18. Stout TE, Lingeman JE, Krambeck AE, et al. A randomized trial evaluating the use of a smart water bottle to increase fluid intake in stone formers. J Ren Nutr 2022 Jul;32(4):389-395. [doi: 10.1053/j.jrn.2021.07.007] [Medline: 35283036]
19. Tasian GE, Ross M, Song L, et al. Ecological momentary assessment of factors associated with water intake among adolescents with kidney stone disease. J Urol 2019 Mar;201(3):606-614. [doi: 10.1016/j.juro.2018.07.064] [Medline: 30076905]
20. DTA's adoption & interpretation of ISO's DTx definition. Digital Therapeutics Alliance. 2023. URL: https://dtxalliance.org/wp-content/uploads/2023/06/DTA_FS_New-DTx-Definition.pdf [accessed 2026-02-11]
21. Ceci L. Number of apps available in leading app stores. Statista. 2026 Jan 13. URL: https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/ [accessed 2026-02-06]
22. Li H, Zheng S, Li D, et al. The establishment and practice of pharmacy care service based on internet social media: telemedicine in response to the COVID-19 pandemic. Front Pharmacol 2021;12:707442. [doi: 10.3389/fphar.2021.707442]
23. Hong Z, Li N, Li D, et al. Telemedicine during the COVID-19 pandemic: experiences from Western China. J Med Internet Res 2020 May 8;22(5):e19577. [doi: 10.2196/19577] [Medline: 32349962]
24. Chen X, Zhou X, Li H, Li J, Jiang H. The value of WeChat application in chronic diseases management in China. Comput Methods Programs Biomed 2020 Nov;196:105710. [doi: 10.1016/j.cmpb.2020.105710] [Medline: 32858284]
25. Stevens DJ, McKenzie K, Cui HW, Noble JG, Turney BW. Smartphone apps for urolithiasis. Urolithiasis 2015 Feb;43(1):13-19. [doi: 10.1007/s00240-014-0738-0] [Medline: 25410731]
26. Karagöz MA, Sarıca K. Patient compliance to dietary recommendations: tips and tricks to improve compliance rates. World J Urol 2023 May;41(5):1261-1268. [doi: 10.1007/s00345-023-04318-x] [Medline: 36786816]
27. Philip-McKenzie Y, Jamnadass E, Hameed BZ, et al. A content analysis of "Water Apps" and prevention of urological diseases: do apps really help? Cent Eur J Urol 2020;73(2):187-192. [doi: 10.5173/ceju.2020.0136] [Medline: 32782839]
28. Winoker JS, Koo K, Matlaga BR. Barriers to the adoption and integration of patient-facing mHealth tools in urology. Nat Rev Urol 2021 May;18(5):247-249. [doi: 10.1038/s41585-021-00444-6] [Medline: 33692501]
29. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27. [doi: 10.2196/mhealth.3422] [Medline: 25760773]
30. Agarwal S, LeFevre AE, Lee J, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
31. Partridge SR, McGeechan K, Hebden L, et al. Effectiveness of a mHealth lifestyle program with telephone support (TXT2BFiT) to prevent unhealthy weight gain in young adults: randomized controlled trial. JMIR Mhealth Uhealth 2015 Jun 15;3(2):e66. [doi: 10.2196/mhealth.4530] [Medline: 26076688]
32. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11(1):97. [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
33. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. Psychol Bull 1982 Jul;92(1):111-135. [Medline: 7134324]
34. Penniston KL, Antonelli JA, Viprakasit DP, et al. Validation and reliability of the Wisconsin Stone Quality of Life Questionnaire. J Urol 2017 May;197(5):1280-1288. [doi: 10.1016/j.juro.2016.11.097] [Medline: 27889419]

XSL•FO
RenderX

35. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
36. Craig CL, Marshall AL, Sjöström M, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB]
37. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
38. Ialongo C. Understanding the effect size and its measures. Biochem Med (Zagreb) 2016;26(2):150-163. [doi: 10.11613/BM.2016.015] [Medline: 27346958]
39. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
40. Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001 Apr 17;134(8):663-694. [doi: 10.7326/0003-4819-134-8-200104170-00012] [Medline: 11304107]
41. Schulz KF, Grimes DA. Generation of allocation sequences in randomised trials: chance, not choice. Lancet 2002 Feb 9;359(9305):515-519. [doi: 10.1016/S0140-6736(02)07683-3] [Medline: 11853818]
42. Bhide A, Shah PS, Acharya G. A simplified guide to randomized controlled trials. Acta Obstet Gynecol Scand 2018 Apr;97(4):380-387. [doi: 10.1111/aogs.13309] [Medline: 29377058]
43. Schulz KF. Randomized controlled trials. Clin Obstet Gynecol 1998 Jun;41(2):245-256. [doi: 10.1097/00003081-199806000-00005] [Medline: 9646957]
44. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. J Pharmacol Pharmacother 2010 Jul;1(2):100-107. [doi: 10.4103/0976-500X.72352] [Medline: 21350618]
45. Novick G. Is there a bias against telephone interviews in qualitative research? Res Nurs Health 2008 Aug;31(4):391-398. [doi: 10.1002/nur.20259] [Medline: 18203128]

Abbreviations

AI: artificial intelligence
CONSORT: Consolidated Standards of Reporting Trials
DTx: digital therapeutic intervention
eHEALS: Electronic Health Literacy Scale
mHealth: mobile health
PHQ-9: Patient Health Questionnaire-9
QoL: quality of life
RCT : randomized controlled trial
UC: urinary calculi
WAFIR: WeChat-Based Applet Fluid Intake Reminder
WISQOL : Wisconsin Stone Quality of Life Questionnaire

Edited by A Schwartz; submitted 06.Jul.2025; peer-reviewed by S Wang; accepted 20.Jan.2026; published 19.Feb.2026.

Please cite as:
Zhou X, Wang T, Chen Q
Developing and Evaluating a WeChat-Based Applet Fluid Intake Reminder on Enhancing Fluid Adherence in Postoperative Patients With Urinary Calculi: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e80214
URL: https://www.researchprotocols.org/2026/1/e80214
doi:10.2196/80214

© Xue Zhou, Tongyao Wang, Qian Chen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

PATH Trial for Examining Yoga as a Strategy to Improve Remote-Based Weight Loss in Adults: Protocol for a Randomized Controlled Trial

Selene Y Tobin[1,2,3], PhD; Sally Sherman[4], PhD; Beth Bock[2], PhD; Shira Dunsiger[5], PhD; J Graham Thomas[1,2], PhD; Stephanie P Goldstein[1,2], PhD; Jessica Unick[1,2], PhD

[1]Weight Control and Diabetes Research Center, The Miriam Hospital, 196 Richmond Street, Providence, RI, United States

[2]Department of Psychiatry and Human Behavior, Brown University, Providence, RI, United States

[3]Department of Health and Kinesiology, California Polytechnic University, San Luis Obispo, CA, United States

[4]Department of Health and Human Development, Physical Activity Research Center, University of Pittsburgh, Pittsburgh, PA, United States

[5]Department of Behavioral and Social Sciences, Brown University, Providence, RI, United States

Corresponding Author:
Jessica Unick, PhD
Weight Control and Diabetes Research Center, The Miriam Hospital, 196 Richmond Street, Providence, RI, United States

Abstract

Background: Over 70% of US adults are overweight or have obesity. Internet-delivered weight loss (IDWL) treatment overcomes many common barriers to in-person weight loss (WL) programs; yet, IDWLs underperform compared to in-person treatment. Yoga is a novel intervention that targets physical and psychological health and can be delivered virtually, increasing translation potential. Yoga has been understudied as an intervention to address barriers to WL despite favorable effects in other health disciplines. Preliminary studies suggest yoga to improve self-regulation and reduce lapses from dietary recommendations, 2 factors integral to long-term WL success.

Objective: The aim of the study is to conduct the first fully powered randomized controlled trial to test the added effect of yoga to IDWL treatment on weight (primary outcome), dietary lapses, potential lapse triggers, and self-regulation mediator variables (secondary outcomes).

Methods: Adults that are overweight or have obesity will be recruited to participate in a 12-month IDWL program. Following a 3-month assessment, eligible participants will be randomized to (1) a yoga intervention, including 14 weeks of group-based, virtual Iyengar yoga delivered twice weekly, followed by 22 weeks of yoga delivered once per week, and then 6 months of no contact, or (2) a contact-matched wellness comparison condition. Assessments will take place at baseline, 3 months, 6 months, 12 months, and 18 months. Several questionnaires, ecological momentary assessments, and objectively measured weight and accelerometry data will be collected. This trial is registered on ClinicalTrials.gov (NCT06166225).

Results: This randomized trial was funded with an original project period of July 2023-June 2028. Enrollment began in February 2024 and is ongoing with the expectation of closing enrollment in June 2026.

Conclusions: This trial will examine the combined effect of yoga and WL treatment among individuals with overweight or obesity and will be the first fully powered trial to assess the combination of virtually delivered yoga and WL beyond 6 months. Findings will inform whether yoga could be considered a valuable intervention component for improving IDWL treatment.

Trial Registration: ClinicalTrials.gov NCT06166225; https://clinicaltrials.gov/study/NCT06166225

International Registered Report Identifier (IRRID): DERR1-10.2196/78973

(JMIR Res Protoc 2026;15:e78973) doi:10.2196/78973

KEYWORDS

internet-delivered weight loss; long-term weight loss; yoga; randomized controlled trial; obesity; overweight

Introduction

Over 70% of US adults are overweight or have obesity [1]. Internet-delivered weight loss (IDWL) treatment overcomes many common barriers to in-person weight loss (WL) programs (eg, travel, childcare, and geographical constraints), and emerging evidence indicates that online delivery of wellness programs is effective and acceptable [2-5]. Nonetheless, IDWL often underperforms when compared to in-person treatment, and weight regain following WL is still highly prevalent [6,7].

XSL•FO
RenderX

Thus, novel approaches for improving long-term IDWL are needed.

Yoga is an ancient practice that encompasses physical and mental elements that promote overall well-being, flexibility, strength, and mental clarity [8]. Yoga is often most recognized for its physical postures, which are associated with movement and physical activity (PA), but its broader benefits extend to emotional and psychological health. Interdisciplinary findings demonstrate the downstream effects of yoga to improve aspects of all 3 dimensions of self-regulation. This includes having a favorable impact on mindfulness and distress tolerance (emotion regulation), inhibitory control (cognitive regulation), self-efficacy, and self-compassion (self-related processes) [9-29]. These effects are of particular interest, as prior obesity medicine research indicates that individuals successful at long-term WL have high self-regulatory capacity and exhibit more positive self-related processing [30-36]. Therefore, yoga may offer a novel pathway to cultivate these weight-related behavioral skills. Yoga may also reduce dietary lapses and improve WL via improvements in psychological factors that are associated with or trigger a lapse. For example, yoga has been shown to improve affect (ie, improved positive affect and reduced negative affect) [37,38], and prior research shows that more negative affect is associated with increased frequency of dietary lapses [31,39,40]. Together, these findings suggest that yoga targets multiple self-regulatory and affective processes, which can exert a synergistic effect on eating behavior and body weight. Thus, the rationale for using yoga to reduce dietary lapses and improve long-term WL in IDWL treatment is strong.

To date, yoga has largely been overlooked within the context of WL treatment, despite positive effects observed in other health disciplines (eg, cardiovascular disease risk, pain, depression, and smoking cessation) [41-44]. While other disciplines focus on the psychological benefits of yoga, the weight control field has traditionally viewed yoga as a *mode* of exercise (characterized by low energy expenditure and classified as light PA that may not elicit a sufficient caloric deficit to contribute to WL), thereby overlooking the cognitive and emotional pathways through which yoga may enhance WL. Although cross-sectional evidence suggests strong associations between yoga and healthy dietary choices, mindful eating, and emotional eating [25,45-47], studies investigating the inclusion of yoga within WL treatment are limited. A 2021 systematic review that explored relationships between yoga participation and reductions in body weight via decreased energy intake or increased PA identified only 2 prior randomized controlled trials (RCTs), which combined yoga with calorie reduction in individuals with overweight or obesity [48]. The first was a 6-month study conducted among 50 individuals with overweight or obesity and compared 2 styles of yoga, restorative hatha and vinyasa. While this study demonstrated yoga participation within a standard behavioral WL intervention to be feasible, no differences in WL were observed between the 2 styles of yoga, as both groups lost 3.4 - 3.8 kg [49]. Further, because it did not have a control condition, any conclusion related to the added effect of yoga on WL is limited. The second study was a 12-week trial conducted in 260 individuals diagnosed with metabolic syndrome and was conducted outside of the United States (India), where yoga practice differs from the United States [50]. Results suggested that a greater proportion of participants randomized to the yoga-based intervention recovered from metabolic syndrome (45.4%) as compared to dietary intervention alone (32.3%). However, energy cost was not controlled for in this study, and observed effects may have been due to elevated PA levels.

We recently published findings from our pilot RCT (n=60), wherein 60 female individuals with overweight and obesity were randomized to a 3-month yoga program or a contact-matched control (cooking or nutrition classes) following 3 months of behavioral WL treatment [4,51]. Adherence (attendance to the yoga classes was 75%) and retention (100%) were high. Yoga participants reported fewer lapses from the recommended diet (eg, via less overeating, stress eating, and loss of control when eating) and less difficulty resisting dietary temptations compared to the control group. Post-hoc analyses revealed that yoga was associated with a larger magnitude of 6-month WL and several self-regulatory skills (mindfulness, distress tolerance, and self-compassion) compared to the control, among those who lost more weight in the first 3 months of treatment (≥5%). These results led to the hypothesis that yoga can not only decrease the likelihood of dietary lapse but also reduce many common factors, which could potentially trigger a dietary lapse contributing to longer-term WL success. Additionally, in this pilot trial, we found online delivery of a yoga program to be feasible. Specifically, due to COVID-19, one cohort in this study received yoga classes remotely (live via videoconferencing), and attendance and satisfaction ratings were comparable to in-person yoga classes administered in the previous cohort. These preliminary findings demonstrate initial feasibility and acceptability of including yoga within the context of a WL program and indicate that yoga can reduce dietary lapses and improve self-regulation, 2 factors important for long-term weight control.

The purpose of this study, PATH Weight Loss Trial (Pathway to Achieving Total Health and Weight Loss), is to extend upon this pilot study and address important limitations in the current literature in several ways. First, this will be the first fully powered RCT to assess the added effect of yoga to IDWL treatment on longer-term WL outcomes (primary outcome is 12-month WL [immediately after the intervention], and secondary outcome is 18-month WL [following a 6-month no intervention period]). Investigation of long-term WL outcomes is currently a high priority of obesity medicine, as high rates of weight regain following WL persist. This trial also extends the literature by examining the effect of yoga on dietary lapses and potential lapse triggers (eg, affect, craving, and dietary temptations) to better understand how yoga may influence eating behaviors within the context of WL treatment. It will also examine potential self-regulatory mechanistic pathways through which yoga may impact WL and dietary lapses, which have not yet been explored within WL treatment. Finally, it is also anticipated that results will provide a higher degree of generalizability than standard weight control literature by including male individuals and greater racial or ethnic diversity than previous literature.

XSL•FO
RenderX

Methods

Study Setting

The PATH Trial is a randomized clinical trial funded by the National Center for Complementary and Integrative Health (R01AT011868) and is registered on ClinicalTrials.gov (NCT06166225). While the trial is entirely remote, the study will be conducted out of the Weight Control and Diabetes Research Center in Providence, Rhode Island. The Weight Control and Diabetes Research Center is a joint research institution affiliated with Brown University Health and The Warren Alpert Medical School at Brown University.

Ethical Considerations

All study procedures were developed in accordance with the Declaration of Helsinki ethical principles and approved by the institutional review board at Brown University Health (IRB # 1947629). Participants will be asked to complete informed consent via REDCap (Research Electronic Data Capture; Vanderbilt University) after attending an online study orientation session with a member of the research team. Participants will be eligible for compensation up to US $75 in the form of a gift card for the completion of assessments at 3, 6, 12, and 18 months. The US $75 compensation is calculated in the following manner: participants will receive US $25 for completing the online questionnaires and wearing the PA monitor; they will also be compensated US $0.50 for each ecological momentary assessment (EMA) survey completed (50 surveys per assessment period) and will receive a US $25 bonus for EMA compliance >80%. No compensation will be provided for the baseline assessment visit. The full study protocol, informed consent document, and study results will be published on ClinicalTrials.gov upon study completion. To protect participant confidentiality, all study data collected via our secure website using an active Secure Sockets Layer or Transport Layer Security certificate (HTTPS), or other platforms, will be securely stored on institutionally managed, access-controlled servers protected by firewalls, multifactor authentication, and regular security monitoring. All physical data will be stored securely in locked cabinets, with access strictly limited to authorized and trained study personnel.

Study Overview

This protocol is reported in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2025 guidelines for randomized trials, and any significant modifications to the protocol will require prior approval from the study sponsor. An overview of the study design is shown in Figure 1. A total of 210 participants will enroll in this randomized controlled clinical trial with a parallel-group design and equal allocation. All participants will receive a 12-month IDWL program, and those completing the 3-month assessment visit will be randomized to receive (1) a 9-month yoga program (YOGA; delivered during months 4 - 12) or (2) a 9-month contact-matched control "Health & Wellness program" (WELLNESS; delivered during months 4 - 12). The choice to delay randomization to YOGA or WELLNESS until after the 3-month assessment, rather than at baseline, was made to give participants time to build foundational WL skills without added burden and to reinforce self-regulatory skills at a time (3 months) when WL-related motivation may begin to dip [7,52]. Additional rationale to delay randomization to YOGA or WELLNESS is provided in the discussion. A full description of all intervention components is detailed below and conforms with the TiDieR (Template for Intervention Description and Replication) checklist for intervention reporting to enhance clarity and replicability.

Yoga and wellness classes will be delivered via videoconferencing software (ie, Health Insurance Portability and Accountability Act [HIPAA]–compliant Zoom; Zoom Video Communications). There will be no intervention during months 13 - 18. Randomization will occur at the 3-month assessment (ie, 3 months into the IDWL) in equal numbers based on a permuted block randomization procedure with small, random sized blocks, stratified by biological sex, racial or ethnic minority status, and percent WL achieved during the first 3 months of the WL program (<5% vs ≥5% WL). The randomization scheme was generated by the study statistician using R Studio (version 3.6.0) and uploaded into REDCap directly (to ensure study staff was kept blinded toward the full schematic). Participants will be recruited in cohorts over the study period. Assessments of weight, aerobic PA minutes, dietary lapses, potential lapse triggers (eg, affect, craving, and dietary temptations), and self-regulation mediator variables will be measured at baseline, 3 months (just prior to randomization), 6 months (mid-treatment), 12 months (postintervention), and 18 months (following a 6-month period of no intervention). At each assessment period, data will be collected via a series of REDCap questionnaires, an Actigraph accelerometer, and a 10-day EMA protocol (ie, 5 surveys per day over 10 days with each survey taking 2 - 5 minutes to complete).

XSL•FO
RenderX

Figure 1. Study design. Participants will complete a 12-month weight loss program. Following 3 months of weight loss, participants will be randomized to either yoga or the wellness comparison condition. Yoga and wellness classes will be held twice weekly for 14 weeks, then reduced to once weekly for 22 weeks, followed by a 6-month period of no contact. Study assessments will take place at baseline, 3 months (immediately following weight loss, prior to intervention randomization), 6 months, 12 months, and 18 months.

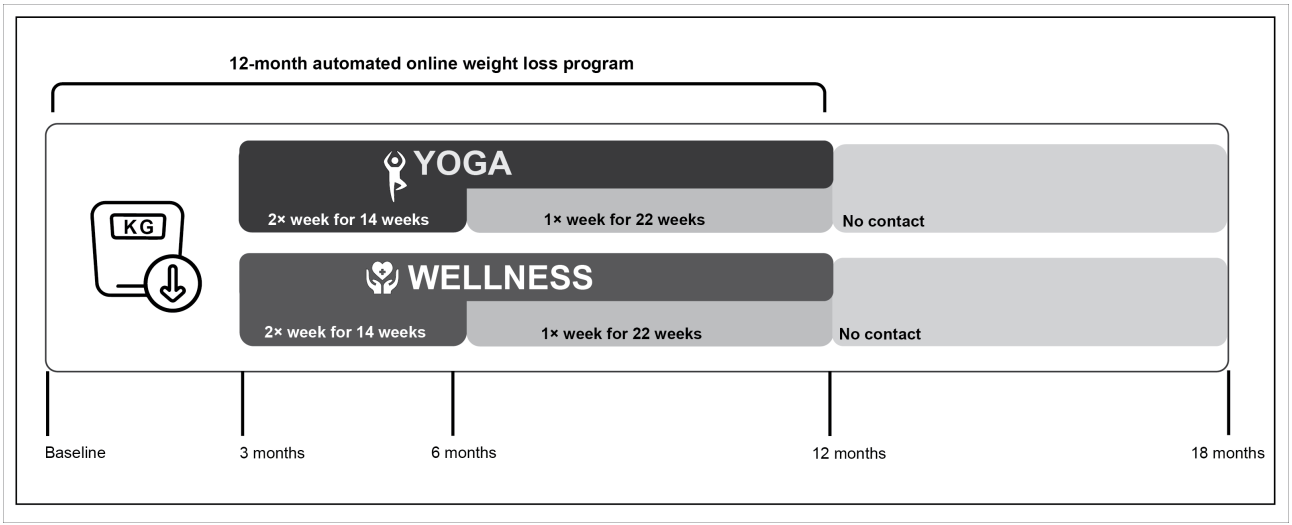

Study Aims

The primary aim of the PATH Trial is to compare the treatment effects of IDWL and YOGA versus IDWL and WELLNESS on weight change at 12 months. The primary weight outcome is absolute weight change from baseline to 12 months. It is hypothesized that WL will be greater in those randomized to the YOGA when compared to WELLNESS.

Secondary aims of this study are (1) to compare YOGA and WELLNESS at 12 months on dietary lapses and potential lapse triggers as assessed via EMA and (2) to examine self-regulation mediators (eg, mindfulness, self-compassion, distress tolerance, self-efficacy, and inhibitory control) of the treatment effect on primary (ie, weight change) and secondary outcomes (ie, dietary lapses) at 12 months.

Exploratory aims are to examine the long-term (18 months) effects of yoga on primary and secondary outcomes, to explore potential moderators (3-month WL, race or ethnicity, biological sex, and BMI) of the treatment effect on weight change, and to explore the effect of yoga on participation in aerobic exercise.

Study Participants, Eligibility Criteria, and Recruitment Procedures

A total of 210 WL-seeking male and female individuals will be recruited for this study. Specifically, participants will be between the ages of 18 and 65 years and have a BMI of 25 - 40 kg/m^2, regardless of race or ethnicity. Participants must own a smartphone (for EMA surveys) and have daily, home internet access that allows for videoconferencing. Exclusion criteria will include currently pregnant, planning to become pregnant within the next 18 months, or pregnant within the past 6 months; current or recent enrollment (<2 years) in a WL program; current or recent yoga, meditation, or mindfulness practice, either at home or in a studio (participants who reported ≥12 sessions within the past 2 years will be followed up with to confirm eligibility); presence of any medical condition for which yoga, exercise, WL, or dietary restriction is contraindicated (Physical Activity Readiness Questionnaire); recent WL (≥4.5 kg within the past 6 months); currently taking WL medications or history of bariatric surgery; inability to walk 2 blocks without stopping; failure to complete all baseline assessment procedures; and baseline EMA compliance <80% (to be used as an indicator of behavioral adherence). To ensure safety, individuals with a history of heart disease, diabetes, cancer, or major orthopedic conditions will be required to provide physician consent prior to enrolling.

Participants will be recruited through online advertisements via social media outlets. Participants will be asked to respond to advertisements by completing an online screener linked to the advertisement. Participants will be recruited in 5 cohorts with a targeted cohort size of 35 - 60 participants. Individuals who meet all inclusion criteria and are eligible to participate in the study will be invited to attend a study orientation (delivered via Zoom), in which the study will be explained in greater detail. Individuals who attend orientation and remain interested in participating in the study will be emailed a link to complete informed consent via REDCap [53]. Data will be collected and managed using the REDCap tools hosted at Brown University Health's Department of Information Services. REDCap is a secure, web-based application designed to support data capture for research studies. Once consent is obtained, participants will begin the baseline assessment, which consists of study questionnaires, an EMA protocol, and wearing an accelerometer. Further, as part of the assessment, participants are required to attend a one-on-one behavioral interview via videoconference with a member of the research staff to ensure that the participant is willing to commit to all study procedures prior to beginning the WL program.

Internet-Based WL Program

All participants will receive a 12-month, fully automated IDWL program, which has been used in previous trials [7,54,55]. Prior to the start of the program, participants will be asked to watch a brief video, which will provide them with the basics for navigating the study website. The goal of the program is to induce a 1 - 2 lb (~0.5 - 1 kg) WL per week. To assist in

XSL•FO
RenderX

achieving this goal, individuals will be given an initial daily calorie intake goal of 1200 - 1800 kcal per day depending upon starting weight. Participants will also be taught to increase or decrease calorie intake as necessary based on WL progress and goals. There will be also a moderate-intensity aerobic exercise (eg, walking, cycling, and elliptical) goal, which progresses to 150 minutes per week. Participants will be instructed to self-monitor body weight, energy intake, and PA daily over the 12-month WL period, as self-monitoring in a behavioral WL program has been shown to be associated with weight change [56]. They will be provided with a smart scale, which will allow for wireless [57] transmission of weight data to the study website. Participants will be given the option to track their calories and exercise minutes in the Fitbit mobile app, in which case data will automatically transfer to the study website through an application programming interface built specifically for this study. However, if they use a different tracking method, they will be asked to manually input their daily calorie intake and exercise minute totals on the study website. Computer-generated, personalized feedback will be provided using these self-report data. Participants will also be instructed to view 10 - to 15-minute multimedia lessons, which are modeled after the Look AHEAD Trial [57]. Example lesson topics including strategies for increasing PA, eating out at restaurants, and problem-solving. Video lessons and automated feedback are provided weekly during the first 3 months of the IDWL program and monthly during months 4 - 12. Participants also have access to healthy recipes and tip sheets (eg, meal planning strategies, eating healthy on a budget, and exercising safely) via the study website. During months 4 - 12, automated text messages, which are informational and motivational in nature, will also be sent to participants twice per month and will differ between yoga and wellness conditions.

YOGA Intervention

Overview

Upon completion of the 3-month assessment visit, participants randomized to YOGA will receive a 9-month fully remote, Iyengar yoga intervention, which will include live, virtual classes (delivered via Zoom with real-time feedback provided) and self-guided, home-based practice using prerecorded videos. All live and prerecorded group classes will be led by a certified Iyengar yoga teacher, who will be trained using a standardized manual of procedure developed by the research team in collaboration with the yoga instructors. Iyengar yoga is a form of hatha yoga, which incorporates breathing, postural, and meditation practices and focuses on the precision in movement, maintaining proper alignment, and attention to subtle aspects of posture and breath. Iyengar yoga was chosen as the style of yoga for this study due to the precision in cueing, the use of props to modify postures for participants based on personal needs, and to minimize the risk of injury. Additionally, Iyengar is believed to have a lower energy expenditure than other more vigorous styles of yoga (eg, vinyasa or power yoga). Furthermore, the yoga curriculum was designed to ensure that it was inclusive to participants of any comfort level with yoga or existing assumptions about the practice. For example, there was no chanting of "OM" in the classes, and if yoga teachers used Sanskrit terms, they would always translate the names of the poses in English.

Introduction to Yoga Content

All participants will be mailed a yoga mat, strap, and 2 yoga blocks to facilitate home-based practice. At the start of the yoga program, participants will be required to complete an "Introduction to Yoga" curriculum (2 classes per week for 2 weeks) prior to engaging in the larger group classes. Each of these 60-minute introductory sessions will provide guidance on how to set up the home environment to foster a regular yoga practice, teach the proper form and alignment for foundational yoga poses, familiarize participants with the format of a yoga practice, and allow participants the opportunity to ask questions. Participants will practice yoga in front of their web camera so that instructors can provide individualized feedback to ensure that poses are performed properly and safely.

Live, Virtual Classes

Following the completion of the "Introduction to Yoga" curriculum, participants will advance to the larger group, 60-minute live classes. These classes will be delivered via Zoom for 36 weeks to guide participants through the yoga practice. Modifications will be provided for each pose to ensure that the level of difficulty is suitable for all individuals. Walls and chairs will be used for stability if needed. Group classes will be held multiple times per week. Participants will be instructed to attend at least 2 classes per week for the first 3 months (following the 2-week introduction, for a total of 14 weeks) and at least 1 class per week until the 12-month study assessment (22 weeks total). To increase translation potential, yoga classes will be open groups, meaning that participants may attend classes on days and times that best fit their needs, and participants from multiple cohorts will attend sessions together.

Self-Guided Yoga Practice

In addition to participating in the live, virtual classes, participants will also be given a self-guided yoga prescription to assist in the development of their personal practice. To facilitate self-guided practice, participants will be instructed to use video recordings from the study website, which will offer classes from the study's Iyengar teachers of varying durations (eg, 15, 30, 45, and 60 minutes). The first 3 months will focus more heavily on live yoga instruction, and the last 6 months will be designed to foster a more autonomous practice. The self-guided prescription is based upon total duration (ie, minutes per week) rather than frequency (ie, days per week) to facilitate greater retention and adherence by allowing individuals the choice to practice yoga less frequently for a longer duration (eg, 1 time per week for 60 minutes) or more frequently for shorter durations (eg, 3 times per week; one, 30-minute class and two, 15-minute classes). Progression toward a total dose (ie, combination of live classes and self-guided practice) of 180 minutes per week of yoga practice is the intervention target for both the first 3 months and the final 6 months of the program.

Verbal Cues for Life Application (Yamas or Niyamas)

Asanas (ie, the physical practice of yoga poses) taught within classes will be infused with verbal teachings of yamas or niyamas (ie, social disciplines to guide relationships with oneself

XSL•FO
RenderX

and others) to help cue participants into how yogic philosophies can overflow into everyday life (ie, how to apply skills practiced within class to personal experiences or struggles encountered outside of class). To ensure translatability of the yoga intervention, verbal cues will not explicitly discuss application to WL but will be broad enough so that participants can apply these principles to their weight control practices. Table 1 describes the 5 yamas and 5 niyamas, which simultaneously target several of the self-regulatory mechanisms examined within this study (eg, distress tolerance, mindfulness, and self-compassion). It is also important to note that many weight control behaviors can relate to several yamas and niyamas at once. For instance, choosing to eat and drink in moderation (vs overeating) so that one has more energy to engage in activities that one values can involve brahmacharya (moderation), aparigraha (nongreed), and satya (truth).

Table . Yamas and niyamas used in yoga.

Practice		Description
Yamas		
	Ahimsa	Nonviolence; making sure not to harm oneself or others; not thinking negative thoughts; kindness and consideration.
	Satya	Truthfulness; being truthful in one's thoughts, feelings, speech, and actions; being honest with oneself and others.
	Asteya	Nonstealing; not coveting; not being jealous; letting go of cravings. The need to "steal" often arises from a lack of faith in oneself.
	Brahmacharya	Moderation and channeling emotions; having moderation in our actions; conserving energy (slowing down); showing restraint.
	Aparigrapha	Nongreed; nonpossessiveness; greedlessness; simplicity; fulfilling needs rather than wants.
Niyamas		
	Saucha	Purification; cleanliness, internally and externally; purifying the body, mind, and surroundings.
	Santosha	Contentment; detaching from desires to cultivate inner peace.
	Tapas	Self-discipline; using self-discipline to burn out impurities of mind and body.
	Syadhyaha	Self-reflection; the practice of introspection; self-study.
	Ishvara Prandihana	Surrender; surrendering to a higher power.

Additional Yoga-Related Website Components

Instructions on how to properly and safely perform each yoga pose will be available on the study website, along with prerecorded videos for self-guided practice. To assist individuals with making the connection between weight control practices and what they are learning "on the yoga mat," a brief statement applying a yama or niyama to a WL behavior will be added to the text messages that are sent out twice per month as a part of the IDWL program. Participants will self-monitor their yoga practice and log the daily number of yoga practice minutes on the study website.

Recommended Yoga Practice Following the 9-Month Yoga Program

The yoga prescription is designed to facilitate an autonomous practice, which could be continued after the intervention. Upon completion of the 9-month yoga intervention, participants will be encouraged to continue their personal yoga practice of at least 180 minutes per week with the tools taught to them during the intervention. Although no intervention will be provided during months 13 - 18, participants will have continued access to the prerecorded videos, and participants will be asked to complete monthly surveys reporting on their yoga practice.

Contact Control: Health and Wellness Program

Upon completion of the 3-month assessment visit, participants randomized to WELLNESS will receive a 9-month fully remote, health and wellness program. Similar to YOGA, wellness classes will be approximately 60 minutes in duration and held multiple times per week. Participants will be instructed to attend 2 classes per week for the first 14 weeks and 1 class per week until the 12-month study assessment (22 weeks). Classes will be taught by individuals with a bachelor's degree or higher in a wellness-related field (eg, nutrition, exercise science, health promotion, and psychology). Classes will consist of lectures, handouts, videos, and discussions addressing a variety of health and lifestyle topics, which include immune health, sleep recommendations, health screenings, healthy aging, vitamins and minerals, and humor and health. WL will not be discussed

XSL•FO
RenderX

within these sessions, and participants will be informed that wellness classes are designed to function independently of the IDWL program. Overall, the WELLNESS condition provides a treatment program, which serves as a contact-matched control for the yoga arm.

Fidelity

All live intervention sessions (YOGA and WELLNESS) will be audio-recorded using a HIPAA-compliant version of Zoom. Recordings will be downloaded weekly and stored on a secure server. On a quarterly basis, 15% of all completed class recordings will be assessed for treatment fidelity via checklists developed by the research team for the YOGA and WELLNESS classes. These checklists include items such as class duration, whether the instructor taught the class as intended, and whether there was an opportunity for participants to ask questions. Fidelity will be assessed by a member of the research team who does not interact with participants. If at any time fidelity is <80%, teachers will be retrained and instructed on how to implement any necessary changes.

Outcome Measures

Overview

Assessments of weight, aerobic exercise minutes, dietary lapses, potential lapse triggers, and self-regulation mediator variables will be measured at baseline, 3 months (just prior to randomization), 6 months (mid-treatment), 12 months (postintervention), and 18 months (following a 6-month period of no intervention). Both questionnaires and EMA will be used. Research staff members will not be blinded to group assignment, and efforts will be made to collect data from all enrolled participants at the time of each assessment, including those who may no longer be attending intervention classes. Brief descriptions of study measures are provided below.

Anthropometric Measures

Weight change throughout the intervention period will be objectively assessed using Wi-Fi and Bluetooth-enabled smart scales (Withings Body), which will be mailed to study participants upon enrollment. At each assessment period, participants will be instructed to weigh first thing in the morning before eating or drinking, without clothing, or in light clothing. They will be asked to weigh 3 consecutive times, and if the maximum difference between any 2 weight measurements is ≤0.5 kg, the mean of the 3 weights will be used in the analysis. If the difference between any 2 weights is >0.5 kg, the weight that is farthest from the median will be dropped, and the mean of the remaining 2 weights will be used. Height will be self-reported, which has high concordance with stadiometer-assessed height measurements [58].

Ecological Momentary Assessment

Overview

Participants will complete brief surveys for 10 days at each assessment period. All entries will be made in response to 5 semirandom prompts delivered throughout the day via text message. Upon receiving a text message with a one-time link to the survey, participants will have 45 minutes to answer the questions. Each survey begins with the following message: “When completing this survey, please read the questions carefully as some may ask about how you feel ‘right now,’ others may ask what you may have experienced ‘since the last survey,’ and some questions may allow you to ‘select all response options that apply’ versus just one option. Please be as honest and accurate as possible. Your genuine input will help us gain valuable insights into your daily experiences.” Each survey will take 2 - 5 minutes to complete.

Dietary Lapses and Potential Lapse Triggers

At each survey, participants will be asked if they ate since the last EMA prompt. If eating was endorsed, a series of questions related to dietary lapse will be asked in addition to questions about potential triggers for lapses. These questions are included in Table 2. When eating is not reported, participants will receive a set of supplemental questions, which do not directly address study aims, to maintain consistency in survey structure and reduce response bias. This approach is consistent with previous work studying dietary lapse [39].

XSL•FO
RenderX

Table . Summary of ecological momentary assessment variables to be collected.

Variable	Measure	Response options
Dietary lapse		
Dietary lapse (type or total)[a]	Since your last survey (even if it was yesterday), at how many meals or snacks did any of the following occur? (a) Ate past the point of feeling full. (b) Ate a bigger portion than you meant to. (c) Had unplanned eating (ie, ate or drank when you do not usually eat or drink and were not making up for a missed meal). (d) Ate a food that you were intending to avoid or reduce. (e) Went over a calorie goal that you had in mind for yourself.	Number of lapse events
Dietary lapse (frequency)[b]	On the previous question, you indicated that since the last prompt, at least one of the following occurred at XX meals or snacks since the last survey (number from above question will be piped in). Now we want to know at how many meals or snacks did each of the following occur: (1) Ate past the point of feeling full. (2) Ate a bigger portion than you meant to. (3) Had unplanned eating (ie, ate or drank when you do not usually eat or drink and were not making up for a missed meal). (4) Ate a food that you were intending to avoid or reduce. (5) Went over a calorie goal that you had in mind for yourself.	Number of each type of lapse
Potential dietary lapse triggers		
Dietary temptation (type)[c]	Since the last survey (even if it was yesterday), have you experienced any of the following temptations to eat or drink, regardless of whether or not you actually acted on the temptation (check all that apply)?	(a) I was tempted to overeat, (b) I had a sudden urge to eat a tempting food or drink that I have been trying to reduce or avoid, (c) I was exposed to a tempting food or drink, (d) none of the above
Dietary temptation (details)[d]	If a, b, or c are selected from the question above, the following questions will be asked: (1) thinking about the most recent temptation you had: what was the intensity of the temptation? (2) Did you eat or drink the tempting food or drink?	(1) 1 - 7 Likert scale; 1=not at all intense, 7=extremely intense; (2) yes or no
Craving	How strong is your craving to eat a specific food right now?	Visual analog scale: 1=low intensity to 10=high intensity
Momentary affective experience (or related interoceptive state indicators)	(1) Right now, I feel calm, tense, upset, relaxed, content, worried, hungry, stressed, tired, bored, anxious, overwhelmed, ashamed, energetic; and (2) Right now, I feel very good, good, fairly good, neutral, fairly bad, bad, very bad	(1) 5 Likert scale; 1=not at all, 5=extremely; (2) −5 to 5 Likert scale; −5=very bad, +5=very good

[a]This will be the question used to calculate the total number of lapses.

[b]This will be the question used to determine the frequency of endorsed types of lapses, if applicable.

[c]This question will be used to assess the type of dietary temptations.

[d]These questions will be used to assess the intensity of and ability to resist dietary temptations.

Questionnaire Measures

At each assessment, a series of questionnaires will be completed via REDCap. At baseline only, demographic information and food insufficiency will be assessed. Questionnaire measures of self-regulatory mechanisms are found in Table 3 and will be completed at all assessment time points by all participants. At follow-up time points, program satisfaction questionnaires will also be completed, which will be used to assess the acceptability of the IDWL, yoga, and wellness programs. Yoga participants will also be asked to complete several additional questionnaires at follow-up related to their perception of the yoga classes, perceived benefits of practicing yoga [59], and frequency and duration of yoga practice.

XSL•FO
RenderX

Table . Validated questionnaires administered at baseline, 3-month, 6-month, and 18-month assessments to assess various aspects of self-regulation.

Questionnaire	Assessment	Scoring
Weight Control Strategies [60]	Self-monitoring, exercise, and diet as weight control strategies	0 - 90; comprised of a 30-item scale with four component solution for the subscales: (1) dietary choices, (2) self-monitoring strategies, (3) physical activity, and (4) psychological coping
Five Facet Mindfulness Questionnaire [61,62]	Mindfulness (including one's ability to observe, describe, and act with awareness, and be nonjudgmental and nonreactive to one's inner experience)	Comprised of 5 subscales: nonjudging (8-40), nonreactivity (7-35), observing (8-40), describing (8-40), and acting with awareness (8-40)
Distress Tolerance Scale [63]	Perceived capacity to tolerate distress	15 - 75; higher scores indicate greater distress tolerance
Self-Compassion Scale [64]	Degree to which individuals act compassionate toward themselves when experiencing a difficult time	26 - 130; higher scores reflect higher levels of self-compassion
Weight Efficacy Lifestyle Questionnaire [65]	Self-efficacy (one's confidence in their ability to avoid overeating when faced with certain situations)	20 - 100; higher scores indicate greater self-efficacy in managing weight-related behaviors
Behavioral Inhibition System (BIS) and Behavioral Activation System (BAS) [66]	Inhibitory control (measures motivation to avoid aversive outcomes)	BIS: 7 - 28; higher BIS scores indicate greater sensitivity to punishment and BAS: 4 - 16; higher BAS scores indicate greater sensitivity to reward
Barratt Impulsiveness Scale [67]	Inhibitory control: impulsiveness and includes cognitive, motor, and nonplanning impulsivity	30 - 150; higher scores reflect higher impulsivity

Aerobic PA

Minutes of aerobic PA will be objectively measured for 10 days, at each assessment period, using the previously validated Actigraph accelerometer [68]. Participants will be mailed an accelerometer, provided with instructions on how to wear the device, and be instructed to wear the Actigraph on their waist during all waking hours, except for water-related activities. They will also be asked to keep a log to record when they put the device on, took it off, and any exercise that they may have performed while not wearing the device. When processing the data, only "valid" days (ie, ≥8 hours of Actigraph wear time) will be considered in the analysis. Minutes per day spent in total and bout-related (≥10-minute bouts) moderate-to-vigorous intensity PA will be calculated using validated methods. One common critique of yoga as a WL intervention is that yoga participation may reduce aerobic PA. By objectively assessing PA, this study will be able to empirically answer this question. Upon study completion, at the time of analysis for this study, we will use all recommended processing methods for all accelerometry data and report findings accordingly.

Adherence Measures

Intervention adherence will include the percentage of IDWL lessons viewed and frequency of self-monitoring energy intake, weight, and aerobic exercise minutes entered into the study website (or transmitted to the study website via the Fitbit app) by participants over the study duration. Attendance at live yoga and wellness classes will also be assessed. Additional adherence metrics in the yoga condition include the number of times each prerecorded yoga video was viewed and weekly minutes per week of yoga practice (reported on the study website). During months 13 - 18, self-reported yoga practice will be assessed via monthly surveys.

Trial Monitoring, Safety, and Adverse Events

Upon study enrollment, participants will be instructed to contact the study staff should they experience any illnesses, injury, or side effects that are distressing to them throughout the 18-month research study. Moreover, each month, randomized participants will be required to complete a brief survey, which asks about any injuries or illnesses over the past month. Should a participant report an injury or illness on this survey, they will be contacted by the research staff to gather more information. Finally, at each assessment period (3, 6, 12, and 18 months), participants will be asked to report any illnesses, injuries, changes in health status, and medications experienced since the previous assessment. All adverse events (AEs; any untoward medical occurrence in a person that is temporarily associated with participation in the clinical study) will be reported by the principal investigator (JU) to the institutional review board, the trial's data safety monitoring board, and funding sponsor. Serious AEs (any AE that results in 1 or more of the following: death, life-threatening event, inpatient hospitalization, persistent or significant disability or incapacity, and important medical event based on appropriate medical judgment) will be reported within 24 hours. All other AEs will be reported twice yearly and will be included within regular reporting to the data safety monitoring board (consisting of external reviewers who have been approved by the study sponsor).

Data Analytic Plan and Sample Size Considerations

Sample Size Considerations

This study is fully powered for the primary aim, which is to examine treatment effects on WL from baseline through follow-up, taking into account the initial 3-month WL period. Power was calculated using a combination of G*Power and MPlus Monte Carlo estimation. Effect sizes were based on a

XSL•FO
RenderX

preliminary study assessing the effects of yoga for WL that showed within the yoga condition, there was significant WL from 3 to 6 months (Cohen *d*=0.45). Effect sizes from both yoga and contact-matched control arms were used to estimate a between-group effect at follow-up to use in the power calculations (difference in median percent WL of 3.5% between conditions among those with high initial WL; 95% CI 0.005-0.139). Assuming a similar effect size, 20% correlation between covariates, a total sample size as low as 146 would yield sufficient power (>80%) for testing between-group effects on WL from baseline. However, the final sample size was inflated to 210 to account for (1) the possibility that the effect size estimates that were obtained from our pilot work and used to determine the sample size were not generalizable, (2) adequately powered secondary aims (secondary outcomes and mediation), and (3) attrition rates similar to what we have seen in past studies by this research team (<20% of randomized sample at the end of treatment).

Analysis and Statistical Methods

Prior to analyses, data will be cleaned and explored, including outlier analysis and censoring of implausible values. Sociodemographics (eg, age, biological sex, and race or ethnicity), baseline weight, height, and BMI will be summarized across the aggregate sample and compared between groups using 2-tailed *t* tests for continuous variables, chi-square analyses for categorical variables, and nonparametric tests as appropriate. The distribution of each of the primary and secondary study outcomes will be assessed using both parametric and graphical methods and transformed as necessary (eg, log transform toward normality). Potential distributions for the outcome variables include normal and zero-inflated distributions. *P* values will be adjusted for multiple comparisons using the Hommel method. All analyses will be on the intent-to-treat sample (everyone randomized at 3 months will be included in the final analysis). Sensitivity analyses will explore effects of those who provide information during the postenrollment or prerandomization period that might influence outcomes (eg, initiation of WL medication).

Primary Aim

We will examine treatment effects on weight change from baseline to 18 months using a longitudinal mixed effects regression model. Absolute weight will be regressed on treatment assigned (YOGA vs WELLNESS), initial WL (from baseline to 3 months), and identified confounders. Model estimates and group comparisons will be generated for the 12-month primary end point, with contrasts specified to allow for understanding group differences in the postrandomization period (eg, 3 - 12 months controlling for baseline to 3 months). Models will include a participant-specific intercept and random slope terms to adjust for repeated measurements within participants over time. Further, we will test for clustering by cohort and adjust, if needed. Mixed models are flexible, in that they allow for a wide range of specifications of the time effect (linear and nonlinear), which is particularly important in WL data, where outcomes do not necessarily follow a linear pattern over time.

Secondary Aims

Using a similar analytic approach as that described for the primary aim, we will examine intervention effects on secondary outcomes, including frequency of dietary lapses and potential lapse triggers (eg, affect, cravings, and ability to resist dietary temptations), as assessed via EMA. Namely, longitudinal mixed effects models will be specified that regress outcomes over time on intervention group (YOGA vs WELLNESS), time, intervention*time, baseline value of the outcome, and confounders. Models will also adjust for the clustering of data within participants. Effects of day and measurement period will be tested, and adjustments made, as needed.

To evaluate the secondary aims of this study, we will examine potential mediators (eg, affect, mindfulness, self-compassion, and distress tolerance) of the treatment effect on primary (ie, weight change) and secondary outcomes (ie, dietary lapses) using a multiple mediator model with bootstrapped SEs (10,000 bootstrapped samples). Models estimate the path coefficients (a path: effects of YOGA vs WELLNESS on changes from baseline in potential mediators and b path: effects of changes from baseline in potential mediators on WL or lapse outcomes) as well as the indirect effect of intervention (ab path: YOGA vs WELLNESS on outcomes through the mediators).

Exploratory Aims

To examine longer-term treatment effects on weight change, we will generate model estimates and perform group comparisons for the 18-month end point using the mixed effects model of weight described earlier (aim 1). We will also examine moderators of the treatment effect on weight change from baseline to 12-month follow-up, including initial WL (weight change from baseline to 3 months), baseline BMI, race or ethnicity, and biological sex, using a similar analytic approach to that described in the primary aim.

Missing Data

Analyses will be on the intent-to-treat sample (everyone randomized will be included in the final analysis) under various assumptions about the missing data mechanism. Sensitivity to these assumptions will be tested. Specifically, we will gather follow-up information and reasons for dropout regardless of protocol completion and censor at the point of loss. We will compare the robustness of our findings using 3 statistical approaches for handling missing data. First, we will use a multiple imputation approach to impute missing outcomes. Next, we will use inverse probability weighting with propensity scores. This is a two-step method: (1) using logistic regression, the probability of missingness is modeled as a function of baseline covariates and baseline values of the outcome and (2) the inverse of the propensity scores (predicted probabilities of dropout from the first step) serves as weights in our regression model of the outcomes. Provided the data are missing at random or that the probability of missingness can be fully explained by observable data, this approach produces asymptotically unbiased estimates. To allow for the possibility that the missing at random assumption may not hold, we will also use a third approach, pattern mixture models, in which the distribution of the outcome

XSL•FO
RenderX

is assumed to follow a mixture of 2 distributions: one for those who complete follow-up and another for those who do not.

Results

This RCT was funded in July 2023 with an original project timeline of July 2023-June 2028. Enrollment began in February 2024 and is ongoing with the expectation of closing enrollment in June 2026. As of the submission of this protocol manuscript, recruitment timelines have been achieved, and 4 of 5 cohorts have launched. Upon study completion, results will be analyzed, and it is expected that the primary outcome findings will be submitted for publication within 6 months of the project end date.

Discussion

Principal Findings

This paper provides a summary of the rationale and methods of the PATH Trial. The primary goals of this randomized trial are to examine whether the addition of yoga to behavioral WL treatment improves long-term WL and to elucidate mechanisms through which yoga impacts weight in WL-seeking adults. These findings will have clinical relevance, as overweight and obesity are global health challenges linked to cardiometabolic risks, and despite effective WL interventions, long-term WL remains difficult [35,36]. Therefore, it is crucial to identify and test novel strategies, such as yoga, which have the potential to support sustained WL.

This study results from a synthesis of literature from a variety of disciplines using yoga interventions (cardiometabolic, pain, and addiction fields) and proposes a unique application of yoga to the area of weight management. While these other disciplines tend to focus on the psychological benefits of yoga for improving health, the weight control field has traditionally viewed yoga as a "mode" of exercise and has failed to consider that yoga is a mind-body intervention that could indirectly impact weight via psychological mechanisms (eg, improvements in self-regulation). Therefore, it is important to delineate that a yoga practice is not an exercise modality that produces effects similar to aerobic exercise, which has a higher caloric expenditure and is commonly recommended in WL programs, and therefore will be unlikely to elicit a comparable exercise-induced energy deficit [69]. Unlike past literature, this trial is uniquely designed and fully powered to evaluate self-regulatory pathways through which yoga impacts dietary lapses and WL, representing the first RCT to examine these mechanistic pathways within the context of WL treatment.

This design was chosen as the best method of answering the questions central to the study. For example, the use of an equal contact-time wellness program for the control arm, delivered through the same live, synchronous videoconferencing modality as yoga, provides a stringent test of the relative efficacy of yoga as a complementary therapy to IDWL. The decision to implement yoga at the start of week 14, rather than week 1, was made so that participants can dedicate their focus to initiating important WL skills and achieving program goals at the start of the program (eg, regular self-weighing, self-monitoring, calorie restriction, and aerobic exercise). Adding yoga at week 1 would further increase participant burden, which is already high at this time. Delaying yoga until week 14 allows additional time for these important WL skills to become more routine, which then allows individuals more time to dedicate to practicing yoga. Further, given that maximal WL in IDWL programs is achieved around 3 months [7,52] (earlier than what is observed with in-person treatment [35]), one way to bolster WL at this time would be to add an intervention component, which can both add novelty and strengthen self-regulation. Research supports this approach, as individuals report that over time, boredom increases; it requires more effort to stay on track, and the cost-benefit ratio for WL increases (ie, the effort required to adhere to WL practices is greater than the positive feedback received for adherence) [35,70-74]. This makes continued diligence to self-regulatory WL behaviors challenging, particularly during times of high stress, emotional difficulties, or when exposed to highly palatable foods. Our pilot study, which also implemented yoga following a period of WL treatment, provided further support for this decision [51].

This study is novel, in that it is one of a handful of studies to combine yoga with a WL program, and it is the first to empirically examine this combination of approaches using fully remote intervention procedures (ie, IDWL combined with live, online yoga classes). We made the decision to combine yoga with IDWL treatment from months 4 to 12 (vs stopping IDWL treatment) for the following reasons. First, most WL programs are 1 year in duration, and there is no added cost to continuing the IDWL program, given that it is fully automated. Additionally, as prior research suggests that yoga reinforces self-regulatory skills taught within IDWL, extending the IDWL treatment would provide an opportunity for participants to "practice" these skills. Finally, this approach allows for examination of the combined effect of yoga and IDWL.

Another novel aspect of this study is the use of technology-based assessment methods (EMA) for capturing participants' dietary lapses and potential lapse triggers (eg, cravings, temptations, and affect) in near-real-time, within a naturalistic setting, while participating in a WL and yoga intervention. We also pair these survey assessment methods with objective monitoring of aerobic PA, so that they occur simultaneously. While pilot work indicates that yoga participation does not reduce aerobic exercise engagement [4], this is a frequent concern regarding the inclusion of yoga into a WL program. Thus, this study will be able to empirically answer whether including yoga reduces engagement in other modalities of exercise while also allowing for novel exploratory analyses related to both the acute and chronic effects of yoga on psychological and dietary variables.

This study will also add to the literature in its ability to examine for whom yoga is most efficacious for producing WL. Examination of yoga within the context of a weight management program is in its infancy, and studies assessing the effect of yoga on body weight have been primarily in the normal weight, non-Hispanic, White female population. This study seeks to enroll ≥35% male individuals and ≥40% individuals from racial and ethnic groups that are commonly underrepresented in WL trials. Thus, we will be in a unique position to explore whether the effect of yoga on WL differs by biological sex and race or

XSL•FO
RenderX

ethnicity or whether there are other significant moderators (eg, BMI, 3-month WL, and socioeconomic status) of the effects of yoga on WL. To our knowledge, this is the first study to examine whether there are specific subgroups of WL-seeking individuals who benefit most from yoga participation.

Conclusions

The proposed study is significant because it is one of the few trials to examine the combined effect of yoga and WL treatment among individuals with overweight or obesity and the first fully powered trial to assess the combination of yoga and WL beyond 6 months. While the yoga field is often criticized for a lack of scientific rigor, this study represents a significant step toward improving the quality of research in this area, as it is empirically guided and thoughtfully designed to ensure rigor and reproducibility of study findings (eg, randomized trial with contact-matched control, manualized protocols, and fidelity checks). This study also uses and tests an easily disseminated intervention model, which combines a fully automated, low-cost IDWL program with live, online yoga classes; thus, translation potential is high. Examination of mediators and moderators of the effect of yoga on WL is also important for improving theoretical models of yoga and self-regulation, optimizing future intervention trials, and enhancing our understanding of how and for whom yoga impacts long-term WL. Ultimately, if primary hypotheses are confirmed, yoga could be considered a valuable intervention component for improving IDWL treatment. In the future, if favorable results are observed, longer-term follow-up would also be warranted to better understand the role of yoga in WL maintenance over periods of time longer than this study is planning. This study has great potential to make a significant contribution to both yoga and WL fields and addresses critical gaps in the current literature.

Funding

This trial is sponsored by the National Center for Complementary & Integrative Health of the National Institutes of Health (award R01AT011868). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

JU, BB, JGT, SPG, and SS were involved in the study conception and design. SD was responsible for the development of the statistical analysis plan. SYT and JU drafted this manuscript, and SS, BB, SD, JGT, and SPG provided critical revision and final approval of the manuscript.

Conflicts of Interest

JU and JGT serve on the Scientific Advisory Board of Medifast, Inc. No other authors have any potential conflicts to report.

Checklist 1
SPIRIT checklist.
[DOCX File, 35 KB - resprot_v15i1e78973_app1.docx]

References

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017-2018. NCHS Data Brief 2020 Feb;2020(360):1-8. [Medline: 32487284]
2. Fowler G. In 2020, we reached peak Internet. Here's what worked—and what flopped. The Keene Sentinel. URL: https://www.keenesentinel.com/in-2020-we-reached-peak-internet-heres-what-worked-and-what-flopped/article_5853c3d2-f829-590a-a048-ffc347a79ae4.html [accessed 2026-01-28]
3. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017 Mar 24;19(3):e90. [doi: 10.2196/jmir.7111] [Medline: 28341617]
4. Unick JL, Dunsiger SI, Bock BC, et al. A randomized trial examining the effect of yoga on dietary lapses and lapse triggers following behavioral weight loss treatment. Obes Sci Pract 2023 Oct;9(5):484-492. [doi: 10.1002/osp4.678]
5. Lufkin B. The evolution of home fitness. British Broadcasting Corporation. URL: https://www.bbc.com/worklife/article/20200504-covid-19-update-quarantine-home-workouts-during-coronavirus [accessed 2026-01-28]
6. Hutchesson MJ, Rollo ME, Krukowski R, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. Obes Rev 2015 May;16(5):376-392. [doi: 10.1111/obr.12268] [Medline: 25753009]

XSL•FO
RenderX

7. Leahey TM, Thomas G, Fava JL, et al. Adding evidence-based behavioral weight loss strategies to a statewide wellness campaign: a randomized clinical trial. Am J Public Health 2014 Jul;104(7):1300-1306. [doi: 10.2105/AJPH.2014.301870] [Medline: 24832424]
8. Aranya SH. Yoga Philosophy of Patanjali with Bhasvati: Univsity of Calcutta; 2000.
9. Van Puymbroeck M, Payne LL, Hsieh PC. A phase I feasibility study of yoga on the physical health and coping of informal caregivers. Evid Based Complement Alternat Med 2007 Dec;4(4):519-529. [doi: 10.1093/ecam/nem075] [Medline: 18227920]
10. Satyapriya M, Nagendra HR, Nagarathna R, Padmalatha V. Effect of integrated yoga on stress and heart rate variability in pregnant women. Int J Gynaecol Obstet 2009 Mar;104(3):218-222. [doi: 10.1016/j.ijgo.2008.11.013] [Medline: 19110245]
11. West J, Otte C, Geher K, Johnson J, Mohr DC. Effects of Hatha yoga and African dance on perceived stress, affect, and salivary cortisol. Ann Behav Med 2004 Oct;28(2):114-118. [doi: 10.1207/s15324796abm2802_6] [Medline: 15454358]
12. Chong CSM, Tsunaka M, Tsang HWH, Chan EP, Cheung WM. Effects of yoga on stress management in healthy adults: a systematic review. Altern Ther Health Med 2011;17(1):32-38. [Medline: 21614942]
13. Sharma M. Yoga as an alternative and complementary approach for stress management: a systematic review. J Evid Based Complementary Altern Med 2014 Jan;19(1):59-67. [doi: 10.1177/2156587213503344] [Medline: 24647380]
14. Ross A, Thomas S. The health benefits of yoga and exercise: a review of comparison studies. J Altern Complement Med 2010 Jan;16(1):3-12. [doi: 10.1089/acm.2009.0044] [Medline: 20105062]
15. Büssing A, Michalsen A, Khalsa SBS, Telles S, Sherman KJ. Effects of yoga on mental and physical health: a short summary of reviews. Evid Based Complement Alternat Med 2012;2012:165410. [doi: 10.1155/2012/165410] [Medline: 23008738]
16. Woodyard C. Exploring the therapeutic effects of yoga and its ability to increase quality of life. Int J Yoga 2011 Jul;4(2):49-54. [doi: 10.4103/0973-6131.85485] [Medline: 22022122]
17. Lee SW, Mancuso CA, Charlson ME. Prospective study of new participants in a community-based mind-body training program. J Gen Intern Med 2004 Jul;19(7):760-765. [doi: 10.1111/j.1525-1497.2004.30011.x] [Medline: 15209590]
18. Curtis K, Osadchuk A, Katz J. An eight-week yoga intervention is associated with improvements in pain, psychological functioning and mindfulness, and changes in cortisol levels in women with fibromyalgia. J Pain Res 2011;4:189-201. [doi: 10.2147/JPR.S22761] [Medline: 21887116]
19. Shelov DV, Suchday S, Friedberg JP. A pilot study measuring the impact of yoga on the trait of mindfulness. Behav Cogn Psychother 2009 Oct;37(5):595-598. [doi: 10.1017/S1352465809990361] [Medline: 19751539]
20. Baird SO, Hopkins LB, Medina JL, Rosenfield D, Powers MB, Smits JAJ. Distress tolerance as a predictor of adherence to a yoga intervention: moderating roles of BMI and body image. Behav Modif 2016 Jan;40(1-2):199-217. [doi: 10.1177/0145445515612401] [Medline: 26530475]
21. Medina J, Hopkins L, Powers M, Baird SO, Smits J. The effects of a hatha yoga intervention on facets of distress tolerance. Cogn Behav Ther 2015;44(4):288-300. [doi: 10.1080/16506073.2015.1028433] [Medline: 25952547]
22. Dick AM, Niles BL, Street AE, DiMartino DM, Mitchell KS. Examining mechanisms of change in a yoga intervention for women: the influence of mindfulness, psychological flexibility, and emotion regulation on PTSD symptoms. J Clin Psychol 2014 Dec;70(12):1170-1182. [doi: 10.1002/jclp.22104] [Medline: 24888209]
23. Gootjes L, Franken I. Cognitive emotion regulation in yogic meditative practitioners: sustained modulation of electrical brain potentials. J Psychophysiol 2011;25:87-94. [doi: 10.1027/0269-8803/a000043]
24. Braun TD, Park CL, Conboy LA. Psychological well-being, health behaviors, and weight loss among participants in a residential, Kripalu yoga-based weight loss program. Int J Yoga Therap 2012;2012(22):9-22. [Medline: 23070668]
25. Braun TD, Park CL, Gorin AA, Garivaltis H, Noggle JJ, Conboy LA. Group-based yogic weight loss with ayurveda-inspired components: a pilot investigation of female yoga practitioners and novices. Int J Yoga Therap 2016 Jan;26(1):55-72. [doi: 10.17761/1531-2054-26.1.55] [Medline: 27797660]
26. Jiang D, Liu Z, Sun G. The effect of yoga meditation practice on young adults' inhibitory control: an fNIRS study. Front Hum Neurosci 2021;15:725233. [doi: 10.3389/fnhum.2021.725233] [Medline: 34588968]
27. Purohit SP, Pradhan B. Effect of yoga program on executive functions of adolescents dwelling in an orphan home: a randomized controlled study. J Tradit Complement Med 2017 Jan;7(1):99-105. [doi: 10.1016/j.jtcme.2016.03.001] [Medline: 28053894]
28. Bilderbeck AC, Farias M, Brazil IA, Jakobowitz S, Wikholm C. Participation in a 10-week course of yoga improves behavioural control and decreases psychological distress in a prison population. J Psychiatr Res 2013 Oct;47(10):1438-1445. [doi: 10.1016/j.jpsychires.2013.06.014] [Medline: 23866738]
29. Hewett ZL, Pumpa KL, Smith CA, Fahey PP, Cheema BS. Effect of a 16-week Bikram yoga program on perceived stress, self-efficacy and health-related quality of life in stressed and sedentary adults: a randomised controlled trial. J Sci Med Sport 2018 Apr;21(4):352-357. [doi: 10.1016/j.jsams.2017.08.006] [Medline: 28866110]
30. Elfhag K, Rössner S. Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. Obes Rev 2005 Feb;6(1):67-85. [doi: 10.1111/j.1467-789X.2005.00170.x] [Medline: 15655039]
31. Grilo CM, Shiffman S, Wing RR. Relapse crises and coping among dieters. J Consult Clin Psychol 1989 Aug;57(4):488-495. [doi: 10.1037//0022-006x.57.4.488] [Medline: 2768608]

XSL•FO
RenderX

32. Westenhoefer J, von Falck B, Stellfeldt A, Fintelmann S. Behavioural correlates of successful weight reduction over 3 y. Results from the Lean Habits Study. Int J Obes Relat Metab Disord 2004 Feb;28(2):334-335. [doi: 10.1038/sj.ijo.0802530] [Medline: 14647175]
33. Bonato DP, Boland FJ. Predictors of weight loss at the end of treatment and 1-year follow-up for a behavioral weight loss program. Int J Eat Disord 1987 Jul;6(4):573-577. [doi: 10.1002/1098-108X(198707)6:4<573::AID-EAT2260060417>3.0.CO;2-L]
34. Maclean PS, Bergouignan A, Cornier MA, Jackman MR. Biology's response to dieting: the impetus for weight regain. Am J Physiol Regul Integr Comp Physiol 2011 Sep;301(3):R581-R600. [doi: 10.1152/ajpregu.00755.2010] [Medline: 21677272]
35. MacLean PS, Wing RR, Davidson T, et al. NIH working group report: innovative research to improve maintenance of weight loss. Obesity (Silver Spring) 2015 Jan;23(1):7-15. [doi: 10.1002/oby.20967] [Medline: 25469998]
36. Wing RR, Phelan S. Long-term weight loss maintenance. Am J Clin Nutr 2005 Jul;82(1 Suppl):222S-225S. [doi: 10.1093/ajcn/82.1.222S] [Medline: 16002825]
37. Danhauer SC, Mihalko SL, Russell GB, et al. Restorative yoga for women with breast cancer: findings from a randomized pilot study. Psychooncology 2009 Apr;18(4):360-368. [doi: 10.1002/pon.1503] [Medline: 19242916]
38. Vadiraja HS, Rao MR, Nagarathna R, et al. Effects of yoga program on quality of life and affect in early breast cancer patients undergoing adjuvant radiotherapy: a randomized controlled trial. Complement Ther Med 2009;17(5-6):274-280. [doi: 10.1016/j.ctim.2009.06.004] [Medline: 19942107]
39. Forman EM, Schumacher LM, Crosby R, et al. Ecological momentary assessment of dietary lapses across behavioral weight loss treatment: characteristics, predictors, and relationships with weight change. Ann Behav Med 2017 Oct;51(5):741-753. [doi: 10.1007/s12160-017-9897-x] [Medline: 28281136]
40. Carels RA, Douglass OM, Cacciapaglia HM, O'Brien WH. An ecological momentary assessment of relapse crises in dieting. J Consult Clin Psychol 2004 Apr;72(2):341-348. [doi: 10.1037/0022-006X.72.2.341] [Medline: 15065966]
41. Desveaux L, Lee A, Goldstein R, Brooks D. Yoga in the management of chronic disease: a systematic review and meta-analysis. Med Care 2015 Jul;53(7):653-661. [doi: 10.1097/MLR.0000000000000372] [Medline: 26035042]
42. Cramer H, Lauche R, Langhorst J, Dobos G. Yoga for depression: a systematic review and meta-analysis. Depress Anxiety 2013 Nov;30(11):1068-1083. [doi: 10.1002/da.22166] [Medline: 23922209]
43. Uebelacker LA, Epstein-Lubow G, Gaudiano BA, Tremont G, Battle CL, Miller IW. Hatha yoga for depression: critical review of the evidence for efficacy, plausible mechanisms of action, and directions for future research. J Psychiatr Pract 2010 Jan;16(1):22-33. [doi: 10.1097/01.pra.0000367775.88388.96] [Medline: 20098228]
44. Sherman KJ, Cherkin DC, Wellman RD, et al. A randomized trial comparing yoga, stretching, and a self-care book for chronic low back pain. Arch Intern Med 2011 Dec 12;171(22):2019-2026. [doi: 10.1001/archinternmed.2011.524] [Medline: 22025101]
45. Carei TR, Fyfe-Johnson AL, Breuner CC, Brown MA. Randomized controlled clinical trial of yoga in the treatment of eating disorders. J Adolesc Health 2010 Apr;46(4):346-351. [doi: 10.1016/j.jadohealth.2009.08.007] [Medline: 20307823]
46. McIver S, McGartland M, O'Halloran P. "Overeating is not about the food": women describe their experience of a yoga treatment program for binge eating. Qual Health Res 2009 Sep;19(9):1234-1245. [doi: 10.1177/1049732309343954] [Medline: 19690205]
47. McIver S. Yoga as a Treatment for Binge Eating: La Trobe University; 2006.
48. Caldwell AE, Purcell SA, Gray B, Smieja H, Catenacci VA. The impact of yoga on components of energy balance in adults with overweight or obesity: a systematic review. Obes Sci Pract 2022 Apr;8(2):219-232. [doi: 10.1002/osp4.552] [Medline: 35388342]
49. Jakicic JM, Davis KK, Rogers RJ, et al. Feasibility of integration of yoga in a behavioral weight-loss intervention: a randomized trial. Obesity (Silver Spring) 2021 Mar;29(3):512-520. [doi: 10.1002/oby.23089] [Medline: 33528905]
50. Yadav R, Yadav RK, Khadgawat R, Pandey RM, Upadhyay AD, Mehta N. Randomized controlled trial of a 12-week yoga-based (including diet) lifestyle vs. dietary intervention on cardio-metabolic risk factors and continuous risk score in Indian adults with metabolic syndrome. Behav Med 2020;46(1):9-20. [doi: 10.1080/08964289.2018.1538098] [Medline: 30615583]
51. Unick JL, Dunsiger SI, Bock BC, Sherman SA, Braun TD, Wing RR. A preliminary investigation of yoga as an intervention approach for improving long-term weight loss: a randomized trial. PLoS ONE 2022;17(2):e0263405. [doi: 10.1371/journal.pone.0263405] [Medline: 35120162]
52. Ross KM, Qiu P, You L, Wing RR. Characterizing the pattern of weight loss and regain in adults enrolled in a 12-week internet-based weight management program. Obesity (Silver Spring) 2018 Feb;26(2):318-323. [doi: 10.1002/oby.22083] [Medline: 29239141]
53. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
54. Unick JL, Leahey T, Kent K, Wing RR. Examination of whether early weight loss predicts 1-year weight loss among those enrolled in an Internet-based weight loss program. Int J Obes (Lond) 2015 Oct;39(10):1558-1560. [doi: 10.1038/ijo.2015.89] [Medline: 25982792]

XSL•FO
RenderX

55. Unick JL, Dorfman L, Leahey TM, Wing RR. A preliminary investigation into whether early intervention can improve weight loss among those initially non-responsive to an internet-based behavioral program. J Behav Med 2016 Apr;39(2):254-261. [doi: 10.1007/s10865-015-9691-9] [Medline: 26518207]
56. Goldstein SP, Goldstein CM, Bond DS, Raynor HA, Wing RR, Thomas JG. Associations between self-monitoring and weight change in behavioral weight loss interventions. Health Psychol 2019 Dec;38(12):1128-1136. [doi: 10.1037/hea0000800] [Medline: 31556659]
57. Look AHEAD Research Group, Wadden TA, West DS, et al. The Look AHEAD study: a description of the lifestyle intervention and the evidence supporting it. Obesity (Silver Spring) 2006 May;14(5):737-752. [doi: 10.1038/oby.2006.84] [Medline: 16855180]
58. Quick V, Byrd-Bredbenner C, Shoff S, et al. Concordance of self-report and measured height and weight of college students. J Nutr Educ Behav 2015;47(1):94-98. [doi: 10.1016/j.jneb.2014.08.012] [Medline: 25316656]
59. Park CL, Elwy AR, Maiya M, et al. The Essential Properties of Yoga Questionnaire (EPYQ): psychometric properties. Int J Yoga Therap 2018 Nov;28(1):23-38. [doi: 10.17761/2018-00016R2] [Medline: 29498893]
60. Pinto AM, Fava JL, Raynor HA, LaRose JG, Wing RR. Development and validation of the weight control strategies scale. Obesity (Silver Spring) 2013 Dec;21(12):2429-2436. [doi: 10.1002/oby.20368] [Medline: 23512914]
61. Baer RA, Smith GT, Lykins E, et al. Construct validity of the five facet mindfulness questionnaire in meditating and nonmeditating samples. Assessment 2008 Sep;15(3):329-342. [doi: 10.1177/1073191107313003] [Medline: 18310597]
62. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
63. Brown RJ, Burton AL, Abbott MJ. The relationship between distress tolerance and symptoms of depression: validation of the Distress Tolerance Scale (DTS) and short-form (DTS-SF). J Clin Psychol 2022 Dec;78(12):2609-2630. [doi: 10.1002/jclp.23370] [Medline: 35510827]
64. Raes F, Pommier E, Neff KD, Van Gucht D. Construction and factorial validation of a short form of the Self-Compassion Scale. Clin Psychol Psychother 2011;18(3):250-255. [doi: 10.1002/cpp.702] [Medline: 21584907]
65. Clark MM, Abrams DB, Niaura RS, Eaton CA, Rossi JS. Self-efficacy in weight management. J Consult Clin Psychol 1991 Oct;59(5):739-744. [doi: 10.1037//0022-006x.59.5.739] [Medline: 1955608]
66. Carver CS, White TL. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS Scales. J Pers Soc Psychol 1994;67(2):319-333. [doi: 10.1037/0022-3514.67.2.319]
67. Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt Impulsiveness Scale. J Clin Psychol 1995 Nov;51(6):768-774. [doi: 10.1002/1097-4679(199511)51:6<768::aid-jclp2270510607>3.0.co;2-1] [Medline: 8778124]
68. Aadland E, Ylvisåker E. Reliability of the Actigraph GT3X+ accelerometer in adults under free-living conditions. PLoS ONE 2015;10(8):e0134606. [doi: 10.1371/journal.pone.0134606] [Medline: 26274586]
69. Donnelly JE. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. Med Sci Sports Exerc 2009;41(7):1532. [doi: 10.1249/MSS.0b013e3181ae46a1]
70. Jeffery RW, Levy RL, Langer SL, et al. A comparison of maintenance-tailored therapy (MTT) and standard behavior therapy (SBT) for the treatment of obesity. Prev Med 2009 Nov;49(5):384-389. [doi: 10.1016/j.ypmed.2009.08.004]
71. Befort CA, Stewart EE, Smith BK, Gibson CA, Sullivan DK, Donnelly JE. Weight maintenance, behaviors and barriers among previous participants of a university-based weight control program. Int J Obes 2008 Mar;32(3):519-526. [doi: 10.1038/sj.ijo.0803769]
72. Davis EM, Clark JM, Carrese JA, Gary TL, Cooper LA. Racial and socioeconomic differences in the weight-loss experiences of obese women. Am J Public Health 2005 Sep;95(9):1539-1543. [doi: 10.2105/AJPH.2004.047050] [Medline: 16118365]
73. Davis R, Cook D, Cohen L. A community resilience approach to reducing ethnic and racial disparities in health. Am J Public Health 2005 Dec;95(12):2168-2173. [doi: 10.2105/AJPH.2004.050146] [Medline: 16304131]
74. Davis SK, Liu Y, Quarells RC, Din-Dzietharn R, Metro Atlanta Heart Disease Study Group. Stress-related racial discrimination and hypertension likelihood in a population-based sample of African Americans: the Metro Atlanta Heart Disease Study. Ethn Dis 2005;15(4):585-593. [Medline: 16259480]

Abbreviations

AE: adverse event
EMA: ecological momentary assessment
HIPAA: Health Insurance Portability and Accountability Act
IDWL: internet-delivered weight loss
PA: physical activity
PATH: Pathway to Achieving Total Health
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TiDieR: Template for Intervention Description and Replication

XSL•FO
RenderX

WL: weight loss

Edited by J Sarvestan; submitted 18.Jun.2025; peer-reviewed by A Martinko, D Guerrero-Magaña; accepted 02.Dec.2025; published 25.Feb.2026.

Please cite as:
Tobin SY, Sherman S, Bock B, Dunsiger S, Thomas JG, Goldstein SP, Unick J
PATH Trial for Examining Yoga as a Strategy to Improve Remote-Based Weight Loss in Adults: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e78973
URL: https://www.researchprotocols.org/2026/1/e78973
doi:10.2196/78973

© Selene Y Tobin, Sally Sherman, Beth Bock, Shira Dunsinger, J Graham Thomas, Stephanie P Goldstein, Jessica Unick. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparative Evaluation of AmnioGuard and Advanced Platelet-Rich Fibrin When Combined With NovaBone Putty in the Regeneration of Human Periodontal Infrabony Defects: Protocol for a Randomized Controlled Clinical Trial

Geeta Bhandari, BDS; Priyanka Jaiswal, MDS; Shweta Bhagat, BDS; Sakshi Kotecha, BDS

Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research (DMIHER(DU)), Sawangi, Wardha, India

Corresponding Author:
Geeta Bhandari, BDS
Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research (DMIHER(DU)), Sawangi, Wardha, India

Abstract

Background: Periodontitis causes progressive attachment loss, which in turn causes tooth mobility and, ultimately, tooth loss. Regenerating and rebuilding the lost periodontal tissue is the goal of periodontal regenerative surgery. Regeneration using AmnioGuard, advanced platelet-rich fibrin (A-PRF), and NovaBone Putty has been demonstrated to be successful in increasing clinical attachment level and decreasing probing pocket depth.

Objective: This study aims to compare and evaluate the effectiveness of AmnioGuard and A-PRF membrane in combination with NovaBone Putty in the regeneration of human periodontal infrabony defects.

Methods: A total of 24 patients will be selected and divided into 2 groups: group A (n=12, 50%) and group B (n=12, 50%). Group A will receive treatment with NovaBone Putty and AmnioGuard, whereas group B will receive treatment with autologous A-PRF and NovaBone Putty. Plaque index, clinical attachment level, probing pocket depth, complete mouth bleeding score, and intraoral periapical radiographs will be statistically analyzed. Clinical and radiographic features will be evaluated before surgery and 6 months following the procedure.

Results: Recruitment began on September 16, 2025, and is ongoing. As of manuscript submission, 4 participants have been enrolled. Results are expected to be published in 2027. All the clinical measurements will be recorded after 6 months. It can be anticipated that AmnioGuard membrane with NovaBone Putty will provide additional benefits in the management of chronic periodontitis by promoting better healing and clinical outcomes.

Conclusions: Within the constraints of this study, the AmnioGuard membrane combined with NovaBone Putty may be more effective in the treatment of chronic periodontitis.

Trial Registration: Clinical Trials Registry - India CTRI/2025/04/084507; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=MTI3MjYw&Enc=&userName=

International Registered Report Identifier (IRRID): DERR1-10.2196/75987

(JMIR Res Protoc 2026;15:e75987) doi:10.2196/75987

KEYWORDS

periodontal regeneration; bone graft biomaterials; guided tissue regeneration; platelet-derived growth factors; clinical periodontology; osteoconductive materials; wound healing modulation; leukocyte-rich fibrin; probing depth reduction; attachment level gain

Introduction

Background

A dysbiotic dental biofilm that gradually destroys bone and periodontal attachment is a hallmark of periodontitis, a chronic, multifactorial, inflammatory illness. This condition negatively affects speech, mastication, aesthetics, psychological well-being, and overall quality of life [1]. On the basis of bone resorption patterns, periodontal defects are traditionally classified as supraosseous (suprabony) or infraosseous (infrabony). Goldman and Cohen [2] further categorized infrabony defects by the number and configuration of residual osseous walls surrounding the periodontal pocket. Infrabony defects pose a significant clinical challenge as they require both structural regeneration and the restoration of periodontal function. Regeneration is defined as “the reproduction or reconstitution of a lost or injured part to restore the architecture and function of the periodontium”

XSL•FO
RenderX

[3]. To achieve this, regenerative periodontal therapy aims to stimulate the formation of new bone, periodontal ligament, and cementum. NovaBone Putty, a next-generation calcium phosphosilicate material derived from bioactive glass, offers multiple advantages for regeneration. It is osteoconductive, osteostimulative, and biocompatible, with hemostatic and anti-inflammatory effects. Its premixed formulation enhances clinical handling and allows for optimal defect adaptation [4]. AmnioGuard, a human amniotic membrane (AM)–based barrier, has also gained attention due to its anti-inflammatory, antifibrotic, antibacterial, and antiscarring properties. Its high laminin subunit alpha-5 content supports epithelial adhesion and accelerates healing, whereas its role as a biological barrier aids in preventing epithelial downgrowth—an essential aspect of guided tissue regeneration (GTR) [5-7]. In recent years, biological membranes have become increasingly important in regenerative therapies. Since their initial application in skin grafting in 1910, their role has expanded across medical and dental fields. In particular, AmnioGuard has shown potential to reduce microbial penetration and modulate wound healing responses [8]. In addition, advanced platelet-rich fibrin (A-PRF), rich in leukocytes, contributes to regeneration by releasing growth factors such as transforming growth factor beta 1, epidermal growth factor, and insulin-like growth factor, all of which enhance fibroblast proliferation and tissue healing [9]. Although several studies have assessed the individual regenerative potential of NovaBone Putty and AmnioGuard, limited clinical data directly compare their effectiveness in managing periodontal infrabony defects. A study conducted by Hazari et al [10] compared NovaBone Putty and platelet-rich fibrin (PRF) and concluded that, in comparison to NovaBone Putty alone, PRF in conjunction with NovaBone Putty produced more favorable findings in terms of relative attachment level gain and higher reduction in probing pocket depth (PPD). A study by Kothiwale et al [11] in which they assessed and compared the effectiveness of bovine-derived xenogenic bone graft (Bio-Oss) and demineralized freeze-dried bone allograft using AM as GTR determined that, while there was no discernible difference between the two materials, both showed a notable improvement in bone fill and percentage gain. Furthermore, evidence regarding clinical outcomes such as probing depth reduction, attachment level gain, and bone fill remains sparse [12]. Therefore, this study aims to address this gap by directly comparing the clinical and radiographic outcomes of NovaBone Putty in combination with A-PRF or AmnioGuard in the treatment of infrabony periodontal defects.

Objectives

This study aims to evaluate and compare the effectiveness of AmnioGuard and A-PRF in combination with NovaBone Putty in terms of gain in clinical attachment level (CAL), reduction in PPD, and radiographic bone fill in infrabony defects before surgery and 6 months after surgery.

Methods

Ethical Considerations

The Institutional Ethics Committee of Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India, approved this study (reference DMIHER(DU)/IEC/2025/507). Participants will be provided with an informed consent form clearly describing the study's aim, procedures, risks, and benefits and their right to opt out at any time. Consent for publication of data will be obtained and used for all analyses, including secondary analyses. We will confirm that the original consent and institutional review board approval explicitly cover the use of primary and secondary data. Collection of data and entering of the data into the database for screening and randomization will be conducted by the primary investigator. The privacy of prospective and enrolled participants' personal information will be upheld to ensure confidentiality before, during, and after the trial. We will ensure that the data collected are fully anonymized, with no personally identifiable information retained or linked to any participant. As there are no interventions involved and only a cross-sectional survey will be conducted in which the data will be collected via an interview and anthropometric measurements, participants will not be provided with any compensation in this study. For any images presented in the manuscript, appropriate consent will be obtained from the participants.

Recruitment

Overview

Participants in this clinical trial will be those receiving treatment at the Department of Periodontics and Implantology of Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, or those referred for treatment there who have periodontal pockets larger than or equal to 4 mm in at least 10 permanent teeth. All clinical procedures will be performed there. It is thought that this will enhance trial adherence and recruitment because patients will already be receiving treatment at this location.

Inclusion Criteria

The inclusion criteria are as follows: (1) at least 1 or 2 interproximal infrabony osseous defects that may be seen on radiography, with a PPD of at least 5 mm and clinical attachment loss of at least 5 mm after starting treatment; (2) clinical and radiographic measurements showing that the intraosseous component of the lesion is ≥3 mm deep (intraoperative assessment will confirm this); (3) defect radiographic base of at least 3 mm coronal to the tooth's apex; and (4) keratinized gingiva surrounding the test teeth of at least 3 mm wide to completely hide the deficiency with soft tissue.

Exclusion Criteria

The exclusion criteria are as follows: (1) patients with evident localized aggressive periodontitis, (2) individuals with poor oral hygiene (plaque index >1), (3) those who have recently smoked more than 10 cigarettes per day or who have used tobacco products, (4) a tooth that has not had enough endodontic or restorative care, (5) a tooth that has a class III or class IV furcation defect and mobility greater than grade II, (6) a history of periodontal surgical treatment in the specific quadrant of the mouth selected for the study, and (7) women who are pregnant or nursing mothers.

XSL•FO
RenderX

Sample Size Estimations

The following calculations were performed to estimate sample size:

- Mean PPD for AmnioGuard and NovaBone Putty (μ_1) before: 8.6 (SD 1.17) mm
- Mean PPD for AmnioGuard and NovaBone Putty (μ_2) after: 6.9 (SD 1.10) mm
- The standard normal deviate (*Z* value) corresponding to a 2-sided significance level (α) of .05, which equals a 95% confidence level, is given by $(1–\alpha/2)=1.96$
- $Z_{(1–\beta}$ [β=.05; power=95%]) = 1.64 ($Z_{(1-\beta)}$ represents the standard normal deviate corresponding to the study power. With β set at 0.05 (power=95%), the Z value used was 1.64. This differs from $Z_{(1-\alpha/2)}$, which corresponds to the significance level ($\alpha = 0.05$) used in the sample size calculation).
- Combining the 2 aforementioned equations yields the following: $n\geq[(Z_{[1-\alpha/2]} + Z_{[1-\beta]})^2 \times (\sigma_1^2 + \sigma_2^2/r)]/(\mu_1 - \mu_2)^2$
- Replacing the aforementioned values into the equation yields the following: $n\geq[(1.96 + 1.64)^2 \times (1.17^2 + 1.10^2/1)]/(8.6 – 6.9)^2=12$

Hence, group 1 and group 2 require a minimum sample size of 12. The minimum total sample size required is 24 [13].

All data will be tested for normality using the Shapiro-Wilk test and for homogeneity of variance using the Levene test to ensure that the assumptions for parametric tests are met.

Randomization

A total of 24 participants will be selected and receive a different treatment based on allocation to group A or B. For participants allocated to group A (n=12), NovaBone Putty and AmnioGuard will be placed into the osseous defect, and for participants allocated to group B (n=12), osseous defects will be treated with NovaBone Putty and then A-PRF membrane. Participants will be randomly allocated into the 2 treatment groups (NovaBone Putty, A-PRF, and AmnioGuard) using a computer-generated randomization sequence. The randomization list was created using an online random number generator, ensuring equal distribution of participants across both groups. Allocation concealment will be maintained by an independent clinician who is not involved in the clinical procedures or outcome assessments. Group assignments will be revealed only after patient enrollment and consent, thereby minimizing selection bias.

Procedure

Overview

Participants will be enrolled and randomly assigned to the interventions by the researchers in charge of the therapy application. Outcome evaluations were performed on the study participants in accordance with the described procedures and outcome measures. The 15-mm University of North Carolina (UNC-15) probe, which measures plaque index, bleeding on probing, probing depth, pocket depth, and clinical attachment loss, will be used to evaluate participants' periodontal clinical characteristics. Every patient will receive lessons on proper brushing technique as well as a demonstration of good oral hygiene practices.

Following a thorough examination and diagnosis, the first course of treatment will involve root planing under local anesthesia, supragingival and subgingival scaling, and recommendations on oral hygiene. If necessary, a coronoplasty will be carried out. Instructions for plaque control will be repeated until the patient's plaque score is less than 1. Six weeks after the initial therapy is finished, a re-evaluation will be conducted to assess the patients' reaction to the treatment and confirm the necessity of periodontal surgery.

A specially made occlusal acrylic stent will be used to standardize the probe's location and angulations. The occlusal stent will be constructed using acrylic on a cast model created from an alginate impression. The occlusal stent covers the occlusal surface of the tooth being treated as well as the occlusal surfaces of at least 1 tooth that is mesially and distally next to the afflicted tooth. The stent will also be expanded to cover the coronal part of the teeth. A reference point (slot) will be made on the stent at the deepest point of the impacted tooth to allow for repeatable periodontal probe placement. The apical border will be linear and serve as a fixed reference point.

Before the surgery, the patients will be told to rinse their mouths for a minute with a 0.2% chlorhexidine gluconate solution. Because the patient will be draped, only the oral cavity will be visible. Asepsis will be maintained during the procedure. Depending on the operation, either nerve block or infiltration anesthesia will be used for the surgical site using a local anesthetic solution consisting of 2% lignocaine and 1:100,000 epinephrine. The standard technique, which involves a periodontal access flap, will start with intracervical (sulcular) incisions performed on the buccal and lingual parts using Bard-Parker number 12 or 15 surgical blades. The incisions will be made as far interproximally as possible to preserve the entire interdental papillae and accomplish primary wound closure. Two teeth will be included in the flap: one mesial and one distal to the tooth associated with the defect. In the event that more access is required, divergent vertical relief incisions will be made one tooth distant from the problem. The full-thickness (mucoperiosteal) flap will be mirrored using a periosteal elevator (24G; Hu-Friedy) to reveal the alveolar bone in the area of the osseous defect. Extreme caution will be taken to avoid flap rupture or papillae loss when removal of granulomatous tissue from the inner aspect of the flap occurs. To create autologous PRF, the protocol calls for obtaining 6 to 10 mL of venous blood from the patient's forearm using an 18-gauge needle in a sterile plastic container. The blood will then be transferred to new, sterile test tubes without the addition of any anticoagulant, and the initial volume of blood will be split equally between 2 tubes that are positioned symmetrically around the rotor axis for proper balancing. The tubes will then be centrifuged for 10 minutes at 3000 revolutions per minute (rpm) to create autologous PRF.

Procedure for the Intervention Group

The defect will be completely isolated and hemostatized. To enable quick flap closure following the graft material

XSL•FO
RenderX

implantation, the flap will be presutured without tying the knot. With minimal pressure, NovaBone Putty and AmnioGuard will be inserted into the test site's osseous defect until the flap is raised, filling it up to the osseous wall's highest coronal level. To ensure primary flap closure, the flap will be coronally adjusted and sutured so that the flap edge is positioned 1 to 2 mm coronal to the cement-enamel junction.

Procedure for the Control Group

Except for packing osseous defects at control sites with NovaBone Putty and covering the bone graft with A-PRF membrane, the surgical procedure for the control group will be the same as the procedure for the intervention group.

Statistical Analysis

A calculation of means and SDs will be conducted, and the mean data will be analyzed using a standard statistical method to determine statistical significance. Unpaired and paired 2-tailed *t* tests will be used to evaluate data at baseline and 6 months for each group. A *P* value greater than .05 will be considered not statistically significant, whereas a *P* value of less than .05 will be deemed significant. The results will be calculated using the SPSS software (version 17; IBM Corp). Group A and group B will be compared based on clinical outcomes of plaque index, full-mouth bleeding score, PPD reduction, and CAL gain using an unpaired *t* test.

Evaluation Outcomes

The primary outcome of this study is radiographic bone fill at 6 months after surgery. This outcome is crucial as it indicates the extent of bone regeneration achieved through the different treatment modalities being tested. The secondary outcomes are PPD reduction, plaque index, papillary bleeding index, and CAL gain.

PPD reduction, which is used to evaluate the decrease in the depth of periodontal pockets, a sign of improved periodontal health, will be measured using an acrylic stent and a UNC-15 probe. Once the acrylic stent is in place, the UNC-15 probe will be positioned at the appropriate angle to access the deepest part of the interproximal pocket within the crevicular space. The point at which the probe makes contact with the stent will be marked in pencil. This will provide a consistent point of reference for subsequent measurements.

Plaque index will be calculated using the Turesky-Gilmore-Glickman modification of the Quigley-Hein index.

Papillary bleeding index will be calculated using the Muhlemann-Son Sulcus Bleeding Index.

CAL gain, which measures the improvement in the attachment of the periodontal tissue to the tooth, indicating successful regeneration, will be measured using a UNC-15 probe positioned at the base of the pocket.

Results

Ethical approval was obtained from the Institutional Ethics Committee of Datta Meghe Institute of Higher Education & Research on February 7, 2025. Recruitment began on September 16, 2025, and is ongoing. As of manuscript submission, 4 participants have been enrolled. Data collection is ongoing, and data analysis has not yet been completed. Results are expected to be published in 2027.

All the clinical measurements will be recorded at baseline and 6 months. Plaque index, which indicates the degree of full-mouth supragingival plaque accumulation, will be used to assess the patient's dental hygiene status and will be measured at the 1-month follow-up. Although a substantial difference between the groups is not anticipated, AmnioGuard and NovaBone Putty should show better results and improvement in clinical outcomes such as CAL gain and PPD reduction.

Discussion

Expected Findings

Regenerative therapy aims to restore the original structure and functionality of the periodontal complex, including the formation of new cementum on the tooth root and new periodontal attachment between newly formed bone and cementum, as well as to regenerate periodontal attachment when periodontal disease causes the attachment apparatus to be lost [14]. The process of GTR, a dental surgical method, uses barrier membranes to guide the creation of new bone and gingival tissue when there are insufficient quantities or dimensions of bone or gingiva for appropriate function, aesthetics, or prosthetic repair. "Guided bone regeneration" is a frequent term used to describe bone-regenerative procedures or ridge augmentation. Certain cell populations in the periodontium can form a new cementum, alveolar bone, and periodontal ligament if given the opportunity to populate the periodontal wound [15]. There are 4 of these progenitor cell types: cells from the epithelium, cells produced from gingival connective tissue, cells produced from alveolar bone, and cells generated from periodontal ligaments. This method has several applications in clinical practice, including diagnosis, treatment planning, and therapeutic management [16]. GTR is a therapeutic approach that permits the restoration of bone, cementum, and periodontal ligament in degranulated osseous lesions [17]. Since its inception, GTR therapy has been applied in a variety of ways. Nonresorbable occlusal barriers were initially used to identify abnormalities of degranulated periodontal infrabony defects. Over time, the phrase "combination GTR" was adopted, and flaws were filled using bone and bone replacements before being covered with a barrier. After some time, bioresorbable barriers replaced nonresorbable ones, and biological growth factors have been used more recently to promote recovery.

Alloplastic bone graft materials are commonly used, and bioactive glass is one of them. NovaBone Putty, made from bioactive glass graft material, has recently been released. Within minutes of implantation, a special surface reaction takes place when bodily fluids come into contact with the bioactive glass substance. The first step is an ionic exchange, in which the material's surface releases cations in return for hydrogen or hydronium ions, creating silanol groups (SiOH) [18]. Through a polycondensation reaction, silanol groups attach to one another to generate a silica-rich gel layer on the specific surface. Silica is essential for the development of bioactive glass's bone

XSL•FO
RenderX

bonding. A site for the redeposition of calcium and phosphorus from the blood and graft material is created by the silica-rich gel [19]. On top of the silica gel layer, a calcium phosphorus layer develops in a matter of hours. As the thickness and size of the layer increase over time, it transforms into a crystalline layer of hydroxycarbonate apatite (HCA), which is the same as bone material. The foundation for the link between this substance and bone is provided by this apatite layer [20]. The main benefit of bioactive glasses is their quick surface reaction rate, which promotes quick tissue attachment. The surface of the silica-rich layer is negatively charged. This raises the electrostatic charges sufficiently to accelerate the absorption of water. Water molecules and the silanol's hydroxyl groups form hydrogen bonds, which gives bioactive glass its cohesiveness. Osteoblastic stem cells, which develop into osteoblasts and form bone, are drawn to the negatively charged surface of the HCA layer by proteins such as growth factors and fibrin, which function as an organic glue [21]. Collagen adheres to the skin and becomes a part of the HCA layer. Collagen expansion up to the junctional epithelium indirectly inhibits apical migration of the junctional epithelium. Cementum repair may be aided by bioactive glass [22].

Other characteristics of AmnioGuard, including low immunogenicity, pain relief and epithelialization promotion, self-adhesiveness, and aesthetics, have also been demonstrated. The GTR procedure is frequently carried out in conjunction with the insertion of bone grafts or bone graft replacements beneath the membrane. AM (AmnioGuard) has been used in mouth cavity repair [23]. The interconnected porous structure promotes blood vessel formation, cell occlusiveness, and biodegradability. Proliferation factors included in the AM promote the proliferation of fibroblasts, which speeds up the creation of granulation tissue [24]. Meanwhile, the AM promotes neovascularization in the surrounding tissues and vascularizes good granulation tissue [25]. Additionally, the AM offers a bioactive matrix enhanced with proteins that promotes cell motility. Therefore, it may be hypothesized that using AM as a GTR membrane could encourage cell migration and wound healing as well as vascularization of the granulation tissue in the defects. The tissue of the AM is antibacterial. β-defensins, a significant class of antimicrobial peptides secreted by epithelial cells and a crucial component of the immune system, are produced by amniotic tissue.

A concentrated mixture of growth factors obtained from platelets is known as a platelet concentrate. Four peripheral blood tubes will be taken and put into a centrifuge that has been preprogrammed. Centrifugation will be carried out using the following two procedures: (1) A-PRF with sterile plain glass–based vacuum tubes (10 mL; 1500 revolutions per minute for 14 minutes) and (2) standard PRF with sterile glass-coated plastic tubes (9 mL; 2700 revolutions per minute for 12 minutes).

Limitations

The follow-up duration is relatively short (3 - 6 months), which may not fully capture the long-term stability and success of the regenerative treatment. Future studies with extended follow-up periods are needed to evaluate the durability of the clinical and radiographic improvements observed. The study will be conducted at a single institution (Sharad Pawar Dental College), which may limit the generalizability of the findings to broader populations with diverse demographic characteristics and clinical backgrounds. Multicenter studies would help improve external validity. The participants in this study may present with varying degrees of periodontal defect severity, which could introduce heterogeneity in treatment responses and influence the observed outcomes. The success of periodontal regenerative therapy is strongly influenced by patient compliance with postoperative oral hygiene instructions and maintenance care. Variability in individual patient cooperation could impact healing and regenerative results.

Conclusions

The combination of AmnioGuard and NovaBone Putty may be more effective in promoting bone regeneration in periodontal osseous defects than the combination of A-PRF and NovaBone Putty.

Acknowledgments

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this article.

Funding

No funding was received for this study.

Conflicts of Interest

None declared.

References

1. Kinane DF, Stathopoulou PG, Papapanou PN. Periodontal diseases. Nat Rev Dis Primers 2017 Jun 22;3:17038. [doi: 10.1038/nrdp.2017.38] [Medline: 28805207]
2. Goldman HM, Cohen DW. The infrabony pocket: classification and treatment. J Periodontol 1958 Oct;29(4):272-291 [FREE Full text] [doi: 10.1902/jop.1958.29.4.272]

XSL•FO
RenderX

3. Glossary of Periodontal Terms, 4th edition: American Academy of Periodontology; 2001. URL: https://books.google.com.ph/books/about/Glossary_of_Periodontal_Terms_2001.html?id=mjV5wgEACAAJ&redir_esc=y&utm_medium=email&utm_source=transaction [accessed 2025-05-25]
4. Hazari V, Choudhary A, Mishra R, Chandrashekar KT, Trivedi A, Pathak PK. Clinical and radiographic analysis of novabone putty with platelet-rich fibrin in the treatment of periodontal intrabony defects: a randomized control trial. Contemp Clin Dent 2021;12(2):150-156. [doi: 10.4103/ccd.ccd_101_20] [Medline: 34220155]
5. Niknejad H, Peirovi H, Jorjani M, Ahmadiani A, Ghanavi J, Seifalian AM. Properties of the amniotic membrane for potential use in tissue engineering. Eur Cell Mater 2008 Apr 29;15:88-99. [doi: 10.22203/ecm.v015a07] [Medline: 18446690]
6. Mamede AC, Carvalho MJ, Abrantes AM, Laranjo M, Maia CJ, Botelho MF. Amniotic membrane: from structure and functions to clinical applications. Cell Tissue Res 2012 Aug;349(2):447-458. [doi: 10.1007/s00441-012-1424-6] [Medline: 22592624]
7. Koob TJ, Lim JJ, Massee M, et al. Angiogenic properties of dehydrated human amnion/chorion allografts: therapeutic potential for soft tissue repair and regeneration. Vasc Cell 2014;6:10. [doi: 10.1186/2045-824X-6-10] [Medline: 24817999]
8. Malhotra C, Jain AK. Human amniotic membrane transplantation: different modalities of its use in ophthalmology. World J Transplant 2014;26:179-186 [FREE Full text] [doi: 10.5500/wjt.v4.i2.111]
9. Ghanaati S, Booms P, Orlowska A, et al. Advanced platelet-rich fibrin: a new concept for cell-based tissue engineering by means of inflammatory cells. J Oral Implantol 2014 Dec;40(6):679-689. [doi: 10.1563/aaid-joi-D-14-00138] [Medline: 24945603]
10. Hazari V, Choudhary A, Mishra R, Chandrashekar KT, Trivedi A, Pathak PK. Clinical and radiographic analysis of NovaBone Putty with platelet-rich fibrin in the treatment of periodontal intrabony defects: a randomized control trial. Contemp Clin Dent 2021;12(2):150-156. [doi: 10.4103/ccd.ccd_101_20] [Medline: 34220155]
11. Kothiwale SV, Anuroopa P, Gajiwala AL. A clinical and radiological evaluation of DFDBA with amniotic membrane versus Bovine derived xenograft with amniotic membrane in human periodontal grade II furcation defects. Cell Tissue Bank 2009 Nov;10(4):317-326. [doi: 10.1007/s10561-009-9126-3] [Medline: 19277899]
12. Mathur A, Bains VK, Gupta V, Jhingran R, Singh GP. Evaluation of intrabony defects treated with platelet-rich fibrin or autogenous bone graft: a comparative analysis. Eur J Dent 2015;9(1):100-108. [doi: 10.4103/1305-7456.149653] [Medline: 25713492]
13. Chakraborthy P, Ravishankar PL, Saravanan AV, et al. Effect of Amnioguard and Biomesh GTR membranes with NovaBone Putty in the treatment of periodontal osseous defects-a communication. Int J Environ Res Public Health 2023 Jan 1;20(1):816. [doi: 10.3390/ijerph20010816] [Medline: 36613137]
14. Miloro M, Ghali GE. Peterson's Principles of Oral and Maxillofacial Surgery: Springer; 2022. [doi: 10.1007/978-3-030-91920-7]
15. Nyman S, Gottlow J, Karring T, Lindhe J. The regenerative potential of the periodontal ligament. J Clinic Periodontology 1982 Jun;9(3):257-265. [doi: 10.1111/j.1600-051X.1982.tb02065.x] [Medline: 6954167]
16. Gottlow J, Nyman S, Lindhe J, Karring T, Wennström J. New attachment formation in the human periodontium by guided tissue regeneration. Case reports. J Clin Periodontol 1986 Jul;13(6):604-616. [doi: 10.1111/j.1600-051x.1986.tb00854.x] [Medline: 3462208]
17. Cortellini P, Pini Prato G, Tonetti MS. Periodontal regeneration of human infrabony defects. II. Re‐entry procedures and bone measures. J Periodontol 1993 Apr;64(4):261-268. [doi: 10.1902/jop.1993.64.4.261] [Medline: 8483088]
18. Ducheyne P, Brown S, Blumenthal N, et al. Bioactive glasses, aluminum oxide, and titanium. Ion transport phenomena and surface analysis. Ann N Y Acad Sci 1988;523:257-261. [doi: 10.1111/j.1749-6632.1988.tb38517.x] [Medline: 3382125]
19. Kitsugi T, Nakamura T, Oka M, Cho SB, Miyaji F, Kokubo T. Bone-bonding behavior of three heat-treated silica gels implanted in mature rabbit bone. Calcif Tissue Int 1995 Aug;57(2):155-160. [doi: 10.1007/BF00298437] [Medline: 7584877]
20. Greenspan DC, Zhong JP, Torre GP. The evaluation of surface structure of bioactive glasses in-vitro. In: Wilson J, Hench LL, Greenspan D, editors. Bioceramics: Proceedings of the 8th International Symposium on Ceramics in Medicine, Ponte Vedra, Florida, USA, November 1995 1995:477-482 URL: https://books.google.co.in/books/about/Bioceramics.html?id=n1lRAAAAMAAJ&redir_esc=y [accessed 2026-02-13]
21. Hench LL. Bioactive ceramics. Ann N Y Acad Sci 1988;523:54-71. [doi: 10.1111/j.1749-6632.1988.tb38500.x] [Medline: 2837945]
22. Wilson J, Low SB. Bioactive ceramics for periodontal treatment: comparative studies in the Patus monkey. J Appl Biomater 1992;3(2):123-129. [doi: 10.1002/jab.770030208] [Medline: 10147709]
23. Ali S, Farooq I, Iqbal K. A review of the effect of various ions on the properties and the clinical applications of novel bioactive glasses in medicine and dentistry. Saudi Dent J 2014 Jan;26(1):1-5. [doi: 10.1016/j.sdentj.2013.12.001] [Medline: 24526822]
24. Mohan R, Bajaj A, Gundappa M. Human amnion membrane: potential applications in oral and periodontal field. J Int Soc Prev Community Dent 2017;7(1):15-21. [doi: 10.4103/jispcd.JISPCD_359_16] [Medline: 28316944]
25. Tunggal P, Smyth N, Paulsson M, Ott MC. Laminins: structure and genetic regulation. Microsc Res Tech 2000 Nov 1;51(3):214-227. [doi: 10.1002/1097-0029(20001101)51:3<214::AID-JEMT2>3.0.CO;2-J] [Medline: 11054872]

XSL•FO
RenderX

Abbreviations

A-PRF: advanced platelet-rich fibrin
AM: amniotic membrane
CAL: clinical attachment level
GTR: guided tissue regeneration
HCA: hydroxycarbonate apatite
PPD: probing pocket depth
PRF: platelet-rich fibrin
rpm: revolutions per minute
UNC-15: 15-mm University of North Carolina probe

Edited by J Sarvestan; submitted 14.Apr.2025; peer-reviewed by AC Das; revised version received 24.Jul.2025; accepted 31.Aug.2025; published 26.Feb.2026.

Please cite as:
Bhandari G, Jaiswal P, Bhagat S, Kotecha S
Comparative Evaluation of AmnioGuard and Advanced Platelet-Rich Fibrin When Combined With NovaBone Putty in the Regeneration of Human Periodontal Infrabony Defects: Protocol for a Randomized Controlled Clinical Trial
JMIR Res Protoc 2026;15:e75987
URL: https://www.researchprotocols.org/2026/1/e75987
doi:10.2196/75987

© Geeta Bhandari, Priyanka Jaiswal, Shweta Bhagat, Sakshi Kotecha. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Preventing and Treating Insomnia Symptoms in Midlife and Older Adults (ASLEEP): Protocol for a Randomized Controlled Trial Using the PROTECT Norge Infrastructure

Jon Arild Aakre[1,2], PhD; Bjørn Bjorvatn[3,4], MD, PhD; Martha Therese Gjestsen[1,2], PhD; Ingvild Dalen[5], PhD; Clive Ballard[6], MD; Ingelin Testad[1,7], PhD

[1]Centre for Age-Related Medicine - SESAM, Stavanger University Hospital, Jan Johnsens gate nr 16, Stavanger, Norway

[2]Department of Clinical Medicine, University of Bergen, Bergen, Norway

[3]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway

[4]Norwegian Competence Center for Sleep Disorders, Haukeland University Hospital, Bergen, Norway

[5]Department of Research, Stavanger University Hospital, Stavanger, Norway

[6]Department of Clinical Biosciences, Faculty of Health and Life Sciences, University of Exeter, Exeter, United Kingdom

[7]Department of Health and Community Sciences, Faculty of Health and Life Sciences, University of Exeter, Exeter, United Kingdom

Corresponding Author:
Jon Arild Aakre, PhD
Centre for Age-Related Medicine - SESAM, Stavanger University Hospital, Jan Johnsens gate nr 16, Stavanger, Norway

Abstract

Background: Sleep is increasingly recognized as a fundamental determinant of health and brain function. Sleep difficulties are common in older adults, with a substantial proportion reporting problems initiating or maintaining sleep, which can negatively affect mental and physical health, cognitive function, and quality of life. Cognitive behavioral therapy for insomnia (CBT-I) is the gold-standard treatment for insomnia disorder; however, its reach is limited due to resource demands and a shortage of professionals that can deliver it. Digitally delivered CBT-I via eHealth platforms increases accessibility and has demonstrable effects but remains limited in many countries.

Objective: The objective of this paper is to describe the protocol for the further development and evaluation of ASLEEP (Preventing and Treating Insomnia Symptoms in Midlife and Older Adults), a tiered, digitally delivered CBT-I intervention designed to reduce insomnia severity and improve related health outcomes in adults aged 50 years and older.

Methods: The project will be conducted in 2 phases. Phase 1 focuses on refining and optimizing ASLEEP, developing an advanced CBT-I course, and integrating a nested trial into PROTECT (Platform for Research Online to investigate Cognition and Genetics in Ageing) Norge, a fully automated digital research platform. Phase 2 is a fully digital, 2-arm, waitlist-controlled randomized controlled trial, with 400 participants randomized 1:1 to the intervention or waitlist control and allocation stratified by age and insomnia severity. Outcomes will be assessed at baseline and at 3, 6, and 12 months, with a 15-month follow-up for the waitlist group. The primary outcome is insomnia severity measured by the Insomnia Severity Index. Secondary outcomes include sleep medication use, depression, anxiety, and cognition.

Results: The project started in January 2026, with funding awarded. As of February 2026, phase 1—intervention optimization and development—is underway. Ethics approval for ASLEEP has not been submitted. Following completion of phase 1, phase 2, which includes a digital randomized controlled trial, will commence; as of February 2026, no participants have been recruited, and data collection and data analysis have not yet started. Short-term data collection is planned to be completed by summer 2028, with results disseminated in winter 2028.

Conclusions: This trial will evaluate the short- and long-term effectiveness of a tiered digital CBT-I intervention for midlife and older adults. By leveraging the PROTECT Norge platform and if effective, ASLEEP may represent a scalable model for low-threshold, accessible prevention and treatment of symptoms of insomnia.

International Registered Report Identifier (IRRID): PRR1-10.2196/81542

(JMIR Res Protoc 2026;15:e81542) doi:10.2196/81542

KEYWORDS

insomnia symptoms; cognitive behavioral therapy for insomnia; eHealth; randomized controlled trial; older adults; sleep intervention; PROTECT Norge

XSL•FO
RenderX

Introduction

The Role of Sleep in Healthy Aging and Brain Health

Sufficient nighttime sleep is vital for maintaining optimal daytime functioning and supporting overall health and quality of life. For most adults, this generally involves obtaining 6 to 9 hours of sleep per night [1]. With increasing age, meeting individual sleep needs in terms of both duration and quality becomes harder [2], and up to 50% of older adults report difficulties initiating or maintaining sleep [3]. Failure to meet sleep needs and chronic sleep deprivation impair mood and emotional regulation; disrupt multiple physiological systems and cognitive functions; weaken immune responses; and increase the risk of several noncommunicable diseases, including cardiovascular disease, diabetes, and obesity, as well as mental health disorders [4]. Beyond these individual consequences, sleep deprivation also carries societal costs, contributing to increased health care expenditures, reduced productivity, and higher rates of workforce absenteeism, collectively amounting to 1% to 3% of gross domestic product in OECD countries [5].

Poor sleep is associated with poorer brain health in midlife, suggesting that sleep is a potential target for maintaining brain health into older age [6]. Experimental studies have shown that sleep deprivation impairs the efficiency of the glymphatic system, a key pathway for clearing neurotoxic waste products from the brain during sleep [7]. This impairment has been associated with β-amyloid deposition in older adults, which may in turn contribute to Alzheimer disease risk [8]. These mechanistic insights provide one of several pathways that may help explain the observed epidemiological association between sleep deprivation in midlife and an increased incidence of dementia later in life [9]. Beyond experimental and epidemiological evidence, sleep is increasingly recognized as a vital public health priority [4] and has been formally positioned as an equal pillar of health alongside physical activity and nutrition [10], underscoring its essential role in maintaining both brain function and overall health and well-being.

When Poor Sleep Develops Into Insomnia: Insomnia Treatment Using Cognitive Behavioral Therapy

Difficulties falling asleep or staying asleep, or waking too early are the defining symptoms of insomnia. Moreover, a symptom burden with a frequency equal to or exceeding 3 days per week, persisting for over 3 months, resulting in daytime impairment, warrants a diagnosis of chronic insomnia [11]. While general sleep deprivation is associated with widespread physiological and cognitive consequences across multiple systems, insomnia is notably characterized by a strong and well-documented bidirectional relationship with mental health conditions, particularly anxiety and depression [12]. According to the Harvey cognitive model, insomnia is not just a consequence of insufficient sleep; rather, maladaptive thoughts, heightened arousal, and counterproductive behaviors interact to maintain the disorder through a feedback loop of worry, arousal, and behaviors that make sleep more difficult [13]. Building on this model, cognitive behavioral therapy for insomnia (CBT-I) is currently the gold-standard treatment for chronic insomnia and is a highly effective and safe treatment [11]. CBT-I combines behavioral strategies (eg, sleep restriction, stimulus control, and relaxation), cognitive techniques (eg, restructuring unhelpful thoughts, mindfulness, and paradoxical intention), and educational elements, including psychoeducation and sleep hygiene, and is usually delivered by trained professionals, such as psychologists or psychiatrists, over several structured sessions tailored to the patient's needs [14]. It is a resource-demanding treatment, and combined with a lack of professionals with the competency to deliver CBT-I, it is usually reserved for those with a severe diagnosis of insomnia [15]. In Norway, Professor Bjørn Bjorvatn has published a CBT-I–based self-help book targeting adults with insomnia. In a randomized controlled trial (RCT) among self-referred participants, the tools and strategies presented in the book improved sleep, reduced use of sleep medication, and slightly decreased depressive symptoms compared with standard sleep hygiene advice [16]. A subsequent RCT among general practitioner–referred patients showed similar benefits, including reductions in sleep medication use and anxiety, as well as improved sleep outcomes [17]. These studies indicate that the self-help book is an effective low-threshold intervention for adults with insomnia in both community and primary care settings.

eHealth Platforms for Increasing the Reach and Dissemination of CBT-I

The continuing digitization of people's daily life has paved the way for new platforms delivering CBT-I via the internet (eCBT-I). eCBT-I offers wide reach and accessibility and serves as an effective alternative or supplement to traditional, in-person-delivered CBT-I [18]. However, access to evidence-based, eCBT-I programs, such as SLEEPIO [19] and Somryst [20], which are available by prescription in some health care systems, remains limited in many countries, including Norway.

The ASLEEP project (Preventing and Treating Insomnia Symptoms in Midlife and Older Adults) aims to enhance access to sleep education and insomnia symptom treatment by leveraging eHealth technologies. The intervention is based on the aforementioned CBT-I–based self-help book [16] and is tailored to the needs of an aging population through a tiered delivery model. The project is supported by the Platform for Research Online to Investigate Cognition and Genetics in Ageing Norge (PROTECT Norge) [21], which was initiated by the Centre for Age-Related Medicine (SESAM) at Stavanger University Hospital in 2020. PROTECT Norge is a digital, prospective research cohort with yearly follow-ups. Focusing on the brain aging, it assesses physical and mental health and lifestyle factors through validated questionnaires, as well as cognitive function using a neuropsychiatric test battery. In addition, the PROTECT infrastructure, originating from University of Exeter and Kings College London in the UK, includes a built-in research infrastructure designed to support large-scale digital RCTs [22,23]. The PROTECT Norge cohort [21] currently consists of 5547 registered individuals, who primarily have been recruited through social media, particularly Facebook. Interested individuals self-refer and are directed to the PROTECT Norge study website [24], where they complete eligibility screening, registration, and informed consent through a secure, fully automated online process. Once enrolled,

XSL•FO
RenderX

participants complete annual validated cognitive and mental health assessments for up to 10 years and provide detailed information on lifestyle, medical, and demographic factors. Participants are also asked to provide consent for contact (C4C), which has been granted by 94% of PROTECT Norge participants [21]. Combining this with the PROTECT research infrastructure, which enables the conduction of nested studies, PROTECT Norge is well placed to deliver large-scale interventions and clinical trials in a time- and resource-efficient manner within the PROTECT Norge study trajectory.

To prevent the development of insomnia and promote good sleep, the ASLEEP intervention offers a basic course (tier 1). Supported by the PROTECT Norge infrastructure, a screening procedure will be developed to flag individuals with symptoms of insomnia of a degree that suggest the need for more intensive eCBT-I treatment and to guide those individuals to an advanced course (tier 2). The PROTECT Norge research infrastructure has already had its feasibility of delivering ASLEEP as a nested study in PROTECT Norge assessed, and the acceptability of the basic course content was also evaluated in the same study by this group (JA Aakre, PhD; I Testad, PhD; MT Gjestsen, PhD; et al; unpublished data; November 2025).

This paper describes the protocol for the ASLEEP project, including further development of the ASLEEP intervention to include an advanced course and its digital delivery (phase 1), as well as the evaluation of its effectiveness (phase 2) through a 2-arm RCT, with a waitlist design. The RCT will be conducted entirely digitally and facilitated by the PROTECT Norge research infrastructure.

The objectives of the ASLEEP project are to:

1. Further develop the ASLEEP intervention, an eCBT-I–based, tiered sleep program, including an advanced course aimed at reducing insomnia symptoms in adults aged 50 years and older.
2. Conduct short- and long-term evaluations using the PROTECT Norge infrastructure to assess the effectiveness of the ASLEEP advanced course on sleep-related outcomes among midlife and older adults.
3. Evaluate the effectiveness of ASLEEP on established outcomes that exist concomitantly or are related to insomnia symptoms, such as medication use, depression, anxiety, and cognitive function.

Hypotheses

Our first hypothesis (H1) is that further development of the ASLEEP intervention—an eCBT-I–based, tiered sleep program—and its subsequent dissemination within the PROTECT Norge cohort will lead to a significant reduction in insomnia symptoms among individuals aged 50 years and older compared with waitlist control group.

Our second hypothesis (H2) is that ASLEEP will result in short-term improvements in sleep outcomes (Insomnia Severity Index [ISI] score and sleep medication use) among midlife and older adults, and these improvements will be sustained in the long-term.

Our third hypothesis (H3) is that ASLEEP will not only improve sleep outcomes but also lead to significant reductions in symptoms of depression, anxiety, and cognitive decline among midlife and older adults compared to a waitlist control group.

Methods

Reporting Guidelines

This protocol is reported in accordance with the Guidance for Reporting Involvement of Patients and the Public–Short Form (GRIPP2-SF) checklist [25] and the SPIRIT 2025 (Standard Protocol Items: Recommendations for Interventional Trials) guidelines (See Checklist 1 and Checklist 2).

Intervention

Following development conducted in phase 1 of the project, the ASLEEP intervention will consist of 2 courses. The basic course has been developed to promote good sleep in a broad group of adults aged 50 years and older (JA Aakre, PhD; I Testad, PhD; MT Gjestsen, PhD; et al; unpublished data; November 2025) and starts off with a video providing instructions on how to complete the digital sleep diary. Participants will then complete a 7-day sleep diary. After completion of the sleep diary, 3 educational modules will be unlocked and delivered through short videos (see Table 1 for video descriptions) focusing on healthy sleep and sleep habits. Then, feedback reports based on recorded sleep through the sleep diary will be made available to participants, with a focus on sleep efficiency. Findings from the acceptability and feasibility study (JA Aakre, PhD; I Testad, PhD; MT Gjestsen, PhD; et al; unpublished data; November 2025) showed that the basic course content was well accepted, with participants describing the videos as clear, systematic, appropriately lengthy, and helpful for understanding sleep and their own sleep patterns. Several participants appreciated the expert-led dialogue format of the videos. The sleep diary was also viewed as meaningful, increasing awareness of sleep habits and prompting daily reflection on factors affecting sleep.

XSL•FO
RenderX

Table . Video descriptions.

Video (length)	Name	Description
Sleep diary instruction (17 min 17 s)	An instruction to keeping digital sleep diary	This video provides step-by-step instructions on how to keep a sleep diary, along with clarification of common misunderstandings people may have when logging their sleep.
Video module 1 (20 min 3 s)	What is sleep and how it is regulated	This video explains what sleep is, its role in the body, and the biological processes that govern it.
Video module 2 (18 min 10 s)	Assessing poor sleep	This video addresses how poor sleep can be assessed, discusses its possible causes, approaches to diagnosis, and provides a general overview of different sleep disorders.
Video module 3 (30 min 10 s)	Treatment of insomnia	This video module focuses on the management of insomnia, including definitions, prevalence, and treatment strategies for both short-term and chronic insomnia, as well as principles of sleep regulation. Additionally, Professor Bjorvatn gives sleep hygiene advice, discussing ways to maintain sufficient sleep, support a healthy circadian rhythm, and minimize evening and nighttime arousal through habits and behavioral practices.

As part of phase 1 of the project, the ASLEEP advanced course will be developed and subsequently offered to participants in a fully digital RCT to assess its effect on sleep outcomes. Eligible participants who have completed the basic course will be funneled into the advanced course if their post–basic course score on the ISI [25] is 10 or higher (defined as symptoms of insomnia) [26].

The advanced course will be developed according to established protocol for CBT-I, drawing on core components such as sleep restriction, stimulus control, cognitive restructuring, sleep-hygiene reinforcement, and relapse prevention [27]. However, unlike other existing digital CBT-I programs (eg, SHUTi, Sleepio), which deliver these components within a single treatment stream, ASLEEP applies a stepped-care approach [28] in which only participants who continue to report insomnia symptoms after completing the basic course progress to the more treatment-intensive modules. The ASLEEP advanced course is expected to consist of a series of weekly digital modules delivered over approximately 6 weeks, although the exact number and pacing of modules will be finalized during the development phase. Consistent with evidence identifying sleep restriction and stimulus control as the most effective components of CBT-I [29], these components will constitute the core of the intervention. Sleep restriction aims to consolidate sleep and strengthen homeostatic sleep pressure by limiting time in bed to approximate actual sleep duration and gradually expanding it as sleep becomes more efficient [28]. Stimulus control complements this by re-establishing the bed and bedroom as cues for sleep rather than wakefulness through structured behavioral instructions, including going to bed only when sleepy, leaving the bed if unable to sleep, and maintaining consistent rise times.

A sleep diary will be completed throughout the advanced course to track progress and to tailor individual intervention recommendations, including ongoing adjustment of the sleep-restriction schedule.

Phase 1: ASLEEP Development, Implementation, and Integration

The ASLEEP project is guided by the updated Medical Research Council (MRC) framework for the development and evaluation of complex interventions, which emphasizes an iterative process integrating development, feasibility testing, evaluation, and implementation, supported by ongoing consideration of context, program theory, stakeholder engagement, and economic factors [30]. In line with this framework, we have already developed (phase 1 of the MRC framework) and assessed the acceptability of the basic course and the feasibility of delivering ASLEEP through the PROTECT Norge infrastructure (phase 2 of the MRC framework) (JA Aakre, PhD; I Testad, PhD; MT Gjestsen, PhD; et al; unpublished data; November 2025), and we are moving on to the next phases of development and evaluation.

Building on the findings from the acceptability and feasibility study showing that some participants missed the opportunity to use their own devices rather than a study-provided iPad (JA Aakre, PhD; I Testad, PhD; MT Gjestsen, PhD; et al; unpublished data; November 2025), the user interface will be optimized to support access across devices, including smartphones, tablets, and personal computers, via both a mobile app or a web browser. This process will be informed by human-centered design to ensure user and stakeholder engagement in the development process. While the MRC framework provides overarching guidance for complex intervention development and evaluation, design thinking adds a complementary methodological perspective to the development phase of the project.

Integration of the ASLEEP study, including registration, consenting and screening, web page text, study documents, and new questionnaires on the PROTECT Norge infrastructure, will undergo multiple rounds of user acceptance testing (UAT) in a process described in a previous publication.

XSL•FO
RenderX

As part of the implementation work, a dialogue with the Western Regional Health Authority digitalization department has been initiated to discuss technical compatibility and potentially building on existing regional digital health infrastructures, thereby supporting potential future integration into existing health care pathways and real-world deployment.

Phase 2: Short- and Long-Term Evaluation of ASLEEP

RCT Design

This study will be an RCT nested within the cohort of PROTECT Norge, a 10-year digital follow-up study with annual data collection since the start in 2020. The treatment arm will be compared with a waitlist control group. Screening, recruitment, and data collection will be conducted through the PROTECT platform. This RCT will be registered in ClinicalTrials.gov.

Comparison Group

Eligible participants that are not randomized to the intervention will go into a waitlist control group and will receive access to the intervention after a predetermined delay.

Recruitment

Participants will be recruited from the PROTECT Norge cohort among individuals who have provided consent for contact regarding future studies. A general invitation will be sent to the whole cohort to perform the basic (tier 1) ASLEEP course. Subsequent scoring of the ISI [26] score and the Global Sleep Assessment Questionnaire [31] will be used to evaluate eligibility for the advanced (tier 2) ASLEEP course and hence the RCT (see further details below).

Eligibility Criteria

Inclusion criteria are as follows: participants must have completed the ASLEEP basic course, have an ISI score of 10 or higher, and have an insomnia duration exceeding 3 months.

Exclusion criteria are as follows: participants with insomnia symptoms that stem from other sleep disorders, including obstructive sleep apnea, restless legs syndrome, parasomnias, and hypersomnia, will be screened for based on selected items (Q2, Q6, Q7, and Q9) from the Global Sleep Assessment Questionnaire [32].

Randomization and Masking

Participants will be randomized in a 1:1 ratio to either the intervention group or the waitlist control group using simple block randomization with varying block sizes of 6 to 10 participants. The allocation sequence will be computer generated in advance by a member of the research team not involved in participant recruitment or follow-up.

Randomization will be performed after completion of baseline assessments and confirmation of eligibility. Allocation will be implemented manually by authorized PROTECT Norge study staff using the pregenerated allocation sequence. Based on group assignment, participants will be provided with a personal access code to be entered on the PROTECT Norge study website. Entry of the code triggers the predefined scheduling of study activities within the platform.

Members of the research team involved in data analysis will be blinded to group allocation until the inferential analyses of the primary 3-month outcomes have been finalized. Study staff with direct participant contact through the PROTECT Norge help desk will not take part in data analysis.

Intervention Fidelity

A core component of the intervention is the completion of a daily sleep diary, which supports both treatment delivery and progress monitoring. To assess adherence, participants will complete a brief questionnaire on diary compliance through their PROTECT Norge study dashboard. The PROTECT Norge infrastructure for real-time tracking of all study activities, including study activity status. Study help desk staff will manually monitor study activity in real time, including diary entries and assessment completion. These data will allow real-time monitoring of adherence. They will be asked to complete and submit this compliance log within a reasonable timeframe following the end of the intervention. For each day of the study, participants will indicate whether the diary was completed in the morning, evening, or both, with additional options for “I do not know” or “Did not complete.” This structured measure of adherence provides insight into participant engagement and supports evaluation of intervention fidelity.

Outcome Assessment Timing

Baseline data will be obtained from all potential participants as part of the screening. At 3 months postrandomization, a similar data collection will be performed. These data will serve as the primary short-term outcome data for the study. For the waitlist group, they will also serve as an updated baseline for their intervention period, which starts immediately after they provided these data (see Textbox 1).

XSL•FO
RenderX

Textbox 1. Overview of data collection time points.

Intervention arm

- T0: Baseline
- T0 to T1: Intervention
- T1 (T0+3 mo): Follow-up (FU) 1
- T2 (T0+6 mo): FU2
- T3 (T0+12 mo): FU3

Control arm

- T0: Baseline
- T0 to T1: Waiting list
- T1 (T0+3 mo): FU1
- T2 (T0+6 mo): Not applicable
- T3 (T0+12 mo): Not applicable
- T0'=T1: Updated baseline
- T0' to T1': Intervention
- T1' (T0'+3 mo): FU1'
- T2' (T0'+6 mo): FU2'
- T3' (T0'+12 mo): FU3'

Outcomes

Study outcomes will be evaluated using validated instruments to capture the effects of the intervention on sleep, cognition, and related mental health domains through PROTECT Norge built-in clinical trials infrastructure. The primary outcome is insomnia severity, assessed using the ISI. All other outcomes are considered secondary outcomes (see Table 2 for details). Use of sleep medication is a secondary outcome and will be captured on a daily interval through the digital sleep diary.

Table . Screening instruments and outcome measures used in the study, assessed through the Platform for Research Online to Investigate Cognition and Genetics in Ageing (PROTECT) Norge infrastructure.

Short name	Long name	Details	Time frame	Assessment points
ISI	Insomnia Severity Index [25]	7 items; scale 0-28, higher scores indicate more severe insomnia symptoms.	Past 2 wk	Screening/baseline + all follow-ups
SPD	Sleep problem duration item	One item assessing duration of sleep problem, scored 0=no sleep problem to 3=3 months or more	Current	Screening
GSAQ	Global Sleep Assessment Questionnaire [32]	Four separate items scored 0-4 will be used to screen for sleep disorders	Past 4 wk	Screening
FLAME	Factors of Longitudinal Attention, Memory and Executive Function [31]	Computerized neuropsychological battery with 8 tests. Gives composite score combining subscores on domains such as attention (speed and accuracy), memory, and executive function; higher scores reflect better overall cognitive performance.	Current	Baseline + all follow-ups
PHQ-9	Patient Health Questionnaire [33]	9 items; scale 0-27, higher scores indicate more severe depressive symptoms.	Past 2 wk	Baseline + all follow-ups
GAD-7	Generalized Anxiety Disorder scale [34]	7 items; scale 0-21, higher scores indicate more severe anxiety symptoms.	Past 2 wk	Baseline + all follow-ups

XSL•FO
RenderX

Statistical Analysis

Descriptive statistics will be presented for baseline demographics, including age, gender, ethnicity, marital status, education, and employment status. Counts and percentages (categorical variables) and either means and SDs (symmetric continuous variables) or medians and IQR (skewed continuous variables) will be presented for the total sample and per intervention arm. Similarly, descriptive statistics will be presented for primary and secondary outcome variables, for all time points.

The primary estimand will be the average treatment effect (ATE), defined in accordance with the International Council for Harmonisation (ICH) E9(R1) guideline on estimands and sensitivity analysis as the adjusted mean difference in insomnia symptoms measured as ISI scores between the intervention and waitlist arms at three months postbaseline. Nonadherence will be handled using the treatment policy strategy (which reflects the intention-to-treat principle under the estimand framework), meaning that the analysis will include all participants according to their randomized allocation, regardless of adherence.

The primary inferential analysis will thus be the comparison of the intervention and the waitlist arms at 3 months post baseline with regard to mean ISI scores. This comparison will be performed using analysis of covariance (ANCOVA), implemented as a linear regression model with follow-up ISI as the dependent variable, intervention arm as the independent variable, and baseline ISI as a covariate. The interaction between intervention arm and baseline ISI will be included to allow for slope heterogeneity. To improve precision, further adjustments will be made for potentially predictive baseline variables, ie, the demographic variables listed above plus baseline measures of depression (Patient Health Questionnaire-9) and anxiety (Generalized Anxiety Disorder-7). All continuous covariates will be entered linearly. The primary contrast will be reported as a marginal (population average) mean difference, obtained via model-based standardization over the joint baseline covariate distribution of the randomized sample.

The estimated ATE will be reported with a 95% CI based on a heteroscedasticity-robust SE and corresponding 2-sided *P* value. An effect size will be calculated by dividing the ATE by the pooled SD of baseline ISI scores.

Missing 3-month ISI and any missing baseline covariates will be handled using multiple imputation by chained equations (MICE). Imputation models will include all variables specified for the substantive model plus auxiliary variables given by 3-month measures of Patient Health Questionnaire-9, Generalized Anxiety Disorder-7, and FLAME. A total of m=100 imputed datasets will be generated, using logistic regression for dichotomous variables and otherwise predictive mean matching (k=5). Each imputed dataset will be analyzed with the primary ANCOVA, and effect estimates and standard errors will be combined using Rubin rules.

Supplementary analyses will assess the complier average causal effect. Sensitivity to missing not at random dropouts will be explored in a tipping-point (delta) analysis, by applying additive shifts to the imputed three months ISI scores separately by arm. In case of substantial skewness of the outcome, explorative sensitivity analyses will apply either Box-Cox transformation of the outcome or other regression models.

For the assessment of long-term effects on insomnia symptoms, there is no control group; thus, no conclusions will be drawn regarding treatment effects. The longitudinal development of ISI scores from baseline to 3, 6, and 12 months will be studied in linear mixed models, including random intercepts and slopes to allow for correlation between repeated measurements. Data from the original intervention arm will be combined with the (3 mo + delayed) data from the original waitlist arm for these analyses. Time since baseline (updated baseline for the waitlist arm) will be treated as a categorical variable.

Secondary outcomes will be analyzed similarly to ISI scores, with assessment of treatment impact at three months and longitudinal descriptive analysis of long-term trajectories, including covariance with insomnia symptoms. However, all secondary analyses will be exploratory.

Power Analysis

At present, more than 3500 active participants with an active consent for contact in the PROTECT cohort. With the planned expansion, approximately 5000 participants are expected to be available at study initiation. Assuming that 15% of these (approximately 750 individuals) will meet eligibility criteria, we aim to recruit 500 participants (approximately 67% of those eligible). Allowing for 20% attrition in both arms prior to the 3-month follow-up, we expect 400 participants to be available for the primary efficacy analysis (200 per arm). This sample size provides more than 80% power to detect a small-to-medium effect size of 0.3 using a 2-sided significance level of α=.05. Applying ANCOVA in the primary analysis is expected to further increase statistical power, depending on the proportion of variance explained (R^2) by baseline variables.

Ethical Considerations

This project will be conducted in accordance with the Declaration of Helsinki [35] and complies with the General Data Protection Regulation (GDPR, 2018) and Stavanger University Hospital's data security policy.

The ASLEEP RCT will be conducted as a nested study within the PROTECT Norge infrastructure, which was launched in 2020 and approved by the Regional Committee for Medical and Health Research Ethics, West (REK vest, reference 2019/478). A separate application for ethical approval for the ASLEEP project will be submitted to REK, and internal approval will be sought at Stavanger University Hospital, including a data protection impact assessment for any new technology used to process participant data. Only participants who have previously provided broad consent (C4C) for future studies within PROTECT Norge will be invited, and all will provide additional written informed consent specific to the ASLEEP RCT study through a secure and validated procedure. Participation is voluntary, and participants may withdraw from the study at any time, without giving a reason. No financial compensation will be provided for participation.

XSL•FO
RenderX

All study data will be stored on secure cloud servers in Norway, with regular backups and monthly data transfers from the PROTECT Norge infrastructure to internal servers at Stavanger University Hospital. Any significant protocol modifications (eg, changes to eligibility criteria, outcomes, or analyses) will be submitted for review by the REK, updated in the trial registry and appropriate internal registries at Stavanger University Hospital. Personally identifiable information will be stored separately from research data, and data will be retained for 10 years following study completion and then permanently anonymized by deleting the reidentification key. Access to personally identifiable information will be restricted to authorized research personnel only.

Specifically for the phase 2 RCT, access to identifiable data will be restricted to authorized study staff with direct participant contact through the PROTECT Norge help desk via a secure, access-controlled management portal requiring authenticated login. Pseudonymized datasets will be used for data analysis by members of the research team involved in data analysis. Study staff responsible for system administration, participant support, and adherence monitoring may have access to operational study data but will not take part in data analysis.

Data quality will be ensured through standard platform-based validation procedures and institutional quality assurance routines.

A waitlist-controlled design has been chosen to ensure that all participants will ultimately receive access to the intervention while enabling robust evaluation. Given the low-risk nature of this behavioral, fully digital intervention, no physical harms are anticipated, but depression symptoms will be regularly monitored to identify any signs of distress or harmful behavior, with follow-up procedures in place to provide appropriate support if needed.

Patient and Public Involvement

Patient and public involvement (PPI) is a guiding pillar at the SESAM and shaping both the design and execution of research. To strengthen this commitment, SESAM established WiseAge in 2015, a dedicated platform that promotes meaningful PPI across all projects, from initial idea to full implementation. The overarching goal is to enhance the quality and relevance of research by grounding it in real-world need [36,37]. To support this, SESAM and WiseAge have developed a practical guideline for embedding PPI throughout the research process, aimed at professionals, researchers, and user representatives alike [38].

In accordance with these guidelines, the ASLEEP project has been presented to the WiseAge community panel, a forum of 15 dedicated members, ensuring broader user dialogue and input. In addition, two user representatives have been part of the ASLEEP project group from the very beginning. Their contributions have been substantial, ranging from feedback on study documents, funding applications, video materials, and sleep reports, to active participation in the development of the ASLEEP basic course and UAT reported in a previous study (JA Aakre, PhD; I Testad, PhD; MT Gjestsen, PhD; et al; unpublished data; November 2025) and dissemination. Building on this, users will remain actively engaged in Phase 1 of the ASLEEP project by being engaged in the workgroup responsible for refining the intervention and designing the delivery software in line with the human-centered design approach. In Phase 2, they will continue their established role by contributing to UAT of the nested RCT conducted on the PROTECT Norge infrastructure, while also supporting project dissemination and recruitment, ensuring continuity of user-centered input throughout the project.

Results

The ASLEEP project started in January 2026, with funding awarded. Under the funding scheme, the DAM Foundation requires the project to be funded and published as a registered report. This paper represents the stage 1 registered report, and the results will be reported in a stage 2 registered report. As of February 2026, intervention development has commenced, initiated by a multistakeholder prototyping process involving user representatives, members of the research group, and representatives from the hospital's Information and Communication Technology (ICT) and Innovation Departments. Subsequently, integration of the ASLEEP study as a nested digital trial within PROTECT Norge will follow. This entails establishing fully digital registration, consent, and screening procedures, updating web page content and study documentation, and scheduling questionnaires for baseline and follow-up assessments for both trial arms. Phase 1 is scheduled for completion in Q4 2027, delivering a fully developed intervention and trial infrastructure in preparation for Phase 2.

Phase 2 will commence in Q4 2027 with implementation of the ASLEEP intervention, recruitment for the RCT, and trial delivery, which will be concluded in 2028 before short-term (3-mo) analysis and subsequent dissemination through stage 2 registered reports. PPI activities have informed the ASLEEP project from its inception, and outcomes of this involvement will be reported.

A dissemination plan will be developed to share the results through relevant channels.

Discussion

Expected Impact

The ASLEEP project represents an innovative, tiered intervention to promote sleep health in midlife and older adults, evaluated as a fully digital, nested RCT within the PROTECT Norge research infrastructure. Sleep health is a multidimensional construct encompassing adequate quality, timing, duration, and regularity of sleep, essential for overall health and well-being [39]. By targeting insomnia symptoms in older adults, ASLEEP aims to address a highly prevalent problem in midlife and older adulthood [3] with significant negative implications for quality of life, physical and mental health [4]. The tiered design has the potential to empower individuals with the capacity and interest to take an active role in prevention, treatment, and follow-up, supporting greater self-management of their sleep health.

Grounded in CBT-I, the project will build further on an existing basic course, which delivers foundational psychoeducation, and

XSL•FO
RenderX

will develop an advanced course for individuals experiencing persistent insomnia symptoms. This advanced course will be designed for delivery via a dedicated app or web platform, ensuring accessibility across devices and integration into participants' daily routines.

The ASLEEP intervention will be evaluated in a nested clinical trial conducted entirely online through the PROTECT Norge platform, using a 2-arm, waitlist-controlled RCT design. The PROTECT Norge help desk will provide technical support, reminders, and study-related communication to minimize barriers to participation and enhance engagement. Embedding the trial within PROTECT enables efficient recruitment, centralized data collection, and long-term follow-up while supporting broad participation from across Norway, including rural and remote areas.

The RCT will evaluate both short- and long-term effects on insomnia symptoms and examine potential benefits for medication use, mental health, cognition, and brain health.

Patient and Public Involvement

PPI has been integral to the ASLEEP project so far, with WiseAge panel input and user representatives embedded in the team to strengthen relevance and usability. These measures have shaped existing study materials, intervention design, and platform development and will continue in line with SESAM's PPI guidelines going forward. What worked well and what did not will be reported, so others can learn from this experience.

Limitations

Although the digital format of the ASLEEP study enables wide reach, adherence to treatment recommendations may be a challenge, as fully automated digital CBT-I programs have shown substantial variation in completion and adherence rates [40]. To address potential challenges with adherence, retention, and data completeness, a 20% dropout rate has been accounted for, and the PROTECT Norge help desk will provide proactive and continuous technical and procedural support, reminders, and guidance to encourage participant engagement throughout the trial.

Recruitment of 400 participants for randomization is ambitious given the current PROTECT Norge research willing cohort of approximately 3500 participants. While this represents a substantial recruitment pool, achieving the target sample size may be challenging; however, the ongoing expansion of the cohort toward 9000 registered participants, in line with PROTECT Norge's established recruitment plan, is expected to strengthen recruitment feasibility.

Participation in PROTECT Norge requires a baseline level of digital competence, which supports delivery of this fully digital intervention but may limit generalizability to the broader population aged 50+ years, particularly those with less confidence in technology or internet use. However, a central element of the ongoing PROTECT Norge cohort expansion is a deliberate effort to broaden cohort diversity through targeted outreach to underrepresented demographic groups, recruitment strategies that extend beyond social media, and partnerships with community organizations and primary care providers. In addition, the PROTECT Norge help desk offers technical support intended to facilitate participation among individuals with lower digital competence. These efforts aim to mitigate selection bias in future phases of the research.

Acknowledgments

We thank Pamela and Tor Cranner for their active contributions as user representatives in the Preventing and Treating Insomnia Symptoms in Midlife and Older Adults (ASLEEP) project. Generative artificial intelligence (ChatGPT, OpenAI; models GPT-5.0, GPT-5.1, and GPT-5.2) was used for proofreading and language editing of this paper.

Funding

The study is funded by the Dam Foundation through the Norwegian Health Association, with additional minor funding support from the Norwegian Competence Center for Sleep Disorders. As a condition of funding from the Dam Foundation, the project is conducted and published as a registered report. Beyond this requirement, the funders had no role in the study design, data collection, data analysis, interpretation of results, writing of the manuscript, or the decision to publish.

Conflicts of Interest

BB is the author of a self-help book on cognitive behavioral therapy for insomnia (CBT-I). The authors declare that this research was carried out without any commercial or financial relationships that could be interpreted as a potential conflict of interest.

Checklist 1
GRIPP 2-SF checklist.
[PDF File, 141 KB - resprot_v15i1e81542_app1.pdf]

Checklist 2
SPIRIT 2025 checklist.
[PDF File, 310 KB - resprot_v15i1e81542_app2.pdf]

XSL•FO
RenderX

References

1. Hirshkowitz M, Whiton K, Albert SM, et al. National Sleep Foundation's sleep time duration recommendations: methodology and results summary. Sleep Health 2015 Mar;1(1):40-43. [doi: 10.1016/j.sleh.2014.12.010] [Medline: 29073412]
2. Institute of Medicine (US) Committee on Sleep Medicine and Research. In: Colten HR, Altevogt BM, editors. Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem: National Academies Press; 2006. URL: https://www.ncbi.nlm.nih.gov/books/NBK19960/ [accessed 2026-02-10]
3. Foley DJ, Monjan AA, Brown SL, Simonsick EM, Wallace RB, Blazer DG. Sleep complaints among elderly persons: an epidemiologic study of three communities. Sleep 1995 Jul;18(6):425-432. [doi: 10.1093/sleep/18.6.425] [Medline: 7481413]
4. Lim DC, Najafi A, Afifi L, et al. The need to promote sleep health in public health agendas across the globe. Lancet Public Health 2023 Oct;8(10):e820-e826. [doi: 10.1016/S2468-2667(23)00182-2] [Medline: 37777291]
5. Hafner M, Stepanek M, Taylor J, Troxel WM, van Stolk C. Why sleep matters - the economic costs of insufficient sleep: a cross-country comparative analysis. Rand Health Q 2017 Jan;6(4):11. [Medline: 28983434]
6. Namsrai T, Ambikairajah A, Cherbuin N. Poorer sleep impairs brain health at midlife. Sci Rep 2023 Feb 1;13(1):1874. [doi: 10.1038/s41598-023-27913-9] [Medline: 36725955]
7. Xie L, Kang H, Xu Q, et al. Sleep drives metabolite clearance from the adult brain. Science 2013 Oct 18;342(6156):373-377. [doi: 10.1126/science.1241224] [Medline: 24136970]
8. Spira AP, Gamaldo AA, An Y, et al. Self-reported sleep and β-amyloid deposition in community-dwelling older adults. JAMA Neurol 2013 Dec;70(12):1537-1543. [doi: 10.1001/jamaneurol.2013.4258] [Medline: 24145859]
9. Sabia S, Fayosse A, Dumurgier J, et al. Association of sleep duration in middle and old age with incidence of dementia. Nat Commun 2021 Apr 20;12(1):2289. [doi: 10.1038/s41467-021-22354-2] [Medline: 33879784]
10. Optimizing brain health across the life course: WHO position paper. World Health Organization. 2022. URL: https://iris.who.int/server/api/core/bitstreams/619fbcac-9fb4-4913-a56c-3be2b55084f2/content [accessed 2026-02-10]
11. Riemann D, Espie CA, Altena E, et al. The European Insomnia Guideline: an update on the diagnosis and treatment of insomnia 2023. J Sleep Res 2023 Dec;32(6):e14035. [doi: 10.1111/jsr.14035]
12. Alvaro PK, Roberts RM, Harris JK. A systematic review assessing bidirectionality between sleep disturbances, anxiety, and depression. Sleep 2013 Jul 1;36(7):1059-1068. [doi: 10.5665/sleep.2810] [Medline: 23814343]
13. Harvey AG. A cognitive model of insomnia. Behav Res Ther 2002 Aug;40(8):869-893. [doi: 10.1016/s0005-7967(01)00061-4] [Medline: 12186352]
14. Perlis ML, Aloia M, Kuhn B, editors. Behavioral Treatments for Sleep Disorders: A Comprehensive Primer of Behavioral Sleep Medicine Interventions: Academic Press; 2011. [doi: 10.1016/C2009-0-62216-9]
15. Riemann D, Baglioni C, Bassetti C, et al. European guideline for the diagnosis and treatment of insomnia. J Sleep Res 2017 Dec;26(6):675-700. [doi: 10.1111/jsr.12594] [Medline: 28875581]
16. Bjorvatn B, Fiske E, Pallesen S. A self-help book is better than sleep hygiene advice for insomnia: a randomized controlled comparative study. Scand J Psychol 2011 Dec;52(6):580-585. [doi: 10.1111/j.1467-9450.2011.00902.x] [Medline: 21790620]
17. Bjorvatn B, Lundetræ RS, Vedaa Ø, Pallesen S, Evanger LN. A randomized controlled trial comparing sleep hygiene advice with a self-help book focusing on cognitive behavioral therapy for insomnia: a study among patients with prescribed hypnotics from the GP. Scand J Prim Health Care 2025 Jun 27:1-10. [doi: 10.1080/02813432.2025.2525423] [Medline: 40577556]
18. Deng W, van der Kleij RMJJ, Shen H, et al. eHealth-based psychosocial interventions for adults with insomnia: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2023 Mar 14;25:e39250. [doi: 10.2196/39250] [Medline: 36917145]
19. Espie CA, Kyle SD, Williams C, et al. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. Sleep 2012 Jun 1;35(6):769-781. [doi: 10.5665/sleep.1872] [Medline: 22654196]
20. Hagatun S, Vedaa Ø, Nordgreen T, et al. The short-term efficacy of an unguided internet-based cognitive-behavioral therapy for insomnia: a randomized controlled trial with a six-month nonrandomized follow -up. Behav Sleep Med 2019;17(2):137-155. [doi: 10.1080/15402002.2017.1301941] [Medline: 28345961]
21. Testad I, Aakre JA, Gjestsen MT, et al. Web-based assessment of cognition and dementia risk factors in over 3000 norwegian adults aged 50 years and older: cross-sectional PROTECT norge study. JMIR Aging 2025 Aug 25;8:e69867. [doi: 10.2196/69867] [Medline: 40854096]
22. Corbett A, Taylor R, Llewellyn D, et al. Impact of vitamin D supplementation on cognition in adults with mild to moderate vitamin D deficiency: outcomes from the VitaMIND randomized controlled trial. J Am Med Dir Assoc 2025 Aug;26(8):105711. [doi: 10.1016/j.jamda.2025.105711] [Medline: 40480279]
23. Corbett A, Williams G, Creese B, et al. Impact of short-term computerized cognitive training on cognition in older adults with and without genetic risk of Alzheimer's disease: outcomes from the START randomized controlled trial. J Am Med Dir Assoc 2024 May;25(5):860-865. [doi: 10.1016/j.jamda.2024.03.008] [Medline: 38642588]
24. PROTECT Norge [Article in Norwegian]. URL: https://www.protect-norge.no [accessed 2026-02-05]
25. Staniszewska S, Brett J, Simera I, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. Res Involv Engagem 2017;3(1):13. [doi: 10.1186/s40900-017-0062-2] [Medline: 29062538]

XSL•FO
RenderX

26. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [doi: 10.1016/s1389-9457(00)00065-4] [Medline: 11438246]
27. Walker J, Muench A, Perlis ML, Vargas I. Cognitive behavioral therapy for insomnia (CBT-I): a primer. Klin Spec Psihol 2022;11(2):123-137. [doi: 10.17759/cpse.2022110208] [Medline: 36908717]
28. Bower P, Gilbody S. Managing common mental health disorders in primary care: conceptual models and evidence base. BMJ 2005 Apr 9;330(7495):839-842. [doi: 10.1136/bmj.330.7495.839] [Medline: 15817554]
29. Steinmetz L, Simon L, Feige B, et al. Network meta-analysis examining efficacy of components of cognitive behavioural therapy for insomnia. Clin Psychol Rev 2024 Dec;114:102507. [doi: 10.1016/j.cpr.2024.102507] [Medline: 39504928]
30. Skivington K, Matthews L, Simpson SA, et al. Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. Health Technol Assess 2021 Sep;25(57):1-132. [doi: 10.3310/hta25570] [Medline: 34590577]
31. Brooker H, Williams G, Hampshire A, et al. FLAME: a computerized neuropsychological composite for trials in early dementia. Alzheimers Dement (Amst) 2020;12(1):e12098. [doi: 10.1002/dad2.12098] [Medline: 33088895]
32. Roth T, Zammit G, Kushida C, et al. A new questionnaire to detect sleep disorders. Sleep Med 2002 Mar;3(2):99-108. [doi: 10.1016/s1389-9457(01)00131-9] [Medline: 14592227]
33. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
34. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
35. World Medical Association. World Medical Association declaration of Helsinki: ethical principles for medical research involving human participants. JAMA 2025 Jan 7;333(1):71-74. [doi: 10.1001/jama.2024.21972] [Medline: 39425955]
36. Beresford P, McLaughlin H. Service user involvement in research: what difference does it make? In: McLaughlin H, Beresford P, Cameron C, Casey H, Duffy J, editors. The Routledge Handbook of Service User Involvement in Human Services Research and Education: Routledge; 2020:515-521 URL: https://www.taylorfrancis.com/chapters/edit/10.4324/9780429433306-54/service-user-involvement-research-peter-beresford-hugh-mclaughlin [accessed 2026-02-10]
37. Kujala S. User involvement: a review of the benefits and challenges. Behav Inf Technol 2003 Jan;22(1):1-16. [doi: 10.1080/01449290301782]
38. SESAMs veileder for brukermedvirkning i forskning [Report in Norwegian]. : SESAM; 2023 URL: https://www.helse-stavanger.no/4a4993/siteassets/avdelinger/sesam/documents/sesam_veileder_for_brukermedvirkning_original.pdf [accessed 2026-01-30]
39. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 1;37(1):9-17. [doi: 10.5665/sleep.3298] [Medline: 24470692]
40. Hwang JW, Lee GE, Woo JH, Kim SM, Kwon JY. Systematic review and meta-analysis on fully automated digital cognitive behavioral therapy for insomnia. NPJ Digit Med 2025 Mar 12;8(1):157. [doi: 10.1038/s41746-025-01514-4] [Medline: 40075149]

Abbreviations

ANCOVA: analysis of covariance
ASLEEP: Preventing and Treating Insomnia Symptoms in Midlife and Older Adults
ATE: average treatment effect
C4C: consent for contact
CBT-I: cognitive behavioral therapy for insomnia
GDPR: General Data Protection Regulation
GRIPP 2-SF: Guidance for Reporting Involvement of Patients and the Public–Short Form
ICH: International Council for Harmonisation
ISI: Insomnia Severity Index
MRC: Medical Research Council
PPI: patient and public involvement
PROTECT Norge: Platform for Research Online to Investigate Cognition and Genetics in Ageing Norge
RCT: randomized controlled trial
REK: Regional Committee for Medical and Health Research Ethics
SESAM: Centre for Age-Related Medicine
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
UAT: user acceptance testing

XSL•FO
RenderX

Edited by J Sarvestan; submitted 13.Sep.2025; peer-reviewed by A Bramoweth; revised version received 16.Jan.2026; accepted 20.Jan.2026; published 02.Mar.2026.

Please cite as:
Aakre JA, Bjorvatn B, Gjestsen MT, Dalen I, Ballard C, Testad I
Preventing and Treating Insomnia Symptoms in Midlife and Older Adults (ASLEEP): Protocol for a Randomized Controlled Trial Using the PROTECT Norge Infrastructure
JMIR Res Protoc 2026;15:e81542
URL: https://www.researchprotocols.org/2026/1/e81542
doi:10.2196/81542

© Jon Arild Aakre, Bjørn Bjorvatn, Martha Therese Gjestsen, Ingvild Dalen, Clive Ballard, Ingelin Testad. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of a Community-Based AI-Assisted Visual Impairment Screening Model for Performance, Operational Efficiency, Acceptability, Feasibility, and Costs: Protocol for a 2-Arm Pragmatic Randomized Controlled Trial

Yibing Chen[1,2], BEng; Kai Hui Koh[3], BSc; Jin Wei Clarine Ho[3], SDCO; Samantha Min Er Yew[1,2], BSc; Jocelyn Hui Lin Goh[1,2,3], BEng; Elaine Lum[4], PhD; Yih Chung Tham[3], PhD

[1]Centre for Innovation and Precision Eye Health, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[2]Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[3]Singapore Eye Research Institute, Singapore National Eye Centre, The Academia, 20 College Road, Discovery Tower Level 6, Singapore, Singapore

[4]Centre for Population Health Research & Implementation, SingHealth, Singapore, Singapore

Corresponding Author:
Yih Chung Tham, PhD
Singapore Eye Research Institute, Singapore National Eye Centre, The Academia, 20 College Road, Discovery Tower Level 6, Singapore, Singapore

Abstract

Background: Visual impairment (VI) affects more than 600 million people globally and significantly reduces quality of life. In Singapore, 20% of adults aged 60 years and older (~180,000 people) have VI, a figure expected to double by 2030 due to population aging. While about half of VI cases are due to uncorrected refractive errors, the rest are caused by age-related diseases. The current traditional screening model is a 2-visit, labor-intensive approach with low follow-up rates and frequent unnecessary referrals. Although AI for Disease-related Visual Impairment Screening Using Retinal Imaging, the deep learning model in this study, has demonstrated strong diagnostic performance in retrospective datasets (area under the curve=0.942), key aspects of real-world implementation such as operational efficiency, patient acceptability, workflow feasibility, and cost remain insufficiently studied. As a result, real-world evidence directly comparing artificial intelligence (AI)–assisted and traditional screening pathways is limited.

Objective: This study aims to evaluate the referral accuracy, operational efficiency, acceptability, feasibility, and cost of an AI–assisted screening model compared with the current traditional screening model.

Methods: This study aims to recruit 1000 participants aged 50 years and older using a 2-arm pragmatic randomized controlled trial design. Participants with presenting visual acuity worse than 6/12 (L2) will be randomized 1:1 into either the AI-assisted or traditional screening arms. In the AI-assisted arm, the AI model will analyze retinal photos on-site, with positive cases referred to an optometrist for secondary evaluation. The AI model, previously developed with promising diagnostic accuracy and further validated using community-acquired data, has been integrated with a custom user interface for use in this study. Traditional screening will include pinhole visual acuity, intraocular pressure, slit lamp examination, auto refraction, and retinal photography. All L2 participants will complete a patient-acceptance questionnaire and undergo assessments to determine ground truth.

Results: The study was funded in 2022. Participant recruitment commenced in July 2024, with 487 participants enrolled as of September 14, 2024. Recruitment is ongoing, with study completion anticipated by March 2026 and data analysis expected to begin in April 2026.

Conclusions: This study will provide critical evidence on the clinical utility, feasibility, and cost analysis of AI-assisted VI screening. Our findings may contribute real-world evidence to inform scalable, sustainable screening strategies that enhance efficiency, accuracy, and health system outcomes.

Trial Registration: ClinicalTrials.gov NCT06877988; https://clinicaltrials.gov/study/NCT06877988

International Registered Report Identifier (IRRID): DERR1-10.2196/74164

(JMIR Res Protoc 2026;15:e74164) doi:10.2196/74164

KEYWORDS

artificial intelligence; visual impairment; pragmatic randomized controlled trial; retinal imaging; prospective studies

XSL•FO
RenderX

Introduction

Background and Rationale

Visual impairment (VI) affects approximately 600 million individuals globally and significantly reduces quality of life, independence, and mobility [1]. In Singapore, one-fifth of adults aged 60 years and older (~180,000 people) experience VI [2]. This figure is expected to nearly double by 2030 due to population aging [1]. Among the VI cases, approximately half result from uncorrected refractive errors, primarily myopia, which can be readily corrected with properly dispensed spectacles available at optical shops. The remaining half are caused by age-related ocular diseases, including cataract, diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma, requiring specialized medical or surgical intervention [3]. Alarmingly, around 70% of older adults with VI in Singapore remain undiagnosed, primarily due to low awareness [4]. Findings from the Singapore Epidemiology of Eye Diseases (SEED) study highlighted that 80% of VI cases among the older adults are treatable [2]. Timely detection through effective vision screening and prompt identification of underlying causes are crucial to prevent disease progression and severe visual loss [5]. Thus, primary prevention and early intervention are essential to mitigate the burden of VI.

The traditional vision screening model commonly used in Singapore involves a 2-visit process. During the first visit, visual acuity (VA) and pinhole tests are conducted, and participants with poor pinhole VA (worse than 6/12) are referred for a second visit on a separate day for additional slit lamp and retinal examinations [6]. While this model provides a structured and thorough screening process, it is relatively manpower-intensive and may require additional effort from participants, potentially impacting follow-up rates and service capacity. As demand for vision screening grows with an aging population, exploring more streamlined and scalable approaches may help enhance accessibility, efficiency, and long-term sustainability.

Artificial intelligence (AI) presents transformative possibilities for health care delivery [7]. Recently, we developed AVIRI (AI for Disease-related Visual Impairment Screening Using Retinal Imaging), a novel deep learning-based algorithm leveraging retinal photography to detect disease-related VI. In a proof-of-concept study published in *Lancet Digital Health* (2021), AVIRI achieved high accuracy (area under the curve [AUC] 0.942) in identifying eye disease-related VI compared with ophthalmologist diagnoses [8]. Although many AI models have shown promising performance on retrospective evaluation, real-world implementation required more than diagnostic accuracy. Successful adoption also depends on workflow feasibility, efficiency, patient acceptability, and cost [9-11]. These aspects of real-world implementations have not been studied in previous AI research. To that end, real-world validation of AI-assisted screening remains limited, with even fewer studies comparing between AI-assisted and traditional screening approaches. To address this gap, we propose an AI-assisted screening model that combines a presenting VA test with AVIRI and will directly compare its performance against traditional screening methods to evaluate efficiency and accuracy in detecting VI.

Hypothesis and Objectives

This study is designed as a hybrid type 1 effectiveness-implementation trial to evaluate and compare the performance, operational efficiency, acceptability, feasibility, and cost differences between the AI-assisted screening model and the traditional screening model in a community screening setting. A pragmatic randomized controlled trial (pRCT) design will be used to better represent the targeted real-world population. This study will focus on the following specific aims:

Aim 1: To evaluate and compare the referral accuracy between the AI-assisted screening model and traditional screening model for disease-related VI.

Hypothesis 1: The AI-assisted screening model will outperform the traditional screening model in detecting and referring VI cases.

Aim 2: To compare the operational efficiency and acceptability between the AI-assisted screening model and the traditional screening model, and to investigate the feasibility of the AI-assisted screening model.

Hypothesis 2: The AI-assisted screening model will demonstrate greater operational efficiency and acceptability compared to the traditional screening model and will be feasible for implementation in community-based settings.

Aim 3: To compare the overall costs of the AI-assisted screening model and the traditional screening model from a health system perspective.

Hypothesis 3: The AI-assisted screening model will be more cost-effective than the current traditional screening model from a health system perspective.

Methods

AI Model Development and Validation

In a proof-of-concept study, AVIRI demonstrated high diagnostic performance, achieving an area under the receiver operating characteristic curve (AUC) of 0.942 (95% CI 0.922 - 0.956), with an accuracy of 87.7%, sensitivity of 94.6%, and specificity of 81.3%, compared with the ground truth of ophthalmologists' clinical diagnoses. External validation using population-based datasets from China, India, and Australia yielded AUCs ranging from 0.859 to 0.935, further supporting AVIRI's robustness across different populations [8]. While these findings highlight the model's ability to detect disease-related VI, its real-world performance in community-based settings with undilated eye examinations has not yet been explored.

To address this gap, a previous retrospective evaluation was conducted using 421 retinal images collected in a community setting with the Crystalvue retinal camera (Crystalvue NFC-700; Crystalvue Medical Corporation). This evaluation was part of pilot testing and was not previously published. Multiple domain adaptation strategies were applied to enhance model generalizability, and a ViT-Base architecture, trained on the

XSL•FO
RenderX

SEED dataset, was selected as the optimal backbone. This adaptation achieved an AUC of 0.853, specificity of 0.688 (when sensitivity fixed at 80%), and F_1 score of 0.347. Additionally, a user interface was developed to facilitate clinical integration, allowing researchers to input clinical data and generate automated diagnostic reports using AVIRI. These findings provide the foundation for further prospective validation, evaluating AVIRI's real-world performance in an AI-assisted vision screening model.

Sample Size Calculation

This study will recruit a total of 1000 participants aged 50 years and older using a 2-arm pRCT design with broad inclusion criteria based solely on age. Of these, 500 participants are targeted to have presenting VA better than 6/12 (L1), a number determined based on operational feasibility in community screenings.

Participants with presenting VA worse than 6/12 (L2) will be randomized to either the AI-assisted or traditional screening model. Sample size estimation for this randomized comparison was based on an assumed accuracy of 70% for the current traditional screening model and prevalence of 18% for any VI, 11% for disease-related VI, and 8% for refractive-error related VI (prevalence estimates were based on the SEED study) [2]. Detecting a clinically meaningful difference of 10% between the 2 screening models with 80% statistical power at a 5% 2-sided type I error requires 250 participants per arm. Therefore, a total of 500 randomized participants (L2) will be included in the primary analysis.

Randomization

VI is defined according to the United States VI definition as having a presenting VA worse than 20/40 (equivalent to 6/12) in the better-seeing eye [3]. Participants with presenting VA of 6/12 or better (L1) will not be randomized. These participants will receive a basic ocular assessment, including fundus photography and anterior eye imaging. The collected images will not be used in the comparative analysis between the AI-assisted and traditional screening arms.

Participants with presenting VA worse than 6/12 (L2) will be randomized into either the AI-assisted screening arm or the traditional screening arm using a 1:1 allocation ratio. Randomization will be conducted at the patient level, ensuring that both eyes of the same participant are assigned to the same study arm. The randomization sequence will be generated by an independent statistician using the "randomizeR" package in R statistical software. This approach ensures an unbiased allocation process while maintaining the integrity of the study design.

Study Setting

This study will be conducted at Pioneer Polyclinic, a public primary care clinic located in the western region of Singapore. Polyclinics in Singapore are government-funded health care institutions that provide subsidized outpatient medical services, including acute care, chronic disease management, health screenings, and preventive care. They serve as a key component of the country's primary health care network and are typically staffed by physicians, nurses, and allied health professionals.

Pioneer Polyclinic is part of the National University Polyclinics (NUP) group under the National University Health System. It serves a multiethnic residential population comprising predominantly older adults and middle-aged individuals, with a substantial proportion managing chronic conditions such as hypertension, diabetes, and hyperlipidemia. The clinic is located in a mature public housing estate, where access to subsidized screening and community-based care is particularly important.

Data Collection

Study Workflow

Recruitment and Study Procedures

Recruitment will take place at Pioneer Polyclinic (National University Polyclinics) over an 18-month period. Each participant will attend a single study visit, which is expected to last approximately 30 to 90 minutes. Participants will be recruited through 3 main channels: (1) review of past study records to identify eligible individuals, (2) word-of-mouth referrals, and (3) display of an electronic brochure (e-brochure) on screens within participating polyclinics. For the first recruitment channel, only individuals who meet the study's eligibility criteria will be invited to participate. Eligible participants must have complete clinical data and retinal imaging records. They must also have previously provided consent to be recontacted.

Both the AI and traditional screening arms will use the same fundus cameras (Crystalvue) and will follow identical equipment calibration protocols according to manufacturer recommendations. All image acquisition will be performed by trained clinical research coordinators who have completed standardized certification in retinal imaging. Refresher training and inter-operator assessments will be conducted prior to and throughout the study to minimize variability.

Control Arm

Participants will first sign informed consent before proceeding with any assessment. After being randomized into the traditional model arm, the screening will follow the standard examination workflow used in the traditional screening model, which consists of pinhole VA, intraocular pressure, auto refraction, slit lamp examination, and retinal photography. The fundus photographs will be reviewed by the onsite optometrists, and any images of insufficient quality will be retaken immediately to improve clarity. If image quality remains inadequate, the participant's ID will be flagged in the study log. Referral decisions will be made by the onsite optometrists, incorporating the findings from all other screening tests. After receiving the referral outcome, participants will be invited to complete a patient-acceptance questionnaire to assess their perceptions and satisfaction with the traditional screening model (Multimedia Appendix 1). Finally, anterior eye photography and subjective refraction will be performed to establish the ground truth status for the referral decision (Figure 1).

XSL•FO
RenderX

Figure 1. The data collection workflow. [a]Data to establish the gold standard. AI: artificial intelligence; AVIRI: AI for disease-related visual impairment screening using retinal imaging; IOP: intraocular pressure; VA: visual acuity.

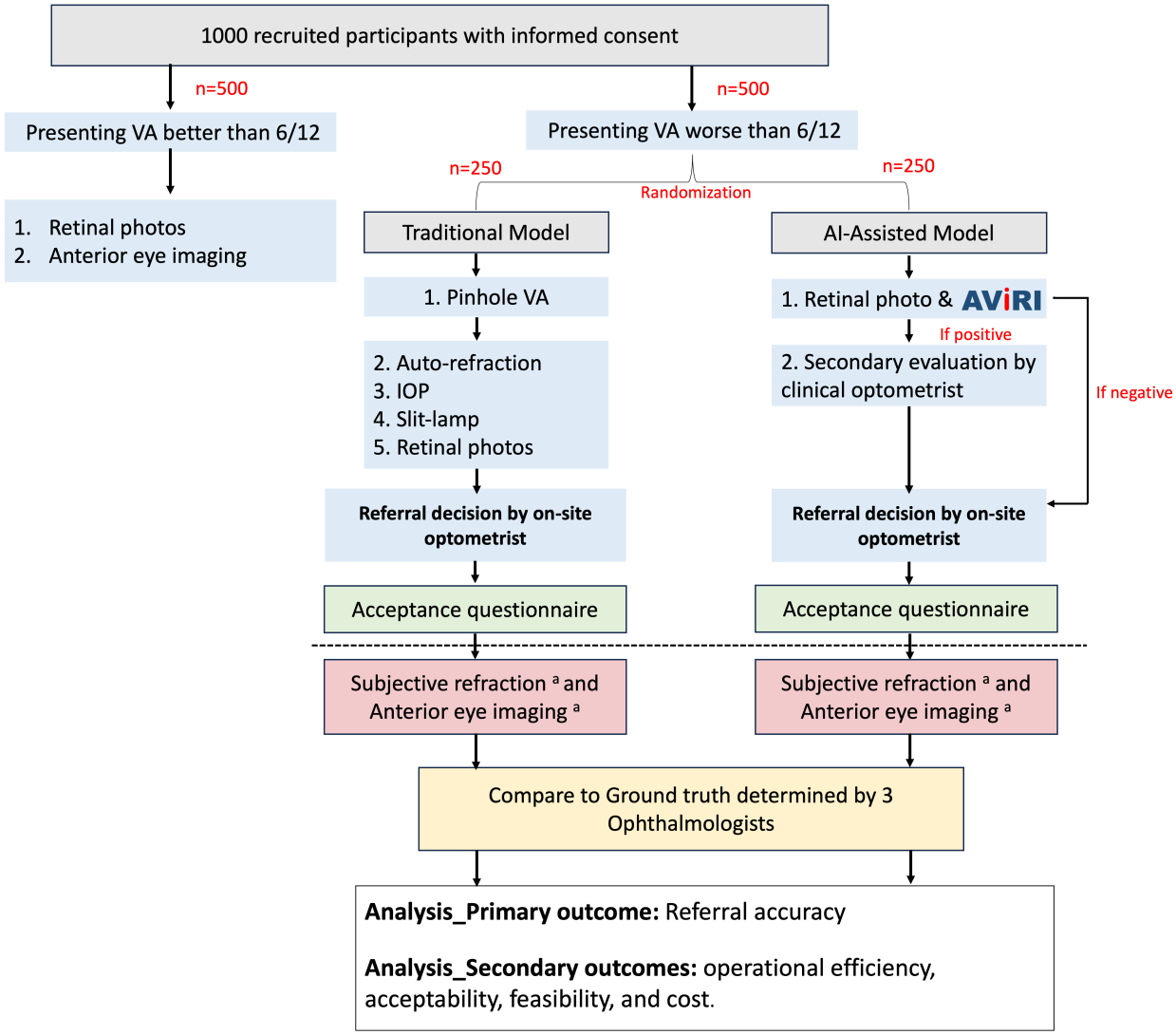

Intervention Arm

Patients randomized to the AI-assisted arm will undergo retinal photography, with their images analyzed on-site by AVIRI. Fundus images will be checked by onsite optometrists, with poor-quality photographs retaken on the spot. If adequate quality still cannot be obtained, the participant's ID will be recorded in the study log. The AVIRI report will also include a quality-check notification indicating when an image does not pass its internal quality assessment, alerting the clinical optometrists that the AI-generated results may be affected. If a case is detected by AVIRI as having disease-related VI, the retinal images and AVIRI report will be sent to an assigned clinical optometrist via the institution's secured Microsoft Teams for further evaluation (Microsoft Corp). Clinical optometrists, who are independent from the onsite screening team, will examine the AI-generated results and make the final referral decision. Only cases identified as positive by the AVIRI model are reviewed by the clinical optometrists, and the referral process follows the same standardized procedure applied across both study arms. Similarly, after the referral decision, participants will complete a patient-acceptance questionnaire to assess their perceptions and satisfaction with the AI-assisted model. Finally, anterior eye photography and subjective refraction will be performed to establish the ground truth status (Figure 1).

Data Quality Assurance

A senior data coordinator will oversee all data entry procedures to ensure consistency and accuracy throughout the process. Each data entry will be independently verified by a separate data entry personnel to minimize errors and discrepancies. Regular checks will be conducted to ensure that all consented, eligible participants' data have been properly entered into the system. Additionally, basic tabulation will be performed against the study screening log, and any inconsistencies will be promptly addressed to maintain the integrity of the data. To ensure consistent image quality, all imaging personnel will undergo standardized training and certification in retinal photography. Fundus cameras will be calibrated according to manufacturer guidelines. Images will be captured following a standardized protocol for field of view, focus, and illumination and reviewed in real time. Ungradable images will be repeated onsite and

XSL•FO
RenderX

flagged in the data log. Image quality will be graded as good, borderline, or poor. A detailed review process will be implemented to guarantee the accuracy and completeness of the dataset before analysis.

Data Entry and Storage

Data will be recorded on paper forms and subsequently entered into a password-protected Excel spreadsheet stored on an encrypted storage drive. Slit lamp and retinal photographs will be deidentified and stored in digital format for grading purposes. A digital audio recorder will be used during the patient acceptance questionnaire to facilitate accurate documentation, analysis, and review of discussions. A distinct section in the consent form will be used to obtain participants' permission for audio recording their responses. The recorded audio data will be transcribed and translated into English (if required) by study team members after the participant's visit. The data entry process will be supervised by a data coordinator, and all entered data will be independently verified by another data entry personnel.

All data files will be encrypted to ensure data security. Access keys for encryption will be securely stored and made available only to study team members. Data access and key management will adhere to Singhealth's data retention policy. Case files will be securely stored with restricted access, in accordance with the study delegation log.

Participants' confidentiality and privacy will be upheld throughout the study, from the initial data collection phase until the secure destruction of personal and sensitive information. Only authorized research team members will have access to research data.

Outcome Measures

Ground Truth Establishment

The ground truth diagnosis will be determined by a panel of 3 board-certified ophthalmologists, who will independently review presenting VA, best-corrected VA data, intraocular pressure measurements, medical history, and anterior and retinal photographs. The reference standard outcome (refer vs not refer) will be determined by majority agreement (2 out of 3 panel experts). Cases without a clear majority (eg, ungradable) will undergo consensus discussion among the experts, and if consensus cannot be reached, a senior ophthalmologist will adjudicate. The adjudicator will not be blinded by the clinical context as it is required to make a referral decision. To ensure the credibility of the pRCT results, the panel of three board-certified ophthalmologists will be blinded to the participants' group assignments to minimize bias, and they will not participate in any study procedures that could compromise or reveal participants' randomized group assignments.

Primary Outcomes

The primary outcome is the rate of correct referrals based on the reference standard. A correct referral was defined as a case in which the AI or optometrist's referral decision (refer vs no refer) matched the ground truth outcome determined by the panel of three board-certified ophthalmologists. To evaluate whether the AI-assisted model offers improved referral accuracy compared to the traditional screening model, performance metrics such as accuracy, AUC, sensitivity, and specificity will be calculated and compared for both models. A calibration plot will be used to assess the agreement between the model's predicted probability of VI and the actual VI outcomes. Performance metrics will also be stratified and analyzed according to the severity of VI. Moderate VI is defined as presenting VA between 6/12 and 6/48, while severe VI is defined as presenting VA of 6/60 or worse [12].

Secondary Outcomes

The secondary outcomes of this pRCT include operational efficiency, acceptability, perceived feasibility, and cost evaluation. These secondary outcomes are important for successful implementation and future scalability.

Operational Efficiency

This evaluation will assess the effectiveness of the screening workflow by measuring the average time required per patient to complete the entire screening process, providing insights into overall efficiency. The screening capacity will be evaluated by calculating the average number of patients screened per hour. In addition to the average number of patients screened per hour and average time per patient, we will break down the total screening time into specific workflow steps (eg, retinal photo taking, results analysis, review, and referral decision making) to allow a more detailed comparison of operational efficiency between the traditional and AI-assisted arms. These metrics will help determine whether the AI-assisted model improves workflow efficiency compared to the current screening approach.

Acceptability

In this study, acceptability is defined as patients' perception of the AI-assisted screening model in terms of its usability, comfort, and overall satisfaction [13]. All participants randomized to either the AI-assisted or traditional screening arm will be invited to complete a structured questionnaire designed to assess their perception and satisfaction with the screening process. The questionnaire consists of 3 items rated on a 5-point Likert scale, evaluating the overall satisfaction with the screening process [14]. Five dichotomous questions are also included to evaluate factors such as awareness of AI and confidence in results. The questionnaire was developed following established psychometric principles to ensure validity and reliability in assessing patient perspectives [15].

To enhance the accuracy of data collection, trained research coordinators will administer the questionnaire with audio recording after obtaining patients' consent. The responses will be transcribed verbatim. This approach ensures consistency in responses, minimizes misinterpretation of questions, and allows coordinators to clarify any concerns raised by participants. Missing questionnaire items will be reported as nonresponses. Given the short questionnaire and supervised administration, minimal missing data are expected. However, sensitivity analyses will be conducted if missingness exceeds 5%. Insights gained from this assessment will help refine the implementation of AI-assisted screening models and address potential barriers to patient acceptance.

XSL•FO
RenderX

Feasibility

Feasibility is defined as the degree to which the AI-assisted screening model can be effectively implemented within the intended setting [13]. Service providers' perceptions of feasibility will be assessed through focus group discussions. The results from evaluations of operational efficiency and patient acceptability, as described earlier, will inform the development of discussion questions based on the constructs of the Consolidated Framework for Implementation Research (CFIR), a well-established framework for assessing the implementation of health care innovations. The CFIR provides a structured approach for understanding the factors that influence successful adoption and integration of new interventions within health care settings [16].

Focus group discussions will be conducted with optometrists, each lasting 60 to 90 minutes, with the objective of achieving both data saturation and meaning saturation [17]. All seven clinical optometrists who reviewed the AVIRI output will be invited to join the focus group discussions. The FGD will enable us to capture diverse perspectives and stakeholder experiences with the AI workflow. The focus group discussions will be audio-recorded, transcribed verbatim, and analyzed using framework analysis [18]. Two trained researchers will conduct both inductive and deductive coding, with a third researcher moderating if discrepancies arise after a consensus discussion. The analysis will be interpreted through the lens of the CFIR framework [16]. This in-depth approach will help identify barriers and enablers to the implementation of the AI-assisted screening model from the service providers' perspectives, facilitate the development of meaningful adjustments to the model, and uncover potential strategies for improving future implementations.

Cost Evaluation

We will assess the incremental cost savings of implementing the AI-assisted screening model compared to the traditional screening model using an Activity-Based Costing (ABC) approach. The analysis will include all nonsunk costs incurred up to the referral decision. These costs include labor, materials, supplies, and amortized expenses for technology and space use. Costs will be grouped according to the key activities of each screening model, such as conducting the screening examinations, running the AVIRI algorithm, and generating on-site test results. Fixed costs will be amortized based on the expected useful life of the involved fixed assets. For clinical assessment costs, we will use nonsubsidized billing rates, as these are expected to closely reflect actual costs [19]. As the study involves only a one-time screening visit without longitudinal follow-up, all cost calculations will reflect a cross-sectional time frame, limited to the immediate costs incurred during the single-visit screening encounter.

Data Analysis Plan

Model Performance

The AI model was based on the ViT-base backbone, which was initialized using pretrained weights obtained through supervised learning on the ImageNet-21k dataset [20]. Fine-tuning was carried out for 40 epochs to allow for sufficient learning while reducing the risk of overfitting. The process was performed using a batch size of 16, with a fixed learning rate of 1×10 , weight decay set at 5×10^{4} , and the default drop path rate applied. To evaluate the effectiveness of the AI-assisted screening model compared to the traditional screening model in detecting VI, both models will be assessed against the ground truth. Overall accuracy, sensitivity, and specificity will be calculated and compared using *z* tests for 2 proportions. Additionally, AUC and the area under the precision-recall curve will be determined for each model. Differences in AUC and area under the precision-recall curve between the two models will be analyzed using bootstrap resampling. All quantitative analyses for model performance will be conducted using R, including diagnostic accuracy metrics and comparative statistical tests.

Acceptability and Operational Efficiency

The acceptance scores, measured as ordinal data, will be compared between the 2 screening models using the Mann-Whitney *U* test. To assess operational efficiency, the average time taken per participant for the entire screening process (efficiency assessment) and the average number of individuals screened per hour (capacity assessment) will be calculated [21]. The time taken for each station will also be assessed. Differences in procedure time and screening volume between the 2 models will be analyzed using *t* tests for mean comparisons. To provide additional insight, subgroup analyses by age and gender will be conducted to further evaluate participant acceptability. Age-based analyses will categorize participants into two groups: those aged 50 years and older, and those aged 70 years and older. Gender-based analyses will compare outcomes between male and female participants. Quantitative data related to acceptability and operational efficiency will be analyzed using R, and qualitative data will be coded and thematically analyzed using NVivo.

Cost Evaluation

A decision tree model will be employed to estimate the net costs and potential cost savings of the AI-assisted screening model compared with the traditional screening model. The model starts with a decision node for the screening method, followed by chance nodes incorporating VI prevalence and the probability of a positive or negative referral decision based on the sensitivity and specificity of each screening approach. Costs will be incorporated at the level of each screening arm to capture the resources required, including labor, equipment, and operational expenses associated with a single screening visit. Downstream costs such as treatment or follow-up are not included, as they are assumed to be equivalent across both arms and therefore do not impact the incremental cost comparison."

Key model parameters will include VI prevalence, screening-related costs, and probability outcomes. The primary outcome will be the patient-level cost incurred under each screening model, restricted to procedures performed up to the referral decision. One-way sensitivity analyses will be conducted on key parameters to assess the robustness of the findings. The underlying hypothesis is that the AI-assisted model is at least as effective as the traditional screening model (Figure 2).

XSL•FO
RenderX

Figure 2. The decision tree structure for cost evaluation. AI: artificial intelligence; Cost C_{AI}: Total screening cost (per patient) in the AI-assisted arm. Cost C_{Trad}: Total screening cost (per patient) in the traditional arm. P_0 = VI prevalence: Probability that a participant has visual impairment. P_1 = P(TP_A): Probability of true positive outcome in AI-assisted arm. P_2 = P(TN_A): Probability of true negative outcome in AI-assisted arm. P_3 = P(TP_T): Probability of true positive outcome in traditional arm. P_4 = P(TN_T): Probability of true negative outcome in traditional arm; VI: visual impairment.

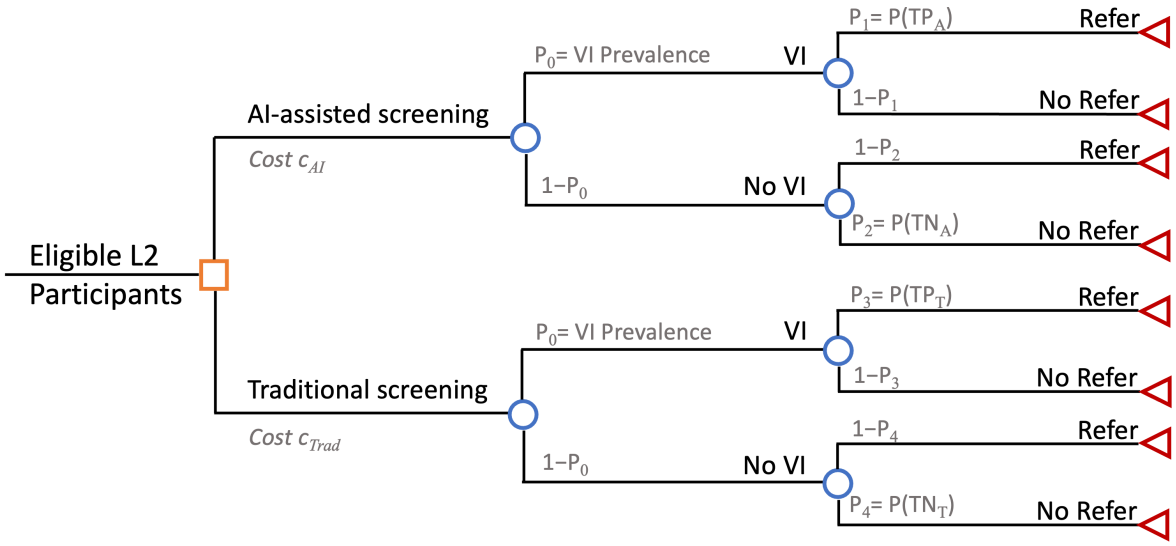

Ethical Considerations

Written informed consent will be obtained from all participants prior to examination, with a witness present, in accordance with the Declaration of Helsinki. Trained research staff certified in Good Clinical Practice (GCP) will administer the consent process. For participants who do not speak English, the consent document will be provided, and the study procedures will be explained in their preferred language (eg, Malay, Hindi, Tamil, or Mandarin) by study staff. An impartial witness, not involved in the study, will also be present to ensure transparency and comprehension. This study will be conducted in accordance with the Declaration of Helsinki, and ethical approval has been obtained from the SingHealth Centralized Institutional Review Board (CIRB) (2024/2297). Any amendments to the study protocol will be submitted for IRB approval. Data confidentiality will be strictly upheld, and no personal identifiers will be included in any study reports or publications. The study is registered on ClinicalTrials.gov under the registration identifier NCT06877988. L1 participants received an SGD 10 (US $7.90) voucher, while L2 participants received an SGD 30 (US $23.70) voucher as compensation for their participation.

Results

This study is funded by the National Medical Research Council's (NMRC) HPHSR Clinician Scientist Award (INV category), awarded in 2022. Participant recruitment and data collection commenced in July 2024. As of September 14, 2024, a total of 487 participants had been recruited. These figures reflect the recruitment status at the time of protocol submission. Recruitment is ongoing, with study completion anticipated by March 2026. Data cleaning and statistical analysis are expected to commence in April 2026, followed by manuscript preparation planned for June 2026.

Discussion

This study aims to evaluate the effectiveness, feasibility, and cost of an AI-assisted VI screening model compared with the traditional screening model in a real-world community setting. Findings from this work are expected to provide useful insights into the integration of AI tools in primary eye care, particularly regarding operational efficiency, patient acceptability, and economic sustainability. This is especially important given the limited real-world prospective evaluations of AI models in community-based screening. The evidence generated will help address this gap and support future policy and implementation decisions.

This study design offers several strengths that enhance its rigor and potential impact. It adopts a 2-arm pRCT design, increasing the relevance and applicability of findings to routine clinical practice. Most prospective studies involving AI in ophthalmology have focused primarily on evaluating the performance of the AI model itself, with few evaluating the actual workflow and adopting a 2-arm RCT that allows a direct comparison between an AI-assisted screening model and the traditional screening model. For example, Rao et al [22] conducted a prospective study on AI-based detection of referable glaucoma, but the investigation was limited to assessing the AI model's performance in a research setting. Similarly, Antaki et al [23] prospectively validated an AI model for detecting diabetic retinopathy, yet their study also focused exclusively on AI performance without evaluating its workflow integration or comparing with traditional screening model. In contrast, the head-to-head comparison in our pRCT provides a more comprehensive evaluation of the value-add of AI model in

XSL•FO
RenderX

real-world eye screening setting [1] Standardized deployment of the AVIRI model and structured data collection protocols support reproducibility. Additionally, the integration of both quantitative metrics and qualitative feedback, including patient satisfaction and acceptability, provides a comprehensive assessment of model performance and implementation feasibility.

While the pRCT design strengthens internal validity, model-level factors also play a role in shaping screening outcomes in practice. Model performance may also be influenced by the model settings, such as the threshold used to detect the disease-related VI. Adjusting the threshold affects the balance between sensitivity, specificity, and the resulting referral decision. Different community settings may have varying levels of follow-up capacity, workforce availability, and tolerance for false positives or false negatives. Recognizing the impact of these settings is important for interpreting study findings and ensuring that the AI model can be appropriately adapted for real-world implementation.

Limitations

Despite these strengths, challenges may arise during the implementation process. As a single-site study, generalizability to other health care settings may be limited. Recruitment of participants could pose logistical difficulties, potentially affecting sample size and study power. Variability in image quality and real-world operational constraints could affect model performance. Additionally, assessing feasibility and cost within the study duration may present constraints, requiring assumptions for long-term projections.

To address these challenges, several mitigation strategies will be employed. While generalizability is limited by the single-site design, detailed documentation of the study context will support future multisite applications. Recruitment efforts will be supported through multiple outreach channels and flexible scheduling to enhance participation. If these challenges are appropriately managed, this study will provide critical evidence on the potential role of AI in community-based VI screening and inform future large-scale implementation efforts.

Conclusion

This study protocol will provide critical insights into the performance, feasibility, and cost of an AI-assisted VI screening model in community settings. By demonstrating its potential advantages over the traditional screening approach, including improved referral accuracy, enhanced operational efficiency, and reduced health care costs, the study outcomes will inform future strategies for implementation. Ultimately, the findings are expected to inform future implementation strategies and contribute to the evidence base for scalable and sustainable approaches to VI screening.

Acknowledgments

Generative artificial intelligence was not used in the generation of this manuscript. The authors acknowledge the use of ChatGPT to assist with improving the clarity and readability of the manuscript. All text was subsequently reviewed, edited, and approved by the authors, who take full responsibility for the content of this publication.

Funding

This study will be supported by the National Medical Research Council's (NMRC) HPHSR Clinician Scientist Award (INV category). YCT is supported by the National Medical Research Council of Singapore (NMRC/MOH/ HCSAINV21nov-0001). These funders had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit for publication.

Data Availability

As this manuscript describes a study protocol, no data have yet been collected or generated. Upon completion of the study, deidentified data will be made available from the corresponding author upon reasonable request, in accordance with institutional and ethical guidelines.

Authors' Contributions

YC, KHK, JHLG, and YCT conceived the study and methodology. YC drafted the protocol and manuscript. KHK and CJWH assisted with the ethics application process and managed project administration. SMEY, JHLG, and EL reviewed the draft manuscript and provided comments. YCT secured funding, provided overall supervision, and revised the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Subject's satisfaction and acceptance evaluation questionnaire.
[DOCX File, 21 KB - resprot_v15i1e74164_app1.docx]

XSL•FO
RenderX

Checklist 1
CONSORT-AI checklist [24].
[DOCX File, 23 KB - resprot_v15i1e74164_app2.docx]

References

1. GBD 2019 Blindness and Vision Impairment Collaborators, Vision Loss Expert Group of the Global Burden of Disease Study. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. Lancet Glob Health 2021 Feb;9(2):e130-e143. [doi: 10.1016/S2214-109X(20)30425-3] [Medline: 33275950]
2. Wong TY, Tham YC, Sabanayagam C, Cheng CY. Patterns and risk factor profiles of visual loss in a multiethnic Asian population: the Singapore epidemiology of eye diseases study. Am J Ophthalmol 2019 Oct;206:48-73. [doi: 10.1016/j.ajo.2019.05.006] [Medline: 31095951]
3. Lim ZW, Chee ML, Soh ZD, et al. Six-year incidence of visual impairment in a multiethnic Asian population: the Singapore epidemiology of eye diseases study. Ophthalmol Sci 2023 Dec;3(4):100392. [doi: 10.1016/j.xops.2023.100392] [Medline: 38025163]
4. Thakur S, Lavanya R, Yu M, et al. Six-year incidence and risk factors for primary open-angle glaucoma and ocular hypertension: the Singapore Epidemiology of Eye Diseases study. Ophthalmol Glaucoma 2024;7(2):157-167. [doi: 10.1016/j.ogla.2023.08.003] [Medline: 37574187]
5. Olawade DB, Weerasinghe K, Mathugamage MDDE, et al. Enhancing ophthalmic diagnosis and treatment with artificial intelligence. Medicina (Kaunas) 2025 Feb 28;61(3):433. [doi: 10.3390/medicina61030433] [Medline: 40142244]
6. Healthy ageing hub. HealthHub. URL: https://www.healthhub.sg/programmes/aap/functional-screening [accessed 2026-01-03]
7. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J 2021 Jul;8(2):e188-e194. [doi: 10.7861/fhj.2021-0095] [Medline: 34286183]
8. Tham YC, Anees A, Zhang L, et al. Referral for disease-related visual impairment using retinal photograph-based deep learning: a proof-of-concept, model development study. Lancet Digit Health 2021 Jan;3(1):e29-e40. [doi: 10.1016/S2589-7500(20)30271-5] [Medline: 33735066]
9. Yin J, Ngiam KY, Teo HH. Role of artificial intelligence applications in real-life clinical practice: systematic review. J Med Internet Res 2021 Apr 22;23(4):e25759. [doi: 10.2196/25759] [Medline: 33885365]
10. Wolff J, Pauling J, Keck A, Baumbach J. The economic impact of artificial intelligence in health care: systematic review. J Med Internet Res 2020 Feb 20;22(2):e16866. [doi: 10.2196/16866] [Medline: 32130134]
11. Alhashmi SFS, Alshurideh M, Al Kurdi B, Salloum SA. A systematic review of the factors affecting the artificial intelligence implementation in the health care sector. 2020 Presented at: Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2020); Apr 8-10, 2020. [doi: 10.1007/978-3-030-44289-7_4]
12. Low vision and vision rehabilitation. American Optometric Association. URL: https://www.aoa.org/healthy-eyes/caring-for-your-eyes/low-vision-and-vision-rehab [accessed 2025-11-24]
13. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
14. Ranganathan P, Caduff C. Designing and validating a research questionnaire - Part 1. Perspect Clin Res 2023;14(3):152-155. [doi: 10.4103/picr.picr_140_23] [Medline: 37554243]
15. Swan K, Speyer R, Scharitzer M, et al. Measuring what matters in healthcare: a practical guide to psychometric principles and instrument development. Front Psychol 2023;14:1225850. [doi: 10.3389/fpsyg.2023.1225850] [Medline: 37790221]
16. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1). [doi: 10.1186/s13012-022-01245-0]
17. Hennink MM, Kaiser BN, Weber MB. What influences saturation? Estimating sample sizes in focus group research. Qual Health Res 2019 Aug;29(10):1483-1496. [doi: 10.1177/1049732318821692] [Medline: 30628545]
18. Ahmed SK, Mohammed RA, Nashwan AJ, et al. Using thematic analysis in qualitative research. J Med Surg Public Health 2025 Aug;6(6):100198. [doi: 10.1016/j.glmedi.2025.100198]
19. Quesado P, Silva R. Activity-based costing (ABC) and its implication for open innovation. J Open Innov Technol Mark Complex 2021 Mar;7(1):41. [doi: 10.3390/joitmc7010041]
20. Dosovitskiy A, et al. An image is worth 16x16 words: transformers for image recognition at scale. Presented at: International Conference on Learning Representations; May 3-7, 2021 URL: https://arxiv.org/pdf/2010.11929/1000 [accessed 2026-02-03]
21. Akinbobola EO, De Domenico F, Manetti S, Noto G. Health technology assessment (HTA) and performance management (PM): a scoping review on the intersecting realms. BMC Health Serv Res 2025 Jul 3;25(1):917. [doi: 10.1186/s12913-025-12879-3] [Medline: 40611138]
22. Rao DP, Shroff S, Savoy FM, et al. Evaluation of an offline, artificial intelligence system for referable glaucoma screening using a smartphone-based fundus camera: a prospective study. Eye (Lond) 2024 Apr;38(6):1104-1111. [doi: 10.1038/s41433-023-02826-z] [Medline: 38092938]

XSL•FO
RenderX

23. Antaki F, Hammana I, Tessier MC, et al. Implementation of artificial intelligence-based diabetic retinopathy screening in a tertiary care hospital in Quebec: prospective validation study. JMIR Diabetes 2024 Sep 3;9:e59867. [doi: 10.2196/59867] [Medline: 39226095]
24. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med 2020 Sep;26(9):1364-1374. [doi: 10.1038/s41591-020-1034-x] [Medline: 32908283]

Abbreviations

AI: artificial intelligence
AUC: area under the curve
AVIRI: AI for Disease-related Visual Impairment Screening Using Retinal Imaging
CFIR: Consolidated Framework for Implementation Research
CONSORT: Consolidated Standards of Reporting Trials
pRCT: pragmatic randomized controlled trial
SEED: Singapore Epidemiology of Eye Diseases
VA: visual acuity
VI: visual impairment

Edited by J Sarvestan; submitted 14.Sep.2025; peer-reviewed by CY Chen, K Jin; accepted 24.Dec.2025; published 02.Mar.2026.

Please cite as:
Chen Y, Koh KH, Ho JWC, Yew SME, Goh JHL, Lum E, Tham YC
Evaluation of a Community-Based AI-Assisted Visual Impairment Screening Model for Performance, Operational Efficiency, Acceptability, Feasibility, and Costs: Protocol for a 2-Arm Pragmatic Randomized Controlled Trial
JMIR Res Protoc 2026;15:e74164
URL: https://www.researchprotocols.org/2026/1/e74164
doi:10.2196/74164

© Yibing Chen, Kai Hui Koh, Jin Wei Clarine Ho, Samantha Min Er Yew, Jocelyn Hui Lin Goh, Elaine Lum, Yih Chung Tham. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Strengthening Care for Children Using a Virtual Integrated General Practitioner–Pediatrician Model of Primary Care (SUSTAIN): Protocol for a Stepped Wedge Cluster Randomized Controlled Trial

Tammy Meyers Morris[1], PhD; Harriet Hiscock[2,3], MD; Karen Wheeler[1], BSc; Corin Miller[1], FRACGP-RG; Nan Hu[1], PhD; Shukri Hassan Shire[1], MStats; Carmen Crespo-Gonzalez[1], PhD; Susan Bullock[1], MClinPsych; Natalie Taylor[4], PhD; Douglas Boyle[5], PhD; Lena Sanci[5], PhD; Kenny Lawson[6,7], PhD; Louisa Adams[8], FRACP; Ken Peacock[8], FRACP; Michael Hodgins[1], PhD; Annemarie Christie[8], MBBS; Raghu Lingam[1], PhD

[1]Population Child Health Research Group, Discipline of Paediatrics & Child Health, School of Clinical Medicine, University of New South Wales, Building C29 HTH Level 3, 55 Botany St Kensington, Sydney, NSW, Australia

[2]Health Services and Economics, Murdoch Children's Research Institute, Melbourne, Australia

[3]Department of Paediatrics, University of Melbourne, Melbourne, Australia

[4]Implementation to Impact, School of Population Health, University of New South Wales, Sydney, Australia

[5]Department of General Practice and Primary Care, University of Melbourne, Melbourne, Australia

[6]Translational Health Research Institute, School of Medicine, Western Sydney University, Sydney, Australia

[7]Faculty of Medicine and Health, School of Clinical Medicine, University of New South Wales, Sydney, Australia

[8]Sydney Children's Hospitals Network, Sydney, Australia

Corresponding Author:
Tammy Meyers Morris, PhD
Population Child Health Research Group, Discipline of Paediatrics & Child Health, School of Clinical Medicine, University of New South Wales, Building C29 HTH Level 3, 55 Botany St Kensington, Sydney, NSW, Australia

Abstract

Background: Australia's health care system is under pressure. Pediatric referrals to public hospital emergency and outpatient departments have increased recently, overburdening emergency services and resulting in extended waiting times for nonurgent pediatric care. Children living outside metropolitan areas are disproportionately affected. Integrated models of care with pediatricians collaborating with general practitioners (GPs) in their practices have been evaluated in the United Kingdom and Australia. Results are promising for quality of care improvement and reducing referrals to hospitals. GPs and pediatricians found the model feasible, knowledge- and confidence-boosting. In-person pediatric-GP support is resource-intensive, limiting scalability and sustainability.

Objective: The SUSTAIN trial is designed to evaluate a digitally delivered, integrated GP-pediatrician model of care. The primary objective is to determine whether the SUSTAIN model reduces GP referrals to hospital emergency departments for children <18 years. Secondary objectives include whether the model improves the delivery of guideline-concordant pediatric care by GPs, enhances GP confidence, and strengthens family trust in primary care. The trial also examines barriers and enablers to the implementation and includes a health economic evaluation comparing intervention costs with standard GP care.

Methods: SUSTAIN uses a stepped wedge cluster randomized controlled trial design to implement a GP-pediatrician integrated model of care delivered digitally. Participating GP practices across metropolitan and nonmetropolitan New South Wales are included and randomly assigned a start time. The intervention consists of 12 months' access to the shared GP-pediatrician consulting sessions with patients younger than 18 years conducted by telehealth, virtual pediatrician–led case discussions, phone/email pediatrician support, and complimentary access to the internationally renowned Sydney Child Health Program learning platform. GP and family surveys are collected at baseline and in the final month of intervention. An implementation evaluation using focus group discussions is conducted with each practice during the intervention and optional GP and family interviews at the end of the intervention. A health economic evaluation will explore the cost-effectiveness of this model of care.

Results: The trial is supported through a 2.5-year New South Wales Ministry of Health Translational Research Grants Scheme. Human Research Ethics Committee approval was obtained in November 2022, and practice recruitment began in March 2023. Data collection commenced for all participating practices from September 1, 2023, with anticipated completion on February 28, 2025. Data analysis will commence from March 2025, with results expected in the first quarter of 2026.

XSL•FO
RenderX

Conclusions: Positive outcomes for the SUSTAIN trial, demonstrating that virtual pediatric support for GPs in both metropolitan and nonmetropolitan areas can strengthen pediatric primary care provision, have the potential to influence future health policy. This innovative approach to integrated care could be rolled out across Australia and other countries with primary care–led health care systems facing similar challenges.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12623000543684; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=385645

International Registered Report Identifier (IRRID): DERR1-10.2196/69728

(JMIR Res Protoc 2026;15:e69728) doi:10.2196/69728

KEYWORDS

general practice; pediatrics; randomized controlled trial; virtual integrated care model; health services; primary care

Introduction

Background

Demands on children's tertiary health services in Australia and other high-income countries are growing [1]. Across Australia, children aged <15 years represent 26% of all lower urgency emergency department (ED) presentations and the highest presentation rates at ED [2]. While low urgency conditions do not necessarily equate with low severity or complexity [2], up to 74% of such presentations could be appropriately managed in a timely and safe manner within primary care [3]. High costs are involved in the hospitalization of children, both for governments and consumers. Alternative ways to strengthen pediatric primary care could avert some of the costs implicated in low-acuity presentations to hospitals [4,5].

Australia's health care system is centralized around primary care provided by general practitioners (GPs). Over time, GPs have seen a diminishing proportion of pediatric cases, reducing their confidence in managing common pediatric conditions [6,7]. Adherence to clinical practice guideline recommendations by GPs has also been found to be reduced for some common conditions [8,9]. As the burden and complexity of disease evolve in high-income settings, increasingly, children present to general practice with medical complexity, including chronic diseases, neurodevelopmental, behavioral, and mental health issues [10-12]. The most common health system concerns expressed by GPs in Australia include an unsustainable workload in the context of this increasing patient complexity and difficulty navigating a fragmented health system [12]. These pressures impact the delivery of high-quality care, highlighting the need to better support GPs in fulfilling their vital role [13]. Challenges faced at the primary care level contribute to health inequities, especially for children living in rural and remote Australia, who experience reduced access and longer waitlists for specialist health services, contributing to diagnostic and treatment delays, adversely impacting health outcomes, and increasing health care costs [14,15].

The need for service redesign and exploration of new approaches to health care delivery for children, young people, and families has been identified to effectively manage health service demands. The primary care workforce requires support to improve care pathways for children, especially those with medical complexity and living in nonmetropolitan areas. Telemedicine and innovative technologies to support the sharing of clinical information have been highlighted as a mechanism to address this need [16].

Our group has implemented and evaluated an integrated GP-pediatrician model of care through several trials: the Strengthening Care for Children (SC4C), based on the UK Children and Young People's Health Partnership [17], and Connecting Care for Children studies. These studies aimed to support GPs to strengthen the delivery of pediatric care, reducing unnecessary tertiary health service use and improving quality of care at the primary care level. In the UK setting, the intervention was implicated in a reduction in new patient hospital appointments, specialty referrals, and ED attendances, which was reported as a result. Families reported an increasing preference to see their GP, whilst GPs reported improved pediatric knowledge and understanding of how to navigate hospital services [18]. In Australia, the SC4C study, using an in-person, GP-pediatrician engagement model with pediatricians supporting GPs in their practices, demonstrated a reduction in referrals to hospital EDs and was found acceptable and suitable for GPs, pediatricians, and families [19]. However, the in-person model included considerable travel distances for pediatricians, even within metropolitan areas, posing equity and scalability challenges [20].

The evaluation design of the trial was discussed and piloted with GPs [19]. Although there was equipoise, one of the key areas discussed was the wish for all GPs to have access to the intervention. A stepped wedge randomized controlled trial (RCT) design was preferred to a 2-arm cluster RCT. The stepped wedge RCT design offers the gold standard of a cluster RCT. The advantage of a stepped wedge trial design ensures that all participating GP practices receive the intervention [21].

Building on learnings from the SC4C study, we propose to evaluate a model of integrated GP-pediatric care that uses a virtual approach to integrating GP-pediatric care, with the aim of demonstrating scalability and sustainability through a more equitably available model of care. The SUSTAIN study also evaluates the renowned Sydney Child Health Program (SCHP), a modular, web-based pediatric training program that has been developed for GPs, supported through the Sydney Children's Hospital Network (SCHN) in New South Wales (NSW), and which has not previously been robustly evaluated [22].

XSL•FO
RenderX

Objectives

Our primary and secondary objectives are to (1) reduce GP referrals to hospital ED and outpatient (OP) clinics for children and young people younger than 18 years and (2) improve GP pediatric care aligned with best practice (BP) guidelines, GP confidence, and increase family trust in primary care, while reducing family preference for specialist pediatrician referral.

Our implementation objective is to examine factors that help or hinder the implementation of SUSTAIN and understand factors associated with acceptability and scalability, and to assess the sustainability of the model.

Our economic objective is to assess the cost-effectiveness of SUSTAIN compared to that of standard GP care by undertaking a health economic evaluation (Table 1).

XSL•FO
RenderX

Table . Trial objectives, data sources, data collection methods, and outcomes.

Primary and secondary objectives	Data sources	Methods of collection	Period of data collection	Outcomes of interest
What is the impact of SUSTAIN on GP[a] referral to hospital OP[b] clinics and EDs?[c] (Primary Outcome)	Referral pop-up supported through GRHANITE technologies on GP desktop	GRHANITE data extraction of referral destination (including no referral) for each pediatric visit to participating GPs.	• Collected as part of pediatric GP consults at the beginning of the control data period and throughout the intervention period.	• Whether or not the GP referred the child to the hospital OP clinic or emergency department.
What is the impact of SUSTAIN on GP quality of care for common childhood conditions?	GP medical records	GRHANITE data extraction of care quality based on the measurement of the CareTrack Kids indicators.	• Collected as part of pediatric GP consults at the beginning of the control data period and throughout the intervention period.	• Whether or not the GP followed clinical guidelines (ie, they did not request unnecessary tests or prescriptions) for 5 common childhood conditions (ie, asthma/wheezing, bronchiolitis, constipation/abdominal pain, upper respiratory infections, and infant crying, and reflux).
What is the impact of SUSTAIN on GPs? (eg, confidence and skills in pediatric care and use of clinical guidelines)	GP web-based survey	Web-based control and intervention surveys completed by GPs via REDCap[d].	• Control surveys are collected in the month prior to the implementation commencing in each practice. Intervention surveys are collected in the last month of the implementation at each practice.	• Changes in the level of confidence in pediatric care, level of knowledge and skill in navigating the health system for children, and reported use of clinical guidelines.
What is the impact of SUSTAIN on patients and family experience?	Family web-based survey	Web-based control and intervention surveys completed by families/caregivers via REDCap.	• Control surveys are collected in the month prior to the implementation commencing in each practice. Intervention surveys are collected in the last month of the implementation at each practice.	• Level of confidence in GP care, level of satisfaction with GP care, desire for referral to specialist care, and preference for GP or specialist review.
Health economic evaluation				
What is the cost of implementing the model of care? What is the cost- effectiveness?	Trial data and supplementary unit costings	Trial data on the model of care, health service activity, and combined with relevant unit costs.	• Data will be collected throughout the model of care.	• Costs of conducting the model of care compared with usual care and costs/cost offsets from changes in OP/ED referrals, MBS,[e] and PBS[f] (including out-of-pocket fees) compared with accessing usual care.
Implementation evaluation				

XSL•FO
RenderX

Primary and secondary objectives	Data sources	Methods of collection	Period of data collection	Outcomes of interest
What are the aspects of the model of care that make it effective or ineffective at producing system change?	Pediatrician collected data, qualitative interviews, and web-based surveys with GPs, general practice managers and administration staff, trial pediatricians, families, and children	Consolidated framework for implementation research; qualitative interviews with families, children, and practitioners; web-based surveys. Pediatricians will collect unidentifiable data on the patient characteristics (eg, age and sex) and nature of pediatric support provided (eg, reason for consult and topic of case study discussion).	• Interim qualitative data will be collected via focus groups with general practice 6 months into the model of care (iterative data collection process). Interviews with GPs, families, and pediatricians will be conducted at the end of the model of care. • Web-based survey data will be collected upon completion of the model of care in each general practice; pediatrician data will be collected as part of the coconsultations, case study discussions, and phone/email support throughout the model of care.	• Identify strategies for successful implementation as well as barriers and facilitators in adoption, delivery, and maintenance to inform future scaling. • Describe the model of care, including number of children seen in GP-pediatrician coconsultations, reasons for coconsultations, number of and reason for phone and email support to the pediatrician, number and topic of case study discussions. • Feasibility/acceptability and appropriateness of the model; adoption and fidelity to the model.
Sustainability				
Explore the sustainability and enduring effects of SUSTAIN post implementation on proportion of GP pediatric referrals to OP clinics or EDs and GP quality of care compared with preintervention GP care.	GP medical records	GRHANITE data extraction of referral destination (including no referral made) for each pediatric visit to participating GPs and of care quality based on the measurement of care provided for the 5 common childhood conditions.	• Data will be collected on pediatric GP consults following completion of the model of care in each general practice (ie, once access to pediatrician support has ceased) until the end of the trial.	• Whether or not the GP refers the child to a hospital OP or ED. • Whether or not the GP followed clinical guidelines for 5 common childhood conditions (as above). • How GP pediatric referrals (1) and quality of care (2) in the sustainability period compared with the intervention period (when the pediatrician was in the clinic). • Economic evaluation results reflecting national rollout in real-world sustainable settings along with budget impact.

[a]GP: general practitioner.

[b]OP: outpatient.

[c]ED: emergency department.

[d]REDCap: Research Electronic Data Capture.

[e]MBS: Medicare Benefits Schedule.

[f]PBS: Pharmaceutical Benefits Scheme.

Methods

Overview

This paper reports the research protocol for the SUSTAIN trial, including how we will partner with primary care, government, and pediatric hospitals to evaluate the effectiveness and cost-effectiveness of SUSTAIN in a cohort of general practices across NSW, using a stepped-wedge cluster RCT. The trial will be conducted according to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist. We will apply the Medical Research Council framework used for developing and evaluating complex interventions [23].

XSL•FO
RenderX

Trial Design

SUSTAIN is a stepped wedge cluster RCT of a virtual GP-pediatrician integrated model of care compared with standard GP care. This trial design has the rigor of a cluster RCT but allows all participating general practices to be exposed to the SUSTAIN care model, providing control, intervention, and postintervention data. This trial design was selected to ensure all practices receive the intervention [21] (Figure 1).

Figure 1. Stepped wedge randomized control trial design.

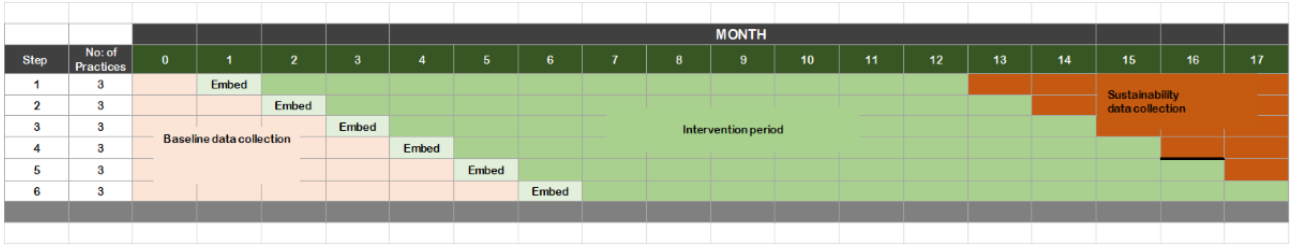

Each general practice provides control period data (referral and medical record data) during the control period (standard GP care, no SUSTAIN model operating). Each month thereafter, 3 practices switch concurrently from control to intervention where a pediatrician supports GPs through telehealth coconsultations, virtual case-based discussions, and being available by phone and email contact in working hours. During the first month of the intervention for each practice, a 1-month transition period is allowed to embed the SUSTAIN model of care into the practice, where data do not contribute to the analysis. This embedding period assists with implementation of the model of care into the practice and allows time to address any procedural issues. After the embedding period, each practice is exposed to an 11-month intervention period. Sustainability data are collected following cessation of the intervention until the intervention is completed for the final practice.

Setting and Participants

SUSTAIN is a multisite trial conducted across GP practices in different local health districts within NSW. GP practices in both metropolitan and nonmetropolitan (regional, rural, and remote) areas of NSW, including areas with higher-than-average proportions of priority populations, will be eligible to participate.

Trial participants will include general practices with consenting GPs across NSW. Families of children <18 years attending the general practice may be asked to complete anonymous surveys regarding their experience of pediatric care in GP services.

Eligible general practices include those responding to an expression of interest and meeting the inclusion criteria (Textbox 1).

Textbox 1. Trial population exclusion and inclusion criteria.

Inclusion criteria

General practices:

- Be located within metropolitan and nonmetropolitan local health districts in New South Wales.
- Have best practice or medical director 3 as their electronic medical record.
- Be accredited or working toward accreditation against Royal College of General Practice standards.
- Agree to install GRHANITE software (for extracting general practitioner [GP] medical record data).

GPs:

- See patients younger than 18 years.
- Provide a minimum of 2 weeks of referral data during the control period.
- Sign participant information consent forms.

Families:

- Caregivers of children younger than 18 years who have received care at the general practice regardless of GP seen, in the prior 3 months for both the control and the intervention period.
- Caregivers with sufficient English to complete the survey.

Exclusion criteria

Families:

- Children or young people who present to the general practice without a parent/guardian.
- Insufficient English.

XSL•FO
RenderX

Recruitment

General Practices and GPs

An expression of interest will be distributed to practice managers, GPs, and practice nurses through the primary health networks (PHNs) [24], Central and Eastern Sydney, South Western Sydney, South Eastern NSW, and directly to other general practices in rural, regional, and remote areas in NSW. Interested general practices will be required to sign a memorandum of understanding with the research team adhering to the requirements of their participation, and sign a license agreement to install the clinical data extraction software tool GRHANITE [25]. GRHANITE is discussed in the Data Management Plan section. Each GP will be required to sign a participant information consent form (Section S1 in Multimedia Appendix 1) to participate in the trial. Any GP wanting to participate during the intervention phase must provide at least 2 weeks of control referral data. Each participating practice will receive a one-off AU $1000 (approximately US $655) payment to support the implementation of the trial into their practice. An ethics-approved poster will be available for practices to promote the GP-pediatrician coconsultation service and inform patients about the study.

Families/Caregivers

Caregivers of patients younger than 18 years who present to a participating general practice are invited to complete an anonymous, web-based survey about their perceptions and experience of the care received during a recent GP consultation (Section S4 in Multimedia Appendix 2). These surveys will be sent by each practice via a broadcast SMS text message on behalf of the research team to families/caregivers of patients younger than 18 years who have received care at the general practice, regardless of GP seen, in the 3 months preintervention (baseline) and in the 3 months before the SUSTAIN model of care (intervention) ends. In the final survey, caregivers will be invited to voluntarily participate in a qualitative interview via telephone/web-based video led by the implementation evaluation team for the trial.

Randomization and Blinding

Participating practices will be randomized for intervention using a predetermined web-based randomization sequence generated by an independent trial statistician. Participating general practices will be randomly ordered in terms of when they switch from standard GP care to the SUSTAIN model of care by the independent statistician. To avoid recruitment bias in clusters, randomization will occur once all general practices have been recruited and enrolled; that is, after all inclusion and exclusion criteria are addressed, and all general practices sign the relevant trial agreements. Following randomization, practices will be unblinded to their allocation status/model of care start date. It is necessary for each general practice to know their randomization position to prepare for intervention, so allocation is not concealed. It is not anticipated that GPs knowing their randomization status will impact/change their standard practice due to their demanding workload. The research team is also unblinded to each practice's allocation status to allow them to effectively engage with each practice to prepare their implementation schedule.

Intervention

SUSTAIN is a 12-month-long virtual model of care offered to all consenting GPs in participating general practices. It consists of four components: (1) virtual GP-led, shared GP-pediatrician consulting sessions with patients younger than 18 years, (2) virtual pediatrician-led case discussion sessions, (3) phone/email pediatrician support, and (4) complimentary access to the internationally renowned SCHP learning platform (Figure 2).

XSL•FO
RenderX

Figure 2. SUSTAIN intervention (model of care components). GP: general practitioner.

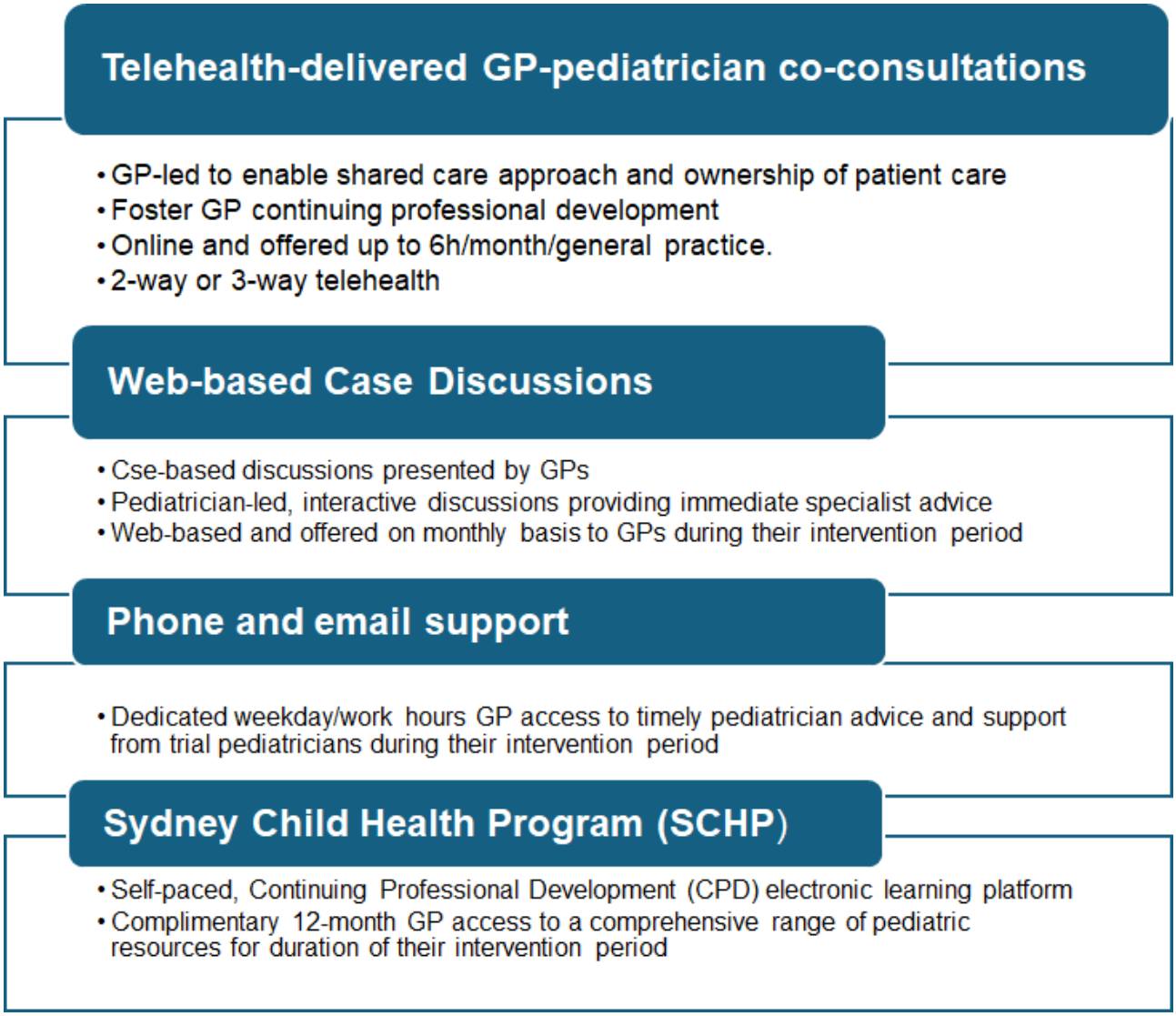

During their 12-month intervention period, GPs may engage with any component of the model of care based on their needs, capacity, and learning style. The SUSTAN model of care will be delivered by a 1.0 FTE pediatrician (can be a shared position). The pediatricians will be employed through the SCHN.

Data Management Plan

All participants (GPs and caregivers) will be assigned a unique numerical identifier (an ID code) generated in REDCap (Research Electronic Data Capture; Vanderbilt University) for use throughout the trial. A single electronic, password-protected database in REDCap will record all general practices, GP details, and survey data. Any data recorded will be stored on the secure University of New South Wales (UNSW) network drive on a password-protected computer and will be made available for importing into the project secure research environment. The project database will only be accessible by the designated research team members (investigators, statistician, and project managers) and hosted on the UNSW secure server, which meets security and ethical confidentiality requirements.

GRHANITE, developed and managed by the Health and Biomedical Informatics Center at the University of Melbourne (UoM), specifically designed for research purposes [25,26], extracts and curates the delivery of deidentified and encrypted primary care data to secure research data storage facilities. Person identifiers are removed from the data, in keeping with data custodian permission, and subject to ethics committee approval and legal agreement. Data extracted by GRHANITE will be stored in a Research Databank on the UoM Research Cloud, physically located within the UoM Secure Data Center. GRHANITE data will be made available to the research team from this databank via a UoM secure virtual research environment that gives researchers access to the data in a controlled manner. Upon completion of GRHANITE data extraction, only deidentified data will be transferred to the project analysis team, with a project ID unique to each patient. GRHANITE is only compatible with BP and medical director (MD) electronic medical records (EMRs), used by ≈90% of Australian general practices [27].

Patient confidentiality will be strictly held in trust by the participating investigators, research staff, and the sponsoring institution and their agents. No information concerning the trial or the data will be released to any unauthorized third party

XSL•FO
RenderX

without the participant's consent and written approval of the sponsoring institution.

All interview data will be transcribed and deidentified for analysis and will be stored in a restricted-access folder on UNSW network drives on a password-protected computer. Any personal information or quotes attributed to individual participants in published form will be anonymized.

Data Collection

GP EMR Data Collection

To measure the primary outcome (ie, referrals to hospitals), pediatric referral data will be routinely collected from the GP EMRs via GRHANITE. This software will be remotely embedded into all participating GP medical software computers (compatible with BP or MD EMR software). For research purposes, GP data will be collected on all patients younger than 18 years. These data will be deidentified at the patient level, although the identity of the GP will be supplied to the research team. Collection of GP EMR referral data will commence for all practices following randomization and will continue for the duration of the trial.

The ability to record referral information, manage deidentification and consent processes, is a principal rationale for the use of the GRHANITE technologies and was demonstrated to be acceptable to GPs in the SC4C study [19].

GP Referrals

Data on GP referrals is not recorded in a standardized fashion in the GP EMR. Therefore, in order to measure the primary outcome, that is, referrals to hospitals, the GRHANITE team developed a tailored referral pop-up window (Section S5 in Multimedia Appendix 3) of common referral options to be completed by GPs following each pediatric consultation. This specific software extracts information regarding whether the GP referred a child and where (eg, no referral, hospital OP, or emergency department, private pediatrician, or allied health). The research team will continuously monitor GP data to promptly identify and address any technical issues posing a risk to data acquisition.

GP Quality of Care

GRHANITE will be used to extract deidentified patient-level data routinely collected on all patients younger than 18 years seen by consented GPs throughout the trial. Deidentified pediatric patient data will be limited to patient demographics, reason for visit, diagnoses, referrals, prescriptions, ordered imaging and pathology testing, and Medicare item billing. The patient-level data will be used to characterize GP pediatric visits (eg, number of children seen, gender of the children, and common diagnoses) and compare quality of care for common conditions before and after the model of care is implemented, based on the CareTrack Kids indicators [28].

A natural language processing algorithm, developed by the Computing and Information Systems at the UoM for the SC4C trial, will be used to automatically transform GP EMR clinical free text of "reason for visit" or diagnosis into structured clinical data, based on the Systematized Nomenclature of Medicine Clinical Terminology [29].

GP Model of Care Engagements

As per our ethics approval, GP engagement with the components of pediatric support, such as coconsults, case discussion, phone/email, and SCHP access, will be recorded by the following processes and stored on the secure UNSW network drive. For coconsults, GPs will record deidentified patient information (age, gender, presenting problem, and reason for coconsult) via the SCHN Microsoft Bookings. For email, phone, and case discussions, pediatricians will record via REDCap each engagement, including GP name and deidentified patient information on the topic or concern discussed. For SCHP usage, the SCHP team will record the GP name and track GP access with the SCHP learning platform.

GP and Family Data Collection

GP and family survey data collection will occur preintervention (baseline) and in the last month of the SUSTAIN model of care (intervention) for each practice.

A web-based GP Survey (Section S3 in Multimedia Appendix 4) will be used to measure GP experience and confidence in pediatric care, completed by all participating GPs via REDCap. Based on the SC4C trial, the survey collects information including GP demographics, factors that impact their decision to refer a pediatric patient, and knowledge and confidence in pediatric care and services. GP name and general practice will be recorded for survey completion tracking, but will not be attached to the survey responses.

A web-based family survey (Section S4 in Multimedia Appendix 2) will be used to measure caregivers' perceptions and experience of pediatric care provided at their general practice. To provide data on a "whole of practice" level, caregivers of a patient <18 years, regardless of whether they attended a consultation with a participating GP, will be eligible to complete the Family Survey via REDCap. The survey collects information on the child/family regarding a GP visit up to 3 months prior. These surveys are voluntary, and responses are anonymous.

No power calculation has been performed, as we will recruit as many families as possible during this timeframe and use all available data.

SUSTAIN GP and family surveys have been adapted from the SC4C study and comprise items generated by the SC4C investigators and drawn from previously published literature [19,30,31]. Data collection timepoints for participant enrollment and interventions can be found in Table 2.

XSL•FO
RenderX

Table . Outline of participant enrollment, intervention, and data collection timepoints.

	Enrollment ($-t_1$)	Allocation (t_0)	Study period			Close out (t_x)
			Control period (t_1)	1-month Embedding period (t_2)	11-month Intervention period (t_3)	
Enrollment						
General practice eligibility screen	✓					
Informed GP[a] consent	✓					
General practice intervention allocation		✓				
Intervention						
Virtual GP-pediatrician coconsultation				✓[b]	✓	
Virtual case discussions				✓[b]	✓	
Pediatrician phone/email support				✓[b]	✓	
GP access to SCHP[c]				✓[b]	✓	
Primary outcome						
GP referral outcomes						
GRHANITE extracted from routinely collected GP EMR[d] data			✓	✓[b]	✓	
Secondary outcomes						
GPs						
Quality of care						
GRHANITE EMR data extraction on care quality based on measurement of the CareTrack Kids indicators: asthma, bronchiolitis, constipation, upper respiratory tract infections, and infant crying and gastroesophageal reflux			✓	✓[b]	✓	
Baseline and intervention /follow-up surveys						
GP demographics			✓		✓	
Confidence in pediatric care and access to pediatric services			✓		✓	

XSL•FO
RenderX

	Enrollment ($-t_1$)	Allocation (t_0)	Study period			Close out (t_x)
			Control period (t_1)	1-month Embedding period (t_2)	11-month Intervention period (t_3)	
Skills to manage child health			✓		✓	
Awareness and use of HealthPathways			✓		✓	
Model feasibility, acceptability, feedback, and patient benefit					✓	
Families/caregivers						
Baseline and intervention /follow-up surveys						
Family demographics			✓		✓	
Confidence in GP care			✓		✓	
Quality of care and interactions with the GP			✓		✓	
Preference for pediatrician referral and GP review			✓		✓	
General practice implementation focus groups				✓		
Implementation evaluation						✓
Economic evaluation						✓
Reporting and manuscript reporting						✓

[a]GP: general practitioner.

[b]Data collected during the embedding period will not be analyzed.

[c]SCHP: Sydney Child Health Program.

[d]EMR: electronic medical record.

Sample Size Calculation

In the SC4C pilot study, there was a 7% reduction in GP referrals to hospital EDs and OPs, but we have powered our trial using a conservative 4% reduction [19,32]. The sample size of 18 practices with at least 60 pediatric consultations per practice per month will have 90% power to detect a 3.7% reduction (from the baseline rate of 10%) in the percentage of children who are referred to OP clinics or EDs following their GP appointment, assuming an intracluster correlation of 0.06 (derived from the pilot data [19]) and 2-sided type I error of 0.05. With the same sample size, we will have 80% power to detect a 5.5% - 6.55% reduction in the proportion of referrals to OP or ED for our planned stratification analysis by practice rurality, GP baseline referring level, years of practice, and gender.

Engagement/Retention Strategies

Several methods will be used to maintain general practice and GP engagement in the trial and minimize “withdrawals” or “lost to follow-up.” These will include email reminders on accessing the SUSTAIN model of care components, “touch point” meetings by the researchers with practices, and quarterly newsletters providing updates and relevant trial information.

Statistical Analysis

We will undertake descriptive analysis to describe the GP referral rate alongside the numbers of all children seen and the

XSL•FO
RenderX

numbers of children referred to OP or ED by month during the study period for each practice.

All available data from each recruited GP and family will be analyzed according to an intention-to-treat principle. For the outcomes of GP referrals identified using the GP EMR data (eg, GP referral to OP clinic or ED), we will conduct mixed effects logistic regression, fitted to data collected during the control and intervention periods. The model will include a fixed effect of group (intervention vs control) and adjust for calendar time (as a continuous variable), and random effects to allow for variation in the outcomes for 3 hierarchical levels, including GP practice, GP within each practice, and patients seen by each GP.

Secondary outcomes collected at the child consultation level will be analyzed similarly, with separate models for the 5 common childhood conditions to measure quality of care. GP survey outcomes will be analyzed using mixed effects logistic regression, again including a fixed effect for group by intervention status and calendar time, and a random effect for GP practice, GPs, and patients.

We will control for baseline levels of GP referral for each practice in the analyses, GP-level factors (eg, rurality of the practice's location, billing type, GP gender, years of practice, and number of children patients seen per week), and patient-level sociodemographic factors (eg, age at the encounter and socioeconomic status of residence). We will also consider whether the intervention effect varies by rurality of practice and GP-level factors (eg, referring level during baseline, years of practice, and GP gender) by the inclusion of interaction terms between group (intervention vs control) and these variables in the regression models.

We will assess whether the correlation structure has been mis-specified. The primary analysis will assume an exchangeable within-cluster correlation structure, meaning that the correlation between any 2 individuals within the same cluster is considered constant. To evaluate the impact of potential misspecification, we will conduct a sensitivity analysis using a block-exchangeable correlation structure within periods. This structure assumes stronger correlations among individuals within the same period, with weaker correlations across different periods. To implement this, we will extend the random-effects model to include an interaction between time and the GP practice random effect.

Prior to analysis, we will understand the mechanisms of the occurrence of missing data in the primary and secondary outcomes. If there is a small number of missing data (<5%), complete case analysis will be presented as the primary analysis. Missing data can occur due to 2 main reasons, considering the current study design: the pop-up failure to launch due to technical issues (nonmonotone missingness), and GP or GP practices withdrawal from the trial (generally monotone missingness, though could also be instances of nonmonotone, for example, personal leave such as maternity leave and retirement). It is reasonable to assume that data missing due to technical failure or personal leave occur missing completely at random. Based on our knowledge, the main reason for GP and GP practice withdrawal is the GP leaving the participating practice. Therefore, it is reasonable to assume these missing data occur as missing at random, and we will undertake multiple imputation to account for the impact of the missing data on the outcomes. It is unlikely that data is "missing not at random (MNAR)" in the nature of this current study, that is, GPs withdraw from the study because of their referral patterns, and this cannot be accounted for by all observed data. If we identify any MNAR, we will not impute the data and will report and discuss the results with limitations and the reasons potentially responsible for the missingness.

We will conduct an intention-to-treat analysis as the primary approach to understand the effect of offering the intervention on referral patterns, under real-world conditions. This means the data will be analyzed based on the information generated from a consenting GP in a randomized GP practice, according to the intervention that is supposed to be delivered in a particular period, irrespective of whether the intervention has been implemented as planned (practices analyzed by randomized step, irrespective of actual start date). Prespecified sensitivity analyses included as-treated and per-protocol analyses, modeling approaches that incorporate the exact calendar date of intervention initiation (to address early or delayed starts), and methods to explore the influence of contamination or noncompliance (for example, complier average causal effect estimates or instrumental variable techniques where appropriate).

Economic Evaluation

From a health sector perspective, a within-trial economic evaluation will be conducted following methodological and reporting guidelines [33-35]. The investment costs of conducting SUSTAIN (including GP training, practice administrative support, pediatrician time, consults, and case-study discussions) will be assessed relative to GP usual care using standard price lists and any patient out-of-pocket payments (including Medicare Benefits Schedule, Pharmaceutical Benefits Schedule, and Independent Hospital Pricing Authority). The costs (and cost offsets) associated with changes in study outcomes (including ED presentations and OP attendances) will then be estimated using similar methods and aligned with the statistical analysis (described above). Three forms of complementary economic analysis and metrics will then be conducted using the information generated. First, the incremental net cost per referral avoided (primary study outcome) will be generated relative to usual care to estimate cost-effectiveness. Second, a return on investment will assess (expected) net cost offsets from avoiding ED/OP relative to the investment in SUSTAIN, and changes in Medicare Benefits Schedule/Pharmaceutical Benefits Scheme use. Probabilistic sensitivity analysis will assess statistical uncertainty in both analyses and the cost-effectiveness acceptability curves generated. Third, a budget impact analysis [36] will estimate the costs of implementing the model at the national level and disaggregated by multiple potential payers (State and Commonwealth). This will inform affordability considerations and combine with wider trial findings to guide translation and sustainable implementation.

XSL•FO
RenderX

Implementation Evaluation

Overview

The mixed methods implementation evaluation for SUSTAIN will be based on the methodology developed for the SC4C trial [37]. We will use the Consolidated Framework for Implementation Research (CFIR) [38] to understand the effectiveness of the SUSTAIN model in driving systems change, while also identifying contextually relevant strategies for successful implementation. Factors (ie, barriers and facilitators) identified as moderating the adoption, delivery, and maintenance will also be used to inform wider state and potentially national rollout. The mixed methods evaluation will allow us to assess implementation metrics as defined by Proctor et al [39] (Table 3).

Table . Trial implementation metrics.

	Questions addressed by each implementation factor
Acceptability	Do practitioners, parents, and children view the SUSTAIN model as agreeable?
Adoption	To what extent do practitioners and parents use the SUSTAIN model?
Appropriateness	Do stakeholders perceive SUSTAIN as relevant and useful?
Fidelity	Is SUSTAIN applied as intended? Are all component parts of the intervention delivered as planned?
Coverage	How many service users of those eligible are reached?
Cost	How much does it cost to successfully implement SUSTAIN?
Sustainability	What are the factors that will allow SUSTAIN to be scaled up further?

A logic model (Figure 3) has been developed to inform the implementation evaluation guided by the implementation research logic model [40]. The logic model encompasses the specific contextual determinants that the implementation evaluation would need to consider, both within (eg, inner context representing individual factors and organizational settings) and external to the sites (eg, area demographics and socioeconomic status). In addition to the contextual factors, the logic represents the measurable intervention characteristics, implementation strategies, mechanisms of action, and outcomes.

Figure 3. Trial logic model. GP: general practitioner; PHN: primary health network; SCHP: Sydney Child Health Program.

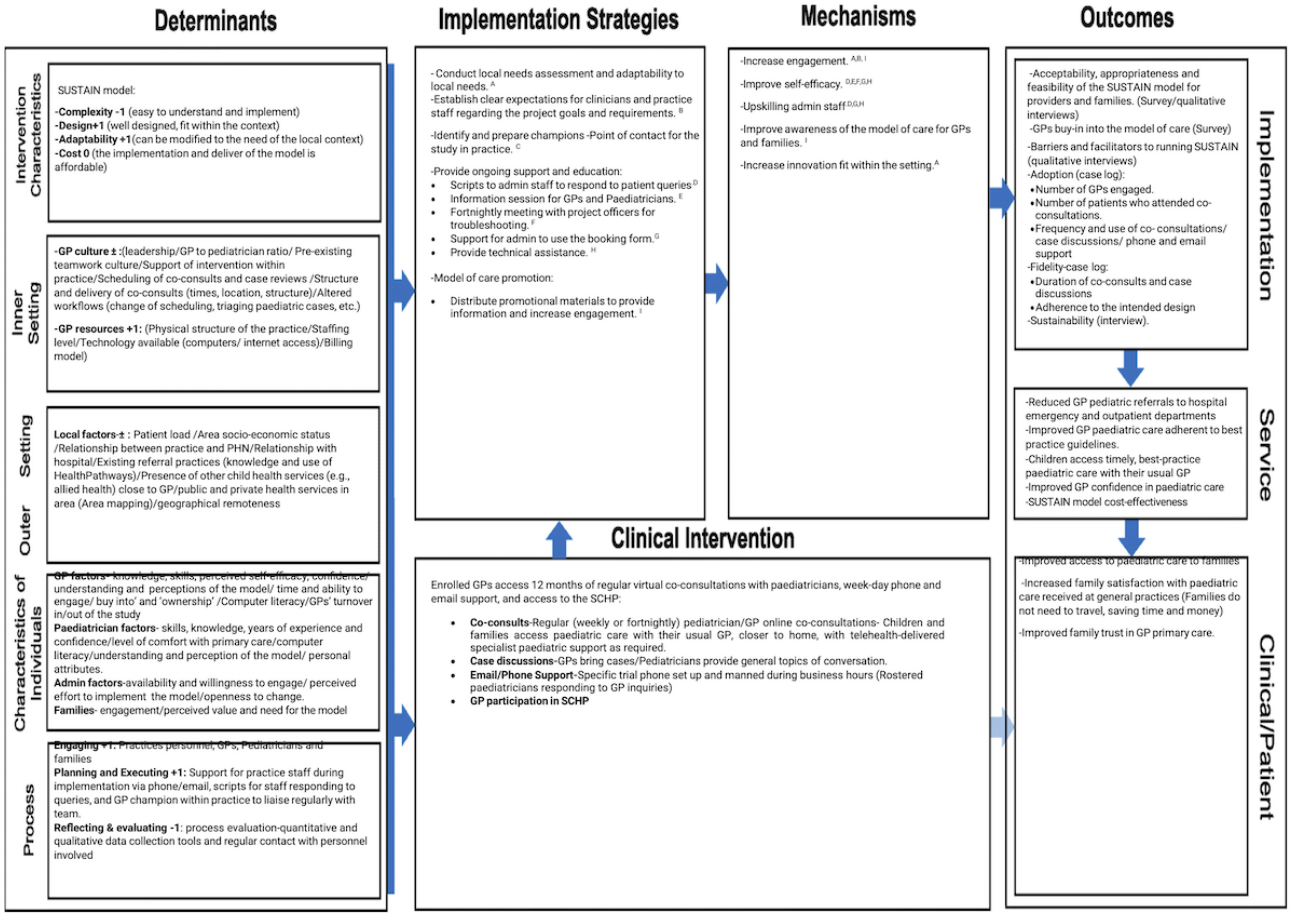

XSL•FO
RenderX

The implementation evaluation will use a mixed methods approach; findings will be triangulated as in our previous studies [41,42].

Surveys with GPs will be conducted at the start of and end of the intervention. GPs survey will include 2 validated instruments: the short intervention acceptability, appropriateness, and feasibility measure, and the NoMAD tool, based on the Normalization Process Theory, which assesses to what extent practitioners "buy into" SUSTAIN [43-45]. The survey will assess their experience of the current model of care and confidence in pediatric care. All GP instruments have been used in our previous SC4C trial and shown to be acceptable and feasible to use.

To determine individuals' knowledge and beliefs about the model of care, relative advantages of the model of care, GP and pediatrician self-efficacy, barriers and facilitators affecting the delivery of the intervention both from an individual and organizational perspective, the appropriateness and acceptability of the intervention; and recommendations for future implementation, we will conduct focus groups or interviews with GPs, practice managers, administrative staff, pediatricians, and families who participated in coconsultations. Interview guides have been derived from the CFIR. Practice managers and administrative staff will specifically be asked about how the model of care affected the normal operation of general practices. GPs and pediatricians will specifically be asked about features of the working relationship, for example, the collaborative nature of the relationship. Interviews with family members will determine their perceptions of the acceptability of the SC4C model and any potential adaptations to the model to make it more acceptable for families of children presenting to GP practices. Additionally, feedback from general practice staff and pediatricians will be routinely collected (eg, through meetings and email) to provide ongoing support during the implementation.

As described above, GPs will be asked to complete a short web-based survey before and after the intervention rollout via REDCap. As part of the implementation evaluation, GPs, practice managers, and administrative staff in participating practices will be contacted to participate in a focus group within each practice, with the option for an individual interview ≈6 months into the intervention and individual interviews with GPs at the end of the intervention at 12 months. Feedback will be used to determine appropriateness and acceptability of the intervention, identify barriers and enablers affecting delivery of the intervention, and sustainability of the model. The focus groups and interviews will be facilitated by investigators responsible for the implementation evaluation, and prior to the focus group or interview, the researcher will describe to participants the reasons for conducting the interview or focus group and provide a participant information sheet and a web-based consent form. The researcher will respond to any participant questions or concerns and inform participants that they can stop the interview at any time and revoke their consent to participate during or after the interview. In this event, interview recordings and transcripts will be removed from the trial and destroyed.

Family surveys conducted during the intervention period will include an item seeking permission to contact them about the opportunity to participate in a voluntary qualitative interview via telephone or web. This opportunity will only be open to families who have participated in a GP-pediatrician coconsultation, which will be determined earlier in the survey. When contacting parents or caregivers, the researcher will describe the reasons for conducting a follow-up interview, will not ask for the identity of their GP, and reassure parents that any information obtained will be confidential and will have no impact on the care their child receives from their GP. They will be asked to sign a web-based consent form.

The researcher will respond to any parent's or caregiver's questions or concerns and inform participants that they are able to stop the interview at any time and revoke their consent to participate during or after the interview. In this event, interview recordings and transcripts will be removed from the trial and destroyed.

Implementation evaluation data collection and analysis will assess the following.

Acceptability, Appropriateness, and Feasibility of the SUSTAIN Model and GPs' Buy-In Into the Model

Questionnaire data and open-ended questions from surveys will be exported into SPSS for analysis. Descriptive statistics will be calculated for all the practices recruited in the trial, including information about the inner and outer context and the intervention use and its acceptability.

Adoption and Fidelity

GRHANITE will allow us to access routinely collected process data from GP systems, allowing us to assess the coverage of the intervention both in terms of the number and characteristics of enrolled GPs, and the number and characteristics of children seen. This is especially important if, as we hypothesize, SUSTAIN will increase access to care for children with increased socioeconomic need. In addition to best describing the model of care, unidentifiable data will be collected on patient characteristics (eg, age and sex), reason for coconsult, topic or concern discussed at case-based discussions, and phone/email support during the 12-month intervention period at each site. This information will be recorded in REDCap and a secure UNSW network drive on a password-protected computer. No identifiable information will be collected, and data will only be analyzed to provide information on patterns of support provided. No data on individual children will be reported. This data will also provide data to understand if the intervention was implemented with fidelity in terms of the number of coconsults offered and the number of case discussions completed.

Barriers, Facilitators, and Strategies for Scaling Up

The CFIR domains allow us to assess not only intervention components but also the context in which primary care is currently situated. The CIFR will inform the development of the interview guides. Purposive sampling will be used to recruit a diverse sample of families, caregivers, and frontline practitioners, including GPs, pediatricians, and practice managers. Data collection for the qualitative interviews will

XSL•FO
RenderX

continue until data saturation (ie, when no new themes pertaining to the research objectives are identified with subsequent interviews). All interviews and focus groups will be audio recorded and transcribed verbatim using Otter (Otter.ai, Inc), a secure transcription software. The data collected will be deidentified in preparation for data analysis. Transcripts will be analyzed thematically using an iterative thematic analysis following Braun and Clarke's [46] approach. All participant interview audio recordings will be destroyed upon completion of the trial. Qualitative software (Nvivo 12; Lumivero) will be used to organize and classify data into emerging themes.

Sustainability

To assess the longer-term impact of the SUSTAIN model, sustainability data, referral data, and quality of care indicators will be collected from the time the initial practices complete the intervention. This data will be analyzed in conjunction with qualitative data collected from GPs and practice staff to assess whether changes in referral practices, adherence to standard practice guidelines, and GP confidence are sustained after the intervention has ceased.

Ethical Considerations

This trial is approved by the Human Research Ethics Committees of SCHN (2022/ETH02068), NSW, Australia. The trial has been registered on the Australian New Zealand Clinical Trials Registry (ANZCTR number ACTRN12623000543684).

Written informed consent is obtained from each GP prior to participating in the SUSTAIN trial (Section S1 in Multimedia Appendix 1). All participants in focus groups or interviews (GPs, practice managers, administration staff, and pediatricians) are asked to complete a web-based consent form via REDCap (Section S6 in Multimedia Appendix 5).

A waiver of consent to collect deidentified data on GP care provided during a GP-pediatrician coconsultation was approved (2022/ETH02068).

Informed consent is obtained from families participating in the family experience surveys by agreeing to this through the web-based survey via REDCap. All families are provided with a plain-language statement about the study, explaining the purpose of the study and offering an opportunity to voluntarily participate in the anonymous survey (Section S2 in Multimedia Appendix 6). If the caregiver agrees to take part in the survey, the caregiver's consent will be indicated via a checkbox at the start of the web-based survey.

Follow-up family surveys will include an item seeking their permission to be contacted about an opportunity to participate in a telephonic or web-based (eg, via Zoom and Teams) interview. Prior to conducting interviews with families, the researcher provides a participant information sheet and consent form for families to consent via web if they agree to participate (Section S7 in Multimedia Appendix 7). The REDCap e-Consent Framework provides a standardized tool to obtain consent and store consent, which automatically generates a "hard-copy" PDF of the signed forms.

Participant confidentiality will strictly be held in trust by the investigators, research staff, and the sponsoring institutions and their agents, and will be extended to cover clinical information relating to participants. The trial protocol, documentation, data, and all other information generated will be held in strict confidence and in password-protected electronic files. No information concerning the trial or the data will be released to any unauthorized third party without prior written approval of the sponsoring institutions. Investigators will have access to the final dataset via permissions maintained by the data managers.

GP practices will receive a one-off AU $1000 (US $655) payment to support implementation of the trial into their practice. GPs and families are not otherwise compensated.

The chief investigators (CIs) will maintain overall accountability, and a project manager and research assistant will provide support and assistance with trial implementation and address any process issues through the duration of the trial. Project team (weekly), research team (monthly), and advisory committee (quarterly) meetings will be held to foster information sharing, problem-solving, and decision-making regarding the trial, as well as ongoing consideration of knowledge translation. The project team includes CIs, trial pediatricians, project manager, research assistant, and implementation evaluation team; research team, in addition to project team, includes economic evaluator, trial statistician, GRHANITE representative; advisory committee, in addition to project and research teams, includes GPs, consumers, and policymakers.

CIs will hold the primary responsibility for publication of the results of the trial in accordance with the trial publication and dissemination plan. The findings from this trial will be reported according to the CONSORT (Consolidated Standards of Reporting Trials) statement guidelines [47].

Results

The trial is supported through a 2.5-year NSW Ministry of Health Translational Research Grants Scheme (TRGS Round 6). This includes direct funding for research staff, as well as in-kind contributions from the SCHN supporting 1.0 FTE pediatrician, estimated to be the appropriate allocation of pediatrician time for the number of practices from our previous experience with the SC4C study. GPs also receive complimentary enrollment in the SCHP, a pediatric web-based modular learning course for GPs. In-kind contributions from PHNs, Central Eastern Sydney Primary Health Network, South Western Sydney Primary Health Network, and South Eastern NSW Primary Health Network also contributed to support recruitment and ongoing engagement with general practices in each PHN region. From March 2023, general practice and GP recruitment commenced across GP practices in metropolitan and nonmetropolitan (regional, rural, and remote) areas in NSW. Data collection then started in all consented practices from September 1, 2023, with anticipated completion by February 28, 2025. The delivery of the virtual 12-month-long pediatric-integrated GP intervention commenced in the first practices from October 2023 and will conclude in the final practices at the end of February 2025. GP and practice staff are engaged in focus group discussions as part of the implementation evaluation midway through the intervention phase. Data analysis, report, and manuscript preparation will begin from

XSL•FO
RenderX

March 2025, with results expected to be available by the first quarter of 2026.

Discussion

Principal Findings

One method to prevent rising pediatric attendance at emergency departments and OP services could be strengthening the quality of pediatric health care provided by GPs closer to where people live and by increasing family confidence in the care provided by GPs for their children. This is particularly important in rural areas of countries like Australia due to the vast geographical regions. The further primary care providers are located outside of metropolitan centers in Australia, the fewer specialist services are available [48].

A primary care workforce operating at the top of the scope of practice is key to maintaining appropriate levels of services within the health care system, regardless of location. The recently released "Unleashing the Power of our Health Workforce—a Scope of Practice Review" reports that almost all health professionals in the primary care sector in Australia face some restrictions or barriers to working at the full scope of practice [49]. The SUSTAIN model, using a virtual care platform to bring GPs and pediatricians together, has the potential to empower GPs, wherever they are located, to play a greater role across the spectrum of pediatric health care from prevention, early diagnosis, management of common problems, through to chronic disease management. For children, this could reduce the need for families to seek specialist referrals for routine issues, potentially saving health care costs and wait times and leading to better long-term health outcomes. GPs may also provide recommendations for interventions for children while they are awaiting pediatrician appointments, potentially averting negative consequences of prolonged nonintervention, such as in neurodevelopmental or behavioral conditions.

Using the same rigorous mixed methods approach as for SC4C, including a stepped wedge trial design, we will be able to conduct an impact, implementation, and economic evaluation of this integrated virtual GP-pediatrician model of care for children.

There are minimal anticipated risks to patient care, as opposed to the potential benefits, with this integrated GP-pediatrician model of care proposed. However, since the model relies on relationship development between the pediatricians and GPs, it is feasible that personality issues may surface with potential negative consequences to the strengthening of confidence and skills that this model aims to enhance in primary care providers. The use of telehealth and virtual modalities for the development of these relationships may also be a challenge, as opposed to in-person models. These difficulties will be explored particularly through the implementation evaluation, where GPs, practice staff, and pediatricians are involved in focus group discussions or interviews to uncover these underlying issues. Using digital tools for health and education delivery has become more technologically stable and acceptable since the COVID-19 pandemic, and in-person modalities are less affordable and scalable. Although the estimated time available for the pediatrician to provide support for the GPs is based on prior projects, it is conceivable that this is under or overestimated. If underestimated, GPs may find it difficult to obtain the support required to assist with behavior change in practice. If overestimated, this could have financial implications for underused specialist time. These issues will be evaluated carefully, both quantitatively and qualitatively, as well as through the economic evaluation, to determine the most cost-effective way of delivering the intervention. The voice of consumers (patients/families) is obtained through anonymized surveys at baseline and postintervention, with an invitation to patients/families to voluntarily participate in an interview to gather their views on pediatric care provided at GP practices that have participated in the SUSTAIN trial.

Limitations

The stepped wedge design is susceptible to trends over calendar time, and delay in the implementation of the model could potentially decrease motivation to participate and subsequently increase withdrawal, although researchers will maintain regular contact with prospective practices through regular "touch points," newsletters, and emails to support motivation and engagement. Results may not be generalizable due to the general practices self-selecting to participate. In a real-world setting, cooperation of GP practices in such a study is crucial to engagement, and we actively target recruitment of a diverse range of practices through PHNs and GP networks, aiming to reach GP practices that care for underrepresented or harder-to-reach population groups, for example, remote, regional, rural general practices. General practices that only use BP or MD EMR software are included for compatibility with GHRANITE, although these EMRs are used by ≈90% of general practices in Australia [27]. Extracted data that are collected from GP EMR may not represent the true scope of care provided by the GPs who also make notes in free text, which could provide a more comprehensive assessment of care quality provided. Nevertheless, the GHRANITE data extraction tool provides a robust mechanism and is the only one of which we are aware that can extract useful data from GP practices in Australia. The voice of consumers relies on voluntary agreement to respond to pre- and postintervention surveys as well as to participate in a postintervention interview. The voice of consumers is important, and we recognize the inherent risk posed by poor participation. Nevertheless, participants are aware that this is voluntary, and their data are provided anonymously. Budget constraints also have meant that family surveys are restricted to English-speaking populations, limiting the generalizability of our findings outside these groups. To highlight the consumer voice, in further implementation and scale-up of this integrated model of care, consumer engagement will need to be a focus, including prioritizing non–English-speaking and other priority population consumers. Financial constraints also limit the collection of sustainability data beyond the period post intervention, which may underestimate the ongoing need for support between GPs and pediatricians in an ongoing manner.

XSL•FO
RenderX

Conclusions

This study provides the potential to deliver quality pediatric care for patients across diverse settings and closer to where they live, in an equitable and scalable fashion. If this innovative integrated care approach is found to be effective and cost-effective, this could be adapted in different settings across Australia and beyond.

Acknowledgments

SUSTAIN is a research collaboration between The University of New South Wales, Sydney Children's Hospital Network, and the University of Melbourne. We thank our partnering primary health networks (Central Eastern Sydney Primary Health Network, South Western Sydney Primary Health Network, and South New South Wales Primary Health Network), all general practices, general practitioners, pediatricians, and families who will be participating in the trial.

Data Availability

The datasets generated or analyzed during this study will be made available from the corresponding author on reasonable request.

Authors' Contributions

The original trial design was conceived by RL, AC, TMM, DB, HH, NH, CCG, KL, MH, NT, LS and CM. NH provided statistical expertise in the trial design, and NH and SHS provided the statistical analysis plan. MH, CCG, NT, and CM provided expertise on the implementation evaluation framework.TMM, KW, AC, KL, LA,, KP, SB and CM contributed to the refinement of the trial protocol paper, and all authors approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
GP participation sheet and GP consent form. GP: general practitioner.
[PDF File, 136 KB - resprot_v15i1e69728_app1.pdf]

Multimedia Appendix 2
Family survey control and intervention period.
[PDF File, 365 KB - resprot_v15i1e69728_app2.pdf]

Multimedia Appendix 3
Sustain Referral popup box.
[PDF File, 81 KB - resprot_v15i1e69728_app3.pdf]

Multimedia Appendix 4
GP survey control and intervention period. GP: general practitioner.
[PDF File, 432 KB - resprot_v15i1e69728_app4.pdf]

Multimedia Appendix 5
GP, pediatrician, practice manager, staff interview information sheet, and consent. GP: general practitioner.
[PDF File, 150 KB - resprot_v15i1e69728_app5.pdf]

Multimedia Appendix 6
Family survey plain-language information sheet.
[PDF File, 138 KB - resprot_v15i1e69728_app6.pdf]

Multimedia Appendix 7
Family interview information sheet and consent.
[PDF File, 125 KB - resprot_v15i1e69728_app7.pdf]

Checklist 1
SPIRIT checklist.
[PDF File, 118 KB - resprot_v15i1e69728_app8.pdf]

XSL•FO
RenderX

References

1. Ruzangi J, Blair M, Cecil E, et al. Trends in healthcare use in children aged less than 15 years: a population-based cohort study in England from 2007 to 2017. BMJ Open 2020 May 5;10(5):e033761. [doi: 10.1136/bmjopen-2019-033761] [Medline: 32371509]
2. Use of emergency departments for lower urgency care. Australian Institute of Health and Welfare (AIHW). 2024. URL: https://www.aihw.gov.au/reports/primary-health-care/use-of-emergency-departments-lower-urgency-care/contents/lower-urgency-care/summary [accessed 2024-11-01]
3. Freed G, Gafforini S, Carson N. Age-related variation in primary care-type presentations to emergency departments. Aust Fam Physician 2015 Aug;44(8):584-588. [Medline: 26510148]
4. Demetriou EA, Boulton KA, Thapa R, et al. Burden of paediatric hospitalisations to the health care system, child and family: a systematic review of Australian studies (1990-2022). Lancet Reg Health West Pac 2023 Nov;40:100878. [doi: 10.1016/j.lanwpc.2023.100878] [Medline: 38116503]
5. Disparities in potentially preventable hospitalisations across australia, 2012-13 to 2017-18. Australian Institute of Health and Welfare (AIHW). 2020. URL: https://www.aihw.gov.au/getmedia/20bc5bf9-d46c-40a7-96c1-d632a1d448bc/aihw-hpf-50.pdf?v=20230605173952&inline=true [accessed 2024-11-01]
6. Freed GL, Spike NA, Sewell JR, et al. Changes in longer consultations for children in general practice. J Paediatrics Child Health 2013 Apr;49(4):325-329. [doi: 10.1111/jpc.12157]
7. Ellis LA, Wiles LK, Selig R, et al. Assessing the quality of care for paediatric depression and anxiety in Australia: a population-based sample survey. Aust N Z J Psychiatry 2019 Oct;53(10):1013-1025. [doi: 10.1177/0004867419866512] [Medline: 31394909]
8. Braithwaite J, Hibbert PD, Jaffe A, et al. Quality of health care for children in Australia, 2012-2013. JAMA 2018 Mar 20;319(11):1113-1124. [doi: 10.1001/jama.2018.0162] [Medline: 29558552]
9. Homaira N, Wiles LK, Gardner C, et al. Assessing appropriateness of paediatric asthma management: a population-based sample survey. Respirology 2020 Jan;25(1):71-79. [doi: 10.1111/resp.13611] [Medline: 31220876]
10. Bayram C, Harrison C, Charles J, Britt H. “The kids are alright” - changes in GP consultations with children 2000-15. Aust Fam Physician 2015 Dec;44(12):877-879. [Medline: 27054203]
11. Charles J, Harrison CM, Britt H. Management of children’s psychological problems in general practice 1970-1971, 1990-1991 and 2008-2009. Aust N Z J Psychiatry 2011 Nov;45(11):976-984. [doi: 10.3109/00048674.2011.610743] [Medline: 22017686]
12. General practice health of the nation 2023. Royal Australian College of General Practitioners (RACGP).: Royal Australian College of General Practitioners (RACGP); 2023. URL: https://www.racgp.org.au/getmedia/122d4119-a779-41c0-bc67-a8914be52561/Health-of-the-Nation-2023.pdf [accessed 2024-11-30]
13. Naehrig D, Glozier N, Klinner C, et al. Determinants of well-being and their interconnections in Australian general practitioners: a qualitative study. BMJ Open 2022 Jul 18;12(7):e058616. [doi: 10.1136/bmjopen-2021-058616] [Medline: 35851015]
14. Australian institute of health and welfare (AIHW). Rural Remote Health. URL: https://www.aihw.gov.au/reports/rural-remote-australians/rural-and-remote-health [accessed 2024-11-01]
15. Arefadib N, Moore T. Reporting the health and development of children in rural and remote Australia Victoria. The Royal Children’s Hospital Melbourne. 2017. URL: https://www.rch.org.au/uploadedFiles/Main/Content/ccchdev/CCCH-Reporting-the-Health-and-Development-of-Children--in-Rural-and-Remote-Aus-September-2017.pdf [accessed 2024-11-01]
16. Henry R. Review of health services for children, young people and families within the NSW health system. NSW Health, Australia. 2019. URL: https://www.health.nsw.gov.au/kidsfamilies/paediatric/Documents/henry-review.pdf [accessed 2024-09-10]
17. Wolfe I, Forman J, Cecil E, et al. Effect of the children and young people’s health partnership model of paediatric integrated care on health service use and child health outcomes: a pragmatic two-arm cluster randomised controlled trial. Lancet Child Adolesc Health 2023 Dec;7(12):830-843. [doi: 10.1016/S2352-4642(23)00216-X] [Medline: 37866369]
18. Montgomery-Taylor S, Watson M, Klaber R. Child health general practice hubs: a service evaluation. Arch Dis Child 2016 Apr;101(4):333-337. [doi: 10.1136/archdischild-2015-308910] [Medline: 26699536]
19. Hiscock H, O Loughlin R, Pelly R, et al. Strengthening care for children: pilot of an integrated general practitioner-paediatrician model of primary care in Victoria, Australia. Aust Health Rev 2020 Aug;44(4):569-575. [doi: 10.1071/AH19177] [Medline: 32045563]
20. Crespo-Gonzalez C, Hodgins M, Zurynski Y, et al. Advancing integrated paediatric care in Australian general practices: qualitative insights from the SC4C GP-paediatrician model of care. PLoS One 2024;19(5):e0302815. [doi: 10.1371/journal.pone.0302815] [Medline: 38771818]
21. Hemming K, Haines TP, Chilton PJ, Girling AJ, Lilford RJ. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. BMJ 2015 Feb 6;350:h391. [doi: 10.1136/bmj.h391] [Medline: 25662947]
22. Sydney child health program - essential paediatrics: the Sydney children’s hospital network. The Sydney Children’s Hospital Network. 2024. URL: https://schneducation.com.au/essential-paediatrics [accessed 2024-09-30]

XSL•FO
RenderX

23. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. Int J Nurs Stud 2024 Jun;154:104705. [doi: 10.1016/j.ijnurstu.2024.104705] [Medline: 38564982]
24. Primary health networks Canberra, Australia. Australian Government DOH, Disability and Ageing. 2024. URL: https://www.health.gov.au/our-work/phn [accessed 2024-10-01]
25. GRHANITE. University of Melbourne. 2024. URL: https://medicine.unimelb.edu.au/school-structure/general-practice-and-primary-care/research/research-groups/health-and-biomedical-research-information-technology-unit/grhanite [accessed 2024-09-30]
26. Nguyen L, Stoové M, Boyle D, et al. Privacy-preserving record linkage of deidentified records within a public health surveillance system: evaluation study. J Med Internet Res 2020 Jun 24;22(6):e16757. [doi: 10.2196/16757] [Medline: 32579128]
27. RACGP digital tools in general practice survey. The Royal Australian College of General Practitioners (RACGP). 2021. URL: https://www.racgp.org.au/getmedia/bfc46468-0480-4944-b436-b1d96065216f/2021-RACGP-Digital-tools-in-general-practice-survey-1.pdf.aspx [accessed 2025-10-08]
28. Wiles LK, Hooper TD, Hibbert PD, et al. Clinical indicators for common paediatric conditions: processes, provenance and products of the CareTrack Kids study. PLoS One 2019;14(1):e0209637. [doi: 10.1371/journal.pone.0209637] [Medline: 30625190]
29. Gaudet-Blavignac C, Foufi V, Bjelogrlic M, Lovis C. Use of the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) for processing free text in health care: systematic scoping review. J Med Internet Res 2021 Jan 26;23(1):e24594. [doi: 10.2196/24594] [Medline: 33496673]
30. Freed GL, Turbitt E, Kunin M, Gafforini S, Sanci L, Spike N. General practitioner perspectives on referrals to paediatric public specialty clinics. Aust Fam Physician 2016 Oct;45(10):747-753. [Medline: 27695726]
31. Freed GL, Turbitt E, Kunin M, Gafforini S, Sanci L, Spike N. Children referred for specialty care: parental perspectives and preferences on referral, follow-up and primary care. J Paediatr Child Health 2017 Jan;53(1):18-25. [doi: 10.1111/jpc.13321] [Medline: 27696573]
32. Khano S, Sanci L, Woolfenden S, et al. Strengthening Care for Children (SC4C): protocol for a stepped wedge cluster randomised controlled trial of an integrated general practitioner-paediatrician model of primary care. BMJ Open 2022 Sep 28;12(9):e063449. [doi: 10.1136/bmjopen-2022-063449] [Medline: 36171040]
33. Chapko MK, Liu CF, Perkins M, Li YF, Fortney JC, Maciejewski ML. Equivalence of two healthcare costing methods: bottom-up and top-down. Health Econ 2009 Oct;18(10):1188-1201. [doi: 10.1002/hec.1422] [Medline: 19097041]
34. Husereau D, Drummond M, Augustovski F, et al. Consolidated health economic evaluation reporting standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. BMC Med 2022 Jan 12;20(1):23. [doi: 10.1186/s12916-021-02204-0] [Medline: 35022047]
35. Torrance G. Methods for the Economic Evaluation of Health Care Programmes, th edition: Oxford Academic; 2005. URL: https://academic.oup.com/book/54294 [accessed 2025-12-09]
36. Petrou P. Budget impact analysis: can we afford the added value? J Med Econ 2021;24(1):487-489. [doi: 10.1080/13696998.2021.1886108] [Medline: 33544634]
37. Hodgins M, Zurynski Y, Burley J, et al. Protocol for the implementation evaluation of an integrated paediatric and primary care model: strengthening care for children (SC4C). BMJ Open 2022 Sep 28;12(9):e063450. [doi: 10.1136/bmjopen-2022-063450] [Medline: 36171022]
38. Consolidated framework for implementation research CFIR research team-center for clinical management research. Consolidated Framework for Implementation Research. 2022. URL: www.cfirguide.org [accessed 2024-11-01]
39. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
40. Smith JD, Li DH, Rafferty MR. The implementation research logic model: a method for planning, executing, reporting, and synthesizing implementation projects. Implement Sci 2020 Sep 25;15(1):84. [doi: 10.1186/s13012-020-01041-8] [Medline: 32988389]
41. Satherley RM, Green J, Sevdalis N, et al. The children and young people's health partnership Evelina London model of care: process evaluation protocol. BMJ Open 2019 Sep 3;9(8):e027302. [doi: 10.1136/bmjopen-2018-027302] [Medline: 31481367]
42. Beames JR, Lingam R, Boydell K, et al. Protocol for the process evaluation of a complex intervention delivered in schools to prevent adolescent depression: the future proofing study. BMJ Open 2021 Jan;11(1):e042133. [doi: 10.1136/bmjopen-2020-042133]
43. May CR, Mair F, Finch T, et al. Development of a theory of implementation and integration: normalization process theory. Implement Sci 2009 May 21;4(29):29. [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
44. Finch TL, Girling M, May CR, et al. Improving the normalization of complex interventions: part 2 - validation of the NoMAD instrument for assessing implementation work based on normalization process theory (NPT). BMC Med Res Methodol 2018 Nov 15;18(1):135. [doi: 10.1186/s12874-018-0591-x] [Medline: 30442094]

XSL•FO
RenderX

45. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
46. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
47. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012 Sep 4;345:e5661. [doi: 10.1136/bmj.e5661] [Medline: 22951546]
48. Baazeem M, Kruger E, Tennant M. Current status of tertiary healthcare services and its accessibility in rural and remote Australia: a systematic review. Health Sciences Review 2024 Jun;11:100158. [doi: 10.1016/j.hsr.2024.100158]
49. Unleashing the potential of our health workforce- scope of practice review canberra, australia. Australian Government Department of Health and Aged Care. 2024. URL: https://www.health.gov.au/sites/default/files/2024-11/unleashing-the-potential-of-our-health-workforce-scope-of-practice-review-final-report_0.pdf [accessed 2024-10-01]

Abbreviations

BP: best practice
CFIR: Consolidated Framework for Implementation Research
CI: chief investigator
CONSORT: Consolidated Standards of Reporting Trials
ED: emergency department
EMR: electronic medical record
GP: general practitioner
MD: medical director
MNAR: missing not at random
NSW: New South Wales
OP: outpatient
PHN: primary health network
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SC4C: Strengthening Care for Children
SCHN: Sydney Children's Hospital Network
SCHP: Sydney Child Health Program
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
UNSW: University of New South Wales
UoM: University of Melbourne

Edited by J Sarvestan; submitted 12.Dec.2024; peer-reviewed by GA Kingston, L Pickard; accepted 27.Oct.2025; published 14.Jan.2026.

Please cite as:
Meyers Morris T, Hiscock H, Wheeler K, Miller C, Hu N, Hassan Shire S, Crespo-Gonzalez C, Bullock S, Taylor N, Boyle D, Sanci L, Lawson K, Adams L, Peacock K, Hodgins M, Christie A, Lingam R
Strengthening Care for Children Using a Virtual Integrated General Practitioner–Pediatrician Model of Primary Care (SUSTAIN): Protocol for a Stepped Wedge Cluster Randomized Controlled Trial
JMIR Res Protoc 2026;15:e69728
URL: https://www.researchprotocols.org/2026/1/e69728
doi:10.2196/69728

© Tammy Meyers Morris, Harriet Hiscock, Karen Wheeler, Corin Miller, Nan Hu, Shukri Hassan Shire, Carmen Crespo-Gonzalez, Susan Bullock, Natalie Taylor, Douglas Boyle, Lena Sanci, Kenny Lawson, Louisa Adams, Ken Peacock, Michael Hodgins, Annemarie Christie, Raghu Lingam. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 14.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical and Cost-Effectiveness of Blended Cognitive Behavioral Therapy or Psychodynamic Therapy Versus Face-to-Face Psychotherapy for Depression (BLENDED Study): Protocol for a Pragmatic, Multicenter, Assessor-Blinded Randomized Controlled Noninferiority Trial

Patrick Luyten[1,2], PhD; Danielle Speybrouck[2], MSc; Peter Martin[3], PhD; Stephan Claes[4], PhD, MD; Nick Midgley[1], PhD; Hester van Eeren[5], PhD; Jan van Busschbach[5], PhD; Eileen Tang[2], PhD

[1]Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom

[2]Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102 box 3722, Leuven, Belgium

[3]Institute of Epidemiology and Health Care, University College London, London, UK

[4]University Psychiatric Center KU Leuven, Universitair Ziekenhuis Leuven, Leuven, Belgium

[5]Erasmus University Rotterdam, Rotterdam, The Netherlands

Corresponding Author:
Patrick Luyten, PhD
Research Department of Clinical, Educational and Health Psychology, University College London, London, United Kingdom

Abstract

Background: Depression is a highly prevalent disorder. Yet, there is still a considerable treatment gap because of capacity issues across clinical services, which create barriers to access to effective psychological therapies. In addition, many individuals with depression do not seek treatment, and waiting lists for psychotherapy are typically very long. Blended psychotherapy, which combines online components and in-person sessions, may help bridge the treatment gap as a cost-effective intervention that complements other types of treatment for depression, as it may reduce therapist time and potentially lower the threshold for people to seek treatment for their depression.

Objective: This study aims to investigate the clinical effectiveness and cost-effectiveness of blended psychodynamic therapy (PDT) and cognitive behavioral therapy (CBT) for depression compared with face-to-face (FTF) PDT and CBT.

Methods: A pragmatic, single-blind, multisite, noninferiority trial will randomize adult patients referred to mental health care centers in Flanders, Belgium, and diagnosed with major depressive disorder (MDD; n=504), to FTF or blended PDT and CBT. The primary outcome is to investigate whether blended therapy for depression (ie, blended PDT and CBT) is noninferior from baseline to 6-month follow-up after treatment termination compared with FTF PDT and CBT in terms of severity of depression assessed with the Beck Depression Inventory-II (BDI-II) based on intention-to-treat analyses. Secondary outcomes include severity of depression as measured with the BDI-II at 1 and 2 years after treatment termination; recovery from depression as assessed with the Structured Clinical Interview for DSM-5 disorders – Clinical Trials Version (SCID-5-CT) and the Patient Health Questionnaire-9 (PHQ-9) at treatment termination and at 6-month, 1-year, and 2-year follow-up; and quality of life as measured with the EuroQoL 5-Dimension 5-Level (EQ-5D-5L) at treatment termination and at 6-month, 1-year, and 2-year follow-up. The feasibility of implementing blended care will be investigated, and health economic analyses will address the cost-effectiveness of blended care versus FTF psychotherapy. Exploratory analyses will focus on possible predictors of treatment outcome and mechanisms of change. Sensitivity analyses will address the potential impact of the COVID-19 pandemic on therapeutic outcomes. Finally, a qualitative substudy aims to address patients' and therapists' subjective experience of blended psychotherapy.

Results: The study was funded in July 2018, and the first patient was included in April 2019. As of September 2025, we have enrolled 463 patients. The first data lock (primary outcome) will take place in October 2025, and the results of the primary outcome are expected in February 2026. The second data lock is expected in March 2027, and the results of the 2-year follow-up are expected in September 2027.

Conclusions: This trial promises to inform decisions concerning the implementation of blended versus FTF therapy for individuals with depression in routine clinical care.

Trial Registration: ClinicalTrials.gov NCT04337242; https://clinicaltrials.gov/study/NCT04337242

International Registered Report Identifier (IRRID): DERR1-10.2196/80511

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e80511) doi:10.2196/80511

KEYWORDS

depression; cognitive behavioral therapy; psychodynamic therapy; blended; cost-effectiveness

Introduction

Background

Depression is a highly prevalent and disabling disorder that is associated with high psychosocial and economic costs. Population-based studies suggest a lifetime prevalence of approximately 15% for unipolar depression and up to 25% in women [1,2], with some evidence suggesting increases in prevalence in recent decades [3], particularly in young people [4-6]. However, there is still a considerable treatment gap because of capacity issues across clinical services, which create barriers to access to effective psychological therapies [7-10]. In addition, studies show that many people with depression do not seek treatment or delay seeking treatment [8,11,12].

There is, therefore, an urgent need to develop and implement (cost-)effective treatments for depression in routine clinical practice and decrease the barriers to treatment. Internet-based therapy, particularly when offered in a blended format, in which online work is integrated into and alternated with in-person face-to-face (FTF) sessions, may help to bridge this treatment gap as a potentially cost-effective treatment that complements other treatments for depression [7,9,10,13-16]. Blended psychotherapy has the potential to lower the threshold for people with depression to seek help, as studies have suggested that people feel that they are more in charge of their own therapeutic process with blended treatment compared with traditional FTF therapy [15,17]. Moreover, blended treatment for depression typically reduces therapist time and thus increases capacity [9,18,19]. Hence, the implementation of blended treatments for depression in routine clinical care may lead to increased availability of psychotherapy for depression and a more effective use of available resources in mental health care.

Although a number of recent high-quality studies and meta-analyses have suggested that blended psychotherapy may be as effective as traditional FTF psychotherapy for depression [9,20-30], there is a need for further research, particularly concerning the (cost-)effectiveness of blended psychotherapy for depression in patients with more complex presentations [17,19,22,31-37]. Moreover, little is known about the effects of blended psychotherapy for depression in the long term [21,22,31,33]. Furthermore, most research in this area has investigated the effectiveness of blended cognitive behavioral therapy (CBT), although there is also growing evidence for the effectiveness of other types of internet-delivered treatments, including psychodynamic therapy (PDT) [38]. Most studies of internet-delivered PDT have focused on adolescents with depression and involved guided internet-delivered PDT, which contained lower levels of therapist guidance than in blended treatment [39-41]. Finally, it is largely unknown which factors may moderate or mediate therapeutic change in blended treatment for depression [15,27,30], and studies have suggested that although therapists and patients both report advantages associated with blended psychotherapy, they also experience disadvantages and barriers in relation to blended psychotherapy for depression in terms of acceptability and suitability [28,42-44].

This study, therefore, set out to investigate the (cost-)effectiveness of blended PDT and CBT for depression compared with FTF PDT and CBT in adults diagnosed with major depressive disorder (MDD) from baseline to 12-month follow-up, in a pragmatic, single-blind, multisite, noninferiority trial in secondary mental health care centers in Belgium. Given the recurring nature of depression and the relatively high relapse rates following brief treatments for depression [45,46], patients will also be followed up until 2 years after the end of treatment. A health economic evaluation will address the potential cost-effectiveness of blended psychotherapy versus FTF psychotherapy for depression. Exploratory analyses will focus on potential moderators and mediators of change. A qualitative substudy will investigate both patients' and therapists' experiences of blended treatment, with a focus on acceptability, credibility, and satisfaction with treatment.

In what follows, we discuss the emerging evidence for the effectiveness of internet-based psychotherapy, with a focus on blended treatments. We then provide an outline of this study, methods, and expected results, reported according to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [47].

Internet-Based Psychotherapy and the Need to Improve Access to Psychological Therapies

Meta-analyses and qualitative reviews suggest that internet-based psychotherapy may be as effective as FTF psychotherapy for MDD both at treatment termination and at follow-up in the medium to long term [15,21-23,27,33,35,37]. Research in this domain to date has mainly focused on 2 types of internet-based psychotherapy, CBT [15,48] and PDT [38,49]. Although few trials have directly compared these 2 types of internet-based therapy, existing studies suggest that they have similar effects [38,40]. These findings are also consistent with findings on FTF psychotherapy, with meta-analyses showing no substantial differences in the effectiveness of brief CBT and PDT for MDD [50-54].

Research findings have also suggested that guided internet-delivered treatments may be associated with better outcomes and lower dropout rates compared with unguided or minimally guided online treatments [17,22,28,55,56]. These findings have sparked increasing interest in blended psychotherapy; this type of internet-delivered treatment may combine the "best of both worlds" in that it involves alternating FTF therapy sessions and online modules. In this format, therapists are able to attend to so-called process-related components of psychotherapy (ie, building a therapeutic alliance, explaining the treatment approach, and addressing specific needs, concerns, and anxieties), which may increase adherence to treatment and reduce dropout [10,28,43,57]. More

XSL•FO
RenderX

content-related components of the treatment can be the focus of the online modules, which may be flexibly integrated among the FTF sessions. Moreover, personal contact with a therapist may broaden the spectrum of patients who can be treated with blended psychotherapy compared with unguided or minimally guided internet-delivered treatments [17,58]. Blended psychotherapy may also foster feelings of autonomy and empower patients to a greater extent than FTF therapy, as the availability of online modules may better enable patients to work on their presenting problems between FTF sessions, which may further facilitate the translation of insights from therapy to their everyday lives [59]. Finally, blended psychotherapy may also be more cost-effective than FTF psychotherapy, as it typically involves less therapist time compared with FTF therapy [9,18,19,60].

More research is needed concerning the (cost-)effectiveness, acceptability, and suitability of blended treatment for depression for a number of reasons. First, the number of studies comparing the effectiveness of blended psychotherapy for depression with FTF psychotherapy is relatively small [21,22,33], and even fewer studies have compared their cost-effectiveness [60]. For instance, a recent meta-analysis [22] included only 9 randomized controlled trials (RCTs) of blended psychotherapy for depression versus control conditions, with considerable heterogeneity among studies in terms of the kind of blended treatments offered (with digital sessions being either integrated or added as supplementary treatment), the ratio of FTF sessions to online modules, and the number of FTF sessions and online modules. Second, the majority of studies of blended psychotherapy for depression have been conducted in controlled research settings, and thus it remains uncertain to what extent these findings can be generalized to more complex patients typically seen in routine clinical care [35,58]. Hence, there is a need for more research on the suitability of internet-based treatments for different types of patients with depression and to identify factors that foster or hinder therapeutic change in these types of treatments [15,37,40,44,61,62]. In this context, it is noteworthy that there is still very little research on predictors of outcome or mechanisms of change in blended versus FTF psychotherapy for depression, although research on these issues could have important implications for assessing the suitability of blended versus FTF psychotherapy for different patients.

Finally, there is an emerging body of evidence from qualitative studies focusing on the subjective experience of both patients and therapists involved in blended care for depression, but more research is needed in this context as well. Studies on internet-based interventions in general tend to highlight the potential importance of human interaction and support in explaining treatment outcomes in blended treatments across different therapeutic modalities [63-65], whereas a lack of understanding, trust, and support has been associated with dropout and lack of improvement [66-71]. Whether this is also the case for blended treatment for depression is largely unknown, given the small number of studies. Moreover, almost all of these studies have examined internet-based CBT [15,38]. Only a few qualitative studies have investigated the role of therapists and the therapeutic relationship in other treatment modalities, such as PDT, and most of these studies have focused on adolescents with depression [62,72,73].

This Study

The primary objective of this study is to investigate whether blended therapy for depression (ie, blended PDT and CBT) is noninferior from baseline to 6-month follow-up after treatment termination compared with FTF PDT and CBT in adults diagnosed with MDD (target n=504). FTF psychotherapy will consist of 16 sessions of PDT or CBT, offered over a period of 6 months. Blended therapy will consist of, on average, 8 FTF sessions, alternating with 8 online modules that patients can complete between FTF sessions, similarly offered over a period of 6 months to control for time effects. The primary outcome assessment will focus on noninferiority in terms of severity of depression assessed with the Beck Depression Inventory-II (BDI-II) [74] from baseline to 6 months after the end of treatment (ie, 12 months after the start of treatment), based on intention-to-treat (ITT) analyses. The noninferiority margin was defined as a difference in effect size (ES) equal to or smaller than 2 points on the BDI-II between blended care and FTF psychotherapy from baseline to 6-month follow-up after treatment termination. Given an estimated baseline SD of 10, this is equivalent to a standardized ES of Cohen $d \leq 0.2$, which is below the minimal meaningful difference in psychosocial interventions for depression proposed as Cohen $d=0.24$ [75].

The secondary objectives of this study are 6-fold. First, the trial aims to investigate the comparative efficacy of blended psychotherapy versus FTF psychotherapy on the following secondary outcomes: severity of depression as measured with the BDI-II at 1 and 2 years after treatment termination; recovery from depression as assessed with the Structured Clinical Interview for DSM-5 disorders – Clinical Trials Version (SCID-5-CT) [76] and the Patient Health Questionnaire-9 (PHQ-9) [77] at treatment termination and at 6-month, 1-year, and 2-year follow-up; and quality of life as measured with the EuroQoL 5-Dimension 5-Level (EQ-5D-5L) [78] at treatment termination and at 6-month, 1-year, and 2-year follow-up.

Second, we aim to investigate the feasibility of implementing blended care in the participating mental health care centers in Belgium, including the acceptability of treatments to both therapists and patients. To this end, we will record the recruitment rate, retention in treatment, treatment adherence, and adherence to the research protocol. Acceptability will be indicated by the number of sessions attended, including the number of individuals who refuse treatment, and feasibility will be indicated by the number of patients who do not comply with the full clinical and research protocol. Patients will also complete measures of credibility of the treatments and satisfaction with the treatments. To measure patient attitudes toward the treatment programs, we will administer the Credibility and Expectancy Questionnaire (CEQ) [79] at baseline and an adapted version of this measure at posttreatment and 6-month follow-up. Satisfaction with treatment will be measured with the Client Satisfaction Questionnaire-8 (CSQ-8) [80]. Treatment integrity of therapists will be assessed on the basis of independent ratings of audio- or video-recorded psychotherapy sessions.

XSL•FO
RenderX

Third, the cost-effectiveness of blended care versus FTF psychotherapy will be investigated based on a societal perspective, taking all relevant costs and effects into account (intervention costs, direct and indirect medical costs, as well as productivity losses and costs incurred elsewhere in the health care system) based on both self-reported data, using an adapted version of the Treatment Inventory of Costs in Patients with psychiatric disorders (TIC-P) [81], and data provided by the Rijksinstituut voor Ziekte- en Invaliditeitsverzekering/Institut National d'Assurance Maladie-Invalidité (RIZIV/INAMI), the Belgian federal national health insurance body.

Fourth, exploratory analyses will focus on possible predictors of treatment outcome and mechanisms of change. Consistent with the focus of this study on patients who are typically referred to specialized mental health care centers, these analyses will focus on patient features that have often been linked to negative outcomes in the treatment of depression [57,82,83], such as severity of depression (assessed with the BDI-II and PHQ-9), dependent and self-critical personality features (assessed with a brief version of the Depressive Experiences Questionnaire [DEQ]) [84], and patients' experience of both types of treatment as assessed with the CEQ [79]. Concerning mechanisms of change, relationships between changes in dependent and self-critical personality features assessed with the DEQ [84] and primary and secondary outcomes will be investigated from both a variable-centered and a person-centered perspective.

Fifth, sensitivity analyses will address the potential impact of the COVID-19 pandemic on therapeutic outcomes.

Sixth, a qualitative substudy aims to address patients' and therapists' subjective experience of blended psychotherapy, with a focus on treatment suitability and the advantages and disadvantages of blended therapy. This qualitative study aims to capture the subjective experiences and understandings of both patients and therapists, including what they may find helpful or hindering in blended psychotherapy. The study thus also includes a strong service-user perspective, which is often neglected in psychotherapy outcome research. Qualitative research regarding the subjective experience of blended psychotherapy for depression may also provide policymakers and clinicians with detailed and in-depth information about how to contextualize, and thus better understand, findings from RCTs, which will help them effectively adapt and implement blended psychotherapy in routine clinical care in the future [85].

Methods

Trial Design

This study is a pragmatic, single-blind, multicenter, randomized controlled, noninferiority, multiple-armed trial, with patients being followed up for up to 2 years after treatment termination. The primary interest lies in comparing the efficacy of blended psychotherapy versus FTF psychotherapy for depression at 6-month follow-up after treatment termination. The overall expectation is that blended psychotherapy is noninferior to FTF psychotherapy but more cost-effective. The noninferiority margin is defined as a small difference in ES between blended and FTF psychotherapy (2 points on the BDI-II), which is deemed clinically meaningful in psychotherapy research [75,86]. Additionally, the study design allows exploration of the relative effectiveness of the 2 different types of psychotherapy (ie, PDT vs CBT) in both forms provided (blended and FTF), although the trial is not powered for this purpose, and this is not the main aim of the study. A qualitative substudy will be embedded within the trial, focusing on the subjective experiences of blended therapy of both therapists and patients. Semistructured interviews, based on an adapted version of the Experience of Therapy Interview [87], will be conducted with 16 patients who found blended psychotherapy to be either helpful or unhelpful. Focus groups will be conducted with participating therapists to explore their experiences of delivering blended psychotherapy. Qualitative data will be analyzed using thematic analysis (TA) based on a combined deductive–inductive approach.

Study Setting

The study will be conducted in specialized, state-funded secondary care mental health care centers (Centra Geestelijke Gezondheidszorg [CGGs]) in Flanders, Belgium. About half of patient referrals to CGGs are made by primary care medical practitioners or specialized mental health professionals, about one-third are self-referrals, and the remaining patients are referred by schools, family, or friends, or by the Department of Justice. We anticipate that a maximum of 15 CGGs will participate in the study. Factors that might hamper accrual are expected to be mainly related to the severity and complexity of the presenting problems of patients at CGGs, with a substantial number of patients needing longer-term treatment and/or counseling. Access to the internet might be a challenge for some depressed patients seen at these centers. However, 99.9% of people in Belgium potentially have access to broadband internet (>30 Mbps). In 2015, when this study was planned, 81.8% of households had access to the internet, and most of those had access to broadband internet. Access to the internet was slightly less prevalent among people living in poverty in Belgium. Patients seeking treatment for depression overall have higher levels of poverty than the general population. This is especially the case for patients referred to CGGs, with an estimated 40% - 45% of these patients meeting criteria for poverty (defined as having low levels of education and being unemployed). However, data from 2015 showed that, of people living in poverty in Belgium, 72% had internet access, and this proportion has increased year on year. For the majority of patients in this study, therefore, internet access is not expected to differ greatly from that of the general population (ie, minimum 72% for patients vs 81.8% in the general population, based on 2015 data), and thus it is likely to have a relatively limited impact on inclusion rates [88].

Participants

In total, 504 patients with MDD will be included. Based on a review of referrals over the 3 years preceding the trial in the participating CGGs, we anticipate a recruitment rate of approximately 168 patients with MDD who can be included per 6 months (ie, approximately 28 patients per month, approximately 34 patients per CGG in the 18-month recruitment period, and approximately 2 patients per month per CGG).

XSL•FO
RenderX

Eligibility Criteria

Inclusion and exclusion criteria for the trial are depicted in Textbox 1. Patients excluded from the trial will receive standard care in one of the participating centers or will be referred to specialist care, as appropriate.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Aged 18-65 years.
- Current diagnosis of major depressive disorder (MDD) with or without dysthymic disorder according to *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*) criteria [89].
- Patient Health Questionnaire-9 (PHQ-9) score ≥10.
- The use of pharmacotherapy for depression and other psychiatric disorders during the intervention is allowed (with the exception of antipsychotic medication for psychotic disorder; see exclusion criteria), as this is a pragmatic trial and many patients seen in routine clinical care use pharmacotherapy.

Exclusion criteria:

- Current psychotic symptoms or bipolar disorder.
- Current use of antipsychotic medication specifically for the treatment of primary psychotic disorder.
- Severe personality disorder (eg, borderline personality disorder).
- Historic or current self-injury/parasuicide of such an extent and/or severity that it may substantially interfere with the ability to engage in brief psychotherapy.
- Current excessive use of drugs/alcohol.
- Lack of Dutch language fluency.
- Clinical contraindication to brief psychotherapy (eg, attachment history characterized by multiple separations, serious ongoing trauma in childhood, or multiple caregivers, suggesting the need for longer-term psychotherapy).
- Evidence of pervasive use of help.
- Highly unstable or insecure life arrangements.
- No access to a computer/internet or computer illiteracy.
- Participation in another depression clinical trial within the past year in which the participant had received CBT or PDT.

For the qualitative study, patients will be recruited based on purposive sampling to identify participants who found blended therapy either helpful or unhelpful. After completing the 6-month follow-up assessment, patients from the blended psychotherapy conditions will be asked to rate the level of helpfulness of blended psychotherapy on a scale from 0 (Not at all helpful) to 100 (Extremely helpful). They will mark their score on a form that is sent to them by email, after completion of a separate informed consent procedure specific to the qualitative study. Patients with scores either below 35 or 65 or higher will be eligible to participate in the qualitative study and will be categorized as those who found blended therapy to be "not helpful" or "helpful," respectively. We will thus recruit an equal number of participants from each group from both blended CBT and blended PDT, to give a total of 16 patients (4 "helpful" and 4 "not helpful" cases each from blended CBT and blended PDT). Therapists for the focus groups will be contacted via email and/or telephone. All therapists who have participated in the trial will be eligible for the qualitative substudy.

Recruitment, Screening, and Gaining Informed Consent

Each participating mental health care center will be responsible for identifying potential patients who meet eligibility criteria for the study. Identification of potential participants will occur at the point of triage or initial assessment, depending on the specific procedures of the participating centers, as part of routine clinical care. Specifically, the clinician conducting initial assessments of patients referred for mood problems will consider the inclusion and exclusion criteria for the study. It is acknowledged that the assessing clinician may identify exclusion criteria or other factors that could deem a patient unsuitable for the study (eg, risk of possible injury to therapists, imminent suicidal risk) even if the patient meets the inclusion criteria. Unless this is the case, the clinician will offer the patient a brief verbal introduction to the study, constituting a preliminary invitation to participate. The clinician will provide the patient with a participant information sheet to read and keep. The participant information sheet will include age- and culturally appropriate material explaining the study and what involvement in the study would entail at that particular site. Clinicians then confirm whether the patient consents to being contacted by a researcher from the research team, who will then discuss the study in more detail with the patient. Patients will be given ample time (minimum 48 hours) to think about the information provided and can request more time to think about the study before consenting to further contact from the researcher. If a patient verbally consents to being contacted by a research assistant (RA), the clinician informs the research team of this decision and provides relevant contact information.

XSL•FO
RenderX

An RA will contact potential participants no less than 48 hours after they receive information about the study from participating centers, to explain what participation might involve and to arrange an appointment at the mental health center. A study identification number will be assigned to the participant, and from this point on, the participant will be referred to only by this number, and their collected data will be labeled with it. After the participant has provided explicit and written informed consent, an RA will administer the SCID-5-CT interview and the PHQ-9 and will assess all other inclusion/exclusion criteria. When all the measures have been completed, and study eligibility has been confirmed at baseline (ie, checks for inclusion/exclusion criteria have been confirmed), an RA will provide the trial coordinator (TC) with the details required to randomize the participant to one of the four conditions.

For the qualitative substudy, an RA will invite (first via email and then followed up by telephone) all patients who received blended treatment in the BLENDED trial who are beyond the 6-month follow-up assessment to find out whether they are interested in participating in an interview study focusing on their experience of blended psychotherapy. A minimum of 48 hours will be allowed between the initial contact with the participant and confirmation of their participation (by telephone or email) and the participant being approached again by the RA. An information leaflet and an additional informed consent form (ICF) will be sent by email to the patient following this first contact. This schedule is to allow sufficient time for individuals to consider and discuss participation in the substudy with any party they may choose to consult. An RA will confirm with the potential participant that this time frame has been sufficient and allow additional time if requested. As the substudy will be conducted online and not at the mental health care centers, patients will be asked in a separate contact to sign and email the ICF to the research team; after receipt, the RA will sign the ICF and return the fully signed ICF to the patient. As mentioned earlier, eligibility (ie, whether they found blended psychotherapy helpful or not) will be determined based on patients' self-reported levels of helpfulness of blended treatment when they are approached to participate in this study. If a patient is eligible for the qualitative study, an RA will contact the patient again to schedule an online interview. For the focus group study with therapists, a similar procedure will be followed to obtain informed consent from the therapists.

Randomization

This trial will use stratified and automated randomization of participants to one of four conditions in permuted blocks built into the electronic Case Report Form (eCRF). Stratification will be carried out in 2 stages. At stage 1, patients will be randomized to therapists within each participating center, using permuted blocks to ensure that the maximum difference of trial caseload for therapists within the same center is on average 2 patients. As each therapist will provide only either CBT or PDT, stage 1 implicitly allocates patients to the type of therapy. At stage 2, patients will be randomized to either FTF or blended care, again using permuted blocks such that the maximum difference in the number of FTF and blended therapy patients for each therapist will be 2 patients at any one point in time. Participant randomization will be initiated centrally by the TC, following the completion of the eligibility assessment (screening) by the RA and of the baseline measures by the participant, and before commencement of any trial intervention. As randomization will be performed using an automated web-based randomization algorithm built into the eCRF, which sends an automated email to the therapist, this will ensure that the RA remains blind to the treatment condition to which the patient is randomized.

This 2-stage stratified randomization scheme ensures balance for the comparison of FTF and blended therapy with respect to participating centers and therapists. The scheme will also promote, although not guarantee, balance with respect to any patient-level prognostic factors whose distribution varies between the patient populations of the centers (eg, age).

Interventions

Both CBT and PDT, as offered in this trial, are based on manualized treatments developed in the context of the Improving Access to Psychological Therapies program in the United Kingdom [90,91]. FTF PDT and CBT consist of 16 FTF sessions, offered over a period of 4 - 6 months. Blended PDT and CBT consist of 8 FTF sessions in combination with 8 online modules offered over a period of 4 - 6 months and thus are similar in length to the FTF psychotherapies. Blended PDT and CBT are based on principles outlined in the manualized FTF psychotherapies and adapted for use in Flanders (Belgium) [17,92] in collaboration with service users and therapists as part of depressiehulp.be, a government-supported online platform that offers psychoeducation and (both unguided and guided) internet-delivered treatment for depression.

FTF sessions in FTF therapy and blended therapy are offered on a weekly and a 2-weekly basis, respectively, although with mutual agreement by the patient and therapist, the session frequency may deviate from the planned weekly or 2-weekly sessions.

The content of FTF psychotherapy and blended therapy is very similar, as the treatments follow the same treatment manuals and treatment phases. The main difference between FTF and blended therapy is that in blended therapy, the therapists will be specifically trained to help motivate and encourage patients to complete the online modules between the FTF sessions, and to take the patient's experience in completing each module as the starting point of the next FTF session. Hence, using the terminology introduced by Erbe et al [23], in this study, blended therapy will be offered in a fully integrated format, with digital modules being a core component of the treatment, FTF and digital sessions alternating, and blended treatment primarily being delivered in a standardized format for each patient, with some room for personalization of the treatment as detailed below.

FTF and blended PDT will be based on dynamic interpersonal therapy (DIT), an integrative evidence-based brief treatment for depression [90,93]. FTF DIT involves 16 weekly sessions of individual therapy, which formulates the presenting symptoms of depression as responses to interpersonal difficulties or perceived threats to attachments (loss and/or separation) and hence also as threats to the self. DIT conceptualizes depression in terms of an underlying temporary disorganization of the

XSL•FO
RenderX

attachment system caused by current relationship problems, which in turn generates a range of distortions in thinking and feelings typical of depression. In DIT, there is a focus on both content and process. The content focus relates to the content of the individual dynamics that are associated with the onset and maintenance of depression. This focus helps the patient understand the connection between their presenting symptoms and what is happening in their relationships by identifying a core, repetitive pattern of relating that becomes the focus of the therapy, which is known as the interpersonal affective focus (IPAF). The process focus involves fostering the patient's capacity for reflective functioning or mentalizing, as depression is typically associated with serious impairments in this capacity. In DIT, therapists use a range of interventions and techniques to restore the capacity for reflective functioning and to foster insight and change in the IPAF, consisting of supportive, mentalizing, insight-oriented, and directive interventions to empower the patient to translate insights obtained in therapy to their everyday life.

DIT consists of 3 phases. In the first phase (sessions 1 - 3), the focus is on engaging and motivating the patient, assessing their presenting problems and dynamics, and collaboratively delineating the focus of the treatment (ie, the IPAF). The middle phase (sessions 4 - 12) focuses on exploring how the patient's typical pattern of relating to the self and others has led to feelings of depression, and how a change in this pattern might improve their mood. The ending phase (sessions 13 - 16) is mainly concerned with empowering the patient to translate the insights they have gained in therapy to their everyday life.

The blended DIT program consists of the same 3 phases as FTF DIT: an exploration phase, a middle or working-through phase, and an ending phase. Instead of having 16 FTF sessions, patients in blended DIT are invited to complete online modules between 8 FTF sessions to explore the typical dynamics that are otherwise explored in the FTF sessions, and to discuss these in subsequent FTF sessions. The structure and phases of blended DIT are discussed in detail elsewhere [92].

FTF CBT involves 16 sessions offered over 4 - 6 months. CBT is based on the assumption that depression is directly related to patterns of thinking [91,94]. Specifically, dysfunctional and often automatic patterns of thinking are assumed to be related to the onset and maintenance of depression. The overall goal of CBT is therefore to help a person recognize such patterns of thought, evaluate their validity, and replace them with more adaptive ways of thinking. In CBT, there is an emphasis not only on thinking (ie, "cognitive"), but also on doing and thus on changing behavior ("behavioral").

Therapeutic sessions in FTF CBT are highly structured, with clear tasks and goals in each session. These tasks and goals draw from more classic cognitive behavioral principles, including psychoeducation about depression, cognitive restructuring (ie, identifying and challenging maladaptive thoughts and beliefs, fostering problem-solving capacities), behavioral activation, and relapse prevention. Therapists are allowed to use these principles flexibly in a manner that is tailored to the needs and preferences of individual patients. In this trial, a focus on more recent principles and techniques that have emerged within the CBT tradition, such as fostering mindfulness and applying principles of acceptance and commitment therapy, is also part of both FTF and blended CBT [91], but these are covered in only one FTF session or one online module.

The structure and focus of FTF CBT and blended CBT are similar. The only difference is that, as with FTF and blended PDT, in blended CBT, 8 of the sessions are replaced by online treatment modules, so that the focus in the FTF sessions is more on attending to the process components of treatment (ie, is the patient motivated? Do they understand the treatment rationale?) and on reviewing and discussing therapeutic progress. The structure and phases of blended CBT in this trial are discussed in detail elsewhere [17].

As a result of the early phase of the COVID-19 pandemic, various restrictions made physical attendance of screening and therapeutic sessions impossible. An amendment to the original trial protocol, replacing FTF in-person sessions with online sessions, was allowed during this phase of the pandemic and particularly during mandated restrictions on meeting in person (so-called "lockdowns;" see protocol amendments in Multimedia Appendix 1). Sensitivity analyses will address the potential effects on treatment outcomes of the fact that some patients had their FTF sessions online (see also the Statistical Analysis Plan section).

As this is a pragmatic trial, if the patient and therapist agree at the end of the treatment program that further treatment is necessary, the treatment can be continued with the same therapist, or the patient can be referred for additional and/or longer-term therapy based on mutual agreement between patient and therapist. In the former case, the number of additional sessions with the same therapist is recorded in the eCRF; in the latter case, additional psychotherapy is monitored using the TIC-P.

Therapists will be certified CBT and PDT psychotherapists (psychiatrists or psychologists), who must have completed a postgraduate psychotherapy training of 3 - 4 years. In addition, they will receive a 4-day training program in manualized FTF and blended CBT or PDT delivered by certified trainers in CBT and PDT, respectively. They will also receive 2-weekly small-group supervision of at least 2 patients by certified CBT and PDT supervisors.

Discontinuation and Withdrawal

Participants can withdraw from the study for any reason at any given time. Therapists can propose to withdraw a patient from the study (1) if they are of the opinion that a patient risks serious deterioration and nonresponse to the treatment to which they have been allocated (ie, the patient needs longer-term treatment), and/or (2) if it becomes clear that the patient does not meet inclusion criteria or meets one or more exclusion criteria, despite initial screening and assessment (eg, the patient does have severe personality disorder), and/or (3) if participation of the patient in the study seriously interferes with treatment (eg, the patient is overly burdened by having to complete assessments). In all of these cases, the therapist needs to discuss the option of withdrawal from the study with the patient and must seek advice

XSL•FO
RenderX

from a site supervisor and from the chief investigator. Patients who withdraw from the study for any reason are not replaced.

Measures

The following measures will be completed by patients during the trial as specified in the CONSORT diagram (Figure 1; also see Table 1). Visit windows are 2 months either side of the ideal follow-up time point.

The primary outcome will be the severity of depression as measured by the BDI-II [74] assessed at 6-month follow-up after treatment termination. The BDI-II consists of 21 items, in line with the depression criteria of the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*) [89]. The BDI-II will also be administered at 1-year and 2-year follow-up for secondary outcome analyses.

Secondary outcomes include:

- Recovery (defined as no longer scoring above the clinical cut-off for complete or partial recovery), assessed at treatment termination and at 6-month, 1-year, and 2-year follow-up by using the SCID-5-CT [76]. The SCID-5-CT is customized to include only those diagnostic elements that are needed to determine whether the patient fulfills the inclusion and exclusion criteria of a particular clinical trial. Recovery will also be assessed with the PHQ-9, a brief, 9-item multipurpose instrument for screening, diagnosing, and monitoring the severity of depression [77]. It incorporates the *DSM-5* diagnostic criteria for depression and is extensively used in research on depression and in routine clinical practice.
- Quality of life as measured with the EQ-5D-5L [78] at baseline, treatment termination, and at 6-month, 1-year, and 2-year follow-up. The EQ-5D-5L contains 5 items, measuring quality of life in 5 dimensions of health (mobility, self-help, habitual activities, pain, and anxiety/depression), each with 5 levels reflecting “no problems,” “slight problems,” “moderate problems,” “severe problems,” and “extreme problems.” The EQ Visual Analog Scale records the respondent’s self-rated health on a 20 cm vertical visual analog scale with endpoints labeled “the best health you can imagine” and “the worst health you can imagine.” The EQ-5D-5L will also be used for the cost-effectiveness component of the study.

As indicators of the acceptability of the treatment, we will also measure treatment expectancy (at baseline), treatment credibility (at follow-up), and satisfaction with treatment:

- Patients’ beliefs about the credibility of treatment and their expectations of the effectiveness of treatment will be measured with the CEQ [79] at baseline, and with a modified version of the CEQ at treatment termination and 6-month follow-up. The CEQ has 6 questions, rated on a 9-point Likert-type scale (items 1 - 3 and item 5) or a 0% - 100% scale (items 4 and 6).
- Satisfaction with treatment will be measured at baseline, treatment termination, and at 6-month, 1-year, and 2-year follow-up, using the CSQ-8 [80]. This questionnaire contains 8 questions that are scored on a 4-point scale.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram of the BLENDED trial. CBT: cognitive behavioral therapy; PDT: psychodynamic therapy; FTF: face-to-face.

XSL•FO
RenderX

Table . Overview of the BLENDED trial procedures based on SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials).

	Study period						
	Enrollment	Allocation	Postallocation				Close-out
Timepoint	$-t_1$	0	During treatment	End of treatment	Follow-up at 6 months	Follow-up at 1 year	Follow-up at 2 years
				6 months after start of treatment (±2 months)	12 months after start of treatment (±2 months)	18 months after start of treatment (±2 months)	30 months after start of treatment (±2 months)
Enrollment							
Eligibility screen: PHQ-9[a]; SCID-5-CT[b]	✓						
Informed consent	✓[c]						
Allocation		✓[d]					
Interventions							
FTF[e] DIT[f]			✓				
Blended DIT			✓				
FTF CBT[g]			✓				
Blended CBT			✓				
Assessments							
Demographics		✓					
Concomitant medications		✓		✓	✓	✓	✓
BDI-II[h]		✓		✓	✓	✓	✓
SCID-5-CT				✓	✓	✓	✓
PHQ-9		✓		✓	✓	✓	✓
TIC-P[i]		✓		✓	✓	✓	✓
EQ-5D-5L[j]		✓		✓	✓	✓	✓
CEQ[k]		✓		✓	✓		
CSQ-8[l]		✓		✓	✓	✓	✓
Safety assessment		✓	✓	✓	✓		
Qualitative interview					✓[m]		

[a]PHQ-9: Patient Health Questionnaire-9.

[b]SCID-5-CT: Structured Clinical Interview for DSM-5 Disorders – Clinical Trials version.

[c]Before eligibility screen.

[d]After baseline assessments.

[e]FTF: face-to-face.

[f]DIT: dynamic interpersonal therapy.

[g]CBT: cognitive behavioral therapy.

[h]BDI-II: Beck Depression Inventory-II.

[i]TIC-P: Treatment Inventory of Costs in Patients with Psychiatric Disorders.

[j]EQ-5D-5L: EuroQoL 5-Dimension 5-Level.

[k]CEQ: Credibility and Expectancy Questionnaire.

[l]CSQ-8: Client Satisfaction Questionnaire-8.

[m]Purposeful sample.

XSL•FO
RenderX

As part of the health economic evaluation, patients will complete the TIC-P [81], which has been used by the authors in previous RCTs [95,96]. The TIC-P is a generic questionnaire designed for self-report by patients with a psychiatric disorder and consists of 2 parts. The first part assesses relevant contacts with health care (eg, contacts with psychotherapists, inpatient care, visits to the primary care medical practitioner, etc). The second part consists of the Short Form-Health and Labor Questionnaire (SF-HLQ), which is designed to collect data on productivity loss resulting from health problems. Additionally, 3 questions are included to measure productivity losses due to reduced efficiency during paid work, if appropriate (see the Health Economic Evaluation section for more details). In addition, direct health-related costs will be calculated based on reimbursement data from RIZIV/INAMI, the national health service body in Belgium. Specifically, after the completion of the study, the sponsor will request approval from the competent chamber of the Information Security Committee to have the relevant study data linked with IMA (InterMutualistisch Agentschap/Agence InterMutualiste) data by a trusted third party (TTP, eHealth platform) using the participant's national number. As soon as the approval has been received, the sponsor will collaborate with the eHealth platform and IMA for the data linkage process. This data linkage is planned to obtain a dataset containing costs related to health care paid by the compulsory health insurance and to compare these costs with self-reported direct and indirect costs assessed by the TIC-P, to arrive at a more comprehensive assessment of the cost-effectiveness of the interventions investigated in this trial.

Serious Adverse Events

Only the serious adverse events (SAEs) listed below, occurring from the time of the participant providing written informed consent until 90 days after cessation of the trial treatment, will be recorded in the eCRF, and the sponsor of the study will be notified within 24 hours of the research staff at the site becoming aware of the event.

For each SAE, the following information will be collected:

- Full details in medical terms and case description.
- Event duration (start and end dates, if applicable).
- Action taken.
- Outcome.
- Seriousness criteria.
- Causality (ie, relatedness to trial intervention/investigation), in the opinion of the research team.
- Severity (mild, moderate, or severe).
- Whether the event would be considered expected or unexpected.

Any change of condition or other follow-up will be recorded, and the sponsor will be notified of such changes as soon as the information is available, preferably within 1 working day. Events will be followed up on until the event has resolved or a final outcome has been reached by recording on a detailed Safety Reporting Form. All SAEs will also be reported to the Trial Steering Committee (TSC). In the event that there are 5 or more SAEs within a period of 6 months, or any other condition that may give rise to serious concerns regarding the occurrence of SAEs (eg, when there is reason to suspect that the experimental condition may lead to more SAEs), an unblinded evaluation by an independent statistician and review by an independent psychiatrist or psychologist may be demanded by the TSC and reported back to the TSC. Depending on the outcome of this independent evaluation, the TSC can advise the research team to end the trial should there be sufficient concern about the welfare of patients in the trial.

The following SAEs will be monitored, recorded, and reported:

- Lethal suicide attempts and/or suicide attempts that require hospitalization.
- Any other behavior that may be related to depression that requires inpatient hospitalization or that may result in persistent or significant disability/incapacity (eg, serious self-harm, overdose).

Blinding

In trials of psychosocial interventions, therapists and patients cannot be blinded to the type of intervention they provide and receive, respectively. To minimize the risk of potential biases in the assessment of outcomes arising from knowledge of treatment allocation, the main outcome data will be collected by an RA who will be unaware of each participant's treatment allocation. Thus, participants will be asked before all postrandomization interviews with the RA not to reveal anything about their treatment. RAs and therapists will not share the same room, and all therapists involved in the trial will be instructed by the TC not to reveal treatment allocations for participants in any communication with the RAs. The RAs will additionally be asked at each assessment point to guess which treatment was given to each participant, so that the effects of possible bias can be examined in the analysis. Finally, to minimize the risk of bias in relation to treatment knowledge becoming apparent in RAs' scoring of interview data, RAs' interviews may be audio-taped, and a random sample may be re-rated by an assessor who has no knowledge of treatment allocation.

First analyses of the trial data will be performed only after all patients have completed the 6-month follow-up. At this point, a data lock will be performed and signed off by the research team and each of the involved mental health care centers, and the locked database will be transferred to an independent trial statistician, as will be attested by (1) the audit trail of the eCRF database and (2) a unique CRC32 hashing code (ie, a checksum algorithm that hashes byte sequences to 32-bit values). All codes will remain unbroken, and all researchers, clinicians, and anyone else involved in the trial will remain blind to treatment outcomes until that time. RAs will remain blind to the treatment condition of patients until all patients have completed the 2-year follow-up.

For the qualitative study, blinding of RAs to treatment condition is not possible, as only patients who have been randomized to the blended condition are interviewed and RAs will also know whether participants have been selected on the basis of how successful they perceived their treatment to have been; it is also highly likely that the latter type of information will be mentioned during the interview. Likewise, RAs will have been in contact with participating therapists, so they will know the therapists

XSL•FO
RenderX

and at least to some degree therapists' opinions and attitudes with regard to blended treatment.

Statistical Plan and Data Analysis

Sample Size and Power Calculation

The primary outcome in this trial is noninferiority of blended care, compared with FTF therapy, in depression as measured by the BDI-II. Sample size and statistical power calculations are based on planned analyses involving the BDI-II. Meta-analytic data suggest a medium ES (Cohen *d*/Hedges *g*=.50) of FTF psychotherapy and blended psychotherapy compared with control conditions, with no differences between FTF or blended PDT and CBT [50,55,97-101]. Our analysis will focus on the difference in effects between blended and FTF treatment and assume that there is no difference between PDT and CBT regarding the effectiveness of either blended or FTF treatment. Thus, the PDT and CBT groups will be combined in the primary analysis.

The tolerance margin for this noninferiority trial was defined to be a maximum of 2 points on the BDI-II. Assuming an SD of 10 points, this is equivalent to a standardized ES of Cohen *d*=0.2, which is considered a minimal meaningful difference in psychosocial interventions [75].

The primary endpoint is the BDI-II score at 6-month follow-up. Noninferiority will be assessed by a mixed-effects analysis of covariance (ANCOVA) that adjusts for therapist variance, baseline BDI-II score, and other baseline covariates. The primary analysis model assumed by the power analysis is:

$$BDIij=(\alpha+uj)+(\beta)BASELINEij+(\gamma)BLENDEDij+(\delta)xijT+\varepsilon ij$$

where:

- BDIij is the BDI-II score at the primary endpoint of patient *i* treated by therapist *j*.
- *j=1, …, J* indicates the therapists.
- *i=1, …, n_j* indicates the patients; n_j denotes the number of patients treated by the jth therapist.
- α is an intercept.
- uj~N(0,σu2) is a random intercept accounting for differences in outcomes between therapists.
- εij~N(0,σ2) is an error term that accounts for between-patient variability in outcomes.
- BASELINEij is the baseline BDI-II score, which is related to BDIij by slope coefficient β.
- BLENDEDij is a dummy variable coded 1=Blended psychotherapy and 0=FTF psychotherapy; the associated coefficient γ estimates the difference in primary outcome between the two groups, adjusted for pretreatment scores and therapist effects.
- δxijT represents further baseline information to be included in the model as controls, where xijT denotes a column vector of covariate values, and δ denotes a row vector of coefficients.

We will conduct a 1-sided test of noninferiority on the coefficient γ, which represents the difference (after adjustment) between the blended and FTF treatment groups. Formally, our hypotheses will be:

- H0: γ≥2
- H1: γ<2

The null hypothesis of inferiority will be assessed by a 1-sided *t* test, as is usual for ANCOVA models. The null hypothesis will be rejected if 1-sided *P*<.025, as recommended for noninferiority trials [102].

Power calculations were performed using the standard formula for a 1-sided *t* test for noninferiority trials [103]. The alternative hypothesis used for power calculations was HA:γ=0, that is, exact equivalence of blended and FTF treatment. The sample size was adjusted to take account of the expected within-patient and within-therapist correlations as follows:

$$n=n*(1-rwithin-patient2)(1-ICCtherapist)$$

where *n** is the sample size required for 80% power without adjustment, and *n* is the actual sample size required in this trial, taking the adjustment into account. Since all other baseline covariates should be approximately equally distributed between randomized groups, their effect on statistical power was assumed to be negligible. Power calculations were confirmed by a user-written simulation in R (R Foundation for Statistical Computing) [104], with 10,000 replications. All simulations agreed with calculations to within 1 percentage point. We made the following assumptions: the within-patient correlation was assumed to be $r_{within\text{-}patient}$=0.6, which is close to the median value of 0.59 estimated by a meta-analysis of within-patient correlations [105]. We further assumed a therapist effect of $ICC_{therapist}$=0.05, which is the current best estimate for therapist effects in trials of psychological treatments [106].

We assumed that 36 therapists would take part in the study. We calculated the smallest sample size needed to achieve at least 80% power, assuming that each therapist treats the same number of patients. This was n=504, that is, 14 patients per therapist. This design has 82% power according to our formula (our simulation estimated the power to be 81.6%). Although in reality, the number of patients is unlikely to be exactly the same for each therapist, a simulation study has shown that even far more extreme deviations from equal numbers than are expected in this trial have little effect on the statistical power of mixed-effects models, especially when the intraclass correlation coefficient is small [107].

Statistical Analysis Plan

Overview

In what follows, the overall data analytic strategy is summarized. A detailed prespecified statistical analysis plan (SAP) for the primary and secondary analyses, as well as moderator and mediator analyses, is provided in Multimedia Appendix 2. This prespecified SAP outlines the analyses concerning primary and secondary outcomes and has not changed substantially compared with the first protocol (version 1.1) approved by the funder on October 25, 2018, that is, before the start of the trial (see also the protocol amendments in Multimedia Appendix 1). These analyses will begin to be performed after all the data from the primary endpoint (6-month follow-up) have been entered, checked, and locked. As outlined earlier, all researchers, clinicians, and anyone else involved in the trial will remain

XSL•FO
RenderX

blinded to treatment outcomes until that time. Further analyses focusing on 1-year and 2-year outcomes will be performed after all the data from the 2-year follow-up have been entered, checked, and locked.

Primary and Secondary Outcome Analysis

The primary analysis will be an ITT analysis, that is, all participants will be analyzed as randomized, even if they dropped out of treatment or changed treatment during the study period. The primary endpoint is 6 months after the end of treatment. The analysis will focus on the difference between the blended and FTF groups at the primary endpoint:

- Adjusting for patients' baseline BDI-II score.
- Adjusting for therapy type (CBT or PDT).
- Taking into account between-therapist variation.
- Taking into account variation between treatment centers (as a fixed effect).

A linear mixed-effects model (3-level ANCOVA, with therapists, time points, and patients as levels) will be estimated to investigate the null hypothesis that blended treatment yields inferior outcomes to FTF psychotherapy by 2 points or more on the BDI-II, after adjusting for pretreatment scores and therapist effects. This is equivalent to a standardized ES of 0.2, given the expected baseline SD of 10 for the BDI-II. For details, see the SAP.

This model can accommodate missing values of the outcome variable at either of the 2 time points included in the analysis. Estimates are valid provided that values are missing at random (MAR) conditional on the other variables included in the model. This procedure has been shown to lead to unbiased treatment effect estimates under the MAR assumption and to more efficient estimates compared with multiple imputation [1,108]. The model will be fitted using restricted maximum likelihood estimation.

Analogous to the primary analysis, we will use a mixed-effects logistic regression model to estimate the difference in recovery rates between blended and FTF psychotherapy, and a mixed-effects linear regression model to estimate the difference in participants' quality of life.

A number of sensitivity analyses will be conducted:

- Using complete cases only. This amounts to estimating Model 1 on participants who provided valid information on the primary outcome at the 6-month follow-up. To be included in this analysis, participants do not need to have provided information on the primary outcome at the end of treatment. Within this sensitivity analysis, item missingness will be handled as described above.
- The assumption of equivalence of effect of blended CBT and PDT will be investigated by estimating a model identical to Model 1 in all respects except that we will allow for an interaction between type of therapy and treatment condition.
    - Whether the conclusions from the primary analysis model (Model 1) would hold if we instead estimated separate treatment effects (blended versus FTF) for CBT and PDT, respectively.
    - Adjusting for concerning imbalances in baseline characteristics.
- Impact of COVID-19. We will document the impact of the COVID-19 pandemic on the delivery of treatments (eg, moving FTF sessions to a remote format), describe the frequency of remote sessions, and conduct exploratory analyses to evaluate whether the difference in effectiveness between the blended and FTF conditions varies depending on whether some FTF sessions were conducted remotely.

Supportive analogous analyses will also be conducted for the 1-year and 2-year follow-up endpoints once they are available after the second database lock, as for the primary endpoint, including all previous endpoints in the linear mixed-effects models.

Furthermore, we will conduct exploratory analyses to determine the shape of the relationship between BDI-II scores and time. On this basis, we will develop a longitudinal mixed-effects model and investigate whether the rate of change in BDI-II differs between (1) blended versus FTF psychotherapy, (2) CBT versus PDT, and (3) the interaction of the two. For the secondary outcomes, we will also explore longitudinal mixed-effects logistic regression to estimate trajectories of recovery rates and longitudinal mixed-effects linear regression to estimate trajectories of quality of life.

Moderators and Mechanisms of Change

We will conduct exploratory analyses to investigate potential treatment effect moderators on (1) the primary outcome (BDI-II) at 6-month follow-up, and at 1-year and 2-year follow-up, (2) the secondary outcomes at 6-month follow-up, and at 1-year and 2-year follow-up, and (3) the difference in effectiveness between blended versus FTF treatment, and PDT versus CBT, at 6-month follow-up, and at 1-year and 2-year follow-up. The following moderators will be considered:

- Severity of depression at baseline (assessed by the BDI-II).
- Expectancy of treatment at baseline (assessed by the CEQ).
- Credibility of treatment at baseline (assessed by the CEQ).
- Depressive personality: dependency at baseline (assessed by the DEQ).
- Depressive personality: self-criticism at baseline (assessed by the DEQ).

To investigate these potential moderator effects, we will include all of these potential moderators as predictors in our primary outcome analysis model and assess 95% CIs of the relevant coefficient.

Exploratory analyses will also focus on potential relationships between changes in dependent and self-critical personality features assessed with the DEQ and primary and secondary outcomes, from both a variable-centered and a person-centered perspective.

Per-Protocol Analysis

We will repeat the primary analysis as a per-protocol analysis instead of an ITT analysis, using either Model 1 or, if missing outcome values are present, the modification of Model 1 described above. Only participants who have completed the treatment will be included in the per-protocol analysis.

XSL•FO
RenderX

Treatment completion is defined as attending at least 8 sessions in the FTF treatments, or attending at least 4 FTF sessions and completing at least 4 modules in the blended treatments.

Treatment Fidelity, Feasibility, Acceptability, and Satisfaction

Analyses on treatment fidelity/integrity, feasibility, acceptability, and satisfaction will focus on recruitment rate, retention in treatment, treatment adherence, and adherence to the research protocol, as well as experienced credibility of the treatments and patient satisfaction.

Fidelity/Integrity

The integrity of the provision of treatment by therapists will be assessed on the basis of independent ratings of recordings of therapeutic sessions based on the DIT Competence Rating Scale [93] for FTF PDT and an adapted version for blended PDT, and on the Cognitive Therapy Adherence and Competency Scale (CTACS) [109] for FTF CBT. An adapted version of the CTACS will be used for assessing adherence to blended CBT.

Feasibility

The feasibility of future studies will be assessed via the proportion of patients who comply with the full clinical and research protocol. Full compliance with the clinical and research protocol is defined as completion of treatment and participation in research interviews at all time points up to the 6-month follow-up. Treatment completion is defined as attending a minimum of 8 sessions in the FTF treatments, or attending at least 4 FTF sessions and completing at least 4 modules in the blended treatments.

Acceptability

Acceptability of the treatment to patients will be measured by:

- Treatment retention: the proportion of patients who complete the treatment as defined above.
- Treatment adherence: the proportion of planned sessions attended.

Treatment Credibility and Expectancy

Treatment credibility and expectancy will be assessed by using the CEQ [79]. We will document baseline differences (FTF versus blended) in both dimensions and use a multilevel model with random intercepts for sites to test for evidence of systematic differences. We will also compare FTF and blended treatments on credibility and expectancy at the end of treatment and at 6-month follow-up, using regression models analogous to the primary analysis model, including random intercepts for therapists and controlling for baseline differences.

Treatment Satisfaction

Satisfaction with treatment will be measured with the CSQ-8 [80]. Again, we will document baseline differences (FTF versus blended) with an adapted version of the CSQ-8 assessing satisfaction with received care before randomization. A multilevel model with random intercepts for sites will be used to test for evidence of systematic differences. We will also compare FTF and blended treatments on treatment satisfaction at the end of treatment and at 6-month follow-up, using regression models analogous to the primary analysis model, including random intercepts for therapists and controlling for baseline differences.

Qualitative Analyses

Qualitative analyses, as detailed below, will also focus on issues related to treatment acceptability and feasibility, as well as patient satisfaction.

Interim Analyses

There will be no planned SAE analyses unless there are 5 or more SAEs within a time span of 6 months or any other condition that gives rise to serious concerns regarding the occurrence of SAEs (eg, when there is reason to suspect that the experimental condition may lead to more SAEs). In this event, see the section on SAEs. Adherence to protocol, including the treatment protocol, will be monitored continuously. Any deviation from the planned statistical analyses will be reported in an amendment after it has been discussed and approved by the TSC.

Health Economic Evaluation

The economic evaluation will be a “piggyback” economic evaluation using outcomes from the trial. Because economic evaluations have a wider time horizon than the trial, and given that beneficial effects can be expected even after 5 years, a Markov model will be used to project future costs and effects using different time horizons, as the higher initial cost of FTF therapy may outweigh the lower cost of blended provision in the long term. The costs of the implementation, costs of usual care, costs in other sectors, patient and family costs, costs of lost production, and judicial costs will be included by using the TIC-P. For the valuation of health care use, standard prices according to published costing guidelines will be used. Medication use and other direct costs will be monitored using national health insurance data provided by RIZIV/INAMI and valued using the listed prices in Belgium. The treatment costs will be estimated using a bottom-up approach. Personnel costs, implementation costs, and overhead costs will be taken into account. In the economic evaluation, outcomes will be expressed in terms of costs per quality-adjusted life year (QALY). Probabilistic analyses that allow multivariate sensitivity analyses will be used [110].

We will also calculate detailed cost-effectiveness plans of the incremental costs in relation to gained QALYs at different willingness-to-pay thresholds, as a basis for future decision-making concerning the implementation of these interventions in mental health care in Belgium.

Data Collection and Data Analysis in the Qualitative Substudy

Data Collection

Semistructured individual interviews with patients will be conducted based on an adaptation of the Experience of Therapy Interview [87], which we have used in similar studies of both FTF and guided internet-based therapy for young people with depression [62,69,87]. Adaptations include questions focusing on patients’ beliefs about the credibility and treatment effects of blended psychotherapy; concerns about privacy and security

XSL•FO
RenderX

related to the online modules, their experience with the online modules, the FTF sessions, and the way in which the sessions and modules were integrated; their therapist and the therapeutic relationship; and their experiences with regard to helping and hindering factors associated with blended therapy more generally. As in the original Experience of Therapy Interview [87], the questions are broad, open-ended, and exploratory, to allow participants to tell the story of their experience. The interview guide also includes open-ended follow-up questions to use when elaboration is needed (eg, "Could you tell me a bit more about that?" "Do you remember any specific impasse in therapy where…?"). The interview guide will be piloted with 2 patients to determine whether any further improvements are needed. Focus groups will be conducted using a semistructured interview guide based on an adaptation of the Experience of Therapy Interview [87], based on previous qualitative interview studies of therapists' experience of providing blended and guided internet-delivered therapy [41,73].

Interviews with patients will be conducted online or in person by a research psychologist with specific training and experience in conducting in-depth, semistructured interviews online. Interviews are expected to take in the range of 1 - 1.5 hours. Focus groups with the therapists will be conducted online or in person by a research psychologist with specific training and experience in conducting qualitative research. Focus groups are expected to take in the range of 2 - 2.5 hours. All interviews and focus groups will be recorded and transcribed verbatim by an independent RA using Amberscript (Amberscript Global BV), an artificial intelligence-powered software that we have used in previous studies, which is able to automatically transcribe audio- and video-recorded interviews and is fully compliant with General Data Protection Regulation policies.

Data Analysis: Thematic Analysis

Data from the qualitative substudy of patients and therapists will be analyzed separately, using TA, a flexible and theory-neutral approach to the analysis of qualitative data. The main aim of TA is to allow the researcher to approach the data with an open mind, and thus to explore and generate rich, context-specific, so-called "thick descriptions" [111] rather than to test specific theories or a priori expectations. Although TA is based on an inductive, bottom-up approach, it is also open to a more top-down, deductive component [112], as in framework analysis [113]. A combined inductive–deductive (or bottom-up and top-down) approach will be adopted in order to remain open to new and potentially unexpected findings, while at the same time exploring the findings from other studies about the subjective experience of blended psychotherapy discussed above.

Consistent with the general approach for the application of TA in psychology [114], a 6-stage approach will be taken for data analysis of the patient interviews. First (stage 1), 2 members of the research team will read all of the transcripts of the interviews to familiarize themselves with the data while taking notes about potentially important findings, themes, or ideas, all the time staying very close to the raw data. Second (stage 2), transcripts of interviews categorized as "helpful" from patients who received blended CBT or blended PDT will be reread by the same 2 members of the research team, and categories that emerge from these data will be coded systematically by both researchers independently; then, transcripts of interviews categorized as "unhelpful" from blended CBT and blended PDT cases will be coded in the same way. After this process, the categories and themes that emerge from these transcripts will be compared between the 2 researchers and discussed. Any differences in interpretation will be resolved through discussion among the 2 researchers and the wider research team. In stage 3, these preliminary qualitative analyses will be discussed by the 2 researchers and collapsed into different subthemes and overarching main themes. Each of these themes will then be labeled and exemplified with quotes from the transcripts (stage 4). This model will then be reviewed by the research team, and any differences in categorizing or labeling will be resolved by consensus, resulting in a final model (stage 5), which will be written up as such (stage 6). A similar data analytic strategy will be followed for the analyses of the focus groups with therapists.

Ethical Considerations

This trial was approved by the Ethics Committee Research UZ/KU Leuven on January 11, 2019 (approval S59765). This trial was registered on ClinicalTrials.gov (NCT04337242). As mentioned in the Recruitment, Screening, and Gaining Informed Consent section, all patients and therapists will be asked to provide explicit and written informed consent. Personal information will be stored in coded formats only to protect privacy in the study's eCRF; only the treatment centers will have direct access to uncoded personal information. Access to all coded data will be restricted to authorized study personnel, and strict confidentiality protocols will be followed. Data will be stored in secure, password-protected electronic systems. Participants will receive no compensation for their participation in the study.

Monitoring and Audit

A full trial monitoring plan was developed and agreed upon by the Trial Management Group and TSC based on the trial risk assessment, which includes regular on-site monitoring. Monitoring is done completely independently from the Trial Management Group and TSC by the UZ Clinical Trials Center (Leuven).

User Involvement

User involvement is a central component in each stage of the planning and implementation of the trial. Both the trial design and the blended modules were the topic of regular feedback from service users and members of various patient organizations. In addition, the blended modules and online platform used in this study have been the topic of extensive feedback by both therapists and service users since their development. Hence, service users and therapists have been involved at each stage of the adaptation and development of the online modules used in this trial. We will continue to involve service users and therapists during the trial, as we aim to investigate patients' attitudes and experiences in relation to online mental health care and the combination of online modules and FTF sessions. Even though blended therapy may have major economic

XSL•FO
RenderX

benefits, in line with our emphasis on the empowerment of patients, we consider the subjective experience of service users and therapists as central in this trial. A service user representative will also be part of the TSC during the trial.

The 2 overarching themes that emerged from previous rounds of feedback and consultations with service users and patient organizations were the urgent need to increase access to effective treatments for depression in Belgium and to increase the diversity of treatments offered, particularly of psychotherapeutic treatments, as patients felt that choice and the match between patient and type of treatment were very important. This trial heeds this call from service users by evaluating the effectiveness of two different evidence-based treatments for depression (CBT and PDT) offered in two different formats (FTF and blended).

Management of the Study

The chief investigator will be responsible for the overall management and coordination of the trial across all participating sites and will ensure that the study adheres to the protocol, timelines, and regulatory requirements. A TSC composed of clinical and scientific experts provides strategic guidance and decision-making throughout the study. The committee ensures the integrity and scientific validity of the trial.

Results

Inclusions started on April 1, 2019, and were seriously delayed because of COVID-19 restrictions as detailed in the protocol amendments (Multimedia Appendix 1). The trial is currently still ongoing; recruitment ended on December 2, 2024. The first data lock (primary outcome) will take place in October 2025, and the results of the primary outcome are expected in February 2026. The second data lock is expected in March 2027, and the results of the 2-year follow-up are expected in September 2027.

Discussion

Principal Findings

Depression is a highly prevalent mental health condition. Unfortunately, there still is a considerable treatment gap in that many patients with depression do not seek treatment, delay treatment seeking, or do not receive effective treatment because of limited treatment capacity. Blended psychotherapy, which combines online work and in-person FTF sessions, may help to bridge this treatment gap as a potentially cost-effective treatment that complements other treatments for depression [7,9,10,13-16]. Yet, there is still a paucity of research in this area [9,20-30], particularly concerning the (cost-)effectiveness of blended psychotherapy for depression in patients with more complex presentations [17,19,22,31-37]. Furthermore, little is known about the long-term effects of blended psychotherapy [21,22,31,33], and most research in this area has investigated the effectiveness of blended CBT, with less research on other types of internet-delivered treatments, such as PDT [38]. Finally, more research is needed on moderators and mediators of change and both patients' and therapists' experiences of blended treatment.

This multisite trial is expected to provide robust evidence on the effectiveness and cost-effectiveness of blended psychotherapy for depression in adult patients routinely referred to mental health care centers in Flanders, Belgium, compared with traditional FTF psychotherapy. It will also be one of the first trials to directly compare the effectiveness of 2 types of blended psychotherapy (ie, blended CBT and blended PDT) and the effectiveness of these treatments over a clinically meaningful follow-up period of 2 years, which is important given the highly recurrent nature of depression. The trial's implications may also extend to improving treatment personalization based on moderator and mediator analyses. Secondary analyses, together with the qualitative study, will provide a strong service user perspective on the experience of blended psychotherapy in general and blended psychotherapy for depression in particular. These latter analyses will provide policymakers, researchers, and clinicians with detailed, in-depth information concerning the feasibility, acceptability, and satisfaction associated with FTF and blended psychotherapy for depression, in order to better contextualize and understand findings from this RCT. Taken together, the results will likely inform best practices in psychotherapy delivery and support the development of evidence-based guidelines.

The study's large sample size and cross-site design promise to enhance the ecological validity of the findings. Other strengths include the study's rigorous methodology, standardized intervention protocols, and recruitment and treatment delivery in real-world settings. However, potential limitations may include variability in implementation fidelity across sites and challenges in maintaining uniformity in therapist adherence. In addition, results may not generalize to other clinical settings and different health care systems. Despite this, the trial is well positioned to contribute meaningfully to the psychotherapy research landscape and provide a foundation for future studies on scalability and comparative effectiveness.

Dissemination Plan

To ensure broad impact, the dissemination plan will first focus on the publication of peer-reviewed publications in high-impact clinical and psychological journals to target academic audiences. In addition, presentations will be delivered at national and international conferences. Executive summaries and policy briefs will be shared with health care providers, mental health organizations, and policy stakeholders to inform practice and service delivery. Findings will be communicated to trial sites and participating therapists through tailored webinars and workshops. To reach the general public and patients, we will develop accessible content such as infographics and blog posts shared via institutional websites and social media platforms.

XSL•FO
RenderX

Acknowledgments

The authors would like to thank all the patients who participated in this study. They are also grateful to the therapists and mental health care centers involved in the trial for their continuous support of the study. They also thank the sponsor of this study and the UZ Leuven Clinical Trials Center. Finally, they thank the team at the Belgian Health Care Knowledge Centre (KCE) for their unwavering support of the study, especially during and in the aftermath of the COVID-19 pandemic-related restrictions.

Funding

This trial is funded by the Belgian Health Care Knowledge Centre (KCE; KCE-17003) and the American Psychoanalytic Association. The sponsor of this trial is the Universitaire Ziekenhuizen KU Leuven, Belgium. The funders have no role in the study design, data collection, analysis, or interpretation. The authors declare that no generative artificial intelligence was used in the creation of this manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request after publication of results focusing on primary and secondary outcomes and endpoints.

Authors' Contributions

PL, ET, SC, and JvB were responsible for the original proposal, securing funding for the trial, and drafting the original protocol. PL, as chief investigator, had overall responsibility for the management of the study, with support from ET and DS as trial coordinators. PL also had responsibility for training and supervision and responsibility for the implementation of the clinical service across each of the sites. ET and DS were trial coordinators throughout the different phases of the trial and coordinated the randomization protocol. PM conducted the sample size calculation and developed the SAP. JvB and HvE developed the health economics analysis plan. PL, DS, and ET designed the qualitative substudy, with the support of NM. PL wrote the original draft of the manuscript. All authors contributed to and approved the final manuscript.

Conflicts of Interest

PL and ET have received funding for research on both psychodynamic and cognitive behavioral treatments and have received payments for delivering training in these treatments. None declared by the other authors.

Multimedia Appendix 1
Protocol amendments.
[DOCX File, 17 KB - resprot_v15i1e80511_app1.docx]

Multimedia Appendix 2
Statistical analysis plan.
[DOCX File, 107 KB - resprot_v15i1e80511_app2.docx]

Checklist 1
SPIRIT checklist.
[DOC File, 123 KB - resprot_v15i1e80511_app3.doc]

References

1. Alonso J, Angermeyer MC, Bernert S, et al. Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004;109(420):21-27. [doi: 10.1111/j.1600-0047.2004.00327.x] [Medline: 15128384]
2. Blazer DG, Kessler RC, McGonagle KA, Swartz MS. The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. AJP 1994 Jul 1;151(7):979-986. [doi: 10.1176/ajp.151.7.979]
3. Moreno-Agostino D, Wu YT, Daskalopoulou C, Hasan MT, Huisman M, Prina M. Global trends in the prevalence and incidence of depression:a systematic review and meta-analysis. J Affect Disord 2021 Feb 15;281:235-243. [doi: 10.1016/j.jad.2020.12.035] [Medline: 33338841]
4. Goodwin RD, Dierker LC, Wu M, Galea S, Hoven CW, Weinberger AH. Trends in U.S. depression prevalence from 2015 to 2020: the widening treatment gap. Am J Prev Med 2022 Nov;63(5):726-733. [doi: 10.1016/j.amepre.2022.05.014]
5. Shorey S, Ng ED, Wong CHJ. Global prevalence of depression and elevated depressive symptoms among adolescents: a systematic review and meta-analysis. Br J Clin Psychol 2022 Jun;61(2):287-305. [doi: 10.1111/bjc.12333] [Medline: 34569066]

XSL•FO
RenderX

6. McGorry PD, Mei C, Dalal N, et al. The Lancet Psychiatry Commission on youth mental health. Lancet Psychiatry 2024 Sep;11(9):731-774. [doi: 10.1016/S2215-0366(24)00163-9]
7. Simon N, Ploszajski M, Lewis C, et al. Internet-based psychological therapies: a qualitative study of National Health Service commissioners and managers views. Psychol Psychother 2021 Dec;94(4):994-1014. [doi: 10.1111/papt.12341] [Medline: 33788999]
8. Cuijpers P, Stringaris A, Wolpert M. Treatment outcomes for depression: challenges and opportunities. Lancet Psychiatry 2020 Nov;7(11):925-927. [doi: 10.1016/S2215-0366(20)30036-5]
9. Clark DM, Wild J, Warnock-Parkes E, et al. More than doubling the clinical benefit of each hour of therapist time: a randomised controlled trial of internet cognitive therapy for social anxiety disorder. Psychol Med 2023 Aug;53(11):5022-5032. [doi: 10.1017/S0033291722002008]
10. Kleiboer A, Smit J, Bosmans J, et al. European COMPARative Effectiveness research on blended Depression treatment versus treatment-as-usual (E-COMPARED): study protocol for a randomized controlled, non-inferiority trial in eight European countries. Trials 2016 Dec;17(1):27488181. [doi: 10.1186/s13063-016-1511-1]
11. Demyttenaere K. The temporal experience in depression: from slowing down and delayed help seeking to the emergency setting and length of treatment. Int Rev Psychiatry 2022;34(7-8):693-698. [doi: 10.1080/09540261.2022.2141103] [Medline: 36786108]
12. The WHO special initiative for mental health (2019-2023): universal health coverage for mental health. : World Health Organization; 2019 URL: https://iris.who.int/server/api/core/bitstreams/16cd81fb-7b2f-47cf-b3c9-734d6d77f9f4/content [accessed 2025-12-22]
13. Leigh E, Nicol-Harper R, Travlou M, Clark DM. Adolescents' experience of receiving internet-delivered cognitive therapy for social anxiety disorder. Internet Interv 2023 Dec;34:100664. [doi: 10.1016/j.invent.2023.100664] [Medline: 37720914]
14. Leigh E, Clark DM. Internet-delivered therapist-assisted cognitive therapy for adolescent social anxiety disorder (OSCA): a randomised controlled trial addressing preliminary efficacy and mechanisms of action. J Child Psychol Psychiatry 2023 Jan;64(1):145-155. [doi: 10.1111/jcpp.13680] [Medline: 35943064]
15. Andersson G, Titov N, Dear BF, Rozental A, Carlbring P. Internet-delivered psychological treatments: from innovation to implementation. World Psychiatry 2019 Feb;18(1):20-28. [doi: 10.1002/wps.20610] [Medline: 30600624]
16. Thew GR. IAPT and the internet: the current and future role of therapist-guided internet interventions within routine care settings. Cogn Behav Therap 2020;13:e4. [doi: 10.1017/S1754470X20000033] [Medline: 34567240]
17. Buelens F, Luyten P, Claeys H, Van Assche E, Van Daele T. Usage of unguided, guided, and blended care for depression offered in routine clinical care: lessons learned. Internet Interv 2023 Dec;34:100670. [doi: 10.1016/j.invent.2023.100670] [Medline: 37767005]
18. Thase ME, McCrone P, Barrett MS, et al. Improving cost-effectiveness and access to cognitive behavior therapy for depression: providing remote-ready, computer-assisted psychotherapy in times of crisis and beyond. Psychother Psychosom 2020;89(5):307-313. [doi: 10.1159/000508143] [Medline: 32396917]
19. Kemmeren LL, van Schaik A, Draisma S, Kleiboer A, Riper H, Smit JH. Effectiveness of blended cognitive behavioral therapy versus treatment as usual for depression in routine specialized mental healthcare: E-COMPARED Trial in the Netherlands. Cogn Ther Res 2023 Jun;47(3):386-398. [doi: 10.1007/s10608-023-10363-y]
20. Thase ME, Wright JH, Eells TD, et al. Improving the efficiency of psychotherapy for depression: computer-assisted versus standard CBT. AJP 2018 Mar;175(3):242-250. [doi: 10.1176/appi.ajp.2017.17010089]
21. Mamukashvili-Delau M, Koburger N, Dietrich S, Rummel-Kluge C. Efficacy of computer- and/or internet-based cognitive-behavioral guided self-management for depression in adults: a systematic review and meta-analysis of randomized controlled trials. BMC Psychiatry 2022 Nov 24;22(1):730. [doi: 10.1186/s12888-022-04325-z] [Medline: 36424570]
22. Ferrao Nunes-Zlotkowski K, Shepherd HL, Beatty L, Butow P, Shaw JM. Blended psychological therapy for the treatment of psychological disorders in adult patients: systematic review and meta-analysis. Interact J Med Res 2024 Oct 29;13:e49660. [doi: 10.2196/49660] [Medline: 39470720]
23. Erbe D, Eichert HC, Riper H, Ebert DD. Blending face-to-face and internet-based interventions for the treatment of mental disorders in adults: systematic review. J Med Internet Res 2017 Sep 15;19(9):e306. [doi: 10.2196/jmir.6588] [Medline: 28916506]
24. Chow DYW, Jiang X, You JHS. Information technology-based versus face-to-face cognitive-behavioural therapy for anxiety and depression: a systematic review and meta-analysis. J Affect Disord 2022 Aug 1;310(429-40):429-440. [doi: 10.1016/j.jad.2022.05.048] [Medline: 35577156]
25. Moshe I, Terhorst Y, Philippi P, et al. Digital interventions for the treatment of depression: a meta-analytic review. Psychol Bull 2021 Aug;147(8):749-786. [doi: 10.1037/bul0000334] [Medline: 34898233]
26. Andrews G, Basu A, Cuijpers P, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55(70-8):70-78. [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
27. Hedman-Lagerlöf E, Carlbring P, Svärdman F, Riper H, Cuijpers P, Andersson G. Therapist-supported Internet-based cognitive behaviour therapy yields similar effects as face-to-face therapy for psychiatric and somatic disorders: an updated

XSL•FO
RenderX

systematic review and meta-analysis. World Psychiatry 2023 Jun;22(2):305-314. [doi: 10.1002/wps.21088] [Medline: 37159350]
28. Omylinska-Thurston J, Aithal S, Liverpool S, et al. Digital psychotherapies for adults experiencing depressive symptoms: systematic review and meta-analysis. JMIR Ment Health 2024 Sep 30;11:e55500. [doi: 10.2196/55500] [Medline: 39348177]
29. Plessen CY, Panagiotopoulou OM, Tong L, Cuijpers P, Karyotaki E. Digital mental health interventions for the treatment of depression: a multiverse meta-analysis. J Affect Disord 2025 Jan 15;369(1031-44):1031-1044. [doi: 10.1016/j.jad.2024.10.018] [Medline: 39419189]
30. Karyotaki E, Efthimiou O, Miguel C, et al. Internet-based cognitive behavioral therapy for depression: a systematic review and individual patient data network meta-analysis. JAMA Psychiatry 2021 Apr 1;78(4):361-371. [doi: 10.1001/jamapsychiatry.2020.4364] [Medline: 33471111]
31. Andersson G, Rozental A, Shafran R, Carlbring P. Long-term effects of internet-supported cognitive behaviour therapy. Expert Rev Neurother 2018 Jan;18(1):21-28. [doi: 10.1080/14737175.2018.1400381] [Medline: 29094622]
32. Edelbluth S, Klein JP, Schwartz B, et al. The long shadow of early-change patterns: a 3-year follow-up after the use of a web-based intervention for mild to moderate depressive symptoms. Cogn Behav Ther 2024 Nov;53(6):681-700. [doi: 10.1080/16506073.2024.2368520] [Medline: 38912859]
33. Mamukashvili-Delau M, Koburger N, Dietrich S, Rummel-Kluge C. Long-term efficacy of internet-based cognitive behavioral therapy self-help programs for adults with depression: systematic review and meta-analysis of randomized controlled trials. JMIR Ment Health 2023 Aug 22;10:e46925. [doi: 10.2196/46925] [Medline: 37606990]
34. Kenter RMF, van de Ven PM, Cuijpers P, et al. Costs and effects of Internet cognitive behavioral treatment blended with face-to-face treatment: results from a naturalistic study. Internet Interv 2015 Mar;2(1):77-83. [doi: 10.1016/j.invent.2015.01.001]
35. Etzelmueller A, Vis C, Karyotaki E, et al. Effects of internet-based cognitive behavioral therapy in routine care for adults in treatment for depression and anxiety: systematic review and meta-analysis. J Med Internet Res 2020 Aug 31;22(8):e18100. [doi: 10.2196/18100] [Medline: 32865497]
36. Owusu JT, Wang P, Wickham RE, Varra AA, Chen C, Lungu A. Real-world evaluation of a large-scale blended care-cognitive behavioral therapy program for symptoms of anxiety and depression. Telemed J E Health 2022 Oct;28(10):1412-1420. [doi: 10.1089/tmj.2021.0590] [Medline: 35263185]
37. Karyotaki E, Ebert DD, Donkin L, et al. Do guided internet-based interventions result in clinically relevant changes for patients with depression? An individual participant data meta-analysis. Clin Psychol Rev 2018 Jul;63:80-92. [doi: 10.1016/j.cpr.2018.06.007] [Medline: 29940401]
38. Lindegaard T, Berg M, Andersson G. Efficacy of internet-delivered psychodynamic therapy: systematic review and meta-analysis. Psychodyn Psychiatry 2020;48(4):437-454. [doi: 10.1521/pdps.2020.48.4.437] [Medline: 33779225]
39. Lindqvist K, Mechler J, Carlbring P, et al. Affect-focused psychodynamic internet-based therapy for adolescent depression: randomized controlled trial. J Med Internet Res 2020 Mar 30;22(3):e18047. [doi: 10.2196/18047] [Medline: 32224489]
40. Mechler J, Lindqvist K, Carlbring P, et al. Therapist-guided internet-based psychodynamic therapy versus cognitive behavioural therapy for adolescent depression in Sweden: a randomised, clinical, non-inferiority trial. Lancet Digit Health 2022 Aug;4(8):e594-e603. [doi: 10.1016/S2589-7500(22)00095-4] [Medline: 35803894]
41. Midgley N, Guerrero-Tates B, Mortimer R, et al. The Depression: Online Therapy Study (D:OTS)—A pilot study of an internet-based psychodynamic treatment for adolescents with low mood in the UK, in the context of the COVID-19 pandemic. Int J Environ Res Public Health 2021 Dec 9;18(24):12993. [doi: 10.3390/ijerph182412993] [Medline: 34948601]
42. Schuster R, Topooco N, Keller A, Radvogin E, Laireiter AR. Advantages and disadvantages of online and blended therapy: replication and extension of findings on psychotherapists' appraisals. Internet Interv 2020 Sep;21:100326. [doi: 10.1016/j.invent.2020.100326] [Medline: 32477885]
43. van der Vaart R, Witting M, Riper H, Kooistra L, Bohlmeijer ET, van Gemert-Pijnen L. Blending online therapy into regular face-to-face therapy for depression: content, ratio and preconditions according to patients and therapists using a Delphi study. BMC Psychiatry 2014 Dec 14;14(355):355. [doi: 10.1186/s12888-014-0355-z] [Medline: 25496393]
44. Van Assche E, Bonroy B, Mertens M, et al. E-mental health implementation in inpatient care: exploring its potential and future challenges. Front Digit Health 2022;4:1027864. [doi: 10.3389/fdgth.2022.1027864] [Medline: 36588747]
45. Steinert C, Hofmann M, Kruse J, Leichsenring F. Relapse rates after psychotherapy for depression – stable long-term effects? A meta-analysis. J Affect Disord 2014 Oct;168(107-18):107-118. [doi: 10.1016/j.jad.2014.06.043]
46. Steinert C, Kruse J, Leichsenring F. Long-term outcome and non-response in psychotherapy: are we short-sighted? Psychother Psychosom 2016;85(4):235-237. [doi: 10.1159/000442262] [Medline: 27230866]
47. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
48. Karyotaki E, Riper H, Twisk J, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry 2017 Apr 1;74(4):351-359. [doi: 10.1001/jamapsychiatry.2017.0044] [Medline: 28241179]

XSL•FO
RenderX

49. Mechler J, Lindqvist K, Philips B, Midgley N, Lilliengren P. Internet-delivered affect-focused psychodynamic therapy for adolescent depression: treatment principles and clinical application in the ERiCA project. J Infant Child Adolesc Psychother 2024 Apr 2;23(2):123-141. [doi: 10.1080/15289168.2024.2339523]
50. Driessen E, Hegelmaier LM, Abbass AA, et al. The efficacy of short-term psychodynamic psychotherapy for depression: a meta-analysis update. Clin Psychol Rev 2015 Dec;42(1-15):1-15. [doi: 10.1016/j.cpr.2015.07.004] [Medline: 26281018]
51. Leichsenring F, Abbass A, Heim N, et al. The status of psychodynamic psychotherapy as an empirically supported treatment for common mental disorders - an umbrella review based on updated criteria. World Psychiatry 2023 Jun;22(2):286-304. [doi: 10.1002/wps.21104] [Medline: 37159376]
52. Barber JP. Research on dynamic therapies. In: Bergin and Garfield's Handbook of Psychotherapy and Behavior Change, 7th edition: John Wiley & Sons, Inc; 2021:387-419.
53. Cuijpers P, Karyotaki E, de Wit L, Ebert DD. The effects of fifteen evidence-supported therapies for adult depression: a meta-analytic review. Psychother Res 2020 Mar;30(3):279-293. [doi: 10.1080/10503307.2019.1649732] [Medline: 31394976]
54. Abbass A, Luyten P, Leichsenring F. Short-term psychodynamic psychotherapy is the first-line option for depression and treatment resistant depression according to available evidence. Can J Psychiatry 2024 Dec;69(12):888-889. [doi: 10.1177/07067437241299487] [Medline: 39558627]
55. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
56. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
57. Luyten P, Fonagy P. Psychodynamic treatments of depression. In: Pettit J, Olino TM, Boyd RC, Chu BC, Hayden PE, Pizzagalli DA, editors. APA Handbook of Depression: American Psychological Association Publishing; 2025, Vol. 2: Minoritized Populations, Lifespan Development, Assessment, and Treatment:579-602.
58. Kemmeren LL, van Schaik DJF, Riper H, Kleiboer AM, Bosmans JE, Smit JH. Effectiveness of blended depression treatment for adults in specialised mental healthcare: study protocol for a randomised controlled trial. BMC Psychiatry 2016 Apr 21;16(1):113. [doi: 10.1186/s12888-016-0818-5] [Medline: 27102812]
59. Fonagy P, Luyten P, Allison E. Epistemic petrification and the restoration of epistemic trust: a new conceptualization of borderline personality disorder and its psychosocial treatment. J Pers Disord 2015 Oct;29(5):575-609. [doi: 10.1521/pedi.2015.29.5.575] [Medline: 26393477]
60. Mitchell LM, Joshi U, Patel V, Lu C, Naslund JA. Economic evaluations of internet-based psychological interventions for anxiety disorders and depression: a systematic review. J Affect Disord 2021 Apr 1;284:157-182. [doi: 10.1016/j.jad.2021.01.092] [Medline: 33601245]
61. Leibovich L, Mechler J, Lindqvist K, Mortimer R, Edbrooke-Childs J, Midgley N. Unpacking the active ingredients of internet-based psychodynamic therapy for adolescents. Psychother Res 2023 Jan 2;33(1):108-117. [doi: 10.1080/10503307.2022.2050829]
62. MacKean M, Lecchi T, Mortimer R, Midgley N. 'I've started my journey to coping better': exploring adolescents' journeys through an internet-based psychodynamic therapy (I-PDT) for depression. Journal of Child Psychotherapy 2023 Sep 2;49(3):432-453. [doi: 10.1080/0075417X.2023.2173271]
63. Housby H, Thackeray L, Midgley N. What contributes to good outcomes? The perspective of young people on short-term psychoanalytic psychotherapy for depressed adolescents. PLoS ONE 2021;16(9):e0257334. [doi: 10.1371/journal.pone.0257334] [Medline: 34559829]
64. Wilmots E, Midgley N, Thackeray L, et al. The therapeutic relationship in cognitive behaviour therapy with depressed adolescents: a qualitative study of good-outcome cases. Psychol Psychother 2020 Jun;93(2):276-291. [doi: 10.1111/papt.12232] [Medline: 31119849]
65. Patel S, Akhtar A, Malins S, et al. The acceptability and usability of digital health interventions for adults with depression, anxiety, and somatoform disorders: qualitative systematic review and meta-synthesis. J Med Internet Res 2020 Jul 6;22(7):e16228. [doi: 10.2196/16228] [Medline: 32628116]
66. Ekeblad A, Dahlman M, Pio M, Malmquist A, Falkenström F. "What shall we focus on?" - A thematic analysis of what characterizes cognitive-behavior therapy sessions with high or low quality of working alliance. Psychother Res 2022 Nov;32(8):1003-1015. [doi: 10.1080/10503307.2022.2074806] [Medline: 35580260]
67. von Below C. "We Just Did Not Get on". young adults' experiences of unsuccessful psychodynamic psychotherapy - a lack of meta-communication and mentalization? Front Psychol 2020;11:1243. [doi: 10.3389/fpsyg.2020.01243] [Medline: 32625146]
68. O'Keeffe S, Martin P, Midgley N. When adolescents stop psychological therapy: rupture-repair in the therapeutic alliance and association with therapy ending. Psychotherapy (Chic) 2020 Dec;57(4):471-490. [doi: 10.1037/pst0000279] [Medline: 31928023]
69. Li ET, Midgley N, Luyten P, Sprecher EA, Campbell C. Mapping the journey from epistemic mistrust in depressed adolescents receiving psychotherapy. J Couns Psychol 2022 Oct;69(5):678-690. [doi: 10.1037/cou0000625] [Medline: 35737539]

XSL•FO
RenderX

70. Løvgren A, Røssberg JI, Engebretsen E, Ulberg R. Improvement in psychodynamic psychotherapy for depression: a qualitative study of the patients' perspective. Int J Environ Res Public Health 2020 Sep 19;17(18):18. [doi: 10.3390/ijerph17186843] [Medline: 32961671]
71. Løvgren A, Røssberg JI, Engebretsen E, Ulberg R. Multiple mutuality. Positions and storylines in adolescent psychodynamic psychotherapy. J Infant Child Adolesc Psychother 2020 Jul 2;19(3):239-254. [doi: 10.1080/15289168.2020.1771530]
72. Mortimer R, Somerville MP, Mechler J, et al. Connecting over the internet: establishing the therapeutic alliance in an internet-based treatment for depressed adolescents. Clin Child Psychol Psychiatry 2022 Jul;27(3):549-568. [doi: 10.1177/13591045221081193] [Medline: 35333646]
73. Lindqvist K, Mechler J, Midgley N, et al. "I didn't have to look her in the eyes"-participants' experiences of the therapeutic relationship in internet-based psychodynamic therapy for adolescent depression. Psychother Res 2024 Jun;34(5):648-662. [doi: 10.1080/10503307.2022.2150583] [Medline: 36473231]
74. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory-II: Psychological Corporation; 1996.
75. Cuijpers P, Turner EH, Koole SL, van Dijke A, Smit F. What is the threshold for a clinically relevant effect? The case of major depressive disorders. Depress Anxiety 2014 May;31(5):374-378. [doi: 10.1002/da.22249] [Medline: 24677535]
76. First MB, Williams JB, Karg RS, S RL. Structured Clinical Interview for DSM-5 Disorders, Clinical Trials Version (SCID-5-CT): American Psychiatric Association; 2015.
77. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
78. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x]
79. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000 Jun;31(2):73-86. [doi: 10.1016/s0005-7916(00)00012-4] [Medline: 11132119]
80. Attkisson C, Greenfield T. The UCSF client satisfaction scales: I. The Client Satisfaction Questionnaire-8. In: Maruish ME, editor. The Use of Psychological Testing for Treatment Planning and Outcomes Assessment: Instruments for Adults, 3rd edition: Lawrence Erlbaum Associates; 2004:799-811.
81. Bouwmans C, De Jong K, Timman R, et al. Feasibility, reliability and validity of a questionnaire on healthcare consumption and productivity loss in patients with a psychiatric disorder (TiC-P). BMC Health Serv Res 2013 Jun 15;13:217. [doi: 10.1186/1472-6963-13-217] [Medline: 23768141]
82. Luyten P, Blatt SJ. Psychodynamic treatment of depression. Psychiatr Clin North Am 2012 Mar;35(1):111-129. [doi: 10.1016/j.psc.2012.01.001] [Medline: 22370494]
83. van der Lem R, de Wever WW, van der Wee NJ, van Veen T, Cuijpers P, Zitman FG. The generalizability of psychotherapy efficacy trials in major depressive disorder: an analysis of the influence of patient selection in efficacy trials on symptom outcome in daily practice. BMC Psychiatry 2012 Nov 8;12(1):192. [doi: 10.1186/1471-244X-12-192] [Medline: 23137143]
84. Fichman L, Koestner R, Zuroff DC. Depressive styles in adolescence: assessment, relation to social functioning, and developmental trends. J Youth Adolescence 1994 Jun;23(3):315-330. [doi: 10.1007/BF01536722]
85. French C, Pinnock H, Forbes G, Skene I, Taylor SJC. Process evaluation within pragmatic randomised controlled trials: what is it, why is it done, and can we find it?-a systematic review. Trials 2020 Nov 9;21(1):916. [doi: 10.1186/s13063-020-04762-9] [Medline: 33168067]
86. Lambert M. The efficacy and effectiveness of psychotherapy. In: Lambert M, editor. Bergin and Garfield's Handbook of Psychotherapy and Behavior Change, 6th edition: John Wiley & Sons; 2013:169-218.
87. Midgley N, Ansaldo F, Target M. The meaningful assessment of therapy outcomes: incorporating a qualitative study into a randomized controlled trial evaluating the treatment of adolescent depression. Psychotherapy (Chic) 2014 Mar;51(1):128-137. [doi: 10.1037/a0034179] [Medline: 24377403]
88. Survey on the use of ICT in households and by individuals. Eurostat. URL: https://ec.europa.eu/eurostat/web/microdata/survey-ict-use-households-individuals [accessed 2025-12-22]
89. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Association; 2013. [doi: 10.1176/appi.books.9780890425596]
90. Lemma A, Target M, Fonagy P. Brief Dynamic Interpersonal Therapy: A Clinician's Guide: Oxford University Press; 2011.
91. Beck AT, Rush AJ, Shaw BF, Emery G, DeRubeis RJ, Hollon SD. Cognitive Therapy of Depression, 2nd edition: The Guilford Press; 2024.
92. Lemma A, Hepworth M, Fonagy P, Luyten P, Abrahams D. Internet-delivered dynamic interpersonal therapy (i-DIT) for depression and anxiety. In: Brief Dynamic Interpersonal Therapy: A Clinician's Guide, 2nd edition: Oxford University Press; 2024:327-346.
93. Lemma A, Hepworth M, Fonagy P, Luyten P, Abrahams D. Brief Dynamic Interpersonal Therapy: A Clinician's Guide, 2nd edition: Oxford University Press; 2024.
94. Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive Therapy of Depression: The Guilford Press; 1979.

XSL•FO
RenderX

95. Laurenssen EMP, Smits ML, Bales DL, et al. Day hospital mentalization-based treatment versus intensive outpatient mentalization-based treatment for patients with severe borderline personality disorder: protocol of a multicentre randomized clinical trial. BMC Psychiatry 2014 Nov 18;14(1):301. [doi: 10.1186/s12888-014-0301-0] [Medline: 25403144]
96. Laurenssen EM, Westra D, Kikkert MJ, et al. Day Hospital Mentalization-Based Treatment (MBT-DH) versus treatment as usual in the treatment of severe borderline personality disorder: protocol of a randomized controlled trial. BMC Psychiatry 2014 Dec;14(1):24886402. [doi: 10.1186/1471-244X-14-149]
97. Cuijpers P, van Straten A, van Oppen P, Andersson G. Are psychological and pharmacologic interventions equally effective in the treatment of adult depressive disorders? A meta-analysis of comparative studies. J Clin Psychiatry 2008 Nov;69(11):1675-1685. [doi: 10.4088/jcp.v69n1102] [Medline: 18945396]
98. Cuijpers P, Cristea IA, Karyotaki E, Reijnders M, Huibers MJH. How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. World Psychiatry 2016 Oct;15(3):245-258. [doi: 10.1002/wps.20346] [Medline: 27717254]
99. Cuijpers P, Berking M, Andersson G, Quigley L, Kleiboer A, Dobson KS. A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. Can J Psychiatry 2013 Jul;58(7):376-385. [doi: 10.1177/070674371305800702] [Medline: 23870719]
100. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: a meta-analysis. PLoS ONE 2010 Oct 13;5(10):e13196. [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
101. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. . 2014 Oct(3) p. 288-295. [doi: 10.1002/wps.20151]
102. Leichsenring F, Luyten P, Hilsenroth MJ, et al. Psychodynamic therapy meets evidence-based medicine: a systematic review using updated criteria. Lancet Psychiatry 2015 Jul;2(7):648-660. [doi: 10.1016/S2215-0366(15)00155-8]
103. Wellek S. Testing Statistical Hypotheses of Equivalence and Noninferiority, 2nd edition: CRC Press; 2010. [doi: 10.1201/EBK1439808184]
104. The R Project for Statistical Computing. URL: https://www.R-project.org [accessed 2025-12-22]
105. Balk EM, Earley A, Patel K, Trikalinos TA, Dahabreh IJ. AHRQ Methods for Effective Health Care Empirical Assessment of Within-Arm Correlation Imputation in Trials of Continuous Outcomes: Agency for Healthcare Research and Quality (US); 2012.
106. Johns RG, Barkham M, Kellett S, Saxon D. A systematic review of therapist effects: a critical narrative update and refinement to review. Clin Psychol Rev 2019 Feb;67(78-93):78-93. [doi: 10.1016/j.cpr.2018.08.004] [Medline: 30442478]
107. Vierron E, Giraudeau B. Sample size calculation for multicenter randomized trial: taking the center effect into account. Contemp Clin Trials 2007 Jul;28(4):451-458. [doi: 10.1016/j.cct.2006.11.003] [Medline: 17188941]
108. Sullivan TR, White IR, Salter AB, Ryan P, Lee KJ. Should multiple imputation be the method of choice for handling missing data in randomized trials? Stat Methods Med Res 2018 Sep;27(9):2610-2626. [doi: 10.1177/0962280216683570] [Medline: 28034175]
109. Barber JP, Liese BS, Abrams MJ. Development of the Cognitive Therapy Adherence and Competence Scale. Psychother Res 2003 Jun 1;13(2):205-221. [doi: 10.1093/ptr/kpg019] [Medline: 22475163]
110. Soeteman DI, Busschbach JJV, Verheul R, Hoomans T, Kim JJ. Cost-effective psychotherapy for personality disorders in the Netherlands: the value of further research and active implementation. Value Health 2011;14(2):229-239. [doi: 10.1016/j.jval.2010.10.020] [Medline: 21296601]
111. Geertz C. The Interpretation of Cultures: Basic Books; 1973.
112. Braun V, Clarke V. Thematic Analysis: A Practical Guide: Sage; 2021.
113. Parkinson S, Eatough V, Holmes J, Stapley E, Midgley N. Framework analysis: a worked example of a study exploring young people's experiences of depression. Qual Res Psychol 2016 Apr 2;13(2):109-129. [doi: 10.1080/14780887.2015.1119228]
114. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

Abbreviations

ANCOVA: analysis of covariance
BDI-II: Beck Depression Inventory-II
CBT: cognitive behavioral therapy
CEQ: Credibility and Expectancy Questionnaire
CGG: Centrum Geestelijke Gezondheidszorg
CONSORT: Consolidated Standards of Reporting Trials
CSQ-8: Client Satisfaction Questionnaire-8
CTACS: Cognitive Therapy Adherence and Competency Scale

XSL•FO
RenderX

DEQ: Depressive Experiences Questionnaire
DIT: dynamic interpersonal therapy
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
eCRF: electronic Case Report Form
EQ-5D-5L: EuroQoL 5-Dimension 5-Level
ES: effect size
FTF: face-to-face
ICF: informed consent form
IPAF: interpersonal affective focus
ITT: intention-to-treat
MAR: missing at random
MDD: major depressive disorder
PDT: psychodynamic therapy
PHQ-9: Patient Health Questionnaire-9
QALY: quality-adjusted life year
RA: research assistant
RCT: randomized controlled trial
RIZIV/INAMI: Rijksinstituut voor Ziekte- en Invaliditeitsverzekering/Institut National d'Assurance Maladie-Invalidité
SAE: serious adverse event
SAP: statistical analysis plan
SCID-5-CT: Structured Clinical Interview for DSM-5 disorders – Clinical Trials Version
SF-HLQ: Short Form-Health and Labor Questionnaire
TA: thematic analysis
TC: trial coordinator
TIC-P: Treatment Inventory of Costs in Patients with psychiatric disorders
TSC: Trial Steering Committee

Edited by A Schwartz; submitted 11.Jul.2025; peer-reviewed by F Leichsenring; revised version received 17.Nov.2025; accepted 29.Nov.2025; published 14.Jan.2026.

Please cite as:
Luyten P, Speybrouck D, Martin P, Claes S, Midgley N, van Eeren H, van Busschbach J, Tang E
Clinical and Cost-Effectiveness of Blended Cognitive Behavioral Therapy or Psychodynamic Therapy Versus Face-to-Face Psychotherapy for Depression (BLENDED Study): Protocol for a Pragmatic, Multicenter, Assessor-Blinded Randomized Controlled Noninferiority Trial
JMIR Res Protoc 2026;15:e80511
URL: https://www.researchprotocols.org/2026/1/e80511
doi:10.2196/80511

© Patrick Luyten, Danielle Speybrouck, Peter Martin, Stephan Claes, Nick Midgley, Hester van Eeren, Jan van Busschbach, Eileen Tang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 14.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessment of Blood Glucose Measurement Using New Noninvasive Technology: Protocol and Methodology

Eka Widrian Suradji[1,2,3*], MD, PhD; Satvinder Singh Dhaliwal[4,5,6,7], PhD; Zhang Li-Feng[8], PhD; Entong Zhou[8], BSc; Amos Lim[8], MSc; Hui Hern Wong[8], MDS; Yao Du[8], PhD; Fredicia Fredicia[9], M Com; Amanda Dianky[3*], MD; Puspasari Pasaribu[3], BSc; Ernest Gamaliel Masayoshi Takashi[2], BM; Justin Yong An Tan[8,10]; Mardi Santoso[11], DTM, MD, Dr Med; Seng Bin Ang[12,13], MBBS

[1]Department of Public Health, Faculty of Medicine and Health Sciences, Krida Wacana Christian University, Arjuna Utara no 6, Kec Kebon Jeruk, West Jakarta, DKI Jakarta, Indonesia

[10]Applied Artificial Intelligence, School of Science, Singapore Institute of Technology, Singapore, Singapore

[11]Department of Internal Medicine, Faculty of Medicine and Health Sciences, Krida Wacana Christian University, DKI Jakarta, Indonesia

[12]Family Medicine Academic Clinical Program, Duke-NUS Medical School, Singapore, Singapore

[13]Family Medicine, KK Women's and Children's Hospital, Singapore, Singapore

[2]Integrated Research Laboratory, Krida Wacana Christian University, DKI Jakarta, Indonesia

[3]LIF Indonesia, DKI Jakarta, Indonesia

[4]Curtin Medical Research Institute, Curtin University, Curtin, Australia

[5]Obstetrics and Gynaecology Academic Clinical Program, Duke-NUS Medical School, Singapore, Singapore

[6]Office of the Provost, Singapore University of Social Sciences, Singapore, Singapore

[7]Health and Social Sciences, Singapore Institute of Technology, Singapore, Singapore

[8]Actxa Pte Ltd, Singapore, Singapore

[9]Computer Science Department, Faculty of Engineering and Computer Science, Krida Wacana Christian University, DKI Jakarta, Indonesia

[*]these authors contributed equally

Corresponding Author:
Eka Widrian Suradji, MD, PhD
Department of Public Health, Faculty of Medicine and Health Sciences, Krida Wacana Christian University, Arjuna Utara no 6, Kec Kebon Jeruk, West Jakarta, DKI Jakarta, Indonesia

Abstract

Background: Diabetes mellitus (DM) is a major noncommunicable disease with a significant increase in prevalence, especially in low- and middle-income countries. The latest International Diabetes Federation Diabetes Atlas (2025) reports that 11.1% of the adult population (20 to 79 years old) is living with diabetes, with over 4 in 10 unaware of their condition. Early diagnosis and treatment of diabetes reduce the risk and slow the progression of debilitating complications, such as amputation, vision loss, renal failure, cardiovascular disease, dementia, some cancers, and infections like tuberculosis and severe COVID-19. Current screening methods for diabetes are invasive and costly. This has limited their utilization, especially in high-density populations and low- and middle-income countries such as Indonesia. Blood Glucose Evaluation and Monitoring (BGEM) is a machine learning algorithm developed by Actxa to analyze photoplethysmography data from wearable devices for diabetic risk assessment. Its noninvasive and user-friendly nature makes it a strong candidate for fulfilling the need for a diabetes screening or monitoring tool.

Objective: The aim of this study is to collect a large and more diverse dataset for the training of BGEM machine learning models. This dataset is intended to improve the model's generalizability and to evaluate its performance across different age groups, racial groups, and skin types, with the goal of enhancing accuracy and robustness for diabetes risk assessment and glucose monitoring.

Methods: Adult participants aged 18 years and above, with either a diabetic or a nondiabetic history, who reside in Greater Jakarta Area, Indonesia, were approached for recruitment. Blood glucose was assessed using laboratory blood analysis from capillary or plasma samples after fasting and at 1, 2, and 3 hours after a meal. BGEM data were also collected at each of these time points. Anthropological measurements with a standardized questionnaire on physical activity, demographic information, respondent's diabetic status, and current medications taken were also collected.

Results: Between June and October 2024, 885 participants were enrolled. Eight photoplethysmography recordings per participant were collected across 4 meal time points using 2 wearable devices in addition to the collection of clinical measurements, blood sampling, and related questionnaires.

XSL•FO
RenderX

Conclusions: This protocol paper outlines the methodology designed for assessing and interpreting participants' blood sugar profiles, especially on demographic variability, in order to evaluate BGEM, a photoplethysmography-based artificial intelligence model designed to estimate blood glucose levels and diabetic risk. The clinical trial was conducted on Indonesian participants with and without diabetes while considering various influencing factors. This dataset is designed to enable assessment of the model's performance across diverse racial, risk factors, and skin-type groups, with the aim of making the model more valid and reliable.

Trial Registration: ClinicalTrials.gov NCT06642467; https://clinicaltrials.gov/study/NCT06642467

International Registered Report Identifier (IRRID): DERR1-10.2196/76558

(JMIR Res Protoc 2026;15:e76558) doi:10.2196/76558

KEYWORDS

diabetes mellitus; noninvasive; blood glucose; screening; machine learning; photoplethysmography; artificial intelligence; wearables; BGEM; Indonesia; Blood Glucose Evaluation and Monitoring

Introduction

Diabetes mellitus (DM) is a major noncommunicable disease that has affected around 500 million people worldwide and continues to rise up to 700 million people globally in 2045 [1]. Other noncommunicable diseases' death risks have decreased by 20% since the turn of the century; DM, in contrast, has increased by 3% worldwide. The latest International Diabetes Federation Diabetes Atlas (2025) reports that 11.1% of the adult population (20 - 79 y) is living with diabetes, with over 4 in 10 unaware of their condition.

Several reports have projected that about 1 in 2 adult patients with diabetes are unaware of their diagnosis [1-3]. Undiagnosed DM has resulted in an increase of cases with complications and elevated the burden of the disease, as younger undiagnosed DM cases have more rapid progression of complications and lower response to general DM treatment. This has been shown to be higher in low- and middle-income countries, where 80% of the DM population lived, thus reflecting the enormous burden of the disease [4,5]. Meanwhile, early diagnosis and treatment of DM reduce the risk and slow the progression of debilitating complications such as amputation, vision loss, renal failure, cardiovascular disease, dementia, some cancers, and infections including tuberculosis and severe COVID-19 [6]. All of the above calls for early screening and frequent monitoring strategies that are both effective and cost-efficient.

Currently available methods to detect DM are hemoglobin A_{1c} (HbA_{1c}) level, fasting plasma glucose, 2-hour glucose tolerance test, and random plasma glucose test. All these methods are invasive; they require trained professionals and laboratory equipment to be performed correctly. The American Diabetes Association has recommended DM screening starting at 35 years and every 3 years after, especially for people with obesity, overweight, and one risk factor [7]. In low- and middle-income countries, this is even harder to implement due to the cost and scarce resources of such screening methods [5,6,8].

Indonesia, a middle-income country with a population of 283 million, faces a big challenge in DM management. It was estimated that DM prevalence was 9.49% (2024) and will increase to 16.09% (2045) without any major effective intervention [9]. As in any other low- and middle-income countries, the cost and availability of DM-related testing have prohibited Indonesia from enacting an effective prevention or screening program; therefore, an affordable, effective, and simple screening or monitoring method for DM will be a boon in the national health program.

Photoplethysmography is a noninvasive optical technology that measures light absorbance to assess blood volume changes. It has been widely integrated into wearable devices, such as smartwatches, to monitor heart rate (HR) and other physiological parameters [10]. Emerging research suggests that photoplethysmography data may also be valuable for detecting DM and estimating glucose levels [11-13]. The noninvasive nature, affordability, and convenience of photoplethysmography make it a promising tool for early detection and preventive screening of DM, especially in large and diverse populations, such as in Indonesia. Despite photoplethysmography technology's potential use in estimating blood glucose and DM prediction, several issues have been reported to influence its performance, such as medication status of subjects [14], motion artifacts [15], and also skin types [16]. This protocol paper will strive to address these issues to improve the reliability (accuracy and robustness) of photoplethysmography-based technology in wearable devices for blood glucose monitoring and DM risk prediction through advanced artificial intelligence (AI) and data analytics.

Blood Glucose Evaluation and Monitoring (BGEM), a machine learning (ML) algorithm developed by Actxa to analyze photoplethysmography data for assessing diabetic risk. It is one of the first ML models deployed in commercial wearable devices. BGEM was trained using data collected from a registered clinical trial (NCT05504096) conducted at KK Women's and Children's Hospital in Singapore, involving 500 Singaporean participants. The reported accuracy of BGEM is 84.7% [12]. The model's performance is sensitive to demographic features, with the inclusion of demographic data significantly enhancing its accuracy [13]. BGEM would benefit from a larger and more diverse training dataset, as the current training data consists primarily of Singaporean participants, predominantly healthy females.

In addition to glycemia prediction, novel methods for diabetic risk assessment and management based on multiple glucose measurements may be developed using BGEM due to its lower cost and noninvasiveness. Glycemic variability, which refers

XSL•FO
RenderX

to fluctuations of blood glucose concentration that happen in a day or similar timeframes on different days, has been shown to be associated with microvascular and macrovascular damage in patients with DM [17]. Glycemic variability has been proposed in multiple studies for DM management; however, difficulties in the measurement and consensus on its optimal value, even with recent more widespread use of continuous glucose monitoring, have limited its use in the general population [17,18], despite the general acceptance that it is a promising indicator for DM control.

Here, we present the protocol for a registered clinical trial (NCT06642467) designed to collect a larger and more robust training dataset to enhance BGEM's performance. This study will collect data to improve current algorithm performance, with an attempt to generalize the diagnostic findings from the previous clinical study to specific groups of participants, with the goal of enhancing accuracy and robustness for glucose monitoring and diabetes risk assessment. The study aims to comprehensively collect photoplethysmography data using 2 different types of wearable devices at multiple time points before and after a breakfast meal. Both capillary and plasma blood samples were collected and subjected to thorough laboratory analysis. Participants from Indonesia were recruited, ensuring a balanced representation in terms of diabetic status (diabetic or nondiabetic) and gender. Additionally, a comprehensive questionnaire was used to gather essential information, capitalizing on the diversity of the training data, including factors such as hypertension, skin type, and other relevant variables, which have been reported to possibly influence either photoplethysmography sensor readings or ML model prediction.

Methods

Participant Recruitment

A total number of 885 participants from adults aged 18 years and above were recruited. These participants reside in the Greater Jakarta Area, Indonesia, and were from the diabetic exercise group, from staff and family members of Krida Wacana Christian University (Ukrida), and from its affiliated organizations. Participants were recruited using an online registration provided by the research team. Exclusion criteria were implemented as follows: BMI >37 kg/m^2, diastolic blood pressure ≥120 mmHg, and systolic blood pressure ≥180 mmHg, wearing a pacemaker, having fever or any other acute critical illness during the data collection, and pregnant female participants. All subjects received an explanation on the study purpose, protocol, and risks, and also signed the informed consent prior to admission to the research study.

Study Design

This study utilizes a cross-sectional design with a short-term follow-up post meal for blood glucose monitoring, from July 30 to October 5, 2024, at the Ukrida Hospital, Jakarta, Indonesia. There was an a priori aim to assess the measurement error of the 4 groups, males and females with diabetes and without diabetes, separately. A sample size of 200 for each group was determined using the methods described by Bland and Altman [19]. Hence, a total sample size of 800 was required to provide a 95% CI of ±0.24×SD. Allowing for about 10% dropout, a total sample size of about 890 subjects was planned for recruitment.

The study will explore BGEM blood glucose prediction capabilities compared to laboratory blood analysis from capillary or plasma blood samples, with the study workflow summarized in Figure 1. The subjects were given an explanation on the study, asked for informed consent, and given a numeric identity tag to ensure anonymity of the dataset after undergoing a minimum of 8 hours of fasting before admittance to the study. Subjects undertook weight, height, waist circumference, and blood pressure measurement using GEA ZT120 body scales, tape measure roll, and Beurer BM26 sphygmomanometer, respectively. BMI scores were calculated from the weight and height of the subjects to evaluate obesity or metabolic condition of the subjects.

XSL•FO
RenderX

Figure 1. A conceptual framework of study design from recruitment, data collection, to discharge for 473 participants who self-acknowledged having diabetes mellitus (DM) and 412 participants who did not (non-DM).

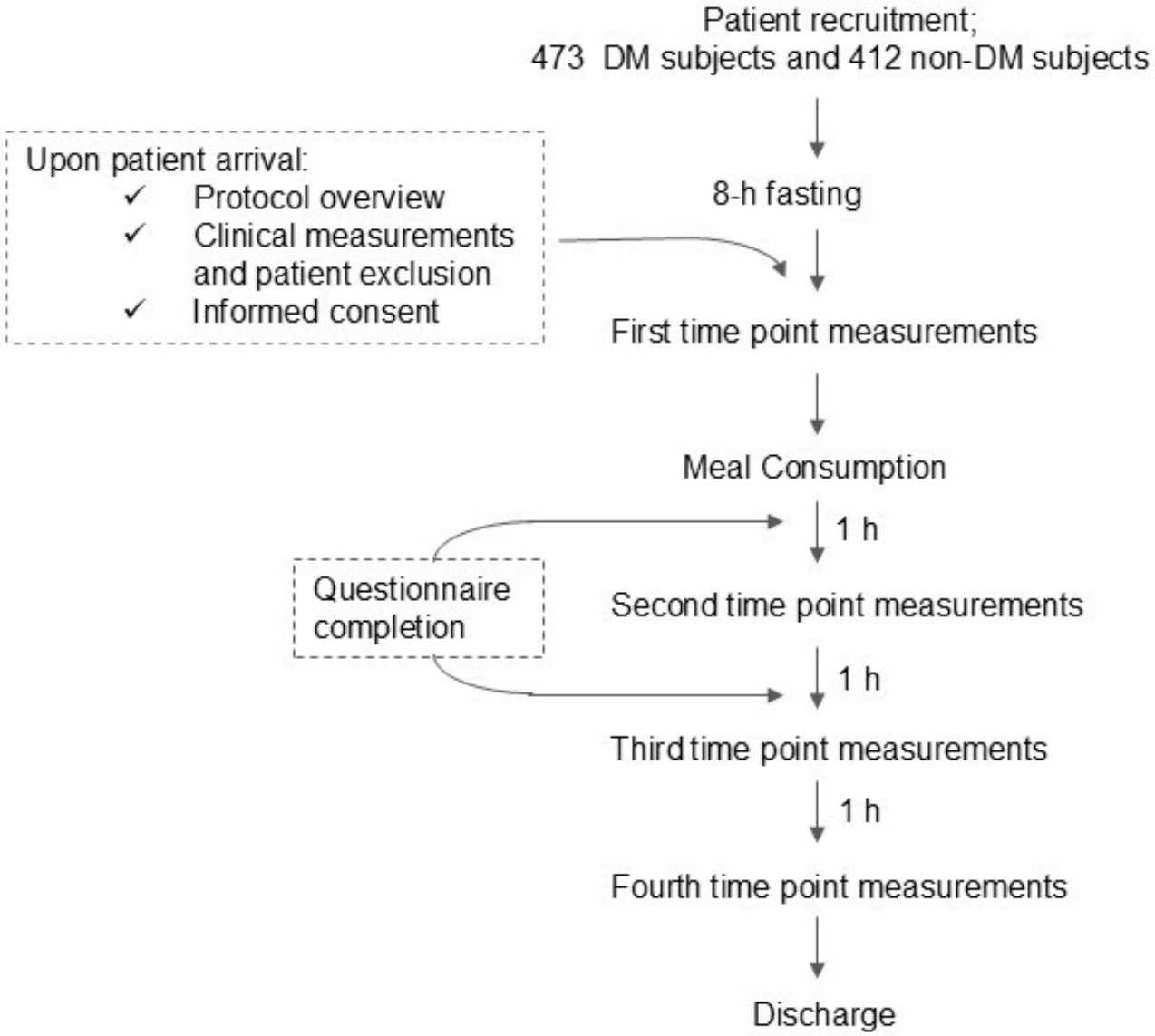

Subsequently, each participant's photoplethysmography data at the fasting time point (the first time point) were collected using both a smart ring (JC Smart Health Ring 2301B) and a smartwatch (Actxa Spark+Series 2), both provided by Actxa Pte Ltd. Simultaneously, blood samples were obtained from 2 sources: capillary blood for glucose concentration measurement using Accu-Chek (Roche) and plasma blood from the cubital vein for laboratory analysis, as described in a later section.

Participants then consumed a 552-kcal meal set with 79.8 g of carbohydrate content prepared by Ukrida Hospital. This meal was chosen to ensure adequate sustenance and satiety after overnight fasting, as sugary drinks alone may not sufficiently fulfill. The nutritional composition of the meal is detailed in Textbox 1 and Table S1 in Multimedia Appendix 1.

XSL•FO
RenderX

Textbox 1. Nutritional value of the set meal for the subjects, given after subjects fast for a minimum of 8 h and undergo first blood sample collection and photoplethysmography reading from wearables.

Energy (kcal): 552.1

Water (g): 241.8

Protein (g): 19

Protein percentage: 13.8

Fat (g): 19.3

Fat percentage: 31.5

Carbohydrate (g): 79.8

Carbohydrate percentage: 57.8

Dietary fiber (g): 5.2

Polyunsaturated fatty acid (g): 1.8

Cholesterol (mg): 33

Vitamin A (µg): 1396.8

Carotene (mg): 0.4

Vitamin E (eq.) (mg): 0.8

Vitamin B1 (mg): 0

Vitamin B2 (mg): 0.1

Vitamin B6 (mg): 0.6

Total folic acid (µg): 31.8

Vitamin C (mg): 10.7

Sodium (mg): 1015.5

Potassium (mg): 549.4

Calcium (mg): 59.2

Magnesium (mg): 66.2

Phosphorus (mg): 231.5

Iron (mg): 2.5

Zinc (mg): 1.6

Participants then underwent measurements at the second time point (1 h postprandial), third time point (2 h postprandial), and fourth time point (3 h postprandial). Photoplethysmography data collection followed the same procedure as the first time point using both the smartphone and the smart ring. For blood sample collection and testing, the third time point followed the same procedure as the first time point, with both capillary and plasma blood samples collected and analyzed. At the second and fourth time points, plasma blood collection was omitted, and only capillary blood samples were obtained and tested. Further details are provided in Table 1.

XSL•FO
RenderX

Table . Blood glucose metabolism-related data collection from photoplethysmography sensors from 2 wearables, a glucometer, and blood samples based on several time points before and after meal.

Measurements	Smart ring[a]	Smartwatch[b]	Glucose meter[c]	Plasma blood[d]
First time point (fasting state)	Yes	Yes	Yes	Yes
Second time point (1 h postprandial)	Yes	Yes	Yes	No
Third time point (2 h postprandial)	Yes	Yes	Yes	Yes
Fourth time point (3 h postprandial)	Yes	Yes	Yes	No

[a]JC smart health ring 2301B.

[b]Actxa Spark + Series 2.

[c]Roche, ACCU-CHEK guide.

[d]Full blood count: Sysmex XN-550. Plasma glucose: Roche, Cobas c111 Analyzer. Hemoglobin A_{1c} level: Finecare FIA Meter III Plus.

During the 1-hour intervals at the last 3 time points, while waiting between data and sample collection, participants were asked to complete the International Physical Activity Questionnaire long form, along with questionnaires related to DM and other factors that may affect photoplethysmography reading and BGEM prediction that follows, such as skin type, treatment status, dominant hand, and race (Table 2).

Table . Baseline study questionnaire and test descriptions collected from the study's participants.

Questionnaire and test	Description of content	References
Demographic	Five items: age, gender, ethnicity, handedness, an Fitzpatrick skin type	[13]
Diabetic status and history	Six items: family history, status, year diagnosed, medication, gestational diabetes, and drink consumption	[13,20-23]
Relevant medical history	Four items: smoking, hypercholesterolemia, hypertension, and related medication	[13,21-23]
Physical activity	International Physical Activity Questionnaire (IPAQ) long form; 27 items	[24]
Clinical measurements	Four items: weight, height, waist circumference, and blood pressure	[13,20,21,23]
Pre-photoplethysmography measurements questions	Seven items: time of medication, time of meal, meal content, physical activity in 24 h, last, duration, and intensity	[12,25]

Ethical Considerations

All protocols were performed under hospital's medical staff supervision at all times. Ethics approval for the study has been obtained from the Ukrida Faculty of Medicine & Health Sciences ethical committee on July 15, 2024 (1779/SLKE/IM/UKKW /FKIK/KEPK /VII/2024).

Blood Sample Analysis

All capillary blood samples were analyzed using the Accu-Chek Guide glucometer (Roche, USA). Cubital plasma blood samples were analyzed for full blood count using the Sysmex XN-550 (Sysmex, Japan), plasma blood glucose using the Cobas C111 Analyzer (Roche, Switzerland), and HbA_{1c} levels using the FineCare FIA Meter III Plus (Gongu Wondfo, China). These instruments were available at the Ukrida Hospital clinical laboratory. All hospital equipment was routinely calibrated according to the provider's manual.

Photoplethysmography Data Collection

In this study, 2 wearable devices were utilized: the Actxa Spark+Series 2 smartwatch and the JC Smart Health Ring 2301B. Both devices are equipped with a green light photoplethysmography sensor that captures photoplethysmography signals at 50 Hz, ensuring adequate signal quality for HR analysis.

During data collection, the smartwatch was worn on the participant's nondominant wrist, positioned at least 2 finger widths away from the wrist bone. The smart ring, selected based on appropriate sizing, was worn at the base of the index finger on the same hand. To ensure optimal signal acquisition, the ring's photoplethysmography sensor was positioned facing the inner side of the finger. Each photoplethysmography recording session lasted for 5 minutes. Participants were instructed to remain seated, maintain a relaxed posture, and minimize

XSL•FO
RenderX

movement, with their nondominant hand resting on a table to reduce motion artifacts.

Photoplethysmography data were collected by the wearable devices and transmitted via Bluetooth to a smartphone. The smartphone then relayed the data via the internet to Actxa's company server for storage and further analysis.

Non-Photoplethysmography Data Preprocessing

All 7 non-photoplethysmography features required for downstream model training—age, gender, body weight, body height, meal timing, diabetes status, and family history of diabetes—were complete and free of errors. Automated blood test results and finger prick glucometer measurements also contained no missing or invalid values.

Medication status, collected via participant self-report, was manually curated by medical personnel. Other questionnaire data will be further cleaned or curated as needed for downstream analyses. Missing values, if encountered, can be addressed by tracing the source information or applying statistical methods, such as decision tree–based imputation.

Photoplethysmography Data Preprocessing

Each photoplethysmography recording was required to be at least 4 minutes in duration. No recordings from this clinical trial were excluded based on this criterion.

One participant was absent for the fourth measurement (3 h after the meal). For this participant, the available time point data were retained for individual data point analyses. However, in the time-series analysis (see the section photoplethysmography Data Feature Extraction), the data were excluded due to the incomplete sequence.

Raw photoplethysmography signals were obtained from the smartwatch (16-bit binary) and smart ring (23-bit binary) devices. Raw photoplethysmography data were preprocessed using a previously developed pipeline incorporating frequency noise filtering and outlier correction [12].

During the signal preprocessing stage, the raw digital signals were first converted to analog values using the following formula:

$Vi=5 \times Signal216$ (for smartwatch data)

$Vi=5 \times Signal223$ (for smart ring data)

A Chebyshev Type II bandpass filter (0.3 - 5 Hz) was applied to remove low- and high-frequency noise, including motion artifacts. Outliers were identified, when necessary, using a z score threshold of 3 SD from the mean and replaced with reasonable estimates via nearest-neighbor interpolation prior to HR variability feature extraction.

Photoplethysmography Data Feature Extraction

To accurately identify key temporal positions of the photoplethysmography waveform, potential systolic peaks were first detected using sign changes in the first-order difference of the photoplethysmography signal, specifically identifying transitions from positive to negative values. The leading and trailing troughs of the pulse wave were then detected by inverting the photoplethysmography signal and applying the same first-order difference and sign change method.

HR was then calculated using the key temporal positions. The measured HR was required to exceed 48 beats per minute. No recordings from this clinical trial were excluded based on this criterion.

Using the identified key temporal positions, a feature extraction module was then applied to the photoplethysmography data fragments to compute a comprehensive set of features. The panel of features was established in our previously reported model, which included 248 features related to HR, HR variability, waveform morphology, energy metrics, complexity measures, and continuous wavelet transform [12].

Prior to model training, feature selection techniques were applied to eliminate redundant features. This process was conducted exclusively on the feature set derived from photoplethysmography data and involved 2 stages. First, a self-correlation matrix was calculated, and features with a Pearson correlation coefficient >0.7 were removed to reduce multicollinearity. Second, the correlation between each remaining feature and the regression target—blood glucose levels measured via finger prick tests under different mealtime conditions—was assessed. A subset of features with the strongest correlations to the target variable was retained, with the final number determined empirically to optimize model performance.

To account for the varying feature magnitudes, Min-Max scaling was applied. Since the regression target (blood glucose values) exhibited a nonnormal distribution, a Box-Cox transformation was used for normalization. Following model prediction, an inverse Box-Cox transformation was performed to restore the predicted blood glucose levels to their original scale.

Model Training

During model training, 10-fold cross-validation was applied to obtain a robust estimate of generalization performance and to monitor potential overfitting. To prevent data leakage, particularly biometric identity-related features potentially embedded in the photoplethysmography data [26], we ensured that fragments from the same participant were not included in both the training and validation sets. Data collected in this clinical trial were used exclusively for model training and validation.

Additionally, given that this study employs short-term, mealtime prospective measurements of blood glucose profiles, we aim to leverage each participant's individual blood glucose fluctuation pattern to explore prediction of diabetes risk or glucose levels and probable glycemia monitoring post meal. To achieve this, the photoplethysmography data from the 4 mealtime points for each participant will be combined into a time series, which will serve as input to a model trained to predict glucose levels or diabetes risk across all 4 points of the meal cycle. This approach allows the longitudinal aspect of the data—the participant-specific fluctuation pattern—to be fully utilized, potentially improving model performance and predictive accuracy.

XSL•FO
RenderX

Statistical Analysis

Demographic data for continuous variables will be presented as mean (SD), and categorical variables will be presented as count (percentage). Univariate analyses will be performed using *t* tests for independent groups for continuous variables and chi-square tests for categorical variables. Discriminations are assessed by plotting the receiver operating characteristic curve and calculating the area under the receiver operating characteristic curve (AUC) or *C* statistic. The accuracy of the prediction model will be analyzed using both Type I and Type II Parkes Error Grids. The sensitivity, specificity, and F_1-score of the BGEM model will also be evaluated and presented to give a clear description of its strength and weakness.

The Parkes Error Grid is used to assess how clinically risky an inaccurate blood glucose reading might be. Each device reading is compared with a reference value and placed into 1 of 5 zones. Zone A means the reading is clinically accurate and would not change treatment. Zone B reflects small errors that might adjust a decision but have little or no impact on outcome. Zone C includes errors likely to affect clinical results. Zone D involves inaccuracies that could lead to clearly risky treatment decisions. Zone E represents dangerous errors that could cause severe or life-threatening consequences if acted upon [27].

The ML model will be trained using Lasso regression, cross-validation, and other ML techniques to find the optimal regularization parameter to ensure that the model generalizes to unseen data. The dataset is initially divided into a training set and a validation set. The training set is used for model development and hyperparameter tuning, while the validation set is reserved for a final, unbiased evaluation of the model's performance. Analysis will ensure that the model does not overfit to the training data and is able to predict outcomes on new data with confidence. Once the optimal regularization parameter is identified through cross-validation, a final regression model is trained on the entire training set using this optimal regularization parameter value.

The blood sugar profiles of subjects with varying demographic characteristics will be analyzed and compared statistically to quantify the similarities and differences between individuals. Profile analysis and correlation and regression techniques will be used for the comparison of the blood sugar profiles, taking into consideration the correlation between the repeated measurements within each individual. Trending analysis will also be performed using mixed effects models and clustering methods, so as to incorporate the correlated repeated measurements at the 4 time points for each participant.

The data will be analyzed using statistical packages, including Stata (version 19) and ML packages R and Python. *P* values <.05 will be considered as statistically significant.

Results

The IRB approval was given on July 15, 2024, after which the data collection was started on July 30, 2024, and finished on October 5, 2024; the analysis was projected to be concluded in the third quarter of 2025 and the results submitted for publication at the end of 2025.

A total of 885 participants (473 with diabetes and 412 without diabetes) were enrolled. Eight photoplethysmography recordings per participant were collected across 4 mealtime points using 2 wearable devices. Participants completed the International Physical Activity Questionnaire (IPAQ) long form and additional questionnaires on diabetes and related risk factors. Capillary and plasma blood tests were performed, and results were recorded.

Discussion

Overview

This study will provide more data to further empower the BGEM AI model that has been developed since 2022 [12].

Photoplethysmography feature data from participants, collected across 4 time points, are intended to support the improvement of model performance across diverse racial groups, risk factors, and skin types in terms of its accuracy and robustness. The study design also addresses issues that have been reported to influence photoplethysmography sensors, such as motion artifacts.

The inclusion of additional clinical data such as anthropometric measurement, HbA_{1c}, hematocrit, and other blood parameters performed in this study will give a more expansive view on factors that may influence photoplethysmography blood glucose predictive ability [14-16,28]. Analysis of these factors will provide insight into whether any of them influence BGEM accuracy or can be used as additional data to improve its performance.

This study is the first to report performance of photoplethysmography-based AI model used in the Indonesian population. Population-wide DM screening in Indonesia posed a problem to any health intervention program due to the challenging geography and population spread across multiple islands and diverse ethnicity of Indonesian people [29,30]. Therefore, this study can become a foundation of BGEM use in DM management, prevention, and screening in Indonesia.

DM, with its huge disease burden and increasing number of afflicted populations, needs new, innovative, and effective means to control and minimize its impact [31,32]. This protocol paper hopes to provide a base method in developing a noninvasive photoplethysmography-based AI model to predict blood glucose level and evaluating DM status of the subject.

Comparison With Previous Work

BGEM use in the relatively large Indonesian population is unique and will become an important stepping stone for developing a noninvasive model that is effective in blood glucose screening and monitoring and DM management. The previous study of BGEM was performed in a female-dominant and mainly non-DM population [13]; thus, this study with a different gender make-up and larger DM population (Table 3) will give a good contrast and data set with different population characteristics that may increase the model performance and generalizability [12,28].

XSL•FO
RenderX

Table . Comparison of basic population characteristics between the previous clinical study at KK Women's Hospital, Singapore, and the current study in Krida Wacana Christian University (Ukrida) Hospital, Indonesia.

Characteristics	KK Women's Hospital	Ukrida Hospital
Age (y), mean (SD)	38.73 (10.60)	44.22 (11.79)
Female proportion, n/N (%)	436/500 (90.3)	535/885 (60.45)
BMI (kg/m^2), mean (SD)	24.4 (5.13)	26.77 (5.08)
Type 2 DM[a] proportion, n/N (%)	157/500 (31.40)	473/885 (53.44)

[a]DM: diabetes mellitus.

The study has a comprehensive 4 time points of data collection, which is more extensive [33,34], and will provide a more complete view of how blood glucose level fluctuates between different state of food intake. The additional 1 and 3 hours after the meal may give a clearer view of blood glucose level changes compared to the standard fasting and 2-hour postmeal that is already widely accepted. Additionally, the use of AI in assessing glycemic variability is very limited [35], and to our knowledge, no study using photoplethysmography-based sensors has been reported despite its strong potential use for DM management. Therefore, this study will also be giving new and novel insight on the topic.

Study Limitations

This protocol has certain limitations. We prioritized achieving a balanced distribution of gender and participants with and without diabetes. With this focus, efforts were made to increase the overall number of participants. As the sample size grew, a broader range of demographic characteristics was naturally captured, reflecting the inherent diversity of the larger Indonesian population. However, no deliberate strategies were employed to design or randomize demographic features; the observed diversity emerged organically from the expanded participant pool. Several demographic features, such as participants aged over 60, are underrepresented; this will undermine the understanding of photoplethysmography in BGEM's predictive ability in elderly population. Although the study includes a diverse range of ethnicities, the overall ethnic composition does not reflect the general proportions of the Indonesian population. This limits the extent to which certain inferences can be drawn from the final data analysis. Future studies with more proportional sampling from the general Indonesian population should address this bias.

Conclusions

This protocol paper outlines the methodology designed for evaluating BGEM, a photoplethysmography-based AI model designed to estimate blood glucose levels and diabetic risk. The clinical trial was conducted on Indonesian participants with and without DM, while considering various influencing factors. The protocol is distinctive as it aims to generate high-quality training data for the AI model, ensuring balanced data structure, sufficient data quantity, diversity, and comprehensive relevant information including unique data on glycemic variability using wearables for downstream analysis and future research.

Acknowledgments

The authors did not use generative AI in any portion of the manuscript.

Funding

This research was funded mainly by Actxa Pte Ltd, with support from LIF Indonesia and Krida Wacana Christian University (Ukrida) Internal Grant, while data collection was performed at Ukrida hospital in Jakarta, Indonesia.

Authors' Contributions

Conceptualization: EWS (lead), SSD (equal), SBA (equal) Data curation: AL (lead), AD (equal), EZ (Supporting), PP (Supporting), EGM (supporting), FF (supporting) Formal analysis: ZLF (lead), DY (equal), SSD (equal), AL (supporting), EZ (supporting), HHW (supporting), JYAT (supporting), FF (supporting) Funding acquisition: EWS Investigation: EWS (lead), AD (equal), AL (supporting), EZ (supporting) Methodology: EWS (lead), MS (equal), AD (equal), SWD (equal), ASB (equal) Project administration: EWS (lead), AD (equal), ZLF (equal), AL (supporting) Resources: EWS (lead), AD (equal), AL (equal), EGM (supporting) Software: ZLF (lead), DY (equal), AL (supporting), EZ (supporting), HHW (supporting), JYAT (supporting) Supervision: SSD (lead), MS (equal), SBA (equal) Validation: ZLF (lead), DY (equal), AL (supporting), EZ (supporting), HHW (supporting), JYAT (supporting) Visualization: EWS (lead), ZLF (equal), AD (supporting), EZ (supporting), PP (supporting), FF (supporting), EGM (supporting) Writing – original draft: EWS (lead), ZLF (equal) Writing – review & editing: SWD (lead), EWS (equal), ZLF (equal), DY (equal),

XSL•FO
RenderX

Conflicts of Interest

SBA is a medical advisor of Actxa Pte Ltd. AL, EZ, and ZFL are employees of Actxa Pte Ltd. JYAT is an intern attached to the study. SSD is a clinical research advisor of Actxa Pte Ltd. Other Ukrida or Living In Fitness authors have no competing interest.

Multimedia Appendix 1
Nutritional value breakdown of each ingredient from the meal set given to break the subjects' fast.
[DOCX File, 24 KB - resprot_v15i1e76558_app1.docx]

References

1. Saeedi P, Petersohn I, Salpea P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. Diabetes Res Clin Pract 2019 Nov;157:107843. [doi: 10.1016/j.diabres.2019.107843] [Medline: 31518657]
2. Ogurtsova K, da Rocha Fernandes JD, Huang Y, et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract 2017 Jun;128:40-50. [doi: 10.1016/j.diabres.2017.03.024] [Medline: 28437734]
3. Magliano DJ, Boyko EJ, Genitsaridi I, Piemonte L, Riley P, Salpea P, editors. IDF Diabetes Atlas, 11th edition: International Diabetes Federation; 2025. URL: https://diabetesatlas.org/media/uploads/sites/3/2025/04/IDF_Atlas_11th_Edition_2025.pdf [accessed 2025-12-31]
4. The Lancet Diabetes Endocrinology. Undiagnosed type 2 diabetes: an invisible risk factor. Lancet Diabetes Endocrinol 2024 Apr;12(4):215. [doi: 10.1016/S2213-8587(24)00072-X] [Medline: 38460527]
5. Manne-Goehler J, Geldsetzer P, Agoudavi K, et al. Health system performance for people with diabetes in 28 low- and middle-income countries: a cross-sectional study of nationally representative surveys. PLoS Med 2019 Mar;16(3):e1002751. [doi: 10.1371/journal.pmed.1002751] [Medline: 30822339]
6. Zhou B, Rayner AW, Gregg EW, Pessôa dos Prazeres TM. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. Lancet 2024 Nov 23;404(10467):2077-2093. [doi: 10.1016/S0140-6736(24)02317-1] [Medline: 39549716]
7. Ali MK, Imperatore G, Benoit SR, et al. Impact of changes in diabetes screening guidelines on testing eligibility and potential yield among adults without diagnosed diabetes in the United States. Diabetes Res Clin Pract 2023 Mar;197:110572. [doi: 10.1016/j.diabres.2023.110572] [Medline: 36775024]
8. Misra A, Gopalan H, Jayawardena R, et al. Diabetes in developing countries. J Diabetes 2019 Jul;11(7):522-539. [doi: 10.1111/1753-0407.12913] [Medline: 30864190]
9. Wahidin M, Achadi A, Besral B, et al. Projection of diabetes morbidity and mortality till 2045 in Indonesia based on risk factors and NCD prevention and control programs. Sci Rep 2024 Mar 5;14(1):5424. [doi: 10.1038/s41598-024-54563-2] [Medline: 38443384]
10. Zanelli S, Ammi M, Hallab M, El Yacoubi MA. Diabetes detection and management through photoplethysmographic and electrocardiographic signals analysis: a systematic review. Sensors (Basel) 2022 Jun 29;22(13):4890. [doi: 10.3390/s22134890] [Medline: 35808386]
11. Gusev M, Poposka L, Spasevski G, et al. Noninvasive glucose measurement using machine learning and neural network methods and correlation with heart rate variability. J Sensors 2020 Jan 6;2020:1-13. [doi: 10.1155/2020/9628281]
12. Shi B, Dhaliwal SS, Soo M, et al. Assessing elevated blood glucose levels through blood glucose evaluation and monitoring using machine learning and wearable photoplethysmography sensors: algorithm development and validation. JMIR AI 2023 Oct 27;2:e48340. [doi: 10.2196/48340] [Medline: 38875549]
13. Ang SB, Chua MT, Shi B, Chan SHC, Liaw CSK, Dhaliwal SS. Utility of photoplethysmography in detecting elevated blood glucose among non-diabetics. Singapore Med J 2024 Sep 17. [doi: 10.4103/singaporemedj.SMJ-2023-156] [Medline: 39287509]
14. Chu J, Yang WT, Lu WR, Chang YT, Hsieh TH, Yang FL. 90% accuracy for photoplethysmography-based non-invasive blood glucose prediction by deep learning with cohort arrangement and quarterly measured HbA1c. Sensors (Basel) 2021 Nov 24;21(23):7815. [doi: 10.3390/s21237815] [Medline: 34883817]
15. Sen Gupta S, Kwon TH, Hossain S, Kim KD. Towards non-invasive blood glucose measurement using machine learning: An all-purpose PPG system design. Biomed Signal Process Control 2021 Jul;68:102706. [doi: 10.1016/j.bspc.2021.102706]
16. Kwon TH, Hossain S, Turja MS, Kim KD. Design and validation of a Monte Carlo method for the implementation of noninvasive wearable devices for HbA1c estimation considering the skin effect. Micromachines (Basel) 2024 Aug 24;15(9):1067. [doi: 10.3390/mi15091067] [Medline: 39337727]
17. Suh S, Kim JH. Glycemic variability: how do we measure it and why is it important? Diabetes Metab J 2015 Aug;39(4):273-282. [doi: 10.4093/dmj.2015.39.4.273] [Medline: 26301188]
18. Bermingham KM, Smith HA, Gonzalez JT, et al. Glycaemic variability, assessed with continuous glucose monitors, is associated with diet, lifestyle and health in people without diabetes. Res Sq 2023 Oct 30:rs.3.rs-3469475. [doi: 10.21203/rs.3.rs-3469475/v1] [Medline: 37961419]

XSL•FO
RenderX

19. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Int J Nurs Stud 2010 Aug;47(8):931-936. [doi: 10.1016/j.ijnurstu.2009.10.001]
20. Zou X, Zhou X, Ji L, et al. The characteristics of newly diagnosed adult early-onset diabetes: a population-based cross-sectional study. Sci Rep 2017 Apr 19;7:46534. [doi: 10.1038/srep46534] [Medline: 28422176]
21. Xiong X, Wei L, Xiao Y, et al. Effects of family history of diabetes on pancreatic β-cell function and diabetic ketoacidosis in newly diagnosed patients with type 2 diabetes: a cross-sectional study in China. BMJ Open 2021 Jan 11;11(1):e041072. [doi: 10.1136/bmjopen-2020-041072] [Medline: 33431489]
22. Liu XL, Willis K, Fulbrook P, Wu CJJ, Shi Y, Johnson M. Factors influencing self-management priority setting and decision-making among Chinese patients with acute coronary syndrome and type 2 diabetes mellitus. Eur J Cardiovasc Nurs 2019 Dec;18(8):700-710. [doi: 10.1177/1474515119863178] [Medline: 31319694]
23. Donazar-Ezcurra M, Lopez-Del Burgo C, Martinez-Gonzalez MA, Basterra-Gortari FJ, de Irala J, Bes-Rastrollo M. Soft drink consumption and gestational diabetes risk in the SUN project. Clin Nutr 2018 Apr;37(2):638-645. [doi: 10.1016/j.clnu.2017.02.005] [Medline: 28262323]
24. Penelitian A, Aritonang JP, Widiastuti IAE, Harahap IL. Gambaran Tingkat Aktivitas Fisik Mahasiswa Pendidikan Dokter Fakultas Kedokteran Universitas Mataram di Masa Pandemi COVID-19 [Description of physical activity level of medical students Faculty of Medicine, University of Mataram in COVID-19 pandemic]. eJournal Kedokteran Indonesia 2022;10(1):58-63. [doi: 10.23886/ejki.10.129.58-63]
25. Koshy AN, Sajeev JK, Nerlekar N, et al. Utility of photoplethysmography for heart rate estimation among inpatients. Intern Med J 2018 May;48(5):587-591. [doi: 10.1111/imj.13777] [Medline: 29722189]
26. Cherry A, Nasser A, Salameh W, Abou Ali M, Hajj-Hassan M. Real-time PPG-based biometric identification: advancing security with 2D gram matrices and deep learning models. Sensors (Basel) 2024 Dec 25;25(1):40. [doi: 10.3390/s25010040] [Medline: 39796830]
27. Pfützner A, Klonoff DC, Pardo S, Parkes JL. Technical aspects of the Parkes error grid. J Diabetes Sci Technol 2013 Sep 1;7(5):1275-1281. [doi: 10.1177/193229681300700517] [Medline: 24124954]
28. Zeynali M, Alipour K, Tarvirdizadeh B, Ghamari M. Non-invasive blood glucose monitoring using PPG signals with various deep learning models and implementation using TinyML. Sci Rep 2025 Jan 2;15(1):581. [doi: 10.1038/s41598-024-84265-8] [Medline: 39753714]
29. Januraga PP, Lukitosari E, Luhukay L, Hasby R, Sutrisna A. Mapping key populations to develop improved HIV and AIDS interventions: multiphase cross-sectional observational mapping study using a district and city approach. JMIR Public Health Surveill 2025 Jan 30;11:e56820. [doi: 10.2196/56820] [Medline: 39883483]
30. Priyanto H, Aulia FA, Kahar H, Faizi M, Marpaung FR, Aryati A. Neonatal thyroid-stimulating hormone reference intervals in multi-ethnics population. Children (Basel) 2025 Jan 17;12(1):104. [doi: 10.3390/children12010104] [Medline: 39857937]
31. Metherall B, Berryman AK, Brennan GS. Machine learning for classifying chronic kidney disease and predicting creatinine levels using at-home measurements. Sci Rep 2025 Feb 5;15(1):4364. [doi: 10.1038/s41598-025-88631-y] [Medline: 39910170]
32. Tang Y, Yang D. Overcoming dietary complexity in type 2 diabetes: influencing factors and coping strategies. Eur J Med Res 2025 Feb 6;30(1):82. [doi: 10.1186/s40001-025-02318-8] [Medline: 39910637]
33. Rodin D, Kirby M, Sedogin N, Shapiro Y, Pinhasov A, Kreinin A. Comparative accuracy of optical sensor-based wearable system for non-invasive measurement of blood glucose concentration. Clin Biochem 2019 Mar;65:15-20. [doi: 10.1016/j.clinbiochem.2018.12.014] [Medline: 30629956]
34. Monte-Moreno E. Non-invasive estimate of blood glucose and blood pressure from a photoplethysmograph by means of machine learning techniques. Artif Intell Med 2011 Oct;53(2):127-138. [doi: 10.1016/j.artmed.2011.05.001] [Medline: 21696930]
35. Zhang K, Qi Y, Wang W, et al. Future horizons in diabetes: integrating AI and personalized care. Front Endocrinol (Lausanne) 2025;16:1583227. [doi: 10.3389/fendo.2025.1583227] [Medline: 40213102]

Abbreviations

AI: artificial intelligence
BGEM: Blood Glucose Evaluation and Monitoring
DM: diabetes mellitus
HbA_{1c}: hemoglobin A_{1c}
HR: heart rate
ML: machine learning
Ukrida: Krida Wacana Christian University

XSL•FO
RenderX

Edited by J Sarvestan; submitted 26.Apr.2025; peer-reviewed by I Rinaldi, MI Rusydi; revised version received 08.Dec.2025; accepted 09.Dec.2025; published 08.Jan.2026.

Please cite as:
Suradji EW, Dhaliwal SS, Li-Feng Z, Zhou E, Lim A, Wong HH, Du Y, Fredicia F, Dianky A, Pasaribu P, Masayoshi Takashi EG, Tan JYA, Santoso M, Ang SB
Assessment of Blood Glucose Measurement Using New Noninvasive Technology: Protocol and Methodology
JMIR Res Protoc 2026;15:e76558
URL: https://www.researchprotocols.org/2026/1/e76558
doi:10.2196/76558

© Eka Widrian Suradji, Satvinder Singh Dhaliwal, Zhang Li-Feng, Entong Zhou, Amos Lim, Hui Hern Wong, Yao Du, Fredicia Fredicia, Amanda Dianky, Puspasari Pasaribu, Ernest Gamaliel Masayoshi, Justin Yong An Tan, Mardi Santoso, Seng Bin Ang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 8.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Personalized Machine Learning Intervention to Improve Sleep Quality Using Wearable Technology in Healthy Middle-Aged Adults From Mexico City: Protocol for a Pilot Randomized Controlled Trial

Rodrigo Quezada Reyes, MDS; Luis A Trejo, PhD

Computer Science Department, School of Engineering and Science, Tecnologico de Monterrey, Carretera al Lago de Guadalupe Km 3.5, Col. Margarita Maza de Juarez, Atizapán de Zaragoza, Estado de México, Mexico

Corresponding Author:
Luis A Trejo, PhD
Computer Science Department, School of Engineering and Science, Tecnologico de Monterrey, Carretera al Lago de Guadalupe Km 3.5, Col. Margarita Maza de Juarez, Atizapán de Zaragoza, Estado de México, Mexico

Abstract

Background: In 2019, global sleep surveys reported that 80% of adults want to improve their sleep quality, and in 2021, 45% were reported to be dissatisfied with their sleep. In 2025, among American adults, 37% reported sleep dissatisfaction and 38% reported not feeling energized after sleep. These findings are consistent with data from the 2016 nationally representative survey of Mexican adults (aged ≥18 years), in which 37% reported sleep problems.

Objective: This protocol describes a pilot randomized controlled trial (RCT) testing whether a single personalized sleep intervention driven by machine learning (ML) using consumer wearable data can improve sleep scores compared with generic sleep hygiene education in healthy middle-aged adults from Mexico City.

Methods: This pilot RCT plans to enroll 32 participants (16 per arm; stratified by sex) in Mexico City. All participants wear Samsung Galaxy Watch 4 devices for 60 days. During days 1 - 30 (baseline), objective sleep data (10 variables related to duration, efficiency, sleep stages, movements, cycles, and recovery metrics) are collected. The control group receives generic sleep hygiene education. The experimental group receives personalized recommendations on day 30 based on the top predictive sleep parameters identified by ML models using Shapley Additive Explanations analysis and recursive feature elimination. The primary outcome is the sleep score (scale 1 - 100; composite device metric) during days 31 - 60, which is analyzed using analysis of covariance, with the baseline sleep score as a single covariate. The secondary outcome is the Pittsburgh Sleep Quality Index (PSQI) global score (scale 0 - 21; subjective validation), which is assessed at baseline, day 30, and day 60.

Results: The study is currently in progress. Recruitment started in August 2024 and ended in July 2025. Data collection is expected to be completed by December 2025. The study will compare 960 nights from the control group with 960 nights from the experimental group to explore whether ML interventions can improve sleep scores using wearable technology, generate a dataset from objective data for iterative model training and analysis, correlate objective and subjective sleep quality metrics, and establish whether a feasible framework for proactive sleep quality approaches can be developed. The results will be available by March 2026, and the findings will be submitted for publication within 6 months of study completion.

Conclusions: This pilot study establishes the feasibility and preliminary effect size for ML-personalized sleep interventions using consumer wearables within a manufacturer-independent personalization framework. The approach combines objective device monitoring with a subjective measure (PSQI) to test whether precision targeting of individual sleep parameters outperforms generic recommendations. If validated, this methodology could advance sleep interventions from universal protocols toward individualized behavioral targeting. The resulting dataset will enable model refinement and provide preliminary evidence for scaling personalized sleep health interventions in healthy populations.

International Registered Report Identifier (IRRID): DERR1-10.2196/76415

(JMIR Res Protoc 2026;15:e76415) doi:10.2196/76415

KEYWORDS

sleep quality; well-being; machine learning; smartwatches; sensors; Pittsburgh Sleep Quality Index; PSQI; wearables

XSL•FO
RenderX

Introduction

Overview

Sleep is essential for health and productivity. Global sleep dissatisfaction has been reported by 35% - 45% of adults, and this is consistent with 37% of Mexican adults reporting sleep problems [1-5]. Sleep benefits are well researched, and they include (1) a healthy body through immune system efficiency [6,7], (2) healthy emotions through psychological well-being [8], and (3) a healthy mind through memory consolidation [9-11]. The negative consequences of poor sleep are also well researched, and they include heart disease [12], metabolic syndrome [13], high blood pressure [14], diabetes [15], obesity [16], common cold [6], higher probability of injury [17], and higher mortality risk [18-20].

Demographic and Geographic Context

This pilot study will be conducted in Mexico City, a metropolitan area with over 21 million residents [21]. Adults aged 30 - 60 years represent over 40% of Mexico's economically active population [22], making this demographic relevant for sleep health research. Urban workers in Mexico City face unique sleep challenges, including long commutes, shift work, and environmental stressors, that may impact sleep quality [5,23,24]. National surveys indicate that 35% - 40% of Mexican adults sleep fewer than 7 hours per night, with 37% reporting sleep problems [4,5]. Mexico City's context provides a useful setting to explore sleep intervention feasibility in a rapidly urbanizing environment [25,26].

Knowledge Gap

Despite growing interest in wearable sleep technology, a significant knowledge gap exists regarding whether machine learning (ML) analysis of individual sleep patterns from commercial devices can generate actionable recommendations that improve objective sleep outcomes. No previous studies have examined whether sleep scores from wearables (Samsung Galaxy Watch [27] or equivalent) can serve as primary outcome measures for personalized ML-derived interventions validated against the Pittsburgh Sleep Quality Index (PSQI) [28]. Existing consumer sleep coaching, such as Samsung's Sleep Coach, categorizes users into predefined "sleep animal" types based on general sleep patterns (sleep time, sleep consistency, and awake time) and provides standardized multiweek programs [29,30] without subjective validation. Recent advances focus on validation studies comparing device accuracy to polysomnography (PSG) and developing ML algorithms for sleep stage classification [31-34]. Samsung has collected over 716 million nights of sleep data (collected between June 2021 and May 2023 via the Galaxy Watch series) globally [35], and research using Samsung Galaxy Watch sleep scores as intervention targets remains unexplored.

Study Uniqueness

This pilot study addresses the gaps in the literature by exploring if wearable sleep scores [27,29,30,36-43] can serve as primary outcome measures for personalized ML-derived interventions and showing if ML analysis of individual patterns can identify the predictors of improvement, with validation against the PSQI (structured 30-day correlation framework) [28], in healthy middle-aged Mexican adults, a population facing unique urban sleep challenges but underrepresented in sleep technology research. This represents a personalized, ML-driven sleep intervention using a consumer wearable sleep data source (sleep metrics) for personalization as well as primary outcome measures (sleep scores) and user subjective assessments, establishing a novel paradigm for sleep health.

Core Methodological Innovation

This pilot study tests a single, focused innovation: a personalization framework that generates behavioral sleep interventions from ML analysis of individual patterns in consumer wearable data. Our innovation is the *methodology*, not the device, theories, or algorithms. We are not validating device accuracy or comparing algorithms. Existing studies have established these foundations [36-43]. We are testing whether personalized behavioral interventions derived from individual ML analysis of the sleep metrics of a consumer device can improve objective sleep scores better than generic sleep hygiene education. This methodology is intentionally device-agnostic. The same personalization framework can be applied to other manufacturers' wearables, providing continuous numeric sleep data and sleep score scales. We perform tests within Samsung's ecosystem to isolate the intervention effect (personalized vs generic) without confounding by device differences. If successful, this will establish a transferable approach that other researchers can replicate across manufacturers, potentially driving the field toward less proprietary and more user-centered sleep optimization tools.

Sleep Measurements

There are 2 ways to measure sleep. The first involves subjective measures of sleep that assess patient-reported satisfaction and functional impact (essential domains reflecting clinical relevance) and have evolved from sleep diaries [44] into validated survey instruments like the gold-standard PSQI [28] for sleep quality and sleep patterns, the Epworth Sleepiness Scale for daytime sleepiness [45], the Patient-Reported Outcomes Measurement Information System for sleep disturbances [46], and the Sleep Health Index for general sleep health [47]. The second involves objective measures of sleep that capture physiological parameters not consciously perceived (sleep parameters [measurable components] such as efficiency, latency, architecture, and fragmentation). Sensors collect objective measures of sleep by measuring physical inputs [31-34,48-51]. The most important biomarkers for sleep tracking are heart rate, breathing rate, blood oxygen flow, and body movements [51,52].

Sleep devices can be subdivided by grade into (1) medical-grade devices (also called clinical-grade devices), such as the gold-standard PSG [53,54]; and (2) commercial-grade devices (also called consumer-grade devices or consumer sleep technology [55]), such as off-the-shelf smartwatches [43]. They can also be subdivided by their body proximity into (1) devices in contact with the body (wearables) like watches, patches, rings, and bands [56]; (2) devices near the body (nearables) like bedside stands and devices in or under the mattress [57]; and (3) devices inside the body (implants) like subcutaneous

XSL•FO
RenderX

electroencephalography devices [58,59]. Crivello et al [60] categorized sleep monitoring technologies by deployment as follows: laboratory-based (eg, PSG and specialized electroencephalography), ambulatory or portable (eg, wearables, on-bed sensors, actigraphy devices, and contactless devices), and consumer integration (eg, smartphone apps and smart home sensors).

In 2019, the wearable sleep tracker segment represented 44% of global smart sleep-tracking products [57]. The most common body locations of wearables are the head, chest, wrist, and finger. The wrist is the most popular location, representing more than 58% of the segment in 2024 [61].

PSG is the gold standard for objective sleep measures, but requires expensive equipment. Type 1 PSG requires technologist labor and lab settings, while type 2 PSG can be performed at home with or without a technologist [53,54,62-66]. The cost of long-term sleep monitoring with type 1 PSG or type 2 PSG is very high [66-68]. However, PSG is instrumental in setting a baseline to compare recent technological advances with wearable objective measurements for identifying an affordable and noninvasive option for continuous monitoring [36,62,69-71].

As the lines between consumer and clinical devices are blurring [55], there is an ongoing debate regarding the trustworthiness of available technologies. In 2018, the American Academy of Sleep Medicine stated that for consumer sleep technologies to provide a diagnosis or treatment, they first need Food and Drug Administration clearance and testing against current gold standards [72]. Studies have reported overall overestimated sleep with these devices compared with PSG due to lower sensitivity to detect wake epochs [73] and more discrepancies in sleep disorder populations [74-76].

Commercial-grade wearable devices should be validated with PSG to ensure reliability. A study in 2017 reported high agreement between a commercial wearable device and PSG in good sleepers but low agreement in patients with insomnia [75]. Another study in 2019 showed comparable results for the rapid eye movement (REM) sleep phase [63]. A study in 2020 concluded that wearables can accurately measure sleep and cardiorespiratory variables in healthy people [77], and a recent study reported the accuracy of a wearable device in sleep stage identification [65].

This study focuses on healthy adults without diagnosed sleep disorders [78-81], as this population represents 63% - 67% of adults [82,83]. In 2004, only 9%-11% of people were considered to have chronic insomnia by the World Health Organization Regional Office for Europe [84]. In 2022, a study reported that 10%-15% of people have chronic insomnia with daytime consequences [85], and another source reported that up to 22% of people meet the insomnia sleep disorder criteria [86]. These findings align with the overall percentage of adults not getting enough sleep (33.3%-36.8%) from 2013 to 2022 [82,83].

Sleep Quantity

Sleep quantity refers to the amount of daily time spent sleeping (sleep duration), measured in hour units. The American Academy of Sleep Medicine, the Sleep Research Society, and the National Sleep Foundation agree that adults (aged 18 - 60 years) need a minimum of 7 hours of sleep each night [87,88], and a sleep duration less than that implies a sleep deprivation condition (insufficient sleep) [18]. The following variables can be considered for assessment: (1) time in bed (TIB), period from lights off to lights on; (2) total sleep time (TST), actual time spent sleeping (sleep quantity) consistently shorter than TIB unless efficiency reaches 100%; and (3) sleep efficiency ([TST / TIB] × 100).

There are four main disadvantages of relying on this metric in isolation: (1) quantity increase does not guarantee a positive correlation with sleep quality (eg, during the 2021 pandemic period, Americans were sleeping more [quantity], but sleep quality declined) [89]; (2) TIB can be misunderstood as TST (these are only equal if 100% sleep efficiency is achieved) [90-92]; (3) sleeping beyond a limit can have negative consequences, such as stroke [93] and intracerebral hemorrhage [94] (the National Sleep Foundation has recommended a maximum sleep duration of 9 hours [88]); and (4) sleeping time cannot provide valuable data on sleep architecture, including time spent in each of the 4 sleep stages, number of sleep cycles, distribution of sleep stages, and transition speed between sleep phases [90-92].

Sleep Quality: Definition and Measurement

Sleep quality or restorative sleep represents a multidimensional construct encompassing both objective parameters and subjective restoration. Nelson et al [95] identified 4 core attributes: sleep efficiency, latency, duration, and wake after sleep onset. Contemporary understanding extends beyond these continuity attributes to include architecture components (adequate deep sleep for physical recovery, sufficient REM sleep for cognitive restoration, and optimal 4 - 6 sleep cycles) and minimal fragmentation [95-99]. An increase in quantity does not guarantee a positive correlation with sleep quality [89]. The 2014 and 2015 survey results by the National Sleep Foundation showed an 11-point delta in sleep duration compared with the sleep quality score and a final 69/100 score for sleep quality [47]. Despite China's records showing that the general population has an average sleep duration of 7.31 hours [100] and Chinese university students reporting an average sleep duration of 7.08 hours [101], studies on the prevalence of poor sleep quality have been published [102].

Sleep Quality in Healthy Populations

Rosipal et al [103] identified objective components for sleep quality indexing in normal sleep. Moreover, O'Donnell et al [104] found significant subjective-objective relationships in healthy older study participants. Even healthy individuals show consequences, including cognitive decrements, mood alterations, fatigue, and behavioral adaptations (increased caffeine use).

Related Studies

In 2018, Guillodo et al [105] concluded that wearables for sleep monitoring are acceptable, and de Arriba-Pérez et al [71] concluded that the use of consumer-grade wearables as sleep indicators is a feasible option, although some require more rigor. In 2023, Lins et al [106] concluded that the photoplethysmography technology of the Samsung Galaxy Watch 4 was comparable to the oscillometric reference method

XSL•FO
RenderX

for blood pressure determination. Moreover, Gaiduk et al [36] reported the feasibility of measuring several relevant sleep characteristics with Samsung Galaxy Watch 4 as a sensor-based approach in a real-life setting, if some level of inaccuracy is accepted. In 2024, Alshamari and Althobaiti [107] compared wearable sleep tracker–related improvements and concluded that a WHOOP device can both improve the perception of quality and provide accurate measures of sleep and cardiorespiratory variables.

Methods

Conformance to Reporting Standards

This protocol conforms to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 Statement. The SPIRIT 2025 checklist of items to address in a randomized trial protocol is presented in Checklist 1.

Trial Design

This is a parallel-group randomized controlled pilot trial with a 1:1 allocation ratio. Participants are stratified by sex (male/female) and randomized to either (1) a control group receiving generic sleep hygiene education or (2) an experimental group receiving personalized ML-derived behavioral recommendations.

The observation period is 60 days, which aligns with PSQI requirements for 30-day data collection periods. PSQI assessments are performed at (1) baseline (prior to device use), (2) day 30 (first monitoring period), and (3) day 60 (postintervention period). This enables subjective-objective correlation while establishing a baseline (days 1 - 30) to be used as a covariate for intervention analysis (days 31 - 60). While 30 days may be minimal for entrenching behavior change, this pilot will establish feasibility and preliminary effect sizes for future extended trials.

Patient and Public Involvement

No patient or public involvement has been planned in the design, conduct, or reporting of this pilot trial owing to resource constraints for this project. Future definitive trials will incorporate patient and public involvement in intervention design, outcome selection, and dissemination strategies.

Study Setting

This pilot study is conducted in Mexico City, a metropolitan area with over 21 million residents. Participants sleep in their habitual home setting while wearing a Samsung Galaxy Watch 4 continuously. All study interactions (consent, PSQI administration, device distribution, intervention delivery, and equipment return) occur in Mexico City through in-person meetings with a researcher.

Eligibility Criteria

Eligibility Criteria Operationalization

The operationalization of the eligibility criteria is as follows:

- Sleep disorder screening: structured interview regarding diagnosed disorders, medications, and medical consultations; self-declaration (no PSG screening)
- Medical/psychiatric assessment: standardized questionnaires (no chart review)
- Nationality and age: birth certificate upon request
- Active workforce: full-time employment (≥40 hours/week) verified through documentation
- Limitations: self-report with documentation appropriate for a community-based pilot

Inclusion Criteria

The inclusion criteria are as follows: (1) Mexican nationality, (2) age 30 - 60 years (minimizes age-related sleep physiology variability), (3) male or female gender, (4) active workforce status during the study, (5) residence in Mexico City during the study, (6) willingness and ability to use the wearable device for 60 days (at least during sleep), and (7) consent for providing sensitive personal data for the experiment.

Exclusion Criteria

The exclusion criteria are as follows: (1) diagnosis of chronic sleep disorders (eg, insomnia) requiring specific treatment, (2) major psychiatric or neurological conditions affecting sleep, (3) substance abuse or significant medical conditions impacting sleep quality, and (4) inability to comply with study requirements (eg, wearing the smartwatch).

Recruitment: Participant Recruitment and Screening

The recruitment methods are as follows: (1) physical (study invitations posted on university and public hospital bulletin boards), (2) digital (study invitations on professional networking platforms like LinkedIn), and (3) referral (snowball sampling through professional networks).

Screening Process

The screening process begins with initial contact from interested participants. Structured remote interviews are then conducted to screen participants based on the inclusion and exclusion criteria. Only participants meeting all criteria proceed to an in-person meeting for informed consent and verification of eligibility documentation. Retention leverages the intrinsic motivation of participants for sleep improvement, along with structured touchpoints at the 3 assessments.

Interventions

Control Intervention: Generic Sleep Hygiene Education

Control group participants receive standardized sleep hygiene education on day 30, covering general recommendations applicable to all adults. Examples include maintaining a consistent sleep-wake schedule and sleeping between 7 and 9 hours per night. These recommendations are not personalized to individual sleep patterns.

Experimental Intervention: Personalized ML-Derived Recommendations

Experimental group participants receive personalized behavioral recommendations from a predefined library (Multimedia Appendix 1) on day 30 based on individual ML model analysis of their baseline sleep data (days 1 - 30).

XSL•FO
RenderX

ML Framework for Personalized Interventions

Regarding data structure and variables, Samsung Galaxy Watch 4 generates 1 target variable (sleep_score) and 10 explanatory variables (sleep_duration, movement_awakening, sleep_cycle, mental_recovery, physical_recovery, efficiency, Transition_N1, Light_N2, Deep_N3, and REM) per sleep session. Through preprocessing, 5 additional features are engineered (real_total_sleep, Transition_pct, Light_pct, Deep_pct, and REM_pct), and temporal (Night_Number) and identifier (Participant_ID) variables are added, creating an 18-column dataset. For ML analysis, 15 variables are used as ML input variables (excluding sleep_score as the target, and Participant_ID and Night_Number as nonpredictive identifiers).

The ML framework serves as the mechanism for generating personalized interventions, not as the study outcome. We use established methods (eg, random forest and SHAP [Shapley Additive Explanations]) documented in ML literature to identify each participant's top predictive sleep parameters. Our innovation is applying these tools to generate behavioral recommendations from consumer device data, not developing new algorithms.

Model Architecture and Development

The 15 Samsung Galaxy Watch 4 explanatory variables and additional engineered features serve as the base dataset. The training-validation framework involves a temporal split using baseline period data (days 1 - 30) with a 70/30 split for training/validation, cross-validation through 5-fold time-series cross-validation to prevent data leakage, and individual modeling for separate models trained for each participant. Models include random forest regression (primary model for interpretability), gradient boosting (extreme gradient boosting [XGBoost]) for comparison, and multiple linear regression as baseline, with selection based on cross-validation root mean square error and R^2. Feature importance and selection involve SHAP values for feature importance ranking, recursive feature elimination to identify the top 5 predictors per participant, and correlation analysis to avoid multicollinearity (threshold: $r>0.8$). Personalized recommendation generation is based on individual feature importance ranking, and participants receive targeted recommendations from a predefined library (Multimedia Appendix 1), with a focus on their top 2 - 4 modifiable sleep parameters.

Intervention Delivery and Verification

Experimental group participants receive personalized recommendations explained by the researcher in person on day 30. Participants can clarify and provide feedback. A closure interview at day 60 verifies engagement and implementation attempts, and includes a subjective effectiveness assessment.

Personalized Sleep Intervention Framework

The intervention content is based on individual ML feature importance ranking, and participants receive nonpharmacological recommendations from a predefined library (Multimedia Appendix 1) targeting their top 2 - 4 predictive sleep parameters [108-116]. Example recommendations for TST optimization (if sleep_duration is a key predictor) are as follows: ensure wake-up time commitment for too much sleep and bedtime commitment for too little sleep. Example recommendations for sleep efficiency enhancement (if movement_awakening is a predictor) are as follows: use white noise, ensure a quiet environment, and restrict caffeine intake (avoid intake ≤6 hours before bedtime). Example recommendations for sleep timing regulation (if sleep_cycle is a predictor) are as follows: ensure a consistent sleep-wake schedule and keep the bedroom cool (around 18 °C - 20 °C).

Experimental group participants receive personalized recommendations, and the top 2 - 4 recommendations are explained by the researcher during the in-person meeting. Adherence tracking involves self-reported compliance assessed at study closure (0% - 100% scale). Control group participants receive generic sleep hygiene education without personalization by the researcher during the in-person meeting.

ML-Derived Heterogeneity by Design

Personalized interventions intentionally create between-participant variability within the experimental group, and this is the intervention's defining feature, not a statistical flaw. Participant A might receive recommendations targeting REM sleep and sleep timing, while participant B might receive advice on deep sleep and wake episodes, based on individual ML model findings. This heterogeneity is the core innovation: targeting individual-specific parameters rather than universal protocols.

With regard to statistical implications, within-group variance in the experimental arm may exceed variance in the control arm. This reduces statistical power but reflects real-world personalization implementation. If personalization is effective, the increased precision of targeting individual needs should overcome variance increases. If personalization shows no benefit despite this heterogeneity, the approach is not superior to simpler generic advice. We do not attempt to statistically "control for" intervention heterogeneity, as it represents the intervention. Exploratory analyses will descriptively examine whether certain recommendation types (eg, duration-focused vs architecture-focused recommendations) show differential effectiveness, informing future research on personalization boundaries.

Intervention Modifications

No modifications to assigned interventions are permitted during the 60-day study period. If participants cannot implement specific recommendations due to unforeseen circumstances, this is documented as nonadherence rather than modification. Adherence tracking captures whether participants attempted to follow recommendations, not whether recommendations were modified.

Concomitant Care Protocol

The aspects related to the concomitant care protocol are as follows:

- Permitted: regular medical care, stable chronic disease management, and emergency care
- Prohibited: new sleep medications, sleep clinic consultations, or other sleep apps or programs

XSL•FO
RenderX

- Documentation: weekly medication diary and immediate reporting of treatment changes
- Safety: participants with sleep deterioration (<30 score for >1 week) withdrawn and referred
- Baseline: usual care patterns documented at enrollment

Adherence Assessment

In the experimental group, adherence to personalized recommendations is assessed through self-report at the day 60 closure interview. Participants rate adherence on a scale from 0% to 100% for each recommendation provided. In the control group, adherence to generic sleep hygiene is assessed in a similar manner. Device wear adherence is monitored through objective data. Participants with missing data for >20% of nights may be excluded and replaced.

Outcomes

Primary Outcome: Mean Samsung Galaxy Watch 4 Sleep Score During the Postintervention Period

We are working within proprietary constraints. We use the Samsung sleep score (1-100 scale) as the primary outcome for the following reasons: (1) real-world relevance: the score is the metric users see and seek to optimize in consumer wearable ecosystems; (2) continuous sensitivity: the scoring scale (1 - 100) detects individual-level changes with precision, which categorical systems cannot achieve; (3) composite integration: the score algorithmically combines multiple sleep dimensions (duration, architecture, efficiency, and fragmentation), providing a holistic optimization target; (4) feasibility for personalization: continuous data for 60 nights enable individual ML models to identify predictive patterns; and (5) validation through correlation: secondary PSQI outcomes confirm that objective score improvements correspond to subjective sleep quality perception.

There are some limitations. The use of a proprietary algorithm limits the reproducibility of exact score computation. Moreover, there is no PSG validation of the composite score specifically (though component parameters are validated [36-43]). Furthermore, the manufacturer-specific metric precludes direct cross-device comparisons. These limitations are inherent to consumer wearable research and do not invalidate the personalization methodology (they contextualize the interpretation). Our framework is designed to work *within* these pragmatic constraints while remaining transferable to other manufacturers' proprietary systems.

Secondary Outcome: PSQI Global Score Measured at Three Time Points

The PSQI (0-21 scale) validates that objective sleep score improvements correspond to the subjective perception of sleep quality and functional impact, ensuring clinical relevance from the patient perspective. The PSQI is not used as a competing primary measure but rather as a tool to provide complementary validation that device-measured changes reflect meaningful improvements in lived experience.

Measurement Time Points

The measurement time points are as follows:

- Baseline (day 0): PSQI
- Days 1-30 (baseline period): continuous Samsung score collection for ML model training and baseline covariate establishment
- Day 30: PSQI; intervention delivery of personalized behavioral recommendations to the experimental group or intervention delivery of standardized sleep hygiene education covering general recommendations applicable to all adults
- Days 31 - 60 (postintervention period): continuous Samsung score collection for the primary outcome
- Day 60: PSQI; closure interview

Exploratory Outcomes

The exploratory outcomes include (1) individual sleep parameter changes (which specific variables improve), (2) adherence to personalized recommendations (self-reported 0% - 100% scale), and (3) behavioral confounders (alcohol consumption, exercise, and stress documented in the closure interview).

Participant Timeline

The measurement time points after eligibility screening, informed consent approval, and randomization in the control and experimental groups are presented in Figures 1 and 2, respectively.

XSL•FO
RenderX

Figure 1. Illustration of the key actions during the 60-day study period in the control group. PSQI: Pittsburgh Sleep Quality Index.

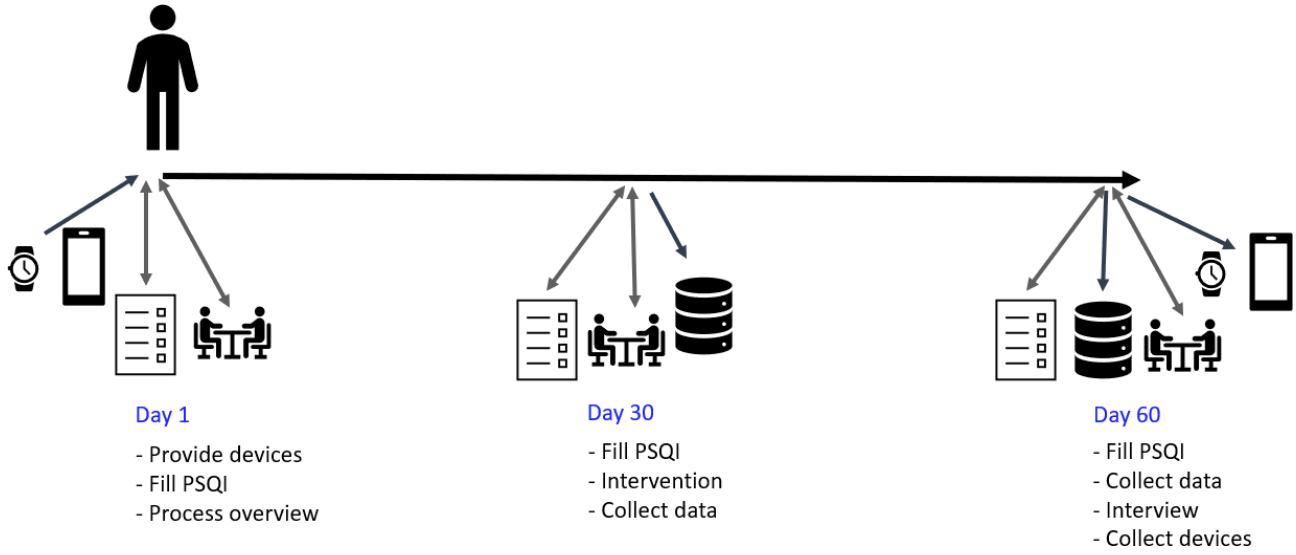

Figure 2. Illustration of key actions during the 60-day study period in the experimental group. PSQI: Pittsburgh Sleep Quality Index.

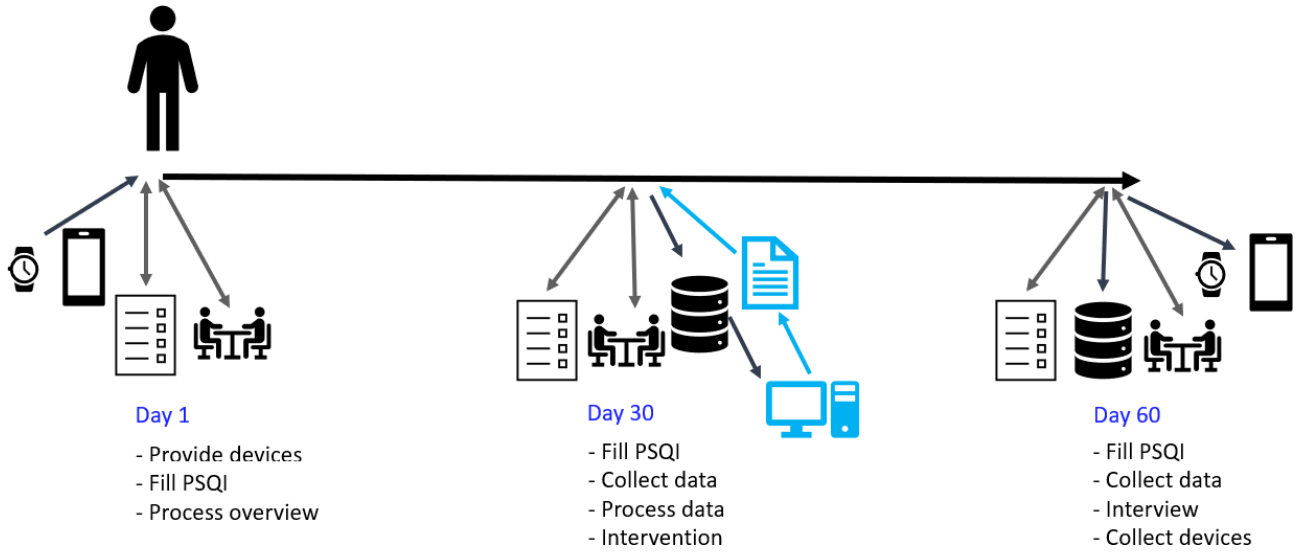

Sample Size: Pilot Feasibility Justification

This pilot trial enrolls 32 participants (16 per arm; stratified by sex), and this is consistent with established pilot study guidelines recommending 12 - 15 participants per arm for feasibility assessment [117-120]. The sample size is explicitly designed for pilot objectives, not definitive efficacy testing.

Pilot Objectives

The pilot objectives are as follows:

1. Feasibility: Can we successfully recruit, enroll, randomize, and retain participants? Can personalized ML interventions be delivered? Do participants engage with recommendations?
2. Preliminary effect size estimation: What magnitude of difference between groups might we observe? What variability exists within groups? What baseline-to-follow-up correlation informs covariate adjustment? The estimates can inform future randomized controlled trial (RCT) sample size calculations.
3. Methodology refinement: What challenges arise in data collection, ML model training, and intervention delivery? What proportion of nights yield usable data? Which sleep parameters show the most predictive value?
4. Hypothesis generation: Does personalization show a promising signal warranting a larger trial? Do certain participant subgroups respond better?

Power Analysis Context: Not Designed for Definitive Testing

The sample size provides adequate power (>80%) only for very large effects (d≥1.0). For moderate effects typical in sleep research (d=0.3 - 0.7), the power is insufficient, which is consistent with this study's role as a feasibility-focused pilot trial. We explicitly acknowledge insufficient power for moderate effects. The power analysis table showing power for different effect sizes is presented in Multimedia Appendix 2. If moderate-to-large effects exist, we may detect a signal, and if effects are small, this pilot will yield null results, informing researchers that larger samples, longer interventions, or different

XSL•FO
RenderX

approaches are needed. Both outcomes are valuable for the advancement of the field.

Sample Size for a Future RCT

Based on pilot effect sizes and variability estimates, a future definitive RCT would require approximately 175 participants per arm (350 participants in total) to detect moderate effects (d=0.5), considering 80% power, α of .05, and 15% attrition. Pilot findings will refine these projections.

Assignment of Interventions (Allocation)

Sequence generation is performed through computer-generated randomization using Python (numpy.random.RandomState with seed=42 for reproducibility). Participants are stratified by sex (male/female) to ensure balance. Block randomization with randomly permuted block sizes (2, 4, 6) is used to maintain unpredictability while ensuring approximately equal group sizes throughout recruitment. The approach is implemented using numpy, hashlib, and pandas libraries.

Allocation Concealment Mechanism

The randomization sequence is stored in a password-protected file accessible only to the researcher. Allocation is revealed only after the participant completes the baseline PSQI and receives the device (day 0), preventing selection bias. Participants are enrolled and allocated in cohorts of 2 - 8 to maintain seasonal balance across the 15-month recruitment window.

Implementation

The researcher generates an allocation sequence and assigns interventions according to the sequence concealed in sealed, sequentially numbered envelopes, which are opened only at the allocation time point. Participants are informed of group assignment on day 30 (intervention delivery) for the experimental group or at close-out for the control group.

Seasonal Control

To account for seasonal or environmental factors over the 15-month recruitment window, device assignment ensures equal representation from both groups at each time point. An even number of participants receive devices simultaneously, preventing differential seasonal effects between groups.

Blinding (Masking) Level

This is an open-label pilot trial with partial blinding where feasible. Participants become aware of group assignment at different time points (experimental group participants are informed on day 30 at intervention delivery; control group participants are informed on day 60 at close-out). Samsung Galaxy Watch 4 automatically records objective sleep data independent of participant or researcher knowledge, eliminating measurement bias for the primary outcome. The researcher analyzing primary outcome data (Samsung scores) remains blinded to group allocation, and unblinding occurs on day 30 to identify participants for analysis in the experimental group in order to provide personalized recommendations.

Justification for Partial Blinding

Complete participant blinding is impossible for providing personalized and generic advice. The primary outcome (device-recorded sleep score) has objective measurement protection. For the secondary outcome (PSQI), assessor blinding is attempted where feasible. This pragmatic approach balances methodological rigor with pilot study feasibility.

Data Collection Quality Control

The following measures are implemented for user-related issues: (1) noncompliance: participants with missing data for >20% of nights are excluded and replaced; (2) charging issues: backup chargers are provided, and daily battery monitoring is required; (3) unauthorized modifications: settings are locked, and violations lead to replacement; and (4) sync delays: daily commitment for cloud backup within 24 hours. The following measures are implemented for device-related issues: (1) fragmented recording: nights without combined scores are excluded; (2) data mismatch: nights with artifact conflicts are excluded; and (3) sync errors: multiple attempts are made with escalating intervals. For sensor inaccuracies, quality control is performed through sleep diaries, and outliers trigger replacement. Excluded participants are replaced from the recruitment pool to maintain the sample size. The expected night exclusion impact is around 18%, with anticipated replacement ensuring the target sample.

Subjective data collection is performed through PSQI forms completed on paper. Information is manually entered by the researcher into a password-protected spreadsheet with self-verification (data are entered twice by the researcher to ensure accuracy). Data are processed through the manufacturer's algorithms to interpret the findings of the 10 explanatory variables and the target variable (Table 1).

XSL•FO
RenderX

Table . Target and explanatory variables collected by the wearable (indispensable for this study).

Variable	Description	Unit	Data type
Explanatory variable			
sleep_duration	Sum of N1, N2, N3, and REM[a]	Minutes	Integer
movement_awakening	Count of movement during sleep	Number	Integer
sleep_cycle	Count of sleep cycles	Number	Integer
mental_recovery	Score assigned according to the REM phase	Score (0 - 100 scale)	Integer
physical_recovery	Score assigned according to the N3 phase	Score (0 - 100 scale)	Integer
efficiency	real_total_sleep by sleep_duration	Percentage	Integer
Transition_N1	Sum of minutes of sleep in the N1 phase	Minutes	Integer
Light_N2	Sum of minutes of sleep in the N2 phase	Minutes	Integer
Deep_N3	Sum of minutes of sleep in the N3 phase	Minutes	Integer
REM	Sum of minutes of sleep in the REM phase	Minutes	Integer
Target variable			
sleep_score	Score assigned for the night	Score (1 - 100 scale)	Integer

[a]REM: rapid eye movement.

Data Preprocessing Pipeline

Data are downloaded from Samsung Cloud onto the participant's mobile device for transfer to the researcher's computer for analysis. Moreover, individual nights are processed, sleep stages are integrated, and participant IDs are assigned. For quality control, nights with missing or corrupted segments are excluded, fragmented sleep data without combined scores are removed, daytime sleep (outside 7 PM to 11 AM) is excluded, nights with ≥30 minutes of stage/total discrepancies are excluded, and outliers >3 SD from baseline are flagged. For feature engineering, additional features (sleep stage percentage and real TST) are derived from the 10 device variables in ML analysis. Completeness of reporting is ensured by documenting the percentage of excluded nights and the reasons for exclusion.

Data Management

Data entered and collected by Samsung Galaxy Watch 4 are automatically synced to Samsung Cloud (encrypted transmission) upon pairing. The PSQI forms are completed on paper, and the information is manually entered by the researcher into a password-protected spreadsheet with self-verification (data entered twice by the researcher to ensure accuracy). All data stored on the researcher's local computer are encrypted, providing additional privacy protection through local-only storage. Moreover, data are stored in 2 separate local files as a backup measure for protection against file corruption. Access control is set, with only the researcher having access to all study data, including identifiable information linking Participant_ID to personal information. No other personnel have access to the data.

Data security involves (1) physical security (paper PSQI forms are stored in a locked filing cabinet at a secure location), (2) digital security (electronic files are password-protected and encrypted, and are stored on the researcher's local computer only [no remote or cloud storage of identifiable data]), and (3) transmission security (Samsung device data download uses encrypted protocols [HTTPS]).

ML-Specific Data Security Measures

ML model training is conducted on the researcher's local computer with no internet connectivity during sensitive data processing to prevent data exfiltration. All access to ML model files and training data is logged within the researcher's system for accountability. Cross-validation and model testing use separate data partitions, and there is no test data contamination of training sets. Validation results are stored separately from production models on the researcher's local system. No commercial ML platforms (eg, AWS SageMaker and Google AutoML) are used to prevent corporate access to participant data. All processing uses open-source libraries (scikit-learn, XGBoost, and SHAP) on local infrastructure.

Data quality assurance includes data completeness checks (missing nights and sync failures), value range checks for physiologically implausible values, and analysis data validation using scripts for checking distributions, outliers, and duplicates.

XSL•FO
RenderX

A retention and destruction policy is implemented. Identifiable data will be retained for 5 years after publication, and then, the data will be securely destroyed (paper shredding and secure overwriting of electronic files). Deidentified data will be retained indefinitely.

Statistical Methods

Primary Analysis: Single-Covariate Analysis of Covariance

The mean sleep score in the postintervention period (days 31 - 60) is compared between groups using analysis of covariance with a single covariate: mean sleep score in the baseline period (days 1 - 30).

$$\text{Model: Postintervention score} = \beta 0 + \beta 1\,(\text{group}) + \beta 2\,(\text{baseline score}) +$$

The control group is coded as "0," and the experimental group is coded as "1." The primary inference is based on the β_1 coefficient (group difference adjusted for baseline). The effect size is calculated as Cohen *d* with 95% CIs. Statistical significance is assessed at α=.05 (1-tailed per directional hypothesis), but the interpretation emphasizes the effect size magnitude and CI width, given the pilot nature of the study.

We do not use multiple covariates (age, sex beyond stratification, and behavioral factors) to avoid overfitting with the small sample size. Sex stratification in randomization addresses balance concerns without requiring covariate adjustment.

Secondary Analysis: PSQI Changes

The PSQI global score change from baseline (day 0) to follow-up (day 60) is compared between groups using the independent samples *t* test. The correlation between the Samsung score change and PSQI change is calculated using Pearson *r*.

Exploratory Analyses: Hypothesis Generation Only

The exploratory analyses involve the following three aspects:

1. Individual parameter changes: Compare the change in specific sleep variables (duration, deep sleep percentage, and REM percentage) between groups using *t* tests (unadjusted for multiple comparisons, given the exploratory nature).
2. Adherence subgroups: In the experimental group, compare outcomes between high (≥80%) and low (<80%) self-reported adherence using the Mann-Whitney *U* test.
3. Predictive feature patterns: Perform a descriptive analysis of which sleep parameters are most frequently identified as the top predictors across participants.

All exploratory analyses are labeled hypothesis-generating, not confirmatory. No adjustments are made for multiple comparisons, given the pilot scope of the study. Regarding the handling of missing data for the primary outcome (sleep score), participants are excluded if sufficient night data are not available in the baseline and postintervention periods, and action on the secondary outcome (PSQI) is derived from the handling of the primary outcome. This protocol contains the complete statistical analysis plan, and no separate statistical analysis plan exists.

Interpretation Guidance

All findings will be interpreted as preliminary, hypothesis-generating signals requiring confirmation in adequately powered trials. We will report effect sizes, CIs, and descriptive patterns rather than definitive claims of efficacy or lack thereof. The pilot's value lies in feasibility demonstration and effect size estimation for future sample size calculations, not in definitive answers about intervention effectiveness.

Data Monitoring Oversight

The researcher monitors study progress, enrollment rates, data quality, and adverse events throughout the study period. No formal data monitoring committee is established, given the pilot scale of the study and the use of a low-risk intervention (behavioral recommendations only).

Interim Analysis

No formal interim analysis is planned. The researcher conducts data quality monitoring but does not analyze outcomes until study completion to prevent bias.

Stopping Guidelines

The study may be stopped early if (1) enrollment fails to reach the target within 18 months, (2) 50% of enrolled participants withdraw, (3) serious adverse events attributable to study procedures occur, or (4) unexpected device failures affect >30% of participants.

Harms: Adverse Event Tracking

The following events are tracked:

- Study-related events: documented at 3 interaction points (baseline, day 30, and day 60)
- Behavioral or environmental factors: captured in the closure interview
- Withdrawal events: managed according to informed consent procedures
- All adverse events: recorded and reported in the final analysis

Behavioral Factor Assessment

The study closure interview captures key behavioral factors affecting sleep (alcohol consumption, exercise, and stress) that may confound the results. Rather than excluding participants with these factors, we document them to understand their impacts on intervention effectiveness. This approach maintains ecological validity while acknowledging potential confounders.

Instruments

Commercial-Grade Wearables

The study uses Samsung Galaxy Watch 4 [27] for sleep tracking. Samsung Galaxy Watch 4 has been selected following a systematic evaluation of available consumer sleep tracking devices. The selection criteria are as follows: validation evidence (published PSG comparison studies), parameter comprehensiveness (10 sleep variables vs 3 - 5 in most competitors), data accessibility for research use, cost feasibility (US $200 - 250 per unit [within budget]), operationalization of sleep quality through a composite sleep score (scale 1-100) that

XSL•FO
RenderX

algorithmically integrates multiple dimensions, and user compliance factors (battery life >24 hours). While alternatives exist (Fitbit: limited architecture data; Apple Watch: iOS requirement; Garmin: higher cost; WHOOP: subscription model), Samsung Galaxy Watch 4 provides an optimal balance of scientific validity, data richness, and practical feasibility for this pilot study [29,30,36-43,71,105-107].

This sleep monitoring system is noninvasive, involves ambulatory or home-based deployment, is contact-based (wrist-worn), has comprehensive architecture analysis parameters, involves real-time analysis, and is a consumer-grade device with clinical validation studies. The sensors include an accelerometer, a gyroscope, a photoplethysmography sensor, a bioelectrical impedance sensor, and an ambient light sensor [27,121]. This positioning offers ecological validity while acknowledging trade-offs in precision compared with laboratory approaches.

Pairing Devices

Mobile phones (1 per wearable) are used for data synchronization.

PSQI Questionnaire

The PSQI is used for evaluation. It is a validated subjective sleep quality assessment [28].

Computer

The researcher's computer is used for data storage, processing, and ML analysis.

Study Timeline

The overall experimentation period is 15 months (all volunteers). The start date is October 4, 2024, and the maximum end date is December 31, 2025. The experiment duration for each volunteer is 60 days.

Artificial Intelligence Accountability and Algorithmic Governance

This study uses ML algorithms to analyze participant sleep data and generate personalized recommendations. The governance framework is presented below.

Algorithmic Transparency Obligations

Participant Information

The participant information aspects are as follows:

- Informed consent explicitly states the use of ML for sleep data.
- Participants are informed that recommendations represent statistical pattern recognition, not clinical judgment.
- Models identify correlations, not causation.
- Recommendations are behavioral suggestions, not medical advice.
- Participants maintain full autonomy to accept, modify, or ignore recommendations.

Explainability

The explainability aspects are as follows:

- ML recommendations are accompanied by plain-language explanations.
- SHAP visualizations show which sleep parameters influence the individual model.
- Example: "Your model identified REM sleep and sleep timing consistency as your two strongest predictors. When you had ≥90 minutes REM and went to bed within 30 minutes of your average time, your scores were typically 8 - 12 points higher."

Human Oversight

The human oversight aspects are as follows:

- ML models suggest recommendations, but the principal researcher reviews all outputs before delivery.
- The human-in-the-loop approach ensures recommendations are safe, coherent, and appropriate.
- Models do not make automated health decisions.

Participant Rights Concerning Algorithmic Outputs

Participant rights are as follows:

- Right to explanation: Participants may request a detailed explanation of how personalized recommendations are generated.
- Right to question: Participants may challenge recommendations they believe are inaccurate or inappropriate. The researcher is available for discussion and can modify recommendations based on participant feedback.
- Right to decline: Participants may decline to implement some or all ML-derived recommendations without penalty or withdrawal from the study. Nonadherence is tracked but does not affect participant standing.
- Right to human review: Participants may request that the principal researcher personally review their data and provide alternative recommendations if dissatisfied with ML outputs.
- Access to model information: Participants may request information about ML methodology (algorithms used, training process, and validation approach) in plain language.

Data Retention and Security for ML Processes

The data retention and security aspects are as follows:

- Raw sleep data: Data from 1920 nights are stored encrypted on the researcher's local computer with password protection and backup. The data are retained for 5 years after publication and are then securely destroyed.
- ML model files: Individual participant models (random forest and XGBoost) are stored on the researcher's local computer separately from identifiable data. Models contain learned patterns but no raw data. The information is retained indefinitely for reproducibility verification. The information can be deleted upon participant request during the identifiable data retention period.
- SHAP values and feature importance rankings: The data are stored as CSV files on the researcher's local computer, linking Participant_ID to rank predictive features. These data have the same retention schedule as the raw data (5 years after publication).

XSL•FO
RenderX

- Transmission security: No ML processing occurs on participant devices or cloud services. All analyses are conducted on the researcher's local computer. No data are transmitted to third parties or commercial ML platforms.

Accountability Framework

The accountability framework involves the following aspects:

- Adverse events from artificial intelligence (AI): If ML-derived recommendations cause participant harm (eg, a recommendation to reduce sleep duration leads to excessive daytime sleepiness affecting work safety), the event is documented as an adverse event. The researcher notifies the ethics committee, and an investigation is conducted to determine if a systematic model error exists.
- Replication and audit: The ML code, trained models, and analysis scripts are made available to the ethics committee and peer reviewers upon request.
- Model performance monitoring: During the study, the principal researcher monitors whether the experimental group shows unexpected negative outcomes (eg, worsening sleep scores), triggering a review of ML recommendation patterns for systematic errors.

Compliance With AI Governance Frameworks

This study follows the principles from the World Health Organization (WHO) guidance on ethics and governance of AI for health (transparency, accountability, and human oversight) and the International Association of Scientific, Technical & Medical Publishers recommendations for AI use in academic manuscript preparation [122].

Ethical Considerations

Research Ethics Approval

This research has been reviewed and approved by the Institutional Committee on Research Ethics of Tecnológico de Monterrey (ethics approval number: CA-EIC-2408-06; approval date: October 2024; review type: full committee review). The study complies with the principles of the Declaration of Helsinki for medical research involving human participants and the international ethical guidelines for health research. There is no compensation for participation in the study.

Protocol Amendments

Any modifications to the eligibility criteria, interventions, outcomes, or ethical procedures will be submitted to the ethics committee for review and approval prior to implementation. Amendments will be documented in protocol version history with the rationale for changes.

Informed Consent Collection

The principal researcher has completed ethical training to obtain informed consent from potential trial participants.

Additional Consent Provisions for the Collection and Use of Participant Data and Biological Specimens in Ancillary Studies

It is not applicable to consider additional consent provisions for the collection and use of participant data and biological specimens in ancillary studies since there are no additional studies.

Confidentiality: Data Protection

Personal data are not shared. The data are only used to allocate participants to the study groups and to ensure consistent assignment of the proper Participant_ID to keep the process accurate and deidentified. The participants can access their personal data or request data correction. Data are protected before, during, and after the trial.

Compensation for Research-Related Injury

In the unlikely event of research-related injury, immediate medical care coordination will be provided by the researcher, participants can use existing health insurance for treatment, no additional compensation will be provided, and the incident will be documented and reported to the ethics committee. The anticipated risk is minimal (behavioral intervention, noninvasive device use, skin irritation from the device, sleep tracking anxiety, and time commitment).

Ancillary and Posttrial Care

Incidental Findings

Participants with incidental findings during the study may be withdrawn and referred for severe sleep deterioration (sleep score <30 for >1 consecutive week). Participants may be withdrawn from the study and referred to sleep medicine services at their medical institutions. Contact information may be provided for sleep specialists in Mexico City. There are no study-related costs for referral (participants use existing health insurance).

Posttrial Care

There is no posttrial provision of interventions. Personalized recommendations are study-specific research interventions, not ongoing clinical care interventions. After study completion, participants may continue using strategies that helped them. There is no ongoing monitoring or intervention delivery. Participants are encouraged to consult health care providers for persistent sleep concerns. No direct medical benefits are guaranteed.

Access Control and Protection Layers for Potential and Excluded Participants

Data collected from potential participants who do not fill out the informed consent form (Multimedia Appendix 3) or the commitment letter (Multimedia Appendix 4) and those who voluntarily withdraw or need to be excluded from the study are immediately and securely deleted.

Access Control and Protection Layers for Enrolled Participants

Only the principal researcher has access to all study data, including identifiable information linking Participant_ID to personal information (mapping). The only risk is the mapping between the participant and the alias, which is kept in physical form at a secure physical location. Accessing it through a digital hacking attack is not possible. Participant data are stored securely and can be accessed only by authorized research

XSL•FO
RenderX

personnel. Data will be retained for 5 years after publication, while the deidentified dataset will be retained indefinitely.

Data security involves (1) physical security (paper PSQI forms are stored in a locked filing cabinet at a secure location), (2) digital security (electronic files are password-protected and encrypted, and are stored on the researcher's local computer only [no remote or cloud storage of identifiable data]), and (3) transmission security (Samsung device data download uses encrypted protocols [HTTPS]).

Anonymization Process (Deidentification)

Anonymization involves assigning an alias to the personal data collected from each participant at different process steps. Email or Samsung accounts do not use personal data. These accounts are only used for data collection, as the approach is already anonymous by design.

An alias is used at each phase of the process. The smartwatch user setup is generic, and the smartwatch collects objective data without the participant's personally identifiable data. The same device and setup are used throughout the study. The pairing device setup, like the user setup, is generic, and thus, the mobile phone collects objective data without the participant's personally identifiable data. PSQI questionnaire data collection uses the Participant_ID alias for subjective data collection without the participant's personally identifiable data. In the last phase, the data are merged into a single dataset, which is one of the research outputs.

ML and AI Elements

Participants receive an explanation that the algorithms will analyze individual sleep patterns; a clarification that recommendations are statistically derived suggestions, not medical diagnoses; a statement that they retain full autonomy to follow or ignore recommendations; a description of SHAP explainability (plain-language explanations of personalized insights); and an assurance of human oversight (the researcher reviews all ML outputs before delivery).

Enhanced AI/ML-Specific Consent Elements

Given the ML component, participants are explicitly informed about how ML models work (pattern recognition in their individual data) and what SHAP analysis means (identifying which factors predict their sleep scores). They are also informed that recommendations are correlational, not causal (eg, "REM correlates with higher scores" not "REM causes better sleep"), and that they have a right to question or decline any recommendation. Furthermore, they are told that the researcher reviews the recommendations before delivery (human-in-the-loop approach) and that no automated health decisions are made.

Results

This protocol describes the planned analytical approach. The actual results, including group comparisons, effect sizes, statistical significance levels, and interpretation of findings, will be reported in a subsequent manuscript following study completion. The study is funded by Instituto Tecnológico y de Estudios Superiores de Monterrey (publishing fee) and *Secretaría de Ciencia, Humanidades, Tecnología e Innovación*, which was formerly called *Consejo Nacional de Ciencia y Tecnología* (wearable and pairing device equipment funding). The funding reference number is REF4661568-251121 (November 24, 2025). The trial has not yet been registered, but we intend to register it on ClinicalTrials.gov.

The study is currently in progress. Recruitment started in August 2024 and ended in July 2025. Data collection is expected to be completed by December 2025. The study will (1) compare 960 nights from the control group with 960 nights from the experimental group to explore whether ML interventions can improve sleep scores using wearable technology, (2) generate a dataset from objective data for iterative model training and analysis, (3) correlate objective and subjective sleep quality metrics, and (4) establish whether a feasible framework for proactive sleep quality approaches can be developed. The final analyzed results incorporating all quality control procedures, statistical analyses, and exploratory investigations will be available by March 2026, and the findings will be submitted for peer-reviewed publication within 6 months of study completion.

Upon completion of data collection by December 2025, the study will generate a comprehensive dataset from 32 participants (16 per arm) wearing Samsung Galaxy Watch 4 devices for 60 consecutive nights each, yielding approximately 1920 nights of sleep data. The final dataset will include 18 variables (Table 2).

XSL•FO
RenderX

Table . Final variables included in this study.

Variable	Description	Unit	Data type
Explanatory variable			
Participant_ID	Unique number assigned to each participant	Number	Integer
Night_Number	Night count	Number	Integer
sleep_duration	Sum of N1, N2, N3, and REM[a]	Minutes	Integer
movement_awakening	Count of movement during sleep	Number	Integer
sleep_cycle	Count of sleep cycles	Number	Integer
mental_recovery	Score assigned according to the REM phase	Score (0 - 100 scale)	Integer
physical_recovery	Score assigned according to the N3 phase	Score (0 - 100 scale)	Integer
efficiency	real_total_sleep by sleep_duration	Percentage	Integer
Transition_N1	Sum of minutes of sleep in the N1 phase	Minutes	Integer
Light_N2	Sum of minutes of sleep in the N2 phase	Minutes	Integer
Deep_N3	Sum of minutes of sleep in the N3 phase	Minutes	Integer
REM	Sum of minutes of sleep in the REM phase	Minutes	Integer
real_total_sleep[b]	Sum of N2, N3, and REM	Minutes	Integer
Transition_pct	Percentage of sleep in the N1 phase	Percentage	Float
Light_pct	Percentage of sleep in the N2 phase	Percentage	Float
Deep_pct	Percentage of sleep in the N3 phase	Percentage	Float
REM_pct	Percentage of sleep in the REM phase	Percentage	Float
Target variable			
sleep_score	Score assigned for the night	Score (1 - 100 scale)	Integer

[a]REM: rapid eye movement.

[b]The “real_total_sleep” variable excludes the N1 transition stage from total sleep duration, as this brief transitional period between wakefulness and consolidated sleep is typically excluded from restorative sleep calculations in sleep research.

The dataset will enable comparisons between the control group (960 nights; participants receiving generic sleep hygiene education on day 30) and the experimental group (960 nights; participants receiving personalized ML-derived recommendations on day 30).

Each night of sleep data will be automatically collected through sensors by Samsung Galaxy Watch 4 and processed through the manufacturer’s proprietary algorithms to generate objective sleep metrics. Data will be organized by participant identifier and temporal sequence to enable both group-level comparisons and individual-level ML model development.

Data completeness evaluation based on pilot testing and literature review suggests that approximately 18% of individual nights may be excluded due to the quality control criteria, with participant replacement ensuring that the final dataset achieves the target 1920 nights across all study participants.

Discussion

Proprietary Algorithm Constraints

The Samsung Galaxy Watch 4 sleep score is a proprietary composite metric with undisclosed algorithmic weightings. This limits reproducibility of the exact score calculation and precludes independent validation against PSG for the composite

XSL•FO
RenderX

metric specifically (though component parameters have validation evidence [36-43]). We acknowledge these constraints openly and position our methodology as working within the pragmatic realities of consumer wearable research.

Our framework demonstrates that effective personalization can be generated despite algorithmic opacity through pattern recognition that transcends the need to understand underlying computations. If reducing prebedtime screen time correlates with a participant's score improvements across 30 nights, that pattern remains actionable regardless of Samsung's computational details. The methodology's value lies in its transferability. Future studies can apply identical approaches to Fitbit, Apple Watch, Garmin, or other platforms, testing whether personalization generalizes across manufacturers.

Potential Outcomes and Interpretation Framework

This pilot study is explicitly designed for feasibility assessment and preliminary effect size estimation, not definitive efficacy testing. As noted in the Methods section, this pilot has limited power for moderate effects. Pilot studies generate hypotheses and refine the methodology, and definitive answers require subsequent adequately powered trials.

Study Limitations

This pilot study has several limitations that contextualize interpretation. The small sample (n=32) provides insufficient power for moderate effects and limits subgroup analyses. The 30-day intervention period may be inadequate for sustained behavior change. The single-device focus prevents cross-platform validation. Geographic and cultural limitations (Mexico City context with unique altitude, pollution, and lifestyle factors) constrain generalizability. The exclusion of sleep disorder populations limits applicability to healthy adults seeking optimization rather than clinical treatment.

The proprietary algorithm dependency limits reproducibility, though we position this as a pragmatic constraint our methodology addresses rather than an insurmountable barrier. Consumer-grade sensor accuracy has inherent limitations compared to PSG, potentially obscuring true intervention effects with small samples.

We do not collect comprehensive outcome data assessing daytime functioning, cognitive performance, mood, physical health markers, or quality of life. The focus remains on sleep score optimization without demonstrating whether improvements translate to meaningful real-world benefits. Future studies should incorporate broader outcomes to establish clinical relevance beyond device metrics.

Study Strengths

Despite these limitations, this pilot study has several strengths. The manufacturer-independent framework provides a transferable methodology applicable across platforms, potentially driving the field toward less proprietary and more user-centered approaches. The dual validation combining objective and subjective measures addresses both physiological and patient-reported domains. The explicit pilot design appropriately positions findings as preliminary and hypothesis-generating rather than making premature efficacy claims. The comprehensive data governance framework (local-only storage, multilayer anonymization, and explicit AI accountability measures) provides a model for ethical wearable research. The transparent reporting approach with public data sharing enables replication and meta-analytic integration regardless of the findings.

Implications for Future Research

If this pilot study demonstrates feasibility and promising effect sizes, subsequent research should prioritize the following: (1) an adequately powered definitive RCT (n=350 total) with extended intervention periods (3 - 6 months); (2) concurrent PSG validation in subsamples to establish measurement precision; (3) multidevice validation testing framework generalizability across manufacturers; (4) comprehensive outcome batteries, including daytime functioning, cognitive performance, and quality of life; (5) cost-effectiveness analyses comparing personalized and generic approaches; and (6) implementation science examining real-world deployment barriers and facilitators.

The pilot dataset will enable iterative ML model refinement, identification of optimal baseline data collection periods (Is 30 days sufficient or should future studies collect data for 60 - 90 days?), and exploration of which sleep parameters show the most reliable personalization effects. These methodological refinements will strengthen subsequent trials.

Conclusions

This pilot study establishes the feasibility and preliminary evidence base necessary for future definitive trials testing whether ML personalization involving consumer wearable data can improve sleep outcomes. The manufacturer-independent framework, dual validation approach (objective device metrics plus subjective PSQI), and explicit pilot design appropriately position this study as hypothesis-generating research that could inform the next generation of personalized digital health interventions for sleep quality optimization in healthy populations.

The proposed approach tests a specific mechanistic hypothesis: precision targeting of high-impact parameters identified through individual ML models may enhance motivation through personalized predictions while reducing cognitive burden compared with generic recommendation lists. Rather than providing universal sleep hygiene advice, this methodology will generate interventions tailored to each participant's unique sleep architecture patterns and predictive factors.

Building upon foundational work establishing wearable feasibility for sleep monitoring [36,71,105-107], this study's planned contribution to the field involves testing whether ML-derived personalization improves outcomes compared with generic sleep hygiene. If the methodology proves feasible and preliminary effect sizes suggest a benefit, the resulting dataset of 1920 nights will enable iterative model refinement and provide evidence for scaling personalized sleep interventions to larger populations.

Several important considerations frame the interpretation of the anticipated findings. The study acknowledges the inherent

XSL•FO
RenderX

limitations of consumer wearable research, particularly reliance on Samsung's proprietary sleep score algorithm. While this limits reproducibility of exact score computation and precludes independent PSG validation of the composite metric, the personalization framework itself is deliberately manufacturer-agnostic. Additional methodological constraints include the single-device focus, the limited geographic and cultural generalizability (Mexico City context with unique altitude, pollution, and lifestyle factors), potential measurement errors inherent to consumer-grade sensors, and the 30-day intervention period, which may be insufficient for sustained behavior change in entrenched sleep habits. The study excludes individuals with diagnosed sleep disorders, limiting the findings to healthy populations seeking sleep quality optimization rather than clinical treatment.

If validated through this pilot study, the methodology could direct sleep optimization from universal protocols toward individualized behavioral targeting, offering a scalable approach to sleep health improvement.

However, if the findings reveal no difference between personalized and generic interventions, there would be valuable evidence that (1) generic interventions may be equally effective and more cost-efficient for healthy populations, (2) the ML personalization approach requires refinement before clinical utility can be established, or (3) consumer wearables may not yet provide sufficient data quality for meaningful pattern detection at the individual level. Both positive and null findings will advance scientific understanding and inform future research directions.

The results will be available by March 2026 and will be disseminated through a peer-reviewed publication regardless of the findings (positive, null, or negative). A deidentified dataset will be part of the study output. This transparency supports scientific rigor and allows the broader research community to build upon the methodological framework regardless of this pilot's specific outcomes.

Acknowledgments

The principal researcher (RQR) would like to acknowledge the scholarships provided by the *Instituto Tecnológico y de Estudios Superiores de Monterrey* for PhD research framework education and the British Council for a research training workshop. We acknowledge the use of generative artificial intelligence (AI) in the preparation of this academic manuscript in the following three areas according to Scientific, Technical & Medical Publishers [122]: (1) refining, correcting, editing, and formatting the manuscript to improve clarity of language (spell, grammar, and tone checking to refine the text written by humans); (2) refining and formatting of data reported in the manuscript (improve clarity or readability using AI tools); and (3) assisting with gathering references (find and identify articles related to the research reported in the manuscript). Generative AI was not used for study design, data collection planning, data analysis, or scientific interpretation. All AI-generated content was reviewed and validated by the authors.

Funding

The *Secretaría de Ciencia, Humanidades, Tecnología e Innovación*, which was formerly called *Consejo Nacional de Ciencia y Tecnología*, funds smartwatch and pairing device purchases. The *Instituto Tecnológico y de Estudios Superiores de Monterrey* funds the publication fee of this article. The funders are not involved in the study design, data collection, analysis, interpretation, or writing of the manuscript.

Data Availability

The data collected are not yet publicly available, as data collection is in progress. The dataset generated during this study will be available upon reasonable request.

Authors' Contributions

Conceptualization: RQR, LAT
Funding acquisition: RQR, LAT
Investigation: RQR
Methodology: RQR, LAT
Project administration: LAT
Resources: RQR
Software: RQR
Supervision: RQR, LAT
Validation: RQR, LAT
Writing – original draft: RQR
Writing – review & editing: RQR, LAT

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Predefined library.
[PDF File, 118 KB - resprot_v15i1e76415_app1.pdf]

Multimedia Appendix 2
Power analysis table showing power for different effect sizes.
[PDF File, 38 KB - resprot_v15i1e76415_app2.pdf]

Multimedia Appendix 3
Informed consent form.
[PDF File, 39 KB - resprot_v15i1e76415_app3.pdf]

Multimedia Appendix 4
Commitment letter.
[PDF File, 33 KB - resprot_v15i1e76415_app4.pdf]

Checklist 1
SPIRIT 2025 checklist.
[PDF File, 104 KB - resprot_v15i1e76415_app5.pdf]

References

1. Philips global sleep survey shows we want better sleep, but only if it comes easily. Philips. 2019. URL: https://www.philips.com/a-w/about/news/archive/standard/news/press/2019/20190307-philips-global-sleep-survey-shows-we-want-better-sleep-but-only-if-it-comes-easily.html [accessed 2025-12-09]
2. Philips survey reveals COVID-19's negative impact on sleep quality and CPAP use. Philips. 2021. URL: https://www.philips.com/a-w/about/news/archive/standard/news/press/2021/20210310-philips-survey-reveals-covid-19-s-negative-impact-on-sleep-quality-and-cpap-use.html [accessed 2025-12-09]
3. National Sleep Foundation's Sleep in America® poll. : National Sleep Foundation; 2025 URL: https://www.thensf.org/wp-content/uploads/2025/03/NSF_SIA_2025-Report_final.pdf [accessed 2025-12-09]
4. Arrona-Palacios A, Gradisar M. Self-reported sleep duration, sleep quality and sleep problems in Mexicans adults: results of the 2016 Mexican National Halfway Health and Nutrition Survey. Sleep Health 2021 Apr;7(2):246-253. [doi: 10.1016/j.sleh.2020.08.006] [Medline: 33097465]
5. Encuesta Nacional de Salud y Nutrición - Medio Camino 2016 [Article in Spanish]. Instituto Nacional de Salud Pública (INSP). 2016. URL: https://ensanut.insp.mx/encuestas/ensanut2016/index.php [accessed 2025-12-09]
6. Cohen S, Doyle WJ, Alper CM, Janicki-Deverts D, Turner RB. Sleep habits and susceptibility to the common cold. Arch Intern Med 2009 Jan 12;169(1):62-67. [doi: 10.1001/archinternmed.2008.505] [Medline: 19139325]
7. Gómez-González B, Domínguez-Salazar E, Hurtado-Alvarado G, et al. Role of sleep in the regulation of the immune system and the pituitary hormones. Ann N Y Acad Sci 2012 Jul;1261:97-106. [doi: 10.1111/j.1749-6632.2012.06616.x] [Medline: 22823399]
8. Denisova K, Motomura Y, Song C. Sleep on it: exploring the psychology of sleep amidst contemporary challenges. Sci Rep 2024 Dec 16;14(1):30501. [doi: 10.1038/s41598-024-82330-w] [Medline: 39681593]
9. Diekelmann S, Born J. The memory function of sleep. Nat Rev Neurosci 2010 Feb;11(2):114-126. [doi: 10.1038/nrn2762] [Medline: 20046194]
10. Feld GB, Born J. Sculpting memory during sleep: concurrent consolidation and forgetting. Curr Opin Neurobiol 2017 Jun;44:20-27. [doi: 10.1016/j.conb.2017.02.012] [Medline: 28278432]
11. Stickgold R. Sleep-dependent memory consolidation. Nature New Biol 2005 Oct 27;437(7063):1272-1278. [doi: 10.1038/nature04286] [Medline: 16251952]
12. Kwok CS, Kontopantelis E, Kuligowski G, et al. Self-reported sleep duration and quality and cardiovascular disease and mortality: a dose-response meta-analysis. J Am Heart Assoc 2018 Aug 7;7(15):e008552. [doi: 10.1161/JAHA.118.008552] [Medline: 30371228]
13. Baek Y, Jeong K, Lee S. Association of sleep timing, sleep duration, and sleep latency with metabolic syndrome in middle-aged adults in Korea: a cross-sectional and longitudinal study. Sleep Health 2025 Feb;11(1):73-79. [doi: 10.1016/j.sleh.2024.06.002] [Medline: 39174451]
14. Wang Q, Xi B, Liu M, Zhang Y, Fu M. Short sleep duration is associated with hypertension risk among adults: a systematic review and meta-analysis. Hypertens Res 2012 Oct;35(10):1012-1018. [doi: 10.1038/hr.2012.91] [Medline: 22763475]
15. Jemere T, Mossie A, Berhanu H, Yeshaw Y. Poor sleep quality and its predictors among type 2 diabetes mellitus patients attending Jimma University Medical Center, Jimma, Ethiopia. BMC Res Notes 2019 Aug 6;12(1):488. [doi: 10.1186/s13104-019-4531-6] [Medline: 31387638]

XSL•FO
RenderX

16. De Bernardi Rodrigues AM, da Silva CDC, Vasques ACJ, et al. Association of sleep deprivation with reduction in insulin sensitivity as assessed by the hyperglycemic clamp technique in adolescents. JAMA Pediatr 2016 May 1;170(5):487-494. [doi: 10.1001/jamapediatrics.2015.4365] [Medline: 26999405]
17. What are sleep deprivation and deficiency? National Heart, Lung, and Blood Institute (NHLBI). URL: https://www.nhlbi.nih.gov/health/sleep-deprivation [accessed 2025-12-09]
18. Li Y, Zhang X, Winkelman JW, et al. Association between insomnia symptoms and mortality. Circulation 2014 Feb 18;129(7):737-746. [doi: 10.1161/CIRCULATIONAHA.113.004500]
19. Parthasarathy S, Vasquez MM, Halonen M, et al. Persistent insomnia is associated with mortality risk. Am J Med 2015 Mar;128(3):268-275. [doi: 10.1016/j.amjmed.2014.10.015] [Medline: 25447616]
20. Vgontzas AN, Liao D, Pejovic S, et al. Insomnia with short sleep duration and mortality: the Penn State cohort. Sleep 2010 Sep;33(9):1159-1164. [doi: 10.1093/sleep/33.9.1159] [Medline: 20857861]
21. World urbanization prospects 2025. United Nations. 2025. URL: https://population.un.org/wup/countryprofiles?country=Mexico&mainCategory=Cities&subCategory=Urbanization%20-%20International%20definition [accessed 2025-12-16]
22. Censo de Población y Vivienda 2020 [Article in Spanish]. Instituto Nacional de Estadística y Geografía. 2020. URL: https://www.inegi.org.mx/temas/estructura/ [accessed 2025-12-09]
23. Jiménez-Genchi A, Caraveo-Anduaga J. Crude and adjusted prevalence of sleep complaints in Mexico City. Sleep Sci 2017;10(3):113-121. [doi: 10.5935/1984-0063.20170020] [Medline: 29410740]
24. Dokkedal-Silva V, Fernandes GL, Tufik S, Andersen ML. The links between commuting time and sleep quality: a trend in modern urban centers. J Clin Sleep Med 2022 Dec 1;18(12):2875-2876. [doi: 10.5664/jcsm.10248] [Medline: 35975557]
25. Sánchez-Cárdenas AG, Navarro-Gerrard C, Nellen-Hummel H, Halabe-Cherem J. Insomnia. A severe health care problem. Rev Med Inst Mex Seguro Soc 2016;54(6):760-769. [Medline: 27819787]
26. Blanco M, Kriber N, Cardinali DP. A survey of sleeping difficulties in an urban Latin American population. Rev Neurol 2004;39(2):115-119. [Medline: 15264159]
27. Smartwatch Galaxy Watch4 color Negro Bluetooth 40 mm. Samsung. URL: https://www.samsung.com/mx/watches/galaxy-watch/galaxy-watch4-black-bluetooth-sm-r860nzkamxo/ [accessed 2025-12-09]
28. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
29. Use Samsung Health sleep monitoring on your Galaxy Watch. Samsung. URL: https://www.samsung.com/us/support/answer/ANS10003657/ [accessed 2025-12-09]
30. Use Samsung Health sleep monitoring on your Samsung Galaxy Watch and Fit. Samsung. URL: https://www.samsung.com/latin_en/support/mobile-devices/use-samsung-health-sleep-monitoring-on-your-samsung-galaxy-watch-and-fit/ [accessed 2025-12-09]
31. Birrer V, Elgendi M, Lambercy O, Menon C. Evaluating reliability in wearable devices for sleep staging. NPJ Digit Med 2024 Mar 18;7(1):74. [doi: 10.1038/s41746-024-01016-9] [Medline: 38499793]
32. Zhai B, Elder GJ, Godfrey A. Challenges and opportunities of deep learning for wearable-based objective sleep assessment. NPJ Digit Med 2024 Apr 4;7(1):85. [doi: 10.1038/s41746-024-01086-9] [Medline: 38575794]
33. Boe AJ, McGee Koch LL, O'Brien MK, et al. Automating sleep stage classification using wireless, wearable sensors. NPJ Digit Med 2019;2(1):131. [doi: 10.1038/s41746-019-0210-1] [Medline: 31886412]
34. Aziz S, A M Ali A, Aslam H, et al. Wearable artificial intelligence for sleep disorders: scoping review. J Med Internet Res 2025 May 6;27:e65272. [doi: 10.2196/65272] [Medline: 40327852]
35. Have we been sleeping well? Samsung answers the age-old question with the global sleep health study. Samsung. 2023. URL: https://news.samsung.com/global/have-we-been-sleeping-well-samsung-answers-the-age-old-question-with-the-global-sleep-health-study [accessed 2025-12-09]
36. Gaiduk M, Seepold R, Martínez Madrid N, Ortega JA. Assessing the feasibility of replacing subjective questionnaire-based sleep measurement with an objective approach using a smartwatch. Sensors (Basel) 2023 Jul 4;23(13):6145. [doi: 10.3390/s23136145] [Medline: 37447992]
37. Sarhaddi F, Kazemi K, Azimi I, et al. A comprehensive accuracy assessment of Samsung smartwatch heart rate and heart rate variability. PLOS ONE 2022;17(12):e0268361. [doi: 10.1371/journal.pone.0268361] [Medline: 36480505]
38. Davoudi A, Wanigatunga AA, Kheirkhahan M, et al. Accuracy of Samsung Gear S smartwatch for activity recognition: validation study. JMIR Mhealth Uhealth 2019 Feb 6;7(2):e11270. [doi: 10.2196/11270] [Medline: 30724739]
39. Theurl F, Schreinlechner M, Sappler N, et al. Smartwatch-derived heart rate variability: a head-to-head comparison with the gold standard in cardiovascular disease. Eur Heart J Digit Health 2023 May;4(3):155-164. [doi: 10.1093/ehjdh/ztad022] [Medline: 37265873]
40. Nissen M, Slim S, Jäger K, et al. Heart rate measurement accuracy of Fitbit Charge 4 and Samsung Galaxy Watch Active2: device evaluation study. JMIR Form Res 2022 Mar 1;6(3):e33635. [doi: 10.2196/33635] [Medline: 35230250]

XSL•FO
RenderX

41. Bennett JP, Liu YE, Kelly NN, et al. Next-generation smart watches to estimate whole-body composition using bioimpedance analysis: accuracy and precision in a diverse, multiethnic sample. Am J Clin Nutr 2022;116(5):1418-1429. [doi: 10.1093/ajcn/nqac200] [Medline: 35883219]
42. Jung H, Kim D, Choi J, Joo EY. Validating a consumer smartwatch for nocturnal respiratory rate measurements in sleep monitoring. Sensors (Basel) 2023 Sep 19;23(18):7976. [doi: 10.3390/s23187976] [Medline: 37766031]
43. Kim D, Joo EY, Choi SJ. Validation of the Samsung smartwatch for sleep–wake determination and sleep stage estimation. J Sleep Med 2023 Apr 1;20(1):28-34. [doi: 10.13078/jsm.230004]
44. Weitzman ED, Czeisler CA, Zimmerman JC, Ronda JM, Knauer RS. Chronobiological disorders: analytic and therapeutic techniques. In: Guilleminault C, editor. Sleeping and Waking Disorders: Indications and Techniques: Addison-Wesley; 1982:297-329.
45. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991 Dec;14(6):540-545. [doi: 10.1093/sleep/14.6.540] [Medline: 1798888]
46. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194. [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
47. Knutson KL, Phelan J, Paskow MJ, et al. The National Sleep Foundation's sleep health index. Sleep Health 2017 Aug;3(4):234-240. [doi: 10.1016/j.sleh.2017.05.011] [Medline: 28709508]
48. Ibáñez V, Silva J, Cauli O. A survey on sleep assessment methods. PeerJ 2018;6:e4849. [doi: 10.7717/peerj.4849] [Medline: 29844990]
49. Brady S, Dunne LE, Lynch A, Smyth B, Diamond D. Wearable sensors? What is there to sense? Stud Health Technol Inform 2005;117:80-88. [Medline: 16282656]
50. Zheng YL, Ding XR, Poon CCY, et al. Unobtrusive sensing and wearable devices for health informatics. IEEE Trans Biomed Eng 2014 May;61(5):1538-1554. [doi: 10.1109/TBME.2014.2309951] [Medline: 24759283]
51. Lawson S, Jamison-Powell S, Garbett A, et al. Validating a mobile phone application for the everyday, unobtrusive, objective measurement of sleep. Presented at: CHI '13: CHI Conference on Human Factors in Computing Systems; Apr 27 to May 2, 2013. [doi: 10.1145/2470654.2481345]
52. Klier K, Wagner M. Agreement of sleep measures-a comparison between a sleep diary and three consumer wearable devices. Sensors (Basel) 2022 Aug 18;22(16):6189. [doi: 10.3390/s22166189] [Medline: 36015949]
53. Robertson B, Marshall B, Carno MA. Polysomnography for the Sleep Technologist: Instrumentation, Monitoring, and Related Procedures: Elsevier Health Sciences; 2013.
54. What to expect at your sleep study. AASM Sleep Education. URL: https://sleepeducation.org/what-to-expect-at-your-sleep-study/ [accessed 2025-12-09]
55. Schutte-Rodin S, Deak MC, Khosla S, et al. Evaluating consumer and clinical sleep technologies: an American Academy of Sleep Medicine update. J Clin Sleep Med 2021 Nov 1;17(11):2275-2282. [doi: 10.5664/jcsm.9580] [Medline: 34314344]
56. Lee J, Kim D, Ryoo HY, Shin BS. Sustainable wearables: wearable technology for enhancing the quality of human life. Sustainability 2016;8(5):466. [doi: 10.3390/su8050466]
57. Global smart sleep tracking products market analysis 2019-2028. : Research Nester Pvt Ltd; 2022 URL: https://www.researchandmarkets.com/reports/5624587/global-smart-sleep-tracking-products-market [accessed 2025-12-09]
58. Sarkez-Knudsen M, Ballegaard M, Piilgaard H, et al. Exploring SLEEPINESS through home monitoring with ultra-long-term subcutaneous EEG and ecological momentary assessment in sleepy treatment naïve obstructive sleep apnea patients starting CPAP treatment—a study protocol article. Front Sleep 2025;3. [doi: 10.3389/frsle.2024.1496923]
59. Djurhuus BD, Viana PF, Ahrens E, et al. Minimally invasive surgery for placement of a subcutaneous EEG implant. Front Surg 2023;10:1304343. [doi: 10.3389/fsurg.2023.1304343] [Medline: 38026479]
60. Crivello A, Barsocchi P, Girolami M, Palumbo F. The meaning of sleep quality: a survey of available technologies. IEEE Access 2019;7:167374-167390. [doi: 10.1109/ACCESS.2019.2953835]
61. Wearable technology market (2025 - 2030). : Grand View Research; 2025 URL: https://www.grandviewresearch.com/industry-analysis/wearable-technology-market [accessed 2025-12-09]
62. Van de Water ATM, Holmes A, Hurley DA. Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography--a systematic review. J Sleep Res 2011 Mar;20(1 Pt 2):183-200. [doi: 10.1111/j.1365-2869.2009.00814.x] [Medline: 20374444]
63. Lee XK, Chee N, Ong JL, et al. Validation of a consumer sleep wearable device with actigraphy and polysomnography in adolescents across sleep opportunity manipulations. J Clin Sleep Med 2019 Sep 15;15(9):1337-1346. [doi: 10.5664/jcsm.7932] [Medline: 31538605]
64. Menghini L, Alschuler V, Claudatos S, et al. Accuracy of a commercial wearable in detecting sleep stages compared to polysomnography in adults: considering sleep classification methods and effects of evening alcohol consumption. Sleep 2020 May 27;43(Supplement_1):A456-A457. [doi: 10.1093/sleep/zsaa056.1187]
65. Fox B, Shihab M, Nassir A, Barnea O, Tal A. Advancing home sleep testing: a comparative analysis of type 2 new HST technology vs. in-lab PSG validation study. Sleep 2024 Apr 20;47(Supplement_1):A136-A137. [doi: 10.1093/sleep/zsae067.0317]

XSL•FO
RenderX

66. Hirshkowitz M. Polysomnography challenges. Sleep Med Clin 2016 Dec;11(4):403-411. [doi: 10.1016/j.jsmc.2016.07.002] [Medline: 28118865]
67. Clinica del sueño [Article in Spanish]. Grupo Infra. URL: https://grupoinfra.com/pagina/servicio/10/Cl%C3%ADnica-del-sue%C3%B1o [accessed 2025-12-09]
68. Clinica de transtornos del sueño [Article in Spanish]. Facultad de Medicina de la Universidad Nacional Autónoma de México. URL: https://clinicadelsueno.facmed.unam.mx/estudios_cts.php [accessed 2025-12-09]
69. Peake JM, Kerr G, Sullivan JP. A critical review of consumer wearables, mobile applications, and equipment for providing biofeedback, monitoring stress, and sleep in physically active populations. Front Physiol 2018;9:743. [doi: 10.3389/fphys.2018.00743] [Medline: 30002629]
70. Chan M, Estève D, Fourniols JY, Escriba C, Campo E. Smart wearable systems: current status and future challenges. Artif Intell Med 2012 Nov;56(3):137-156. [doi: 10.1016/j.artmed.2012.09.003] [Medline: 23122689]
71. de Arriba-Pérez F, Caeiro-Rodríguez M, Santos-Gago JM. How do you sleep? Using off the shelf wrist wearables to estimate sleep quality, sleepiness level, chronotype and sleep regularity indicators. J Ambient Intell Human Comput 2018 Aug;9(4):897-917. [doi: 10.1007/s12652-017-0477-5]
72. Khosla S, Deak MC, Gault D, et al. Consumer sleep technology: an American Academy of Sleep Medicine position statement. J Clin Sleep Med 2018 May 15;14(5):877-880. [doi: 10.5664/jcsm.7128] [Medline: 29734997]
73. Baron KG, Duffecy J, Berendsen MA, Cheung Mason I, Lattie EG, Manalo NC. Feeling validated yet? A scoping review of the use of consumer-targeted wearable and mobile technology to measure and improve sleep. Sleep Med Rev 2018 Aug;40:151-159. [doi: 10.1016/j.smrv.2017.12.002] [Medline: 29395985]
74. de Zambotti M, Claudatos S, Inkelis S, Colrain IM, Baker FC. Evaluation of a consumer fitness-tracking device to assess sleep in adults. Chronobiol Int 2015;32(7):1024-1028. [doi: 10.3109/07420528.2015.1054395] [Medline: 26158542]
75. Kang SG, Kang JM, Ko KP, Park SC, Mariani S, Weng J. Validity of a commercial wearable sleep tracker in adult insomnia disorder patients and good sleepers. J Psychosom Res 2017 Jun;97:38-44. [doi: 10.1016/j.jpsychores.2017.03.009] [Medline: 28606497]
76. Markwald RR, Bessman SC, Reini SA, Drummond SPA. Performance of a portable sleep monitoring device in individuals with high versus low sleep efficiency. J Clin Sleep Med 2016 Jan;12(1):95-103. [doi: 10.5664/jcsm.5404] [Medline: 26285110]
77. Berryhill S, Morton CJ, Dean A, et al. Effect of wearables on sleep in healthy individuals: a randomized crossover trial and validation study. J Clin Sleep Med 2020 May 15;16(5):775-783. [doi: 10.5664/jcsm.8356] [Medline: 32043961]
78. Sateia MJ. International classification of sleep disorders-third edition: highlights and modifications. Chest 2014 Nov;146(5):1387-1394. [doi: 10.1378/chest.14-0970] [Medline: 25367475]
79. Thorpy MJ. Classification of sleep disorders. Neurotherapeutics 2012 Oct;9(4):687-701. [doi: 10.1007/s13311-012-0145-6] [Medline: 22976557]
80. The AASM International Classification of Sleep Disorders – Third Edition Text Revision (ICSD-3-TR): American Academy of Sleep Medicine; 2023. URL: https://aasm.org/clinical-resources/international-classification-sleep-disorders/ [accessed 2025-12-09]
81. Diagnostic classification of sleep and arousal disorders. 1979 first edition. Association of Sleep Disorders Centers and the Association for the Psychophysiological Study of Sleep. Sleep 1979;2(1):1-154. [doi: 10.1093/sleep/2.1.1] [Medline: 531417]
82. FastStats: sleep in adults. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/sleep/data-research/facts-stats/adults-sleep-facts-and-stats.html [accessed 2025-12-09]
83. Behavioral Risk Factor Surveillance System (BRFSS). Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/brfss/index.html [accessed 2025-12-09]
84. WHO technical meeting on sleep and health: Bonn Germany, 22–24 January 2004. : World Health Organization; 2004 URL: https://iris.who.int/handle/10665/349782 [accessed 2025-12-09]
85. Risk factors, comorbidities, and consequences of insomnia in adults. UpToDate. URL: https://www.uptodate.com/contents/risk-factors-comorbidities-and-consequences-of-insomnia-in-adults [accessed 2025-12-09]
86. What are sleep disorders? American Psychiatric Association. URL: https://www.psychiatry.org/patients-families/sleep-disorders/what-are-sleep-disorders [accessed 2025-12-09]
87. Watson NF, Badr MS, Belenky G, et al. Recommended amount of sleep for a healthy adult: a joint consensus statement of the American Academy of Sleep Medicine and Sleep Research Society. Sleep 2015 Jun 1;38(6):843-844. [doi: 10.5665/sleep.4716] [Medline: 26039963]
88. Hirshkowitz M, Whiton K, Albert SM, et al. National Sleep Foundation's updated sleep duration recommendations: final report. Sleep Health 2015 Dec;1(4):233-243. [doi: 10.1016/j.sleh.2015.10.004] [Medline: 29073398]
89. Breaking report: new insights on sleep health during the pandemic - a National Sleep Foundation analysis of over 12,000 Americans in the Sleep Health Index® from 2019 to 2021. : National Sleep Foundation; 2022 URL: https://www.thensf.org/wp-content/uploads/2022/03/National-Sleep-Foundation_Sleep-Health-in-COVID-19-Pandemic-Report.pdf [accessed 2025-12-09]

XSL•FO
RenderX

90. McClintock KT, Mercurio AK, Reilly DS, Cosgrove M, Rachler R. Sleep: architecture, deprivation, and propofol-induced sleep. AANA J 2024 Dec 1;92(6):459-470. [Medline: 39601855]
91. Moszczynski A, Murray BJ. Neurobiological aspects of sleep physiology. Neurol Clin 2012 Nov;30(4):963-985. [doi: 10.1016/j.ncl.2012.08.001] [Medline: 23099125]
92. Sletten TL, Weaver MD, Foster RG, et al. The importance of sleep regularity: a consensus statement of the National Sleep Foundation sleep timing and variability panel. Sleep Health 2023 Dec;9(6):801-820. [doi: 10.1016/j.sleh.2023.07.016] [Medline: 37684151]
93. Lee HC, Seo JG. The effect of long sleep on the recurrence of ischemic stroke. J Sleep Res 2018 Oct;27(S2). [doi: 10.1111/jsr.83_12766]
94. Kim TJ, Kim CK, Kim Y, et al. Long sleep duration as a risk factor for intracerebral hemorrhage: a nationwide, multicenter case-control study. Stroke 2016 Feb;47(suppl_1). [doi: 10.1161/str.47.suppl_1.wp369]
95. Nelson KL, Davis JE, Corbett CF. Sleep quality: an evolutionary concept analysis. Nurs Forum 2022 Jan;57(1):144-151. [doi: 10.1111/nuf.12659] [Medline: 34610163]
96. About sleep. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/sleep/about/ [accessed 2025-12-09]
97. Adjaye-Gbewonyo D, Ng AE, Black LI. Sleep difficulties in adults: United States, 2020 (NCHS data brief no. 436). Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/nchs/products/databriefs/db436.htm [accessed 2025-12-09]
98. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 1;37(1):9-17. [doi: 10.5665/sleep.3298] [Medline: 24470692]
99. Lee S, Kim JH, Chung JH. The association between sleep quality and quality of life: a population-based study. Sleep Med 2021 Aug;84:121-126. [doi: 10.1016/j.sleep.2021.05.022] [Medline: 34147026]
100. Wang S, Li B, Wu Y, et al. Relationship of sleep duration with sociodemographic characteristics, lifestyle, mental health, and chronic diseases in a large Chinese adult population. J Clin Sleep Med 2017 Mar 15;13(3):377-384. [doi: 10.5664/jcsm.6484] [Medline: 27998377]
101. Li L, Wang YY, Wang SB, et al. Sleep duration and sleep patterns in Chinese university students: a comprehensive meta-analysis. J Clin Sleep Med 2017 Oct 15;13(10):1153-1162. [doi: 10.5664/jcsm.6760] [Medline: 28859718]
102. Chen P, Lam MI, Si TL, et al. The prevalence of poor sleep quality in the general population in China: a meta-analysis of epidemiological studies. Eur Arch Psychiatry Clin Neurosci 2024 Oct;274(7):1-14. [doi: 10.1007/s00406-024-01764-5] [Medline: 38429554]
103. Rosipal R, Lewandowski A, Dorffner G. In search of objective components for sleep quality indexing in normal sleep. Biol Psychol 2013 Sep;94(1):210-220. [doi: 10.1016/j.biopsycho.2013.05.014] [Medline: 23751915]
104. O'Donnell D, Silva EJ, Münch M, Ronda JM, Wang W, Duffy JF. Comparison of subjective and objective assessments of sleep in healthy older subjects without sleep complaints. J Sleep Res 2009 Jun;18(2):254-263. [doi: 10.1111/j.1365-2869.2008.00719.x] [Medline: 19645969]
105. Guillodo E, Lemey C, Simmonet M, Ropars J, Berrouiguet S. Sleep monitoring and wearables: a systematic review of clinical trials and future applications (Preprint). JMIR Preprints. Preprint posted online on Dec, 2018. [doi: 10.2196/preprints.13044]
106. Lins LFTDS, do Nascimento EGC, da Silva Júnior JA, de Medeiros Fernandes TAA, de Andrade MF, de Mesquita Andrade C. Accuracy of wearable electronic device compared to manual and automatic methods of blood pressure determination. Med Biol Eng Comput 2023 Oct;61(10):2627-2636. [doi: 10.1007/s11517-023-02869-0]
107. Alshamari MA, Althobaiti MM. Usability evaluation of wearable smartwatches using customized heuristics and system usability scale score. Future Internet 2024;16(6):204. [doi: 10.3390/fi16060204]
108. Baranwal N, Yu PK, Siegel NS. Sleep physiology, pathophysiology, and sleep hygiene. Prog Cardiovasc Dis 2023;77:59-69. [doi: 10.1016/j.pcad.2023.02.005] [Medline: 36841492]
109. Circadian rhythm. Cleveland Clinic. URL: https://my.clevelandclinic.org/health/articles/circadian-rhythm [accessed 2025-12-09]
110. De Pasquale C, El Kazzi M, Sutherland K, et al. Sleep hygiene - what do we mean? A bibliographic review. Sleep Med Rev 2024 Jun;75:101930. [doi: 10.1016/j.smrv.2024.101930] [Medline: 38761649]
111. Hirohama K, Imura T, Hori T, Deguchi N, Mitsutake T, Tanaka R. The effects of nonpharmacological sleep hygiene on sleep quality in nonelderly individuals: a systematic review and network meta-analysis of randomized controlled trials. PLOS ONE 2024;19(6):e0301616. [doi: 10.1371/journal.pone.0301616] [Medline: 38837997]
112. McAlpine T, Mullan B, Clarke PJF. Re-considering the role of sleep hygiene behaviours in sleep: associations between sleep hygiene, perceptions and sleep. IntJ Behav Med 2024 Oct;31(5):705-717. [doi: 10.1007/s12529-023-10212-y]
113. How sleep works - your sleep/wake cycle. National Heart, Lung, and Blood Institute. URL: https://www.nhlbi.nih.gov/health/sleep/sleep-wake-cycle [accessed 2025-12-09]
114. Alanazi EM, Alanazi AMM, Albuhairy AH, Alanazi AAA. Sleep hygiene practices and its impact on mental health and functional performance among adults in Tabuk City: a cross-sectional study. Cureus 2023 Mar;15(3):e36221. [doi: 10.7759/cureus.36221] [Medline: 37069886]

XSL•FO
RenderX

115. Desai D, Momin A, Hirpara P, Jha H, Thaker R, Patel J. Exploring the role of circadian rhythms in sleep and recovery: a review article. Front Neurosci 2024;18:1221196. [doi: 10.7759/cureus.61568]
116. Circadian rhythm. Sleep Foundation. URL: https://www.sleepfoundation.org/circadian-rhythm [accessed 2025-12-09]
117. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239. [doi: 10.1136/bmj.i5239] [Medline: 27777223]
118. Kuula L, Halonen R, Kajanto K, et al. The effects of presleep slow breathing and music listening on polysomnographic sleep measures - a pilot trial. Sci Rep 2020 May 4;10(1):7427. [doi: 10.1038/s41598-020-64218-7] [Medline: 32366866]
119. Wu D, Dang Y, Wang J, Cui Z. Subdural effusion evolves into chronic subdural hematoma after deep brain stimulation surgery: case report and review of the literature. Brain Sci 2022 Oct 10;12(10):1375. [doi: 10.3390/brainsci12101375] [Medline: 36291309]
120. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
121. User manual - English (LTN). 05/2024. rev.1.1. Samsung. 2024. URL: https://downloadcenter.samsung.com/content/UM/202405/20240527192412628/SM-R86X_R87X_R88X_R89X_R90X_R91X_R92X_UM_LTN_Android_D0_Eng_Rev.1.1_240516.pdf [accessed 2025-12-09]
122. Recommendations for a classification of AI use in academic manuscript preparation. The International Association of Scientific, Technical & Medical Publishers (STM). 2025. URL: https://stm-assoc.org/document/recommendations-for-a-classification-of-ai-use-in-academic-manuscript-preparation/ [accessed 2025-12-09]

Abbreviations

AI: artificial intelligence
ML: machine learning
PSG: polysomnography
PSQI: Pittsburgh Sleep Quality Index
RCT: randomized controlled trial
REM: rapid eye movement
SHAP: Shapley Additive Explanations
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TIB: time in bed
TST: total sleep time
XGBoost: extreme gradient boosting

Edited by J Sarvestan; submitted 23.Apr.2025; peer-reviewed by CH Fung, F Gomez; revised version received 18.Nov.2025; accepted 19.Nov.2025; published 06.Jan.2026.

Please cite as:
Quezada Reyes R, Trejo LA
Personalized Machine Learning Intervention to Improve Sleep Quality Using Wearable Technology in Healthy Middle-Aged Adults From Mexico City: Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2026;15:e76415
URL: https://www.researchprotocols.org/2026/1/e76415
doi:10.2196/76415

© Rodrigo Quezada Reyes, Luis A Trejo. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Tailored Individual Follow-Ups Versus a One-Day Group Course in Patients With Long COVID (Post– COVID-19 Condition): Protocol for a Randomized Controlled Trial

Marte Wilson[1], PT, MSc; Synne Garder Pedersen[1,2], PT, PhD; Nina Langeland[3,4,5], MD, PhD; Rebecca Jane Cox[3,6], PhD; Pål Aukrust[7,8], MD, PhD; Tuva B Dahl[7], PhD; Axel Sandvig[9], MD, PhD; Maja Wilhelmsen[1,2,10], MD, PhD

[1]Department of Rehabilitation, University Hospital of North Norway, Tromsø, Norway

[10]Department of Clinical Medicine, UiT The Arctic University of Norway, Postboks 6050 Stakkevollan, Tromsø, Norway

[2]Institute of Health and Society, Research Centre for Habilitation and Rehabilitation Model and Services (CHARM), Faculty of Medicine, University of Oslo, Oslo, Norway

[3]Department of Clinical Science, University of Bergen, Bergen, Norway

[4]Department of Medicine, Haukeland University Hospital, Bergen, Vestland, Norway

[5]Norwegian Institute for Public Health, Oslo, Norway

[6]Influenza Centre, Department of Microbiology, Haukeland University Hospital, Bergen, Vestland, Norway

[7]Research Institute of Internal Medicine, Rikshospitalet, Oslo University Hospital, Oslo, Norway

[8]Faculty of Medicine, University of Oslo, Oslo, Norway

[9]Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway

Corresponding Author:
Maja Wilhelmsen, MD, PhD
Department of Rehabilitation, University Hospital of North Norway, Tromsø, Norway

Abstract

Background: The high prevalence of patients with post–COVID-19 condition, also called long COVID, even among those with mild initial disease, may have a large impact on both the individual and society. Disability in everyday life, reduced health-related quality of life and work capacity, strain on the health care system, and substantial socioeconomic costs are associated with long COVID. More research to investigate the effectiveness of rehabilitation services is warranted.

Objective: This study aims to examine the effectiveness of tailored individual follow-ups versus a 1-day group course in patients with long COVID. Additionally, the feasibility and use of a mobile app for self-monitoring goal achievement will be assessed.

Methods: This is a single-center, parallel-group, superiority randomized controlled trial with a 1:1 allocation ratio. A total of 62 outpatients aged 18‐65 years with long COVID will be randomized to either a rehabilitation program with individual follow-up consultations or a 1-day self-management group course. The individual intervention incorporates setting goals, teaching cognitive behavioral strategies, energy management (pacing), and a supervised gradual increase in both physical and cognitive activities tailored to individual tolerance levels. The primary outcome is the between-group difference in health-related quality of life, measured using the EQ-5D-5L index at 6 months. Secondary outcomes include improvements in symptoms, work participation, neurocognitive function, and app usability, assessed at 3, 6, and 12 months, depending on the outcome measure.

Results: Data enrollment started in October 2023. A total of 62 participants were included by November 2024. Data collection is planned to be completed in November 2025.

Conclusions: Long COVID poses substantial challenges for both individuals and society, underscoring the need for effective rehabilitation strategies. This study will provide valuable insights into the benefits of an individualized outpatient rehabilitation program. The results from this clinical trial will help guide future treatment recommendations and may improve long-term outcomes for affected patients. Additionally, the study will generate important knowledge about neuropsychological function and digital self-management tools in long COVID rehabilitation.

Trial Registration: ClinicalTrials.gov NCT06085911; https://www.clinicaltrials.gov/study/NCT06085911

International Registered Report Identifier (IRRID): DERR1-10.2196/74113

(JMIR Res Protoc 2026;15:e74113) doi:10.2196/74113

XSL•FO
RenderX

KEYWORDS

postacute COVID-19 syndrome; graded activity; fatigue; rehabilitation; cognitive training; telerehabilitation; randomized controlled trial; cognitive behavioral therapy; exercise; mobile apps

Introduction

Symptoms and Consequences

By the end of 2023, more than 277 million confirmed SARS-CoV-2 infections were reported in Europe [1]. According to the World Health Organization (WHO), 10% to 20% of those who have been infected develop post–COVID-19 condition, also called long COVID. This condition is defined as "individuals with a history of probable or confirmed SARS-CoV-2 infection, usually 3 months from the onset, with symptoms that last for at least 2 months and cannot be explained by an alternative diagnosis. Common symptoms include, but are not limited to, fatigue, shortness of breath, and cognitive dysfunction, and generally have an impact on everyday functioning" [2]. Long-lasting symptoms affect not only hospitalized patients, but worryingly also nonhospitalized people after a mild infection [3].

Long COVID is a complex multisystem secondary condition [4] and a new disease of partly unknown pathogenesis [5-7]. Long-term symptoms following infection are similar to those found in conditions where the central nervous system is affected [5,6] and consist of key symptoms similar to chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) [8,9]. One of the main complaints reported by patients with long COVID is impaired cognitive function [9-13], which includes self-reported fatigue, memory deficits, difficulty concentrating, and difficulty in information processing. A variety of other persistent symptoms are reported, including headache, dyspnea, cough, chest pain, myalgia, joint pain, impaired mobility, sleep disorders, depression, and anxiety [10]. Symptoms of long COVID may lead to disability in everyday life, reduced health-related quality of life (HRQoL) [14,15], strain on the health care system, and substantial socioeconomic costs [15]. Moreover, long COVID is strongly associated with a reduction in workability [16].

Treatment Content

Different treatment options are delivered to patients with long COVID: drug interventions, dietary interventions, or behavioral interventions [17]. In patients with complex illness, behavioral interventions are often an important factor stimulating recovery [18,19]. Behavioral interventions can vary in intensity (eg, number of consultations), content (eg, graded activity or aerobic training), modality (eg, group-based or individual), and may be mono- or multidisciplinary.

Cognitive behavioral therapy (CBT) is based on the theory that our body (symptoms), thoughts, feelings, and behaviors are interconnected. Adjustments in behavior and thoughts can be helpful for somatic symptoms, and adjustments can be implemented through home-based exercises [18]. Results of a randomized controlled trial (RCT) using CBT have shown significantly less severe fatigue in patients with long COVID across follow-up assessments compared to patients receiving standard care [20]. Participants received substantial follow-ups over 17 weeks, and the control group received care as usual. All secondary outcomes, including social and physical functioning, somatic symptoms, and concentration, also favored patients receiving CBT; however, HRQoL was not evaluated. Patients in this study were mostly nonhospitalized [20]. A recent review recommends behavioral interventions based on CBT as management of long COVID [17].

A previous systematic review reported that regular physical exercises increase physical performance, neurocognitive function as well as fatigue and HRQoL [3] in patients with long COVID. In a cohort study by Daynes et al [21], patients with long COVID showed improvements in all those parameters listed above after supervised exercise training (aerobic and strength training) in combination with education and pacing advice twice a week for 6 weeks. However, this study had only 30 participants and did not include a control group.

A case-control study by Palladini et al [22] detected improved cognitive performance in hospitalized patients with long COVID, with a correlation to measures of quality of life. The study intervention included cognitive exercises of the most affected cognitive domains and counseling follow-ups for 2 months. Furthermore, cognitive exercise in patients with long COVID is supported by several other studies [5,9,23].

According to the WHO, patients with long COVID may experience worsening of symptoms following minimal cognitive, physical, emotional, and social activity, or activity that could previously be tolerated [24]. A result of this is often a behavioral challenge where patients exceed their own tolerance limit, followed by complete rest, an all-or-nothing pattern. Others may generally fear activity because it may lead to exacerbation of symptoms, also called fear-avoidance. Such patterns are often seen in CFS/ME [25]. Graded fixed incremental exercise therapy is debated in the literature because it might cause PEM, which is associated with poor health outcomes [26-29]. Pacing, on the other hand, is how to balance activities and rest to avoid exacerbation of symptoms, and thus not fixed incremental steps [24]. In patients with long COVID, individualized activity plans with education on pacing should therefore be provided. It is documented as beneficial to gradually introduce activity to restore patients to previous levels of activity, including daily routines [5,30]. Interventions that combine pacing and graded escalation of physical and cognitive activities are called graded activity [25]. A recent review (2023) found graded activity based on elements from CBT to be effective in reducing fatigue and improving physical functioning in CFS/ME [25].

Findings from previous studies indicate that long COVID is a complex condition that can be treated by a graded activity intervention. Daily routines and regular physical and cognitive exercise may be 3 important focus areas in this behavioral management of long COVID.

XSL•FO
RenderX

Treatment Modality

Evidence from recent studies supports international guidelines [24,31] in terms of individual [17], multidisciplinary, and multimodal [15,21,32] rehabilitation approaches in long COVID.

In patients with CFS/ME, a self-management program delivered as group education without follow-up consultations did not have any effect compared to usual care on fatigue and physical functioning [33]. In contrast, a recent review on CFS/ME found that individual clinical follow-ups have an effect on fatigue and physical functioning [25]. Further, a digital individual self-efficacy program based on a holistic approach showed several benefits in supporting physical health and mental well-being in patients with long COVID [34].

Telerehabilitation, including video consultations, has also shown positive results in long COVID [35-37]. To ensure continuity in the rehabilitation process and promote self-management of individual rehabilitation goals [38], mobile apps have also shown effectiveness in long COVID [37,39,40].

Knowledge Gap and Study Aims

Until today, rehabilitation studies in patients with long COVID are limited and show a substantial heterogeneity in treatment content, modality, and outcome measures, making it difficult to conclude the effectiveness [3,15,17]. Common to many interventions is that they provide clinical follow-ups over time, from 6 to 17 weeks [3,15,20,22,35]. It is still unclear how treatment content should be combined and which modality is most effective.

Recent comprehensive reviews propose conducting high-quality RCT studies to gain more insight into strategies for structuring the rehabilitation and examining the effectiveness of the management of long COVID [3,9,15,17,41]. It is unclear what treatment option is best to achieve the goal of restoring patients with long COVID to their previous levels of functioning [3], everyday function, including HRQoL [15], cognitive function [9], and workability [5-7]. More research is warranted to evaluate cognitive symptoms in patients with long COVID, and most conducted studies have used self-reported outcome measures [8,10]. Supplementing assessments with standardized neuropsychological tests may refine research findings.

To the best of our knowledge, no RCTs have so far investigated the management of long COVID with neurocognitive complaints comparing 2 rehabilitative interventions: a graded activity intervention with tailored individual follow-ups versus a one-day group course. Both intervention groups will be recommended to use the mobile app *My COVID Rehabilitation* as a supplement to the intervention for self-management of personal goals. Our RCT addresses the present knowledge gap, aiming to add robust evidence on what may be an effective rehabilitation approach for long COVID. The main research question is do tailored individual follow-ups give better outcomes in patients with long COVID compared to a 1-day group course? Additionally, the study will generate important knowledge about neuropsychological function and a digital self-management tool in long COVID rehabilitation.

Methods

This study protocol adheres to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines (Checklist 1), providing reporting recommendations for protocols of randomized trials [42].

Study Setting

This is a single-center, two-armed, parallel-group, superiority RCT with a 1:1 allocation ratio. Assessments are at baseline and at 3, 6, and 12 months. Participants will receive either individual rehabilitation (intervention group) or a 1-day group course (standard treatment). The study is embedded in ordinary clinical care at the University Hospital of North Norway (UNN). No blinding is performed except for a blinded statistical evaluation.

Participants

A total of 62 adult participants will be included in this trial at the outpatient regional long COVID clinic at UNN. Patients are first thoroughly examined by their general practitioners to exclude other underlying diseases explaining current symptoms. Patients with symptoms of long COVID are referred to our regional long COVID clinic where clinical assessment is done primarily in a video consultation. The assessment is done by a bi-disciplinary team led by a doctor, supplemented by either a physiotherapist or a psychologist based on the medical history. During the video consultation, patients are briefly informed about the study, and consent for a SMS text message invitation is obtained.

Eligibility Criteria

Individuals will be included if they had (1) symptoms attributable to long COVID affecting daily activities according to the WHO's definition, (2) a positive COVID-19 test via a home test or polymerase chain reaction test, (3) neurocognitive symptoms, (4) age between 18 and 65 years, and (5) ability and willingness to provide informed consent. Exclusion criteria were as follows: (1) patients who did not want to comply with the planned physical study visits, (2) patients who were unable to complete surveys in Norwegian, and (3) patients with known chronic neurocognitive disease before COVID-19 or other diseases that can explain current symptoms. The inclusion age is set to 18 to 65 years, as one of the outcome measures in the study is return to work.

After assessment in the COVID clinic, eligible patients will receive an SMS text message invitation with a link to an information sheet and consent form (Multimedia Appendix 1). Patients are encouraged to contact the research coordinator by phone for supplementary information and possibly provide oral consent. Written informed consent is collected at the first visit. Each participant has the right to withdraw from the study at any time. Any adverse events will be recorded. The research coordinator is responsible for the enrollment process.

Randomization

Participants will be block randomized with the computer program Research Electronic Data Capture (REDCap) either to

XSL•FO
RenderX

the intervention group (n=31) or standard treatment group (n=31), presented in Figure 1.

Figure 1. Study design and overview of data collection. WHO: World Health Organization.

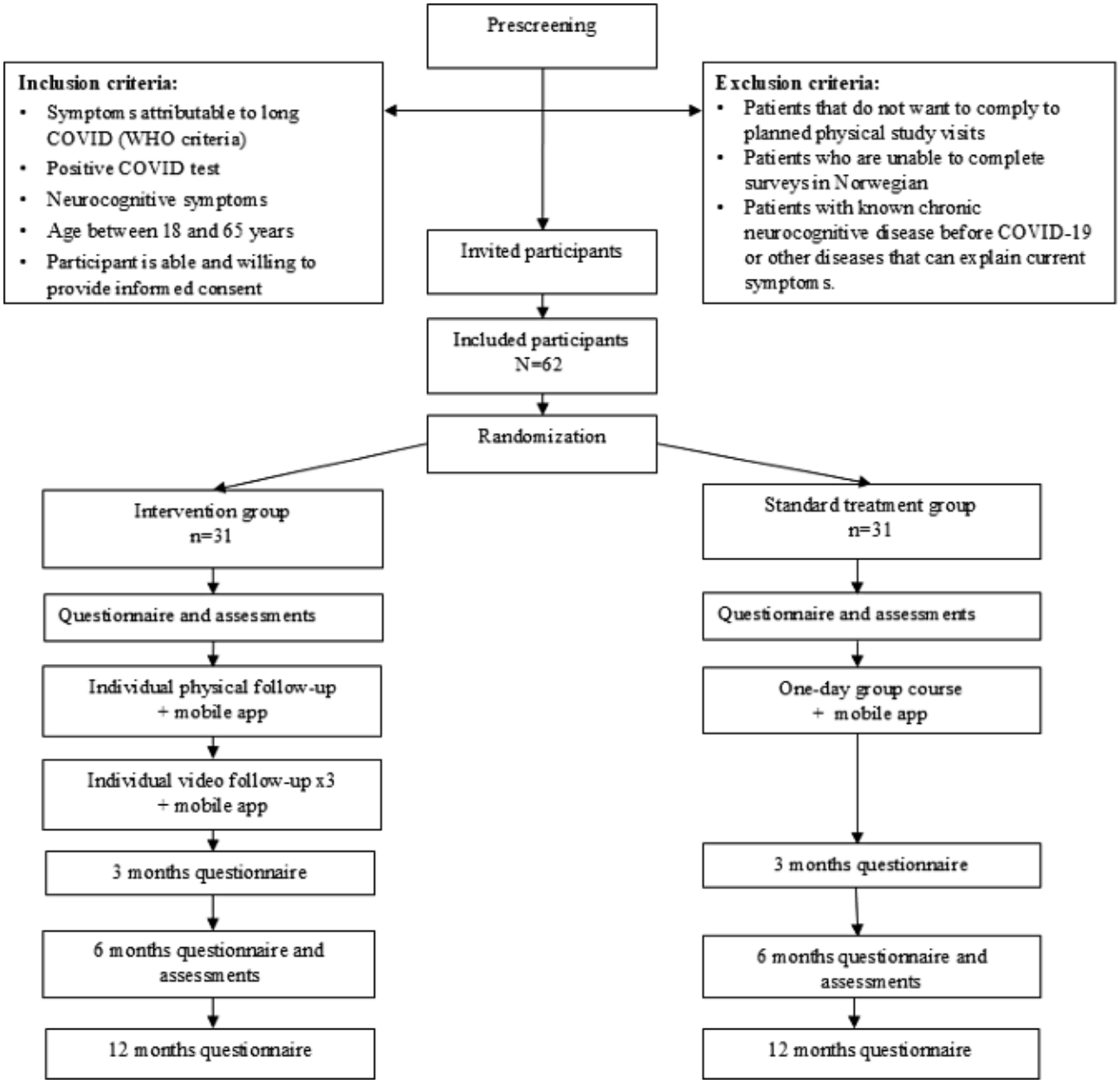

Standard Treatment: One-Day Group Course

Standard treatment at UNN is a one-day multidisciplinary group course. The duration is 6 hours, including breaks. Each group consists of 8 to 10 participants. The aim of the course is to improve self-management of long COVID. The course provides education about (1) sustaining factors of long COVID based on elements from a CBT framework, (2) activity regulation, (3) the importance of daily routines, and the advantages of both (4) physical and (5) cognitive exercise. Information is provided by a doctor, a psychologist, and an occupational therapist or a physiotherapist. Exchange of experience between participants is also facilitated. No follow-up consultations are provided. Information about the mobile app supplement is offered; a mobile app called *My COVID Rehabilitation*. The participants can voluntarily download the mobile app as a supplement, and personal goals can be noted and adjusted.

Study Intervention

The intervention in this study is delivered individually by a physiotherapist following the study instructions described in Textbox 1. The physiotherapist may consult the multidisciplinary team if necessary. The intervention begins with a physical consultation where symptoms, function in everyday life, sustaining factors, and resources are thoroughly assessed. Education about long COVID based on elements from a CBT framework [18] is provided. Tailored to the participant's life, a personalized rehabilitation plan is created, consisting of home-based exercises to facilitate a combination of pacing and graded escalation of physical and cognitive activities. The home-based exercises have 3 focus areas: daily routine, physical activity, and cognitive training. The plan consists of weekly personal goals to guide the participant's progress. Three video-follow-up consultations are given to support change and

XSL•FO
RenderX

to encourage participants to adjust personal goals in step with recovery within the 3 focus areas. Supported by the physiotherapist, the personalized goal setting is then verbalized and adjusted during the follow-up consultations. In total, 4 consultations are given within 6 - 8 weeks. Voluntarily, the participants can download the *My COVID Rehabilitation* mobile app as a supplement, where the goals can be noted and adjusted.

Textbox 1. Overview of the content of the individual study intervention.

- Part 1: Education about long COVID and coping strategies
  - Session 1: Physical consultation (1.5 h) where symptoms, function in everyday life, sustaining factors and resources are mapped out. Provide education about long COVID based on elements of cognitive behavioral therapy. Information about treatment recommendations emphasizing adjustments of behavior. Tailored goal-setting work based on the 3 focus areas: daily routine (eg, sleep rhythm), physical activity (eg, housework or run), and cognitive exercises (reading comics).
  - *Home assignment:* To implement behavioral change according to participants' personal goals, voluntarily noted in the *My COVID Rehabilitation* mobile app.
- Part 2: Behavioral change and home exercises
  - Session 2 and session 3: Video follow-up (60 min). Current status and review of home exercises after previous sessions. Tailored education about how adjustments in behavior may gain a positive health effect. Reflect on goal achievement and adjustments of goals.
  - *Home assignment*: Participants are encouraged to continue to work on their personal goals, voluntarily noted in the *My COVID Rehabilitation* mobile app.
- Part 3: Summary and further plan
  - Video follow-up (60 minutes). Current status and review of home exercises after previous session. Continued tailored education about how adjustments in behavior may gain a positive health effect. Reflect on goal achievement and adjustments of goals. Discussion about how to maintain and secure further progress.
  - *Home assignment:* Participants are encouraged to continue to work on their tailored personal goals, voluntarily noted in the *My COVID Rehabilitation* mobile app.

Data Collection

Questionnaires will be sent to participants by email and captured by REDCap. The questionnaires are web-based and answered online. Reminders will be sent, if needed, by email once, SMS text messaging twice, and phone call once to improve adherence. If the participants have any questions regarding the questionnaires, they will be offered help by the research coordinator. The questionnaires will capture patient-reported outcome measurements at baseline and at 3, 6, and 12 months, including EQ-5D-5L index, EQ-5D-5L VAS, Chalder Fatigue Scale, number of symptoms, work participation, and feasibility of mobile app. At baseline, the patient also reports demographic data and relevant background information. Data is stored securely, and only the research team has access to the data. An overview of the measures and timepoints for assessments is presented in Table 1.

XSL•FO
RenderX

Table . Outcome measurements.

Type of outcome	Measurement	Timepoints
Health-related quality of life		
Primary outcome	• EQ-5D-5L index [43]	6 mo
Secondary outcome (1)	• EQ-5D-5L index [43]	0, 3, 12 mo
Secondary outcome (2)	• EQ VAS[a] [43]	0, 3, 6, 12 mo
Number of and grading of symptoms		
Secondary outcome	• Self-completed questionnaire	0, 3, 6, 12 mo
Work participation		
Secondary outcome	• Self-completed questionnaire	0, 3, 6, 12 mo
Fatigue		
Secondary outcome	• Chalder Fatigue Scale (11 items) [44]	0, 3, 6, 12 mo
Utility and degree of mobile app usage		
Secondary outcome	• Self-completed questionnaire	3, 6, 12 mo
Neuropsychological function		

Secondary outcome	• Assessment from 4 cognitive domains:
	  • Episodic memory [45-47]
	  • Working memory: [48]
	  • Executive functions: [49,50]
	  • Processing speed [51]
		0, 6 mo

[a]VAS: visual analog scale.

Primary Outcome: HRQoL Measured With EQ-5D-5L

The primary outcome is HRQoL at the 6-month follow-up, assessed using the well-known EQ-5D-5L that gives a total index score [43,52]. EQ-5D-5L is a validated generic scale [52] and has also been applied and shown to be responsive to long COVID [37,53,54]. The 5 dimensions include mobility, self-care, usual activity, pain and discomfort, and anxiety and depression. Each dimension is assessed with 1 question, which has 5 levels of severity (“no,” “slight,” “moderate,” “severe,” and “extreme” problems). The Norwegian score has a range from −0.453 to 1, where 0 equals death and 1 is perfect health [55].

Secondary Outcomes

Demographics

At baseline, demographic data are reported, including height, weight, time since infection, number of vaccines, medication, comorbidities, and highest level of education and occupational status.

Health-Related Quality of Life

HRQoL measured by the EQ-5D-5L index at 0, 3, and 12 months and the EQ visual analog scale (VAS) are part of the secondary outcomes. EQ-VAS instructs respondents to rate their overall health on a scale from 0 (the worst imaginable health) to 100 (the best imaginable health). This provides a global assessment of health status [56].

Fatigue

Fatigue is investigated using the Chalder Fatigue Scale as part of the questionnaire package. The fatigue scale is valid and reliable as a generic self-administered questionnaire for measuring the extent and severity of physical and mental fatigue. Each of the 11 items is answered on a 4-point scale ranging from asymptomatic to maximum symptomology [44], which are summed to a total score ranging from 0 to 33. The fatigue scale has previously been used in research on long COVID [57,58].

Symptoms

Symptom burden is explored through the presence of 13 common symptoms of long COVID (binary): fever, tingling, headache, dizziness, palpitations, nausea, cough, vomiting, diarrhea, altered taste and smell, sleep disturbances, chest pain, and muscle and joint pain. Participants also rate the severity of dyspnea and feeling of depression on a 4-point scale. The questions have previously been used in COVID research and have been shown to detect change [59].

Work Participation

Absence from work will be recorded as yes or no (binary) questions. In addition, the number of days of absence from work and the percentage of the absence are self-reported.

XSL•FO
RenderX

Mobile App

App use and perceived usefulness of the mobile app will be registered. Participants will be asked if they have used the mobile app. If yes, they will rate the frequency and usefulness on a 4-point scale.

Health Care Services

Participants will be asked to report contacts with the health care services and any use of medication.

Neurocognitive Functioning

A comprehensive neuropsychological assessment will be conducted at baseline and repeated at 6 months to evaluate 4 key cognitive domains. The following measures from standardized tests will be used to measure aspects of the 4 selected cognitive domains.

1. Episodic memory: the average score from the total number of correct words recalled in the delayed recall trial of the Rey Auditory Verbal Learning Test [45,46] and the Rey Complex Figure Test [47].
2. Working memory: average score from the Letter-number-sequencing and Spatial Span subtests of the Wechsler Memory Scale-III [48].
3. Executive functions: average completion time for the trial 1-5 of the Trail Making Test and the Inhibition trial of the Color-Word Interference Test, both from the Delis-Kaplan Executive Function System [49,50].
4. Processing speed: average reaction time from the Conners Continuous Performance Test, 3rd edition [51].

Raw test scores are standardized using age-corrected published population norms. When available, published normative data from Scandinavian populations are prioritized. Scores will be reversed where necessary (reaction time) to ensure that higher scores consistently indicate better performance. Each assessment session will last 90 minutes, and the tests will be administered by a psychologist or trained health care personnel under supervision.

mHealth Supplement: COVID Rehabilitation Mobile App

As part of this study, and at the request of our user representative, a mobile app was developed. The app has been named *My COVID Rehabilitation* (Norwegian: *Min COVID Rehabilitering*). Both groups will be invited to download the mobile app. The mobile app is a digital pamphlet with relevant information, including the possibility to make personal notes and goals based on activities from the 3 focus areas. Users will receive a daily reminder of their goals. After completing the activity, a simple digital reward is provided. The mobile app is available on the App Store and Google Play, and only the participant will have access to personal notes and goals in the mobile app on their phone. The mobile app has been developed in collaboration with user representatives, the multidisciplinary health care team at UNN, and system developers at Information and Communication Technology, Northern Norway Regional Health Authority. In the development process, feedback was given at several time points and adjustments made on both content, phrasing, and visual presentation.

Explorative Outcomes

This trial is embedded in ordinary clinical care at UNN with the last author (M Wilhelmsen) as project leader. At the same time, the study is part of a national research collaboration called ReCover. Participants are therefore invited to participate in voluntary supplement investigations at baseline and at 6-month follow-up. Written informed consent, including the collection, storage, and use of participant data, is obtained at the first visit. Explorative investigations of immunological, inflammatory, and metabolic changes of markers in plasma/serum and whole blood RNA and EEG/MRI are done to explore changes in brain volume, density of white matter fiber tracts, and brain activity. The ReCover study will explore if immune profile and changes in the brain are modulated by the intervention. Details on these supplementary investigations will be provided in other articles from the ReCover research group.

Statistical Analyses

Sample Size

Sample size justification was made in collaboration with a statistician. Due to the lack of studies with a similar design, we could not have completely precise estimates of the between-group difference in EQ-5D-5L index and standard deviation. We based our sample size estimation on a study by Espinoza-Bravo et al [37] that closely resembled ours. Espinoza-Bravo et al [37] compared 2 interventions consisting of home-based exercise with digital follow-ups in both groups. They found a mean difference between groups of (SD) 0.06 on EQ-5D-5L index and a standard deviation (SD) of 0.1 at the end of the intervention. There is no established minimal clinically significant difference (MCID) for the EQ-5D-5L index in long COVID [17]. However, for a respiratory condition (chronic obstructive pulmonary disease), the estimated MCID is 0.05 [60], and this has been used in previous research in long COVID [37]. In stroke rehabilitation, the MCID for the EQ-5D-5L index is found to be 0.10 [61]. In our study, we expect a higher mean difference (SD 0.07) in the EQ-5D-5L index between groups than in the study by Espinoza-Bravo et al [37] because we have a larger difference in treatment design. We also consider that the difference of 0.07 could be clinically significant and therefore based the sample size calculation on a mean difference of 0.07 and an SD of 0.1. We have a one-sided hypothesis where we assume that the intervention group will have a higher EQ-5D-5L index than the control group after 6 months. With a 15% (9/62) expected dropout, a power of 80%, and a P value of .05, the required sample size is estimated to be 31 participants in each group. The power estimation was run in SAS Enterprise Guide 8.3 using the POWER procedure. Other RCTs in populations with long COVID have demonstrated differences between groups with similar sample sizes [62,63].

Data Analyses

The analysis plan was made in collaboration with a statistician. The primary outcome, between-group difference in EQ-5D-5L index score at 6 months follow-up, will be analyzed by a statistician blinded to the intervention groups. Prior to statistical analysis, we will run an exploratory analysis of the outcome to check for data distribution and the presence of outliers, and to

XSL•FO
RenderX

check for associations between the outcomes and covariates. To explore the effect of improved HRQoL, we will use linear models adjusted for baseline HRQoL, age of patients, and gender. The response variable in the model will be HRQoL at the 6-month follow-up. The linear model will be checked using common diagnostic tools (eg, distribution of residuals and influential observations). We will also run an exploratory analysis of the secondary outcomes. Next, because secondary outcomes will be measured repeatedly at 0, 3, 6, and 12 months, we will apply statistical models that are suitable for longitudinal data (eg, generalized estimating equation or mixed models class). Common model diagnostics will also be used for secondary outcomes.

Results for the primary outcome analysis were reported as linear regression coefficients with 95% confidence interval and *P* value. The standard cut-off of $P<.05$ will be used for statistical significance levels. We will evaluate if the difference is clinically and statistically significant. For secondary outcomes, we will report appropriate effect measures with 95% CIs.

For neurocognitive function, to facilitate comparison across cognitive domains, all scores will be converted to *z* scores ($z = (x - \mu)/\sigma$). Each individual cognitive test score will be described by mean and SD, and the cutoff for cognitive impairment on any test or domain score will be set to 1 SD or less from the normative mean.

We will analyze significant between-group differences in cognitive domains at baseline to ensure comparability between the intervention and the control groups. We will also examine the mean changes in the 4 cognitive domains by calculating each participant's average change in z-scores from T1 to T2 for each domain. These change scores are then compared between the intervention and control groups using significance testing and effect sizes [partial eta-squared (η^2)]. Additionally, we will compute the Reliable Change Index (RCI) for each participant in both groups to categorize individual performance as stable, deteriorating, or improving based on clinically significant individual changes defined by the RCI. The RCI will be calculated using test-retest reliability estimates from published, normative samples [64].

Ethical Considerations

The Regional Committee for Medical and Health Research Ethics in northern Norway (number 587293) approved this study based on protocol version 5, April 12, 2024.

The study has been registered at ClinicalTrials.gov under the identifier NCT06085911. This study will be conducted according to the ethical principles of the Declaration of Helsinki (2013) [65]. Written informed consent will be obtained from all included participants before baseline assessments. Consent and personal data are stored using a linkage key with deidentified data on a secure, closed research server. All personal data will be deidentified during analysis and dissemination of results. Participants do not receive any compensation for participating in the project, and they can unconditionally withdraw at any time.

Results

The data collection (October 2023-December 2025) is conducted with grants from KlinBeForsk (grant 34476). Recruitment of participants and data collection started in November 2023. The data collection, including 12 months of follow-ups for all participants, is estimated to finish in November 2025. Recruitment of 62 participants was completed in October 2024. Results are expected to be published in a scientific journal in 2025 or 2026.

Discussion

The present study investigates the effectiveness of tailored individual follow-ups versus a one-day group course in patients suffering from long COVID. Both groups are based on elements of CBT. Individual follow-ups with home-based exercises focusing on graded activity are expected to be more effective than a one-day group course. The main aim is to explore if the study intervention has better results of HRQoL at 6 months follow-up compared to a one-day group course. Other secondary outcomes will also be explored at different time points. Both intervention groups will be recommended to use the mobile app *My COVID Rehabilitation* as a supplement to the intervention for self-management of personal goals.

A strength of this study is that it is embedded in ordinary clinical care. This strengthens external validity. Further, this is a single-center study, and the lack of blinding for group allocation is considered a limitation. The inclusion age is set to 18 - 65 years as one of the outcome measures in the study is return to work, which excludes older people and might therefore be a limitation. Long COVID is a complex condition, and therefore, both groups receive treatment combining several elements. The study will therefore not be able to identify which of the treatment elements were effective. Also, the study cannot conclude about the effectiveness of the mobile app as it is a voluntary element in both groups. The main outcome, EQ-5D-5L index, is recognized as a generic HRQoL measure and has proven to be responsive in long COVID populations [37,53]. Recent studies have questioned the instrument's ability to capture fatigue as a particularly dominant symptom [66]. However, to compensate for this potential weakness, the Chalder Fatigue Scale, which is a more specific measure of subjective fatigue, is also included as an outcome measure.

Rehabilitation RCTs in patients with long COVID are limited and show a heterogeneity in both treatment content, treatment approach, and outcome measures, making it difficult to conclude the effectiveness [3,15,17]. Common to many interventions is that they provide clinical follow-ups over time, often 6 to 12 weeks [3,15,22,35]. It is still unclear how treatment content should be combined and which modality is most effective. We argue that long COVID is a complex condition and thereby may demand treatment with more than 1 focus area and individual follow-ups to reduce the impact of long COVID. This study will provide knowledge about rehabilitation interventions in long COVID, their effectiveness, and whether individual follow-ups are better than a 1-day group course. This study has potentially important implications for patients with long COVID,

XSL•FO
RenderX

health services, and society. Individual follow-ups are expected to lead to increased adherence in terms of the individual rehabilitation plan within the 3 focus areas and thereby achieving better results in terms of less symptom burden, increased function, work participation, and lower costs for society. Additionally, the study will generate important knowledge about neuropsychological function and digital self-management tools in long COVID rehabilitation.

Acknowledgments

The University Hospital of North Norway, our sponsor, has provided clinical resources, for example, support from the Department of Clinical Research Tromsø, that has been valuable. The authors thank all the participants enrolled in the study and all the clinical staff who assisted in this study. Academic feedback from professor Audny Anke, neuropsychologist Marte Ørbo, and psychologist Bjørn Eivind Kirsebom has been of great importance. Two user representatives, Hans Richard Bråten and Bjørn Egil Bjørnstad LHL (Norwegian Heart, Lung and Stroke Association), have given valuable input. No generative artificial intelligence was applied in the writing process.

Authors' Contributions

Conceptualization: M Wilhelmsen, NL, PA, AS, TBD, RJC
Data curation: M Wilson
Funding acquisition: M Wilhelmsen, NL, PA, AS, TBD, RJC
Writing – original draft: M Wilson and M Wilhelmsen
Writing – review & editing: all authors
M Wilson, SGP, and M Wilhelmsen have contributed to the implementation phase of the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Consent form.
[PDF File, 182 KB - resprot_v15i1e74113_app1.pdf]

Checklist 1
SPIRIT checklist.
[PDF File, 239 KB - resprot_v15i1e74113_app2.pdf]

References

1. Post COVID-19 condition. World Health Organization. 2022. URL: https://www.who.int/europe/news-room/fact-sheets/item/post-covid-19-condition [accessed 2022-12-07]
2. Soriano JB, Murthy S, Marshall JC, Relan P, Diaz JV, Definition WCC. A clinical case definition of post-COVID-19 condition by a Delphi consensus. Lancet Infect Dis 2022 Apr;22(4):e102-e107. [doi: 10.1016/S1473-3099(21)00703-9] [Medline: 34951953]
3. Sick J, König D. Exercise training in non-hospitalized patients with post-COVID-19 syndrome-a narrative review. Healthcare (Basel) 2023 Aug 12;11(16):37628475. [doi: 10.3390/healthcare11162277] [Medline: 37628475]
4. Greenhalgh T, Knight M, A'Court C, Buxton M, Husain L. Management of post-acute covid-19 in primary care. BMJ 2020 Aug 11;370:m3026. [doi: 10.1136/bmj.m3026] [Medline: 32784198]
5. Fine JS, Ambrose AF, Didehbani N, et al. Multi-disciplinary collaborative consensus guidance statement on the assessment and treatment of cognitive symptoms in patients with post-acute sequelae of SARS-CoV-2 infection (PASC). PM R 2022 Jan;14(1):96-111. [doi: 10.1002/pmrj.12745] [Medline: 34902226]
6. Rolin S, Chakales A, Verduzco-Gutierrez M. Rehabilitation strategies for cognitive and neuropsychiatric manifestations of COVID-19. Curr Phys Med Rehabil Rep 2022;10(3):182-187. [doi: 10.1007/s40141-022-00352-9] [Medline: 35602927]
7. Chuang HJ, Lin CW, Hsiao MY, Wang TG, Liang HW. Long COVID and rehabilitation. J Formos Med Assoc 2024 Jan;123 Suppl 1:S61-S69. [doi: 10.1016/j.jfma.2023.03.022] [Medline: 37061399]
8. Weigel B, Eaton-Fitch N, Thapaliya K, Marshall-Gradisnik S. Illness presentation and quality of life in myalgic encephalomyelitis/chronic fatigue syndrome and post COVID-19 condition: a pilot Australian cross-sectional study. Qual Life Res 2024 Sep;33(9):2489-2507. [doi: 10.1007/s11136-024-03710-3] [Medline: 38961009]
9. Möller M, Borg K, Janson C, Lerm M, Normark J, Niward K. Cognitive dysfunction in post-COVID-19 condition: mechanisms, management, and rehabilitation. J Intern Med 2023 Nov;294(5):563-581. [doi: 10.1111/joim.13720] [Medline: 37766515]

XSL•FO
RenderX

10. Andronescu LR, Richard SA, Scher AI, et al. SARS-CoV-2 infection is associated with self-reported post-acute neuropsychological symptoms within six months of follow-up. PLoS ONE 2024;19(4):e0297481. [doi: 10.1371/journal.pone.0297481] [Medline: 38626117]
11. Groff D, Sun A, Ssentongo AE, et al. Short-term and long-term rates of postacute sequelae of SARS-CoV-2 Infection: a systematic review. JAMA Netw Open 2021 Oct 1;4(10):e2128568. [doi: 10.1001/jamanetworkopen.2021.28568] [Medline: 34643720]
12. Barbara C, Clavario P, De Marzo V, et al. Effects of exercise rehabilitation in patients with long coronavirus disease 2019. Eur J Prev Cardiol 2022 May 25;29(7):e258-e260. [doi: 10.1093/eurjpc/zwac019] [Medline: 35078233]
13. Ceban F, Ling S, Lui LMW, et al. Fatigue and cognitive impairment in post-COVID-19 syndrome: a systematic review and meta-analysis. Brain Behav Immun 2022 Mar;101(93–135):93-135. [doi: 10.1016/j.bbi.2021.12.020] [Medline: 34973396]
14. Malesevic S, Sievi NA, Baumgartner P, et al. Impaired health-related quality of life in long-COVID syndrome after mild to moderate COVID-19. Sci Rep 2023 May 12;13(1):7717. [doi: 10.1038/s41598-023-34678-8] [Medline: 37173355]
15. Fugazzaro S, Contri A, Esseroukh O, et al. Rehabilitation interventions for post-acute COVID-19 syndrome: a systematic review. Int J Environ Res Public Health 2022 Apr 24;19(9):5185. [doi: 10.3390/ijerph19095185] [Medline: 35564579]
16. Kerksieck P, Ballouz T, Haile SR, et al. Post COVID-19 condition, work ability and occupational changes in a population-based cohort. Lancet Reg Health Eur 2023 Jun 23;31(100671):100671. [doi: 10.1016/j.lanepe.2023.100671] [Medline: 37366496]
17. Zeraatkar D, Ling M, Kirsh S, et al. Interventions for the management of long covid (post-covid condition): living systematic review. BMJ 2024 Nov 27;387:e081318. [doi: 10.1136/bmj-2024-081318] [Medline: 39603702]
18. Beck JS. Cognitive Behavior Therapy: Basics and Beyond, 2nd edition: Guilford Press; 2011.
19. White PD, Goldsmith KA, Johnson AL, et al. Comparison of adaptive pacing therapy, cognitive behaviour therapy, graded exercise therapy, and specialist medical care for chronic fatigue syndrome (PACE): a randomised trial. Lancet 2011 Mar 5;377(9768):823-836. [doi: 10.1016/S0140-6736(11)60096-2] [Medline: 21334061]
20. Kuut TA, Müller F, Csorba I, et al. Efficacy of cognitive-behavioral therapy targeting severe fatigue following coronavirus disease 2019: results of a randomized controlled trial. Clin Infect Dis 2023 Sep 11;77(5):687-695. [doi: 10.1093/cid/ciad257] [Medline: 37155736]
21. Daynes E, Gerlis C, Chaplin E, Gardiner N, Singh SJ. Early experiences of rehabilitation for individuals post-COVID to improve fatigue, breathlessness exercise capacity and cognition - a cohort study. Chron Respir Dis 2021;18:14799731211015691. [doi: 10.1177/14799731211015691] [Medline: 33957805]
22. Palladini M, Bravi B, Colombo F, et al. Cognitive remediation therapy for post-acute persistent cognitive deficits in COVID-19 survivors: a proof-of-concept study. Neuropsychol Rehabil 2023 Aug;33(7):1207-1224. [doi: 10.1080/09602011.2022.2075016] [Medline: 35583357]
23. Maggio MG, De Luca R, Manuli A, Calabrò RS. The five “W” of cognitive telerehabilitation in the Covid-19 era. Expert Rev Med Devices 2020 Jun;17(6):473-475. [doi: 10.1080/17434440.2020.1776607] [Medline: 32476504]
24. Clinical management of COVID-19: living guidance. : World Health Organization; 2022 URL: https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2022-1 [accessed 2022-09-15]
25. Casson S, Jones MD, Cassar J, et al. The effectiveness of activity pacing interventions for people with chronic fatigue syndrome: a systematic review and meta-analysis. Disabil Rehabil 2023 Nov;45(23):3788-3802. [doi: 10.1080/09638288.2022.2135776] [Medline: 36345726]
26. Myalgic encephalomyelitis (or encephalopathy)/chronic fatigue syndrome: diagnosis and management. National Institute for Health and Care Excellence. 2021. URL: https://www.nice.org.uk/guidance/ng206 [accessed 2025-12-23]
27. Appelman B, Charlton BT, Goulding RP, et al. Muscle abnormalities worsen after post-exertional malaise in long COVID. Nat Commun 2024 Jan 4;15(1):17. [doi: 10.1038/s41467-023-44432-3] [Medline: 38177128]
28. Davis HE, Assaf GS, McCorkell L, et al. Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. EClinicalMedicine 2021 Aug;38:101019. [doi: 10.1016/j.eclinm.2021.101019] [Medline: 34308300]
29. Tryfonos A, Pourhamidi K, Jörnåker G, et al. Functional limitations and exercise intolerance in patients with post-COVID condition: a randomized crossover clinical trial. JAMA Netw Open 2024 Apr 1;7(4):e244386. [doi: 10.1001/jamanetworkopen.2024.4386] [Medline: 38573638]
30. Herrera JE, Niehaus WN, Whiteson J, et al. Multidisciplinary collaborative consensus guidance statement on the assessment and treatment of fatigue in postacute sequelae of SARS-CoV-2 infection (PASC) patients. PM R 2021 Sep;13(9):1027-1043. [doi: 10.1002/pmrj.12684] [Medline: 34346558]
31. COVID-19 rapid guideline: managing the long-term effects of COVID-19. National Institute for Health and Care Exellence. 2024. URL: https://www.nice.org.uk/guidance/NG188 [accessed 2025-12-23]
32. Nopp S, Moik F, Klok FA, et al. Outpatient pulmonary rehabilitation in patients with long COVID improves exercise capacity, functional status, dyspnea, fatigue, and quality of life. Respiration 2022;101(6):593-601. [doi: 10.1159/000522118] [Medline: 35203084]

XSL•FO
RenderX

33. Pinxsterhuis I, Sandvik L, Strand EB, Bautz-Holter E, Sveen U. Effectiveness of a group-based self-management program for people with chronic fatigue syndrome: a randomized controlled trial. Clin Rehabil 2017 Jan;31(1):93-103. [doi: 10.1177/0269215515621362] [Medline: 26672998]
34. Al-Jabr H, Windle K, Clifton A, Thompson DR, Castle DJ, Ski CF. Patient experiences of the long COVID-optimal health programme: a qualitative interview study in community settings. BJGP Open 2024 Apr;8(1):37907337. [doi: 10.3399/BJGPO.2023.0137] [Medline: 37907337]
35. Oliveira MR, Hoffman M, Jones AW, Holland AE, Borghi-Silva A. Effect of pulmonary rehabilitation on exercise capacity, dyspnea, fatigue, and peripheral muscle strength in patients with post-COVID-19 syndrome: a systematic review and meta-analysis. Arch Phys Med Rehabil 2024 Aug;105(8):1559-1570. [doi: 10.1016/j.apmr.2024.01.007] [Medline: 38311096]
36. Valverde-Martínez M, López-Liria R, Martínez-Cal J, Benzo-Iglesias MJ, Torres-Álamo L, Rocamora-Pérez P. Telerehabilitation, a viable option in patients with persistent post-COVID syndrome: a systematic review. Healthcare (Basel) 2023 Jan 7;11(2):187. [doi: 10.3390/healthcare11020187] [Medline: 36673555]
37. Espinoza-Bravo C, Arnal-Gómez A, Martínez-Arnau FM, et al. Effectiveness of functional or aerobic exercise combined with breathing techniques in telerehabilitation for patients with long COVID: a randomized controlled trial. Phys Ther 2023 Nov 4;103(11):pzad118. [doi: 10.1093/ptj/pzad118] [Medline: 37658773]
38. Frisk B, Jürgensen M, Espehaug B, et al. A safe and effective micro-choice based rehabilitation for patients with long COVID: results from a quasi-experimental study. Sci Rep 2023 Jun 9;13(1):9423. [doi: 10.1038/s41598-023-35991-y] [Medline: 37296140]
39. Samper-Pardo M, León-Herrera S, Oliván-Blázquez B, Méndez-López F, Domínguez-García M, Sánchez-Recio R. Effectiveness of a telerehabilitation intervention using ReCOVery APP of long COVID patients: a randomized, 3-month follow-up clinical trial. Sci Rep 2023 May 16;13(1):7943. [doi: 10.1038/s41598-023-35058-y] [Medline: 37193738]
40. Bileviciute-Ljungar I, Apelman A, Braconier L, Östhols S, Norrefalk JR, Borg K. A first randomized eight-week multidisciplinary telerehabilitation study for the post-COVID-19 condition: improvements in health- and pain-related parameters. J Clin Med 2025 Jan 14;14(2):486. [doi: 10.3390/jcm14020486] [Medline: 39860492]
41. Chandan JS, Brown KR, Simms-Williams N, et al. Non-pharmacological therapies for post-viral syndromes, including long COVID: a systematic review. Int J Environ Res Public Health 2023 Feb 16;20(4):3477. [doi: 10.3390/ijerph20043477] [Medline: 36834176]
42. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
43. EuroQol G. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9] [Medline: 10109801]
44. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. J Psychosom Res 1993;37(2):147-153. [doi: 10.1016/0022-3999(93)90081-p] [Medline: 8463991]
45. Stricker NH, Christianson TJ, Lundt ES, et al. Mayo normative studies: regression-based normative data for the auditory verbal learning test for ages 30-91 years and the importance of adjusting for sex. J Int Neuropsychol Soc 2021 Mar;27(3):211-226. [doi: 10.1017/S1355617720000752] [Medline: 32815494]
46. Espenes J, Eliassen IV, Öhman F, et al. Regression‐based normative data for the rey auditory verbal learning test in Norwegian and Swedish adults aged 49–79 and comparison with published norms. Clin Neuropsychol 2023 Aug;37(6):1276-1301. [doi: 10.1080/13854046.2022.2106890] [Medline: 35968846]
47. Öhman F, Eckerström M, Hessen E, et al. Demographically adjusted rey–osterrieth complex figure test norms in a Swedish and Norwegian cohort aged 49–77 years and comparison with north american norms. Scand J Psychol 2024 Apr;65(2):168-178. [doi: 10.1111/sjop.12966] [Medline: 37721999]
48. Wechsler D. Wechsler Adult Intelligence Scale--Third Edition. APA PsycTests. 1997. URL: https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft49755-000 [accessed 2025-12-23]
49. Delis D, Kaplan E, Kramer J. Delis-Kaplan Executive Function System (DKEFS). Scientific Research. 2001. URL: https://www.scirp.org/reference/referencespapers?referenceid=2055045 [accessed 2025-12-23]
50. Espenes J, Lorentzen IM, Eliassen IV, et al. Regression-based normative data for the D-KEFS Color-Word Interference Test in Norwegian adults ages 20–85. Clin Neuropsychol 2024 Jul 3;38(5):1227-1255. [doi: 10.1080/13854046.2023.2276967]
51. Conners CK. Conners Continuous Performance Test 3rd Edition (Conners CPT 3). ePROVIDE. 2000. URL: https://eprovide.mapi-trust.org/instruments/conners-continuous-performance-test-3rd-edition [accessed 2025-12-23]
52. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
53. Del Corral T, Fabero-Garrido R, Plaza-Manzano G, Navarro-Santana MJ, Fernández-de-Las-Peñas C, López-de-Uralde-Villanueva I. Minimal clinically important differences in EQ-5D-5L index and VAS after a respiratory muscle training program in individuals experiencing long-term post-COVID-19 symptoms. Biomedicines 2023 Sep 13;11(9):37760964. [doi: 10.3390/biomedicines11092522] [Medline: 37760964]
54. Wang J, Goodfellow H, Walker S, et al. Trajectories of functional limitations, health-related quality of life and societal costs in individuals with long COVID: a population-based longitudinal cohort study. BMJ Open 2024 Nov;14(11):e088538. [doi: 10.1136/bmjopen-2024-088538] [Medline: 39537389]

XSL•FO
RenderX

55. Garratt AM, Stavem K, Shaw JW, Rand K. EQ-5D-5L value set for Norway: a hybrid model using cTTO and DCE data. Qual Life Res 2025 Feb;34(2):417-427. [doi: 10.1007/s11136-024-03837-3] [Medline: 39565555]
56. Brooks R. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72. [doi: 10.1016/0168-8510(96)00822-6] [Medline: 10158943]
57. Blomberg B, Mohn KGI, Brokstad KA, et al. Long COVID in a prospective cohort of home-isolated patients. Nat Med 2021 Sep;27(9):1607-1613. [doi: 10.1038/s41591-021-01433-3] [Medline: 34163090]
58. Stavem K, Ghanima W, Olsen MK, Gilboe HM, Einvik G. Prevalence and determinants of fatigue after COVID-19 in non-hospitalized subjects: a population-based study. Int J Environ Res Public Health 2021 Feb 19;18(4):2030. [doi: 10.3390/ijerph18042030] [Medline: 33669714]
59. Ertesvåg NU, Iversen A, Blomberg B, et al. Post COVID-19 condition after delta infection and omicron reinfection in children and adolescents. EBioMedicine 2023 Jun;92(104599):104599. [doi: 10.1016/j.ebiom.2023.104599] [Medline: 37149931]
60. Nolan CM, Longworth L, Lord J, et al. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. Thorax 2016 Jun;71(6):493-500. [doi: 10.1136/thoraxjnl-2015-207782] [Medline: 27030578]
61. Chen P, Lin KC, Liing RJ, Wu CY, Chen CL, Chang KC. Validity, responsiveness, and minimal clinically important difference of EQ-5D-5L in stroke patients undergoing rehabilitation. Qual Life Res 2016 Jun;25(6):1585-1596. [doi: 10.1007/s11136-015-1196-z] [Medline: 26714699]
62. Longobardi I, Goessler K, de Oliveira Júnior GN, et al. Effects of a 16-week home-based exercise training programme on health-related quality of life, functional capacity, and persistent symptoms in survivors of severe/critical COVID-19: a randomised controlled trial. Br J Sports Med 2023 Oct;57(20):1295-1303. [doi: 10.1136/bjsports-2022-106681] [Medline: 37164620]
63. Suthanawarakul B, Promma N, Iampinyo P, Saraboon C, Wattanaprateep J, Waongenngarm P. Effects of physical therapy via telerehabilitation on cardiopulmonary, physical and psychological function in patients with coronavirus disease 2019: a randomised controlled trial. J Telemed Telecare 2024 Dec 23;23:39716851. [doi: 10.1177/1357633X241303804] [Medline: 39716851]
64. Kiselica AM, Karr JE, Mikula CM, et al. Recent advances in neuropsychological test interpretation for clinical practice. Neuropsychol Rev 2024 Jun;34(2):637-667. [doi: 10.1007/s11065-023-09596-1] [Medline: 37594687]
65. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
66. Carlile O, Briggs A, Henderson AD, et al. Impact of long COVID on health-related quality-of-life: an openSAFELY population cohort study using patient-reported outcome measures (OpenPROMPT). Lancet Reg Health Eur 2024 May;40:100908. [doi: 10.1016/j.lanepe.2024.100908] [Medline: 38689605]

Abbreviations

CBT: cognitive behavioral therapy
CFS: chronic fatigue syndrome
HRQoL: health-related quality of life
ME: myalgic encephalomyelitis
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
UNN: University Hospital of North Norway
WHO: World Health Organization

Edited by J Sarvestan; submitted 19.Mar.2025; peer-reviewed by K Windle, P Zeier; accepted 10.Nov.2025; published 05.Jan.2026.

Please cite as:
Wilson M, Pedersen SG, Langeland N, Cox RJ, Aukrust P, Dahl TB, Sandvig A, Wilhelmsen M
Tailored Individual Follow-Ups Versus a One-Day Group Course in Patients With Long COVID (Post– COVID-19 Condition): Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e74113
URL: https://www.researchprotocols.org/2026/1/e74113
doi:10.2196/74113

© Marte Wilson, Synne Garder Pedersen, Nina Langeland, Rebecca Jane Cox, Pål Aukrust, Tuva B Dahl, Axel Sandvig, Maja Wilhelmsen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Jan.2026. This is an

XSL•FO
RenderX

open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Improving Outpatient Psychotherapy for Adults With Major Depressive and Anxiety Disorders Using Web-Based High-Frequency Monitoring and Feedback in Autosystemic Hypnotherapy: Protocol for a Two-Arm ABAB Crossed-Therapist Randomized Clinical Implementation Trial

Günter Schiepek[1,2], Prof Dr; Stephanie Wackernagel[2], MSc, Dipl-Des

[1]Institute of Synergetics and Psychotherapy Research, Paracelsus Medical University, Salzburg, Austria

[2]Faculty of Psychology and Education, Ludwig-Maximilians-Universität München, Munich, Germany

Corresponding Author:
Stephanie Wackernagel, MSc, Dipl-Des
Faculty of Psychology and Education
Ludwig-Maximilians-Universität München
Leopoldstraße 13
Munich, 80802
Germany
Phone: 49 15141620226
Email: s.wackernagel@campus.lmu.de

Abstract

Background: In recent years, routine outcome monitoring has been increasingly complemented by routine process monitoring in psychotherapy and other health care settings. Various approaches to therapy feedback exist, differing in assessment frequency, integration into the therapeutic process, and degree of personalization. In this study, we will use a procedure of high-frequency assessment through daily self-ratings, a standard process questionnaire, alongside a personalized questionnaire derived from case formulation, and frequent feedback interviews using visual diagrams to mirror the ongoing therapeutic processes.

Objective: This study aims to investigate the effectiveness of combining routine process monitoring with hypno-psychotherapy (autosystemic hypnotherapy) by comparing it to autosystemic hypnotherapy without process feedback in the outpatient treatment of mood disorders. It also seeks to examine process-outcome relationships and mechanisms of change through high-frequency self-assessments and session-based feedback.

Methods: This study is a randomized controlled trial with 2 arms, using within-therapist randomization (ABAB design) in outpatient psychotherapy. Participants are recruited offline via routine intake procedures. A total of 100 patients will be randomly assigned to one of the two conditions following a waiting period. The inclusion criterion is the existence of any mood disorder (major depressive disorder or anxiety disorder), assessed via a clinical interview. Each therapist treats patients in both conditions. Outcomes will be measured at 4 time points: after diagnosis confirmation, postwaiting period, posttreatment, and a 6-month follow-up. Primary and secondary outcomes, including symptom severity, will be assessed using questionnaires. Data collection also includes patient and therapist session evaluations using the Bern Patient and Therapist Session Questionnaire. In the feedback condition, therapists conduct frequent interviews using time-series data generated from daily self-assessments using the synergetic navigation system, including the Therapy Process Questionnaire and an individualized measure based on case conceptualization.

Results: While this study is ongoing, the primary aim is to assess the effects of the feedback condition on therapeutic outcomes, including symptom reduction and patient motivation. This study will also explore how dynamic monitoring and feedback influence the therapeutic alliance and session-level improvements. It is expected that the feedback condition will lead to improvements in symptom severity and therapeutic engagement compared to the nonfeedback condition. Recruitment is ongoing, with 22 participants enrolled. The training of therapists and the data collection began in 2022. Data collection will end and study findings will be published in 2027. The German Society for Auto-Systemic Hypnotherapy is funding these training courses.

Conclusions: This study combines effect and process measures within a feedback condition, compared to a nonfeedback condition. It incorporates dynamic process assessment to explore change mechanisms by analyzing patterns of time-series data

XSL•FO
RenderX

and session ratings by patients and therapists. The approach provides insights into how continuous feedback and tailored monitoring influence therapeutic progress and outcomes.

Trial Registration: OSF Registries osf.io/z2efy; https://doi.org/10.17605/OSF.IO/Z2EFY

International Registered Report Identifier (IRRID): DERR1-10.2196/78166

(JMIR Res Protoc 2026;15:e78166) doi:10.2196/78166

KEYWORDS

major depressive disorder; anxiety disorders; outpatient psychotherapy; autosystemic hypnotherapy; randomized clinical trial; routine process monitoring; high-frequency monitoring; synergetic navigation system; therapist effects; feedback effects; mobile phone

Introduction

Effects of Therapy Monitoring and Feedback

For many years, monitoring and feedback procedures have been implemented in mental health care and in psychotherapy. Some researchers and practitioners identify these procedures as part of "good practice" [1] and even as the "most significant innovation in therapy practice within the past 20 years" [2]. Studies (case studies, implementation reports, and randomized controlled trials [RCTs]) tried to investigate the effects of monitoring and feedback in different mental health settings, applied to the treatment of many different diagnoses. Meta-analyses report small to medium overall effect sizes and showed that the feedback significantly reduces the odds of treatment dropout [3,4]. With respect to the effects, it has to be noted that the methodological quality of the studies is different and often poor [5], for example, pragmatic trials were realized with outcome assessments which were not independent of the process measures [6], and studies often are statistically underpowered (small, expected effect sizes require larger sample sizes) [7]. Other criteria concern the proof of using the feedback as intended—such as the manipulation check in an experiment—which, in case of not being taken for granted, may reduce the effect size of interventions compared to treatment as usual [8], and the implemented spectrum of outcome and progress measures [9].

To obtain a deeper insight into the mechanisms of feedback procedures, it should be considered that there are different approaches to realize it. Depending on the paradigmatic frame of the procedure, which can be linear and standardized or nonlinear [10,11], there may be different mechanisms of intervening and creating effects. The linear approach usually focuses on standard tracks or expected treatment responses [12], based on infrequent measures (eg, session by session). The nonlinear frame is based on the conceptualization of psychotherapy as a self-organizing and chaotic process [10,11] and does not presume any assumptions on the shape of trajectories [13]. Presuming standard tracks, pathways of clinical change are expected to run on these tracks, with deviations being at risk of deteriorations or unfavorable developments [14,15]. In case of "not on track" courses, clinical support tools are offered [16]. Different from this, nonlinear monitoring and feedback is based on more frequent and equidistant measures (eg, daily assessments) and, in consequence, on continuous feedback interviewing and process reflection (eg, in every session or every second session). Sampling rates with frequent and equidistant measures allow for different options of analyzing the pathways (time series) of change [17]. No specific incidents (eg, "not on track" states) are required for executing feedback talks. A focus of interest is on occurring critical instabilities, which are valued as precursors of upcoming pattern transitions [18].

Different from high-frequency process monitoring, the so-called routine outcome monitoring [19] focuses on outcome criteria, which are usually given by standard outcome questionnaires such as Outcome Questionnaire-45.2, Symptom Checklist-90-Revised, Depression Anxiety Stress Scales, or Beck Depression Inventory. The sampling rate is low, for example, each month or lower. In consequence, no real time series can be created. In contrast, the process monitoring approach as applied in this study includes a spectrum of common factors that have been proven to be important for creating successful treatments [20]. A Therapy Process Questionnaire (TPQ) was developed for daily self-assessments, which are not restricted to in-session evaluation but can be carried out in the real-world setting of the patients (eco-systemic approach). It includes factors as "well-being and positive emotions," "emotional and problem intensity," "motivation for change," "insight/confidence/therapeutic progress," and "mindfulness/self-care" [21].

Another difference exists between standardized and personalized procedures, with standardized approaches using predefined questionnaires, whereas personalized measures usually are based on case formulations as the idiographic system modeling and can be implemented by a questionnaire editor, as it is available in the synergetic navigation system (SNS) [22]. As practical experience shows, compliance rates can be optimized by personalized approaches of monitoring and feedback because they focus on patients' own goals and topics [23]. The items of the questionnaires in use are created by the patients themselves, improving the experienced meaningfulness and thereby, the validity of the measures.

In consequence, there may be different mechanisms of creating effects and benefits by different approaches. Complex high-frequency procedures need intense training for the professionals, closer contact with the patients, a higher frequency of feedback sessions, and a higher degree of personal involvement of the patients, with corresponding effects on emotions, self-experience, and behavior by daily self-assessments ("measurement is intervention"). Clinicians also report on different attitudes to critical periods (instabilities)

XSL•FO
RenderX

which are more desirable and less threatening if you see psychotherapy as a self-organizing process running through cascades of pattern transitions, and open-mindedness to different patterns of change. Other mechanisms may be more similar between different approaches, for example, that feedback may provide new information (though it may be various types of information that seem to be important), that feedback enhances patient-clinician communication (though it may focus on different topics), or that it may improve the therapeutic alliance [24]. Successful strategies of implementation into the respective setting of application, training procedures with different contents and messages concerning linear or nonlinear processes, and the clinician's attitude toward the usefulness of the procedure seem to be important [4].

The research question of the study presented in this paper focuses on the effects of personalized high-frequency feedback based on case formulation (idiographic system modeling) in the context of a nonlinear dynamic systems approach combined with a hypno-psychotherapy called autosystemic hypnotherapy (ASH; original German term: Autosystemhypnose), compared with treatment as usual, in this case, ASH only.

Hypnotherapy and the Concept of ASH

Hypnotherapy is an evidence-based treatment method that has gained increasing recognition in psychological and medical settings. Its effectiveness is particularly well established in the treatment of anxiety disorders (ADs): a meta-analysis of 17 studies (N=690) [25] reported medium to large effect sizes, showing that patients who received hypnosis for anxiety demonstrated a weighted mean effect size of 0.79 (95% CI 0.61-0.97) posttreatment, suggesting that the average hypnosis patient improved by more than 79% (based on Cohen *d* interpretation) compared to control participants. At follow-up (7 studies, N=215), the mean effect size was 0.99 (95% CI 0.74-1.24), reflecting an improvement greater than 84% of controls. Hypnotherapy has also proven highly effective for depression. In an RCT [26], hypnotherapy was found to be noninferior to cognitive behavioral therapy in treating mild to moderate depression. After 12 months, the response rate (defined as ≥50% symptom reduction) was 44.6% (33/74) of patients who responded to hypnotherapy (compared to 30/78, 38.5%, for cognitive behavioral therapy), indicating sustained therapeutic effects.

The therapeutic efficacy of hypnotherapy is based on its structured therapeutic process and its neurobiologically grounded mechanisms. Hypnotherapy typically follows four evidence-based phases [27]:

- Prehypnotic preparation: therapist and patient explore beliefs, previous experiences, and treatment goals (eg, "reducing panic attacks from daily to ≤1 per week") while clarifying misconceptions about hypnosis. This builds trust and aligns expectations.
- Hypnotic induction: using focusing techniques or verbal guidance, the therapist induces a state of focused and relaxed attention. Critical reasoning is bypassed, and a trance state is typically reached within 5-10 minutes.
- Therapeutic suggestions: this phase involves tailored therapeutic suggestions and metaphors designed to elicit emotional, psychological, or physiological changes aligned with the treatment goals. These use the brain's heightened suggestibility during trance to reduce symptoms and promote adaptive experiences.
- Posthypnotic integration: before returning to normal awareness, patients receive posthypnotic suggestions to reinforce and integrate change into daily life. They are encouraged to apply trance-derived insights to real-world situations.

Recent neuroimaging research has provided compelling evidence that hypnosis induces distinct and measurable changes in brain activity. Using multimodal imaging, the research team [28] demonstrated that neural state changes occur synchronously with specific hypnotic induction instructions, indicating that hypnosis produces targeted modulations of brain function rather than nonspecific relaxation effects. In clinical populations, hypnosis has been shown to markedly inhibit the neural response of fear-related circuits, supporting its therapeutic efficacy in anxiety regulation:

- Amygdala inhibition: functional magnetic resonance imaging studies show reduced activity in fear-processing centers during anxiety-focused hypnotherapy, correlating with symptom improvement.
- Suppression of the default mode network: trance reduces self-referential thought, helping patients disengage from rumination patterns.
- Enhanced top-down regulation: increased connectivity between prefrontal and limbic regions allows for improved cognitive restructuring of emotional responses.

This dual foundation—systematic procedural structure and targeted neuroplasticity—explains hypnotherapy's superiority over placebo in meta-analyses [25,26] and supports its role as a method for modulating emotional and cognitive processes through attentional and self-regulatory mechanisms.

ASH [29] is a variant of hypnotherapy that integrates self-organizing systems theory and synergetic principles with classical hypnotherapeutic elements. As in traditional hypnotherapy, the therapist uses induction and suggestion, but ASH specifically addresses the client's internal "autosystem" (self-regulatory system) and supports the development of self-observation and resource orientation. While ASH incorporates the established components of hypnotherapy (induction, suggestion, and integration), it adds additional distinctive features:

- Diagnostic precision through hypnotic analysis: virtual "system scans" during trance enable the identification of unconscious blocks (eg, a patient visualizing their depression as "a black hole sucking away motivation" reveals implicit helplessness schemas), as well as exploration of the adaptive function of symptoms (eg, chronic pain maintaining caregiver attention).
- Resolution of multilayered resistances: the therapist guides the patient from addressing superficial objections ("therapy takes too long") toward deeper levels of processing in which core conflicts ("if I recover, my partner will leave me") are uncovered and reframed through trance-based cognitive restructuring.

XSL•FO
RenderX

While ASH has an established theoretical foundation and clinical tradition in German-speaking countries [30], no controlled efficacy studies exist to date. Accordingly, the present project aims to evaluate the effectiveness of ASH within a controlled design, building upon the broader evidence base for hypnotherapy.

Major Depression and ADs

The planned study will apply feedback-informed psychotherapy in an outpatient setting to patients with major depressive disorder (MDD) and the spectrum of ADs, including MDD with anxiousness disorders. MDD is one of the most important mental health issues all over the world. It affects about 3.8% of the global population, or about 280 million people of all ages, with an increasing trend [31]. MDD is about 50% more common among women than among men. More than 700,000 people die due to depression-related suicide every year. MDD is characterized by a mental state of low mood and aversion to activity. It affects the thoughts, behavior, feelings, and sense of well-being of the person concerned. Experiences that would normally bring a person pleasure or joy give reduced pleasure or joy, and the afflicted person often experiences a loss of motivation or interest in many different activities. It may feature sadness, difficulty in thinking and concentration, or a significant increase or decrease in appetite and time spent sleeping. People experiencing MDD may have feelings of dejection or hopelessness and may experience suicidal thoughts. MDD is one of the most common comorbidities of many chronic diseases, including cancer and cardiovascular, metabolic, inflammatory, and neurological disorders. It is also a frequent comorbidity of other mental disorders, such as AD, borderline personality disorder, or obsessive-compulsive disorder. AD is the world's most common mental disorder, affecting 301 million people in 2019 [32]. People with an AD may experience excessive fear or worry about a specific situation (for example, a panic attack or social situation) or, in the case of generalized anxiety disorder (GAD), about a broad range of everyday situations. They typically experience these symptoms over an extended period, at least several months. Usually, they avoid situations that make them anxious. Other symptoms of AD may include trouble concentrating or making decisions, feeling irritable, tense, or restless, experiencing nausea or abdominal distress, heart palpitations, sweating, trembling or shaking, trouble sleeping, or a sense of impending danger, panic, or doom. AD increases the risk for depression and substance use disorders, as well as the risk of suicidal thoughts and behaviors.

There are different kinds of AD, including GAD, characterized by persistent and excessive worry about daily activities or events, panic disorder, social AD, specific phobias, and others. Psychological mechanisms of GAD illustrate the comorbidity between anxiety and depression, for example, uncontrollable worry, constituting maladaptive strategies such as avoiding anxiety-related situations and emotional states. Individuals enduring GAD show deficits in detecting and regulating emotional states, which might accelerate a positive feedback loop between general stress symptoms and pathological worrying [33]. Experiential avoidance might lead to a restriction in proactive behavior because individuals will be focused on preventing negative events and maintaining safety [34].

Therapist Effects in Randomized Clinical Trials

In this study, a randomized allocation of the treatment conditions will be realized among the therapists. Although blinded allocation of patients and health professionals is an important claim in double-blind randomized pharmacological trials, in evidence-based human interventions, for example, psychotherapy, health professionals cannot be blinded, because hopefully they are well-informed and fully aware of the interventions they apply [20,35]. This implies that personal preferences and the degree of identification with the treatment procedures that are conducted by the therapist (so-called allegiance effects) may affect the quality of the treatment as well as the outcome. However, therapist effects seem to be an important general (nonspecific) factor, which also will have consequences for the quality of processes and outcomes. Research on common factors revealed that therapist effects have bigger effects than treatment differences, in naturalistic designs, but even in RCTs, which usually are designed to minimize therapist effects (eg, by implementing treatment protocols) [20,36]. In consequence, crossed-therapist designs were suggested for naturalistic intervention studies [37,38], where a therapist is allocated to two or more treatment conditions, and therefore, potential differences in overall therapist effectiveness can be estimated across conditions. To estimate therapists' personal treatment preferences, which may impact their effectiveness [8,39], these preferences should explicitly be assessed [20,38,40].

Aims of This Study

This study addresses different research questions. There is some concern regarding the implementation of the procedure into the routine practice of outpatient psychotherapists, which is challenging because of the complexity of the routine process monitoring procedure. It encompasses an assessment of the patient's personal resources, a case formulation, and, based on this, the development of a personalized process questionnaire. Regular feedback interviews during the sessions are scheduled with reference to the visualized results. Other research questions focus on the outcome of this study's conditions and on the patterns of change that may characterize successful or unsuccessful therapies. Potential therapists' implementation effects will be examined based on an ABAB therapist allocation (see the Study Design section). The main research questions are as follows:

1. Practicability: Can high-frequency process monitoring (daily self-assessments) and feedback interviewing (at least every second session) be implemented in outpatient routine psychotherapy, that is, in naturalistic settings? To what extent is the routine practice of outpatient psychotherapy supported by process monitoring? What can be learned from this study for the implementation of process monitoring into practice and training?
2. Treatment outcomes: What is the efficacy of ASH compared with a waiting period control? What are the differences in the treatment efficacy of the two randomized implementation conditions, ASH versus ASH supported by personalized process monitoring and feedback in self-reported primary and secondary outcomes?

XSL•FO
RenderX

3. Session effects and change dynamics: What are the differences between the implemented conditions in the experienced quality of the sessions (eg, micro-outcome) and in the therapeutic alliance? Does the process feedback support the therapeutic alliance? With reference to the time series created by the daily self-assessments: Are there specific features of the change dynamics (eg, discontinuous pattern transitions and critical instabilities) related to the outcome?
4. Therapist effects: Are there any differences across therapists (therapist effects)? Is there an interaction between potential therapists' preferences for a particular implementation condition and outcome?

Methods

Study Design

This study is designed as an RCT with 2 arms, using within-therapist randomization in the real-world clinical practice of routine outpatient psychotherapy in Germany.

Participants are recruited offline through routine intake procedures when presenting for outpatient psychotherapy with participating therapists. Patients who are suspected by the therapist to meet criteria for MDD, AD, or MDD with anxious disorders complex will be briefly screened for their ability and willingness to use a smartphone. If eligible, they will be informed about this study and its requirements by the respective therapist and asked whether they are interested in participating in a research project. Afterward, they will be informed of the procedure and the requirements of this study by the respective therapist. They obtain a flyer with basic information and, in case of concrete interest, will be instructed in detail and obtain an explanation to its full extent.

Patients who have consented to participate in this study will undergo a structured diagnostic interview (Structured Clinical Interview for *DSM-5* [*Diagnostic and Statistical Manual of Mental Disorders* {Fifth Edition}] – Clinician Version) [41] to confirm the inclusion criteria (diagnoses as listed in the Participants section, see below). The interview will be conducted remotely by an independent, trained clinical psychologist who is otherwise not involved in this study. The Structured Clinical Interview for *DSM-5* allows a final decision on including or excluding the patient. In addition to diagnostic confirmation, baseline information on participants' current medication, concurrent treatment, and symptom severity will be systematically assessed. These variables will be considered as covariates in the main outcome analyses to account for potential prognostic influences. In addition, sensitivity analyses will be conducted to examine whether excluding participants with concurrent treatments or psychopharmacological medication changes affects the robustness of the findings. Following inclusion, a waiting period of 1 (minimum) to 6 months (maximum) begins for each patient, depending on the therapist's capacity to initiate treatment.

After the waiting period, a randomization procedure decides on the allocation of a patient to condition A (ASH) or B (ASH+synergetic process management [SPM]). To ensure methodological rigor and internal validity, patients will be randomly assigned to one of the two treatment conditions using therapist-specific block randomization. The predefined block-randomized allocation sequences (eg, ABAB, BABA, AABB, BBAA, ABBA, and BAAB) will be generated in advance by the principal investigator for each therapist to ensure a balanced distribution of conditions within therapists. Patients will be assigned to intervention conditions in the order of study inclusion according to the next position in the respective therapist's allocation sequence. The allocation sequence will remain concealed from the therapists until each new patient is assigned.

After the randomization, the treatment period starts, whose length depends on the concrete requirements of each patient and must be decided by the patient and therapist together (shared decision-making). In contrast to manualized treatments (adherence to protocol), this is a usual procedure in naturalistic settings. A follow-up assessment will be conducted 6 months after the end of therapy.

There are four assessment points: at the beginning of the waiting period (T1), at the start of the treatment (T2), at the end of the treatment (T3), and at follow-up (T4). The design of this study is illustrated in Figure 1.

Figure 1. Design of this study. ASH: autosystemic hypnotherapy; SPM: synergetic process management; T1: beginning of the waiting period; T2: at the start of the treatment; T3: at the end of the treatment; T4: at follow-up.

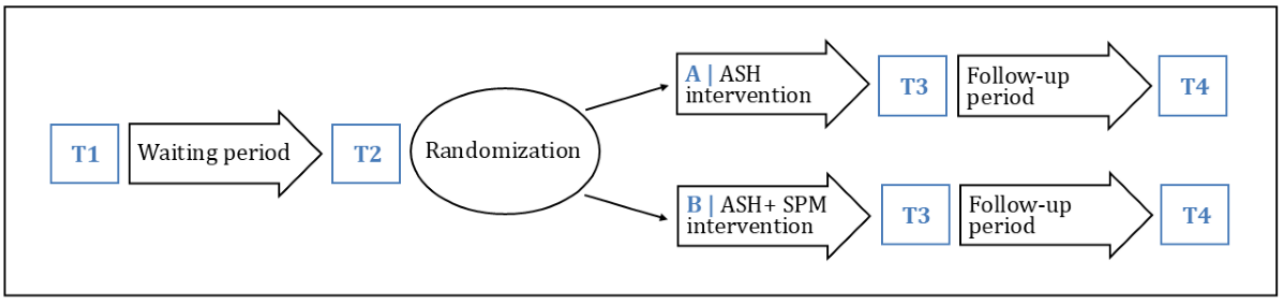

Ethical Considerations

This study received ethical approval from the Ethics Committee of the Faculty of Psychology and Educational Sciences at Ludwig Maximilian University, Munich (approval number: not specified in protocol; approval letter dated November 8, 2021). The use of the SNS was additionally approved by the Ethics Committee of the Province of Salzburg (No. 415-E/1068/3–2009).

XSL•FO
RenderX

All participants provided written informed consent before the intake assessment (T1). The consent form detailed the study's purpose, procedures, potential benefits/risks, and data anonymization. Participants were explicitly informed that (1) participation is voluntary and may be withdrawn at any time without consequence, (2) declining or withdrawing imposes no detriment to their clinical care, and (3) no financial compensation or gratuities are offered.

All collected data are subject to data protection regulations and will be published in anonymized form, ensuring participant confidentiality. These procedures are in accordance with the Declaration of Helsinki [42].

Participants

As this study involves self-assessments via a smartphone application of the SNS, participants are required to have basic smartphone and internet literacy. This is assessed informally during the initial instruction and onboarding process by the respective therapist.

The patients who will be included in this study should be diagnosed with a spectrum of depression and ADs as listed in the *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) [43], which is still valid in Germany. These are F32 depressive episode, F33 recurrent depressive disorder, F34 persistent mood disorders, F40 phobic ADs (eg, agoraphobia, social phobia, and specific phobias), F41 other ADs (eg, panic disorder, GAD, mixed anxiety and depressive disorder), F43 reaction to severe stress and adjustment disorders, and F45 somatoform disorders. At least one of these categories should be identified as the main diagnosis; other diagnoses (comorbidities) will be accepted. The reason for focusing on these diagnoses is their high prevalence rates and in consequently, the frequency of occurrence in routine outpatient psychotherapy. Exclusion criteria are aged younger than 18 years, acute suicidality, mental confusion, intoxication by drugs, substance use, and severe neurological diseases.

Treatment Protocol

This study comprises 2 conditions. One is ASH, which is a version of existing hypno-psychotherapy. It comprises exercises in imagination, activating inner parts (ego-states), and training in mentalization and relaxation. Exercises should induce altered states of consciousness, and triggered ideomotor movements should allow some direct "communication" with "the unconscious." Resources are activated by the imagination of symbolic persons, figures, or other objects, which might help to protect or support the person to realize changes in their life, to prepare decisions, or to accept and overcome traumatic experiences.

The other condition of this study combines ASH with an advanced method of personalized routine process monitoring, the SPM [44]. The SPM includes (1) an interview for activating personal resources and by this, induces a mental resource state; (2) a method of case conceptualization, which creates a qualitative network model of the most important psychological and social conditions (variables) of a problem [45,46]; and from this, (3) develops a personalized process questionnaire (each item of the questionnaire represents a variable of the model) in close cooperation between the patient and the therapist.

The monitoring is realized by the internet- and server-based device SNS, which collects, stores, analyzes, and visualizes the incoming data from the patient's smartphone [44]. Patients access the SNS as part of their therapeutic process, supported by their therapist, and use the app in their daily life outside the therapy sessions (ecological setting). No payment is required, and internet access is assumed through patients' own smartphones. Patients complete one self-assessment per day via the SNS app. They can activate optional reminder notifications in the SNS app to support adherence to daily self-assessments. However, there are no additional prompts such as emails, SMS text messaging, or phone calls from this study's team.

A final but important part of the SPM is (4) a regular feedback interview on the visualized data of the ongoing process. The feedback interviews take place in standard face-to-face therapy sessions. It can be based on several heuristics, which define important conditions for self-organized pattern transitions to be realized during the change process of a patient, the so-called "generic principles" [44]. The design and functionality of SPM and the SNS are based on the theoretical framework of nonlinear dynamic systems theory applied to psychotherapy. This approach conceptualizes therapeutic change as a self-organizing process characterized by dynamic instabilities, pattern transitions, and emergent reorganization. In the SNS, this framework is operationalized by the continuous self-assessment of individualized process variables, visualized as time series (see Figure 2 for a screenshot of the visualized data in SNS). The system does not rely on standardized change trajectories but allows for the identification of nonlinear phenomena such as phase transitions, critical fluctuations, and synchronization across process dimensions. These visualized patterns are reflected upon during regular feedback interviews and serve as the basis for process-oriented decision-making in therapy. The feedback is thus not triggered by deviations from a normative path (eg, "not on track") but by the recognition of system-dynamic signals emerging from the individual course of therapy. A recent case study illustrates how ASH with SPM can be integrated [47].

XSL•FO
RenderX

Figure 2. Screenshot of the visualized data in the device SNS. DASS-21: Depression Anxiety Stress Scales – 21 Items; SNS: synergetic navigation system; TPB-R: Therapy Process Questionnaire (German: Therapie-Prozessbogen-Revidiert).

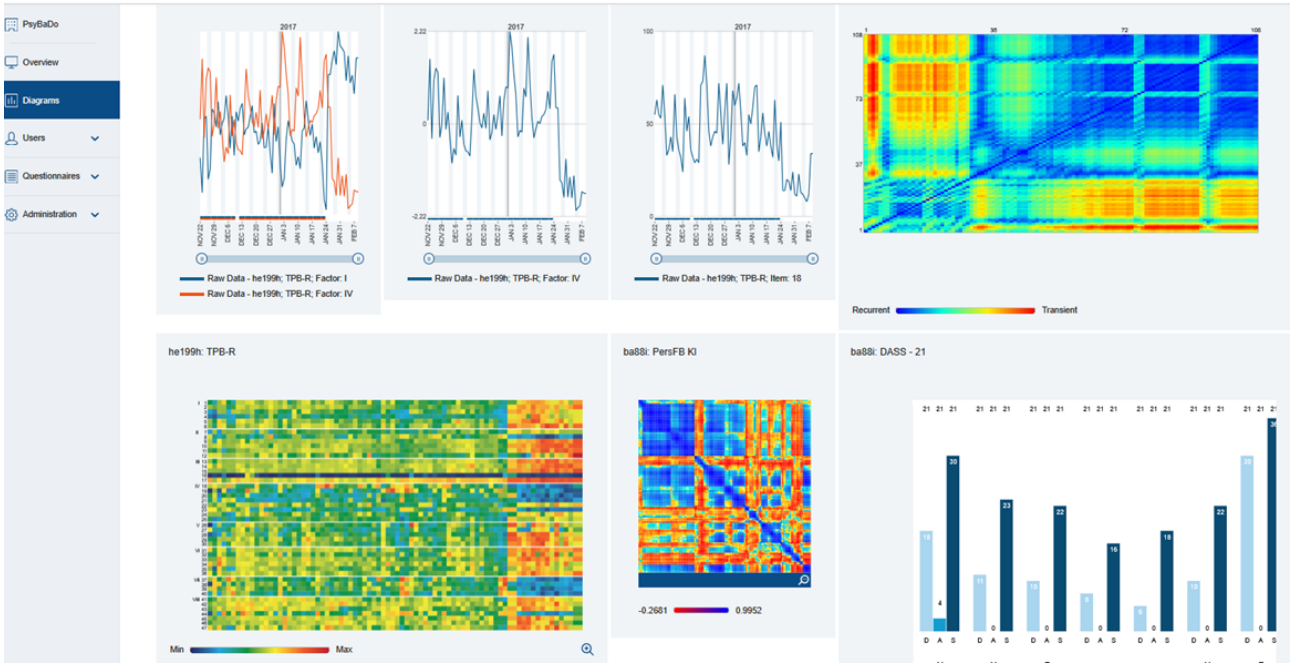

In the ASH+SPM condition, the personalized questionnaire of each patient will be combined with the standardized and validated TPQ [21], both administered at the same sampling rate (once per day).

Implementation Check

At the end of each session, therapists will estimate to what extent the features and criteria of the respective treatment condition were implemented in the current session. For this purpose, a list of criteria was compiled for each condition. Examples of the ASH criteria are “in this session we applied methods from ASH, like analysis of ambiguity and resistance,” “in this session we activated inner parts (problem or solution-related parts) and archetypes,” “in this session, we used symbolizations,” “in this session, we used metaphors,” “in this session, I supported the patient to apply autosystemic hypnotherapy techniques autonomously.” Examples of the SPM list are “in this session, we reflected the process by using its visualization in the SNS,” “this session was planned and realized with reference to the generic principles,” “in this session, the resources of the patient were important (eg, by conducting a resource interview),” or “in this session, we created a system model or referred to this.” Response options range from “not at all” to “very clearly.”

Beyond these kinds of session ratings by the therapists, regular supervision meetings will be offered to all therapists. These meetings are intended to support the therapists in their work, to reflect on the ongoing process, for example, on difficult situations during the treatments, such as lack of progress and breaks in the alliance, and to obtain an impression from a third-party point of view on which the criteria of the treatment conditions were fulfilled.

Therapists

All therapists to be included are trained and experienced in ASH, with a professional background in medicine or psychology. All are approbated physicians or psychotherapists according to the German psychotherapy law, with at least some years of practical experience in outpatient psychotherapy. For acquiring competencies in SPM, 8 online training modules on the components of the SPM will be performed, particularly idiographic system modeling and case conceptualization, developing personalized process questionnaires, feedback interviewing, technical handling of the SNS, and understanding and interpreting the diagrams as presented by the SNS. At least 15 therapists will participate in this study, each delivering both conditions.

Assessments

To evaluate treatment effectiveness, both primary and secondary outcome measures will be assessed at all four measurement points (T1-T4), with participants completing the assessments via online questionnaires in the SNS app outside the therapy sessions.

Primary outcomes focus on symptom reduction in depression and ADs and will be measured using (1) the short version of the Symptom Checklist – 14 Items [48]; and (2) the Depression Anxiety Stress Scales – 21 Items, short version [49,50].

Secondary outcomes include measures of psychosocial functioning, emotion regulation, and quality of life: (1) the General Self-Efficacy Scale [51], (2) the Intercultural Quality of Life Comic [52], and (3) the Difficulties in Emotion Regulation Scale [53,54].

Process and moderator variables will be collected throughout the intervention.

At T2, the motivation for psychotherapy will be measured via a psychotherapy motivation questionnaire (Fragebogen zur Psychotherapiemotivation – 23 Items, FPTM-23) [55] to check for bias between treatment conditions. After each therapy session, the Bern Patient and Therapist Session Questionnaire

XSL•FO
RenderX

[56] will assess therapeutic alliance and session quality (therapist and patient reports).

By this, potentially important mediating and moderating variables of the treatment conditions should be incorporated in the analysis, for example, the course of the alliance, the session-by-session micro-outcome, motivation for change, activation of resources, and others.

Additionally, the tool SNS will be used for personalized feedback and monitoring in the ASH+SPM condition, enabling daily assessment of intraindividual change processes based on the TPQ for outpatient psychotherapy [21] in combination with a personalized process questionnaire developed individually with each patient.

Power Considerations

The estimation of the sample size by power analysis refers to two research questions: (1) Is there a difference between the waiting period and the treatment? (2) Is there a difference between the two conditions of the treatment? The statistical procedure for answering these questions is ANOVA with repeated measures, with the groups (conditions) as the between-subject factor and the comparison of the waiting period (T1-T2, level 1) with the overall treatment (T2-T3, level 2) as the within-subject factor. The power calculation for estimating the sample size was realized by the software package G*Power (Heinrich-Heine-Universität Düsseldorf) [57].

The power estimation of the within-subjects calculation (waiting period versus overall treatment, independent of the condition) was based on the reported average effect size of psychotherapy outcomes from studies included in the meta-analysis (Cohen *d*=0.68) [58]. Using an effect size converter, this d value was transformed into Cohen f (f=0.34), which is needed for calculations in G*Power. The power was defined at 0.95, and the correlation between the measurement points was left at a default value of *r*=0.50. Sphericity was assumed. For two measurement points within the sample, it was estimated that the null hypothesis of no difference (no change) could be rejected with a probability of 96% if a sample size of 32 patients is included.

For the power calculation on the effects of the conditions (between-group comparison), a lower threshold value of the mean effect size was supposed at f=0.25, because of the lack of data in this case. The power was defined at 0.80. All other values and assumptions were the same for the between-subject factor as for the intersubject factor. It was estimated that the null hypothesis of no difference (no change) could be rejected with a probability of 80% if a sample size of 98 patients (49 in each condition) were included.

Based on the estimated sample sizes of 32 patients (hypothesis 1) and 98 patients (hypothesis 2), we defined the overall sample size of this study as 100 patients (n=50 in condition A: ASH; n=50 in condition B: ASH+SPM). The allocation ratio is 1:1 with 50 patients per arm.

Statistical Analysis

The statistical procedure for answering the outcome-related questions of this study is ANOVA with repeated measures, with the groups (conditions) as the between-subject factor and the comparison of the waiting period (T1-T2, level 1) with the overall treatment (T2-T3, level 2) as the within-subject factor. The preference ratings of the therapists in the ASH+SPM condition can be used as covariates in a 1-factor ANOVA with repeated measures.

The therapy effects can be predicted by linear regression models and multilevel regression models with conditions, therapy motivation (patients), alliance rates (patients), and the quality of the implementation as predictors. In the ASH+SPM condition, the preference of implementing the SPM (therapists) and the compliance rates (patients) of the process monitoring can also be used as predictors.

The research questions on the dynamics of change (eg, Are pattern transitions related to the outcome?) will be analyzed by applying the Pattern Transition Detection Algorithm [59] and methods of recurrence quantification analysis [60,61]. The respective quantifiers will be correlated with the outcome measures (T2-T3 and T2-T4).

Results

Although the trial is ongoing, recruitment has already begun. As of the time of submission, 22 participants have been enrolled and have completed the diagnostic eligibility assessment. Data collection started in 2022 following the training of participating therapists in the use of the SNS, which was funded by the German Society for Autosystemic Hypnotherapy in 2022 and 2025. Data collection will continue until all follow-up assessments (T4) are completed, which is projected for the first quarter of 2027. No outcome analyses have been conducted to date. Study findings are expected to be disseminated in the second quarter of 2027.

Discussion

Principal Findings

This study tries to combine the strictness of an RCT with the ecological validity of routine care (outpatient settings). The aims of this study focus on the effectiveness of ASH and of ASH improved by personalized high-frequency monitoring and feedback, which was technologically implemented by the SNS. This kind of comparison has never been conducted in an RCT under real-world conditions, and the combination of outcome and process research from the perspective of nonlinear complex systems [10,62] is an innovative step in psychotherapy research.

Besides the frequency of patients' self-assessments, the frequency and quality of feedback sessions (participative interviewing and shared decision-making) also matter. Studies have shown that information from feedback is often not used in clinical practice. Although routine outcome monitoring is mandatory in many countries, measuring outcomes (or even processes) does not necessarily mean that this information is being used to inform or to advance the treatment. Even in RCTs, up to half of the clinicians do not use progress feedback to inform treatment decisions [8,63]. In a specific RCT [64], one third of the clinicians reportedly never viewed any progress feedback reports. In a content analysis of written notes [65],

XSL•FO
RenderX

clinicians included in the study incorporated feedback data in only about one-fifth of the sessions. A review addressing the barriers to implementing precision methods as feedback in personalizing psychological therapies, and possible solutions, has been published recently [66]. This study investigates an optimal way of implementing feedback in clinical practice (outpatient therapy) and integrating it into treatment as usual.

The combined structure of this study's design should allow differentiation between patient effects, therapist effects, and different degrees of preferences for the feedback method (SPM). In contrast to pharmaceutical trials, patients and therapists know what they should do and are proactively involved in the treatment. This involvement is not a bias that has to be eliminated; it might rather be an active ingredient of a successful psychotherapy in which the therapists and the patients take a proactive and collaborative role in the treatment plan.

Bias Minimization

At the level of patients and therapists, biases should be minimized. Patients will be randomly assigned to conditions to reduce systematic selection biases in participant characteristics. The inclusion and exclusion criteria allow for a relatively homogeneous group of individuals with depression-anxiety spectrum diagnoses, and by measuring the psychotherapy motivation at the beginning of the treatment period (T2), motivation biases can be considered (statistically balanced). At the therapist level, a potential bias due to the therapist's preferences is a concern because of outcome expectations, which can be impactful for patients and for therapists. In the proposed design, potential therapist effects can be considered (statistically balanced), and the present ABAB crossed therapist design can consider the preferences of realizing the SPM procedure.

Limitations

Despite all the measures taken to optimize the design and minimize biases, there are some limitations due to the naturalistic setting of this study. On the one hand, there are differences in the waiting time, which can be statistically taken into account as covariates in the ANOVA, but cannot be eliminated. The same applies to therapist effects, which can be regarded as a confounding factor, but are not completely eliminated. One reason for this is that it is not yet clear which psychological mechanisms are behind the effects of the therapists, which are more important even in controlled studies than the differences between the treatment conditions [20,36]. Another limitation concerns the implementation check, which is given by self-assessments of the therapists after each session, not by video-based coding of the sessions by independent and blinded raters. Independent session ratings of the quality and the accuracy of the conditions should be realized in future studies, because this can be supposed to have an impact on the outcome, especially in feedback-based therapy sessions.

Strengths

This study will have a number of strengths. One is the combined process and outcome research, with outcome measures at 4 time points (T1-T4, see the Study Design section) and process monitoring at 2 different sampling rates: session by session with the Bern Patient and Therapist Session Questionnaire in both conditions and high-frequency equidistant time sampling (day by day) in the ASH+SPM condition. Given the 2 sampling rates, the course of micro-outcomes and alliance rates can be aligned (synchronized) with the nonlinear features of change dynamics (eg, critical instabilities, pattern transitions, and within-person synchronization) as they can be derived from the high-frequency time series data. The combination of outcome and process research is still unusual and innovative in RCTs.

Another strength is the strictness of an RCT design implemented in a naturalistic setting, which improves the ecological validity of this study and allows for the transfer to similar settings in outpatient psychotherapy. Third, we will recruit experienced therapists trained in both ASH and SPM (including process monitoring and feedback). As a result, both conditions are performed highly competently and qualify as bona fide therapies. Fourth, we will conduct an implementation review in both study conditions (with the restrictions as noted in the Limitations section). Fifth, in addition to the hypotheses about differential treatment effects (see the Aims of This Study section), mechanisms of change can also be examined.

Acknowledgments

We are grateful to all patients and to the therapists who will participate in this study. The power estimation was realized by Dr Kathrin Viol. The conception of this study and the recruitment of therapists were supported by Viola Hoffknecht, MSc. We are grateful for the support of Marc-Alexander Bäumer, BSc, in data management and therapist training.

Funding

There is no external funding for this study itself. The therapists' training in using the synergetic navigation system (SNS), as well as the usage fees for the SNS device, were funded by the Deutsche Gesellschaft für Autosystemhypnose (German Society for Autosystemic Hypnotherapy).

Authors' Contributions

Conceptualization: GS
Data curation: SW
Formal analysis: SW
Funding acquisition: GS

XSL•FO
RenderX

Investigation: GS, SW
Methodology: GS, SW
Project administration: SW
Resources: GS
Supervision: GS
Visualization: GS, SW
Writing – original draft: GS, SW
Writing – review & editing: SW

Conflicts of Interest

GS and his team at the Center for Complex Systems (CCSys GmbH) developed and provided the SNS for this study.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V1.6).
[PDF File (Adobe PDF File), 10150 KB - resprot_v15i1e78166_app1.pdf]

Multimedia Appendix 2
Informed consent form.
[PDF File (Adobe PDF File), 249 KB - resprot_v15i1e78166_app2.pdf]

References

1. Lambert MJ. Maximizing psychotherapy outcome beyond evidence-based medicine. Psychother Psychosom 2017;86(2):80-89. [doi: 10.1159/000455170] [Medline: 28183083]
2. McLeod J. Science and psychotherapy: developing research-based knowledge that enhances the effectiveness of practice. Trans Anal J 2017;47(2):82-101. [doi: 10.1177/0362153717694885]
3. de Jong K, Conijn JM, Gallagher RA, Reshetnikova AS, Heij M, Lutz MC. Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: a multilevel meta-analysis. Clin Psychol Rev 2021;85:102002 [FREE Full text] [doi: 10.1016/j.cpr.2021.102002] [Medline: 33721605]
4. de Jong K, Douglas S, Wolpert M, Delgadillo J, Aas B, Bovendeerd B, et al. Using progress feedback to enhance treatment outcomes: a narrative review. Adm Policy Ment Health 2025;52(1):210-222. [doi: 10.1007/s10488-024-01381-3] [Medline: 38733413]
5. Kendrick T, El-Gohary M, Stuart B, Gilbody S, Churchill R, Aiken L, et al. Routine use of patient reported outcome measures (PROMs) for improving treatment of common mental health disorders in adults. Cochrane Database Syst Rev 2016;7(7):CD011119 [FREE Full text] [doi: 10.1002/14651858.CD011119.pub2] [Medline: 27409972]
6. Østergård OK, Randa H, Hougaard E. The effect of using the partners for change outcome management system as feedback tool in psychotherapy-a systematic review and meta-analysis. Psychother Res 2020;30(2):195-212. [doi: 10.1080/10503307.2018.1517949] [Medline: 30213240]
7. Barkham M. Smaller effects matter in the psychological therapies: 25 years on from Wampold et al. (1997). Psychother Res 2023;33(4):530-532 [FREE Full text] [doi: 10.1080/10503307.2022.2141589] [Medline: 36525633]
8. de Jong K, van Sluis P, Nugter M, Heiser W, Spinhoven P. Understanding the differential impact of outcome monitoring: therapist variables that moderate feedback effects in a randomized clinical trial. Psychother Res 2012;22(4):464-474. [doi: 10.1080/10503307.2012.673023] [Medline: 22468992]
9. Boswell J. Monitoring processes and outcomes in routine clinical practice: a promising approach to plugging the holes of the practice-based evidence colander. Psychother Res 2020;30(7):829-842. [doi: 10.1080/10503307.2019.1686192] [Medline: 31672104]
10. Haken H, Schiepek G. Synergetik in der Psychologie: Selbstorganisation Verstehen und Gestalten. 2., Korrigierte Auflage 2010. Göttingen: Hogrefe; 2010.
11. Schiepek GK, Viol K, Aichhorn W, Hütt MT, Sungler K, Pincus D, et al. Psychotherapy is chaotic-(not only) in a computational world. Front Psychol 2017;8:379 [FREE Full text] [doi: 10.3389/fpsyg.2017.00379] [Medline: 28484401]
12. Finch AE, Lambert MJ, Schaalje BG. Psychotherapy quality control: the statistical generation of expected recovery curves for integration into an early warning system. Clin Psychol Psychother 2001;8(4):231-242. [doi: 10.1002/cpp.286]
13. Schiepek G, Gelo O, Viol K, Kratzer L, Orsucci F, de Felice G, et al. Complex individual pathways or standard tracks? A data‐based discussion on the trajectories of change in psychotherapy. Couns Psychother Res 2020;20(4):689-702. [doi: 10.1002/capr.12300]
14. Hannan C, Lambert MJ, Harmon C, Nielsen SL, Smart DW, Shimokawa K, et al. A lab test and algorithms for identifying clients at risk for treatment failure. J Clin Psychol 2005;61(2):155-163. [doi: 10.1002/jclp.20108] [Medline: 15609357]

XSL•FO
RenderX

15. Lutz W, Lambert MJ, Harmon SC, Tschitsaz A, Schürch E, Stulz N. The probability of treatment success, failure and duration—what can be learned from empirical data to support decision making in clinical practice? Clin Psychol Psychother 2006;13(4):223-232. [doi: 10.1002/cpp.496]
16. Harmon SC, Lambert MJ, Smart DM, Hawkins E, Nielsen SL, Slade K, et al. Enhancing outcome for potential treatment failures: therapist–client feedback and clinical support tools. Psychother Res 2007;17(4):379-392. [doi: 10.1080/10503300600702331]
17. Seizer L, Schiepek G, Cornelissen G, Löchner J. A primer on sampling rates of ambulatory assessments. Psychol Methods 2024. [doi: 10.1037/met0000656] [Medline: 38815067]
18. Olthof M, Hasselman F, Strunk G, van Rooij M, Aas B, Helmich MA, et al. Critical fluctuations as an early-warning signal for sudden gains and losses in patients receiving psychotherapy for mood disorders. Clin Psychol Sci 2019;8(1):25-35. [doi: 10.1177/2167702619865969]
19. Wampold BE. Routine outcome monitoring: coming of age--with the usual developmental challenges. Psychotherapy (Chic) 2015;52(4):458-462. [doi: 10.1037/pst0000037] [Medline: 26641376]
20. Wampold BE, Imel ZE. The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work (2nd Edition). New York: Routledge/Taylor & Francis Group; 2015.
21. Schiepek G, Stöger-Schmidinger B, Kronberger H, Aichhorn W, Kratzer L, Heinz P, et al. The therapy process questionnaire - factor analysis and psychometric properties of a multidimensional self-rating scale for high-frequency monitoring of psychotherapeutic processes. Clin Psychol Psychother 2019;26(5):586-602 [FREE Full text] [doi: 10.1002/cpp.2384] [Medline: 31153157]
22. Schiepek GK, Stöger-Schmidinger B, Aichhorn W, Schöller H, Aas B. Systemic case formulation, individualized process monitoring, and state dynamics in a case of dissociative identity disorder. Front Psychol 2016;7:1545 [FREE Full text] [doi: 10.3389/fpsyg.2016.01545] [Medline: 27812338]
23. Michaelis R, Edelhäuser F, Hülsner Y, Trinka E, Schiepek G. Personalized high-frequency monitoring of a process-oriented psychotherapeutic approach to seizure disorders: treatment utilization and participants' feedback. Psychotherapy (Chic) 2022;59(4):629-640. [doi: 10.1037/pst0000430] [Medline: 35201832]
24. Moltu C, Veseth M, Stefansen J, Nøtnes JC, Skjølberg Å, Binder P, et al. This is what I need a clinical feedback system to do for me: a qualitative inquiry into therapists' and patients' perspectives. Psychother Res 2018;28(2):250-263. [doi: 10.1080/10503307.2016.1189619] [Medline: 27219820]
25. Valentine KE, Milling LS, Clark LJ, Moriarty CL. The efficacy of hypnosis as a treatment for anxiety: a meta-analysis. Int J Clin Exp Hypn 2019;67(3):336-363. [doi: 10.1080/00207144.2019.1613863] [Medline: 31251710]
26. Fuhr K, Meisner C, Broch A, Cyrny B, Hinkel J, Jaberg J, et al. Efficacy of hypnotherapy compared to cognitive behavioral therapy for mild to moderate depression - results of a randomized controlled rater-blind clinical trial. J Affect Disord 2021;286:166-173. [doi: 10.1016/j.jad.2021.02.069] [Medline: 33725616]
27. Rosendahl J, Alldredge C, Haddenhorst A. Meta-analytic evidence on the efficacy of hypnosis for mental and somatic health issues: a 20-year perspective. Front Psychol 2024;14:1330238 [FREE Full text] [doi: 10.3389/fpsyg.2023.1330238] [Medline: 38268815]
28. Halsband U, Wolf TG. Current neuroscientific research database findings of brain activity changes after hypnosis. Am J Clin Hypn 2021;63(4):372-388. [doi: 10.1080/00029157.2020.1863185] [Medline: 33999768]
29. Renartz G. Autosystemhypnose - Wie sie wirkt und warum: Die selbstorganisatorische Hypnotherapiestrategie. Berlin, Denmark: Epubli; 2017.
30. Renartz G. Das unbewusste weiß, was zu tun ist. Exp Klin Hypnose 1996;12:123-136.
31. Depression. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2025-11-28]
32. Anxiety disorders. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/anxiety-disorders [accessed 2025-11-28]
33. Mennin DS, McLaughlin KA, Flanagan TJ. Emotion regulation deficits in generalized anxiety disorder, social anxiety disorder, and their co-occurrence. J Anxiety Disord 2009;23(7):866-871 [FREE Full text] [doi: 10.1016/j.janxdis.2009.04.006] [Medline: 19464142]
34. Strauman TJ, Socolar Y, Kwapil L, Cornwell JF, Franks B, Sehnert S, et al. Microinterventions targeting regulatory focus and regulatory fit selectively reduce dysphoric and anxious mood. Behav Res Ther 2015;72:18-29 [FREE Full text] [doi: 10.1016/j.brat.2015.06.003] [Medline: 26163353]
35. Barlow DH, Bullis JR, Comer JS, Ametaj AA. Evidence-based psychological treatments: an update and a way forward. Annu Rev Clin Psychol 2013;9:1-27. [doi: 10.1146/annurev-clinpsy-050212-185629] [Medline: 23245338]
36. Wampold BE. How important are the common factors in psychotherapy? An update. World Psychiatry 2015;14(3):270-277 [FREE Full text] [doi: 10.1002/wps.20238] [Medline: 26407772]
37. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, CONSORT group, Pragmatic Trials in Healthcare (Practihc) group. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ 2008;337:a2390 [FREE Full text] [doi: 10.1136/bmj.a2390] [Medline: 19001484]

XSL•FO
RenderX

38. Falkenström F, Markowitz JC, Jonker H, Philips B, Holmqvist R. Can psychotherapists function as their own controls? Meta-analysis of the crossed therapist design in comparative psychotherapy trials. J Clin Psychiatry 2013;74(5):482-491 [FREE Full text] [doi: 10.4088/JCP.12r07848] [Medline: 23146326]
39. Lutz W, Rubel J, Schiefele A, Zimmermann D, Böhnke JR, Wittmann WW. Feedback and therapist effects in the context of treatment outcome and treatment length. Psychother Res 2015;25(6):647-660. [doi: 10.1080/10503307.2015.1053553] [Medline: 26218788]
40. Flückiger C, Wolfer C, Held J, Hilpert P, Rubel J, Allemand M, et al. How to customize a bona fide psychotherapy for generalized anxiety disorder? A two-arms, patient blinded, ABAB crossed-therapist randomized clinical implementation trial design [IMPLEMENT 2.0]. BMC Psychiatry 2018;18(1):86 [FREE Full text] [doi: 10.1186/s12888-018-1666-2] [Medline: 29614982]
41. Beesdo-Baum K, Zaudig M, Wittchen HU. Strukturiertes Klinisches Interview für DSM-5®-Störungen - Klinische Version. Göttingen: Hogrefe; 2019.
42. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
43. The ICD-10 classification of mental and behavioural disorders: diagnostic criteria for research. World Health Organization. Geneva: World Health Organization; 1993. URL: https://www.who.int/publications/i/item/9241544554 [accessed 2025-12-18]
44. Schiepek G, Eckert H, Aas B, Wallot S, Wallot A. Integrative psychotherapy: a feedback-driven dynamic systems approach. Boston, MA: Hogrefe International Publishing; 2015.
45. Schiepek G. A dynamic systems approach to clinical case formulation. Eur J Psychol Assess 2003;19(3):175-184. [doi: 10.1027//1015-5759.19.3.175]
46. Schiepek G, Siebert-Blaesing B, Hausner MB. Systemische Fallkonzeption. Idiographische Systemmodellierung und Personalisierte Prozessbegleitung. Göttingen: Hogrefe; 2025.
47. Hoffknecht V. Ambulante langzeittherapie mit unterstützung durch das synergetische prozessmanagement. In: Schiepek G, Siebert-Blaesing B, Hausner MB, editors. Systemische Fallkonzeption. Idiographische Systemmodellierung und Personalisierte Prozessbegleitung. Göttingen: Hogrefe; 2025:56-72.
48. Franke G. Mini-SCL. Hogrefe. 2017. URL: https://www.testzentrale.de/shop/mini-symptom-checklist.html [accessed 2025-11-28]
49. Lovibond SH, Lovibond PF. Manual for the Depression Anxiety & Stress Scales. 2nd ed. Sydney: Psychology Foundation; 1995.
50. Nilges P, Essau C. Die Depressions-Angst-Stress Skalen. Schmerz 2015;29(6):649-657. [doi: 10.1007/s00482-015-0019-z] [Medline: 26205682]
51. Schwarzer R, Jerusalem M. SWE. Skala zur Allgemeinen Selbstwirksamkeitserwartung. Trier: ZPID (Leibniz Institute for Psychology) – Open Test Archive; 2003.
52. Donovan JL, Kay HE, Peters TJ, Abrams P, Coast J, Matos-Ferreira A, et al. Using the ICSOoL to measure the impact of lower urinary tract symptoms on quality of life: evidence from the ICS-'BPH' Study. International Continence Society--Benign Prostatic Hyperplasia. Br J Urol 1997;80(5):712-721. [Medline: 9393291]
53. Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. J Psychopathol Behav Assess 2004;26(1):41-54. [doi: 10.1023/B:JOBA.0000007455.08539.94]
54. Ehring T, Fischer S, Schnülle J, Bösterling A, Tuschen-Caffier B. Characteristics of emotion regulation in recovered depressed versus never depressed individuals. Pers Individ Dif 2008;44(7):1574-1584. [doi: 10.1016/j.paid.2008.01.013]
55. Schulz H, Nübling R, Rüddel H. Entwicklung einer kurzform eines fragebogens zur psychotherapiemotivation. Verhaltenstherapie 1995;5:89-95. [doi: 10.1159/000258901]
56. Flückiger C, Regli D, Zwahlen D, Hostettler S, Caspar F. Der berner patienten- und therapeutenstundenbogen 2000: Ein instrument zur erfassung von therapieprozessen. Z Klin Psychol Psychother 2010;39(2):71-79. [doi: 10.1026/1616-3443/a000015]
57. Faul F, Erdfelder E, Lang A, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
58. Cuijpers P, van Straten A, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. Psychol Med 2010;40(2):211-223. [doi: 10.1017/S0033291709006114] [Medline: 19490745]
59. Viol K, Schöller H, Kaiser A, Fartacek C, Aichhorn W, Schiepek G. Detecting pattern transitions in psychological time series - a validation study on the Pattern Transition Detection Algorithm (PTDA). PLoS One 2022;17(3):e0265335 [FREE Full text] [doi: 10.1371/journal.pone.0265335] [Medline: 35275971]
60. Wallot S, Roepstorff A, Mønster D. Multidimensional Recurrence Quantification Analysis (MdRQA) for the analysis of multidimensional time-series: a software implementation in MATLAB and its application to group-level data in joint action. Front Psychol 2016;7:1835 [FREE Full text] [doi: 10.3389/fpsyg.2016.01835] [Medline: 27920748]
61. Webber CL, Marwan N. Recurrence quantification analysis: theory and best practices. Berlin: Springer; 2015.

XSL•FO
RenderX

62. Gelo OCG, Salvatore S. A dynamic systems approach to psychotherapy: a meta-theoretical framework for explaining psychotherapy change processes. J Couns Psychol 2016;63(4):379-395. [doi: 10.1037/cou0000150] [Medline: 27177027]
63. Simon W, Lambert MJ, Harris MW, Busath G, Vazquez A. Providing patient progress information and clinical support tools to therapists: effects on patients at risk of treatment failure. Psychother Res 2012;22(6):638-647. [doi: 10.1080/10503307.2012.698918] [Medline: 22755547]
64. Bickman L, Kelley SD, Breda C, de Andrade AR, Riemer M. Effects of routine feedback to clinicians on mental health outcomes of youths: results of a randomized trial. Psychiatr Serv 2011;62(12):1423-1429. [doi: 10.1176/appi.ps.002052011] [Medline: 22193788]
65. Casline E, Woodard G, Patel ZS, Phillips D, Ehrenreich-May J, Ginsburg GS, et al. Characterizing measurement-based care implementation using therapist report. Evid Based Pract Child Adolesc Ment Health 2023;8(4):549-559 [FREE Full text] [doi: 10.1080/23794925.2022.2124555] [Medline: 38031580]
66. Deisenhofer A, Barkham M, Beierl ET, Schwartz B, Aafjes-van Doorn K, Beevers CG, et al. Implementing precision methods in personalizing psychological therapies: barriers and possible ways forward. Behav Res Ther 2024;172:104443 [FREE Full text] [doi: 10.1016/j.brat.2023.104443] [Medline: 38086157]

Abbreviations

AD: anxiety disorder
ASH: autosystemic hypnotherapy (original German term: Autosystemhypnose)
DSM-5: Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition]
GAD: generalized anxiety disorder
ICD-10: International Statistical Classification of Diseases, Tenth Revision
MDD: major depressive disorder
RCT: randomized controlled trial
SNS: synergetic navigation system
SPM: synergetic process management
TPQ: Therapy Process Questionnaire

Edited by J Sarvestan; submitted 28.May.2025; peer-reviewed by A Hafkenscheid, KW Tay; comments to author 28.Aug.2025; accepted 17.Nov.2025; published 07.Jan.2026.

Please cite as:
Schiepek G, Wackernagel S
Improving Outpatient Psychotherapy for Adults With Major Depressive and Anxiety Disorders Using Web-Based High-Frequency Monitoring and Feedback in Autosystemic Hypnotherapy: Protocol for a Two-Arm ABAB Crossed-Therapist Randomized Clinical Implementation Trial
JMIR Res Protoc 2026;15:e78166
URL: https://www.researchprotocols.org/2026/1/e78166
doi:10.2196/78166
PMID:

©Günter Schiepek, Stephanie Wackernagel. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 07.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

An Online Gamified Pain Management Program for Adults With Chronic Pain: Protocol for a Randomized Controlled Trial

Mun Yee Mimi Tse[1], Prof Dr; Jiafan He[1], MSc; Tyrone Tai On Kwok[1], Prof Dr; Sukki Ho[2], Prof Dr

[1]School of Nursing and Health Sciences, Hong Kong Metropolitan University, Kowloon, China (Hong Kong)

[2]Hong Kong Polytechnic University, Kowloon, China (Hong Kong)

Corresponding Author:
Jiafan He, MSc
School of Nursing and Health Sciences
Hong Kong Metropolitan University
Jockey Club Institute of Healthcare, Hong Kong Metropolitan University
1 Sheung Shing Street, Homantin
Kowloon, 999077
China (Hong Kong)
Phone: 852 56152810
Email: s1350029@live.hkmu.edu.hk

Abstract

Background: Although digital interventions have been widely used in chronic pain management, evidence on their effectiveness has demonstrated heterogeneity. They have low retention rates, which are exacerbated by a lack of feedback, social interaction, and engaging content, contributing to boredom and higher attrition. Gamification offers a possibility to overcome these barriers and enhance user motivation and engagement.

Objective: This study aims to investigate the effects of an online gamified pain management program in a sample of adults with chronic noncancer pain.

Methods: A 2-arm, single-blind randomized controlled trial will recruit 172 participants through Chinese hospitals and online platforms. Eligible participants will be randomized to 2 groups: a 10-week gamified pain management program for the experimental group and routine online pain management materials for the control group. Primary outcomes, including pain intensity and pain interference, and secondary outcomes, including fear avoidance, emotion, exercise adherence, quality of life, and user satisfaction, will be assessed at baseline, after the intervention (10 weeks), and at a 3-month follow-up. A mixed-design ANOVA and intention-to-treat analysis will be used to evaluate intervention effects.

Results: The trial is registered, with recruitment ongoing from July 2025 to August 2026. A total of 172 participants (86 per group) will be recruited. As of November 17, 2025, we have recruited 19 participants. Results are expected to be published in March 2027

Conclusions: The findings will address a critical gap by examining a novel gamified intervention within a theory-driven design to provide an engaging solution to improve self-management in people with chronic pain.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400094247; https://www.chictr.org.cn/hvshowproject.html?id=266029&v=1.0

International Registered Report Identifier (IRRID): DERR1-10.2196/80542

(JMIR Res Protoc 2026;15:e80542) doi:10.2196/80542

KEYWORDS

pain; pain management; gamification; digital health; self-determination theory; randomized controlled trial

Introduction

The Global Burden of Disease Study highlights the substantial prevalence and impact of pain globally [1]. Chronic pain, defined as pain persisting for more than 3 months, significantly affects various aspects of daily life, including work performance, physical activity, and social functioning. It is often associated with psychological disorders such as anxiety and depression, sleep disturbances, substance use, cognitive impairment, and reduced work productivity [2].

XSL•FO
RenderX

Existing guidelines advocate for biopsychosocial treatments that incorporate education, exercise, and psychological interventions to reduce pain in both the short and long term. These approaches emphasize the importance of physical activity and engagement in meaningful daily activities [3,4]. However, the International Association for the Study of Pain has concluded that the training provided to health professionals in pain management is insufficient to achieve effective outcomes [5]. Additionally, patients with chronic pain often face substantial challenges in maintaining physical activity, as they may feel both physically inactive and emotionally hesitant due to the fear of exacerbating pain [6,7].

Online platforms have revolutionized the management of chronic pain by providing accessible alternatives to conventional medical treatments [8]. These platforms allow patients to engage with pain management strategies at their own pace, overcoming barriers such as geographic isolation and socioeconomic disadvantage [9]. Studies suggest that telerehabilitation saves costs and that web-based exercise interventions improve home exercise adherence and confidence [10,11]. Despite their potential, existing online pain management programs often suffer from low retention rates, a problem that is exacerbated by a lack of feedback, social interaction, and engaging content [12]. Gamification has emerged as a promising solution to these challenges, offering an approach that can enhance user engagement by incorporating game elements that motivate and sustain participation [13].

Gamification leverages game mechanics to influence user behavior, improving engagement, motivation, and overall health outcomes [14]. In health care, gamification has proven effective in addressing various health issues and promoting physical activity by making exercise and learning both enjoyable and motivating [15,16]. The use of motivational affordances within gamification has been shown to positively influence psychological factors, such as motivation, attitude, and enjoyment [17], making it a potentially powerful tool for behavior change. Regarding chronic pain, a systematic review demonstrated that gamified exercise interventions, including those using platforms such as Kinect Sport, Nintendo Wii, and virtual reality, can effectively reduce pain by providing interactive scenarios and rewards [18].

Despite these benefits, existing research has not fully explored gamification within the biopsychosocial framework for chronic pain management. Most studies have focused primarily on increasing physical activity, while neglecting the critical educational and psychological components that are essential for comprehensive pain management [18]. As fear of pain is associated with avoidance behaviors, educational interventions, such as those based on pain neurophysiology, can help patients better understand pain and modify their attitudes toward it, potentially leading to improved outcomes [19]. The ability of patients to accept pain sensations, rather than attempting to fight them, has been associated with greater reductions in pain interference [20]. Recent studies have also highlighted the effectiveness of biopsychosocial interventions that combine physical exercise and psychological support to reinforce positive behaviors and challenge negative pain-related perceptions [21,22].

To address this gap in the literature, an online gamified pain management program, based on an escape room task, will be developed and augmented with game elements. The aim of this study is to investigate the effects of online gamified pain management in a sample of adults with chronic noncancer pain via a randomized controlled trial. To this end, the objectives will be to (1) examine the effects of the gamified pain management intervention on the levels of pain intensity, pain interference, psychological outcomes, and quality of life among adults; (2) assess exercise adherence and game experience in the gamified home-based management program; and (3) identify the relationships between the effects of the intervention and sociodemographic factors.

Methods

Study Design and Setting

This study is a 2-arm, single-blind (analysts are blinded), randomized controlled trial. A total of 172 individuals with chronic pain will be recruited and randomly assigned to either gamified pain management program (intervention group) or a digital resource package (control group). The interventions for both groups will be delivered online. Participants will complete outcome assessments at 3 time points: baseline (T0), immediately after the intervention (T1), and at a 3-month follow-up (T2). This study protocol follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [23].

Sample Size

A previous power analysis was used for sample size calculations using G*Power software (version 3.1; Heinrich-Heine-Universität Düsseldorf) [24]. To establish the effect size for the sample size calculation, we synthesized evidence from multiple sources. A meta-analysis analyzing the effects of serious games on pain intensity in patients with chronic pain reported a mean difference of −0.62 (95% CI −1.15 to −0.10) for pain reduction [18]. To supplement this, we conducted another meta-analysis of 12 relevant studies examining the effects of gamification on pain intensity in adults. The results yielded a mean difference of −0.86 (95% CI −1.61 to −0.11) [25-36]. However, given that our digital intervention comprises mixed physical, psychological, and educational components, we also incorporated evidence from 2 other meta-analyses including online pain management and multidisciplinary programs, with effect size of −0.30 (95% CI −0.50 to −0.10) [12] and −0.57 (95% CI −0.63 to −0.50) [37].

Therefore, to ensure robust and conservative statistical power, we selected a medium effect size of 0.50 to reduce potential bias or heterogeneity [38]. The following parameters were used: power (1−β)=0.90, effect size=0.50, and α=.05 (2 tailed). Taking a 20% margin for dropout and missing data, a total of 172 participants will be recruited, with 86 allocated to the experimental group and 86 to the control group.

Participants and Recruitment

Participants recruitment will be conducted through both on-site and online approaches. A series of offline educational lectures accompanied by recruitment flyers or posters will be held at

XSL•FO
RenderX

outpatient pain clinics at Tianhe District People's Hospital and community service centers in Guangdong province to attract potential participants. Online recruitment will be conducted via social media platforms, including WeChat (version 8.0.66; Tencent Holdings Limited), Xiaohongshu (version 9.14; Xingyin Information Technology Limited), and TikTok (version 6.9.0; ByteDance), by posting advertisements. All recruitment materials will include a brief introduction to the study aim, implementation process, and a survey hyperlink on the Wenjuanxing platform [39] (Ranxing Information Technology Co, Ltd), which will direct participants to the consent form and eligibility screening.

To be eligible to participate in the study, the participants must meet all the following inclusion criteria: (1) be aged 18 years or older, (2) be able to read and understand Mandarin Chinese, (3) own a smartphone with internet access, (4) have a previously diagnosed chronic noncancer pain condition lasting more than 3 months [40], and (5) report a score of ≥2 on the visual analogue scale [41].

Participants will be excluded if they meet one or more of the following exclusion criteria: (1) associated pathologies that make it impossible to perform physical exercise (eg, myopathies, neurological diseases, cardiac diseases, pregnancy, pulmonary diseases, infection, or fracture) [42], (2) undergoing surgery or invasive treatments within the past 3 months [43], or (3) concurrent participation (or participation within the preceding 3 months) in a supervised exercise program or multidisciplinary treatment.

Procedure and Randomization

The eligibility screening survey will automatically classify individuals after clicking the Agree button as they agree to provide consent. Screening questions will consist of pain intensity, pain duration, health history, recent treatment enrollment, and demographic data. The survey will be designed with conditional logic such that individuals who do not meet the eligibility criteria will be automatically redirected to the end of the survey and prompted to submit their responses. This approach means individuals who are not eligible will neither complete the full survey nor engage in the intervention. Conversely, eligible individuals will proceed to complete the remaining sections, which constitute the full baseline assessment. At the end of the questionnaire, interested participants may provide their WeChat account to receive group allocation and intervention information.

Following baseline assessment, enrolled participants will be randomly assigned in a 1:1 ratio to either the intervention or the control group. A computer-generated randomization sequence using blocks of sizes 2 and 4, stratified by age and sex, will be used. The randomization process will be managed by an independent researcher (JH), and allocation numbers will be stored securely in a password-protected database. Allocation will be revealed to the intervention deliverer (Chen Z, a PhD student) only after a participant is enrolled. The intervention deliverer will be blinded to data assessment and data analysis.

After randomization, participants will be notified of their group allocation and provided with preassigned WeChat log-in credentials. Those in the intervention group will receive a direct message containing a hyperlink to access the gamified pain management platform, while those in the control group will receive a digital resource package. After the intervention, T1 (10 weeks immediately after intervention) and T2 (3-month follow-up) assessments will be collected via the Wenjuanxing platform [39], and semistructured interviews will be conducted to collect individual's perceptions and experiences of the intervention. Participants who report the most, moderate, or least change in pain intensity will be invited to participate in the interviews.

Study Intervention

The design of the gamified pain management is based on established gamification principles and frameworks [44,45] and is theoretically grounded in self-determination theory (SDT), which supports individual autonomy, competence, and relatedness to improve intrinsic motivation and sustain engagement [46]. The intervention was co-designed by an expert in chronic pain (MYMT) and an expert in digital health (TTOK), with contributions from the lead researcher (JH), a pain clinic clinician, a physiotherapist with more than 5 years of experience, and 3 internet technology specialists. The intervention is delivered via the web-based platform VxEditor [47], which is accessible on both smartphones and computers, over a maximum of 10-week period, allowing participants to complete the program at their own pace. In addition, participants can choose their specific pain site via the platform, which tailors exercises accordingly and thus improve autonomy.

The gamified pain management is designed as an "escape room" game. Participants assume the role of a hero (avatar) guiding others to escape from a metaphorical "pain room," which fosters a sense of relatedness. The intervention consists of 8 core "rooms" (sessions). Participants must complete various challenges, including unlocking 10 clues and accumulating at least 60 points to progress from one session to the next, which is designed to foster a sense of competence. Participants are awarded 5 points for completing each exercise task and 2 points for correct quiz answers, while skipping an exercise earns no points. To further support competence and reinforce the knowledge of pain, immediate and automatic feedback is provided, offering affirmation and encouragement for correct quiz answers and explanatory feedback for incorrect answers. Four optional hidden sessions are also available, offering additional challenges for highly engaged users. A detailed overview of the gamification design is provided in Table 1.

XSL•FO
RenderX

Table 1. Gamification design of the interventions.

Game elements	Applicable gamified interventions
Anonymity	Users are anonymized and represented by a unique code.
Avatar	Users assume the role of a hero (avatar) in the game.
Narrative	A background story (escape from the pain room) is presented to engage users.
Tutorials	Introductory tutorials establish an emotional connection with users and provide step-by-step instructions.
Choices	Users can select different games based on their specific pain site.
Feedback	Users receive notifications upon completion of exercises and correct quiz answers.
Points	Users earn points for completing exercises and quizzes and for progressing to the next sessions.
Progression and leaderboard	A progress number indicates the level of effort achieved.
Challenges and rewards	Completing challenges unlocks hidden games with additional exercise opportunities.

The contents of the intervention consist of exercises complemented with education and relaxation practices, which are identified as “clues” within each session. Education topics are presented through animated videos and written notes, focusing on pain mechanisms, neurobiology and neurophysiology, the biopsychosocial model of pain, and management strategies [48]. Quizzes are used to assess and reinforce knowledge from the educational content. A total of 28 prerecorded exercise videos demonstrate stretching, strengthening, balance, and resistance exercises [49], tailored to specific pain areas (eg, neck, shoulder, back, knee, or head). The program recommends at least 60 minutes of exercise per week, and the completion of each exercise is verified via an in-video dialogue prompt that appears midway through the video, which participants must click to earn points. The relaxation practices are presented through recorded audios that guide participants in breathing exercises and muscle relaxation [50]. In addition, participants are required to solve real-world case puzzles by applying their knowledge to determine the mechanisms of chronic pain, which is a key task for session completion. Details of these components are shown in Table 2.

XSL•FO
RenderX

Table 2. The gamified pain management intervention components.

Session	Clue components
1	• Exercise: three 6-minute balancing exercises • Educational information: pain mechanisms • Relaxation audio: breathing exercises • Quiz and puzzle
2	• Exercise: three 6-minute stretching exercises for different pain sites • Educational information: etiology of pain • Relaxation audio: breathing exercises • Quiz and puzzle
3	• Exercise: three 6-minute core strengthening exercises • Educational information: neurobiology and neurophysiology of pain • Relaxation audio: muscle relaxation • Quiz and puzzle
4	• Exercise: three 6-minute stretching exercises for different pain sites • Educational information: biopsychosocial model of chronic pain • Relaxation audio: muscle relaxation • Quiz and puzzle
5	• Exercise: three 6-minute core strengthening exercises • Educational information: attention system and pain and distraction from pain • Relaxation audio: body scan • Quiz and puzzle
6	• Exercise: three 6-minute balancing exercises for different pain sites • Educational information: medication use and drug abuse • Relaxation audio: body scan • Quiz and puzzle
7	• Exercise: three 6-minute core strengthening exercises • Educational information: physical activity for chronic pain • Relaxation audio: relaxation music • Quiz and puzzle
8	• Exercise: three 6-minute resistance exercises • Educational information: nonpharmacological strategies • Relaxation audio: outdoor relaxation • Quiz
Four hidden sessions	• Exercise: three 20-minute stretching and resistance exercises

The gamified pain management platform automatically captures and stores detailed engagement data on a secure, cloud-based back end. The latest room successfully completed is recorded, allowing research staff to track progress. Time stamps for the first log-in and last interaction with the platform are also generated to calculate the total time spent. Quiz accuracy rates and total points accumulated are stored on the backend, enabling monitoring by the research team. To promote adherence, interim reminders will be sent via WeChat biweekly to participants who have not progressed at the expected pace. After completing the program, participants are encouraged to continue using the platform for practice. Screenshots of the gamified pain management are provided in Figures 1-4.

XSL•FO
RenderX

Figure 1. Home page interface of the gamified pain management intervention.

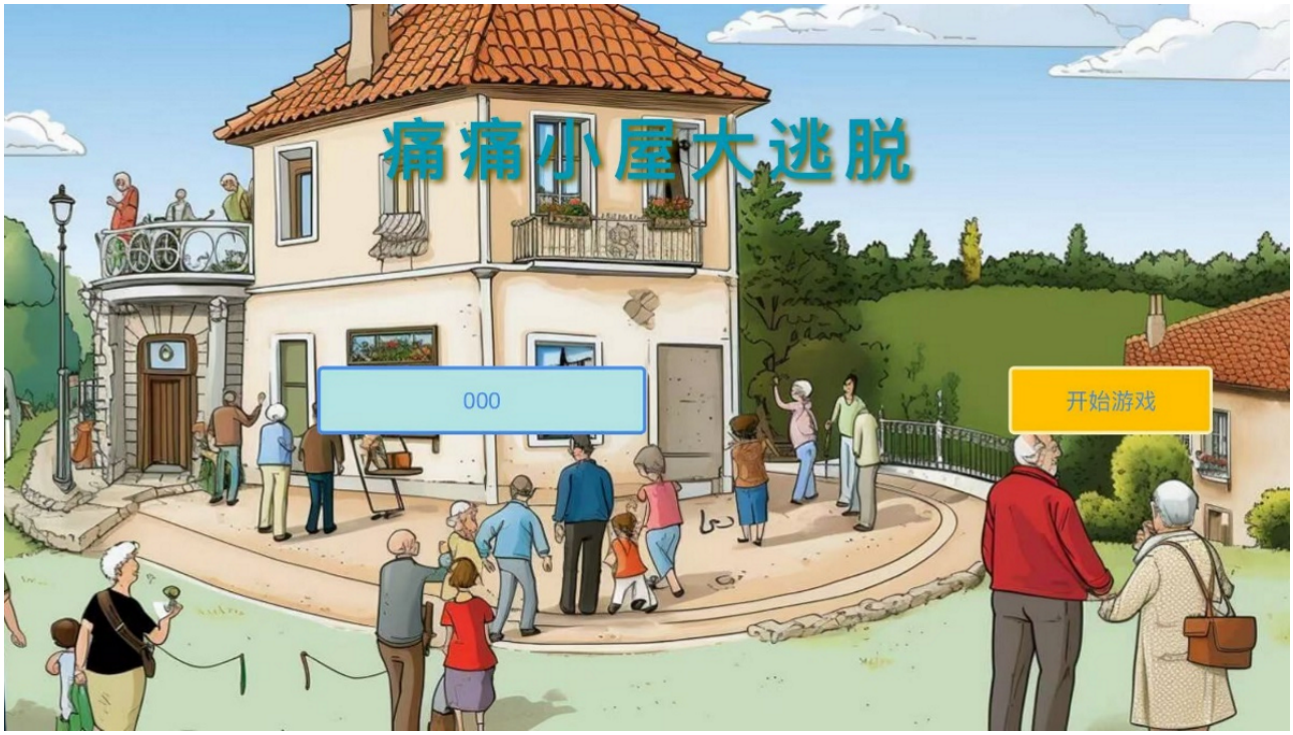

Figure 2. Choice feature interface.



XSL•FO
RenderX

Figure 3. Room background and points feature interface.



Figure 4. Quiz function interface.

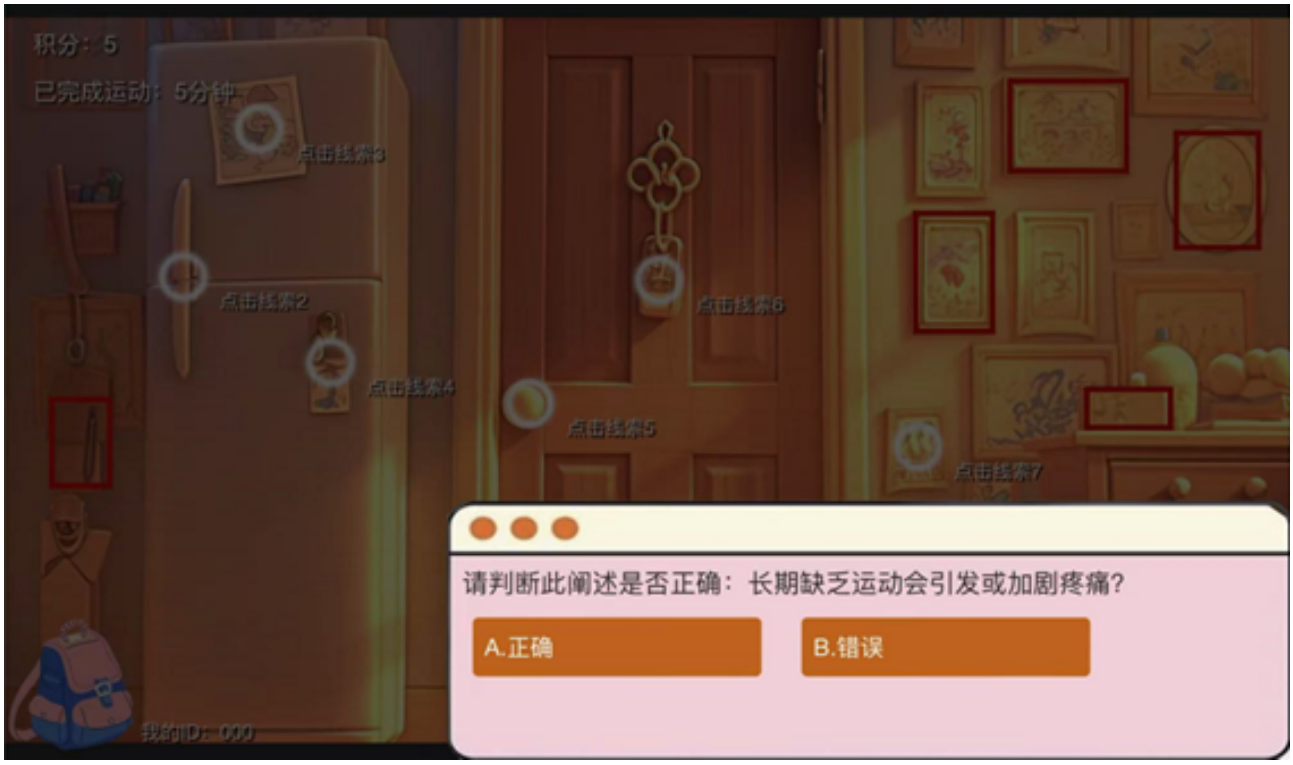

Control Group

In the control group, participants will receive a digital pain management resource package for a 10-week period, containing the same components as the gamified pain management group but delivered in a nongamified, standard digital format. The package contains educational content, exercise videos, educational videos, and relaxation audios that are identical in content and volume to those used in the gamified pain management group, delivered via an online PDF handbook with hyperlinks through WeChat. These multimedia resources are hosted on WeChat and the Xiaohongshu official account to ensure accessibility. In addition, participants will receive interim reminders biweekly via WeChat, encouraging them to read the handbook and engage in exercises, relaxation, and educational

XSL•FO
RenderX

activities to promote adherence throughout the intervention period.

Outcome Measures

The T0 (baseline before intervention), T1 (immediately after intervention), and T2 (the 3-month follow-up) assessments will be collected via the Wenjuanxing platform [39].

Primary Outcome

Pain intensity and pain interference are assessed using the Brief Pain Inventory between baseline, T1, and T2, rated from 0 (no pain) to 10 (worst pain). Pain severity is calculated based on pain experienced in the past 24 hours and "current pain." Pain interference includes the degree to which pain interferes with general activity, mood, walking, work, social relationships, sleep, and enjoyment of life. The Chinese version has good reliability, with Cronbach α=0.89 [51].

Secondary Outcomes

Study Instruments

The Fear Avoidance Beliefs Questionnaire is used to measure fear of pain and consequent avoidant behaviors related to physical activity. The questionnaire consists of 16 items; each rated on a 7-point Likert scale ranging from 0 (completely disagree) to 6 (completely agree). The Cronbach α of the Chinese version ranges from 0.75 to 0.85 [52].

The Pain Catastrophizing Scale is used to measure pain catastrophizing beliefs, including 3 subscales, rumination, magnification, and helplessness, rated from 0 to 4. Higher scores indicate stronger pain catastrophizing beliefs. The Cronbach α of the Chinese version is 0.91 [53].

The Godin-Shephard leisure-time physical activity questionnaire is used to measure exercise adherence, including frequency of strenuous, moderate, and mild exercise per week. The percentage of exercises completed out of the total number of prescribed exercises is used to measure total exercise adherence [54].

The Patient Health Questionnaire-9 is used to assess the symptoms of depression, scored on a 4-point Likert scale from 0 (not at all) to 3 (nearly every day). Higher scores indicate greater severity of depressive symptoms. The Cronbach α of the Chinese version of this questionnaire is 0.86 [55].

The Generalized Anxiety Disorder-7 is used to measure anxiety, including 7 items rated from 0 (not at all) to 3 (nearly every day). The Chinese version has excellent internal reliability, with a Cronbach α of 0.91 [56].

The 3-level version of the EuroQol-5D is used to assess an individual's perceptions of quality of life, including mobility, self-care, usual activities, pain and discomfort, and anxiety and depression. Each dimension is evaluated with 3 levels of severity. Higher scores indicate a lower level of quality of life. The Chinese version has been treated as the standard in China [57].

The Game User Experience Satisfaction Scale-18, an 18-item scale is used to evaluate overall user satisfaction with the game, rated from 0 (strongly disagree) to 7 (strongly agree), with good internal reliability (Cronbach α=0.93) [58]. Higher scores indicate a higher level of satisfaction.

Intraintervention Variables (Collected During or Immediately After the Intervention)

Engagement metrics are assessed through session-specific satisfaction scores (Satisfaction scores for sessions 1 to 8) and quiz accuracy, recorded via the gamified pain management platform. Satisfaction scores are derived from the sum of 3 items rated on a 5-point Likert scale following the completion of each session. These items evaluate the following: (1) understandability ("Are the materials in this session easy for you to understand?"), (2) comprehensibility ("How much do you understand the materials in this session?"), and (3) applicability ("To what extent do you think you would apply this session to future self-management?"). Quiz accuracy is calculated as the proportion of correctly answered questions in each session's embedded quiz.

Treatment duration is automatically calculated as the total number of weeks between a participant's first and last log-in on the gamified pain management platform.

Semistructured Interviews

Semistructured interviews are conducted by JH after completion of the intervention to measure participants' perceptions, attitudes, and experiences after participating in the intervention. Interviews will be conducted via a private online Tencent Meeting room (version 3.37.10.401; Tencent Meeting) [59], ensuring that participants' confidentiality and privacy are protected. In addition, perceived motivators and barriers related to gamification elements will be explored. Each interview will last 40 to 60 minutes. Examples of semistructured interview questions are provided in Textbox 1.

Textbox 1. Examples of interview questions.

1. How has your pain intensity changed since the start of the gamified pain management intervention?
2. What changes have occurred in your perception of pain since you started receiving the gamified pain management intervention?
3. What aspects of gamified pain management do you think are most beneficial? How about the game elements?
4. What challenges or difficulties have you encountered when using the gamified pain management intervention?
5. Which game element do you find most useful, and why?
6. Do you plan to continue using the GPM intervention after this study? If not, what barriers are preventing you from continuing?
7. What suggestions do you have for improving the gamified pain management intervention?

XSL•FO
RenderX

Statistical Analysis

SPSS (version 25.0; IBM Inc) will be used for all statistical analyses. An intention-to-treat analysis will be used, whereby all participants will be analyzed in the context of the groups to which they had been originally assigned. Descriptive statistics for feasibility will include the following: (1) the percentage of participants recruited from the total number approached; (2) the percentage of participants (among those recruited) who are retained in the study at the end of 3 months in each arm and overall; and (3) engagement metrics, including completion status, duration of the intervention, and rate of accuracy on quizzes.

To detect potential differences in sociodemographic and baseline clinical characteristics, the χ^2 test with continuity correction (or the 2-sided Fisher exact test when appropriate) will be used for categorical variables, and the *t* test will be used for continuous variables. The normality of distributions will be confirmed using the Shapiro-Wilk test before statistical analysis. Repeated measures ANOVA (mixed design: within and between groups) will be used to determine the overall effectiveness of the intervention. Missing values will be addressed using intention-to-treat principles with maximum likelihood estimation, assuming data are missing at random. Statistical significance will be set at $P<.05$.

The interviews will be transcribed verbatim and reviewed. Open and axial coding will be used for data identification and integration, resulting in different meaning categories derived using grounded theory methods. NVivo (QSR International Inc) will be used for the coding process and information handling.

Ethical Considerations

This study has been carried out in accordance with the guidelines of the Hong Kong Metropolitan University Research Ethics Committee (HE-SF2024/34) and was registered in the Chinese Clinical Trial Registry (ChiCTR2400094247).

All potential participants will be provided with a detailed digital informed consent form. This form outlines the purpose of the study, procedures, potential risks and benefits, confidentiality measures, voluntary participation, the right to withdraw at any time without penalty, and contact information for the principal investigator. Participants will be required to click the Yes button below the statement “I have read and understood the above information, and I voluntarily agree to participate in this study” to indicate their consent to participate.

To ensure privacy and confidentiality, all data will be anonymized by assigning a unique reference number, which will be used to link all data collected from the game and surveys, rather than using participants’ names. All data will be securely stored in an encrypted folder, with backups stored on a separate encrypted external hard drive to further protect participant confidentiality. Only the research team will have access to the information.

During the design of this study, care was taken to ensure that participants would not be exposed to any physical harm or discomfort. To prevent potential harm, all participants will be required to undergo prescreening to assess any existing health conditions that might contraindicate physical activity, as advised by their health care providers. Participants with such conditions will be deemed ineligible to participate in the study.

Results

Study Timeline

Recruitment was initiated in July 2025, and enrollment is ongoing. As of November 17, 2025, we have recruited 19 participants. Data analysis has not yet started. Final data collection is expected to be completed by August 2026, and the results are expected to be published in March 2027. Figure 5 shows the trial flow diagram.

XSL•FO
RenderX

Figure 5. Trial flow diagram.

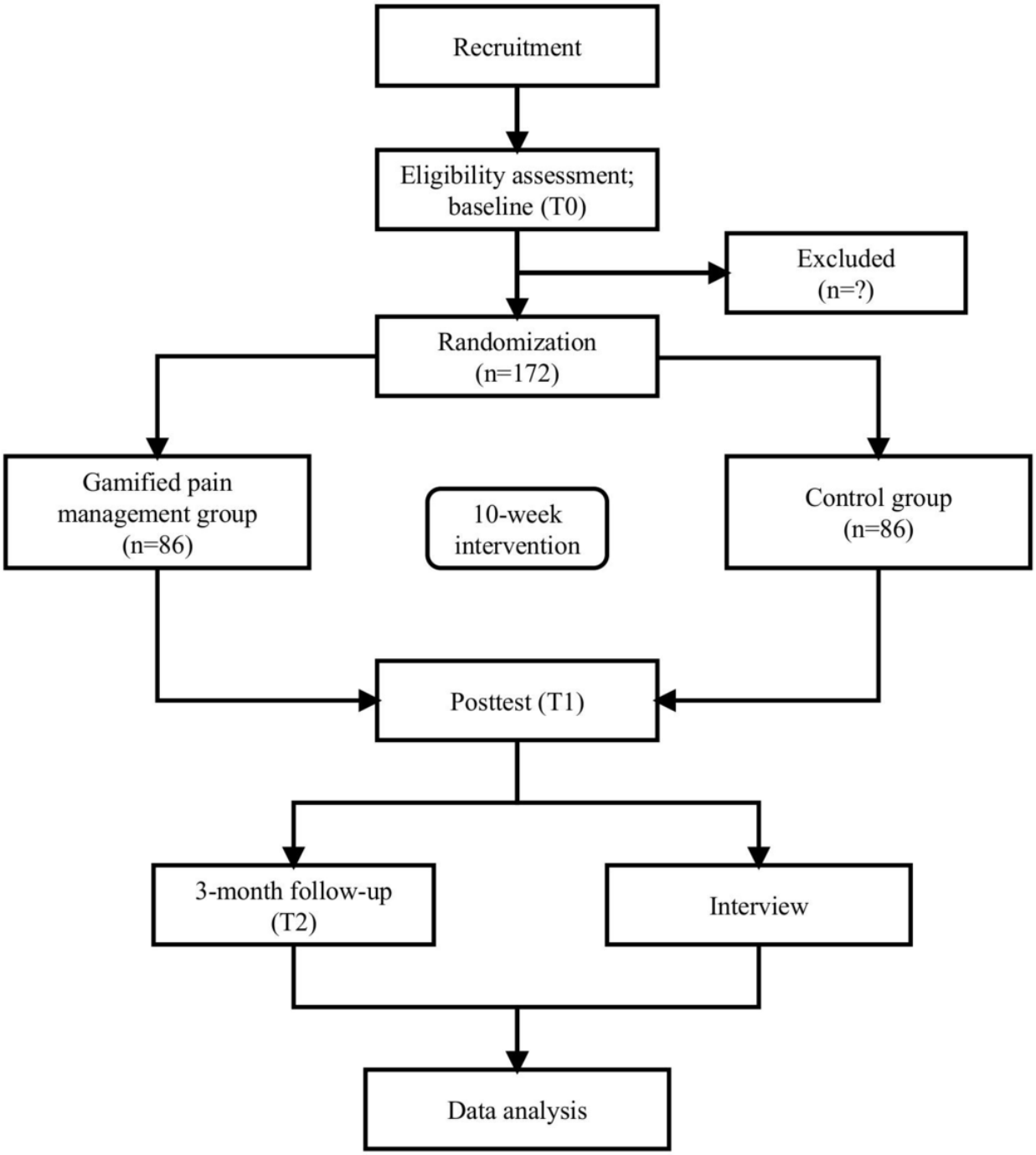

Anticipated Findings

We hypothesize that participants in the gamified pain management intervention will report a substantial reduction in pain intensity and pain interference compared with the control group at both postintervention and 3-month follow-up. Furthermore, the gamified pain management group is expected to exhibit greater improvements in secondary outcomes, including emotional distress, fear-avoidance beliefs, and pain catastrophizing, as well as higher engagement and exercise adherence. In addition, recovery is expected to be associated with sex, age, and socioeconomic status.

Discussion

Principal Findings

This study has demonstrated how gamification can be applied in health care. Gamification has been explored in various studies targeting different populations with diverse health goals. For example, a study conducted in 2024 on the use of virtual reality exergames in pain management showed positive effects for individuals with chronic lower back pain [60]. Given the insufficient adherence and motivation in chronic pain management, this study goes beyond the existing state of the

XSL•FO
RenderX

art by developing a gamified pain management intervention, which reinforces patients' motivation for self-management through an engaging approach combined with real-time feedback on achievements.

In the gamified pain management intervention, we will collect data through log-in time points, quiz responses, points accumulation, and weekly question-and-answer sessions on the gamified pain management platform. The information obtained is used to personalize physical exercises according to participants' pain locations, thereby enhancing engagement. In addition, the intervention enhances user engagement and motivation by leveraging key principles of SDT. The escape room narrative and avatar role fulfill the need for relatedness by creating an immersive, purpose-driven experience that connects users to a larger mission. The points system, progress tracking, and tiered challenges address competence by providing clear goals, immediate feedback, and achievable milestones that validate skill development. Flexible pacing, pain-specific exercise choices, and optional hidden rooms support autonomy, allowing users to tailor the intervention to their preferences and capabilities. Together, these game mechanics, reinforced by educational content and physical practices, transform pain management into an intrinsically motivating process, consistent with studies demonstrating a moderate fit between SDT and the promotion of exercise and peer support for patients with chronic pain [61]. Furthermore, in the context of gamification, SDT provides a valuable lens for understanding how gamification elements influence motivation and sustain behavior change [62].

This randomized controlled trial was specifically designed to address key questions in the field of gamification research, particularly regarding the efficacy of gamification in pain management and the optimal design of such interventions. To our knowledge, this is the first randomized controlled trial combining gamified designs with education, exercise, and relaxation techniques aimed at chronic pain reduction. Randomization was implemented to prevent selection bias and control for confounding variables. Furthermore, while most existing studies focus primarily on pain intensity outcomes, this study incorporated psychological, perceptual, and satisfaction-related outcome measures to provide a more comprehensive evaluation of the intervention's effectiveness.

Limitations

Several limitations of this study warrant discussion. First, the research was conducted exclusively among adults, which may limit the generalizability of the findings. Future studies should aim to expand the applicability of the results by examining the impact of gamification on other populations, such as children, and across various settings. Second, at the time of designing this study, social interaction features were not incorporated into the gamified intervention. Given that loneliness is a common issue among individuals with chronic pain [63], the absence of a social context may have hindered the potential development of social and emotional skills and reduced opportunities to mitigate social isolation. Future research should explore the role of social interaction within gamification to determine whether it enhances engagement and outcomes.

In addition, although a click function was integrated into the exercise videos to monitor participant adherence, this method has limitations. Peer collaboration could potentially offer more benefits in terms of increasing participant engagement and adherence to the intervention. Future studies might consider incorporating more interactive elements, such as peer interactions, to improve the overall effectiveness of the gamified approach. All outcomes rely on self-reported measures or platform-collected data. As a pragmatic trial, medical diagnoses and pain medication details were not clinically verified, which may introduce reporting bias. Participants were recruited online, potentially favoring technologically adept individuals.

Finally, the distribution inequities of digital health care between rural and urban areas [64] and between younger and older adults in China limit the broader feasibility of this study, especially regarding the acceptability of digital games among older adults. Future studies should consider supporting initiative and confidence while preserving intrinsic capacity among older adults and individuals with digital illiteracy [65,66] by providing on-site training and multiple resources, including video-, paper-, or audio-based materials on technology use. These factors may limit generalizability to populations with lower digital literacy or limited internet access. In addition, due to the inherently visible nature of gamified interventions, participants in the intervention group would inevitably recognize their exposure to the gamified program, which may introduce expectation bias or placebo effects.

Conclusions

This study aims to determine the impacts of online GPM on adults' pain functioning and well-being. By leveraging a rigorous single-blind, randomized controlled trial design with standardized digital delivery, we provide a framework for assessing the efficacy of gamification in chronic pain management. Our findings may offer critical evidence on how gamified interventions can enhance engagement, adherence, and clinical outcomes, potentially informing future digital health strategies for chronic pain. Furthermore, the implementation of a fully remote trial demonstrates the feasibility and scalability of virtual interventions while ensuring accessibility for diverse populations. Ultimately, this research may help establish gamification as an effective tool for improving self-management in chronic pain, with implications for both clinical practice and digital therapeutic development.

Acknowledgments

The authors would like to thank all the physicians and clinicians who provided support and consultation during the revision of the intervention design. The authors also disclose the use of ChatGPT (OpenAI) [67] for refining the language and improving the readability of the manuscript.

XSL•FO
RenderX

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

MYMT and TTOK conceptualized the study and provided overall project supervision. JH designed the methodology and is responsible for data analysis, visualization, and drafting the initial manuscript. Furthermore, all authors (including SH) contributed to reviewing and editing the manuscript and provided final approval.

Conflicts of Interest

None declared.

References

1. GBD 2017 Mortality Collaborators. Global, regional, and national age-sex-specific mortality and life expectancy, 1950-2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet 2018 Nov 10;392(10159):1684-1735 [FREE Full text] [doi: 10.1016/S0140-6736(18)31891-9] [Medline: 30496102]
2. Hadi MA, McHugh GA, Closs SJ. Impact of chronic pain on patients' quality of life: a comparative mixed-methods study. J Patient Exp 2019 Jun;6(2):133-141 [FREE Full text] [doi: 10.1177/2374373518786013] [Medline: 31218259]
3. Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of Physicians, Denberg TD, et al. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the american college of physicians. Ann Intern Med 2017 Apr 04;166(7):514-530 [FREE Full text] [doi: 10.7326/M16-2367] [Medline: 28192789]
4. Low back pain and sciatica in over 16s: assessment and management. National Institute for Health and Care Excellence. URL: https://www.nice.org.uk/guidance/ng59/resources/low-back-pain-and-sciatica-in-over-16s-assessment-and-management-pdf-1837521693637 [accessed 2025-05-29]
5. Fishman SM, Young HM, Lucas Arwood E, Chou R, Herr K, Murinson BB, et al. Core competencies for pain management: results of an interprofessional consensus summit. Pain Med 2013 Jul;14(7):971-981 [FREE Full text] [doi: 10.1111/pme.12107] [Medline: 23577878]
6. Aiyegbusi OL, Cruz Rivera S, Oliver K, Manna E, Collis P, King-Kallimanis BL, et al. The opportunity for greater patient and public involvement and engagement in drug development and regulation. Nat Rev Drug Discov 2023 May 28;22(5):337-338. [doi: 10.1038/d41573-023-00031-x] [Medline: 36854837]
7. Powell SM, Larsen CA, Phillips SM, Pellegrini CA. Exploring beliefs and preferences for reducing sedentary behavior among adults with symptomatic knee osteoarthritis or knee replacement. ACR Open Rheumatol 2021 Jan 05;3(1):55-62 [FREE Full text] [doi: 10.1002/acr2.11216] [Medline: 33400397]
8. Moreno-Ligero M, Moral-Munoz JA, Salazar A, Failde I. mHealth intervention for improving pain, quality of life, and functional disability in patients with chronic pain: systematic review. JMIR Mhealth Uhealth 2023 Feb 02;11:e40844 [FREE Full text] [doi: 10.2196/40844] [Medline: 36729570]
9. Vassilev I, Rogers A, Sanders C, Kennedy A, Blickem C, Protheroe J, et al. Social networks, social capital and chronic illness self-management: a realist review. Chronic Illn 2011 Mar;7(1):60-86. [doi: 10.1177/1742395310383338] [Medline: 20921033]
10. Fatoye F, Gebrye T, Fatoye C, Mbada CE, Olaoye MI, Odole AC, et al. The clinical and cost-effectiveness of telerehabilitation for people with nonspecific chronic low back pain: randomized controlled trial. JMIR Mhealth Uhealth 2020 Jun 24;8(6):e15375 [FREE Full text] [doi: 10.2196/15375] [Medline: 32357128]
11. Bennell KL, Marshall CJ, Dobson F, Kasza J, Lonsdale C, Hinman RS. Does a web-based exercise programming system improve home exercise adherence for people with musculoskeletal conditions?: a randomized controlled trial. Am J Phys Med Rehabil 2019 Oct;98(10):850-858. [doi: 10.1097/PHM.0000000000001204] [Medline: 31021823]
12. Chew MT, Chan C, Kobayashi S, Cheng HY, Wong TM, Nicholson LL. Online pain management programs for chronic, widespread musculoskeletal conditions: a systematic review with meta-analysis. Pain Pract 2023 Jul 13;23(6):664-683. [doi: 10.1111/papr.13227] [Medline: 37051894]
13. Fernandes LG, Devan H, Fioratti I, Kamper SJ, Williams CM, Saragiotto BT. At my own pace, space, and place: a systematic review of qualitative studies of enablers and barriers to telehealth interventions for people with chronic pain. Pain 2022 Feb 01;163(2):e165-e181. [doi: 10.1097/j.pain.0000000000002364] [Medline: 34433776]
14. Sailer M, Homner L. The gamification of learning: a meta-analysis. Educ Psychol Rev 2019 Aug 15;32(1):77-112. [doi: 10.1007/s10648-019-09498-w]

XSL•FO
RenderX

15. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
16. Ning Y, Jia Z, Zhu R, Ding Y, Wang Q, Han S. Effect and feasibility of gamification interventions for improving physical activity and health-related outcomes in cancer survivors: an early systematic review and meta-analysis. Support Care Cancer 2022 Dec 30;31(1):92. [doi: 10.1007/s00520-022-07550-0] [Medline: 36585478]
17. Hamari J, Koivisto J, Sarsa H. Does gamification work?: a literature review of empirical studies on gamification. In: Proceedings of the 47th Hawaii International Conference on System Sciences. 2014 Presented at: HICSS '14; January 6-9, 2014; Waikoloa, HI p. 3025-3034 URL: https://ieeexplore.ieee.org/document/6758978 [doi: 10.1109/hicss.2014.377]
18. Saragih ID, Suarilah I, Saragih IS, Lin Y, Lin C. Efficacy of serious games for chronic pain management in older adults: a systematic review and meta-analysis. J Clin Nurs 2024 Mar 30;33(3):1185-1194. [doi: 10.1111/jocn.17012] [Medline: 38291564]
19. Moseley L. Unraveling the barriers to reconceptualization of the problem in chronic pain: the actual and perceived ability of patients and health professionals to understand the neurophysiology. J Pain 2003 May;4(4):184-189 [FREE Full text] [doi: 10.1016/s1526-5900(03)00488-7] [Medline: 14622702]
20. Veehof MM, Trompetter HR, Bohlmeijer ET, Schreurs KM. Acceptance- and mindfulness-based interventions for the treatment of chronic pain: a meta-analytic review. Cogn Behav Ther 2016 Jan 28;45(1):5-31 [FREE Full text] [doi: 10.1080/16506073.2015.1098724] [Medline: 26818413]
21. Hochheim M, Ramm P, Amelung V. The effectiveness of low-dosed outpatient biopsychosocial interventions compared to active physical interventions on pain and disability in adults with nonspecific chronic low back pain: a systematic review with meta-analysis. Pain Pract 2023 Apr 09;23(4):409-436. [doi: 10.1111/papr.13198] [Medline: 36565010]
22. Hrkać A, Bilić D, Černy-Obrdalj E, Baketarić I, Puljak L. Comparison of supervised exercise therapy with or without biopsychosocial approach for chronic nonspecific low back pain: a randomized controlled trial. BMC Musculoskelet Disord 2022 Nov 08;23(1):966 [FREE Full text] [doi: 10.1186/s12891-022-05908-3] [Medline: 36348309]
23. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 08;346(jan08 15):e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
24. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/BF03193146]
25. Nambi G, Abdelbasset WK, Elsayed SH, Alrawaili SM, Abodonya AM, Saleh AK, et al. Comparative effects of isokinetic training and virtual reality training on sports performances in university football players with chronic low back pain-randomized controlled study. Evid Based Complement Alternat Med 2020 Jun 16;2020(1):2981273 [FREE Full text] [doi: 10.1155/2020/2981273] [Medline: 32617104]
26. Nambi G, Abdelbasset WK, Alrawaili SM, Alsubaie SF, Abodonya AM, Saleh AK. Virtual reality or isokinetic training; its effect on pain, kinesiophobia and serum stress hormones in chronic low back pain: a randomized controlled trial. Technol Health Care 2021 Jan 15;29(1):155-166. [doi: 10.3233/thc-202301]
27. Carvalho MS, Carvalho LC, Menezes FD, Frazin A, Gomes ED, Iunes DH. Effects of exergames in women with fibromyalgia: a randomized controlled study. Games Health J 2020 Oct;9(5):358-367. [doi: 10.1089/g4h.2019.0108] [Medline: 32379982]
28. Tuck N, Pollard C, Good C, Williams C, Lewis G, Hames M, et al. Active virtual reality for chronic primary pain: mixed methods randomized pilot study. JMIR Form Res 2022 Jul 13;6(7):e38366 [FREE Full text] [doi: 10.2196/38366] [Medline: 35830224]
29. Groenveld TD, Smits ML, Knoop J, Kallewaard JW, Staal JB, de Vries M, et al. Effect of a behavioral therapy-based virtual reality application on quality of life in chronic low back pain. Clin J Pain 2023 Jun 01;39(6):278-285 [FREE Full text] [doi: 10.1097/AJP.0000000000001110] [Medline: 37002877]
30. Lin YT, Lee WC, Hsieh RL. Active video games for knee osteoarthritis improve mobility but not WOMAC score: a randomized controlled trial. Ann Phys Rehabil Med 2020 Nov;63(6):458-465 [FREE Full text] [doi: 10.1016/j.rehab.2019.11.008] [Medline: 31981832]
31. Park JH, Lee SH, Ko DS. The effects of the Nintendo Wii exercise program on chronic work-related low back pain in industrial workers. J Phys Ther Sci 2013 Aug;25(8):985-988 [FREE Full text] [doi: 10.1589/jpts.25.985] [Medline: 24259899]
32. Tejera D, Beltran-Alacreu H, Cano-de-la-Cuerda R, Leon Hernández JV, Martín-Pintado-Zugasti A, Calvo-Lobo C, et al. Effects of virtual reality versus exercise on pain, functional, somatosensory and psychosocial outcomes in patients with non-specific chronic neck pain: a randomized clinical trial. Int J Environ Res Public Health 2020 Aug 16;17(16):5950 [FREE Full text] [doi: 10.3390/ijerph17165950] [Medline: 32824394]
33. Eccleston C, Fisher E, Liikkanen S, Sarapohja T, Stenfors C, Jääskeläinen SK, et al. A prospective, double-blind, pilot, randomized, controlled trial of an "embodied" virtual reality intervention for adults with low back pain. Pain 2022 Sep 01;163(9):1700-1715 [FREE Full text] [doi: 10.1097/j.pain.0000000000002617] [Medline: 35324507]
34. Mueller J, Niederer D, Tenberg S, Oberheim L, Moesner A, Mueller S. Acute effects of game-based biofeedback training on trunk motion in chronic low back pain: a randomized cross-over pilot trial. BMC Sports Sci Med Rehabil 2022 Nov 13;14(1):192 [FREE Full text] [doi: 10.1186/s13102-022-00586-z] [Medline: 36372870]

XSL•FO
RenderX

35. Matheve T, Bogaerts K, Timmermans A. Virtual reality distraction induces hypoalgesia in patients with chronic low back pain: a randomized controlled trial. J Neuroeng Rehabil 2020 Apr 22;17(1):55 [FREE Full text] [doi: 10.1186/s12984-020-00688-0] [Medline: 32321516]
36. Villafaina S, Collado-Mateo D, Domínguez-Muñoz FJ, Fuentes-García JP, Gusi N. Benefits of 24-week exergame intervention on health-related quality of life and pain in women with fibromyalgia: a single-blind, randomized controlled trial. Games Health J 2019 Dec 01;8(6):380-386. [doi: 10.1089/g4h.2019.0023] [Medline: 31259617]
37. Lewis GN, Bean DJ. What influences outcomes from inpatient multidisciplinary pain management programs?: a systematic review and meta-analysis. Clin J Pain 2021 Jul 01;37(7):504-523. [doi: 10.1097/AJP.0000000000000941] [Medline: 33883414]
38. Lachenbruch PA, Cohen J. Statistical power analysis for the behavioral sciences (2nd ed.). J Am Stat Assoc 1989 Dec;84(408):1096. [doi: 10.2307/2290095]
39. Home page. Changsha Ranxing Information Technology. URL: https://www.wjx.cn/ [accessed 2025-05-29]
40. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). Pain 2019 Jan;160(1):19-27. [doi: 10.1097/j.pain.0000000000001384] [Medline: 30586067]
41. Ersek M, Turner JA, Cain KC, Kemp CA. Results of a randomized controlled trial to examine the efficacy of a chronic pain self-management group for older adults [ISRCTN11899548]. Pain 2008 Aug 15;138(1):29-40 [FREE Full text] [doi: 10.1016/j.pain.2007.11.003] [Medline: 18086516]
42. Ibrahim AA, Akindele MO, Ganiyu SO. Effectiveness of patient education plus motor control exercise versus patient education alone versus motor control exercise alone for rural community-dwelling adults with chronic low back pain: a randomised clinical trial. BMC Musculoskelet Disord 2023 Feb 23;24(1):142 [FREE Full text] [doi: 10.1186/s12891-022-06108-9] [Medline: 36823567]
43. Chimenti RL, Post AA, Rio EK, Moseley GL, Dao M, Mosby H, et al. The effects of pain science education plus exercise on pain and function in chronic Achilles tendinopathy: a blinded, placebo-controlled, explanatory, randomized trial. Pain 2023 Jan 01;164(1):e47-e65 [FREE Full text] [doi: 10.1097/j.pain.0000000000002720] [Medline: 36095045]
44. Schmidt-Kraepelin M, Thiebes S, Tran MC, Sunyaev A. What's in the game? Developing a taxonomy of gamification concepts for health apps. In: Proceedings of the 51st Hawaii International Conference on System Sciences. 2018 Jan 03 Presented at: HICSS '18; January 3-6, 2018; Koloa, HI p. 1217-1226 URL: https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/93fa9001-cd42-4379-b4d7-607c2d223d40/content [doi: 10.24251/hicss.2018.150]
45. Buckley J, DeWille T, Exton C, Exton G, Murray L. A gamification–motivation design framework for educational software developers. J Educ Technol Syst 2018 Jun 20;47(1):101-127. [doi: 10.1177/0047239518783153]
46. Koestner R, Hope N. A self-determination theory approach to goals. In: Gagné M, editor. The Oxford Handbook of Work Engagement, Motivation, and Self-Determination Theory. Oxford, UK: Oxford University Press; 2014:400-413.
47. Home page. VX. URL: https://editor.ivx.cn/ [accessed 2025-05-29]
48. McMahon SB, Koltzenburg M, Tracey I, Turk D. Wall & Melzack's Textbook of Pain. New York, NY: Saunders; 2013.
49. Webb J, Stewart D. Let's Move with Leon. A randomised controlled trial of a UK digital intervention to improve physical activity in people with a musculoskeletal condition. Public Health 2023 Apr;217:125-132 [FREE Full text] [doi: 10.1016/j.puhe.2023.01.030] [Medline: 36878121]
50. Vambheim SM, Kyllo TM, Hegland S, Bystad M. Relaxation techniques as an intervention for chronic pain: a systematic review of randomized controlled trials. Heliyon 2021 Aug;7(8):e07837 [FREE Full text] [doi: 10.1016/j.heliyon.2021.e07837] [Medline: 34485731]
51. Wang S, Yao S, Shang L, Xu C, Ma J. Validation of the Chinese version of the Brief Pain Inventory in patients with knee osteoarthritis. J Orthop Surg Res 2023 Sep 23;18(1):720 [FREE Full text] [doi: 10.1186/s13018-023-04218-1] [Medline: 37742029]
52. Pei L, Xia J, Yan J. Cross-cultural adaptation, reliability and validity of the Chinese version of the Fear Avoidance Beliefs Questionnaire. J Int Med Res 2010 Dec 01;38(6):1985-1996 [FREE Full text] [doi: 10.1177/147323001003800612] [Medline: 21227002]
53. Xiong Y. Validation of a simplified Chinese version of the pain catastrophizing scale and an exploration of the factors predicting catastrophizing in pain clinic patients. Pain Phys 2015 Nov 15;6(18):E1059-E1072. [doi: 10.36076/ppj.2015/18/e1059]
54. Amireault S, Godin G, Lacombe J, Sabiston CM. Validation of the Godin-Shephard leisure-time physical activity questionnaire classification coding system using accelerometer assessment among breast cancer survivors. J Cancer Surviv 2015 Sep;9(3):532-540. [doi: 10.1007/s11764-015-0430-6] [Medline: 25666749]
55. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry 2014 Sep;36(5):539-544. [doi: 10.1016/j.genhosppsych.2014.05.021] [Medline: 25023953]
56. Zeng QZ, He YL, Liu H, Miao JM, Chen JX. Reliability and validity of Chinese version of the Generalized Anxiety Disorder 7-item (GAD-7) scale in screening anxiety disorders in outpatients from traditional Chinese internal department. Chin Ment Health J 2013;27(3):163-168 [FREE Full text]

XSL•FO
RenderX

57. Liu GG, Wu H, Li M, Gao C, Luo N. Chinese time trade-off values for EQ-5D health states. Value Health 2014 Jul;17(5):597-604 [FREE Full text] [doi: 10.1016/j.jval.2014.05.007] [Medline: 25128053]
58. Phan MH, Keebler JR, Chaparro BS. The development and validation of the Game User Experience Satisfaction Scale (GUESS). Hum Factors 2016 Dec 27;58(8):1217-1247. [doi: 10.1177/0018720816669646] [Medline: 27647156]
59. Home page. Tencent Meeting. URL: https://meeting.tencent.com/ [accessed 2025-05-29]
60. Harvie DS, Kelly J, Kluver J, Deen M, Spitzer E, Coppieters MW. A randomized controlled pilot study examining immediate effects of embodying a virtual reality superhero in people with chronic low back pain. Disabil Rehabil Assist Technol 2024 Apr 18;19(3):851-858 [FREE Full text] [doi: 10.1080/17483107.2022.2129846] [Medline: 36256688]
61. Stenberg N, Gillison F, Rodham K. How do peer support interventions for the self-management of chronic pain, support basic psychological needs? A systematic review and framework synthesis using self-determination theory. Patient Educ Couns 2022 Nov;105(11):3225-3234 [FREE Full text] [doi: 10.1016/j.pec.2022.07.017] [Medline: 35985906]
62. Blok AC, Valley TS, Abbott P. Gamification for family engagement in lifestyle interventions: a systematic review. Prev Sci 2021 Oct 30;22(7):831-844 [FREE Full text] [doi: 10.1007/s11121-021-01214-x] [Medline: 33786746]
63. Lee S, Small BJ, Cawthon PM, Stone KL, Almeida DM. Social activity diversity as a lifestyle factor to alleviate loneliness and chronic pain. J Psychosom Res 2023 Sep;172:111434 [FREE Full text] [doi: 10.1016/j.jpsychores.2023.111434] [Medline: 37422980]
64. Wan X, Ao Y, Huang Z, Yu M. Digital health services and rural healthcare access: evidence from China. Health Policy Technol 2026 Feb;15(1):101123. [doi: 10.1016/j.hlpt.2025.101123]
65. Möller J, Stara V, Amabili G, Barbarossa F, Riccardi GR, Martella C, et al. Toward innovation in healthcare: an analysis of the digital behavior of older people in Europe and Japan for the introduction of a technological coaching system. Healthcare (Basel) 2024 Jan 08;12(2):143 [FREE Full text] [doi: 10.3390/healthcare12020143] [Medline: 38255032]
66. Yongai J, Wenbo H, Yang F. Internet use and the life of older adults aged 50 and above in digital era: findings from a national survey. Popul Res 2024;48(1):40-55 [FREE Full text]
67. ChatGPT. OpenAI. URL: https://chatgpt.com/ [accessed 2025-05-29]

Abbreviations

SDT: self-determination theory
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by J Sarvestan; submitted 12.Jul.2025; peer-reviewed by G Amabili; comments to author 28.Oct.2025; accepted 28.Nov.2025; published 13.Jan.2026.

Please cite as:
Tse MYM, He J, Kwok TTO, Ho S
An Online Gamified Pain Management Program for Adults With Chronic Pain: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e80542
URL: https://www.researchprotocols.org/2026/1/e80542
doi:10.2196/80542
PMID:

©Mun Yee Mimi Tse, Jiafan He, Tyrone Tai On Kwok, Sukki Ho. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Evaluation of a Tailored Multifaceted Pharmaceutical Care Intervention to Optimize Chronic Obstructive Pulmonary Disease Management: Protocol for a Cluster Randomized Controlled Trial

Xuedi Ma[1], PhD; Xiaocong Li[2], PhD; Xinyi Li[2], MPharm; Hui Chen[2], MD; Yang Wang[1], PhD; Xuan Jia[1], BM; Yanyan Zhao[1], MD; Wei Li[1], PhD; Lihong Liu[2], PhD

[1]Fuwai Hospital, National Center for Cardiovascular Diseases, Peking Union Medical College & Chinese Academy of Medical Sciences, Beijing, China
[2]Department of Pharmacy, China-Japan Friendship Hospital, Beijing, China

Corresponding Author:
Lihong Liu, PhD
Department of Pharmacy, China-Japan Friendship Hospital
No. 2, Yinghuayuan East Street, Chaoyang District
Beijing, 100029
China
Phone: 86 010 84205559
Email: llh-hong@outlook.com

Abstract

Background: Despite the widespread use of inhalation therapy, patients with chronic obstructive pulmonary disease (COPD) frequently experience suboptimal disease control due to medication nonadherence, improper inhaler use technique, and inappropriate device selection, which collectively impair health-related quality of life (HRQoL). Pharmacist-led interventions may help address these gaps. Interventions based on the information-motivation-behavioral skills model and supported by digital tools can improve adherence and self-management. This study evaluates the efficacy of a multifaceted pharmaceutical care intervention for COPD delivered through digital tool support.

Objective: The primary objective is to compare the change in HRQoL, measured using the St George's Respiratory Questionnaire, between the intervention and control groups from baseline to 12 months. Secondary objectives are to assess changes in medication adherence (Test of Adherence to Inhalers), quality of life (EQ-5D-5L), COPD-related medical costs, and patient-reported pharmacy service experience.

Methods: This 1-year cluster randomized controlled trial evaluates a multifaceted pharmaceutical care intervention in adults with moderate to very severe COPD (Global Initiative for Chronic Obstructive Lung Disease, stages 2-4) who have confirmed suboptimal inhaler practice but remain matched to an appropriate inhaler device based on peak inspiratory flow rate measured using a digital tool. In total, 34 hospital-based cough and wheeze pharmaceutical care clinics were recruited and randomized (1:1) to either an intervention or a control group. The target sample size is 15 patients per site. Participants in the intervention group will receive tailored support using electronic adherence monitoring, inhaler use technique assessments, and peak inspiratory flow rate to optimize device selection and self-management. Participants in the control group will receive usual pharmaceutical care. Descriptive statistics will be used to summarize participant characteristics and outcomes. Linear mixed effects models will be used to compare primary and secondary outcomes. Subgroup analyses will explore effects by age, sex, education level, place of residence, and smoking status. Pharmacy service survey data will be analyzed qualitatively.

Results: The trial was registered on July 15, 2024. Recruitment started on November 9, 2024, and enrollment was completed by December 31, 2025. As of December 31, 2025, we enrolled 454 participants, of whom 16 (3.5%) had completed the 12-month follow-up. The trial is expected to be completed by December 31, 2026, with results planned for publication in 2027.

Conclusions: This multifaceted, pharmacist-led pharmaceutical care intervention may provide a scalable model for improving COPD management and HRQoL. If effective, the digitally supported program, grounded in the information-motivation-behavioral skills model, could be implemented in more than 1000 cough and wheeze pharmaceutical care clinics nationwide.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400086943; https://tinyurl.com/75a9phbw

International Registered Report Identifier (IRRID): DERR1-10.2196/82806

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e82806) doi:10.2196/82806

KEYWORDS

cluster randomized controlled trial; chronic obstructive pulmonary disease; pharmaceutical care; health-related quality of life; medication adherence

Introduction

Background and Rationale

Chronic obstructive pulmonary disease (COPD) poses a substantial global public health burden because of its high prevalence and economic impact [1]. In China, approximately 100 million individuals are affected [2], and the country faces the largest projected COPD-related economic burden globally from 2020 to 2050 [3]. Without effective interventions, new COPD cases and deaths in China may increase 1.5-fold over the next 25 years [4].

Inhalation therapy, the cornerstone of COPD management, relies critically on correct device technique and sustained medication adherence to achieve optimal disease control [5,6]. However, suboptimal practices persist: inhaler use technique errors occur in 50% to 100% of patients [7-10], nonadherence rates range from 50% to 80% [7-10], and inappropriate device selection affects more than 80% of patients [11]. These factors collectively undermine treatment efficacy, leading to poor disease control and impaired health-related quality of life (HRQoL) [12-14]. Therefore, it is essential to ensure the appropriate choice of inhalation devices, provide training on inhaler use, and promote medication adherence in COPD management.

Pharmacist-led interventions offer promise in addressing these gaps. Meta-analyses confirm their effectiveness in improving inhalation techniques (risk ratio=1.85, 95% CI 1.57-2.17; *P*<.0001) and medication adherence (risk ratio=1.41, 95% CI 1.24-1.61; *P*<.0001) [15-19]. In addition, this approach has been found to be cost-effective, with a mean cost saving of £671.59 (US $906.65) per patient (95% CI −£1584.73 [−US $2139.39] to −£68.14 [−US $91.99]; *P*<.0001; GBP £1 [US $1.35]) [20]. Notably, interventions grounded in the information-motivation-behavioral skills (IMB) model that facilitates health behaviors show enhanced adherence outcomes [21,22]. Consequently, IMB model−based interventions may be worthy of clinical promotion and application [15]. Concurrently, digital tools (eg, real-time adherence monitoring and telehealth) have demonstrated efficacy in promoting self-management and HRQoL among patients with COPD [23-26].

Although previous studies have explored the clinical effectiveness and cost-effectiveness of medication management for COPD, several aspects need to be addressed in the future studies. First, there is limited research that explored the impact of medication management on HRQoL among patients with COPD in the hospital-based pharmaceutical clinic setting; existing studies have focused primarily on behavioral outcomes [15-20]. If such pharmacist-led interventions are to be implemented in clinical settings, it is crucial to comprehensively understand their impact on patients' health. Second, insufficient attention has been paid to device-patient mismatch despite its high prevalence [27,28]. Third, integrated interventions that combine the IMB theory with digital tools for real-world implementation remain scarce. To improve disease management and deliver patient-centered care that enhances patients' well-being, it is essential to leverage these advancements in medication management and to tailor approaches to address the most prevalent medication use issues.

Objective

This cluster randomized controlled trial evaluates a digitally enhanced, multifaceted pharmaceutical care intervention delivered across 34 cough and wheeze pharmaceutical care (CWPC) clinics. This will be a parallel-group superiority trial with 1:1 allocation to the intervention and control arms. This study aims to assess the impact of intervention on HRQoL in patients with moderate to very severe COPD who have confirmed suboptimal inhaler practice but remain matched to an appropriate inhaler device. It addresses evidence gaps through theory-driven, digitally enhanced management of medication adherence, inhaler use technique, and device selection. The findings aim to support the broader implementation of multifaceted pharmaceutical care models for COPD across more than 1000 CWPC clinics in 29 provinces and municipalities in China.

Methods

Study Setting

The multicenter trial will be conducted in 34 certified CWPC clinics embedded in secondary or tertiary hospitals across mainland China. Participating sites are required to (1) hold CWPC accreditation with at least 2 licensed pharmacists trained in COPD management, (2) have access to all guideline-recommended inhalation therapies, and (3) have the capacity to implement protocolized interventions. A complete list of participating hospitals and their provinces and municipalities is provided in Table S1 in Multimedia Appendix 1. Thus, these accredited CWPC clinics, staffed by well-trained pharmacists, are well positioned to facilitate effective participant recruitment.

Eligibility Criteria

Recruited participants must meet the eligibility criteria listed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for participants.

Inclusion criteria (all criteria must be met simultaneously)

- Aged 18 years or older
- Diagnosed with chronic obstructive pulmonary disease, with severity of airflow obstruction from level 2 to level 4 (according to the Global Initiative for Chronic Obstructive Lung Disease grades) [6], defined as a forced expiratory volume in 1 second <80% of the predicted value after the use of bronchodilators (pulmonary function test results within 3 months could be used for screening)
- Suboptimal inhaler practice confirmed by at least 1 incorrect technique identified using a standard inhaler checklist
- Appropriate device-patient match (peak inspiratory flow rate–guided selection using digital sensors)
- Signed the informed consent form

Exclusion criteria

- Inability to obtain informed consent from the patient or family members
- Comorbidities precluding participation, including terminal illness (life expectancy <6 months), active tuberculosis or lung cancer, or cognitive or physical impairment affecting inhaler use
- Pregnancy or lactation
- Concurrent enrollment in other clinical trials
- Inability to complete the 12-month follow-up (eg, relocation plans)
- Other conditions in which the pharmacist who evaluated the participant is unable to comply with the trial

Interventions

Intervention Arm

The intervention was systematically developed using the IMB model and comprises 7 core components mapped to this theoretical framework. The 7 components are introduction to therapeutic regimens, diversified information support, inhaler use technique training, medication adherence management, medication management using a diary booklet, follow-up management, and interactive question-and-answer sessions (Table S2 in Multimedia Appendix 1). The digitally enhanced, multifaceted pharmaceutical care intervention is delivered during scheduled CWPC clinic visits at baseline and at months 1, 6, and 9 (Table 1). Each face-to-face intervention visit lasts approximately 30 minutes and follows a structured checklist (Table S2 in Multimedia Appendix 1).

XSL•FO
RenderX

Table 1. Schedule of enrollment, interventions, and assessments (SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials).

Time point	Baseline (−60 to 0 d; T0)	Month 1 (−14 to +14 d; T1)	Month 3 (−14 to +14 d; T2)	Month 6 (−14 to +14 d; T3)	Month 7 (−14 to +14 d; T4)	Month 9 (−14 to +14 d; T5)	Month 10 (−14 to +14 d; T6)	Month 12 (−14 to +14 d; T7)
Enrollment								
Eligibility screen	✓							
Informed consent	✓							
Allocation (cluster level)	✓							
Interventions								
Intervention arm	✓	✓		✓		✓		
Control arm: standard CWPC[a] pharmaceutical care	✓	✓		✓		✓		
Assessments								
Visit date	✓	✓	✓	✓	✓	✓	✓	✓
Telephone follow-up for questionnaire administration			✓		✓		✓	
Basic information	✓							
Medical history	✓							✓
Physical examination	✓							
Pulmonary function test	✓							✓
AECOPD[b] status	✓	✓		✓		✓		✓
SGRQ[c]	✓	✓		✓		✓		✓
TAI-10[d]	✓	✓		✓		✓		✓
EQ-5D-5L[e]	✓							✓
COPD[f] medication use in the past 2 wk	✓	✓		✓		✓		✓
COPD medication regimen at current visit	✓	✓		✓		✓		✓
Other concomitant medications in the past 2 wk	✓	✓		✓		✓		✓
Treatment costs of COPD	✓	✓		✓		✓		✓
Disease status		✓		✓		✓		✓
Pharmacy service survey								✓

[a]CWPC: cough and wheeze pharmaceutical care.

[b]AECOPD: acute exacerbations of chronic obstructive pulmonary disease.

[c]SGRQ: St George's Respiratory Questionnaire.

[d]TAI-10: Test of Adherence to Inhalers.

[e]EQ-5D-5L: European Quality of Life-5 Dimensions-5 Level version.

[f]COPD: chronic obstructive pulmonary disease.

Clinical pharmacists at intervention sites will receive standardized training before study initiation (Figure 1), including instruction on the trial rationale and objectives, COPD pharmacotherapy and inhaler use, and detailed guidance on implementing each intervention component. Pharmacists in the intervention arm will also be trained to use the digital inhaler sensor and the peak inspiratory flow rate assessment system to support inhaler use technique training (Figure 2) and device selection (see method 1 in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 1. Overview of the study stages and implementation timeline. IMB: information-motivation-behavioral skills; SOP: standard operating procedure.

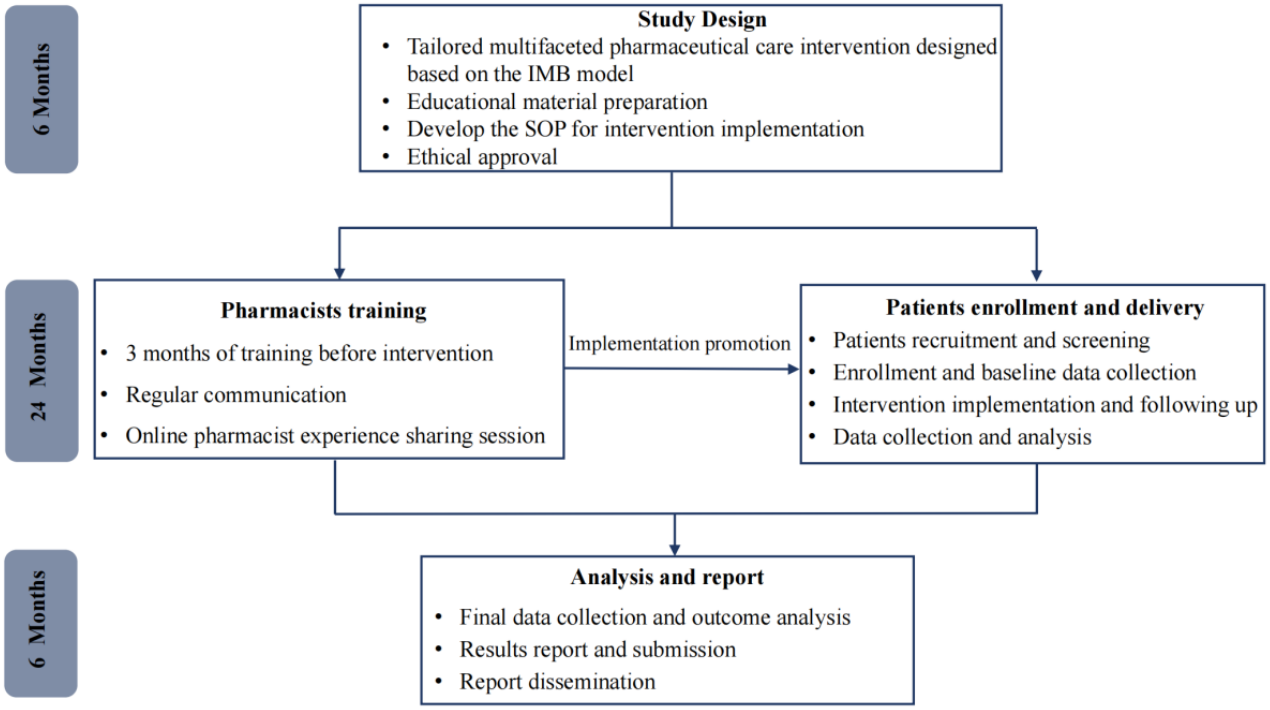

Figure 2. Illustration of the selection of digital inhaler sensor devices for inhaler use technique training.



Control Arm: Standard CWPC Pharmaceutical Care

Sites allocated to the control arm will continue to provide standard CWPC pharmaceutical care according to existing practice. Pharmacists at participating sites will be trained on the standardized operating procedures (SOPs) corresponding to their allocation (Table S3 in Multimedia Appendix 1). At baseline and at months 1, 6, and 9, pharmacists will typically spend approximately 10 minutes per consultation (Table 1). Standard care includes (1) introduction to the therapeutic regimen (drug names, indications, and basic dosing instructions); (2) counseling on dosage and administration; (3) reminders regarding key drug precautions and management of common adverse drug reactions; and (4) demonstration and training in correct inhaler use technique using routine educational tools.

XSL•FO
RenderX

Strategies to Improve and Monitor Adherence to Intervention Protocols

To promote consistency of care within each study arm and to minimize contamination between arms, SOPs have been developed for both the intervention and control pharmaceutical care pathways (Tables S2 and S3 in Multimedia Appendix 1). Pharmacists at all participating sites will be trained in the SOPs corresponding to their allocation.

During the trial, completion of each digital-enabled pharmaceutical care component (intervention arm) or standard care element (control arm) will be documented in the electronic data capture (EDC) system using structured checklists (Tables S2 and S3 in Multimedia Appendix 1), which serve as fidelity-monitoring tools. The coordinating center will conduct periodic remote audits of EDC entries and, when necessary, on-site or online supervision visits to provide feedback and reinforce protocol adherence. For patients, adherence to scheduled visits will be supported through automated appointment reminders delivered via the WeChat mini-program.

Concomitant Care and Cointerventions

All COPD pharmacological treatments, including the choice and adjustment of inhalation therapies, will be prescribed at the discretion of the treating physicians in accordance with current Global Initiative for Chronic Obstructive Lung Disease (GOLD) guidelines and national guidelines. Other routine medical care for COPD and comorbid conditions (eg, vaccinations, pulmonary rehabilitation, and management of cardiovascular disease or diabetes) will be allowed and recorded as concomitant treatments.

Study Outcomes

Primary Outcome

The St George's Respiratory Questionnaire (SGRQ) will be used to assess HRQoL. The primary end point is the between-group difference in change from baseline to 12 months in the SGRQ total score, measured at baseline and 12 months after initiation of the intervention. The SGRQ was originally developed by Jones et al [29] and consists of 50 items divided into 3 sections: symptoms, activities, and impact. The scoring range is from 0 to 100, with a lower score indicating better health status in patients with COPD. SGRQ will be administered at baseline and at months 1, 6, 9, and 12 (Table 1).

Secondary Outcomes

Medication Adherence

The Test of Adherence to Inhalers (TAI-10) will be used to evaluate adherence to inhalation therapy. The main adherence outcome will be the between-group difference in change from baseline to 12 months in the TAI-10 score. The TAI-10 questionnaire was designed to assess the adherence to inhaler use in patients with asthma or COPD [30] and consists of 12 items. Items 1 to 10 assess the patient's perspective and are scored on a 5-point scale (1=worst and 5=best). Items 11 and 12 reflect health care professional assessments and are scored on a 2-point scale (1=bad and 2=good). TAI assessments will be conducted according to the schedule shown in Table 1.

Generic Quality of Life

Generic health status will be assessed using the EQ-5D-5L, which covers 5 dimensions (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression), each rated on 5 levels from *no problems* to *extreme problems* [31]. We calculated the EQ-5D-5L index using the value sets developed by Luo et al [32], with 1 representing the value of instrument-defined full health. The main EQ-5D-5L outcome will be the between-group difference in change from baseline to 12 months in the utility index. Measurement time points are shown in Table 1.

Pharmacy Service Survey

A pharmacy service survey will be used to assess patients' perceived improvements in COPD-related and medication-related knowledge, self-reported quality of life, mastery of correct inhaler use after participation, and adequacy of communication with pharmacists (Table S4 in Multimedia Appendix 1). This survey will be administered at month 12 (Table 1).

Medical Cost

The medical cost will be collected using a microcosting approach [33] from the patient-level direct medical cost perspective over the 12-month trial period. At each scheduled visit, COPD-related direct medical expenses paid by patients will be recorded, including registration fees, examination costs, medication expenses, and other COPD-related medical service fees, when applicable. The primary economic outcome will be cumulative COPD-related direct medical costs per participant over 12 months, with the detailed data collection schedule provided in Table 1.

Participant Timeline

The participant timeline is summarized in Table 1, which outlines the schedule of enrollment, assigned care delivery, and outcome assessments over the 12-month study period. A schematic overview of the study stages and timing is provided in Figure 1.

Sample Size

The sample size was determined for the primary outcome, change in SGRQ total score from baseline to 12 months, using a cluster randomized superiority design for 2 means (PASS, version 2021; NCSS LLC). We assumed a between-group difference (intervention minus control) in SGRQ change of 4 units (the minimal clinically important difference), a common SD of 10 units, an intracluster correlation coefficient of 0.01, an average cluster size of 15 participants per community-based primary care clinic, 80% power, and a 2-sided α of .05. Under these assumptions, 8 clinics per arm (16 in total) were required. Allowing for 20% loss to follow-up, we planned to recruit at least 10 clinics per arm.

Ultimately, 34 clinics (17 per arm) were enrolled, with a target of 15 participants per clinic (planned N=510). Assuming up to 20% attrition (≥408 analyzable participants, approximately 204 per arm), this sample provides >90% power to detect a 4-unit difference even if the intracluster correlation coefficient increases to 0.03. Details of the sample size formulas and

XSL•FO
RenderX

calculations are provided in Multimedia Appendix 1 (see methods 2).

Randomization and Blinding

Hospitals (clusters) will be stratified by their level (secondary or tertiary). Within each stratum, hospitals will be treated as clusters and randomized to either the intervention or the control group in a 1:1 ratio using simple random sampling. The randomization sequence will be generated by an independent statistician using a computerized random number generator. Randomization will be implemented centrally by the coordinating center. The allocation list will be kept in a password-protected file accessible only to the independent statistician and the designated trial coordinator. At each site, pharmacists will screen and enroll individual participants after cluster allocation has been assigned. Cluster allocation will be revealed to each participating hospital only after the data analysis has been completed.

Because of the nature of our interventions, the trial was registered as "single blind" (ChiCTR2400086943) presented in Multimedia Appendix 2. Pharmacists delivering the intervention cannot be blinded, and full participant blinding is not feasible because of differences in content and intensity of care. However, participants will not be explicitly informed of their clinic's allocation. Outcome assessors and data analysts will remain blinded to group allocation. Group codes will be masked in the database, and analyses will be conducted using anonymized group labels (group A and group B) until the primary analysis is finalized. Unblinding of outcome assessors or data analysts is not planned during the trial.

Data Collection and Management

Outcome, baseline, and other trial data will be collected at prespecified time points (Table 1). Data sources include hospital medical records and participant-reported questionnaires, administered in person at baseline and months 1, 6, 9, and 12, and by telephone at months 3, 7, and 10. Data will be recorded in an electronic case report form and stored in the EDC system by trained site pharmacists. EDC access will be role based and password protected.

Data quality will be promoted through standardized training on questionnaire administration and electronic case report form completion, the use of predefined variable definitions and built-in logic checks in the EDC system, and routine source data verification by an independent pharmacist at each site, supplemented by periodic central data monitoring and query resolution by the coordinating center. Standardized and validated instruments (SGRQ, TAI, and EQ-5D-5L) will be used as described earlier, and study-specific forms (pharmacy service survey) are provided in Table S4 in Multimedia Appendix 1.

To promote retention and complete follow-up, appointment reminders will be sent via the WeChat mini-program, and missed visits will be followed up by telephone. Participants who discontinue or deviate from the assigned care pathway will be encouraged to remain in the study for outcome assessment.

Statistical Analysis

Baseline Characteristics

Baseline characteristics will be summarized by study arm using mean (SD) or median (IQR) for continuous variables and counts (%) for categorical variables. All statistical tests will be 2 sided, with a significance level of P=.05. Effect estimates will be reported with 95% CIs.

Analysis Populations

The primary analysis will follow the intention-to-treat principle, including all enrolled participants who provided informed consent, and analyzed according to the allocation of their recruiting clusters, regardless of intervention exposure or protocol deviations.

Primary Outcome Analysis

The primary end point (between-group difference in change from baseline to 12 months in SGRQ total score) will be analyzed using a linear mixed effects model accounting for clustering by site and adjusting for baseline SGRQ, cluster effects, and prespecified baseline characteristics. Treatment effect will be expressed as the adjusted between-group difference in change.

Secondary and Other Outcome Analysis

Continuous secondary outcomes (eg, TAI-10 and EQ-5D-5L utility index) will be analyzed using linear mixed effects models, adjusting for baseline value and cluster effects. Results of the pharmacy service survey will be summarized using descriptive statistics (proportions).

Subgroup Analyses

Subgroup analyses for the primary outcome will be conducted based on the following variables: age (<65 vs ≥65 years), sex (male vs female), education level (high school or below vs college or university or above), place of residence (rural vs urban), COPD severity (moderate vs severe vs very severe), smoking status (current or former smoker vs never smoker), and baseline number of acute exacerbations of COPD (0 vs ≥1). Interaction terms for group and these covariates will be included in the analysis.

Missing Data and Sensitivity Analyses

In the primary analysis, missing data for the primary outcome (SGRQ) will be handled using the last observation carried forward method. To assess the robustness of the primary outcome to different assumptions about missing data and analysis populations, several sensitivity analyses will be conducted, including (1) reanalysis of the primary outcome in the per-protocol population, (2) complete-case analysis without imputation, and (3) multiple imputation for missing SGRQ data.

Economic Evaluation

The EQ-5D-5L questionnaire will be used to obtain the health utility values for each participant before and after the intervention [32,34]. These values will be used to calculate the change in quality-adjusted life years (QALYs) over the 12-month follow-up for both groups [32,34]. The incremental cost-effectiveness ratio will be calculated as incremental cost

XSL•FO
RenderX

divided by incremental QALYs. The willingness-to-pay per QALY threshold was set at 128,000 RMB (US $36,364) per QALY, equivalent to 1.76 times the gross domestic product per capita in China [35]. A 1-way sensitivity analysis will be conducted to assess the robustness of the model and the reliability of the results. All analyses will be performed using SAS (version 9.4; SAS Institute Inc).

Data Monitoring and Quality Assurance

The trial will be overseen by a steering committee convened by the coordinating center (China-Japan Friendship Hospital), comprising clinical experts in COPD, pharmacologists, and epidemiologists. Given the low-risk nature of the intervention (pharmacist-led care enhancement) and the absence of planned interim efficacy analyses, a separate data monitoring committee will not be established. Trial monitoring will be conducted by the coordinating center's monitoring team, which is independent of site investigators and reports to the steering committee.

The EDC system will automatically validate the input data. If the recorded values fall outside the acceptable range, a prompt will appear, requesting the user to double-check the entry. Each research center will have an independent pharmacist responsible for conducting source data verification. Additionally, the lead organization will provide real-time reminders to each center regarding participant enrollment and follow-up implementation plans. If any issues are identified during monitoring visits, the lead team will assist the site in resolving them. The lead organization will also conduct audits to assess the overall quality and completeness of the data, review source documents, interview investigators and coordinators, and ensure that clinical centers comply with the study protocol requirements.

Because the intervention is a pharmacist-led care enhancement, it is expected to pose minimal risk, and serious adverse events are unlikely. Nevertheless, potential risks will be reviewed throughout the trial. Study staff will be trained to identify and document any adverse events and to report them in accordance with each hospital's standard procedures and ethical requirements.

Ethical Considerations

This project has received approval from the Clinical Research Ethics Committee at the China-Japan Friendship Hospital, China, on May 20, 2024 (2024-KY-154). Researchers will provide participants or their legal representatives with an easily understandable informed consent form, approved by the ethics committee and allow them sufficient time to consider participation. Participants may not be enrolled in the study until a signed consent form is obtained from the participant or their legal representative. During the study, participants will be provided with all updated versions of the consent form and any relevant written information. The informed consent form will be retained as an essential document of the clinical trial for future reference.

All participant data were de-identified in the EDC system. Specifically, participants were identified using unique numerical codes rather than full personal identifiers. Each participant received compensation of CNY 200 (USD 28.66) for their time and travel expenses associated with study participation. Participants were informed during the informed consent process that they had the right to withdraw from the study at any time, for any reason, and without penalty.

Dissemination Policy

The study findings will be disseminated to participating pharmacists, referring physicians, patients, and the broader medical community through peer-reviewed publications.

Results

Overview

The trial was registered on July 15, 2024 (ChiCTR2400086943). The study was funded in December 2024. Recruitment started on November 15, 2024, and enrollment was completed by December 31, 2025. As of December 31, 2025, we enrolled 454 participants, of whom 16 (3.5%) had completed the 12-month follow-up. The trial is expected to be completed by December 31, 2026, with results expected to be published in 2027 (Figure 1).

Trial Status

The trial enrollment was completed by December 31, 2025, and the trial is now in the follow-up stage.

Discussion

Anticipated Findings

COPD remains a major public health challenge. The GOLD guideline highlights the importance of patient education, with particular emphasis on appropriate inhaler use technique and adherence to regular maintenance therapy [6]. Despite these recommendations, suboptimal disease control often persists due to 3 interrelated issues: medication nonadherence, inappropriate inhaler use technique, and inappropriate device selection [7-11]. These shortcomings not only exacerbate symptoms and hasten disease progression but also diminish HRQoL and increase health care facility use [12-14]. Against this backdrop, this study seeks to determine whether a multifaceted, pharmacist-led intervention, supported by digital tools and based on the IMB model, can mitigate these challenges in routine care. We anticipate that our digital-enabled pharmaceutical care program, which combines structured, IMB-guided pharmaceutical care with digital support via a WeChat mini-program and inhaler-integrated digital sensors, might yield clinically significant improvements in HRQoL.

Comparison With Previous Work

Previous research has shown that pharmacist-led interventions can enhance critical aspects of inhalation therapy, including medication adherence and inhaler use technique [15-20]. IMB-based strategies have further demonstrated improved adherence in other chronic disease contexts by emphasizing clear information, meaningful motivation, and practical behavioral skills [21,22]. In this study, the *information* component involves interactive question-and-answer sessions, visual aids, and online educational resources designed to facilitate accessible and comprehensible knowledge. The *motivation* component enhances patient autonomy through

XSL•FO
RenderX

regular follow-ups, feedback, and reinforcement of self-efficacy. The *behavioral skills* component underscores standardized training, digital sensor–based reminders, and real-time corrective guidance, all aimed at sustaining appropriate medication practices by promoting both correct medication administration and device appropriateness. Each of these 3 IMB dimensions is reinforced by quantifiable evaluation metrics, helping pharmacists deliver more precisely targeted management and education.

Meanwhile, digital tools such as real-time adherence monitoring and telehealth modules hold promise for improving self-management and HRQoL in patients with COPD [23-26]. In this intervention, *inhalable drug digital sensors* provide voice prompts to guide patients in the correct use of inhalers and record real-time inhalation behaviors. Pharmacists can intervene promptly when they identify incorrect use or recommend more suitable inhalation devices. In addition, the study expands beyond traditional face-to-face consultations by offering health knowledge updates, allowing participants to access project-specific information on COPD and medication use at home rather than in person. This approach leverages emerging technologies to establish a new model for pharmacist-led, multifaceted services and lays the groundwork for future digitally integrated pharmaceutical interventions.

Strengths and Limitations

This trial has several anticipated strengths. It uses a multicenter, cluster randomized design across secondary and tertiary hospitals, which enhances external validity in higher-level care settings and reduces contamination between the intervention arm and control arm. The intervention is explicitly theory driven, grounded in the IMB model, and delivered by pharmacists who receive standardized training, which supports both intervention fidelity and future scalability. Furthermore, the study uses a comprehensive set of outcomes, including the SGRQ, TAI-10, EQ-5D-5L, and a pharmacy service survey, to capture not only clinical and behavioral effects but also patient-reported experiences with pharmacist-led care.

However, this study also possesses some limitations. First, our findings mainly reflect patients with COPD attending secondary and tertiary hospitals whose inhaler use technique is suboptimal and who are matched to an appropriate inhaler device, which may limit the generalizability of the results to other care settings. Future investigations should explore primary care settings to validate the intervention's impact across multiple levels of the health care system. Second, we did not measure participants' digital literacy, which could understate differences in engagement or create suboptimal outcomes among individuals who are less comfortable with technology. Finally, because this is a complex pharmacist-led behavioral intervention, neither the content nor the intensity of care can be fully masked to participants. Although the trial was registered as *single blind*, participants in the intervention arm are likely to be aware of receiving additional education and digital tools, and the primary and key secondary outcomes are all patient reported (SGRQ, TAI-10, EQ-5D-5L, and the pharmacy service survey). This may introduce performance and reporting bias. To mitigate this limitation, we used cluster randomization at the clinic level, applied an identical follow-up schedule and basic medication education and inhaler use training in both arms, and blinded the outcome assessors and statisticians.

Conclusions

By facilitating the deep involvement of pharmacists in the medication management process for patients with COPD, the goal of the study is to improve medication adherence and correct inhaler use, thereby enhancing patient outcomes and reducing the disease burden associated with COPD. The study findings will be shared through peer-reviewed publications. If the study results demonstrate a positive impact of this digitally enhanced, multifaceted pharmaceutical intervention on COPD prognosis and disease burden, they will support the widespread adoption of this pharmaceutical management model in more than 1000 CWPC clinics across China, offering a novel approach to medication management for patients with COPD.

Acknowledgments

The funding organizations were not involved in the study design, data collection, data analysis, data interpretation, or manuscript preparation.

Funding

This study was supported by the Clinical and Translational Medical Research Program of the Chinese Academy of Medical Sciences (2024-I2M-C&T-B-092; which also serves as the trial sponsor), the National High-Level Hospital Clinical Research Funding (2023-GSP-GG-1), and the Special Project for the Development of Characteristic Disciplines of Fuwai Hospital, Chinese Academy of Medical Sciences (2022-FWTS08).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the inclusion of personal medication-related information but are available from the corresponding author on reasonable request.

XSL•FO
RenderX

Authors' Contributions

LL and WL designed the research. XM, Xiaocong Li, and Xinyi Li wrote the manuscript. HC, YW, XJ, and YZ provided expertise in pharmaceutical practice. All authors made critical revisions and approved the final manuscript as submitted.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Trial stages, digital inhaler training, and trial registration dataset.
[DOCX File , 208 KB - resprot_v15i1e82806_app1.docx]

Multimedia Appendix 2
Trial registration: dataset.
[DOCX File , 19 KB - resprot_v15i1e82806_app2.docx]

References

1. GBD 2021 Forecasting Collaborators. Burden of disease scenarios for 204 countries and territories, 2022-2050: a forecasting analysis for the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):2204-2256 [FREE Full text] [doi: 10.1016/S0140-6736(24)00685-8] [Medline: 38762325]
2. Wang C, Xu J, Yang L, Xu Y, Zhang X, Bai C, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. Lancet 2018 Apr 28;391(10131):1706-1717. [doi: 10.1016/S0140-6736(18)30841-9] [Medline: 29650248]
3. Chen S, Kuhn M, Prettner K, Yu F, Yang T, Bärnighausen T, et al. The global economic burden of chronic obstructive pulmonary disease for 204 countries and territories in 2020-50: a health-augmented macroeconomic modelling study. Lancet Glob Health 2023 Aug;11(8):e1183-e1193 [FREE Full text] [doi: 10.1016/S2214-109X(23)00217-6] [Medline: 37474226]
4. Hu W, Fang L, Zhang H, Ni R, Pan G. Global disease burden of COPD from 1990 to 2019 and prediction of future disease burden trend in China. Public Health 2022 Jul;208:89-97. [doi: 10.1016/j.puhe.2022.04.015] [Medline: 35728417]
5. Chen D, Long H, Li S, Chen Y. Interpretation of global strategy for the diagnosis, treatment, management and prevention of chronic obstructive pulmonary disease 2024 report. Chin Gen Pract 2024;27(13):1533-1543 [FREE Full text] [doi: 10.12114/j.issn.1007-9572.2023.0867]
6. Venkatesan P. GOLD COPD report: 2024 update. Lancet Respir Med 2024 Jan;12(1):15-16. [doi: 10.1016/S2213-2600(23)00461-7] [Medline: 38061380]
7. Chrystyn H, van der Palen J, Sharma R, Barnes N, Delafont B, Mahajan A, et al. Device errors in asthma and COPD: systematic literature review and meta-analysis. NPJ Prim Care Respir Med 2017 Apr 03;27(1):22 [FREE Full text] [doi: 10.1038/s41533-017-0016-z] [Medline: 28373682]
8. Ding N, Zhang W, Wang Z, Bai C, He Q, Dong Y, et al. Prevalence and associated factors of suboptimal daily peak inspiratory flow and technique misuse of dry powder inhalers in outpatients with stable chronic airway diseases. Int J Chron Obstruct Pulmon Dis 2021;16:1913-1924 [FREE Full text] [doi: 10.2147/COPD.S311178] [Medline: 34188467]
9. Bogart M, Stanford RH, Laliberté F, Germain G, Wu JW, Duh MS. Medication adherence and persistence in chronic obstructive pulmonary disease patients receiving triple therapy in a USA commercially insured population. Int J Chron Obstruct Pulmon Dis 2019;14:343-352 [FREE Full text] [doi: 10.2147/COPD.S184653] [Medline: 30863037]
10. Peng Y, Duan J, Li X, Zeng Y, Zhou Z, Deng M, et al. Adherence to inhaled therapy in patients with COPD associated to pneumoconiosis. Int J Chron Obstruct Pulmon Dis 2021;16:2697-2706 [FREE Full text] [doi: 10.2147/COPD.S327686] [Medline: 34611398]
11. Grandmaison G, Grobéty T, Vaucher J, Hayoz D, Suter P. Prevalence of critical errors and insufficient peak inspiratory flow in patients hospitalized with COPD in a department of general internal medicine: a cross-sectional study. Chronic Obstr Pulm Dis 2024 Jul 25;11(4):406-415 [FREE Full text] [doi: 10.15326/jcopdf.2024.0505] [Medline: 38901836]
12. Molimard M, Raherison C, Lignot S, Balestra A, Lamarque S, Chartier A, et al. Chronic obstructive pulmonary disease exacerbation and inhaler device handling: real-life assessment of 2935 patients. Eur Respir J 2017 Feb;49(2):1601794 [FREE Full text] [doi: 10.1183/13993003.01794-2016] [Medline: 28182569]
13. Turégano-Yedro M, Trillo-Calvo E, Navarro I Ros F, Maya-Viejo JD, González Villaescusa C, Echave Sustaeta JM, et al. Inhaler adherence in COPD: a crucial step towards the correct treatment. Int J Chron Obstruct Pulmon Dis 2023;18:2887-2893 [FREE Full text] [doi: 10.2147/COPD.S431829] [Medline: 38059011]
14. Usmani OS, Lavorini F, Marshall J, Dunlop WC, Heron L, Farrington E, et al. Critical inhaler errors in asthma and COPD: a systematic review of impact on health outcomes. Respir Res 2018 Jan 16;19(1):10 [FREE Full text] [doi: 10.1186/s12931-017-0710-y] [Medline: 29338792]

XSL•FO
RenderX

15. Jia X, Zhou S, Luo D, Zhao X, Zhou Y, Cui YM. Effect of pharmacist-led interventions on medication adherence and inhalation technique in adult patients with asthma or COPD: a systematic review and meta-analysis. J Clin Pharm Ther 2020 Oct 27;45(5):904-917. [doi: 10.1111/jcpt.13126] [Medline: 32107837]
16. Abdulsalim S, Unnikrishnan MK, Manu MK, Alrasheedy AA, Godman B, Morisky DE. Structured pharmacist-led intervention programme to improve medication adherence in COPD patients: a randomized controlled study. Res Social Adm Pharm 2018 Oct;14(10):909-914 [FREE Full text] [doi: 10.1016/j.sapharm.2017.10.008] [Medline: 29104008]
17. Rodrigues AT, Romano S, Romão M, Figueira D, Bulhosa C, Madeira A, et al. Effectiveness of a pharmacist-led intervention on inhalation technique for asthma and COPD patients: the INSPIRA pilot cluster-randomized controlled trial. Respir Med 2021 Aug;185:106507 [FREE Full text] [doi: 10.1016/j.rmed.2021.106507] [Medline: 34166959]
18. Nguyen T, Truong MT, Lam DN, Nguyen HT, Huynh AM, Duong VK, et al. Effectiveness of clinical pharmacist intervention on medication adherence in patients with chronic obstructive pulmonary disease - a randomized controlled study. Patient Educ Couns 2024 Jan;118:108037. [doi: 10.1016/j.pec.2023.108037] [Medline: 37931404]
19. Torres-Robles A, Benrimoj SI, Gastelurrutia MA, Martinez-Martinez F, Peiro T, Perez-Escamilla B, et al. Effectiveness of a medication adherence management intervention in a community pharmacy setting: a cluster randomised controlled trial. BMJ Qual Saf 2022 Feb;31(2):105-115 [FREE Full text] [doi: 10.1136/bmjqs-2020-011671] [Medline: 33782092]
20. Khdour MR, Agus AM, Kidney JC, Smyth BM, McElnay JC, Crealey GE. Cost-utility analysis of a pharmacy-led self-management programme for patients with COPD. Int J Clin Pharm 2011 Aug 4;33(4):665-673. [doi: 10.1007/s11096-011-9524-z] [Medline: 21643784]
21. Chen Q, Wang H, Wang Y, Wang Z, Zhao D, Cai Y. Exploring effects of self-management on glycemic control using a modified information-motivation-behavioral skills model in type 2 diabetes mellitus patients in Shanghai, China: a cross-sectional study. J Diabetes 2018 Sep 23;10(9):734-743. [doi: 10.1111/1753-0407.12655] [Medline: 29457694]
22. To KW, Lee IF, Choi KC, Cheung YT, Yu DS. An information-motivation-behavioural-based model and adherence to inhalation therapy and other health outcomes in patients with chronic obstructive pulmonary disease: a pilot randomized controlled trial. Int J Nurs Pract 2020 Apr;26(2):e12799. [doi: 10.1111/ijn.12799] [Medline: 31903654]
23. Hale EM, Greene G, Mulvey C, Mokoka MC, van Boven JF, Cushen B, et al. Use of digital measurement of medication adherence and lung function to guide the management of uncontrolled asthma (INCA Sun): a multicentre, single-blinded, randomised clinical trial. Lancet Respir Med 2023 Jul;11(7):591-601. [doi: 10.1016/S2213-2600(22)00534-3] [Medline: 36963417]
24. Farmer A, Williams V, Velardo C, Shah SA, Yu LM, Rutter H, et al. Self-management support using a digital health system compared with usual care for chronic obstructive pulmonary disease: randomized controlled trial. J Med Internet Res 2017 May 03;19(5):e144 [FREE Full text] [doi: 10.2196/jmir.7116] [Medline: 28468749]
25. Park SK, Bang CH, Lee SH. Evaluating the effect of a smartphone app-based self-management program for people with COPD: a randomized controlled trial. Appl Nurs Res 2020 Apr;52:151231. [doi: 10.1016/j.apnr.2020.151231] [Medline: 31955942]
26. Wang L, Guo Y, Wang M, Zhao Y. A mobile health application to support self-management in patients with chronic obstructive pulmonary disease: a randomised controlled trial. Clin Rehabil 2021 Jan;35(1):90-101. [doi: 10.1177/0269215520946931] [Medline: 32907384]
27. Xin C, Xia Z, Jiang C, Lin M, Li G. The impact of pharmacist-managed clinic on medication adherence and health-related quality of life in patients with COPD: a randomized controlled study. Patient Prefer Adherence 2016;10:1197-1203 [FREE Full text] [doi: 10.2147/PPA.S110167] [Medline: 27468229]
28. Tommelein E, Mehuys E, Van Hees T, Adriaens E, Van Bortel L, Christiaens T, et al. Effectiveness of pharmaceutical care for patients with chronic obstructive pulmonary disease (PHARMACOP): a randomized controlled trial. Br J Clin Pharmacol 2014 May;77(5):756-766 [FREE Full text] [doi: 10.1111/bcp.12242] [Medline: 24117908]
29. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. Am Rev Respir Dis 1992 Jun;145(6):1321-1327. [doi: 10.1164/ajrccm/145.6.1321] [Medline: 1595997]
30. Plaza V, Fernández-Rodríguez C, Melero C, Cosío BG, Entrenas LM, de Llano LP, et al. Validation of the 'Test of the Adherence to Inhalers' (TAI) for asthma and COPD patients. J Aerosol Med Pulm Drug Deliv 2016 Apr;29(2):142-152 [FREE Full text] [doi: 10.1089/jamp.2015.1212] [Medline: 26230150]
31. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
32. Luo N, Liu G, Li M, Guan H, Jin X, Rand-Hendriksen K. Estimating an EQ-5D-5L value set for China. Value Health 2017 Apr;20(4):662-669 [FREE Full text] [doi: 10.1016/j.jval.2016.11.016] [Medline: 28408009]
33. Ricci de Araújo T, Papathanassoglou E, Gonçalves Menegueti M, Grespan Bonacim CA, Lessa do Valle Dallora ME, de Carvalho Jericó M, et al. Critical care nursing service costs: comparison of the top-down versus bottom-up micro-costing approach in Brazil. J Nurs Manag 2021 Sep 09;29(6):1778-1784. [doi: 10.1111/jonm.13313] [Medline: 33772914]
34. EQ-5D-5L. EuroQol Research Foundation. URL: https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-5l/ [accessed 2025-02-05]

XSL•FO
RenderX

35. Xu L, Chen M, Angell B, Jiang Y, Howard K, Jan S, et al. Establishing cost-effectiveness threshold in China: a community survey of willingness to pay for a healthylife year. BMJ Glob Health 2024 Jan 09;9(1):e013070 [FREE Full text] [doi: 10.1136/bmjgh-2023-013070] [Medline: 38195152]

Abbreviations

COPD: chronic obstructive pulmonary disease
CWPC: cough and wheeze pharmaceutical care
EDC: electronic data capture
GOLD: Global Initiative for Chronic Obstructive Lung Disease
HRQoL: health-related quality of life
IMB: information-motivation-behavioral skills
QALY: quality-adjusted life year
SGRQ: St George's Respiratory Questionnaire
SOP: standardized operating procedure
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TAI-10: Test of Adherence to Inhalers

Edited by J Sarvestan; submitted 22.Aug.2025; peer-reviewed by T Yang, J Edwards; comments to author 08.Dec.2025; revised version received 27.Dec.2025; accepted 27.Dec.2025; published 15.Jan.2026.

Please cite as:
Ma X, Li X, Li X, Chen H, Wang Y, Jia X, Zhao Y, Li W, Liu L
Evaluation of a Tailored Multifaceted Pharmaceutical Care Intervention to Optimize Chronic Obstructive Pulmonary Disease Management: Protocol for a Cluster Randomized Controlled Trial
JMIR Res Protoc 2026;15:e82806
URL: https://www.researchprotocols.org/2026/1/e82806
doi:10.2196/82806
PMID:

©Xuedi Ma, Xiaocong Li, Xinyi Li, Hui Chen, Yang Wang, Xuan Jia, Yanyan Zhao, Wei Li, Lihong Liu. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Impact of the Population Medicine Multimorbidity Intervention in Xishui County (POPMIX) on People at High Risk for Chronic Obstructive Pulmonary Disease Who Experience Mental Health Symptoms: Protocol for the POPMIX-MH Cluster Randomized Controlled Trial

Yuhao Liu[1*], BA; Wenjin Chen[1*], BMed; Shiyu Zhang[2*], BSc; Yingping Wang[3*], BMed; Zhoutao Zheng[1], BMed; Ke Huang[4,5,6], MD; Xingyao Tang[4,5], MMSc; Zhong Cao[7], PhD; Xunliang Tong[8,9], MD; Lei Tang[10], PhD; Jinghan Zhao[7], MA; Liu He[1], BMed; Lirui Jiao[11], MPH; Tianying Zhao[3], BM; Yingchi Luo[3], BMed; Qiande Lai[12], BMed; Xiangqin Lyu[12], BMed; Qiushi Chen[13], PhD; Aditi Bunker[7], PhD; Sebastian Vollmer[14], PhD; Pascal Geldsetzer[15,16], PhD; Dean Jamison[17,18], PhD; Till Bärnighausen[1,7,19], MD, ScD; Ting Yang[4,5,6], MD; Simiao Chen[1,6,7], ScD; Chen Wang[1,4,5,6], MD; POPMIX Group[20]

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[2]School of Health Policy and Management, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[3]Center for Disease Control and Prevention of Xishui County, Zunyi, China

[4]Department of Pulmonary and Critical Care Medicine, China-Japan Friendship Hospital, Beijing, China

[5]National Center for Respiratory Medicine, Beijing, China

[6]State Key Laboratory of Respiratory Health and Multimorbidity, Beijing, China

[7]Heidelberg Institute of Global Health, Faculty of Medicine and University Hospital, Heidelberg University, Heidelberg, Germany

[8]Department of Pulmonary and Critical Care Medicine, Beijing Hospital, Beijing, China

[9]Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

[10]Guizhou Medical University, Guiyang, China

[11]Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[12]Department of Pulmonary and Critical Care Medicine, People's Hospital of Xishui County, Zunyi, China

[13]The Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, PA, United States

[14]Department of Economics and Centre for Modern Indian Studies, University of Göttingen, Göttingen, Germany

[15]Division of Primary Care and Population Health, Department of Medicine, Stanford University, Stanford, CA, United States

[16]Chan Zuckerberg Biohub San Francisco, San Francisco, CA, United States

[17]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

[18]Institute for Global Health Sciences, University of California, San Francisco, San Francisco, CA, United States

[19]Department of Global and Population Health, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, United States

[20]see Acknowledgments

[*]these authors contributed equally

Corresponding Author:
Chen Wang, MD
School of Population Medicine and Public Health
Chinese Academy of Medical Sciences & Peking Union Medical College
Beijige Santiao 31
Dongcheng District
Beijing,
China
Phone: 86 010 65120552
Email: wangchen@pumc.edu.cn

XSL•FO
RenderX

Abstract

Background: Chronic obstructive pulmonary disease (COPD) and mental health conditions represent intersecting public health challenges, especially in resource-limited rural China. Existing care models often neglect the psychosocial needs of populations at high risk for COPD, resulting in limited effectiveness of prevention and management strategies. This study evaluates an integrated intervention designed to improve both mental and physical health outcomes among high-COPD-risk individuals with mental health symptoms, using a population medicine framework.

Objective: This study aims to evaluate the effect of an integrated, population medicine–based multimorbidity intervention package among high-COPD-risk individuals with mental health symptoms in Xishui County, Guizhou Province, China.

Methods: We are conducting a 12-month, 2-arm cluster randomized controlled trial across 26 townships in Xishui County, Guizhou, China. A total of 44,000 residents aged ≥35 years were screened using the Chronic Obstructive Pulmonary Disease Screening Questionnaire, identifying 10,000 individuals at high risk of COPD. Among them, 3807 individuals with Warwick-Edinburgh Mental Well-Being Scale scores below 45 were enrolled as participants. Intervention components include digital cognitive behavioral therapy–based mental health support, community screening, chronic disease management, patient education, digital follow-up, and team-based care. The primary outcomes are depressive symptoms (9-item Patient Health Questionnaire), anxiety symptoms (7-item General Anxiety Disorder), and mental well-being (Warwick-Edinburgh Mental Well-Being Scale). Secondary outcomes are control of chronic diseases, physiological and functional indicators such as lung function, health-related quality of life, mental and behavioral health, health care utilization, knowledge of COPD and asthma, productivity loss, and care cascade indicators for chronic conditions.

Results: Data collection for the POPMIX-MH trial began in June 2024. Baseline, 3-month, and 6-month assessments have been completed, and the 12-month follow-up assessments are planned to be completed in March 2026.

Conclusions: This study is the first to integrate psychological support, chronic disease management, and community-based screening into a single scalable intervention package targeting multimorbidity in China. It tests the feasibility of applying population medicine principles, emphasizing integrated, preventive, and population-level care, within primary care systems in low-resource settings. By targeting both mental and physical health, it redefines chronic care beyond traditional organ-specific approaches.

Trial Registration: ClinicalTrials.gov NCT06458218; https://clinicaltrials.gov/ct2/show/NCT06458218

International Registered Report Identifier (IRRID): DERR1-10.2196/85853

(JMIR Res Protoc 2026;15:e85853) doi:10.2196/85853

KEYWORDS

population medicine; COPD; depression; anxiety; multimorbidity; digital health; China; cluster randomized trial

Introduction

Background

The increasing complexity of chronic disease epidemiology, characterized by the convergence of biomedical, behavioral, psychological, and structural risk factors, has challenged the effectiveness of conventional, disease-centered care models [1,2]. Population medicine has therefore been proposed as an approach that emphasizes preventive, integrated, population-level care to realign health systems with the growing burden of multimorbidity and health inequalities [3]. The Lancet Commission's "50-by-50" goal highlights tobacco-related noncommunicable diseases (NCDs) as a priority area for action and estimates that they account for approximately 1 year of the 3.5-year life expectancy gap between China and the North Atlantic regions [4,5]. These patterns underscore the need to address tobacco-related NCDs at the population level through early identification and comprehensive disease management.

Chronic obstructive pulmonary disease (COPD) is the fourth leading cause of death and disability-adjusted life years globally and is projected to affect nearly 600 million people by 2050, with a disproportionate rise among women and populations in low- and middle-income countries [6,7]. In China, the disease accounted for over 11 million deaths in 2021 and continues to impose a substantial economic burden, with estimated costs exceeding Int $1.3 trillion between 2020 and 2050 [6,8]. Despite its high burden, COPD remains severely underdiagnosed and undertreated, particularly in primary care and low-resource settings [9]. Recent research has revealed substantial gaps across the COPD care cascade in China, particularly in early detection, formal diagnosis, treatment initiation, and symptom control. Data from the "Happy Breathing Programme," a large-scale population-based COPD screening and management initiative in China, reveal major gaps along the care cascade: among 29,201 patients with COPD, 41.0% had previously undergone spirometry testing, 17.6% had been formally diagnosed, and 8.5% were currently receiving treatment. Among those receiving treatment, 4.6% had mild or no exacerbations in the prior year, and 3.9% had experienced no exacerbations [10]. The severity of this care cascade breakdown underscores the urgent need for proactive, community-level interventions aimed at identifying and supporting high-risk individuals before irreversible disease progression occurs.

COPD rarely occurs in isolation but typically forms part of complex multimorbidity profiles that worsen quality of life and clinical outcomes [11]. Among the various comorbid conditions, mental health disorders, particularly depression and anxiety,

XSL•FO
RenderX

are highly prevalent yet often overlooked [11]. Meta-analytic and large cohort evidence suggests that the prevalence of depression in people with COPD is around 25%-30% compared with about 10% in the general adult population [12], corresponding to roughly a 2- to 3-fold higher risk of depression and anxiety than in the general population [13,14]. A national multicenter study found that patients with COPD with comorbid depression had a markedly greater symptom burden and poorer health-related quality of life, yet only 39% were receiving antidepressant treatment and 22% had access to mental health counseling [15].

Current COPD prevention and management frameworks, in both clinical guidelines and public health programs, still focus mainly on biomedical risk factors such as lung function decline and tobacco use, while mental health receives little attention, particularly among high-risk individuals who do not yet meet diagnostic criteria and are therefore rarely screened or systematically assessed for psychological distress [16-20]. Existing evidence on COPD largely comes from diagnosed patients, among whom depression and anxiety are highly prevalent yet frequently underrecognized and undertreated, suggesting that similar or greater unmet mental health needs may exist in undiagnosed high-risk populations who remain outside the formal care system [21,22]. At the same time, most prevention efforts for COPD continue to target physiological indicators, and only a few models address multimorbidity and psychosocial needs in an integrated manner, even though integrated disease management that combines respiratory care, education, and mental health support has been shown to improve self-management, reduce exacerbations, and enhance quality of life [18,19,23-25]. Cognitive behavioral therapy (CBT) is an effective and scalable intervention for anxiety and depression that can be delivered digitally via smartphones, yet such integrated, technology-enabled approaches have rarely been applied to individuals at high risk for COPD; our multicomponent intervention, which incorporates digital CBT, spirometry-based risk stratification, community education, and behavioral self-management support, is designed to address this gap in resource-limited settings [23,25-29].

Objectives

This research forms part of the Population Medicine Multimorbidity Intervention in Xishui County (POPMIX) project, which focuses on creating and assessing evidence-informed, multicomponent approaches for managing chronic conditions and their multimorbidity in resource-limited settings. This trial targets individuals at high risk of COPD who also exhibit mental health symptoms. We will examine whether a population-based, integrated intervention for multiple long-term conditions affects the primary outcomes—symptoms of depression, symptoms of anxiety, and overall mental well-being. Additionally, we will evaluate the intervention's impact on several secondary outcomes, including the number of controlled multimorbid chronic conditions; forced expiratory volume in 1 second (FEV_1); quality-of-life measures; patient knowledge of COPD and asthma; indicators of COPD symptoms; levels of smoking dependence; care cascade indicators for COPD, asthma, hypertension, and type 2 diabetes; patterns of health care services utilization; indicators of health risk behaviors; and productivity loss.

Methods

Trial Design

This investigation employed a parallel, 2-arm, stratified cluster randomized controlled trial (cRCT) design, implemented in Xishui, Guizhou, China (Figure 1). The study protocol was developed in strict accordance with the guidance of the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2025 Statement [30]. Compliance with SPIRIT 2025 is detailed in Multimedia Appendix 1, which includes the completed checklist for this protocol.

XSL•FO
RenderX

Figure 1. Trial flowchart. COPD: chronic obstructive pulmonary disease; COPD-SQ: COPD Screening Questionnaire. See Multimedia Appendix 4 for a higher resolution version of this figure.

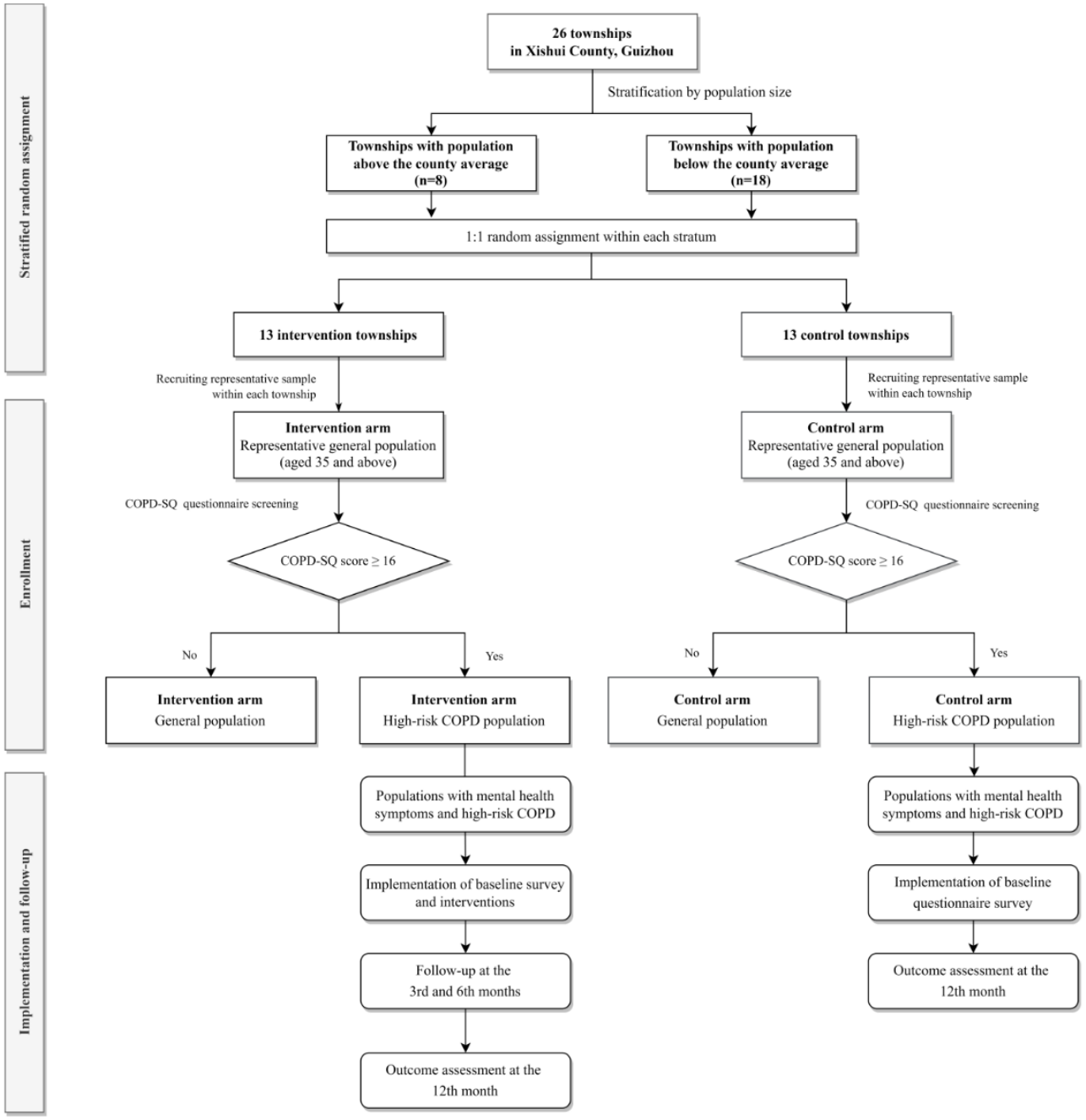

Townships were stratified according to whether they were above or below the county average population size. Within each stratum, townships were randomly assigned (1:1) to the intervention or control arm using computer-generated randomization. Participants were recruited from a government-provided roster of permanent residents aged 35 years or older as of May 10, 2024, with individuals enrolled from each township.

The COPD Screening Questionnaire (COPD-SQ) was used to identify individuals at high risk of COPD. Developed by Zhou et al [31], this instrument comprises 7 items: age, cumulative smoking history, BMI, cough, breathlessness, family history of respiratory diseases, and exposure to cooking-related smoke. It generates a total score ranging from 0 to 38, with higher scores indicating greater COPD risk. Validated through multiple epidemiological surveys and community health screenings, the COPD-SQ is recommended for primary care–based COPD screening in China. For this study, participants scoring 16 or higher on the COPD-SQ were considered eligible for inclusion, thereby ensuring that the cohort comprised individuals identified as high risk for COPD.

Setting

This study was implemented across 26 township-level clusters in Xishui County, a mountainous area in northern Guizhou Province, China. A rural county designated as China's National Comprehensive Primary Health Experimental Area, Xishui presents a unique juxtaposition of policy-driven innovation and socioeconomic challenges. With a gross domestic product per capita (48%) below the national average in 2024 and limited

XSL•FO
RenderX

human capital development, the county exemplifies a resource-constrained rural setting. Notably, Guizhou Province also has one of the highest prevalences of smoking in China (at 37.9%), driving a high regional burden of tobacco-related NCDs such as COPD [32]. These conditions, combined with the county's experimental role in health system transformation, have established Xishui as an ideal setting in which to evaluate scalable multimorbidity interventions tailored to underserved populations.

Trial Participants (Inclusion and Exclusion Criteria)

Participants were eligible for inclusion if they were aged 35 years or older and met the study definition of a high-COPD-risk individual with mental health symptoms. Such individuals were defined as having a COPD-SQ score of 16 or higher, indicating elevated risk for COPD, and a Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) score of 45 or lower, reflecting reduced mental well-being. To ensure stability and adherence to the intervention, only participants who had lived in the same township for at least 3 months before enrollment and who planned to remain in the township for the subsequent 12 months were included. This residency requirement helped limit population mobility and reduce potential dilution or contamination of intervention effects.

Individuals exhibiting severe cognitive disorders that substantially compromised comprehension, decision-making capacity, or intervention adherence were excluded. Likewise, participants lacking full independence in performing activities of daily living were ineligible. These criteria ensured that enrolled individuals could actively participate in the intervention and follow-up procedures while reducing potential confounding factors.

Intervention

Participants in the intervention group were provided with access to a multicomponent intervention package. This package integrates key strategies, including community-based risk screening for COPD and mental health problems, digital CBT delivered via the EmoEase program, risk-stratified management of chronic respiratory and cardiometabolic conditions, smoking cessation support and health education, and a pay-for-population incentive mechanism for primary care providers within routine primary care (Multimedia Appendix 2). The intervention package was designed in collaboration with policy makers, aligning with the specific needs and challenges of Xishui County. Policy makers primarily focused on the overall effectiveness of the intervention package rather than on individual components. Accordingly, the integrated design of the package is highly relevant to the local context, aiming to address broad community health goals rather than isolating the effects of each subintervention.

Developed using iterative prototyping and stakeholder input, the strategies within this package are all established as cost-effective and are specifically promoted by the Lancet Commission on Investing in Health [4]. Table 1 summarizes the distinct eligibility criteria applicable to each intervention. The control group receives usual care only. In Xishui, usual care largely follows a traditional, reactive model, in which residents seek care on their own initiative when they experience symptoms rather than through proactive population management. As a national pilot county for comprehensive primary health care reform, Xishui has relatively favorable reimbursement policies for chronic diseases. A visualized schematic flow of the integrated intervention logic is presented in Figure 2. This figure maps the progression from initial risk screening through intervention stratification and follow-up for multiple chronic conditions. The illustration also demonstrates how the EmoEase program and chronic disease pathways are embedded within a broader population medicine framework.

Figure 2. Integrated pathway of the multicomponent intervention package. CBT: cognitive behavioral therapy; COPD: chronic obstructive pulmonary disease; COPD-SQ: COPD Screening Questionnaire; CT: computed tomography; WEMWBS: Warwick-Edinburgh Mental Well-Being Scale. See Multimedia Appendix 4 for a higher resolution version of this figure.

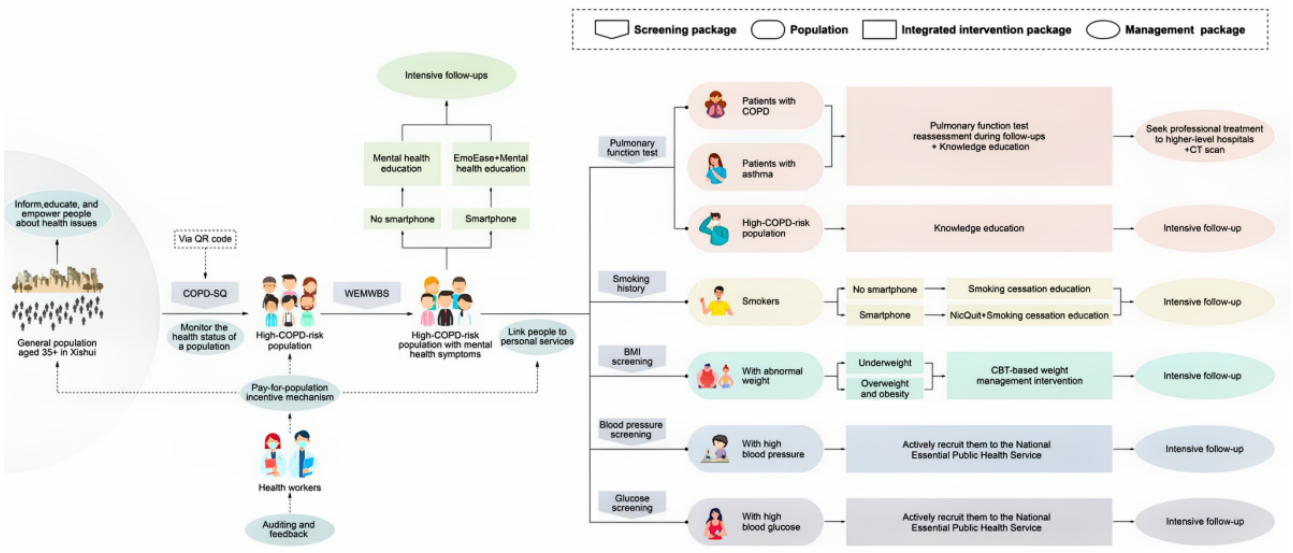

XSL•FO
RenderX

Table 1. Eligibility criteria for interventions.

Population level and target population	Intervention	Eligibility criteria
General population		
General population	Health education	N/A[a]
General population	Online screening with the COPD[b] Screening Questionnaire, followed by WEMWBS screening for identified high-COPD-risk individuals	Permanent residents aged 35 years and above
Patients with COPD or high-COPD-risk population		
High-COPD-risk population	Community-based spirometry pulmonary function tests, interpretation of results, and health education	Individuals with COPD-SQ[c] score ≥16
High-COPD-risk population with mental health symptoms	Digital mental health interventions and health education	Individuals with WEMWBS[d] score <45 who own a smartphone
High-COPD-risk population with mental health symptoms	Health education for mental health symptoms	Individuals with WEMWBS score <45
Smokers within the high-COPD-risk population with mental health symptoms	Digital interventions for smoking cessation and mental health, along with health education	Currently smoking or have quit within the last 6 months and own a smartphone
Patients with COPD or high-COPD-risk population with mental health symptoms who fail to conduct pulmonary function tests	Encouragement to conduct a computed tomography scan and seek professional medical treatment	Individuals with FEV_1[e]/FVC[f] <0.7 after bronchodilation, or high-COPD-risk population who fail to conduct a pulmonary function test for any reason
High-COPD-risk population with mental health symptoms, and hypertension	Hypertension management and education	Systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg
High-COPD-risk population with mental health symptoms, and type 2 diabetes mellitus	Diabetes management and education	Fasting blood glucose ≥7.0 mmol/L or random blood glucose ≥11.1 mmol/L
High-COPD-risk population with mental health symptoms, and abnormal weight	Weight abnormality interventions	Individuals with a BMI < 18.5 kg/m^2or a BMI ≥ 24.0 kg/m^2
Smokers within the high-COPD-risk population with mental health symptoms	Health education for smokers to support smoking cessation	Currently smoking or have quit within the last 6 months
Health providers		
Health providers	Intrinsic incentive mechanism	Primary care providers who engage with the investigation and intervention
Health providers	Extrinsic incentive mechanism	Primary care providers who engage with the investigation and intervention

[a]N/A: not applicable.

[b]COPD: chronic obstructive pulmonary disease.

[c]COPD-SQ: COPD Screening Questionnaire.

[d]WEMWBS: Warwick-Edinburgh Mental Well-Being Scale.

[e]FEV_1: forced expiratory volume in 1 second.

[f]FVC: forced vital capacity.

In addition to the intervention components and care pathways described in Figure 2, implementation of the multicomponent intervention package was supported by a multilevel delivery structure and a performance-linked incentive mechanism. Figure 3 illustrates the organizational structure, population stratification, incentive design, and distribution of responsibilities across administrative levels (county, township, village, and household) for delivering the intervention in Xishui. This figure also specifies the eligibility criteria for each target subpopulation, clarifying how risk stratification and service delivery were operationalized in the field.

XSL•FO
RenderX

Figure 3. Implementation structure and execution mechanism of the multicomponent intervention package. CDC: Centers for Disease Control and Prevention; COPD: chronic obstructive pulmonary disease; COPD-SQ: COPD Screening Questionnaire; ECRHS: European Community Respiratory Health Survey; FEV1: forced expiratory volume in 1 second; FVC: forced vital capacity; PCCM: pulmonary and critical care medicine; PFT: pulmonary function test; WEMWBS: Warwick-Edinburgh Mental Well-Being Scale. See Multimedia Appendix 4 for a higher resolution version of this figure.

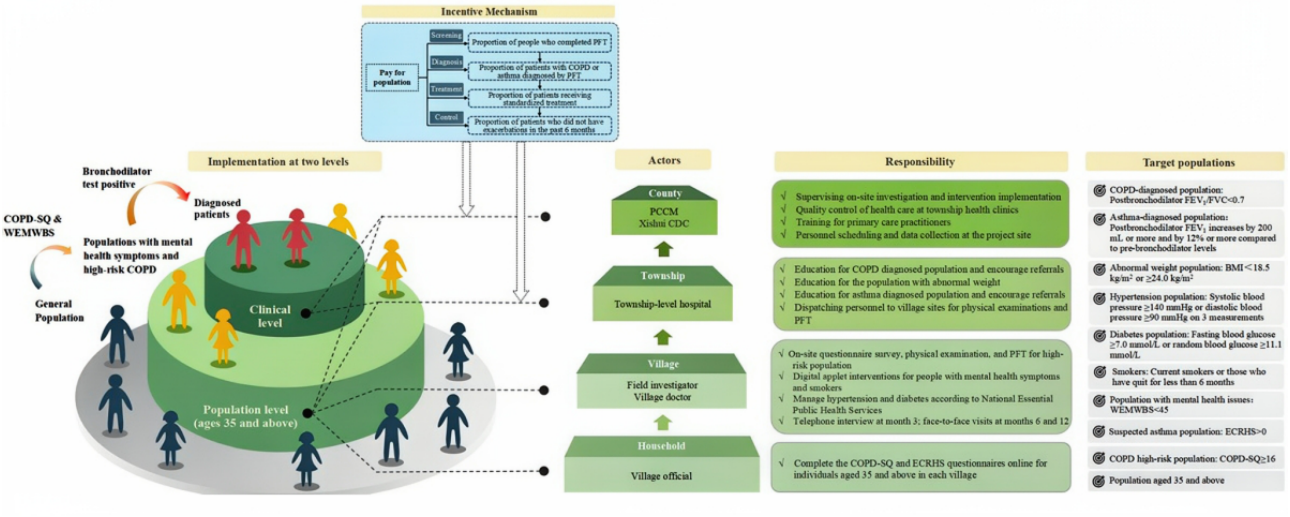

Specific Interventions

Overview

Specific interventions as part of the multicomponent intervention package are described in the following sections.

Health Education

Health education is provided universally to all permanent residents living in townships randomized to the intervention arm. The content includes basic information about COPD; mental health; and common risk factors such as tobacco exposure, poor diet, and lack of physical activity. Educational materials are distributed in both printed and digital formats, such as videos/messages distributed via WeChat (Tencent Holdings Limited). Health workers conduct door-to-door education sessions and use public broadcasts and posters in villages to raise awareness and promote behavior change. These activities aim to improve population-level health literacy and early symptom recognition.

Psychological Health Questionnaire Screening

To facilitate the early detection of psychological distress, we implement mental health questionnaire screening at baseline. All high-COPD-risk populations aged 35 years and above are invited to complete a brief mental health assessment in the baseline questionnaire, which consists of the WEMWBS to measure overall mental well-being, the 9-item Patient Health Questionnaire (PHQ-9) to assess depressive symptoms, and the 7-item Generalized Anxiety Disorder (GAD-7) scale to assess anxiety symptoms. Screening results are used to identify individuals who may benefit from subsequent digital mental health interventions and tailored follow-up support.

Online Screening for COPD

To facilitate the early identification of the high-COPD-risk individuals, the study implements an online screening mechanism using digital tools. A QR code linking to the COPD-SQ is distributed to all eligible representative permanent residents aged 35 years and above. Community health workers assist participants in completing the online forms when needed. This approach enables broad population coverage, immediate risk assessment, and real-time referral for spirometry testing and education. Screening results are stored in the digital health system to support follow-up planning.

Community-Based Spirometry Pulmonary Function Testing and Health Education

High-COPD-risk individuals in the intervention group receive real-time pop-up alerts directing them to a designated community gathering place for spirometry testing, which is conducted using the BH-AX-MAPG spirometry equipment (BreathHome). Prebronchodilator spirometry is conducted at baseline and 12 months for the high-COPD-risk population; postbronchodilator spirometry is conducted at baseline and 12 months for patients with COPD; and prebronchodilator spirometry is conducted at 6 months for high-COPD-risk participants with prebronchodilator FEV/FVC<70% at baseline. Individuals who screen positive for COPD at baseline are referred to the county hospital for computed tomography and a definitive diagnosis. In addition, they receive health education on the risks of COPD and strategies for its prevention and management. This education is delivered verbally by primary health care providers and supplemented with printed materials distributed by trained community health workers or general practitioners.

Mental Health Digital Health Interventions

A CBT-based digital mental health intervention, EmoEase, a WeChat (Tencent Holdings Limited)-based application, is offered to individuals experiencing mental health symptoms (WEMWBS score <45) who also have a smartphone [33]. The intervention includes structured psychoeducational content, mood tracking, and guided CBT exercises designed to enhance emotional regulation and reduce symptoms of anxiety and depression [33]. The content is tailored to highlight the interplay between mental health and respiratory symptoms, aiming to

XSL•FO
RenderX

improve coping strategies, promote treatment adherence, and support sustainable behavior change. The program is self-paced and optimized for accessibility in low-resource settings, with additional support provided by primary health care providers to encourage ongoing engagement.

Health Education for High-COPD-Risk Individuals With Mental Health Symptoms

Health education is offered to high-COPD-risk individuals with co-occurring mental health symptoms, defined by a WEMWBS score below 45. This health education includes guidance on managing mental health symptoms and is delivered verbally by trained community health workers or general practitioners, supplemented with printed materials for distribution.

Smoking Cessation Digital Health Interventions

NicQuit is a WeChat-based digital smoking cessation intervention that includes CBT modules focused on smoking cessation strategies, methods for coping with triggers, and reinforcement techniques to maintain abstinence. It is designed for smokers who are currently smoking or who have quit within the last 6 months. The intervention targets individuals familiar with smartphone technology, ensuring accessibility and usability. Personalized notifications and reminders are delivered through the WeChat platform to encourage regular engagement with the cessation plan and maintain adherence.

Encouragement to Conduct Computed Tomography Scan and Seek Professional Medical Treatment

Participants diagnosed with COPD or asthma through spirometry—defined as those with a postbronchodilator FEV_1-to-forced vital capacity (FVC) ratio of <0.7 for COPD and those with a ≥200 mL and ≥12% improvement in FEV_1 postbronchodilation for asthma—are encouraged to seek professional medical treatment at higher-level hospitals for further diagnosis and management. High-risk participants who do not undergo pulmonary function testing for any reason are encouraged to receive a computed tomography scan and seek formal diagnosis and medical treatment at a higher-level hospital.

Hypertension and Diabetes Management

The goal of this intervention is to actively enroll high-COPD-risk individuals with mental health symptoms and blood pressure higher than 140/90 mmHg or random blood glucose higher than 11.1 mmol/L (or fasting blood glucose ≥7.0 mmol/L) into the National Essential Public Health Service in China. These participants are also provided with health education on abnormal blood pressure or blood glucose [34,35].

Weight Abnormality Interventions

Individuals with a BMI <18.5 kg/m^2 (underweight) or a BMI ≥24.0 kg/m^2 (overweight and obesity) are considered to have weight abnormalities. CBT-based motivational interviewing is used to guide participants in self-identifying weight-related barriers. During the intervention, participants are asked CBT-informed questions designed to encourage active reflection on the inconveniences and conveniences of being underweight or overweight.

Health Education for Smokers to Support Smoking Cessation

Smokers in the intervention group receive targeted health education to reinforce the importance of smoking cessation. This education focuses on the health risks associated with smoking and the benefits of quitting, providing evidence-based information and practical advice. The health education is delivered verbally by primary health care providers and supplemented with printed posters, online videos/messages for distribution.

Pay-for-Population Mechanism

A novel pay-for-population mechanism has been introduced to actively engage primary care providers in the intervention by aligning their compensation with population health goals. This approach ties provider payment to 4 key stages of care: screening, diagnosis, treatment, and control. Health providers in township hospitals are rewarded based on the proportion of residents aged 35 years and above who complete pulmonary function testing; the proportion of individuals diagnosed with COPD among those identified as high risk in the initial screening (defined as COPD-SQ score >16); the proportion of patients with confirmed COPD receiving standardized inhaled treatment relative to all confirmed cases; and the proportion of patients with confirmed COPD who have not experienced acute exacerbations in the past 6 months. The county hospital's respiratory department and the county Center for Disease Control and Prevention are assessed using the same 4 indicators, with data aggregated across all 13 intervention townships. In addition to offering providers extrinsic, results-based financial incentives, the program provides specialized training and capacity-building opportunities to appeal to intrinsic motivation, support proactive service delivery, and foster a strong sense of responsibility for population health (see Multimedia Appendix 2 for details).

Importantly, the POPMIX-MH intervention does not increase the number of primary care providers or extend their contracted working hours. All activities are delivered by the existing workforce within routine service schedules, with capacity enhanced through focused training on COPD risk assessment, mental health screening, and population-based follow-up. These activities have not reduced access to routine care for nonparticipants; on the contrary, opportunistic screening has enabled providers to identify additional residents with previously unmet health needs and to offer appropriate counseling or referral when indicated.

To enhance clarity, Table 2 summarizes all intervention and control activities at baseline, during the 3- and 6-month follow-ups, and at the 12-month final assessment.

XSL•FO
RenderX

Table 2. Components and timeline of the intervention and control conditions.

Time point	Intervention group	Control group
Baseline	• Online spirometry questionnaire screening • Psychological health questionnaire screening • Collection of demographic and health information • Distribution of the mental health intervention and other chronic obstructive pulmonary disease–related interventions (see Figure 1, Table 1 for details)	• Online spirometry questionnaire screening • Psychological health questionnaire screening • Collection of demographic and health information
Month 3	• Follow-up assessment of intervention adherence and psychological health status; delivery of booster sessions	• N/A[a]
Month 6	• Follow-up assessment of intervention adherence and psychological health status; delivery of booster sessions	• N/A
Month 12	• Final outcome assessment	• Final outcome assessment

[a]N/A: not applicable.

Outcomes

We appraise primary and secondary outcomes among high-COPD-risk individuals with concurrent psychological distress symptoms. Following intention-to-treat allocation to the intervention or control arm, participants are tracked longitudinally for 12 months. Outcome and covariate data are collected at 3 months (via telephone for the intervention group only), at 6 and 12 months (via in-person visits) after enrollment. Specific secondary end points (chronic disease knowledge, health service use, and count of controlled chronic conditions) are measured solely at 12 months. Table 3 presents the definitions and measurement procedures for the primary outcomes; secondary outcomes are listed in Table S1 in Multimedia Appendix 3.

Table 3. Primary outcomes of the cluster randomized controlled trial.

Primary outcomes	Description
Depression symptoms	• Definition: Emotional disorders, including sadness, loss, and anger • Functional form: Continuous • Measurement: Score on the 9-item Patient Health Questionnaire items [36], which ranges from 0 to 27, with higher scores representing more severe depression symptoms
Anxiety symptoms	• Definition: Unpleasant state of inner turmoil • Functional form: Continuous • Measurement: Score on the 7-item General Anxiety Disorder scale [37], which ranges from 0 to 21, with higher scores representing more severe anxiety symptoms
Warwick-Edinburgh Mental Well-Being Scale	• Definition: A score reflecting the overall mental health state • Functional form: Continuous • Measurement: Score on the Warwick-Edinburgh Mental Well-Being Scale [38], which ranges from 14 to 70, with lower scores representing worse general mental health

Timeline

Trial initiation occurred on June 17, 2024, with recruitment spanning townships in Xishui County. Participants in the intervention group will complete structured assessments at 4 time points: baseline (in person), 3-month follow-up (telephone), 6-month follow-up (in person), and 12-month follow-up (in person). Postconsent baseline assessments include comprehensive physical examinations and pulmonary function tests. Pulmonary function test protocols are stratified: prebronchodilator spirometry is conducted at baseline and 12 months for the high-COPD-risk population; postbronchodilator spirometry is conducted at baseline and 12 months for patients with COPD; and prebronchodilator spirometry is conducted at 6 months for high-COPD-risk participants with prebronchodilator FEV/FVC<70% at baseline. All baseline procedures are supervised by trained site personnel to ensure protocol adherence. Subsequent follow-ups include a telephone-based 3-month assessment (tracking intervention compliance and health status) and in-person 6- and 12-month visits replicating baseline protocols. Control group activities include baseline questionnaires, and 12-month on-site surveys. The prospective cohort timeline and data collection schedule are detailed in Table S2 in Multimedia Appendix 3.

Sample Size

The study encompassed all 26 townships in Xishui County, which were allocated 1:1 to the intervention or control group through stratified randomization. We enrolled a population-representative sample of 44,000 individuals aged 35 years or older. COPD risk screening was conducted universally using the COPD-SQ, administered digitally via QR

XSL•FO
RenderX

code. Based on an anticipated 10% attrition rate and an estimated high-COPD-risk prevalence of 25% in this age group, we projected 10,000 high-risk participants.

Three primary outcomes were prespecified: overall mental well-being measured by the WEMWBS, depressive symptoms measured by the PHQ-9, and anxiety symptoms measured by the GAD-7. The sample size calculation was based on the WEMWBS, which served as the primary well-being end point. To control for multiple comparisons across the 3 primary outcomes, a Bonferroni correction was applied by dividing the significance threshold by 3. The parameters used for the WEMWBS-based sample size calculation, including an SD of 10.59 and an anticipated mean difference of 5.95, were conservatively derived from a prior longitudinal observational study of online psychiatric assessments [39]. Accounting for a 2-sided α of 5%, 80% power, an intraclass correlation coefficient (ICC) of 0.05 to adjust for clustering, an anticipated 10% nonresponse rate, and a 10% loss to follow-up, the calculation yielded a minimum required sample of 2610 high-COPD-risk participants with multimorbid mental health symptoms. To date, 3807 participants have been enrolled, exceeding the initial target and allowing for more robust statistical analyses of the outcomes.

Power Calculations

In our power calculation, we calculated the minimum detectable differences for the primary outcomes (WEMWBS, PHQ-9, and GAD-7). The ICC estimate was informed by previous literature, which suggests a likely ICC range of 0.03-0.05 for these outcomes [40,41]; we therefore selected 0.05 as a reasonable and conservative estimate. Using this ICC value, together with a 2-sided α of 0.05 and 80% power, the study is able to detect a standardized effect size of Cohen *d*=0.425. This corresponds to a small effect size and to the following raw score differences (Table 4): 4.497 for WEMWBS, 1.606 for PHQ-9, and 2.293 for GAD-7. These calculations indicate that the study is adequately powered to detect small but meaningful differences in these outcomes. Moreover, based on previous research reporting an effect size of 5.95 on the WEMWBS, we consider the study to be well powered to detect clinically relevant changes in these measures [39].

Table 4. Minimum detectable differences in WEMWBS[a], PHQ-9[b], and GAD-7[c].

N_actual	Intraclass correlation coefficient	Cohen *d*	WEMWBS	PHQ-9	GAD-7
3807	0.050	0.425	4.497	1.606	2.293

[a]WEMWBS: Warwick-Edinburgh Mental Well-Being Scale.

[b]PHQ-9: 9-item Patient Health Questionnaire.

[c]GAD-7: 7-item Generalized Anxiety Disorder.

Blindness

In this study, we employed an open-label design, in which both participants and health care providers were aware of group assignments. However, to reduce potential bias, several measures were implemented. Although participants and health care providers knew the intervention group allocation, the principal investigator remained blinded to group assignments. Given the trial design, principal investigator blinding represents the best possible approach and helps ensure impartiality in monitoring and data analysis.

Despite these efforts, complete blinding of both physicians and participants was not feasible due to practical constraints. Physicians, as key members responsible for administering the intervention, must be aware of group assignments to provide appropriate care. Additionally, participants are likely to be aware of their group assignment, particularly in psychological or behavioral interventions. Traditional clinical trials often highlight the risk of bias associated with a lack of blinding, especially for self-reported outcomes. However, open-label trials are frequently used in population medicine and real-world research settings, where blinding participants and health care providers is often impractical.

Given that our study includes 3 primary outcomes and 34 secondary outcomes, of which 8 are based on validated self-reported scales and 8 are based on objective assessments, we made efforts to include as many objective measures as possible to reduce bias. This approach enhances the objectivity of the data and minimizes reliance on self-reported outcomes.

Additionally, recognizing that social desirability bias may affect self-reported outcomes, we plan to conduct qualitative interviews as part of the follow-up assessments. These interviews will provide insights into the potential impact of social desirability bias on participants' responses and help clarify how such biases may have influenced the results. By doing so, we aim to ensure a more comprehensive and balanced interpretation of the findings, while acknowledging the potential influence of bias on self-reported data.

Assignment of the Intervention

This stratified cRCT assigned participants at the township level to either a multicomponent intervention or a control arm, with randomization stratified by geographic region to balance community characteristics. Using a computer-generated algorithm developed by an independent statistician not involved in recruitment or intervention delivery, 13 townships were allocated to each study arm. Given the open-label design, blinding was not feasible, and both participants and providers were aware of group assignments. Control group participants received no additional interventions beyond ongoing usual care throughout the trial period.

Adherence reinforcement strategies are embedded across all follow-up phases. A structured 3-month telephone consultation reinforces participation. In-person 6- and 12-month visits provide individualized feedback on key health indicators (eg,

XSL•FO
RenderX

spirometry results, mental health scores), benchmarking progress against baseline values. Adherence is quantitatively tracked using validated questionnaires at each interval to assess protocol compliance (eg, app engagement frequency). Automated reminders and tailored counseling iteratively address adherence gaps. Each follow-up thus serves as an opportunity to enhance participant motivation, resolve barriers, and refine engagement strategies.

Data Collection Plan

Data are collected across multiple domains to comprehensively assess the effects of the multicomponent intervention (see Table 1). At baseline, all participants completed a structured questionnaire that includes measures of mental health status (PHQ-9 for depression, GAD-7 for anxiety, and the Warwick-Edinburgh Mental Well-Being Scale for overall mental health), sociodemographic characteristics, occupational and environmental exposures, health-related behaviors, chronic disease history, and respiratory symptoms. Physiological and biochemical indicators include measurements of height, weight, BMI, waist circumference, blood pressure, heart rate, capillary blood glucose, and oxygen saturation.

Lung function is assessed using portable spirometry (FEV_1, FVC, and FEV_1/FVC) at multiple time points. In the high-COPD-risk population, prebronchodilator spirometry was performed at baseline and is repeated at months 6 (patients with COPD) and 12 to monitor changes in lung function. Among participants with diagnosed COPD with mental health symptoms, postbronchodilator spirometry was conducted at baseline and is repeated at month 12 to assess reversibility and disease progression.

Subsequent follow-up collections are scheduled at 3 months (via telephone), 6 months (in-person), and 12 months (in-person) for the intervention group and only at 12 months for the control group (in-person). These include assessments of mental health (PHQ-9, GAD-7, and WEMWBS), health-related quality of life (EQ-5D-5L), and Work Productivity and Activity Impairment-General Health. Complementary assessments encompass treatment adherence, control of COPD symptoms, frequency of exacerbations, physical and biochemical parameters (BMI, blood pressure, and blood glucose), and lung function, as previously specified. Lifestyle practices, including physical activity levels, alcohol intake, and dietary patterns (sugar consumption, salted vegetable intake, and fresh vegetable consumption), are additionally documented using participant self-reported questionnaires. Data on health care resource utilization, such as outpatient visits, hospital admissions, and medication use, are also systematically recorded.

All data collection is conducted by trained research personnel in accordance with standardized operational protocols. Collected data are entered into a secure cloud-based electronic data capture (EDC) platform equipped with automated validation checks and quality assurance procedures. These safeguards are designed to ensure data accuracy and minimize transcription errors.

Data Management

This research employs a cloud-based EDC platform to facilitate immediate data entry, automated validation checks, and secure data storage. Trained field researchers collect information using tablet-based digital questionnaires, with responses synchronized to a central repository in real time. The EDC system incorporates built-in safeguards, including range validations, logical consistency checks, and alerts for incomplete entries, to reduce data entry errors. Any anomalies identified by these mechanisms are promptly reviewed by the study's data management team.

Data collection instruments follow standardized, precoded formats to streamline demographic, clinical, and behavioral assessments. Biometric measurements (eg, lung function, blood pressure, BMI, and blood glucose) are entered directly into the system. Pulmonary function data from portable spirometry devices are transmitted to the EDC platform, where they undergo automated quality control procedures to verify data validity.

To protect participant confidentiality, all data are deidentified using numeric coding systems and stored in a password-protected database reserved exclusively for analytical purposes by the research team. Access to this database is strictly limited to a small group of authorized personnel, including the principal investigator, co-investigators, biostatisticians, and data analysts.

Statistical Analysis

Primary analyses will estimate the effect of the intervention on the 3 primary mental health outcomes (WEMWBS, PHQ-9, and GAD-7 scores at 12 months) using generalized linear mixed models under the intention-to-treat principle. Before model fitting, baseline characteristics of participants in the intervention and control arms, including WEMWBS, PHQ-9, and GAD-7 scores, will be summarized to assess postrandomization balance and identify any substantial differences that may warrant additional covariate adjustment. Models will include a fixed effect for trial arm, visit, and trial arm-by-visit interaction; random intercepts for township (cluster) and participant to account for clustering and repeated measurements; and adjustment for the baseline value of the corresponding outcome and a set of prespecified baseline covariates. These covariates will include age, sex, smoking status, and other relevant demographic and health variables that may influence the outcomes. Continuous secondary outcomes will be analyzed using analogous linear mixed models, and binary outcomes using logistic mixed models with the same random-effects structure and covariate set.

To control for multiple comparisons across the 3 primary outcomes, we will apply the Bonferroni correction, dividing the significance threshold by 3 to reduce the family-wise error rate. For secondary outcomes, we will apply the Benjamini-Hochberg procedure at a 5% false discovery rate. False discovery rate control provides a principled and widely accepted way to limit the overall proportion of false positives, while preserving greater statistical power compared with more stringent approaches such as the Bonferroni correction.

Although the primary analysis focuses on the overall intervention package, we will also conduct additional secondary analyses using the Regression Discontinuity Design. This approach was used to explore the distributional effects of the intervention, with a specific focus on the WEMWBS outcome.

XSL•FO
RenderX

We applied a threshold of 45 for WEMWBS scores, dividing participants into those just above and just below this cut-off. The Regression Discontinuity Design methodology allows us to estimate the local average treatment effect at this threshold, helping us assess the effects of the intervention on individuals near the cut-off and providing a more precise estimate of the intervention's impact. By analyzing participants around this threshold, we evaluate the intervention's effect in a more localized and targeted manner.

Monitoring

An independent Data and Safety Monitoring Board (DSMB) was appointed to ensure scientific rigor and impartial oversight of the trial's safety and progress. The DSMB convened 3 meetings in which baseline and follow-up descriptive analyses for the intervention arm were evaluated for process monitoring, as well as a formal interim analysis in February 2026. At this meeting, the research team reviewed the conduct of the trial, including adherence to intervention components, logistical issues, and the completeness and quality of the collected data. Our biostatisticians also presented emerging effect estimates for the intervention and control arms to the steering committee. Stopping rules were established regarding safety and study completion:

- If deaths or severe adverse effects were reported, these would be immediately conveyed to the DSMB. The DSMB would determine whether the deaths or adverse effects were related to the intervention and advise whether the study should be stopped immediately.
- In February 2026, during the planned interim analysis, the trial was approaching completion, as field workers confirmed that the majority of eligible participants had been contacted. At this stage, the results were assessed to determine the study's conclusion based on the following:
    - Null findings, suggesting that the intervention had no detectable effect and that the study should be stopped.
    - Statistically significant findings, indicating that the study had achieved its primary objectives and should be stopped.
    - Emerging but statistically insignificant effects, suggesting that the study might require further observation, if applicable. However, the study would be terminated if field workers confirmed that all eligible participants had been contacted and no additional individuals remained for follow-up assessments.

Ethics and Dissemination

Research Ethics Approval

This study is registered on ClinicalTrials.gov under the identifier NCT06458218, with registration completed on June 9, 2024. It was initially approved by the Ethics Committee of Peking Union Medical College (approval number CAMS&PUMC-IEC-2024-043). A continuing ethics review was completed in June 2025, and updated approval was granted under the number CAMS&PUMC-IEC-2025-064. The study is conducted in accordance with the Declaration of Helsinki and good clinical practice. Participation in this study is entirely voluntary. The study was fully explained to participants, who were required to provide written informed consent before enrollment. Participants were required to declare that they freely and voluntarily agreed to participate in the study and to complete all required questionnaires. Participants were informed that they may withdraw from the study at any time without any repercussions. They may withdraw without providing a reason, and their data will be excluded upon request.

Plans for Communicating Important Protocol Amendments to Relevant Parties

The Ethics Committee of Peking Union Medical College was approached and informed of all protocol revisions requiring its review and approval. After securing Ethics Committee clearance but before implementation, all study personnel received detailed briefings on the amended protocol.

In this trial, an unavoidable amendment was made to the secondary outcome measures. Although outcome modifications are generally discouraged, the change was necessitated by a nonnegotiable regulatory restriction. Specifically, we initially planned to measure health care utilization (ie, outpatient visits, inpatient visits, and medical expenditure) via electronic health records from medical institutions in Xishui County. However, due to unexpected government regulations, we were unable to reach an agreement to obtain these data. This amendment was approved by the Ethics Committee, and all research personnel were notified before implementation to ensure consistency in data collection and analysis.

Consent and Withdrawal

Our study adheres to the ethical principles outlined in the Declaration of Helsinki and the Belmont Report, ensuring that all participants are treated with respect for autonomy, beneficence, and justice. Before enrollment, all eligible participants received detailed written study information and an informed consent form explaining the study's purpose, procedures, potential risks, and benefits. Trained study investigators provided verbal explanations, in the presence of a witness when required, ensuring that participants fully understood their rights before providing consent. Participants were given sufficient time to consider their decision and had the opportunity to ask questions before signing the consent form. For illiterate participants, a thumbprint and a witness signature were obtained as a formal record of consent.

Participants provided written consent for the collection and analysis of biological samples (if applicable) and other relevant study data. They were informed that participation in the study is voluntary and that they have the right to withdraw consent at any time without facing any negative consequences related to health care access or other personal matters. In the event of withdrawal, participants may request the deletion of their collected data and the destruction of any stored biological samples, which will be carried out in accordance with ethical guidelines.

Throughout the study, participants will be informed about any abnormal findings identified from data collection. When relevant, they will be referred to appropriate health care facilities for further evaluation and treatment. Study participation will not interfere with any ongoing medical care, and participants

XSL•FO
RenderX

will continue receiving their routine health care services. The intervention is designed to be low risk, ensuring that no harm is caused to participants. If any damage occurs to intervention components (eg, monitoring devices or other study-related materials), repairs or replacements will be provided as necessary.

Eligible households were fully informed about the potential risks and benefits of participation. No direct financial compensation was provided for participation, but material costs for intervention-related maintenance (if applicable) may be covered. Participants have the freedom to refuse or discontinue participation at any point without justification. The study team will track and document reasons for withdrawal and report findings in subsequent publications. No invasive procedures will be conducted beyond standard diagnostic assessments, and no additional interventions will be performed outside the scope of this study.

Confidentiality

To ensure the privacy and confidentiality of participants, all data collected in this study are entered into the EDC system via tablets and securely shared through the EDC platform. All data are encrypted and stored on a secure server at the Information Center of the Chinese Academy of Medical Sciences, in strict compliance with national data protection regulations. Personal identifiers are stored separately from health data, using a coded numerical system for analysis. Only authorized research personnel have access to identifiable data, and all information is stored in a password-protected database with restricted access. Before any publication or public dissemination, all participant information will be fully anonymized. The study strictly adheres to ethical guidelines, ensuring that participants' rights and confidentiality are protected throughout the research process.

Results

Data collection for this trial began in June 2024 and is planned to end by March 2026. Baseline assessments, 3-month follow-ups, and 6-month follow-ups have been completed, and 12-month follow-up assessments are currently underway as scheduled. Primary and secondary outcome analyses are expected to be completed in 2026, with the main trial results anticipated to be disseminated in 2027.

Discussion

Overview

This protocol reports a large cluster randomized trial designed to evaluate whether a multicomponent, population medicine–oriented intervention can improve both mental health and COPD-related outcomes among adults at high risk for chronic respiratory disease in a resource-limited rural setting. We anticipate that integrating digital psychological support, community-based screening and follow-up, and a pay-for-population incentive mechanism into routine primary care will lead to meaningful improvements in depression, anxiety, and overall mental well-being, while simultaneously strengthening COPD case-finding, diagnosis, treatment uptake, and disease control. These hypotheses provide the conceptual basis for the discussion below and frame the potential contributions of the POPMIX-MH trial.

This trial makes several contributions that distinguish it from previous COPD- or mental health–related cRCTs. First, it focuses on a high-risk but understudied population in rural China—namely, individuals at high risk for COPD who also experience psychological distress—and evaluates an integrated package that jointly targets mental and respiratory health. Second, the intervention operationalizes population-based chronic care by combining community screening, digital CBT, chronic disease management, and a pay-for-population incentive model within routine primary care. Third, unlike most COPD trials that focus primarily on physiological indicators, this study positions mental well-being and common mental disorder symptoms as primary outcomes, thereby broadening the goals of chronic disease management. Together, these features provide rare real-world evidence on a scalable, multimorbidity-oriented intervention in a resource-limited setting.

Conceptual Contribution and Trial Rationale

Guided by the principles of population medicine, this study is the first to operationalize the modular, cost-effective population health intervention strategies proposed by the Lancet Commission on Investing in Health in a real-world setting—Xishui County, Guizhou Province, China [4]. Targeting individuals with mental health symptoms at high risk for COPD, the intervention integrates multiple components, including digital psychological support, health education, lung function screening, and structured follow-up. This design moves beyond the conventional disease-specific chronic care model and embodies the population medicine philosophy of addressing population-level health rather than individual illness alone [3]. As a designated national pilot zone for integrated primary health reform, Xishui County offers an ideal setting for testing the feasibility and policy relevance of such innovative strategies in a resource-limited rural context, marking the first real-world implementation of the Lancet Commission on Investing in Health framework.

Strengths and Limitations

Unlike most chronic disease intervention studies that primarily focus on physiological outcomes, such as lung function or exacerbation rates, the POPMIX-MH trial places mental health at the core of its outcome framework [42,43]. Using standardized instruments such as the PHQ-9, GAD-7, and WEMWBS, the study comprehensively assesses changes in depression, anxiety, and overall mental well-being. This approach directly responds to the high prevalence of underrecognized psychological distress among individuals at high risk for COPD, challenging the long-standing emphasis on organ-specific metrics in chronic disease care [44,45]. By combining digital CBT interventions (via EmoEase) with personalized health education, the program not only improves psychological outcomes but also enhances participants' understanding of their condition and self-management capabilities—factors that may, in turn, contribute to better physical health. Conceptually, this expands the definition of health in chronic care; operationally, it provides a scalable model for embedding mental health support into primary care systems.

XSL•FO
RenderX

While mental health outcomes constitute the primary end points of this study, a comprehensive set of secondary outcomes, such as COPD, asthma, hypertension, and diabetes care cascade performance; health care resource utilization; knowledge; high-risk behaviors and lifestyles; and socioeconomic profiles, has been included to capture the broader impact of the intervention across multiple dimensions of population health, including physical, mental, social, and environmental well-being [46].

The multicomponent intervention strategies employed in this study have all been recognized by the Lancet Commission on Investing in Health as cost-effective and high-priority [4]. In addition to digital health interventions, the package includes population-based screening, intensive counseling and follow-up, and a pay-for-population incentive mechanism for community health workers. Among these, population-based COPD screening, although proven to be cost-effective through the simulation model, has rarely been implemented at scale in real-world settings [47]. One of the key objectives of this study is to pilot this screening approach in rural regions such as Xishui County, enabling early identification and timely referral of suspected patients to higher-level hospitals for formal diagnosis and standardized treatment. Intensive counseling and regular follow-up are also core components of this study and are designed to ensure adherence to and the sustainability of the intervention. With ongoing psychological support, lifestyle guidance, and disease monitoring, participants are better equipped to maintain positive behavioral changes throughout the intervention period. These activities also represent fundamental responsibilities of community health workers and primary care physicians, as the long-term management of chronic and multimorbid conditions relies heavily on patients' compliance with medical advice and consistent engagement with follow-up care. To motivate active service delivery and transform traditional care models, this study implemented a pay-for-population incentive mechanism. Under this system, performance-based payments are tied to key health indicators, such as screening rates, diagnosis rates, treatment uptake, and disease control. This model encourages a shift from conventional "clinic-based passive care" to "proactive community-based service," prompting health care workers to go beyond the clinic to conduct screening and health promotion directly within the population.

Policy Implications

Our trial has strong policy implications. First, this study aims to systematically assess the feasibility and contextual adaptability of a multicomponent, population-based intervention within primary and community health systems in China and beyond. Unlike traditional single-disease, facility-centered models of chronic care, our approach emphasizes real-world integration of psychological support, behavior modification, disease screening, and workforce incentives into existing community health infrastructure. Experiences from this trial may offer practical guidance for scaling up community-based population health strategies in resource-limited and high-burden contexts similar to Xishui County, where the study is being conducted. Second, this study seeks to address a historically underappreciated area in chronic disease management: the widespread but unrecognized mental health burden among high-COPD-risk populations. By incorporating tools such as the PHQ-9, GAD-7, and WEMWBS into screening and follow-up workflows, we aim to support comprehensive assessment of participants' psychological status and apply scalable digital CBT-based interventions. This represents not only a technical innovation but also a conceptual shift, positioning mental health as a core component of chronic care, equal in priority to physiological outcomes.

Most importantly, we hope this study will catalyze a paradigm shift in the role of primary and community health workers. In conventional health systems, frontline providers often operate passively, waiting for patients to present at clinics. This trial introduces mechanisms such as pay-for-population, performance-linked process metrics, and home-based outreach to actively engage providers in population screening, health promotion, and continuous care. This transformation—from a patient-centered, reactive model to a proactive, population-centered approach—is at the heart of population medicine. Although this trial is conducted in a single rural county in southwest China with a high burden of COPD risk and common mental health problems, our first priority is to establish the effectiveness of the intervention within Xishui itself through a science-first, real-world evaluation. Given economic, cultural, and health-system differences across regions, adaptations will likely be required for implementation elsewhere. Nevertheless, we believe that the population-centered approach tested in this study has the potential to reshape public health governance not only in similar resource-limited primary care settings in China but also in other low- and middle-income countries striving for equitable, sustainable health system reform.

Dissemination

Upon completion of the trial, participant-related information and results will be provided to participants themselves or their families. The trial has gained public attention since it was first introduced in the Lancet Commission on Investing in Health. We will disseminate the results of the trial in international peer-reviewed journals and at other renowned international academic conferences. Authorship will be determined according to the International Committee of Medical Journal Editors (ICMJE) criteria. If any professional medical writers are used, their involvement will be transparently disclosed in the publication.

Conclusion

This trial represents a pioneering effort to integrate population medicine principles into the management of common mental health conditions in resource-limited rural settings. By combining community-based screening, digital CBT interventions, and team-based care supported by performance-linked incentives, the study aims to improve early detection, enhance access to effective treatment, and strengthen continuity of care. The findings are expected to provide robust evidence for scalable, cost-effective mental health service models and to inform national and international policy development, promoting equitable and sustainable improvements in population mental health.

XSL•FO
RenderX

Acknowledgments

This work was supported by the People's Government of Xishui County, the Health Bureau of Xishui County, the Youth League Committee of Guizhou Medical University, Medieco Group Co. einPaper Team, the Finance Bureau of Xishui County, the Health Insurance Bureau of Xishui County, the Bureau of Sports and Education of Xishui County, the Bureau of Statistics of Xishui County, the Xishui Sub-Bureau of the Ecology and Environment Bureau of Zunyi City, the Water Authority of Xishui County, the Bureau of Housing and Urban-Rural Development of Xishui County, the Meteorological Bureau of Xishui County, the Centers for Disease Control and Prevention of Xishui County (Health Surveillance Station of Xishui County), the People's Hospital of Xishui County, the Traditional Chinese Medicine Hospital of Xishui County, the Maternal and Child Health Care Hospital of Xishui County, and the People's Governments (Offices) of 26 Townships (Subdistricts) and Village Committees of Xishui County. Y Liu acknowledges the support of the China Scholarship Council Program (Project ID 202406210195).

The following are the collaborators of this work: The People's Government of Xishui County, Health Bureau of Xishui County, Youth League Committee of Guizhou Medical University, Medieco Group Co. einPaper Team, Finance Bureau of Xishui County, Health Insurance Bureau of Xishui County, Bureau of Sports and Education of Xishui County, Bureau of Statistics of Xishui County, Xishui Sub-Bureau of Ecology and Environment Bureau of Zunyi City, Water Authority of Xishui County, Bureau of Housing and Urban-Rural Development of Xishui County, Meteorological Bureau of Xishui County, Centers for Disease Control and Prevention of Xishui County (Health Surveillance Station of Xishui County), People's Hospital of Xishui County, Traditional Chinese Medicine Hospital of Xishui County, Maternal and Child Health Care Hospital of Xishui County, The People's Governments (Offices) of 26 Townships (Subdistricts), and Village Committees of Xishui County.

The following individuals should also be considered study collaborators: Chun Zhang, Yuzhu Ye, Mingqiang Hou, Shengbo Liao, Yuan Zhou, Dong Liu, Chentao Zhong, Yushuang Zhao, Hongxia Mu, Min Liu, Qiong Lyu, Zhengyu Zhang, Weiwei Li, Jian Wang, Tiansheng Lan, Tingyu Liu, Xin Zhou, Bie Yu, Xianping Wang, Yansong Luo, Wenwu Jiang, Wen Zhang, Dongmei Wu, Qingping Luo, Dajun Rao, Cunkun Yang, Feng Zhao, Qin Wang, Linlin Zhang, Can Li, Min Chen, Yufeng Huang, Guanghai Jian, Li Wang, Yunjiang Zheng, Jianghong Linghu, Li Zhong, Xiaoling Ma, Bin Gui, Yan Cheng, Yu Yang, Jing Yan, Xixi Liao, Mingguo Xu, Zhijun Zhao, Qian Han, Panyan Wu, Dandan Tang, Qingqing Wang, Li Feng, Lilan Wang, Lei Si, Peng Zeng, Jianxiu Tian, Weili Fan, Xin Luo, Sumei Feng, Rong Ding, Huadan Yi, Fengxue Lyu, Ludan Yuan, Yuan Liu, Xueshuang Wang, Aihong Duan, Yang Zhao, Yuping Luo, Yicheng Li, Anchun Liang, Chengmei Li, Hu Tang, Anfeng Liu, Xu Wang, Yi Zhong, Wen Zheng, Benwei Zhao, Qianmei Zhang, Xiaojing Tu, Huanhuan Lu, Xiumei Zhao, Yan Chen, Xuemei Yuan, Yanhong Mu, Yao Zhao, Juan Zhang, Qinhui Li, Fuhui Wang, Xiaomin Mu, Yan Hu, Liuyang Chen, Gang Lei, Zhufei Hu, Xiangyong Tang, Yan Luo, Shunxia Zhou, Ting Pan, Yuanyin Zhou, Xiaoqin Ren, Yaling Zhou, Jing Zeng, Keping Hu, Xiaoyu Liu, Deguo Yang, Qiong Wan, Cai Zhang, Heng He, Chengping Wang, Yuanzhong Wang, Shuyin Wang, Xuhong Lei, Yushen Yuan, Tailiang Luo, Dazhou Zhu, Anrong Luo, Xianmei Hu, Ya Zhang, Ling Yuan, Senlin Luo, Weili Ma, Yonghua Cao, Yunpeng Cao, Taojin Chen, Wen Chen, Zhiqun Chen, Zhongyi Chen, Huagang Cheng, Xianpan Feng, Taimei Gao, Lizhu Jian, Renxian Jiang, Shengmei Kang, Hang Lei, Shanshan Liu, Ruxun Lu, Xiaojiao Luo, Dongshun Lyu, Guanghong Meng, Jianping Mu, Xuerou Mu, Hongxia Qian, Huimin Ren, Juan Wang, Li Wang, Loutao Wang, Qin Wang, Wei Wang, Ya Wang, Yan Wang, Dongjie Wu, Xiaoxia Xu, Jing Yang, Qian Yu, Dan Yuan, Yanlu Yuan, Yu Yuan, Hongying Yue, Yuanmin Yue, Li Zhang, Zhizhong Zhang, Jiaqin Zhao, Yuanxin Zhao, Yue Zhong, Lejie Zhou, Lufang Yang, Xue Chen, Lihui Ren, Guixiang Gong, Ping Shi, Jiangshun Tan, Zhonghang Yu, Lijun Mu, Meilin Chen, Su Yu, Lin Li, Xianbi Mu, Jianxia Wu, Mei Zhao, Jiangmei Chen, Jing Huang, Yanfang Li, Huayan Yang, Jiaheng Yang, Xingqi Yuan, Yunhao Zhou, Mengting Ye, Junlin Zhang, Taijun Luo, Shuai Wang, Ju Hu, Gang Yuan, Shili Mu, Zhongxiang Yuan, Huili He, Jiali Li, Chunlan Liu, Kailiang Cheng, Quan Liu, Huade Yang, Shaojun Zhang, Bo Chen, Guishan Ying, Xumei Zhao, Zhi He, Gan Feng, Huanhuan Guo, Hongling Yuan, Fuyou Yuan, Wei Wang, Lang Chen, Liping Zhao, Jiajun Ning, Jie Feng, Yuxian Chen, Jin Chen, Jing Qi, Yuanting Li, Lingli Zhao, Guo Wang, Ziyi Wang, Huaishu Zhong, Jie Wang, Gang Chen, Xueqin Mu, Yuanfeng Gong, Xinan Lu, Zhengli Chen, Tianwei Chen, Xiaoping Chen, Runqin Mu, Xiaoyi Deng, Yu Duan, Xiujuan Dong, Wangyong Yan, Jiayong Rao, Mingfang Wang, Xiangcai Chen, Zhongquan Wang, Lanfang Zhang, Jianghua Luo, Kaijie Huang, Ye Chen, Jin'e Li, Li Huang, Ju Chen, Hao Huang, Qingqun Long, Yan Chen, Ling Jiang, Man Ding, Xiaolin Yu, Tongjie Wei, Zhengqin Liu, Kaifu Si, Yuxian Mu, Yan Huang, Qian Yang, Tao Xu, Xiaorong Yuan, Cailun Zhao, Minli Chen, Rongfang Luo, Ling Chen, Ying Zhang, Guiwei Yuan, Qianyi Shui, Yong Wu, Qian Huang, Changli Wu, Nan Zhang, Chaojiang Wu, Defei Chen, Wu Xu, Linhui He, Wanlun Zhang, Jinsong Chen, Jin Yuan, Rujun Li, Minghe Wang, Chunmei Zhang, Yueshan Wang, Lang Liao, Anqian Feng, Dengyang Wang, Zelin Deng, Hupiao Yang, Kaimeng You, Lijuan Deng, Qiong Huang, Xiaoqing Zhang, Zhengqian Zhu, Xiaoyan Liao, Yongbo Song, Yushui Yuan, Qinghong Cai, Lu Liu, Hong Yu, Xue Luo, Xiaoshan Huang, Qin Hu, Dian Yuan, Zengye Zhao, Jiemin Zeng, Zhaoyan Huang, Peng Luo, Taibo Luo, Wenwu Wu, Guangni Liu, Gengning Zhang, Mingyuan Ren, Mei Yang, Zhongjin Wang, Rongping Yuan, Yunfeng Yuan, Anquan Yuan, Yunqiang Chen, Hongxian Liu, Xingyu He, Tongxiang Zhang, Mingkai Xu, Minqing Hu, Chunyi Qian, Yunfen Zhang, Yueguang Lei, Jun Wang, Xiaoju Yuan, Changxian Yang, Ai Xiang, Minjing Zhao, Xiaomei Ren, Guimei Li, Huiying Ren, Ju Teng, Yu Chen, Jing Yuan, Shunfeng Mu, Xiangyu Zhang, Ye Li, Yuwei Sun, Shangjing Wang, Fangping He, Guoxiang Wang, Wei Cai, Weijiang Tian, Tingfen Cai, Zhao Chen, Dakai Mu, Yufu Mu, Cai Cheng, Xiaofang Pan, Fang Wang, Mingqian Zhao, Qijun Liu, Bangqi Yang, Ye Yu, Ting Min, Xiaoya Wang, Jun Li, Yang Zhang, Yan Wang, Lihong Zhao, Zhihua Luo,

XSL•FO
RenderX

Hao Yuan, Huihua Ruan, Jian Wu, Chengfen Wang, Dezhi Zhang, Yunfang Zhang, Weining Ao, Pan Zhao, Yunxia Liang, Yicheng Zhao, Xiaosong Mu, Lyufang Luo, Ziyi Zhang, Wei Huang, Jun Cao, Huan Li, Jun Liu, Liye Yuan, Na Zhao, Ju Wang, Qiuyan Duan, Weizhong Chen, Lirong Zhao, Jiangbo Wu, Liqin Deng, Kunyu Wang, Zhengguo Yan, Chaolian Wang, Can Liu, Quan Cao, Gang Ma, Youyong Lu, Yuan Zhang, Kaijing Xiang, Yixiang Wang, Xingye Zhu, Jiamei Wang, Demin Hu, Dongmin Li, Peng Chen, Chao Tian, Lu Wang, Mei Wang, Lulu Xiao, Lingling Yu, Min Wei, Chengcheng Yang, Limei Zhang, Rong Zhao, Chongguang Zhong, Yan Xiao, Jian Luo, Hongfa Chen, Pengju Dai, Jin Liu, Yingtao Luo, Guangjie Ma, Shiyou Mu, Zhengli Rao, Deqin Wen, Guowang Wu, Yong Hu, Xiaorong Shi, Shaofen Huang, Ya Chen, Wenmao Zhang, Min Chen, Xin Luo, Hongqiong Ruan, Liming Liu, Lihui Fan, Xiaoyan Zou, Huizhi He, Xinggao Yang, Linfei Wang, Hong Wang, Wenxiang Li, Jialian Wu, Jinyong Yang, Jianying Wu, Guangliang Liu, Xiaolin Xiong, Gang Yang, Ping Feng, Cheng Luo, Qiwei Zhang, Qihui Yang, Taoshan Ni, Rensong Tian, Li Zhang, Yuanmei Zhang, Hui Bi, Bijue Huang, Yu Yan, Shan Chen, Xinghui Chen, Jiangfei Wang, Guanglun Zou, Junting Liu, Rongjiang Zhao, Li Yuan, Ruyin Yuan, Jianbing Li, Aiping Wang, Daijie Wang, Mei Yang, Kaiyan Si, Peng Wang, Yan Linghu, Tuye Yuan, Zhiyuan Yuan, Weiwei Qian, Daimei Wang, Rongqin Ren, Dong Dai, Qianhua Deng, Yan Xu, Jiali Zhang, Qingyan Bai, Yingwen Zhang, Xingke Wang, Qiang Liu, Liyuan Feng, Fancha Wang, Shiwen Wang, Xiangguang Liu, Runping Liu, Xiaoying Wang, Daizhi Wang, Huihai Cao, Jiang Yuan, Qinhui Cao, Min Luo, Hao Hou, Lu Kong, Zhu Ren, Yu Zhou, Xiaoyan Yuan, Hong Liu, Shixian Gong, Yonghong Zhao, Rusheng Wang, Ji Chen, Fuping Bai, Yang Zhou, Wenbi Chen, Yu Sun, Qunying Xu, Gang Zheng, Yong Wang, Juan Chen, Ling Zhong, Ji Ren, Qiling Zhu, Zheng Yang, Yongsong Chen, Zhonghui Huang, Xiaojiang Zhang, Junyan Yu, Panting Yu, Qun Wang, Qimei Yu, Hao Zhou, Yuanmei Lu, Yiyu Wang, Min Huang, Feng Wu, Jiangyu Yuan, Guangjun Xiong, Jian Zhang, Hong Zhou, Fenfang Liu, Yuqin Chen, Xiaoli Ye, Qingli Luo, Qijie Dai, Yunxian Feng, Tianmei Huang, Yunfeng Xia, Lei Zhao, Anshui Wang, Fuxiu Yang, Minqin Hu, Tingting Yuan, Song Zhou, Lin Yuan, Mingjian Yuan, Maozhao Wen, Hongming Zhao, Hongbo Zhao, Yuanchao Zhao, Binyuan Yuan, Hua Xiong, Qiyong Yuan, Qin Yuan, Xingqin Zhao, Qifu Luo, Qigui Luo, Tulin Yuan, Xianglin Huang, Chun Luo, Dehong Wang, Rongwei Zhang, Xingji Chen, Zhong'en Wang, Deping Yuan, Qihua Luo, Huahua Li, Yueming Wang, Jian Chen, Wei Teng, Jisheng Wang, Mingli He, Shan Wu, Junjun Xia, Tingxian Mao, Tingmin Luo, Sihong Yi, Ruixue Li, Kang Yang, Chengshang Ao, Xu Zhong, Yang Liu, Langsha Chen, Huan Luo, Qingxia Wang, Chen Chen, Wenqiang Wu, Yumei Wu, Limei Feng, Caixian Lei, Kai Li, Youqun Chen, Lulu Zhao, Han Ren, Juhang Yang, Shanshan Luo, Xiaoyi Zhang, Huiqin Yuan, Li Chen, Guishuang Liu, Jiakuan Ren, Dongqiao Jiang, Yi Yuan, Dengmin Ren, Xiaojiang Mu, Xue Yang, Wuyao He, Wanyi Hu, Piaopiao Huang, Guijiang Li, Jinman Li, Suosuo Li, Yiling Lu, Yuqi Lu, Huan Luo, Qingqing Luo, Wenqin Mu, Daolin Qin, Jinlan Quan, Yingying Tang, Junbao Wang, Lingling Wang, Rui Xu, Chenxi Yang, Dengqin Yang, Liqin Yang, Qingting Zhang, Lei Zhao, Lin Zhao, Jingmei Zhou, Lijuan Zuo, Xiaoning Xia, Yujia Xie, Hangyu Chen, Li Yang, Yu Zhao, Sisi Li, Juan Pan, Ju Yang, Minjia Zhao, Meiyuan Zhou, Guimei Hu, Chunqin Huang, Haiyan Long, Zeyuan Luo, Donghua Pan, Xuelang Ying, Yanyan Zhang, Yuting Zhang, Hong Zhao, Yunli Zhou, Bijin Zhu, Kang Chen, Kaidi Ding, Jiacheng Dong, Tai Fu, Qin Li, Zaijing Liu, Siqi Huang, Lingli Ou, Zhuoting Tan, Yan Tang, Qingzhong Wu, Xinyi Xu, Bingyang Xu, Yulin Zhang, Shuangshuang Zhou, Zena Zhu, Jiancao Zuo, Guangxianfeng Chen, Jiahui Luo, Zhiqi Duan, Shangrong Gu, Binyao Ou, Xingtai Zhao, Yanran He, Xingli Liao, Shunxin Shi, Cheng Wang, Jiayu Wang, Rongmei Wang, Ruirui Zhang, Yuhan Zhou, Serena Luo, Jiwei Jin, Xiaoqi Shi, Ziqianqian Yang, Xiaoxue An, Jiwei Jin, Mingli Huang, Yinghua Mu, Xiaozheng Wang, Songlin Zhang, Xue Zhou, Ruirui Ao, Shaohuan Bao, Yating Jiang, Yijun Liu, Zhirui Nie, Dingyuan Rao, Jiangqin Wan, Xue Wang, Die Wei, Lijin Xie, Jianan Xu, Xuerong Zhang, Die Deng, Jingyu Fu, Shunli Guo, Binhan He, Sitong Liu, Lin Liu, Linqu Qin, Teng Tian, Kehao Wang, Qirui Wu, Laitian Ye, Xiao Zhang, Tiantian Zhou, Sitong Zhou, Jiajia Liu, Huali Yin, Yue Luo, Jianxiu Bai, Yuchan Li, Xinglian Zhang, Shenghui Zhu, Xue Chen, Dan Jiang, Xiaoqin Yang, Changnan Zhao, Fangjing Geng, Jia Li, Jie Wang, Wenfan Hu, Xingling Li, Yumei Lu, Zhuying Long, Junxue Qian, Guoxian Shi, Ruoxi Wang, Guangjian Wu, Guangyan Zhang, Rongrong Zhou, Hongyu Hu, Zhaoyi Ren, Lingxuan Shi, Meng Zeng, Shengsheng Li, Ying Qin, Yi Zhou, Tianyu Liu, Jia He, Jinan Li, Mingyuan Mao, Jiayi Wang, Ronghui Ma, and Bin Li.

Disclaimer

The authors declare the use of generative artificial intelligence in the research and writing process. According to the GAIDeT (Guideline for Reporting Evidence-based Practice Educational Interventions and Teaching) standards (2025), the following tasks were delegated to generative artificial intelligence tools under full human supervision: proofreading and editing, adapting and adjusting emotional tone, reformatting, and quality assessment. The generative artificial intelligence tools used were ChatGPT 5.1 (OpenAI) and DeepSeek (High-Flyer). Responsibility for the final manuscript lies entirely with the authors. Generative artificial intelligence tools are not listed as authors and do not bear responsibility for the final outcomes.

Data Availability

The datasets used or analyzed in this study are available from the principal investigator (CW) on reasonable request.

Funding

This study was supported by the Noncommunicable Chronic Diseases-National Science and Technology Major Project (2024ZD0522300), the CAMS Innovation Fund for Medical Sciences (2023-I2M-2-001), the Non-Profit Central Research Institute

XSL•FO
RenderX

Fund of the Chinese Academy of Medical Sciences (grant 2022-ZHCH330-01), the State Key Laboratory Special Fund (2060204), and the Chinese Academy of Medical Sciences (CAMS) Innovation Fund for Medical Sciences (2022-I2M-CoV19-003). Additional funding was provided by the Tencent Sustainable Social Value Inclusive Health Lab (SSVPJ202307060001), the EU Horizon Europe Programme (HORIZON-MSCA-2021-SE-01;101086139-PoPMeD-SuSDeV), and the China Medical Board (grant 22-469 to SC).

Authors' Contributions

Conceptualization: CW, TY, SC, Y Liu, KH, TB
Data curation: KH, X Tong, Y Liu, WC
Investigation: Y Liu, WC, SZ, YW, ZZ, X Tang, ZC, JZ, LH, LJ, TZ, Y Luo, QL, X Lyu
Methodology: CW, TY, SC, Y Liu, WC, SZ, YW, TB, PG, AB
Project administration: CW, TY, SC, LT
Supervision: CW, TY, SC, TB
Validation: CW, TY, SC
Visualization: Y Liu, KH, ZZ, SZ
Writing – original draft: Y Liu, WC, SZ
Writing – review & editing: All
All authors contributed to revising the manuscript and to this protocol in accordance with the International Committee of Medical Journal Editors' guidelines. Y Liu, WC, SZ, and YW contributed equally to this work and are designated as cofirst authors. TB is designated as the cosenior author. CW, SC, and TY are the corresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2025 checklist.
[DOCX File , 32 KB - resprot_v15i1e85853_app1.docx]

Multimedia Appendix 2
The pay-for-population incentive mechanism.
[DOCX File , 22 KB - resprot_v15i1e85853_app2.docx]

Multimedia Appendix 3
Secondary outcomes and trial timeline.
[DOCX File , 27 KB - resprot_v15i1e85853_app3.docx]

Multimedia Appendix 4
Full size versions of Figures 1–3.
[ZIP File (Zip Archive), 3637 KB - resprot_v15i1e85853_app4.zip]

References

1. Murray CJL. The global burden of disease study at 30 years. Nat Med 2022 Oct;28(10):2019-2026. [doi: 10.1038/s41591-022-01990-1] [Medline: 36216939]
2. GBD 2019 Risk Factors Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1223-1249 [FREE Full text] [doi: 10.1016/S0140-6736(20)30752-2] [Medline: 33069327]
3. Wang C, Chen S, Shan G, Leng Z, Bärnighausen T, Yang W. Strengthening population medicine to promote public health. Chin Med J (Engl) 2022 May 20;135(10):1135-1137 [FREE Full text] [doi: 10.1097/CM9.0000000000002221] [Medline: 35788092]
4. Jamison DT, Summers LH, Chang AY, Karlsson O, Mao W, Norheim OF, et al. Global health 2050: the path to halving premature death by mid-century. The Lancet 2024 Oct;404(10462):1561-1614. [doi: 10.1016/s0140-6736(24)01439-9]
5. Verguet S, Bolongaita S, Chang AY, Cardoso DS, Stevens GA. The economic value of reducing mortality due to noncommunicable diseases and injuries. Nat Med 2024 Nov 27;30(11):3335-3344. [doi: 10.1038/s41591-024-03248-4] [Medline: 39333314]
6. GBD 2021 Causes of Death Collaborators. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease

XSL•FO
RenderX

Study 2021. Lancet 2024 May 18;403(10440):2100-2132 [FREE Full text] [doi: 10.1016/S0140-6736(24)00367-2] [Medline: 38582094]
7. de Oca MM, Perez-Padilla R, Celli B, Aaron SD, Wehrmeister FC, Amaral AFS, et al. The global burden of COPD: epidemiology and effect of prevention strategies. The Lancet Respiratory Medicine 2025 Aug;13(8):709-724. [doi: 10.1016/s2213-2600(24)00339-4]
8. Chen S, Kuhn M, Prettner K, Yu F, Yang T, Bärnighausen T, et al. The global economic burden of chronic obstructive pulmonary disease for 204 countries and territories in 2020–50: a health-augmented macroeconomic modelling study. The Lancet Global Health 2023 Aug;11(8):e1183-e1193. [doi: 10.1016/s2214-109x(23)00217-6]
9. GBD 2015 Chronic Respiratory Disease Collaborators. Global, regional, and national deaths, prevalence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Respir Med 2017 Sep;5(9):691-706 [FREE Full text] [doi: 10.1016/S2213-2600(17)30293-X] [Medline: 28822787]
10. Wang C, Qi W, Yang T, Jiao L, Chen Q, Huang K, et al. The care cascade of chronic obstructive pulmonary disease in China: a cross-sectional study of individual-level data at enrolment into the national 'Happy Breathing' Programme. EClinicalMedicine 2024 Aug;74:102597 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102597] [Medline: 39114273]
11. Fabbri LM, Celli BR, Agustí A, Criner GJ, Dransfield MT, Divo M, et al. COPD and multimorbidity: recognising and addressing a syndemic occurrence. The Lancet Respiratory Medicine 2023 Nov;11(11):1020-1034. [doi: 10.1016/s2213-2600(23)00261-8]
12. Matte DL, Pizzichini MM, Hoepers AT, Diaz AP, Karloh M, Dias M, et al. Prevalence of depression in COPD: a systematic review and meta-analysis of controlled studies. Respir Med 2016 Aug;117:154-161 [FREE Full text] [doi: 10.1016/j.rmed.2016.06.006] [Medline: 27492526]
13. Tsai T, Livneh H, Lu M, Tsai P, Chen P, Sung F. Increased risk and related factors of depression among patients with COPD: a population-based cohort study. BMC Public Health 2013 Oct 19;13:976 [FREE Full text] [doi: 10.1186/1471-2458-13-976] [Medline: 24138872]
14. Thapa N, Maharjan M, Shrestha TM, Gauchan S, Pun P, Thapa YB. Anxiety and depression among patients with chronic obstructive pulmonary disease and general population in rural Nepal. BMC Psychiatry 2017 Dec 12;17(1):397 [FREE Full text] [doi: 10.1186/s12888-017-1550-5] [Medline: 29233103]
15. Wang JG, Bose S, Holbrook JT, Nan L, Eakin MN, Yohannes AM, et al. Clinical characteristics of patients with COPD and comorbid depression and anxiety: data from a national multicenter cohort study. J COPD F 2025;12(1):33-42. [doi: 10.15326/jcopdf.2024.534]
16. Barrueco E, Hernández-Mezquita MA, Hidalgo-Sierra V, Cordovilla R, Olivera-Pueyo J, Galán J. Diagnosis of anxiety in COPD patients: usefulness of the HADS test in primary care and pulmonology services. J Pers Med 2024 Jul 02;14(7):713 [FREE Full text] [doi: 10.3390/jpm14070713] [Medline: 39063967]
17. Pauwels R. Global initiative for chronic obstructive lung diseases (GOLD): time to act. Eur Respir J 2001 Dec;18(6):901-902 [FREE Full text] [Medline: 11829093]
18. Fabbri L, Hurd S, GOLD Scientific Committee. Global strategy for the diagnosis, management and prevention of COPD: 2003 update. Eur Respir J 2003 Jul 01;22(1):1-2 [FREE Full text] [doi: 10.1183/09031936.03.00063703] [Medline: 12882441]
19. Han MK, Ye W, Wang D, White E, Arjomandi M, Barjaktarevic IZ, et al. Bronchodilators in tobacco-exposed persons with symptoms and preserved lung function. N Engl J Med 2022 Sep 29;387(13):1173-1184. [doi: 10.1056/nejmoa2204752]
20. US Preventive Services Task Force (USPSTF), Siu AL, Bibbins-Domingo K, Grossman DC, Davidson KW, Epling JW, et al. Screening for chronic obstructive pulmonary disease: US Preventive Services Task Force recommendation statement. JAMA 2016 Apr 05;315(13):1372-1377 [FREE Full text] [doi: 10.1001/jama.2016.2638] [Medline: 27046365]
21. Yohannes AM, Alexopoulos GS. Depression and anxiety in patients with COPD. Eur Respir Rev 2014 Sep;23(133):345-349 [FREE Full text] [doi: 10.1183/09059180.00007813] [Medline: 25176970]
22. Ye Z, Chen L. Research progress of chronic obstructive pulmonary disease with anxiety and depression. Chin J Respir Crit Care 2019 Sep;18:491-494. [doi: 10.7507/1671-6205.201811011]
23. Lei J, Huang K, Pan J, Li W, Niu H, Ren X, et al. The national COPD screening programme in China: rationale and design. ERJ Open Res 2023 Mar 02;9(2):00597-2022 [FREE Full text] [doi: 10.1183/23120541.00597-2022] [Medline: 37101739]
24. Poot C, Meijer E, Kruis A, Smidt N, Chavannes N, Honkoop P. Integrated disease management interventions for patients with chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2021 Sep 08;9(9):CD009437 [FREE Full text] [doi: 10.1002/14651858.CD009437.pub3] [Medline: 34495549]
25. Kaczkurkin AN, Foa EB. Cognitive-behavioral therapy for anxiety disorders: an update on the empirical evidence. Dialogues in Clinical Neuroscience 2022 Apr 01;17(3):337-346. [doi: 10.31887/dcns.2015.17.3/akaczkurkin]
26. Chambless DL, Ollendick TH. Empirically supported psychological interventions: controversies and evidence. Annu Rev Psychol 2001 Feb;52(1):685-716. [doi: 10.1146/annurev.psych.52.1.685] [Medline: 11148322]
27. Karyotaki E, Efthimiou O, Miguel C, Bermpohl FMG, Furukawa TA, Cuijpers P, Individual Patient Data Meta-Analyses for Depression (IPDMA-DE) Collaboration, et al. Internet-based cognitive behavioral therapy for depression: a systematic

XSL•FO
RenderX

review and individual patient data network meta-analysis. JAMA Psychiatry 2021 Apr 01;78(4):361-371 [FREE Full text] [doi: 10.1001/jamapsychiatry.2020.4364] [Medline: 33471111]
28. Espie CA, Emsley R, Kyle SD, Gordon C, Drake CL, Siriwardena AN, et al. Effect of digital cognitive behavioral therapy for insomnia on health, psychological well-being, and sleep-related quality of life: a randomized clinical trial. JAMA Psychiatry 2019 Jan 01;76(1):21-30 [FREE Full text] [doi: 10.1001/jamapsychiatry.2018.2745] [Medline: 30264137]
29. Han MK, Ye W, Kim D, Woodruff P, Pulmonary Trials Cooperative Investigators. Design of the redefining therapy in early COPD study. Chronic Obstr Pulm Dis 2020 Oct;7(4):382-389 [FREE Full text] [doi: 10.15326/jcopdf.7.4.2020.0157] [Medline: 32989941]
30. Chan A, Boutron I, Hopewell S, Moher D, Schulz K, Collins G, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. The Lancet 2025 May;405(10491):e19-e27. [doi: 10.1016/S0140-6736(25)00770-6]
31. Zhou Y, Chen S, Tian J, Cui J, Li X, Hong W, et al. Development and validation of a chronic obstructive pulmonary disease screening questionnaire in China. Int J Tuberc Lung Dis 2013 Dec 01;17(12):1645-1651. [doi: 10.5588/ijtld.12.0995] [Medline: 24200283]
32. Zhang M, Yang L, Wang L, Jiang Y, Huang Z, Zhao Z, et al. Trends in smoking prevalence in urban and rural China, 2007 to 2018: findings from 5 consecutive nationally representative cross-sectional surveys. PLoS Med 2022 Aug;19(8):e1004064 [FREE Full text] [doi: 10.1371/journal.pmed.1004064] [Medline: 36006870]
33. Chen S, Chen W, Li Y, Yu Y, Chen Q, Jiao L, et al. Effectiveness and cost-effectiveness of an integrated digital psychological intervention (EmoEase) in Chinese chronic obstructive pulmonary disease patients: study protocol of a randomized controlled trial. Digit Health 2024;10:20552076241277650 [FREE Full text] [doi: 10.1177/20552076241277650] [Medline: 39381816]
34. World Health Organization. Package of Essential Noncommunicable (PEN) Disease Interventions for Primary Health Care in Low-Resource Settings. Geneva, Switzerland: World Health Organization; Oct 29, 2014.
35. World Health Organization. Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycemia. Geneva, Switzerland: World Health Organization; Apr 21, 2006.
36. Wang W, Bian Q, Zhao Y, Li X, Wang W, Du J, et al. Reliability and validity of the Chinese version of the Patient Health Questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry 2014 Sep;36(5):539-544. [doi: 10.1016/j.genhosppsych.2014.05.021] [Medline: 25023953]
37. Tong X, An D, McGonigal A, Park S, Zhou D. Validation of the Generalized Anxiety Disorder-7 (GAD-7) among Chinese people with epilepsy. Epilepsy Res 2016 Feb;120:31-36. [doi: 10.1016/j.eplepsyres.2015.11.019] [Medline: 26709880]
38. Dong A, Zhang X, Zhou H, Chen S, Zhao W, Wu M, et al. Applicability and cross-cultural validation of the Chinese version of the Warwick-Edinburgh mental well-being scale in patients with chronic heart failure. Health Qual Life Outcomes 2019 Mar 29;17(1):55 [FREE Full text] [doi: 10.1186/s12955-019-1120-2] [Medline: 30922342]
39. Mirea D, Martin-Key NA, Barton-Owen G, Olmert T, Cooper JD, Han SYS, et al. Impact of a web-based psychiatric assessment on the mental health and well-being of individuals presenting with depressive symptoms: longitudinal observational study. JMIR Ment Health 2021 Feb 22;8(2):e23813 [FREE Full text] [doi: 10.2196/23813] [Medline: 33616546]
40. Shackleton N, Hale D, Bonell C, Viner R. Intraclass correlation values for adolescent health outcomes in secondary schools in 21 European countries. SSM Popul Health 2016 Dec;2:217-225 [FREE Full text] [doi: 10.1016/j.ssmph.2016.03.005] [Medline: 29349141]
41. Elley CR, Kerse N, Chondros P, Robinson E. Intraclass correlation coefficients from three cluster randomised controlled trials in primary and residential health care. Aust N Z J Public Health 2005 Oct;29(5):461-467 [FREE Full text] [doi: 10.1111/j.1467-842x.2005.tb00227.x] [Medline: 16255449]
42. Wedzicha JA, Seemungal TA. COPD exacerbations: defining their cause and prevention. The Lancet 2007 Sep;370(9589):786-796. [doi: 10.1016/s0140-6736(07)61382-8]
43. Pauwels RA, Löfdahl C, Laitinen LA, Schouten JP, Postma DS, Pride NB, et al. Long-term treatment with inhaled budesonide in persons with mild chronic obstructive pulmonary disease who continue smoking. N Engl J Med 1999 Jun 24;340(25):1948-1953. [doi: 10.1056/nejm199906243402503]
44. Bury M. Chronic illness as biographical disruption. Sociol Health Illn 1982 Jul 28;4(2):167-182 [FREE Full text] [doi: 10.1111/1467-9566.ep11339939] [Medline: 10260456]
45. Engel GL. The need for a new medical model: a challenge for biomedicine. Science 1977 Apr 08;196(4286):129-136. [doi: 10.1126/science.847460] [Medline: 847460]
46. Wang C, Chen S, Shao R, Yang W. Redefining human health: physical wellbeing, mental wellbeing, social wellbeing, and environmental wellbeing. Chin Med J (Engl) 2023 Oct 20;136(20):2395-2396 [FREE Full text] [doi: 10.1097/CM9.0000000000002817] [Medline: 37649389]
47. Chen Q, Fan Y, Huang K, Li W, Geldsetzer P, Bärnighausen T, et al. Cost-effectiveness of population-based screening for chronic obstructive pulmonary disease in China: a simulation modeling study. Lancet Reg Health West Pac 2024 May;46:101065 [FREE Full text] [doi: 10.1016/j.lanwpc.2024.101065] [Medline: 38721063]

Abbreviations

CBT: cognitive behavioral therapy

XSL•FO
RenderX

COPD: chronic obstructive pulmonary disease
COPD-SQ: COPD Screening Questionnaire
cRCT: cluster randomized controlled trial
DSMB: Data and Safety Monitoring Board
EDC: electronic data capture
FEV1: forced expiratory volume in 1 second
FVC: forced vital capacity
GAD-7: 7-item Generalized Anxiety Disorder
ICC: intraclass correlation coefficient
NCD: noncommunicable disease
PHQ-9: 9-item Patient Health Questionnaire
POPMIX: Population Medicine Multimorbidity Intervention in Xishui County
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
WEMWBS: Warwick-Edinburgh Mental Well-Being Scale

Edited by J Sarvestan; submitted 14.Oct.2025; peer-reviewed by PH Chang, Z Hao, F Bao, Z Jin, S Wang; comments to author 27.Nov.2025; accepted 18.Dec.2025; published 06.Mar.2026.

Please cite as:
Liu Y, Chen W, Zhang S, Wang Y, Zheng Z, Huang K, Tang X, Cao Z, Tong X, Tang L, Zhao J, He L, Jiao L, Zhao T, Luo Y, Lai Q, Lyu X, Chen Q, Bunker A, Vollmer S, Geldsetzer P, Jamison D, Bärnighausen T, Yang T, Chen S, Wang C, POPMIX Group
Impact of the Population Medicine Multimorbidity Intervention in Xishui County (POPMIX) on People at High Risk for Chronic Obstructive Pulmonary Disease Who Experience Mental Health Symptoms: Protocol for the POPMIX-MH Cluster Randomized Controlled Trial
JMIR Res Protoc 2026;15:e85853
URL: https://www.researchprotocols.org/2026/1/e85853
doi:10.2196/85853
PMID:

©Yuhao Liu, Wenjin Chen, Shiyu Zhang, Yingping Wang, Zhoutao Zheng, Ke Huang, Xingyao Tang, Zhong Cao, Xunliang Tong, Lei Tang, Jinghan Zhao, Liu He, Lirui Jiao, Tianying Zhao, Yingchi Luo, Qiande Lai, Xiangqin Lyu, Qiushi Chen, Aditi Bunker, Sebastian Vollmer, Pascal Geldsetzer, Dean Jamison, Till Bärnighausen, Ting Yang, Simiao Chen, Chen Wang, POPMIX Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 06.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effects of 5-Minute Light-Intensity Physical Activity Breaks (Sit-Cycling and Stand-Twisting) on Cognitive Function, Sleepiness, and Back Pain in Physically Inactive University Students: Protocol for a Mixed-Design Laboratory Study

Jiameng Ma[1,2], PhD; Buan Kiong Terence Chua[1], MSc; You Hui Jacqueline Chin[1], BSc (Hons); Wei Peng Teo[1,3], PhD; Hyunshik Kim[1,2], PhD; Seow Ting Low[1], BSc (Hons); Dan Li[1], PhD; Yong Hwa Michael Chia[1], PhD

[1]Physical Education and Sports Science, National Institute of Education, Nanyang Technological University, Singapore, Singapore

[2]Faculty of Sports Science, Sendai University, Shibata-machi, Miyagi-ken, Japan

[3]Science of Learning in Education Centre, National Institute of Education, Nanyang Technological University, Singapore, Singapore

Corresponding Author:
Yong Hwa Michael Chia, PhD
Physical Education and Sports Science
National Institute of Education
Nanyang Technological University
General Office, Level 3
1 Nanyang walk
Singapore, 637616
Singapore
Phone: 65 6790 3719
Email: michael.chia@nie.edu.sg

Abstract

Background: University students may remain sedentary for large parts of the day, spending >9 hours per day sitting. Prolonged sitting is associated with increased risk for adverse health outcomes in adulthood. Light-intensity physical activity (LIPA) can be integrated into university students' daily routines to counteract the negative effects of prolonged sitting on cognitive function and musculoskeletal health.

Objective: This protocol describes a trial conducted in Singapore to evaluate the acute (immediate) and short-term delayed effects of two 5-minute LIPA modalities (sit-cycling [S-C] and stand-twisting [S-T]) on neurocognitive and somatic outcomes (prefrontal cortical oxygenation, cognitive performance, subjective sleepiness, and back pain) in physically inactive university students compared to a modality-matched passive control condition.

Methods: The study comprised a main study (N=22) and a methodological substudy (n=6, 27.3%). In the substudy, students' oxygen uptake was measured during 5-minute bouts of S-C and S-T to verify that both modalities meet the operational definition of LIPA. In the main study, physically inactive university students were allocated in a 1:1 ratio to either S-C or S-T (between-participant factor). Each participant completed two laboratory visits within their assigned modality: an experimental LIPA visit and a corresponding modality-matched passive control visit of equal duration (passive sitting for S-C and S-T), enabling within-person intervention-control comparisons. The primary endpoint is the change in bilateral dorsolateral prefrontal cortex (DLPFC) oxygenated hemoglobin (O2Hb) from baseline to immediately after the intervention during cognitive task performance (intervention vs control), measured with functional near-infrared spectroscopy (fNIRS). Secondary outcomes include O2Hb persistence at 20 minutes postintervention, cognitive performance (n-back test and Stroop Color and Word Test [SCWT]), subjective sleepiness, and back pain.

Results: The study was funded by the Sendai University International Collaborative Research Grant in April 2023, and the trial was registered with ClinicalTrials.gov on November 21, 2024. Participant recruitment and data collection were completed in December 2024. The substudy results showed that both 5-minute S-C and S-T bouts fell within the LIPA range of <3 metabolic equivalents (METs) and could be implemented as brief, standardized LIPA breaks in the main trial. For the main trial, data cleaning and preliminary analyses are ongoing, and the outcomes will be reported in a subsequent outcomes paper (anticipated in late 2026).

XSL•FO
RenderX

Conclusions: This protocol provides a prespecified framework to test whether brief 5-minute LIPA bouts may be a practical strategy to interrupt prolonged sitting and acutely improve prefrontal oxygenation and related neurocognitive and somatic outcomes in physically inactive university students.

Trial Registration: ClinicalTrials.gov NCT06703424; https://clinicaltrials.gov/study/NCT06703424

International Registered Report Identifier (IRRID): PRR1-10.2196/82510

(JMIR Res Protoc 2026;15:e82510) doi:10.2196/82510

KEYWORDS

physically inactive; sit-cycling; stand-twisting; cognitive function; daytime sleepiness; back pain

Introduction

Background

It is common for university students to remain sedentary for large portions of the day, spending over 9 hours per day sitting [1]. Prolonged sitting is associated with increased risk for adverse health outcomes in adulthood, including cardiovascular disease [2], various forms of cancer, and type 2 diabetes [3]. In this context, understanding sedentary behavior among university students in Singapore is particularly important, as lifestyle patterns in a highly urbanized environment may differ from those in other regions.

Population-based studies in Singapore indicate that 47.7% of adults spend at least 7 hours per day in sedentary activities. Notably, the highest prevalence (59.8%) is observed among individuals aged 18-34 years [4], a demographic that closely aligns with the typical age range of university students. This pattern raises concerns about the potential health implications of prolonged sitting in this age group.

High levels of sedentary behavior appear to persist despite a reported increase in leisure-time sports and exercise participation among Singaporean citizens and residents, as documented in the National Sports Participation Survey [5]. This apparent paradox suggests that even when individuals engage in exercise, they may still spend substantial portions of the day sedentary, potentially undermining the health benefits of physical activity.

Meta-analytic evidence further indicates that the health risks linked to prolonged sitting are largely independent of an individual's overall physical activity level, including among those who meet established physical activity guidelines [2,3]. Therefore, sedentary behavior should be addressed as a distinct risk factor for later ill health, rather than being considered solely in relation to insufficient physical activity.

The Movement Ecology Framework (MEF) conceptualizes human movement as a natural and essential component of health and well-being, emphasizing that even low-intensity, nonexercise movements (eg, fidgeting, stretching, or trunk-twisting) can contribute to metabolic and brain health [6]. Anchored in the MEF, this study will focus on light-intensity physical activity (LIPA) as a practical approach that can be integrated into university students' daily routines to counteract the negative effects of prolonged sitting on cognitive function and musculoskeletal health, potentially improving sleep-related outcomes and reducing back pain [7].

Interventional studies suggest that incorporating short physical activity breaks can effectively reduce excessive sitting among university students [8]. Such interventions may mitigate risks associated with prolonged sedentary behavior and contribute to improvements in physical health [9,10], mental health [11,12], and work productivity [13,14]. These findings underscore the importance of developing strategies that encourage regular movement throughout the day.

The association between sedentary behavior and cognitive function is increasingly recognized. Neuroimaging studies have shown that prolonged sitting correlates with reduced structural thickness in brain areas, such as the medial temporal lobe, which may impair cognitive abilities [15]. Systematic reviews also suggest that sedentary behavior is associated with lower cognitive function [16,17], although some studies report weaker associations [18]. Experimental evidence indicates that short physical activity breaks may offset some negative effects of prolonged sitting by enhancing cerebral blood flow [19], particularly to the frontal lobe, which supports complex cognitive processes [20-25]. In addition, sedentary behavior may adversely affect mood and physical health, contributing to fatigue, back pain, and hypertension [26-28].

Intervention studies in older adults suggest that interrupting sitting with brief walks can increase brain-derived neurotrophic factor and improve cognitive function [29]. However, similar strategies have not consistently demonstrated significant benefits in university students, potentially due to differences in exercise intensity requirements or the relatively higher baseline physical activity levels common in this population [30-32]. Although sit-cycling (S-C) has been used as a method to interrupt prolonged sitting [33], stand-twisting (S-T) has, to the best of our knowledge, not been used as a modality for breaking up prolonged sitting.

Research Intention

S-C and S-T are cost-effective and space-efficient LIPA modalities that can be implemented in typical university settings, making them practical strategies for increasing gross motor movement and interrupting prolonged sitting. The intervention duration in this study was set at 5 minutes to balance physiological plausibility with real-world feasibility. Brief LIPA breaks of this duration are implementable in university classrooms, libraries, and study spaces with minimal disruption to academic routines, reflecting realistic conditions under which sedentary-behavior interruptions could be adopted. Mechanistically, a short bout of light movement may elicit rapid changes in autonomic activation and cerebral

XSL•FO
RenderX

perfusion/oxygenation, which may be accompanied by short-term changes in executive performance, while minimizing fatigue or musculoskeletal discomfort that could influence subjective symptom ratings.

Study Purpose

A two-arm study using sequential (quasi-random) allocation was designed to examine the acute (immediate) and short-term delayed effects of a single 5-minute bout of either S-C or S-T on the prespecified outcomes, relative to a modality-matched 5-minute passive control (passive sitting for S-C and S-T). This is a mixed within-between laboratory study in which participants were assigned to one modality only (S-C or S-T; between-participant factor) and completed two visits within that modality—an intervention visit and a modality-matched passive control visit—constituting a two-period within-participant crossover for the condition with repeated assessments over time; intervention-control effects will be evaluated within person, whereas modality effects will be compared between participants (ie, no crossover between modalities). The study consisted of two parts: an oxygen uptake measurement phase (*substudy*) and an intervention phase (*main study*). In essence, the substudy focused on the measurement of oxygen uptake during a single 5-minute bout of S-C or S-T, while the main study examined the transient and delayed effects of a single 5-minute bout of S-C or S-T on oxygenated hemoglobin (O_2Hb) levels in the bilateral dorsolateral prefrontal cortex (DLPFC), cognitive function, blood pressure (BP), and self-reported daytime sleepiness and back pain compared to 5 minutes of passive sitting. These passive comparators were designed to match exposure time and posture but were not intended as sham interventions.

Standard Protocol Items: Recommendations for Interventional Trials

The study adhered to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [34]. This protocol describes the design and development of the intervention study using 5-minute bouts of S-C and S-T and involves an investigation into the oxygen uptake during 5 minutes each of S-C and S-T compared to the passive rest states of sitting. Findings from this study will add knowledge to the literature about the effectiveness of novel LIPA, such as S-C and S-T, in reducing sedentary behavior and its impact on performance tasks and somatic symptoms. Future research can adapt the described protocol for application in different desk-bound or movement-restricted occupational settings in larger-scale studies (eg, for deskbound administrative staff or classroom-based schoolteachers) to interrupt prolonged occupational sitting.

Study Objectives and Hypotheses

Substudy: Measurement of Oxygen Uptake in 5 Minutes of S-C and S-T and Intensity Validation

The objective of the substudy is to quantify oxygen uptake during a single 5-minute bout of S-C or S-T performed at a self-selected pace and to determine whether both modalities meet the operational definition of LIPA (<3 metabolic equivalents [METs]). We hypothesized that S-C against a negligible or minimal applied resistance and S-T at a self-determined cadence (about 2 seconds per full twist to the left and to the right) are LIPA-compliant and incur an energy output of <3 METs. Physical activities involving <3 METs are typically classified as LIPA. Data from an earlier study using the same S-C equipment showed that the intensity of S-C at a self-selected pace in older working adults is of light intensity [33]. Moreover, the use of waist-twisters at public outdoor parks are considered LIPA [35].

Main Study: Examining the Effect of Light-Intensity S-C and S-T

Prefrontal hemodynamic responses will be evaluated using O_2Hb, with the primary endpoint defined as the within-person intervention-control difference in the change in bilateral DLPFC O_2Hb levels (region-of-interest average) from baseline (t_0) to immediately after the intervention (posttest 1, t_1) during cognitive task performance. Key secondary outcomes include persistence of the O_2Hb response 20 minutes postintervention (posttest 2, t_2) and cognitive performance (n-back test and Stroop Color and Word Test [SCWT]; accuracy/response time, as applicable). Deoxygenated hemoglobin (HHb) will be analyzed as a supportive physiological indicator, and all remaining physiological and symptom-based measures have been prespecified as secondary/exploratory outcomes.

The objectives of the main study are as follows:

- To compare changes in prefrontal cortical hemodynamics (O2 Hb and HHb) between each intervention and its corresponding control across baseline (t0), posttest 1 (t1), and posttest 2 (t2)
- To measure changes in O2 Hb and HHb levels in the bilateral DLPFC at baseline, during, immediately after, and 20 minutes after 5 minutes each of S-C and S-T compared to 5 minutes of passive sitting
- To measure the cognitive test performance (n-back test and SCWT) and BP at baseline, immediately after, and 20 minutes after 5 minutes each of S-C and S-T compared to 5 minutes of passive sitting

The secondary objectives of the main study are as follows:

- To examine changes in cognitive performance (working memory and inhibitory control; n-back test and SCWT) from t0 to t1 and t2 under each intervention condition relative to its modality-matched control
- To examine changes in BP/heart rate (HR) from t0 to t1 and t2 relative to the modality-matched control
- To examine changes in self-reported daytime sleepiness and back pain from t0 to t2 under each intervention condition relative to its modality-matched control, as specified in the assessment schedule

Study Hypotheses

We hypothesize that compared with the modality-matched passive control, the 5-minute LIPA intervention will produce a greater t_0-to-t_1 change in DLPFC O_2Hb during cognitive task performance (primary endpoint). We further hypothesize that O_2Hb changes may persist at t_2 and that cognitive performance

XSL•FO
RenderX

(n-back test and SCWT accuracy/response time) may improve following the intervention (key secondary outcomes).

The hypotheses are as follows:

- Hypothesis (H)1 (hemodynamics): Compared with the modality-matched control (passive sitting), both S-C and S-T are associated with higher DLPFC O2 Hb (primary) and corresponding changes in HHb (supportive) during cognitive task performance at t1 and/or t2 relative to t0.
- H2 (cognition): Compared with the corresponding modality-matched control, both S-C and S-T will result in improved working memory and inhibitory control performance (eg, n-back test and SCWT outcomes) at t1 and/or t2 relative to t0.
- H3 (physiological and symptoms): Compared with the corresponding modality-matched control, both S-C and S-T will lead to favorable acute changes in BP/HR and reduced subjective daytime sleepiness and back pain at t2 relative to t0.

Methods

Ethical Considerations

The study was reviewed and approved by the Institutional Review Board (IRB) of Nanyang Technological University (approval number IRB-2023-939). All participants provided written informed consent prior to enrollment, and all procedures were conducted in accordance with the Declaration of Helsinki. Participation was voluntary, and participants could withdraw at any time without penalty. Data were collected and stored in a deidentified form on password-protected devices, with access restricted to the research team. Participants were compensated with cash payment (US $30 for full compliance) for their participation.

Study Design

The trial was registered on November 21, 2024 (ClinicalTrials.gov identifier: NCT06703424), and participant recruitment commenced only after registration. This manuscript is presented as a protocol with trial status, describing the prespecified protocol together with the completed trial timeline/status.

Details of this protocol are reported in accordance with SPIRIT guidelines [34]. SPIRIT provides a comprehensive checklist to ensure clarity and completeness in trial protocol reporting, and the SPIRIT items addressed in this protocol are summarized in Table 1. Outcomes were assessed at baseline (t_0), posttest 1 (t_1), and posttest 2 (t_2) to characterize both acute and short-term delayed effects of a brief LIPA bout. The immediate postintervention window was intended to capture rapid physiological and neurocognitive responses (eg, transient changes in prefrontal oxygenation and executive performance), whereas the 20-minute follow-up evaluated whether any effects persisted beyond the immediate period or attenuated toward baseline. The 20-minute interval was selected as a pragmatic postbreak window that fits within a typical study or class session and aligns with commonly used observation periods for short-lived responses after brief activity.

XSL•FO
RenderX

Table 1. Timeline of enrollment, interventions, and assessments in the main study.

SPIRIT[a] items	Enrollment	Pre-intervention	During the intervention	Postintervention	
	$-t_1$	t_0[b]	t_{during}[c]	t_1[d]	t_2[e]
Enrollment					
Eligibility screen	✓	—[f]	—	—	—
Informed consent	✓	—	—	—	—
Allocation (S-C[g] or S-T[h])	✓	—	—	—	—
Familiarization	—	✓	—	—	—
Interventions					
S-C or S-T	—	—	✓	—	—
Waitlist control	—	—	✓	—	—
Assessments					
fNIRS[i]	—	✓	✓	✓	✓
n-Back test	—	✓	—	✓	✓
SCWT[j]	—	✓	—	✓	✓
BP[k]	—	✓	—	✓	✓
Self-rating sleepiness	—	✓	—	—	✓
Self-rating back pain	—	✓	—	—	✓

[a]SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials.

[b]t_0: baseline measurements (pretest).

[c]t_{during}: time during the intervention (the intervention is 5 minutes long).

[d]t_1: immediately after the exercise.

[e]t_2: 20 minutes after t_1.

[f]Not applicable.

[g]S-C: sit-cycling.

[h]S-T: stand-twisting.

[i]fNIRS: functional near-infrared spectroscopy.

[j]SCWT: Stroop Color and Word Test.

[k]BP: blood pressure.

Study Setting and Participants

The trial was conducted at a university in Singapore. This manuscript is intended as a study protocol with trial status information, accompanied by supportive methodological evidence from a MET validation substudy. Prospective participants were students from the same university, aged between 21 and 30 years, who do not meet physical activity guidelines [36]. An a priori target sample size of 28 participants (n=14, 50%, per modality group) was estimated using G*Power 3.1 for a repeated-measures analysis with within-between factors (α=.05; 1 − β=0.80; Cohen f=0.25). Owing to recruitment constraints, 22 (78.6%) participants were ultimately enrolled (n=11, 50%, per modality group). This shortfall is acknowledged in the *Discussion* section as a limitation, with potential implications for statistical power, particularly for small-to-moderate effects. A suggested sample of 6 (27.3%) participants who met the selection criteria was recruited based on a G*power test (α=.05; power: S-C, 99%; S-T, 86%) for the substudy.

Recruitment

For recruitment, study advertisements were displayed within the university campus and posted in student messaging groups on online messaging channels, such as WhatsApp and Telegram. Interested students contacted the research team, and a research team member explained the study's purpose, the requirements for participation, and potential risks and benefits and addressed questions in person. Afterward, prospective participants were screened for eligibility. Those who met the inclusion criteria and were willing to participate signed an informed consent form that confirmed participation in the study.

The inclusion criteria were as follows:

- Age between 21 and 30 years
- Right-handed

XSL•FO
RenderX

- Physically inactive: less than 150 minutes of moderate-intensity physical activity (MPA) or less than 75 minutes of vigorous-intensity physical activity (VPA) per week in the past month [36]
- Cleared for LIPA based on the Get Active Questionnaire (GAQ) [37] (no positive responses on the GAQ indicating the need for medical clearance or medical clearance obtained, if required) and willing to undergo fNIRS
- Not taking any antihypertensive medication at present

The criterion of not meeting physical activity guidelines (<150 minutes of MPA or <75 minutes of VPA per week in the past month) was appraised by the researcher during recruitment based on each prospective participant's response to the following questions [36]:

- How many days in a week do you engage in MPA or VPA?
- How many minutes on average per day do you engage in physical activity at this level?
- How many days per week do you engage in muscle-building or strength related exercise?

Examples and characteristics of MPA and VPA were shared in helping participants not familiar with MPA or VPA.

Reasons for exclusion included not meeting the physical inactivity criterion, not meeting safety criteria on the GAQ, or declining participation.

Procedural Flow

A screening log was maintained to document responses to advertisements, eligibility assessment outcomes, reasons for exclusion, and final enrollment. A total of 22 eligible participants provided written informed consent and were enrolled. Following allocation, each participant completed two laboratory visits within the assigned modality (an experimental LIPA visit and a modality-matched passive control visit), as shown in Figure 1. At each visit, outcomes were assessed at baseline (t_0), during the 5-minute activity period (t_{during}; intervention or control), posttest 1 (t_1), and posttest 2 (t_2). The number of participants completing each visit and each assessment (t_0, t_{during}, t_1, and t_2), as well as the number in final analytic sample, are presented in Figure 1.

Figure 1. Revised study design and procedural flow. BP: blood pressure; fNIRS: functional near-infrared spectroscopy; GAQ: Get Active Questionnaire; HR: heart rate; RA: rating of alertness.

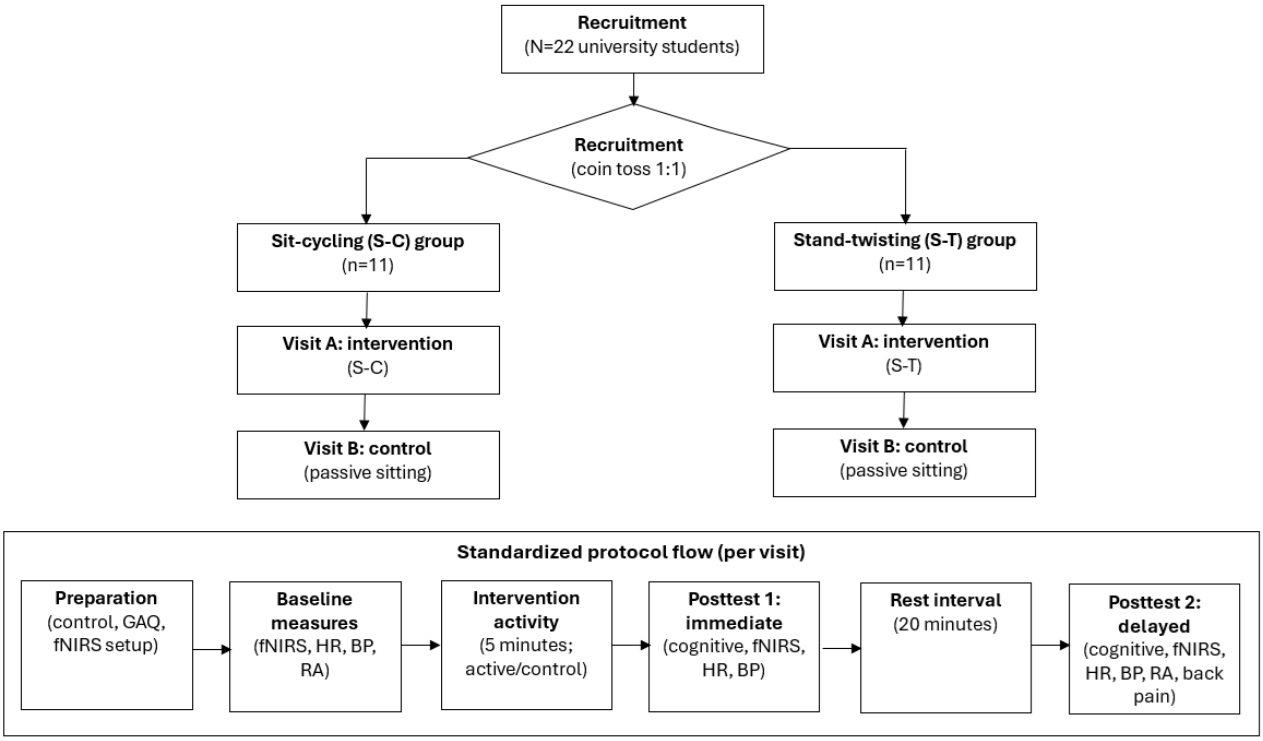

Data Collection Procedures: Assessment Time Points and Visit Structure

Within their assigned modality, each participant completed two laboratory visits (intervention and modality-matched passive control), forming a within-participant crossover for condition. As mentioned before, at each visit, outcomes were assessed at baseline (t_0) after a 3-minute rest period, during the 5-minute intervention (t_{during}), posttest 1 (t_1), and posttest 2 (t_2) 20 minutes after t_1. Unless otherwise specified, all time points (t_0, t_{during}, t_1, t_2) refer to assessments during a visit. Subjective outcomes (sleepiness and back pain) were assessed only at t_0 and t_2 to minimize interruption during cognitive testing and to capture baseline and delayed symptom responses; they were not assessed at t_1. Baseline measures (t_0) included cognitive testing (n-back test and SCWT) and physiological measures (functional near-infrared spectroscopy [fNIRS], HR, BP, and rating of alertness [RA]). The intervention or control condition was delivered for 5 minutes, during which monitoring measures (eg, rating of perceived exertion [RPE] checks) were recorded (t_{during}). The equipment included an S-C apparatus (AIBI EZ Tone Chair) and an S-T apparatus (AIBI Massage Body Twister)

XSL•FO
RenderX

[38,39]. Immediately after the 5-minute bout, participants completed the immediate postintervention assessments (t_1), including cognitive outcomes, fNIRS, HR, and BP measurements. This was followed by delayed assessments at t_2 after a rest interval, which included cognitive outcomes, fNIRS, HR, BP, RA, and back pain. To minimize procedural variability, all sessions were conducted in the same laboratory setting using a standardized script and identical equipment setup, and outcome assessors followed a prespecified checklist for each time point.

Protocol Design for the Substudy

A total of 6 (27.3%) participants for the substudy were drawn from the participant pool (N=22) for the main study. The 6 participants visited the exercise physiology laboratory on a designated day. Upon their arrival, a research team member explained the process to them. The participants were then allocated to either the S-C group or the S-T group, after which they spent approximately 5 minutes familiarizing themselves with either the S-C or the S-T equipment. During the indirect calorimetry measurement, a mouthpiece connected to a metabolic cart (Parvo Medics True One 2400) was attached to the participant's head. The participant breathed through the mouthpiece while remaining seated for 5 minutes. The metabolic cart collected breath-by-breath expired air and recorded the rate of oxygen consumption (VO_2) and METs. After baseline measurements were recorded, the participants rested for about 5 minutes. Subsequently, they underwent an indirect calorimetry session while doing S-C or S-T at a self-selected pace with minimal or negligible applied resistance for 5 minutes each. The total duration of the session averaged approximately 40 minutes. Figure 2 shows a schematic illustration of the substudy.

Figure 2. Schematic representation of the substudy. MET: metabolic equivalent; VO2: oxygen consumption.

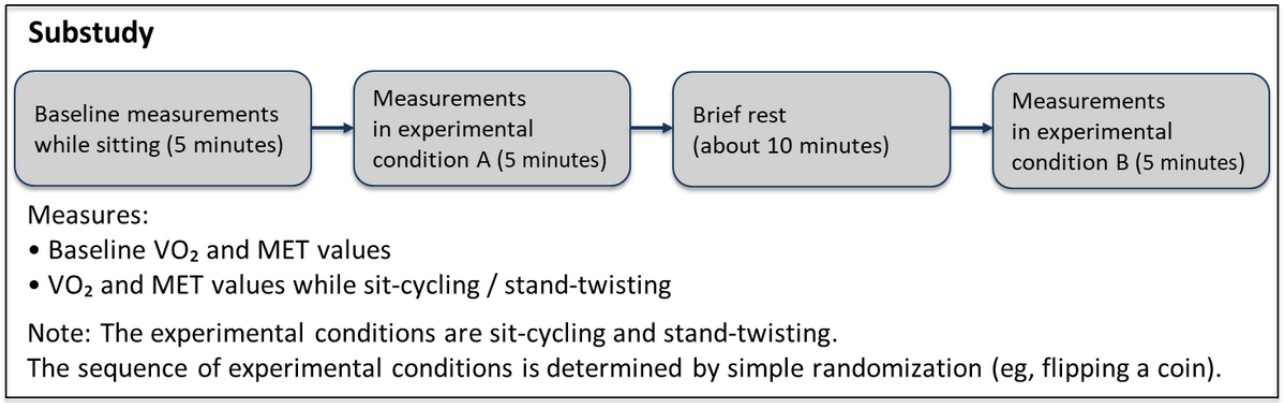

Protocol Design for the Main Study

Participants were allocated by the researcher to either the S-C or the S-T group in a 1:1 ratio using a coin toss for the first assignment (random start), followed by a prespecified alternating allocation sequence based on enrollment order. Within each assigned modality, outcomes were assessed repeatedly across prespecified time points (t_0, t_1, t_2) and across two occasions (the experimental LIPA condition and the corresponding modality-matched passive control condition). As allocation concealment was not implemented and subsequent assignments were predictable, this approach is described as sequential (quasi-random) allocation rather than full randomization. Figure 3 shows the timeline of enrollment, intervention, and assessment phases of the main study.

XSL•FO
RenderX

Figure 3. Schematic representation of the main study. BP: blood pressure; fNIRS: functional near-infrared spectroscopy; HR: heart rate.

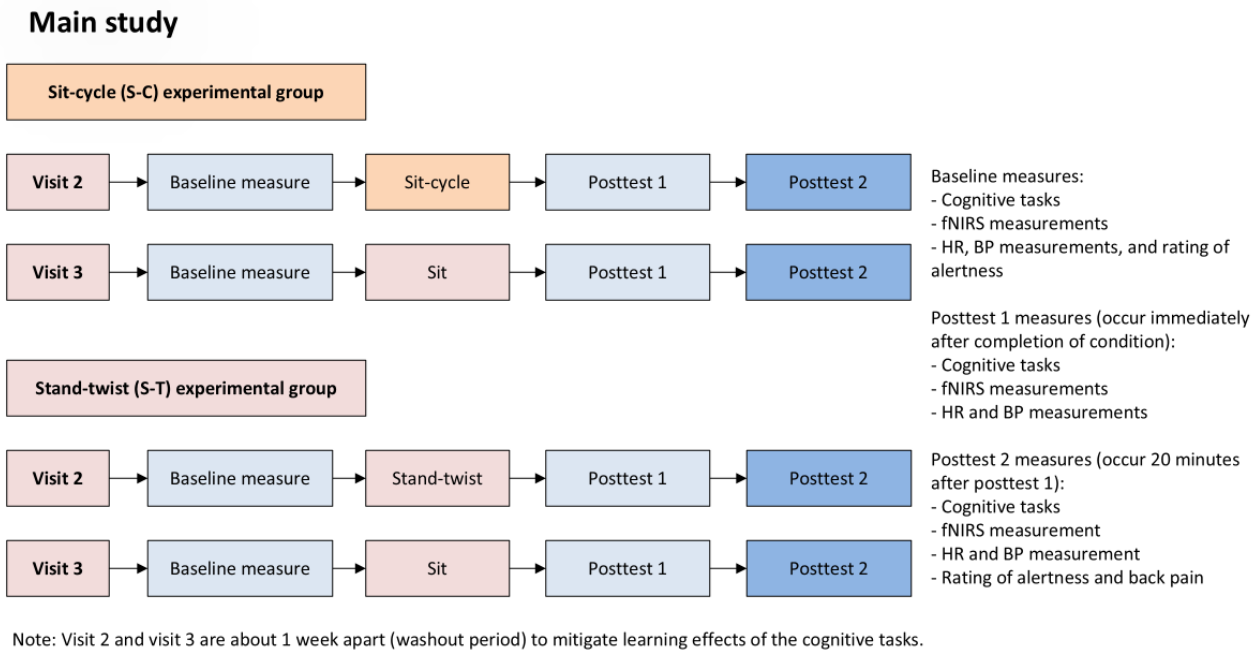

Note: Visit 2 and visit 3 are about 1 week apart (washout period) to mitigate learning effects of the cognitive tasks.

Intervention

S-C Experimental Condition

Participants assigned to the S-C intervention group cycled on the EZ Tone Chair at a self-selected cadence against a minimal applied resistance for 5 minutes. The chair measures 57 × 39 × 76 cm^3 in dimension and is equipped with toe cap foot pedals. Previous research in the same institution used it as an exercise tool for reducing sedentary behavior in desk-bound participants [33]. In addition, participants were asked to rate their exertion levels on the 1-10 Borg RPE scale at every 1-minute interval.

S-T Experimental Condition

Participants assigned to the S-T intervention group stood on the AIBI Massage Body Twister and twisted their torso along the transverse plane in an alternating clockwise and counterclockwise direction with even arm swings for balance and at a comfortable cadence (about 2 seconds for a full left-right trunk twist) for 5 minutes. The twister is a compact piece of equipment with a diameter of 35 cm and is portable, making it convenient for use in any setting. It has been used as an intervention method for reducing sitting time and investigating the weight loss effects in participants with obesity [40]. In addition, participants were asked to rate their exertion levels on the 1-10 Borg RPE scale at every 1-minute interval.

Intervention Delivery and Fidelity

Interventions were delivered under direct supervision of a trained research team member to ensure standardization and participant safety. Both S-C and S-T groups received the same brief standardized instructions regarding posture, cadence, and safety before they started the 5-minute bout. Perceived exertion was recorded at 1-minute intervals using the Borg RPE scale to monitor intensity and support fidelity to LIPA.

Participants in the control groups spent a modality-matched passive control period with no structured physical activity (passive sitting on a chair or standing still) for the same 5-minute duration. Outcome assessments were conducted at identical time points as in the interventions. The modality-matched passive control conditions were selected to match the session duration and posture as closely as possible; however, they were not designed as sham interventions and therefore may not have adequately controlled for posture-related arousal, expectancy effects, or other context-related cues (eg, awareness of movement and perceived exertion).

Any protocol deviations and adverse events were documented.

Key Outcome Measures

Primary and Secondary Outcomes

The primary endpoint of the main trial is the within-person intervention-control difference in the change in O_2Hb in the bilateral DLPFC during cognitive task performance from baseline (t_0) to posttest 1 (t_1). The DLPFC O_2Hb metric is defined as the bilateral DLPFC region-of-interest average, calculated as the baseline-corrected mean O_2Hb during cognitive task performance at each time point.

Key secondary outcomes include (1) persistence of the O_2Hb response posttest 2 (t_2), (2) cognitive performance on the n-back test and the SCWT (accuracy and response time, as applicable), (3) HHb in the same DLPFC region (supportive physiological indicator), (4) BP and HR, and (5) symptom-based outcomes (eg, subjective daytime sleepiness and back pain measures).

Functional Near-Infrared Spectroscopy

fNIRS is based on the principle that most biological tissues are transparent to near-infrared light (NIR), while O_2Hb and HHb better absorb NIR light in the 700-900 nm spectrum. fNIRS uses the modified Beer-Lambert law to measure the absorption

XSL•FO
RenderX

rate of these hemoglobin species at two different NIR wavelengths between transmitter and receiver pairs placed on the scalp. In this study, an eight-channel portable continuous-wave NIRS system (Octamon, Artinis Medical Systems) was used to measure fNIRS signals during the n-back test and the SCWT. Hemodynamic responses were recorded from the bilateral DLPFC using eight transmitters and two receivers (one receiver receiving from four transmitters on each hemisphere; see Figure 4) over the forehead region, with each receiver measuring time-multiplexed NIR light from four surrounding transmitters that were 3 cm apart, resulting in eight data channels. All signals were sampled at 10 Hz. A MATLAB-based toolbox was used to accurately estimate the spatial brain mapping of the electroencephalography (EEG) 10-20 system onto associated Brodmann areas, identifying the left and the right DLPFC as the regions of interest. Following data collection, the fNIRS data will be processed using the HOMER3 MATLAB toolbox with the following pipeline: (1) conversion to optical density changes, (2) removal of low signal-to-noise channels, (3) motion artifact detection and correction, (3) bandpass filtering, and (4) conversion to concentration changes using the modified Beer-Lambert law, with block averaging to obtain the average change across three trials for each cognitive task condition.

Figure 4. Illustration of the optode montage used indicating brain measurements using fNIRS (blue color: transmitters; yellow color: receivers). fNIRS: functional near-infrared spectroscopy.

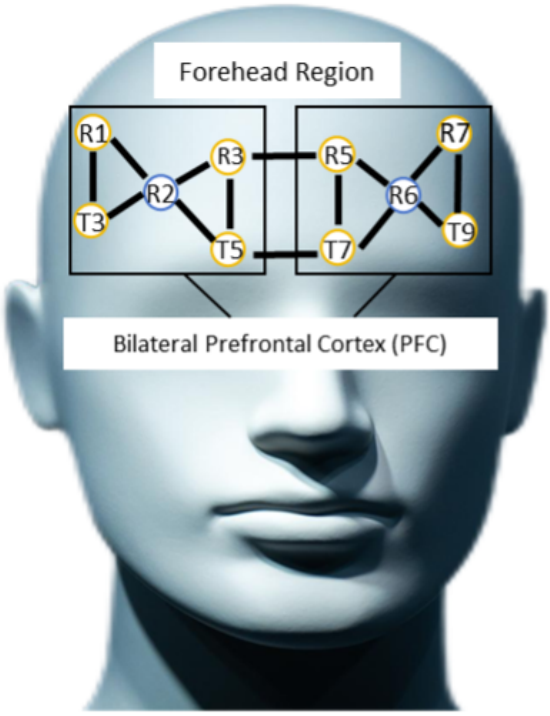

n-Back Test

The n-back test is a cognitive assessment tool used to evaluate working memory capacity and cognitive inhibition [41]. It has also been widely used as a training tool for enhancing working memory [42,43]. During the n-back test, participants were shown a sequence of visual stimuli in the form of digits from 0 to 9 and were asked to indicate whether the current stimulus shown matched the one presented n characters earlier by pressing the designated letters on the keyboard (eg, letter "F" to indicate a match and "J" to indicate a mismatch). For this study, the 1-back and 2-back tests were adopted. In the 1-back test, participants were required to recall and indicate whether the current digit shown matched the previous digit, whereas in the 2-back test, participants indicated whether the current digit shown matched the second-last digit presented earlier. The n-back test consists of three blocks of 1- and 2-back tasks. A practice block was added to the baseline pretest, where participants were given 20 stimuli (letters) to familiarize themselves with the process before moving on to the actual test. The actual test consisted of three blocks with 30 stimuli (letters). The test began with a centered fixation cross on-screen for 500 ms, followed by the stimulus in that location for 1000 ms, followed by a 500 ms interstimulus interval. The *accuracy* (number of correct recalls) and *response time* (speed of recall) were assessed at the end of each block. The study design entailed participants to undergo the n-back test thrice (baseline, posttest 1, and posttest 2) on two occasions (experimental condition and waitlist control condition). Changes in the n-back test score from baseline to posttests 1 and test 2 were measured.

Stroop Color and Word Test

The SCWT is a cognitive test widely used to assess the ability of an individual to inhibit cognitive interference that occurs when the processing of a specific stimulus feature is impeded by the simultaneous processing of another attribute of the same stimulus [44]. In this study, the test consisted of a series of color words (eg, "red," "blue," "green," and "yellow") printed in colored ink, with the font color sometimes matching the word and sometimes not. We adopted the incongruent condition, where participants were shown a sequence of stimuli in which the word text matched or did not match the color of the ink (eg,

XSL•FO
RenderX

the word "red" could be printed in blue ink). Participants were required to indicate whether the word text matched the color of the ink by pressing the designated letter on the keyboard (eg, letter "F" to indicate a match and "J" to indicated a mismatch). The color of the words was displayed on the screen one at a time for 2000 ms. Participants were presented with six blocks of words, alternating between congruent and incongruent words. The test typically consists of multiple blocks (three or four) and records the *accuracy* of each response (number of correct texts and colors) and the *reaction time* (time taken to input a response). A practice block was added to the baseline pretest for familiarization like in the n-back task. Participants underwent the test thrice at different time points (ie, baseline, posttest 1, and posttest 2) on two occasions (experimental condition and waitlist control condition), and changes in the SCWT scores from baseline to posttests 1 and test 2 were determined.

BP and HR Measurement

Participants' BP and HR were measured using the Vivo Watch BP HC-A04 (ASUSTek Computer Inc) worn on the wrist. With a pair of electrocardiography (ECG) and photoplethysmography sensors embedded in the watch, the Vivo Watch BP HC-A04 provides systolic and diastolic BP and HR readings. Previous studies have evaluated the validity of smartwatch-based BP-monitoring features, and validation evidence has also been reported for an ASUS VivoWatch series device against reference sphygmomanometer measurements [45]. Accordingly, BP readings in this study will be interpreted primarily for research monitoring (ie, within-participant comparisons across time points) rather than for clinical diagnosis. Measurements of BP and HR were taken at baseline, posttest 1, and posttest 2, and changes in the BP and HR readings from baseline to posttests 1 and 2 were determined.

Stanford Sleepiness Scale

The Stanford Sleepiness Scale (SSS) is a self-rating tool designed to evaluate changes in sleepiness on a 7-point scale, applicable at any time of day [44,46]. In this study, participants were asked to rate their level of sleepiness upon visiting the test site at baseline and posttest 2 by choosing a number and description that best described their state. The scale ranges from "feeling active, vital, alert, or wide awake" to "no longer fighting sleep, sleep onset soon, or having dream-like thoughts." The SSS correlates well with task performance related to alertness [47], and it has been previously used to gather subjective sleepiness data among workers [48].

Oswestry Low Back Disability Questionnaire

The Oswestry Low Back Disability Questionnaire (OLBDQ) is a widely used tool for assessing functional abilities relating to low back pain. The self-reported questionnaire has been validated and its reliability tested, and it is used in different populations [49]. In this study, participants completing the OLBDQ were asked about how low back pain affects their ability to manage their everyday life. The different contexts in everyday life included in the questionnaire are (1) pain intensity, (2) personal care (washing, dressing), (3) lifting objects, (4) walking, (5) sitting, (6) standing, (7) sleeping, (8) social life, and (9) traveling. For example, regarding pain intensity, participants were asked to select the statement that best described their problem: "I have no pain at the moment," "The pain is very mild at the moment," "The pain is moderate at the moment," The pain is fairly severe at the moment," "The pain is very severe at the moment," or "The pain is the worst imaginable at the moment." For each section, the total possible score is 5, and each statement is marked with a score; for example, the first statement is marked with a section score of 0, the second statement is marked with a section score of 1, and so on until the sixth statement is marked with a section score of 5. There are nine application sections, yielding a total possible score of 45. A higher score is interpreted as more severe low back functional disability. In this study, the OLBDQ was administered at t_0 and t_2 during each visit to capture baseline and delayed responses; it was not administered at t_1.

Blinding

Owing to the behavioral nature of the intervention conditions, blinding of participants and intervention personnel was not feasible. To minimize performance and detection bias, all sessions follow a standardized protocol with scripted instructions, and objective outcomes were obtained using instrument-based or computer-administered procedures (eg, fNIRS-derived hemodynamic responses and computerized cognitive tasks). Data processing and analyses will follow prespecified procedures.

Statistical Analysis

Descriptive statistics will be reported as the mean (SD) for approximately normally distributed continuous variables and as the median (IQR) for skewed variables, while categorical variables will be summarized as counts and percentages. The primary endpoint is the within-person intervention-control difference in the change in bilateral DLPFC O_2Hb levels during cognitive task performance from baseline (t_0) to posttest 1 (t_1), while HHb will be analyzed as a supportive indicator. The DLPFC O_2Hb metric will be the bilateral DLPFC region-of-interest average (baseline-corrected mean O_2Hb during the cognitive task at each time point). The primary analysis will test the within-person intervention-control difference in the change in DLPFC O_2Hb from t_0 to t_1 (primary endpoint) using a prespecified linear mixed-effects model, with participant as a random effect and condition (intervention vs control), time, and assigned modality (S-C vs S-T) as fixed effects, including condition × time (and, where relevant, condition × time × modality) interaction terms. The primary endpoint will be evaluated at a two-sided α=.05. For key secondary outcomes (O_2Hb persistence at t_2 and cognitive performance measures), we will apply Holm-Bonferroni adjustment across the prespecified key secondary tests (O_2Hb persistence at t_2 and the prespecified cognitive performance endpoints [accuracy and response time, as applicable]); all other outcomes will be interpreted as secondary/exploratory. The visit order/session will be included as a covariate to address potential practice or order effects. The primary contrast of interest will test the condition difference in the change from t_0 to t_1 (ie, condition × time contrast corresponding to $t_0 \rightarrow t_1$). Prespecified secondary contrasts will examine persistence at t_2 (eg, $t_0 \rightarrow t_2$). Model

XSL•FO
RenderX

assumptions will be assessed using standard diagnostics (eg, residual distributions and influence). Because the large multimodal model (LMM) is used for primary analyses, no sphericity assumption is required. Missing data will be handled using maximum likelihood estimation in the LMM under a missing-at-random assumption; the extent and patterns of missingness will be described. For secondary outcomes (cognitive test performance, BP, subjective sleepiness, and back pain), analogous LMMs will be fitted using the same modeling framework (participant as a random effect; condition, time, and modality as fixed effects, as appropriate). To address multiplicity across key secondary outcomes, we will control the family-wise error rate using a Holm-Bonferroni procedure, while maintaining a two-sided α of .05 for the primary endpoint. Where distributional concerns remain, nonparametric analyses (eg, Friedman test, Wilcoxon signed-rank test, Mann-Whitney U test, as appropriate) will be used as sensitivity analyses to assess robustness. Correlation analyses will use Pearson's r for approximately normally distributed variables and Spearman's ρ otherwise. Comparisons of outcomes by sex will be treated as exploratory and not powered, given the sample size (N=22), and will be reported with effect sizes and 95% Cis, where applicable. All tests will be two-sided. The primary endpoint will be evaluated at α=.05; key secondary outcomes will be interpreted using Holm-Bonferroni–adjusted inference, and all other analyses will be considered exploratory.

Results

Trial Status

The study protocol was approved by the IRB of Nanyang Technological University (approval number: IRB-2023-939) and was funded by the Sendai University International Collaborative Research Grant (grant number: CER-2023) in April 2023. Recruitment and data collection were completed in December 2024. A total of 22 physically inactive university students were enrolled and allocated to either the S-C or the S-T arm, with outcomes compared against the corresponding modality-matched passive control condition (passive sitting for S-C and S-T). Data cleaning and preliminary analyses are ongoing, and the complete outcomes will be reported in a subsequent results paper (anticipated in late 2026).

Methodological Substudy Results: MET Validation for Intervention Intensity

A methodological substudy was conducted to confirm that the two 5-minute activity modalities (S-C and S-T) meet the operational definition of LIPA (<3 METs), thereby justifying their classification for the main trial. The mean MET value was 0.98 during 5 minutes of quiet seated rest during the modality-matched passive control stage and 2.71 during S-C (paired *t* test, *P*<.05), indicating higher energy expenditure during S-C than during seated rest. Similarly, the mean MET value was 1.19 during 5 minutes of quiet standing and 2.26 during S-T (paired *t* test, *P*<.05), indicating higher energy expenditure during S-T than during quiet standing. As a safety check, paired-sample *t* tests showed no significant pre-post differences in BP or HR following either 5-minute bout. Collectively, these findings indicate that both 5-minute S-C and S-T bouts fell within the LIPA range (<3 METs) and could be implemented as brief, standardized LIPA breaks in the main trial. No analyses of the main trial outcomes (fNIRS-derived prefrontal cortex hemodynamics, cognitive performance, BP, subjective sleepiness, and back pain) are reported in this protocol paper; these results will be presented in a separate outcomes manuscript once analyses are finalized.

Discussion

Summary

This manuscript primarily serves as a study protocol with trial status information, detailing the rationale, design, procedures, and planned analyses for examining the acute effects of two 5-minute LIPA breaks (S-C and S-T) on neurocognitive and somatic outcomes among physically inactive university students. In addition, we reported findings from a methodological substudy undertaken to support the main trial by confirming that both S-C and S-T can be operationalized as LIPA (<3 METs). Taken together, this protocol provides a transparent framework for replication and for interpreting subsequent outcomes reporting from the main trial.

The primary aim of this protocol was to establish that a 5-minute bout of S-C or S-T constitutes LIPA. Consistent with our hypothesis, 5 minutes of both S-C and S-T were deemed as LIPA. These findings provide methodological support for the main study, which is designed to investigate the acute effects of LIPA on cognitive function and brain activity. Establishing the intensity of these modalities is an essential methodological step, as the primary trial aims to evaluate whether a brief LIPA bout can elicit measurable changes in prefrontal cortex hemodynamics and cognitive performance, alongside changes in subjective sleepiness and back pain. If effective, such short activity breaks may represent a pragmatic strategy to interrupt prolonged sitting in student populations who are at elevated risk of sustained sedentary exposure.

From a broader perspective, the importance of interrupting prolonged sitting, even among physically active individuals, has been emphasized [50]. In parallel, interventions using digital health technologies to reduce sedentary behavior in workplaces and to promote behavior change are gaining popularity [51]. Although much of the existing evidence on these digital interventions focuses on general health outcomes, evidence relating to brain activity and cognitive function remains limited. Research investigating how short physical activity breaks influence cerebral blood oxygenation and cognitive performance in adults is still in its infancy, and key parameters, such as the optimal exercise modality, bout duration, and interruption frequency, that are feasible in real-world settings are yet to be established. In this context, by integrating objective indices (eg, prefrontal cortex oxygenation) with behavioral and symptom-based outcomes (eg, working memory, inhibitory control, sleepiness, and back pain), this trial is expected to contribute to a more integrative understanding of how short, low-intensity movement may influence both neurocognitive and musculoskeletal-related indicators in young adults. Moreover, this protocol provides a practical framework to inform future studies seeking to elucidate whether brief S-C

XSL•FO
RenderX

and S-T breaks, compared with prolonged sitting, can meaningfully reduce sedentary exposure and support cognitive functioning and overall well-being.

This protocol is underpinned by the MEF, which provides a robust foundation for understanding the multifaceted role of movement in health promotion. By valuing all movement behaviors, including LIPA, the MEF encourages a holistic approach to health that may be particularly relevant for university students, who are vulnerable to sedentary lifestyles [6,52].

Strengths and Limitations

The strengths of this protocol include the pragmatic nature of the intervention (brief, low-cost, and implementable in common study settings), clearly specified eligibility criteria, standardized delivery of activity bouts, and prespecified outcome assessment time points. These features are intended to enhance transparency and reproducibility and to facilitate translation into real-world university contexts, such as classrooms, libraries, and other desk-bound study spaces.

A further strength of this protocol is the assessment schedule incorporating immediate and short-term follow-up. By evaluating outcomes at t_0, t_1, and t_2, the study can differentiate transient responses from effects that persist for a meaningful period within a typical academic session, thereby improving interpretability and translational relevance.

Several limitations should be acknowledged when interpreting this protocol and the subsequent outcomes reporting. First, the achieved sample size (N=22) was smaller than the a priori target (N=28), which may reduce statistical power, particularly for detecting small-to-moderate effects. Accordingly, findings should be interpreted cautiously, with emphasis on effect sizes and Cis, in addition to statistical significance. Second, this is a nonrandomized study in which modality assignment followed a sequential (quasi-random) alternating schedule and allocation concealment was not implemented; therefore, a risk of selection bias cannot be ruled out despite consecutive enrollment and the use of a prespecified allocation schedule. In addition, because the protocol involves brief behavioral tasks assessed repeatedly over time, performance-related influences (eg, differential engagement, learning, or fatigue across repeated assessments) may affect outcomes, particularly for subjective measures, such as sleepiness and back pain. Third, blinding was not feasible for participants or intervention personnel due to the nature of the activity conditions, which may increase susceptibility to expectancy and performance effects, particularly for subjective outcomes, such as sleepiness and back pain. Although objective outcomes (eg, prefrontal cortex hemodynamics and computerized cognitive task performance) are relatively less susceptible to reporting bias, nonblinding may still influence participant engagement and effort during assessments. Fourth, although the analytical plan specifies a primary endpoint and prespecified multiplicity control for secondary outcomes, the assessment of multiple outcomes still warrants cautious interpretation to mitigate the risk of inflated type I errors. Finally, as the study was conducted within a single university context, generalizability may be limited.

Future Directions

Future research should build on this protocol by implementing a fully randomized allocation procedure with allocation concealment; incorporating a sham or active control condition that better matches posture and expectancy; and recruiting larger samples, preferably across multiple sites, to improve precision and generalizability. In addition, preregistering a single primary endpoint (or a clearly ranked hierarchy of outcomes), applying a priori multiplicity control strategies, and considering linear mixed modeling approaches for repeated measures would further strengthen inferential robustness. Longer-term and ecologically embedded evaluations (eg, repeated breaks across a study day or semester) may also clarify whether acute responses translate into sustained benefits in cognitive performance and musculoskeletal symptoms.

Acknowledgments

Financial and material support were funded by Sendai University International Collaborative Research Grant (grant number: CER-2023) for this collaborative research between Singapore (National Institute of Education) and Japan (Sendai University). The funding source played no role in the study design; data collection, analysis, and interpretation; writing of the manuscript; or the decision to submit the paper for publication.

Data Availability

Research data collected in this study will be deposited to the data repository of the authors' university, as required by the university.

Authors' Contributions

MCYH is the principal investigator of the study. The study was collectively conceived and designed by MCYH, TWP, TCBK, JC, MJ, KHS, LST, and LD. The first full draft of the manuscript was drafted by MJ and TCBK. MCYH and coauthors provided a critical review and revised the drafts. Financial and material support (honorariums to participants and ASUS Vivowatch BP) were secured by MJ. All authors reviewed, contributed, and approved the final draft.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Castro O, Bennie J, Vergeer I, Bosselut G, Biddle SJH. How sedentary are university students? A systematic review and meta-analysis. Prev Sci 2020 Apr 23;21(3):332-343 [FREE Full text] [doi: 10.1007/s11121-020-01093-8] [Medline: 31975312]
2. Ekelund U, Brown WJ, Steene-Johannessen J, Fagerland MW, Owen N, Powell KE, et al. Do the associations of sedentary behaviour with cardiovascular disease mortality and cancer mortality differ by physical activity level? A systematic review and harmonised meta-analysis of data from 850 060 participants. Br J Sports Med 2019 Jul 10;53(14):886-894 [FREE Full text] [doi: 10.1136/bjsports-2017-098963] [Medline: 29991570]
3. Biswas A, Oh PI, Faulkner GE, Bajaj RR, Silver MA, Mitchell MS, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. Ann Intern Med 2015 Jan 20;162(2):123-132 [FREE Full text] [doi: 10.7326/M14-1651] [Medline: 25599350]
4. Lau JH, Nair A, Abdin E, Kumarasan R, Wang P, Devi F, et al. Prevalence and patterns of physical activity, sedentary behaviour, and their association with health-related quality of life within a multi-ethnic Asian population. BMC Public Health 2021 Oct 25;21(1):1939 [FREE Full text] [doi: 10.1186/s12889-021-11902-6] [Medline: 34696751]
5. Sport participation in Singapore hits all-time high; walking, jogging and calisthenics top 3 activities. The Straits Times. 2023 Jun 27. URL: https://www.straitstimes.com/sport/sport-participation-at-an-all-time-high-in-singapore-gradual-increase-since-pandemic [accessed 2025-08-12]
6. Chaput J, Carson V, Gray C, Tremblay M. Importance of all movement behaviors in a 24 hour period for overall health. Int J Environ Res Public Health 2014 Dec 04;11(12):12575-12581 [FREE Full text] [doi: 10.3390/ijerph111212575] [Medline: 25485978]
7. McGregor DE, Carson V, Palarea-Albaladejo J, Dall PM, Tremblay MS, Chastin SFM. Compositional analysis of the associations between 24-h movement behaviours and health indicators among adults and older adults from the Canadian Health Measure Survey. Int J Environ Res Public Health 2018 Aug 18;15(8):1779 [FREE Full text] [doi: 10.3390/ijerph15081779] [Medline: 30126215]
8. Salmon J. Novel strategies to promote children's physical activities and reduce sedentary behavior. J Phys Act Health 2010 Nov;7(S3):S299-S306 [FREE Full text] [doi: 10.1123/jpah.7.s3.s299] [Medline: 21116014]
9. Loh R, Stamatakis E, Folkerts D, Allgrove JE, Moir HJ. Effects of interrupting prolonged sitting with physical activity breaks on blood glucose, insulin and triacylglycerol measures: a systematic review and meta-analysis. Sports Med 2020 Feb 24;50(2):295-330 [FREE Full text] [doi: 10.1007/s40279-019-01183-w] [Medline: 31552570]
10. Kerr J, Crist K, Vital DG, Dillon L, Aden SA, Trivedi M, et al. Acute glucoregulatory and vascular outcomes of three strategies for interrupting prolonged sitting time in postmenopausal women: a pilot, laboratory-based, randomized, controlled, 4-condition, 4-period crossover trial. PLoS One 2017 Nov 30;12(11):e0188544 [FREE Full text] [doi: 10.1371/journal.pone.0188544] [Medline: 29190761]
11. Hyunshik K, Jiameng M, Sunkyoung L, Ying G. Change in Japanese children's 24-hour movement guidelines and mental health during the COVID-19 pandemic. Sci Rep 2021 Nov 26;11(1):22972 [FREE Full text] [doi: 10.1038/s41598-021-01803-4] [Medline: 34836992]
12. Wei J, Xie L, Song S, Wang T, Li C. Isotemporal substitution modeling on sedentary behaviors and physical activity with depressive symptoms among older adults in the U.S.: the National Health and Nutrition Examination Survey, 2007-2016. J Affect Disord 2019 Oct 01;257:257-262 [FREE Full text] [doi: 10.1016/j.jad.2019.07.036] [Medline: 31301629]
13. Ma J, Ma D, Li Z, Kim H. Effects of a workplace sit–stand desk intervention on health and productivity. Int J Environ Res Public Health 2021 Nov 04;18(21):11604 [FREE Full text] [doi: 10.3390/ijerph182111604] [Medline: 34770116]
14. Hadgraft NT, Brakenridge CL, Dunstan DW, Owen N, Healy GN, Lawler SP. Perceptions of the acceptability and feasibility of reducing occupational sitting: review and thematic synthesis. Int J Behav Nutr Phys Act 2018 Sep 18;15(1):90 [FREE Full text] [doi: 10.1186/s12966-018-0718-9] [Medline: 30227861]
15. Siddarth P, Burggren AC, Eyre HA, Small GW, Merrill DA. Sedentary behavior associated with reduced medial temporal lobe thickness in middle-aged and older adults. PLoS One 2018 Apr 12;13(4):e0195549 [FREE Full text] [doi: 10.1371/journal.pone.0195549] [Medline: 29649304]
16. Falck RS, Davis JC, Liu-Ambrose T. What is the association between sedentary behaviour and cognitive function? A systematic review. Br J Sports Med 2017 May 06;51(10):800-811 [FREE Full text] [doi: 10.1136/bjsports-2015-095551] [Medline: 27153869]
17. Voss MW, Carr LJ, Clark R, Weng T. Revenge of the “sit” II: does lifestyle impact neuronal and cognitive health through distinct mechanisms associated with sedentary behavior and physical activity? Ment Health Phys Act 2014 Mar;7(1):9-24 [FREE Full text] [doi: 10.1016/j.mhpa.2014.01.001]
18. Dillon K, Morava A, Prapavessis H, Grigsby-Duffy L, Novic A, Gardiner PA. Total sedentary time and cognitive function in middle-aged and older adults: a systematic review and meta-analysis. Sports Med Open 2022 Oct 12;8(1):127 [FREE Full text] [doi: 10.1186/s40798-022-00507-x] [Medline: 36224459]
19. Carter SE, Draijer R, Holder SM, Brown L, Thijssen DHJ, Hopkins ND. Regular walking breaks prevent the decline in cerebral blood flow associated with prolonged sitting. J Appl Physiol 2018 Sep 01;125(3):790-798 [FREE Full text] [doi: 10.1152/japplphysiol.00310.2018] [Medline: 29878870]

XSL•FO
RenderX

20. Tari B, Vanhie JJ, Belfry GR, Shoemaker JK, Heath M. Increased cerebral blood flow supports a single-bout postexercise benefit to executive function: evidence from hypercapnia. J Neurophysiol 2020 Sep 01;124(3):930-940 [FREE Full text] [doi: 10.1152/jn.00240.2020] [Medline: 32755360]
21. Byun K, Hyodo K, Suwabe K, Ochi G, Sakairi Y, Kato M, et al. Positive effect of acute mild exercise on executive function via arousal-related prefrontal activations: an fNIRS study. NeuroImage 2014 Sep;98:336-345 [FREE Full text] [doi: 10.1016/j.neuroimage.2014.04.067] [Medline: 24799137]
22. Ichinose Y, Morishita S, Suzuki R, Endo G, Tsubaki A. Comparison of the effects of continuous and intermittent exercise on cerebral oxygenation and cognitive function. Adv Exp Med Biol 2020;1232:209-214 [FREE Full text] [doi: 10.1007/978-3-030-34461-0_26] [Medline: 31893412]
23. Herold F, Wiegel P, Scholkmann F, Müller NG. Applications of functional near-infrared spectroscopy (fNIRS) neuroimaging in exercise–cognition science: a systematic, methodology-focused review. J Clin Med 2018 Nov 22;7(12):466 [FREE Full text] [doi: 10.3390/jcm7120466] [Medline: 30469482]
24. Hillman C, Pontifex M, Raine L, Castelli D, Hall E, Kramer A. The effect of acute treadmill walking on cognitive control and academic achievement in preadolescent children. Neuroscience 2009 Mar 31;159(3):1044-1054 [FREE Full text] [doi: 10.1016/j.neuroscience.2009.01.057] [Medline: 19356688]
25. Chandrasekaran B, Pesola AJ, Rao CR, Arumugam A. Does breaking up prolonged sitting improve cognitive functions in sedentary adults? A mapping review and hypothesis formulation on the potential physiological mechanisms. BMC Musculoskelet Disord 2021 Mar 12;22(1):274 [FREE Full text] [doi: 10.1186/s12891-021-04136-5] [Medline: 33711976]
26. Giurgiu M, Koch ED, Ottenbacher J, Plotnikoff RC, Ebner-Priemer UW, Reichert M. Sedentary behavior in everyday life relates negatively to mood: an ambulatory assessment study. Scand J Med Sci Sports 2019 Sep 23;29(9):1340-1351 [FREE Full text] [doi: 10.1111/sms.13448] [Medline: 31044456]
27. Hanna F, Daas RN, El-Shareif TJ, Al-Marridi HH, Al-Rojoub ZM, Adegboye OA. The relationship between sedentary behavior, back pain, and psychosocial correlates among university employees. Front Public Health 2019 Apr 9;7:80 [FREE Full text] [doi: 10.3389/fpubh.2019.00080] [Medline: 31024881]
28. Lavôr LCDC, Sousa RRD, Rodrigues LARL, Rodrigues Filho ODS, Paiva ADA, Frota KDMG. Prevalence of arterial hypertension and associated factors: a population-based study. Rev Assoc Med Bras 2020 May;66(5):630-636 [FREE Full text] [doi: 10.1590/1806-9282.66.5.630] [Medline: 32638969]
29. Wheeler MJ, Dunstan DW, Smith B, Smith KJ, Scheer A, Lewis J, et al. Morning exercise mitigates the impact of prolonged sitting on cerebral blood flow in older adults. J Appl Physiol 2019 Apr 01;126(4):1049-1055 [FREE Full text] [doi: 10.1152/japplphysiol.00001.2019] [Medline: 30730813]
30. Chang Y, Labban J, Gapin J, Etnier J. The effects of acute exercise on cognitive performance: a meta-analysis. Brain Res 2012 May 09;1453:87-101 [FREE Full text] [doi: 10.1016/j.brainres.2012.02.068] [Medline: 22480735]
31. Maasakkers CM, Melis RJF, Kessels RPC, Gardiner PA, Olde Rikkert MGM, Thijssen DHJ, et al. The short-term effects of sedentary behaviour on cerebral hemodynamics and cognitive performance in older adults: a cross-over design on the potential impact of mental and/or physical activity. Alzheimers Res Ther 2020 Jun 22;12(1):76 [FREE Full text] [doi: 10.1186/s13195-020-00644-z] [Medline: 32571399]
32. Wu Y, Van Gerven PWM, de Groot RHM, Eijnde BO, Winkens B, Savelberg HHCM. Effects of breaking up sitting with light-intensity physical activity on cognition and mood in university students. Scand J Med Sci Sports 2023 Mar;33(3):257-266 [FREE Full text] [doi: 10.1111/sms.14277] [Medline: 36434768]
33. Chia M, Chen B, Suppiah H. Office sitting made less sedentary?a future-forward approach to reducing physical inactivity at work. Montenegrin J Sports Sci Med 2015;4(2):5-10 [FREE Full text]
34. Chan A, Tetzlaff J, Altman D, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 05;158(3):200-207 [FREE Full text] [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
35. Chow HW, Ho CH. Does the use of outdoor fitness equipment by older adults qualify as moderate to vigorous physical activity? PLoS One 2018;13(4):e0196507 [FREE Full text] [doi: 10.1371/journal.pone.0196507] [Medline: 29709035]
36. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020 Dec 25;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
37. Ho CL, Anantharaman V. Relevance of the Get Active Questionnaire for pre-participation exercise screening in the general population in a tropical environment. Healthcare (Basel) 2024 Apr 10;12(8):815 [FREE Full text] [doi: 10.3390/healthcare12080815] [Medline: 38667577]
38. AIBI EZ Tone Chair. AIBI. URL: https://aibifitness.com/consumer/aibi-ez-tone-chair-ab-ut2610.html [accessed 2026-03-07]
39. AIBI Massage Body Twister. AIBI. URL: https://aibifitness.com/consumer/figure-twister.html [accessed 2026-03-07]
40. Cicek G, Imamoglu O, Gullu A, Celik O, Ozcan O, Gullu E, et al. The effect of exercises on left ventricular systolic and diastolic heart function in sedentary women: step-aerobic vs core exercises. J Exerc Sci Fit 2017 Dec;15(2):70-75 [FREE Full text] [doi: 10.1016/j.jesf.2017.07.002] [Medline: 29541135]
41. Kirchner WK. Age differences in short-term retention of rapidly changing information. J Exp Psychol 1958 Apr;55(4):352-358 [FREE Full text] [doi: 10.1037/h0043688] [Medline: 13539317]

XSL•FO
RenderX

42. Soveri A, Antfolk J, Karlsson L, Salo B, Laine M. Working memory training revisited: a multi-level meta-analysis of n-back training studies. Psychon Bull Rev 2017 Aug 23;24(4):1077-1096 [FREE Full text] [doi: 10.3758/s13423-016-1217-0] [Medline: 28116702]
43. Barth B, Mayer-Carius K, Strehl U, Wyckoff SN, Haeussinger FB, Fallgatter AJ, et al. A randomized-controlled neurofeedback trial in adult attention-deficit/hyperactivity disorder. Sci Rep 2021 Aug 19;11(1):16873 [FREE Full text] [doi: 10.1038/s41598-021-95928-1] [Medline: 34413344]
44. Kikuchi S, Tsutsui N, Nishizawa Y, Tsuchiya K, Shimoda K, Hirao K, et al. Habituation of brain activity with repetition in color and picture–word Stroop tests. Ann Biomed Eng 2024 Aug 15;52:2088-2100 [FREE Full text] [doi: 10.1007/s10439-024-03509-w] [Medline: 38622484]
45. Liu Y, Chen L, Guo Y, Lin H. Validation of blood pressure measurement using a smartwatch in patients with acute ischemic stroke. Medicine (Baltimore) 2025 Aug 15;104(33):e42899 [FREE Full text] [doi: 10.1097/MD.0000000000042899] [Medline: 40826739]
46. Hoddes E, Dement WC, Zarcone V. The development and use of the Stanford Sleepiness Scale. Psychophysiology 1972;9(2):150. [doi: 10.1111/j.1469-8986.1972.tb01281.x]
47. Hoddes E, Zarcone V, Smythe H, Phillips R, Dement WC. Quantification of sleepiness: a new approach. Psychophysiology 1973 Jul 30;10(4):431-436 [FREE Full text] [doi: 10.1111/j.1469-8986.1973.tb00801.x] [Medline: 4719486]
48. Kato C, Shimada J, Hayashi K. Sleepiness during shift work in Japanese nurses: a comparison study using JESS, SSS, and actigraphy. Sleep Biol Rhythms 2012;10(2):109-117 [FREE Full text] [doi: 10.1111/j.1479-8425.2011.00528.x]
49. Baradaran A, Ebrahimzadeh MH, Birjandinejad A, Kachooei AR. Cross-cultural adaptation, validation, and reliability testing of the Modified Oswestry Disability Questionnaire in Persian population with low back pain. Asian Spine J 2016 Apr;10(2):215-219 [FREE Full text] [doi: 10.4184/asj.2016.10.2.215] [Medline: 27114759]
50. Grace M, Dempsey P, Sethi P, Mundra PA, Mellett NA, Weir JM, et al. Breaking up prolonged sitting alters the postprandial plasma lipidomic profile of adults with type 2 diabetes. J Clin Endocrinol Metab 2017 Jun 01;102(6):1991-1999 [FREE Full text] [doi: 10.1210/jc.2016-3926] [Medline: 28323950]
51. Wilson SL, Crosley-Lyons R, Junk J, Hasanaj K, Larouche ML, Hollingshead K, et al. Effects of increased standing and light-intensity physical activity to improve postprandial glucose in sedentary office workers: protocol for a randomized crossover trial. JMIR Res Protoc 2023 Aug 23;12:e45133 [FREE Full text] [doi: 10.2196/45133] [Medline: 37610800]
52. Arnold M, Vernikos J. Mobility–non-exercise physical activity (NEPA) is the key to health and longevity. J Nov Physiother Phys Rehabil 2020 Mar;7(1):007-008 [FREE Full text] [doi: 10.17352/2455-5487.000067]

Abbreviations

BP: blood pressure
DLPFC: dorsolateral prefrontal cortex
fNIRS: functional near-infrared spectroscopy
GAQ: Get Active Questionnaire
Hhb: deoxygenated hemoglobin
HR: heart rate
IRB: Institutional Review Board
LIPA: light-intensity physical activity
LMM: large multimodal model
MEF: Movement Ecology Framework
MET: metabolic equivalent
MPA: moderate-intensity physical activity
NIR: near-infrared light
O_2 Hb: oxygenated hemoglobin
OLBDQ: Oswestry Low Back Disability Questionnaire
RA: rating of alertness
RPE: rating of perceived exertion
S-C: sit-cycling
SCWT: Stroop Color and Word Test
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SSS: Stanford Sleepiness Scale
S-T: stand-twisting
VO2: oxygen consumption
VPA: vigorous-intensity physical activity

XSL•FO
RenderX

Edited by J Sarvestan; submitted 19.Aug.2025; peer-reviewed by S Rostami, A Samuel; comments to author 24.Nov.2025; revised version received 24.Feb.2026; accepted 25.Feb.2026; published 17.Mar.2026.

Please cite as:
Ma J, Chua BKT, Chin YHJ, Teo WP, Kim H, Low ST, Li D, Chia YHM
Effects of 5-Minute Light-Intensity Physical Activity Breaks (Sit-Cycling and Stand-Twisting) on Cognitive Function, Sleepiness, and Back Pain in Physically Inactive University Students: Protocol for a Mixed-Design Laboratory Study
JMIR Res Protoc 2026;15:e82510
URL: https://www.researchprotocols.org/2026/1/e82510
doi:10.2196/82510
PMID:

©Jiameng Ma, Buan Kiong Terence Chua, You Hui Jacqueline Chin, Wei Peng Teo, Hyunshik Kim, Seow Ting Low, Dan Li, Yong Hwa Michael Chia. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparison of Uncemented and Hybrid Hip Arthroplasty: Protocol for a Brazilian Randomized Controlled Trial

Bruno Gonçalves Schroder e Souza[1], BSc, MD, PhD; Isabella Dias Lacerda[1]; Matheus Malta Vasconcelos[1]; Nathália Lacerda Furtado[1]; Julia Machado Vieira[2], BSc; Iasmim Sand Ferreira de Souza[2], BSc; Marco Antônio Perígolo Magalhães[2], BSc; Leonã Aparecido Homem do Amaral[2], BSc, PG; Luiz Guilherme Vidal Assad[2], BSc, PG; Carlos Henrique Cavaglieri Silveira Silva[2], BSc; Carolina Tsen[3], BPT, PhD; Valdeci Manoel de Oliveira[1], BSc, MD, PhD

[1]Faculdade de Ciências Médicas e da Saúde de Juiz de Fora – Suprema, Alameda Salvaterra, 200, Juiz de Fora, Brazil

[2]Hospital e Maternidade Therezinha de Jesus, Juiz de Fora, Brazil

[3]Universidade Federal de São Carlos, São Carlos, Brazil

Corresponding Author:
Bruno Gonçalves Schroder e Souza, BSc, MD, PhD
Faculdade de Ciências Médicas e da Saúde de Juiz de Fora – Suprema, Alameda Salvaterra, 200, Juiz de Fora, Brazil

Abstract

Background: Total hip arthroplasty is a highly successful procedure for treating hip arthritis, improving patients' pain, function, and quality of life (QoL). Scarce publications on total hip arthroplasty performance using Brazilian-manufactured materials report results comparable to those from other countries.

Objective: The aim of this study is to compare the clinical, radiographic, functional, pain, and QoL outcomes of patients who underwent surgery with a national hybrid prosthesis (with femoral cementation) versus patients who underwent surgery with a national uncemented prosthesis.

Methods: This study is a single-center, single-surgeon, single-approach, pragmatic, double-blinded, and prospective randomized trial. A total of 120 patients will be enrolled and randomly allocated in a 1:1 ratio to 2 groups: the hybrid group and the uncemented group. The incidence of complications (during and after surgery) and the restoration of normal anatomical parameters on postoperative radiographs will be assessed using radiographic parameters. Participants' QoL, joint mobility, function, and satisfaction will be evaluated using the 12-Item Short Form Health Survey version 2 questionnaire, hip range of motion, the Harris Hip Score, and a numeric rating scale. Prosthesis survival will be analyzed using the annual revision surgery rate from 1 to 5 years, and up to 10 years.

Results: The project did not receive external funding. Data collection began in October 2024 and is ongoing, with completion expected in December 2029. As of April 2025, 72 participants have been enrolled. Preliminary data analysis has been initiated and is ongoing. The first results are expected to be published in the first half of 2026.

Conclusions: This study is a pragmatic clinical trial that uses blinding to evaluate national implants for both hybrid and uncemented hip arthroplasty. Successful completion of this study may provide clinical evidence on the performance of national implants and identify a preferred implant construct (hybrid or uncemented), if any, for hip replacement in Brazil.

Trial Registration: Brazilian Registry of Clinical Trials RBR-263nf22; https://ensaiosclinicos.gov.br/rg/RBR-263nf22

International Registered Report Identifier (IRRID): DERR1-10.2196/79721

(JMIR Res Protoc 2026;15:e79721) doi:10.2196/79721

KEYWORDS

hip; osteoarthritis; surgical procedures; operative; arthroplasty; replacement; randomized controlled trial

Introduction

Total hip arthroplasty (THA) is a highly successful procedure for treating hip arthritis, improving patients' pain, function, and quality of life (QoL) [1,2]. However, THA is still vulnerable to complications such as fractures, loosening, infection, as well as other biological issues [3]. Because patients who undergo THA face the risk of cumbersome and costly revision procedures, the life expectancy and durability of the implant are important issues [3]. Different surgical techniques and implants, such as cemented and uncemented stems, are used to perform THA [4]. However, despite the good results of this procedure, no single surgical technique or type of implant has been proven superior for all patients [4].

To date, considerable controversy persists regarding the best type of surgical implant. While in North America, there is a

XSL•FO
RenderX

preference for noncemented implants [4], this is mainly based on the findings of a 17-year follow-up randomized controlled trial (RCT) published by Corten et al [5] in 2011. In other countries, especially in Europe, the preference for cemented femoral implants has been maintained. The latter is supported by at least 6 RCTs that show no differences between these types of implants [4], as well as reports of a greater incidence of surgical complications with uncemented implants, especially in older patients [6].

Worldwide, the increasing use of uncemented implants has been described as the "paradox of the decline of cemented implants" [7]. In fact, Troelsen et al [8] in 2013 analyzed 7 national registries, including Australia, Canada, England, Denmark, New Zealand, Norway, and Sweden, in which the use of uncemented prostheses showed a steady increase in the use of this kind of prostheses [8]. Recent studies have highlighted an increased "cementless" trend in some countries between 2010 and 2020; however, this trend appeared to be reversing in others, such as England and Wales, Australia, New Zealand, and Finland [9]. Those authors agree that survival was significantly better for cemented prostheses in all registers in people aged older than 70 to 75 years [8,9]. For other age groups, survival was better for cemented prostheses or equivalent, except in the Australian register. Using another approach to the problem, a meta-analysis including 27 studies conducted by Phedy et al [10] reported very similar results, showing equivalence of both types of implants in the younger patients and superiority of cemented fixation in older patients.

While most RCTs in this area of research have traditionally focused on implant survival and revision rates [4], patient-centered outcomes, QoL, and early complications have become a current hot topic. Hybrid THA yields similar short- to mid-term patient-reported outcome measures (Harris Hip Score [HHS], WOMAC, EQ-5D) compared with cemented and uncemented constructs. A 2025 systematic review of 357,748 THAs found no difference in patient-reported outcome measures between hybrid and cemented THA and no clear advantage of uncemented constructs [11]. However, the incidence of periprosthetic fracture was reported to be significantly higher with uncemented stems, especially in older adults. Moreover, a slightly higher incidence of dislocations was also reported for uncemented stems [12]. Regarding pain, direct elective THA comparisons are limited, but evidence consistently shows more early pain and slower recovery with uncemented femoral fixation due to press-fit stem insertion forces. A meta-analysis of RCTs in THA showed better short-term pain scores with cemented fixation [13]. Additionally, a large National Health Service cost-effectiveness analysis (n≈300,000) found higher postoperative EQ-5D scores with hybrid THA (0.81 vs 0.80 for uncemented), and for most patient groups, hybrid prostheses were found to be the most cost-effective [14].

Few publications on THA performance using Brazilian-manufactured materials have reported results comparable to those from other countries; however, the level of evidence and sample sizes of these studies are limited [15-17]. Additionally, early failures of previous national models, commonly used in Brazil until the 1990s, became the reason for persistent questioning about the safety and durability of Brazilian-made implants [18]. To date, no level 1 study evaluating the performance of such implants in Brazil has been identified.

This study aims to compare the clinical outcomes (function, pain, complications, QoL, and need for revisions) and radiographic outcomes of patients treated with a national hybrid prosthesis (with femoral cementation and uncemented acetabular cup) versus those receiving national uncemented prostheses of the same brand.

Methods

Study Design and Setting

This study was written according to the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines (Checklist 1) for protocols of randomized clinical trials. The proposal is a single-center, single-surgeon, single-approach, pragmatic, double-blinded, prospective randomized trial to analyze the differences between uncemented hip replacement and the hybrid (femoral cemented) hip replacement.

Two different types of hip implants will be used for comparison: the Maxima Cemented Vincula Stem (Víncula, Rio Claro, SP, Brazil) and the Taper Uncemented Primary Stem (Víncula, Rio Claro, SP, Brazil). Patients and evaluators will be blinded to group allocation: the hybrid group (HG) and the uncemented group (UG). A total of 120 patients will be enrolled in this study and randomly allocated to each group in a 1:1 ratio.

Recruitment and Preoperative Assessment

All the patients ascribed to a waiting list for hip arthroplasty in a hip reconstruction ambulatory of a tertiary teaching hospital in Juiz de Fora-MG, in southeast Brazil (Hospital e Maternidade Therezinha de Jesus) will form the recruitment base for this study. This study was approved by an institutional review board, registered in the Brazilian Registry of Clinical Trials (RBR-263nf22), and assigned the Universal Trial Number U1111-1313-6912. All eligible patients will be informed verbally about the study and its objectives and will be invited to participate (Figure 1). All exclusions will be recorded and reported. Patients who consent to participate will be invited for a consultation when they sign a written informed consent form and will be assigned a sequential study number.

XSL•FO
RenderX

Figure 1. Flowchart of the study protocol. mHHS: modified Harris Hip Score; mo: months; SF-12: 12-Item Short Form Health Survey; wk: weeks; y: years; *: clinical evaluator; **: surgeon; ***: radiographic evaluator; ****: data curator (supervisor).

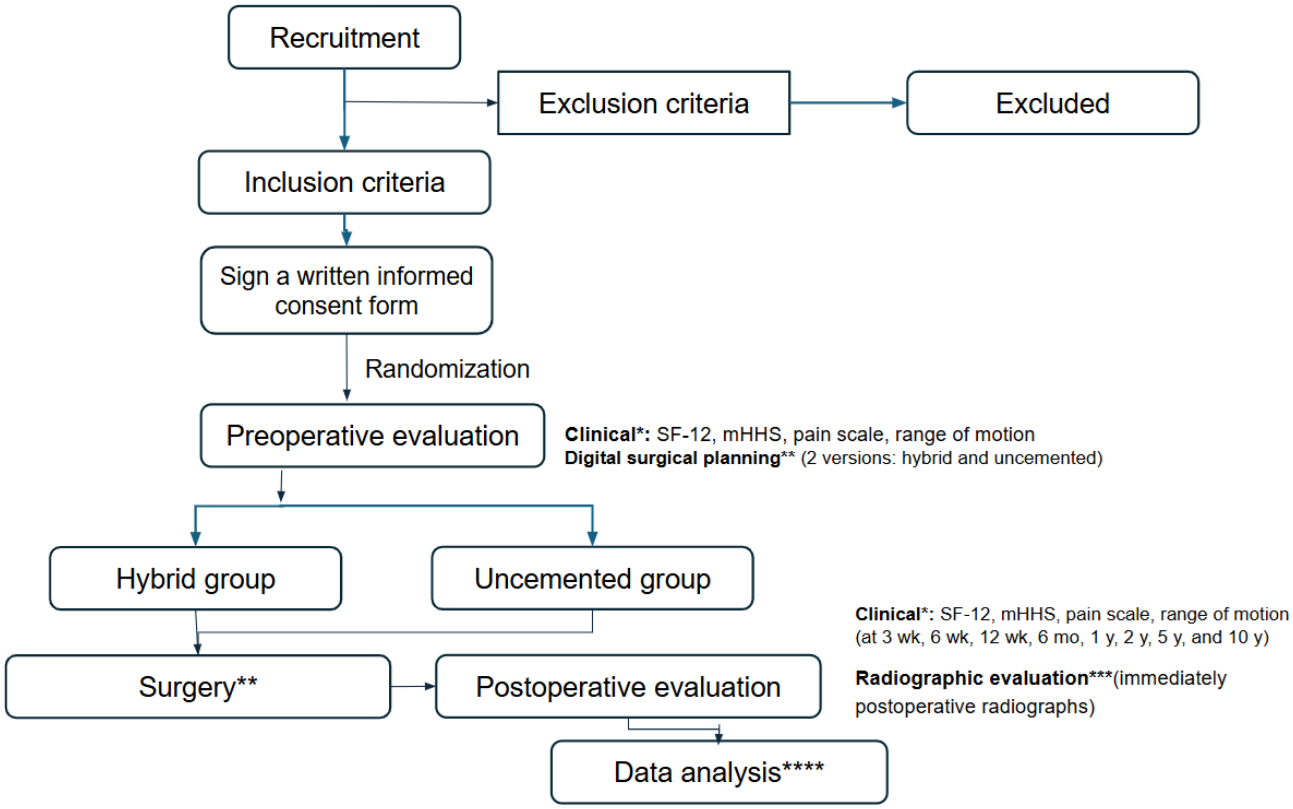

Standard digital anteroposterior pelvis radiographs with calibration spheres will be obtained. A clinical interview and thorough preoperative evaluation will be undertaken by a medical doctor (orthopedic surgery resident). Functional, pain, and QoL questionnaires will be administered at this point, and surgical appointments will be scheduled (Multimedia Appendix 1).

Patient assignment and preoperative evaluations will be conducted in monthly blocks of 12 patients for practical and logistic purposes. Random allocation of patients to each group, the HG and the UG, will be performed at this time using double blinding.

At this stage, the surgeon will be blinded to group allocation and clinical assessment and will perform surgical planning for both uncemented and hybrid procedures using Peekmed software (Peekmed Inc, Braga, Portugal). Peekmed is a highly sensitive tool currently in use that focuses on digital surgical planning to evaluate the accuracy of corrections and to reproduce patient-specific anatomy. Data from both surgical plans will be extracted by the same surgeon using a specially designed digital form (Multimedia Appendices 2 and 3). Reports of the surgical plans will be printed and kept until surgery; however, only the correct plan, as decided by the random group allocation, will be provided to the surgeon by the assistant researcher before surgery.

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria are shown in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Age between 40 and 80 years
- Clinical and radiographic diagnosis of hip arthritis (primary or secondary), unilateral or bilateral
- Both sexes

Exclusion criteria

- History of pathological fractures or active, uncontrolled metabolic and/or rheumatic diseases
- Severe hip deformities with extensive acetabular defects, such as severe hip dysplasia
- Previous hip infection

XSL•FO
RenderX

Withdrawal From the Study

Participants who wish to withdraw after the operation, as instructed in the preparticipation recommendations, may continue with usual treatment at the hospital. In the case of withdrawal at any point after surgery, the data will be considered missing, and data imputation techniques will be evaluated. Data from participants who discontinue the study will remain available in the electronic dataset during revision and will be presented in the results. The study will follow the Consolidated Standards of Reporting Trials (CONSORT) flow chart and tables.

Randomization

Participants will be randomized and allocated to 2 different groups. On the day the participant signs the written consent form, each participant will receive a unique identification number that will be randomized later. Randomization will be performed in 10 random blocks of 12 participants each using Randomized.org. A computer-generated list will assign participants to either (i) the HG or (ii) the UG. An assistant researcher who is not involved in other study roles will carry out this allocation process. Data related to the randomization list will be kept confidential by the researcher in charge of the randomization. Consequently, both the researcher responsible for outcome evaluations and the participants will be blinded to group allocation.

Blinding

Before surgery, the surgeon will be blinded to group allocation and will generate radiographic measurements and surgical planning for both uncemented and hybrid surgeries for each patient. All participants will be blinded to the technique utilized. However, a record for medical use will be available, upon request, in the hospital medical file. The surgeon cannot be blinded during the surgical procedure because this is a clinical trial of surgical treatment techniques. To minimize observer bias, all postoperative clinical and functional evaluations will be performed by independent researchers blinded to participant group allocation and postoperative radiographs. Postoperative radiographs will be analyzed by a separate researcher with experience in hip surgery, who will be blinded to preoperative measurements, surgical planning, and postoperative clinical and functional evaluations.

Therefore, blinding was applied in all 3 stages and across all 3 levels, which is known as triple-blinding. The surgeon is blinded to the preoperative evaluation and performs surgical planning without knowledge of which implant will be used. Blinding is broken at the time of surgery. Participants remain blinded throughout the study. Two types of evaluators are involved: an imaging evaluator and a clinical evaluator. The postoperative evaluator, using the X-ray, is blinded to the preoperative evaluators and the postoperative clinical evaluation. The postoperative clinical evaluator is blinded to both group evaluation and radiographic examination. A description of each researcher involved in the study is provided below in Table 1.

Table . Description of each role in the study.

Research role	Action	How was blinding implemented	Breach of blinding
Surgeon	Independent surgical planning for every patient, generating two archives: hybrid construct and uncemented construct	The surgeon received the digital radiograph of each patient and was unaware of the allocation group	Blinding was relieved only at the time of surgery, when the surgeon received only the applicable copy of the digital plan previously generated, according to the random allocation
Patient	Remained blinded to the allocation group throughout the study	The patient received no information regarding group allocation. The discharge summary from the hospital contained only a reference that he was participating in the THA[a] study and that details about the implant were available on file upon request	No cases of unblinding were observed
Evaluators			
Clinical evaluators	Applied questionnaires and physical exam, blinded to the allocation group and postoperative radiographs	This group of doctors was called in exclusively for this task during the study	No cases of unblinding were observed
Radiographic evaluators	An orthopedic surgeon analyzed exclusively postoperative radiographs using a custom extraction questionnaire	This evaluator had no access to preoperative radiographs, surgical planning, allocation group, clinical evaluations, or the patient themselves	No cases of unblinding were observed

[a]THA: total hip arthroplasty.

XSL•FO
RenderX

Interventions

Initial Common Procedure for Both Groups

Patients will be admitted on the day before surgery and will undergo 8 hours of fasting before surgery. Antibiotic prophylaxis (cefazolin 2 g) will be administered 1 hour before anesthetic incision (clindamycin 600 mg will be used as an alternative in cases of allergy). Patients will receive 1 g of tranexamic acid administered intravenously.

Surgery will be performed under neuraxial block (spinal anesthesia) with sedation. Clinical measurement of limb length discrepancy will be estimated by the surgeon with the patient in a supine position and recorded, along with all surgical details, in a specific digital form (Multimedia Appendix 4).

On a regular surgical table, the patient will be positioned in the lateral decubitus, with a hip-positioning device supporting the sacrum and both anterior superior iliac spines. Limb length will be reassessed at this position.

Routine aseptic cleansing and preparation will be performed, and sterile surgical fields will be positioned. The limb to be operated on will be degermed with polyvinylpyrrolidone iodine solution (chlorhexidine will be used instead for patients with allergy), and excess solution will be removed with alcohol. Antisepsis with alcoholic polyvinylpyrrolidone iodine solution will be performed by the surgeon, and sterile drapes will be positioned.

The same surgeon will conduct the surgery on every patient using a standard posterolateral approach. The piriformis tendon, along with other external rotators, will be divided from the femur and tagged with an Ethibond® 5 sutures for later transosseous reinsertion. The capsule will be opened with an L-shaped incision and tagged with an Ethibond® 5 running suture for later reinsertion. Posterior hip dislocation and femoral neck osteotomy, according to the surgical plan, will then be executed with a power saw.

Preparation of the acetabulum will begin with full exposure using a long Homann retractor placed anteriorly and inferiorly. A 4.0 Steinman pin will be placed superiorly for soft tissue retraction. Another 3.5 Steinman pin will be placed between the hip capsule and the posteroinferior labrum in the ischium to complete exposure.

The acetabular labrum, periacetabular osteophytes, round ligament and pulvinar tissue, and curtain osteophytes exposing the fovea will be removed. A successive reaming of the acetabulum will be done with a high-torque drill to the planned diameter and adequate bone contact. The prosthesis will be tested to verify positioning and adequate press fit (with a target inclination of 45° and anteversion of 20°). Subsequently, an uncemented acetabular component (Phenon poly II, Vincula, Rio Claro, Brazil) will be inserted through impaction using the press-fit technique. Additional fixation with 2 to 3 acetabular screws will be used as needed by the surgeon, depending on the perception of the press-fit and bone quality. Insertion of a highly cross-linked polyethylene liner will be performed through impaction, followed by confirmation of system stability.

After performing the procedure on the acetabulum, patients will undergo the femoral procedure according to the group to which they were allocated.

Femoral Procedure in the HG

Patients allocated for hybrid arthroplasty will undergo a surgical procedure consisting of positioning the femur by internal rotation of 90° and flexion of the hip. The Schumacher-type retractors will be positioned under the anterior surface of the femoral neck, and the remaining capsule and tendons of the piriform fossa will be resected. Following this, the metaphyseal bone with a box-type chisel will be removed, and a cylindrical femoral starting broach will be used to enlarge the canal. According to the surgical plan, the femoral canal will be prepared with successive femoral rasps until adequate stability and restoration of the anatomical parameters of offset and length are achieved.

Following the surgical plan, a femoral neck and head test will be done. Hip reduction, stability test at the extremes of the range of motion (ROM), assessment of limb length, soft tissue tension, and dislocation of the test prosthesis will be done, and the hip reduced to assess stability, limb length restoration, and soft tissue tension. After ensuring adequate size, the surgeon will dislocate the hip joint again and mark it with electrocautery for guidance during definitive implant positioning. Further removal of bone at the Gruen zone I will be undertaken with a curette for adequate cement penetration. A cement restrictor will then be placed, obliterating the femoral canal 2 cm distal to the tip of the femoral stem. Copious irrigation of the canal and thorough cleansing and drying will be performed to remove all blood and medullary content. Cementation of the femoral canal using a retrograde technique with a cement pistol and two cements will be done under pressure. The polished double wedge cemented stem (Máxima, Vincula, Rio Claro, Brazil) will be introduced at a constant speed with a distal centralizer to the planned position. The target anteversion will be 15° with central alignment to the femoral axis.

Femoral Procedure in the UG

Patients allocated to uncemented arthroplasty will undergo a surgical procedure consisting of positioning the femur in 90° of internal rotation with hip flexion. Schumacher-type retractors will be placed beneath the anterior surface of the femoral neck, and the remaining capsule and tendons of the piriform fossa will be resected. Subsequently, the metaphyseal bone will be removed using a box-type chisel, and a cylindrical femoral starting broach will be used to enlarge the canal. The femoral canal will be reamed with successive femoral rasps, according to the plan, until adequate stability and restoration of the anatomical parameters of offset and length.

A testing femoral head will be assembled, and the hip will be reduced. Hip stability, limb length restoration, and soft tissue tension will be assessed. After ensuring adequate size, the surgeon will dislocate the hip joint again and mark it with electrocautery for guidance during definitive implant positioning. The uncemented titanium triple wedge plasma-sprayed femoral stem (Taper, Vincula, Rio Claro, Brazil) will be introduced by hammering with a press-fit technique

XSL•FO
RenderX

down to the planned position. The target anteversion will be 15° with central alignment to the femoral axis.

Final Common Procedure for Both Groups

A metallic femoral head will be introduced and impacted on the Morse taper of the stem. The hip will then be reduced, followed by a final check of stability at the extremes of the ROM, as well as evaluation of limb length and soft tissue tension.

For closure, 2 holes will be drilled in the greater trochanter using a 2.5-mm drill for reinsertion of the external rotators and posterior hip capsule. Ethibond sutures will be tightened for a transosseous repair of the external rotator capsule in anatomical position. The fascia lata will be closed with X-shaped Vicryl® 2.0 separate stitches. The subcutaneous layer will be sutured with inverted Vicryl® 2.0 separate stitches. Finally, a suture of the skin with intradermal stitches (continuous suture) of Nylon 2.0 and a sterile dressing with gauze and adhesive tape will be done.

Postoperative Care

In both groups, early active motion of the operated limb will be permitted as soon as motor function recovers from anesthesia. All participants will be instructed to begin immediate knee flexion and extension exercises, ankle and foot movements, and active calf pump exercises. No orthotic devices or physical positional measures will be utilized to restrict hip motion. However, all patients will be advised to avoid flexion, adduction, and internal rotation of the operated hip, or sitting in low chairs for at least 3 months. Walking with 2 crutches with weight-bearing as tolerated will be stimulated on the first postoperative day until the follow-up visit at 21 days, and patients will be discharged for ambulatory care as soon as ambulation is possible with comfort. Sports and activities on extreme hip motions will be discouraged for at least 3 months.

Primary Outcomes

The primary outcomes of this study will be patient-reported function and complications.

Clinical evaluation will comprise the measurement of hip ROM to assess joint mobility, and the HHS questionnaire will be applied to assess function. Hip ROM will be measured using the goniometer, with the patient lying supine, to determine joint mobility. Hip function will be assessed using the HHS, which is a THA-specific questionnaire composed of patient-reported function and physician-assessed joint ROM, including pain, function, deformity, and mobility assessment [19,20]. The HHS is scored up to 100 points, where each section has a different score, and after applying the questionnaire, all scores are summed up. Scores below 70 are considered low, scores between 70 and 80 are considered fair, scores between 80 and 90 are considered good, and scores between 90 and 100 are considered excellent [19].

The incidence of complications (during and after surgery) will be extracted from an intraoperative questionnaire completed by the surgeon, an immediate postoperative questionnaire reported by the patient, and a special severe adverse effect questionnaire, which will be filled whenever the patient requires readmission or additional surgery.

Secondary Outcomes

Secondary outcomes will include comparisons of QoL, joint mobility, function, satisfaction, and pain before and after the surgery, as well as radiographic parameters extracted from the immediate postoperative radiograph. The annual frequency of revision surgeries will be recorded.

QoL will be evaluated using 12-Item Short Form Health Survey version 2 questionnaire [21]. QoL is defined as an individual's perception of their position in life in the context of the culture and value systems in which they live and about their goals, expectations, standards, and concerns [21]. The 12-Item Short Form Health Survey is an instrument for measuring QoL composed of 12 items that assess eight different dimensions of influence on QoL, considering the individual's perception of aspects of their health in the last 4 weeks [21]. Each dimension's score is summed up, and the final score can be between 0 and 100; the higher the score, the better the QoL [22].

Participant satisfaction assessment will be measured using a 10-point numerical rating scale ranging from 1 (very dissatisfied) to 10 (completely satisfied) [23].

Other aspects of health, such as pain, will be evaluated through the Visual Analog Scale. The Visual Analog Scale provides a simple and efficient measure of pain intensity and is used when a fast pain index that can be assigned to a numerical value is needed [24].

Radiographic parameters will include the ability to restore anatomical normality parameters on postoperative radiographs according to the surgical plan (namely the limb length discrepancy, the horizontal off-set of the hip, difference between the back of the prosthesis and the tip of the trochanter, inclination of the acetabular component, and the difference between offset obtained to the normal side) all of this using the Peekmed surgical planning software.

The annual frequency of revision surgery will be observed starting from 1 year after surgery and annually until 5 years, and up to 10 years.

All questionnaires will be administered by examiners blinded to the surgical procedure. Questionnaires will be completed using digital devices (computer or tablet).

Data Collection and Access

Participant data will be collected using study forms (Multimedia Appendices 1-4) and stored on a designated Google Drive, which will be used as the study repository. A copy of all the files will be stored on an external hard drive for backup. The research supervisor will audit completion of the electronic spreadsheet and will be responsible for data safety and accuracy. Incorrect or missing data will be assessed by this researcher, who will not be involved in other tasks in the study, and corrected when necessary. Data from participants who discontinue the study will remain available in Google Drive for the reviewers and will be presented in the results, CONSORT

XSL•FO
RenderX

flowchart, and tables. Study data will be stored for at least 5 years after completion of the study.

Level of Pragmatism

The level of pragmatism is shown in Figure 2 and detailed in Table 2.

Figure 2. ePragmatism level wheel. 1: very explanatory; 2: rather explanatory; 3: equally explanatory/pragmatic; 4: rather pragmatic; 5: very pragmatic.

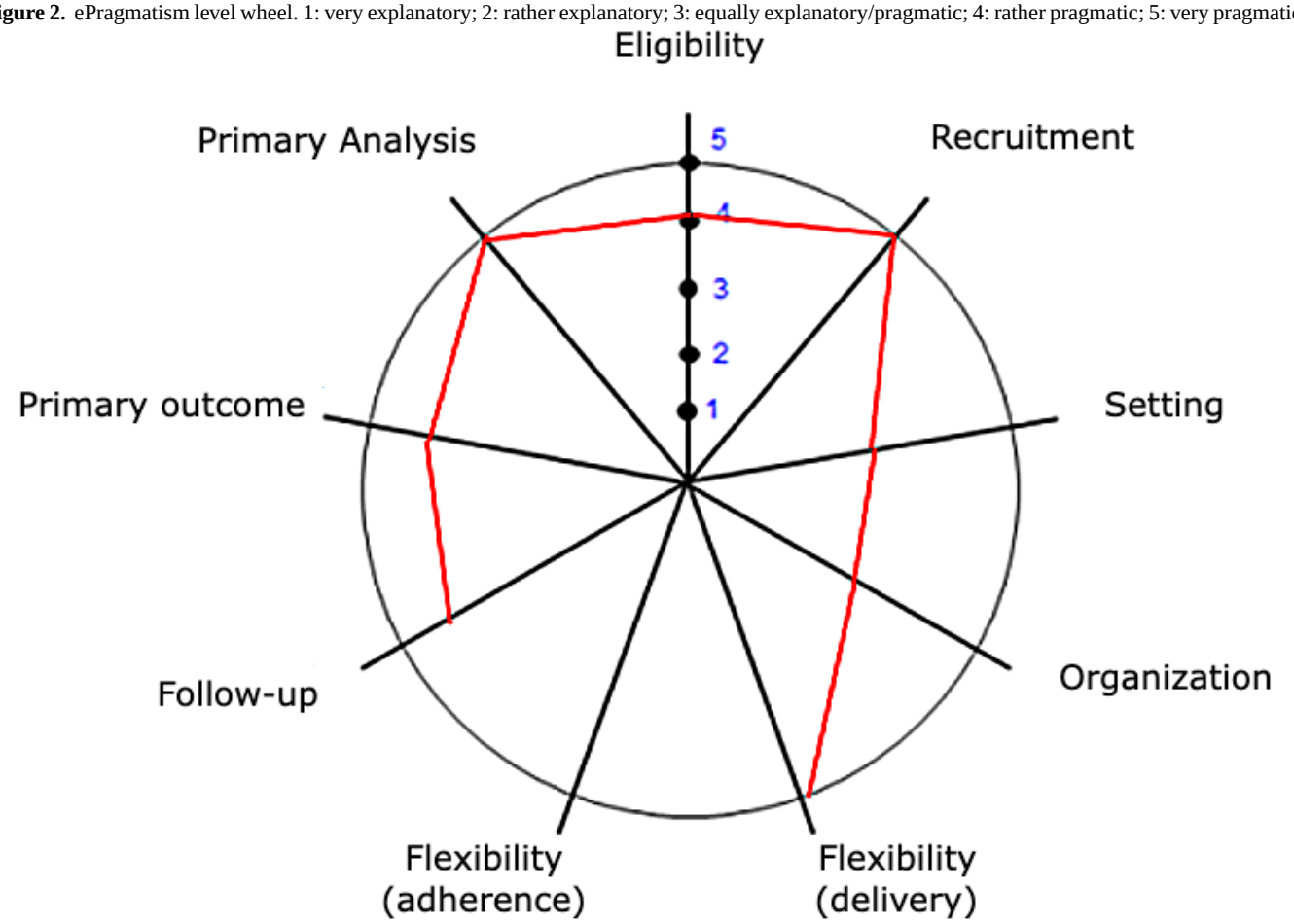

XSL•FO
RenderX

Table . Domains and score of pragmatism level.[a]

Domains	Score	Rationale
1. Eligibility—Who is selected to participate in the trial?	4	All patients already had an indication for total hip arthroplasty in the outpatient clinic. Exclusions were limited to cases in which the disease could compromise the results because of acetabular outcome (such as severe hip dysplasia) or when there was a known indication for other specific implants (eg, ceramic femoral heads for patients younger than 40 years old).
2. Recruitment—How are participants recruited into the trial?	5	Patients were selected from the public waiting list and referred by the health department and/or spontaneous demand. The study was announced, but the waiting list at the beginning of the study already included the entire projected population.
3. Setting—Where is the trial being done?	3	Despite being a tertiary teaching hospital, currently, this is the most common setting to offer this treatment in the public system (Sistema Único de Saúde) in Brazil
4. Organization—What expertise and resources are needed to deliver the intervention?	3	Because of the study, a dedicated set of instruments was kept in the hospital for consistency, which does not usually occur in the routine care. However, the surgical team, operating rooms, and available equipment are the same as the usual care.
5. Flexibility (delivery)—How should the intervention be delivered?	5	Surgery will be undertaken as planned, in the same manner as the usual technique used at the hospital over the last 10 years
6. Flexibility (adherence)—What measures are in place to make sure participants adhere to the intervention?	—	Surgical intervention (to be left blank)
7. Follow-up—How closely are participants followed up?	4	Patient visits to the clinic will follow the same pattern as the usual care. Because diverse questionnaires need to be applied, each visit is expected to be longer than routine consultations
8. Outcome—How relevant is it to participants?	4	Pain, satisfaction, quality of life, function, and need for revision are highly relevant to patients. However, radiographic parameters will also be measured, which are mostly important only for surgeons
9 Analysis—To what extent are all data included?	5	After surgery, all data will be collected and analyzed. An intention-to-treat approach is planned if necessary

[a]1: very explanatory; 2: rather explanatory; 3: equally explanatory / pragmatic; 4: rather pragmatic; 5: very pragmatic.

Sample Size

The sample size was calculated using the Statulator Sample Size Calculation tool [25]. To disprove the null hypothesis of equality, assuming a pooled SD of 10.5 units for the HHS [6], the study would require a sample size of 60 participants per each group (ie, a total sample size of 120 for equal group sizes) to achieve a power of 80% and a significance level of .05 (2-sided) for detecting a true difference in means of 5.39 units between the test and the reference groups (ie, 56.9 vs 62.3).

Statistical Analysis

Results will be presented descriptively, with continuous data expressed as means (SD) and medians, and categorical data will be presented as means or percentages. For inferential analysis, the normality of the variables will be assessed using the Kolmogorov-Smirnov test. Dichotomous variables will be analyzed using the Fisher exact test and chi-square test, and continuous variables will be analyzed using the *t* test. Demographic data will be analyzed using ANOVA test, and the chi-square test will be used for the categorical data. In cases of non-normal distribution, appropriate nonparametric tests will be used according to the nature of the variable.

Primary outcomes, complications, and reoperations will be analyzed at 48 weeks. Paired and subgroup analyses are foreseen between periods of 3, 6, 24, and 48 weeks for the described outcomes. Results will be reported with 95% CI.

All statistical tests will be 2-sided, with a significance level of P=.05. Data will be analyzed using the Online Statistics

XSL•FO
RenderX

Calculator, DATAtab/Numiqo (2025) [26]. Data from participants who receive additional treatments (eg, revision surgery for fractures) during study participation will be retained for intention-to-treat analysis.

Ethical Considerations

This study was approved by the Institutional Review Board of the Faculty of Medical and Health Sciences of Juiz de Fora–Suprema (CAAE: 82413124.4.0000.5103). Any important protocol modifications will be communicated to the ethics committee. Study results will be communicated through the paper publication and will be explained to participants who wish to know their individual results. No compensation will be provided to any participant. All digital records will be stored on a private, password-protected Google Drive. Access to each specific folder will be granted only to the researchers to the extent of their participation. Only the study supervisor will be granted full access to the whole database during the study. Participants must sign a written informed consent form before they are invited for a consultation.

Results

The study started in October 2024. As of April 2025, 72 participants have been enrolled and operated on, corresponding to 60% of the target population. As the study timeline is being maintained, the first data analysis of the complete population was performed in November 2025, reporting population characteristics, radiographic and intraoperative findings, and early complications. The first results are expected to be published in the first half of 2026.

Discussion

Background and Rationale

In Brazil, up to 75% of the population relies exclusively on the public health system (Sistema Único de Saúde [SUS]). From 2008 to 2015, 166,365 total hip replacements were performed only in SUS [27]. After 2012, the number of uncemented and hybrid constructs surpassed the cemented replacements in Brazil [19].

Designed as a universal, integral, and equitable system, SUS should consider the cost-effectiveness of the treatments provided [27]. While total hip replacement has been proven cost-effective in many countries around the world [28], such studies were carried out using data from implants that are either unavailable in Brazil or are much more expensive due to importation fees and taxes. Therefore, although modern national implants have become available and good clinical outcomes have been reported in some studies [15-17]here are no level 1 studies currently available, to date, comparing such implants. Additionally, this study introduces an innovation by being the first to use Peekmed, an artificial intelligence–based preoperative planning software for orthopedic surgery.

Besides clinical information as to whether the Brazilian population should have better outcomes using hybrid or uncemented total hip replacement, the conclusion of this study may also be instrumental for future economic studies applicable to SUS and the private health sector in Brazil.

This study is expected to have several limitations: the fact that it was conducted in one hospital and by one surgeon may raise questions about external validity and reproducibility in other scenarios. On the other hand, this feature provides better homogeneity, increases internal validity, and consistency of results. To minimize this external validity limitation, we used a pragmatic approach to mitigate the potential bias of a very controlled setting. Another potential conflict is that the company that makes the implants provided them at a discount rate and gave equipment to the study team. Even though it is called an “unrestricted agreement,” this still creates a risk of potential conflict of interest because the results could benefit that company. The researchers received no personal benefits, salaries, or incentives, which is relevant for this subject. Additionally, each extract of information was collected independently by the researchers, and the complete database was only accessible to one independent data curator (the research supervisor), who was only responsible for that task and had no role in data collection. An additional measure to reduce bias was the use of implants from the same manufacturer, thereby minimizing the risk of conflict of interest in the comparison between implants. Nevertheless, a potential influence on the overall results cannot be entirely excluded, as a purchase discount was applied. The blinding of the surgeon, the evaluators, and the patients in different stages of the study also reinforces the commitment to diminish biases and increase the independence of the results.

Conclusion

This study is a prospective, randomized, pragmatic clinical trial that uses blinding to evaluate national implants for both hybrid arthroplasty and uncemented arthroplasty. It is the first study with adequate power to prospectively evaluate the outcomes of each surgical procedure in Brazil. Successful completion of this study may provide clinical evidence on the performance of national implants, as well as showing the preferred implant construct (hybrid or uncemented), if any, for hip replacement in Brazil.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. All implants used in this study were provided to the hospital (Hospital e Maternidade Therezinha de Jesus) by the industry (Vincula, Rio Claro, Brazil) as an unrestricted funding agreement. The institution (SUPREMA) also received donations of informatic materials for use in this research as part of this agreement. The medical students participating in the study received scholarships from the Scientific Initiation Program from the Institution (SUPREMA).

XSL•FO
RenderX

Authors' Contributions

BGSeS was assigned as the surgeon and conceptualized the whole project and contributed to the general writing of this study and the final revision before submission.
IDL was assigned to collect clinical data in the study and contributed to writing questionnaires for this protocol.
MMV was assigned to collect clinical data in the study and contributed to writing questionnaires for this protocol and application of the PRECIS-2 tool.
NLF was assigned to collect clinical data in the study and contributed to writing questionnaires for this protocol.
JMV was assigned to collect clinical data in the study and contributed to writing questionnaires for this protocol.
ISFdS was assigned to collect clinical data in the study and contributed to writing questionnaires for this protocol.
LGVA was assigned to work on postoperative radiographic analysis and contributed to the writing of the methods section.
VMdO revised the manuscript and approved for submission.
CHCSS was assigned to collect clinical data in the study and contributed to writing of the results and discussion section.
LAHdA was assigned to work as the study supervisor and contributed to the general writing of this study.
CT contributed to the general writing of this study and the pre-revision before the final revision.
MAPM was assigned to work with the randomization process and contributed to the writing of the methods section.

Conflicts of Interest

BGSeS's institution received the implants for this study at a discounted rate, as stated in the manuscript. CT is a member of the company that provides implants. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Functional, pain, and quality of life questionnaires.
[PDF File, 372 KB - resprot_v15i1e79721_app1.PDF]

Multimedia Appendix 2
Surgical plan 1.
[PDF File, 149 KB - resprot_v15i1e79721_app2.PDF]

Multimedia Appendix 3
Surgical plan 2.
[PDF File, 152 KB - resprot_v15i1e79721_app3.PDF]

Multimedia Appendix 4
Clinical measurement of limb length discrepancy.
[PDF File, 126 KB - resprot_v15i1e79721_app4.PDF]

Checklist 1
SPIRIT checklist.
[DOCX File, 35 KB - resprot_v15i1e79721_app5.docx]

References

1. Learmonth ID, Young C, Rorabeck C. The operation of the century: total hip replacement. Lancet 2007 Oct 27;370(9597):1508-1519. [doi: 10.1016/S0140-6736(07)60457-7] [Medline: 17964352]
2. Jones CA, Voaklander DC, Johnston DW, Suarez-Almazor ME. Health related quality of life outcomes after total hip and knee arthroplasties in a community based population. J Rheumatol 2000 Jul;27(7):1745-1752. [Medline: 10914862]
3. Evans JT, Evans JP, Walker RW, Blom AW, Whitehouse MR, Sayers A. How long does a hip replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. Lancet 2019 Feb 16;393(10172):647-654. [doi: 10.1016/S0140-6736(18)31665-9] [Medline: 30782340]
4. Matar HE, Platt SR, Board TN, Porter ML. Overview of randomized controlled trials in primary total hip arthroplasty (34,020 patients): what have we learnt? J Am Acad Orthop Surg Glob Res Rev 2020 Aug;4(8):e20. [doi: 10.5435/JAAOSGlobal-D-20-00120] [Medline: 32852916]
5. Corten K, Bourne RB, Charron KD, Au K, Rorabeck CH. What works best, a cemented or cementless primary total hip arthroplasty?: minimum 17-year followup of a randomized controlled trial. Clin Orthop Relat Res 2011 Jan;469(1):209-217. [doi: 10.1007/s11999-010-1459-5] [Medline: 20625950]
6. Clement ND, van der Linden M, Keating JF. Higher rate of complications with uncemented compared to cemented total hip arthroplasty for displaced intracapsular hip fractures: a randomised controlled trial of 50 patients. Eur J Orthop Surg Traumatol 2021 Apr;31(3):587-594. [doi: 10.1007/s00590-020-02808-x] [Medline: 33068165]

XSL•FO
RenderX

7. Dujardin F, De Sulauze P, Lalevée M, Matsoukis J, Curado J. The paradox of the decline of cemented femoral stems. Arthroplast Today 2022;19:101057. [doi: 10.1016/j.artd.2022.10.007] [Medline: 36465698]
8. Troelsen A, Malchau E, Sillesen N, Malchau H. A review of current fixation use and registry outcomes in total hip arthroplasty: the uncemented paradox. Clin Orthop Relat Res 2013 Jul;471(7):2052-2059. [doi: 10.1007/s11999-013-2941-7] [Medline: 23539124]
9. Bunyoz KI, Malchau E, Malchau H, Troelsen A. Has the use of fixation techniques in THA changed in this decade? The uncemented paradox revisited. Clin Orthop Relat Res 2020 Apr;478(4):697-704. [doi: 10.1097/CORR.0000000000001117] [Medline: 31899744]
10. Phedy P, Ismail HD, Hoo C, Djaja YP. Total hip replacement: a meta-analysis to evaluate survival of cemented, cementless and hybrid implants. World J Orthop 2017 Feb 18;8(2):192-207. [doi: 10.5312/wjo.v8.i2.192] [Medline: 28251071]
11. Pearce A, Butcher A, Hébert-Losier K. Outcomes of cemented and hybrid primary total hip arthroplasty for osteoarthritis: a systematic review with narrative synthesis. Arch Orthop Trauma Surg 2025 Jul 28;145(1):388. [doi: 10.1007/s00402-025-06007-3] [Medline: 40719919]
12. Lindberg-Larsen M, Petersen PB, Jørgensen CC, Overgaard S, Kehlet H, Lundbeck Foundation Center for Fast-track Hip and Knee Arthroplasty Collaborating Group. Postoperative 30-day complications after cemented/hybrid versus cementless total hip arthroplasty in osteoarthritis patients > 70 years. Acta Orthop 2020 Jun;91(3):286-292. [doi: 10.1080/17453674.2020.1745420] [Medline: 32285735]
13. Abdulkarim A, Ellanti P, Motterlini N, Fahey T, O'Byrne JM. Cemented versus uncemented fixation in total hip replacement: a systematic review and meta-analysis of randomized controlled trials. Orthop Rev (Pavia) 2013 Mar 15;5(1):e8. [doi: 10.4081/or.2013.e8] [Medline: 23705066]
14. Pennington M, Grieve R, Sekhon JS, Gregg P, Black N, van der Meulen JH. Cemented, cementless, and hybrid prostheses for total hip replacement: cost effectiveness analysis. BMJ 2013 Feb 27;346:f1026. [doi: 10.1136/bmj.f1026] [Medline: 23447338]
15. Loures EA, Vitoi JRB, Loures DN, Pereira VC, Mendes AF, Romero V. Preliminary report on Phenom® femoral component in total hip replacement: the correlation between outcome scores in a cross-sectional study. Open J Orthop 2021;11(4):110-125. [doi: 10.4236/ojo.2021.114011]
16. Diesel CV, Ribeiro TA, Macedo CAS, Galia CR. Results of medium-term survival of the non-cemented logical femoral stem. Rev Bras Ortop (Sao Paulo) 2019 Jul;54(4):453-458. [doi: 10.1055/s-0039-1688755] [Medline: 31435114]
17. Loures EDA, Leite ICG. Análise da qualidade de vida de pacientes osteoartrósicos submetidos à artroplastia total do quadril [Article in Portuguese]. Rev bras ortop 2012;47(4):498-504. [doi: 10.1590/S0102-36162012000400017]
18. Picado CHF, Martins LGG, Garcia FL, Kanabushi C. Avaliação radiográfica do acetábulo rosqueado tipo CO-10 [Article in Portuguese]. Rev Bras Ortop 2004;39:189-195 [FREE Full text]
19. Guimarães RP, Viamont-Guerra MR, Antonioli E, Lenza M. Total hip arthroplasty in the public health system of São Paulo: comparing types of fixation. Acta Ortop Bras 2022;30(5):e251150. [doi: 10.1590/1413-785220223005e251150] [Medline: 36451786]
20. Edwards PK, Queen RM, Butler RJ, Bolognesi MP, Lowry Barnes C. Are range of motion measurements needed when calculating the Harris Hip Score? J Arthroplasty 2016 Apr;31(4):815-819. [doi: 10.1016/j.arth.2015.10.016] [Medline: 26639985]
21. Damásio BF, Andrade TF, Koller SH. Psychometric properties of the Brazilian 12-Item Short-Form Health Survey Version 2 (SF-12v2). Paidéia (Ribeirão Preto) 2015;25(60):29-37. [doi: 10.1590/1982-43272560201505]
22. Silveira MF, Almeida JC, Freire RS, Haikal DS, Martins A. Propriedades psicométricas do instrumento de avaliação da qualidade de vida: 12-item health survey (SF-12) [Article in Portuguese]. Ciênc saúde colet 2013 Jul;18(7):1923-1931. [doi: 10.1590/S1413-81232013000700007] [Medline: 23827896]
23. Okafor L, Chen AF. Patient satisfaction and total hip arthroplasty: a review. Arthroplasty 2019 Sep 2;1(1):6. [doi: 10.1186/s42836-019-0007-3] [Medline: 35240763]
24. Begum MR, Hossain MA. Validity and reliability of Visual Analogue Scale (VAS) for pain measurement. J Med Case Rep Rev 2022;2(11) [FREE Full text]
25. Sample size calculator for comparing paired proportions. Statulator. URL: https://statulator.com/SampleSize/ss2PP.html [accessed 2026-02-06]
26. Statistics calculator. Numiqo. URL: https://numiqo.com/statistics-calculator/charts [accessed 2026-02-06]
27. Gonçalves Schroder e Souza B, Coelho Pereira FJ, Gomes Tabet C, Ribeiro Monte LF, De Oliveira VM, Chaoubah A. Artroplastias de quadril no Sistema Único de Saúde: análise dos dados brasileiros de 2008 a 2015. HU rev 2019;45(2):185-194. [doi: 10.34019/1982-8047.2019.v45.26274]
28. Daigle ME, Weinstein AM, Katz JN, Losina E. The cost-effectiveness of total joint arthroplasty: a systematic review of published literature. Best Pract Res Clin Rheumatol 2012 Oct;26(5):649-658. [doi: 10.1016/j.berh.2012.07.013] [Medline: 23218429]

XSL•FO
RenderX

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
HG: hybrid group
HHS: Harris Hip Score
QoL: quality of life
RCT: randomized controlled trial
ROM: range of motion
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SUS: Sistema Único de Saúde
THA: total hip arthroplasty
UG: uncemented group

Edited by J Sarvestan; submitted 26.Jun.2025; peer-reviewed by ME Maitland; revised version received 07.Jan.2026; accepted 07.Jan.2026; published 16.Mar.2026.

Please cite as:
Souza BGSE, Lacerda ID, Vasconcelos MM, Furtado NL, Vieira JM, Souza ISFD, Magalhães MAP, Amaral LAHD, Assad LGV, Silva CHCS, Tsen C, Oliveira VMD
Comparison of Uncemented and Hybrid Hip Arthroplasty: Protocol for a Brazilian Randomized Controlled Trial
JMIR Res Protoc 2026;15:e79721
URL: https://www.researchprotocols.org/2026/1/e79721
doi:10.2196/79721

© Bruno Gonçalves Schroder e Souza, Isabella Dias Lacerda, Matheus Malta Vasconcelos, Nathália Lacerda Furtado, Julia Machado Vieira, Iasmim Sand Ferreira de Souza, Marco Antônio Perígolo Magalhães, Leonã Aparecido Homem do Amaral, Luiz Guilherme Vidal Assad, Carlos Henrique Cavaglieri Silveira Silva, Carolina Tsen, Valdeci Manoel de Oliveira. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical Efficacy Evaluation of Tuina Combined Medicated Oil in the Treatment of Nonspecific Low Back Pain: Protocol for a Randomized Controlled Trial

Man Yuan[1*], PhD; Guangxin Guo[2*], PhD; Siyu Wang[2*], MD; Fei Yao[2], PhD; Guimao Wang[2*], PhD; Hongxi Xu[1*], PhD

[1]School of Pharmacy, Shanghai University of Traditional Chinese Medicine, 1200 Cailun Road, Pudong, Shanghai, China

[2]Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Nunmber 274 Zhijiang Middle Road, Jing'an District, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Hongxi Xu, PhD
School of Pharmacy, Shanghai University of Traditional Chinese Medicine, 1200 Cailun Road, Pudong, Shanghai, China

Abstract

Background: Nonspecific low back pain (NSLBP) is a significant global public health concern that affects the health and well-being of individuals across different age groups, limiting their daily activities and reducing their quality of life. As Tuina (Chinese therapeutic massage) therapy and medicated oil are widely used in China, it is necessary to design a randomized clinical trial to assess the effectiveness of Tuina combined medicated oil (TNO) in treating NSLBP.

Objective: This study aims to evaluate the efficacy and safety of Tuina therapy combined with medicated oil in treating NSLBP.

Methods: One hundred participants will be enrolled and randomly allocated to either a TNO group (n=50) or a Tuina combined water group (n=50). Treatment will last for 4 weeks, with sessions 3 times a week, followed by a 4-week follow-up. The visual analog scale score is the primary outcome; secondary outcomes include the evaluation of treatment Japanese Orthopaedic Association scores, infrared thermography, muscle tension tests, and tenderness scores. All adverse reactions will be recorded.

Results: The trial commenced in June 2023 and is expected to conclude in September 2025. In June 2025, key preliminary steps were completed, and the ethical review and clinical trial registration were concluded. Recruitment is proceeding as planned, with 100 participants enrolled to date. Data collection is underway, while formal data analysis has not yet commenced.

Conclusions: The project's purpose is to evaluate the effectiveness of TNO in alleviating pain and improving lumbar function in patients with NSLBP.

Trial Registration: ChiCTR ChiCTR2300076144; https://www.chictr.org.cn/hvshowproject.html?id=247247

International Registered Report Identifier (IRRID): DERR1-10.2196/81887

(JMIR Res Protoc 2026;15:e81887) doi:10.2196/81887

KEYWORDS

Tuina; medicated oil; non-specific low back pain; clinical effect; randomized controlled trial

Introduction

Nonspecific low back pain (NSLBP) clinical symptoms include low back pain, muscle tension, stiffness below the rib margin and above the gluteal sulcus, leg pain, numbness, and weakness. NSLBP is one of the leading causes of disability among workers in high-income and transitional countries, ranking sixth in overall disease burden [1,2]. NSLBP is a significant global public health issue that impacts individuals, their health, and society. It often recurs and imposes a burden on the economy. It is the most expensive medical condition in the United States, estimated at US $134.5 billion in 2016 [3,4].

According to an epidemiological survey [5], the incidence rate of NSLBP increased by 18% from 2006 to 2016. With the change of people's lives and learning styles, low back pain is increasingly affecting younger individuals. A study of 404,206 adolescents from 28 countries found that 37% experienced low back pain every month or more often [6]. Risk factors associated with NSLBP include obesity, lack of exercise, poor occupational or lifestyle habits, depression, and other social and psychological factors [7]. Maintaining a normal weight, good physical health, proper posture, and regular exercise may reduce the risk, but evidence-based support is lacking [8,9].

NSLBP accounts for the highest proportion of low back pain [1]. NSLBP affects individuals of different age groups, limits

XSL•FO
RenderX

their daily activity function, and reduces their quality of life [10,11]. There is no specific treatment for NSLBP; the purpose of therapy is to alleviate pain and its consequences, including any disabilities related to NSLBP. Traditional Chinese medicine (TCM) external treatment methods, Tuina, offer unique advantages in clinical applications, as they are noninvasive, have relatively low treatment costs, are simple to apply, and are highly acceptable [12-14]. Clinical studies on treating lower back pain with Tuina, both domestically and internationally, have demonstrated that massage therapy effectively reduces muscle and myofascial tension, promoting local metabolism [12,13,15,16].

Currently, the treatment methods for patients with NSLBP are divided into surgical and nonsurgical therapies. Nonsurgical therapies include cognitive behavioral therapy, exercise, manual therapy, physical therapy, acupuncture, and pharmacological therapy [17-20]. The North American Spine Society indicates that there is currently insufficient evidence to recommend or oppose a fusion surgery for lower back pain. There is also insufficient evidence to suggest that imaging results of lumbar fusion are associated with better clinical outcomes, with a level I recommendation [14]. Different treatment methods have their advantages and disadvantages. Surgical therapy has high costs, while traumatic treatment is insufficient, and its traumatic nature limits its effectiveness. Nonsurgical therapies, such as exercise therapy, can be challenging to adhere to. The practical and convenient treatment methods must be explored in clinical practice. Tuina acts on the body's surface through mechanical stimulation as an alternative and complementary therapy. Patients easily accept Tuina therapy and feel highly comfortable. Medicated oil is an effective, safe, and easy-to-use treatment [21,22]. However, clinical trials of Tuina combined medicated oil (TNO) for treating lower back pain generally have a low level of evidence-based methods, lack a high-quality randomized controlled trial design, and are not included in clinical diagnosis and treatment guidelines.

Tuina has unique advantages as one of the traditional external treatment methods in TCM. Previous studies have confirmed its role in various waist-related diseases such as NSLBP, lumbar disc herniation, and acute lumbar sprain [23-26]. Tuina therapy emphasizes the balance of muscles and bones, prioritizing tendons and the presence of tendons that can lead to dislocation of bones, both of which are primary pathogenic factors. Tuina has a holistic view of TCM in the practice of "principles, methods, formulas, and techniques," and has the characteristics of benign and bidirectional regulation in the treatment mechanism. It has been widely applied in clinical practice to treat musculoskeletal diseases.

The Flying Eagle Wood Lok Medicated Oil (FEMO) used in this study has effects that include warming meridians, dispelling cold, wind, and dampness, promoting blood circulation, and relieving pain. Medicated oil is a traditional Chinese patent medicine and a simple preparation made according to the principles of TCM. It exhibits characteristics of percutaneous penetration, rapid onset, and minimal gastrointestinal side effects.

The main components of FEMO consist of *Panax notoginseng* (Burk.) F.H. Chen, *Aconitum carmichaelii* Debx., *Arisaema erubescens* (Wall.) Schott, *Pinellia ternata* (Thunb.) Breit, *Carthamus tinctorius* L., methyl salicylate (wintergreen oil), menthol, camphor, eucalyptus oil, lavender oil, rose oil, and olive oil [27,28]. Modern medical research has shown that medicated oil has anti-inflammatory, analgesic, and preventive effects on pressure ulcers [29]. It is applied to treat muscle soreness, rheumatic bone pain, old and new sprains, arthritis pain, mosquito bites, sprain swelling and pain, low back and leg tendon pain, bone spurs, sciatica, and other diseases. NSLBP is one of its indications. For usage, follow the instructions and apply Tuina techniques, including finger pressure, for 15 - 20 minutes to enhance the effectiveness of the drugs.

Although Tuina is widely used in the prevention and treatment of NSLBP, the current literature on its efficacy often lacks a rigorous experimental design, significantly weakening the persuasiveness of the research results. Although clinical trials have demonstrated effectiveness, no relevant evidence exists to explain the mechanism of action, which also applies to traditional Chinese patent medicines and simple preparations when rubbed with Tuina. This study is based on combining manual therapy with the external application of medicated oil in treating NSLBP, to fully leverage the complementary advantages of manual therapy combined with external treatment drugs. This study aims to provide an effective method for treating NSLBP and to objectively evaluate the clinical application of medicated oil by combining TCM Tuina with medicated oil.

Methods

Ethical Considerations

This study has been approved by the Ethics Committee of Shanghai Municipal Hospital of Traditional Chinese Medicine (2023SHL-KY-96-01). All participants provided written informed consent before participation. Participant privacy and confidentiality were strictly protected throughout this study; all data were anonymized and stored securely, and no personally identifiable information was disclosed. Participants received no compensation.

Study Design

The research is a single-site study, parallel-controlled, randomized controlled trial. We will recruit 100 patients with NSLBP from the Shanghai Municipal Hospital of Traditional Chinese Medicine. Participants will receive informed consent during the recruitment process and will have equal opportunities to be randomly allocated to the TNO or Tuina combined water (TNW) groups. Due to limitations in intervention methods, outcome evaluators and statisticians will be kept blinded to the treatment assignment. Evaluators will evaluate and analyze the results at 5 time points (preintervention, 1, 2, and 4 wk after intervention, and 1-mo follow-up). Data management and statistical analysis will be performed at the Shanghai TCM Hospital Research Department. Figure 1 provides a visual representation of the research design, while Figure 2 displays the research schedule.

XSL•FO
RenderX

Figure 1. The trial flowchart. This study is a randomized controlled trial. A total of 100 patients will be randomly divided into 2 groups, the TNO group and the TNW group. Each group has the same number of patients, and the trial will include a 4-week treatment period. During the treatment process, patients in the TNO group received massage combined with medicated oil intervention, while patients in the TNW group received massage combined with water intervention. Two outcome evaluations will be conducted at 5 time points: baseline, the first week, the second week, the fourth week of treatment, and the eighth week of follow-up. JOA: Japanese Orthopaedic Association scores; TNO: Tuina combined medicated oil; TNW: Tuina combined water; VAS: visual analog scale.

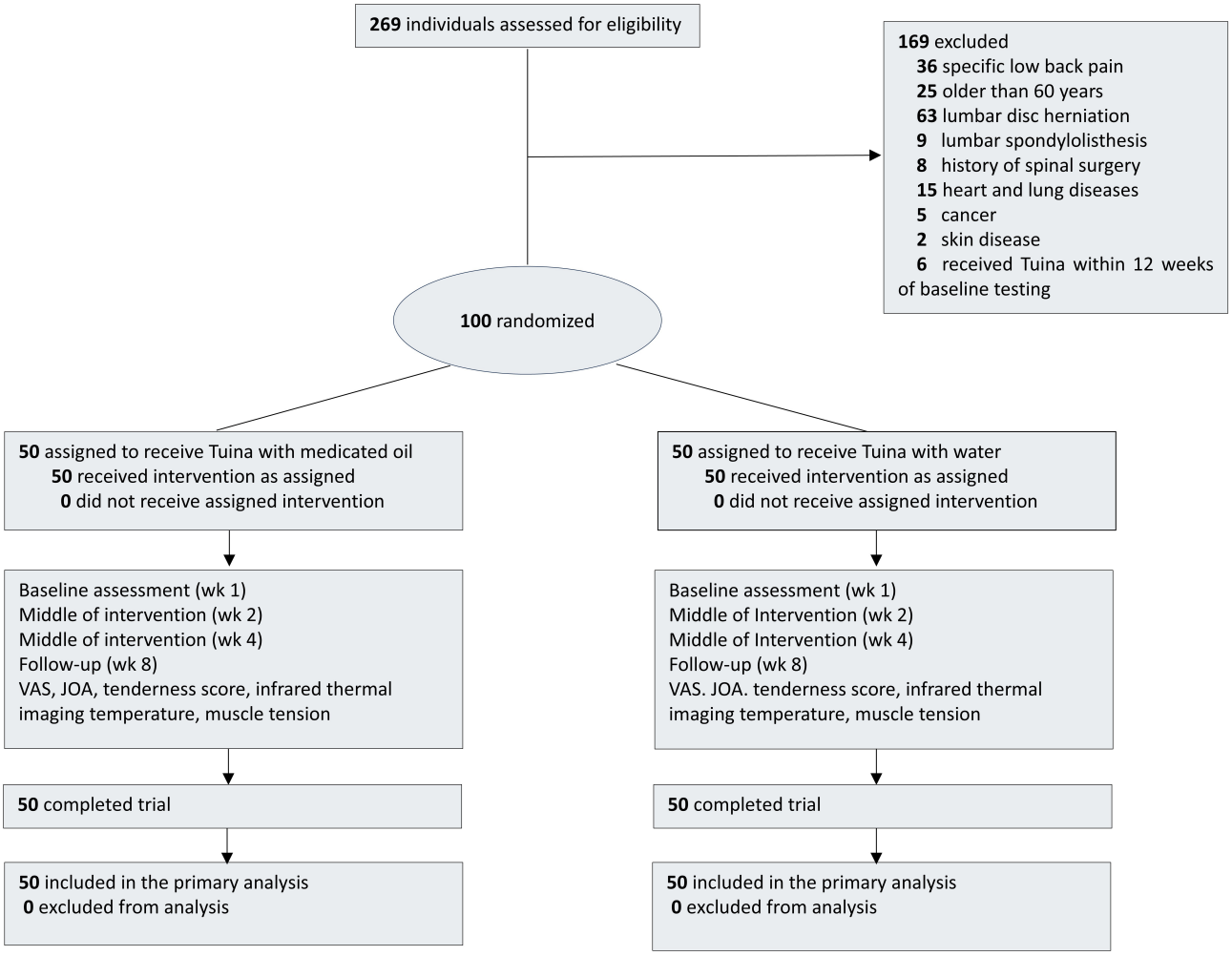

XSL•FO
RenderX

Figure 2. This study's schedule shows the time points (wk [W]) for enrollment and assessment.

Time	W-0	W-1	W-2	W-4	W-8
Period	Baseline	Intervention	Intervention	Intervention	Follow-up
Enrollment	✓				
Eligibility	✓				
Demography	✓				
Informed consent	✓				
Sign the informed consent	✓				
Medical history	✓				
Physical examination	✓				
Randomization	✓				
Interventions					
Tuina combined medicated oil (TNO) group (n=50)		3 sessions per week of Tuina combined medicated oil (TNO)			
Tuina combined water (TNO) group (n=50)		3 sessions per week of Tuina combined water (TNW)			
Assessments					
Visual analog scale (VAS)	✓	✓	✓	✓	✓
Japanese Orthopaedic Association scores (JOA)	✓	✓	✓	✓	✓
Tenderness score	✓	✓	✓	✓	✓
Infrared thermal imaging	✓	✓	✓		
Muscle tension test	✓	✓	✓		
Trial evaluation		✓	✓	✓	✓
Patients compliance		✓	✓	✓	✓
Safety evaluation		✓	✓	✓	✓
Adverse events		✓	✓	✓	✓
Analysis	✓			✓	✓

Sample Size Calculation

Based on the preliminary experimental research results, the sample size was estimated, and G*Power (version 3.1; Heinrich-Heine-Universität Düsseldorf) software was used to calculate the sample size for each group. The sample size was calculated to detect a between-group difference in the change

XSL•FO
RenderX

in visual analog scale (VAS) score from baseline to week 4 (end of treatment), based on pilot posttreatment data. No power calculations were performed for week 8 outcomes, which are considered exploratory.

For grasp 1, the testing level was set at α=.05 and β=.95, and the required sample size was calculated to be 42 per group, considering a 10% dropout rate. According to 1:1 parallel grouping, VAS's mean for the 2 pre-experimental groups was 3.57 (SD 1.35) in the TNO group and 2.61 (SD 1.39) in the TNW group. Using the above parameters, 50 patients with NSLBP should be included in each of the 2 groups, yielding a mean difference of 0.96 and a pooled SD of 1.37 (corresponding to a Cohen *d* effect size of 0.70), indicating a moderate-to-large treatment effect. The high statistical power was chosen to ensure robust internal validity, particularly given the modest, single-center study design and potential overestimation of the effect size from pilot data.

Participants Recruitment

This study used a randomized controlled, single-blind trial method, with participants recruited from the outpatient departments of Shanghai Municipal Hospital of TCM. A total of 100 patients were randomly assigned to an observation group (TNO) and a control group (TNM), with 50 patients in each group (Textbox 1).

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Meeting the diagnostic criteria for the nonspecific low back pain set by the North American Spine Society and Chinese expert consensus [30,31].
- Being between the ages of 20 and 59 years.
- During the trial period of the research method, immediately stopping any therapies that may affect the determination of research effectiveness.
- Participating in this study, cooperating with the doctor's arrangements, and signing an informed consent form.

Exclusion criteria

- Excluding specific low back pain lesions such as lumbar disc herniation and lumbar spondylolisthesis, and those related to bone joints and the spinal canal.
- Being patients with severe heart, lung, liver, and kidney diseases.
- Being individuals undergoing other clinical trials.
- Being individuals with a history of spinal surgery, spinal injury or deformity, or fracture.
- Being patients with combined acute infections, cancer, rheumatic and immune diseases, and mental illnesses.
- Being individuals with other primary functional disorders.

Dropout Criteria

The exclusion criteria were (1) not persisting in treatment or not following the treatment plan, (2) voluntarily requesting withdrawal, (3) having other serious concurrent diseases during the treatment process, and (4) using other treatment methods in combination.

Washout Period and Concomitant Treatments

No formal washout period was required before enrollment. To minimize potential confounding, all concomitant treatments used before enrollment, including physical therapy and medications, were comprehensively documented at baseline. During this trial period, participants were required to discontinue any additional therapies that could interfere with outcome assessment immediately, as specified in the inclusion criteria.

Randomization and Allocation Concealment

Using SPSS (version 27.0; IBM Corp) for Windows software, a random number table was generated and randomly grouped into 2 sets in order of enrollment: the observation group, treated with a combination of Tuina and medicated oil; and the control group received the same Tuina technique, but the activating oil was changed to water, forming a control group with the treatment group.

Blinding

Due to the nature of manual therapy combined with topical application, complete blinding of all parties is inherently challenging. Several measures were implemented to maximize blinding feasibility. They were the following: (1) participants neither handled nor observed the bottles during treatment, and allocation will be concealed from all the participants; (2) the patients will be informed that they will receive one of two effective interventions, which will be randomly assigned after enrollment; (3) this study's coordinator will notify participants of the randomized results and schedule any necessary interventions; (4) the group allocation will be kept concealed from data managers, statisticians, and evaluators throughout the procedure to minimize bias; (5) medicated oil and water are packed in bottles of the same shape, size, and color, and the bottles are encoded according to a random sequence; and (6) the bottle cap of water is soaked in activating oil to replicate its odor. In practical Tuina applications, apart from its odor, the medicated oil used in this study exhibits physical properties comparable to those of water, including fluidity, spreadability, viscosity, and skin contact characteristics. The Tuina practitioner and the cutaneous sensation experienced by participants during manipulation are broadly similar between the 2 intervention media in handling properties, which reduces the likelihood that therapists can reliably distinguish group allocation during

XSL•FO
RenderX

treatment based on touch alone, thereby further limiting potential operator-expectation bias.

Observation Group (TNO Group)

The dedicated attending physician of the Tuina Department is responsible for the treatment of Tuina. Tuina doctors have graduated from the Tuina College of Shanghai University of TCM, majoring in acupuncture, moxibustion, and Tuina, and have been occupied with clinical Tuina work for 5 - 8 years. We finished standardized training on the techniques before treatment begins and ensured patients are treated only after passing the assessment. According to the diagnostic and treatment standards, the Tuina techniques established by the expert group will be used for intervention. The specific plan is as follows:

1. Rubbing tendons is applied to the paraspinal erector spinae muscle of the lumbar spine using the palm root kneading method. Up and down reciprocating operation is performed in 3 minutes. Apply, in 3 minutes, the thumb Tuina method to the attachment area of the iliac crest muscle, the affected buttocks, popliteal fossa, and calf, focusing on the areas where the patient feels sore and swollen.
2. Apply the thumb folding finger flicking technique to the lumbar quadratus muscle, transverse process of the third lumbar vertebra, and Ashi acupoint (tenderness point or local cord-like or nodular area), with 3 - 5 flicks per area, reaching the affected area. To loosen adhesions and relieve spasms for 2 minutes.
3. Grab the affected buttocks and calf muscles for 2 minutes.
4. The patient is advised to take a standing position and perform extreme movements such as waist flexion, lateral flexion to the left and right, and forward flexion to the left and right for 2 minutes. This is to stretch the muscles in the lower back.
5. The patient is prone when using medicated oil for external rubbing. According to the instructions, take 8 - 10 drops of the medicated oil. First, apply the medicated oil evenly to the affected area using the palm-rubbing method, and then take 6 - 8 drops of the medicated oil. Use the palm rubbing method to horizontally rub the lumbar and sacral areas and vertically rub the Du meridian and foot sun meridian tendons, each about 5 - 7 times, with a degree of heat transmission. At the pain point, use techniques combined with activating meridian oil to perform Tuina and pluck for 8 minutes.

The treatment frequency is 3 times a week, for 20 minutes each time.

The treatment cycle is for 4 weeks, with a total of 12 treatments. The specific action characteristics are shown in Figures 3A-3E.

XSL•FO
RenderX

Figure 3. (A) Lumbar flexion, (B) left lumbar flexion, (C) right lumbar flexion, (D) left anterior lumbar flexion, and (E) right anterior lumbar flexion.

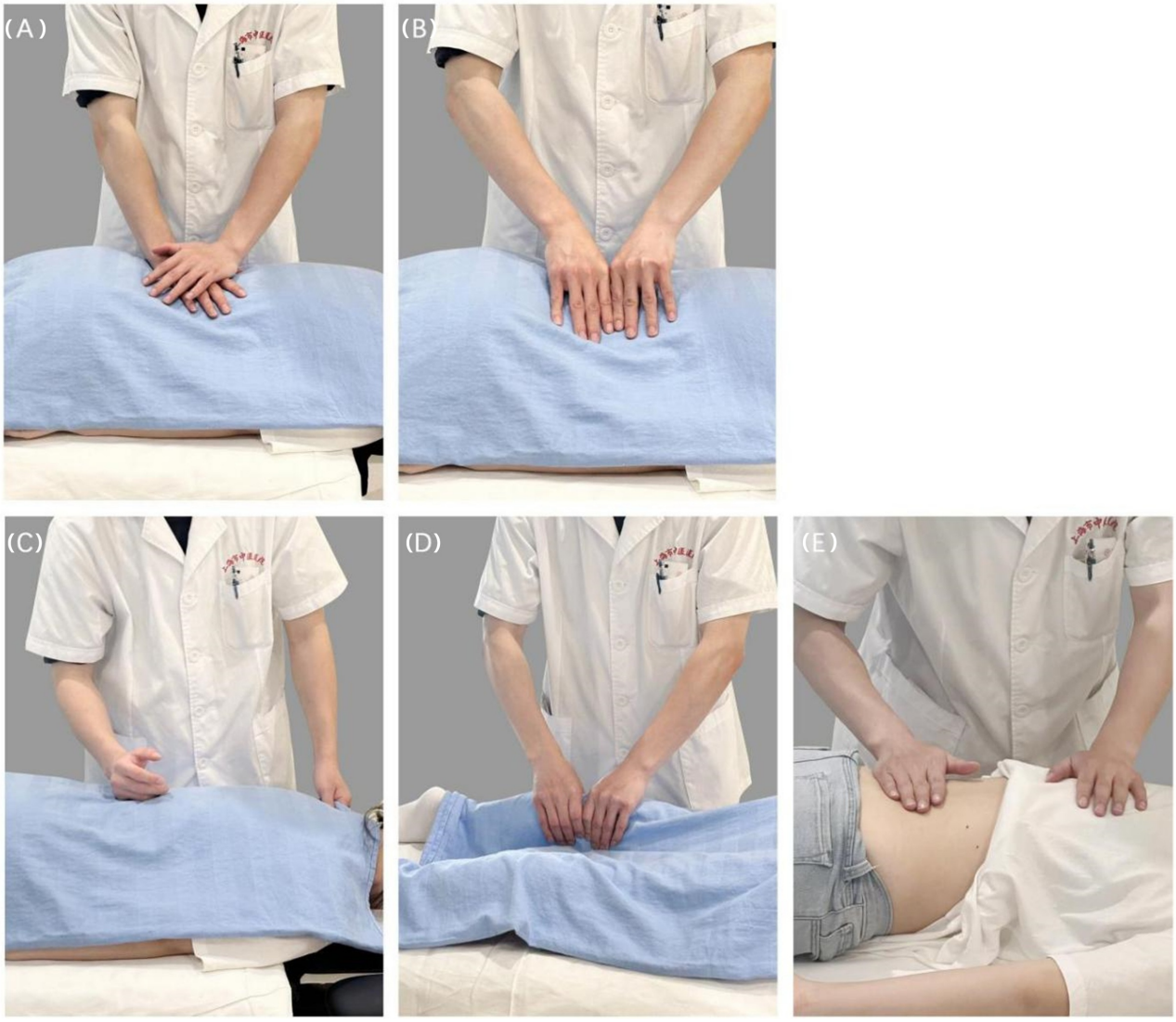

Medicated Oil Drug Information

The brand was FEMO (executive standard number: JZ2017004).

The product registration number was ZC20170003, with a specification of 50 mL.

The medicinal ingredients were *P notoginseng* (Burk) F.H. Chen, *A carmichaelii* Debx., *A erubescens* (Wall.) Schott, *P ternata* (Thunb.) Breit, *C tinctorius* L., methyl salicylate (wintergreen oil), menthol, camphor, eucalyptus oil, lavender oil, rose oil, and olive oil.

FEMO is a government-approved proprietary Chinese medicine manufactured under good manufacturing practice standards. The alkaloids from *A carmichaelii* are strictly controlled during processing to ensure compliance with safety requirements. The product is approved for external use and has been lawfully marketed in Hong Kong (registration number: HKC-00369).

Control Group (TNW Group)

Using the Tuina technique in the above plan, we performed step 5. When we applied the medicated oil externally, we replaced the medicated oil with clean water to form a control.

The treatment frequency was 3 times a week, for 20 minutes each time. The treatment cycle was 4 weeks.

Outcome Measurement

Through the combination of VAS, Japanese Orthopaedic Association scores (JOA), tenderness score, infrared thermal imaging, and muscle tension testing, this study comprehensively evaluates the effectiveness of the intervention, gathering both subjective (patient-reported) and objective (physiological) outcomes. The primary end point is the change in VAS score from baseline to week 4 (end of treatment). VAS is assessed at baseline and weeks 1, 2, 4, and 8; however, the week 4 assessment serves as the prespecified primary analysis time point.

Primary Outcome

The VAS scoring standard, also known as the pain level scoring system, is a visual analog scale method used to evaluate the severity of pain. A 10-cm-long straight line is used, with both ends representing 0 points for painless and 10 points for extreme pain. The scale is: 0 is no pain, 1 - 3 indicates mild pain, 4 - 6 moderate pain, and 7 - 10 severe pain. After introducing the

XSL•FO
RenderX

key points, the patient will self-evaluate the degree of pain based on their feelings [32].

Secondary Outcome Measurements

About JOA

The JOA scale evaluates both subjective symptoms and objective signs. The total score of the scale is 29 points, assessed across 5 aspects and 25 items [33]. The score is directly proportional to the function. The larger the scale score, the milder the subjects' symptoms.

Infrared Thermal Imaging

Using infrared thermal imaging (FLIR ONE Pro, FLIR Systems, Inc) measurement method: maintain indoor temperature at 26 ℃±2 ℃, humidity at 45%-60%, maintain stable indoor air without convection, and avoid direct wind blowing to the examination site or between the subject and the camera. The subject shall not take any medication that affects body temperature orally within 24 hours, shall not undergo physical therapy, and shall not drink alcohol. The subject should lie prone, rest for a moment, maintain a certain balance in body temperature, wait for the skin to dry and become sweat-free, and then expose the tested area for 5 minutes before taking photos. The corresponding colors on the infrared image from high temperature to low temperature are white, deep red, red, light red, yellow, green, light green, light blue, deep blue, and black. Identify the pain points in the patient's area and mark them accordingly. Measure the temperature of the pain points in the pain area before enrollment, at week 1, at week 2, and after week 4 of treatment. The location of the pain points in the pain area is determined based on palpation, patient feedback, and the first infrared thermography image. A low temperature in the pain area before treatment indicates poor local blood circulation and a decreased pain threshold. A high temperature at pain points after treatment indicates good local blood circulation and an increased pain threshold. The testing unit we used was ℃.

Muscle Tension Test

We use a muscle hardness tester (OE-220 from Ito Ultra Short Wave Co, Ltd)×1 unit. Our test method is that the subject lies prone, fully exposing the skin on the waist and back, and naturally relaxes their hands, placing them on both sides of the body. Mark 8 points, which are 1.5 and 3 cun to the left and right of the L2 spinous process, and 1.5 and 3 cun left and right of the L4 spinous process. Press vertically and slowly on the surface of the measured muscle to prompt the patient to stop pressing. The obtained data will be displayed on the display screen. If the data are correct, press the record button to record the data. After 3 measurements, the display screen directly shows the average test value. The range of values is 0 - 99.99, where a small value represents high muscle fiber elasticity and low muscle tension. In contrast, a significant value represents low muscle fiber elasticity and high muscle tension, expressed as a percentage of the total muscle tension. Of measurement frequency, the two groups were measured before enrollment, at week 1, at week 2, and after treatment at week 4.

Tenderness Score

A normal score of 0 is defined as no significant pain during heavy pressure or maximum activity.

A mild score of 1 point is defined as when under heavy pressure, the patient reports pain, but passive activity is not restricted as such.

A moderate score of 2 is defined as when under heavy pressure, the patient reports pain and frowns, indicating discomfort, with mild restricted mobility.

A severe score of 3 points is defined as when under heavy pressure, the patient reports pain and withdrawal, and passive activity is severely restricted.

For measurement frequency, the 2 groups were measured before enrollment, at week 1, at week 2, at week 4, and after treatment at week 8.

Adherence

At the start of this study, a group of individuals with an interest in Tuina treatment of NSLBP will be evaluated to assess their suitability for participation in the experiment. The assessment will involve collecting their medical history, conducting specialized examinations, and administering auxiliary tests under the supervision of doctors. Participants can choose whether to participate after fully understanding the grouping and procedures. Researchers will complete standardized training courses and regularly review the course videos to monitor the accuracy of the intervention measures. Before and after each treatment, a dedicated research assistant is responsible for evaluating and recording data, which is recorded in the case report form (CRF). Any exceptional circumstances that arise during treatment should be promptly reported to the Tuina doctor and documented in the CRF.

Safety Evaluation

The category and severity of adverse events (AEs) after each intervention are recorded in detail on the CRF, including drug allergies, subcutaneous bleeding, joint injuries, hypertension, headache, dizziness, tinnitus, fever, and other relevant conditions. The incidence of AEs will be reported as the proportion of AEs to the total number of treatments. If the AE does not improve after sufficient rest, participants should receive timely medical treatment. Suppose a serious AE occurs, the chief investigator and the ethics committee will be well-informed within 24 hours of the event. They will be knowledgeable for additional assessment and treatment and decide whether to terminate the trial. Researchers also continue to monitor patients who have experienced AE until the issue is settled, especially individuals who withdrew due to AE. If AE is proven to be related to the intervention measures in this study, participants will receive appropriate medical and economic compensation.

The medicated oil used in this study contains ingredients with known warming or irritant properties (methyl salicylate, menthol, and processed *Aconitum* species). A multilevel monitoring strategy has been implemented to ensure participant safety. First, FEMO is a government-approved proprietary

XSL•FO
RenderX

Chinese medicine manufactured under good manufacturing practice standards, with strict control of potentially toxic constituents, and is approved exclusively for external use.

During each treatment session, participants are actively observed by trained Tuina physicians for any signs of abnormal local skin reactions (excessive erythema, burning sensation, pain, or discomfort at the application site). The observations are systematically documented using tenderness scores, muscle tension tests, and structured patient feedback and are collected at predefined time points.

Infrared thermal imaging is used as an objective tool to monitor local skin temperature changes. Moderate temperature elevation is considered a physiological response reflecting improved microcirculation; disproportionate or excessive temperature increases may indicate abnormal inflammatory or irritant responses.

All AEs, particularly skin-related reactions, are recorded in detail after each intervention. Participants are instructed to report any discomfort immediately. If adverse reactions occur, appropriate medical evaluation and intervention are initiated, and events are followed until resolution. Serious AEs are reported to the principal investigator and ethics committee within 24 hours, in accordance with the predefined safety management protocol.

Quality Control

Quality control will be conducted under the guidance of an independent oversight committee to eradicate potential biases and guarantee the quality of the experiment. The steering committee comprises investigators with expertise in Tuina trials, statistical design, medication, and AEs. They are accountable for monitoring data, identifying issues, reviewing collected data, and controlling biases. The protocol will remain unchanged during the 8-week treatment period, to ensure consistency in delivery.

All physicians and evaluators are required to complete standard training. Additionally, experienced Tuina physicians will be required to complete standard training courses with nationwide recognized academic experts who have 5 - 8 years of clinical experience in Tuina before the trial commences. The treatment plan should strictly follow the protocol of Tuina treatment steps, including the selection of Tuina techniques, intensity, time, sequence, and NSLBP protection measures.

Randomly assign to 2 groups simultaneously to avoid differences in the degree of NSLBP due to treatment physician and allocation bias. Participants need to cooperate in completing the entire treatment plan and follow-up.

Data Collecting and Monitoring

Clinical trial data will be obtained through CRFs, paper questionnaires, or online data collection applications and stored securely. Confidentiality will be maintained by assigning an encrypted numerical code to each individual. The accuracy of data input has been double-checked. The data will be locked and analyzed by 2 statisticians under the supervision of a third statistician and administrator. Then, record this trial's data on the subwebsite of the Chinese Clinical Trial Center [34], the electronic data management system. After the data input is completed, the electronic database will be closed. Management personnel must report weekly to the steering committee on the progress of data monitoring, including the accuracy and reliability of the data. Experimental audits encompass regular, independent reviews of core experimental processes and documents, including data and scheduling monitoring, to ensure the integrity of experimental results. Data monitoring is the responsibility of the Research Department at Shanghai Municipal Hospital of TCM. The supervisor will verify that all AEs are recorded in the correct format and that they are consistent with the protocol definition. Monitoring personnel will verify the following variables for all patients: name abbreviation, date of birth, gender, signed informed consent form, eligibility criteria, randomization date, treatment allocation, AEs, and end point. The arrangement of monitoring visits will depend on the patient's registration status, the site's status, and other commitments. Investigators must meet with monitors at all times. If problems are discovered during the visit (ie, insufficient personnel for research and lack of research documents), the supervisor will assist in resolving the problem on-site. We will access electronic monitoring to review source files and confirm AEs.

Statistical Analysis

Although the sample size calculation was based on a simplified between-group comparison at the primary postintervention time point (wk 4), the primary outcome analysis will use a longitudinal modeling framework to account for repeated measurements appropriately. A linear mixed-effects model will be used, including fixed effects for treatment group, time, and group-by-time interaction, with a random intercept specified for each participant. An unstructured covariance matrix will be prespecified to model within-subject correlations across repeated VAS measurements (baseline, and wk 1, 2, 4, and 8). Sensitivity analyses using alternative covariance structures, including first-order autoregressive and compound symmetry matrices, will be conducted to assess the robustness of the results. The data analysis was performed using SPSS (version 27.0) statistical software, and the testing criteria were as follows: α=.05, with P<.05 considered statistically significant.

Given the longitudinal design with repeated measurements collected at 5 time points (baseline, and wk 1, 2, 4, and 8), all primary and secondary continuous outcomes will be analyzed rather than with independent *t* tests or simple ANOVA, which accounts for within-subject correlation and avoids violations of the independence and sphericity assumptions inherent in traditional methods.

In each linear mixed-effects model, fixed effects will include treatment group (TNO vs TNW), time (modeled as a categorical variable), and the group-by-time interaction, with a random intercept specified for each participant. An unstructured covariance matrix will be prespecified for the repeated measures, with sensitivity analyses using alternative covariance structures (autoregressive and compound symmetry) to assess robustness. The primary estimand is the between-group difference in change from baseline to week 4 (end of treatment), which is defined a

XSL•FO
RenderX

priori as the primary time point based on the sample size calculation.

To control multiplicity, a hierarchical testing strategy will be applied. The primary end point (change in VAS from baseline to wk 4) will be tested at a 2-sided α level of .05. Secondary outcomes (JOA score, tenderness score, muscle tension, and infrared thermography) will be tested sequentially only if the preceding end point reaches statistical significance, thereby controlling the family-wise error rate. Longitudinal group-by-time interactions across all postbaseline time points will be considered exploratory.

Clinically meaningful improvement will be evaluated using responder analyses in addition to continuous outcomes. A reduction of ≥2 points on the VAS and an improvement of ≥4 points on the JOA score will be defined a priori as the minimal clinically significant difference. The proportions of responders will be compared between groups using generalized estimating equations, with effect sizes reported as risk ratios and 95% CIs.

Analysis of the Primary Outcome

The primary outcomes will be evaluated at baseline, intervention 1 week, 2 weeks, 4 weeks, and follow-up. We will conduct a sensitivity analysis on missing data (intention-to-treat analysis). The primary end point is the average change in the VAS scale from baseline to week 4 (primary end point) and week 8 (durability assessment). The baseline covariates include height, gender, age, cardiovascular history, and weight. The entire outcome scale will be measured at baseline. When the distribution is normal, nonpaired *t* tests are used to compare 2 groups, and 1-way ANOVA and repeated measures ANOVA are used to compare multiple groups. Sensitivity analysis using multiple input techniques (chain equation inputs) will be included to evaluate the impact of missing data on the main results. In addition, we will compare participants randomly assigned to the TNO and TNW groups, who received treatment considered to have high fidelity allocation. The percentage of missing data is 20%, and 5 input datasets will be used for analysis, which is the recommended minimum quantity. Before conducting any analysis, investigate any missing data patterns, identify the reasons for the missing data, and investigate the patterns to determine the validity of the hypothesis.

Analysis of Secondary Outcomes

Refer to the primary outcome analysis and conduct a secondary analysis of the outcomes. We will test the interarm differences for each specified end point. AE analysis was performed using the chi-square or Fisher exact test.

Results

This trial commenced in June 2023 and is expected to conclude in September 2025. As of June 2025, the initial steps have been completed. Specifically, the ethical review and clinical trial registration are finalized. Participant recruitment is progressing well, with 100 individuals enrolled so far. Data collection is ongoing, and formal data analysis has not yet been initiated ().

Discussion

Low back pain is the most prevalent chronic condition among the working population and a leading cause of disability and absenteeism. More than 80% of people will experience back pain at some point in their lifetime [35]. Studies have shown that low back pain can significantly impact daily activities, work performance, and sexual function [11,36]. Patients with low back pain frequently experience degeneration in the structure of the lumbar paravertebral muscles due to protective postures or muscle disuse due to the pain, along with increased inflammatory mediators.

Studies have demonstrated aerobic metabolic adaptations during trunk stretching in patients with chronic low back pain, who exhibit reduced mechanical efficiency, muscle endurance, and strength compared to healthy individuals [37]. The current study used a muscle hardness tester to monitor changes in muscle tension in patients with NSLBP before and after Tuina treatment combined with medicated oil. Current evidence suggests a correlation between trunk muscle strength, lumbar spine mobility, and the risk of low back pain. It has been noted that decreasing muscle tension in the waist can effectively alleviate low back pain symptoms, enhance joint mobility, and improve daily functioning [38]. Some studies have found that massage therapy can effectively relax the muscles and intervertebral joints surrounding the lumbar spine, correct spinal dysfunction, alter spinal biomechanics, and promote recovery from lower back pain [23,39-41]. In addition to patient-reported pain and functional outcomes, the current study incorporates infrared thermography and muscle hardness as objective secondary measures to enhance the clinical interpretability of treatment effects. In NSLBP, abnormal local skin temperature and increased paraspinal muscle stiffness are commonly associated with impaired microcirculation, persistent muscle spasm, and heightened nociceptive sensitivity. Improvements in skin temperature following intervention are interpreted as indicators of enhanced local blood flow and metabolic activity, which are physiologically linked to pain modulation and tissue recovery [42,43]. Similarly, reductions in muscle hardness reflect decreased muscle hypertonicity and improved elasticity, changes that are closely related to enhanced lumbar mobility and functional performance [37,44]. Although the measures of scales are not direct clinical end points, they provide objective evidence of underlying physiological recovery that supports and complements clinically meaningful improvements observed in pain intensity (VAS) and functional status (JOA). The multidimensional outcome framework helps bridge the gap between subjective symptom relief and objective biological change, thereby strengthening the clinical relevance and mechanistic interpretation of TNM in the management of NSLBP.

Infrared thermal imaging is widely used in cardiovascular medicine, sports medicine, ophthalmology, urology, and orthopedics [45-49] to visualize and observe temperature changes in skin, subcutaneous tissues, and blood vessels. The FLIR ONE Pro is a new type of infrared camera that is more convenient, affordable, and portable. A study has shown that people with lower back pain experience a significant increase

XSL•FO
RenderX

in local infrared temperature after massage therapy compared to before treatment [50]. However, no research uses infrared thermography to evaluate local microcirculation changes in patients with NSLBP treated with TNM. Studies indicate significant differences between infrared thermal imaging methods in the average and highest skin temperature parameter values, reminding us that it is crucial to unify equipment and parameters for the observation results in the next step of conducting multicenter clinical studies.

The American College of Physicians clinical practice guidelines recommend using nonpharmacological therapies to treat NSLBP [51]. Tuina has unique advantages as one of the traditional external treatments of TCM. Several studies have demonstrated the benefits of manual therapy for both acute and chronic lower back pain. It can relieve muscle spasms, promote local metabolism, absorb damaged tissue, and eliminate edema [52]. A systematic analysis shows moderate-quality evidence that spinal manipulative therapy has similar effects on short-term pain relief as other recommended treatments. Medicated oil is a kind of Tuina ointment preparation widely used in clinical practice. It is mainly used for bruises and various trauma and pain syndromes. Promoting blood circulation and relieving pain are common indications of preparations [53].

A study has found that using local rose essential oil has a therapeutic effect on women with pregnancy-related lower back pain. After 4 weeks of treatment, compared with the base oil group and the no-intervention group, the rose essential oil group showed a significant reduction in pain intensity and improvement in daily activities [38]. Cayenne, devil's claw, willow bark, comfrey, Brazilian arnica, and lavender essential oil have limited evidence of efficacy but suggest a possible benefit [54,55].

Some implementation difficulties need to be addressed in the research plan; we have designed corresponding measures to address these issues. First, to objectively evaluate the therapeutic effect more accurately, we used a muscle tension tester and an infrared thermal imaging device to assess the efficacy of TNM in treating patients with NSLBP. Under the premise of applying a unified standard Tuina method, we compared the effects of medicated oil and water on the experiment. Then, to promote and use the standardized technology of NSLBP clinical massage, a preliminary project launch meeting will be held to discuss the treatment plan with multiple experts in the field, simplifying the process to make it highly operable. To ensure the accuracy of the experimental intervention, all personnel involved in the project, including those implementing the treatment and evaluating the results, have received separate training. Each research group member is clear about their responsibilities, cooperates effectively, and completes the project on time. We standardized the appearance, specifications, and odor of the bottles to eliminate the placebo effect and compared them with water and activating oil, which will benefit clinical promotion and application.

The results are evaluated in 2 stages: the effects of lower back pain, motor function, muscle hardness, and temperature on 2 groups of patients with NSLBP were examined during treatment. To observe changes in pain and lumbar function during the follow-up period to evaluate the dependence of pain and movement disorders on muscle function and temperature.

A limitation of this study is that longitudinal correlation parameters could not be incorporated into the initial sample size calculation due to the lack of multiple time points pilot data. However, empirical intraclass correlation coefficients derived from the completed trial will be reported to inform the design of future multicenter randomized studies.

Nevertheless, we acknowledge that partial unblinding of therapists or participants cannot be entirely excluded, which represents a standard limitation in manual therapy trials. The assessor's blinding, standardized intervention delivery, and inclusion of objective outcome measures (infrared thermography and muscle tension testing) are emphasized to mitigate potential bias due to expectations.

The combination of Tuina and medicated oil therapy for NSLBP is easy to operate. If time and space are limited, medicated oil can be used alone to apply to the waist. This study is a suitable option for patients with NSLBP who opt for conservative treatment, aiming to investigate the effects of Tuina Gao Mo on pain, functional activity, and emotional state. If this study's findings are compelling, this approach can be further clinically promoted and applied.

Acknowledgments

We acknowledge the Shanghai Municipal Hospital of Traditional Chinese Medicine project number K231201N. The authors thank all the investigators, study coordinators, study participants, and their families for their support of this trial.

Funding

This study was partially funded by Hong Kong Europharm Laboratoires Company Limited (E4-D23044) in China. The sponsor, Europharm Laboratoires Company Limited, did not participate in this experiment, data collection, or statistical processing. The final test results and any report release are independent of the sponsor.

Authors' Contributions

MY, GG, SW, and GW had full access to all data in this study and were responsible for the integrity and accuracy of the data analysis. MY, GG, SW, and GW contributed equally to this work. MY, FY, GW, and HX handled the concept and design. All authors worked on the acquisition, analysis, or interpretation of the data. MY, GG, SW, and HX drafted the manuscript. All

XSL•FO
RenderX

authors critically revised the manuscript for important intellectual content. MY, GG, SW, and GW analyzed the statistics. GG and FY obtained the funding. MY and HX supervised this study.Fei Yao, PhD, Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, No. 274 Zhijiang Middle Road, Jing'an District, Shanghai, 200071,China; 86 137 6462 8380, doctoryaofei@126.com
Man Yuan, Guangxin Guo, Siyu Wang, and Guimao Wang contributed equally to this work and are co-first authors.

Conflicts of Interest

None declared.

References

1. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet 2017 Feb 18;389(10070):736-747. [doi: 10.1016/S0140-6736(16)30970-9] [Medline: 27745712]
2. Becker BA, Childress MA. Nonspecific low back pain and return to work. Am Fam Physician 2019 Dec 1;100(11):697-703. [Medline: 31790184]
3. Ma VY, Chan L, Carruthers KJ. Incidence, prevalence, costs, and impact on disability of common conditions requiring rehabilitation in the United States: stroke, spinal cord injury, traumatic brain injury, multiple sclerosis, osteoarthritis, rheumatoid arthritis, limb loss, and back pain. Arch Phys Med Rehabil 2014 May;95(5):986-995. [doi: 10.1016/j.apmr.2013.10.032] [Medline: 24462839]
4. Dieleman JL, Cao J, Chapin A, et al. US health care spending by payer and health condition, 1996-2016. JAMA 2020 Mar 3;323(9):863-884. [doi: 10.1001/jama.2020.0734] [Medline: 32125402]
5. Vos T, Abajobir AA, Abate KH, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep;390(10100):1211-1259. [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28711585]
6. Swain MS, Henschke N, Kamper SJ, Gobina I, Ottová-Jordan V, Maher CG. An international survey of pain in adolescents. BMC Public Health 2014 May 13;14(1):447. [doi: 10.1186/1471-2458-14-447] [Medline: 24885027]
7. Vlaeyen JWS, Maher CG, Wiech K, et al. Low back pain. Nat Rev Dis Primers 2018 Dec 13;4(1):52. [doi: 10.1038/s41572-018-0052-1] [Medline: 30546064]
8. Hartvigsen J, Hancock MJ, Kongsted A, et al. What low back pain is and why we need to pay attention. Lancet 2018 Jun 9;391(10137):2356-2367. [doi: 10.1016/S0140-6736(18)30480-X] [Medline: 29573870]
9. Bontrup C, Taylor WR, Fliesser M, et al. Low back pain and its relationship with sitting behaviour among sedentary office workers. Appl Ergon 2019 Nov;81:102894. [doi: 10.1016/j.apergo.2019.102894] [Medline: 31422243]
10. Evans R, Haas M, Schulz C, Leininger B, Hanson L, Bronfort G. Spinal manipulation and exercise for low back pain in adolescents: a randomized trial. Pain 2018 Jul;159(7):1297-1307. [doi: 10.1097/j.pain.0000000000001211] [Medline: 29596158]
11. Grabovac I, Dorner TE. Association between low back pain and various everyday performances : activities of daily living, ability to work and sexual function. Wien Klin Wochenschr 2019 Nov;131(21-22):541-549. [doi: 10.1007/s00508-019-01542-7] [Medline: 31493101]
12. Kong LJ, Fang M, Zhan HS, et al. Tuina-focused integrative chinese medical therapies for inpatients with low back pain: a systematic review and meta-analysis. Evid Based Complement Alternat Med 2012;2012:578305. [doi: 10.1155/2012/578305] [Medline: 23346207]
13. Tang S, Qian X, Zhang Y, Liu Y. Treating low back pain resulted from lumbar degenerative instability using Chinese Tuina combined with core stability exercises: a randomized controlled trial. Complement Ther Med 2016 Apr;25:45-50. [doi: 10.1016/j.ctim.2016.01.001] [Medline: 27062947]
14. Shi Y, Guo W, Yu L. Non-specific low back pain: interpretation of North American Spine Society (NASS) guidelines for evidence-based medicine. Zhongguo Xiu Fu Chong Jian Wai Ke Za Zhi 2021 Oct 15;35(10):1336-1340. [doi: 10.7507/1002-1892.202103131] [Medline: 34651490]
15. Daneau C, Cantin V, Descarreaux M. Effect of massage on clinical and physiological variables during muscle fatigue task in participants with chronic low back pain: a crossover study. J Manipulative Physiol Ther 2019 Jan;42(1):55-65. [doi: 10.1016/j.jmpt.2018.12.001] [Medline: 30955910]
16. Wen Yicheng LB. Clinical observation of 60 cases of sacroiliac joint induced lower back pain treated with meridian massage technique Chinese. J Ethn Folk Med 2023;32(24):71-74. [doi: 10.3969/j.issn.1007-8517.2023.24.zgmzmjyyzz2023240]
17. Savigny P, Watson P, Underwood M, Guideline Development Group. Early management of persistent non-specific low back pain: summary of NICE guidance. BMJ 2009 Jun 4;338:b1805. [doi: 10.1136/bmj.b1805] [Medline: 19502217]
18. Yabuki S, Ip AKK, Tam CK, et al. Evidence-based recommendations on the pharmacological management of osteoarthritis and chronic low back pain: an Asian consensus. Asian J Anesthesiol 2019 Jun 1;57(2):37-54. [doi: 10.6859/aja.201906_57(2).0003] [Medline: 31382324]
19. Nim CG, Kawchuk GN, Schiøttz-Christensen B, O'Neill S. Changes in pain sensitivity and spinal stiffness in relation to responder status following spinal manipulative therapy in chronic low back pain: a secondary explorative analysis of a randomized trial. BMC Musculoskelet Disord 2021 Jan 6;22(1):23. [doi: 10.1186/s12891-020-03873-3] [Medline: 33407345]

XSL•FO
RenderX

20. Kikuchi S. The recent trend in diagnosis and treatment of chronic low back pain. Spine Surg Relat Res 2017;1(1):1-6. [doi: 10.22603/ssrr.1.2016-0022] [Medline: 31440605]
21. Kongsted A, Kent P, Axen I, Downie AS, Dunn KM. What have we learned from ten years of trajectory research in low back pain? BMC Musculoskelet Disord 2016 May 21;17:220. [doi: 10.1186/s12891-016-1071-2] [Medline: 27209166]
22. He JY. A prospective multicenter randomized study on the combined treatment of cold-dampness type chronic non-specific low back pain with “wood lock medicated balm + multifunctional physiotherapy instrument”. : Global Thesis; 2024 URL: https://www.globethesis.com/?t=2544306917971669 [accessed 2026-02-05]
23. Kong LJ, Fang M, Zhan HS, et al. Chinese massage combined with herbal ointment for athletes with nonspecific low back pain: a randomized controlled trial. Evid Based Complement Alternat Med 2012;2012:695726. [doi: 10.1155/2012/695726] [Medline: 23258996]
24. Rubinstein SM, de Zoete A, van Middelkoop M, Assendelft WJJ, de Boer MR, van Tulder MW. Benefits and harms of spinal manipulative therapy for the treatment of chronic low back pain: systematic review and meta-analysis of randomised controlled trials. BMJ 2019 Mar 13;364:l689. [doi: 10.1136/bmj.l689] [Medline: 30867144]
25. Chen ZH, Zheng QK, Chen SJ, Li X, Lian XW. Warming acupuncture combined with “three steps and seven methods” of tuina for chronic nonspecific low back pain of yang deficiency and cold-dampness blockage: a randomized controlled trial. Zhongguo Zhen Jiu 2022 May 12;42(5):505-510. [doi: 10.13703/j.0255-2930.20210622-k0001] [Medline: 35543940]
26. Ma Q, Zhou X, Liang H, et al. Tuina versus physiotherapy or a combination of both for the management of chronic nonspecific low back pain: a randomized controlled trial. Complement Ther Med 2024 Nov;86:103098. [doi: 10.1016/j.ctim.2024.103098] [Medline: 39395482]
27. Wang S, Yuan M, Guo G, et al. Evaluation of Tuina combined with medicated oil in non-specific low back pain treatment: a randomized clinical trial. J Pain Res 2025;18:5781-5796. [doi: 10.2147/JPR.S554329] [Medline: 41209746]
28. Dai T, Wang Y, Zhang W, et al. Biologically active guanidine alkaloids. Acta Mater Med 2025;4(4):575-611. [doi: 10.15212/AMM-2024-0088]
29. Zhibin X. Preparation, quality analysis, and activity evaluation of an anti-inflammatory and analgesic traditional Chinese medicine called medicated oil. 2014.
30. Kreiner DS, Matz P, Bono CM, et al. Guideline summary review: an evidence-based clinical guideline for the diagnosis and treatment of low back pain. Spine J 2020 Jul;20(7):998-1024. [doi: 10.1016/j.spinee.2020.04.006] [Medline: 32333996]
31. Consensus among Chinese experts in the diagnosis and treatment of acute/chronic non-specific low back pain. Chin J Spinal Spinal Cord 2016;26(12):1134-1138. [doi: 10.3969/j.issn.1004-406X.2016.12.16]
32. He S, Renne A, Argandykov D, Convissar D, Lee J. Comparison of an emoji-based visual analog scale with a numeric rating scale for pain assessment. JAMA 2022 Jul 12;328(2):208-209. [doi: 10.1001/jama.2022.7489] [Medline: 35819433]
33. Nakamura M, Miyamoto K, Shimizu K. Difference in evaluation of patients with low back pain using the Japanese Orthopaedic Association Score for Back Pain and the japanese version of the Roland-Morris Disability Questionnaire. J Orthop Sci 2009 Jul;14(4):367-373. [doi: 10.1007/s00776-009-1348-5] [Medline: 19662468]
34. ResMan. URL: http://www.medresman.org.cn/login.aspx [accessed 2026-01-29]
35. Freburger JK, Holmes GM, Agans RP, et al. The rising prevalence of chronic low back pain. Arch Intern Med 2009 Feb 9;169(3):251-258. [doi: 10.1001/archinternmed.2008.543] [Medline: 19204216]
36. Pieber K, Stein KV, Herceg M, Rieder A, Fialka-Moser V, Dorner TE. Determinants of satisfaction with individual health in male and female patients with chronic low back pain. J Rehabil Med 2012 Jul;44(8):658-663. [doi: 10.2340/16501977-1010] [Medline: 22729793]
37. Anthierens A, Thevenon A, Olivier N, Mucci P. Paraspinal muscle oxygenation and mechanical efficiency are reduced in individuals with chronic low back pain. Sci Rep 2024 Feb 28;14(1):4943. [doi: 10.1038/s41598-024-55672-8] [Medline: 38418858]
38. Shirazi M, Mohebitabar S, Bioos S, et al. The effect of topical Rosa damascena (Rose) oil on pregnancy-related low back pain: a randomized controlled clinical trial. J Evid Based Complementary Altern Med 2017 Jan;22(1):120-126. [doi: 10.1177/2156587216654601] [Medline: 27335145]
39. Paige NM, Miake-Lye IM, Booth MS, et al. Association of spinal manipulative therapy with clinical benefit and harm for acute low back pain: systematic review and meta-analysis. JAMA 2017 Apr 11;317(14):1451-1460. [doi: 10.1001/jama.2017.3086] [Medline: 28399251]
40. Mo Z, Li D, Zhang R, Chang M, Yang B, Tang S. Comparisons of the effectiveness and safety of Tuina, acupuncture, traction, and Chinese herbs for lumbar disc herniation: a systematic review and network meta-analysis. Evid Based Complementary Alternat Med 2019;2019:6821310. [doi: 10.1155/2019/6821310] [Medline: 31015852]
41. Chen XM, Wen Y, Chen S, et al. Traditional Chinese manual therapy (Tuina) reshape the function of default mode network in patients with lumbar disc herniation. Front Neurosci 2023;17:1125677. [doi: 10.3389/fnins.2023.1125677] [Medline: 37008205]
42. Nahm FS. Infrared thermography in pain medicine. Korean J Pain 2013 Jul;26(3):219-222. [doi: 10.3344/kjp.2013.26.3.219] [Medline: 23861995]
43. Lubkowska A, Pluta W. Infrared thermography as a non-invasive tool in musculoskeletal disease rehabilitation—the control variables in applicability—a systematic review. Appl Sci (Basel) 2022;12(9):4302. [doi: 10.3390/app12094302]

XSL•FO
RenderX

44. Wu Z, Ye X, Ye Z, et al. Asymmetric biomechanical properties of the paravertebral muscle in elderly patients with unilateral chronic low back pain: a preliminary study. Front Bioeng Biotechnol 2022 Feb 10;10:814099. [doi: 10.3389/fbioe.2022.814099]
45. Zhu W, Yang Y, Jiang J, et al. Value of the combination of a smartphone-compatible infrared camera and a hand-held doppler ultrasound in preoperative localization of perforators in flaps. Heliyon 2023 Jun;9(6):e17372. [doi: 10.1016/j.heliyon.2023.e17372]
46. Machado Á, Priego-Quesada JI, Jimenez-Perez I, Gil-Calvo M, Carpes FP, Perez-Soriano P. Influence of infrared camera model and evaluator reproducibility in the assessment of skin temperature responses to physical exercise. J Therm Biol 2021 May;98:102913. [doi: 10.1016/j.jtherbio.2021.102913] [Medline: 34016340]
47. Antwi A, Nti AN, Ritchey ER. Thermal effect on eyelid and tear film after low-level light therapy and warm compress. Clin Exp Optom 2024 Apr;107(3):267-273. [doi: 10.1080/08164622.2023.2206950] [Medline: 37156225]
48. Brooks JP, Brooks JM, Seals T. Smartphone thermal imaging in the detection of testicular ischemia. Urology 2021 Nov;157:233-238. [doi: 10.1016/j.urology.2021.02.031] [Medline: 33689763]
49. Ahn SM, Chun JH, Hong S, et al. The value of thermal imaging for knee arthritis: a single-center observational study. Yonsei Med J 2022 Feb;63(2):141-147. [doi: 10.3349/ymj.2022.63.2.141] [Medline: 35083899]
50. Yan Binhong NK. Effects of traditional Chinese medicine Tuina on lumbar muscle tissue temperature and serum 5-HT and β - EP in LBP patients. J Hainan Med Coll 2022;28(3):204-208. [doi: 10.13210/j.cnki.jhmu.20210919.001]
51. Qaseem A, Wilt TJ, McLean RM, et al. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017 Apr 4;166(7):514-530. [doi: 10.7326/M16-2367] [Medline: 28192789]
52. Will JS, Bury DC, Miller JA. Mechanical low back pain. Am Fam Physician 2018 Oct 1;98(7):421-428. [Medline: 30252425]
53. Farley J, Taylor-Swanson L, Koppenhaver S, Thackeray A, Magel J, Fritz JM. The effect of combining spinal manipulation and dry needling in individuals with nonspecific low back pain. J Pain 2024 Aug;25(8):104506. [doi: 10.1016/j.jpain.2024.03.002] [Medline: 38484853]
54. Chou R. Low back pain. Ann Intern Med 2021 Aug;174(8):ITC113-ITC128. [doi: 10.7326/AITC202108170] [Medline: 34370518]
55. Smith CA, Levett KM, Collins CT, Dahlen HG, Ee CC, Suganuma M. Massage, reflexology and other manual methods for pain management in labour. Cochrane Database Syst Rev 2018 Mar 28;2018(3):29589380. [doi: 10.1002/14651858.CD009290.pub3]

Abbreviations

AE: adverse event
CRF: case report form
FEMO: Flying Eagle Wood Lok Medicated Oil
JOA: Japanese Orthopaedic Association scores
NSLBP: nonspecific low back pain
TCM: traditional Chinese medicine
TNO: Tuina combined medicated oil
TNW: Tuina combined water
VAS: visual analog scale

Edited by J Sarvestan; submitted 06.Aug.2025; peer-reviewed by TTN Hang Nguyen; revised version received 22.Dec.2025; accepted 31.Dec.2025; published 11.Mar.2026.

Please cite as:
Yuan M, Guo G, Wang S, Yao F, Wang G, Xu H
Clinical Efficacy Evaluation of Tuina Combined Medicated Oil in the Treatment of Nonspecific Low Back Pain: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e81887
URL: https://www.researchprotocols.org/2026/1/e81887
doi:10.2196/81887

© Man Yuan, Guangxin Guo, Siyu Wang, Fei Yao, Guimao Wang, Hongxi Xu. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 11.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Elastic Tape Associated With Pulmonary Rehabilitation in Male Individuals With Chronic Obstructive Pulmonary Disease: Protocol for a 2-Arm, Assessor-Blinded Randomized Controlled Trial

Thiago Fernandes Pinto[1*], PT, MSc; Juliana de Melo Batista dos Santos[1,2*], PE, PhD; Estéfane Caroline Monteiro Reis[1*], PT; Cibele C Berto Marques da Silva[1*], PT, MSc; Fabiano Francisco de Lima[1*], PT, PhD; Regina Maria Carvalho-Pinto[1,3*], MD, PhD; Celso Ricardo Fernandes Carvalho[1,3*], PT, PhD

[1]Department of Physical Therapy, Medical School, University of São Paulo, Rua Cipotânea, 51, Cidade Universitária, São Paulo, Brazil

[2]Health Sciences Post-Graduation Program, Santo Amaro University - UNISA, São Paulo, Brazil

[3]Pulmonary Division, Heart Institute (InCor), Clinics Hospital, Medical School, University of São Paulo, São Paulo, Brazil

[*]all authors contributed equally

Corresponding Author:
Celso Ricardo Fernandes Carvalho, PT, PhD
Department of Physical Therapy, Medical School, University of São Paulo, Rua Cipotânea, 51, Cidade Universitária, São Paulo, Brazil

Abstract

Background: Individuals with severe chronic obstructive pulmonary disease (COPD) may exhibit thoracoabdominal asynchrony, which reduces ventilatory efficiency. A novel intervention using elastic tape (ET) applied to the chest wall has been shown to acutely reduce thoracoabdominal asynchrony and dyspnea during exercise among individuals with COPD. We hypothesize that using ET in pulmonary rehabilitation (PR) may increase the benefits of PR in this population.

Objective: This study aims to evaluate the additional effects of ET on exercise capacity, symptoms of anxiety and depression, health-related quality of life, and physical activity in daily life among male individuals with moderate to very severe COPD who are undergoing PR.

Methods: This is a protocol for a randomized, controlled, 2-arm, parallel, blinded assessor clinical trial. Individuals will be followed for 8 weeks, twice a week, with PR sessions lasting approximately 1 hour. Health status (COPD Assessment Test), health-related quality of life (Chronic Respiratory Questionnaire), and psychological distress (Hospital Anxiety and Depression Scale) will be assessed before and after the intervention. Then, exercise capacity will be assessed via the incremental shuttle walking test and endurance shuttle walking test, and participants will use a triaxial accelerometer (ActiGraph) for 7 days to assess physical activity in daily life. Subsequently, individuals will be randomized into ET or sham groups; both groups will complete a PR program (2 times per week for 8 weeks). The ET group will receive applications of ET, whereas the sham group will receive a nonelastic tape. Data will be presented as means and SDs or medians and IQRs. Intergroup comparisons will be performed using a 2-way ANOVA, followed by the Bonferroni post hoc correction test, or the Kruskal-Wallis test, followed by the Dunn post hoc test. The threshold for statistical significance will be set at 5%.

Results: The clinical trial registration was approved in June 2023. Recruitment and data collection for the trial are ongoing; as of November 2025, a total of 10 individuals have been recruited, and the results are expected to be available by the end of November 2026.

Conclusions: We hypothesize that the use of ET can enhance the benefits of PR in individuals with moderate to very severe COPD and increase exercise capacity and quality of life, as well as reduce symptoms of anxiety and depression.

Trial Registration: ClinicalTrials.gov NCT05939999; https://clinicaltrials.gov/study/NCT05939999

International Registered Report Identifier (IRRID): DERR1-10.2196/75029

(JMIR Res Protoc 2026;15:e75029) doi:10.2196/75029

KEYWORDS

chronic obstructive pulmonary disease; COPD; adhesive elastic tape; pulmonary rehabilitation; dyspnea; exercise capacity; physical capacity

XSL•FO
RenderX

Introduction

Chronic obstructive pulmonary disease (COPD) is a preventable and treatable disease characterized by persistent and progressive airflow limitations associated with a chronic inflammatory response of the airways [1]. In the PLATINO (Projeto Latino-Americano de Investigação em Obstrução Pulmonar) study conducted in 2005, the prevalence of COPD was 15.8% (95% CI 13.5% - 18.1%) using the fixed ratio definition. COPD showed a positive association with age and smoking and an inverse association with BMI; thus, we understand that most individuals with COPD in the Brazilian population are not obese [2]. These findings were consistent with a study published in 2019 [3], which showed that individuals with moderate to very severe COPD had a slight reduction in BMI. The same 2019 study [3] showed that there was a similar prevalence for both sexes. Individuals with COPD experience destruction of the small airways and decreased elastic recoil, leading to air trapping at rest due to the loss of the intrinsic tendency of the lungs to deflate. Movements of the thoracic and abdominal compartments may be impaired and may become worse with exertion [4]. Therefore, the rib cage and abdomen may move uncoordinately; this phenomenon is called thoracoabdominal asynchrony (TAA) [5].

TAA occurs in individuals with severe airway obstruction, leading to a mechanical disadvantage, increased breathing work, dyspnea symptoms, limited exercise, and reduced physical activity in daily life (PADL) [6,7]. To our knowledge, few studies have found interventions to minimize TAA. Yamaguti et al [8] found that a diaphragmatic breathing retraining program could reduce TAA among individuals with COPD. Recently, Pinto et al [9] accomplished a novel intervention using elastic tape (ET) in the chest wall and demonstrated that ET reduces TAA, tidal volume, minute volume, and dyspnea during isoload exercise among individuals with severe to very severe COPD. Individuals also demonstrated improvements in maximal oxygen consumption during cardiopulmonary exercise testing and a reduction in sedentary time during daily life. More recently, Santos et al [10] used ET for 3 weeks in patients with moderate to very severe COPD and reported reduced dyspnea symptoms as well as improved health status and health-related quality of life (HRQoL). Although these studies used ET, neither did so in conjunction with pulmonary rehabilitation (PR), and both studies had an intervention exposure duration of a maximum of 21 days. Furthermore, there is no evidence of ET being used in PR as an adjuvant strategy.

PR is an effective and successful treatment for individuals with persistent symptoms and/or excessive use of health care resources [11]. The European Respiratory Society and the American Thoracic Society define PR as a comprehensive intervention based on a thorough patient assessment followed by patient-tailored therapies, which include but are not limited to exercise training, education, and behavior changes designed to improve the physical and psychological conditions of people with chronic respiratory disease and to promote long-term adherence to health-enhancing behaviors [12,13]. The core of PR is exercise training, which directly addresses deconditioning and its association with decreases in symptoms such as dyspnea and fatigue, which are the primary reasons for referral. Exercise training programs also have additional positive effects on comorbidities, HRQoL, and general health [13]. However, the additional effects of ET associated with PR have not been studied.

Although previous studies have investigated the use of ET in isolation [9,10], and considering that ET has been shown to reduce TAA [9], our hypothesis is that the application of ET during a PR program may confer additional positive effects in individuals with greater disease severity regarding exercise capacity, symptoms of anxiety and depression, and HRQoL in individuals with moderate to very severe COPD undergoing PR.

Methods

Study Design and Participants

This is a protocol for a randomized, controlled, 2-arm, parallel, assessor-blinded clinical trial. Individuals with COPD will be screened and recruited at the pulmonology outpatient clinic of a tertiary hospital. The study protocol was developed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist guidelines.

Analysis of the Population

The inclusion criteria will be as follows: diagnosis of moderate to very severe COPD [1], clinical stability (ie, no exacerbations for at least 30 days), male sex, nonobese status (BMI ≤29.9 kg/m^2), absence of musculoskeletal limitations, and no participation in another research protocol or a PR program in the past 6 months. The exclusion criteria will be decompensated cognitive, neurological, and/or cardiological dysfunction and receiving home oxygen therapy. An intention-to-treat analysis will be performed using the patient's most recent assessment in case of study dropout or absence of data and will be treated by imputing the last observation carried forward [14].

Ethical Considerations

Patients will be recruited from the pulmonology outpatient clinic of the Hospital das Clínicas de São Paulo, São Paulo, Brazil, and will be included in the study after signing the informed consent form. This study will be performed in accordance with the Declaration of Helsinki. The study was registered at ClinicalTrials.gov (NCT05939999), and the Hospital Research Ethics Committee approved the study (55617321.2.0000.0068). All participants will be informed about the possibility of withdrawing at any time without incurring any loss of treatment at the institution. Participants will receive an identification code according to the order of inclusion in the study. The data will be deidentified by excluding the main identifying elements such as name, address, date of birth, phone number, email, medical record number, institutional identification number, and social security number. In addition, participants will be informed about the voluntary nature of their participation; this means that there will be no compensation of any kind for participation.

Randomization and Allocation Concealment

The randomization will be generated using a computer (Sealed Envelope [15]). An investigator who is not involved in patient

XSL•FO
RenderX

recruitment, assessment, or intervention will perform group allocation. Group allocation will be concealed via consecutively numbered, sealed, opaque envelopes. After randomization, both groups (ET vs sham) will not be informed of their allocation, and the same instructions on perception of tape use will be given to both groups.

Adverse Events

The individual's skin condition will be assessed at every session during the PR program to immediately identify any allergic reactions to the tape in both groups. If any skin changes are identified, such as redness or itching, the tape will be removed immediately and will not be reapplied until the condition has been completely resolved. If the condition does not resolve within 72 hours, even after removing the tape, the individual will be referred for evaluation by a dermatologist and will be removed from the study.

Experimental Design

Clinical and anthropometric data will be collected during the participant interview. Participants will complete questionnaires to assess health status (COPD Assessment Test [CAT]), HRQoL (Chronic Respiratory Questionnaire [CRQ]), and anxiety and depression symptoms (Hospital Anxiety and Depression Scale [HADS]). Then, a complete pulmonary function test (body plethysmography), the incremental shuttle walking test (ISWT), and the endurance shuttle walking test (ESWT) will be carried out. Finally, individuals will use a triaxial accelerometer (ActiGraph GT9X) for 7 days to assess their PADL. Lastly, individuals will be randomly assigned to 1 of 2 groups: the ET group or the sham group.

At baseline and postintervention, assessments will be performed by a researcher who is not involved in individual recruitment or intervention. After inclusion, individuals will begin an 8-week PR program twice a week (Figure 1). Individuals in the ET group will have the ET placed before the beginning of the PR program, and the ET will be replaced every 7 days throughout the 8-week period. The sham group will receive a nonelastic tape of the same color and size at the same position in the chest wall at the beginning of the PR program, and the tape will be replaced every 7 days, similar to the ET group. At the end of the PR program, participants will repeat the same assessments, except for the ISWT. The ESWT will be performed at the same level established by the ISWT during the initial evaluation (Figure 2).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) 2025 flow diagram showing the enrollment, allocation, follow-up, and analysis phases of the randomized controlled trial. Participants were assessed for eligibility, randomized into 2 groups—elastic tape+pulmonary rehabilitation (intervention group [IG]) or pulmonary rehabilitation alone (control group [CG])—and reevaluated after 8 weeks. The number of patients shown in the figure is the planned number based on the sample size. Outcome assessments included the Hospital Anxiety and Depression Scale (HADS), Chronic Respiratory Questionnaire (CRQ), and COPD Assessment Test (CAT). ESWT: endurance shuttle walking test; ISWT: incremental shuttle walking test.

XSL•FO
RenderX

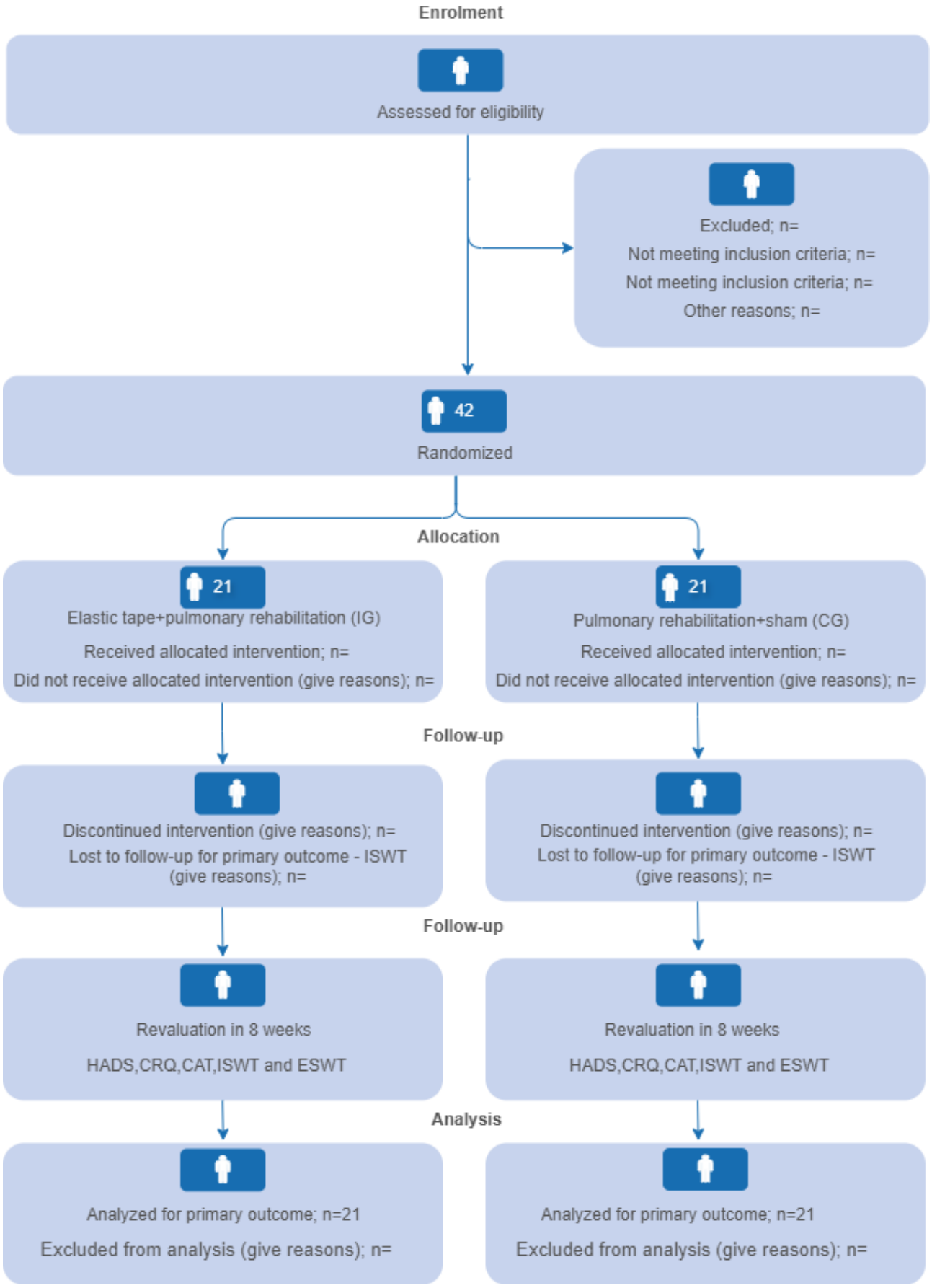

XSL•FO
RenderX

Figure 2. Schematic representation of the study design. Assessments were performed at baseline and after 8 weeks of intervention. 1RM: 1 repetition maximum test; CAT: COPD Assessment Test; CRQ: Chronic Respiratory Questionnaire; ET: elastic tape; ESWT: endurance shuttle walking test; HADS: Hospital Anxiety and Depression Scale; ISWT: incremental shuttle walking test; PAL: physical activity level; PR: pulmonary rehabilitation; PFT: pulmonary function test; V: visit.

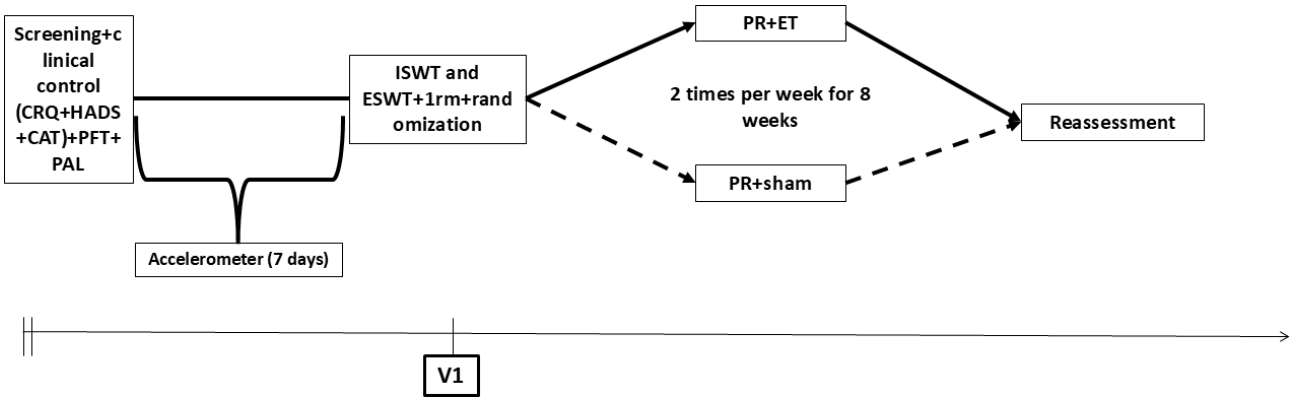

Intervention

ET Overview

The ET (Kinesio Brand Tape, Tex Gold), which is made of cotton and elastic material and is skin-colored and waterproof, will be applied by the same trained professional throughout the study to avoid intervention fidelity bias, with the participant in a relaxed, supine position. The individual will lie down on a stretcher so that their torso is extended, thus facilitating the application of the ET [9]. The ET will be positioned at the origins and insertions of the following muscle groups: rectus abdominis, internal oblique, and internal intercostal (Figure 3). These muscle groups were chosen because they are the main muscles involved in forced expiration during exercise, and the tape will always be placed with the maximum possible tension, which is 200% of its original size. The tape will be placed at the beginning of the session and replaced every 7 days. If at least 2 cm of tape is detached, the ET will be replaced regardless of the length of time it has been in place.

Figure 3. Illustration of the elastic tape placement protocol. Tape was applied to key expiratory muscle groups—rectus abdominis, internal oblique, and intercostal muscles—using standardized positioning and tension to ensure reproducibility and comfort during the pulmonary rehabilitation sessions.

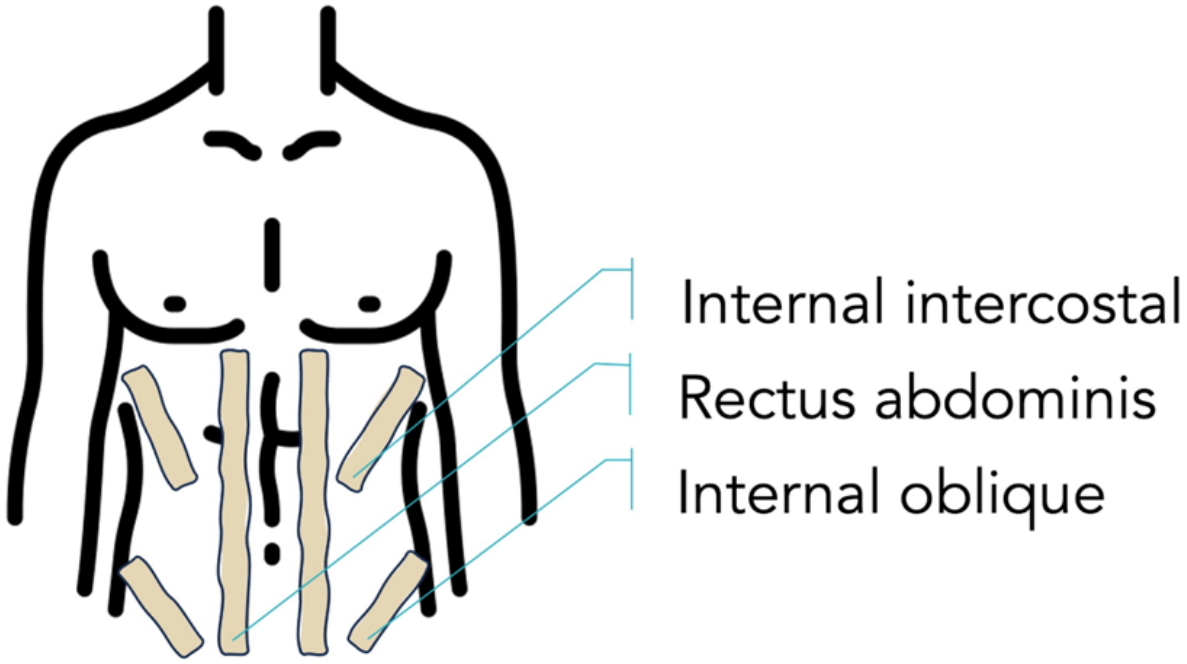

Sham Group

The sham group will receive a nonelastic tape (3M Micropore) placed using an identical procedure to the ET and replaced every 7 days. If at least 2 cm of tape is detached, the tape will be replaced regardless of the length of time it has been in place.

PR Program

On the first day, at the start of PR, individuals will be briefed on the benefits of PR, such as reduced dyspnea, additional positive effects on comorbidities, improved HRQoL and overall health, and increased exercise capacity [13]. Additionally, a WhatsApp contact will be provided on the same day for ongoing communication regarding health-related issues. Both groups will undergo an 8-week PR program, including 1-hour sessions

XSL•FO
RenderX

twice a week, supervised by a physiotherapist with experience in PR, carried out in the hospital's outpatient clinic. The program will be individualized and divided into 30 minutes of aerobic training on an ergometric treadmill (Technogym Jog 500) and 30 minutes of resistance training using free weights (dumbbells and barbells) for the upper limbs and gym machines for the lower limbs (LLs). Aerobic training will be performed with a heart rate (HR) corresponding to 50% to 80% of the maximum HR according to the formula proposed by Karvonen et al ($HR_{training}$ =HR_{rest}+ [HR_{max}–HR_{rest}]×% intensity), where $HR_{training}$ is training heart rate, HR_{rest} is resting heart rate, HR_{max} is maximum heart rate, and % intensity is the percentage of training intensity. The intensity will be increased weekly by 5% of the maximum HR until reaching 80% of the maximum HR. Furthermore, this progression of treadmill load will also be guided by the scale of subjective perception of exertion (modified Borg scale) [16]. The training will be carried out in such a way that during each session, the participant must remain on the Borg scale between scores 4 and 6 [17]. The increase in intensity by 5% of the maximum HR will be adjusted after 2 sessions and/or if the score on the Borg scale is less than 4. If the individual reports fatigue greater than 6 on the modified Borg scale, the intensity will be reduced and, in the next session, the increase in intensity will be tried again. The individual may interrupt the activity if they feel symptoms or intense respiratory discomfort, resuming it after improvement. HR, peripheral oxygen saturation (SpO_2), and the subjective perception of exertion will be monitored during all exercise. Oxygen supplementation will be provided as needed to maintain SpO_2 of 90% or greater [17].

Resistance training will consist of exercises for the upper limbs through lifting weights with dumbbells and for the LLs via weight machines. Large muscle groups will be trained, including the biceps, deltoids, quadriceps, and hamstrings. Training will begin with 3 sets of 8 repetitions of each exercise at a workload of 50% of the one-repetition maximum [18]. The number of repetitions will subsequently increase to 3 sets of 10 and, later, 12 repetitions. Load progression will be carried out when the individual tolerates performing more than 12 repetitions for each set without compensatory movements and/or symptoms of intolerance. The workload will be increased according to individual tolerance.

Resistance training for the upper limbs will include shoulder abduction and elbow flexion. Resistance training for the LLs will include knee extension and hip extension using "leg press" equipment.

No written material will be provided to either group. Transportation expenses will be subsidized for all individuals to reduce absenteeism and increase feasibility and adherence to the protocol. This financial resource will be funded for this research project. For analysis of the results, individuals must complete at least 80% (13 sessions) of the program, which will be monitored using the PR monitoring form, where data on HR, SpO_2, the Borg scale, speed, and treadmill incline will be tabulated every 5 minutes, from the beginning to the end of the training session for all sessions.

The acceptability of ET use will be monitored throughout the session based on a visual inspection performed by the physiotherapist at the beginning of the session and during ET replacement every 7 days to immediately identify any skin changes that the individual may experience due to continued ET use.

To ensure or increase adherence to the 8-week PR program, the importance and potential benefits of PR, as established in the literature, will be explained to each participant individually, regardless of the group they are allocated to after randomization. Throughout the PR program, individuals will have a WhatsApp contact available to clarify and assist with any questions regarding their COPD-related health status. In cases of clinical worsening, the individual will be taken to the hospital emergency room for immediate medical care.

Assessment Tools

Primary Outcome Measure: Functional Exercise Capacity

The ISWT will be performed to assess functional exercise capacity. In this test, individuals will be instructed to walk in a 10 m corridor delimited by cones (with a 9 m distance between them), and the speed will be determined by a beep recorded on the sound box, which increases the speed by 0.17 m/s every minute, with a total of up to 15 stages. They will also be instructed to walk at a speed that allows them to reach the cones when the beep sounds, and at the end of each minute, an additional signal will be given to alert them to the increase in speed [19]. The test will end when the participant is more than 0.5 m away from the cone, reports chest pain, reports severe dyspnea, reports fatigue, reports exhaustion, or requests to stop. The distance walked at the end of the test will be registered for subsequent analyses [20]. Before and after the exercise tests, SpO_2 and HR will be assessed via a portable pulse oximeter positioned on the individual's dominant hand (WristOx2 3150; Nonin). Blood pressure (BP) will be assessed via a stethoscope (Littman) and a sphygmomanometer calibrated on the dominant arm (Aneroid Premium - G-Tech).

The criteria for stopping the test will include diastolic BP (DBP) above 140 mm Hg, sustained elevation of systolic BP (SBP) above 240 mm Hg, motor incoordination, feelings of imbalance, mental confusion, clinical manifestations of respiratory distress that worsen with increasing load, dyspnea disproportionate to the intensity of the effort, progressive claudication of the LL, and limb exhaustion [12]. Individuals will be instructed not to consume beverages containing caffeine for at least 12 hours before the test and not to perform intense exercise 24 hours before.

The ESWT will be performed 40 minutes after the end of the ISWT, using the same structure; that is, the volunteers will be instructed to walk in a 10m corridor limited by cones (with a distance of 9 m between them), and the speed will be determined by a sound signal recorded by the cell phone. In this case, the speed will be determined using 85% of the maximum speed reached in the ISWT and will be constant throughout the test [12]. Individuals will be monitored before and after the test,

XSL•FO
RenderX

and the interruption criteria will be the same as those used in the ISWT procedure.

Secondary Outcome Measures: Health-Related Quality of Life (Chronic Respiratory Questionnaire – CRQ)

This questionnaire is already validated and translated into Portuguese [21]. It consists of 20 questions divided into 4 domains: fatigue (4 questions), dyspnea (5 questions), emotional function (7 questions), and self-control (4 questions). The responses will be assessed via a Likert scale, with 7 alternatives ranging from "never" to "always," whereas the dyspnea domain includes an individualized assessment in which the patient chooses, from a list of 26 daily activities, the 5 activities that cause the most dyspnea and then rates each one again via a Likert scale with 7 alternatives ranging from "unbearable shortness of breath" to "no shortness of breath." The results will be expressed as the mean score of each domain and the total score, with higher scores indicating better HRQoL.

Anxiety and Depression Symptoms

The HADS will be used to assess anxiety and depression symptoms. The questionnaire was developed to identify symptoms of anxiety and depressive mood and has been translated, validated, and published in Portuguese by Botega et al [22]. It consists of 14 multiple-choice questions divided into 2 subscales, anxiety (HADS-A) and depression (HADS-D) symptoms, with 7 questions each. The domains are categorized by the severity of the stress present: none, 0 - 7; probable, 8 - 10; and present, ≥11 [23,24].

Health Status

Health status will be assessed by the CAT, a specific questionnaire for COPD that assesses the impact of the disease's symptoms [1]. It consists of 8 items related to health conditions and has been applied and validated in several languages, including Portuguese [25].

Sample Characterization

Initial Assessment

Individuals will undergo an initial interview to provide information on several aspects: identification and contact data, medical history (eg, depression, osteoarthritis or arthrosis, osteoporosis, diabetes mellitus, and arterial hypertension), and current medications.

Lung Function

Pulmonary function parameters will be obtained from the individual's medical records within the previous 6 months. The interpretation will be performed according to the standards of the American Thoracic Society and European Respiratory Society [26] and normal values relative to the Brazilian population [27].

Physical Activity in Daily Life

The PADL will be quantified using a triaxial accelerometer (GT9X; ActiGraph) for sample characterization. The equipment will be placed on the individual's waist, and the individual will be advised to wear it for 7 days and only remove it to sleep and shower. They will be instructed to maintain their lifestyle habits. Seven full days of monitoring will be considered to evaluate the PADL: 5 weekdays and 2 weekend days. Zero counting time (considered the period in which the participant did not use the accelerometer) or days with less than 8 hours of device use will be excluded [28]. The results will be expressed as the time of sedentary behavior, steps per day, and time spent in moderate to vigorous physical activity.

Statistical Analysis

Data normality will be assessed using the Shapiro-Wilk test, and homoscedasticity will be assessed using the Levene test. Continuous variables will be presented as mean (95% CI) or as medians and IQRs, according to their distribution. Intergroup comparisons will be performed using 2-way ANOVA for repeated measures with Bonferroni post hoc correction or the Kruskal-Wallis test with Dunn test when assumptions are violated. The proportion of individuals not reaching the minimal clinically important difference will be compared using the chi-square test. Within-group effect sizes (Cohen *d*) will be classified as small (0.2), moderate (0.5), or large (0.8). The Spearman correlation coefficient will be used to analyze the associations between changes in ESWT time and changes in CAT, HADS-A, HADS-D, and CRQ scores. Analyses will be conducted according to the intention-to-treat principle, using GraphPad Prism version 5.0, with significance set at $P<.05$.

Sample Calculation

The sample size calculation was based on the primary outcome, based on a previous study showing a mean difference of 65 (SD 65) seconds between groups in the ESWT [29]. Thus, considering an alpha of .05 and a beta of .80, the estimated sample size will be 17 individuals per group. To account for potential losses and to allow for data imputation under the intention-to-treat principle, the sample size was increased by 20% in each group. Therefore, the final estimated sample size was 21 participants per group.

Results

The clinical trial registration was approved in June 2023. Recruitment and data collection for the trial are ongoing; as of November 2025, a total of 10 individuals have been recruited, and the results are expected by November 2026.

Discussion

Anticipated Findings

The use of ET during rest periods does not alter lung volume; however, during exercise, individuals with COPD under the same workload (isoload) have shown a notable decrease in tidal volume, minute volume, dyspnea, and TAA. Furthermore, individuals using ET have shown a significant increase in oxygen consumption, workload, and maximal exercise duration compared with those not using ET. Additionally, ET increased the amount of time spent in moderate and vigorous physical activity, along with a marked reduction in sedentary time. As the measurements were randomly performed with or without ET, these data strongly suggest that ET attenuated the ventilatory

XSL•FO
RenderX

demand required for exercise performance across individuals [9].

Recently, Santos et al [10] demonstrated that ET use for 21 consecutive days does not alter PADL but reduces dyspnea and improves health status and HRQoL in individuals with moderate to very severe COPD in the short term. Therefore, this new, low-cost intervention also improves COPD symptoms during its continuous use. The use of ET may enhance the already-known effects of the PR program, such as increased exercise capacity and improved quality of life among individuals with COPD.

One of the strengths of the proposed study is the use of ET throughout the PR period (16 sessions over 2 months), which may improve the HRQoL of individuals with COPD. Furthermore, this is a randomized, controlled, 2-arm, and blinded study, making it the first study to evaluate the potential effects of ET throughout a PR program. Among the study's limitations is the inclusion of only nonobese male individuals, which followed the same design as previously published studies [9,10]; however, our goal was to minimize bias, so we did not include obese individuals due to the protective effect of obesity on the respiratory mechanics of patients with COPD [30]. Therefore, future studies should include obese individuals and both male and female individuals. The single-center design and the inclusion of individuals who do not use oxygen therapy will make it difficult to generalize the interpretation of the data at the end of the study. The secondary outcomes, such as HRQoL, anxiety, and depression, were classified as exploratory. Therefore, these outcomes were used for a more comprehensive understanding of the intervention's effects; however, the statistical analysis for these aspects may not have sufficient statistical power to yield definitive conclusions, and the results regarding secondary outcomes should be interpreted with due caution. For this reason, the results for these outcomes should be interpreted as preliminary insights, suggesting the need for future investigations into these outcomes, especially in female participants, as COPD affects both sexes similarly.

Conclusions

We hypothesize that the use of ET can enhance the benefits of PR in individuals with moderate to very severe COPD and increase exercise capacity and quality of life, as well as reduce symptoms of anxiety and depression.

Acknowledgments

The authors declare that no artificial intelligence tools were used in the development of the study's concept, research questions, objectives, methods, or references.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The data supporting this study's findings are available from the corresponding author, CRFC, upon reasonable request.

Authors' Contributions

TFP, JMBDS, ECMR, CCBMS, and CRFC contributed to the conception and planning of the study and to the initial data collection. TFP, JMBDS, ECMR, CCBMS, FFL, RMCP, and CRFC reviewed the manuscript's writing and/or provided critical revisions of its intellectual content. All authors read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

1. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. : Global Initiative for Chronic Obstructive Lung Disease; 2023 URL: https://www.knowledge-action-portal.com/en/content/global-strategy-diagnosis-management-and-prevention-chronic-obstructive-pulmonary-disease-0 [accessed 2026-02-06]
2. Menezes AMB, Jardim JR, Pérez-Padilla R, et al. Prevalence of chronic obstructive pulmonary disease and associated factors: the PLATINO Study in São Paulo, Brazil. Cad Saude Publica 2005;21(5):1565-1573. [doi: 10.1590/s0102-311x2005000500030] [Medline: 16158163]
3. Marchioro J, Gazzotti MR, Moreira GL, et al. Anthropometric status of individuals with COPD in the city of São Paulo, Brazil, over time - analysis of a population-based study. J Bras Pneumol 2019 Jul 29;45(6):e20170157. [doi: 10.1590/1806-3713/e20170157] [Medline: 31365731]
4. Kawachi S, Fujimoto K. Efficacy of tiotropium and olodaterol combination therapy on dynamic lung hyperinflation evaluated by hyperventilation in COPD: an open-label, comparative before and after treatment study. Int J Chron Obstruct Pulmon Dis 2019 May 27;14:1167-1176. [doi: 10.2147/COPD.S201106] [Medline: 31213796]

XSL•FO
RenderX

5. Cano Porras D, Lunardi AC, Marques da Silva CCB, et al. Comparison between the phase angle and phase shift parameters to assess thoracoabdominal asynchrony in COPD patients. J Appl Physiol (1985) 2017 May 1;122(5):1106-1113. [doi: 10.1152/japplphysiol.00508.2016] [Medline: 28183817]
6. Priori R, Aliverti A, Albuquerque AL, Quaranta M, Albert P, Calverley PMA. The effect of posture on asynchronous chest wall movement in COPD. J Appl Physiol (1985) 2013 Apr;114(8):1066-1075. [doi: 10.1152/japplphysiol.00414.2012] [Medline: 23412901]
7. Romagnoli I, Gigliotti F, Lanini B, et al. Chest wall kinematics and breathlessness during unsupported arm exercise in COPD patients. Respir Physiol Neurobiol 2011 Sep 15;178(2):242-249. [doi: 10.1016/j.resp.2011.06.014] [Medline: 21729772]
8. Yamaguti WP, Paulin E, Salge JM, Chammas MC, Cukier A, Carvalho CR. Diaphragmatic dysfunction and mortality in patients with COPD. J Bras Pneumol 2009 Dec;35(12):1174-1181. [doi: 10.1590/s1806-37132009001200003] [Medline: 20126918]
9. Pinto TF, Fagundes Xavier R, Lunardi AC, et al. Effects of elastic tape on thoracoabdominal mechanics, dyspnea, exercise capacity, and physical activity level in nonobese male subjects with COPD. J Appl Physiol (1985) 2020 Sep 1;129(3):492-499. [doi: 10.1152/japplphysiol.00690.2019] [Medline: 32702276]
10. Santos EAD, Pinto TF, Xavier RF, et al. Elastic tape reduces dyspnea and improves health status in the short term in nonobese COPD males: a randomized controlled trial. Respir Med 2024 Jan;221:107459. [doi: 10.1016/j.rmed.2023.107459] [Medline: 37944827]
11. Cruz J, Brooks D, Marques A. Walk2Bactive: a randomised controlled trial of a physical activity-focused behavioural intervention beyond pulmonary rehabilitation in chronic obstructive pulmonary disease. Chron Respir Dis 2016 Feb;13(1):57-66. [doi: 10.1177/1479972315619574] [Medline: 26703921]
12. Holland AE, Spruit MA, Troosters T, et al. An official European Respiratory Society/American Thoracic Society technical standard: field walking tests in chronic respiratory disease. Eur Respir J 2014 Dec;44(6):1428-1446. [doi: 10.1183/09031936.00150314] [Medline: 25359355]
13. Troosters T, Janssens W, Demeyer H, Rabinovich RA. Pulmonary rehabilitation and physical interventions. Eur Respir Rev 2023 Jun 7;32(168):220222. [doi: 10.1183/16000617.0222-2022] [Medline: 37286219]
14. Moher D, Jones A, Lepage L, CONSORT Group (Consolidated Standards for Reporting of Trials). Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. JAMA 2001 Apr 18;285(15):1992-1995. [doi: 10.1001/jama.285.15.1992] [Medline: 11308436]
15. Sealed envelope. Randomisation and online databases for clinical trials. URL: https://www.sealedenvelope.com/ [accessed 2026-02-13]
16. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [Medline: 7154893]
17. Spruit MA, Singh SJ, Garvey C, et al. An official American Thoracic Society/European Respiratory Society statement: key concepts and advances in pulmonary rehabilitation. Am J Respir Crit Care Med 2013 Oct 15;188(8):e13-e64. [doi: 10.1164/rccm.201309-1634ST] [Medline: 24127811]
18. Storer TW. Exercise in chronic pulmonary disease: resistance exercise prescription. Med Sci Sports Exerc 2001 Jul;33(7 Suppl):S680-S692. [doi: 10.1097/00005768-200107001-00006] [Medline: 11462077]
19. Singh SJ, Morgan MD, Scott S, Walters D, Hardman AE. Development of a shuttle walking test of disability in patients with chronic airways obstruction. Thorax 1992 Dec;47(12):1019-1024. [doi: 10.1136/thx.47.12.1019] [Medline: 1494764]
20. Dourado VZ, Guerra RLF, Tanni SE, Antunes LDO, Godoy I, Antunes LC. Reference values for the incremental shuttle walk test in healthy subjects: from the walk distance to physiological responses. J Bras Pneumol 2013;39(2):190-197. [doi: 10.1590/s1806-37132013000200010] [Medline: 23670504]
21. Moreira GL, Pitta F, Ramos D, et al. Portuguese-language version of the Chronic Respiratory Questionnaire: a validity and reproducibility study. J Bras Pneumol 2009 Aug;35(8):737-744. [doi: 10.1590/s1806-37132009000800004] [Medline: 19750325]
22. Botega NJ, Bio MR, Zomignani MA, Garcia C Jr, Pereira WA. Mood disorders among inpatients in ambulatory and validation of the anxiety and depression scale HAD. Rev Saude Publica 1995 Oct;29(5):355-363. [doi: 10.1590/s0034-89101995000500004] [Medline: 8731275]
23. Cullen K, Talbot D, Gillmor J, et al. Effect of baseline anxiety and depression symptoms on selected outcomes following pulmonary rehabilitation. J Cardiopulm Rehabil Prev 2017 Jul;37(4):279-282. [doi: 10.1097/HCR.0000000000000258] [Medline: 28640172]
24. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S531-S543. [doi: 10.1249/01.mss.0000185657.86065.98] [Medline: 16294116]
25. Spelta F, Fratta Pasini AM, Cazzoletti L, Ferrari M. Body weight and mortality in COPD: focus on the obesity paradox. Eat Weight Disord 2018 Feb;23(1):15-22. [doi: 10.1007/s40519-017-0456-z] [Medline: 29110280]
26. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. Eur Respir J 2005 Aug;26(2):319-338. [doi: 10.1183/09031936.05.00034805] [Medline: 16055882]
27. Pereira CA, Neder JA. Diretrizes para testes de função pulmonar. J Bras Pneumol 2002;28(3):S1-S238 [FREE Full text]

XSL•FO
RenderX

28. Craig CL, Marshall AL, Sjöström M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
29. Pepin V, Laviolette L, Brouillard C, et al. Significance of changes in endurance shuttle walking performance. Thorax 2011 Feb;66(2):115-120. [doi: 10.1136/thx.2010.146159] [Medline: 21148135]
30. Ora J, Laveneziana P, Wadell K, Preston M, Webb KA, O'Donnell DE. Effect of obesity on respiratory mechanics during rest and exercise in COPD. J Appl Physiol (1985) 2011 Jul;111(1):10-19. [doi: 10.1152/japplphysiol.01131.2010] [Medline: 21350021]

Abbreviations

BP: blood pressure
COPD: chronic obstructive pulmonary disease
CRQ: Chronic Respiratory Questionnaire
ESWT: endurance shuttle walking test
ET: elastic tape
HADS: Hospital Anxiety and Depression Scale
HR: heart rate
HRQoL: health-related quality of life
ISWT: incremental shuttle walking test
LL: lower limb
PADL: physical activity in daily life
PLATINO: Projeto Latino-Americano de Investigação em Obstrução Pulmonar
PR: pulmonary rehabilitation
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SpO_2: peripheral oxygen saturation
TAA: thoracoabdominal asynchrony

Edited by J Sarvestan; submitted 06.Apr.2025; peer-reviewed by J Edwards, N Atilla; revised version received 27.Dec.2025; accepted 29.Dec.2025; published 27.Feb.2026.

Please cite as:
Pinto TF, dos Santos JDMB, Reis ECM, Marques da Silva CCB, de Lima FF, Carvalho-Pinto RM, Carvalho CRF
Effects of Elastic Tape Associated With Pulmonary Rehabilitation in Male Individuals With Chronic Obstructive Pulmonary Disease: Protocol for a 2-Arm, Assessor-Blinded Randomized Controlled Trial
JMIR Res Protoc 2026;15:e75029
URL: https://www.researchprotocols.org/2026/1/e75029
doi:10.2196/75029

© Thiago Fernandes Pinto, Juliana de Melo Batista dos Santos, Estéfane Caroline Monteiro Reis, Cibele C Berto Marques da Silva, Fabiano Francisco Lima, Regina Maria Carvalho-Pinto, Celso Ricardo Fernandes Carvalho. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Kidney-Tonifying, Phlegm-Resolving, and Blood Stasis–Removing Therapy for Multiple Myeloma: Protocol for a Randomized Controlled Trial on Epigenetic and Immune Modulation

Xiaoqi Sun, MSc; Yongming Zhou, MSc; Yujue Wang, BSc; Youya Dai, BSc; Wenwei Zhu, PhD; Hailin Chen, PhD

Department of Hematology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, 110 Ganhe Road, Hongkou District, Shanghai, China

Corresponding Author:
Hailin Chen, PhD
Department of Hematology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, 110 Ganhe Road, Hongkou District, Shanghai, China

Abstract

Background: Multiple myeloma (MM) is characterized by kidney deficiency, phlegm, and blood stasis as core findings, specifically in Traditional Chinese Medicine (TCM), and the kidney-tonifying, phlegm-resolving, and blood stasis–removing (KPR) method is a fundamental therapeutic approach for MM in TCM. Western medicine primarily focuses on targeted immunotherapy or chemotherapy for MM treatment, whereas TCM characterizes MM through distinct pathological patterns that directly correspond to immune microenvironment dysregulation. Emerging evidence implicates the PHD finger protein 19 (PHF19)/enhancer of zeste homolog 2 (EZH2)/trimethylated histone H3 at lysine 27 (H3K27me3) epigenetic axis in immune microenvironment dysregulation and MM progression. Notably, TCM “blood stasis” correlates with hypoxia-induced immune gene silencing in MM bone marrow, and KPR (a clinically validated TCM decoction with 16 herbs) acts on this axis via its active components that regulate EZH2 and epigenetic function, merging TCM syndrome differentiation with modern epigenetics. We have designed a randomized controlled trial (RCT) to investigate the mechanism of action and safety of the KPR method in MM.

Objective: This RCT aims to assess whether a KPR herbal formula combined with standard bortezomib-based therapy improves the immune microenvironment via the PHF19-EZH2-H3K27me3 epigenetic axis to restore immune function in MM, providing a mechanistic basis for integrating TCM into evidence-based oncology care in relapsed or refractory patients.

Methods: This is a single-center, prospective RCT involving patients with MM. It has been designed to test the hypothesis that the KPR formula epigenetically regulates the PHF19-EZH2-H3K27me3 axis to improve the immune microenvironment. Patients are randomly assigned in a 1:1:1 ratio to 3 groups (blank control group, Western medicine control group, and integrated TCM and Western medicine treatment group). All patients undergo 12 weeks of treatment and a 6-month follow-up. The primary outcome is the $CD3^+$ T-cell ratio in bone marrow/peripheral blood, which is detected by flow cytometry. The secondary outcomes include quantified TCM syndrome scores, Western medicine efficacy evaluation criteria, complete blood count, bone marrow morphology, blood and urine immunoglobulin levels, quantitative M protein levels, free light chain levels, β2-microglobulin levels, and whole-body imaging findings. Statistical analysis involves linear mixed models for longitudinal data and Bonferroni correction to verify KPR’s immunomodulatory effects via the targeted epigenetic axis.

Results: This study was funded in November 2023. Recruitment was initiated in March 2025 and is expected to be completed in February 2026. As of October 2025, 41 patients have been enrolled. Data collection is projected to end in October 2026. Data analysis has not yet been initiated, and the results are expected to be published in 2027.

Conclusions: This unique mechanistic RCT evaluating a TCM formula targeting the PHF19-EZH2-H3K27me3 axis in patients with MM will establish a biomarker-driven framework for integrating TCM with immunotherapy, offering novel strategies for treatment-refractory patients.

Trial Registration: International Traditional Medicine Clinical Trial Registry ITMCTR2025000671; https://itmctr.ccebtcm.org.cn/mgt/project/view/4375050683463103813 and ITMCTR2025000449; https://itmctr.ccebtcm.org.cn/mgt/project/view/4779140633660123982

International Registered Report Identifier (IRRID): DERR1-10.2196/86322

(JMIR Res Protoc 2026;15:e86322) doi:10.2196/86322

KEYWORDS

myeloma; herbal medicine; randomized controlled trial; epigenetic regulation; tumor immune microenvironment

XSL•FO
RenderX

Introduction

Background

Multiple myeloma (MM) is a hematological malignancy driven by clonal plasma cell proliferation, which damages multiple organ systems. Its main manifestations include bone destruction, renal impairment, anemia, hypercalcemia, and infection [1]. Since 2010, immunotherapy has initiated a revolutionary era of treatment for MM [2,3]. Currently, many novel targets for the treatment of specific MM subtypes have been clinically identified, significantly improving progression-free survival and overall survival [4-9]. However, MM remains incurable, and drug resistance inevitably develops [10-12].

PHD finger protein 19 (PHF19) expression strongly correlates with MM progression and shows superior predictive value over the high-risk gene *MMSET*. Knockout of PHF19 causes myeloma cell lines to enter a state of reduced proliferation. As a key component of polycomb repressive complex 2 (PRC2) [13], PHF19 is expressed across all MM subgroups and is particularly overexpressed in high-risk patients with MM. Existing studies have shown that increased histone methylation, especially H3K27 trimethylation, is closely associated with disease aggressiveness [14]. PHF19 regulates histone methylation to modulate chromatin activity [15] and facilitates B cell differentiation into plasma cells in germinal centers [16]. The mechanisms of PHF19 in MM biology require further investigation. Another critical methyltransferase in MM is enhancer of zeste homolog 2 (EZH2), a PRC2 subunit. EZH2 primarily methylates H3K27, silencing target genes and contributing to MM onset, progression, and poor prognosis [17]. While proteasome inhibitor therapy can counteract its adverse effects, EZH2's negative prognostic impact appears treatment-independent [18]. Beyond driving tumorigenesis via histone methylation, EZH2 also disrupts immune homeostasis to shape the tumor microenvironment (TME), promoting progression [19]; however, its specific effects on the TME in MM require further exploration [20]. Functionally, PHF19 is closely linked to EZH2, regulating PRC2/EZH2 recruitment and activity to modulate genes critical for cell growth and differentiation. PHF19 thereby influences cell cycle and genetic stability genes, correlating strongly with high-risk MM behavior [21].

In summary, PHF19 and EZH2 are closely related. As a key epigenetic methyltransferase, EZH2 participates in the pathogenesis within the TME. These findings suggest that the triad of PHF19, EZH2, and the immune microenvironment interacts and contributes to the pathological mechanisms of multidrug resistance in MM. However, related research on this topic has not been thoroughly investigated.

In ancient Chinese medicine, MM was not specifically named but was categorized based on symptoms such as "bone bi syndrome" (骨痹), "low back pain" (腰痛), "consumptive disorder" (虚劳), and "bone erosion" (骨蚀). At the 2008 Special Symposium on TCM Disease Names for Hematological Diseases, a consensus was reached to designate "marrow tumor" (骨髓瘤) as the Traditional Chinese Medicine (TCM) disease name for MM [22]. However, physicians have held varying views on its etiology and pathogenesis [23,24]. The primary pathogenic factors are generally attributed to "toxin" (毒), "deficiency" (虚), "phlegm" (痰), and "blood stasis" (瘀). Toxin invasion weakens the body's vital energy (正气), leading to the generation of phlegm and stasis; this interplay of underlying deficiency and superficial pathogenic factors creates a vicious cycle of mutual reinforcement. Yueyang Hospital of Integrated Traditional Chinese and Western Medicine (affiliated to the Shanghai University of Traditional Chinese Medicine) has explored treatment approaches combining Western medicine with the "method of tonifying the kidney, resolving phlegm, and removing blood stasis." The basic formula consists of the following components: *Taxillus chinensis* (24 g), *Eucommia ulmoides* (24 g), *Epimedium* herb (24 g), stir-fried *Atractylodes macrocephala* rhizome (12 g), *Scrophularia ningpoensis* root (12 g), processed *Pinellia ternata* rhizome (18 g), *Prunella vulgaris* spike (15 g), white mustard seed (9 g), *Fritillaria thunbergii* bulb (18 g), oyster shell (30 g), *Hedyotis diffusa* herb (30 g), *Scutellaria barbata* herb (30 g), *Duchesnea indica* herb (15 g), *Eupolyphaga sinensis* insect (9 g), processed *Bombyx batryticatus* larva (9 g), and processed *Glycyrrhiza uralensis* root (6 g). Preliminary qualitative mass spectrometric analysis by our research group identified active components, including naringenin, icariin, and formononetin, which demonstrate therapeutic efficacy against MM [25-27]. Rooted in clinical experience, this approach addresses the core pathogenesis and variable manifestations of MM, showing good patient acceptance.

However, the mechanism by which the kidney-tonifying, phlegm-resolving, and blood stasis–removing (KPR) method combined with a basic Western medicine regimen treats MM has not been thoroughly investigated. Therefore, we propose the following hypothesis: The KPR method may act on PHF19 to regulate EZH2/trimethylated histone H3 at lysine 27 (H3K27me3) epigenetically, ultimately improving the immune microenvironment to inhibit immune escape and exert therapeutic effects, and thereby enhancing the clinical efficacy of MM.

Aims and Research Questions

We propose a multilevel mechanistic framework bridging TCM syndrome differentiation with PHF19-EZH2-H3K27me3–mediated immune dysregulation in MM.

First, we think kidney deficiency (肾虚) indicates PHF19 transcriptional activation. In TCM, the "kidney" governs the "marrow" and governs "essence" (精), analogous to hematopoietic stem cell niche integrity [28]. We think that kidney deficiency–induced stress signals may transcriptionally activate PHF19. Herbs, such as *Scutellaria barbata* and *Scutellaria baicalensis*, may stabilize HIF-1α degradation through prolyl hydroxylase domain (PHD) inhibition, thereby suppressing PHF19 overexpression in MM cells [29-31].

Second, we think phlegm indicates EZH2 catalytic activity. "Phlegm" in MM corresponds to pathological glycoprotein-rich microenvironment and extracellular matrix remodeling. Phlegm-induced transforming growth factor-β (TGF-β) signaling enhances EZH2 phosphorylation at Ser21, increasing its

XSL•FO
RenderX

methyltransferase activity [32,33]. Phlegm-resolving herbs, such as rhizome of lagehead atractylodes, may inhibit TGF-β phosphorylation, blocking EZH2 recruitment to the PRC2 complex and thus reducing H3K27me3 deposition on immune gene loci [34,35].

Third, blood stasis may indicate H3K27me3-mediated immune gene silencing. Stasis reflects hypoxia, acidosis, and inflammatory cytokine entrapment in MM bone marrow. This microenvironment induces DNA hypermethylation at CpG islands of T-cell exhaustion genes, which synergizes with H3K27me3 and irreversibly silences effector function [36,37]. Stasis-eliminating herbs, such as *Hedyotis diffusa* and ground beetle *Eupolyphaga sinensis*, may inhibit EZH2 interaction, preventing heterochromatin propagation.

Finally, we think the KPR formula induces immune reprogramming through polypharmacological effects, such as direct PHF19 siRNA-like activity via herb-derived microRNAs [38-40] and gut microbiome modulation by herbs to increase *Lactobacillus* [41].

This hypothesis integrates TCM “holistic network regulation” with molecular epigenetics, providing a falsifiable mechanistic scaffold for our randomized controlled trial (RCT).

Methods

Study Location and Recruitment Procedure

The study is conducted at the Department of Hematology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, a tertiary hospital with integrated TCM-Western medicine expertise, ensuring targeted recruitment of patients with MM seeking care here.

Recruitment follows a systematic, multichannel procedure. Public advertisements are posted on electronic display screens in hematology clinic waiting areas and MM/hematological disease patient education WeChat groups. Eligible interested individuals contact the dedicated research team via the advertised phone number or in-person inquiries during clinic visits, undergo preliminary screening with a standardized checklist (eg, age and MM diagnosis), and complete a comprehensive formal evaluation (Western medical MM diagnosis confirmation, TCM syndrome differentiation, and exclusion criteria review), and those meeting all criteria are enrolled after receiving detailed study information and signing a written informed consent form. The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist is provided in Checklist 1. The study flowchart is shown in Figure 1.

XSL•FO
RenderX

Figure 1. Study flowchart. EZH2: enhancer of zeste homolog 2; H3K27me3: trimethylated histone H3 at lysine 27; KPR: kidney-tonifying, phlegm-resolving, and blood stasis–removing; MM: multiple myeloma; NK: natural killer; PHF19: PHD finger protein 19.

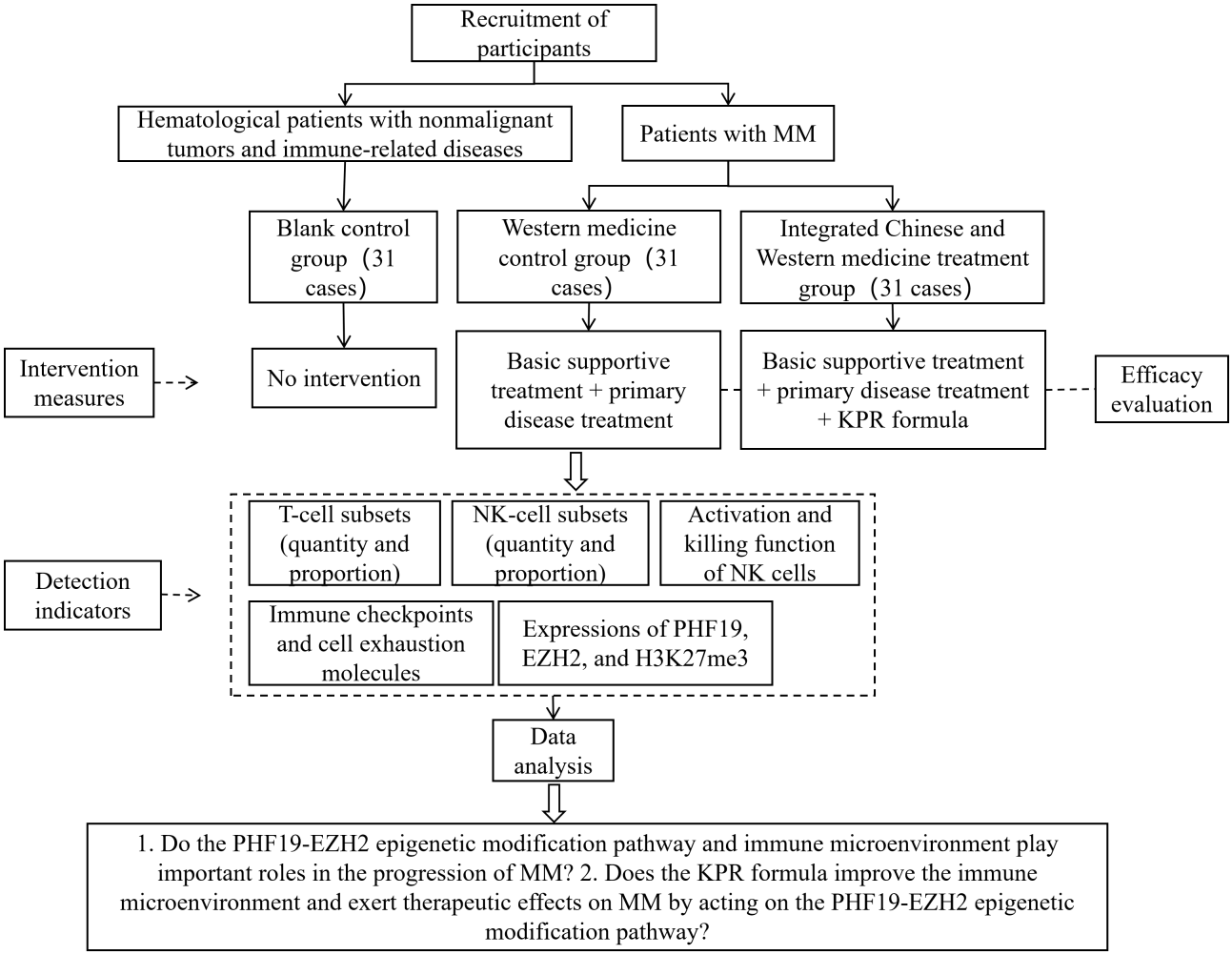

Source of Cases and Study Population

A total of 62 patients with MM meeting the diagnostic criteria have been enrolled from the outpatient and inpatient departments of hematology at Yueyang Hospital of Integrated Traditional Chinese and Western Medicine. Additionally, 31 patients with hematological diseases (excluding malignancies and immune-related disorders) have been included in a blank control group.

Diagnostic Criteria

Western Medicine Diagnostic Criteria

Diagnosis was based on the following criteria: a bone marrow monoclonal plasma cell proportion ≥10% and/or histopathological confirmation of plasmacytoma, with at least one of the SLiM-CRAB features (Multimedia Appendix 1).

TCM Diagnostic Criteria

The condition falls under the category of “bone bi syndrome” (骨痹) in TCM. The TCM syndrome differentiation was formulated with reference to the Diagnostic and Therapeutic Criteria for Internal Medicine Diseases and Syndromes in TCM issued by the National Administration of TCM [42], combined with the common clinical manifestations of patients with MM. The TCM syndrome diagnostic criteria for kidney deficiency, phlegm, and blood stasis syndrome (Multimedia Appendix 2) have been applied, requiring the presence of all major symptoms and at least 2 secondary symptoms, and confirmation via tongue and pulse examinations. All 3 TCM practitioners have undergone standardized training. Interrater reliability for syndrome diagnosis is κ=0.78 (substantial agreement) [22-24]. TCM diagnosis retains inherent subjectivity despite standardized criteria and substantial interrater reliability (κ=0.78). We have mitigated this by correlating syndrome improvements with objective molecular markers (PHF19 and H3K27me3).

Inclusion, Exclusion, and Discontinuation Criteria

The inclusion, exclusion, and discontinuation criteria are provided in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion, exclusion, and discontinuation criteria.

Inclusion criteria

- Diagnosis of multiple myeloma (MM) according to Western medicine diagnostic criteria
- Diagnosis of kidney deficiency, phlegm, and blood stasis syndrome (TCM syndrome) based on TCM diagnostic criteria
- Age ≥18 years or ≤80 years, with no gender restriction
- Willingness to receive the treatment regimen and sign an informed consent form

Exclusion criteria

- Allergic constitution or hypersensitivity to MM therapeutic agents
- Severe cardiovascular diseases, including myocardial infarction or heart failure of grade 3 or higher
- Severe abnormalities in liver or renal function (alanine aminotransferase >5 times the upper limit of normal; glomerular filtration rate <15 mL/min)
- Clinically significant systemic diseases, such as osteoporosis, bone metastatic carcinoma, or lumbar tuberculosis
- Presence of ≥2 complications (eg, bone disease, infection, hemorrhage, and anemia)
- Alcoholism, drug abuse, mental disorders, or intellectual disability
- Pregnancy or lactation

Discontinuation, termination, and attrition criteria

- Inability to tolerate the trial drug
- Development of serious adverse events or adverse reactions related to the treatment regimen
- Withdrawal by the patient (eg, due to poor efficacy or intolerable adverse reactions)
- Poor compliance (eg, failure to take medication as prescribed, addition of other medications that can affect efficacy assessment, or failure to attend follow-up visits)
- Loss to follow-up

Randomization and Blinding

Randomization numbers are assigned based on the order of patient visits, and participants are grouped using the random number table method.

The KPR herbal decoction is prepared using herbs supplied by Shanghai Hongqiao TCM Co, Ltd, and uniformly decocted by the pharmacy of Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, under standardized parameters (200 g of herbs boiled in 1600 mL of water for 45 minutes at 100 °C; concentrated to 200 mL per pouch). All batches are required to meet safety standards for heavy metals and pesticides. Patients receive sequential batches per enrollment order, with batch numbers documented for potential sensitivity analysis. A placebo, primarily composed of maltodextrin and lactose, is color-matched with caramel to mimic the appearance and weight of the active decoction.

This is a single-blind trial. Participants remain blinded to allocation; the placebo matches the KPR decoction in appearance, taste, and packaging (maltodextrin base with added natural coloring and bitter flavoring). Outcome assessors (laboratory staff and statisticians) are blinded, while TCM practitioners who dispense decoctions and adjust prescriptions based on syndrome differentiation are not blinded for clinical feasibility. Blinding integrity will be assessed using the Bang Index at weeks 4 and 12.

Grouping and Interventions

Overview

A total of 62 patients with MM are randomly divided into the following 2 groups: Western medicine control group and integrated TCM-Western medicine treatment group, with 31 patients in each group. The general clinical data of the 2 groups are comparable. Additionally, 31 patients with nonmalignant and nonimmune hematological diseases are included in a blank control group (Multimedia Appendix 3).

Given that MM is a life-threatening malignancy, a pure placebo arm would be ethically indefensible according to the Helsinki Declaration §33 [43]. Our design includes the following study arms:

- Blank control (arm 1): Patients with nonmalignant hematological diseases providing baseline immune reference values, which are essential for the mechanistic interpretation of PHF19-EZH2 dynamics in cancer versus normal states.
- Western medicine control (arm 2): Patients with MM receiving standard-of-care bortezomib-based therapy, which serves as an active comparator, enabling the isolation of the additive effect of KPR therapy on immune reprogramming.
- Integrated TCM-Western medicine (arm 3): Patients with MM receiving the test intervention.

This design directly addresses our hypothesis: Does KPR therapy augment immune recovery beyond what is achieved by

XSL•FO
RenderX

standard therapy alone? A placebo-KPR group would answer a different question (placebo effect magnitude) but would not inform clinical integration.

Blank Control Group

Patients receive treatment for underlying diseases as needed, with no additional interventions administered.

Western Medicine Control Group

Patients receive risk-stratified supportive care according to Padua thrombotic risk assessment scores (Multimedia Appendix 4) combined with standard antimyeloma therapy (Multimedia Appendix 5).

Integrated TCM-Western Medicine Treatment Group

Patients receive basic supportive care, treatment for the primary disease, and KPR therapy. The modification of prescriptions based on syndrome differentiation is shown in Multimedia Appendix 6.

Data Collection and Management

Data are collected from daily symptom records and examination forms completed by participants, as well as from case report forms (CRFs) and TCM syndrome scores documented by medical staff and laboratory results obtained from Yueyang Hospital of Integrated Traditional Chinese and Western Medicine.

CRFs are double-entered and cross-checked by trained clinical research coordinators/data managers. Paper CRFs are transcribed within 7 days. Monitors conduct on-site source data verification. The system is preset with logical rules: key indicators are 100% verified, and nonkey indicators are spot-checked at a rate of ≥20%. Data discrepancies are automatically flagged, generating a query list sent to investigators for resolution within 5 days, and unresolved issues are escalated to monitors. After prelocking, the database is reviewed by the Clinical Trial Data Monitoring Committee. Final locking requires confirmation by the sponsor, investigators, and statisticians, with data use exclusively for statistical analysis. Electronic data will be encrypted, backed up daily, and retained for ≥5 years. Paper documents will be stored in locked cabinets, archived within 3 months after the trial, and retained for ≥10 years.

Investigators have completed Good Clinical Practice (GCP) and protocol training before enrollment. Routine monitoring includes on-site visits every 4 weeks to verify CRF quality and source data. Centralized monitoring via the electronic data capture system enables real-time detection of anomalies. Annual third-party audits assess GCP compliance.

Before unblinding, a comprehensive review of the data is conducted to evaluate the appropriateness of defining the intention-to-treat population, modified intention-to-treat population, per-protocol population, and safety analysis population. This assessment requires joint approval from the data manager, study promoter, and core researchers. Unblinding is completed after database locking, formally initiating the statistical analysis and results reporting phase.

Assessment: Primary and Secondary Endpoints

Primary Endpoints

Flow cytometry is used to detect the $CD3^+$ T-cell ratio (primary outcome) in bone marrow and peripheral blood, and assess other T-cell subsets, including $CD4^+$ T cells, $CD8^+$ T cells, and $CD4^+CD25^+FOXP3^+$ regulatory T cells (Tregs). Moreover, the immune checkpoint molecules PD-1, LAG-3, and TIM-3 are evaluated. Furthermore, natural killer (NK) cell subsets, including CD3 $CD56^+$ NK cells, $CD56^{high}CD16$ NK cells, and $CD56^{dim}CD16^+$ NK cells, are assessed, and the presence of the activation molecule NKG2D on NK cells and the expression levels of its ligands (MICA and MICB) on $CD34^+$ cells are determined. In addition, the expression levels of NK cytotoxic activity markers CD107a, interferon-γ, granzyme B, and perforin are evaluated.

Secondary Endpoints

For quantifying the TCM syndrome score, syndromes (main symptoms, secondary symptoms, and tongue and pulse manifestations) are observed and recorded once before treatment and at 12 weeks after treatment. The efficacy is graded as follows: (1) complete recovery (disappearance or near-disappearance of clinical symptoms and signs, with a reduction in the syndrome score of ≥95%), (2) marked improvement (significant improvement in clinical symptoms and signs, with a reduction in the syndrome score of 70% - 95%), (3) effective (improvement in clinical symptoms and signs, with a reduction in the syndrome score of 30% - 70%), and (4) ineffective (no significant improvement in clinical symptoms and signs, with a reduction in the syndrome score of <30%). The nimodipine method is used for calculation, and the formula is as follows: ([pretreatment score − posttreatment score] / pretreatment score) × 100%.

For the assessment of Western medicine, efficacy is evaluated according to the 2024 Revised Chinese Guidelines for the Diagnosis and Treatment of MM at 12 weeks after treatment and is categorized as follows: (1) stringent complete response, (2) complete response, (3) very good partial response, (4) partial response, (5) minimal response, (6) stable disease, (7) progressive disease, and (8) clinical relapse.

The main symptoms include generalized bone pain, fever, fatigue, dysuria, nausea, vomiting, numbness, and tingling in the hands and feet. The secondary symptoms include soreness and weakness in the lower back and knees, shortness of breath, dizziness, tinnitus, palpitations, chest tightness, hemorrhage, edema, and loss of appetite. The tongue and pulse manifestations include tortuous sublingual veins, a pale and dull tongue, a white and slippery coating, and a deep, weak pulse. Observations and records are made before and after treatment.

Blood tests include complete blood count and C-reactive protein. Bone marrow assessments include bone marrow cytology, bone marrow biopsy for pathological analysis, chromosome analysis, flow cytometry, and abnormal gene testing. Blood and urine are assessed for immunoglobulins, quantitative M protein, free light chains, and β2-microglobulin. Furthermore, whole-body radiography, computed tomography, magnetic resonance

XSL•FO
RenderX

imaging, and positron emission tomography-computed tomography are performed. Observations are made before treatment and at 12 weeks after treatment.

Safety Evaluation

Safety evaluations, including urinalysis/stool routine, liver/renal function tests, and electrocardiography, are conducted at baseline and at 12 weeks after treatment.

Adverse Event Record

Adverse event time, type (eg, adverse reaction and device malfunction), and severity (mild, moderate, or severe); basic subject information (name abbreviation and ID); management measures (eg, drug discontinuation and symptomatic treatment); and outcomes (recovery, persistence, or death) of adverse events or serious adverse events (SAEs) are recorded.

Adverse events are documented within 24 hours using standardized forms, and these are reviewed and signed by the principal investigator with clinical research coordinator assistance and Ethics Committee oversight. SAEs are reported to the Ethics Committee and regulatory authorities within 24 hours. Non-SAEs are reported quarterly. Both paper (signed) and electronic (encrypted) records will be archived for ≥5 years. Potential adverse reactions include Western medicine–related bone marrow suppression (anemia, leukopenia, thrombocytopenia, hemorrhage, and infection/sepsis), gastrointestinal toxicity (nausea, vomiting, abdominal pain, diarrhea, and mucosal ulcers), hepatotoxicity (elevated transaminases, fatty liver, and jaundice), dermatological reactions (dermatitis, drug rash, and phlebitis), and disease progression/relapse, as well as KPR method–related gastrointestinal side effects (mild diarrhea, nausea, and anorexia from cold/bitter herbs like *Oldenlandia diffusa* and *Scutellaria barbata*) and allergic reactions (rashes, papules, and erythema) to herbs such as *Eupolyphaga sinensis* and *Bombyx batryticatus*.

Sample Size Calculation

This study is an RCT, and the $CD3^+$ T-cell ratio of participants is the observed outcome indicator. Based on a comprehensive review of the literature, the $CD3^+$ T-cell ratio in the Western medicine control group is determined to be 50.15 (SD 4.27). It is estimated that the $CD3^+$ T-cell ratio in the integrated TCM-Western medicine treatment group could increase by 3.86 points [44,45]. With a 2-tailed α of .05 and a power of 90%, the sample size is calculated using the following formula:

$$n=2(Z\alpha+Z\beta)2*\sigma2\delta$$

where $Z_{\alpha}/2$ is the critical value for a 2-tailed test at α=.05 (1.96), Z_{β} is the critical value for a power of 90% (1.28), σ is the SD (4.27), and δ is the difference in means (3.86).

Using the CD3+ T-cell ratio as the observed outcome indicator, with a 2-tailed α of .05 and a power of 90%, the required sample size per group is 26 cases. Considering 1:1:1 randomization, 26 cases are needed per group. Accounting for a 15% loss to follow-up and refusal, a minimum of 31 cases per group (total ≥93 cases) are to be enrolled.

While the MM immune microenvironment is indeed complex, we have selected the $CD3^+$ T-cell ratio as the sole primary outcome for sample size calculation based on the following integrated rationale.

$CD3^+$ T cells have a central integrator function. They serve as the master regulator of adaptive antitumor immunity. Their abundance in bone marrow directly correlates with clinical response to immunotherapy and overall survival in patients with MM (ρ=0.71; *P*<.01) [1]. As a composite marker, $CD3^+$ T-cell recovery integrates successful reprogramming of multiple upstream pathways (Treg suppression, checkpoint blockade, and NK-T cell crosstalk), making it a more robust single indicator than isolated subset markers.

$CD3^+$ T cells have a direct mechanistic linkage to the PHF19-EZH2 axis. PHF19 overexpression specifically drives T-cell exhaustion through H3K27me3-mediated silencing of T-cell effector genes [40]. Our preclinical data show that PHF19 knockdown increases $CD3^+$ T-cell infiltration by 4.2-fold, mirroring the hypothesized 3.86-point clinical effect. Thus, $CD3^+$ T cells represent the most proximal and relevant indicator of our targeted epigenetic intervention.

According to CONSORT (Consolidated Standards of Reporting Trials) guidelines, a single primary outcome avoids multiplicity inflation (type I error) and ensures adequate power for the core hypothesis. The secondary endpoints characterize the full immune landscape without requiring separate powering, and if used, they would require an over 3-fold larger sample size, compromising the feasibility of this early phase mechanistic trial.

Food and Drug Administration guidance for immuno-oncology mechanistic trials supports the use of a single validated immune parameter for powering early phase studies, with composite endpoints reserved for confirmatory trials [46,47].

Statistical Analysis

To address the study hypothesis that the KPR formula epigenetically regulates the PHF19-EZH2-H3K27me3 axis to improve the immune microenvironment, descriptive statistics, including number, percentage, mean, and SD, will be used to describe variables as appropriate. Normally distributed continuous data will be summarized as mean (SD), while nonnormally distributed continuous data will be presented as median and IQR. Categorical data will be expressed as frequency and percentage.

For the primary outcome ($CD3^+$ T-cell ratio) and key mechanistic indicators (PHF19, EZH2, and H3K27me3 expression levels), one-way ANOVA will be used for 3-group comparisons (blank control, Western medicine control, and integrated TCM-Western medicine treatment) if the data are normally distributed and homogeneous in variance; otherwise, the Kruskal-Wallis H test will be applied. Post hoc pairwise comparisons will use Bonferroni correction to control for type I error. For other continuous outcomes with normal distribution, the independent samples *t* test will be used for 2-group comparisons, and for outcomes with nonnormal distribution,

XSL•FO
RenderX

the Mann-Whitney *U* test will be used. For categorical outcomes, the chi-square test or Fisher exact test will be used when expected cell counts are less than 5. Intragroup comparisons before and after treatment will utilize the paired samples *t* test for normally distributed data and the Wilcoxon signed-rank test for nonnormally distributed data.

Longitudinal data analysis across multiple time points will be performed using linear mixed models or generalized estimating equations to account for correlations between repeated measurements from the same participant. These models will include "treatment group" (fixed effect), "time point" (fixed effect), and "group×time interaction" (fixed effect) to test whether the KPR formula induces sustained improvements in the primary outcome and mechanistic indicators beyond standard therapy, efficiently handling unequal time intervals and missing data under the missing-at-random assumption.

When multiple hypothesis tests are performed simultaneously, Bonferroni correction will be applied to control for type I error risk by adjusting the significance level. The pattern and mechanism of missing data will first be assessed, and if missing data are minimal (less than 5%), complete-case analysis may be used, while larger proportions will be handled using multiple imputation methods to create several complete datasets for analysis with pooled results. Moreover, sensitivity analyses will be conducted to evaluate the impact on conclusions.

To further validate the mechanistic link in the hypothesis, Pearson or Spearman correlation analysis will be used to explore the association between PHF19/EZH2/H3K27me3 expression levels and the $CD3^+$ T-cell ratio in the integrated TCM-Western medicine group, verifying whether the KPR formula exerts immunomodulatory effects through the targeted epigenetic axis.

All analyses will be performed on an intention-to-treat basis. A P value of <.05 is considered to indicate statistical significance, unless otherwise adjusted for multiple comparisons.

Ethical Considerations

Ethics Committee Approval

This study has been approved by the Ethics Committee of Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine (ethics approval number: 2025-006; approval date: February 10, 2025). The study protocol is conducted in accordance with the Declaration of Helsinki and relevant international ethical guidelines for human subjects research.

Informed Consent

The study protocol and potential risks/benefits are fully explained to all eligible participants. Written informed consent is obtained from each participant before enrollment and specimen collection. Participants have the right to withdraw from the study at any time without penalty, and their decision will not affect their subsequent medical care.

Privacy and Confidentiality

All participant information is deidentified using coded ID numbers, with no personal identifiers (eg, name and contact information) linked to study records. Documents containing personal identifiers (eg, informed consent forms and locator forms) are stored separately from deidentified study data in locked cabinets with restricted access. Electronic data are encrypted, password-protected, and backed up daily. Laboratory specimens, reports, and administrative forms are labeled only with coded IDs. All study-related records (paper and electronic) will be retained for ≥5 years (electronic) or ≥10 years (paper), as required by regulatory standards, and will not be disclosed to third parties without written permission from participants.

Participant Compensation

No financial or material compensation is provided to participants for their involvement in this study.

Patient and Public Involvement

Patients and/or the public are not involved in the design, conduct, reporting, or dissemination plans of this research.

Results

The study was funded in November 2023, with recruitment of eligible patients initiated in March 2025. Eligibility assessment targeted patients with MM meeting Western medical diagnostic criteria and TCM kidney deficiency, phlegm, and blood stasis syndrome criteria. As of October 2025, 41 patients with MM have been enrolled, with follow-ups scheduled according to the study protocol. Recruitment is projected to be completed by February 2026, with a target enrollment of ≥93 patients. Data collection is expected to conclude in October 2026, and the study results are anticipated to be published in 2027.

Discussion

Anticipated Main Findings

We hypothesize that KPR therapy will significantly enhance the $CD3^+$ T-cell ratio by 3.86 points versus bortezomib-based standard therapy, selected as the primary endpoint for integrating Treg suppression, checkpoint modulation, and NK crosstalk. Beyond this, we anticipate improvements in (1) immunological parameters (reduced Treg proportions, decreased checkpoint molecule expression, and enhanced NK cytotoxicity), (2) clinical outcomes (better TCM syndrome scores and higher stringent complete response/complete response rates), and (3) safety (no excess adverse events). Confirmation would provide the first clinical evidence that TCM-mediated epigenetic modulation of the PHF19-EZH2-H3K27me3 axis reprograms the MM immune microenvironment.

Principal Findings in Relation to Prior Research

Our results, if positive, would suggest that KPR therapy can counteract EZH2-driven immune suppression. This builds directly on preclinical observations indicating that PHF19 overexpression elevates immune checkpoint molecules and Treg cell frequencies [13,14,48]. These findings form the mechanistic basis for our hypothesis that a multiherb formula can epigenetically modulate the MM microenvironment, potentially shifting it from an immune-evasive state to a more immune-permissive state. The anticipated improvement in $CD3^+$ T-cell function is particularly significant, given that T-cell

XSL•FO
RenderX

exhaustion constitutes a major barrier to immunotherapy efficacy in MM [1]. Our 3-arm design (lacking an independent KPR placebo) is justified by the pragmatic framework requiring standard therapy and a mechanistic focus on augmentation. The Western medicine control serves as a benchmark to attribute improvements to KPR's epigenetic modulation vs spontaneous recovery, aligning with FDA integrative oncology guidance [44].

Comparison With Prior Work

Unlike prior TCM MM studies focused on palliation or synergistic cytotoxicity [22-24], our study is the first to test a biomarker-driven mechanism (PHF19-EZH2-H3K27me3 axis). If validated, KPR may offer a multitarget, lower-toxicity alternative to single-agent epigenetic inhibitors (eg, EZH2 inhibitors) [18], and our blank control group (nonmalignant hematological diseases) provides a critical immunological baseline rarely established in MM trials.

Strengths and Limitations

The key strengths of our study include its prospective randomized controlled design, single-blind allocation, and comprehensive immune monitoring using flow cytometry. The novel hypothesis connecting TCM principles with epigenetic-immune crosstalk represents a potential paradigm shift in integrative oncology research. However, several limitations must be acknowledged. For the first time, this study proposes the concept that Chinese herbal medicine with the therapeutic principle of "tonifying the kidney, resolving phlegm, and eliminating stasis" improves the clinical efficacy of MM by acting on PHF19 expression to mediate epigenetic regulation via the EZH2/H3K27me3 pathway, ultimately enhancing the immune microenvironment. This provides new therapeutic ideas and directions for MM treatment. There is currently a lack of international consensus on the use of TCM for the treatment of MM. Powering the analysis solely on the $CD3^+$ T-cell ratio may not capture the full immune complexity, and this is mitigated by comprehensive secondary profiling and planned mediation analysis, though future trials should consider composite immune scores for powering. Additionally, a single-blind design is necessary because practitioners cannot be blinded due to TCM syndrome differentiation requirements, which may introduce performance bias. To mitigate this, we will use blinded outcome assessors who are independent of the treatment team and unaware of group allocation. Furthermore, a standardized base formula will be used to minimize variability in TCM prescriptions.

Future Directions

This protocol lays the groundwork for several important translational next steps. Positive findings would warrant validation in larger, multicenter cohorts with extended follow-up to assess the impacts on progression-free and overall survival. Mechanistic studies using techniques, such as mass spectrometry, could help identify the active components within the KPR formula that directly modulate the PHF19-EZH2 interaction. Combination studies with novel immunotherapies (eg, bispecific antibodies and chimeric antigen receptor [CAR] T cells) could explore whether KPR therapy helps overcome resistance mechanisms by restoring T-cell fitness. Finally, developing a "PHF19 expression signature" could enable better patient stratification for personalized TCM-integrated treatment approaches.

Dissemination Plan

Consistent with open-science principles, we plan to disseminate our findings through multiple channels as follows: (1) publication in peer-reviewed journals within the fields of oncology and integrative medicine; (2) presentation at major international conferences (eg, American Society of Hematology Annual Meeting and International Myeloma Society Workshop); (3) deposition of deidentified immune monitoring datasets in public repositories (eg, Zenodo) to facilitate secondary analyses; and (4) collaboration with TCM practitioner networks to develop evidence-based clinical practice guidelines, should the efficacy of the approach be demonstrated.

Conclusions

This protocol describes the first rigorous investigation of whether a TCM formula can epigenetically reprogram the MM immune microenvironment via the PHF19-EZH2-H3K27me3 axis. By bridging mechanistic biomarker research with clinical outcomes, this study may establish a scientifically grounded framework for integrating TCM into evidence-based MM management, offering new hope for patients who develop resistance to conventional therapies.

Acknowledgments

No generative artificial intelligence (AI) tools (eg, ChatGPT, Gemini, and Llama) were used in any part of the manuscript generation process. All content of the manuscript, including study design, literature review, data analysis planning, and writing of the text, tables, and figures, was independently completed by the authors. Generative AI was not used for drafting, editing, revising, or refining the manuscript or for supporting statistical analysis or interpretation of the results. All authors attest to the originality and integrity of the work.

Funding

This work was supported by the Construction Project of National Famous Traditional Chinese Medicine Expert Inheritance Studio by the State Administration of Traditional Chinese Medicine, Shanghai Clinical Key Specialty Construction Project, High-level Key Discipline Construction Project of Traditional Chinese Medicine of the State Administration of Traditional Chinese Medicine, Peak Plateau Team Project of Shanghai University of Traditional Chinese Medicine, Shanghai Famous Traditional Chinese Medicine Academic Experience Research Studio Construction Project, and Shanghai Baiyulan Talent Plan Pujiang Project, as

XSL•FO
RenderX

well as the Natural Science Foundation of China (grant numbers: "National Medical Education Letter (2022) No. 75," shslczdzk05201, zyyzdxk-202365, 30304114341, SHGZS-2017019, 23PJD095, and 82205003).

Data Availability

This is a randomized controlled trial protocol. Upon trial completion and result publication, individual deidentified participant data, statistical code, and related study materials will be retained by the corresponding author (HC; email: chenhailinyyyy@163.com) in compliance with privacy and ethical guidelines. Requests for data access should be submitted to the corresponding author with a detailed research proposal. Eligible requests will be approved following review and signing of a data use agreement. All supplementary materials are provided as separate files with the manuscript.

Authors' Contributions

XS has conceived and designed the study. HC is responsible for the coordination of the study. YZ has developed the study design and revised the protocol. YW and YD are participating in the work of enrolling the patients and collecting the data. WZ sought funding and ethical approval. All authors have contributed to the writing of the manuscript and have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Descriptions of CRAB and SLiM.
[DOCX File, 9 KB - resprot_v15i1e86322_app1.docx]

Multimedia Appendix 2
Syndrome differentiation of kidney deficiency and phlegm stasis pattern in Traditional Chinese Medicine.
[DOCX File, 9 KB - resprot_v15i1e86322_app2.docx]

Multimedia Appendix 3
Group differences.
[DOCX File, 9 KB - resprot_v15i1e86322_app3.docx]

Multimedia Appendix 4
Padua score.
[DOCX File, 9 KB - resprot_v15i1e86322_app4.docx]

Multimedia Appendix 5
Basic disease treatment.
[DOCX File, 9 KB - resprot_v15i1e86322_app5.docx]

Multimedia Appendix 6
Modification of prescriptions based on syndrome differentiation.
[DOCX File, 9 KB - resprot_v15i1e86322_app6.docx]

Checklist 1
SPIRIT checklist.
[PDF File, 84 KB - resprot_v15i1e86322_app7.pdf]

References

1. Cowan AJ, Green DJ, Kwok M, et al. Diagnosis and management of multiple myeloma: a review. JAMA 2022 Feb 1;327(5):464-477. [doi: 10.1001/jama.2022.0003] [Medline: 35103762]
2. Verkleij CPM, Korst CLBM, van de Donk NWCJ. Immunotherapy in multiple myeloma: when, where, and for who? Curr Opin Oncol 2020 Nov;32(6):664-671. [doi: 10.1097/CCO.0000000000000677] [Medline: 32852308]
3. Ohmine K, Uchibori R. Novel immunotherapies in multiple myeloma. Int J Hematol 2022 Jun;115(6):799-810. [doi: 10.1007/s12185-022-03365-1] [Medline: 35583724]

XSL•FO
RenderX

4. Moreau P, Attal M, Hulin C, et al. Bortezomib, thalidomide, and dexamethasone with or without daratumumab before and after autologous stem-cell transplantation for newly diagnosed multiple myeloma (CASSIOPEIA): a randomised, open-label, phase 3 study. The Lancet 2019 Jul;394(10192):29-38. [doi: 10.1016/S0140-6736(19)31240-1]
5. Dimopoulos MA, Richardson PG, Bahlis NJ, et al. Addition of elotuzumab to lenalidomide and dexamethasone for patients with newly diagnosed, transplantation ineligible multiple myeloma (ELOQUENT-1): an open-label, multicentre, randomised, phase 3 trial. Lancet Haematol 2022 Jun;9(6):e403-e414. [doi: 10.1016/S2352-3026(22)00103-X] [Medline: 35550060]
6. Onda Y, Kanda J, Kaneko H, et al. Real-world effectiveness and safety analysis of carfilzomib-lenalidomide-dexamethasone and carfilzomib-dexamethasone in relapsed/refractory multiple myeloma: a multicenter retrospective analysis. Ther Adv Hematol 2022;13:20406207221104584. [doi: 10.1177/20406207221104584] [Medline: 35785245]
7. Hill E, Morrison C, Kazandjian D. Daratumumab: a review of current indications and future directions. Semin Oncol 2022 Feb;49(1):48-59. [doi: 10.1053/j.seminoncol.2022.01.008] [Medline: 35184871]
8. Frampton JE. Isatuximab: a review of its use in multiple myeloma. Targ Oncol 2021 Sep;16(5):675-686. [doi: 10.1007/s11523-021-00827-0]
9. Morabito F, Zamagni E, Conticello C, et al. Adjusted comparison between elotuzumab and carfilzomib in combination with lenalidomide and dexamethasone as salvage therapy for multiple myeloma patients. Eur J Haematol 2022 Mar;108(3):178-189. [doi: 10.1111/ejh.13723] [Medline: 34716957]
10. Gozzetti A, Ciofini S, Sicuranza A, et al. Drug resistance and minimal residual disease in multiple myeloma. Cancer Drug Resist 2022;5(1):171-183. [doi: 10.20517/cdr.2021.116] [Medline: 35582527]
11. Uckun FM. Cancer drug resistance in multiple myeloma. Cancer Drug Resist 2022;5(2):271-276. [doi: 10.20517/cdr.2022.32] [Medline: 35800374]
12. Das S, Juliana N, Yazit NAA, Azmani S, Abu IF. Multiple myeloma: challenges encountered and future options for better treatment. Int J Mol Sci 2022 Jan 31;23(3):1649. [doi: 10.3390/ijms23031649] [Medline: 35163567]
13. Ren Z, Ahn JH, Liu H, et al. PHF19 promotes multiple myeloma tumorigenicity through PRC2 activation and broad H3K27me3 domain formation. Blood 2019 Oct 3;134(14):1176-1189. [doi: 10.1182/blood.2019000578] [Medline: 31383640]
14. Mason MJ, Schinke C, Eng CLP, et al. Multiple Myeloma DREAM Challenge reveals epigenetic regulator PHF19 as marker of aggressive disease. Leukemia 2020 Jul;34(7):1866-1874. [doi: 10.1038/s41375-020-0742-z] [Medline: 32060406]
15. Cai L, Rothbart SB, Lu R, et al. An H3K36 methylation-engaging Tudor motif of polycomb-like proteins mediates PRC2 complex targeting. Mol Cell 2013 Feb 7;49(3):571-582. [doi: 10.1016/j.molcel.2012.11.026] [Medline: 23273982]
16. Ning F, Wang C, Niu S, Xu H, Xia K, Wang N. Transcription factor Phf19 positively regulates germinal center reactions that underlies its role in rheumatoid arthritis. Am J Transl Res 2018;10(1):200-211. [Medline: 29423005]
17. Schütt J, Nägler T, Schenk T, Brioli A. Investigating the interplay between myeloma cells and bone marrow stromal cells in the development of drug resistance: dissecting the role of epigenetic modifications. Cancers (Basel) 2021 Aug 13;13(16):4069. [doi: 10.3390/cancers13164069] [Medline: 34439223]
18. Pawlyn C, Bright MD, Buros AF, et al. Overexpression of EZH2 in multiple myeloma is associated with poor prognosis and dysregulation of cell cycle control. Blood Cancer J 2017 Mar 31;7(3):e549. [doi: 10.1038/bcj.2017.27] [Medline: 28362441]
19. Sun S, Yu F, Xu D, Zheng H, Li M. EZH2, a prominent orchestrator of genetic and epigenetic regulation of solid tumor microenvironment and immunotherapy. Biochim Biophys Acta Rev Cancer 2022 Mar;1877(2):188700. [doi: 10.1016/j.bbcan.2022.188700] [Medline: 35217116]
20. García-Ortiz A, Rodríguez-García Y, Encinas J, et al. The role of tumor microenvironment in multiple myeloma development and progression. Cancers (Basel) 2021 Jan 9;13(2):217. [doi: 10.3390/cancers13020217] [Medline: 33435306]
21. Davies FE, Pawlyn C, Usmani SZ, et al. Perspectives on the risk-stratified treatment of multiple myeloma. Blood Cancer Discov 2022 Jul 6;3(4):273-284. [doi: 10.1158/2643-3230.BCD-21-0205] [Medline: 35653112]
22. Wei W, Qihai D, Qinyuan C, Zhimin L, Weimin C. 多发性骨髓瘤骨病的中西医研究进展 (Advances in Traditional Chinese and Western medicine research on multiple myeloma bone disease) [Article in Chinese]. Hunan Journal of Traditional Chinese Medicine 2020;36(3):161-164 [FREE Full text] [doi: 10.16808/j.cnki.issn1003-7705.2020.03.065]
23. Dan W, Ying Z. 多发性骨髓瘤临床治疗及中医辨治策略探讨 (Discussion on clinical treatment and TCM diagnosis and treatment strategies for multiple myeloma) [Article in Chinese]. Hubei Journal of Traditional Chinese Medicine 2018;40(8):10-13 [FREE Full text]
24. Yaling H, Songshan L, Hong C. 浅谈中医对多发性骨髓瘤的认识 [Article in Chinese]. 世界最新医学信息文摘 2019;19(20):237-243 [FREE Full text] [doi: 10.19613/j.cnki.1671-3141.2019.20.131]
25. Jung YY, Lee JH, Nam D, et al. Anti-myeloma effects of icariin are mediated through the attenuation of JAK/STAT3-dependent signaling cascade. Front Pharmacol 2018;9:531. [doi: 10.3389/fphar.2018.00531] [Medline: 29899697]
26. Torcasio R, Gallo Cantafio ME, Veneziano C, et al. Targeting of mitochondrial fission through natural flavanones elicits anti-myeloma activity. J Transl Med 2024 Feb 27;22(1):208. [doi: 10.1186/s12967-024-05013-0] [Medline: 38413989]
27. Kim C, Lee SG, Yang WM, et al. Formononetin-induced oxidative stress abrogates the activation of STAT3/5 signaling axis and suppresses the tumor growth in multiple myeloma preclinical model. Cancer Lett 2018 Sep 1;431:123-141. [doi: 10.1016/j.canlet.2018.05.038] [Medline: 29857127]

XSL•FO
RenderX

28. Xinyi T, Jun W, Jie W, et al. 基于单细胞水平的补肾中药调控骨重建作用机制探讨肾主骨生髓的科学内涵 (Exploring scientific connotation of kidney dominating bone and producing marrow regulating mechanism of bone remodeling by kidney-tonifying Chinese Medicine based on single cell level) [Article in Chinese]. Chinese Archives of Traditional Chinese Medicine 2024;42(12):171-175 [FREE Full text] [doi: 10.13193/j.issn.1673-7717.2024.12.035]
29. Yang X, Chen W, Sun H, et al. Baicalin promotes anti-tumor immunity in hepatocellular carcinoma through HIF-1α/lactate/CXCL9 axis. Biochem Pharmacol 2025 Nov;241:117157. [doi: 10.1016/j.bcp.2025.117157] [Medline: 40659129]
30. Shiau AL, Shen YT, Hsieh JL, Wu CL, Lee CH. Scutellaria barbata inhibits angiogenesis through downregulation of HIF-1 α in lung tumor. Environ Toxicol 2014 Apr;29(4):363-370. [doi: 10.1002/tox.21763] [Medline: 22331677]
31. Borsi E, Perrone G, Terragna C, et al. HIF-1α inhibition blocks the cross talk between multiple myeloma plasma cells and tumor microenvironment. Exp Cell Res 2014 Nov 1;328(2):444-455. [doi: 10.1016/j.yexcr.2014.09.018] [Medline: 25257607]
32. Adhikari S, Singh V, Nandi S, et al. UBR7 in concert with EZH2 inhibits the TGF-β signaling leading to extracellular matrix remodeling. Cell Rep 2024 Jul 23;43(7):114394. [doi: 10.1016/j.celrep.2024.114394] [Medline: 38923455]
33. Li Y, Gao Z, Wang Y, et al. Lysine methylation promotes NFAT5 activation and determines temozolomide efficacy in glioblastoma. Nat Commun 2023;14(1):4062. [doi: 10.1038/s41467-023-39845-z]
34. Li W, Xiang X, Cao N, et al. Polysaccharide of atractylodes macrocephala koidz activated T lymphocytes to alleviate cyclophosphamide-induced immunosuppression of geese through novel_mir2/CD28/AP-1 signal pathway. Poult Sci 2021 Jul;100(7):101129. [doi: 10.1016/j.psj.2021.101129] [Medline: 34058564]
35. Li L, Liu J, Xue H, et al. A TGF-β-MTA1-SOX4-EZH2 signaling axis drives epithelial-mesenchymal transition in tumor metastasis. Oncogene 2020 Mar;39(10):2125-2139. [doi: 10.1038/s41388-019-1132-8] [Medline: 31811272]
36. Li Y, Wang J, Zhou L, et al. DNMT1 inhibition reprograms T cells to NK-like cells with potent antitumor activity. Sci Immunol 2025 Mar 21;10(105):eadm8251. [doi: 10.1126/sciimmunol.adm8251] [Medline: 40117344]
37. Wang X, Wang Y, Xie M, et al. Hypermethylation of CDKN2A CpG island drives resistance to PRC2 inhibitors in SWI/SNF loss-of-function tumors. Cell Death Dis 2024 Nov 5;15(11):794. [doi: 10.1038/s41419-024-07109-3] [Medline: 39500892]
38. Coira IF, Rincón R, Cuendet M. The multiple myeloma landscape: epigenetics and non-coding RNAs. Cancers (Basel) 2022 May 10;14(10):2348. [doi: 10.3390/cancers14102348] [Medline: 35625953]
39. Xu H, Hu YW, Zhao JY, et al. MicroRNA-195-5p acts as an anti-oncogene by targeting PHF19 in hepatocellular carcinoma. Oncol Rep 2015 Jul;34(1):175-182. [doi: 10.3892/or.2015.3957] [Medline: 25955388]
40. Ji Y, Fioravanti J, Zhu W, et al. miR-155 harnesses Phf19 to potentiate cancer immunotherapy through epigenetic reprogramming of CD8+ T cell fate. Nat Commun 2019 May 14;10(1):2157. [doi: 10.1038/s41467-019-09882-8] [Medline: 31089138]
41. Kumar A, Raj A, Chaman Lal K. Intratumoral microbe correlated with expression of DNA methylation genes in hepatocellular carcinoma. Dig Dis Sci 2025 Aug 27:MID. [doi: 10.1007/s10620-025-09351-x] [Medline: 40864369]
42. 骨痹的诊断依据、证候分类、疗效评定——中华人民共和国中医药行业标准《中医内科病证诊断疗效标准》(ZY/T001.1-94) [Article in Chinese]. 辽宁中医药大学学报 2017;19(1):224 [FREE Full text]
43. Declaration of Helsinki – Ethical principles for medical research involving human participants. World Medical Association. 2024. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki/World [accessed 2026-02-20]
44. Huiying X, Dongju H, Yongmin W, et al. 活血逐瘀方联合化疗治疗多发性骨髓瘤临床疗效及对机体免疫功能的影响 (Clinical observation of Huoxue Zhuyu decoction combined with chemotherapy in treatment of multiple myeloma and its effect on immune function) [Article in Chinese]. Chinese Archives of Traditional Chinese Medicine 2020;38(6):160-164 [FREE Full text] [doi: 10.13193/j.issn.1673-7717.2020.06.038]
45. Jianling Q, Fei Q. 补肾活血方辨治多发性骨髓瘤化疗引起的周围神经病变的疗效观察 (Efficacy of the Bushen Huoxue recipe in treating chemotherapy-induced peripheral neuropathy in multiple myeloma) [Article in Chinese]. Shanghai Medical and Pharmaceutical Journal 2025;46(11):45-49 [FREE Full text]
46. Wages NA, Chiuzan C, Panageas KS. Design considerations for early-phase clinical trials of immune-oncology agents. J Immunother Cancer 2018 Aug 22;6(1):81. [doi: 10.1186/s40425-018-0389-8] [Medline: 30134959]
47. Shin W, Kim MG, Kim A. Comparison of international guidelines for early-phase clinical trials of cellular and gene therapy products. Transl Clin Pharmacol 2022 Mar;30(1):13-23. [doi: 10.12793/tcp.2022.30.e2] [Medline: 35419313]
48. You Z, Xiangyin K, Jingting J. 表观遗传修饰在肿瘤免疫治疗中的意义 (Significance of the epigenetic modifications in cancer immunotherapy) [Article in Chinese]. Chinese Journal of Cancer Biotherapy 2016;23(5):727-732 [FREE Full text] [doi: 10.3872/j.issn.1007-385X.2016.05.024]

Abbreviations

CRF: case report form
EZH2: enhancer of zeste homolog 2
GCP: Good Clinical Practice
H3K27me3: trimethylated histone H3 at lysine 27
KPR: kidney-tonifying, phlegm-resolving, and blood stasis–removing

XSL•FO
RenderX

MM: multiple myeloma
NK: natural killer
PHF19: PHD finger protein 19
PRC2: polycomb repressive complex 2
RCT: randomized controlled trial
SAE: serious adverse event
TCM: Traditional Chinese Medicine
TGF-β: transforming growth factor-β
TME: tumor microenvironment
Treg: regulatory T cell

Edited by A Schwartz; submitted 22.Oct.2025; peer-reviewed by J Hou, Y Guo; revised version received 26.Jan.2026; accepted 29.Jan.2026; published 05.Mar.2026.

Please cite as:
Sun X, Zhou Y, Wang Y, Dai Y, Zhu W, Chen H
Kidney-Tonifying, Phlegm-Resolving, and Blood Stasis–Removing Therapy for Multiple Myeloma: Protocol for a Randomized Controlled Trial on Epigenetic and Immune Modulation
JMIR Res Protoc 2026;15:e86322
URL: https://www.researchprotocols.org/2026/1/e86322
doi:10.2196/86322

© Xiaoqi Sun, Yongming Zhou, Yujue Wang, Youya Dai, Wenwei Zhu, Hailin Chen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Validating a Multibiomarker Panel for the Assessment of Quantity and Quality of Plant Foods in the Diet (PLAENTI): Protocol for a Parallel Group–Designed Randomized Controlled Trial

Victor Schmalle[1], MSc; Julia Renz[1], MSc; Paola G Ferrario[1], Dr.; Stephanie Seifert[1], Dr.; Ann Katrin Engelbert[1], Dr.; Oliver Wittek[1], Dr.; Benedikt Merz[1], Dr.; Lorraine Brennan[2], PhD; Claudine Manach[3], Dr.; Manuela J Rist[1], Dr.; Achim Bub[1], Prof. Dr.

[1]Department of Physiology and Biochemistry of Nutrition, Max Rubner-Institut, Haid-und-Neu-Straße 9, Karlsruhe, Germany

[2]Institute of Food and Health, School of Agriculture and Food Science, University College Dublin (UCD), Dublin, Ireland

[3]Human Nutrition Unit, INRAE, Université Clermont Auvergne, Clermont-Ferrand, France

Corresponding Author:
Victor Schmalle, MSc
Department of Physiology and Biochemistry of Nutrition, Max Rubner-Institut, Haid-und-Neu-Straße 9, Karlsruhe, Germany

Abstract

Background: Although a high intake of plant foods is often considered healthy, some plant foods can be detrimental to health. Reliable dietary assessment is crucial to examine the relationship between diet and disease. Current dietary assessment methods rely on self-reported intake data, which are subject to bias. Objective measurement using biomarkers of food intake could mitigate this problem. However, single biomarkers of food intake have limitations as well. Combining several biomarkers of food intake into a multibiomarker panel could attenuate these limitations and allow for an accurate, objective dietary assessment.

Objective: The PLAENTI study aims to validate a multibiomarker panel for the assessment of quantity and quality of plant foods in the diet.

Methods: PLAENTI is a randomized controlled trial with 4 arms in a parallel design. Metabolically healthy adults (≥18 years old) were enrolled in the study. The study consisted of 1 week of run-in, with a standardized diet low in healthful plant foods for all participants; 2 weeks of a dietary intervention according to the assigned arm; and 1 week of washout, during which participants returned to their habitual diet. During the intervention, the participants' diet consisted of either a low, medium, or high proportion of healthful plant foods or a high proportion of unhealthful plant foods in the diet according to the assigned arm. The arm that received a high proportion of healthful plant foods served as the control. All food was provided based on energy-adjusted menu plans. During the visits, anthropometry and body composition were assessed, and blood samples were collected. Throughout the study, participants collected multiple urine samples (24-hour urine, evening and morning spot urine) and stool samples. Blood and urine samples will be analyzed by liquid chromatography-mass spectrometry to determine biomarker levels for the validation of a multibiomarker panel.

Results: After receiving approval from the ethics committee, recruitment began, and the first screening visit took place in November 2023. Between January and August 2024, of the 66 enrolled participants, 59 (31 female, 28 male) successfully completed the study, and their urine, blood, and stool samples are available for analysis. PLAENTI was conducted in 5 waves with a maximum of 16 participants enrolled in each wave. The mean age of the study population was 45.5 (SD 18.4) years, the mean BMI was 24.8 (SD-3.9) kg/m², and the mean total energy expenditure was 2464 (SD 440) kcal.

Conclusions: PLAENTI was conducted in a highly controlled and standardized manner, yielding samples and data that will be used to examine whether the quantity and quality of plant foods in the diet can be assessed using a multibiomarker panel. Successful validation of the multibiomarker panel would enable its application for objective dietary assessment.

Trial Registration: German Clinical Trials Register DRKS00032738; https://drks.de/search/en/trial/DRKS00032738

International Registered Report Identifier (IRRID): DERR1-10.2196/77571

(JMIR Res Protoc 2026;15:e77571) doi:10.2196/77571

KEYWORDS

biomarkers of food intake; dietary assessment; multibiomarker panel; plant-based diet; randomized controlled trial; targeted metabolomics; standardized diet

XSL•FO
RenderX

Introduction

A high intake of plant foods within a diet is often considered healthy. However, there are also plant foods that can be detrimental to health. For this reason, plant-based diet indices have been developed, which distinguish between "healthful" and "unhealthful" plant foods and form 3 categories: a general plant-based diet index, a healthful plant-based diet index, and an unhealthful plant-based diet index [1,2]. This distinction between healthful (eg, whole grains, unprocessed fruits and vegetables, nuts, and seeds) and unhealthful (eg, refined grains, sugar-sweetened beverages, and unfavorable vegetable oils) plant foods made it possible to obtain more detailed data on the relationship between diet and disease [1,2]. To date, nutritional epidemiology has relied on self-reported intake data, collected by tools such as food frequency questionnaires (FFQs), 24-hour recalls, and weighed or estimated diet records, which are prone to bias [3-6]. Self-reported data are subjective and influenced by factors such as the social desirability of food items, how individuals estimate portion sizes, the training of subjects, and the interview itself [3,7]. Well-known inaccuracies of plant food assessment include overreporting of fruit and vegetable intake and underreporting of oil and nut consumption [8-11]. The reliability of data may be enhanced by repeated measurements, improvement of assessment tools, or country-specific adaptation [7,12]. On the other hand, repeated measurements or higher-validity tools, such as 7-day dietary records, will increase the burden on participants [12].

There is growing evidence that biomarkers of food intake (BFIs) are a valuable tool for improving the accuracy of dietary assessment by introducing objective measurement [6,13-17]. However, BFIs also have limitations. There are overlaps between different foods with the same biomarker, and only a few food-specific biomarkers are known. Additionally, the concentration of some BFIs in biospecimens can vary between individuals, depending on factors such as age, sex, physical activity, and body composition [18]. Validation of candidate BFIs is of key importance and still scarce [15,19]. Combining several biomarkers into a multibiomarker panel (MBP) may mitigate the limitations of single BFIs and self-reported intake data [13,20]. It has been shown to increase the specificity and robustness of dietary assessment for individual foods (eg, banana, cocoa, and walnut) and for food groups such as total fruit intake [21-24]. However, dietary patterns, rather than individual nutrients or foods, have become a prominent topic in epidemiology, and several associations have been discovered between dietary patterns and noncommunicable diseases [25-28]. To apply BFIs and MBPs in epidemiological research, it is necessary to study and validate BFIs for the habitual diets of individuals or dietary patterns within populations. Accordingly, MBPs that include specific BFIs, BFIs for food groups, and BFIs that overlap between several food items could be promising to cover dietary patterns [29-34].

Cuparencu et al [35] recently reviewed validation criteria for BFIs and the general concepts of objective dietary assessment and listed known BFI candidates and their utility levels, as well as potential MBPs. The review also proposes research gaps. These include the identification of new MBPs and the discrimination of, for example, animal versus plant foods or healthful versus unhealthful foods. Additionally, more data on the interplay between BFI kinetics, frequency of food consumption, sampling schemes, sampling time windows, and different biospecimens are needed [35]. Furthermore, most studies still focus on biomarker discovery, often with untargeted analytical methods [33,36]. Beckmann et al [37] measured an MBP consisting of 54 BFI candidates in the urine samples of 15 participants who received a 3-day menu plan. The researchers were able to discriminate between the 3 dietary patterns that mimicked common UK diets [37]. Ultimately, more research is needed to unlock the full potential of BFIs and MBPs. Intervention studies specifically designed to measure BFIs in various samples from a diverse group of individuals consuming exact amounts and types of food, while considering dose- and time-response aspects, are necessary.

The JPI HDHL–funded project Plant*Intake* (acronym for "Combining biomarker panels and dietary intake data for improved assessment of healthful/unhealthful plant food intake") aims to adapt the plant-based diet indices described by Satija et al [1] to better fit to European diets and common foods (European plant-based diet indices). Furthermore, an MBP reflecting the quantity and nutritional quality (healthful vs unhealthful) of plant foods in the diet will be derived using intake data and samples from previous studies. Validation of MBPs is paramount for future use. Therefore, the developed MBP will be validated with data from the intervention study PLAENTI ("Validating a multibiomarker panel for the assessment of quantity and quality of plant foods in the diet"), which is described in this study protocol.

The aim of the PLAENTI study is to validate an MBP for the assessment of the quantity and nutritional quality of plant foods in an omnivorous diet. Accordingly, the primary objective of the validation study PLAENTI is to test whether the quantity of plant foods within the diet can be correctly estimated by measuring levels of selected urinary and blood biomarkers that are part of the MBP. A secondary objective is to test whether a difference in nutritional quality (healthful vs unhealthful) of plant foods in the diet can also be detected with the MBP. The effect of potential confounders (age, sex, physical activity, and body composition) on the precision of estimating plant food intake using the MBP will be assessed. Finally, the MBP will be determined in both 24-hour urine and spot urine samples. This will allow a comparison of the different sample types (24-hour urine collection, evening and morning spot urine, and different pools of these spot urines) to assess their reliability in reflecting dietary patterns through biomarker analysis. To address these objectives, participants received a 2-week dietary intervention with different proportions and nutritional qualities of plant foods, while urine and blood samples were collected at several time points.

Methods

Study Design

The PLAENTI study was conducted at the Study Center for Human Nutrition of the Max Rubner-Institut, Federal Research Institute of Nutrition and Food in Karlsruhe, Germany, from

XSL•FO
RenderX

January 9 to August 15, 2024. PLAENTI was designed as a randomized controlled trial with 4 arms. A parallel-group design was chosen to efficiently conduct the study and allow for direct comparison of results between the arms. A total of 66 participants were enrolled and randomized to 1 of 4 arms, stratified by sex. For logistic reasons, the study was conducted in 5 waves, each with a maximum of 16 participants. Each wave lasted a total of 4 weeks and started with a 1-week run-in period during which all participants consumed the same diet consisting of a low proportion of healthful plant foods to achieve comparable baseline levels of biomarkers. For the 2-week intervention period, participants were assigned to 1 of 4 interventions and received a diet that varied in the quantity and nutritional quality of plant foods to mimic different dietary patterns. The intervention was followed by a 1-week washout period, during which the participants returned to their habitual diet.

The study was carried out in an outpatient setting. Participants collected urine and stool samples and prepared and consumed food at home. Visits to the study center took place once a week on an identical schedule for all participants. Participants were asked to fast for at least 10 hours beforehand and underwent a series of standardized examinations and measurements described in detail below. Standard operating procedures were developed and tested for feasibility for all interviews, examinations, measurements, sample collections, sample preparations, and storage. Interviewers and examiners were extensively trained before and supervised during data collection. An overview of the study procedure is shown in Figure 1.

Figure 1. Overview of the study procedure. The figure shows the time sequence of run-in, intervention (either low, medium, or high in healthful or high in unhealthful plant foods), and washout, with corresponding visits to the study center and sample collections at home, as well as the schedule of visits to the study center and deliveries (A, B, and C). BIA: bioelectrical impedance analysis; FFQ: food frequency questionnaire; GI questionnaire: gastrointestinal symptom questionnaire.

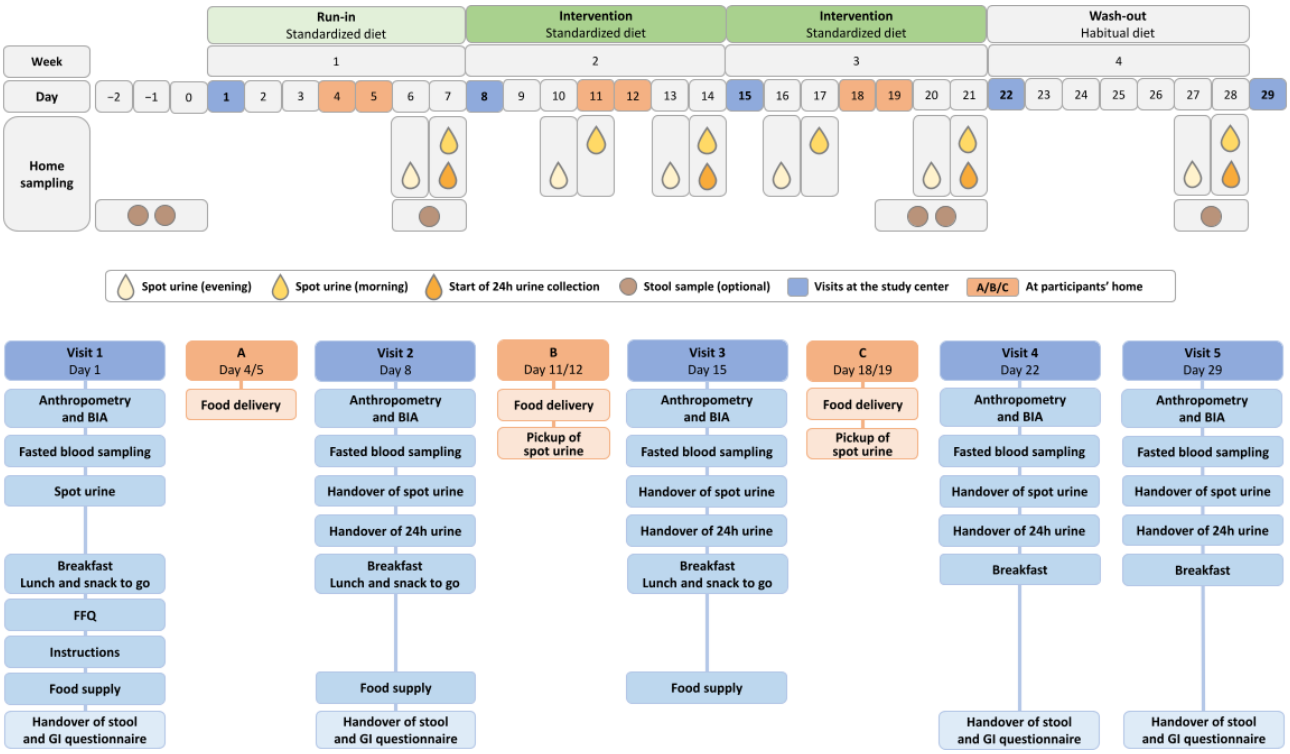

Prior to PLAENTI, an internal feasibility study called prePLAENTI was conducted (unpublished; registered at the German Clinical Trials Register: DRKS00031484, available at [38]). The purpose of prePLAENTI was to evaluate the acceptability of the standardized diets for participants and the feasibility of the planned study procedures. Participant feedback and process evaluation by the study center were used to ensure that PLAENTI can be conducted efficiently.

Recruitment

Female and male volunteers aged 18 years or older were eligible to participate. Detailed inclusion and exclusion criteria are listed in Textbox 1. Medical history was obtained using a health status questionnaire, and a basic physical examination was performed.

XSL•FO
RenderX

Textbox 1. Eligibility criteria of the PLAENTI study.

Inclusion criteria

- Females and males, 18 years or older
- Volunteers who are able to comply with the study terms
- Parameters of liver and kidney function, carbohydrate and lipid metabolism, showing no pathological significance
- Nonsmokers
- Omnivorous diet
- Written informed consent to participate in the study

Exclusion criteria

- Diseases affecting the absorption, digestion, metabolism, or excretion of nutrients
- Acute or chronic infectious diseases
- Tumors requiring therapy
- History of bariatric surgery
- History of gastrointestinal tract surgery affecting the absorption, digestion, metabolism, or excretion of nutrients
- History of pancreas surgery
- Cholecystectomy
- Acute or regular intake of medications resulting in energy or nutrient malabsorption
- Allergy or intolerance to any food or food ingredient in the study diet
- Use of dietary supplements <2 weeks prior to the study period
- Use of antibiotics <3 months prior to the study period
- Blood donation during or <4 weeks prior to the study period
- Pregnancy and breastfeeding
- Illegal drug use
- Alcohol or drug abuse
- Participation in other clinical trials that may affect the outcome of this study
- Individuals living in a care facility
- Individuals deprived of their liberty by a court or administrative order
- Individuals who are not expected to comply with the terms of the study

Participants were recruited in Karlsruhe, Germany, and the surrounding area. Posters and flyers were distributed at local and regional universities and public institutions (eg, libraries), doctors' offices, pharmacies, and health care facilities. In addition, recruitment was carried out via local/regional print media and social media in accordance with the requirements of the "Social Media Guideline" [39]. In addition, invitation emails were sent to potentially eligible participants listed in the study center's participant database to inform them about the study. First, persons interested in study participation were informed about the nature and scope of the trial in a detailed telephone screening. Then, potential participants were asked to visit the study center for a screening that included basic physical examinations and a detailed information session.

Allocation

The allocation sequence was generated by a biostatistician who was not involved in any further steps during the ongoing study and had no direct contact with the participants. Participants were randomly assigned to 1 of the 4 intervention arms with a 1:1:1:1 allocation by repeatedly using the "sample" function of the R software (R Core Team). Specifically, cluster randomization was used with clusters of different sizes. The clusters were heterosexual couples on the one hand and individual persons on the other. We randomized within waves for logistic reasons and stratified by sex to assess effects for females and males independently. The result of randomization was a sequential list of intervention arms with the corresponding study identifier of participants. The randomization plan is reproducible.

Blinding

Due to the nature of the dietary intervention, neither the participants nor the research staff responsible for carrying out the intervention were blinded to the intervention allocation. Only the study staff responsible for data collection, sample processing, and analysis were blinded.

XSL•FO
RenderX

Intervention

During the intervention, participants received an omnivorous diet that varied in the quantity and nutritional quality of plant foods according to their assigned arm: (1) low proportion of healthful plant foods (LH), (2) medium proportion of healthful plant foods (MH), (3) high proportion of healthful plant foods (HH), or (4) high proportion of unhealthful plant foods (HU). The intervention HH served as the control, as the diet implemented here comes closest to the intake of plant foods recommended by international dietary guidelines (eg, the "healthy reference diet" proposed by the EAT-*Lancet* Commission on Healthy Diets from Sustainable Food Systems [40]).

The intervention arms were designed to mimic dietary patterns similar to those described by the original plant-based diet indices [1]. To better reflect European eating habits, slight adjustments were made to the categorization of plant foods as healthful or unhealthful, and plant-based alternatives were included. Whole grains, fruits, vegetables, legumes, nuts, vegetable oils, tea, and coffee were categorized as healthful. Refined grains, fried potato products, fruit juices, sugar-sweetened beverages, sweets, and desserts were categorized as unhealthful. As literature on the nutritional quality of potatoes does not distinguish between different preparation methods, there is a lack of evidence to classify cooked, baked, or roasted potatoes as healthful or unhealthful. For this reason, this food group was categorized as "neutral." Plant-based alternatives were also classified as neutral due to their heterogeneous nature and inconsistent evidence on health associations. Animal-based foods were also viewed as neutral, and no BFI candidates for animal food were considered for the intended biomarker analysis.

For this study, a total of 35 base menu plans with an average energy content of 2000 (SD 50) kcal were calculated with the software PRODI (version 7.2, Nutri-Science GmbH). These base menu plans were later energy-adjusted to create the menu plans for the participants as described in the following section. Each menu plan consisted of 4 meals. In total, 7 menu plans were calculated for the run-in period and 7 for each intervention arm. During both intervention weeks, the same 7 daily menu plans were used. However, the sequence of these menu plans differed between the first and the second week, resulting in a variation in the foods consumed in relation to the sampling times. This should help to test whether the MBP is robust and can correctly assign participants to their intervention based on the overall dietary pattern, rather than by single dishes.

The main metric used to calculate the menu plans was the proportion of plant food servings while adhering as closely as possible to a macronutrient ratio of 50/30/20 energy percent for carbohydrates/fat/protein. Table 1 shows the average proportions of plant foods in the final menu plans of the run-in period and the 4 arms during the intervention. The base menu plans and the definition of serving sizes can be found in Multimedia Appendix 1.

Table . Proportions of plant foods in the calculated menu plans. Proportions of plant foods (in %) in the diet of the run-in and the 4 arms during the intervention by the metrics of servings, weight, and calories.

Metrics	Run-in	LH[a]	MH[b]	HH[c]	HU[d]
Proportion of plant foods of total servings (in %)	39	43	62	84	82
Proportion of plant foods of total weight (in %)	51	47	61	86	79
Proportion of plant foods of total calories (in %)	50	50	67	87	85

[a]LH: low proportion of healthful plant foods.

[b]MH: medium proportion of healthful plant foods.

[c]HH: high proportion of healthful plant foods.

[d]HU: high proportion of unhealthful plant foods.

For each meal, a time window for consumption was defined: breakfast (6 AM to 10 AM), lunch (11:30 AM to 2:30 PM), snack (10 AM to 5 PM), and dinner (5 PM to 8:30 PM). Alcohol, extra sugar, and dietary supplements were prohibited, while coffee, black tea, and green tea were limited to a total of two cups (=300 mL) per day. All participants could freely consume 100 mL of cow milk daily to add to coffee or tea.

Prior to each of the 5 study waves, the base menu plans were individually energy-adjusted to the calculated total energy expenditure (TEE) of each participant (see the following section), and the corresponding recipes were created. If the given portion sizes were too large, the participants were allowed to reduce the amounts of ingredients evenly to maintain the proportion of plant foods. In this case, they had to weigh back the leftover food. For this purpose, all participants were provided with kitchen scales from the same manufacturer in order to ensure comparability. If the participants were still hungry, they could eat an additional snack consisting of white bread and cream cheese, with low impact on plant biomarkers. Any other modifications to the intervention were prohibited.

Participants prepared and consumed meals and collected urine and stool samples at home, carrying out the intervention as part of their daily life. For this purpose, the study center fully supplied them with all the necessary food in the required

XSL•FO
RenderX

quantities and provided them with printed and digital recipe folders containing all the information, instructions, recipes, and daily logs. All recipes included a list of the individual energy-adjusted values of the ingredients and step-by-step instructions on how to cook or prepare the meals. The daily logs had to be completed during the run-in and the intervention. Here, participants were asked to record the exact time they ate each meal, the amount and type of tea or coffee they drank, and how much of the additional snack, if any, they consumed. There was also space for notes, mostly about permitted deviations from the recipes, such as the exact amount of food that was not eaten and weighed back.

Participants were closely monitored throughout the study. Personal contact was the main mode of communication, with 2 contacts per week (visits to the study center and food deliveries to the participants' homes) as well as the possibility of contacting the study center by email on weekdays or the study hotline (SMS text message and phone call) at any time.

Energy Adjustment

The menu plans were linearly energy-adjusted using Excel (Microsoft Corp) to create the individual menu plans for the participants. Total energy expenditure (TEE, in kcal) of each participant (see section Total Energy Expenditure Calculation) divided by the total kcal of the base menu plan yields a factor. This factor was then multiplied with the weight (in grams) of each food item in the base menu plan so that the plant proportions and macronutrient ratios are not altered. Resulting values ≤10 were mathematically rounded in increments of 1; values >10 in increments of 5; values >50 in increments of 10. The food items pretzel rolls (frozen), pretzels (frozen), schnitzel (vegan and pork, frozen), eggs, and crispbread were rounded to whole pieces. Wraps, bread rolls, and bagels were rounded to 0.5 piece increments.

Food Logistics

Food purchase, preparation of food packages, and delivery to the participants were carried out based on the energy-adjusted menu plans by the study center staff. Food was purchased at wholesale markets and supermarkets. The selection of food items remained the same for each of the 5 waves and was not adapted to the seasons. Food items that are not available year-round, such as strawberries, were only used frozen.

All shelf-stable foods were delivered to participants before the first study day. Frozen foods were delivered once a week. Fruits, vegetables, and refrigerated foods (dairy and meat products) were supplied twice a week, at the visit to the study center and via home delivery.

Anthropometry and Body Composition Assessment

Anthropometric data were assessed at all visits using standard methods. Measurements were taken in underwear, without shoes. Waist and hip circumferences were measured to the nearest 0.1 cm using the ergonomic measuring tape seca 201 (seca GmbH & Co KG). Three measurements were taken for each circumference, and the averages were recorded for analysis. Body composition was assessed by bioelectrical impedance analysis (BIA) using the seca mBCA 515. Body weight was measured to the nearest 0.05 kg, also using the seca mBCA 515, and height was measured to the nearest 0.01 m, using the seca 285 height measuring device. Weight and height were used to calculate the respective BMI (in kg/m^2).

Total Energy Expenditure Calculation

The TEE was calculated to enable adjustment of the menu plans to the individual energy requirements of each participant. To determine the TEE, resting energy expenditure (REE) and activity energy expenditure (AEE) were calculated. For REE calculation, the fat-free mass (FFM), which was determined as part of the BIA measurement, was used in the formula of Armbruster et al [41]:

$$\text{REE (kcal)} = 21.14 \times \text{FFM (kg)} + 345.5$$

The AEE was determined using accelerometers (move 4, movisens GmbH). The accelerometers were placed on the participants' hips at the screening visit and were then worn for 1 week. The software *SensorManager* (version 1.16.1, movisens GmbH) was used to program the sensor and read out the acquired data, and the software *DataAnalyzer* (version 1.15.1, movisens GmbH) was used to analyze the data and calculate the required parameters.

The participants' TEE was calculated with the estimated REE, the AEE measured by accelerometry, and the diet-induced thermogenesis (DIT), which is assumed as 10% of the sum of REE and AEE [42]:

$$\text{TEE (kcal)} = \text{REE (kcal)} + \text{AEE (kcal)} + \text{DIT (kcal)}$$

Habitual Diet Assessment (30-Day Food Frequency Questionnaire)

The validated 57-item FFQ applied in the German Health Interview and Examination Survey for Adults 2008-2011 (DEGS) of the Robert Koch-Institut was used to assess the habitual diet of the participants [9]. The FFQ includes questions about the frequency of consumption and usual portion sizes of different food groups consumed by the participants in the 4 weeks prior to study entry. The questionnaire was given and explained to the participants at visit 1 at the study center, with the request to complete it at home and return it at visit 2.

Urine and Blood Samples for Biomarker Analysis

Analytical Method

During the study, several urine and blood samples were collected to measure BFIs by liquid chromatography-mass spectrometry (LC-MS). The LC-MS method is being developed independently of the PLAENTI study in the Plant*Intake* project. It aims to quantitatively analyze the BFIs described in the list of biomarkers that can be found in Multimedia Appendix 2. The samples of PLAENTI will be analyzed using the finalized analytical method.

Blood Samples

Blood samples were collected at every visit after at least 10 hours of fasting. Participants were placed on an examination couch in the supine position. After cutaneous disinfection, a tourniquet was moderately applied (safety blood collection device: Venofix Safety, B. Braun SE), and blood samples were

XSL•FO
RenderX

taken from an antecubital vein by experienced medical study staff. Blood was drawn into ethylenediaminetetraacetic acid (EDTA)- or lithium-heparin-containing S-Monovette tubes (Sarstedt AG & Co KG) for plasma collection and into anticoagulant-containing serum-gel S-Monovette tubes (Sarstedt AG & Co KG) for serum collection.

After venipuncture, the filled tubes for plasma preparation were immediately centrifuged (20 °C, 10 minutes, 2500× *g*, Heraeus Multifuge X3R, Thermo Fisher Scientific) and kept on ice until aliquoted into microtubes. Erythrocytes (from lithium-heparin tubes) were washed with 0.9% sodium chloride solution and pipetted into CryoPure tubes (Sarstedt AG & Co KG). Tubes for serum preparation were allowed to clot for 30 minutes at room temperature, then centrifuged (20°C, 10 minutes, 2500×*g*) and kept on ice until aliquoted into microtubes or CryoPure tubes. Microtubes containing EDTA plasma, lithium-heparin-plasma, or serum were frozen at −20°C and then stored at −70°C until analysis. CryoPure tubes containing erythrocytes or serum for long-term storage were frozen at −20°C and then stored in the cryogenic storage tanks at −194°C.

Urine Samples

For urine sampling, participants received detailed oral instructions from trained study staff and a folder with written manuals for the collection procedures, including a calendar of sampling times. They were also given a log to record the exact time of sampling, any difficulties, and, for females, whether they were menstruating.

The 24-hour urine samples were collected from the morning before until the morning of visits 2, 3, 4, and 5. Participants were asked to start the urine collection after their first morning void, which was collected as a morning spot urine sample. Participants were provided with 3 urine collection containers (2 L capacity each, Sarstedt AG & Co KG) and cool bags with thermal packs. During collection, participants were asked to store their urine at home in the refrigerator or in a cool bag with the prechilled thermal packs to keep the collected urine permanently cool. This cool bag was also used to ensure the chilled transport of the collected urine to the study center on visit days. At the study center, the volume of the collected 24-hour urine was determined. The urine was mixed, and 15 mL of it was pipetted into a tube, centrifuged (4 °C, 10 minutes, 2500× *g*), and kept on ice until aliquoted into microtubes or CryoPure tubes. Microtubes containing 24-hour urine were frozen at −20°C and then stored at −70°C until analysis. CryoPure tubes for long-term storage of aliquots were frozen at −20°C and then stored in cryogenic storage tanks at −194°C.

Spot urine samples were collected from participants at home the evening and morning before the start of the 24-hour urine collection. Additional spot urine samples were collected halfway through each intervention week, also in the evening and the following morning (see Figure 1). The evening sample had to be the last urine before going to bed, and the morning sample had to be the first morning void. Spot urine samples were collected in 100 mL polypropylene collection cups with an integrated transfer unit in the screw caps and directly transferred to 10 mL vacuum tubes (V-Monovette, Sarstedt AG & Co KG). Participants were asked to store the filled vacuum tubes in the refrigerator immediately after sampling. A cool bag with prechilled thermal packs was used to transport the samples to the study center. At the study center, the vacuum tubes were centrifuged (4°C, 10 minutes, 2500× *g*) and kept on ice until aliquoted into microtubes or CryoPure tubes. Microtubes containing spot urine were frozen at −20°C and then stored at −70°C until analysis. CryoPure tubes for long-term storage of aliquots were frozen at −20°C and then stored in cryogenic storage tanks at −194°C.

Stool Samples

The collection of stool samples was optional for participants. Their willingness to provide a stool sample was queried at the screening visit. Participants who agreed to provide stool samples were given detailed oral instructions and a written manual on the sampling procedure and were provided with stool collection kits containing a fecal sample collection aid (servoprax GmbH), a collection cup (Sarstedt AG & Co KG), a disposable spatula, and a 60 mL twist-seal bag (VWR International GmbH), as well as a thermal pack for transport, gloves, and zip-log bags for packaging. Participants collected two cherry-sized stool samples (approximately 5 g) on different days in a 72-hour window before visit 1 and visit 4, and one cherry-sized stool sample in a 48-hour window before visit 2 and visit 5. Samples were frozen at −20°C immediately after collection at the participants' home, delivered frozen to the study center, and then stored at −70 °C until aliquoted. Samples were thawed and manually mixed inside the collection bags before taking aliquots of approximately 500 mg each for further processing. Aliquots were frozen immediately at −70°C until further analysis.

Participants who collected stool samples also completed a short questionnaire on gastrointestinal symptoms at visits 1, 2, 4, and 5, concerning the previous week or 2 weeks, respectively. The questionnaire covers stool frequency as well as the occurrence and frequency of the following symptoms: abdominal pain, heartburn, burping, watery stool or diarrhea, lumpy or hard stool, constipation, feeling full after food intake, gastrointestinal sounds, flatulence, nausea, and increased or decreased hunger or appetite.

Stool samples were collected for exploratory analyses. First, we can examine the effects of different quantities and qualities of plant foods on the gut microbiota. Another approach that can complement PLAENTI is to detect plant taxa in stool samples by genetic analysis, as described by Petrone et al [43].

Compliance

Care was taken to ensure that only the most conscientious individuals were included in an attempt to keep the dropout rate as low as possible. During the telephone screening and the information session at the screening visit, potential participants were given detailed information about the study procedure and, in particular, the effort required to take part in the study. In addition to the inclusion and exclusion criteria, potential participants were asked about other issues such as food aversions, usual eating habits, and the compatibility of their daily routine with participation.

To improve participants' compliance during the ongoing study, a number of measures were taken. As detailed in the

XSL•FO
RenderX

"Intervention" subsection, food consumption was monitored using daily logs. Participants had to record their exact mealtimes, as well as the exact amount of uneaten food, in these daily logs. There was also a urine collection log to document the time of sampling. There was very close personal contact between the study staff and the participants during the weekly visits and food deliveries. In addition, the team of the study center was always available by email or telephone to directly respond to any questions or concerns that participants might have and to provide them with optimum support. Noncompliance occurred when relevant samples were missing, instructions regarding the intervention were not followed, or medical events (medication, illness) conflicted with participation. Noncompliance led to exclusion.

Adverse Events

Safety monitoring included the recording, assessment, and follow-up of adverse events and the determination of the following safety parameters in blood at the first and last visit: gamma glutamyl transferase, glutamic oxaloacetic transaminase, bilirubin (total), creatinine, uric acid, glucose, triglycerides, and total cholesterol (all in serum), and a blood count. Only commercially available food was used, which was fully provided by the study center. Allergies to any food or food ingredient in the study diet were assessed in advance and resulted in exclusion (see Textbox 1). Therefore, no serious adverse events related to the consumption of the intervention diet were expected. Participants with unacceptable adverse events or acute illnesses that interfered with study participation were excluded from the study. This decision was made individually by the study physicians.

Statistics

Outcomes

The primary outcome of the study is an MBP measured in urine and blood samples from participants in the arms LH, MH, and HH at defined time points during the 2-week intervention period. The analysis metric for the primary outcome is the change in levels of biomarkers contained in the MBP from visit 2 to visit 4, considered as a combined panel value. The variables will be aggregated and described by average values and SDs.

The secondary outcome is the MBP measured in urine and blood samples at defined time points in the two arms, HH and HU. The analysis metric for the secondary outcome is also the change in levels of biomarkers contained in the MBP from visit 2 to visit 4, considered as a combined panel value. The variables will be aggregated and described by average values and SDs.

Moreover, the comparability of the biomarker levels in the different types of urine samples (24-hour urine, spot urine, and pooled spot urine) will be investigated to gain insight into the applicability of spot urine sampling in epidemiological studies. Other outcomes of an exploratory nature will also be measured. Metabolomic analyses will be performed by measuring different classes of metabolites in the urine and blood samples. For microbiome parameters, the analysis metrics are alpha diversity (number and uniformity of species distribution within a microbiome), beta diversity (dissimilarity of the composition of different microbiomes), and abundance of individual taxa in different microbiomes. The results of the FFQ will be used to characterize the participants' habitual diet.

Sample Size

The MBP to be validated in this study will be developed at a later stage of the Plant*Intake* project, so it was not possible to calculate the sample size prior to submission of the project proposal. As a Williams test is performed in the main statistical analysis, this was used as the basis for a sensitivity analysis. The calculation was performed using the R package "MCPAN" (R version 4.1.0, R Core Team). The option "Type: Williams" was used. Specifically, for a 3-arm design (for the primary outcome) with n=15 per arm and expected values for the MBP of 1 (LH), 2 (MH), and 3.8 (HH), and a homoscedastic variance of 2.5, a power of 0.81 was obtained [21,44-46]. A total of 60 participants were expected to complete the study, with even distribution between the groups.

Based on experience with previous studies, including a study with low, medium, and high intake of fruits and vegetables by Watzl et al [47] and the more recent JPI FoodBAll study (Universal Trial Number: U1111-1177-1536), a dropout rate of 12% was assumed. Therefore, in order to analyze a total of 60 participants, the goal was to enroll 68 participants in total.

Statistical Methods

The aim of the primary statistical analysis is to detect a trend in the MBP as a function of the proportion of healthful plant foods consumed by the participants. Here, the values of the MBP determined at visit 4 are adjusted with the baseline values from visit 2. The MBP values are compared between the interventions using the Williams test. The comparison is made for the 3 diets with different proportions of healthful plant foods (low, medium, or high). For simplicity, it is assumed that the primary outcome is normally distributed and has a homogeneous variance. The statistical hypothesis is that the primary outcome will follow a trend for the ordered categories of low, medium, and high intake of healthful plant foods. Group HH, with a high proportion of healthful plant foods in the diet, serves as a control. The following 3 average values of the MBP are assumed for the 3 intervention groups: $\mu 1$, $\mu 2$, and $\mu 3$. All intervention groups have a comparable sample size. An average different response is expected in the group with low and medium consumption compared to the high consumption (control) of healthful plant foods. The global null hypothesis is formulated as $H0: \mu 1=\mu 2=\mu 3$ versus the alternative hypothesis $H1: \mu 3\geq\mu 2\geq\mu 1$, with at least one strict inequality. When modeling the MBP, adjustment is made for the following potential confounding variables: age, sex, body composition (assessed by BIA), and physical activity (assessed by accelerometry). The MBP is considered to be validated if a significant monotonic trend is observed ($P<.05$).

As the main secondary analysis, mean MBP values will be compared between the HH and HU intervention arms, characterized by similar proportions of plant foods but differing in nutritional quality (healthful vs unhealthful). Group differences will first be assessed using an independent-samples *t* test, followed by linear regression models incorporating the same set of adjusting variables.

XSL•FO
RenderX

As a further aim, biomarker levels will be compared between different urine samples (24-hour urine, single spot urine, and pooled spot urine). Temporal variability during the 2-week intervention will be characterized descriptively and by variance components. For this purpose, mixed-effects linear models with biomarker levels as the dependent variable, urine sample type and time as fixed effects, and participants as random effects will be applied. From these models, within- and between-participant variance and intraclass correlation coefficients will be estimated. Additional exploratory analyses can be conducted on metabolomic profiles, metabolic pathways, and further physiological or health-related variables. These outcomes will be analyzed using linear regression models. To account for multiple testing in high-dimensional metabolomics, *P* values will be adjusted using the Benjamini-Hochberg procedure.

In the case of missing data, the first step is to investigate what type of missing data occurs. If biomarker levels are below a reliable detection limit, these are modeled as "censored data." A distinction will also be made as to whether the missing data are random or not, and they will be treated accordingly. We will not include the data of participants in subsequent analyses if an unacceptable adverse event occurred or if they were noncompliant.

Data Management

The consortium partners have developed a data management plan that is updated regularly. Data will be published in accordance with the Code of Conduct of the German Research Foundation (DFG) "Guidelines for Safeguarding Good Research Practice" [48] and regardless of the magnitude and direction of the effects. We aim to publish and disseminate data and associated metadata in anonymized form according to the FAIR (Findability, Accessibility, Interoperability, and Reuse) guidelines.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki. The described study procedures are in accordance with the protocol submitted to the ethics committee prior to recruitment. It was registered in the German Clinical Trials Register (DRKS00032738; available at [49]) and approved by the Ethics Committee of the State Medical Chamber of Baden-Wu¨rttemberg, Stuttgart, Germany, in September 2023 (F-2023-085). The Universal Trial Number is U1111-1298-1562. Consent to participate in the study, the processing of personal data, and the analysis of pseudonymized research data were documented by signing the informed consent form. A copy of the signed consent form was kept by the study management. The consent form was approved by the ethics committee. The participants received €50 (approximately US $59.5) as compensation for their time and effort.

Personal information about potential and enrolled participants is subject to medical confidentiality. The Max Rubner-Institut adheres to the rules of the State Medical Chamber, the General Data Protection Regulation, and the Federal Data Protection Act when processing data. Participant data and sample material will only be passed on to directly involved project partners in pseudonymized form. A list of the names assigned to study identifiers is stored separately at the study center and is subject to technical and organizational measures to ensure that the personal data cannot be assigned to the study participant by unauthorized persons. Thereby, only those responsible at the trial site have access to the personal data.

Results

Study Timeline

PLAENTI was conducted in 5 waves, from January 9 (first participant, first visit) to August 15 (last participant, last visit), 2024, each with a maximum of 16 participants, covering each intervention arm. After the final approval from the ethics committee was received on September 14, 2023, recruitment of potential participants was started on November 3, 2023, and the first screening visit took place on November 20, 2023. As the study was conducted in waves, participants were recruited continuously until the last wave began. As part of the recruitment process, a total of 153 telephone screenings were conducted with interested individuals. A total of 91 potential participants attended the screening visit, of which 66 participants were enrolled in the study. Five of them took part in the first study wave (January 9 to February 8), 14 in the second (February 20 to March 21), 15 in the third (April 9 to May 8), 16 in the fourth (June 4 to July 4), and 16 in the fifth wave (July 15 to August 15). In total, 7 participants (5 males and 2 females) were excluded after allocation for various reasons. Accordingly, 59 participants successfully completed the PLAENTI study, and their blood, urine, and stool samples are available for analysis. Detailed information about the participant flow is shown in Figure 2.

XSL•FO
RenderX

Figure 2. Participant flow. This flowchart shows the retention of participants from recruitment to completion of the study, including reasons for exclusion in all steps. Furthermore, the numbers of female and male participants of those who successfully completed the study are given. HH: high proportion of healthful plant foods; HU: high proportion of unhealthful plant foods; LH: low proportion of healthful plant foods; MH: medium proportion of healthful plant foods.

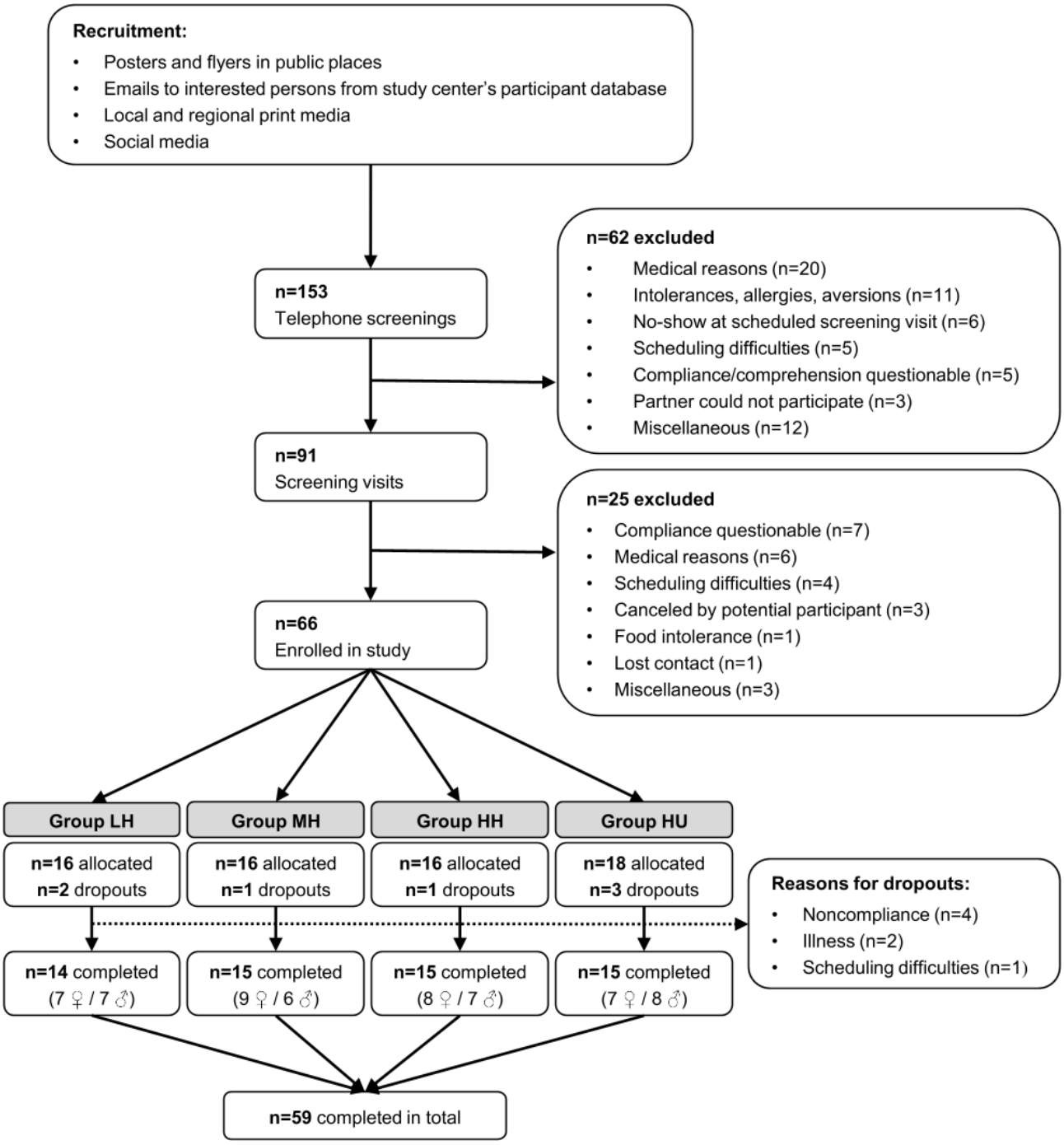

Study Population

Baseline characteristics of the study population are given in Table 2. The calculated mean REE of the participants was 1679 (SD 278) kcal, the measured mean AEE was 785 (SD 205) kcal, and the calculated mean TEE was 2464 (SD 440) kcal. The proportion of female and male participants was 52.5% and 47.5%, respectively.

XSL•FO
RenderX

Table . Baseline characteristics of the study population. Number of females and males as well as characteristics on the first study day for the total study population and the intervention arms.

Characteristics	Total (N=59)	LH[a] (n=14)	MH[b] (n=15)	HH[c] (n=15)	HU[d] (n=15)
Female, n	31	7	9	8	7
Male, n	28	7	6	7	8
Age (years), mean (SD)	45.5 (18.4)	50.9 (19.3)	46.9 (20.1)	42.6 (17.6)	42.0 (14.8)
Age (years), range	19 - 86	19 - 76	21 - 86	22 - 69	20 - 67
BMI (kg/m²), mean (SD)	24.8 (3.9)	25.7 (3.2)	24.0 (3.8)	24.2 (4.3)	25.3 (3.9)
TEE[e] (kcal), mean (SD)	2464 (440)	2402 (510)	2312 (338)	2551 (517)	2572 (272)

[a]LH: low proportion of healthful plant foods.

[b]MH: medium proportion of healthful plant foods.

[c]HH: high proportion of healthful plant foods.

[d]HU: high proportion of unhealthful plant foods.

[e]TEE: total energy expenditure.

Dietary Intervention

The calculation of menu plans resulted in a macronutrient proportion and fiber content of the diets as shown in Table 3.

Table . Macronutrient proportions and fiber content of the diets in the run-in period and the 4 intervention arms. The data shown are average values of the macronutrients and the fiber content of the menu plans calculated for the run-in and each arm.

Metrics	Run-in	LH[a]	MH[b]	HH[c]	HU[d]
Carbohydrates (in % of energy)	41	40	46	49	55
Fat (in % of energy)	35	35	33	34	31
Protein (in % of energy)	24	25	21	17	14
Fiber (total value in grams)	27	24	32	42	18

[a]LH: low proportion of healthful plant foods.

[b]MH: medium proportion of healthful plant foods.

[c]HH: high proportion of healthful plant foods.

[d]HU: high proportion of unhealthful plant foods.

Discussion

Addressing Research Gaps

Thanks to its carefully conceived experimental design, the PLAENTI study will be able to validate an MBP that reflects the quantity and quality of plant foods within omnivorous diets. Additionally, the study can address various aspects of BFI research, including several of the described validation criteria for BFIs. For example, aspects of dose-response can be assessed by comparing the 3 intervention arms with different quantities of healthful plant foods in the diet. Aspects of time-response can be explored by comparing biomarker levels at several time points during the intervention. Since data on physical activity, TEE, age, sex, and body composition were recorded, the effect of known confounders on BFIs can be considered. Furthermore, collecting 24-hour and spot urine samples, as well as blood samples, allows for a direct comparison of the informative value of data and burden on study participants of different sample types. The research gaps that PLAENTI can address align closely with the roadmap for using BFIs as assessment tools proposed by Cuparencu et al [35].

Study Design

A crossover design may be considered advantageous for validation studies due to the reduced interindividual variability [50]. On the downside, a crossover design would significantly increase the duration of a standardized diet, which may result in higher dropout rates and a reduced likelihood of good compliance [51]. Because of this, we chose a parallel-group design in PLAENTI, even though this might lead to differences in baseline biomarker levels or confounders between participants. However, the run-in period will have reduced these baseline variations. The 2-week intervention is expected to be

XSL•FO
RenderX

sufficient to establish a stable biomarker pattern, as the sampling windows for BFIs with long elimination half-lives are approximately 6 days [35].

We decided not to adjust macronutrient or fiber content, since our aim was not to investigate the health effects of the intervention, but instead we intended to mimic realistic dietary patterns. This will also increase the acceptability of the dietary intervention. The macronutrient proportions, as shown in Table 3, varied between the intervention arms, as expected. There are also differences in fiber content between arms, particularly between the HH and HU arms. This is naturally due to unhealthful plant foods having lower fiber contents than the healthful plant foods. To make the dietary intervention more acceptable in the HU arm, up to 10% of total servings were allocated to fresh fruits and vegetables instead of just unhealthful plant foods. This decision was based on the findings of the feasibility study prePLAENTI, which showed that participants found an extreme diet containing no healthful plant foods less acceptable (data not published; registered at the German Clinical Trial Register: DRKS00031484). Additionally, we prioritized fruits and vegetables that are commonly consumed and well accepted in the European Union. While this approach may limit overall diversity, it increases compliance and helps to obtain robust results for these plant foods. Eventually, we want to validate whether the MBP developed in the Plant*Intake* project can reflect common European diets. Overall, the study design of PLAENTI should be feasible for many research facilities. Ultimately, more whole-supply studies are needed to advance nutrition research. Exact intake data in the context of a complex diet allow us to consider matrix effects and biological relevance to obtain the most robust results regarding nutrition, foods, and diets. By providing this extensive protocol, we aim to facilitate the implementation of whole-supply studies by others, thereby advancing the quality and standard of nutritional interventions.

Comparison to Prior Work

In related research, participants are often recruited from previous studies [16,34] or characterized by specific traits. For example, previous studies have only included postmenopausal females [36] or individuals with untreated prehypertension [52]. In other cases, samples and intake data from past cohorts are used [33]. In PLAENTI, we tried to best represent the general population to allow for a reliable validation of the MBP. To achieve this, we stratified by sex and included a wide age range.

Another important consideration is the accurate determination of the TEE. This ensures that participants receive the appropriate amount of food that meets their caloric needs. Ideally, the TEE would be determined directly using doubly labeled water, as Playdon et al [36] did. However, this is a time-consuming and expensive method, and therefore, we calculated it using REE and AEE instead. For the calculation of the REE, we used the formula of Armbruster et al [41], which was developed at the Max Rubner-Institut, and its applicability has been demonstrated [41]. The AEE was determined individually by accelerometry to provide more precise results. This is particularly important given the expected wide range of AEE levels, such as between older individuals and highly active young individuals. Often, the AEE is estimated by using the physical activity level (PAL). For example, Oude Griep et al [52] used a fixed PAL of 1.4 for all participants, which is comparable to the mean TEE/REE ratio of 1.47 of the individualized values in PLAENTI. Although our method did not allow for an exact determination, it provided a very good estimate of TEE. Just as there are different possible approaches to determining TEE, there are also different approaches to adjust menu plans to TEE values. One common approach is to use menu plans with predefined energy level increments and match participants to the plan that most closely aligns with their individual caloric needs [52-54]. Our individual approach is more accurate with the downside of increased logistical effort.

In a study design with a free-living population, compliance is one of the most critical factors, and a multitude of considerations must be made. In PLAENTI, we did not measure compliance objectively, but we implemented several methods to improve it. The most significant measure is the accurate supply and delivery of all required foods. The aforementioned manuals and schedules will have improved compliance, and the logs deliver a trust-based control. It was important to consider where burden was necessary for accurate results while making study participation feasible and acceptable. Measures described in the literature to facilitate compliance include providing 90% of TEE through strict menu plans and 10% from a restricted free-choice food list [52], prescribing or allowing a certain amount of specific foods within a habitual diet [55], or calculating menu plans based on each participant's habitual diet to provide an individualized controlled version of it [36]. In PLAENTI, due to the validation purpose of the study, we chose a rather strict approach, where all foods in the menu plans were mandatory. Based on participant feedback from prePLAENTI, we permitted a total of up to two cups of coffee, black tea, or green tea per day to allow for some flexibility. Alcohol and dietary supplements were prohibited. The approach of allowing limited flexibility for coffee consumption was also used by Oude Griep et al [52], while they and Playdon et al [36] also allowed some leeway for alcohol consumption. The unique design of PLAENTI offers new methods and considerations for implementation in future whole-supply studies.

Strengths

One of the major strengths of the PLAENTI study is that the study team purchased and distributed all of the food, thereby standardizing the diet. Although this was time-consuming and labor-intensive, it offered numerous advantages. First, participants were provided with the exact amount of food they needed. This helps reduce food waste, increase adherence to energy-adjusted menu plans, and prevent overconsumption. Furthermore, food safety (eg, cold chain) and quality could be ensured. In the study of Oude Griep et al [52], food was ordered online and delivered directly to the participants by grocers. While this approach is easier to implement, the study team loses precise control over food quality and is unable to split units of certain food items. In contrast, Playdon et al [36] provided precooked, portioned, weighed, refrigerated, or frozen food 3 times per week. Any unconsumed food was returned. While this approach is highly standardized and the best for ensuring compliance, it also requires the most effort, as well as appropriate cooking facilities and staff capacity.

XSL•FO
RenderX

Another strength of PLAENTI is the collection of different biological samples (24-hour urine, spot urine, and blood) at multiple time points. This increases the likelihood of covering a wide variety of BFIs, including those with short elimination half-lives, and allows for comparison of the informative values of different sample types. Furthermore, recording meal consumption and sampling times makes it possible to examine the temporal relationship between food intake and sample collection in the context of the analytical detection of BFIs. Several publications suggest that the applicability of spot urine samples may be comparable to 24-hour urine collections, and that both may be superior to blood samples [37,56,57]. Spot urine has the advantage of reduced burden, cost, and logistical effort compared to 24-hour urine collections [37,56]. We used vacuum tubes for spot urine sample collection to ensure optimal stability of the compounds [37]. The system was easy to use and allowed hygienic collection by the participants. Overall, PLAENTI may provide evidence for the use of spot urine samples as an alternative to 24-hour urine collections in BFI research. This could enable the widespread use of BFIs in nutrition-related studies.

Limitations

One limitation of the PLAENTI study is that selection bias may have been introduced by the requirement that participants live within a reasonable driving distance (up to 40 km) of the study center for food deliveries. Additionally, the topic of the study may have particularly attracted individuals interested in nutrition. According to the eligibility criteria, participants were nonsmokers and metabolically healthy, which likely created a bias toward healthier lifestyles. This may also be reflected in the lower average BMI (24.8 kg/m²) of the study population compared to the general population of Germany (26.0 kg/m²) [58]. In this regard, the study population may differ from the general population. However, we do not expect this to significantly impact the results of the study.

Another limitation is that the composition of food varies, and secondary plant metabolites of fruits and vegetables are particularly affected [59-61]. This is also evident from the fact that metabolite concentrations can vary significantly even among apples from the same tree [62]. Since PLAENTI was divided into 5 waves from January to August, seasonal variations of the food metabolome may affect the results obtained. However, this also allows testing whether the MBP is robust under real conditions. By covering all intervention arms in each study wave, the effect of food seasonality on the comparability of results should be reduced [51]. To minimize the influence of other factors on food composition, food was purchased from the same grocers. Through access to the wholesale market, fruits and vegetables were mostly purchased in units from a single farm, which should reduce differences in the food metabolome. By providing detailed recipes (including preparation method, cooking time, and heat settings) and supplying food twice a week, we expect to have minimized the effects of different storage and processing. As not all foods are permanently available over an 8-month period, we replaced unavailable foods with the most comparable alternative. This mainly concerned processed animal products, which were not the focus of this trial.

Conclusion and Prospects

A total of 59 participants successfully completed the PLAENTI study, and their urine, blood, and stool samples are available for analysis to validate the MBP. The study was conducted in a highly standardized setting to validate the MBP in terms of robustness, dose-response, sampling time, and reproducibility. Successful validation of the MBP would enable its application as a tool for objective dietary assessment in epidemiology. This, in turn, could contribute to a better understanding of the relationship between diet and disease.

Acknowledgments

We want to sincerely thank Martina Broßart, Tanja Gadau, Eva Hoch, Petra Kunzmann, and Gisela Schultheiss for medical examinations, sample management, and sample processing; Nelly Birk, Sophie Büttner-Koch, and Ulrike Rossi for recruitment, telephone screening, and for taking care of volunteers; Selina Busch, Cornelia Hodapp, and Heike Schäfer for assistance in food logistics and supply; Hanna Haidar for support regarding the food frequency questionnaire; our interns Til Becker, Paula Lauff, Sophia Veit, and Lara Walz for support in various tasks; Ann-Kathrin Henk, MD, and Seher Emir-Stückler, MD, for taking medical care of participants; and all past and present members of the PlantIntake consortium. We would like to especially thank the volunteers for their motivated and conscientious participation in the study.

Funding

This project has received funding from the Federal Ministry of Research, Technology and Space (BMFTR), represented by the Project Management Agency in the German Aerospace Center (DLR-PT), the French Agency of National Research (ANR, grant 21-HDH2-0001-03), and the Health Research Board (HRB) under the umbrella of the European Joint Programming Initiative "A Healthy Diet for a Healthy Life" (JPI HDHL) and of the ERA-NET cofund HDHL-INTIMIC (GA No. 727565 of the EU Horizon 2020 Research and Innovation Programme).

Data Availability

In accordance with the informed consent of the PLAENTI participants, pseudonymized study data will only be provided to directly involved project partners. Data will be processed and published upon obtaining the results of the ongoing analyses.

XSL•FO
RenderX

Authors' Contributions

MJR, AB, SS, PGF, and AKE developed the study concept and design; VS, JR, OW, BM, LB, and CM contributed to the study design. VS, JR, SS, and OW were responsible for study organization and data acquisition. VS and JR drafted the manuscript. PGF provided statistical expertise. All authors contributed to the refinement of the study protocol and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Menu plans.
[XLSX File, 67 KB - resprot_v15i1e77571_app1.xlsx]

Multimedia Appendix 2
Biomarker list.
[XLSX File, 26 KB - resprot_v15i1e77571_app2.xlsx]

Checklist 1
SPIRIT checklist.
[PDF File, 232 KB - resprot_v15i1e77571_app3.pdf]

References

1. Satija A, Bhupathiraju SN, Rimm EB, et al. Plant-based dietary patterns and incidence of type 2 diabetes in US men and women: results from three prospective cohort studies. PLoS Med 2016 Jun;13(6):e1002039. [doi: 10.1371/journal.pmed.1002039] [Medline: 27299701]
2. Satija A, Bhupathiraju SN, Spiegelman D, et al. Healthful and unhealthful plant-based diets and the risk of coronary heart disease in U.S. adults. J Am Coll Cardiol 2017 Jul 25;70(4):411-422. [doi: 10.1016/j.jacc.2017.05.047] [Medline: 28728684]
3. Bailey RL. Overview of dietary assessment methods for measuring intakes of foods, beverages, and dietary supplements in research studies. Curr Opin Biotechnol 2021 Aug;70:91-96. [doi: 10.1016/j.copbio.2021.02.007] [Medline: 33714006]
4. Kipnis V, Midthune D, Freedman L, et al. Bias in dietary-report instruments and its implications for nutritional epidemiology. Public Health Nutr 2002 Dec;5(6A):915-923. [doi: 10.1079/PHN2002383] [Medline: 12633516]
5. Bingham SA, Gill C, Welch A, et al. Comparison of dietary assessment methods in nutritional epidemiology: weighed records v. 24 h recalls, food-frequency questionnaires and estimated-diet records. Br J Nutr 1994 Oct;72(4):619-643. [doi: 10.1079/bjn19940064] [Medline: 7986792]
6. Ravelli MN, Schoeller DA. Traditional self-reported dietary instruments are prone to inaccuracies and new approaches are needed. Front Nutr 2020;7:90. [doi: 10.3389/fnut.2020.00090] [Medline: 32719809]
7. Buso ME, Boshuizen HC, Naomi ND, et al. Relative validity of habitual sugar and low/no-calorie sweetener consumption assessed by food frequency questionnaire, multiple 24-h dietary recalls and urinary biomarkers: an observational study within the SWEET project. Am J Clin Nutr 2024 Feb;119(2):546-559. [doi: 10.1016/j.ajcnut.2023.11.019] [Medline: 38043866]
8. Bogers RP, Dagnelie PC, Westerterp KR, et al. Using a correction factor to correct for overreporting in a food-frequency questionnaire does not improve biomarker-assessed validity of estimates for fruit and vegetable consumption. J Nutr 2003 Apr;133(4):1213-1219. [doi: 10.1093/jn/133.4.1213] [Medline: 12672945]
9. Haftenberger M, Heuer T, Heidemann C, Kube F, Krems C, Mensink GBM. Relative validation of a food frequency questionnaire for national health and nutrition monitoring. Nutr J 2010 Sep 14;9:36. [doi: 10.1186/1475-2891-9-36] [Medline: 20840739]
10. Steinemann N, Grize L, Ziesemer K, Kauf P, Probst-Hensch N, Brombach C. Relative validation of a food frequency questionnaire to estimate food intake in an adult population. Food Nutr Res 2017;61(1):1305193. [doi: 10.1080/16546628.2017.1305193] [Medline: 28469546]
11. Garcia-Aloy M, Hulshof PJM, Estruel-Amades S, et al. Biomarkers of food intake for nuts and vegetable oils: an extensive literature search. Genes Nutr 2019;14(1):7. [doi: 10.1186/s12263-019-0628-8] [Medline: 30923582]
12. Yuan C, Spiegelman D, Rimm EB, et al. Relative validity of nutrient intakes assessed by questionnaire, 24-hour recalls, and diet records as compared with urinary recovery and plasma concentration biomarkers: findings for women. Am J Epidemiol 2018 May 1;187(5):1051-1063. [doi: 10.1093/aje/kwx328] [Medline: 29036411]
13. Garcia-Aloy M, Rabassa M, Casas-Agustench P, Hidalgo-Liberona N, Llorach R, Andres-Lacueva C. Novel strategies for improving dietary exposure assessment: multiple-data fusion is a more accurate measure than the traditional single-biomarker approach. Trends Food Sci Technol 2017 Nov;69:220-229. [doi: 10.1016/j.tifs.2017.04.013]

XSL•FO
RenderX

14. Brouwer-Brolsma EM, Brennan L, Drevon CA, et al. Combining traditional dietary assessment methods with novel metabolomics techniques: present efforts by the Food Biomarker Alliance. Proc Nutr Soc 2017 Nov;76(4):619-627. [doi: 10.1017/S0029665117003949] [Medline: 29137687]
15. Dragsted LO, Gao Q, Scalbert A, et al. Validation of biomarkers of food intake—critical assessment of candidate biomarkers. Genes Nutr 2018 Dec;13(1):14. [doi: 10.1186/s12263-018-0603-9]
16. Hu Y, Wang M, Willett WC, et al. Calibration of citrus intake assessed by food frequency questionnaires using urinary proline betaine in an observational study setting. Am J Clin Nutr 2024 Jul;120(1):178-186. [doi: 10.1016/j.ajcnut.2024.05.011] [Medline: 38762186]
17. Kisi SB, Filskov Petersen C, Sand Andersen R, et al. Validity and reproducibility of a web-based dietary assessment tool: a cross-sectional study in an adult Danish population. J Nutr Sci 2025;14:e40. [doi: 10.1017/jns.2025.10010] [Medline: 40568623]
18. Rist MJ, Roth A, Frommherz L, et al. Metabolite patterns predicting sex and age in participants of the Karlsruhe Metabolomics and Nutrition (KarMeN) study. PLoS ONE 2017;12(8):e0183228. [doi: 10.1371/journal.pone.0183228] [Medline: 28813537]
19. Goerdten J, Muli S, Rattner J, et al. Identification and replication of urine metabolites associated with short-term and habitual intake of sweet and fatty snacks in European children and adolescents. J Nutr 2024 Nov;154(11):3274-3285. [doi: 10.1016/j.tjnut.2024.09.026] [Medline: 39332769]
20. Low DY, Mina TH, Sadhu N, et al. Metabolic variation reflects dietary exposure in a multi-ethnic Asian population. Nat Metab 2025 Sep;7(9):1939-1954. [doi: 10.1038/s42255-025-01359-x] [Medline: 40962943]
21. Vázquez-Manjarrez N, Weinert CH, Ulaszewska MM, et al. Discovery and validation of banana intake biomarkers using untargeted metabolomics in human intervention and cross-sectional studies. J Nutr 2019 Oct 1;149(10):1701-1713. [doi: 10.1093/jn/nxz125] [Medline: 31240312]
22. Garcia-Aloy M, Llorach R, Urpi-Sarda M, et al. A metabolomics-driven approach to predict cocoa product consumption by designing a multimetabolite biomarker model in free-living subjects from the PREDIMED study. Mol Nutr Food Res 2015 Feb;59(2):212-220. [doi: 10.1002/mnfr.201400434] [Medline: 25298021]
23. Garcia-Aloy M, Llorach R, Urpi-Sarda M, et al. Novel multimetabolite prediction of walnut consumption by a urinary biomarker model in a free-living population: the PREDIMED study. J Proteome Res 2014 Jul 3;13(7):3476-3483. [doi: 10.1021/pr500425r] [Medline: 24882253]
24. McNamara AE, Walton J, Flynn A, Nugent AP, McNulty BA, Brennan L. The potential of multi-biomarker panels in nutrition research: total fruit intake as an example. Front Nutr 2020;7:577720. [doi: 10.3389/fnut.2020.577720] [Medline: 33521031]
25. Cena H, Calder PC. Defining a healthy diet: evidence for the role of contemporary dietary patterns in health and disease. Nutrients 2020 Jan 27;12(2):334. [doi: 10.3390/nu12020334] [Medline: 32012681]
26. Poulsen SK, Due A, Jordy AB, et al. Health effect of the New Nordic Diet in adults with increased waist circumference: a 6-mo randomized controlled trial. Am J Clin Nutr 2014 Jan;99(1):35-45. [doi: 10.3945/ajcn.113.069393] [Medline: 24257725]
27. Siervo M, Lara J, Chowdhury S, Ashor A, Oggioni C, Mathers JC. Effects of the Dietary Approach to Stop Hypertension (DASH) diet on cardiovascular risk factors: a systematic review and meta-analysis. Br J Nutr 2015 Jan 14;113(1):1-15. [doi: 10.1017/S0007114514003341] [Medline: 25430608]
28. Estruch R, Ros E, Salas-Salvadó J, et al. Primary prevention of cardiovascular disease with a Mediterranean diet supplemented with extra-virgin olive oil or nuts. N Engl J Med 2018 Jun 21;378(25):e34. [doi: 10.1056/NEJMoa1800389] [Medline: 29897866]
29. Owen EJ, Patel S, Flannery O, Dew TP, O'Connor LM. Derivation and validation of a total fruit and vegetable intake prediction model to identify targets for biomarker discovery using the UK National Diet and Nutrition Survey. J Nutr 2021 Apr 8;151(4):962-969. [doi: 10.1093/jn/nxaa406] [Medline: 33484153]
30. Massey LK. Food oxalate: factors affecting measurement, biological variation, and bioavailability. J Am Diet Assoc 2007 Jul;107(7):1191-1194. [doi: 10.1016/j.jada.2007.04.007] [Medline: 17604750]
31. Brouwer-Brolsma EM, Brandl B, Buso MEC, Skurk T, Manach C. Food intake biomarkers for green leafy vegetables, bulb vegetables, and stem vegetables: a review. Genes Nutr 2020 Apr 9;15(1):7. [doi: 10.1186/s12263-020-00667-z] [Medline: 32272877]
32. Chen GC, Chai JC, Xing J, et al. Healthful eating patterns, serum metabolite profile and risk of diabetes in a population-based prospective study of US Hispanics/Latinos. Diabetologia 2022 Jul;65(7):1133-1144. [doi: 10.1007/s00125-022-05690-w] [Medline: 35357561]
33. Li Y, Xu Y, Sayec ML, et al. Development of metabolic signatures of plant-rich dietary patterns using plant-derived metabolites. Eur J Nutr 2024 Nov 28;64(1):29. [doi: 10.1007/s00394-024-03511-x] [Medline: 39604558]
34. Silva AM, Levy J, De Carli E, et al. Biomarker panels for fruit intake assessment: a metabolomics analysis in the ELSA-Brasil study. Metabolomics (Los Angel) 2024 Jul 29;20(4):88. [doi: 10.1007/s11306-024-02145-8] [Medline: 39073486]
35. Cuparencu C, Bulmuş-Tüccar T, Stanstrup J, La Barbera G, Roager HM, Dragsted LO. Towards nutrition with precision: unlocking biomarkers as dietary assessment tools. Nat Metab 2024 Aug;6(8):1438-1453. [doi: 10.1038/s42255-024-01067-y] [Medline: 38956322]

XSL•FO
RenderX

36. Playdon MC, Tinker LF, Prentice RL, et al. Measuring diet by metabolomics: a 14-d controlled feeding study of weighed food intake. Am J Clin Nutr 2024 Feb;119(2):511-526. [doi: 10.1016/j.ajcnut.2023.10.016] [Medline: 38212160]
37. Beckmann M, Wilson T, Lloyd AJ, et al. Challenges associated with the design and deployment of food intake urine biomarker technology for assessment of habitual diet in free-living individuals and populations—a perspective. Front Nutr 2020;7:602515. [doi: 10.3389/fnut.2020.602515] [Medline: 33344495]
38. DRKS00031484. German Clinical Trials Register. URL: https://drks.de/search/en/trial/DRKS00031484 [accessed 2026-02-07]
39. Richtlinie für Untersuchungen in den und unter Einsatz der Sozialen Medien (Soziale Medien Richtlinie) [in German]. Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.. 2021 Mar. URL: https://www.adm-ev.de/wp-content/uploads/2024/09/RL-Untersuchung-Soziale-Medien.pdf [accessed 2026-02-06]
40. Willett W, Rockström J, Loken B, et al. Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. Lancet 2019 Feb 2;393(10170):447-492. [doi: 10.1016/S0140-6736(18)31788-4] [Medline: 30660336]
41. Armbruster M, Rist M, Seifert S, et al. Metabolite profiles evaluated, according to sex, do not predict resting energy expenditure and lean body mass in healthy non-obese subjects. Eur J Nutr 2019 Sep;58(6):2207-2217. [doi: 10.1007/s00394-018-1767-1] [Medline: 29974226]
42. Westerterp KR. Diet induced thermogenesis. Nutr Metab (Lond) 2004 Aug 18;1(1):5. [doi: 10.1186/1743-7075-1-5] [Medline: 15507147]
43. Petrone BL, Aqeel A, Jiang S, et al. Diversity of plant DNA in stool is linked to dietary quality, age, and household income. Proc Natl Acad Sci U S A 2023 Jul 4;120(27):e2304441120. [doi: 10.1073/pnas.2304441120] [Medline: 37368926]
44. Bub A, Kriebel A, Dörr C, et al. The Karlsruhe Metabolomics and Nutrition (KarMeN) study: protocol and methods of a cross-sectional study to characterize the metabolome of healthy men and women. JMIR Res Protoc 2016 Jul 15;5(3):e146. [doi: 10.2196/resprot.5792] [Medline: 27421387]
45. Auclair S, Chironi G, Milenkovic D, et al. The regular consumption of a polyphenol-rich apple does not influence endothelial function: a randomised double-blind trial in hypercholesterolemic adults. Eur J Clin Nutr 2010 Oct;64(10):1158-1165. [doi: 10.1038/ejcn.2010.135] [Medline: 20683465]
46. Bader M, Lang T, Lang R, Hofmann T. Synephrine as a specific marker for orange consumption. J Agric Food Chem 2017 Jun 14;65(23):4853-4858. [doi: 10.1021/acs.jafc.7b01941] [Medline: 28532150]
47. Watzl B, Kulling SE, Möseneder J, Barth SW, Bub A. A 4-wk intervention with high intake of carotenoid-rich vegetables and fruit reduces plasma C-reactive protein in healthy, nonsmoking men. Am J Clin Nutr 2005 Nov;82(5):1052-1058. [doi: 10.1093/ajcn/82.5.1052] [Medline: 16280438]
48. Guidelines for Safeguarding Good Research Practice: Code of Conduct. : Deutsche Forschungsgemeinschaft; 2025 Jan 27. [doi: 10.5281/zenodo.14281892]
49. DRKS00032738. German Clinical Trials Register. URL: https://drks.de/search/en/trial/DRKS00032738 [accessed 2026-2-7]
50. Sibbald B, Roberts C. Understanding controlled trials: crossover trials. BMJ 1998 Jun 6;316(7146):1719. [doi: 10.1136/bmj.316.7146.1719] [Medline: 9614025]
51. Lichtenstein AH, Petersen K, Barger K, et al. Perspective: design and conduct of human nutrition randomized controlled trials. Adv Nutr 2021 Feb 1;12(1):4-20. [doi: 10.1093/advances/nmaa109] [Medline: 33200182]
52. Oude Griep LM, Frost G, Holmes E, Wareham NJ, Elliott P. Systems approach to investigate the role of fruit and vegetable types on vascular function in pre-hypertensive participants: protocol and baseline characteristics of a randomised crossover dietary intervention. Nutrients 2024 Sep 1;16(17):2923. [doi: 10.3390/nu16172923] [Medline: 39275238]
53. White K, Wu B, Pilla SJ, et al. Design and implementation of a time-restricted eating intervention in a randomized, controlled eating study. Nutrients 2023 Apr 20;15(8):1978. [doi: 10.3390/nu15081978] [Medline: 37111198]
54. Miller ER, Erlinger TP, Young DR, et al. Results of the diet, exercise, and weight loss intervention trial (DEW-IT). Hypertension 2002 Nov;40(5):612-618. [doi: 10.1161/01.hyp.0000037217.96002.8e] [Medline: 12411452]
55. Hill EB, Tang M, Long JM, et al. mini-MED: study protocol for a randomized, multi-intervention, semi-controlled feeding trial of a Mediterranean-amplified vs. habitual Western dietary pattern for the evaluation of food-specific compounds and cardiometabolic health. Trials 2024 Feb 2;25(1):101. [doi: 10.1186/s13063-024-07939-8] [Medline: 38302990]
56. Wilson T, Garcia-Perez I, Posma JM, et al. Spot and cumulative urine samples are suitable replacements for 24-hour urine collections for objective measures of dietary exposure in adults using metabolite biomarkers. J Nutr 2019 Oct 1;149(10):1692-1700. [doi: 10.1093/jn/nxz138] [Medline: 31240300]
57. Playdon MC, Sampson JN, Cross AJ, et al. Comparing metabolite profiles of habitual diet in serum and urine. Am J Clin Nutr 2016 Sep;104(3):776-789. [doi: 10.3945/ajcn.116.135301] [Medline: 27510537]
58. Körpermaße nach Altersgruppen und Geschlecht. Statistisches Bundesamt. 2019. URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Gesundheitszustand-Relevantes-Verhalten/Tabellen/liste-koerpermasse.html#104708 [accessed 2026-02-06]
59. Amarowicz R, Carle R, Dongowski G, et al. Influence of postharvest processing and storage on the content of phenolic acids and flavonoids in foods. Mol Nutr Food Res 2009 Sep;53(Suppl 2):S151-S183. [doi: 10.1002/mnfr.200700486] [Medline: 19065581]

XSL•FO
RenderX

60. Raffo A, Malfa GL, Fogliano V, Maiani G, Quaglia G. Seasonal variations in antioxidant components of cherry tomatoes (Lycopersicon esculentum cv. Naomi F1). J Food Compost Anal 2006 Feb;19(1):11-19. [doi: 10.1016/j.jfca.2005.02.003]
61. Serrano-García I, Domínguez-García J, Hurtado-Fernández E, et al. Assessing the RP-LC-MS-based metabolic profile of Hass avocados marketed in Europe from different geographical origins (Peru, Chile, and Spain) over the whole season. Plants (Basel) 2023 Aug 20;12(16):3004. [doi: 10.3390/plants12163004] [Medline: 37631215]
62. Awad MA, de Jager A, van Westing LM. Flavonoid and chlorogenic acid levels in apple fruit: characterisation of variation. Sci Hortic 2000 Mar;83(3-4):249-263. [doi: 10.1016/S0304-4238(99)00124-7]

Abbreviations

AEE: activity energy expenditure
BFI: biomarker of food intake
BIA: bioelectrical impedance analysis
DIT: diet-induced thermogenesis
EDTA: ethylenediaminetetraacetic acid
FFM: fat-free mass
FFQ: food frequency questionnaire
HH: high proportion of healthful plant foods
HU: high proportion of unhealthful plant foods
LC-MS: liquid chromatography-mass spectrometry
LH: low proportion of healthful plant foods
MBP: multibiomarker panel
MH: medium proportion of healthful plant foods
PAL: physical activity level
REE: resting energy expenditure
TEE: total energy expenditure

Edited by A Schwartz; submitted 15.May.2025; peer-reviewed by A Renteria-Mexia, IA Brownlee, S Casperson; revised version received 19.Dec.2025; accepted 29.Dec.2025; published 27.Feb.2026.

Please cite as:
Schmalle V, Renz J, Ferrario PG, Seifert S, Engelbert AK, Wittek O, Merz B, Brennan L, Manach C, Rist MJ, Bub A
Validating a Multibiomarker Panel for the Assessment of Quantity and Quality of Plant Foods in the Diet (PLAENTI): Protocol for a Parallel Group–Designed Randomized Controlled Trial
JMIR Res Protoc 2026;15:e77571
URL: https://www.researchprotocols.org/2026/1/e77571
doi:10.2196/77571

© Victor Schmalle, Julia Renz, Paola G Ferrario, Stephanie Seifert, Ann Katrin Engelbert, Oliver Wittek, Benedikt Merz, Lorraine Brennan, Claudine Manach, Manuela J Rist, Achim Bub. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Motor Imagery and Action Observation in Breast Cancer Survivors: Protocol for a Randomized Controlled Trial

María Pilar Arnal-Vallés[1*], MECCN, MSc; Nelia Soto-Ruiz[1,2*], PhD; Ana Beatriz Bays-Moneo[1*], PhD; Cristina García-Vivar[1,2*], PhD; Paula Escalada-Hernández[1,2*], PhD

[1]Department of Health Sciences, Public University of Navarre (UPNA), Avda. Barañain S/N., Pamplona, Spain

[2]Navarre Institute of Health Research (IdiSNA), Pamplona, Spain

[*]all authors contributed equally

Corresponding Author:
Nelia Soto-Ruiz, PhD
Department of Health Sciences, Public University of Navarre (UPNA), Avda. Barañain S/N., Pamplona, Spain

Abstract

Background: Breast cancer is the most common type of cancer in women worldwide, and its incidence is increasing. Although breast cancer survival is slowly increasing, related sequelae can persist after the disease has been treated. The main physical symptoms associated with breast cancer survival include pain, lymphedema, and associated functional limitations. Although multiple treatments are available for alleviating symptoms in breast cancer survivors, their effectiveness remains limited. Motor imagery (MI) and action observation (AO) therapies, which are based on the theory of motor simulation and are used in multiple fields with satisfactory results, have been proposed as alternatives for treating pain and improving mobility and strength.

Objective: This study aims to design, implement, and evaluate the effectiveness of a program combining MI and AO therapies to improve functionality and mobility and alleviate pain and lymphedema of the affected upper limb in women who have survived breast cancer.

Methods: A randomized controlled clinical study will be conducted in a sample of 108 participants who have experienced breast cancer and, as a result, have pain in the affected extremity, lymphedema, or loss of strength and/or mobility. The intervention group will include 54 participants managed with the MI and AO program (a combination of MI, AO, and mobility exercises), while the control group will consist of 54 women performing mobility exercises alone. Pain intensity, muscle strength, joint range, limb diameter, fear of movement, and imagery capability will be evaluated.

Results: The intervention is expected to yield improvements in pain intensity, joint range, muscle strength, and symptoms associated with lymphedema, among other outcomes. The study was funded in December 2023. The number of participants recruited as of manuscript submission is approximately 80, and data analysis has not yet started. These results will be published in 2026.

Conclusions: The implementation of an intervention based on MI and AO has the potential to positively impact female breast cancer survivors who face physical and psychological sequelae that interfere with their daily lives.

Trial Registration: ClinicalTrials.gov NCT07067710; https://clinicaltrials.gov/study/NCT07067710

International Registered Report Identifier (IRRID): DERR1-10.2196/85469

(JMIR Res Protoc 2026;15:e85469) doi:10.2196/85469

KEYWORDS

breast cancer; survivors; survivorship; rehabilitation; motor imagery; action observation; lymphedema; pain; mobility; functionality; strength; upper limb; kinesiophobia; range of motion; imagery ability; randomized; trial; protocol; exercises; oncology

Introduction

Worldwide, breast cancer accounts for approximately 23.8% of all malignant tumors in women, which establishes it as the most prevalent cancer in this population, with incidence rates continuing to rise [1]. Data from the International Agency for Research on Cancer of the World Health Organization indicate that, in 2022, a total of 2,296,840 new cases of breast cancer were diagnosed among women worldwide [1], with the 5-year global prevalence projected to rise to 8,178,393 cases [1].

The incidence of breast cancer continues to increase, and as mortality decreases, more patients live with the disease for longer, causing the prevalence to also increase [1,2]. According to the GLOBOCAN database, in 2020, the number of deaths worldwide among women with breast cancer was 684,996, representing 15.5% of all deaths from cancer. Improvements in

XSL•FO
RenderX

treatments and prevention strategies have allowed for a reduction in mortality, mainly in African countries, but global lifestyle changes have driven an increase in disease prevalence [3]. As a result, the number of breast cancer survivors worldwide is increasing [2], creating significant challenges for long-term follow-up and the management of persistent symptoms [4].

Many breast cancer survivors continue to live with physical and emotional sequelae as a result of the treatments received and the oncological process itself. These late and persistent side effects not only compromise quality of life and functionality but may also prevent survivors from integrating the cancer experience into their daily lives [5]. Among the most common physical sequelae are pain, lymphedema, and postmastectomy syndrome. The latter is characterized by a burning sensation, tightness, stabbing, and pain in the chest, armpit, upper medial part of the arm, and surgical scar, characteristics similar to those of phantom limb syndrome. The likelihood of developing lymphedema is greater in women undergoing axillary lymphadenectomy, sentinel node biopsy, or mastectomy [6]. Lymphedema causes a decrease in the functionality of the affected arm, negatively impacting quality of life. This condition is estimated to affect 35% of patients and is associated with pain and functional limitations in activities of daily living (ADLs) [7].

Although it is one of the most frequent symptoms, pain secondary to breast cancer has not been sufficiently investigated [8]. It has been ascribed to a number of etiologies, such as cervical radiculopathy, rotator cuff tendinopathy, postmastectomy syndromes, or neuropathies, as well as preexisting comorbidities (back pain, arthritis, osteoarthritis, and axillary lymphadenectomy accompanied by chemotherapy and/or radiotherapy), young age (18-39 years), fibromyalgia, and depression or anxiety about the possibility of experiencing chronic pain [9]. There is existing evidence suggesting that breast cancer is associated with pain that can become chronic, affecting both patients and survivors.

Various interventions have been developed in an attempt to improve functionality and indirectly reduce pain [10]. Nonpharmacological strategies used in physiotherapy include multicomponent training (aerobic exercise, flexibility, and strength), neuromuscular bandaging, manual lymphatic drainage, and multimodal therapy. However, in many cases, the pain persists after the therapies have been completed [11].

In light of this situation, it is necessary to assess the effectiveness of emerging therapies, such as motor imagery (MI) and action observation (AO). Both interventions are based on simulation theory, which suggests that imagined or observed actions activate brain regions similar to those involved in the actual execution of the movement [12]. Specifically, MI is based on the fact that imagining the performance of an action stimulates the same parts of the brain map involved in performing the action itself [12]. This therapy has proven useful in various fields, such as sports, psychology, rehabilitation, medicine, and education [13]. Studies have reported beneficial effects of MI on the range of motion, strength, pain, and functionality of the affected limb [14-20]. For example, in the case of phantom limb syndrome, patients have reported a significant reduction in pain, which even disappeared in some cases [15]. There is recent randomized controlled trial evidence suggesting that MI-based interventions can increase physical activity levels and self-efficacy in women with breast cancer, supporting the potential of MI as a behavioral and rehabilitation strategy [21].

AO, meanwhile, is based on the hypothesis that, when observing the performance of an action or movement in the first or third person, the neural circuits involved in the execution of the action are activated [12]. This mechanism has been closely linked to the mirror neuron system, which responds during both execution and AO, providing a neurophysiological substrate for observational learning [22,23]. In other words, observing an action engages brain processes similar to those used when performing the action itself [12]. Similar to MI, AO has been applied in different disciplines and has shown positive results in the reorganization of action, motor control, and functional rehabilitation [24].

Individually, both MI and AO have demonstrated effectiveness in improving muscle strength, joint range, pain intensity, and functionality [15,19,24]. However, there is current evidence suggesting that their combined use produces greater activation of motor brain regions than when applied individually [25], although studies assessing the effectiveness of this combined therapy remain scarce [26]. A systematic review and meta-analysis published in 2025 has reported that combining both therapies effectively improves motor functionality, gait speed, upper-extremity functionality, pain intensity, and fear of movement, especially in the upper extremities [27]. Despite their promising potential, the application of these techniques within oncology—and especially among breast cancer survivors—has been insufficiently investigated, leaving a significant gap in the evidence base.

Therefore, the aim of this project is to design, implement, and evaluate the effectiveness of a program combining MI and AO therapies to improve functionality and mobility and alleviate pain and lymphedema of the affected upper limb in women who have survived breast cancer.

Methods

Design

A randomized controlled trial will be conducted to test the following hypothesis: a combination of MI and AO therapies is effective for improving mobility and alleviating pain and lymphedema of the upper limb in breast cancer survivors.

The protocol follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [28] and has been registered on ClinicalTrials.gov (NCT07067710). Future publications of the study results will be reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) guidelines [29].

Setting

The study will be carried out in Navarra, an autonomous community and province in northern Spain. Information about the study will be disseminated through social networks and

XSL•FO
RenderX

newspapers, and collaboration will be requested from SARAY, the Navarra Breast Cancer Association, a nonprofit regional association that promotes comprehensive care for women who have experienced breast cancer and provides biopsychosocial support through different services (eg, psychology, physiotherapy, and accompaniment).

Study Participants and Eligibility Criteria

Female survivors of breast cancer who reside in Navarra will be invited to participate in the study. The inclusion criteria for participants in the study will be being female, being aged over 18 years, having a diagnosis of breast cancer, having completed any active cancer treatments (chemotherapy, radiotherapy, and/or immunotherapy), and being disease free at the time of data collection. Women undergoing hormonal therapy after treatment completion will also be included as hormonal treatment is not considered an active cancer therapy and is commonly prescribed among breast cancer survivors. Conversely, participants will be excluded if they have a diagnosis of cancer other than breast cancer or evidence of recurrence or metastasis requiring new treatment or are undergoing active cancer therapy (for recurrent or newly diagnosed disease).

Sample Size

To determine the sample size, we assumed a type I error of 0.05 and a type II error of 0.20 (80% power). The calculation was based on a minimum detectable difference of 1.5 points in pain intensity measured using the numerical rating scale (NRS). This minimum difference aligns with established minimal clinically important difference values for the NRS in breast cancer–related pain contexts, which range from 1.39 to 2.0 points [30-33].

An estimated SD of 2.61 was selected based on a conservative approach considering the variability reported in breast cancer populations using the NRS, where SDs range from 1.4 to 2.8 [34-37]. This conservative estimate ensures adequate statistical power even if data dispersion is greater than anticipated. Considering an estimated 10% dropout, the required sample size was 108 women, with 54 allocated to the control group (CG) and 54 allocated to the intervention group (IG). The power calculations were performed using the GRANMO online program (version 7.12; Programa de Registre Gironí del Cor).

Randomization, Allocation, and Blinding

The participants in the study will be randomly assigned to 1 of 2 groups: the IG (MI+AO+mobility exercises) or the CG (mobility exercises alone). All the participants will perform mobility exercises of the upper limb involving flexion, extension, abduction, adduction, internal rotation, and external rotation for approximately 20 minutes between 1 and 3 times a day 3 times a week for the duration of the intervention (6 weeks).

The randomization process will be carried out using the Research Randomizer tool and an allocation ratio of 1:1 with the aim of guaranteeing impartiality and validity in the study and ensuring a balanced distribution, simple randomization, and a reduction in biases while preserving the integrity and reliability of the results. The randomization list will be created by a member of the research team without contact with participants or access to personal or clinical information.

This is a single-blind study, as neither participants nor the researchers delivering the intervention can be blinded. Blinding will be maintained for the researchers responsible for collecting outcome data, administering evaluation instruments, and performing the statistical analyses.

Intervention: the IM-OA23 Program

The study intervention is part of the IM-OA23 program, a structured initiative designed to integrate MI and AO into rehabilitation strategies. Support for the intervention will be presented in video format and will be accompanied by instructions for performing the MI and OA therapies as well as the additional mobility exercises, with a video prepared to help participants perform the therapies.

The intervention will last 6 weeks, during which the content of the program will be progressively modified. During the first 2 weeks, the following anatomical movements will be visualized and imagined, with an approximate duration of 20 seconds per movement: (1) arm flexion from 0 to 180°, (2) arm extension from 0 to −10°, (3) arm abduction from 0 to 180°, (4) arm adduction from 0 to 45°, (5) internal arm rotation from 0 to 45° with the elbow flexed at 90°, (6) external arm rotation from 0 to 45° with the elbow flexed at 90°, and (7) forward and backward arm circumductions (Figures 1 and 2 [38]). In the third and fourth weeks, the participants will progress to performing the anatomical movements described above with a small weight, for example, a 1-kg dumbbell. The videos will depict a woman dressed in black in the frontal and lateral planes performing the described anatomical movements. In the final 2 weeks (fifth and sixth), the content will focus on ADLs through the following tasks: (1) reaching for a jar, (2) raising and lowering blinds, (3) drawing a curtain, (4) changing a light bulb, (5) stirring a pot, (6) scratching one's back, (7) stretching, (8) putting on and taking off a coat, (9) hugging someone, (10) cleaning a window pane, (11) turning the steering wheel of a car, and (12) throwing a ball.

XSL•FO
RenderX

Figure 1. Anatomical movements performed [38].

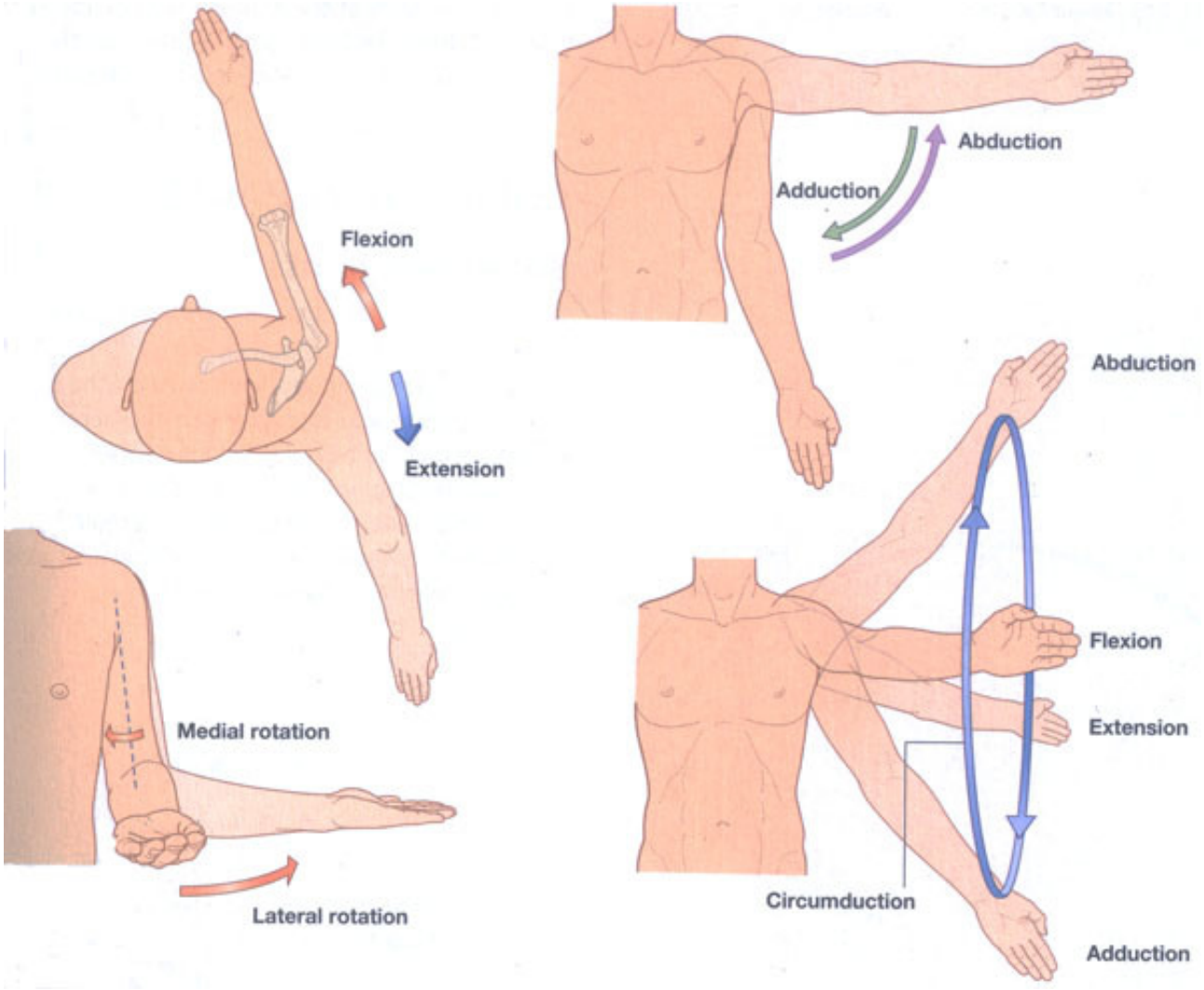

XSL•FO
RenderX

Figure 2. Audiovisual resources developed. The images illustrate the exercises assigned for weeks 1 and 2.

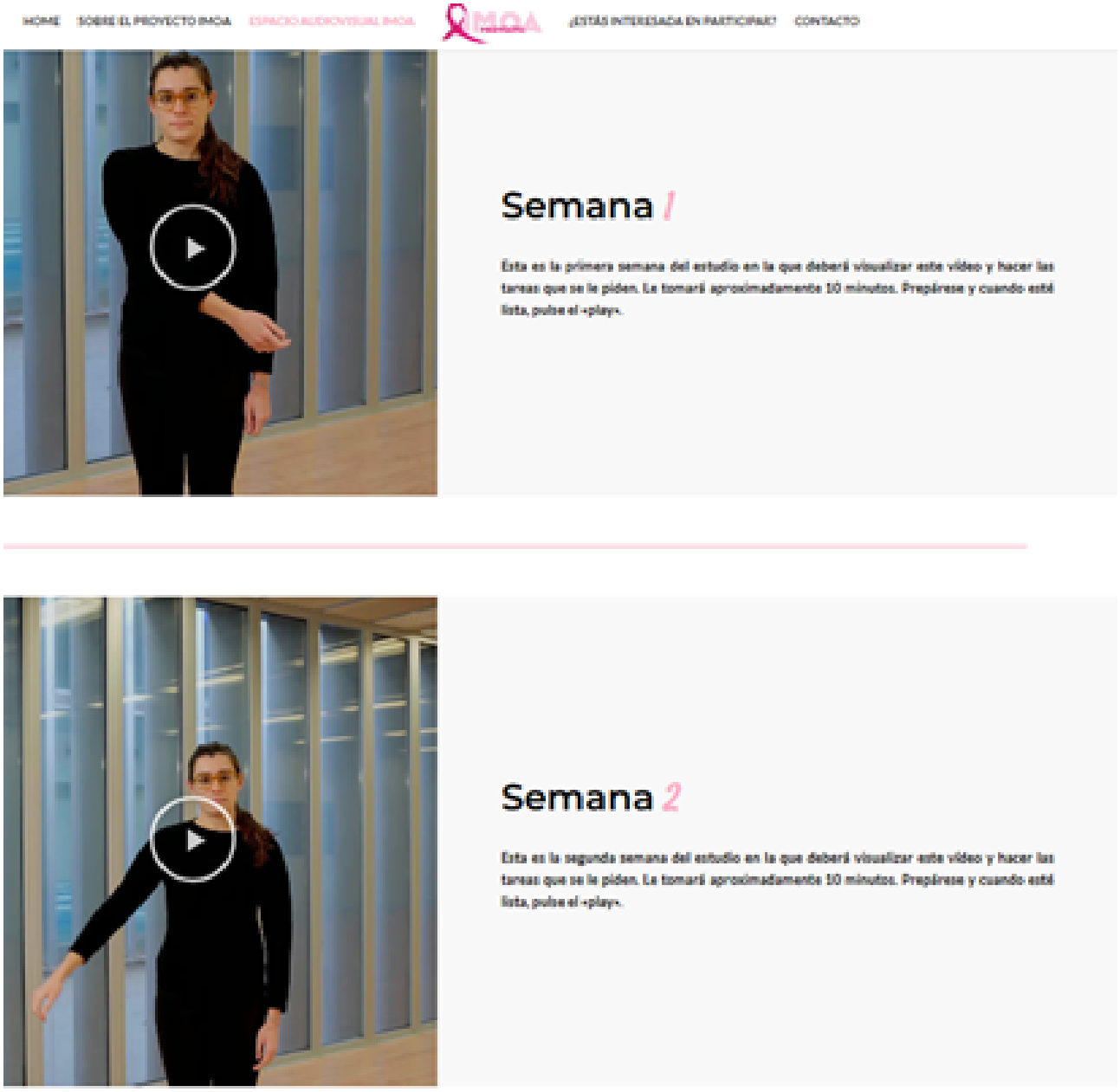

For each of the movements and ADLs, AO will be performed first, followed by MI. After the completion of the MI and AO blocks, the participants will also carry out mobility exercises involving repeating the movements they saw in the videos.

As indicated, all the participants will perform upper-extremity mobility exercises involving flexion, extension, abduction, adduction, internal rotation, and external rotation. They will perform the mobility exercises for approximately 20 minutes per session between 1 and 3 times per day on 3 different days per week throughout the 6-week intervention period. The minimum dose will be 1 time per day on 3 different days per week, and the maximum dose will be 3 times per week on 3 different days per week.

Each participant will be instructed on how to access the audiovisual resources developed specifically for this study and hosted on a dedicated website (Figure 3). They will also receive an information sheet describing the exercises and a field diary to record the frequency of intervention practice during the established weeks. At the end of the explanation, the participant will be scheduled for data collection at 6 weeks (T1) and 3 months (T2).

XSL•FO
RenderX

Figure 3. Web page created for the IM-OA23 project. The image shows the project name accompanied by the following phrase: "Innovating in breast cancer care to improve quality of life. Learn more."



The videos will be structured as follows: first, a black screen will appear, and a female voice will provide the instructions for performing the therapies. Next, the video will display the movement chosen for the AO therapy. The black screen will then appear again, and 2 acoustic signals will play, indicating the start and end of the MI task. During the MI phase, the participants will be instructed to perform kinesthetic MI from a first-person perspective, imagining themselves executing the observed movement and focusing on the sensations associated with the action. The same process will be subsequently repeated with the next selected movement. The videos will last between 6 and 10 minutes. After viewing the videos and performing the MI and AO therapies, the participants will actively perform the exercises displayed previously in the screens.

Control Condition

The CG will perform the same movement exercises as the IG but without the addition of MI or AO therapies. All the participants will perform mobility exercises of the upper limb involving flexion, extension, abduction, adduction, internal rotation, and external rotation for approximately 20 minutes 3 times a day 3 times a week for the duration of the intervention (6 weeks).

Once participants consent and are allocated to the CG, sociodemographic and clinical data will be collected, and the remaining variables will be measured (T0). Finally, participants will receive detailed instructions on the exercises, including the number of series and repetitions per series, along with a diary to record exercise frequency throughout the established 6 weeks. Follow-up assessments will be scheduled at 6 weeks (T1) and 3 months (T2).

Variables and Measurement Instruments

The independent variable will be the IM-OA23 program, which will combine MI and AO therapies. There will be 6 dependent variables.

Pain is the primary outcome variable. It will be evaluated using the NRS, one of the most widely recommended and commonly used tools for assessing pain intensity in a simple, quick, and reliable way. It is an objective scale that quantifies the intensity of pain, similar to the visual analog scale. The NRS is an 11-point scale presented as a horizontal line, with 0 indicating "no pain" and 10 indicating "the worst pain imaginable" or "the worst pain that can exist." Participants will be asked to report their average pain or pain experienced in the previous 24 hours [39].

Joint range, the degree of movement of the shoulder joint, will be measured using a goniometer. Flexion, abduction, and internal and external rotation will be measured using previously established reference points [40].

Muscular strength, the grip strength of the affected limb, will be measured using a dynamometer (Jamar). The measurement will be made with the participant in a sitting position with the elbow at 90° flexion, the shoulder close to the torso, and a neutral wrist position. The maximum achievable force will be recorded from the best value of 3 attempts [41,42].

Fear of movement will be measured with the Tampa Scale of Kinesiophobia, adapted to Spanish. The scale consists of 11 items rated on a 4-point Likert scale (1-4), yielding a total score from 11 to 44, with higher scores indicating greater fear of movement [43,44].

The diameter of the affected limb will be measured using anthropometric tape with a precision of 1 mm. The measurement will be made at the midpoint between the acromion and the olecranon, leaving the tape perpendicular to the extremity without compressing the area. The midpoint will be located at 90° elbow flexion, and subsequently, the measurement will be obtained with the elbow in extension; the result will be recorded in centimeters [45].

Imagination capability will be assessed using the Movement Imagery Questionnaire–Revised, Second Edition, an 8-item instrument. The questionnaire requires 2 tasks: first, creating a drawing or a mental image of a movement and, second, attempting to feel the sensation of performing the movement

XSL•FO
RenderX

without physically executing it. The difficulty of each item will be rated on a 7-point scale ranging from 1 ("very difficult to see/feel") to 7 ("very easy to see/feel"). The total score is obtained by adding the scores of all the questions. Higher scores correspond to greater ease of imagination and perceived vividness [46].

Data Collection

As previously described in the Setting section, participants will be enrolled through the dissemination of study information via various media, including the press, social networks, and SARAY. The study information will include a Google Forms questionnaire in which participants can express their interest in taking part and provide their contact information. All respondents to the questionnaire will be invited to an individual meeting with the research team, during which study information will be provided, questions will be addressed, and written informed consent will be obtained upon confirmation of participation. Subsequently, the planning of the intervention and data collection will be explained, and the assigned group will be indicated to each participant.

In both the IG and CG, outcomes will be assessed at 3 time points. The preintervention time point (T0) will coincide with the initial meeting and study inclusion, during which sociodemographic and clinical data will be collected, along with the other study variables. The postintervention time point will take place at 6 weeks (T1). Follow-up (T2) will subsequently be carried out 3 months after the end of the intervention to determine whether the effects have persisted over time (Figure 4). Pain will be considered the primary outcome, and the primary time point to evaluate the effectiveness will be T1. To collect the data, an electronic questionnaire will be designed with the help of the Microsoft Forms tool.

Figure 4. Timeline of the study. AO: action observation; BCLE-SEI-ES: Spanish version of the Breast Cancer and Lymphedema Symptom Experience Index; MI: motor imagery; MIQ-R: Movement Imagery Questionnaire–Revised; NPSI: Neuropathic Pain Symptom Inventory; NRS: numerical rating scale; TSK-11SV: Spanish version of the Tampa Scale of Kinesiophobia.

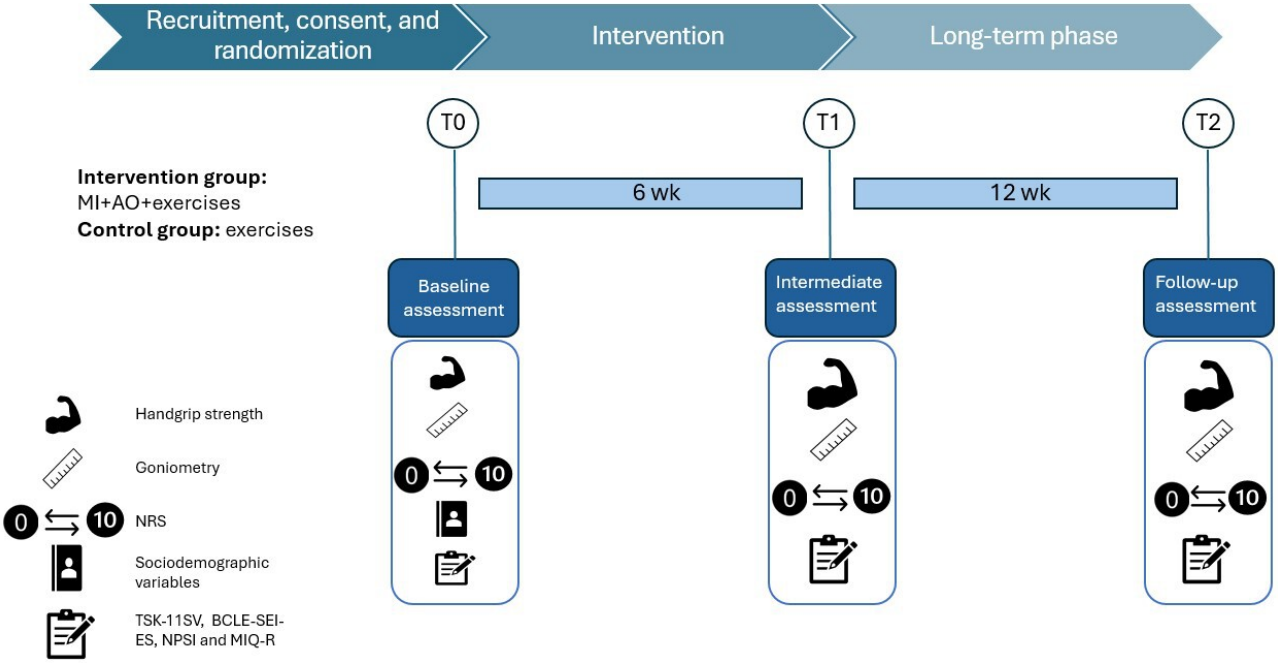

The IG will participate in the IM-OA23 program.

Data Analysis

First, a descriptive analysis of all study variables will be performed. Continuous variables will be summarized using means and SDs (or medians and IQRs), and categorical variables will be summarized using frequencies and percentages.

The normality of outcome distributions will be assessed for each study group and time using the Kolmogorov-Smirnov test with Lilliefors correction. If normality assumptions are not satisfied, equivalent nonparametric approaches will be considered.

Baseline comparability between groups will be assessed using a 2-tailed Student *t* test (or Wilcoxon rank-sum test if normality is not met) for continuous variables and chi-square or Fisher exact tests for categorical variables.

Second, inferential analyses will be carried out following an intention-to-treat approach whereby all randomized participants will be included in the analyses focused on between-group comparisons (intervention vs control) based on pretest-posttest change scores. Missing data will be imputed and included in all analyses.

The primary analysis will consist of evaluating the effect of the intervention on pain at T0, T1, and T2 comparing changes between the IG and CG from baseline using a linear mixed-effects model.

For all outcome variables meeting normality assumptions, longitudinal trajectories will be evaluated using linear mixed-effects models with participant-specific random intercepts for within-subject correlation. The models will include fixed effects for group (control vs intervention), time (baseline, T1, and T2), and the group × time interaction. Standardized effect sizes (Cohen *d*) will be calculated from the adjusted

XSL•FO
RenderX

between-group differences in change from baseline estimated by the mixed-effects models. Effect size will be interpreted as very small for values of less than 0.20, small for values of 0.20 to 0.49, moderate for values of 0.50 to 0.79, and large for values of 0.80 or more.

For variables not meeting normality assumptions, appropriate nonparametric methods will be considered. Between-group comparisons of change scores at each follow-up will be performed using the Wilcoxon rank-sum test.

In all cases, the significance level will be set at .05. The data will be analyzed using the SPSS program (version 27; IBM Corp).

Ethical Considerations

This project has been approved by the Ethics Committee for Drug Research of Navarra (PI_2023/81), and the necessary institutional authorization to carry out the study has been obtained . Participants will receive information about the project, their degree of involvement, and the legal requirements we will follow in an information sheet, as well as an informed consent form to be signed if they agree to voluntarily participate. Participants will not receive any compensation for participating in the study. The study will exhaustively comply with all the ethical-legal requirements in force related to research involving humans [47]. In order to maintain privacy and confidentiality, all data will be obtained under a personal and unique code. In the event that the IM-OA23 program is shown to be effective in improving pain, mobility, and lymphedema in breast cancer survivors, once the research is completed, breast cancer survivors from the CG will be invited to participate in the full program, and the resources used for the IG will be provided. In addition, the feasibility of incorporating the tool into the national health system will be assessed, with the aim of making it accessible to the entire population.

Results

The intervention is expected to yield improvements in pain intensity, joint range, muscle strength, and symptoms associated with lymphedema, among other outcomes. The study was funded in December 2023. The number of participants recruited as of manuscript submission is approximately 80, and data analysis has not yet started. Data collection is expected to be finished by June 2026, and the results will be published in 2026.

Discussion

Expected Findings

This randomized controlled trial will investigate the potential effect of the combination of MI and AO therapies in improving the functionality of the affected upper limb in breast cancer survivors. Although previous studies have demonstrated the effectiveness of combining these therapies in improving pain, fear of movement, and other outcomes in conditions such as phantom limb syndrome, stroke, or chronic pain, no studies to date have examined their impact in female breast cancer survivors with upper-extremity symptoms. This study seeks to address that gap. Given the similarities of symptoms in breast cancer survivors to those observed in other conditions for which MI and AO therapies have proven effective, the potential benefits for quality of life and treatment appear promising.

Implications for Clinical Practice

The results of this study have the potential to generate evidence supporting new therapeutic approaches for addressing functional limitations and mobility problems of the affected limb in women with breast cancer—a population in which these symptoms are common and frequently insufficiently treated.

This project is of particular interest given the growing number of female breast cancer survivors, many of whom face limitations in recovery and reintegration into daily activities due to disease-related consequences.

The implementation of the IM-OA23 program could directly benefit female breast cancer survivors with upper-limb sequelae. Aligned with current priorities for survivorship care outlined in the Strategic Framework of Primary and Community Care and the Comprehensive Plan of Care for Long-Term Cancer Survivors of the Spanish Society of Medical Oncology [48], the IM-OA23 program also represents an innovative approach with potential for significant social and economic impact nationally and internationally.

Short-term benefits may include improved pain management, increased mobility, and reduced lymphedema. If effective, the program could offer breast cancer survivors additional tools to alleviate pain and lymphedema while improving upper-limb mobility. This could reduce the need for pain medication and improve the ability to perform daily activities such as dressing, grooming, and reaching for objects while also helping prevent complications such as lymphedema-related infections.

Evaluating the medium-term impacts of the program would generate valuable knowledge on the translation of research into health care practice. Such assessments could clarify the program's sustainability, its integration into clinical workflows, and its potential to inform evidence-based guidelines for survivorship care.

In the long term, the results of this study could have a significant impact on alleviating symptoms affecting the lives of breast cancer survivors. At the same time, it may contribute to improving the quality and continuity of survivorship care while offering a cost-effective tool to improve patient management and expand the resources available to health care professionals.

In addition, if effective, the program could be proposed as an open access resource at the international level—adapted into different languages—to be accessible not only to rehabilitation professionals but also directly to patients. Such an approach would facilitate global reach, allowing for its use from anywhere while enabling the monitoring of symptom evolution and adherence to the intervention and the timely resolution of emerging issues.

Finally, this study highlights the need to prioritize research on women and the health challenges they face. Beyond its immediate contributions, it has the potential to open new avenues for future research aimed at understanding and

XSL•FO
RenderX

addressing symptoms that emerge with the chronification of specific pathologies.

Study Limitations

This study has several limitations. First, the follow-up period of the intervention is short, restricted to 3 months. At this time point, the long-term effects of the intervention may not yet have fully emerged, potentially underestimating its sustained impact. However, it should be noted that the selected follow-up times are based on previous studies analyzing the combination of MI and AO therapies. It should also be noted that follow-up is constrained by the funding period and by the practical and organizational limitations of participants. Therefore, future research should include long-term evaluations of the intervention's outcomes.

Second, given the behavioral nature of the program, it will not be possible to blind either the participants or the researchers delivering the intervention, which may introduce expectation bias. In addition, some outcomes (such as the number of repetitions and the frequency of exercises) rely on self-reported diaries, increasing the risk of reporting bias and raising potential concerns about adherence.

Third, the aim of the study is to evaluate the combination of MI and AO therapies, and although it would also be interesting to compare the effects of this combination with those of the therapies alone, due to the difficulty in recruiting volunteers, we will not be able to conduct a study in which we can compare the combination of MI and OA and the effects of the use of MI or OA alone.

Conclusions

The implementation of an intervention based on MI and AO has the potential to positively impact female breast cancer patients who face physical and psychological sequelae that interfere with their daily lives. This study has the potential to contribute important evidence on the effectiveness of combining MI and AO while also offering pioneering insights into the implementation of these therapies within oncology. By extending the use of these therapies to breast cancer survivorship, the findings of this study could open up new avenues for innovative, nonpharmacological strategies to improve long-term care and quality of life in this growing patient population.

Funding

This study was supported by the Department of Health of the Government of Navarra (Spain) under the call for grants for the promotion of research projects of 2023 (December 2023; grant 1100/2023).

Data Availability

All data generated or analyzed during this study are included in this published article.

Conflicts of Interest

None declared.

References

1. Cancer today. International Agency for Research on Cancer. URL: https://gco.iarc.who.int/today/ [accessed 2025-08-28]
2. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
3. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
4. Nardin S, Mora E, Varughese FM, et al. Breast cancer survivorship, quality of life, and late toxicities. Front Oncol 2020;10:864. [doi: 10.3389/fonc.2020.00864] [Medline: 32612947]
5. Monográfico SEOM de largos supervivientes en cáncer 1a parte. Sociedad Española de Oncología Médica. 2012. URL: https://seom.org/seomcms/images/stories/recursos/infopublico/publicaciones/MONOGRAFICO_SEOM_LARGOS_SUPERVIVIENTES_I.pdf [accessed 2026-03-18]
6. Nesvold IL, Dahl AA, Løkkevik E, Marit Mengshoel A, Fosså SD. Arm and shoulder morbidity in breast cancer patients after breast-conserving therapy versus mastectomy. Acta Oncol 2008;47(5):835-842. [doi: 10.1080/02841860801961257] [Medline: 18568481]
7. Dawes DJ, Meterissian S, Goldberg M, Mayo NE. Impact of lymphoedema on arm function and health-related quality of life in women following breast cancer surgery. J Rehabil Med 2008 Aug;40(8):651-658. [doi: 10.2340/16501977-0232] [Medline: 19020699]
8. Stubblefield MD, Custodio CM. Upper-extremity pain disorders in breast cancer. Arch Phys Med Rehabil 2006 Mar;87(3 Suppl 1):S96-S99. [doi: 10.1016/j.apmr.2005.12.017] [Medline: 16500198]

XSL•FO
RenderX

9. Schou Bredal I, Smeby NA, Ottesen S, Warncke T, Schlichting E. Chronic pain in breast cancer survivors: comparison of psychosocial, surgical, and medical characteristics between survivors with and without pain. J Pain Symptom Manage 2014 Nov;48(5):852-862. [doi: 10.1016/j.jpainsymman.2013.12.239] [Medline: 24703940]
10. Chan DN, Lui LY, So WK. Effectiveness of exercise programmes on shoulder mobility and lymphoedema after axillary lymph node dissection for breast cancer: systematic review. J Adv Nurs 2010 Sep;66(9):1902-1914. [doi: 10.1111/j.1365-2648.2010.05374.x] [Medline: 20626480]
11. Forsythe LP, Alfano CM, George SM, et al. Pain in long-term breast cancer survivors: the role of body mass index, physical activity, and sedentary behavior. Breast Cancer Res Treat 2013 Jan;137(2):617-630. [doi: 10.1007/s10549-012-2335-7] [Medline: 23242613]
12. Jeannerod M. Neural simulation of action: a unifying mechanism for motor cognition. Neuroimage 2001 Jul;14(1 Pt 2):S103-S109. [doi: 10.1006/nimg.2001.0832] [Medline: 11373140]
13. Schuster C, Hilfiker R, Amft O, et al. Best practice for motor imagery: a systematic literature review on motor imagery training elements in five different disciplines. BMC Med 2011 Jun 17;9:75. [doi: 10.1186/1741-7015-9-75] [Medline: 21682867]
14. Di Rienzo F, Blache Y, Kanthack TFD, Monteil K, Collet C, Guillot A. Short-term effects of integrated motor imagery practice on muscle activation and force performance. Neuroscience 2015 Oct 1;305:146-156. [doi: 10.1016/j.neuroscience.2015.07.080] [Medline: 26241339]
15. Herrador Colmenero L, Perez Marmol JM, Martí-García C, et al. Effectiveness of mirror therapy, motor imagery, and virtual feedback on phantom limb pain following amputation: a systematic review. Prosthet Orthot Int 2018 Jun;42(3):288-298. [doi: 10.1177/0309364617740230] [Medline: 29153043]
16. Hoyek N, Di Rienzo F, Collet C, Hoyek F, Guillot A. The therapeutic role of motor imagery on the functional rehabilitation of a stage II shoulder impingement syndrome. Disabil Rehabil 2014;36(13):1113-1119. [doi: 10.3109/09638288.2013.833309] [Medline: 24575717]
17. La Touche R, Fernández Pérez JJ, Martínez García S, Cuenca-Martínez F, López-de-Uralde-Villanueva I, Suso-Martí L. Hypoalgesic effects of aerobic and isometric motor imagery and action observation exercises on asymptomatic participants: a randomized controlled pilot trial. Pain Med 2020 Oct 1;21(10):2186-2199. [doi: 10.1093/pm/pnaa015] [Medline: 32142135]
18. Mateo S, Di Rienzo F, Bergeron V, Guillot A, Collet C, Rode G. Motor imagery reinforces brain compensation of reach-to-grasp movement after cervical spinal cord injury. Front Behav Neurosci 2015 Sep 11;9:234. [doi: 10.3389/fnbeh.2015.00234] [Medline: 26441568]
19. Salik Sengul Y, Kaya N, Yalcinkaya G, Kirmizi M, Kalemci O. The effects of the addition of motor imagery to home exercises on pain, disability and psychosocial parameters in patients undergoing lumbar spinal surgery: a randomized controlled trial. Explore (NY) 2021;17(4):334-339. [doi: 10.1016/j.explore.2020.02.001] [Medline: 32147444]
20. Suso-Martí L, La Touche R, Angulo-Díaz-Parreño S, Cuenca-Martínez F. Effectiveness of motor imagery and action observation training on musculoskeletal pain intensity: a systematic review and meta-analysis. Eur J Pain 2020 May;24(5):886-901. [doi: 10.1002/ejp.1540] [Medline: 32031724]
21. Robin N, Rieusset L, Ruart S, Ben-Rais A, Bourhis-Espiand V. Influence of text messages promoting mental imagery on self-reported physical activity in women with breast cancer: a randomized controlled study. J Bodyw Mov Ther 2024 Oct;40:797-803. [doi: 10.1016/j.jbmt.2024.06.005] [Medline: 39593679]
22. Proverbio AM, Zani A. Mirror neurons in action: erps and neuroimaging evidence. In: Boggio PS, Wingenbach TS, Da Silveira Coêlho ML, Comfort WE, Murrins Marques L, Alves MV, editors. Social and Affective Neuroscience of Everyday Human Interaction: Springer; 2023:65-84. [doi: 10.1007/978-3-031-08651-9_5]
23. Oberman LM, Pineda JA, Ramachandran VS. The human mirror neuron system: a link between action observation and social skills. Soc Cogn Affect Neurosci 2007 Mar;2(1):62-66. [doi: 10.1093/scan/nsl022] [Medline: 18985120]
24. Ertelt D, Small S, Solodkin A, et al. Action observation has a positive impact on rehabilitation of motor deficits after stroke. Neuroimage 2007;36 Suppl 2:T164-T173. [doi: 10.1016/j.neuroimage.2007.03.043] [Medline: 17499164]
25. Eaves DL, Riach M, Holmes PS, Wright DJ. Motor imagery during action observation: a brief review of evidence, theory and future research opportunities. Front Neurosci 2016;10:514. [doi: 10.3389/fnins.2016.00514] [Medline: 27917103]
26. Choi JB, Yang SW, Ma SR. The effect of action observation combined with motor imagery training on upper extremity function and corticospinal excitability in stroke patients: a randomized controlled trial. Int J Environ Res Public Health 2022 Sep 23;19(19):12048. [doi: 10.3390/ijerph191912048] [Medline: 36231353]
27. Arnal-Vallés MP, Escalada-Hernández P, Bays-Moneo AB, et al. Effectiveness of the combined use of motor imagery and action observation therapies in improving functionality: a systematic review and meta-analysis. NeuroRehabilitation 2025 Dec;57(4):435-452. [doi: 10.1177/10538135251382909] [Medline: 41056136]
28. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346:e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
29. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010 Mar 24;8:18. [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]

XSL•FO
RenderX

30. Kang M, Rice D, Helsby N, Somogyi A, Kluger M, Chiang D. Establishing the patient acceptable symptom state for patient-reported pain outcomes 6 months after breast cancer surgery. Pain Rep 2025;10(4):e1297. [doi: 10.1097/PR9.0000000000001297] [Medline: 40524741]
31. Kendrick DB, Strout TD. The minimum clinically significant difference in patient-assigned numeric scores for pain. Am J Emerg Med 2005 Nov;23(7):828-832. [doi: 10.1016/j.ajem.2005.07.009] [Medline: 16291435]
32. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008 Feb;9(2):105-121. [doi: 10.1016/j.jpain.2007.09.005] [Medline: 18055266]
33. Harrington S, Gilchrist L, Sander A. Breast cancer EDGE Task Force outcomes: clinical measures of pain. Rehabil Oncol 2014;32(1):13-21. [Medline: 25346950]
34. Sancar M, Ozturk EC. The assessment of neuropathic pain in patients with lymphedema secondary to breast cancer. Pain Physician 2025 Nov;28(6):545-550. [Medline: 41337767]
35. Cho Y, Do J, Jung S, Kwon O, Jeon JY. Effects of a physical therapy program combined with manual lymphatic drainage on shoulder function, quality of life, lymphedema incidence, and pain in breast cancer patients with axillary web syndrome following axillary dissection. Support Care Cancer 2016 May;24(5):2047-2057. [doi: 10.1007/s00520-015-3005-1] [Medline: 26542271]
36. Carretti G, Mirandola D, Maestrini F, et al. Quality of life improvement in breast cancer survivors affected by upper limb lymphedema through a novel multiperspective physical activity methodology: a monocentric pilot study. Breast Cancer 2022 May;29(3):437-449. [doi: 10.1007/s12282-021-01322-0] [Medline: 35025064]
37. Scaturro D, Vitagliani F, Mangano MS, et al. Effect of telerehabilitation and outpatient rehabilitation in women with breast cancer. J Funct Morphol Kinesiol 2023 Jul 27;8(3):105. [doi: 10.3390/jfmk8030105] [Medline: 37606400]
38. Drake RL, Vogl AW, Mitchell AM. Gray's Anatomy for Students 5th Edition: Elsevier; 2023.
39. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken) 2011 Nov;63 Suppl 11:S240-S252. [doi: 10.1002/acr.20543] [Medline: 22588748]
40. Taboadela CH. Goniometría: Una Herramienta para la Evaluación de las Incapacidades Laborales: Asociart ART; 2007.
41. Ha YC, Yoo JI, Park YJ, Lee CH, Park KS. Measurement of uncertainty using standardized protocol of hand grip strength measurement in patients with sarcopenia. J Bone Metab 2018 Nov;25(4):243-249. [doi: 10.11005/jbm.2018.25.4.243] [Medline: 30574469]
42. Roberts HC, Denison HJ, Martin HJ, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. Age Ageing 2011 Jul;40(4):423-429. [doi: 10.1093/ageing/afr051] [Medline: 21624928]
43. Gómez-Pérez L, López-Martínez AE, Ruiz-Párraga GT. Psychometric properties of the Spanish version of the Tampa Scale for Kinesiophobia (TSK). J Pain 2011 Apr;12(4):425-435. [doi: 10.1016/j.jpain.2010.08.004] [Medline: 20926355]
44. Hudes K. The Tampa Scale of Kinesiophobia and neck pain, disability and range of motion: a narrative review of the literature. J Can Chiropr Assoc 2011 Sep;55(3):222-232. [Medline: 21886284]
45. Canda Moreno AS. Variables Antropométricas de la Población Deportista Española: Consejo Superior de Deportes; 2012.
46. Gregg M, Hall C, Butler A. The MIQ-RS: a suitable option for examining movement imagery ability. Evid Based Complement Alternat Med 2010 Jun;7(2):249-257. [doi: 10.1093/ecam/nem170] [Medline: 18955294]
47. Ley 14/2007, de 3 de julio, de investigación biomédica. URL: https://www.boe.es/buscar/pdf/2007/BOE-A-2007-12945-consolidado.pdf [accessed 2026-03-24]
48. Recomendaciones para la mejora de la atención a pacientes largos supervivientes de cáncer en el Sistema Nacional de Salud. Ministerio de Sanidad. 2025. URL: https://www.sanidad.gob.es/areas/calidadAsistencial/estrategias/cancer/docs/Recomendaciones_largo_superviviente._ACCESIBLE.pdf [accessed 2026-03-18]

Abbreviations

ADL: activity of daily living
AO: action observation
CG: control group
CONSORT: Consolidated Standards of Reporting Trials
IG: intervention group
MI: motor imagery
NRS: numerical rating scale
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

XSL•FO
RenderX

Edited by J Sarvestan; submitted 09.Oct.2025; peer-reviewed by N Robin; revised version received 23.Feb.2026; accepted 24.Feb.2026; published 30.Mar.2026.

Please cite as:
Arnal-Vallés MP, Soto-Ruiz N, Bays-Moneo AB, García-Vivar C, Escalada-Hernández P
Motor Imagery and Action Observation in Breast Cancer Survivors: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e85469
URL: https://www.researchprotocols.org/2026/1/e85469
doi:10.2196/85469

© María Pilar Arnal-Vallés, Nelia Soto-Ruiz, Ana Beatriz Bays-Moneo, Cristina García-Vivar, Paula Escalada-Hernández. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Ayurvedic Nutritional Supplements and Yoga Protocol in Reducing the Incidence and Severity of Acute Mountain Sickness (AYAMS Study): Study Protocol for an Open-Label Randomized Controlled Trial

Amit Kumar Rai[1], MD; Sophia Jameela[2], MD; Arun Kumar Yadav[3], MD, PhD; Rajneesh Kumar Joshi[3], MD, PhD; Pallavi Mundada[2], MD; Azeem Ahmad[2], MD; Bidhan Mahajon[4], MD, PhD; Bhagwan Sahai Sharma[2], MD; Shruti Khanduri[2], MD; Babita Yadav[2], MD, PhD; Arunabh Tripathi[2], PhD; Rakesh Kumar Rana[2], PhD; Bhogavalli Chandrasekhararao[2], MD; Narayanam Srikanth[2], MD; Raghavendra Rao[5], PhD; Rabinarayan Acharya[2], MD, PhD

[1]Ayurvedic and Unani Tibbia College and Hospital, New Delhi, India

[2]Central Council for Research in Ayurvedic Sciences, New Delhi, India

[3]Office of the DGAFMS, Fifth Floor, Block A, Defence Officer Complex, Africa Avenue, New Delhi, India

[4]All India Institute of Ayurveda, New Delhi, India

[5]Central Council for Research in Yoga and Naturopathy, New Delhi, India

Corresponding Author:
Arun Kumar Yadav, MD, PhD
Office of the DGAFMS, Fifth Floor, Block A, Defence Officer Complex, Africa Avenue, New Delhi, India

Abstract

Background: Acute mountain sickness (AMS) poses a unique and formidable challenge to healthy personnel at high altitudes.

Objective: This randomized controlled trial (RCT) protocol aimed to assess the effectiveness of Ayurvedic nutritional supplements in conjunction with a yoga protocol in reducing the incidence and severity of AMS among healthy personnel stationed in the challenging high-altitude landscapes of the western Himalayas.

Methods: The proposed open-label, parallel-group RCT was conducted in apparently healthy individuals of any gender aged 18 - 50 years. The study was conducted at two distinct high-altitude stages within the western Himalayan region: stage I, situated at an elevation ranging from 9000 to 12,000 feet and stage II, spanning 12,000 to 15,000 feet. A total of 1660 participants (n=830 per stage) underwent random assignment in a 1:1 ratio to receive either the existing acclimatization schedule (AS) for high altitude (control group) or the Ayush intervention (Ayush group) in addition to the AS. The participants in the Ayush group received Ayurvedic nutritional supplements, including the Ayur-nutri kit (Ayush poshak yoga, 25 g, and Ayush cardiac care tea, 125 mL), twice daily, along with a yoga protocol (60 min daily) for 120 days. The primary outcome was the incidence of AMS, assessed using the 2018 Lake Louise Scoring System, and the proportion of participants with a Lake Louise Scoring System score of 6 or higher during the first 7 days from baseline. The secondary outcome measures included the proportion of participants with thromboembolic events; changes in coagulation and hemostasis activation markers and proinflammatory markers; and changes in self-reported negative emotional states (depression, anxiety, and stress), sleep quality, and overall quality of life (assessed through Depression Anxiety Stress Scale-21 items, Pittsburgh Sleep Quality Index, and 12-item short-form, respectively) on day 60 and day 120 from baseline.

Results: The study was funded in March 2023. The data collection was completed in December 2023. A total of 1660 participants were enrolled in the study. The analysis of the study data is in progress. The study outcomes are expected to be published by December 2026.

Conclusions: This RCT was the first of its kind to explore the potential benefits of using Ayurvedic nutritional supplements and a yoga protocol in conjunction with the standard AS to reduce the occurrence and severity of AMS among healthy personnel. The outcomes of this trial can aid in better acclimatization and resilience among healthy personnel at high altitudes.

Trial Registration: Clinical Trial Registry of India CTRI/2023/03/051028; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=ODI3Njg=&Enc=&userName=

International Registered Report Identifier (IRRID): DERR1-10.2196/81567

(JMIR Res Protoc 2026;15:e81567) doi:10.2196/81567

XSL•FO
RenderX

KEYWORDS

acute mountain sickness; acclimatization; high altitude; Terminalia arjuna; Withania somnifera; yoga

Introduction

Background

High-altitude illness (HAI) is the collective term used for various debilitating health conditions that can affect individuals who are not acclimatized and ascend rapidly to high altitudes within 6 to 12 hours of their ascent to high altitude [1,2]. Conventionally, 2500 meters has been used as the threshold for HAI [2]. The spectrum of HAI encompasses various conditions that can range from mild to severe, among which the most typical are acute mountain sickness (AMS), high-altitude cerebral edema, and high-altitude pulmonary edema (HAPE). High-altitude cerebral edema and HAPE occur much less frequently than AMS but are potentially fatal. The risk factors for HAI are the rate of ascent, high altitude reached (especially the sleeping altitude), previous high-altitude exposure, and individual susceptibility [1,3,4]. AMS is a common clinical presentation in individuals venturing over 2500 meters and above [5]. The prevalence and severity of AMS rise with increasing altitude. It occurs in approximately 10% to 25% of nonacclimatized individuals at 2500 meters, which may increase to 50% to 85% at an altitude of 4500 to 5500 meters [2,6-8]. The pathophysiology of AMS includes relative hypoventilation, impaired gas exchange, increased sympathetic activity, fluid retention and redistribution, and raised intracranial pressure in case of moderate to severe AMS [1,5]. AMS presents with a group of nonspecific symptoms such as headache (primary defining symptom), gastrointestinal symptoms (anorexia, nausea, or vomiting), fatigue and/or weakness, dizziness, or disturbed sleep [9].

A gradual ascent to enable acclimatization is usually impractical logistically for individuals susceptible to AMS, particularly in recreational, tactical, and exigency situations. Considering this limitation, several clinical studies have been conducted over the last 2 decades to explore the efficacy of conventional medications in reducing the incidence of AMS, followed by a few systematic reviews [10-13]. The interventions, such as acetazolamide, dexamethasone, and ibuprofen, have shown promising results in reducing the severity of AMS; however, they have specific adverse effects, limiting their routine use. Acetazolamide has reported adverse effects, including headache, nausea, polyuria, and dysgeusia [11,12]. Similarly, dexamethasone and ibuprofen have shown adverse effects such as gastrointestinal symptoms, headache, skin rashes, edema, and mood changes (including depression) [12]. Considering this, exploring effective, safe, and readily available interventions for prophylaxis and management of AMS is essential.

Ayurveda, the traditional Indian medicine, aims to improve the quality and span of life, primarily emphasizing disease prevention and health promotion by strengthening body tissues to withstand both exogenous and endogenous stressors. The same can be achieved by modulating diet and lifestyle and by using Ayurvedic nutritional supplements (containing herbal ingredients) such as *Rasayana*, which help restore the body's equilibrium. The herbal ingredients used in Ayurvedic nutritional supplements are known for their adaptogenic, antioxidant, and immunomodulatory properties [14-16]. A holistic approach involving the use of Ayurvedic nutritional supplements and a yoga protocol may have significant scope for reducing AMS, effectively addressing psychological and physiological well-being, considering that while the Ayurveda supplements enable the body to cope with the physical stressors, the yoga protocol can enhance the mental resilience and focus, which are crucial for operating in challenging conditions.

Rationale for Study Interventions

The Ayur-nutri kit comprises Ayush poshak yoga, a nutritional supplement, and Ayush cardiac care tea (CCT), an herbal drink, having ingredients such as *Terminalia arjuna* (Roxb.) Wight & Arn, *Crocus sativus* L., *Withania somnifera* (L.) Dunal, *Tinospora cordifolia* (Willd.) Miers ex Hook.f, *Zingiber officinale* Roscoe, *Piper longum* L., *Piper nigrum* L., *Chlorophytum borivilinium* Santapau & R.R.Fern., etc, along with a wide variety of nuts and seeds such as almonds, walnuts, cashews, pistachios, pumpkin seeds, and melon seeds. The herbs in these nutritional supplements have cardioprotective, neuroprotective, hepatoprotective, antihypertensive, antidyslipidemic, anti-inflammatory, antiatherogenic, antiplatelet aggregation, antithrombotic, and anticoagulant activities [17-21]. A clinical study on these nutritional supplements, conducted among volunteers of the Antarctica expedition, reported significant improvements in stress-related changes in immunoglobulins and antioxidant enzymes, alongside an enhanced quality of life. This suggests that they are effective in managing cold stress and were also found to be safe [22]. Preliminary studies have shown that yoga contributes to acclimatization at high altitude and plays a crucial role in maintaining physical fitness and mental resilience [23]. Respiratory adaptations induced by yoga may offer an efficient strategy to cope with high altitude–induced hypoxia [24,25].

Based on the evidence stated above, this study has been conceptualized to assess the comparative effectiveness of incorporating a combined protocol of Ayurvedic nutritional supplements and yoga alongside the existing acclimatization regimen for preventing AMS among healthy personnel at high-altitude terrains.

Objectives

The present randomized controlled trial (RCT) was designed to assess the effectiveness of a combination of Ayurvedic nutritional supplements and yoga protocol in reducing the incidence and severity of AMS compared to the standard acclimatization schedule (AS) at high-altitude areas in the western Himalayan region. The key secondary objectives of this study included evaluating the impact of the Ayush intervention protocol (Ayur-nutri kit + yoga) on subjective and assessment parameters, including coagulation and hemostasis activation markers, proinflammatory markers, negative emotional states (depression, anxiety, and stress), sleep quality, and overall quality of life after 120 days. Another secondary

XSL•FO
RenderX

objective was assessing the safety of Ayush interventions in AMS.

Methods

Study Design and Setting

The study protocol has been drafted in accordance with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines [26]. This clinical study was an open-label, randomized, controlled parallel-group trial. High-altitude areas are classified into three stages: stage I (9000 - 12,000 feet), stage II (12,000 - 15,000 feet), and stage III (more than 15,000 feet) [27]. The study was conducted at stage I and II high-altitude areas in the western Himalayan region.

Study Participants

Apparently healthy personnel of any gender aged 18 - 50 years, deemed physically fit for sojourn to high-altitude areas (in the western Himalayan regions) and willing to provide written informed consent, were included in the study. The study population consisted of individuals who were sojourning at high altitude for the first time and those who resumed sojourns after a break.

Individuals with a history of underlying cardiovascular disease or a clinically significant electrocardiogram abnormality, cerebrovascular disease, pulmonary disease, neurological disease, psychiatric illness, major venous thromboembolism (deep vein thrombosis, pulmonary embolism, or cerebrovascular thrombosis), sickle-cell disease, metabolic disorders, malignancy, glaucoma; a history of pulmonary edema or severe physiological response to high altitude (such as HAPE); evidence of uncontrolled diabetes mellitus (glycated hemoglobin more than 8.0%), uncontrolled hypertension (more than 140/90 mm Hg) even after medications, abnormal hepatic function (aspartate transaminase and/or alanine transaminase more than 2 times the upper limit of normal), or abnormal renal function (serum creatinine more than 1.2 mg%); or BMI ≥30 kg/m^2 were excluded from the study. Similarly, those with a history of any surgery within the past 6 months; transfusion of blood or any blood product within 90 days before screening; regular use of any concomitant conventional medication; a history of chain smoking, alcohol use disorder, or substance abuse; and the presence of any other clinical condition which the investigator thinks may jeopardize the safety of the study participants were not considered for the study.

Study Intervention

The eligible participants in the Ayush group (AG) received a comprehensive Ayush-AMS protocol, which included the use of an Ayur-nutri kit and a yoga protocol, along with the existing AS, for 120 days while staying in high-altitude regions [27].

The Ayur-nutri kit contains 2 herbal products, viz, an Ayurvedic nutritional supplement, Ayush poshak yoga, and an herbal tea, Ayush CCT, developed by the Central Council for Research in Ayurvedic Sciences, Ministry of Ayush, Government of India, to address nutritional supplementation requirements and to maintain healthy well-being at high-altitude regions. The Ayush poshak yoga powder contains a diverse blend of nuts and seeds, including cashews, almonds, pistachios, walnuts, pumpkin seeds, and cucumber seeds. It also contains herbs such as *W. somnifera* (L.) Dunal, *P. longum* L., *T. cordifolia* (Willd.) Miers ex Hook.f, *P. nigrum* L., *Z. officinale* Roscoe., *Foeniculum vulgare* Mill., renowned for their nutritional and therapeutic benefits (Table 1). The Ayush CCT is a novel herbal tea that combines the heritage of traditional tea with a blend of herbal ingredients, including *T. arjuna* (Roxb.). Wight & Arn, saffron, cardamom, and cinnamon that can promote cardiovascular health, overall well-being, and healthy circulation (Table 1). Ayush CCT was made into an infusion by dipping one Ayush CCT tea bag in 125 mL of hot water for 5 minutes, allowing for an appropriate steeping time, and then used warm. Ayush poshak yoga was given in a dose of 25 g twice daily (morning and evening, 1 h before meals), along with 125 mL of Ayush CCT.

XSL•FO
RenderX

Table . Details of ingredients of Ayush poshak yoga and Ayush CCT[a].

Ayur-nutri kit and ingredients (botanical/English name)	Part used
Ayush poshak yoga	
Anacardium occidentale L. (Cashew nut)	Kernel
Prunus amygdalus Stokes (Almond)	Kernel
Pistacia vera L. (Pistachio)	Kernel
Juglans regia L. (Walnut)	Cotyledons
Papaver somniferum L.	Seed
Chlorophytum borivilianum Santap & R.R.Frn	Tuberous root
Cucumis melo L.	Seed
Cucumis sativus L.	Seed
Citrullus vulgaris Schrad.	Seed
Cucumis melo var. *utilissimus* Dutt & Fueller	Seed
Zingiber officinale Rosc	Rhizome
Piper longum L.	Fruit
Piper nigrum L.	Fruit
Withania somnifera (L.) Dunal.	Root
Foeniculum vulgare Mill.	Fruit
Tinospora cordifolia (Willd.) Miers	Stem
Sugar candy	
Excipients, preservatives, flavoring agents (approved)	
Ayush CCT	
Camellia sinensis (L.) O.Kuntze	Leaf
Elettaria cardamomum Maton	Seed
Cinnamomum verum Breyn	Stem bark
Crocus sativus L. (Saffron)	Style and stigma
Terminalia arjuna (Roxb.) Wight & Am	Stem bark

[a]CCT: cardiac care tea.

The yoga protocol (High-Altitude Illness Prevention Protocol of Yoga) has been tailor-made for high altitude by the Central Council for Yoga & Naturopathy, Ministry of Ayush, Government of India. The participants performed the yoga protocol under the supervision of a qualified yoga instructor. It involved sessions beginning with breathing exercises, moving onward to performing Asanas in standing posture (05 Asanas), followed by Asanas in supine (05), prone (02), and sitting postures (02). The yoga session concluded with a sequence of Pranayama and meditation or breathing exercises, totaling approximately 60 minutes each day. The yoga schedule followed an alternating pattern throughout the 120 days (Table 2). Incorporating diverse asanas within varying postures, strategically arranged in specific patterns, ensured that each sequence effectively addressed certain aspects, such as the targeted engagement of distinct muscle groups and nurturing mental well-being.

XSL•FO
RenderX

Table . Details of yoga protocol (High-Altitude Illness Prevention Protocol of Yoga).

Type of yoga	Yoga practices (alternate day schedule)		Rounds	Duration[a] (min)
	Day 1	Day 2		
Breathing exercises	Hands stretch breathing	Hands stretch breathing	—[b]	10 min
	Hands in and out breathing	Hands in and out breathing	9	
	Yogic breathing	Yogic breathing	9	
	Ankle stretch breathing	Ankle stretch breathing	9	
	Kapalabhati kriya	*Kapalabhati kriya*	9	
	Yogic breathing	Yogic breathing	3	
Asanas: standing	*Surya namaskar*	*Surya namaskar*	4	6 min
	Ardha kati chakrasana	*Trikonasana*	1	10 min
	Padahasthasana	*Parshvakonasana*	1	
	Ardhachakrasana	*Vrikshasana*	1	
	Trikonasana	*Veerabhadrasana*	1	
Asanas: supine	*Shavasana*—quick relaxation technique	*Shavasana*—quick relaxation technique	1	5 min
	Uttanapadasana	*Pawanamuktasana*	1	5 min
	Sarvangasana	*Suptapadangushtasana*		
	Matsyasana	*Viparaitakarni*		
Asanas: prone	*Bhujangasana*	*Dhanurasana*	1	4 min
	Shalabhasana	*Makarasana*		
Asanas: sitting	*Shashankasana*	*Ardhamatsyendrasana*	1	4 min
	Ustrasana	*Gomukhasana*		
Pranayama	*Nadishuddi*	*Nadishuddi*	9	6 min
	Bhramari	*Bhramari*	9	
Meditation or relaxation	*Om* meditation or guided meditation	*Om* meditation or guided meditation	1	10 min
	Shavasana	*Shavasana*	1	

[a]Total duration: 60 min.

[b]Not applicable.

The participants in the control group (CG) followed the existing AS for high altitude. The research team, comprising a medical expert, an Ayurveda practitioner, and a yoga instructor, administered the Ayush interventions at the study sites throughout the study. The participants were monitored at frequent intervals at both study sites.

Ayush CCT was provided by the Indian Medicines Pharmaceutical Corporation Limited, India, and Ayush poshak yoga by MARC Laboratories Ltd, India.

Discontinuation of the Study Interventions

If any participant developed adverse effects, such as gastrointestinal symptoms, allergic reactions, or changes in biochemical assessment parameters, the administration of the Ayur-nutri kit was temporarily halted, and the participant was closely monitored. If symptoms recurred after reintroducing the nutritional supplements, the participant was withdrawn from the study after assessing the causality of adverse event (AE) or adverse drug reaction. All such events were recorded and reported in the AE or adverse drug reaction reporting format.

Compliance With the Study Interventions

All participants in the AG group were provided with an information leaflet containing instructions for the use (dose, frequency, and time of administration) and storage of the Ayur-nutri kit. The participants were also issued a compliance form during the baseline visit and subsequent follow-up visits to self-report their consistent or irregular use of Ayush interventions in the Ayur-nutri kit and to record any missed doses, along with remarks for missing doses, which enabled the assessment of adherence to the dosing pattern as per the study protocol. During each follow-up visit, the participants were asked to return the used, unused, or partially used sachets of the Ayur-nutri kit to the investigators to assess adherence and cross-check with the participant's self-reported compliance form. The participants were advised to maintain an open communication channel with the research team to address any

XSL•FO
RenderX

issues they faced, including missed doses, adverse effects, and general health-related queries.

Adherence to the 60-minute daily yoga protocol was systematically monitored through instructor-maintained attendance logs at each site and also from the participant's compliance diary. Daily session attendance was recorded for every participant throughout the intervention period.

The participants who did not adhere to the study protocol, or did not have 80% or more compliance, or those who developed any study-related AEs due to which the participant preferred to withdraw from the study, or those who withdrew their voluntary consent for participation in the study, were withdrawn from the study. If participants consented to data collection during the scheduled follow-ups or at the end of the 120 days, data were collected and recorded in the case report form (CRF). The data from completed assessments, available until the withdrawal, were used for analysis.

Concomitant Care

The investigators monitored the participants for any concomitant treatment they received during the study period. In the event of any AEs, the use of rescue medication was permitted under the investigator's discretion. All instances of concomitant care were carefully documented in the CRF. However, any study participant prescribed a conventional medication for a longer duration (more than 6 consecutive days) during the study period must discontinue use of the Ayush intervention during that period. This threshold distinguishes brief, incidental, or rescue medication use from sustained therapy that could potentially interact with or confound the effects of the study intervention. All such instances were documented in the CRF. If any participant developed a serious AE or a treatment-emergent AE during the study period, the participant was withdrawn from the study and given appropriate incidental care at the hospital. The government authorities and the Ethics Committee were notified within 2 working days, along with an appropriate justification.

Outcome Measures

The primary outcome was the incidence of AMS, assessed using the 2018 Lake Louise Scoring System (LLSS). AMS is defined as an LLSS score of 3 or more points from the 4 rated symptoms (headache, nausea or vomiting, fatigue, and dizziness or light-headedness), including at least 1 point from headache [9]. The grading of LLSS score is mild AMS 3 - 5 points, moderate AMS 6 - 9 points, and severe AMS 10 - 12 points [9]. The participants were assessed daily for the first 7 days at the stage I high-altitude areas using the LLSS score to record the incidence and evaluate the severity of AMS. The primary outcome will be reported as the proportion of participants who display symptoms of AMS and the proportion of participants with a score of more than 6 as per the LLSS. The participants enrolled for stage II were again assessed daily during their first 4 days at that altitude for the incidence and severity of AMS.

The secondary outcome measures included the proportion of participants who developed thromboembolic events (deep-vein thrombosis, pulmonary embolism, cerebral venous thrombosis, myocardial infarction, etc) during the study, change in the levels of markers of coagulation and hemostasis activation (total red blood cell count, hemoglobin, packed cell volume, D-dimer, lactate dehydrogenase, creatine kinase, creatine kinase-myocardial band, fibrinogen, activated partial thromboplastin time, prothrombin time), changes in the levels of proinflammatory markers (high sensitivity C-reactive protein, interleukin-6, tumor necrosis factor-α); self-reported changes in negative emotional states of depression, anxiety, and stress assessed through Depression Anxiety Stress Scale-21 item (DASS-21); change in sleep quality assessed through the Pittsburgh Sleep Quality Index (PSQI), and quality of life assessed through the 12-item short-form (SF-12) questionnaire. The secondary outcomes were evaluated on day 60 and day 120 from baseline. However, the participants enrolled for the stage II high-altitude area were also assessed on day 10 (day 4 of the AS at that altitude) and on the day of descent, in case of early descent before day 120.

Resting heart rate, systolic and diastolic blood pressure, respiratory rate, and oxygen saturation (SpO2) were also documented daily until day 7, alongside LLSS at stage I high altitude, and until day 10 in participants at stage II high altitude. These measurements were then taken on scheduled follow-up visits.

Safety Assessment

The safety of the Ayush interventions was determined by recording the incidence of AEs, if any, during scheduled follow-up visits in a structured format. All AEs during the study were recorded and monitored as per ICH (International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use) Good Clinical Practice guidelines. Safety was also evaluated by assessing liver function tests, kidney function tests, and lipid profiles on days 60 and 120 from baseline.

Sample Size

Based on the findings from a previous study on similar conditions, the anticipated incidence of AMS within the CG was 25%, while the assumed incidence in the AG was 18% [7]. To achieve a statistical power of 90% and a CI of 95%, an initial sample size of 722 per group was calculated. Accounting for an expected attrition rate of 15%, the final sample size for each group was determined to be 830 participants (N=1660 in total). The sample size of 1660 participants was stratified into stage I high altitude and stage II high altitude. A total of 830 participants were included from stage II high altitude, while the remaining 830 were from stage I high altitude. The participants studied at each of the 2 altitude stages were randomly allocated to the AG and CG.

Recruitment of Study Participants

During the study, the database of healthy personnel in high-altitude areas of the western Himalayan region was used to identify and recruit potential subjects. The investigators at the study site screened subjects based on the defined inclusion and exclusion criteria. The eligible participants were allocated to either of the 2 groups based on a computer-generated randomization schedule. The screening process was continued until the target sample size for the study was achieved.

XSL•FO
RenderX

Randomization

Block randomization with unequal block sizes randomly assigned participants to either the AG or CG with a 1:1 allocation to minimize predictability. Block randomization was implemented separately for stage I and stage II participants to ensure balanced allocation within each altitude-exposure setting. The SPSS version 28.0 (IBM) was used to generate the random number sequences. An independent statistician not involved in the participants' enrollment and assessment generated the randomization sequence.

Allocation Concealment

Sequentially numbered, opaque, sealed envelopes were used to ensure allocation concealment. The participant's enrollment number was printed at the top of the envelope, and a slip indicating the participant's allocated group was kept inside. After completing all baseline assessments, the research staff provided the sealed envelope to the eligible participant. The participant opened the envelope and was then assigned to the group as indicated on the slip inside. The opened envelope and the printed slip were attached to the participant's CRF for record and trial monitoring.

Participants assigned to the 2 study groups were housed in separate barracks throughout the study to prevent any protocol violations and minimize the risk of contamination bias. Further, yoga sessions and supplement distribution were conducted at designated areas accessible only to the intervention group, with attendance monitored by site instructors. The CG did not receive access to these sessions or materials, reducing the likelihood of unintentional participation. Study staff members were instructed to avoid sharing intervention-related information with control-group participants.

Blinding

The study is open-label; however, to minimize any potential bias, the research staff assigned to assess the study outcomes will be blinded to the allocation of participants to the study groups. In addition, laboratory technicians analyzing blood samples were blinded through the use of coded specimens. The intervention team (Ayurveda practitioner and yoga instructor) had no role in outcome assessments, data entry, or laboratory procedures. The assessment team functioned independently and interacted with participants only at scheduled assessment time points. They were instructed not to inquire about participants' allocation or intervention experiences. Each participant was assigned a unique study ID; all CRFs, questionnaires, and lab requisitions contained only this ID and not the allocation label. Further, statistical analysis will be conducted by an independent data analysis team not involved in intervention delivery or participant contact.

Data Collection

The baseline demographics, clinical, and physical examination aspects were collected by qualified Ayush study personnel and reported in a CRF designed for this purpose. Subjective and objective outcome assessments were conducted according to the protocol (Table 3).

XSL•FO
RenderX

Table . Study schedule.

	Screening	Baseline	Day 1 to 7	Day 10 (for participants at stage II HA[a])	Day 30	Day 60	In case of early descent	Day 120
Written informed consent	✓							
Eligibility evaluation	✓							
Demographics and medical history		✓						
Clinical assessment	✓	✓	✓	✓	✓	✓	✓	✓
Vital parameters (resting heart rate, systolic and diastolic blood pressure, respiratory rate, SpO_2[b])	✓	✓	✓	✓	✓	✓	✓	✓
CBC[c]	✓			✓		✓	✓	✓
LFT[d], KFT[e]	✓					✓	✓	✓
HbA_{1c}[f] and ECG[g]	✓							
Lipid profile		✓				✓	✓	✓
LDH[h], CK[i], CK-MB[j], fibrinogen, aPTT[k], prothrombin time, hs-CRP[l], IL-6[m], TNF-α[n]		✓		✓		✓	✓	✓
Lake Louise Scoring System for AMS[o] assessment		✓	✓	✓				
Assessment for thromboembolic events						✓	✓	✓
DASS-21[p], PSQI[q], SF-12[r] questionnaire		✓		✓		✓	✓	✓
Rescue medication			✓	✓	✓	✓	✓	✓
Assessment of adverse events			✓	✓	✓	✓	✓	✓
Assessment of drug compliance			✓	✓	✓	✓	✓	✓

[a]HA: high altitude.

[b]SpO_2: peripheral capillary oxygen saturation.

XSL•FO
RenderX

[c]CBC: complete blood count.
[d]LFT: liver function test.
[e]KFT: kidney function test.
[f]HbA_{1c}: glycated hemoglobin.
[g]ECG: electrocardiogram.
[h]LDH: Lactate Dehydrogenase.
[i]CK: creatine kinase.
[j]MB: myocardial band.
[k]aPTT: Activated Partial Thromboplastin Time.
[l]hs-CRP: high-sensitivity C-reactive protein.
[m]IL-6: interleukin-6.
[n]TNF-α: tumor necrosis factor-alpha.
[o]AMS: acute mountain sickness.
[p]DASS-21: Depression Anxiety Stress Scale-21 items.
[q]PSQI: Pittsburgh Sleep Quality Index.
[r]SF-12: 12-item short-form.

The incidence of AMS and its severity were assessed using the LLSS AMS score, according to which AMS is diagnosed by having a headache score of at least 1 point and a total score of at least 3 points [9]. This was assessed daily for the first day in participants at stage I high altitude and for an additional 4 days in participants who progressed to stage II high altitude. The self-reported emotional states of depression, anxiety, and stress were reported by adding scores of the relevant items in the 3 subscales of the DASS-21 [28]. The quality of sleep and overall quality of life were assessed through validated scales, the PSQI and the SF-12 questionnaire, respectively [29-34]. The secondary outcomes, including DASS-21, PSQI, and SF-12, were assessed at baseline and at follow-ups on days 60 and 120.

The serum sample for the objective assessment parameters, such as coagulation and hemostasis activation markers, proinflammatory markers, LFT and RFT, etc, was collected at baseline, day 60, and day 120 and transported to an NABL-accredited laboratory, and the data received from the lab were entered into the CRFs and electronic CRFs. The research team underwent training on the study protocol, standard operating procedures for the conduct of the study, storage, and dispensing of interventions, handling of biological samples, data collection, and recording to ensure compliance with GCP principles while ensuring participant safety, data accuracy, and reliability.

Data Management

Data management in this clinical study adhered to stringent guidelines to ensure the accuracy, reliability, and integrity of the collected information. Upon the participants' assessments, the research team promptly entered the data into CRFs and electronic CRFs. Source documents and CRFs were securely stored in restricted-access areas, limited solely to the study team. Electronic CRFs were password-protected and stored in secure, access-restricted computer systems.

Data entered by the study personnel was subjected to meticulous cross-verification by the principal investigators at the study site, ensuring the reliability of the data. Rigorous quality measures were implemented, such as regular audits, to identify and address any discrepancies in the data. The data management practices adhered to regulatory guidelines and ethical principles, prioritizing the protection of participants' rights.

Statistical Methods

The categorical data will be presented as numbers (percentages) and compared between groups using the χ^2 test, while within-group comparisons will be performed by using the McNemar or Cochran Q-test. Continuous data following a normal distribution will be reported as mean (SD), and between-group comparisons will be performed using an independent sample *t*-test. Within-group comparisons for normal data will be performed using a paired sample *t*-test/repeated measures ANOVA. Nonnormal data will be reported as median (first quartile, third quartile), and between-group comparisons will be conducted using the Mann-Whitney test. Within-group comparison will be done using the Wilcoxon signed-rank test or Friedman test. Baseline variables, such as age, prior altitude exposure, and hydration-related indicators, will be examined descriptively and adjusted in multivariable models where relevant. Regression analysis will be performed to study the effect of confounding variables on the incidence of AMS. The data of participants having early descent will be analyzed with censored data models. A *P* value less than .05 will be considered significant. The SPSS version 28.0 is being used for statistical analysis.

The modified intention-to-treat analysis approach will be applied to handle missing data. The missing data for all participants for whom data are available for at least one visit post-baseline will be imputed. The last observation carried forward method will be used for imputing the missing values.

Monitoring

The Data and Safety Monitoring Board monitored the study for quality and regulatory compliance. The Data and Safety Monitoring Board reviewed the study's progress every 3 months until the end of the study period.

Trial Audit

The on-site monitoring visit by an independent committee constituted by the sponsor was planned to ensure that the study procedures and data collection processes adhered to existing

XSL•FO
RenderX

regulatory standards and to verify the accuracy, completeness, legibility, and timeliness of the reported study data.

Ethical Considerations

The Institutional Ethics Committee has approved the study protocol to ensure compliance with ethical standards and safeguard the rights and well-being of the participants. The study has been registered prospectively at the Clinical Trial Registry of India (CTRI/2023/03/051028). The study was conducted in accordance with the principles of the Declaration of Helsinki, the ICMR's National Ethical Guidelines for Biomedical and Health Research on Human Participants (2017) [35], and the ICH Good Clinical Practices guidelines [36]. No substantial amendments in the study protocol were made that may affect patient safety or study integrity. Before undergoing any study-related procedure, potential participants received a participant information sheet in their native language or Hindi. The participant information sheet comprehensively outlined the various aspects of the study, equipping participants with the necessary information to make an informed decision about participating in the study. Written consent was obtained through the consent form, which was signed by both the participant and the study personnel delegated to the task. No financial compensation will be provided to participants for their involvement in the study. To ensure participant safety, clinical trial insurance has been secured to cover any unexpected adverse events or study-related injuries.

The results and findings of the study will be disseminated in accordance with best practices in scientific publishing, ensuring that the knowledge generated benefits the scientific community and the public at large through publication in peer-reviewed, indexed medical journals, and presentations at national and international conferences.

Confidentiality

All the relevant study data were stored securely at the study site with password-protected access systems in areas with limited access. To maintain participant confidentiality, a coded enrollment identification number was used to identify all laboratory specimens, reports, data collection, and relevant forms. All records containing names or other personal identifiers, such as informed consent forms, were stored separately from the study records and identified by a code number in a restricted-access area.

Ancillary and Posttrial Care

No ancillary studies were proposed for the present clinical study. If required, participants were provided with routine medical care after completing the study period.

Results

The study was funded in March 2023. The data collection was completed in December 2023. A total of 1660 participants were enrolled in the study. The analysis of the study data is in progress. The study outcomes are expected to be published by December 2026.

Discussion

Novel Aspects of the Study

High-altitude regions pose unique physiological and psychological challenges. The present publication outlines a groundbreaking RCT designed to assess the impact of holistic Ayush interventions, consisting of Ayurvedic nutritional supplements (Ayur-nutri kit) and Yoga, in conjunction with the existing AS, on reducing the incidence of AMS, while also addressing hemostatic and vascular function, psychological well-being, and overall quality of life among personnel at high altitude.

A conventional AS is routinely followed to prevent HAI. It is primarily based on a strategy that combines the conventional staged and graded ascents practiced by mountaineers [27]. The 3-stage AS extends up to 14 days to reach stage III high-altitude regions (more than 15,000 feet) [37]. Stage I acclimatization lasts for 6 days for stage I high altitude (9000 - 12000 feet), whereas an additional 4 days each are required to achieve stage II (12,000 - 15,000 feet) and stage III (more than 15,000 feet) acclimatization [27,37]. During reinduction to high-altitude regions, after 10 to 30 days of absence from high altitude, 4 days at each stage need to be spent, while in the case of more than 30 days, a full AS as a fresh inductee needs to be complied [37]. However, effective interventions for preventing HAI with a better safety profile must be explored to facilitate better acclimatization, shorter ascent times to reach high-altitude regions, and optimal use of limited resources. It could also have logistical implications, such as the early availability of healthy personnel for high-altitude regions.

The preliminary evidence suggests that the proposed Ayush interventions, viz Ayurvedic nutritional supplements and yoga protocols are safe and can potentially alleviate the pathophysiology of HAI, aid in better acclimatization, and reduce the incidence and severity of AMS [22-25]. Furthermore, the herbal ingredients of the Ayur-nutri kit possess cardioprotective, antidyslipidemic, antihypertensive, anti-inflammatory, antiatherogenic, antiplatelet aggregation, antithrombotic, and anticoagulant properties [17-21]. Therefore, these supplements may also have potential benefits in maintaining healthy circulation, thereby possibly reducing thromboembolic events at high altitude by showing favorable effects on markers of coagulation and hemostasis activation, as well as proinflammatory markers. The integration of nuts and seeds in the Ayur-nutri kit provides nutrient-dense nourishment that ensures an adequate supply of essential vitamins, minerals, and healthy fats, while serving as an excellent energy source, providing sustained fuel for physical activities and supporting cardiovascular health. The 60-minute yoga protocol, tailored for high altitude, can offer a wide range of benefits to enhance resilience and promote physical and psychological well-being. Yoga cultivates mental and emotional well-being, which is vital for personnel coping with challenging and unpredictable high-altitude conditions. Meditation and relaxation techniques can improve sleep, which is often disrupted at high altitudes. Better sleep quality supports cognitive function, mood stability, and overall well-being. A well-structured Ayush-AMS protocol

XSL•FO
RenderX

that addresses the unique needs and challenges of high altitude can promote holistic well-being.

The present RCT represents an innovative approach by incorporating a holistic strategy in mitigating the challenges posed by AMS and the various physiological and psychological stressors associated with high altitude. The comprehensive Ayush-AMS regimen has the potential to inform future protocols in HAI and offer valuable contributions to the broader field of high-altitude medicine.

Strengths and Limitations

The present RCT protocol has particular key strengths. First, the use of nutritional supplements enriched with Ayurveda herbs, which are easy to use, palatable, and safe, with the potential to address AMS and related conditions. Second, the study has a large sample size, validated subjective assessment tools, and biomarkers to address various circulation and hemostasis-related parameters. This protocol also has a few limitations, including the open-label design of the study. However, to remove any potential bias, the research staff assigned to assess the study outcomes will be unaware of the allocation of participants to the study groups. Furthermore, there is an unequal attention between groups (due to oral interventions and supervised yoga routines in the AG) that may influence participants' perceptions. Additionally, the altitude-exposure stage may influence the onset patterns of symptoms. Also, as the study is exploratory in nature for the secondary outcomes, it lacks power for multiple hypothesis testing across biomarkers, and no hierarchical testing plan or multiplicity adjustment was prespecified.

Expected Outcomes

The outcomes of the present RCT are expected to facilitate the optimal use of limited resources through better acclimatization, reduced ascent time, and the rapid induction of healthy personnel at high-altitude regions during exigency situations.

Acknowledgments

The authors acknowledge the support provided by Padmashri Vaidya Rajesh Kotecha, Secretary, Ministry of Ayush, who has been instrumental in conceptualizing the study and facilitating collaboration. Their heartfelt thanks to Dr Anant Mohan, Professor and Head, Department of Pulmonary, Critical Care and Sleep Medicine, All India Institute of Medical Sciences, New Delhi, for his meticulous review of the study protocol and valuable inputs. Generative artificial intelligence was not used to draft or revise any section of the manuscript.

Funding

The Central Council for Research in Ayurvedic Sciences (CCRAS), Ministry of Ayush, Government of India, funded the study. The funding agency also provided technical support in conceptualizing and designing the study, and played a role in data collection, management, analysis, and interpretation; writing the clinical study report; and the decision to submit the final clinical study report for publication in an indexed medical journal.

Data Availability

The data related to the outcomes of this study are available from the funding agency upon reasonable request.

Authors' Contributions

Conceptualization: AKR, SJ, AKY, BCR, NS
Funding acquisition: RA
Investigation: AKY
Methodology: AKR, SJ, AKY, RKJ, PM, AA, BM, BSS, SK, BY, BC, RR
Project administration: AKR, AKY, BSS, SK, BY, NS
Resources: BSS, SK, BY, RA
Statistical methods: RKR, AT
Supervision: BC, NS, RA
Writing – original draft: AKR
Writing – review & editing: SJ, AKY, BSS, SK, BY, BC, NS, RA

Conflicts of Interest

The authors declare that they have no known competing interests. However, the study protocol, data collection procedures, and analytical framework were jointly developed by the funding agency, CCRAS, and collaborators following internal scientific and ethical review processes. Although CCRAS supported the study infrastructure and data management, the investigators retained full academic independence in interpreting the study findings.

References

1. Basnyat B, Murdoch DR. High-altitude illness. Lancet 2003 Jun 7;361(9373):1967-1974. [doi: 10.1016/S0140-6736(03)13591-X] [Medline: 12801752]

XSL•FO
RenderX

2. Bärtsch P, Swenson ER. Clinical practice: acute high-altitude illnesses. N Engl J Med 2013 Jun 13;368(24):2294-2302. [doi: 10.1056/NEJMcp1214870] [Medline: 23758234]
3. Waeber B, Kayser B, Dumont L, Lysakowski C, Tramèr MR, Elia N. Impact of study design on reported incidences of acute mountain sickness: a systematic review. High Alt Med Biol 2015 Sep;16(3):204-215. [doi: 10.1089/ham.2015.0022] [Medline: 26230550]
4. Schneider M, Bernasch D, Weymann J, Holle R, Bartsch P. Acute mountain sickness: Influence of susceptibility, preexposure, and ascent rate. Med Sci Sports Exerc 2002 Dec;34(12):1886-1891. [doi: 10.1097/00005768-200212000-00005] [Medline: 12471292]
5. Hackett PH, Roach RC. High-altitude illness. N Engl J Med 2001 Jul 12;345(2):107-114. [doi: 10.1056/NEJM200107123450206] [Medline: 11450659]
6. Maggiorini M, Bühler B, Walter M, Oelz O. Prevalence of acute mountain sickness in the Swiss Alps. BMJ 1990 Oct 13;301(6756):853. [doi: 10.1136/bmj.301.6756.853] [Medline: 2282425]
7. Honigman B, Theis MK, Koziol-McLain J, et al. Acute mountain sickness in a general tourist population at moderate altitudes. Ann Intern Med 1993 Apr 15;118(8):587-592. [doi: 10.7326/0003-4819-118-8-199304150-00003] [Medline: 8452324]
8. Karinen H, Peltonen J, Tikkanen H. Prevalence of acute mountain sickness among Finnish trekkers on Mount Kilimanjaro, Tanzania: an observational study. High Alt Med Biol 2008;9(4):301-306. [doi: 10.1089/ham.2008.1008] [Medline: 19115914]
9. Roach RC, Hackett PH, Oelz O, et al. The 2018 Lake Louise acute mountain sickness score. High Alt Med Biol 2018 Mar;19(1):4-6. [doi: 10.1089/ham.2017.0164] [Medline: 29583031]
10. Sridharan K, Sivaramakrishnan G. Pharmacological interventions for preventing acute mountain sickness: a network meta-analysis and trial sequential analysis of randomized clinical trials. Ann Med 2018 Mar;50(2):147-155. [doi: 10.1080/07853890.2017.1407034] [Medline: 29166795]
11. Gao D, Wang Y, Zhang R, Zhang Y. Efficacy of acetazolamide for the prophylaxis of acute mountain sickness: a systematic review, meta-analysis, and trial sequential analysis of randomized clinical trials. Ann Thorac Med 2021;16(4):337-346. [doi: 10.4103/atm.atm_651_20] [Medline: 34820021]
12. Dumont L, Mardirosoff C, Tramèr MR. Efficacy and harm of pharmacological prevention of acute mountain sickness: quantitative systematic review. BMJ 2000 Jul 29;321(7256):267-272. [doi: 10.1136/bmj.321.7256.267] [Medline: 10915127]
13. Pandit A, Karmacharya P, Pathak R, Giri S, Aryal MR. Efficacy of NSAIDs for the prevention of acute mountain sickness: a systematic review and meta-analysis. J Community Hosp Intern Med Perspect 2014;4(4). [doi: 10.3402/jchimp.v4.24927] [Medline: 25317267]
14. Rege NN, Thatte UM, Dahanukar SA. Adaptogenic properties of six rasayana herbs used in Ayurvedic medicine. Phytother Res 1999 Jun;13(4):275-291. [doi: 10.1002/(SICI)1099-1573(199906)13:4<275::AID-PTR510>3.0.CO;2-S] [Medline: 10404532]
15. Kumar D, Arya V, Kaur R, Bhat ZA, Gupta VK, Kumar V. A review of immunomodulators in the Indian traditional health care system. J Microbiol Immunol Infect 2012 Jun;45(3):165-184. [doi: 10.1016/j.jmii.2011.09.030] [Medline: 22154993]
16. Govindarajan R, Vijayakumar M, Pushpangadan P. Antioxidant approach to disease management and the role of “Rasayana” herbs of Ayurveda. J Ethnopharmacol 2005 Jun 3;99(2):165-178. [doi: 10.1016/j.jep.2005.02.035] [Medline: 15894123]
17. Mukherjee PK, Banerjee S, Biswas S, Das B, Kar A, Katiyar CK. Withania somnifera (L.) Dunal - modern perspectives of an ancient Rasayana from Ayurveda. J Ethnopharmacol 2021 Jan 10;264:113157. [doi: 10.1016/j.jep.2020.113157] [Medline: 32783987]
18. Wakade AS, Shah AS, Kulkarni MP, Juvekar AR. Protective effect of Piper longum L. on oxidative stress induced injury and cellular abnormality in adriamycin induced cardiotoxicity in rats. Indian J Exp Biol 2008 Jul;46(7):528-533. [Medline: 18807757]
19. Lee W, Yoo H, Ku SK, Kim JA, Bae JS. Anticoagulant activities of piperlonguminine in vitro and in vivo. BMB Rep 2013 Oct;46(10):484-489. [doi: 10.5483/bmbrep.2013.46.10.028] [Medline: 24148768]
20. Shadrack K, Faraj A, Alex MK, Kenneth N. Anti-thrombotic effect of Zingiber officinale (ginger) in Sprague Dawley rats. Int J Res Med Sci 2019 Sep;7(9):3239-3245. [doi: 10.18203/2320-6012.ijrms20193905]
21. Salve BA, Tripathi RK, Petare AU, Raut AA, Rege NN. Effect of Tinospora cordifolia on physical and cardiovascular performance induced by physical stress in healthy human volunteers. Ayu 2015;36(3):265-270. [doi: 10.4103/0974-8520.182751] [Medline: 27313412]
22. Bansal P, Sannd R, Srikanth N, Lavekar GS. Effect of traditionally designed nutraceutical on stress induced immunoglobulin changes at Antarctica. Afr J Biochem Res 2009;3(4):084-088 [FREE Full text]
23. Himashree G, Mohan L, Singh Y. Yoga practice improves physiological and biochemical status at high altitudes: a prospective case-control study. Altern Ther Health Med 2016 Sep;22(5):53-59. [Medline: 27622961]
24. Bernardi L, Passino C, Spadacini G, et al. Reduced hypoxic ventilatory response with preserved blood oxygenation in yoga trainees and Himalayan Buddhist monks at altitude: evidence of a different adaptive strategy? Eur J Appl Physiol 2007 Mar;99(5):511-518. [doi: 10.1007/s00421-006-0373-8] [Medline: 17206440]

XSL•FO
RenderX

25. Spatenkova V, Bednar R, Oravcova G, Melichova A, Kuriscak E. Yogic breathing in hypobaric environment: breathing exercising and its effect on hypobaric hypoxemia and heart rate at 3,650-m elevation. J Exerc Rehabil 2021 Aug;17(4):270-278. [doi: 10.12965/jer.2142324.162] [Medline: 34527639]
26. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346:e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
27. Sikri G, Kotwal A, Singh SP, et al. Is it time to revise the acclimatization schedule at high altitude? Evidence from a field trial in western Himalayas. Med J Armed Forces India 2019 Jul;75(3):251-258. [doi: 10.1016/j.mjafi.2018.01.001] [Medline: 31388226]
28. Moya E, Larson LM, Stewart RC, Fisher J, Mwangi MN, Phiri KS. Reliability and validity of Depression Anxiety Stress Scale (DASS)-21 in screening for common mental disorders among postpartum women in Malawi. BMC Psychiatry 2022 May 24;22(1):352. [doi: 10.1186/s12888-022-03994-0] [Medline: 35606733]
29. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
30. Manzar MD, Moiz JA, Zannat W, et al. Validity of the Pittsburgh Sleep Quality Index in Indian university students. Oman Med J 2015 May;30(3):193-202. [doi: 10.5001/omj.2015.41] [Medline: 26171126]
31. Wang L, Wu YX, Lin YQ, et al. Reliability and validity of the Pittsburgh Sleep Quality Index among frontline COVID-19 health care workers using classical test theory and item response theory. J Clin Sleep Med 2022 Feb 1;18(2):541-551. [doi: 10.5664/jcsm.9658] [Medline: 34534069]
32. Ruotolo I, Berardi A, Sellitto G, et al. Criterion validity and reliability of SF-12 Health Survey Version 2 (SF-12v2) in a student population during COVID-19 pandemic: a cross-sectional study. Depress Res Treat 2021;2021:6624378. [doi: 10.1155/2021/6624378] [Medline: 34394986]
33. Shou J, Ren L, Wang H, et al. Reliability and validity of 12-item Short-Form health survey (SF-12) for the health status of Chinese community elderly population in Xujiahui district of Shanghai. Aging Clin Exp Res 2016 Apr;28(2):339-346. [doi: 10.1007/s40520-015-0401-9] [Medline: 26142623]
34. Huo T, Guo Y, Shenkman E, Muller K. Assessing the reliability of the short form 12 (SF-12) health survey in adults with mental health conditions: a report from the wellness incentive and navigation (WIN) study. Health Qual Life Outcomes 2018 Feb 13;16(1):34. [doi: 10.1186/s12955-018-0858-2] [Medline: 29439718]
35. National ethical guidelines for biomedical and health research involving human participants. : Indian Council of Medical Research; 2017 URL: https://ethics.ncdirindia.org/asset/pdf/ICMR_National_Ethical_Guidelines.pdf [accessed 2026-03-31]
36. ICH harmonised guideline: guideline for good clinical practice E6(R3). : International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use; 2025 URL: https://database.ich.org/sites/default/files/ICH_E6%28R3%29_Step4_FinalGuideline_2025_0106.pdf [accessed 2026-03-31]
37. Syed VS, Sharma S, Singh RP. Determinants of acclimatisation in high altitude. Med J Armed Forces India 2010 Jul;66(3):261-265. [doi: 10.1016/S0377-1237(10)80052-8] [Medline: 27408314]

Abbreviations

AE: adverse event
AG: Ayush group
AMS: acute mountain sickness
AS: acclimatization schedule
CCRAS: Central Council for Research in Ayurvedic Sciences
CCT: cardiac care tea
CG: control group
CRF: case report form
HAI: high-altitude illness
HAPE: high-altitude pulmonary edema
ICH: International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
LLSS: Lake Louise Scoring System
PSQI: Pittsburgh Sleep Quality Index
RCT: randomized controlled trial
SF-12: 12-item short-form
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

XSL•FO
RenderX

Edited by J Sarvestan; submitted 30.Jul.2025; peer-reviewed by K Dwivedi, RB Bhandari, S Bhinde; revised version received 17.Nov.2025; accepted 05.Dec.2025; published 08.Apr.2026.

Please cite as:
Rai AK, Jameela S, Yadav AK, Joshi RK, Mundada P, Ahmad A, Mahajon B, Sharma BS, Khanduri S, Yadav B, Tripathi A, Rana RK, Chandrasekhararao B, Srikanth N, Rao R, Acharya R
Effectiveness of Ayurvedic Nutritional Supplements and Yoga Protocol in Reducing the Incidence and Severity of Acute Mountain Sickness (AYAMS Study): Study Protocol for an Open-Label Randomized Controlled Trial
JMIR Res Protoc 2026;15:e81567
URL: https://www.researchprotocols.org/2026/1/e81567
doi:10.2196/81567

© Amit Kumar Rai, Sophia Jameela, Arun Kumar Yadav, Rajneesh Kumar Joshi, Pallavi Mundada, Azeem Ahmad, Bidhan Mahajon, Bhagwan Sahai Sharma, Shruti Khanduri, Babita Yadav, Arunabh Tripathi, Rakesh Kumar Rana, Bhogavalli Chandrasekhararao, Narayanam Srikanth, Raghavendra Rao, Rabinarayan Acharya. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effects of Orexin Receptor Antagonism During Early Withdrawal in Stimulant Use Disorder: Protocol for a Proof-of-Concept Study

Heather Webber[1], PhD; Nikky Tran[1], BS; Rita Blake[2], BA; Sarah Goldman[2], MSW; Trevor George[3], ND; Bill Bailey[3], BBA; Robert Suchting[1], PhD; Jin Yoon[1], PhD; Jessica Vincent[1], BS; Michael Weaver[1], MD; Joy Schmitz[1], PhD; Scott Lane[1], PhD

[1]Faillace Department of Psychiatry and Behavioral Sciences, McGovern Medical School, University of Texas Health Science Center at Houston, 1941 East Rd, BBSB 1st floor, Houston, TX, United States

[2]The Right Step - Houston, Houston, TX, United States

[3]Cenikor Foundation, Houston, TX, United States

Corresponding Author:
Heather Webber, PhD
Faillace Department of Psychiatry and Behavioral Sciences, McGovern Medical School, University of Texas Health Science Center at Houston, 1941 East Rd, BBSB 1st floor, Houston, TX, United States

Abstract

Background: Stimulant use disorders (StimUDs), including cocaine and methamphetamine use disorders, remain a major public health concern in the United States with no US Food and Drug Administration (FDA)–approved pharmacological treatments. The early recovery period following stimulant use is marked by sleep disturbances, heightened stress reactivity, and dysregulated reward processing, contributing to high rates of resumption of use. Evidence from animal and human studies supports the orexin neuropeptide system as a promising therapeutic target for minimizing these disruptive mechanisms. Suvorexant, an FDA-approved dual orexin receptor antagonist for insomnia, has previously been shown to improve sleep, reduce stress, and attenuate craving in the first-in-human study with non–treatment-seeking individuals with cocaine use disorder.

Objective: The primary objective of this randomized open-label clinical trial is to evaluate the effects of suvorexant on sleep, stress, reward processing, and craving during early abstinence in treatment-seeking individuals with StimUD. This proof-of-concept trial represents the first mechanistic evaluation of suvorexant in a residential sample of individuals with StimUD.

Methods: A total of 40 participants, recruited from a residential treatment facility, will be randomized (1:1) to either 7 days of nightly suvorexant (10 mg) or treatment as usual. Participants' sleep will be monitored using an activity tracker watch, and they will complete a sleep diary each night to confirm bedtimes and wake-up times. Participants in the suvorexant group will receive the study medication for 7 nights. On premedication and postmedication study days, participants will be transported to the Center for Neurobehavioral Research on Addiction to complete the study sessions. At both study sessions, participants will complete self-report and safety measures (safety measures for the suvorexant group only) followed by administration of a stress reactivity test and an electroencephalographic assessment.

Results: The study was funded in January 2025 and is currently ongoing. We began enrollment of the first participant prior to funding in July 2024. At the time of manuscript submission (September 2025), a total of 8 participants had been enrolled in the study toward the planned enrollment of 40 participants. This study is expected to conclude in January 2027.

Conclusions: Findings from this study will provide initial evidence for the mechanisms involved in suvorexant for StimUD, with the long-term goal of repurposing orexin-based therapy to regulate neurobehavioral mechanisms and inform future medication development for StimUD.

Trial Registration: ClinicalTrials.gov NCT06444256; https://clinicaltrials.gov/study/NCT06444256

International Registered Report Identifier (IRRID): DERR1-10.2196/83842

(JMIR Res Protoc 2026;15:e83842) doi:10.2196/83842

KEYWORDS

stimulants; substance use disorder; suvorexant; treatment; orexin; sleep; stress; reward

XSL•FO
RenderX

Introduction

Stimulant use disorders (StimUDs), including cocaine and methamphetamine use disorders, represent a growing public health crisis in the United States. Unlike opioid use disorder, for which multiple pharmacotherapies exist, StimUDs remain without a US Food and Drug Administration (FDA)–approved medication, posing a significant challenge for clinical management. Historically, pharmacological interventions have targeted monoaminergic systems—particularly dopaminergic pathways—with limited efficacy [1]. This therapeutic gap underscores the urgent need for novel, mechanism-based treatments that address the neurobiological underpinnings of stimulant addiction and withdrawal.

One promising target is the orexin (hypocretin) system, a neuropeptide network originating in the lateral hypothalamus that plays a central role in regulating arousal, stress, and reward processing. Converging evidence from preclinical and translational studies implicates orexin signaling in the maintenance of drug-seeking behavior, particularly under conditions of heightened stress or disrupted sleep [2]. Of particular relevance to early abstinence is the role of orexin in sleep-wake regulation. Sleep disturbances are a hallmark of stimulant withdrawal and are strongly predictive of resumption of use [3]. Therefore, pharmacological modulation of orexin receptors may offer a dual mechanism of action—ameliorating sleep disruption while attenuating drug-motivated behavior.

Dual orexin receptor antagonists, such as suvorexant, have demonstrated efficacy in improving sleep architecture and reducing hyperarousal in individuals with insomnia. In a recent proof-of-concept study, suvorexant was shown to improve sleep quality, reduce stress, and attenuate craving in non–treatment-seeking individuals with cocaine use disorder [4]. Building on these findings, this study aims to evaluate the effects of suvorexant during early withdrawal in treatment-seeking individuals with StimUD, including those using both cocaine and methamphetamine, as they transition into residential addiction treatment.

To extend the mechanistic understanding of suvorexant's effects, we incorporate electroencephalography (EEG) measures of brain function. Two neural markers are of particular interest. First, resting-state alpha power (8 - 13 Hz), typically observed over posterior scalp regions during relaxed wakefulness, is sensitive to sleep pressure and may serve as an objective index of daytime neural deactivation [5]. Preliminary data from our group suggest that increased alpha power is associated with poorer subjective sleep quality in people who use stimulants. Second, the reward positivity (RewP) component—an event-related potential elicited by positive prediction errors—indexes reward sensitivity and is often blunted in individuals with substance use disorders [6]. Given the orexin system's role in (1) modulating mesolimbic dopamine pathways and (2) normalizing sleep architecture, suvorexant may enhance neural responsiveness to nondrug stimuli (potentially both cues and rewards), thereby supporting recovery-related motivational processes.

The primary objective of this translational study is to evaluate the effects of suvorexant on sleep during early abstinence from stimulants—a critical window in which effective withdrawal management may reduce the risk of return to stimulant use. The secondary objective is to examine whether suvorexant will also improve reward processing and reduce stress and craving during early abstinence. By integrating behavioral, subjective, and neurophysiological measures, this investigation represents the first mechanism-focused clinical evaluation of orexin antagonism in StimUD. We hypothesize that suvorexant, relative to treatment as usual (TAU), will improve sleep outcomes during early abstinence (aim 1). We also hypothesize that the secondary outcomes (EEG resting alpha power, acute stress reactivity, non–drug-reward processing, and drug demand) will improve from pretreatment to posttreatment in the suvorexant group compared with the TAU group (aim 2). Finally, we will examine treatment engagement and the effects of the treatment group (suvorexant vs TAU) on posttreatment (days 13 - 30) residential program length of stay and completion rates (aim 3). We hypothesize that participants in the suvorexant group, compared with the TAU group, will demonstrate longer posttreatment lengths of stay and higher residential program completion rates.

Methods

Study Design

We will use an open-label, randomized 2-group design. Participants (n=40) will be randomized to receive 7 days of suvorexant (10 mg) or TAU, consisting of supportive care and symptomatic medication. Figure 1 outlines the schedule of events. Informed consent, eligibility criteria, and baseline sleep data (via actigraphy) will be assessed on day 0 of the study via telephone or video call. Day 0 will ideally occur around 4 days (but within the first week) following admission to the residential program. On day 1 and day 8, participants will be transported to the outpatient treatment research clinic at the Center for Neurobehavioral Research on Addiction (CNRA) to complete the study sessions. At each study session, participants will complete self-report and safety measures (safety measures for the suvorexant group only) followed by administration of the stress reactivity test and then the EEG assessment. Details regarding each of these procedures are provided in subsequent sections. Participants in the suvorexant group will begin receiving the study medication, administered by residential nurses, from day 1 up until day 8 (ie, nights 1 - 7). All participants will remain in the residential program until either regular discharge at approximately 30 days ("successful completion") or early departure (eg, patient-directed discharge or leaving against medical advice). This design captures the critical first 2 weeks of stimulant withdrawal, when sleep disruptions and other withdrawal-related symptoms are typically most pronounced [7].

XSL•FO
RenderX

Figure 1. Study flow diagram. Consent and randomization will occur virtually. Laboratory visits will occur at the Center for Neurobehavioral Research on Addiction. All other activities will occur at the inpatient facilities. NSF: National Sleep Foundation; TAU: treatment as usual.

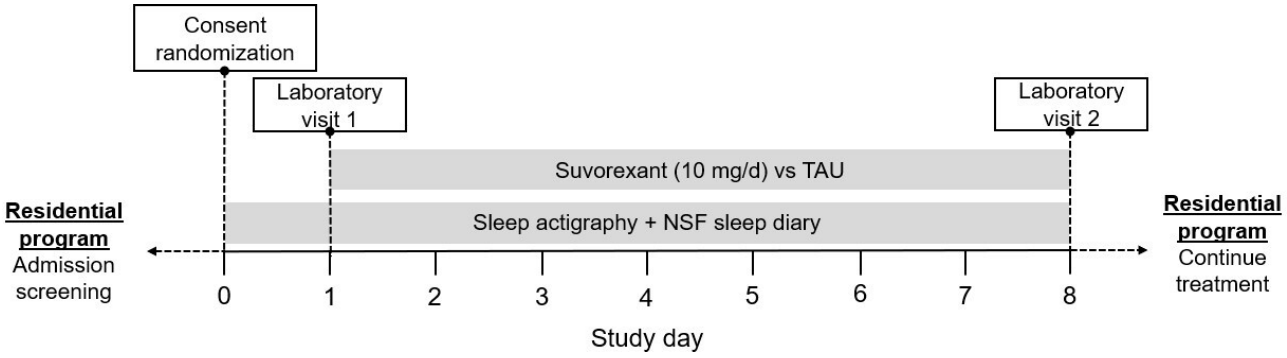

Recruitment

Participants will be recruited directly from 2 local inpatient residential treatment facilities in the greater Houston area (The Right Step Houston and Cenikor), both of which provide similar therapeutic programming and services. Patients meeting the preliminary eligibility criteria will be referred to the CNRA by the residential programs. Additional recruitment will include online advertisements and outreach through the University of Texas Health Science Center at Houston (UTHealth Houston), Department of Psychiatry Recruitment Registry. Individuals recruited from outside the residential programs must voluntarily seek admission to one of the residential facilities before participating. Eligibility will be assessed using residential intake records (including a history and physical or psychological assessment) and a phone-based screening. Eligible candidates will participate in a phone interview including the consent process and administration of the structured clinical interview for *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*) [8] by trained CNRA staff (ie, licensed professional counselor, doctoral student, or postdoctoral fellow). The principal investigator (PI; author HW) and the medical director (author MW) will review intake and phone-interview information to determine final participant eligibility.

Participants

Participants will be eligible for the study if they are aged between 18 and 65 years and meet the *DSM-5* criteria for a primary StimUD, specifically related to cocaine or methamphetamine. All participants must be fluent in English and capable of fully understanding and providing informed consent for study procedures. These inclusion criteria ensure that participants have the condition of interest and can engage meaningfully with study requirements.

Participants will be excluded if they meet any of the following criteria: (1) diagnosis of an opioid use disorder of any severity; (2) diagnosis of greater than moderate substance use disorder for any substance other than cocaine or methamphetamine; (3) currently receiving medication treatment for withdrawal from any substance; (4) presence of medical conditions that contraindicate the use of suvorexant (eg, severe pulmonary disease, severe cardiovascular disease or clinically abnormal electrocardiogram, severe liver or kidney disease, or sleep disorder—particularly narcolepsy); (5) use of medications known to interact with suvorexant (eg, monoamine oxidase inhibitors, anticonvulsants, haloperidol, phenothiazines, anesthetics, and any sedative); (6) pregnancy or breastfeeding; (7) BMI >30 kg/m^2 (women only); (8) current *DSM-5* psychiatric disorder or neurological disease requiring ongoing treatment that would make participation unsafe; (9) history of seizure disorder; and (10) history of head injury with loss of consciousness within the past 5 years.

Randomization

Participants will be randomly assigned in a 1:1 allocation ratio to each treatment group using block randomization. Allocation will use permuted blocks with randomly varying per-arm sizes to minimize predictability while preserving balance. The randomization sequence will be generated in the R Statistical Computing Environment (R Foundation for Statistical Computing) via the blockrand package [9]. A blinded code will be provided to an unblinded study team member who will execute the script, export the allocation file, and upload it to the project's randomization module in REDCap (Research Electronic Data Capture). The full table will be access restricted such that staff members will see only the next assignment when needed. A read-only copy will be kept safe on a secure drive for emergency unblinding.

Intervention

Suvorexant is a dual orexin receptor antagonist approved by the FDA for the treatment of insomnia. In this study, we will administer 10 mg of suvorexant for 7 nights. Two prior studies from our group suggested that suvorexant could be safely administered to individuals with cocaine use disorder [4] and methamphetamine use disorder [10]. The half-life is about 12 hours on average, and steady-state plasma concentrations occur in 3 days with daily administration [11]. A nurse at the residential programs will administer the medication to participants in the suvorexant group each evening before bedtime, at the same time nightly (around 10 PM), with a medication log to monitor compliance.

All participants will receive TAU, consisting of standard programming from the inpatient facilities, including a combination of individual therapy, group therapy, 12-step meetings, case management, and recreational therapy. Participants in the suvorexant group will receive suvorexant and will be asked to avoid medications listed in the exclusion criteria, including sleep-related medications. Participants in the TAU group will continue their usual medication regimen as

XSL•FO
RenderX

prescribed by the inpatient facility staff. A comprehensive list of medications prescribed for each patient will be collected. If possible, statistical analyses will attempt to control for other medication exposure.

Sleep Assessment

Sleep will be monitored using an activity tracker watch (wGT3X-BT; ActiGraph, LLC) from night 0 to the morning of day 8. Data will be downloaded via the ActiLife software onto a UTHealth Houston computer and securely stored on IT-protected servers. Outcome measures derived from the actigraphy data will include total sleep duration and wake after sleep onset, two widely recognized indicators of sleep quality [12]. To supplement the actigraphy data, participants will complete the National Sleep Foundation Sleep Diary each night to confirm bedtimes and wake-up times for quality control purposes. Additional self-report sleep measures will be collected as described in subsequent sections.

Study Visit Measures

Self-Report and Behavioral Assessments

Sleep

The Pittsburgh Sleep Quality Index total summary score will be used to assess baseline sleep quality [13-15]. We will use the Insomnia Severity Index (ISI) [16] and Patient-Reported Outcomes Measurement Information System–Short Form (PROMIS-SF; Sleep Disturbance and Sleep-Related Impairment items) [17] as the main outcome variables. ISI and PROMIS-SF will be assessed on day 5 and day 12 at the laboratory visits to capture changes in sleep over the past 7 days.

Stress

Subjective stress levels will be assessed with the Depression, Anxiety, and Stress Scale-21 Items (DASS-21), a validated and reliable measure of core symptoms associated with depression, anxiety, and stress [18]. We will also use a visual analog scale (VAS; “Please rate your current stress level”: 0=no stress, 10=extreme stress) to assess current subjective stress levels [19].

Craving and Demand

The Stimulant Craving Questionnaire (STCQ) will be used to assess current craving [20]. The STCQ is a 10-item inventory that is useful in measuring craving across both stimulant types (cocaine and methamphetamine). The drug purchasing task will be used to assess demand for stimulants. During the hypothetical drug purchasing task, participants indicate how much of the drug they would use based on increasing prices [21-25]. Participants will complete a version of the drug purchasing task based on their primary stimulant use disorder (cocaine or methamphetamine).

Safety, Side Effects, and Other

A 31-item questionnaire will be administered to participants in the suvorexant group only to assess side effects and adverse events. The Columbia Suicide Severity Rating Scale [26] will be administered at each laboratory visit, and the Beck Depression Inventory [27] will also be included in the self-report battery. Additionally, at the first in-person visit, the Addiction Severity Index will also be administered to assess substance use severity and collect relevant demographic and addiction-related information [28].

Laboratory Assessments

Sleep

Resting-state EEG will be collected on laboratory days to assess the waking effects of sleep quality. The outcome variable will be alpha power (8 - 12 Hz), which reflects a state of wakeful rest [5,29]. Both eyes open and eyes closed conditions will be measured and analyzed separately. Resting alpha power will be collected at each laboratory visit day. Refer to the EEG Protocol section for details.

Stress

The socially evaluated cold pressor test will be used to measure stress reactivity. This task reliably activates the sympathetic nervous system and the hypothalamic-pituitary-adrenal (HPA) axis, as evidenced by increases in heart rate (HR), blood pressure (BP), and cortisol levels [30-33]. The social evaluative component is used to enhance HPA-axis activation. Participants will face a computer monitor showing their face using a webcam and a confederate wearing a white laboratory coat and clipboard will also be in the room. Participants will be informed that the camera is recording their facial reactions for later analysis and that the physician in the white laboratory coat will also be taking notes on their stress responses. No facial recordings or actual notes will be taken. During the cold pressor test (CPT), participants will submerge their arm in an ice-water bath. Physiological measures of stress reactivity will include HR, BP, and salivary cortisol, consistent with previous studies from our group [19,34]. Saliva samples will be collected before and after the CPT using the Cortisol-Salivette system (Sarstedt) and analyzed with the Cortisol ELISA Kit (Enzo Life Sciences), per manufacturer instructions. Subjective stress will be assessed before and after the CPT using a single-item VAS question, and HR and BP will also be recorded before, during, and after the CPT.

Reward Processing

EEG will also be recorded during the doors task to elicit the RewP [6]. In this task, participants are shown 2 doors, one of which has a hypothetical prize behind it. Participants guess which door hides the prize and then receive feedback indicating the outcome (up arrow=monetary win, down arrow=monetary loss). The RewP will be quantified as the mean amplitude in response to monetary feedback. Refer to the EEG Protocol section for details.

EEG Protocol

On laboratory days, participants will be seated comfortably in front of a monitor, fitted with the EEG cap, and the electrodes will be prepared with conductive gel. EEG will be collected using a 64-channel actiCAP electrode cap, amplified with BrainAmp MR (Brain Products GmbH) and digitized using BrainVision Recorder (Brain Products GmbH). The sampling rate will be 250 Hz, and data will be filtered with a 0.1 Hz high-pass filter and a 100 Hz low-pass filter.

XSL•FO
RenderX

Resting-state EEG will be collected prior to the doors task. We will collect resting-state EEG for 3 minutes with eyes open and eyes closed.

EEG data reduction will be performed with Brain Vision Analyzer 2 (Brain Products GmbH), MATLAB (MathWorks), and BESA (BESA GmbH). Resting-state data will be analyzed using the fast Fourier transform. The outcome of interest will be alpha power (8 - 12 Hz). The doors task data will be segmented (from 200 ms prestimulus to 800 ms poststimulus) by the following conditions: wins and losses. The RewP will be defined as the mean poststimulus amplitude over the frontal electrodes.

Dissemination Plan

The PI will be responsible for keeping the study record up to date on a yearly basis and for posting the results of the clinical trial on ClinicalTrials.gov. Findings will also be disseminated through conference presentations and publications in peer-reviewed scientific journals.

Statistical Analysis Plan

Analyses will be conducted within a generalized linear mixed modeling framework, which provides the flexibility to accommodate non-Gaussian outcomes via link functions and repeated or multilevel measurements via random effects. Traditional methods (eg, both 1-tailed and 2-tailed *t* tests and ANOVA) are subsumed as special cases, and the framework supports further extensions for nonlinear effects, time-varying predictors, and multivariate outcomes. For aims 1 (sleep) and 2 (all secondary measures), analyses will model each outcome variable as a function of the time and treatment interaction, controlling for main effects (fixed effects: outcome=$\beta_0 + \beta_1 \times$ treatment + $\beta_2 \times$ time + $\beta_3 \times$ treatment × time), with random effects to handle correlated observations. For aim 3 (treatment completion), analyses will model each outcome as a function of treatment.

All models will use distribution families and link functions as appropriate for each outcome (eg, dichotomous outcomes: binomial, with logit link). Longitudinal models evaluating change via interaction terms (eg, treatment × time) will be reduced to main effects if change is unsupported, and both models will be reported. Model fit information criteria (Akaike information criterion; leave-one-out information criterion) will be used to select structural components including multilevel (eg, random intercepts and slopes) and nonlinear terms. Analyses will be performed in R [35] via packages lme4 [36], rstan [37], and brms [38]. Potential confounders will be identified empirically [39,40] (ie, associated with both predictor and outcome) and modeled with and without adjustment; models with unchanged inference will favor parsimony. Sample characteristics (eg, sex, age, and education) will be evaluated as potential confounders. Analyses will be run in parallel using frequentist and Bayesian models. Frequentist results yield the probability of the data (or data more extreme) given the null hypothesis, whereas Bayesian results directly yield the probability of an alternative hypothesis [41,42]. Frequentist models will use α=.05 with false discovery rate correction for exploratory tests. Bayesian priors will be weakly informative (β~ N(0,1); sd~ Half-Cauchy(0,1); correlation matrices ~ LKJ(2)). Bayesian models will be evaluated via posterior probability (PP) threshold guidelines in the literature [43,44], suggesting that PP=75% to 90% indicates moderate evidence, PP=91% to 96% indicates strong evidence, and PP=97% or above indicates very strong to extreme evidence. Posterior predictive checking, scale reduction factors ("rhat"), and effective sample size will be used to evaluate Bayesian assumptions. Evaluation for frequentist models will use graphical evidence and statistical tests. Generalized linear mixed models inherently accommodate outcome values missing at random; other missing values will be handled by imputation or pattern-mixture modeling. Analyses will follow intention-to-treat principles.

The current proof-of-concept study is designed primarily to estimate a plausible range of treatment effects rather than to provide definitive hypothesis tests. To characterize expected inferential performance at the proposed sample size (n=40; 20 per arm), Monte Carlo simulation–based design analyses were conducted (1000 iterations). Data were generated under a linear mixed model with a treatment × time interaction, AR(1) correlated repeated measures (ρ=0.50), and the weakly informative priors specified earlier. At this sample size, assuming 10% attrition (n=36), simulations demonstrate at least an 80% probability (analogous to statistical power) of achieving moderate evidence (PP ≥75%) for standardized average treatment effects, unfolding in linear fashion over time, as small as Cohen *d*=0.26, and 80% probability of strong evidence (PP ≥91%) at Cohen *d*=0.39. Wide credible intervals are anticipated at this sample size, and results will be interpreted accordingly as preliminary effect size estimates to inform future definitive trials.

Ethical Considerations

The study protocol was approved by the Committee for the Protection of Human Subjects at UTHealth, Houston, Texas (HSC-MS-24 - 0330). Consent will be obtained virtually at the inpatient facility by a member of the CNRA team via voice or video call. Interested participants will e-sign the consent via REDCap. A copy of the signed consent can be printed or emailed to the participant. The PI will be responsible for ensuring compliance with the policies of the local institutional review board (the UTHealth Houston Committee for the Protection of Human Subjects). All participant data will be kept confidential. Data and safety oversight will include staff training, weekly audits of data collection and entry, medical screening with results reviewed by an on-site nurse or physician, use of standardized assessments, continued medical monitoring during the study, and procedures to monitor medication compliance. A data safety monitoring board will provide additional, independent oversight of data related to participant safety. All adverse events will be recorded for duration, severity, and frequency, and an evaluation will be made as to the event's relationship to the study drug, classified as unrelated, unlikely, possible, or probable. Participants will be compensated for their time as follows: US $5 for the screening and consent, US $40 for laboratory visits, US $10 for questionnaires, US $1 per day for medication adherence, and US $50 completion bonus.

XSL•FO
RenderX

Results

The study was funded in January 2025 and is currently ongoing. We began enrollment of the first participant prior to funding in July 2024. As of the submission of the manuscript (September 2025), a total of 8 participants had been enrolled in the study toward the planned enrollment of 40 participants. This study is expected to conclude in January 2027. No outcome data are currently available.

Discussion

Summary

Due to its mechanism of action, suvorexant may aid in preventing resumption of use in patients recently abstinent from stimulants via improvements in sleep, stress, and reward processing. The current open-label, randomized clinical trial will assess this hypothesis by randomizing patients with StimUD to receive either suvorexant or TAU for 7 nights at the local residential treatment facilities. We anticipate participants will demonstrate good adherence to suvorexant, and we will observe these improvements in both subjective and objective measures compared with the TAU group. We also anticipate that these mechanisms will be associated with increased treatment engagement demonstrating longer posttreatment lengths of stay and higher residential program completion rates.

Orexin receptor antagonism has shown promise in reducing stimulant-seeking behavior in preclinical studies, but there has been limited research in humans. This study builds upon the human literature that has examined the effects of suvorexant in StimUD, which has yielded mixed findings. Previously, in the preliminary proof-of-concept study, we found initial human evidence that dual orexin receptor antagonism with suvorexant improved multiple mechanistic processes in cocaine use disorder [4]. In contrast, in an acute cocaine self-administration study, Stoops et al [45] found that suvorexant increased the participants' motivation for cocaine rather than decreasing it. While the methods varied between these studies and could account for the differences in results, this study will add to the growing literature in the area.

There are several strengths and seminal innovations to the current proof-of-concept trial design, but it is not without limitations. This will be the first study to examine the effects of suvorexant during early withdrawal from stimulant use (both cocaine and methamphetamine) in treatment-seeking patients entering residential treatment for StimUD. Furthermore, the well-validated actigraphy-focused sleep assessment will contribute to our understanding of the mechanistic effects of suvorexant in this population. Another first will be leveraging EEG to assess the effects of suvorexant on the waking brain through both resting-state alpha power and the RewP. Finally, this trial will potentially provide effect size estimates of the efficacy of suvorexant regarding the treatment of StimUD, by assessing inpatient treatment completion rates. Despite these strengths, this study will be limited by the open-label design. This could result in the subjective outcomes (eg, sleep quality ratings, stress ratings, and craving scores) being vulnerable to placebo effects, expectancy bias, and demand characteristics. These outcomes will be interpreted cautiously in the statistical analysis. To overcome these concerns, we have included several validated objective outcomes (eg, actigraphy, EEG, and physiological measures), which are less susceptible to participant expectations and will be prioritized when drawing conclusions about treatment effects. Findings from this study will be exploratory and future blinded, placebo-controlled trial designs will establish definitive efficacy.

Conclusions and Future Directions

Notably, this trial was designed as a proof-of-concept study primarily concerned with deriving a preliminary range of plausible effect sizes for mechanistic effects and the effect on treatment completion. The study sample size was selected based on available funding and time to complete the study tasks. Therefore, the results of this study are meant to guide future fully powered trials rather than provide accurate efficacy estimates. A further limitation of this study is the lack of a placebo control and the open-label design. This design introduces potential biases into subjective measures due to participants' knowledge that they are receiving the treatment. As replication in a larger controlled trial is still needed, a negative finding would not rule out efficacy. Positive findings will also need to be confirmed in a larger sample. A TAU control condition was selected as the next best option and will provide information regarding how add-on suvorexant compares to supportive care and symptomatic medication alone. Critically, the use of several objective measures of main outcomes (actigraphy, HR, BP, cortisol, and EEG) will partially offset the influence of the open-label medication. This translational research approach will be the first mechanism-focused clinical investigation of suvorexant for the treatment of StimUD. FDA approval of suvorexant for the treatment of primary insomnia helps pave the way for rapidly advancing the repurposing of orexin-based therapies for drug addiction treatment.

Funding

This work was supported by the Brain and Behavior Research Foundation Young Investigator Grant (32442). HEW was also supported by NIDA K01DA058765.

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Valentino RJ, Volkow ND. Drugs, sleep, and the addicted brain. Neuropsychopharmacology 2020 Jan;45(1):3-5. [doi: 10.1038/s41386-019-0465-x] [Medline: 31311031]
2. Matzeu A, Martin-Fardon R. Understanding the role of orexin neuropeptides in drug addiction: preclinical studies and translational value. Front Behav Neurosci 2022;15:787595. [doi: 10.3389/fnbeh.2021.787595] [Medline: 35126069]
3. Roehrs TA, Roth T. Sleep disturbance in substance use disorders. Psychiatr Clin North Am 2015 Dec;38(4):793-803. [doi: 10.1016/j.psc.2015.07.008] [Medline: 26600109]
4. Suchting R, Yoon JH, Miguel GGS, et al. Preliminary examination of the orexin system on relapse-related factors in cocaine use disorder. Brain Res 2020 Mar 15;1731:146359. [doi: 10.1016/j.brainres.2019.146359] [Medline: 31374218]
5. Knyazev GG, Slobodskoj-Plusnin JY, Bocharov AV, Pylkova LV. The default mode network and EEG α oscillations: an independent component analysis. Brain Res 2011 Jul 21;1402:67-79. [doi: 10.1016/j.brainres.2011.05.052] [Medline: 21683942]
6. Proudfit GH. The reward positivity: from basic research on reward to a biomarker for depression. Psychophysiology 2015 Apr;52(4):449-459. [doi: 10.1111/psyp.12370] [Medline: 25327938]
7. Mancino MJ, Gentry BW, Feldman Z, Mendelson J, Oliveto A. Characterizing methamphetamine withdrawal in recently abstinent methamphetamine users: a pilot field study. Am J Drug Alcohol Abuse 2011 Mar;37(2):131-136. [doi: 10.3109/00952990.2010.543998] [Medline: 21219261]
8. First MB, Williams JB, Karg RS, Spitzer RL. Structured Clinical Interview for DSM-5 Disorders - Research Version: American Psychiatric Association; 2015.
9. Snow G. blockrand: randomization for block randomized clinical trials. The Comprehensive R Archive Network. 2020. URL: https://cran.r-project.org/web/packages/blockrand/blockrand.pdf [accessed 2026-03-17]
10. Webber H, Badawi J, Calvillo D, et al. Preliminary examination of orexin receptor antagonism with suvorexant in individuals with methamphetamine use disorder: a case series study. J Addict Dis 2026;44(1):88-95. [doi: 10.1080/10550887.2024.2434298] [Medline: 39731508]
11. Bennett T, Bray D, Neville MW. Suvorexant, a dual orexin receptor antagonist for the management of insomnia. P T 2014 Apr;39(4):264-266. [Medline: 24757363]
12. Ohayon M, Wickwire EM, Hirshkowitz M, et al. National Sleep Foundation's sleep quality recommendations: first report. Sleep Health 2017 Feb;3(1):6-19. [doi: 10.1016/j.sleh.2016.11.006] [Medline: 28346153]
13. Tang J, Liao Y, He H, et al. Sleeping problems in Chinese illicit drug dependent subjects. BMC Psychiatry 2015 Feb 19;15:28. [doi: 10.1186/s12888-015-0409-x] [Medline: 25884573]
14. Mahfoud Y, Talih F, Streem D, Budur K. Sleep disorders in substance abusers: how common are they? Psychiatry (Edgmont) 2009 Sep;6(9):38-42. [Medline: 19855859]
15. Brooks AT, Krumlauf MC, Whiting BP, Clark RJ, Wallen GR. Are you sleeping? Pilot comparison of self-reported and objective measures of sleep quality and duration in an inpatient alcoholism treatment program. Subst Abuse 2012;6:135-139. [doi: 10.4137/SART.S10385] [Medline: 23115479]
16. Bastien CH, Vallières A, Morin CM. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001 Jul;2(4):297-307. [doi: 10.1016/s1389-9457(00)00065-4] [Medline: 11438246]
17. Yu L, Buysse DJ, Germain A, et al. Development of short forms from the PROMIS™ sleep disturbance and Sleep-Related Impairment item banks. Behav Sleep Med 2011 Dec 28;10(1):6-24. [doi: 10.1080/15402002.2012.636266] [Medline: 22250775]
18. Antony MM, Bieling PJ, Cox BJ, Enns MW, Swinson RP. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. Psychol Assess 1998;10(2):176-181. [doi: 10.1037/1040-3590.10.2.176]
19. Vujanovic AA, Wardle MC, Liu S, Dias NR, Lane SD. Attentional bias in adults with cannabis use disorders. J Addict Dis 2016;35(2):144-153. [doi: 10.1080/10550887.2015.1116354] [Medline: 26566718]
20. Northrup TF, Green C, Walker R, Greer TL, Trivedi MH. On the invariance of the Stimulant Craving Questionnaire (STCQ) across cocaine and methamphetamine users. Addict Behav 2015 Mar;42:144-147. [doi: 10.1016/j.addbeh.2014.11.020] [Medline: 25462663]
21. MacKillop J, Goldenson NI, Kirkpatrick MG, Leventhal AM. Validation of a behavioral economic purchase task for assessing drug abuse liability. Addict Biol 2019 Mar;24(2):303-314. [doi: 10.1111/adb.12592] [Medline: 29314376]
22. Webber HE, Yoon JH, de Dios C, et al. Assessing cocaine motivational value: comparison of brain reactivity bias toward cocaine cues and cocaine demand. Exp Clin Psychopharmacol 2023 Aug;31(4):861-867. [doi: 10.1037/pha0000622] [Medline: 36480395]
23. Yoon JH, Suchting R, de Dios C, et al. Decreased cocaine demand following contingency management treatment. Drug Alcohol Depend 2021 Sep 1;226:108883. [doi: 10.1016/j.drugalcdep.2021.108883] [Medline: 34198136]

XSL•FO
RenderX

24. Yoon JH, Suchting R, McKay SA, et al. Baseline cocaine demand predicts contingency management treatment outcomes for cocaine-use disorder. Psychol Addict Behav 2020 Feb;34(1):164-174. [doi: 10.1037/adb0000475] [Medline: 31233323]
25. Yoon JH, Suchting R, Cassidy RN, et al. Assessment of demand for methamphetamine and cigarettes among individuals with methamphetamine use disorder. Exp Clin Psychopharmacol 2021 Aug;29(4):334-344. [doi: 10.1037/pha0000367] [Medline: 32202813]
26. Posner K, Brown GK, Stanley B, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011 Dec;168(12):1266-1277. [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
27. Beck AT, Steer RA, Brown GK. Beck Depression Inventory, 2nd edition: Guilford Press; 1996.
28. McLellan AT, Kushner H, Metzger D, et al. The fifth edition of the Addiction Severity Index. J Subst Abuse Treat 1992;9(3):199-213. [doi: 10.1016/0740-5472(92)90062-s] [Medline: 1334156]
29. Jann K, Dierks T, Boesch C, Kottlow M, Strik W, Koenig T. BOLD correlates of EEG alpha phase-locking and the fMRI default mode network. Neuroimage 2009 Apr 15;45(3):903-916. [doi: 10.1016/j.neuroimage.2009.01.001] [Medline: 19280706]
30. Velasco M, Gómez J, Blanco M, Rodriguez I. The cold pressor test: pharmacological and therapeutic aspects. Am J Ther 1997 Jan;4(1):34-38. [doi: 10.1097/00045391-199701000-00008] [Medline: 10423589]
31. McRae AL, Saladin ME, Brady KT, Upadhyaya H, Back SE, Timmerman MA. Stress reactivity: biological and subjective responses to the cold pressor and Trier Social stressors. Hum Psychopharmacol 2006 Aug;21(6):377-385. [doi: 10.1002/hup.778] [Medline: 16915579]
32. Lovallo W. The cold pressor test and autonomic function: a review and integration. Psychophysiology 1975 May;12(3):268-282. [doi: 10.1111/j.1469-8986.1975.tb01289.x] [Medline: 1153632]
33. Geliebter A, Gibson CD, Hernandez DB, et al. Plasma cortisol levels in response to a cold pressor test did not predict appetite or ad libitum test meal intake in obese women. Appetite 2012 Dec;59(3):956-959. [doi: 10.1016/j.appet.2012.08.025] [Medline: 22983369]
34. Dias NR, Schmitz JM, Rathnayaka N, et al. Anti-saccade error rates as a measure of attentional bias in cocaine dependent subjects. Behav Brain Res 2015 Oct 1;292:493-499. [doi: 10.1016/j.bbr.2015.07.006] [Medline: 26164486]
35. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2024. URL: https://www.R-project.org/ [accessed 2026-03-17]
36. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. arXiv. Preprint posted online on Jun 23, 2014. [doi: 10.48550/arXiv.1406.5823]
37. Stan Development Team. RStan: the R interface to Stan. The Comprehensive R Archive Network. 2025. URL: https://cran.r-project.org/web/packages/rstan/vignettes/rstan.html [accessed 2026-03-17]
38. Bürkner PC. brms: an R package for Bayesian multilevel models using Stan. J Stat Softw 2017;80(1):1-28. [doi: 10.18637/jss.v080.i01]
39. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. Lancet 2000 Mar 25;355(9209):1064-1069. [doi: 10.1016/S0140-6736(00)02039-0] [Medline: 10744093]
40. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med 2002 Oct 15;21(19):2917-2930. [doi: 10.1002/sim.1296] [Medline: 12325108]
41. McElreath R. Statistical Rethinking: A Bayesian Course with Examples in R and Stan: CRC Press; 2016.
42. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis, 3rd edition: CRC Press; 2013.
43. Jeffreys H. The Theory of Probability, 3rd edition: Oxford University Press; 1998.
44. Lee MD, Wagenmakers EJ. Bayesian Cognitive Modeling: A Practical Course: Cambridge University Press; 2013.
45. Stoops WW, Strickland JC, Hatton KW, et al. Suvorexant maintenance enhances the reinforcing but not subjective and physiological effects of intravenous cocaine in humans. Pharmacol Biochem Behav 2022 Oct;220:173466. [doi: 10.1016/j.pbb.2022.173466] [Medline: 36152876]

Abbreviations

BP : blood pressure
CNRA: Center for Neurobehavioral Research on Addiction
CPT: cold pressor test
DASS-21: Depression, Anxiety, and Stress Scale-21 Items
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
EEG: electroencephalography
FDA: Food and Drug Administration
HPA: hypothalamic-pituitary-adrenal
HR: heart rate
ISI : Insomnia Severity Index

XSL•FO
RenderX

PI: principal investigator
PP: posterior probability
PROMIS-SF: Patient-Reported Outcomes Measurement Information System–Short Form
REDCap: Research Electronic Data Capture
RewP: reward positivity
STCQ : Stimulant Craving Questionnaire
StimUD: stimulant use disorder
TAU : treatment as usual
UTHealth Houston: University of Texas Health Science Center at Houston
VAS: visual analog scale

Edited by J Sarvestan; submitted 09.Sep.2025; peer-reviewed by CS Freet; revised version received 10.Feb.2026; accepted 18.Feb.2026; published 08.Apr.2026.

Please cite as:
Webber H, Tran N, Blake R, Goldman S, George T, Bailey B, Suchting R, Yoon J, Vincent J, Weaver M, Schmitz J, Lane S
The Effects of Orexin Receptor Antagonism During Early Withdrawal in Stimulant Use Disorder: Protocol for a Proof-of-Concept Study
JMIR Res Protoc 2026;15:e83842
URL: https://www.researchprotocols.org/2026/1/e83842
doi:10.2196/83842

© Heather Webber, Nikky Tran, Rita Blake, Sarah Goldman, Trevor George, Bill Bailey, Robert Suchting, Jin Yoon, Jessica Vincent, Michael Weaver, Joy Schmitz, Scott Lane. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Spinal Manipulative Therapy on a Healthy Population: Protocol for a Randomized Blinding Feasibility Trial

Margaret Sliwka*, DC; Tyson Perez*, DC, PhD; Ahmed Qazi*, MS, PhD; Stephanie Sullivan, DC, PhD

Dr. Sid E. Williams Center for Chiropractic Research, Life University, 1429 Lucile Ave, Building 900, Ste 910, Marietta, GA, United States

*these authors contributed equally

Corresponding Author:
Margaret Sliwka, DC
Dr. Sid E. Williams Center for Chiropractic Research, Life University, 1429 Lucile Ave, Building 900, Ste 910, Marietta, GA, United States

Abstract

Background: Few manual chiropractic high velocity, low amplitude (HVLA) type shams have been validated in research. The proposed project is a randomized controlled trial (RCT) designed to assess a novel, full-spine, manual sham chiropractic maneuver and its blinding success.

Objective: This study aims to evaluate the blinding integrity of participants receiving a genuine or sham chiropractic maneuver. We will also be evaluating the effects of genuine chiropractic treatments relative to sham chiropractic treatments by measuring several neurophysiological mechanisms.

Methods: Participants (N=60) will be recruited from in and around Marietta, Georgia, United States. They will undergo a chiropractic physical exam and health history review with a licensed chiropractor and be randomized to either a sham or genuine chiropractic group (1:1 ratio). Participants, outcome assessors, and data analysts will be blinded to group allocation. The genuine group will receive diversified HVLA chiropractic spinal manipulative therapy (SMT), while the sham group will receive a novel chiropractic HVLA-emulating therapy. Each participant will attend 2 sessions spaced 1 week apart. Assessments will consist of blinding surveys postsessions (both visits) and presession (second visit). Additionally, we will measure gait parameters. The primary outcome of interest is blinding measured via the Bang Blinding Index (Bang BI). Secondary aims include measuring blinding via the James Blinding Index (James BI) and exploring the potential differential effects of genuine versus sham SMT on gait parameters.

Results: This study is being supported and internally funded by the Dr. Sid E. Williams Center for Chiropractic Research at Life University in Marietta, Georgia, United States. Our study was prospectively registered on clinicaltrials.gov (NCT06931600) on March 28, 2025, and the first participant was enrolled on May 5, 2025. As of December 10, 2025, 18 participants have completed the trial. Projections for completing data collection, data analysis, and manuscript submission are Summer 2026, Fall 2026, and Winter 2027, respectively.

Conclusions: The significance of the current RCT will be in its ability to inform whether our novel, full-spine, manual sham SMT protocol successfully blinds participants, suggesting feasibility for future clinical trials, as well as assessing for secondary outcome measures between groups.

Trial Registration: ClinicalTrials.gov NCT06931600; https://clinicaltrials.gov/study/NCT06931600 and Universal Trial Number (UTN) U1111-1321-6250; https://trialsearch.who.int/Trial2.aspx?TrialID=NCT06931600

International Registered Report Identifier (IRRID): DERR1-10.2196/85956

(JMIR Res Protoc 2026;15:e85956) doi:10.2196/85956

KEYWORDS

double-blind; chiropractic; manual therapy; randomized controlled trial; gait analyses

Introduction

Randomized controlled trials (RCTs) are widely regarded as essential in evidence-based health care. To accurately evaluate treatment effects, RCTs often rely on sham control interventions. These controls are designed to mimic the nonspecific aspects of treatment, such as provider attention, touch, and therapeutic interaction, without delivering the active therapeutic component [1]. By using shams, researchers are better able to distinguish between the specific effects of an intervention versus those changes resulting from other factors relating to psychological response or contextual markers. When it comes to spinal manipulative therapies (SMTs), the treatment itself cannot be

XSL•FO
RenderX

easily masked, and sham interventions can therefore play a crucial role in minimizing bias.

Developing credible sham interventions for SMTs is a challenging endeavor. SMTs and other manual therapies usually involve direct physical contact, as well as distinct movements or methodologies, making it difficult to design a control that is both convincing to participants and yet lacking any kind of specific therapeutic effect. Therapist touch, timing, and intention are hard to simulate without unintentionally introducing some active input [2,3]. Additionally, the therapeutic environment, which involves verbal cues, confidence of the provider, and ritual, has previously been shown to produce clinical outcomes that are independent of the actual physical intervention [4,5]. These complexities make it difficult to create a standardized sham SMT protocol and can distort the interpretation of treatment effectiveness.

In chiropractic research, several sham SMT procedures have been tested. These range from superficial contact without thrust to zeroed-out mechanical devices or maneuvers mimicking SMT (also known as a chiropractic "adjustment") but applied in nontherapeutic directions [6,7]. Despite these efforts, systematic reviews have found significant variation in how convincing and inert these controls are [8-10]. Many trials also fail to report on or verify whether participants were truly blinded and unaware of their intervention assignment, undermining the strength of the study findings [11,12]. Consequently, there remains a need for standardized, validated sham SMT interventions that are both perceptually credible and limiting in physiological input for the participants.

Two widely referenced tools in manual therapy studies are the Bang Blinding Index (Bang BI) and the James Blinding Index (James BI), which offer distinct approaches to accurately measuring blinding [13,14]. The Bang BI examines blinding separately for each treatment group and evaluates both the degree and direction of potential bias. In contrast, the James BI provides an overall interpretation of blinding by comparing correct guesses to what would be expected by chance, where "don't know" responses are regarded as optimal [13,14]. As such, the Bang BI helps with the identification of group-specific bias patterns, whereas the James BI provides a more overall view of blinding success [13-15]. When used together, these 2 indices can strengthen confidence in a study's internal validity by offering complementary interpretations of the results. Recent studies have evaluated the use of these 2 indices and recognized the feasibility, utility, and importance of them in manual therapy trials [15,16]. These studies also highlighted the importance of testing blinding in manual therapy research, especially in chiropractic trials where tactile and treatment context cues may influence participants' expectations. Blinding efficacy is fundamental to maintaining the validity of clinical trials, especially when comparing the differences between sham and genuine SMTs.

In addition to blinding efficacy, we are also interested in exploring the possible neurophysiological mechanisms potentially responsive to change following chiropractic care. It is therefore relevant to quantify and assess movement-related measures such as gait parameters to elucidate the definitive effects of such therapeutic treatments. Contemporary evidence involving genuine SMTs has provided evidence to suggest such therapies result in improvements in several gait parameters, including step length and gait symmetry [17]. However, there is limited data comparing gait outcomes between genuine SMTs and shams. Therefore, we aim to expand the literature on the potential impacts of SMT on movement-related outcomes within the context of a putatively robust study design.

In this RCT, we aim to evaluate several things. Our primary outcome measure will consist of evaluating the blinding integrity of participants (as measured via the Bang BI). Our secondary outcome measures will be to evaluate the blinding integrity of participants (as measured by James BI), as well as to investigate the potential effects of genuine SMT relative to sham SMT via gait kinematics and biomechanics.

Methods

Trial Design

This study is a randomized, blinded (participants, outcome assessors, and data analysts), sham-controlled, parallel-group (2-arm) trial with an allocation ratio of 1:1.

Study Population and Setting

Our target population is healthy adults living in and around Marietta, Georgia, United States. This study will take place at Life University's Dr. Sid E Williams Center for Chiropractic Research (CCR) located in Marietta.

Eligibility Criteria

The detailed eligibility criteria are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion requirements for participant selection in this study.

Eligibility criteria were defined to specify the inclusion and exclusion requirements for participant selection in this study:

- 18 - 60 years of age
- No history of stroke or transient ischemic attack or the symptoms, including:
- Dizziness or vertigo
- Tinnitus (ringing in the ears)
- Visual, sensory, or motor disturbances
- No new pattern headache complaint
- No recent whiplash injury (within 3 months)
- No spinal fractures or dislocations
- No disc problems with radiating symptoms to the arms or legs
- No severe degenerative joint disease in the spine
- No connective tissue disorders
- No primary fibromyalgia
- No metabolic or metaplastic bone disease
- No history of cervical, thoracic, or lumbar spine surgery
- No uncontrolled high blood pressure or vascular disease
- No current use of anticoagulant therapy
- No diagnosed condition that causes fainting during postural changes, such as Postural Orthostatic Tachycardia Syndrome (POTS) or orthostatic hypotension
- No current use of short-acting benzodiazepines, including midazolam and triazolam
- No change in medications in the past 6 weeks or intentions to change medications during the study
- Inability to walk unassisted on a treadmill
- No pacemakers or known heart conditions that influence the electrical or mechanical function of the heart, such as severe heart valve disease
- Not a doctor of chiropractic or a doctor of chiropractic student enrolled in the 4th quarter or above
- No present or self-reported pregnancy
- No chiropractic care in the 2 weeks prior to participation
- Did not participate in our prior feasibility study (Perceptions and experiences following a single session of simulated or genuine high velocity, low amplitude [HVLA] manual chiropractic adjustments)

Explanation for the Choice of Comparators

We have developed a manual, full-spine, sham SMT for use in future RCTs and are endeavoring to test its credibility. A credible sham control will allow SMT researchers to elucidate any potential specific effects and help address concerns of generally weak methodological designs in SMT-related trials [18,19].

Recruitment

Participants will be recruited via targeted advertising (eg, social media advertisements, emails, and flyers) with an invitation to participate in a Life University trial involving SMT. Recruitment materials will briefly describe the study and direct potential participants to a short online screening form via a link or QR code. Individuals who complete the online screen and meet the eligibility requirements will be contacted via email and/or phone and asked to attend an in-person session at Life University's Dr. Sid E Williams CCR. Recruitment will continue until our target sample size is met and is expected to take up to 12 months. To help foster recruitment efforts, participants will receive up to US $75 in gift cards total following their participation as reimbursement for their time.

Participant Timeline

This trial will require 2 sessions, one week apart. The participants will receive the same intervention (ie, genuine or sham SMT) at both sessions. The study flow is outlined in Figure 1.

XSL•FO
RenderX

Figure 1. Study flow chart diagram. SMT: Spinal Manipulative Therapy.

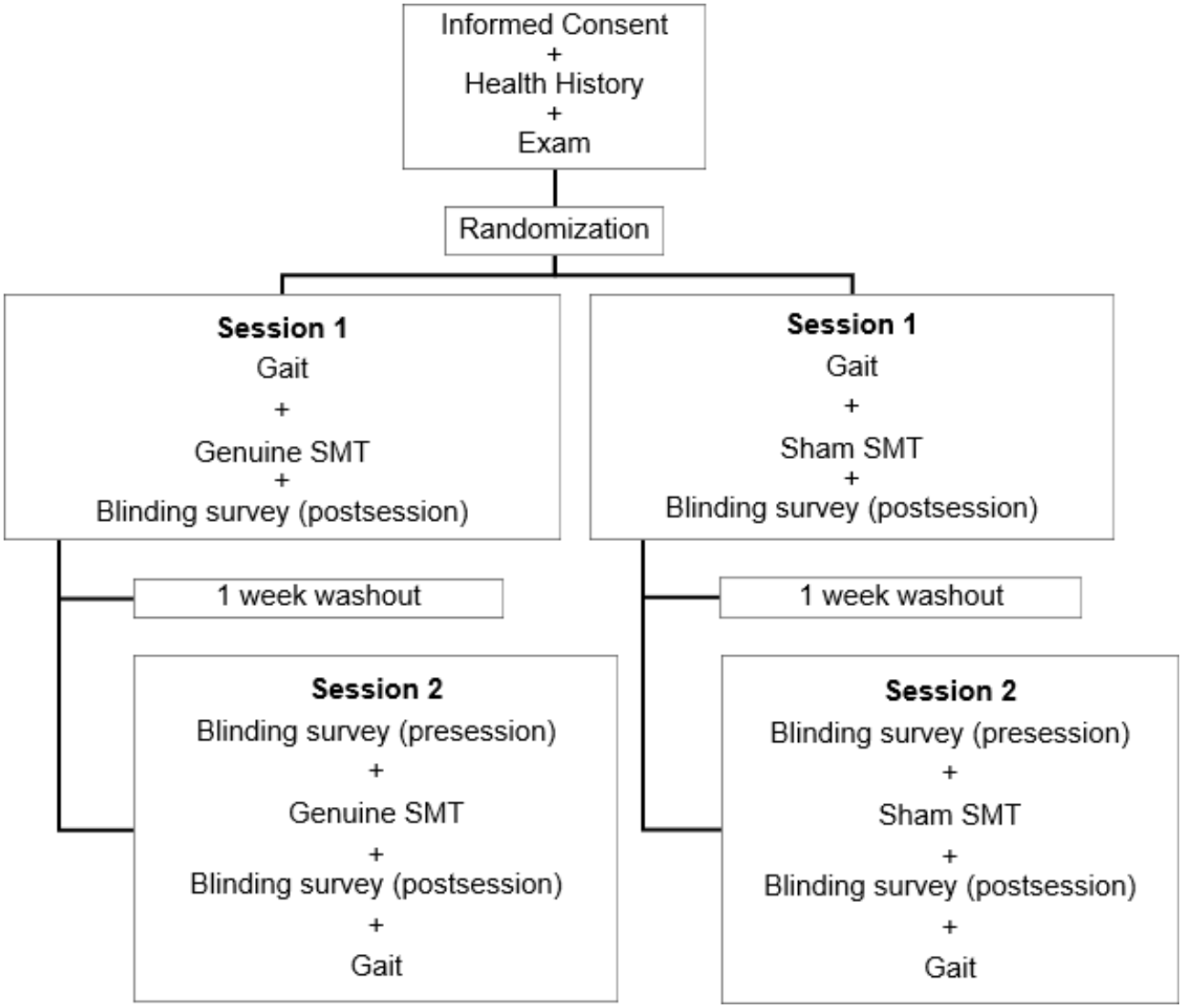

Demographic Information, Health History, and Anthropometric Measurements

Participants will be given a quiet room and asked to provide their demographic information and health history on an electronic tablet. Participants will then have their anthropometric measurements (ie, height and weight) recorded.

Chiropractic Exam

During their initial session, participants will be escorted to an exam room and introduced to their attending chiropractor (MS or AJ). MS or AJ will discuss the participant's health history. Orthopedic and/or neurological assessments may be performed to rule out red flags. If during the physical exam MS or AJ feels X-rays are indicated, finds contraindications to chiropractic care (ie, red flags), or fails to locate any spinal segmental dysfunctions, the candidate will be excluded from study participation, receive compensation, and may be referred to an appropriate medical provider. If no spinal segmental dysfunctions are found in one or 2 spinal regions but there is at least one spinal region with a segmental dysfunction present, the participant will be allowed to proceed with the study, and only the spinal segments found will be treated accordingly.

To determine the presence or absence of a spinal segmental dysfunction, MS will assess each spinal region in the following ways:

- Cervical spine (C1): the patient will be seated at the end of the table with their arms relaxed at their sides. The examiner will use their left hand to hold the patient's head to direct their motion, while their right hand will be performing the following motions: rotation and extension (the examiner contacts the posterior aspect of the C1 transverse process with their right hand and moves the patient through extension and right and left rotation, only once on each side). Coupled motion of C1 is to be used instead of flexion and extension, since the lack of a C1 spinous process makes a flexion and extension assessment of C1 movement challenging. Restriction is indicated at the end of the coupled rotation and extension on the side where there is a "blocked" end feel to the movement (the joint does not glide as far as on the other side).
- Cervical spine (C2-C7): the patient will be seated at the end of the table with their arms relaxed at their sides. The examiner will use their left hand to hold the patient's head to direct their motion, while their right hand will be performing the following motions: rotation and extension

XSL•FO
RenderX

(the examiner contacts the articular pillars bilaterally with the index finger and thumb of their right hand and moves the patient through left and right rotation and extension, feeling for motion of the spinal level). Spinal levels C2 through C7 will be individually assessed and motioned only once on each side. Restriction is indicated when the patient is in rotation and extension on either the left, right, or both sides and there is a "blocked" end feel to the movement (the joint does not glide as far as at other levels).

- Thoracic and lumbar spine: the patient will be seated at the end of the table with their arms relaxed at their sides. The examiner will be stabilizing their right elbow on the patient's right shoulder to direct their motion, while their left thumb will be pushing along the right lateral aspect of each spinous process to ensure the following motions: extension, rotation, and lateral flexion (as the examiner pushes against the right lateral aspect of the spinous process, they are using the patient's right shoulder as a lever to assist in the correct movement of the thoracolumbar spine). The examiner will repeat this for the opposite side (standing on the patient's left, using the left elbow to direct motion, and using the right thumb to palpate the left lateral spinous processes). Restriction is indicated when the patient is in extension, rotation, and lateral flexion on either the left, right, or both sides and there is a "blocked" end feel to the movement (the joint does not glide as far as at other levels).
- Pelvis (Thompson sacral check): the participant will be lying prone on the table with their arms relaxed to the sides. The examiner will stabilize the base of the participant's sacrum with the base of their palm and ask the participant to raise one leg straight up from the hip as high as they can without bending the knee. After repeating this procedure for the opposite leg, the sacral listing is determined via the Thompson chiropractic protocol, whereby the leg that raises higher is the side of anterior sacral displacement and has a listing of "SAR" if that is the right leg or "SAL" if that is the left leg.
- Pelvis (prone leg check): the participant will be lying prone on the table with their arms relaxed to the sides. The examiner will assess the leg length of the participant in this position. If there is a shorter leg, a listing of the posterior inferior ilium on that side will be interpreted.
- Tenderness to palpation: as the examiner is motioning the patient, they will simultaneously note if there is any discomfort at the levels where motion restriction is found, as this is further indication of a spinal segmental dysfunction [20].
- Soft tissue markers: in addition to noting any soft tissue changes while motioning the patient, the examiner will instruct the patient to lie prone on a flat chiropractic bench and then palpate for soft tissue markers. These soft tissue markers, such as localized edema or taut muscle fibers, will be assessed in the cervical, thoracic, and lumbar spine by gently palpating lateral to the spinous processes in the paraspinal musculature in each region. The sacroiliac joints will be gently palpated in this position as well. The suboccipital muscles will also be palpated along the occipital ridge bilaterally. These soft tissue changes can also indicate a potential spinal segmental dysfunction [20].

Once the determination of a dysfunctional segment in the cervical, thoracic, and lumbopelvic regions has been made, the clinician will notate said listings in the participant file.

Randomization

Using simple randomization via a box draw, participants will be randomized to one of two arms: (1) genuine high velocity, low amplitude (HVLA) cervical, thoracic, and/or lumbopelvic SMT, or (2) sham HVLA cervical, thoracic, and/or lumbopelvic SMT.

Sequence Generation

Prior to participant recruitment, 60 blank, opaque, unidentifiable, sealed envelopes will be prepared with an equal distribution of genuine versus sham chiropractic SMT by a lab staff member with no direct contact with participants.

Concealment Mechanism

To ensure concealment, the envelope allotments for randomization will be kept in the central office in a box that only the clinicians (MS and AJ) will be able to access.

Implementation (Enrollment and Assignment)

The clinicians (MS and AJ) are responsible for participant enrollment. Following the initial chiropractic exam and immediately prior to the delivery of the participant's first sham or real SMT, the participant will be permitted to choose their own envelope from a box, though MS or AJ will be the one to open the envelope to review the group assignment. This assignment will be recorded by MS or AJ in a binder that will only be accessible to her throughout the course of the trial.

Who Will Be Blinded?

Participants, outcome assessors, and data analysts will be unaware of group assignments. The clinicians (MS and AJ) will not be blinded. To improve participant blinding, we will endeavor to make all aspects of the sham sessions identical to the genuine sessions, including the audible drop of the chiropractic table in the cervical, thoracic, and lumbopelvic regions.

Unblinding Procedure

Group assignments will be disclosed to trial participants 30 days after the completion of the study.

Assessments

Participants will be asked to attend 2 separate sessions approximately one week apart. Automated text reminders will also be sent via ChiroHD software (ChiroHD Inc) ~24 hours and~3 hours prior to each session.

To assess gait, participants will be asked to walk on a motorized treadmill at their preferred walking speed (PWS). Participants will be fitted with 16 inertial measurement units at different anatomical landmarks. Inertial measurement units will then be calibrated using commercial motion-capture equipment (NORAXON MyoMotion, MR2; Noraxon Inc). Participants will be fitted with a pair of slip-resistant socks prior to testing. They will then be instructed to step onto the nonmoving treadmill to begin the PWS assessment. Participants will start by walking on the treadmill at a speed of 0.5 mph. A trained

XSL•FO
RenderX

research investigator will then increase the speed by 0.1 mph increments every 2 - 4 seconds until the participant reports that they are walking at their PWS. Then, in 0.1 mph increments, 1 mph is added to the speed, followed by a decrease at a rate of 0.1 mph every 2 - 4 sec until they re-establish their PWS. This procedure will be repeated 3 times in total. The speeds will be averaged to determine their PWS. The PWS assessment will only be performed at baseline. Participants will also be blinded to the digital speed display while walking on the treadmill. Once the PWS has been established, participants will be asked to walk on the treadmill for 5 minutes at their PWS. The treadmill will be used to measure forces during walking, as well as several traditional gait parameters (eg, step time, stride time, and steps per minute in milliseconds). Gait assessment during the follow-up session will be identical, except for determining PWS.

Interventions

The clinicians (MS and AJ) are licensed chiropractors with >20 years of combined experience in providing HVLA treatments. MS was also involved in the creation of the sham protocol. Both clinicians undertook extensive training in its delivery over a ~3 month period.

Genuine Chiropractic Treatment

The genuine cervical intervention will be performed in the predetermined location using a supine position and a diversified maneuver (see Figure 2). The thoracic spine will be treated in a prone position using a double hypothenar thrust maneuver (see Figure 3). This will be done with or without an additional drop in the thoracic region at the discretion of the chiropractor and is based on muscle tone in the region. The lumbopelvic region will receive either a prone drop maneuver with 3 - 4 drops or a side posture push, depending on clinical findings (see Figure 4). The spinal levels and listings to be treated will be based on examination findings. Only one SMT will be performed in each of the 3 regions assessed (cervical, thoracic, and lumbopelvic). Following the genuine chiropractic treatment, the spine will be reassessed for intersegmental motion restrictions in a postcheck palpation of the spine of the participant. It will be noted in the session notes if the pre- and posttreatment objective findings are the same or different. The levels to be treated may be different at each of the 2 sessions, as determination is based on the palpatory findings at each visit.

XSL•FO
RenderX

Figure 2. Genuine cervical spinal manipulative therapy (supine cervical set).



XSL•FO
RenderX

Figure 3. Genuine thoracic spinal manipulative therapy (double hypothenar with drop).



XSL•FO
RenderX

Figure 4. Genuine lumbopelvic spinal manipulative therapy (prone drop or side posture push).



Sham Chiropractic Treatment

For the sham cervical SMT, the participant will be placed in a supine position on a cervical drop piece with the chiropractor's forearm under the participant's neck and contact lateral to the cervical spine (see Figure 5). This will only be performed unilaterally over the side with the predetermined restriction in the articular pillars. The thoracic sham (see Figure 6) and the lumbopelvic sham (see Figure 7) will be performed with the participant in the prone position. Contact for the thoracic sham will be medial to the medial scapular border, and contact for the lumbopelvic sham will be over the lateral gluteal muscles, opposite to the side of the chiropractor's stance. The chiropractor will contact close to the predetermined site of intersegmental motion restriction in the spine, but no manual thrust will be directed into the spine. Instead, the chiropractor will apply a downward pressure to the drop piece with their forearm (cervical) or knee (thoracic and lumbopelvic), which will elicit a noise from the table that mimics a genuine drop SMT, as well as induced movement of the participant's body. Only a single drop will be made in all regions to limit any potential joint motion with the procedure. Following the sham chiropractic treatment, the spine will be reassessed for intersegmental motion restrictions and objective findings in a postcheck, allowing for the comparison with pretreatment findings. Similar to the genuine treatment, the sham will be performed at levels determined at that particular session.

XSL•FO
RenderX

Figure 5. Sham cervical spinal manipulative therapy (supine drop with forearm assist from 2 angles).

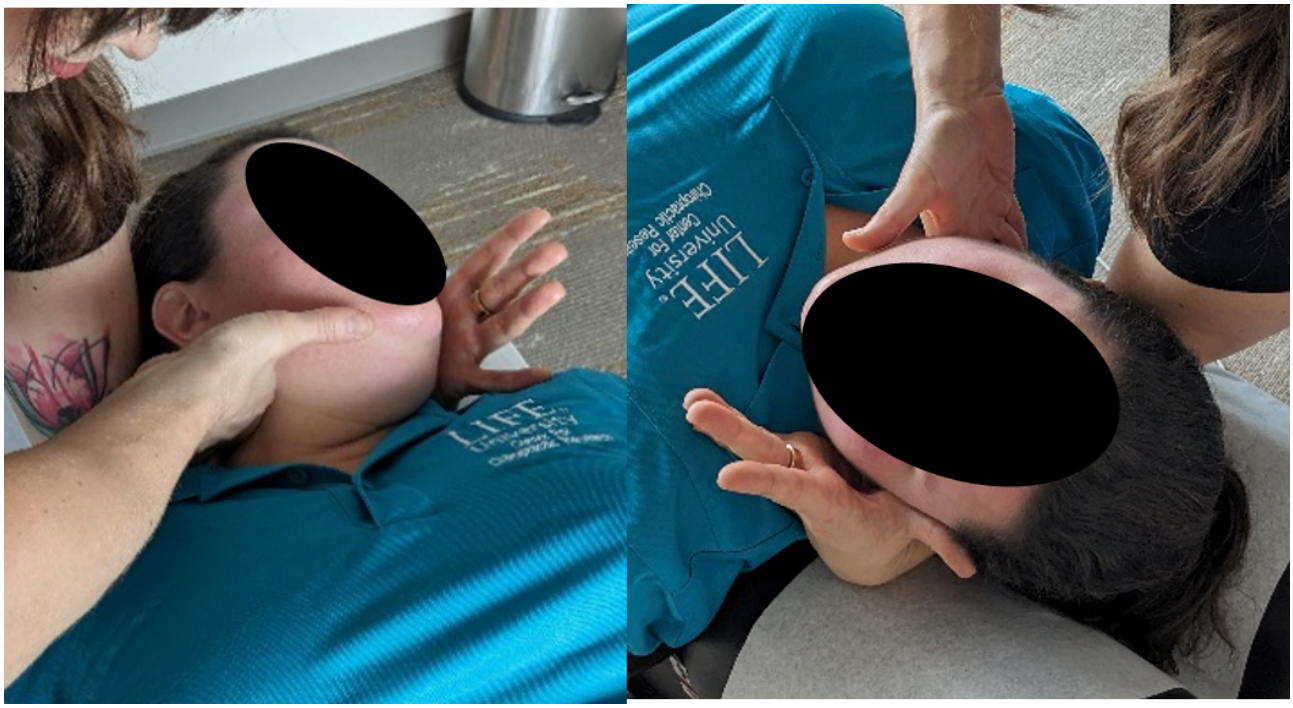

XSL•FO
RenderX

Figure 6. Sham thoracic spinal manipulative therapy (prone drop with knee assist).



XSL•FO
RenderX

Figure 7. Sham lumbopelvic spinal manipulative therapy (prone drop with knee assist).



Table 1 provides a detailed description of the trial interventions using the Template for Intervention Description and Replication (TIDieR) [21]. Immediately following each session, as well as at the beginning of the second session, blinding will be assessed via a 4-item electronic survey taken with the computer program Jotform. Additionally, within 30 days of the trial's completion, participants will be informed of their group allocation via email.

XSL•FO
RenderX

Table . Template for intervention description and replication (TIDieR).

Item	Description
Brief name: provide the name or a phrase that describes the intervention.	Drop table-assisted genuine or sham spinal manipulative therapies (SMTs)
Why: describe any rationale, theory, or goal of the elements essential to the intervention.	Using shams in manual therapy (MT) studies allows for comparison of neurological measures (ie, gait parameters) to expose any definitive effects of therapeutic treatments. Blinding efficacy is fundamental to maintaining the validity of said shams. To date, there are limited manual high velocity, low amplitude (HVLA)–mimicking shams that have been validated in MT research.
What (materials): Describe any physical or informational materials used in the intervention, including those provided to participants or used in intervention delivery or training of intervention providers. Provide information on where the materials can be accessed (eg, online appendix or URL).	—[a]
What (procedures): describe each of the procedures, activities, and/or processes used in the intervention, including any enabling or support activities.	• Genuine treatment: only one SMT will be performed in each of the 3 regions assessed (cervical, thoracic, and lumbopelvic). The genuine cervical intervention will be performed in the predetermined location using a supine position and a diversified maneuver. The thoracic spine will be treated in a prone position using a double hypothenar thrust maneuver. This will be done with or without an additional drop in the thoracic region at the discretion of the chiropractor and was based on muscle tone in the region. The lumbopelvic region will receive either a prone drop maneuver with 3-4 drops or a side posture push, depending on clinical findings. • Sham treatment: only one sham will be performed in each of the 3 regions assessed (cervical, thoracic, and lumbopelvic). For the sham cervical SMT, the participant will be placed in a supine position on a cervical drop piece with the chiropractor's forearm under the participant's neck and contact lateral to the cervical spine. This will only be performed unilaterally over the side with the predetermined restriction in the articular pillars. The thoracic sham and the lumbopelvic sham will be performed with the participant in the prone position. Contact for the thoracic sham will be medial to the medial scapular border, and contact for the lumbopelvic sham will be over the lateral gluteal muscles, opposite to the side of the chiropractor's stance. The chiropractor will contact close to the predetermined site of intersegmental motion restriction in the spine, but no manual thrust will be directed into the spine. Instead, the chiropractor will apply a downward pressure to the drop piece with their forearm (cervical) or knee (thoracic and lumbopelvic), which will elicit a noise from the table that mimics a genuine drop SMT, as well as induced movement of the participant's body. Only a single drop will be made in all regions to limit any potential joint motion with the procedure.
Who Provided: for each category of intervention provider (eg, psychologist and nursing assistant), describe their expertise, background, and any specific training given.	Georgia-licensed Doctors of Chiropractic (DCs) with 20+ years of combined experience in the delivery of HVLA SMT and ~3 months of sham protocol training.
How: describe the modes of delivery (eg, face-to-face or by some other mechanism, such as internet or telephone) of the intervention and whether it was provided individually or in a group.	Genuine and sham SMTs will be performed individually and face-to-face.
Where: describe the type(s) of location(s) where the intervention occurred, including any necessary infrastructure or relevant features.	Sessions will take place at the Dr Sid E. Williams Center for Chiropractic Research (CCR), Life University, Marietta, Georgia, United States.
When and how much: describe the number of times the intervention was delivered and over what period including the number of sessions, their schedule, and their duration, intensity, or dose.	Over the course of a ~1-week period, participants will attend 2 sessions (first ~2 hours, second ~1 hour) scheduled at approximately the same time of day.
Tailoring: if the intervention was planned to be personalized, titrated, or adapted, then describe what, why, when, and how.	Genuine and sham SMTs will be delivered to areas of spinal segmental dysfunction as determined by the chiropractic clinician using an established, reliable detection protocol.
Modifications: if the intervention was modified during the study, describe the changes (what, why, when, and how).	—
How well (planned): if intervention adherence or fidelity was assessed, describe how and by whom, and if any strategies were used to maintain or improve fidelity, describe them.	Participant adherence to the intervention schedule will be monitored by the chiropractic clinician. Attempts will be made to mitigate adherence issues via automated SMS text messaging reminders sent ~24 and ~3 hours prior to each session.

XSL•FO
RenderX

Item	Description
How well (actual): if intervention adherence or fidelity was assessed, describe the extent to which the intervention was delivered as planned.	—

[a]Not applicable.

Criteria for Discontinuing or Modifying Allocated Interventions

Participants will be advised that they are able to withdraw at any time without giving a reason or may be withdrawn by the investigators if they experience significant adverse effects that are deemed detrimental to their well-being, as determined by the clinician.

Strategies to Improve Adherence to Interventions

We will attempt to mitigate adherence issues via automated SMS text messaging reminders sent ~24 hours and ~3 hours prior to each session.

Plans to Promote Participant Retention and Complete Follow-Up

Once enrolled, every reasonable effort will be made to follow participants throughout the entirety of the study period via ongoing email and SMS text messaging correspondence. In the event of premature discontinuation of the study for any reason, participants will be made aware that all data collected up to the point of withdrawal may be used for analyses, though they will have the right to pull their data if so desired.

Relevant Concomitant Care Permitted or Prohibited During the Trial

Participants will be asked to avoid any outside manual therapies (eg, chiropractic and osteopathy) from 14 days prior to their initial session until the end of the trial period. Any introduction of manual therapies during this period will render participants ineligible.

Measurements (All Assessments Will Be Performed by the Data Team)

Primary Outcome

Blinding will be assessed immediately following each session and directly prior to the second session via a brief electronic survey whereby participants will be asked (1) their perceived group allocation (“genuine adjustments,” “sham adjustments,” and “I don’t know”), (2) if their group assignment was revealed to them in any way (“No,” “Yes - explain”), (3) if they felt any changes in their body following their sessions (“No,” “Yes-explain,” and “I don’t know”), and (4) if they have received chiropractic care prior to this study (“No” and “Yes-for how long-what kind”). This data will be stored in Jotform. The Bang BI is a continuous value from −1 to 1 computed on each trial arm [22]. Per the index developers, an index score of 1 indicates that all responses are correct, suggesting complete unblinding. An index of −1 denotes that all responses are incorrect, suggesting either complete blinding or unblinding in the opposite direction if participants have some bias leading them to believe they are in the opposite arm. An index of 0 indicates half of the responses are correct and half are incorrect, which suggests random guessing. Arm-level Bang BIs between −0.3 and 0.3 are commonly interpreted as “successful blinding” [14-16]. Successful blinding might also be inferred in cases of “wishful thinking,” whereby trial participants tend to believe they received genuine treatment irrespective of group assignment, resulting in Bang BIs for the genuine and sham groups approaching 1 and −1, respectively (summed Bang BI ≈0) [23].

Secondary Outcomes

James BI

The James BI will also be used for evaluating the blinding success of the participants based on their electronic survey answers as noted above. The James BI is calculated based on the proportion of participants who correctly or incorrectly guess their treatment group, where a “don’t know” response is considered a successful blinding outcome [22]. This index is scored on a scale of 0-1; a score of 0 indicates complete lack of blinding, while a value of 1 indicates perfect blinding. Generally, acceptable blinding is indicated when values are between 0.5 and 1 [13].

Gait

Secondary outcomes will assess the effects of spinal manipulative therapy on spatiotemporal gait characteristics. Gait will be evaluated using inertial measurement motion capture units and an instrumented force-plate treadmill (Noraxon USA Inc). Spatiotemporal gait parameters will include step length, step time, cadence, and velocity. Gait events will be identified using vertical ground reaction force thresholds implemented within Noraxon’s analytical pipeline. Where applicable, parameters will be processed bilaterally and averaged to generate representative values for each participant. Gait will be evaluated at baseline and at the end of the second session to assess changes over time, as well as differences between participants receiving genuine SMT versus sham SMT.

Sample Size

As recommended in the literature [24], a precision-based approach (ie, 95% CI width) was used to determine the sample size. Using a custom R script (R Foundation for Statistical Computing) developed by Muñoz Laguna et al [15] for a similar study [16], it was estimated that a sample size of ~25 participants per arm would yield Bang BI 95% CI widths of 0.45. Allowing for ~15% loss to follow-up, we computed an adjusted sample size of ~30 participants per arm (ie, 25/[1.0 - 0.15]=~30).

Statistical Considerations (Primary Outcome)

The primary focus will be on arm-specific blinding, which will be reported using Bang BI point estimates and 95% CIs computed using a custom script written in “R” (v4.4.3) [25], which incorporates the “BI” package [26].

XSL•FO
RenderX

Statistical Considerations (Secondary Outcomes)

As our sample size determination was based on our primary outcome (ie, Bang BI), estimates derived from secondary end points will be considered exploratory and interpreted with caution. Additionally, to guard against over-interpretation of results, we will only perform estimation and not perform null hypothesis significance testing. James BI point estimates and 95% CIs will be constructed using a custom R script incorporating the "BI" package. For group-level analysis of other secondary outcomes, mixed models will be run via custom R scripts incorporating the lme4 package [27]. Mixed models account for nonindependent data (eg, repeated measures) and have been shown to provide robust estimates in longitudinal designs even with high levels of missing data [28]. Between-group change score differences over time will be evaluated using linear mixed models with random intercepts adjusting for age, sex, and baseline values.

```
lmm=lmer(y~ybaseline+1+time*group+ybaseline+age+sex+(1 | ID),data=d)
```

Assumption checks will be performed via "visualize()" in the flexplot package [29,30]. If model assumptions are severely violated, data were evaluated via generalized linear mixed models incorporating a gamma distribution and a log link.

```
glmm=glmer(y~ybaseline+1+time*group+ybaseline+age+sex+(1 | ID),data=d,family=Gamma(link=log))
```

Between-group change score difference point estimates and 95% CIs will be computed and plotted using the emmeans package [31]. Gait parameters include multiple correlated measures. Therefore, Bonferroni correction will be applied to account for multiplicity and control the family-wise type I error rate.

Methods in Analysis to Handle Protocol Nonadherence and Any Statistical Methods to Handle Missing Data

We will report the number and percentages of withdrawals in each of the groups. Based on the trialists' experience, loss to follow-up is expected to be no more than 15%, which has been factored into our sample size calculation. Regarding secondary outcomes, mixed models have been shown to provide robust estimates in longitudinal designs even with high levels of missing data.

Adverse Event Reporting and Harms

The principal investigator (PI) will be responsible for ensuring participants' safety and for reporting anticipated adverse events, unanticipated adverse events (UAEs), unanticipated problems (UPs), and serious adverse events (SAEs) to the institutional review board (IRB). Once the PI is made aware of any SAEs, UAEs, or UPs by study personnel, she will prepare and send a report to the IRB within 5 business days. Anticipated adverse events will be reported in the annual continuing review application. In response to adverse events (AEs) and UPs, the PI may choose to incorporate addendums into the study protocol to mitigate risks. If the PI determines that there is imminent danger to the trial participants, she may terminate the study only following consultation with other trialists. If this occurs, the IRB and participants will be immediately notified, and all data collected up to the point of suspension or termination may be used in the final analyses. A list of any reported AEs, UAEs, UPs, and SAEs will be logged and maintained by a lab team member.

Data Management and Processing

To ensure data integrity, trialists review protocols monthly to ensure the procedures for storing and sharing data are being followed. The results from the trial will be reported in aggregate, and the data will be retained for a minimum of 10 years.

Access to Data

The final trial dataset will be password protected and housed locally at the CCR. Oversight officials (eg, Life University's IRB) may also be given access to the files upon request.

Plans to Give Access to the Full Protocol, Participant-Level Data, and Statistical Code

The full trial protocol will be submitted for publication to a peer-reviewed, open-access journal prior to the completion of data collection and analyses. In line with scientific imperatives of increased transparency, reproducibility, and interpretations of trials, statistical codes and deidentified data sets will be provided to interested researchers upon request.

Oversight and Monitoring

Composition of the Data Monitoring Committee, Its Role, and Reporting Structure

No independent data and safety monitoring board was required for this trial. Per the CCR's standard policies developed to ensure good trial practices and identify emerging trends, our lab has in place a data and safety monitoring plan whereby trialists review the data collection procedures, AEs, and trial progress once monthly.

Frequency and Plans for Auditing Trial Conduct

Auditing of trial conduct will be performed once monthly by the trialists.

Plans for Communicating Important Protocol Amendments to Relevant Parties (eg, Trial Participants and Ethical Committees)

Substantive protocol amendments that may impact the conduct of the study will be agreed upon by the research team, submitted to the IRB for approval, reported in the trial registry, reflected in the participant-facing materials (eg, informed consent documents), and outlined in the final manuscript.

Dissemination Policy

Every effort will be made to minimize the interval between the completion of data collection and the release of study results. We estimate this process to take up to 12 months. Irrespective of the magnitude or direction of effect, results from the study will be submitted for publication in an international peer-reviewed scientific journal, presented at scientific conferences, and may form part of future grant applications.

Ethical Considerations

Ethical approval was obtained from the IRB of Life University (2025-03-27-Sliwka; approved on March 27, 2025; 1269 Barclay Cir, Marietta, GA 30060). This is a protocol, and no participants

XSL•FO
RenderX

have been recruited to date for the sake of this manuscript. Participants will receive up to US $75 in gift cards to compensate for their time (a US $25 gift card after the first session and a US $50 gift card after the second session). Photo releases for images in this manuscript have been procured.

Written informed consent will be collected at Life University's CCR. At the initial meeting, a research assistant will (1) provide each potential participant with a paper copy of the informed consent form written in English, (2) outline the key aspects of the trial, (3) ask if they have any questions about the trial, and (4) request the participant's signature. Participants will be informed that they may withdraw at any time without giving a reason and that all data collected up to the point of withdrawal may be used in the final analysis.

All information generated in this study will be considered highly confidential and is not to be shared with any persons not directly concerned with the trial. For deidentification purposes, participants will be assigned unique study numbers upon enrollment. Apart from the health history forms, all electronic records will be identified solely using assigned study numbers and stored locally in a password-protected database. Any records that contain personal identifiers (ie, informed consent forms, health history forms) will be stored on-site in a locked office or locally in a password-protected database accessible only to the researchers directly involved in the trial.

Results

This study is being supported and internally funded by the Dr. Sid E. Williams Center for Chiropractic Research at Life University in Marietta, Georgia, United States. Our study was prospectively registered on clinicaltrials.gov (NCT06931600) on March 28, 2025, and the first participant was seen on May 5, 2025. As of February 11, 2026, 25 participants have been enrolled. Projections for completing data collection, data analysis, and manuscript submission are Summer 2026, Fall 2026, and Winter 2027, respectively.

Discussion

Overview

Though health care clinical trials tend to rely on RCTs for treatment analysis, creating an inert yet believable sham for SMT trials has proven challenging [1,3]. Despite attempts using techniques ranging from light touch to simulated thrusts, there is no universally recognized manual sham for SMT trials, and few sham techniques account for participant blinding [6-12].

Principal Findings

To address the lack of a standardized manual SMT sham, our lab will launch a study testing a novel full-spine manual SMT sham. Our main goal is to assess the blinding success of this novel sham. Secondary aims are to explore the possible effect(s) of SMT on gait parameters by comparing movement characteristics and gait-related motor control between individuals receiving genuine SMT and those receiving sham SMT.

Comparison to Prior Work

This multisession trial is the second step in our lab's quest to develop a standardized sham protocol. Previously, a feasibility study was conducted in our lab that assessed blinding success following a single session of the same sham and genuine chiropractic techniques mentioned in this protocol. Bang BI point estimates suggested a high level of unblinding in the genuine group but successful blinding in the sham, both immediately posttreatment as well as 48 hours later. We hypothesize similar results from this study.

Strengths and Limitations

One potential limitation of our protocol is that we are not undertaking a crossover design. By limiting participants to one study arm, they will not receive both genuine and sham SMT, therefore not controlling for between-subject variability. However, since manual therapies may have long-lasting neurophysiological effects [32], a parallel study can avoid carryover effects. When it comes to gait, there is a high variability due to individual differences, which we will account for using a generalized linear mixed model. Another potential limitation relates to our recruitment of participants from a geographic area with a population that is more familiar with chiropractic care due to the local chiropractic college. We will exclude participants with substantial chiropractic education to help mitigate this potential bias.

Future Directions

Ultimately, this line of research is endeavoring to develop a novel manual SMT sham protocol that can be used to improve the quality of future chiropractic RCTs that use clinical measures as primary outcomes.

Dissemination Plan

Regardless of outcomes, the findings will be published in an international peer-reviewed journal within 12 months of trial completion.

Acknowledgments

Amber Jackson, DC, a chiropractor at Life University, will also be contributing to this study by providing interventions. Generative AI was not used in the production of this manuscript.

Funding

There were no outside sources of funding for this protocol and its pilot trial. Internal funding was provided by Dr Sid E. Williams Center for Chiropractic Research (CCR), Life University, Marietta, GA, USA.

XSL•FO
RenderX

Data Availability

In line with scientific imperatives of increased transparency, reproducibility, and interpretations of trials, statistical codes and deidentified datasets will be provided to interested researchers upon request.

Authors' Contributions

Conceptualization: MS, TP, AQ, SS
Data curation: TP, AQ
Formal analysis: TP, AQ
Funding acquisition: SS
Investigation: MS
Methodology: MS, TP, AQ, SS
Project administration: MS
Resources: MS
Software: TP, AQ
Supervision: MS, SS
Validation: MS, TP, AQ
Visualization: MS, TP, AQ
Writing – original draft: MS, TP, AQ
Writing – review & editing: MS, TP, AQ, SS

Conflicts of Interest

None declared.

References

1. Karanicolas PJ, Farrokhyar F, Bhandari M. Practical tips for surgical research: blinding: who, what, when, why, how? Can J Surg 2010 Oct;53(5):345-348. [Medline: 20858381]
2. D'Alessandro G, Ruffini N, Iacopini A, Annoni M, Kossowsky J, Cerritelli F. Overcoming placebo-related challenges in manual therapy trials: the 'whats and hows' and the 'touch equality assumption' proposals. Int J Osteopath Med 2021 Dec;42:5-10. [doi: 10.1016/j.ijosm.2021.10.003]
3. Hage R, Demoulin C, J. Puentedura E, Brismée JM. Sham treatment? Should it be redefined for its use in manual therapy clinical studies? J Man Manip Ther 2025 Jan 2;33(1):1-3. [doi: 10.1080/10669817.2025.2448921]
4. Lavazza C, Galli M, Abenavoli A, Maggiani A. Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pairwise meta-analysis. BMJ Open 2021 May 4;11(5):e045106. [doi: 10.1136/bmjopen-2020-045106] [Medline: 33947735]
5. Hancock MJ, Maher CG, Latimer J, McAuley JH. Selecting an appropriate placebo for a trial of spinal manipulative therapy. Aust J Physiother 2006;52(2):135-138. [doi: 10.1016/s0004-9514(06)70049-6] [Medline: 16764551]
6. Vernon H, Puhl A, Reinhart C. Systematic review of clinical trials of cervical manipulation: control group procedures and pain outcomes. Chiropr Man Therap 2011 Jan 11;19(1):3. [doi: 10.1186/2045-709X-19-3] [Medline: 21247413]
7. Lougee H, Johnston RG, Thomson OP. The suitability of sham treatments for use as placebo controls in trials of spinal manipulative therapy: a pilot study. J Bodyw Mov Ther 2013 Jan;17(1):59-68. [doi: 10.1016/j.jbmt.2012.06.005] [Medline: 23294685]
8. Machado LAC, Kamper SJ, Herbert RD, Maher CG, McAuley JH. Imperfect placebos are common in low back pain trials: a systematic review of the literature. Eur Spine J 2008 Jul;17(7):889-904. [doi: 10.1007/s00586-008-0664-3] [Medline: 18421484]
9. Puhl AA, Reinhart CJ, Doan JB, Vernon H. The quality of placebos used in randomized, controlled trials of lumbar and pelvic joint thrust manipulation-a systematic review. Spine J 2017 Mar;17(3):445-456. [doi: 10.1016/j.spinee.2016.11.003] [Medline: 27888138]
10. Ruddock JK, Sallis H, Ness A, Perry RE. Spinal manipulation vs sham manipulation for nonspecific low back pain: a systematic review and meta-analysis. J Chiropr Med 2016 Sep;15(3):165-183. [doi: 10.1016/j.jcm.2016.04.014] [Medline: 27660593]
11. Hawk C, Long CR, Reiter R, Davis CS, Cambron JA, Evans R. Issues in planning a placebo-controlled trial of manual methods: results of a pilot study. J Altern Complement Med 2002 Feb;8(1):21-32. [doi: 10.1089/107555302753507159] [Medline: 11890430]
12. Ernst E, Harkness E. Spinal manipulation: a systematic review of sham-controlled, double-blind, randomized clinical trials. J Pain Symptom Manage 2001 Oct;22(4):879-889. [doi: 10.1016/s0885-3924(01)00337-2] [Medline: 11576805]
13. Bang H, Flaherty SP, Kolahi J, Park J. Blinding assessment in clinical trials: a review of statistical methods and a proposal of blinding assessment protocol. Clin Res Regul Aff 2010 Jun;27(2):42-51. [doi: 10.3109/10601331003777444]

XSL•FO
RenderX

14. Reid E, Kamlin OF, Orsini F, et al. Success of blinding a procedural intervention in a randomised controlled trial in preterm infants receiving respiratory support. Clin Trials 2023 Oct;20(5):479-485. [doi: 10.1177/17407745231171647] [Medline: 37144610]
15. Muñoz Laguna J, Kurmann A, Hofstetter L, et al. Feasibility of blinding spinal manual therapy interventions among participants and outcome assessors: protocol for a blinding feasibility trial. Pilot Feasibility Stud 2024 May 2;10(1):70. [doi: 10.1186/s40814-024-01492-6] [Medline: 38698433]
16. Muñoz Laguna J, Kurmann A, Hofstetter L, et al. "Which treatment do you believe you received?" A randomised blinding feasibility trial of spinal manual therapy. Chiropr Man Therap 2025 Jan 14;33(1):4. [doi: 10.1186/s12998-024-00561-0] [Medline: 39810207]
17. Yanuck SB, Fox SK, Harting BR, Motyka TM. Effect of manual manipulation on mechanical gait parameters. J Osteopath Med 2024 Oct 1;124(10):437-446. [doi: 10.1515/jom-2023-0203] [Medline: 38807459]
18. Aspinall SL, Nim C, Hartvigsen J, et al. Waste not, want not: call to action for spinal manipulative therapy researchers. Chiropr Man Therap 2024 May 14;32(1):16. [doi: 10.1186/s12998-024-00539-y] [Medline: 38745213]
19. Fernández-de-las-Peñas C, Alonso-Blanco C, San-Roman J, Miangolarra-Page JC. Methodological quality of randomized controlled trials of spinal manipulation and mobilization in tension-type headache, migraine, and cervicogenic headache. J Orthop Sports Phys Ther 2006 Mar;36(3):160-169. [doi: 10.2519/jospt.2006.36.3.160] [Medline: 16596892]
20. Bergmann TF, Peterson DH. Chiropractic Technique: Principles and Procedures, 3rd edition: Mosby; 2010.
21. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 7;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
22. Bang H, Ni L, Davis CE. Assessment of blinding in clinical trials. Control Clin Trials 2004 Apr;25(2):143-156. [doi: 10.1016/j.cct.2003.10.016] [Medline: 15020033]
23. Bang H. Random guess and wishful thinking are the best blinding scenarios. Contemp Clin Trials Commun 2016 Aug 15;3:117-121. [doi: 10.1016/j.conctc.2016.05.003] [Medline: 27822568]
24. Bland JM. The tyranny of power: is there a better way to calculate sample size? BMJ 2009 Oct 6;339:b3985. [doi: 10.1136/bmj.b3985] [Medline: 19808754]
25. R Core Team. R: a language and environment for statistical computing (version 4.5.1). R Foundation for Statistical Computing. 2024. URL: https://www.R-project.org [accessed 2026-02-09]
26. Schwartz M, Mercaldo N. BI: blinding assessment indexes for randomized, controlled, clinical trials. R package ver100. 2021. URL: https://cran.r-project.org/web/packages/BI/index.html [accessed 2026-02-28]
27. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Softw 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
28. Chakraborty H, Gu H. A Mixed Model Approach for Intent-to-Treat Analysis in Longitudinal Clinical Trials with Missing Values: RTI Press; 2009. [doi: 10.3768/rtipress.2009.mr.0009.0903]
29. Fife DA, Longo G, Correll M, Tremoulet PD. A graph for every analysis: mapping visuals onto common analyses using flexplot. Behav Res Methods 2021 Oct;53(5):1876-1894. [doi: 10.3758/s13428-020-01520-2] [Medline: 33634423]
30. Fife D. Flexplot: graphically-based data analysis. Psychol Methods 2022 Aug;27(4):477-496. [doi: 10.1037/met0000424] [Medline: 34843276]
31. Lenth R. Emmeans: estimated marginal means, aka least-squares means. R package version 185. 2023. URL: https://CRAN.R-project.org/package=emmeans [accessed 2026-02-28]
32. Bialosky JE, Bishop MD, Price DD, Robinson ME, George SZ. The mechanisms of manual therapy in the treatment of musculoskeletal pain: a comprehensive model. Man Ther 2009 Oct;14(5):531-538. [doi: 10.1016/j.math.2008.09.001] [Medline: 19027342]

Abbreviations

Bang BI: Bang Blinding Index
CCR: Center for Chiropractic Research
HVLA: high velocity, low amplitude
IRB: institutional review board
James BI: James Blinding Index
PI: principal investigator
PWS: preferred walking speed
RCT: randomized controlled trial
SAE: serious adverse event
SMT: spinal manipulative therapy
TIDieR: Template for Intervention Description and Replication
UAE: unanticipated adverse event
UP: unanticipated problem

XSL•FO
RenderX

Edited by J Sarvestan; submitted 22.Oct.2025; peer-reviewed by H Bang; revised version received 17.Dec.2025; accepted 27.Dec.2025; published 27.Mar.2026.

Please cite as:
Sliwka M, Perez T, Qazi A, Sullivan S
Spinal Manipulative Therapy on a Healthy Population: Protocol for a Randomized Blinding Feasibility Trial
JMIR Res Protoc 2026;15:e85956
URL: https://www.researchprotocols.org/2026/1/e85956
doi:10.2196/85956

© Margaret Sliwka, Tyson Perez, Ahmed Qazi, Stephanie Sullivan. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Tunnel Anastomosis vs Double-Tract Jejunal Interposition Reconstruction After Proximal Gastrectomy: Protocol for a Multicenter Prospective Randomized Controlled Trial

Qingyu Xie[1*], MSc; Rui Peng[1*], PhD; Chao Yue[1*], PhD; Wei Wei[1*], MSc; Lingli Huang[1*], PhD; Xiaoxiao Wang[2], PhD; Haitian Wang[1], MSc; Liang Chen[1], MSc; Rongmin Gu[1], BSc; Huanqiu Chen[1], BSc; Xuezhi Ming[1], PhD; Xu Wen[1], PhD; Weiguo Xu[1], PhD; Guangli Sun[1], PhD; Hao Fan[1], PhD; Zhe Wang[1], BSc; Longhao Yang[1], BSc; Xiaohua Zhou[3*], PhD; Xiaoyu Wu[2*], PhD; Jin Zhou[4*], PhD; Meng Wang[5*], PhD; Hao Xu[6*], PhD; Gang Li[1*], PhD

[1]The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, No. 42, Baiziting, Nanjing, China

[2]The Affiliated Hospital of Nanjing University of Chinese Medicine, Nanjing, China

[3]Jiangsu University Affiliated Gaochun Hospital, Nanjing, China

[4]The First Affiliated Hospital of Soochow University, Suzhou, China

[5]Nanjing University Medical School Affiliated Drum Tower Hospital, Nanjing, China

[6]The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

[*]these authors contributed equally

Corresponding Author:
Gang Li, PhD
The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, No. 42, Baiziting, Nanjing, China

Abstract

Background: Tunnel anastomosis is a novel anastomotic technique for digestive tract reconstruction following proximal gastrectomy. A previous retrospective study by our team demonstrated its favorable antireflux effect; therefore, we hypothesize that tunnel anastomosis is noninferior to double-tract jejunal interposition reconstruction in preventing postoperative reflux esophagitis, and we will conduct this prospective study to further validate this assumption.

Objective: In this study, we will prospectively compare tunnel anastomosis with the currently more prevalent double-tract jejunal interposition reconstruction technique to further validate its safety and efficacy.

Methods: This is a multicenter prospective randomized controlled study that will enroll 240 patients who will undergo proximal gastrectomy. The study will be divided into 2 groups: the tunnel anastomosis group and the double-tract jejunal interposition reconstruction group, with 120 patients in each group. Patients will undergo clinical assessments and complete questionnaires preoperatively, as well as at the 3rd, 6th, and 12th months postoperatively. The primary end point is the incidence of reflux esophagitis within 1 year. The secondary end points include perioperative safety, postoperative quality of life, and postoperative nutritional status.

Results: Recruitment of patients commenced in March 2022 and is scheduled to conclude in February 2027. The follow-up for all enrolled patients will be completed by February 2028.

Conclusions: To our knowledge, this is the first prospective study on this technique, aiming to provide novel insights into the methods of digestive reconstruction following proximal gastrectomy.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2200057397; https://www.chictr.org.cn/showprojEN.html?proj=154418

International Registered Report Identifier (IRRID): DERR1-10.2196/82712

(JMIR Res Protoc 2026;15:e82712) doi:10.2196/82712

KEYWORDS

gastric cancer; proximal gastrectomy; tunnel anastomosis; double-tract jejunal interposition reconstruction; antireflux

XSL•FO
RenderX

Introduction

Gastric cancer is a type of cancer with high morbidity and mortality in China, and the incidence of proximal gastric cancer is currently increasing [1-3]. Surgical treatment for proximal gastric cancer involves either total gastrectomy or proximal gastrectomy. Research has shown that proximal gastrectomy has the same therapeutic potential and efficacy as total gastrectomy, with advantages in terms of postoperative weight loss, dumping syndrome, anemia, and nutritional supplementation [4-10]. However, the issue of reflux esophagitis following proximal gastrectomy remains a significant challenge for surgeons [11]. Numerous studies have been conducted on antireflux techniques for digestive reconstruction, including double-tract jejunal interposition reconstruction and the double-flap technique (Kamikawa anastomosis), which have exhibited promising antireflux effects. Nevertheless, there is still no consensus on which reconstruction method should be considered the standard [12-17].

Tunnel anastomosis is a novel technique for digestive tract reconstruction following proximal gastrectomy. Owing to its preservation of an intact muscular flap, it ensures superior blood supply to the anastomotic site. A previous study by our team demonstrated that the antireflux effects of this technique are favorable [18]. Double-tract jejunal interposition reconstruction is currently the mainstream method for digestive tract reconstruction following proximal gastrectomy and is widely recognized by most experts [19]. The incidence of postoperative reflux esophagitis following this procedure is approximately 10%. Therefore, we selected this technique as the control group to further validate the effectiveness of TA.

We hypothesize that tunnel anastomosis is noninferior to double-tract jejunal interposition reconstruction in preventing postoperative reflux esophagitis. The objective of this study is to further validate the surgical safety and antireflux effect of tunnel anastomosis and assess the efficacy and ability of this technique to improve the quality of life of patients following proximal gastrectomy. It is hoped that this research will contribute to enhancing the theoretical and clinical practice foundations for refining surgical treatment strategies for upper gastric cancer.

Methods

Ethical Considerations

This study has been approved by the ethics committee of the Affiliated Cancer Hospital of Nanjing Medical University (2021-091-01). All patients will be fully informed of the precautions by a professional physician and provided with a written informed consent form before participation. All research information will be kept strictly at the study site and these data will be anonymized. This information will not be published outside the research without the consent of the patients.

Study Design and Participants

This multicenter prospective study will be conducted at the Affiliated Cancer Hospital of Nanjing Medical University, First Affiliated Hospital of Nanjing Medical University, Affiliated Drum Tower Hospital of Nanjing University Medical School, First Affiliated Hospital of Soochow University, Affiliated Hospital of Nanjing University of Chinese Medicine, and Affiliated Gaochun Hospital of Jiangsu University. Figure 1 illustrates the flow of the study.

The inclusion criteria for this study are as follows: (1) age range of 18 to 80 years, with no sex preference; (2) histopathological confirmation of adenocarcinoma via endoscopic biopsy; (3) tumor located in the upper third of the stomach without esophageal involvement, largest diameter of ≤4 cm, and preoperative clinical stage cT1-4aN0M0; (4) Eastern Cooperative Oncology Group Performance Status score of 0 to 1; (5) no surgical contraindications identified through comprehensive preoperative evaluations; and (6) voluntary and informed consent signed by the patient or their legal representative. The exclusion criteria are as follows: (1) pregnant or lactating women; (2) presence of severe psychiatric disorders; (3) intraoperative findings indicating tumor invasion into the esophagus or unsuitability for proximal gastrectomy as determined by the primary surgeon; (4) preoperative or intraoperative discovery of distant organ metastasis or extensive peritoneal implantation metastasis; (5) presence of concurrent or metachronous malignancies, including other organ tumors; (6) incomplete radical surgery, including patients who underwent palliative tumor resection; (7) a history of gastrointestinal surgery; (8) a history of neoadjuvant radiotherapy or chemotherapy; (9) serious concomitant diseases that may make the survival period <5 years; and (10) cases considered unsuitable by the investigator.

XSL•FO
RenderX

Figure 1. Flow diagram for the study.

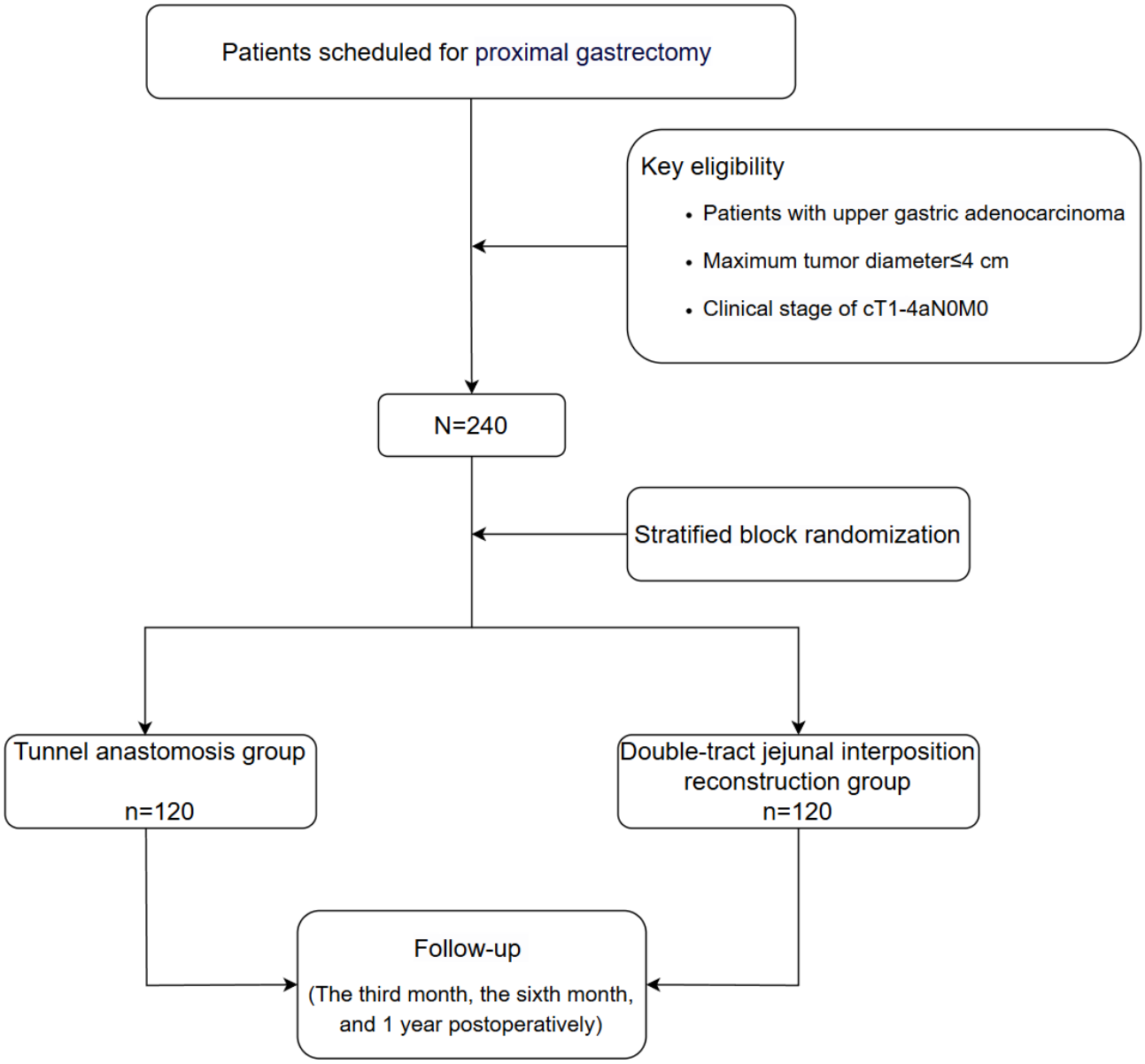

Postrecruitment Withdrawals and Exclusions

Patients can withdraw from this study at any time. For patients who withdraw, the information collected before withdrawal will be used for the final analysis unless they request that their information be deleted.

Preoperative Clinical Staging

Before surgery, all patients will be required to undergo an endoscopic examination with biopsy, endoscopic ultrasonography, a plain scan and a contrast-enhanced scan of the whole abdomen, and an upper gastrointestinal series. This comprehensive assessment aims to evaluate the gastric cancer lesion and abdominal lymph nodes thoroughly, with the objective of obtaining an accurate preoperative clinical diagnosis and tumor staging.

Randomization

Patients can only be randomized into different groups once they have signed the informed consent form, completed baseline assessments, and been confirmed as fully meeting the inclusion criteria. Eligible patients will be randomly assigned to 1 of the following 2 groups at a 1:1 ratio: the tunnel anastomosis group or the double-tract jejunal interposition reconstruction group. The randomization process will use a stratified block randomization method, with disease stage (categorized as early stage or advanced stage) serving as the stratification factor. Patient group assignment will be determined based on random numbers generated by the R software program (version 4.0.2; R Foundation for Statistical Computing).

Surgical Procedure

Overview

Both groups will undergo proximal gastrectomy with radical lymph node dissection (open, laparoscopic, or robot-assisted). cT1N0 patients will undergo D1+ lymph node dissection (1, 2, 3a, 4sa, 4sb, 7, 8a, 9, and 11p), whereas the remaining patients will undergo D2 lymph node dissection (D1+ and 11d). The surgeon will be required to hold the qualification of a chief physician and strictly adhere to the protocol procedures for

XSL•FO
RenderX

reconstruction. All surgical procedures will be video or photo recorded.

Tunnel Anastomosis Technique

First, a linear cutting stapler will be used to transect the esophagus and create a gastric tube. Second, a 3-cm transverse incision will be made in the anterior wall of the remnant stomach, approximately 3 to 4 cm from the upper edge between the greater and lesser curvatures, reaching but not incising the muscular layer. Third, another parallel incision of equal length 3.5 cm distal to the first incision will be made. Fourth, between these 2 incisions, the connective tissue between the submucosa and muscular layers will be dissected via an electrosurgical knife, creating a tunnel flap of approximately 3 × 3.5 cm. Fifth, the posterior wall of the esophagus, located 5 cm from the residual stump, will be sutured with 4 stitches to the gastric wall at the upper edge of the seromuscular flap. Sixth, the esophageal stump will be pulled through the tunnel, and the anterior wall of the esophagus will be sutured with 4 to 5 stitches to the upper edge of the gastric seromuscular flap. Seventh, the submucosa and mucosal layers of the stomach will be incised at the lower incision of the tunnel to prepare for anastomosis with a caliber of 3 cm. Eighth, the esophageal stump will be opened using an ultrasonic knife, and the posterior wall of the esophagus will be sutured to the gastric mucosa and submucosa. Ninth, the anterior wall of the esophagus will be sutured to the full layer of the stomach. Finally, the lower edge of the seromuscular flap and the seromuscular layer of the remnant stomach will be sutured (Figure 2) [18].

XSL•FO
RenderX

Figure 2. Surgical steps for tunnel anastomosis.

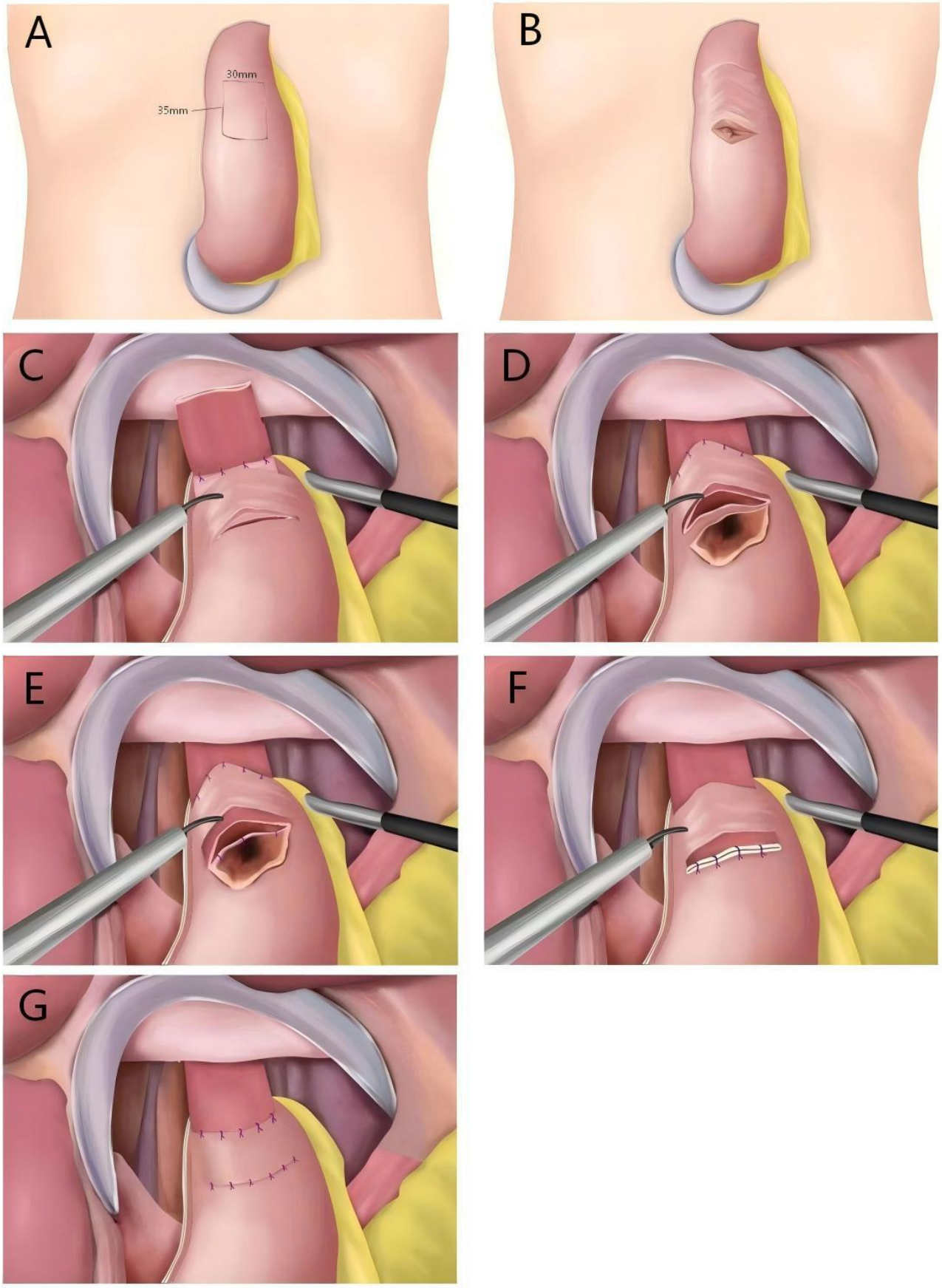

XSL•FO
RenderX

Double-Tract Jejunal Interposition Reconstruction Technique

First, mesenteric vessels will be ligated 15 to 20 cm from the Treitz ligament. Second, the distal jejunum will be pulled anterior to the transverse colon and anastomosed with the esophagus. Third, approximately 10 to 15 cm below the esophagojejunal anastomosis, the distal jejunum will be anastomosed with the anterior wall of the remnant stomach. A 6-cm linear stapler will be used to create an anastomosis with a size of 4 cm. Subsequently, at 30 to 35 cm from the gastrojejunal anastomosis, a second anastomosis will be performed between the proximal and distal jejunum.

Follow-Up

Follow-up will be conducted by dedicated follow-up staff. All patients will undergo follow-up visits at 3 months, 6 months, and 1 year postoperatively. At these visits, hematological examinations and postoperative quality of life assessments (using the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 [EORTC QLQ-C30] and 36-Item Short Form Health Survey) will be performed. Nutritional status will be assessed based on hematological test results, changes in body weight, and the prognostic nutritional index (PNI). In addition, reflux symptoms will be assessed via the Visick score. At the 1-year follow-up, gastroscopy will be performed to assess reflux esophagitis according to the Los Angeles (LA) classification.

Assessment of Outcomes

Primary End Points

The primary end point is the incidence of reflux esophagitis, which will be determined via gastroscopy results 1 year postoperatively. To determine the severity of reflux esophagitis, the modified LA classification system for reflux esophagitis will be used, which is considered a reliable method for categorizing reflux esophagitis. The modified LA classification criteria are as follows [20]:

- Grade N—normal mucosa
- Grade M—minimal changes to the mucosa, such as erythema and/or whitish turbidity
- Grade A—nonconfluent mucosal breaks of <5 mm in length
- Grade B—nonconfluent mucosal breaks of >5 mm in length
- Grade C—confluent mucosal breaks <75% circumferential
- Grade D—confluent mucosal breaks >75% circumferential

Secondary End Points

The secondary study end points include perioperative safety, postoperative nutritional status, and postoperative quality of life.

Perioperative safety primarily encompasses the duration of surgery, amount of intraoperative blood loss, and postoperative complication status. Nutritional status and quality of life following surgery will be assessed at the 3rd, 6th, and 12th postoperative months. Postoperative nutritional status will be determined based on changes in patient weight, hemoglobin levels, total protein levels, albumin levels, total lymphocyte count, and PNI. Additionally, patients' quality of life will be evaluated via the EORTC QLQ-C30 and daily food intake frequency. The Visick score will be used for assessing patients' symptoms and quality of life [21]:

- Grade I—patients are asymptomatic or have only mild symptoms that do not significantly affect their quality of life. Surgical outcomes are considered excellent.
- Grade II—patients experience mild symptoms that have minimal impact on their daily activities. Overall, patients are satisfied with the surgical outcome.
- Grade III—patients have moderate to severe symptoms that affect their daily lives. These symptoms may require medical treatment or lifestyle modifications. The surgical outcome is considered less favorable.
- Grade IV—patients have severe symptoms that significantly impact their quality of life. The surgical procedure has failed to achieve the desired outcome. Further surgical intervention or other treatments may be necessary.

Power and Sample Size

This study is designed as a prospective randomized controlled trial with a noninferiority objective. It comprises 2 groups: the experimental group (tunnel anastomosis group) and the control group (double-tract jejunal interposition reconstruction group).

The reported incidence of reflux esophagitis for DTJIR in the literature varies widely, ranging from 1.7% to 25%. However, most contemporary studies report rates concentrated at approximately 10% [22-24]. On the basis of retrospective data from our center, the incidence of reflux esophagitis for TA is approximately 5%. Therefore, the estimated incidence of reflux esophagitis was 10% in the DTJIR group and 5% in the TA group. With a significance level set at α=.025 (one sided), a power of $1 - \beta$ of 0.80, a noninferiority margin of 0.05 (given the reported incidence of reflux esophagitis with the current surgical options, a noninferiority margin of ≤5% was deemed clinically acceptable as such a difference would not offset the significant benefits associated with the new technique), a 1:1 ratio between the experimental and control groups, and an anticipated dropout rate of 10%, the sample size was calculated via the PASS software (NCSS, LLC). The resulting sample sizes for both groups were 120 patients each.

Therefore, a total of 240 patients will be included in this study, with 120 patients in the experimental group and 120 patients in the control group.

Data Collection

Basic patient information and clinical data will be collected before discharge, including identity card number; gender; age; contact information; date of operation; surgical approach; duration of surgery; intraoperative blood loss; pathological tumor, node, metastasis staging; number of lymph nodes dissected; number of positive lymph nodes; length of postoperative hospital stay; and postoperative complications. Each patient will be followed up on at the 3rd, 6th, and 12th months postoperatively, and the following data will be collected regarding their postoperative quality of life: daily meal frequency, Visick score, and postoperative quality of life score (using the EORTC QLQ-C30), as well as data on their postoperative nutritional status, including weight change,

XSL•FO
RenderX

hemoglobin level, total protein level, albumin level, total lymphocyte count, and PNI.

The basic information and clinical data for each patient collected before discharge will be classified, stored, and entered into the database. The results of each follow-up will be meticulously recorded and entered into the database alongside the other data. All electronic data will be backed up to prevent loss.

Statistical Analysis

Data analysis will be conducted by professional statisticians. In addition to the overall comparative analysis conducted between the TA and DTJIR groups, we will further perform subgroup analyses stratified by disease stage and surgical approach.

Intention-to-treat and per-protocol approaches will be applied for efficacy analysis. An as-treated approach will be used for safety analysis. For continuous variables, a normality test will be performed first. For those that conform to a normal distribution, the 2-tailed *t* test will be used for statistical analysis (all continuous values will be expressed as means and SDs). Those not conforming to a normal distribution will be presented as quartiles and rank means, and the Mann-Whitney *U* test will be used to calculate the *P* value to compare differences between groups. For categorical variables, the chi-square test will be used for statistical analysis. Repeated-measure data will be analyzed using a mixed-effects model. Subgroup analyses will be conducted by disease stage (early vs advanced); surgical approach (open, laparoscopic, or robot-assisted); age; macroscopic types; and tumor, node, metastasis stage. All *P* values calculated in the analysis will be 2-sided, and $P<.05$ will be considered statistically significant. If missing data constitute more than 5%, multiple imputations will be conducted. Statistical analyses will be performed via SPSS (IBM Corp), Prism (GraphPad Software), or PASS.

Results

The first patient was enrolled in March 2022, and the recruitment of all 240 patients is expected to be completed by February 2027. Follow-up for these patients is scheduled to conclude in February 2028.

Discussion

This study may demonstrate that tunnel anastomosis is an alternative method of digestive tract reconstruction following proximal gastrectomy that is as effective as DTJIR. Its antireflux effect may make it one of the most widely used methods of digestive tract reconstruction after proximal gastrectomy.

Proximal gastrectomy is gaining increasing acceptance among the medical community due to its advantages in postoperative nutrition [25]. However, for patients undergoing proximal gastrectomy, postoperative reflux esophagitis is a significant issue. Various anastomotic techniques have been investigated in an attempt to address this problem, yet a standard approach remains elusive [24]. In pursuit of a superior method for digestive tract reconstruction, we modified the Kamikawa anastomosis to develop tunnel anastomosis and conducted a retrospective study. The retrospective study found that tunnel anastomosis is as effective as DTJIR at preventing reflux esophagitis [15,18]. To further validate the preventive effect against reflux esophagitis of this technique, we will conduct this prospective study with the aim of providing a novel reference for digestive tract reconstruction following proximal gastrectomy. Compared with our previous study, this research has a clearly defined objective and rigorous structure. It can also reduce recall bias and control for confounding factors more effectively, yielding more reliable outcomes. Achieving strict adherence to the study protocol will also take longer.

To our knowledge, this is the first prospective study on this technique. To ensure the quality of the enrolled patients, we used stringent inclusion and exclusion criteria in this study. As this is a multicenter study, surgeons are required to possess extensive surgical experience to ensure the quality of the surgical procedures. The operative processes of patients will be documented through video and photographic records. We have also meticulously designed the subsequent follow-up and data analysis protocols. This well-considered design leads us to expect a genuine and reliable conclusion.

This study also has certain inherent limitations. It is not possible to blind patients, surgeons, radiologists, or clinical assessors in this trial. Both the physicians and patients have a clear understanding of the surgical procedure that will be performed. To minimize the bias introduced by the inability to implement blinding, the follow-up and data analysis will be conducted by different people. Meanwhile, as a prospective study, it may have the capacity to mitigate the impact of such bias on the results. The primary end point of this study is the 1-year incidence rate of reflux esophagitis. However, this fails to reflect the long-term antireflux effect. In the future, we will conduct longer-term follow-up to investigate the long-term efficacy of tunnel anastomosis. To increase the dissemination of our research, we will also present our study and its results at a wide range of academic conferences.

Acknowledgments

No artificial intelligence tools or services were used during the preparation of this work.

Funding

This work was supported by the Yishan Research Project of Jiangsu Cancer Hospital (YSZD202404); the Qunfeng Project of Jiangsu Cancer Hospital (GFXK202503); the National Natural Science Foundation of China (82203226, 82303602, and 82404002); the Research Project of Jiangsu Cancer Hospital (ZL202101, XHMS202401, ZJ202220, ZJ202203, and RCQY202409); the

XSL•FO
RenderX

Excellent Talent Program of Jiangsu Cancer Hospital (2017YYCJH-08); the Jiangsu Provincial Association for Science and Technology Youth Talent Lifting Project (JSTJ-2023-WJ013); the Jiangsu Province "333 Project" High-Level Talent Training Program (2024-3-2249); and the Jiangsu Provincial Health Commission Research Project (MQ2024056).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

GL, CY, and RP were primarily responsible for the study design. GL designed the tunnel anastomosis technique. QX and RP drafted and revised the manuscript. X Wang and RP were responsible for the sample size calculation and randomization. GL, HX, MW, JZ, X Wu, and XZ were responsible for the administrative oversight of their respective centers. QX, RP, CY, WW, LH, HW, LC, RG, HC, XM, X Wen, WX, GS, HF, ZW, LY, XZ, X Wu, JZ, MW, HX, and GL were actively involved in the execution of the research and data collection, analysis, and interpretation. All the authors critically reviewed the manuscript, provided final approval of the version to be published, and reached a consensus on the journal for submission.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[PDF File, 112 KB - resprot_v15i1e82712_app1.pdf]

References

1. He S, Xia C, Li H, et al. Cancer profiles in China and comparisons with the USA: a comprehensive analysis in the incidence, mortality, survival, staging, and attribution to risk factors. Sci China Life Sci 2024 Jan;67(1):122-131. [doi: 10.1007/s11427-023-2423-1] [Medline: 37755589]
2. Park DJ, Park YS, Ahn SH, Kim HH. Laparoscopic proximal gastrectomy as a surgical treatment for upper third early gastric cancer. Korean J Gastroenterol 2017 Sep 25;70(3):134-140. [doi: 10.4166/kjg.2017.70.3.134] [Medline: 28934829]
3. Yang WJ, Zhao HP, Yu Y, et al. Updates on global epidemiology, risk and prognostic factors of gastric cancer. World J Gastroenterol 2023 Apr 28;29(16):2452-2468. [doi: 10.3748/wjg.v29.i16.2452] [Medline: 37179585]
4. Fu J, Li Y, Liu X, et al. Clinical outcomes of proximal gastrectomy with gastric tubular reconstruction and total gastrectomy for proximal gastric cancer: a matched cohort study. Front Surg 2023 Jan 12;9:1052643. [doi: 10.3389/fsurg.2022.1052643] [Medline: 36713677]
5. Hirata Y, Kim HI, Grotz TE, Matsuda S, Badgwell BD, Ikoma N. The role of proximal gastrectomy in gastric cancer. Chin Clin Oncol 2022 Oct;11(5):39. [doi: 10.21037/cco-22-82] [Medline: 36336898]
6. Ikeguchi M, Kader A, Takaya S, et al. Prognosis of patients with gastric cancer who underwent proximal gastrectomy. Int Surg 2012;97(3):275-279. [doi: 10.9738/CC150.1] [Medline: 23113860]
7. Lee I, Oh Y, Park SH, Kwon Y, Park S. Postoperative nutritional outcomes and quality of life-related complications of proximal versus total gastrectomy for upper-third early gastric cancer: a meta-analysis. Sci Rep 2020 Dec 8;10(1):21460. [doi: 10.1038/s41598-020-78458-0] [Medline: 33293604]
8. Park DJ, Han SU, Hyung WJ, et al. Effect of laparoscopic proximal gastrectomy with double-tract reconstruction vs total gastrectomy on hemoglobin level and vitamin B12 supplementation in upper-third early gastric cancer: a randomized clinical trial. JAMA Netw Open 2023 Feb 1;6(2):e2256004. [doi: 10.1001/jamanetworkopen.2022.56004] [Medline: 36790808]
9. Yamasaki M, Takiguchi S, Omori T, et al. Multicenter prospective trial of total gastrectomy versus proximal gastrectomy for upper third cT1 gastric cancer. Gastric Cancer 2021 Mar;24(2):535-543. [doi: 10.1007/s10120-020-01129-6] [Medline: 33118118]
10. Yuan Z, Cui H, Xu Q, et al. Total versus proximal gastrectomy for proximal gastric cancer after neoadjuvant chemotherapy: a multicenter retrospective propensity score-matched cohort study. Int J Surg 2024 Feb 1;110(2):1000-1007. [doi: 10.1097/JS9.0000000000000927] [Medline: 38085808]
11. Rosa F, Quero G, Fiorillo C, et al. Total vs proximal gastrectomy for adenocarcinoma of the upper third of the stomach: a propensity-score-matched analysis of a multicenter western experience (on behalf of the Italian Research Group for Gastric Cancer-GIRCG). Gastric Cancer 2018 Sep;21(5):845-852. [doi: 10.1007/s10120-018-0804-3] [Medline: 29423892]
12. Ahn SH, Jung DH, Son SY, Lee CM, Park DJ, Kim HH. Laparoscopic double-tract proximal gastrectomy for proximal early gastric cancer. Gastric Cancer 2014;17(3):562-570. [doi: 10.1007/s10120-013-0303-5] [Medline: 24052482]
13. Huang QZ, Wang PC, Chen YX, Lin S, Ye K. Comparison of proximal gastrectomy with double-flap technique and double-tract reconstruction for proximal early gastric cancer: a meta-analysis. Updates Surg 2023 Dec;75(8):2117-2126. [doi: 10.1007/s13304-023-01638-w] [Medline: 37728858]

XSL•FO
RenderX

14. Isobe T, Hashimoto K, Kizaki J, et al. Reconstruction methods and complications in proximal gastrectomy for gastric cancer, and a comparison with total gastrectomy. Kurume Med J 2014;61(1-2):23-29. [doi: 10.2739/kurumemedj.MS64003] [Medline: 25152248]
15. Saze Z, Kase K, Nakano H, et al. Functional benefits of the double flap technique after proximal gastrectomy for gastric cancer. BMC Surg 2021 Nov 5;21(1):392. [doi: 10.1186/s12893-021-01390-1] [Medline: 34740344]
16. Wang S, Lin S, Wang H, et al. Reconstruction methods after radical proximal gastrectomy: a systematic review. Medicine (Baltimore) 2018;97(11):e0121. [doi: 10.1097/MD.0000000000010121] [Medline: 29538208]
17. Ying K, Bai W, Yan G, Xu Z, Du S, Dang C. The comparison of long-term oncological outcomes and complications after proximal gastrectomy with double tract reconstruction versus total gastrectomy for proximal gastric cancer. World J Surg Oncol 2023 Mar 23;21(1):101. [doi: 10.1186/s12957-023-02985-z] [Medline: 36949503]
18. Peng R, Shi Y, Zhang H, et al. Tunnel anastomosis: a modified flap technique in esophagogastrostomy as a novel antireflux technique after proximal gastrectomy. J Gastrointest Surg 2025 Jan;29(1):101871. [doi: 10.1016/j.gassur.2024.10.026] [Medline: 39490562]
19. Tang HP, Zhu HX, Lu GP, Peng ZQ, Chen ZK, Wang MC. Technical variety of anastomotic techniques used in proximal gastrectomy with double-tract-reconstruction - a narrative review. Langenbecks Arch Surg 2024 May 2;409(1):148. [doi: 10.1007/s00423-024-03339-3] [Medline: 38695994]
20. Miwa H, Yokoyama T, Hori K, et al. Interobserver agreement in endoscopic evaluation of reflux esophagitis using a modified Los Angeles classification incorporating grades N and M: a validation study in a cohort of Japanese endoscopists. Dis Esophagus 2008;21(4):355-363. [doi: 10.1111/j.1442-2050.2007.00788.x] [Medline: 18477259]
21. Muraoka A, Kobayashi M, Kokudo Y. Laparoscopy-assisted proximal gastrectomy with the hinged double flap method. World J Surg 2016 Oct;40(10):2419-2424. [doi: 10.1007/s00268-016-3510-5] [Medline: 27094564]
22. Sun KK, Wu YY. Current status of laparoscopic proximal gastrectomy in proximal gastric cancer: technical details and oncologic outcomes. Asian J Surg 2021 Jan;44(1):54-58. [doi: 10.1016/j.asjsur.2020.09.006] [Medline: 32981822]
23. Li B, Wang Y, Li B, Shan F, Li Z. Short-term outcomes and long-term quality of life of reconstruction methods after proximal gastrectomy: a systematic review and meta-analysis. BMC Cancer 2024 Jan 10;24(1):56. [doi: 10.1186/s12885-024-11827-4] [Medline: 38200411]
24. Shaibu Z, Chen Z, Mzee SA, Theophilus A, Danbala IA. Effects of reconstruction techniques after proximal gastrectomy: a systematic review and meta-analysis. World J Surg Oncol 2020 Jul 16;18(1):171. [doi: 10.1186/s12957-020-01936-2] [Medline: 32677956]
25. Ahn SH, Lee JH, Park DJ, Kim HH. Comparative study of clinical outcomes between laparoscopy-assisted proximal gastrectomy (LAPG) and laparoscopy-assisted total gastrectomy (LATG) for proximal gastric cancer. Gastric Cancer 2013 Jul;16(3):282-289. [doi: 10.1007/s10120-012-0178-x] [Medline: 22821182]

Abbreviations

DTJIR: double-tract jejunal interposition reconstruction
EORTC QLQ-30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
LA: Los Angeles
PNI: prognostic nutritional index
TA: tunnel anastomosis

Edited by J Sarvestan; submitted 20.Aug.2025; peer-reviewed by F Ma, J Hong; accepted 05.Jan.2026; published 27.Mar.2026.

Please cite as:
Xie Q, Peng R, Yue C, Wei W, Huang L, Wang X, Wang H, Chen L, Gu R, Chen H, Ming X, Wen X, Xu W, Sun G, Fan H, Wang Z, Yang L, Zhou X, Wu X, Zhou J, Wang M, Xu H, Li G
Tunnel Anastomosis vs Double-Tract Jejunal Interposition Reconstruction After Proximal Gastrectomy: Protocol for a Multicenter Prospective Randomized Controlled Trial
JMIR Res Protoc 2026;15:e82712
URL: https://www.researchprotocols.org/2026/1/e82712
doi:10.2196/82712

© Qingyu Xie, Rui Peng, Chao Yue, Wei Wei, Lingli Huang, Xiaoxiao Wang, Haitian Wang, Liang Chen, Rongmin Gu, Huanqiu Chen, Xuezhi Ming, Xu Wen, Weiguo Xu, Guangli Sun, Hao Fan, Zhe Wang, Longhao Yang, Xiaohua Zhou, Xiaoyu Wu, Jin Zhou, Meng Wang, Hao Xu, Gang Li. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License

XSL•FO
RenderX

(https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Enhanced Family Education on BMI Changes in Children and Adolescents With Overweight or Obesity: Study Protocol for a City-Wide Cluster Randomized Controlled Trial

Rui Hang Zhang[1,2*], BMed; Ting Yu Lu[1,2*], BMed; Si Han Hou[1,2*], BMed; Ya Bin Qu[3], MPhil; Qiu Xia Chen[3], MPhil; Meng Li[3], MPhil; Jiao Wang[1,2], PhD; Lin Xu[1,2,4,5], PhD

[1]School of Public Health, Sun Yat-sen University, Rm 103, School of Public Health Building, No. 74 Zhongshan 2nd Road, Yuexiu District, Guangzhou, China

[2]Greater Bay Area Public Health Research Collaboration, Guangzhou, China

[3]Guangdong Provincial Center for Disease Control and Prevention, Guangzhou, China

[4]Department of Applied Health Sciences, University of Birmingham, Birmingham, United Kingdom

[5]School of Public Health, University of Hong Kong, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Lin Xu, PhD
School of Public Health, Sun Yat-sen University, Rm 103, School of Public Health Building, No. 74 Zhongshan 2nd Road, Yuexiu District, Guangzhou, China

Abstract

Background: The prevalence of obesity among children and adolescents has become a critical global public health issue, particularly in China, with significant increases observed over the past few decades. Despite regular surveillance for overweight and obesity in the past decades, there have been no annual reports provided to parents, resulting in a significant underrecognition of this issue. Early intervention and enhanced family-based educational intervention are necessary to address this growing problem.

Objective: This protocol aims to evaluate the effectiveness of an enhanced family-based educational intervention in reducing overweight and obesity among school-age children and adolescents.

Methods: This study uses a stratified cluster randomized controlled trial design involving 40 schools from 4 cities in Guangdong Province, selected based on economic levels and overweight and obesity prevalence. Schools will be randomly assigned to intervention or control groups. The intervention group receives an enhanced family-based educational intervention on obesity prevention, including health reports and educational materials. The control group continues with regular practices, including routine physical examinations, general health education activities, and school-based physical activity programs. The intervention spans 9 months, followed by a 3-month follow-up. Data on BMI, waist circumference, and waist-to-height ratio are collected at baseline and after 12 months to assess the potential effectiveness of the intervention.

Results: This trial was funded in January 2024 and registered in the Chinese Clinical Trial Registry on November 5, 2024. Recruitment was completed in December 2024, with 20 schools enrolled in each of the intervention and control groups. Baseline data collection was completed during the 2024 fall semester, and follow-up data collection continued through December 2025. Data analysis will start after completion of the 12-month assessments. The study findings are expected to be published in 2026.

Conclusions: This study protocol addresses the urgent need for effective interventions to combat the increasing prevalence of childhood and adolescent obesity in Guangdong Province. Given the significant underrecognition of the need for consistent reporting to parents, the proposed enhanced family-based educational intervention aims to fill this gap by raising parental awareness and promoting healthier lifestyles among children and adolescents. If successful, this approach could significantly reduce the risk of overweight and obesity in Chinese populations, which account for approximately one-fifth of the world's population. The findings will provide insights into the efficacy of family-centered interventions and underscore the importance of integrating routine parental reporting into existing surveillance programs.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400091909; https://www.chictr.org.cn/showprojEN.html?proj=239755

International Registered Report Identifier (IRRID): DERR1-10.2196/86508

(JMIR Res Protoc 2026;15:e86508) doi:10.2196/86508

XSL•FO
RenderX

KEYWORDS

enhanced family education; obesity; school; China; randomized controlled trial

Introduction

The global prevalence of obesity among children and adolescents has increased nearly 10-fold over the past 4 decades, from 0.7% in 1975 to 5.6% in 2016 among girls and from 0.9% to 7.8% among boys in the same period [1,2]. This prevalence continues to rise, with more than 390 million children and adolescents aged 5 to 19 years being overweight, including 160 million who are living with obesity as of 2022 [3]. While the prevalence of obesity has plateaued at high levels in high-income countries, it is still accelerating in low- and middle-income countries, including China [2,4]. The mean prevalence of overweight and obesity among Chinese children and adolescents increased significantly from 5.3% in 1995 to 20.5% in 2014 [5], and more recent data reported by the latest round of China Chronic Disease and Nutrition Surveillance (2015-2019) indicated that the prevalence among children and adolescents aged 6 to 17 years was approximately 19% for overweight and obesity as of 2019 [6]. Childhood and adolescent obesity can affect both physical and psychological health [3], academic attainment, and quality of life [7,8]. In the longer term, it can also increase the risk of cardiometabolic diseases [9], musculoskeletal problems [10], and cancers in adulthood [11].

Despite the alarming rise in childhood and adolescent obesity, many Chinese parents and caregivers hold the belief that a chubby child is a sign of good health and nutrition, reflecting social and cultural factors where a fuller body is seen as a sign of prosperity and well-being [12]. This perception is compounded by the common belief that children who are overweight will naturally shed the excess weight as they grow older [13]. However, without proper intervention, children who are overweight are more likely to become adults who are overweight or obese, carrying the associated health risks into later life [14,15]. The issue is further complicated by the misconception that obesity primarily affects adults, leading to insufficient attention to the dietary and physical activity needs for children and adolescents [16,17]. Many parents are not fully aware of the serious health risks linked to childhood and adolescent obesity, such as diabetes, hypertension, and psychological issues [18-20]. This lack of awareness contributes to a lack of urgency in addressing and managing weight issues among young people [19,20].

Existing literature underscores the need for early intervention to curb the trajectory of obesity from childhood into adulthood [15,20,21]. Studies have shown that interventions targeting children and adolescents can effectively reduce BMI and improve motor skills [22-24].

Despite regular surveillance and monitoring of overweight and obesity in China, there has been a significant underrecognition of these issues due to the absence of consistent reporting to parents. Furthermore, evidence suggests that interventions with significant family involvement achieve better outcomes in reducing childhood and adolescent obesity [25]. However, the effectiveness of enhanced family-based educational interventions aimed at children and adolescents with overweight or obesity has not yet been fully evaluated. Notably, most studies evaluating obesity interventions in China have not adequately measured the degree of family involvement during implementation [26-32]. Furthermore, these studies were confined to single, high-income cities, such as Shanghai [31], Nanjing [28,32], Guangzhou [30], and Beijing [29], raising concerns about their generalizability to broader, more diverse populations across the country. Therefore, these gaps highlight the need for enhanced family-based educational interventions to raise awareness and promote healthier lifestyles among children and adolescents, together with their families.

This study aims to evaluate the effectiveness of enhanced family-based educational interventions in reducing the prevalence of overweight and obesity among school-aged children and adolescents in Guangdong Province. By integrating routine parental reporting into existing surveillance programs and providing comprehensive family-centered interventions, this study seeks to address the critical gaps in awareness and management of childhood and adolescent overweight and obesity. The findings will contribute valuable insights into effective strategies for obesity prevention and control, ultimately improving the health outcomes for children and adolescents in China.

Methods

Study Design

This study will use a multistage cluster randomized controlled trial (RCT) design. All cities in Guangdong Province will be divided into 2 economic strata. The 2 cities with the highest prevalence of overweight or obesity will be selected from each stratum. Schools within these cities will be randomly assigned to either the intervention group or the control group. The intervention group will receive enhanced family-based educational interventions, while the control group will continue with regular practices. The study will involve baseline data collection, 4 rounds of interventions, and follow-up assessments over 12 months. Figure 1 shows the flowchart of the study design.

XSL•FO
RenderX

Figure 1. Overview of the study design. KAP: knowledge, attitude, and practice.

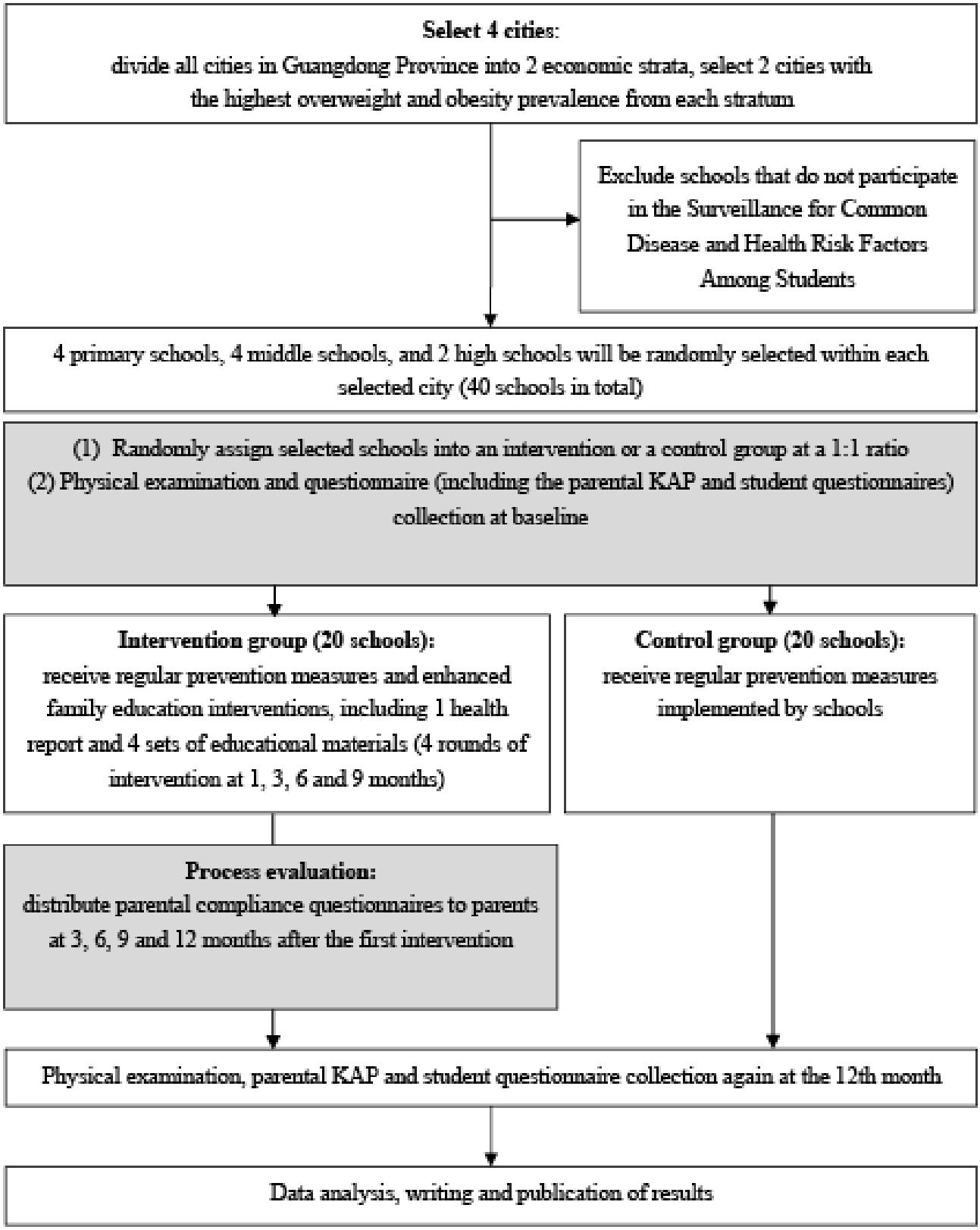

Setting

The study will be conducted in Guangdong Province, which includes 21 prefecture-level cities. Within these cities, a total of 40 schools (including primary, middle, and high schools) will be selected based on their participation in an initiative called Surveillance for Common Disease and Health Risk Factors Among Students in Guangdong Province. The selected schools will represent a diverse socioeconomic and demographic student population to ensure the generalizability of the study findings. The intervention will be implemented in both urban and rural settings to capture the varying dynamics of childhood and adolescent obesity across different environments.

XSL•FO
RenderX

Stratification, Randomization, and Allocation Concealment

Stratified randomization will be used, considering the comparability between the intervention and control groups. First, all cities in Guangdong Province will be divided into 2 economic strata. The 2 cities with the highest prevalence of overweight or obesity will be selected from each stratum. Second, schools will be stratified according to school types, with 4 primary schools, 4 middle schools, and 2 high schools being selected in each city. Subsequently, the selected schools within each stratum will be randomly assigned in a 1:1 ratio to either the intervention or control group. To ensure rigorous allocation concealment, a random allocation sequence will be generated by an independent statistician blinded to the participating schools. Opaque, sealed envelopes containing group assignments will be prepared by staff not involved in intervention implementation, and allocations will be disclosed only after completion of school enrollment.

Inclusion and Exclusion Criteria

Participants in this study will be school-aged children and adolescents from Guangdong Province who meet the following criteria: aged between 8 and 17 years and enrolled in grades 4 and 5 (primary school), grades 7 and 8 (middle school), and grades 10 and 11 (high school). They must be identified as overweight or obese based on the Screening for Overweight and Obesity Among School-Age Children and Adolescents (WS/T 586 - 2018) standards during the baseline assessment. Additionally, both students and their parents or guardians must provide informed consent to participate in the study.

Participants will be excluded if they meet any of the following criteria: refusal by students or their parents or guardians to participate in the study; presence of mental health conditions, intellectual disabilities, or other factors that may impede understanding or cooperation with the study procedures; and presence of severe illnesses or undergoing major medical treatments, such as dialysis or chemotherapy, which may affect the study outcomes.

Sample Size

Assuming an effect size (Cohen *d*) of 0.1, an SD of 0.05, a significance level (α) of .05, and a power (1-β) of .9, with an intracluster correlation coefficient of 0.01, an average cluster size (*m*) of 160 students, and an overweight and obesity prevalence rate of 20%, the required sample size *n* for each cluster (school) was calculated using the following Cohen *d* formula:

$$(1)n=(z1-a2+z1-\beta d)2\times 2\times (SD)2$$

Considering that the sample size for a cluster RCT needs to be adjusted based on the intracluster correlation coefficient, the adjusted sample size n was calculated. By multiplying the adjusted sample size *n* by the number of clusters (schools), the total sample size was estimated to be 539 students. Accounting for a 20% dropout rate, the total sample size required is approximately 674 students.

Intervention and Control

The intervention group will receive an enhanced family-based educational intervention in addition to the regular prevention measures implemented by schools. Each student and their parents will receive 1 health report (Multimedia Appendix 1) and 4 sets of educational materials, which will be developed by the Guangdong Provincial Center for Disease Control and Prevention and delivered by schools. Each set includes an enhanced health education booklet, a family weight management action guideline, and an advocacy letter (Multimedia Appendix 2). Details of these intervention components are shown in Table 1. Participants in the control group will only receive the regular prevention measures implemented by their schools. These measures may include routine physical examinations, general health education activities, and school-based physical activity programs, but no additional educational materials or personalized health reports will be provided to the control group.

XSL•FO
RenderX

Table . Description of the enhanced family-based educational intervention components.

Intervention component	When it was delivered	What was delivered
Health report	Within 1 month of the baseline survey (ie, the first round of the intervention)	Information on the student's height, weight, waist circumference, BMI, and their overweight or obesity status
Advocacy letter	Within 1 month of the baseline survey (ie, the first round of the intervention), and at 3, 6, and 9 months after the first intervention	Information on the impact of being overweight and obese on the physical and brain development of students and the “Three 1/2s” initiative: increasing the daily intake of fresh fruits and vegetables by half, reducing the daily intake of sugary drinks and fried foods by half, and increasing daily outdoor physical activity duration by half
Enhanced health education booklet	Within 1 month of the baseline survey (ie, the first round of the intervention), and at 3, 6, and 9 months after the first intervention	Information on the short-term and long-term health risks associated with childhood and adolescent overweight and obesity, healthy dietary habits to prevent obesity, physical activity recommendations, BMI calculation methods, and screening standards for childhood and adolescent overweight and obesity
Family weight management action guideline	Within 1 month of the baseline survey (ie, the first round of the intervention), and at 3, 6, and 9 months after the first intervention	Details of the implementation of the “Three 1/2s” initiative, methods for monitoring children and adolescents' growth, and monthly checklists for tracking adherence to the intervention

Study Outcomes

Primary outcomes are the changes in BMI and standardized BMI (zBMI) over a 12-month period. These outcomes will provide a direct measure of the intervention’s effectiveness in reducing overweight and obesity among the participants. The change in BMI will be assessed from baseline to the end of the study, while the change in zBMI will account for age and sex differences, offering a more nuanced understanding of the weight status changes in the school-aged children and adolescents.

Secondary outcomes include the incidence and recovery rates of both general and central obesity; changes in standardized waist-to-height ratio (zWHtR) and standardized waist circumference (zWC); and changes in parental knowledge, attitude, and practice (KAP) scores (Multimedia Appendix 3) [33]. Specifically, the incidence of general obesity will be determined by newly diagnosed students with obesity at the end of the study among participants who are not obese at baseline, and the recovery rate will be calculated based on the number of students who transition from obesity to nonobesity status. Similarly, the incidence and recovery rates of central obesity will be assessed. Additional secondary outcomes involve changes in zWHtR and zWC over 12 months. The staff who measure the children’s weight and height will be blinded to the school allocation. Furthermore, KAP scores will be assessed at the start of the intervention and at 3, 6, 9, and 12 months after the intervention to gauge improvements in health literacy regarding childhood and adolescent overweight and obesity. These comprehensive outcomes will help elucidate the effectiveness of the enhanced family-based educational interventions and their impact on health behaviors and literacy.

Process evaluation will be used to evaluate the feasibility and acceptability of the trial by assessing parental compliance. Specifically, parent questionnaires (Multimedia Appendix 4) are administered at 3, 6, 9, and 12 months after the intervention to evaluate adherence to dietary and physical activity recommendations.

Data Collection

Data collection will include both questionnaire surveys and anthropometric measurements, carried out at the beginning and end of the study. The 2024 Student Health Status and Influencing Factors Survey, integrated with the provincial Surveillance for Common Disease and Health Risk Factors Among Students project, will be administered before the study begins and after it concludes. Both students and their parents will complete the survey, which is provided by national authorities and includes versions for primary and secondary school students. The survey covers general demographic characteristics, dietary habits, physical activity, smoking and drinking behaviors, use of eyes outside of school, electronic screen use, outdoor activities, and sleep habits. Additionally, parental KAP questionnaires for the intervention group will be collected at baseline and 3, 6, 9, and 12 months, while those for the control group will be obtained at baseline and after 12 months. Nutritional status monitoring will include measurements of height, weight, and waist circumference following standardized procedures with light clothing and without shoes. BMI will be calculated as weight in kilograms divided by height in meters squared (kg/m^2). Waist circumference will be measured horizontally at the midpoint between the last rib and the top of the iliac crest. General obesity will be determined by sex- and age-specific BMI cutoff values according to the Screening Standard for Overweight and Obesity Among School-Age Children and Adolescents (WS/T 586-2018). Central obesity will be based on the 90th percentile of waist circumference by sex and age according to the High Waist Circumference Screening Threshold for Children and Adolescents Aged 7-18 Years (WS/T 611-2018). Data collected from paper-based questionnaires will be independently

XSL•FO
RenderX

double-entered by 2 trained university students and subsequently entered into an electronic database.

Statistical Analysis

The collected data will be analyzed to evaluate the effectiveness of the enhanced family-based educational interventions. Descriptive statistics will summarize baseline characteristics of the intervention and control groups, with means (SDs) or medians (IQRs) for continuous variables and frequencies (percentages) for categorical variables. Comparisons between groups will be made using independent sample 2-tailed *t* tests or Wilcoxon rank-sum tests for continuous variables and chi-square tests or Fisher exact tests for categorical variables. Multivariable regression analyses will adjust for potential confounders and assess the intervention's independent effects on primary and secondary outcomes, including changes in BMI, zBMI, zWHtR, and zWC over 12 months. Parental KAP scores will be the sum of scores in each questionnaire entry, which will be scored on a scale of 1 to 4 depending on the option. KAP scores from the intervention group will be assessed every 3 months, and the difference between KAP scores at the start and the end of the study will be used to evaluate changes in health literacy.

To address the handling of missing data in our study, we will use a combination of intention-to-treat, per-protocol, and complete case analyses. The intention-to-treat analysis will include all participants as randomized to account for nonadherence and dropouts, thereby preserving the randomization's integrity. The per-protocol analysis will focus on participants who complete the intervention according to the study protocol, offering insights into the efficacy under ideal adherence. Additionally, we will conduct a complete case analysis, which considers only those participants with complete data on all variables of interest. This approach will allow us to maintain the rigor of our statistical evaluation while acknowledging the limitations associated with missing data.

Ethical Considerations

This study protocol was reviewed and approved by the institutional review board of the School of Public Health, Sun Yat-sen University (approval number 2024-160). Informed consent will be obtained from all participants and their parents or guardians prior to enrollment in the study. All procedures will be conducted in accordance with the ethical standards of the responsible committee and the World Medical Association Declaration of Helsinki. Participant confidentiality will be ensured by assigning unique study IDs, and privacy will be strictly protected. Any participant who experiences harm related to trial participation will receive appropriate medical care and compensation.

Results

This trial was funded in January 2024. The study recruitment is scheduled to be completed before December 2024. The intervention will be integrated with the 2024 Surveillance for Common Disease and Health Risk Factors Among Students in Guangdong Province initiative. Baseline data will be collected during the 2024 and 2025 fall semester. Within 1 month, the local Center for Disease Control and Prevention will provide the intervention group with health reports and educational materials (including an enhanced health education booklet, a family weight management action guideline, and an advocacy letter). Educational materials will also be sent to the intervention group at 3, 6, and 9 months after the first intervention. Health reports are provided at the first intervention. The control group will continue with regular school-based prevention measures. The total intervention duration lasts 9 months, with 4 rounds of intervention. Table 2 shows the timeline of the study implementation. Data collection is scheduled to conclude in December 2025. Data analysis will begin after completion of the 12-month assessments. At the time of manuscript submission, baseline data collection was complete, with 20 schools enrolled in each of the intervention and control groups. Descriptive and analytic results are expected to be published in 2026.

XSL•FO
RenderX

Table . Timeline of the study implementation.

Task	November-December 2024		January-December 2025											
	November	December	January	February	March	April	May	June	July	August	September	October	November	December
Ethics approval and clinical trial registration	✓[a]													
Sampled school and class determination	✓[a]													
Staff training	✓[a]													
Student survey questionnaire distribution		✓[b]												✓[b]
Parental KAP[c] questionnaire distribution		✓[b]			✓[d]			✓[d]			✓[d]			✓[b]
Student health check-up		✓[b]												✓[b]
Health report distribution		✓[d]												
Enhanced health education booklet distribution		✓[d]			✓[d]			✓[d]			✓[d]			
Advocacy letter distribution		✓[d]			✓[d]			✓[d]			✓[d]			

XSL•FO
RenderX

Task	November-December 2024		January-December 2025										
	November	December	January	February	March	April	May	June	July	August	September	October	November	December
Family weight management action guideline distribution		✓[d]			✓[d]			✓[d]			✓[d]			
Parental compliance questionnaire distribution					✓[d]			✓[d]			✓[d]			✓[d]

[a]Preliminary preparation work.

[b]Tasks for all participants.

[c]KAP: knowledge, attitude, and practice.

[d]Tasks only for the intervention group.

Discussion

Hypothesized Findings

The enhanced family-based educational intervention described here is anticipated to be effective for obesity control, reducing the prevalence of overweight and obesity among children and adolescents in Guangdong Province. We hypothesize that, compared with the control group, the intervention group will show significantly more favorable changes in BMI, zBMI, zWHtR, and zWC over the 12-month period. We also anticipate a significantly lower incidence of general obesity and a higher recovery rate in the intervention group. Furthermore, the intervention is expected to improve parental health literacy. This study integrates surveillance feedback with an enhanced family-based educational intervention, aiming to generate critical evidence to inform obesity prevention and control policies in China.

Comparison to Prior Work

Previous intervention studies with family involvement have been predominantly conducted in single, high-income cities [28-32], limiting the generalizability of their findings to populations with different economic levels in China. Furthermore, these studies overlooked standardized family guidelines and the timely feedback of surveillance results to parents [34], neglecting elements essential for sustained engagement. To address these gaps, this study uses a multicity cluster RCT design across cities of varying economic levels. Additionally, we will provide structured educational materials, including the actionable “Three 1/2s” initiatives and child health reports for parents, aiming to facilitate practical implementation and enhance family engagement.

Strengths and Limitations

This study has some strengths, including randomly selecting 4 cities from 2 economic strata to enhance representativeness. Additionally, schools will serve as the unit of intervention, rather than individuals, to minimize contamination between groups and improve implementation feasibility. Several limitations should be acknowledged. First, although a relatively short follow-up period may limit the detection of significant intervention effects on the control of overweight or obesity, we implemented 4 rounds of enhanced intervention, with continuous process evaluations and monitoring of parental health literacy. Second, although the study protocol does not include tailored, culturally adaptive materials, analyzing the data for differential effects across socioeconomic subgroups may provide insights to inform future refinements of intervention strategies. Finally, the anthropometric measurements will be performed only at baseline and at the end of the study based on the provincial Surveillance for Common Disease and Health Risk Factors Among Students project. The 2-point measurements restrict our ability to capture the dynamic trajectories of nutritional status during the study. Future studies incorporating more detailed and frequent measurements are warranted to better characterize the temporal trajectories of the intervention effects.

Future Directions

This study aims to assess the effectiveness of a cost-effective family-based educational intervention to address childhood obesity in China. If proven effective, this framework could provide a scalable solution for nationwide implementation. Future research may explore tailored strategies across diverse socioeconomic and urban-rural contexts to enhance intervention effectiveness. In addition, identifying the optimal frequency of parental reporting would be valuable for maximizing long-term adherence.

XSL•FO
RenderX

Dissemination Plan

Study results will be published in peer-reviewed journals and presented at conferences and relevant stakeholder engagement activities. No information that can identify participants will be included in any publication. Participants who explicitly express a wish to be informed about the study outcome will be contacted and provided with an article or poster containing a lay summary of findings.

Conclusions

This cluster RCT will evaluate the effectiveness of an enhanced family-based educational intervention in children and adolescents with overweight or obesity. By integrating parental reporting into the existing provincial health surveillance system, the study aims to bridge the critical gap between school-based screening and family-based management, addressing widespread underrecognition of childhood obesity. The findings will provide evidence on the efficacy and feasibility of this scalable, family-based model. If proven effective, this approach could be incorporated into national surveillance programs, offering a sustainable public health strategy to curb the growing obesity epidemic.

Acknowledgments

The authors would like to thank the participating schools, students, and their families for their cooperation. They also thank the Guangdong Provincial Center for Disease Control and Prevention for their support in facilitating this study.

Funding

This study is funded by the Department of Education of Guangdong Province (grant 51000-42240006). The funding body had no role in the design of the study; the collection, analysis, and interpretation of data; and the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study will be available from the corresponding author on reasonable request.

Authors' Contributions

RHZ, TYL, SHH, YBQ, QXC, ML, JW, and LX have made substantial contributions to conception and design, acquisition of funding, and interpretation of data. RHZ, TYL, SHH, JW, and LX designed the research plan. RHZ, TYL, and SHH drafted the paper. JW and LX revised it critically for important intellectual content. All authors read the final manuscript and accepted responsibility to submit for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Overweight and obesity screening report.
[DOCX File, 20 KB - resprot_v15i1e86508_app1.docx]

Multimedia Appendix 2
“Wise drinking and smart movement” advocacy.
[DOCX File, 15 KB - resprot_v15i1e86508_app2.docx]

Multimedia Appendix 3
Parental knowledge, attitude, and practice survey.
[DOCX File, 18 KB - resprot_v15i1e86508_app3.docx]

Multimedia Appendix 4
Parental compliance evaluation questionnaire.
[DOCX File, 17 KB - resprot_v15i1e86508_app4.docx]

Checklist 1
SPIRIT 2025 checklist.
[PDF File, 244 KB - resprot_v15i1e86508_app5.pdf]

References

XSL•FO
RenderX

1. Tenfold increase in childhood and adolescent obesity in four decades: new study by Imperial College London and WHO. : World Health Organization; 2017 Oct 11 URL: https://www.who.int/news/item/11-10-2017-tenfold-increase-in-childhood-and-adolescent-obesity-in-four-decades-new-study-by-imperial-college-london-and-who [accessed 2026-02-23]
2. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in body-mass index, underweight, overweight, and obesity from 1975 to 2016: a pooled analysis of 2416 population-based measurement studies in 128·9 million children, adolescents, and adults. Lancet 2017 Dec 16;390(10113):2627-2642. [doi: 10.1016/S0140-6736(17)32129-3] [Medline: 29029897]
3. Obesity and overweight. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2026-02-23]
4. Zhang XY, Yang WY, Yang J, et al. Epidemiological characteristics of childhood obesity in eastern China, 2017-2019. World J Pediatr 2019 Aug;15(4):412-414. [doi: 10.1007/s12519-019-00298-4] [Medline: 31407142]
5. Dong Y, Jan C, Ma Y, et al. Economic development and the nutritional status of Chinese school-aged children and adolescents from 1995 to 2014: an analysis of five successive national surveys. Lancet Diabetes Endocrinol 2019 Apr;7(4):288-299. [doi: 10.1016/S2213-8587(19)30075-0] [Medline: 30902266]
6. Press briefing for the report on Chinese residents' chronic diseases and nutrition. The State Council Information Office of the People's Republic of China. 2020 Apr. URL: https://www.sohu.com/a/441742585_120056612 [accessed 2026-03-05]
7. Diao H, Wang H, Yang L, Li T. The impacts of multiple obesity-related interventions on quality of life in children and adolescents: a randomized controlled trial. Health Qual Life Outcomes 2020 Jul 6;18(1):213. [doi: 10.1186/s12955-020-01459-0] [Medline: 32631401]
8. Kang NR, Kwack YS. An update on mental health problems and cognitive behavioral therapy in pediatric obesity. Pediatr Gastroenterol Hepatol Nutr 2020 Jan;23(1):15-25. [doi: 10.5223/pghn.2020.23.1.15] [Medline: 31988872]
9. Salama M, Balagopal B, Fennoy I, Kumar S. Childhood obesity, diabetes and cardiovascular disease risk. J Clin Endocrinol Metab 2023 Nov 17;108(12):3051-3066. [doi: 10.1210/clinem/dgad361] [Medline: 37319430]
10. Firman N, Wilk M, Harper G, Dezateux C. Are children with obesity at school entry more likely to have a diagnosis of a musculoskeletal condition? Findings from a systematic review. BMJ Paediatr Open 2022 Aug;6(1):e001528. [doi: 10.1136/bmjpo-2022-001528] [Medline: 36053659]
11. Weihe P, Spielmann J, Kielstein H, Henning-Klusmann J, Weihrauch-Blüher S. Childhood obesity and cancer risk in adulthood. Curr Obes Rep 2020 Sep;9(3):204-212. [doi: 10.1007/s13679-020-00387-w] [Medline: 32519271]
12. Wu Y, Ma X, Fraser WD, et al. Caregivers' perceptions, challenges and service needs related to tackling childhood overweight and obesity: a qualitative study in three districts of Shanghai, China. BMC Public Health 2021 Apr 21;21(1):768. [doi: 10.1186/s12889-021-10744-6] [Medline: 33882878]
13. Schwartz MB, Puhl R. Childhood obesity: a societal problem to solve. Obes Rev 2003 Feb;4(1):57-71. [doi: 10.1046/j.1467-789x.2003.00093.x] [Medline: 12608527]
14. Singh AS, Mulder C, Twisk JWR, van Mechelen W, Chinapaw MJM. Tracking of childhood overweight into adulthood: a systematic review of the literature. Obes Rev 2008 Sep;9(5):474-488. [doi: 10.1111/j.1467-789X.2008.00475.x] [Medline: 18331423]
15. Rundle AG, Factor-Litvak P, Suglia SF, et al. Tracking of obesity in childhood into adulthood: effects on body mass index and fat mass index at age 50. Child Obes 2020 Apr;16(3):226-233. [doi: 10.1089/chi.2019.0185] [Medline: 32191541]
16. Haqq AM, Kebbe M, Tan Q, Manco M, Salas XR. Complexity and stigma of pediatric obesity. Child Obes 2021 Jun;17(4):229-240. [doi: 10.1089/chi.2021.0003] [Medline: 33780639]
17. Abbas N, Rouaiheb H, Saliba J, El-Bikai R. Childhood obesity: facts and parental perceptions. World Acad Sci J 2023;5(6):38. [doi: 10.3892/wasj.2023.215]
18. Bass R, Eneli I. Severe childhood obesity: an under-recognised and growing health problem. Postgrad Med J 2015 Nov;91(1081):639-645. [doi: 10.1136/postgradmedj-2014-133033] [Medline: 26338983]
19. Etelson D, Brand DA, Patrick PA, Shirali A. Childhood obesity: do parents recognize this health risk? Obes Res 2003 Nov;11(11):1362-1368. [doi: 10.1038/oby.2003.184] [Medline: 14627757]
20. Golan M. Parents as agents of change in childhood obesity--from research to practice. Int J Pediatr Obes 2006;1(2):66-76. [doi: 10.1080/17477160600644272] [Medline: 17907317]
21. Simmonds M, Burch J, Llewellyn A, et al. The use of measures of obesity in childhood for predicting obesity and the development of obesity-related diseases in adulthood: a systematic review and meta-analysis. Health Technol Assess 2015 Jun;19(43):1-336. [doi: 10.3310/hta19430] [Medline: 26108433]
22. Jacob CM, Hardy-Johnson PL, Inskip HM, et al. A systematic review and meta-analysis of school-based interventions with health education to reduce body mass index in adolescents aged 10 to 19 years. Int J Behav Nutr Phys Act 2021 Jan 4;18(1):1. [doi: 10.1186/s12966-020-01065-9] [Medline: 33397403]
23. Ells LJ, Rees K, Brown T, et al. Interventions for treating children and adolescents with overweight and obesity: an overview of Cochrane reviews. Int J Obes (Lond) 2018 Nov;42(11):1823-1833. [doi: 10.1038/s41366-018-0230-y] [Medline: 30301964]

XSL•FO
RenderX

24. Han A, Fu A, Cobley S, Sanders RH. Effectiveness of exercise intervention on improving fundamental movement skills and motor coordination in overweight/obese children and adolescents: a systematic review. J Sci Med Sport 2018 Jan;21(1):89-102. [doi: 10.1016/j.jsams.2017.07.001] [Medline: 28728887]
25. Kobes A, Kretschmer T, Timmerman G, Schreuder P. Interventions aimed at preventing and reducing overweight/obesity among children and adolescents: a meta-synthesis. Obes Rev 2018 Aug;19(8):1065-1079. [doi: 10.1111/obr.12688] [Medline: 29671938]
26. Liu Z, Gao P, Gao AY, et al. Effectiveness of a multifaceted intervention for prevention of obesity in primary school children in China: a cluster randomized clinical trial. JAMA Pediatr 2022 Jan 1;176(1):e214375. [doi: 10.1001/jamapediatrics.2021.4375] [Medline: 34747972]
27. Zhou HR, Wang WW, Yang HF, Qi SX, Wang CC, Hong X. The effects of a one year school-based comprehensive lifestyle intervention among Chinese kids against obesity (CLICK-obesity) in Nanjing, China. J Public Health (Berl) 2023 Jul;31(7):1071-1079. [doi: 10.1007/s10389-021-01646-3]
28. Wang Z, Xu F, Ye Q, et al. Childhood obesity prevention through a community-based cluster randomized controlled physical activity intervention among schools in China: the health legacy project of the 2nd World Summer Youth Olympic Games (YOG-Obesity study). Int J Obes (Lond) 2018 Apr;42(4):625-633. [doi: 10.1038/ijo.2017.243] [Medline: 28978975]
29. Liu Z, Li Q, Maddison R, et al. A school-based comprehensive intervention for childhood obesity in China: a cluster randomized controlled trial. Child Obes 2019;15(2):105-115. [doi: 10.1089/chi.2018.0251] [Medline: 30565955]
30. Li B, Pallan M, Liu WJ, et al. The CHIRPY DRAGON intervention in preventing obesity in Chinese primary-school--aged children: A cluster-randomised controlled trial. PLoS Med 2019 Nov;16(11):e1002971. [doi: 10.1371/journal.pmed.1002971] [Medline: 31770371]
31. Cao ZJ, Wang SM, Chen Y. A randomized trial of multiple interventions for childhood obesity in China. Am J Prev Med 2015 May;48(5):552-560. [doi: 10.1016/j.amepre.2014.12.014] [Medline: 25891054]
32. Xu F, Ware RS, Leslie E, et al. Effectiveness of a randomized controlled lifestyle intervention to prevent obesity among Chinese primary school students: CLICK-Obesity study. PLoS ONE 2015 Oct 28;10(10):e0141421. [doi: 10.1371/journal.pone.0141421] [Medline: 26510135]
33. Ma J. Development and evaluation of the parental health literacy scale for childhood obesity [thesis]. : Anhui Medical University; 2018 URL: https://kns.cnki.net/kcms2/article/
[accessed 2026-03-05]
34. Zhang T, Cai L, Jing J, Ma L, Ma J, Chen Y. Parental perception of child weight and its association with weight-related parenting behaviours and child behaviours: a Chinese national study. Public Health Nutr 2018 Jun;21(9):1671-1680. [doi: 10.1017/S136898001800006X] [Medline: 29510769]

Abbreviations

KAP: knowledge, attitude, and practice
RCT: randomized controlled trial
zBMI: standardized BMI
zWC: standardized waist circumference
zWHtR: standardized waist-to-height ratio

Edited by J Sarvestan; submitted 25.Oct.2025; peer-reviewed by J Kain; accepted 05.Feb.2026; published 26.Mar.2026.

Please cite as:
Zhang RH, Lu TY, Hou SH, Qu YB, Chen QX, Li M, Wang J, Xu L
Impact of Enhanced Family Education on BMI Changes in Children and Adolescents With Overweight or Obesity: Study Protocol for a City-Wide Cluster Randomized Controlled Trial
JMIR Res Protoc 2026;15:e86508
URL: https://www.researchprotocols.org/2026/1/e86508
doi:10.2196/86508

© Rui Hang Zhang, Ting Yu Lu, Si Han Hou, Ya Bin Qu, Qiu Xia Chen, Meng Li, Jiao Wang, Lin Xu. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Sitting Baduanjin on Cancer-Related Fatigue in Patients With Advanced Cancer: Protocol for a Randomized Controlled Trial

Qinghui Zhang[1], MSc; Xi Chen[2], PhD; Yinglong Duan[3], MSc; Jianfei Xie[3], PhD; Yongyi Chen[1], PhD; Yonghong Hu[1], MSc; Qinqin Cheng[1], PhD

[1]Hunan Cancer Hospital, No. 283, Yuelu District, Changsha, Hunan, China

[2]The Hong Kong Polytechnic University, Hong Kong, China

[3]The Third Xiangya Hospital, Central South University, Changsha, China

Corresponding Author:
Qinqin Cheng, PhD
Hunan Cancer Hospital, No. 283, Yuelu District, Changsha, Hunan, China

Abstract

Background: Approximately 60% of patients with advanced cancer experience the distress of cancer-related fatigue (CRF), which significantly worsens their daily function and quality of life. Baduanjin has been regarded as a promising nonpharmacological intervention for alleviating CRF, anxiety, and depression and improving quality of life. Owing to varying degrees of CRF in patients with advanced cancer, patients may have insufficient endurance, making it difficult to implement standing Baduanjin. However, relevant evidence on sitting Baduanjin for CRF in patients with advanced cancer is lacking.

Objective: This study aims to design a sitting Baduanjin intervention and explore the efficacy and safety of sitting Baduanjin in reducing CRF among patients with advanced cancer.

Methods: This study will be a single-blind pilot randomized controlled trial. Patients with CRF will be enlisted from a tertiary cancer hospital in China. The participants (N=98) will be randomly assigned to either a sitting Baduanjin group or a control group at a 1:1 ratio using a block-randomized scheme. The participants in the sitting Baduanjin group will undergo a 16-week sitting Baduanjin intervention in addition to standard care, whereas the participants in the control group will receive standard care in the form of a booklet on the self-management of cancer symptoms. CRF will constitute the primary outcome, whereas anxiety, depression, and quality of life will serve as the secondary outcomes. These outcomes will be assessed at baseline (T0), 12 weeks (T1), and 16 weeks (T2).

Results: The study was approved by the Medical Ethics Review Committee of Hunan Cancer Hospital in March 2024 (ethics approval number 2024.50). Before this full-scale randomized controlled trial, we conducted a pilot study, which demonstrated the feasibility and acceptability of sitting Baduanjin for patients with advanced cancer, with potential benefits for relieving fatigue.

Conclusions: This randomized trial will evaluate the effectiveness of sitting Baduanjin exercises in alleviating CRF among patients with advanced cancer. If proven effective, it will provide a promising alternative intervention for patients with advanced cancer.

Trial Registration: Chinese Clinical Trial Register ChiCTR-2400092148; https://www.chictr.org.cn/showproj.html?proj=242990

International Registered Report Identifier (IRRID): IRRID: DERR1-10.2196/84925

(JMIR Res Protoc 2026;15:e84925) doi:10.2196/84925

KEYWORDS

sitting Baduanjin; patients with advanced cancer; cancer-related fatigue; quality of life; study protocol

Introduction

Cancer-related fatigue (CRF) is a distressing, lasting, subjective state of physical, emotional, and/or cognitive tiredness that is not in proportion to recent activities. It is related to cancer or associated treatments and is frequently accompanied by dysfunction [1]. Approximately 60% of patients in the advanced stage of cancer experience CRF [2]. CRF is one of the most prevalent symptoms manifested by patients with advanced cancer, and it can considerably deteriorate their quality of life and daily functioning [3]. Furthermore, it constitutes one of the primary reasons for the disruption of treatment or the deterioration of the disease [4]. Therefore, relieving CRF for patients with advanced cancer is critical.

XSL•FO
RenderX

At present, there is no acknowledged standard for the treatment of CRF [5] as its pathogenesis remains unclear. It could be associated with proinflammatory cytokines, growth factors, modulation of the circadian rhythm, disruption of the adrenal axis, serotonin imbalance, afferent activation of the vagus nerve, or the generation or use of abnormal adenosine triphosphate [6]. Previous research has shown that aerobic exercise (including tai chi and yoga) might be an effective CRF intervention and prove to be more efficacious than existing pharmaceutical treatments in alleviating CRF [7,8]. The current guidelines for the treatment of CRF recommend exercise-based activities rather than pharmaceutical interventions [9]. Aerobic exercise mitigates CRF both during and after treatment, particularly for those with solid tumors [10], and contributes to alleviating the cognitive and physiological consequences of fatigue [11]. Despite the existing evidence on the positive effects of aerobic exercise, many patients with advanced cancer cannot adhere to or perform exercise while standing due to weakness [12]. Therefore, it is necessary to find an appropriate exercise for those patients with advanced cancer who need to exercise while sitting.

Baduanjin, a traditional Chinese mind-body workout consisting of just 8 simple movements, is a proper moderate-intensity exercise for CRF in the treatment of advanced cancer as it is easy to master, consumes less energy, and requires no specific equipment [13]. Previous research indicates that Baduanjin is capable of alleviating the fatigue symptoms of patients with diverse diseases [14,15]. From the viewpoint of traditional Chinese medicine, CRF belongs to the category of deficiency and is marked by insufficient visceral functions and deficiencies in *qi*, blood, *yin*, and *yang*. Baduanjin can relieve CRF by facilitating the smooth flow of meridians, enhancing *qi*, improving blood circulation, regulating internal organ functions, boosting the immune system, and eliminating harmful toxins [16,17].

However, the current research on Baduanjin for patients with advanced cancer mainly focuses on standing Baduanjin. Due to varying degrees of CRF in patients with advanced cancer, patients may have insufficient endurance, making it difficult to implement standing Baduanjin. Therefore, sitting Baduanjin is more suitable for these patients. Currently, there is little research on sitting Baduanjin. One study developed a sitting Baduanjin intervention for patients with sepsis undergoing noninvasive ventilation [18]. The findings suggest that sitting Baduanjin could improve patients' muscle strength and activities of daily living [18]. Another study developed a sitting Baduanjin intervention for older adults and found that it improved their balance and motor function [19]. Previously, the evidence on the benefits of Baduanjin mainly came from studies on patients with breast cancer and other populations. Before this trial, the efficacy of Baduanjin in relieving fatigue in patients with advanced cancer had not been confirmed. Due to the differences in symptom burden, there is currently no direct empirical verification of the impact of the previously developed sitting Baduanjin interventions on patients with advanced cancer. This emphasizes a significant need to investigate the effects of performing sitting Baduanjin on CRF in patients with advanced cancer.

Methods

Study Design

This research is a single-blinded, parallel randomized controlled trial (RCT) with 2 arms. An overall diagram of the research is presented in Figure 1. The schedule of the study adheres to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [20].

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram of the study.

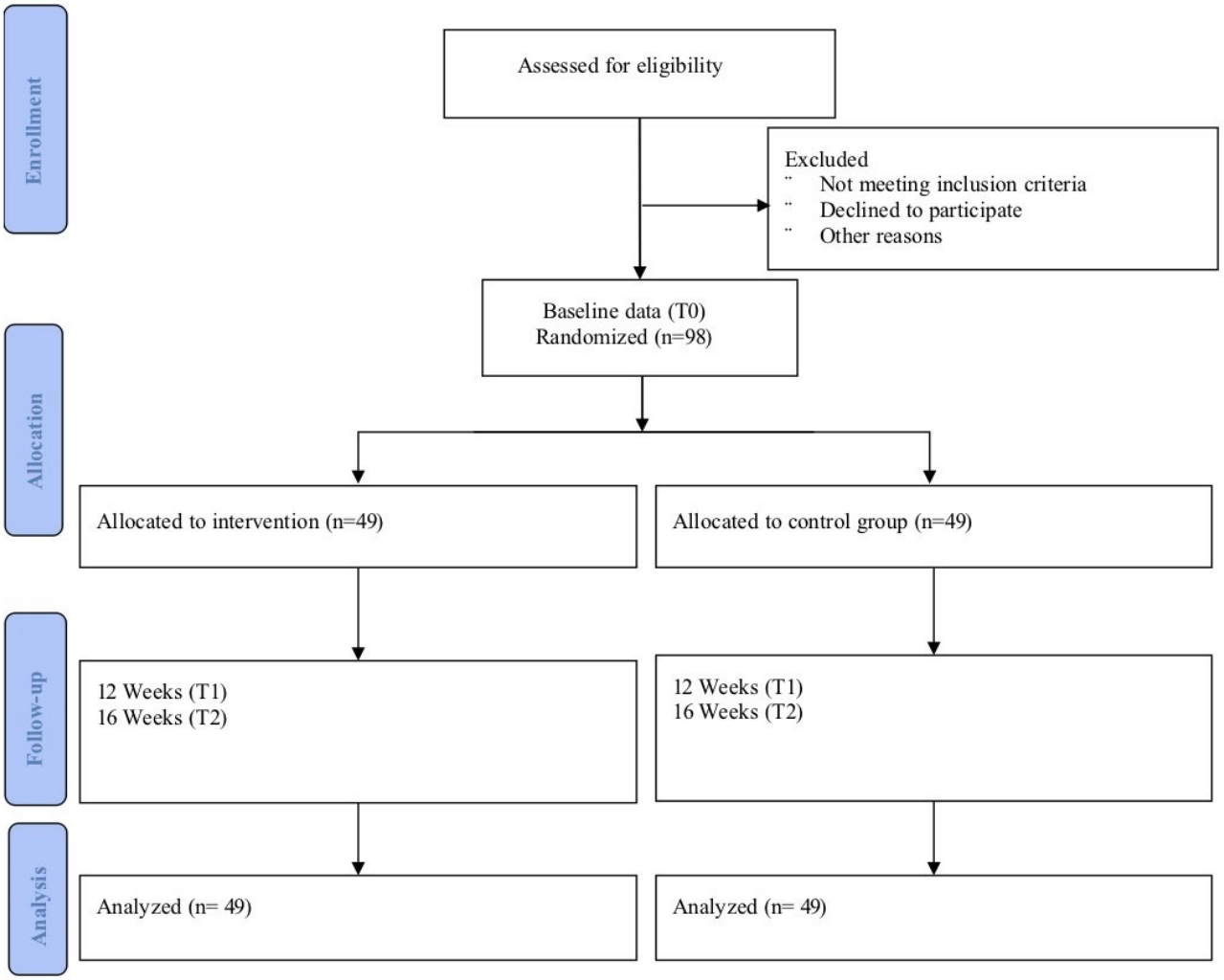

Study Setting and Participants

Participants will be enlisted from a tertiary cancer hospital in Hunan, China. Patients with cancer will be eligible for enlistment if they meet the following criteria: (1) diagnosis of stage III or IV cancer; (2) age of >18 years; (3) being informed of their illness; (4) capability of communicating and understanding the research; (5) willingness and ability to provide written informed consent for study participation; and (6) Revised Piper Fatigue Scale (RPFS) score of ≥4, indicating moderate to severe fatigue. Patients will be excluded in case they meet any of the following exclusion criteria: (1) currently using medications for fatigue, sleep disturbance, or depression, such as antidepressants, psychostimulants, or hypnotics; (2) being too frail to participate in physical activities due to terminal stages of chronic illnesses or having cognitive impairment and/or severe mental illness; (3) having serious exercise contraindications, such as cardiovascular or joint problems; and (4) participating in other exercise research programs.

Sample Size Calculation

The sample size was computed using the G*Power software (version 3.10). No previous research reports have emerged regarding the effect of sitting Baduanjin on CRF in patients with advanced cancer. Therefore, the effect value was set to 0.50. In the G*Power software, the *t* test for the mean difference between 2 independent samples was selected. With an effect size of 0.69 [21], a statistical power of 0.85, and a 2-tailed test, the sample size for each group was calculated to be 39. Thus, considering a conservative anticipation of a 20% dropout rate, the final sample size will be 49 in each group, with a total of 98 participants.

Recruitment

Patients with advanced cancer will be approached and recruited in a tertiary cancer hospital by a research nurse. This research nurse will screen patients with advanced cancer according to the inclusion and exclusion criteria. Information on this study, encompassing the purpose, the intervention, and the possible benefits and harms, will then be offered to the potential participants to obtain their informed consent. In addition, recruitment posters will be displayed in the study hospital to expand recruitment. After the participants' consent is obtained, the researchers will invite the participants to provide baseline data using paper questionnaires.

Randomization and Allocation Concealment

The participants will be randomly allocated into 2 groups in a 1:1 ratio using a block-randomized scheme with block sizes of 4 or 6. One set of randomization sequences will be generated

XSL•FO
RenderX

through an online randomizer [22] based on the estimated sample size. To guarantee allocation concealment, the randomization sequences will be produced by a research assistant who will not be involved in any other part of this study. The randomization sequences will then be put into sequential, sealed, opaque envelopes and kept by another research assistant who is not involved in the study. The randomization sequences will be inaccessible to other research staff. Once an eligible participant gives consent to take part in the study and finishes the baseline assessment, the research nurse in charge of recruitment will call the research assistant to disclose the allocation result.

Blinding

Owing to the nature of the sitting Baduanjin intervention, it will be impossible to blind the study investigators and participants. Consequently, blinding will only be implemented for the outcome assessors in this RCT to prevent potential detection bias during data collection. The outcome assessors will not be involved in the participant recruitment process. To ensure objective and unbiased results, the people in charge of data collection and analysis will not be informed of the group details.

Experimental Group

The participants in the experimental group will receive the sitting Baduanjin program. The sitting Baduanjin intervention was developed according to the characteristics of CRF under the guidance of the *zangfu* theory of traditional Chinese medicine and the *qi*-blood meridian theory. In addition, we improved the original standing Baduanjin workout by integrating relevant movements. This plan was revised through 2 rounds of expert consultation, resulting in the improved sitting Baduanjin intervention.

The sitting Baduanjin workout consists of 8 movements, which involve the upper limbs, head and neck, and spine, as well as organized breathing. A preparation and closing posture are added before and after the 8 movements to prepare the patient and ensure their safety [23]. The movements of the sitting Baduanjin workout were tested among patients with advanced cancer for preliminary feasibility and safety. The detailed movements of sitting Baduanjin are shown in Table 1.

XSL•FO
RenderX

Table . The movements of sitting Baduanjin.

Movement	Main points of the movement
Preparation posture	• The patient sits on a chair with their legs naturally open or on the bed. • With the body straight and facing the front, the patient feels calm, at ease, and levitated at the Baihui acupoint. The tongue is pressed against the upper jaw. • The patient puts their palms in front of the abdomen, closes their eyes, and breathes slightly and peacefully.
First movement: “hands in the sky to support Sanjiao”	• When inhaling, the patient lifts their hands from their chest to the top of their head. • When exhaling, the patient drops their hands from both sides of their body to their waist and holds them in front of their abdomen.
Second movement: “drawing the bow to shoot the hawk”	• The patient reaches their arms out to the sides with their palms down as if they were holding a bow and arrow. They inhale and then exhale as they draw their arms back, bending their elbows and bringing their hands toward their chest. • They hold for a moment and then release by straightening their arms and returning to the starting position.
Third movement: “lifting the arms to regulate the spleen and stomach”	• The patient raises the left palm and rotates it through the chest, then brings it upward in front of their face, followed by an inward rotation of the arms to lift it to the upper left side of the head. Simultaneously, they rotate the right palm inward and press it down beside the right hip with their fingertips facing forward. • The patient then drops their left hand, lifts their right hand simultaneously, and holds their 2 hands in front of their abdomen. • They repeat the same movement on the opposite side.
Fourth movement: “looking backward to relieve the fatigues and 7 injuries”	• The patient rotates both arms outward with the palms facing outward, turns their head to the left and back, pauses for a moment, and looks diagonally behind to the left. • They rotate both arms inward, press them against their hips, move their fingers forward, and look straight ahead. • They repeat the same movement on the opposite side.
Fifth movement: “clearing the heart fire by shaking the head and wagging the tail”	• The patient lifts both palms in front of the abdomen and chest, extending them above their head with their fingers facing each other. • They lower their arms to both sides, placing their palms on top of their thighs. • They raise the center of gravity slightly, then shift it to the right, tilt their upper body to the right, bend down, and visually observe the surface of their right foot. • They shift the center of gravity to the left while rotating their upper body from right to front and left, visually observing their right heel. • They shift the center of gravity to the right, shake their head backward, slightly retract their jaw, and look ahead. • The right-side action is the same as the left-side action but in the opposite direction.
Sixth movement: “move 2 hands down to the leg to strengthen the kidney and lumbar area”	• The patient raises both arms forward and upward, with their palms facing forward. • They press their palms against their chest with the elbows bent, palms facing downward, and fingers facing each other. • They rotate both arms outward with their palms facing upward and then insert the palm fingers under the armpits. • They rub their palms inward along the spine to the kidney area and then rub them back and forth. • They return their hands to the front of the abdomen through the body.

XSL•FO
RenderX

Movement	Main points of the movement
Seventh movement: “tighten fists to gather strength”	• The patient holds both fists at the waist, with the thumb inside and the fist facing upward, and looks straight ahead. • They inhale and hold their fists, exhale and clench their left fist with a furious gaze, keep their eyes wide open, grip firmly, inhale, and retract. • They exhale and clench their right fist with a furious gaze, keep their eyes wide open, grip firmly, inhale, and retract.
Eighth movement: “rotate shoulders to enhance *yang* and *qi*”	• The patient lowers both arms naturally, contracts the abdomen, and expands the chest. • They inhale to lift their shoulders and rotate them inward and exhale to lower their shoulders. • They inhale to lift their shoulders and rotate them outward and exhale to lower their shoulders.
Closing posture	• The patient wipes their face, combs their hair, rubs their ears, and taps their left and right elbow creases.

The program lasts for a total of 16 weeks [12]. In the first week, the participants will attend a training session to learn about sitting Baduanjin by watching a training video, which explains the exercise methods and precautions of sitting Baduanjin. A trained research nurse will then guide the participants to complete sitting Baduanjin according to the video and adjust incorrect movements individually, if any. The training session will include no more than 8 patients. From week 2 to week 16, the participants will perform sitting Baduanjin exercises by themselves once daily for approximately 20 to 30 minutes, including 5 minutes of preparation posture, 12 minutes of sitting Baduanjin, and 5 minutes of closing posture.

To enhance participant compliance, we have established a WeChat group for the experimental group to provide remote supervision. Videos of the sitting Baduanjin exercises and relevant theoretical knowledge will be pushed to patients via WeChat. Participants in the experimental group are required to practice daily, take photos as proof, and upload them via a check-in mini program. This process serves to document their practice duration and facilitates self-monitoring. Their caregivers are also involved to encourage participants to persist in their exercises. During home care, the main responsibility of the caregiver is to remind the participants to do the Baduanjin exercises on time and submit proof of having done the exercises via WeChat. We conduct follow-up visits and phone calls with patients twice a week via WeChat video to inquire about factors that may hinder their activities and resolve them promptly. In addition, participants in the experimental group are requested not to discuss their practice with each other to minimize interference.

Control Group

Patients randomized to the control group are provided with usual care. They will receive regular health education on exercise and a healthy lifestyle, including seated stretching and breathing exercises. After discharge, they will be provided with material on how to exercise. Participants will be asked to avoid any other traditional Chinese sports activities during the study, with reminders provided at every assessment interval. Considering ethical aspects, once the effectiveness of sitting Baduanjin is established, we will provide it for free to the control group if they are interested.

Outcome Measurements

Primary Outcomes

The primary outcome of this RCT is CRF. The severity of CRF in the daily functioning of the participants will be measured by using the Chinese version of the RPFS. The original scale was developed by Piper [24] in 1989, with its revised Chinese version developed by So et al [25] in 2003. The questionnaire includes 22 items that assess the general level of fatigue using a visual analog scale with values from 0 to 10. It also assesses 4 facets of fatigue: severity, emotional, sensory, and cognitive. A score of 0 indicates no fatigue symptoms, mild fatigue is defined as scores between 1 and 3, moderate fatigue is defined as scores between 4 and 6, and severe fatigue is defined as scores between 7 and 10. The RPFS exhibits Cronbach α coefficients between 0.89 and 0.93 for each subscale and 0.91 for the overall scale [26].

Secondary Outcomes

The secondary outcomes include patients’ anxiety, depression, and quality of life.

Anxiety

Secondary outcomes encompass anxiety as gauged using the mean score on the 7-item Generalized Anxiety Disorder Scale (GAD-7) [27]. The GAD-7 was developed by Spitzer et al [28] in 2006 for measuring the level of anxiety in patients over the previous 2 weeks. The GAD-7 is a gold-standard measuring instrument for generalized anxiety disorder [27]. It is a rapid, user-friendly, compact, and self-executed screening and diagnostic instrument. The GAD-7 is computed by allocating scores of 0, 1, 2, and 3 to the response categories of “not at all,” “several days,” “more than half the days,” and “nearly every day,” respectively. The total score ranges from 0 to 21. Scores of 5, 10, and 15, respectively, denote cutoff points for mild, moderate, and severe anxiety. The Chinese version of the GAD-7 shows solid internal consistency reliability (Cronbach α ranging from 0.715 to 0.931), along with construct validity and convergent validity [29].

XSL•FO
RenderX

Depression

Depression will be gauged by means of the Patient Health Questionnaire–9 (PHQ-9) [30]. The PHQ-9 is a self-administered diagnostic tool for the severity of depression [30]. It is determined by attributing scores of 0, 1, 2, and 3 to the response categories of "not at all," "several days," "more than half the days," and "nearly every day," respectively. The overall score ranges between 0 and 27. Scores of 5, 10, 15, and 20 stand for the cutoff points for mild, moderate, moderately severe, and severe depression, respectively. The Chinese version of the PHQ-9 exhibits internal consistency reliability (with Cronbach α ranging from 0.729 to 0.891), along with construct validity and convergent validity [31].

Quality of Life

Quality of life will be evaluated using the Functional Assessment of Cancer Therapy–General (FACT-G) total score. The scale encompasses 27 items across 4 dimensions: physiological status, social or family status, and functional status, each containing 7 items, and emotional status, containing 6 items. Each item is scored from 0 to 4, with higher scores denoting better quality of life [32]. The FACT-G is accessible in a simplified Chinese version, and adequate psychometric properties have been reported among patients with advanced cancer. The Chinese version of the FACT-G exhibits excellent internal consistency reliability (with Cronbach α from 0.89 to 0.98), along with construct validity and convergent validity [33].

Other Measurements

In addition, patient demographics (including gender, age, educational background, marital status, employment status, and household income) and clinical characteristics (such as diagnosis, stage, course of disease, and current treatment) will be collected at baseline.

Data Collection

The baseline assessments will be conducted by the research nurse who is in charge of participant recruitment and random assignment to groups. All participants will be requested to complete the RPFS, GAD-7, PHQ-9, FACT-G, and demographic data sheet on-site at baseline (T0). The reassessment of the outcomes will be conducted at 12 (T1) and 16 (T2) weeks after baseline. Follow-up data will be gathered via telephone by a research assistant who will be blinded to the participants' group allocation. The schedule of study outcome assessments is outlined in Table 2.

XSL•FO
RenderX

Table . SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) diagram of enrollment, treatment, and assessments over the intervention period.

	Study period				
	Enrollment	Baseline assessment	Allocation	Intervention week 12 (T1)	Intervention week 16 (T2)
Recruitment					
Eligibility screen	✓				
Informed consent	✓				
Randomization					
Allocation			✓		
Interventions					
Experimental group					
Control group					
Patient-reported outcomes					
Demographic characteristics		✓			
RPFS[a]		✓		✓	✓
GAD-7[b]		✓		✓	✓
PHQ-9[c]		✓		✓	✓
FACT-G[d]		✓		✓	✓
Adverse effects or events				✓	✓
Adherence to exercise					

[a]RPFS: Revised Piper Fatigue Scale.

[b]GAD-7: 7-item Generalized Anxiety Disorder Scale.

[c]PHQ-9: Patient Health Questionnaire–9.

[d]FACT-G: Functional Assessment of Cancer Therapy–General.

Data Analysis

Statistical analyses will be carried out using SPSS Statistics for Windows (version 23.0; IBM Corp). Demographic data will be summed and presented using appropriate descriptive statistics, such as frequency distributions and means and SDs. The homogeneity of baseline characteristics between the 2 groups will be evaluated using the independent 2-tailed *t* test, chi-square test, or Wilcoxon rank sum test as appropriate. A generalized estimating equation (GEE) model will be implemented for repeated multivariate analysis between the 2 study groups regarding the total scores and domain scores on the RPFS, GAD-7, PHQ-9, and FACT-G. The analysis will be conducted in accordance with the intention-to-treat principle. In case of missing data at random, the missingness will be left unaddressed because the GEE model could yield valid results through analyzing the observed data [34]. Otherwise, multiple imputation will be used for the missing data. Sensitivity analyses will be conducted by comparing results across standard GEE and multiple imputation–based GEE. The significance level for identifying statistical differences will be set at $P<.05$.

Ethical Considerations

The study will comply with the principles of the Declaration of Helsinki and the International Code of Ethics for Biomedical Research Involving Human Subjects. The study was approved by the Medical Ethics Review Committee of Hunan Cancer Hospital in March 2024 (ethics approval number 2024.50) and registered in the Chinese Clinical Trial Register (ChiCTR-2400092148). Written consent will be obtained from each participant before data collection. Participation in the study will not cause any harm to the patients. The patients will be notified that their participation is voluntary and that they have the right to withdraw from the study at any time without providing any explanation or facing any penalty. Data collection from the participants will adhere to the principles of confidentiality and anonymity. In addition, small gifts will be given to patients (eg, towels and toothbrushes worth

XSL•FO
RenderX

approximately ¥20 [US $2.87]) to thank them for the time taken to fill out the questionnaires at each assessment. The records and all information gathered during the study will be stored in a locked cabinet. Electronic data will be stored in a password-protected storage device. Only the researchers will be granted access to the data, which will be merely used for research purposes. The data will be deleted 5 years after the completion of the study.

Results

The study was approved by the Medical Ethics Review Committee of Hunan Cancer Hospital in March 2024 (ethics approval number 2024.50). Before this full-scale RCT, we conducted a pilot study. A total of 24 participants were recruited between March 2025 and April 2025. After the intervention, 18 participants, comprising 9 from the intervention group and 9 from the control group, completed the follow-up evaluation. The retention rate was 75%. The pilot study demonstrated the feasibility and acceptability of sitting Baduanjin for patients with advanced cancer, with potential benefits for relieving fatigue.

Discussion

Hypothesized Findings

Sitting Baduanjin helps control breathing, align the body, and soothe the mind while softly adjusting the functions of internal organs [35]. Therefore, it is supposed to relieve patients' CRF, anxiety, and depression and improve their quality of life. As there is limited research on sitting Baduanjin, this study will contribute to the knowledge on sitting Baduanjin in traditional Chinese medicine. A meticulously designed RCT will be used to evaluate the impacts of the sitting Baduanjin program on alleviating CRF among patients with advanced cancer. If the effects of the sitting Baduanjin program are positive, this study will offer an appropriate Chinese traditional exercise approach for patients with advanced cancer who have insufficient endurance for standing exercises. In addition, the results will enrich the scientific and practical knowledge of interventions for relieving CRF among patients with advanced cancer.

Sitting Baduanjin's effectiveness could be due to its unique features. First, the sitting Baduanjin program outlined in this protocol will be developed with consideration of the fatigue of patients with advanced cancer in accordance with the *zangfu* theory of traditional Chinese medicine and the *qi*-blood meridian theory. Second, patients with advanced cancer can easily access it. Sitting Baduanjin can be performed in a seated position, which suits their physical abilities and satisfies their exercise demands. Third, sitting Baduanjin involves meditation and deep breathing as a mind-body exercise, which might help decrease CRF [36]. Fourth, participants are encouraged by their caregivers to keep up with their exercises. Videos of sitting Baduanjin and theoretical knowledge were created to enhance patients' adherence to exercise. This study will use daily exercise logs via an app to monitor home practice and record patients' self-reported experiences with sitting Baduanjin. Furthermore, weekly conversations through WeChat or phone calls will be held to boost exercise self-efficacy.

This protocol has several potential limitations. First, because of the nature of the intervention, it will not be possible to blind the participants. This has the potential to introduce performance bias. Nevertheless, outcome assessors will be kept blind to minimize bias in the evaluation of the effects of the intervention. Second, due to the restricted study sites, the sample in this study might not be fully representative of patients with advanced cancer experiencing CRF. As patients will only be recruited from a tertiary cancer hospital, the results might not apply to all patients with advanced cancer.

Future Directions

There are several strategies that can be used to enhance the reliability and comprehensiveness of the research on the effects of sitting Baduanjin in patients with advanced cancer experiencing CRF. First, it is crucial to increase the sample size and study duration by carrying out multicenter studies with more participants and longer intervention periods. Second, enhancing the research design by applying more rigorous randomization and blinding techniques is essential for reducing bias and guaranteeing the validity of the findings. Third, enhancing intervention strategies by integrating nutritional advice, psychological support, and other pertinent techniques can further boost the intervention's effectiveness.

Funding

This protocol is funded by the Hunan Provincial Administration of Traditional Chinese Medicine Project (B2024108) and supported by High-Level Talent Support Program of Hunan Cancer Hospital (20250731-1053). The funding sources played no role in the design of the study.

Authors' Contributions

QC was accountable for the initial development and design of the study. XC, JX, and YD provided guidance on the selection of appropriate methodology design and statistical analysis approaches. QZ wrote the original draft of the manuscript. QC, YC, and YH made significant contributions to revising the draft. All authors collaboratively contributed to and approved the final manuscript. YH is the co-corresponding author and can be reached via email at 307929759@qq.com.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Schroder J, Mackenzie L. Outcomes related to activity performance and participation of non-pharmacological cancer-related fatigue interventions. OTJR (Thorofare N J) 2022 Jan;42(1):50-64. [doi: 10.1177/15394492211029214] [Medline: 34315289]
2. Mücke M, Mochamat, Cuhls H, et al. Pharmacological treatments for fatigue associated with palliative care. Cochrane Database Syst Rev 2015 May 30;2015(5):CD006788. [doi: 10.1002/14651858.CD006788.pub3] [Medline: 26026155]
3. Muthanna FM, Hassan BA, Karuppannan M, Ibrahim HK, Mohammed AH, Abdulrahman E. Prevalence and impact of fatigue on quality of life (QOL) of cancer patients undergoing chemotherapy: a systematic review and meta-analysis. Asian Pac J Cancer Prev 2023 Mar 1;24(3):769-781. [doi: 10.31557/APJCP.2023.24.3.769] [Medline: 36974528]
4. Bower JE. Cancer-related fatigue--mechanisms, risk factors, and treatments. Nat Rev Clin Oncol 2014 Oct;11(10):597-609. [doi: 10.1038/nrclinonc.2014.127] [Medline: 25113839]
5. Meneses-Echávez JF, González-Jiménez E, Ramírez-Vélez R. Effects of supervised multimodal exercise interventions on cancer-related fatigue: systematic review and meta-analysis of randomized controlled trials. Biomed Res Int 2015;2015:328636. [doi: 10.1155/2015/328636] [Medline: 26167483]
6. Sleight AG, Crowder SL, Skarbinski J, et al. A new approach to understanding cancer-related fatigue: leveraging the 3P model to facilitate risk prediction and clinical care. Cancers (Basel) 2022 Apr 14;14(8):1982. [doi: 10.3390/cancers14081982] [Medline: 35454890]
7. Mitchell SA, Hoffman AJ, Clark JC, et al. Putting evidence into practice: an update of evidence-based interventions for cancer-related fatigue during and following treatment. Clin J Oncol Nurs 2014;18 Suppl:38-58. [doi: 10.1188/14.CJON.S3.38-58] [Medline: 25427608]
8. Hilfiker R, Meichtry A, Eicher M, et al. Exercise and other non-pharmaceutical interventions for cancer-related fatigue in patients during or after cancer treatment: a systematic review incorporating an indirect-comparisons meta-analysis. Br J Sports Med 2018 May;52(10):651-658. [doi: 10.1136/bjsports-2016-096422] [Medline: 28501804]
9. Thong MS, van Noorden CJ, Steindorf K, Arndt V. Correction to: cancer-related fatigue: causes and current treatment options. Curr Treat Options Oncol 2022 Mar;23(3):450-451. [doi: 10.1007/s11864-021-00916-2] [Medline: 35230612]
10. Cramp F, Byron-Daniel J. Exercise for the management of cancer-related fatigue in adults. Cochrane Database Syst Rev 2012 Nov 14;11(11):CD006145. [doi: 10.1002/14651858.CD006145.pub3] [Medline: 23152233]
11. Mustian KM, Alfano CM, Heckler C, et al. Comparison of pharmaceutical, psychological, and exercise treatments for cancer-related fatigue: a meta-analysis. JAMA Oncol 2017 Jul 1;3(7):961-968. [doi: 10.1001/jamaoncol.2016.6914] [Medline: 28253393]
12. Campbell KL, Winters-Stone KM, Wiskemann J, et al. Exercise guidelines for cancer survivors: consensus statement from international multidisciplinary roundtable. Med Sci Sports Exerc 2019 Nov;51(11):2375-2390. [doi: 10.1249/MSS.0000000000002116] [Medline: 31626055]
13. Gong X, Rong G, Wang Z, Zhang A, Li X, Wang L. Baduanjin exercise for patients with breast cancer: a systematic review and meta-analysis. Complement Ther Med 2022 Dec;71:102886. [doi: 10.1016/j.ctim.2022.102886] [Medline: 36122861]
14. Liu H, Liu S, Xiong L, Luo B. Efficacy of Baduanjin for treatment of fatigue: a systematic review and meta-analysis of randomized controlled trials. Medicine (Baltimore) 2023 Aug 11;102(32):e34707. [doi: 10.1097/MD.0000000000034707] [Medline: 37565842]
15. Kuo CC, Wang CC, Chang WL, Liao TC, Chen PE, Tung TH. Clinical effects of Baduanjin Qigong exercise on cancer patients: a systematic review and meta-analysis on randomized controlled trials. Evid Based Complement Alternat Med 2021;2021:6651238. [doi: 10.1155/2021/6651238] [Medline: 33880125]
16. Ye J, Simpson MW, Liu Y, et al. The effects of Baduanjin Qigong on postural stability, proprioception, and symptoms of patients with knee osteoarthritis: a randomized controlled trial. Front Med (Lausanne) 2019;6:307. [doi: 10.3389/fmed.2019.00307] [Medline: 31998728]
17. Ye J, Zheng Q, Zou L, et al. Mindful exercise (Baduanjin) as an adjuvant treatment for older adults (60 years old and over) of knee osteoarthritis: a randomized controlled trial. Evid Based Complement Alternat Med 2020 Jun 14;2020:9869161. [doi: 10.1155/2020/9869161] [Medline: 32617115]
18. Chen MG, Wang F, Huang L, et al. Effect of Sitting Baduanjin exercise on early rehabilitation of sepsis patients with non-invasive ventilation: a randomized controlled trial. BMC Complement Med Ther 2024 Sep 6;24(1):330. [doi: 10.1186/s12906-024-04626-8] [Medline: 39243078]
19. Bao X, Qiu QX, Shao YJ, Quiben M, Liu H. Effect of Sitting Ba-Duan-Jin exercises on balance and quality of life among older adults: a preliminary study. Rehabil Nurs 2020;45(5):271-278. [doi: 10.1097/rnj.0000000000000219] [Medline: 30730382]
20. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
21. Lu Y, Qu HQ, Chen FY, et al. Effect of Baduanjin Qigong exercise on cancer-related fatigue in patients with colorectal cancer undergoing chemotherapy: a randomized controlled trial. Oncol Res Treat 2019;42(9):431-439. [doi: 10.1159/000501127] [Medline: 31266043]
22. Research Randomizer. URL: https://www.randomizer.org/ [accessed 2026-02-02]

XSL•FO
RenderX

23. Rau KM, Shun SC, Hung SH, et al. Management of cancer-related fatigue in Taiwan: an evidence-based consensus for screening, assessment and treatment. Jpn J Clin Oncol 2023 Jan 6;53(1):46-56. [doi: 10.1093/jjco/hyac164] [Medline: 36354095]
24. Piper BF. Piper fatigue scale available for clinical testing. Oncol Nurs Forum 1990;17(5):661-662. [Medline: 2251183]
25. So WK, Dodgson J, Tai JW. Fatigue and quality of life among Chinese patients with hematologic malignancy after bone marrow transplantation. Cancer Nurs 2003 Jun;26(3):211-219. [doi: 10.1097/00002820-200306000-00006] [Medline: 12832954]
26. So WK, Tai JW. Fatigue and fatigue-relieving strategies used by Hong Kong Chinese patients after hemopoietic stem cell transplantation. Nurs Res 2005;54(1):48-55. [doi: 10.1097/00006199-200501000-00007] [Medline: 15695939]
27. Löwe B, Decker O, Müller S, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
28. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
29. Qingzhi Z, Yanling H, Han L, et al. Reliability and validity of Chinese version of the Generalized Anxiety Disorder 7-item (GAD-7) scale in screening anxiety disorders in outpatients from traditional Chinese internal department. Chin Ment Health J 2013;27(3):163-168 [FREE Full text]
30. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. Med Care 2004 Dec;42(12):1194-1201. [doi: 10.1097/00005650-200412000-00006] [Medline: 15550799]
31. Baoquan M, Aihong Z, Feng L, Jianping J. Clinical application of the Patient Health Questionnaire - 9 (PHQ - 9). J Neurol Ment Health 2013;13(6):569-572 [FREE Full text]
32. Victorson D, Barocas J, Song J, Cella D. Reliability across studies from the functional assessment of cancer therapy-general (FACT-G) and its subscales: a reliability generalization. Qual Life Res 2008 Nov;17(9):1137-1146. [doi: 10.1007/s11136-008-9398-2] [Medline: 18841493]
33. Lili Z, Mingqiang H, Zhanbin C, Feifei J, Junli Q. Applicability study of the Chinese version of the FACT - G scale in the quality - of - life evaluation of cancer patients. Chin J Health Stat 2022;39(2):243-245 [FREE Full text]
34. Seaman S, Copas A. Doubly robust generalized estimating equations for longitudinal data. Stat Med 2009 Mar 15;28(6):937-955. [doi: 10.1002/sim.3520] [Medline: 19153970]
35. Zhao W, Zhang S, Xie H, et al. Sitting Baduanjin combined with acupoint massage on frailty in patients undergoing maintenance hemodialysis: protocol for a randomized controlled trial. JMIR Res Protoc 2025 Nov 21;14:e72352. [doi: 10.2196/72352] [Medline: 41269893]
36. Liu C, Qin M, Zheng X, Chen R, Zhu J. A meta-analysis: intervention effect of mind-body exercise on relieving cancer-related fatigue in breast cancer patients. Evid Based Complement Alternat Med 2021;2021:9980940. [doi: 10.1155/2021/9980940] [Medline: 34285706]

Abbreviations

CRF: cancer-related fatigue
FACT-G: Functional Assessment of Cancer Therapy–General
GAD-7: 7-item Generalized Anxiety Disorder Scale
GEE: generalized estimating equation
PHQ-9: Patient Health Questionnaire–9
RCT: randomized controlled trial
RPFS: Revised Piper Fatigue Scale
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by J Sarvestan; submitted 27.Sep.2025; peer-reviewed by LM Wintner; revised version received 20.Dec.2025; accepted 07.Jan.2026; published 07.Apr.2026.

Please cite as:
Zhang Q, Chen X, Duan Y, Xie J, Chen Y, Hu Y, Cheng Q
Effects of Sitting Baduanjin on Cancer-Related Fatigue in Patients With Advanced Cancer: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84925
URL: https://www.researchprotocols.org/2026/1/e84925
doi:10.2196/84925

XSL•FO
RenderX

© Qinghui Zhang, Xi Chen, Yinglong Duan, Jianfei Xie, Jianfei Xie, Yongyi Chen, Yonghong Hu, Qinqin Cheng. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 7.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Treating Chronic Pain by Modulating Phenomenological and Psychophysiological Indices of Self-Transcendence: Single-Blind Randomized Controlled Clinical Trial Protocol

Kennedi Childs[1*], BS; Carter Minnick[1], BA; Geraldine Martorella[2], BSCN, MSN, PhD; Adam W Hanley[1*], PhD

[1]Complementary Health Innovation Lab, College of Nursing, Florida State University, 2010 Levy Ave Ste. B0217, Tallahassee, FL, United States

[2]Institute on Digital Health and Innovation, College of Nursing, Florida State University, Tallahassee, FL, United States

[*]these authors contributed equally

Corresponding Author:
Adam W Hanley, PhD
Complementary Health Innovation Lab, College of Nursing, Florida State University, 2010 Levy Ave Ste. B0217, Tallahassee, FL, United States

Abstract

Background: Chronic musculoskeletal pain (CMP) imposes a significant psychological, physical, and emotional burden on millions of adults in the United States. A challenge in treating CMP is its tendency to become enmeshed with a person's sense of self. Mindfulness-based interventions (MBIs) offer a promising approach to help patients disentangle pain from their sense of self. However, rigorous research is needed to determine which styles of mindfulness practice most effectively promote self-transcendence, a candidate mechanism for durable pain relief.

Objective: This study aims to evaluate whether adding direct pointing instruction to traditional mindful breathing enhances self-transcendence and pain relief. We will assess the effects of 2 mindfulness training protocols, mindful breathing (MB) and MB plus direct pointing (DP) instruction, on subjective and neurophysiological markers of self-transcendence and on pain-related outcomes.

Methods: In this study, we will conduct a 3-arm randomized controlled trial with 173 adults with CMP, assigned to one of the following groups: (1) 4, 30-minute sessions of traditional mindful breathing instruction, (2) 4, 30-minute sessions of MB+DP instruction, or (3) a waitlist control group. Self-transcendence will be assessed during the first and final intervention sessions using validated self-report measures and psychophysiological indices, including electroencephalography (EEG; theta activity) and functional near-infrared spectroscopy (fNIRS; default mode network [DMN] activity). Primary outcomes include changes in self-transcendence, theta activity, and DMN activity. Additional outcomes include acute pain intensity and pain-related functional interference, measured at baseline through 3-month follow-up.

Results: Data collection began in October 2024 and will conclude by May 2027. As this is a protocol paper, no results are yet available. We will report recruitment rates, baseline characteristics, and initial outcome trends in the full trial results paper.

Conclusions: This will be the first randomized controlled trial to examine whether direct pointing instruction enhances analgesia by eliciting self-transcendence. Findings may advance understanding of the mechanisms underlying mindfulness-based pain relief and inform the development of more targeted MBIs for chronic pain. Results will be disseminated through peer-reviewed publications, conference presentations, and stakeholder engagement, with potential implications for clinical practices.

Trial Registration: OSF Registries OSF.IO/VQFWA; https://doi.org/10.17605/OSF.IO/VQFWA

International Registered Report Identifier (IRRID): DERR1-10.2196/82362

(JMIR Res Protoc 2026;15:e82362) doi:10.2196/82362

KEYWORDS

chronic pain; mindfulness-based intervention; self-transcendence; pain; perception of pain

Introduction

Background

Chronic musculoskeletal pain (CMP) is among the most prevalent and disabling health conditions worldwide, affecting up to one-third of the global population and significantly diminishing quality of life and workplace productivity [1-3]. In the United States, CMP is a major contributor to health care expenditures and long-term disability [1,2,4]. Given the well-documented risks associated with pharmacologic treatments, such as gastrointestinal complications from nonsteroidal anti-inflammatory drugs and the potential for

XSL•FO
RenderX

addiction and overdose with opioid analgesics, guidelines from the American College of Physicians recommend nonpharmacologic treatments, such as massage, acupuncture, and mindfulness, as preferred first-line therapies for managing CMP [5]. Nonpharmacologic approaches not only offer a safer alternative to pharmacologic treatment but also more directly address the complex biopsychosocial factors that contribute to chronic pain [6]. Among the notable psychological consequences of chronic pain is its potential to alter an individual's sense of self [7-9]. Persistent pain may become enmeshed with personal identity through repeated and simultaneous activation of self- and pain-related schemas (eg, "I am [ie, self-schema] in pain [ie, pain schema]") [7]. Over time, this repeated activation leads to a pain-laden self-concept [8,10,11] which may be reflected in corresponding changes within the brain's default mode network (DMN) [12-14]. The DMN is a system involved in both self-representation and pain processing, potentially trapping individuals in a self-perpetuating "pain perception box" [7,10,11].

Mindfulness-based interventions (MBIs) offer a promising means of disentangling chronic pain from the sense of self, thereby disrupting the maladaptive cycle that sustains the pain individuals endure [9]. Rather than providing mere distraction or placebo effects, mindfulness engages specific neural and psychological mechanisms that fundamentally alter pain perception and appraisal [15-17]. Central to this process is a reduction in self-referential processing, a shift supported by neuroimaging evidence showing decoupling between the thalamus and DMN hubs such as the precuneus and medial prefrontal cortex during mindfulness practice [9,15,18,19]. This decoupling may represent a loosening of the neural circuits that enmesh pain in the sense of self, facilitating a more open, present-centered awareness [9,15,18,19]. The Mindfulness-to-Meaning Theory deepens this perspective by suggesting that mindfulness initiates a recursive cycle of positive psychological processes, including decentering, attentional broadening, cognitive reappraisal, and savoring, that synergistically promote self-transcendent states [9]. These transient states, characterized by diminished self-salience and expansive positive affect (eg, awe, bliss, and compassion), weaken the cognitive and emotional enmeshment between the individual and their pain experience [9,20]. Through momentary dissolution of the subject-object dichotomy, self-transcendence can reduce the centrality of pain in consciousness and cultivate a renewed sense of meaning [9,20]. In this way, mindfulness-induced self-transcendence may not only reshape the brain's functional architecture but may also reframe the lived experience of pain, offering a path toward relief that is both neurologically grounded and psychologically liberating [9].

Emerging empirical evidence supports this theoretical framework. Two randomized controlled trials (RCTs) have examined the therapeutic value of self-transcendent states elicited during brief (ie, 15-20-minute), single-session mindfulness-based interventions (MBIs) for adults undergoing knee or hip replacement surgery (n=196, n=118) [21]. In both RCTs, the MBIs incorporated direct pointing (DP), an ancient yet underexplored contemplative technique designed to evoke self-transcendence by guiding attention away from self-specifying information and toward the nature of awareness itself [22-24]. Specifically, DP instruction encourages deidentification with thoughts, emotions, sensations, and sensory input as a means of accessing a nondual form of awareness in which the usual subject-object distinction dissolves [22-24]. Across both RCTs, groups involving mindfulness paired with DP consistently evoked the highest levels of self-transcendence relative to the psychoeducation control group, which in turn predicted immediate pain relief and improved postoperative outcomes, including reduced pain intensity and interference during the first month after surgery and better physical functioning 6 weeks after surgery [21,25,26]. These findings highlight the potential of DP to enhance the analgesic effects of brief MBIs by targeting self-related mechanisms of pain. However, because all mindfulness groups in these RCTs included DP instruction, the specific contribution of this technique remains to be isolated.

The therapeutic value of self-transcendence has also been examined in 4 RCTs evaluating the 8-week Mindfulness-Oriented Recovery Enhancement (MORE) program as a treatment for adults with chronic pain receiving long-term opioid therapy (n=95, n=187, n=62, and n=165) [27-30]. MORE incorporates DP throughout its protocol and has been shown to increase self-transcendent experiences both in daily life [27,30] and during lab-based, self-guided mindfulness practice [28,29]. In these RCTs, self-transcendent experiences were associated with reductions in pain, opioid misuse, and opioid dosing out to 9-month follow-up, as well as increases in frontal midline theta activity, which is thought to be a neurophysiological marker of meditative absorption and self-dissolution [27-30]. Collectively, these findings suggest that self-transcendence, along with its neurophysiological correlates, may serve as a core mechanism through which mindfulness improves chronic pain-related outcomes. Yet, due to the multimodal nature of MORE, the specific effects of DP remain confounded and warrant further investigation.

Aims

To isolate the specific effect of DP to mindfulness-based pain relief, the Pointing Beyond project will randomize 173 adults with CMP into one of three groups: (1) 4 weekly 30-minute sessions of traditional mindful breathing (MB) instruction, (2) 4 weekly 30-minute sessions of MB combined with DP (MB+DP) instruction, or (3) a waitlist control group (WLC). The specific aims are:

Aim 1

Evaluate between-group differences in changes in self-transcendence from Session 1 to Session 4 across the MB, MB+DP, and WLC groups, using a validated self-report instrument (state version of the Nondual Awareness Dimensional Assessment [NADA-S]; primary outcome) and 2 psychophysiological indices (secondary outcomes), including frontal midline theta activity (via electroencephalography [EEG]) [28,29] and a functional near-infrared spectroscopy (fNIRS) [31] derived proxy default mode network (DMN) activity.

XSL•FO
RenderX

Aim 2

Examine the associations among phenomenological and psychophysiological indices of self-transcendence by testing within- and between-session correlations between self-reported self-transcendent state (NADA-S), frontal midline theta activity (EEG), and fNIRS-derived proxy DMN activity across intervention sessions and study groups.

Aim 3

Determine whether increases in self-transcendence mediate treatment-related improvements in pain outcomes by testing indirect effects of group assignment (MB, MB+DP, and WLC) on (1) acute pain intensity at the end of Session 4 (secondary outcome) [32] and (2) pain-related outcomes (eg, functional interference) through 3-month follow-up (tertiary outcomes) [33-38], via changes in (1) self-reported self-transcendence (NADA-S), (2) frontal midline theta activity (EEG), and (3) fNIRS-derived DMN activity measured during Session 4, controlling for presession pain and baseline levels of the mediators.

Hypotheses

The following are our hypotheses:

Aim 1

We hypothesize that, from Session 1 to Session 4, both mindfulness groups (MB and MB+DP) will show greater increases in self-transcendence, indexed by higher NADA-S scores and greater frontal midline theta activity (EEG), and greater decreases in DMN activity, indexed by reduced fNIRS-derived fronto-temporal-parietal connectivity, compared to WLC group. We further hypothesize that MB+DP will produce larger increases in these phenomenological and neurophysiological indices of self-transcendence, and larger reductions in DMN activity, relative to MB alone.

Aim 2

We hypothesize that phenomenological and psychophysiological indices of self-transcendence will be significantly and coherently related across sessions and groups, such that a higher self-reported self-transcendent state (NADA-S) will be associated with greater frontal midline theta activity (EEG) and lower fNIRS-derived proxy DMN activity within sessions and across Sessions 1-4. We further hypothesize that these associations will be the strongest in the MB+DP group, intermediate in the MB group, and the weakest or absent in the WLC group, reflecting graded coupling between subjective experience and neural indices of self-transcendence as a function of mindfulness instruction type.

Aim 3

We hypothesize that treatment-related improvements in pain outcomes will be mediated by increases in self-transcendence and its psychophysiological correlates. Specifically, we expect that assignment to the mindfulness groups, particularly MB+DP, will be associated with lower acute pain intensity at the end of Session 4 and reduced pain-related functional interference through 3-month follow-up via (1) increases in self-reported self-transcendent state (NADA-S), (2) increases in frontal midline theta activity (EEG), and (3) reductions in fNIRS-derived proxy DMN activity measured during Session 4, controlling for presession pain and baseline levels of each mediator [39,40].

Methods

Overview

This will be a single-blind, randomized, 3-arm, parallel-group, clinical trial (2:2:1 allocation ratio). The planned flow diagram for this single-site RCT is presented in Figure 1. The protocol is reported according to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials; Checklist 1) [41] checklist. This study was prospectively registered with the Open Science Framework on April 25, 2023.

XSL•FO
RenderX

Figure 1. Diagram of planned study flow. DP: direct pointing; MB: mindful breathing; WLC: waitlist control.

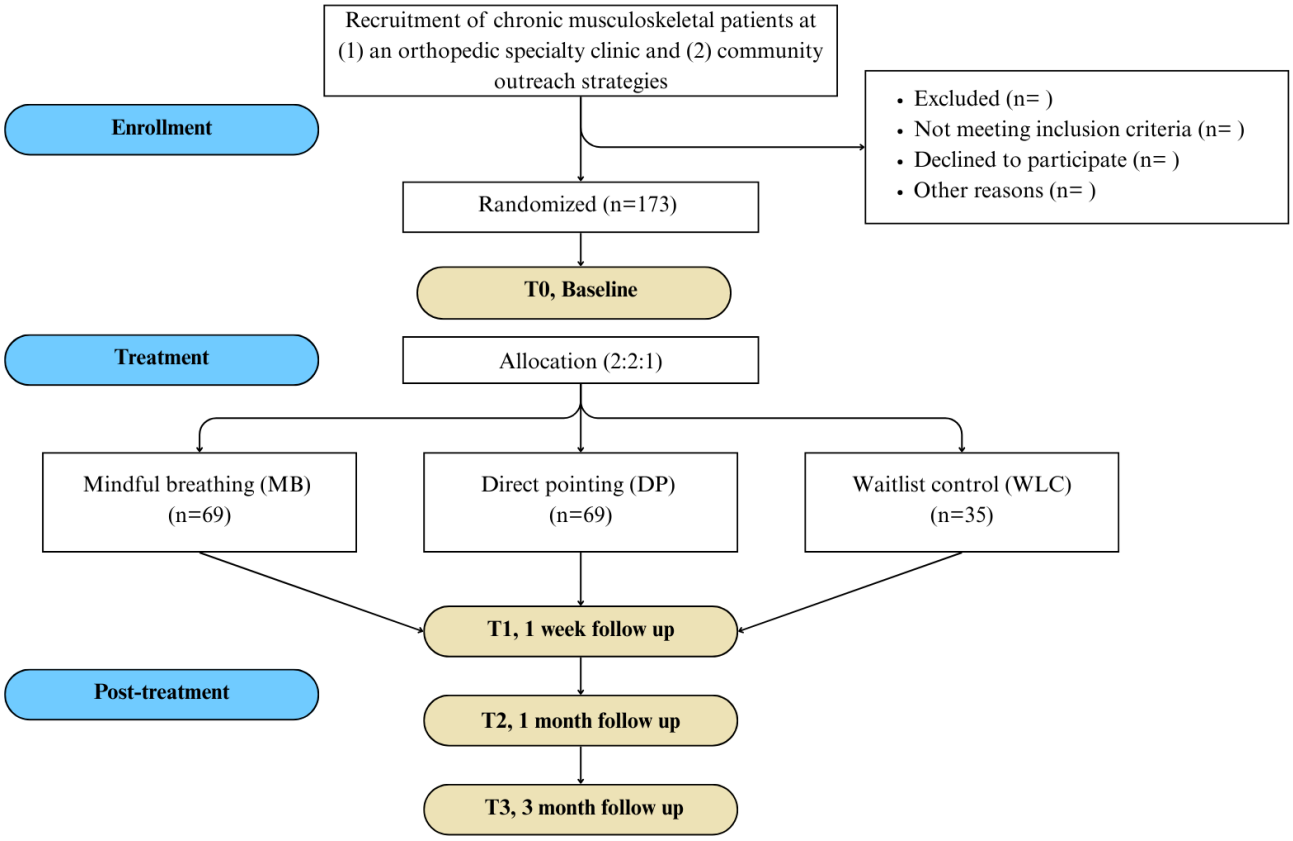

Participants

We will recruit and enroll 173 adults with CMP who meet established eligibility criteria. Sample size estimation was conducted in 2 stages, reflecting an expansion from a sole focus on mechanistic sensitivity to the joint consideration of mechanistic and clinical relevance. The first stage was designed to detect changes in self-transcendence in order to elucidate an emerging therapeutic mechanism of MBIs, whereas the second stage broadened the scope of the study to contribute to a growing body of research evaluating the effects of brief MBIs on chronic pain.

First, our original power analysis using G*Power (Heinrich Heine University Düsseldorf) indicated that a total sample of 138 participants would provide sufficient power (β=.80) to detect changes in self-reported and psychophysiological indices of self-transcendence [42]. This target reflected a required analyzable sample of 110, inflated by 25% to account for expected attrition. To promote robust detection of mechanistic effects, we specified a conservative effect size (Cohen *d*=0.30), representing half the smallest effect observed in our preliminary data (Cohen *d*=0.60). This approach was intended to support the more stringent comparison between the MB and MB+DP groups, while incorporating corrections for multiple comparisons.

Second, to enhance the clinical interpretability of the intervention effects, we added a waitlist control group. This adjustment was made prior to participant enrollment. Inclusion of this no-treatment comparison group provides a benchmark against which to evaluate the specific impact of the MBIs on both mechanistic and clinical outcomes, relative to the natural progression of symptoms in the absence of additional intervention. This control arm was designed to be half the size of each active treatment group (n=35), resulting in a total sample size of 173 and preserving a 2:2:1 randomization ratio. This total sample size (n=173) is consistent with a G*Power estimate for a 3 (group)×2 (time) mixed design, assuming a small effect (Cohen *d*=0.30), 80% power, and a Bonferroni-adjusted alpha of .0125 for 4 pairwise comparisons (ie, self-reported self-transcendence; 2 psychophysiological indices of self-transcendence, including frontal midline theta activity and DMN activity, and acute pain). Beyond these primary group-level effects, the planned sample size also supports Aims 2 and 3. Specifically, a G*Power analysis (2-tailed, α=.0167, corrected for multiple comparisons; power=0.80) indicated that 112 participants are sufficient to detect a moderate association (*r*=0.30) among phenomenological and psychophysiological indices of self-transcendence, well below the planned sample of n=173. Finally, this sample size is adequate for the planned mediation analyses; the parallel multiple-mediator model for acute pain includes 19 estimated parameters, yielding approximately 9.1 observations per parameter with n=173, consistent with the commonly cited 5 - 10 observations per parameter convention.

Eligibility Criteria

Participants will be eligible for inclusion in the study if they meet the following criteria (Textbox 1):

XSL•FO
RenderX

Textbox 1. Participant inclusion and exclusion criteria.

The following criteria were used to determine participant eligibility for the study.

Inclusion criteria

- Age 18 years or older
- Diagnosis of chronic musculoskeletal pain (CMP) by a licensed medical provider
- Self-reported average pain intensity of ≥3 on a 0–10 numeric rating scale, experienced on most days over the past 3 months
- Ability to understand spoken and written English
- Ability to complete online self-report questionnaires using a tablet or personal computer

Exclusion criteria

- Inability to provide informed consent due to cognitive, psychiatric, or physical impairment
- Previous formal mindfulness training (eg, completion of Mindfulness-Based Stress Reduction or equivalent structured programs)
- Diagnosis of cancer at the time of enrollment
- Active psychosis, suicidal ideation with intent or plan, or severe substance use disorder, as determined by a structured clinical interview (Mini-International Neuropsychiatric Interview [MINI])
- Presence of an unstable medical condition that could interfere with study participation or safety
- Surgery within the past 3 months
- Receipt of pain-relieving injections (eg, corticosteroid or hyaluronic acid) within the past 3 months

Relevant concomitant care is restricted during the trial. Patients are welcome to continue their usual medical care but are asked to refrain from beginning any new treatments during the study. Cortisone and hyaluronate injections are prohibited during the trial.

Study Settings

All mindfulness training sessions and psychophysiological assessments will be conducted at the Complementary Health Innovation Lab (CHIL), located within Florida State University's (FSU) College of Nursing. CHIL is equipped with a dedicated psychophysiology laboratory that integrates EEG and fNIRS systems, allowing for simultaneous data acquisition via a unified interface.

Sessions involving psychophysiological data collection (Sessions 1 and 4) will be conducted in a low-light environment to optimize the quality of fNIRS signal acquisition and more closely approximate real-world mindfulness settings. During these sessions, the participant, study therapist, and trained research assistant (RA) will be present. For Sessions 2 and 3, which do not involve psychophysiological data collection, only the participant and the therapist will be in the lab room. The treatment environment remains identical across all sessions, with the sole distinction being that the simultaneous EEG and fNIRS recording cap is worn only during the psychophysiology sessions.

Self-report measures will be administered during sessions, including assessments of self-transcendence and acute pain. All in-session measures will be completed using a tablet-based survey via the Qualtrics (Qualtrics International Inc) platform. Broader clinical assessments, including baseline and follow-up surveys at 1 week, 1 month, and 3 months, will be conducted remotely through Qualtrics at a time and location convenient for each participant.

Assignment of Interventions and Blinding

Participants will be randomly assigned to intervention groups using a computer-generated randomization sequence created in Microsoft Excel. Randomization is stratified by opioid use (yes or no) to ensure balance across treatment arms on this clinically relevant variable. Within each stratum, participants are allocated using permuted block randomization with a fixed block size of 6 to maintain allocation concealment and prevent prediction of assignment. Treatment allocation is generated by personnel uninvolved in assessment or treatment and concealed on the private, password-protected computer of the uninvolved research member.

To preserve blinding, the randomization sequence will not be accessible to either assessment staff or the principal investigator (PI) until the study's completion. Treatment allocation, for active treatment groups, will only be revealed to the interventionist immediately before the participant's first treatment appointment. Outcome assessor blinding will further be ensured through the use of online, asynchronous self-report surveys, preventing any direct interaction between study personnel and participants.

Unblinding will only occur in exceptional circumstances where knowledge of the assigned intervention is essential for the clinical safety of the participant or to address a serious adverse event (AE). The request for unblinding must be made by the PI or a designated senior clinical investigator after determining that unblinding is necessary for participant safety or ethical oversight. The PI will document the rationale for unblinding and notify the institutional review board (IRB) if the event involves a serious adverse reaction or protocol deviation that may impact participant safety or data integrity.

XSL•FO
RenderX

Recruitment Plan

Participants will be recruited through a combination of clinical and community-based outreach strategies to ensure a diverse sample of adults with CMP. Primary recruitment will occur through an orthopedic specialty clinic located in the Southeastern United States. Potential participants will be identified through direct referrals from clinicians or may self-refer after viewing recruitment materials displayed in the clinic. To supplement clinic-based recruitment, we will implement community outreach strategies designed to engage individuals with CMP who may not be actively seeking specialty care. These strategies will include the distribution of digital and print advertisements across local community centers and libraries. Targeted social media campaigns (eg, Facebook and Instagram [Meta Platforms]) will further expand recruitment reach. All advertisements will include a brief study description, eligibility criteria, and a link to a secure online screening form. Interested individuals will be directed to complete a brief web-based prescreening questionnaire to determine initial eligibility. Those who appear eligible will be contacted by study staff for a structured telephone interview to confirm inclusion and exclusion criteria.

AE Reporting and Harms

Participant safety and data integrity will be monitored on an ongoing basis by the PI and designated study team members. Any reported events, whether physical, psychological, or behavioral, will be documented in a standardized case report form and evaluated for severity, duration, and potential relationship to the study intervention. The PI will assess each AE for relatedness and seriousness and determine whether the event qualifies as an unanticipated problem involving risk to participants.

Any AE or unanticipated problem involving risk to participants will be reported in accordance with the local IRB policies and applicable regulatory requirements. Nonserious but study-related AEs will be recorded and reviewed regularly by the study team to identify any emerging patterns that could indicate risks associated with the intervention or trial procedures. Participants who experience distress or negative effects related to the intervention (eg, heightened anxiety) will be provided with appropriate support. All documentation of AEs will be maintained in a secure, confidential manner and reviewed periodically to ensure prompt and appropriate response.

Patient and Public Involvement Statement

Patients and the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Interventions

Participants in the 2 active treatment arms (MB and MB+DP) will attend 4 weekly, 30-minute, manualized mindfulness training sessions. Each session will be delivered in person by a PhD-level interventionist. In addition to the in-session mindfulness practice, all participants will be assigned daily at-home mindfulness practice throughout the 4-week intervention period. The at-home mindfulness practice will be a 15-minute, audio-recorded mindfulness practice recorded by a clinician and sent to participants via email. Each participant in the active treatment groups will receive a 15-minute audio recording that corresponds to the specific mindfulness practice style of the group to which they were randomized.

Each session will follow a consistent structure. The first session will begin with a 5-minute introduction to the intervention. In Sessions 2 through 4, the initial 5 minutes will be used to review participants' experiences with their at-home mindfulness practice. All sessions will then include a 15-minute guided mindfulness practice and conclude with 10 minutes of therapist-led inquiry and reflective discussion. This structure is informed by established and empirically supported brief mindfulness intervention models [43]. To ensure treatment fidelity, all sessions will be audio-recorded. A study supervisor will review session recordings and provide regular supervision, ensuring adherence to the treatment manual and consistency in intervention delivery across participants. The specific elements of each session in both programs are described in Table 1.

XSL•FO
RenderX

Table . Content and practice overview of each session of mindful breathing (MB) and mindful breathing+direct pointing (MB+DP) active treatment groups.

Program session	MB[a]	MB+DP[b]	Total session duration
1	• Brief introduction and pain psychoeducation (5 minutes) • Formal practices: • Mindfulness of breath (15 minutes) • Therapeutic processing (10 minutes)	• Brief introduction and pain psychoeducation (5 minutes) • Formal practices: • Mindfulness of breath (10 minutes) • 5 minutes of direct pointing instruction • Therapeutic processing (10 minutes)	• 30 minutes
2	• Brief welcome + review of at-home practice (5 minutes) • Formal practices: • Mindfulness of breath (15 minutes) • Therapeutic processing (10 minutes)	• Brief welcome + review of at-home practice (5 minutes) • Formal practices: • Mindfulness of breath (10 minutes) • 5 minutes of direct pointing instruction • Therapeutic processing (10 minutes)	• 30 minutes
3	• Brief welcome + review of at-home practice (5 minutes) • Formal practices: • Mindfulness of breath (15 minutes) • Therapeutic processing (10 minutes)	• Brief welcome + review of at-home practice (5 minutes) • Formal practices: • Mindfulness of breath (10 minutes) • 5 minutes of direct pointing instruction • Therapeutic processing (10 minutes)	• 30 minutes
4	• Brief welcome + review of at-home practice (5 minutes) • Formal practices: • Mindfulness of breath (15 minutes) • Therapeutic processing (10 minutes)	• Brief welcome + review of at-home practice (5 minutes) • Formal practices: • Mindfulness of breath (10 minutes) • 5 minutes of direct pointing instruction • Therapeutic processing (10 minutes)	• 30 minutes

[a]MB: mindful breathing.

[b]MB+DP: mindful breathing+direct pointing.

Mindful Breathing (MB)

Participants assigned to the MB group will engage in a focused attention mindfulness practice adapted from validated brief mindfulness protocols [43,44]. During the 15-minute guided practice, participants will be instructed to maintain attention on the breath, particularly sensations at the nostrils and abdomen, and to gently return their focus to the breath when distracted by thoughts, emotions, or bodily sensations. When pain arises, they will be guided to acknowledge it nonjudgmentally, without resistance or elaboration, before returning their attention to the breath.

Mindful Breathing+Direct Pointing (MB+DP)

Participants assigned to the MB+DP group will receive the same initial instructions as the MB group for the first 10 minutes of practice. During this time, foundational mindfulness skills are cultivated to support receptivity to the DP instruction. In the final 5 minutes of the guided practice, the therapist will deliver DP instructions, an introspective, contemplative technique derived from Mahamudra and Dzogchen traditions [22,23]. DP aims to shift attention from the contents of awareness (eg, pain and thoughts) to the nature of awareness itself [22,23]. Instructions will guide participants to experientially investigate the spacious, nonconceptual qualities of awareness, often described in classical sources as possessing inherent clarity and

XSL•FO
RenderX

bliss [22,23]. The DP phase is designed to induce a transient, self-transcendent state in which rigid identification with self-referential content (eg, "I am in pain") may dissolve.

Waitlist Control (WLC)

The WLC group was included to provide a reference for the natural progression of symptoms in the absence of additional intervention. While this group allows for assessment of change over time relative to no treatment, it does not control for nonspecific factors such as attention, expectancy, or facilitator contact. Participants assigned to the WLC group will not receive any active intervention during the initial 4-week study period but will complete all study assessments on the same schedule as those in the treatment arms. Specifically, they will undergo baseline evaluation and complete identical outcome measures, such as self-reported self-transcendent state, acute pain ratings, EEG, and fNIRS, at time points matched to Sessions 1 and 4 of the intervention groups. They will also complete the full set of follow-up assessments at 1 week, 1 month, and 3 months postintervention to evaluate potential naturalistic changes in clinical outcome variables over time. Following the final assessment, participants in the waitlist group will be offered the opportunity to participate in a 2-hour MBI for chronic pain.

Measures

Outcomes will be evaluated at two primary timescales: (1) during and immediately surrounding the mindfulness training sessions to assess mechanistic changes, including self-reported self-transcendence and psychophysiological responses, and (2) at baseline and follow-up intervals to examine the durability of clinical effects. Psychophysiological data (EEG and fNIRS), along with session-specific self-report measures, will be collected during the first and final training sessions (Sessions 1 and 4) to assess within-participant changes over the course of the intervention (see Table 2). Broader clinical outcomes, including pain interference, pain catastrophizing, and sleep disturbance, will be assessed remotely via online surveys at baseline and again at 1-week, 1-month, and 3-month postbaseline to evaluate the persistence of any treatment-related improvements (Table 3).

Table . State survey: overview of measures and corresponding measurement time points.

Measure	Target concept	Session 1	Session 2	Session 3	Session 4
NADA-S[a]	Self-transcendence	✓			✓
MPoD[b]	Decentering	✓			✓
Pain intensity (single item)	Pain intensity	✓			✓
EEG[c]	Frontal midline theta power	✓			✓
fNIRS[d]	Fronto-temporal-parietal cortical proxy of DMN[e] activity; cerebral cortex	✓			✓

[a]NADA-S: Nondual Awareness Dimensional Assessment–State.

[b]MPoD: Metacognitive Processes of Decentering Scale.

[c]EEG: electroencephalogram.

[d]fNIRS: functional near-infrared spectroscopy.

[e]DMN: default mode network.

XSL•FO
RenderX

Table . Trait survey: overview of measures and corresponding measurement time points.

Measure	Target concept	T0[a]	T1[b]	T2[c]	T3[d]
BPI[e] – Functional Interference	Functional Interference	✓	✓	✓	✓
NIH HEAL CDE[f] Demographics	Demographic Information	✓			
BPI-PI[g]	Pain Intensity	✓	✓	✓	✓
Physical function	Physical function	✓	✓	✓	✓
PCS[h]	Pain catastrophizing	✓	✓	✓	✓
PHQ-2[i]	Depression	✓	✓	✓	✓
GAD-2[j]	Anxiety	✓	✓	✓	✓
PROMIS[k] Sleep Disturbance – Short Form 6a	Sleep	✓	✓	✓	✓

[a]T0: 1-week pretreatment (Baseline).

[b]T1: 1-week post treatment.

[c]T2: 1-month post treatment.

[d]T3: 3-month post treatment.

[e]BPI: Brief Pain Inventory.

[f]NIH HEAL CDE: The Helping to End Addiction Long-term Initiative Common Data Elements.

[g]PI: Pain intensity

[h]PCS: Pain Catastrophizing Scale.

[i]PHQ-2: Patient Health Questionnaire 2-item.

[j]GAD-2: Generalized Anxiety Disorder 2-item.

[k]PROMIS: Patient-Reported Outcomes Measurement Information System.

Primary Outcome

Primary Self-Report Outcome

The self-transcendent state will be assessed immediately before and after the first and final treatment sessions using the state version of the Nondual Awareness Dimensional Assessment (NADA-S). This validated 3-item measure captures momentary fluctuations in nondual awareness. Two items assess reductions in self-salience, and one assesses the presence of bliss. All items are rated on an 11-point numeric rating scale ranging from 0 (“not at all”) to 10 (“extremely”).

Primary Psychophysiological Outcomes

Frontal midline theta power will be measured by EEG. In Sessions 1 and 4, participants in the active treatment groups (MB and MB+DP) will complete psychophysiological data collection during a 5-minute resting baseline prior to treatment and throughout the 15-minute mindfulness practice. During the same sessions, participants in the WLC group will undergo a 5-minute resting baseline followed by a 15-minute time-matched control period. During all psychophysiological assessments, EEG will be continuously recorded from 10 midline scalp sites (Fz, F3, F4, FC1, FC2, FCz, Cz, CP1, CP2, and PZ) using an active sensor cap with Ag-AgCl electrodes (actiCap GmbH). In addition, vertical electrooculograms will be recorded. All recordings will be collected with an actiCHamp amplifier (Brain Products GmbH). Data will be acquired at a sampling rate of 500 Hz, a resolution of 0.489 µV and an amplification cutoff of 140 Hz, with impedances kept below 10 kilohms.

Connectivity in a validated fronto-temporal-parietal cortical proxy of DMN activity (ie, connectivity in-between channels belonging to the medial prefrontal cortex and the left and right MTG and STG, left and right posterior temporoparietal lobe, left and right temporoparietal junction will be measured by fNIRS during a 5-minute resting baseline period before treatment and during the 15-minute mindfulness practice. Hemodynamic responses to neuroactivation via oxy-, deoxy-, and total hemoglobin changes in the cerebral cortex will be measured with the NIRx NIRSport 2, a user-friendly, modular, and robust wireless fNIRS platform. The automated signal optimization algorithm ensures optimal signal quality before a measurement is started. Once data are being recorded, oxygenated hemoglobin and hemoglobin concentration changes can be visualized in real-time in several display modes. In addition, high-end whole-head visualizations are immediately available.

Acute Clinical Outcome

Acute pain intensity will be measured immediately before and after Sessions 1 and 4 using a single-item numeric rating scale: “How intense is your pain right now?” Responses are scored on an 11-point scale from 0 (“no pain”) to 10 (“worst pain imaginable”).

XSL•FO
RenderX

Tertiary and Secondary Outcomes

Tertiary Clinical Outcomes

The following pain-related outcomes will be assessed using validated instruments from the Helping to End Addiction Long-term Initiative (NIH HEAL) Common Data Elements (CDE) battery at pretreatment and posttreatment time points (see Table 4): Brief Pain Inventory, Patient-Reported Outcomes Measurement Information System (PROMIS) Physical Functioning Short Form 6b, Brief Pain Catastrophizing Scale (PCS), Patient Health Questionnaire-2 (PHQ-2), Generalized Anxiety Disorder 2-Item Scale (GAD-2), and PROMIS Short Form v1.0–Sleep Disturbance 6a.

XSL•FO
RenderX

Table . Schedule of enrollment, interventions, and assessments (SPIRIT [Standard Protocol Items: Recommendations for Interventional Trials] diagram).

Description	**Study phase**			In-lab **procedures**				Follow-**u**p **p**ost-**t**reatment		
	Enrollment	Baseline	Allocation	Week 1	Week 2	Week 3	Week 4	1-week	1-month	3-month
Time point	–T1[a]	T0[b]						T1[c]	T2[d]	T3[e]
Enrollment										
Eligibility form	✓									
Informed consent	✓									
MINI[f]	✓									
Randomization										
Allocation			✓							
Study arm										
MB[g]				✓	✓	✓	✓	✓	✓	✓
MB+DP[h]				✓	✓	✓	✓	✓	✓	✓
WLC[i]				✓			✓	✓	✓	✓
Assessments										
NADA-S[j]				✓			✓			
EEG[k]				✓			✓			
fNIRS[l]				✓			✓			
MPoD[m]				✓			✓			
Acute pain intensity				✓			✓			
At-home mindfulness practice (MB and MB+DP only)				✓	✓	✓	✓			
BPI[n]		✓						✓	✓	✓
Physical function		✓						✓	✓	✓
PCS[o]		✓						✓	✓	✓
PHQ-2[p]		✓						✓	✓	✓
GAD-2[q]		✓						✓	✓	✓
PROMIS[r] sleep disturbance		✓						✓	✓	✓
Pain catastrophizing		✓						✓	✓	✓

[a]–T1: Enrollment activities

XSL•FO
RenderX

[b]T0: 1-week pretreatment (baseline).
[c]T1: 1-week posttreatment.
[d]T2: 1-month posttreatment.
[e]T3: 3-month posttreatment.
[f]MINI: Mini-International Neuropsychiatric Interview.
[g]MB: mindful breathing.
[h]MB+DP: mindful breathing+direct pointing.
[i]WLC: waitlist control.
[j]NADA-S: Nondual Awareness Dimensional Assessment-State.
[k]EEG: electroencephalogram.
[l]fNIRS: functional near-infrared spectroscopy.
[m]MPoD: Metacognitive Processes of Decentering.
[n]BPI: Brief Pain Inventory.
[o]PCS: Pain Catastrophizing Scale.
[pq]PHQ-2: Patient Health Questionnaire 2-item. GAD-2: Generalized Anxiety Disorder 2-item.
[r]PROMIS: Patient-Reported Outcomes Measurement Information System.

Pain

Participants will self-report their pain intensity, treatments for pain relief, effectiveness of those treatments, and pain interference with the 14-item BPI. Statements include “Please rate your pain by indicating the one number that best describes your pain at its WORST in the past 24 HOURS” with an 11-point Likert-like scale (0=no pain, 10=pain as bad as you can imagine). Five items measure recent and immediate pain intensity. One item asks participants about their pain treatments, and the next immediate item asks how much relief the treatments offered (from 0% to 100%). The next 7 items measure pain interference in the participant’s daily life with an 11-point Likert-like scale (0=does not interfere, 10=completely interferes).

Physical Function

Participants will self-report their physical function with the PROMIS Physical Function (PROMIS Physical Functioning Short Form 6b) questionnaire [45]. The PROMIS Physical Function questionnaire is a 6-item measure using a 5-point Likert scale (5=not at all, 1=cannot do) with questions such as “Does your health now limit you in doing two hours of physical labor?” Other questions, such as “Are you able to do chores such as vacuuming or yard work?,” use a 5-point Likert scale (5=without any difficulty, 1=unable to do).

Pain Catastrophizing

Participants will be asked about their pain catastrophizing with the Brief Pain Catastrophizing Scale (PCS) [36]. The Brief PCS is a validated, 5-item questionnaire on which participants rate their thoughts and feelings about pain on a 5-point Likert scale, ranging from 0 (“Not at all”) to 4 (“All the time”). Statements include “It’s awful, and I feel that it overwhelms me.”

Depression

Depressive symptoms were measured using the Patient Health Questionnaire-2 (PHQ-2), a 2-item questionnaire that serves as a brief screener for depression [35]. The PHQ-2 prompts patients to reflect on their mood over the previous 2 weeks, asking them how often they experienced “little interest or pleasure in doing things” or felt “down, depressed, or hopeless.” A 4-point Likert scale ranging from 0 (not at all) to 3 (nearly every day) is used to report the frequency of depressive symptoms, and a total score is determined from summing these points together. A total score of 3 or greater supports the need for further screening and indicates an increased likelihood of major depressive disorder.

Anxiety

Anxiety symptoms were measured using the Generalized Anxiety Disorder 2-Item Scale (GAD-2), a 2-item screener for generalized anxiety disorder [46]. Similar to the PHQ-2, the GAD-2 asks patients about how often they experienced certain symptoms over the past 2 weeks, including “feeling nervous, anxious, or on edge” and “not being able to stop or control worrying.” Both items are scored using a 4-point Likert scale that spans from 0 (not at all) to 3 (nearly every day). Points from these 2 items are added together to calculate a total score, with a total score of 3 or above being suggestive of generalized anxiety disorder.

Sleep

Sleep disturbance was assessed with the PROMIS Sleep Disturbance Scale (PROMIS Short Form v1.0–Sleep Disturbance 6a), a 6-item questionnaire that evaluates a patient’s overall sleep experience [33]. The first item evaluates a patient’s sleep quality over the past 7 days. Responses are scored on a 5-point Likert scale ranging from 1 (very poor) to 5 (very good). The other 5 items ask about the depth and restfulness of a patient’s sleep over the past 7 days. Sample items include “My sleep was refreshing” and “I tried hard to get to sleep.” As with the first item, these responses are also scored on a 5-point Likert scale ranging from 1 (not at all) to 5 (very much).

Process Measures

Engagement with at-home mindfulness practice will be monitored using a Qualtrics-based tracking system. Each participant will receive a unique survey link embedded with their participant ID to record the frequency and duration of practice. At the beginning of Sessions 2, 3, and 4, participants will also self-report their adherence to at-home exercises, including at-home mindfulness practice duration and any notable experiences.

XSL•FO
RenderX

As a manipulation check, state mindfulness will be assessed using the Metacognitive Processes of Decentering (MPoD) scale [47] administered immediately before and after Sessions 1 and 4. The MPoD captures momentary shifts in decentering, a key psychological mechanism thought to mediate mindfulness-related outcomes.

Participants will receive modest financial compensation at the baseline assessment and 3 follow-up time points. These compensation procedures are designed to recognize participant contributions while maintaining ethical standards for noncoercive remuneration.

Statistical Analysis

Analysis Plan

The protocol will favor an intention-to-treat approach for the analysis of results. Data analysis will be performed in SPSS (IBM Corp). Baseline demographic information will be summarized using descriptive statistics. Continuous variables will be summarized with means and SDs. Categorical variables will be summarized with counts, proportions, and medians (IQR). The distribution of continuous variables will be assessed using graphical inspection and the Shapiro-Wilk test for normality. If the assumption of normality is not met, equivalent nonparametric approaches or data transformation will be used. The multiple testing problem will be accounted for using Bonferroni correction. All hypothesis tests will be 2-sided, and the familywise Type I error rate will be controlled at α=.05 using Bonferroni-adjusted significance thresholds. If needed, statistical analysis will also be conducted after controlling for baseline demographics and clinical characteristics (eg, age, sex, and race).

The effects of treatment on continuous variables will be evaluated using linear mixed models with fixed effects consisting of a time factor (before to after the experimental session; before treatment to 3-month follow-up) and a between-participants treatment factor (MB vs MB+DP vs WLC, as appropriate). If necessary, we will adjust for randomization imbalance by co-varying baseline variables. Little's Missing Completely at Random (MCAR) test will be used to determine if data are missing completely at random. We will use maximum likelihood estimation procedures to deal with missing data according to an intent-to-treat philosophy that is robust against common patterns of missing data. The maximum likelihood estimation is based on all data observations; no values are deleted or imputed.

All EEG analyses will be performed in MATLAB (Matrix Laboratory; MathWorks) using scripts implementing the EEGLAB toolbox [48]. First, epochs of 300 s will be created for each participant. Then, the signal will be notch-filtered at 60 Hz and low-pass filtered at 40 Hz cutoff using a fourth-order Butterworth filter and high-pass filtered at 0.1 Hz cutoff using a second-order Butterworth filter. Filtered data will be passed through the PREP pipeline [49] which detrends the data, applies a notch filter tapering off the harmonics of 60 Hz, rereferences the data to the linked earlobe, and identifies and interpolates bad channels. Interpolation in low-density montages has been shown to produce valid data [49,50]. In the power spectral density analyses, average power spectra in the theta (4-8 Hz) band will be calculated using Welch periodogram across frontal midline sites.

All fNIRS data will be preprocessed in MATLAB as follows: detected NIR light will be transformed into O_2 Hb and HHb concentrations using the modified Beer–Lambert law. Motion artifacts will be corrected using temporal derivative distribution repair, followed by band-pass filtering (0.01-0.1 Hz). Low-amplitude motion artifacts will be corrected using the algorithm of Cui et al [51], and single artifact-laden channels will be interpolated. A Gaussian PCA kernel will be applied to remove global signal components. Finally, signals will be converted to *z* scores to facilitate comparisons across participants.

Plans to Promote Participant Retention and Complete Follow-Up

To promote participant retention and ensure complete follow-up, the study will implement a structured, participant-centered engagement strategy grounded in best practices for behavioral trial management. All follow-up assessments (at 1-week, 1-month, and 3-month posttreatment) will be conducted remotely via secure, web-based surveys to enhance flexibility and reduce burden. A dedicated RA will serve as each participant's primary point of contact, maintaining regular communication and providing reminders and technical support based on participant preference (eg, phone, text, or email). Participants will receive modest compensation at baseline, each follow-up, and upon treatment completion, consistent with ethical standards for noncoercive remuneration.

Adherence will be further supported by weekly audio-guided mindfulness recordings and at-home practice diaries. Flexible scheduling will be offered for in-person EEG and fNIRS assessments to accommodate work and caregiving responsibilities. Missed sessions or assessments will be followed up within 48 hours, with up to 3 recontact attempts using alternate methods. Participants may re-engage at the next appropriate activity without penalty and may request a personalized summary of their self-report data poststudy. This approach aims to promote retention and data completeness while minimizing participant burden.

Ethical Considerations

The Florida State University IRB approved all study procedures prior to participant enrollment (STUDY00004506). All procedures comply with institutional and international ethical standards. AEs will be monitored and reported per IRB policy. This study was prospectively registered on Open Science Framework (OSF): DOI 10.17605/OSF.IO/3UY2C. Trial results will be disseminated via peer-reviewed publications, academic conferences, and stakeholder engagement efforts. A summary of findings will also be offered to participants following study completion.

All participants provided informed consent prior to the participation in the study, in accordance with institutional guidelines and the 1964 Declaration of Helsinki. Any publications resulting from this research will not contain any identifiable information in any form. A model consent form

XSL•FO
RenderX

will be made available from the corresponding author on reasonable requests to qualified individuals.

After eligibility has been established, individuals who express interest in participating in the study will complete the informed consent process under the guidance of a trained RA, in accordance with IRB requirements and ethical research standards. Participants will have the option to receive a copy of the signed consent form via email for their records. All consent documentation will be stored securely in compliance with institutional data protection policies.

Data will be deidentified and stored on secure servers at CHIL. Any paper data (questionnaires, sociodemographic, and clinical info) will have identifiers removed and banked for a minimum of 3 years. Deidentified data will be available to the research team, including the PI (AWH), the research coordinator, and the RAs. Any electronic data will be stored on a password-protected computer. Computers will be stored in locked rooms. Any data collected online will be stored on an encrypted data server at the College of Nursing. Any downloaded data will be stored on a password-protected computer. The RA supervised by the PI will be responsible for data receipt. AWH will be responsible for data analysis and results interpretation.

Results

At the time of the initial manuscript submission, this protocol is Version 4, dated May 20, 2025. Recruitment for the trial began on October 28, 2024, is ongoing, and is expected to be completed by May 2027. Final results will be reported in accordance with CONSORT (Consolidated Standards of Reporting Trials) guidelines and submitted for peer-reviewed publication. Any substantial protocol amendments will be documented with updated version numbers and dates and communicated to relevant regulatory bodies, investigators, and trial registries as required.

Discussion

CMP continues to pose a significant public health burden, in part due to its capacity to become enmeshed with the sense of self. The present randomized controlled clinical trial is designed to move beyond symptom reduction alone by examining whether altering the self-referential structure of pain perception, via DP instruction, can produce meaningful and durable analgesic effects. DP may offer relief by shifting how individuals relate to pain, reducing identification with painful sensations and loosening the pain-laden self-concept that can sustain the pain they endure. The Pointing Beyond project is designed to make 4 key contributions to the field of mindfulness-based pain relief.

The first contribution concerns measurement. By combining momentary self-report with concurrent psychophysiological indices, frontal midline theta activity via EEG and DMN dynamics via fNIRS, this trial moves beyond the retrospective and aggregate outcome data that have characterized most prior work on self-transcendence in mindfulness research. Real-time triangulation of subjective experience and neural activity allows for a more precise characterization of when and how self-transcendent states arise during practice and whether their phenomenological and neurobiological signatures are coherently coupled. This multimodal approach is essential because self-report measures of transcendence are inherently subjective and susceptible to demand characteristics; convergence with independent psychophysiological indices would substantially strengthen causal inference about the mechanism.

Second, a central contribution is the dismantling design itself. Comparing MB alone against MB+DP against a waitlist control isolates the specific contribution of direct pointing instruction to both self-transcendence and analgesia. Prior trials incorporating DP, including RCTs of the MORE program and brief perioperative MBIs, confounded DP with other intervention components, leaving its independent effect unexamined. If MB+DP produces meaningfully greater self-transcendence and pain relief than MB alone, this would provide the first direct evidence that DP adds analgesic value beyond standard mindfulness instruction. This finding would have immediate implications for how brief MBIs are designed and delivered in clinical settings.

Further, this trial advances the goals of personalized medicine by investigating whether different mindfulness practice instructions produce distinct phenomenological experiences and clinical outcomes. Comparing MB and MB+DP at the level of both phenomenological experience and neural activity advances the goals of mechanism-based treatment personalization. If the two conditions diverge in their effects on self-transcendence but converge on pain outcomes, this would suggest that multiple pathways can produce equivalent clinical benefit and that individual preference and acceptability, rather than specific instructional content, should guide practice selection. Conversely, if MB+DP produces superior outcomes on both mechanistic and clinical indices, this would support a more targeted prescriptive approach, particularly for patients whose pain experience is strongly identity-bound. Either pattern of results would meaningfully inform clinical decision-making in a field that currently lacks guidance on how to match mindfulness instruction style to patient presentation.

The fourth contribution concerns scalability. Demonstrating that just four 30-minute sessions can produce durable reductions in pain interference through three-month follow-up would represent a meaningful addition to the feasibility literature on brief MBIs. In contrast to multiweek programs like MBSR or MORE, which impose substantial time and access burdens, this trial evaluates the effects of just 2 hours of in-person training (across 4 30-minute sessions) paired with 15 minutes of daily practice. An minfulness intervention of this brevity could realistically be integrated into primary care, rehabilitation settings, or digital health platforms. Short, scalable interventions are especially vital for reaching populations with limited access to traditional programs, such as rural residents or those with time, mobility, or financial constraints.

Several challenges warrant consideration. Capturing transient self-transcendent states during brief interventions requires sensitive instrumentation and careful interpretation. While fNIRS offers valuable ecological validity, its spatial resolution limits assessment of deeper DMN structures such as the posterior

XSL•FO
RenderX

cingulate cortex. Moreover, ensuring fidelity in delivering nuanced practices such as DP necessitates skilled facilitators and robust supervision protocols. Despite these limitations, the study's emphasis on real-time neural and experiential assessment in naturalistic settings enhances its translational relevance.

The Pointing Beyond project represents a theoretically grounded and methodologically rigorous effort to identify the specific mechanisms through which mindfulness produces pain relief. By isolating the contribution of direct pointing instruction, triangulating subjective and neurobiological indices of self-transcendence, and testing their mediating role in both acute and sustained pain outcomes, this trial directly contributes to the scientific understanding of the experiential and neural processes through which mindfulness relieves CMP.

Acknowledgments

The authors acknowledge that AI tools were used under STM Category 1 (refinement, correction, editing, or formatting) to improve the clarity and readability of the manuscript [49]. All content and interpretation were generated by the authors.

Funding

This work was supported by the Tiny Blue Dot Foundation.

Data Availability

The datasets used and/or analyzed during this study will be made available from the corresponding author on reasonable requests to qualified individuals.

Authors' Contributions

Conceptualization: AWH
Funding acquisition: AWH
Investigation: KC
Methodology: AWH
Project administration: KC, AWH
Supervision: AWH
Writing – original draft: KC
Writing – review & editing: KC, CM, GM, AWH

Conflicts of Interest

None declared.

Checklist 1
SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist.
[PDF File, 300 KB - resprot_v15i1e82362_app1.pdf]

References

1. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet 2020 Oct;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9]
2. Sebbag E, Felten R, Sagez F, Sibilia J, Devilliers H, Arnaud L. The world-wide burden of musculoskeletal diseases: a systematic analysis of the World Health Organization Burden of Diseases Database. Ann Rheum Dis 2019 Jun;78(6):844-848. [doi: 10.1136/annrheumdis-2019-215142] [Medline: 30987966]
3. Treede RD, Rief W, Barke A, et al. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). PAIN 2019 Jan;160(1):19-27. [doi: 10.1097/j.pain.0000000000001384] [Medline: 30586067]
4. Dieleman JL, Cao J, Chapin A, et al. US health care spending by payer and health condition, 1996-2016. JAMA 2020 Mar 3;323(9):863-884. [doi: 10.1001/jama.2020.0734] [Medline: 32125402]
5. Qaseem A, Wilt TJ, McLean RM, et al. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the American College of Physicians. Ann Intern Med 2017 Apr 4;166(7):514-530. [doi: 10.7326/M16-2367] [Medline: 28192789]
6. Driscoll MA, Edwards RR, Becker WC, Kaptchuk TJ, Kerns RD. Psychological interventions for the treatment of chronic pain in adults. Psychol Sci Public Interest 2021 Sep;22(2):52-95. [doi: 10.1177/15291006211008157] [Medline: 34541967]
7. Pincus T, Morley S. Cognitive-processing bias in chronic pain: a review and integration. Psychol Bull 2001 Sep;127(5):599-617. [doi: 10.1037/0033-2909.127.5.599] [Medline: 11548969]

XSL•FO
RenderX

8. Garland EL, Froeliger B, Zeidan F, Partin K, Howard MO. The downward spiral of chronic pain, prescription opioid misuse, and addiction: cognitive, affective, and neuropsychopharmacologic pathways. Neurosci Biobehav Rev 2013 Dec;37(10 Pt 2):2597-2607. [doi: 10.1016/j.neubiorev.2013.08.006] [Medline: 23988582]
9. Garland EL, Fredrickson BL. Positive psychological states in the arc from mindfulness to self-transcendence: extensions of the mindfulness-to-meaning theory and applications to addiction and chronic pain treatment. Curr Opin Psychol 2019 Aug;28:184-191. [doi: 10.1016/j.copsyc.2019.01.004] [Medline: 30763873]
10. Baliki MN, Geha PY, Apkarian AV, Chialvo DR. Beyond feeling: chronic pain hurts the brain, disrupting the default-mode network dynamics. J Neurosci 2008 Feb 6;28(6):1398-1403. [doi: 10.1523/JNEUROSCI.4123-07.2008] [Medline: 18256259]
11. Baliki MN, Mansour AR, Baria AT, Apkarian AV. Functional reorganization of the default mode network across chronic pain conditions. PLOS ONE 2014;9(9):e106133. [doi: 10.1371/journal.pone.0106133] [Medline: 25180885]
12. Baliki MN, Petre B, Torbey S, et al. Corticostriatal functional connectivity predicts transition to chronic back pain. Nat Neurosci 2012 Jul 1;15(8):1117-1119. [doi: 10.1038/nn.3153] [Medline: 22751038]
13. Kucyi A, Moayedi M, Weissman-Fogel I, et al. Enhanced medial prefrontal-default mode network functional connectivity in chronic pain and its association with pain rumination. J Neurosci 2014 Mar 12;34(11):3969-3975. [doi: 10.1523/JNEUROSCI.5055-13.2014] [Medline: 24623774]
14. Jiang Y, Oathes D, Hush J, et al. Perturbed connectivity of the amygdala and its subregions with the central executive and default mode networks in chronic pain. Pain 2016 Sep;157(9):1970-1978. [doi: 10.1097/j.pain.0000000000000606] [Medline: 27168362]
15. Zeidan F, Vago DR. Mindfulness meditation-based pain relief: a mechanistic account. Ann N Y Acad Sci 2016 Jun;1373(1):114-127. [doi: 10.1111/nyas.13153] [Medline: 27398643]
16. Zeidan F, Emerson NM, Farris SR, et al. Mindfulness meditation-based pain relief employs different neural mechanisms than placebo and sham mindfulness meditation-induced analgesia. J Neurosci 2015 Nov 18;35(46):15307-15325. [doi: 10.1523/JNEUROSCI.2542-15.2015] [Medline: 26586819]
17. Zeidan F, Grant JA, Brown CA, McHaffie JG, Coghill RC. Mindfulness meditation-related pain relief: evidence for unique brain mechanisms in the regulation of pain. Neurosci Lett 2012 Jun 29;520(2):165-173. [doi: 10.1016/j.neulet.2012.03.082] [Medline: 22487846]
18. Sezer I, Pizzagalli DA, Sacchet MD. Resting-state fMRI functional connectivity and mindfulness in clinical and non-clinical contexts: a review and synthesis. Neurosci Biobehav Rev 2022 Apr;135:104583. [doi: 10.1016/j.neubiorev.2022.104583] [Medline: 35202647]
19. Riegner G, Posey G, Oliva V, Jung Y, Mobley W, Zeidan F. Disentangling self from pain: mindfulness meditation-induced pain relief is driven by thalamic-default mode network decoupling. PAIN 2023 Feb 1;164(2):280-291. [doi: 10.1097/j.pain.0000000000002731] [Medline: 36095039]
20. Yaden DB, Haidt J, Hood RW, Vago DR, Newberg AB. The varieties of self-transcendent experience. Rev Gen Psychol 2017 Jun;21(2):143-160. [doi: 10.1037/gpr0000102]
21. Hanley AW, Garland EL. Self-transcendence predicts better pre- and postoperative outcomes in two randomized clinical trials of brief mindfulness-based interventions. Mindfulness (N Y) 2022 Jun;13(6):1532-1543. [doi: 10.1007/s12671-022-01896-6]
22. Namgyal DT. Mahamudra: The Moonlight -- Quintessence of Mind and Meditation, 2nd edition: Wisdom Publications; 2006.
23. Brown DP. Pointing Out the Great Way: The Stages of Meditation in the Mahamudra Tradition: Wisdom Publications; 2006.
24. Wallis CD. The Recognition Sūtras: A Complete Translation and Explanation of the 1,000-Year-Old Spiritual Masterpiece the Pratyabhijñā-Hr daya by Rājānaka K emarāja: Mattamayūra Press; 2017.
25. Hanley AW, Gililland J, Erickson J, et al. Brief preoperative mind-body therapies for total joint arthroplasty patients: a randomized controlled trial. Pain 2021 Jun 1;162(6):1749-1757. [doi: 10.1097/j.pain.0000000000002195] [Medline: 33449510]
26. Hanley AW, Gililland J, Garland EL. To be mindful of the breath or pain: comparing two brief preoperative mindfulness techniques for total joint arthroplasty patients. J Consult Clin Psychol 2021 Jul;89(7):590-600. [doi: 10.1037/ccp0000657] [Medline: 34165999]
27. Garland EL, Hanley AW, Riquino MR, et al. Mindfulness-oriented recovery enhancement reduces opioid misuse risk via analgesic and positive psychological mechanisms: a randomized controlled trial. J Consult Clin Psychol 2019 Oct;87(10):927-940. [doi: 10.1037/ccp0000390] [Medline: 31556669]
28. Garland EL, Hanley AW, Hudak J, Nakamura Y, Froeliger B. Mindfulness-induced endogenous theta stimulation occasions self-transcendence and inhibits addictive behavior. Sci Adv 2022 Oct 14;8(41):eabo4455. [doi: 10.1126/sciadv.abo4455] [Medline: 36223472]
29. Hudak J, Hanley AW, Marchand WR, Nakamura Y, Yabko B, Garland EL. Endogenous theta stimulation during meditation predicts reduced opioid dosing following treatment with mindfulness-oriented recovery enhancement. Neuropsychopharmacology 2021 Mar;46(4):836-843. [doi: 10.1038/s41386-020-00831-4] [Medline: 32919401]

XSL•FO
RenderX

30. Garland EL, Jinpa T. Mindfulness-induced self-transcendence promotes universal love with consequent effects on opioid misuse. Behav Res Ther 2024 Apr;175:104494. [doi: 10.1016/j.brat.2024.104494] [Medline: 38395015]
31. Hudak J, Rosenbaum D, Barth B, Fallgatter AJ, Ehlis AC. Functionally disconnected: a look at how study design influences neurofeedback data and mechanisms in attention-deficit/hyperactivity disorder. PLOS ONE 2018;13(8):e0200931. [doi: 10.1371/journal.pone.0200931] [Medline: 30096196]
32. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short - Form McGill Pain Questionnaire (SF - MPQ), Chronic Pain Grade Scale (CPGS), Short Form - 36 Bodily Pain Scale (SF - 36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken) 2011 Nov;63(S11):S240-S252 [FREE Full text] [doi: 10.1002/acr.20543]
33. Yu L, Buysse DJ, Germain A, et al. Development of short forms from the PROMIS. Behav Sleep Med 2011;10(1):6-24. [doi: 10.1080/15402002.2012.636266] [Medline: 22250775]
34. Kroenke K, Spitzer RL, Williams JBW, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 6;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
35. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
36. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychol Assess 1995;7(4):524-532. [doi: 10.1037/1040-3590.7.4.524]
37. Hung M, Clegg DO, Greene T, Saltzman CL. Evaluation of the PROMIS physical function item bank in orthopaedic patients. J Orthop Res 2011 Jun;29(6):947-953. [doi: 10.1002/jor.21308] [Medline: 21437962]
38. Tan G, Jensen MP, Thornby JI, Shanti BF. Validation of the Brief Pain Inventory for chronic nonmalignant pain. J Pain 2004 Mar;5(2):133-137. [doi: 10.1016/j.jpain.2003.12.005] [Medline: 15042521]
39. Garland EL, Hanley A, Farb NA, Froeliger BE. State mindfulness during meditation predicts enhanced cognitive reappraisal. Mindfulness (N Y) 2015 Apr 1;6(2):234-242. [doi: 10.1007/s12671-013-0250-6] [Medline: 26085851]
40. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008 Feb;9(2):105-121. [doi: 10.1016/j.jpain.2007.09.005] [Medline: 18055266]
41. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomized trials. Nat Med 2025 Jun;31(6):1784-1792. [doi: 10.1038/s41591-025-03668-w] [Medline: 40295741]
42. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
43. Zeidan F, Gordon NS, Merchant J, Goolkasian P. The effects of brief mindfulness meditation training on experimentally induced pain. J Pain 2010 Mar;11(3):199-209. [doi: 10.1016/j.jpain.2009.07.015] [Medline: 19853530]
44. Zeidan F, Adler-Neal AL, Wells RE, et al. Mindfulness-meditation-based pain relief is not mediated by endogenous opioids. J Neurosci 2016 Mar 16;36(11):3391-3397. [doi: 10.1523/JNEUROSCI.4328-15.2016] [Medline: 26985045]
45. Jensen RE, Potosky AL, Reeve BB, et al. Validation of the PROMIS physical function measures in a diverse US population-based cohort of cancer patients. Qual Life Res 2015 Oct;24(10):2333-2344. [doi: 10.1007/s11136-015-0992-9] [Medline: 25935353]
46. Sapra A, Bhandari P, Sharma S, Chanpura T, Lopp L, Sapra A. Using Generalized Anxiety Disorder-2 (GAD-2) and GAD-7 in a primary care setting. Cureus 2020 May 21;12(5):e8224. [doi: 10.7759/cureus.8224] [Medline: 32582485]
47. Hanley AW, Bernstein A, Nakamura Y, et al. The metacognitive processes of Decentering Scale: development and initial validation of trait and state versions. Psychol Assess 2020 Oct;32(10):956-971. [doi: 10.1037/pas0000931] [Medline: 32700920]
48. Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 2004 Mar 15;134(1):9-21. [doi: 10.1016/j.jneumeth.2003.10.009] [Medline: 15102499]
49. Bigdely-Shamlo N, Mullen T, Kothe C, Su KM, Robbins KA. The PREP pipeline: standardized preprocessing for large-scale EEG analysis. Front Neuroinform 2015;9:16. [doi: 10.3389/fninf.2015.00016] [Medline: 26150785]
50. Soler A, Muñoz-Gutiérrez PA, Bueno-López M, Giraldo E, Molinas M. Low-Density EEG for neural activity reconstruction using multivariate empirical mode decomposition. Front Neurosci 2020;14:175. [doi: 10.3389/fnins.2020.00175] [Medline: 32180702]
51. Cui J, Xu L, Bressler SL, Ding M, Liang H. BSMART: A Matlab/C toolbox for analysis of multichannel neural time series. Neural Netw 2008 Oct;21(8):1094-1104 [FREE Full text] [doi: 10.1016/j.neunet.2008.05.007]

Abbreviations

AE: adverse event
BPI: Brief Pain Inventory

XSL•FO
RenderX

CDE: Common Data Elements
CHIL: Complementary Health Innovation Lab
CMP: Chronic musculoskeletal pain
CON: College of Nursing
DMN: Default mode network
DP: direct pointing
EEG: Electroencephalography
fNIRS: functional near-infrared spectroscopy
FSU: Florida State University
GAD-2: Generalized Anxiety Disorder 2-item
IRB: Institutional Review Board
ITT: Intention-to-treat
MAR: Missing at random
MATLAB: Matrix Laboratory
MB: mindful breathing
MBI: Mindfulness-based interventions
MICE: Multiple imputations by chained equations
MINI: Mini-International Neuropsychiatric Interview
MORE: Mindfulness-Oriented Recovery Enhancement
MPoD: Metacognitive Processes of Decentering Scale
NADA-S: Nondual Awareness Dimensional Assessment – State
NIH HEAL: The Helping to End Addiction Long-term Initiative
NMAR: not missing at random
PCS: Pain Catastrophizing Scale
PHQ-2: Patient Health Questionnaire 2-item
PI: principal investigator
PROMIS: Patient-Reported Outcomes Measurement Information System
RA: research assistant
RCT: Randomized controlled trial
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
WLC: waitlist control

Edited by J Sarvestan; submitted 13.Aug.2025; peer-reviewed by N Lecy; revised version received 22.Jan.2026; accepted 23.Jan.2026; published 23.Apr.2026.

Please cite as:
Childs K, Minnick C, Martorella G, Hanley AW
Treating Chronic Pain by Modulating Phenomenological and Psychophysiological Indices of Self-Transcendence: Single-Blind Randomized Controlled Clinical Trial Protocol
JMIR Res Protoc 2026;15:e82362
URL: https://www.researchprotocols.org/2026/1/e82362
doi:10.2196/82362

© Kennedi Childs, Carter Minnick, Geraldine Martorella, Adam W Hanley. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Online Personalized Normative Feedback to Foster Intention to Change and Help Seeking in Young Adults With Disordered Gambling and Trading Behaviors: Protocol for a Randomized Controlled Trial

Ainhoa Coloma-Carmona[1,2,3], MA, PhD; José Luis Carballo[1,2,3], PhD; Fernando Miró-Llinares[4], PhD; Virtudes Pérez-Jover[1,3], PhD

[1]Center for Applied Psychology, Miguel Hernández University of Elche, Avenida de la Universidad s/n, Elche, Spain

[2]Brief Intervention and Addictions Research Group (IBREA), Health Psychology Department, Miguel Hernández University of Elche, Elche, Spain

[3]Alicante Institute for Health and Biomedical Research (ISABIAL), Alicante, Spain

[4]CRÍMINA Research Center for the Study and Prevention of Crime, Miguel Hernández University of Elche, Elche, Spain

Corresponding Author:
Ainhoa Coloma-Carmona, MA, PhD
Center for Applied Psychology, Miguel Hernández University of Elche, Avenida de la Universidad s/n, Elche, Spain

Abstract

Background: Despite the availability of treatment options, help-seeking rates among individuals with gambling problems remain low. To reach a broader population of those affected by disordered gambling, online and self-guided interventions have been developed. Personalized normative feedback (PNF) is one of the most widely used strategies for preventing gambling issues among young adults. However, most studies on PNF efficacy focus solely on its impact on the intensity and severity of gambling behavior, without exploring its potential to increase intention to change and help-seeking behaviors. Furthermore, there is a lack of studies assessing the efficacy of PNF in addressing emerging online gambling-like behaviors, such as betting within video games or excessive financial trading of high-risk assets (eg, cryptocurrencies), which have been linked to gambling disorder.

Objective: This study aims to (1) quantify intention to change and the prevalence of help-seeking behaviors in young adults in Spain with disordered gambling or trading behaviors and (2) assess the efficacy of online PNF in increasing these behaviors.

Methods: A randomized controlled trial using a Solomon 3-group design was conducted with a sample of emerging adults aged 18 to 34 years. The study included 3 assessments: a pretest, an immediate posttest, and a 12-week follow-up assessment. Participants were randomized into 1 of 3 conditions, with the intervention group receiving online PNF. The primary outcomes were intention to change and help-seeking behaviors. Secondary outcomes included gambling and trading behaviors (intensity, frequency, and severity) and their longitudinal trajectories. Individual, interpersonal, and contextual factors will be assessed to identify the profile of individuals most likely to benefit from this intervention.

Results: The study was funded in December 2023 by the Spanish Ministry of Social Rights, Consumer Affairs and 2030 Agenda (SUBV23/00004) and approved by the Ethics Committee of the Miguel Hernández University. The first study assessment was conducted between December 2024 and January 2025. A total of 1889 people completed the eligibility assessment, of whom 1112 (58.9%) met the inclusion criteria (gambling or trading within the past 60 days) and were enrolled in the trial. The 12-week follow-up was finalized in March 2025, with 666 completing the final assessment (59.9% retention). Data collection has been completed, statistical analyses are ongoing, and primary results are expected to be published in March 2026.

Conclusions: By examining motivational outcomes rather than behavioral change alone, this trial addresses a key gap in the literature on digital interventions for gambling-related harm. Findings are expected to inform the development of scalable prevention and early intervention strategies targeting gambling and gambling-like behaviors in young adults.

Trial Registration: ClinicalTrials.gov NCT06681103; https://clinicaltrials.gov/study/NCT06681103

International Registered Report Identifier (IRRID): DERR1-10.2196/73155

(JMIR Res Protoc 2026;15:e73155) doi:10.2196/73155

KEYWORDS

young adults; gambling; trading; gambling disorder; trading disorder; brief intervention; normative feedback; clinical trial; study protocol

XSL•FO
RenderX

Introduction

Background

The digitalization of gambling markets has substantially transformed the gambling landscape, contributing to increased participation in online gambling and heightened exposure to gambling-related harms [1,2]. A recent systematic review and meta-analysis of representative samples across 68 countries found that 46.2% of adults gambled in the past 12 months, with 8.7% engaging in any risky gambling behavior, and that the prevalence of problem gambling was the highest among individuals participating in online casino or slot machine gambling (15.8%) [3]. In Spain, online gambling participation has tripled since 2013, and the prevalence of problem gambling among individuals who gamble online is also markedly higher than among those who gamble exclusively offline (18.4% and 4.3%, respectively) [4]. These patterns are particularly pronounced among young adults, an age group characterized by high exposure to digital environments and increased vulnerability to addictive behaviors [1,4-6].

Beyond traditional gambling products, digitalization has facilitated the expansion of gambling-like products that encourage monetary spending across multiple online contexts [2,6-9]. Within digital environments, gambling-like practices have been most extensively studied in the context of video games. In particular, loot boxes, mechanisms that involve spending real money on randomized virtual items or rewards, have been consistently associated with increased gambling severity and gambling-related harms among adolescents and young adults [10-14]. More recently, attention has also been drawn to skin gambling, which involves wagering of in-game items with real-world monetary value and has been suggested in recent studies to have stronger associations with gambling severity than loot boxes [6,15,16].

While younger adolescents (11 - 14 years) represent the largest group of video game users, recent European data indicate that the greatest growth in gaming participation has occurred among individuals aged 25 to 34 years, underscoring the relevance of this age group for research on gambling-like digital spending [17].

In parallel, this same cohort has become increasingly active on online platforms and mobile apps that allow speculative trading of financial assets, such as cryptocurrencies and other high-risk instruments [7,18,19]. Unlike traditional long-term investing, speculative trading is characterized by short-term profit seeking, frequent transactions, and exposure to highly volatile markets, features that closely resemble those of gambling activities [9,20-23]. Moreover, trading platforms often incorporate *gamblified* design elements, such as frequent push notifications or visual cues emphasizing gains, which may reinforce risky decision-making and promote these activities as a form of entertainment [24-28].

An accumulating body of research indicates that speculative trading is associated with problem gambling, impulsivity, gambling-related cognitive distortions, and other established risk factors for gambling disorder [23,29-35]. Evidence also suggests a substantial overlap between gambling and trading behaviors. Previous studies have found that a large proportion of regular gamblers engage in financial trading, and that problem gambling severity is the strongest predictor of intense and maladaptive trading behaviors [32,34-36]. Importantly, rather than the use of specific financial instruments alone, the combination of speculative trading with multiple forms of gambling, particularly in online contexts, appears to identify subgroups with elevated risk profiles and greater gambling-related harm [23].

Despite this growing evidence base, gambling and trading behaviors are rarely examined jointly within intervention research, and trading is typically not systematically considered within prevention and treatment frameworks targeting gambling-related harms. A further challenge in addressing gambling-related problems is the extremely low rate of help seeking among affected individuals. A recent meta-analysis estimated that only 0.25% of people with gambling problems have ever sought formal help, a pattern observed consistently across types of gambling, countries, and treatment formats [37]. Available Spanish data on help seeking are largely derived from clinical samples already receiving specialized treatment and may therefore underestimate unmet need in the general population [38]. Commonly reported barriers to treatment include low motivation for change, stigma, and a preference for self-managing gambling-related difficulties [39-41]. These barriers are especially relevant for young adults, who are less likely to access traditional treatment services despite experiencing elevated levels of gambling-related risk [42-44].

To overcome these limitations, internet-based and self-guided interventions have been increasingly developed to reach individuals who are not in contact with specialized services [45-48]. Personalized normative feedback (PNF) is one of the most widely studied brief motivational interventions in this context, and it constitutes the main type of intervention used in the prevention of problem gambling among young adults [5]. As a core component of motivational interviewing approaches, PNF provides individualized feedback that compares an individual's behavior with the observed behavior of a relevant reference group, such as peers, with the aim of highlighting discrepancies, fostering self-reflection, and supporting motivation to make informed decisions about problematic behavior [49].

Although meta-analytic evidence indicates that PNF produces only small reductions in gambling frequency and severity, these modest behavioral effects suggest that its primary utility may lie in enhancing motivation to change and promoting help seeking, rather than serving as a stand-alone behavioral intervention [50]. However, most existing PNF studies have focused on behavioral outcomes, with comparatively little attention to motivational end points such as intention to change or engagement with support resources. Moreover, PNF interventions have rarely been examined in relation to emerging gambling-like behaviors, including speculative trading or other forms of digital monetary spending. This gap is particularly relevant given that cryptocurrency trading is increasingly reported in gambling helpline contacts [51], underscoring the

XSL•FO
RenderX

need for low-threshold interventions capable of addressing this broader risk landscape.

Aims of the Project

The objective of this project was 2-fold: (1) to explore intention to change and help-seeking behavior among young adults with gambling and/or trading problems, given the heightened vulnerability of this age group [5,6], and (2) to evaluate the efficacy of an online PNF intervention in increasing intention to change and help-seeking behaviors in this population. As a secondary objective, the study examined whether exposure to PNF was associated with changes in gambling and trading behaviors (eg, frequency and intensity) over the follow-up period. Finally, the study compared the effects of PNF between individuals who engage in gambling or trading without meeting the criteria for a disorder (selective prevention) and those meeting the criteria for a gambling or trading disorder (indicated prevention).

Methods

Ethical Considerations

The study procedures were conducted in accordance with the Declaration of Helsinki and were approved by the Committee of Research and Ethics of the Miguel Hernández University of Elche (DPS.ACC.150523). All participants provided informed consent before participation. Data were collected through a secure web-based platform and were provided to the research team in anonymized form by the panel provider, with no access to personally identifiable information, thereby ensuring participant privacy and confidentiality. Participants received nonmonetary incentives in the form of redeemable points, in accordance with the panel provider's standard procedures. The study protocol was prospectively registered in ClinicalTrials.gov (NCT06681103) on November 7, 2024, before participant enrollment.

Study Design

This study was designed in accordance with SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines for clinical trial protocols (refer to the SPIRIT checklist in Checklist 1). A randomized controlled trial based on an adapted Solomon 3-group design was conducted, consisting of 2 control groups and 1 intervention condition. This design structure allows for the independent assessment of the immediate effects of the intervention on the primary motivational outcomes (ie, intention to change and help-seeking behaviors) while controlling for potential pretest sensitization (testing effects).

The study included a pretest, an immediate posttest (administered directly after the intervention or initial assessment where applicable), and a 12-week follow-up assessment to evaluate the maintenance of effects 3 months after the intervention. Although the classic Solomon design includes a fourth group that receives the intervention without a pretest [52], a pretest assessment is technically required in this study to generate the PNF for each participant. Accordingly, the adaptation proposed by Moreira et al [53] for studies applying PNF was adopted, omitting the intervention group with only a posttest.

Within the single intervention group (PNF), participants were randomly assigned to 1 of 2 predefined feedback assessment sequences (PNF-A vs PNF-B). Both subgroups received identical PNF content and completed the same primary and secondary outcome measures. The distinct feature between PNF-A and PNF-B lay in the timing of the postintervention assessment of gambling and trading behaviors (ie, frequency and intensity). In PNF-A, all posttest measures were completed after receiving PNF, whereas in PNF-B, participants completed posttest assessment of gambling and trading behaviors before receiving PNF and subsequently completed the posttest assessment of primary motivational outcomes after PNF.

This sequencing was intended to support secondary analyses examining potential reporting reactivity (eg, social desirability) in self-reported behavioral outcomes [54] and involved no alteration to the interventions' delivery, content, or the assessment of the primary outcomes.

Accordingly, the study comprised three experimental conditions, with an additional within-intervention assessment sequencing: (1) a control group with pretest and posttest assessments (assessment group); (2) a control group with posttest assessment only (delayed assessment group); and (3) an intervention group receiving PNF, with pretest and posttest assessments and randomized assessment sequencing (PNF-A vs PNF-B; Figure 1).

XSL•FO
RenderX

Figure 1. Flowchart of the adapted Solomon 3-group randomized controlled trial design, illustrating the assessment sequence across groups, including the intervention condition with randomized assessment sequencing (personalized normative feedback [PNF]-A and PNF-B). O1–O11 indicate assessment time points, and X denotes the intervention (PNF).

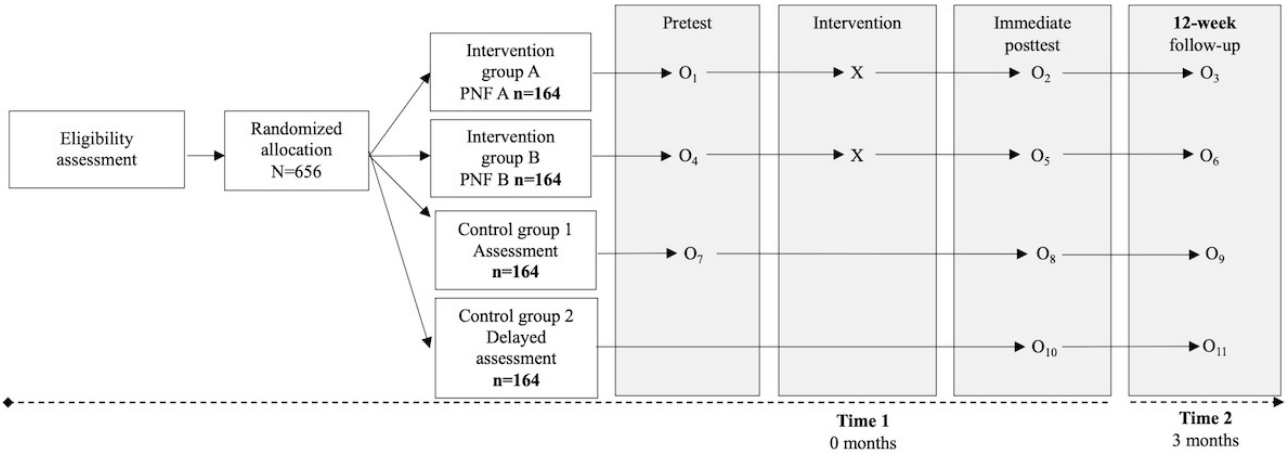

Participant Recruitment and Randomization

Recruitment targeted emerging adults aged 18 to 34 years. This specific age range was selected as it represents the demographic with the highest participation rates in the target behaviors, particularly in high-risk financial assets (eg, cryptocurrencies) and online gambling products [4,7]. In addition, this period of early adulthood is characterized by heightened exposure to digital environments and the adoption of novel gambling-like practices, which are less prevalent in older age groups and have been repeatedly examined in prior gambling research focusing on young and emerging adults [5,6,18,55].

Given the specific online nature of the behaviors assessed and the objectives of the project, recruitment was conducted in an online setting using an online panel managed by an independent Spanish agency specialized in ad hoc quantitative and qualitative research. Although nonprobability-based, this method was prioritized to effectively reach a geographically diverse group of young adults, as online panels offer the advantage of reaching a wider and more diverse group of young people compared with data collection in more specific contexts, such as university settings [56]. To contextualize sample representativeness, sociodemographic characteristics were benchmarked against a prior Spanish study using random digital sampling methods [7].

Participants' eligibility was assessed in an initial screening, which included sociodemographic variables and measures of psychological distress, suicidal behavior, substance use, and gambling and trading participation in the past 60 days. Participants who met the inclusion criteria proceeded to the randomization stage. Eligible participants were those who (1) reported engaging in trading or gambling activities (land-based or online) at least once per month over the past 60 days; (2) were aged between 18 and 34 years; (3) resided in Spain; and (4) provided informed consent to participate in initial online assessment, posttest, and follow-up. Exclusion criteria included (1) currently undergoing psychological or psychiatric treatment at the time of assessment and (2) scoring 2 or more incorrect responses on the Oviedo Infrequency Scale-Revised (E Fonseca-Pedrero et al, unpublished document, 2008) indicating potentially random or dishonest responding.

Participant allocation to each condition was conducted via a computer-generated allocation algorithm integrated into the online survey platform. To ensure equal sizes across the 4 arms, the algorithm used a least-fill dynamic allocation method. This procedure automatically assigned each incoming participant to the experimental condition with the fewest enrolled participants at the time of entry, ensuring a balanced distribution of total participants across groups independent of baseline characteristics. In cases where 2 or more conditions had the same lowest number of participants, allocation was determined by simple random selection among those conditions.

Sample Size Estimation

Given the lack of prior studies examining the impact of online PNF on intention to change and help seeking among individuals with gambling or trading problems, sample size estimation was based on previous trials evaluating PNF effects on gambling-related outcomes in young adults. Assuming an α error of 5%, an effect size f=0.15, and a correlation between baseline and follow-up measures of 0.5 [39,57], it was estimated that a total of 164 participants would be required to achieve a statistical power (1–β) of 90% for detecting differences within and between groups on a single outcome.

Given the evaluation of multiple outcomes, a minimum effective sample size of 328 was required. To account for potential attrition, the effective sample was adjusted using the following formula: final sample size=effective sample size/(1–loss to follow-up rate). Given the online panel setting, a conservative loss-to-follow-up rate of 50% was assumed before study launch. This adjustment yielded a target enrollment of 656 participants in total, corresponding to a minimum of 164 (25%) participants per randomized condition at baseline.

Blinding

A double-blind design was used, wherein participants were unaware of their assigned experimental condition and were blinded to the study hypotheses regarding the efficacy of the intervention. In accordance with recommendations by Boutron et al [58], the outcome evaluation was conducted with researchers blinded to allocation to minimize the risk of detection bias.

XSL•FO
RenderX

Intervention

The online PNF intervention provided participants with a brief report summarizing the results of their baseline assessment. This PNF included (1) an overview of the gambling and trading activities in which the participant engaged, including frequency of engagement and amounts spent; and (2) assessment scores on disordered trading (Trading Disorder Scale) and gambling (Problem Gambling Severity Index), along with interpretative information indicating levels of severity. Each result was accompanied by graphical representations, allowing participants to compare their scores with population-based norms from the Spanish demographic of the same sex and age group [3].

To mitigate the potential for increased gambling or trading behaviors in individuals whose current engagement levels are below the normative average, we incorporated the recommendations put forth by Newall et al [59] by presenting the normative data in conjunction with extreme values and relevant percentages.

Eligibility Assessment

The eligibility assessment was conducted at baseline and included sociodemographic information such as sex, age, education, employment status, and monthly income. Psychological distress was assessed using the Kessler Psychological Distress Scale [60], which measures symptoms of nervousness, hopelessness, depression, and feelings of worthlessness over the past 4 weeks. A single item assessed current or past psychiatric or psychological treatment. Suicidal behavior was evaluated using the 5-item Paykel Suicide Scale [61] in its Spanish version [62], capturing death thoughts, ideation, and suicide attempts in the past year.

Substance use indicators were collected using items adapted from the Encuesta sobre Alcohol y otras Drogas en España (EDADES) survey [63]. Alcohol risk consumption was assessed using the 3-item version of the Alcohol Use Disorders Identification Test–Consumption [64]. Gambling and trading activities during the past 60 days were identified via ad hoc items based on the EDADES survey [63], capturing engagement in traditional gambling formats (eg, bingo, lotteries, and casino), emerging gambling-like activities (eg, buying loot boxes, e-sports betting, and skin betting), and speculative trading (eg, cryptocurrency trading, foreign exchange, and commodities).

Primary Outcomes

Intention to change gambling or trading behavior was assessed through 3 ad hoc items. Participants answered a dichotomous item (yes or no) adapted to their chosen behaviors during the screening phase (eg, “In the last year, have you thought about reducing or quitting [financial asset trading/gambling]?”). Participants then reported the likelihood of engaging in gambling or trading over the next 12 months on a scale from 0% (very unlikely) to 100% (very likely). A third item asked participants to specify their actual plan regarding their behavior (not reducing or stopping, reducing or stopping in the following 6 months, and reducing or stopping in the following 30 days), with the latter 2 items administered at baseline, postintervention assessment, and follow-up.

Intention to seek treatment was also measured (at pretest, posttest, and at 12-week follow-up) using a likelihood item (0% - 100%). Help-seeking behaviors were additionally assessed at baseline using a question adapted from the study by Rodda et al [65] about prior help seeking (eg, “Have you ever sought professional help or advice, support from family or friends, or taken any self-directed actions to reduce your [gambling/trading] behavior?”). If answered affirmatively, participants completed an adaptation of the Help-Seeking Questionnaire [65], which includes items assessing self-directed, distance-based, and face-to-face help-seeking behaviors (eg, reading forums, browsing informational websites, and talking to a gambling counselor).

Secondary Outcomes

Gambling and trading behaviors were assessed for frequency and intensity using items adapted from the EDADES survey [63] to capture gambling frequency and spending in both online and offline contexts. Trading behavior was assessed with items adapted from the study by Delfabbro et al [29], covering trading frequency, trading types, and frequency of monitoring financial markets. In the baseline and posttest assessments, these behaviors were measured retrospectively for the past year, while the follow-up covered the previous 3 months. Amateur or professional investment use was assessed with a dichotomous item (yes or no) regarding professional trading, as well as items on the use of broker platforms or apps.

Problem gambling was assessed using the 9-item Spanish version of the Problem Gambling Severity Index (PGSI) [66,67], categorizing participants into 4 groups: nonproblematic (PGSI=0), low-risk (PGSI=1 - 2), moderate-risk (PGSI=3 - 7), and problem gambling (PGSI≥8). Additionally, the Gambling Disorder Identification Test (GDIT) [68,69] was administered at follow-up as part of a planned validation process. The GDIT is a 14-item Likert scale instrument evaluating gambling behaviors, symptoms, and negative consequences. This scale allows for the classification of participants into the following categories: nongambling, recreational gambling, problem gambling, and pathological gambling (mild, moderate, or severe).

Disordered trading was measured with the Trading Disorder Scale [34], a validated 13-item instrument assessing problematic trading behaviors. The scale is based on research criteria proposed by Guglielmo et al [70], which are derived from selected *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*) criteria for gambling disorder and internet gaming disorder [71] to conceptualize trading-related problems. This scale captures key aspects of behavioral addiction, such as tolerance, withdrawal, loss of control, craving, social impairment, and persistence despite negative consequences [34].

Other Measures

Gambling- and trading-related harms were measured using the Short Gambling Harm Screen [72], a brief 10-item binary-response tool adapted to capture financial, emotional or psychological, and relational harms from gambling and trading. Perceived self-efficacy to resist gambling and trading urges was

XSL•FO
RenderX

measured using an adaptation of the Brief Situational Confidence Questionnaire [73], which has been psychometrically validated for gambling-related behaviors in an adult general population sample [74]. This adapted version includes 2 additional items assessing confidence in resisting the urge to gamble when faced with situations involving financial pressure or the influence of alcohol or drugs, complementing the original 8 items focused on unpleasant emotions, physical discomfort, pleasant emotions, social pressure, and tests of control. Peer and family gambling and trading behaviors were assessed through ad hoc items. These items inquired about friends' and family members' involvement in financial trading and gambling. Finally, a 5-point Likert scale was used to evaluate satisfaction with and usefulness of the online PNF intervention (only in intervention groups).

Data Analysis Plan

The data of the project will be made available upon reasonable request, supporting adherence to Open Science principles and ensuring transparency throughout the research process. Data will be analyzed using SPSS (version 28.0; IBM Corp), Mplus (version 8.7; Muthén & Muthén), and R (R Foundation for Statistical Computing). Descriptive analyses and group comparisons will be performed using chi-square tests for categorical variables and 2-tailed Student *t* test for independent samples, and ANOVA (including multivariate analysis of variance [MANOVA] and multivariate analysis of covariance [MANCOVA]) for continuous variables with a 95% confidence level. Data distribution will be tested with Kolmogorov-Smirnov and Shapiro-Wilk tests. If the assumption of normality is not met, nonparametric alternatives will be used. Effect sizes will be reported using appropriate statistics (eg, odds ratio, Cohen *d*, Rosenthal r, omega squared, or Cramer V). To control for type I error due to multiple comparisons, the Bonferroni correction will be applied. Results will be weighted by age, gender, and region.

To examine gambling and trading trajectories, primary longitudinal analyses will focus on generalized linear mixed models, which are preferred over traditional repeated-measures ANOVA due to their ability to handle missing data and accommodate nonnormal outcome distributions. Mixed latent class analyses will be used to identify latent subpopulations characterized by different patterns of gambling and trading intensity and severity over time. These analyses will be considered exploratory and contingent on data structure and model fit.

Mediation and moderation analyses will be conducted as secondary and exploratory analyses to examine potential mechanisms and conditional effects of the PNF intervention. These analyses will be implemented using the PROCESS macro for SPSS and structural equation modeling in Mplus to estimate direct and indirect effects, with continuous variables centered to reduce multicollinearity [75]. Bias-corrected 95% CIs will be generated from 10,000 bootstrap random samples.

During the evaluation of the PNF, observed variables will be controlled to ensure balanced baseline characteristics between control and experimental groups, thus confirming randomization integrity. Additionally, a comparative analysis will be conducted between participants who meet inclusion criteria (ie, engaging in gambling or trading activities in the past 60 days) and those who do not, assessing differences in sociodemographic variables, psychological distress, suicidal behavior, substance use, and gambling participation reported during eligibility assessment.

Questionnaires not previously validated in the Spanish population or for the constructs of interest will undergo psychometric evaluation, including item analysis, reliability, and validity testing. Item analysis will include item-test correlations, discarding items with discrimination indices below 0.30 [76]. Reliability analysis will be assessed using omega coefficients; a total omega ≥0.70 will be considered acceptable [77], while hierarchical omega will require a minimum of 0.50 per subscale [78]. Dimensionality and validity will be assessed through confirmatory factor analysis. Goodness-of-fit tests will be conducted, using the following indices and interpretations for the assessment of model fit: the comparative fit index, the relative goodness-of-fit index, the root mean square error, the Tucker-Lewis index, and the chi-square statistic corrected for degrees of freedom. A satisfactory model fit will be considered when the comparative fit index, Tucker-Lewis index, and relative goodness-of-fit index values are greater than 0.90; the root mean square error value is less than or equal to 0.05; and the chi-square statistic corrected for degrees of freedom values is close to 0 [79]. Measurement invariance will be tested across age, sex, and gambling or trading activity subgroups to confirm the stability of the factor structure across groups.

Results

The study was funded in December 2023. Baseline data collection was conducted between December 2024 and January 2025. A total of 1889 individuals completed the eligibility assessment, of whom 1112 (58.9%) met the inclusion criteria and were enrolled in the trial. The 12-week follow-up was finalized in March 2025, with 666 (59.9% retention) participants completing the final assessment. As of January 2026, data collection has been completed, and statistical analyses are ongoing. Primary results are expected to be published in March 2026.

Discussion

This study aims to address critical gaps in understanding emerging gambling and trading behaviors among young adults, a population that is increasingly involved in such activities, which have well-documented associations with psychological distress and gambling-related harm. Although generalizability may be influenced by the recruitment methodology, the online design of the study is particularly well suited to examining behaviors that occur primarily in digital contexts (eg, cryptocurrency trading and skin betting) [29,56,80] and enables timely and efficient assessment of motivational and behavioral outcomes over time [81].

To ensure data quality within this digital framework, sample characteristics were benchmarked against national reference data to contextualize representativeness [7]. Furthermore, to address limitations associated with self-reports, including social

XSL•FO
RenderX

desirability bias, validated self-administered questionnaires and anonymized assessments were prioritized, and the Oviedo Infrequency Scale was included to detect and exclude random or dishonest responses. The use of multiple instruments for each measured variable will also allow for cross-verification, improving the reliability of the results. A key methodological strength is the use of an adapted Solomon 3-group experimental design, which enhances internal validity by allowing for the independent isolation of intervention effects from potential pretest sensitization.

This study represents a pioneering application of online PNF by extending it to the context of gambling-like practices, such as high-risk trading. To the best of our knowledge, this is one of the first randomized trials to assess the efficacy of PNF in increasing intention to change and help-seeking behaviors specifically within these emerging high-risk digital behaviors.

By examining the profiles of individuals who exhibit low motivation to seek help, the study also intends to provide insights that can contribute to the refinement of treatment and prevention strategies. However, several limitations should be acknowledged. First, longitudinal studies using online panels are susceptible to high attrition rates. To mitigate this, a conservative oversampling strategy (assuming 50% loss to follow-up) was used, and missing data will be handled via generalized linear mixed models. Second, as the sample is restricted to Spanish residents, cross-cultural generalizability may be limited. Finally, given the novelty of behaviors such as cryptocurrency trading, some measures were adapted ad hoc. Rigorous psychometric analyses are planned to validate these instruments as part of the study's contributions.

Despite these limitations, findings may inform regulatory policies for trading platforms and gambling-like activities within video games to better protect vulnerable users. Excessive engagement in these activities has been associated with other addictive behaviors, such as disordered gambling and substance use, as well as emotional difficulties [82]. Furthermore, increased engagement in financial markets has been linked to rising suicide rates among both amateur and professional investors [70]. Consequently, the outcomes of the study are expected to advance public health by elucidating the psychological and social consequences of these converging behaviors and supporting the development of effective, scalable interventions aimed at reducing harm and facilitating help seeking in this population.

Funding

This project has been funded by the Spanish Ministry of Social Rights, Consumer Affairs and 2030 Agenda (SUBV23/00004). The funder has no role in the study design, conduct, or reporting.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AC-C was the principal investigator for this project and had overall responsibility for the trial, including leading the design of all aspects of the study. JLC provided mentorship on the study design and the selection of measures. JLC, FM-L, and VP-J were investigators on the study and contributed to the grant application. AC-C drafted the initial manuscript, and all authors contributed to subsequent revisions. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT 2025 checklist.
[PDF File, 183 KB - resprot_v15i1e73155_app1.pdf]

References

1. Wardle H, Degenhardt L, Marionneau V, et al. The Lancet Public Health Commission on gambling. Lancet Public Health 2024 Oct 16;9(11):e950-e994. [doi: 10.1016/S2468-2667(24)00167-1] [Medline: 39491880]
2. Guillou-Landreat M, Gallopel-Morvan K, Lever D, Le Goff D, Le Reste JY. Gambling marketing strategies and the internet: what do we know? A systematic review. Front Psychiatry 2021;12:583817. [doi: 10.3389/fpsyt.2021.583817] [Medline: 33716809]
3. Tran LT, Wardle H, Colledge-Frisby S, et al. The prevalence of gambling and problematic gambling: a systematic review and meta-analysis. Lancet Public Health 2024 Aug;9(8):e594-e613. [doi: 10.1016/S2468-2667(24)00126-9] [Medline: 39025095]
4. Spanish Observatory on Drugs and Addictions. Report on behavioral addictions and other addictive disorders 2024: admissions to treatment for behavioral addictions indicator. : Government Delegation for the National Plan on Drugs; 2025

XSL•FO
RenderX

URL: https://pnsd.sanidad.gob.es/profesionales/sistemasInformacion/sistemaInformacion/pdf/2024_OEDA_AdiccionesComportamentales_ExecutiveSummary.pdf [accessed 2026-01-19]
5. Grande-Gosende A, López-Núñez C, García-Fernández G, Derevensky J, Fernández-Hermida JR. Systematic review of preventive programs for reducing problem gambling behaviors among young adults. J Gambl Stud 2020 Mar;36(1):1-22. [doi: 10.1007/s10899-019-09866-9] [Medline: 31168687]
6. Wardle H, Tipping S. The relationship between problematic gambling severity and engagement with gambling products: longitudinal analysis of the Emerging Adults Gambling Survey. Addiction 2023 Jun;118(6):1127-1139. [doi: 10.1111/add.16125] [Medline: 36606732]
7. Coloma-Carmona A, Carballo JL, Miró-Llinares F, Aguerri JC. Place your investments or place your bets? Relationship between gambling addiction and trading. The Social Observatory - La Caixa Foundation. 2024. URL: https://elobservatoriosocial.fundacionlacaixa.org/ [accessed 2026-01-19]
8. Brock T, Johnson M. The gamblification of digital games. J Consum Cult 2021 Feb;21(1):3-13. [doi: 10.1177/1469540521993904]
9. Newall PW, Weiss-Cohen L. The gamblification of investing: how a new generation of investors is being born to lose. Int J Environ Res Public Health 2022 Apr 28;19(9):5391. [doi: 10.3390/ijerph19095391] [Medline: 35564790]
10. Garea SS, Drummond A, Sauer JD, Hall LC, Williams MN. Meta-analysis of the relationship between problem gambling, excessive gaming and loot box spending. Int Gambl Stud 2021 Sep 2;21(3):460-479. [doi: 10.1080/14459795.2021.1914705]
11. Griffiths MD. Is the buying of loot boxes in video games a form of gambling or gaming? Gaming Law Rev 2018 Feb;22(1):52-54. [doi: 10.1089/glr2.2018.2216]
12. Raneri PC, Montag C, Rozgonjuk D, Satel J, Pontes HM. The role of microtransactions in internet gaming disorder and gambling disorder: a preregistered systematic review. Addict Behav Rep 2022 Jun;15:100415. [doi: 10.1016/j.abrep.2022.100415] [Medline: 35434248]
13. González-Cabrera J, Basterra-González A, Montiel I, Calvete E, Pontes HM, Machimbarrena JM. Loot boxes in Spanish adolescents and young adults: relationship with internet gaming disorder and online gambling disorder. Comput Human Behav 2022 Jan;126:107012. [doi: 10.1016/j.chb.2021.107012]
14. Zendle D. Problem gamblers spend less money when loot boxes are removed from a game: a before and after study of Heroes of the Storm. PeerJ 2019;7:e7700. [doi: 10.7717/peerj.7700] [Medline: 31681510]
15. Greer N, Rockloff M, Hing N, Browne M, King DL. Skin gambling contributes to gambling problems and harm after controlling for other forms of traditional gambling. J Gambl Stud 2023 Mar;39(1):225-247. [doi: 10.1007/s10899-022-10111-z] [Medline: 35217933]
16. Kim HS, Leslie RD, Stewart SH, et al. A scoping review of the association between loot boxes, esports, skin betting, and token wagering with gambling and video gaming behaviors. J Behav Addict 2023 Jun 29;12(2):309-351. [doi: 10.1556/2006.2023.00013] [Medline: 37159337]
17. All about video games – European key facts. Video Games Europe. 2023. URL: https://www.videogameseurope.eu/ [accessed 2026-01-19]
18. Delfabbro P, King D. The evolution of young gambling studies: digital convergence of gaming, gambling and cryptocurrency technologies. Int Gambl Stud 2023 Sep 2;23(3):491-504. [doi: 10.1080/14459795.2023.2171469]
19. Guzmán A, Pinto-Gutiérrez C, Trujillo MA. Trading cryptocurrencies as a pandemic pastime: COVID-19 lockdowns and bitcoin volume. Mathematics 2021;9(15):1771. [doi: 10.3390/math9151771]
20. Arthur JN, Williams RJ, Delfabbro PH. The conceptual and empirical relationship between gambling, investing, and speculation. J Behav Addict 2016 Dec;5(4):580-591. [doi: 10.1556/2006.5.2016.084] [Medline: 27929350]
21. Davies N, Ferris S. Cryptocurrency and new financial instruments: unquantified public health harms. Lancet Public Health 2022 Aug;7(8):e655. [doi: 10.1016/S2468-2667(22)00173-6] [Medline: 35907417]
22. Mosbey A, Delfabbro P, King D. The harmful consequences of cryptocurrency speculation and associated risk factors. Int J Ment Health Addiction 2024. [doi: 10.1007/s11469-024-01405-x]
23. Coloma-Carmona A, Carballo JL, Miró-Llinares F, C Aguerri J. Not all traders gamble, but some gamblers trade: a latent class analysis of trading and gambling behaviors among retail investors. Public Health (Fairfax) 2025 Jul;244:105742. [doi: 10.1016/j.puhe.2025.105742] [Medline: 40373541]
24. Andrade M, Newall PW. Cryptocurrencies as gamblified financial assets and cryptocasinos: novel risks for a public health approach to gambling. Risks 2023;11(3):49. [doi: 10.3390/risks11030049]
25. Macey J, Hamari J, Adam M. A conceptual framework for understanding and identifying gamblified experiences. Comput Human Behav 2024 Mar;152:108087. [doi: 10.1016/j.chb.2023.108087]
26. Mosenhauer M, Newall PW, Walasek L. The stock market as a casino: associations between stock market trading frequency and problem gambling. J Behav Addict 2021 Sep 28;10(3):683-689. [doi: 10.1556/2006.2021.00058] [Medline: 34587115]
27. Gao X, Lin TC. Do individual investors treat trading as a fun and exciting gambling activity? Evidence from repeated natural experiments. Rev Financ Stud 2015 Jul 1;28(7):2128-2166. [doi: 10.1093/rfs/hhu075]
28. Prachayanant P, Kraiwanit T, Chutipat V. Cryptocurrency gamification: having fun or making money. J Gov Regul 2023;12(2):184-193. [doi: 10.22495/jgrv12i2art17]

XSL•FO
RenderX

29. Delfabbro P, King D, Williams J, Georgiou N. Cryptocurrency trading, gambling and problem gambling. Addict Behav 2021 Nov;122:107021. [doi: 10.1016/j.addbeh.2021.107021] [Medline: 34171583]
30. Grall-Bronnec M, Sauvaget A, Boutin C, et al. Excessive trading, a gambling disorder in its own right? A case study on a French disordered gamblers cohort. Addict Behav 2017 Jan;64:340-348. [doi: 10.1016/j.addbeh.2015.12.006] [Medline: 26712561]
31. Granero R, Tárrega S, Fernández-Aranda F, et al. Gambling on the stock market: an unexplored issue. Compr Psychiatry 2012 Aug;53(6):666-673. [doi: 10.1016/j.comppsych.2011.12.004] [Medline: 22300903]
32. Loscalzo Y, Rogier G, Velotti P. Problematic trading: a systematic review of theoretical considerations. Front Psychiatry 2025;16:1505012. [doi: 10.3389/fpsyt.2025.1505012] [Medline: 40364998]
33. Coloma-Carmona A, Carballo JL, Sancho-Domingo C, Rodríguez-Espinosa S, Miró-Llinares F, C Aguerri J. Engagement in new gambling practices and its association with gambling disorder, impulsivity, and cognitive distortions: findings from a nationally representative sample of Spanish gamblers. J Gambl Stud 2025 Sep;41(3):1211-1227. [doi: 10.1007/s10899-024-10361-z] [Medline: 39373778]
34. Coloma-Carmona A, Carballo JL, Miró-Llinares F, Aguerri JC, Griffiths MD. Development and validation of the trading disorder scale for assessing problematic trading behaviors. J Behav Addict 2025 Jul 2;14(2):941-958. [doi: 10.1556/2006.2025.00019] [Medline: 40126562]
35. Johnson B, Co S, Sun T, et al. Cryptocurrency trading and its associations with gambling and mental health: a scoping review. Addict Behav 2023 Jan;136:107504. [doi: 10.1016/j.addbeh.2022.107504] [Medline: 36174424]
36. Leslie RD, Shaw CA, McGrath DS. Correlates of gambling behaviours among day traders: evidence from a national study. J Gambl Stud 2025 Mar;41(1):51-66. [doi: 10.1007/s10899-024-10314-6] [Medline: 38744751]
37. Bijker R, Booth N, Merkouris SS, Dowling NA, Rodda SN. Global prevalence of help-seeking for problem gambling: a systematic review and meta-analysis. Addiction 2022 Dec;117(12):2972-2985. [doi: 10.1111/add.15952] [Medline: 35830876]
38. Directorate General for the Regulation of Gambling. Study and analysis of risk factors for gambling disorder in a Spanish clinical population. Ministry of Finance and Civil Service. 2017. URL: https://www.ordenacionjuego.es [accessed 2026-01-19]
39. Cunningham JA, Hodgins DC, Toneatto T. Pilot study of an Internet-based personalized feedback intervention for problem gamblers. J Gambl Iss 2011;26(26):4. [doi: 10.4309/jgi.2011.26.2]
40. Menchon JM, Mestre-Bach G, Steward T, Fernández-Aranda F, Jiménez-Murcia S. An overview of gambling disorder: from treatment approaches to risk factors. F1000Res 2018;7:434. [doi: 10.12688/f1000research.12784.1] [Medline: 30090625]
41. Wall H, Molander O, Marionneau V. Help-seeking and help expectations for gambling problems at different severity levels. Discov Public Health 2025;22(1):447. [doi: 10.1186/s12982-025-00853-7]
42. Ledgerwood DM, Arfken CL, Wiedemann A, Bates KE, Holmes D, Jones L. Who goes to treatment? Predictors of treatment initiation among gambling help-line callers. Am J Addict 2013 Jan;22(1):33-38. [doi: 10.1111/j.1521-0391.2013.00323.x] [Medline: 23398224]
43. Gainsbury S, Hing N, Suhonen N. Professional help-seeking for gambling problems: awareness, barriers and motivators for treatment. J Gambl Stud 2014 Jun;30(2):503-519. [doi: 10.1007/s10899-013-9373-x] [Medline: 23494244]
44. Luquiens A, von Hammerstein C, Benyamina A, Perney P. Burden and help-seeking behaviors linked to problem gambling and gaming: observational quantitative and qualitative analysis. JMIR Ment Health 2021 Nov 26;8(11):e26521. [doi: 10.2196/26521] [Medline: 34842562]
45. Berge J, Abrahamsson T, Lyckberg A, Franklin K, Håkansson A. A normative feedback intervention on gambling behavior-a longitudinal study of post-intervention gambling practices in at-risk gamblers. Front Psychiatry 2022;13:602846. [doi: 10.3389/fpsyt.2022.602846] [Medline: 35432012]
46. Bücker L, Gehlenborg J, Moritz S, Westermann S. A randomized controlled trial on a self-guided Internet-based intervention for gambling problems. Sci Rep 2021 Jun 22;11(1):13033. [doi: 10.1038/s41598-021-92242-8] [Medline: 34158553]
47. Luquiens A, Tanguy ML, Lagadec M, Benyamina A, Aubin HJ, Reynaud M. The efficacy of three modalities of internet-based psychotherapy for non-treatment-seeking online problem gamblers: a randomized controlled trial. J Med Internet Res 2016 Feb 15;18(2):e36. [doi: 10.2196/jmir.4752] [Medline: 26878894]
48. Neighbors C, Rodriguez LM, Rinker DV, et al. Efficacy of personalized normative feedback as a brief intervention for college student gambling: a randomized controlled trial. J Consult Clin Psychol 2015 Jun;83(3):500-511. [doi: 10.1037/a0039125] [Medline: 26009785]
49. Carballo JL, Coloma-Carmona A. Brief intervention in addictive behaviors: promoting change. Pap Psicol 2025;46(2):76-85. [doi: 10.70478/pap.psicol.2025.46.11]
50. Saxton J, Rodda SN, Booth N, Merkouris SS, Dowling NA. The efficacy of personalized normative feedback interventions across addictions: a systematic review and meta-analysis. PLoS ONE 2021;16(4):e0248262. [doi: 10.1371/journal.pone.0248262] [Medline: 33793583]
51. Marionneau V, Kristiansen S, Wall H. Harmful types of gambling: changes and emerging trends in longitudinal helpline data. Eur J Public Health 2024 Apr 3;34(2):335-341. [doi: 10.1093/eurpub/ckae024] [Medline: 38389465]
52. Campbell DT, Stanley JC. Experimental and Quasi-Experimental Designs for Research: Houghton Mifflin; 1963.

XSL•FO
RenderX

53. Moreira MT, Oskrochi R, Foxcroft DR. Personalised normative feedback for preventing alcohol misuse in university students: solomon three-group randomised controlled trial. PLoS ONE 2012;7(9):e44120. [doi: 10.1371/journal.pone.0044120] [Medline: 22984466]
54. Cunningham JA, Wong HTA. Assessing the immediate impact of normative drinking information using an immediate post-test randomized controlled design: implications for normative feedback interventions? Addict Behav 2013 Jun;38(6):2252-2256. [doi: 10.1016/j.addbeh.2013.02.005] [Medline: 23507459]
55. Wardle H, Kesaite V, Tipping S, McManus S. Changes in severity of problem gambling and subsequent suicide attempts: a longitudinal survey of young adults in Great Britain, 2018-20. Lancet Public Health 2023 Mar;8(3):e217-e225. [doi: 10.1016/S2468-2667(23)00008-7] [Medline: 36841562]
56. Wardle H. The emerging adults gambling survey: study protocol. Wellcome Open Res 2020;5:102. [doi: 10.12688/wellcomeopenres.15969.1] [Medline: 32596512]
57. Martens MP, Arterberry BJ, Takamatsu SK, Masters J, Dude K. The efficacy of a personalized feedback-only intervention for at-risk college gamblers. J Consult Clin Psychol 2015 Jun;83(3):494-499. [doi: 10.1037/a0038843] [Medline: 25664642]
58. Boutron I, Guittet L, Estellat C, Moher D, Hróbjartsson A, Ravaud P. Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. PLoS Med 2007 Feb;4(2):e61. [doi: 10.1371/journal.pmed.0040061] [Medline: 17311468]
59. Newall PW, Rockloff M, Hing N, et al. Designing improved safer gambling messages for race and sports betting: what can be learned from other gambling formats and the broader public health literature? J Gambl Stud 2023 Jun;39(2):913-928. [doi: 10.1007/s10899-023-10203-4] [Medline: 36961657]
60. Kessler RC, Andrews G, Colpe LJ, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [doi: 10.1017/s0033291702006074] [Medline: 12214795]
61. Paykel ES, Myers JK, Lindenthal JJ, Tanner J. Suicidal feelings in the general population: a prevalence study. Br J Psychiatry 1974 May;124:460-469. [doi: 10.1192/bjp.124.5.460] [Medline: 4836376]
62. Fonseca-Pedrero E, Pérez de Albéniz A. Evaluación de la conducta suicida en adolescentes: a propósito de la Escala Paykel de Suicidio. Pap Psicol 2020;41(1):106-115. [doi: 10.23923/pap.psicol2020.2928]
63. Encuesta sobre alcohol y otras drogas en españa (EDADES), 1995–2024. Ministry of Health. 2023. URL: https://pnsd.sanidad.gob.es/profesionales/sistemasInformacion/sistemaInformacion/encuestas_EDADES.htm [accessed 2026-01-19]
64. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol use disorders identification test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
65. Rodda SN, Dowling NA, Lubman DI. Gamblers seeking online help are active help-seekers: Time to support autonomy and competence. Addict Behav 2018 Dec;87:272-275. [doi: 10.1016/j.addbeh.2018.06.001] [Medline: 29935737]
66. Ferris J, Wynne H. The Canadian problem gambling index: final report. Canadian Centre on Substance Abuse. 2001. URL: https://www.greo.ca/Modules/EvidenceCentre/files/Ferris%20et%20al(2001)The_Canadian_Problem_Gambling_Index.pdf [accessed 2025-05-25]
67. Lopez-Gonzalez H, Estévez A, Griffiths MD. Spanish validation of the problem gambling severity index: a confirmatory factor analysis with sports bettors. J Behav Addict 2018 Sep 1;7(3):814-820. [doi: 10.1556/2006.7.2018.84] [Medline: 30264603]
68. Molander O, Volberg R, Sundqvist K, Wennberg P, Månsson V, Berman AH. Development of the Gambling Disorder Identification Test (G-DIT): protocol for a Delphi method study. JMIR Res Protoc 2019 Jan 8;8(1):e12006. [doi: 10.2196/12006] [Medline: 30622097]
69. Molander O, Wennberg P, Berman AH. The Gambling Disorders Identification Test (GDIT): psychometric evaluation of a new comprehensive measure for gambling disorder and problem gambling. Assessment 2023 Jan;30(1):225-237. [doi: 10.1177/10731911211046045] [Medline: 34617456]
70. Guglielmo R, Ioime L, Janiri L. Is pathological trading an overlooked form of addiction? Addict Health 2016 Jul;8(3):207-209. [Medline: 28496959]
71. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Association; 2013.
72. Browne M, Goodwin BC, Rockloff MJ. Validation of the Short Gambling Harm Screen (SGHS): a tool for assessment of harms from gambling. J Gambl Stud 2018 Jun;34(2):499-512. [doi: 10.1007/s10899-017-9698-y] [Medline: 28578519]
73. Breslin FC, Sobell LC, Sobell MB, Agrawal S. A comparison of a brief and long version of the situational confidence questionnaire. Behav Res Ther 2000 Dec;38(12):1211-1220. [doi: 10.1016/s0005-7967(99)00152-7] [Medline: 11104185]
74. Coloma-Carmona A, Carballo JL, Dowling NA. Psychometric validation and measurement invariance of the Brief Situational Confidence Questionnaire for Gambling (BSCQ-G) in a sample of adult gamblers from the general population. Psychol Addict Behav 2025 Jun 5. [doi: 10.1037/adb0001076] [Medline: 40471818]
75. Aiken LS, West SG, Reno RR. Multiple Regression: Testing and Interpreting Interactions: Sage Publication; 1991.
76. Ebel RL, Frisbie DA. Essentials of Educational Measurement: Prentice-Hall; 1991.

XSL•FO
RenderX

77. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39-50. [doi: 10.1177/002224378101800104]
78. Watkins MW. The reliability of multidimensional neuropsychological measures: from alpha to omega. Clin Neuropsychol 2017;31(6-7):1113-1126. [doi: 10.1080/13854046.2017.1317364] [Medline: 28429633]
79. Yu CY. Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. : University of California; 2002 URL: https://search.gesis.org/publication/zis-Yu2002Evaluating [accessed 2025-05-25]
80. Oksanen A, Sirola A, Savolainen I, Kaakinen M. Gambling patterns and associated risk and protective factors among Finnish young people. Nordisk Alkohol Nark 2019 Apr;36(2):161-176. [doi: 10.1177/1455072518779657] [Medline: 32934557]
81. Callegaro M, Baker RP, Bethlehem J, Göritz AS, Krosnick JA, Lavrakas PJ. Online Panel Research: A Data Quality Perspective: John Wiley & Sons; 2014.
82. Oksanen A, Hagfors H, Vuorinen I, Savolainen I. Longitudinal perspective on cryptocurrency trading and increased gambling problems: a 3 wave national survey study. Public Health (Fairfax) 2022 Dec;213:85-90. [doi: 10.1016/j.puhe.2022.10.002] [Medline: 3639568]

Abbreviations

***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
EDADES: Encuesta sobre Alcohol y Drogas en España
GDIT: Gambling Disorder Identification Test
MANCOVA: multivariate analysis of covariance
MANOVA: multivariate analysis of variance
PGSI: Problem Gambling Severity Index
PNF: personalized normative feedback
PSS: Paykel Suicide Scale
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by J Sarvestan; submitted 26.Feb.2025; peer-reviewed by S Rodda; revised version received 18.Jan.2026; accepted 19.Jan.2026; published 23.Feb.2026.

Please cite as:
Coloma-Carmona A, Carballo JL, Miró-Llinares F, Pérez-Jover V
Online Personalized Normative Feedback to Foster Intention to Change and Help Seeking in Young Adults With Disordered Gambling and Trading Behaviors: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e73155
URL: https://www.researchprotocols.org/2026/1/e73155
doi:10.2196/73155

© Ainhoa Coloma-Carmona, José Luis Carballo, Fernando Miró-Llinares, Virtudes Pérez-Jover. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Diffusion Tensor Imaging in the Diagnosis of Perianal Abscess: Protocol for a Single-Blind Randomized Controlled Trial

Yuting Ma*, MMed; Pingping Mei*, MMed; Xiutian Guo, PhD; Yan Chen, PhD

Department of Anorectal, Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Lane 68, Ronglian Road, Jiading Area, Shanghai, China

*these authors contributed equally

Corresponding Author:
Yan Chen, PhD
Department of Anorectal, Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Lane 68, Ronglian Road, Jiading Area, Shanghai, China

Abstract

Background: Perianal abscesses are common anorectal conditions that often necessitate surgical intervention. Accurate preoperative assessment is crucial for effective treatment and reducing recurrence rates. Diffusion tensor imaging (DTI) is a valuable method for visualizing the degree of infection and infiltration, the extent of abscess formation, and the relationship between perianal abscess erosion. However, there is currently a lack of specific studies focusing on perianal abscesses.

Objective: The objective of this study is to investigate the diagnostic utility of DTI in the preoperative assessment of perianal abscesses. By enhancing the precision of preoperative evaluation, we aim to minimize damage to the anal sphincter, reduce the recurrence rate, and improve the prognosis and quality of life for patients.

Methods: This study adopts a randomized, prospective, single-blind design. Ninety participants are being randomized into 2 groups: a DTI group and a conventional magnetic resonance imaging group. A structured report is being completed based on imaging of the perianal abscesses in terms of location, number, specific pus cavity alignment, thickness, and relationship to the surrounding muscles. In addition, the patient condition is being assessed, and corresponding surgical treatment is being performed. If the patient's blood routine shows infection, cefuroxime sodium combined with metronidazole is being administered intravenously as an anti-infective therapy. Postoperatively, the primary variable is being assessed for recurrence rate at 6 months, and the secondary variables, including postoperative pain scores on days 3 and 7, peripheral blood inflammatory factors, and assessment of anal function, are being evaluated. Normally distributed continuous data will be presented as mean (SD) and analyzed using independent or paired *t* tests. Non-normally continuous data will be analyzed with rank-sum tests. Categorical data will be expressed as frequency (%) and compared using a chi-square test or appropriate nonparametric tests. Ordinal data will be analyzed using the Ridit test. A *P* value <.05 will be considered statistically significant.

Results: This study is funded by Science and Technology Commission of Shanghai Municipality Science and Technology Program (grant 23Y11920800). Patient recruitment was initiated in April 2025. As of January 2026, 37 participants have been enrolled, and data collection is scheduled to be completed in October 2026.

Conclusions: DTI technique can be used to gain a deeper understanding of the relationship between internal orifice, the degree of infected infiltration, the extent of the abscess, and the involvement of the perianal tissues and muscles in patients with perianal abscess. Deep pelvic floor DTI reveals the complex 3D structure of the pelvic floor in perianal abscesses through structured reports, which may provide new insights into the diagnosis of perianal abscesses.

Trial Registration: Itmctr.ccebtcm.org.cn ITMCTR2025000068; https://itmctr.ccebtcm.org.cn/mgt/project/view/6847137108505689154

International Registered Report Identifier (IRRID): DERR1-10.2196/83449

(JMIR Res Protoc 2026;15:e83449) doi:10.2196/83449

KEYWORDS

diffusion tensor imaging; magnetic resonance imaging; perianal abscess; single-blind method; pelvic floor

XSL•FO
RenderX

Introduction

Clinical Background and Diagnostic Challenges

Perianal abscesses are acute suppurative infections caused by obstruction of the anal gland opening and bacterial infection in the perianal tissues or spaces. They are among the most common anorectal diseases [1], particularly affecting young adults aged 20 to 40 years, with a higher prevalence in males than females [2]. Between 30% and 70% of patients with abscesses are often accompanied by anal fistula, while among patients without fistulas, the probability of developing an anal fistula is approximately 30% to 50% within months to years after abscess drainage [3]. Reducing recurrence rate and fistula rate of perianal abscesses remains a significant challenge. The recurrence rate after surgical intervention remains substantial, underscoring the need for improved diagnostic and therapeutic strategies [4].

The initial diagnosis of perianal abscesses is often based on clinical symptoms and anal palpation. However, this approach has limitations in dealing with multiple, high-grade perianal abscesses. In such cases, preoperative imaging, particularly magnetic resonance imaging (MRI), plays a pivotal role in accurate diagnosis and treatment [5,6]. MRI has a sensitivity of 96% and specificity of 97% for diagnosing abscesses [7]. Accurate assessment of the location, range, and depth of the abscess using MRI can help health care professionals develop more reasonable and effective surgical plans, including selecting the appropriate incision location, accurately positioning the internal opening, and employing the correct drainage method. This comprehensive approach aims to remove the infection as thoroughly as possible, reduce the risk of recurrence and fistula rate, and protect the anal function and quality of life of patients [8].

Diffusion-weighted imaging (DWI) is a variant of conventional MRI based on different rates of water diffusion in different tissues, offering unparalleled sensitivity to water motion within tissue structures [9,10]. Some studies have demonstrated the valuable role of DWI in the diagnosis and differential diagnosis of anal fistula and perianal abscess [11-13]. DWI exhibits high specificity in distinguishing abscesses from inflammatory masses, surpassing traditional MRI [12].

Diffusion tensor imaging (DTI) extends the principles of DWI by modeling water diffusion as a 3D tensor, thereby quantifying both the magnitude and directionality (anisotropy) of diffusion within tissues [14]. Originally developed for neural tractography, DTI has shown promise in musculoskeletal and pelvic floor imaging by revealing microstructural integrity and architectural disruption [15]. A pivotal study by Wang et al [16] demonstrated the feasibility of DTI in evaluating perianal fistula, suggesting that parameters such as fractional anisotropy (FA) and apparent diffusion coefficient could reflect tissue characteristics relevant to surgical planning. However, while this and other emerging studies hint at DTI's potential, the existing literature remains strikingly sparse regarding its specific application to perianal abscesses. Current evidence is largely extrapolated from fistula research or limited to technical feasibility reports. Critical gaps persist: there is a lack of standardized, quantitative DTI protocols for preoperatively mapping abscess extent, infection infiltration, and precise involvement of the anal sphincter complex and levator ani muscle. Moreover, it remains unclear whether the microstructural metrics provided by DTI correlate with surgical outcomes, such as recurrence rates or postoperative anal function.

Objectives of the Study

The primary goal of this study is to use DTI to delineate the relationship between the internal opening, the degree of infection and infiltration, the extent of the abscess, and the erosion of perianal tissues and muscles in patients with perianal abscesses. This approach aims to develop an accurate imaging evaluation method for diagnosing the disease, explore the application value of DTI in the diagnosis of perianal abscesses, and improve the accuracy of preoperative evaluation, as well as reduce the damage to the anal sphincter. Ultimately, these efforts aim to reduce the recurrence rate and improve the prognosis and quality of life for patients.

Methods

Study Design

This study adopts a randomized, controlled, single-blind design to compare the structural characteristics of perianal abscesses and the recurrence rate at 6 months after surgery between the DTI group and the conventional MRI group. Figure 1 illustrates the study flowchart. The schedule of enrollment, interventions, and assessments, in accordance with the SPIRIT (Standard Protocol Items Recommendations for Interventional Trials) guidelines, is summarized in Table 1.

XSL•FO
RenderX

Figure 1. Flowchart of the trial. DTI: diffusion tensor imaging; IL-1β: interleukin-1β; IL-6: interleukin-6; MRI: magnetic resonance imaging; TNF-α: tumor necrosis factor-α.

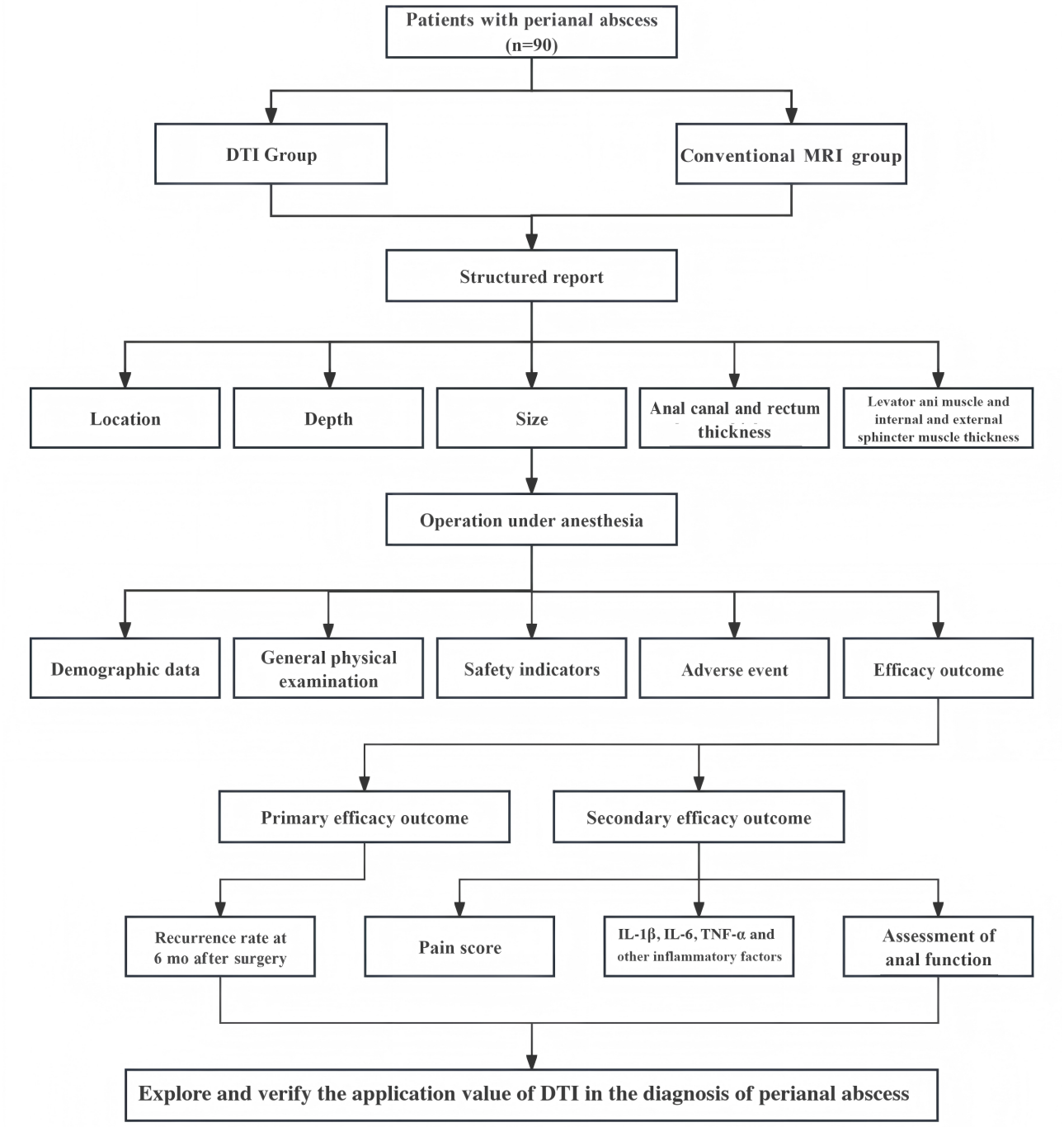

XSL•FO
RenderX

Table . The Standard Protocol Items Recommendations for Interventional Trials (SPIRIT) schedule table.

	Study period						
	Enrollment	Allocation	Postallocation				Close-out
Timepoint	$-t_1$	0	t_1 (preoperation)	t_2 (immediately after operation)	t_3 (the third day after the operation)	t_4 (the seventh day after the operation)	t_5 (6 mo after the operation)
Enrollment							
Eligibility screen	✓						
Informed consent	✓						
Allocation		✓					
Interventions							
Diffusion tensor imaging			✓				
Magnetic resonance imaging			✓				
Assessments							
Baseline variables: demographic data, vital signs testing, safety observation indicators			✓				
Recurrence rate							✓
Structured report				✓			
Postoperative pain score					✓	✓	
Detection of inflammatory factors			✓		✓		
Assessment of anal function			✓				✓

Randomization

In this study, 90 patients with perianal abscess are sequentially numbered from 1 to 90 in order of enrollment. SPSS statistical software is used to generate 90 random 3-digit numbers, which are then sorted in ascending order. The first 45 numbers are assigned to the experimental group, and the remaining 45 to the control group. To implement allocation concealment, the randomization results are sealed in opaque, sequentially numbered envelopes by an independent research coordinator who is not involved in participant recruitment, clinical assessment, or data analysis. Once a patient meets the eligibility criteria and provides written informed consent, the independent coordinator opens the corresponding envelope according to the patient's number and informs the radiology technician of the group assignment so that the appropriate MRI scanning protocol can be performed. The surgeon then reviews the imaging report generated based on this group assignment before surgery and formulates the surgical plan accordingly.

Blinding

This study adopts a single-blind design in which participants remain unaware of their group assignments. Surgeons are necessarily unblinded due to the preoperative review of imaging reports. To minimize detection bias, all postoperative outcome assessments, including the evaluation of recurrence, pain scoring, anal function measurement, and inflammatory marker testing, are performed by independent assessors who remain blinded to group allocation throughout the trial.

Sample Size Calculation

The sample size was calculated based on data from a preliminary pilot study involving 20 patients with perianal abscesses (10 in each preliminary subgroup). Previous studies have reported an overall postoperative recurrence rate for perianal abscess of approximately 11.7% [17]. In the pilot study, the 6-month postoperative recurrence rate was 15% among patients who underwent conventional MRI–guided surgical treatment, while the recurrence rate was 5% in those who received preoperative DTI assessment combined with targeted surgical intervention. Using a 2-sided α level of .05 and a β level of .20, the sample size for the intervention (DTI group) and control (conventional

XSL•FO
RenderX

MRI group) was estimated using Power Analysis and Sample Size software (version 11), resulting in N1=N2=41. Considering a 10% potential attrition rate during follow-up, the required sample size for each group was adjusted to 45 (41/0.9), leading to a total enrollment of 90 participants for the formal study.

Study Population

Patients with perianal abscesses who have surgical indications and meet the inclusion criteria are being recruited from the Anorectal Department of the Shanghai Hospital of Traditional Chinese Medicine between April 2025 and October 2026.

The diagnostic criteria are formulated with reference to the diagnostic part of perianorectal abscess in Chinese Anorectal Diseases [18] and the diagnostic part of perianorectal abscess issued by the State Administration of Traditional Chinese Medicine in 2012 [19]. The systemic symptoms may include chills, high fever, general weakness, and lack of appetite. Additionally, the patient may experience anal swelling, difficulty, or pain when defecating. The presence of hard nodules or lumps around the anal canal, localized pressure or fluctuating sensation, and increased skin temperature. Palpation of the affected area reveals a slightly hard or fluctuating soft mass and a worsening of the pain when pressure is applied. Congested and swollen anal sinus with purulent secretion upon pressing. A puncture examination is conducted, whereby pus is successfully extracted from the most prominent swelling and congestion of the rectal mucosa observed under an anoscope. This procedure proves invaluable in facilitating an accurate diagnosis. A blood routine examination reveals an increase in the total number of white blood cells and neutrophils.

The inclusion criteria are as follows: (1) informed consent and signed informed consent form; (2) meeting the diagnostic criteria of perianal abscess according to both Western medicine and traditional Chinese medicine syndrome of heat and toxin; (3) aged 15 to 80 years; (4) willing to participate and cooperate in completing the required surgical treatment; (5) agree to participate in this experimental study without accepting any other treatment options. In contrast, the exclusion criteria are as follows: (1) patients younger than 15 years old or older than 80 years old; (2) patients who refuse to participate in the study; (3) patients with severe primary diseases of the liver, kidney, hematopoietic system, endocrine system, cardiovascular and cerebrovascular system, nervous system, tuberculosis, vertebral malformation, malignant tumors, or mental illness; (4) patients with a suspected history of abuse of sedative-hypnotics, opioid analgesics, or alcohol. Patients with skin diseases or infectious diseases; (5) individuals with metal implants who cannot undergo MRI; (6) pregnant women, women preparing to become pregnant, or lactating women; (7) individuals who are not suitable or unable to use this treatment for other reasons; (8) individuals with mental disorders, language disorders, motor dysfunction, or psychological diseases.

Ethical Considerations

The trial is being conducted in accordance with the ethical standards of the Declaration of Helsinki. Ethical approval has been obtained from the Ethics Committee of Shanghai Municipal Hospital of Traditional Chinese Medicine (2023SHL-KYYS-104). Written informed consent is being obtained from all participants after a detailed explanation of the study objectives, procedures, potential risks, and rights (including the right to withdraw at any time without prejudice to clinical treatment), ensuring full comprehension before enrollment. To protect participant privacy and confidentiality, all collected data are being deidentified immediately after acquisition (personal identifiers are replaced with unique study serial numbers) and stored in password-protected electronic databases or locked file cabinets with access restricted to authorized personnel only. No financial compensation is provided to participants for participation in this trial.

Treatment Plan

Both groups undergo an identical perianal MRI protocol on the same 3.0-T scanner using a 16-channel abdominal phased-array coil, with all participants scanned in the supine position. The sole distinction between the 2 imaging protocols lies in the diffusion-weighted sequence: the control group receives conventional DWI, whereas the intervention group undergoes DTI. Crucially, both groups receive intravenous gadolinium-based contrast agent (GD-DTPA) and subsequent identical postcontrast sequences. All remaining sequences, including T1-weighted, T2-weighted, and fat-suppressed T2-weighted imaging, are standardized across groups.

Conventional MRI Group

The conventional MRI group underwent imaging using a 3.0-T Philips Ingenia magnetic resonance system equipped with a 16-channel abdominal phased-array coil, with patients scanned in the supine position. The scanning protocol includes the following sequences and parameters. Sagittal fast spin-echo T2-weighted imaging (FSE T2WI) is acquired with an echo time (TE) of 85 ms, repetition time (TR) of 4000 ms, slice thickness of 3.5 mm, slice spacing of 0.5 mm, and matrix size of 416 × 224. Oblique axial FSE T2WI and oblique axial fat-suppressed T2WI are obtained using identical parameters: TE 85 ms, TR 4000 ms, slice thickness 3.5 mm, slice spacing 0.5 mm, and matrix 416 × 224. DWI is performed in the oblique axial plane with a TE of 71 ms, TR of 2000 ms, slice thickness of 5 mm, slice spacing of 1 mm, and a b-value of 800 s/mm^2. For T1-weighted imaging (T1WI), oblique axial FSE T1WI is acquired with minimum TE, a TR of 400 ms, slice thickness of 3.5 mm, slice spacing of 0.5 mm, and matrix size of 416 × 224. Following intravenous administration of gadolinium-based contrast agent (GD-DTPA), spoiled gradient–recalled echo sequences are performed in oblique axial, oblique coronal, and sagittal orientations, with minimum TE, TR of 200 ms, flip angle of 70°, slice thickness of 3.5 mm, slice spacing of 0.5 mm, and matrix dimensions of 320 × 256.

The DTI Group

The DTI group is scanned using a 3.0-T Philips Ingenia magnetic resonance imager with a 16-channel abdominal phased array coil, with all participants positioned supine during imaging. The scanning protocol comprises several sequences conducted with specific parameters. Turbo spin-echo T1WI is performed with a TR of 600 ms, TE of 10 ms, a slice thickness of 5 mm, and a total of 20 slices. For transverse axial imaging,

XSL•FO
RenderX

both turbo spin-echo T2-weighted and fat-suppressed T2-weighted sequences are acquired using a TR of 1560 ms, TE of 80 ms, slice thickness of 5 mm, and 20 slices. Sagittal fat-suppressed T2WI is carried out with a field of view of 152 × 179 mm, slice thickness of 5 mm, and 20 slices. DTI is conducted using a spin-echo echo-planar imaging sequence with a single excitation in the transverse axis, employing a TR of 3250 ms, TE of 48 ms, 32 diffusion directions, b values of 0 and 600 s/mm², slice thickness of 5 mm, 20 slices, and a total scan time of 5 minutes and 47 seconds.

Perioperative Management

All surgical procedures in this trial are being performed by surgeons who hold senior professional titles and possess extensive surgical experience. Prior to surgery, preoperative communication is conducted between the medical team and each patient to ensure a clear understanding of the procedural steps and to facilitate cooperation during treatment. Preoperative preparations include an 8-hour fasting and 6-hour fluid restriction period before the procedure. Patients are instructed to take laxatives on the evening prior to surgery and to undergo a cleansing enema on the morning of the operation.

On the day of surgery, perianal hair is shaved and the local area is thoroughly cleansed. Patient positioning—whether lithotomy, prone jackknife, or lateral decubitus—is determined based on the specific location and characteristics of the abscess. Intravenous anesthesia is administered routinely, with general anesthesia reserved for special clinical circumstances.

The surgical approach is decided following a comprehensive evaluation that incorporates the patient's clinical symptoms, findings from specialized physical examination, and results from auxiliary imaging. The selected intervention is a one-time radical resection of the perianal abscess, aiming to completely address the primary lesion and achieve a curative outcome.

Postoperative care begins on the first day after surgery, with wound dressings being changed twice daily at 8:00 AM and 6:00 PM. During each dressing change, iodophor-soaked cotton balls are used to gently cleanse the pus cavity, wound surface, and any necrotic tissue surrounding the drainage line or tube. Additional wound management interventions, such as cavity irrigation, negative pressure suction, or pad compression, are applied as clinically indicated. In cases where postoperative blood tests suggest infection, intravenous administration of cefuroxime sodium combined with metronidazole is initiated as anti-infective therapy.

Outcomes

The following data are being systematically collected and analyzed throughout the trial.

Demographic and Baseline Data

Data on age, gender, race, height, weight, medical history, and medication history are being recorded for all participants.

General Physical Examination

Vital signs are being monitored and documented as part of the initial and follow-up assessments.

Safety Monitoring

Safety is being evaluated through a series of routine tests, including complete blood count, urinalysis, stool analysis, liver and kidney function panels, and electrocardiograms.

Primary Efficacy Outcome

This study assesses the recurrence rate of perianal abscess as the primary endpoint, evaluated at 6 months (±7 d) after surgery. Recurrence is defined by the reappearance of clinical symptoms (eg, perianal pain, swelling, or purulent discharge) along with confirmation on a follow-up MRI scan—specifically the DTI sequence—of a new or persistent fluid collection suggestive of an abscess. For the purpose of this primary outcome, the presence of a newly developed anal fistula also constitutes a recurrence event. All assessments for recurrence are performed by 2 independent, experienced radiologists. These radiologists are blinded to the patient's group allocation and baseline imaging findings. Any discrepancy between their assessments is resolved either by consensus or by a third senior radiologist.

Secondary Efficacy Outcomes

Structured Imaging Report

The structured report (see Table S1 in Multimedia Appendix 1) is completed following the surgical procedure and serves as a synthesis of direct intraoperative findings and preoperative DTI results. During surgery, an experienced colorectal surgeon meticulously explores the lesion and documents the key anatomical conditions. These surgical observations are then integrated with the preoperative DTI interpretations provided by a senior radiologist to finalize the report collaboratively. It is important to clarify that this report is compiled postintervention and therefore does not inform real-time intraoperative decisions. To ensure consistency, all participating surgeons and radiologists received standardized training on the reporting items and definitions prior to the study. To assess inter-rater reliability, a subset of the completed reports will undergo independent review by a second surgeon who remains blinded to the initial assessment.

Postoperative Pain Score

Pain scores: The visual analog scale is being used to measure postoperative pain on days 3 and 7 [20]. Participants are asked to draw a horizontal line on a piece of paper, with 0 at one end representing no pain and 10 at the other end representing severe pain. Participants then mark a point on the line corresponding to the intensity of their pain sensation. The distance from point 0 to the marked point is measured to determine the pain intensity score.

Detection of Inflammatory Factors

Peripheral blood samples are collected to measure the levels of interleukin (IL)-1β, IL-6, tumor necrosis factor-α, IL-2, IL-4, IL-5, IL-8, IL-10, IL-12p70, IL-17A, interferon-γ, and interferon-α2. All sample processing and testing are conducted in strict accordance with the standardized operating procedures of the hospital's clinical laboratory. Antibody chips are being employed to covalently bind highly specific inflammatory factor-capture antibodies onto chemically modified solid-phase chip substrates. Signal amplification is achieved through the

XSL•FO
RenderX

use of high-affinity paired secondary antibodies, enabling the effective detection of trace cytokines. Furthermore, a standard curve is constructed using the detection signals of gradient dilutions of standards, allowing for accurate quantitative detection of inflammatory factors in human samples by standard curve fitting. All laboratory procedures are carried out by trained staff who are blinded to group assignments. Intra-assay and inter-assay coefficients of variation are consistently controlled below 10%, ensuring the reliability of the analytical results.

Assessment of Anal Function

The Wexner Anal Incontinence Score [21] is being used to assess the patient's anal function before and 6 months after surgery. This score evaluates anal function and records incontinence in terms of dry stools, loose stools, gas, need for pads, and lifestyle changes (See Table S2 in Multimedia Appendix 1).

Assessments are conducted by trained research assistants who are blinded to group allocation and baseline scores. All assessors received standardized training prior to study initiation to ensure consistent administration and interpretation of the Wexner scoring system. At each assessment time point, the Wexner questionnaire is administered via a structured interview format. To further mitigate potential assessment bias, baseline scores are recorded prior to randomization and will be incorporated as covariates in the statistical models used to compare postoperative functional outcomes between the 2 groups.

Follow-Up Plan

To ensure complete data collection, all participants will be contacted according to a standardized follow-up schedule. Automated text message reminders will be sent 2 days before each scheduled visit (postoperative days 3 and 7, and month 6), followed by a confirmatory telephone call 1 day prior. For the 6-month assessment, participants who are unable to attend an in-person will be offered a structured telephone interview to evaluate symptoms and functional status; if recurrence is suspected during the interview, an urgent in-person clinical evaluation, including DTI reassessment, will be arranged.

Adverse Events Monitoring and Reporting

Adverse events (AEs) are defined as any unfavorable medical occurrence experienced by trial participants, irrespective of causal relationship to the study intervention. This includes newly emerging symptoms, clinically relevant abnormal laboratory findings, or exacerbations of preexisting medical conditions. A serious adverse event (SAE) is defined as any AE that fulfills at least one of the following criteria: results in death, is life-threatening, necessitates inpatient hospitalization or prolongation of existing hospitalization, results in persistent or significant disability/incapacity, or constitutes a congenital anomaly or birth defect.

All AEs will be graded according to the Common Terminology Criteria for Adverse Events, version 5.0, as follows: grade 1 (mild: asymptomatic or mild symptoms, no intervention indicated); grade 2 (moderate: minimal intervention indicated, interferes with activities of daily living); grade 3 (severe: medically significant but not immediately life-threatening, hospitalization indicated); grade 4 (life-threatening: urgent intervention required); and grade 5 (death). All AEs are systematically recorded at each study visit and during scheduled follow-up interactions. Nonserious AEs are documented in the case report forms and will be summarized in the final study report. SAEs must be reported to the principal investigator and the Institutional Ethics Committee within 24 hours of awareness, with a comprehensive written report submitted within 7 calendar days, followed by continuous documentation until event resolution or stabilization.

An independent Data and Safety Monitoring Board has been established to conduct periodic reviews of SAEs, interim safety data, and overall trial safety. The Ethics Committee maintains the authority to request supplementary information, recommend protocol modifications, or suspend the study based on safety evaluations. All investigators and study personnel have completed standardized training in AE/SAE identification, documentation, and reporting procedures prior to study initiation.

Data Management and Quality Assurance

All study data will be managed through an electronic data capture system equipped with built-in logic checks and mandatory field validation to ensure data integrity, accuracy, and consistency. After entry, data will undergo dual independent verification, with systematic reviews conducted by data managers to identify outliers and address missing values. Data and Safety Monitoring Board will conduct periodic assessments of safety data, protocol compliance, and overall data quality. All study personnel have completed standardized training to maintain consistency and procedural rigor throughout data collection and evaluation processes.

Statistical Plan

All analyses are prespecified. The primary analysis will follow the intention-to-treat principle. For the primary outcome, the 6-month recurrence rate will be analyzed using binary logistic regression, with results reported as an adjusted odds ratio and 95% CI. Any clinically relevant baseline imbalances will be included as covariates, and missing recurrence data will be handled using multiple imputation. Secondary outcomes will be analyzed as follows: categorical items from the structured report will be compared using chi-square test or Fisher exact test; continuous structured report items will be compared using independent samples *t* tests or Mann-Whitney *U* tests, depending on normality; postoperative visual analog scale pain scores will be evaluated with a linear mixed model to assess group, time, and interaction effects, with results expressed as estimated marginal means and mean differences with 95% CI; and inflammatory markers and Wexner incontinence scores will be compared at follow-up using *t* tests or Mann-Whitney *U* tests. To control for type I error across secondary outcomes, the false discovery rate procedure will be applied. Prespecified sensitivity analyses using the per-protocol population and subgroup analyses stratified by abscess complexity will also be performed. All statistical analyses will be conducted using SPSS software (version 22.0), and statistical significance will be defined as a 2-sided P value < .05 .

XSL•FO
RenderX

Management of Protocol Amendments

Any proposed amendment to the study protocol must be submitted for review and approval by the Ethics Committee prior to implementation. Approved amendments will be promptly reflected in the public clinical trial registry, and all investigators and relevant stakeholders will be formally informed. No protocol amendment shall adversely affect the rights, safety, or well-being of already enrolled participants. Any modification deemed to be of potential benefit to participants may only be implemented with Ethics Committee approval and, where applicable, following informed consent from the affected participants.

Management of Participant Withdrawal

Participants retain the right to withdraw from the study at any stage without condition, and such withdrawal will not affect their future medical care. If a participant withdraws, data collected prior to withdrawal may be retained for use in the intention-to-treat analysis, provided that the participant has given consent for such use. If consent is not provided, the participant's data will be deleted and excluded from the final analysis. The reason for withdrawal will be documented and reported. In cases where withdrawal is due to an AE, appropriate follow-up will be conducted until the event is resolved or stabilized.

Results

The trial was registered on January 10, 2025, with the International Traditional Medicine Clinical Trial Registry (ITMCTR2025000068). The study obtained ethical approval in 2023. Participant recruitment began in April 2025 and is expected to continue until October 2026. Completion of the 6-month postoperative end point follow-up is projected by April 2027. Preliminary data analysis is planned for the second quarter of 2027, and the full study report anticipated to be submitted and published before the end of 2027. As of January 2026, a total of 37 participants had been enrolled. Results will be submitted for publication in a peer reviewed international journal. Findings will also be presented at relevant national and international conferences.

Discussion

This study is designed to test the hypothesis that preoperative DTI, compared with conventional MRI, can provide superior visualization of perianal abscess extent and muscular involvement, thereby guiding more precise surgical excision and leading to a lower 6-month recurrence rate.

Currently, there is only 1 comparative study between DTI and MRI for perianal fistula [16], and no specific studies have been reported on the location depth, size, thickness of anal canal and rectum, and thickness of the levator ani muscle and internal and external sphincter of perianal abscesses. This protocol builds upon the feasibility shown in fistula imaging but shifts the focus to preoperative mapping of abscess morphology and infiltration—a critical gap in current clinical practice.

Perianal abscess is a common disease in the anorectal department, and surgery remains the most effective and complete cure treatment [22]. However, the relationship between the extent of abscess involvement and surrounding inflammatory infiltration is not determined before the operation, which can increase the risk of postoperative recurrence [23]. Although conventional MRI sequences have some diagnostic value in determining the extent of disease, they have limitations in distinguishing abscesses from inflammatory masses [24,25]. Therefore, a contrast agent that shows the characteristic edge enhancement of an abscess is often used in combination. However, there are increasing concerns regarding contrast injection, especially in patients with renal insufficiency, and the association between contrast exposure and nephrogenic systemic fibrosis has been emphasized [26]. Therefore, the search for alternative imaging techniques that do not use contrast agents has been an important topic, especially for high-risk patients, while ensuring that diagnostic value is maintained. DTI can clearly show the degree of infection infiltration, the extent of the abscess, and the relationship between the erosion of tissue and muscle around the pelvic floor [27,28]. This makes DTI essential for reducing the chance of recurrence. In addition, the fractional anisotropy and apparent diffusion coefficient values of the DTI parameters may play a role in assessing the extent and activity of the perianal abscess.

In this trial, 90 participants are being enrolled and divided into 2 groups, a DTI group and a conventional MR group, based on randomization performed before enrollment. "The primary variable, recurrence rate, and secondary variables, including structured reports, pain scores on postoperative days 3 and 7, peripheral inflammatory factors, and assessment of anal function, are being assessed at 6 months after surgery." To ensure accurate and specific information, a structured report on the extent of perianal abscess infection based on DTI images is being prepared by both anorectal physicians and radiologists. Surgeons experienced in examining and operating on the lesions can provide valuable insights into the actual situation of the patient during surgery. By using a structured report, we aim to determine whether the test under consideration is effective and which of the 2 tests is more effective.

Although the recurrence rate of perianal abscesses after surgery is influenced by numerous factors, there is no doubt that the accuracy of the examination can greatly improve the success rate of surgery [29]. Currently, surgical treatment is the main method for perianorectal abscesses in clinical medicine. However, the location, number, specific shape of the abscess cavity, and its relationship with surrounding muscles of perianorectal abscess have a direct impact on the choice of surgical methods, which in turn will influence its prognosis [30]. Therefore, it is very important to pay attention to the preoperative diagnosis of such patients for the implementation of targeted and effective surgical treatment. Therefore, by using DTI MRI imaging sequences, we can scan the internal opening, the degree of infection and infiltration, the extent of abscess, and the relationship between the erosion of perianal tissues and muscles, to explore a new imaging method for the evaluation and diagnosis of perianal abscess. This approach can explore the application value of DTI in the diagnosis of perianal abscess.

XSL•FO
RenderX

It can improve the accuracy of preoperative evaluation, reduce damage to the anal sphincter, reduce the recurrence rate, and improve the quality of life of patients.

DTI brings additional challenges to the diagnosis of perianal abscess. First, there is no uniform standard for measuring the thickness and ratio of the relevant muscles before surgery for perianal abscess, as well as the location, depth, and size of the abscess. To address this issue, relevant researchers will be trained by the director of radiology, relevant literature will also be reviewed to formulate unified standards. Second, monitoring MRI and operating room schedules is needed to identify potential study patients and appoint someone to manage the imaging data. Despite the challenges of such data collection efforts, many of the challenges of integrating research into clinical workflows can be overcome. DTI is a useful clinical tool for surgical planning and intraoperative guidance. By using DTI technology, we can gain a deeper understanding of the relationship between the internal orifice, the degree of infection infiltration, the extent of abscess, and the erosion of perianal tissues and muscles. These data may provide a basis for surgical treatment, and we hope that it will form a new technical standard for the diagnosis of perianal abscesses in the future.

It is hoped that the dataset being collected in this study will represent the objective parameters of the perianal muscular tissues and also effectively respond to the patient population we are eventually serve. As expected, due to regional ethnic differences, more data will need to be collected in the future, but it is expected that based on this study, the draft technical standards of DTI for perianal abscess diseases will be established, and new techniques for diagnosis and evaluation of perianal abscess will be developed and formed. This will serve the surrounding hospitals and reduce the postoperative recurrence rate and reoperation rate of perianal abscess patients.

Acknowledgments

We would like to express our heartfelt gratitude to the participants who will take part in the study.

Funding

The authors declare that financial support was received for the research and/or publication of this article. This study was financially supported by Science and Technology Commission of Shanghai Municipality Science and Technology Program (23Y11920800), Training Program for High-caliber Talents of Clinical Research at Affiliated Hospitals of SHUTCM (2025LCRC12), the National Natural Science Foundation of China (82405390; 82305230), Traditional Chinese Medicine Inheritance and Innovation Studio (2025CXGZS-18) and the Key Specialty of Shanghai Municipal Hospital of Traditional Chinese Medicine.

Data Availability

Data sharing is not applicable to this article as no datasets have been generated or analyzed yet.

Authors' Contributions

Conceptualization: YC, XG
Funding acquisition: YC, XG
Investigation: YM, PM
Project administration: YC, XG
Supervision: YC, XG
Writing – original draft: YM, PM
Writing – review & editing: YM, PM, YC, XG

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables of the DTI structural report of perianal abscess and the Wexner anal incontinence score.
[DOCX File, 15 KB - resprot_v15i1e83449_app1.docx]

References

1. Sahnan K, Adegbola SO, Tozer PJ, Watfah J, Phillips RK. Perianal abscess. BMJ 2017 Feb 21;356:j475. [doi: 10.1136/bmj.j475] [Medline: 28223268]
2. Wang D, Yang G, Qiu J, et al. Risk factors for anal fistula: a case-control study. Tech Coloproctol 2014 Jul;18(7):635-639. [doi: 10.1007/s10151-013-1111-y] [Medline: 24452294]
3. Gaertner WB, Burgess PL, Davids JS, et al. The American Society of Colon and Rectal Surgeons Clinical Practice Guidelines for the management of anorectal abscess, fistula-in-ano, and rectovaginal fistula. Dis Colon Rectum 2022 Aug 1;65(8):964-985. [doi: 10.1097/DCR.0000000000002473] [Medline: 35732009]

XSL•FO
RenderX

4. Vogel JD, Johnson EK, Morris AM, et al. Clinical practice guideline for the management of anorectal abscess, fistula-in-ano, and rectovaginal fistula. Dis Colon Rectum 2016 Dec;59(12):1117-1133. [doi: 10.1097/DCR.0000000000000733] [Medline: 27824697]
5. Amato A, Bottini C, De Nardi P, et al. Evaluation and management of perianal abscess and anal fistula: SICCR position statement. Tech Coloproctol 2020 Feb;24(2):127-143. [doi: 10.1007/s10151-019-02144-1] [Medline: 31974827]
6. Villa C, Pompili G, Franceschelli G, et al. Role of magnetic resonance imaging in evaluation of the activity of perianal Crohn's disease. Eur J Radiol 2012 Apr;81(4):616-622. [doi: 10.1016/j.ejrad.2011.01.046] [Medline: 21316171]
7. Beets-Tan RG, Beets GL, van der Hoop AG, et al. Preoperative MR imaging of anal fistulas: does it really help the surgeon? Radiology 2001 Jan;218(1):75-84. [doi: 10.1148/radiology.218.1.r01dc0575] [Medline: 11152782]
8. Dong S, Chen B, Zhang J. Study on the factors influencing the prognosis after perianal abscess surgery. BMC Gastroenterol 2023 Sep 27;23(1):334. [doi: 10.1186/s12876-023-02959-1] [Medline: 37759161]
9. Gullo RL, Partridge SC, Shin HJ, Thakur SB, Pinker K. Update on DWI for breast cancer diagnosis and treatment monitoring. AJR Am J Roentgenol 2024 Jan;222(1):e2329933. [doi: 10.2214/AJR.23.29933] [Medline: 37850579]
10. Alyami AS. Imaging of ulcerative colitis: the role of diffusion-weighted magnetic resonance imaging. J Clin Med 2024 Sep 2;13(17):5204. [doi: 10.3390/jcm13175204] [Medline: 39274415]
11. Hori M, Oto A, Orrin S, Suzuki K, Baron RL. Diffusion‐weighted MRI: a new tool for the diagnosis of fistula in ano. J Magn Reson Imaging 2009 Nov;30(5):1021-1026. [doi: 10.1002/jmri.21934] [Medline: 19856434]
12. Dohan A, Eveno C, Oprea R, et al. Diffusion-weighted MR imaging for the diagnosis of abscess complicating fistula-in-ano: preliminary experience. Eur Radiol 2014 Nov;24(11):2906-2915. [doi: 10.1007/s00330-014-3302-y] [Medline: 25038854]
13. Bakan S, Olgun DC, Kandemirli SG, et al. Perianal fistula with and without abscess: assessment of fistula activity using diffusion-weighted magnetic resonance imaging. Iran J Radiol 2015 Oct;12(4):e29084. [doi: 10.5812/iranjradiol.29084] [Medline: 26715982]
14. Tae WS, Ham BJ, Pyun SB, Kang SH, Kim BJ. Current clinical applications of diffusion-tensor imaging in neurological disorders. J Clin Neurol 2018 Apr;14(2):129-140. [doi: 10.3988/jcn.2018.14.2.129] [Medline: 29504292]
15. Zijta FM, Froeling M, Nederveen AJ, Stoker J. Diffusion tensor imaging and fiber tractography for the visualization of the female pelvic floor. Clin Anat 2013 Jan;26(1):110-114. [doi: 10.1002/ca.22184] [Medline: 23168612]
16. Wang Y, Gu C, Huo Y, et al. Diffusion tensor imaging for evaluating perianal fistula: feasibility study. Medicine (Baltimore) 2018;97(29):e11570. [doi: 10.1097/MD.0000000000011570]
17. JS HU, WANG W, LM SUN. Analysis of influencing factors of anal fistula formation after perianal abscess operation with different operation methods. Chin J Surg Integr Trad West Med 2023;29(1):24-28 [FREE Full text]
18. Huang NJ. Chinese Journal of Colorectal Diseases [Book in Chinese]: Shandong Science and Technology Press; 1996. URL: https://book.m.supfree.net/book.asp?id=120911 [accessed 2026-02-18]
19. State Administration of Traditional Chinese Medicine (SACM). Traditional Chinese Medicine Disease and Syndrome Diagnosis and Treatment Standards [Book in Chinese]: Nanjing University Press; 1994. URL: https://books.google.co.kr/books/about/?id=0EIftwAACAAJ [accessed 2026-02-18]
20. Kelly AM. The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain. Emerg Med J 2001 May;18(3):205-207. [doi: 10.1136/emj.18.3.205] [Medline: 11354213]
21. Agachan F, Chen T, Pfeifer J, Reissman P, Wexner SD. A constipation scoring system to simplify evaluation and management of constipated patients. Dis Colon Rectum 1996 Jun;39(6):681-685. [doi: 10.1007/BF02056950] [Medline: 8646957]
22. Qiheng L, Maonan W, Siqi L, Tao M. Progress of Chinese and Western medicine treatment of perianal abscess [Article in Chinese]. Jilin Med J 2024;45(5):1201-1203 [FREE Full text]
23. Amato A, Bottini C, De Nardi P, et al. Evaluation and management of perianal abscess and anal fistula: a consensus statement developed by the Italian Society of Colorectal Surgery (SICCR). Tech Coloproctol 2015 Oct;19(10):595-606. [doi: 10.1007/s10151-015-1365-7] [Medline: 26377581]
24. Hopkins KL, Li KC, Bergman G. Gadolinium-DTPA-enhanced magnetic resonance imaging of musculoskeletal infectious processes. Skeletal Radiol 1995 Jul;24(5):325-330. [doi: 10.1007/BF00197059] [Medline: 7570151]
25. Kan JH, Young RS, Yu C, Hernanz-Schulman M. Clinical impact of gadolinium in the MRI diagnosis of musculoskeletal infection in children. Pediatr Radiol 2010 Jul;40(7):1197-1205. [doi: 10.1007/s00247-010-1557-2] [Medline: 20180105]
26. Daftari Besheli L, Aran S, Shaqdan K, Kay J, Abujudeh H. Current status of nephrogenic systemic fibrosis. Clin Radiol 2014 Jul;69(7):661-668. [doi: 10.1016/j.crad.2014.01.003] [Medline: 24582176]
27. Zijta FM, Froeling M, van der Paardt MP, et al. Feasibility of diffusion tensor imaging (DTI) with fibre tractography of the normal female pelvic floor. Eur Radiol 2011 Jun;21(6):1243-1249. [doi: 10.1007/s00330-010-2044-8] [Medline: 21197534]
28. Zijta FM, Lakeman MME, Froeling M, et al. Evaluation of the female pelvic floor in pelvic organ prolapse using 3.0-Tesla diffusion tensor imaging and fibre tractography. Eur Radiol 2012 Dec;22(12):2806-2813. [doi: 10.1007/s00330-012-2548-5] [Medline: 22797954]
29. Chun CW, Jung JY, Baik JS, Jee WH, Kim SK, Shin SH. Detection of soft-tissue abscess: comparison of diffusion-weighted imaging to contrast-enhanced MRI. J Magn Reson Imaging 2018 Jan;47(1):60-68. [doi: 10.1002/jmri.25743] [Medline: 28432835]

XSL•FO
RenderX

30. van Koperen PJ, Wind J, Bemelman WA, Bakx R, Reitsma JB, Slors JFM. Long-term functional outcome and risk factors for recurrence after surgical treatment for low and high perianal fistulas of cryptoglandular origin. Dis Colon Rectum 2008 Oct;51(10):1475-1481. [doi: 10.1007/s10350-008-9354-9] [Medline: 18626715]

Abbreviations

AE: adverse event
DTI: diffusion tensor imaging
DWI: diffusion-weighted imaging
FSE: fast spin-echo
GD-DTPA: Gadolinium-Diethylenetriamine Pentaacetate
IL: interleukin
MRI: magnetic resonance imaging
SAE: serious adverse event
SPIRIT: Standard Protocol Items Recommendations for Interventional Trials
T1WI: T1-weighted imaging
T2WI: T2-weighted imaging
TE: echo time
TR: repetition time

Edited by A Schwartz; submitted 03.Sep.2025; peer-reviewed by N Nazari; revised version received 19.Jan.2026; accepted 29.Jan.2026; published 25.Feb.2026.

Please cite as:
Ma Y, Mei P, Guo X, Chen Y
Diffusion Tensor Imaging in the Diagnosis of Perianal Abscess: Protocol for a Single-Blind Randomized Controlled Trial
JMIR Res Protoc 2026;15:e83449
URL: https://www.researchprotocols.org/2026/1/e83449
doi:10.2196/83449

© Yuting Ma, Pingping Mei, Xiutian Guo, Yan Chen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Urban Care Farming to Enhance Quality of Life Among Older Adults: Protocol for a Waitlist Randomized Trial

Cynthia Chen[1,2], PhD; Jocelin Lam[1], MCSt; Zhenye Shi[1], MSc; Katika Akksilp[3], PhD, MD; Su Aw[1], PhD; Mary Foong-Fong Chong[1], PhD; Choon Nam Ong[1], PhD; Angelia Sia[4], PhD; Leng Leng Thang[5], PhD; Xin Kai Tham[6], BLArch; JunXiang Pong[7], MBA; Kartini Omar[4], BSc; Joan Hung[4], BCom; Marek Kukumberg[8], PhD; Roger Ho[9], MD; Elizabeth Diehl[10], RLA, HTM

[1]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, 12 Science Drive 2, 09-01T, Singapore, Singapore

[10]Department of Environmental Horticulture, University of Florida, Gainesville, FL, United States

[2]Schaeffer Center for Health Policy and Economics, University of Southern California, Los Angeles, CA, United States

[3]Health Intervention and Technology Assessment Program, Bangkok, Thailand

[4]Centre for Urban Greenery and Ecology Research, National Parks Board, Singapore, Singapore, Singapore

[5]Faculty of Arts & Social Sciences, National University of Singapore, Singapore, Singapore

[6]Terrapy SG, Singapore, Singapore

[7]Hort Therapeutics, Singapore, Singapore

[8]Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[9]Division of Life Science (LIFS), Hong Kong University of Science and Technology (HKUST), Clear Water Bay, Hong Kong, China (Hong Kong)

Corresponding Author:
Cynthia Chen, PhD
Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, 12 Science Drive 2, 09-01T, Singapore, Singapore

Abstract

Background: Population aging poses challenges to health systems and costs, and evidence shows that older adults spend a long time in ill health. Improving healthspan, time spent in good health, allows older adults to contribute and improve in their quality of life. Active and healthy aging are crucial to improving healthspan. Urban care farming (UCF) is a behavioral intervention that is purported to enhance active and healthy aging.

Objective: This trial evaluates the effectiveness of a care farming intervention in improving the quality of life and biopsychosocial health outcomes of older participants.

Methods: We conducted a parallel group, 2-arm pragmatic waitlist randomized trial with a 1:1 allocation, in which participants were randomized into either the intervention or waitlist control arm. Community-dwelling participants aged 50-85 years, without any mobility issues, were recruited. Participants in the intervention arm commenced the 24-week UCF program, while waitlist control participants received no intervention during this period. The primary (World Health Organization Quality of Life−brief version) and secondary outcomes were collected at baseline, 6th month, and 12th month after the intervention group completed the trial. Secondary outcomes include objectively measured physiological outcomes, cognition, frailty, and self-reported psychosocial outcomes. Intervention effects were estimated using mixed-effects difference-in-differences regression to account for repeated measurements.

Results: The randomized controlled trial commenced in April 2024, with the intervention group starting first. By April 2024, we had enrolled 137 participants at commencement, with 67 participants randomized to the intervention group and 70 to the control group. The intervention arm started in April 2024 and concluded in September 2024. Baseline data were collected in March 2024, and 6-month follow-up data were collected in September 2024. The waitlist control participants began the UCF intervention at the end of September 2024 and concluded in April 2025. Data collection for the 12-month follow-up concluded in May 2025. Analysis of the baseline and 6-month follow-up data is still ongoing.

Conclusions: The outcomes of this study will contribute to the understanding of UCF on quality of life and health. This trial has potential positive implications for public health, as it utilizes a robust research design and methods to provide empirical insights into the multifaceted health benefits of the multicomponent UCF intervention. This trial could also serve as a model for future intervention research on scalable, community-based programs. Taken together, the UCF content and the outcomes, process, and economic evaluations completed through this study could inform scalable models of the UCF intervention, with potential implications for public health strategies to address health issues related to population aging.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT06277583; https://clinicaltrials.gov/study/NCT06277583
International Registered Report Identifier (IRRID): DERR1-10.2196/78584

(JMIR Res Protoc 2026;15:e78584) doi:10.2196/78584

KEYWORDS

biopsychosocial health; complex intervention; healthy and active aging; quality of life; urban care farming

Introduction

Globally, population aging has raised concerns about the growing burden on health care systems and rising health care costs. While life expectancy has increased, individuals are living longer in poorer health [1]. Aging brings about physiological changes, such as the loss of muscle mass, which leads to functional decline and reduced brain volume [2,3], contributing to cognitive impairment [4]. Emotional and psychosocial well-being are also impacted, with older adults experiencing higher rates of depression [5] and social isolation [6]. In Singapore, 1 in 5 older adults aged 75 years and above shows signs of depression [7], and the proportion who committed suicide rose from 23% in 2000 to 30% in 2014 [8]. Furthermore, many mental health conditions remain underdiagnosed and untreated, largely due to stigma and delayed health-seeking behavior [9]. Active and healthy aging are crucial to improving healthspan. Care farming is a behavioral intervention that is purported to enhance active and healthy aging.

Care farming is a nature-based therapeutic intervention that integrates agricultural and farming practices to positively influence participants in multiple health domains [10-12]. Nature-based intervention (NBI) has garnered traction as a green social prescribing activity that addresses upstream social determinants of health [13]. The nature-based biopsychosocial resilience theory (NBRT) provides a conceptual foundation for understanding how exposure to nature can promote health and well-being [14]. The NBRT posits that humans have an inherent affinity for the natural environment, through which they derive restorative benefits that support biopsychosocial resilience and overall health. According to White and colleagues, biopsychosocial health restoration occurs through increased exposure to nature, which reduces stress (eg, by mitigating pollution through the use of trees) while simultaneously strengthening resources such as physical fitness, social connections, and self-esteem. The NBRT also aligns with the World Health Organization's quality of life (QOL) definition, which emphasizes health and illness as outcomes of complex interactions among biological, psychological, and social factors.

The World Health Organization defines QOL as an individual's perception of their position in life, considered within the context of their culture, values, goals, expectations, standards, and concerns [15]. To capture this multidimensional construct of health, the WHOQOL (World Health Organization Quality of Life) instruments assess multiple domains, including physical, psychological, social, and environmental aspects of well-being. From this perspective, a comprehensive understanding of a person's health or QOL requires the consideration of biological processes and lifestyle; cognitive and emotional functioning; coping capacities; and the social context of relationships, support systems, and culture. Given the multidimensional nature of WHOQOL, its domains align closely with the components of the NBRT, offering a coherent framework for evaluating health and well-being. In line with this, active aging aims to maintain functional ability, promote participation, and enhance the QOL as people age [16,17].

Growing evidence suggests that adopting healthy behaviors can lower disease risk, enhance physical and mental health function, and delay the onset of care dependency [18]. Care farming is valued for its multidimensional impact on biopsychosocial health and QOL. It is an intervention with the goal of promoting holistic health through activities by increasing light-to-moderate physical activity; reducing stress, anxiety, and negative mood; enhancing positive mood; fostering social interaction and sense of purpose; stimulating cognition; and encouraging healthier eating habits [19-24]. Unlike many NBIs, which can be undertaken individually, care farming places a stronger emphasis on fostering community engagement and collaboration. Hence, we hypothesize that care farming better addresses psychological and social relationship domains of QOL by fostering a sense of purpose and social connectedness through collective cultivation, responsibility, and group harvesting and communal cooking. Care farming is therefore a highly complex intervention consisting of interrelated components that target multiple domains of well-being. This is consistent with the UK Medical Research Council's framework for complex interventions, which considers the interaction of multiple components, variability in outcomes, and context-sensitive implementation [25].

Although care farming is garnering attention as a green care intervention, the empirical evidence remains limited. Systematic reviews by Murray et al [12] and Coventry et al [26] highlight limitations in current literature, concluding that there is mixed evidence showing that care farming or NBIs improved individuals' health, well-being, and QOL. Murray et al also found that only 2 care farming studies employed randomized controlled methods, with the majority of reviewed studies being pre-post and qualitative studies conducted with small sample sizes. Moreover, care farming tasks are typically tailored to the needs and capabilities of participants, shaped by the type of farm and the resources available. Such flexibility poses challenges for research, as the heterogeneity of care farming interventions complicates the identification of mechanisms of change and limits both replicability and generalizability. Both reviews also noted that care farming is grouped with other NBIs, making it difficult to attribute observed outcomes specifically to farming-related activities. These limitations underscore a clear literature gap and the need for a robust, well-designed experimental study with structured tasks to establish the effectiveness of care farming as a health intervention.

XSL•FO
RenderX

There is a strong theoretical and empirical rationale to assess its potential impacts across physical, psychological, cognitive, and behavioral domains. This study aims to (1) develop a structured care farming intervention, which we termed urban care farming (UCF) and (2) evaluate the acceptability and effectiveness of a structured UCF program for older adults in Singapore using a waitlist randomized controlled trial. Specifically, this outcome evaluation seeks to answer the following research questions:

1. Effectiveness on primary outcome: Does participation in UCF improve overall QOL, as measured by the WHOQOL-BREF (World Health Organization Quality of Life−brief version), relative to a waitlist control?
2. Effects on secondary and exploratory outcomes: What are the effects of UCF on physical activity, psychosocial well-being, dietary behaviors, cognitive function, and biomarkers?

These research questions were guided by the intervention's logic model (Figure S1 in Multimedia Appendix 1), which outlines hypothesized pathways from program activities to short- and medium-term outcomes. This intervention targets multiple dimensions of health, including physical, psychological, social, and cognitive outcomes, aligning with both the WHOQOL framework and biopsychosocial model [27,28]. In addition, subsequent, separate evaluations of UCF will also incorporate process and cost-effectiveness assessments to further examine implementation effectiveness and identify challenges. These evaluations include the following:

1. Process evaluation will be conducted to assess the quality of urban care farming (UCF) intervention delivery, including fidelity, dose delivered and received, acceptability, and to identify the mechanisms and contextual factors that facilitate or hinder outcomes.
2. Cost-utility analysis will be conducted from the provider and societal perspective to assess the potential value of scaling up the UCF intervention, using quality-adjusted life years (QALYs) as the primary effectiveness metric. This economic evaluation will be exploratory in nature and designed to inform future trial design and policy translation.

Methods

Study Design

This UCF intervention was a parallel group, 2-arm, pragmatic waitlist randomized trial with 1:1 allocation ratio to either intervention or waitlist control. The intervention group completed the first 24-week UCF program, followed by a 24-week observation period. The control arm served as a waitlist group, receiving no intervention during the initial 24 weeks but undergoing the same intervention in the following 24 weeks after the intervention group completed. Participants in both arms completed assessment at baseline, 6-month, and 12-month follow-up. We reported using the standard protocol items: recommendation for interventional trials (Standard Protocol Items: Recommendations for Interventional Trials [SPIRIT]) checklist (Checklist 1) [29].

UCF Intervention

The UCF intervention was designed as a structured 24-week care farming curriculum and delivered by certified horticulturists and trained farmers. It was intended to engage older adults with the natural environment through therapeutic farming (eg, nature walks around sensory gardens and reflecting on nature landscape) and gardening activities (eg, plant cultivation and gardening) in green spaces, with the goal of enhancing QOL across physical, psychological, and social domains of health. Both the intervention and waitlist control arms participated in a weekly UCF program, lasting between 2.5 and 3 hours per session over 24 weeks. The program included activities such as fertilization, sowing, watering, transplanting, maintaining, and harvesting vegetables and herbs. The UCF intervention included structured, facilitated sessions involving a range of agricultural and food-growing activities, such as seedling propagation, composting, harvesting of edible greens, and hands-on food preparation using farm-grown produce. These went beyond casual gardening and were intentionally designed to promote physical activity, encourage teamwork and autonomy, and stimulate cognitive and sensory engagement. Each participant selected 1 weekday to attend weekly. All sessions offered each week were identical in content. For example, group 1 attended a Monday session for 24 weeks. This arrangement was necessary as the intervention site could not accommodate all 60-70 participants, depending on the study arm. Furthermore, smaller group sizes promoted meaningful discussion and the forming of social networks. Treatment adherence was monitored using attendance sheets. Additionally, treatment fidelity was used to monitor and ensure the consistency of intervention content delivery.

Trainers facilitated the 24-week UCF intervention, engaging them with horticultural and urban care farming activities to acquire relevant skills and knowledge. To ensure consistent content and activities delivery, trainers were provided training materials, which listed the objectives and activities (Supplementary 1 in Multimedia Appendix 1). Participants received guidance on fundamental planting techniques, information on different varieties of plants, and their nutritional value. UCF-integrated activities related to gardening and farming that fostered social connections and promoted psycho-emotional well-being were also included in the intervention. Additionally, participants gained insights into the benefits of naturally grown food without chemical additives and were encouraged to make healthier dietary choices. This intervention included other activities such as cooking, nutrition, and arts and crafts. Table 1 presents the overview of the activities that participants engaged in the intervention.

A small incentive of S$10 (US$7.35) [30] per session was reimbursed to participants to cover travel costs to the intervention site. Participants were given vegetable packs fortnightly in addition to crops they had personally cultivated. This hands-on experience provided farm-to-table experience. The practice also encouraged participants to reflect on the quality of their produce.

XSL•FO
RenderX

Study Setting

The study was held in the western region of Singapore, at a national garden equipped with growing edibles and therapeutic gardens, and a rooftop community garden. Enrolled participants began the intervention at the national garden, which includes foundational knowledge and hands-on training on plants and farming. As part of the intervention, they transitioned to a community rooftop garden, collaborating to grow, maintain, and harvest plants.

The UCF intervention was designed as a structured 24-week care farming curriculum and delivered by certified horticulturists and trained farmers. It was intended to engage older adults with the natural environment through therapeutic farming (eg, nature walks around sensory gardens and reflecting on nature landscape) and gardening activities (eg, plant cultivation and gardening) in green spaces, with the goal of enhancing QOL across physical, psychological, and social domains of health. Both the intervention and waitlist control arms participated in a weekly UCF program, lasting between 2.5 and 3 hours per session over 24 weeks. The program included activities such as fertilization, sowing, watering, transplanting, maintaining, and harvesting vegetables and herbs. The UCF intervention included structured, facilitated sessions involving a range of agricultural and food-growing activities, such as seedling propagation, composting, harvesting of edible greens, and hands-on food preparation using farm-grown produce. These went beyond casual gardening and were intentionally designed to promote physical activity, encourage teamwork and autonomy, and stimulate cognitive and sensory engagement. Each participant selected 1 weekday to attend weekly. All sessions offered each week were identical in content. For example, group 1 attended a Monday session for 24 weeks. This arrangement was necessary as the intervention site could not accommodate all 60-70 participants, depending on the study arm. Furthermore, smaller group sizes promoted meaningful discussion and the forming of social networks. Treatment adherence was monitored using attendance sheets. Additionally, treatment fidelity was used to monitor and ensure the consistency of intervention content delivery.

Inclusion and Exclusion Criteria

Participants who enrolled in the study met these inclusion criteria: (1) aged 50 and 85 years, (2) willing to take part in either the intervention or waitlist control arm, (3) agree for their data to be used for research purposes, and (4) ambulant to be able to carry out gardening tasks (eg, watering the plants). Participants were excluded if they had medical conditions that could compromise or interfere with their ability to participate, upper or lower disabilities that limit their mobility, or were unwilling to provide their medical or health records for research.

Ethical Considerations

The study has been approved by the National University of Singapore Institutional Review Board (NUS-IRB-2023 - 191). The approval has also been granted for the pilot study of this project. Any modification to the approved protocol will be submitted for review by the ethics committee. If there are any amendments, changes will be reflected in the trial registry. All participants provided written consent prior to the participation. Data collected was anonymized and kept confidential. Consent from participants was obtained to publish the results from deidentified data. To acknowledge the time and effort required, participants were provided with a S$30 (US $22.05) [30] honorarium for each data collection session (approximately 210 min). The amount was set at a level appropriate to compensate participants without constituting undue inducement, consistent with ethical guidelines for research involving human participants. Participants' data (eg, case record forms, laboratory tests, information sheets, and consents) are stored in a locked cabinet in a researcher's office. All other data (eg, biomarker results, and responses to surveys) have been deidentified. All personal identifier data (eg, names and contact information) are kept separate from the research data by a trusted third party. All data collection has been completed, and a link (otherwise known as code key) was passed to the trusted third party so the research team only works with deidentified data. After the study has been completed, all personal data will be kept for a maximum of 5 years. Research and biological data will be kept for a maximum of 10 years. During the study, the data will only be accessible to the research team. The research team has exclusive rights to the deidentified data for 36 months after the trial is completed.

Dissemination

In addition to disseminating our research findings to the funder of this study, we will disseminate our findings to other countries, the study participants, and the research community. We will also follow the authorship guidelines of the International Committee of Medical Journal Editing (ICMJE) [31].

XSL•FO
RenderX

Table . Summary of indoor and outdoor care farming activities across the 24-week intervention.

No. of sessions	Program description	Activity breakdown	Goals of intervention

15 sessions	• Participants engaged in indoor plant- and farming or gardening-related knowledge-sharing sessions led by trained horticulturist farmers. Participants learned basic skills and knowledge in maintaining a garden and acquired skills in pruning, soil mixing, transplanting, and composting, as well as knowledge of the properties of plants and vegetables. Example topics included composting, the properties and nutritional value of plants and vegetables, and gardening techniques. A few indoor sessions also included microgreen planting, which participants brought home to tend to, as well as nature-art activities designed to foster creative expression and social interaction.
	• Following the indoor activities, participants attended outdoor hands-on sessions where they engaged in green maintenance in the designated garden or farm area and practiced the gardening techniques introduced during the indoor sessions. Several outdoor sessions included therapeutic walks, during which participants took part in guided educational strolls through therapeutic gardens to learn about the unique properties of local plants and to strengthen their connection with nature. During these walks, participants engaged in sensory-based activities involving touch, scent, and visual observation, which created opportunities to reconnect with local flora and fauna and to reminisce.
	• Several of the 15 sessions at the national garden focused on nutrition, cooking, and consumption of herbs, vegetables, and fruits. Participants learned about the nutritional value and importance of consuming more plant-based diets. Live cooking demonstrations were provided, and participants sampled the cooked vegetables. Participants were encouraged to share their recipes to promote greater social interaction.
		• Indoor and outdoor activities:
		  • 20 minutes of briefing of daily activities
		  • 60 minutes of indoor gardening and sharing of farming tips
		  • 60 minutes of outdoor gardening activities with short breaks in between
		  • 10 minutes of debriefing
		• Nutrition and cooking activities:
		  • 20 minutes of briefing of daily activities
		  • 50 minutes of gardening, harvesting, and preparation
		  • 60 minutes of cooking and demonstration, food sharing, and social interaction
		  • 20 minutes of debriefing
			• Increase physical activity
			• Improve gross and fine motor skills
			• Strengthen working memory
			• Gain a sense of achievement
			• Care for plants—promote sense of purpose, meaningfulness, and hope
			• Improve nutrition knowledge and attitude toward vegetable consumption
			• Improve interpersonal interaction

XSL•FO
RenderX

No. of sessions	Program description	Activity breakdown	Goals of intervention
8 sessions	• These 8 sessions were held at the community garden near the national garden. Care farming activities involved participants working collaboratively to plan, carry out, and care for crops from initial germination through to harvest in the community garden. Participants also prepared meals together using the produce they planted and harvested.	• 15 minutes of briefing of daily activities • 120 minutes of gardening and horticultural activities with short breaks in between • 15 minutes of debriefing	• Increase physical activity • Improve physical performance • Strengthen working memory • Gain a sense of achievement • Care for plants—promote sense of purpose and meaningfulness • Improve gross and fine motor skills • Improve community engagement and social interaction
1 session	• Participants organized a collective indoor community event to showcase their harvest and promote community gardening activities. Participants from different session days came together to attend this event. Trainers assisted participants in preparing for the community event.	• 150 minutes of preparation, sharing, and exchange among participants and urban farming experts and horticulturists	• Improve place attachment • Improve community engagement and social interaction

Outcome Measures

Outcome domains were selected based on the intervention's theoretical framework and logic model (Figure S1 in Multimedia Appendix 1) and included primary outcomes of QOL, along with secondary and measures capturing psychosocial, behavioral, cognitive, and biological responses.

Primary Outcome

The primary outcome of this study was participants' QOL using the WHOQOL-BREF (World Health Organization Quality of Life−brief version), a multidimensional QOL tool developed by the World Health Organization. The instrument consists of 25 questions evaluating 4 domains of QOL, namely physical (7 items), psychological (6 items), social relationships (2 items), and environment (8 items). One item in the social dimension—"How satisfied are you with your sex life?"—was removed because it was considered culturally sensitive for this study population. The items that were negatively worded had their scores inversed. In addition, there were also 2 questions about general QOL and health satisfaction questions. The physical health domain measures individuals' biological well-being related to the body and its functions. The psychological domain examines the aspects of mental well-being and cognitive functioning. Social relationships were based on interpersonal connections and support systems. The environment domain covers factors related to surroundings and external conditions [32]. Higher scores in each of the 4 domains indicate better QOL. The score of the 4 domains was summed together, where a higher score indicated a higher QOL. The validity and reliability of the WHOQOL-BREF were tested in Singapore in Chinese, English, and Malay. The instrument is valid and reliable for assessing the QOL among English-, Chinese-, and Malay-speaking people in Singapore [33].

Secondary Outcomes

Physiological Outcomes

The physiological component included changes in cardiometabolic outcomes. The cardiometabolic outcomes consisted of the following physical measurements: (1) waist circumference and hip circumference, (2) systolic and diastolic blood pressure, (3) resting heart rate, (4) body weight, and (5) height and BMI. Weight and height were measured without shoes. BMI was calculated by dividing weight in kilograms by the square of height in meters. Systolic and diastolic blood pressures were measured after the participants had rested for at least 5 minutes [34]. Waist circumference (cm) was measured at the middle point between the lowest rib in the standing position and the iliac crest.

Frailty

Participants' frailty was assessed using the Fried frailty phenotype [35], which consists of 5 components: unintentional weight loss, exhaustion, weakness, slowness, and low physical activity. Each of the components was calculated using the computation methods described in [35, 36]. Weight loss was a self-reported item to determine if participants had lost 4.5 kg or more in the past year. Exhaustion was assessed through 2 self-reported feelings of fatigue and low energy. Physical activity was evaluated based on self-reported items on walking for 30 minutes or more a day and the amount of time spent sitting. Weakness was measured using grip strength, with scores computed according to gender and BMI-specific cut-offs, where values below the threshold indicated weakness. Slowness was assessed using a 5-m walking speed test, with thresholds for slowness determined according to gender and height; scores above the threshold indicated slowness. Frailty scores were computed and categorized into non-frail (none or 1 criterion met), pre-frail (2 criteria met), and frail (3 or more criteria met) [35, 36].

XSL•FO
RenderX

Biomarkers

Fasting blood was collected and sent to a certified laboratory. Around 10 mL of blood was collected in each data collection. In addition, clinical bioprofiling, such as the assessment of plasma glucose, hemoglobin A_{1c}, and lipid panels (total cholesterol, triglyceride, low-density lipoprotein cholesterol, and high-density lipoprotein cholesterol), and serum creatinine were examined. Plasma biomarkers were analyzed using commercially available enzyme-linked immunosorbent assay and bead-based multiplexing Luminex system immunoassay kits. Evaluated analytes involved several functions: inflammation, immune system adaptation, senescence, and cellular signaling. More specifically, the targeted biomarkers assessed cytokines involved in innate immune functions (interleukin-1 receptor antagonist, interleukin-1 beta, interleukin-6, interferon-gamma-induced protein 10, tumor necrosis factor alpha), adaptive immunity (interferon gamma, interleukin-10, interleukin-13, interleukin-17A), macrophage inflammatory markers involved in aging altering cellular communication and inflammatory state (macrophage inflammatory protein-3 beta, stromal cell-derived factor-1 alpha, matrix metalloproteinase-3, matrix metalloproteinase-9, monocyte chemoattractant protein-1, plasminogen activator inhibitor-1, granulocyte-macrophage colony-stimulating factor, vascular endothelial growth factor A), and chemokines (eotaxin, interleukin-8). Study participants with abnormal biomarkers were encouraged to consult with a physician.

Psychosocial Outcomes

A range of psychosocial measures was adopted. To measure social well-being, the Lubben social network scale [37] was used. The De Jong Gierveld Loneliness Scale [38], which has 6 items, was used to measure social and emotional loneliness. The Brief Sense of Community 8-item scale was employed to assess participants' perceptions of utilizing community resources to fulfill their needs [39]. Older adults' resilience was measured using the Brief Resilience Scale [40]. The 10-item Perceived Stress Scale was adopted to measure psychological stress [41].

Cognitive Outcomes

The Montreal Cognitive Assessment (MoCA) was adopted to evaluate cognitive function, with assessment categorized into memory, executive functioning, attention, language, visuospatial skills, and orientation domains [42]. The maximum total score was 30, with a cutoff of 26 or above indicating a normal cognitive function. Additionally, functional near-infrared spectroscopy (fNIRS) was collected. fNIRS measures hemodynamic changes in the cerebral cortex over time, reflecting neuronal activity via neurovascular coupling. Near-infrared light penetrates tissues to the cortex, where it is absorbed by oxyhemoglobin and deoxy-hemoglobin. Increased neuronal activity will lead to a net rise in oxy-hemoglobin and a slight decrease in deoxy-hemoglobin, making oxy-hemoglobin a marker of cortical activity. Participants were asked to perform a semantic verbal fluency task during each fNIRS measurement. The verbal fluency task consisted of alternating blocks of 4 control periods and 3 task periods. Each control and task period was 30 seconds long. Participants were instructed to say the names of the weekdays consecutively and generate words that belong to a particular category (eg, fruits, flowers, professions) during task periods [43]. Categories differed between task periods and fNIRS measurements. The total number of words generated per category was noted.

Dietary Outcomes

Participants' attitude and behavior toward healthier diet consumption, for example, consumption of fruits and vegetables, were measured. Participants' healthy eating behavior was also assessed by indicating the frequency of consuming food from various groups (eg, vegetables, fruits, fried snacks, sweet snacks, fast food, sweet drinks, fruit juice) and their willingness to consume fruits and vegetables. Attitude toward healthy eating was measured by an 11-item scale [44].

General Life Satisfaction

A single-item measure, "In general, how satisfied are you with your life as a whole?", was adopted. Participants ranked a 10-point scale, with 1 being "very unsatisfied" and 10 being "very satisfied." The single-item satisfaction with life measure is a validated instrument that performs similarly to the multiple-item satisfaction with life scale [45].

Exploratory Outcomes

Physical activity was measured with Fitbits. The intervention group was given a Fitbit Inspire 2 or Inspire 3 device to track their step count throughout the trial. Fitbit is an activity device that uses a tri-axis accelerometer to provide real-time feedback and allows self-monitoring of steps. Participants were encouraged to wear the device as frequently as possible to obtain the most accurate results. Fitbit intraday data were downloaded, providing minute-by-minute data that was available through a web application programming interface.

Sample Size

This study aimed to compare the QOL between the treatment and waitlist control participants, using a 1:1 allocation ratio. Based on a prior study of the overall WHOQOL measure, if the true difference-in-differences estimate between experimental and control means is 4.93 with a standard deviation of 9.29 [46], a sample of 57 intervention and 57 control subjects is required to reject the null hypothesis based on a 2-tailed significance level of 5% and power of 80%. We inflated the sample by 20% to account for dropouts and the potential loss of participants due to noncompliance and faulty measurements. Accounting for all these factors, the total sample size required for 2 arms is 138 participants.

Recruitment

Singapore residents were invited to participate in this study. We adopted a 2-prong recruitment strategy. First, we collaborated with local community partners and senior activity centers (SACs) to recruit potential participants. The findings from our feasibility study suggest that older adults are more inclined to participate when approached through familiar and trusted sources. Concurrently, we simultaneously conducted open recruitment by promoting the study via social media and word of mouth. Although this second approach engaged young-old adults, some have encouraged their parents to join the study. Our recruitment approaches ensured effective outreach

XSL•FO
RenderX

to our target audience of older adults. Recruitment occurred between February and March 2024. During the recruitment phase, the research team was in close contact with SACs and community partners to facilitate recruitment. During recruitment, interested participants, recruited from either partnered SACs or online recruitment, provided their contact details on a sign-up form. Subsequently, a research staff member contacted these participants to share more information about the study and screened them first for their age eligibility. Those who met the criteria were invited to a briefing session where more detailed information about the research was presented. Informed consent was obtained from participants who agreed to participate and were enrolled in this research study. Participation was voluntary, and participants could withdraw from the study at any time without reason or penalty.

Blinding and Assignment of Groups

Participants were randomized into either of the groups based on a computer-generated random generator using Stata 14 [47]. A first block of 130 participant IDs was randomly generated, followed by a second block of 20 due to additional recruitment. At baseline data collection, all participants and data collectors were blinded to group allocation. Randomization and allocation results were shared only at the last station after written informed consent and the completion of all baseline data to ensure allocation concealment. To maintain neutrality, the research staff who generated the randomization were independent of the staff responsible for revealing group allocations to participants. As the nature of the intervention required participants to be involved in the study, both participants and research staff were eventually aware of group allocation. All outcome assessments were conducted by trained assessors who were independent of intervention delivery. Assessors were blinded to participant group allocation at baseline and 6-month and 12-month follow-up, where feasible. However, full blinding could not always be maintained during follow-up, particularly when participants disclosed their participation during data collection. To mitigate potential bias, the assessors were trained to ensure accurate and objective data collection, and validated instruments were used to reduce subjectivity in cognitive and psychosocial outcomes.

Data Collection and Analysis

Data Collection

Baseline data were collected after informed consent was obtained. Table 2 outlines the measures and instruments that were collected in each of the 3 assessment time points. At baseline (T0) and 6-month follow-up (T1), participants completed the survey, blood test, fNIRS, MoCA, and physiological assessments at the National University of Singapore. At the 12-month follow-up (T2), all the assessments completed at T0 and T1 were repeated, except for the blood test and fNIRS. Participants were invited to return to the same location for T2 data collection. However, we anticipated that some participants may be unwilling or unable to return in person. For those who did not attend the in-person assessment, a survey link was sent to them to allow them to self-administer remotely. The rest of the assessment was not collected from these participants. To maximize the data collected, participants were reimbursed a small amount of S$30 (US $22.05) [30] each for T0 and T1 assessments, and S$20 (US $14.70) [30] for T2 in-person assessment or S$10 (US $7.35) [30] for completing the T2 online self-administered survey. The physiological assessments were collected by trained research assistants, while blood samples were collected by certified phlebotomists or registered nurses. The survey was available in English and Chinese and was self-administered by participants; however, trained research assistants assisted participants who were illiterate in either of the languages to complete the survey. fNIRS data were collected by a trained engineer.

XSL•FO
RenderX

Table . Overview of study outcomes and measurement instruments used.

	Feasibility	Baseline	6-month	12-month
Demographics				
• Age	✓	✓	✓	
• Gender	✓	✓	✓	
• Ethnicity	✓	✓	✓	
• Marital status	✓	✓	✓	
• Education	✓	✓	✓	
• Pre-existing diseases, self-reported		✓	✓	✓
• Smoking and alcohol consumption	✓	✓	✓	
• Community participation	✓	✓	✓	
Quality of life: WHOQOL-BREF[a]		✓	✓	✓
EQ-5D-5L	✓	✓	✓	✓
Psychosocial				
• Lubben social network scale	✓	✓	✓	✓
• De Jong Gierveld loneliness scale	✓	✓	✓	✓
• Brief sense of community		✓	✓	✓
• Brief resilience scale	✓	✓	✓	✓
• Perceived stress scale	✓	✓	✓	✓
• Satisfaction with life	✓	✓	✓	✓
Physical activity				
• Fitbit tracking	✓	✓	✓	
Dietary				
• Attitude toward consumption of fruits and vegetables	✓	✓	✓	✓
• Fruits and vegetable consumption frequency (behavior)	✓	✓	✓	✓
Cognition				
• Functional near-infrared spectroscopy		✓	✓	
• Montreal cognitive assessment	✓	✓	✓	✓

XSL•FO
RenderX

	Feasibility	Baseline	6-month	12-month
Physiological outcomes				
• Body weight and height	✓	✓	✓	✓
• Waist and hip circumference	✓	✓	✓	✓
• Blood pressure	✓	✓	✓	✓
• Resting heart rate	✓	✓	✓	✓
• Fried phenotype of frailty		✓	✓	✓
Biomarkers				
• Fasting plasma glucose, HbA_{1c}[b]		✓	✓	
• Fasting insulin		✓	✓	
• Lipid profile: triglyceride, total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol		✓	✓	
• Cytokines: interleukin-1 receptor antagonist, interleukin-1 beta, interleukin-6, interferon-gamma-induced protein 10, tumor necrosis factor alpha		✓	✓	
• Chemokines: eotaxin, interleukin-8		✓	✓	
• Adaptive immunity: interferon gamma, interleukin-10, interleukin-13, interleukin-17A		✓	✓	
• Macrophage inflammatory markers: macrophage inflammatory protein-3 beta, stromal cell-derived factor-1 alpha, matrix metalloproteinase-3, matrix metalloproteinase-9, monocyte chemoattractant protein-1, plasminogen activator inhibitor-1, granulocyte-macrophage colony-stimulating factor, vascular endothelial growth factor A		✓	✓	

[a]WHOQOL-BREF: World Health Organization Quality of Life−brief version.

[b]HbA_{1c}: hemoglobin A_{1c}.

XSL•FO
RenderX

Data Management

All registered participants' information was deidentified and given a unique identification number. This identification number was used on all data collected. All contact information of potential participants collected via in-person recruitment or online platforms was stored electronically and password protected. All survey data, whether completed in-person or remotely, were collected via a secure online link and stored electronically. Physiological data were initially recorded on paper and entered into an electronic database. fNIRS and MoCA assessments were conducted in person, and the resulting scores were entered and stored electronically. Blood samples were collected during in-person visits, deidentified, and immediately sent to accredited laboratories for analysis. The results were then stored electronically by the respective laboratories. All electronic records were stored on a secure institutional server, which was protected by password access and restricted to authorized personnel only. All personal information of potential or enrolled participants will be kept a maximum of 5 years before being discarded. The research data used in the publication will be kept for a maximum of 10 years before being discarded. If participants provide informed consent for future studies, their biological samples will be securely stored for a maximum of 10 years. Otherwise, all biological materials will be discarded immediately after the analyses for this study are completed.

Statistical Analysis

The primary analysis, intention-to-treat, will be conducted using a linear mixed effects model with overall WHOQOL score as the dependent variable and the interaction effect of intervention allocation and time as the independent variable. Domain-specific scores will also be analyzed. To address potential bias arising from missing data due to attrition or incomplete responses, our primary longitudinal analyses will employ linear mixed-effects models, which are robust to unbalanced data structures and can handle missing observations under the missing-at-random assumption. While listwise deletion may be used for cross-sectional comparisons or descriptive analyses, this will be clearly indicated. Secondary outcomes will be analyzed using similar methods. These analyses will be conducted at 6-month follow-up to investigate the effectiveness of the intervention as well as post-intervention behavioral maintenance at 12 months. Sensitivity analysis using the per-protocol approach will also be carried out. Baseline characteristics of participants who discontinued or were lost to follow-up will be compared to those who completed the study to investigate differential dropout between intervention and control groups.

Due to potential imbalances between the groups, further analysis techniques, such as adjustment or matching of baseline characteristics, will be used in the difference-in-difference analysis. Specifically, we will adjust or match on key baseline covariates (eg, age, gender, years of education, marital status, and housing type) to improve group comparability at baseline in case of any imbalance despite randomization. This approach improves covariate balance and enhances the comparability of groups prior to applying linear mixed models for outcome analyses. Conducting both intention-to-treat and per-protocol analyses for this behavioral intervention allowed comparison of the results and provided a comprehensive understanding of the treatment effect. All evaluations were and will be performed using a 2-sided test at the 5% level of significance.

Method Monitoring

There was no data monitoring board. The study team had complied with data collection and management procedures approved by the National University of Singapore ethics committee (NUS-IRB-2023 - 191) and abided by the rules of medical confidentiality. We recognized the importance of safety monitoring in community-based interventions involving older adults. Accordingly, safety precautions and procedures were implemented in line with the university's ethics committee. Trainers were briefed on these safety procedures, and, during sessions, they consistently reminded participants to exercise caution and check for potential tripping hazards while engaging in activities. Additionally, weekly post-intervention debriefings with trainers were conducted to identify any safety concerns or adverse events related to the outdoor farming sessions (Supplementary 2 in Multimedia Appendix 1). No adverse events were reported throughout the intervention, suggesting that it was conducted safely and with appropriate precautions in place. There was no independent audit of the trial.

To ensure consistency in intervention delivery, all trainers were provided with a structured facilitator manual that outlined session-by-session objectives, activities, and key content to be delivered. Teaching presentation slides were also shared with all trainers. Additionally, training sessions were conducted prior to the start of the intervention to align all trainers on core delivery principles, safety protocols, and communication style. To ensure fidelity, staff completed checklists after each session to document adherence to the protocol and note any deviations. Additionally, a member of the research team attended a random sample of sessions for verification purposes. Weekly debriefing meetings were conducted among the research team and trainers to determine whether the intervention was executed as planned, create feedback loops, and allow the early identification of problems and problem-solving to enhance implementation. We have included this fidelity assessment document in Supplementary 3 in Multimedia Appendix 1.

Economic Evaluation

A cost-utility analysis will be conducted from the provider and societal perspective to assess the potential value of scaling up the UCF intervention, using QALYs as the primary effectiveness metric. This economic evaluation will be exploratory in nature and designed to inform future trial design and policy translation.

Resource use associated with the UCF intervention will be systematically recorded, including staff time (trainers), materials (gardening supplies), venue rental, and transportation. Unit costs will be derived from internal budget records and prevailing trainers' wage rates. Any health care utilization during the intervention period (eg, GP visits, hospitalizations) will be tracked using participant self-reports and costed using standard unit prices obtained from publicly available Ministry of Health fee benchmarks or institutional databases.

Effectiveness will be expressed in terms of QALYs, derived from utility scores measured using the EQ-5D-5L instrument

XSL•FO
RenderX

at baseline and postintervention. Area under the curve methods will be applied to calculate QALYs over the 24-week period. The decision tree model will be developed using the TreeAge software to compare the two scenarios: (1) the waitlist control group and (2) the urban care farming program. An incremental cost-effectiveness ratio (ICER) will be calculated as the ratio of the change in cost over QALY. The outputs of these analyses will be costs, QALY, and the incremental cost-effectiveness ratio to reflect the program's cost-effectiveness. As Singapore does not have an official willingness-to-pay threshold for QALYs, a cost-effectiveness acceptability curve will also be plotted to show the probability that the intervention is cost-effective at varying willingness-to-pay thresholds.

Results

This randomized controlled trial (RCT) commenced in April 2024, with the intervention group undergoing the intervention for the first 24 weeks, which ended in September 2024. Prior to the start of the data collection and intervention, participants were screened for eligibility, and informed consent was obtained. A total of 137 participants were enrolled, with 67 randomized to the intervention group and 70 to the control group. Baseline data were collected in March 2024, and the 6-month follow-up was completed in September 2024. The waitlist control group began the 24-week UCF intervention in late September 2024 and completed it in April 2025, with a 3-week break over Christmas, New Year, and Lunar New Year period. Data collection for the 12-month follow-up concluded in May 2025. Analysis of the baseline and 6-month follow-up data is currently ongoing. The participant timeline is presented as a flowchart in Figure 1.

XSL•FO
RenderX

Figure 1. Illustration of the participant flow including the time schedule of enrollment, intervention, and assessments. Sample sizes (n) reported above were based on intention-to-treat random allocation. A "dropout" was defined as a participant who voluntarily withdrew from the study or could not be recontacted, resulting in missing data.

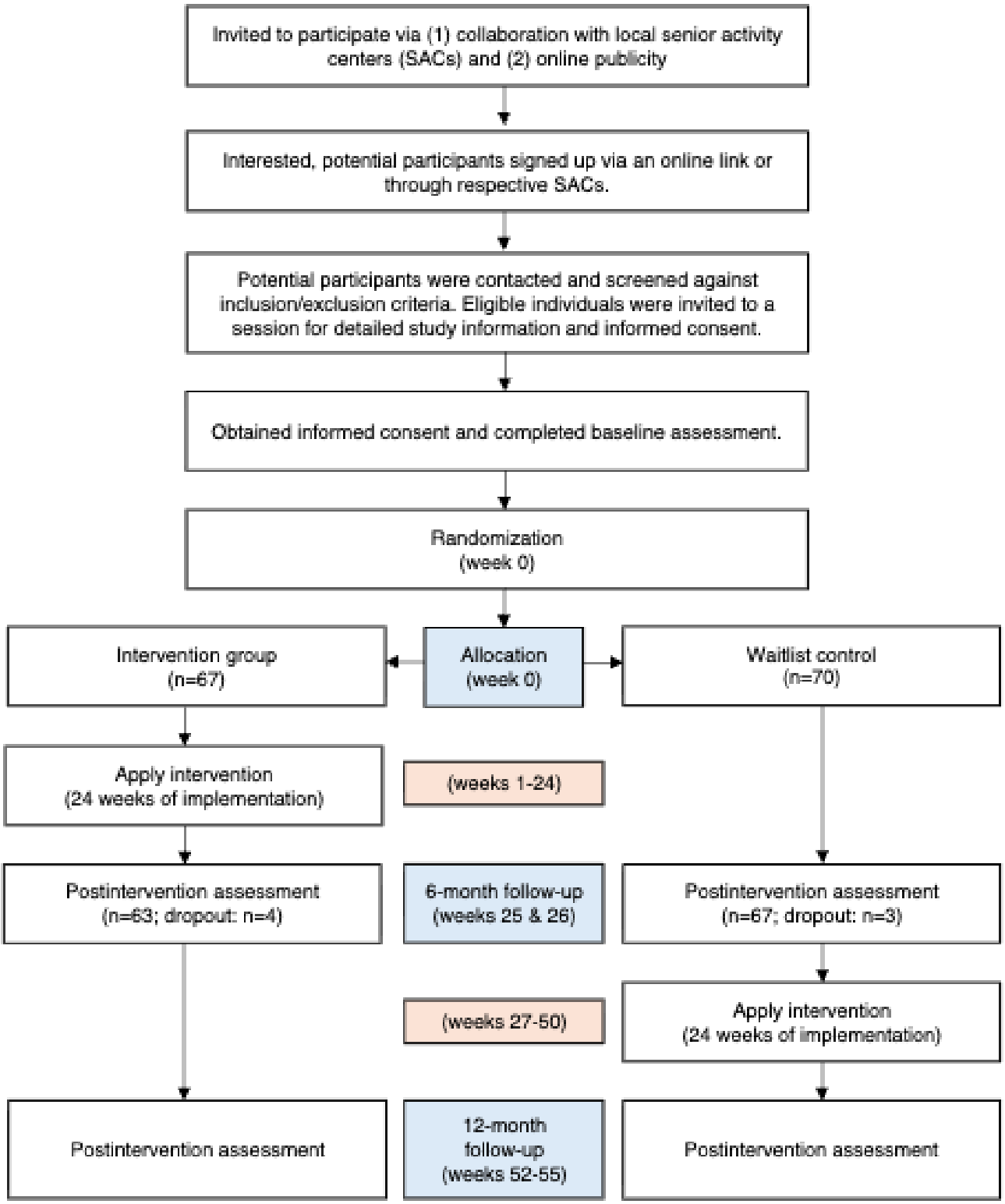

Discussion

This paper describes the design of a waitlist randomized trial to evaluate the effectiveness of UCF, a complex behavioral intervention, in improving the QOL among older adults. UCF leverages the principles of biophilia and therapeutic horticulture to promote active and healthy aging. Previous studies have suggested that nature-based interventions may provide benefits such as improved mood [48] and increased physical activity [21]. While pilot studies have been conducted to illustrate health benefits in Singapore [49,50], these studies are typically underpowered and not designed with scalability in mind, leaving gaps in understanding their long-term impact and cost-effectiveness. Moreover, the existing evidence remains mixed, highlighting the need for well-powered,

XSL•FO
RenderX

methodologically robust trials to assess the effectiveness of care farming as a public health intervention.

Our study aims to address these gaps through a robust RCT design, incorporating a waitlist control group to strengthen causal inference. By including objective physiological outcomes (eg, BMI, blood pressure), validated psychological and cognitive measures, and QOL assessments aligned with the WHOQOL framework, the trial aims to generate comprehensive and policy-relevant evidence. This study is also timely and relevant to Singapore's aging population and growing interest in a community-driven approach to aging in place. With growing public interest in urban farming, this intervention leverages that momentum to promote health through nature-based, skill-building activities that may be sustained beyond the program period. While our study has been designed for scale-up, this study primarily serves to evaluate acceptability and preliminary outcomes. Further research, including a planned cost-effectiveness and process evaluation, will be necessary to assess implementation feasibility and inform national policy translation.

Nevertheless, this study has several limitations. The effect size used for sample size estimation was based on a prior WHOQOL study, which may not fully reflect the characteristics of the Singaporean older population or the unique aspects of the UCF intervention. The calculated sample size also includes a 20% buffer to account for expected attrition in long-term community interventions. Sensitivity analyses for secondary outcomes were not included. One methodological challenge is estimating the effect sizes of repeated measures data. To address this, the study will employ linear mixed-effects models and robust standard errors to derive reliable difference-in-difference estimates. This statistical approach enables the evaluation of longitudinal data, accounting for within-subject correlation and heterogeneity in responses over time. Attrition over the 24-week period may limit the ability to evaluate longer-term outcomes, though mitigation strategies such as regular participant engagement and reminders have been implemented. Next, the use of self-reported measures may introduce recall and reporting bias; however, the inclusion of biomarkers and clinical assessments partially mitigates this. Finally, the findings may be less generalizable to more frail or homebound populations, as the study sample consists of relatively healthy, community-dwelling individuals who are able to travel and commit to a 24-week intervention. As such, the study sample is likely skewed toward relatively healthy and motivated individuals, which may limit the generalizability of our findings. While the results offer important insights into the potential benefits of urban care farming, future work should explore adaptations of the intervention to accommodate more vulnerable older adults and assess its effectiveness in those populations. For future trials, we also recommend a more structured approach to adverse event monitoring, including standardized incident reporting forms and real-time logs, to further strengthen safety assurance and protocol reproducibility.

In conclusion, this RCT evaluates the impact of a multicomponent UCF behavioral intervention on QOL, across physical, psychological, social, and environmental domains. We will also assess whether greater compliance with UCF sessions improves QOL, mental health, social outcomes, physical activity, and dietary attitudes toward consuming fruits and vegetables. If effective, this intervention could serve as a model for future community−based public health research targeting the QOL and well-being of older adults. The findings will also contribute to future health economic analyses and support evidence-based policy formulation under the national Age Well SG initiative.

Acknowledgments

We thank Stella Shwe for the helpful discussions and administrative support that have contributed to this work. We also thank Guo Xueying for her Python programming support on the collection of Fitbit intraday minute-by-minute data through the web application programming interface. We also thank the trainers, Fion Tham, Mei Kheng Lee, and Josephine Tan, for providing engaging and informative hands-on training sessions. We would also like to thank the Jurong Lake Gardens team from NParks, Kamisah Atan, Margaret Sin, and Pearlina Sim for going the extra mile to ensure participants in this study had access to gardens and learning materials. We are also grateful to all the participants who generously shared their time and experiences for this research. We also acknowledge the Healthy Longevity Translational Research Programme Stem Cell Core Facility for their collaboration and support in biomarker analyses.

Funding

This study was supported by the Singapore Ministry of Health's National Medical Research Council under its HPHSR Clinician Scientist Award (HCSAINV22jul-0005, ID: MOH-001304). The funders had no role in study design, data collection or analysis, preparation of the manuscript, or decision to publish. The views and opinions expressed in this article are those of the authors and do not necessarily reflect those of funders or institutions of belonging.

Data Availability

The research team has exclusive rights to the deidentified data for 36 months after the trial is completed. After that, the data are available from the corresponding author upon reasonable request.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: CC.
Data curation: CC, ZS, JL.
Funding acquisition: CC.
Investigation: CC.
Methodology: CC, ZS, SA, MFFC, CNO, AS, LLT, XKT, JP, KO, JH, MK, RH, and ED.
Project administration: CC, JP.
Resources: XKT, JH, MK, RH, ED.
Supervision: CC.
Visualization: CC, ZS, JL.
Writing (original draft): CC, JL, and KA.
Writing (review & editing): CC, ZS, JL, KA, SA, MFFC, CNO, AS, LLT, XKT, JP, KO, JH, MK, RH, and ED.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Logic model (Figure S1), sample curriculum (Supplementary 1), and implementation monitoring sample documents (Supplementary 2 and 3) for urban care farming intervention (UCF).
[DOCX File, 160 KB - resprot_v15i1e78584_app1.docx]

Checklist 1
SPIRIT checklist.
[DOCX File, 49 KB - resprot_v15i1e78584_app2.docx]

References

1. Garmany A, Yamada S, Terzic A. Longevity leap: mind the healthspan gap. NPJ Regen Med 2021 Sep 23;6(1):57. [doi: 10.1038/s41536-021-00169-5] [Medline: 34556664]
2. Volpi E, Nazemi R, Fujita S. Muscle tissue changes with aging. Curr Opin Clin Nutr Metab Care 2004 Jul;7(4):405-410. [doi: 10.1097/01.mco.0000134362.76653.b2] [Medline: 15192443]
3. Scahill RI, Frost C, Jenkins R, Whitwell JL, Rossor MN, Fox NC. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. Arch Neurol 2003 Jul;60(7):989-994. [doi: 10.1001/archneur.60.7.989] [Medline: 12873856]
4. Peters R. Ageing and the brain. Postgrad Med J 2006 Feb;82(964):84-88. [doi: 10.1136/pgmj.2005.036665] [Medline: 16461469]
5. Weyerer S, Eifflaender-Gorfer S, Köhler L, et al. Prevalence and risk factors for depression in non-demented primary care attenders aged 75 years and older. J Affect Disord 2008 Dec;111(2-3):153-163. [doi: 10.1016/j.jad.2008.02.008] [Medline: 18372049]
6. Nicholson NR. A review of social isolation: an important but underassessed condition in older adults. J Prim Prev 2012 Jun;33(2-3):137-152. [doi: 10.1007/s10935-012-0271-2] [Medline: 22766606]
7. Paulo DA. “Like a knife poking my heart”: loss, loneliness and the killing pain of elderly depression. Channel NewsAsia. 2018. URL: https://www.channelnewsasia.com/cnainsider/elderly-depression-lonely-dementia-chronic-illness-treatment-924151 [accessed 2025-12-24]
8. More seniors in Singapore taking own lives. The Straits Times. 2015. URL: https://www.straitstimes.com/singapore/more-seniors-in-singapore-taking-own-lives [accessed 2025-12-24]
9. Mental health of older adults. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/mental-health-of-older-adults [accessed 2025-12-24]
10. Gigliotti CM, Jarrott SE. Effects of horticulture therapy on engagement and affect. Can J Aging 2005;24(4):367-377. [doi: 10.1353/cja.2006.0008]
11. Oh KH, Shin WS, Khil TG, Kim DJ. Six-step model of nature-based therapy process. Int J Environ Res Public Health 2020 Jan 21;17(3):685. [doi: 10.3390/ijerph17030685] [Medline: 31973063]
12. Murray J, Wickramasekera N, Elings M, et al. The impact of care farms on quality of life, depression and anxiety among different population groups: a systematic review. Campbell Syst Rev 2019 Dec;15(4):e1061. [doi: 10.1002/cl2.1061] [Medline: 37131853]
13. Darcy PM, Armitt H, Hurd A, Paton LW, White PCL, Coventry PA. Green social prescribing: a before and after evaluation of a novel community‐based intervention for adults experiencing mental health problems. Health Soc Care Community 2025 Jan;2025(1):2016261. [doi: 10.1155/hsc/2016261]

XSL•FO
RenderX

14. White MP, Hartig T, Martin L, et al. Nature-based biopsychosocial resilience: an integrative theoretical framework for research on nature and health. Environ Int 2023 Nov;181:108234. [doi: 10.1016/j.envint.2023.108234] [Medline: 37832260]
15. WHOQOL-BREF. World Health Organization. 2012. URL: https://www.who.int/tools/whoqol/whoqol-bref [accessed 2025-12-27]
16. Paúl C, Ribeiro O, Teixeira L. Active ageing: an empirical approach to the WHO model. Curr Gerontol Geriatr Res 2012;2012:382972. [doi: 10.1155/2012/382972] [Medline: 23193396]
17. Rudnicka E, Napierała P, Podfigurna A, Męczekalski B, Smolarczyk R, Grymowicz M. The World Health Organization (WHO) approach to healthy ageing. Maturitas 2020 Sep;139:6-11. [doi: 10.1016/j.maturitas.2020.05.018] [Medline: 32747042]
18. Ageing and health. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health [accessed 2025-12-24]
19. Kim SO, Wu Y, Lee SM, Park SA. Frontal cortical activity during care farming activities of adults using electroencephalogram. Am Soc Hortic Sci 2024;59(10):1443-1450. [doi: 10.21273/HORTSCI17890-24]
20. Ura C, Okamura T, Yamazaki S, et al. Rice farming care as a novel method of green care farm in East Asian context: an implementation research. BMC Geriatr 2021 Apr 9;21(1):237. [doi: 10.1186/s12877-021-02181-2] [Medline: 33836669]
21. Park SA, Oh SR, Lee KS, Son KC. Electromyographic analysis of upper limb and hand muscles during horticultural activity motions. Am Soc Hortic Sci 2013;23(1):51-56. [doi: 10.21273/HORTTECH.23.1.51]
22. Leck C, Upton D, Evans N. Growing well-beings: the positive experience of care farms. Br J Health Psychol 2015 Nov;20(4):745-762. [doi: 10.1111/bjhp.12138] [Medline: 25810265]
23. Kaley A, Hatton C, Milligan C. Therapeutic spaces of care farming: transformative or ameliorating? Soc Sci Med 2019 Apr;227:10-20. [doi: 10.1016/j.socscimed.2018.05.011] [Medline: 29754783]
24. Finnanger-Garshol B, Pedersen I, Patil G, Eriksen S, Ellingsen-Dalskau LH. Emotional well-being in people with dementia—a comparative study of farm-based and regular day care services in Norway. Health Soc Care Community 2022 Sep;30(5):e1734-e1745. [doi: 10.1111/hsc.13601] [Medline: 34622521]
25. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. Int J Nurs Stud 2024 Jun;154:104705. [doi: 10.1016/j.ijnurstu.2024.104705] [Medline: 38564982]
26. Coventry PA, Brown JE, Pervin J, et al. Nature-based outdoor activities for mental and physical health: systematic review and meta-analysis. SSM Popul Health 2021 Dec;16:100934. [doi: 10.1016/j.ssmph.2021.100934] [Medline: 34646931]
27. Thomson LJ, Morse N, Elsden E, Chatterjee HJ. Art, nature and mental health: assessing the biopsychosocial effects of a "creative green prescription" museum programme involving horticulture, artmaking and collections. Perspect Public Health 2020 Sep;140(5):277-285. [doi: 10.1177/1757913920910443] [Medline: 32449492]
28. Fernández-Ballesteros R, Sánchez-Izquierdo M, Santacreu M. Active aging and quality of life. In: Rojo-Pérez F, Fernández-Mayoralas G, editors. Handbook of Active Ageing and Quality of Life: From Concepts to Applications: Springer International Publishing; 2021:15-42. [doi: 10.1007/978-3-030-58031-5_2]
29. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
30. Exchange rates: USD to SGD. Monetary Authority of Singapore. URL: https://eservices.mas.gov.sg/Statistics/msb/ExchangeRates.aspx [accessed 2026-01-12]
31. Defining the roles of authors and contributors. International Committee of Medical Journal Editing (ICMJE). 2020. URL: https://icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html [accessed 2025-12-24]
32. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
33. Cheung YB, Yeo KK, Chong KJ, Khoo EY, Wee HL. Reliability and validity of the English-, Chinese- and Malay-language versions of the World Health Organization Quality of Life (WHOQOL-BREF) questionnaire in Singapore. Ann Acad Med Singap 2017 Dec;46(12):461-469. [Medline: 29355283]
34. Pickering TG, Hall JE, Appel LJ, et al. Recommendations for blood pressure measurement in humans and experimental animals: Part 1: blood pressure measurement in humans: a statement for professionals from the Subcommittee of Professional and Public Education of the American Heart Association Council on High Blood Pressure Research. Hypertension 2005 Jan;45(1):142-161. [doi: 10.1161/01.HYP.0000150859.47929.8e] [Medline: 15611362]
35. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
36. Bambrick P, Phelan N, Grant E, et al. Diet and Exercise for FRAILty (DEFRAIL): protocol for a study to examine the effect of a novel community-based group exercise and nutritional intervention, designed to reverse frailty in older adults. BMJ Open 2021 Jun 15;11(6):e042408. [doi: 10.1136/bmjopen-2020-042408] [Medline: 34130954]

XSL•FO
RenderX

37. Lubben J, Blozik E, Gillmann G, et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. Gerontologist 2006 Aug;46(4):503-513. [doi: 10.1093/geront/46.4.503] [Medline: 16921004]
38. Gierveld JDJ, Tilburg TV. A 6-item scale for overall, emotional, and social loneliness: confirmatory tests on survey data: confirmatory tests on survey data. Res Aging 2006;28(5):582-598. [doi: 10.1177/0164027506289723]
39. Peterson NA, Speer PW, McMillan DW. Validation of A brief sense of community scale: Confirmation of the principal theory of sense of community. Journal Community Psychology 2008 Jan;36(1):61-73. [doi: 10.1002/jcop.20217]
40. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. Int J Behav Med 2008;15(3):194-200. [doi: 10.1080/10705500802222972] [Medline: 18696313]
41. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
42. Nasreddine ZS, Phillips NA, Bédirian V, et al. Montreal Cognitive Assessment (MOCA). APA PsycTests. 2005. URL: https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft27279-000 [accessed 2026-02-05]
43. Heinzel S, Metzger FG, Ehlis AC, et al. Aging-related cortical reorganization of verbal fluency processing: a functional near-infrared spectroscopy study. Neurobiol Aging 2013 Feb;34(2):439-450. [doi: 10.1016/j.neurobiolaging.2012.05.021] [Medline: 22770542]
44. Saba A, Sinesio F, Moneta E, et al. Measuring consumers attitudes towards health and taste and their association with food-related life-styles and preferences. Food Qual Prefer 2019 Apr;73:25-37. [doi: 10.1016/j.foodqual.2018.11.017]
45. Cheung F, Lucas RE. Assessing the validity of single-item life satisfaction measures: results from three large samples. Qual Life Res 2014 Dec;23(10):2809-2818. [doi: 10.1007/s11136-014-0726-4] [Medline: 24890827]
46. Chan HS, Chu HY, Chen MF. Effect of horticultural activities on quality of life, perceived stress, and working memory of community-dwelling older adults. Geriatr Nurs 2022;48:303-314. [doi: 10.1016/j.gerinurse.2022.10.016] [Medline: 36347114]
47. Stata. URL: https://www.stata.com/ [accessed 2025-12-24]
48. Soga M, Gaston KJ, Yamaura Y. Gardening is beneficial for health: a meta-analysis. Prev Med Rep 2017 Mar;5:92-99. [doi: 10.1016/j.pmedr.2016.11.007] [Medline: 27981022]
49. Ng KST, Sia A, Ng MKW, et al. Effects of horticultural therapy on Asian older adults: a randomized controlled trial. Int J Environ Res Public Health 2018 Aug 9;15(8):1705. [doi: 10.3390/ijerph15081705] [Medline: 30096932]
50. Wong GCL, Ng TKS, Lee JL, et al. Horticultural therapy reduces biomarkers of immunosenescence and inflammaging in community-dwelling older adults: a feasibility pilot randomized controlled trial. J Gerontol A Biol Sci Med Sci 2021 Jan 18;76(2):307-317. [doi: 10.1093/gerona/glaa271] [Medline: 33070170]

Abbreviations

fNIRS: functional near-infrared spectroscopy
MoCA: Montreal Cognitive Assessment
NBI: nature-based intervention
NBRT: nature-based biopsychosocial resilience theory
QALY: quality-adjusted life year
QOL: quality of life
RCT: randomized controlled trial
SAC: senior activity center
UCF: urban care farming
WHOQOL: World Health Organization Quality of Life
WHOQOL-BREF: World Health Organization Quality of Life−brief version

Edited by J Sarvestan; submitted 05.Jun.2025; peer-reviewed by A AL-Asadi, J Hawkins; revised version received 16.Oct.2025; accepted 05.Nov.2025; published 25.Feb.2026.

Please cite as:
Chen C, Lam J, Shi Z, Akksilp K, Aw S, Chong MFF, Ong CN, Sia A, Thang LL, Tham XK, Pong J, Omar K, Hung J, Kukumberg M, Ho R, Diehl E
Urban Care Farming to Enhance Quality of Life Among Older Adults: Protocol for a Waitlist Randomized Trial
JMIR Res Protoc 2026;15:e78584
URL: https://www.researchprotocols.org/2026/1/e78584
doi:10.2196/78584

XSL•FO
RenderX

© Cynthia Chen, Jocelin Lam, Zhenye Shi, Katika Akksilp, Su Aw, Mary Foong-Fong Chong, Choon Nam Ong, Angelia Sia, Leng Leng Thang, Xin Kai Tham, JunXiang Pong, Kartini Omar, Joan Hung, Marek Kukumberg, Roger Ho, Elizabeth Diehl. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of the Effectiveness of Injectable Platelet-Rich Fibrin as an Adjuvant to One-Stage Full-Mouth Disinfection in Patients with Stage II and III Periodontitis: Protocol for a Randomized Clinical Trial

Mahima Kothekar, BDS; Pavan Bajaj, MDS; Sneha Dare, BDS; Shivani Thakre, BDS

Department of Periodontics and Implantology, Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Corresponding Author:
Mahima Kothekar, BDS
Department of Periodontics and Implantology, Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Abstract

Background: Periodontitis is a chronic inflammatory disease that leads to the progressive destruction of the tooth-supporting structures, including the periodontal ligament, alveolar bone, and gingival tissues, leading to tooth mobility, ultimately resulting in potential tooth loss if left untreated. The new classification of periodontitis helps in establishing an appropriate diagnosis and planning treatment according to disease severity. One-stage full-mouth disinfection (OS-FMD), using chlorhexidine, has shown better outcomes than traditional quadrant-wise therapy. Platelet-Rich Fibrin offers enhancement in healing outcomes and regeneration due to sustained release of growth factors. Injectable platelet-rich fibrin (i-PRF) shows promising results in promoting tissue regeneration, reducing inflammation, and improving periodontal therapy outcomes. To date, no clinical study has been carried out for the assessment of the efficacy of i-PRF as an adjuvant in OS-FMD.

Objective: This study aims to assess the efficacy of i-PRF as an adjuvant in OS-FMD therapy based on its clinical outcomes in terms of plaque index (PI), papillary bleeding index (PBI), probing depth (PD), and clinical attachment loss (CAL) in patients with stage II and stage III periodontitis.

Methods: This randomized clinical trial will include 26 systemically healthy patients diagnosed with stage II and III periodontitis, selected from the Outpatient Department of Periodontics at Sharad Pawar Dental College, Sawangi (Meghe), Wardha. Participants will be randomly assigned to one of 2 groups using a parallel-arm design to ensure unbiased allocation. The control group will undergo OS-FMD therapy involving subgingival scaling and root planing for the entire dentition, which will be performed within 24 hours, supplemented by application of chlorhexidine intraorally, including mouth rinsing, pocket irrigation, and tongue cleansing. The test group will receive the same OS-FMD protocol as the control group, in addition to subgingival delivery of i-PRF in all periodontal pockets 1 week post therapy. Assessment of clinical parameters, including PI, PBI, PD, and CAL, will be done at baseline, 3 months, and 6 months. To evaluate intragroup and intergroup differences, statistical analysis will be conducted using appropriate methods, including the Wilcoxon signed-rank test. A P value <.05 will be considered statistically significant.

Results: The study was enrolled in June 2025 and is scheduled to conclude post assessments and analyses by the end of 2026. The accessibility of the study results is anticipated in early 2027.

Conclusions: This study underscores the effectiveness of i-PRF as an adjuvant in OS-FMD therapy on the basis of assessment of clinical parameters. We hypothesize that the use of i-PRF as an adjuvant to OS-FMD will result in superior clinical outcomes beyond the antimicrobial benefits achieved with chlorhexidine in standard OS-FMD, as evidenced by CAL gain, reduction in PD, and reduction in the scores of PI and PBI, due to regenerative and anti-inflammatory properties of i-PRF along with enhanced healing potential.

Trial Registration: Clinical Trials Registry India CTRI/2025/06/088718; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=MTMzNTAy&Enc=&userName=

International Registered Report Identifier (IRRID): DERR1-10.2196/81032

(JMIR Res Protoc 2026;15:e81032) doi:10.2196/81032

KEYWORDS

one-stage full-mouth disinfection; OS-FMD; full mouth disinfection; FMD; injectable platelet rich fibrin; i-PRF; adjuvant

XSL•FO
RenderX

Introduction

Background

Periodontitis is a chronic, inflammatory disease caused by multiple factors that target the tooth supporting structures, including the gingiva, periodontal ligament, alveolar bone, and cementum. It is initiated by dysbiotic microbial biofilms and sustained by a destructive host immune response, ultimately resulting in pocket formation, clinical attachment loss (CAL), and resorption of alveolar bone. Irrespective of age, gender, and ethnicity, it affects individuals across all demographics and is driven by specific microorganisms or the group of microorganisms further responsible for the hard and soft tissue destruction of the periodontium. The new classification system has provided an enhancement in diagnosis and treatment by incorporating staging and grading criteria, which help practitioners in the assessment of disease severity and plan appropriate treatment. Stage II periodontitis, according to the current classification of periodontal and peri-implant diseases and conditions (2017), involves moderate symptoms, such as 5 mm probing depth (PD) and 3 to 4 mm CAL, involving horizontal bone loss without loss of tooth due to periodontitis. Stage III is more severe, with ≥6 mm PD, ≥5 mm CAL, ≥3 mm vertical bone loss, class II or III furcation involvement, and tooth loss due to periodontitis of less than or equal to 4 teeth [1]. Periodontal therapy primarily aims to halt the inflammation by minimizing the microbial load of periodontal pathogens. Its success largely depends upon the effective and complete removal of both subgingival and supragingival bacterial biofilms, smear layer, bacterial endotoxins, and contaminated radicular cementum [2,3].

Periodontal disease could be managed by performing quadrant-wise scaling and root planing (SRP) in multiple 7 - to 14-day spaced visits [4]. This conventional strategy of management of periodontal disease was reevaluated when the concept of full-mouth disinfection (FMD) was introduced in the early 1990s [5]. FMD is a treatment approach for periodontitis that aims to manage periodontal pockets and prevent reinfection from other reservoirs of bacteria in the oral cavity, such as the tongue, tonsils, and mucous membrane. It involves performing SRP in a single session across the entire mouth, using antibacterial agents such as chlorhexidine. The effectiveness of this method has been evaluated in multiple clinical studies [6]. Chlorhexidine adheres well to teeth and oral tissues, showing high substantivity and exhibiting minimal irritation. Several investigations resulted in a reduction of bacterial cross-contamination after performing full-mouth SRP along with disinfection within the span of 24 hours and have offered better clinical results and microbiological outcomes compared to traditional stepwise treatment protocol [7-9].

Platelet concentrates have gained much popularity in regenerative dentistry and medicine due to their excellent healing properties and ability to promote neovascularization. While platelet-rich plasma (PRP) is obtained after centrifugation of blood with the inclusion of anticoagulants which was found to impair cellular inflammatory activation, ultimately delaying the healing process and compromising the regenerative potential, the newer generation of platelet concentrates called platelet-rich fibrin has eliminated the use of anticoagulants showing superior wound healing when compared with PRP, as demonstrated in numerous studies [10]. Injectable platelet-rich fibrin (i-PRF) was discovered by using the low-speed centrifugation concept, in which the quality, applicability, and effectiveness of platelet-rich fibrin were enhanced by reducing centrifugation speed [11].

The platelet-rich fibrin in an injectable form is liquid in consistency, rich in fibrinogen and thrombin, that remains in an injectable form for 15 to 20 minutes before converting into a fibrin matrix. It offers more sustained and controlled release of growth factors than PRP, leading to enhanced healing outcomes. It can be used in periodontal pockets and in periodontal regenerative procedures by mixing with bone grafts to create “sticky bone” to improve the stability of the graft. There are many studies showing its better regenerative potential, enhanced cell migration, higher transforming growth factor-β and collagen expression, antimicrobial properties, reducing inflammation, and promoting regeneration of tissues in gingival and periodontal therapy [12].

One-stage full-mouth disinfection (OS-FMD) with chlorhexidine is an established protocol; its outcomes are limited to antimicrobial effects without significant regenerative potential. The addition of i-PRF to the therapy will introduce a biologically active, autologous adjunct capable of delivering growth factors and regenerative cells in a minimally invasive manner. Integration of i-PRF into the OS-FMD protocol for stage II and III periodontitis, our study seeks to address the gap between infection control and tissue regeneration.

Primary Objectives

To evaluate and compare the outcomes of i-PRF as an adjuvant in OS-FMD and OS-FMD therapy alone by assessment of clinical parameters in patients with stage II and stage III periodontitis after 3 and 6 months in terms of reduction in PD and gain in CAL.

Secondary Objectives

To evaluate and compare the outcomes of i-PRF as an adjuvant in OS-FMD and OS-FMD therapy alone by assessment of changes in the scores of plaque index (PI) and papillary bleeding index (PBI) post therapy in patients with stage II and stage III periodontitis after 3 and 6 months.

Methods

Trial Design

This study is designed as a randomized controlled prospective parallel-arm clinical trial to enable a direct comparison between the clinical outcomes of i-PRF as an adjuvant in OS-FMD therapy and OS-FMD therapy alone in patients with stage II and III periodontitis.

Setting of the Study

In this study, approximately 30 healthy patients will be selected with no history of systemic disease, with stage II and III periodontitis. Recruitment of the patients will be conducted

XSL•FO
RenderX

from the Outpatient Department of Periodontics at Sharad Pawar Dental College, Sawangi (Meghe), Wardha. Prior to participation, information about the study design, purpose, and methodology will be provided to all subjects and written informed consent will be obtained. Eligibility of the participants will be based upon specific inclusion criteria.

The elaboration of the protocol is according to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Studies) guidelines and using the SPIRIT checklist and was registered at Clinical Trials Registry India (CTRI/2025/06/088718).

Eligibility Criteria and Recruitment

Following clinical evaluation, 30 systemically healthy patients in total, with stage II and III periodontitis, must meet the criteria precisely listed in Textbox 1.

Textbox 1. Inclusion and exclusion criteria for eligibility of the patients in the study.

Inclusion criteria

- Generalized interproximal attachment loss affecting at least 3 teeth, excluding incisors and first molars
- Periodontal tissue destruction disproportionate to the amount of microbial deposits present
- Patients presenting with at least 20 natural teeth during examination
- Patients exhibiting probing depth and clinical attachment loss ≥4 mm at 2 or more sites of 12 teeth minimum
- The participants in the study should be healthy and free from any systemic condition
- Patients who had not undergone periodontal therapy within the past 6 months and were not undergoing or had not recently received antibiotic therapy
- Patients showing motivation and willingness to undergo therapy and comply with follow-up, respectively

Exclusion criteria

- Patients with known or suspected hypersensitivity or allergy
- Patients who had received the periodontal therapy in the past 6 months
- Patients presenting with any systemic disorder known to influence periodontal health
- Patients presenting with active infectious conditions other than periodontitis
- Patients with a habit of chronic alcoholism
- Patients having a history of tobacco consumption in any form, including smoking
- Immunocompromised individuals and those presenting with systemic health conditions
- Pregnant or lactating women

Outcomes

One week after performing OS-FMD therapy, subgingival delivery of i-PRF will be done in all periodontal pockets in the test group. Both groups will be followed up after 3 months and 6 months for clinical evaluation to assess the outcomes of the therapy. Reduction in PD and gain in CAL will be recorded using the University of North Carolina 15 mm (UNC-15) periodontal probe, PI, according to the Turesky–Gilmore–Glickman modification of the Quigley Hein Index (1970) and PBI according to the index given by Rebelo and Queiroz [13].

Sample Size Calculation

The formula used for sample size calculation will be $n \geq [(Z_{(1-\alpha/2)}+Z_{(1-\beta)})^2 * (\sigma_1^2+\sigma_2^2/r)] / (\mu_1-\mu_2)^2$, with given values: $Z_{(1-\alpha/2)}$ for a 2-tailed 95% confidence level=1.96 and $Z_{(1-\beta)}$ for 99% power=2.33, where mean ($\mu 1$) and SD ($\sigma 1$) in group 1 will be 4.36 and 0.87 respectively, mean ($\mu 2$) and SD ($\sigma 2$) in group 2 will be 3.04 and 0.66 respectively, and ratio (r)=1. The required sample size is 13 participants per group (n=13).

The sample size was determined on the basis of earlier studies involving the evaluation of clinical parameters in OS-FMD therapy. Considering the outcome variable of PD and the mean difference reported in existing literature, a sample size of 26 participants in total (13 participants per group) was considered to be adequate to achieve sufficient statistical power [6].

To uphold the scientific validity of the study, the sample size was increased to 15 participants per group. This decision was made in anticipation of potential dropouts due to unforeseen reasons or noncompliance from the side of patients with follow-up protocols. By adopting this adjusted sample size, the study aims to preserve adequate statistical power, ensuring the reliability and integrity of the final outcomes, even in the event of participant loss (Figure 1).

XSL•FO
RenderX

Figure 1. A simplified outline of the study protocol based on the CONSORT (Consolidated Standards of Reporting Trials) diagram. CAL: clinical attachment loss; i-PRF: injectable platelet-rich fibrin; PBI: papillary bleeding index; PD: probing depth; PI: plaque index.

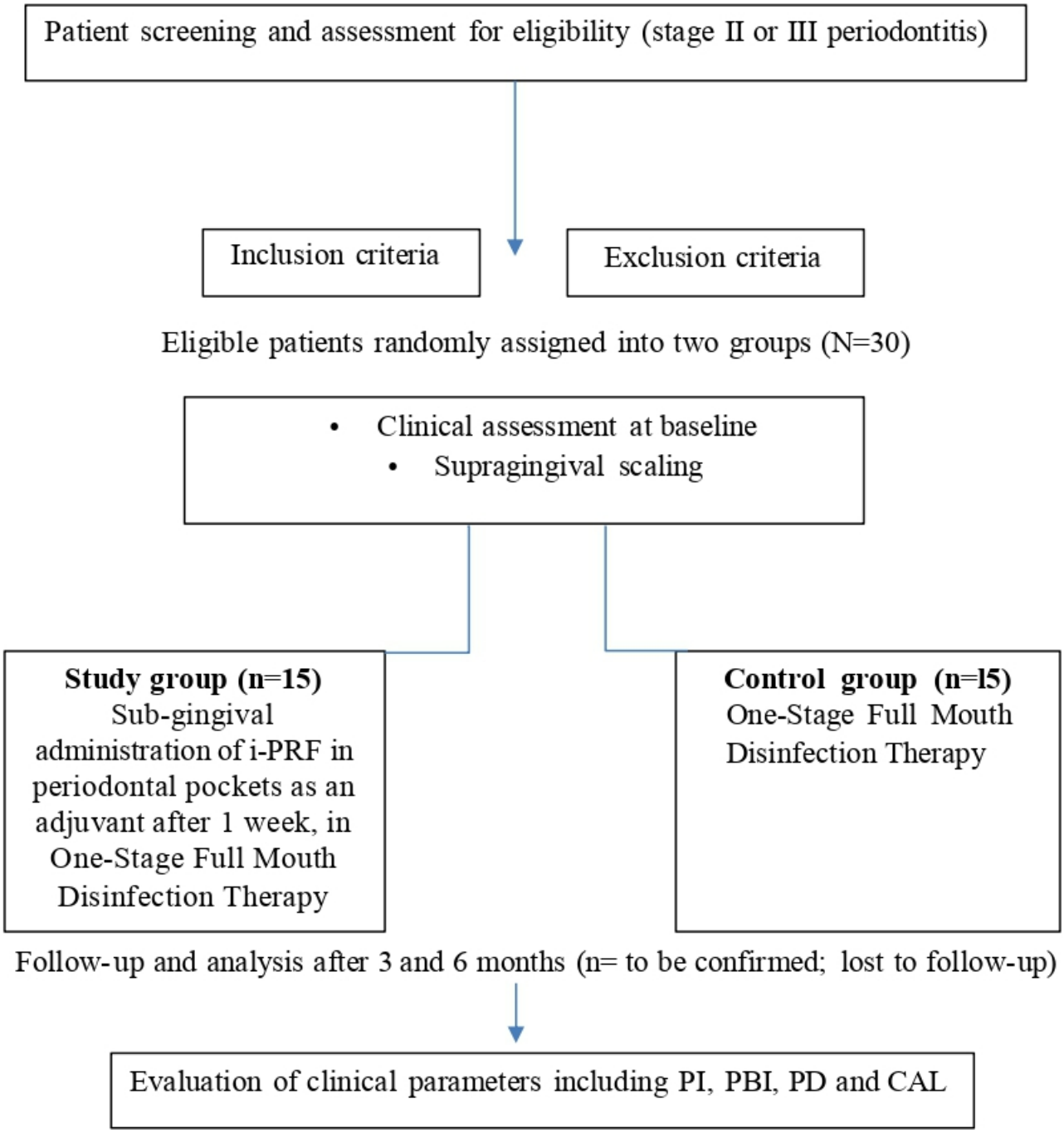

Recruitment

Patients in the Outpatient Department of Periodontics and Implantology, Sharad Pawar Dental College and Hospital, Sawangi (Meghe), Wardha, were recruited.

Randomization and Masking

The eligible patients will be randomly assigned to either test group (i-PRF as an adjuvant in OS-FMD) or the control group (OS-FMD alone) in patients with stage II and III periodontitis. All participants will undergo a periodontal health assessment via clinical examination. Informed consent will be obtained from all participants in written format following a thorough explanation of the nature, purpose, potential benefits, and risks of the intervention. Along with obtaining informed consent, a complete and detailed medical history will also be obtained from the patients after receiving a thorough explanation of the proposed intervention.

Central randomization will be carried out for this study by using a secure computerized system, with assigning a unique study code to each participant, in order to ensure confidentiality. Data presented to the operators and research evaluators will be anonymized and presented using this coded format. During the phase of data analysis, the statistician will remain blinded to group allocation to minimize bias and enhance the objectivity and reliability of the results.

XSL•FO
RenderX

Clinical Procedure

Information concerning dietary habits, systemic health status, oral hygiene practices, and gingival and periodontal conditions, and other routine clinical parameters, will be documented in the specially designed clinical chart, along with obtaining thorough medical and past dental history from the patients. Clinical examination of the patients will be performed under adequate illumination using a mouth mirror and a UNC-15 periodontal probe.

Test Group

In the test group, patients will receive OS-FMD 1 week after the completion of supragingival scaling to allow for initial plaque control and gingival stabilization. Under local anesthesia, full-mouth subgingival SRP will be performed using an ultrasonic scaler and curettes, respectively, within 24 hours. Mechanical debridement will be complemented by the application of chlorhexidine to disinfect the intraoral niches.

The patient will be instructed to brush the dorsum of the tongue with 1% chlorhexidine gel for 1 minute, immediately following each instrumentation session. Tongue cleaning will be followed by rinsing with 0.2% chlorhexidine solution 2 times, each lasting for a minute. The pharyngeal region will then be sprayed with 0.2% chlorhexidine spray, followed by irrigation of all the periodontal pockets with 0.2% chlorhexidine solution thrice, within 10 minutes. Additionally, the prescription to use 0.2% chlorhexidine mouthwash twice daily will be given to the patient for 1 week following FMD [14]. One week postoperatively, i-PRF will be administered subgingivally into all periodontal pockets.

Preparation of I-PRF

The preparation of i-PRF will be conducted in a sterile environment with careful handling to maintain biosafety. Blood will be withdrawn only once per patient in the test group**,** with no repeated phlebotomy, thereby minimizing participant discomfort and procedural risks. All phlebotomy procedures will be performed by trained personnel under strict aseptic conditions using sterile, disposable equipment to minimize infection risk. Participants will be closely monitored during and after venipuncture, and standard protocols for managing potential complications such as hematoma, vasovagal episodes, or allergic reactions will be in place. In case of emergency, immediate medical management will be carried out and referral, if necessary, and the case will be reported to the Data Monitoring Committee (DMC) and the Institutional Ethics Committee (IEC) according to the severity of adverse events (AEs).

10 ml of venous blood will be drawn immediately, just before subgingival administration, through antecubital venipuncture under strict aseptic conditions. i-PRF is a second-generation platelet concentrate, which is a blood-derived autologous biomaterial that possesses a 3D fibrin network, while maintaining its fluid consistency [15]. The obtained blood without the use of anticoagulants is then transferred into a sterile glass test tube. The whole blood will be subjected to low-speed centrifugation at 700 rpm (60×g) at room temperature for 3 minutes using the Dental Laboratory Centrifuge by Labtech Disposables. The low-speed centrifugation technique allows the separation of the components of blood, allowing the formation of the upper layer of liquid, which is rich in fibrinogen and growth factors, collected as i-PRF. In contrast to PRP, which requires multiple centrifugation steps and the addition of anticoagulants, for i-PRF, a single spin is required without the use of anticoagulants. This allows preservation of higher concentrations of growth factors and regenerative cells, enhancing its efficacy in the promotion of tissue regeneration and healing [12].

Control Group

In the control group, patients will receive OS-FMD 1 week after the completion of supragingival scaling. Under local anesthesia, full-mouth subgingival SRP will be performed using an ultrasonic scaler and curettes, respectively, within 24 hours. Mechanical debridement will be complemented by the application of chlorhexidine to disinfect the intra-oral niches.

The patient will be instructed to brush the dorsum of the tongue with 1% chlorhexidine gel for 1 minute, immediately following each instrumentation session. Tongue cleaning will be followed by rinsing with 0.2% chlorhexidine solution 2 times, each lasting for a minute. The pharyngeal region will then be sprayed with 0.2% chlorhexidine spray, followed by irrigation of all the periodontal pockets with 0.2% chlorhexidine solution thrice, within 10 minutes. Additionally, the prescription to use 0.2% chlorhexidine mouthwash twice daily will be given to the patient for 1 week following FMD [14].

Clinical Measurements

Clinical Indices

In order to ensure consistency, reduce interoperator variability, and minimize bias, all clinical recordings will be performed by the same calibrated examiner who is blinded to group allocation. The mean values of these indices will be calculated for outcome assessment.

The status of oral hygiene will be evaluated using the Full Mouth Plaque Index, based on the Turesky-Gilmore-Glickman modification of the Quigley Hein Index (1970) by which assessment of plaque will be carried out on the buccal or labial and lingual surfaces of all teeth after application of a disclosing agent [16].

For the measurement of gingival inflammation, the full-mouth PBI by Rebelo and Queiroz [13] will be recorded. Recordings will be obtained by carefully inserting the periodontal probe into the gingival sulcus at the base of the papilla on the mesial aspect, then gently moving it coronally toward the tip of the papilla. The procedure will be repeated on the distal aspect of the same papilla. The intensity of bleeding provoked by probing will be assessed using a standardized scale ranging from 0 to 4 to evaluate gingival inflammation [16].

Probing Measurements

To evaluate treatment outcomes, clinical parameters will be assessed using the UNC-15 probe (Hu-Friedy) to measure PD and CAL. For recording probing measurements, the deepest

XSL•FO
RenderX

probing site will be evaluated and recorded. Measurements will be taken at baseline, and subsequently at 3 and 6 months.

Statistical Analysis

The values of mean (SD) will be calculated for each of the clinical periodontal parameters, including PI, PBI, PD, and CAL. To assess the significance of observed differences using standard methods, statistical analysis will be performed. To compare baseline values of the findings to those values at 6 months within both groups, the Wilcoxon signed-rank test will be used. Comparisons of PI, PBI, CAL, and PD will be made at baseline, after 3 months of therapy, and then after 6 months. A *P* value less than .05 will be considered statistically significant, while a *P* value greater than .05 will be considered statistically nonsignificant.

Ethical Considerations

Ethical approval for the study is obtained from the IEC (reference number DMIHER(DU)/IEC/2025/751) of Datta Meghe Institute of Higher Education and Research (deemed to be a university). The trial has been registered prospectively with the Clinical Trials Registry of India (CTRI registration no CTRI/2025/06/088718). All the participants will be provided with the sheet consisting of information regarding the purpose, process involved, and specifications of the study. The queries of participants will be addressed, if any. Informed consent in written format will be obtained from the participants before starting this trial. Written informed consent will be obtained from the patients for the publication of their data. The participants are free to withdraw at any time during the study. Clinical assessment will be done by taking aseptic precautions into consideration, without any harm to the patient. Collection of blood sample, transferring it into a glass test tube, and preparation of i-PRF will be done with aseptic precautions and expertise without any harm to the patient. The privacy and confidentiality will be protected by deidentifying the study data. Participants will not receive any compensation for their participation in this study.

Results

The study received approval from the IEC in June 2025. Enrollment of participants began in September 2025 and will last up to March 2026. The study is scheduled to conclude by the end of 2026 after assessments and analyses. The accessibility of the study results is anticipated in early 2027. The results of this study will highlight the therapeutic efficacy of i-PRF as an adjunct to OS-FMD therapy, evaluated on the basis of clinical parameters including PI, PBI, PD, and CAL.

The dissemination plans for the study results are via multiple channels, including research reports, peer-reviewed publications, and presentations at national and international conferences. A primary paper, focusing on the outcomes aligned with the study objectives, will be submitted to a peer-reviewed journal with a high impact factor for publication.

Discussion

Anticipated Findings

Periodontal therapy plays a vital role in the prevention of progression of disease and inflammation of the gingiva and supporting periodontal structures, including the periodontal ligament and alveolar bone. If the condition is left untreated, it can cause progressive attachment loss, resorption of alveolar bone, tooth mobility, and ultimately loss of the tooth. The primary aim of periodontal therapy is to stop the disease from progressing by eliminating microbial load in the form of biofilm or calculus, controlling inflammation, and restoring periodontal health through nonsurgical therapy initially and surgical interventions further, if required. OS-FMD provides a comprehensive approach by treating the entire oral cavity within 24 hours, targeting not only periodontal pockets but also other oral niches in order to prevent cross-contamination and reinfection. The OS-FMD results in a reduction of the overall treatment time and microbial load more efficiently than conventional quadrant-wise therapy. The therapy leads to enhancement of patient compliance, limits the recolonization of pathogens between treated and untreated sites, and serves as an economical option. Over 2 decades of research have shown that FMD can achieve clinical outcomes comparable to or better than conventional approaches, especially when tailored with adjunctive therapies. Chlorhexidine was most effective as an adjuvant, especially when combined with FMD [5].

Bollen et al [17] in their pilot study compared standard periodontal treatment comprising multiple sessions of SRP with FMD in patients with chronic periodontitis. Intensive scaling, subgingival irrigation, and tongue brushing using 1% chlorhexidine gel and 0.2% chlorhexidine rinses within 24 hours were carried out in the test group and simply SRP in the control group, followed by oral hygiene instructions to both groups. Plaque samples showed a greater reduction of harmful bacteria in the test group, especially around multi-rooted teeth, lasting up to 8 months. Anaerobic microbial flora remained lower in the test group than in the control. A significant increase in beneficial microflora was observed in the test group up to 4 months. FMD showed superior short-term microbiological outcomes.

Stein et al [18] in their randomized clinical trial compared four treatment groups: quadrant scaling and root planing (Q-SRP), full-mouth scaling, FMD, and FMD with erythritol air-polishing (FMDAP). All methods resulted in a reduction in PD and gain in CAL, along with improvement in the values of Balance of Payments over the period of 6 months. The FMDAP group showed the greatest reduction in PD for both moderate and deep pockets compared to Q-SRP. FMD also performed better than Q-SRP in deep pockets, but not as well as FMDAP. Conclusively, all the full-mouth approaches were more time-efficient than Q-SRP, with FMDAP being most effective in terms of pocket closure, especially in single-rooted teeth.

Platelet concentrates have gained popularity in the field of periodontal regeneration due to their wide range of use. Along with having the advantage of being autologous, it resulted in enhanced soft and hard tissue healing and amplified regenerative

XSL•FO
RenderX

potential with uncomplicated preparation techniques. Differences in the preparation protocol have led to variable results in the literature [19].

i-PRF is a second-generation platelet concentrate that has shown significant regenerative potential in the field of dentistry and medicine. i-PRF differs from traditional PRP in that it is prepared without the use of anticoagulants, enabling a sustained and gradual release of growth factors, such as platelet-derived growth factor, transforming growth factor-β, and vascular endothelial growth factor. This promotes enhanced cell proliferation, migration, angiogenesis, and tissue regeneration, leading to overall enhancement in clinical parameters [10].

Thamaraiselvan et al [20] in their randomized parallel arm clinical trial evaluated the clinical parameters, such as PI, PBI, PD, and CAL, along with microbiological assessment at baseline, 6 weeks, and 12 weeks in patients with periodontitis randomly allocated into 3 groups with group A receiving ciprofloxacin-loaded i-PRF after SRP, group B receiving local drug delivery with i-PRF, and group C receiving SRP alone. The results demonstrated greater improvements in clinical parameters and a reduction in the microbial load as compared to other groups. The study concluded that ciprofloxacin-loaded i-PRF is a promising, autologous, and biocompatible vehicle for local drug delivery, combining antimicrobial efficacy of ciprofloxacin with regenerative potential, favorable tissue integration capability of i-PRF and its liquid injectable form offering the additional advantage of being flowable at placement before slowly polymerizing to gel in situ, thus adapting well to periodontal pocket. anatomy and minimizing the chances of displacement, thereby enhancing periodontal healing outcomes beyond SRP alone.

Miron et al [12] stated the natural clotting tendency of i-PRF post injection and effective delivery in periodontal pockets due to its injectable form, where it can demonstrate superior outcomes in terms of reduction in pocket depth, periodontal regeneration, gain in CAL, wound healing, and gingival tissue thickening. It also exhibits antimicrobial and anti-inflammatory properties, modulating polarization of macrophages toward a pro-healing M2 phenotype.

In this study, we hypothesize that the incorporation of i-PRF as an adjunct to OS-FMD will result in superior improvements in clinical outcomes such as clinical attachment gain and PD reduction, along with decreased PI and PBI scores due to its regenerative and anti-inflammatory properties along with enhanced healing potential, beyond the antimicrobial benefits achieved with chlorhexidine in the standard OS-FMD protocol. In addition, the preparation of i-PRF is a cost-effective, clinically feasible, minimally invasive, autologous, chairside procedure that can be realistically integrated into routine clinical practice.

The outcomes of this study will throw light on the therapeutic efficacy of an adjunctive use of i-PRF to OS-FMD therapy. Current protocols available in the literature for OS-FMD in the management of periodontitis primarily recommend SRP with the use of antimicrobials such as chlorhexidine according to the standard protocol given by Quirynen and colleagues [21]. By incorporating i-PRF, our trial will explore a regenerative adjunct that could potentially enhance healing outcomes and reduce disease recurrence. Integrating such biologically based approaches into established frameworks may provide clinicians with a more comprehensive, evidence-supported strategy, bridging antimicrobial therapy with regenerative potential. We believe that by establishing strong clinical evidence first, our study will provide a sound basis for future research to explore microbiological and biochemical pathways in greater detail. We also highlight that while this trial focuses on systemically healthy patients, the protocol may serve as a foundation for future studies in more diverse populations, thereby broadening its clinical relevance. The overall findings of the study will provide a strong groundwork for broader and more mechanistic investigations in the future.

Declarations

Oversight and Monitoring

Composition of the coordinating center and trial steering committee consisting of the following: (1) the head of the department of periodontics, (2) postgraduate guide, (3) research scientist, (4) principal investigator, (5) statistician, and (6) data manager.

Composition of the DMC, its Role, and Reporting Structure

The composition of DMC is as follows: senior faculty members who are not directly involved in this trial and 1 external expert in periodontology.

AE Reporting and Harms

All AEs will be actively monitored and documented throughout the study. Mild, self-limiting events such as transient discomfort or swelling will be managed conservatively and recorded in the case report forms. Any moderate or severe AEs will be immediately reported to the principal investigator and the DMC, with appropriate medical care provided. Serious AEs, if any, will be reported to the IEC within 24 hours, and participants will be withdrawn from the study if deemed necessary for safety.

Frequency and Plans for Auditing Trial Conduct

The trial conducted will be reviewed by the project management group that meets every month. The trial steering group and the independent data monitoring and ethics committee will meet to review and conduct the trial period until the trial is complete.

Dissemination Plans

Dissemination plans include publication, research reports, presentations at national and international conferences, and results will also be shared with professional organizations and guideline committees to support translation into future periodontal practice recommendations.

Conclusions

This study highlights the clinical effectiveness of i-PRF as an adjunct to OS-FMD in the management of periodontitis. By comparing adjunctive use of i-PRF in OS-FMD with OS-FMD therapy alone, the study aims to assess improvements in clinical parameters, including CAL, PD, PI, and PBI. The findings of the study are expected to provide evidence regarding the potential regenerative and anti-inflammatory benefits of i-PRF,

XSL•FO
RenderX

thereby contributing to improved therapeutic strategies and clinical decision-making in periodontal care.

Acknowledgments

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Funding

This study has not received external funding. Institutional resources and infrastructure of the Department of Periodontics and Implantology, Sharad Pawar Dental College, are being used to support the conduct of the trial.

Data Availability

Data and materials will be made available on request to authors.

Authors' Contributions

Conception and design, article writing, data acquisition and collection, analysis and interpretation, final approval of the version to be submitted: MK
Supervision and mentoring, conception and design, analysis and interpretation, final approval of the version to be submitted: PB
Data acquisition and collection, article writing, analysis and interpretation: SD
Article writing, data acquisition and collection, final approval of the version to be submitted: ST

Conflicts of Interest

None declared.

Checklist 1
SPIRIT Checklist.
[PDF File, 477 KB - resprot_v15i1e81032_app1.pdf]

References

1. Papapanou PN, Sanz M, Buduneli N, et al. Periodontitis: consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri-Implant Diseases and Conditions. J Periodontol 2018 Jun;89 Suppl 1(S1):S173-S182. [doi: 10.1002/JPER.17-0721] [Medline: 29926951]
2. Cugini MA, Haffajee AD, Smith C, Kent RL, Socransky SS. The effect of scaling and root planing on the clinical and microbiological parameters of periodontal diseases: 12-month results. J Clin Periodontol 2000 Jan;27(1):30-36. [doi: 10.1034/j.1600-051x.2000.027001030.x] [Medline: 10674959]
3. O'Leary TJ. The impact of research on scaling and root planing. J Periodontol 1986 Feb;57(2):69-75. [doi: 10.1902/jop.1986.57.2.69] [Medline: 3514839]
4. Westfelt E, Rylander H, Dahlén G, Lindhe J. The effect of supragingival plaque control on the progression of advanced periodontal disease. J Clin Periodontol 1998 Jul;25(7):536-541. [doi: 10.1111/j.1600-051x.1998.tb02484.x] [Medline: 9696252]
5. Pockpa AD, Soueidan A, Louis P, Coulibaly NT, Badran Z, Struillou X. Twenty years of full-mouth disinfection: the past, the present and the future. Open Dent J 2018;12:435-442. [doi: 10.2174/1874210601812010435] [Medline: 29988213]
6. Sethiya KR, Dhadse P, Bajaj P, Chimote M, Subhadarsanee C, Hassan S. Comparative evaluation of the effectiveness of sodium hypochlorite and chlorhexidine in one-stage full-mouth disinfection treatment of patients with generalized aggressive periodontitis. Journal of Datta Meghe Institute of Medical Sciences University 2021;16(4):728-733. [doi: 10.4103/jdmimsu.jdmimsu_385_21]
7. Quirynen M, Mongardini C, de Soete M, et al. The rôle of chlorhexidine in the one-stage full-mouth disinfection treatment of patients with advanced adult periodontitis. Long-term clinical and microbiological observations. J Clin Periodontol 2000 Aug;27(8):578-589. [doi: 10.1034/j.1600-051x.2000.027008578.x] [Medline: 10959784]
8. Bollen CM, Mongardini C, Papaioannou W, Van Steenberghe D, Quirynen M. The effect of a one-stage full-mouth disinfection on different intra-oral niches. Clinical and microbiological observations. J Clin Periodontol 1998 Jan;25(1):56-66. [doi: 10.1111/j.1600-051x.1998.tb02364.x] [Medline: 9477021]
9. Vandekerckhove BN, Bollen CM, Dekeyser C, Darius P, Quirynen M. Full- versus partial-mouth disinfection in the treatment of periodontal infections. Long-term clinical observations of a pilot study. J Periodontol 1996 Dec;67(12):1251-1259. [doi: 10.1902/jop.1996.67.12.1251] [Medline: 8997670]

XSL•FO
RenderX

10. Miron RJ, Gruber R, Farshidfar N, Sculean A, Zhang Y. Ten years of injectable platelet-rich fibrin. Periodontol 2000 2024 Feb;94(1):92-113. [doi: 10.1111/prd.12538] [Medline: 38037213]
11. Choukroun J, Ghanaati S. Reduction of relative centrifugation force within injectable platelet-rich-fibrin (PRF) concentrates advances patients' own inflammatory cells, platelets and growth factors: the first introduction to the low speed centrifugation concept. Eur J Trauma Emerg Surg 2018 Feb;44(1):87-95. [doi: 10.1007/s00068-017-0767-9] [Medline: 28283682]
12. Miron RJ, Fujioka-Kobayashi M, Hernandez M, et al. Injectable platelet rich fibrin (i-PRF): opportunities in regenerative dentistry? Clin Oral Investig 2017 Nov;21(8):2619-2627. [doi: 10.1007/s00784-017-2063-9] [Medline: 28154995]
13. Rebelo MA, Queiroz AC. Gingival indices: state of art. In: Panagakos FS, Davies RM, editors. Gingival Diseases-Their Aetiology, Prevention and Treatment: IntechOpen Limited; 2011:41-54 URL: https://www.intechopen.com/books/352 [accessed 2026-02-11]
14. Rathod A, Jaiswal P, Kale B, Masurkar D. Comparative evaluation of the effectiveness of triphala and chlorhexidine in full-mouth disinfection treatment of periodontitis in type 2 diabetes patients. J Contemp Dent Pract 2023 Oct 1;24(10):798-801. [doi: 10.5005/jp-journals-10024-3573] [Medline: 38152913]
15. Shashank B, Bhushan M. Injectable platelet-rich fibrin (PRF): the newest biomaterial and its use in various dermatological conditions in our practice: a case series. J Cosmet Dermatol 2021 May;20(5):1421-1426. [doi: 10.1111/jocd.13742] [Medline: 32996229]
16. Fischman SL. Clinical index systems used to assess the efficacy of mouthrinses on plaque and gingivitis. J Clin Periodontol 1988 Sep;15(8):506-510. [doi: 10.1111/j.1600-051x.1988.tb01022.x] [Medline: 3053792]
17. Bollen CM, Vandekerckhove BN, Papaioannou W, Van Eldere J, Quirynen M. Full- versus partial-mouth disinfection in the treatment of periodontal infections. A pilot study: long-term microbiological observations. J Clin Periodontol 1996 Oct;23(10):960-970. [doi: 10.1111/j.1600-051x.1996.tb00519.x] [Medline: 8915027]
18. Stein JM, Yekta-Michael SS, Schittenhelm F, et al. Comparison of three full-mouth concepts for the non-surgical treatment of stage III and IV periodontitis: a randomized controlled trial. J Clin Periodontol 2021 Dec;48(12):1516-1527. [doi: 10.1111/jcpe.13548] [Medline: 34517434]
19. Agrawal AA. Evolution, current status and advances in application of platelet concentrate in periodontics and implantology. World J Clin Cases 2017 May 16;5(5):159-171. [doi: 10.12998/wjcc.v5.i5.159] [Medline: 28560233]
20. Thamaraiselvan M, Jayakumar ND. Efficacy of injectable platelet-rich fibrin (i-PRF) as a novel vehicle for local drug delivery in non-surgical periodontal pocket therapy: a randomized controlled clinical trial. J Adv Periodontol Implant Dent 2024;16(2):94-102. [doi: 10.34172/japid.2024.021] [Medline: 39758269]
21. Quirynen M, Bollen CM, Vandekerckhove BN, Dekeyser C, Papaioannou W, Eyssen H. Full- vs. partial-mouth disinfection in the treatment of periodontal infections: short-term clinical and microbiological observations. J Dent Res 1995 Aug;74(8):1459-1467. [doi: 10.1177/00220345950740080501] [Medline: 7560400]

Abbreviations

AE: adverse event
CAL: clinical attachment loss
DMC: Data Monitoring Committee
FMD: full mouth disinfection
FMDAP: full-mouth disinfection with erythritol air-polishing
i-PRF: injectable platelet-rich fibrin
IEC: Institutional Ethics Committee
OS-FMD: one-stage full-mouth disinfection
PBI: papillary bleeding index
PD: probing depth
PI: plaque index
PRP: platelet-rich plasma
SPIRIT: Standard Protocol Items: Recommendations for Interventional Studies
SRP: scaling and root planing
UNC-15: University of North Carolina 15 mm periodontal probe

XSL•FO
RenderX

Edited by J Sarvestan; submitted 21.Jul.2025; peer-reviewed by MS Mahmood; revised version received 26.Sep.2025; accepted 29.Sep.2025; published 19.Feb.2026.

Please cite as:
Kothekar M, Bajaj P, Dare S, Thakre S
Evaluation of the Effectiveness of Injectable Platelet-Rich Fibrin as an Adjuvant to One-Stage Full-Mouth Disinfection in Patients with Stage II and III Periodontitis: Protocol for a Randomized Clinical Trial
JMIR Res Protoc 2026;15:e81032
URL: https://www.researchprotocols.org/2026/1/e81032
doi:10.2196/81032

© Mahima Kothekar, Pavan Bajaj, Sneha Dare, Shivani Thakre. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Immediate and Sustained Effects of Intensive Equine-Assisted Physiotherapy Based on Neuroproprioceptive "Facilitation and Inhibition" on Psychomotor Development, Clinical Functions, Quality of Life, and Molecular Biological Indicators in Children With Spinal Muscular Atrophy: Protocol for a Crossover Randomized Controlled Trial

Katerina Marikova[1], MSc; Jindra Reissigova[2], PhD; Miloslav Vilimek[3], PhD; Marie Cerna[4], PhD; Marketa Pokorna[4], PhD; Kamila Rasova[1], PhD

[1]Department of Rehabilitation, Third Faculty of Medicine, Charles University and University Hospital Kralovske Vinohrady, Ruská 87, Prague, Czech Republic

[2]The Czech Academy of Sciences, Institute of Computer Science, Prague, Czech Republic

[3]College of Polytechnics Jihlava, Jihlava, Czech Republic

[4]Department of Medical Genetics, Third Faculty of Medicine, Charles University, Prague, Czech Republic

Corresponding Author:
Kamila Rasova, PhD
Department of Rehabilitation, Third Faculty of Medicine, Charles University and University Hospital Kralovske Vinohrady, Ruská 87, Prague, Czech Republic

Abstract

Background: Spinal muscular atrophy (SMA) is a rare neuromuscular disease and the most common genetic cause of infant death. Although pharmacological treatment improves survival rates and functional capacity, physiotherapy remains a key component of care. A newly developed innovative equine-assisted physiotherapy method based on neuroproprioceptive "facilitation and inhibition" principles (NEUROEQUIP-SMA) is hypothesized to improve the quality and extent of motor development in children with SMA.

Objective: The aim of this study is to assess the efficacy of NEUROEQUIP-SMA compared with standard individual outpatient physiotherapy based on the same neuroproprioceptive "facilitation and inhibition" principles (SMA-SOC-N) and to evaluate its effects on functional outcomes and quality of life (QoL). In addition, the response of molecular biomarkers to treatment will be assessed.

Methods: In this crossover randomized controlled trial, 20 children with SMA types I-III, aged 2 - 9 years, will participate in two 6-day therapy programs (A and B) of equal duration (50 min per day) and separated by a 10-week washout period. Each child will be randomly assigned to receive the therapies in either the AB or BA sequence. Therapy A consists of a newly developed NEUROEQUIP-SMA (15 min twice daily), whereas therapy B involves SMA-SOC-N (30 min once daily). Both programs include therapeutic horse grooming (20 min a day). The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND) was selected as the primary outcome measure. Secondary outcomes included motor coordination assessed through 3D motion analysis, muscle fatigue, spirometry, and standardized clinical tests and rating scales, as well as monitoring of psychomotor development (via home-video analysis) and QoL (via questionnaires). Molecular biomarkers will be analyzed from blood samples. The immediate effect of the intervention will be evaluated for most outcomes, while psychomotor development and QoL will be monitored 28 days after therapy as longer-term outcomes. Treatment effect sizes will be reported alongside *P* values to illustrate the magnitude of changes in the outcomes. The study was approved by the Ethics Committee of the Third Faculty of Medicine, Charles University, under the number UK3LF/658559/12025.

Results: The study is designed for 20 participants. Data collection will begin in February 2026 and will be completed in May 2026. Data analyses are planned for autumn 2026, and study results are expected to be available in 2027. A paired *t* test comparing the primary outcome (CHOP INTEND) between treatments NEUROEQUIP-SMA and SMA-SOC-N in 20 children will have 80% power to detect moderate-to-large effect size (Cohen *d*=0.66) at a 5% significance level.

XSL•FO
RenderX

Conclusions: This trial will be the first study to evaluate the effects of NEUROEQUIP-SMA in children with SMA. If preliminary findings confirm a benefit, this physiotherapy approach may represent a promising adjunct to care for the generation of children undergoing gene therapy.

Trial Registration: ClinicalTrials.gov NCT07336602; https://clinicaltrials.gov/study/NCT07336602

International Registered Report Identifier (IRRID): PRR1-10.2196/83266

(JMIR Res Protoc 2026;15:e83266) doi:10.2196/83266

KEYWORDS

spinal muscular atrophy; genetic therapy; equine-assisted therapy; child; pediatrics; physical therapy modalities; long noncoding RNA; neuroproprioceptive facilitation; neuroproprioceptive inhibition; The Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders

Introduction

Background

Spinal muscular atrophy (SMA) is an autosomal recessive neuromuscular disorder caused by the loss or mutation of the *SMN1* gene, with retention of its paralog *SMN2*. Despite the latest treatment options, it is the leading genetic cause of infant mortality under 2 years of age and affects approximately 1 in 10,000 live births in the Czech Republic [1].

The primary characteristic of SMA is the degeneration of alpha motor neurons in the anterior horn of the spinal cord [2]. This leads to a wide range of clinical manifestations, from severe neonatal weakness and respiratory failure to mild proximal weakness presenting in adulthood [3,4]. Although new pharmacological advancements are altering SMA phenotypes and subsequently the clinical manifestation of the disease [5,6], multidisciplinary supportive care remains essential to achieve optimal outcomes [2,7], with physiotherapy being an integral part of it.

The perspective on physiotherapy has been rapidly evolving in recent years. In the past, it was prohibited, as it was assumed that it might accelerate the degeneration of diseased muscles [8,9]. This noninterventionist approach proved to be ineffective, as it deepened muscle weakness and fatigue, leading to mobility impairments [2,10]. Over time, interventions such as positioning, breathing exercises, range-of-motion activities, and targeted strengthening began to be used to prevent complications (eg, contractures and deformities). Nowadays, the focus has shifted more toward enhancing functional mobility, using techniques such as balance training, developmental activities, and gait rehabilitation [11-14].

Experimental research [15] showed a close link between neuron-specific protection and their activation levels by different exercise parameters. Exercise-induced neuroprotection was independent of SMN protein expression and was associated with specific metabolic and behavioral adaptations, including a reduction in muscle fatigability. This new insight into the motor units' adaptations significantly influenced and motivated us to develop a new active physiotherapy approach for children with SMA. We decided to use the principles of neuroproprioceptive "facilitation and inhibition" in therapy, in which it is assumed that, through the combination of appropriate afferent stimuli, the interneuronal system is activated, allowing the transmission of missing signals and enabling the activation of movements that the patient would otherwise be unable to consciously perform. This approach is used in the Czech Republic for neurologically impaired individuals, such as poststroke individuals [16] or people with multiple sclerosis [17]. However, it has not yet been studied in children with SMA.

We have developed a unique method of equine-assisted physiotherapy based on neuroproprioceptive "facilitation and inhibition" (NEUROEQUIP-SMA). Standard Equine-Assisted Physiotherapy has the potential to influence multiple domains, including posture, balance, and coordination, through the dynamic and repetitive movement provided by the horse [18-20]. It has been demonstrated that, in children with cerebral palsy, it shows benefits such as improved postural control and reduced spasticity [18,21,22]. In children with SMA [23], it leads to physical improvements (enhanced muscle strength and better balance) and psychological improvements (self-confidence, self-esteem, pride, independence, and a sense of achievement).

This study primarily aims to assess the effectiveness of 2 physiotherapy approaches: NEUROEQUIP-SMA and a standard physiotherapy protocol based on neuroproprioceptive "facilitation and inhibition" principles (SMA-SOC-N). We hypothesize that NEUROEQUIP-SMA will produce greater improvements in functional and postural outcomes in children with SMA due to:

1. The rider moving in harmony with the horse's motion rather than resisting it as in conventional stabilization techniques.
2. Simultaneous activation of multiple muscle chains, enabling coordinated engagement of weakened muscles.
3. Afferent input combination that initiates movement from the pelvis, lowers the center of gravity, supports the diaphragm in an expiratory position, relaxes the shoulder girdle, and improves overall stability.
4. Increased sensory input intensity and greater movement complexity, further promoting neuromuscular activation.

Primary Hypothesis

Immediately after the 6-day intensive program, we expect improvement in both groups in the Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders (CHOP INTEND); however, we anticipate more pronounced effects in the NEUROEQUIP-SMA group.

XSL•FO
RenderX

Secondary Hypotheses

We anticipate improvements in both groups; however, more pronounced effects are expected in the NEUROEQUIP-SMA group. Specifically, we hypothesize:

1. Reduced undesirable muscle fatigue, assessed using surface electromyography (sEMG) immediately after completion of the program.
2. Improved coordination of the torso and cervical spine, enhanced weight-shift control, and optimized diaphragmatic breathing in the abdominal region immediately after completion of the program.
3. Stabilized or improved motor function, as measured by the Motor Function Measure–20 (MFM-20), immediately after completion of the program.
4. Improved respiratory function and spinal alignment, reflected in spirometry and scoliometric outcomes immediately after completion of the program.
5. Improved trunk control and selective motor control, as demonstrated by performance in the Segmental Assessment of Trunk Control (SATCo), Trunk Control Measurement Scale (TCMS), Modified Pediatric Reach Test, and Selective Control Assessment of the Lower Extremity (SCALE) immediately after completion of the program.
6. Enhanced quality of life assessed using the International Classification of Functioning, Disability and Health (ICF)-based Documentation Form for the Brief Category of Children with Cerebral Palsy, the Strengths and Difficulties Questionnaire, and the Pediatric Quality of Life Inventory 4.0 Generic Core Scales, 28 days after completion of the program.
7. Improved quality and range of psychomotor development, evaluated through home video recordings conducted 28 days after completion of the program.

The Monitoring of Neurophysiological Mechanisms of Therapy Effects

As a complementary and innovative approach to document the benefits of equine-related activities, monitoring will be complemented by analysis of long noncoding RNAs (lncRNAs) expression in peripheral blood. These 200+ nucleotide long transcripts (eg, SMN-AS1) are involved in the regulation of gene expression, for example, through association with polycomb repressive complex 2 [24]. Their deregulated levels are associated with many pathological conditions including neurodegenerative diseases. Changes in the expression of selected lncRNAs in body fluids may also reflect changes in the body in response to therapeutic intervention [25]. We propose a set of selected lncRNAs as a biomarker of improvement status in SMA patients after the inclusion of intensive therapy. Based on previous studies, we selected the lncRNAs (MALAT1 and PARTICLE), which interact with PCR2 [26]. This makes them similar to SMN-AS1, which inhibits SMN2 expression by binding PCR2 [24]. MEG3, MALAT1, and NEAT1 are lncRNAs involved in motor neuron development [27], which are affected in this disease. Deregulated levels of the lncRNAs, GAS5, and H19 have been detected in the blood of patients with neurodegenerative disease [25]. Analysis of selected lncRNAs in the blood of SMA patients before and after intensive therapy could provide novel diagnostic and prognostic biomarkers and therapeutic targets and allow for an understanding of the neurophysiological effects of the treatment. It is anticipated that measurable alterations in lncRNA expression will be detectable immediately following completion of the 6-day therapeutic program.

The aim of this protocol is to present the design and procedures of the randomized crossover study and to specify how the effectiveness of a unique method of equine-assisted physiotherapy based on neuroproprioceptive “facilitation and inhibition” will be tested against standard physiotherapy based on the same principles in children with SMA.

Methods

Study Design

This study is designed as a single-blind randomized crossover-controlled trial. The protocol was substantially revised after the first peer review round, and the trial was newly registered on ClinicalTrials.gov on November 11, 2025 (Unique Protocol ID: 4424; NCT07336602). Participants will be children with SMA types I, II, and III, aged 2 to 9 years. Every child will undergo two 6-day therapeutic programs, administered in 2 separate periods. During each period, children receive 30 minutes of individual treatment and 20 minutes of therapeutic horse grooming per day. In total, 2 different concepts (A and B) will be used for the individual therapy. Therapy A is a newly developed NEUROEQUIP-SMA, while therapy B, the active comparator, is SMA-SOC-N (Table 1). Participants will be randomized in a 1:1 ratio to receive both therapies in either the AB or BA sequence, over two periods separated by a 10-week washout period.

XSL•FO
RenderX

Table . AB and BA sequence comparison.

Sequence	Sequence AB	Sequence BA
Frequency of treatment	• 50 minutes per day	• 50 minutes per day
Content of physiotherapy	• Period 1: equine-assisted physiotherapy based on neuroproprioceptive "facilitation and inhibition" • Period 2: standard individual outpatient physiotherapy	• Period 1: standard individual outpatient physiotherapy • Period 2: equine-assisted physiotherapy based on neuroproprioceptive "facilitation and inhibition"
Length of therapy	• Period 1: 15 minutes twice a day • Period 2: 30 minutes per day	• Period 1: 30 minutes per day • Period 2: 15 minutes twice a day
Length of therapeutic grooming	• 20 minutes per day	• 20 minutes per day
Environment	• Period 2: treatment room • Period 1: riding hall or outdoor terrain	• Period 1: treatment room • Period 2: riding hall or outdoor terrain
Involvement of specialists	• Period 1: collaboration of physiotherapist with a team including a horse trainer and leader • Period 2: physiotherapist only	• Period 1: physiotherapist only • Period 2: collaboration of physiotherapist with a team including a horse trainer and leader

The 6-day duration of each period was chosen to capture an immediate (short-term) response while avoiding undue fatigue in the children. In addition, the short, intensive therapeutic window allows for standardized application and limits the influence of concurrent rehabilitation. The long-term effect is captured 28 days after posttherapy as evidence of maintenance of effect and molecular adaptation.

The immediate effect of the 6-day intervention will be evaluated before and at the end of the program, focusing on functional performance (CHOP INTEND and MFM-20), muscle fatigue (sEMG), standardized postural and trunk control tests (TCMS, Modified Pediatric Reach Test, SCALE, and SATCo), respiratory function (spirometry), scoliotic progression (scoliometric measurements), and molecular biological changes (lncRNAs). 3D motion analysis will be performed using a 16-camera Qualisys Motion Capture System.

The long-term effect of the intervention (28 d after program completion) will be assessed to ensure clinical relevance and sustained benefit, focusing on quality of life (QoL) (standardized questionnaires), and psychomotor development (analyzed via home-video recordings).

The washout interval is a period during which participants do not receive any study-controlled therapy, only their usual care or ongoing therapeutic procedures independent of this study. The washout interval between Period 1 and Period 2 was set at 10 weeks. This duration was determined based on clinical experience and parental feedback, which suggested that short-term benefits of intensive physiotherapy can last approximately 6 weeks. The extended 4-week period beyond that allows sufficient time for residual therapeutic effects to dissipate while maintaining feasibility for participating families. After the completion of Period 2, another washout interval will follow (Figure 1 and Table 2).

XSL•FO
RenderX

Figure 1. Study design.

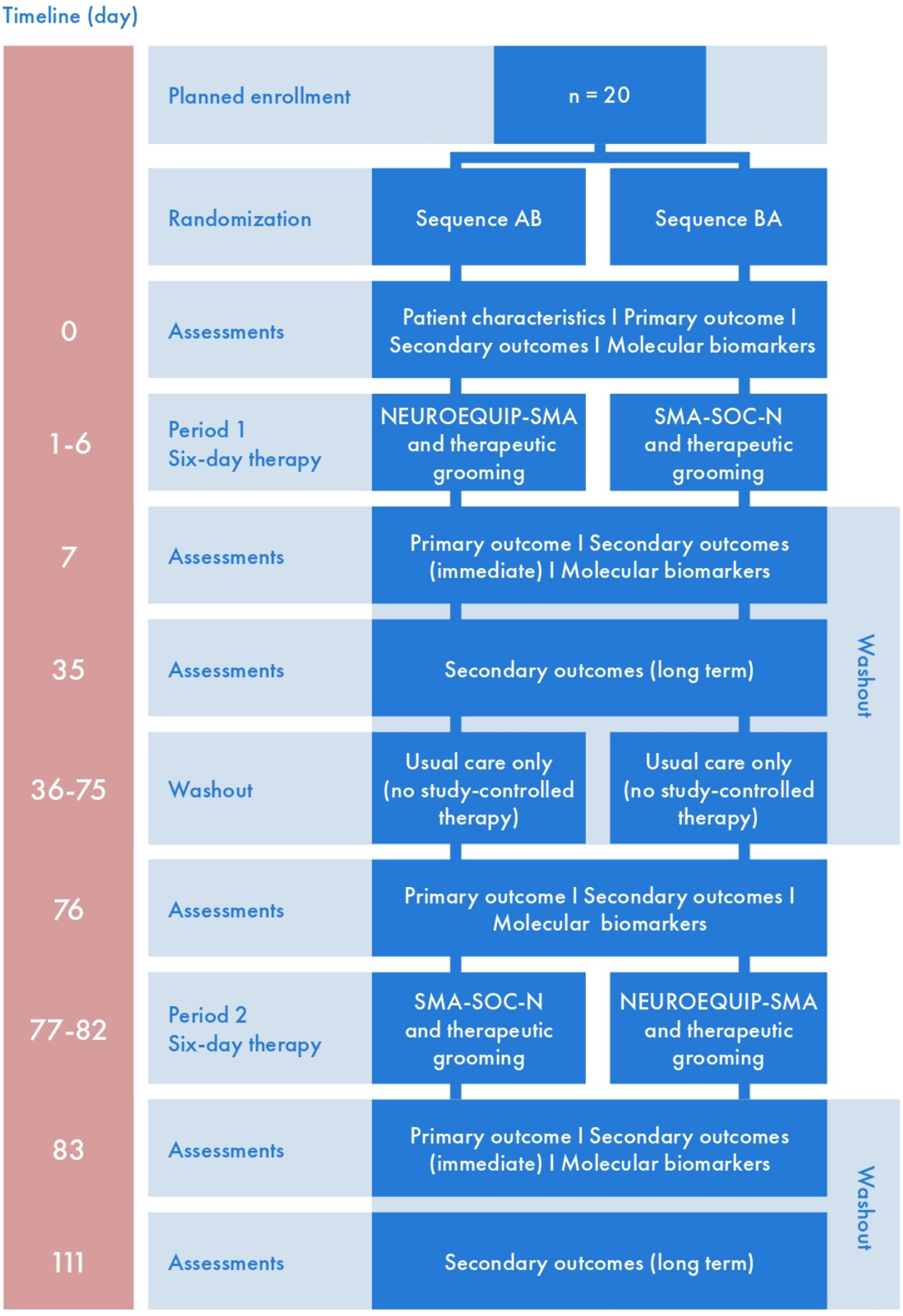

XSL•FO
RenderX

Table . Schedule of assessments and interventions.

Timeline (day)		0	1 - 6	7	8 - 34	35	36 - 75	76	77 - 82	83	84 - 110	111
	Before therapy	Pretherapy	Six-day therapy	Postthera-py		28 days postthera-py		Pretherapy	Six-day therapy	Postthera-py		28 days postthera-py
Enrollment	✓											
Eligibility screen	✓											
Informed consent	✓											
Allocation	✓											
Therapy												
Sequence AB			✓						✓			
Sequence BA			✓						✓			
Examination												
Patient characteristics		✓										
Primary outcome: CHOP INTEND[a]		✓		✓				✓		✓		
Secondary outcomes (immediate): MFM-20[b], sEMG[c], 3D-M[d], SATCo[e], TCMS[f], mPRT[g], SCALE[h], SPM[i], SCM[j]		✓		✓				✓		✓		
Secondary outcomes (long-term): ICF[k], SDQ[l], PedsQL[m], and HVR[n]		✓				✓		✓				✓

XSL•FO
RenderX

Timeline (day)		0	1 - 6	7	8 - 34	35	36 - 75	76	77 - 82	83	84 - 110	111
	Before therapy	Prethera-py	Six-day therapy	Postthera-py		28 days postthera-py		Prethera-py	Six-day therapy	Postthera-py		28 days postthera-py
Molecu-lar biomark-ers		✓		✓				✓		✓		
Washout period				✓	✓	✓	✓	✓		✓	✓	✓

[a]CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders.

[b]MFM-20: Motor Function Measure-20.

[c]sEMG: surface electromyography.

[d]3D-M: 3D motion.

[e]SATCo: Segmental Assessment of Trunk Control.

[f]TCMS: Trunk Control Measurement Scale.

[g]mPRT: Modified Pediatric Reach Test.

[h]SCALE: Selective Control Assessment of the Lower Extremity.

[i]SPM: spirometry measurement.

[j]SCM: scoliometer measurement.

[k]ICF: International Classification of Functioning, Disability, and Health.

[l]SDQ: Strengths and Difficulties Questionnaire.

[m]PedsQL: pediatric quality of life inventory.

[n]HVR: home video recording.

Study Setting

Clinical examinations will be conducted at the College of Polytechnics Jihlava, which is equipped with advanced technology, including a 16-channel surface EMG system (DELSYS), spirometry, and Qualisys Motion Capture System with 16 cameras. The Mirákl Hippotherapy Center, a healthcare facility offering year-round Equine-Assisted Physiotherapy programs for children with special needs, has extensive experience and resources to support the study. LncRNA analysis will be performed at the Department of Medical Genetics, Third Faculty of Medicine, Charles University, which houses a state-of-the-art molecular biology laboratory specializing in genomic and biomarker research. The Institute of Computer Science at the Czech Academy of Sciences will contribute data management and statistical data analysis. Additionally, the Third Faculty of Medicine, Charles University, will provide institutional support for clinical research coordination, ensuring efficient execution and fostering interdisciplinary collaboration.

Research and Management Team

The team includes coordinators, molecular biologists, specialists in 3D motion analysis and sEMG, physiotherapists, including those with expertise in Equine-Assisted Physiotherapy, horse trainers, and statisticians. The study coordinator will oversee recruitment, enrollment, eligibility assessment, and study conduct while providing participants with study details. All personnel involved in examinations will have no role in treatment allocation. A blinded assessor will perform clinical measurements, evaluate video recordings of clinical and functional tests, and collect questionnaire data. Molecular biologists will process peripheral blood samples for lncRNA analysis. Specialists will oversee 3D motion analysis, sEMG, and data processing. The statistician will manage dataset preparation, do statistical hypothesis testing, and interpret their results.

All therapists are trained to deliver interventions according to the protocol and working under the supervision of a senior therapist. Reproducibility is ensured by therapeutic manuals covering NEUROEQUIP-SMA, standard physiotherapy SMA-SOC-N, and therapeutic grooming with pre-defined parameter ranges. Outcome assessors undergo standardized training and calibration for clinical scales and clinical tests, spirometry, and scoliometer and instrumental measurements (eg, 3D motion capture and sEMG). Primary and secondary outcome video assessors are blinded to allocation and have ≥5 years of clinical experience with children with SMA.

Sample Size

The study plans to include 20 children (10 per group), with each child receiving both therapies in the crossover design. When comparing the primary outcome (CHOP INTEND scores) between the NEUROEQUIP-SMA and standard therapies, a 2-sided paired *t* test with 80% power at a 5% significance level can detect an effect size of Cohen *d*=0.66, which is considered a moderate-to-large effect. The calculation of the detectable effect size for n=20, assuming normally distributed within-subject outcome differences, was performed using the pwr.t.test function from the R package *pwr* (version 1.3.0; R Foundation for Statistical Computing) [28].

XSL•FO
RenderX

Recruitment

Participants will be simultaneously recruited from the database of the patient organization SMÁci, z. s., which registers 90% of children with SMA in the Czech Republic, while the remaining participants will be sourced from the database of the Mirákl Hippotherapy Center, o.p.s. All parents of children who meet the entry criteria will be contacted. Children will participate in 2 therapeutic cycles according to randomization, each time in a group of 10 children (the first 10 children enrolled will be assigned to the first therapeutic cycle). As of November 7, 2025, the patient organization has 215 registered members, including 45 children aged 2 - 9.

Eligibility Criteria

Participants will be eligible if they are between 2 and 9 years old and have a diagnosed SMA type I, II, or III with a stable health condition for at least 6 months, no other serious illnesses, and no intensive therapy 10 weeks before study intervention. Exclusion criteria will include hip luxation, allergies to horses or the stable environment, and an overwhelming fear of horses.

Randomization and Blinding

The parents of children participating in the study will be fully informed. Upon giving written informed consent, their children will be randomly assigned to one of the treatment sequences. Block randomization with variable block sizes of 2 and 4 will be used to achieve balanced sequence sizes (AB and BA) within each therapeutic cycle of 20 children, due to organizational considerations. Allocation concealment will be maintained through opaque sealed envelopes. The randomization sequence will be generated using the online software Sealed Envelope [29]. Outcome assessors and investigators will remain blinded to minimize potential bias in outcome assessments. All examinations will take place in the same examination room under identical conditions, regardless of the assigned intervention. However, blinding will not be feasible for therapists, children with SMA, or their parents due to the nature of the interventions. Parents and children will not be told which of the therapies is considered the main intervention and which is the active comparator. To reduce bias, children and parents will be strongly advised not to disclose the allocated therapy during both the baseline and final assessments. While this is the planned procedure, it is acknowledged that practical factors, such as participant availability, medical status, or logistical constraints, may necessitate adjustments to the randomization or blinding processes to ensure the study's feasibility and ethical conduct.

Retention and Withdrawal and Participation Conditions

To support adherence, therapists will use their expertise and provide effective instruction to tailor each therapy session to the individual needs of the child, thereby enhancing adherence. Reasons for participant dropout will include discontinuing therapy for more than 2 missed physiotherapy sessions or significant noncooperation from the child. Participants will have the option to withdraw from the study at any time.

Permitted and Prohibited Concomitant Care

During the 6-day intervention period, participants will not undergo any additional rehabilitation therapies. During the entire washout period, only standard outpatient physiotherapy will be allowed, with a maximum frequency of once per week.

Intervention

The intervention group will undergo equine-assisted physiotherapy based on neuro-proprioceptive "facilitation and inhibition" (NEUROEQUIP-SMA; Figure 2A), while the active comparator group will receive standard individual outpatient physiotherapy based on neuroproprioceptive "facilitation and inhibition" (SMA-SOC-N; Figure 2B). Both groups will also receive therapeutic grooming (Figure 2C).

Figure 3 shows how the intervention group differs from the active comparator group in addressing functional problems identified in both groups.

XSL•FO
RenderX

Figure 2. Photographs illustrating different therapeutic modalities.



XSL•FO
RenderX

Figure 3. Comparison of functional problem management between intervention and active comparator groups. C-spine: cervical spine; UE: upper extremities; LE: lower extremities; SMA: spinal muscular atrophy.

Functional problems in children with SMA	SMA-SOC-N	NEUROEQUIP-SMA
Contractures	Stretching; positioning and releasing the chest; facilitation-inhibition techniques for normalizing tone	Relaxation of the chest with the help of movement in the transverse plane, adjustment of external rotation in the hips with the support of movement in the sagittal plane
Limited lung capacity	Breathing exercises; facilitating exhalation chest position	Differentiation in the transverse plane–supporting exhalation
Hip joint dislocations	Positioning; hip centration; facilitatory-inhibitory inputs for pelvic stability	Differentiation in the sagittal plane–pubic bone area relaxation
Scoliosis	Active exercises with facilitation supporting symmetry of weight transfer and postural correction	Differentiation in the frontal plane–weight transfer symmetry
Weakness of flexors and extensors (C-spine)	Active exercise with facilitated movement guidance	Assisted sitting position–correction of the chest area
Upper and lower limb strength	Active exercise with facilitated movement guidance	Differentiation in the sagittal plane–lowering the center of gravity to the pelvic area
Movement quality	Assisted locomotion/movement with manual facilitation of selective patterns	Activation of the pelvis as a movement initiator
Movement quantity	Exercise on unstable surfaces; facilitating variability	Using the horse's back as an unstable platform
Development of walking	Facilitative inputs to the scapula and pelvis to activate the crossed UE–LE pattern. Proprioceptive stimulation of the sole of the foot (stimulation of support and unwinding)	Support for trunk differentiation with a more pronounced pelvic retroversion setting, a combination of pelvic movement in the transverse and frontal planes

Equine-Assisted Physiotherapy Based on Neuroproprioceptive "Facilitation and Inhibition"

Equine-Assisted Physiotherapy Based on Neuroproprioceptive "Facilitation and Inhibition" (NEUROEQUIP-SMA) differs from conventional hippotherapy in the way the horse's movement stimulus is used and therapeutically processed. While conventional hippotherapy relies mainly on passive adaptation to the rhythmic movement of the horse, NEUROEQUIP-SMA is an individualized therapeutic method that activates proper motor function through the child's consciously controlled response to the horse's back movement. The goal is not to stabilize against the movement, but to accompany it, allowing natural postural reactions, primarily from the pelvic area. The child's movement is secondary, naturally distributed, and rhythmically coordinated throughout the body. This method actively incorporates neuroproprioceptive "facilitation and inhibition," using targeted afferent input, pelvic movement initiation, lowered center of gravity, and coordinated activation of muscle chains, which together enhance postural control and re-activate physiologic motor patterns.

The therapy, moreover, involves careful selection of the horse, therapeutic equipment, appropriate terrain, horse-leading technique, stride length, movement speed, horse head position, child's position, use of positioning aids, and the physiotherapist's manual contacts. The horse is chosen based on specific movement patterns that facilitate the desired motor responses. The pace, stride length, and direction of the horse's movement are individually adapted to the child's needs, with the therapist actively facilitating movement and correcting the child's position using manual contacts or various types of positioning aids (Figure 4). The child is placed in various positions (sitting, lying, and assisted sitting) according to therapeutic goals and current capabilities.

XSL•FO
RenderX

The intervention is performed by a physiotherapist who is specially trained and certified in equine-assisted physiotherapy. Professional competence is obtained by completing a course accredited by the Ministry of Health of the Czech Republic [30,31] and working under professional supervision. The physiotherapist is part of a multidisciplinary team alongside a horse trainer and a horse leader. The horse must be specially trained for therapeutic purposes.

Therapy sessions are conducted either in a therapeutic riding hall or in an outdoor arena with equivalent surface conditions and a calm environment. The intervention is individually tailored based on the clinical assessment of the child, their current motor abilities, and therapeutic goals. Adjustments to meet therapy objectives may include changing the horse, modifying stride length, movement speed, horse-leading method, use of aids, child's position, or the intensity of the therapist's facilitation. Therapeutic tolerance is monitored through control or final examinations and feedback from parents. The success of the intervention is assessed based on the achievement of short-term therapy goals.

Figure 4. Example of a key therapeutic element: pelvic positioning correction. (A) Corrective position with correction by the therapist; the child is seated on the ischial tuberosity with external rotation of the hip joints. (B) The child presents with a pathological posture characterized by an anterior pelvic tilt and internal rotation of the hip joints.



Standard Individual Ambulatory Physiotherapy Based on Neuroproprioceptive "Facilitation and Inhibition" Principles (SMA-SOC-N)

Therapy follows the recommendations of the International Coordinating Committee for SMA Clinical Trials–Standard of Care, so it is individualized to address each child's specific needs, taking into account their motor abilities and disease progression. The primary goals are to maintain or improve muscle strength, prevent contractures and deformities, support postural control, and optimize respiratory function. Therapeutic interventions include regular positioning, passive and active range of motion exercises, facilitation of functional movement patterns, and targeted training aimed at enhancing both stability and mobility [12].

The facilitation of functional movement patterns uses the principles of neuroproprioceptive "facilitation and inhibition," which are commonly used in neurorehabilitation in the Czech Republic. It involves a combination of appropriate afferent stimuli designed to increase the excitability of motoneurons (lowering their excitability threshold) so that impulses from the central nervous system can be effectively transmitted, triggering neuronal discharge and muscle contraction. The goal is to maintain optimal conductivity and function of motor pathways, ensure appropriate excitability of control and regulatory structures, and adjust the activation level at the synapses of motor neurons in the anterior horns of the spinal cord as needed [32]. Their inclusion also allows for a better comparison of the effects between the active comparator and intervention groups, as both approaches use a similar degree of neuroproprioceptive "facilitation and inhibition" within motor patterns.

XSL•FO
RenderX

Therapeutic Grooming

Therapeutic horse care in this protocol refers to a social activation service that takes place as structured, therapist-led work with a horse. Its goal is to support the child's sensory and emotional regulation while safely and comfortably introducing them to the horse and stable environment. The intervention includes repeatable horse care activities (eg, brushing, mane combing, safe cleaning, and short horse leading), which are carried out according to our workplace's internal standardized methodology. These activities allow children to engage naturally with the horse in a safe and supportive environment, promoting both physical and psychological development.

The therapy is guided by a trained therapist, who tailors the intervention to the individual needs of the child. This process involves selecting a suitable horse, adapting the environment, choosing appropriate aids, and applying specific touch techniques, all aimed at maximizing the therapeutic benefits. The goal extends beyond merely brushing the horse's coat to stimulating sensory perception, regulating emotions and muscle tone, and enhancing body awareness.

Sessions take place in a calm and welcoming environment, where external noise and stressors are minimized, facilitating relaxation and allowing the child to focus on the present moment. The effectiveness of the intervention is evaluated through observations of the child's interaction with the horse, emotional state before and after therapy, changes in motor skills, and long-term improvements in self-confidence, communication, and stress management.

Overall, therapeutic grooming offers a unique blend of physical care and emotional well-being. Through their connection with the horse, participants experience deep relaxation, inner balance, and the development of vital life skills.

Therapeutic cleaning is included in both groups to control the effect of interaction with the horse and staying in the stable or riding hall environment.

Data Collection and Management

To support data consistency, instruments will be calibrated regularly. A statistician and coordinator will perform data checks. To ensure data collection, an Excel form will be created, and basic statistical reports will be generated regularly to check the completeness and validity of the data. Personal data will be pseudonymized using unique subject codes and stored on password-protected computers. Only the research team will have access to the identifiable data, and all records will be kept confidential in accordance with GDPR regulations. Anonymized data may be used for analysis and shared only as permitted by the informed consent.

Outcome Measures

Patient characteristics will be collected, including age, gender, information about the disease and social anamnesis, last 5 results from the expanded Hammersmith Functional Motor Scale [33], date of last intensive hippotherapy (with the condition of more than 10 wk), the start and type of medication or rehabilitation, and use of compensatory aids prior to the program.

Primary Outcomes: CHOP INTEND

A clinical assessment tool used to evaluate motor function in infants with neuromuscular disorders measures their ability to perform specific movements and assesses the severity of motor impairments. Measurements are recorded during the test and are later evaluated by 2 independent physiotherapists (PP and MH) [34]. We will expand the standardized assessment with structured qualitative observations (compensatory strategies, movement initiation, body stabilization at the beginning of the movement, movement fluency, respiratory effort, etc). A higher score indicates a larger scope of motor skills.

Secondary Outcomes (Immediate)

3D Motion-Based Evaluation of Postural Control, Breathing, and Coordination

Overview

All 3 tests share a common setup, in which the child sits straddling a cylinder with sensors attached. The cylinder is positioned within a calibrated 3D space and firmly secured to prevent movement. The child sits with feet on the ground and hips flexed to at least 90 degrees. In total, 18 markers are attached bilaterally at anatomically significant points, including the inferior angle of the scapula, sacroiliac joint, distal end of the clavicle, temporomandibular joint, dorsum of the hand, 4th rib (1-quarter proximal to the sternum), 10th rib (at the midpoint), top of the scapula, and hip. Additional markers are placed unilaterally at C7 (the thoracolumbar junction [Th/L]), L4 (the distal end of the sternum), and the umbilicus (Figure 5). Marker motion is captured using 16 Qualisys Motion Capture System cameras, allowing calculation of their spatial coordinates, from which position, velocity, and acceleration data are derived. Each measurement session lasts 30 seconds.

XSL•FO
RenderX

Figure 5. Locations of the 18 markers used for motion analysis.

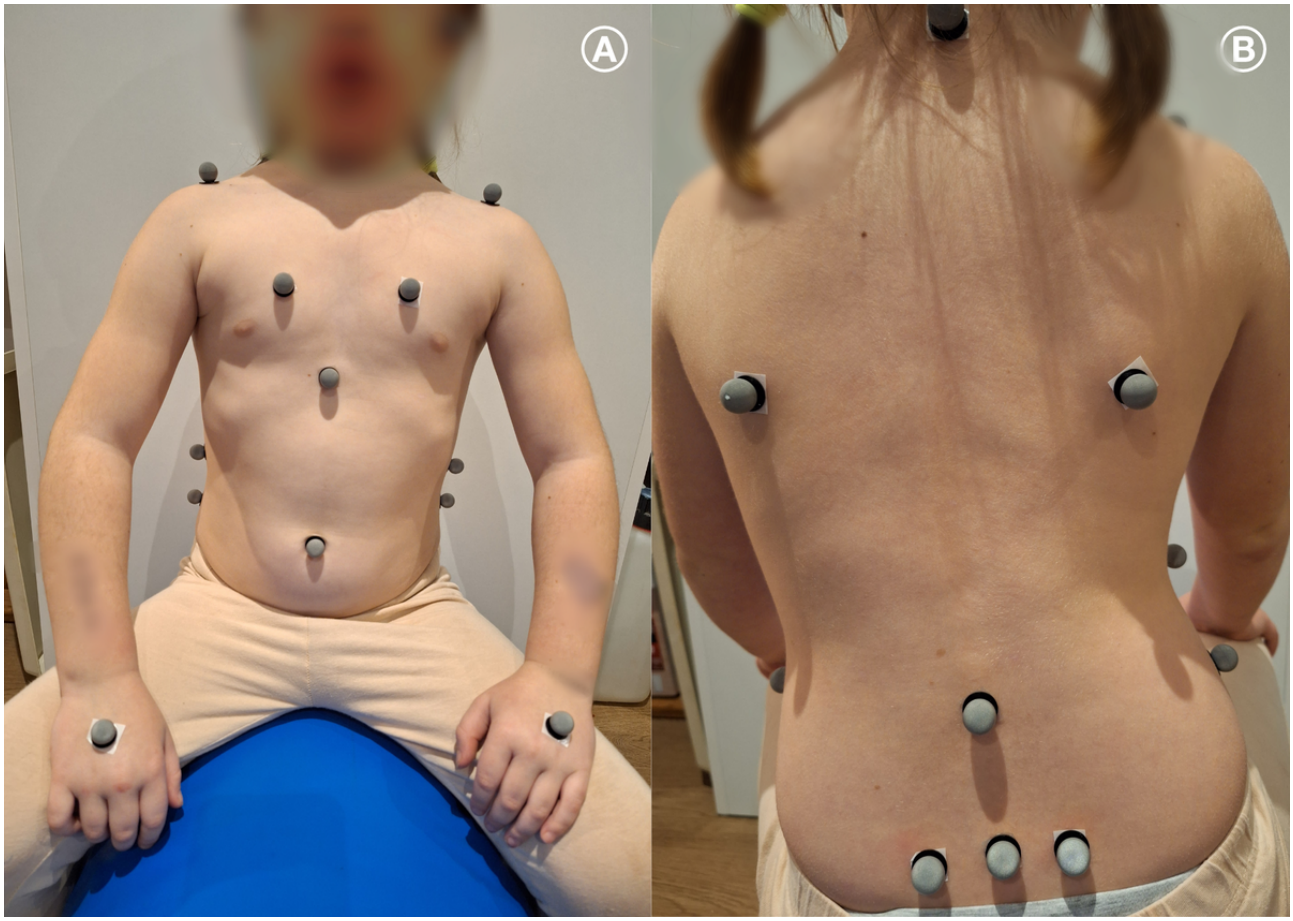

Test of Trunk and Cervical Spine Coordination Ability

The child performs repeated back-and-forth upper limb movements while sitting on the cylinder, typically 3‐6 repetitions. Key variables include the extent of thrust from the body's center, the point at which cervical spine extension begins as compensation, and the coordination between markers located on the dorsum of the hand, C7 vertebra, and temporomandibular joint. A higher functional level is reflected by the child's ability to synchronize cervical spine movements with trunk and upper limb actions over a longer period without compensatory cervical spine extensions (Figure 6).

XSL•FO
RenderX

Figure 6. Movements of the upper limbs while sitting on the cylinder: (A) forward and backward and (B) reaching sideways.



XSL•FO
RenderX

Abdominal Breathing Ability Test

Over a 30-second period, the child is instructed to breathe deeply, aiming for 8 - 12 respiratory cycles. Analysis focuses on the displacement of markers on the distal clavicles, distal sternum, L4, and the umbilicus during inspiration and expiration. The objective is to detect shifts toward an abdominal breathing pattern, with improved function indicated by increased distance between the Th/L and L4 markers relative to those around the umbilicus.

Ability to Work With the Center of Gravity

During the measurement, the child is asked to extend 1 upper limb sideways, and, depending on cooperation, may be encouraged to reach for a toy. Movements are performed 3 to 6 times within the 30-second recording period. This test evaluates stability and control over the center of gravity, with progress reflected in greater deviations of the markers from their initial positions after therapy (Figure 6).

Muscle Fatigue

Muscle fatigue will be measured using sEMG. A 16-channel surface EMG system from DELSYS will be used to measure the activity of the m. obliquus externus abdominis for 30 seconds. The child will lie on their back, and the therapist will present a toy near the opposite knee to encourage the child to reach for it. The movement will be performed first in an isometric contraction, followed by a repetitive movement. We expect the child to be able to repeat the movement 5 - 10 times during the 30-second measurement period. The EMG data will be processed using standard methods and normalized to the maximum EMG observed during isometric contraction to enable between-muscle comparisons (right and left side of the child). A lower median frequency and decreasing EMG amplitude during repeated movement indicate improved function, as the muscle performs the task more efficiently and with less effort. Additionally, a higher number of repetitions without significant changes in EMG characteristics reflects better function, suggesting functionally stronger or more coordinated muscle activation.

SCALE

The SCALE test is a standardized clinical test used to assess selective voluntary motor control in the lower extremities. It assesses the ability to isolate joint movements during 5 tasks for each lower extremity. Each joint movement is scored from 0 to 2, and a higher total score indicates better selective motor control [35].

The Modified Pediatric Reach Test

The Functional Reach Test is a simple, validated measure of dynamic sitting or standing balance that evaluates the maximal distance an individual can reach forward beyond arm's length while maintaining a fixed base of support. It is a reliable indicator of trunk stability and balance control in children with various neuromuscular conditions [36]. Longer reach distances indicate better dynamic postural control and balance performance [37].

Spirometry Measurement

Spirometry will be performed according to the international standards of the European Respiratory Society [38]. The patient will be fitted with a nose clip, then they will breathe into the spirometer through the mouthpiece while sitting on the cylinder. Each child receives the same instructions: Take a deep breath and then exhale as quickly and forcefully as possible. The test will be repeated 3 times in a row. It will be performed using the Vitalograph Asma-1 device. Forced expiratory volume in 1 second and peak expiratory flow will be measured [39]. Higher values indicate better function.

Scoliometer Measurement

The measurement will be performed using a scoliometer according to the standardized procedure for assessing the angle of trunk rotation (ATR). The scoliometer is a noninvasive diagnostic tool based on the principle of a spirit level, which allows for the orientation detection of trunk asymmetry and possible rotation of the vertebrae. The child being examined will sit on a cylinder with the upper limbs relaxed along the body. The child will then be asked to bend forward as much as possible while sitting. The scoliometer is placed on the spine so that the center of the device exactly copies the spinous processes. The measurement will be performed at 2 levels: Th/L transition (thoracolumbar transition) and Level 2/3 of the thoracic spine. At each level, the ATR value will be recorded as well as the side to which the trunk rotation is directed. Each measurement will be performed 3 times in succession, and the average of the values will be used. A higher ATR value indicates greater axial rotation of the trunk and thus a higher degree of possible scoliotic curvature [40].

MFM-20

The MFM-20 is a validated quantitative scale assessing motor skills in children with neuromuscular disorders aged 2 - 7 years. It contains 20 items covering standing, transfer, axial and proximal motor functions, and distal motor control. The total score is expressed as a percentage of the maximum possible score, with higher scores indicating better motor performance. The MFM-20 has been validated in SMA and Duchenne muscular dystrophy and is sensitive to changes over time [33].

SATCo

The SATCo is a standardized test designed to identify the highest level of segmental trunk control in sitting. It assesses static, active, and reactive control at 7 levels of trunk support. Although originally developed for children with cerebral palsy, it has been used in other pediatric neuromotor populations to evaluate trunk stability and postural development [41]. Higher scores represent a greater degree of segmental trunk control [42].

TCMS

The TCMS is a clinical assessment tool that evaluates trunk control during static and dynamic sitting balance tasks. It measures selective movement control in the trunk and coordination between upper and lower body segments. Although validated primarily in children with cerebral palsy, it is also used to assess postural control in children with other neuromotor

XSL•FO
RenderX

disorders such as SMA [43]. The total score ranges from 0 to 58, with higher scores indicating better trunk control [44].

Secondary Outcomes (Long-Term)

Quality of Life Questionnaire

Parents of children with SMA will complete a special questionnaire. The ICF score set, specifically the ICF-based Documentation Form for the category "Children with Cerebral Palsy Brief" (for children under 6 y old), will be used to evaluate the QoL [45]. Although the ICF-based Documentation Form was originally developed for children with cerebral palsy, it was selected in this trial due to its better applicability in clinical practice and its ability to capture functional and quality-of-life domains relevant also to children with SMA.

The Pediatric Quality of Life Inventory Generic Core Scales

Parents of children with SMA will also complete the Pediatric Quality of Life Inventory Generic Core Scales - Parent Proxy Report. The assessment tool, designed to evaluate health-related QoL in children, consists of 23 items across 4 domains (physical, emotional, social, and school functioning). It is used to monitor longitudinal changes in perceived QoL and to evaluate the effectiveness of clinical and rehabilitative interventions. Higher scores indicate better QoL [46].

Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire is a widely used behavioral screening tool assessing psychological attributes across 5 domains: emotional symptoms, conduct problems, hyperactivity or inattention, peer relationship problems, and prosocial behavior. It is suitable for children aged 2 to 17 years and provides a total difficulties score reflecting overall psychosocial well-being. Although originally validated in the general pediatric population, it has been successfully applied in children with chronic or neuromuscular disorders to assess emotional and social adaptation related to health status. Higher total scores indicate greater difficulties, while higher prosocial subscale scores indicate better social functioning [47].

Home Video Recording

Parents will send home video recordings of their child's movement behavior. The video recordings will be evaluated by 2 independent physiotherapists who will assess quality and extent of the children's movement. Movement of greater quality or scope indicates better function.

Molecular Biomarkers

For lncRNA analysis, peripheral blood will be collected in the evening before and after the completion of the rehabilitation course, directly at the accommodation site for families with children with SMA. Given the collection site, blood samples will be immediately frozen using dry ice and transported to the laboratory, where they will be stored at -80°C. Rapid freezing eliminates RNA degradation as much as possible. RNA will be isolated from gently thawed samples [48] using the NucleoSpin RNA blood kit (Macherey-Nagel), which is specially designed for RNA isolation from frozen blood. The expected average yield is 7 µg per 1 ml of blood. Total RNA will be transcribed into cDNA using reverse transcriptase. The expression of selected lncRNAs will be detected by qPCR using hydrolysis probes designed specifically for the selected lncRNAs. Relative quantification with selected reference genes will be used for expression analysis. Reference genes will be selected during qPCR protocol optimization using the R program and recommended packages (eg, NormFinder). All qPCR experiments will be done in accordance with MIQE (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) guidelines.

Statistical Analysis and Software

Data will be presented as means (SDs), medians (IQRs), or as absolute numbers with percentages, as appropriate. The quantitative outcomes will be expressed as differences relative to the corresponding pretherapy values (post – pre), while qualitative outcomes will be categorized as improved, unchanged, or worsened compared to pretherapy. For each quantitative outcome, the treatment effect will be evaluated by testing within-subject differences between NEUROEQUIP-SMA therapy (A) and SMA-SOC-N therapy (B) using a paired *t* test. Although no period effect (eg, due to disease progression or season) is expected, it will be assessed by comparing within-subject differences (Period 1 vs Period 2) using a paired *t* test. The study was designed to avoid carryover effects through a long washout period; nonetheless, carryover effects will be explored by comparing between-subject differences in outcome sums (Sequence AB vs Sequence BA) using a 2-sample *t* test. If the normality assumption is not met, the exact Wilcoxon signed-rank test will be used instead of the paired *t* test, and the exact Wilcoxon rank-sum test instead of the 2-sample *t* test. Normality will be assessed both graphically (eg, Q–Q plot) and using the Shapiro–Wilk test. Similarly, for qualitative outcomes, the effects of treatment, period, and carryover will be assessed using the McNemar test for within-subject differences and Fisher exact test for between-subject differences.

When conducting multiple comparisons, the achieved levels of significance (*P* values) will be adjusted using a correction method, such as the Holm-Bonferroni method. All tests will be 2-sided, with *P* values less than .05 considered statistically significant. Effect sizes will be reported alongside *P* values to provide insight into the magnitude of observed changes. Sensitivity analyzes for outliers will be included, comparing results with and without these subjects. Data analysis will be performed using R software (R Development Core Team) [49]. Missing data will not be imputed due to the small sample size, which could make imputation unstable; instead, an available-case analysis will be performed, and the pattern of missing data will be explored. No interim analyzes are planned. The trial will be conducted in full as described, unless terminated early for ethical or safety reasons upon recommendation of the principal investigator or the ethics committee.

Harms

The physiotherapists are qualified professionals who actively approach all subjects with respect to their individual impairment. The horses performing the therapy in the study have specialized examinations, and the trainers are professionals with many years of experience to sufficiently ensure safety during

XSL•FO
RenderX

NEUROEQUIP-SMA and therapeutic care. The program is individualized to avoid burdening or harming the patient. All adverse events (AEs) must be reported to the principal investigator and the appropriate specialist. All AEs will be monitored and systematically recorded by the treating physiotherapists during the study using a predefined reporting form. AEs will be classified according to severity (mild, moderate, severe) and potential relationship to the intervention (unrelated, possibly related, and definitely related) and will be reviewed by the principal investigator and a designated clinical specialist. Serious AEs will be reported immediately to the ethics committee. In the event of any injury related to the intervention, appropriate clinical care will be provided, but no financial compensation will be provided.

Ethical Considerations

The multicentric Ethics Committee of the Third Faculty of Medicine of Charles University has approved the study under the code UK3LF/658559/12025, based on submitted informed consent forms. The results of the study will be disseminated to the public and the scientific and medical community (published in journal articles and conferences). Anonymized data and outcome data may be published for research purposes (for meta-analysis, systematic reviews, etc) on request. This study does not include a Data Monitoring Committee, as the intervention is short-lasting, low risk, and nonpharmacological in nature. Oversight is ensured by the principal investigator and the ethics committee. Any modifications to the study design will be submitted to the ethics committee for review, and the informed consent form will be revised accordingly. No monetary compensation was provided to participants. Written informed consent for publication of identifiable images was obtained from the legal guardians of all participating children.

Results

All potentially eligible study participants will be contacted in December 2025. Data collection will take place from February 2026 to May 2026. The start of statistical analyzes is planned for September 2026. In the crossover design, 20 children are planned to receive both therapies over 2 periods, enabling within-subject comparisons. The washout period of sufficient length will minimize potential carryover effect between periods. A paired *t* test comparing the primary outcome between NEUROEQUIP-SMA and SMA-SOC-N therapies in 20 children will have 80% power to detect moderate-to-large effect size (Cohen *d*=0.66) at a 5% significance level. Analyses of other outcomes are exploratory in nature and aimed at informing future hypothesis-driven research.

Although period and carryover effects on the outcome are not expected, they will also be assessed alongside the treatment effects; however, all these effects will be analyzed separately due to the small sample size. Along with *P* values, effect sizes will be reported to reflect the magnitude of changes. We would like to publish the collected data at the beginning of 2027.

Discussion

Anticipated Findings

This protocol presents the design of a clinical trial comparing an innovative NEUROEQUIP-SMA approach with standard physiotherapy in children with SMA. This is the first study of its kind to evaluate the effect of EAT based on the principles of neuroproprioceptive “facilitation and inhibition.” We plan to support the results with molecular data from lncRNA analysis. Based on previous studies [25], we predict that the expression levels of selected lncRNAs (GAS5, H19, MALAT1, MEG3, NEAT1, and PARTICLE) will change as a result of therapy. These lncRNAs have not been sufficiently studied in relation to SMA, but based on their mechanism of action and detection in other diseases, we postulate their deregulation in the blood of patients with SMA. This study may be the first step toward understanding the involvement of lncRNAs in the pathophysiology of this disease and their use as prognostic markers [48]. The uniqueness of the study also lies in the specific target group: children who have undergone gene therapy. We anticipate that intensive EAT based on the principles of neuroproprioceptive “facilitation and inhibition” will be proven and well tolerated in children with SMA. We also expect that the protocol will allow for the capture of changes at the molecular, biomechanical, and clinical levels.

In the Czech Republic, the Nusinersen substitution therapy has been included in the system of reimbursed care as regular reimbursement since 2018 [50], while the Zolgensma gene therapy was approved by the Ministry of Health of the Czech Republic in April 2020 before registration as preauthorized use [51]. Physiotherapists are thus in current practice dealing with diverse groups of children: those treated with substitution therapy for a long time, those treated with a combination of both approaches, and those who have received gene therapy very early. We have a new generation of children who can be expected to slow or stop the progression of the disease [52,53].

Based on these changes, there is a need to revise existing rehabilitation strategies in clinical practice. Some physiotherapists are recommending an increase in the frequency and intensity of therapy [54] and new standards of care are needed to better reflect patient expectations and variable motor skills [55]. Although pharmacological treatment significantly improves survival, motor function often remains limited and is accompanied by pathological compensatory mechanisms [12,56,57]. Targeted correction of these functions with appropriate rehabilitation approaches is therefore essential. For this reason, an intensive six-day course and the therapeutic neuroproprioceptive principle were chosen.

Comparison With Prior Work

There is currently only 1 study examining the effect of EAT in children with SMA. Lemke et al [23] focused on subjective assessment of the effects of EAT using a questionnaire survey. In contrast, our study evaluates the direct response to a therapeutic intervention, which we are able to observe at the clinical, biomechanical, and molecular levels, including comparison with a standardized SMA-SoC procedure.

XSL•FO
RenderX

Research on conventional rehabilitation approaches for children with SMA has so far focused mainly on maintaining existing function, preventing contractures, and promoting self-sufficiency [6,12,58]. Most of these were designed before the onset of gene therapy, when disease progression was considered inevitable and often lacked targeted monitoring of compensatory mechanisms. In our study, we augmented standardized tests with qualitative assessments to capture potential re-education of pathological patterns. In the era of genetic treatments, we consider the monitoring of movement quality to be essential, and the existing literature has not yet addressed this milestone.

Cammarano et al [59] summarize that rehabilitation care for SMA, whether before or after genetic treatment, lacks standardization, high-quality randomized trials, and comprehensive multidomain evaluations. Our study addresses this requirement: it includes children treated with both replacement and genetic therapy.

Limitations of This Study

As the participants are children, the results are partly affected by their level of cooperation and willingness to participate. Although we carefully selected the assessment tools to minimize this source of variability, some degree of bias cannot be excluded. Due to the extensive test battery and possible fatigue effects, we administer each test only once the child is able to respond appropriately. Limitations can also be seen in the selection of the quality-of-life test. Although the ICF tool is practical and most commonly used, it was originally validated for cerebral palsy, which may reduce its specificity for SMA.

The functional focus, specificity, precision, and technique inherent in the NEUROEQUIP-SMA methodology are key to the overall therapeutic effect and can account for minor differences in results. Since SMA is a rare genetic disease, participants were recruited from across the country. The number of participants is limited by the rarity of the disease, which limits the statistical power and generalizability of the results. Bias may also occur due to the established randomization process, where therapists and families could not be blinded. Due to travel restrictions, the long-term effects of the NEUROEQUIP-SMA intervention can only be partially assessed. Furthermore, given the severity of the disease, it is likely that children will engage in further physiotherapy interventions approximately 4 weeks after completing our treatment program. Although the small sample size limits the use of more complex statistical models, such as linear mixed models, the results will provide a solid basis for future research.

Given the crossover design, potential attrition before the second half of the study is considered a methodological risk. To mitigate this risk, families will be engaged through regular communication, stable scheduling, and motivation supported by the same therapeutic team across both parts of the study.

Conclusions

The proposed evaluation methods provide a comprehensive assessment of the NEUROEQUIP-SMA approach and can serve as a foundation for future comparisons of different methodologies in equine-assisted therapy. The findings of this study have the potential to inform and shape rehabilitation strategies for children with SMA receiving genetic or substitution pharmacotherapy. The results will be shared with the professional and lay public and translated into clinical practice. We hope that this protocol will spark a discussion about possible changes in intervention for children with SMA on genetic therapy and highlight the need for larger and longer-term multicenter studies to verify durability of effect, clinical efficacy, and transferability to practice.

Acknowledgments

We would like to sincerely thank all participating children and their families for their trust, commitment, and cooperation throughout the study. Our thanks also go to the entire staff of the Mirákl Hipporehabilitation Center for their dedicated help during the intensive therapy blocks. We especially appreciate the invaluable contribution of the therapy horses, whose responsiveness and cooperation were essential for the success of the intervention. OpenAI tool ChatGPT was used to correct grammatical errors and improve the clarity of text composed. The views expressed are those of the authors and do not necessarily reflect those of their affiliated institutions. The study is not intended to replace clinical judgment or individualized treatment recommendations. The authors declare that all statements made in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or the publisher, editors, and reviewers. No patients or public were involved in the design, development, or management of the trial.

Funding

The study is funded by the Grant Agency of Charles University (4424; January 2024), the Cooperation program (Neurosciences), and SVV 260533/SVV/2024. The study sponsors financially supported the implementation of the study, including material and operating costs. They did not influence the study design, data collection, management, analysis, interpretation, writing of reports, or the decision to submit for publication. The researchers retained full authority over these activities without any contribution from the sponsor.

Data Availability

The data collected during this study have not yet been analyzed. After the completion of data analysis, anonymized datasets will be made available upon reasonable request. Researchers interested in accessing the data may contact the corresponding author, and the data will be shared for scientific purposes.

XSL•FO
RenderX

The authors attest that they obtained written consent from the study participants, or their legally authorized representatives or guardians, to publish their likenesses in this open access article.

Authors' Contributions

Conception and design of the study: KŘ and KM
Methodology of laboratory assessment: MV
Methodology of therapy content: KŘ and KM
Project administration: KŘ, KM, MP, and MV
Conceptualization and methods of lncRNA: MP and MČ
Data set preparation, data management, and statistical analysis: JR
Guarantor of the article: KŘ.

Conflicts of Interest

None declared.

Checklist 1
CONSORT-EHEALTH checklist (V1.6.1).
[PDF File, 1316 KB - resprot_v15i1e83266_app1.pdf]

Checklist 2
SPIRIT checklist.
[DOCX File, 35 KB - resprot_v15i1e83266_app2.docx]

References

1. Voight S, Arya K. Considerations for treatment in clinical care of spinal muscular atrophy patients. Children (Basel) 2024 Apr 20;11(4):495. [doi: 10.3390/children11040495] [Medline: 38671712]
2. García BEC, Santos CCC, Tapia CPF, et al. Spinal muscular atrophy clinical features, classification, natural history, genetics, diagnosis, complications and treatment of the disease. EPRA 2023:191-196. [doi: 10.36713/epra12994]
3. Arnold WD, Kassar D, Kissel JT. Spinal muscular atrophy: diagnosis and management in a new therapeutic era. Muscle Nerve 2015 Feb;51(2):157-167. [doi: 10.1002/mus.24497] [Medline: 25346245]
4. Rouault F, Christie-Brown V, Broekgaarden R, et al. Disease impact on general well-being and therapeutic expectations of European Type II and Type III spinal muscular atrophy patients. Neuromuscul Disord 2017 May;27(5):428-438. [doi: 10.1016/j.nmd.2017.01.018] [Medline: 28237437]
5. Nishio H, Niba ETE, Saito T, Okamoto K, Takeshima Y, Awano H. Spinal muscular atrophy: the past, present, and future of diagnosis and treatment. Int J Mol Sci 2023 Jul 26;24(15):11939. [doi: 10.3390/ijms241511939] [Medline: 37569314]
6. Mercuri E, Sumner CJ, Muntoni F, Darras BT, Finkel RS. Spinal muscular atrophy. Nat Rev Dis Primers 2022 Aug 4;8(1):52. [doi: 10.1038/s41572-022-00380-8] [Medline: 35927425]
7. Davidson JE, Farrar MA. The changing therapeutic landscape of spinal muscular atrophy. Aust J Gen Pract 2022;51(1-2):38-42. [doi: 10.31128/AJGP-03-21-5924] [Medline: 35098272]
8. Bose M, Parab SD, Patil SM, Pandey NA, Pednekar GV, Saini SS. Exploring spinal muscular atrophy and its impact on functional status: Indian scenario. Indian J Public Health 2019;63(3):254-257. [doi: 10.4103/ijph.IJPH_239_18] [Medline: 31552858]
9. Iannaccone ST, White M, Browne R, Russman B, Buncher R, Samaha FJ. Muscle fatigue in spinal muscular atrophy. J Child Neurol 1997 Aug;12(5):321-326. [doi: 10.1177/088307389701200507] [Medline: 9378900]
10. Lewelt A, Krosschell KJ, Stoddard GJ, et al. Resistance strength training exercise in children with spinal muscular atrophy. Muscle Nerve 2015 Oct;52(4):559-567. [doi: 10.1002/mus.24568] [Medline: 25597614]
11. Yi YG, Shin HI, Jang DH. Rehabilitation of spinal muscular atrophy: current consensus and future direction. J Genet Med 2020 Dec 31;17(2):55-61. [doi: 10.5734/JGM.2020.17.2.55]
12. Mercuri E, Finkel RS, Muntoni F, et al. Diagnosis and management of spinal muscular atrophy: part 1: recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. Neuromuscul Disord 2018 Feb;28(2):103-115. [doi: 10.1016/j.nmd.2017.11.005] [Medline: 29290580]
13. Dunaway S, Montes J, McDermott MP, et al. Physical therapy services received by individuals with spinal muscular atrophy (SMA). In: Matthews DJ, editor. J Pediatr Rehabil Med 2016;9(1):35-44. [doi: 10.3233/PRM-160360] [Medline: 26966799]
14. Jędrzejowska M, Kostera-Pruszczyk A. Spinal muscular atrophy - new therapies, new challenges. Neurol Neurochir Pol 2020;54(1):8-13. [doi: 10.5603/PJNNS.a2019.0068] [Medline: 31922583]
15. Chali F, Desseille C, Houdebine L, et al. Long-term exercise-specific neuroprotection in spinal muscular atrophy-like mice. J Physiol 2016 Apr 1;594(7):1931-1952. [doi: 10.1113/JP271361] [Medline: 26915343]

XSL•FO
RenderX

16. Řasová K, Martinková P, Vařejková M, et al. COMIRESTROKE-a clinical study protocol for monitoring clinical effect and molecular biological readouts of COMprehensive Intensive REhabilitation program after STROKE: a four-arm parallel-group randomized double blinded controlled trial with a longitudinal design. Front Neurol 2022;13:954712. [doi: 10.3389/fneur.2022.954712] [Medline: 36388199]
17. Řasová K, Bučková B, Prokopiusová T, et al. A Three-Arm Parallel-group Exploratory Trial documents balance improvement without much evidence of white matter integrity changes in people with multiple sclerosis following two months ambulatory neuroproprioceptive "facilitation and inhibition" physical therapy. Eur J Phys Rehabil Med 2021;57(6). [doi: 10.23736/S1973-9087.21.06701-0]
18. Plotas P, Papadopoulos A, Apostolelli EM, et al. Effects of hippotherapy on motor function of children with cerebral palsy: a systematic review study. Ital J Pediatr 2024 Sep 19;50(1):188. [doi: 10.1186/s13052-024-01715-9] [Medline: 39300490]
19. Wood W, Alm K, Benjamin J, et al. Optimal terminology for services in the United States that incorporate horses to benefit people: a consensus document. J Altern Complement Med 2021 Jan;27(1):88-95. [doi: 10.1089/acm.2020.0415] [Medline: 33252244]
20. Stern C, Chur‐hansen A. An umbrella review of the evidence for equine‐assisted interventions. Aust J Psychol 2019 Dec 1;71(4):361-374. [doi: 10.1111/ajpy.12246]
21. Pantera E, Froment P, Vernay D. Does hippotherapy improve the functions in children with cerebral palsy? Systematic review based on the international classification of functioning. J Integr Complement Med 2022 Sep;28(9):705-720. [doi: 10.1089/jicm.2021.0417] [Medline: 35671520]
22. Hyun C, Kim K, Lee S, Ko N, Lee IS, Koh SE. The short-term effects of hippotherapy and therapeutic horseback riding on spasticity in children with cerebral palsy: a meta-analysis. Pediatr Phys Ther 2022 Apr 1;34(2):172-178. [doi: 10.1097/PEP.0000000000000880] [Medline: 35184078]
23. Lemke D, Rothwell E, Newcomb TM, Swoboda KJ. Perceptions of equine-assisted activities and therapies by parents and children with spinal muscular atrophy. Pediatr Phys Ther 2014;26(2):237-244. [doi: 10.1097/PEP.0000000000000027] [Medline: 24675128]
24. d'Ydewalle C, Ramos DM, Pyles NJ, et al. The antisense transcript SMN-AS1 regulates SMN expression and is a novel therapeutic target for spinal muscular atrophy. Neuron 2017 Jan 4;93(1):66-79. [doi: 10.1016/j.neuron.2016.11.033] [Medline: 28017471]
25. Pokorná M, Černá M, Boussios S, Ovsepian SV, O'Leary VB. lncRNA biomarkers of glioblastoma multiforme. Biomedicines 2024 Apr 23;12(5):932. [doi: 10.3390/biomedicines12050932] [Medline: 38790894]
26. O'Leary VB, Ovsepian SV, Carrascosa LG, et al. PARTICLE, a triplex-forming long ncRNA, regulates locus-specific methylation in response to low-dose irradiation. Cell Rep 2015 Apr 21;11(3):474-485. [doi: 10.1016/j.celrep.2015.03.043] [Medline: 25900080]
27. Vangoor VR, Gomes-Duarte A, Pasterkamp RJ. Long non-coding RNAs in motor neuron development and disease. J Neurochem 2021 Mar;156(6):777-801. [doi: 10.1111/jnc.15198] [Medline: 32970857]
28. pwr: basic functions for power analysis. R Project. URL: http://cran.r-project.org/web/packages/pwr/index.html [accessed 2026-01-06]
29. Create a blocked randomisation list. Sealed envelope. URL: https://www.sealedenvelope.com/simple-randomiser/v1/lists [accessed 2025-05-15]
30. Ministry of Health of the Czech Republic. Decision on the extension of accreditation nomzdr 39110/2019-7/ONP [internal document provided to the authors' institution]. : Ministry of Health; 2019.
31. Ministry of Health of the Czech Republic. Decision on the granting of accreditation no.55296/2014-3/ONP [internal document provided to the authors' institution]. : Ministry of Health; 2014.
32. Řasová K, Bučková B, Prokopiusová T, et al. A three-arm parallel-group exploratory trial documents balance improvement without much evidence of white matter integrity changes in people with multiple sclerosis following two months ambulatory neuroproprioceptive "facilitation and inhibition" physical therapy. Eur J Phys Rehabil Med 2021 Dec;57(6):889-899. [doi: 10.23736/S1973-9087.21.06701-0] [Medline: 33565742]
33. Mazzone E, De Sanctis R, Fanelli L, et al. Hammersmith Functional Motor Scale and Motor Function Measure-20 in non ambulant SMA patients. Neuromuscul Disord 2014 Apr;24(4):347-352. [doi: 10.1016/j.nmd.2014.01.003] [Medline: 24491485]
34. Glanzman AM, Mazzone E, Main M, et al. The Children's Hospital of Philadelphia infant test of neuromuscular disorders (CHOP INTEND): test development and reliability. Neuromuscul Disord 2010 Mar;20(3):155-161. [doi: 10.1016/j.nmd.2009.11.014] [Medline: 20074952]
35. Fowler EG, Staudt LA, Greenberg MB, Oppenheim WL. Selective Control Assessment of the Lower Extremity (SCALE): development, validation, and interrater reliability of a clinical tool for patients with cerebral palsy. Dev Med Child Neurol 2009 Aug;51(8):607-614. [doi: 10.1111/j.1469-8749.2008.03186.x] [Medline: 19220390]
36. Bahar A, Ozgen A, Ayse K. Reliability of balance tests in children with duchenne muscular dystrophy. Sci Res Essays 2011;6(20):4428-4431. [doi: 10.5897/SRE10.1185]
37. Bartlett D, Birmingham T. Validity and reliability of a pediatric reach test. Pediatr Phys Ther 2003;15(2):84-92. [doi: 10.1097/01.PEP.0000067885.63909.5C] [Medline: 17057438]

XSL•FO
RenderX

38. Miller MR, Hankinson J, Brusasco V, et al. Standardisation of spirometry. Eur Respir J 2005 Aug;26(2):319-338. [doi: 10.1183/09031936.05.00034805] [Medline: 16055882]
39. Quanjer PH, Stanojevic S, Cole TJ, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J 2012 Dec;40(6):1324-1343. [doi: 10.1183/09031936.00080312] [Medline: 22743675]
40. Adamczewska K, Wiernicka M, Malchrowicz-Mośko E, Małecka J, Lewandowski J. The angle of trunk rotation in school children: a study from an idiopathic scoliosis screening. prevalence and optimal age screening value. Int J Environ Res Public Health 2019 Sep 16;16(18):3426. [doi: 10.3390/ijerph16183426] [Medline: 31527403]
41. Sakanaka TE, Butler P, Kulshrestha R, Willis T, Loram I. Children with spinal muscular atrophy treated with disease-modifying therapies, do they attain neutral vertical head and trunk control? Rehabilitation Medicine and Physical Therapy. . [doi: 10.1101/2025.02.24.25321598]
42. Butler PB, Saavedra S, Sofranac M, Jarvis SE, Woollacott MH. Refinement, reliability, and validity of the segmental assessment of trunk control. Pediatr Phys Ther 2010;22(3):246-257. [doi: 10.1097/PEP.0b013e3181e69490] [Medline: 20699770]
43. Bulut N, Aydın Yağcıoğlu G, Yardımcı-Lokmanoğlu BN, Alemdaroğlu-Gürbüz İ, Yılmaz Ö, Karaduman A. Reliability and validity of the Trunk Control Measurement Scale in children with duchenne muscular dystrophy. Percept Mot Skills 2024 Dec;131(6):2167-2181. [doi: 10.1177/00315125241286622] [Medline: 39308035]
44. Heyrman L, Molenaers G, Desloovere K, et al. A clinical tool to measure trunk control in children with cerebral palsy: the Trunk Control Measurement Scale. Res Dev Disabil 2011;32(6):2624-2635. [doi: 10.1016/j.ridd.2011.06.012] [Medline: 21757321]
45. Comprehensive ICF core set for children and youth with CP. ICF Research Branch. URL: https://www.icf-research-branch.org/icf-core-sets?task=download.send&id=210:comprehensive-icf-core-set-for-children-and-youth-with-cp&catid=20 [accessed 2025-03-27]
46. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the pediatric quality of life inventory version 4.0 generic core scales in healthy and patient populations. Med Care 2001 Aug;39(8):800-812. [doi: 10.1097/00005650-200108000-00006] [Medline: 11468499]
47. Goodman R. The Strengths and Difficulties Questionnaire: a research note. J Child Psychol Psychiatry 1997 Jul;38(5):581-586. [doi: 10.1111/j.1469-7610.1997.tb01545.x] [Medline: 9255702]
48. Yamagata H, Kobayashi A, Tsunedomi R, et al. Optimized protocol for the extraction of RNA and DNA from frozen whole blood sample stored in a single EDTA tube. Sci Rep 2021 Aug 23;11(1):17075. [doi: 10.1038/s41598-021-96567-2] [Medline: 34426633]
49. The R Project for Statistical Computing. URL: https://www.R-project.org/ [accessed 2025-12-24]
50. Community update on access to Spinraza in Europe. SMA Europe. URL: https://www.sma-europe.eu/news/community-update-on-access-to-spinraza-in-europe [accessed 2025-08-04]
51. Czech Republic allows pre-authorized use of Zolgensma. Eversana. URL: https://www.eversana.com/2020/04/23/czech-republic-preauthorized-zolgensma [accessed 2025-08-04]
52. Mahajan R. Onasemnogene abeparvovec for spinal muscular atrophy: the costlier drug ever. Int J Appl Basic Med Res 2019;9(3):127-128. [doi: 10.4103/ijabmr.IJABMR_190_19] [Medline: 31392173]
53. Zolgensma data, including patients with more severe SMA at baseline, further demonstrate therapeutic benefit, including prolonged event-free survival, increased motor function, and milestone achievement. NOVARTIS. URL: https://www.novartis.com/news/media-releases/zolgensma-data-including-patients-more-severe-sma-baseline-further-demonstrate-therapeutic-benefit-including-prolonged-event-free-survival-increased-motor-function-and-milestone-achievement [accessed 2025-08-04]
54. Trenkle J, Brugman J, Peterson A, Roback K, Krosschell KJ. Filling the gaps in knowledge translation: physical therapy recommendations for individuals with spinal muscular atrophy compared to standard of care guidelines. Neuromuscul Disord 2021 May;31(5):397-408. [doi: 10.1016/j.nmd.2021.02.011] [Medline: 33741230]
55. Mortenson P, Cielecka J, Harrison E, Zwicker JG. Rehabilitation practices for childhood spinal muscular atrophy. Disabil Rehabil 2025 Oct;47(21):5538-5547. [doi: 10.1080/09638288.2025.2484345] [Medline: 40163566]
56. Jablonka S, Hennlein L, Sendtner M. Therapy development for spinal muscular atrophy: perspectives for muscular dystrophies and neurodegenerative disorders. Neurol Res Pract 2022 Jan 4;4(1):2. [doi: 10.1186/s42466-021-00162-9] [Medline: 34983696]
57. Song W, Ke X. Rehabilitation management for patients with spinal muscular atrophy: a review. Orphanet J Rare Dis 2025 Jul 10;20(1):352. [doi: 10.1186/s13023-025-03888-w] [Medline: 40640932]
58. Shin HI. Rehabilitation strategies for patients with spinal muscular atrophy in the era of disease-modifying therapy. Ann Rehabil Med 2024 Aug;48(4):229-238. [doi: 10.5535/arm.240046] [Medline: 39210748]
59. Cammarano S, Chirico VA, Giardulli B, Mazzuoccolo G, Ruosi C, Corrado B. Physical and respiratory rehabilitation in spinal muscular atrophy: a critical narrative review. Appl Sci (Basel) 2025;15(8):4398. [doi: 10.3390/app15084398]

XSL•FO
RenderX

Abbreviations

AE: adverse event
CHOP INTEND: Children's Hospital of Philadelphia Infant Test of Neuromuscular Disorders
ICF: International Classification of Functioning, Disability, and Health
lncRNA: long noncoding RNA
MFM-20: Motor Function Measure-20
QoL: quality of life
SCALE: Selective Control Assessment of the Lower Extremity
sEMG: surface electromyography
SMA: spinal muscular atrophy
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by J Sarvestan; submitted 01.Sep.2025; peer-reviewed by D Ravnik, RQ Wang; accepted 10.Dec.2025; published 19.Feb.2026.

Please cite as:
Marikova K, Reissigova J, Vilimek M, Cerna M, Pokorna M, Rasova K
Immediate and Sustained Effects of Intensive Equine-Assisted Physiotherapy Based on Neuroproprioceptive "Facilitation and Inhibition" on Psychomotor Development, Clinical Functions, Quality of Life, and Molecular Biological Indicators in Children With Spinal Muscular Atrophy: Protocol for a Crossover Randomized Controlled Trial
JMIR Res Protoc 2026;15:e83266
URL: https://www.researchprotocols.org/2026/1/e83266
doi:10.2196/83266

© Katerina Marikova, Jindra Reissigová, Miloslav Vilimek, Marie Cerna, Marketa Pokorna, Kamila Rasova. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy of Acupuncture for Mild to Moderate Depression in Older People: Protocol for a Randomized Controlled Trial

Qingnan Fu[1*], MM; Kaihui Xiao[2*], MD; Jie Zhang[1], MM; Yang Li[1], MD; Yuxian Wang[1], MD; Molin Jiang[1], MD; Zengqi Man[3], MM; Jing Yang[1], MM; Wei Lu[1], MD

[1]Department of Psychosomatic Medicine, Beijing Traditional Chinese Medicine Hospital Affiliated to Capital Medical University, 23 Back Street, Art Museum, Dongcheng District, Beijing, China

[2]Neurology Ward 1, Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing, China

[3]Acupuncture and Massage College, Capital Medical University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Wei Lu, MD
Department of Psychosomatic Medicine, Beijing Traditional Chinese Medicine Hospital Affiliated to Capital Medical University, 23 Back Street, Art Museum, Dongcheng District, Beijing, China

Abstract

Background: Selective serotonin reuptake inhibitors are first-line antidepressants; however, only approximately 60% of patients can benefit from them. There is insufficient evidence for using acupuncture for symptom relief or for improving tolerance to selective serotonin reuptake inhibitors.

Objective: This randomized controlled trial aims to assess the effects of acupuncture combined with citalopram hydrobromide on mild to moderate depression in older people.

Methods: This study is a 2-arm, parallel, randomized controlled trial. A total of 132 participants aged 60 to 80 years diagnosed with major depressive disorder were divided into an acupuncture and medication group or a medication group. Participants in both groups take citalopram hydrobromide at a dose of up to 20 mg daily for 12 weeks. The acupuncture and medication group additionally receives 36 sessions of acupuncture treatment over 12 weeks. The primary outcome is the response rate of the 17-item Hamilton Depression Scale at the twelfth week. The secondary outcomes include changes in scores on the 17-item Hamilton Depression Scale and Mini-Mental State Examination at various time points. Adverse events will be recorded in detail.

Results: The study commenced on June 30, 2023, and as of October 17, 2024, a total of 132 participants had been enrolled. Data collection has been completed. Currently, data analysis is in progress, with preliminary findings anticipated to be available by October 2025. The findings of this study are expected to be submitted for publication in 2026.

Conclusions: This pilot study is expected to provide critical insights into the feasibility of integrating acupuncture with standard medication for managing mild to moderate depression in older people. By generating preliminary evidence on its potential benefits, the study aims to inform the design and sample size estimation of future multicenter trials, potentially advancing nonpharmacological treatment options for depression.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2300072740; https://www.chictr.org.cn/showproj.html?proj=198821

International Registered Report Identifier (IRRID): DERR1-10.2196/79327

(JMIR Res Protoc 2026;15:e79327) doi:10.2196/79327

KEYWORDS

depression; mild to moderate depression in older people; acupuncture; randomized controlled trial; protocol

Introduction

Depression is a common mental disorder caused by an interaction of social, psychological, and biological factors that can occur in anyone [1]. The clinical symptoms include anhedonia, fatigue, and negative emotions [2,3], as well as increased vulnerability to physical illnesses such as stroke, pain, and cancer [4,5], which may contribute to poor responses to pharmacological treatments. The World Health Organization predicted that it would become the greatest cause of global disease burden by 2030 [6]. Depression is a major issue for older adults, and its prevalence is rising [7]. Old age is the time of life when emotional fragility is accentuated. In addition to neurobiological changes in the brain, aging inevitably entails significant losses over time, not only in terms of individuals' emotions but also in terms of their physical condition and social status. Depression is the most common psychological disorder

XSL•FO
RenderX

among people older than 65 years and affects approximately 15% of this age group [8].

Currently, drug therapy and psychotherapy are the main forms of treatment for depression, with various complementary therapies serving as an adjunct. The standard drug treatments (mainly selective serotonin reuptake inhibitors [SSRIs]), such as paroxetine, sertraline, and citalopram, are not entirely satisfactory. The response rate to SSRIs, defined as at least a 50% improvement from baseline, is only approximately 60%, with documented side effects such as nausea, weight gain, and reduced sexual function, which limit patient compliance with the treatment [9,10]. As a special group, older adults require safer support options due to their low tolerance to medications, poor compliance with psychotherapy, and higher rates of chronic physical illnesses [11]. Therefore, new strategies to enhance the efficacy of antidepressants in the early stage of treatment are urgently needed.

Acupuncture treatment for depression has been provided alone or together with antidepressants in China and worldwide for years. Growing evidence indicates that acupuncture combined with drug treatment is more effective than drug treatment alone, is safe and well tolerated, and has an early onset of action [12]. However, systematic reviews have shown the potential benefits of acupuncture as an add-on therapy, but they have failed to obtain high-certainty evidence. The main limits of past studies have been a lack of large sample sizes, a high risk of bias, and inconsistent results [13-15].

This randomized controlled trial (RCT) is designed to evaluate the difference in efficacy between citalopram hydrobromide alone versus citalopram hydrobromide together with acupuncture for mild to moderate depression in older people.

Methods

Study Design

This prospective RCT was performed at the Department of Psychosomatic Medicine in Beijing Traditional Chinese Medicine Hospital Affiliated to Capital Medical University between January 2022 and December 2024. The protocol was developed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 checklist [16] and the STRICTA (Standards for Reporting Interventions in Clinical Trials of Acupuncture) guidelines [17]. The trial was registered with the Chinese Clinical Trial Registry (ChiCTR2300072740) in June 2023, prior to the enrollment of the first participant (on June 30, 2023). A flowchart of the trial design is shown in Figure 1, and the study timeline is presented in Table 1.

XSL•FO
RenderX

Figure 1. Flowchart of the trial design. HAMD-17: 17-item Hamilton Depression Scale; MMSE: Mini-Mental State Examination; TESS: Treatment Emergent Symptom Scale.

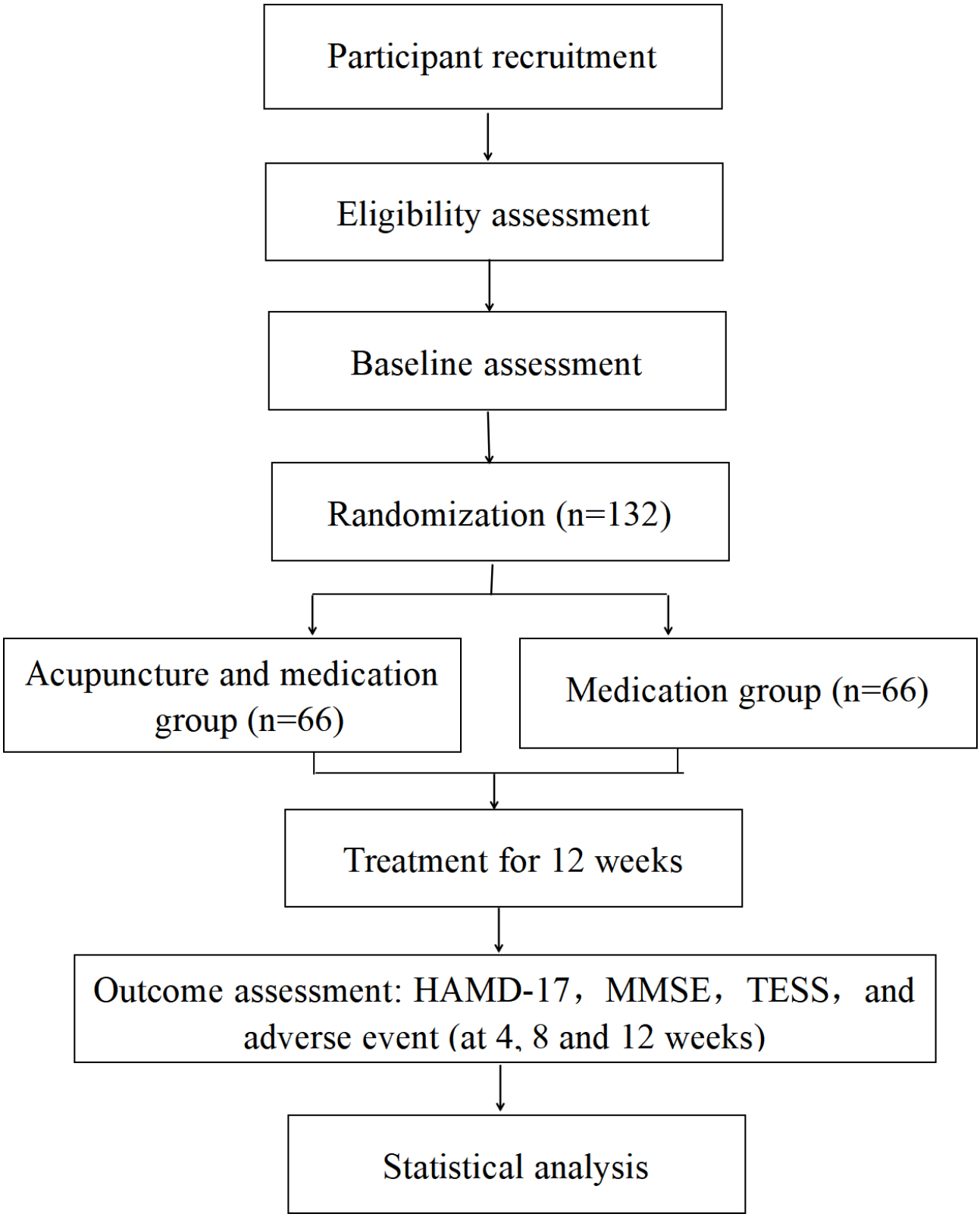

XSL•FO
RenderX

Table . Schedule of recruitment, interventions, and outcome measurements.

	Treatment period			
	Baseline	Week 4	Week 8	Week 12
Enrollment				
Screening	✓			
Signing of informed consent	✓			
Randomization	✓			
Intervention				
Acupuncture and medication group				
Medication group				
Outcome measurements				
Primary outcome				
HAMD-17[a]	✓	✓	✓	✓
Secondary outcomes				
MMSE[b]	✓	✓	✓	✓
TESS[c]	✓	✓	✓	✓
Routine blood test, routine urine test, routine stool test, liver function test, renal function test, and an ECG[d]	✓			✓
Adverse events		✓	✓	✓

[a]HAMD17: 17-item Hamilton Depression Scale.

[b]MMSE: Mini-Mental State Examination.

[c]TESS: Treatment Emergent Symptom Scale.

[d]ECG: electrocardiogram.

Study Setting and Recruitment

A total of 132 participants were recruited through posters and hospital websites, and they were allocated to 2 parallel treatment groups using a 1:1 allocation ratio. Participants were assessed at the following time points: baseline, the middle of the treatment (4 and 8 weeks after treatment starts), and the end of the treatment (12 weeks after treatment starts).

Inclusion Criteria

The inclusion criteria were as follows: (1) participants who met the criteria for major depressive disorder according to the *International Classification of Diseases, Tenth Revision*; (2) men or women aged 60 to 80 years; (3) a 17-item Hamilton Depression Rating Scale (HAMD-17) score of 7 to 24; (4) participants who were consistently not using antidepressants or had already stopped taking antidepressants for more than 4 weeks; (5) participants who did not receive acupuncture treatment in the last 4 weeks and did not participate in any other ongoing research; and (6) participants who provided informed consent.

Exclusion Criteria

Participants who reported any of the following conditions were excluded: (1) fundamental communication problems, (2) high risk of suicide or self-injurious behavior, (3) dementia, (4) a history of bipolar disorder, (5) a history of schizophrenia or psychotic symptoms, (6) current or lifetime alcohol abuse, and (7) severe cardiovascular diseases or poor liver or kidney function.

Randomization and Blinding

Eligible participants were randomly assigned to either the acupuncture and medication group or the medication group in a 1:1 ratio. To ensure equal distribution between treatment groups, the random block size was set to a fixed size of 6. An independent statistician generated a block randomization schedule using SAS (version 9.4; SAS Institute Inc). The randomization scheme, protected by strict viewing permission, is kept as a blind code by a member of staff who is not involved in this study. Baseline assessors were responsible for randomization. They obtained a random number and group assignment over the telephone. Patients and acupuncturists were not blinded. Outcome assessors and statisticians who perform

XSL•FO
RenderX

the statistical analyses are blinded to group assignment. Participants' allocated intervention will not be revealed until the statistical analysis is completed.

Interventions

Acupuncture and Medication Group

The selection of acupoints was based on our previous study on major depression, clinical experience, and expert opinion. Treatment was performed by licensed acupuncturists who had at least 5 years of experience in acupuncture. All acupuncturists were trained on how to locate acupoints, puncture, and manipulate needles before the trial. Sterile disposable acupuncture needles (length: 25 - 40 mm, diameter: 0.25 mm; Hua Tuo) will be used. Acupuncture treatments consisted of 36 sessions with a 30-minute duration over 12 weeks (3 sessions per week, ideally every other day). Acupuncture was discontinued if patients experienced any adverse events (AEs).

Participants allocated to the acupuncture and medication group received treatment using needles inserted at the prespecified acupuncture points. The protocol, including obligatory and additional acupoints, was developed from the clinical experience of acupuncture experts. The obligatory acupoints included bilateral *feishu* (BL13), *xinshu* (BL15), *ganshu* (BL18), *pishu* (BL23), *shenshu* (BL35), and *geshu* (BL17). According to different symptoms, additional acupoints could be chosen individually: *fengchi* (G20) and *taichong* (LR3) for dizziness, *shenmen* (TF4) and *sishencong* (Ex-HN1) for sleep disorders, *touwei* (SRT8) and *shuaigu* (GB8) for headache, *zusanli* (ST36) and *sanyinjiao* (SP6) for lassitude, *hegu* (LI4) and *daling* (PC7) for anxiety, *taibai* (SP3) and *gongsun* (SP4) for inappetence, and *neiguan* (PC6) and *danzhong* (RN17) for heart palpitations and chest tightness. All acupoints were localized according to the World Health Organization Standard Acupuncture Locations [18] (Table 2 and Figure 2). Manipulations of twirling, lifting, and thrusting were performed on all needles to reach *deqi* (a composite sensation including soreness, numbness, distention, and heaviness), which is believed to be an essential component of acupuncture efficacy.

XSL•FO
RenderX

Figure 2. Acupoints selected for use in the study were localized according to the World Health Organization Standard Acupuncture Locations.

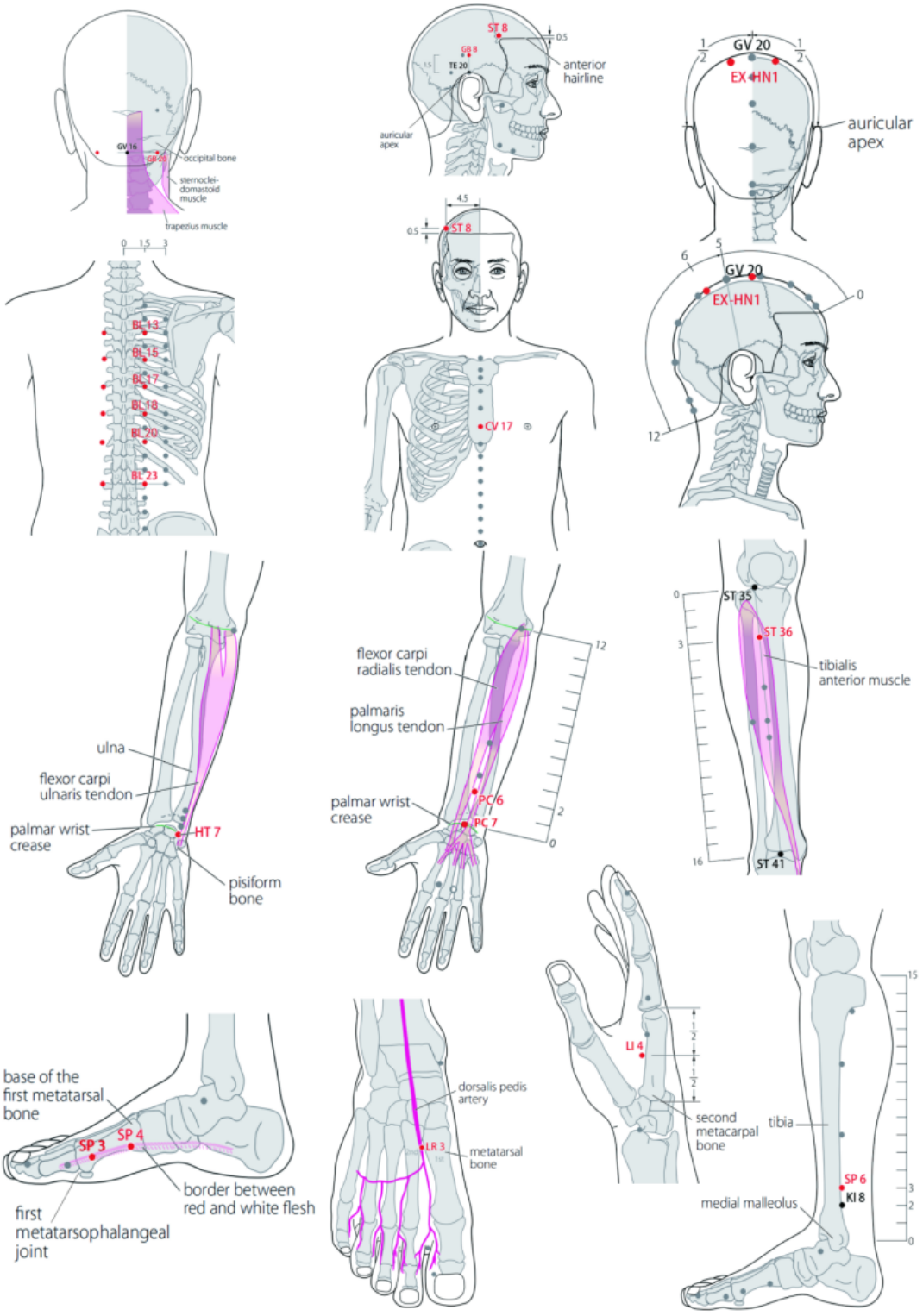

XSL•FO
RenderX

Table . Acupoints selected for use in the study.

Acupoints	Location
Ganshu (BL18)	In the upper back region, at the same level as the inferior border of the spinous process of the ninth thoracic vertebra, 1.5 B-*cun* lateral to the posterior median line.
Shenshu (BL23)	In the lumbar region, at the same level as the inferior border of the spinous process of the second lumbar vertebra, 1.5 B-*cun* lateral to the posterior median line.
Pishu (BL20)	In the upper back region, at the same level as the inferior border of the spinous process of the 11th thoracic vertebra, 1.5 B-*cun* lateral to the posterior median line.
Xinshu (BL15)	In the upper back region, at the same level as the inferior border of the spinous process of the fifth thoracic vertebra, 1.5 B-*cun* lateral to the posterior median line.
Feishu (BL13)	In the upper back region, at the same level as the inferior border of the spinous process of the third thoracic vertebra, 1.5 B-*cun* lateral to the posterior median line.
Geshu (BL17)	In the upper back region, at the same level as the inferior border of the spinous process of the seventh thoracic vertebra, 1.5 B-*cun* lateral to the posterior median line.
Fengchi (G20)	In the anterior region of the neck, inferior to the occipital bone, in the depression between the origins of the sternocleidomastoid and the trapezius muscles.
Taichong (LR3)	On the dorsum of the foot, between the first and second metatarsal bones, in the depression distal to the junction of the bases of the 2 bones, over the dorsalis pedis artery.
Shenmen (TF4)	On the anteromedial aspect of the wrist, radial to the flexor carpi ulnaris tendon, on the palmar wrist crease.
Sishencong (Ex-HN1)	On the parietal region, 1 *cun* anterior, posterior, and lateral to *baihui*; 4 acupoints.
Touwei (SRT8)	On the head, 0.5 B-*cun* directly superior to the anterior hairline at the corner of the forehead, 4.5 B-*cun* lateral to the anterior median line.
Shuaigu (GB8)	On the head, directly superior to the auricular apex, 1.5 B-*cun* superior to the temporal hairline.
Zusanli (ST36)	On the anterior aspect of the leg, on the line connecting ST35 with ST41, 3 B-*cun* inferior to ST35.
Sanyinjiao (SP6)	On the tibial aspect of the leg, posterior to the medial border of the tibia, 3 B-*cun* superior to the prominence of the medial malleolus.
Hegu (LI4)	On the dorsum of the hand, radial to the midpoint of the second metacarpal bone.
Daling (PC7)	On the anterior aspect of the wrist, between the tendons of palmaris longus and the flexor carpi radialis, on the palmar wrist crease.
Taibai (SP3)	On the medial aspect of the foot, in the depression proximal to the first metatarsophalangeal joint, at the border between the red and white flesh.
Gongsun (SP4)	On the medial aspect of the foot, anteroinferior to the base of the first metatarsal bone, at the border between the red and white flesh.
Neiguan (PC6)	On the anterior aspect of the forearm, between the tendons of the palmaris longus and the flexor carpi radialis, 2 B-*cun* proximal to the palmar wrist crease.
Danzhong (RN17)	In the anterior thoracic region, at the same level as the fourth intercostal space, on the anterior median line.

Participants took citalopram hydrobromide up to 20 mg daily. Patients' usual treating clinicians adjusted doses based on their routine practices.

Medication Group

Participants in this group took the same medication at the same dose and schedule as the acupuncture and medication group. Additional treatment for depression was not provided.

XSL•FO
RenderX

Outcome Measurements

Follow-up examinations were performed at baseline and at 4, 8, and 12 weeks after treatment initiation in the 2 groups.

Primary Outcome

The primary outcome measurement of this study is the clinical response rate according to the total HAMD-17 score at the end of the study. Response is defined as a 50% or greater decrease from baseline in the total HAMD-17 score at the end of the study (week 12). The HAMD-17 is a multiple-item questionnaire used to assess the severity of depression and as a guide to evaluate recovery. The range of the HAMD-17 score is 0 to 52, with 0 to 6 indicating normal depression, 7 to 17 indicating mild depression, 18 to 24 indicating moderate depression, and 25 or more indicating severe depression [19].

Secondary Outcomes

Hamilton Depression Scale

Changes in the total HAMD-17 score from baseline are assessed.

Mini-Mental State Examination

The Mini-Mental State Examination (MMSE) is used to screen for possible cognitive impairment. The MMSE contains 30 items that assess a range of cognitive functions, such as orientation, short-term memory, language, comprehension, attention, and calculation. Total scores on the MMSE range from 0 to 30. Participants with scores of 23 or below are considered likely to have a cognitive impairment [20].

Treatment Emergent Symptom Scale

The Treatment Emergent Symptom Scale is commonly used to measure the presence and intensity of psychotropic medication side effects [21,22]. Both intensity (from 0 to 4, with a higher score correlating with a worse condition) and intervention (from 0 to 6) are documented in the scale.

Safety Assessment

A routine blood test, routine urine test, routine stool test, liver function test, renal function test, and an electrocardiogram were administered for safety outcomes. These biological indicators were monitored during the screening period and after 12 weeks of treatment.

The most common AEs related to acupuncture treatment include localized hematoma, localized infection, broken needles, fainting, bleeding, nausea, dizziness, vomiting, palpitations, and severe pain. Detailed information on AEs and serious AEs (SAEs), including the name, onset and end date, intensity, relationship with acupuncture, and outcome, is recorded. If the treatments cause SAEs, researchers will immediately report them to the medical ethics committee of Beijing Traditional Chinese Medicine Hospital and suspend the study.

Sample Size Calculation

The sample size was calculated based on a noninferiority design using PASS (version 15.0; NCSS, LLC). The primary objective is to demonstrate that acupuncture combined with medication is not inferior to medication alone. We conservatively assumed a baseline rate of 30% for the control group (medication alone). This assumption aligns with large-scale clinical trials such as the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study [23], which reported a remission rate of approximately 30% to 33% for patients treated with citalopram. For the intervention group, based on our pilot data and clinical experience, we anticipated a response rate of 40%. It is important to note that while this anticipation was used for power calculation to ensure adequate sensitivity, the hypothesis testing remains focused on noninferiority. The noninferiority margin (Δ) was set at −15%. This margin was chosen based on the risk-benefit profile: given that acupuncture is a physical therapy with a superior safety profile compared to pharmacotherapy, a wider margin is clinically acceptable. With a 1-sided significance level (α) of .025 and 80% power, a sample size of 57 patients per group was required. Assuming a dropout rate of 15%, we aimed to recruit a total of 132 participants (66 per group).

Statistical Analysis

The statistical analysis is performed by an independent statistician who is not aware of the group allocation. SAS (version 9.4) is used for data analysis. The level of significance is set at α <.05 with a 2-sided test. Continuous data are represented as the mean (SD) or median (IQR), whereas categorical data are represented as percentages.

All efficacy analyses are performed using the intent-to-treat approach. For the intent-to-treat analysis, the population consists of all patients who have been randomized. Missing data are imputed using the multiple imputation method. Continuous variables are compared using the Student *t* test or Wilcoxon rank sum test, as appropriate. Categorical variables are compared using the Fisher exact test or the Wilcoxon rank sum test, as appropriate.

Our primary analysis for between-group differences uses a repeated measures ANOVA (or a mixed effects model) to account for the multiple time points (weeks 0, 4, 8, and 12), with “group” as the between-subject factor and “time” as the within-subject factor. This model appropriately handles the longitudinal nature of the data. We have also specified that the primary end point is the difference at week 12.

A sensitivity analysis is being conducted for the primary outcome using the per-protocol population, including only those patients who complete at least 80% of allocated treatments and have no major protocol violations (taking other drugs during the trial or not completing the case report form as required).

Data Monitoring and Quality Control

All researchers, including acupuncturists, outcome assessors, and statisticians, receive training regarding data management. Monitors check case report forms once every week as well as the acupuncture operation during the treatment period. Dropouts and withdrawals, including the reasons, are documented in detail throughout the trial. Participants’ information is stored in locked file cabinets at the study sites with limited access; only investigators have the right to access the data.

XSL•FO
RenderX

Data Availability Statement

All data will be preserved for at least 5 years after publication. The principal investigator will have access to the final trial dataset, and trial data will be made available to readers upon reasonable request by contacting the corresponding author. Information on patients will remain anonymous, including name, age, and telephone number.

Patient and Public Involvement Statement

Patients were not involved in the design, conduct, or reporting of this trial. They will be evaluated for their ability to complete the study treatment based on their time and the degree of mobility difficulty. Specifically, we consulted patients before the trial regarding the burden of 36 sessions. On the basis of their feedback, we implemented the following measures: (1) flexible scheduling—avoiding early morning sessions and allowing makeup visits, (2) caregiver engagement—actively involving family members to support adherence and (3) outcome relevance—confirming that our chosen outcomes aligned with patients' daily living priorities. A journal article manuscript will be prepared to present the results after the trial is completed, and a brief summary of the results in plain language will be provided to all participants. No attempt will be made to assess the burden of the intervention on the patients themselves.

Ethical Considerations

The trial protocol (version 2.0; September 19, 2022) was planned in accordance with the Declaration of Helsinki and has been approved by the ethics committee of Beijing Traditional Chinese Medicine Hospital Affiliated to Capital Medical University (2022BL02-061-01). The trial was registered in June 2023 at the Chinese Clinical Trial Registry (ChiCTR2300072740). Before each participant enrolled in this study, the investigator provided a complete and comprehensive written introduction to the purpose, procedures, and potential risks of the study. All the participants were fully informed about this trial and given enough time to inquire about details and decide whether to participate at the first visit. Participants were asked to sign the informed consent form if they agreed to participate. Participants were informed of their right to withdraw from the study at any time. If a participant withdrew during the course of the trial, we reserved the right to retain the data obtained for statistical analysis. All data are anonymized. All participants were entitled to take free citalopram hydrobromide (up to 20 mg) daily for 12 weeks, and participants in the acupuncture and medication group were entitled to 36 free acupuncture treatments. No other compensation, such as for transportation or lost wages, was provided. Trial results will be published in peer-reviewed journals and will be disseminated to the media and the general public.

Results

The study commenced on June 30, 2023, and as of October 17, 2024, a total of 132 participants had been enrolled. Data collection has been completed. Currently, data analysis is in progress, with preliminary findings anticipated to be available by October 2025. The findings of this study are expected to be submitted for publication in 2026.

Discussion

Anticipated Findings

This is an RCT designed to investigate the efficacy of acupuncture for mild to moderate depression in older people over a 12-week period. The trial is designed to meet the methodological demands of adequate power and allocation concealment, following the Good Clinical Practice guideline [24]. This study is defined as a "pilot" or "preliminary" trial because it is the first step in our research program to evaluate the specific integration of acupuncture with citalopram for the older population. Its main purposes are to generate effect size estimates for sample size calculation of future multicenter trials and to test the feasibility of the protocol (including recruitment and retention) in a single center. The results of this study may provide further evidence for the effectiveness of acupuncture in promoting depressive emotion and improving cognitive function in older people with mild to moderate depression.

Acupuncture has been used to treat mental disorders since antiquity in China. Results from animal experiments suggest a multitarget antidepressant effect of acupuncture, which may be related to amino acid metabolism and inflammatory pathways, especially the toll-like receptor signaling pathway, tumor necrosis factor signaling pathway, and nuclear factor κ light chain enhancer of activated B cells signaling pathway [25]. The literature has emphasized that acupuncture combined with antidepressant pharmacological treatment not only enhances the improvement of primary and secondary depressive symptoms but also reduces the side effects of medical treatment, which is the main cause of high dropout rates for drug treatment [26]. Emotional and cognitive disorders are common psychological features underlying the depressive symptoms of depression. It has been found that those with depression have difficulty shifting attention from negative information, which leads to persistent negative emotional rumination. Depression in older adults is frequently comorbid with cognitive decline ("pseudodementia"), and improvement in depressive symptoms often correlates with cognitive recovery. The existing studies on depression are mainly based on responses to clinical scales to assess depressive symptoms but have not focused on the effects of acupuncture on cognition in those with depression. However, the mechanisms of how acupuncture regulates attentional bias in those with depression are still not clear. Therefore, this trial is designed to explore the clinical efficacy and safety of acupuncture therapy for mild to moderate depression in older people. Moreover, we will use indices reflective of cognitive function, such as the MMSE, as secondary outcomes, with the expectation that they will provide reliable clinical evidence for acupuncture treatment that helps improve depression, which is a key concern for the older population.

However, there are several limitations that need to be acknowledged. First, the single-center experimental design will result in a single sample, with limited representativeness, and possible experimental bias. More influencing factors should be considered, and the result should be further verified and explored in a large sample population. Second, blinding of acupuncturists is not possible due to the unique nature of acupuncture

XSL•FO
RenderX

procedures. To reduce the impact, we provided detailed training to the acupuncturists, including standardized communication with patients. Third, the decision not to use a sham control is supported by public concerns on the methodology of sham or placebo acupuncture design, as well as the substantial risk of failure in blinding of participants, who are usually very experienced with acupuncture treatment in Chinese medical culture and settings. Finally, without a follow-up period, it will be impossible to assess how long the therapeutic effect of acupuncture lasts. However, as a preliminary trial, our primary goal is to establish immediate efficacy and feasibility at the end of the 12-week treatment. This phased approach aligns with standard research frameworks for complex interventions [27]. Previous pilot studies [28] have similarly focused on end-of-treatment outcomes to validate protocols before extending to longitudinal follow-up. Future confirmatory trials will include a 3-month follow-up.

Conclusions

This pilot study is expected to provide critical insights into the feasibility of integrating acupuncture with standard medication for managing mild to moderate depression in older people. By generating preliminary evidence on its potential benefits, the study aims to inform the design and sample size estimation of future multicenter trials, potentially advancing nonpharmacological treatment options for depression.

Acknowledgments

The authors thank all participants for their involvement in the study, as well as the principal investigators and other investigators who contributed to the study through administrative, technical, or material support; patient recruitment; and data collection. The authors also declare that no generative AI technologies were used in the writing or preparation of this manuscript.

Funding

This randomized controlled trial was funded by the Beijing Municipal Administration of Hospitals Incubating Program (PZ2022011). The funding source had no role in the study design; data collection, management, analysis, or interpretation; manuscript preparation; or the decision to submit the study for publication.

Authors' Contributions

Conceptualization: QNF (lead), WL (equal), JZ and JY.
Data curation: ZQM.
Formal analysis: QNF (lead), KHX (supporting).
Funding acquisition: QNF.
Investigation: JZ.
Methodology: WL.
Project administration: QNF (lead), WL (equal), JZ (supporting).
Resources: QNF.
Supervision: JZ.
Validation: WL, JZ.
Visualization: QNF (lead), KHX (supporting).
Writing—original draft: QNF (lead), QNF (supporting).
Writing—review and editing: QNF (lead), WL (supporting), JZ (supporting).

Conflicts of Interest

None declared.

References

1. Ren X, Yu S, Dong W, Yin P, Xu X, Zhou M. Burden of depression in China, 1990-2017: findings from the Global Burden of Disease Study 2017. J Affect Disord 2020 May 1;268:95-101. [doi: 10.1016/j.jad.2020.03.011] [Medline: 32158012]
2. Battle DE. Diagnostic and Statistical Manual of Mental Disorders (DSM). Codas 2013;25(2):191-192. [doi: 10.1590/s2317-17822013000200017] [Medline: 24413388]
3. Ménard C, Hodes GE, Russo SJ. Pathogenesis of depression: insights from human and rodent studies. Neuroscience 2016 May 3;321:138-162. [doi: 10.1016/j.neuroscience.2015.05.053] [Medline: 26037806]
4. Zhang Y, Chen Y, Ma L. Depression and cardiovascular disease in elderly: current understanding. J Clin Neurosci 2018 Jan;47:1-5. [doi: 10.1016/j.jocn.2017.09.022] [Medline: 29066229]
5. Bener A, Dafeeah EE, Abou-Saleh MT, Bhugra D, Ventriglio A. Co-morbidity between major depression and schizophrenia: prevalence and clinical characteristics. Psychiatr Danub 2020;32(1):78-83. [doi: 10.24869/psyd.2020.78] [Medline: 32303036]
6. Penner-Goeke S, Binder EB. Epigenetics and depression. Dialogues Clin Neurosci 2019 Dec;21(4):397-405. [doi: 10.31887/DCNS.2019.21.4/ebinder] [Medline: 31949407]

XSL•FO
RenderX

7. Lu J, Xu X, Huang Y, et al. Prevalence of depressive disorders and treatment in China: a cross-sectional epidemiological study. Lancet Psychiatry 2021 Nov;8(11):981-990. [doi: 10.1016/S2215-0366(21)00251-0] [Medline: 34559991]
8. García Herrera JM, Nogueras Morillas V, Muñof Cobos F, Morales Asencio JM. Guía de práctica clínica para el tratamiento de la depresión en atención primaria. : Servicio Andaluz de Salud; 2011 URL: https://www.sspa.juntadeandalucia.es/servicioandaluzdesalud/sites/default/files/sincfiles/wsas-media-pdf_publicacion/2020/GPC_488_Depresion_AP_compl.pdf [accessed 2025-12-16]
9. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet 2009 Feb 28;373(9665):746-758. [doi: 10.1016/S0140-6736(09)60046-5] [Medline: 19185342]
10. Gartlehner G, Hansen RA, Morgan LC, Thaler K, Lux LJ, Van Noord M. Second-generation antidepressants in the pharmacologic treatment of adult depression: an update of the 2007 comparative effectiveness review [internet]. : Agency for Healthcare Research and Quality (US); 2011 URL: https://pubmed.ncbi.nlm.nih.gov/22299185/ [accessed 2025-12-16]
11. Särkämö T, Sihvonen AJ. Golden oldies and silver brains: deficits, preservation, learning, and rehabilitation effects of music in ageing-related neurological disorders. Cortex 2018 Dec;109:104-123. [doi: 10.1016/j.cortex.2018.08.034] [Medline: 30312779]
12. Wang Z, Wang X, Liu J, et al. Acupuncture treatment modulates the corticostriatal reward circuitry in major depressive disorder. J Psychiatr Res 2017 Jan;84:18-26. [doi: 10.1016/j.jpsychires.2016.09.014] [Medline: 27693978]
13. Chan YY, Lo WY, Yang SN, Chen YH, Lin JG. The benefit of combined acupuncture and antidepressant medication for depression: a systematic review and meta-analysis. J Affect Disord 2015 May 1;176:106-117. [doi: 10.1016/j.jad.2015.01.048] [Medline: 25704563]
14. Smith CA, Armour M, Lee MS, Wang LQ, Hay PJ. Acupuncture for depression. Cochrane Database Syst Rev 2018 Mar 4;3(3):CD004046. [doi: 10.1002/14651858.CD004046.pub4] [Medline: 29502347]
15. Zhang ZJ, Chen HY, Yip KC, Ng R, Wong VT. The effectiveness and safety of acupuncture therapy in depressive disorders: systematic review and meta-analysis. J Affect Disord 2010 Jul;124(1-2):9-21. [doi: 10.1016/j.jad.2009.07.005] [Medline: 19632725]
16. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346:e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
17. MacPherson H, Altman DG, Hammerschlag R, et al. Revised STandards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement. J Altern Complement Med 2010 Oct;16(10):ST1-ST14. [doi: 10.1089/acm.2010.1610] [Medline: 20954957]
18. Qiu T, Li L. Discussion on the Chinese edition of the WHO standard acupuncture point locations in the Western Pacific Region. Zhongguo Zhen Jiu 2011 Sep;31(9):827-830. [Medline: 21972638]
19. Hamilton M. The Hamilton Rating Scale for Depression. In: Sartorius N, Ban TA, editors. Assessment of Depression: Springer; 1986.
20. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [doi: 10.1016/0022-3956(75)90026-6] [Medline: 1202204]
21. Shan X, Zhao W, Qiu Y, et al. Preliminary clinical investigation of combinatorial pharmacogenomic testing for the optimized treatment of depression: a randomized single-blind study. Front Neurosci 2019 Sep 13;13:960. [doi: 10.3389/fnins.2019.00960] [Medline: 31572113]
22. Zheng W, Xiang YQ, Ng CH, Ungvari GS, Chiu HF, Xiang YT. Extract of Ginkgo biloba for tardive dyskinesia: meta-analysis of randomized controlled trials. Pharmacopsychiatry 2016 May;49(3):107-111. [doi: 10.1055/s-0042-102884] [Medline: 26979525]
23. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry 2006 Jan;163(1):28-40. [doi: 10.1176/appi.ajp.163.1.28] [Medline: 16390886]
24. Grimes DA, Hubacher D, Nanda K, Schulz KF, Moher D, Altman DG. The Good Clinical Practice guideline: a bronze standard for clinical research. Lancet 2005;366(9480):172-174. [doi: 10.1016/S0140-6736(05)66875-4] [Medline: 16005342]
25. Wang Y, Jiang H, Meng H, et al. Antidepressant mechanism research of acupuncture: insights from a genome-wide transcriptome analysis of frontal cortex in rats with chronic restraint stress. Evid Based Complement Alternat Med 2017;2017:1676808. [doi: 10.1155/2017/1676808] [Medline: 29098013]
26. Yang NN, Lin LL, Li YJ, et al. Potential mechanisms and clinical effectiveness of acupuncture in depression. Curr Neuropharmacol 2022;20(4):738-750. [doi: 10.2174/1570159X19666210609162809] [Medline: 35168522]
27. Mischoulon D, Brill CD, Ameral VE, Fava M, Yeung AS. A pilot study of acupuncture monotherapy in patients with major depressive disorder. J Affect Disord 2012 Dec 10;141(2-3):469-473. [doi: 10.1016/j.jad.2012.03.023] [Medline: 22521855]
28. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655. [doi: 10.1136/bmj.a1655] [Medline: 18824488]

XSL•FO
RenderX

Abbreviations

AE: adverse event
HAMD-17: 17-item Hamilton Depression Scale
MMSE: Mini-Mental State Examination
RCT: randomized controlled trial
SAE: serious adverse event
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SSRI: selective serotonin reuptake inhibitor
STRICTA: Standards for Reporting Interventions in Clinical Trials of Acupuncture

Edited by J Sarvestan; submitted 23.Jun.2025; peer-reviewed by JF Tu; revised version received 07.Dec.2025; accepted 08.Dec.2025; published 30.Jan.2026.

Please cite as:
Fu Q, Xiao K, Zhang J, Li Y, Wang Y, Jiang M, Man Z, Yang J, Lu W
Efficacy of Acupuncture for Mild to Moderate Depression in Older People: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e79327
URL: https://www.researchprotocols.org/2026/1/e79327
doi:10.2196/79327

© Qingnan Fu, Kaihui Xiao, Jie Zhang, Yang Li, Yuxian Wang, Molin Jiang, Zengqi Man, Jing Yang, Wei Lu. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Enhancing Upper Secondary Students' Situational Engagement and Cognitive Prerequisites of Learning Through the Physically Active Academic Lessons Intervention: Protocol for a Mixed Methods Cluster Randomized Individual Crossover Trial

Heidi J Syväoja[1], PhD; Susanna Takalo[2], PhD; Tuomas Kukko[1], MSc; Nina Salmela[3,4], MA; Harto Hakonen[1], MSc; Janne Kulmala[1], MSc; Heidi Lindfors[1], MSc; Hermanni Oksanen[1], MSc; Pekka Räsänen[5,6], PsyLic; Kati Mäkitalo[3], PhD; Tuija H Tammelin[1], PhD

[1]Likes, School of Health and Social Studies, Jamk University of Applied Sciences, Piippukatu 2, Jyväskylä, Central Finland, Finland

[2]Early Childhood and Music Education Studies (CHIMES), Faculty of Education and Psychology, University of Oulu, Oulu, Finland

[3]Learning and Learning Processes Research Unit, Faculty of Education and Psychology, University of Oulu, Oulu, Finland

[4]City of Oulu, Oulu, Finland

[5]Epilepsia Helsinki, Department of Pediatric Neurology, HUS Helsinki University Hospital, Helsinki, Finland

[6]Turku Research Institute for Learning Analytics, Faculty of Science, University of Turku, Turku, Finland

Corresponding Author:
Heidi J Syväoja, PhD
Likes, School of Health and Social Studies, Jamk University of Applied Sciences, Piippukatu 2, Jyväskylä, Central Finland, Finland

Abstract

Background: Internationally, physical activity is successfully integrated into academic lessons in primary schools, showing promising results on cognition and student engagement. However, there is a lack of knowledge about its effects and feasibility for individual situated learning processes in upper secondary school.

Objective: This protocol describes the design and methods of the Physically Active Academic Lessons (PAAL) study, a mixed methods, cluster randomized, individual crossover trial. The PAAL study aims to examine the acute effects of physically active academic lessons on cognitive prerequisites of learning (alertness and executive functions) and situational engagement (behavioral, cognitive, and emotional engagement; disaffection; competence experiences; and help seeking), as well as factors modifying these effects (physical and mental load and perceived physical and academic competence). Further, subject teachers' and students' experiences and perceptions of physically active academic lessons in general upper secondary school will be explored.

Methods: The first part of the PAAL study involves exploring subject teachers' experiences of facilitators, barriers, usefulness, and the meaning of physically active academic lessons for the situational learning process through semistructured interviews with 14 teachers. The second part consists of a cluster-randomized individual crossover trial including 168 students in mathematics and foreign language lessons, followed by interviews with 30 students.

Results: Funding for the study was obtained in May 2023. Ethical approval for the teacher interviews was granted in September 2023, and for the student trial in December 2023. Data collection was completed between October 2023 and November 2024. Data analysis is ongoing. The findings of the study will provide essential evidence-based information on physically active classroom practices that support teachers and schools in implementing pedagogical methods that enhance student learning and well-being in upper secondary schools.

Conclusions: The background, design, content of the intervention, and methods of the PAAL study are presented. This study aims to address a gap in the literature regarding the feasibility and effectiveness of physically active methods during academic lessons in upper secondary school.

Trial Registration: ISRCTN 63809854; http://www.controlled-trials.com/ISRCTN63809854/63809854

International Registered Report Identifier (IRRID): RR1-10.2196/84601

(JMIR Res Protoc 2026;15:e84601) doi:10.2196/84601

XSL•FO
RenderX

KEYWORDS

physically active learning; physically active break; student engagement; executive functions; acute effects; experience; cluster randomized individual crossover trial

Introduction

Background

In a well-functioning educational environment, young people are not only performing academically but are also satisfied with their lives and feeling good physically and mentally [1]. However, young people's mental health disorders have been predicted to be a leading cause of disability in high-income countries [2]. For example, in the United States, rates of major depressive episodes increased by 52% from 2005 to 2017 among adolescents aged 12 to 17 years and by 63% from 2009 to 2017 among young adults aged 18 to 25 years [3]. Likewise, in Finland, a recently published student barometer shows a decline in upper secondary school student well-being, with up to 62% finding studying is mentally taxing [4]. School burnout can, at worst, lead to academic failure, school dropout, and other mental health issues [5,6]. In addition to mental well-being, the physical health and functional capacity of young people are also a concern, as globally about 80% of students aged 11 to 17 years do not engage in sufficient physical activity (PA) for their health [7]. In Finland, general upper secondary students aged 16 to 18 years remain sedentary for 9 to 10 hours per day [8], most of which occurs in educational institutions. Prolonged sedentariness is harmful to both physical and mental health and may interfere with learning. In contrast, increased PA positively impacts physical health, strengthens mental health, and promotes brain health, cognitive function, and academic performance [9].

Current research supports increasing PA during the school day, including during academic lessons in children [10-12]. In school-based interventions, PA has been implemented into academic lessons in various ways, such as through PA breaks and physically active learning (PAL) [12,13]. PA breaks are intended to interrupt prolonged sedentary behavior by increasing PA [14,15] and to create optimal conditions for learning after the break by enhancing cognition, such as attention and executive functions [16], as well as classroom behavior [12,14,15] and learning outcomes [16]. Executive functions, a group of higher-order cognitive processes involved in self-regulatory and goal-directed actions, include key components such as inhibition (eg, the suppression of inappropriate or distracting responses), working memory, and cognitive flexibility [17]. Executive functions are strong predictors of academic performance [18].

In addition to these effects, PAL has been shown to engage students in learning by increasing their enjoyment of the learning process [10,11] and enhancing memory recall by constructing rich and elaborative representations [19]. PAL increases PA without reducing active academic learning time [20], as PA is integrated into the academic learning content of subjects other than physical education [21]. Student engagement is one of the key learning prerequisites in terms of students' involvement and commitment to school and their own learning [22,23], and can be seen as consisting of several components, such as behavioral (eg, student conduct and on-task behavior), emotional (eg, students' emotions felt towards learning, teachers, and classmates), and cognitive engagement (eg, students' motivational goals and self-regulated learning) [22,23], as well as social engagement [24]. Student engagement is context-dependent and can be promoted through internal and external changes in the context and environment [22,25].

Prior Work

Preliminary findings suggest that the effects of acute PA on older adolescents' and young adults' cognitive functions [26,27] and student engagement [28,29] are similar to those in childhood. However, due to limited repeated studies, inconsistency, and low methodological quality of available evidence, the benefits of PA among older adolescents have not been confirmed in an academic context [14,30,31]. More evidence-based information is needed on how physically active approaches during lessons acutely affect the prerequisites of learning compared with typical academic lessons without PA. Furthermore, not all young people benefit from PA in the same way [31]. While late adolescence is a vulnerable period for coping with mental, physical, and academic demands [32,33], mental load, including stress and anxiety, has been shown to impair executive functions [34,35] and efforts to be physically active [36]. Furthermore, perceived physical and academic competence predict participation in PA [37,38] and learning outcomes [39], while previously acquired skills, such as motor competence, may modify the effects of PAL [40]. Therefore, physical and mental load, as well as physical and academic competence, may influence the effectiveness of physically active pedagogical approaches. Further research is needed to clarify the role of potential moderating factors in the outcomes of physically active academic lessons.

Teachers across all subjects play a crucial role in implementing PAL and PA breaks. However, many teachers are not trained to teach through PA and often possess limited knowledge and skills regarding physically active classroom practices [41]. Previous research suggests that teachers interpret PAL through the lens of their professional identity and established teaching practices [42]. Therefore, it is important to understand subject teachers' long-term experiences of implementing PAL in authentic academic lessons. Most existing PAL research has focused on the perspectives of subject teachers working with pupils under the age of 16 years [41,43]. There is currently a lack of research on general upper secondary school teachers' experiences regarding the implementation of PAL and PA breaks.

Learning is context dependent, yet only a limited number of studies have examined situational variation in student engagement. Research in lower and upper secondary education indicates that students respond to contextual changes. Therefore, it is important to explore students' personal observations and experiences in upper secondary lessons [44] to gain a deeper understanding of how to design effective and sustainable

XSL•FO
RenderX

interventions [45] that influence their behavior and engagement in certain situations. Elementary school students have found physically active breaks and PAL to be enjoyable and beneficial for learning, with breaks offering a refreshing pause from schoolwork and PAL creating engaging and meaningful learning experiences in primary and lower secondary education among children aged 7 to 15 years [46-49]. Recent studies by Fred et al [50] and Romar et al [51] are among the few that have explored upper secondary students' experiences of PA breaks and PAL, which students report as a welcome change from sitting, enhancing concentration, supporting learning, and making lessons more enjoyable—thereby promoting engagement. However, because most evidence on classroom-based PA interventions comes from primary school settings, and the few upper secondary studies involve small samples with limited generalizability, further research is needed to amplify students' voices and explore situational variation in engagement [44,51-53].

Goal of the Study

The Physically Active Academic Lessons (PAAL) study evaluated teachers' and students' experiences and perceptions during physically active academic lessons in a general upper secondary school. Specifically, we investigated subject teachers' perceptions and experiences regarding the usefulness of physically active classroom practices implemented in different ways throughout the academic year. This included both the benefits and challenges associated with PAL and PA breaks, experiences of applying physically active classroom practices during various phases of teaching within academic lessons, and perceptions of how these methods support the achievement of learning objectives across different school subjects. In addition, we investigated students' perceptions and experiences on students' situational engagement (behavioral, emotional, cognitive, and social engagement) and their relation to physically active academic lessons, as well as their views on the usefulness of PAL and PA breaks. Moreover, we examined the acute effects of PAL and PA breaks on students' situational engagement and cognitive prerequisites of learning. We especially examined how PAL and PA breaks affect students' alertness, executive functions, behavioral, cognitive, and emotional engagement, as well as disaffection, competence experiences, and help seeking during academic lessons compared to lessons without PA. Furthermore, we examined how current physical and mental load or perceived physical and academic competence modify the acute effects of PAL and PA breaks on situational engagement and cognitive prerequisites of learning.

Methods

Study Design

This protocol (version 1) followed the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement (see the SPIRIT checklist in Checklist 1) and describes the design and methods of the PAAL study. The PAAL study consists of teacher interviews (Part 1), a cluster randomized individual crossover trial among students (Part 2), and their interviews (Part 3). The overall study design is presented in Figure 1.

Figure 1. The overall study design of the Physically Active Academic Lessons (PAAL) study. The PAAL study consists of teacher interviews (Part 1), a cluster-randomized individual crossover trial among students (Part 2), and their interviews (Part 3).

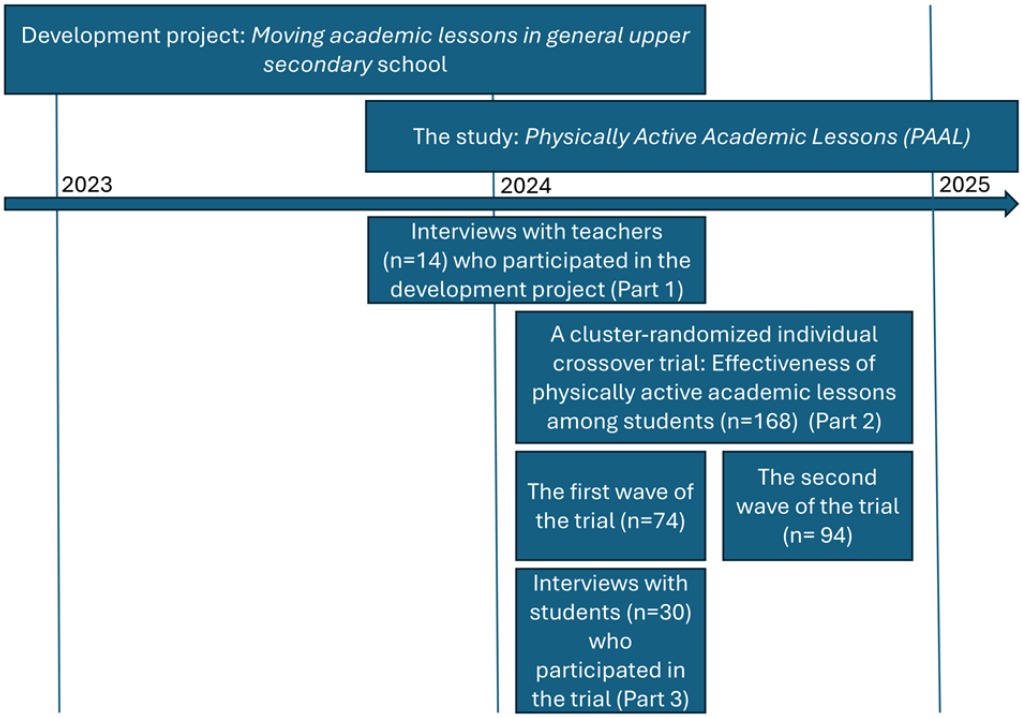

XSL•FO
RenderX

Teacher Interviews (Part 1)

Content of the Development Project Before Teacher Interviews

The evaluation of subject teachers' experiences and perceptions of physically active academic lessons in general upper secondary school was carried out in the context of the development project M*oving Academic Lessons in General Upper Secondary School* (2022 - 2024). The aim of the development project was to promote the learning, well-being, and health of general upper secondary school students by combining PA with academic lessons. This development project was funded by the Ministry of Education and Culture and implemented for general upper secondary school teachers in Finland in cooperation with a national action program, Finnish Schools on the Move.

Fourteen subject teachers (n=14) from 8 different schools accepted an invitation to participate in the development project as PAL pioneer teachers in spring 2022. Participants represented 16 different teaching subjects, and their teaching experience in general upper secondary education ranged from 2.5 to 29 years, with a mean of 12 years. The subject teachers were asked to use and create a variety of teaching methods to suit PA with their own subject lessons' topics and contents, and implement PA breaks during academic lessons. The development project was reviewed to inform the formulation of teacher interview questions (Part 1 of the PAAL study) and the planning of teaching methods for the intervention (Part 2 of the PAAL study).

Participants

Fourteen PAL pioneer teachers involved in the development project were invited to participate in the PAAL study teacher interviews conducted between October 2023 and May 2024, and all invited teachers (13 females and 1 male; ages 30 - 62 y) accepted the invitation. Participants received correct and sufficient information about the study, including its objectives and the intended use of the interview data.

Ethical Considerations

Participation was voluntary, and participants were informed of their right to withdraw from the study at any time. Written informed consent was obtained from all participants prior to the interviews. Privacy and confidentiality were assured, and no financial or other compensation was offered. All identifying details have been omitted from the published written descriptions to protect the anonymity of the participants.

Before data collection (September 2023), the need for an ethical review of this part of the study was requested from the Ethics Committee of Human Sciences of the University of Oulu. According to the statement of the Ethics Committee of Human Sciences of the University of Oulu, an ethical review is not required when research involving human participants complies with the 2019 guidelines of the Finnish National Board on Research Integrity (TENK) [54]. This study was conducted in accordance with TENK's guidelines. In addition, before teacher data collection, the study was approved by the Education and Culture Services of the municipality where the data were collected (OUKA/11121/07.01.04.02/2023) and fulfilled the ethical standards for empirical research.

Methods for Teacher Interviews

The participating subject teachers were interviewed after implementing PAL and PA breaks in their teaching for more than 12 months. The interview questions were finalized after a pilot interview. The list of final interview themes was shared with participants 2 days before the interview. Interviews were conducted at times and locations convenient for the teachers: at the participant's school, at the university, or at the interviewer's home. Teachers were asked to describe their experiences with PAL and PA breaks. The questions were organized around the main themes (see *Teacher Interviews* in Multimedia Appendix 1), which were developed through discussions with PAL pioneer teachers and reflections from the development project. In addition, similar overarching themes have been identified to summarize teachers' perspectives on implementing PAL in primary schools [55]. Although the interviews adhered to the main themes, the order of questions varied between interviews based on the flow of the conversation between the interviewer and the interviewee. The semistructured interviews were recorded and transcribed during the autumn of 2023 and the spring of 2024 in Northern Finland. Interviews lasted 63 to 130 minutes (mean 95 min) and were done individually. Together, there were 365 pages of long transcribed text. PAL and PA breaks were treated as distinct classroom activity types. The interview data were analyzed using thematic analysis [56]. An inductive approach was conducted through 6 phases [56] (see *Analysis of Teacher Interviews* in Multimedia Appendix 1) to gain insight into why teachers integrate PAL and PA breaks into academic lessons. In addition to the inductive approach, a deductive approach was used to identify teachers' perceptions related to successful PAL adoption and implementation.

Acute Effects of Physically Active Academic Lessons (Part 2)

Study Design and Participants

The effectiveness of physically active academic lessons among students was examined using a cluster randomized individual crossover trial conducted in two waves. In spring 2024, a total of 4 teachers of mathematics and foreign languages from 5 general upper secondary schools in Northern Ostrobothnia, who had participated in both the development project and the teacher interviews, were invited and agreed to implement the intervention. In autumn 2024, 2 schools from Central Finland were contacted, and a total of 5 teachers of mathematics and foreign languages from these schools expressed interest and agreed to implement the intervention. From a total of 17 eligible course groups of first- and second-year students, 168 students volunteered to participate in the study. These groups included 10 from schools in Northern Ostrobothnia (4 in mathematics and 6 in foreign languages), comprising 242 students, of whom 74 volunteered, and 7 from schools in Central Finland (4 in mathematics and 3 in foreign languages), comprising 200 students, of whom 94 volunteered. Students had the opportunity to withdraw from the study at any stage, and any data collected

XSL•FO
RenderX

up to the point of withdrawal were retained as part of the research material. Eight students dropped out of the study.

Students in the same course group formed the clusters (mean cluster size 10, ranging from 1 to 18), which were randomized to undergo the 3 different treatments (lessons) in a different order. As there are 6 possible permutations of the 3 treatments, each cluster was allocated one of these sequences using a random number generator that sampled integers uniformly from 1 to 6. This procedure ensured randomized assignment of treatment order across the groups. After the trial, 30 voluntary students were invited to student interviews. This study followed the Consolidated Standards of Reporting Trials (CONSORT) statement (2025). The CONSORT flow diagram of the trial is presented in Figure 2.

XSL•FO
RenderX

Figure 2. CONSORT flow diagram illustrating the study design of a cluster-randomized individual crossover trial among students. PA: physical activity; PAL: physically active learning.

XSL•FO
RenderX

Ethical Considerations

The study protocol was approved by the Ethics committee of Jamk University of Applied Sciences (Ethical review statement 472673, 21 December 2023). The study was prospectively registered (ISRCTN63809854, registered 11 January 2024 [57]).

Informed consent was obtained from the students and, in the case of minors, also from their guardians. This included consent for data collection as described in the information sheet and for the use and processing of personal data in accordance with the privacy notice. The consent form further included consent to being contacted after the trial for the purpose of inviting the students to a potential interview. Participation in the study was voluntary, and no financial or other compensation was offered for participation in the trial. Students who voluntarily chose to participate in the post-trial interview were offered a €10 gift card (approximately US $11) as compensation.

With regard to privacy and confidentiality, all study data were pseudonymized prior to analysis, with personal identifiers replaced by ID numbers and accessible only to data managers/statisticians. All data are stored on secure, access-controlled local servers at Jamk University of Applied Sciences with daily backups. Access is restricted to authorized research team members and managed by the principal investigator. No identifiable information will be reported.

Intervention

The first wave of the trial was delivered in spring 2024, and the second wave of the trial was delivered in autumn 2024. The trial focused on 2 subjects: mathematics and foreign languages, representing different types of subjects. Students participated in 3 different treatments (lessons) in random order. All lessons followed the National Core Curriculum for General Upper Secondary Education, but the teaching methods were different:

1. Treatment 1: Physically active mathematics or foreign language lessons with PAL (20 min PA is integrated into learning goals)
2. Treatment 2: Physically active mathematics or foreign language lessons with PA breaks (two 5-min breaks, including PA not related to learning goals)
3. Treatment 3: Typical, traditional math/foreign language lessons without PA (control lesson).

A description of the PAL methods and PA breaks used for treatment 1 and 2 is described in more detail (Textbox 1). The content of these methods was constructed in collaboration with general upper secondary school teachers, teacher educators, PA and learning researchers, and students.

Textbox 1. Description of the content of the physically active learning (Treatment 1) and physical activity breaks used (Treatment 2).

Physically active learning

Twenty minutes of PA (including instructions) were integrated into mathematics and foreign language lessons using station work. Station work was selected as the PAL method based on the results of a development project, "Moving academic lessons in general upper secondary school," in which subject teachers identified it as a suitable approach across subjects. In this method, students solved subject-related tasks at different stations. The criteria for implementing station work were developed collaboratively by general upper secondary school teachers, teacher educators, and PA and learning researchers. These criteria included the number and location of stations, the type of PA integrated into tasks, and the transitions between stations.

Based on these criteria, 6-8 stations per lesson were set up in classrooms and/or other indoor spaces (eg, corridors), excluding outdoor areas. Transitions between stations could include climbing stairs, and students were instructed to walk briskly between stations. Approximately half of the station tasks involved PA, such as responding while moving. The physical activities were designed to be versatile and included balance, locomotor skills, manipulative skills, and strength (eg, squats, boxing, and precision throws). Most station tasks were completed in small groups, with some tasks performed in pairs.

Using these criteria, teachers planned lesson-specific tasks for each station with support from the research team. In addition to the station-based tasks, teachers planned the remainder of the lesson, which mainly consisted of teacher-led instruction, assigning and checking homework, and individual work using mathematics study books. The mathematics or foreign language tasks, including PA, were implemented in the middle of a 75-minute lesson, starting 25 minutes after the lesson began.

Teachers were instructed to rehearse the PAL (with different academic content) with the students to familiarize themselves and the students with the methods before the measurements.

Physical activity breaks

Two 3-minute PA breaks were added to mathematics and foreign language lessons. At the teachers' request, prerecorded exercise videos were selected as the simplest format for implementing the PA breaks. Two different breaks were constructed from the Students on the Move program's break exercise cards [58]. The breaks were planned to be versatile, including balance, locomotor skills, and strength exercises, and five cards per break were selected based on students' preferences.

The PA breaks were implemented in the classroom, with the first break taking place 25 minutes after the lesson began and the second break 50 minutes after the lesson began (Break 1: [59]; Break 2: [60]). The breaks were mainly performed individually with the whole class participating, while the final exercises of each break were completed in pairs. The PA breaks did not include mathematics or foreign language learning goals, and teachers designed the lesson content apart from the breaks.

Before the measurements, teachers were instructed to rehearse the PA breaks with the students to familiarize both themselves and the students with the procedures.

Researchers held a 2-hour meeting to present the study, PAL method principles, and PA breaks to the teachers implementing the intervention. The PAL method (treatment 1) was similar for every group, but the lesson-specific content was designed by the teacher. This ensured that the lesson content was a part of the course learning objectives. Teachers had the opportunity to

XSL•FO
RenderX

ask for assistance from research team members and were required to submit lesson plans. The PA breaks (treatment 2) were the same for every group, and teachers only needed to play them on the screen for the students. Because only some of the teachers and students were familiar with the different physically active classroom practices, the teachers were instructed to rehearse the PAL method and PA breaks with the students to familiarize themselves and the students with the methods before the measurements.

Given that the physically active lessons adhered to the national curriculum and consisted of tasks typical of students' everyday lives, participation did not pose any risk beyond that of a regular school day. Despite the study's hypotheses, PAL and PA breaks may cause negative effects, such as reduced situational engagement, alertness, or executive functions. These effects were measured and analyzed as part of the intervention outcomes, and subgroup analyses may reveal whether modifying factors are associated with perceived disadvantages.

The fidelity of the intervention was monitored by having each lesson observed by trained personnel, who recorded the course of the lesson and any deviations or unusual events. Apart from the addition of the second wave, no changes were made to the study protocol after trial registration, and no major deviations from the protocol occurred during the study.

Measurements

The majority of the measurements were performed as part of the lessons conducted by trained personnel. In the baseline measurement session, students' executive functions (inhibition and working memory) and academic skills such as reading and math fluency and spatial visualization were measured. In addition, the devices (Axivity AX3 and Firstbeat Bodyguard 3 monitoring device) and instructions for measuring PA, sedentary time, stress, recovery, and sleep were distributed, and instructions for self-monitoring usual daily and hourly variations in alertness were given. Finally, personalized answer links to complete the background survey were sent to the students via email. Students completed the questionnaire in their free time.

Students wore the devices for two 4-to-5-day measurement periods (including the treatments). The first measurement period started straight after the baseline measurement session, or at the latest, in the morning 3 days before the first intervention lesson, and it ended in the morning after the first intervention lesson. The second measurement period started in the morning of the day before the second intervention lesson and ended in the morning after the last intervention lesson. Students also filled out an online alertness questionnaire 4 times per day (9 AM, 11 AM, 1 PM, and 3 PM) for 3 days. The timeline of the measurements is presented in Table 1.

Table . Example of a timetable of the measurements[a].

	Day 1	Day 2	Day 3	Day 4	Day 5
Week 1	Baseline measurements. The first measurement period starts (devices on)	Alertness self-monitored at 9 AM, 11 AM, 1 PM, and 3 PM (devices on)	Alertness self-monitored at 9 AM, 11 AM, 1 PM, and 3 PM (devices on)	Intervention lesson 1 (devices on)	Devices removed
Week 2	Second measurement period starts (devices on). Alertness self-monitored at 9 AM, 11 AM, 1 PM, and 3 PM	Intervention lesson 2 (devices on)	Devices on	Intervention lesson 3 (devices on)	Devices removed

[a]For one group, the measurements lasted two weeks. During the first week of the study, the devices (Firstbeat Life and Axivity) were distributed during baseline measurements and were put on the same or the following day. The devices were worn for 3 - 4 days continuously and were removed the day after the first intervention lesson. During the second week of the study, the devices were put on the day before the second intervention lesson and removed the day after the third intervention lesson. The measurement period in the second week lasted 2 - 4 days continuously, depending on the group's schedule.

Measurements during the intervention lessons included a short initial survey, an executive function test, and an alertness questionnaire at the beginning of the lesson; 2 alertness questions during the lesson; and an alertness questionnaire, an executive function test, and a situational engagement questionnaire at the end of the lesson (Figure 3). The initial survey (2 minutes to complete), conducted on students' own computers through Visual Learning Environment (ViLLE), a widely used digital assessment and learning platform, included questions about the quality of the previous night's sleep, as well as the consumption of breakfast, lunch, caffeine, nicotine products, and PA before the treatment. Students also wore the Axivity AX3 and Firstbeat Bodyguard 3 monitors to assess PA, standing, sitting, stress, and recovery during the intervention lessons.

XSL•FO
RenderX

Figure 3. The structure of the intervention lessons. Measurements during the intervention lessons included a short initial survey, an executive function test, and an alertness questionnaire at the beginning of the lesson; two alertness questions during the lesson; and an alertness questionnaire, an executive function test, and a situational engagement questionnaire at the end of the lesson.

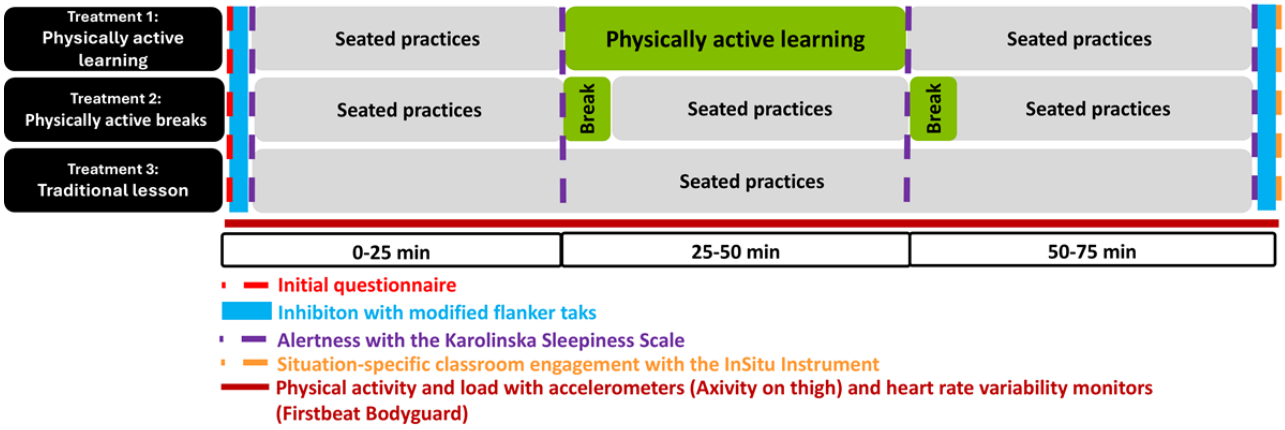

Main Outcomes

Alertness

Changes in alertness were assessed using the validated Karolinska Sleepiness Scale [61,62]. The Karolinska Sleepiness Scale is a single-item measure of subjective situational alertness on a 9-point scale (1=extremely alert, 9=extremely sleepy, fighting sleep). Students completed it on paper, and it took 0.5 minutes to complete.

Executive Function—Inhibition

Changes in executive function performance, particularly in inhibition, were measured during the treatments. Inhibition was assessed using a computerized, modified version of the Eriksen flanker task [63], in which participants completed up to 160 trials involving 5 horizontally aligned blue arrows presented in either congruous (the middle arrow facing the same direction as the flanking arrows) or incongruous (the middle arrow facing the opposite direction) conditions. Students were instructed to respond to the direction that the middle arrow was facing as quickly and accurately as possible. A fixed interstimulus interval of 1000 ms separated the trials. Accuracy (%) and reaction time for correct responses (ms) in congruent and incongruent trials were assessed. Students completed the flanker task on their own computers via the ViLLE platform, and it took 3 minutes. The task, although built on a different technology, mimics the NIH Toolbox version of the Flanker task, which has demonstrated developmental sensitivity and good test–retest reliability in both children and adults [64-66].

Situational Engagement

Situational engagement was assessed with the validated InSituations (InSitu) instrument [25]. The InSitu consists of 17 items rated on a 5-point scale (1=not at all; 5=very much) measuring (1) behavioral and cognitive engagement (7 items; eg, “How well did you concentrate during the lesson?”), (2) emotional engagement (3 items, eg, “How much did you like this lesson?”), (3) disaffection (3 items, eg, “How much did you do other things than the ongoing tasks and instruction?”), (4) competence experiences (2 items, eg, “How easy was the lesson for you?”), and (5) help seeking (2 items, eg, “How much did you ask for help from the teacher/another adult during the lesson?”). Students completed the InSitu on their own computers through the ViLLE platform, and it took 2 to 3 minutes to complete.

Modifying Factors

Physical Load

To monitor physiological stress and recovery, students recorded a single-channel electrocardiogram to measure heart rate variability (HRV). HRV is measured as the time interval, in milliseconds, between two consecutive R-wave peaks (RR-Interval) in the electrocardiogram. The measurement was done using a Bodyguard 3 (Firstbeat Technologies Oy, Jyväskylä, Finland) and the Firstbeat Life app. The device was attached with self-adhesive single-use electrodes (Arbo H92SG). Participants wore the device around the clock, except for showering and water exercise. In the app, the students logged their activities in the supporting electronic diary.

Data were processed using Firstbeat research software. Firstbeat Life analysis is based on heart rate analysis software that identifies different physiological changes through HRV. Physiological recovery, stress, sleep, and PA were collected [67]. A score ranging from 0 to 100 was generated for each of

XSL•FO
RenderX

the different physiological states for each day (morning to morning). A higher number represents a better score.

Mental Load

Mental load was assessed using questionnaires measuring study burnout and engagement as part of the background survey. The Student Burnout Inventory-9 [68] assesses study burnout through 9 items measuring exhaustion, cynicism, and inadequacy at school, rated on a 6-point scale (1=strongly disagree; 6=strongly agree). Meanwhile, Energy, Dedication, and Absorption-9 [69] assesses study engagement through 9 items, rated on a 7-point scale (0=not at all, 7=daily).

Physical Competence

Perceived physical competence was measured as part of the background survey using a questionnaire that included a modified Finnish version of the Perceived Physical Competence Scale for Children [70] and questions about the ability and desire to develop one's own physical ability [71]. Altogether, they comprise 12 items rated on a 5-point scale.

Academic Competence

Perceived academic competence was assessed using a questionnaire including self-reported academic achievement score and self-perceptions (in math and foreign language) [72]. Perceived academic achievement score was assessed with a question "How would you rate your current academic performance in mathematics?" rated on a 4-point scale (1=excellent, 4=poor). Self-efficacy in math and foreign language was measured with 4 items (eg, "Mathematics/foreign language is an easy subject") with a 5-point Likert scale (1=completely agree, 5=completely disagree).

Other Study Variables

PA and Sedentary Time

To measure PA and sedentary time, students wore an accelerometer (AX3 [73]) on the anterior midline of either thigh. The accelerometer was attached to the skin using Opsite Flexifix transparent adhesive film (8 × 10 cm), with a piece of mesh between the monitor and skin to enhance breathability [74]. Envelopes including the monitor, detailed instructions including video link for how to attach the monitor, 5 pieces of adhesive film, and mesh swabs were delivered to participants face-to-face in a classroom. Data were collected with a 100 Hz sampling rate and a dynamic range of ±8 g.

Raw data were downloaded using OMGUI (version 1.0.0.45; Open Lab) [75]. ActiPASS (version 1.62) [76] was used to perform activity classification. Data were processed using the software's default ProPASS settings, which have been validated to work well with children's data as well as adults' [77]. The start and end times, time at school, and intervention class times for each participant's data were imported to ActiPASS using a diary file to inspect different time domains separately. The autocalibration function of the software was used. Activity was classified based on movement type: sleep, lie, sitting, standing, walking, running, stair-walking, cycling, but also based on the intensity of movement: sedentary, light, moderate, and vigorous PA.

Cognitive Functions and Academic Skills

To address baseline cognitive functions and academic skills, performance in the following tasks was evaluated through the ViLLE platform: the Choice Reaction Time task (as a warm-up task), the Sentence Reading Fluency task (assessing reading fluency), the Calculations task (assessing arithmetic fluency), the Digitized Corsi Blocks task (assessing working memory), the Spatial Relations subtask of the Woodcock-Johnson III Cognitive Assessment (assessing spatial visualization ability), and the Flanker task (assessing inhibition). Completing the tests took approximately 20 minutes. The tasks are described in more detail in Table S1 in Multimedia Appendix 1.

Background variables

The background survey was conducted electronically using Webropol software (Webropol Ltd, Helsinki, Finland), consisted of 55 questions and took a median time of 20 minutes to complete. The collected data included a comprehensive list of background variables: personal identifying information, age and sex, perceived health, height, weight, physical fitness, several questions on PA, sleep and wake-up times, nutrition, family form, parental socioeconomic questions, background academic performance in comprehensive school, and language syllabi in comprehensive school. Regarding the studies in upper secondary school, the questionnaire measured, for example, perceived performance and motivation in studies and diagnosed learning difficulties.

Statistical Approaches

Analysis Plan

Generalized linear mixed-effects models will be applied to study the acute effects of the intervention on outcome measures. A cluster-randomized individual crossover design is assumed, that is, individuals serve as their own controls when the clusters (school classes) are crossed over the three intervention arms. This is the most efficient of cluster-randomized study designs [78]. The data will be analyzed according to intention-to-treat principles. Data lost due to participation withdrawals, student absences from measurement lessons, and missingness caused by other reasons will be multiply imputed in the analysis of primary outcome variables.

Within the analysis of main outcomes of classroom interventions (alertness, executive function, and situational engagement), the first analyses will quantify the main intervention effects on the outcomes. The main effects in generalized linear mixed-effects models will be adjusted for 3 levels of confounding factors. First, the factors that are consequences of the study design are taken into account (season, lesson time, subject, and group ID as a random intercept). Second, the student-level constant factors (sex, learning-related disabilities) will be included in the models. Finally, the third set of models will be additionally adjusted for individual lesson-specific factors (prelesson activities, baseline alertness, quality and duration of preceding sleep, preceding meals, smoking and caffeine intake, and menstruation status). The further manuscripts will study modifying factors for the intervention effects. The studied modifying factors are planned to include current physical and mental load, perceived physical competence, and perceived academic competence. The models

XSL•FO
RenderX

will include corresponding or less extensive adjustments based on the observations of analysis of main intervention effects. Mplus and R statistical packages will be used.

Sample Size Determination

Prerecruitment power calculations were conducted to determine the required sample size for the study. Required statistical power (1-β) was set to 90%, and significance level (α) was set at .05. The study was planned to detect medium intervention main effects (Cohen *d*=0.50) in the primary outcome variables: situational engagement, alertness, and inhibition. Definition of the desired minimum detectable effect size was, as usual, a compromise between academic desire and study costs and supported by published works on the outcomes (eg, alertness [79,80], inhibition [81,82]).

The required sample size *n* for an individually randomized controlled trial would be

$$nIRCT=(Z1-\alpha/2+Z1-\beta)2\cdot 2\sigma 2\Delta 2,$$

where Zx is the *x*th quantile of the standard normal distribution, α is the size of the test (ie, level of significance), β is 1 – required level of statistical power, Δ is the minimum detected difference in the levels of outcome, and σ2 is the assumed overall variance of the outcome variable.

In educational interventions, participants are most often divided into school classes, and therefore, individual randomization for intervention groups is rarely feasible. This cluster-randomization weakens the design generally, and the required sample size increases by a factor called the design effect.

In our study design, individuals serve as their own controls by visiting the two intervention treatments and a control treatment in randomized order. The design effect gets a form of

$$DECRIC=2(1-\rho 1-\rho 2),$$

where ρ1 is the assumed intra-subject correlation and ρ2 is the assumed intra-cluster correlation.

Assuming the variance structure as observed in the pilot study (Matikkavire), ρ1=0.31 and ρ2=0.01, the design effect is 21-0.31-0.01=1.36 and the required sample is n=115 individuals to be crossed over the treatments. By assuming 33 students per class and a 70% student participation rate, we need to recruit five classes. To balance the design, we decided to recruit six classes such that two classes would have been randomized to each of the three intervention arms, where the order of treatments is rotated. As the number of volunteering participants per group remained lower than expected, more groups were invited, and eventually, the recruitment continued in a second wave in the autumn of 2024 to obtain the desired statistical power.

The forthcoming subgroup analyses (eg, sex-specific analyses) and moderation analyses (moderating factors will include physiological and psychological state, and physical and academic competence of the student) have less statistical power to detect medium interaction effects. Statistical power is approximated for subgroups by standard methods as discussed above, and for more complex moderation analyses by, eg, post data-collection Monte Carlo simulation. The post hoc power calculations are not used to determine sample size, but to support the confidence of the interpretation of the results of the moderation models.

Student interviews (Part 3)

The evaluation of students' experiences and perceptions of physically active academic lessons in a general upper secondary school will be conducted in the context of the cluster-randomized individual crossover trial described above. Of the 74 students recruited in the first wave of the trial, 59 consented (21 from mathematics lessons and 38 from foreign language lessons) to being invited for an interview (Figure 1). The aim is to interview as wide a range of students as possible; therefore, students will be selected in different ways. While one subject teacher proposed potential interviewees to the researcher, most students expressed their willingness to participate during their initial face-to-face contact with the researcher following the intervention. Thirty students were contacted and agreed to participate in the interviews in spring 2024. Ten students participated in the trial during mathematics lessons (5 taking the short syllabus and 5 the long syllabus), and 20 students during foreign language lessons (8 in English, 13 in Swedish, and 1 in German).

All 30 students were interviewed after participating in all three different lessons (treatments; see Figure 3). Ten students will be interviewed virtually via Microsoft Teams, and 20 students will be interviewed in person at their schools using an Olympus LS-PS-1 audio recorder. Semistructured interviews were chosen as a method to enable open and spontaneous responses and to capture the essence of students' experience [83]. The questions are organized around the main themes (see Student interviews in the Supplemental material in Multimedia Appendix 1). The duration of the interviews will range from 37 to 70 minutes. The student responses will be transcribed and analyzed by using the engagement framework [23] and thematic analysis [56]. Thematic analysis is suitable for identifying patterns within and across data in relation to participants' lived experience, views, and perspectives when the study aims to understand what participants think, feel, and do. Both deductive (literature-driven dimensions) and inductive (emerging from participants' data) approaches will be used. The main focus of using the situational engagement framework is to analyze students' responses to PAL. In addition, responses to PA breaks allow for a comparison of how PAL and PA breaks influence situational engagement.

Results

Data collection has been completed. Fourteen teachers participated in the interviews (Part 1), 168 students participated in a cluster-randomized individual crossover trial in mathematics and foreign language lessons (Part 2), and 30 students participated in post-trial interviews (Part 3). The participating students received personalized feedback on their PA, physiological recovery, stress, and sleep measurements. In addition, teachers who conducted the intervention lessons received school-specific summaries, including information on students' levels of PA, sleep patterns, and engagement during lessons. The results of the study will be reported in separate manuscripts, and the first findings based on the data analysis

XSL•FO
RenderX

are expected by spring 2026. The results will be disseminated through plain-language summaries tailored for teachers and schools, published reports in the trial registry, and presentations at scientific and national seminars aimed at researchers and practitioners.

Discussion

Principal Results

This multidisciplinary mixed methods study aims to evaluate the experiences and perceptions of both teachers and students during physically active academic lessons in a general upper secondary school. The research interest is especially in facilitators, barriers, usefulness, and meaning of physically active academic lessons for the situational learning process. Additionally, the purpose is to examine the immediate effects of PAL and PA breaks on cognitive prerequisites for learning such as alertness and executive functions, and situational engagement, including behavioral, emotional, and cognitive engagement, as well as disaffection, competence experiences, and help-seeking among secondary school students. The study also aims to explore factors that modify these effects, such as physical and mental load, as well as perceived physical and academic competence. Aligning with previous findings, we hypothesize that this study will confirm the benefits [26-29], feasibility, and usability [41,43,50,51] of physically active classroom practices in general upper secondary education. Moreover, we expect the study to provide new insights into the role of PAL and PA breaks in situational learning processes in authentic educational settings, considering both teacher and student perspectives as well as potential moderating factors.

Comparison With Prior Work

This is one of the first studies to explore experiences related to teaching, learning processes, and optimal instructions during physically active academic lessons in general upper secondary schools, while also examining the immediate effects of PAL and PA breaks on situational engagement and cognitive prerequisites for learning. Furthermore, the strength of this study lies in its combination of solid knowledge in both teaching and PA, allowing it to examine the phenomenon from both the student's and the subject teacher's perspectives at different levels of action (collected experiences and measured prerequisites for learning). A further strength of the study is that PA was measured from the thigh, enabling the distinction between sitting and standing during lessons. Additionally, HRV was monitored online to assess the balance between stress and recovery during classroom activities, which adds a novel dimension to the research.

Implications and Future Research

The study's findings will offer crucial, evidence-based insights into active classroom practices, aiding subject teachers and teacher education programs. This support helps teachers and schools adopt effective and innovative instructional strategies, pedagogical models, and designs, ultimately enhancing student learning and well-being in upper secondary education. Future research should study the phenomenon in other educational settings for older students, such as vocational secondary education and higher education, and across different cultural contexts. Research on the acute effects and students' experiences should be extended to other subjects beyond mathematics and languages.

Limitations

The study has faced certain limitations. The participation rate in the study was lower than desired, and possibly, less active or academically weaker students did not participate in the study, which may weaken the generalizability of the study results. The generalizability of teachers' experiences is limited, as the findings are based on fourteen interviews from one municipality. While the interviews offer varied insights into integrating PA into different subjects, the experiences of individual subjects often rely on just one or two teachers. The teachers who participated were interested in the topic, but the views of less interested teachers were not captured. However, a couple of teachers who withdrew from the development project were also interviewed, which helps to understand why implementing PAL and PA breaks can be challenging.

Conclusions

This study aims to address a gap in the literature by evaluating the feasibility and effectiveness of integrating physically active methods into academic lessons in upper secondary schools. Using a mixed-methods cluster-randomized individual crossover trial, the study seeks to enhance students' learning and well-being.

Acknowledgments

We used Microsoft's generative AI tool, Copilot, only to improve the readability and language of the manuscript. The authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Funding

This study was funded by the Research Council of Finland under Grant 355350.

Data Availability

The anonymized datasets generated or analyzed during this study will be available from the authors [HJS, THT, KM, and ST], upon reasonable request. The Data Management Plan for the study has been created using the DMPTuuli service and made publicly available.

XSL•FO
RenderX

Authors' Contributions

The authors HJS, THT, KM, and ST conceived the original concept and study design. TK performed sample size calculations and planned statistical data analyses. HH performed the randomization of the sample. HJS and PR planned cognitive function and academic performance measurements. HJS, HL, and HO were responsible for all measurements. HL and HO were responsible for the physiological stress and recovery measurements and analyses. JK planned accelerometer-based physical activity measurements, and JK, HH, and TK conducted the analysis. ST and NS were responsible for teacher and student interviews. All authors drafted the first version of the manuscripts and critically reviewed the manuscript for important intellectual content. All authors read and approved the final version of the manuscript and agreed to be accountable for all aspects of the work, ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplemental material on teacher and student interviews and measurements of cognitive and academic skills.
[PDF File, 160 KB - resprot_v15i1e84601_app1.pdf]

Checklist 1
SPIRIT 2025 checklist.
[PDF File, 149 KB - resprot_v15i1e84601_app2.pdf]

References

1. PISA 2015 Results (Volume III) Students' Well-Being: Organisation for Economic Co-operation and Development; 2017. URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2017/04/pisa-2015-results-volume-iii_g1g787cb/9789264273856-en.pdf [accessed 2026-01-23] [doi: 10.1787/9789264273856-en]
2. Global health risks: mortality and burden of disease attributable to selected major risks. : World Health Organization; 2009 URL: https://iris.who.int/server/api/core/bitstreams/50e6ba96-c5c3-4e1d-b635-f111bb74f4bf/content [accessed 2026-01-23]
3. Twenge JM, Cooper AB, Joiner TE, Duffy ME, Binau SG. Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. J Abnorm Psychol 2019 Apr;128(3):185-199. [doi: 10.1037/abn0000410] [Medline: 30869927]
4. Student barometer in general upper secondary school [Article in Finnish]. The Finnish Upper Secondary Students' Union. 2022. URL: https://lukio.fi/seravo-start/english/ [accessed 2026-01-23]
5. Salmela-Aro K, Savolainen H, Holopainen L. Depressive symptoms and school burnout during adolescence: evidence from two cross-lagged longitudinal Studies. J Youth Adolescence 2009 Nov;38(10):1316-1327. [doi: 10.1007/s10964-008-9334-3]
6. Madigan DJ, Curran T. Does burnout affect academic achievement? A meta-analysis of over 100,000 students. Educ Psychol Rev 2021 Jun;33(2):387-405. [doi: 10.1007/s10648-020-09533-1]
7. Guthold R, Stevens GA, Riley LM, Bull FC. Global trends in insufficient physical activity among adolescents: a pooled analysis of 298 population-based surveys with 1·6 million participants. Lancet Child Adolesc Health 2020 Jan;4(1):23-35. [doi: 10.1016/S2352-4642(19)30323-2] [Medline: 31761562]
8. Finland's report card 2022 on physical activity for children and youth. 2022. LIKES research reports on physical activity and health 407. : JAMK University of Applied Sciences; 2022 URL: https://www.activehealthykids.org/wp-content/uploads/2022/03/Finland-report-card-long-form-2022.pdf [accessed 2026-01-23]
9. Chaput JP, Willumsen J, Bull F, et al. 2020 WHO guidelines on physical activity and sedentary behaviour for children and adolescents aged 5-17 years: summary of the evidence. Int J Behav Nutr Phys Act 2020 Nov 26;17(1):141. [doi: 10.1186/s12966-020-01037-z] [Medline: 33239009]
10. Norris E, van Steen T, Direito A, Stamatakis E. Physically active lessons in schools and their impact on physical activity, educational, health and cognition outcomes: a systematic review and meta-analysis. Br J Sports Med 2020 Jul;54(14):826-838. [doi: 10.1136/bjsports-2018-100502] [Medline: 31619381]
11. Bedard C, St John L, Bremer E, Graham JD, Cairney J. A systematic review and meta-analysis on the effects of physically active classrooms on educational and enjoyment outcomes in school age children. PLoS ONE 2018;14(6):e0218633. [doi: 10.1371/journal.pone.0218633]
12. Watson A, Timperio A, Brown H, Best K, Hesketh KD. Effect of classroom-based physical activity interventions on academic and physical activity outcomes: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017 Aug 25;14(1):114. [doi: 10.1186/s12966-017-0569-9] [Medline: 28841890]
13. Sneck S, Viholainen H, Syväoja H, et al. Effects of school-based physical activity on mathematics performance in children: a systematic review. Int J Behav Nutr Phys Act 2019 Nov 21;16(1):109. [doi: 10.1186/s12966-019-0866-6] [Medline: 31752903]

XSL•FO
RenderX

14. Daly-Smith AJ, Zwolinsky S, McKenna J, Tomporowski PD, Defeyter MA, Manley A. Systematic review of acute physically active learning and classroom movement breaks on children's physical activity, cognition, academic performance and classroom behaviour: understanding critical design features. BMJ Open Sport Exerc Med 2018;4(1):e000341. [doi: 10.1136/bmjsem-2018-000341] [Medline: 29629186]
15. Masini A, Marini S, Gori D, Leoni E, Rochira A, Dallolio L. Evaluation of school-based interventions of active breaks in primary schools: a systematic review and meta-analysis. J Sci Med Sport 2020 Apr;23(4):377-384. [doi: 10.1016/j.jsams.2019.10.008] [Medline: 31722840]
16. de Greeff JW, Bosker RJ, Oosterlaan J, Visscher C, Hartman E. Effects of physical activity on executive functions, attention and academic performance in preadolescent children: a meta-analysis. J Sci Med Sport 2018 May;21(5):501-507. [doi: 10.1016/j.jsams.2017.09.595] [Medline: 29054748]
17. Diamond A. Executive functions. Annu Rev Psychol 2013;64:135-168. [doi: 10.1146/annurev-psych-113011-143750] [Medline: 23020641]
18. Best JR, Miller PH, Naglieri JA. Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. Learn Individ Differ 2011 Aug;21(4):327-336. [doi: 10.1016/j.lindif.2011.01.007] [Medline: 21845021]
19. Mavilidi MF, Ruiter M, Schmidt M, et al. A narrative review of school-based physical activity for enhancing cognition and learning: the importance of relevancy and integration. Front Psychol 2018;9(OCT):2079. [doi: 10.3389/fpsyg.2018.02079] [Medline: 30464752]
20. Bartholomew JB, Jowers EM. Physically active academic lessons in elementary children. Prev Med 2011 Jun;52 Suppl 1(Suppl):S51-S54. [doi: 10.1016/j.ypmed.2011.01.017] [Medline: 21281672]
21. Daly-Smith A, Quarmby T, Archbold VSJ, et al. Implementing physically active learning: Future directions for research, policy, and practice. J Sport Health Sci 2020 Jan;9(1):41-49. [doi: 10.1016/j.jshs.2019.05.007] [Medline: 31921479]
22. Fredricks JA, Blumenfeld PC, Paris AH. School engagement: potential of the concept, state of the evidence. Rev Educ Res 2004 Mar;74(1):59-109. [doi: 10.3102/00346543074001059]
23. Fredricks JA, Filsecker M, Lawson MA. Student engagement, context, and adjustment: addressing definitional, measurement, and methodological issues. Learn Instr 2016 Jun;43:1-4. [doi: 10.1016/j.learninstruc.2016.02.002]
24. Xerri MJ, Radford K, Shacklock K. Student engagement in academic activities: a social support perspective. High Educ 2018 Apr;75(4):589-605. [doi: 10.1007/s10734-017-0162-9]
25. Vasalampi K, Muotka J, Pöysä S, Lerkkanen MK, Poikkeus AM, Nurmi JE. Assessment of students' situation-specific classroom engagement by an InSitu Instrument. Learn Individ Differ 2016 Dec;52(2):46-52. [doi: 10.1016/j.lindif.2016.10.009]
26. Haverkamp BF, Wiersma R, Vertessen K, van Ewijk H, Oosterlaan J, Hartman E. Effects of physical activity interventions on cognitive outcomes and academic performance in adolescents and young adults: a meta-analysis. J Sports Sci 2020 Dec;38(23):2637-2660. [doi: 10.1080/02640414.2020.1794763] [Medline: 32783695]
27. Verburgh L, Königs M, Scherder EJA, Oosterlaan J. Physical exercise and executive functions in preadolescent children, adolescents and young adults: a meta-analysis. Br J Sports Med 2014 Jun;48(12):973-979. [doi: 10.1136/bjsports-2012-091441] [Medline: 23467962]
28. Mavilidi MF, Mason C, Leahy AA, et al. Effect of a time-efficient physical activity intervention on senior school students' on-task behaviour and subjective vitality: the 'burn 2 learn' cluster randomised controlled trial. Educ Psychol Rev 2021 Mar;33(1):299-323. [doi: 10.1007/s10648-020-09537-x]
29. Robinson KJ, Lubans DR, Mavilidi MF, et al. Effects of classroom-based resistance training with and without cognitive training on adolescents' cognitive function, on-task behavior, and muscular fitness. Front Psychol 2022;13:811534. [doi: 10.3389/fpsyg.2022.811534] [Medline: 35386901]
30. Donnelly JE, Hillman CH, Castelli D, et al. Physical activity, fitness, cognitive function, and academic achievement in children: a systematic review. Med Sci Sports Exerc 2016 Jun;48(6):1197-1222. [doi: 10.1249/MSS.0000000000000901] [Medline: 27182986]
31. Singh AS, Saliasi E, van den Berg V, et al. Effects of physical activity interventions on cognitive and academic performance in children and adolescents: a novel combination of a systematic review and recommendations from an expert panel. Br J Sports Med 2019 May;53(10):640-647. [doi: 10.1136/bjsports-2017-098136] [Medline: 30061304]
32. Jafflin K, Pfeiffer C, Bergman MM. Effects of self-esteem and stress on self-assessed health: a Swiss study from adolescence to early adulthood. Qual Life Res 2019 Apr;28(4):915-924. [doi: 10.1007/s11136-018-2059-1] [Medline: 30560538]
33. Matud MP, Díaz A, Bethencourt JM, Ibáñez I. Stress and p[sychological distress in emerging adulthood: a gender analysis. J Clin Med 2020 Sep 4;9(9):1-11. [doi: 10.3390/jcm9092859] [Medline: 32899622]
34. Girotti M, Adler SM, Bulin SE, Fucich EA, Paredes D, Morilak DA. Prefrontal cortex executive processes affected by stress in health and disease. Prog Neuropsychopharmacol Biol Psychiatry 2018 Jul 13;85:161-179. [doi: 10.1016/j.pnpbp.2017.07.004] [Medline: 28690203]
35. Dong Z, Wang P, Xin X, et al. The relationship between physical activity and trait anxiety in college students: the mediating role of executive function. Front Hum Neurosci 2022;16:1009540. [doi: 10.3389/fnhum.2022.1009540] [Medline: 36211122]

XSL•FO
RenderX

36. Stults-Kolehmainen MA, Sinha R. The effects of stress on physical activity and exercise. Sports Med 2014 Jan;44(1):81-121. [doi: 10.1007/s40279-013-0090-5] [Medline: 24030837]
37. Timo J, Sami YP, Anthony W, Jarmo L. Perceived physical competence towards physical activity, and motivation and enjoyment in physical education as longitudinal predictors of adolescents' self-reported physical activity. J Sci Med Sport 2016 Sep;19(9):750-754. [doi: 10.1016/j.jsams.2015.11.003] [Medline: 26671711]
38. Lindner KJ. The physical activity participation–academic performance relationship revisited: perceived and actual performance and the effect of banding (academic tracking). Pediatr Exerc Sci 2002;14(2):155-169. [doi: 10.1123/pes.14.2.155]
39. Ferla J, Valcke M, Schuyten G. Judgments of self-perceived academic competence and their differential impact on students' achievement motivation, learning approach, and academic performance. Eur J Psychol Educ 2010 Dec;25(4):519-536. [doi: 10.1007/s10212-010-0030-9]
40. Sneck S, Syväoja H, Järvelä S, Tammelin T. More active lessons: teachers' perceptions of student engagement during physically active maths lessons in Finland. Education Inquiry 2023 Oct 2;14(4):458-479. [doi: 10.1080/20004508.2022.2058166]
41. Schmidt SK, Bratland-Sanda S, Bongaardt R. Secondary school teachers' experiences with classroom-based physically active learning: "I'm excited, but it's really hard". Teach Teach Educ 2022 Aug;116:103753. [doi: 10.1016/j.tate.2022.103753]
42. Teslo S, Thurston M, Lerum Ø, et al. Teachers' sensemaking of physically active learning: a qualitative study of primary and secondary school teachers participating in a continuing professional development program in Norway. Teach Teach Educ 2023 Jun;127:104113. [doi: 10.1016/j.tate.2023.104113]
43. Lerum Ø, Bartholomew J, McKay H, et al. Active smarter teachers: primary school teachers' perceptions and maintenance of a school-based physical activity intervention. Transl J ACSM 2019 Sep 1;4(17):141-147. [doi: 10.1249/TJX.0000000000000104]
44. Pöysä S, Vasalampi K, Muotka J, Lerkkanen MK, Poikkeus AM, Nurmi JE. Variation in situation-specific engagement among lower secondary school students. Learn Instr 2018 Feb;53:64-73. [doi: 10.1016/j.learninstruc.2017.07.007]
45. Fröberg A, Jonsson L, Berg C, et al. Effects of an empowerment-based health-promotion school intervention on physical activity and sedentary time among adolescents in a multicultural area. Int J Environ Res Public Health 2018 Nov 13;15(11):2542. [doi: 10.3390/ijerph15112542] [Medline: 30428548]
46. Mullins NM, Michaliszyn SF, Kelly-Miller N, Groll L. Elementary school classroom physical activity breaks: student, teacher, and facilitator perspectives. Adv Physiol Educ 2019 Jun 1;43(2):140-148. [doi: 10.1152/advan.00002.2019] [Medline: 30933539]
47. Howie EK, Newman-Norlund RD, Pate RR. Smiles count but minutes matter: responses to classroom exercise breaks. Am J Health Behav 2014 Sep;38(5):681-689. [doi: 10.5993/AJHB.38.5.5] [Medline: 24933137]
48. Sneck S, Järvelä S, Syväoja H, Tammelin T. Pupils' experiences and perceptions of engagement during the Moving Maths programme. Educ 3 13 2022 Apr 3;50(3):419-434. [doi: 10.1080/03004279.2020.1857816]
49. Riley N, Lubans D, Holmes K, Hansen V, Gore J, Morgan P. Movement-based mathematics: enjoyment and engagement without compromising learning through the EASY minds program. EURASIA J MATH SCI T 2017;13(6):1653-1673. [doi: 10.12973/eurasia.2017.00690a]
50. Fred N, Mikkonen K, Pramila-Savukoski S, Hylkilä K, Kuivila HM. Upper secondary school students' experiences of how exercise breaks affect their well-being and ability to study: a qualitative study. Int J Res Educ Sci 2023;9(3):715-731. [doi: 10.46328/ijres.3210]
51. Romar JE, Enlund M, Lind S, Björkgren M. Movement integration in academic classrooms; a focus on secondary students' experiences. J Phys Educ Sport 2023;23(4):918-928. [doi: 10.7752/jpes.2023.04115]
52. Mulhearn SC, Kulinna PH, Webster C. Stakeholders' perceptions of implementation of a comprehensive school physical activity program: a review. Kinesiol Rev (Champaign) 2020;9(2):159-169. [doi: 10.1123/kr.2019-0045]
53. Marttinen RHJ, McLoughlin G, Fredrick R, Novak D. Integration and physical education: a review of research. Quest 2017 Jan 2;69(1):37-49. [doi: 10.1080/00336297.2016.1150864]
54. The ethical principles of research with human participants and ethical review in the human sciences in Finland. : Finnish National Board on Research Integrity TENK publications; 2019 URL: https://tenk.fi/sites/default/files/2021-01/Ethical_review_in_human_sciences_2020.pdf [accessed 2026-02-02]
55. Daly-Smith A, Morris JL, Norris E, et al. Behaviours that prompt primary school teachers to adopt and implement physically active learning: a meta synthesis of qualitative evidence. Int J Behav Nutr Phys Act 2021 Nov 20;18(1):151. [doi: 10.1186/s12966-021-01221-9] [Medline: 34801039]
56. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
57. Physically active academic lessons in general upper secondary school - student study. ISRCTN. URL: https://doi.org/10.1186/ISRCTN63809854 [accessed 2026-01-29]
58. Liikkuva opiskelu. URL: https://liikkuvaopiskelu.fi/tukimateriaali/tulostettavat-taukoliikuntakortit/ [accessed 2026-01-29]
59. Taukoliikuntavideoita fyysisesti aktiivisiin oppitunteihin 1. URL: https://www.youtube.com/watch?v=85R7H8-WFZo [accessed 2026-01-29]

XSL•FO
RenderX

60. Taukoliikuntavideoita fyysisesti aktiivisiin oppitunteihin 2. URL: https://www.youtube.com/watch?v=YfX_4p3VIGY [accessed 2026-01-29]
61. Akerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. Int J Neurosci 1990 May;52(1-2):29-37. [doi: 10.3109/00207459008994241] [Medline: 2265922]
62. Shochat T, Santhi N, Herer P, Dijk DJ, Skeldon AC. Sleepiness is a signal to go to bed: data and model simulations. Sleep 2021 Oct 11;44(10):zsab123. [doi: 10.1093/sleep/zsab123] [Medline: 33991415]
63. Eriksen BA, Eriksen CW. Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept Psychophys 1974 Jan;16(1):143-149. [doi: 10.3758/BF03203267]
64. Zelazo PD, Anderson JE, Richler J, et al. NIH toolbox cognition battery (CB): validation of executive function measures in adults. J Int Neuropsychol Soc 2014 Jul;20(6):620-629. [doi: 10.1017/S1355617714000472] [Medline: 24960301]
65. Zelazo PD, Anderson JE, Richler J, Wallner-Allen K, Beaumont JL, Weintraub S. II. NIH toolbox cognition battery (CB): measuring executive function and attention. Monogr Soc Res Child Dev 2013 Aug;78(4):16-33. [doi: 10.1111/mono.12032] [Medline: 23952200]
66. Weintraub S, Dikmen SS, Heaton RK, et al. Cognition assessment using the NIH Toolbox. Neurology (ECronicon) 2013 Mar 12;80(11_supplement_3):S54-S64. [doi: 10.1212/WNL.0b013e3182872ded]
67. Stress and recovery analysis method based on 24-hour heart rate variability. : Firstbeat Technologies Ltd; 2014 URL: https://assets.firstbeat.com/firstbeat/uploads/2015/10/Stress-and-recovery_white-paper_20145.pdf [accessed 2023-01-23]
68. Salmela-Aro K, Kiuru N, Leskinen E, Nurmi JE. School burnout inventory (SBI). Eur J Psychol Assess 2009 Jan;25(1):48-57. [doi: 10.1027/1015-5759.25.1.48]
69. Salmela-Aro K, Upadaya K. The schoolwork engagement inventory. Eur J Psychol Assess 2012 Sep;28(1):60-67. [doi: 10.1027/1015-5759/a000091]
70. Lintunen T. Perceived physical competence scale for children. Scand J Sport Sci 1987;9:57-64.
71. Hirvensalo M, Jaakkola T, Sääkslahti A, Lintunen T. Koettu liikunnallinen pätevyys ja koetut esteet [Book in Finnish]. In: Lasten Ja Nuorten Liikuntakäyttäytyminen Suomessa LIITU-Tutkimuksen Tuloksia 2016: Valtion liikuntaneuvoston julkaisuja; 2016, Vol. 4:36-40 URL: https://www.liikuntaneuvosto.fi/wp-content/uploads/2019/09/LIITU_2016.pdf [accessed 2026-01-23]
72. Metsämuuronen J. Challenges of the Fennema-Sherman test in the international comparisons. IJPS 2012;4(3):1-22. [doi: 10.5539/ijps.v4n3p1]
73. Axivity. URL: https://axivity.com/product/ax3 [accessed 2026-01-29]
74. Schneller MB, Bentsen P, Nielsen G, et al. Measuring children's physical activity: compliance using skin-taped accelerometers. Med Sci Sports Exerc 2017 Jun;49(6):1261-1269. [doi: 10.1249/MSS.0000000000001222] [Medline: 28181981]
75. Jackson D. OMGUI (version 1.0.0.45). Github. 2024. URL: https://github.com/openmovementproject/openmovement/wiki/AX3-GUI [accessed 2026-01-26]
76. Hettiarachchi P, Johansson P. ActiPASS (version 2025.04). Zenodo. 2025. URL: https://zenodo.org/records/15181460 [accessed 2026-01-23]
77. Lendt C, Hettiarachchi P, Johansson PJ, et al. Assessing the accuracy of activity classification using thigh-worn accelerometry: a validation study of ActiPASS in school-aged children. J Phys Act Health 2024 Oct 1;21(10):1092-1099. [doi: 10.1123/jpah.2024-0259] [Medline: 39159934]
78. Rutterford C, Copas A, Eldridge S. Methods for sample size determination in cluster randomized trials. Int J Epidemiol 2015 Jun;44(3):1051-1067. [doi: 10.1093/ije/dyv113] [Medline: 26174515]
79. Campbell IG, Burright CS, Kraus AM, Grimm KJ, Feinberg I. Daytime sleepiness increases with age in early adolescence: a sleep restriction dose-response study. Sleep 2017 May 1;40(5):zsx046. [doi: 10.1093/sleep/zsx046] [Medline: 28419388]
80. Sallinen M, Härmä M, Akila R, et al. The effects of sleep debt and monotonous work on sleepiness and performance during a 12-h dayshift. J Sleep Res 2004 Dec;13(4):285-294. [doi: 10.1111/j.1365-2869.2004.00425.x] [Medline: 15560763]
81. Hillman CH, Pontifex MB, Raine LB, Castelli DM, Hall EE, Kramer AF. The effect of acute treadmill walking on cognitive control and academic achievement in preadolescent children. Neuroscience 2009 Mar 31;159(3):1044-1054. [doi: 10.1016/j.neuroscience.2009.01.057] [Medline: 19356688]
82. Stroth S, Kubesch S, Dieterle K, Ruchsow M, Heim R, Kiefer M. Physical fitness, but not acute exercise modulates event-related potential indices for executive control in healthy adolescents. Brain Res 2009 May 7;1269:114-124. [doi: 10.1016/j.brainres.2009.02.073] [Medline: 19285042]
83. Hirsjärvi S, Hurme H. Tutkimushaastattelu: Teemahaastattelun Teoria Ja Käytäntö: Helsinki University Press; 2000. URL: https://primo.uef.fi/discovery/fulldisplay?docid=alma9925364370966&context=L&vid=358FIN_UOEF:VU1&lang=fi&adaptor=Local%20Search%20Engine&tab=Everything&query=sub,exact,kasvatusfilosofia [accessed 2026-01-23]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials

XSL•FO
RenderX

HRV: heart rate variability
InSitu: InSituations
PA: physical activity
PAAL: Physically Active Academic Lessons
PAL: physically active learning
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
ViLLE: Visual Learning Environment

Edited by S Khan; submitted 27.Oct.2025; peer-reviewed by J Tejada, L Hollarder; revised version received 15.Dec.2025; accepted 19.Dec.2025; published 03.Feb.2026.

Please cite as:
Syväoja HJ, Takalo S, Kukko T, Salmela N, Hakonen H, Kulmala J, Lindfors H, Oksanen H, Räsänen P, Mäkitalo K, Tammelin TH Enhancing Upper Secondary Students' Situational Engagement and Cognitive Prerequisites of Learning Through the Physically Active Academic Lessons Intervention: Protocol for a Mixed Methods Cluster Randomized Individual Crossover Trial
JMIR Res Protoc 2026;15:e84601
URL: https://www.researchprotocols.org/2026/1/e84601
doi:10.2196/84601

© Heidi J Syväoja, Susanna Takalo, Tuomas Kukko, Nina Salmela, Harto Hakonen, Janne Kulmala, Heidi Lindfors, Hermanni Oksanen, Pekka Räsänen, Kati Mäkitalo, Tuija H Tammelin. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 3.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Acupuncture-Combined Tuina on Patients With Knee Osteoarthritis: Protocol for a Randomized Controlled Trial

Chendong Gu[1*], MSc; Dong Wei[2*], PhD; Kaiyue Zhang[3*], MSc; Jinxi Ren[4], BSc; Yihui Qiao[4], BSc; Fengyu Zhang[5], BSc; Yiting Zhu[4], BSc; Zhixian Xu[3], MSc; Yinlong Cao[3], MSc; Shuyi Yao[6], BSc; Hanrui Guo[7], BSc; Yihang Wang[8], MSc; Huan Wang[9], MSc; Lin Wang[10], MSc; Quanliang Wang[11], MSc; Guangxin Guo[2,3,8], PhD

[1]Department of Acupuncture and Moxibustion Massage, Qingpu Traditional Chinese Medicine Hospital, Shanghai, China

[10]Shanghai Pudong New Area People's Hospital, Shanghai, China

[11]Department of Acupuncture-Moxibustion and Tuina, Henan Integrative Medicine Hospital, Zhengzhou, China

[2]Department of Tuina, Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, 274 Zhijiang Middle Road, Jing'an District, Shanghai, Shanghai, China

[3]School of Acupuncture-Moxibustion and Tuina, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[4]School of Integrative Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[5]Yueyang Clinical Medical College, Shanghai University of Traditional Chinese Medicine, Shangai, China

[6]School of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[7]The Third Affiliated Hospital of Henan University of Traditional Chinese Medicine, Henan, China

[8]Shuguang Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[9]Department of Traditional Chinese Medicine, Clinical Medicine, Chongqing University School of Medicine, Chongqing, China

[*]these authors contributed equally

Corresponding Author:
Guangxin Guo, PhD
Department of Tuina, Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, 274 Zhijiang Middle Road, Jing'an District, Shanghai, Shanghai, China

Abstract

Background: Knee osteoarthritis (KOA) is a prevalent degenerative joint disease that is characterized by joint pain, stiffness, and dysfunctional impairment, imposing a substantial medical burden annually. Tuina, a widely used noninvasive traditional Chinese medicine therapy commonly used for KOA, has been used for the management of this condition; however, its additive benefit to acupuncture remains unclear.

Objective: The objective of the study is to assess whether the effectiveness of acupuncture-combined Tuina is superior to that of acupuncture in the management of KOA.

Methods: This single-center, parallel-group, randomized controlled trial aims to enroll 60 patients with KOA, who will be randomly assigned to either a Tuina group (acupuncture-combined Tuina) or an acupuncture group (30 patients per group). Baseline assessments will include demographic and clinical evaluations: the visual analog scale from the short-form McGill Pain Questionnaire, pain threshold, muscle tension, 10-meter walking test, Western Ontario and McMaster Universities Osteoarthritis Index, and multimodal magnetic resonance imaging. Over a 6-week intervention, the Tuina group will receive Tuina on the basis of acupuncture, that is, acupuncture-combined Tuina, while the acupuncture group will undergo just acupuncture, with both groups continuing standard care as prescribed. Posttreatment, clinical outcomes and safety will be reassessed using baseline indicators. A 12-week follow-up will include all clinical evaluations. Assessments will be conducted by blinded assessors, and statistical analyses will be conducted by independent, blinded analysts. Outcomes will evaluate clinical pain and functional differences between groups and will elucidate the underlying cerebral mechanisms.

Results: This study was funded in August 2024. The experimental plan will begin on December 31, 2025, and end on May 4, 2027.

Conclusions: This trial aims to verify whether acupuncture-combined Tuina receives better efficacy than single acupuncture, as well as to explore the neuroimaging mechanisms that are clinically affected, thus providing scientific evidence for the treatment of clinical patients with KOA.

Trial Registration: International Traditional Medicine Clinical Trial Registry ITMCTR2024000635; https://itmctr.ccebtcm.org.cn/mgt/project/view/8026608627430517572

International Registered Report Identifier (IRRID): PRR1-10.2196/84082

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e84082) doi:10.2196/84082

KEYWORDS

knee osteoarthritis; Tuina; acupuncture; pain; stiffness; multimodal magnetic resonance imaging

Introduction

Knee osteoarthritis (KOA) is a prevalent and debilitating condition characterized by joint pain, stiffness, and restricted mobility, significantly impairing daily functioning and work participation [1-3]. In the United Kingdom, employees with arthritis have a 1.35% higher rate of sickness absence than their counterparts without arthritis [4]. The prevalence of KOA is 8.1% in China and 12.1% in the United States and Europe, with rates expected to rise as global populations age [5,6]. Currently, there is no treatment that can completely stop or reverse the progression of osteoarthritis [7,8]. Due to painfulness from the knee joint, patients usually use painkillers, but there is a risk of overuse and side effects of analgesic drugs [7-9]. The effectiveness of acupuncture in treating KOA in improving pain has been preliminarily verified [10,11]. Tuina, a traditional Chinese medicine manipulation therapy, is applied in the clinical treatment of KOA. However, whether the clinical efficacy of acupuncture-combined Tuina is more effective than single acupuncture remains to be proven.

Recent advances in neuroimaging, particularly the use of multimodal magnetic resonance imaging (multi-MRI), have provided powerful tools for investigating changes in brain activity involved in the pathogenesis of KOA, complementing research on biomechanics, aging, and genetic factors [12-14]. Pain, the hallmark symptom of KOA, results from complex interactions between brain regions responsible for pain transmission and modulation [15-17]. Emerging evidence suggests that KOA-associated pain is linked to functional activation and structural alterations in the cerebral cortex and subcortical regions [18,19]. Studies have identified increased resting-state functional connectivity between the ventrolateral periaqueductal gray and the bilateral thalamus in patients with KOA with knee pain compared to healthy controls [20]. Additionally, gray matter (GM) volume reductions in the bilateral insula and hippocampus have been observed [21]. Both static and dynamic functional network connectivity have been shown to correlate with clinical symptoms in KOA [22], highlighting the complex neural mechanisms underlying the condition.

Acupuncture, a core component of traditional Chinese medicine, is recommended in the Chinese Guidelines for Diagnosis and Treatment of Osteoarthritis (2024 Edition) as an effective treatment for alleviating pain and improving joint function in patients with KOA [23]. Studies have demonstrated that acupuncture effectively reduces joint pain and enhances functionality [24,25], with its therapeutic benefits lasting up to 4.5 months compared to sham acupuncture [26]. Evidence suggests [27-30] that acupuncture alleviates KOA by modulating central neural pathways, with its mechanisms varying based on the treatment modality and patient pathology, as reflected in specific cerebral alterations and activation patterns.

Tuina, a noninvasive Chinese therapeutic massage, is widely used for managing KOA due to its direct effects on the musculoskeletal system [31]. Through techniques such as pressing and kneading, Tuina alleviates pain and stiffness, while also improving knee stability, muscle strength, and joint mechanics, thus supporting long-term recovery [23]. Unlike acupuncture, which focuses on neurophysiological pathways, Tuina emphasizes physical manipulation of the musculoskeletal system to address knee joint dysfunction and muscle tension, enhancing joint stability and optimizing movement patterns [32,33]. The process of Tuina tends to be more acceptable to patients, as it is tender, comfortable, and associated with no injury, which is able to effectively relieve the pain, negative emotion, and dysfunction in patients with KOA.

As a result, to compare whether acupuncture-combined Tuina is better than single acupuncture, as well as to reveal the mechanism of analgesic effect that acupuncture-combined Tuina imposes to the brain center, an optimal, efficient, randomized comparative trial will be carried out, aiming to provide more scientific evidence for treating patients with KOA clinically.

Methods

Design

This is a single-center, parallel, randomized controlled trial. On the basis of the criteria established by the American College of Rheumatology [34], a total of 60 participants diagnosed with KOA will be recruited as eligible participants. These participants will be randomly assigned to 2 groups, each consisting of 30 individuals: the Tuina group (acupuncture-combined Tuina) and the acupuncture group. The intervention duration for both groups will be 30 minutes, administered 2 times weekly. The treatment period will last for 6 weeks, followed by a 12-week follow-up period.

Outcome measurements and magnetic resonance imaging scans will be assessed at baseline and at 6 weeks posttreatment, followed by a 12-week clinical follow-up period. Once data collection is completed, analyses will be conducted to evaluate changes in clinical indices and cerebral multi-MRI parameters across groups. The trial flowchart and study schedule for data collection are illustrated in Figure 1 and Table 1, respectively.

XSL•FO
RenderX

Figure 1. Flowchart of the trial. This study is a randomized controlled trial using multimodal magnetic resonance imaging (multi-MRI) and plans to enroll 60 patients with knee osteoarthritis. These patients will be equally allocated to 2 groups: the Tuina group and the acupuncture group, with 30 patients in each group. The trial will include a 6-week treatment period. During the treatment, patients in the Tuina group will receive 12 Tuina sessions based on acupuncture, whereas those in the acupuncture group will only receive acupuncture. Outcome assessments and multi-MRI scans will be conducted at 2 time points: baseline and the end of the treatment (6 weeks later). Subsequently, only clinical outcomes will be evaluated during the 12-week follow-up period.

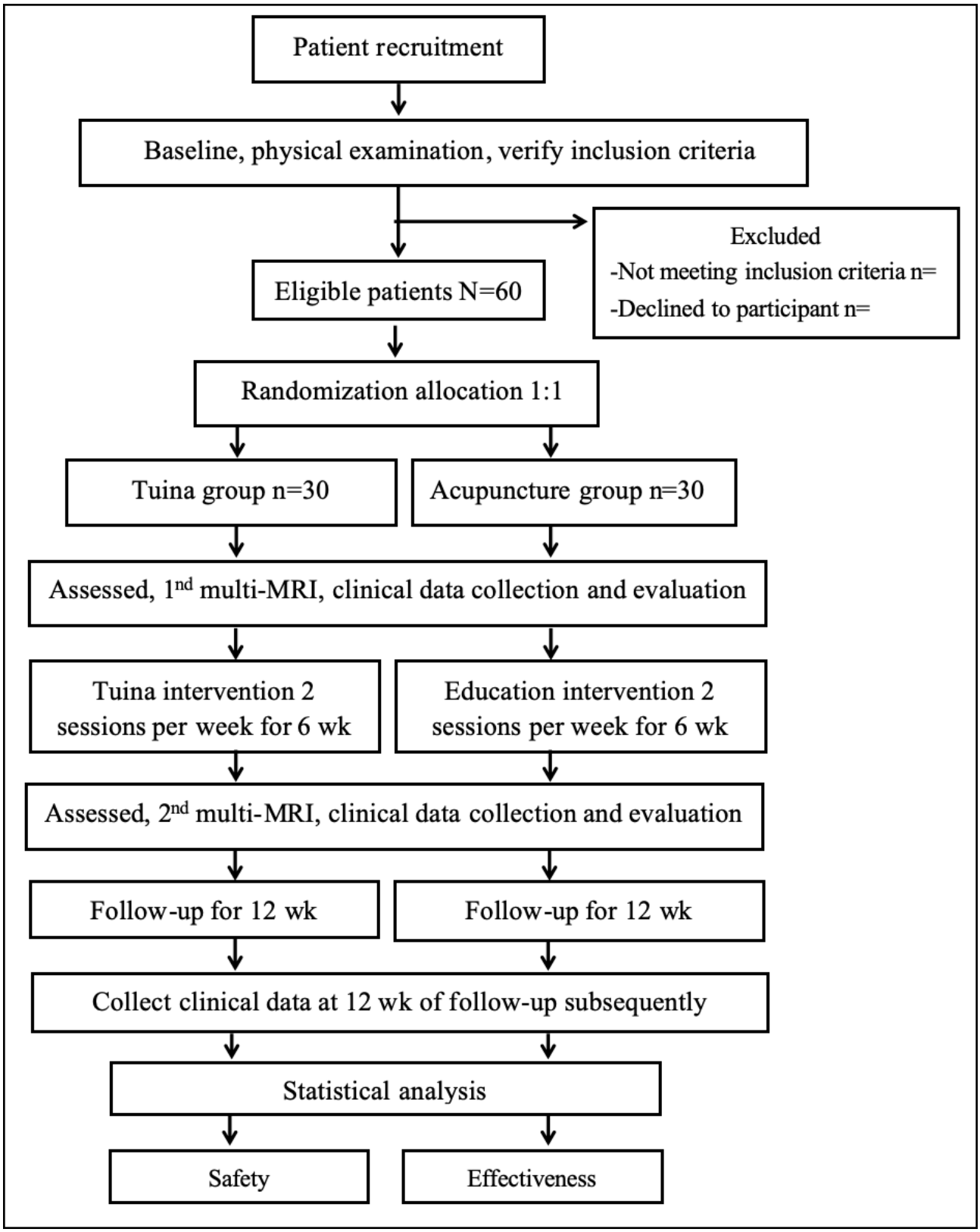

XSL•FO
RenderX

Table . Study schedule for data collection[a].

Period	−2 weeks	0 weeks	0 - 6 weeks	6 weeks	6 - 18 weeks	18 weeks
	Screening	Baseline	Intervention	Assessing	Follow-up	Assessing
Enrollment						
Eligibility	✓[b]					
Demography	✓					
Informed consent	✓					
Sign the informed consent		✓				
Medical history	✓		✓	✓	✓	✓
Physical examination	✓		✓	✓	✓	✓
Randomization		✓				
Interventions						
Tuina group (n=30)			✓			
Acupuncture group (n=30)			✓			
Assessments						
VAS[c]	✓	✓		✓		✓
SF-MPQ[d]	✓	✓		✓		✓
Pain threshold		✓		✓		✓
Muscle tension		✓		✓		✓
10MWT[e]		✓		✓		✓
WOMAC[f]		✓		✓		✓
Clinical efficacy				✓		✓
Multi-MRI[g]		✓		✓		
Trial evaluation						
Patients' compliance				✓		✓
Safety evaluation				✓		✓
Credibility test				✓		✓
Adverse events				✓		✓
Analysis		✓	✓	✓	✓	✓

[a]The informed consent and examination will be conducted after recruitment. Then, matched patients with knee osteoarthritis will be randomized into 2 groups and receive treatment. Both clinical outcomes and multimodal magnetic resonance imaging scans will be performed at 2 time points. Only clinical data will be collected at the 12-week follow-up subsequently. Adverse events will be recorded in the case report form at any time during the study.

[b]✓: used to indicate that the inspection item needs to be executed at the corresponding time point.

[c]VAS: visual analog scale.

[d]SF-MPQ: short-form McGill Pain Questionnaire.

[e]10MWT: 10-meter walking test.

[f]WOMAC: the Western Ontario and McMaster Universities Osteoarthritis Index.

[g]Multi-MRI: multimodal magnetic resonance imaging.

XSL•FO
RenderX

Participants

All patients will be recruited from the Qingpu Traditional Chinese Medicine Hospital, Shanghai, through a combination of posters, online announcements, and promotional pamphlets. The inclusion and exclusion criteria for participants are presented in Textbox 1.

Textbox 1. The inclusion and exclusion criteria for participants.

Inclusion criteria

- Meet the diagnostic criteria for knee osteoarthritis set by the American College of Rheumatology [34].
- Be aged between 40 and 70 years, with left knee pain mainly or left knee pain only, as well as right-hand dominance.
- Have a Kellgren-Lawrence radiological [35] grade of I-II.
- Present with a visual analog scale score >3 from the short-form McGill Pain Questionnaire and a disease duration of at least 3 mo.
- Voluntarily agree to participate in this study and sign the informed consent form.

Exclusion criteria

- Have a history of trauma or knee surgery.
- Have tumors, tuberculosis, osteomyelitis, or other diseases affecting the knee.
- Exhibit severe hepatic or renal dysfunction, severe cardiovascular diseases, diabetes mellitus, or mental illnesses that may interfere with the treatment of Tuina.
- Experience pain in the knee area or throughout the body due to other diseases.
- Have magnetic resonance imaging (MRI) contraindications, such as claustrophobia, pacemakers, defibrillators, heart stents, or intrauterine devices or any other conditions that prevent safe MRI scanning.
- Have skin lesions in the periarticular area of the knee joint.

Sample Size Calculation

This study is designed as a superiority trial, with the sample size calculated using PASS 15 (NCSS, LLC). On the basis of relevant literature, a 2-point difference on the visual analog scale (VAS) score was defined as the superiority margin [36]. According to our preliminary experiments, the mean difference between the acupuncture group and the Tuina group (acupuncture-combined Tuina) is 3.25, with an SD of 1.5 in both groups. Using a superiority test with an α of .025 and a power of 0.8, and accounting for a 20% dropout rate, the calculated sample size is 30 cases per group, resulting in a total of 60 cases required.

Randomization and Allocation Concealment

This study uses blocked randomization to ensure unbiased group allocation. The statistical software SPSS 21.0 will be used to randomly assign participants to either the Tuina group or the acupuncture group. The randomization process will be performed by an independent statistician who will generate 60 random numbers of varying sizes using a computer. This statistician will have no involvement in participant recruitment, assessment, treatment administration, or data analysis, ensuring the independence and impartiality of the process.

Participants will be sequentially numbered from 1 to 60, and blocked randomization will be applied with a block size of 4. Within each block, treatment allocation (Tuina or acupuncture) will be determined by computer-generated random numbers, with smaller numbers assigned to Tuina and larger numbers assigned to acupuncture. The allocation process within each block will be dynamic and unpredictable, ensuring an equal distribution of participants between the 2 groups in a 1:1 ratio.

To ensure allocation concealment, the randomization scheme will be enclosed in opaque, sealed envelopes, numbered sequentially according to participant enrollment. Each envelope will contain the allocation result for the corresponding participant, whereas the outside will remain blank to prevent unintentional disclosure. These envelopes will remain sealed until the treatment allocation is determined, and they will be stored securely to prevent premature access to the allocation information. The assigned treatment (acupuncture-combined Tuina group or the acupuncture group) will be revealed only by opening the envelope corresponding to the participant's number, ensuring the integrity of allocation concealment and enhancing the scientific rigor of the study.

Blinding

This clinical trial will implement blinding for assessment to minimize potential biases. Statistical analysts will be blinded to participant recruitment, allocation, and treatment details to ensure the integrity and reliability of the study results. They will not have access to raw data until after the data collection process is completed and quality control measures are in place.

Due to the nature of Tuina and acupuncture interventions, achieving full blinding of patients and therapists is not feasible [37]. To minimize potential bias, therapists will not be involved in participant information collection, assessment, or subsequent data analysis. They will also refrain from disclosing information regarding the specific benefits or risks of the treatments to patients [38]. Additionally, to mitigate bias, outcome assessors and statisticians will remain blinded to group allocation throughout the study.

XSL•FO
RenderX

Interventions

Overview

Participants in this study will receive a total of 12 treatment sessions over a 6-week period. The interventions will be administered by therapists with at least 6 years of professional experience in Tuina and acupuncture. All therapists will undergo standardized training to ensure consistency in treatment delivery, focusing on session duration, target treatment areas, and the specific techniques used.

To maintain the study's integrity, patients will be prohibited from using concomitant medications or receiving alternative interventions during the treatment period. Celecoxib (Pfizer) will be allowed for pain management, but only if necessary and not within 48 hours before outcome measurements. During the follow-up phase, any medication use or additional interventions will be documented thoroughly to ensure an accurate assessment of the effects of Tuina and acupuncture, thereby isolating their impact on patient outcomes. This approach will guarantee the validity and reliability of the study findings.

Acupuncture Group

In this study, acupuncture treatment will follow a standardized protocol based on traditional Chinese medicine principles, expert consensus, and clinical experience [39]. Participants will receive acupuncture in a private room, with the affected knee being the sole treatment focus for those with osteoarthritis symptoms. Sterile, single-use needles (0.25 mm gauge, 25 - 40 mm in length, Hwato) will be used.

Before the study, all acupuncturists will undergo comprehensive training on standardized procedures, including precise acupoint location and needle manipulation techniques. The acupuncture protocol will involve five core acupoints: Dubi (ST35), Neixiyan (EX-LE5), Ququan (LR8), Xiyangguan (GB33), and an Ashi point, with up to 3 supplementary points chosen based on the affected meridian (Figure 2). Ashi points are defined as tender points within the knee joint region, marked by sensations of soreness, distension, or pain.

Figure 2. Location of acupoints for the acupuncture group. Core acupoints: Dubi (ST35), Neixiyan (EX-LE5), Ququan (LR8), and Xiyangguan (GB33).

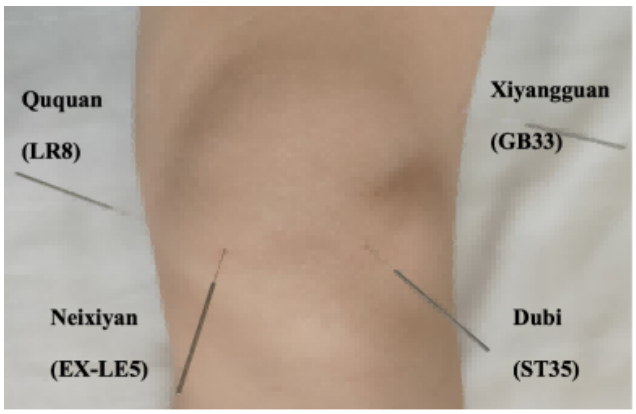

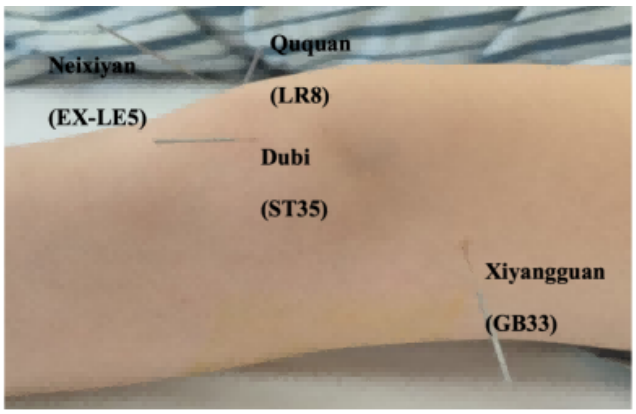

Achieving the De qi sensation, a combination of soreness, numbness, distension, and heaviness, will be required for every session. Participants will receive two 30-minute acupuncture sessions per week for 6 weeks, totaling 12 sessions. Sessions will be spaced at least 48 hours apart.

Tuina Group

Overview

Participants randomized to the Tuina intervention arm will undergo integrated acupuncture and Tuina in private treatment settings, following the same standardized acupuncture protocol as that administered in the acupuncture group. The Tuina protocol will include manual therapy, covering Tuina techniques applied to both sides of the affected limb, as well as passive range-of-motion exercises for the knee joint [40,41].

To ensure consistency and reproducibility across all sessions and practitioners, the core Tuina manipulations (rolling, kneading, and pressing) are formulated primarily according to the *Tuina Therapy–14th Five-Year Plan Textbook* and previous research and will be executed with a fixed order listed below.

Step 1: Rolling in Supine Position

The practitioner's right hand should exert force on the dorsal ulnar side, stick it to the anterior thigh, and roll back and forth through continuous movement of wrist flexion and extension and forearm pronation and supination. The frequency will be approximately 120 times per minute. The manipulation will be repeated from top to bottom 3 to 5 times [42].

Step 2: Kneading and Pressing Technique

This step focuses on the application of kneading and pressing techniques.

- Kneading: use the left or right thrusting thread or the fingertip of the thumb to perform rhythmic spiral movements.
- Pressing: using their thumb or index finger, the physician should press the corresponding acupoints with the middle finger or index finger for 5 to 10 seconds.

Both kneading and pressing techniques will be applied to Heding (EX-LE2), Neixiyan (EX-LE4), Dubi (ST35), Yanglingquan (GB34), Yinlingquan (SP9), Xuehai (SP10), and Ashi points, taking the patient's feeling of soreness, distension, and pain as the degree of tolerance without exerting too much force, aiming to elicit the De qi sensation, which is indicative of effective

XSL•FO
RenderX

manipulation. Following this, patients will transition to a prone position.

Step 3: Rolling in Prone Position

The practitioner will then use the back of the hand to perform rolling motions over the posterior thigh, popliteal fossa, and posterior leg. The operational standard of rolling is identical to the rolling technique in step 1 (Figure 3). This sequence aims to relax the muscles and stimulate the collaterals.

XSL•FO
RenderX

Figure 3. Demonstration of Tuina. (A) Initially, patients will be positioned supine, with the practitioner standing beside them. The practitioner will use the back of the hand to perform rolling motions over the anterior thigh. (B) Subsequently, kneading and pressing techniques will be applied to key acupoints, including Heding (EX-LE2), Neixiyan (EX-LE4), Dubi (ST35), Yanglingquan (GB34), Yinlingquan (SP9), Xuehai (SP10), and Ashi points. (C) The practitioner will then use the back of the hand to perform rolling motions over the posterior thigh, popliteal fossa, and posterior leg. (D) Finally, one hand will be used to apply pressure to the knee joint, while the other hand holds the ankle joint to facilitate passive flexion and extension exercises for the lower limbs for 1 min.

XSL•FO
RenderX

Step 4: Kneading and Passive Joint Manipulating

This step combines kneading with passive joint manipulation.

1. Kneading: The thumb will be used to apply pressure from the posterior thigh to the knee joint, whereas the other 4 fingers and the palm help stabilize the thigh.
2. Passive joint manipulating: When one hand is manipulating the kneading technique, the other hand holds the ankle joint to facilitate passive flexion and extension exercises for the lower limbs for 1 minute.

Kneading and pressing techniques will be applied to Weizhong (BL40) and Chengshan (BL57). In total, participants will receive two 30-minute sessions per week for 6 weeks, totaling 12 sessions, with a minimum interval of 48 hours between treatments.

Follow-Up

Upon completion of the 6-week intervention period, all patients will immediately transition into a 12-week unsupervised follow-up phase. At the end of the 18-week clinical period, patients with KOA will undergo assessments for all clinical evaluations, with the exception of except for the multi-MRI outcomes.

Multi-MRI Examination Procedure

Multi-MRI data will be acquired using a 3.0-T magnetic resonance scanner (SIEMENS MAGNETOM Verio syngo MR B17), equipped with a 32-channel phase-array head coil, at the Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University.

Before the magnetic resonance imaging (MRI) scan, participants will undergo a pre-MRI evaluation procedure, which includes instructions to rest for 10 minutes, relax, and maintain calmness throughout the session. During the scan, participants will be requested to close their eyes, wear earplugs, keep their heads stationary, remain relaxed and awake, and avoid thinking of anything specific.

The imaging protocol will include the following steps:

- High-resolution T1-weighted image: A spin-echo sequence will be used to acquire a transverse sagittal scan with a flip angle of 9°, a repetition time (TR) of 1900 ms, an echo time (TE) of 2.93 ms, a field of view of 256×256 mm², 160 slices, and a slice thickness of 1 mm.
- Diffusion tensor imaging (DTI): DTI will be performed using axial DTI mapping sequences with a TR of 10,000 ms, a TE of 89 ms, a matrix size of 240×240, a slice thickness of 2 mm, B-values of 0 and 1000 s/mm², and 30 directions.
- Blood oxygenation level–dependent resting-state functional imaging: An echo-planar imaging sequence will acquire blood oxygenation level dependent–resting-state functional images with a coronal axial scan, 33 slices, a slice thickness of 4 mm, a TE of 30 ms, a TR of 2000 ms, a field of view of 220×220 mm², a voxel size of 3.4×3.4×4.0 mm³, a flip angle of 90°, and a scan duration of 8 minutes and 8 seconds, resulting in 240 volumes.

Following the MRI procedure, participants will be given an additional 5-minute rest period and may leave if they experience no discomfort.

Outcome Measurements

Assessment Tools and Schedule

Clinical outcomes will be evaluated using 2 self-report questionnaires, 4 instrument-based indicators measured at baseline and after 6 weeks of treatment, and multi-MRI data. At the conclusion of the full 18-week clinical period, all clinical assessments, excluding multi-MRI outcomes, will be performed. The responses will provide valuable information on pain levels, sensory experiences, emotional states, and overall quality of life.

Primary Outcomes

The VAS score is based on a 10-cm long straight line, with the starting end marked as 0 points, indicating no pain, and the end marked as 10 points, indicating unbearable severe pain. Patients will select a point on the line that represents their current level of pain. A lower total score indicates reduced pain severity, reflecting the effectiveness of the interventions. This VAS is distinct from the VAS component within the short-form McGill Pain Questionnaire (SF-MPQ), which is used solely for screening purposes and for baseline characterization.

Secondary Outcomes

Short-Form McGill Pain Questionnaire

The SF-MPQ will be used to assess the multidimensional experience of pain. It consists of 2 primary components.

Descriptive Pain Rating

This component comprises a list of 15 pain descriptors, categorized into 11 items evaluating sensory qualities of pain (eg, throbbing and shooting) and 4 items evaluating affective qualities (eg, tiring and fearful). Each descriptor is rated on an intensity scale ranging from 0 (none) to 3 (severe). The scores are summed to yield a sensory score (range 0 - 33), an affective score (range 0 - 12), and a total pain rating index (range 0 - 45).

Present Pain Intensity

The present pain intensity is a single-item, 6-point verbal rating scale (0=no pain, 5=excruciating) that provides a global assessment of current pain intensity [43].

The pain intensity scores derived from the SF-MPQ will not be used as outcome measures in the longitudinal analysis.

Pain Threshold

Pain threshold will be assessed at specific knee points—EX-LE02 (Heding), EX-LE04 (Neixiyan), EX-LE05 (Xiyan), and the area beneath the tip of the patella—using a pressure algometer. This measure will evaluate the sensitivity to mechanical pressure pain in patients with KOA [44,45].

Muscle Tension

Muscle tension will be assessed in the quadriceps and gastrocnemius muscles, providing insights into muscle stiffness in patients with KOA [46].

XSL•FO
RenderX

10-Meter Walking Test

The 10-meter walking test (10MWT) will assess walking speed by timing the patients over a 10-m distance, with a focus on the middle 6 m to account for acceleration and deceleration. The test will begin when the toe of the leading foot crosses the 2-m mark and will end when the toe crosses the 8-m mark. Three trials will be conducted, and the average time will be recorded.

Western Ontario and McMaster Universities Arthritis Index

The Western Ontario and McMaster Universities Arthritis Index (WOMAC) questionnaire will evaluate symptoms of KOA in 3 domains: pain (5 questions), stiffness (2 questions), and functional limitations during daily activities (17 questions) [47].

Clinical Efficacy

Clinical efficacy will be categorized into 4 levels: cure, markedly effective, effective, and ineffective [48].

- Cure: complete resolution of pain and symptoms, with normal joint activity
- Markedly effective: resolution of pain and symptoms, with no limitation in joint activity
- Effective: significant reduction in pain and symptoms, with slight joint activity limitation
- Ineffective: no significant improvement in pain, symptoms, or joint activity

Multi-MRI Outcomes

Functional Imaging

Functional connectivity analysis will be used to assess brain network activity and interregional communication, highlighting potential neural mechanisms underlying pain perception and motor function in KOA.

Diffusion Imaging

Fractional anisotropy and mean diffusivity will be measured to assess white matter (WM) integrity, providing insights into neural pathway alterations in patients with KOA.

Structural Imaging

GM density, cortical thickness, and volumes of subcortical nuclei will be evaluated to investigate brain structural changes related to KOA.

Safety Evaluation and Adverse Events

Participant safety will be actively monitored throughout the trial. At each treatment session and assessment visit, practitioners will systematically inquire about and record any adverse events (AEs). The following treatment-related AEs will be proactively monitored:

- For acupuncture: significant needle pain, bleeding, hematoma or bruising (>2 cm), dizziness or fainting, and local infection [49]
- For Tuina: transient (>24 h) increases in pain, significant bruising, skin irritation, and unexpected joint discomfort [41]
- General: any other unexpected AE reported by the participant [42]

All AEs will be recorded in the case report form (CRF) with details on severity, duration, action taken, and outcome. The causality relationship to the intervention will be assessed by the investigator. The incidence and type of AEs will be summarized and compared between the 2 treatment groups as a key safety outcome. All AEs will be managed, recorded, and reported in accordance with the study protocol and the requirements of the Ethics Committee.

Data Management

Baseline data will be collected using CRFs, which will include clinical observations, outcome measures, AEs, and safety evaluations. Outcome assessors, who are independent of the data management team, will ensure impartiality in evaluations, completing forms based on anonymous participant identifiers to maintain blinding. Completed CRFs will be entered into a Microsoft Excel database by 2 independent data administrators who are blinded to group allocations and trained in data monitoring. An independent data manager will oversee data entry into the International Traditional Medicine Clinical Trial Registry.

Data will be stored securely in both physical and digital formats, with access limited to the data manager. Upon trial completion, anonymization procedures will be implemented to preserve participant confidentiality, and an independent data analyst will perform statistical analysis on the anonymized data, ensuring objectivity and transparency.

Quality Control

Rigorous quality control measures will be enforced by the steering committee. All researchers will undergo training in trial methodologies and monitoring techniques before participation. Any changes to the study protocol will be communicated to both the steering and ethics committees to ensure adherence to the original plan. The committees will also oversee recruitment, the intervention protocols (Tuina and acupuncture), and the quality of data collection.

Statistical Analysis

All statistical analyses will be conducted using IBM SPSS 27.0 and R software (version 4.3.0), with a 2-sided P value of $<.05$ considered statistically significant. Analyses will commence after the completion of follow-up for the last participant and will adhere to the intention-to-treat principle.

Changes in the primary outcome (VAS) and key secondary outcomes (WOMAC, 10MWT, pain threshold, and muscle tension) over time will be analyzed using linear mixed-effects models. These models will directly test whether the pattern of change from baseline through the 6-week and 18-week assessments differs between the Tuina and acupuncture groups (ie, a group×time interaction). If this interaction is significant, post hoc tests (ie, Bonferroni) will be conducted to identify the time points at which the groups differ.

The mixed-effects models appropriately handle missing data under the assumption that they are missing at random. A sensitivity analysis using multiple imputation will also be conducted to confirm the robustness of the primary results. Baseline characteristics will be summarized using means (SDs)

XSL•FO
RenderX

or medians (IQRs). Differences between the 2 groups at baseline will be tested using *t* tests or Mann-Whitney *U* tests for continuous variables and chi-square tests for categorical variables. Correlations between changes in different outcomes will be assessed using Pearson or Spearman coefficients.

Analysis of Multi-MRI Data

Structural MRI data will be processed using FSL tools (FMRIB Software Library). SIENAX will calculate normalized volumes of neocortical GM, total GM, and WM [50], while subcortical region volumes (eg, hippocampus, thalamus) will be estimated using FMRIB's integrated registration and segmentation tool [51]. The cortical thickness will be measured at each vertex using FreeSurfer [52].

The DTI pipeline will follow FSL's diffusion toolbox to generate fractional anisotropy and mean diffusivity images, with voxelwise tract-based spatial statistics applied for group comparisons. Resting-state functional magnetic resonance imaging data preprocessing and functional connectivity analyses will be performed using SPM12 software within MATLAB 2013b (MathWorks). This rigorous, standardized approach ensures reliability and accuracy in statistical and neuroimaging analyses.

Patient and Public Involvement

The design, conduct, reporting, and dissemination plans of this study will be carried out without the involvement of any patients or members of the public.

Ethical Considerations

This study was approved by the Ethics Committee of Qingpu Traditional Chinese Medicine Hospital, Shanghai (2024BL0098), and is registered with the International Traditional Medicine Clinical Trial Registry (ITMCTR2024000635). Written informed consent will be obtained from all participants before the commencement of data collection. In the event of adverse reactions during the study, participants will receive appropriate medical care tailored to their specific needs. The findings from this trial will be shared at relevant conferences focusing on KOA, ranging from local to international levels. A detailed manuscript will be prepared and submitted to a peer-reviewed journal for publication, ensuring the results reach the broader scientific community. The primary outcomes will also be communicated directly to participants. Additionally, tailored dissemination efforts, including professional workshops, public lectures, and online platforms, will target researchers, clinicians, and health care providers to maximize the practical application of the study's findings. This approach will ensure transparency and accessibility, irrespective of the study's results.

This study does not collect any information that can directly identify the personal identity of participants (such as name, ID number, address, telephone number, etc.). All research data (including clinical scale scores, behavioral test results, and deidentified neuroimaging data) will be identified and processed with study numbers after collection. The data will be stored on a password-protected independent research computer, accessible only to members of our research team with necessary knowledge. When the research results are published, no personal information that may be traced to the participants will be disclosed. Participation in this study is entirely voluntary and does not provide any form of economic reward or compensation to participants.

Results

This study was funded in August 2024. This experiment is currently not open for recruitment. The experimental plan will begin on December 31, 2025, and will end on May 4, 2027.

Discussion

Anticipated Findings

The primary characteristics of KOA are pain, stiffness, and limited mobility, which significantly impair patients' quality of life and daily activities [53,54]. Pain management constitutes a core component of KOA treatment. On note, patients may continue to experience pain even after surgical intervention, as emphasized in both the Osteoarthritis Research Society International guidelines (2014) [55] and the Chinese Guidelines for the Diagnosis and Treatment of Osteoarthritis (2018) [7]. Tuina and acupuncture have been widely used in the treatment of clinical symptoms, such as pain in KOA [56-59]. Although more evidence is still needed to support the effectiveness of these therapies in relieving KOA symptoms, Chinese Tuina, particularly, as a typical representative of noninvasive and nonpharmacological therapies in traditional Chinese medicine, has received increasing attention and verification for its clinical effects in reducing pain and improving function in patients with KOA [56-60]. Therefore, this study will evaluate the clinical efficacy of acupuncture-combined Tuina in the treatment of KOA, through clinical scale assessment, behavioral assessment, and the application of multimodal neuroimaging techniques. It will also deeply explore its central nervous system analgesic mechanisms, providing more evidence for its clinical application in KOA.

The integration of validated clinical scales and behavioral tests in this trial provides a multidimensional framework for evaluating KOA interventions. The VAS, 36-Item Short Form Health Survey, and WOMAC are accepted to evaluate patients' pain intensity and life standards. Although subjective in nature, these scales mentioned earlier are widely used and universally recognized treatments [17,18]. Complementarily, the WOMAC index dissects symptom impact into pain, stiffness, and functional domains, offering granularity beyond global pain scores [47]. The inclusion of the WOMAC scale aligns with international guidelines and better reflects the comprehensive clinical assessment of KOA symptoms. In terms of behavioral assessments, the 10MWT is used to quantify walking speed, which serves as a proxy for real-world mobility limitations in patients with KOA [61]. Reduced walking time correlates not only with biomechanical improvement but also with decreased fear of movement, reflecting neuromotor integration [46,53]. Pain threshold (via algometry) and muscle tension measurements target peripheral mechanisms. Elevated pressure pain thresholds may indicate reduced peripheral sensitization, while diminished quadriceps tension suggests reduced muscle coordination—both

XSL•FO
RenderX

potentially mediated by descending pain modulation pathways [44,62]. Although this pragmatic study comparing 2 active treatment approaches could not implement blinding for participants and struggled to fully distinguish the specific biomechanical effects of massage from nonspecific factors (such as increased therapeutic attention and contact time), it directly addresses a core question in clinical practice: whether combined massage therapy yields additional benefits in real-world treatment scenarios.

This study used multi-MRI technology to explore the central analgesic brain mechanisms of Tuina intervention for KOA. This approach was chosen due to its advantages of high spatial resolution, absence of radiation, fast imaging speed, and noninvasiveness [63-65], enabling the assessment of brain GM structure, brain function, and WM integrity. KOA pain may be associated with morphological changes in subcortical structures and nuclei, representing part of the neuroimaging pathogenesis [66,67]. DTI can be used to quantify microstructural and macrostructural changes in WM [68]. Functional magnetic resonance imaging allows for in-depth analysis of patients' pain perception and brain regulatory mechanisms for motor control. It is widely applied in research on the central mechanisms of Tuina analgesia [69-73] and has preliminarily identified functional changes in key brain regions, such as the amygdala, hypothalamus, nucleus accumbens, and hippocampus [74-76].

Current research has confirmed that both acupuncture and Tuina can improve clinical symptoms in patients with KOA, such as pain and dysfunction. The onset of clinical efficacy may be related to improvements in brain functional activity [40,62,77-81]. However, there is a lack of comprehensive, multidimensional research integrating brain GM structure, WM structure, and brain functional behavior for the combined treatment of KOA with Tuina and acupuncture.

The study's extended follow-up period, including assessments at 12 weeks posttreatment, enhances its design by providing insights into the long-term effects of Tuina and acupuncture for KOA. Rigorous blinding protocols for outcome assessors and statisticians further ensure the reliability and validity of the findings. Unified training for all practitioners and assessors promotes consistency in intervention delivery, data collection, and analysis, thereby enhancing methodological robustness [82].

Study Strengths and Limitations

This study will be a single-center, strictly randomized controlled clinical trial comparing Tuina combined with acupuncture versus acupuncture alone. For the first time, it will use multi-MRI to explore the central nervous system's analgesic mechanisms of Tuina combined with acupuncture in treating KOA. The study will implement blinding measures for evaluators and statisticians, and the clinical observation period is relatively long, including a 6-week intervention period and a 12-week follow-up period.

This study has certain limitations. Due to the nature of the therapies, complete blinding of participants and therapists to the treatment assignment was not feasible. More specifically, a key consideration is that both groups received active acupuncture, while the Tuina group received additional manual therapy. This difference in treatment intensity and contact time means that nonspecific factors may contribute to the outcomes observed in the combined treatment group, and their effects cannot be isolated from the specific biomechanical effects of Tuina. This is a recognized challenge in trials comparing complex, multicomponent nonpharmacological interventions. Nevertheless, by implementing blinding of outcome assessors and statisticians and by using objective neuroimaging biomarkers, we have sought to mitigate bias and gain insights into the central effects of the treatment package. The findings from this first-of-its-kind multimodal neuroimaging study will provide crucial preliminary evidence and generate specific hypotheses to guide the design of future, more controlled studies aimed at dissecting the specific mechanisms of Tuina.

Summary

This randomized controlled trial investigates the efficacy of acupuncture-combined Tuina for treating KOA and explores its neuroimaging-based central analgesic mechanisms, thereby providing a scientific basis for its clinical application.

Acknowledgments

The authors would like to express their heartfelt gratitude to all trial participants, including study coordinators, participants, and their families, for their invaluable contributions and unwavering dedication to this trial. Generative artificial intelligence tools were used only for language polishing and structural improvements under author supervision. All content was reviewed and verified by the authors before submission.

Funding

This work was supported by the Traditional Chinese Medicine Research Project of Shanghai Municipal Health Commission (2022QN066), National Natural Science Foundation of China (82474666, 82105042), Fang Min Shanghai Famous Traditional Chinese Medicine Inheritance Studio (YC-2023 - 0133), and the 18th batch of Innovation and Entrepreneurship Projects of Shanghai University of Traditional Chinese Medicine (202510268035). The funders have no business with the design and conduct of the trial.

Data Availability

Data sharing is not applicable to this article as no datasets have been generated or analyzed yet.

XSL•FO
RenderX

Authors' Contributions

GG, QW, LW, and CG designed the study protocol and prepared the initial draft of the manuscript. CG, DW, KZ, JR, YQ, FZ, YZ, ZX, YC, SY, HG, YW, HW, LW, QW, and GG contributed to the revision of the manuscript. HG, YW, and HW are responsible for overseeing the management of the study. All authors have reviewed and approved the final version of the manuscript.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Checklist 1
SPIRIT 2025 checklist.
[PDF File, 63 KB - resprot_v15i1e84082_app1.pdf]

References

1. Guo G, Wu B, Xie S, et al. Effectiveness and safety of Tai Chi for chronic pain of knee osteoarthritis: a protocol for systematic review and meta-analysis. Medicine (Baltimore) 2022 Jan 14;101(2):e28497. [doi: 10.1097/MD.0000000000028497] [Medline: 35029201]
2. Tore NG, Oskay D, Haznedaroglu S. The quality of physiotherapy and rehabilitation program and the effect of telerehabilitation on patients with knee osteoarthritis. Clin Rheumatol 2023 Mar;42(3):903-915. [doi: 10.1007/s10067-022-06417-3] [Medline: 36279075]
3. Xu H, Zhao C, Guo G, et al. The effectiveness of Tuina in relieving pain, negative emotions, and disability in knee osteoarthritis: a randomized controlled trial. Pain Med 2023 Mar 1;24(3):244-257. [doi: 10.1093/pm/pnac127] [Medline: 35997590]
4. Whittaker W, Higgerson J, Eden M, Payne K, Wilkie R, Verstappen SM. Effects of employees living with an “arthritis” on sickness absence and transitions out of employment: a comparative observational study in the UK. RMD Open 2024 Nov 29;10(4):e004817. [doi: 10.1136/rmdopen-2024-004817] [Medline: 39615887]
5. Ni F, Zhang Y, Peng X, Li J. Correlation between osteoarthritis and monocyte chemotactic protein-1 expression: a meta-analysis. J Orthop Surg Res 2020 Nov 10;15(1):516. [doi: 10.1186/s13018-020-02045-2] [Medline: 33168099]
6. Guo G, Wang Y, Xu X, et al. Effectiveness of Yijinjing exercise in the treatment of early-stage knee osteoarthritis: a randomized controlled trial protocol. BMJ Open 2024 Mar 7;14(3):e074508. [doi: 10.1136/bmjopen-2023-074508] [Medline: 38453194]
7. Joint Surgery Group, Society of Osteology, Chinese Medical Association. Guidelines for the Diagnosis and Treatment of Osteoarthritis (2018 Edition) [Article in Chinese]. Chin J Orthop 2018;38(12):705-715. [doi: 10.3760/cma.j.issn.0253-2352.2018.12.001]
8. Moseng T, Vliet Vlieland TPM, Battista S, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis: 2023 update. Ann Rheum Dis 2024 May 15;83(6):730-740. [doi: 10.1136/ard-2023-225041] [Medline: 38212040]
9. Huang D, Feng Z, Hu R, Joint Surgery Group, Society of Osteology, Chinese Medical Association. Guidelines for the Diagnosis and Treatment of Osteoarthritis. Chin J Pain Med 2024;20(3):323-338. [doi: 10.3760/cma.j.cn101658-20240618-00074]
10. Lee SK, Oh EJ, Bae GH, et al. Integrative Korean medicine treatment for inpatients with osteoarthritis of the knee: an observational study. Medicine (Baltimore) 2024 Aug 9;103(32):e39183. [doi: 10.1097/MD.0000000000039183] [Medline: 39121270]
11. Liu CY, Tu JF, Lee MS, et al. Is acupuncture effective for knee osteoarthritis? A protocol for a systematic review and meta-analysis. BMJ Open 2022 Jan 11;12(1):e052270. [doi: 10.1136/bmjopen-2021-052270] [Medline: 35017242]
12. Tschopp M, Pfirrmann CWA, Brunner F, et al. Morphological and quantitative parametric MRI follow-up of cartilage changes before and after intra-articular injection therapy in patients with mild to moderate knee osteoarthritis: a randomized, placebo-controlled trial. Invest Radiol 2024 Sep 1;59(9):646-655. [doi: 10.1097/RLI.0000000000001067] [Medline: 38421679]
13. Chang SH, Mori D, Kobayashi H, et al. Excessive mechanical loading promotes osteoarthritis through the gremlin-1–NF-κB pathway. Nat Commun 2019 Mar 29;10(1):1442. [doi: 10.1038/s41467-019-09491-5] [Medline: 30926814]
14. Lu F, Liu P, Zhang Q, Wang W, Guo W. Association between the polymorphism of IL-17A and IL-17F gene with knee osteoarthritis risk: a meta-analysis based on case-control studies. J Orthop Surg Res 2019 Dec 16;14(1):445. [doi: 10.1186/s13018-019-1495-0] [Medline: 31842922]
15. Yue H, Li Y, Ma J, et al. Effect of Tai Chi on knee pain and muscle strength in middle-aged and older adults with knee osteoarthritis: a randomized controlled trial protocol. BMC Complement Med Ther 2023 Jul 20;23(1):256. [doi: 10.1186/s12906-023-04070-0] [Medline: 37474949]

XSL•FO
RenderX

16. Kim JA, Bosma RL, Hemington KS, et al. Sex-differences in network level brain dynamics associated with pain sensitivity and pain interference. Hum Brain Mapp 2021 Feb 15;42(3):598-614. [doi: 10.1002/hbm.25245] [Medline: 33068500]
17. Löfgren M, Opava CH, Demmelmaier I, et al. Pain sensitivity at rest and during muscle contraction in persons with rheumatoid arthritis: a substudy within the physical activity in rheumatoid arthritis 2010 study. Arthritis Res Ther 2018 Mar 15;20(1):48. [doi: 10.1186/s13075-018-1513-3] [Medline: 29544539]
18. Dahmani D, Taik FZ, Berrichi I, Fourtassi M, Abourazzak FE. Impact of central sensitization on pain, disability and psychological distress in patients with knee osteoarthritis and chronic low back pain. BMC Musculoskelet Disord 2023 Nov 10;24(1):877. [doi: 10.1186/s12891-023-07019-z] [Medline: 37950225]
19. Soni A, Wanigasekera V, Mezue M, et al. Central sensitization in knee osteoarthritis: relating presurgical brainstem neuroimaging and PainDETECT-based patient stratification to arthroplasty outcome. Arthritis Rheumatol 2019 Apr;71(4):550-560. [doi: 10.1002/art.40749] [Medline: 30295432]
20. Zhou J, Zeng F, Cheng S, et al. Modulation effects of different treatments on periaqueductal gray resting state functional connectivity in knee osteoarthritis knee pain patients. CNS Neurosci Ther 2023 Jul;29(7):1965-1980. [doi: 10.1111/cns.14153] [Medline: 36890655]
21. Guo H, Wang Y, Qiu L, et al. Structural and functional abnormalities in knee osteoarthritis pain revealed with multimodal magnetic resonance imaging. Front Hum Neurosci 2021;15:783355. [doi: 10.3389/fnhum.2021.783355] [Medline: 34912202]
22. Cheng S, Zeng F, Zhou J, et al. Altered static and dynamic functional brain network in knee osteoarthritis: a resting-state functional magnetic resonance imaging study: static and dynamic FNC in KOA. Neuroimage 2024 Apr 15;292:120599. [doi: 10.1016/j.neuroimage.2024.120599] [Medline: 38608799]
23. Osteoarthrosis Committee of Chinese Aging Well Association. Guideline for diagnosis and non-surgical treatment of early-stage knee osteoarthritis (2024 edition). Zhonghua Yi Xue Za Zhi 2024 Aug 13;104(31):2895-2909. [doi: 10.3760/cma.j.cn112137-20240503-01035] [Medline: 39118337]
24. Huang C, Wang Z, Xu X, Hu S, Zhu R, Chen X. Does acupuncture benefit delayed-onset muscle soreness after strenuous exercise? A systematic review and meta-analysis. Front Physiol 2020;11:666. [doi: 10.3389/fphys.2020.00666] [Medline: 32765287]
25. Lin JG, Kotha P, Chen YH. Understandings of acupuncture application and mechanisms. Am J Transl Res 2022;14(3):1469-1481. [Medline: 35422904]
26. Chen H, Shi H, Gao S, et al. Durable effects of acupuncture for knee osteoarthritis: a systematic review and meta-analysis. Curr Pain Headache Rep 2024 Jul;28(7):709-722. [doi: 10.1007/s11916-024-01242-6] [Medline: 38635021]
27. Qu Y, Peng Y, Xiong Y, Dong X, Ma P, Cheng S. Acupuncture-related therapy for knee osteoarthritis: a narrative review of neuroimaging studies. J Pain Res 2024;17:773-784. [doi: 10.2147/JPR.S450515] [Medline: 38435748]
28. Gollub RL, Kirsch I, Maleki N, et al. A functional neuroimaging study of expectancy effects on pain response in patients with knee osteoarthritis. J Pain 2018 May;19(5):515-527. [doi: 10.1016/j.jpain.2017.12.260] [Medline: 29325883]
29. Chen X, Spaeth RB, Freeman SG, et al. The modulation effect of longitudinal acupuncture on resting state functional connectivity in knee osteoarthritis patients. Mol Pain 2015 Oct 29;11:67. [doi: 10.1186/s12990-015-0071-9] [Medline: 26511911]
30. Park JY, Kim YK, Kim SY, et al. Acupuncture modulates brain neural activity in patients: a systematic review and meta-analysis. Orient Pharm Exp Med 2017 Jun;17(2):111-126. [doi: 10.1007/s13596-017-0266-x]
31. Deng X, Zhu H, Shi L, et al. Comparison of the efficacy of acupuncture with tuina with acupuncture-only in the treatment of peripheral facial paralysis: a network meta-analysis. Intern Emerg Med 2024 Apr;19(3):839-858. [doi: 10.1007/s11739-024-03562-2] [Medline: 38483737]
32. Zhao RL, Ma PH, Liu BY, et al. Short-term and long-term effectiveness of acupuncture and Tuina on knee osteoarthritis: study protocol for a randomized controlled trial. Front Neurol 2023;14:1301217. [doi: 10.3389/fneur.2023.1301217] [Medline: 38152644]
33. Liu Z, Wang H, Yu T, et al. A review on the mechanism of Tuina promoting the recovery of peripheral nerve injury. Evid Based Complement Alternat Med 2021;2021:6652099. [doi: 10.1155/2021/6652099] [Medline: 34285705]
34. Altman RD. Classification of disease: osteoarthritis. Semin Arthritis Rheum 1991 Jun;20(6 Suppl 2):40-47. [doi: 10.1016/0049-0172(91)90026-v] [Medline: 1866629]
35. Boer CG, Radjabzadeh D, Medina-Gomez C, et al. Intestinal microbiome composition and its relation to joint pain and inflammation. Nat Commun 2019 Oct 25;10(1):4881. [doi: 10.1038/s41467-019-12873-4] [Medline: 31653850]
36. Knight J, Rangnekar R, Richardson D, McIlroy S, Bell D, Ahmed AI. Does low-back pain improve after decompressive spinal surgery? A prospective observational study from the British Spine Registry. J Neurosurg Spine 2023 Oct 1;39(4):583-592. [doi: 10.3171/2023.5.SPINE23116] [Medline: 37486861]
37. van der Ende NA, Roozenbeek B, Broderick JP, Khatri P, Lingsma HF, Dippel DW. Blinding of outcome assessors and its association with outcome in a randomized open-label stroke trial. Int J Stroke 2023 Jun;18(5):562-568. [doi: 10.1177/17474930221131706] [Medline: 36169032]
38. Vela MB, Erondu AI, Smith NA, Peek ME, Woodruff JN, Chin MH. Eliminating explicit and implicit biases in health care: evidence and research needs. Annu Rev Public Health 2022 Apr 5;43(1):477-501. [doi: 10.1146/annurev-publhealth-052620-103528] [Medline: 35020445]

XSL•FO
RenderX

39. Sun N, Wang LQ, Shao JK, et al. An expert consensus to standardize acupuncture treatment for knee osteoarthritis. Acupunct Med 2020 Oct;38(5):327-334. [doi: 10.1177/0964528419900789] [Medline: 32309995]
40. Guo G, Kong Y, Zhu Q, et al. Cerebral mechanism of Tuina analgesia in management of knee osteoarthritis using multimodal MRI: study protocol for a randomised controlled trial. Trials 2022 Aug 19;23(1):694. [doi: 10.1186/s13063-022-06633-x] [Medline: 35986403]
41. Ma J, Guo G, Yue H, et al. Tuina on knee pain and functional decline of lower limbs for patients with mild-to-moderate knee osteoarthritis in Shanghai: protocol for a multicentre, assessor-blinded, randomised controlled trial. BMJ Open 2024 Jun 12;14(6):e083440. [doi: 10.1136/bmjopen-2023-083440] [Medline: 38866576]
42. Liu K, Zhan Y, Zhang Y, et al. Efficacy and safety of Tuina (Chinese therapeutic massage) for knee osteoarthritis: a randomized, controlled, and crossover design clinical trial. Front Med (Lausanne) 2023;10:997116. [doi: 10.3389/fmed.2023.997116] [Medline: 37089612]
43. Wang HD, Wang HY, Yang LY, Liu JJ, Zhang T, Zheng J. Acupotomy at trigger points combined with meridian sinew manipulation in treatment of knee osteoarthritis: a randomized controlled trial. Zhen Ci Yan Jiu 2025 Jul 25;50(7):822-828. [doi: 10.13702/j.1000-0607.20240549] [Medline: 40691033]
44. Arnstad ED, Iversen JM, Uglem M, et al. Pain sensitivity in young adults with juvenile idiopathic arthritis: a quantitative sensory testing study. Arthritis Res Ther 2020 Nov 5;22(1):262. [doi: 10.1186/s13075-020-02345-2] [Medline: 33153493]
45. Pomares FB, Roy S, Funck T, et al. Upregulation of cortical GABAA receptor concentration in fibromyalgia. Pain 2020 Jan;161(1):74-82. [doi: 10.1097/j.pain.0000000000001707] [Medline: 31569142]
46. Sánchez-Sánchez J, García-Unanue J, Hernando E, et al. Repeated sprint ability and muscular responses according to the age category in elite youth soccer players. Front Physiol 2019;10:175. [doi: 10.3389/fphys.2019.00175] [Medline: 30894819]
47. Sit RWS, Wu RWK, Rabago D, et al. Efficacy of intra-articular hypertonic dextrose (prolotherapy) for knee osteoarthritis: a randomized controlled trial. Ann Fam Med 2020 May;18(3):235-242. [doi: 10.1370/afm.2520] [Medline: 32393559]
48. Li S, Huang M, Wu G, et al. Efficacy of Chinese herbal formula Sini Zuojin decoction in treating gastroesophageal reflux disease: clinical evidence and potential mechanisms. Front Pharmacol 2020;11:76. [doi: 10.3389/fphar.2020.00076] [Medline: 32174826]
49. Bäumler P, Zhang W, Stübinger T, Irnich D. Acupuncture-related adverse events: systematic review and meta-analyses of prospective clinical studies. BMJ Open 2021 Sep 6;11(9):e045961. [doi: 10.1136/bmjopen-2020-045961] [Medline: 34489268]
50. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. Neuroimage 2004;23 Suppl 1:S208-S219. [doi: 10.1016/j.neuroimage.2004.07.051] [Medline: 15501092]
51. Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. Neuroimage 2011 Jun 1;56(3):907-922. [doi: 10.1016/j.neuroimage.2011.02.046] [Medline: 21352927]
52. FreeSurfer. URL: http://surfer.nmr.mgh.harvard.edu [accessed 2026-01-20]
53. Aw NMY, Yeo SJ, Wylde V, et al. Impact of pain sensitisation on the quality of life of patients with knee osteoarthritis. RMD Open 2022 Mar;8(1):e001938. [doi: 10.1136/rmdopen-2021-001938] [Medline: 35296529]
54. Mo L, Jiang B, Mei T, Zhou D. Exercise therapy for knee osteoarthritis: a systematic review and network meta-analysis. Orthop J Sports Med 2023 May;11(5):23259671231172773. [doi: 10.1177/23259671231172773] [Medline: 37346776]
55. McAlindon TE, Bannuru RR, Sullivan MC, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthritis Cartilage 2014 Mar;22(3):363-388. [doi: 10.1016/j.joca.2014.01.003] [Medline: 24462672]
56. Wang M, Liu L, Zhang CS, et al. Mechanism of traditional Chinese medicine in treating knee osteoarthritis. J Pain Res 2020;13:1421-1429. [doi: 10.2147/JPR.S247827] [Medline: 32606908]
57. Atalay SG, Durmus A, Gezginaslan Ö. The effect of acupuncture and physiotherapy on patients with knee osteoarthritis: a randomized controlled study. Pain Physician 2021 May;24(3):E269-E278. [Medline: 33988943]
58. Luo X, Liu J, Li Q, et al. Acupuncture for treatment of knee osteoarthritis: a clinical practice guideline. J Evid Based Med 2023 Jun;16(2):237-245. [doi: 10.1111/jebm.12526] [Medline: 36999342]
59. Su X, Song P, Meng F, Xu W, Xing H, Zhang H. Tuina manipulation to reduce inflammation and cartilage loss in knee osteoarthritis rats. J Vis Exp 2023 Jun 30;2023(196). [doi: 10.3791/65451] [Medline: 37458470]
60. Fan Y, Li Z, Zhang H, et al. Valgus knee bracing may have no long-term effect on pain improvement and functional activity in patients with knee osteoarthritis: a meta-analysis of randomized trials. J Orthop Surg Res 2020 Sep 1;15(1):373. [doi: 10.1186/s13018-020-01917-x] [Medline: 32873332]
61. Lozano-Meca J, Montilla-Herrador J, Gacto-Sánchez M. Gait speed in knee osteoarthritis: a simple 10-meter walk test predicts the distance covered in the 6-minute walk test. Musculoskelet Sci Pract 2024 Aug;72:102983. [doi: 10.1016/j.msksp.2024.102983] [Medline: 38851178]
62. Liu ZF, Wang HR, Yu TY, Zhang YQ, Jiao Y, Wang XY. Tuina for peripherally-induced neuropathic pain: a review of analgesic mechanism. Front Neurosci 2022;16:1096734. [doi: 10.3389/fnins.2022.1096734] [Medline: 36620462]
63. Huang W, Pach D, Napadow V, et al. Characterizing acupuncture stimuli using brain imaging with FMRI--a systematic review and meta-analysis of the literature. PLoS ONE 2012;7(4):e32960. [doi: 10.1371/journal.pone.0032960] [Medline: 22496739]

XSL•FO
RenderX

64. Boly M, Gosseries O, Massimini M, Rosanova M. Chapter 2 - functional neuroimaging techniques. In: Laureys S, Gosseries O, Tononi G, editors. The Neurology of Conciousness, 2nd edition: Academic Press; 2016:31-47. [doi: 10.1016/B978-0-12-800948-2.00002-9]
65. Caria A, Sitaram R, Birbaumer N. Real-time fMRI: a tool for local brain regulation. Neuroscientist 2012 Oct;18(5):487-501. [doi: 10.1177/1073858411407205] [Medline: 21652587]
66. May A. Chronic pain may change the structure of the brain. Pain 2008 Jul;137(1):7-15. [doi: 10.1016/j.pain.2008.02.034] [Medline: 18410991]
67. Mao CP, Bai ZL, Zhang XN, Zhang QJ, Zhang L. Abnormal subcortical brain morphology in patients with knee osteoarthritis: a cross-sectional study. Front Aging Neurosci 2016;8:3. [doi: 10.3389/fnagi.2016.00003] [Medline: 26834629]
68. Bonney PA, Conner AK, Boettcher LB, et al. A simplified method of accurate postprocessing of diffusion tensor imaging for use in brain tumor resection. Oper Neurosurg 2017 Feb 1;13(1):47-59. [doi: 10.1227/NEU.0000000000001181] [Medline: 28931252]
69. Cai RL, Shen GM, Wang H, Guan YY. Brain functional connectivity network studies of acupuncture: a systematic review on resting-state fMRI. J Integr Med 2018 Jan;16(1):26-33. [doi: 10.1016/j.joim.2017.12.002] [Medline: 29397089]
70. Kang BX, Ma J, Shen J, et al. Altered brain activity in end-stage knee osteoarthritis revealed by resting-state functional magnetic resonance imaging. Brain Behav 2022 Jan;12(1):e2479. [doi: 10.1002/brb3.2479] [Medline: 34967156]
71. Zhang H, Chen H, Wang H, et al. Effect of Chinese Tuina massage therapy on resting state brain functional network of patients with chronic neck pain. J Tradit Chin Med Sci 2015 Jan;2(1):60-68. [doi: 10.1016/j.jtcms.2015.10.001]
72. Li Z, Sun X, Yan J. Study on brain FMRI of the mechanism of Tuina analgesia. Presented at: 7th Asian-Pacific Conference on Medical and Biological Engineering IFMBE Proceedings; Apr 22-25, 2008. [doi: 10.1007/978-3-540-79039-6_174]
73. Davis KD, Moayedi M. Central mechanisms of pain revealed through functional and structural MRI. J Neuroimmune Pharmacol 2013 Jun;8(3):518-534. [doi: 10.1007/s11481-012-9386-8] [Medline: 22825710]
74. Wang L, Song M, Jiang T, Zhang Y, Yu C. Regional homogeneity of the resting-state brain activity correlates with individual intelligence. Neurosci Lett 2011 Jan 25;488(3):275-278. [doi: 10.1016/j.neulet.2010.11.046] [Medline: 21108990]
75. Berridge KC, Kringelbach ML. Pleasure systems in the brain. Neuron 2015 May 6;86(3):646-664. [doi: 10.1016/j.neuron.2015.02.018] [Medline: 25950633]
76. Bellec P, Messé A, Coynel D, Perlbarg V, Benali H, Marrelec G. Resting-state brain networks in functional magnetic resonance imaging (MRI). In: Brain Mapping: From Neural Basis of Cognition to Surgical Applications: Springer; 2011:361-372. [doi: 10.1007/978-3-7091-0723-2_28]
77. Wang Z, Xu H, Wang Z, et al. Effects of externally-applied, non-pharmacological Interventions on short- and long-term symptoms and inflammatory cytokine levels in patients with knee osteoarthritis: a systematic review and network meta-analysis. Front Immunol 2023;14:1309751. [doi: 10.3389/fimmu.2023.1309751] [Medline: 38155966]
78. Liu ZF, Zhang Y, Liu J, et al. Effect of traditional Chinese non-pharmacological therapies on knee osteoarthritis: a narrative review of clinical application and mechanism. Orthop Res Rev 2024;16:21-33. [doi: 10.2147/ORR.S442025] [Medline: 38292459]
79. Colombo B, Rocca MA, Messina R, Guerrieri S, Filippi M. Resting-state fMRI functional connectivity: a new perspective to evaluate pain modulation in migraine? Neurol Sci 2015 May;36 Suppl 1(S1):41-45. [doi: 10.1007/s10072-015-2145-x] [Medline: 26017510]
80. Ma SJ, Zhang JP, Hua XY, Wu JJ, Zheng MX, Xu JG. Tuina therapy promotes behavioral improvement and brain plasticity in rats with peripheral nerve injury and repair. Brain Behav 2023 Sep;13(9):e3174. [doi: 10.1002/brb3.3174] [Medline: 37522806]
81. Liang SH, Zhao WJ, Yin JB, et al. A neural circuit from thalamic paraventricular nucleus to central amygdala for the facilitation of neuropathic pain. J Neurosci 2020 Oct 7;40(41):7837-7854. [doi: 10.1523/JNEUROSCI.2487-19.2020] [Medline: 32958568]
82. Forbes D. Blinding: an essential component in decreasing risk of bias in experimental designs. Evid Based Nurs 2013 Jul;16(3):70-71. [doi: 10.1136/eb-2013-101382] [Medline: 23696228]

Abbreviations

10MWT: 10-meter walking test
AE: adverse event
CRF: case report form
DTI: diffusion tensor imaging
GM: gray matter
KOA: knee osteoarthritis
MRI: magnetic resonance imaging
multi-MRI: multimodal magnetic resonance imaging
SF-MPQ: short-form McGill Pain Questionnaire
TE: echo time

XSL•FO
RenderX

TR: repetition time
VAS: visual analog scale
WM: white matter
WOMAC: the Western Ontario and McMaster Universities Osteoarthritis Index

Edited by J Sarvestan; submitted 25.Sep.2025; peer-reviewed by X Li; revised version received 20.Dec.2025; accepted 22.Dec.2025; published 05.Feb.2026.

Please cite as:
Gu C, Wei D, Zhang K, Ren J, Qiao Y, Zhang F, Zhu Y, Xu Z, Cao Y, Yao S, Guo H, Wang Y, Wang H, Wang L, Wang Q, Guo G
Effects of Acupuncture-Combined Tuina on Patients With Knee Osteoarthritis: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84082
URL: https://www.researchprotocols.org/2026/1/e84082
doi:10.2196/84082

© Chendong Gu, Dong Wei, Kaiyue Zhang, Jinxi Ren, Yihui Qiao, Fengyu Zhang, Yiting Zhu, Zhixian Xu, Yinlong Cao, Shuyi Yao, Hanrui Guo, Yihang Wang, Huan Wang, Lin Wang, Quanliang Wang, Guangxin Guo. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effect of Du-Moxibustion Combined With Spine-Pinching Therapy on Cognitive Frailty in Older Adults With Prefrailty: Protocol for a Randomized Controlled Trial

Suzhen Liu[1*], BMedSci, CMP; Mengxia Qi[2*], BMedSci; Hua-Fang Li[1], BSc; Ziqiu Ye[3], MMed; Yu Xu[1], MMed; Xiangying Yang[1], MD, PhD

[1]The Geriatrics Department, Affiliated Hangzhou First People's Hospital, School of Medicine, Westlake University, No. 261 Huansha Road, Shangcheng District, Hangzhou, Zhejiang, China

[2]School of Nursing, Zhejiang Chinese Medical University, Hangzhou, Zhejiang, China

[3]The Fourth Clinical Medical College, Zhejiang Chinese Medical University, Hangzhou, Zhejiang, China

[*]these authors contributed equally

Corresponding Author:
Xiangying Yang, MD, PhD
The Geriatrics Department, Affiliated Hangzhou First People's Hospital, School of Medicine, Westlake University, No. 261 Huansha Road, Shangcheng District, Hangzhou, Zhejiang, China

Abstract

Background: Cognitive frailty (CF), characterized by the coexistence of mild cognitive impairment and physical frailty in the absence of dementia or other neurodegenerative diseases, is a significant risk factor for dementia and functional decline in older adults. Although Du-moxibustion has shown potential benefits in improving CF, the effect of spine-pinching therapy remains underexplored.

Objective: This study aims to evaluate, for the first time, the efficacy and safety of Du-moxibustion combined with spine-pinching therapy in older adults with prefrailty and CF.

Methods: This is a prospective, single-center, randomized, single-blind, 4-arm parallel controlled trial. A total of 156 older adults with prefrailty and CF will be recruited and randomly assigned to 1 of 4 groups: routine care group, Du-moxibustion group, spine-pinching group, or combined intervention group. The intervention will last for 8 weeks. The primary outcome is the change in Montreal Cognitive Assessment score. Secondary outcomes include the Fried frailty phenotype, Barthel Index (activities of daily living), 15-item Geriatric Depression Scale, 36-item Short Form Health Survey, global frailty status, and traditional Chinese medicine syndrome scores. Outcome assessments will be performed at baseline (wk 0), midintervention (wk 6), postintervention (wk 8), and follow-up (wk 12).

Results: Recruitment for this study is scheduled to commence in March 2026 and will end in June 2027 (recruitment and intervention). All follow-up and data collection activities will be finalized by October 2027. Results are anticipated to be completed in the first quarter of 2028.

Conclusions: This study is expected to provide high-quality evidence for the clinical efficacy of Du-moxibustion combined with spine-pinching therapy in managing CF and contribute to the integration of traditional Chinese medicine external therapies in the promotion of healthy aging. Although the single-blind design may introduce expectancy bias, strict randomization procedures and standardized interventions will enhance the reliability and scientific rigor of the results.

Trial Registration: International Traditional Medicine Clinical Trial Registry ITMCTR2025001589; https://itmctr.ccebtcm.org.cn/mgt/project/view/1957679735857348608

International Registered Report Identifier (IRRID): PRR1-10.2196/82683

(JMIR Res Protoc 2026;15:e82683) doi:10.2196/82683

KEYWORDS

cognitive frailty; Du-moxibustion; older adults; prefrailty; randomized controlled trial; RCT; spine-pinching therapy

Introduction

Cognitive frailty (CF), defined as the coexistence of physical frailty and mild cognitive impairment (MCI) in the absence of dementia, has been increasingly recognized as a critical geriatric syndrome associated with accelerated functional decline, disability, hospitalization, and mortality [1]. Recent epidemiological studies report that the prevalence of CF ranges

XSL•FO
RenderX

from 1% to 12% in community-dwelling older adults and is substantially higher among hospitalized populations [2]. With the rapid demographic shift in China, the number of older adults is projected to exceed 280 million by 2030, suggesting that the burden of CF will continue to rise [3].

Prefrailty, a transitional and potentially reversible stage within the frailty continuum, affects more than half of older adults and represents an important window for targeted preventive intervention [4,5]. However, evidence indicates that prefrailty is frequently underrecognized and undertreated in clinical settings, resulting in missed opportunities to delay or reverse progression toward frailty and cognitive decline [6,7]. Current Western medicine approaches for managing CF—including pharmacological treatments such as cholinesterase inhibitors and nonpharmacological strategies such as cognitive training, physical exercise, and psychosocial support—offer limited benefits on combined cognitive and physical outcomes, and pharmacological approaches often carry risks of adverse events in older adults [8,9]. These limitations highlight the need for safe, multidimensional, and holistic interventions specifically tailored to the population with prefrailty.

Traditional Chinese medicine (TCM) emphasizes individualized, syndrome-based treatment and has shown potential in improving both cognitive performance and frailty-related physical deficits. External TCM therapies such as Du-moxibustion and spine-pinching (niefu) therapy are believed, according to TCM theory, to regulate physiological balance, support spleen-kidney function, and enhance vitality. Emerging evidence suggests that Du-moxibustion may improve cognitive function, physical frailty symptoms, gait performance, and inflammatory biomarkers in older adults with frailty or cognitive impairment [10-12]. Spine-pinching therapy, widely used in TCM rehabilitation to modulate meridian pathways and autonomic function, has demonstrated preliminary benefits in improving gastrointestinal function, physical vitality, and neurocognitive-related symptoms [13,14]. Nevertheless, despite their complementary theoretical basis, no randomized controlled trial (RCT) has evaluated the combined effects of Du-moxibustion and spine-pinching therapy in older adults with prefrailty-related CF.

To address this gap, this study will conduct an RCT to evaluate the efficacy and safety of Du-moxibustion combined with spine-pinching therapy in older adults diagnosed with prefrailty and CF, based on both Western diagnostic criteria and a specific TCM syndrome pattern (“spleen–kidney deficiency with Du meridian obstruction”). This trial aims to provide high-quality evidence supporting an integrative, multidimensional intervention strategy for improving cognition, frailty status, emotional well-being, and quality of life in this growing vulnerable population.

Methods

Ethical Considerations

This study has been approved by the ethics committee of Hangzhou First People’s Hospital (approval 2025ZN211-1), in compliance with the Declaration of Helsinki (Edinburgh 2000 Amendment) and medical ethical requirements. Concurrently, this trial is registered with the International Traditional Medicine Clinical Trial Registry under the registration number ITMCTR2025001589. Participants received comprehensive study information, including the study’s purpose, potential risks (eg, burns or emotional distress), benefits, and their right to withdraw at any time, and provided written informed consent prior to screening. Serious adverse events will be reported to the ethics committee within 24 hours and monitored by the Data and Safety Monitoring Committee (DSMC) throughout the 8-week intervention and 12-week follow-up. Patients or their legal representatives were informed in an accessible language about the study objectives, methods, and procedures. During the preparation of this study, generative artificial intelligence tools were used to assist with translation and linguistic refinement. Results will be submitted to peer-reviewed SCI (Science Citation Index) journals and presented at conferences on geriatrics, TCM, and nursing. Anonymized data will be uploaded to an open-access repository for validation and future research.

Study Design

This study is designed as a prospective, single-center, randomized, single-blind, 4-armed parallel controlled clinical trial. A total of 160 eligible patients will be randomly assigned in a 1:1:1:1 ratio to one of four groups via a centralized randomization system: routine care group, Du-moxibustion group, spine-pinching group, and combined intervention group (Du-moxibustion+spine-pinching). The intervention will last for 8 weeks. The routine care group will receive health education once per week, while the other 3 groups will receive treatment twice per week: Du-moxibustion (30 min/session), spine-pinching (15 - 20 min/session), and combined therapy (45 - 50 min/session). Outcomes will be assessed at baseline, midintervention, postintervention, and follow-up.

Participants

Source of Participants

Participants will be recruited from the geriatrics and neurology outpatient clinics of Hangzhou First People’s Hospital between March 2026 and June 2027. Recruitment will be conducted through outpatient screening and community-based advertisements.

Inclusion Criteria

Participants must meet all of the inclusion criteria summarized in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion criteria

- Aged ≥60 years
- Fried frailty phenotype score between 1 and 2, indicating a prefrail state
- Diagnosed with cognitive frailty, as indicated by a Montreal Cognitive Assessment score of <26
- Diagnosed with the traditional Chinese medicine (TCM) syndrome of "spleen-kidney deficiency with Du meridian obstruction," according to the Standards for TCM Disease and Syndrome Diagnosis and Efficacy Evaluation [15] (see "Diagnostic Criteria for CF" section for diagnostic criteria: ≥2 primary symptoms, ≥2 secondary symptoms, and corresponding tongue and pulse signs)
- Able and willing to provide written informed consent and comply with the study protocol

Diagnostic Criteria for CF

Participants will be screened based on both Western medicine and TCM diagnostic criteria.

Western Medicine Diagnostic Criteria

Based on the definition by the International Academy on Nutrition and Aging or the International Association of Gerontology and Geriatrics [1], the diagnostic criteria for CF include: (1) patient or family member reports subjective cognitive decline; (2) no diagnosis of dementia, with a Clinical Dementia Rating score of 0.5; and (3) Mini–Mental State Examination score ≤27 or MoCA score <26.

These criteria align with the international consensus definition of CF—coexistence of physical frailty and MCI without dementia.

TCM Diagnostic Criteria

According to the Standards for TCM Disease and Syndrome Diagnosis and Efficacy Evaluation [15], patients must meet the following criteria for the TCM syndrome of "spleen-kidney deficiency with Du meridian obstruction":

1. Primary symptoms: fatigue, soreness, and weakness of the waist and knees, poor appetite, and insomnia;
2. Secondary symptoms: dizziness, abdominal distension, lower back pain, abdominal cold pain, tinnitus, edema, pale or flushed complexion, aversion to cold or 5-center heat, spontaneous or night sweating, dry mouth or thirst, frequent urination or nocturia, and loose or dry stools;
3. Tongue and pulse signs: pale or swollen tongue, or red tongue with scant coating; weak, deep-slow, or thin-rapid pulse.

A diagnosis can be made when the patient presents with at least 2 primary symptoms, 2 secondary symptoms, and characteristic tongue and pulse signs.

Exclusion Criteria

Participants will be excluded if they meet any of the criteria summarized in Textbox 2.

Textbox 2. Exclusion criteria

- Diagnosed with Alzheimer disease or other forms of dementia
- Suffering from severe cardiovascular, cerebrovascular, hepatic, renal, or psychiatric disorders
- Received Du-moxibustion, spine-pinching, or other external traditional Chinese medicine therapies in the past 6 months
- Allergic to moxa, ginger, or other topical substances used in the intervention, or have skin ulcers, infections, or other conditions that contraindicate treatment
- Unable to maintain prone or lateral positions due to severe spinal deformities, arthritis, or other conditions
- Participated in other exercise, cognitive training, or intervention studies within the past 3 months

Criteria for Withdrawal or Discontinuation

Participants will be withdrawn or excluded from analysis under the following circumstances: (1) voluntary withdrawal or loss to follow-up before completing the intervention, (2) serious adverse events or complications during the study, (3) poor compliance leading to discontinuation of treatment, and (4) development of severe illness during the study period that requires termination of participation.

Sample Size Calculation

This study uses the MoCA score as the sole primary outcome measure, focusing specifically on the change from baseline to week 8. The primary comparison of interest is the difference in therapeutic efficacy between the combined intervention group and the routine care group.

Based on previous literature concerning nonpharmacological TCM interventions for cardiovascular and cognitive conditions, we hypothesize that the improvement in MoCA scores in the combined intervention group after 8 weeks will exceed that of the routine care group by a clinically meaningful difference ($\triangle\mu$) of approximately 3 points. We acknowledge that the estimated effect size (standardized mean difference ≈0.7-0.9) cited in similar studies may appear optimistic. However, this anticipated 3-point improvement is supported by the moderate-to-strong effect trend consistently reported in meta-analyses of comparable TCM external therapies, justifying its use for the primary calculation.

Assuming an SD (σ) of approximately 4, a one-way ANOVA was used to estimate the required effective sample size. The significance level (α) was set at .05, and statistical power ($1-\beta$)

XSL•FO
RenderX

at 0.80. The initial calculation indicated a required effective sample size of 33 participants per group.

Considering a potential dropout rate of 15%, the required number of participants to be recruited per group is calculated as approximately 33/(1 - 0.15)=38.82. The final sample size was therefore rounded up to 39 participants per group to ensure adequate retention and statistical power, resulting in a total sample size of 156 participants.

Randomization, Allocation, and Blinding

Randomization and Allocation Concealment

Randomization will be conducted independently by a research assistant at the study center. Eligible participants will be randomly assigned in a 1:1:1:1 ratio to the routine care group, Du-moxibustion group, spine-pinching group, or combined intervention group. An independent statistician will generate a stratified random number list using SAS 9.3 software (SAS Institute Inc), stratified by age and MoCA score, with a block size of 6 or 9. Randomization will be implemented through an online central randomization system. Each participant will be assigned a unique identification number to ensure allocation concealment and prevent selection bias.

Blinding

This study uses a single-blind design. Outcome assessors and data analysts will be blinded to group allocation, whereas intervention practitioners cannot be blinded due to the nature of Du-moxibustion and spine-pinching procedures. Participants will be informed they are receiving treatment but will not be told the specific group or intervention type (eg, Du-moxibustion or spine-pinching) to minimize expectation bias. All assessors will undergo standardized training in MoCA and Fried phenotype scoring to ensure objective and consistent evaluations. Statistical analysis will be performed by an independent third party, and raw data will be anonymized. In the case of emergencies (eg, serious adverse events), unblinding may occur, with reasons, timing, and actions recorded in the case report form (CRF). Final unblinding will occur after study completion and data analysis.

Intervention Methods

All participants will receive standardized routine care as a basic treatment. Interventions will be administered by licensed TCM practitioners or trained technicians according to standardized procedures to ensure consistency. During the intervention period, participants will be instructed to avoid receiving any additional external TCM treatments or other therapies that may affect CF.

Routine Care Group

Participants in the control group will receive 30-minute health education sessions twice per week, conducted by registered nurses in classrooms or consultation rooms. Education will be delivered using Microsoft PowerPoint presentations, video demonstrations, and interactive Q&A sessions. Printed health education manuals will also be provided. The curriculum includes:

1. Dietary guidance: low-salt, low-fat diet with balanced nutrition (eg, omega-3-rich fish, fruits, and vegetables);
2. Exercise advice: moderate aerobic exercise (eg, 30 min of brisk walking or Tai Chi daily);
3. Cognitive health strategies: memory training techniques (eg, associative memory and number games) and psychological health management.
4. Frequency and duration: twice per week for 8 consecutive weeks (total of 16 sessions).
5. Safety: any psychological or emotional discomfort (eg, anxiety) reported after sessions will be recorded. No invasive procedures are involved.

Du-Moxibustion Group

Details pertaining to the participants in the Du-moxibustion group are as follows:

- Treatment site: this implies the Du meridian line, from Dazhui (GV14) to Yaoshu (GV2), covering thoracic vertebrae T1 to lumbar vertebrae L5, as illustrated in Figure 1.
- Materials: these include a herbal powder (equal parts of astragalus, codonopsis, chuanxiong, and carthamus), fresh ginger paste (thickness: 1.5 cm), mulberry paper, moxa cones (1.5 cm diameter, approximately 2 g each), alcohol swabs, and an ignition device.
- Procedure: these include (1) the participant lying in the prone position; (2) the back is cleaned and dried; (3) the herbal powder is applied along the Du meridian, covered with mulberry paper; (4) a 1.5 centimeter layer of ginger paste is spread over the area; (5) Moxa cones are placed in shallow depressions formed in the ginger layer; (6) the cones are ignited and burned for approximately 30 minutes, and (7) the residues are cleaned and the skin is inspected for irritation or burns.
- Environment: the treatment room must be well-ventilated and equipped with smoke evacuation systems.
- Frequency and duration: the frequency is twice per week for 8 consecutive weeks (total of 16 sessions). The session duration of Du-moxibustion was set at 30 minutes based on previous clinical trials, in which moxibustion treatments typically lasted 20 to 30 minutes, ensuring both sufficient stimulation and patient safety [16].
- Safety monitoring: skin reactions (eg, redness and burns) will be observed after each session. Cold compresses or suspension of treatment will be applied if necessary. All findings will be recorded in the CRF.

XSL•FO
RenderX

Figure 1. Acupoints and treatment area for Du-moxibustion along the Du meridian (GV14–GV2).

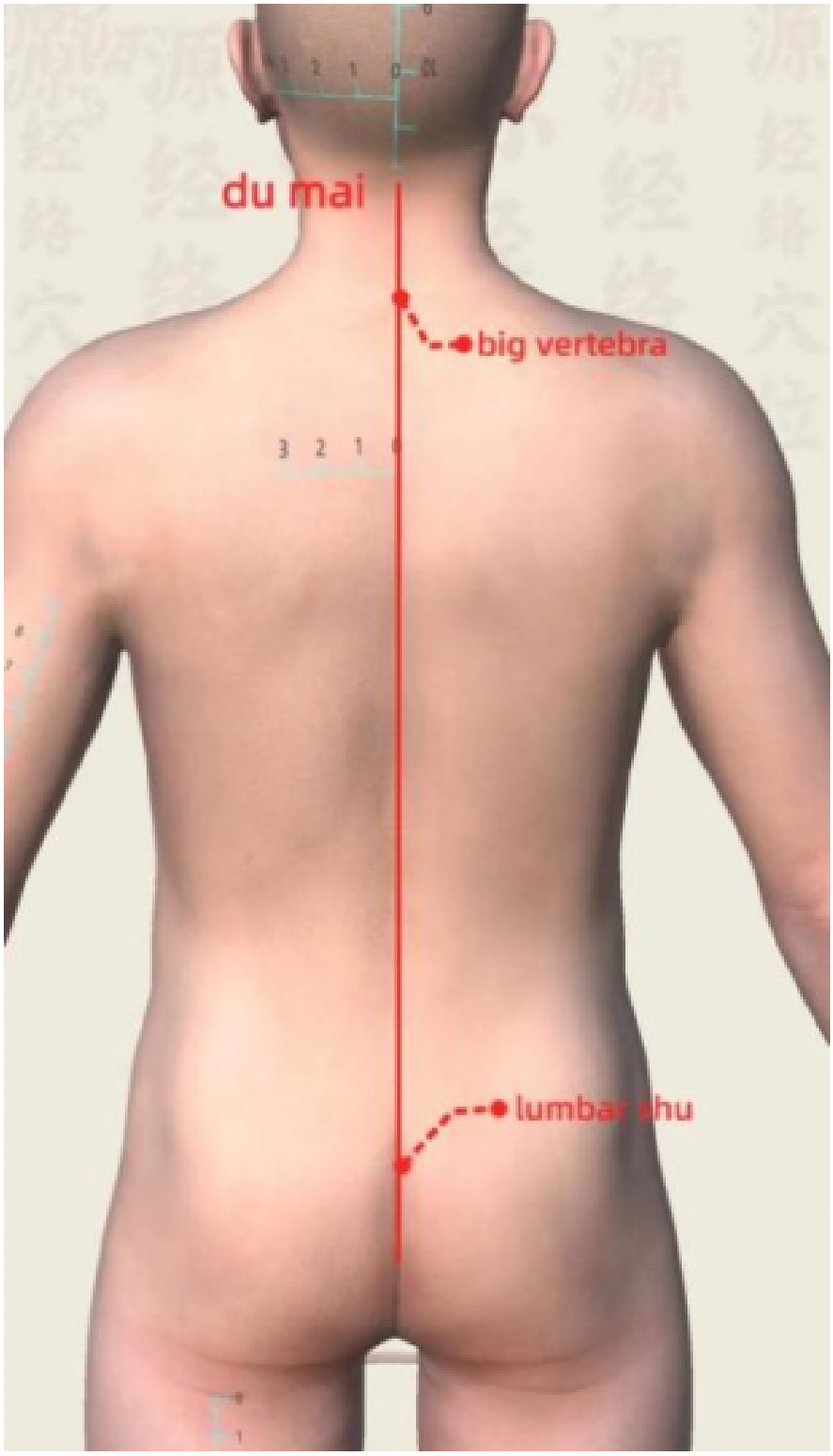

Spine-Pinching Group

Details pertaining to the participants in the spine-pinching group are as follows:

- Position: the position should be prone, with the back exposed.
- Method: starting from the sacrococcygeal region along the midline of the spine, pinch the skin and move upward toward the seventh cervical vertebra. The procedure is performed along 3 lines centered on the spine (left bladder meridian, right bladder meridian, and Du meridian). For each line, perform 3 pinches, followed by lifting (pinch 3 times and lift once) for 4 to 5 cycles. Acupressure is applied to the Xinshu (HT7), Shenshu (KI3), Pishu (SP3), Weishu (ST25), Ganshu (LV3), and Sanjiaoshu (SJ5) points (1 min per acupoint). The acupoint locations are shown in Figure 2, and the manipulation technique is illustrated in Figure 3. The total treatment duration is 15 to 20 minutes.

XSL•FO
RenderX

- Frequency and duration: the frequency is twice per week for 8 weeks, totaling 16 sessions.
- Safety: after each intervention, skin tenderness or muscle discomfort will be monitored and recorded in the CRF.

Figure 2. Acupoint locations used in spine-pinching therapy.

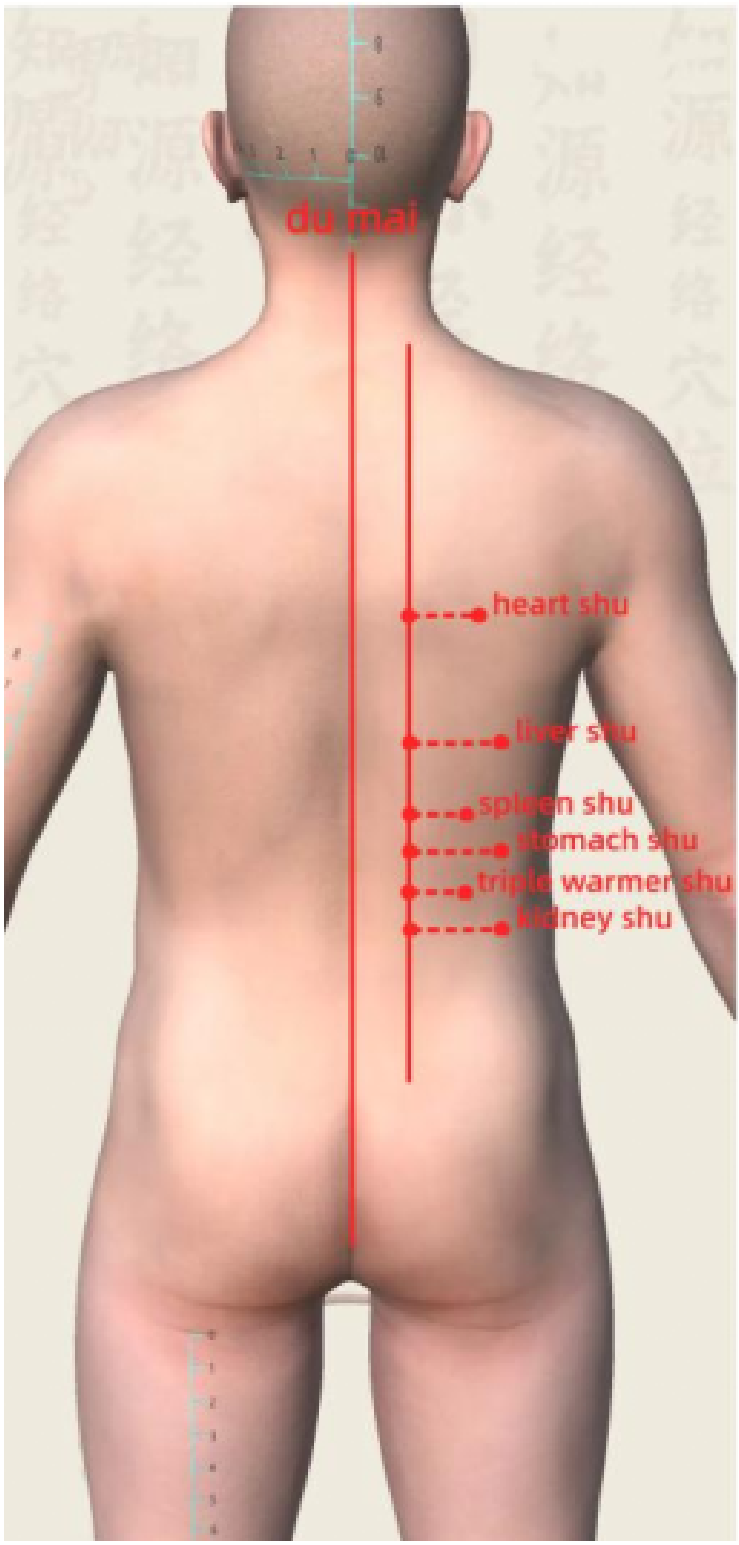

XSL•FO
RenderX

Figure 3. Manipulation technique of spine-pinching therapy.



Combined Intervention Group

Details pertaining to the participants in the combined intervention group are as follows:

- Intervention order: following the "first unblock, then replenish" principle, spine-pinching therapy is performed first (to unblock meridians and regulate qi and blood), followed by Du-moxibustion (to warm and regulate the Du meridian and tonify the kidney and marrow).
- Procedure: the spine-pinching procedure is the same as described in the "Spine-Pinching Group" section, and the Du-moxibustion procedure is the same as described in the "Du-Moxibustion Group" section. The total treatment time is 45 to 50 minutes.
- Environment: the treatment room should have proper ventilation and be equipped with smoke-extraction devices.
- Frequency and duration: the frequency is twice per week for 8 weeks, totaling 16 sessions.
- Safety: after each intervention, skin redness, burns, or tenderness will be monitored. If necessary, cold compresses or suspension of treatment will be applied, and the results will be recorded in the CRF.

Intervention Adherence Management

A dedicated team member will record the execution of interventions and patient feedback. Prior to each session, the research nurse will inform, supervise, and reassure the participants. If a participant misses 2 consecutive sessions or completes less than 80% of the interventions, the reasons will be analyzed, and a decision will be made regarding their inclusion in the final analysis.

Outcome Measures

Overview

During the screening phase, all participants will complete a questionnaire collecting demographic data such as gender, age, nationality, education level, occupation, marital status, as well as disease-related information such as onset time and duration. Prior to randomization, baseline measurements of MoCA and the Fried frailty phenotype will be taken. Clinical evaluations of both primary and secondary outcomes will be conducted at baseline (week 0), midintervention (week 6), postintervention (week 8), and follow-up (week 12). The timeline of the study is depicted in Table 1.

XSL•FO
RenderX

Table . SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines for the schedule of enrollment, interventions, and assessments.

	Baseline	Treatment phase			Baseline
	Week –2～0	Week 0	Week 6	Week 8	Week 12
Ethics and informed consent					
Ethics committee approval	✓				
Signature of informed consent		✓			
Screening and randomization					
Inclusion or exclusion criteria assessment	✓				
Random grouping		✓			
Allocation hiding		✓			
Pretest					
4 cases of pretest	✓				
Intervention					
Conventional control group (2 sessions/week, 30 min each)		✓	✓	✓	
Du-moxibustion group (2 sessions/week, 30 min each)		✓	✓	✓	
Spine-pinching group (2 sessions/week, 15 - 20 min each)		✓	✓	✓	
Combined intervention group		✓	✓	✓	
Outcome assessment (4 sessions/week total)					
MoCA[a]		✓	✓	✓	✓
Fried		✓	✓	✓	✓
Barthel		✓	✓	✓	✓
SF-36[b]		✓	✓	✓	✓
GDS-15[c]		✓	✓	✓	✓
Traditional Chinese medicine syndrome scores		✓	✓	✓	✓
Safety control					
Adverse event monitoring		✓	✓	✓	✓
Serious adverse event report		✓	✓	✓	✓
Quality control					
Research personnel training	✓				

XSL•FO
RenderX

	Baseline	Treatment phase			Baseline
Compliance management		✓	✓	✓	

[a]MoCA: Montreal Cognitive Assessment.

[b]SF-36: 36-item Short Form Health Survey.

[c]GDS-15: 15-item Geriatric Depression Scale.

Primary Outcome Measures: Cognitive Function

Cognitive function will be assessed using the Chinese version of the MoCA. The total score ranges from 0 to 30, with higher scores indicating better cognitive function. The primary outcome is the change in MoCA scores from baseline. Evaluations will take place at weeks 0, 6, 8, and 12.

Secondary Outcome Measures

Physical Frailty

Physical frailty will be assessed using the Fried Frailty Phenotype Scale (0‐5 points). A score of 1 to 2 indicates prefrailty, and ≥3 indicates frailty. The scale includes measures of weight loss, fatigue, grip strength (measured using a standard dynamometer), walking speed (measured by a 4-m walking test), and physical activity reduction. Evaluations will be conducted by assessors at weeks 0, 6, 8, and 12. The main focus will be to observe whether participants transition from "prefrailty" to "nonfrailty."

Activities of Daily Living

Activities of daily living will be assessed using the Barthel Index (0‐100 points, with higher scores indicating better function). The index includes 10 items related to basic daily activities such as feeding, bathing, and dressing. Assessments will be made at weeks 0, 6, 8, and 12.

Quality of Life

Quality of life will be measured using the Short Form Health Survey, self-administered by participants. It includes 8 dimensions, with scores ranging from 0 to 100. The results are summarized into 2 components: physical health component summary and mental health component summary, with higher scores indicating better quality of life. Evaluations will occur at weeks 0, 6, 8, and 12.

Depression Status

Depression will be assessed using the 15-item Geriatric Depression Scale, with a total score ranging from 0 to 15. A score ≥5 suggests the presence of depressive symptoms, and higher scores indicate greater depression severity. The scale will be administered through an interview conducted by an assessor. Evaluations will occur at weeks 0, 6, 8, and 12.

TCM Syndrome Score

The TCM syndrome score will be based on the syndrome differentiation criteria for "spleen-kidney deficiency with Du meridian obstruction." The scale includes items such as dizziness, memory loss, physical weakness, insomnia, and soreness in the waist and knees. A 4-point Likert scale will be used, with higher scores indicating more severe symptoms. Evaluations will be conducted at weeks 0, 6, 8, and 12.

Evaluation Time Points

All outcome measures will be assessed at the following time points: baseline (week 0), midintervention (week 6), postintervention (week 8), and follow-up (week 12).

Data Management

CRFs will be completed on paper by data entry officers, and the data will be independently entered into a password-protected Microsoft Excel database by 2 trained personnel. Any discrepancies will be cross-checked to ensure accuracy. All trial documents—including informed consent forms, CRFs, and intervention records—will be securely stored in encrypted electronic databases and locked cabinets, under the supervision of the study coordinator. Adherence to the intervention (eg, attendance rate and missed sessions) will be documented in the CRFs and collected at weeks 0, 6, 8, and 12. Participant confidentiality will be maintained during result dissemination. At least one audit by the hospital's ethics committee or research oversight body will be conducted to assess data completeness and regulatory compliance.

Statistical Analysis

Overview

Data analysis will be conducted using SPSS (version 27.0; IBM Corp). The primary analysis will be based on the intention-to-treat principle, supplemented by per-protocol analysis. Missing data will be imputed using the last observation carried forward method. Results will be expressed as mean (SD) or median (IQR), as appropriate.

Primary Outcome Analysis

Analysis of the sole primary outcome is focused on the difference in MoCA scores at week 8 between the combined intervention group and the routine care group. We will use ANCOVA for this comparison, adjusting for the baseline MoCA score as a covariate, to derive the primary efficacy conclusion (significance level set at $P<.05$).

Secondary and Exploratory Analysis

All secondary outcomes (eg, Fried Frailty Phenotype, Barthel Index, and 36-item Short Form Health Survey), as well as all pairwise comparisons between the 4 intervention groups, will be considered exploratory analyses. Continuous variables will be tested for normality using the Shapiro-Wilk test. Normally distributed variables will be compared using 1-way ANOVA, followed by the Bonferroni correction for post hoc comparisons. Non-normally distributed variables will be analyzed using the Kruskal-Wallis test, followed by the Dunn test for pairwise comparisons. For changes over multiple time points (weeks 0, 6, 8, and 12), repeated measures ANOVA or the Friedman test (for nonnormal data) will be used for within-group and

XSL•FO
RenderX

between-group comparisons over time. The categorical variables (eg, frailty status) will be analyzed using chi-square tests or the Fisher exact test.

Mechanism Verification and Interaction Test

To formally explore whether the combined therapy offers a synergistic effect (superior to the simple sum of the individual therapies), we will conduct a formal interaction analysis using a 2×2 factorial structure (Du-moxibustion factor×spine-pinching factor) within the ANCOVA framework. Furthermore, preplanned contrast analyses will be performed to test specific hypotheses, including the superiority of the combined intervention group over the individual Du-moxibustion group and the individual spine-pinching group. These results will also be treated as exploratory findings.

Strategy for Controlling Multiple Comparison Risk

To strictly control the risk of type I error inflation due to multiple comparisons, this study will only perform statistical inference and power evaluation on the primary outcome comparison. Formal multiple comparison adjustments will not be applied to the secondary outcomes or multitime-point comparisons. These results will be considered exploratory findings, used only to describe trends and generate future research hypotheses, and will be interpreted with extreme caution, not as definitive conclusions.

Safety Monitoring

Adverse events will be monitored and recorded after each intervention session, including moxibustion burns, skin allergies, moxibustion syncope, and localized tenderness in the intervention groups, as well as psychological or emotional discomfort (eg, anxiety) in the control group. Adverse events will be classified by severity: mild (no treatment required), moderate (intervention paused; eg, cold compress for burns), and severe (intervention terminated and referral required; eg, hospitalization). The research nurse will document the time, symptoms, management, and outcomes in the CRF. Serious adverse events, such as hospitalization or life-threatening conditions, will be reported to the ethics committee within 24 hours. The DSMC will oversee trial safety and may trigger unblinding if necessary. Monitoring will cover the 8-week intervention and 12-week follow-up periods.

Quality Assurance

Prior to participant enrollment, all research staff will undergo standardized training on trial implementation. This includes procedures for moxibustion and spinal pinching techniques, precautions, intervention frequency and duration, questionnaire administration and scoring, and outcome measurement protocols. The trial will adhere to principles of randomization and blinding. Randomization will be performed by noninvestigator personnel using statistical software, with strict allocation concealment to minimize selection bias. Participants and operators will remain blinded to group assignments to prevent contamination and ensure methodological rigor. A pilot trial will be conducted with four prefrail participants to refine intervention protocols. Data collection will avoid leading questions and ensure completeness. Double data entry will be used to prevent omissions and errors.

Quality Control

Before the trial begins, all personnel, including operators, assessors, and data entry staff, will receive training on moxibustion and spinal pinching techniques, safety precautions, and the 8-week treatment schedule (1 session/wk for control group; 2 sessions/wk for intervention groups). Standardized assessments of MoCA and Fried frailty scores, as well as questionnaire distribution and verification, will be covered through both theoretical and practical sessions. Randomization and blinding protocols outlined in the “Randomization, Allocation, and Blinding” section will be followed. A noninvestigator will generate the randomization sequence using SAS 9.3 and perform allocation concealment. Participants will be unaware of their group assignments or intervention types. Although operators will not be blinded, they will not participate in assessments. A pilot study with 4 prefrail older participants will be conducted to refine the procedures (eg, frequency and technical steps). Data collection will avoid suggestive language, and questionnaires will be reviewed on-site. Double-checking by 2 personnel will ensure data quality. The DSMC and ethics committee will oversee trial quality through regular audits.

Results

Recruitment for this study is scheduled to commence in March 2026 for eligible older participants. The 8-week intervention for the final participant is slated for completion in June 2027. All follow-up and data collection activities will be finalized by October 2027. Data analysis and the reporting of primary results are anticipated to be completed in the first quarter of 2028. A flow diagram of the trial process is shown in Figure 4.

XSL•FO
RenderX

Figure 4. Trial flow diagram of the study. EFS: Edmonton Frail Scale; GDS-15: 15-item Geriatric Depression Scale; MoCA: Montreal Cognitive Assessment; SF-36: 36-item Short Form Health Survey; TCM: traditional Chinese medicine.

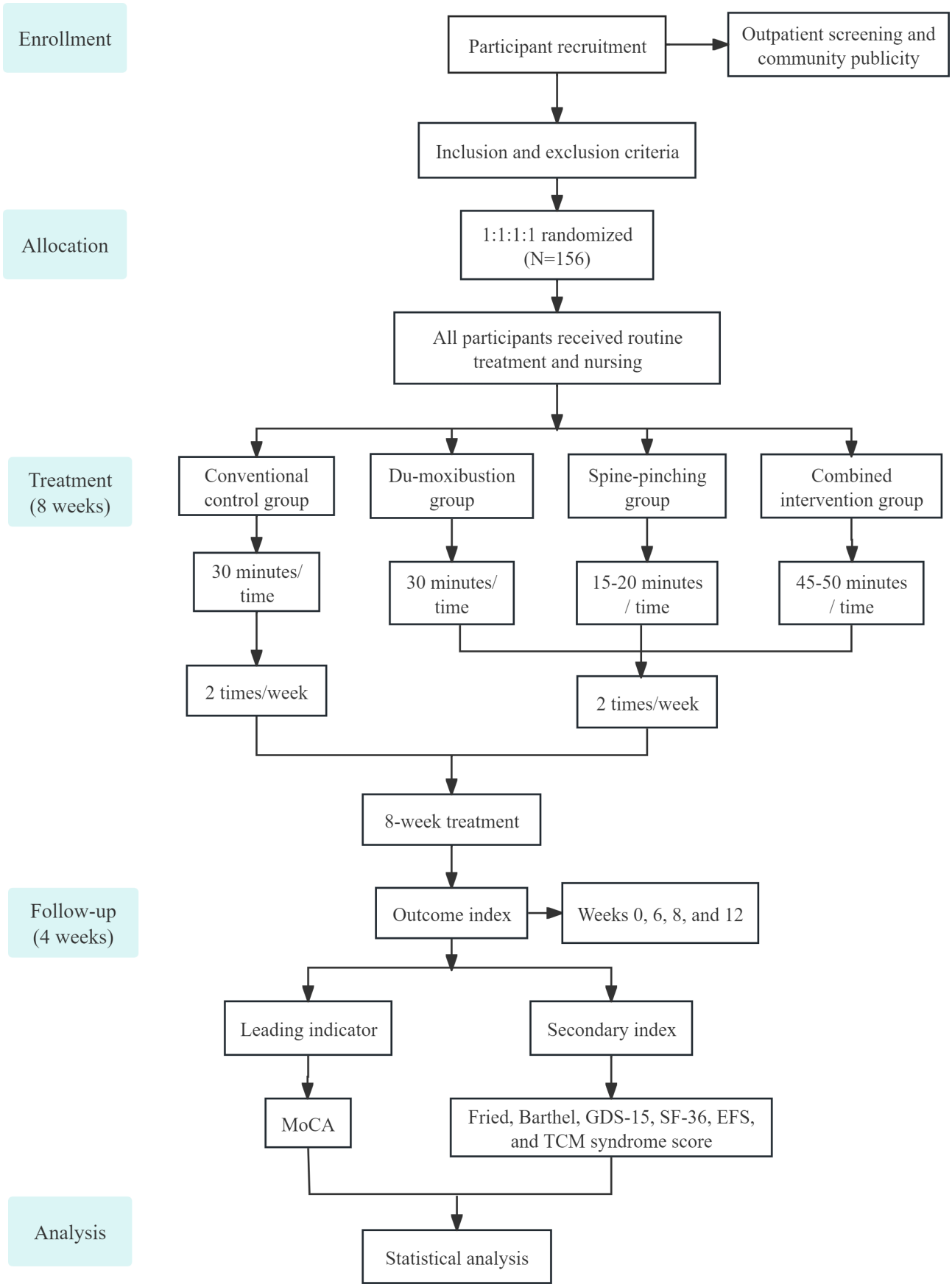

XSL•FO
RenderX

Discussion

This RCT is designed to evaluate the combined effects of Du-moxibustion and spinal pinching therapy on CF in prefrail older adults. By targeting both cognitive and physical decline at an early, potentially reversible stage, this study addresses a critical gap in the current evidence base, where few interventions simultaneously engage multidomain mechanisms relevant to CF [17]. Findings from this work are expected to contribute novel insights into integrative, nonpharmacological strategies for early prevention of cognitive deterioration.

CF—a clinical construct characterized by the coexistence of MCI and physical frailty—has been recognized as a precursor to dementia, disability, and mortality [18]. Despite increasing research interest, available treatments remain limited. Pharmacological therapies have shown only modest benefits for cognitive outcomes and carry risks of adverse effects [19], while exercise-based or nutritional interventions often exhibit suboptimal adherence and do not fully address the complex interplay between neurocognitive, musculoskeletal, and systemic vulnerability [20]. Consequently, there is a growing need for multidimensional, low-risk strategies suitable for early-stage intervention.

TCM offers a conceptual and therapeutic framework aligned with the multidomain nature of CF. According to the TCM theory, cognitive and physical decline in older adults frequently arises from spleen-kidney deficiency and impaired circulation in the Du meridian [21]. External therapies such as moxibustion and spinal pinching are believed to “warm yang,” “nourish the brain,” regulate gastrointestinal function, and restore systemic vitality [22]. Importantly, emerging biomedical evidence provides preliminary support for these mechanisms. Studies indicate that moxibustion may modulate neuroinflammatory pathways, enhance mitochondrial activity, promote neurotrophic factor expression, and regulate autonomic function [23]. At the peripheral level, moxibustion has been shown to improve gastrointestinal motility, reduce systemic inflammation, and enhance microcirculation—processes closely related to frailty progression [24].

Spinal-pinching therapy, though less extensively studied, has demonstrated potential neuromodulatory effects. Mechanistic studies suggest that stimulation along the paraspinal region may influence vagal tone, gastrointestinal-brain axis communication, and immune regulation [25]. These actions are consistent with improvements in appetite, digestion, and physical strength described in clinical practice. Therefore, combining Du-moxibustion and spinal pinching may produce synergistic effects by simultaneously engaging central regulatory pathways (eg, neuroendocrine-immune modulation) and peripheral support systems (eg, gastrointestinal function and muscular vitality) [26]. The principle of “unblocking before tonifying” further provides a rationale for integrating these therapies to enhance functional recovery in individuals with prefrailty.

The methodological design of this study strengthens its scientific rigor. The 4-arm randomized controlled design allows differentiation of the specific and additive effects of Du-moxibustion, spinal pinching therapy, and their combination—a level of comparative evidence rarely achieved in existing TCM research [27]. Standardized operating procedures, assessor blinding, clearly defined intervention parameters, and multidimensional outcome measures contribute to internal validity. The chosen treatment dose (30 min per session, twice weekly for 8 wk) aligns with parameters commonly used in moxibustion trials [28], ensuring sufficient thermal stimulation of the governor vessel while supporting participant adherence and safety. The inclusion of a 12-week posttreatment follow-up enables assessment of both immediate and sustained effects, addressing a common limitation in TCM clinical trials [29].

Nevertheless, several limitations must be acknowledged. First, complete participant blinding is infeasible due to the sensory characteristics of external TCM therapies, raising the possibility of expectation effects. Although assessor blinding mitigates this concern, residual bias cannot be entirely excluded. Second, the single-center nature of the trial may restrict the diversity of participants, limiting generalizability across different healthcare settings and cultural contexts. Third, stringent eligibility criteria requiring prefrailty, CF, and a specific TCM syndrome pattern (“spleen–kidney deficiency with Du meridian obstruction”) enhance internal validity but narrow the applicable population. Many older adults with cognitive or physical impairment may present with alternative syndrome patterns, and future studies should consider multiple TCM subtypes to broaden applicability.

In addition, although this study includes comprehensive functional and symptomatic assessments, biological indicators are absent. The lack of mechanistic biomarkers limits the ability to delineate physiological pathways underlying treatment effects. Incorporating inflammatory markers (eg, interleukin-6, tumor necrosis factor-α, and C-reactive protein), neurotrophic factors (eg, brain-derived neurotrophic factor), mitochondrial function indicators, or neuroimaging modalities (eg, functional magnetic resonance imaging, functional near-infrared spectroscopy, and electroencephalography) in future research may help clarify the neurobiological basis of TCM external therapies [30]. Multicenter studies with larger, more diverse samples and comparative evaluation of different syndrome patterns would further enhance generalizability and allow more robust conclusions regarding clinical applicability [31].

In conclusion, the integrated application of Du-moxibustion and spinal pinching therapy represents a promising, multidomain intervention for early management of CF. By combining central neuroregulatory effects with improvements in peripheral physiological function, this approach aligns with both TCM principles and emerging biomedical evidence [32]. Although limitations exist, the results of this trial may provide foundational evidence supporting scalable, culturally rooted, and mechanistically plausible strategies for promoting healthy aging and delaying cognitive and physical decline.

XSL•FO
RenderX

Acknowledgments

The authors confirm that all artificial intelligence–assisted content was thoroughly reviewed, edited, and verified by them. The authors retain full responsibility for the intellectual integrity, accuracy, and final content of this submission.

Funding

This work was supported by the Hangzhou Municipal Health Science and Technology Plan Project (A20252157), the Hangzhou Municipal Medical and Health Science and Technology Project (ZD20250236), and the Clinical Research Program of Traditional Chinese Medicine Project (2026ZL0607).

Data Availability

No new data were created or analyzed in this study. Data sharing is not applicable to this paper.

Authors' Contributions

Conceptualization: SL
Data curation: ZY
Formal analysis: HFL, YX
Funding acquisition: HFL
Investigation: MQ, YX
Methodology: SL, ZY, HFL, YX
Project administration: XY
Writing – original draft: SL
Writing – review & editing: HFL, XY, SL, MQ, ZY, YX
Authors SL and MQ contributed equally and are considered co–first authors.

Conflicts of Interest

None declared.

References

1. Kelaiditi E, Cesari M, Canevelli M, et al. Cognitive frailty: rational and definition from an (I.A.N.A./I.A.G.G.) international consensus group. J Nutr Health Aging 2013 Sep;17(9):726-734. [doi: 10.1007/s12603-013-0367-2] [Medline: 24154642]
2. Vargas-Torres-Young DA, Salazar-Talla L, Cuba-Ruiz S, Urrunaga-Pastor D, Runzer-Colmenares FM, Parodi JF. Cognitive frailty as a predictor of mortality in older adults: a longitudinal study in Peru. Front Med (Lausanne) 2022;9:910005. [doi: 10.3389/fmed.2022.910005] [Medline: 35814770]
3. Lobanov-Rostovsky S, He Q, Chen Y, et al. Growing old in China in socioeconomic and epidemiological context: systematic review of social care policy for older people. BMC Public Health 2023 Jun 30;23(1):1272. [doi: 10.1186/s12889-023-15583-1] [Medline: 37391766]
4. Science and Technology Management and Evaluation Committee of the Chinese Association of Rehabilitation Medicine, Rehabilitation Branch of the China Association for the Promotion of Anti-Aging, Rehabilitation Assessment Group of the Physical Medicine and Rehabilitation Branch of the Chinese Medical Association, Vascular Cognitive Impairment Branch of the Chinese Stroke Association, Geriatric Rehabilitation Professional Committee of the Beijing Association of Rehabilitation Medicine. Chinese expert consensus on rehabilitation of cognitive frailty in older adults (2023) [Article in Chinese]. Chin Med J 2023;58(9):949-953. [doi: 10.3969/j.issn.1008-1070.2023.09.007]
5. Jing X, Hongyan M, Minhao S, Zhiyu M, Kewen N. Systematic review of risk prediction models for cognitive frailty among Chinese older adults [Article in Chinese]. J Nurs Sci 2024;39(2):23-27. [doi: 10.3870/j.issn.1001-4152.2024.02.023]
6. Mazya AL, Garvin P, Ekdahl AW. Outpatient comprehensive geriatric assessment: effects on frailty and mortality in old people with multimorbidity and high health care utilization. Aging Clin Exp Res 2019 Apr;31(4):519-525. [doi: 10.1007/s40520-018-1004-z] [Medline: 30039453]
7. Juan F, Jianping R, Lixian R. Screening and non-pharmacological interventions for prefrailty in older adults: advances in research [Article in Chinese]. Chin J Health Manage 2022;16(3):212-216. [doi: 10.3760/cma.j.cn115624-20211022-00634]
8. Li Y, Yun J, Xie C, et al. The efficacy of health promotion interventions for pre-frail older adults: a systematic review. J Am Med Dir Assoc 2025 Sep;26(9):105755. [doi: 10.1016/j.jamda.2025.105755] [Medline: 40690947]
9. Sobrinho ADS, de Paula Venancio RC, da Silva Rodrigues G, et al. Systematic review of interventions for pre-frail and frail older adults: evidence from clinical trials on frailty levels. Arch Gerontol Geriatr 2025 Jul;134:105851. [doi: 10.1016/j.archger.2025.105851] [Medline: 40262339]
10. O'Mahony D. STOPP/START criteria for potentially inappropriate medications/potential prescribing omissions in older people: origin and progress. Expert Rev Clin Pharmacol 2020 Jan;13(1):15-22. [doi: 10.1080/17512433.2020.1697676] [Medline: 31790317]

XSL•FO
RenderX

11. Yu L, Tang X, Gong D, Man C, Fan Y. Prefrailty prevalence and all-cause mortality risk in older adults from the general population: a meta-analysis of prospective cohort studies. Geriatr Gerontol Int 2025 Sep;25(9):1167-1177. [doi: 10.1111/ggi.70133] [Medline: 40696888]
12. Eidam A, Durga J, Bauer JM, et al. Interventions to prevent the onset of frailty in adults aged 60 and older (PRAE-Frail): a systematic review and network meta-analysis. Eur Geriatr Med 2024 Oct;15(5):1169-1185. [doi: 10.1007/s41999-024-01013-x] [Medline: 39060779]
13. Pei H, Ma L, Cao Y, et al. Traditional Chinese medicine for Alzheimer's disease and other cognitive impairment: a review. Am J Chin Med 2020;48(3):487-511. [doi: 10.1142/S0192415X20500251] [Medline: 32329645]
14. Sala G, Gobet F. Cognitive training does not enhance general cognition. Trends Cogn Sci 2019 Jan;23(1):9-20. [doi: 10.1016/j.tics.2018.10.004] [Medline: 30471868]
15. Sang BS, Deng WP, Zhao Z. Introduction to Standardization of Traditional Chinese Medicine [Book in Chinese]: China TCM Publishing House; 2013.
16. Lu S, Wang B, Wang J, et al. Moxibustion for the treatment of cancer and its complication: efficacies and mechanisms. Integr Cancer Ther 2023;22:15347354231198089. [doi: 10.1177/15347354231198089] [Medline: 37746720]
17. Sugimoto T, Arai H, Sakurai T. An update on cognitive frailty: Its definition, impact, associated factors and underlying mechanisms, and interventions. Geriatr Gerontol Int 2022 Feb;22(2):99-109. [doi: 10.1111/ggi.14322] [Medline: 34882939]
18. Qiu Y, Li G, Zheng L, et al. Relationship between cognitive frailty and mortality in older adults: a systematic review and meta-analysis. J Am Med Dir Assoc 2023 Nov;24(11):1637-1644. [doi: 10.1016/j.jamda.2023.08.001] [Medline: 37660724]
19. Livingston G, Huntley J, Liu KY, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet standing Commission. Lancet 2024 Aug 10;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
20. Chen LK, Hwang AC, Lee WJ, et al. Efficacy of multidomain interventions to improve physical frailty, depression and cognition: data from cluster-randomized controlled trials. J Cachexia Sarcopenia Muscle 2020 Jun;11(3):650-662. [doi: 10.1002/jcsm.12534] [Medline: 32134208]
21. Qi Z, Wu BL, Chen C, et al. Symptoms based on deficiency syndrome in traditional Chinese medicine might be predictor of frailty in elderly community dwellers. Evid Based Complement Alternat Med 2021;2021:9918811. [doi: 10.1155/2021/9918811] [Medline: 34484408]
22. Ding X, Ma X, Meng P, Yue J, Li L, Xu L. Potential effects of traditional Chinese medicine in anti-aging and aging-related diseases: current evidence and perspectives. Clin Interv Aging 2024;19:681-693. [doi: 10.2147/CIA.S447514] [Medline: 38706635]
23. Aum S, Choe S, Cai M, Jerng UM, Lee JH. Moxibustion for cognitive impairment: a systematic review and meta-analysis of animal studies. Integr Med Res 2021 Jun;10(2):100680. [doi: 10.1016/j.imr.2020.100680] [Medline: 33747784]
24. Bao CH, Wang CY, Li GN, et al. Effect of mild moxibustion on intestinal microbiota and NLRP6 inflammasome signaling in rats with post-inflammatory irritable bowel syndrome. World J Gastroenterol 2019 Aug 28;25(32):4696-4714. [doi: 10.3748/wjg.v25.i32.4696] [Medline: 31528095]
25. Bonaz B, Sinniger V, Pellissier S. Vagal tone: effects on sensitivity, motility, and inflammation. Neurogastroenterol Motil 2016 Apr;28(4):455-462. [doi: 10.1111/nmo.12817] [Medline: 27010234]
26. Osadchiy V, Martin CR, Mayer EA. The gut-brain axis and the microbiome: mechanisms and clinical implications. Clin Gastroenterol Hepatol 2019 Jan;17(2):322-332. [doi: 10.1016/j.cgh.2018.10.002] [Medline: 30292888]
27. Shi F, Wen H, Liu Y, et al. Comparative clinical efficacy of acupuncture combined with manipulation and other non-pharmacological interventions in the treatment of lumbar disc herniation: a prospective, multi-arm, randomized, open-label, blinded endpoint trial. Front Med (Lausanne) 2025;11:1507115. [doi: 10.3389/fmed.2024.1507115] [Medline: 39830381]
28. Chen X, Liu F, Lin N, et al. Warming needle moxibustion for Alzheimer's disease: a systematic review of randomized controlled trials. Geriatr Nurs 2022;43:219-226. [doi: 10.1016/j.gerinurse.2021.12.004] [Medline: 34953330]
29. Bao C, Wu L, Shi Y, et al. Long-term effect of moxibustion on irritable bowel syndrome with diarrhea: a randomized clinical trial. Therap Adv Gastroenterol 2022;15:17562848221075131. [doi: 10.1177/17562848221075131] [Medline: 35222693]
30. Tran Van Hoi E, De Glas NA, Portielje JEA, et al. Biomarkers of the ageing immune system and their association with frailty—a systematic review. Exp Gerontol 2023 Jun 1;176:112163. [doi: 10.1016/j.exger.2023.112163] [Medline: 37028607]
31. Canevelli M, Arisi I, Bacigalupo I, et al. Biomarkers and phenotypic expression in Alzheimer's disease: exploring the contribution of frailty in the Alzheimer's disease neuroimaging initiative. Geroscience 2021 Apr;43(2):1039-1051. [doi: 10.1007/s11357-020-00293-y] [Medline: 33210215]
32. Zheng L, Wang C, Qiu Y, et al. Effectiveness of interventions in older adults with cognitive frailty: a systematic review and meta-analysis of randomised controlled trials. Age Ageing 2022 Dec 5;51(12):afac286. [doi: 10.1093/ageing/afac286] [Medline: 36571775]

Abbreviations

CF: cognitive frailty

XSL•FO
RenderX

CRF: case report form
DSMC: Data and Safety Monitoring Committee
MCI: mild cognitive impairment
MoCA: Montreal Cognitive Assessment
RCT: randomized controlled trial
SCI: Science Citation Index
TCM: traditional Chinese medicine

Edited by J Sarvestan; submitted 20.Aug.2025; peer-reviewed by T Li; revised version received 10.Dec.2025; accepted 12.Dec.2025; published 10.Feb.2026.

Please cite as:
Liu S, Qi M, Li HF, Ye Z, Xu Y, Yang X
Effect of Du-Moxibustion Combined With Spine-Pinching Therapy on Cognitive Frailty in Older Adults With Prefrailty: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e82683
URL: https://www.researchprotocols.org/2026/1/e82683
doi:10.2196/82683

© Suzhen Liu, Mengxia Qi, Hua-Fang Li, Ziqiu Ye, Yu Xu, Xiangying Yang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy of Tuina for Myopia in Children: Protocol for a Randomized Controlled Trial

Shouyao Zhang[1], MM; Yanju Han[2], MM; Lijuan Zhang[3], BM; Fang Tan[4], BM; Xueping Huang[5], MMed; Can Zhang[6], PhD; Zheng Yao[7], PhD; Lei Xiong[1], PhD; Xiantao Tai[6], Prof Dr

[1]The First Clinical Medical College, Yunnan University of Traditional Chinese Medicine, No.1076, Yuhua Road, Chenggong District, Kunming, China

[2]The Second Affiliated Hospital, Yunnan University of Traditional Chinese Medicine, Kunming, China

[3]Department of Tuina, Shiping Hospital of Chinese Medicine, Shiping, China

[4]Department of Tuina, Baoshan Hospital of Chinese Medicine, Baoshan, China

[5]Department of Tuina, Zhaotong Hospital of Chinese Medicine, Zhaotong, China

[6]The Second Clinical Medical College, Yunnan University of Traditional Chinese Medicine, No.1076, Yuhua Road, Chenggong District, Kunming, China

[7]The Basic Medical College, Yunnan University of Traditional Chinese Medicine, Kunming, China

Corresponding Author:
Xiantao Tai, Prof Dr
The Second Clinical Medical College, Yunnan University of Traditional Chinese Medicine, No.1076, Yuhua Road, Chenggong District, Kunming, China

Abstract

Background: Myopia has emerged as a major threat to the visual health of adolescents worldwide. Early intervention can effectively slow down the progression of myopia in adolescents. *Tuina* (also known as Tui Na), a significant therapeutic method in traditional Chinese medicine, has shown promising clinical efficacy in delaying the progression of myopia; however, it lacks robust, large-scale, and standardized randomized controlled trials.

Objective: This study aims to explore the efficacy and safety of *tuina* therapy in managing myopia in adolescents, thereby providing solid evidence for the application of *tuina* in the clinical treatment of myopia.

Methods: This study is a multicenter randomized controlled clinical trial. A total of 62 children with myopia will be recruited from 4 hospitals and randomly assigned in a 1:1 ratio to a *tuina* experimental group and a drug-positive control group (tropicamide eye drops). Treatments in each group will be administered 3 times per week for a total of 8 weeks. The *tuina* experimental group will receive 20 minutes of *tuina* therapy per session, while the drug-positive control group will receive tropicamide eye drops administered every other day, with 2 drops per session. The primary outcome measures include uncorrected visual acuity and axial length, while secondary outcome measures include refractive power and accommodative amplitude. Data will be collected at baseline (week 0), on the day of completion of weeks 4 and 8 of treatment, and at the end of the 10-week follow-up period. Adverse events will be monitored and recorded throughout the study. Statisticians will be blinded. Data will be analyzed using SPSS version 28.0.

Results: This study has been funded, and recruitment began in June 2025. As of December 2025, 29 participants have been enrolled, with 16 allocated to the *tuina* group and 13 to the drug-positive control group. Recruitment is expected to continue until October 2026. Final manuscript submission is anticipated by December 2026.

Conclusions: This study aims to evaluate the efficacy and safety of *tuina* therapy in the treatment of adolescents with myopia. We hypothesize that the therapeutic effect of *tuina* therapy is noninferior to that of tropicamide eye drops, with the additional advantages of fewer side effects and stable long-term efficacy, thereby providing reliable evidence and support for the application of *tuina* therapy in the management of myopia in adolescents.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2600116606; https://www.chictr.org.cn/bin/project/edit?pid=263583

International Registered Report Identifier (IRRID): DERR1-10.2196/79324

(JMIR Res Protoc 2026;15:e79324) doi:10.2196/79324

KEYWORDS

myopia; pediatric tuina therapy; adolescents; randomized controlled trials; protocol; tuina; Tui Na

XSL•FO
RenderX

Introduction

Background

Myopia, a common ocular disorder predominantly developing during childhood and adolescence, is characterized by excessive elongation of the eye globe, leading to the focusing of distant object images in front of the retina and thereby causing blurred distance vision. With continuous advancement of socioeconomic levels [1], the widespread prevalence of mobile electronic devices [2], and a lack of outdoor activities [3], the problem of myopia among adolescents has become increasingly severe. Currently, the global population with myopia is approximately 2 billion individuals (28.3% of the global population), and it is estimated that by 2050, the prevalence of myopia will increase to 4.76 billion individuals, approximately half of the global population [4]. Among school-aged children (6 - 19 y old), the prevalence of myopia in East Asian regions is the highest (73%), which is significantly higher than that in other regions of the world, including North America (42%), Europe (40%), South America (10%), and Africa (3.4% - 4%) [5]. Taking China as an example, more than 80% of high school students suffer from myopia [6], which not only has a significant impact on children's academic performance and future career choices [7] but also serves as a major cause of incurable vision impairments, such as glaucoma, retinal detachment, and macular degeneration [8]. Owing to the extremely high prevalence rate, myopia has caused tremendous losses to socioeconomic development. A research report stated that in 2019, productivity loss caused by severe visual impairment and blindness was approximately US $94.5 billion, and it is expected to rise to US $229.3 billion by 2050 [9]. Therefore, researchers and clinicians are eager for effective treatment methods that can delay or even prevent the progression of myopia in children.

Currently, common interventions for myopia include surgical and nonsurgical methods. Refractive surgery is used to correct the vision; however, high costs [10] and postoperative complications, such as dry eye syndrome, are drawbacks that are difficult to ignore in surgical treatments [11]. Nonsurgical interventions include pharmacological and nonpharmacological treatments, with pharmacological therapy commonly using low-concentration atropine; however, photophobia [12] and dry eye syndrome are common complications. Tropicamide eye drops are a cholinergic receptor antagonist medication. Compared with atropine, tropicamide has the advantages of rapid metabolism, fewer adverse reactions, higher safety, and higher patient compliance and adherence. Evidence suggests that, compared with atropine treatment, tropicamide eye drops can improve the clinical treatment outcomes and enhance the quality of life in adolescents with myopia [13].

Optical therapies, such as myopia-control glasses, including highly aspheric lenslets, spectacle lenses with slightly aspheric lenslets, and single-vision spectacle lenses, are widely applied methods for myopia control and can effectively slow the progression of myopia and axial elongation [14]. However, a large-scale network meta-analysis indicated that the GRADE (Grading of Recommendations Assessment, Development, and Evaluation) evidence level for myopia-control spectacles is relatively low, and their true efficacy requires further investigation [15]. Corneal reshaping lenses can correct myopia to some extent but are associated with adverse reactions, such as corneal infections [16], and many children find these lenses difficult to tolerate. Some studies have shown that more than half of children were unable to complete the treatment [17]. Moreover, high costs necessitate long-term consideration when choosing contact lenses. A study on the cost of myopia treatment over 5 years found that corneal reshaping lenses were the most expensive, with an approximate cost of US $15,347 [10]. Recent research has found that repetitive low-intensity red light therapy can help improve axial length (AL), control vision in children with high myopia, and increase macular choroidal thickness [18]; however, there is still a certain rebound effect after the cessation of repetitive low-intensity red light treatment [19].

Tuina (also known as Tui Na), one of the nonpharmaceutical therapies characteristic of traditional Chinese medicine (TCM), can be traced back to the Qin and Han dynasties (99 BC-26 BC) [20,21] and has a history of more than 2000 years. Massage, acupressure, acupoint eye exercises, and other techniques all fall under the category of *tuina*. Existing research has demonstrated that localized *tuina* therapy around the eyes can promote meridian circulation, enhance blood circulation, and alleviate spasms of the ciliary and extraocular muscles [22-24]. These effects may be associated with improvements in retinal thickness, choroidal thickness, perfusion area, and vascular density through acupoint massage around the eyes [25]. Tian et al [26] evaluated the impact of eye exercises on AL in adolescents and found that the frequency of eye exercises was conducive to controlling the trend of axial elongation. Moreover, a large-scale meta-analysis suggested that acupoint stimulation around the eyes has a moderate protective effect in controlling myopia; however, incorrect execution and attitudes may have a negative impact on vision health [27]. This indicates that the correct method of stimulating ocular acupoints is of paramount importance for the recovery of vision.

The holistic concept of TCM emphasizes the dialectical relationship between local functions and the whole body, with the various parts of the human body being inseparable in structure and mutually influential in function. Normal vision is a result of the proper functioning of all organs. A sole focus on local manipulation neglects the holistic perspective of TCM, potentially affecting the final clinical efficacy. Therefore, we have developed an integrative approach combining overall regulation with local treatment, known as selective spinal *tuina* manipulation therapy, as a potential intervention for myopia in adolescents. However, high-quality medical evidence supporting this approach is currently lacking.

Objectives

This study aims to conduct a high-quality randomized controlled trial to objectively evaluate the efficacy and safety of selective spinal *tuina* manipulation therapy in the treatment of childhood myopia. We hypothesize that selective spinal manipulation therapy can improve vision in children with myopia and is safer than drug therapy, with no side effects and more stable long-term efficacy.

XSL•FO
RenderX

Methods

Study Design

This study is designed as a multicenter randomized controlled noninferiority trial to be conducted from June 2025 to October 2026. We plan to recruit 62 participants from 4 hospitals in Yunnan Province. Eligible participants will be randomly assigned in a 1:1 ratio to the *tuina* experimental group and the positive drug control group (tropicamide eye drops). Each participant will undergo an 8-week treatment period followed by a 2-week follow-up. Baseline information and outcomes will be assessed at baseline, at the completion of weeks 4 and 8 of treatment, and at the end of the 10-week follow-up (Figure 1). This trial adheres to the guidelines of the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials).

XSL•FO
RenderX

Figure 1. Flowchart of this study.

XSL•FO
RenderX

Participants

Recruitment Strategy

This study will recruit 62 participants through recruitment notices in communities and hospitals from the Second Affiliated Hospital of Yunnan University of Traditional Chinese Medicine, Baoshan Hospital of Chinese Medicine, Shiping Hospital of Chinese Medicine, and Zhaotong Hospital of Chinese Medicine. Recruitment will be conducted by physicians from the *tuina* departments of hospitals. Appropriate compensation will be provided to participants upon completion of the entire procedure. Moreover, the parents or legal guardians of the participants have the right to withdraw from this study at any time and for any reason without facing discrimination or difficulties in the hospital. Participants must be children with myopia and must meet the following inclusion criteria.

Diagnostic Criteria

Myopia will be diagnosed according to the standards outlined in the White Paper on Myopia Control and Prevention Research established by the International Myopia Institute under the World Health Organization [28,29], as well as the "Diagnostic and Therapeutic Criteria for Internal Medicine Diseases in Traditional Chinese Medicine" (ZY/T001.1 - 94), a standard of the People's Republic of China for Traditional Chinese Medicine [30].

1. Mild myopia: −0.50D to −3.00D
2. Moderate myopia: −3.00D to −6.00D
3. High myopia: above −6.00D

Inclusion Criteria

The inclusion criteria are as follows:

1. Children who undergo routine ophthalmic examination before enrollment and meet the diagnostic criteria for simple mild-to-moderate myopia in both TCM and Western medicine
2. Age range: 6 to 12 years, gender-neutral
3. Guardians and children who sign the informed consent forms
4. Guardians and children who volunteer to participate in clinical observation, comply with the physician's arrangements, and cooperate with treatment

Exclusion Criteria

The exclusion criteria are as follows:

1. Presence of ocular inflammation, facial dermatitis, or infectious lesions
2. Patients with severe systemic diseases, such as cardiovascular, cerebrovascular, hepatic, renal, hematological, and psychiatric disorders
3. Individuals with a history of congenital hereditary high myopia
4. Patients with pathological myopia fundus changes and/or significant visual impairment, or other concomitant ophthalmic diseases that may affect the assessment of efficacy
5. Children suffering from neurological and psychiatric disorders (eg, cerebral palsy, autism, spina bifida, anorexia nervosa)

Sample Size

Using G*Power software to calculate the sample size, based on the preliminary pilot study results, the expected efficacy rate in the manipulation treatment group is above 80% ($P<.05$). Therefore, we set an expected effect size of 0.8. To detect between-group differences with 80% statistical power ($\beta=.20$), we choose a 2-tailed α level of .05. With an allocation ratio of 1:1, the required sample size was 52 participants (26 per group). Considering a 20% expected dropout rate, the final sample size was set at 62 participants (31 per group).

Randomization and Blinding

The head of each hospital and the relevant department will recruit participants. Eligible children with myopia will be randomly assigned on a 1:1 basis to either a *tuina* therapy intervention group or a drug positive control group. Randomization will be conducted by the head of each hospital using a computer-generated random number generator. Upon the entry of qualified trial participants, the hospital head will assign them to the respective groups according to the randomly generated numbers and follow the corresponding medical instructions without any modification. There are 2 types of interventions in this study. It is difficult to implement blinding for both the therapists and the parents or legal guardians of the children. Nevertheless, the outcome assessors will use blinding for the randomization, intervention procedures, and result analyses.

Intervention

Selection of Therapists and Participants

To ensure that all *tuina* manipulation in this study conforms to standardized procedures, the therapists involved possess physician qualifications and a master's degree in pediatric *tuina* and have undergone unified training in the *tuina* techniques used in this study. Only those who successfully complete the training are eligible to participate. During the course of the experiment, children with myopia are generally not allowed to use any other interventions. If participants adopt other interventions during this period, they will be defined as noncompliant and withdrawn from the experiment.

Tuina Experimental Group

Participants will receive 20-minute pediatric *tuina* treatments. The treatments will be administered 3 times a week for a total of 8 weeks. A professor specializing in pediatric *tuina* will determine the specific techniques. The techniques will require a light and gentle touch, steady and substantial. Children will be placed in a quiet and warm treatment room with a temperature maintained between 25 °C and 28 °C, ensuring they are in a state of calm. The specific steps are as follows: (1) Step 1: holistic regulation; (2) Step 2: selective stimulation of back-sensitized acupoints; (3) Step 3: symptomatic treatment; and (4) Step 4: harmonization of Yin and Yang.

XSL•FO
RenderX

First, based on the holistic concept and the principle of adhering to the disease mechanism, we will apply *tuina* to the Governor Vessel (GV) and the Bladder Meridian of Foot-Taiyang (BL):

- The child will be placed in a prone position, with the practitioner standing on the right side of the child. They will use the thumb to knead the entire spine in the direction of the GV from top to bottom, 3 times for 1 minute.
- The practitioner will knead the first and second lateral lines of the bladder meridian from top to bottom using the index and middle fingers, 3 times for 1 minute.
- The practitioner will use the thumb and index finger to pinch and lift the skin of the spinal column surface, performing this action from bottom to top 3 times in 1 minute.
- The practitioner will rub the acupoints of BL 23, GV 4, and BL 31 - 34, using the palm or the thenar eminence to firmly press against the skin and perform rapid linear reciprocating rubbing movements for 1 minute (Figure 2).

XSL•FO
RenderX

Figure 2. Holistic regulation. Step 1 of selected spinal manipulative therapy. BL: Bladder Meridian of Foot-Taiyang; GV: Governor Vessel.

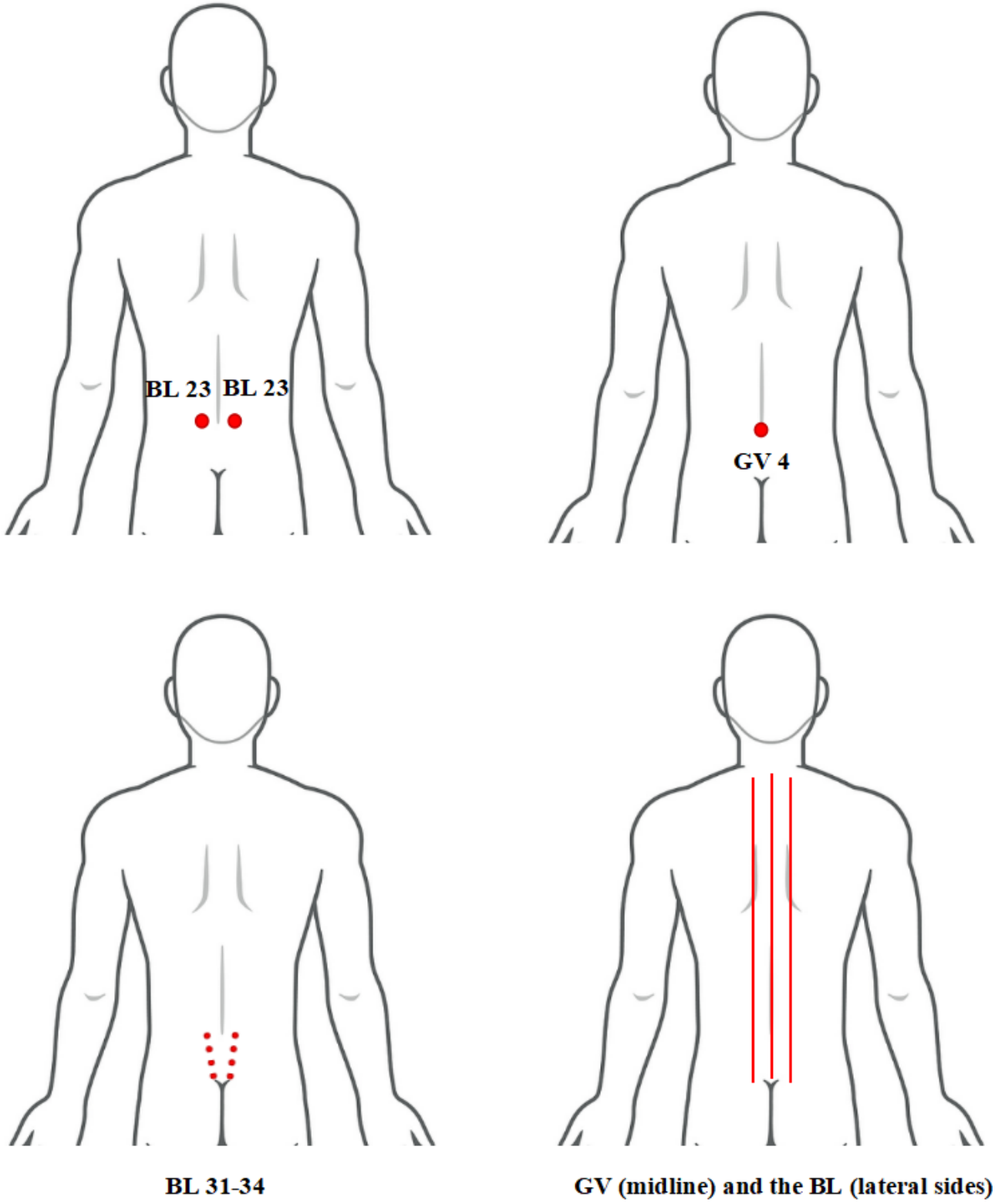

Second, the child will remain in a prone position, and the practitioner will stand on the right side. The practitioner will perform palpation of the cervical and dorsal regions to identify positive reaction points (linear or nodular reaction points). We will confirm the sensitized acupoints using infrared thermography. The practitioner will perform targeted stimulation using kneading manipulation for 2 minutes (Multimedia Appendix 1).

Third, for symptoms of dry eyes and blurred vision, the practitioner will apply the kneading technique to BL 1, BL 2, Eight Extraordinary Meridians-Head and Neck (EX-HN) 4, Triple Energizer Meridian of Hand-Shaoyang (TE) 23, Stomach Meridian of Foot-Yangming (ST) 1, and ST 2, with each acupoint treated for 1 minute, totaling 6 minutes. They will gently rub both palms together to generate heat and then apply the warmth to the eyes for 1 minute (Figure 3).

XSL•FO
RenderX

Figure 3. Symptomatic treatment. Step 3 of selected spinal manipulative therapy. BL: Bladder Meridian of Foot-Taiyang; EX-HN: Eight Extraordinary Meridians-Head and Neck; ST: Stomach Meridian of Foot-Yangming; TE: Triple Energizer Meridian of Hand-Shaoyang.

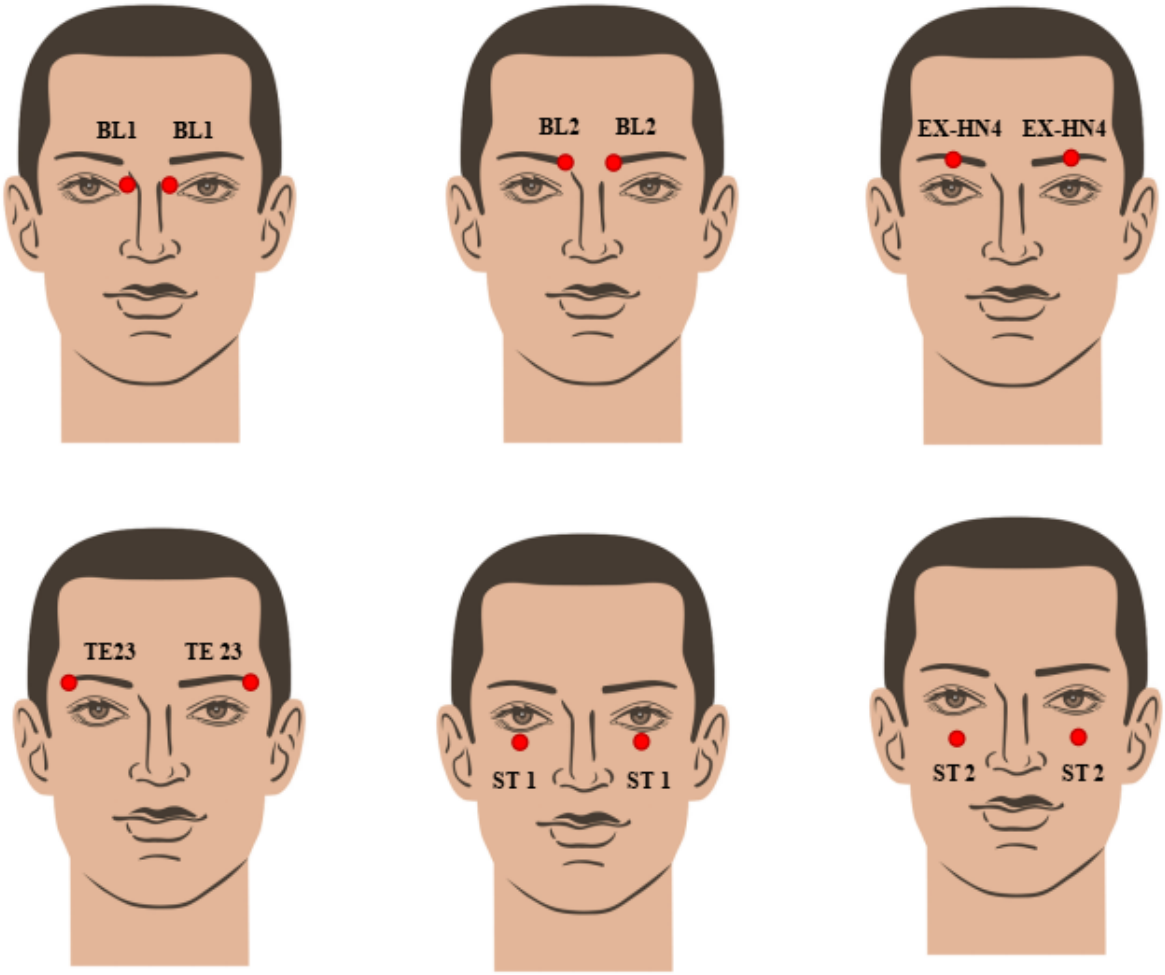

Fourth, the child will assume a supine position, and the practitioner will stand on the side of the child's head. The practitioner will alternately push the "Tianmen" and the "Kangong" with both thumbs 50 times for each and knead EX-HN5 for 1 minute with the middle fingers, totaling 2 minutes. The practitioner then will use the thumbs to knead Large Intestine Meridian of Hand-Yangming (LI) 20 and the palm to rub GV 22 and GV 20 circularly for 1 minute each, totaling 3 minutes. We will then place the patient in a sitting position, and the practitioner will lightly pat the back with a hollow palm, performing this 3 times from top to bottom, for 1 minute. The thumb and index finger, along with the middle finger, will pinch the BL 17 3 times for 1 minute (Figure 4).

XSL•FO
RenderX

Figure 4. Harmonization of Yin and Yang. Step 4 of selected spinal manipulative therapy. EX-HN: Eight Extraordinary Meridians-Head and Neck; GV: Governor Vessel; LI: Large Intestine Meridian of Hand-Yangming.

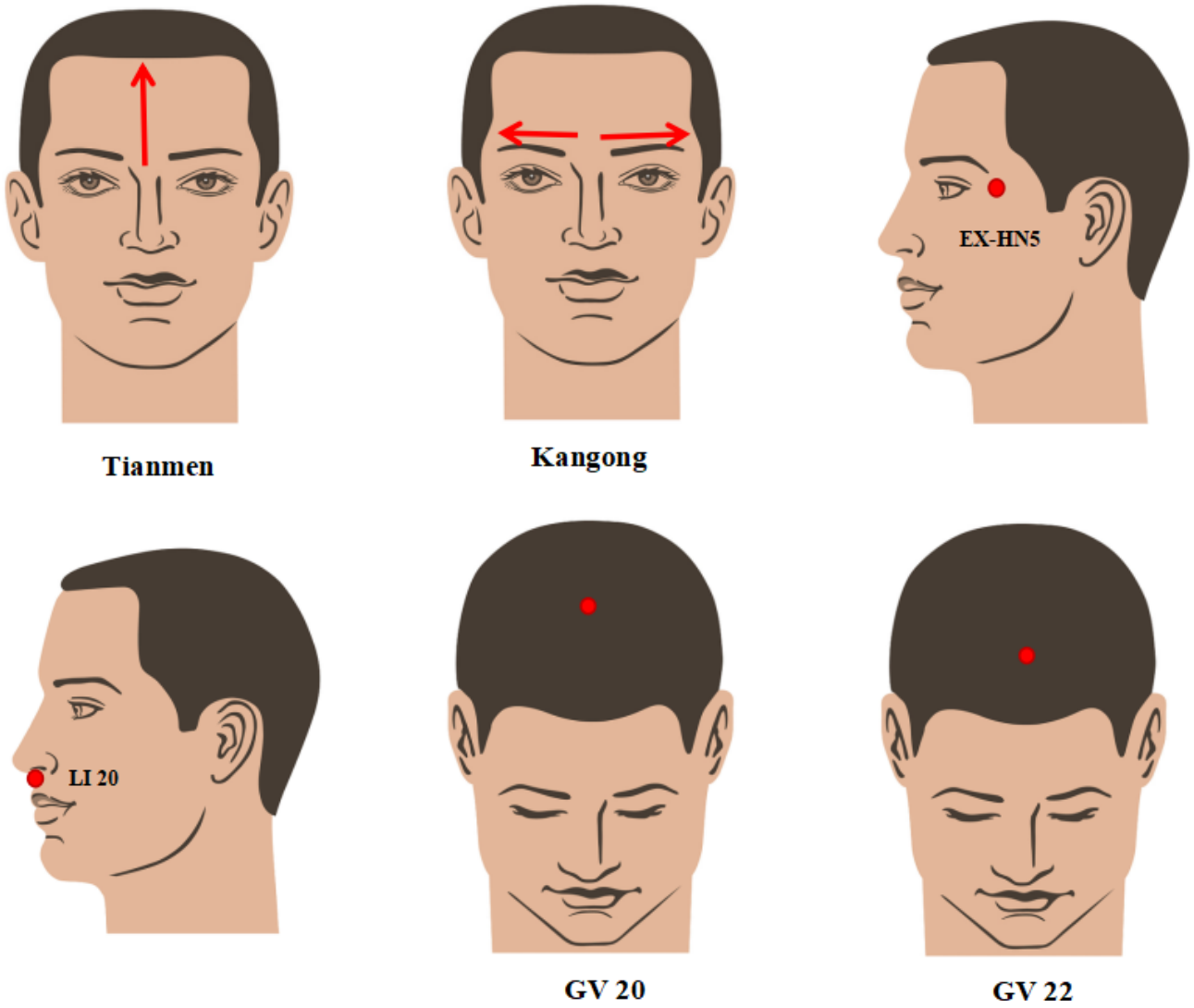

For the location and manipulation details of each acupoint, see Multimedia Appendix 1.

Drug-Positive Control Group

The positive-drug control group will receive treatment with compound tropicamide eye drops (Shenyang Xingqi Co., Ltd.; product specification: 1 ml per vial) at a frequency of once every 2 days before bedtime, with 2 drops per administration. The group will continue the treatment for a duration of 8 weeks.

Outcomes Measures

Primary Outcome

Uncorrected Visual Acuity

We use an international standard visual acuity chart to assess visual acuity, record uncorrected visual acuity (UCVA) results in decimal notation, and subsequently convert them into logMAR format for statistical purposes. The complete disappearance of symptoms such as visual fatigue and exotropia, the restoration of distant vision, and the achievement of a visual acuity of 5.0 on the visual acuity chart are indicative of a cure; the basic disappearance of symptoms with an improvement of 3 to 5 lines on the visual acuity chart is indicative of marked efficacy; a significant alleviation of symptoms with an improvement of 2 lines on the visual acuity chart is indicative of efficacy; no change in symptoms with an improvement of 1 line on the visual acuity chart is indicative of inefficacy. The brightness of the visual acuity chart box must meet the latest national standards, which stipulate that the brightness of the visual acuity chart box should not be less than 200 cd/m^2. We check the brightness of the examination room, as it affects the size of the pupil and use moderate appropriate lighting, with a diameter of the pupil around 3 mm being optimal.

Axial Length

We measure AL using ophthalmic A-ultrasound, obtain 5 readings with differences between them all less than 0.05 mm, and take the average. Normal AL for adolescents ranges from 22 to 24 mm, and values outside this range represent refractive errors.

Secondary Measures

Spherical Equivalent

One drop of 1% cyclopentolate is instilled at 0, 5, and 20 minutes to induce cycloplegia in both eyes. Fifteen minutes

XSL•FO
RenderX

after instillation, cycloplegia affects the pupillary light reflex and the size of the pupil. Automatic cycloplegic refraction is performed using an automatic refractor (KR-8800, Topcon). Three readings are obtained for each eye, and the average is calculated until the desired precision (spherical and cylindrical refractive power of 0.25 D and axis of 5°) is achieved. We calculate the spherical equivalent (SE) refraction as the sum of the spherical value and half the cylindrical value.

Amplitude of Accommodation

We determine the ability of the lens to change refractive power in response to near stimuli using the negative lens method. Refractive status is corrected using a comprehensive refractor (RK-F2, Canon, Japan), with the near visual target fixed at 40 cm. We instruct the participants to fixate on the line above the best vision on the near visual acuity chart and gradually add negative lenses in increments of −0.25 D until the participant experiences persistent blur, at which point we restore the previous degree. The amplitude of accommodation (AMP) is calculated as the total degree of added negative lenses (taking the positive value) plus 2.50 D (Figure 5).

Figure 5. Study schedule of measurements at different time points. AL: axial length; AMP: amplitude of accommodation; SE: spherical equivalent; UCVA: uncorrected visual acuity.

Time point	Study period (d)							
	Baseline	Treatment					Follow-up	
	0	1	14	28	42	56	63	70
Enrollment								
Sign informed consent	✓							
Inclusion/exclusion criteria	✓							
Allocation	✓							
Intervention								
Tuina		⟷	⟷	⟷	⟷	⟷		
Tropicamide		⟷	⟷	⟷	⟷	⟷		
Measurement								
UCVA	✓			✓		✓		✓
AL	✓			✓		✓		✓
SE	✓			✓		✓		✓
AMP	✓			✓		✓		✓
Safety								
Adverse events		✓	✓	✓	✓	✓	✓	✓

Data Analysis

All data will be analyzed using SPSS version 28.0 and GraphPad Prism version 10.3.1 statistical software. For normally distributed data, results will be described as mean (SD) and analyzed using the *t* test. For nonnormally distributed data, results will be described as median (IQR) and analyzed using the Wilcoxon rank sum test for within-group comparisons and the Mann-Whitney *U* test for between-group comparisons. A 2-sided *P* value of .05 will be considered statistically significant. Missing data will be handled using expectation maximization, last observation carried forward, or listwise deletion.

Adverse Events Reporting and Compliance Enhancement

From week 0 to week 10, detailed records of adverse events (AEs) will be documented, including type, occurrence time, severity, and duration. In the event of visual impairment, skin injuries, fainting, infections, or other serious AEs, treatment will be discontinued, and the participant will be transported to the nearest hospital. A detailed AE report will be submitted to the Ethics Committee. At the conclusion of the follow-up period, children are entitled to receive 3 sessions of pediatric *tuina* therapy at no cost; however, those who withdraw prematurely are not eligible for this benefit.

Data Management and Quality Control

In this study, the basic information of each participant includes gender, age, height, weight, medical history, current condition, and contact information for parents or legal guardians. Researchers promptly complete the case report form (CRF) for each participant, including UCVA, SE, AL, AMP, concomitant medications, dropout status, and AEs (such as blurred vision, photophobia, etc). Regardless of whether these AEs are related to the treatment of this study, participants will receive appropriate management and timely documentation, which we

XSL•FO
RenderX

record in detail in the CRF. Four rounds of outcome data are collected from baseline to the end of follow-up. Completed CRF forms from 4 hospitals are sent to the Department of Tuina Basic Education of the Second Clinical Medical College of Yunnan University of Chinese Medicine. Two members who are unaware of the grouping status will conduct double data entry and aggregate the data into an electronic dataset. Simultaneously, the data administrator monitors and checks for any procedural errors. Unblinding is permissible after the completion of the experiment. The Data Monitoring Committee is independent of the sponsor and has no conflicts of interest.

Ethical Considerations

The trial has been registered on the Chinese Clinical Trial Registry (ChiCTR2600116606). At the onset of this study, we submitted an application to the Chinese Clinical Trial Registry Center; however, the review process took an extended period of time, resulting in registration occurring only after the recruitment and data collection had commenced. This study was approved by the Ethics Committee of the Second Affiliated Hospital of Yunnan University of Chinese Medicine (2024‐050) and conforms to the Declaration of Helsinki. It will be ensured that participants actively provide informed consent or discuss and make decisions with trusted family members. All participants' personal identification information, including names, telephone numbers, identification numbers, and medical records, is anonymized to prevent information leakage. Researchers store all participants' data in a dedicated cabinet and retain them for at least 5 years after publication. The Ethics Committee of Yunnan University of Chinese Medicine will regularly review the progress of the trial and supervise the collection, distribution, and confidentiality of the data. The committee can implement modifications or termination of the trial. The use of all personal photographs has been approved by the individuals concerned and their guardians, and informed consent forms have been signed (Multimedia Appendix 2). Upon completion of the treatment and follow-up, the participants will be provided with 3 iterations of *tuina* therapy at no cost. Participants who withdraw before completion will not be eligible for this benefit. All treatments and tests received by participants are free of charge. Participants will be allowed to withdraw at any time, and their subsequent treatment will not be affected in any way.

Results

As of December 2025, 29 participants have been enrolled, with 16 allocated to the *tuina* experimental group and 13 allocated to the drug-positive control group. The current data indicate that *tuina* therapy is helpful in enhancing UCVA and improving AL of eyes, with no AEs reported, suggesting that *tuina* therapy is an effective alternative treatment. The enrollment process will continue until October 2026. We will submit the final manuscript before December 2026, following which we will disseminate the final results in accordance with the CONSORT checklist.

Discussion

Expected Findings

This study aims to evaluate the efficacy and safety of *tuina* therapy through a randomized controlled clinical trial. We anticipate that children who receive *tuina* therapy will show improvements in AL, UCVA, SE, and AMP that are not inferior to those achieved through drug therapy. Furthermore, we expect that *tuina* therapy will be safer, free of adverse effects, and exhibit more stable long-term efficacy.

The high prevalence and severe complications of myopia have inflicted significant damage on the visual health of adolescents, imposing a substantial burden on both society and the economy. Therefore, the urgent need for an effective method to prevent and control myopia is evident. *Tuina* is a potential alternative therapy for myopia. However, current research on the application of *tuina* therapy for myopia predominantly focuses on local ocular areas, neglecting the holistic perspective, which to some extent limits its therapeutic efficacy. The holistic perspective is a fundamental concept in TCM, which views the human body as an integrated whole. This perspective posits that the human body is a harmonious entity, with the heart as the master and the 5 organs at the center, interconnected through the meridian system. This holistic entity is indivisible in terms of structural integration, functionally coordinated, and pathologically interrelated. Consequently, in the rehabilitation of myopia, it is imperative not only to address local symptoms around the eyes but also to promote the overall regulation of the child's body.

During our long-term clinical practice, we have observed that children with myopia exhibit corresponding abnormalities at specific acupoints on their backs, such as BL 13, BL 15, BL 18, BL 20, BL 23, and their surroundings. These abnormalities include the presence of cord-like or nodular reactive substances, or localized temperature decreases. Selectively stimulating these abnormal reaction points on the back acupoints (or regions) can effectively improve the visual acuity of children. Further research has revealed that the acupoints of the first lateral line of the BL are highly congruent with the dermatomal projections of the spinal nerve roots on the surface of the body. Based on these findings, we have integrated the holistic perspective of TCM with the anatomical principles of spinal nerves in Western medicine, thereby establishing "selective spinal *tuina* manipulation therapy." This technology system consists of 4 steps: the first and fourth steps focus on overall regulation, while the second and third steps emphasize localized treatment. This approach achieves a combined model of diagnosis and treatment that integrates both global and local care, leading to the restoration of vision.

The first step is comprehensive treatment to achieve overall rehabilitation objectives. The principle underlying this approach is the stimulation of the GV and the Bladder Meridian, which is closely associated with the regulation of visceral functions. According to 《*Yin Hai Zhi Nan*》, "In treating the eyes, one cannot afford to overlook the Bladder Meridian." The Bladder Meridian is the meridian with the highest number of acupoints in the human body. Each internal organ corresponds to a specific

XSL•FO
RenderX

acupoint on this meridian. By stimulating the acupoints on this meridian, one can regulate the body's neural and humoral functions, enhance the body's *qi*, and achieve the purpose of preventing disease [31]. The second step involves selective stimulation of back sensitized acupoints or regions to achieve personalized and precise rehabilitation objectives. The sensitization of acupoints refers to the transition from a "quiescent" (physiological state) to an "activated" (pathological state) condition, primarily manifesting as abnormalities in morphological, thermal, and electrophysiological indicators. By stimulating abnormal back acupoints, the regulation of visceral functions can be achieved. The third step involves addressing common symptoms associated with myopia, such as blurred vision and visual fatigue, through local massage stimulation around the eye area. This can help to relax muscles, promote blood circulation, and enhance vision. A meta-analysis on the selection rules of acupoints for myopia indicates that local eye acupoints such as BL 1, BL 2, and ST 2, which have branches of the supraorbital nerve, oculomotor nerve, and facial nerve, as well as the orbital vessels and nerves, may promote increased blood flow to the eyes, enhance orbital blood supply, and stimulate orbital nerves to regulate, thereby coordinating the contraction and relaxation of eye muscles and exerting regulatory effects on myopia [32,33]. Fourth, harmonizing *yin* and *yang* involves stimulating acupoints along the GV and related specific points on the head and face, further promoting overall rehabilitation. Stimulating acupoints on the head and face can enhance cognitive function and calm the mind, while rubbing the back along the GV and Bladder Meridian can regulate the flow of *qi* throughout the body, adjusting the functions of the organs. By integrating these methods, the circulation of blood and *qi* can be promoted, and *yin* and *yang* can be harmonized, ultimately achieving the rehabilitation of myopia.

Innovations and Limitations

It should be noted that this study also has certain limitations: First, given that *tuina* therapy relies on the proficient mastery of techniques and acupoints by the practitioner, it is difficult to implement a blinded design. As a form of physical contact between the practitioner and the patient, *tuina* therapy, whether or not it targets specific acupoints, inevitably stimulates the skin, muscles, and peripheral nerves. Therefore, it is also challenging to blind the patients. Consequently, a sham *tuina* group was not included. However, relatively standardized operational protocols and procedures have been established, and all participating practitioners must strictly adhere to them in an effort to minimize such differences. Nevertheless, completely eliminating such differences is challenging. If subsequent analyses reveal that these differences significantly impact the results, they will be considered incorporating them into our analytical models to ensure the accuracy and reliability of our findings. Second, due to financial and human resource constraints, this study does not conduct a multicenter trial involving the entire country. Future work will aim to expand the scope of the study.

Conclusions

This study will investigate the efficacy and safety of *tuina* therapy in the prevention and treatment of myopia in adolescents, providing evidence for clinical decision-making and supporting the inheritance and development of *tuina* therapy. It is noteworthy that previous studies have not investigated the intervention effects of *tuina* therapy on myopia from the perspective of TCM holistic philosophy. This study will be the first to explore the effects of *tuina* therapy on myopia in adolescents from the perspective of TCM holistic philosophy. This study could support *tuina* therapy as an effective method for managing myopia in adolescents, offering potential benefits to patients and their caregivers.

Acknowledgments

We extend our heartfelt gratitude to the Second Affiliated Hospital of Yunnan University of Chinese Medicine, Baoshan Hospital of Chinese Medicine, Shiping Hospital of Chinese Medicine, and Zhaotong Hospital of Chinese Medicine for their tremendous support in this study.

Funding

The Yunnan Province Innovation Team for Acupuncture and Tuina Prevention and Treatment of Encephalopathy supports this work (grant 202405AS350007). The research has received funding for a period of 2 years, commencing from November 2024. The funding amount is 200,000 RMB (US $1=6.9518 RMB, according to the exchange rate between RMB and USD on February 1, 2026). The funder had no involvement in the study 18 design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study will be available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: SZ, XT
Data curation: YH, LZ, FT, XH
Funding acquisition: XT.
Supervision: XT, LX

XSL•FO
RenderX

Writing – original draft: SZ
Writing – review & editing: CZ, ZY

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details of *tuina* manipulation.
[DOCX File, 11481 KB - resprot_v15i1e79324_app1.docx]

Multimedia Appendix 2
Informed consent.
[DOCX File, 377 KB - resprot_v15i1e79324_app2.docx]

References

1. Galdón A, Vila-Vidal N, El Gharbi M, et al. The impact of socioeconomic status on visual acuity changes in schoolchildren: a one-year follow-up. Children (Basel) 2024 Oct 9;11(10):1226. [doi: 10.3390/children11101226] [Medline: 39457191]
2. Enthoven CA, Polling JR, Verzijden T, et al. Smartphone use associated with refractive error in teenagers: the myopia app study. Ophthalmology 2021 Dec;128(12):1681-1688. [doi: 10.1016/j.ophtha.2021.06.016] [Medline: 34245754]
3. Kido A, Miyake M, Watanabe N. Interventions to increase time spent outdoors for preventing incidence and progression of myopia in children. Cochrane Database Syst Rev 2024 Jun 12;6(6):CD013549. [doi: 10.1002/14651858.CD013549.pub2] [Medline: 38864362]
4. Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology 2016 May;123(5):1036-1042. [doi: 10.1016/j.ophtha.2016.01.006] [Medline: 26875007]
5. Grzybowski A, Kanclerz P, Tsubota K, Lanca C, Saw SM. A review on the epidemiology of myopia in school children worldwide. BMC Ophthalmol 2020 Jan 14;20(1):27. [doi: 10.1186/s12886-019-1220-0] [Medline: 31937276]
6. Liu Y, Qin R, Xiong Y, et al. Reference values of non-cycloplegic spherical equivalent for screening and predicting myopia among children and adolescents - China, 2020-2024. China CDC Wkly 2025 Feb 28;7(9):298-303. [doi: 10.46234/ccdcw2025.048] [Medline: 40104755]
7. Zhao C, Cai C, Ding Q, Dai H. Efficacy and safety of atropine to control myopia progression: a systematic review and meta-analysis. BMC Ophthalmol 2020 Dec 7;20(1):478. [doi: 10.1186/s12886-020-01746-w] [Medline: 33287746]
8. Haarman AEG, Enthoven CA, Tideman JWL, Tedja MS, Verhoeven VJM, Klaver CCW. The complications of myopia: a review and meta-analysis. Invest Ophthalmol Vis Sci 2020 Apr 9;61(4):49. [doi: 10.1167/iovs.61.4.49] [Medline: 32347918]
9. Holy C, Kulkarni K, Brennan NA. Predicting costs and disability from the myopia epidemic – a worldwide economic and social model. Invest Ophthalmol Vis Sci 2019;60(9):5466 [FREE Full text]
10. Agyekum S, Chan PP, Adjei PE, et al. Cost-effectiveness analysis of myopia progression interventions in children. JAMA Netw Open 2023 Nov 1;6(11):e2340986. [doi: 10.1001/jamanetworkopen.2023.40986] [Medline: 37917061]
11. Giri P, Azar DT. Risk profiles of ectasia after keratorefractive surgery. Curr Opin Ophthalmol 2017 Jul;28(4):337-342. [doi: 10.1097/ICU.0000000000000383] [Medline: 28594648]
12. Yam JC, Zhang XJ, Zhang Y, et al. Effect of low-concentration atropine eyedrops vs placebo on myopia incidence in children: the LAMP2 randomized clinical trial. JAMA 2023 Feb 14;329(6):472-481. [doi: 10.1001/jama.2022.24162] [Medline: 36786791]
13. Chunyan Z, Xin J, Li Z, Yanyan F, Huimin K. The impact of compound tropicamide versus atropine on myopia progression in adolescents. Xiandaiyixue Yu Jiankangyanjiu 2023;7(20):17-19 [FREE Full text] [doi: 10.3969/j.issn.2096-3718.2023.20.006]
14. Bao J, Yang A, Huang Y, et al. One-year myopia control efficacy of spectacle lenses with aspherical lenslets. Br J Ophthalmol 2022 Aug;106(8):1171-1176. [doi: 10.1136/bjophthalmol-2020-318367] [Medline: 33811039]
15. Wolffsohn JS, Whayeb Y, Logan NS, Weng R, International Myopia Institute Ambassador Group. IMI-Global trends in myopia management attitudes and strategies in clinical practice-2022 update. Invest Ophthalmol Vis Sci 2023 May 1;64(6):6. [doi: 10.1167/iovs.64.6.6] [Medline: 37126357]
16. Kam KW, Yung W, Li GKH, Chen LJ, Young AL. Infectious keratitis and orthokeratology lens use: a systematic review. Infection 2017 Dec;45(6):727-735. [doi: 10.1007/s15010-017-1023-2] [Medline: 28534320]
17. Lawrenson JG, Shah R, Huntjens B, et al. Interventions for myopia control in children: a living systematic review and network meta-analysis. Cochrane Database Syst Rev 2023 Feb 16;2(2):CD014758. [doi: 10.1002/14651858.CD014758.pub2] [Medline: 36809645]
18. Xiong R, Zhu Z, Jiang Y, et al. Longitudinal changes and predictive value of choroidal thickness for myopia control after repeated low-level red-light therapy. Ophthalmology 2023 Mar;130(3):286-296. [doi: 10.1016/j.ophtha.2022.10.002] [Medline: 36240954]

XSL•FO
RenderX

19. Xiong R, Zhu Z, Jiang Y, et al. Sustained and rebound effect of repeated low-level red-light therapy on myopia control: a 2-year post-trial follow-up study. Clin Exp Ophthalmol 2022 Dec;50(9):1013-1024. [doi: 10.1111/ceo.14149] [Medline: 36054314]
20. Gang G. Huangdi Neijing: Yuelu Press; 2021.
21. Lei X, Xiantao T. Textbook of Pediatric Tuina for Traditional Chinese Medicine Pediatrics: A Planned Textbook for Higher Medical Education in Traditional Chinese Medicine: China Press of Traditional Chinese Medicine; 2022.
22. Yifei S, Dan L, Min Z, et al. Clinical effect of mingmu guiyuan tuina in treatment of simple myopia in school-age children. Liaoning Journal of Traditional Chinese Medicine 2025;52(4):91-96. [doi: 10.13192/j.issn.1000-1719.2025.04.024]
23. Mei W, Su W, Liang J, Yu J. Teaching the “Four Ming Acupressure Points” massage method for adolescent myopia prevention and treatment. Asian J Surg 2023 Dec;46(12):5985-5986. [doi: 10.1016/j.asjsur.2023.09.019] [Medline: 37723039]
24. Pang Y, Wang K, Zhang M, Yu J. Treatment of 17 cases of anisometropic amblyopia in children with tuina at “Simingxue” combined with pressing therapy at auricular points. World J Acupunct Moxibustion 2021 Apr;31(2):151-153. [doi: 10.1016/j.wjam.2021.03.003]
25. Liu Y, Li W, Li L, et al. Effects of Chinese acupoint eye exercise on ocular biological parameters. Trans Vis Sci Tech 2025 Jul 1;14(7):20. [doi: 10.1167/tvst.14.7.20]
26. Tian FF, Liu LJ, Guo Y, et al. Effects of eye exercises on axial eye elongation in junior students. Zhonghua Liu Xing Bing Xue Za Zhi 2021 Sep 10;42(9):1621-1627. [doi: 10.3760/cma.j.cn112338-20201118-01338] [Medline: 34814593]
27. Tang J, Pei Y, Wang J, et al. The association between Chinese eye exercises and myopia in children and adolescents: a systematic review and meta-analysis. Front Public Health 2023;11:950700. [doi: 10.3389/fpubh.2023.950700] [Medline: 36969666]
28. Flitcroft DI, He M, Jonas JB, et al. IMI - Defining and classifying myopia: a proposed set of standards for clinical and epidemiologic studies. Invest Ophthalmol Vis Sci 2019 Feb 28;60(3):M20-M30. [doi: 10.1167/iovs.18-25957] [Medline: 30817826]
29. Ruihua W, Daqian L, Nan J, Bei D. Interpretation of the International Myopia Institute white papers focusing on myopia prevention and control. Recent Advances in Ophthalmology 2019;39(8):701-713. [doi: 10.13389/j.cnki.rao.2019.0162]
30. Diagnostic criteria, syndrome classification, and efficacy evaluation for diseases that are close to the patient and distant from the patient—Standards for Diagnosis and Efficacy Evaluation of Diseases in Internal Medicine of Traditional Chinese Medicine (ZY/T001.1-94) of the People's Republic of China. J Liaoning Univ Tradit Chin Med 2019;21(1) [FREE Full text]
31. Qiongjiao H, Hongbo L. A brief discussion on the relationship between bladder meridian of foot sun and the prevention of diseases. DAZHONG KEJI 2021;23(8):81-83. [doi: 10.3969/j.issn.1008-1151.2021.08.021]
32. Shuqin W, Dongyan G, Chunyang Z. Discussion on the location of anatomical landmarks of cuanzhu(bl2). Chinese Acupuncture & Moxibustion 2022;42(1):96-98. [doi: 10.13703/j.0255-2930.20201120-k0003]
33. Xiaoyan T, Kai R, Song L, et al. Observations on the efficacy of acupuncture at periocular extraordinary points in treating juvenile myopia. Shanghai J Acupunct Moxibustion 2010;29(10):643-645. [doi: 10.3969/j.issn.1005-0957.2010.10.643]

Abbreviations

AE: adverse event
AL: axial length
AMP: amplitude of accommodation
BL: Bladder Meridian of Foot-Taiyang
CRF: case report form
EX-HN: Eight Extraordinary Meridians-Head and Neck
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
GV: Governor Vessel
LI: Large Intestine Meridian of Hand-Yangmin
SE: spherical equivalent
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
ST: Stomach Meridian of Foot-Yangming
TCM: traditional Chinese medicine
TE: Triple Energizer Meridian of Hand-Shaoyang
UCVA: uncorrected visual acuity

XSL•FO
RenderX

Edited by A Schwartz; submitted 19.Jun.2025; peer-reviewed by C Yang, X Zhang, Y Zhang; revised version received 13.Jan.2026; accepted 15.Jan.2026; published 12.Feb.2026.

Please cite as:
Zhang S, Han Y, Zhang L, Tan F, Huang X, Zhang C, Yao Z, Xiong L, Tai X
Efficacy of Tuina for Myopia in Children: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e79324
URL: https://www.researchprotocols.org/2026/1/e79324
doi:10.2196/79324

©Shouyao Zhang, Yanju Han, Lijuan Zhang, Fang Tan, Xueping Huang, Can Zhang, Zheng Yao, Lei Xiong, Xiantao Tai. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical Efficacy and Safety Assessment of Specific-Mode Electroacupuncture Stimulation Combined With Paclitaxel for Recurrent Malignant Gliomas: Study Protocol for a Single-Arm Trial

Zhaoxing Jia[1*], MD; Tianxiang Jiang[1*], MMed; Yiqing Zhang[1*], MMed; Qianyue Chen[1*], MMed; Zhong Di[2], MD; Qi Yuan[3], MMed; Kecheng Qian[1], MD; Lin Gan[2], MD; Congcong Ma[2], MD; Xianming Lin[2], MD

[1]The Third Clinical College of Zhejiang University of Traditional Chinese Medicine, Hangzhou, China

[2]Department of Acupuncture, The Third Affiliated Hospital of Zhejiang Chinese Medical University, No. 219 Moganshan Road, Xihu District, Hangzhou, Zhejiang, China

[3]Medicine Hangzhou TCM Hospital Affiliated to Zhejiang Chinese Medical University, Hangzhou, China

[*]these authors contributed equally

Corresponding Author:
Xianming Lin, MD
Department of Acupuncture, The Third Affiliated Hospital of Zhejiang Chinese Medical University, No. 219 Moganshan Road, Xihu District, Hangzhou, Zhejiang, China

Abstract

Background: Despite advances in surgical resection, radiotherapy, and chemotherapy, the prognosis of recurrent malignant gliomas (rMG) remains poor, with limited efficacy of conventional treatments due to the blood-brain barrier (BBB) hindering drug delivery to the tumor site. Studies have demonstrated that albumin-bound paclitaxel (ABX), while potent in vitro, is restricted in its intravenous use due to BBB limitations. To overcome this, specific-mode electrical stimulation (SMES) has shown promise in transiently opening the BBB, enhancing the accumulation of ABX in glioma tumors. Therefore, this protocol designs a single-center, single-arm, prospective phase II clinical trial aiming to evaluate the safety and clinical efficacy of SMES combined with ABX (SMES+ABX) for treating rMG.

Objective: This study primarily evaluates the safety of SMES+ABX therapy in treating patients with rMG and assesses whether it can improve the 4-month progression-free survival (4m-PFS) rate, while providing data support for future large-scale clinical trials.

Methods: In this study, 20 eligible patients will receive intravenous ABX (135 - 175 mg/m²) per 21-day cycle for 6 cycles, combined with SMES for BBB modulation. A Simon 2-stage design will be employed, with the primary end point being the 4m-PFS. Secondary end points include adverse events, disease control rate, objective response rate, duration of disease control, duration of response, Neurological Assessment in Neuro-Oncology score, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30, progression-free survival, and overall survival.

Results: The results will determine the 4m-PFS rate, overall safety profile, secondary efficacy outcomes, and patient-reported quality of life measures. The data will be analyzed upon trial completion. Patient enrollment is scheduled to begin in May 2025. The treatment and primary efficacy assessment phases are anticipated to be completed by January 2027 (allowing for staggered enrollment and a 4-month treatment period for the last enrolled patient). The final survival follow-up for all patients is anticipated to be completed by January 2028 (ie, 1 year after the last patient completes treatment). Data management is currently ongoing, and formal statistical analyses have not yet been performed.

Conclusions: This study aims to evaluate the efficacy and safety of SMES combined with ABX in the treatment of rMG. If successful, the combination could offer a promising therapeutic strategy for this challenging patient population.

Trial Registration: ClinicalTrials.gov NCT06818331; https://clinicaltrials.gov/study/NCT06818331

International Registered Report Identifier (IRRID): DERR1-10.2196/84593

(JMIR Res Protoc 2026;15:e84593) doi:10.2196/84593

KEYWORDS

ABX; albumin-bound paclitaxel; BBB; blood-brain barrier; glioma; recurrent; single-arm trial; SMES; specific-mode electrical stimulation

XSL•FO
RenderX

Introduction

Glioblastoma multiforme is the most aggressive primary malignant brain tumor in adults. Even with multimodal treatment, the median overall survival (OS) is only 12‐15 months [1,2]. Nearly all patients experience recurrence, and effective treatment options for recurrence are extremely limited [3,4].

Albumin-bound paclitaxel (ABX), a microtubule-stabilizing agent, exerts antitumor effects by disrupting mitotic spindle dynamics and inducing apoptosis, making it a cornerstone therapy for diverse solid malignancies. In vitro studies have demonstrated the potent antitumor activity of ABX in glioma cell lines (eg, U87 and U251), with mean IC50 values 1400-fold lower than temozolomide [5-8]. However, clinical trials have revealed the limited efficacy of intravenous ABX monotherapy in recurrent glioblastoma [9,10], primarily attributable to insufficient drug accumulation in tumor tissue due to blood-brain barrier (BBB) exclusion [11,12], thereby precluding therapeutic efficacy.

To overcome these limitations, various ABX delivery strategies have been developed, including drug structural modifications and carrier-mediated delivery systems to enhance BBB penetration [13-16]. Recent breakthroughs in focused ultrasound-induced BBB opening have demonstrated 3‐to 5-fold increases in cerebral ABX concentration in animal models, with significant survival extension. The first human trial (NCT04528680, 2021) preliminarily confirmed the safety of this approach, showing local disease control in 42% of the patients [17,18]. However, despite improved drug delivery to the brain, these strategies face critical challenges, including inconsistent delivery efficiency, safety concerns, and high treatment costs, which limit their clinical translation. Therefore, there is an urgent need for safer, more operable, and cost-effective BBB-penetrating ABX delivery strategies.

Our preliminary studies demonstrated that specific-mode electroacupuncture stimulation (SMES; 2/100 Hz, 3 mA, 6 s on/6 s off, 40 min) effectively opens the BBB, facilitating macromolecule penetration, including Evans blue, fluorescein isothiocyanate–dextran, and neural growth factor [19,20]. The underlying mechanism of SMES in BBB modulation involves the alteration of tight junctions between endothelial cells, transiently disrupting the integrity of the BBB without causing permanent damage or inflammation [21]. This controlled disruption of the BBB effectively enhances the delivery of ABX to glioma tissue, resulting in a significantly higher accumulation at the tumor site than conventional intravenous administration, thereby exerting potent antitumor effects [22].

Although our study demonstrated SMES-mediated macromolecular drug delivery across the BBB, enhancing intracerebral accumulation and therapeutic efficacy, prospective clinical trials are required to validate the clinical benefits of SMES-induced BBB opening for ABX delivery in recurrent malignant gliomas (rMG). Accordingly, we initiated a single-arm, single-center phase II clinical trial to evaluate the safety and efficacy of SMES+ABX therapy for rMG.

Methods

Study Design

This was a single-center, single-arm, open-label phase II clinical trial conducted at the Third Affiliated Hospital of Zhejiang Chinese Medical University. The protocol was approved by the Institutional Review Board on January 15, 2025 (approval number: ZSLL-KY-2024-079-01). The study procedure included a brief baseline assessment completed within 48 h after enrollment, an 18-week active treatment phase (a total of 6 cycles), and a subsequent 1-year survival follow-up phase. Patient enrollment is scheduled to begin in May 2025, and the final survival follow-up of all patients is expected to be completed by January 2028. Eligible patients with malignant glioma will receive SMES+ABX therapy, with a 4-month progression-free survival (4m-PFS) rate serving as the primary end point. The secondary end points include adverse events (AEs), disease control rate (DCR), objective response rate (ORR), duration of disease control, duration of response, Neurological Assessment in Neuro-Oncology (NANO) score, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (EORTC QLQ-C30), progression-free survival (PFS), and OS. Exploratory biomarker end points include multiomics profiling (transcriptomics, proteomics, and metabolomics) of serial peripheral blood samples collected at baseline and during treatment, aiming to identify potential circulating biomarkers associated with treatment response or resistance mechanisms. The study protocol adheres to the Standards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA 2010) and SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [23,24].

Sample Size

The sample size was determined using a Simon 2-stage design, with the primary efficacy end point being the 4m-PFS rate. Previous studies have shown that the 4-month PFS rate for temozolomide in patients with recurrent World Health Organization (WHO) grade 4 glioblastoma does not exceed 25.7% [25,26]. This study set a target threshold of 60% for the 4m-PFS rate based on 3 aspects. First, preliminary exploratory pilot trials conducted by our team suggested activity signals indicating disease stability beyond 4 months in a small number of patients. Second, preclinical studies have confirmed that the combination of SMES and ABX can increase intratumoral drug concentration by approximately 5.5 times and significantly extend animal survival. In addition, focused ultrasound combined with ABX technology, based on a similar principle, achieved a 42% DCR in the first human trial. Together, this evidence supports setting an efficacy goal higher than that of conventional chemotherapy in this study, with the aim of screening treatment strategies with sufficient potential for subsequent confirmatory trials to be conducted. According to calculations using the R software (R Foundation for Statistical Computing; version 4.3.1; 1-sided α=.05, power=80%), a total of 20 patients are planned to be enrolled (including an approximate 20% dropout cushion). In the first stage, 9 patients will be enrolled; if fewer than 3 achieve PFS at 4 months, the trial will be terminated early. If the standard is met, 7 additional

XSL•FO
RenderX

patients will be recruited in the second stage. If the final overall 4m-PFS rate is less than 56.3% (9/16 patients), the treatment regimen will be considered ineffective.

Participants and Recruitment Strategy

Patients with rMG will be recruited from the Zhejiang Chinese Medical University, its Third Affiliated Hospital, and other collaborating medical institutions. Recruitment will be primarily conducted through outpatient clinics and inpatient units of the oncology, neurosurgery, and rehabilitation medicine departments. Prior to enrollment, all potential participants will be thoroughly informed of the potential benefits and associated risks related to the exercise intervention component of the study. Written informed consent will be obtained from each participant, who will also be explicitly advised of their right to withdraw from the study at any time without providing a reason for doing so. Upon the completion of the intervention, the participants will be debriefed on the study findings. The research outcomes will be disseminated through publications in peer-reviewed journals and by presenting them at academic conferences.

rMG Diagnostic Criteria

In this study, postoperative recurrence of malignant glioma was defined as meeting all criteria: (1) WHO grade IV classification per the Chinese Anti-Cancer Association Guidelines for Integrated Diagnosis and Treatment of Glioma (V2.0_2025, dated January 10, 2025) [27]; (2) prior surgical resection; and (3) neuroradiologically confirmed recurrence through cranial magnetic resonance imaging (MRI), which may show either the progression of original lesions or the emergence of new lesions, as adjudicated by board-certified neuro-oncologists. Participant selection will follow the predefined inclusion or exclusion criteria. Textbox 1 details the inclusion, exclusion, withdrawal, dropout, and termination criteria for this study.

XSL•FO
RenderX

Textbox 1. Inclusion, exclusion, withdrawal, dropout, and termination criteria.

1. Inclusion criteria

- Histologically confirmed World Health Organization grade IV glioma according to the Chinese Anti-Cancer Association Guidelines for Integrated Diagnosis and Treatment of Glioma (V2.0_2025, dated January 10, 2025)
- Radiologically confirmed disease progression per the RANO 2.0 criteria following surgical resection, with ≥1 measurable contrast-enhancing lesion
- Age 18-70 years (inclusive), any gender
- Dexamethasone dose for mass effect: <6 mg daily (stable for 7 days) or <6 mg average during tapering. Nonmass-effect steroid use was permitted
- Functional status: Karnofsky Performance Status Scale score ≥40
- Adequate organ function (within 14 days): (1) hemoglobin ≥90 g/L; (2) white blood cell ≥3.0×109/L; (3) absolute neutrophil count ≥1500/μL (white blood cell×neutroph); (4) platelets ≥100×109/μL; (5) total bilirubin ≤5×upper limit of normal (ULN); (6) aspartate aminotransferase ≤3×ULN (with bilirubin ≤3×ULN); (7) creatinine ≤1.5 mg/dL or estimated glomerular filtration rate 30-90 mL/min
- Tolerability of electroacupuncture and expected compliance with treatment
- Conscious with preserved:
  - Nociception or discrimination
  - Basic communication capacity
- Willingness to provide written informed consent

1. Exclusion criteria

- Uncontrolled seizure disorder
- Concurrent participation in other interventional trials or within 30 days post participation
- Currently receiving albumin-bound paclitaxel or similar drug treatment
- Severe allergy to albumin-bound paclitaxel or similar compounds
- Pregnant or breastfeeding women
- Diseases affecting cognitive function, such as congenital dementia, alcoholism, drug abuse, or psychotropic substance abuse
- Skin infections at the acupuncture sites
- Patients with conductive foreign objects in their bodies
- Contraindications to gadolinium-enhanced magnetic resonance imaging
- Other acute or chronic diseases, mental illnesses, or abnormal laboratory test values that may increase risks associated with study participation or study drug administration, or interfere with interpretation of study results, and individuals determined by the investigator as ineligible for study participation
- Concurrent other types of antitumor treatments during the trial, such as chemotherapy, radiotherapy, targeted therapy, or immunotherapy

1. Withdrawal, drop-out, and termination criteria

- Individuals who were enrolled but subsequently found not to satisfy the predefined inclusion criteria
- Participants with major protocol deviations related to safety assessments
- Participants who developed serious adverse events or medical complications that necessitated discontinuation of trial participation

Interventional Methods

Enrolled patients will receive intravenous ABX at a dose of 135 - 175 mg/m² on day 1 of each 21-day treatment cycle, with concurrent SMES intervention (Figure 1).

XSL•FO
RenderX

Figure 1. Schematic overview of the experimental protocols. The treatment regimen consisted of 6 cycles, each lasting 21 days, with combined specific-mode electroacupuncture stimulation (SMES) and albumin-bound paclitaxel (ABX) therapy administered on day 1 of each cycle. Adverse events (AEs) were assessed on days 7, 14, and 21 of each cycle. Efficacy was evaluated at the end of each cycle. Postintervention monitoring continued for 1 year, including the monthly follow-up assessments. 4m-PFS: 4-month progression-free survival; DCR: disease control rate; DDC: duration of disease control; DOR: duration of response; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30; NANO: Neurological Assessment in Neuro-Oncology; ORR: objective response rate; OS: overall survival; PVBC: peripheral blood samples collected; PFS: progression-free survival.

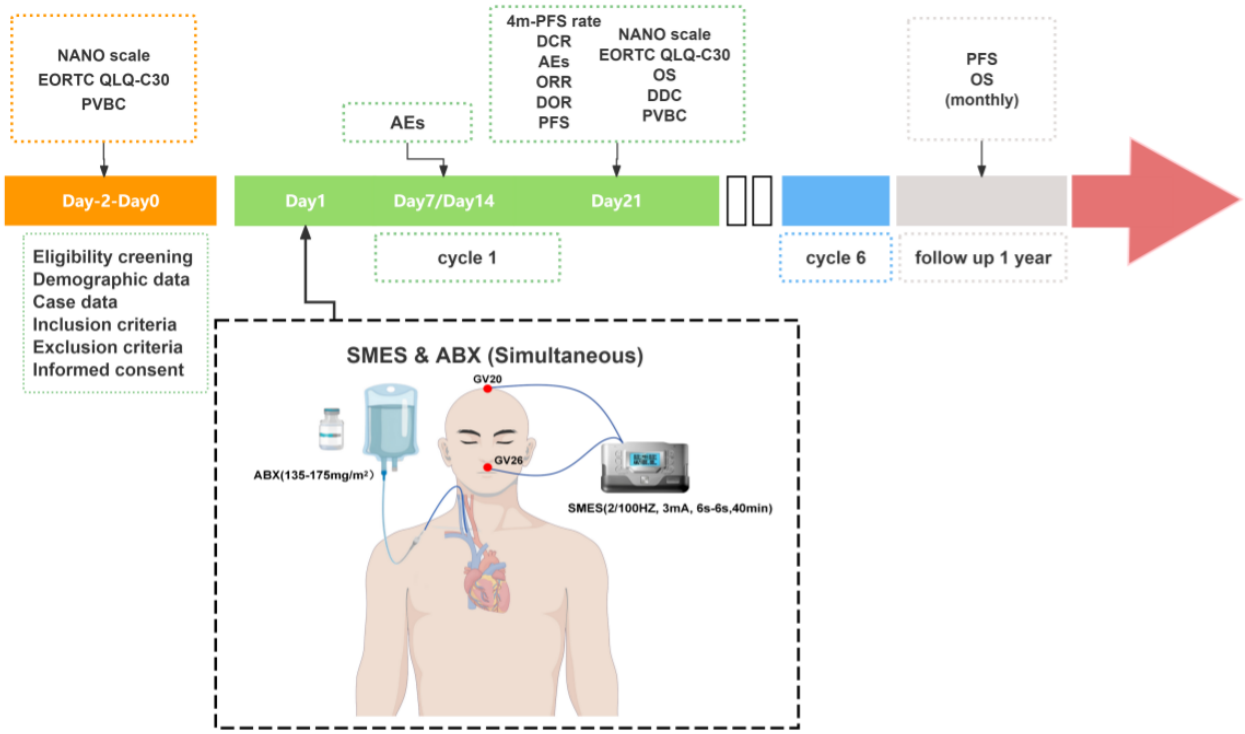

ABX Treatment Protocol

ABX was administered intravenously at a dose of 175 mg/m² over 40 min. A dose reduction to no less than 135 mg/m² was required in cases of (1) severe neutropenia (absolute neutrophil count <500/mm³ for ≥7 d) or (2) grade ≥3 sensory neuropathy. Treatment was withheld in cases of grade 3 neurotoxicity until resolution to grade ≤2.

SMES Therapeutic Protocol

The patient was placed in the supine position with full-body relaxation. The skin surrounding GV20 (Baihui) and GV26 (Shuigou) was disinfected. Acupuncture was then administered at DU20, and a 0.25 mm × 40 mm filiform needle was inserted horizontally at a depth of 15 - 20 mm. At GV26, a 0.18 mm × 25 mm needle was obliquately inserted toward the nasal septum at a depth of 9 - 15 mm. The needles were then connected to an electroacupuncture device (HANS-200; Jisheng Medical Technology). Electrical stimulation was delivered in a dense-disperse wave mode (2/100 Hz) at 1.5 - 3.0 mA, which was adjusted to elicit mild muscle twitching near GV26. The treatment lasted 40 minutes, with a 6 seconds on/6 seconds off duty cycle controlled by an external device. The start and end times of SMES device operation were synchronized with the initiation and completion of ABX infusion.

Assessment and Follow-Up

Tumor response will be evaluated according to the RANO 2.0 criteria [28], incorporating MRI findings and clinical outcomes to determine therapeutic efficacy. Baseline assessments will be conducted within 2 days before treatment initiation, followed by serial evaluations at 3-week intervals, resulting in a total of 7 assessments. AEs will be graded according to CTCAE v5.0. Additionally, the causality between AEs and the investigational drug will be assessed to identify treatment-related AEs. Furthermore, patient-reported outcomes, including quality of life and performance status, will be systematically evaluated.

Post-study follow-up will be conducted for 1 year with monthly telemedicine consultations (via telephone or WeChat) to document survival outcomes. As illustrated in Figure 1, a comprehensive analysis will be performed for each parameter with predefined time frames.

Concomitant Medications

Concomitant medications for AE management were permitted during the trial, including acid-suppressing agents for gastric protection, hepatoprotective drugs for liver function abnormalities, low-dose corticosteroids for allergy prophylaxis, recombinant human granulocyte colony-stimulating factor (rhG-CSF) or pegylated G-CSF to stimulate leukopoiesis, and recombinant thrombopoietin to stimulate thrombopoiesis. Detailed records were maintained for all concomitant medications, including the generic name, indication, single dose (with units), administration frequency, route, start date, and current status (ongoing or discontinued).

XSL•FO
RenderX

Rationale for ABX Dose Selection

Although the manufacturer-recommended maximum dose of ABX for breast cancer is 260 mg/m², no clinical guidelines exist for its use in malignant gliomas. Previous studies have demonstrated that ultrasound-mediated BBB disruption significantly enhances ABX delivery in glioma models, achieving 3 - to 5-fold higher brain concentrations than controls [29]. Similarly, our findings indicated that SMES transiently disrupts the BBB, resulting in a 5.5-fold increase in ABX tumor accumulation relative to controls [22]. Importantly, a phase I clinical trial confirmed the safety of 260 mg/m² ABX when combined with repeated ultrasound-induced BBB opening [17].

However, our preliminary data revealed that SMES combined with ABX doses exceeding 175 mg/m² induced severe dose-limiting toxicities, primarily myelosuppression, resulting in significantly reduced treatment tolerability. Given the discrepancy between preclinical evidence and early clinical safety data, we selected a reduced ABX dose range (135 - 175 mg/m²) to ensure patient safety while maintaining the therapeutic feasibility. The dosing strategy was determined by integrating the enhanced drug exposure resulting from BBB disruption technology with a rigorous assessment of clinical tolerability, thereby ensuring scientific validity and clinical relevance.

Baseline Evaluation and Outcome Measure

Following screening and enrollment, baseline assessments will be conducted within 48 h of enrollment. Comprehensive data collection will include (1) informed consent documentation, (2) medical history, (3) physical examination, (4) laboratory tests, (5) MRI scans (contrast-enhanced brain MRI, magnetic resonance spectroscopy, perfusion-weighted imaging, etc), (6) NANO scale, (7) EORTC QLQ-C30, (8) tumor characteristics (pathological and genomic analyses), (9) detailed demographic information, and (10) peripheral venous blood collection (≈5 mL per cycle) for biobanking and optional future exploratory multiomics studies will be performed. Further specifications are provided in Table 1 and Textbox 2.

XSL•FO
RenderX

Table . SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) schedule of enrollment, interventions, and assessments.

Study period	Baseline (day 2 to day 0)	Intervention period (treatment cycles 1 through 6)				Follow-up (1 y)
		Day 1	Day 7	Day 14	Day 21	
Eligibility screening	✓					
Demographic data	✓					
Case data	✓					
Inclusion criteria	✓					
Exclusion criteria	✓					
Informed consent	✓					
SMES+ABX[a]		✓				
Outcome assessment						
4m-PFS rate[b]					✓	
DCR[c]					✓	
AEs[d]			✓	✓	✓	
ORR[e]					✓	
NANO[f] scale	✓				✓	
EORTC QLQ-C30[g]	✓				✓	
PFS[h]					✓	✓ [i]
OS[j]					✓	✓ [i]
DOR[k]					✓	
DDC[l]					✓	
PVBC[m]	✓				✓	

[a]SMES+ABX: specific-mode electroacupuncture stimulation+albumin-bound paclitaxel.

[b]4m-PFS rate: 4-month progression-free survival rate.

[c]DCR: disease control rate.

[d]AE: adverse event.

[e]ORR: objective response rate.

[f]NANO: Neurological Assessment in Neuro-Oncology.

[g]EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30.

[h]PFS: progression-free survival.

[i]Monthly follow-up assessments were conducted during the observation period.

[j]OS: overall survival.

[k]DOR: duration of response.

[l]DDC: duration of disease control.

[m]PVBC: peripheral venous blood collection.

XSL•FO
RenderX

Textbox 2. Primary end point and secondary end points.

Primary end point

- The primary objective of this study is to determine the 4-month progression-free survival rate, defined as the proportion of participants in the full analysis set (FAS) population who remained progression-free beyond 4 months. This end point will be assessed in this study.

Secondary end points

- The disease control rate was defined as the proportion of patients in the FAS population who achieved complete response (CR), progressive disease (PR), or stable disease, as assessed per the RANO 2.0 criteria.
- The objective response rate was calculated as the percentage of FAS patients exhibiting CR or PR, evaluated according to the RANO 2.0 guidelines.
- Overall survival is measured from the initiation of specific-mode electroacupuncture stimulation+albumin-bound paclitaxel (SMES+ABX) therapy until death or censoring at the last follow-up.
- PFS was defined as the time from the first SMES+ABX administration to disease progression (per RANO 2.0), intolerable adverse events, or death, whichever occurred first.
- Duration of response was defined, for patients who achieved an objective response (complete or partial response) to SMES+ABX therapy, as the time from the first documented date of objective response to the date of first documented progressive disease or death from any cause, whichever occurred first.
- Duration of disease control was defined as the time interval from the first radiologically confirmed achievement of disease control (CR, PR, or stable disease) to the first documented progressive disease or death from any cause (whichever occurred first) in patients who received SMES+ABX combination therapy and achieved disease control.
- The NANO scale is a standardized quantitative instrument that evaluates neurofunctional status in patients with neuro-oncology through the systematic assessment of 8 key domains (e.g. motor, sensory, and language functions).
- The European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 questionnaire is a globally validated multidimensional instrument for assessing cancer-related quality of life across 5 functional domains (physical, role, cognitive, emotional, and social) and symptom scales.
- Safety assessments include treatment-emergent adverse events, treatment-related adverse events, and serious adverse events. Treatment-emergent adverse events were defined as adverse events occurring between treatment initiation and study completion. Treatment-related adverse events are adverse events specifically attributable to investigational drugs. Serious adverse events were classified as events resulting in death, life-threatening illness, disability, hospitalization or prolonged hospitalization, or persistent or severe incapacity.

Analysis Populations

To comprehensively evaluate the efficacy and safety of this single-arm trial and address possible protocol deviations, three analysis populations were predefined in this study. The first is the full analysis set (FAS), in which all patients who signed the informed consent received at least 1 complete cycle of SMES+ABX combination therapy and underwent at least 1 valid posttreatment tumor efficacy assessment. As a single-arm exploratory trial, the FAS is the core population for the primary efficacy analysis in this study (in particular, for decision-making regarding the primary end point of the 4-month PFS rate), aiming to reflect the potential effect of the treatment regimen in eligible patients as fully as possible. The second analysis population is the safety set, in which all patients who signed the informed consent form received at least 1 complete cycle of SMES+ABX combination therapy and underwent at least 1 posttreatment safety assessment. This group will be used for all the safety analyses. The third analysis population is the per protocol set, which is a subset of the FAS, consisting of patients who strictly adhered to the study protocol throughout the treatment period. The per protocol set will serve as a supportive analysis, used to evaluate the treatment effect under ideal compliance, and to compare with FAS results for sensitivity analysis.

Statistical Analysis Methods

Efficacy assessment will use an analysis framework based on the RANO 2.0 criteria. The primary end point (4m-PFS rate) will be evaluated strictly according to the predefined decision rules of the Simon 2-stage design, with its point estimate and corresponding 95% exact binomial CI (Clopper-Pearson method) calculated. Time-to-event end points (PFS, OS, duration of response, and duration of disease control) will be descriptively analyzed using the Kaplan-Meier method, with median survival times and their 95% CIs reported. An exploratory univariate Cox proportional hazards model was used to assess the association between baseline prognostic factors and survival outcomes. Categorical efficacy end points (such as ORR and DCR) will be reported as frequencies (percentages) with 95% CIs. Continuous variables (such as NANO scores and QLQ-C30 scale scores) will be summarized at each visit using descriptive statistics (mean, standard deviation), and their longitudinal changes will be presented using figures or charts. Given the small sample size, any complex statistical models for analyzing longitudinal changes (such as linear mixed models) will be regarded as exploratory analyses. Safety data will be graded according to the CTCAE v5.0 criteria, and the incidence of AEs and their 95% CIs will be calculated using the Clopper-Pearson method.

Considering that this study is a small-sample, single-arm, exploratory phase II trial, all statistical analyses focused on describing the point and interval estimates for efficacy and

XSL•FO
RenderX

safety. Primary conclusions will be drawn based on the decision rules for the primary end point (4m-PFS) outlined in the Simon 2-stage design. All multivariate model analyses and subgroup analyses for secondary end points are exploratory and hypothesis-generating in nature; their results should be interpreted with caution and require validation in future large-scale confirmatory studies. Owing to the limited sample size and low statistical power, the analysis of secondary end points in this study was primarily intended to describe trends and generate hypotheses rather than to conduct confirmatory testing. Statistical analyses will be performed using R 4.3.1 (survival and lme4 packages) and the SAS software (version 9.4; SAS Institute Inc).

Ethical Considerations

This study was approved by the Institutional Review Board of the Third Affiliated Hospital of Zhejiang Chinese Medical University (approval no. ZSLL-KY-2024-079-01). Before enrollment, potential participants were fully informed of the study protocol, expected benefits, and potential risks using standardized procedures. Separate optional consent was obtained for the collection and long-term storage of peripheral blood samples (≈5 mL per cycle) for future exploratory multiomics research (eg, transcriptomics, proteomics, metabolomics). Participation in this ancillary component is voluntary and does not affect the eligibility for the main trial. Written informed consent was obtained after ensuring a complete understanding of the study. The trial adhered to the principles of the Declaration of Helsinki. All research data were deidentified and stored in encrypted databases to ensure participant confidentiality and data security. The study protocol was registered at ClinicalTrials.gov (Identifier: NCT06818331) and was continuously monitored by an independent data monitoring committee (IDMC). In addition, if there are any changes to the research protocol, we will submit a written application to the research ethics committee, and the committee members will decide whether the protocol needs to be amended.

Data Management and Monitoring

The research data will be entered into a dedicated database by trained independent personnel according to the study protocol. A multitiered quality control strategy will be implemented to ensure data accuracy, completeness, and consistency of the data. This includes the following: (1) source data verification (100% verification of source documentation for the primary end point [4m-PFS rate]; a minimum of 20% random sampling of secondary end point and safety data for source data verification); (2) medical review (all serious AEs and protocol deviation records will be reviewed by the study physicians); and (3) logical consistency checks (automated verification of data logic and visit timeline adherence using predefined scripts). All data queries will be generated and tracked for resolution using an electronic system. All study-related original data will be stored at the Third Affiliated Hospital of Zhejiang Chinese Medical University.

Any deviations from the approved protocol will be documented in a dedicated protocol deviation log. Deviations will be categorized as minor or major, based on their potential impact on participant safety or data integrity. All major deviations will be promptly reported to the principal investigator and the IDMC.

An independent IDMC will be established to safeguard participant rights, data reliability, and scientific integrity. The committee will consist of independent experts not involved in the trial, including specialists in neuro-oncology, clinical statistics, medical ethics, acupuncture therapeutics, and neuroradiology. The IDMC will review unblinded cumulative safety and efficacy data upon completion of stage 1 of the Simon 2-stage design to assess the risk-benefit profile and provide recommendations for proceeding to stage 2. Following trial completion, the IDMC will review the final data to ensure result integrity. All major protocol deviations and AEs will be reported to the committee, which has the authority to recommend early trial termination based on safety or efficacy concerns. The study will also undergo ongoing oversight and periodic review by the Institutional Review Board of the hospital.

Results

Patient enrollment is scheduled to begin in May 2025. The treatment and primary efficacy assessment phases are anticipated to be completed by January 2027 (allowing for staggered enrollment and a 4-month treatment period for the last enrolled patient). The final survival follow-up for all patients is anticipated to be completed by January 2028 (ie, 1 year after the last patient completes treatment). Data management is currently ongoing, and formal statistical analyses have not yet been performed.

Discussion

Dilemmas and Clinical Needs in the Treatment of rMG

Malignant gliomas, especially the most aggressive form, glioblastoma (WHO grade IV), continue to have an extremely poor prognosis despite the use of standard multimodal treatments, including maximal safe resection, radiotherapy, and temozolomide chemotherapy [30]. The BBB is one of the major factors contributing to unfavorable outcomes, as it severely restricts the delivery of potentially effective chemotherapeutic agents to infiltrating tumor cells [31]. ABX is an effective microtubule stabilizer that has demonstrated efficacy in various solid tumors and exhibits strong cytotoxic effects against glioma cell lines in vitro [32-34]. However, its clinical application in gliomas has been hampered by insufficient accumulation in the brain following systemic administration [14,15]. Current research shows that 3 strategies—nanoformulation engineering, coadministration with BBB modulators, and physical or chemical disruption of the BBB—can enhance the ability of drugs to cross the BBB [35-37]. Nevertheless, there are many challenges in clinical translation owing to issues such as technical complexity, high costs, and safety concerns. Thus, developing strategies that can open the BBB in a brief, safe, and controllable manner is a key challenge in neuro-oncology. The SMES combined with ABX therapy explored in this study aims to provide a novel, noninvasive, cost-effective, and easy-to-operate solution to address this challenge.

XSL•FO
RenderX

Innovation and Design Rationality

This study innovatively applies a physical BBB modulation technology, SMES, in a prospective manner to the chemotherapy treatment of patients with rMG. The theoretical foundation is based on our previous series of studies, which demonstrated that electroacupuncture stimulation with specific parameters (2/100 Hz, 3 mA, 6 s on/6 s off) can safely increase BBB permeability by transiently and reversibly regulating tight junction proteins (such as ZO-1 and Occludin), thereby promoting the delivery of macromolecules (such as Evans blue and fluorescein isothiocyanate–dextran), and even therapeutic proteins (such as nerve growth factor) into the brain [19,20]. More importantly, in rat models of glioma, SMES significantly increased the accumulation of ABX in tumor tissue (by approximately 5.5 times) and enhanced its antitumor efficacy [22]. These studies provide robust preclinical evidence for translational research.

In terms of study design, we adopted a single-center, single-arm, Simon 2-stage phase II clinical trial, which is a rational choice for evaluating the preliminary efficacy and safety of an innovative therapy. The primary end point was set as 4m-PFS, aligning with conventions in clinical trials of recurrent high-grade gliomas and enabling a rapid and effective assessment of whether the treatment brings about clinically meaningful disease control [25,26]. The secondary end points encompass multiple dimensions, including efficacy (such as ORR, DCR, and OS), neurological function (NANO scale), patient-reported quality of life (EORTC QLQ-C30), and safety, aiming to comprehensively evaluate the SMES+ABX therapy. The chosen ABX dose (135 - 175 mg/m²) was determined based on mechanistic, safety, and clinical practice considerations. The upper limit (175 mg/m²) was set as the starting dose based on dose-limiting toxicity (mainly grade 3 - 4 myelosuppression) observed in our previous studies, ensuring a safe initiation. The lower limit (135 mg/m²) was established considering the following: preclinical studies showed that SMES could increase ABX concentration in tumors by approximately 5.5 times [22]; from a drug exposure perspective, even with a reduced systemic dose, enhanced delivery via SMES could still achieve effective local concentrations; published clinical trials combining focused ultrasound BBB opening with ABX have confirmed the safety of a 260 mg/m² dose [17], so 135 mg/m² ensures a sufficient safety margin; and this dose also conforms to the standard oncology dose reduction (approximately 25%), which facilitates clinical management while retaining clear pharmacological activity after reduction. This dosing range reflects a cautious balance between therapeutic potential and patient safety under the SMES-enhanced delivery mechanism.

Explanation of Expected Outcomes and Potential Mechanisms

If this study meets its predefined primary end point (ie, ≥9 out of 16 patients achieve 4m-PFS, with an overall response rate of ≥56.3%), it will strongly suggest that the SMES combined with the ABX regimen has potential advantages in controlling the progression of rMG. The reason for such positive outcomes may lie in the SMES, through its specific electrophysiological stimulation, acting on the trigeminal nerve-cerebrovascular system, resulting in local cerebral vasodilation and hemodynamic changes, which in turn lead to a transient and reversible structural remodeling of the tight junctions between endothelial cells [21,38-43]. This change in BBB permeability allows more ABX to cross the barrier, achieving effective concentrations at the tumor site, thereby exerting strong cytotoxic effects.

Regarding the secondary end points, the regulatory effect of SMES on local cerebral hemodynamics [39-43], combined with the increased concentration of ABX at the tumor site [22], may allow the patients' neurological status (NANO score) to remain stable or even improve. Moreover, the maintenance or improvement of neurological function, together with symptom alleviation brought about by disease control, may be reflected in stable or improved quality of life scores (EORTC QLQ-C30). Safety was the top priority in this study. As a modified electroacupuncture therapy, SMES should have a favorable safety profile; AEs are still expected to be mainly due to the known toxicities of ABX (such as sensory neuropathy and myelosuppression). However, it is essential to closely monitor whether increased ABX concentrations in the brain induced by SMES affect central nervous system toxicity.

Limitations

This study had several inherent design limitations. First, the single-arm design lacks concurrent randomized controls; therefore, efficacy assessments rely on historical controls, which may introduce selection bias and be influenced by advances in treatment over time, thus limiting the certainty of the conclusions. Second, although the Simon 2-stage design was used to improve exploratory efficiency, the limited sample size (planned for 16 cases, with up to 20 cases to account for dropouts) may result in insufficient statistical power for secondary end points and make it difficult to identify rare AEs or differences in efficacy among the subgroups. Third, patients with rMG exhibit heterogeneity in molecular pathology and treatment history, which may potentially affect their treatment response. Fourth, the planned multiomics analyses of serial peripheral blood samples are exploratory and hypothesis-generating; their technical feasibility, biological interpretability, and clinical relevance remain to be determined, and any findings will require independent validation in larger cohorts. However, phase II clinical trials are essentially exploratory rather than confirmatory in nature; this study aims to provide preliminary evidence for the potential benefits of the SMES combined with ABX regimen. If the results are positive, a phase III randomized controlled trial will be required to further validate its efficacy.

Conclusion

This study adopted the Simon 2-stage design to conduct a single-arm phase II clinical trial enrolling 16 patients, aiming to evaluate the efficacy and safety of SMES combined with ABX in the treatment of rMG. The primary end point of the study is 4m-PFS, while the secondary end points include ORR, DCR, OS, NANO score, and EORTC QLQ-C30. Safety assessments include treatment-related AEs and serious AEs. If successful, this study will provide an innovative, noninvasive, and cost-effective treatment strategy for this disease, which has

XSL•FO
RenderX

an extremely poor prognosis, and will lay a foundation for subsequent studies.

Acknowledgments

We extend our sincere gratitude to the physicians from the Departments of Acupuncture, Oncology, Radiology, and Neurosurgery at the Third Affiliated Hospital of Zhejiang Chinese Medical University, and to CIXINGTECHNOLOGY (China) for their expert guidance and collaboration in study protocol development.

Funding

This work was supported by the National Natural Science Foundation of China (grant no. 82474626), the Joint Science and Technology Project of the National Administration of Traditional Chinese Medicine and Zhejiang Provincial Administration of Traditional Chinese Medicine (grant no. GZY-ZJ-KJ-24021), the 2024 General Scientific Research Project of Zhejiang Provincial Department of Education—Special Program for Reform in Professional Degree Graduate Education Models (grant no. Y202456327), and the 2024 Zhejiang University Students' Science and Technology Innovation Activity Plan (New Talent Program; grant no. 2024R410C066).

Authors' Contributions

ZJ contributed to conceptualization, investigation, formal analysis, and original draft preparation. TJ and YZ were responsible for data curation and manuscript drafting. QC conducted visualization and investigation tasks. ZD and QY provided resources, supervision, and validation. KQ and LG performed visualization and participated in manuscript review and editing. CM and XL oversaw study design, secured funding, managed resources, and supervised manuscript review and editing. All authors approved the final version for submission.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist for protocol of a clinical trial.
[DOCX File, 30 KB - resprot_v15i1e84593_app1.docx]

References

1. Stupp R, Mason WP, van den Bent MJ, et al. Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. N Engl J Med 2005 Mar 10;352(10):987-996. [doi: 10.1056/NEJMoa043330] [Medline: 15758009]
2. Stupp R, Taillibert S, Kanner A, et al. Effect of tumor-treating fields plus maintenance temozolomide vs maintenance temozolomide alone on survival in patients with glioblastoma: a randomized clinical trial. JAMA 2017 Dec 19;318(23):2306-2316. [doi: 10.1001/jama.2017.18718] [Medline: 29260225]
3. Jones C, Baker SJ. Unique genetic and epigenetic mechanisms driving paediatric diffuse high-grade glioma. Nat Rev Cancer 2014 Oct;14(10):651-661. [doi: 10.1038/nrc3811] [Medline: 25230881]
4. Ostrom QT, Patil N, Cioffi G, Waite K, Kruchko C, Barnholtz-Sloan JS. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013-2017. Neuro Oncol 2020 Oct 30;22(12 Suppl 2):iv1-iv96. [doi: 10.1093/neuonc/noaa200] [Medline: 33123732]
5. Semreen AM, Alsoud LO, Semreen MH, et al. Multi-omics analysis revealed significant metabolic changes in brain cancer cells treated with paclitaxel and/or topotecan. Heliyon 2024 Nov 15;10(21):e39420. [doi: 10.1016/j.heliyon.2024.e39420] [Medline: 39524752]
6. Liu X, Wang J, Wu LJ, Trinh B, Tsai RYL. IMPDH inhibition decreases TERT expression and synergizes the cytotoxic effect of chemotherapeutic agents in glioblastoma cells. Int J Mol Sci 2024 May 30;25(11):5992. [doi: 10.3390/ijms25115992] [Medline: 38892179]
7. Parkin HC, Shopperly LK, Perez MR, Willerth SM, Manners I. Uniform block copolymer nanofibers for the delivery of paclitaxel in 2D and 3D glioblastoma tumor models. Biomater Sci 2024 Oct 8;12(20):5283-5294. [doi: 10.1039/d4bm00480a] [Medline: 39246052]
8. Barretina J, Caponigro G, Stransky N, et al. Addendum: the Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature New Biol 2019 Jan;565(7738):E5-E6. [doi: 10.1038/s41586-018-0722-x] [Medline: 30559381]
9. Chang SM, Kuhn JG, Robins HI, et al. A phase II study of paclitaxel in patients with recurrent malignant glioma using different doses depending upon the concomitant use of anticonvulsants: a North American Brain Tumor Consortium report.

XSL•FO
RenderX

Cancer 2001 Jan 15;91(2):417-422. [doi: 10.1002/1097-0142(20010115)91:2<417::aid-cncr1016>3.0.co;2-9] [Medline: 11180089]
10. Chang SM, Kuhn JG, Rizzo J, et al. Phase I study of paclitaxel in patients with recurrent malignant glioma: a North American Brain Tumor Consortium report. J Clin Oncol 1998 Jun;16(6):2188-2194. [doi: 10.1200/JCO.1998.16.6.2188] [Medline: 9626220]
11. Heimans JJ, Taphoorn MJ, Vermorken JB, et al. Does paclitaxel penetrate into brain tumor tissue? J Natl Cancer Inst 1995 Dec 6;87(23):1804-1805. [doi: 10.1093/jnci/87.23.1804-a] [Medline: 7473842]
12. Authier N, Gillet JP, Fialip J, Eschalier A, Coudore F. Assessment of neurotoxicity following repeated cremophor/ethanol injections in rats. Neurotox Res 2001 Jul;3(3):301-306. [doi: 10.1007/BF03033269] [Medline: 15111255]
13. Yan H, Miao H, Hu J, et al. Oxidative stress induced paclitaxel-derived carbon dots inhibit glioblastoma proliferation and EMT process. J Nanobiotechnology 2025 Apr 23;23(1):310. [doi: 10.1186/s12951-025-03406-x] [Medline: 40269908]
14. Kim YB, Lee SH, Kasala D, et al. Potent therapeutic efficacy of intranasally deliverable paclitaxel modified with pH-sensitive and PEGylated polymeric micelle against glioblastoma. J Control Release 2025 Jun 10;382:113711. [doi: 10.1016/j.jconrel.2025.113711] [Medline: 40204132]
15. Slika H, Shahani A, Gattu K, et al. Intracranial nanogel pellets carrying temozolomide and paclitaxel for adjuvant brain cancer therapy. Mol Pharm 2025 Jan 6;22(1):131-141. [doi: 10.1021/acs.molpharmaceut.4c00708] [Medline: 39666995]
16. Belyaeva AA, Averchuk AS, Rozanova NA, et al. Thermosensitive injectable fibrillar gels based on cellulose nanocrystals grafted with poly(N-isopropylacrylamide) as biocompatible brain implants. Carbohydr Polym 2024 Dec 15;346:122596. [doi: 10.1016/j.carbpol.2024.122596] [Medline: 39245487]
17. Sonabend AM, Gould A, Amidei C, et al. Repeated blood-brain barrier opening with an implantable ultrasound device for delivery of albumin-bound paclitaxel in patients with recurrent glioblastoma: a phase 1 trial. Lancet Oncol 2023 May;24(5):509-522. [doi: 10.1016/S1470-2045(23)00112-2] [Medline: 37142373]
18. Habashy KJ, Dmello C, Chen L, et al. Paclitaxel and carboplatin in combination with low-intensity pulsed ultrasound for glioblastoma. Clin Cancer Res 2024 Apr 15;30(8):1619-1629. [doi: 10.1158/1078-0432.CCR-23-2367] [Medline: 38295144]
19. Zhang S, Gong P, Zhang J, et al. Specific frequency electroacupuncture stimulation transiently enhances the permeability of the blood-brain barrier and induces tight junction changes. Front Neurosci 2020;14:582324. [doi: 10.3389/fnins.2020.582324] [Medline: 33122995]
20. Zhao Y, Mao X, Wang H, et al. The influence of electronic acupuncture at a specific frequency in facilitating the passage of NGF through the blood-brain barrier and its effect on learning and memory in MCAO/R rats. J Integr Neurosci 2022 Apr 8;21(3):79. [doi: 10.31083/j.jin2103079] [Medline: 35633160]
21. Gong P, Zhang S, Ren L, et al. Electroacupuncture of the trigeminal nerve causes N-methyl-D-aspartate receptors to mediate blood-brain barrier opening and induces neuronal excitatory changes. Front Cell Neurosci 2022;16:1020644. [doi: 10.3389/fncel.2022.1020644] [Medline: 36313622]
22. Ma C, Ye Q, Qian K, et al. Anti-glioma effect of paclitaxel mediated by specific mode electroacupuncture stimulation and the related role of the Hedgehog pathway. Brain Res Bull 2024 Jul;213:110985. [doi: 10.1016/j.brainresbull.2024.110985] [Medline: 38806118]
23. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg 2012;10(1):28-55. [doi: 10.1016/j.ijsu.2011.10.001] [Medline: 22036893]
24. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
25. Khan RB, Raizer JJ, Malkin MG, Bazylewicz KA, Abrey LE. A phase II study of extended low-dose temozolomide in recurrent malignant gliomas. Neuro Oncol 2002 Jan;4(1):39-43. [doi: 10.1093/neuonc/4.1.39] [Medline: 11772431]
26. McCracken DJ, Celano EC, Voloschin AD, Read WL, Olson JJ. Phase I trial of dose-escalating metronomic temozolomide plus bevacizumab and bortezomib for patients with recurrent glioblastoma. J Neurooncol 2016 Oct;130(1):193-201. [doi: 10.1007/s11060-016-2234-6] [Medline: 27502784]
27. Chinese Anti-Cancer Association guidelines for integrated diagnosis and treatment of glioma (V2.0_2025) [Article in Chinese]. : Chinese Anti-Cancer Association; 2025 URL: https://wap.cacakp.com/appguide/pdfDetail/id/415 [accessed 2026-01-14]
28. Wen PY, van den Bent M, Youssef G, et al. RANO 2.0: update to the response assessment in neuro-oncology criteria for high- and low-grade gliomas in adults. J Clin Oncol 2023 Nov 20;41(33):5187-5199. [doi: 10.1200/JCO.23.01059] [Medline: 37774317]
29. Zhang DY, Dmello C, Chen L, et al. Ultrasound-mediated delivery of paclitaxel for glioma: a comparative study of distribution, toxicity, and efficacy of albumin-bound versus cremophor formulations. Clin Cancer Res 2020 Jan 15;26(2):477-486. [doi: 10.1158/1078-0432.CCR-19-2182] [Medline: 31831565]
30. Vargas López AJ, Fernández Carballal C, Valera Melé M, Rodríguez-Boto G. Survival analysis in high-grade glioma: the role of salvage surgery. Neurologia (Engl Ed) 2023;38(1):21-28. [doi: 10.1016/j.nrleng.2020.04.032] [Medline: 36464224]
31. Idbaih A, Canney M, Belin L, et al. Safety and feasibility of repeated and transient blood-brain barrier disruption by pulsed ultrasound in patients with recurrent glioblastoma. Clin Cancer Res 2019 Jul 1;25(13):3793-3801. [doi: 10.1158/1078-0432.CCR-18-3643] [Medline: 30890548]

XSL•FO
RenderX

32. Gornstein E, Schwarz TL. The paradox of paclitaxel neurotoxicity: mechanisms and unanswered questions. Neuropharmacology 2014 Jan;76 Pt A:175-183. [doi: 10.1016/j.neuropharm.2013.08.016] [Medline: 23978385]
33. Dmello C, Sonabend A, Arrieta VA, et al. Translocon-associated protein subunit SSR3 determines and predicts susceptibility to paclitaxel in breast cancer and glioblastoma. Clin Cancer Res 2022 Jul 15;28(14):3156-3169. [doi: 10.1158/1078-0432.CCR-21-2563] [Medline: 35552677]
34. Barretina J, Caponigro G, Stransky N, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature New Biol 2012 Mar 28;483(7391):603-607. [doi: 10.1038/nature11003] [Medline: 22460905]
35. Malekpour MR, Hosseindoost S, Madani F, Kamali M, Khosravani M, Adabi M. Combination nanochemotherapy of brain tumor using polymeric nanoparticles loaded with doxorubicin and paclitaxel: an in vitro and in vivo study. Eur J Pharm Biopharm 2023 Dec;193:175-186. [doi: 10.1016/j.ejpb.2023.11.002] [Medline: 37926270]
36. Park J, Aryal M, Vykhodtseva N, Zhang YZ, McDannold N. Evaluation of permeability, doxorubicin delivery, and drug retention in a rat brain tumor model after ultrasound-induced blood-tumor barrier disruption. J Control Release 2017 Mar 28;250:77-85. [doi: 10.1016/j.jconrel.2016.10.011] [Medline: 27742444]
37. Zeng W, Kwan Law BY, Wai Wong VK, et al. HM30181A, a potent P-glycoprotein inhibitor, potentiates the absorption and in vivo antitumor efficacy of paclitaxel in an orthotopic brain tumor model. Cancer Biol Med 2020 Nov 15;17(4):986-1001. [doi: 10.20892/j.issn.2095-3941.2020.0128] [Medline: 33299648]
38. Ma C, Gan L, Wang H, et al. Transcriptomic analysis of rat cerebral cortex reveals the potential mechanism of electroacupuncture opening blood brain barrier. Front Neurosci 2022;16:834683. [doi: 10.3389/fnins.2022.834683] [Medline: 35281512]
39. Qian YF, Fan XN, Li YJ, et al. Effects of acupuncture at different acupoints on cerebral blood flow in cerebral ischemia model rats. Zhongguo Zhen Jiu 2009 Mar;29(3):213-216. [Medline: 19358505]
40. Tang ECH, Lau AYL, Au D, et al. Effects of acupuncture upon cerebral hemodynamics in cerebral small vessel disease: a pilot study. Cereb Circ Cogn Behav 2023;4:100168. [doi: 10.1016/j.cccb.2023.100168] [Medline: 37397268]
41. Hori E, Takamoto K, Urakawa S, Ono T, Nishijo H. Effects of acupuncture on the brain hemodynamics. Auton Neurosci 2010 Oct 28;157(1-2):74-80. [doi: 10.1016/j.autneu.2010.06.007] [Medline: 20605114]
42. Schwarz G, Litscher G, Sandner-Kiesling A. Pseudoparadoxical dissociation of cerebral oxygen saturation and cerebral blood flow velocity after acupuncture in a woman with cerebrovascular dementia: a case report. Neurol Res 2004 Sep;26(6):698-701. [doi: 10.1179/016164104225015912] [Medline: 15327762]
43. Kovalenko O, Chizhykova M, Dzhura N. Evaluation of the effects of acupuncture on cerebral hemodynamics in patients suffering from cerebrovascular pathology with duplex ultrasound Doppler imaging. J US China Med Sci 2018 Feb 28;15(1). [doi: 10.17265/1548-6648/2018.01.001]

Abbreviations

4m-PFS: 4-month progression-free survival
ABX: albumin-bound paclitaxel
AE: adverse event
BBB: blood-brain barrier
DCR: disease control rate
EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30
FAS: full analysis set
IDMC: independent data monitoring committee
MRI: magnetic resonance imaging
NANO: Neurological Assessment in Neuro-Oncology
ORR: objective response rate
OS: overall survival
PFS: progression-free survival
rMG: recurrent malignant gliomas
SMES: specific-mode electroacupuncture stimulation
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
WHO: World Health Organization

XSL•FO
RenderX

Edited by J Sarvestan; submitted 22.Sep.2025; peer-reviewed by B Li, R Akhavan-Sigari; accepted 31.Dec.2025; published 29.Jan.2026.

Please cite as:
Jia Z, Jiang T, Zhang Y, Chen Q, Di Z, Yuan Q, Qian K, Gan L, Ma C, Lin X
Clinical Efficacy and Safety Assessment of Specific-Mode Electroacupuncture Stimulation Combined With Paclitaxel for Recurrent Malignant Gliomas: Study Protocol for a Single-Arm Trial
JMIR Res Protoc 2026;15:e84593
URL: https://www.researchprotocols.org/2026/1/e84593
doi:10.2196/84593

© Zhaoxing Jia, Tianxiang Jiang, Yiqing Zhang, Qianyue Chen, Zhong Di, Qi Yuan, Kecheng Qian, Lin Gan, Congcong Ma, Xianming Lin. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 29.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Jiedu Xiaozhen Granules for Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor–Mediated Skin Toxicity: Protocol for a Randomized Controlled Trial

Shoujiang Hao[1*], BMed; Xiaoli Li[2*], MMed; Shulan Hao[2], MMed; Xiaoying Zhang[2], MMed; Xiaojun Qi[1], MMed; Gang Jin[3], MMed; Fangfang Shen[4*], MMed; Likun Liu[2*], MMed

[1]Shanxi University of Traditional Chinese Medicine, 89 Section 1 Jinci Road, Taiyuan, China

[2]Shanxi Province Hospital of Traditional Chinese Medicine, 46 Bingzhou West Street, Yingze District, Taiyuan, Shanxi, China

[3]The Second Affiliated Hospital of Shanxi Medical University, Taiyuan, China

[4]Shanxi Provincial Cancer Hospital, Taiyuan, China

[*]these authors contributed equally

Corresponding Author:
Likun Liu, MMed
Shanxi Province Hospital of Traditional Chinese Medicine, 46 Bingzhou West Street, Yingze District, Taiyuan, Shanxi, China

Abstract

Background: Epidermal growth factor receptor tyrosine kinase inhibitors (EGFR-TKIs) are widely used in the treatment of non–small cell lung cancer due to their precision, efficiency, and ease of use. However, skin rashes induced by EGFR-TKIs are the most common and earliest form of skin toxicity, often affecting the quality of life and treatment compliance of patients and leading to early discontinuation of therapy. These skin reactions may even impact cancer outcomes. In clinical practice, traditional Chinese medicine detoxification granules have shown effectiveness in relieving skin discomforts such as itching, pain, and burning caused by EGFR-TKI therapy. A prior single-arm trial investigating the treatment of targeted drug–induced rashes demonstrated a sustained improvement in rash symptoms with an effectiveness rate of 80.77% and was well tolerated by patients.

Objective: As an exploratory clinical study, this randomized controlled trial will preliminarily evaluate the potential efficacy and safety of jiedu xiaozhen (JDXZ) granules in managing EGFR-TKI–related skin toxicities.

Methods: This randomized controlled trial will be conducted at Shanxi Provincial Hospital of Traditional Chinese Medicine. A total of 94 patients with confirmed epidermal growth factor receptor gene–mutated non–small cell lung cancer who developed rashes after EGFR-TKI treatment will be enrolled. Patients will be randomly assigned to either a JDXZ traditional Chinese medicine group (group J) or a urea ointment group (group U). The primary outcome will be the severity of the rash as assessed using the National Cancer Institute Common Terminology Criteria for Adverse Events grading. Secondary outcomes will include the WoMo (Wollenberg and Moosmann) score, numerical rating scale, Dermatology Life Quality Index scale, European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 score, median progression-free survival, and changes in the levels of fibroblast growth factor 7 and hepatocyte growth factor in the blood. Adverse reactions will be recorded throughout the study. Data will be analyzed using SPSS.

Results: The clinical trial registration was completed in October 2024. This study is currently underway. As of December 1, 2025, a total of 81 eligible participants had been enrolled, all of whom were assigned to groups following the randomization principle. Among them, 42 participants were allocated to the JDXZ group (with an additional 2 participants pending enrollment), and 39 to the control group. Based on the current progress, the estimated trial completion date has been extended to January 31, 2026.

Conclusions: The results of this study may help develop an effective treatment for EGFR-TKI–mediated rashes. The findings will be published in academic journals upon the completion of the trial.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400086657; https://www.chictr.org.cn/showprojEN.html?proj=234631

International Registered Report Identifier (IRRID): DERR1-10.2196/79579

(JMIR Res Protoc 2026;15:e79579) doi:10.2196/79579

KEYWORDS

traditional Chinese medicine; non–small cell lung cancer; targeted drugs; jiedu xiaozhen granules; skin toxicity

XSL•FO
RenderX

Introduction

Patients with non–small cell lung cancer (NSCLC) who harbor epidermal growth factor receptor (*EGFR*) mutations represent a unique subgroup [1]. Women, nonsmokers, and patients with adenocarcinoma are more likely to carry *EGFR* mutations, with an incidence of 30% to 35% in East Asian populations [2]. In the Chinese NSCLC population, the prevalence of *EGFR* mutations is 28.2%, rising to 50.2% in those with lung adenocarcinoma [3]. *EGFR* tyrosine kinase inhibitors (EGFR-TKIs) selectively block the *EGFR* signaling pathway, thereby improving survival in patients with NSCLC carrying *EGFR* mutations [4]. It is categorically endorsed as a grade IA recommendation for the postoperative, locally advanced, and advanced first-line treatment of EGFR-positive NSCLC [5]. However, as *EGFR* is also expressed in skin keratinocytes, sebaceous glands, hair follicles, and perinuclear tissues [6], inhibiting *EGFR* activity can trigger a cascade of intracellular signaling pathways. The integrity of the human epidermis is compromised, the proliferation of epidermal cells is perturbed, the sebaceous gland lipid secretion is disrupted, and the skin's hydration retention function is impaired, culminating in the manifestation of adverse cutaneous reactions [7].

According to phase 3 clinical trials approved by the US Food and Drug Administration, the incidence of skin rash associated with gefitinib use was 66% in patients from East Asia, whereas 3.1% experienced grade 3 rashes. Erlotinib use resulted in a 73% rash incidence, with 32% being grade 3; afatinib use led to rashes in 80.8% of patients, with 14.6% experiencing grade 3 rashes. In China, icotinib use caused rashes in 15.5% of patients, with 14.9% experiencing grade 3 rashes [8]. Results from the FLAURA2 trial found that the incidence of grade 3 adverse effects was higher in the combination therapy group than in the monotherapy group (ranging from 11% to 54%), with serious adverse effects increasing from 5% to 19%. Osimertinib, another EGFR-TKI, was associated with skin toxicity in approximately 24% of patients [9].

The most common and earliest cutaneous side effects of anti-*EGFR* therapies are papules and pustules distributed in seborrheic areas such as the face, scalp, and upper trunk, sometimes extending to the limbs, palms, and soles. These reactions typically occur within 1 to 2 weeks of starting targeted therapy and are dose dependent. It is noteworthy that this issue is encountered in a minimum of 75% of the patient population [10]. The severity of the rash is often positively correlated with the response to treatment of the patients [11]. Approximately 35% of patients initially manifest dry, pruritic dermatitis affecting the extremities [12,13]. The onset of sebum deficiency in eczema is typically concurrent with the exacerbation of the condition, culminating in a pustular rash secondary to either *Staphylococcus aureus* or herpes simplex infection [14], which may or may not be accompanied by hyperpigmentation and telangiectasia [15]. Unlike acne vulgaris [16], these skin reactions are characterized by erythematous papules and pustules without acne or cysts. Other associated symptoms include dry skin, burning sensations, itching, pain, and changes to the hair and nails (eg, paronychia) [17]. To manage these side effects, patients should avoid contact with irritants, particularly in areas prone to rashes. Daily skin care should include the use of moisturizers and sun protection and the avoidance of extreme temperatures. Wearing comfortable clothing and shoes and using nonirritating bath products is also recommended [18]. It is imperative to mitigate the adverse cutaneous reactions that are elicited.

At present, the assessment of targeted drug–induced cutaneous toxicities is conducted in accordance with the 2017 National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE). Various guidelines, including those from the Multinational Association of Supportive Care in Cancer skin toxicity study group, the National Comprehensive Cancer Network, and the European Society for Medical Oncology, offer different expert opinions and consensuses on managing skin toxicity [11,19-21]. However, these guidelines lack comprehensive, evidence-based recommendations [8,10]. As patients often experience dry, itchy skin that compromises their skin barrier, most treatment strategies focus on symptom relief, and there are discrepancies in clinical practice regarding medication use [8,10].

The dermatological adverse effects associated with *EGFR* inhibitors are transient and tend to diminish with extended treatment duration. Resolution of the rash is typically complete upon cessation of therapy, and it is seldom constrained by the dosage or duration of treatment. Xerosis and pruritus are commonly attributable to a compromised skin barrier function. The predominant therapeutic modalities used for the management of grade 1 or 2 cutaneous toxicity encompass topical treatments such as 2.5% hydrocortisone cream, topical antibiotics (including erythromycin cream, clindamycin emulsion, metronidazole cream, and tacrolimus cream), and moisturizers (such as urea cream and fusidic acid cream). For more severe cases, oral medications may be considered, such as doxycycline (100 mg twice daily) or minocycline (50 - 100 mg twice daily). Oral antihistamines, including levocetirizine (5 mg daily), desloratadine (5 mg daily), diphenhydramine (25 - 50 mg 3 times daily), hydroxyzine (25 mg 3 times daily), or fexofenadine (60 mg 3 times daily), can help manage pruritus. For grade 3 rashes, oral tretinoic acid (0.3 - 0.5 mg per kilogram) or prednisone (0.5 mg per kilogram) may be prescribed. Gamma-aminobutyric acid receptor agonists such as gabapentin (300 mg every 8 hours) or pregabalin (50 - 75 mg every 8 hours) can also be considered for symptom relief. In severe cases, the dosage of targeted drugs should be reduced according to clinical guidelines, and for grade 4 rashes, oral steroid use may be escalated, and targeted therapy should be discontinued as soon as possible. Additionally, other studies have explored alternative treatments. For example, a study by Nagase et al [22] found that rapid desensitization therapy was effective in treating severe rashes caused by molecularly targeted drugs in patients with NSCLC. Furthermore, vitamin K_1 has shown potential in preventing rashes associated with *EGFR* inhibitors, possibly reducing both the incidence and severity of acne-like rashes [18,21].

Retinoic acid drugs are commonly used in clinical practice, but higher doses can exacerbate skin and mucosal dryness. Additionally, due to their photosensitivity, patients need to avoid

XSL•FO
RenderX

radiation exposure as these factors can worsen rashes [23]. Doxycycline has been found to reduce the severity of rashes induced by EGFR-TKIs [18]. However, in clinical practice, doxycycline and minocycline are not frequently used for treatment; instead, they are primarily used in preventive studies [18,24-26]. The findings of Deplanque et al [25] also confirmed that, while doxycycline may not lower the incidence of erlotinib-induced folliculitis, it does reduce the severity of the rash. It is important to note the potential side effects of tetracycline antibiotics, which include anorexia, nausea, vomiting, diarrhea, esophagitis, and esophageal ulcers [27]. Systemic steroid hormones are usually not the first line of treatment for *EGFR* inhibitor–associated rashes due to their own side effects [19], such as endocrine disorders, osteoporosis, and weakened immunity, which can have serious consequences [28].

Due to the complex nature of infections and the standardized use of antibiotics, it is important to remain vigilant about antibiotic resistance. Glucocorticoids, while potent in treating skin toxicity, should only be considered when other treatment options are ineffective, making both antibiotic therapy and glucocorticoid therapy less ideal for routine clinical use. Patients receiving molecular-targeted therapies are often highly sensitive to irritants and allergens, complicating clinical treatment decisions. Most of the patients with lung cancer and *EGFR* mutations who experience cutaneous adverse reactions to EGFR-TKIs do not seek treatment due to a lack of awareness of or confidence in existing treatment options. Research reveals that EGFR-TKI skin toxicity stems from a compromised skin barrier linked to abnormal keratinocyte differentiation, apoptosis, and inflammatory infiltration [29]. To guarantee the rational deployment and therapeutic efficacy of clinical pharmacotherapy, it is advisable to use mild emollients, which incorporate active constituents such as urea, ceramide, plant sterols, and linolenic acid, among others, for the purpose of local hydration; moisturization; anti-inflammatory action; antipruritic effects; and amelioration of cutaneous symptoms such as dryness, itching, and pain [29]. Consequently, the identification of novel and well-tolerated therapeutic modalities to address drug-induced dermatotoxicity continues to pose a significant challenge for the oncological community.

Traditional Chinese medicine (TCM) active ingredients and preparations are increasingly used in skin disease treatment as modern pharmacology and clinical trials advance [30]. Jiedu xiaozhen (JDXZ) granules, a blend of *Astragalus membranaceus*, *Isatis indigotica Fort*, and *Concha margaritifera*, show promise in managing EGFR-TKI–induced skin toxicity. Author LL, with 40 years' experience in TCM, has shown JDXZ granules' efficacy and safety for treating EGFR-TKI–related skin toxicity. The oral JDXZ granules in their single-arm trial [31] effectively alleviated systemic symptoms, sustaining rash improvement and reducing posttargeting recurrence with minimal side effects. In clinical observation, targeted drug–related skin toxicity showed sustained improvement with an effectiveness rate of 80.77% and was well tolerated by patients. This exploratory randomized controlled trial (RCT) will preliminarily evaluate the efficacy and safety of JDXZ granules to explore a potential management strategy for skin toxicity induced by EGFR-TKIs.

Methods

Study Design

This study is a multicenter, prospective, randomized controlled clinical trial in which participants will receive either enhanced TCM or positive control drugs. The trial will involve 94 patients with a confirmed diagnosis of *EGFR* gene mutation–positive NSCLC who developed skin rashes after using EGFR-TKIs. The participants will be divided into 2 groups: the treatment group (receiving oral JDXZ granules) and the control group (receiving topical urea ointment). Effectiveness will be assessed by comparing baseline conditions with the results at 2 weeks after treatment. Each patient will undergo 4 scheduled visits: at baseline and on days 3, 7, and 14 (Figure 1).

XSL•FO
RenderX

Figure 1. Clinical trial flowchart. CRF: case report form; DLQI: Dermatology Life Quality Index; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30; JDXZ: jiedu xiaozhen; laboratory examination: changes in blood levels of factors such as fibroblast growth factor 7 and hepatocyte growth factor; NCI-CTCAE: National Cancer Institute Common Terminology Criteria for Adverse Events; NRS: numerical rating scale; PFS: progression-free survival (time during which patients remain free from disease progression; follow-up every 3 months); TCM: traditional Chinese medicine; WoMo (Wollenberg and Moosmann) score: ranges from 0 to 100 and is calculated using the formula final score=1/4A + 1/4B + 10/3C).

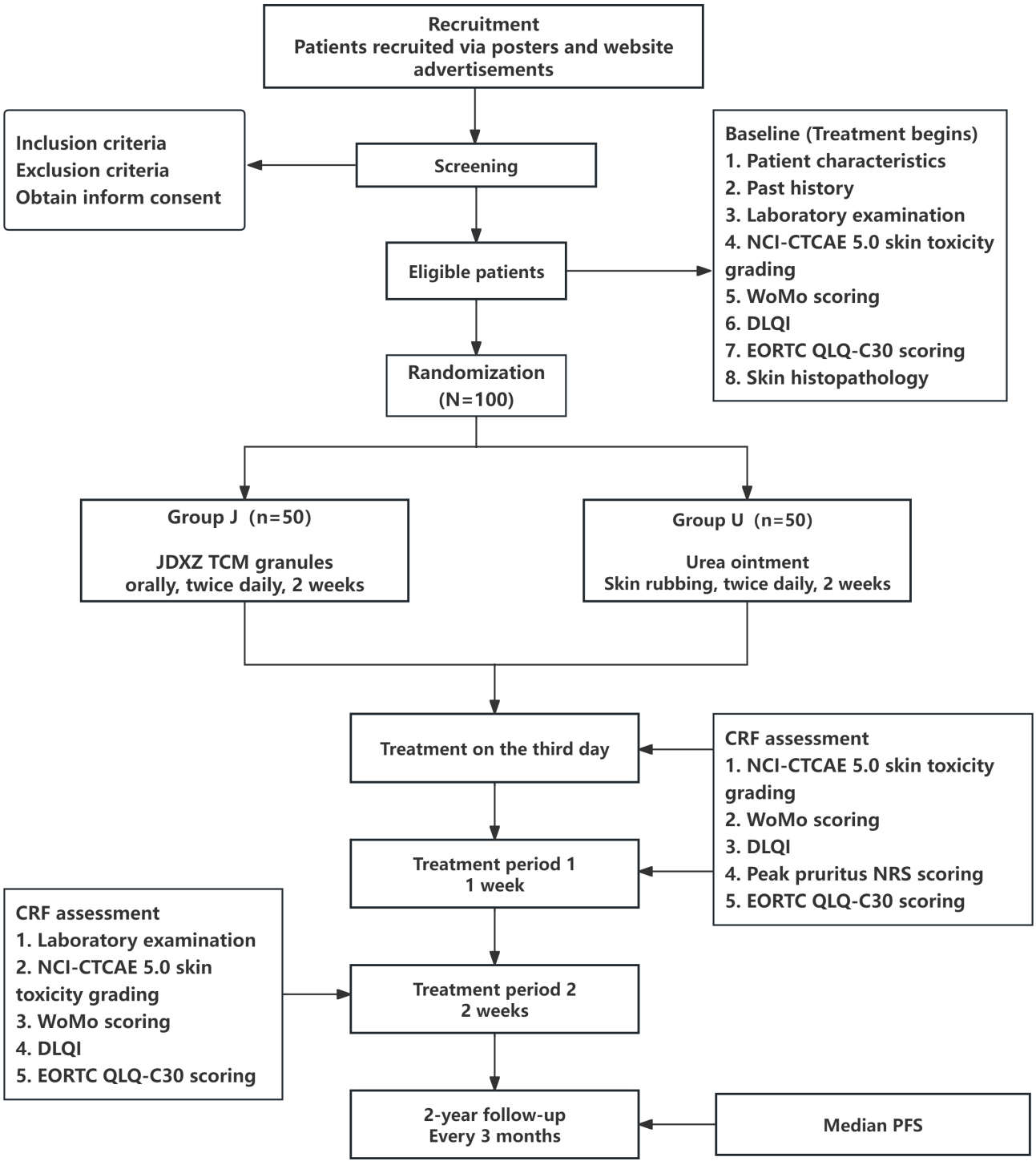

Study Setting

Participants will be recruited from 3 centers: Shanxi Provincial Hospital of Traditional Chinese Medicine, Shanxi Cancer Hospital, and the Second Affiliated Hospital of Shanxi Medical University. Adequate sample sizes will be obtained through various hospital departments and online platforms. Inclusion criteria will be verified, ensuring that participants have relatively complete medical records and have signed the informed consent form.

Participants

Recruitment Strategy

Patients will be enrolled from 3 clinical centers, most of which have been designated as affiliated hospitals of the Lung Cancer Mutation Consortium. Recruitment advertisements will be

XSL•FO
RenderX

posted in hospital bulletins, social media, and medical forums. The enrollment procedure will be streamlined to facilitate expeditious and hassle-free completion by the patients, thereby mitigating potential obstacles to their participation.

Eligibility Criteria

Inclusion Criteria

Inclusion criteria are as follows: (1) patients diagnosed with NSCLC carrying *EGFR* gene mutations who have developed skin rashes following EGFR-TKI treatment, with rash severity graded as 1 to 3 according to the NCI-CTCAE 5.0 standards; (2) patients aged ≥18 years, with no gender restrictions, and who have a Karnofsky Performance Status score of at least 60; (3) patients with adequate hematological, hepatic, renal, and metabolic function as indicated by specific blood count and organ function parameters (white blood cells of $>3 \times 10^9$ per liter, absolute neutrophil count of $\geq 1.5 \times 10^9$ per liter, platelet count of $\geq 100 \times 10^9$ per liter, hemoglobin of >9 g/dL, serum creatinine of ≤1.5 times the upper limit of normal [ULN], bilirubin of ≤1.5 times the ULN, aspartate aminotransferase or alanine transaminase of ≤2.5 times the ULN, if liver metastases are present, AST/ALT ≤ 5 × ULN, and serum electrolyte within normal range); and (4) patients who are willing to receive the study treatment and sign the informed consent form with good compliance.

Exclusion Criteria

Exclusion criteria are as follows: (1) impaired visceral function; (2) a history of significant cardiovascular diseases (including myocardial infarction, unstable angina pectoris, symptomatic congestive heart failure, and severe uncontrolled arrhythmia) within the previous year, interstitial lung disease (such as pneumonia or pulmonary fibrosis), or evidence of interstitial lung disease on a baseline chest computerized tomography scan; (3) concurrent antitumor treatments; (4) rashes unrelated to EGFR-TKI drugs or recent rash-related treatments; (5) participation in other clinical trials within 4 weeks prior; (6) diagnosis of dialectical spleen or stomach deficiencies or cold syndromes as per TCM; (7) known allergies to trial medications; and (8) pregnant or lactating women and those planning to become pregnant within 6 months after the end of treatment.

Withdrawal Criteria

Patients will be withdrawn from the study if they experience serious adverse events, such as nausea, vomiting from oral TCM, or liver and kidney insufficiency, making further observation unsuitable. Additionally, patients who fail to complete the treatment and observation period outlined in the protocol for any reason will be considered for withdrawal. Patients unwilling to continue treatment and requesting to exit the trial will also be withdrawn. Finally, any disease progression that affects or interferes with the study outcomes will result in the withdrawal of the patient from the trial.

Randomization and Blinding

The randomization sequence will be generated by the data manager. The patients will be randomly assigned to either the treatment group or the control group in a 1:1 ratio. Blinding methods will not be used in this clinical trial. Although blinding of participants is not feasible due to the nature of the intervention, the following blinding strategies will be implemented in the study design to minimize potential bias: health care providers administering the treatment, outcome assessors, and data analysts will all be kept blinded to group allocation. Specifically, the treatment providers responsible for assigning the interventions will not participate in subsequent efficacy or safety evaluations. During follow-up visits, clinical assessors will perform dermatological lesion observation and grading without knowledge of the participants' group assignments. For the final statistical analysis, analysts will process data using anonymized codes, and blinding will only be lifted after the statistical analysis is completed. This approach was designed to ensure objectivity and accuracy in outcome evaluation and data interpretation.

Interventions

In the treatment group, after EGFR-TKI treatment, patients will receive daily oral TCM-based granule treatment. The prescription consists of *A membranaceus* (30 g), *I indigotica Fort* (15 g), and *C margaritifera* (15 g). The granules will be taken twice daily, in the morning and evening, following meals for 7 days per treatment course, with 2 consecutive courses administered. In the control group, patients will apply the urea ointment to the affected areas daily, 1 to 3 times per day, with each course lasting 7 days, also for 2 consecutive courses. All medications were supplied by Shanxi Provincial Hospital of Traditional Chinese Medicine. Investigators will assess patients on days 0, 3, 7, and 14 (Table 1). During the observation period, all patients will avoid contact with irritants, particularly in areas prone to rash, and will use sunscreen daily to reduce the incidence of skin rashes.

XSL•FO
RenderX

Table . Schedule of the jiedu xiaozhen granule clinical trial.

	Visit 0 (day 0; ±2 days)	Visit 1 (day 3; ±2 days)	Visit 2 (day 7; ±2 days)	Visit 3 (day 14; ±2 days)	Visit 4 (3, 6, 9, 12, 15, 18, 21, and 24 months)
Signature of the informed consent form	✓				
Inclusion and exclusion criteria for verification	✓				
Randomization grouping	✓				
Release of experimental drugs	✓				
Recovery of the experimental drugs				✓	
Basic data collection					
Demographic data	✓				
Personal history and allergy history	✓				
Study disease–related history	✓				
Evaluation of the condition and its effectiveness[a]	✓	✓	✓	✓	
History of other diseases	✓				
Concomitant medications or treatment	✓	✓	✓	✓	
Security observations					
Liver and kidney function examination	✓			✓	
Untoward effect		✓	✓	✓	
Efficacy observation					
Clinical efficacy rate (NCI-CTCAE[b] version 5.0)	✓	✓	✓	✓	
WoMo score	✓	✓	✓	✓	
Peak pruritus NRS[c]	✓	✓	✓	✓	
DLQI[d]	✓	✓	✓	✓	
EORTC QLQ-C30[e] scale (quality of life measurement)	✓	✓	✓	✓	
Median PFS[f]					✓
Laboratory examination (FGF7[g] and HGF[h] in human blood)	✓	✓		✓	
Skin histopathological examination	✓				
Evaluation of medication compliance		✓	✓	✓	
Adverse events and complications		✓	✓	✓	
Serious adverse events		✓	✓	✓	

XSL•FO
RenderX

[a]A systematic and comprehensive evaluation of patients with non–small cell lung cancer was conducted in accordance with the Response Evaluation Criteria in Solid Tumors (RECIST 1.1), encompassing changes in pulmonary lesions and metastatic foci, as well as indicators such as tumor marker.

[b]NCI-CTCAE: National Cancer Institute Common Terminology Criteria for Adverse Events.

[c]NRS: numerical rating scale.

[d]DLQI: Dermatology Life Quality Index.

[e]EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30.

[f]PFS: progression-free survival.

[g]FGF7: fibroblast growth factor 7.

[h]HGF: hepatocyte growth factor.

To uphold the integrity of the study, participants are permitted to engage in fundamental skin care regimens throughout the trial and will maintain diligent sun protection measures and refrain from indulging in excessive cleansing or the application of abrasive cosmetics on a daily basis. Nevertheless, it is imperative to abstain from the ingestion of any other pharmacological agents or the implementation of interventions that could potentially disrupt the therapeutic process for the rash, encompassing, though not restricted to, hormonal medications and immunosuppressants. Should any significant cutaneous discomfort or additional complications arise, it is essential to promptly notify the attending physician for a timely evaluation and modification of the treatment plan if necessary.

Assessments

Primary Outcomes

In conformity with the NCI-CTCAE version 5.0 and the efficacy assessment criteria in alignment with the *Guiding Principles for Clinical Research on New Chinese Medicine Drugs* [32], the criterion for improvement was defined as a posttreatment rash grade of 0, denoting complete recovery. A ≥2-grade improvement in the rash classification was defined as a marked effect. A single-grade improvement in the rash classification was deemed effective. Should these criteria not be satisfied and the rash either remain unaltered in severity or worsen by more than 1 grade, the treatment will be regarded as ineffective. The following formula will be used: clinical efficacy rate=(cured + significantly improved + improved cases)/total cases × 100% .

Secondary Outcomes

The secondary outcomes include the WoMo (Wollenberg and Moosmann) score, which ranges from 0 to 100 and is calculated using the formula final score= 1/4A + 1/4B + 10/3C. Acnelike rash scores of 0 to 20 are considered mild, scores of 21 to 40 are considered moderate, and scores of 41 to 100 are considered severe (Multimedia Appendix 1). Other outcomes include the peak pruritus numerical rating scale, which measures the intensity of pruritus; the dermatology Life Quality Index, which assesses the impact of skin disease on quality of life; the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 scale (quality of life measurement), which evaluates overall quality of life in patients with cancer; median progression-free survival, which represents the time during which patients remain free from disease progression (follow-up every 3 months); a laboratory examination to measure changes in blood levels of factors such as fibroblast growth factor 7 and hepatocyte growth factor, and a skin histopathological examination to evaluate skin tissue changes after the treatment.

Adverse Events

In conformity with the criteria delineated by the Ministry of Health of China's Adverse Drug Reaction Monitoring Center, as documented in the *Guiding Principles for Clinical Research of New Chinese Medicine Drugs*, the potential relationship between adverse events and the use of the investigational drug will be appraised using a 5-tiered categorization scheme encompassing "certain," "probable," "possible," "remote," and "unrelated." The specific evaluation metrics are meticulously outlined in the adverse event table, which is an integral component of the study's records (Multimedia Appendix 2).

Researchers must prompt patients to honestly report medication effects, steer clear of suggestive questions, and document any adverse reactions on the adverse event form. Follow-up, including detailed records and laboratory test normalization, should continue until symptoms resolve via hospital visits, outpatient checks, or other methods based on reaction severity (Multimedia Appendix 2).

Upon detecting adverse reactions, the researchers will diagnose, treat, and decide whether to halt observation based on the condition. For serious events, the study unit must act promptly to ensure participant safety. Researchers will submit the serious adverse event report form to the oncology department, the ethics committee, and all participating units of the hospital within 24 hours, signing and dating it.

Data Collection and Management

Informed consent will be enhanced by detailing the study's purpose, procedure, potential discomfort, and risks, ensuring voluntary participation. Effective physician-patient communication will be fostered, promptly addressing patient needs and resolving treatment issues to boost compliance. A follow-up plan will be developed to minimize participant burden, and phones and SMS text messages, among other tools, will be used to increase follow-up rates.

Data will be collected and recorded in a case report form by trained researchers. Baseline assessment data will be collected on the sociodemographics of the patients (including gender, age, educational level, religion, marital status, occupation, income, economic burden of illness, caregiver during illness, and smoking and drinking history), as well as disease duration and any previous treatment with TCM. Medical records of clinical trial participants, both inpatient and outpatient, should be stored at the hospital to create an electronic case database.

Researchers must document medical records concurrently with treatment to ensure timeliness, completeness, and accuracy. Any corrections should be underlined and dated by the

XSL•FO
RenderX

researchers without altering the original entry. The CRF Questionnaire and Adverse Event Recording Survey, completed by both the patient and physician, adheres to objectivity and sincerity. Objectivity pertains to neutral instrument design and data authenticity, mitigating researcher bias. Respondent sincerity reflects truthful participation, minimizing social desirability bias. Their synergy ensures data reliability and study validity. Value ranges and logic will be regularly verified. For uncertainties, the question list will be completed, and researchers will address the listed queries. Main concerns include the standardization of the components of Chinese herbal preparations and their quality; potential interference from concomitant medications; and possible subjectivity in skin toxicity assessment indicators. Upon finishing the data verification report, the database will be secured. Quality control documents will be kept, including records of data consistency checks, value range and logic verifications, blind audits, and researcher communications.

Statistical Analysis

Data will be analyzed using SPSS Statistics for Windows (version 26.0; IBM Corp). Continuous data following a normal distribution will be expressed as means and SDs and compared within groups using a paired 2-tailed *t* test or between groups using an independent-sample 2-tailed *t* test. Data will be presented as frequencies and percentages, with chi-square tests for group comparisons and rank sum tests for ranked data. For continuous data that are not normally distributed, between-group comparisons will be performed using the Mann-Whitney *U* test, with data described as medians and IQRs. For measurements involving multiple time points, nonparametric methods such as the Friedman test will be used to ensure robust and reliable findings. A *P* value of <.05 will be considered statistically significant.

Sample Size

On the basis of a small-sample clinical trial, the effectiveness rate of the treatment group was 80.77% [31]. Additionally, a previous randomized controlled double-blind clinical study had a response rate of 45% [33]. Using the PASS tool (version 21), with a significance level of .05 and a power of 0.9, the total sample size was calculated as 78 participants. Accounting for a 20% dropout rate, the required sample size was determined to be 94 patients, with 47 in each group.

Ethical Considerations

This RCT has received joint ethics approval from the institutional review boards of Shanxi Provincial Hospital of Traditional Chinese Medicine, Shanxi Cancer Hospital, and the Second Affiliated Hospital of Shanxi Medical University (approval 2024 KY-07026 and 2024KYY1-07014). Participants who sign the informed consent form will be enrolled in the study, with the right to withdraw at any point during the trial. Finally, the results of this RCT will be published in a relevant academic journal on completion of the trial irrespective of whether the RCT yields negative or positive outcomes. This study has been registered with the Chinese Clinical Trial Registry (ChiCTR2400086657).

Trial Monitoring

Paper details will be managed by a designated individual and stored securely. Access to the e–medical case system will be restricted to authorized users. Before the commencement of the project, the researchers underwent rigorous training on the experimental protocol, and the uniformity of the quantitative metrics for symptoms and signs was meticulously verified. Any modifications to this protocol are subject to the prior approval of the ethics committee at Shanxi Provincial Hospital of Traditional Chinese Medicine to ensure the trial's probity and adherence to ethical standards.

Results

This study is currently underway. As of December 1, 2025, a total of 81 eligible participants had been enrolled, all of whom were assigned to groups following the randomization principle. Among them, 42 participants were allocated to the JDXZ group (with an additional 2 participants pending enrollment), and 39 to the control group. Based on the current progress, the estimated trial completion date has been extended to January 31, 2026.

Discussion

Currently, the investigation into skin-related toxicities caused by EGFR-TKIs remains in the exploratory phase, with its findings yet to be fully substantiated, thereby necessitating further comprehensive research. EGFR-TKIs have the potential to elicit the aggregation of inflammatory cells within the context of normal skin. A study by Satoh et al [34] showed EGFR-TKIs' synergy with skin *Bacillus* acnes to boost interleukin (IL)-36γ in keratinocytes. EGFR-TKIs suppressed *EGFR* signaling and upregulated KLF4, whereas *Bacillus* acnes triggered the activation of the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB). KLF4 and NF-κB collaborate to upregulate IL-36γ promoter activity, thereby inducing neutrophil recruitment via IL-8 and worsening skin inflammation. EGFR-TKIs not only suppress the proliferative activity of basal keratinocytes but also facilitate their transition toward terminal differentiation, ultimately leading to xerosis of the skin. Erlotinib, as demonstrated by Ikarashi et al [35], markedly lowered aquaporin-3 expression in keratinocytes via the extracellular signal–regulated kinase pathway. Aquaporin-3 is crucial for skin hydration; reduced expression may impair this function. EGFR-TKIs have the potential to disrupt the intimate associations between basal keratinocytes, consequently impairing the epidermal barrier's functionality. Gefitinib impairs the epidermal barrier by reducing claudin-1 and claudin-4 expression and increasing claudin-2 levels, as discovered by Fang et al [36].

Astragaloside IV, the main active compound of Astragalus, can inhibit the production of pro-inflammatory cytokines and the expression of TLR4 and its downstream signaling molecules, NF-κB, iNOS (inducible nitric oxide synthase), and COX-2 (cyclooxygenase-2) proteins [30]. *Astragalus* polysacharin, a monomeric *Astragalus* extract, curbs inflammatory cytokine (tumor necrosis factor, IL-6, and monocyte chemoattractant protein 1) production in lipopolysaccharide-activated RAW264.7 macrophages through NF-κB and mitogen-activated protein

XSL•FO
RenderX

kinase (extracellular signal–regulated kinase and c-Jun N-terminal kinase) pathways [37]. Compounds 4 (3H) -quinazolinone (I) and 2,4 (1H, 3H) -quinazolindione (II) from purified *I indigotica Fort* effectively eliminate *S aureus*, *Escherichia coli*, *Bacillus subtilis*, and *Salmonella* [38]. Min et al [39] found that *I indigotica Fort* extract eases atopic dermatitis–like inflammation by suppressing NF-κB pathway–induced cytokine and chemokine production. Chymotrypsin severs desmosome links in skin keratinocytes, speeding cell shedding and boosting epithelial metabolism [40]. *C margaritifera*'s bioactives boost chymotrypsin, stimulate skin cell growth, and enhance collagen synthesis for skin regeneration and elasticity [40]. Reports confirm the synergistic effects of *A membranaceus*, *I indigotica Fort*, and *C margaritifera*, deeming JDXZ granules safe and effective for inflammation, antibacterial protection, and skin protection.

In the treatment of NSCLC with EGFR-TKIs, the presence of a rash has been shown to significantly predict the effectiveness of the treatment, particularly in patients with unknown *EGFR* mutation status [17]. As a result, clinicians must closely monitor skin toxicity throughout the entire course of treatment and manage it appropriately to prevent treatment discontinuation or dose reduction of EGFR-TKIs. Building on the clinical promise of JDXZ granules for EGFR-TKI-induced skin toxicity demonstrated in this RCT, a follow-up RCT is warranted to rigorously assess the treatment's efficacy and safety, with the goals of reducing skin toxicity and enhancing patient quality of life. In this study, the severity of the rash will be graded according to the NCI-CTCAE 5.0 criteria. Assessment tools such as the itch numerical rating scale, the Dermatology Life Quality Index, and the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 will be used. Additionally, serum levels of fibroblast growth factor 7 [41] and hepatocyte growth factor [42] will be incorporated as biomarkers, with WoMo [43,44] scores and dermatopathological findings serving as secondary end points. This approach provides a more comprehensive and updated perspective on EGFR-TKI–associated cutaneous toxicities.

However, this nonblinded study has limitations, first, in its level of evidence, with potential issues such as unevenly designed control groups and subjective bias in patient-reported outcomes (eg, itching and pain). Furthermore, the collection of long-term medication safety data remains insufficient. Should this study show promising signs, a more rigorously designed multicenter, large-sample, randomized, double-blind, double-dummy, active- or placebo-controlled clinical trial will be necessary for further validation. Such a step is critical to elevate the findings to a higher evidence level and support their potential inclusion in clinical guidelines. Future research should also focus on identifying predominant responders to JDXZ granules based on specific TCM syndrome types or biomarkers, enabling more precise application. Additionally, exploring synergistic effects with conventional Western treatments (such as topical corticosteroids or oral antibiotics) may help reduce the dosage and duration of antibiotic or corticosteroid use, thereby minimizing potential side effects.

Funding

This study was supported by the Construction Project of the China National Innovation Center for Traditional Chinese Medicine (grant 202203) and the Medical Science and Technology Innovation Team Building Task of Shanxi Province (grant 2020TD04). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

Conceptualization: XL, LL, GJ, and FS
Data curation: Shoujiang Hao and XQ
Methodology: XL and XZ
Supervision: LL, Shulan Hao, and XL
Writing—original draft: XL and Shoujiang Hao
Writing—review and editing: XL and Shoujiang Hao

Conflicts of Interest

None declared.

Multimedia Appendix 1
The WoMo (Wollenberg and Moosmann) rating scale.
[DOCX File, 17 KB - resprot_v15i1e79579_app1.docx]

Multimedia Appendix 2
Instructions for adverse events.
[DOCX File, 13 KB - resprot_v15i1e79579_app2.docx]

References

XSL•FO
RenderX

1. Liu Y. Small cell lung cancer transformation from EGFR-mutated lung adenocarcinoma: a case report and literatures review. Cancer Biol Ther 2018 Jun 3;19(6):445-449. [doi: 10.1080/15384047.2018.1435222] [Medline: 29461911]
2. Zhang YL, Yuan JQ, Wang KF, et al. The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. Oncotarget 2016 Nov 29;7(48):78985-78993. [doi: 10.18632/oncotarget.12587] [Medline: 27738317]
3. Chinese Society of Clinical Oncology (CSCO) Non-small Cell Lung Cancer Committee, Anti-cancer Drug Safety Management Committee (ASMC). Consensus on application of third-generation EGFR-TKI in EGFR mutated NSCLC (2022 version). Zhongguo Fei Ai Za Zhi 2022 Sep 20;25(9):627-641. [doi: 10.3779/j.issn.1009-3419.2022.101.47] [Medline: 36172727]
4. Wu JY, Yu CJ, Chang YC, Yang CH, Shih JY, Yang PC. Effectiveness of tyrosine kinase inhibitors on "uncommon" epidermal growth factor receptor mutations of unknown clinical significance in non-small cell lung cancer. Clin Cancer Res 2011 Jun 1;17(11):3812-3821. [doi: 10.1158/1078-0432.CCR-10-3408] [Medline: 21531810]
5. Hsu WH, Yang JC, Mok TS, Loong HH. Overview of current systemic management of EGFR-mutant NSCLC. Ann Oncol 2018 Jan 1;29(suppl_1):i3-i9. [doi: 10.1093/annonc/mdx702] [Medline: 29462253]
6. Peng Y, Li Q, Zhang J, et al. Update review of skin adverse events during treatment of lung cancer and colorectal carcinoma with epidermal growth receptor factor inhibitors. Biosci Trends 2019 Jan 22;12(6):537-552. [doi: 10.5582/bst.2018.01246] [Medline: 30555112]
7. Xueqi T, Zhe W, Lijing J, et al. Status quo analysis of outcome indexes and evaluation tool in treatment of skin rash caused by EGFR-TKI with traditional Chinese medicine. World Sci Technol Mod Tradit Chin Med 2023;25(6):2147-2158.
8. Chinese Society of Lung Cancer, Chinese Anti-Cancer Association. EGFR-TKI ADR management Chinese expert consensus. Zhongguo Fei Ai Za Zhi 2019 Feb 20;22(2):57-81. [doi: 10.3779/j.issn.1009-3419.2019.02.01] [Medline: 30827323]
9. Planchard D, Jänne PA, Cheng Y, et al. Osimertinib with or without chemotherapy in EGFR-mutated advanced NSCLC. N Engl J Med 2023 Nov 23;389(21):1935-1948. [doi: 10.1056/NEJMoa2306434] [Medline: 37937763]
10. Macdonald JB, Macdonald B, Golitz LE, LoRusso P, Sekulic A. Cutaneous adverse effects of targeted therapies: part I: inhibitors of the cellular membrane. J Am Acad Dermatol 2015 Feb;72(2):203-218. [doi: 10.1016/j.jaad.2014.07.032] [Medline: 25592338]
11. Lacouture ME, Sibaud V, Gerber PA, et al. Prevention and management of dermatological toxicities related to anticancer agents: ESMO Clinical Practice Guidelines☆. Ann Oncol 2021 Feb;32(2):157-170. [doi: 10.1016/j.annonc.2020.11.005] [Medline: 33248228]
12. Saltz LB, Meropol NJ, Loehrer PJ Sr, Needle MN, Kopit J, Mayer RJ. Phase II trial of cetuximab in patients with refractory colorectal cancer that expresses the epidermal growth factor receptor. J Clin Oncol 2004 Apr 1;22(7):1201-1208. [doi: 10.1200/JCO.2004.10.182] [Medline: 14993230]
13. Shah NT, Kris MG, Pao W, et al. Practical management of patients with non-small-cell lung cancer treated with gefitinib. J Clin Oncol 2005 Jan 1;23(1):165-174. [doi: 10.1200/JCO.2005.04.057] [Medline: 15557594]
14. Galimont-Collen AF, Vos LE, Lavrijsen AP, Ouwerkerk J, Gelderblom H. Classification and management of skin, hair, nail and mucosal side-effects of epidermal growth factor receptor (EGFR) inhibitors. Eur J Cancer 2007 Mar;43(5):845-851. [doi: 10.1016/j.ejca.2006.11.016] [Medline: 17289377]
15. Chang GC, Yang TY, Chen KC, Yin MC, Wang RC, Lin YC. Complications of therapy in cancer patients: case 1. Paronychia and skin hyperpigmentation induced by gefitinib in advanced non-small-cell lung cancer. J Clin Oncol 2004 Nov 15;22(22):4646-4648. [doi: 10.1200/JCO.2004.02.168] [Medline: 15542815]
16. Guttman-Yassky E, Mita A, De Jonge M, et al. Characterisation of the cutaneous pathology in non-small cell lung cancer (NSCLC) patients treated with the EGFR tyrosine kinase inhibitor erlotinib. Eur J Cancer 2010 Jul;46(11):2010-2019. [doi: 10.1016/j.ejca.2010.04.028] [Medline: 20621734]
17. Petrelli F, Borgonovo K, Cabiddu M, Lonati V, Barni S. Relationship between skin rash and outcome in non-small-cell lung cancer patients treated with anti-EGFR tyrosine kinase inhibitors: a literature-based meta-analysis of 24 trials. Lung Cancer 2012 Oct;78(1):8-15. [doi: 10.1016/j.lungcan.2012.06.009] [Medline: 22795701]
18. Yuan C, Wang B. Acneiform eruption induced by molecularly targeted agents in antineoplastic therapy: a review. J Cosmet Dermatol 2023 Aug;22(8):2150-2157. [doi: 10.1111/jocd.15704] [Medline: 36924348]
19. Lacouture ME, Anadkat MJ, Bensadoun RJ, et al. Clinical practice guidelines for the prevention and treatment of EGFR inhibitor-associated dermatologic toxicities. Support Care Cancer 2011 Aug;19(8):1079-1095. [doi: 10.1007/s00520-011-1197-6] [Medline: 21630130]
20. Burtness B, Anadkat M, Basti S, et al. NCCN Task Force Report: management of dermatologic and other toxicities associated with EGFR inhibition in patients with cancer. J Natl Compr Canc Netw 2009 May;7 Suppl 1:S5-S21. [doi: 10.6004/jnccn.2009.0074] [Medline: 19470276]
21. Ocvirk J, Heeger S, McCloud P, Hofheinz RD. A review of the treatment options for skin rash induced by EGFR-targeted therapies: evidence from randomized clinical trials and a meta-analysis. Radiol Oncol 2013 May 21;47(2):166-175. [doi: 10.2478/raon-2013-0014] [Medline: 23801914]
22. Nagase M, Ohshima N, Kawashima M, et al. Usefulness of rapid desensitization therapy for severe rash caused by molecularly targeted drugs used in the treatment of non-small-cell lung cancer. Intern Med 2020 Dec 15;59(24):3201-3205. [doi: 10.2169/internalmedicine.5048-20] [Medline: 32788540]

XSL•FO
RenderX

23. Sah JF, Eckert RL, Chandraratna RA, Rorke EA. Retinoids suppress epidermal growth factor-associated cell proliferation by inhibiting epidermal growth factor receptor-dependent ERK1/2 activation. J Biol Chem 2002 Mar 22;277(12):9728-9735. [doi: 10.1074/jbc.M110897200] [Medline: 11788593]
24. Vaubel J, Livingstone E, Schadendorf D, Zimmer L. Retarded low‐dose doxycycline for EGFR or MEK inhibitor–induced papulopustular rash. J Eur Acad Dermatol Venereol 2014 Dec;28(12):1685-1689. [doi: 10.1111/jdv.12365] [Medline: 24422792]
25. Deplanque G, Gervais R, Vergnenegre A, et al. Doxycycline for prevention of erlotinib-induced rash in patients with non-small-cell lung cancer (NSCLC) after failure of first-line chemotherapy: a randomized, open-label trial. J Am Acad Dermatol 2016 Jun;74(6):1077-1085. [doi: 10.1016/j.jaad.2016.01.019] [Medline: 26946985]
26. Goulden V, Glass D, Cunliffe WJ. Safety of long-term high-dose minocycline in the treatment of acne. Br J Dermatol 1996 Apr;134(4):693-695. [doi: 10.1111/j.1365-2133.1996.tb06972.x] [Medline: 8733373]
27. Drugs and Lactation Database (LactMed®): National Institute of Child Health and Human Development; 2006. URL: https://www.ncbi.nlm.nih.gov/books/NBK501922/ [accessed 2026-01-12]
28. Oray M, Abu Samra K, Ebrahimiadib N, Meese H, Foster CS. Long-term side effects of glucocorticoids. Expert Opin Drug Saf 2016;15(4):457-465. [doi: 10.1517/14740338.2016.1140743] [Medline: 26789102]
29. Papoui E, Papastavrou E, Merkouris A, Charalambous A. The extent to which the last decade has yielded additional treatment options for EGFR-associated rash besides classic treatment with antibiotics and corticosteroids - a systematic review. Eur J Oncol Nurs 2021 Feb;50:101896. [doi: 10.1016/j.ejon.2021.101896] [Medline: 33493993]
30. Yang J, Guo J, Tang P, et al. Insights from traditional Chinese medicine for restoring skin barrier functions. Pharmaceuticals (Basel) 2024 Sep 5;17(9):1176. [doi: 10.3390/ph17091176] [Medline: 39338338]
31. Liu L, Li X, Hao S, et al. A TCM composition containing large green leaves and its application in the treatment of targeted drug-related skin rash: CN114788847B.x. Chinese Patent Literature. 2023 Apr 7. URL: http://epub.cnipa.gov.cn/patent/CN114788847A [accessed 2025-05-25]
32. General Principles for Clinical Research of New Traditional Chinese Medicine Drugs. National Medical Products Administration. URL: http://www.nmpa.gov.cn/xxgk/ggtg/qtggtg/2015110312187.html [accessed 2015-11-03]
33. Yu R, Jia L, Pei Y, et al. A randomized, controlled, double-blind clinical study of the treatment of hand and foot skin reactions induced by multi-kinet inhibitors of antitumor targeted drugs (LC09 granules) combined with urea ointment. J Trad Chin Med 2020. [doi: 10.13288/j.11-2166/r.2020.23.012]
34. Satoh TK, Mellett M, Meier-Schiesser B, et al. IL-36γ drives skin toxicity induced by EGFR/MEK inhibition and commensal Cutibacterium acnes. J Clin Invest 2020 Mar 2;130(3):1417-1430. [doi: 10.1172/JCI128678] [Medline: 31805013]
35. Ikarashi N, Kaneko M, Watanabe T, et al. Epidermal growth factor receptor tyrosine kinase inhibitor erlotinib induces dry skin via decreased in aquaporin-3 expression. Biomolecules 2020 Apr 3;10(4):545. [doi: 10.3390/biom10040545] [Medline: 32260143]
36. Fang H, Wang Y, Xu L, et al. EGFR inhibitor gefitinib regulates barrier function in human epidermal keratinocytes via the modulation of the expression of claudins. Int J Mol Med 2019 Mar;43(3):1522-1530. [doi: 10.3892/ijmm.2018.4046] [Medline: 30628660]
37. Chen G, Jiang N, Zheng J, et al. Structural characterization and anti-inflammatory activity of polysaccharides from Astragalus membranaceus. Int J Biol Macromol 2023 Jun 30;241:124386. [doi: 10.1016/j.ijbiomac.2023.124386]
38. Liao BL, Pan YJ, Zhang W, Pan LW. Four natural compounds separated from folium isatidis: crystal structures and antibacterial activity. Chem Biodivers 2018 Jul;15(7):e1800152. [doi: 10.1002/cbdv.201800152] [Medline: 29696796]
39. Min GY, Kim TI, Kim JH, Cho WK, Yang JH, Ma JY. Anti-atopic effect of isatidis folium water extract in TNF-α/IFN-γ-induced HaCaT cells and DNCB-induced atopic dermatitis mouse model. Molecules 2023 May 8;28(9):3960. [doi: 10.3390/molecules28093960] [Medline: 37175371]
40. Liu Y, Yan X. In vitro effects of scallop shell aqueous extract on proteases and fibroblasts. Chin J Exp Med 2013;15:190-193.
41. Hampton T. New insight on preventing EGFR inhibitor-induced adverse effects. JAMA 2020 Mar 3;323(9):814. [doi: 10.1001/jama.2020.0812] [Medline: 32125381]
42. Hichert V, Scholl C, Steffens M, et al. Predictive blood plasma biomarkers for EGFR inhibitor-induced skin rash. Oncotarget 2017 May 23;8(21):35193-35204. [doi: 10.18632/oncotarget.17060] [Medline: 28456787]
43. Wollenberg A, Moosmann N, Klein E, Katzer K. A tool for scoring of acneiform skin eruptions induced by EGF receptor inhibition. Exp Dermatol 2008 Sep;17(9):790-792. [doi: 10.1111/j.1600-0625.2008.00715.x] [Medline: 18503553]
44. Gaiser MR, Lorenzen S, Merx K, et al. Evaluation of EGFR inhibitor-mediated acneiform skin toxicity within the double-blind randomized EVITA trial: a thorough gender-specific analysis using the WoMo score. Cancer Med 2019 Aug;8(9):4169-4175. [doi: 10.1002/cam4.2132] [Medline: 31199595]

Abbreviations

EGFR: epidermal growth factor receptor
EGFR-TKI: epidermal growth factor receptor tyrosine kinase inhibitor
IL: interleukin

XSL•FO
RenderX

JDXZ: jiedu xiaozhen
NCI-CTCAE: National Cancer Institute Common Terminology Criteria for Adverse Events
NF-κb: nuclear factor kappa-light-chain-enhancer of activated B cells
NSCLC: non–small cell lung cancer
RCT: randomized controlled trial
TCM: traditional Chinese medicine
ULN: upper limit of normal
WoMo: Wollenberg and Moosmann

Edited by J Sarvestan; submitted 24.Jun.2025; peer-reviewed by C Lu; revised version received 29.Dec.2025; accepted 29.Dec.2025; published 22.Jan.2026.

Please cite as:
Hao S, Li X, Hao S, Zhang X, Qi X, Jin G, Shen F, Liu L
Jiedu Xiaozhen Granules for Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitor–Mediated Skin Toxicity: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e79579
URL: https://www.researchprotocols.org/2026/1/e79579
doi:10.2196/79579

© Shoujiang Hao, Xiaoli Li, Shulan Hao, Xiaoying Zhang, Xiaojun Xiaojun Qi, Gang Jin, Fangfang Shen, Likun Liu. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Normalization of Gastrointestinal Symptoms in Adults With Constipation With Daily Green Kiwifruit Consumption: Protocol for an Open-Label Intervention Study

Jasjot Maggo[1,2*], MMed, MPhil, PhD; Hwei Min Ng[1,2*], PhD; Simone Birgit Bayer[1,2], PhD; Catherine L Wall[1,2], PhD; Caroline L Hoad[3], PhD; Luca Marciani[3], PhD; Jane Mullaney[2,4,5], PhD; Diana Cabrera[2,6], PhD; Karl Fraser[2,4,5], PhD; Janine M Cooney[2,7], PhD; Catrin S Günther[2,7], PhD; Tania Trower[2,7], MSc; Jeffery Tang[2,8], PhD; Olivier Gasser[2,8], PhD; Amber Milan[2,9], PhD; Warren C McNabb[2,5], PhD; Robin Spiller[1], MD, PhD; Tamlin Conner[10], PhD; Chris Frampton[1], PhD; Meika Foster[2,5,11], PhD; Nicole C Roy[2,5,12†], PhD; Richard B Gearry[1,2,13], MBChB, PhD

[1]Department of Medicine, Christchurch, University of Otago, 2 Riccarton Avenue, Christchurch, New Zealand

[10]Department of Psychology, University of Otago, Christchurch, New Zealand

[11]Edible Research, Rangiora, New Zealand

[12]Department of Human Nutrition, University of Otago, Dunedin, New Zealand

[13]Te Whatu Ora Waitaha, Canterbury, Christchurch, New Zealand

[2]High Value Nutrition National Science Challenge, Auckland, New Zealand

[3]NIHR (National Institute for Health and Care Research) Nottingham Digestive Diseases Biomedical Research Unit, Nottingham University Hospitals NHS (National Health Service) Trust and the University of Nottingham, Nottingham, United Kingdom

[4]Grasslands, AgResearch, Palmerston North, New Zealand

[5]Riddet Institute, Massey University, Palmerston North, New Zealand

[6]Plant and Food Research, Palmerston North, New Zealand

[7]Ruakura Research Centre, Plant and Food Research, Hamilton, New Zealand

[8]Malaghan Institute of Medical Research, Wellington, New Zealand

[9]Liggins Institute, University of Auckland, Auckland, New Zealand

[†]deceased

[*]these authors contributed equally

Corresponding Author:
Richard B Gearry, MBChB, PhD
Department of Medicine, Christchurch, University of Otago, 2 Riccarton Avenue, Christchurch, New Zealand

Abstract

Background: Irritable bowel syndrome with constipation (IBS-C) and functional constipation (FC) have significant personal, health care, and social impacts, affecting patients' quality of life. Treatment of these conditions is challenging. While green kiwifruit is a promising natural alternative to laxatives, its effectiveness in managing abdominal pain and the underlying mechanism of action is yet to be substantiated.

Objective: This study investigates the effect of consuming 2 green kiwifruit daily for 4 weeks (the habitual serving) on abdominal pain and discomfort in individuals with IBS-C and FC.

Methods: This study is a 2-arm parallel, open-label, placebo-controlled randomized study. This study's duration was 9 weeks, with a 3-week lead-in phase, a 4-week intervention phase, and a 2-week follow-up phase. A total of 60 participants with IBS-C and FC were randomized to consume either 2 Zespri green kiwifruit (*Actinidia deliciosa* "Hayward," ~150 g per serving, ~90 kcal) or maltodextrin (calorie-matched to the fruit, ~25 g per serving, ~90 kcal) per day for 4 weeks. The participants completed validated questionnaires assessing digestive and general health and well-being parameters, underwent magnetic resonance imaging to determine colon physiological measures, ingested a blue food dye, provided blood and fecal samples to measure microbial, immunological, and biochemical parameters, and ingested wireless motility devices (selected participants only) to assess physiological processes.

Results: Recruitment for this study began in May 2021 and was completed in May 2022. A total of 63 participants were randomized, and 57 were analyzed using intention-to-treat analysis. Data analysis is complete, and full results are expected to be published in a peer-reviewed journal by April 2026.

XSL•FO
RenderX

Conclusions: This study aims to evaluate the effectiveness of green kiwifruit consumption in managing abdominal pain in individuals with IBS-C and FC. It will provide new insights into the mechanisms behind the habitual consumption of green kiwifruit for digestive comfort in this population.

Trial Registration: Australian New Zealand Clinical Trial Registry ACTRN12621000621819; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=381723

International Registered Report Identifier (IRRID): RR1-10.2196/75286

(JMIR Res Protoc 2026;15:e75286) doi:10.2196/75286

KEYWORDS

functional constipation; constipation-predominant irritable bowel syndrome; green kiwifruit; abdominal pain; colonic volume; gut microbiome; metabolites; randomized control trial; mobile phone

Introduction

Background

Chronic constipation affects about 14% of the population and has a substantial socioeconomic burden [1]. The symptoms of constipation include infrequent bowel movements, excessive straining, hard feces, and unsatisfactory defecation experience, with or without abdominal pain [1,2].

The 2 main constipation subtypes are irritable bowel syndrome with constipation (IBS-C) and functional constipation (FC). The key difference between the 2 is the presence of abdominal pain for patients with IBS-C [3]. In addition to constipation symptoms, these conditions are present with abdominal discomfort, bloating, reflux, indigestion, nausea, and vomiting [4]. They are frequently associated with psychological comorbidity, including mood disorders and anxiety, impacting quality of life [2,5,6]. Additionally, many individuals report extra-intestinal symptoms such as somatic pain, fatigue, and fibromyalgia [7-10].

Maintaining healthy dietary habits is essential for managing constipation. International guidelines [11-13] recommend drinking plenty of fluids and incorporating high-fiber foods (eg, fruits, vegetables, whole-grain bread, and cereals) into the diet as a first-line treatment for constipation. Moreover, many individuals prefer to self-manage their constipation symptoms with dietary management as opposed to pharmacological drugs [14].

Green kiwifruit (*Actinidia deliciosa* “Hayward”) is known for its laxative effects [15]. Several interventional studies of fresh fruit have reported improvements in bowel movement frequency, consistency, and ease of defecation [16-22]. Three studies involving participants with constipation have reported a significant improvement in complete spontaneous bowel movements [17,20,21]. Two recent studies demonstrated improved abdominal pain in people with constipation [20,21]. Additionally, some short-term studies have shown that consuming 2 servings of green kiwifruit can improve sleep quality, lipid profile, blood pressure, and platelet aggregation [23-25].

Green kiwifruit is rich in dietary fiber (2% - 3%), which is primarily derived from plant cell walls. The fiber has unique hydration properties, such as water retention, swelling, and viscosity [15,26-28]. A recent magnetic resonance imaging (MRI) study by Wilkinson-Smith et al [29], 2019, found that eating 2 servings of green kiwifruit twice a day for 3 days increased the water content of both the small bowel and the chyme in the ascending colon and raised colonic volume. These results support earlier in vitro research, which suggested that the polysaccharides in kiwifruit cell walls have a significant capacity for swelling and water retention. This indicates that kiwifruit may exert both osmotic and bulking effects in the colon [28,30].

Additionally, in vitro and animal studies have shown that the proteolytic enzyme actinidin from green kiwifruit aids gastric and small intestinal protein digestion [31,32], gastric emptying [33,34], and stimulates gut motility [35]. Only 1 human study has been published on the impact of green kiwifruit on the gut microbiome [36]. The study found that healthy individuals who consumed freeze-dried green kiwifruit (equivalent to 2 kiwifruit servings) for 4 days experienced a favorable shift in the abundance of lactic acid bacteria, such as lactobacilli and bifidobacteria, in fecal matter [36]. Moreover, there was a trend towards a decreased abundance of *Clostridium* and *Bacteroides* genera [36].

There is promising evidence from human, animal, and in vitro studies regarding the laxative effect of green kiwifruit. However, data on the effects of sustained consumption of 2 green kiwifruit daily on abdominal pain and the possible mechanisms of action in individuals with constipation are limited. This underscores the need for a comprehensive study to evaluate both the effectiveness and the underlying mechanisms involved.

Hypotheses and Objectives

This study was based on the hypothesis that consuming 2 green kiwifruits daily over 4 weeks would improve abdominal pain.

Secondary hypotheses were that constipation and other associated symptoms improved. Additionally, the symptom improvement is due to improved physiological functions such as changes in colonic volume and water content, gut motility, gas fermentation profile, and biological functions such as the gut microbiome composition and the host-microbial metabolite profile.

The primary aim of this study was to assess the impact of consuming 2 green kiwifruits daily over 4 weeks on abdominal pain in participants with either FC or IBS-C.

XSL•FO
RenderX

The secondary aims were to determine the impact of consuming green kiwifruit daily over four weeks on (1) constipation and its associated symptoms using validated questionnaires; (2) bowel habits using a daily bowel movement diary; (3) mental health and general well-being parameters using validated questionnaires; (4) dietary fiber intake using validated questionnaires; (5) colonic volume and water content of colonic chyme using a validated MRI protocol; (6) whole gut transit time using blue dye; (7) a relative abundance of individual taxa, predictive function (gene abundances), and diversity indices of the fecal gut microbiota using shotgun metagenomic sequencing, fecal and plasma metabolites using liquid chromatography-mass spectrometry (LCMS); and (8) regional and whole gut transit time and gas fermentation profiles using Smartpill (Medtronic Inc) and Atmo gas-sensing capsules (Atmo Biosciences; Atmo capsule) in a subset of participants.

Methods

Study Design

This study was a 4-week, 2-arm, parallel, open-label, negative-controlled, randomized design investigating the daily impact of 2 Zespri green kiwifruit compared to calorie-matched maltodextrin on abdominal pain in 60 adults with IBS-C or FC. Thirty of these participants were invited to participate in the wireless motility devices (WMDs) substudy.

This study's duration was 9 weeks: 21 days (3 weeks) lead-in phase, 28 days (4 weeks) intervention phase, and 14 days (2 weeks) follow-up phase (Figure 1). The lead-in phase started with participant enrollment and finished a week after the baseline visit. The week between the baseline visit and the intervention phase was designed to give sufficient time for the WMDs to be eliminated from the body [37]. The lead-in phase was followed by the intervention phase, where participants were either randomized to kiwifruit or maltodextrin. This was followed by a 2-week follow-up phase. This study was conducted according to the CONSORT (Consolidated Standards of Reporting Trials) and the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines and adhered to the superiority study framework [38,39].

Figure 1. Study flowchart. MRI: magnetic resonance imaging; WMD: wireless motility device.

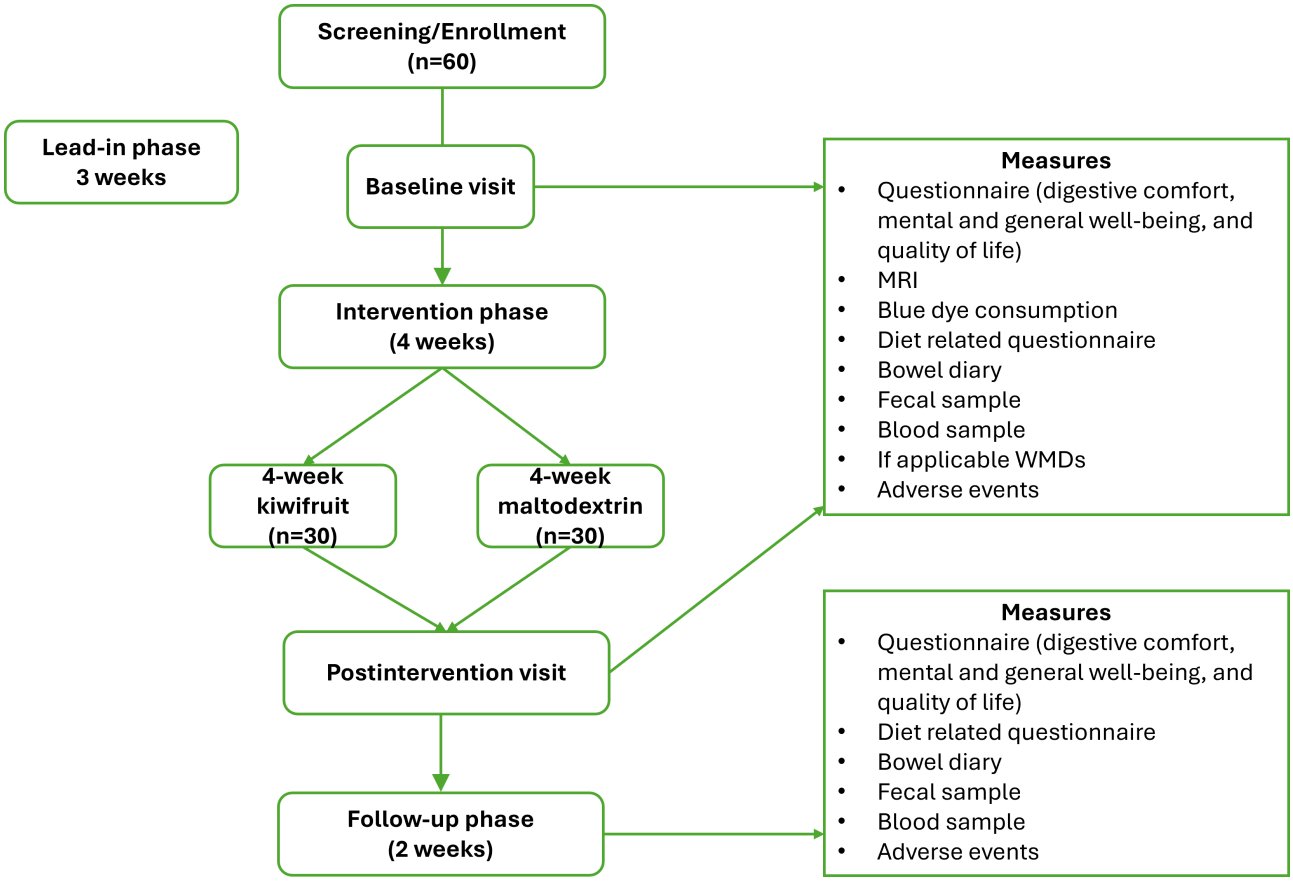

Ethical Considerations

This study was approved by the Health and Disability Ethics Committee, New Zealand (NZ/21/NTB/96) May 14, 2021 and registered with the Australian New Zealand Clinical Trial Registry (ACTRN12621000621819) on May 24, 2021. The University of Otago Māori Health Advancement review process endorsed this study. Before entering this study, eligible participants provided informed consent, and additional informed consent was obtained for the WMDs substudy.

The researchers involved in data collection were experienced in maintaining the security and confidentiality of health information. This study did not report any identifying information about participants, such as names, birth dates, or

XSL•FO
RenderX

images. Identifying details were collected during enrollment and accessed through REDCap (Research Electronic Data Capture; Vanderbilt University) only by the research team.

Participants received supermarket vouchers worth NZD $300 (US $185) as a thank you for completing the study. Additionally, the participants selected for the substudy received NZD $100 (US $62) to compensate for their time and any inconvenience.

Recruitment

Participant recruitment started in May 2021 and ended in May 2022 in Christchurch, New Zealand. Due to slow recruitment at this study's start, limited availability of MRI assessments, and a COVID-19 pandemic lockdown (August 2021-September 2021), study recruitment occurred over 2 kiwifruit growing seasons. Full results are expected to be published in peer-reviewed journals by the April 2026.

Eligibility Criteria

Overview

All interested volunteers were assessed for eligibility based on the following criteria (Textbox 1).

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Adults (18-65 years).
- BMI between 18-35 kg/m^2.
- Function constipation or constipation-predominant irritable bowel syndrome as per ROME IV diagnostic criteria.

Exclusion criteria

- Significant gastrointestinal disorder.
- Alarm features associated with bowel habits.
- Known systematic disease that impacts the gastrointestinal tract directly or through medication.
- Significant chronic diseases.
- Allergy to kiwifruit or latex.
- Inability to access magnetic resonance imaging machine.
- Inability to swallow pills or stop laxatives.
- Irritable bowel syndrome severity score of over 300.
- Significant deviation for regular blood test.
- Blood glucose level of >6 mmol.

Inclusion Criteria

Adults (18 - 65 y) with a BMI between 18 and 35 kg/m^2, with either FC or IBS-C as per Rome IV diagnostic criteria, shown in Textbox 2, were recruited.

XSL•FO
RenderX

Textbox 2. Rome IV diagnostic inclusion criteria for functional constipation (FC) and irritable bowel syndrome with constipation (IBS-C). Source: this textbox was generated using the study by Drossman [3].

FC (criteria fulfilled for the last 3 months with symptom onset at least 6 months before diagnosis)

- Must include two or more of the following:
  - Straining during more than 25% of defecation.
  - Lumpy or hard feces (Bristol Stool Form Scale: 1 or 2) more than 25% of defecations.
  - Sensation of incomplete evacuation of more than 25% of defecations.
  - Sensation of anorectal obstruction or blockage of more than 25% of defecations.
  - Manual maneuvers to facilitate more than 25% of defecations.
  - Fewer than three spontaneous bowel movements per week.
  - Loose feces are rarely present without the use of laxatives.
  - Insufficient criteria for irritable bowel syndrome.

Constipation-predominant irritable bowel syndrome (criteria fulfilled for the last 3 months, with symptom onset at least 6 months before diagnosis)

- Recurrent abdominal pain on average at least 1 day/week in the last 3 months, associated with two or more of the following criteria:
  - Related to defecation.
  - Associated with a change in the frequency of feces.
  - Associated with a change in the form (appearance) of feces.
  - Criteria for constipation-predominant subtype.
  - More than ¼ (25%) of bowel movements with Bristol Stool Scale types 1 or 2.
  - Less than ¼ (25%) of bowel movements with Bristol Stool Scale feces types 6 or 7.

Exclusion Criteria

Participants were excluded if (1) they had a significant gastrointestinal disorder other than IBS-C (eg, diverticulitis, celiac disease, inflammatory bowel disease, or previous bowel resection); (2) alarm features associated with bowel habits such as a recent change in bowel habits (less than three months), rectal bleeding, sudden weight loss, fecal occult blood, anemia, anal fissures, bleeding hemorrhoids, and family history of early onset gastrointestinal cancer or inflammatory bowel disease; (3) any known systemic disease that impacts the gastrointestinal tract directly or through medication use; and (4) significant chronic diseases such as cardiovascular disease, cancer, renal failure, or neurological diseases. Additionally, people with a known allergy to kiwifruit or latex, inability to access an MRI machine, inability to swallow pills, inability or unwillingness to stop laxative use for 7 days (except for rescue medication bisacodyl suppositories), and a failure to comply with study procedures were also excluded. An irritable bowel syndrome severity score of over 300 and significant deviations in complete blood count, renal function, liver function, C-reactive protein, and fasting blood glucose of >6 mmol also resulted in exclusion.

Nonexclusion Criteria

Participants diagnosed with diabetes or conditions requiring selective serotonin reuptake inhibitors, tricyclic antidepressants, or nonsteroidal antiinflammatory drugs were recruited if the health condition and the medication use had been stable for over 3 months.

Screening and Enrollment

Eligibility parameters were collected in 2 steps: participants completed a prescreening online questionnaire and an in-person screening visit. The prescreening questionnaire included questions based on inclusion and exclusion criteria, including questions related to the Rome IV diagnostic criteria (Textbox 2). Based on these replies, participants were diagnosed with either FC or IBS-C. Participants who met all study criteria were invited for a personal screening visit to provide consent and a blood sample to exclude any significant abnormalities.

Upon enrolling in this study, each participant was sent an individualized REDCap survey link hosted by the University of Otago, New Zealand, to complete 2 online enrollment questionnaires. One of the questionnaires was the Modified Hunter New England Health Survey [40], which includes the validated 12-item Short-Form Health Survey Version 2 (SF-12v2) and selected question domains from the New South Wales Population Health Survey.

The SF-12v2 questionnaire was used to evaluate the general health status of the participants, measuring eight domains of functioning and well-being, including physical functioning, role limitations due to physical problems, bodily pain, general health perception, energy and vitality, social functioning, role limitation due to emotional problems, and mental health well-being of the participants. The SF-12v2 questionnaire was scored by using published algorithms [41] to produce the physical component summary-12 and mental component summary-12. Participants

XSL•FO
RenderX

also completed the Economic Living Standard Index short form. This validated questionnaire assesses the standard of living and socioeconomic background, ranging from "very good" to "severe hardship" [42].

Study Intervention

Participants were randomly assigned to consume Zespri 2 green kiwifruit (*Actinidia deliciosa* cultivar "Hayward," ~150 g per serving, ~90 kcal) per day or the calorie-matched negative control (Maltodextrin, 25 g per dose, ~90 kcal) for 4 weeks. The participants were instructed to consume 2 kiwifruit without skin or maltodextrin (available as powder) in 1 daily dose by incorporating it into their morning meal or breakfast. The intervention provided in this study was in addition to their usual diet.

Participants were provided with a minimum of a 14-day supply of kiwifruit on the day of the baseline visit and had their supply replenished for the subsequent 14 days. A few additional days' supply was provided to the participants to ensure good-quality fruit was available, given that fresh fruit is a perishable product. Providing a 14-day supply was considered reasonable, as the shelf life of kiwifruit is typically around 16 - 20 days at room temperature.

The kiwifruit was supplied by Zespri International Limited for the duration of this study.

The negative control was commercially sourced digestible maltodextrin (Davis Food Ingredients). This low-sweet polymer of D-dextrin was selected as a placebo as it can easily match the calorie content of the kiwifruit. Maltodextrin is also completely digested into glucose and actively absorbed in the small intestine [43], which does not increase small- or large-intestinal water content [29]. Participants receiving maltodextrin were asked not to consume kiwifruit during this study's period.

Randomization and Blinding

Following enrollment, participants were randomly allocated to the kiwifruit or maltodextrin group using randomized permuted blocks (block size four) drawn from sealed opaque envelopes. Due to the nature of the intervention, that is, consumption of whole fruit, blinding the participant and the administering researcher was not possible. The researchers involved in the analysis were blinded to the group allocation.

Data and Sample Collection

Figure 2summarizes the timing of measures. Web-based questionnaires were administered to participants at different stages of this study, including screening, enrollment, baseline, postintervention, and follow-up assessments. Individualized REDCap survey links were sent to participants to complete questionnaires. A paper copy of the questionnaires was available on request.

XSL•FO
RenderX

Figure 2. Overview of the schedule of screening, enrollment, interventions, and assessments. FFQ: Food Frequency Questionnaire; GSRS: Gastrointestinal Symptom Rating Scale; IBS-QoL: Irritable Bowel Syndrome Quality of Life Survey; MFI-SF: Multidimensional Fatigue Inventory Short Form; MRI: magnetic resonance imaging; PROMIS: Patient-Reported Outcomes Measurement Information System; PROMIS-GI: Gastrointestinal Domain of the Patient-Reported Outcomes Measurement Information System; SAGIS: Structured Assessment of Gastrointestinal Symptoms; WEMWBS: Warwick-Edinburgh Mental Well-Being Scales; WHO-5: World Health Organization - Five Question Well-Being Index; WMD: wireless motility device. *Participants began intervention 7 days after the baseline visit to ensure the excretion of the WMDwireless motility device.**A subset of participants completed this assessment.

	Screening	Enrollment	Baseline Visit	Start Intervention*	Post-Intervention	Follow Up
Time point	Day X -21 or more	Day -14 days	Day -7	Day 0	Day 28	Day 42
Screening and enrolment						
Eligibility Screen	X					
Informed Consent	X					
Blood Panel	X					X
General health and standard of living		X				
Allocation			X			
Group						
Green Kiwifruit				----------		
Maltodextrin				----------		
Assessment						
Bowel Diary		----------				
Questionnaire						
Rome IV criteria	X		X		X	X
Digestive comfort (GSRS, SAGIS, PROMIS-GI)			X		X	X
Mental well-being (PROMIS, Anxiety and Depression, WHO-5, WEMWBS, MFI-SF)			X		X	X
Quality of life (IBS-QoL)						
Diet-related (3-day food diary, FFQ)			X		X	
Physiome						
MRI			X	X		
Blue dye			X	X		
WMD**			X	X		
Blood Sample			X	X		X
Fecal Sample			X	X		X
Adverse Events		----------				

Participants underwent an MRI assessment, ingested blue food dye, and provided blood and fecal samples at baseline and postintervention. In addition, participants completed a three-day food diary at baseline and postintervention. They also completed a follow-up visit 2 weeks postintervention, providing blood and fecal samples. Participants recorded daily bowel movement diaries using a REDCap-based app from enrollment until the follow-up period. A subset of participants also ingested WMDs at baseline and postintervention.

This study's participants were given a cooler bag that contained a fecal collection kit and an ice pack. They were instructed to collect 2 fecal samples in the provided 60 mL containers (2 containers were provided) ≤24 hours before clinic visits at baseline, postintervention, and follow-up. One sample was stored at 4 °C in their home refrigerator and the other in their freezer (−18 °C) and transported within 24 hours in the provided cooler bag.

Extra fecal collection kits were offered to participants with less frequent bowel movements (occurring once in 3 days or less).

XSL•FO
RenderX

They were encouraged to collect fecal samples 48 to 72 hours before and up to 24 hours after the clinic visit. Participants were also instructed to label the container with the collected time and date and record it in the bowel movement diary app. The collected samples were kept chilled until aliquoted in the laboratory. The home-frozen fecal sample was stored at −80 °C until the meta-transcriptomic analysis. The chilled fecal sample was aliquoted (1 g each) for analyses of microbial composition, microbial gene abundances, metabolome, bile acid, organic acid, and immune marker analyses. All aliquots were snap-frozen in liquid nitrogen before storage at −80 °C until analysis.

The participants also provided a blood sample (described below) at the in-person screening visit, baseline, postintervention, and follow-up. At the screening visit, blood samples were collected for complete blood count in a 4 mL ethylenediaminetetraacetic acid (EDTA) vacutainer and an 8 mL heparin tube for standard biochemistry analyses (liver function, renal function, electrolytes, C-reactive protein, and blood glucose) at Canterbury Health Laboratory (Christchurch, New Zealand). These analyses were repeated at follow-up visits before participants were discharged from this study.

At baseline and postintervention, blood samples were collected in one 4 mL EDTA tube, three 9 mL EDTA tubes, and one 6 mL heparin tube. The 4 mL EDTA and 6 mL heparin tubes were centrifuged for 5 minutes at 2000 x g with full acceleration and no deceleration. The plasma was aliquoted (4×500 µL) into labeled Eppendorf tubes and stored at −80 °C until metabolome, targeted metabolite, and immune profile analyses were performed. Peripheral blood mononuclear cells (PBMCs) were isolated from approximately 30 mL of whole blood (3×9 mL EDTA vacutainers) using the SepMate technique [44]. Briefly, the blood was split into two 15 mL centrifuge tubes and diluted in a 1:2 v/v ratio by adding 15 mL of Dulbecco's phosphate-buffered saline containing 2% heat-inactivated fetal bovine serum to each tube. Next, the diluted blood was transferred into two SepMate tubes containing 15 mL of Lymphoprep. The SepMate tubes were centrifuged at 1200 x g for 10 minutes. After centrifugation, the top layer was poured into a new tube. This tube was centrifuged at 700 x g for 8 minutes, and the supernatant was removed. The pelleted PBMC was washed with 10 mL of Dulbecco's phosphate-buffered saline containing 2% heat-inactivated fetal bovine serum and centrifuged again at 300 x g for 8 minutes, aspirating the supernatant while leaving the pelleted PBMC intact. Finally, the PBMC pellet was resuspended in 3 mL of a freezing mixture of 90% heat-inactivated fetal bovine serum and 10% (v/v) dimethyl sulfoxide and aliquoted into three cryovials. The cryovials were placed into a CoolCell LX freezer container and stored at −80 °C overnight. The following day, all the frozen cryovials were transferred to liquid nitrogen storage at −80 °C.

Primary Outcome

The self-administered Gastrointestinal Symptom Rating Scale (GSRS, AstraZeneca) [45], a 15-item questionnaire, assessed common gastrointestinal symptoms, including abdominal pain, constipation, diarrhea, reflux, and indigestion. The abdominal pain domain of GSRS was used to determine the primary outcome of abdominal pain and discomfort relief. The GSRS is a 7-point Likert-type scale, where 1 represents the absence of symptoms, and 7 represents severe discomfort. The scale can be used as a whole, or the mean value for items under each symptom domain can be assessed to give a subscale [45]. The mean change in symptom scores between baseline and postintervention measurements was compared to evaluate the effectiveness of the intervention.

Secondary Outcome

Physiological Parameters

About MRI

The MRI images were acquired using a MAGNETOM Skyra 3T scanner (Siemens Healthcare) by Pacific Radiology, and MRI image analysis was completed by members of the Nottingham Biomedical Research Center using a noninvasive MRI technique described elsewhere [29]. Briefly, participants were instructed to eat a standardized dinner the day before and fast overnight. They were then positioned supine in the scanner, and the colon was imaged using a variety of breath-holding scanning sequences.

The colonic volumes were assessed using a coronal 3D DIXON dual-echo sequence with a field of view (400×400 mm^2) positioned to visualize the large intestine, from the splenic flexure down to all the colonic regions in the pelvis. Slices (48 in number) of 3.5 mm thickness covered the abdomen from anterior to posterior. The in-plane reconstructed resolution was 0.78×0.78 mm^2 with a matrix of 512 ×512. Four image types were derived from the data, which provided images of water only, fat only, fat and water in phase, and fat and water out of phase. Individual regional colon volumes were manually segmented on each slice using Medical Image Processing, Analysis, and Visualization software (National Institute of Health) [46].

The colon of the participants was divided into 4 segments: ascending colon, transverse colon, descending colon, sigmoid colon, and rectum region. Each colon segment was identified within each coronal image slice, and a region of interest (ROI) was drawn around it, building a 3D representation of the morphology. The volume of each segment was measured by summing the areas of the corresponding ROI and multiplying by slice thickness. The total colon volume was calculated by summing the volumes of the 4 separate segments.

The relaxation time (T1), a time constant determining the rate at which magnetically excited water hydrogen protons in the colonic chyme take to return to equilibrium, was measured for the ascending and descending colon and reflects the water content of the chyme [47]. Two sagittal-oblique True Fast Imaging With Steady-State Free Precession (TRUFI) sequences were used to localize the ascending and descending colon, respectively.

Each colonic segment was scanned using an inversion recovery single-slice sagittal oblique TRUFI sequence, using 8 different inversion times ranging from 100 - 5000 ms. Each different inversion time required a separate breath-hold, with at least a 15-second gap between each to allow full signal recovery before

XSL•FO
RenderX

the subsequent acquisition. The field of vision was 284×400 mm^2 with a slice thickness of 7 mm, with a reconstructed in-plane resolution of 1.56×1.56 mm^2 and a matrix size of 182×256.

ROIs were drawn on all images in the chyme at 3 locations in each segment using custom-written software in Matlab (The MathWorks Inc). The mean signal in each ROI was measured at each inversion time. The data for each ROI was fitted to a model of the expected signal intensities in the inversion recovery–TRUFI sequence to determine the T1 as previously described [47]. Single values of T1 for the ascending and descending colon, respectively, were calculated from the average T1 of the 3 regions in that segment.

Blue Food Dye

Participants ingested Royal Blue Liqua-gel (Chefmaster) food coloring (12 drops/1.5 g) in water to measure the total gut transit time [48]. The time of ingestion was recorded, and participants were requested to log the first appearance of "blue/green feces" on the app-based bowel diary. The date and time (h, min) of the blue color's appearance were recorded through this app. The time of ingestion and exit was used to calculate the whole gut transit time.

Patient Reported Outcome

Constipation and Associated Symptoms

Constipation and its associated symptoms were assessed by validated questionnaires, including GSRS, the Rome IV questionnaire [3] Structured Assessment of Gastrointestinal Symptoms (Universities of Queensland and Newcastle) [49] and the Gastrointestinal Domain of the Patient-Reported Outcomes Measurement Information System (Health Measures) [50]. The Rome IV questionnaire classifies participants into IBS-C and FC groups and evaluates for the shift into a different classification during this study's period. Structured Assessment of Gastrointestinal Symptom measures the intensity and impact of 22 upper and lower gastrointestinal symptoms and the presence of selected nongastrointestinal symptoms. Four study-relevant domains (disrupted swallowing, constipation, belly pain, and gas, bloating, or flatulence) of the Gastrointestinal Domain of the Patient-Reported Outcomes Measurement Information System were administered to assess gastrointestinal symptoms and group effects.

Bowel Movement Diary, Constipation, and Associated Symptoms

A REDCap app-based daily bowel habit diary captured data regarding bowel movement frequency, spontaneity, completeness, ease of defecation, use of manual maneuvers, Bristol Stool Scale, and menstruation (if applicable). The participants were asked to report the following after each bowel movement. This provided a comprehensive record of bowel habits and was used to determine the number of complete spontaneous bowel movements per week, defined as a complete evacuation not induced by laxatives, rescue medications, manual maneuvers, or enemas [21]. Other parameters include total bowel movement per week, average Bristol Stool Scale, and straining. In addition, this app was used to collect the time of excretion of blue food dye.

Mental Health and General Well-Being

General and mental well-being parameters were assessed by the World Health Organization - Five Question Well-Being Index (WHO-5) [51], the Warwick-Edinburgh Mental Well-Being Scale [52], the Multidimensional Fatigue Inventory Short Form [53], and Patient-Reported Outcomes Measurement Information System (PROMIS) anxiety and depression domain questionnaires [54].

The 5-item WHO-5 assesses recent experiences of well-being related to physical vitality, cheerfulness, and interest. The 14-item Warwick-Edinburgh Mental Well-Being Scale assesses psychological functioning (optimism, autonomy, agency, curiosity, clarity of thought, and positive relationships) and positive affect (confidence, feeling relaxed, cheerful, and having the energy to spare). The 30-item Multidimensional Fatigue Inventory Short Form measures various dimensions of fatigue (general fatigue, physical fatigue, emotional fatigue, mental fatigue, and vigor). The 8-item PROMIS depression questionnaire Version 8a assesses common symptoms of depression, such as worthlessness, depressed mood, and anhedonia. The 8-item PROMIS anxiety questionnaire Version 8a assesses common symptoms of anxiety, such as nervous mood, worries, and tension. In addition, the 34-item Irritable Bowel Syndrome - Quality of Life Survey measured the impact of irritable bowel syndrome and its management [55].

All questionnaires queried symptoms over the past week following their standard timeframes. The WHO-5 timeframe was modified from the past 2 weeks to the past week to maintain consistency with the other questionnaires.

Dietary Intake

All enrolled participants completed a nonconsecutive 3-day food diary before each study visit, that is, at baseline, postintervention, and follow-up. The food diary included written prompts for detailed information about when, where, and with whom food was consumed, the type of food or drink, brand details, preparation or cooking methods, quantity, and a section for recipe details. Participants were also asked to record if their intake differed from their normal eating routines. At each clinic visit, this study's dietitian reviewed each food diary. Two nutrition graduates trained in dietary assessment entered the food diaries into a dietary analysis software, FoodWorks Online Professional (Version 1.0, Xyris Software, 2021). The NZ Food Composition Data (FOODfiles 2018 – Version 1.0) was used to estimate energy and nutrient intake.

In addition to the food diaries, participants completed a fiber-specific Food Frequency Questionnaire (FFQ) before each clinic visit. This FFQ assessed inulin and oligofructose consumption from 9 commonly eaten foods: wheat, onion, garlic, banana, leek, rye, asparagus, chicory, and barley [56]. The FFQ data were entered into a spreadsheet. Once all the data were collected, the spreadsheet was processed by a bespoke algorithm to calculate inulin and oligofructose consumption and the relative contributions of each food item.

XSL•FO
RenderX

Biological Measurements

Fecal Microbiome

The taxonomic composition and gene abundances of the fecal microbiome were analyzed by shotgun metagenomics. Auckland Genomics constructed libraries for metagenomics sequencing of the extracted DNA at the University of Auckland, New Zealand. Microbial metagenomic libraries were then generated using the Seqwell PurePlex DNA library preparation approach [57]. Briefly, DNA was extracted from fecal samples using ZYMOBiomics DNA and ribonucleic acid Miniprep kits following the manufacturer's instructions, with the addition of bead beating on a BioSpec Mini-Beadbeater 96 set to 4 minutes. DNA was measured using the Qubit high-sensitivity kit. DNA was normalized to 5 ng/µl using the Eppendorf Epmotion robot. The DNA was prepared for Illumina shotgun sequencing using the Seqwell Pureplex Unique Dual Index library preparation kit (Part No. 301067 for Set 1A; 301068 for Set 2B). The samples were shotgun sequenced using the Illumina Novaseq 6000 platform using 2 ×150 bp paired-end sequencing at Novogene in Singapore.

The raw data quality was checked using FastQC (V0.11.9). The software Trimmomatic (V0.36) [58] was then used for the removal of adapters, low-quality (Phred scores <30), and short (<36 bp) sequencing reads. Read pairs were aligned to the human reference genome (RefSeq: GCF_000001405) [59] using the "mem" algorithm of Burrow-Wheeler Aligner (V0.7.17-r1188) [60]. The unmapped sequencing reads were converted into fastq files using the "fastq" function of samtools (V1.8) [61]. Read pairs were joined using paired-end read merger (V0.9.6) [62] with default settings. Read pairs that did not join were pasted together with a string of N's using the "fuse" function from the *BBMAP* package (V38.22 - 0) [63]. Joined and fused reads from different lanes from the same sample were compiled into a final "clean" read sample file. Metagenomic functions were obtained through the "blastx" function of Diamond (version 0.9.22) [64], mapping the reads against the "non-redundant" National Center for Biotechnology Information database [65]. Megan (version 6 ultimate edition) [66] was used to assign putative functions to the alignment files produced by Diamond. A trained statistician (AgResearch) performed this alignment.

Differential abundance was performed using DESeq2 in R software (version 1.40.2; R Foundation for Statistical Computing). This package analyzes differentially expressed taxa based on a negative binomial distribution. The effect of groups was assessed by comparing between time points and groups. The top 10 differentially abundant taxa were extracted based on false discovery rate (FDR) adjusted *P* values for both baseline and postintervention for each group. Taxa that changed between the two time points within each group were identified by finding the set difference between the top taxa between baseline and groups postintervention. Log2 fold change values were extracted for the identified changed taxa in each group. Enrichment of functional gene attributes arranged hierarchically with level 1 (broadest level of function, eg, metabolism and cellular processes), level 2 (specific functions, eg, carbohydrate metabolism and amino acid metabolism), and level 3 (detailed pathways, eg, glycolysis and glycan degradation), was analyzed.

Plasma and Fecal Metabolome

Fecal samples were freeze-dried and extracted using a previously described method [67] with minor modifications. Briefly, 50 mg of lyophilized fecal samples were homogenized with a ceramic bead for 1 minute, and then 400 µL of 75% methanol/MilliQ water was added. The mixture was vortexed (30 s), sonicated (2 min), and transferred to ice for 10 minutes. Next, 1 mL of methyl tert-butyl ether was added, and the resulting mixture was incubated for 1 hour at 450 rpm. Then, 550 µL of MilliQ water was added, and the mixture was incubated (10 min) before being centrifuged at 14,000 x g for 25 minutes to separate the aqueous (lower) and organic (upper) phases. The polar and lipid extracts were evaporated under a stream of nitrogen at room temperature, and the resulting dried samples were stored at –80 °C until subjected to LCMS analysis. The metabolite profiling analyses were conducted using high-resolution LCMS on a Shimadzu Q-Tof 9030 with electrospray [68]. The polar extract was analyzed using hydrophilic interaction LC [68], and semipolar metabolites were resolved using reverse-phase liquid chromatography [69]. The lipidomic methodology was used to analyze the phase extract [70].

Polar, semipolar, and nonpolar metabolites were extracted from plasma using a previously described method [69]. Briefly, polar metabolites were extracted from 50 µL of plasma with 450 µL of prechilled acetonitrile: water (9:1 v/v). The mixture was shaken for 60 seconds before being centrifuged at 14,000 g at 4 °C for 10 minutes, and then the extract was placed into a high-pressure liquid chromatography (HPLC) vial for analysis. For semipolar metabolite extraction, 400 µL of ice-cold chloroform:methanol (1:1 v/v) was added to 50 µL of plasma, the mixture was then vortexed (30 s) and incubated (1 h) at −20 °C [71]. Then, 200 µL of MilliQ water was added to the mixture, vortexed, and centrifuged at 14,000 g (10 min) at 4 °C. The supernatants were evaporated to dryness under a stream of nitrogen and stored at −80 °C until analysis.

On the analysis day, the semipolar extracts were thawed at room temperature (18±2 °C) and redissolved in acetonitrile:water (1:9 v/v). The mixture was vortexed (1 min) and centrifuged (14,000 g, 4 °C, 10 min). The extract was transferred into an HPLC vial for analysis. For lipid extraction, 10 µL of plasma was mixed with 95 µL of butanol:methanol (1:1 v/v) and spiked with 5 µL of internal standard SPLASH mix (Avanti Lipids) [70]. The mixture was then vortexed (1 min), sonicated (60 min) at room temperature, and centrifuged at 14,000 g (10 min) at 20 °C. All extracts were transferred into an HPLC vial and stored at −80 °C until LCMS analysis. The plasma polar, semipolar, and nonpolar extracts were analyzed using the LCMS methods described for the fecal samples [70].

All LCMS data files were obtained in Data Independent Acquisition mode, facilitating mass spectrometry (MS)/MS spectral annotation and identification. The resulting data files were processed using MS-DIAL and publicly available databases [72]. MS-DIAL performs peak detection, retention time alignment, grouping, gap filling, run order, and batch correction

XSL•FO
RenderX

using the LOWESS algorithm and then searches and annotates the aligned peaks found in the collected Data Independent Acquisition MS/MS spectral data.

The lipidome results were matched against the built-in lipid library containing 257,000 silico-generated MS/MS lipid fragmentation spectra. Semipolar and polar compounds were searched against the MS/MS public library, which included 13,303 unique compounds. This data matrix was subsequently exported to filter unreliably measured peaks using the quality control samples (relative SD>0.3). Furthermore, any peaks that occurred in the samples and blanks that were less than 3 times the size of the average peak in the blank were removed. The resultant data matrix was ready for downstream statistical analyses.

Plasma and Fecal Bile Acid

Extraction methods for bile acid analysis followed those previously outlined. Briefly, 50 mg of wet fecal samples were extracted with 100 µL ice-cold methanol containing internal standards (10,000 nM of d_5-CA and d_5-CDCA). After homogenizing the mixture for 30 s, it was incubated at −4 °C for 30 minutes and then centrifuged at 18,000 g for 20 minutes. The supernatant (20 µL) was then dissolved in 80 µL of 0.1% aqueous formic acid solution and vortexed before being analyzed using LCMS on an AB Sciex LCMS/MS QTRAP 6500+ system (Sciex) coupled to an ExionLC (Shimadzu) [73]. Bile acid peak areas were measured, and absolute concentrations (ng/g dry fecal matter) were calculated for each bile acid based on the instrument response of known concentrations of each bile acid relative to the deuterated bile acid internal standards.

Plasma and Fecal Organic Acids

Aliquots of 200 µL of heparin plasma and 1 g of fecal samples were sent to Plant & Food Research to measure organic acid concentrations using an LCMS method. Fourteen linear and branched organic acids (carbons 1 to 6) were derivatized with an MS probe and stable isotope techniques as previously described [74] with modifications [75,76]. The names and corresponding acronyms of organic acids are presented in Table 1. Plasma organic acids were measured using targeted LCMS on an AB Sciex LCMS/MS QTRAP 7500 instrument equipped with a Turbo V ion source and atmospheric pressure chemical ionization probe (Sciex) coupled to a Nexera UHPLC (Shimadzu). Fecal organic acids were measured using targeted LCMS on an AB Sciex LCMS/MS QTRAP 5500 instrument equipped with a Turbo V ion source and atmospheric pressure chemical ionization probe (Sciex) coupled to an Exion UHPLC (Shimadzu). Labeled internal standards for each organic acid were used to ensure accurate quantification.

Table . Plasma and fecal organic acids were analyzed using LCMS with the corresponding acronym.

Full name	Acronym
Formic acid	FA
Lactic acid	LA
Acetic acid	AA
Propanoic acid	PA
Isobutyric acid (2-methylpropanoic acid)	IBA
Butyric acid (butanoic acid)	BA
Succinic acid	SucA
2-methyl butyric acid (2-methylbutanoic acid)	2MBA
Isovaleric acid (3-methylbutanoic acid)	IVA
Valeric acid (pentanoic acid)	VA
3-methyl valeric acid	3MVA
4-methyl valeric acid (isocaproic acid)	4MVA
Caproic acid	CA
Hexanoic acid	HA

[a]LCMS: liquid chromatography-mass spectrometry.

Plasma Neurotransmitters

Two aliquots of 500 µL of heparin plasma were sent to Plant and Food Research for quantitative measurement of neurotransmitters from the tyrosine and tryptophan metabolic pathways identified as important to the Gut-Brain Axis using two separate LCMS methods as previously described [77]. Method A uses an MS-probe and stable isotope coding derivatization LCMS method to quantitatively convert neurotransmitters and related compounds into their corresponding acetate or ester, thereby increasing their analytical sensitivity. Method B is used to quantify kynurenic acid, quinolinic acid, and xanthurenic acid, which are poorly derivatized using method A. These three analytes and their labeled internal standards are measured directly by LCMS and are underivatized. The names and corresponding acronyms of neurotransmitters are presented in Table 2. Plasma neurotransmitters were measured using targeted LCMS on an AB Sciex LCMS/MS QTRAP 7500 instrument equipped with

XSL•FO
RenderX

a Turbo V ion source and APCI probe (Sciex) coupled to a Nexera UHPLC (Shimadzu).

Table . Plasma neurotransmitters analyzed using LCMS[a] with the corresponding acronym.

Full name	Acronym
Serotonin	5HT
L-kynurenine	KYN
Tryptophan	TRP
Quinolinic acid	QA
DL- 5 hydroxytryptophan	5HTP-DL
Xanthurenic acid	XA
Kyruneric acid	KA
Hydroxy-indoleacetic acid	5-HIAA

[a]LCMS: liquid chromatography-mass spectrometry.

Plasma Immune Markers

Two EDTA plasma and 3 aliquots of isolated PBMC were sent to the Malaghan Institute of Medical Research for plasma cytokine and immune profile analysis. The concentration of circulating cytokines, including IL-1β, IL-6, IL-8, IL-10, IL-12p70, IL-17A, IL-18, IL-23, IL-33, interferon-α2, interferon-γ, tumor necrosis factor-α, and monocyte chemoattractant protein-1 was measured using Human Inflammation Panel 1 LegendPlex assay (Cat #: 740809, BioLegend) as per manufacturer instructions. The PBMCs were counted using flow cytometry with a commercial staining kit to capture their activation state.

About WMDs

During the main study, participants were invited to ingest a Smartpill and an Atmo capsule. The dual WMD ingestion protocol was similar to a previously described protocol [78]. Participants consumed a standardized cereal bar (Smart bar, Medtronic Inc) followed by the WMD with water. They fasted for 6 hours postingestion and wore transponders across their body (or close to their body at night) until WMD excretion. The participants were asked to enter additional bowel movement data on the Smartpill and Atmo capsule transponders for the duration of the substudy. Participants with gut transit times exceeding the capsule's or transponder's battery life (~5 d) were requested to visualize their feces to confirm the excretion of the capsule. If the excretion of one or both capsules could not be determined from the transponder data or visual confirmation, an abdominal radiological examination was conducted.

Smartpill measures anatomical landmarks and regional transit time via intraluminal temperature, pH, and pressure changes as it passes through the gastrointestinal tract [79]. The novel Atmo capsule measures anatomical landmarks and colonic fermentation by measuring oxygen, hydrogen, carbon dioxide, volatile organic compounds, capsule orientation, environmental electromagnetic properties, and temperature changes [78]. The analyses included regional transit time and whole-gut transit time measured by SmartPill and Atmo capsule, and colonic gas (hydrogen and carbon dioxide) profiles.

Smartpill transponder data were downloaded and analyzed using the proprietary software MotiliGI (version 2.5, Medtronic Inc). Atmo capsule transponder data was collected in real time on the smartphone app and transmitted to the Atmo cloud (hosted by Atmo Bioscience) for review and analysis.

Adverse Events

All participants were informed about potential side effects. They were encouraged to record any symptoms and contact their health care provider if needed.

Compliance

In this study, compliance was not formally measured. Participants received small batches of the intervention throughout this study's period and were encouraged to continue consuming the interventional product. The food diaries collected during the postintervention visit provided insights into the consumption of kiwifruit and maltodextrin.

Sample Size

This study was powered to detect an effect size of 1 within the abdominal pain domain of the GSRS, 5% significance (α), and 80% power (β)=16 participants regarding a difference between the groups based on GSRS data obtained from studies with gold kiwifruit [80-82]. In addition to the primary outcome, this study was intentionally powered to ensure the detection of differences between groups in colonic volume as measured by MRI as a secondary outcome. Based on a previous study with 300 g of kiwifruit, a parallel design with 5% significance (α) and 90% power (β), the minimum sample size was determined as 22 participants per group [29]. To account for the lower daily kiwifruit intake (150 g) over 4 weeks and an estimated dropout rate of 10%, a total of 30 participants needed to be recruited for each group.

Statistical Analysis

This study was conducted as a superiority trial. The Guidelines of the Committee for Proprietary Medicinal Products (now termed Committee for Medicinal Products for Human Use) require intention to treat analysis [83]. All data from randomized participants who received at least one intervention dose and had at least one postintervention measurement were analyzed. All

XSL•FO
RenderX

statistical analyses were performed using SPSS software (version 29.0; IBM Corp) or R software (version 4.3.1; R Core Team 2021) by blinded researchers under the guidance of an independent biostatistician.

Baseline characteristics were presented using descriptive statistics. The mean and SD, or median and range, were used to describe continuous variables. Frequencies and percentages were used to describe categorical variables. The group effects (ie, the mean change from baseline between groups) were assessed using (parametric) *t* tests and (nonparametric) Mann-Whitney or Kruskal-Wallis tests for symmetrically and asymmetrically distributed data. The mean, SD, and CIs were used to describe continuous variables. Categorical variables were assessed using chi-square tests (or Fisher exact tests for small samples). A 2-tailed statistical significance $P<.05$ showed statistical significance. A P value of <.1 would be considered a trend for most variables.

To analyze the fecal microbiome measures, both univariate and multivariate statistical methods were used to evaluate variations in alpha and beta diversity and differences in the relative abundance of taxa and gene abundance differences between groups. Differences in alpha diversity were assessed using the nonparametric Wilcoxon signed rank test with continuity correction, comparing baseline and postintervention data, as well as between groups postintervention. The Benjamini-Hochberg method was applied to control for the FDR associated with multiple hypothesis testing for variables in microbiome analyses. Adjusted P values were calculated to mitigate the inflated type I error rates arising from numerous comparisons. A $P<.05$ and FDR of $P<.1$ were considered significant, with a probable biologically significant difference (trend) at unadjusted $P\leq.1$ indicating a trend.

Boxplots were created using ggplot2 to compare groups for each α diversity (Chao1 and Shannon indices) across time points. The statistical significance of microbial community composition was assessed using the permutational multivariate ANOVA and the Analysis of Dissimilarity (number of permutations: 999), which uses permutation testing and F tests to evaluate differences in beta diversity. Differential in microbial gene expression was analyzed using quasi-likelihood F tests, controlling for dispersion uncertainty. Genes with an absolute log-fold change greater than 1.1 were deemed differentially expressed.

As mentioned above, for plasma or fecal metabolome measures, the comparisons were between each group and baseline and between groups postintervention. The partial least squares projection-discriminant analysis (PLS-DA) model was performed to identify metabolites using SIMCA, version 17 (Umetrics). The PLS-DA models were subject to 100-fold permutations to evaluate performances and validated using the predictive ability of the model (Q2) and the ANOVA testing of cross-validated predictive residuals [84]. The most discriminating features were selected for building PLS-DA models. These models were again subjected to prediction accuracy and overfitting assessment by the corresponding tests (cross-validated predictive residuals) and comparison of the Q2 afterward. Features were selected based on their variable importance in the projection score. Univariate statistical analysis and heatmaps were performed using Metaboanalyst 5.0 [85]. The metabolite and lipid relative intensity differences between groups were tested using the *t* test or the Wilcoxon rank-sum test. The multiple testing corrections were controlled using FDR correction [86].

Results

Recruitment began in May 2021 and concluded in May 2022. Of the 63 participants who were randomized, 6 dropped out before receiving the intervention. Consequently, 57 participants were analyzed based on the intention-to-treat principle.

The average age of the cohort was 42.39 (SD 12.83, range: 18 - 65) years, and their average BMI was 26.27 (SD 4.26) kg/m^2. Of the 57 participants, 86% (n=49) were females, and 79% (n=45) were New Zealand Europeans. There were more individuals with FC (n=39, 68.4%) than IBS-C (n=8, 31.6%).

Data analysis is complete, and full results have been submitted to a peer-reviewed journal, with an expected publication date of April of 2026. Meta-transcriptomics and PBMC analysis were not performed due to poor sample quality.

Discussion

Anticipated Findings

Several studies have demonstrated that the consumption of green kiwifruit improves laxation. However, there are few robust randomized controlled trials assessing the efficacy and mechanism of green kiwifruit ingestion for people with constipation. To our knowledge, this is the first study to evaluate the effectiveness of consuming 2 green kiwifruit daily for 4 weeks on abdominal pain as a primary end point and on colonic volume and chyme water content in individuals with constipation. In addition to physiome outcomes, the fecal microbiota composition and gene abundance, plasma and fecal metabolites, and plasma immune markers were also analyzed to understand the underlying mechanism.

Improvement in pain was chosen as a primary outcome to assess whether kiwifruit offers benefits beyond laxation. Specifically, the goal was to evaluate kiwifruit's suitability as a natural dietary intervention for individuals who experience both pain and constipation. Several dietary fiber interventions, including wheat bran, psyllium, and inulin, are known to alleviate constipation but are not preferred by patients because they may exacerbate pain [20,87].

In addition, this study included a range of patient-reported outcomes to evaluate nongastrointestinal symptoms and mental well-being, aiming to determine whether kiwifruit positively impacts health beyond its nutritional composition. A recent study showed that habitual kiwifruit consumption improved the quality of life in people with constipation, supporting that kiwifruit provides benefits that extend beyond mere nutrition [21].

This study used an MRI protocol to measure colonic volume and water content after habitual kiwifruit consumption over 4

XSL•FO
RenderX

weeks. This approach replicates a more natural setting where kiwifruit is integrated into the regular diet of individuals with constipation. A previous MRI study used serial imaging to examine the effects of a high dose of kiwifruit in healthy adults, providing valuable insights into the potential mechanism [29]. Additionally, this study also takes into account the secondary outcomes of colonic volume and water content when calculating the sample size, aiming to optimize its clinical utility and reduce the chances of false positive results [88].

This study was an open-label design. Both participants and researchers were aware of the group assignments for kiwifruit and maltodextrin. Open-label studies are known to have potential biases in reporting and analysis, so several measures were implemented to minimize these risks [89]. Various questionnaires were used to reduce these risks by asking the same questions in different forms and styles. Additionally, only two researchers were informed about the group allocations, and multiple endpoints were assessed, including both subjective and objective outcome measures. Participants completed the subjective end points independently, without the presence of a researcher. Additionally, a blinded outcome analysis was conducted, and the trial's protocol and statistical plan were finalized before this study commenced [90].

The inclusion criteria could be expanded to include individuals aged 65 years and older. This is a significant group worldwide [91,92]. Older adults experience changes in their gastrointestinal function, which alter taste, motility, hormone secretion, and anorexia [93], resulting in their exclusion from many studies. However, it is a growing population with a known increased prevalence of bowel irregularity. Future research could assess the effectiveness and the change in gastrointestinal function and fecal microbiota parameters following high-fiber food interventions. Such dietary interventions could be particularly beneficial for those experiencing bowel irregularities due to polypharmacy, defined as taking more than 5 medications [94].

Conclusions

This study was the first to investigate the effectiveness of habitual consumption of kiwifruit (2 servings daily for 4 weeks) on abdominal pain, colonic volume, and colonic chyme water content. This study also assessed patient-reported outcomes and objective regional physiological, biochemical, and microbial changes that may aid digestion, absorption, and fermentation in order to understand the mechanism of action.

Acknowledgments

Acknowledgments are also due to Dr. Karin Schofield for project management duties, as well as this study's participants, the University of Otago Christchurch AskOtago receptionists, Stewart St Medical Specialists receptionists, the University of Otago Department of Human Nutrition administrators, Research and Enterprise Department administrators, and Divisions of Health Sciences and Sciences staff, and drivers who helped the research team around the logistics of this study.

Funding

The study received primary funding from the High-Value Nutrition National Science Challenge via the New Zealand Ministry of Business, Innovation and Employment, contract number (UOAX1902). Zespri contributed approximately 10% of the co-funding and in-kind support. The funders were not involved in the study design, data collection, data analysis or interpretation of the results, manuscript writing, or the decision to publish the results.

Authors' Contributions

Conceptualization: AM (supporting), NCR (lead), RBG (lead)
Data curation: J Maggo (equal), HMN (equal)
Formal analysis: J Maggo (equal), HMN (equal), CLW (supporting), CLH (supporting), CSG (supporting), CF (supporting), DC (supporting), J Mullaney (supporting), JC (supporting), JT
(supporting), LM (supporting), KF (supporting), MF (supporting), OG (supporting), RS (supporting)
Funding acquisition: NCR (equal), RBG (equal)
Investigation: J Maggo (equal), HMN (equal), CSG (supporting), DC (supporting), J Mullaney
(supporting), JC (supporting), JT (supporting), KF (supporting), OG (supporting), SBB (supporting), TC (supporting), TT (supporting)
Methodology: NCR (lead), RBG (lead), CLH (supporting), DC (supporting), J Mullaney (supporting), JC (supporting), JT (supporting), LM (supporting), KF (supporting), OG (supporting), RS (supporting), TC (supporting), TT (supporting)
Project administration: J Maggo (equal), HNM (equal)
Supervision: CLW (supporting), CSG (supporting), CF (supporting), DC (supporting), J Mullaney (supporting), JC (supporting), KF (supporting), MF (supporting), NCR (equal), RBG (equal), SBB (supporting)
Validation: SBB
Writing – original draft: J Maggo (equal), HMN (equal)
Writing – review & editing: J Maggo (equal), HMN (equal), CLW (supporting), CSG (supporting), CF (supporting), DC (supporting), J Mullaney (supporting), JC (supporting), JT (supporting), KF (supporting), MF (supporting), NCR (supporting), OG (supporting), RBG (supporting), SBB (supporting), TT (supporting), WCM (supporting)

XSL•FO
RenderX

All authors have read and agreed to the published version of this paper.

Conflicts of Interest

The authors declare that the research was carried out without any commercial or financial affiliations that could be perceived as a potential conflict of interest. RBG sits on the Science Advisory Board for Zespri. RBG and SBB have received travel and research grants from Zespri International. J Maggo has received a travel grant from Atmo Bioscience.

References

1. Sharma A, Rao S. Constipation: pathophysiology and current therapeutic approaches. Handb Exp Pharmacol 2017;239:59-74. [doi: 10.1007/164_2016_111] [Medline: 28185025]
2. Camilleri M, Ford AC, Mawe GM, et al. Chronic constipation. Nat Rev Dis Primers 2017 Dec 14;3:17095. [doi: 10.1038/nrdp.2017.95] [Medline: 29239347]
3. Drossman DA. Functional gastrointestinal disorders: history, pathophysiology, clinical features, and Rome IV. Gastroenterology 2016 May;150(6):1262-1279. [doi: 10.1053/j.gastro.2016.02.032]
4. Wong MYW, Hebbard G, Gibson PR, Burgell RE. Chronic constipation and abdominal pain: independent or closely interrelated symptoms? J Gastroenterol Hepatol 2020 Aug;35(8):1294-1301. [doi: 10.1111/jgh.14970] [Medline: 31900961]
5. Silk DB. Impact of irritable bowel syndrome on personal relationships and working practices. Eur J Gastroenterol Hepatol 2001 Nov;13(11):1327-1332. [doi: 10.1097/00042737-200111000-00011]
6. Canavan C, West J, Card T. Review article: the economic impact of the irritable bowel syndrome. Aliment Pharmacol Ther 2014 Nov;40(9):1023-1034. [doi: 10.1111/apt.12938] [Medline: 25199904]
7. Vakil N, Stelwagon M, Shea EP, Miller S. Symptom burden and consulting behavior in patients with overlapping functional disorders in the US population. United Euro Gastroenterol J 2016 Jun;4(3):413-422. [doi: 10.1177/2050640615600114] [Medline: 27403308]
8. Choung RS, Rey E, Locke GR III, et al. Chronic constipation and co‐morbidities: a prospective population‐based nested case‐control study. United Eur Gastroenterol J 2016 Feb;4(1):142-151. [doi: 10.1177/2050640614558476] [Medline: 26966534]
9. Aziz I, Palsson OS, Törnblom H, Sperber AD, Whitehead WE, Simrén M. The prevalence and impact of overlapping Rome IV-diagnosed functional gastrointestinal disorders on somatization, quality of life, and healthcare utilization: a cross-sectional general population study in three countries. Am J Gastroenterol 2018;113(1):86-96. [doi: 10.1038/ajg.2017.421]
10. Ohlsson B. Extraintestinal manifestations in irritable bowel syndrome: a systematic review. Therap Adv Gastroenterol 2022;15:17562848221114558. [doi: 10.1177/17562848221114558] [Medline: 35967918]
11. Eating and activity guidelines for New Zealand adults: updated 2020. : Wellington: Ministry of Health: Ministry of Health; 2020 URL: https://www.tewhatuora.govt.nz/assets/For-the-health-sector/Health-sector-guidance/Active-Families/eating-activity-guidelines-new-zealand-adults-updated-2020-oct22.pdf [accessed 2025-12-09]
12. Constipation. Constipation Auckland, New Zealand Health Te Whatu Ora. 2022. URL: https://www.tewhatuora.govt.nz/for-health-professionals/nutrition-education-resources [accessed 2025-12-09]
13. McKenzie YA, Bowyer RK, Leach H, et al. British Dietetic Association systematic review and evidence‐based practice guidelines for the dietary management of irritable bowel syndrome in adults (2016 update). J Human Nutrition Diet 2016 Oct;29(5):549-575. [doi: 10.1111/jhn.12385]
14. Bellini M, Tonarelli S, Barracca F, et al. Chronic constipation: is a nutritional approach reasonable? Nutrients 2021;13(10):3386. [doi: 10.3390/nu13103386]
15. Richardson DP, Ansell J, Drummond LN. The nutritional and health attributes of kiwifruit: a review. Eur J Nutr 2018 Dec;57(8):2659-2676. [doi: 10.1007/s00394-018-1627-z] [Medline: 29470689]
16. Rush EC, Patel M, Plank LD, Ferguson LR. Kiwifruit promotes laxation in the elderly. Asia Pac J Clin Nutr 2002;11(2):164-168. [doi: 10.1046/j.1440-6047.2002.00287.x] [Medline: 12074185]
17. Chan AOO. Increasing dietary fiber intake in terms of kiwifruit improves constipation in Chinese patients. WJG 2007;13(35):4771. [doi: 10.3748/wjg.v13.i35.4771]
18. Chang CC, Lin YT, Lu YT, Liu YS, Liu JF. Kiwifruit improves bowel function in patients with irritable bowel syndrome with constipation. Asia Pac J Clin Nutr 2010;19(4):451-457. [Medline: 21147704]
19. Caballero N, Benslaiman B, Ansell J, Serra J. The effect of green kiwifruit on gas transit and tolerance in healthy humans. Neurogastroenterol Motil 2020 Sep;32(9):e13874. [doi: 10.1111/nmo.13874] [Medline: 32431019]
20. Chey SW, Chey WD, Jackson K, Eswaran S. Exploratory comparative effectiveness trial of green kiwifruit, psyllium, or prunes in US patients with chronic constipation. Am J Gastroenterol 2021 Jun 1;116(6):1304-1312. [doi: 10.14309/ajg.0000000000001149] [Medline: 34074830]
21. Gearry R, Fukudo S, Barbara G, et al. Consumption of 2 green kiwifruits daily improves constipation and abdominal comfort-results of an international multicenter randomized controlled trial. Am J Gastroenterol 2023 Jun 1;118(6):1058-1068. [doi: 10.14309/ajg.0000000000002124] [Medline: 36537785]
22. Ishida F, Hu A, Yamaguchi T, Naraoka Y, Kobayashi H. The effects of green kiwifruit ingestion on digestive health, blood flow, skin health, and the autonomic nervous system. Health (London) 2021;13(6):647-659. [doi: 10.4236/health.2021.136049]

XSL•FO
RenderX

23. Gammon CS, Kruger R, Conlon CA, von Hurst PR, Jones B, Stonehouse W. Inflammatory status modulates plasma lipid and inflammatory marker responses to kiwifruit consumption in hypercholesterolaemic men. Nutr Metab Cardiovasc Dis 2014 Jan;24(1):91-99. [doi: 10.1016/j.numecd.2013.06.001]
24. Karlsen A, Svendsen M, Seljeflot I, et al. Kiwifruit decreases blood pressure and whole-blood platelet aggregation in male smokers. J Hum Hypertens 2013 Feb;27(2):126-130. [doi: 10.1038/jhh.2011.116]
25. Svendsen M, Tonstad S, Heggen E, et al. The effect of kiwifruit consumption on blood pressure in subjects with moderately elevated blood pressure: a randomized, controlled study. Blood Press 2015 Feb;24(1):48-54. [doi: 10.3109/08037051.2014.976979] [Medline: 25483553]
26. Carnachan SM, Bootten TJ, Mishra S, Monro JA, Sims IM. Effects of simulated digestion in vitro on cell wall polysaccharides from kiwifruit (Actinidia spp.). Food Chem 2012 Jul;133(1):132-139. [doi: 10.1016/j.foodchem.2011.12.084]
27. Monro JA, Mishra S, Venn B. Baselines representing blood glucose clearance improve in vitro prediction of the glycaemic impact of customarily consumed food quantities. Br J Nutr 2010 Jan;103(2):295-305. [doi: 10.1017/S0007114509991632] [Medline: 19930760]
28. Mishra S, Monro J. Kiwifruit remnants from digestion in vitro have functional attributes of potential importance to health. Food Chem 2012 Dec 15;135(4):2188-2194. [doi: 10.1016/j.foodchem.2012.06.102] [Medline: 22980789]
29. Wilkinson-Smith V, Dellschaft N, Ansell J, et al. Mechanisms underlying effects of kiwifruit on intestinal function shown by MRI in healthy volunteers. Aliment Pharmacol Ther 2019 Mar;49(6):759-768. [doi: 10.1111/apt.15127] [Medline: 30706488]
30. Sims IM, Monro JA. Fiber: composition, structures, and functional properties. Adv Food Nutr Res 2013;68:81-99. [doi: 10.1016/B978-0-12-394294-4.00005-5] [Medline: 23394983]
31. Kaur L, Rutherfurd SM, Moughan PJ, Drummond L, Boland MJ. Actinidin enhances gastric protein digestion as assessed using an in vitro gastric digestion model. J Agric Food Chem 2010 Apr 28;58(8):5068-5073. [doi: 10.1021/jf903332a]
32. Kaur L, Rutherfurd SM, Moughan PJ, Drummond L, Boland MJ. Actinidin enhances protein digestion in the small intestine as assessed using an in vitro digestion model. J Agric Food Chem 2010 Apr 28;58(8):5074-5080. [doi: 10.1021/jf903835g]
33. Montoya CA, Hindmarsh JP, Gonzalez L, Boland MJ, Moughan PJ, Rutherfurd SM. Dietary actinidin from kiwifruit (Actinidia deliciosa cv. Hayward) increases gastric digestion and the gastric emptying rate of several dietary proteins in growing rats. J Nutr 2014 Apr;144(4):440-446. [doi: 10.3945/jn.113.185744] [Medline: 24431326]
34. Rutherfurd SM, Montoya CA, Zou ML, Moughan PJ, Drummond LN, Boland MJ. Effect of actinidin from kiwifruit (Actinidia deliciosa cv. Hayward) on the digestion of food proteins determined in the growing rat. Food Chem 2011 Dec;129(4):1681-1689. [doi: 10.1016/j.foodchem.2011.06.031]
35. Montoya CA, Rutherfurd SM, Olson TD, et al. Actinidin from kiwifruit (Actinidia deliciosa cv. Hayward) increases the digestion and rate of gastric emptying of meat proteins in the growing pig. Br J Nutr 2014 Mar 28;111(6):957-967. [doi: 10.1017/S0007114513003401] [Medline: 24252432]
36. Lee YK, Low KY, Siah K, Drummond LM, Gwee KA. Kiwifruit (Actinidia deliciosa) changes intestinal microbial profile. Microb Ecol Health Dis 2012;23. [doi: 10.3402/mehd.v23i0.18572] [Medline: 23990838]
37. Saad RJ, Hasler WL. A technical review and clinical assessment of the wireless motility capsule. Gastroenterol Hepatol (N Y) 2011 Dec;7(12):795-804. [Medline: 22347818]
38. Moher D, Schulz KF, Altman DG, CONSORT. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. BMC Med Res Methodol 2001;1:2. [doi: 10.1186/1471-2288-1-2] [Medline: 11336663]
39. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
40. Koloski NA, Jones M, Talley NJ. Evidence that independent gut-to-brain and brain-to-gut pathways operate in the irritable bowel syndrome and functional dyspepsia: a 1-year population-based prospective study. Aliment Pharmacol Ther 2016 Sep;44(6):592-600. [doi: 10.1111/apt.13738] [Medline: 27444264]
41. Ware J, Ma K, Tuner-bowker D, Gandek B. Version 2 of the SF12 health survey. 2002 URL: https://www.center-tbi.eu/crf/Tree/SF12v2.pdf [accessed 2025-12-19]
42. Jensen J, Spittal M, Krishnan V. ELSI short form: user manual for a direct measure of living standards. : Ministry of Social Development Wellington, New Zealand; 2005 URL: http://msd.govt.nz/about-msd-and-our-work/publications-resources/monitoring/living-standards/elsi-short-form.html [accessed 2025-12-09]
43. Hofman DL, van Buul VJ, Brouns FJPH. Nutrition, health, and regulatory aspects of digestible maltodextrins. Crit Rev Food Sci Nutr 2016 Sep 9;56(12):2091-2100. [doi: 10.1080/10408398.2014.940415] [Medline: 25674937]
44. Efthymiou A, Mureanu N, Pemberton R, et al. Isolation and freezing of human peripheral blood mononuclear cells from pregnant patients. STAR Protoc 2022 Mar 18;3(1):101204. [doi: 10.1016/j.xpro.2022.101204] [Medline: 35252886]
45. Svedlund J, Sjödin I, Dotevall G. GSRS--a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci 1988 Feb;33(2):129-134. [doi: 10.1007/BF01535722] [Medline: 3123181]
46. McAuliffe MJ, Lalonde FM, McGarry D, Gandler W, Csaky K, Trus BL. Medical image processing, analysis and visualization in clinical research. Presented at: 14th IEEE Symposium on Computer-Based Medical Systems CBMS 2001; Jul 26-27, 2001. [doi: 10.1109/CBMS.2001.941749]

XSL•FO
RenderX

47. Dellschaft N, Murray K, Ren Y, et al. Assessing water content of the human colonic chyme using the MRI parameter T1: a key biomarker of colonic function. Neurogastroenterol Motil 2025 Apr;37(4):e14999. [doi: 10.1111/nmo.14999] [Medline: 39789944]
48. Asnicar F, Leeming ER, Dimidi E, et al. Blue poo: impact of gut transit time on the gut microbiome using a novel marker. Gut 2021 Sep;70(9):1665-1674. [doi: 10.1136/gutjnl-2020-323877]
49. Koloski NA, Jones M, Hammer J, et al. The validity of a new Structured Assessment of Gastrointestinal Symptoms Scale (SAGIS) for evaluating symptoms in the clinical setting. Dig Dis Sci 2017 Aug;62(8):1913-1922. [doi: 10.1007/s10620-017-4599-6]
50. Spiegel BMR, Hays RD, Bolus R, et al. Development of the NIH Patient-Reported Outcomes Measurement Information System (PROMIS) gastrointestinal symptom scales. Am J Gastroenterol 2014;109(11):1804-1814. [doi: 10.1038/ajg.2014.237]
51. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176. [doi: 10.1159/000376585] [Medline: 25831962]
52. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes 2007 Nov 27;5(1):63. [doi: 10.1186/1477-7525-5-63] [Medline: 18042300]
53. Jelsness-Jørgensen LP, Moum B, Grimstad T, et al. The Multidimensional Fatigue Inventory (MFI-20): psychometrical testing in a Norwegian sample of inflammatory bowel disease (IBD) patients. Scand J Gastroenterol 2022 Jun;57(6):683-689. [doi: 10.1080/00365521.2022.2029939] [Medline: 35076321]
54. Clover K, Lambert SD, Oldmeadow C, et al. PROMIS depression measures perform similarly to legacy measures relative to a structured diagnostic interview for depression in cancer patients. Qual Life Res 2018 May;27(5):1357-1367. [doi: 10.1007/s11136-018-1803-x] [Medline: 29423755]
55. Drossman DA, Patrick DL, Whitehead WE, et al. Further validation of the IBS-QOL: a disease-specific quality-of-life questionnaire. Am J Gastroenterol 2000 Apr;95(4):999-1007. [doi: 10.1111/j.1572-0241.2000.01941.x] [Medline: 10763950]
56. Dunn S, Datta A, Kallis S, Law E, Myers CE, Whelan K. Validation of a Food Frequency Questionnaire to measure intakes of inulin and oligofructose. Eur J Clin Nutr 2011 Mar;65(3):402-408. [doi: 10.1038/ejcn.2010.272] [Medline: 21179053]
57. Labant M. NGS sample prep: enduring basics and evolving enzyme kits. Genet Eng Biotechnol News 2023 Nov 1;43(11):54-56. [doi: 10.1089/gen.43.11.20]
58. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014 Aug 1;30(15):2114-2120. [doi: 10.1093/bioinformatics/btu170]
59. Genome. NCBI-NIH. URL: https://www.ncbi.nlm.nih.gov/datasets/genome/?taxon=9606 [accessed 2025-12-23]
60. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv. Preprint posted online on May 26, 2013. [doi: 10.48550/arXiv.1303.3997]
61. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. Bioinformatics 2009 Aug 15;25(16):2078-2079. [doi: 10.1093/bioinformatics/btp352]
62. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End Read Merger. Bioinformatics 2014 Mar 1;30(5):614-620. [doi: 10.1093/bioinformatics/btt593]
63. Bushnell B. BBMap: a fast, accurate, splice-aware aligner. 2014 Presented at: 9th Annual Genomics of Energy & Environment Meeting; Mar 17-20, 2014.
64. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods 2015 Jan;12(1):59-60. [doi: 10.1038/nmeth.3176] [Medline: 25402007]
65. NCBI Resource Coordinators. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res 2013;41(D1):D8-D20. [doi: 10.1093/nar/gks1189]
66. Huson DH, Beier S, Flade I, et al. MEGAN community edition - interactive exploration and analysis of large-scale microbiome sequencing data. PLoS Comput Biol 2016 Jun;12(6):e1004957. [doi: 10.1371/journal.pcbi.1004957] [Medline: 27327495]
67. Chen S, Hoene M, Li J, et al. Simultaneous extraction of metabolome and lipidome with methyl tert-butyl ether from a single small tissue sample for ultra-high performance liquid chromatography/mass spectrometry. J Chromatogr A 2013 Jul;1298:9-16. [doi: 10.1016/j.chroma.2013.05.019]
68. Abshirini M, Cabrera D, Fraser K, et al. Mass spectrometry-based metabolomic and lipidomic analysis of the effect of high fat/high sugar diet and GreenshellTM mussel feeding on plasma of ovariectomized rats. Metabolites 2021 Oct 31;11(11):754. [doi: 10.3390/metabo11110754] [Medline: 34822412]
69. Fraser K, Roy NC, Goumidi L, et al. Plasma biomarkers and identification of resilient metabolic disruptions in patients with venous thromboembolism using a metabolic systems approach. ATVB 2020 Oct;40(10):2527-2538. [doi: 10.1161/ATVBAHA.120.314480]
70. Huynh K, Barlow CK, Jayawardana KS, et al. High-throughput plasma lipidomics: detailed mapping of the associations with cardiometabolic risk factors. Cell Chem Biol 2019 Jan;26(1):71-84. [doi: 10.1016/j.chembiol.2018.10.008]
71. Xu J, Begley P, Church SJ, et al. Graded perturbations of metabolism in multiple regions of human brain in Alzheimer's disease: snapshot of a pervasive metabolic disorder. Biochim Biophys Acta (BBA) - Mol Basis Dis 2016 Jun;1862(6):1084-1092. [doi: 10.1016/j.bbadis.2016.03.001]

XSL•FO
RenderX

72. Tsugawa H, Cajka T, Kind T, et al. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat Methods 2015 Jun;12(6):523-526. [doi: 10.1038/nmeth.3393]
73. James SC, Fraser K, Young W, et al. Concentrations of fecal bile acids in participants with functional gut disorders and healthy controls. Metabolites 2021;11(9):612. [doi: 10.3390/metabo11090612]
74. Han J, Lin K, Sequeira C, Borchers CH. An isotope-labeled chemical derivatization method for the quantitation of short-chain fatty acids in human feces by liquid chromatography–tandem mass spectrometry. Anal Chim Acta 2015 Jan;854:86-94. [doi: 10.1016/j.aca.2014.11.015]
75. Cooney JM, Trower TM. MS-probes and stable isotope coding for targeted analysis of short chain fatty acids (SCFAs) in biological samples. Presented at: Metabolomics 2017 (13th Annual Conference of Metabolomics Society); Jun 25-29, 2017.
76. Fristedt R, Ruppert V, Trower T, Cooney J, Landberg R. Quantitation of circulating short-chain fatty acids in small volume blood samples from animals and humans. Talanta 2024 May;272:125743. [doi: 10.1016/j.talanta.2024.125743]
77. Lomiwes D, Günther CS, Bloor SJ, et al. Identification of sarmentosin as a key bioactive from blackcurrants (Ribes nigrum) for Inhibiting platelet monoamine oxidase in humans. J Agric Food Chem 2024 Jul 31;72(30):16777-16789. [doi: 10.1021/acs.jafc.4c03802]
78. Thwaites PA, Yao CK, Maggo J, et al. Comparison of gastrointestinal landmarks using the gas-sensing capsule and wireless motility capsule. Aliment Pharmacol Ther 2022 Nov;56(9):1337-1348. [doi: 10.1111/apt.17216] [Medline: 36082475]
79. Vilz TO, Pantelis D, Lingohr P, et al. SmartPill® as an objective parameter for determination of severity and duration of postoperative ileus: study protocol of a prospective, two-arm, open-label trial (the PIDuSA study). BMJ Open 2016 Jul 8;6(7):e011014. [doi: 10.1136/bmjopen-2015-011014] [Medline: 27401360]
80. Eady SL, Wallace AJ, Butts CA, et al. The effect of “Zesy002” kiwifruit (Actinidia chinensis var. chinensis) on gut health function: a randomised cross-over clinical trial. J Nutr Sci 2019;8:e18. [doi: 10.1017/jns.2019.14] [Medline: 31080591]
81. Eady SL, Wallace AJ, Hedderley DI, Bentley-Hewitt KL, Butts CA. The effects on immune function and digestive health of consuming the skin and flesh of Zespri® SunGold kiwifruit (Actinidia Chinensis var. Chinensis ’Zesy002’) in healthy and IBS-constipated individuals. Nutrients 2020 May 18;12(5):1453. [doi: 10.3390/nu12051453] [Medline: 32443433]
82. Bayer SB, Heenan P, Frampton C, et al. Two gold kiwifruit daily for effective treatment of constipation in adults-a randomized clinical trial. Nutrients 2022 Oct 6;14(19):4146. [doi: 10.3390/nu14194146] [Medline: 36235798]
83. Committee for Proprietary Medicinal Products (CPMP). Points to consider on switching between superiority and non‐inferiority. Brit J Clinical Pharma 2001 Sep;52(3):223-228. [doi: 10.1046/j.1365-2125.2001.01397-3.x]
84. Eriksson L, Trygg J, Wold S. CV‐ANOVA for significance testing of PLS and OPLS® models. J Chemom 2008 Nov;22(11-12):594-600. [doi: 10.1002/cem.1187]
85. Pang Z, Chong J, Zhou G, et al. MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights. Nucleic Acids Res 2021 Jul 2;49(W1):W388-W396. [doi: 10.1093/nar/gkab382] [Medline: 34019663]
86. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
87. El-Salhy M, Ystad SO, Mazzawi T, Gundersen D. Dietary fiber in irritable bowel syndrome (review). Int J Mol Med 2017 Sep;40(3):607-613. [doi: 10.3892/ijmm.2017.3072] [Medline: 28731144]
88. Harahsheh Y, Ho KM. Optimizing the utility of secondary outcomes in randomized controlled trials. J Emerg Crit Care Med 2020;5:8-8. [doi: 10.21037/jeccm-20-136]
89. Roydhouse JK, Fiero MH, Kluetz PG. Investigating potential bias in patient-reported outcomes in open-label cancer trials. JAMA Oncol 2019 Apr 1;5(4):457. [doi: 10.1001/jamaoncol.2018.6205]
90. Higgins KM, Levin G, Busch R. Considerations for open-label randomized clinical trials: design, conduct, and analysis. Clin Trials 2024 Dec;21(6):681-688. [doi: 10.1177/17407745241244788] [Medline: 38618711]
91. Parr-Brownlie LC, Waters DL, Neville S, Neha T, Muramatsu N. Aging in New Zealand: ka haere ki te ao pakeketanga. Gerontologist 2020 Jul 15;60(5):812-820. [doi: 10.1093/geront/gnaa032] [Medline: 32359062]
92. Sander M, Oxlund B, Jespersen A, et al. The challenges of human population ageing. Age Ageing 2015 Mar;44(2):185-187. [doi: 10.1093/ageing/afu189] [Medline: 25452294]
93. Bhutto A, Morley JE. The clinical significance of gastrointestinal changes with aging. Curr Opin Clin Nutr Metab Care 2008 Sep;11(5):651-660. [doi: 10.1097/MCO.0b013e32830b5d37] [Medline: 18685464]
94. O’Brien L, Wall C, Wilkinson TJ, Gearry RB. Chronic diarrhoea in older adults and the role of dietary interventions. Nutr Healthy Aging 2022 Oct 6;7(1-2):39-50. [doi: 10.3233/NHA-220152]

Abbreviations

Atmo capsule: Atmo gas-sensing capsule
CONSORT: Consolidated Standards of Reporting Trials
EDTA: ehylenediaminetetraacetic acid
FC: functional constipation
FDR: false discovery rate
FFQ: Food Frequency Questionnaire

XSL•FO
RenderX

GSRS: Gastrointestinal Symptom Rating Scale
HPLC: high-pressure liquid chromatography
IBS-C: irritable bowel syndrome with constipation
LCMS: liquid chromatography-mass spectrometry
MRI: magnetic resonance imaging
MS: mass spectrometry
PBMC: peripheral blood mononuclear cell
PLS-DA: partial least squares projection-discriminant analysis
PROMIS: Patient-Reported Outcomes Measurement Information System
REDCap: Research Electronic Data Capture
ROI: region of interest
SF-12v2: Short-Form Health Survey Version 2 (12-item)
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TRUFI: True Fast Imaging With Steady-State Free Precession
WHO-5: World Health Organization - Five Question Well-Being Index
WMD: wireless motility device

Edited by J Sarvestan; submitted 31.Mar.2025; peer-reviewed by S Nandurkar; revised version received 19.Nov.2025; accepted 20.Nov.2025; published 13.Jan.2026.

Please cite as:
Maggo J, Ng HM, Bayer SB, Wall CL, Hoad CL, Marciani L, Mullaney J, Cabrera D, Fraser K, Cooney JM, Günther CS, Trower T, Tang J, Gasser O, Milan A, McNabb WC, Spiller R, Conner T, Frampton C, Foster M, Roy NC, Gearry RB
Normalization of Gastrointestinal Symptoms in Adults With Constipation With Daily Green Kiwifruit Consumption: Protocol for an Open-Label Intervention Study
JMIR Res Protoc 2026;15:e75286
URL: https://www.researchprotocols.org/2026/1/e75286
doi:10.2196/75286

© Jasjot Maggo, Hwei Min Ng, Simone Birgit Bayer, Catherine L Wall, Caroline L Hoad, Luca Marciani, Jane Mullaney, Diana Cabrera, Karl Fraser, Janine Cooney, Catrin S Günther, Tania Trower, Jeffery Tang, Oliver Gasser, Amber Milan, Warren C McNabb, Robin Spiller, Tamlin Conner, Chris Frampton, Meika Foster, Nicole C Roy, Richard B Gearry. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Curative Effect and Autonomic Nerve Function of Patients With Primary Insomnia of Liver Depression and Spleen Deficiency Type Based on the Acupoint Selection of Meridian Theory: Protocol for a Randomized Controlled Trial

Su Fu, BA; Kang He*, BA; Chuanlong Zhou*, MA

Department of Acupuncture and Moxibustion, The Third Affiliated Hospital of Zhejiang Chinese Medical University, 219 Moganshan Road, Hangzhou, Zhejiang, China

*these authors contributed equally

Corresponding Author:
Chuanlong Zhou, MA
Department of Acupuncture and Moxibustion, The Third Affiliated Hospital of Zhejiang Chinese Medical University, 219 Moganshan Road, Hangzhou, Zhejiang, China

Abstract

Background: Primary insomnia often corresponds to the syndrome of liver depression and spleen deficiency in traditional Chinese medicine. This study evaluates a new acupuncture protocol derived from the "Jing Bie" theory. Tianrong (TE16) and Tianyou (SI17) were selected to regulate the Shaoyang meridian, soothe the liver, regulate qi (the basic life energy or life force to maintain the physiological function of the body according to traditional Chinese medicine), strengthen the spleen, and tranquilize the mind, so as to restore sleep.

Objective: The objective of this study is to evaluate the clinical efficacy of this regimen and to explore its effects on autonomic nervous system function and central nervous system chemistry.

Methods: This is a single-blind randomized controlled trial. A total of 96 patients with primary insomnia (liver depression and spleen deficiency pattern) will be recruited and randomly assigned to the treatment group or the control group. At the same time, 48 healthy volunteers will be recruited as the healthy control group.

Results: This study was funded in November 2023. Recruitment and data collection began in January 2024 and are currently underway. As of December 2025, a total of 40 participants have been enrolled, of whom 4 have withdrawn. Recruitment is projected to conclude by December 2026. Data analysis will be performed after the completion of recruitment. The results are expected to be published in summer 2027.

Conclusions: This study integrates a traditional Chinese medicine framework with modern physiological measurements. The aim is to provide evidence for targeted acupuncture strategies by linking clinical improvement to autonomic nervous system balance and neurochemical changes and to elucidate potential mechanisms to provide a nondrug treatment option for insomnia.

Trial Registration: ClinicalTrials.gov NCT06782815; https://clinicaltrials.gov/study/NCT06782815

International Registered Report Identifier (IRRID): DERR1-10.2196/84122

(JMIR Res Protoc 2026;15:e84122) doi:10.2196/84122

KEYWORDS

liver stagnation and spleen deficiency insomnia; acupuncture; randomized controlled trial; autonomic nerve function; Jing Bie theory; traditional Chinese medicine

Introduction

Insomnia is a subjective experience in which patients are not satisfied with the length and/or quality of sleep, and it affects their daytime social function [1]. Long-term insomnia can lead to fatigue, mood swings, and cognitive impairment [2]. Studies have shown that acupuncture is effective in treating insomnia [3]. "Jing Bie" is a longitudinal branch of the 12 meridians, which goes deep into the body cavity to strengthen the connection between the 2 meridians in the deep part of the trunk. The theory of Jing Bie is recorded in "Lingshu · Jing Bie," and the distribution characteristics of Jing Bie are "leaving, entering, leaving and combining." Six acupoints, such as Tianrong (TE16) and Tianyou (SI17), are located in the neck [4]. In the theory of traditional Chinese medicine (TCM), Jing Bie not only strengthens the close relationship between the limbs and the viscera, the head, and the 5 sense organs but also plays a role in connecting the 2 meridians, reconciling the "ying" and "wei,"

XSL•FO
RenderX

and promoting the balance of qi (the basic life energy or life force to maintain the physiological function of the body) and blood in the head [5]. From the perspective of modern medical anatomy, the 6 acupoints on the neck are mostly located near the sternocleidomastoid muscle, where the superior cervical ganglion and stellate ganglion of the sympathetic trunk are deeply distributed. Intervention in these parts can regulate sympathetic and parasympathetic nerve function and relieve symptoms related to autonomic nervous system (ANS) disorders. For example, stellate ganglion block (SGB) is often used clinically to treat sympathetic-parasympathetic imbalances such as migraine, cluster headache, and insomnia [6,7]. SGB can inhibit central and peripheral activity; jointly control the glands, blood vessels, and muscle movement in the region; and regulate sympathetic nerve activity from a pathologically hyperactive state to normal levels and maintain stability. Repeated SGB can interfere with autonomic nerve activity, thereby modifying autonomic nerve function [8,9]. Therefore, the acupoints selected based on the Jing Bie theory may also have a good effect on autonomic nerve regulation because they are anatomically adjacent to the stellate ganglion.

A study that analyzed the effects of acupuncture on the ANS showed that acupuncture can reduce sympathetic nerve activity under stress conditions and balance the ANS [10]. In a review that included 14 studies on the effects of acupuncture on heart rate variability (HRV), it was reported that acupuncture can reduce the low-frequency power (LF) and the LF to high-frequency power (HF) ratio (LF/HF) in patients and may function to balance the sympathetic and parasympathetic nervous systems [11]. A previous study by the project leader found that acupuncture at the acupoints Tianrong (TE16) and Fengchi (GB20) can improve the autonomic nerve function of patients with migraine. Although studies have shown that acupuncture has a regulatory effect on autonomic nerves, there is still a lack of in-depth discussion on the mechanism [12]. At present, most studies on acupuncture treatment of insomnia do not strictly distinguish TCM syndromes, and it is difficult to accurately provide effective treatment options for specific syndromes. This study focuses on insomnia of liver depression and spleen deficiency type. It is expected to clarify the curative effect of acupuncture on insomnia of this type through rigorous randomized controlled trials (RCTs) and provide more targeted evidence for clinical treatment with TCM.

HRV reflects the functional status of sympathetic and parasympathetic nerves, which offers a window into the functional status of human ANS [13]. Therefore, it is believed that HRV reflects the mechanism of excessive arousal from the changes in autonomic nerve function, and improving HRV indices and reducing excessive arousal may be a new idea for the treatment of insomnia. We designed this study based on the theory of Jing Bie selection to explore the treatment of insomnia through the analysis of HRV.

In “Huangdi Neijing,” there is a discussion on the disharmony between ying and wei in insomnia. According to the theory of branches of 12 meridians, selection of acupoints on the neck can regulate the qi of the 2 meridians; the theory also explains the principle of ANS intervention from the perspective of modern medicine. Acupuncture intervention can regulate autonomic nerve function. In a preliminary study on autonomic nerve intervention in migraine, we found that acupuncture at neck acupoints has a certain therapeutic effect. In modern medical research, HRV, which reflects autonomic nerve function, can well reflect the clinical situation of patients with primary insomnia-hyperarousal, and it is believed that primary insomnia can be treated by improving HRV indices. Therefore, in this study, we aim to examine whether, according to the theory of meridian acupoint intervention, the curative effect of the treatment of primary insomnia can be improved through autonomic nerve intervention.

Methods

Study Setting and Design

This study is a single-center, prospective, single-blind RCT. A total of 96 patients with primary insomnia of liver depression and spleen deficiency type will be recruited from the Department of Acupuncture and Moxibustion of the Third Affiliated Hospital of Zhejiang Chinese Medical University in Zhejiang province, China. Participants will be randomly assigned (1:1) to the acupuncture group based on the theory of Jing Bie selection or the conventional acupuncture control group. The intervention will be carried out for 4 weeks, followed by follow-up.

The specific follow-up arrangement will be evaluated in the fourth week of the treatment period, and the patients’ recovery and treatment efficacy will be assessed in depth at key follow-up time points. In both groups, the treatment will be administered 3 times a week for 4 weeks. At the same time, 48 healthy volunteers without sleep disorders will be included as the healthy control group without any intervention measures, which will be used to compare and analyze the changes in relevant indicators to evaluate the effect of acupuncture treatment more comprehensively. The flowchart of the study process is shown in Figure 1, and the trial schedule of enrollment, treatments, and assessments is displayed in Table 1. The reporting of this study protocol conforms to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines (Checklist 1).

XSL•FO
RenderX

Figure 1. Study flowchart. DBAS-16: 16-item Dysfunctional Beliefs and Attitudes About Sleep scale, GABA: γ-aminobutyric acid, Glu: glutamic acid; HRV: heart rate variability; ISI: Insomnia Severity Index; MRI: magnetic resonance imaging; PSG: polysomnography; PSQI: Pittsburgh Sleep Quality Index; SF-36: 36-item Short Form Health Survey.

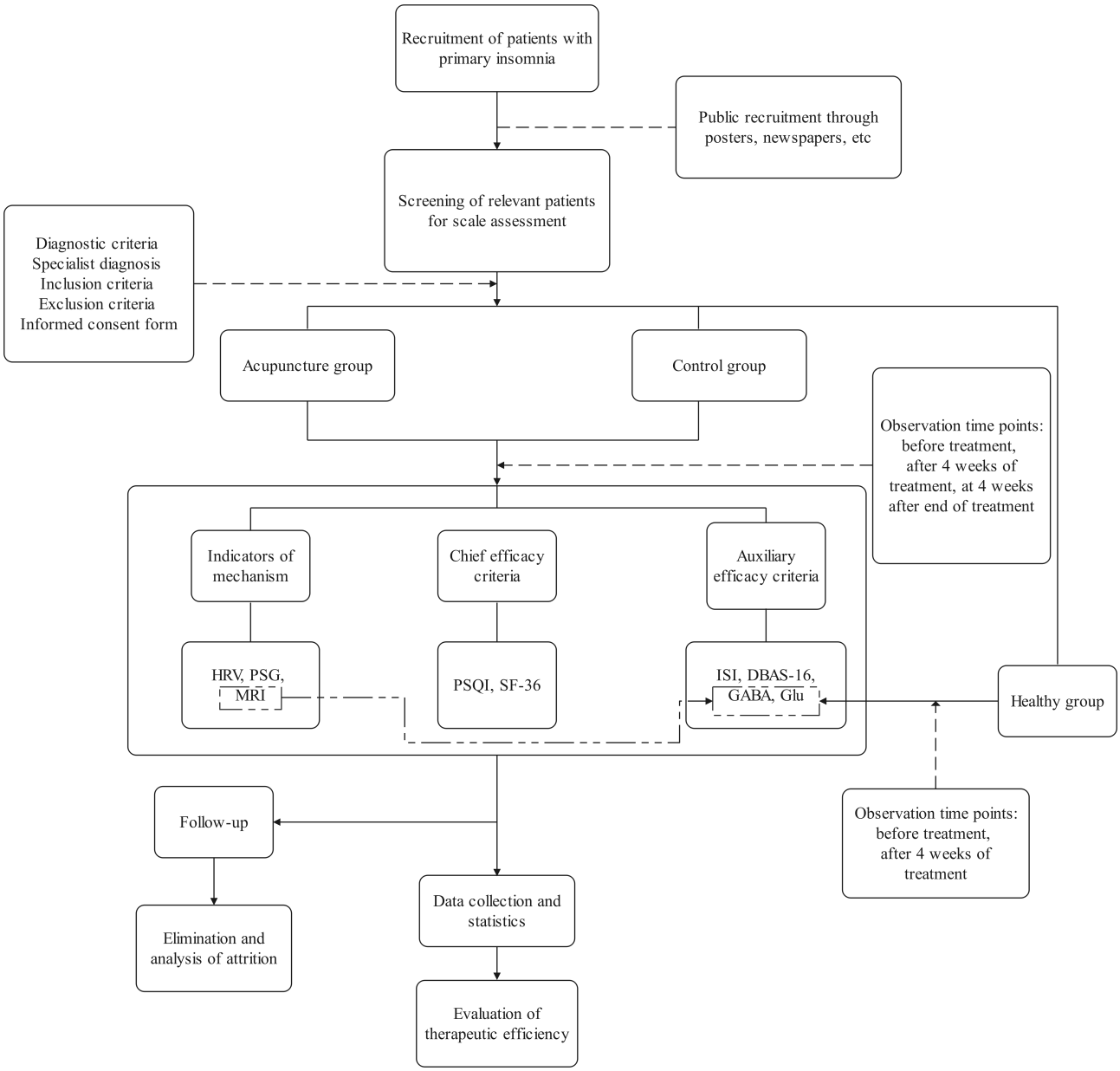

XSL•FO
RenderX

Table . Schedule of study assessments and outcome measures.

Research stage	Before treatment	Treatment stage	Follow-up stage
Assessment	First	Second	Third
Time (weeks)	0	4	8
Informed consent	✓	—[a]	—
Inclusion/exclusion criteria	✓	—	—
Demographic data	✓	—	—
Vital signs	✓	—	—
Medical history record	✓	—	—
Physical examination	✓	—	—
Medication situation	✓	—	—
Indicator of mechanism			
HRV[b]	✓	✓	✓
PSG[c]	✓	✓	✓
MRI[d]	✓	✓	—
Primary outcome indicator			
PSQI[e]	✓	✓	✓
SF-36[f]	✓	✓	✓
Secondary outcome indicator			
ISI[g]	✓	✓	✓
DBAS-16[h]	✓	✓	✓
GABA[i]	✓	✓	—
Glu[j]	✓	✓	—
Safety evaluation	✓	✓	✓

[a]Not applicable.

[b]HRV: heart rate variability.

[c]PSG: polysomnography.

[d]MRI: magnetic resonance imaging.

[e]PSQI: Pittsburgh Sleep Quality Index.

[f]SF-36: 36-item Short Form Health Survey.

[g]ISI: Insomnia Severity Index.

[h]DBAS-16: 16-item Dysfunctional Beliefs and Attitudes About Sleep scale.

[i]GABA: γ-aminobutyric acid.

[j]Glu: glutamic acid.

Participant Recruitment

Participants in this study will be recruited from the outpatient Department of Acupuncture and Moxibustion of our hospital, through posters placed on the hospital's announcement board and other ways. The recruitment information will detail the research content, purpose, process, and corresponding contact information.

The participants will be asked to first contact the research assistant for preliminary screening. Participants who may be eligible will be registered, and a specialist will be appointed to check whether they meet the inclusion and exclusion criteria of this study (Textbox 1). Only patients who meet all the inclusion criteria and do not meet any exclusion criteria will be invited to participate in the study, and they will be given a detailed introduction to the specifics of this RCT. All participants are required to sign informed consent forms and provide comprehensive personal information.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Patients with symptoms in line with the diagnostic criteria of insomnia of liver depression and spleen deficiency type in traditional Chinese medicine
- Patients with symptoms in line with the western diagnostic criteria for insomnia [14]
- Male or female patients aged 18 to 40 years
- Patients with a disease course of more than 2 weeks

Exclusion criteria (patients who meet any one of the following criteria will be excluded)

- Patients with insomnia caused by a variety of organic diseases or neurological or psychiatric disorders
- Pregnant or lactating patients or those with a drug allergy
- Patients with more serious primary diseases such as cardiovascular and cerebrovascular diseases, gastrointestinal diseases, endocrine system diseases, or psychiatric disorders
- Patients with coagulation dysfunction and bleeding tendency

The baseline assessment will be performed by the same physician, and the results will be recorded in detail in the case report form. All participants will receive treatment for 4 weeks and undergo follow-up observation after treatment. The evaluation will be done at 3 time points: before treatment, after 4 weeks of treatment, and at 4 weeks after end of treatment. Each participant will receive a schedule of treatment dates and a schedule of follow-up appointments developed by research assistants who will be blinded to the grouping information.

In addition, all participants will be required to keep a sleep diary throughout the treatment and follow-up periods, recording the date of each insomnia episode, the length of time required to fall asleep, the number of times they wake up at night, the difficulty of falling asleep again, and the subjective feelings after waking up in the morning. Moreover, they will be asked to include sleep-related details as well as the name, dose, and time of use of other drugs (such as sleeping pills) during the trial and their own feelings after taking the drug.

Ethical Considerations

The study was approved by the Medical Ethics Committee of the Third Affiliated Hospital of Zhejiang Chinese Medical University (approval no. ZSLL-KY-2024-032-01). If the main protocol needs to be modified, all members of the research team will participate in the process and submit the final modified protocol to the Medical Ethics Committee for review. This approval ensures that the study adheres to ethical standards and protects the rights and safety of participants. Patients and/or the public were involved in the design, conduct, reporting, or dissemination plans of this research. All participants are required to provide written informed consent to participate in this study. The consent process explicitly emphasizes the voluntary nature of participation and the right to withdraw at any time without penalty. To ensure privacy, all study data will be deidentified for analysis, with personal information stored separately under strict access controls. No financial compensation will be provided to participants.

Randomization and Blinding

This is a single-blind RCT. In total, 96 patients with insomnia who meet the inclusion criteria will be randomly assigned (1:1) to the acupuncture group and control group. Statisticians who are blinded to the participants' recruitment or intervention process will use the block randomization method to set the block size to 4 and generate a random distribution sequence. The assigned sequence will be hidden using serially numbered, opaque, sealed envelopes. After a participant has provided informed consent and completed the baseline assessment, the therapist will open the next serially numbered envelope in the presence of the participant to reveal the group assignment. Given the nature of acupuncture interventions, acupuncturists who perform the treatment cannot be blinded to group assignment. To ensure the integrity of the single-blind design, the following steps will be taken. Participants will not be informed of their specific group assignment but will be informed that they are receiving 1 of the 2 active acupuncture treatments for insomnia. Outcome evaluators and statisticians who perform data analysis will be blinded to group assignment. All data collection forms and electronic databases will use only participant identification codes and will not indicate group assignment. After all analyses have been completed, the allocation list will be securely maintained by the primary investigator until the database is locked. A total of 48 healthy volunteers without sleep disorders will be included as a nonrandomized observation control group without any intervention. The purpose of this group is to establish standardized physiological and questionnaire-based baseline values, which will be compared with the data of the patient groups. Given the nonrandomized allocation of healthy controls, statistical comparisons between patients and healthy controls would be considered to be observational in nature. To address the potential confounding factors, in appropriate cases, covariate adjustment and sensitivity analysis will be used to enhance causal inference to explain the differences between the patient groups and the healthy control group.

To ensure successful blinding, acupuncture will be performed for patients from both groups using several common approaches: skin sterilization with 75% alcohol; usage of disposable and sterile steel needles (1.5-inch needle, 0.25×40 mm); and

XSL•FO
RenderX

insertion of needles into the skin to a depth of 10 to 15 mm, left for 30 minutes. Acupuncture treatment will be performed 3 times a week, and a total of 12 acupuncture or conventional acupuncture operations will be performed in 4 weeks. In addition, the external conditions such as the nursing process and the environment during the treatment will be kept consistent for the 2 groups of patients, so as to minimize the interference of additional factors on the results of the study.

Interventions

Overview

All participants will receive acupuncture based on the Jing Bie theory or conventional acupuncture at a frequency of 3 times a week. To ensure the accuracy and consistency of the treatment administered, all acupuncture physicians participating in this study are required to have passed the clinical standard operation training in this subject and have the corresponding professional qualifications. Moreover, any possible adverse reactions during the study will be recorded in detail.

During the study period, the use of any prescription or nonprescription drugs for insomnia will be prohibited in principle. Participants who are required to use drugs for other diseases, which may affect the central nervous system, or have a severe insomnia problem will be processed and recorded as a “significant program deviation” or “withdrawal from the study” event. For those who do need medication, their complete medication information will be included in the subsequent statistical analysis as a key covariate. All participants are required to maintain a relatively stable work and rest schedule during the study period and to record information such as their bedtime, wake-up time, and occurrence of night awakening in a daily sleep diary. The researchers will review the diary every week to identify and guide individuals who experience excessive daytime sleepiness (>1 h) or significant changes in work and rest schedule (>2 h). Such behavioral changes will be included as covariates in the analysis.

Acupuncture Group

This group will include 48 patients. The patients will be placed in a supine position. Acupuncture will be performed bilaterally at Tianrong (SI17) points, with the needle inserted backward obliquely at 45±5° to a depth of 1.0 to 1.5 inches, reinforcing and reducing horizontally, and the needle insertion will be stopped when the patient feels a heat sensation in the auricle.

Acupuncture will be performed bilaterally at Tianyou (TE16) points, with the needle pointing toward the tip of the nose. The needle will be inserted to a depth of 0.5 to 1.0 inch using flat reinforcing and reducing maneuvers, until the sense of soreness and distension obviously stops; the needle will be retained for 30 minutes. The location of acupuncture points is shown in Table 2.

Table . Acupoint selection and stimulation parameters for the acupuncture group.

Acupoint	Location	Manipulation
Tianyou (TE16)	On the lateral side of the neck, located in the depression at the posterior border of the sternocleidomastoid muscle, level with the mandibular angle.	The needle is directed toward the tip of the nose, the insertion depth is 0.5 - 1.0 inch, the reinforcing and reducing maneuvers are flat, needle insertion is stopped when soreness is obvious, and the needle is retained for 30 min.
Tianrong (SI17)	In the lateral aspect of the neck and in the depression on the anterior border of the sternocleidomastoid muscle and posterior to the mandibular angle	The needle is inserted backward obliquely at 45±5° to a depth of 1.0 - 1.5 inches, with reinforcing and reducing maneuvers, and needle insertion is stopped when the patient feels a heat sensation in the auricle.

Control Group

This group will include 48 patients. The patients will be placed in a supine position. Acupuncture will be performed at Baihui (DU20), Sishencong (EX-HN1), bilateral Qimen (LR13), Zhangmen (LR14), Hegu (LR14), and Taichong (LR3), with reinforcing and reducing maneuvers. The needle insertion will be stopped when acid swelling is obvious, and the needle will be retained for 30 min. The location of acupuncture points is shown in Textbox 2.

XSL•FO
RenderX

Textbox 2. Acupoint selection and stimulation parameters for the control group.

Acupoints and location

- Baihui (DU20): on the head, 5 inches straight up the hairline or the midpoint of the line between the 2 ear tips
- Sishencong (EX-HN1): on the top of the head, a total of 4 points located 1 inch anterior, posterior, and bilateral to the Baihui
- Qimen (LR13): on the chest, in the sixth intercostal space, approximately 4 inches lateral to the anterior median line (midline) and level with the nipple.
- Zhangmen (LR14): on the lateral abdomen, below the free extremity of the 11th rib
- Hegu (LI4): on the dorsum of the hand, between the first and second metacarpals, at the midpoint of the radial side of the second metacarpal
- Taichong (LR3): on the dorsal side of the foot, in the posterior depression of the first metatarsal space

Manipulation

- Tonifying and reducing, until the soreness and distension obviously stops; the needle is retained for 30 minutes

Healthy Group

This group will include 48 healthy volunteers without sleep disorders, who will not undergo any treatment.

Outcome Measures

Primary Outcome Measures

Sleep Quality Assessed Using the Pittsburgh Sleep Quality Index

The patients will be assessed using the Chinese version of the Pittsburgh Sleep Quality Index (PSQI) questionnaire at each observation point, and the clinical efficacy will be assessed. The questionnaire consists of 19 self-evaluation items and 5 other evaluation items. Only the 19 self-evaluation questions will be scored, which comprise 7 factors, each scored on a scale of 0 to 3 points (0: no difficulty, 3: serious difficulty). The sum of all factor scores ranges from 0 to 21, with a total score of 21 indicating serious difficulty in all respects. The reference values for adult sleep quality problems in China are as follows: PSQI ≥8 indicates poor sleep quality, whereas PSQI ≤7 indicates good sleep quality; the higher the score, the worse the sleep quality.

Quality of Life Assessed Using the 36-Item Short Form Health Survey

The patients will be evaluated using the 36-item Short Form Health Survey (SF-36) at each observation point. The SF-36 is a general scale developed by the American Medical Research Group to evaluate the quality of life. It is widely recognized and used worldwide. This scale has 8 dimensions to evaluate health-related quality of life, which belong to 2 categories of physical health and mental health, namely physiological function, role-physical, bodily pain, general health, vitality, social function, role-emotional, and mental health. In addition, the SF-36 scale includes another indicator, health change, which is used to evaluate health changes over the past year.

Secondary Outcome Measures

Insomnia Severity Index

The patients will be evaluated using the Insomnia Severity Index (ISI) at each observation point. The degree of insomnia is determined according to the total score: a score between 0 and 7 indicates no clinically significant insomnia, a score between 8 and 14 indicates subthreshold insomnia, a score between 15 and 21 indicates clinical insomnia (moderate to severe), and a score between 22 and 28 indicates clinical insomnia (severe).

16-Item Dysfunctional Beliefs and Attitudes About Sleep Scale Score

The items in the 16-item Dysfunctional Beliefs and Attitudes About Sleep scale (DBAS-16) comprise 4 factors: potential consequences of insomnia, worry about sleep, sleep expectations, and cognition of insomnia drugs. The lower the total score, the more likely it is that the participants' cognition of sleep is biased.

γ-Aminobutyric Acid

γ-Aminobutyric acid (GABA) is used as a secondary outcome indicator for insomnia treatment, mainly based on its physiological role as a central inhibitory neurotransmitter. Studies have shown that the hyperarousal state of patients with insomnia is closely related to a decrease in GABA levels in brain regions such as the anterior cingulate gyrus and occipital cortex, and GABA concentration is negatively correlated with objective sleep indicators (such as awake time during sleep). Magnetic resonance spectroscopy can noninvasively and quantitatively determine the concentration of GABA in specific brain regions in vivo, thereby directly assessing whether the treatment improves sleep by enhancing the key mechanism of inhibitory neurotransmission.

Glutamic Acid

Glutamic acid (Glu) will be used as a secondary outcome indicator due to its key role in maintaining cortical excitability. Insomnia is often accompanied by excessive activation of the brain, which is related to abnormal activity of the glutamatergic system. Studies have shown that changes in the Glu levels in specific brain regions of patients with insomnia are associated with subjective sleep quality, anxiety, and anhedonia. Magnetic resonance spectroscopy can provide direct and quantitative information on Glu at the brain region level and thus will be used to verify whether the treatment restores the excitatory-inhibitory balance of the brain network by regulating excitatory neurotransmission.

XSL•FO
RenderX

Indicators of Mechanism

Heart Rate Variability

The following HRV indices will be assessed: very LF (vLF), LF, HF, total power (TP), and LF/HF. Additionally, the parasympathetic nerve index (Log L×T) and sympathetic nerve index (Log L/T) will be calculated, where L denotes LF and T denotes TP, to evaluate ANS activity.

Polysomnography

The participants will be monitored by polysomnography at the TCM Diagnosis and Treatment Center of Sleep Disorders in the Department of Acupuncture and Moxibustion of our hospital after 4 weeks of treatment and at 4 weeks after the treatment ends. The monitoring was performed using a Philips polysomnography system (model: Alice 6). Positioning and interpretation will be performed according to Berry et al [15]. The changes in the sleep parameters of patients will be analyzed using the data collected by Sleepware G3 (Philips Respironics), including sleep efficiency, the proportion of each sleep stage, sleep latency, and other sleep indicators corresponding to other stages.

Magnetic Resonance Imaging

In this study, a GE Discovery MR750 1.5T MRI scanner will be used for imaging. During the test, the participant will be placed in a supine position, with their head held stable with a fixation frame. Moreover, the participant will be asked to wear earplugs to reduce stimulation due to equipment noise, so that the participant can relax as much as possible. First, a routine sequence axial scan will be performed to exclude organic lesions in the brain or the presence of artifacts; GABA and Glu experimental data will be collected again if there is no definite abnormality in the brain.

Safety Evaluation

Adverse events during the study will be recorded and reported, and the patients will be treated in a timely and reasonable manner. The researchers will explain to the patients that they (or their families) are required to truthfully communicate the changes in their condition after treatment. The physicians will avoid guiding questions. During the observation of a curative effect, the researchers will pay attention to identify any adverse reactions.

Adverse events such as bleeding, hematoma, fainting, pain at the acupuncture site, and elevated blood pressure caused by acupuncture at any time will be recorded. Regardless of whether the adverse reactions or adverse events are related to the treatment method of this study, they will be recorded in detail. When these conditions occur, the patients will be recorded and treated properly in time until they completely return to normal, and the causes of adverse reactions will be analyzed.

Data Management

The researchers will fill in the collected data into the case report forms according to the requirements of the research program. At the end of the study, the researchers will submit the case report forms of all patients included in this study to the data management center. These case report forms are required to be complete and signed. An internal data consistency check will be performed on all case report forms. Any inconsistencies identified will be logged, and a formal query will be generated and sent to the principal investigator for clarification and resolution.

Sample Size and Statistical Methods

The main outcome measure of this study will be the total PSQI score at 4 weeks after treatment. The sample size was calculated on the basis of the expected difference between the test acupuncture group and the conventional acupuncture group in this index.

The expected effect size was determined as follows: studies have shown that acupuncture can significantly reduce PSQI scores compared with conventional therapy, with a combined mean difference of 2.0 to 3.5 points [16]. With reference to an RCT of similar design [17], we conservatively estimated that the mean difference (Δ) of PSQI between the 2 groups after treatment was 2.5 points, and the combined SD was approximately 4.0 points. According to this calculation, the standardized effect size is 0.625 (Δ/SD=2.5/4.0). Using G*Power software (version 3.1; Heinrich-Heine University Dusseldorf), setting α=.05 (bilateral) and test efficiency (1-β)=80%, according to a 1:1 distribution, the minimum sample size of each group was calculated to be 43. Considering a dropout rate of approximately 10%, 48 patients will be included in each group, and a total sample size of 96 patients was finally determined.

SPSS software (version 22.0; IBM Corp) will be used to analyze the collected data. All data will be expressed as mean (SD). For measurement data of the 2 groups, normally distributed data will be analyzed by independent 2-tailed *t* test and paired 2-tailed *t* test. If the data are not normally distributed, the Mann-Whitney rank sum test will be used. One-way ANOVA will be used for comparison between groups. Count data will be analyzed using a chi-square test, and rank data will be compared using the rank sum test and other analysis methods. A *P* value of ≤.05 will be considered statistically significant. To control for potential confounding, the use of concomitant medications and measures derived from sleep diaries will be incorporated as covariates into a generalized linear model for sensitivity analysis of the primary outcomes.

Final Report and Publication

After the end of the study, the main researchers will take the lead to cooperate with the researchers at each clinical department to write the research summary report, and the researchers at each clinical department will sign the research summary report. The research report will include a description of the research purpose, the methods used in the research, and the results and conclusions.

Results

This study was funded in November 2023. Recruitment and data collection began in January 2024. The results will be shared with participants, researchers, and the public via conferences, publications, and online platforms. The findings will be submitted to a peer-reviewed journal. Participant recruitment

XSL•FO
RenderX

is underway and will continue until the target sample size of 96 patients with primary insomnia of liver depression and spleen deficiency type and 48 healthy controls is reached. Participant recruitment is scheduled to be completed by the end of 2026. Each enrolled patient will undergo the 4-week intervention followed by a 4-week follow-up period as per the study protocol. As of December 2025, a total of 40 participants have been enrolled, of whom 4 have withdrawn. Data collection is expected to conclude by the end of 2026, with full data analysis and manuscript preparation completed by summer 2027.

Discussion

Study Rationale

In this study, an RCT was designed to evaluate the efficacy of acupuncture at specific meridians and the acupoints SI17 and TE16 on patients with primary insomnia of liver depression and spleen deficiency type and to conduct a parallel assessment of ANS regulation using HRV and neurochemical indicators.

The selection of the cervical acupoints SI17 and TE16 is based on the classical meridian divergence (Jing Bie) theory of TCM [5]. According to "Lingshu · Jing Bie," although SI17 belongs to the small intestine meridian, it is located at the intersection of the liver and gallbladder meridians. Therefore, acupuncture at SI17 can soothe and regulate the liver and gallbladder qi. Although TE16 belongs to the triple energizer meridian, it is located at the intersection of the triple energizer, pericardium, and gallbladder meridians. Acupuncture at TE16 can regulate the qi of the Shaoyang meridian, relieve the stagnation of the upper energizer and tranquilize the mind, and provide a theoretical basis for related diseases. In TCM, the neck is a key region where multiple yang meridians converge [18]. As pivotal acupoints in the neck, SI17 and TE16 are essential for regulating the mind and balancing yin and yang, and their clinical values stem from the deep regulation of ying-wei qi and blood by the hand Taiyang and hand Shaoyang meridians. Specifically, the Tianrong point follows the path of "entering the armpit and walking the heart" by the hand Taiyang meridian, which has the effect of dredging the neck qi and blood and calming the heart and nerves. Its anatomical deep layer corresponds to the cervical sympathetic ganglion, which can inhibit the excessive excitement of the sympathetic nerve. The Tianyou acupoint, with the function of the hand Shaoyang meridian "scattered in the chest," can regulate the triple energizer qi, straighten out the upper energizer, and affect the activity of the vagus nerve. According to the "Lingshu · Jing Bie," these divergences "branch from the main channels, enter the body cavity, link exterior and interior, and harmonize *ying* (nutritive) and *wei* (defensive) energies." The yang meridians are different from the original meridians, and the yin meridians are different from the yang meridians in the exterior and interior, which strengthens the close relationship between the limbs and the zang-fu organs, head, and face. In TCM, insomnia of the liver depression and spleen deficiency type is considered to stem from disrupted qi flow and impaired balance between ying and wei [19,20]. The mechanism of acupuncture involves regulating the relationship between yin and yang, the unity of opposites, to restore the harmonious flow and balance of qi in the whole body [21]. On the basis of this principle, specific acupoints on the neck are stimulated to dredge the Shaoyang meridian and reconcile ying and wei so as to promote yang qi into yin and support natural sleep. This follows the classical TCM principle that "acupuncture treats areas along the pathway of the meridian."

Notably, this traditional rationale aligns with modern neuroanatomical insights. The cervical regions where these meridian divergences run are anatomically close to major sympathetic ganglia, such as the stellate ganglion [22,23]. Therefore, acupuncture at SI17 and TE16 may modulate local sympathetic nerve activity, in a way similar to SGB—a procedure known to influence autonomic function [24-26]. This integrated view connects the TCM concepts of qi regulation and ying-wei harmony with the modern understanding of ANS balance, offering a coherent mechanistic framework for acupuncture.

Proposed Mechanism and Empirical Testing

The proposed mechanism can be traced through specific physiological pathways. The anatomical proximity of these points to the cervical sympathetic chain means that stimulation may send signals through the spinal cord to higher brain regions that regulate autonomic function, such as the brain stem and hypothalamus [27,28]. This influence on the central nervous system can rebalance the sympathetic and parasympathetic activity, a change measurable through HRV [29-32]. Thus, HRV provides an objective link between the TCM idea of restored ying-wei harmony and modern metrics of ANS balance. Beyond this neural pathway, the intervention likely also involves neuroendocrine adjustments [33-35]. Chronic insomnia and the TCM pattern of liver qi stagnation are both associated with overactivity of the body's stress response system, the hypothalamic-pituitary-adrenal axis [36-38]. By mitigating excessive sympathetic drive, acupuncture may help calm the hypothalamic-pituitary-adrenal axis hyperactivity [39], potentially leading to lower cortisol levels and a more sleep-conducive internal milieu [40-42]. In summary, the treatment is hypothesized to work through a convergent network, modulating both neural excitability and endocrine function to address the multifaceted pathophysiology of insomnia.

To empirically test this mechanistic model, our study employs a multimethod assessment strategy. In addition to standard patient-reported outcomes (PSQI and ISI), we will measure HRV to directly quantify ANS changes [43]. We will also analyze the levels of key neurotransmitters—GABA and Glu. We hypothesize that successful treatment will be correlated with a shift toward a less excited central nervous system state, reflected in an increased GABA to glutamate ratio [18]. This combined biochemical and physiological evidence aims to clarify how stimulation at these specific points translates into clinical improvement, thereby strengthening the scientific rationale for this targeted acupuncture protocol.

This study incorporates a parallel cohort of healthy controls to establish normative baselines for HRV and relevant neurochemical markers. These reference values will enable the quantitative assessment of the severity of ANS dysfunction in patients with insomnia and provide a benchmark for evaluating postintervention recovery. This comparative design strengthens

XSL•FO
RenderX

the robustness of the findings and supports a multidimensional evaluation of treatment efficacy.

To comprehensively capture both objective and subjective dimensions of sleep and autonomic function, the primary outcome indicators include the PSQI and SF-36. The secondary outcome indicators include the ISI and DBAS-16 as well as GABA and Glu levels. These 2 neurotransmitters were selected as biomarkers because they play an important role in sleep-wake regulation and the high arousal state characterized by primary insomnia. Assessing these neurochemicals will help to elucidate the mechanisms underlying treatment response.

Methodologically, the trial employs a single-blind, randomized design with block randomization to minimize selection bias. The sample size was determined by an a priori power calculation to ensure adequate statistical power, and all acupuncture procedures will be performed adhering to a strict standardized protocol to maintain treatment fidelity across participants. Focusing exclusively on patients with the liver depression-spleen deficiency pattern will reduce clinical heterogeneity, allowing a more precise investigation of acupuncture mechanisms within this defined TCM syndrome framework.

Limitations

Several limitations warrant consideration. Achieving complete blinding in acupuncture trials remains inherently challenging, and the participant age range (18 - 40 y) may constrain the generalizability of the findings to older populations. Nevertheless, the protocol aligns with the CONSORT (Consolidated Standards of Reporting Trials) guidelines through structured data collection, systematic adverse event monitoring, and prespecified statistical analysis—features that collectively uphold the internal validity and interpretability of the study outcomes.

This study has some limitations. First, the scope of sample selection is limited. The inclusion criteria limited participant age to 18 to 40 years and the course of the disease to more than 2 weeks. Although a narrow age range is conducive to controlling confounding age factors, it cannot represent the situation of patients with insomnia in other age groups, which limits the universality of the research results and may lead to age-related bias in the conclusions. Second, there are limitations in the blinding design. Due to the significant difference between the intervention strategies of the Jing Bie theory acupuncture group and the traditional acupuncture group, the blinding protocol for acupuncture physicians faces challenges in practice. Therefore, this study adopts a single-blind design. Although this may, in theory, introduce operator bias, several measures will be taken to maintain methodological rigor: the acupuncture physician will follow strict standardized operating procedures; the outcome evaluators and data analysts will remain blinded to the grouping; and the participants will know only that they received active acupuncture intervention and will not know the specific grouping. These measures aim to minimize potential bias while maintaining the authenticity of clinical practice. Future studies may consider using sham acupuncture or noninvasive needles in the control group to achieve double blinding, although such methods also have inherent limitations in simulating real acupuncture effects.

In summary, this study will explore a new path for acupuncture treatment of insomnia based on the theory of Jing Bie and establish acupoint selection (targeting cervical regions to regulate autonomic function) and evaluation (integrated comprehensive evaluation program using PSQI, HRV, and neurochemical biomarkers), providing a reference for subsequent research. However, the limitations of sample representativeness and blinding design suggest that future research needs to be improved in terms of expanding the age range of the sample and optimizing the strategy of implementation of blinding, so as to further verify and expand the findings of this study.

Conclusion

This study is expected to fill the following evidence gaps.

Efficacy Evidence of Specific Syndromes

At present, most of the studies on acupuncture treatment of insomnia do not strictly distinguish TCM syndromes. This study will focus on insomnia with liver depression and spleen deficiency and provide high-quality RCT evidence for specific syndromes.

Mechanism of ANS Regulation

Previous studies have not sufficiently discussed the mechanism of acupuncture regulating autonomic nerve function. This study will provide a more comprehensive explanation for the mechanism of acupuncture in improving insomnia through an in-depth analysis of HRV and other indicators.

Long-Term Efficacy Evaluation

Most acupuncture studies only focus on short-term efficacy. This study will use a 4-week follow-up to evaluate the long-term effect of acupuncture and provide a more reliable basis for clinical application.

Changes in Neurotransmitter Levels

In this study, GABA and Glu levels will be determined to fill the evidence gap of acupuncture regulation of neurotransmitters in patients with insomnia and to provide a new direction for the neurobiological mechanism of acupuncture treatment of insomnia.

Acknowledgments

We are particularly grateful to anyone who provided helpful advice on the design of the trial. Meanwhile, we appreciate the help from the patients participating in the trial, including those who contributed toward the trial but do not meet the criteria for authorship. In this study, the generation process of the manuscript (including the Discussion section) did not use any generative artificial intelligence tools for content creation, text writing, or opinion generation.The two authors, Kang He and Chuanlong Zhou, have the same contribution as corresponding authors.

XSL•FO
RenderX

Funding

This work was supported by the Zhejiang Provincial Health Commission (grant numbers 2024ZL531 and 2025ZL368).

Authors' Contributions

SF played a role in the conceptualization, formal analysis, investigation, methodology, software, visualization, and writing (original draft and editing). KH determined research directions and methods and guided the entire research process. CZ led the overall design of the research, proposed research questions, determined research directions and methods, and guided the entire research process.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[DOCX File, 30 KB - resprot_v15i1e84122_app1.docx]

References

1. Buysse DJ. Insomnia. JAMA 2013 Feb 20;309(7):706-716. [doi: 10.1001/jama.2013.193] [Medline: 23423416]
2. Fortier-Brochu E, Beaulieu-Bonneau S, Ivers H, Morin CM. Insomnia and daytime cognitive performance: a meta-analysis. Sleep Med Rev 2012 Feb;16(1):83-94. [doi: 10.1016/j.smrv.2011.03.008] [Medline: 21636297]
3. Yao R, Chen M, Li Q. Research trend of acupuncture in insomnia treatment: analysis of knowledge map based on CiteSpace [Article in Chinese]. Neural Inj Funct Reconstr 2021;16(3):146-149 [FREE Full text]
4. Huang Y, Yi G, Zhou B, et al. Exploring the application of the “root, flow, injection, and entry” and “liaison” theories in the diagnosis and treatment of cervical spondylosis with radiculopathy [Article in Chinese]. Zhen Jiu Lin Chuang Za Zhi 2016;32:79-82 [FREE Full text] [doi: 10.19917/j.cnki.1005-0779.2016.03.027]
5. Gao B, Cheng Y, Zeng YL, Niu SP, Li YM, Wang Y. Analysis of the divergent meridians of twelve meridians [Article in Chinese]. Zhongguo Zhen Jiu 2020;40:887-890. [Medline: 32869601]
6. Tian Y, Hu Y, Hu T, et al. Stellate ganglion block therapy for complex regional pain syndrome: a systematic review and meta-analysis. Pain Physician 2024 May;27(4):175-184. [Medline: 38805523]
7. Shang K, Kong Q. Research progress in stellate ganglion block and regulation of autonomic nervous functions [Article in Chinese]. Sichuan Da Xue Xue Bao Yi Xue Ban 2024 Nov 20;55(6):1604-1609. [doi: 10.12182/20241160510] [Medline: 39990849]
8. Kobayashi H, Hayashi M, Kobayashi S, Kabuto M, Handa Y, Kawano H. Effect of endothelin on the canine basilar artery. Neurosurgery 1990 Sep;27(3):357-361. [doi: 10.1097/00006123-199009000-00003] [Medline: 2234327]
9. Palmer RM, Ferrige AG, Moncada S. Nitric oxide release accounts for the biological activity of endothelium-derived relaxing factor. Nature 1987;327(6122):524-526. [doi: 10.1038/327524a0] [Medline: 3495737]
10. Park SU, Jung WS, Moon SK, et al. Effects of acupuncture on autonomic nervous system in normal subjects under mental stress. J Korean Oriental Med 2008:107-115 [FREE Full text]
11. Chung JWY, Yan VCM, Zhang H. Effect of acupuncture on heart rate variability: a systematic review. Evid Based Complement Alternat Med 2014;2014:819871. [doi: 10.1155/2014/819871] [Medline: 24693326]
12. Wang Q, Gao L, Liu C, et al. Treatment of 66 cases of intractable insomnia with the Yuan-Luo-Back-Shu point combination method [Article in Chinese]. J Chin Acupunct 2022;42:369-370. [doi: 10.13703/j.0255-2930.20210302-k0001]
13. Stein PK, Bosner MS, Kleiger RE, Conger BM. Heart rate variability: a measure of cardiac autonomic tone. Am Heart J 1994 May;127(5):1376-1381. [doi: 10.1016/0002-8703(94)90059-0] [Medline: 8172068]
14. American Academy of Sleep Medicine. International Classification of Sleep Disorders, 2nd edition: American Academy of Sleep Medicine; 2005.
15. Berry RB, Brooks R, Gamaldo CE, Harding SM, Marcus C, Vaughn BV. The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications, Version 2.6: American Academy of Sleep Medicine; 2020. URL: https://aasm.org/clinical-resources/scoring-manual [accessed 2026-01-12]
16. Schiller J, Niederer D, Kellner T, et al. Effects of acupuncture and medical training therapy on depression, anxiety, and quality of life in patients with frequent tension-type headache: a randomized controlled study. Cephalalgia 2023 Jan;43(1):3331024221132800. [doi: 10.1177/03331024221132800] [Medline: 36622877]
17. Guo J, Wang LP, Liu CZ, et al. Efficacy of acupuncture for primary insomnia: a randomized controlled clinical trial. Evid Based Complement Alternat Med 2013;2013:163850. [doi: 10.1155/2013/163850] [Medline: 24159338]
18. Chen Z, Wang Z, Gao C, Zheng H, Cai T, Liu J. Exploration of tuina for the treatment of acute pharyngitis based on “Tianyou (TE 16) and five regions” in Huangdi Neijing [Article in Chinese]. Zhongguo Zhen Jiu 2024 Oct 12;44(10):1192-1196. [doi: 10.13703/j.0255-2930.20231128-k0010] [Medline: 39401819]

XSL•FO
RenderX

19. Zhang B, Wang Q, Zhang Y, et al. Treatment of insomnia with traditional Chinese medicine presents a promising prospect. Phytother Res 2025 Apr 18. [doi: 10.1002/ptr.8495] [Medline: 40251853]
20. Zhang Q, Fan Y, Song Q. Application of Xiao Chai Hu Tang in paediatric diseases based on the principle of “Shao Yang is the Pivot”. MEDS Chinese Medicine 2023;5(6):48-54. [doi: 10.23977/medcm.2023.050610]
21. Guo Y, Sun T, Qiu F, et al. Electroacupuncture combined with cognitive rehabilitation outperforms cognitive rehabilitation alone in treating post-stroke cognitive impairment: a randomized controlled trial. Front Neurol 2025;16:1507475. [doi: 10.3389/fneur.2025.1507475] [Medline: 39944540]
22. Jiang JF. Regional heterogeneity of acupoints stimulation in the modulation of autonomic nerve system. World J Acupunct Moxibustion 2024 Apr;34(2):77-82. [doi: 10.1016/j.wjam.2024.04.002]
23. Yuan X, Tang Y, Xie X, et al. The effect of stellate ganglion block on postoperative recovery quality in peri-menopausal women undergoing gynecological laparoscopic surgery. Front Med 2025;12. [doi: 10.3389/fmed.2025.1561512]
24. Li Z, Wang L, Cong P, et al. Ultrasound-guided stellate ganglion block with cervicotemporal myofascial trigger point acupuncture for insomnia with comorbid anxiety and depression: a retrospective observational study. Nat Sci Sleep 2025;17:2131-2143. [doi: 10.2147/NSS.S551748] [Medline: 40933856]
25. Shan HH, Chen HF, Ni Y, Yang JX, Zhou XL. Effects of stellate ganglion block through different approaches under guidance of ultrasound. Front Surg 2021;8:797793. [doi: 10.3389/fsurg.2021.797793] [Medline: 35111806]
26. Wang J, Chen Z, Huang Y, et al. Nomogram model based on clinical factors and autonomic nervous system activity for predicting residual renal function decline in patients undergoing peritoneal dialysis. Front Neurosci 2024;18:1429949. [doi: 10.3389/fnins.2024.1429949]
27. Wang LH, Ding WQ, Sun YG. Spinal ascending pathways for somatosensory information processing. Trends Neurosci 2022 Aug;45(8):594-607. [doi: 10.1016/j.tins.2022.05.005] [Medline: 35701247]
28. Jang I, Cho K, Moon S, et al. A study on the central neural pathway of the heart, Nei-Kuan (EH-6) and Shen-Men (He-7) with neural tracer in rats. Am J Chin Med 2003;31(4):591-609. [doi: 10.1142/S0192415X03001314] [Medline: 14587882]
29. Wang G, Tian Y, Jia S, Zhou W, Zhang W. Acupuncture regulates the heart rate variability. J Acupunct Meridian Stud 2015 Apr;8(2):94-98. [doi: 10.1016/j.jams.2014.10.009] [Medline: 25952126]
30. Ma Y, Mullington JM, Wayne PM, Yeh GY. Heart rate variability during sleep onset in patients with insomnia with or without comorbid sleep apnea. Sleep Med 2024 Oct;122:92-98. [doi: 10.1016/j.sleep.2024.07.034] [Medline: 39137665]
31. Ma Y, Wayne P, Yeh G. 0417 Pre-sleep wakefulness in patients with insomnia: heart rate variability during sleep onset. Sleep 2023 May 29;46(Supplement_1):A184-A185. [doi: 10.1093/sleep/zsad077.0417]
32. Zhao W, Jiang B. Heart rate variability in patients with insomnia disorder: a systematic review and meta-analysis. Sleep Breath 2023 Aug;27(4):1309-1313. [doi: 10.1007/s11325-022-02720-0] [Medline: 36219384]
33. Gompf HS, Anaclet C. The neuroanatomy and neurochemistry of sleep-wake control. Curr Opin Physiol 2020 Jun;15:143-151. [doi: 10.1016/j.cophys.2019.12.012] [Medline: 32647777]
34. Wu Y, Wang L, Yang F, Xi W. Neural circuits for sleep-wake regulation. Adv Exp Med Biol 2020;1284:91-112. [doi: 10.1007/978-981-15-7086-5_8] [Medline: 32852742]
35. Dijk DJ, Landolt HP. Sleep physiology, circadian rhythms, waking performance and the development of sleep-wake therapeutics. Handb Exp Pharmacol 2019;253:441-481. [doi: 10.1007/164_2019_243] [Medline: 31254050]
36. Nicolaides NC, Vgontzas AN, Kritikou I, Chrousos G. HPA axis and sleep. In: Feingold KR, Adler RA, Ahmed SF, editors. Endotext: MDText.com, Inc; 2000. [Medline: 25905298]
37. Dressle RJ, Feige B, Spiegelhalder K, et al. HPA axis activity in patients with chronic insomnia: a systematic review and meta-analysis of case-control studies. Sleep Med Rev 2022 Apr;62:101588. [doi: 10.1016/j.smrv.2022.101588] [Medline: 35091194]
38. Xie Z, Xie H, Peng X, et al. The antidepressant-like effects of Danzhi Xiaoyao San and its active ingredients. Phytomedicine 2023 Oct;119:155015. [doi: 10.1016/j.phymed.2023.155015] [Medline: 37597362]
39. Li P, Zhao J, Wei X, et al. Acupuncture may play a key role in anti-depression through various mechanisms in depression. Chin Med 2024 Oct 4;19(1):135. [doi: 10.1186/s13020-024-00990-2] [Medline: 39367470]
40. Zheng JY, Zhu J, Wang Y, Tian ZZ. Effects of acupuncture on hypothalamic-pituitary-adrenal axis: current status and future perspectives. J Integr Med 2024 Jul;22(4):445-458. [doi: 10.1016/j.joim.2024.06.004] [Medline: 38955651]
41. Cui L, Li S, Wang S, et al. Major depressive disorder: hypothesis, mechanism, prevention and treatment. Signal Transduct Target Ther 2024 Feb 9;9(1):30. [doi: 10.1038/s41392-024-01738-y] [Medline: 38331979]
42. Ntoumas I, Karatzaferi C, Boubougiatzi F, et al. Massage positively influences daytime brain activity and reduces arousal state in poor sleepers: a randomized controlled trial. BMC Complement Med Ther 2025 Jul 28;25(1):290. [doi: 10.1186/s12906-025-05022-6] [Medline: 40722019]
43. Liu CR, Yang CY, Sharma D, et al. Associations between sleep duration and autonomic nervous system regulation in patients with probable Alzheimer’s disease: a cross-sectional pilot study. Clocks Sleep 2024 Sep 24;6(4):533-545. [doi: 10.3390/clockssleep6040035] [Medline: 39449309]

XSL•FO
RenderX

Abbreviations

ANS: autonomic nervous system
CONSORT: Consolidated Standards of Reporting Trials
DBAS-16: 16-item Dysfunctional Beliefs and Attitudes About Sleep scale
Glu: glutamic acid
HF: high-frequency power
HRV: heart rate variability
ISI: Insomnia Severity Index
LF: low-frequency power
Log L/T: sympathetic nerve index
Log L×T: parasympathetic nerve index
PSQI: Pittsburgh Sleep Quality Index
RCT: randomized controlled trial
SF-36: 36-item Short Form Health Survey
SGB: stellate ganglion block
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TCM: traditional Chinese medicine
TP: total power
vLF: very low-frequency power

Edited by J Sarvestan; submitted 15.Sep.2025; peer-reviewed by X Li; revised version received 25.Dec.2025; accepted 26.Dec.2025; published 22.Jan.2026.

Please cite as:
Fu S, He K, Zhou C
Curative Effect and Autonomic Nerve Function of Patients With Primary Insomnia of Liver Depression and Spleen Deficiency Type Based on the Acupoint Selection of Meridian Theory: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84122
URL: https://www.researchprotocols.org/2026/1/e84122
doi:10.2196/84122

© Su Fu, Kang He, Chuanlong Zhou. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effect of Virtual Reality–Based Social Cognitive Training for Autistic Adults: Protocol for STEPS (Social Cognitive Training Enhancing Pro-Functional Skills) Randomized Clinical Trial

Johannes Andresen[1,2], MSc; Alberte C E Jeppesen[1,2], MSc; Anne Sofie Due[1], MSc; Lise Sandvig Mariegaard[1], MSc; Rizwan Parvaiz[3], MD; Carsten Hjorthøj[4,5], PhD; Amy Elizabeth Pinkham[6], PhD; Merete Nordentoft[4,7], DrMSc; Gasper Letnar[8], MSc; Jens Richardt Møllegaard Jepsen[9,10], PhD; Louise Birkedal Glenthøj[1,2], DrMSc

[1]VIRTU Research Group, Mental Health Center Copenhagen, Copenhagen University Hospital – Mental Health Services CPH, Gentofte Hospitalsvej 3A, 4th Floor, Copenhagen, Denmark

[10]Center for Neuropsychiatric Schizophrenia Research, Mental Health Center Glostrup, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark

[2]Department of Psychology, University of Copenhagen, Copenhagen, Denmark

[3]Department of ADHD and Autism, Mental Health Center Glostrup, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark

[4]Copenhagen Research Center for Mental Health (CORE), Mental Health Center Copenhagen, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark

[5]Department of Public Health, University of Copenhagen, Copenhagen, Denmark

[6]School of Behavioural and Brain Sciences, The University of Texas at Dallas, Richardson, TX, United States

[7]Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

[8]Sanos ApS, Søborg, Denmark

[9]Child and Adolescent Mental Health Center, Copenhagen University Hospital – Mental Health Services CPH, Copenhagen, Denmark

Corresponding Author:
Louise Birkedal Glenthøj, DrMSc
VIRTU Research Group, Mental Health Center Copenhagen, Copenhagen University Hospital – Mental Health Services CPH, Gentofte Hospitalsvej 3A, 4th Floor, Copenhagen, Denmark

Abstract

Background: Autistic adults constitute a growing and largely overlooked population with limited clinical and research resources. Social cognitive impairments are key deficits faced by this population, significantly impacting social interactions, educational and vocational functioning, and quality of life. Interventions targeting social cognition in autistic adults have shown promising results. Recent studies investigating the effect of virtual reality (VR)–based interventions for autistic adults have provided preliminary evidence supporting the feasibility and effectiveness of using this innovative technology. These studies indicate that VR interventions can enhance functional and social skills and improve specific neurocognitive and social cognitive functions. However, large-scale randomized clinical trials are urgently needed to fully assess the effectiveness of VR-based interventions for autistic adults.

Objective: This protocol aims to provide a comprehensive description of the design and methodology of the STEPS (Social Cognitive Training Enhancing Pro-Functional Skills) trial.

Methods: STEPS is a clinical, randomized, assessor-blinded, parallel-group superiority trial. A total of 140 participants will be allocated to receive either virtual reality–based social cognitive training (VRSCT) + treatment as usual (TAU) or TAU alone. The experimental group will receive 12 weekly 1-hour sessions of VRSCT, aiming at improving psychosocial functioning and social cognition through exposure to virtual social environments. The intervention comprises 3 core modules, namely emotions, social understanding, and complex social interactions. The exact content and duration of TAU received by each participant will be mapped and documented upon trial completion. Assessments will be conducted at baseline, at cessation of the intervention (3 months post baseline), and at 6 months post baseline.

Results: Participant enrollment began in May 2024. As of February 2025 (initial manuscript submission), 34 participants had been enrolled, increasing to 97 participants as of December 2025. Completion of enrollment is expected in April 2026. Data analysis is expected to begin in October 2026 following the final 6-month follow-up assessment. Results are anticipated in December 2026 and will be disseminated through peer-reviewed publications.

Conclusions: To our knowledge, STEPS is the hitherto largest randomized clinical trial globally investigating the effect of VRSCT for autistic adults. The results of this innovative intervention approach may significantly advance research in the field

XSL•FO
RenderX

of autism. VRSCT holds potential to improve psychosocial functioning, quality of life, and co-occurring clinical symptoms, and to reduce social cognitive deficits in autistic adults. Establishing evidence-based interventions is crucial for addressing the debilitating psychosocial challenges faced by this population, especially considering the absence of established gold-standard treatments.

Trial Registration: ClinicalTrials.gov NCT06438536; https://clinicaltrials.gov/study/NCT06438536

International Registered Report Identifier (IRRID): DERR1-10.2196/72854

(JMIR Res Protoc 2026;15:e72854) doi:10.2196/72854

KEYWORDS

autism spectrum disorder; virtual reality; psychosocial functioning; social skills; social cognitive training; emotion recognition; theory of mind

Introduction

Terminology

In this paper, we use the term autistic adults, in line with international recommendations [1] to refer to adults on the autism spectrum. Additionally, we use the term autism spectrum condition as an equivalent to autism spectrum disorder to minimize discriminatory language.

Autism

According to the diagnostic criteria described in the *DSM-5-TR* (*Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision*) by the American Psychiatric Association, autism is defined as a neurodevelopmental condition characterized by persistent deficits in social communication, such as difficulties with socioemotional reciprocity, communication, and the development, maintenance, and understanding of relationships. This is accompanied by restricted, repetitive patterns of behavior, interest, or activities. These clinical characteristics must be present from early development (but may not become fully manifest until social demands exceed limited capacities or may be masked by learned strategies in later life), cause significant impairment in psychosocial functioning, and cannot be better explained by other intellectual or developmental conditions [2]. Autism is a lifelong neurodevelopmental variation that is estimated to occur in 1% - 2% of the global population [3]. The prevalence rate is increasing [4,5], as well as the lifelong economic impact associated with the condition. In 2019, the estimated lifetime social cost of an individual on the autism spectrum was approximately US $3.6 million in the United States, with the largest costs associated with lost productivity and adult care [6]. These costs should not be attributed to individuals on the autism spectrum but should rather be regarded as a consequence of inadequate support and accommodations in educational and occupational settings and a lack of effective interventions that address the specific challenges faced by the autistic population. These obstacles are particularly pronounced for autistic adults, as many experience a lack of appropriate interventions and support [7], are vastly underrepresented in the workforce, and are at heightened risk of depression and suicidal actions compared to the general population [8-11]. Despite the critical need for interventions and support systems aimed at improving outcomes for autistic adults, research on autism has primarily focused on children and adolescents, overlooking the growing population of adults who face extensive unmet treatment needs [12].

Social Cognition in Autistic Adults

Social cognition can be defined as "(…) the mental operations that underlie social interactions, including perceiving, interpreting, and generating responses to the intentions, dispositions, and behaviors of others." [13]. As this definition implies, social cognition is a multifaceted and complex construct, underpinned by different functions, such as emotion processing and recognition, theory of mind (ToM), empathy, social perception, social attribution, and social awareness. It is well established that autistic adults, on average, show reduced performance on social cognitive tasks, including facial emotion processing, prosody interpretation, mental state labeling, social perception, social attribution, and social understanding, compared to neurotypicals. However, social cognitive profiles in autism present differently from person to person, and individual profiles can be dynamic and may shift over time [14].

Social cognitive impairments are key deficits faced by autistic adults [15], which severely impact social interactions, hinder educational and vocational functioning, and compromise quality of life [16-18]. The social cognitive deficits in autism encompass both the more basic functions, such as recognizing emotions or making appropriate eye contact, and higher-order functions, such as inferring others' mental states (ToM) or understanding subtle aspects of social communication [19,20]. These functions are vital to understanding and reacting adequately to dynamic, complex social situations in all aspects of life, from making friends and building intimate relationships in private life to being able to participate appropriately in educational or professional settings.

Interventions Targeting Social Cognitive Impairments

Targeting social cognitive deficits in autistic adults without using virtual reality (VR) technology has been shown to improve psychosocial functioning [21]. These interventions often use computer- or group-based exercises to improve different aspects of social cognition (eg, perception of social cues, emotion recognition, ToM, and social skills) through repeated practice and skills-based training. While preliminary evidence supports the effectiveness of targeted non-VR-based social cognitive interventions for autistic adults [22], the existing studies are limited in number and are subject to significant methodological limitations, such as not being randomized clinical trials or

XSL•FO
RenderX

having small sample sizes (n≤17) categorizing them as pilot studies. Previous studies of social cognition training for autistic adults typically target multiple domains, such as emotion processing and recognition, ToM, and social perception and attribution. Hypothesized mechanisms of change include repeated practice, gradual and errorless learning, compensatory strategy training, and supportive contexts that promote generalization [22-24]. Specifically for VR-based interventions, additional mechanisms of change include immersion in ecologically valid scenarios, graded exposure, individually tailored training, real-time feedback, control of sensory load, and structured transfer to daily life [25,26].

VR in Mental Health

The use of VR holds great potential for mental health interventions by enabling researchers and clinicians to evoke psychological reactions to lifelike scenarios in real-time within a controlled laboratory environment. VR serves as a powerful tool for delivering exposure therapy, as it creates a safe, flexible, and controlled space for immersive exposure to challenging situations. This approach allows individuals to practice confronting real-world challenges without facing the consequences of real-life situations, while also reducing the social pressure often associated with face-to-face interactions [27]. As a result, VR exposure is considered an effective tool for treating various psychological challenges, with minimal risks of adverse side effects [28]. Additionally, creating in vivo exposure to challenging social situations in daily life is often difficult to arrange and nearly impossible to control and can be overly stressful for participants, making engagement in such settings challenging. Interventions using VR exposure can overcome these obstacles by offering complete control over various scenarios and their elements, leading to greater user acceptance. As a result, participants may be more inclined to engage in exposure therapy when it is provided in a VR setting [29]. Additionally, VR enables the creation of highly personalized interventions by allowing control over the level of exposure, tailoring scenarios to individual needs, and making real-time adjustments.

VR-based Interventions for Autistic Adults

By using immersive VR exposure in interventions targeting social cognitive challenges in autistic adults, it becomes possible to assess real reactions and responses to life-like challenging social situations, allowing access to the real-time cognition of the participant in such situations. This approach might be of particular importance for interventions aimed at individuals on the autism spectrum, since self-awareness and self-insight are often compromised by alexithymia (difficulties in identifying and describing one's own emotional state) [30]. Alexithymia can hinder the accurate assessment of psychosocial difficulties in interventions, as more traditional approaches often require individuals to describe their own emotions retrospectively. Therefore, using VR exposure might enhance assessment accuracy and, in turn, enable more personalized and effective interventions. Furthermore, individuals on the autism spectrum may prefer technology-based interventions [23], and preliminary evidence indicates that VR-based interventions lead to faster symptom reduction than traditional psychological interventions [31,32].

Given that social cognitive differences in autistic adults often involve emotion processing and recognition, ToM, social understanding, and motivation [14,15], VR-based interventions are theoretically well-suited to target these domains. The ability to simulate complex, dynamic social interactions in a controlled, repeatable, and safe environment enables targeted practice of interpreting and responding to social cues, adjusting difficulty as skills improve [24,33]. By providing immediate feedback, opportunities for graded exposure, and the capacity to individualize scenarios, VR can support the gradual enhancement of underlying neuropsychological processes essential for real-world social functioning.

Research on VR-based interventions for individuals on the autism spectrum has gained significant momentum in recent years. Most previous studies have investigated interventions targeting specific functional domains, such as job interviewing skills, driving skills, and adaptive social skills [26,34]. While some studies have examined the impact of VR-based interventions on cognition in individuals on the autism spectrum, these studies have primarily addressed specific neurocognitive domains and daily life skills [35], with comparatively fewer studies investigating social cognition [24,25]. The results of these studies have provided preliminary evidence supporting the feasibility of VR-based interventions for individuals on the autism spectrum. These studies have yielded promising results, indicating positive effects on specific functional and social skills, as well as neurocognitive domains like attention, memory, and executive functioning and social cognitive domains including emotion recognition and ToM. However, the findings are limited by small sample sizes, lack of randomized, clinical trials, and insufficient long-term follow-up. Additionally, most research has focused on children and adolescents. Therefore, large-scale, methodologically rigorous trials are urgently needed to investigate the effectiveness of VR-based interventions for autistic adults to draw more firm conclusions.

Objectives

The aim is to conduct the hitherto largest clinical randomized trial globally, comparing VR-based social cognitive training (VRSCT) + treatment as usual (TAU) to TAU for autistic adults. The objective is to examine whether VRSCT is effective in improving psychosocial functioning, quality of life, and co-occurring clinical symptoms and to reduce social cognitive deficits in autistic adults. Additionally, we will examine whether the VRSCT is cost-effective.

We hypothesize the following:

1. VRSCT+TAU will be superior to TAU in improving psychosocial functioning, quality of life, co-occurring clinical symptoms, and reducing social cognitive deficits in autistic adults.
2. VRSCT will be cost-effective in improving daily life functioning and reducing social cognitive deficits in autistic adults.

XSL•FO
RenderX

Trial Design

STEPS (Social Cognitive Training Enhancing Pro-Functional Skills) is a randomized, assessor-blinded, parallel-group superiority clinical trial. A total of 140 participants will be allocated to receive either VRSCT+TAU or TAU alone. Assessments will be carried out at baseline, at cessation of intervention (3 months post baseline), and at 6 months post baseline. A flow diagram of the trial is presented in the Results section. A stratified block randomization with a concealed randomization sequence will be conducted with a 1:1 allocation. Independent assessors, blinded to the treatment, will evaluate outcomes. The analysis of outcomes will be carried out according to the intention-to-treat principle [36].

Methods

Study Setting

The trial will enroll participants from public and private outpatient facilities in Denmark that specialize in the clinical assessment and diagnosis of autism in adulthood. The clinical staff at these sites, including nurses, psychologists, doctors, and social workers, will approach relevant participants and refer them to the project group. Additionally, relevant participants from community-based social support services in the Capital Region of Denmark will be referred to the trial by the multidisciplinary staff at these locations. Participants will also be recruited via TrialTree, an online recruitment platform that specializes in connecting interested participants with targeted and relevant clinical research projects. Furthermore, self-referrals will be accepted, provided that the eligibility criteria for participation are met. Assessment, randomization, experimental intervention, and analysis of outcomes will be carried out by the VIRTU Research Group located at Mental Health Center Copenhagen, 2900 Hellerup, Denmark.

Eligibility Criteria

Participants must provide written, informed consent prior to the initiation of any assessment or intervention in the STEPS trial. Inclusion and exclusion criteria are listed in Textbox 1.

Textbox 1. Inclusion and exclusion criteria of the STEPS (Social Cognitive Training Enhancing Pro-Functional Skills) trial.

Inclusion criteria

- Age ≥18 years
- Diagnosis of autism (*ICD-10* [*International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*]: F84.0, F84.1, F84.5, and F84.8)
- T-score of ≥60 on the Social Responsiveness Scale Total Score, Second Edition, Adult Form, Self-report

Exclusion criteria

- A diagnosis of organic brain disease (*ICD-10*: F00-F09)
- Intellectual disability (IQ <70)
- Inadequate command of spoken Danish or English for engaging in the assessments and intervention
- Co-occurring attention-deficit/hyperactivity disorder with medically untreated symptoms
- Immediate risk of homicide or suicide

The study is conducted in Denmark, where the *ICD-10* (*International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* [37], remains the current standard for diagnostic classification in clinical practice. As no autism diagnostic assessment will be conducted as part of the trial, participants are required to provide documentation confirming a clinical diagnosis of autism based on *ICD-10* criteria (*ICD-10* codes, namely F84.0, F84.1, F84.5, and F84.8). Diagnostic assessments based on autism-specific instruments, for example, Adult Asperger Assessment [38], Autism Diagnostic Observation Schedule, Second Edition (ADOS-2 [39]), and Autism Diagnostic Interview - Revised (ADI-R [40]), preferably supplemented by comprehensive developmental history and information from multiple informants (eg, self-report and informant-report) and from different modalities (eg, observation, interviews, and questionnaires), as recommended for improving autism diagnostic procedure, especially relevant for girls and women [41], are accepted. In cases where a diagnosis was confirmed without using formal assessment tools, documentation will be accepted only if the evaluation was conducted by an experienced clinician and is adequately documented.

To ensure participants experience significant deficiencies in social interactions, the Social Responsiveness Scale, Second Edition, Adult Form, Self-report (SRS-2A) will be used to assess the level of social impairments. A sex-normed T-score of ≥60 on the SRS-2A total score was chosen as the cutoff since a T-score of 60 is the lower threshold for mild deficiencies in social behaviors that are clinically significant and may lead to mild or moderate difficulties in daily social interactions [42]. Estimated full-scale IQ will be estimated using the 2 subtests, Vocabulary and Matrix Reasoning, from the Danish version of the Wechsler Adult Intelligence Scale, Fourth Edition (WAIS-IV) [43], which are known to correlate strongly with full-scale IQ [44]. Participants with co-occurring attention-deficit/hyperactivity disorder (ADHD) must demonstrate a preliminary satisfactory treatment response to ADHD medication at the time of enrollment. This determination will be based on mutual agreement between the participant and the clinician responsible for their ADHD treatment, confirming that targeted ADHD symptoms are well managed. Additionally,

XSL•FO
RenderX

no changes in the type or dosage of ADHD medication are permitted within the 4 weeks preceding enrollment in the trial. The type and dosage of ADHD medicine will be thoroughly mapped and documented at enrollment, and at follow-up assessments. Further details of this criterion are displayed in the "Relevant Concomitant Care and Interventions Permitted or Prohibited During the Trial" section. Additionally, general psychopathology will be assessed with the Danish version of the Mini-International Neuropsychiatric Interview (MINI) [45].

Interventions

Explanation for the Choice of Comparator

This trial compares the effects of VRSCT+TAU to TAU alone offered to autistic adults in Denmark. TAU usually includes short-term psychoeducation provided by specialized facilities for autistic adults, following psychiatric assessment and diagnosis. Additionally, TAU often involves supportive counselling or practical assistance in daily life situations, delivered in community settings. The exact content and duration of TAU received by each participant will be mapped and documented upon trial completion, enabling examination of potential between-group differences regarding the content and duration of TAU. This definition of TAU as a comparator was chosen, although its specific components are not strictly defined, because it represents the most common intervention currently offered to autistic adults in Denmark. To accurately assess the efficacy of VRSCT, it is essential to compare it against a standard treatment that represents routine, real-world care provided to this population.

Intervention Description

Participants in the experimental group will receive 12 weekly, 1-hour sessions of VRSCT, with the primary aim of enhancing psychosocial functioning and social cognition through exposure to social environments in VR. Figure 1 displays an example of the VRSCT in STEPS.

Relatives are invited to participate in the intervention. Participation of relatives is dependent on the participant's consent. Relatives will participate in 1 session at the beginning of the intervention to add information to the assessment of the participants' challenges. Relatives will also participate in 1 session at the end of the intervention to support transfer to everyday life. If participants decline involvement of relatives, this will be registered, and the intervention will consist of 10 individual sessions. This 12-session format is in line with previous positive findings of other short-term interventions (10 - 14 sessions) targeting social cognition in autistic adults [46,47].

Figure 1. Virtual reality–based social cognitive training in STEPS (Social Cognitive Training Enhancing Pro-Functional Skills); a participant (portrayed by a colleague from the VIRTU Research Group) practices conversing with a stranger during a virtual bus ride.



XSL•FO
RenderX

The VRSCT in the trial is a manualized cognitive behavioral therapy–based intervention targeting social cognition. The intervention consists of 3 modules, namely emotions, social understanding (including ToM), and complex social interactions. Within these modules, the intervention is highly personalized to meet the specific needs of each participant. The intervention uses the Social Worlds VR software developed by CleVR. Originally developed for psychotic disorders, the Social Worlds software has been adapted to meet the needs of other clinical populations, including personality disorders and autism spectrum conditions [48]. The software enables exposure to different social environments, such as a bus, café, street, supermarket, park, classroom, waiting room, home environment, and an office setting. In these environments, social interactions and role-plays, facilitated by the therapist, can take place. Additionally, the software allows for perspective switching, enabling the therapist to record and replay role-plays the participant engages in. These interactions can be replayed from multiple perspectives (eg, from another person involved or an observer), allowing the participant to experience the situation from someone else's point of view and thereby stimulate ToM development. By addressing the participant's individual social interaction challenges through exposure to VR-based social situations and role-plays, key social cognitive functions can be targeted. These include the speed and accuracy of emotion identification in complex, life-like situations, distinguishing relevant from irrelevant social information, and ToM. Additionally, daily social communication and functioning are trained by engaging participants in virtual encounters with avatars (eg, initiating conversations, engaging in small talk, and navigating misunderstandings and setting boundaries). The *S*ocial Worlds software allows for a highly individualized and tailored approach by having various daily life scenarios and 100 different avatars. The highly modifiable social interactions in VR allow for addressing social cognitive impairments across a broad range of severities. During virtual engagements, the therapist facilitates cognitive behavioral therapy–based dialogues to enhance psychosocial functioning, develop new coping strategies, and reduce avoidant behavior, including social isolation. In cocreation with the participant, the therapist facilitates individually tailored dialogues and role-plays, based on the participant's goals and challenges, and selects the most appropriate avatars in terms of appearance and voice, rather than relying on preprogrammed scenarios or fixed tasks. The VR intervention was co-developed with autistic adults and is continuously tailored to individual goals, sensory needs, and communication styles, supporting skills and confidence in situations that participants themselves find meaningful. In this way, the intervention focuses on the goals of participants rather than enforcing conformity to neurotypical standards.

Fidelity

All sessions of VRSCT will be audiotaped, and a selected number of sessions will be rated by an independent rater to ensure fidelity to the treatment protocol. Fidelity will be assessed by an independent rater with clinical expertise in autism, using a manual developed specifically for the STEPS intervention and based on key elements from the Cognitive Therapy Scale - Revised (CTS-R [49]). A total of 10 participant courses will be randomly selected across the recruitment period (3 from the early phase, 4 from the middle phase, and 3 from the late phase) to ensure representation across all trial phases and across therapists delivering the intervention. From each selected course, 3 sessions (one from the early, one from the middle, and one from the late phase of the course) will be randomly chosen for fidelity rating, resulting in 30 rated sessions in total. Time spent in VR will gradually increase during the first sessions and continue to build up across the intervention. The exact time spent in VR will be documented for each session.

Comparison Group

Participants in the comparison group will receive TAU alone, which typically comprises short-term psychoeducation provided by the clinical staff at the psychiatric facilities for autistic adults, following psychiatric assessment and diagnosis. Additionally, supportive counselling or practical assistance in daily life situations can be provided in community settings by social workers.

Criteria for Discontinuing or Modifying Allocated Interventions

Participants can choose to discontinue the interventions at any given time without any impact on their future psychiatric care. If participants experience a deterioration in their mental health condition during the intervention, this will be considered, and the intervention will be adjusted accordingly. Likewise, follow-up assessments and individual sessions in the intervention may be postponed if participants are unable to attend due to hospitalization. Any psychiatric hospitalization and its extent will be mapped upon trial completion.

Strategies to Improve Adherence to Intervention Protocols

Exposure to challenging social environments is a key component of the experimental intervention. Participants may encounter social situations that can provoke discomfort and anxiety. To prevent overwhelming the participant and reduce the risk of dropout, it is essential that the intervention be tailored to each participant's optimal level of challenge. This ensures that difficult situations are addressed without exceeding their coping abilities. The therapists will maintain this balance through continuous dialogue with participants about their experiences and by promoting adherence to the intervention.

Relevant Concomitant Care and Interventions Permitted or Prohibited During the Trial

The prevalence of co-occurring psychiatric diagnoses is considered high among autistic adults, with ADHD (25.7%) being the most prevalent comorbid condition, followed by mood disorders (18.8%) and anxiety disorders (17.8%) [50]. The estimated prevalence of ADHD in the total autistic population is varying widely in the literature. Recent meta-analyses estimate the pooled point prevalence rate of ADHD in the autistic population to range from 28% to approximately 38% [51,52]. To enhance the ecological validity of the trial, co-occurring psychiatric conditions and associated medical treatments will not constitute exclusion criteria apart from medically untreated ADHD. As previously mentioned, participants with co-occurring ADHD must demonstrate a preliminary satisfactory treatment

XSL•FO
RenderX

response to ADHD medication at the time of enrollment. This determination will be based on mutual agreement between the participant and the clinician responsible for their ADHD treatment, confirming that targeted ADHD symptoms are well managed. Additionally, no changes in the type or dosage of ADHD medication are permitted within the 4 weeks preceding enrollment in the trial. This criterion is essential due to the relatively high prevalence of co-occurring ADHD in individuals on the autism spectrum. Further, many adults are often being assessed for (and diagnosed with) ADHD while undergoing evaluation for autism. As a result, many potential participants may be titrating ADHD medication while being referred to the trial. To be able to effectively distinguish the effects of the trial intervention from those of ADHD medication, this criterion is essential. Furthermore, all psychopharmacological treatments will be thoroughly mapped and documented upon trial completion.

Provisions for Ancillary and Posttrial Care

All participants have access to the Danish public health care system before, during, and after participation in the trial. The experimental intervention is considered an add-on treatment to the standard treatment. If posttrial care is needed, it will be provided by the health care system and will follow national standards. If participants experience a significant deterioration in their mental health during the trial, the project may initiate a referral for outpatient psychiatric follow-up and treatment. Additionally, participants in the trial are covered by the Danish Patient Compensation Association, which evaluates and covers injuries resulting from health care treatment.

Outcomes

Table 1 provides an overview of the outcome measures and their assessment time points.

XSL•FO
RenderX

Table . Schedule of enrollment, intervention, and assessments

Content	Enrollment	Baseline assessment	Allocation	Intervention	Follow-up assessment 3 months post baseline	Follow-up assessment 6 months post baseline
Time points		t_0			t_1	t_2
Information about the trial	✓					
Informed consent	✓					
Eligibility evaluation	✓					
SRS-2A[a]	✓					
Vocabulary and Matrix Reasoning (WAIS-IV[b])	✓					
MINI[c]	✓					
Randomization			✓			
Interventions						
Comparison group (TAU[d])				✓		
Experimental group (VRSCT[e]+TAU[d])				✓		
Assessments						
Primary outcome						
SRS-2A[a]		✓ (administered at enrollment)			✓	✓
Secondary outcomes						
HiSoC[f]		✓			✓	✓
SRS-2AI[g]		✓			✓	✓
SIAS[h]		✓			✓	✓
ERT[i] from CANTAB[j]		✓			✓	✓
EQ[k]		✓			✓	✓
Exploratory outcomes						
PSP[l]		✓			✓	✓
BRIEF-A[m]		✓			✓	✓
GSE[n]		✓			✓	✓
WHO-5[o]		✓			✓	✓
EQ-5D[p]		✓			✓	✓
SR-AS[q]		✓				
RBQ-3[r]		✓				
GAFS-8[s]		✓			✓	✓
CAT-Q[t]		✓			✓	✓
ASRS[u]		✓			✓	✓

XSL•FO
RenderX

Content	Enrollment	Baseline assessment	Allocation	Intervention	Follow-up assessment 3 months post baseline	Follow-up assessment 6 months post baseline
BDI-II[v]		✓			✓	✓
TASIT-S[w]		✓			✓	✓
DACOBS[x]		✓			✓	✓
SSPA[y]		✓			✓	✓
Other measures						
CSQ[z]					✓	✓
NEQ[aa]					✓	✓
MPS-MSDV[ab] (Only experimental group, 3 and 10 weeks post baseline)				✓		
SSQ[ac] (Only experimental group, 3 and 10 weeks post baseline)				✓		
RTQ[ad] (Only experimental group, 1 week post baseline)				✓		

XSL•FO
RenderX

Content	Enrollment	Baseline assessment	Allocation	Intervention	Follow-up assessment 3 months post baseline	Follow-up assessment 6 months post baseline
I-VAS[ae] (Only experimental group, 12 weeks post baseline)				✓		

[a]SRS-2A: Social Responsiveness Scale, Second Edition, Adult Form, Self-report.
[b]WAIS-IV: Wechsler Adult Intelligence Scale, Fourth Edition.
[c]MINI: Mini-International Neuropsychiatric Interview.
[d]TAU: treatment as usual.
[e]VRSCT: virtual reality-based social cognitive training.
[f]HiSoC: High-Risk Social Challenge task.
[g]SRS-2AI: Social Responsiveness Scale, Adult Form, Informant report.
[h]SIAS: Social Interaction Anxiety Scale.
[i]ERT: Emotion Recognition Task.
[j]CANTAB: Cambridge Neuropsychological Test Automated Battery.
[k]EQ: Empathy Quotient.
[l]PSP: Personal and Social Performance Scale.
[m]BRIEF-A: Behavior Rating Inventory of Executive Function - Adult Version, Self-Report.
[n]GSE: General Self-Efficacy Scale.
[o]WHO-5: WHO-5 Well-Being Index.
[p]EQ-5D: EuroQOL Five Dimensions Questionnaire.
[q]SR-AS: Sensory Reactivity in Autism Spectrum questionnaire .
[r]RBQ-3: Repetitive Behavior Questionnaire-3.
[s]GAFS-8: General Alexithymia Factor Score.
[t]CAT-Q: Camouflaging Autistic Traits Questionnaire.
[u]ASRS: Adult ADHD Self-Rating Scale.
[v]BDI-II: Beck Depression Inventory-II.
[w]TASIT-S: The Awareness of Social Inference Test short form.
[x]DACOBS: Davos Assessment of the Cognitive Biases Scale.
[y]SSPA: Social Skills Performance Assessment.
[z]CSQ: Client Satisfaction Questionnaire.
[aa]NEQ: Negative Effects Questionnaire.
[ab]MPS-MSDV: Multimodal Presence Scale, Modified Shortened Danish Version.
[ac]SSQ: Simulator Sickness Questionnaire.
[ad]RTQ: Readiness for Therapy Questionnaire.
[ae]I-VAS: Insight - Visual Analog Scale

Primary Outcome

The primary outcome is the change of social impairment from baseline to intervention cessation, measured at 3-month follow-up. At follow-up assessments, the observation period will comprise the last 2 weeks, instead of the previous 6 months, as at baseline assessment. The SRS-2A will be used to measure the level of social impairment [42]. The SRS-2A is a 65-item self-report rating scale designed to assess social impairments related to autism spectrum conditions [53]. SRS-2A consists of 65 items, scored on a 4-point Likert Scale (0 - 3). Each item is scaled from 0 (never true) to 3 (almost always true), generating a total score ranging from 0 to 195. The SRS-2A provides a continuous measure of social impairment, where a higher score reflects a greater degree of difficulties with social interactions in everyday life. The SRS-2 has been extensively used in autism research across different age groups and has demonstrated strong validity and reliability in both children [54,55] and adults [56-58].

Secondary Outcomes

Secondary outcomes will assess changes in performance-based social skills, informant-reported social functioning, social anxiety, emotion recognition, and empathy:

1. Social skills performance will be measured using the High-Risk Social Challenge (HiSoC) task, a standardized performance-based task assessing social interpersonal skills [59]. HiSoC scoring comprises 3 primary factors, namely affect, odd behavior and language, and social interpersonal skills. Additionally, an overall impression of the performance is rated. These 3 factors and the general impression are added to calculate the HiSoC total score.

XSL•FO
RenderX

2. Social impairment will be measured using the Social Responsiveness Scale, Adult Form, Informant report (SRS-2AI) [42]. The SRS-2AI, a 65-item rating scale, assesses the level of social impairment based on informant observation. Items are rated on a 4-point Likert Scale, from 0 (never true) to 3 (almost always true), generating a total score range of 0-195. The items, while similar in content to the self-report version, are adapted for the informant's perspective. The SRS-2AI provides a continuous measure of social impairment, with higher scores indicating more severe difficulties with social functioning in everyday life.
3. Social anxiety levels will be measured by the Social Interaction Anxiety Scale (SIAS) [60]. The SIAS consists of 20 items rated on a 5-point Likert scale, from 0 (not at all characteristic of me) to 4 (extremely characteristic of me). Items 5, 9, and 11 are positively worded items and require reverse scoring. The total SIAS, ranging from 0 to 80, is calculated by summing all item ratings, with higher scores indicating greater social anxiety experienced in social interaction.
4. Emotion recognition ability will be assessed using the Emotion Recognition Task (ERT) from the Cambridge Neuropsychological Test Automated Battery (CANTAB) [61]. The ERT outcome measures include the percentage and count of correctly and incorrectly identified emotions displayed in images of facial expressions, as well as overall response latencies. These measures can be assessed across individual emotions or as an aggregate across all emotions.
5. The level of empathy will be assessed using the Empathy Quotient (EQ), a 60-item questionnaire designed to measure empathy in adults with normal intelligence [62]. Each item is scored on a 4-point scale (strongly agree, slightly agree, slightly disagree, and strongly disagree), yielding a possible score from 0 to 80, with higher scores reflecting higher levels of empathy.

Exploratory Outcomes

Overview

Below, exploratory outcomes are grouped into the following categories: psychosocial functioning and quality of life, clinical symptoms, social cognition and performance, and other measures.

Psychosocial Functioning and Quality of Life

This will be measured using the following outcomes: (1) the Personal and Social Performance scale (PSP) [63] to assess daily life social functioning; (2) the Behavior Rating Inventory of Executive Function - Adult Version, Self-Report (BRIEF-A) [64,65] to assess everyday executive function; (3) the General Self-Efficacy Scale (GSE) [66] to assess perceived self-efficacy; (4) WHO-5 Well-Being Index (WHO-5) [67] to assess subjective, psychological well-being; and (5) the EuroQOL 5D questionnaire (EQ-5D) [68,69] to assess health-related quality of life and to conduct cost-effectiveness analysis.

Clinical Symptoms

This will be measured using the following outcomes: (1) Sensory Reactivity in Autism Spectrum (SR-AS) questionnaire [64] to assess sensory reactivity; (2) the Repetitive Behavior Questionnaire-3 (RBQ-3) [70] to assess restricted and repetitive behaviors and interests; (3) the General Alexithymia Factor Score (GAFS-8) [71] to assess alexithymia; (4) the Camouflaging Autistic Traits Questionnaire (CAT-Q) [72] to assess social camouflaging behaviors; (5) the Adult ADHD Self-Report Scale (ASRS) [73] to assess ADHD symptoms; and (6) the Beck Depression Inventory-II (BDI-II) [74] to assess depressive symptoms.

Social Cognition and Performance

This will be measured by the following outcomes: (1) the Awareness of Social Inference Test short form (TASIT-S), part 2 Social Inference, Minimal) [75,76] to assess ToM and complex emotion comprehension; (2) the Davos Assessment of the Cognitive Biases Scale (DACOBS) [77] to assess cognitive biases; and (3) the Social Skills Performance Assessment (SSPA) [78] to assess social communication skills (competence and appropriateness of conversational skills).

Other Measures

These measures related to participants' experiences of the intervention will be assessed using the following measures: (1) the Client Satisfaction Questionnaire (CSQ) [79] to assess satisfaction with the assigned intervention; the Negative Effects Questionnaire (NEQ) [80] to assess negative events and outcomes; (2) the Multimodal Presence Scale, Modified Shortened Danish Version (MPS-MSDV) [81] to assess involvement and immersion into the VR environments (only administered to participants in the experimental group); (3) the Simulator Sickness Questionnaire (SSQ) [82] to assess simulator sickness related to being in VR (only administered to participants in the experimental group); (4) the Readiness for Therapy Questionnaire (RTQ) [83] to assess attitudes and readiness for therapy aimed at resolving problems; and (5) an Insight-Visual Analog Scale (I-VAS) to assess level of insight into challenges related to social interaction. Participants and therapists will indicate, on a scale from 1 to 10, the extent to which the intervention has increased the participant's awareness of their challenges with social interaction, with 1 representing no increased awareness and 10 indicating the highest possible increase in awareness. I-VAS will be administered by the therapists in the intervention and only to participants in the experimental group upon completion of the intervention.

Sample Size and Power Calculation

The sample size calculation is based on detecting a between-group difference in the primary outcome, the SRS-2A, at cessation of the intervention (3 months post baseline). Previous studies have reported considerable variation in the SD of the social responsiveness scale total score when assessing social impairment in autistic adults [25,56,58]. As a result, we have not been able to determine the minimal clinically important difference. Therefore, we have based our sample size calculation on a medium effect size (Cohen d=0.5), which corresponds to a difference of 0.5 SD between the experimental group (VRSCT+TAU) and the comparison group (TAU). With 80% power and a 2-sided significance level of α=.05, 63 participants are required in each group to detect this difference. To account for an anticipated 10% dropout rate, we plan to recruit a total of 140 participants. Sample size calculation is based on a

XSL•FO
RenderX

medium effect size (Cohen *d*=0.5) in accordance with established practice in clinical trial design. However, it may not reflect the smallest change considered clinically meaningful. Clinical significance will therefore be interpreted alongside observed effect sizes and CIs in subsequent reporting to ensure meaningful interpretation of trial outcomes.

Recruitment

The project group has established strong collaborations with several public and private outpatient facilities in Denmark, which are specializing in the clinical assessment and diagnosis of autism in adulthood. Based on the collaboration agreement, the multidisciplinary staff at these sites will routinely approach relevant participants and refer them to the VIRTU Research Group. To foster collaboration and ensure ongoing referrals to the trial, staff from the VIRTU Research Group will regularly attend conferences and meetings at the various recruitment sites and invite clinicians to discuss potential participants.

Allocation

Sequence Generation

Participants who provide signed informed consent and meet the eligibility criteria will be randomly allocated to receive either TAU or VRSCT+TAU upon completion of the baseline assessment. Randomization will be performed using a centralized, automated, computer-generated randomization module in the online REDCAP (Research Electronic Data Capture; Vanderbilt University) system (refer to "Data Management" section). A stratified block randomization with a concealed randomization sequence will be used, with block sizes that are variable and unknown to the investigators. Randomization will be stratified by site and sex. The allocation of the randomized intervention will remain concealed to investigators and assessors until the statistical analyses of outcome data have been completed.

Concealment Mechanism

When a participant has completed the baseline assessment, the trial assessor, who is responsible for conducting the assessment, will email one of the trial therapists responsible for generating the allocation sequence. The assessor will then receive a confirmation email acknowledging receipt only. The trial therapist has exclusive access to the randomization module in REDCAP (refer to "Data Management" section), ensuring that assessors are blinded to intervention allocation.

Implementation

Allocation sequences are generated by one of the trial therapists. Enrollment and assignment to the interventions are carried out by the therapists who are conducting the interventions in the experimental group.

Blinding

Following the allocation to one of the interventions, the outcome assessors will be blinded to the participants' group assignment. Due to the nature of the intervention in the experimental group (VRSCT), participants, therapists conducting the intervention, and the clinicians managing TAU will not be blinded to the received interventions. Participants will be instructed prior to follow-up assessments not to disclose their allocation to the outcome assessors. To maintain blinding, assessments and interventions will be conducted at separate locations, ensuring that the assessors remain unaware of group assignment. Assessors will be instructed to notify their supervisors immediately if they believe blinding has been broken. If unblinding of the outcome assessor occurs, another outcome assessor blind to the allocation of intervention will conduct the subsequent assessments.

Data Collection

Assessment and Collection of Outcomes

Assessments will be conducted at baseline (t_0), 3 months post baseline (t_1), and 6 months post baseline (t_2). An overview of assessments and outcome measures is provided in Table 1. Outcome assessors will be psychologists who will receive comprehensive training before conducting assessments. To ensure consistency and maintain data quality, assessors will participate in monthly supervision sessions focused on the administration and scoring of assessments. The following rater-based assessments, HiSoC, PSP, and SSPA, will be video- and audio-recorded to enable subsequent scoring. To ensure high interrater reliability, a structured calibration procedure has been implemented. First, the initial 20 consecutive assessments of each rater-based measure will be scored through consensus by the 2 outcome assessors together with a senior researcher and clinical expert in autism. This process is used to establish a shared understanding of scoring criteria and interpretation. Subsequently, the next 10 consecutive assessments will be independently rated by both outcome assessors, and intraclass correlation coefficients will be calculated to evaluate interrater agreement. Provided that acceptable reliability is achieved (intraclass correlation coefficient ≥0.75) [84], the assessors will continue to rate independently. To prevent rater drift and ensure continued data quality throughout the trial, monthly consensus rating sessions will be conducted involving both assessors and the senior researcher. Further, the assessment of co-occurring psychopathology using the MINI will be scored through consensus. Diagnoses based on the MINI will be diagnosed according to both *ICD-10* and *ICD-11* (*International Classification of Diseases, Eleventh Revision*) criteria. All training, supervision, and reliability procedures will be overseen by a senior researcher and clinical expert in autism.

Plans to Promote Participant Retention and Complete Follow-Up

To enhance participant retention and ensure follow-up assessments, several initiatives will be taken. Participants will be contacted by phone approximately 1 week before the planned follow-up assessments, and travel expenses for public transportation will be compensated to reduce the financial burden associated with participation in the trial. Additionally, assessors can rearrange scheduled assessment times to accommodate participants' needs. In cases where participants leave the trial prematurely, efforts will be made to complete follow-up assessments. If participants are unable to complete the full assessment battery, priority will be given to obtaining the primary and secondary outcomes.

XSL•FO
RenderX

Data Management

Data from evaluation of eligibility and assessments during the trial will be entered directly into the electronic case report form using the data entry system REDCAP. REDCAP is an electronic data capture tool hosted by the Center for IT, Medical Technology and Telephony (CIMT) in the Capital Region of Denmark. When electronic entry is not possible, data will be collected on paper and later transferred into REDCAP. Data on paper will be stored locally and secured. Using REDCAP, surveys can securely be distributed and returned via E-boks, a public Danish digital mailing system that all citizens possess, reducing the risk of data loss and unauthorized access. Data for each participant will be linked to a unique serial number, with only assigned researchers having access to REDCAP. REDCAP includes a complete audit trail of data transactions, detailed user rights, and access control management, ensuring compliance with Danish legislation (the Act on Processing Personal Data). Research data will be exported from REDCap without personal identifiers. Data will be exported to well-known software packages (eg, SPSS [IBM Corp], SAS [SAS Institute], Stata [StataCorp LLC], and R [R Foundation for Statistical Computing]) and placed in logged folders on a network drive under the control of the CIMT. A data manager will ensure that all variables are properly defined with variable and value labels. All derived variables will be properly defined, and algorithms will be kept in special files. All data will be examined carefully to identify errors in data entry.

Statistical Analysis of Primary and Secondary Outcomes

Analyses will be performed on the intention-to-treat population for the primary and secondary end points, that is, analyzing all randomized participants according to the arm they were randomized to. The analysis method will be ANCOVA (analysis of covariance) of change from baseline with treatment as a factor and baseline value as a covariate at the corresponding postbaseline visit, that is, third month. Additionally, stratification variables and baseline values of social anxiety, alexithymia, and social function will be included as covariates. Treatment difference point estimates will be reported together with 95% CIs and *P* value. Missing data will be handled using multiple imputation. For the exploratory end points, a modified intention-to-treat set will be used, requiring at least 1 postbaseline measurement for the inclusion in the analysis for each end point. The analysis method will be the linear mixed effects model. Results will be reported for both postbaseline visits. Details of model specifications, imputation methods and assumptions, sensitivity analyses, and subgroup analyses will be provided in the forthcoming statistical analysis plan, if applicable. Results will be reported with a 95% CI and *P* value, using a 5% significance level. Baseline demographic and clinical characteristics will be summarized for each group, using descriptive statistics (eg, means, SDs, medians, and ranges for continuous variables and frequencies and proportions for categorical) to assess comparability between groups at baseline. Efficacy analysis will be conducted by a blinded and independent statistician with no contact with the participants in the trial. Reporting will follow CONSORT (Consolidated Standards of Reporting Trials) guidelines [85].

To address potential confounding from changes in psychotropic medication, we will conduct sensitivity analyses to explore whether medication changes influence the primary and secondary outcomes. Medication type and dosage are documented at both follow-up time points (3 and 6 months post baseline). To account for variability in TAU, we document the type, frequency, and duration of TAU received during two predefined periods (1) from enrollment to the 3-month follow-up and (2) from 3- to 6-month follow-up. If systematic between-group differences in TAU exposure emerge, summary TAU variables (eg, number of sessions and service type) will be included as covariates in a sensitivity analysis to adjust for potential confounding. Additionally, exploratory analyses may be conducted to examine whether specific aspects of TAU moderate the intervention effect. In addition, exploratory subgroup analyses will be conducted to examine whether selected baseline characteristics moderate the effect of the intervention on primary and secondary outcomes. Analyses will be based on several prespecified variables, including sex, alexithymia, social anxiety, depression, and ADHD.

Data Monitoring and Oversight

Trial Oversight and Management

The trial will be overseen by a Project Management Group (PMG). The PMG will comprise the principal investigator of the study, who is also the trial manager and daily leader of the trial, the principal therapist of the trial, and the supervisor of the assessors. The PMG will hold monthly meetings throughout the recruitment period. There will be no data monitoring committee. The PMG will oversee trial safety and the status of recruitment and retention.

Adverse Events

Although VR interventions are generally well tolerated with minimal to no severe side effects or adverse events [28], including among autistic adults [25], some individuals may still experience cybersickness, which is not considered to be serious. Therefore, we do not expect to encounter any serious side effects or adverse events. Side effects and adverse events will be monitored and recorded throughout the study, as well as any complaints about the intervention. The following are considered as adverse events: (1) hospital admissions, (2) suicide attempts, (3) any violent incident necessitating police involvement (whether survivor or accused), (4) self-harming behavior, and (5) all deaths. Any adverse events will be reported to the Committee on Health Research Ethics of the Capital Region Denmark.

Auditing Trial Conduct

The leading members of the PMG will continuously ensure that the trial is conducted according to protocol, that data is collected accurately and reliably, and that the safety and rights of the participants are protected. Any necessary revisions to the processes will be discussed at the monthly PMG meetings.

XSL•FO
RenderX

Ethical Considerations

Research Ethics Approval

The study was approved by the Committee on Health Research Ethics in the Capital Region of Denmark (Project ID H-23055504) and by the Danish Data Protection Agency at the Capital Region of Denmark (project ID p-2023 - 14488).

Protocol Amendments

Protocol amendments must be approved by the Committee on Health Research Ethics of the Capital Region Denmark. Any modifications to the protocol will be documented in the trial registration record (NCT06438536).

Consent

Trained psychologists from the VIRTU Research Group will obtain signed informed consent for participation in the STEPS trial. Referred participants will initially receive information about the trial written in layman's terms, covering the trial's background, objectives, design, intervention, randomization process, risks, confidentiality, and legal rights. Next, an appointment will be arranged with a trial researcher to ensure participants understand the information provided and accept random assignment to either VRCST+TAU or to TAU alone. If they decide to proceed, they will sign a consent form confirming they have received adequate written and verbal information about the trial's purpose, method, benefits, and risks, and agree to participate. Written, informed consent to participate will be obtained from all participants, and participants have the right to withdraw from the study at any time and for any reason. Written information about the trial aimed at the participants and the informed consent form is available from the corresponding author.

Privacy and Confidentiality

This protocol does not include any personal data or identifying images of participants, nor will such data be included in any subsequent publications reporting the trial results. All data presented in future papers will be fully anonymized to ensure participant confidentiality.

Dissemination

The results of the trial will be shared with the scientific community through presentations at national and international conferences. The findings will also be published in scientific papers submitted to relevant academic journals. Additionally, results will be communicated to trial participants. The implications of the results will be discussed with health care professionals and other stakeholders at clinical conferences. Furthermore, the findings will be presented to the public through meetings with interest groups and to the wider public at local meetings or seminars.

Compensation Details

Participants will not receive financial compensation. However, transportation expenses will be reimbursed, and meals will be provided on assessment days to ensure participant comfort and minimize any financial burden associated with participation.

Results

Funding was secured in October and November 2022. The first participant was enrolled in May 2024. Figure 2 provides a flow diagram of the trial. As of February 2025 (initial manuscript submission), 34 participants had been enrolled, increasing to 97 participants as of December 2025. Completion of enrollment is expected in April 2026. Data analysis is expected to begin in October 2026 following the final 6-month follow-up assessment. Results are anticipated in December 2026 and will be disseminated at international conferences and through peer-reviewed publications by summer 2027.

XSL•FO
RenderX

Figure 2. Flow diagram of the STEPS (Social Cognitive Training Enhancing Pro-Functional Skills) trial. TAU: treatment as usual; VRSCT: virtual reality-based social cognitive training.

XSL•FO
RenderX

Discussion

Anticipated Findings and Implications

Autistic adults constitute a growing and largely overlooked population, for whom few clinical and research resources have been allocated. Due to the limited evidence on effective interventions for autistic adults, no national clinical guidelines have been developed for psychological interventions aimed at this population in Denmark. Sparse evidence from profunctional studies conducted in England or Sweden is often referenced when attempting to guide psychological treatment approaches for this group [86]. To our knowledge, STEPS is the hitherto largest randomized clinical trial globally investigating the effect of VRSCT for autistic adults. The results of this innovative intervention approach may significantly advance research in the field of autism. If the VRSCT intervention in the STEPS trial proves effective, it has the potential to improve psychosocial functioning and quality of life, co-occurring clinical symptoms, and reduce social cognitive deficits in autistic adults. The intervention was co-developed in collaboration with autistic adults to ensure relevance and alignment with neuro-affirming principles. It is tailored to each participant's goals, sensory needs, and communication preferences and aims to enhance comfort, confidence, and agency in social situations that participants themselves identify as meaningful. Rather than imposing neurotypical norms of interaction, the approach supports self-defined outcomes and promotes the transfer of skills into everyday life in ways that respect and value neurodivergent ways of relating. Establishing evidence-based interventions is essential to address the debilitating psychosocial challenges encountered by this population, especially considering the absence of established gold-standard interventions. Enhancing profunctional psychosocial skills can significantly impact the lives of these individuals by enabling them to build more meaningful and intimate relationships, better maintain personal and professional relations, and improve educational and occupational functioning. Developing effective interventions for autistic adults can further contribute to filling the critical gap in available support and thereby contribute to alleviating the personal and social consequences and the economic impact associated with autism. If proven effective, the manualized VRSCT intervention from this study could be implemented in mental health services across Denmark and potentially on an international scale, supported by training and ongoing supervision. Additionally, the intervention could be adapted to fit the needs of other clinical populations experiencing social cognitive deficits, such as autistic adolescents, individuals with bipolar disorder, or individuals with psychosis [87]. This could help address the increasing demand for novel, effective interventions for psychiatric conditions globally.

Limitations

In this trial, eligibility was based on a clinical diagnosis of autism according to *ICD-10* criteria. Although widely used in clinical and research settings, the diagnostic criteria in *ICD-10* and *DSM-5-TR* have been criticized for overemphasizing observable behaviors and reflecting a prototypical male presentation of autism. These limitations can obscure how autism manifests differently across sex and age and may overlook how traits are masked by compensatory strategies, potentially contributing to the underdiagnosis of autism in females and the delayed identification of autism in adults [41,88]. While this may also affect the representativeness of the study sample by leading to fewer females being included, the final sex distribution will help determine the extent to which this concern is reflected in the sample.

The participants in the comparison group will receive TAU, administered across multiple sites. It is expected that there will be variations in the content and duration of TAU among participants. Although the specific details of TAU for each participant will be documented at the end of the trial, this variability remains a limitation. The inconsistency in TAU introduces less control compared to an active control uniformly administered in the same setting as the experimental intervention. However, TAU was chosen as the comparator because it reflects current clinical practice available for autistic adults in Denmark. Additionally, there is no established "gold standard" intervention specifically targeting social cognitive deficits in autistic adults, which otherwise could have been an obvious choice of comparator. Further, given the absence of a "gold standard" comparator and the choice of TAU as the control condition, it is important to note that the immersive and novel nature of the VR intervention may contribute to increased participant expectancy in the intervention group compared with TAU, which does not involve comparable technology. While a sham or low-dose VR control condition was considered to address this technology-expectancy bias, it was not feasible within the scope and resources of this study. To mitigate this risk, we included both a blinded, performance-based outcome measure (HiSoC) and an informant-rated version of the Social Responsiveness Scale (SRS-2AI) as secondary outcomes. This approach allows for triangulation across self-report, informant-report, and blinded assessment. Due to the exclusion criteria regarding autistic adults with medically untreated ADHD and autistic adults with intellectual disability, the results from this trial may not be generalizable to these populations. However, the findings may still provide useful indications of the potential effects for autistic adults with ADHD, and the intervention could eventually be adapted for autistic adults with intellectual disability. Finally, as the present trial does not include follow-up assessment beyond 6 months, the absence of longer-term follow-up (eg, at 12 months) limits conclusions regarding sustained effects. As the effects of VR-based interventions may diminish over time, and follow-up assessments beyond 3 months are often lacking [89], future trials should address this.

Conclusion

Autistic adults constitute a growing and largely overlooked population with limited evidence-based interventions available. To our knowledge, STEPS is the largest randomized clinical trial globally investigating the effect of VRSCT for autistic adults. If proven effective, VRSCT has the potential to improve psychosocial functioning and quality of life and reduce co-occurring clinical symptoms and social cognitive deficits in autistic adults, thereby potentially contributing to filling the critical gap in available support. The results of the trial should

XSL•FO
RenderX

be interpreted in light of key limitations, including heterogeneity in TAU, potential technology-related expectancy effects, and follow-up limited to 6 months.

In light of these considerations, the findings of the trial are expected to advance research in the field of autism and may inform future clinical practice and subsequent trials, including in other clinical populations experiencing social cognitive challenges.

Acknowledgments

The authors express their sincere gratitude to the study participants and the staff at the recruitment sites for their assistance in referring individuals to the trial. We extend our thanks to the focus group, consisting of autistic adults, with lived experiences, whose insights were invaluable in developing the intervention manual. Finally, we acknowledge and thank all contributors to the development of the STEPS (Social Cognitive Training Enhancing Pro-Functional Skills) trial who may not have been specifically mentioned here.

Funding

The trial is funded by the Lundbeck Foundation (ID R396-2022-88) and the Independent Research Fund Denmark (ID 2100-00012B). The funders did not take part in the study design. The funders will not participate in data collection, management, analysis, or interpretation; writing the report; or deciding to submit the report for publication.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ongoing data collection and because the dataset is not yet finalized, but will be deposited as a fully anonymized dataset with the Danish National Archives for scientific sharing purposes upon completion of the study. Queries regarding data access can be directed to the corresponding author.

Authors' Contributions

Conceptualization: LSM, RP, AEP, MN, JRMJ, LBG
Data curation: JA, ACEJ
Formal analysis: CH
Funding acquisition: LBG
Investigation: JA, ACEJ, ASD, LSM
Methodology: JA, ACEJ, ASD, LSM, RP, CH, AEP, GL, JRMJ, LBG
Project administration: JA, ACEJ, LBG
Resources: JA, ACEJ, LBG
Supervision: CH, JRMJ, LBG
Validation: JA, ACEJ, JRMJ
Visualization: JA
Writing – original draft: JA
Writing – review & editing: JA, ACEJ, ASD, LSM, RP, CH, AEP, MN, GL, JRMJ, LBG

Conflicts of Interest

AEP has received consulting fees and travel reimbursements from Boehringer Ingelheim (BI). BI has no role in the current work, and the information presented here is not related to any BI products or services.

Checklist 1
SPIRIT checklist.
[PDF File, 237 KB - resprot_v15i1e72854_app1.pdf]

References

1. Bottema-Beutel K, Kapp SK, Lester JN, Sasson NJ, Hand BN. Avoiding ableist language: suggestions for autism researchers. Autism Adulthood 2021 Mar 1;3(1):18-29. [doi: 10.1089/aut.2020.0014] [Medline: 36601265]
2. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR): American Psychiatric Association; 2022. [doi: 10.1176/appi.books.9780890425787]
3. Elsabbagh M, Divan G, Koh YJ, et al. Global prevalence of autism and other pervasive developmental disorders. Autism Res 2012 Jun;5(3):160-179. [doi: 10.1002/aur.239] [Medline: 22495912]
4. Zablotsky B, Black LI, Blumberg SJ. Estimated prevalence of children with diagnosed developmental disabilities in the United States, 2014–2016. NCHS Data Brief 2017 Nov(291):1-8. [Medline: 29235982]

XSL•FO
RenderX

5. Schendel DE, Thorsteinsson E. Cumulative incidence of autism into adulthood for birth cohorts in Denmark, 1980–2012. JAMA 2018 Nov 6;320(17):1811-1813. [doi: 10.1001/jama.2018.11328] [Medline: 30398592]
6. Cakir J, Frye RE, Walker SJ. The lifetime social cost of autism: 1990–2029. Res Autism Spectr Disord 2020 Apr;72:101502. [doi: 10.1016/j.rasd.2019.101502]
7. Camm-Crosbie L, Bradley L, Shaw R, Baron-Cohen S, Cassidy S. "People like me don't get support": Autistic adults' experiences of support and treatment for mental health difficulties, self-injury and suicidality. Autism 2019 Aug;23(6):1431-1441. [doi: 10.1177/1362361318816053] [Medline: 30497279]
8. Burke SL, Li T, Grudzien A, Garcia S. Brief report: improving employment interview self-efficacy among adults with autism and other developmental disabilities using virtual interactive training agents (ViTA). J Autism Dev Disord 2021 Feb;51(2):741-748. [doi: 10.1007/s10803-020-04571-8] [Medline: 32642959]
9. Cassidy S, Bradley P, Robinson J, Allison C, McHugh M, Baron-Cohen S. Suicidal ideation and suicide plans or attempts in adults with Asperger's syndrome attending a specialist diagnostic clinic: a clinical cohort study. Lancet Psychiatry 2014 Jul;1(2):142-147. [doi: 10.1016/S2215-0366(14)70248-2] [Medline: 26360578]
10. Cassidy S, Bradley L, Shaw R, Baron-Cohen S. Risk markers for suicidality in autistic adults. Mol Autism 2018;9(1):42. [doi: 10.1186/s13229-018-0226-4] [Medline: 30083306]
11. Hirvikoski T, Mittendorfer-Rutz E, Boman M, Larsson H, Lichtenstein P, Bölte S. Premature mortality in autism spectrum disorder. Br J Psychiatry 2016 Mar;208(3):232-238. [doi: 10.1192/bjp.bp.114.160192] [Medline: 26541693]
12. Sasson NJ, Morrison KE, Kelsven S, Pinkham AE. Social cognition as a predictor of functional and social skills in autistic adults without intellectual disability. Autism Res 2020 Feb;13(2):259-270. [doi: 10.1002/aur.2195] [Medline: 31419075]
13. Fett AKJ, Viechtbauer W, Dominguez MDG, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci Biobehav Rev 2011 Jan;35(3):573-588. [doi: 10.1016/j.neubiorev.2010.07.001] [Medline: 20620163]
14. Bölte S. Social cognition in autism and ADHD. Neurosci Biobehav Rev 2025 Feb;169:106022. [doi: 10.1016/j.neubiorev.2025.106022] [Medline: 39832687]
15. Frith U, Frith C. The social brain: allowing humans to boldly go where no other species has been. Philos Trans R Soc Lond B Biol Sci 2010 Jan 12;365(1537):165-176. [doi: 10.1098/rstb.2009.0160] [Medline: 20008394]
16. Schmidt L, Kirchner J, Strunz S, et al. Psychosocial functioning and life satisfaction in adults with autism spectrum disorder without intellectual impairment. J Clin Psychol 2015 Dec;71(12):1259-1268. [doi: 10.1002/jclp.22225] [Medline: 26406481]
17. Mason D, McConachie H, Garland D, Petrou A, Rodgers J, Parr JR. Predictors of quality of life for autistic adults. Autism Res 2018 Aug;11(8):1138-1147. [doi: 10.1002/aur.1965] [Medline: 29734506]
18. Bishop-Fitzpatrick L, Mazefsky CA, Eack SM. The combined impact of social support and perceived stress on quality of life in adults with autism spectrum disorder and without intellectual disability. Autism 2018 Aug;22(6):703-711. [doi: 10.1177/1362361317703090] [Medline: 28666391]
19. Uljarevic M, Hamilton A. Recognition of emotions in autism: a formal meta-analysis. J Autism Dev Disord 2013 Jul;43(7):1517-1526. [doi: 10.1007/s10803-012-1695-5] [Medline: 23114566]
20. Sasson NJ, Pinkham AE, Carpenter KLH, Belger A. The benefit of directly comparing autism and schizophrenia for revealing mechanisms of social cognitive impairment. J Neurodev Disord 2011 Jun;3(2):87-100. [doi: 10.1007/s11689-010-9068-x] [Medline: 21484194]
21. Keshavan MS, Vinogradov S, Rumsey J, Sherrill J, Wagner A. Cognitive training in mental disorders: update and future directions. Am J Psychiatry 2014 May;171(5):510-522. [doi: 10.1176/appi.ajp.2013.13081075] [Medline: 24700194]
22. Bishop-Fitzpatrick L, Minshew NJ, Eack SM. A systematic review of psychosocial interventions for adults with autism spectrum disorders. J Autism Dev Disord 2013 Mar;43(3):687-694. [doi: 10.1007/s10803-012-1615-8] [Medline: 22825929]
23. Odom SL, Thompson JL, Hedges S, et al. Technology-aided interventions and instruction for adolescents with autism spectrum disorder. J Autism Dev Disord 2015 Dec;45(12):3805-3819. [doi: 10.1007/s10803-014-2320-6] [Medline: 25468409]
24. Kandalaft MR, Didehbani N, Krawczyk DC, Allen TT, Chapman SB. Virtual reality social cognition training for young adults with high-functioning autism. J Autism Dev Disord 2013 Jan;43(1):34-44. [doi: 10.1007/s10803-012-1544-6] [Medline: 22570145]
25. van Pelt BJ, Nijman SA, van Haren NEM, et al. Dynamic interactive social cognition training in virtual reality (DiSCoVR) for adults with autism spectrum disorder: a feasibility study. Res Autism Spectr Disord 2022 Aug;96:102003. [doi: 10.1016/j.rasd.2022.102003]
26. Carnett A, Neely L, Gardiner S, Kirkpatrick M, Quarles J, Christopher K. Systematic review of virtual reality in behavioral interventions for individuals with autism. Adv Neurodev Disord 2023 Sep;7(3):426-442. [doi: 10.1007/s41252-022-00287-1]
27. Veling W, Moritz S, van der Gaag M. Brave new worlds--review and update on virtual reality assessment and treatment in psychosis. Schizophr Bull 2014 Nov;40(6):1194-1197. [doi: 10.1093/schbul/sbu125] [Medline: 25193975]
28. Rus-Calafell M, Garety P, Sason E, Craig TJK, Valmaggia LR. Virtual reality in the assessment and treatment of psychosis: a systematic review of its utility, acceptability and effectiveness. Psychol Med 2018 Feb;48(3):362-391. [doi: 10.1017/S0033291717001945] [Medline: 28735593]

XSL•FO
RenderX

29. Thompson A, Gleeson J, Alvarez-Jimenez M. Should we be using digital technologies in the treatment of psychotic disorders? Aust N Z J Psychiatry 2018 Mar;52(3):225-226. [doi: 10.1177/0004867418757920] [Medline: 29424235]
30. Kinnaird E, Stewart C, Tchanturia K. Investigating alexithymia in autism: a systematic review and meta-analysis. Eur Psychiatry 2019 Jan;55:80-89. [doi: 10.1016/j.eurpsy.2018.09.004] [Medline: 30399531]
31. Freeman D, Bradley J, Antley A, et al. Virtual reality in the treatment of persecutory delusions: randomised controlled experimental study testing how to reduce delusional conviction. Br J Psychiatry 2016 Jul;209(1):62-67. [doi: 10.1192/bjp.bp.115.176438] [Medline: 27151071]
32. Percie du Sert O, Potvin S, Lipp O, et al. Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: a pilot clinical trial. Schizophr Res 2018 Jul;197:176-181. [doi: 10.1016/j.schres.2018.02.031] [Medline: 29486956]
33. Georgescu AL, Kuzmanovic B, Roth D, Bente G, Vogeley K. The use of virtual characters to assess and train non-verbal communication in high-functioning autism. Front Hum Neurosci 2014;8:807. [doi: 10.3389/fnhum.2014.00807] [Medline: 25360098]
34. Tan BL, Shi J, Yang S, et al. The use of virtual reality and augmented reality in psychosocial rehabilitation for adults with neurodevelopmental disorders: a systematic review. Front Psychiatry 2022;13:1055204. [doi: 10.3389/fpsyt.2022.1055204] [Medline: 36590624]
35. Shahmoradi L, Rezayi S. Cognitive rehabilitation in people with autism spectrum disorder: a systematic review of emerging virtual reality-based approaches. J Neuroeng Rehabil 2022 Aug 18;19(1):91. [doi: 10.1186/s12984-022-01069-5] [Medline: 35982460]
36. Gupta SK. Intention-to-treat concept: a review. Perspect Clin Res 2011 Jul;2(3):109-112. [doi: 10.4103/2229-3485.83221] [Medline: 21897887]
37. Bertelsen A. Psykiske Lidelser Og Adfærdsmæssige Forstyrrelser: Klassifikation Og Diagnostiske Kriterier [Book in Danish], 1st edition: Munksgaard; 2004. URL: https://munksgaard.dk/products/psykiske-lidelser-og-adfardsmassige-forstyrrelser-bog-15975-9788762800427 [accessed 2025-12-17]
38. Baron-Cohen S, Wheelwright S, Robinson J, Woodbury-Smith M. The Adult Asperger Assessment (AAA): a diagnostic method. J Autism Dev Disord 2005 Dec;35(6):807-819. [doi: 10.1007/s10803-005-0026-5] [Medline: 16331530]
39. Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop SL. Autism Diagnostic Observation Schedule, Second Edition (ADOS-2) Manual (Part 1) Modules 1–4: Western Psychological Services; 2012.
40. Rutter M, Le Couteur A, Lord C. Autism Diagnostic Interview-Revised. Western Psychological Services. 2003. URL: https://www.wpspublish.com/adi-r-autism-diagnostic-interviewrevised.html [accessed 2025-12-17]
41. Cook J, Hull L, Mandy W. Improving diagnostic procedures in autism for girls and women: a narrative review. Neuropsychiatr Dis Treat 2024;20:505-514. [doi: 10.2147/NDT.S372723] [Medline: 38469208]
42. Bruni TP. Test review: Social Responsiveness Scale–Second Edition (SRS-2). J Psychoeduc Assess 2014 Jul;32(4):365-369. [doi: 10.1177/0734282913517525]
43. Wechsler D. Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV): Technical and Interpretive Manual: Pearson; 2008. URL: https://www.pearsonassessments.com/store/en/usd/p/100000392.html [accessed 2025-12-17]
44. Axelrod BN. Validity of the Wechsler abbreviated scale of intelligence and other very short forms of estimating intellectual functioning. Assessment 2002 Mar;9(1):17-23. [doi: 10.1177/1073191102009001003] [Medline: 11911230]
45. Sheehan DV, Lecrubier Y, Sheehan KH, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20: 22-33:quiz. [Medline: 9881538]
46. Bölte S, Feineis-Matthews S, Leber S, Dierks T, Hubl D, Poustka F. The development and evaluation of a computer-based program to test and to teach the recognition of facial affect. Int J Circumpolar Health 2002;61 Suppl 2:61-68. [doi: 10.3402/ijch.v61i0.17503] [Medline: 12585821]
47. Gantman A, Kapp SK, Orenski K, Laugeson EA. Social skills training for young adults with high-functioning autism spectrum disorders: a randomized controlled pilot study. J Autism Dev Disord 2012 Jun;42(6):1094-1103. [doi: 10.1007/s10803-011-1350-6] [Medline: 21915740]
48. CleVR. URL: https://www.clevr.net/ [accessed 2024-10-06]
49. Blackburn IM, James IA, Milne DL, et al. The revised Cognitive Therapy Scale (CTS-R): psychometric properties. Behav Cogn Psychother 2001 Oct;29(4):431-446. [doi: 10.1017/S1352465801004040]
50. Lugo-Marín J, Magán-Maganto M, Rivero-Santana A, et al. Prevalence of psychiatric disorders in adults with autism spectrum disorder: a systematic review and meta-analysis. Res Autism Spectr Disord 2019 Mar;59:22-33. [doi: 10.1016/j.rasd.2018.12.004]
51. Rong Y, Yang CJ, Jin Y, Wang Y. Prevalence of attention-deficit/hyperactivity disorder in individuals with autism spectrum disorder: a meta-analysis. Res Autism Spectr Disord 2021 May;83:101759. [doi: 10.1016/j.rasd.2021.101759]
52. Micai M, Fatta LM, Gila L, et al. Prevalence of co-occurring conditions in children and adults with autism spectrum disorder: a systematic review and meta-analysis. Neurosci Biobehav Rev 2023 Dec;155:105436. [doi: 10.1016/j.neubiorev.2023.105436] [Medline: 37913872]
53. Constantino JN, Gruber CP. Social Responsiveness Scale Second Edition (SRS-2): Manual: Western Psychological Services; 2012. URL: https://www.wpspublish.com/srs-2-social-responsiveness-scale-second-edition.html [accessed 2025-12-17]

XSL•FO
RenderX

54. Constantino JN, Davis SA, Todd RD, et al. Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. J Autism Dev Disord 2003 Aug;33(4):427-433. [doi: 10.1023/a:1025014929212] [Medline: 12959421]
55. Bölte S, Poustka F, Constantino JN. Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). Autism Res 2008 Dec;1(6):354-363. [doi: 10.1002/aur.49] [Medline: 19360690]
56. Bölte S. Brief Report: the Social Responsiveness Scale for Adults (SRS-A): initial results in a German cohort. J Autism Dev Disord 2012 Sep;42(9):1998-1999. [doi: 10.1007/s10803-011-1424-5] [Medline: 22183423]
57. Chan W, Smith LE, Hong J, Greenberg JS, Mailick MR. Validating the social responsiveness scale for adults with autism. Autism Res 2017 Oct;10(10):1663-1671. [doi: 10.1002/aur.1813] [Medline: 28639377]
58. Takei R, Matsuo J, Takahashi H, Uchiyama T, Kunugi H, Kamio Y. Verification of the utility of the social responsiveness scale for adults in non-clinical and clinical adult populations in Japan. BMC Psychiatry 2014 Nov 18;14(1):302. [doi: 10.1186/s12888-014-0302-z] [Medline: 25403232]
59. Gibson CM, Penn DL, Prinstein MJ, Perkins DO, Belger A. Social skill and social cognition in adolescents at genetic risk for psychosis. Schizophr Res 2010 Sep;122(1-3):179-184. [doi: 10.1016/j.schres.2010.04.018] [Medline: 20570111]
60. Mattick RP, Clarke JC. Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. Behav Res Ther 1998 Apr;36(4):455-470. [doi: 10.1016/s0005-7967(97)10031-6] [Medline: 9670605]
61. Sahakian BJ, Owen AM. Computerized assessment in neuropsychiatry using CANTAB: discussion paper. J R Soc Med 1992 Jul;85(7):399-402. [Medline: 1629849]
62. Baron-Cohen S, Wheelwright S. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. J Autism Dev Disord 2004 Apr;34(2):163-175. [doi: 10.1023/b:jadd.0000022607.19833.00] [Medline: 15162935]
63. Morosini PL, Magliano L, Brambilla L, Ugolini S, Pioli R. Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale (SOFAS) to assess routine social functioning. Acta Psychiatr Scand 2000 Apr;101(4):323-329. [doi: 10.1034/j.1600-0447.2000.101004323.x] [Medline: 10782554]
64. Elwin M, Schröder A, Ek L, Kjellin L. Development and pilot validation of a sensory reactivity scale for adults with high functioning autism spectrum conditions: Sensory Reactivity in Autism Spectrum (SR-AS). Nord J Psychiatry 2016;70(2):103-110. [doi: 10.3109/08039488.2015.1053984] [Medline: 26158770]
65. Jones CRG, Livingston LA, Fretwell C, et al. Measuring self and informant perspectives of Restricted and Repetitive Behaviours (RRBs): psychometric evaluation of the Repetitive Behaviours Questionnaire-3 (RBQ-3) in adult clinical practice and research settings. Mol Autism 2024 Jun 6;15(1):24. [doi: 10.1186/s13229-024-00603-7] [Medline: 38845057]
66. Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. J Psychol 2005 Sep;139(5):439-457. [doi: 10.3200/JRLP.139.5.439-457] [Medline: 16285214]
67. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176. [doi: 10.1159/000376585] [Medline: 25831962]
68. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
69. EQ-5D-5L. EuroQol. URL: https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-5l/ [accessed 2024-09-19]
70. Jones CRG, Livingston LA, Fretwell C, et al. Measuring self and informant perspectives of Restricted and Repetitive Behaviours (RRBs): psychometric evaluation of the Repetitive Behaviours Questionnaire-3 (RBQ-3) in adult clinical practice and research settgs. Mol Autism 2024 Jun 6;15(1):24. [doi: 10.1186/s13229-024-00603-7] [Medline: 38845057]
71. Williams ZJ, Gotham KO. Improving the measurement of alexithymia in autistic adults: a psychometric investigation of the 20-item Toronto Alexithymia Scale and generation of a general alexithymia factor score using item response theory. Mol Autism 2021 Aug 10;12(1):56. [doi: 10.1186/s13229-021-00463-5] [Medline: 34376227]
72. Hull L, Mandy W, Lai MC, et al. Development and Validation of the Camouflaging Autistic Traits Questionnaire (CAT-Q). J Autism Dev Disord 2019 Mar;49(3):819-833. [doi: 10.1007/s10803-018-3792-6] [Medline: 30361940]
73. Kessler RC, Adler L, Ames M, et al. The World Health Organization Adult ADHD Self-Report Scale (ASRS): a short screening scale for use in the general population. Psychol Med 2005 Feb;35(2):245-256. [doi: 10.1017/s0033291704002892] [Medline: 15841682]
74. Beck AT, Steer RA, Brown GK. BDI-II, Beck Depression Inventory: Manual: Psychological Corporation; 2019.
75. McDonald S, Flanagan S, Rollins J, Kinch J. TASIT: A new clinical tool for assessing social perception after traumatic brain injury. J Head Trauma Rehabil 2003;18(3):219-238. [doi: 10.1097/00001199-200305000-00001] [Medline: 12802165]
76. Honan CA, McDonald S, Sufani C, Hine DW, Kumfor F. The awareness of social inference test: development of a shortened version for use in adults with acquired brain injury. Clin Neuropsychol 2016 Feb;30(2):243-264. [doi: 10.1080/13854046.2015.1136691] [Medline: 26918817]
77. van der Gaag M, Schütz C, Ten Napel A, et al. Development of the Davos assessment of cognitive biases scale (DACOBS). Schizophr Res 2013 Mar;144(1-3):63-71. [doi: 10.1016/j.schres.2012.12.010] [Medline: 23332365]
78. Patterson TL, Moscona S, McKibbin CL, Davidson K, Jeste DV. Social skills performance assessment among older patients with schizophrenia. Schizophr Res 2001 Mar 30;48(2-3):351-360. [doi: 10.1016/s0920-9964(00)00109-2] [Medline: 11295387]

XSL•FO
RenderX

79. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann 1979;2(3):197-207. [doi: 10.1016/0149-7189(79)90094-6] [Medline: 10245370]
80. Rozental A, Kottorp A, Forsström D, et al. The Negative Effects Questionnaire: psychometric properties of an instrument for assessing negative effects in psychological treatments. Behav Cogn Psychother 2019 Sep;47(5):559-572. [doi: 10.1017/S1352465819000018] [Medline: 30871650]
81. Makransky G, Lilleholt L, Aaby A. Development and validation of the Multimodal Presence Scale for virtual reality environments: a confirmatory factor analysis and item response theory approach. Comput Human Behav 2017 Jul;72:276-285. [doi: 10.1016/j.chb.2017.02.066]
82. Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. Int J Aviat Psychol 1993 Jul;3(3):203-220. [doi: 10.1207/s15327108ijap0303_3]
83. Ghomi M, Wrightman M, Ghaemian A, Grey N, Pickup T, Richardson T. Development and validation of the Readiness for Therapy Questionnaire (RTQ). Behav Cogn Psychother 2020 Nov 17;49(4):1-13. [doi: 10.1017/S1352465820000764] [Medline: 33198833]
84. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
85. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. BMJ 2025 Apr 14;389:e081123. [doi: 10.1136/bmj-2024-081123] [Medline: 40228833]
86. Kaufmanas CH, Beyer J, Jørgensen MH, Kaas M. People with autism: social interventions that work. current knowledge to support the development and planning of municipal services [Report in Danish]. : Danish Social Services Agency (Socialstyrelsen) URL: https://www.sbst.dk/media/11550/Mennesker%20med%20autisme%20-%20sociale%20indsatser%20der%20virker_Socialstyrelsen_web.pdf [accessed 2025-12-17]
87. Deste G, Vita A, Nibbio G, et al. Autistic symptoms in people with schizophrenia: neurocognitive, socio-cognitive, clinical and real-world functional characteristics of individuals without autistic features. Schizophr Res 2021 Oct;236:12-18. [doi: 10.1016/j.schres.2021.07.037] [Medline: 34364032]
88. Lai MC, Baron-Cohen S. Identifying the lost generation of adults with autism spectrum conditions. Lancet Psychiatry 2015 Nov;2(11):1013-1027. [doi: 10.1016/S2215-0366(15)00277-1] [Medline: 26544750]
89. Bell IH, Pot-Kolder R, Rizzo A, et al. Advances in the use of virtual reality to treat mental health conditions. Nat Rev Psychol 2024 Jul 10;3(8):552-567. [doi: 10.1038/s44159-024-00334-9]

Abbreviations

AAA: Adult Asperger Assessment
ADHD: attention-deficit/hyperactivity disorder
ADI-R: Autism Diagnostic Interview - Revised
ADOS-2: Autism Diagnostic Observation Schedule, Second Edition
ANCOVA: analysis of covariance
ASRS: Adult ADHD Self-Report Scale
BDI-II: Beck Depression Inventory-II
BRIEF-A: Behavior Rating Inventory of Executive Function - Adult Version, Self-Report
CANTAB: Cambridge Neuropsychological Test Automated Battery
CAT-Q: Camouflaging Autistic Traits Questionnaire
CIMT: Center for IT, Medical Technology and Telephony
CONSORT: Consolidated Standards of Reporting Trials
CSQ: Client Satisfaction Questionnaire
CTS-R: Cognitive Therapy Scale - Revised
DACOBS: Davos Assessment of Cognitive Biases Scale
***DSM-5-TR*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision*
EQ: Empathy Quotient
ERT: Emotion Recognition Task
GAFS-8: General Alexithymia Factor Score
GSE: General Self-Efficacy Scale
HiSoC: High-Risk Social Challenge task
I-VAS: Insight-Visual Analog Scale
***ICD-10*:** *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*
***ICD-11*:** *International Classification of Diseases, Eleventh Revision*
MINI: Mini-International Neuropsychiatric Interview
MPS-MSDV: Multimodal Presence Scale, Modified Shortened Danish Version
NEQ: Negative Effects Questionnaire

XSL•FO
RenderX

PMG: Project Management Group
PSP: Personal and Social Performance Scale
RBQ-3: Repetitive Behavior Questionnaire-3
REDCAP: Research Electronic Data Capture
RTQ: Readiness for Therapy Questionnaire
SIAS: Social Interaction Anxiety Scale
SR-AS: Sensory Reactivity in Autism Spectrum
SRS-2A: Social Responsiveness Scale, Second Edition, Adult Form, Self-report
SRS-2AI: Social Responsiveness Scale, Adult Form, Informant report
SSPA: Social Skills Performance Assessment
SSQ: Simulator Sickness Questionnaire
STEPS: Social Cognitive Training Enhancing Pro-Functional Skills
TASIT-S: The Awareness of Social Inference Test - Short Form
TAU: treatment as usual
ToM: theory of mind
VR: virtual reality
VRSCT: virtual reality-based social cognitive training
WAIS-IV: Wechsler Adult Intelligence Scale, Fourth Edition
WHO-5: WHO-5 Well-Being Index

Edited by J Sarvestan; submitted 21.Feb.2025; peer-reviewed by B Bao, C Abbatantuono, F O'Donald; accepted 01.Oct.2025; published 05.Jan.2026.

Please cite as:
Andresen J, Jeppesen ACE, Due AS, Mariegaard LS, Parvaiz R, Hjorthøj C, Pinkham AE, Nordentoft M, Letnar G, Jepsen JRM, Glenthøj LB
The Effect of Virtual Reality–Based Social Cognitive Training for Autistic Adults: Protocol for STEPS (Social Cognitive Training Enhancing Pro-Functional Skills) Randomized Clinical Trial
JMIR Res Protoc 2026;15:e72854
URL: https://www.researchprotocols.org/2026/1/e72854
doi:10.2196/72854

© Johannes Andresen, Alberte C E Jeppesen, Anne Sofie Due, Lise Sandvig Mariegaard, Rizwan Parvaiz, Carsten Hjorthøj, Amy Elizabeth Pinkham, Merete Nordentoft, Gasper Letnar, Jens Richardt Møllegaard Jepsen, Louise Birkedal Glenthøj. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effects of Multi-Theory Model–Based Behavior Change Intervention with Staircase Approach on Sedentary Lifestyle Among Community-Dwelling Older Adults: Study Protocol for a Randomized Controlled Trial

Subinuer Tuerdi[1*], MMed; HuiLing Yang[2*], MMed; Li He[1], MMed; Rong Yan[1], MMed; YaoYi Cui[1], MMed; XingHui Wang[3], MMed; ShanShan Dong[4], MMed; JiaYu Yang[1], MMed; KeQiang Zhang[5], MMed; Feng Li[1], MD; YueWei Li[1], MD

[1]School of Nursing, Jilin University, 965 Xinjiang St, Changchun, China

[2]Operating Room, Second Affiliated Hospital of Jilin University, Changchun, Jilin Province, China

[3]Nursing Department, The First Hospital of Hebei Medical University, Hebei, China

[4]Hepatopancreatobiliary Surgery Department, General External Center, First Hospital of Jilin University, Changchun, Jilin Province, China

[5]Medical Department, Second Hospital of Jilin University, Changchun, Jilin Province, China

[*]these authors contributed equally

Corresponding Author:
YueWei Li, MD
School of Nursing, Jilin University, 965 Xinjiang St, Changchun, China

Abstract

Background: Sedentary lifestyles, as a nursing diagnosis, are prevalent in the life patterns of community-dwelling older adults, who have been shown to be the most sedentary and physically inactive subgroup. Prolonged low energy expenditure during waking hours leads to an increased risk of noncommunicable diseases and impairs physical functioning in older adults, negatively impacting their health outcomes. Therefore, interventions targeting changes in sedentary lifestyles are necessary to improve health behaviors in older adults.

Objective: This study aimed to investigate the effect of the multi-theory model–based stepped behavior change intervention on sedentary lifestyle among community-dwelling older adults.

Methods: This study is designed as a community-based, parallel-arm, assessor-blinded randomized controlled trial. Sixty participants were enrolled in this study and randomly assigned to the intervention group (received theory-based sedentary lifestyle change intervention) and the control group (received conventional behavioral change advice). End points were collected at baseline, immediately after the end of the intervention, week 12, and week 18. The primary endpoint is the change in self-reported sedentary time. Secondary endpoints include the changes in step count, time spent in light-intensity physical activity and moderate-to-vigorous-intensity physical activity, Measuring Change in Physical Activity Questionnaire score, Short Physical Performance Battery score, BMI, waist circumference, blood pressure, and Short-Form Health Survey-12-Version 2 score.

Results: The study was initiated in May 2024. As of May 2025, the project had completed participant recruitment and data collection. The final manuscript with findings is expected to be submitted for publication in September 2025.

Conclusions: This study uses a multi-theory model as its theoretical framework and adopts stepped sedentary lifestyle change as the intensity gradient of exercise behavior, creating a behavioral change pattern from sitting to standing and movement to light-intensity physical activity, to moderate-to-vigorous-intensity physical activity, and ultimately improving the sedentary lifestyle and obtaining health benefits, to provide community-dwelling older adults with individualized behavior change programs that are easy to adhere to and more applicable to daily activity patterns.

Trial Registration: China Clinical Trial Registry ChiCTR2400082225; https://www.chictr.org.cn/showproj.html?proj=223351

International Registered Report Identifier (IRRID): DERR1-10.2196/81284

(JMIR Res Protoc 2026;15:e81284) doi:10.2196/81284

KEYWORDS

sedentary behavior; sedentary lifestyle; physical activity; older adults; behavior change; MTM; lifestyle intervention; multi-theory model

XSL•FO
RenderX

Introduction

According to the United Nations World Population Prospects 2022 report, the average life expectancy at birth is expected to reach 77.2 years [1]. By the end of 2024, adults aged ≥65 years accounted for 15.6% of China's total population [2]. However, increased longevity is often accompanied by more health problems and age-related diseases. As people age, their risk of disability and noncommunicable diseases rises, placing a heavy burden on both individuals and health care systems [3]. Sedentary behavior (SB) and physical inactivity (PI) are the two most significant public health issues among the modifiable factors. Preventing noncommunicable diseases and promoting healthy aging by addressing these issues are top priorities for organizations such as the World Health Organization [4-6].

According to the World Health Organization Guidelines on Physical Activity and Sedentary Behavior, older adults should restrict their sedentary time (ST), replacing ST with any intensity of physical activity (PA) provides health benefits. To mitigate the negative health effects of excessive SB, older adults should aim to do more than the suggested levels of moderate-to-vigorous-intensity physical activity (MVPA) [7]. Despite the well-known benefits of being physically active, SB and PI are still prevalent among older adults in China, and a large proportion of older adults do not meet the essential amount of PA required to maintain their wellness [8,9]. The North American Nursing Diagnostic Association has termed such behavioral patterns characterized by low–energy-expenditure waking-time activities as sedentary lifestyles and classified them as nursing diagnoses in the field of health promotion [10]. Both SB and PI can be considered as a sedentary lifestyle according to its defining features. Older adults are the most sedentary and physically inactive subpopulation [11,12].

Research shows that even if some seniors meet the recommended 150 minutes of MVPA per week (non-PI), they may still be sedentary, which can have a negative impact on metabolism and overall health [13]. SB is defined as any waking behavior characterized by an energy expenditure of 1.5 METs or lower while sitting, reclining, or lying [14]. Statistics show that the average ST of Chinese older adults is 9.6 hours per day, indicating that they spend a large proportion of their waking hours in SB [8]. Prolonged sitting time can result in decreased physical function, raised depression and anxiety, and an increased risk of chronic disease and all-cause mortality, all of which can have a significant impact on older adults' health-related quality of life [15-21]. Therefore, effective interventions are needed to limit ST in older adults.

Common interventions targeting SB in community-dwelling older adults can be divided into two types: interventions that focus on reducing SB and approaches that target increased PA to reduce SB [22]. Research has shown that the former approach can lead to substantial reductions in ST [23]. Addressing prolonged SB independently of PA has many beneficial health outcomes, such as reduced risk of all-cause mortality and improved cardiometabolic risk factors [24]. However, these interventions may have several limitations. First, ST cannot be reduced endlessly. Findings from meta-analyses have suggested that total ST should be limited to less than 4 hours per day for those who are not physically active [25,26]. This goal may be difficult to achieve for the majority of older adults with chronic diseases, impaired cardiovascular health, or pain-related functional limitations [27]. Furthermore, sedentary habits in contexts such as reading and socializing should not be discouraged due to the mental health advantages and promotion of cognitive functioning in the aging process [28,29]. These limitations may also apply to other types of SB interventions. Second, the observed effects of SB reduction interventions in cardiometabolic biomarkers are generally inferior to those reported after exercise training interventions [30]. Interventions targeting the reduction of SB also include methods of interrupting ST—breaking and replacing ST with frequent brief periods of PA (ie, nonexercise daily living activities) to mitigate the adverse effects of prolonged SB [31]. Studies suggest that replacing ST with light-intensity physical activity (LPA) is associated with risk biomarkers such as waist circumference, high-density lipoprotein cholesterol, triglycerides, and insulin [32-34]. Where clinically meaningful improvements in these biomarkers occur, frequent and long-term interventions or a combination of PA interventions that replace SB with higher-intensity PA are needed [23].

For the second approach, the isotemporal substitution model posits that increasing the duration of one behavior leads to a decrease in the duration of one or more of the remaining behaviors. Therefore, it is feasible to increase an individual's PA duration to decrease their SB [35]. Robust findings still suggest that sedentary individuals may benefit from PA interventions [24,36]. High sitting time is significantly associated with an increased risk of all-cause mortality and cardiovascular disease, which is especially pronounced in physically inactive individuals (those who do not meet the recommended minimum level of >150 min MVPA/wk), and the increased risk of ST can be effectively attenuated by engaging in PA above the guideline-recommended level [26,36,37]. The second intervention, aimed at increasing PA, also has some limitations, with low adherence and insignificant effects on reducing SB [23,38]. First, to attenuate the harms of prolonged SB, 60 to 75 minutes per day of MVPA is required. However, when the majority of individuals do not meet the recommended guidelines (ie, 30 min/d of MVPA, 5 d/wk), an additional 30 to 45 minutes per day of MVPA is impractical, which affects adherence to the intervention [26,27]. Second, despite participating in regular PA, the adverse health impacts of prolonged ST can diminish the benefits of physical exercise [39,40]. In conclusion, focusing only on one of the two interventions may not produce the desired results. Thus, the optimal prescription for changing a sedentary lifestyle should be based on the interaction between SB and PA and to find a balance between the w elements to better promote a healthy lifestyle [25].

Dunstan et al [25] in their study "Sitting Less and Moving More" suggest that sedentary lifestyle can be addressed through a "staircase approach" (Figure 1). This is a promising but understudied approach. It involves modest transition steps: first emphasizing the reduction of SB, decreasing overall sitting time, and breaking up ST by increasing standing and activity;

XSL•FO
RenderX

then gradually increasing LPA to provide a "preparation base" for the long-term transition to higher-intensity PA, for example, walking, engaging in family leisure activities, flexibility training, strength training, etc; finally, gradually engaging in sufficient MVPA, such as brisk walking or jogging, dancing, tai chi, handball, water sports, etc, to meet the amount of PA above the guideline level [41]. Compared to accumulating weekly amounts of MVPA, interventions of the beginning stage aimed at reducing SB may seem more appealing and be easier for older adults to integrate into their daily lives [42]. Reducing SB may be an essential initial step toward implementing sustainable changes in movement patterns, which would support higher levels of overall PA and thus be beneficial to health.

Figure 1. The "sit less, move more" strategy is based on the staircase approach.

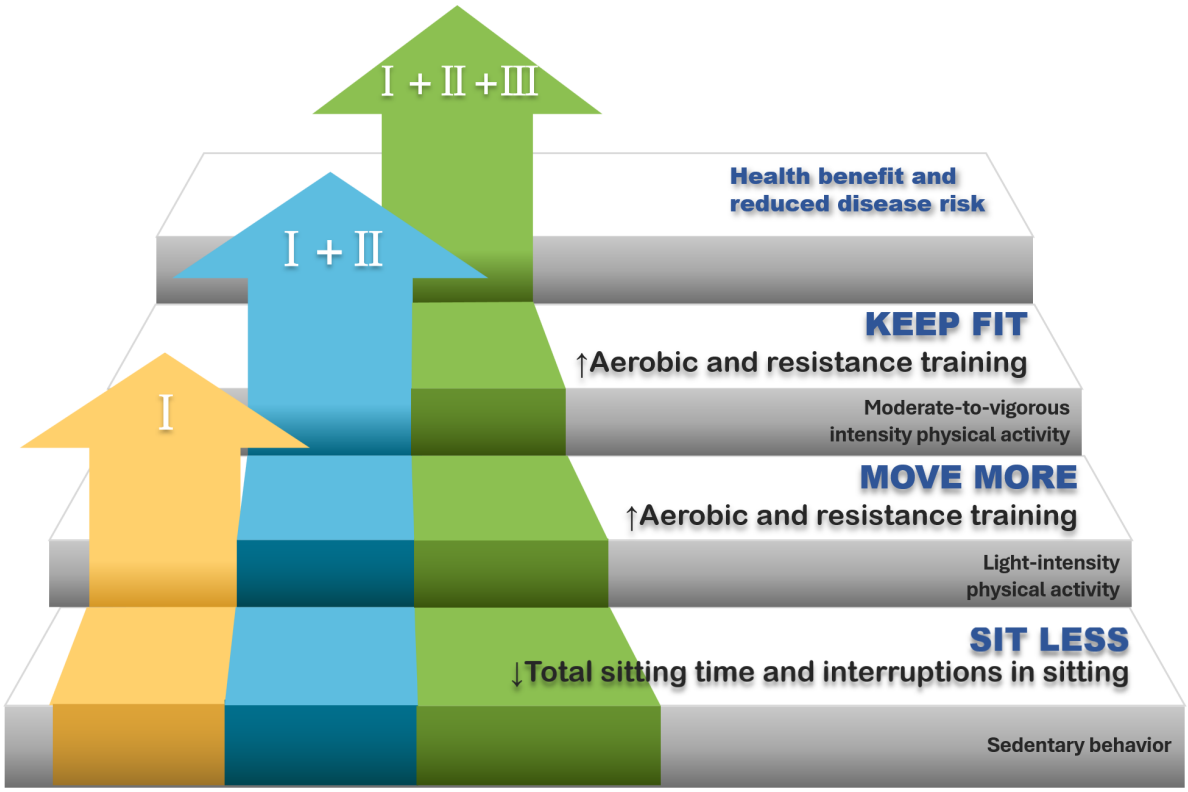

As older adults have more time to establish links between contextual cues and SB and because most of them engage in SB without being aware of it, they are considered to have the most potential to develop strong sedentary habits [43,44]. Therefore, breaking habitual SB and sustaining health behavior change over time is considered one of the greatest obstacles to this intervention. The multi-theory model (MTM) of health behavior change can be used to explain and predict the initiation and maintenance of health behavior change and can be applied at the individual, population, and community levels [45]. It consists of 2 components: the initiation and the maintenance of health behavior change. The initiation of health behavior change is the process of transitioning from one behavior to another, which means a one-time behavior change. It includes 3 components: participatory dialogue (interactive communication about the pros and cons of health behavior change), behavioral confidence (a way to gauge a person's level of certainty that they will participate in changing a health behavior in the future), and changes in physical environment (the environment in which the behavior is enacted). The maintenance of health behavior change is the same as the long-term performance of the health behavior change. It involves emotional transformation (which refers to overcoming self-doubt, inertia, and impulsivity, focusing one's feelings and emotions on health behavior change, directing oneself moving toward a goal), practice for change (which emphasizes reflective behavior, including continuously and prudently considering behavior change, combined with the ongoing correction to remove ineffective strategies and address barriers), and changes in social environment (building social support in the environment, using positive relationships to achieve desired behaviors). As a comprehensive theoretical model, MTM has been successfully applied in the initiation and maintenance of health behavior changes, such as PA, vaccination, and smoking and alcohol cessation [46-48].

Behavior change techniques (BCTs) are an observable, replicable, and nonreducible component of an intervention [49]. To evaluate interventions based on the MTM, multiple BCTs were selected to map onto the structure of the theoretical model. This study seeks to better describe and make transparent the intervention by combining the theoretical structure with the behavioral change techniques.

Using the theoretical framework of the MTM and the corresponding BCTs, a sedentary lifestyle intervention program for older adults based on the "staircase approach" was constructed. An experimental study will be conducted to verify whether a sedentary lifestyle modification intervention based on the MTM and the staircase approach is more effective than conventional education in reducing ST and improving PA levels,

XSL•FO
RenderX

physical function, anthropometric parameters, blood pressure, and quality of life among community-dwelling older adults.

Methods

Study Design

The study protocol adhered to the SPIRIT guidelines (Checklist 1) [50]. The Ethics Committee of the School of Nursing, Jilin University, approved the ethical clearance for the project (2023122001). The trial was registered in the China Clinical Trial Registry on March 25, 2024 (ChiCTR2400082225). This is a multicenter, assessor-blinded, 18-week, two-arm, parallel-group, randomized controlled trial examining the effects of an MTM-based sedentary lifestyle change intervention for older adults on SB, PA, and health outcomes compared with usual education.

Participants

Participants for this study were recruited through posters and leaflets distributed at 2 community health service centers in Changchun City, Jilin Province, China, from March 2024 to October 2024. All participants signed informed consent before the baseline outcome assessment. The following are the inclusion and exclusion criteria for participants:

Inclusion criteria:

1. Older adults aged ≥65 years.
2. Residents living in the community for ≥6 months.
3. Sedentary lifestyle (ie, more than 6 h of time awake spent in a sitting or reclining posture per day and PI: MVPA ≦150 min/wk) for at least 6 months.
4. Able to walk 100 m without assistance.
5. Informed consent and voluntary participation in this study.

Exclusion criteria:

1. Severe physical illness or long-term bedridden.
2. Unable or unwilling to communicate with researchers.
3. Participating in other PA or SB intervention programs.
4. Severe cardiovascular and cerebral vascular diseases that require restriction of activities.
5. Severe muscle weakness, skeletal diseases, or deformities that lead to restriction of PA.
6. Depression, schizophrenia, and other psychiatric disorders or dementia that led to severe cognitive impairment.
7. History of central nervous dysfunction, such as mild hemiplegia, spinal cord disease, and cerebral ataxia.

Study Interventions

Using the MTM as a framework and based on the review of relevant guidelines and literature, the sedentary lifestyle intervention program for community-dwelling older adults was developed (Table 1). Multimedia Appendix 1 shows the mapping of the sedentary lifestyle change interventions in the MTM, and Multimedia Appendix 2 shows the mapping of the sedentary lifestyle change interventions in the “staircase approach.” Participants in the intervention group will receive a sedentary lifestyle modification intervention based on the MTM and the staircase approach. Participants in the usual care group will receive a group education session (non-stepwise) on changing sedentary lifestyles and will be given a healthy behavior manual, an activity diary, and a pedometer.

XSL•FO
RenderX

Table . The multi-theory model (MTM)–based sedentary lifestyle intervention program for community-dwelling older adults.

Timing and interventions	MTM constructs	BCTs[a]	Content of interventions	Duration/frequency
Week 1 - 2, Phase I				
Discussion of advantages and disadvantages	Participatory dialogue/emotional transformation	• Pros and cons	• In small groups of 2 - 3 people, participants engage in face-to-face communication with health educators to discuss the advantages and disadvantages of sedentary lifestyles and to identify the main factors contributing to sedentary lifestyles, concerns regarding behavior change and the home environment in order to provide personalized advice.	20 min/group
Thematic session on sedentary behavior change	Behavioral confidence	• Information about health consequences • Credible sources • Instruction on how to perform the behavior • Problem solving	• This session will focus on the recognition and prevalence of SB[b], the harms of being sedentary, the benefits of reducing sitting time, etc • Peer sharing: sharing the adverse effects of long-term SB through the experiences of others, video cases, etc • Recommendations for reducing SB—interrupting or replacing SB with standing or LPA[c], household domain physical activity in different settings where it occurs, etc • Give personalized guidance on the main causes of SB for each older adult. Co-discuss the barriers that participants may encounter during the process of SB change and provide practical solutions (after discussion of advantages and disadvantages).	60 min

XSL•FO
RenderX

Timing and interventions	MTM constructs	BCTs[a]	Content of interventions	Duration/frequency
Healthy Behavior Manual	Behavioral confidence/changes in physical environment	• Information about health consequences • Demonstration of the behavior • Instruction on how to perform the behavior	• The manual contains the same information as presented in the two sessions: • Common patterns of SB, their negative effects, the advantages of reducing SB, and advice on how to break prolonged sitting. • The benefits of engaging in regular PA[d], tips on how to do it, such as aerobic, resistance, balance, and flexibility training, as well as the related graphic explanations.	One for each person
Sedentary behavior interruption intervention	Changes in physical environment/practice for change	• Prompts/cues • Behavior substitution	• Participants will receive the Cube Timer with a 5-15-25-45 min preset countdown function. It is a simple, easy-to-carry, and button-free time reminder tool. The researchers recommend that participants anticipate the duration of occurrences of SBs before they proceed with them and turn the selected time face up to start timing (choose less than 25 min if possible). When the time is up, the timer emits a continuous beep, indicating that the participant is getting up and moving around, and then the cycle of timekeeping is repeated, thus interrupting the SB. • Encourage participants to set at least one if-then planning for themselves and stick to it over the next few weeks, for example, stand when speaking on the phone, put the remote control in the distance when watching TV[e], get off the bus one stop earlier, stand during TV adverts, drink plenty of water to promote urination (if applicable), etc	Breaks the sitting time every 5/15/25/45 min
Week 3 - 4, Phase II				

XSL•FO
RenderX

Timing and interventions	MTM constructs	BCTs[a]	Content of interventions	Duration/frequency
Warning stickers	Changes in physical environment	• Prompts/cues	• The stickers were predesigned by the research team, including short and eye-catching sentences, and their content was developed based on motivational phrases and health-related proverbs that resonate well with older adults in China, aiming to effectively raise awareness of sedentary risks and encourage behavior change. Participants can post them where they regularly experience SB. Examples of the sticker messages include: “Flowing water never stagnates, a moving body stays strong,” “Walk a hundred steps after meals, live to ninety-nine,” “Feeling stuck to the chair? Stand up, take a walk, and break up with it!,” “Start your day with Baduanjin, end it with Tai Chi grace.”	Three stickers per person
Pedometer-based interventions	Changes in physical environment/emotional transformation/practice for change	• Self-monitoring of behavior • Graded tasks • Material reward		One for each person

XSL•FO
RenderX

Timing and interventions	MTM constructs	BCTs[a]	Content of interventions	Duration/frequency
			• Pedometers will be used for self-monitoring of PA and to record step counts. The intervention consisted of the following: • Educating participants on the benefits of walking and the necessity of maintaining a certain number of steps every day, • Participants will be instructed to record their daily step count in an activities diary, create a desired step goal, and review the objective to detect any gaps and regulate their activity level. The step goal is personalized to the participant. For older adults with less than 5000 steps at baseline, it is advised to increase the objective by at least 500 steps each week until it is maintained at more than 3000 steps from baseline. • Participants with high completion of the activity diary and who achieved and maintained the weekly target ST[f], step count, and PA were materially rewarded at the end of the intervention and at the end of the follow-up period.	
Activities diary	Emotional transformation/practice for change	• Self-monitoring of behavior • Action planning • Review behavior goals • Goal setting		One for each person

XSL•FO
RenderX

Timing and interventions	MTM constructs	BCTs[a]	Content of interventions	Duration/frequency
			• Participants can set weekly goals for sedentary time (min/day), step count (steps/day), and MV-PA[g] time (min/week) in the goal-setting table of the diary. Goals can be updated every 7 days. • Participants can record their daily sitting time (total time, the longest continuous sitting time, prolonged sedentary bouts) in the corresponding context of the SB recording form. • The diary includes a free recording section, where participants can record daily action plans, problems encountered, and experiences gained during the process of behavioral change. • The activity diary is also used to review the participant's behavioral goals, overcome self-inertia, focus on the individual's emotions on healthy behavior change, and direct them toward their goals. The activities diary emphasizes reflective behavior, continual critical consideration of behavior change, and removal of ineffective strategies.	
Week 5 - 6, Phase III				

XSL•FO
RenderX

Timing and interventions	MTM constructs	BCTs[a]	Content of interventions	Duration/frequency
Thematic sessions on physical activity	Behavioral confidence	• Information about health consequences • Demonstration of the behavior • Instruction on how to perform the behavior • Graded tasks	• This session highlights the following: • The significance of PA, the risks associated with PI[h], and the advantages of physical exercise for health • Forms of exercise, time, control of the intensity, conditions in which exercise should not be performed, warming up and stretching safely • The main points of exercise will be presented through text, diagrams, videos, and demonstrations, following the guidelines of the WHO[i] on physical activity and sedentary behavior and the expert consensus of the International Conference on Frailty and Sarcopenia Research. Through a combination of structured and integrated lifestyle activities, multicomponent exercise forms are recommended, including aerobic, plyometric, balance, and flexibility training (VIVIFRAIL). • To achieve optimal levels of adherence, it is recommended to start with a single form of exercise and allow older adults to gradually adapt to the new form of exercise before adding other components.	90 min
Introduction to health and fitness resources	Changes in physical environment	• Reconstructing the physical environment	• To improve recognition and utilization of health resources, participants will be introduced to local gym facilities, medical resources, and recreational facilities (stadiums, badminton courts, table tennis tables, etc) available in the neighborhood.	30 min
Social support	Changes in social environment	• Social reward • Reconstructing the physical environment		20 min

XSL•FO
RenderX

Timing and interventions	MTM constructs	BCTs[a]	Content of interventions	Duration/frequency
			• Contacting the patient's family members or partners and seeking their supervision and cooperation so that the patient can receive their encouragement and support. Call on community health service providers to join in the support of sedentary lifestyle change for the participants. • Publicize the benefits of sedentary lifestyle change by posters in order to dispel the community's concerns about PA in older adults.	
Week 7 - 18				
Telephone follow-up	Changes in social environment	• Problem solving	• During the follow-up period, two telephone consultations will be conducted to gather information on the implementation of behavioral change, provide advice to participants on how to overcome barriers to behavioral change, and encourage the continuation of good lifestyle habits.	20 min

[a]BCT: behavior change technique.

[b]SB: sedentary behavior.

[c]LPA: light-intensity physical activity.

[d]PA: physical activity.

[e]TV: television.

[f]ST: sedentary time.

[g]MVPA: moderate-to-vigorous-intensity physical activity.

[h]PI: physical inactivity.

[i]WHO: World Health Organization.

The program for the intervention group will follow the staircase approach in the rhythm of the reduction of ST and interruption of SB through standing and moving, followed by gradual replacement of SB with LPA, and ultimately leading to increased PA and modification of a sedentary lifestyle. During weeks 1 - 2, "Phase I: Sitting Less" will be implemented. This phase includes a discussion of advantages and disadvantages, thematic session on SB change, Healthy Behavior Manual, and SB interruption interventions. This component involves recognition of SB and reduction and interruption of SB through standing and activities. "Phase II: Moving More" will be implemented during weeks 3 - 4. The interventions will include warning stickers, pedometer-based interventions, and activities diary. This component involves the gradual replacement of SB with LPA. "Phase III: Increasing Physical Activity and Modifying Sedentary Lifestyle" will be implemented during weeks 5 - 6. The interventions will include thematic sessions on PA, introduction to health and fitness resources, and social support. This intervention aims to increase and maintain the amount of MVPA among participants. They will be encouraged to engage in all stages of the intervention, and data will also be incorporated if the participants stay at any stage of the staircase lifestyle change intervention, provided that they have attended phase I (weeks 1 - 2); otherwise, they are deemed to have withdrawn.

XSL•FO
RenderX

Study Measurements

End Points

The duration of the intervention is 6 weeks with 12 weeks of follow-up. Outcome measures were collected at baseline T1, end of intervention (week 6) T2, week 12 (T3), and week 18 (T4; Figure 2). The primary endpoint is the change in self-reported ST (total ST and SB characteristics, such as screen-based ST, longest continuous ST, and number of sedentary bouts ≤30 min). Secondary endpoints include the changes in step count (steps/d), time spent in LPA (min/wk), time spent in MVPA (min/wk), the Short Physical Performance Battery (SPPB) score, blood pressure, Measuring Change in Physical Activity Questionnaire (MCPAQ) score, waist circumference, BMI, and quality of life (SF-12v2 score) (Table 2).

XSL•FO
RenderX

Figure 2. CONSORT flow diagram.

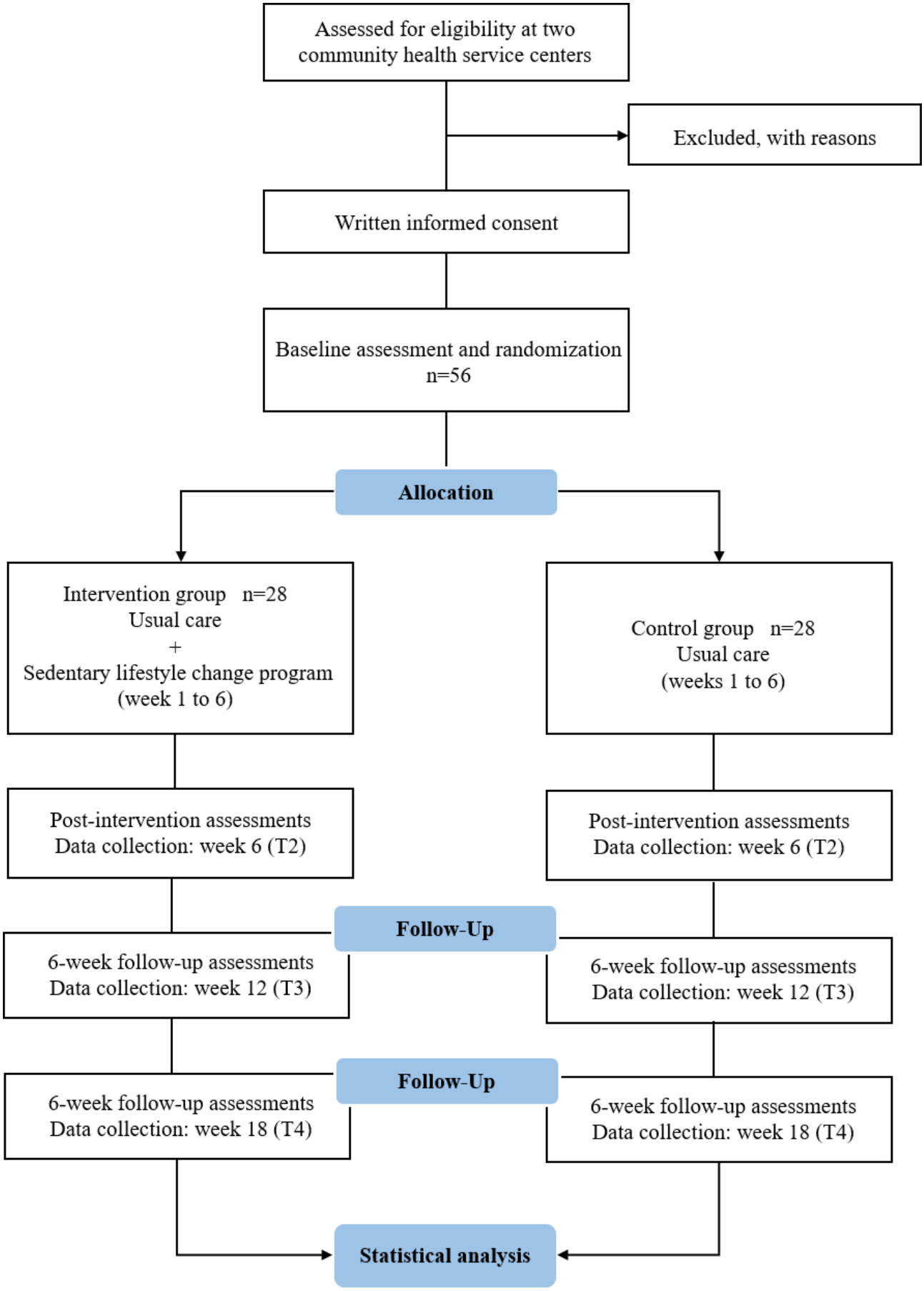

XSL•FO
RenderX

Table . Participant schedule of enrollment, interventions, and assessments.

	Study period				
	Enrollment	Baseline assessment/allocation	Post allocation		
Timepoint	T0	T1	T2 6 wk	T3 12 wk	T4 18 wk
Eligibility screen	×				
Informed consent	×				
Allocation		×			
Interventions					
CON group					
INT group					
Assessments					
Sociodemographics		×			
Primary outcomes					
Sedentary time (total sedentary time and sedentary behavior characteristics)		×	×	×	×
Secondary outcomes					
Step count (Omron HJ-204)		×	×	×	×
Time spent in LPA/MVPA (IPAQ-S-C)		×	×	×	×
Measuring Change in Physical Activity questionnaire (MC-PAQ)		×	×	×	×
Physical function (SPPB)		×	×	×	×
Quality of life (SF-12)		×	×	×	×
BMI		×	×	×	×
Waist circumference		×	×	×	×
Blood pressure		×	×	×	×

[a]SF-12: Short-Form Health Survey-12.

Measurements

A self-developed questionnaire will be used to investigate the sociodemographic and clinical information of the participants, including age, gender, ethnicity, educational level, marital status, illness, residence, risk of falls, frailty status, current employment status, and availability of exercise areas in the vicinity of the residence.

ST and PA

The International Physical Activity Questionnaire Short version Chinese (IPAQ-S-C) was used to assess participants' levels of PA and SB over the past 7 days. It consists of vigorous intensity, moderate intensity, walking (light intensity), and ST, with a total of 7 items. Qu et al [51] conducted a cross-cultural adaptation to form a Chinese version of the IPAQ. The reliability and validity of the Chinese version of the questionnaire were favorable. Total ST was recorded as the average amount of time participants spent in sitting, lying, and semi-lying postures each waking hour. When participants have difficulty recalling total ST, recall will be aided by possible sedentary scenarios in a "typical day," such as calculating the total amount of ST spent daily in transportation, hobbies, watching TV, smartphones and computers, socializing, reading, eating, and any other SBs.

SB characteristics, including screen-based ST, longest continuous ST, and number of sedentary bouts lasting ≤30 minutes, were obtained through a self-developed questionnaire.

XSL•FO
RenderX

Screen-based ST was recorded as ST occurring in situations such as watching television and using mobile phones, computers, and other similar electronic devices.

Step Count

Participants' PA will be monitored by using Omron HJ-204 pedometers, which collect the average number of steps taken over 7 days, in steps per day. Participants will be asked to wear the pedometer over their hip and to record the number of steps in the activities diary before going to bed.

Measuring Change in PA

The MCPAQ was developed by Nahar et al [52] based on an MTM with 29 items containing 2 subscales, initiation and maintenance. The questionnaire is based on a 5-point Likert scale, with values ranging from 0 (never) to 4 (completely sure). The initiation model consists of 3 constructs: participatory dialogue (10 items), behavioral confidence (5 items), and changes in physical environment (3 items). The maintenance model consists of emotional transformation (3 items), practice for change (3 items), and changes in social environment (3 items). The higher the total score for each item, the greater the likelihood that participants will initiate or maintain changes in PA. Yang et al [53] conducted a cross-cultural adaptation to develop the Chinese version of the MCPAQ. The Chinese version of the MCPAQ has shown good reliability and validity in Chinese older adults with hypertension. The total scale Cronbach α=0.911, the subscales ranged from 0.729 to 0.943, and the test-retest reliability was 0.898.

Physical Function

SPPB will measure participants' physical function [54], which combines balance, gait speed, and leg strength into a single score from 0 (worst) to 12 (best). Meaningful clinical change is considered 1 point for the SPPB [55]. A total score of <10 is considered frailty, as well as a high risk of disability and falls. For increased PA, participants were categorized based on SPPB scores at baseline to recommend a personalized Vivifrail multicomponent exercise program at phase III [56].

Quality of Life

Participants' quality of life will be assessed with the Chinese version of the Short-Form Health Survey-12-Version 2 (SF-12v2), one of the common measurement tools for evaluating health-related quality of life from the patient's perspective [57]. The scale has demonstrated high reliability and validity in evaluating the health status of Chinese older adults [58,59]. The scale consists of 12 items, including 8 dimensions, which are physical functioning (2 items), role physical (2 items), bodily pain (1 item), general health (1 item), vitality (1 item), social functioning (1 item), role emotional (2 items), and mental health (2 items). The first 4 dimensions are within the physical health composite scale, and the rest are part of the mental health composite scale. It is recommended to use the total scores of the physical and mental composite scale, with a mean score of 50 (SD 10).

Anthropometric Parameters

Anthropometric parameters included height, weight, waist circumference, BMI, and blood pressure. Each data collection will be repeated twice using the same measuring tool, and the average value recorded will be rounded to one decimal.

Blood Pressure

Avoid strenuous exercise or workout within 1 hour before measurement. The participants will sit still for 5 to 10 minutes; the mercury sphygmomanometer is used to measure the systolic and diastolic blood pressure.

Adherence

The weekly sessions and interventions will be repeated 2 to 4 times in the community health center to avoid scheduling conflicts for participants and to ensure compliance. The percentage of participants who attend sessions and activities will be recorded (total number of activities/actual number of sessions attended). Participants' satisfaction with the intervention will be assessed by rating the study on a scale of 1 to 10.

Adverse Events

Adverse events will be assessed through participant self-reporting. The number of falls, hospitalizations in the last 5 months, and emergency room visits will be recorded for participants during the study.

Sample Size Calculation

The sample size was calculated using G*Power software (version 3.1.9.6) [60]. The sample size estimation refers to the standardized deviation of the self-reported ST among community-dwelling older adults obtained from a similar previous study [42]. A total of 48 older adults (24 per group) will be required to detect the effect size of 0.84, with a power of 80% ($1-\beta$=0.80) and a level of significance of 5% (α=.05, 2-tailed) by unpaired *t*-test. Using an SD of 142.98 minutes per day and an expected reduction of 120 minutes per day self-reported sitting time, a sample size of 24 patients per arm is needed. Taking a dropout of 15% into consideration, we will have at least 56 participants in our study.

Randomization and Blinding

Participants will be randomized in a 1:1 ratio into a behavioral intervention group, which receives a staircase MTM-based SB change intervention and a usual care group, which receives recommendations for reducing SBs. Randomization will be stratified by center, and each center will be grouped by age 65 to 75 and ≥75 years (as older adults aged ≥75 y are more likely to experience mobility impairment due to chronic conditions, resulting in higher ST compared to those aged 65 - 75 y) [19,61]. Using a permuted block randomization, with a block size of 4, allocation sequences will be generated online using a website (www.randomizer.org) by an independent professional who will not be involved in subject recruitment, intervention implementation, or outcome assessment for this study, and investigators will be informed of group assignment by telephone.

Investigators

The research team consisted of nursing specialists, community nurses, rehabilitation therapists, and postgraduate nursing students. Before implementing the intervention, all staff received comprehensive training from the principal investigator and

XSL•FO
RenderX

participated in simulation training to grasp the important elements of the intervention and effectively implement it.

Dropout and Withdrawal

Participants can withdraw at any time for any reason. They can also choose to remove their data from the study. Participants will be regarded as having withdrawn if they request to leave the trial or are lost to follow-up. Participants in the intervention group will be regarded as having withdrawn if they no longer participate in the behavior change intervention that differed from those in the control group. As this research assumes that each stage of the staircase intervention generates health effects, participants will be allowed to skip a stage but will be expected to complete the first stage and be considered to have withdrawn if they do not.

Statistical Analyses

Data analyses follow intention-to-treat principles, using generalized estimating equation models to assess differential changes in the outcome variables between the 2 groups from baseline to 12 weeks (group×time interaction), with adjustment for potential covariates as appropriate (ie, baseline group differences at 2-sided $P<.25$). Effect sizes for continuous outcome variables were estimated using Cohen *d* statistic based on between-group mean differences from baseline T1 to T2, T3, and T4 end points with cutoffs set at 0.2 (small), 0.5 (medium), and 0.8 (large) [62]. Each generalized estimating equation model includes the main effects of group (intervention vs control), time (T2, T3, T4 vs baseline), and 2-way interaction effects (group×time). The group difference in the change from baseline to 6, 12, or 18 weeks between the 2 groups is verified when the 2-way interaction effects (group×time) was statistically significant. All statistical analyses are 2-sided and performed using IBM SPSS 25.0 software. *P* values <.05 is considered statistically significant.

Continuous variables that conform to normal distribution will be expressed as mean (SD) using an independent samples *t* test. Continuous variables that are not normally distributed and ranked variables will use the Mann-Whitney *U* test, expressed as median (M) and quartiles (P_2 -P). Categorical table variables will be statistically described by frequency and percentage (%) using χ^2 test or Fisher exact test.

Adverse Events and Audit

The research team will collect, report, and assess adverse events to determine their severity and whether they are related to this trial. Serious adverse events and reactions will be reported to the Ethics Committee and appropriate measures will be taken (if present). Audits will be conducted by the School of Nursing, Jilin University.

Data Collection, Monitoring, Management, and Storage

The principal investigator will input the collected data into a computerized database. Participants' names will be anonymized and replaced with codes. All other study data will be securely stored on encrypted computer drives in the Research Centre of Jilin University School of Nursing. Only the research team will have access to the data. Data will be securely archived at the end of the study and stored for up to 3 years.

Ethical Considerations

The design and procedures of this controlled trial followed the Declaration of Helsinki. The study obtained approval from the Ethics Committee of the School of Nursing, Jilin University on December 20, 2023 (grant 2023122001). Written informed consent will be obtained from each participant before baseline measurements. After the follow-up data collection was completed and in accordance with their expressed preferences and needs, participants in the control group were provided with the intervention components and material support they had missed. The study's implementation adheres strictly to the principles of beneficence, respect for human dignity, justice, and respect for truth and fact.

Results

The study was initiated in May 2024. As of May 2025, the project had completed participant recruitment (60 participants) and data collection. The final manuscript with findings is expected to be submitted for publication in September 2025. This protocol is published to ensure methodological transparency and to invite peer review focused on study design.

Discussion

Anticipated Findings

We hypothesize that community-dwelling older adults receiving the MTM-based stepped intervention for sedentary lifestyle change will demonstrate significantly greater improvements across multiple outcomes compared to those in the conventional health education group. Specifically, the intervention group is anticipated to exhibit a pronounced reduction in total ST, longest continuous ST, and screen-based ST. Concurrently, we expect increases in daily step counts, frequency of sedentary breaks, and time spent in MVPA. These behavioral changes are further projected to translate into meaningful health benefits, including improved quality of life, as well as modest reductions in blood pressure, BMI, and waist circumference.

Comparison to Previous Work

The evidence indicates that older adults spend approximately 65% to 80% of their waking hours sedentary [63]. In a survey of Chinese older adults, 26.1% participated in regular PA (more than 3 times a week, each time for more than 30 min of MVPA) [64], suggesting that sedentary lifestyles may be prevalent among older adults. The long-term presence of low energy expenditure during waking hours is detrimental to physical health, mental health, and quality of life and exacerbates the adverse outcomes [36]. Therefore, interventions targeting changes in sedentary lifestyles are necessary to improve health outcomes in older adults. Previous research has confirmed the separate benefits of reducing SB and increasing PA. However, the feasibility and efficacy of a comprehensive, stepped intervention that systematically integrates both strategies to modify sedentary lifestyles in older adults have not been adequately examined.

Most current interventions for PI in older adults involve the traditional "elevator approach," that is, interventions aimed

XSL•FO
RenderX

directly at increasing PA, and tend to ignore the potential SB of the target population [41]. Despite increased PA, older adults may still engage in SB [65,66]. As the health effects of SB are independent of PA, too much cumulative SB can still harm overall health and make it difficult to produce the desired health benefits [67,68]. Similarly, modifying SB alone—without emphasizing the quantity of MVPA—the presence of PI still leads to an inability to efficiently achieve the desired health benefits [36]. Our study addresses this critical gap by proposing an integrated, staircase approach-based sedentary lifestyle intervention that simultaneously targets SB reduction and PA promotion, guided by the MTM.

The MTM suggests that for initiation of behavior change, individuals perceive that the advantages of behavior change outweigh the disadvantages, and they have behavioral self-confidence and are supported by the physical environment. To maintain behavior change, it is necessary to translate one's feelings into goals, to constantly strive for change, and to have the support of the social environment [45]. Rather than simply applying behavioral change theory, our approach was informed by the specific characteristics of the older adults (eg, cognition, environment, social norms), leading to a targeted selection of components and design of strategies. According to qualitative studies, older adults are often unaware of the benefits of reducing SB, resulting in low self-efficacy [69,70]. Therefore, there is a need to provide knowledge-based sessions on SB for older adults, emphasizing its advantages and disadvantages and the importance of behavioral change and offering practical tools [43]. For instance, the cube timer was selected as a tool to interrupt SB. Interventions aimed at breaking up ST through standing and LPA throughout the day may be an effective and more feasible option for improving health [42,71]. This initial step within the staircase approach laid the foundation for subsequently introducing higher-intensity PA, thereby demonstrating a systematic implementation that transcends previously fragmented applications. Higher daily step counts are associated with higher LPA or SB ratios and greater time spent on MVPA [72]. This study also added a pedometer-based intervention that included sessions on the benefits of walking and recommendations for self-monitoring daily steps with a pedometer. Pedometer and cube timer serve as low-technological-threshold, cost-efficient tools that may present a pragmatic alternative to many digital interventions by overcoming the technological usage barriers commonly encountered by older adults [43].

Distinct from prior research predominantly focused on individual-level behavior change, this study also specifically focuses on the impact of sociocultural environments on SB in older adults. Consequently, our intervention design incorporates components aimed at mobilizing social support to address this documented, yet frequently overlooked challenge in existing interventions. Existing sociocultural expectations and social norms are supportive of older adults' seated position, considering it "safe" for them, and supporting their inactivity as a "benevolent" act or, in some settings, standing up and moving around when everyone is sitting is considered "embarrassing" behavior [28,69,73]. These viewpoints may hinder older adults' capacity to change their SB in community settings. Therefore, our intervention integrates strategies to engage participants' close social networks and community members, creating a supportive environment that aligns with the changes in the social environment construct of the MTM.

Strengths and Limitations

This research has several strengths. First, our study takes community-dwelling older adults as the research population; applying a sedentary lifestyle modification intervention program based on the MTM model, and through the chosen behavior-change techniques, acting on the participants' behavior, physical and social environments, addresses the problems of difficulty in breaking habitual SB and low adherence to behavioral modification and ultimately achieves the gradual transformation of SB into PA of different intensities and long-term maintenance of behavioral change. Second, this study uses the "staircase approach" as an activity intensity gradient for modifying sedentary lifestyles. It facilitates a transition from SBs to standing and light activities, then to LPA, and further to MVPA, ultimately establishing a sustainable behavior change model that delivers health benefits. The program provides community-dwelling older adults with sedentary lifestyles a personalized behavior change program that is easier to adhere to and better suited to their daily activity patterns.

There are also several limitations. The modest sample size in this study may constrain the generalizability and statistical power of the outcomes. Moreover, the reliance on self-reported measures for SB introduces potential recall bias. Future investigations should utilize objective monitoring devices, such as accelerometers, to enhance data accuracy and adopt multicenter designs with larger cohorts to validate the external validity of the findings. Extending both the intervention and follow-up durations is also recommended to better evaluate the long-term sustainability of the intervention effects.

Future Directions

This study offers valuable practical implications for nursing practice in health management of community-dwelling older adults. For older adults with sedentary lifestyles, the MTM-based stepped intervention demonstrates potential in systematically reducing SB and promoting MVPA through a progressive intensity gradient, thereby facilitating health benefits. It is recommended that community health centers incorporating such interventions into routine health management programs, promoting the use of low-cost prompting tools (eg, cube timer), combined with personalized goal-setting and social support, is essential to enhance participant engagement and adherence. Furthermore, future initiatives may explore the integration of digital tools (eg, mobile app reminders) with the stepped intervention model to improve both the accessibility and scalability of the intervention.

Dissemination

Participants will be informed of the results of the study through a summary report. Study results will be disseminated through journal publications. The anonymized participant dataset will be retained in an appropriate data archive for 3 years after the study's end for sharing. Without participant approval, the results

XSL•FO
RenderX

of this study and any information submitted by participants will not be disclosed to any third party.

Acknowledgments

We sincerely thank everyone who contributed to the development of this study. We also want to express gratitude to the Health Commission of Jilin Province for providing support with this research.

Funding

This project was supported by the Health Science and Technology Capacity Improvement Project of Jilin Province (project ID 2024A031) and the National Key Research and Development Plan "Active Health and Aging Science and Technology Response" Key Special Project (No. 2022YFC3601305). The major funding body is the Health Commission of Jilin Province, and this protocol (Chinese version) has been peer-reviewed as part of the funding process. The funding had no role in the planning of the study design and will not be involved in data collection and analysis, decision to publish or preparation of the manuscript.

Authors' Contributions

All authors were involved in the design or implementation of this study. ST wrote the initial draft of the manuscript and YHL provided overall guidance on the design of the experiments. CYY, HL, and YR helped to correct this script and prepared figures and tables. ZKQ, WXH, DSS, YJY and LF designed the intervention and the study. LYW is the main supervisor of the principal investigator and provides overall guidance to develop the study protocol and write the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The mapping of the sedentary lifestyle change interventions in the multi-theory model.
[DOCX File, 71 KB - resprot_v15i1e81284_app1.docx]

Multimedia Appendix 2
The mapping of the sedentary lifestyle change interventions in the "staircase approach."
[PDF File, 211 KB - resprot_v15i1e81284_app2.pdf]

Checklist 1
SPIRIT 2013 checklist.
[DOC File, 126 KB - resprot_v15i1e81284_app3.doc]

References

1. World population prospects. Population Division - United Nations. URL: https://population.un.org/wpp [accessed 2025-12-13]
2. Wang Y. National bulletin on the development of ageing undertakings. 2024. URL: https://www.gov.cn/lianbo/bumen/202507/content_7033724.htm [accessed 2025-11-12]
3. Bruins MJ, Van Dael P, Eggersdorfer M. The role of nutrients in reducing the risk for noncommunicable diseases during aging. Nutrients 2019 Jan 4;11(1):85. [doi: 10.3390/nu11010085] [Medline: 30621135]
4. McPhee JS, French DP, Jackson D, Nazroo J, Pendleton N, Degens H. Physical activity in older age: perspectives for healthy ageing and frailty. Biogerontology 2016 Jun;17(3):567-580. [doi: 10.1007/s10522-016-9641-0] [Medline: 26936444]
5. Global action plan on physical activity 2018–2030: more active people for a healthier world. : World Health Organization; 2018 URL: https://www.who.int/publications-detail-redirect/9789241514187 [accessed 2025-12-13]
6. Izquierdo M, Merchant RA, Morley JE, et al. International Exercise Recommendations in Older Adults (ICFSR): expert consensus guidelines. J Nutr Health Aging 2021;25(7):824-853. [doi: 10.1007/s12603-021-1665-8] [Medline: 34409961]
7. WHO guidelines on physical activity and sedentary behaviour. : World Health Organization; 2020 URL: https://www.who.int/publications/i/item/9789240015128 [accessed 2025-12-13]
8. Song Y, Wang Z. Study on the relationship of built-in environmental factors to outdoor physical activity and sedentary time of the elderly in urban areas. J Chengdu Sport Univ 2019;45:113-120. [doi: 10.15942/j.jcsu.2019.04.018]
9. Liu Q, Liu F, Li J, et al. Sedentary behavior and risk of incident cardiovascular disease among Chinese adults. Sci Bull (Beijing) 2020 Oct 30;65(20):1760-1766. [doi: 10.1016/j.scib.2020.05.029] [Medline: 36659249]
10. Herdman TH, Kamitsuru S, editors. NANDA International, Inc Nursing Diagnoses: Definitions and Classification, 2020-2023, 12th edition: Thieme; 2021. [doi: 10.1055/b000000888]

XSL•FO
RenderX

11. Harvey JA, Chastin SFM, Skelton DA. Prevalence of sedentary behavior in older adults: a systematic review. Int J Environ Res Public Health 2013 Dec 2;10(12):6645-6661. [doi: 10.3390/ijerph10126645] [Medline: 24317382]
12. Sun F, Norman IJ, While AE. Physical activity in older people: a systematic review. BMC Public Health 2013 May 6;13:449. [doi: 10.1186/1471-2458-13-449] [Medline: 23648225]
13. Owen N, Healy GN, Matthews CE, Dunstan DW. Too much sitting: the population health science of sedentary behavior. Exerc Sport Sci Rev 2010 Jul;38(3):105-113. [doi: 10.1097/JES.0b013e3181e373a2] [Medline: 20577058]
14. Tremblay MS, Aubert S, Barnes JD, et al. Sedentary behavior research network (SBRN) - terminology consensus project process and outcome. Int J Behav Nutr Phys Act 2017 Jun 10;14(1):75. [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
15. Dempsey PC, Matthews CE, Dashti SG, et al. Sedentary behavior and chronic disease: mechanisms and future directions. J Phys Act Health 2020 Jan 1;17(1):52-61. [doi: 10.1123/jpah.2019-0377] [Medline: 31794961]
16. de Rezende LFM, Rey-López JP, Matsudo VKR, do Carmo Luiz O. Sedentary behavior and health outcomes among older adults: a systematic review. BMC Public Health 2014 Apr 9;14(1):333. [doi: 10.1186/1471-2458-14-333] [Medline: 24712381]
17. da Silva VD, Tribess S, Meneguci J, et al. Association between frailty and the combination of physical activity level and sedentary behavior in older adults. BMC Public Health 2019 Jun 7;19(1):709. [doi: 10.1186/s12889-019-7062-0] [Medline: 31174515]
18. Silveira EA, Mendonça CR, Delpino FM, et al. Sedentary behavior, physical inactivity, abdominal obesity and obesity in adults and older adults: a systematic review and meta-analysis. Clin Nutr ESPEN 2022 Aug;50:63-73. [doi: 10.1016/j.clnesp.2022.06.001] [Medline: 35871953]
19. Kim Y, Lee E. The association between elderly people's sedentary behaviors and their health-related quality of life: focusing on comparing the young-old and the old-old. Health Qual Life Outcomes 2019 Jul 26;17(1):131. [doi: 10.1186/s12955-019-1191-0] [Medline: 31349858]
20. Hermelink R, Leitzmann MF, Markozannes G, et al. Sedentary behavior and cancer-an umbrella review and meta-analysis. Eur J Epidemiol 2022 May;37(5):447-460. [doi: 10.1007/s10654-022-00873-6] [Medline: 35612669]
21. Yerramalla MS, van Hees VT, Chen M, Fayosse A, Chastin SFM, Sabia S. Objectively measured total sedentary time and pattern of sedentary accumulation in older adults: associations with incident cardiovascular disease and all-cause mortality. J Gerontol A Biol Sci Med Sci 2022 Apr 1;77(4):842-850. [doi: 10.1093/gerona/glac023] [Medline: 35094083]
22. Chastin S, Gardiner PA, Harvey JA, et al. Interventions for reducing sedentary behaviour in community-dwelling older adults. Cochrane Database Syst Rev 2021 Jun 25;6(6):CD012784. [doi: 10.1002/14651858.CD012784.pub2] [Medline: 34169503]
23. Prince SA, Saunders TJ, Gresty K, Reid RD. A comparison of the effectiveness of physical activity and sedentary behaviour interventions in reducing sedentary time in adults: a systematic review and meta-analysis of controlled trials. Obes Rev 2014 Nov;15(11):905-919. [doi: 10.1111/obr.12215] [Medline: 25112481]
24. Li S, Lear SA, Rangarajan S, et al. Association of sitting time with mortality and cardiovascular events in high-income, middle-income, and low-income countries. JAMA Cardiol 2022 Aug 1;7(8):796-807. [doi: 10.1001/jamacardio.2022.1581] [Medline: 35704349]
25. Dunstan DW, Dogra S, Carter SE, Owen N. Sit less and move more for cardiovascular health: emerging insights and opportunities. Nat Rev Cardiol 2021 Sep;18(9):637-648. [doi: 10.1038/s41569-021-00547-y] [Medline: 34017139]
26. Ekelund U, Steene-Johannessen J, Brown WJ, et al. Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. Lancet 2016 Sep 24;388(10051):1302-1310. [doi: 10.1016/S0140-6736(16)30370-1] [Medline: 27475271]
27. Raffin J, de Souto Barreto P, Le Traon AP, Vellas B, Aubertin-Leheudre M, Rolland Y. Sedentary behavior and the biological hallmarks of aging. Ageing Res Rev 2023 Jan;83:101807. [doi: 10.1016/j.arr.2022.101807] [Medline: 36423885]
28. Matson TE, Renz AD, Takemoto ML, McClure JB, Rosenberg DE. Acceptability of a sitting reduction intervention for older adults with obesity. BMC Public Health 2018 Jun 7;18(1):706. [doi: 10.1186/s12889-018-5616-1] [Medline: 29879948]
29. Dogra S, Dunstan DW, Sugiyama T, Stathi A, Gardiner PA, Owen N. Active aging and public health: evidence, implications, and opportunities. Annu Rev Public Health 2022 Apr 5;43:439-459. [doi: 10.1146/annurev-publhealth-052620-091107] [Medline: 34910580]
30. Glenney SS, Brockemer DP, Ng AC, Smolewski MA, Smolgovskiy VM, Lepley AS. Effect of exercise training on cardiac biomarkers in at-risk populations: a systematic review. J Phys Act Health 2017 Dec 1;14(12):968-989. [doi: 10.1123/jpah.2016-0631] [Medline: 28682654]
31. Aunger JA, Doody P, Greig CA. Interventions targeting sedentary behavior in non-working older adults: a systematic review. Maturitas 2018 Oct;116:89-99. [doi: 10.1016/j.maturitas.2018.08.002] [Medline: 30244786]
32. Hadgraft NT, Winkler E, Climie RE, et al. Effects of sedentary behaviour interventions on biomarkers of cardiometabolic risk in adults: systematic review with meta-analyses. Br J Sports Med 2021 Feb;55(3):144-154. [doi: 10.1136/bjsports-2019-101154] [Medline: 32269058]

XSL•FO
RenderX

33. Duvivier B, Schaper NC, Koster A, et al. Benefits of substituting sitting with standing and walking in free-living conditions for cardiometabolic risk markers, cognition and mood in overweight adults. Front Physiol 2017;8:353. [doi: 10.3389/fphys.2017.00353] [Medline: 28642713]
34. Le Roux E, De Jong NP, Blanc S, Simon C, Bessesen DH, Bergouignan A. Physiology of physical inactivity, sedentary behaviours and non-exercise activity: insights from the space bedrest model. J Physiol 2022 Mar;600(5):1037-1051. [doi: 10.1113/JP281064] [Medline: 33501660]
35. Dumuid D, Pedišić Ž, Stanford TE, et al. The compositional isotemporal substitution model: a method for estimating changes in a health outcome for reallocation of time between sleep, physical activity and sedentary behaviour. Stat Methods Med Res 2019 Mar;28(3):846-857. [doi: 10.1177/0962280217737805] [Medline: 29157152]
36. Zhou W, Yan W, Wang T, et al. Independent and joint association of physical activity and sedentary behavior on all-cause mortality. Chin Med J (Engl) ;134(23):2857-2864. [doi: 10.1097/CM9.0000000000001730]
37. Katzmarzyk PT, Ross R, Blair SN, Després JP. Should we target increased physical activity or less sedentary behavior in the battle against cardiovascular disease risk development? Atherosclerosis 2020 Oct;311:107-115. [doi: 10.1016/j.atherosclerosis.2020.07.010] [Medline: 32773106]
38. Barone Gibbs B, Brach JS, Byard T, et al. Reducing sedentary behavior versus increasing moderate-to-vigorous intensity physical activity in older adults. J Aging Health 2017 Mar;29(2):247-267. [doi: 10.1177/0898264316635564] [Medline: 26944808]
39. Akins JD, Crawford CK, Burton HM, Wolfe AS, Vardarli E, Coyle EF. Inactivity induces resistance to the metabolic benefits following acute exercise. J Appl Physiol (1985) 2019 Apr 1;126(4):1088-1094. [doi: 10.1152/japplphysiol.00968.2018] [Medline: 30763169]
40. Wheeler MJ, Dunstan DW, Ellis KA, et al. Effect of morning exercise with or without breaks in prolonged sitting on blood pressure in older overweight/obese adults. Hypertension 2019 Apr;73(4):859-867. [doi: 10.1161/HYPERTENSIONAHA.118.12373] [Medline: 30782027]
41. Mama SK, Soltero EG, Joseph RP. Reducing sedentary behavior and increasing physical activity among low active, underserved adults: a Staircase approach. J Phys Act Health 2023 Aug 1;20(8):677-679. [doi: 10.1123/jpah.2023-0171] [Medline: 37279898]
42. Crombie KM, Leitzelar BN, Almassi NE, Mahoney JE, Koltyn KF. The feasibility and effectiveness of a community-based intervention to reduce sedentary behavior in older adults. J Appl Gerontol 2022 Jan;41(1):92-102. [doi: 10.1177/0733464820987919] [Medline: 33504249]
43. Compernolle S, De Cocker K, Cardon G, De Bourdeaudhuij I, Van Dyck D. Older adults' perceptions of sedentary behavior: a systematic review and thematic synthesis of qualitative studies. Gerontologist 2020 Nov 23;60(8):572-582. [doi: 10.1093/geront/gnz127] [Medline: 31670766]
44. Maher JP, Conroy DE. A dual-process model of older adults' sedentary behavior. Health Psychol 2016 Mar;35(3):262-272. [doi: 10.1037/hea0000300] [Medline: 26690644]
45. Sharma M, Awan A, Kapukotuwa S. Mini review: possible role of the multi-theory model of health behavior change in designing substance use prevention and treatment interventions. Front Public Health 2024;12:1298614. [doi: 10.3389/fpubh.2024.1298614] [Medline: 38496384]
46. Batra K, Sharma M, Dai CL, Khubchandani J. COVID-19 booster vaccination hesitancy in the United States: a Multi-Theory-Model (MTM)-based national assessment. Vaccines (Basel) 2022 May 11;10(5):758. [doi: 10.3390/vaccines10050758] [Medline: 35632514]
47. Sharma M, Khubchandani J, Nahar VK. Applying a new theory to smoking cessation: case of multi-theory model (MTM) for health behavior change. Health Promot Perspect 2017;7(2):102-105. [doi: 10.15171/hpp.2017.18] [Medline: 28326291]
48. Sharma M, Asare M, Lakhan R, Kanekar A, Nahar VK, Moonie S. Can the Multi-Theory Model (MTM) of health behavior change explain the intent for people to practice meditation? J Evid Based Integr Med 2021;26:2515690X211064582. [doi: 10.1177/2515690X211064582] [Medline: 34898284]
49. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
50. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346:e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
51. Qu NN, Li KJ. Study on the reliability and validity of international physical activity questionnaire (Chinese Vision, IPAQ). Zhonghua Liu Xing Bing Xue Za Zhi 2004 Mar;25(3):265-268. [Medline: 15200945]
52. Nahar VK, Sharma M, Catalano HP, Ickes MJ, Johnson P, Ford MA. Testing multi-theory model (MTM) in predicting initiation and sustenance of physical activity behavior among college students. Health Promot Perspect 2016;6(2):58-65. [doi: 10.15171/hpp.2016.11] [Medline: 27386419]
53. Yang J, Luo Y, Jiang X, Qiu Q, Zhong D, Zhou L. Reliability and validity test of Chinese version of measuring change in physical activity questionnaire in hypertensives. Chin Nurs Manag 2022;22:43-48. [doi: 10.3969/j.issn.1672-1756.2022.01.010]

XSL•FO
RenderX

54. Guralnik JM, Simonsick EM, Ferrucci L, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
55. Guralnik JM, Ferrucci L, Pieper CF, et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci 2000 Apr;55(4):M221-M231. [doi: 10.1093/gerona/55.4.m221] [Medline: 10811152]
56. Exercise for elderly adults. Vivifrail. URL: https://vivifrail.com/tw/%e7%b4%a2%e5%bc%95 [accessed 2024-10-08]
57. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
58. Lam ETP, Lam CLK, Fong DYT, Huang WW. Is the SF-12 version 2 Health Survey a valid and equivalent substitute for the SF-36 version 2 Health Survey for the Chinese? J Eval Clin Pract 2013 Feb;19(1):200-208. [doi: 10.1111/j.1365-2753.2011.01800.x] [Medline: 22128754]
59. Su SW, Wang D. Reliability and Validity of the Chinese Version of SF-12v2 among the Elderly with Community Home-based Care. Chin Gen Pract 2018;21:4262-4265. [doi: 10.12114/j.issn.1007-9572.2018.34.020]
60. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
61. Lally M, Valentine-French S. Lifespan Development a Psychological Perspective, 4th edition: College of Lake County; 2021. URL: https://open.umn.edu/opentextbooks/textbooks/540 [accessed 2025-12-13]
62. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Routledge; 1988. [doi: 10.4324/9780203771587]
63. Harvey JA, Chastin SFM, Skelton DA. How sedentary are older people? A systematic review of the amount of sedentary behavior. J Aging Phys Act 2015 Jul;23(3):471-487. [doi: 10.1123/japa.2014-0164] [Medline: 25387160]
64. The national fitness monitoring center released the 2020 survey bulletin on the status of national fitness activities. : General Administration of Sport of China; 2022 URL: https://www.sport.gov.cn/n315/n329/c24335053/content.html [accessed 2025-12-13]
65. Katzmarzyk PT. Physical activity, sedentary behavior, and health: paradigm paralysis or paradigm shift? Diabetes 2010 Nov;59(11):2717-2725. [doi: 10.2337/db10-0822] [Medline: 20980470]
66. Dogra S, Stathokostas L. Sedentary behavior and physical activity are independent predictors of successful aging in middle-aged and older adults. J Aging Res 2012;2012:190654. [doi: 10.1155/2012/190654] [Medline: 22997579]
67. Biswas A, Oh PI, Faulkner GE, et al. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. Ann Intern Med 2015 Jan 20;162(2):123-132. [doi: 10.7326/M14-1651] [Medline: 25599350]
68. Bankoski A, Harris TB, McClain JJ, et al. Sedentary activity associated with metabolic syndrome independent of physical activity. Diabetes Care 2011 Feb;34(2):497-503. [doi: 10.2337/dc10-0987] [Medline: 21270206]
69. Greenwood-Hickman MA, Renz A, Rosenberg DE. Motivators and barriers to reducing sedentary behavior among overweight and obese older adults. Gerontologist 2016 Aug;56(4):660-668. [doi: 10.1093/geront/gnu163] [Medline: 26035881]
70. Van Dyck D, Mertens L, Cardon G, De Cocker K, De Bourdeaudhuij I. Opinions toward physical activity, sedentary behavior, and interventions to stimulate active living during early retirement: a qualitative study in recently retired adults. J Aging Phys Act 2017 Apr;25(2):277-286. [doi: 10.1123/japa.2015-0295] [Medline: 27759483]
71. Duvivier B, Schaper NC, Hesselink MKC, et al. Breaking sitting with light activities vs structured exercise: a randomised crossover study demonstrating benefits for glycaemic control and insulin sensitivity in type 2 diabetes. Diabetologia 2017 Mar;60(3):490-498. [doi: 10.1007/s00125-016-4161-7] [Medline: 27904925]
72. Amagasa S, Fukushima N, Kikuchi H, et al. Older adults' daily step counts and time in sedentary behavior and different intensities of physical activity. J Epidemiol 2021 May 5;31(5):350-355. [doi: 10.2188/jea.JE20200080] [Medline: 32536640]
73. Chastin SFM, Fitzpatrick N, Andrews M, DiCroce N. Determinants of sedentary behavior, motivation, barriers and strategies to reduce sitting time in older women: a qualitative investigation. Int J Environ Res Public Health 2014 Jan 7;11(1):773-791. [doi: 10.3390/ijerph110100773] [Medline: 24402064]

Abbreviations

BCTs: behavior change techniques
IPAQ: International Physical Activity Questionnaire
LPA: light-intensity physical activity
MCPAQ: Measuring Change in Physical Activity Questionnaire
MTM: multi-theory model
MVPA: moderate-to-vigorous-intensity physical activity
PA: physical activity
PI: physical inactivity
SB: sedentary behavior
SF-12v2: The Short-Form Health Survey-12 version 2

XSL•FO
RenderX

SPPB: Short Physical Performance Battery
ST: sedentary time

Edited by A Mavragani; submitted 25.Jul.2025; peer-reviewed by T Yu; revised version received 16.Nov.2025; accepted 28.Nov.2025; published 06.Jan.2026.

Please cite as:
Tuerdi S, Yang H, He L, Yan R, Cui Y, Wang X, Dong S, Yang J, Zhang K, Li F, Li Y
The Effects of Multi-Theory Model–Based Behavior Change Intervention with Staircase Approach on Sedentary Lifestyle Among Community-Dwelling Older Adults: Study Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e81284
URL: https://www.researchprotocols.org/2026/1/e81284
doi:10.2196/81284

© Subinuer Tuerdi, HuiLing Yang, Li He, Rong Yan, YaoYi Cui, XingHui Wang, ShanShan Dong, JiaYu Yang, KeQiang Zhang, Feng Li, YueWei Li. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Ayurvedic Therapeutic Regimen as an Add-On to Optimized Conventional Management of Parkinson Disease: Protocol for an Exploratory Randomized Controlled Trial Evaluating Clinical, Cortical Excitability, Neuroimmune, and Autonomic Function Parameters

Umesh Chikkanna[1], MD; Pramod Kumar Pal[2], DM; Sophia Jameela[3], MD; Rohan R Mahale[2], DM; Nitish Kamble[2], DM; Venkataram Shivakumar[1], PhD; Bharath Holla[1], PhD; Ravi Yadav[2], DM; Vikram V Holla[2], DM; Talakad Narasappa Sathyaprabha[4], PhD; Kaviraja Udupa[4], PhD; Srinibash Sahoo[5], PhD; Manjunath M Venkataswamy[6], PhD; Selva Ganapathy[7], PhD; Chandra Shekara Rao[3], MD; Babita Yadav[3], MD; Sreenithya M[1], MD; Shivarama Varambally[8], MD, DSC; Sulochana Bhat[5], MD; Narayanam Srikanth[3], MD; Kishore Kumar Ramakrishna[1], PhD

[1]Department of Integrative Medicine, National Institute of Mental Health and Neurosciences, Hosur Main Road, Bengaluru, India

[2]Department of Neurology, National Institute of Mental Health and Neurosciences, Bengaluru, India

[3]Central Council for Research in Ayurvedic Sciences, New Delhi, India

[4]Department of Neurophysiology, National Institute of Mental Health and Neurosciences, Bengaluru, India

[5]Central Ayurveda Research Institute, Bengaluru, India

[6]Department of Neurovirology, National Institute of Mental Health and Neurosciences, Bengaluru, India

[7]Department of Physiotherapy, National Institute of Mental Health and Neurosciences, Bengaluru, India

[8]Department of Psychiatry, National Institute of Mental Health and Neurosciences, Bengaluru, India

Corresponding Author:
Umesh Chikkanna, MD
Department of Integrative Medicine, National Institute of Mental Health and Neurosciences, Hosur Main Road, Bengaluru, India

Abstract

Background: Parkinson disease (PD), a progressive neurodegenerative disorder, lacks disease-modifying treatments. Current therapies focus on symptomatic relief, highlighting the need for adjunctive neuroprotective strategies. Ayurveda, a holistic system, shows promise in improving PD clinical outcomes.

Objective: This assessor-blinded randomized controlled study aims to systematically evaluate the efficacy of an add-on Ayurveda therapeutic regimen compared to conventional treatment as usual (TAU) in improving the clinical outcomes of PD.

Methods: A total of 80 patients with PD, diagnosed according to UK Parkinson's Disease Society Brain Bank criteria, will be randomized into two groups: TAU and add-on Ayurveda. The intervention group will receive Ayurvedic therapy alongside conventional treatment for 180 days, whereas the control group will continue TAU. Assessments will occur at baseline and at 60-, 120-, and 180-day follow-ups, evaluating motor and nonmotor symptoms. Transcranial magnetic stimulation, heart rate variability, and pulmonary function tests will assess cortical excitability, autonomic function, and pulmonary function, respectively. Immunological parameters, including cytokine levels and telomere length, will be analyzed at baseline and at 180 days to explore disease-modifying effects. Liver and renal function tests will monitor safety.

Results: As of 2025, a total of 259 patients with PD have been screened for eligibility, of whom 58 participants have been successfully enrolled in the trial. Among these, 33 participants have completed the intervention and follow-up assessments, 14 have discontinued participation, and 11 are currently continuing participation in the study. Recruitment and follow-up are ongoing, and the trial is scheduled for completion in September 2026.

Conclusions: This study aims to address the lack of mechanistic evidence and robust data on Ayurveda in PD. By systematically evaluating clinical efficacy and potential biomechanisms, the findings will provide preliminary evidence for Ayurvedic interventions, potentially paving the way for their integration into comprehensive PD management.

Trial Registration: Clinical Trial Registry of India CTRI/2022/01/039803; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=NjQ4MTU=&Enc=&userName=

International Registered Report Identifier (IRRID): DERR1-10.2196/83336

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e83336) doi:10.2196/83336

KEYWORDS

Ayurveda; Parkinson disease; transcranial magnetic stimulation; autonomic function test; immunoparameters

Introduction

Background

Parkinson disease (PD), a progressive neurodegenerative disorder characterized by motor symptoms such as tremor and rigidity, affects diverse populations worldwide [1]. First described in 1817, PD's prevalence increases with age, impacting approximately 1% of individuals aged >65 years [2,3]. Regional variations in prevalence exist. Indian studies show a crude prevalence of 14.1 per 100,000 in rural Kashmir, with rates rising to 247 per 100,000 in those aged >60 years [4]. Conversely, Bangalore reported a lower crude prevalence of 27 per 100,000 [3]. Notably, a higher rate of 328.3 per 100,000 was observed among the Parsi community in Mumbai [5]. Early-onset PD, occurring before the age of 40 years, accounts for 3% to 5% of cases, and PD is generally twice as common in men [6]. Beyond motor symptoms, PD presents nonmotor challenges such as sleep disturbances and cognitive impairment. Autonomic dysfunction also increases with age and disease progression. Patients face a 6-fold increased risk of dementia [7].

Standard PD treatment focuses on symptom management using medications such as levodopa and dopamine agonists [8]. However, these treatments do not halt disease progression and can cause side effects [9,10]. Invasive procedures such as deep-brain stimulation are expensive and inaccessible to many, particularly in India. Consequently, many patients with PD explore complementary and alternative medicine (CAM), including Ayurveda [11]. Studies indicate that nearly half of patients with PD use CAM, typically educated, younger, urban individuals with longer disease duration [11].

While CAM, including Ayurvedic interventions, shows potential for improving motor function, robust clinical evidence is lacking. Existing research consists of limited clinical trials, pilot studies, and case reports. A systematic review of published clinical trials evaluating Ayurvedic interventions in PD suggests that therapies such as Mucuna pruriens and various Panchakarma procedures may offer potential benefits in symptom management [12]. However, the evidence is insufficient to reach any reliable conclusion. There is a need for randomized controlled trials (RCTs) with objective measures to evaluate the mechanisms of Ayurvedic treatments. Specifically, the impact of comprehensive Ayurvedic therapies on cortical activity, neuroimmune function, autonomic function, telomere length, and neuropsychological function remains unexplored. Investigating these areas will provide crucial evidence for the efficacy and safety of Ayurveda in PD and shed light on the neurobiological basis of its effects.

Ayurveda and PD

PD can be comprehended within the paradigm of Vāta Vyādhi (neurological disorders) as delineated in Ayurveda. It is primarily attributed to Vāta Dosha Prakopa (vitiation of Vata Dosha, a bodily humor that is primarily associated with movement), which manifests as tremors, rigidity, and bradykinesia. The chronicity and progressive nature of the disease necessitate a multidimensional therapeutic approach involving Vāta-pacifying strategies, Rasāyana (rejuvenating measures), and Panchakarma (detoxification therapies) for symptomatic relief and improving disease progression [13].

Objectives

This assessor-blinded randomized controlled study aims to systematically evaluate the efficacy of an add-on Ayurveda therapeutic regimen compared to conventional treatment as usual (TAU) in improving the clinical outcomes of PD.

The secondary objectives are to explore the effects of add-on Ayurveda therapy in comparison to conventional TAU on multiple physiological and functional parameters, namely cortical excitability, assessed through single (motor threshold [MT], resting cortical, and silent period [SP]) and paired pulse (short intracortical inhibition [SICI], long-interval cortical inhibition [LICI], and intracortical facilitation [ICF]) transcranial magnetic stimulation (TMS) measures; and immune parameters, including Th1/Th2/Th17/T-regulatory cell population, plasma C-reactive protein, plasma levels of Th1/Th2/Th17 pathway cytokines, and telomere length.

Hypothesis

This study hypothesizes that Ayurveda treatment, when integrated with standard TAU, will be safe and effective in improving clinical and cognitive outcomes in patients with idiopathic PD. It is expected that these therapies will enhance the optimal response to levodopa, potentially leading to better symptomatic control. Ayurveda interventions are hypothesized to modulate cortical excitability measures, which will correlate with improvements in both clinical and cognitive parameters. Furthermore, it is postulated that add-on Ayurveda therapies may induce a significant shift in the Th1/Th2/Th17/T regulatory cell population, along with alterations in plasma levels of pro-inflammatory cytokines and telomere length, which will further correlate with clinical, cognitive, and neurophysiological outcomes. Finally, it is anticipated that autonomic dysfunction, as measured through heart rate variability (HRV), will show significant improvement following Ayurveda interventions, aligning with overall clinical benefits observed in patients with PD.

Methods

Study Design Description

This study is designed as an assessor-blinded RCT to be conducted at the National Institute of Mental Health and Neurosciences (NIMHANS), Bengaluru. The study protocol was developed collaboratively by investigators from the NIMHANS and the Central Council for Research in Ayurvedic Sciences (CCRAS) based on existing clinical and preclinical

XSL•FO
RenderX

evidence, and it underwent expert peer review before funding approval.

Trial Design

This study is a prospective, assessor-blinded, exploratory RCT with a parallel-group design and a 1:1 allocation ratio. It is designed as a superiority trial to evaluate the potential benefits of an Ayurvedic therapeutic regimen as an add-on to optimized conventional management in patients with PD, compared to optimized conventional management alone. The trial will be conducted at a single tertiary care center (NIMHANS, Bengaluru, India) over a period of 3 years (Figure 1).

Figure 1. Study design. TAU: treatment as usual.

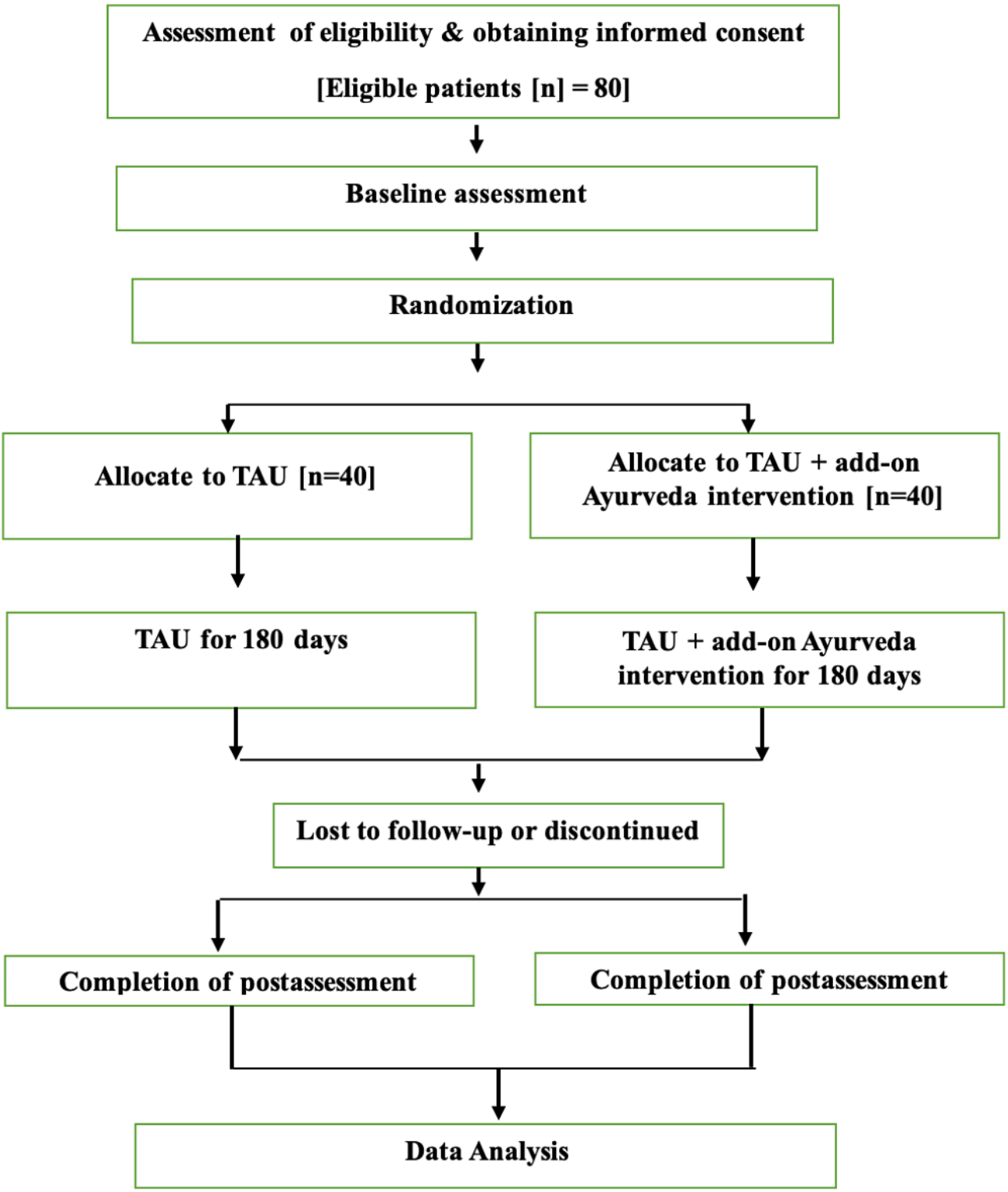

Participants will be followed for 24 weeks post-randomization. The study aims to generate preliminary evidence on clinical efficacy and mechanistic insights related to cortical excitability, neuroimmune function, and autonomic function. Participants will be randomized into either the add-on Ayurveda intervention

XSL•FO
RenderX

arm or the TAU control arm, with each participant undergoing a 6-month study period.

Participants will be recruited from both the outpatient department (OPD) and inpatient department services of the Departments of Neurology and Integrative Medicine at the NIMHANS, Bengaluru, India. NIMHANS is a tertiary care neuropsychiatric and neurosciences hospital and Institute of National Importance under the Ministry of Health and Family Welfare, Government of India. The study will be conducted within NIMHANS using its specialized neurology and integrative medicine infrastructure for recruitment, intervention delivery, and follow-up assessments.

Those meeting the inclusion criteria will be recruited after obtaining written informed consent. Informed consent will be obtained by the senior research fellow (SRF) possessing an MD (Ayurveda) qualification, who has been trained in Good Clinical Practice (GCP) [14] and institutional ethical guidelines for human research. The study procedures, potential benefits, risks, and the right to withdraw at any stage will be clearly explained in the patient's language, and a copy of the consent document will be provided (Multimedia Appendix 1).

To promote adequate enrollment, participants will be recruited through advertisements in print and institutional media, as well as by prescreening of hospital health records to identify potentially eligible individuals. Efforts will also be made to minimize participant burden by providing clear information, convenient scheduling of assessments, and coordinated support from study staff during recruitment and follow-up.

Selection of Participants

Inclusion Criteria

Participants eligible for the study will be those diagnosed with Idiopathic PD based on the UK Parkinson's Disease Society Brain Bank criteria [15]. The study will include individuals in Hoehn and Yahr stages 2.5 to 4, with a disease duration of more than 3 years, experiencing severe motor fluctuations and suboptimal response to standard treatment, including dyskinesia and gait freezing. The eligible age range for participation will be 40 to 70 years. Participants must be on a stable dopaminergic regimen for ≥2 weeks before baseline.

Exclusion Criteria

Participants with Parkinsonism spectrum disorders or other neurodegenerative conditions, such as motor neuron diseases and multiple sclerosis, will be excluded from the study. Individuals with major cardiovascular disorders or those with hepatic, renal, or uncontrolled pulmonary dysfunction will not be eligible. Patients receiving concomitant medications that may interfere with Parkinsonian symptoms, including anticoagulants, neuroleptics, metoclopramide, Compazine, beta-blockers, fluoxetine, clozapine, quetiapine, olanzapine, buspirone, antipsychotics, and central nervous system stimulants such as sodium valproate, will also be excluded. Those with cognitive impairment, defined by a Mini-Mental State Examination score of <18, as well as individuals with evidence of malignancy or a history of substance abuse, will not be considered for participation. Additionally, patients who have participated in any other clinical trial within the past 6 months will be excluded to ensure unbiased study outcomes.

Sample Size

This study has been designed as an exploratory trial primarily aimed at assessing feasibility and generating preliminary clinical evidence on the effects of a multicomponent Ayurvedic add-on therapy in PD. The sample size estimation was based on parameters from a prior exploratory study [16], which reported a mean change in the Unified Parkinson's Disease Rating Scale (UPDRS) score of 1.0 with an SD of 2.0 and a correlation coefficient of 0.60. Using these assumptions, a 2-tailed alpha of 0.05, and 80% power, the required sample size was calculated to be 28 participants per group. Allowing for a 20% attrition rate, a total of 40 participants per group (80 overall) will be recruited.

This sample size is expected to provide preliminary insights into the effects of add-on Ayurveda therapy in PD and to serve as a basis for future large-scale confirmatory trials.

Assignment of Interventions

Participants will be prescreened from the OPD and inpatient department of consultant neurologists at NIMHANS by the designated study team. Following prescreening, the principal investigator (PI) will conduct formal screening procedures and enroll eligible participants. Allocation to study arms will be performed using a computer-generated randomization sequence prepared by an independent statistician from the funding agency (CCRAS). Allocation concealment will be ensured using the sequentially numbered, opaque, sealed envelopes approach to prevent selection bias during participant enrollment. Once the patient is enrolled, the study PI shall provide the sequentially numbered envelope with the group allocation concealed inside to the participant who will open it.

The Ayurvedic interventions will be administered by the PI, who is a qualified Ayurveda physician. Standard conventional care will continue to be managed by the participant's treating neurologist, as per routine clinical practice. However, the assessor conducting the assessments will remain blinded to group allocation to minimize bias.

Intervention

The interventions used in this study consist of two arms: group A (control group) receiving conventional TAU and group B (intervention group) receiving add-on Ayurveda therapy along with TAU of PD, prescribed by the treating neurologist.

Group A (Conventional TAU—Control Group)

Participants will receive pharmacological treatment as prescribed by the treating neurologist. This would include tablet levodopa and carbidopa 125 mg or tablet entacapone 200 mg or tablet pramipexol 1 mg or tablet selegiline 5 mg as appropriate per the assessment of the consultant neurologist.

Group B (Conventional TAU With Add-On Ayurveda—Intervention Group)

Participants will receive Ayurveda intervention as an adjunct to the conventional treatment for a period of 180 days,

XSL•FO
RenderX

comprising 3 cycles of Ayurveda therapies at 60-day intervals. The regimen is structured as follows:

- Preparatory phase (*Deepana-Pachana* therapy): To enhance digestion and metabolism, participants will be administered *Chitrakadi Vati* (250 mg per tablet), 2 tablets with warm water, 3 times daily before food for a duration of 5 days.
- Shodhana therapy (purification phase—*Basti* [rectal administration of medicine] therapy): This includes a Yoga Basti regimen, wherein a combination of *Niruha Basti* (therapeutic decoction enema) and *Anuvasana Basti* (oil enema) will be administered in a prespecified regimen in the form of *Basti* comprising one *Sneha Basti* initially, 3 decoction *Bastis* alternately followed by 3 *Sneha Bastis*, and 1 *Sneha Basti* in the end aimed at pacifying *Vāta Dosha*.
  - Niruha Basti: Decoction of *Erandamoola* (made with the root of *Ricinus Communis* Linn [450 mL]) administered rectally on an empty stomach as per the protocol for *Yoga Basti*.
  - *Anuvasana Basti*: *Bala Taila* (90 mL) administered immediately after food on 5 occasions as per the protocol for *Yoga Basti*.
- Shamana therapy (pacificatory phase): Following purification, the participants will receive oral Ayurveda formulations for symptomatic relief and alleviation of vitiated *Vata Dosha* for 47 days in each cycle as follows:
  - *Mashabaladi Kwatha* (decoction)—15 mL, with 2 pinches of asafoetida and rock salt administered with warm water 3 times daily.
  - *Kalyanaka Ghrita* (medicated ghee)—20 mL, taken early in the morning on an empty stomach.
- The complete Ayurveda intervention regimen will be repeated thrice over the study period, with 3 cycles of 60 days each, totaling 180 days.

Usual concomitant care (if any), including physiotherapy and occupational therapy, will be permitted as part of standard management and will be recorded as concomitant therapy. However, the use of any additional Ayurvedic therapies or herbs not specified in the study protocol will be prohibited during the trial to avoid confounding effects. All concomitant medications, including any dose modifications of standard medications or rescue interventions required for symptom management, will be documented at each study visit in the case report forms.

Strategies for Improving the Fidelity of Care Providers and Adherence of Participants

The Ayurvedic intervention is standardized and will be administered uniformly as per the protocol. Participants will be screened for suitability before Panchakarma therapy, and the intervention will be withheld in cases of intolerance, fever, respiratory infection, or any other clinical contraindication. To ensure uniformity, all procedures will follow standard operating protocols, and participants will receive patient information leaflets detailing the interventions.

As panchakarma therapies are delivered in an inpatient or closely supervised setting by trained professionals, adherence will be monitored daily through direct supervision and treatment logs. Participants are required to maintain a compliance form to record the intake of both standard care and add-on oral Ayurvedic medications, which will be reviewed by the study team during follow-up visits. Furthermore, to promote participant retention and minimize missing data, follow-up will be reinforced through periodic telephonic contact between scheduled visits.

Outcome Measures

The primary outcome is the change in the total score of the Movement Disorder Society-UPDRS (MDS-UPDRS) after 6 months of treatment, compared with the baseline. MDS-UPDRS Part III will be assessed in both the ON and OFF medication states. The OFF state assessment will be conducted after a 12-hour withdrawal from the patient's last dopaminergic medication, whereas the ON state assessment will be performed 1 hour after the administration of the first morning dose of dopaminergic medication. The OFF-state evaluation (after 12 h withdrawal) reflects the underlying disease severity, whereas the ON-state evaluation (1-h postmorning dose) reflects treatment responsiveness. The total MDS-UPDRS change is thus intended to integrate both clinical states, which is relevant for this exploratory phase of the Ayurvedic add-on intervention assessment.

The secondary outcome measures include a range of neurophysiological, biochemical, cognitive, and functional parameters. Cortical excitability will be assessed using TMS, evaluating both single-pulse (MT, resting cortical, and SP) and paired-pulse (SICI, LICI, and ICF) measures. Autonomic function will be analyzed through HRV, whereas pulmonary function tests will assess respiratory capacity.

Exploratory mechanistic outcomes include alterations in immunological markers, including changes in Th1/Th2/Th17/T regulatory cell populations, plasma levels of pro-inflammatory cytokines, and telomere length, which will be measured to evaluate the immunomodulatory and anti-inflammatory effects of the intervention.

Functional and motor outcomes will be examined through reductions in dyskinesia severity using the Unified Dyskinesia Rating Scale and improvements in gait parameters, measured using the 6-minute walk test and gait speed (10-m walk test). Cognitive function will be assessed using the Montreal Cognitive Assessment, and postural stability will be evaluated through the Balance Error Scoring System. Changes in Scales for Outcomes in Parkinson's disease-Sleep will be monitored to assess improvements in sleep disturbances associated with PD.

Safety assessments will include monitoring for treatment-emergent adverse events throughout the study period. Liver function tests, renal function tests, and complete blood count will be performed at baseline and at the end of the study (day 180) to detect any potential systemic adverse effects. For participants undergoing Basti therapy, adverse events (AEs) will be monitored during the in-patient care.

Participants in the TAU arm will continue with their dopaminergic medications. At baseline, stabilized or optimized dopaminergic doses will be recorded, and any subsequent changes during the trial will be documented along with dates and reasons. Levodopa equivalent daily dose (LEDD) will be

XSL•FO
RenderX

calculated at each visit (baseline and follow-ups) using standardized conversion factors and tracked longitudinally.

All clinical and neurophysiological outcome assessments will be performed by a trained SRF with an MD (Ayurveda) qualification, and the routine assessments will be done by the neurologists involved in the study before prescribing the study TAU interventions or deciding changes in interventions. The SRF underwent a 3-month structured training program under the supervision of a neurologist and neurophysiologist to ensure competency in administering both clinical and neurophysiological assessments. The same SRF will also be responsible for data collection throughout the study to maintain consistency.

All outcome measures are evaluated using validated clinical scales and standardized laboratory instruments, as detailed in the manuscript, each with established reliability and validity. Before the commencement of data collection, all laboratory instruments and procedures will be calibrated and standardized. To promote data quality, assessors received formal training, and interrater reliability for the primary outcome measure (MDS-UPDRS Part III) will be established in at least 20% of participants.

Ethical Considerations

Overview

The trial has been approved by the NIMHANS Human Ethics Committee for Research in AYUSH and Integrative Medicine (NIMHANS/HECAIM/5th/MEETING/2021 - 2022, dated January 7, 2022). All procedures in the trial will comply with the ethical guidelines outlined by the Indian Council of Medical Research and the Declaration of Helsinki [14,17]. Participants will be recruited only after obtaining written informed consent, ensuring that they fully understand the study's objectives, interventions, potential risks, and benefits. The study is covered by trial insurance from the Oriental Insurance Company. Confidentiality of the study participants will be ensured by anonymization of the data.

Consent

Informed consent will be obtained by the SRF possessing an MD (Ayurveda) qualification, who has been trained in GCP [14] and institutional ethical guidelines for human research. The SRF has prior experience in conducting clinical studies and in obtaining informed consent in accordance with ethical standards.

The consent process will be conducted in the participant's preferred language, ensuring adequate explanation of study objectives, procedures, potential risks, and benefits. Ample time will be provided for participants to ask questions and decide about participation voluntarily. The entire process will be documented, and a copy of the signed consent form will be provided to each participant.

Any AEs occurring during the course of the trial will be promptly reported to the Data and Safety Monitoring Board (DSMB) as per the approved protocol. Participants experiencing an AE will receive appropriate and immediate clinical care. Each event will be independently reviewed by a committee comprising experts from both the Ayurveda system of medicine and modern medical sciences.

If, upon thorough evaluation, the AE is determined to be related to the investigational intervention, compensation will be provided to the participant in accordance with regulatory requirements through the study's approved insurance coverage.

Brief Description of Some Relevant Assessments

Clinical Assessments

Movement Disorder Society–Unified Parkinson's Disease Rating Scale

This scale assesses motor and nonmotor PD symptoms. It includes four parts: nonmotor experiences of daily living (part I), motor experiences of daily living (part II), motor examination (part III), and motor complications (part IV). Parts IA, IB, and II are conducted in the "ON" state, whereas Part III is assessed in both "ON" and "OFF" states. Part IV integrates patient-reported information with clinical observations [18].

Unified Dyskinesia Rating Scale

This tool evaluates involuntary movements (dyskinesia) associated with PD treatment. It includes historical (on-dyskinesia and off-dystonia) and objective (impairment and disability) sections, focusing on choreic and dystonic movements [19].

Scales for Outcomes in Parkinson's Disease–Sleep

This assesses nighttime and daytime sleep problems over the past month. It includes subscales for nighttime sleep and daytime sleep, with a global assessment of nocturnal sleep quality [20].

Montreal Cognitive Assessment

This tool evaluates cognitive function, assessing memory, language, executive functions, visuospatial skills, attention, abstraction, and orientation. It provides a Derived Memory Index score to predict Alzheimer dementia conversion [21,22].

10-m Walk Test and 6-Minute Walk Test

These assess gait. The 10-m walk test measures gait speed, whereas the 6-minute walk test evaluates aerobic capacity and endurance. Both are performed in the "OFF" state [23,24].

Balance Error Scoring System

This measures postural stability using the Biodex balance system. It involves 3 stances (double-leg, single-leg, tandem) performed on firm and foam surfaces with eyes closed. Errors are scored, with lower scores indicating better balance [25].

Transcranial Magnetic Stimulation

This noninvasive technique assesses motor cortical function. TMS is conducted through "Magventure R30 with magoption" equipment using the MCB-70 coil. It measures parameters such as resting MT (RMT), central motor conduction time, SP, SICI, LICI, and ICF. It is performed in the "OFF" state [26].

Heart Rate Variability

This assesses autonomic function using the "Chronovisor" HRV machine. Time-domain (SDNN, pNN50, RMSSD) and frequency-domain (low frequency, very low frequency, ultra

XSL•FO
RenderX

low, high frequency, and low frequency–to–high frequency ratio) measures are acquired in the "OFF" state.

Pulmonary Function Test

This evaluates respiratory capacity using spirometry. Lung volume measures (forced vital capacity, forced expiratory volume in 1 s, and forced expiratory volume in 1 s–to–forced vital capacity ratio) and respiratory muscle strength (MIP and MEP) are assessed in the "OFF" state.

Neuroimmune Parameters

Objectives

The objectives are to assess (1) the proportions and activation status of human peripheral blood Th1/Th2/Th17 cells and (b) the estimation of cytokine by Cytometric Bead Array.

Peripheral Blood Mononuclear Cell Isolation and Cryopreservation

Peripheral blood mononuclear cells (PBMCs) and plasma samples will be separated from peripheral whole blood samples (10 mL) by density gradient centrifugation and cryopreserved.

Procedure

Staining With Fluorescent Antibodies for Immunophenotyping of Th1/Th2/Th17 Cells

Cryo-preserved PBMCs will be processed for multiparametric assessment of the frequencies of major adaptive immune T cells, namely Th1, Th2, Th17, and Treg cells. Briefly, 1 million PBMCs for each panel of immune surface markers will be stained with LIVE/DEAD Fixable green (Life Technologies) to exclude dead cells and with respective antibodies for 20 and 30 minutes, respectively, at 4°C. Cells will be washed and resuspended in 0.5% paraformaldehyde for fixation. Finally, the cells will be washed and resuspended in phosphate-buffered saline (containing 1% fetal bovine serum). An unstained control will be included with all samples. Stained samples will be acquired using a FACSLyric (BD Biosciences) flow cytometer. Analysis is done by using FlowJo software. Quality control of flow cytometry analysis will be ensured by monitoring the performance of the instrument using Cytometer Setup and Tracking beads. Compensation for fluorescence spectral overlaps will be calculated and applied for the analysis of data for each of the different antibody cocktails. Data will be acquired through the BD FACSuite interface and analyzed on FlowJo software (Table 1).

Table . Design of fluorescent antibody panel based on the lasers and detection filters available in the FACSLyric flow cytometer for the identification of Th1, Th2, and Th17 cells.

Lasers and antibody marker	Fluorochrome
Violet laser (405 nm)	
Anti-CD4	BV 421
Anti-CD25	BV 605
Anti-CD127	BV 750
Blue laser (488 nm)	
Anti-CCR6[CD 196]	Alexa Fluor 488
Anti-CD8	PE
Anti-CCR4[CD 194]	PE CY 7
Anti-FoxP3	BB 700
Red laser (640 nm)	
Anti-CD3	AF 647
Anti-CXCR3[CD 183]	BD Horizon R 718

Estimation of Cytokine by Cytometric Bead Array

Cryopreserved plasma specimens will be processed according to the manufacturer's instructions for the detection and quantification of the proposed cytokines using the Cytometric Bead Array kit (BD Biosciences). Data will be acquired on FACS Verse flow cytometer (BD Biosciences) and analyzed on FCAP Array software (BD Biosciences Table 2).

XSL•FO
RenderX

Table . Participant enrollment and schedule of assessment.

Assessments	Baseline	60th day	120th day	180th day
MDS-UPDRS[a]	✓	✓	✓	✓
UDyRS[b]	✓	✓	✓	✓
SCOPA-S[c]	✓	✓	✓	✓
MoCA[d]	✓	✓	✓	✓
6-minute walk test	✓	✓	✓	✓
Gait speed (10-m walk test)				
BESS[e]	✓	✓	✓	✓
Laboratory parameters				
Hematology				
Complete hemogram	✓	—[f]	—	✓
ESR[g]	✓	—	—	✓
Biochemistry				
Liver function test	✓	✓	—	✓
Renal function test	✓	✓	—	✓
Lipid profile	✓	—	—	✓
HbA_{1c}[h]	✓	—	—	✓
FBS[i]	✓	—	—	✓
Sr. Uric acid	✓	—	—	✓
Inflammatory markers				
Plasma cytokine assay	✓	—	—	✓
Immunophenotyping for Th1/Th2/Th17/T-regulatory cell population	✓	—	—	✓
TMS[j]				
Single pulse measures	✓	—	—	✓
Paired pulse measures	✓	—	—	✓
Neurophysiological parameters	✓	—	—	✓
HRV[k]	✓	—	—	✓
PFT[l]	✓	—	—	✓
Concomitant medication	✓	✓	✓	✓
Rescue medication	—	✓	✓	✓

[a]MDS-UPDRS: Movement Disorder Society–Unified Parkinson's Disease Rating Scale.

[b]UDyRS: Unified Dyskinesia Rating Scale.

[c]SCOPA-S: Scales for Outcomes in Parkinson's Disease–Sleep.

[d]MoCA: Montreal Cognitive Assessment.

[e]BESS: Balance Error Scoring System.

[f]Not available.

[g]ESR: erythrocyte sedimentation rate.

[h]HbA_{1c}: glycated hemoglobin.

[i]FBS: fetal bovine serum.

[j]TMS: transcranial magnetic stimulation.

[k]HRV: heart rate variability.

XSL•FO
RenderX

[l]PFT: pulmonary function test.

Plans for Data Storage, Handling, and Availability

Research investigators and study personnel will receive comprehensive, documented training on the protocol. Access to CRFs and electronic CRFs will be limited to trained individuals to ensure standardized data entry. MDS-UPDRS scores will be recorded on paper and cross-verified with electronic entries to minimize errors. Participant data will be coded, securely stored with password protection, and assigned unique identification codes for pseudonymization. The data analyst will be blinded to group allocation. PIs will ensure secure data sharing mechanisms for privacy and confidentiality.

Upon completion of the study, the deidentified research data will be submitted to and maintained jointly by CCRAS and the NIMHANS. The dataset will be stored in the research data repository managed by CCRAS. Access to the data can be requested by contacting the funding agency through the official CCRAS website [27] and submitting a formal request describing the intended use. Data will be shared upon reasonable request, subject to institutional approvals and applicable data sharing regulations. No public code repository is applicable for this protocol.

Statistical Analysis

All analyses will be performed using SPSS software. The primary outcome, change in MDS-UPDRS Part III score from baseline to 24 weeks, will be analyzed using a mixed model for repeated measures with treatment, time, and treatment×time interaction as fixed effects and baseline values as covariates. LEDD will be included as a time-varying covariate to account for dopaminergic medication effects. Results will be reported as least-square mean differences with 95% CIs. The main analysis will follow a modified intention-to-treat approach including all randomized participants who receive at least one dose of the assigned intervention, whereas a per-protocol analysis will be performed as a sensitivity check on participants who complete ≥80% of scheduled therapies and medication adherence.

Missing data will be handled under the assumption of data missing at random using the mixed model for repeated measures model, which uses all available data without imputation. Additional sensitivity analyses, including multiple imputation and complete case analyses, will be performed to evaluate the robustness of results. The secondary outcomes will be analyzed using appropriate mixed or generalized linear models and interpreted as exploratory. Descriptive statistics will summarize categorical variables as frequencies and percentages, and continuous variables as mean (SD), median (IQR), and range.

Predefined subgroup analyses will explore whether treatment effects differ by baseline Hoehn and Yahr stage (≤3.0 vs >3.0), disease duration (<5 vs ≥5 y), age (<65 vs ≥65 y), baseline LEDD (≤ median vs >median), and presence of dyskinesia. Interaction terms will be tested using mixed effects models, and the results will be interpreted as exploratory. Sensitivity analyses will include assessing participants with stable dopaminergic doses (<10% LEDD change) and testing the interaction between treatment allocation and LEDD change to explore medication-dependent effects. Safety analyses will be descriptive, summarizing AEs and serious AEs by treatment group.

Monitoring and Oversight

DSMB constituted by the CCRAS, comprising experts in neurology, clinical trials, ethics, and Ayurveda, to oversee participant safety and trial integrity. The NIMHANS Ethics Committee will provide continuous ethical oversight. CCRAS will conduct independent monitoring for protocol, GCP, and regulatory compliance, ensuring objectivity. As an exploratory trial, no interim analysis is planned. However, the DSMB will periodically review safety data, including SAEs and potential risks, to ensure participant well-being. The DSMB, in consultation with the Institutional Ethics Committee and CCRAS, will make independent recommendations regarding trial continuation, modification, or termination. This process ensures ongoing participant safety and upholds the trial's scientific and ethical integrity through consistent monitoring and safety oversight.

Harms: AE Collection, Assessment, Reporting, and Management

Participants will be assessed for AEs at each scheduled follow-up visit (days 60, 120, and 180) and during routine hospital visits. Each reported AE will be documented and categorized based on severity, relatedness to the intervention, and expectedness, following standardized guidelines. Laboratory safety assessments, including liver function tests, RFTs, and complete blood count, will be performed at baseline and at the final study visit (day 180) to monitor any potential systemic effects. All recorded AEs will be reviewed by the DSMB, which will assess their clinical significance and provide recommendations for participant safety. SAEs will be promptly reported to the NIMHANS Human Ethics Committee and the sponsor, CCRAS, in accordance with regulatory guidelines.

For participants who discontinue or deviate from the intervention protocol, no further outcome assessments will be conducted, as the primary outcome (MDS-UPDRS Part III) requires evaluation in both ON and OFF medication states, and other assessments (TMS, pulmonary function tests, Montreal Cognitive Assessment, 6-m walk test, and Balance Error Scoring System) are performed in the OFF state. The reasons for nonadherence (eg, adverse effects, lack of perceived benefit, or personal constraints) and nonretention (eg, withdrawal or loss to follow-up) will be systematically documented either during the next scheduled visit or through telephonic communication. These details will be recorded in the case report forms to enable transparent reporting and analysis of participant flow and data completeness.

Protocol Amendments

The PI will be responsible for communicating all significant protocol amendments. Substantial amendments will require approval from the accredited medical ethics committee before implementation, whereas nonsubstantial amendments will be

XSL•FO
RenderX

documented and filed without separate notification. Any modifications impacting participant safety or study procedures will be promptly conveyed to enrolled participants, and, if necessary, a new informed consent process will be initiated.

Protocol Version

The study was sanctioned on April 9, 2021 (3-29/2021-CCRAS/Admn/Coll).

Results

As of August 2025, a total of 259 patients with PD have been screened for eligibility, of whom 58 participants have been successfully enrolled in the trial; among these, 33 participants have completed the intervention and follow-up assessments, 14 have discontinued participation, and 11 are currently continuing in the study. Recruitment and follow-up are ongoing, and the trial is scheduled for completion in September 2026.

Plan to Disseminate Trial Results

The study results will be submitted for publication in peer-reviewed, open-access journals and presented at relevant national and international conferences. A summary of trial outcomes will also be posted on the Clinical Trial Registry of India. Dissemination materials will be jointly reviewed and approved by the PIs, the CCRAS, and the NIMHANS before submission or public release. The anticipated time frame for dissemination is within 12 months following study completion and final data verification.

Discussion

Anticipated Findings

The biomechanisms of conventional medicines are well established. The biosynthesis and metabolism of dopaminergic drugs, dopamine agonists, monoamine oxidase B inhibitors, catechol-O-methyl transferase inhibitors, anticholinergics, and amantadine used in the management of PD are established and well understood [8]. Levodopa-based preparations are designed to replace the dopamine in the depleted striatum. Levodopa is able to cross the blood-brain barrier and converts into the neurotransmitter dopamine by DOPA decarboxylase [28]. Dopamine agonists stimulate the activity of the dopamine system by binding to dopaminergic receptors. Monoamine oxidase B inhibitors work by inhibiting the enzymes involved in dopamine metabolism, which preserves the levels of endogenous dopamine [29]. Catechol-O-methyl transferase inhibitors offer a therapeutic means of preserving endogenous dopamine levels by reducing its breakdown [30]. Anticholinergics act through nondopaminergic mechanisms. These reduce the activity of the neurotransmitter acetylcholine by acting as antagonists at cholinergic receptors and offer some benefits in improving rigidity and tremor [31]. Amantadine acts as a weak glutamate antagonist at the N-methyl-D-aspartate receptor and is useful in the treatment of rigidity and rest tremor, and it can limit the severity of levodopa-induced dyskinesias [32].

Despite Ayurveda's long-standing clinical application in managing neurodegenerative disorders, there is a lack of robust scientific evidence elucidating its mechanisms of action in comparison to conventional medicine. This study aims to bridge this gap by systematically evaluating the efficacy and safety of add-on Ayurveda therapy in PD and exploring its potential biomechanisms. By incorporating clinical, neurophysiological, immunological, and autonomic function assessments, this trial is expected to provide scientific validation for Ayurveda-based interventions in PD management.

The findings from this study could serve as critical evidence supporting the integration of Ayurveda in contemporary neurological care, potentially influencing clinical guidelines and policy decisions related to neurodegenerative disorders. Moreover, this research will provide preliminary data to guide future large-scale, multicentric trials and mechanistic studies exploring Ayurveda's role in neuroprotection and disease modification. Establishing the scientific basis for Ayurveda in PD management may also contribute to its global acceptance, fostering interdisciplinary research and collaboration in integrative neurology.

Ancillary and Posttrial Care

Participants completing the trial will continue to receive routine medical care through the OPD of NIMHANS. Clinical trial insurance has been secured to provide coverage for any unforeseen AEs, ensuring comprehensive risk mitigation and participant safety.

Acknowledgments

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this article.

The study protocol was jointly formulated by principal findings from the National Institute of Mental Health and Neurosciences and the Central Council for Research in Ayurvedic Sciences (CCRAS) and was peer-reviewed by the concerned subject experts before approval of the grant to the investigators. The primary and secondary data collected by the investigators will be analyzed, and the final report will be prepared and published jointly by the National Institute of Mental Health and Neurosciences and CCRAS. The coordinating center has procured and provided the trial drug and has formed the independent Data and Safety Monitoring Board.

A co-investigator from CCRAS is actively involved in the study, and the project was reviewed and vetted by the Internal Screening Committee of CCRAS before initiation. As a co-sponsor and collaborator, CCRAS retains the right to review and provide

XSL•FO
RenderX

comments on the draft manuscript before submission. This review is intended to ensure accuracy, consistency, and alignment with institutional standards; however, CCRAS has no influence over the conduct, data collection, analysis, or interpretation of study results, which remain the responsibility of the study investigators.

No generative artificial intelligence tools were used in the preparation, writing, or editing of this manuscript. All contents were developed by the authors.

Funding

This study is fully funded by the Central Council for Research in Ayurvedic Sciences, an autonomous body under the Ministry of AYUSH, Government of India (grant F.No.3-29/2021-CCRASAdmn/Coll, dated April 9, 2021). The Central Council for Research in Ayurvedic Sciences headquarters are located at 61 - 65 Institutional Area, Opp. D Block, Janakpuri, New Delhi 110058110058. The study has been funded for a 3-year period beginning in April 2021, with an extension approved until September 2026. The total sanctioned amount is INR 8,058,308 (approximately US $96,000).

Authors' Contributions

Conceptualization: UC, PKP, BH, KKR, SJ
Data curation: SV, BH, SJ
Formal analysis: SV, BH, KKR, SJ, KU, MMV, TNS, CSR, SB
Funding acquisition: UC, PKP, BH, KKR, SJ
Investigation: SV, BH, KKR, SJ, UC, MMV, KU, SG, SM
Methodology: UC, SJ, BY, BH, KKR
Project administration: NS
Resources: UC, PKP, NK, RRM, VVH, RY, KU, SG, MMV, SJ, CSR, SB
Supervision: UC, VS, BH, SJ, KKR, KU, MMV, CSR, SB
Validation: SJ, MMV, KU, TNS, KKR, CSR, SB
Writing – original draft: UC, SV, BH, KKR, SJ, KU, MMV, CSR, SM
Writing – review & editing: BH, KKR, UC, SJ, KU, MMV, CSR, SB
Visualization: CSR, SJ, NS, KKR, UC

Conflicts of Interest

All authors are employees of the National Institute of Mental Health and Neurosciences, Bengaluru, and the Central Council for Research in Ayurvedic Sciences under the respective Ministries of Health and Family Welfare and AYUSH, Government of India. The authors declare that they have no personal, financial, or other competing interests that could have influenced the conduct or reporting of this research. None of the authors serves on the editorial board of JMIR Publications or has any other conflicts to disclose.

Multimedia Appendix 1
Patient information sheet and consent forms.
[PDF File, 239 KB - resprot_v15i1e83336_app1.pdf]

Checklist 1
SPIRIT checklist.
[PDF File, 183 KB - resprot_v15i1e83336_app2.pdf]

References

1. Jankovic J. Parkinson’s disease: clinical features and diagnosis. J Neurol Neurosurg Psychiatry 2008 Apr;79(4):368-376. [doi: 10.1136/jnnp.2007.131045] [Medline: 18344392]
2. Pearce JM. Aspects of the history of Parkinson’s disease. J Neurol Neurosurg Psychiatry 1989 Jun;Suppl(Suppl):6-10. [doi: 10.1136/jnnp.52.suppl.6] [Medline: 2666579]
3. Singhal B, Lalkaka J, Sankhla C. Epidemiology and treatment of Parkinson’s disease in India. Parkinsonism Relat Disord 2003 Aug;9(Suppl 2):105-109. [doi: 10.1016/s1353-8020(03)00024-5] [Medline: 12915075]
4. Razdan S, Kaul RL, Motta A, Kaul S, Bhatt RK. Prevalence and pattern of major neurological disorders in rural Kashmir (India) in 1986. Neuroepidemiology 1994;13(3):113-119. [doi: 10.1159/000110368] [Medline: 8015664]
5. Bharucha NE, Bharucha EP, Bharucha AE, Bhise AV, Schoenberg BS. Prevalence of Parkinson’s disease in the Parsi community of Bombay, India. Arch Neurol 1988 Dec;45(12):1321-1323. [doi: 10.1001/archneur.1988.00520360039008] [Medline: 3264148]

XSL•FO
RenderX

6. Cerri S, Mus L, Blandini F. Parkinson's disease in women and men: what's the difference? J Parkinsons Dis 2019;9(3):501-515. [doi: 10.3233/JPD-191683] [Medline: 31282427]
7. Padovani A, Costanzi C, Gilberti N, Borroni B. Parkinson's disease and dementia. Neurol Sci 2006 Mar;27 Suppl 1:S40-S43. [doi: 10.1007/s10072-006-0546-6] [Medline: 16708183]
8. Church FC. Treatment options for motor and non-motor symptoms of Parkinson's disease. Biomolecules 2021 Apr 20;11(4):612. [doi: 10.3390/biom11040612] [Medline: 33924103]
9. Müller T, Russ H. Levodopa, motor fluctuations and dyskinesia in Parkinson's disease. Expert Opin Pharmacother 2006 Sep;7(13):1715-1730. [doi: 10.1517/14656566.7.13.1715] [Medline: 16925499]
10. Nutt JG. Motor fluctuations and dyskinesia in Parkinson's disease. Parkinsonism Relat Disord 2001 Oct;8(2):101-108. [doi: 10.1016/s1353-8020(01)00024-4] [Medline: 11489675]
11. Pandit AK, Vibha D, Srivastava AK, Shukla G, Goyal V, Behari M. Complementary and alternative medicine in Indian Parkinson's disease patients. J Tradit Complement Med 2016 Oct;6(4):377-382. [doi: 10.1016/j.jtcme.2015.03.009] [Medline: 27774422]
12. Chikkanna U, Venkatram S, Holla B, et al. Exploring Ayurveda's potential in Parkinson's disease: a comprehensive narrative. Cureus 2025 Oct;17(10):e93992. [doi: 10.7759/cureus.93992] [Medline: 41210049]
13. Agnivesa. Chikitsa sthana, vatavyadhichikitsa. In: Charaka Samhita: Chaukambha Sanskrit Pratishthan; 2017:616 URL: https://niimh.nic.in/ebooks/ecaraka/?mod=read [accessed 2026-01-06]
14. Subject: draft guideilnes on good clinical practices - regarding. ICMR Bioethics Unit. 2024 Sep 12. URL: https://ethics.ncdirindia.org/asset/pdf/Indian_GCP_guideline.pdf [accessed 2026-01-06]
15. Rizzo G, Copetti M, Arcuti S, Martino D, Fontana A, Logroscino G. Accuracy of clinical diagnosis of Parkinson disease: a systematic review and meta-analysis. Neurology (ECronicon) 2016 Feb 9;86(6):566-576. [doi: 10.1212/WNL.0000000000002350] [Medline: 26764028]
16. Takahashi M, Shimokawa T, Koh J, et al. Efficacy and safety of istradefylline in patients with Parkinson's disease presenting with postural abnormalities: results from a multicenter, prospective, and open-label exploratory study in Japan. J Neurol Sci 2022 Jan 15;432:120078. [doi: 10.1016/j.jns.2021.120078] [Medline: 34923334]
17. World Medical Association. World Medical Association declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
18. Goetz CG, Stebbins GT, Tilley BC. Calibration of unified Parkinson's disease rating scale scores to movement disorder society-unified Parkinson's disease rating scale scores. Mov Disord 2012 Sep 1;27(10):1239-1242. [doi: 10.1002/mds.25122] [Medline: 22886777]
19. Goetz CG, Stebbins GT, Chung KA, et al. Which dyskinesia scale best detects treatment response? Mov Disord 2013 Mar;28(3):341-346. [doi: 10.1002/mds.25321] [Medline: 23390076]
20. Martínez Martín P, Cubo Delgado E, Aguilar Barberà M. Estudio piloto sobre una medida específica para los trastornos del sueño de la enfermedad de Parkinson: SCOPA-Sueño [A pilot study on a specific measure for sleep disorders in Parkinson's disease: SCOPA-Sleep]. Rev Neurol 2006;43(10):577-583. [doi: 10.33588/rn.4310.2006307]
21. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Sep(9):695-699. [doi: 10.1111/jgs.15925] [Medline: 15817019]
22. Kobayashi E, Himuro N, Takahashi M. Clinical utility of the 6-min walk test for patients with moderate Parkinson's disease. Int J Rehabil Res 2017 Mar;40(1):66-70. [doi: 10.1097/MRR.0000000000000205] [Medline: 27977464]
23. Duncan RP, Combs-Miller SA, McNeely ME, et al. Are the average gait speeds during the 10 meter and 6 minute walk tests redundant in Parkinson disease? Gait Posture 2017 Feb;52:178-182. [doi: 10.1016/j.gaitpost.2016.11.033] [Medline: 27915221]
24. Bell DR, Guskiewicz KM, Clark MA, Padua DA. Systematic review of the balance error scoring system. Sports Health 2011 May;3(3):287-295. [doi: 10.1177/1941738111403122] [Medline: 23016020]
25. Kim HM, Nazor C, Zabetian CP, et al. Prediction of cognitive progression in Parkinson's disease using three cognitive screening measures. Clin Park Relat Disord 2019;1:91-97. [doi: 10.1016/j.prdoa.2019.08.006] [Medline: 32368733]
26. Udupa K. Transcranial magnetic stimulation in exploring neurophysiology of cortical circuits and potential clinical implications. IJPP 2021;64:244-257. [doi: 10.25259/IJPP_90_2020]
27. Central council for research in ayurvedic sciences. URL: https://www.ccras.nic.in [accessed 2026-01-06]
28. Beckers M, Bloem BR, Verbeek MM. Mechanisms of peripheral levodopa resistance in Parkinson's disease. NPJ Parkinsons Dis 2022 May 11;8(1):56. [doi: 10.1038/s41531-022-00321-y] [Medline: 35546556]
29. Riederer P, Müller T. Monoamine oxidase-B inhibitors in the treatment of Parkinson's disease: clinical-pharmacological aspects. J Neural Transm (Vienna) 2018 Nov;125(11):1751-1757. [doi: 10.1007/s00702-018-1876-2] [Medline: 29569037]
30. Rivest J, Barclay CL, Suchowersky O. COMT inhibitors in Parkinson's disease. Can J Neurol Sci 1999 Aug;26 Suppl 2:S34-S38. [doi: 10.1017/s031716710000007x] [Medline: 10451758]
31. Gonzalez-Latapi P, Bhowmick SS, Saranza G, Fox SH. Non-dopaminergic treatments for motor control in Parkinson's disease: an update. CNS Drugs 2020 Oct;34(10):1025-1044. [doi: 10.1007/s40263-020-00754-0] [Medline: 32785890]
32. Carrillo-Mora P, Silva-Adaya D, Villaseñor-Aguayo K. Glutamate in Parkinson's disease: role of antiglutamatergic drugs. Basal Ganglia 2013 Dec;3(3):147-157. [doi: 10.1016/j.baga.2013.09.001]

XSL•FO
RenderX

Abbreviations

AE: adverse event
CAM: complementary and alternative medicine
CCRAS: Central Council for Research in Ayurvedic Sciences
DSMB: Data and Safety Monitoring Board
GCP: Good Clinical Practice
HRV: heart rate variability
ICF: intracortical facilitation
LEDD: levodopa equivalent daily dose
LICI: long-interval cortical inhibition
MDS-UPDRS: Movement Disorder Society–Unified Parkinson's Disease Rating Scale
MT: motor threshold
NIMHANS: National Institute of Mental Health and Neurosciences
OPD: outpatient department
PD: Parkinson disease
PI: principal investigator
RCT: randomized controlled trial
SICI: short-interval cortical inhibition
SP: silent period
SRF: senior research fellow
TAU: treatment as usual
TMS: transcranial magnetic stimulation

Edited by J Sarvestan; submitted 01.Sep.2025; peer-reviewed by K Kumar, S Ogundepo; revised version received 27.Oct.2025; accepted 14.Nov.2025; published 09.Jan.2026.

Please cite as:
Chikkanna U, Pal PK, Jameela S, Mahale RR, Kamble N, Shivakumar V, Holla B, Yadav R, Holla VV, Sathyaprabha TN, Udupa K, Sahoo S, Venkataswamy MM, Ganapathy S, Rao CS, Yadav B, M S, Varambally S, Bhat S, Srikanth N, Ramakrishna KK
Ayurvedic Therapeutic Regimen as an Add-On to Optimized Conventional Management of Parkinson Disease: Protocol for an Exploratory Randomized Controlled Trial Evaluating Clinical, Cortical Excitability, Neuroimmune, and Autonomic Function Parameters
JMIR Res Protoc 2026;15:e83336
URL: https://www.researchprotocols.org/2026/1/e83336
doi:10.2196/83336

© Umesh Chikkanna, Pramod Kumar Pal, Sophia Jameela, Rohan R Mahale, Nitish Kamble, Venkataram Shivakumar, Bharath Holla, Ravi Yadav, Vikram V Holla, Talakad Narasappa Sathyaprabha, Kaviraja Udupa, Srinibash Sahoo, Manjunath M Venkataswamy, Selva Ganapathy, Chandra Shekara Rao, Babita Yadav, Sreenithya M, Shivarama Varambally, Sulochana Bhat, Narayanam Srikanth, Kishore Kumar Ramakrishna. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effects of Oliceridine Versus Sufentanil on Postoperative Recovery Quality During Hysteroscopy Under Laryngeal Mask Airway Anesthesia: Protocol for a Single-Blind and Randomized Controlled Trial

YuHang Shou[1,2], BM; JianSheng Luo[1], MM; XianJie Zhang[1], MM; WenHu Zhai[1], MM; Jia Han[1], MM

[1]Department of Anesthesiology, Deyang People's Hospital, Deyang, China

[2]Department of Clinical Medicine, School of Medical and Life Sciences, Chengdu University of Traditional Chinese Medicine, Chengdu, China

Corresponding Author:
Jia Han, MM
Department of Anesthesiology
Deyang People's Hospital
No 173, Section 1, Taishan North Road, Jingyang District
Sichuan
Deyang, 618000
China
Phone: 86 13518278867
Email: will418@126.com

Abstract

Background: Hysteroscopy, the gold standard for diagnosing and treating intrauterine pathologies, has shown substantial increase in its adoption in clinical practice. Nevertheless, early postoperative pain and opioid-related adverse effects remain critical determinants of recovery quality. Oliceridine—a novel G protein-biased μ-opioid receptor agonist—demonstrates an improved therapeutic range over conventional opioids in preliminary studies.

Objective: This study aims to evaluate whether oliceridine enhances recovery quality while reducing opioid-related complications compared to sufentanil in patients undergoing hysteroscopy under general anesthesia.

Methods: This single-center randomized controlled trial will enroll 120 patients undergoing hysteroscopy under general anesthesia with 1:1 randomization to sufentanil- or oliceridine-based analgesia. The primary outcome is early recovery quality assessed by the 15-item Quality of Recovery scale at 24 hours after the surgery, while secondary outcomes include hemodynamic fluctuations during induction, total intraoperative opioid consumption and supplemental bolus frequency, proportion requiring vasoactive agents, incidence of respiratory depression in postanesthesia care unit, postoperative extubation time, opioid-related adverse events within 24 hours, and Visual Analog Scale pain scores at 30 minutes, 4 hours, 8 hours, and 24 hours postextubation.

Results: This study received approval from the Medical Ethics Committee of Deyang People's Hospital, Deyang, China, on April 16, 2025 (approval 2025-03-009-K01). Participant recruitment is anticipated to be completed by December 2025. Data analysis, manuscript preparation, and submission for publication are expected to be completed by February 2026.

Conclusions: The successful completion of this trial will generate evidence regarding whether oliceridine enhances recovery quality while reducing opioid-related complications compared to sufentanil in patients undergoing hysteroscopy under general anesthesia.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2500104024; https://www.chictr.org.cn/showproj.html?proj=275501

International Registered Report Identifier (IRRID): DERR1-10.2196/84521

(JMIR Res Protoc 2026;15:e84521) doi:10.2196/84521

KEYWORDS

oliceridine; hysteroscopy; quality of recovery; opioid-related adverse events; sufentanil; general anesthesia

XSL•FO
RenderX

Introduction

Hysteroscopy is currently recognized as the gold standard for diagnosing and treating intrauterine pathologies, having revolutionized gynecological diagnostics and operative interventions. This modality enables both diagnostic evaluations and surgical treatments through a minimally invasive approach, characterized by rapid postoperative recovery. Advancements in hysteroscopic technologies and instrumentation have substantially increased its adoption across diverse patient populations [1,2]. Pain remains a leading cause of procedural failure during hysteroscopy. To optimize surgical conditions and enhance satisfaction for both patients and surgeons, hysteroscopic procedures in China are predominantly performed under general anesthesia. A nationwide survey revealed that 63.8% of the hysteroscopic surgeries used anesthesia in 2021, underscoring its critical role in contemporary practice [3]. European data corroborate this pattern, with a retrospective analysis confirming that a proportion of hysteroscopic procedures also necessitate general anesthesia for successful completion [4].

To date, no robust evidence exists to define optimal anesthesia protocols for pain management in hysteroscopy, and standardized approaches remain underexplored. For hysteroscopic procedures anticipated to be lengthy or involving significant tissue trauma, the 2020 Chinese Expert Consensus on Anesthesia Management recommends general anesthesia with laryngeal mask airway (LMA) to ensure procedural tolerance [5]. Current clinical practice predominantly uses conventional opioids such as sufentanil for intraoperative analgesia. However, sufentanil administration is associated with multiple opioid-related adverse effects, including nausea, vomiting, excessive sedation, respiratory depression, drug dependence, and opioid-induced hyperalgesia, which collectively compromise early postoperative recovery [6].

Oliceridine is a novel G protein-biased μ-opioid receptor agonist that selectively activates G protein signaling while minimizing β-arrestin recruitment. This pharmacological profile confers a mechanistic advantage: selectively engaging intracellular analgesic pathways while avoiding those mediating adverse effects [7,8]. As such, oliceridine has garnered emerging clinical interest for hysteroscopic analgesia.

The 15-item Quality of Recovery scale (QoR-15) is a multidimensional construct evaluating patient-centered recovery, encompassing physical, physiological, psychological, social, and functional domains [9,10]. The QoR-15—validated for reliability and clinical utility—quantifies recovery across 5 dimensions: physical comfort, emotional status, psychological support, independence, and pain experience. Postoperative pain intensity, nausea/vomiting incidence, and opioid consumption profiles constitute key determinants of patient recovery quality [11,12].

Therefore, this study uses QoR-15 as the primary endpoint to evaluate oliceridine's efficacy within LMA-based general anesthesia for hysteroscopy. We aim to determine whether oliceridine enhances recovery quality while reducing opioid-related complications compared to sufentanil in patients undergoing hysteroscopy under general anesthesia.

Methods

Study Design

This prospective, single-center, single-blind, randomized controlled study will be conducted at Deyang People's Hospital, Deyang, China. The protocol received ethical approval from the institutional ethics committee (2025-03-K01) and was registered at the Chinese Clinical Trial Registry on June 10, 2025. Written informed consent will be obtained from all participants or their legal guardians. Eligible patients will be randomly allocated to either the sufentanil (S) group or oliceridine (O) group. The participant recruitment timeline is shown in Figure 1. A time schedule for the enrollment, interventions, and assessments is detailed in Table 1.

XSL•FO
RenderX

Figure 1. Flowchart for this study.

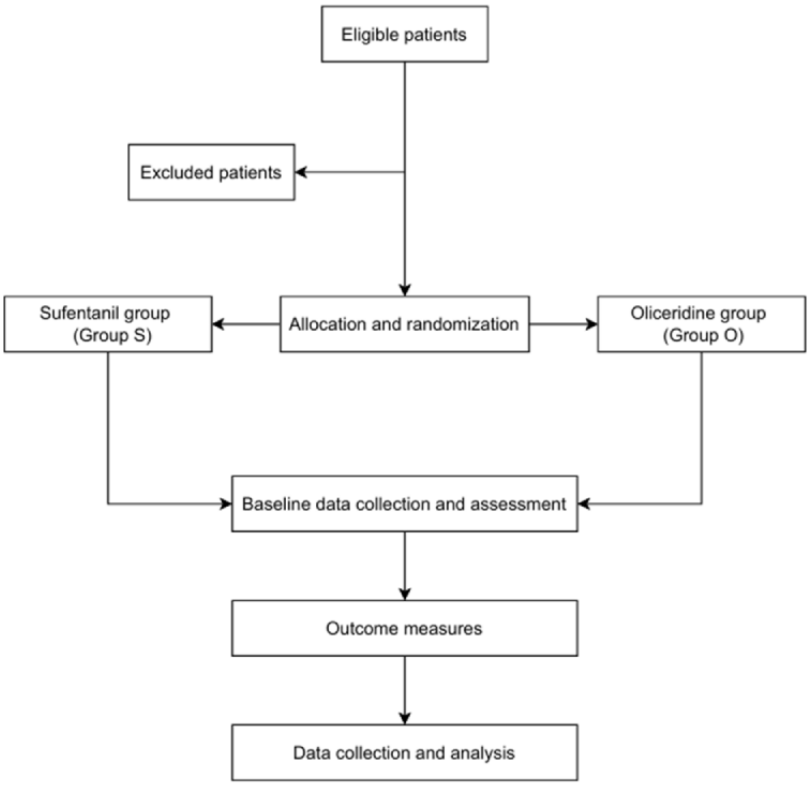

Table 1. The study schedule of the enrollment, interventions, and assessments.

Timepoint	Enrollment phase (day 1)	Allocation (before surgery)	Surgery (during anesthesia)	Follow-up (within 24 h after surgery)
Enrollment				
Eligibility screening	✓			
Informed consent	✓			
Allocation		✓		
Interventions				
Sufentanil			✓	
Oliceridine			✓	
Assessments				
Baseline data	✓			
Primary outcome				✓
Secondary outcome			✓	✓

Eligibility Criteria

Inclusion criteria are as follows: (1) patients scheduled for hysteroscopic surgery under LMA general anesthesia who were assessed by physicians as unable to tolerate the pain stimuli associated with surgical manipulation, (2) patients aged 18 to 65 years, (3) patients with American Society of Anesthesiologists physical status I-III, and (4) patients who voluntarily sign informed consent.

Exclusion criteria are as follows: (1) patients with American Society of Anesthesiologists physical status IV or higher (indicating severe systemic disease, functional incapacity, including but not limited to New York Heart Association class III or higher heart failure, myocardial ischemia, severe reduction in cardiac ejection fraction, third-degree atrioventricular block, stage 3 hypertension, respiratory failure, severe chronic obstructive pulmonary disease, severe pulmonary dysfunction, or decompensated hepatic or renal insufficiency); (2) patients with known allergy to any component of the investigational drugs or contraindications to their use; (3) patients unable to cooperate (eg, psychiatric disorders, anticipated difficulty with neuropsychological assessment, or communication barriers due to pronunciation or dialect); (4) patients with a history of alcohol

XSL•FO
RenderX

or substance abuse; (5) patients with concurrent participation or participation in another clinical trial within 12 weeks prior to this trial; (6) patients with BMI>30 kg/m²; (7) patients unable to comprehend the study procedures or refusal to participate; and (8) patients with any condition deemed by the investigator to pose an unacceptable risk or to contraindicate participation, such as anticipated difficult airway management.

Randomization and Blinding

Block randomization with block sizes of 4 or 6 will be performed using a computer-generated sequence. An independent researcher not involved in the trial will prepare sequentially numbered, opaque, sealed envelopes containing group assignments. Participants will be allocated in a 1:1 ratio to either the experimental or control group by opening these envelopes in a sequential order at the time of enrollment. Allocation cards will be sealed inside opaque envelopes to ensure concealment. The study medications (oliceridine or sufentanil) will be prepared by a dedicated research nurse as visually identical clear solutions in 10-mL syringes. These prepared syringes will then be provided to the anesthesiologist, who will remain unaware of the treatment allocation. Blinding will be maintained for the participants, the anesthesiologist, and the outcome assessors throughout the trial to minimize assessment bias. The blinding will be broken only in the event of severe adverse events. Study participants, their legally authorized representatives, the anesthesiologist, and outcome assessors will be blinded to group allocation throughout the trial. Unblinding the allocation is generally not required per protocol. Emergency unblinding will be performed exclusively to determine rescue medications for critical adverse events or during other medical emergencies

Interventions

Explanation for the Choice of Comparators

This study evaluates the impact of oliceridine (group O) on early postoperative recovery quality in patients undergoing hysteroscopic surgery under general anesthesia. The sufentanil-based regimen (group S) serves as the comparator because of its status as the conventional analgesic approach for such procedures in current clinical practice.

Both groups will receive identical anesthesia induction and maintenance protocols but with divergent analgesic regimens. Group O will be administered oliceridine at a dose of 60 µg/kg (diluted to 10 mL with normal saline) as a slow intravenous bolus for induction. Supplemental doses of 30 µg/kg will be provided if intraoperative analgesia is deemed inadequate, defined as mean arterial pressure and/or heart rate exceeding 20% above baseline values. Group S will receive sufentanil at a dose of 0.3 µg/kg (similarly diluted to 10 mL) as an induction bolus, with supplemental doses of 0.15 µg/kg administered under the same hemodynamic criteria.

Anesthesia Protocol

Upon arrival in the operating room, standard monitoring, including pulse oximetry, invasive radial arterial blood pressure, and electrocardiography will be initiated. Thirty minutes prior to induction, the attending anesthesiologist will administer intravenous penehyclidine hydrochloride (0.01 mg/kg) and ondansetron (4 mg). Anesthesia induction will begin with a slow intravenous bolus of either sufentanil or oliceridine according to group assignment. One minute later, propofol (2 mg/kg) and cisatracurium (0.15 mg/kg) will be administered. After satisfactory neuromuscular blockade is achieved, an LMA will be inserted. Mechanical ventilation will be initiated with a tidal volume of 6-8 mL/kg predicted body weight and positive end-expiratory pressure of 5 cmH_2O. Anesthesia will be maintained using a continuous infusion of propofol (4-12 mg/kg/h), with supplemental opioids administered based on hemodynamic criteria (mean arterial pressure or heart rate >20% above baseline). The propofol infusion rate will be titrated to maintain a bispectral index value between 40 and 60. Neuromuscular blockade will be maintained with intermittent boluses of cisatracurium (one-third of the induction dose). If reversal is required, neostigmine (0.02 mg/kg) and atropine (0.01 mg/kg) will be administered. After surgery, all patients will be transferred to the postanesthesia care unit.

Criteria for Discontinuing or Modifying Allocated Interventions

The criteria for discontinuing are as follows: (1) severe anesthesia/surgical complications (eg, anaphylaxis, shock, cardiac arrest, malignant arrhythmias, aspiration requiring intervention, malignant hyperthermia), (2) protocol deviation in anesthesia/surgical technique, (3) withdrawal of consent, (4) anticipated difficult airway, and (5) failed LMA placement.

Strategies to Improve Adherence to Interventions

The principal investigator will conduct preoperative evaluations against inclusion/exclusion criteria one day prior to surgery. During informed consent, the investigator will detail the study procedures to participants and legally authorized representatives, outlining participant responsibilities.

Relevant Concomitant Care Permitted or Prohibited During the Trial

The provision of additional care during the trial followed institutional protocols.

Outcomes

The primary outcome is defined as the quality of postoperative recovery at 24 hours after surgery and will be assessed using QoR-15 (score range: 0-150; higher scores indicate better recovery). Secondary efficacy endpoints are (1) hemodynamic changes during induction (mean arterial pressure/heart rate at T0 [pre-drug], T1 [pre-LMA], T2 [post-LMA immediate], T3 [5-min post-LMA]); (2) total intraoperative opioid consumption and supplemental bolus frequency; (3) proportion requiring vasoactive agents; (4) incidence of postanesthesia care unit respiratory depression (defined as respiratory rate <6 bpm, end-tidal carbon dioxide amplitude change >50 mm Hg or <30 mm Hg, or waveform loss >20 s); (5) time to extubation; (6) opioid-related adverse events within 24 hours; and (7) Visual Analog Scale pain scores at 30 minutes, 4 hours, 8 hours, and 24 hours postextubation.

XSL•FO
RenderX

Data Collection and Management

Baseline data will include demographic characteristics (age, sex, BMI), the preoperative QoR-15 score, and comorbidities with diagnostic verification. Intraoperative records will include vital signs, anesthetic agents/dosages, surgical duration, and proportion requiring vasopressor administration. In the postanesthesia care unit, time to extubation, early postoperative Visual Analog Scale pain scores, and incidence of respiratory depression will be documented. Postoperative recovery quality will be assessed at 24 hours by using QoR-15, with 24-hour adverse events systematically recorded. Predefined subgroup analyses stratified by age will be conducted for primary and secondary outcomes, applying identical statistical approaches as specified for the main analyses.

Source data will be documented using standardized case report forms. To minimize data entry errors, independent dual-entry verification will be performed. Physical records will be stored securely in locked cabinets with restricted access within the Department of Anesthesiology. Electronic data will be hosted on encrypted, cloud-based platforms.

Data Reporting Guidelines

The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) reporting guidelines will be used for this paper (Multimedia Appendix 1).

Sample Size

An initial retrospective review of 10 patients who underwent hysteroscopy under sufentanil-based general anesthesia resulted in a mean QoR-15 score of 129.8 (SD 4.92). Based on published literature, the minimal clinically important difference for QoR-15 is 6 [13]. We hypothesize that oliceridine will improve the postoperative QoR-15 score by this minimal clinically important difference (6 points). Sample size was calculated using PASS 15 software with the following parameters: a standard deviation of 11.5 [13], a 2-sided α of 0.05, and 90% power. The calculation indicated that 48 patients are required per group. To account for an estimated 20% attrition rate, the total sample size was set at 120 patients (60 per group).

Recruitment Plan

Based on institutional surgical volume projections, patient recruitment is scheduled to commence in July 2025 and is anticipated to be completed within a 5-month period to enroll the target cohort of 120 participants. Eligible patients will be identified primarily through active screening of institutional surgical schedules for planned hysteroscopies. Preoperative screening for eligible candidates will be conducted by the investigators. Written informed consent will be obtained from all participants or their legally authorized representatives either before or on the day of surgery. The consent process, approved by the institutional review board, ensures comprehensive disclosure of the potential benefits, risks, and alternative treatments. Documentation will confirm that all participants have fully understood the information before providing their signature.

Statistical Analysis

The primary analysis will be conducted on a modified intention-to-treat population, defined as all randomized participants who receive the study drug (oliceridine or sufentanil) at anesthesia induction. The per-protocol population will consist of modified intention-to-treat participants who complete the 24-hour follow-up assessment without major protocol deviations, serving as the population for supportive analyses. All participants who receive the study drug will be included in the safety analysis set.

Primary/secondary outcomes will be analyzed using SPSS software (version 27.0; IBM Corp), with continuous variables first assessed for normality via Kolmogorov-Smirnov testing; normally distributed data will be expressed as mean (SD), nonnormal data as median (IQR), and categorical variables as counts (%). Quantitative between-group comparisons will employ independent-samples *t* tests or Mann-Whitney *U* tests, while qualitative comparisons will use chi-square, continuity-corrected chi-square, or Fisher exact tests, with all tests being 2-tailed (statistical significance threshold: $P<.05$).

Methods to Handle Protocol Nonadherence and Any Statistical Methods to Handle Missing Data

Prior to enrollment, comprehensive study details will be disclosed to potential participants, with written informed consent obtained from patients and their legally authorized representatives. Investigators may discontinue participation for substantial protocol deviations (eg, nonattendance at scheduled visits), initiating replacement recruitment according to the randomization sequence.

Plans to Promote Participant Retention and Complete Follow-Up

Investigators will maintain ongoing communication with the participants throughout the trial to address concerns while ensuring protocol adherence. Participant involvement may be discontinued by investigators in cases of nonadherence to scheduled visits or substantial protocol deviations, with all collected data excluded from the final analysis per intention-to-treat principles.

Ethical Considerations

This study received approval from the institutional ethics committee of Deyang People's Hospital (approval 2025-03-009-K01). All procedures strictly adhered to legislative and institutional requirements, with no involvement of vulnerable populations. Eligible patients will undergo preoperative screening. Written informed consent will be obtained from all participants or their legal guardians prior to or on the day of surgery. We ensure comprehensive disclosure of potential benefits and risks, with full participant comprehension verified before consent signing. Participants will be informed that participation is voluntary and that they may withdraw at any time without penalty. Each participant will receive ¥100 (US $14.12) as compensation. All investigators adhere strictly to confidentiality protocols and ethical standards throughout the trial. Personally identifiable data will be maintained with maximized confidentiality.

XSL•FO
RenderX

Confidential information undergoes secure processing/storage protocols, with exclusively anonymized datasets shared for analysis.

Dissemination Plans

Study findings will be submitted for publication in peer-reviewed anesthesiology journals. Access to trial data and protocols is exclusively restricted to the trial leader. No personnel may access participant data without documented prior approval from the principal investigator.

Results

The first patient was enrolled on July 11, 2025. As of September 2025, a total of 55 participants has been enrolled. Data analysis, manuscript preparation, and submission for publication are expected to take place throughout the first quarter of 2026.

Discussion

Anticipated Findings

Hysteroscopy serves as a minimally invasive modality for diagnosing and treating intrauterine pathologies, with indispensable roles in managing endometrial polyps, abnormal uterine bleeding, intrauterine adhesions, uterine septa, endometrial hyperplasia, and early-stage endometrial cancer [14-16]. Procedures requiring prolonged operative time or large instrumentation necessitate anesthesia in operating room settings [17]. In China, LMA-based general anesthesia is currently a prevalent approach for hysteroscopic surgery [18]. Conventional analgesic protocols predominantly use traditional opioids such as sufentanil. However, these agents frequently induce respiratory/circulatory depression, nausea/vomiting, excessive sedation, immunosuppression, opioid-induced hyperalgesia, constipation, and abdominal distension [19-22]. Although effective for anesthesia, these complications prolong hospitalization, increase perioperative management complexity, and compromise postoperative recovery quality. Consequently, reducing opioid-related adverse events is crucial for enhancing recovery.

Oliceridine is a synthetic μ-opioid receptor agonist distinguished from conventional opioids by its G protein-biased selectivity, which mediates potent analgesia while substantially reducing β-arrestin recruitment—a signaling pathway strongly associated with opioid-related adverse events. In contrast, traditional opioids (eg, morphine, fentanyl) nonselectively activate both G protein and β-arrestin pathways upon μ-receptor binding [23,24]. Clinical trial evidence confirms that oliceridine significantly reduces nausea, vomiting, and respiratory complications compared to morphine [25-28]. Therefore, we postulate that this mechanistic advantage of oliceridine may similarly confer clinical benefits when compared with sufentanil.

QoR-15 was selected to assess postoperative recovery quality as a validated patient-reported outcome measure evaluating 5 key recovery dimensions following anesthesia and surgery. Its established reliability and practical clinical utility underpin widespread adoption [11]. The Chinese version has been formally validated, demonstrating robust reliability, validity, and user-friendliness [29]. These properties support its extensive application in clinical trials for postoperative recovery assessment.

Limitations

No prior studies have established whether oliceridine's mechanistic advantages translate to enhanced recovery quality in hysteroscopic surgery under general anesthesia. However, there are a number of potential limitations in the study design: (1) single-center design—though protocol standardization ensures internal validity, generalizability may be limited; (2) exclusive focus on 24-hour recovery without long-term follow-up; and (3) exclusion of high-risk populations (BMI>30 kg/m²), who exhibit heightened susceptibility to opioid-related complications and may theoretically derive the greatest benefit from oliceridine.

Acknowledgments

We are grateful to all the patients and families who participated in this study. We also thank all the anesthesiologists and surgeons who contributed to this study.

Data Availability

Deidentified participant data generated in this study will be made accessible upon reasonable request to the corresponding author following paper publication.

Authors' Contributions

JH and YHS conceptualized and designed the trial and drafted the initial protocol. YHS prepared the preliminary manuscript. JH and XJZ finalized the manuscript. JSL and WHZ were responsible for participant recruitment and allocation. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

SPIRIT checklist.
[DOCX File , 29 KB - resprot_v15i1e84521_app1.docx]

References

1. Vitale SG, Giannini A, Carugno J, van Herendael B, Riemma G, Pacheco LA, et al. Hysteroscopy: where did we start, and where are we now? The compelling story of what many considered the "Cinderella" of gynecological endoscopy. Arch Gynecol Obstet 2024 Oct;310(4):1877-1888. [doi: 10.1007/s00404-024-07677-x] [Medline: 39150502]
2. Buzzaccarini G, Alonso Pacheco L, Vitagliano A, Haimovich S, Chiantera V, Török P, et al. Pain management during office hysteroscopy: an evidence-based approach. Medicina (Kaunas) 2022 Aug 20;58(8):1132 [FREE Full text] [doi: 10.3390/medicina58081132] [Medline: 36013599]
3. Huang L, Yu Q, Zhu Z, Huang P, Ding X, Ma X, et al. The current situation of anaesthesia for hysteroscopy in mainland China: a national survey. J Pers Med 2023 Sep 26;13(10):1436 [FREE Full text] [doi: 10.3390/jpm13101436] [Medline: 37888047]
4. Biela M, Doniec J, Szafarowska M, Sobocinski K, Kwiatkowski A, Kamiński P. Is every patient eligible to have an office hysteroscopy? A retrospective analysis of 1301 procedures. Wideochir Inne Tech Maloinwazyjne 2020 Jun;15(2):337-345 [FREE Full text] [doi: 10.5114/wiitm.2019.89609] [Medline: 32489495]
5. No authors listed. Anesthesia. Anesthesiology 1969;30(1):11. [doi: 10.1097/00000542-196901000-00006]
6. Adams TJ, Aljohani DM, Forget P. Perioperative opioids: a narrative review contextualising new avenues to improve prescribing. Br J Anaesth 2023 Jun;130(6):709-718 [FREE Full text] [doi: 10.1016/j.bja.2023.02.037] [Medline: 37059626]
7. Liu Y, Zhu Y, Fu H. Tolerability of different doses of oliceridine versus traditional opioids in acute pain management: a systematic review and meta-analysis. Sci Rep 2025 Apr 03;15(1):11470 [FREE Full text] [doi: 10.1038/s41598-025-95978-9] [Medline: 40181183]
8. Viscusi ER. A critical review of oliceridine injection as an IV opioid analgesic for the management of severe acute pain. Expert Rev Neurother 2022 Jun;22(6):419-426 [FREE Full text] [doi: 10.1080/14737175.2022.2072731] [Medline: 35502668]
9. Wessels E, Perrie H, Scribante J, Jooma Z. Quality of recovery in the perioperative setting: a narrative review. J Clin Anesth 2022 Jun;78:110685. [doi: 10.1016/j.jclinane.2022.110685] [Medline: 35190342]
10. Markos Z, Amsalu H, Mekuanint A, Yirga S, Simegn AE, Kassegn A. Assessment of intermediate phase post anesthesia quality of recovery and its affecting factors. BMC Anesthesiol 2024 Sep 28;24(1):342 [FREE Full text] [doi: 10.1186/s12871-024-02696-2] [Medline: 39342095]
11. Kleif J, Waage J, Christensen K, Gögenur I. Systematic review of the QoR-15 score, a patient- reported outcome measure measuring quality of recovery after surgery and anaesthesia. Br J Anaesth 2018 Jan;120(1):28-36 [FREE Full text] [doi: 10.1016/j.bja.2017.11.013] [Medline: 29397134]
12. Stark P, Myles P. Burke JA: Development and psychometric evaluation of a postoperative quality of recovery score: the QoR-15. Anesthesiology 2013 2013 Jun 23;118(6):1332-1340. [doi: 10.1097/aln.0b013e318289b84b] [Medline: 23411725]
13. Myles P. Myles DB: an updated minimal clinically important difference for the QoR-15 scale. Anesthesiology 2021 Nov 12;135(5):1. [doi: 10.1097/aln.0000000000003977] [Medline: 34543410]
14. No authors listed. The use of hysteroscopy for the diagnosis and treatment of intrauterine pathology: ACOG committee opinion summary, number 800. Obstet Gynecol 2020 Mar;135(3):754-756. [doi: 10.1097/AOG.0000000000003713] [Medline: 32080048]
15. No authors listed. ACOG technology assessment no. 13 summary: hysteroscopy. Obstet Gynecol 2018 May;131(5):1. [doi: 10.1097/AOG.0000000000002629] [Medline: 29683907]
16. Yen C, Chou H, Wu H, Lee C, Chang T. Effectiveness and appropriateness in the application of office hysteroscopy. J Formos Med Assoc 2019 Nov;118(11):1480-1487 [FREE Full text] [doi: 10.1016/j.jfma.2018.12.012] [Medline: 30611636]
17. De Silva PM, Smith PP, Cooper NAM, Clark TJ, Royal College of ObstetriciansGynaecologists. Outpatient hysteroscopy: (green-top guideline no. 59). BJOG 2024 Dec;131(13):e86-e110. [doi: 10.1111/1471-0528.17907] [Medline: 39160077]
18. ji M, Zhang YZ, Wu Y. Deng X-F: Clinical trial of cis-atracurium in the treatment of patients with hysteroscopic surgery. The Chinese Journal of Clinical Pharmacology 2024;40(18):2640-2644. [doi: 10.13699/j.cnki.1001-6821.2024.18.007]
19. Mauermann E, Clamer D, Ruppen W, Bandschapp O. Association between intra-operative fentanyl dosing and postoperative nausea/vomiting and pain: a prospective cohort study. Eur J Anaesthesiol 2019 Nov;36(11):871-880. [doi: 10.1097/EJA.0000000000001081] [Medline: 31567573]
20. Hoshijima H, Takeuchi R, Kuratani N, Nishizawa S, Denawa Y, Shiga T, et al. Incidence of postoperative shivering comparing remifentanil with other opioids: a meta-analysis. J Clin Anesth 2016 Aug;32:300-312. [doi: 10.1016/j.jclinane.2015.08.017] [Medline: 26432635]
21. de Boer HD, Detriche O, Forget P. Opioid-related side effects: postoperative ileus, urinary retention, nausea and vomiting, and shivering. A review of the literature. Best Pract Res Clin Anaesthesiol 2017 Dec;31(4):499-504. [doi: 10.1016/j.bpa.2017.07.002] [Medline: 29739538]
22. Forget P. Opioid-free anaesthesia. Why and how? A contextual analysis. Anaesth Crit Care Pain Med 2019 Apr;38(2):169-172. [doi: 10.1016/j.accpm.2018.05.002] [Medline: 29775728]

XSL•FO
RenderX

23. Bergese SD, Brzezinski M, Hammer GB, Beard TL, Pan PH, Mace SE, et al. ATHENA: a phase 3, open-label study of the safety and effectiveness of oliceridine (TRV130), a G-protein selective agonist at the μ-opioid receptor, in patients with moderate to severe acute pain requiring parenteral opioid therapy. JPR 2019 Nov;12:3113-3126. [doi: 10.2147/jpr.s217563]
24. Gan T, Wase L. Oliceridine, a G protein-selective ligand at the μ-opioid receptor, for the management of moderate to severe acute pain. Drugs Today (Barc) 2020 Apr;56(4):269-286. [doi: 10.1358/dot.2020.56.4.3107707] [Medline: 32309822]
25. Soergel D, Subach R, Burnham N, Lark M, James I, Sadler B, et al. Biased agonism of the μ-opioid receptor by TRV130 increases analgesia and reduces on-target adverse effects versus morphine: a randomized, double-blind, placebo-controlled, crossover study in healthy volunteers. Pain 2014 Sep;155(9):1829-1835. [doi: 10.1016/j.pain.2014.06.011] [Medline: 24954166]
26. Viscusi E, Webster L, Kuss M, Daniels S, Bolognese J, Zuckerman S, et al. A randomized, phase 2 study investigating TRV130, a biased ligand of the μ-opioid receptor, for the intravenous treatment of acute pain. Pain 2016 Jan;157(1):264-272. [doi: 10.1097/j.pain.0000000000000363] [Medline: 26683109]
27. Singla N, Minkowitz H, Soergel D, Burt D, Subach RA, Salamea M, et al. JPR 2017 Oct;10:2413-2424. [doi: 10.2147/jpr.s137952]
28. Singla NK, Skobieranda F, Soergel DG, Salamea M, Burt DA, Demitrack MA, et al. APOLLO‐2: A Randomized, Placebo and Active‐Controlled Phase Study Investigating Oliceridine (130), a G Protein–Biased Ligand at the μ‐Opioid Receptor, for Management of Moderate to Severe Acute Pain Following Abdominoplasty. Pain Practice 2019 Jun 24;19(7):715-731. [doi: 10.1111/papr.12801]
29. Bu X, Zhang J, Zuo Y. Validation of the Chinese version of the Quality of Recovery-15 score and its comparison with the Post-Operative Quality Recovery Scale. Patient 2016 Jun;9(3):251-259. [doi: 10.1007/s40271-015-0148-6] [Medline: 26518201]

Abbreviations

LMA: laryngeal mask airway
QoR-15: 15-item Quality of Recovery scale
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by J Sarvestan; submitted 21.Sep.2025; peer-reviewed by Y Gong, W Wu; comments to author 03.Dec.2025; revised version received 10.Dec.2025; accepted 11.Dec.2025; published 02.Jan.2026.

Please cite as:
Shou Y, Luo J, Zhang X, Zhai W, Han J
Effects of Oliceridine Versus Sufentanil on Postoperative Recovery Quality During Hysteroscopy Under Laryngeal Mask Airway Anesthesia: Protocol for a Single-Blind and Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84521
URL: https://www.researchprotocols.org/2026/1/e84521
doi:10.2196/84521
PMID:

©YuHang Shou, JianSheng Luo, XianJie Zhang, WenHu Zhai, Jia Han. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 02.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Cluster Randomized Controlled Trial to Promote Physical Activity Among Low-Resourced Mothers in New York City: Protocol for the Free Time for Wellness Effectiveness Trial

Jessica Watterson[1], PhD; Kate Magsamen-Conrad[2], PhD; Kajal Gokal[3], PhD; Ingrid Oakley-Girvan[4], PhD; Jennifer S Hirsch[5], PhD; Min Qian[5], PhD; Richard Buchsbaum[5], MS; Charlene Niles[6]; Kendra Van Horn[7], MPH; Stephanie Moshier[8], BA; Andreina Martinez[5]; Avery Garcia Flores[5]; Sherece Laine[5], MPH; Marilisa Vega[5], BA; Lauren C Houghton[5,9], PhD

[1]Institute for Health Transformation, School of Health and Social Development, Deakin University, Geelong, Australia

[2]Department of Communication Studies, University of Iowa, Iowa City, IA, United States

[3]School of Sport, Exercise and Health, Loughborough University, Loughborough, United Kingdom

[4]Public Health Institute, Oakland, CA, United States

[5]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, United States

[6]Jovie, New York, NY, United States

[7]New York City Department of Parks and Recreation, New York, NY, United States

[8]West Side Campaign Against Hunger, New York, NY, United States

[9]Herbert Irving Comprehensive Cancer Center, Columbia University, New York, NY, United States

Corresponding Author:
Lauren C Houghton, PhD
Department of Epidemiology
Mailman School of Public Health
Columbia University
722 West 168th Street
New York, NY, 10032
United States
Phone: 1 212 342 0246
Email: lh2746@cumc.columbia.edu

Abstract

Background: Physical inactivity is pervasive and prevalent in the United States, particularly among women of low socioeconomic position and women with children. Structural and social barriers make active leisure time a rare commodity, creating a pressing health issue because physical inactivity increases the risk of chronic diseases and poor health.

Objective: The broad objective of this study is to test the effectiveness of Free Time for Wellness, a multilevel intervention to increase physical activity among low-resourced mothers.

Methods: This study comprises a 3-arm parallel cluster randomized controlled trial (RCT) with low-resourced mothers living in New York City. We will randomize fitness class sites (clusters) into arm A (contact control), receipt of free weekly fitness classes; arm B, receipt of free childcare combined with free weekly fitness classes; and arm C, receipt of free childcare combined with free weekly fitness classes and peer support activities. Over 2 years, we will recruit a pilot wave followed by 7 additional waves, totaling 720 participants into 24 fitness classes. Physical activity is the primary outcome, measured using accelerometers, but secondary outcomes also include physical activity assessed using a self-reported questionnaire and attendance data. We will assess additional secondary outcomes (eg, health status, depression, and anxiety) and mediators or moderators (eg, social support and cohesion) with a baseline and follow-up questionnaire. The intention-to-treat analysis will use linear mixed-effects models to assess the main intervention effects on physical activity outcomes and other secondary outcomes. Ethnographic methods will examine how intersecting forms of social identity shape women's experiences of physical activity and understand how real-world conditions shape the intervention implementation.

Results: The study received funding from the US National Institute of Health, covering the period of time from April 1, 2023, through March 31, 2028. We received initial institutional review board approval in August 2023. The study is active and recruiting

XSL•FO
RenderX

participants. As of the day of manuscript submission, we have enrolled 471 participants. Data collection is anticipated to occur until September 2026 for primary completion. The estimated study completion date is December 2026. Dissemination of the results will take place with participants, community members, partners, and researchers through methods such as reports, websites, events, and academic publications and conferences.

Conclusions: This cluster RCT tests whether access to childcare (an understudied structural barrier) and social support can increase physical activity. The study design and outcomes integrate ethnographic methods with a cluster RCT to better understand mechanisms and the impact of intersecting factors such as race or ethnicity, culture, gender, and socioeconomic position. The study leverages widely accessible, existing resources to promote physical activity and foster social support with the ultimate goal of assessing the effect of childcare access on parental health.

Trial Registration: ClinicalTrials.gov NCT06654843; https://clinicaltrials.gov/study/NCT06654843

International Registered Report Identifier (IRRID): DERR1-10.2196/71381

(JMIR Res Protoc 2026;15:e71381) doi:10.2196/71381

KEYWORDS

public health; physical activity; low-resourced; mothers; chronic disease epidemiology; low income; chronic disease; childcare

Introduction

More than 46% of US men and women do not achieve the recommended physical activity guidelines for aerobic activity each week, and 76% do not meet the guidelines for both aerobic and muscle-strengthening activity [1]. This is important because physical inactivity contributes to an increased risk of chronic diseases and poor health. Physical inactivity is particularly prevalent in women of low socioeconomic position (SEP; 60% are inactive), suggesting that there are structural barriers to being physically active. Proposed mechanisms through which low SEP may influence physical inactivity include increased biological stress, increased risk of poor health due to reduced access to health care, and reduced access to other resources such as physical activity facilities and gymnasium memberships [2].

Our previous research identified that an additional barrier for women who are mothers is having the free time to be physically active [3]. This is also supported by other research showing that women with children tend to be less physically active than women of the same age without children, due to structural and social barriers that reduce active leisure time [4,5]. For resource-limited mothers in particular, social and environmental barriers to physical activity include family expectations, lack of transportation, insufficient financial resources, unsafe environments, and a lack of childcare [4]. Childcare and household duties cause significant stress and exhaustion for mothers, with mothers describing parenting as “relentless” (yet also acknowledging its rewards) [6]. There is a large existing body of literature examining wellness interventions for mothers. However, most studies focus only on promoting physical activity during pregnancy or the postpartum period (and not later in motherhood), when being physically active without childcare is easier because young babies can be included in the mother’s activities, such as going for stroller walks. This focus is illustrated by the large number of systematic reviews focused on interventions in these periods [6]. In addition, while a lack of childcare is a well-known barrier to mothers’ physical activity, backed by research, very few studies explicitly address this barrier as part of an intervention [4,7] For example, in a systematic review of group-based physical activity interventions for postpartum women, none of the studies provided childcare, although many studies have identified it as a barrier to being physically active [8].

This study will add to a small but important body of literature on how mothers’ health can be improved by addressing the unique challenges that mothers face in finding time and space for their health and well-being. Some limited evidence suggests that physical activity opportunities for mothers, when children are being cared for by others, can lead to reduced stress and improved BMI, lower resting heart rate, and normal blood pressure in comparison to controls [7,9]. An older systematic review of physical activity and parenthood found that “social support and childcare were paramount in alleviating physical activity barriers, especially for mothers” [8].

To address the need for mothers to find time and space for physical activity, we will test a first-of-its-kind intervention that we co-designed [3] and piloted with mothers [10], which provides free childcare in combination with a novel peer support system. Based on prior theory and our pilot work [11,12], we hypothesize that free childcare and peer support will increase physical activity, and no previous health interventions have studied this. Expanding research on social cohesion suggests that individual adoption of health behavior is much more likely when participants receive social reinforcement from multiple neighbors in their social network [13,14]. For example, a systematic review of randomized controlled trials (RCTs) of group-based physical activity for mothers of young children found evidence of increases to self-reported physical activity and psychological well-being [15]. Further, while resource-limited mothers reported that peer support encourages their participation in physical activity [5], another study found that higher-income mothers were significantly more likely to report friends or family offering to be physically active with them [6]. Our co-designed intervention, the Free Time for Wellness (FT4W) program, will generate novel access to individual- and community-level resources and will leverage peer support among mothers in urban, racially segregated, low-income communities in New York City.

The broad objective of this study is to test the effectiveness of FT4W, an innovative multilevel physical activity intervention to increase physical activity and ultimately reduce the risk of

XSL•FO
RenderX

chronic diseases and improve mothers' well-being. We will do so via the following specific aims:

- Aim 1: conduct a parallel cluster RCT to evaluate the effectiveness of providing free childcare at free fitness classes to increase physical activity among low-resourced, historically minoritized mothers.
- Aim 2: test if peer support enhances the effect of the childcare intervention on physical activity by including another intervention arm promoting social cohesion.
- Aim 3: integrate ethnographic methods throughout the RCT to understand how individual-, community-, and institutional-level factors modify and influence the effect of the intervention, with particular attention to intersectional inequalities.

The saying "it takes a village" captures an existing cultural phenomenon of mothers supporting mothers through in-person and virtual "mom groups" [12]. FT4W builds from this concept, combining cultural context, the science of social cohesion, co-design methodology, and mixed methods to bolster mothers' support networks and help promote their wellness. If effective, the FT4W program will illustrate how supporting mothers at the community level can increase physical activity, thus reducing chronic disease disparities in communities with high social inequality by providing resources, building social cohesion, and freeing up time for mothers to focus on their own wellness.

Methods

Study Design

This study comprises a 3-arm parallel cluster RCT with low-resourced mothers living in New York City. The trial has been planned, and this protocol is reported in line with the CONSORT (Consolidated Standards of Reporting Trials) statement [16].

The complete FT4W program includes free childcare at weekly fitness classes and peer support activities. We will randomize fitness class sites (clusters) into arm A, a contact control, receiving only the free weekly fitness classes; arm B, receiving free childcare provided at the same location as free weekly fitness classes; or arm C, receiving free childcare at free weekly fitness classes, plus peer support activities.

Theoretical Framework and Logic Model

The FT4W program is founded on social cognitive theory, which posits that human behavior is the product of the dynamic interplay of personal, behavioral, and environmental influences [17]. Social cognitive theory emphasizes reciprocal determinism in interactions between people and their environments. The majority of physical activity interventions use social cognitive theory as indicated by the theory and technique tool developed by the Human Behaviour Change Project [18]. We used the theory and technique tool to align the FT4W program components and mediating pathways with behavior change techniques and mechanisms of action substantiated in behavior change literature and expert consensus. Studies that use behavior change techniques of restructuring the physical environment and peer support show that the mechanisms of action are environmental context and resources (aspects of a person's situation or environment that discourage or encourage the behavior) and social influences (interpersonal processes that can cause oneself to change one's thoughts, feelings, or behaviors) [18]. For our study, we will measure social influence via social cohesion.

Our logic model (Figure 1) illustrates the 2 complementary components of the FT4W program and mediating pathways in which the intervention can be effective. We hypothesize that providing free childcare during free fitness classes can provide environmental resources that participants can use to increase their physical activity. We hypothesize that this effect will be stronger in participants who also receive peer support, the intervention component that creates social cohesion. We also hypothesize that the effect of the intervention will differ based on race and ethnicity and other social stratifiers (migration status, SEP, and gender or sexual orientation).

Figure 1. Logic model for Free Time for Wellness program.

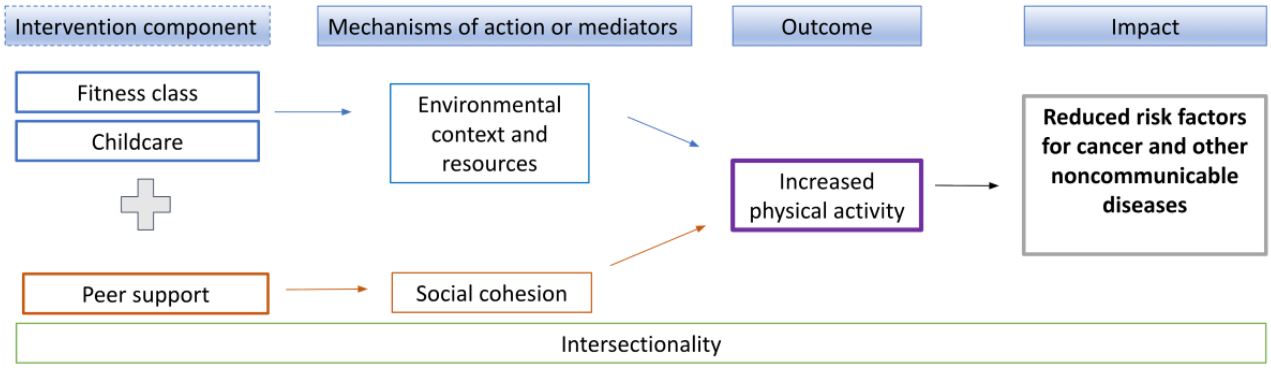

Patient and Public Involvement

The intervention was developed in this New York City–based study through a co-design process with women from Washington Heights, New York City [3,10]. We established collaborations with 3 community partner organizations (Shape Up NYC, Jovie, and the West Side Campaign Against Hunger [WSCAH]). These partners participate in research team meetings and in the study design. Partner details are provided as follows:

- Shape Up NYC: New York City Department of Parks and Recreation programs include Shape Up NYC, a free group fitness program offering 87 fitness classes at 60 locations

XSL•FO
RenderX

across the 5 New York City boroughs. Shape Up NYC will determine and schedule available study sites and classes in coordination with the intervention.
- Jovie: Jovie is a professional childcare company that provides care for families and businesses. We partnered with Jovie during the feasibility study, and they will continue providing childcare.
- WSCAH: WSCAH is one of the largest food pantries that serves New York City and is distinctly committed to providing healthy food through a customer-choice supermarket-style pantry. WSCAH will coordinate volunteer activities and food delivery for participants in the arm receiving peer support.

These partners are members of our community advisory board (CAB). The CAB also includes 4 community liaisons and 2 community champions. The community liaisons are individuals who represent or serve our target communities, the community partners represent our organizational partners, and the community champions are local women who will help facilitate peer support activities.

The main CAB functions are to ensure that the study's research strategies and design respect and serve our participants and make sense "in the real world." The CAB will meet quarterly to troubleshoot any problems with the trial procedures. They will also bring community concerns or ideas to the research team.

Participants

Eligibility Criteria

The trial will enroll mothers aged 18 years and older, who understand or speak English or Spanish, have children younger than 12 years of age (because younger children require more care, and we are focused on preventing chronic disease among premenopausal women), own a mobile phone, and live in nearby zip codes surrounding Shape Up NYC sites. Mothers must also meet income requirements, with total household income less than or equal to 165% of the area median income (calculated by household size).

Clusters

We will recruit 720 participants into 24 Shape Up NYC classes, across 1 pilot wave (n=90) and 7 study waves (n=630) over 2 years (Figure 2). Shape Up NYC class sites that allow research and have adjacent space for childcare will be eligible for selection. We will randomize each wave of 3 sites into 1 of 3 arms, using computer-generated random numbers, and run the program for 12 weeks. We will first randomize the class sites and then recruit participants into them while blinded.

Figure 2. Staggered recruitment.

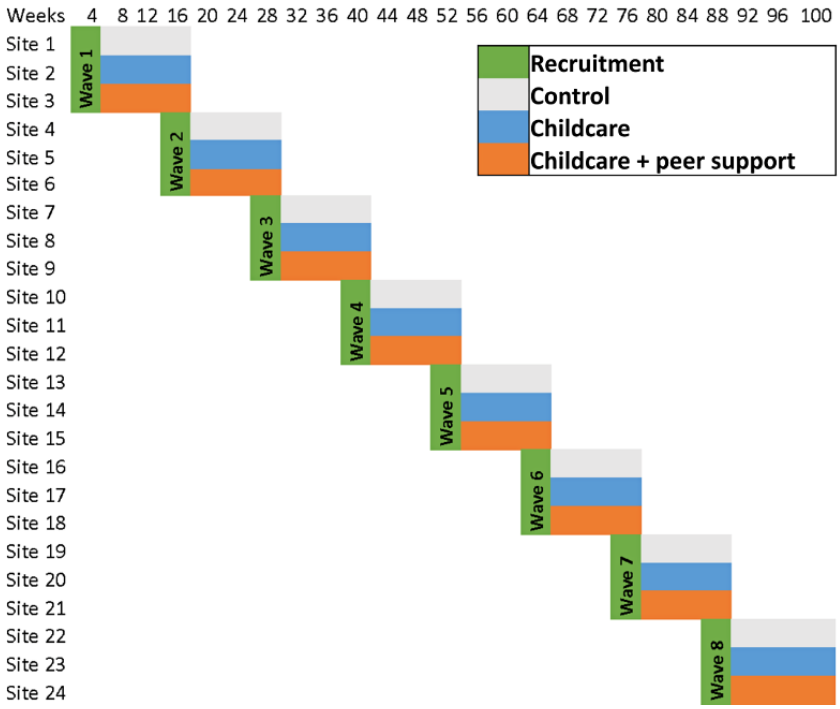

For each wave, 30 participants will be recruited into each class (n=90 per wave). This process will be repeated for each wave with 3 classes, some of which will be located at the same sites as previous classes in previous waves. We will stratify randomization by geographic location and across seasons. While we cannot stratify on individual-level factors in a cluster RCT, stratification by geographic location will help to capture many social demographic factors at the individual level (eg, race or ethnicity and income level). In addition, Shape UP NYC trains fitness instructors at all sites following the same curriculum, and all arm C participants will have the same community champions (described more in the Peer Support Activities (Arm C Only) section).

Blinding

The biostatistician will randomize deidentified fitness class sites into arms and will keep this blinded from the principal investigator (PI: LCH) and project coordinators while recruitment occurs. All the study team members will be blinded

XSL•FO
RenderX

except for the senior ethnographer, field ethnographer, community champions, and community partners.

Sample Size

We powered the study to compare the main intervention effects on the primary outcome, which is the average time spent in moderate to vigorous physical activity (MVPA) per week. We plan to enroll 7 sites per study arm (excluding the pilot wave), with about 30 participants with up to 45 children per site. We assume an attrition rate of 30% based on our feasibility study as well as another similar study [19] and an additional 25% accelerometer nonadherence rate based on previous studies [19,20], yielding an average 15 participants per site. Previous studies show that the intraclass correlation coefficient is ~0.15, and we expect a coefficient of variation of 0.5 in the number of participants across different sites. For aim 1, with 80% power, the minimum detectable effect size for the comparison of intervention arm B versus control arm A on average time spent in MVPA is Cohen *d*=0.68 at 5% significance. For aim 2, we will use the Bonferroni correction to adjust for multiple comparisons of arm C versus arm B and arm C versus arm A. With 80% power, the minimum detectable effect size for each comparison is Cohen *d*=0.75 at 5% significance. A previous physical activity intervention study with mothers led to an 80-minute increase in average time spent in MVPA [20]. Specifically, at posttest, the experimental group performed a mean of 154.8 (SD 45.3) minutes of physical activity, whereas the control group reported a mean of 74.6 (SD 44.4) minutes. This yields an effect size of *d*=1.8; thus, we are adequately powered for aims 1 and 2. Finally, quantitative analyses are not fully appropriate to examine intersectionality (eg, too reductionist), rendering a power analysis inappropriate; instead, we will integrate the statistical analysis with qualitative data to evaluate aim 3.

Recruitment

We will use multiple recruitment methods to identify and recruit participants. We will distribute paper and digital flyers and advertise FT4W on social media platforms including Facebook and Instagram. WSCAH, Jovie, and their partner community-based organizations will circulate our recruitment flyer to their constituents via email. Our project coordinator will do direct in-person outreach at community events and schools. The project coordinators will follow up with individuals interested in participating in the study by phone, email, or SMS text message.

To mask the purpose of the study, we will tell potential participants that the study seeks “to find the best way to connect moms living in the same neighborhood.” Eligible participants will provide informed consent after reviewing the institutional review board–approved document that explains the study and their involvement in it by providing their digital signature.

Interventions

Fitness Classes (All Arms)

Shape Up NYC offers classes that occur weekly for 1 hour free of charge to consenting participants in all 3 arms. Trained Shape Up NYC staff or volunteers lead the fitness classes, which include activities such as yoga, Zumba, and dance fitness. We will invite participants to attend the same scheduled class each week for the duration of their involvement in the 12-week study.

Childcare (Arms B and C)

Arms B and C will receive childcare from Jovie in an adjacent space to the fitness class.

Peer Support Activities (Arm C Only)

Overview

In addition to the fitness class and childcare, arm C will also receive peer support activities. Peer support activities will consist of group smartphone messaging facilitated by community champions, volunteer activities at food pantries, free deliveries of food pantry groceries, and group play dates at local playgrounds.

We structured group activities within the context of intergroup contact theory [21], social penetration theory [22], updated meta-analyses of this work [23,24], and research team members’ prior experience [25-27]. In short, positive communication and disclosure are essential to relationship building, which builds social cohesion. Sharing personal information (disclosure) and spending time together yield the largest effects in meta-analyses—these forms of intimacy grow over the course of developing relationships [22], and growth is more likely when anxiety is reduced and comfort is increased [28]. Anxiety reduction is the most important factor during the initial stages of developing social cohesion, hence our strategic structuring of peer support activities. Our intervention includes a task-based interaction (volunteer activities) where cognitive demand is low and interactional roles are clear. The focus is on the task, rather than developing relationships, reducing the pressure of interaction. As the volunteer activities continue, they also offer the opportunity for group members to enhance empathy, another salient component of cross-group relational development. Mediated interaction (group messaging) supports new relational development through a disinhibiting effect, lowering inhibitions and increasing comfort and disclosure [29].

Community Champions

Two bilingual (Spanish or English) community champions will foster social support and participation among arm C participants.

Weekly Playdates

Participants in arm C will be offered weekly group playdates for a total of 12 during the intervention. We will invite participants to bring their children to playgrounds in Shape Up NYC neighborhoods, where mothers can collectively watch their children play but also chat and reinforce social bonds. The first 2 playdates will have community champion–facilitated activities to help mothers get to know each other and to reduce the social effort and anxiety often associated with meeting new people.

Food Pantry Volunteering

We will invite arm C participants to volunteer as a group to help package food boxes during monthly events with WSCAH at one of their main locations. There is no childcare available for this activity, so participants will be told that they should not

XSL•FO
RenderX

bring their children. Most volunteering opportunities are during the day (around 10 AM) on weekdays, and shifts are a maximum of 2-3 hours.

Free Grocery Delivery

In the first week of the intervention, arm C participants will also be provided with a link to WSCAH's enrollment to opt into the home delivery of free groceries once per month for the 12-week intervention. There are no eligibility criteria to receive food from WSCAH, and participants will be eligible even if they do not volunteer.

Text Messages (All Arms)

Arms A and B

Arms A and B will receive individual text messages with a reminder for their weekly fitness class only.

Arm C

Arm C will be assigned to a messaging group with other mothers through WhatsApp, where community champions will "nudge" peer support and participation in activities through group messages. Community champions will follow scripted messages that remind participants when their next fitness class or peer support activity takes place. In addition, they will send weekly (prescripted) conversation prompts to "nudge" social interaction and will be able to "go off script" in response to organic group interactions. Participants will be allocated to groups first by language preference (ie, some groups may be in Spanish and some in English). The goals of the text messaging groups are (1) to encourage relationship-building, (2) to encourage participation in activities, and (3) to provide social support of various types (emotional, informational, etc).

Outcomes

Primary Outcome: Physical Activity

The primary outcome will be total minutes of time spent in MVPA per week. We will objectively assess weekly MVPA, total physical activity, sedentary time, and sleep using a blinded research-grade wrist-worn accelerometer (Axivity AX3) in all groups.

Secondary Outcomes

Self-Reported Physical Activity

To overcome missing accelerometer data, we will gather self-reported physical activity data using the validated International Physical Activity Questionnaire (IPAQ) [30]. IPAQ asks participants to report the types and frequency of physical activities they have participated in over the past 7 days. IPAQ is reported as a score used to categorize people as having high, moderate, or low levels of physical activity.

Health Status and Well-Being

We will measure health status and well-being using 2 standardized surveys: EuroQol 5 Dimension 5 Level [31] and ICEpop Capability Measure for Adults [32]. EuroQol 5 Dimension 5 Level is a generic measure of health status that includes 5 dimensions: mobility, self-care, usual activities, pain, and anxiety or depression. ICEpop Capability Measure for Adults will capture the broader concept of well-being via 5 dimensions: feeling settled and secure; love, friendship, and support; being independent; achievement and progress; and enjoyment and pleasure.

Implementation Outcomes

Maintenance and Sustainability

Recent advancements in implementation science distinguish maintenance from sustainability by defining the amount of time, with maintenance lasting 6 months or more and sustainability beginning at 1 year or more. We will assess maintenance and sustainability at the organizational level through interviews with 6 stakeholders, 3 from Shape UP NYC and 1 each from Jovie, WSCAH, and the study team. We will use the Extension of the Reach, Effectiveness, Adoption, Implementation, and Maintenance framework to guide our questions.

Fidelity

The fidelity of the intervention will be measured during the 12-week trial periods in each wave. We will measure the dose delivered in each cluster by tracking the number of fitness classes and peer support activities offered, childcare availability, and the number and type of texts sent in each group. We will also measure the dose received according to attendance to all components of the intervention (fitness class, childcare, play dates, and volunteer activities) and the type and number of text message engagements. The field ethnographer will also observe how the fitness instructor delivers the class (physical activity level) and the interactions among participants before, during, and after the class (social cohesion).

Attendance

An additional process outcome will be attendance at fitness classes. Participation rates are arguably the most important process outcome of interest when evaluating the success of an intervention like FT4W. Shape Up NYC takes attendance at each fitness class, and community champions will take attendance for the peer support activities to understand participation, retention, and attrition levels.

Mechanism of Action (Effect Mediators) and Modifiers

Resources

We will ask participants to report resources they have access to at baseline and follow-up. We will ask the same FT4W-specific questions as in the feasibility study [10] and use the following constructs from the Social Determinants of Health Collection of the PhenX Toolkit, a set of standardized measurement protocols used in health equity research [33,34]: health insurance coverage and food insecurity. We will also measure access to childcare [35].

Social Cohesion and Support

We will measure participants' perception of social cohesion in their neighborhood with 4 previously validated, standardized questions from the 2018 National Health Interview Survey using a 5-point Likert scale [36]. Questions ask about participants' perceptions of help availability, neighbors' accountability, trust in neighbors, and close-knit neighbors in their neighborhood. We will also administer the Medical Outcomes Study Social Support Survey, which asks participants to rate how often

XSL•FO
RenderX

different types of social support are available to them when needed, such as someone to confide in or someone to help with chores if sick [37]. Finally, we will also administer the Sense of Community Index-2 at follow-up only (specifically asking about the FT4W group), which captures group belonging by examining the intersection of interpersonal and community influence in the sociocultural environment, byproducts of human relationships that are difficult to capture by measuring social networks and peer support. Previous evidence demonstrates that the Sense of Community Index-2 strongly predicts behaviors such as participation [38,39].

Intersecting Social Stratifiers

Intersectionality considers the interaction of different social stratifiers (eg, race or ethnicity, indigeneity, gender, class, sexuality, geography, age, disability or ability, migration status, and religion) and underlying multilevel power structures [33]. We will administer questionnaires at baseline and follow-up and use the responses for individual quantitative data. At baseline, we will assess individual social stratifiers using measures from the Social Determinants of Health Collection of the PhenX Toolkit [40]: ethnicity and race, biological sex assigned at birth, gender identity, sexual orientation, educational attainment (individual), health insurance coverage, and current employment status. We will also ask about migration status, language preference, number and age of children, as well as the degree to which their neighborhood supports physical activity through the Physical Activity Neighborhood Environment Survey [41]. While previous health disparity research considers forms of social stratification in isolation or in an additive manner [40], intersectionality is multilevel and calls for analysis that is more than just a sum of its parts. Thus, quantitative methods are limited. We will use qualitative methods to explore how intersecting factors at the individual, community, and institutional levels modify effectiveness and implementation. We will specifically explore how SEP, race, and ethnicity intersect when it comes to the FT4W experience.

Data Collection

Quantitative Data Collection

Baseline Assessment

Participants will receive individualized links via text message to complete the online baseline questionnaire, which assesses participant demographics, physical activity, health status and well-being, social cohesion, access to childcare, and individual- and community-level social determinants of health.

Study staff will send a preprogrammed Axivity AX3 watch mailed directly to the participant following completion of the baseline survey (week 0). Participants will be asked to wear the Axivity watch on their nondominant wrist for 24 hours per day for 8 full days and to complete the accompanying sleep diary every day, either on paper or via text message, based on their preference. Instructions on how to use the watch and complete the diary will be included in the parcel via written materials and a brief video. Following 8 days of wear, the participant will post the accelerometer (free of charge in prestamped envelopes) back to the project coordinator, who will check the data for completeness. If the data do not meet our quality control checks and there are at least 2 weeks remaining before the first scheduled class, then we will mail another accelerometer to the participant. A previous study conducted with the same population found nonadherence to waist accelerometers to be 35% [42]. Data from the National Health and Nutrition Examination Survey also show that adherence to wrist accelerometers is higher (70%-80%) than waist accelerometers (40%-70%); therefore, we expect adherence to wrist accelerometers to be 75% in our study [43]. To increase adherence, we will send automated text reminders to confirm receipt of the accelerometer, to complete the accompanying sleep diary every day, either on paper or via text message, and to return the accelerometer.

Follow-Up Assessment

At week 10, study staff will again mail Axivity AX3 watches to participants and links to complete the online follow-up questionnaire, following the same procedure as baseline. We will consider them lost to follow-up if we do not receive a response by week 16.

Qualitative Data Collection

Overview

Ethnographic methods will be used to examine how intersecting forms of social inequality shape women's experiences of physical activity in the context of their broader lives (studying the experiences of research participants) and to understand how real-world conditions shape the intervention implementation (studying the workings of the research team and the impacts of the changing social context on study implementation). The field ethnographer will be bilingual and of the same approximate age as the research participants and will be supervised by the team's senior ethnographer.

Ethnographic Research With Study Participants and Intervention Activities

The field ethnographer will conduct participant observation, which will entail participating in at least 6 fitness classes (2 in each arm) and 4 peer support activities in each study wave throughout the trial. Through participant observation, the ethnographer will use a structured note-taking form to record observations about group dynamics, social environment, expressions of cultural norms, expressions of frustration at challenges of getting to exercise class, and other individual comments or social interactions relevant to understanding challenges women with children face in engaging in physical activity, including related to childcare, as well as how the intervention assets (childcare and social support) do or do not facilitate their participation in Shape Up NYC classes. For each instance of participant observation, the ethnographer will also record the time of day, weather, and time of year to assess temporal impacts on participation. The ethnographer will take both brief field notes on their phone during childcare drop-off, before and after classes, and during social support activities, expanding these into more complete notes after leaving the study site. Notes will describe the physical conditions, the effect of mother-child dyads at drop-off and pickup, how participants engage with each other and the instructor, and what is said or

XSL•FO
RenderX

expressed through body language about the classes and activities.

Group message data for arm C will serve as an additional source of qualitative data. Through content analysis, researchers will assess topics of discussion, jokes, expressions of friendliness, the development of connections between research participants, the sharing of information, and other content types to shed light on what social support consists of and how it might affect participation in physical activity. We will also quantify the number of interactions in the chat per wave.

Ethnographic Case-Control Research

Through comparing participants who engage in high and low rates of physical activity, ethnographic case-control research may shed light on otherwise unmeasured factors that affect the outcome as well as social, cultural, and structural barriers to physical activity. Purposive sampling for the ethnographic case-control study will use data from accelerometers to identify women with low and high levels of each outcome and then recruit individuals who meet the criteria for diversity specified in Table 1. We will select pairs of cases and controls within arms and match on race and ethnicity. We will select pairs from all arms within the same wave.

Table 1. Ethnographic case-control sample.

Race and ethnicity	High physical activity, n	Low physical activity, n
Black		
Non-Hispanic	4	4
Hispanic	4	4
White		
Hispanic	4	4

After the week 12 follow-up, the field ethnographer will conduct semistructured interviews with those recruited for this purposive sample to gain further insight into participants' views and experiences with FT4W and specifically to understand differences in individual characteristics, social context, or access to resources, which might lead the intervention to be impactful for some and not for others. Interviews will be audio-recorded and transcribed. We expect to complete ~48 interviews or until we reach data saturation as recommended for this type of study [44].

Development and Fidelity of Intervention Protocol

Ethnographic research conducted by the senior ethnographer, with a focus on the development of the intervention itself, will generate an understanding of how factors internal to the team (team process, particularities, and power dynamics) and factors external to the team (the political and social context) shape the development and implementation of the intervention protocol, including fidelity to the proposed intervention.

The study itself is an object of inquiry. To do this, the senior ethnographer will participate in all study team and CAB meetings and use a checklist to track development, fidelity, and changes to the protocol based on real-world demands. In addition, the senior ethnographer will conduct key informant interviews with study team members and community partners to learn about community organizations' experience of the intervention implementation, including policy factors and other potential challenges to sustainability. Finally, over the duration of the project, the senior and field ethnographers will compile information about how the broader social context (eg, budget cuts due to the migrant crisis in New York City) shapes the implementation of the project.

Data Monitoring

Data monitoring by the study team will occur at least on a biannual basis (ie, every 6 months). Participant enrollment (including compliance with protocol enrollment criteria), data completeness, and the status of all enrolled participants will be reviewed by the PI and supported by the study team.

This study will be stopped prior to its completion if (1) the intervention is associated with adverse effects that call into question the safety of the intervention, (2) difficulty in study recruitment or retention will significantly impact the ability to evaluate the study end points, (3) any new information becomes available during the trial that necessitates stopping the trial, or (4) other situations occur that might warrant stopping the trial.

Analyses

Quantitative Analysis

The primary analyses will be on the intention-to-treat sample: all randomized participants according to the assigned treatment. Missing data on outcome variables will be dealt with by using multiple imputation or inverse probability weighting of cases with complete data, where we will calculate weights based on the probability of a participant being a completer versus a dropout. We will conduct a sensitivity analysis to provide a range of plausible effect estimates that could arise due to nonignorable missing data.

For aims 1 and 2, we will use linear mixed-effects models (LMMs) to assess the main intervention effects on physical activity outcomes and other secondary outcomes. The outcome measure of interest will be the dependent variable; the treatment arm (categorical variable with 3 categories) will be the predictor. Cluster-specific random intercept will account for intraclass correlation. We will further adjust the model for baseline covariates that are not balanced across treatment groups (if any). Comparison of treatment arms can be tested by forming contrasts of regression coefficients in the models.

We will conduct mediation analysis to test whether environmental resources and social cohesion, described under

XSL•FO
RenderX

working mechanisms, mediate the intervention effects. For each mediation analysis, we will conduct 2 LMM analyses. The first model will use the outcome measure as a dependent variable and treatment arm and mediator as predictors. The second model will use the mediator as a dependent variable and treatment arm as the predictor. Both models will include a cluster-specific random intercept to account for intraclass correlation. Causal mediation effects will be tested using the R package *mediation*. For aim 3, moderator analysis will examine whether intervention effects differ across social stratification parameters (stratifier) at the individual and community levels. We will conduct a similar LMM analysis as in the main effect analysis, with the exception that each stratifier and its interaction with the treatment arm will be included in the LMM. The moderator effect will be examined by testing the stratifier by treatment interaction effect coefficients.

While 21 clusters are on par with the average cluster size in cluster RCTs, we still need to correct for the small sample since type I error can be up to ~7% for studies with 20-40 clusters [45]. For our continuous outcomes, residual (restricted) maximum likelihood will be used [46]. We will also adjust the test statistic to a 2-tailed *t* test with 2K–2 degrees of freedom if cluster structures are balanced [47]. If cluster sizes vary, alternative small sample corrections will be used. As a sensitivity analysis, cluster-level analysis will also be conducted. We will perform a 2-tailed *t* test on the cluster-level means of physical activity.

Qualitative Analysis

We will thematically code the ethnographers' field notes and transcribe participant interviews for both deductive and inductive codes. A codebook will be created using thematic coding by reviewing field notes and transcripts, grouping information into meaningful categories, and creating descriptive labels. The ethnographer will apply the codes to all text. The PI will double-code 10% of the content. We will develop themes from the codes and assess where themes converge and diverge across the ethnographic case-control sample.

Quantitative and Qualitative Data Integration

Mixed methods analysis will integrate the cluster RCT results with the ethnographic data to better understand the underlying processes explaining why the intervention is or is not effective (Figure 3). Quantitatively, we will assess effectiveness and mediation. We will also test for additive interaction by stratifying the average effect from statistical models by race and ethnicity and income. However, because intersectionality is not additive, we will heavily rely on qualitative data analysis to understand how intersecting forms of social inequality shape participants' physical activity, environmental resources, and social cohesion, all components that influence effectiveness as indicated in our logic model (Figure 1).

Figure 3. Embedded mixed method design. FT4W: Free Time for Wellness; RCT: randomized controlled trial.

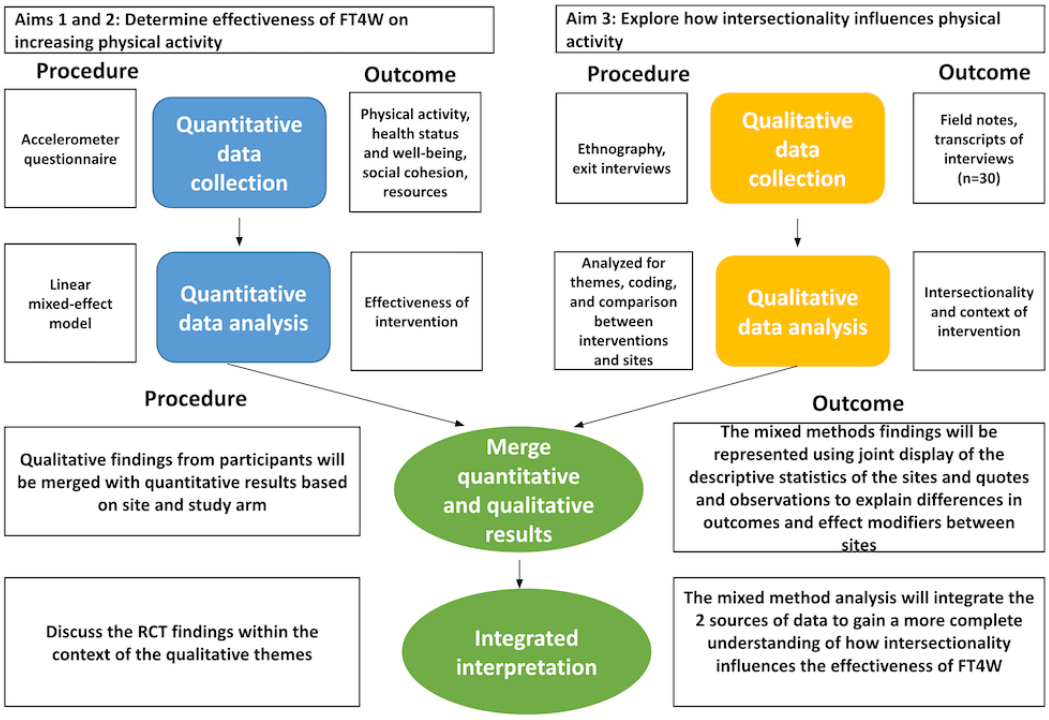

The quantitative and qualitative data will be integrated in several ways. First, we will examine where qualitative themes from the structured interviews converge and diverge based on accelerometer, social cohesion, and demographic data. We will use a joint display of cluster-level descriptive statistics and corresponding quotes to explain differences in outcomes and effect modifiers between clusters. Second, the participant observation of the intervention activities will reveal the unmeasured elements of the intervention that might affect either effectiveness or fidelity across clusters and arms. We will use narrative discussion to merge quantitative and qualitative data at the community level. Third, the institutional ethnography will provide preliminary implementation outcomes that we will use to develop future quantitative measures of fidelity, maintenance, and sustainability.

XSL•FO
RenderX

Timeline

Table 2 summarizes the schedule of enrollment, interventions, and assessments for 1 wave of study participants. As outlined in Figure 2, there will be a pilot wave followed by 7 waves.

Table 2. Schedule of enrollment, interventions, and assessments for 1 wave.

	Study period												
	Allocation	Enrollment	Postallocation										Closeout
		0	W1[a]	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	
Enrollment															
Eligibility screen		✓													
Informed consent		✓													
Allocation	✓	✓													
Interventions															
Arm A (fitness class control)			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Arm B (fitness class and childcare)			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Arm C (fitness class, childcare, and peer support activities)			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Assessments															
Baseline accelerometer (AX3) measurement		✓													
Baseline questionnaire		✓													
Attendance data collection			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Follow-up accelerometer (AX3) measurement												✓	✓	✓	
Follow-up questionnaire												✓	✓	✓	
Interviews (for purposively selected participants only)															✓

[a]W: wave.

Ethical Considerations

Columbia University's institutional review board granted ethics approval (reference AAAU8303). All participants will provide written informed consent prior to participation, and the informed consent materials are provided in Multimedia Appendix 1. Only the PI, project coordinator, and the ethnographers will have access to identifiable data. All other research team members will have access to only deidentified data. Data will be stored on Columbia University computer systems in a custom-built database. Enrolled participants will receive a gift card after completing the baseline survey and returning the accelerometer and another gift card at week 12 follow-up after completing the survey and returning the accelerometer. Interviewees will receive another gift card in addition to the trial incentive.

Results

The study received funding from the US National Institute of Health covering the period of time from April 1, 2023, through March 31, 2028. We received initial institutional review board approval in August 2023. The study is active and recruiting participants. As of the day of manuscript submission, we have enrolled 471 participants. Data collection is anticipated to occur until September 2026 for primary completion. The estimated study completion date is December 2026.

In line with our community-focused approach to research, we plan to communicate our study results back to the participants, partners, and New York City community using appropriate methods, such as reports, social media, or events. In addition, we will disseminate our approach and our findings to other researchers to help inform the further development and expansion of the FT4W intervention and similar interventions to promote physical activity among mothers. This will be done through the publication of this protocol, followed by reporting on any changes that take place during implementation, as well as the qualitative and quantitative results through academic publications and conferences.

Discussion

Anticipated Findings

The anticipated results from this study will increase our understanding of the potential roles of childcare and peer support in promoting physical activity among low-resourced mothers in New York City. The mixed methods results will also help us to better understand the mechanisms and the impact of

XSL•FO
RenderX

intersecting factors such as race or ethnicity, culture, gender, and SEP on mothers' physical activity. Finally, implementation results will help us to better understand how this program could be scaled up in the future, if successful, and how future similar programs should be designed.

Strengths and Limitations

This study has several strengths. First, this cluster RCT tests a structural barrier to physical activity for mothers, lack of childcare, which is understudied. It also integrates ethnographic methods with a cluster RCT to better understand mechanisms and the impact of intersecting factors such as race or ethnicity, culture, gender, and SEP. Finally, the study leverages widely accessible, existing resources to promote physical activity (free Shape Up NYC fitness classes) and foster social support (WhatsApp groups to connect mothers living in the same area).

Despite these strengths, the study also has some limitations. First, it is limited to participants in New York City, which may limit its generalizability to other settings. The intervention includes options for Spanish-speaking participants, but not other languages, despite the diversity of languages spoken in New York City. This limitation, as well as the sample size, may result in an inability to capture a wide diversity of experiences. In addition, participants will only be followed for 12 weeks, so the study results will not be able to provide insight into potential long-term effects and sustainability. In addition, the use of accelerometers to collect physical activity data was not tested in our earlier feasibility study due to resource constraints. Finally, the use of existing Shape Up NYC fitness classes creates some potential for contamination across arms, but locations within a wave will be selected carefully to minimize this risk.

Future Directions

The goal of the trial is to demonstrate that childcare and social cohesion are effective so that they can be implemented long-term in the future. If the intervention is effective, then the New York City local government will have data to justify financial support to implement the childcare component of their existing free fitness program. The implementation results from this study will help to identify what would be needed at the individual, implementation, and organizational levels for the intervention to be maintained and sustained over time.

Conclusions

This cluster RCT tests a structural barrier to physical activity for mothers, lack of childcare, which is understudied. The study design and outcomes are geared toward integrating ethnographic methods with a cluster RCT to better understand mechanisms and the impact of intersecting factors such as race or ethnicity, culture, gender, and SEP. The study leverages widely accessible, existing resources to promote physical activity and foster social support with the ultimate goal of assessing the effect of childcare access on parental health.

Acknowledgments

This work is supported by the National Institute on Minority Health and Health Disparities (grant 1R01MD018609-01). This work is sponsored by Angela R Fernandez, MPH, LCSW, PhD.

Authors' Contributions

All authors contributed to the development of the study methodology outlined in this protocol. JW drafted the manuscript using some relevant content written for other unpublished documents (proposals, internal protocols, and institutional review board application) by all members of the authorship team (eg, MQ wrote the section on quantitative data analysis, JSH wrote the sections on ethnographic methods, and LCH wrote the outcomes and data collection sections). All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Informed consent materials.
[DOCX File , 23 KB - resprot_v15i1e71381_app1.docx]

References

1. National Center for Health Statistics. Early release of selected estimates based on data from the 2018 National Health Interview Survey. Center for Disease Control and Prevention. 2019. URL: https://archive.cdc.gov/#/details?url=https://www.cdc.gov/nchs/nhis/releases/released201905.htm [accessed 2025-09-13]
2. McNeill LH, Kreuter MW, Subramanian S. Social environment and physical activity: a review of concepts and evidence. Soc Sci Med 2006 Aug;63(4):1011-1022. [doi: 10.1016/j.socscimed.2006.03.012] [Medline: 16650513]
3. Houghton LC, Gibbons MP, Shekelle J, Oakley-Girvan I, Watterson JL, Magsamen-Conrad K, et al. Free Time For Wellness: a co-designed intervention utilizing social networks to encourage physical activity for cancer prevention among low resourced mothers. BMC Public Health 2021 Oct 07;21(1):1805 [FREE Full text] [doi: 10.1186/s12889-021-11775-9] [Medline: 34620141]

XSL•FO
RenderX

4. Brown PR, Brown WJ, Miller YD, Hansen V. Perceived constraints and social support for active leisure among mothers with young children. Leis Sci 2001 Jul;23(3):131-144. [doi: 10.1080/014904001316896837]
5. Mansfield ED, Ducharme N, Koski KG. Individual, social and environmental factors influencing physical activity levels and behaviours of multiethnic socio-economically disadvantaged urban mothers in Canada: a mixed methods approach. Int J Behav Nutr Phys Act 2012;9:42 [FREE Full text] [doi: 10.1186/1479-5868-9-42] [Medline: 22500882]
6. Jackson D, Mannix J. Mothering and women's health: I love being a mother but ... there is always something new to worry about. Aust J Adv Nurs 2024;20(3):30-37. [doi: 10.37464/2003.203.2037]
7. Currie J. Anxiolytic benefits of exercise classes for mothers. J Womens Health Care 2018;7:2 [FREE Full text] [doi: 10.4172/2167-0420.1000442]
8. Bellows-Riecken KH, Rhodes RE. A birth of inactivity? A review of physical activity and parenthood. Prev Med 2008;46(2):99-110. [doi: 10.1016/j.ypmed.2007.08.003] [Medline: 17919713]
9. Currie J. Motherhood, stress and the exercise experience: freedom or constraint? Leis Stud 2004;23(3):225-242. [doi: 10.1080/0261436042000251987]
10. Jones C, Gibbons M, Magsamen-Conrad K, Ulanday KT, Watterson J, Oakley-Girvan I, et al. "Mother's health and well-being matters: is a mediated social cohesion public health intervention feasible?". Am J Health Promot 2022;36(3):410-420 [FREE Full text] [doi: 10.1177/08901171211055317] [Medline: 34907785]
11. Schneider S, Diehl K, Görig T, Schilling L, De Bock F, Hoffmann K, et al. Contextual influences on physical activity and eating habits—options for action on the community level. BMC Public Health 2017;17(1):760 [FREE Full text] [doi: 10.1186/s12889-017-4790-x] [Medline: 28964266]
12. Oakley-Girvan I, Watterson JL, Jones C, Houghton LC, Gibbons MP, Gokal K, et al. Use of social media for cancer prevention through neighborhood social cohesion: protocol for a feasibility study. JMIR Res Protoc 2021;10(7):e28147 [FREE Full text] [doi: 10.2196/28147] [Medline: 34328445]
13. Centola D. The spread of behavior in an online social network experiment. Science 2010 Sep 03;329(5996):1194-1197. [doi: 10.1126/science.1185231] [Medline: 20813952]
14. Beck AM, Eyler AA, Aaron Hipp J, King AC, Tabak RG, Yan Y, et al. A multilevel approach for promoting physical activity in rural communities: a cluster randomized controlled trial. BMC Public Health 2019;19(1):126 [FREE Full text] [doi: 10.1186/s12889-019-6443-8] [Medline: 30700262]
15. Peralta LR, Cotton WG, Dudley DA, Hardy LL, Yager Z, Prichard I. Group-based physical activity interventions for postpartum women with children aged 0-5 years old: a systematic review of randomized controlled trials. BMC Womens Health 2021;21(1):435 [FREE Full text] [doi: 10.1186/s12905-021-01581-1] [Medline: 34963456]
16. Cuschieri S. The CONSORT statement. Saudi J Anaesth 2019 Apr;13(Suppl 1):S27-S30 [FREE Full text] [doi: 10.4103/sja.SJA_559_18] [Medline: 30930716]
17. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
18. Johnston M, Carey R, Connell BL, Johnston D, Rothman A, de BM. Development of an online tool for linking behavior change techniques and mechanisms of action based on triangulation of findings from literature synthesis and expert consensus. Transl Behav Med 2021;11(5):1049-1065. [doi: 10.1093/tbm/ibaa050]
19. Audrey S, Procter S, Cooper A, Mutrie N, Hollingworth W, Davis A. Employer schemes to encourage walking to work: feasibility study incorporating an exploratory randomised controlled trial. Public Health Res 2015;3(4):1-60. [Medline: 25763450]
20. Fahrenwald NL, Atwood JR, Walker SN, Johnson DR, Berg K. A randomized pilot test of "Moms on the Move": a physical activity intervention for WIC mothers. Ann Behav Med 2004;27(2):82-90. [doi: 10.1207/s15324796abm2702_2] [Medline: 15026292]
21. Allport G. The nature of prejudice. In: Unabridged, 25th Anniversary Edition. Reading, MA: Addison-Wesley Pub Co; 1979:537.
22. Altman I, Taylor DA. Social Penetration: The Development of Interpersonal Relationships. New York, NY: Holt, Rinehart and Winston; 1973:240.
23. Davies K, Tropp LR, Aron A, Pettigrew TF, Wright SC. Cross-group friendships and intergroup attitudes: a meta-analytic review. Pers Soc Psychol Rev 2011;15(4):332-351. [doi: 10.1177/1088868311411103] [Medline: 21844287]
24. Sillars AL, Vangelisti AL. Communication: basic properties and their relevance to relationship research. In: Vangelisti AL, Perlman D, editors. The Cambridge Handbook of Personal Relationships. 2nd Edition. Cambridge: Cambridge University Press; 2018.
25. Magsamen-Conrad K, Dillon JM, Billotte Verhoff C, Joa CY. Toward a theory of HealthIT adoption across the lifespan: findings from five years in the community. Health Commun 2020;35(3):308-321. [doi: 10.1080/10410236.2018.1563027] [Medline: 30652503]
26. Magsamen-Conrad K, Dillon J, Hanasono L, Valdez P. Developing an intergroup communication intervention curriculum: enhancing workforce skills across generations. In: Developing Workforce Diversity Programs, Curriculum, and Degrees in Higher Education. New York City, NY: IGI Global; 2016.

XSL•FO
RenderX

27. Magsamen-Conrad M, Tetteh D, Lee YI, Smith M. A case study exploring attitudes about people with disability: evaluating the intergroup communication intervention. J Psychol Psychother 2017;7(4):1-8 [FREE Full text] [doi: 10.4172/2161-0487.1000321]
28. Tropp L. The role of trust in intergroup contact: its significance and implications for improving relations between groups. In: Improving Intergroup Relations. New Jersey: John Wiley & Sons, Ltd; 2008.
29. Joinson AN. Chapter 4—disinhibition and the internet. In: Psychology and the Internet. Burlington: Academic Press; 2007:75-92.
30. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
31. Helping the world make better health decisions. EuroQol. URL: https://euroqol.org/ [accessed 2024-07-09]
32. Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Qual Life Res 2012;21(1):167-176 [FREE Full text] [doi: 10.1007/s11136-011-9927-2] [Medline: 21598064]
33. Hendershot T, Pan H, Haines J, Harlan WR, Marazita ML, McCarty CA, et al. Using the PhenX Toolkit to add standard measures to a study. Curr Protoc Hum Genet 2015;86:1.21.1-1.21.17 [FREE Full text] [doi: 10.1002/0471142905.hg0121s86] [Medline: 26132000]
34. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. The PhenX Toolkit: get the most from your measures. Am J Epidemiol 2011;174(3):253-260 [FREE Full text] [doi: 10.1093/aje/kwr193] [Medline: 21749974]
35. Childcare arrangements. PhenX Toolkit. URL: https://www.phenxtoolkit.org/protocols/view/300301?origin=search [accessed 2024-04-18]
36. NHIS: 1997-2018. Centers for Disease Control and Prevention. 2021. URL: https://archive.cdc.gov/#/details?url=https://www.cdc.gov/nchs/nhis/1997-2018.htm [accessed 2024-07-09]
37. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32(6):705-714. [doi: 10.1016/0277-9536(91)90150-b] [Medline: 2035047]
38. Sawyer A, Ucci M, Jones R, Smith L, Fisher A. Simultaneous evaluation of physical and social environmental correlates of physical activity in adults: a systematic review. SSM Popul Health 2017;3:506-515 [FREE Full text] [doi: 10.1016/j.ssmph.2017.05.008] [Medline: 29349241]
39. Gilbert KL, Quinn SC, Goodman RM, Butler J, Wallace J. A meta-analysis of social capital and health: a case for needed research. J Health Psychol 2013;18(11):1385-1399 [FREE Full text] [doi: 10.1177/1359105311435983] [Medline: 23548810]
40. Larson E, George A, Morgan R, Poteat T. 10 Best resources on… intersectionality with an emphasis on low- and middle-income countries. Health Policy Plan 2016 Oct;31(8):964-969. [doi: 10.1093/heapol/czw020] [Medline: 27122486]
41. Sallis J, Kerr J, Carlson J, Norman G, Saelens B, Durant N, et al. Evaluating a brief self-report measure of neighborhood environments for physical activity research and surveillance: Physical Activity Neighborhood Environment Scale (PANES). J Phys Act Health 2010;7(4):533-540. [doi: 10.1123/jpah.7.4.533] [Medline: 20683096]
42. Cato MS, Wyka K, Ferris EB, Evenson KR, Wen F, Dorn JM, et al. Correlates of accelerometry non-adherence in an economically disadvantaged minority urban adult population. J Sci Med Sport 2020;23(8):746-752 [FREE Full text] [doi: 10.1016/j.jsams.2020.01.013] [Medline: 32085979]
43. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]
44. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017 Mar;27(4):591-608 [FREE Full text] [doi: 10.1177/1049732316665344] [Medline: 27670770]
45. Elff M, Heisig JP, Schaeffer M, Shikano S. Multilevel analysis with few clusters: improving likelihood-based methods to provide unbiased estimates and accurate inference. Br J Polit Sci 2020;51(1):412-426. [doi: 10.1017/s0007123419000097]
46. Satterthwaite FE. An approximate distribution of estimates of variance components. Biom Bull 1946;2(6):110. [doi: 10.2307/3002019]
47. Li P, Redden DT. Small sample performance of bias-corrected sandwich estimators for cluster-randomized trials with binary outcomes. Stat Med 2015;34(2):281-296 [FREE Full text] [doi: 10.1002/sim.6344] [Medline: 25345738]

Abbreviations

CAB: community advisory board
CONSORT: Consolidated Standards of Reporting Trials
FT4W: Free Time for Wellness
IPAQ: International Physical Activity Questionnaire
LMM: linear mixed-effects model
MVPA: moderate to vigorous physical activity
PI: principal investigator
RCT: randomized controlled trial
SEP: socioeconomic position

XSL•FO
RenderX

WSCAH: West Side Campaign Against Hunger

Edited by J Sarvestan; submitted 17.Jan.2025; peer-reviewed by S Huang, L Zhang, M Gasmi ; comments to author 26.Mar.2025; revised version received 26.Jun.2025; accepted 25.Aug.2025; published 13.Jan.2026.

Please cite as:
Watterson J, Magsamen-Conrad K, Gokal K, Oakley-Girvan I, Hirsch JS, Qian M, Buchsbaum R, Niles C, Van Horn K, Moshier S, Martinez A, Garcia Flores A, Laine S, Vega M, Houghton LC
Cluster Randomized Controlled Trial to Promote Physical Activity Among Low-Resourced Mothers in New York City: Protocol for the Free Time for Wellness Effectiveness Trial
JMIR Res Protoc 2026;15:e71381
URL: https://www.researchprotocols.org/2026/1/e71381
doi:10.2196/71381
PMID:

©Jessica Watterson, Kate Magsamen-Conrad, Kajal Gokal, Ingrid Oakley-Girvan, Jennifer S Hirsch, Min Qian, Richard Buchsbaum, Charlene Niles, Kendra Van Horn, Stephanie Moshier, Andreina Martinez, Avery Garcia Flores, Sherece Laine, Marilisa Vega, Lauren C Houghton. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Efficacy, Safety, and Economic Impact of Cytisinicline Maintenance Therapy in Patients Who Are Candidates for Smoking Cessation: Protocol for a Phase IV, Multicenter, Randomized, Open-Label, Controlled, Parallel Clinical Trial (CITISILONG Trial)

Carlos Rabade Castedo[1,2], PhD; Ana Estany-Gestal[3*], PhD; Carlos A Jiménez Ruiz[4*], PhD; José Ignacio de Granda-Orive[5,6*], PhD; Juan Antonio Riesco-Miranda[7,8,9*], MSc; María Isabel Cristóbal Fernández[4*], MSc; Angela Ramos-Pinedo[10*], MSc; Jaime Signes-Costa Miñana[11,12], PhD; María Inmaculada Gorordo-Unzueta[13,14*], MSc; Agustin Valido-Morales[15], PhD; Jacobo Sellarés -Torres[16*], PhD; Eva Cabrera-César[17*], PhD; Alejandro Frino-García[16], MSc; Luis Valdés Cuadrado[1,2,18], PhD

[1]Department of Pulmonology, Complejo Hospitalario Universitario de Santiago, Santiago de Compostela, Spain
[2]Health Research Institute, Santiago de Compostela, La Coruña, Spain
[3]Research Methodology Unit, Health Research Institute, Santiago de Compostela, Galicia, Spain
[4]Smoking Cessation Unit, Hospital Clínico San Carlos, Madrid, Spain
[5]Department of Pulmonology, Hospital Universitario 12 De Octubre, Madrid, Spain
[6]Department of Medicine, Universidad Complutense de Madrid, Madrid, Spain
[7]Department of Pulmonology, Hospital San Pedro de Alcántara, Cáceres, Spain
[8]Extremadura University Institute of Biosanitary Research, Cáceres, Spain
[9]Centre for Networked Research on Respiratory Diseases (CIBERES), Madrid, Spain
[10]Department of Pulmonology, Hospital Universitario Fundación Alcorcón, Madrid, Spain
[11]Department of Pulmonology, Hospital Clínico Universitario de Valencia, Valencia, Spain
[12]INCLIVA Health Research Institute, Valencia, Spain
[13]Department of Pulmonology, Hospital Universitario Galdakao-Usánsolo, Bilbao, Spain
[14]Biobizkaia Health Research Institute, Bilbao, Spain
[15]Department of Pulmonology, Hospital Universitario Virgen Macarena, Valencia, Spain
[16]Department of Pulmonology, Hospital Clínic de Barcelona, Barcelona, Spain
[17]Department of Pulmonology, Hospital Clínico Universitario Virgen de la Victoria, Málaga, Spain
[18]Department of Medicine, Universidade de Santiago de Compostela, Santiago de Compostela, Spain
[*]these authors contributed equally

Corresponding Author:
Carlos Rabade Castedo, PhD
Department of Pulmonology
Complejo Hospitalario Universitario de Santiago
Travesía da Choupana S/N
Santiago de Compostela, 15706
Spain
Phone: 34 981950000 ext 250000
Email: crabcas1@gmail.com

Abstract

Background: Cytisinicline has proven to be an effective, efficient, and safe molecule in smoking cessation. However, the established 25-day regimen could be insufficient in a high percentage of smokers, so it is necessary to study maintained therapies of this drug.

XSL•FO
RenderX

Objective: This study aims to compare the efficacy of the cytisinicline regimen used in routine clinical practice versus 2 maintained regimens of 50 and 75 days, respectively. In addition, the safety and economic impact of each regime will be determined.

Methods: A prospective, multicenter, open-label, controlled, parallel, phase IV clinical trial of 402 smoker patients prepared to quit smoking. The study was conducted in 10 hospitals in Spain. A control group is compared to 2 intervention groups in which the duration of the drug is increased without increasing its dose, administering half and all, respectively, of an additional marketed container that includes 100 tablets. Thus, participants will be randomized to three groups in a 1:1:1 ratio to receive cytisinicline: (1) a control group treated with cytisinicline according to the usual clinical guidelines and product information (25 days); (2) a group with a 50-day cytisinicline regimen (an additional 25 days at a dose of 1.5 mg every 12 hours), seeking to increase its efficacy while minimally impacting adherence; and (3) a group with a 75-day regimen (an additional 50 days at a dose of 1.5 mg every 12 hours), attempting to increase its efficacy, although the longer duration of the drug may threaten adherence. Efficacy in the 3 arms will be analyzed through sustained abstinence at 6 and 12 months, point abstinence rate assessed every 7 days, and abstinence rate from Day 25 to Day 50 and from Day 25 to Day 75 in the 3 study arms. (1) The variation in withdrawal and craving symptoms in the 3 groups, (2) safety through the percentage of adverse events in the 3 treatment arms, and (3) economic impact by evaluating the cost-effectiveness and cost-utility ratios of the 2 prolonged regimens versus the usual clinical cytisinicline regimen. To calculate the differences between the 3 groups for each outcome variable, a univariate analysis will be performed. Statistically significant variables will be included in a multivariate model.

Results: Recruitment for the trial and patient enrollment were completed in November 2026. Follow-up of all participants will extend to December 2027.

Conclusions: In conclusion, this study evaluates the optimization of cytisinicline in daily clinical practice, increasing the benefits of its pharmaceutical properties without affecting patient safety. All of this will improve the effectiveness of smoking cessation by reducing the number of smokers, which implies lower morbidity and mortality and lower costs associated with smoking.

Trial Registration: European Clinical Trials Register 2024-518936-36-00; https://euclinicaltrials.eu/ctis-public/view/2024-518936-36-00

International Registered Report Identifier (IRRID): PRR1-10.2196/76815

(JMIR Res Protoc 2026;15:e76815) doi:10.2196/76815

KEYWORDS

cytisinicline; smoking cessation; maintenance therapy; efficacy; safety

Introduction

Background

Cytisine (natural alkaloid), also called cytisinicline (marketed form), is extracted from a golden rain tree (*Cytisus Laburnum*) and has adverse effects similar to those reported for nicotine [1,2]. It is a partial agonist of the α4 β2 nicotinic receptors, which, when stimulated, lead to a moderate release of dopamine, contributing to the control of withdrawal symptoms. On the other hand, its antagonistic action blocks the receptor, preventing nicotine from binding to it, thus reducing the reward produced by smoking [1-4]. This duality makes cytisinicline a drug of choice for the treatment of smoking. Current scientific evidence has shown it to be a safe and effective smoking cessation molecule [5].

According to the data sheet [6] the adverse effects attributable to the use of cytisinicline, at standard doses, are mild and infrequent (<20%), with the most frequent (<20%) being gastrointestinal, dry mouth, nausea, dyspepsia, and abdominal pain. Cytisinicline is marketed as an over-the-counter drug in other European countries such as Poland, where it has been used by more than 7 million smokers, and pharmacovigilance data over the more than 50 years it has been on the market have not identified any serious or life-threatening adverse effects in smokers [7-17].

West et al [10] demonstrated that cytisinicline triples the chances of smoking cessation, and in 2 meta-analyses [11,12], where more than 4000 individuals were studied, a relative risk of 1.59 (95% CI 1.43-1.75) and 1.74 (95% CI 1.38-2.19) was found compared to placebo. These data are corroborated by new meta-analyses and systematic reviews [13-15].

In trials where it was compared with other drugs, cytisinicline demonstrated greater efficacy, measured in time to first relapse, compared with nicotine patch monotherapy [18]. Courtney et al [19], in a noninferiority trial, demonstrated that point abstinence at 4 weeks was 42.5% in the cytisinicline group vs 32.3% for varenicline, although noninferiority of cytisinicline vs varenicline could not be demonstrated at 6 months.

In Spain, cytisinicline is introduced as a smoking cessation drug in 2022 and starts to be funded by the National Health System in early 2023. The drug has demonstrated a high level of patient satisfaction and tolerability [20].

In addition, several authors have proposed various strategies to increase the efficacy/effectiveness of cytisinicline without increasing adverse effects. Rigotti et al [21], in a randomized, double-blind, placebo-controlled trial involving 810 smokers, compared 2 different durations of cytisinicline (6 or 12 weeks) versus placebo, with follow-up up to 24 weeks. For the 6-week course of cytisinicline, continuous abstinence rates were 25.3% vs 4.4% during Weeks 3-6 (odds ratio [OR] 8, 95% CI 3.9-16.3;

XSL•FO
RenderX

P<.001) and 8.9% vs 2.6% during Weeks 3-24 (OR 3.7, 95% CI 1.5-10.2; *P*=.002). For the 12-week course of cytisinicline, continuous abstinence rates were 32.6% vs 7% during weeks 9-12 (OR 6.3, 95% CI 3.7-11.6; *P*<.001) and 21.1% vs 4.8% during weeks 9-24 (OR 5.3, 95% CI 2.8-11.1; *P*<.001). Nausea, abnormal dreams, and insomnia occurred in less than 10% of each group. Sixteen participants (2.9%) discontinued cytisinicline due to an adverse event [21,22]. The Smoking cessation with cytisinicline (SMILE) clinical trial [23] was conducted in patients in a lung cancer screening program. The study aimed to compare a 40-day cytisinicline program with an 84-day program. However, the study did not reach an adequate sample size to demonstrate a significant benefit associated with prolonged treatment. However, the data showed that the adjusted OR of 12-month continuous abstinence between standard and prolonged treatment was higher with prolonged treatment: OR 1.5 (95% CI 1-2.3).

Another clinical trial evaluated maintenance therapy with cytisinicline (cytisinicline at a standard dose for 25 days plus two 1.5 mg tablets per day for maintenance from day 26 until 12 weeks of treatment) versus varenicline at a standard dose for 12 weeks, with no difference at 6 months follow-up [24]. That is, a prolonged 3-month course of cytisinicline could be as effective as other smoking cessation drugs such as varenicline without being associated with adverse effects.

In many countries, cytisinicline is prescribed as a single 25-day course. Studies to date show that the current 25-day cytisinicline regimen is better tolerated and has fewer side effects [20]. However, there is a patient profile in whom this 25-day cytisinicline regimen is not sufficient to control cravings and withdrawal symptoms, increasing the likelihood of relapse, reducing the chances of quitting smoking, and discouraging these smokers from making another serious quit attempt with an effective drug. Recent data from Spain on experience in routine clinical practice with a 25-day cytisinicline regimen show a high rate of relapse after completing treatment. Twenty percent of smokers treated with cytisinicline at this regimen resume smoking from the end of treatment until 2 months after follow-up [25]. Establishing an optimization strategy based on prolonging treatment could increase the drug's efficacy without being associated with a greater number of side effects, generating a favorable economic impact by increasing the number of smokers who permanently quit. We hypothesize that there is a correlation between increasing the duration of cytisinicline and its efficacy, without being associated with a greater number of adverse events/side effects.

Therefore, we aim to compare the efficacy, safety, and economic impact of this treatment regimen used in routine clinical practice in Spain versus 2 proposed extended treatment regimens (double and triple duration). Different drugs are not compared, but rather 3 different treatment regimens (25, 50, and 75 days) of the same drug (cytisinicline), one of which is the treatment regimen used in daily clinical practice.

Objectives

The overall objective of this study is to evaluate the efficacy, safety, and economic impact of 2 extended cytisinicline regimens (50 and 75 days duration) and compare them with the standard regimen (25 days).

The primary objective is to determine the efficacy of the maintained cytisinicline regimens (50 and 75 days) versus the standard 25-day regimen.

The secondary objectives are (1) to determine the percentage and number of adverse effects in each of the 3 treatment arms and (2) to analyze the economic impact of the maintained regimens versus the standard regimen in terms of cost-effectiveness and cost-utility.

Methods

Trial Design

Phase IV randomized, open-label, controlled, parallel, multicenter clinical trial in patients who are candidates for smoking cessation to demonstrate the superiority of maintenance therapy with cytisinicline compared to the standard regimen.

The trial was designed according to the standardized guide for interventional trial protocols, SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials statement [Multimedia Appendix 2]). A total of 402 patients from 10 smoking cessation units in Spain will be randomized. Recruitment will last for one year, with a follow-up of 52 weeks per patient recruited. Candidate patients who sign the informed consent form will be randomized to 3 possible arms in a ratio of 1:1:1.

Arm A: control group. Participants randomized to this group will be treated with cytisinicline as stated in the data sheet. They must remain smoke-free from day 5 of treatment.

Arm B: maintained therapy 1. Following the standard regimen, these patients will maintain treatment for an additional 25 days at a dose of 1.5 mg twice daily.

Arm C: maintained therapy 2. Following the standard regimen, these patients will maintain treatment for an additional 50 days at a dose of 1.5 mg twice daily (see Table 1).

At the end of the treatments, participants will continue in the follow-up phase, without treatment, until completing 52 weeks in the study (see Figure 1).

XSL•FO
RenderX

Table 1. Dosage guidelines in each of the 3 treatment arms.

Days of treatment	Recommended dosage	Maximum daily dose	Recommendations
Arm A: control group			
From the 1st to the 3rd day	1 tablet every 2 hours	6 tablets	—[a]
From the 4th to the 12th day	1 tablet every 2.5 hours	5 tablets	No smoking from the 5th day
From the 13th to the 16th day	1 tablet every 3 hours	4 tablets	No smoking
From the 17th to the 20th day	1 tablet every 5 hours	3 tablets	No smoking
From the 21st to the 25th day	2 tablets a day	Up to 2 tablets a day	No smoking
Arm B: maintained therapy 1			
From the 1st to the 3rd day	1 tablet every 2 hours	6 tablets	—
From the 4th to the 12th day	1 tablet every 2.5 hours	5 tablets	No smoking from the 5th day
From the 13th to the 16th day	1 tablet every 3 hours	4 tablets	No smoking
From the 17th to the 20th day	1 tablet every 5 hours	3 tablets	No smoking
From the 21st to the 25th day	2 tablets a day	Up to 2 tablets a day	No smoking
From the 26th to the 50th day	2 tablets a day	Up to 2 tablets a day	No smoking
Arm C: maintained therapy 2			
From the 1st to the 3rd day	1 tablet every 2 hours	6 tablets	—
From the 4th to the 12th day	1 tablet every 2.5 hours	5 tablets	No smoking from the 5th day
From the 13th to the 16th day	1 tablet every 3 hours	4 tablets	No smoking
From the 17th to the 20th day	1 tablet every 5 hours	3 tablets	No smoking
From the 21st to the 25th day	2 tablets a day	Up to 2 tablets a day	No smoking
From the 26th to the 50th day	2 tablets a day	Up to 2 tablets a day	No smoking

[a]Not applicable.

Figure 1. Trial design. Recruitment will last for one year, with a follow-up of 52 weeks per patient recruited.

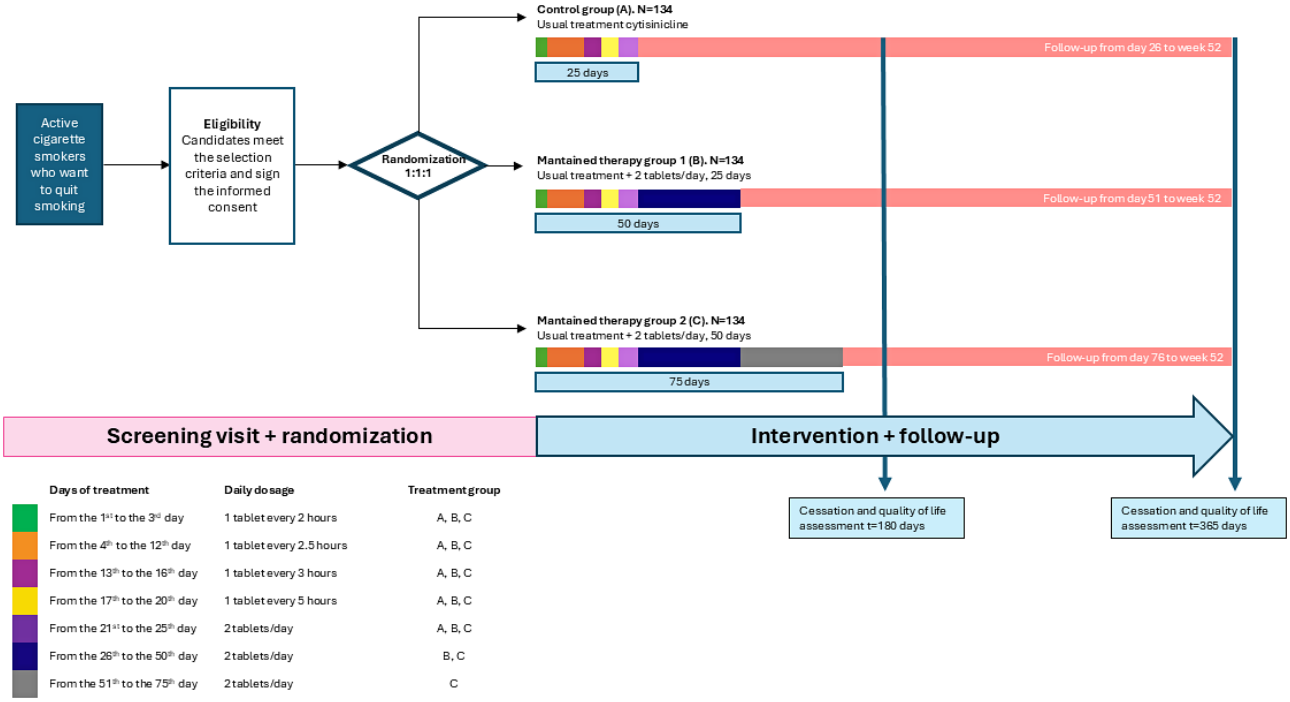

Study Population

This study will be conducted in the Smoking Cessation Units (SCU) of 10 Spanish hospitals with different levels of complexity of care that provide care to active smokers who wish to quit smoking. An SCU is defined as a health care service integrated within a health area that promotes, coordinates, and

XSL•FO
RenderX

implements actions to prevent and treat smoking in that area, in direct collaboration with the area's health care services. The main objectives of SCUs are clinical care, education, research, and management [26]. The SCUs are accredited by the Spanish Society of Pneumology and Thoracic Surgery and by a public agency. Their geographical distribution is representative of the Spanish population (see Table 2).

Table 2. Smoking Cessation Units (SCU) of 10 Spanish hospitals, with different levels of complexity of care, that provide care to active smokers who wish to quit smoking.

Recruitment centers	Estimated number of patients
University Clinical Hospital of Santiago de Compostela	42
Galdakao Hospital, Bilbao	30
San Pedro de Alcántara Hospital, Cáceres	60
San Carlos Clinical Hospital, Madrid	90
October 12 University Hospital, Madrid	40
Clinic Hospital, Barcelona	30
Alcorcón Foundation Hospital, Madrid	30
Virgen de la Macarena University Hospital, Sevilla	40
University Clinical Hospital of Valencia	20
Virgen de la Victoria Hospital, Málaga	20

The sample size of the Cytisinicline Treatment for Smoking Cessation–Long-term Study (CITISILONG) trial (N=402) is based on a minimum number of patients for each of the treatment arms. The sample size calculation is given in the Statistical Analysis section of the Methods section.

Recruitment

A recruitment period of 3 weeks is envisaged, where the persons in charge of the SCU participating in the study will identify patients who come to the practice with the intention to quit smoking and verify and select individuals who meet the abovementioned selection criteria. Written informed consent will be obtained at the screening visit. Once the consent is signed, the patient will have the first trial visit within a maximum of 15 days. Recruitment will be noncompetitive, and each site is allocated a predetermined number of patients (see Figure 1).

Eligibility Criteria

Male and female smokers over the age of 18 years who have smoked more than 10 cigarettes per day during the past year and during the month before the screening visit are eligible to participate in the study. Smokers must also be prepared and motivated to quit smoking (score of 7 or higher on the visual analog scale of motivation). Smokers who have received pharmacological treatment for nicotine dependence (nicotine replacement therapy, varenicline, bupropion, or cytisinicline) within the previous 3 months are excluded from participating in the clinical trial.

In addition, exclusion criteria (Textbox 1) have been designed, taking into account the considerations of the drug label. All patients meeting all inclusion criteria and no exclusion criteria are consecutively invited to participate until the calculated sample size is obtained in each center.

XSL•FO
RenderX

Textbox 1. Exclusion criteria for the study.

- Renal and/or hepatic insufficiency.
- Over 65 years of age (due to limited clinical experience).
- Hypersensitivity to the active substance or to any of the excipients.
- Unstable angina.
- Recent personal history of acute myocardial infarction or recent stroke (less than 8 weeks).
- Clinically relevant arrhythmias or uncontrolled hypertension.
- History of psychiatric illness with and without treatment (panic disorder, psychosis, bipolar disorder, and depression) within the last 12 months.
- History of alcohol or other drug abuse within the last 12 months.
- Use of tobacco products other than cigarettes, e-cigarettes, or marijuana in the past month and did not agree to abstain from using these products during study participation.
- Patients being treated with antituberculosis drugs.
- Patients being treated with drugs metabolized by CYP1A1 (cytochrome *P* 450 family 1, subfamily A, polypeptide 1) and CYP1A2 (cytochrome *P* 450 family 1, subfamily A, polypeptide 2), especially those with a narrow therapeutic margin (eg, theophylline, tacrine, clozapine, and ropinirole).
- Pregnant or breastfeeding smokers.
- Women on systemic hormonal contraceptives.

Patient Enrollment and Randomization

All participants who consent to participate and who meet the eligibility criteria will be included in the trial and randomized. They are recorded in the data collection booklet designed specifically for this purpose.

Randomization will be performed in 2 phases. First, using a simple randomization procedure, participants are assigned to 1 of the 3 arms in a ratio of 1:1:1 (This is already mentioned in Study design).

Then, as noncompetitive recruitment is proposed, the sites are randomized (1 to 8, unrestricted) so that randomization determines where to start distributing the randomized sequence in the first phase. A randomization list will be generated using SAS 9 software (SAS Institute Inc) for Windows. This process will be carried out at the coordinating center, and each participating center will be provided with its own list according to the number of patients allocated. The aim is to achieve a balanced randomization in the treatment groups in the overall sample. There are no blinding techniques, as this is an open-label study. Blind-breaking procedures are therefore not applicable.

Intervention

Cytisinicline is administered as 1.5 mg tablets to all patients for 25, 50, or 75 days depending on arm allocation. The drug is produced and marketed by Adamed Pharma SA ul Karowa 31a 00-324 Warsaw (Poland).

According to the data sheet for cytisinicline, the D-day will be set as the date from which the smoker intends to quit smoking. This date cannot be later than the fifth day after the start of treatment with cytisinicline. This D-day should be decided by the patient with his doctor before starting treatment.

Criteria for Discontinuation of Treatment/Intervention

A patient will be withdrawn from the study when the investigator becomes aware that a participant is experiencing a serious or intolerable adverse event that prevents continuation of the drug study.

The patient may also be excluded from the study at the investigator's request (eg, if the investigator believes that the participant's health could be compromised by continued treatment with the study drug). Noncompliance is not grounds for discontinuing treatment.

If a patient becomes pregnant during the study treatment period, treatment with the study drug must be discontinued. However, the participant must remain in the study to continue with safety and other study assessments.

Discontinuation may also occur at the participant's request if they do not wish to continue drug treatment.

Treatment Adherence Strategies

Adherence to treatment will be determined by checking blister packs at each visit and patient declaration.

Prohibited, Permitted, and Rescue Concomitant Medication

The following are prohibited during the research period (from the signing of informed consent until the end of the follow-up period): antituberculosis drugs, systemic hormonal contraceptives, nicotine-containing products, such as nicotine replacement therapy, and drugs metabolized by the CYP1A2 (cytochrome *P* 450 family 1, subfamily A, polypeptide 2) and CYP1A1 (cytochrome *P* 450 family 1, subfamily A, polypeptide 2) enzymes, especially those with a narrow therapeutic index (eg, theophylline, tacrine, clozapine, and ropinirole). All concomitant medications not previously mentioned are permitted. Rescue medication is not required during the study.

Posttest Care Provisions

Once the study is completed, as this is a low-intervention trial, posttrial care is not required that differs from standard clinical practice. In this case, if the participant quits smoking, he/she

XSL•FO
RenderX

will be discharged from the SCU; if not, he/she will remain a patient in the SCU, and an alternative smoking cessation treatment (change of intervention format or combination of drugs) will be considered.

Outcomes

A full set of outcomes will be analyzed to obtain clinical information on efficacy, safety, and economic impact of maintained cytisinicline therapies versus the standard cytisinicline regimen. Outcomes are classified into efficacy, safety, and economic evaluation objectives.

Specific Outcomes Efficacy

Different outcomes will be used to measure the efficacy of cytisine in each arm of the clinical trial.

The sustained abstinence rate will be determined in each of the three treatment arms at 6 and 12 months, based on carbon monoxide co-oximetry results in ppm. The correlation between the sustained abstinence rate in each of the three arms studied at 6 and 12 months, based on co-oximetry results, and the self-reported abstinence at 6 and 12 months will be calculated.

Additionally, point abstinence will be analyzed every 7 days in each of the 3 arms studied, based on the results obtained in the CO tests.

The percentage increase in abstinence rates will be compared with the increase in the duration of cytisinicline use. The number needed to quit smoking for one patient in each treatment group will also be calculated.

Interval abstinence will also be determined in this clinical trial. The abstinence rate will be analyzed from Day 25 to Day 50 in all 3 study groups and from Day 25 to Day 75 in all 3 study arms. The abstinence rate will also be determined from Day 25 (usual arm), Day 50 (extended arm 1), and Day 75 (extended arm 2) up to 6 and 12 months.

Relapse rates will be analyzed in each treatment arm. The frequency and percentage of relapses will be determined in those who quit smoking from Day 25 (usual arm), Day 50 (extended arm 1), and Day 75 (extended arm 2) up to 6 and 12 months in all 3 arms.

Craving and withdrawal symptoms will be assessed during the clinical trial. The mean Minnesota Nicotine Withdrawal Scale (MNWS) score and the mean craving score will be calculated at the initial in-person visit in all 3 arms, and their change from this visit will be measured on days 25, 50, and 75 in all 3 study arms.

Predictors of continued abstinence at 6 and 12 months of treatment and of relapse (clinical, anthropometric, demographic, and smoking-related variables) will be determined in all study arms.

Quality of life in each group will be calculated using the EQ-5D-5L scale score at baseline, 50 days, 90 days, 6 months, and 1 year.

Specific Outcomes (Safety)

The impact of morbidity and mortality in each treatment arm will be evaluated, as well as the percentage of side effects reported by participants, comparing their prevalence in each group throughout the study (1 year). The percentage of adverse effects at 25, 50, and 75 days will also be determined in participants, comparing their prevalence in each group.

Furthermore, the prevalence and severity of side effects for each regimen will be compared throughout the study (1 year).

Finally, the results of cardiac, hepatic, and renal variables will be compared between each group as indicators of treatment safety in each group at 25, 50, and 75 days of treatment and throughout the entire study period (1 year).

Specific Outcomes (Economic Evaluation)

The economic analysis of 2 prolonged cytisinicline regimens vs a standard regimen will be performed by calculating their cost-effectiveness and cost-utility ratios. The effect of treatment adherence on the cost-effectiveness and cost-utility ratios of the considered therapeutic alternatives will also be evaluated.

The effect of the different predictors of sustained abstinence at 6 and 12 months (clinical, anthropometric, and demographic variables) will be evaluated in relation to cost-effectiveness and cost-utility in the three arms of the clinical trial.

The future effects (15, 30, and 50 years after smoking cessation) will be simulated in each of the treatment arms.

Finally, a budget impact model will be constructed to analyze the potential savings for the National Health System for each of the 3 therapeutic options proposed in this clinical trial.

Schedule of Participants

Table 3 and Figure 1 show the screening visit, follow-up visits (face-to-face and telephone), tests and quizzes to be carried out at each of the visits, and the interventions.

XSL•FO
RenderX

Table 3. Visiting calendar. The screening visit, follow-up visits (face-to-face and telephone), tests and quizzes to be carried out at each of the visits, as well as the intervention.

Number, day, and type of visit	Visit 1 (day 1; in person)	Visit 2 (day 7; in person)	Visit 3 (day 14; telephone)	Visit 4 (day 21; in person)	Visit 5 (day 25; telephone)	Visit 6 (day 50; telephone)	Visit 7 (day 75; telephone)	Visit 8 (day 90; in person)	Visit 9 (day 120; in person)	Visit 10 (day 150; in person)	Visit 11 (week 24; in person)	Visit 12 (week 36; telephone)	Visit 13 (week 48; in person)	Visit 14 (week 52; in person)
Variable registration														
Sociodemographic	✓													
Anthropometric	✓													
Diagnostic	✓													
UISPM[a] Test	✓													
Motivation	✓													
Reward	✓													
Spirometry	✓													
Advice	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Efficacy evaluation														
Co-oximetry	✓	✓		✓				✓	✓	✓	✓		✓	✓
Statement	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Security assessment														
Adverse events		✓	✓		✓	✓	✓							
Quality of life assessment														
EQ-5D	✓				✓	✓	✓		✓	✓	✓			✓

[a]UISPM: Test of the Specialized Smoking Unit of the Community of Madrid.

Data Collection Variables

With regard to patient information, a database will be designed specifically for this proposal. The medical and demographic data are as follows: age (years), sex (male/female), education level (no formal education, secondary education, or higher education) and social status (low, medium, or high). We establish cut-off points for the patient's socioeconomic level according to the OECD (Organisation for Economic Co-operation and Development) based on the approximate monthly income level (less than €1000 (€1=US $1.16–$1.17) per month = low socioeconomic level, from €1000 to €3000 per month = medium socioeconomic level, and more than €3000 per month = high socioeconomic level). Questionnaires carried out in the SCUs are on current consumption, accumulated consumption, motivation, the quitting phase, and nicotine dependence. Smoking consumption characteristics are current number of cigarettes per day (categorized as ≥20 cigarettes per day), packs-year, previous quit attempts (previous attempts, previous attempts in the last year, previous attempts with drug treatment, and previous attempts with varenicline), and graded degree of smoking (mild, moderate, severe, or very severe) [27,28].

The different characteristics and clinical profiles of each smoker will be determined by the Test of the Specialized Smoking Unit of the Community of Madrid. It consists of 7 domains (stimulation, automaticity, sedation, psychological, gestural and social dependence, and automaticity). It allows for a more detailed identification of the smoker's specific characteristics to optimize psychological and pharmacological treatment [27].

Nicotine dependence will be determined using the Fagerström test. It provides information about the patient's physical dependence on nicotine. It is a questionnaire of 6 questions. A score of 7 or more indicates high nicotine dependence [29].

The reward associated with tobacco use will be measured in the clinical trial. We will use the reward test for this purpose. It indicates whether the smoker smokes for pleasure (positive reward), smokes to control withdrawal symptoms (negative reward), or both (mixed reward) [30].

The analysis of withdrawal symptoms in smokers will be performed using the MNWS. It allows for the detection of the intensity of withdrawal symptoms and their variation throughout the smoking cessation follow-up process [30].

Other qualities of the smoker, such as the degree of motivation and self-efficacy, will be evaluated by the Analogue Visual Scale. This scale numerically identifies a smoker's level of motivation and self-efficacy to quit smoking. A score above 7 indicates high motivation and self-efficacy to quit smoking [27].

XSL•FO
RenderX

The EQ-5D scale will be used in this clinical trial to assess the quality of life of smokers and its changes during the smoking cessation process. This scale measures the health status of smokers using 5 dimensions, including mobility, self-care, activities of daily living, pain and discomfort, and anxiety and depression. Participants score from 0 to 100 in each domain [31].

The urge to smoke in this study will be analyzed using the Craving Rate Score. This scale assesses anxiety or intense desire to smoke at baseline and during follow-up. It is scored from 0 to 4 [32].

The additional tests that will be performed on smokers participating in this trial include electrocardiogram, carbon monoxide levels, complete blood count, and a biochemical profile that includes liver enzymes, electrolytes, and renal function tests. Co-oximetry allows for the objective identification of smoking status, as well as confirmation of abstinence at each participant visit. A participant is identified as a nonsmoker when their CO levels measured by co-oximetry are below 10 ppm (Textbox 2).

Textbox 2. Efficacy and safety variables of cytisinicline.

To analyze the efficacy of cytisinicline we used the following variables:

1. Continuous abstinence rate at six months of treatment
2. Efficacy of the maintained cytisinicline regimens (50 and 75 days) compared to the usual 25-day regimen calculated every 7 days
3. Intervention interval effectiveness
    - Abstinence rate between days 25 and 50
    - Abstinence rate between days 25 and 75
    - Abstinence rate between days 25 and 6 months calculated every 7 days
4. Analysis of withdrawal syndrome
    - MNWS score at each visit
    - Variation in MNWS score between visits
    - Mean MNWS score at each visit
5. Craving Analysis
    - SRC score at each visit
    - Change in SRC score at each visit
    - Average SRC score at each visit

To analyze the safety of cytisinicline, we used the following variables:

1. Number and percentage of self-reported adverse events at the baseline visit
2. Number and percentage of self-reported adverse events on days 7, 15, 25, 50, and 75
3. Number and percentage of mild, moderate, or severe adverse events on days 7, 15, 25, 50, and 75
4. Number and percentage of serious adverse events on days 7, 15, 25, 50, and 75
5. Number and percentage of the following adverse events: nausea, vomiting, constipation, abdominal distension, abdominal pain, epigastric pain, nightmares, insomnia, parasomnias, headache, and dizziness
6. Percentage of patients with electrocardiographic abnormalities at the initial visit
7. Percentage of patients with electrocardiographic alterations at 25, 50, and 75 days
8. Percentage of patients with electrocardiographic alterations at 25, 50, and 75 days
9. Percentage of patients with abnormal liver enzyme levels at the baseline visit
10. Percentage of patients with abnormal liver enzyme levels at 25, 50, and 75 days
11. Percentage of patients with abnormal electrolyte levels at the baseline visit
12. Percentage of patients with abnormal electrolyte levels at 25, 50, and 75 days
13. Percentage of patients with abnormal renal function at the baseline visit
14. Percentage of patients with abnormal renal function at 25, 50, and 75 days

XSL•FO
RenderX

Trial Monitoring

As this is a low-intervention trial, the monitoring and quality control processes of the generated database will be carried out by the principal investigators of each site or by a collaborating investigator of the team to whom this task is delegated.

Pharmacovigilance

As this is a low-intervention study, the expected adverse events are those reflected in the drug label. The definition of adverse events, the classification of their severity, and the actions to be taken by the principal investigator and sponsor in the event of an adverse event are included in Multimedia Appendix 1. In all cases, the sponsor should promptly communicate to the investigators any important information that could adversely affect the safety of the participants or the conduct of the trial. The communication of such information should be concise and practical. The communication shall be in accordance with the criteria and procedure specified in the European Commission's guidelines. The sponsor shall report to the Spanish Agency for Medicinal Products and Health Products all suspected serious and unexpected adverse reactions associated with the investigational medicinal products of which he/she has become aware. An adverse event is an undesirable experience that affects a patient during a clinical trial or study, whether or not it is considered to be related to the investigational product.

A serious adverse event is one that is life-threatening, causes persistent disability, or requires prolonged hospitalization of the patient. Apart from these, congenital anomalies and malignant neoplasms shall always be considered serious.

A nonserious adverse event is any adverse event that does not meet any of the above criteria.

An unexpected adverse event is an experience that is not described in the investigator's manual (by nature, severity, or frequency occurred in the clinical trial).

Statistical Analysis

Sample Size Determination

Based on clinical experience, it is estimated that the efficacy of cytisinicline prescribed in accordance with the data sheet is 30% (P^1=.30), and that this efficacy will be greater if the maintenance regimen is extended over time. In this sense, it is hypothesized that the intermediate maintenance regimen could achieve an efficacy of 45% (P^2=.40) and with the longest regimen, 50% sustained abstinence could be achieved (P^3=.50). To obtain 80% power (pow=0.80) to detect differences at a 5% significance level (α=.05), 110 patients would need to be included in each arm. A loss rate of 21% is assumed, giving a final sample size of 402 patients (N=402).

Population of Analysis

Efficacy and safety studies shall be conducted in the entire clinically evaluable population.

Analysis of Demographic and Baseline Characteristics

Demographic and baseline characteristics are summarized for each study group using descriptive statistics. For continuous variables, data on the number of observations, mean, SD, minimum and maximum, or median and IQR will be presented, depending on the fit of the variable to a normal distribution. Discrete variables will be presented with frequency and percentage. Appropriate inference tests chosen according to the nature of the variables were performed to compare these characteristics between the groups under study.

Efficacy Analysis

To determine the main objective of the work, the efficacy variable will be constructed as a dichotomous variable, resulting from an algorithm that will be constructed for this purpose once all the data from the study have been obtained. To study the causality between efficacy and treatment prolongation, a multivariate logistic regression will be constructed. To determine candidate variables, candidate variables will first be tested in a univariate regression. Those variables whose statistical significance has a *P* value of <.20 will become part of a multivariate model. The model that best explains causality will be determined using a measure of relative statistical model quality, the Akaike information criterion. The risk measure to be calculated is the relative risk with its 95% CI. Interim follow-up assessments will be analyzed using a survival analysis to determine the duration of abstinence after the smoking cessation date.

Safety Analysis

The frequency and percentage of occurrence of adverse reactions shall be determined and categorized according to type/intensity. Calculate whether there are differences in each group compared to what is expected according to the product label.

To examine differences in the data between groups defined by the presence/absence of adverse events, the chi-square test will be applied for qualitative variables and the *t* test or the Mann-Whitney *U* test for quantitative variables. Finally, the causal relationship between safety, defined by the absence of adverse events, and treatment prolongation will be verified using multivariate logistic regression. To determine candidate variables, those with a statistical significance of *P*<.20 will be tested and included in a multivariate model.

All analyses will be performed with the statistical package SPSS (version 19.0; IBM Corp), and statistically significant values will be considered those whose *P* value is <.05.

Economic Impact Analysis

For the economic impact analysis, direct costs (medical and nonmedical) and indirect costs will be included.

Health outcomes will be quantified in terms of quality-adjusted life years in our economic evaluation. The health-related quality of life index (utilities) for these outcomes has been estimated using the EQ-5D-3L data for smokers and ex-smokers from the 2017 National Health Survey.

Effectiveness results will be quantified in terms of years of life lost prematurely compared to the average life expectancy for men and women.

The cost-utility analysis of administering cytisinicline at the standard regimen versus 2 sustained cytisinicline therapies will be performed using the incremental cost-utility ratio. This will

XSL•FO
RenderX

be calculated as the incremental cost in Euro per additional unit of quality of life gained by one alternative compared to the other. Cost differences between the 2 alternatives/differences in quality-adjusted life years between the 2 alternatives

The cost-effectiveness analysis of administering cytisinicline at the standard regimen versus 2 cytisinicline-based therapies will be performed using the incremental cost-effectiveness ratio. This will be calculated as the incremental cost in Euro per life-year gained for one alternative compared to the other.

Differences in costs and life-years lost between the 2 alternatives will be estimated. All results will be presented for different time horizons, including the short-time (5 years), medium-time (10 years), and long-time (20 years).

To assess the uncertainty of the results, univariate and multivariate deterministic sensitivity analyses will be conducted under the following assumptions: the first scenario, univariate deterministic sensitivity analysis, will consider reducing only the efficacy rate of cytisinicline by 50% while maintaining the discount rate at 3% as in the baseline analysis. The second scenario, multivariate deterministic sensitivity analysis, will consider reducing the efficacy rate by 50% but increasing the discount rate to 6%.

Ethical Considerations

The trial will be conducted in accordance with the requirements documented in the latest version of the Declaration of Helsinki (Helsinki, 2024) and with the Standards of Good Clinical Practice. The trial is authorized for conduct by the Galician (Spain) Committee on Ethics in Medicines Research with code 2025/043 and by the Spanish Agency for Medicines and Medical Devices. It also has the agreement of the participating centers. This trial has been registered in the European Clinical Trials Register under EUDRACT code 2024-518936-36-00.

The researchers at each center will provide each participant with a patient information sheet, approved by the clinical research ethics committee, informing them about relevant study data and requesting access to their clinical and personal data recorded in their medical history and disclosed during visits. Permission is also requested for tests, examinations, and questionnaires, and the patient is informed that these actions are part of their daily clinical practice. In addition, the researcher will verbally inform the patient about the trial and answer their questions to ensure they understand the implications of their participation. The patient will be informed of their right to withdraw consent at any time and their right to have their data deleted. The patient will be informed of the possibility of being withdrawn from the study for safety reasons (an adverse event caused by the study medication) or for failure to comply with established procedures.

No patient may be included who has not given written informed consent to participate in the trial. Secondary analyses using existing data are included in the primary consent and do not require additional consent.

The research team and the sponsor of this study expressly undertake not to disclose the identity of the patients treated and to respect the confidentiality rules regarding the data and information to which they have access by participating in the trial.

In order to guarantee the confidentiality of the trial data, only the sponsor and members of the research team, the Research Ethics Committee for Medicinal Products, and the relevant health authorities will have access to them.

The research team undertakes to collect, record, and report the data accurately and correctly. The sponsor and the research team guarantee the confidentiality of the participants' data and will ensure compliance at all times with the provisions of EU (European Union) Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation), and Organic Law 3/2018, of 5 December, on the Protection of Personal Data and Guarantee of Digital Rights.

Patients will be identified by a number in the database that guarantees complete data confidentiality, particularly the participants' identity. Data will be sent to the study sponsor in pseudonymous form so that it is not possible to identify the patient in any way. The trial guarantees that patients cannot be identified in any of the images or supplementary material in the manuscripts obtained after this clinical trial. All information obtained as a result of this study will be considered confidential. The investigator may only report on the development and results of the study to the sponsor, the ethics committee, and the relevant health authorities.

The participant will not receive any remuneration for participating in this trial. Expenses incurred during visits related to the study will also not be covered.

Results

Trial recruitment and patient registration will conclude in November 2026. Follow-up for all participants will continue until December 2027. At the time of this writing, we have 12 preselected patients at the sponsoring center (Smoking Unit, University Hospital Complex of Santiago de Compostela, Spain) who meet the exclusion and inclusion criteria of the study and who will begin the initial visit in the coming days.

Discussion

Principal Findings

Extended treatment regimens with cytisinicline of 50 and 75 days may be more effective than the 25-day treatment regimen currently used in clinical practice. Furthermore, the percentage of adverse effects and the economic cost should not be higher with extended treatment regimens. These expected results will confirm the current scientific evidence on this drug. For example, a multicenter observational study conducted in Spain with more than 300 patients treated with cytisinicline for 25 days shows a relapse rate of 20% in the 2 months following the end of treatment. For every point increase in craving at the end of treatment, the success rate of smoking cessation decreases by 13% at 6 months of follow-up [25]. In other words, the intensity of craving at 4 weeks is inversely correlated with the

XSL•FO
RenderX

drug's efficacy. Increasing the number of days of drug administration could correct this intense desire to smoke.

Two placebo-controlled clinical trials conducted in the United States compared a cytisinicline regimen at a dose of 3 mg every 8 hours for 6 weeks versus 12 weeks. These trials showed an increase in the number of smokers who quit smoking when the treatment regimen was extended beyond 6 weeks. Furthermore, the increased number of treatment days allowed a group of smokers who were gradually reducing their consumption to quit smoking, thus increasing the percentage of smokers who quit. In other words, there will be a group of smokers who will require a longer time to quit smoking [21,22]. No greater number of adverse events were observed with these prolonged regimens. In Europe, a clinical trial in a lung cancer screening program demonstrated that a prolonged cytisinicline regimen with doses different from those specified in the product information—an 84-day regimen—was more effective than the 40-day regimen, without a greater number of side effects [23].

As strengths, it is the first study to compare 3 different cytisinicline regimens: a standard 25-day regimen used in routine clinical practice approved in the product information versus 2 prolonged treatments of double and triple duration without modifying the dose or frequency of the same, unlike other studies that evaluate a higher dose, a different frequency of administration, or compare cytisinicline with other drugs [21-23].

This is a low-intervention study due to several characteristics: (1) studies of extended regimens exist in the European Union [23]. These analyze the efficacy and safety of extended cytisinicline treatments of 40 and 84 days. We also have experience with extended doses of 1.5 mg every 12 hours for 3 months [24]; (2) modifications to drug administration are minimal. One of the study arms corresponds to the regimen used in daily clinical practice; and (3) patient selection was performed according to the drug's prescribing information. All of this ensures its feasibility.

Unlike other studies that propose a substantial change in the dose and frequency of the molecule, this one uses extended treatment regimens of an additional 25 and 50 days, with minimal modification to the commercially available drug formulation. Thus, in Spain, each pack of cytisinicline contains 100 tablets, so one-third of patients will be randomized to take half the tablets contained in the additional pack (2 tablets a day for an additional 25 days) and another third to take all of them in that pack (2 tablets a day for an additional 50 days). Furthermore, these regimens, with their lower administration frequency, could be more adherent than regimens of 3 times a day for 12 weeks.

At the time of this writing in Europe, the same dose and duration of cytisine are used for all smokers. However, smokers represent a heterogeneous patient population with differences in dependence, emotionality, motivation, and comorbidities. This study, using a multivariate model, allows the identification of the group of patients who would require extended treatment with cytisinicline, personalizing its use in routine practice.

It is a multicenter study conducted in different geographical areas of Spain, with a representative sample. Moreover, it is carried out by a research group with experience in the diagnosis and treatment of smoking cessation. Patients treated in 10 specialized smoking cessation units accredited by scientific societies are included. Furthermore, this is an independent trial, free from the financial interests of the pharmaceutical industry, which enhances the rigor of the results. Funding for disseminating the results will be obtained through grants or public funds.

This trial incorporates the calculation of the economic impact of extended doses, determining the savings for the National Health System that these regimens could generate by increasing the number of ex-smokers.

The results of the study will be published in indexed scientific journals and presented at national and international conferences.

Limitations

This study is not placebo-controlled; therefore, the effect of the intervention (extended cytisinicline treatment) cannot be accurately determined. The results of the intervention could be attributed not only to the drug but also to external factors. However, the use of a placebo could compromise the low level of intervention in this trial, according to regulatory authorities in Spain and the European Union. Furthermore, the synthesis of the placebo would hinder the feasibility of this study. Also, not including a placebo in the 25-day and 50-day treatment arms could provide results more aligned with routine clinical practice.

On the other hand, the efficacy and safety of different durations of cytisinicline treatment are not compared to other drugs for the treatment of nicotine dependence, such as varenicline or nicotine replacement therapy. This is because this is not the ultimate goal of this study. However, such a comparative analysis will be carried out in subsequent studies. The selected patient population consists of smokers attending SCUs with high nicotine dependence and associated comorbidities and does not represent the general smoking population. This selection bias may influence the greater efficacy of extended cytisinicline treatment. However, these are units of varying complexity, and smoking cessation can also be addressed in individuals with less nicotine dependence.

Because this is a low-intervention clinical trial, smokers older than 65 years and those with serious psychiatric or cardiovascular comorbidities are not included. Therefore, we cannot generate scientific evidence on extended treatment in these vulnerable populations.

Abstinence outcomes are objectively determined during in-person visits using co-oximetry. However, other more precise methods, such as cotinine determination in body fluids, are not used. Nevertheless, this test is accepted as a powerful marker for verifying abstinence, establishing a carbon monoxide (CO in ppm) cutoff point of less than 10 ppm [27]. Although different monitors may vary in their sensitivity to CO, combining this measurement with patient self-reporting improves the predictive capacity of co-oximetry.

XSL•FO
RenderX

Furthermore, nonpharmacological interventions may differ between centers, which could affect the results. However, a standardized model of psychological counseling is established across the centers, based on behavior modification, avoidance, substitution, and cognitive techniques.

Adherence was determined by patient self-report and the number of tablets in the blister pack, not by measuring the drug in bodily fluids. However, the high frequency of follow-up visits in the study allows for better identification of adherence to cytisinicline.

Conclusions

Cytisinicline, when used in a standard regimen, triples the likelihood of smoking cessation, as it is a well-tolerated drug with high adherence. However, the percentage of smokers who quit is less than 50%. Therefore, it is essential to design new administration methods and therapeutic regimens.

An optimization strategy for this drug, such as increasing its duration, enhances its benefits without compromising patient safety, representing a new opportunity for smokers to quit.

Furthermore, this study identifies the population of smokers susceptible to a prolonged treatment regimen with cytisinicline. All of this will lead to a reduction in the number of smokers, decreasing, in the short to medium term, morbidity, mortality, and the direct and indirect costs associated with smoking.

Funding

It is independent research, receiving no funding from the pharmaceutical industry or other public or private companies.

Efficacy, safety and economic impact of cytisinicline maintenance therapy in patients who are candidates for smoking cessation. Protocol for a phase IV, multicenter, randomized, open-label, controlled, parallel trial (the CITISILONG trial).

Data Availability

The datasets generated or analyzed during this study are not publicly available due to their nature as protected health information. Access to this data is restricted by the participating institutions.

Authors' Contributions

All authors have contributed to the preparation of this manuscript.
Conceptualization: CRC, AE-G, CAJR, LVC
Data curation: CRC, AE-G, CAJR, LVC
Formal analysis: CRC, AE-G, CAJR, JIG-O
Investigation: CRC, AE-G, CAJR, JIG-O, JAR-M, MICF, AR-P, JSC-M, MIG-U, AV-M, EC-C, AF-G, LVC, JS-T
Methodology: AE-G
Project administration: CRC, AE-G, CAJR
Resources: CRC
Software: AE-G
Supervision: CAJR, JIG-O, JAR-M
Validation: CRC, AE-G, CAJR, JIG-O, JAR-M, MICF, AR-P, JSC-M, MIG-U, AV-M, EC-C, AF-G, LVC, JS-T
Visualization: CRC, AE-G, CAJR, JIG-O, JAR-M, MICF, AR-P, JSC-M, MIG-U, AV-M, EC-C, AF-G, LVC, JS-T
Writing—original draft: PSC (lead), MSH, WLO, BKS (support): CRC, AE-G, CAJR, JIG-O
Writing—review & editing: all authors: CRC, CAJR

Conflicts of Interest

AR-P has received fees for lectures, scientific advice, participation in clinical studies, and manuscript preparation from Aflofarm, AstraZeneca, Chiesi, FAES Farma, Gebro, Neuraxpharm, Pfizer, and Zambon. JS-T has received fees from Aflofarm, Boehringer Ingelheim, Neuroxpharma, and Adamed for participation on advisory boards and at conferences, and has received grants from Boehringer Ingelheim. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Definition and classification of adverse events. Actions to be taken by the principal investigator and the sponsor in the event of an adverse event.
[DOCX File , 16 KB - resprot_v15i1e76815_app1.docx]

Multimedia Appendix 2
SPIRIT 2025 checklist of items to address in protocol for a phase IV, multicentre, randomised, open-label, controlled, parallel, multicentre clinical trial that evaluates efficacy, safety and economic impact of citisinicline maintenance therapy in patients who are candidates for smoking cessation CITISILONG trial.

XSL•FO
RenderX

[DOCX File , 34 KB - resprot_v15i1e76815_app2.docx]

References

1. Observatorio Español de las Drogas y las Adicciones. Report 2021. Alcohol, tobacco and illegal drugs in Spain. Madrid: Ministry of Health, Government Delegation for the National Plan on Drugs; 2021:243.
2. Bartusik D, Aebisher D, Tutka P. A review of the organic synthesis and medicinal applications of the natural product cytisine. MOCR 2016;1(1). [doi: 10.22606/mocr.2016.11002]
3. Wang X, Yang J, Huang P, Wang D, Zhang Z, Zhou Z, et al. Cytisine: state of the art in pharmacological activities and pharmacokinetics. Biomed Pharmacother 2024;171:116210 [FREE Full text] [doi: 10.1016/j.biopha.2024.116210] [Medline: 38271893]
4. Tutka P, Zatoński W. Cytisine for the treatment of nicotine addiction: from a molecule to therapeutic efficacy. Pharmacol Rep 2006;58(6):777-798 [FREE Full text] [Medline: 17220536]
5. Rollema H, Shrikhande A, Ward KM, Tingley FD, Coe JW, O'Neill BT, et al. Pre-clinical properties of the alpha4beta2 nicotinic acetylcholine receptor partial agonists varenicline, cytisine and dianicline translate to clinical efficacy for nicotine dependence. Br J Pharmacol 2010;160(2):334-345 [FREE Full text] [doi: 10.1111/j.1476-5381.2010.00682.x] [Medline: 20331614]
6. Rábade-Castedo C, de Granda-Orive JI, Riesco-Miranda JA, De Higes-Martínez E, Ramos-Pinedo, Cabrera-César E, et al. Clinical practice guideline of Spanish society of pneumology and thoracic Surgery (SEPAR) on pharmacological treatment of tobacco dependence 2023. Arch Bronconeumol 2023;59(10):651-661 [FREE Full text] [doi: 10.1016/j.arbres.2023.07.024] [Medline: 37567792]
7. Cytisinicline 1. 5 mg film-coated tablets EFG. Centro de Información online de Medicamentos (CIMA). URL: https://cima.aemps.es [accessed 2025-12-30]
8. Etter J, Lukas RJ, Benowitz NL, West R, Dresler CM. Cytisine for smoking cessation: a research agenda. Drug Alcohol Depend 2008;92(1-3):3-8. [doi: 10.1016/j.drugalcdep.2007.06.017] [Medline: 17825502]
9. Vinnikov D, Brimkulov N, Burjubaeva A. A double-blind, randomised, placebo-controlled trial of cytisine for smoking cessation in medium-dependent workers. J Smok Cessat 2012;3(1):57-62. [doi: 10.1375/jsc.3.1.57]
10. West R, Zatonski W, Cedzynska M, Lewandowska D, Pazik J, Aveyard P, et al. Placebo-controlled trial of cytisine for smoking cessation. N Engl J Med 2011;365(13):1193-1200. [doi: 10.1056/NEJMoa1102035] [Medline: 21991893]
11. Hajek P, McRobbie H, Myers K. Efficacy of cytisine in helping smokers quit: systematic review and meta-analysis. Thorax 2013;68(11):1037-1042. [doi: 10.1136/thoraxjnl-2012-203035] [Medline: 23404838]
12. Tutka P, Vinnikov D, Courtney RJ, Benowitz NL. Cytisine for nicotine addiction treatment: a review of pharmacology, therapeutics and an update of clinical trial evidence for smoking cessation. Addiction 2019;114(11):1951-1969. [doi: 10.1111/add.14721] [Medline: 31240783]
13. Puljević C, Stjepanović D, Meciar I, Kang H, Chan G, Morphett K, et al. Systematic review and meta-analyses of cytisine to support tobacco cessation. Addiction 2024;119(10):1713-1725. [doi: 10.1111/add.16592] [Medline: 38965792]
14. De Santi O, Orellana M, Di Niro CA, Greco V. Evaluation of the effectiveness of cytisine for the treatment of smoking cessation: a systematic review and meta-analysis. Addiction 2024;119(4):649-663. [doi: 10.1111/add.16399] [Medline: 38161271]
15. Ofori S, Lu C, Olasupo OO, Dennis BB, Fairbairn N, Devereaux PJ, et al. Cytisine for smoking cessation: a systematic review and meta-analysis. Drug Alcohol Depend 2023;251:110936. [doi: 10.1016/j.drugalcdep.2023.110936] [Medline: 37678096]
16. Poland. In: National Accounts Statistics: Main Aggregates and Detailed Tables 2016. New York, NY: United Nations iLibrary; 2017:491-531.
17. Livingstone-Banks J, Fanshawe TR, Thomas KH, Theodoulou A, Hajizadeh A, Hartman L, et al. Nicotine receptor partial agonists for smoking cessation. Cochrane Database Syst Rev 2023;5(5):CD006103 [FREE Full text] [doi: 10.1002/14651858.CD006103.pub8] [Medline: 37142273]
18. Walker N, Howe C, Glover M, McRobbie H, Barnes J, Nosa V, et al. Cytisine versus nicotine for smoking cessation. N Engl J Med 2014;371(25):2353-2362 [FREE Full text] [doi: 10.1056/NEJMoa1407764] [Medline: 25517706]
19. Courtney RJ, McRobbie H, Tutka P, Weaver NA, Petrie D, Mendelsohn CP, et al. Effect of cytisine vs varenicline on smoking cessation: a randomized clinical trial. JAMA 2021;326(1):56-64 [FREE Full text] [doi: 10.1001/jama.2021.7621] [Medline: 34228066]
20. Jiménez-Ruiz CA, Riesco Miranda JA, Martínez Picó A, de Simón Gutiérrez R, Signes-Costa Miñana J, Díaz Maroto Muñoz JL, et al. The DESTINA study: an observational cross-sectional study to evaluate patient satisfaction and tolerability of cytisine for smoking cessation in Spain. Arch Bronconeumol 2023;59(4):270-272. [doi: 10.1016/j.arbres.2022.11.015] [Medline: 36549936]
21. Rigotti NA, Benowitz NL, Prochaska J, Leischow S, Nides M, Blumenstein B, et al. Cytisinicline for smoking cessation: a randomized clinical trial. JAMA 2023;330(2):152-160 [FREE Full text] [doi: 10.1001/jama.2023.10042] [Medline: 37432430]

XSL•FO
RenderX

22. Rigotti NA, Benowitz NL, Prochaska JJ, Rubinstein M, Clarke A, Blumenstein B, et al. Cytisinicline for smoking cessation: The ORCA phase 3 replication randomized clinical trial. JAMA Intern Med 2025;185(6):648-655. [doi: 10.1001/jamainternmed.2025.0628] [Medline: 40257755]
23. Pastorino U, Ladisa V, Trussardo S, Sabia F, Rolli L, Valsecchi C, et al. Cytisine therapy improved smoking cessation in the randomized screening and multiple intervention on lung epidemics lung cancer screening trial. J Thorac Oncol 2022;17(11):1276-1286 [FREE Full text] [doi: 10.1016/j.jtho.2022.07.007] [Medline: 35908731]
24. Walker N, Smith B, Barnes J, Verbiest M, Parag V, Pokhrel S, et al. Cytisine versus varenicline for smoking cessation in New Zealand indigenous Māori: a randomized controlled trial. Addiction 2021;116(10):2847-2858 [FREE Full text] [doi: 10.1111/add.15489] [Medline: 33761149]
25. Ramos-Pinedo A, de Granda-Orive JI, Cristóbal-Fernández MI, Rábade-Castedo C, Pérez-Fernández E, Vaquero-Lozano P, et al. Analysis of Effectiveness of Clinical Psychological Care Protocol for Smoking Cessation With Cytisinicline. Arch Bronconeumol 2025. [doi: 10.1016/j.arbres.2025.08.002] [Medline: 40935698]
26. Rábade-Castedo C, Estrada Riolobos G, Cebrián S, Díaz-Maroto JL, Gaztelurrutia Lavesa L, González-Béjar M, et al. [Expert consensus for the organization of smoking management in Spain]. Semergen 2025 Oct;51(7):102484. [doi: 10.1016/j.semerg.2025.102484] [Medline: 40138881]
27. Ramos PA, Prieto GE. Clinical study of the smoker. In: Jiménez-Ruiz CA, Fagerström KO, editors. Treatise on Tobacco Addiction (Tratado de tabaquismo). Madrid: Grupo Aula Médica, SL; 2011:295-304.
28. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983;51(3):390-405. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
29. Fagerström K. Determinants of tobacco use and renaming the FTND to the Fagerstrom Test for cigarette dependence. Nicotine Tob Res 2012;14(1):75-78. [doi: 10.1093/ntr/ntr137] [Medline: 22025545]
30. Pomerleau OF, Fagerström KO, Marks JL, Tate JC, Pomerleau CS. Development and validation of a self-rating scale for positive- and negative-reinforcement smoking: the Michigan Nicotine Reinforcement Questionnaire. Nicotine Tob Res 2003;5(5):711-718. [doi: 10.1080/1462220031000158627] [Medline: 14577987]
31. Herdman M, Badia X, Berra S. [EuroQol-5D: a simple alternative for measuring health-related quality of life in primary care]. Aten Primaria 2001;28(6):425-430 [FREE Full text] [doi: 10.1016/s0212-6567(01)70406-4] [Medline: 11602124]
32. Heishman SJ, Singleton EG, Pickworth WB. Reliability and validity of a short form of the tobacco craving questionnaire. Nicotine Tob Res 2008;10(4):643-651 [FREE Full text] [doi: 10.1080/14622200801908174] [Medline: 18418787]

Abbreviations

CITISILONG: Cytisine Treatment for Smoking Cessation–Long-term study
CYP1A1: cytochrome *P* 450 family 1, subfamily A, polypeptide 1
CYP1A2: cytochrome *P* 450 family 1, subfamily A, polypeptide 2
EU: European Union
MNWS: Minnesota Nicotine Withdrawal Scale
OECD: Organisation for Economic Co-operation and Development
OR: odds ratio
SCU: Smoking Cessation Unit
SMILE: smoking cessation with cytisinicline
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials Statement

Edited by A Schwartz; submitted 04.May.2025; peer-reviewed by H Tateno, K Perkins; comments to author 29.Sep.2025; accepted 18.Dec.2025; published 23.Jan.2026.

Please cite as:
Rabade Castedo C, Estany-Gestal A, Jiménez Ruiz CA, de Granda-Orive JI, Riesco-Miranda JA, Cristóbal Fernández MI, Ramos-Pinedo A, Signes-Costa Miñana J, Gorordo-Unzueta MI, Valido-Morales A, Sellarés -Torres J, Cabrera-César E, Frino-García A, Valdés Cuadrado L
Efficacy, Safety, and Economic Impact of Cytisinicline Maintenance Therapy in Patients Who Are Candidates for Smoking Cessation: Protocol for a Phase IV, Multicenter, Randomized, Open-Label, Controlled, Parallel Clinical Trial (CITISILONG Trial)
JMIR Res Protoc 2026;15:e76815
URL: https://www.researchprotocols.org/2026/1/e76815
doi:10.2196/76815
PMID:

©Carlos Rabade Castedo, Ana Estany-Gestal, Carlos A Jiménez Ruiz, José Ignacio de Granda-Orive, Juan Antonio Riesco-Miranda, María Isabel Cristóbal Fernández, Angela Ramos-Pinedo, Jaime Signes-Costa Miñana, María Inmaculada Gorordo-Unzueta,

XSL•FO
RenderX

Agustin Valido-Morales, Jacobo Sellarés -Torres, Eva Cabrera-César, Alejandro Frino-García, Luis Valdés Cuadrado. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effectiveness and Safety of Kangjia Decoction Granules for the Treatment of Hashimoto Thyroiditis: Protocol for a Randomized, Double-Blinded, Placebo-Controlled, Multicenter Clinical Trial

Duanchun Zhang[1,2*], MM; Dan Zhang[1*], MM; Xiaoxiao Qu[1,3*], MD; Huihong Cao[2], MM; Yanming He[1], MD

[1]Department of Endocrinology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[2]Department of Endocrinology, Shanghai Minhang of Integrated Traditional Chinese and Western Medicine, Shanghai, China

[3]Department of Critical Care Medicine, Shuguang Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Huihong Cao, MM
Department of Endocrinology
Shanghai Minhang of Integrated Traditional Chinese and Western Medicine
Number 155 Jianchuan Road, Minhang District
Shanghai, 200437
China
Phone: 86 18017595587
Email: rosyfry@126.com

Abstract

Background: Hashimoto thyroiditis (HT) is a chronic inflammation of the thyroid gland mediated by autoimmune disorders, often leading to hypothyroidism and a significant reduction in a patient's quality of life. At the time of this writing, there is a lack of effective clinical treatments for early-stage HT. Kangjia decoction granules (KDGs) were developed based on clinical experience and results analysis, showing promising outcomes in improving antibody levels and quality of life in patients with HT. However, there is a lack of further evaluation of the efficacy and safety of KDGs.

Objective: This pilot study aims to further understand and validate the efficacy and safety of KDGs for treating HT through clinical research and comprehensively assess the benefits of this intervention for patients.

Methods: This study is a multicenter, randomized, double-blind, placebo-controlled clinical trial. Participants meeting the HT diagnostic criteria will be randomly allocated to the intervention and control groups (n1=n2=70). The intervention group will receive KDG treatment, whereas the control group will receive a placebo treatment. All participants will undergo treatment for 3 months. Changes in antithyroid peroxidase antibody (TPOAb) levels will be the primary outcome. Secondary outcomes include antithyroglobulin antibodies (TGAb), thyrotropin, also known as thyroid stimulating hormone (TSH), triiodothyronine (T3), thyroid hormone (T4), serum free triiodothyronine (FT3), serum free thyroxine (FT4), thyroid ultrasonography, *IL17* mRNA and *FOXP3* mRNA, traditional Chinese medicine (TCM) syndrome efficacy scores, and quality of life scale scores. Throughout the treatment and follow-up periods, safety indicators, such as routine blood and urine tests, hepatic and renal function, electrocardiography, and major adverse reactions, will be monitored.

Results: The research protocol and informed consent form received approval from the Clinical Research Ethics Committee of Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, affiliated with Shanghai University of Traditional Chinese Medicine, on December 14, 2022 (Approval No. 2022-123). Participant recruitment commenced in June 2023. All intervention and concurrent data collection activities were scheduled for completion by October 2025. Data management is still ongoing; therefore, data analysis has not yet been performed.

Conclusions: This study's findings will offer initial clinical evidence regarding the efficacy of the TCM compound KDGs in modulating peripheral immunity in patients with HT, decreasing autoimmune antibody levels, ameliorating TCM syndromes, and enhancing quality of life. These results will serve as a basis for future large-scale trial designs.

Trial Registration: China Clinical Trials Registry ChiCTR2300070184; https://www.chictr.org.cn/showprojEN.html?proj=189169

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/80993

(JMIR Res Protoc 2026;15:e80993) doi:10.2196/80993

KEYWORDS

kangjia decoction granules; Hashimoto thyroiditis; evidence-based clinical trial; randomized controlled trial; protocol

Introduction

Hashimoto thyroiditis (HT), also known as chronic lymphocytic thyroiditis, is a chronic thyroid inflammation mediated by autoimmune disorders [1]. The prevalence of HT in the female population worldwide currently exceeds 10% and is increasing annually, showing a trend toward younger ages [2]. Patients with HT experience diffuse thyroid changes and have elevated levels of thyroid peroxidase antibodies (TPOAb) and thyroglobulin antibodies (TGAb) in the serum [3,4]. The levels of these antibodies to some extent reflect the severity of HT and serve as important indicators for evaluating its occurrence, development, and prognosis. In the early stages, patients with HT may show no obvious symptoms or may experience discomfort or swelling in the neck. As the disease progresses, some patients may exhibit hyperthyroidism symptoms, such as palpitations, sweating, and irritability. In later stages, the disease can lead to hypothyroidism, manifesting as delayed reactions, impaired memory, and other symptoms that significantly impact the quality of life. Studies have also demonstrated a close association between this disease and papillary thyroid cancer, as well as reproductive health issues, such as premature birth and fetal deformities. Therefore, HT is receiving increasing attention globally [5-7].

The pathogenesis of HT remains unclear. Most studies suggest that its onset is due to the disruption of the local immune balance in the thyroid gland, leading to the body's mistaken recognition of its own thyroid cell structure as an antigen, triggering the production of antibodies and initiating the formation of antigen–antibody complexes that lead to inflammation. Recent research has indicated that an imbalance between helper T cell 17 (Th17) and regulatory T cells (Tregs) is closely associated with the occurrence and development of HT [8,9]. Th17 cells mediate inflammation, whereas Tregs mediate immune suppression. Th17 and Treg cells work together to maintain a balance in the immune microenvironment of the body. During the progression of HT, Th17 cells increase in number, and their secretion of cytokines, such as interleukin (IL)-17, is elevated, whereas the number of Tregs and their secretion of FOXP3 are relatively reduced [10]. As HT advances, an imbalanced immune response and dysregulated recognition processes initiate apoptosis, leading to increased follicular cell apoptosis, irreversible fibrosis of follicular structures, reduced effective synthesis of thyroid hormone cells, and, ultimately, hypothyroidism [11]. At this stage, most patients require lifelong thyroid hormone replacement. Even when thyroid hormone levels are restored to normal, patients often experience symptoms such as fatigue, lower limb edema, low mood, and anxiety, which severely affect their quality of life [12,13].

A current clinical shortcoming is the lack of treatment options for patients with early-stage HT, especially for immune dysregulation-related causes. Treatment is typically initiated only when patients develop hypothyroidism, hyperthyroidism, or other complications when levothyroxine replacement or symptomatic treatment is provided. There is, therefore, an urgent need in clinical practice for treatment approaches that target overall immune regulation as the treatment goal, not only to address the disease itself but also to improve the overall quality of life of patients.

Recent research has shown that traditional Chinese medicine (TCM) plays an important role in the treatment of HT through its multitarget actions [14]. The research team integrated, analyzed, and statistically processed clinical data from hundreds of patients in the early stages of the disease and found that certain combinations of Chinese herbal medicines effectively reduced TPOAb and TGAb levels in patients with HT and significantly improved their clinical symptoms. Based on clinical experience and result analysis, the research team formulated the kangjia decoction granules (KDGs; Astragali Radix 18 g, Prunella vulgaris 15 g, Atractylodis Macrocephalae Rhizoma 6 g, Forsythiae Fructus 6 g, Rehmannia 6 g, and Cyperus Rotundus 6 g) for the treatment of HT.

This randomized, parallel, placebo-controlled, double-blind, and multicenter clinical trial aims to evaluate the effectiveness and safety of the KDGs and confirm their ability to regulate the immune system, treat HT, delay or prevent the occurrence of hypothyroidism, and improve quality of life.

Methods

Study Design

This is a randomized, double-blind, placebo-controlled, multicenter clinical trial in which the participants are randomized in a 1:1 ratio to either the treatment group (KDGs group) or the control group (placebo group). Each group will undergo a 12-week treatment. See Multimedia Appendix 1 for the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2025 checklist [15].

We will record all baseline observations during the screening period. Throughout the treatment period, participants will have monthly follow-up visits and complete TCM evidence-based efficacy and quality of life scale score assessments. TPOAb, TGAb, thyroid stimulating hormone (TSH), triiodothyronine (T3), thyroid hormone (T4), serum free triiodothyronine (FT3), serum free thyroxine (FT4), thyroid ultrasonography, and *IL17* mRNA and FOXP3 mRNA levels will be examined and recorded before and after treatment. The schedule of participants is shown in Table 1.

XSL•FO
RenderX

Table 1. Study schedule of assessments.

Item	Baseline (day 7~0)	0 weeks (day 0; visit 1)	4 weeks (day 28; visit 2)	8 weeks (day 56; visit 3)	12 weeks (day 84; visit 4)
Informed consent	✓				
Inclusion/exclusion criteria	✓				
Demographic information	✓				
History taking and recording	✓	✓	✓	✓	✓
Physical examination	✓	✓	✓	✓	✓
Traditional Chinese medicine evidence-based efficacy score		✓	✓	✓	✓
Quality of life scale score		✓	✓	✓	✓
Electrocardiogram	✓				✓
Blood test	✓				✓
Urinalysis	✓				✓
Liver function	✓				✓
Kidney function	✓				✓
TPOAb[a]	✓				✓
TGAb[b]	✓				✓
TSH[c]	✓				✓
T3[d]	✓				✓
T4[e]	✓				✓
FT3[f]	✓				✓
FT4[g]	✓				✓
IL[h]-17/FOXP3	✓				✓×
Thyroid ultrasound	✓				✓
Record of study drug usage	✓	✓	✓	✓	✓
Adverse event record	When necessary	When necessary	When necessary	When necessary	When necessary
Record of combined medication use	✓	✓	✓	✓	✓
Analysis of causes of shedding	✓	✓	✓	✓	✓

[a]TPOAb: thyroid peroxidase antibody.

[b]TGAb: thyroglobulin antibodies.

[c]TSH: thyroid stimulating hormone.

[d]T3: triiodothyronine.

[e]T4: thyroid hormone.

[f]FT3: free triiodothyronine.

[g]FT4: free thyroxine.

[h]IL: interleukin.

XSL•FO
RenderX

Participant Screening and Selection

Per the 2018 China Guidelines for Diagnosis and Treatment of Thyroid Diseases [16], HT is identified by a diffuse, firm goiter, often with isthmus or pyramidal lobe enlargement, and positive serum TPOAb or TGAb. Diagnosis can also be confirmed through fine needle aspiration cytology. Outpatients or inpatients meeting the diagnostic criteria for HT will be assessed for inclusion and exclusion (Textbox 1). Patients will not be recruited if they do not meet any of the inclusion criteria or meet any of the refusal or termination criteria. Participants are required to provide written informed consent, cooperate with the physician's treatment and follow-up, and provide information within the scope of ethical approval. Patients who do not meet these requirements will be excluded from the study. During this period, the participant's medical history, physical examinations, blood samples, and medical records will be collected and preserved.

Textbox 1. Inclusion, exclusion, and rejection criteria and termination.

Inclusion criteria:

1. Patients meet the diagnostic criteria of Hashimoto thyroiditis based on the guidelines for the diagnosis and treatment of thyroid diseases in China;

2. The Traditional Chinese medicine dialectical belongs to the syndrome of liver depression and spleen deficiency;

3. Age ≥ 18 and ≤ 65 years; gender not restricted;

4. Triiodothyronine, thyroid hormone, free triiodothyronine, and free thyroxine in the normal range of the tested values, thyroid stimulating hormone<10 mIU/L;

5. Aspartate transaminase (AST) and alanine transaminase (ALT) ≤ 3 × ULN creatinine ≤ 1.5 ULN;

6. The patients understand and agree to participate in this study and voluntarily sign the informed consent.

Exclusion criteria:

1. Patients with poorly controlled blood pressure on medication, systolic blood pressure160 mm Hg, or diastolic blood pressure>100 mm Hg;

2. Patients who have endured or are enduring hyperthyroidism or hypothyroidism or patients with hormone supplementation therapy after thyroid tumor surgery;

3. Patients who had myocardial infarction, heart failure, or unstable angina within six months prior to enrollment;

4. Patients with a history of stroke within 6 months prior to enrollment;

5. Pregnant or lactating women;

6. Patients with chronic alcoholism, drug dependence, or mental illness;

7. Patients with a history of tumors;

8. Patients with known hypersensitivity to the test drug or components thereof;

9. Participation or ongoing participation in other interventional clinical studies within one month prior to enrollment;

10. Those who, in the opinion of the investigator, should not participate in this clinical study.

Criteria for strike-out, discontinuance, and shedding:

1. Strike-out criteria

1.1 Misscheduling and misinclusion;

1.2 Those who, in the opinion of the investigator, are not suitable for continued participation in this study.

2. Discontinuation criteria

2.1 Adverse events or serious adverse events occurring during the trial, making the participant unsuitable for continued participation in the study;

2.2 Serious deterioration of the patient's disease or the occurrence of certain comorbidities, complications, and special physiological changes, leading to unsuitability for continued participation in the study.

3. Desertion criteria

3.1 Participants who show poor compliance, which is not in accordance with the provisions of the treatment or the follow-up investigation;

3.2 Existence of incomplete medical records, affecting the evaluation of efficacy;

3.3 Participants withdraw on their own.

The study aims to recruit 140 participants across 3 units. These 3 units include primary coordinating and implementing units. Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, affiliated with Shanghai University of Traditional Chinese Medicine, will recruit 80 participants. Longhua Hospital and Putuo Hospital, both affiliated with the Shanghai University of Traditional Chinese Medicine, will each recruit 30 participants. Each unit is responsible for specific aspects of the study, including patient recruitment, specimen collection, and data collection. Imaging examinations will be conducted in the department of each unit. Laboratory testing and data analysis

XSL•FO
RenderX

will be overseen by the Clinical Research Center of Yueyang Hospital.

To ensure participant safety, blinding, data quality, and compliance with the research protocol, all personnel directly involved in this study will receive comprehensive training and adhere to the Good Clinical Practice (GCP) guidelines.

Blinding

Blinding preservation: a two-tier blinding approach was implemented, with the first tier involving the assignment of group codes corresponding to each drug number (Group A or Group B), and the second tier involving the concealment of the actual drug types for the two groups. Sealed blinding documentation was duplicated and entrusted to the clinical research unit for safekeeping.

Blinding procedures: this study used a double-blinded method. Both blinding processes were overseen by the designated personnel responsible for maintaining the blinding. The first blinding involved the specification of Groups A and B, whereas the second blinding involved specifying the test and control groups. Blinding procedures were overseen by the project leader of the clinical study unit, the principal investigator, and the statistician. Randomization order not available to participant registrars and intervention assigners

Emergency blinding: emergency blinding was permitted in the event of serious complications or adverse events during the study that could impact conduct and treatment decisions. In such cases, the researcher, project leader, and clinical supervisor were required to be present, and detailed records of the reason, time, and location of blinding were documented and signed. All data were preserved.

Sample Size Calculation

This study was based on a clinical trial of a Chinese herbal medicine compound KDGs, for the treatment of HT. The study followed a randomized, double-blind, placebo-controlled design, with a change in the TPOAb level after 12 weeks of treatment serving as the primary efficacy endpoint. Using a 1:1 allocation ratio, the study participants were randomly assigned to either the TCM (KDGs) or the placebo group (placebo granules).

To determine the sample size, we used a clinical study sample size calculation formula:



Considering the mean change in TPOAb levels before and after treatment in the treatment group (μ1=120, SD σ1=200) and the control group (μ2=20, SD σ2=200) based on previous study results [17], the study aimed to achieve a power of 80% (1 - β=.80), with a significance level of 0.05 (two-sided) and a β of .20. The calculated sample size was 64, and with a 10% dropout rate, 140 participants were required for the study, with 70 individuals in each group.

Interventions

The KDGs Group

One packet of KDGs (Astragali Radix 18 g, Prunella vulgaris 15 g, Atractylodis Macrocephalae Rhizoma 6 g, Forsythiae Fructus 6 g, Rehmannia 6 g, and Cyperus Rotundus 6 g) is to be taken twice daily, mixed with warm water. The time of administration is half an hour after the morning and evening meals. The treatment will last for 3 months.

Placebo Group

One packet of placebo granules is to be taken twice daily, mixed with warm water. The time of administration is half an hour after the morning and evening meals. The treatment will last for 3 months.

The participants will return any unused pills and bottles during each follow-up visit. Unused pills will be counted and documented in the corresponding case report. KDGs and placebo granules will be provided by Shanghai Chiatai Qingchunbao Co Ltd.

Investigation Product

Manufacturing

The 6 herbal ingredients are to be extracted twice, each time with a 12-fold quantity of water, for 1.5 hours per extraction. The combined extracts are filtered, and the filtrate is concentrated under reduced vacuum to produce a dense extract with a relative density of 1.22-1.29 (at 60 °C). Subsequently, 1 g of sucralose, 4.4 g of caramel flavor essence, and a sufficient quantity of dextrin are added to the dense extract, followed by granulation and drying, to yield a final product of 1000 g.

Quality Control

The KDGs are manufactured and controlled according to the Chinese Pharmacopoeia and internal quality standards. The quality assessment methods are fully validated. Each batch, as documented in the Certificate of Analysis (No. 23W-004), must meet all release specifications outlined in the following table before clinical use (Multimedia Appendix 2. Stability data confirm the product remains within specifications under recommended storage conditions throughout the study.

Placebo Verification

A placebo similarity assessment was conducted to evaluate the blinding integrity of the investigational product. Using a standardized manual scoring method, KDGs (Batch: 20230401) and its matched placebo (Batch: 20230401) were determined to be highly comparable in terms of clarity, odor, and color. However, the similarity in taste was assessed to be less comparable. This finding has been documented and considered in the overall blinding strategy for the trial Multimedia Appendix 3.

Laboratory Methods for Biomarker Analysis: Measurement of IL-17 and FOXP3 mRNA Expression

1. Sample processing

Total RNA will be extracted from 250 μL of citrate-anticoagulated peripheral blood using Trizol reagent.

XSL•FO
RenderX

Briefly, the lysate will be stored at –80 °C until processing. After thawing, chloroform will be added for phase separation, followed by RNA precipitation with isopropanol. The RNA pellet will be washed with 75% ethanol, air-dried, and finally dissolved in RNase-free water.

1. cDNA synthesis.
2. cDNA was synthesized from total RNA using a commercial reverse transcription kit, following the manufacturer's instructions. The reaction mix was gently vortexed, centrifuged, and then incubated in a thermal cycler to complete the reverse transcription.
3. Quantitative polymerase chain reaction (PCR)
4. Gene expression was quantified by real-time PCR using SYBR Green chemistry. Reactions were set up in triplicate for each cDNA sample in a total volume of 20 µL, containing diluted cDNA, master mix, and specific primers. Amplification was carried out on a real-time PCR detection system with standard cycling parameters, including a melt curve stage for amplicon validation. For detailed primer sequences, please refer to
Table 2
.

Table 2. Primer sequence.

Gene	Primer sequence (5'-3')
GAPDH Forward primer (Human）	TGGGCAAGGTCATCCCTGAG
GAPDH Reverse primer (Human）	GCGTCAAAGGTGGAGGAGTG
IL17A Forward primer (Human）	TCCCACGAAATCCAGGATGC
IL17A Reverse primer (Human）	GCACTTTGCCTCCCAGATCA
FOXP3 Forward primer (Human）	TCCAGGACAGGCCACATTTC
FOXP3 Reverse primer (Human）	GCCACGTTGATCCCAGGT

Data Collection and Management

Data Collection

Data will be collected from the participants' daily symptom records, case report form (CRF) records of the trial group's medical staff, and patients' medical records. The data will be collected and managed using the REDCap (Research Electronic Data Capture; Vanderbilt University) electronic data capture tool at the Yueyang Integrated Traditional Chinese and Western Medicine Hospital of Shanghai University of Traditional Chinese Medicine. REDCap is a secure web software platform designed to support research data collection. This platform provides (1) an intuitive interface for data entry validation, (2) audit trails to track data manipulation and export procedures, (3) automated export procedures for seamless data download into popular statistical packages, and (4) procedures for data integration and interoperability with external resources.

Data Entry and Data Query Forms

The responsibility for data entry and management lies with the designated data administrator, who will establish the database and validation program. The data will then be input and verified by 2 trained data-entry operators. After reviewing the CRF, the identified input errors will be rectified until the input data matches the CRF. Any queries related to the CRF will be conveyed to the investigator using a data query table. These data query tables must be processed, signed, and dated by the researcher in a timely manner. Resolved data issues must be returned to the data manager, who will adjust and validate the data based on the feedback, update the database accordingly, and reevaluate the necessary data problem table.

Database Locking

Once the accuracy of the established database has been confirmed by blind checking, the database will be locked. Thus, the locked data cannot be altered. If any data revisions are deemed necessary after locking, an official statement signed jointly by the applicant, principal investigator, experimental project manager, statistician, and data manager must be provided. Furthermore, revisions will be made to the statistical analyses. The Data and Safety Monitoring Board will consist of an internal medicine physician, a sonographer, a medical statistician, and an ethicist and will conduct assessments throughout the study.

Electronic Data Encryption

This study will use REDCap for electronic data encryption. REDCap uses a comprehensive encryption strategy to ensure end-to-end data security. All data transmissions between browsers and the REDCap application are protected by Transport Layer Security/Secure Sockets Layer encryption. Data at rest is stored in a secured network environment with regular backups. For mobile data collection, devices should be password-protected to enable full device encryption. Additionally, application programming interface keys should not be hard-coded in scripts but stored securely using encrypted keyrings, with proper configuration to prevent accidental exposure of credentials in unencrypted files.

Participant Compliance

Informed Consent

A thorough process must be followed to secure participant understanding and cooperation. The investigational product and required laboratory tests must be provided free of charge.

Compliance Monitoring

The medication measurement should be used to monitor adherence. Compliance should be calculated as follows: (actual medication dose administered / planned medication dose) × 100%. Compliance between 80% and 120% is deemed acceptable.

XSL•FO
RenderX

Management of Poor Compliance

All concomitant medications in the eCRF should be recorded. Follow-up for participants with poor therapeutic response or compliance must be intensified (<80% or >120%).

Evaluation

Primary Outcome

- Difference in TPOAb levels before and after treatment.

Secondary Outcomes

- TGAb, TSH, T3, T4, FT3, FT4, thyroid ultrasound,
  IL17 mRNA,
  FOXP3 mRNA (peripheral blood),
  TCM syndrome score, quality of life scale score.

Security Outcome

- Blood routine, urine routine, hepatic and renal function, and electrocardiogram.

Adverse Event Observation and Recording

Following entry into the clinical trial, the patient's vital signs, symptoms, any adverse events related to the disease, and laboratory parameters will be recorded. Thus, the occurrence of adverse events may not necessarily be related to the study drugs. Adverse events were assessed and graded according to the Common Terminology Criteria for Adverse Events (CTCAE) version 5.0. Grade 1 (mild): asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated. Grade 2 (moderate): minimal, local, or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living. Grade 3 (severe or medically significant): medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care activities of daily living. Grade 4 (life-threatening): urgent intervention indicated. Grade 5 (death): death related to the adverse event. The initial report of any serious adverse event must be made within 24 hours of discovery to the principal investigator, Dr He Yanming. The principal investigator is responsible for ensuring the subsequent report is escalated to the Research Administration Office at Yueyang Hospital of Integrated Traditional Chinese and Western Medicine (telephone: 65161782, extension Research Admin). The final mandatory report must be submitted to the Drug Research Supervision Department of the National Medical Products Administration (telephone: 010-68313344-1013). Any causal relationship between the adverse event and the investigational drug will be determined by classification. All serious adverse events were followed until resolution or until the event was considered stable.

In the case of a serious adverse event in a participant, the details of the medication must be disclosed, and unblinding can only be performed with the consent of the principal investigator. Once an emergency letter has been read, the case will be considered a dropout.

Evaluation Index Detection Method

Serological analysis was conducted at the laboratory department of Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, affiliated with Shanghai University of Traditional Chinese Medicine. Blood-related parameters were assessed using a Roche automatic electrochemiluminescence immunoassay analyzer (Roche Diagnostics GmbH) and its corresponding kits. Patients refrained from eating after 8 PM the evening before blood collection and fasted from 8 PM to 9 AM on the day of collection. Subcenter serum samples were transported from a 4 ˚C incubator to the main laboratory department within a 2-hour timeframe.

Thyroid ultrasound assessment: thyroid color Doppler ultrasound evaluations were conducted in the B ultrasound facility at each research center. Measurements of the width, thickness, and isthmus thickness of the left and right thyroid lobes were taken and documented.

IL17 and *FOXP3* mRNA levels were assessed in venous blood samples collected from all patients before and after treatment. Specifically, 5 mL of venous blood was drawn into EDTA tubes, from which 250 µL was aliquoted into 1.5 mL EP tubes. Subsequently, 750 µL of Trizol was added to each tube, and the samples were stored at –80 ˚C for subsequent analysis. The quantification of *IL17* and *FOXP3* mRNA levels in peripheral blood was conducted using fluorescence quantitative PCR.

The TCM symptom score scale (see Multimedia Appendix 4 for the Symptom Score Scale) was revised for HT of liver depression and spleen deficiency type based on clinical syndrome characteristics. This revision, guided by the Beijing expert consensus on HT treatment (2021) [18], involved input from three highly experienced TCM internal medicine clinicians and preliminary investigation findings. The scale encompassed 7 items: main symptoms (fatigue, depression, irritability), secondary symptoms (abdominal distension, pain, excessive breathing, insomnia), and tongue and pulse characteristics (white or greasy fur, stringy or deep pulse). Patients were assessed for symptom severity and assigned scores, with higher scores indicating greater improvement in TCM syndromes.

The Thyroid-Specific Patient-Reported Outcome questionnaire-39-item (ThyPRO-39) scale (see Multimedia Appendix 5 for the Quality of Life Scale) is a reliable and valid instrument designed to assess the quality of life in patients with benign thyroid conditions [19]. It comprises 39 items categorized into 13 subscales, covering somatic symptoms (goiter, hyperthyroidism, hypothyroidism, and eye symptoms), physiological and psychosocial symptoms (fatigue, cognition, anxiety, depression, emotional susceptibility, social activities, daily life, and appearance), and overall quality of life. Each item is scored on a 5-point Likert scale. Subscale scores are calculated by averaging the item scores, dividing by 4, and multiplying by 100, resulting in a standard score ranging from 0 to 100. Higher scores on the ThyPRO-39 indicate a lower quality of life [20].

XSL•FO
RenderX

Statistical Analysis Method

Statistical Software

For statistical analysis of data in this study, SAS 9.2 (SAS Institute) statistical software was used.

Fundamental Principles

In this study, all statistical inferences were made using a 2-sided test at a significance level of α=.05. The confidence intervals of the study parameters were estimated using 2-sided 95% CIs to indicate the statistical significance of differences.

Handling Missing Data and Sensitivity Analysis Plan for Missing Data

Missing data will be handled using Multiple Imputation under the assumption of Missing at Random. Multiple (eg, m=20) complete datasets will be generated by imputing missing values using a multivariate model incorporating key baseline and outcome variables. The planned primary and secondary analyses will be performed separately on each imputed dataset. Final estimates (eg, treatment effects, confidence intervals, and *P* values) will be obtained by pooling the results from all m datasets using Rubin rules. A sensitivity analysis assuming a Missing Not at Random scenario will be conducted to assess the robustness of the conclusions. Safety evaluations were not performed for missing data.

To ensure the robustness of the conclusions from the primary analysis, the following prespecified sensitivity analyses will be performed regarding missing data. First, the results from the primary analysis using multiple imputation will be compared with the results from an analysis based on the complete-case dataset. Second, a worst-case scenario analysis will be conducted. This analysis will assume that all participants in the intervention group with missing primary outcome data experienced treatment failure (eg, no symptomatic improvement or no reduction in antibodies), while all participants with missing data in the control group experienced treatment success. The primary outcome will be re-evaluated under this extreme assumption. These sensitivity analyses aim to assess whether the study conclusions are sensitive to the methods used for handling missing data. Consistency between the results of these analyses and the primary analysis would provide stronger support for the reliability of the final conclusions.

Dropout Analysis

The chi-square test was used to compare the overall dropout rate between the 2 groups. Differences in demographic characteristics, treatment received, and other variables of the dropout participants from the experimental and control groups were described and analyzed to understand the reasons for dropout.

Statistical Description

In this study, quantitative variables were summarized using their mean, SD, minimum, maximum, Q1, median, and Q3. Confidence intervals were provided when necessary. Quantitative variables were expressed as “mean and SD” if they followed a normal distribution and “median and IQR” if they did not follow a normal distribution. Qualitative variables are summarized in terms of frequency distribution and composition ratio.

Inferential Statistics

Baseline data, efficacy indicators, and safety indicators underwent inferential statistical analysis. Baseline data: statistical inferential analysis was conducted based on the type of variable to compare the comparability of the participants in the experimental and control groups. Quantitative variables were analyzed using 2-tailed *t* tests for normally distributed data and nonparametric rank-sum tests for nonnormal distribution. Qualitative variables were analyzed using chi-square tests. Efficacy indicators: comparative analysis of quantitative variables within groups was conducted using paired *t* tests for normally distributed data and paired nonparametric rank-sum tests for nonnormal distribution. Differences in efficacy indicators before and after the intervention were calculated separately for the experimental and control groups, and *t* tests or rank-sum tests were used to assess the statistical significance of these differences. Analysis of covariance, which considers the center effect, was used to provide the mean (or adjusted mean), difference in means, and 95% CIs for each group. Paired chi-square tests were used for qualitative variables, and the Wilcoxon rank-sum test was used for ranked variables. If the baseline comparability of the experimental and control groups was poor, general linear models and logistic regression models were applied for adjustment. Safety analysis: the chi-square test was used to compare the incidence of adverse events between the 2 groups, and a list was used to describe the adverse events that occurred in the trial. Intergroup and intragroup comparisons of quantitative indicators were performed using the corresponding differential tests. Laboratory test results were analyzed for changes in normal/abnormal values before and after the study, and a list was provided for abnormal values.

Clinical Trial End

The trial will end once 140 participants have been recruited, and all participants have completed 12 weeks of treatment and 12 weeks of follow-up.

Ethical Considerations

The research protocol and informed consent form received approval from the Clinical Research Ethics Committee of Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, affiliated with Shanghai University of Traditional Chinese Medicine, on December 14, 2022 (Approval No 2022-123). Before participating in the research, we will obtain written informed consent from all participants. We will ensure voluntary participation and maximum confidentiality of the information provided during the interviews. The trial will be supervised by the Ethics Committee. Any moderate or severe adverse reactions will be reported to the Ethics Committee. Participants who experience adverse reactions during the trial will receive free access to medical care.

All research-related information will be securely stored at the research site, and all participant information will be stored in a locked filing cabinet with restricted access. To ensure participant confidentiality, all laboratory samples, reports, data collection,

XSL•FO
RenderX

processes, and management forms will be identified using coded ID numbers. All records containing names or other personal information (such as informed consent forms) will be stored separately from the coded research records. All local databases will be protected using password-protected access systems. Forms, lists, logs, appointment books, and other lists linking participant identification to other identifying information will be stored in restricted access areas in individually locked files.

The relevance and necessity of the research questions, research design, and patient-centered documents (including informed consent forms, participant information sheets, symptom diaries, all follow-up forms, and promotional materials) have been subjected to review by the public. The public engagement group comprised individuals with experienced and inexperienced backgrounds. Efforts will be made to collaborate with patient advocacy groups to ensure that easily comprehensible abstracts of research findings are shared with participants and the broader patient community. The intervention burden will be comprehensively and reasonably assessed through communication between participants and researchers. While the results will not be communicated directly to the participants, the final outcomes will be disseminated. Throughout the follow-up period, all the examination results will be communicated to each participant.

All study data, including source documents, CRFs, and electronic databases, will be retained by the Sponsor for a period of at least 5 years after the completion or termination of the study. Upon expiration of the retention period, electronic data will be permanently deleted using irreversible methods, and paper records will be confidentially destroyed and incinerated. A certificate of disposal will be issued and maintained in the study files.

Blood samples collected during the study will be stored under coded identifiers at Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, affiliated with Shanghai University of Traditional Chinese Medicine. For specimens authorized for future research, as permitted by the informed consent form, these specimens will be retained indefinitely in a secured biobank for use in future research projects. Any future use will be subject to approval by a Research Ethics Board. When selecting eligible participants, researchers must provide detailed information about the clinical trial, including the purpose of the trial, trial procedures, potential benefits and risks, and the rights and responsibilities of participants, so that participants can fully understand and have sufficient time to consider. After the participants' questions have been answered satisfactorily, they will sign informed consent forms and receive the physician's contact details so that they can contact the physician at any time if there are any changes in their condition.

Randomization

After screening, SAS software was used to generate random numbers, and a sequential randomization plan was used to allocate the study participants to the experimental and control groups to ensure group balance. The investigator sequentially assigned a drug number to each participant based on the enrollment order and dispensed the corresponding test drug number that could not be selected arbitrarily. The number of assigned drugs remained constant throughout the study.

Results

Participant recruitment commenced in June 2023. All intervention and concurrent data collection activities are scheduled for completion by May 2025, with the 12-week follow-up assessment postintervention expected to conclude by the end of May 2025. Data management has been finalized, and data analysis is presently in progress. For data inquiries, kindly reach out to the primary author. The conclusive outcomes of the research will be disseminated in a journal indexed in the Science Citation Index. The study protocol is illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Flow chart of the study procedure. HT: Hashimoto thyroiditis; KDGs: kangjia decoction granules; TCM: traditional Chinese medicine;.

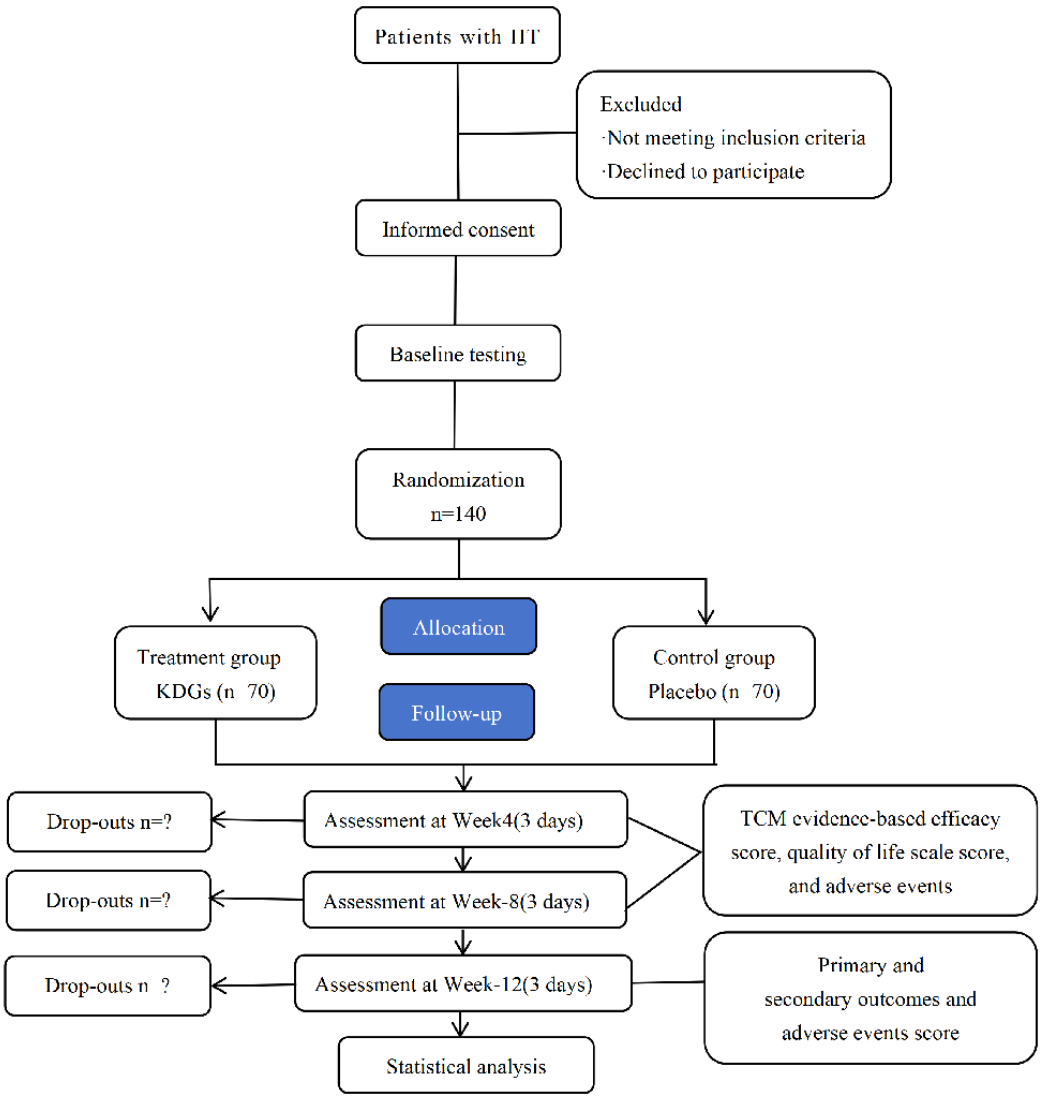

Discussion

Principal Findings

This study anticipates that the KDGs can effectively reduce TPOAb and TGAb levels in patients with HT of the liver depression and spleen deficiency pattern, without significant adverse reactions. It is also expected to ameliorate patients' clinical symptoms and signs, thereby improving their quality of life. Furthermore, KDGs may decrease the expression of *IL17* mRNA and increase the expression of *FOXP3* mRNA in patient serum, helping to improve the immune microenvironment and regulate immune balance.

During the progression of HT, aberrant immune responses and dysregulated recognition processes modulate the induction of apoptosis, leading to increased follicular cell death, irreversible fibrosis of the follicular architecture, and diminished capacity for thyroid hormone synthesis, ultimately resulting in hypothyroidism. Consequently, most patients require lifelong thyroid hormone replacement. Even with normalized thyroid hormone levels, patients often experience symptoms such as fatigue, lower limb edema, depression, and anxiety, which severely impact their quality of life. In conventional Western medicine, the primary treatment for HT relies on levothyroxine replacement to correct thyroid dysfunction. However, this approach lacks effective intervention for the early-stage immune imbalance in HT and shows limited efficacy in reducing autoantibodies (TPOAb and TGAb). Traditional Chinese Medicine identifies "liver depression and spleen deficiency" as a core pathogenesis of HT. Previous clinical observations and studies have demonstrated the efficacy of the KDGs granules in ameliorating HT symptoms and reducing autoantibody levels. KDGs is known for its properties of soothing the liver, strengthening the spleen, clearing heat, and dispersing stagnation. Animal studies have further validated its potential to lower TPOAb and TGAb levels and enhance thyroid function in model rats.

XSL•FO
RenderX

Although numerous clinical studies on Chinese medicine exist, there is a general lack of high-quality, large-sample, randomized, double-blind controlled trials. To address these gaps, this study adopts a "disease-pattern integrated" model, focusing on the core TCM pattern of "liver depression and spleen deficiency," and uses a fixed herbal formula for intervention. Through a rigorous randomized, double-blind, controlled trial design, it aims to provide high-level evidence for TCM pattern-based treatment. Regarding efficacy evaluation, most existing studies primarily focus on changes in thyroid function and antibody titers, with insufficient exploration of the underlying mechanisms of immune imbalance [21-23]. This study innovatively incorporates key markers of Th17/Treg cell balance—the proinflammatory cytokine IL-17 and the transcription factor FOXP3—as core outcome measures. This allows for an in-depth investigation of the drug's potential immunomodulatory targets, effectively complementing traditional evaluation models. In terms of study design, addressing limitations such as small sample sizes and lack of blinding in some previous herbal medicine research, this study uses a multicenter, randomized, double-blind, controlled trial design to minimize bias and enhance the reliability and generalizability of the findings.

This study protocol possesses several notable strengths. First, it provides a novel theoretical integration of the TCM principle of "soothing the liver and strengthening the spleen" with contemporary immunoregulatory theory. The pathophysiology of HT involves a dysregulated neuro-endocrine–immune network that manifests in systemic symptoms, a concept well-aligned with the TCM framework. We propose a mechanistic correspondence: "Liver depression" may correlate with stress-axis (hypothalamic-pituitary-adrenal axis) dysfunction and neurotransmitter disturbances underlying depression and anxiety, while "spleen deficiency" may relate to gut-barrier dysfunction, metabolic dysregulation, and inflammation-induced sarcopenia, reflecting the TCM concept of the "spleen governing the muscles." The therapeutic principle of soothing the liver and strengthening the spleen thus represents a multitarget, systems-level intervention aimed at restoring immune-neuro–metabolic homeostasis. This integrative model offers a coherent pathophysiological rationale for targeting the co-occurrence of emotional, cognitive, and muscular symptoms in HT, underscoring the holistic perspective of TCM and providing a scientific foundation for the study hypothesis.

Second, the selection of outcome measures is comprehensive and mechanistically informed.Beyond assessing conventional HT biomarkers (TPOAb and TGAb), the protocol includes key mediators of the Th17/Treg pathway (*IL17* and *FOXP3* mRNA expression). This dual approach allows for a systematic evaluationof the intervention's effects on both downstream autoantibody production and upstream immunologic imbalance, offering deeper insights into its potential mechanism of action.

Third, the methodology is robust. Using a multicenter, randomized, double-blind, placebo-controlled designminimizes bias, enhances generalizability, and is conducive to generating high-level clinical evidence, which addresses a common limitation in prior TCM interventional research.

Finally, the study addresses a clear unmet clinical need. It directly targets the persistent challenge of reducing thyroid autoantibodies and modulating the underlying immune dysregulation in HT, a gap in current conventional management. Positive results could offer a novel adjunctive strategy aimed at modifying the immune microenvironment and improving patient outcomes.

However, this study also has several limitations. First, Chinese herbal formulas are characterized by multiple components and targets. The specific active material basis of KDGs and its detailed mechanisms of action cannot be fully elucidated merely through changes in clinical indicators. Moreover, although the TCM pattern diagnosis of "liver depression and spleen deficiency" can be quantified using scales, a degree of subjectivity remains, which may affect the homogeneity of the enrolled patients regarding their pattern presentation. Second, HT is a chronic disease requiring a prolonged period for immune modulation. This study's duration may be insufficient to fully evaluate the long-term efficacy of the intervention, the stability of antibody levels over time, and its long-term impact on disease progression (eg, changes in thyroid volume and incidence of clinical hypothyroidism). Finally, the inference of immunological changes in this study is primarily based on the detection of cytokines and transcription factors in peripheral blood, which constitutes indirect evidence. It cannot directly reflect the immune cell infiltration and inflammatory microenvironment within the thyroid gland itself, thereby limiting the depth of mechanistic interpretation.

Based on the positioning and potential findings of this study, future research directions may include, at the mechanistic level, further studies that could use flow cytometry to directly analyze the proportions of Th17 and Treg cells in patient peripheral blood. Using animal or cell models, research could delve deeper into the regulatory effects of KDGs and its active components on key signaling pathways related to Th17/Treg cell differentiation and function, such as RORγt, STAT3, and TGF-β/Smad. Furthermore, research dimensions should be expanded, for instance, by conducting long-term follow-up studies to assess the intervention's long-term benefit in delaying progression to clinical hypothyroidism.

To facilitate translation, should the positive outcomes of this trial be confirmed, future studies could validate its efficacy and safety in real-world clinical settings. Research could also explore its synergistic effects when combined with Western medications (eg, selenium yeast and levothyroxine), aiming to establish a comprehensive management pathway for HT integrating TCM and Western medicine characteristics.

Conclusions

This study's findings will offer initial clinical evidence regarding the efficacy of the TCM compound KDGs in modulating peripheral immunity in patients with HT, decreasing autoimmune antibody levels, ameliorating TCM syndromes, and enhancing quality of life. These results will serve as a basis for future large-scale trial designs.

XSL•FO
RenderX

Acknowledgments

This work was supported by the Three Year Action Plan for Promoting Clinical Skills and Clinical Innovation in Municipal Hospitals (grant SHDC2022CRD009). The authors thank their colleagues from Shanghai University of Traditional Chinese Medicine, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai Longhua Hospital and Shanghai Putuo Hospital for their full support in patient recruitment and treatment. The authors are also grateful for the efforts and assistance of the researchers involved in this trial.

Funding

This work was supported by the project A Multicenter Clinical Study on the Intervention of “Treating Liver and Spleen Simultaneously” Chinese Medicine in Hashimoto’s Thyroiditis (Grant No. SHDC2022CRD009); the project The Effect of Runzao Yishen Formula on Vascular Endothelial Function in Patients with Type 2 Diabetes Mellitus of Kidney-Yin Deficiency Type (Grant No. 2024MW74); and the “Siming Fund” of Shuguang Hospital, Shanghai University of Traditional Chinese Medicine (Grant No. SGKJ-202524). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

Conceptualization: YH and HC
Data curation: Duanchun Zhang, Dan Zhang, and XQ
Formal analysis: Duanchun Zhang and Dan Zhang
Funding acquisition: HC and YH
Investigation: Duanchun Zhang, Dan Zhang, and XQ
Methodology: HC and YH
Project administration: Dan Zhang, HC, and YH
Resources: HC and YH
Software: Duanchun Zhang and YH
Supervision: HC and YH
Validation: HC and YH
Visualization: Duanchun Zhang, Dan Zhang, and XQ
Writing – original draft: Duanchun Zhang and XQ
Writing – review & editing: All authors

Conflicts of Interest

None declared.

Multimedia Appendix 1
Reporting checklist for protocol of a clinical trial. Based on the SPIRIT 2025 guidelines.
[DOCX File , 31 KB - resprot_v15i1e80993_app1.docx]

Multimedia Appendix 2
Certificate of Analysis.
[DOCX File , 16 KB - resprot_v15i1e80993_app2.docx]

Multimedia Appendix 3
Placebo Similarity Evaluation Trial for Kangjia Decoction Granules.
[DOCX File , 23 KB - resprot_v15i1e80993_app3.docx]

Multimedia Appendix 4
Symptom Score Scale (TCM Version).
[DOCX File , 23 KB - resprot_v15i1e80993_app4.docx]

Multimedia Appendix 5
Questionnaire on quality of life related to thyroid function (ThyPRO39se).
[DOCX File , 23 KB - resprot_v15i1e80993_app5.docx]

XSL•FO
RenderX

References

1. Vargas-Uricoechea H. Molecular mechanisms in autoimmune thyroid disease. Cells 2023;12(6):918 [FREE Full text] [doi: 10.3390/cells12060918] [Medline: 36980259]
2. Hu X, Chen Y, Shen Y, Tian R, Sheng Y, Que H. Global prevalence and epidemiological trends of Hashimoto's thyroiditis in adults: a systematic review and meta-analysis. Front Public Health 2022;10:1020709 [FREE Full text] [doi: 10.3389/fpubh.2022.1020709] [Medline: 36311599]
3. Guan H, de Morais NS, Stuart J, Ahmadi S, Marqusee E, Kim MI, et al. Discordance of serological and sonographic markers for Hashimoto's thyroiditis with gold standard histopathology. Eur J Endocrinol 2019;181(5):539-544. [doi: 10.1530/EJE-19-0424] [Medline: 31536967]
4. Barić A, Brčić L, Gračan S, Škrabić V, Brekalo M, Šimunac M, et al. Thyroglobulin antibodies are associated with symptom burden in patients with Hashimoto's thyroiditis: a cross-sectional study. Immunol Invest 2019;48(2):198-209 [FREE Full text] [doi: 10.1080/08820139.2018.1529040] [Medline: 30332318]
5. Ragusa F, Fallahi P, Elia G, Gonnella D, Paparo SR, Giusti C, et al. Hashimotos' thyroiditis: epidemiology, pathogenesis, clinic and therapy. Best Pract Res Clin Endocrinol Metab 2019;33(6):101367. [doi: 10.1016/j.beem.2019.101367] [Medline: 31812326]
6. Rahman S, Archana A, Jan AT, Dutta D, Shankar A, Kim J, et al. Molecular insights into the relationship between autoimmune thyroid diseases and breast cancer: a critical perspective on autoimmunity and ER stress. Front Immunol 2019;10:344 [FREE Full text] [doi: 10.3389/fimmu.2019.00344] [Medline: 30881358]
7. Ferrari S, Bocci G, Di Desidero T, Elia G, Ruffilli I, Ragusa F, et al. Lenvatinib exhibits antineoplastic activity in anaplastic thyroid cancer in vitro and in vivo. Oncol Rep 2018;39(5):2225-2234. [doi: 10.3892/or.2018.6306] [Medline: 29517103]
8. Chaker L, Razvi S, Bensenor IM, Azizi F, Pearce EN, Peeters RP. Hypothyroidism. Nat Rev Dis Primers 2022;8(1):30. [doi: 10.1038/s41572-022-00357-7] [Medline: 35589725]
9. Sekiya T, Kagawa S, Masaki K, Fukunaga K, Yoshimura A, Takaki S. Regulation of peripheral Th/Treg differentiation and suppression of airway inflammation by Nr4a transcription factors. iScience 2021;24(3):102166 [FREE Full text] [doi: 10.1016/j.isci.2021.102166] [Medline: 33665581]
10. Chen Z, Wang Y, Ding X, Zhang M, He M, Zhao Y, et al. The proportion of peripheral blood Tregs among the CD4+ T cells of autoimmune thyroid disease patients: a meta-analysis. Endocr J 2020;67(3):317-326 [FREE Full text] [doi: 10.1507/endocrj.EJ19-0307] [Medline: 31827051]
11. Salazar-Viedma M, Vergaño-Salazar JG, Pastenes L, D'Afonseca V. Simulation model for Hashimoto autoimmune thyroiditis disease. Endocrinology 2021;162(12) [FREE Full text] [doi: 10.1210/endocr/bqab190] [Medline: 34496027]
12. Batóg G, Dołoto A, Bąk E, Piątkowska-Chmiel I, Krawiec P, Pac-Kożuchowska E, et al. The interplay of oxidative stress and immune dysfunction in Hashimoto's thyroiditis and polycystic ovary syndrome: a comprehensive review. Front Immunol 2023;14:1211231 [FREE Full text] [doi: 10.3389/fimmu.2023.1211231] [Medline: 37588599]
13. Groenewegen KL, Mooij CF, van Trotsenburg AP. Persisting symptoms in patients with Hashimoto's disease despite normal thyroid hormone levels: does thyroid autoimmunity play a role? A systematic review. J Transl Autoimmun 2021;4:100101 [FREE Full text] [doi: 10.1016/j.jtauto.2021.100101] [Medline: 34027377]
14. Yuan J, Qi S, Zhang X, Lai H, Li X, Xiaoheng C, et al. Local symptoms of Hashimoto's thyroiditis: a systematic review. Front Endocrinol (Lausanne) 2022;13:1076793 [FREE Full text] [doi: 10.3389/fendo.2022.1076793] [Medline: 36743914]
15. Chan A, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomized trials. JAMA 2025;334(5):435-443. [doi: 10.1001/jama.2025.4486] [Medline: 40294593]
16. Compilation group of Guidelines for Diagnosis and Treatment of Thyroid Diseases in China, Endocrinology Society of Chinese Medical Association. China guidelines for diagnosis and treatment of thyroid diseases-thyroiditis. Chin J Intern Med 2008;9:5. [doi: 10.1002/agm2.12165]
17. Cao H, Zhao H, Fu M, He Y, Yang H. effect of chinese herbal compound granules on antibody levels and thyroid function in patients with Hashimoto's thyroiditis. Chin J Gerontol 2016;36(09):2108-2110 [FREE Full text] [doi: 10.1177/0002764289032005003]
18. Ding Z. Beijing expert consensus on diagnosis and treatment of Hashimoto thyroiditis with integration of traditional and western medicine (2021, Beijing). China Med Herald 2022;19(34):4-7. [doi: 10.20047/j.issn1673-7210.2022.34.01]
19. Watt T, Bjorner JB, Groenvold M, Cramon P, Winther KH, Hegedüs L, et al. Development of a short version of the thyroid-related patient-reported outcome ThyPRO. Thyroid 2015;25(10):1069-1079. [doi: 10.1089/thy.2015.0209] [Medline: 26214034]
20. Wong CKH, Choi EPH, Woo YC, Lang BHH. Measurement properties of ThyPRO short-form (ThyPRO-39) for use in Chinese patients with benign thyroid diseases. Qual Life Res 2018;27(8):2177-2187. [doi: 10.1007/s11136-018-1857-9] [Medline: 29671250]
21. Zhang M, Zhang Y. Effect of Qi-boosting phlegm-transforming and goiter-dispersing on serum thyroglobulin antibody and thyroperoxydase antibody in Hashimoto's thyroiditis patients. J Tradit Chin Med 2007;48(5):414-415. [doi: 10.13288/j.11-2166/r.2007.05.013]

XSL•FO
RenderX

22. Fei Z, Ma C, Chen D. Clinical study on modified yanghe decoction combined with levothyroxine sodium in treating 40 cases of hypothyroidism due to Hashimoto's thyroiditis. Jiangsu J Tradit Chin Med 2018;50(7):36-39. [doi: 10.3969/j.issn.1672-397X.2018.07.013]
23. Chen W, Zhang C, Wang S, Song Y, Gao L, Zhang X. Yiqi Yangyin therapy on Hashimoto's thyroiditis of Qi and Yin deficiency syndrome. Hebei J Tradit Chin Med 2021;43(5):731-735. [doi: 10.3969/j.issn.1002-2619.2021.05.006]

Abbreviations

CRF: case report form
CTCAE: Common Terminology Criteria for Adverse Events
FT3: free triiodothyronine
FT4: free thyroxine
GCP: Good Clinical Practice
HT: Hashimoto thyroiditis
IL: interleukin
KDG: kangjia decoction granule
PCR: polymerase chain reaction
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
T3: triiodothyronine
T4: thyroid hormone
TCM: traditional Chinese medicine
TGAb: thyroglobulin antibodies
Th17: T cell 17
ThyPRO-39: Thyroid-Specific Patient-Reported Outcome questionnaire-39-item
TPOAb: thyroid peroxidase antibody
Treg: regulatory T cell
TSH: thyroid stimulating hormone

Edited by J Sarvestan; submitted 23.Jul.2025; peer-reviewed by YW Lai; comments to author 28.Oct.2025; accepted 25.Dec.2025; published 26.Jan.2026.

Please cite as:
Zhang D, Zhang D, Qu X, Cao H, He Y
Effectiveness and Safety of Kangjia Decoction Granules for the Treatment of Hashimoto Thyroiditis: Protocol for a Randomized, Double-Blinded, Placebo-Controlled, Multicenter Clinical Trial
JMIR Res Protoc 2026;15:e80993
URL: https://www.researchprotocols.org/2026/1/e80993
doi:10.2196/80993
PMID:

©Duanchun Zhang, Dan Zhang, Xiaoxiao Qu, Huihong Cao, Yanming He. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Real-World Effectiveness and Noninferiority Evaluation and Comparison of Messenger RNA–Based and Protein-Based COVID-19 Vaccines: Protocol for the BEEHIVE Randomized Study With a Hybrid Effectiveness Design

Sarang K Yoon[1], DO, MOH; German L Ellsworth[2], MD, MPH, MOH; Steph Battan-Wraith[3], MSPH, PhD; Andrew L Phillips[1], MD, MOH; Rebecca V Fink[3], MPH; Joshua Griffin[4], MHA; Elizabeth A K Rowley[3], DPH, MS; Jacob McKell[5], BS; Ashley S Smith[3], MPH, MLT; Riley Campbell[1], BS; Jesse Williams[1], BS; Sarah W Ball[3], MS, SCD, MPH; Hongwei Zhao[2], SCD; Brandy Warren[6], MPH; Matthew D Rousculp[6], MPH, PhD; Matthew S Thiese[2], MSPH, PhD

[1]Division of Occupational and Environmental Health, School of Medicine, University of Utah, Salt Lake City, UT, United States

[2]Rocky Mountain Center for Occupational and Environmental Health, University of Utah, Salt Lake City, UT, United States

[3]Westat (United States), Rockville, MD, United States

[4]The Rocky Mountain Center for Occupational and Environmental Health, School of Medicine, University of Utah, Salt Lake City, UT, United States

[5]The Rocky Mountain Center for Occupational and Environmental Health, University of Utah, Salt Lake City, UT, United States

[6]Real-World Evidence, Novavax (United States), Gaithersburg, United States

Corresponding Author:
Sarang K Yoon, DO, MOH
Division of Occupational and Environmental Health
School of Medicine
University of Utah
50 North Medical Drive
Salt Lake City, UT, 84132
United States
Phone: 1 801 213 4500
Email: sarang.yoon@hsc.utah.edu

Abstract

Background: Surveillance of COVID-19 vaccine effectiveness (VE) was extensive upon vaccine introduction; however, it declined after the withdrawal of pandemic status in May 2023. Continued monitoring of updated vaccine formulations is needed to ensure the maintenance of VE in the face of evolving viral strains.

Objective: The Booster Epidemiological Evaluation of Health, Illness and Vaccine Efficacy (BEEHIVE) study (NCT06065176), a randomized trial with a hybrid design, was developed to assess the real-world VE of the 2023-2024 Pfizer–BioNTech and Novavax COVID-19 vaccine formulations targeting the XBB.1.5 SARS-CoV-2 variant.

Methods: This study was designed to enroll approximately 1500 participants aged ≥18 years from the Salt Lake City, Utah, area who had previously received ≥2 doses of an authorized messenger RNA (mRNA)–based COVID-19 vaccine but had not received a dose of the 2023-2024 formulation. The study used a randomized, hybrid design comprising 2 blinded groups assigned to receive the 2023-2024 formula of either the Novavax COVID-19 vaccine or the Pfizer–BioNTech COVID-19 vaccine and a nonrandomized, observational control group of volunteers who chose not to receive a 2023-2024 vaccine dose during the study. Follow-up lasted 24 weeks and included symptom surveys and self-administered COVID-19 antigen testing, both occurring weekly. The primary aim was to compare VE (defined as prevention of symptomatic SARS-CoV-2 infection) between study-vaccinated participants and the control group. The secondary aim was to determine the relative VE of the Pfizer–BioNTech mRNA and Novavax 2023-2024 COVID-19 vaccines. Secondary objectives included assessing how the number of previous COVID-19 vaccinations impacted VE of the 2023-2024 COVID-19 vaccines; identifying predictors and associated factors for asymptomatic versus symptomatic infection and/or prolonged or severe illness; examining factors associated with post–COVID-19

XSL•FO
RenderX

conditions; and evaluating participants' knowledge, attitudes, and practices related to COVID-19 vaccination. Participant engagement was maintained via online and text-based reminders and surveys, as well as researcher follow-up.

Results: Participants were recruited from November 2023 through March 2024, with 452 and 457 participants randomized to the Novavax and Pfizer–BioNTech vaccine groups, respectively, and 279 participants enrolled in the control group. SARS-CoV-2 variants from the XBB, JN.1, KP.2, and KP.3 lineages were in circulation in the United States and the Utah region during data collection. The study ended on September 9, 2024, with results expected to be published in 2026.

Conclusions: Data from this study will provide valuable real-world VE data for a dose of the Novavax COVID-19 vaccine or the Pfizer–BioNTech COVID-19 vaccine after an mRNA-based COVID-19 primary series.

Trial Registration: ClinicalTrials.gov NCT06065176; https://www.clinicaltrials.gov/study/NCT06065176

International Registered Report Identifier (IRRID): RR1-10.2196/80858

(JMIR Res Protoc 2026;15:e80858) doi:10.2196/80858

KEYWORDS

protein; COVID-19; SARS-CoV-2; omicron; surveillance; vaccine effectiveness; clinical trial

Introduction

Background

Early in the COVID-19 pandemic, there was vigilant surveillance by regulatory authorities to monitor SARS-CoV-2 cases and assess the impact of COVID-19 vaccines on controlling viral spread. When the public health emergency status was not renewed in May 2023, surveillance mechanisms greatly diminished or stopped entirely [1]. With the continued evolution of immune-evasive variants, COVID-19 remains a public health concern [1]. Maintaining surveillance of the vaccine effectiveness (VE) of updated COVID-19 vaccines targeting new strains is informative for ongoing vaccination campaigns and for the development of new vaccine formulations [2].

There are 3 COVID-19 vaccines approved for use by the US Food and Drug Administration (FDA) [3-5]. Of these, 2 are messenger RNA (mRNA)–based vaccines (Pfizer–BioNTech COVID-19 vaccine and Moderna COVID-19 vaccine) and 1 is an adjuvanted recombinant spike protein–based vaccine (Novavax COVID-19 vaccine). A key difference among these formulations is the use of mRNA encoding the SARS-CoV-2 spike protein in the Pfizer–BioNTech and Moderna vaccines [3,4], compared with the use of recombinant spike protein with Matrix-M adjuvant in the protein-based Novavax vaccine [5-7]. Prototype monovalent COVID-19 vaccines of each of these formulations, targeting the ancestral Wuhan strain, have undergone a series of updates since their initial authorization (ie, bivalent Wuhan and BA.4/5, monovalent XBB.1.5, and monovalent JN.1 lineage) to target circulating lineages of relevance at the time [8].

There are limited clinical and real-world data on the VE of XBB.1.5–based vaccines [9,10]. Notably, existing reports focus on assessments of symptomatic infection over a limited period. A better understanding is needed of the protection these vaccine formulations may provide against both symptomatic and asymptomatic infection for several months after vaccination.

Objectives

The Booster Epidemiological Evaluation of Health, Illness and Vaccine Efficacy (BEEHIVE) study (NCT06065176) was developed as a randomized trial with a hybrid design that includes a randomized comparison between active vaccine groups and an observational control group that did not receive a dose of the study vaccine. The goal was to assess the VE of the 2023-2024 formulations (against XBB.1.5) for the Pfizer–BioNTech and Novavax COVID-19 vaccines in a real-world setting [11]. Importantly, this study was designed with a comprehensive surveillance structure to identify SARS-CoV-2 infections, regardless of symptom presence, over a 6-month period following study vaccination. The study protocol and end points are described here, along with the novel recruitment methods and strategies used to maintain participant engagement.

Methods

Study Design

The BEEHIVE study was designed to enroll approximately 1500 participants. Participants were aged ≥18 years; resided in Salt Lake City, Utah (including a 60-mile radius of surrounding areas); had previously received ≥2 doses of an mRNA-based COVID-19 vaccine authorized by the US FDA; and met the eligibility criteria described in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Aged ≥18 years
- Previous receipt of ≥2 doses of a US Food and Drug Administration (FDA)–authorized messenger RNA–based COVID-19 vaccine
- Comfortable reading and responding to SMS text messaging and emails sent in English (on own or with interpreter assistance)
- Plan to remain in the greater Salt Lake City, Utah, area for the next 12 months
- Daily access to the internet (via cell phone, laptop, desktop, or tablet) and a phone with text messaging capabilities
- Willingness to complete the following:
  - Weekly symptom and illness surveillance (sent via text and email)
  - Enrollment, midstudy, and end-of-study surveys
- Willingness to be contacted periodically by study staff via text, email, and/or telephone as part of study activities
- Willingness to self-collect rapid antigen tests (approved by FDA Emergency Use Authorization for COVID-19 detection) weekly, when prompted for study purposes, if experiencing a qualifying symptomatic illness, or upon rapid antigen test confirmation of an asymptomatic infection and to report the results via the study portal
- Willingness to attend an in-person visit to receive a supply of rapid antigen tests and training on their use (all participants) and to receive an updated 2023-2024 COVID-19 vaccine (randomized participants)

Exclusion criteria

- Lives with another person who is already enrolled in the Booster Epidemiological Evaluation of Health, Illness and Vaccine Efficacy (BEEHIVE) study, as reported on the eligibility survey
- Previous hypersensitivity reaction to a COVID-19 vaccine, as reported on the eligibility survey
- COVID-19 infection (infection confirmed via self-report of a positive real-time reverse transcription polymerase chain reaction assay and/or rapid antigen test) ≤90 days before enrollment
- Receipt of a COVID-19 vaccine ≤90 days before enrollment
- Participation in other vaccine or investigational product trials
- Medical history of immunosuppression
- Receipt of Johnson & Johnson COVID-19 vaccine before study enrollment
- Receipt of any investigational prevention therapies for SARS-CoV-2 infections, such as prophylactic antiviral medications or other immune system–modifying interventions, ≤90 days before trial vaccine administration
- Unwillingness to provide electronic consent or to self-report occupation, work responsibilities, or previous COVID-19 illness

Study groups comprised 2 randomized, blinded groups who received the 2023-2024 formulation of either the Novavax COVID-19 vaccine or the Pfizer–BioNTech COVID-19 vaccine and a nonrandomized, observational control group of volunteers who participated in the study but did not elect to receive a study vaccine or any 2023-2024 COVID-19 vaccine formulation during the surveillance period (Figure 1A).

XSL•FO
RenderX

Figure 1. Study design: (A) recruitment, screening, and study schema, and (B) decision tree for symptom and rapid antigen test monitoring. For this study, the interpretation of results is based on the participant's status (eg, asymptomatic or has COVID-19–like illness symptoms) on the first day (day 0) of each testing. Additional tests on days 1 and 3 were required if a participant had symptoms on day 0 and if a test was positive on day 0, regardless of symptoms status. To be considered a true negative case, all 3 tests must result as negative. FDA: Food and Drug Administration; mRNA: messenger RNA.

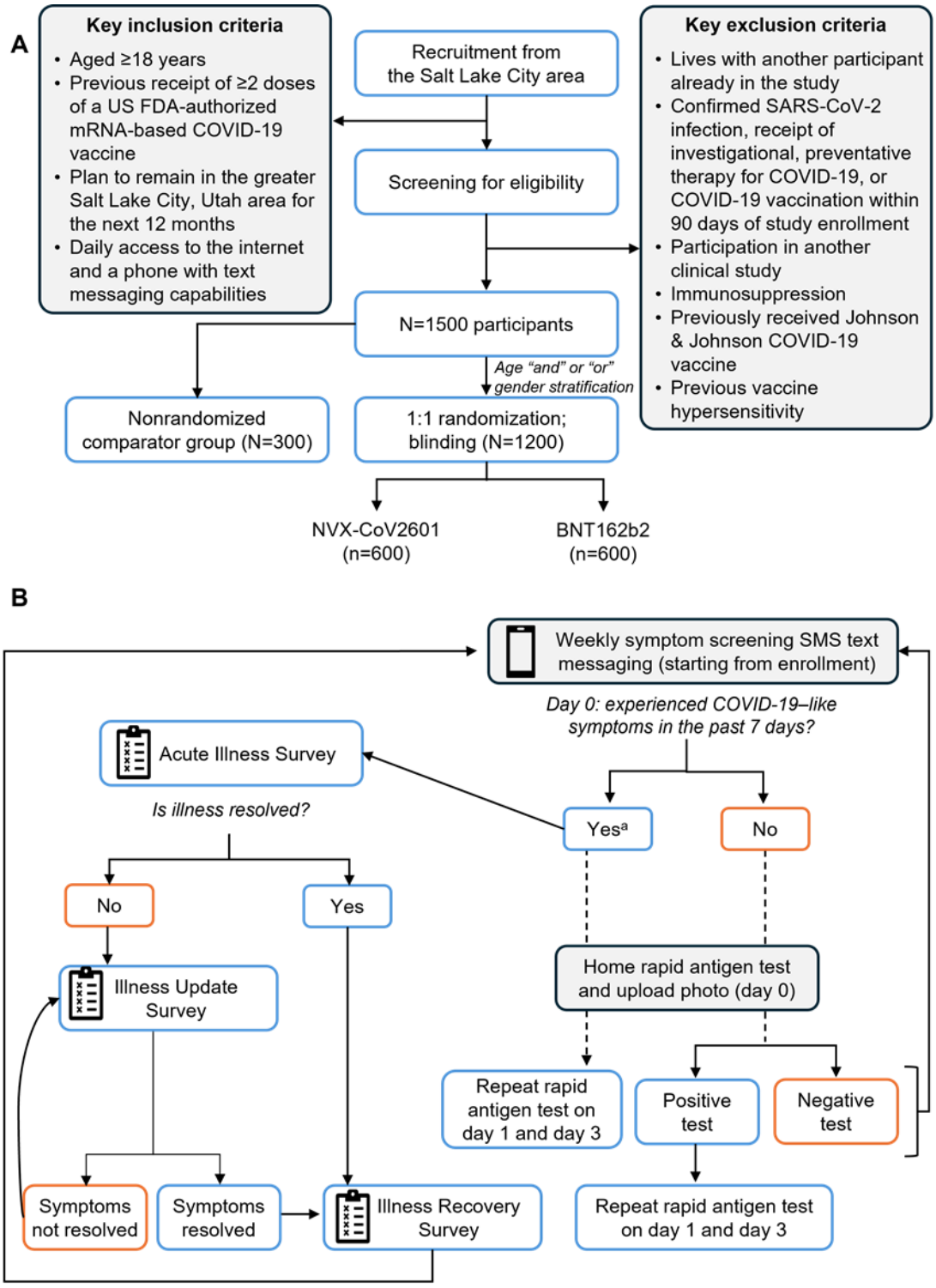

Participant Recruitment

Several unique mechanisms were used to facilitate participant recruitment. Radio and television advertisements were deployed in the Salt Lake City area, and flyers for the BEEHIVE study were distributed to patient groups representative of populations that could meet eligibility criteria (Multimedia Appendix 1). Participants from previous COVID-19 studies (ie, RECOVER [Research on the Epidemiology of SARS-CoV-2 in Essential Response Personnel] and PROTECT [Pediatric Research Observing Trends and Exposures in COVID-19 Timelines]) [12,13] or an influenza vaccine study (ie, RAIVEN [Randomized

XSL•FO
RenderX

Assessment of Influenza Vaccine Efficacy Network]) [14] who consented to be contacted for future studies were sent recruitment invitations.

Additional outreach efforts included joining occupational health vaccination campaigns and other existing mass-vaccination events, advertising through social media and local news outlets, staffing booths at community events, and deploying mobile vaccination stations. Collaborations were established with local community groups, clinics, businesses, universities, older adult centers, and fire and police departments. Potential participants received a link to the online eligibility survey via email and were informed of their eligibility status upon survey completion. After completion of consent and Health Insurance Portability and Accountability Act (HIPAA) authorization forms, eligible participants were invited to create an online portal account, complete the enrollment survey, and schedule an in-person enrollment visit online (Multimedia Appendix 2).

Vaccination History

For participants who provided consent to having their vaccine records obtained by the study staff, the state immunization information system or registry was accessed to obtain COVID-19 and influenza vaccination status and history. Participants were also asked to self-report their COVID-19 and influenza vaccination history and to provide vaccine documentation to the study.

Study Objectives

There were 2 aims of the BEEHIVE study. The primary aim was to compare VE between study-vaccinated participants and a control group who declined receipt of an updated COVID-19 vaccine (2023-2024 formula) in preventing symptomatic SARS-CoV-2 infection. The secondary aim was to determine the relative VE (rVE) of the 2023-2024 formulation of the Pfizer–BioNTech mRNA and Novavax COVID-19 vaccines in preventing symptomatic SARS-CoV-2 infection.

Secondary objectives included evaluating the effect of the number of previous COVID-19 vaccinations on the rVE of the Novavax COVID-19 vaccine. Additional secondary objectives were to assess predictors of asymptomatic versus symptomatic SARS-CoV-2 infection; clinical characteristics and outcomes associated with COVID-19, including illness duration and severity; sociodemographic and health characteristics associated with prolonged or severe illness; the impact of COVID-19 on indicators of functioning (eg, missed work, ability to complete normal work and home activities, and working while ill); the proportion of COVID-19 illnesses that were medically attended; and factors associated with seeking medical care and treatment. Illness characteristics and duration for primary infection versus reinfection will be compared, and the effect modification of VE by sociodemographic characteristics, occupation, health status, and other risk factors will be evaluated.

The goals of additional outcomes analyses are to assess whether the 2023-2024 vaccine formulation modifies illness severity and duration among participants with breakthrough infection. Data related to the incidence of post–COVID-19 condition (PCCs) were collected, along with identification of factors associated with PCC symptoms; assessment of whether these symptoms differed between randomized and control study groups; evaluation of whether PCC symptoms were modified by the number of previous vaccine doses; and assessment of the impact of PCC symptoms on overall functioning and mental health. PCCs were defined as any self-reported condition (continuous and relapsing or remitting) occurring after COVID-19 and persisting for at least 1 month after the initial illness, including fatigue, memory problems, dyspnea, sleep problems, joint pain, and/or newly developed medical conditions (eg, myocarditis, heart failure, cardiac dysrhythmias, stroke, chronic kidney diseases, neurologic conditions, and diabetes) [15]. PCCs were grouped by symptom category (eg, cardiovascular, respiratory, and neurological) and by number of symptoms (≥1 or ≥2). Finally, the knowledge, attitudes, and practices (KAPs) of participants related to 2023-2024 vaccines are to be characterized, and associations between KAPs and subsequent vaccination behaviors (including vaccine refusal, hesitancy, or incomplete adherence to vaccination recommendations) will be examined.

Ethical Oversight

Before study implementation, the protocol, informed consent form, participant education and recruitment materials, data collection instruments, and other protocol-related documents were approved by the University of Utah and Westat institutional review boards (I00172738 and 8084.01, respectively). In-person signatures were obtained for informed consent and HIPAA authorization forms. Participants were able to access copies of the consent forms through their portal accounts and received a signed hard copy. All participants provided consent and were allowed to withdraw at any time.

The research data platform and study database were maintained by Westat, Inc. To protect confidentiality, we used assigned identification numbers rather than personal information on study forms. We stored data in locked files and/or secured computers. For information from this study presented publicly or published in a medical journal, individuals will not be identified by name or by any other personally identifiable information. Participants received compensation for study activities (refer to the Remuneration and Engagement section for additional details).

Randomization and Blinding

A computerized random-number generator produced the randomized list for the active vaccinated arms before study enrollment. Randomly permuted blocks (designated A-F) of size 6, with 3 assignments to each vaccine type per block, were generated to achieve 1:1 allocation. Randomization was stratified by sex and age (18-49 years and ≥50 years). Eligible participants who provided consent indicating agreement to receive an updated 2023-2024 vaccine were able to schedule their first study visit and vaccination appointment.

Participants were randomized to either the Novavax COVID-19 vaccine group or the Pfizer–BioNTech COVID-19 vaccine group; however, based on appointment selection and attendance, some participants who consented may not have completed randomization or received a study vaccine. Enrolled participants who provided consent to the study but declined receipt of the 2023-2024 formulation were included in the nonrandomized

XSL•FO
RenderX

control group until enrollment was full. Participants who changed their minds at the time of their first study visit were able to switch groups (ie, initially agreed to be randomized to receive a vaccine but then switched to the nonrandomized control group, or vice versa) at the time of their appointment.

Randomized participants, study investigators, and study personnel were blinded to the type of vaccine received during the study, and multiple procedures were performed to maintain blinding. Specifically, vaccine packaging, administration, and documentation practices were standardized to avoid any distinguishable characteristics between the 2 vaccines. Additionally, all study staff, except for licensed health care providers who administered the vaccines, were blinded and received training on the importance of consistent data collection and maintaining blinding. Vaccines were prepared in a designated space that was obscured from participants and study personnel.

Study personnel who administered the vaccines were unblinded and underwent specialized training to ensure that vaccine type was not inadvertently disclosed. These personnel were not involved in data collection or analysis and interacted with the participants only to confirm their identity and administer the vaccine. The Westat randomization coordinator and limited data management staff were also unblinded to facilitate development and maintenance of the randomization list and systems. These staff were not involved in data collection, data analysis, or direct interaction with participants. Participants and study investigators were unblinded at the end of data collection and at the conclusion of the study, respectively. All study-related documents and samples contain a unique identifier for each participant to preserve blinding.

Interventions

Randomized participants were to receive a single dose of the study vaccine (0.3 ml for Pfizer–BioNTech or 0.5 ml for Novavax) administered in the deltoid muscle and were to be monitored for 15 minutes after vaccination for any allergic reactions. An investigational new drug application was submitted to and approved by the US FDA for use of the Novavax COVID-19 vaccine in this study, as it was authorized under an Emergency Use Authorization at the time of the study [5].

Surveillance: Surveys

Surveys were developed based on a pre-existing framework used in prospective clinical trials evaluating influenza and COVID-19 vaccines [13,14,16]. In the enrollment survey, the participants self-reported their sociodemographic characteristics, health status and behaviors, occupational status and employment history, chronic medical conditions, COVID-19 and influenza vaccination history, and previous SARS-CoV-2 infections.

Postvaccination surveys were shared 24 hours, 48 hours, and 6 days after study vaccination to obtain information on potential vaccine side effects. Active surveillance of outcomes occurred for participants on the same day each week, beginning the week following the enrollment and study vaccination visit, and continued for 24 weeks (Figure 1B).

Symptoms were assessed weekly through surveillance and illness messaging surveys for 24 weeks after enrollment and study vaccination. Surveillance included weekly subjective assessment of COVID-19–like illness (CLI), which was defined as experiencing at least 1 of the following symptoms: chills, malaise, fatigue, headache, cough, shortness of breath, sore throat, runny nose or nasal congestion, nausea or vomiting, diarrhea, muscle or body aches, or change in smell or taste.

Follow-up surveys were deployed at designated time points and allowed participants to provide sociodemographic and illness updates, assess KAPs, and collect information on PCCs (Textbox 2).

Textbox 2. Focus areas of follow-up surveys.

- Update information on occupation and work responsibilities; health status; potential exposures to COVID-19 at work, home, and in the community; and use of personal protective equipment in these settings
- Report a qualifying illness and/or submit illness specimens for any reason and provide an estimate of when the illness event occurred
- Assess participants' knowledge, attitudes, and practices regarding COVID-19 vaccines; intention to be vaccinated; and, after vaccination, recall of overall health status on the days following vaccination
- Report symptoms related to post–COVID-19 conditions (eg, symptoms lasting longer than 1 month), if applicable

During the surveillance period, if a participant indicated a CLI symptom on a weekly survey or reported an acute illness between their designated weekly surveillance contacts, they were directed to the Acute Illness Survey (Figure 1B). In this survey, they were asked to identify qualifying symptoms from a detailed checklist, confirm the date of symptom onset, and indicate whether the symptoms were ongoing. If the participant did not confirm symptoms of CLI, they returned to weekly surveillance. If the symptoms were ongoing, participants were prompted to collect a specimen and submit the results according to study procedures. Illness Update and Illness Recovery Surveys were deployed to monitor symptoms until a participant reported ≥90% recovery (Figure 1B). Survey data were collected electronically via a text messaging interface and online surveys (Multimedia Appendix 2).

Surveillance: Testing

All participants received OHC COVID-19 Antigen Self-Test kits (Osang Healthcare) for weekly use, along with instructions on how to use the kit, interpret results, and upload results to the study portal (Multimedia Appendix 1). Assay sensitivity is 91% and relative specificity is 99%, as described by the manufacturing company. Anterior nasal swab specimens were self-collected weekly on the same day (day 0), and a photograph of the results was uploaded regardless of symptom status (Figure 1B).

XSL•FO
RenderX

If a participant reported an acute illness and/or collected an illness specimen on or outside their assigned routine specimen day, they were asked to collect a second specimen on the next day (day 1) and a third specimen 2 days later (day 3). For this study, the interpretation of test results was based on the participant's symptom status on the first day of testing (day 0) and whether they did or did not have CLI symptoms.

Safety Assessments

Safety assessments included monitoring and recording of solicited (local and systemic reactogenicity events) and unsolicited adverse events (AEs), serious AEs, and AEs of special interest (AESIs). Solicited AEs were collected through the Postvaccination Survey. Unsolicited AEs could be reported at any time from study vaccination through completion of the last study-related procedure.

The relation of each event to the study vaccine (related or not related), whether the event was expected or unexpected, whether the event was serious or not, and the intensity of AEs and serious AEs were investigator determined. Severity was categorized as mild (grade 1), moderate (grade 2), severe (grade 3), life-threatening (grade 4), or death (grade 5). Protocol-defined AESIs were myocarditis and pericarditis and were collected by participant report, with verification through medical record review and follow-up interviews by study clinicians to obtain additional details.

Participants received an informational sheet describing symptoms of myocarditis and pericarditis, a list of nearby medical facilities for care seeking, and a letter for their evaluating physician, which included recommendations for clinical management. Probable and confirmed cases of AESIs were defined as previously described [17]. Safety reporting related to the investigational new drug status of the Novavax COVID-19 vaccine included sharing suspected unexpected serious adverse reactions and relevant medication errors with the appropriate regulatory authorities.

Statistical Analyses

The primary population for VE analyses will be the modified intent-to-treat (mITT) population, which includes all randomized participants who received a study vaccine and responded to ≥1 active surveillance survey, and all participants in the control group. Any participant with a CLI-associated SARS-CoV-2 infection confirmed by COVID-19 rapid antigen test 0 to 6 days after vaccination will be excluded from the mITT population. Analyses will also be performed with the per-protocol population, which includes participants who met eligibility criteria, received a study vaccine per protocol, participated in study surveillance by responding to at least 1 surveillance contact during the SARS-CoV-2 circulation period (active surveillance), and did not receive another SARS-CoV-2 vaccine outside of the study during the study period. If participants were in the nonvaccinated control group but subsequently received a 2023-2024 updated vaccine, they will be censored at that point. Any participants withdrawing early from the study will be censored at that time and included in analyses. There will be no imputation for missing data.

Using publicly available data [18] on COVID-19 cases in Utah from October 1, 2022, to May 28, 2023, the average rate of reported cases in fully vaccinated individuals from fall 2022 through spring 2023 was 39.64 per 100,000 person-days. On the basis of previously reported survey results of clinic and laboratory testing rates for COVID-19 [19] and the use of at-home test kits in this study, the estimated case reporting was 43.1% higher. Using these values and adjustments to the observed rates of infection in Utah in fall 2022 leads to a conservative estimate of the rate of CLI-associated SARS-CoV-2 infection in the control group, confirmed by COVID-19 rapid antigen test, of 80 per 100,000 person-days. With 1500 participants enrolled (1200 randomized and 300 nonrandomized control), assuming ≤15% attrition, there would have been 80% power to detect a VE of ≥49.4% for the primary aim (ie, VE between randomized and control participant groups), examined with a 2-sided alpha 0.05 superiority test.

For the secondary aim, noninferiority of the Novavax COVID-19 vaccine versus the Pfizer–BioNTech COVID-19 vaccine will be declared if the lower one-sided 95% CI for rVE is >50%. With 600 participants per vaccine group and an assumed infection rate of 60 per 100,000 person-days in the Pfizer–BioNTech COVID-19 vaccine arm and a rVE of 10% for the Novavax COVID-19 vaccine versus the Pfizer–BioNTech COVID-19 vaccine (ie, Novavax>Pfizer–BioNTech) [20,21], there would be 80% power for the noninferiority objective. The positive rVE used in the sample size calculation is based on estimates from large retrospective studies [20,21], which showed that the Novavax COVID-19 vaccine provides added protection against medically attended and severe SARS-CoV-2 infection. Sample size calculations for study aims were performed in PASS 2023 (NCSS, LLC) using the Cox proportional hazards model to estimate hazard ratios. Certain secondary objectives will be underpowered if there are insufficient numbers of nonsymptomatic cases, reinfections, and/or PCCs during the study period.

Participant demographics and characteristics will be summarized descriptively in the mITT population. The Cox proportional hazards model will be used to analyze time-to-event data to evaluate VE in preventing symptomatic COVID-19 infection. The observation period started 7 days after the receipt of the updated 2023-2024 vaccine for the randomized groups and at the enrollment date for the control group. Event onset time is defined as the earlier of the date of a positive rapid antigen test or the date of CLI symptom onset reported within a week of a positive rapid antigen test. Covariates of age (continuous variable) and vaccination dose history (categorical variable) will be included in the model to control for confounding. Additional demographic variables will be considered for inclusion if they change the hazard ratio for the treatment variable by ≥10%. Subgroup analyses based on age group and immune challenge (ie, most recent COVID-19 infection or vaccination) were conducted to assess potential effect modification across age and immune-challenge subgroups. All analyses will be performed using SAS (version 9.4 or higher; SAS Institute Inc), or R (version 4.1.2 or higher; R Foundation for Statistical Computing).

XSL•FO
RenderX

Remuneration and Engagement

Participants received compensation in the form of Amazon, Inc. gift cards, with the value proportional to the number of study activities completed (Multimedia Appendix 1). Newsletters were shared throughout the study with investigators and participants, including information such as enrollment numbers, instructions on how to remain compliant and up-to-date with study protocol and surveys, and updates on COVID-19 activity in the area. Study branding was consistently presented from recruitment throughout the study.

Results

Participant Enrollment and Data Collection

Participant screening occurred from November 17, 2023, through March 14, 2024. Monitoring occurred from enrollment through the end of the study on September 9, 2024. A total of 1188 participants were enrolled: 909 in the randomized vaccine groups (Novavax COVID-19 vaccine, 452/909, 49.7%; Pfizer–BioNTech COVID-19 vaccine, n=457, 50.3%) and 279 in the nonrandomized control group. Data analyses are ongoing, and results are expected to be published in 2026.

Vaccination Timing and SARS-CoV-2 Circulation

Vaccination and data collection during the study occurred when XBB.1.5-, JN.1-, KP.2-, and KP.3-lineage SARS-CoV-2 variants were circulating in the United States and the Utah region (Figure 2 [22]). Study vaccinations were relatively evenly distributed throughout the enrollment period, with most occurring when XBB.1.5 and JN.1 represented ≥50% of SARS-CoV-2 strains circulating in Utah (January 19, 2023, and December 30, 2023, respectively).

Figure 2. Circulating SARS-CoV-2 strains during the Booster Epidemiological Evaluation of Health, Illness and Vaccine Efficacy (BEEHIVE) study [22] (adapted from nexstrain.org under the CC-BY-4.0 license).

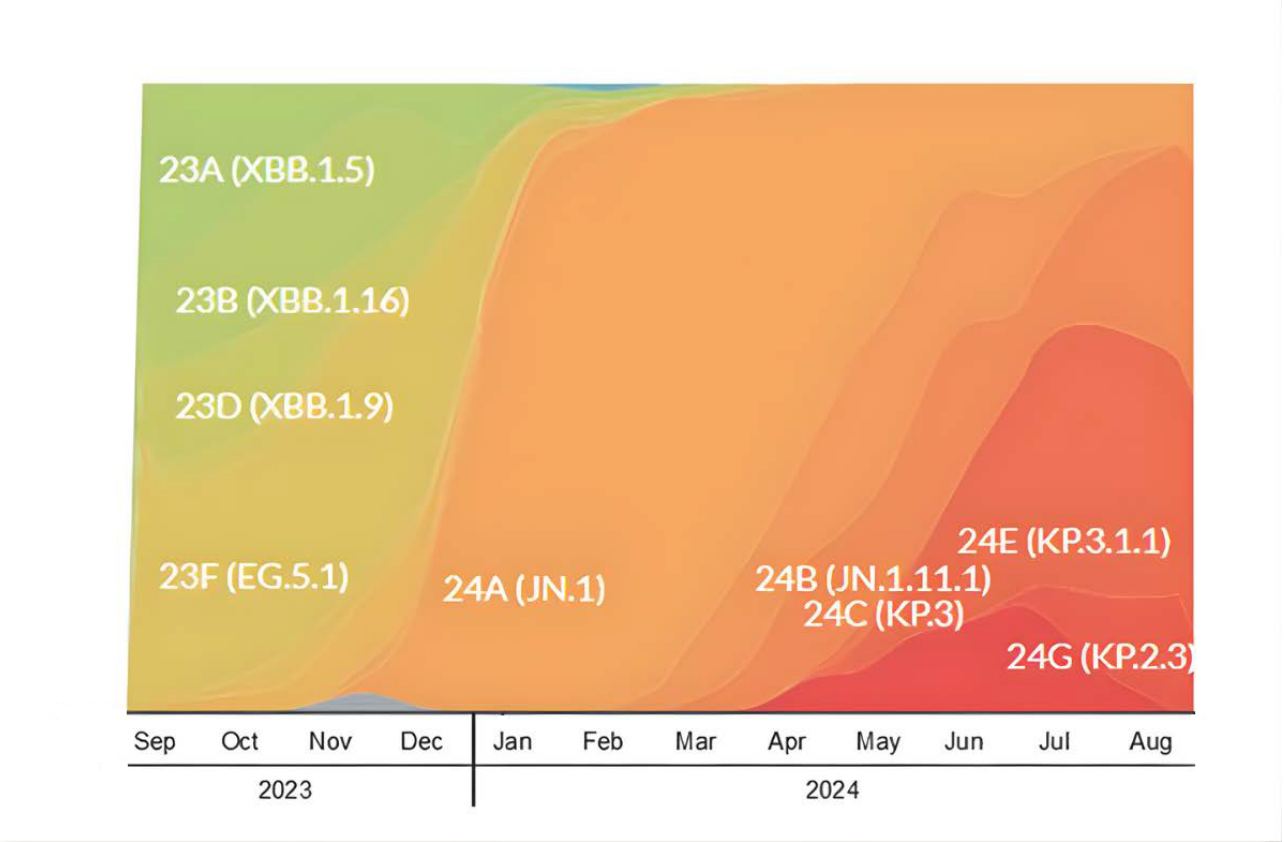

Discussion

The BEEHIVE study will provide real-world VE data for the Novavax COVID-19 vaccine when administered as an additional dose after at least a primary series of an mRNA-based COVID-19 vaccine and will compare these VE data with those of the Pfizer–BioNTech COVID-19 mRNA vaccine. Outcomes may provide information on the benefits of heterologous dosing or the use of mixed vaccine formulations, for administration beyond primary vaccination.

Access to participant cohorts from previous studies allowed recruitment of individuals who were more likely to participate and facilitated access to sociodemographic information and vaccination history. Mobile vaccine clinics, attendance at community events, and collaborations with civic groups were more effective in engaging typically underrepresented communities (eg, older adults and racial and ethnic minority individuals) [23,24] than standard clinical trial site strategies. Electronic surveillance and the use of online and text-based surveys supported efficient data collection and sustained participant engagement throughout the 24-week surveillance period.

A potential limitation of this study includes the underrepresentation of participants in the older adult categories. Additionally, the variability in SARS-CoV-2 circulation and

XSL•FO
RenderX

the introduction of new variant strains during the study period would impact infection rates and symptom severity. Notably, the testing kits, which were identical for all study participants, had received Emergency Use Authorization from the US FDA for Omicron and subvariant testing. The single-site nature of the trial may restrict geographic diversity among the participants; however, the feasibility of a real-world study design was established and can be adapted for future multisite studies. Because rapid antigen tests were used as part of the real-world study design, repeat testing procedures were included in the design to mitigate concerns regarding the absence of more robust clinical assays (eg, polymerase chain reaction). Another consideration is that the control group that opted out of receiving a study vaccination may introduce selection bias. Briefly, participants who declined the recommended vaccine may have differed demographically or behaviorally from those who chose to receive a study vaccine [25].

Participant enrollment (n=1188) was lower than the target of 1500 described in sample size calculations. We acknowledge that the study is underpowered due to recruitment challenges, and the implications of the smaller sample size on effect detectability will be discussed in the final report. Additionally, assumptions of superior rVE of Novavax compared with mRNA vaccines based on medically attended COVID-19 infection may not directly translate to protection against symptomatic SARS-CoV-2 infection. Future noninferiority studies should consider sample size calculations assuming an rVE of 0 to ensure a more robust study design.

Continued surveillance of COVID-19 cases and VE data for vaccine formulations updated based on predominant SARS-CoV-2 strains is necessary to inform public health policy and management. The BEEHIVE study will provide valuable real-world VE data for populations representative of individuals who have received the 2023-2024 updated COVID-19 vaccine, including those who received a dose of the Novavax COVID-19 vaccine after receipt of ≥2 mRNA-based COVID-19 vaccines.

Acknowledgments

The authors thank the study participants who volunteered for this study. The authors also thank Chengbin Wang, MD, PhD, and Mark Thompson, PhD, for their initial support. Medical writing and editorial support were provided by Kelly M Fahrbach, PhD, CMPP, and Ebenezer M Awuah-Yeboah, BS, of Ashfield MedComms, an Inizio company. Additional editorial support was provided by Anar Murphy, PhD, CMPP, of Novavax, Inc.

Data Availability

The datasets generated or analyzed during this study are available in the ClinicalTrials.gov database [11].

Funding

Funding was provided by Novavax, Inc.

Authors' Contributions

Conceptualization: SKY, SB-W, EAKR, SWB, MST, ALP, GLE, ASS
Data curation: HZ, EAKR, SB-W, JM, RC, JW
Formal analysis: SKY, SB-W, EAKR, RVF, SWB, HZ
Funding acquisition: BW, MDR, SKY, MST, ALP, SWB
Investigation: SKY, GLE, ALP, MST, JM, RC, JW, JG
Methodology: SKY, SB-W, EAKR, SWB, MST, ALP, GLE, ASS, HZ, JM, JG
Project administration: SKY, SB-W, RVF, SWB, MST, ALP, GLE, RC, JW, JM, JG, ASS
Resources: BW, SKY, MST, ALP, GLE, ASS, SWB
Software: SKY, ALP, GLE, JM, RC, JW, SB-W, SWB, RVF, EAKR, HZ
Supervision: SKY, MDR, JM, JG, SWB, SB-W, RVF, ASS
Validation: HZ, EAKR, SB-W, SKY, MDR
All authors reviewed and edited manuscript drafts and approved the final version for submission.

Conflicts of Interest

SKY has received funding from Novavax, Inc, for this study; received funding from the US Centers for Disease Control and Prevention (CDC) for previous COVID-19 vaccine studies; and served as an advisory board member for Pfizer–BioNTech without compensation for that role. GLE, ALP, JG, JM, RC, JW, and HZ have received institutional grant funding from Novavax, Inc. SB-W, RVF, EAKR, ASS, and SWB are employees of Westat, Inc, which has received funding from the CDC for vaccine effectiveness studies. BW and MDR are employees of Novavax, Inc. MST has received funding from the American College of Occupational and Environmental Medicine and the National Institute for Occupational Safety and Health outside the scope of this work.

Multimedia Appendix 1

XSL•FO
RenderX

Recruitment and engagement materials.
[PDF File (Adobe PDF File), 1286 KB - resprot_v15i1e80858_app1.pdf]

Multimedia Appendix 2
Electronic interfaces.
[PDF File (Adobe PDF File), 1128 KB - resprot_v15i1e80858_app2.pdf]

References

1. Silk BJ, Scobie HM, Duck WM, Palmer T, Ahmad FB, Binder AM, et al. COVID-19 surveillance after expiration of the public health emergency declaration - United States, May 11, 2023. MMWR Morb Mortal Wkly Rep 2023 May 12;72(19):523-528 [FREE Full text] [doi: 10.15585/mmwr.mm7219e1] [Medline: 37167154]
2. Recommendation for the 2023-2024 formula of COVID-19 vaccines in the U.S. US Food and Drug Administration. URL: https://www.fda.gov/media/169591/download [accessed 2024-03-27]
3. Spikevax (COVID-19 vaccine, mRNA); 2024-2025 formula. Moderna, Inc. 2024 Sep. URL: https://static.modernatx.com/pm/6cef78f8-8dad-4fc9-83d5-d2fbb7cff867/5efa7d9d-05e8-46b5-945a-637c2867bd00/5efa7d9d-05e8-46b5-945a-637c2867bd00_viewable_rendition__v.pdf [accessed 2024-11-06]
4. Comirnaty (COVID-19 vaccine, mRNA); 2024-2025 formula. Pfizer–BioNTech. 2024 Oct. URL: https://labeling.pfizer.com/ShowLabeling.aspx?id=16351&format=pdf [accessed 2024-11-06]
5. Package Insert and Patient Package Insert - NUVAXOVID. US Food and Drug Administration. URL: https://www.fda.gov/media/186544/download?attachment [accessed 2025-05-29]
6. Reimer JM, Karlsson KH, Lövgren-Bengtsson K, Magnusson SE, Fuentes A, Stertman L. Matrix-M™ adjuvant induces local recruitment, activation and maturation of central immune cells in absence of antigen. PLoS One 2012 Jul 23;7(7):e41451 [FREE Full text] [doi: 10.1371/journal.pone.0041451] [Medline: 22844480]
7. Stertman L, Palm AE, Zarnegar B, Carow B, Lunderius Andersson C, Magnusson SE, et al. The Matrix-M™ adjuvant: a critical component of vaccines for the 21 century. Hum Vaccin Immunother 2023 Dec 31;19(1):2189885 [FREE Full text] [doi: 10.1080/21645515.2023.2189885] [Medline: 37113023]
8. Katella K. Comparing the COVID-19 vaccines: how are they different? Yale Medicine. 2025 May 20. URL: https://www.yalemedicine.org/news/covid-19-vaccine-comparison [accessed 2025-05-29]
9. Lin DY, Du Y, Xu Y, Paritala S, Donahue M, Maloney P. Effectiveness of XBB.1.5 vaccines against Omicron subvariants. Med Res Arch 2024;12(8):1-11 [FREE Full text] [doi: 10.18103/mra.v12i8.5740]
10. Link-Gelles R, Ciesla AA, Mak J, Miller JD, Silk BJ, Lambrou AS, et al. Early estimates of updated 2023-2024 (Monovalent XBB.1.5) COVID-19 vaccine effectiveness against symptomatic SARS-CoV-2 infection attributable to co-circulating Omicron variants among immunocompetent adults - increasing community access to testing program, United States, September 2023-January 2024. MMWR Morb Mortal Wkly Rep 2024 Feb 01;73(4):77-83 [FREE Full text] [doi: 10.15585/mmwr.mm7304a2] [Medline: 38300853]
11. The Efficacy of the 2023-2024 updated COVID-19 vaccines against COVID-19 infection (BEEHIVE). National Library of Medicine. 2023 Nov 28. URL: https://clinicaltrials.gov/study/NCT06065176 [accessed 2024-11-06]
12. Burns J, Rivers P, LeClair LB, Jovel KS, Rai RP, Lowe AA, et al. Pediatric Research Observing Trends and Exposures in COVID-19 Timelines (PROTECT): protocol for a multisite longitudinal cohort study. JMIR Res Protoc 2022 Jul 28;11(7):e37929 [FREE Full text] [doi: 10.2196/37929] [Medline: 35635842]
13. Edwards LJ, Fowlkes AL, Wesley MG, Kuntz JL, Odean MJ, Caban-Martinez AJ, et al. Research on the epidemiology of SARS-CoV-2 in essential response personnel (RECOVER): protocol for a multisite longitudinal cohort study. JMIR Res Protoc 2021 Dec 03;10(12):e31574 [FREE Full text] [doi: 10.2196/31574] [Medline: 34662287]
14. Grant L, Whitaker JA, Yoon SK, Lutrick K, Bhargava SP, Brown C, Randomized Assessment of Influenza Vaccine Efficacy Network (RAIVEN). Relative effectiveness and immunogenicity of quadrivalent recombinant influenza vaccine versus egg-based inactivated influenza vaccine among adults aged 18-64 years: results and experience from a randomized, double-blind trial. Open Forum Infect Dis 2024 Oct;11(10):ofae559. [doi: 10.1093/ofid/ofae559] [Medline: 39416990]
15. Mak J, Khan S, Britton A, Rose S, Gwynn L, Ellingson K, et al. Association of messenger RNA coronavirus disease 2019 (COVID-19) vaccination and reductions in post COVID conditions following severe acute respiratory syndrome coronavirus 2 infection in a US prospective cohort of essential workers. J Infect Dis 2025 Mar 17;231(3):665-676. [doi: 10.1093/infdis/jiae556] [Medline: 39531735]
16. Thompson MG, Burgess JL, Naleway AL, Tyner H, Yoon SK, Meece J, et al. Prevention and attenuation of COVID-19 with the BNT162b2 and mRNA-1273 vaccines. N Engl J Med 2021 Jul 22;385(4):320-329 [FREE Full text] [doi: 10.1056/NEJMoa2107058] [Medline: 34192428]
17. Gargano JW, Wallace M, Hadler SC, Langley G, Su JR, Oster ME, et al. Use of mRNA COVID-19 vaccine after reports of myocarditis among vaccine recipients: update from the advisory committee on immunization practices - United States, June 2021. MMWR Morb Mortal Wkly Rep 2021 Jul 09;70(27):977-982 [FREE Full text] [doi: 10.15585/mmwr.mm7027e2] [Medline: 34237049]

XSL•FO
RenderX

18. COVID-19 data - coronavirus dashboard. Utah Department of Health and Human Services. URL: https://coronavirus-dashboard.utah.gov/risk.html [accessed 2023-09-18]
19. Luisi N, Sullivan PS, Sanchez T, Bradley H, Fahimi M, Shioda K, et al. Use of COVIDTests.gov at-home test kits among adults in a national household probability sample - United States, 2022. MMWR Morb Mortal Wkly Rep 2023 Apr 21;72(16):445-449 [FREE Full text] [doi: 10.15585/mmwr.mm7216a6] [Medline: 37079516]
20. Gwak E, Choe SA, Bolormaa E, Choe YJ, Wang C, Fix J, et al. Relative effectiveness of homologous NVX-CoV2373 and BNT162b2 COVID-19 vaccinations in South Korea. Vaccine 2025 Jan 01;43(Pt 1):126503 [FREE Full text] [doi: 10.1016/j.vaccine.2024.126503] [Medline: 39522327]
21. Gwak E, Choe SA, Kim K, Bolormaa E, Fix J, Vadivale M, et al. Real-world effectiveness of NVX-CoV2373 and BNT162b2 mRNA COVID-19 vaccination in South Korea. Vaccine 2025 Aug 30;62:127496 [FREE Full text] [doi: 10.1016/j.vaccine.2025.127496] [Medline: 40683136]
22. Genomic epidemiology of SARS-CoV-2 with subsampling focused on North America since pandemic start. Nextstrain. 2025 May 11. URL: https://nextstrain.org/ncov/gisaid/north-america/all-time?dmax=2024-09-01&dmin=2023-09-22 [accessed 2025-05-12]
23. National Academies of Sciences, Engineering, and Medicine, Policy and Global Affairs, Committee on Women in Science, Engineering, and Medicine, Committee on Improving the Representation of Women and Underrepresented Minorities in Clinical Trials and Research. Why diverse representation in clinical research matters and the current state of representation within the clinical research ecosystem. In: Bibbins-Domingo K, Helman A, editors. Improving Representation in Clinical Trials and Research: Building Research Equity for Women and Underrepresented Groups. Washington, DC: National Academies Press; 2022:23-46.
24. van Marum RJ. Underrepresentation of the elderly in clinical trials, time for action. Br J Clin Pharmacol 2020 Oct 09;86(10):2014-2016 [FREE Full text] [doi: 10.1111/bcp.14539] [Medline: 32909294]
25. Groom HC, Biel FM, Crane B, Sun E, Georgescu JP, Weintraub ES, et al. Disparities in COVID-19 vaccination receipt by race, ethnicity, and social determinants of health among a large patient population in a network of community-based healthcare centers. Vaccine 2024 Oct 24;42(24):126288. [doi: 10.1016/j.vaccine.2024.126288] [Medline: 39241356]

Abbreviations

AE: adverse event
AESI: adverse event of special interest
BEEHIVE: Booster Epidemiological Evaluation of Health, Illness and Vaccine Efficacy
CLI: COVID-19–like illness
FDA: Food and Drug Administration
HIPAA: Health Insurance Portability and Accountability Act
KAP: knowledge, attitudes, and practices
mITT: modified intent-to-treat
mRNA: messenger RNA
PCC: post–COVID-19 condition
rVE: relative vaccine effectiveness
VE: vaccine effectiveness

Edited by J Sarvestan; submitted 18.Jul.2025; peer-reviewed by B Cowling, CA Guay, A Afzalian; comments to author 10.Sep.2025; revised version received 18.Dec.2025; accepted 19.Dec.2025; published 27.Jan.2026.

Please cite as:
Yoon SK, Ellsworth GL, Battan-Wraith S, Phillips AL, Fink RV, Griffin J, Rowley EAK, McKell J, Smith AS, Campbell R, Williams J, Ball SW, Zhao H, Warren B, Rousculp MD, Thiese MS
Real-World Effectiveness and Noninferiority Evaluation and Comparison of Messenger RNA–Based and Protein-Based COVID-19 Vaccines: Protocol for the BEEHIVE Randomized Study With a Hybrid Effectiveness Design
JMIR Res Protoc 2026;15:e80858
URL: https://www.researchprotocols.org/2026/1/e80858
doi:10.2196/80858
PMID:

©Sarang K Yoon, German L Ellsworth, Steph Battan-Wraith, Andrew L Phillips, Rebecca V Fink, Joshua Griffin, Elizabeth A K Rowley, Jacob McKell, Ashley S Smith, Riley Campbell, Jesse Williams, Sarah W Ball, Hongwei Zhao, Brandy Warren, Matthew D Rousculp, Matthew S Thiese. Originally published in JMIR Research Protocols (https://www.researchprotocols.org),

XSL•FO
RenderX

27.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Acupuncture Prescription Based on Thermal-Sensitized Acupoints for the Treatment of Knee Osteoarthritis: Protocol for a Randomized Controlled Pilot Trial

Xue-Zhou Wang[1,2*], MS; Tao-Tao Lv[1,3*], PhD; Kua-Yue Zhang[1,3], MS; Zhuo-Ya Chen[1,3], BA; Li-Na Qin[1,3], Prof Dr Med; Wei-Juan Gang[2], Prof Dr Med; Wei-Heng Chen[1,3], Prof Dr Med; Bao-Hong Mi[1,3], PhD

[1]The Third Affiliated Hospital of Beijing University of Chinese Medicine, Beijing, China

[2]Institute of Acupuncture and Moxibustion, China Academy of Chinese Medical Sciences, Baijing, China

[3]Engineering Research Center of Chinese Orthopaedics and Sports Rehabilitation Artificial Intelligent, Ministry of Education, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Bao-Hong Mi, PhD
The Third Affiliated Hospital of Beijing University of Chinese Medicine
51 Xiaoguan Street
Andingmenwai, Chaoyang District
Beijing, 100029
China
Phone: 86 17703359107
Email: mibaohong12587@163.com

Abstract

Background: Treatment options for knee osteoarthritis (KOA) remain limited. Although previous studies suggest that acupuncture can alleviate pain and functional impairment associated with KOA, there is significant variation in acupoint selection across studies and a lack of objective criteria to guide this choice. Infrared thermography (IRT) is recognized as a reliable method for identifying inflammatory regions in KOA. Our previous research revealed abnormal skin temperature at specific acupoints and surrounding areas in the lower limbs of patients with KOA, indicating potential treatment targets. However, there is still insufficient evidence to support their clinical application.

Objective: This pilot trial aims to assess the feasibility of conducting large-scale studies in the future and to provide preliminary evidence of the effects of acupuncture applied at IRT-guided thermal-sensitized acupoints (TAs) for the treatment of KOA.

Methods: This will be a randomized, controlled pilot trial with participant blinding. A total of 60 patients with KOA will be randomly allocated in a 1:1 ratio to either the TA group or the conventional acupoint group. Both groups will receive 3 acupuncture sessions per week for 4 weeks. In the TA group, acupoint selection will be based on points with abnormal skin temperature identified by IRT using real-time surface projection technology. In the conventional acupoint group, acupoint selection will follow the established prescription validated in previous studies. The primary outcome will be feasibility, assessed using a traffic light system to evaluate recruitment, retention, intervention completion, and completion of other outcome measures. Secondary outcomes will include the visual analog scale for pain, the Western Ontario and McMaster Universities Osteoarthritis Index function subscale, the credibility and expectancy questionnaire, skin temperature, and safety.

Results: The study has been approved by the ethics committee of the Third Affiliated Hospital of Beijing University of Chinese Medicine (BZYSY-2024KYKTPJ-44) and registered in the International Traditional Medicine Clinical Trial Registry. The first participant was enrolled on February 25, 2025, and a total of 30 participants had been enrolled by January 1, 2026. All enrollment and follow-up are expected to be completed by June 2026.

Conclusions: The findings of this study will provide evidence on the feasibility and preliminary clinical effects to support the future adoption of acupuncture prescriptions based on TAs for the treatment of KOA.

Trial Registration: International Traditional Medicine Clinical Trial Registry ITMCTR2025000019; https://itmctr.ccebtcm.org.cn/mgt/project/view/-5870917416012215224

International Registered Report Identifier (IRRID): DERR1-10.2196/81837

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e81837) doi:10.2196/81837

KEYWORDS

knee osteoarthritis; acupuncture; infrared thermography; acupoint; pilot

Introduction

Knee osteoarthritis (KOA) is the most prevalent form of osteoarthritis, primarily characterized by chronic pain and functional impairment [1]. Current treatment options remain limited. Although pharmacological therapies may provide short-term symptom relief, their long-term use is often constrained by potential cardiovascular and gastrointestinal risks [2]. Acupuncture is a widely used nonpharmacological intervention and has been proven to be superior to sham acupuncture in certain studies [3,4]. However, these benefits are not consistently observed, particularly in studies conducted outside China [5,6]. Consequently, the American College of Rheumatology and the American Academy of Orthopaedic Surgeons [7,8] have generally assigned acupuncture a limited strength of recommendation in clinical practice guidelines.

The effects of acupuncture can be attributed to both acupoint stimulation and patient expectations [9]. Acupoints are central to acupuncture treatment; however, there is considerable variation in acupoint selection across studies on KOA. Guided by traditional acupuncture theory, more than 40 acupoints have been considered for the treatment of KOA [10,11]. The lack of clear criteria for acupoint selection has raised concerns regarding the adequacy of acupuncture protocols, as most studies fail to provide objective justification for their choice of acupoints [12]. The importance of the many available acupoints in traditional acupuncture theory remains unclear, as meta-analyses have shown no correlation between the number of acupoints used and clinical effect [9]. Experimental studies further support this observation, as animal studies using only 1 to 2 fixed acupoints often conclude that acupuncture is effective [13,14]. One possible explanation is that the effects of acupuncture may not depend on specific acupoints per se but rather on broader subcutaneous tissue stimulation.

Inflammatory diseases are closely related to thermoregulation. The severity of KOA has been shown to correlate with the temperature of specific areas surrounding the knee joint [15,16]. When abnormal temperature changes occur at or near acupoints, it is referred to as acupoint sensitization. These abnormal skin temperature variations may be associated with autonomic nervous system function and metabolic activity, with certain acupoints exhibiting more pronounced temperature changes than nonacupoint regions [17]. Recent anatomical studies [18] have demonstrated a notable overlap between acupoints and perforating cutaneous blood vessels. Using infrared thermography (IRT), it is possible to detect temperature changes in these blood vessels beneath the acupoints [19]. Our previous research [20] indicates that, compared with individuals without KOA, patients with KOA exhibit thermal sensitization at certain acupoints of the lower limbs, showing characteristics of localized and regional thermal consistencies. In the treatment of peripheral facial paralysis, acupuncture intervention at these temperature-abnormal points is effective [21]. However, in the case of KOA, there is currently limited evidence supporting the effectiveness of this acupuncture approach.

Compared with conventional acupoint (CA) selection, thermal-sensitized acupoints (TAs) may offer improved therapeutic effects. Under IRT guidance, these acupoints can be flexibly selected according to individual differences, while the procedure remains easy to standardize. IRT is convenient and noninvasive, which may result in high acceptability without imposing additional burden on participants. Therefore, this study aims to conduct a pilot trial comparing acupuncture using IRT projection technology for acupoint selection with CA therapy for KOA. The goal is to explore the feasibility of conducting a large-scale clinical trial and to assess differences in effectiveness between the two acupoint selection methods in the treatment of KOA.

Methods

Study Design

This will be a single-center randomized controlled pilot trial conducted at the Third Affiliated Hospital of Beijing University of Chinese Medicine. Participants will be randomly assigned to the TA group and the CA group in a 1:1 ratio. This study will be reported according to the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines [22] (Multimedia Appendix 1).

Participants

Before any study-specific procedures are initiated, individuals expressing interest in participation will receive a comprehensive explanation of the study aims, procedures, and relevance. Information regarding potential benefits, risks, and discomforts will be communicated clearly to support informed decision-making. Participation is voluntary, and participants may withdraw at any time. Individuals who decline participation will be informed of available conventional treatment options upon request. Written informed consent will be obtained before eligibility screening.

As this study compared acupuncture applied at TAs with that applied at CAs, the inclusion and exclusion criteria were based on those used in our previous studies evaluating the effects of conventional acupuncture [23]. The inclusion criteria are as follows: (1) age 45 to 75 years, (2) meeting the American College of Rheumatology diagnostic criteria for KOA [24], (3) reporting knee pain for more than 3 months, (4) having undergone radiological examination within the past 6 months indicating a Kellgren-Lawrence grade II or III [25], (5) having a score ≥40 mm in the visual analog scale (VAS) [26] during the past week, and (6) being willing to sign the informed consent.

Patients will be excluded if they meet any of the following conditions: (1) a history of knee joint surgery or plans for knee joint surgery (eg, knee replacement or knee arthroscopy); (2)

XSL•FO
RenderX

knee pain attributed to other pathologies (eg, meniscus tear, rheumatoid arthritis, joint cavity infection, malignancy, gout, or lumbosacral disease with symptom of lower extremity); (3) a history of arthroscopy within the past 1 year or intra-articular injection within the past 6 months in the evaluating knee; (4) receipt of acupuncture in either knee within the past 6 months; (5) severe acute or chronic organic diseases or neuropsychiatric disorders; (6) disorders of coagulopathy (eg, hemophilia); (7) presence of a pacemaker or epilepsy; (8) pregnancy, planned pregnancy, or lactation; and (9) participation in other clinical studies within the past 1 month.

Randomization and Blinding

The randomization sequence will be generated by an independent statistician using Stata (version 17.0; StataCorp). To conceal the contents, opaque envelopes will be used for distribution. These envelopes will be stored and managed by independent researchers not involved in other aspects of the trial. After including participants who met the criteria and obtaining their informed consent forms, random envelopes will be opened sequentially. The acupuncturists will be aware of group assignments, whereas participants and outcome evaluators will remain blinded.

Interventions

Overview

The difference between the two intervention groups will be limited to the selection of acupoints only. All participants will undergo 4 weeks of acupuncture intervention accompanied by electrical stimulation. The treatment will be administered 3 times per week, with each session lasting 30 minutes. Ideally, sessions will be scheduled every other day, and administration of 2 sessions within a 24-hour period will be prohibited. Over a period of 4 weeks, a total of 12 treatments will be provided. Participants will assume a supine position on the treatment bed, and sterile, single-use needles will be used.

After the needle is inserted, the acupuncturists will perform lifting-insertion and twisting-rotating operations for at least 10 seconds to achieve the *deqi* sensation. The electroacupuncture instrument will then be connected, and a sparse-dense wave will be applied. Current intensity will be adjusted according to the participant's comfort level. The intervention will be delivered by 3 acupuncturists with more than 3 years of clinical experience.

During the trial, participants will be discouraged from undergoing any other treatments that might affect their symptoms, including medications (such as nonsteroidal anti-inflammatory drugs, opioids, etc) and other physical therapies. Participants may use diclofenac sodium enteric–coated tablets (Voltaren, Beijing Novartis Pharma Co., Ltd) for severe pain that cannot be relieved, except within 48 hours before any assessment. If rescue medication is required, participants will receive up to 6 tablets, take 1 tablet 3 times daily, and discontinue use once symptoms improve, returning any remaining tablets to the study staff. Any treatment other than the trial intervention will be recorded, including timing and dose.

TA Group

Acupoint selection in the TA group is based on the points with the highest skin temperature in 8 specific lower limb regions (Figure 1). Our previous studies have indicated that, in patients with KOA, regional temperature increases occur in the lower limb acupoint areas as shown in Figure 1. These increases present a diffuse pattern resembling a "surface spread," suggesting the presence of inflammatory reactions near the acupoint regions [20]. The areas with the highest temperatures are likely to correspond to locations of the most intense inflammatory responses. Therefore, these temperature-abnormal points will be selected as intervention acupoints in the TA group.

Figure 1. The 8 regions on the lower limbs used for acupoint selection.

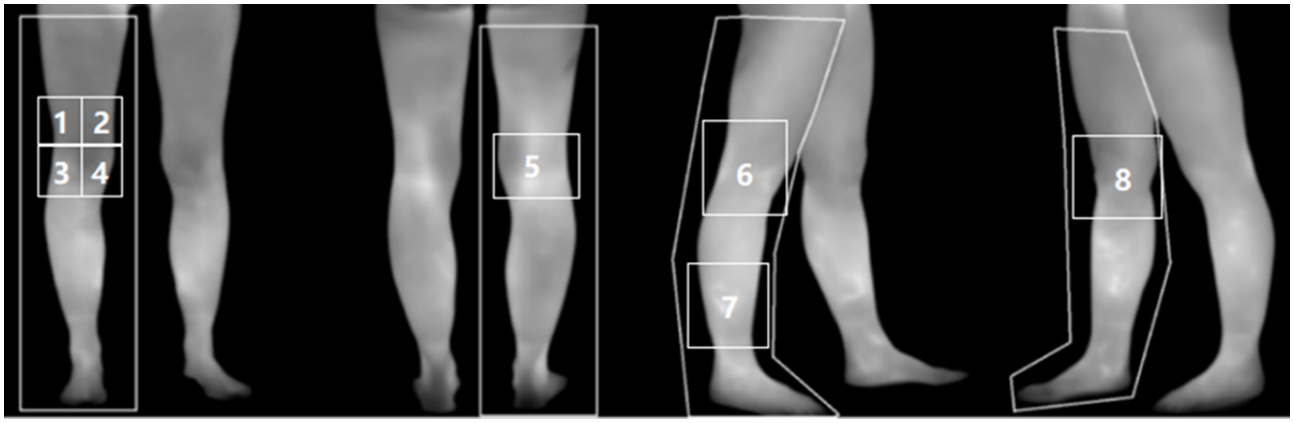

Acupoint selection will be performed using a system established by the research team, consisting of a custom-designed IRT image projection device (patent authorization number CN214748459U) and an auxiliary acupoint selection software system (patent authorization number CN114582489B). This system can project infrared images onto the body surface in real time (Figures 2A and 2B), monitor temperature extremes within the region, and visually present the data (Figure 2C).

XSL•FO
RenderX

Figure 2. Infrared thermographic projection system for acupoint selection. (A) Infrared thermographic projection device; (B) real-time infrared image projected onto the body surface; (C) skin temperature recognition system. The squares indicate the automatically identified points with the highest skin temperature within the region.

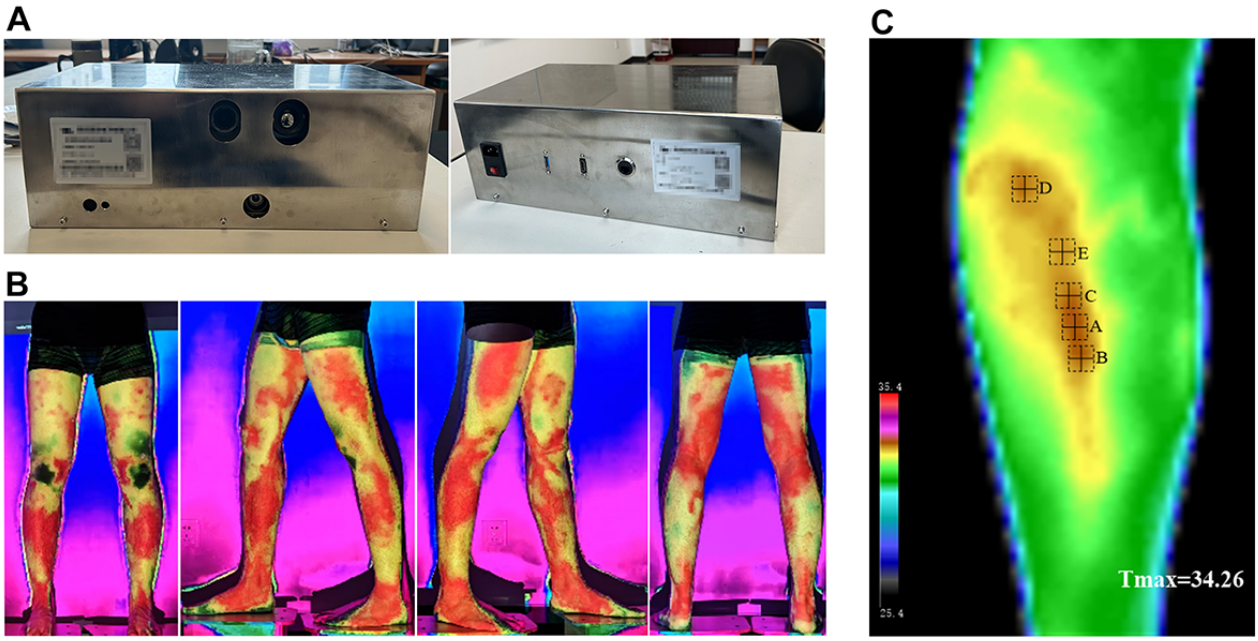

The parameters of the IRT device are as follows: (1) a resolution of 640×512, (2) a temperature sensitivity of 0.05 ˚C, (3) a field of view of 33.5°×26.9°, and (4) manufactured by Airui Optoelectronics Technology Co Ltd. The projector parameters are as follows: (1) a luminance of 1400 ANSI lumens, (2) a lens resolution of 1920×1080 (4K resolution), (3) the HK-610 model, and (4) an IRT image projection distance of 1 to 2 meters.

The background filtering method adopted is a bimodal filter. This method iteratively and statistically analyzes the 2D data of the image, calculates the distribution frequency curve for any temperature value, and applies a Gaussian smoothing filter. The image filtering threshold is automatically determined by locating the 2 peaks of the distribution curve, thereby filtering the raw image data and retaining only the target region of interest for the study. The maximum regional temperature value is calculated using an improved maximum value method.

Participants in the TA group will undergo IRT inspection before the first intervention session to determine the initial intervention points. Thereafter, IRT inspection will be repeated before the first session of each subsequent treatment week to update the acupoints accordingly. Therapists will record the treatment sites using photographs to ensure consistency within each treatment cycle.

CA Group

The CA group acupoint selection adopts a previously validated acupuncture intervention protocol [3]. It includes 5 fixed acupoints as follows: Dubi (ST35), Neixiyan (EX-LE5), Ququan (LR8), Xiyangguan (GB33), and an Ashi point (the point at which the patient experiences the most pain). Additionally, 3 optional acupoints will be selected based on the acupuncturist's diagnosis from a pool of 22 possible acupoints. Detailed acupoint selection information is provided in Table S1 in Multimedia Appendix 2.

Participants in the CA group will undergo IRT inspection before the first intervention session each week, consistent with the procedure used in the TA group. However, these assessments will be intended solely for blinding purposes and will not inform the selection of acupuncture points.

Outcomes

Primary Outcome

The primary outcome will be feasibility. Feasibility will be assessed throughout the trial period, including recruitment rate, retention rate, intervention adherence, and outcome assessment. A traffic light system will be used to evaluate feasibility (Table 1). Outcomes exceeding the upper threshold will be considered feasible; those falling between the upper and lower thresholds will be interpreted as having limited feasibility, and protocol modifications for each outcome will be considered in future studies to enhance feasibility to an acceptable level. If any outcome falls below the lower threshold, the trial will be considered infeasible. In such cases, the results will primarily be used to analyze the reasons for trial failure.

XSL•FO
RenderX

Table 1. Feasibility assessment using a traffic light system.

Dimension	Lower threshold (%)	Upper threshold (%)
Recruitment rate	30	60
Retention rate[a]	70	80
Intervention completion rate[b]	70	90
Outcome completion rate		
VAS[c] score	60	70
WOMAC[d] function	60	70
Credibility and expectancy questionnaire	60	70
Skin temperature	60	70

[a]Including follow-up.

[b]Defined as completing at least 80% (10 sessions) of the intervention with no major protocol deviations.

[c]VAS: visual analog scale.

[d]WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

Secondary Outcomes

Pain

Knee pain will be assessed using the VAS, a 100-mm scale in which one end represents “no pain” and the other represents “the worst pain imaginable” [26]. The VAS will be assessed at baseline and at weeks 2, 4, 8, and 12.

Function

Functional impairment will be evaluated using the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) function subscale [27], which consists of 17 items rated on a 0 to 4 scale, with total scores ranging from 0 to 68. Higher scores indicate worse knee function. Assessments will be conducted at baseline and at weeks 2, 4, 8, and 12.

Expectancy

Expectancy will be measured using the credibility and expectancy questionnaire [28], which consists of 6 items. The first 3 items assess the credibility of the intervention, and the remaining 3 assess participants’ expectancy regarding the subjective effect. The questionnaire will be administered within 5 minutes after the first treatment session. Standardized scoring will be applied by adjusting each item score to have a mean of 0 and an SD of 1. The credibility score will be calculated as the sum of the standardized scores of the first 3 items, and the expectancy score, as the sum of the last 3 items.

Skin Temperature

Skin temperature at acupoints on the lower limb region will be measured using a portable medical IRT device. The measurement method will be the same as that used in our previous study [29] (Multimedia Appendix 2).

Safety

The safety of both interventions will be evaluated based on adverse events. Any adverse events will be documented within 24 hours of occurrence, and their relationship to the intervention will be determined by the research team. When necessary, the researcher responsible for maintaining the randomization sequence will perform unblinding. Serious adverse events will be referred to the principal investigator for further evaluation and management. Common adverse events related to acupuncture include pain, local hematoma, infection, and dizziness.

Feasibility, skin temperature, and safety outcomes will be assessed and documented by investigators. For participant-reported outcomes, investigators will first provide standardized explanations of questionnaire items, after which participants will complete the questionnaires independently, with minimal interaction to reduce potential influence. During the follow-up period, the questionnaire assessments will be conducted by telephone to improve retention.

Sample Size

As this is a pilot study without previous evidence to estimate the difference between the current two groups, we selected a sample size of 30 participants per group based on the CI upper limit method [30]. Simulation studies suggest that this sample size provides stable estimates of variability, resulting in an 80% to 90% likelihood that a subsequent definitive trial will reach adequate power [31]. Assessment of treatment effects will be exploratory and intended to inform a future confirmatory trial, with CI bounds used because point estimates have limited precision in small samples [32]. In total, 60 patients will be recruited.

Statistical Analysis

All statistical analyses will be performed using SPSS (version 20.0; IBM Corp). Multiple imputation will be used to address missing outcome data, with each missing value imputed 5 times and the mean of these imputations used for analysis. The intention-to-treat dataset will be used as the primary dataset. Feasibility outcomes will be primarily reported as observed proportions. The 95% CIs for feasibility outcomes will be calculated using Wilson score intervals and presented as a sensitivity analysis. Between-group differences in VAS scores, the WOMAC functional subscale, and credibility and expectancy scores will be assessed using either 2-tailed independent-sample

XSL•FO
RenderX

t tests or Mann-Whitney *U* tests, as appropriate. Linear regression analyses will be conducted to examine the associations between within-group changes in VAS scores and skin temperature and their corresponding baseline values. Assessment of treatment effects will be exploratory. Results from the VAS and WOMAC will be used to estimate the effect size range for calculating the sample size of a definitive trial. Effect size will be quantified using Cohen *d*, with values of approximately 0.2, 0.5, and 0.8 representing small, medium, and large effects, respectively. All statistical tests will be 2-sided, with a significance level set at *P*=.05.

Ethical Considerations

The study protocol was approved by the research ethics committee of the Third Affiliated Hospital of Beijing University of Chinese Medicine (BZYSY-2024KYKTPJ-44) and registered in the International Traditional Medicine Clinical Trial Registry (ITMCTR2025000019). Recruitment information will be disseminated through the acupuncture department and online media platforms. Participants will provide written informed consent. Original data will be stored using paper case report forms. Deidentified data will be used for storage and analysis.

Results

Figure 3 shows the study flow diagram. The study was approved by the research ethics committee on November 15, 2024, and registered in the trial registry on January 6, 2025. Major funding was obtained on January 1, 2025. As the first participant was enrolled on February 25, 2025, a total of 30 participants had been enrolled by January 1, 2026. All enrollment and follow-up are expected to be completed by June 2026. Any protocol amendments will be documented and reported. The study results will be published in a peer-reviewed journal.

Figure 3. Flowchart of the study. CA: conventional acupoint; TA: thermal-sensitized acupoint.

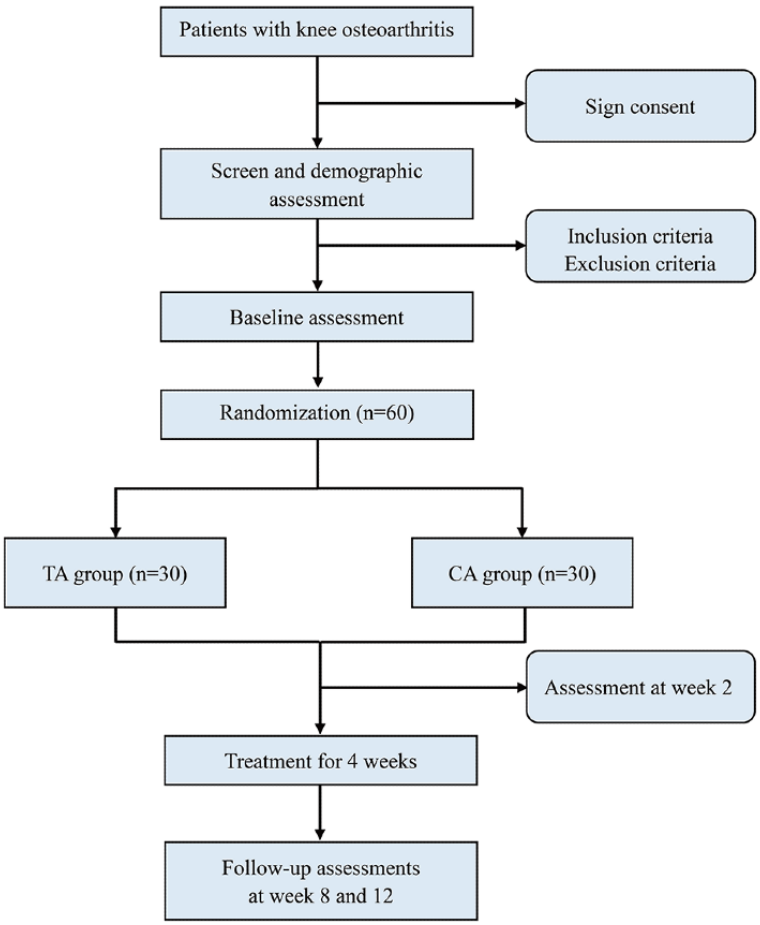

Discussion

Anticipated Findings

This study focuses on the treatment sites in acupuncture therapy, which were relatively underexplored in previous studies. A significant issue in acupuncture research is the substantial variability in acupoint selection, which often lacks objective justification. One reason for the limited acceptability of acupuncture is the reliance on acupoints derived from traditional theories, for which the biological rationale remains unclear [33]. Although the use of *ashi* points in traditional acupuncture or dry needling is supported by some biological rationale, these points are primarily chosen based on the patient's subjective pain perception.

IRT can accurately identify areas of inflammation in the knee joint, and acupuncture guided by IRT offers a more objective approach to determining treatment sites. Compared with other acupuncture studies using TAs [34,35], a potential advantage of this study lies in individualized treatment guided by objective IRT measurements. Rather than directly adopting the TAs

XSL•FO
RenderX

suggested in previous studies, treatment sites will be selected in a targeted manner according to individual differences. This acupuncture method, based on a simple imaging technique, could enhance the acceptability and dissemination of acupuncture in nontraditional Chinese medicine health care settings.

A limitation of this study is the lack of sufficient preliminary evidence to calculate the sample size, which means that the evaluation of efficacy differences between the two acupuncture therapies is a secondary aim. As a pilot trial, the study uses feasibility as the primary outcome. This primary outcome will be used to assess the multidimensional feasibility of conducting a large-scale study. Clinical effect and safety will be secondary outcomes, providing preliminary evidence. Another limitation is that, as a single-center study, patient characteristics may be limited.

Conclusions

The findings of this study will provide feasibility data and preliminary evidence of clinical effect to support future clinical adoption of acupuncture prescription based on TAs for the treatment of KOA.

Funding

This study was funded by the Fundamental Research Funds for the Central Universities (2024-JYB-XJSJJ-032 and 2025-JYB-JBGS-017) and the National Natural Science Foundation of China (82104693). The funding organizations had no involvement in the study design, study conduct, data analysis, or reporting of the trial.

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Authors' Contributions

BHM and WHC conceived the study. XZW and TTL initiated the study design. KYZ, ZYC, LNQ, and WJG contributed to methodological improvements of the protocol. XZW and TTL drafted the manuscript, and all the authors critically revised it. BHM secured funding. KYZ and ZYC obtained ethics approval. All authors contributed to the refinement of the study protocol and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines checklist.
[DOCX File , 35 KB - resprot_v15i1e81837_app1.docx]

Multimedia Appendix 2
Information on acupoint selection methods.
[DOCX File , 34 KB - resprot_v15i1e81837_app2.docx]

References

1. Sharma L. Osteoarthritis of the knee. N Engl J Med 2021 Jan 07;384(1):51-59. [doi: 10.1056/NEJMcp1903768] [Medline: 33406330]
2. Bennell KL, Hunter DJ, Hinman RS. Management of osteoarthritis of the knee. BMJ 2012 Jul 30;345:e4934. [doi: 10.1136/bmj.e4934] [Medline: 22846469]
3. Tu JF, Yang JW, Shi GX, Yu ZS, Li JL, Lin LL, et al. Efficacy of intensive acupuncture versus sham acupuncture in knee osteoarthritis: a randomized controlled trial. Arthritis Rheumatol 2021 Mar;73(3):448-458. [doi: 10.1002/art.41584] [Medline: 33174383]
4. Liu CY, Duan YS, Zhou H, Wang Y, Tu JF, Bao X, et al. Clinical effect and contributing factors of acupuncture for knee osteoarthritis: a systematic review and pairwise and exploratory network meta-analysis. BMJ Evid Based Med 2024 Nov 22;29(6):374-384 [FREE Full text] [doi: 10.1136/bmjebm-2023-112626] [Medline: 39486882]
5. Hinman RS, McCrory P, Pirotta M, Relf I, Forbes A, Crossley KM, et al. Acupuncture for chronic knee pain: a randomized clinical trial. JAMA 2014 Oct 01;312(13):1313-1322. [doi: 10.1001/jama.2014.12660] [Medline: 25268438]
6. Scharf HP, Mansmann U, Streitberger K, Witte S, Krämer J, Maier C, et al. Acupuncture and knee osteoarthritis: a three-armed randomized trial. Ann Intern Med 2006 Jul 04;145(1):12-20. [doi: 10.7326/0003-4819-145-1-200607040-00005] [Medline: 16818924]

XSL•FO
RenderX

7. Kolasinski SL, Neogi T, Hochberg MC, Oatis C, Guyatt G, Block J, et al. 2019 American College of Rheumatology/Arthritis Foundation guideline for the management of osteoarthritis of the hand, hip, and knee. Arthritis Rheumatol 2020 Feb;72(2):220-233 [FREE Full text] [doi: 10.1002/art.41142] [Medline: 31908163]
8. Brophy RH, Fillingham YA. AAOS clinical practice guideline summary: management of osteoarthritis of the knee (nonarthroplasty), third edition. J Am Acad Orthop Surg 2022 May 01;30(9):e721-e729. [doi: 10.5435/JAAOS-D-21-01233] [Medline: 35383651]
9. Gang WJ, Xiu WC, Shi LJ, Zhou Q, Jiao RM, Yang JW, FAMOUS Group. Factors Associated with the Magnitude Of acUpuncture treatment effectS (FAMOUS): a meta-epidemiological study of acupuncture randomised controlled trials. BMJ Open 2022 Aug 29;12(8):e060237 [FREE Full text] [doi: 10.1136/bmjopen-2021-060237] [Medline: 36038176]
10. Sun N, Wang LQ, Shao JK, Zhang N, Zhou P, Fang SN, et al. An expert consensus to standardize acupuncture treatment for knee osteoarthritis. Acupunct Med 2020 Oct 20;38(5):327-334. [doi: 10.1177/0964528419900789] [Medline: 32309995]
11. Gong Z, Liu R, Yu W, Wong TK, Guo Y, Sun Y. Acutherapy for knee osteoarthritis relief in the elderly: a systematic review and meta-analysis. Evid Based Complement Alternat Med 2019;2019:1868107 [FREE Full text] [doi: 10.1155/2019/1868107] [Medline: 30906410]
12. Shi LJ, Tian ZY, Hu XY, Xiu WC, Jiao RM, Hu XY, et al. Approach to assess adequacy of acupuncture in randomized controlled trials: a systematic review. Chin J Integr Med 2023 Aug;29(8):730-737. [doi: 10.1007/s11655-023-3691-3] [Medline: 36790553]
13. Tan Q, Cai Z, Li J, Li J, Xiang H, Li B, et al. Imaging study on acupuncture inhibiting inflammation and bone destruction in knee osteoarthritis induced by monosodium iodoacetate in Rat model. J Pain Res 2022;15:93-103 [FREE Full text] [doi: 10.2147/JPR.S346242] [Medline: 35046719]
14. Xu W, Xiao Y, Zhao M, Zhu J, Wang Y, Wang W, et al. Effective treatment of knee osteoarthritis using a nano-enabled drug acupuncture technology in mice. Adv Sci (Weinh) 2023 Oct;10(28):e2302586 [FREE Full text] [doi: 10.1002/advs.202302586] [Medline: 37555294]
15. Warashina H, Hasegawa Y, Tsuchiya H, Kitamura S, Yamauchi K, Torii Y, et al. Clinical, radiographic, and thermographic assessment of osteoarthritis in the knee joints. Ann Rheum Dis 2002 Sep;61(9):852-854 [FREE Full text] [doi: 10.1136/ard.61.9.852] [Medline: 12176819]
16. De Marziani L, Boffa A, Angelelli L, Andriolo L, Di Martino A, Zaffagnini S, et al. Infrared thermography in symptomatic knee osteoarthritis: joint temperature differs based on patient and pain characteristics. J Clin Med 2023 Mar 16;12(6):2319 [FREE Full text] [doi: 10.3390/jcm12062319] [Medline: 36983319]
17. Sun M, Yang M, Rong J, Ma X, Zheng H, Cai D, et al. Trigger points and sensitized acupoints: same book, different covers? Acupunct Herb Med 2021;1(2):74-80. [doi: 10.1097/HM9.0000000000000018] [Medline: 2233705]
18. Meng S, Tang Z, Zhu M, Liang X, Ding Z, Wang Z, et al. Perforating cutaneous vessels: a key feature of acupoints - anatomical evidence from five-Shu acupoints in the upper limbs. Clin Anat 2024 Jan;37(1):33-42. [doi: 10.1002/ca.24077] [Medline: 37340879]
19. Álvarez-Prats D, Carvajal-Fernández O, Valera Garrido F, Pecos-Martín D, García-Godino A, Santafe MM, et al. Acupuncture points and perforating cutaneous vessels identified using infrared thermography: a cross-sectional pilot study. Evid Based Complement Alternat Med 2019;2019:7126439 [FREE Full text] [doi: 10.1155/2019/7126439] [Medline: 31015854]
20. Tu JF, Wang XZ, Yan SY, Wang YR, Yang JW, Shi GX, et al. Thermal sensitization of acupoints in patients with knee osteoarthritis: a cross-sectional case-control study. J Integr Med 2025 May;23(3):289-296 [FREE Full text] [doi: 10.1016/j.joim.2025.03.004] [Medline: 40210574]
21. Zhang D. A method of selecting acupoints for acupuncture treatment of peripheral facial paralysis by thermography. Am J Chin Med 2007;35(6):967-975. [doi: 10.1142/S0192415X07005430] [Medline: 18186583]
22. Chan AW, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ 2025 Apr 28;389:e081477. [doi: 10.1136/bmj-2024-081477] [Medline: 40294953]
23. Wang XZ, Wang RK, Liu Q, Shi GX, Mi BH, Liu CZ, et al. Electroacupuncture versus exercise in patients with knee osteoarthritis: study protocol for a randomized controlled trial. PLoS One 2024;19(6):e0305105 [FREE Full text] [doi: 10.1371/journal.pone.0305105] [Medline: 38861546]
24. Hochberg MC, Altman RD, Brandt KD, Clark BM, Dieppe PA, Griffin MR, et al. Guidelines for the medical management of osteoarthritis. Part II. Osteoarthritis of the knee. American College of Rheumatology. Arthritis Rheum 1995 Nov 09;38(11):1541-1546. [doi: 10.1002/art.1780381104] [Medline: 7488273]
25. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957 Dec;16(4):494-502 [FREE Full text] [doi: 10.1136/ard.16.4.494] [Medline: 13498604]
26. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, European Palliative Care Research Collaborative (EPCRC). Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. J Pain Symptom Manage 2011 Jun;41(6):1073-1093 [FREE Full text] [doi: 10.1016/j.jpainsymman.2010.08.016] [Medline: 21621130]
27. Xie F, Li SC, Goeree R, Tarride JE, O'Reilly D, Lo N, et al. Validation of Chinese Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) in patients scheduled for total knee replacement. Qual Life Res 2008 May;17(4):595-601. [doi: 10.1007/s11136-008-9340-7] [Medline: 18415706]

XSL•FO
RenderX

28. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000 Jun;31(2):73-86. [doi: 10.1016/s0005-7916(00)00012-4] [Medline: 11132119]
29. Mi BH, Wang XZ, Yang JW, Shi GX, Zhang WZ, Jin LN, et al. Thermographic evaluation of acupoints in lower limb region of individuals with osteoarthritis: a cross-sectional case-control study protocol. PLoS One 2023;18(4):e0284381 [FREE Full text] [doi: 10.1371/journal.pone.0284381] [Medline: 37058539]
30. Browne RH. On the use of a pilot sample for sample size determination. Stat Med 1995 Sep 15;14(17):1933-1940. [doi: 10.1002/sim.4780141709] [Medline: 8532986]
31. Ying X, Freedland KE, Powell LH, Stuart EA, Ehrhardt S, Mayo-Wilson E. Determining sample size for pilot trials: a tutorial. BMJ 2025 Aug 08;390:e083405. [doi: 10.1136/bmj-2024-083405] [Medline: 40780848]
32. Ying X, Robinson KA, Ehrhardt S. Re-evaluating the role of pilot trials in informing effect and sample size estimates for full-scale trials: a meta-epidemiological study. BMJ Evid Based Med 2023 Nov 22;28(6):383-391. [doi: 10.1136/bmjebm-2023-112358] [Medline: 37491141]
33. Li F, He T, Xu Q, Lin L, Li H, Liu Y, et al. What is the Acupoint? A preliminary review of Acupoints. Pain Med 2015 Oct;16(10):1905-1915. [doi: 10.1111/pme.12761] [Medline: 25975413]
34. Wu Z, Hu Z, Xu Z, Xiao T, Huang Q, Zhou X, et al. Efficacy and safety of acupuncture at sensitized acupoints for knee osteoarthritis: protocol for a multicenter, single-blind randomized controlled trial. JMIR Res Protoc 2025 Sep 24;14:e77336 [FREE Full text] [doi: 10.2196/77336] [Medline: 40990366]
35. Liu J, Li Y, Li L, Luo X, Li N, Yang X, et al. Effects of acupuncture at acupoints with lower versus higher pain threshold for knee osteoarthritis: a multicenter randomized controlled trial. Chin Med 2022 Jun 08;17(1):67 [FREE Full text] [doi: 10.1186/s13020-022-00626-3] [Medline: 35676697]

Abbreviations

CA: conventional acupoint
IRT: infrared thermography
KOA: knee osteoarthritis
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TA: thermal-sensitized acupoint
VAS: visual analog scale
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

Edited by J Sarvestan; submitted 05.Aug.2025; peer-reviewed by JF Tu; comments to author 28.Nov.2025; revised version received 28.Dec.2025; accepted 29.Dec.2025; published 28.Jan.2026.

Please cite as:
Wang XZ, Lv TT, Zhang KY, Chen ZY, Qin LN, Gang WJ, Chen WH, Mi BH
Acupuncture Prescription Based on Thermal-Sensitized Acupoints for the Treatment of Knee Osteoarthritis: Protocol for a Randomized Controlled Pilot Trial
JMIR Res Protoc 2026;15:e81837
URL: https://www.researchprotocols.org/2026/1/e81837
doi:10.2196/81837
PMID:

©Xue-Zhou Wang, Tao-Tao Lv, Kua-Yue Zhang, Zhuo-Ya Chen, Li-Na Qin, Wei-Juan Gang, Wei-Heng Chen, Bao-Hong Mi. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 28.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Home-Based Intervention to Improve Adherence to the 24-Hour Movement Guidelines in Young Children: Protocol for a Mobile App–Based Randomized Control Trial

Chelsea L Kracht[1], PhD; Jerica M Berge[2], PhD; Monique LeBlanc[3], PhD; Robert L Newton Jr[4], PhD; Ryan E Rhodes[5], PhD; Grace Bolamperti[1], BA; Madigan Snodgrass[1], BA; Leanne M Redman[6], PhD

[1]Department of Internal Medicine, University of Kansas Medical Center, Kansas City, KS, United States
[2]Department of Family Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States
[3]Department of Psychology, Southeastern Louisiana University, Hammond, LA, United States
[4]Population and Public Health Division, Pennington Biomedical Research Center, Baton Rouge, LA, United States
[5]School of Exercise Science, Physical and Health Education, University of Victoria, Victoria, BC, Canada
[6]Clinical Sciences Division, Pennington Biomedical Research Center, Baton Rouge, LA, United States

Corresponding Author:
Chelsea L Kracht, PhD
Department of Internal Medicine
University of Kansas Medical Center
3901 Rainbow Blvd
Kansas City, KS, 66160
United States
Phone: 1 9135881655
Email: ckracht@kumc.edu

Abstract

Background: One in 10 preschoolers (aged 3-4 y) meet the three 24-hour Movement Guidelines, that is, (1) physical activity, (2) sedentary screen time, and (3) sleep.

Objective: The overarching aim of this study is to evaluate the effectiveness and feasibility of a 12-week mobile health home-based intervention on 24-hour movement behaviors in preschoolers who meet few guidelines (zero or 1 guideline).

Methods: We will conduct a 12-week randomized controlled trial with a wait-list control in 80 families (40 per arm). Preliminary studies in this population informed intervention app content, features, and app development. Behavior change theories, including transfer theory and the multi-process action control framework, helped inform content presentation and topics. Primary outcomes include device-based and parent-report measures of 24-hour movement behaviors, and the secondary outcome is the feasibility and acceptability of the app. Exploratory outcomes include preschoolers' cognitive and motor skills, changes within the home environment, and behavioral control processes.

Results: This 2-phase study (K99/R00) received initial funding in March 2022, and preliminary studies were concluded in December 2023. The main grant received institutional review board approval in April 2024, and the grant funding began in May 2024. The study was registered in Clinical Trials in October 2024 and enrolled its first participant in January 2025. As of October 2025, the study has enrolled 39 families. We anticipate the trial will be completed in late 2026.

Conclusions: This research is designed to test a novel approach to improve all three 24-hour movement behaviors in preschoolers in home settings by using a mobile app. Results from this study will have implications for future 24-hour movement interventions, our understanding of improving all 3 behaviors, and ultimately, improvements in preschoolers' health.

Trial Registration: Clinicaltrials.gov NCT06667661; https://clinicaltrials.gov/study/NCT06667661

International Registered Report Identifier (IRRID): DERR1-10.2196/75621

(JMIR Res Protoc 2026;15:e75621) doi:10.2196/75621

KEYWORDS

screen time; sitting time; movement behavior; sleep; physical activity; adherence

XSL•FO
RenderX

Introduction

In 2019, the World Health Organization created the 24-hour Movement Guidelines [1], which provide recommendations for preschoolers' (aged 3-4 y) physical activity (PA; ≥3 h/d of total PA, including 1 hour of moderate-to-vigorous physical activity [MVPA]), sleep (10-13 h/d), and sedentary behavior (≤1 h/d of sedentary screen time) to promote physical and mental development [2-6]. Only 15% of preschoolers meet all 3 recommendations, and thus many preschoolers are at risk for compromised development and health [7,8]. For example, in a prospective cohort study, preschoolers who met zero or 1 guideline had lower motor skills 1 year later compared with those who met all 3 guidelines [5]. These alarming rates of poor adherence reveal an urgent need for effective interventions that support parents in promoting adherence to all 3 recommendations and thus prevent adverse health outcomes over the long term.

Current interventions overlook the importance of the home and family setting that influences when, how, and why all three 24-hour movement behaviors occur. To the authors' knowledge, there is one 24-hour movement intervention that primarily focused on the parent, rather than the home environment, and demonstrated limited effectiveness [9]. The home environment has a significant contextual influence on child's 24-hour movement behaviors, as illustrated by poor adherence rates to the 24-hour Movement Guidelines during the COVID-19 pandemic [10]. Facilitating changes to multiple movement behaviors within the home environment can rely on behavior change theory, and gaps between intention and behavior often prevent behavior change [11]. To change behaviors, attitudes, habits, and behavioral regulation patterns must be formed [12], including action plans (eg, planning and enlisting support) [13]. One framework that supports motivational (eg, attitude and perceptions of control), self-regulatory (eg, action plans), and habit theory is the multi-process action control (M-PAC) framework. The M-PAC framework combines aspects of behavior change together to form and translate intentions into long-term behavior change [14,15]. The M-PAC model can be applied to PA approaches. For example, family planning for PA can result in changes to family PA [16]. Bridging the intention-to-behavior gap in child movement behaviors also provides an unprecedented opportunity to leverage the transfer theory of behavior change, whereby changes in one behavior lead to changes in another behavior [17,18]. Habit formation, M-PAC, and transfer theory can be used in a 24-hour movement behavior intervention to change multiple behaviors in young children for long-term adherence to all 3 recommendations.

A mobile health (mHealth) modality (eg, mobile apps) can address behavioral patterns by delivering context-based information to parents. Parent-based mobile apps have been effective in reducing preschooler sugar-sweetened beverage consumption [19] but have yet to be translated to improving 24-hour movement behaviors. This paucity is demonstrated in a recent systematic review of digital health interventions for preschool PA, where the authors found 8 interventions in total, and only 1 was a parent-based mobile health app [20]. This mobile health app demonstrated effectiveness on child health behaviors, including PA, suggesting these modalities can be used to reach parents and impact child health behaviors [21]. Together, these modalities provide an opportunity to build upon previous work and change movement behaviors, especially for those at risk for poor adherence to the 24-hour Movement Guidelines.

The overarching aim is to conduct a study to evaluate the effectiveness and feasibility of a 12-week mHealth home-based intervention on 24-hour movement behaviors in preschoolers who meet few guidelines (zero or 1 guideline). Primary outcomes include device-based and self-report measures of 24-hour movement behaviors, and the secondary outcome is the feasibility and acceptability of the app. Exploratory outcomes include preschoolers' cognitive and motor skills, changes within the home environment, and behavioral control processes. We hypothesize that children will increase the number of recommendations obtained, with 35% of preschoolers meeting 24-hour Movement Guidelines (all 3 recommendations), based on other routine-focused interventions [22,23]. We also hypothesize parents will report the intervention is feasible (≥4.0/5.0 on Likert scale).

Methods

Study Design

This study, *Shining Star*, is a randomized controlled trial with a waitlist control design. This protocol follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) reporting guidelines for randomized trials (Multimedia Appendix 1).

Participants, Eligibility, and Recruitment

To be included, parents must report that their preschooler meets zero or 1 guideline for PA, sleep, and screen time as defined by the World Health Organization at the screening call [1]. Zero or 1 guideline was chosen as existing prospective studies demonstrate meeting 1 or zero guidelines is associated with lower motor skills and cognitive development compared with meeting all 3 guidelines [5,24]. Based on work conducted before the COVID-19 pandemic, we anticipated 15%-20% of preschoolers would qualify [5], with similar proportions found in a preschooler-focused systematic review [4]. Adherence to the 24-hour Movement Guidelines by parent report was chosen to reduce the burden, align with a survey study that informed analysis [25], and ensure the inclusion of the target population. The PA guideline is defined as ≥180 min/d of total PA, including ≥60 min/d of MVPA. The sleep guideline is defined as 10-13 h/d. The sedentary behavior guideline is defined ≤1 h/d of sedentary screen time.

Parents must also agree to use their mobile phone for the study duration (approximately 12 weeks), download the mobile app onto their phone, be fluent in English, and be willing to travel to the research site to complete study visits. Exclusion criteria are mobility limitations as reported by the parent, meeting 2 or 3 of the 24-hour movement guidelines, sibling or household member participating in the study, and parents reporting the preschooler will not wear a device on their wrist.

XSL•FO
RenderX

This study seeks 80 families, with 40 families per arm. Parents of preschoolers (aged 3-4 y) are recruited from a metropolitan area of a Midwest US state. We plan to recruit through the institution's research-based list-serves (including local children's hospitals), a local advertisement through social media, blogs, magazines, community-based efforts (eg, childcare centers and local events), and at local medical centers. Recruitment efforts include flyers and paid advertisements.

Preliminary Studies

In total, 2 separate studies were conducted to inform this intervention and mobile app. First, there was a cross-sectional study of 42 parents of preschoolers (24/42, 57% non-Hispanic White and 21/42, 51% male) who met few guidelines (zero or one of the 24-hour Movement Guidelines) who completed semistructured interviews; this information was used to inform mobile app content and describe existing resources and interest in a mobile app. For this cross-sectional study, we obtained written informed consent. The second study included 4 focus groups of 18 parents of preschoolers (10/18, 55% non-Hispanic White and 7/18, 38% male) who also met a few guidelines. We obtained verbal consent before participation for the second study. These focus groups aimed to solicit feedback on the core components and initial design of a 24-hour movement behavior intervention delivered through a mobile app.

In semistructured interviews, parents were asked about barriers and facilitators to their preschooler's and family's 24-hour movement behaviors, existing use of mobile apps and other resources, and interest in mobile app features to promote any of the preschoolers' 24-hour movement behaviors. Interview transcripts were reviewed to identify common barriers and facilitators for each behavior within the home environment, and app interest. These barriers and facilitators were aligned with existing multilevel frameworks [26,27], which included levels at the child, family, home environment, and macroenvironment. For study content, we specifically focused on factors that may inform a 12-week home-based intervention, rather than policy or system-level changes.

The nonmodifiable and modifiable factors are displayed in Figure 1 and described in detail in Table S1 in Multimedia Appendix 2. Nonmodifiable factors of family structure, income or occupation, weather, social stigma, and childcare arrangement were macroenvironment factors that impacted all 3 levels. A modifiable factor that impacted both the family and home environment, and all movement behaviors was scheduling and time management. In general, the remaining single behavior modifiable factors were interrelated with other behavior factors, especially between screen time and sleep. For example, parents may report that allowing screen time (screen time parenting practice) is part of the bedtime routine and helps their child wind down (sleep-promoting practice). These nonmodifiable and modifiable items then informed content matter and behavioral targets for the intervention. For example, in the week 9 lesson, Indoor Activities for MVPA, the content matter discusses the benefits of being outdoors (modifiable: outdoors) and preparing and planning for being active in less-than-desirable weather (nonmodifiable: weather).

Figure 1. Conceptual model of nonmodifiable and modifiable factors that influence child's 24-hour movement behaviors within the home (each ring represents a level of influence). PA: physical activity; ST: screen time. Solid boxes represent nonmodifiable factors. Hashed boxed represent modifiable factors that influence all behaviors.

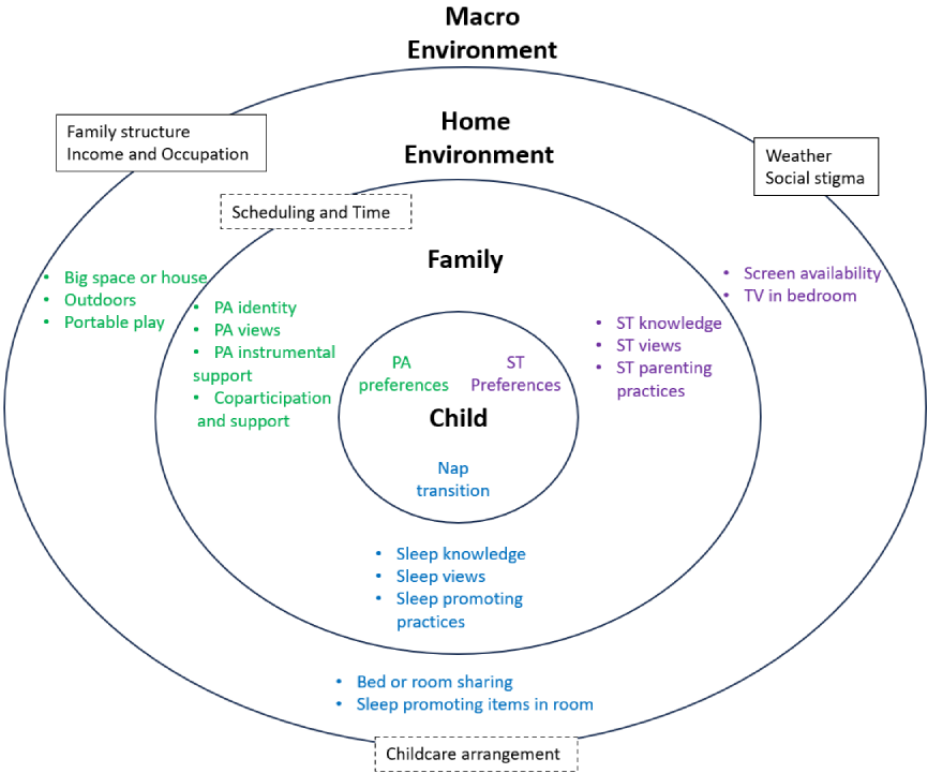

XSL•FO
RenderX

App Development

Study 1: Semistructured Interviews

Most parents (25/42, 59%) reported that they did not use apps to support any of the child's 24-hour movement behaviors, but did use apps to manage a home or family schedule (27/42, 64%), namely, a calendar app or shared to-do list. Parents reported many sources of information on each 24-hour movement behavior, which included the internet, family, and friends (Table S2 in Multimedia Appendix 2). Some reported using no resources for any of their preschooler's 24-hour movement behaviors (10/42, 23%). The majority (36/42, 85%) expressed interest in receiving information on their preschooler's PA, sleep, and screen time from a mobile app. Few (5/42, 12%) were unsure about their interest in a mobile app as they did not want more device notifications. Only 1 participant expressed that this app would not change their behavior.

Study 2: Focus Groups

Among 2 focus groups (n=10), parents were interested in the rationale for changing behaviors, opportunities to track their child's behaviors, gamification, and relating to other parents during the program. Importantly, parents were interested in information for changing child's PA and screen time to improve child's sleep, rather than focusing on all 3 behaviors at one time. Parents were interested in having opportunities to use preplanned and self-selected action plans to facilitate change. Some parents mentioned they would like a child-facing component for the app; this feature was ultimately not included in the final mobile app as it may indirectly impact preschooler's sedentary screen time.

In our third and fourth focus groups (n=8), we presented the created mobile app "*Shining Star*" to families for their feedback on design, features, and intervention components. The app was designed using the Pathverse mobile app, a no-code platform for researchers to create mobile health apps to facilitate behavior change [28]. Families liked the length and design of lessons, various options to set action plans, internal gamification, and the anonymous chat forum to connect with other parents. Overall, there was excitement for the mHealth app.

Intervention

Content

The 12-week intervention follows a multibehavior change process or transfer theory, whereby changing one behavior may lead to changes in another behavior [17,18]. This multibehavior framework was derived from semistructured interviews and was supported by focus group results, where parents were interested in learning about 1 behavior at a time, rather than all 3 at once. The first few lessons focus mainly on sleep and screen time activities based on the following: (1) focus group participants were interested in learning more about sleep, (2) many semistructured interview participants reported screen time during the bedtime routine, and (3) most participants would likely not meet the screen time guidelines and so may benefit from these early changes. After the focus on sleep and screen time during the first 6-8 weeks, the final weeks are focused on positive PA changes with the framing of making "daytime" changes to improve sleep. Given that PA is one of the more difficult behaviors to change, but parents recognize its impact on sleep, this behavior was chosen to be presented last.

As illustrated in Table 1, content was created based on behavior change techniques [29] and mechanism of action (MoA) [30,31] that would facilitate habit formation and align with the M-PAC framework. Accordingly, parents were provided with content that addressed reflective, regulatory, and reflexive processes. To address each part of M-PAC, aligning BCTs demonstrating links to appropriate MoA was identified through an online tool (The Theory and Techniques Tool). Using week 1 as an example, BCTs of "BCIO007063: information on health consequences" and "BCIO007068: salience of health consequences" were chosen for the MoAs of knowledge, attitude toward health behavior, and perceived susceptibility; these BCTs and MoA align with the M-PAC construct "instrumental attitude." Later in week 3, "BCIO007080: prompts and cues" then aligns with the MoA of behavioral cuing and the M-PAC construct of habit. Each week, parents have the option to set goals and action plans supporting the M-PAC construct of behavioral regulation. Parents also have the option to receive points for completing the lesson and meeting a goal, which aligns with BCIO0007257: provide positive material consequence for behavior, and the MoA of reinforcement.

XSL•FO
RenderX

Table 1. Shining Star intervention content and techniques to facilitate behavior change.

Phase, week, focus, behavior, and major topic	BCT[a] label and ID [29]	Mechanism of action [30,31]	Multiprocess action control construct	Conceptual model items addressed
Night				
Week 1: all behaviors—24-h movement to improve sleep	• 007063: Information on health consequences • 007068: Salience of consequences • 007024: Self-monitoring of behavior	• Knowledge • Attitude toward health behavior • Perceived susceptibility • Behavioral regulation	• Instrumental attitude • Behavioral regulation	• ST[b] knowledge • PA[c] views • Sleep knowledge • Social stigma • Sleep promoting items in room • TV in bedroom • Screen availability • Sleep parenting practices • Portable play items • Outdoors • Big space or house
Week 2: sleep—bedtime routines (who, what, and when)	• 007065: Inform about emotional consequences BCT • 007058: Instruction on how to perform the behavior • 007055: Demonstration of the behavior	• Knowledge • Skill • Capabilities	• Affective judgment • Perceived capability	• Sleep knowledge • Sleep promoting practices • ST knowledge • ST parenting practices
Week 3: sleep—how to make the new bedtime routine, routine	• 007096: Context specific repetition of behavior • 007094: Practice behavior • 007080: Prompt intended action	• Behavioral cuing (problem solving)	• Habit formation • Behavioral regulation	• Scheduling time, chaos • Sleep promoting practices
Week 4: sleep or screen time—bedroom environment	• 050348 Restructuring the physical environment • 050331: Remove averse stimulus • 007156: Add objects to the environment • 007080 Prompts and cues	• Environmental context and resources • Behavioral cuing	• Perceived opportunity • Affective judgment • Habit	• Television in bedroom • Sleep promoting items in room • Bed or room sharing • Screen availability
Week 5: sleep—parent sleep habits	• 007063: Information on health consequences • 007068: Salience of consequences • 007057: Draw attention to incompatible beliefs • 007159 Affirm valued self-identity	• Knowledge • Attitude toward health behavior • Perceived susceptibility	• Instrumental attitude • Identity	• Sleep knowledge • Sleep views
Week 6: sleep—how family members can support goals	• 050346: Directly restructure the social environment BCT • 007030: Advise to seek instrumental support • 0007257: Provide positive material consequence for behavior	• Social context and resources • Behavioral cuing • Social influences • Reinforcement	• Perceived opportunity • Identity • Behavioral regulation • Affective judgment	• Sleep promoting practices • Bed or room sharing • ST parenting practices • Screen availability • Childcare arrangement • Family structure

XSL•FO
RenderX

Phase, week, focus, behavior, and major topic	BCT[a] label and ID [29]	Mechanism of action [30,31]	Multiprocess action control construct	Conceptual model items addressed
Week 7: sleep—reassessment of goals	• 007063: Information on health consequences • 007068: Salience of consequences • 007058: Instruction on how to perform the behavior • 007024: Self-monitoring of behavior	• Knowledge • Attitude toward health behavior • Perceived susceptibility • Skill • Capabilities • Behavioral regulation	• Instrumental attitude • Perceived capability • Behavioral regulation • Perceived opportunity	• ST knowledge • PA views • Sleep knowledge • Social stigma • Sleep promoting practices • ST parenting practices • Sleep promoting items in room • Television in bedroom • Screen availability • Sleep parenting practices • Portable play items • Outdoors
Day				

XSL•FO
RenderX

Phase, week, focus, behavior, and major topic	BCT[a] label and ID [29]	Mechanism of action [30,31]	Multiprocess action control construct	Conceptual model items addressed
Week 8: screen time—screen-limiting during the day	• 007063: Information on health consequences • 007068: Salience of consequences • 007058: Instruct on how to perform behavior • 007055: Demonstrate the behavior	• Knowledge • Attitude toward health behavior • Perceived susceptibility • Skill • Belief about capabilities	• Instrumental attitude • Perceived capability	• ST knowledge • ST views • ST parenting practices • Screen availability • Television in bedroom • Portable play items • Outdoors • Big space or house
Week 9: PA—indoor activities for MVPA[d]	• 007058: Instruct on how to perform behavior • 007055: Demonstrate the behavior • 050348 Restructuring the physical environment	• Skill • Belief about capabilities • Environmental context and resources • Behavioral cuing	• Perceived capability • Behavioral regulation • Perceived opportunity	• Weather • Big space or house • PA instrumental support • Coparticipation • ST parenting practices • PA or ST preferences • Portable play items • Outdoors • Television in bedroom
Week 10: PA—being an active family	• 007064: inform about social consequences • 007057: Draw attention to incompatible beliefs • 007159 Affirm valued self-identity	• Knowledge • Belief about capabilities • Attitude toward behavior	• Instrumental attitude • Identity	• PA instrumental support • Coparticipation • PA or ST preferences • PA identity
Week 11: PA—setting up for success or finding “time” for PA	• 007064: inform about social consequences • 007096: Context specific repetition of behavior • 007094: Practice Behavior • 007080: Prompt intended action • 007058: Instruct on how to perform behavior	• Knowledge • Belief about capabilities • Behavioral cuing • Behavioral regulation	• Instrumental attitude • Identity • Habit formation	• PA instrumental support • Coparticipation • PA identity • PA or ST preferences • Scheduling time • Childcare arrangement
Week 12: PA—overall review	• 007063: Information on health consequences • 007068: Salience of consequences • 007058: Instruction on how to perform the behavior • 007024: Self-monitoring of behavior • 0007257: Provide positive material consequence for behavior	• Knowledge • Attitude toward health behavior • Perceived susceptibility • Skill • Capabilities • Behavioral regulation • Reinforcement	• Instrumental attitude • Perceived capability • Behavioral regulation • Affective judgment	• ST knowledge • PA views • Sleep knowledge • Social stigma • Sleep promoting practices • ST parenting practices • PA identity • PA instrumental support • Coparticipation • Sleep promoting items in room • Television in bedroom • Screen availability • Portable play items • Outdoors • Big space or house
All weeks				

XSL•FO
RenderX

Phase, week, focus, behavior, and major topic	BCT[a] label and ID [29]	Mechanism of action [30,31]	Multiprocess action control construct	Conceptual model items addressed
Weeks 1-12	• 007003: Set behavior goal	• Intention	• Behavioral regulation	• Various
Weeks 1-12	• 007008: Goal strategize BCT	• Beliefs about capabilities • Behavioral regulation	• Behavioral regulation	• Various
Weeks 1-12	• 007010: Action planning	• Behavioral cuing	• Behavioral regulation	• Various
Weeks 1-12	• 0007257: Provide positive material consequence for behavior	• Reinforcement	• Affective judgment	• Various

[a]BCT: behavior change technique.

[b]ST: screen time.

[c]PA: physical activity.

[d]MVPA: moderate-to-vigorous physical activity.

App Features

Using the Pathverse platform [28], we created weekly lessons and behavior-related goals. Each lesson is concise for readability (<500 characters) with a quiz, a suggested weekly goal, and links to additional resources. Features within the app to facilitate behavior change included gamification of completing lessons and weekly goals, an optional tracker for child behaviors (eg, sleep and going outside), an optional anonymous chat form to relate to other parents with preset chat topics (eg, local events), and a resources tab with additional research-based information on the importance of 24-hour movement behaviors. A screenshot of the app and the table of contents are presented in Figure 2.

The Pathverse platform is a secure and Health Insurance Portability and Accountability Act (HIPAA)–compliant platform with end-to-end encryption to deliver health content, monitor health behaviors, and support behavior change. To protect child and family information, the platform uses a confidential generic email address supplied by the study team. Content posted within the app is monitored regularly, with concerning or irrelevant posts addressed in a timely manner. No concerning posts have been detected by October 2025. The Pathverse app (which houses the intervention and its content) is available for download on the App Store and is compatible with both iOS (Apple Inc) and Android (Google), and can be downloaded on many digital devices (eg, cellphones and tablets). For this study, all parents are asked to download it onto their personal cellphones, and no parents, as of October 2025, have opted to download the app onto another device.

XSL•FO
RenderX

Figure 2. Screenshots of the Shining Star study welcome page and table of contents.

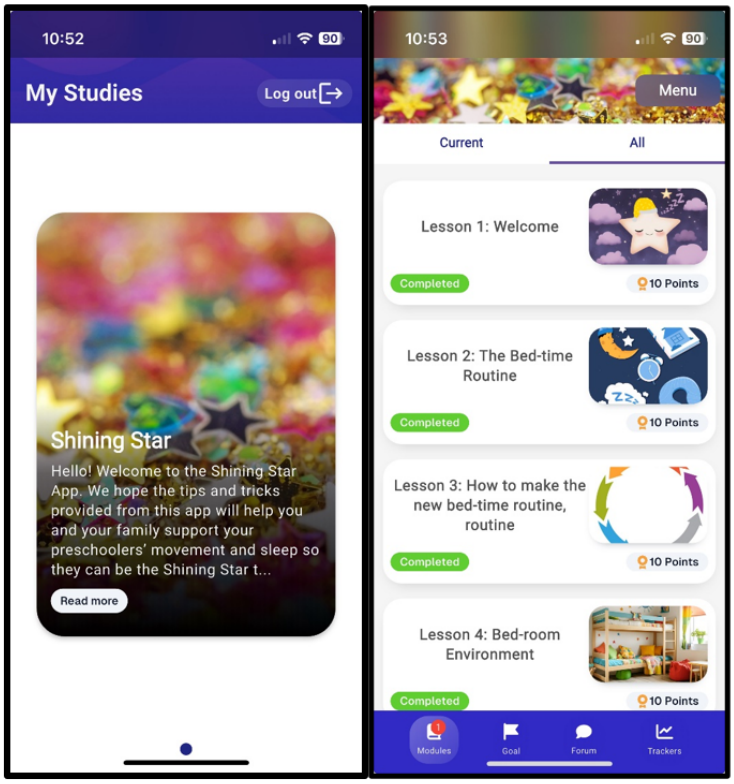

Procedures

An overview of enrollment, intervention, and assessments is provided in Figure 3. Harms and adverse events will be assessed at each visit. Study assessments are low-risk; thus, we anticipate few study-related harms. As this study is ongoing, assessments are described in the present tense.

Figure 3. Schedule of enrollment, intervention, and assessments during study. Feasibility and acceptability will be assessed weekly through the mobile app, and control participants are given access to the intervention after completing all week 12 measures, including 14-day accelerometry. M-PAC: multiple process action control; NIH: National Institutes of Health.

Construct	Measure	Screening	Baseline/ Week 0 Visit 1	Week 6 Visit 2	Week 12 Visit 3
Enrollment and Intervention Information					
Eligibility Screen		✓			
Informed Consent		✓			
Randomization			✓		
Intervention			✓ →	→	✓
24-Hour Movement Behaviors					
Physical Activity	Questionnaire	✓		✓	✓
	Accelerometry		✓	✓	
Sleep	Questionnaire	✓		✓	✓
	Accelerometry		✓	✓	✓
Sedentary screen-time	Questionnaire	✓		✓	✓
Child Development					
Growth	Anthropometrics		✓	✓	✓
Behavioral Problems and Executive Function	Questionnaires		✓	✓	✓
Cognition and Motor Skills	NIH Toolbox		✓	✓	✓
Other					
Demographics	Questionnaires	✓	✓		
Home, Parent, and Child level correlates	Questionnaires		✓	✓	✓
Parent Habit and M-PAC components	Questionnaires		✓	✓	✓
Usability of mobile app	System Usability Scale			✓	✓
Exit Survey	Questionnaire				✓

Screening

Interested parents complete a web screening tool using a secure online platform [32], which includes information on the preschooler's age and sex, language fluency, and interest and ability to download the app. Potential participants who met screening criteria are contacted for a phone call, during which a trained staff member inquires about the ability to travel to

XSL•FO
RenderX

visits, the preschooler's ability to wear a device on their wrist, and the preschooler's habitual PA, sedentary screen time, and sleep. Eligible families are then scheduled for their baseline visit (week 0).

Baseline Visit (Week 0)

At this visit (approximately 60-90 min), the study is described in a child-friendly manner to the preschooler, and the parent provides written informed consent. After consent is obtained, preschooler's height and weight are measured. The preschooler completes developmental tests with a trained researcher, and the parent completes questionnaires. Finally, an accelerometer is secured to the preschooler's wrist to wear for 2 weeks. The parent is instructed to return the accelerometer, an activity or sleep log, and a home inventory of equipment related to 24-hour movement behaviors after 2 weeks. At the end of the visit, the family is randomized.

If randomized to the intervention, the *Shining Star* mobile app is downloaded onto the parent's phone. A trained staff member provides a tutorial to explain the mobile app use and set mobile app reminders. The parent's first weekly lesson and app features are available that day.

Follow-up Visits (Week 6 and Week 12)

At the second visit (approximately 60-90 min; week 6), families return to complete preschooler anthropometrics, preschooler development tests, and parent questionnaires. Procedures are completed similarly to the baseline visit. Preschoolers are fitted with an accelerometer, and parents are asked to return the device, activity or sleep log, and home inventory after 2 weeks of wear. The trained staff answer any questions on the mobile app for families in the intervention group.

At the final visit (approximately 60-90 min; week 12), families return to complete preschooler anthropometrics, preschooler development tests, and parent questionnaires. Procedures are completed similarly to the baseline visit. Preschoolers are fitted with an accelerometer, and parents are asked to return the device, activity or sleep log, and home inventory after 2 weeks of wear. Parents in the intervention are allowed to delete the mobile app or continue viewing the content. Parents randomized to the control are allowed to download the mobile app, receive a tutorial, and access all mobile app lessons and features.

Measures

Primary Outcome

Our primary outcome is adherence to all 3 of the 24-hour Movement Guidelines. Preschoolers' PA and sleep were assessed in 15-second epochs using the Actigraph GT3X+ accelerometer. Preschoolers wear this device on their wrist using a Velcro band, only removing it for water-based activities (eg, bath). Accelerometer data will be evaluated using the Sadeh algorithm to detect sleep periods and compared with the parent-report of sleep [33]. Periods of ≥20 minutes of continuous zeroes during the day will be removed as nonwear time [34,35]. Validated cut points for wrist measurement in this age range will be used for sedentary behavior (0-328 counts per 5 seconds), light PA (329-1392 counts per 5 seconds), and MVPA (≥1393 counts per 5 seconds) [36]. Total PA is the sum of light PA and MVPA. Preschoolers with ≥3 days (at least 1 weekday and 1 weekend day) of ≥10 hours of wear, and 3 days of sleep, will be included for analysis, similar to another wrist-worn protocol [37], and is the estimated minimum for accelerometer wear in this age range [38,39]. To capture adherence to the sedentary screen time guideline, parents are asked, "On an average 24-hour period, how much time does your child spend using electronic screen devices (such as a smartphone, tablet, video game, or watch television or movies, videos on the internet) while they are sitting or lying down?" with a free response option. This question is conducted at screening and subsequent visits and was chosen to align with surveillance efforts of movement behaviors [40]. This questionnaire reflects the previous week before accelerometer wear. The sum of guidelines obtained based on these measures, and if the participant meets all 3 guidelines, will be used for analysis.

Secondary Outcome

Feasibility and acceptability are assessed by a weekly survey within the app. The parent reports how difficult or easy it was to achieve the goal, such as removing TV from the bedroom (Likert scale responses: range 1-5). Parents also rate their satisfaction with the app with regard to design, appeal, and functionality, and complete the System Usability Scale at week 6 and week 12, which can be used to assess overall usability of the app [41]. At the final visit, participants randomized to the intervention complete a 15-item survey to assess the acceptability and usability of app features and content, and overall participation. Responses from these questions will be combined with app usage data to evaluate the performance of these components for future iterations.

Exploratory Outcomes

Parent Report of 24-Hour Guideline Adherence

Parents will report how many days per week (range: 0-7) their child obtains at least 180 min/d of total PA and at least 60 min/d of MVPA. Staff members describe total PA as including any activity, such as slow or easy movement, based on direct report observations of preschooler's movement [42]. These may include walking, light dancing, playing catch, and easy bike riding. MVPA includes any "huff and puff" activity that increases their heart rate, such as playing a sport, more intense dancing, and running. If parents have trouble discerning days, the staff will ask them about the past week. Furthermore, parents will be asked on average how much overnight sleep and time their child spends napping, in separate questions with free response options, which are summed for a total amount of sleep during the day. The sleep questions have demonstrated moderate correlations with accelerometry measures [43]. Parents are also asked about screen time (refer to "Primary Outcomes" section). We chose these questions based on national surveys of child PA [44] and surveillance efforts of child movement behaviors [40,45]. This information is gathered during the phone screening process, where trained staff members ensure parents consider both weekdays and weekends, and any changes due to nontypical weeks, to obtain accurate habitual behavior amounts. Screening questions related to total PA, MVPA, and sleep are repeated at week 6 and week 12.

XSL•FO
RenderX

Anthropometrics

The child's height (cm) and weight (kg) are objectively measured using standard operating procedures with a mounted stadiometer and portable scale. BMI (weight/height^2) will be calculated and compared with national growth standards to calculate BMI *z* score and categorize weight status [46].

Cognitive and Motor Skills

These constructs are measured using the recent update of the National Institutes of Health Toolbox V3 [47]. Preschoolers complete the cognitive tasks validated for this age range and consolidated from the previous version (V2) [48,49]. The tests and associated constructs are picture vocabulary test (language), visual reasoning (nonverbal and visual reasoning), picture sequence memory test (episodic memory), and speed matching (processing speed) [50]. The preschooler also completes the National Institutes of Health Toolbox Early Childhood Motor Battery, which includes the tests and constructs as follows: 9-hole pegboard dexterity test (fine motor), grip strength (muscular strength), standing balance (balance), and 2-minute walk endurance test (muscular endurance). The preschooler completed the cognitive measures on a tablet, while the motor skills were completed with a trained administrator.

Behavioral Problems and Executive Function

Preschoolers' executive function and behavioral problems are assessed during this study. Parents complete questions with Likert scale responses from the Behavioral Assessment System for Children (3rd edition) related to child behavioral problems, and the Behavioral Rating Inventory of Executive Function Preschool for executive function [51], which will then be used to calculate the preschool child's T-score for analysis. Both questionnaires are valid tools and have shown to be useful tools for assessing child development [51,52].

Descriptive Characteristics

Home, parent, and child-level correlates, along with parent habit and M-PAC components, are collected**.** Home space and equipment are assessed by a parent report inventory of preschooler-appropriate availability and accessibility of PA (eg, jump rope) and media (eg, television) equipment, room design, and sleep components. This updated inventory is based on validated inventories in parents of older children [53] and preschoolers [54,55]. Household organization was assessed using the Confusion, Hubbub, and Order Scale, which is a 15-item measure assessing disturbance, organization, and noise levels in the home. It has been validated in parents of young children [56].

Parent-level correlates include parent support for PA, parent screen time practices, and bedtime and family routines. Parents complete a 4-item questionnaire to assess parental support of child PA based on a validated questionnaire in this age range [57], which includes core support elements, including encouragement, logistics, and coparticipation [58]. Parents report their perceived influence on their child's screen media use and limit setting of screen media from a validated questionnaire for parents of preschool children (12 questions total). These scales had high internal consistency (0.84-0.86) and correlation with child BMI and screen time [59]. Parents are asked 7 questions on family regularity, including having a bedtime ritual and regularity of family meals over the last week, based on a validated questionnaire for preschoolers [60].

Child-level correlates include solely child preferences. Parents complete questions, including "*when outside, my child prefers*" with a list of activity options, "*my child's physical activity is limited due to their interest or motivation*," and their child's preference for PA in their free time, with Likert scale options for each question. These scales had high internal consistency (0.60) and correlation with child screen time [59].

Finally, parental support for PA and sleep and limiting screen time is examined using self-report items of habit formation to establish routines for each behavior [61], and an adapted questionnaire measuring intention and behavioral support for all 3 movement behaviors based on the reflective, regulatory, and reflexive parts of the M-PAC model [11].

Stipends and Incentives

Intervention participants can redeem points achieved through engagement with the mobile app for up to US $75 worth of incentives. Examples of incentives are activities to do at home to encourage PA (eg, child wearable device) and sleep (eg, sound machine), and reduce screen time (eg, puzzles) [11].

Analytical Approach

Randomization and Blinding

Participants are randomized to intervention or control in a 1:1 ratio with block randomization stratified by age and sex, with a random allocation process developed by a biostatistician. Due to the nature of the study, staff administering the first visit are unblinded to group assignment.

Aim 1: Meeting 24-Hour Movement Guidelines

We will compare the proportion of intervention and control group participants meeting all 3 recommendations at 12 weeks using a chi-square analysis.

Aim 2: Feasibility

We will average the Likert scale responses (ranges 1.0-5.0) across all weeks completed (up to 12). Feasibility and acceptability will be achieved if the average response is ≥4.0/5.0 for each category.

Sample Size and Power Calculation

Our analysis is based on a previous home-based routine-based multibehavior intervention in preschool children, which sought to improve similar behaviors, and reported an increase in the proportion of children meeting the screen time guideline increased from 5% to 30% over 12 weeks [23]. Based on our past studies, we estimate that few will meet the screen time guideline [25,62], like the beginning metric of the routine-based study. Given the overlap in content between the routine-based study and this study, we hypothesize that we can also improve the screen time guideline and potentially the remaining behaviors within the same time frame. A minimally important difference in 24-hour movement behaviors is not clear as of October 2025, as there are no other 24-hour movement guideline adherence studies. A sample size of 30 families per arm will

XSL•FO
RenderX

provide 80% power to detect a meaningful difference in the proportion of control (1/30, 5%, no change from baseline) and intervention group (10/30, 35%) participants meeting all 3 recommendations at the 12-week visit. Assuming a dropout rate of 25%, we will recruit 40 families per arm (80 total). This power and difference can also be achieved with a minimum of 24 participants per group; while 30 participants per group will have 63.5% power to detect a difference between 5% and 25%, meeting all 3 recommendations; and a 99% power to detect a difference between 5% and 20%, meeting all 3 recommendations. Statistical significance will be set at P=.05 for all aim-related analyses.

Exploratory Aims

We will also conduct an intent-to-treat analysis using a log-binomial regression or Poisson with robust variance to adjust for baseline recommendation number and randomization strata (age and sex). Linear regression will also be used to compare changes in exploratory outcomes between treatment and control groups with adjustment for baseline values. If demographic differences, such as income, are found between the 2 groups, we will conduct a sensitivity analysis with these variables for adjustment. These exploratory outcomes include, in priority order: number of recommendations obtained (range: 0-3), individual 24-hour movement behaviors, anthropometry, early child motor battery scores, early child cognitive battery scores, executive functioning scores, and behavioral problem scores. Descriptive characteristics will also be explored, such as PA and accessibility score, household chaos score, parental support of PA, parental perceived influence on screen time, family regularity index, child preferences for PA, and theoretical component changes (eg, intention and behavioral support for PA). We will also explore the ideal change in behaviors using isotemporal substitution analysis, whereby evaluating substitution of one behavior (eg, sedentary behavior) for another (eg, MVPA) with change in child developmental outcomes.

Furthermore, we will examine the relationship between meeting specific recommendations and child outcomes, with a mediator of theoretical component changes (eg, habit formation scores). Mediation analysis will be conducted using the PROCESS vs 3.5 macro with 10,000 bootstrap intervals with unstandardized estimates and CIs [63]. These exploratory outcome testing will be subject to the Benjamini-Hochberg adjustment for multiple comparisons [64].

As for missing data, we will describe differences in baseline characteristics between participants with complete versus incomplete follow-up. Sensitivity analyses based on multiple imputation methods will be conducted to examine the robustness of our findings with respect to missing data.

Recruitment Plans

We will use various existing modalities of recruitment in the selected area to ensure the recruitment amount is achieved. As part of this 12-week study, participants are contacted every 2 weeks or up to 1 month to schedule visits and pick up accelerometers. All available data will be sought in the case of discontinuation after visit 1, with the priority of 24-hour movement behavior data (questionnaire and accelerometry), other questionnaires, then anthropometry, cognition and motor skill assessments.

Data Management Plan

All participants will be given an ID as part of the screening process. This ID will be stored in a secure database at the University of Kansas Medical Center. Only approved project staff will have access to the server. Data obtained during study visits and as part of the mobile app use will be reported as tabular data and stored in a secure electronic data capture system. All subject-level clinical data will be preserved and shared. Shared data will be deidentified, and the original data will be maintained at the investigator's institution. The minimal risk of this study and its short-term nature precluded appointment of a data monitoring committee.

Ethical Considerations

The University of Kansas Medical Center Institutional Review Board approved this study in April 2024 (approval 00160250). The Pennington Biomedical Research Center Institutional Review Board approved both the studies that were conducted for this protocol (approvals 2021-053 and 2023-011). The informed consent specified that every effort would be made to maintain participant confidentiality, such as deidentifying private information, and participants had the option to opt out of the study. Participants are paid up to US $200 for completing all study visits and receive small child toys (<US $5) at each complete visit.

Results

This study recruitment began in January 2025 and is anticipated to conclude in December 2026. As of October 2025, 39 families have been enrolled and randomized into the study, with 19 families in the intervention arm and 20 in the wait-list control arm. As of October 2025, there has been one loss after randomization. Furthermore, 28 families have completed their final visit, and no adverse events have been reported. Descriptive and analytic results are expected to be published in the year 2027 and reported within the trial registry.

Discussion

Principal Findings

This study aims to improve adherence to all three 24-hour Movement Guidelines in preschoolers who meet few guidelines using a theoretically informed, home-based, mobile app. Our preliminary studies in the target population demonstrate multilevel considerations to improve all 3 behaviors within the home, which can be improved through existing frameworks of multibehavior change, habit formation, behavior adoption, and behavior maintenance. Using mHealth technology allows information to be provided to parents within the most applicable setting for 24-hour movement behaviors, the home environment. Together, a rigorously designed behavior change intervention with thorough preliminary work has resulted in the *Shining Star* study.

This study's use of a mobile app presents challenges and opportunities to impact preschooler's 24-hour movement

XSL•FO
RenderX

behaviors. The majority of parents in our focus groups were interested in receiving information about their child's 24-hour movement behaviors from a mobile app. The mobile app platform may reduce travel burden and reach parents in the home setting; the mobile app may also promote additional parental screen time. Given parental screen time is an established correlate of child screen time [59,65] and identified as a factor in this study's conceptual model, we did not include the parent-suggested child-facing component in the app. Rather, the intervention content and features focus on creating habits, goals, and planning to improve each 24-hour movement behavior; these components have all been associated with reducing child screen time in a meta-analysis of 204 interventions [66]. Nonetheless, parental screen time practices are collected as part of this study to examine whether the current app impacted a key parent 24-hour movement behavior.

Comparison With Previous Work

To the authors' knowledge, there is only one other 24-hour movement-focused intervention, as the current 24-hour movement scientific landscape is predominantly observational studies [9]. That study was also a 12-week randomized controlled trial with wait list control in 147 young children, and it demonstrated significant decreases in screen time but no intervention effect in other behaviors [9,67]. Major differences between the past trial and this trial include (1) population (this trial: 3-4 y who meet 0 or 1 guideline, past trial: 3-5 y who meet 0-2 guidelines), (2) intervention approaches (multibehavior and transfer theory vs all 3 behaviors at once and integrated), (3) outcomes (adherence to 24-hour Movement Guidelines vs Activity Sleep Index), and (4) intervention methods (all mobile app vs internet-based education and online workshops) [9]. It is unclear if this study's different populations, intervention approaches, modalities, and outcomes will result in significant results, but there will still be a noticeable contrast.

Strengths and Limitations

The strengths of this study are a theoretically based intervention in an important population and outcome for improving child health outcomes. This study uses a mobile app, which also allows for measuring engagement and usage of the intervention. Limitations of the study pertain to measurement and generalizability. A potential limitation is the use of parent self-report of child 24-hour movement behaviors for inclusion, which may overreport or underreport certain behaviors, and a lack of overlap in screen time measure questions with accelerometer wear. This study addresses this limitation by using validated individual questions and having thoroughly trained staff conduct these assessments, so accurate measures can be obtained. Furthermore, this use of parent-reported guideline adherence for screening and accelerometer data for outcome measures may misclassify participants and potentially weaken effect estimates. Currently, there is no full 24-hour movement behavior self or parent-report measure for this age range, and there are a limited number of valid self-report tools for 24-hour movement behaviors in general [68]. Another limitation is the generalizability of this study. This intervention is based on preliminary data in the United States, and sleep habits and screen time may significantly differ across the globe. Thus, content and potential results may not translate to other populations [69]. Moreover, this study solely includes English-speaking individuals, as app content was not translated to other languages before study commencement.

Implications

If successful, this study has multiple implications for future research and practice. First, these results may posit using similar transfer theory approaches for improving all 3 movement behaviors, rather than all at once as conducted in another 24-hour movement trial in preschoolers [9]. Second, these results may support the use of mHealth apps for improving children's movement behaviors and adopting other evidence-based information to such platforms to reach parents. Finally, significant results support testing this intervention in a larger trial of diverse participants, as children from low-income and minority backgrounds are at greater risk for not meeting these recommendations, as they are more likely to have excess screen time [70], short sleep [71], and less PA relative to counterparts [72]. Findings from the larger trial and this study would then inform meaningful changes in preschoolers meeting 24-hour movement recommendations and translating this mobile app in pediatricians' and sleep health professionals' practices.

Conclusions

In summary, this paper describes the creation of a novel approach to improve all three 24-hour movement behaviors in home settings using a mobile app. This study will use a unique approach to improve multiple levels of influence and ultimately preschooler behavior. Results from this study have implications for future 24-hour movement interventions, our understanding of improving all 3 behaviors, and ultimately, improvements in preschoolers' health.

Acknowledgments

We are grateful to the families who participated in the preliminary studies and this study, and Jesse Lyle's contribution to the initial study conduct. Generative artificial intelligence was not used in any portion of this paper.

Funding

Funding for CLK was provided by the National Institutes of Health (K99HD107159, R00HD107159, and P20GM144269). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

XSL•FO
RenderX

Data Availability

Preliminary study data analyzed for this paper is available upon reasonable request. Final data generated from this study is available from authors upon reasonable request and will be submitted to a National Institutes of Health–approved repository.

Authors' Contributions

CLK, JMB, ML, RLNJ, RER, and LMR contributed to study conceptualization, methodology, investigation, funding acquisition, and project administration. CLK contributed to writing the original draft. GB and MS contributed to data curation. All authors contributed to writing – review and editing. All authors have read, provided feedback, and agreed to the final version of the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
SPIRIT checklist.
[DOCX File , 29 KB - resprot_v15i1e75621_app1.docx]

Multimedia Appendix 2
Components on the 24-hour movement behaviors in the home environment conceptual model and resources to promote adequate 24-hour movement behaviors provided by parents.
[PDF File (Adobe PDF File), 153 KB - resprot_v15i1e75621_app2.pdf]

References

1. Ansari M. WHO Guidelines on Physical Activity, Sedentary Behaviour and Sleep for Children Under 5 Years of Age. Geneva: World Health Organization; 2019.
2. Mendoza JA, Zimmerman FJ, Christakis DA. Television viewing, computer use, obesity, and adiposity in US preschool children. Int J Behav Nutr Phys Act 2007;4:44 [FREE Full text] [doi: 10.1186/1479-5868-4-44] [Medline: 17894878]
3. Marques A, Ramirez-Campillo R, Gouveia É, Ferrari G, Tesler R, Marconcin P, et al. 24-h movement guidelines and overweight and obesity indicators in toddlers, children and adolescents: a systematic review and meta-analysis. Sports Med Open 2023;9(1):30 [FREE Full text] [doi: 10.1186/s40798-023-00569-5] [Medline: 37184735]
4. Feng J, Zheng C, Sit CH, Reilly JJ, Huang WY. Associations between meeting 24-hour movement guidelines and health in the early years: a systematic review and meta-analysis. J Sports Sci 2021;39(22):2545-2557 [FREE Full text] [doi: 10.1080/02640414.2021.1945183] [Medline: 34176439]
5. Kracht CL, Webster EK, Staiano AE. Relationship between the 24-Hour movement guidelines and fundamental motor skills in preschoolers. J Sci Med Sport 2020;23(12):1185-1190 [FREE Full text] [doi: 10.1016/j.jsams.2020.06.021] [Medline: 32653249]
6. Hjorth MF, Chaput J, Damsgaard CT, Dalskov S, Andersen R, Astrup A, et al. Low physical activity level and short sleep duration are associated with an increased cardio-metabolic risk profile: a longitudinal study in 8-11 year old Danish children. PLoS One 2014;9(8):e104677 [FREE Full text] [doi: 10.1371/journal.pone.0104677] [Medline: 25102157]
7. Chong KH, Suesse T, Cross PL, Ryan ST, Aadland E, Aoko O, et al. Pooled analysis of physical activity, sedentary behavior, and sleep among children from 33 countries. JAMA Pediatr 2024;178(11):1199-1207. [doi: 10.1001/jamapediatrics.2024.3330] [Medline: 39348138]
8. Kracht CL, Burkart S, Groves CI, Balbim GM, Pfledderer CD, Porter CD, et al. 24-hour movement behavior adherence and associations with health outcomes: an umbrella review. J Act Sedentary Sleep Behav 2024;3(1):25. [doi: 10.1186/s44167-024-00064-6] [Medline: 39399355]
9. Feng J, Huang WY, Sit CH. Effectiveness of a parent-focused intervention targeting 24-h movement behaviors in preschool-aged children: study protocol for a randomized controlled trial. Front Public Health 2022;10:870281 [FREE Full text] [doi: 10.3389/fpubh.2022.870281] [Medline: 35677761]
10. Guan H, Okely AD, Aguilar-Farias N, Del Pozo Cruz B, Draper CE, El Hamdouchi A, et al. Promoting healthy movement behaviours among children during the COVID-19 pandemic. Lancet Child Adolesc Health 2020;4(6):416-418 [FREE Full text] [doi: 10.1016/S2352-4642(20)30131-0] [Medline: 32458805]
11. Rhodes RE, Berry T, Faulkner G, Latimer-Cheung AE, O'Reilly N, Tremblay MS, et al. Application of the multi-process action control framework to understand parental support of child and youth physical activity, sleep, and screen time behaviours. Appl Psychol Health Well Being 2019;11(2):223-239. [doi: 10.1111/aphw.12150] [Medline: 30743309]
12. Lally P, van Jaarsveld CHM, Potts HWW, Wardle J. How are habits formed: modelling habit formation in the real world. Euro J Social Psych 2009;40(6):998-1009. [doi: 10.1002/ejsp.674]

XSL•FO
RenderX

13. Rhodes R, Cox A, Sayar R. What predicts the physical activity intention-behavior gap? A systematic review. Ann Behav Med 2022;56(1):1-20. [doi: 10.1093/abm/kaab044] [Medline: 34231844]
14. Rhodes RE. Multi-process action control in physical activity: a primer. Front Psychol 2021;12:797484 [FREE Full text] [doi: 10.3389/fpsyg.2021.797484] [Medline: 34975698]
15. Rhodes RE. The evolving understanding of physical activity behavior: a multi-process action control approach. In: Advances in Motivation Science. San Diego, CA, US: Elsevier Academic Press; 2017:171-205.
16. Rhodes RE, Blanchard CM, Quinlan A, Naylor P, Warburton DE. Family physical activity planning and child physical activity outcomes: a randomized trial. Am J Prev Med 2019;57(2):135-144. [doi: 10.1016/j.amepre.2019.03.007] [Medline: 31248744]
17. Lippke S, Nigg CR, Maddock JE. Health-promoting and health-risk behaviors: theory-driven analyses of multiple health behavior change in three international samples. Int J Behav Med 2012;19(1):1-13 [FREE Full text] [doi: 10.1007/s12529-010-9135-4] [Medline: 21234735]
18. Champion KE, Gardner LA, McGowan C, Chapman C, Thornton L, Parmenter B, Health4Life Team, et al. A web-based intervention to prevent multiple chronic disease risk factors among adolescents: co-design and user testing of the health4life school-based program. JMIR Form Res 2020;4(7):e19485 [FREE Full text] [doi: 10.2196/19485] [Medline: 32720898]
19. Nezami BT, Ward DS, Lytle LA, Ennett ST, Tate DF. A mHealth randomized controlled trial to reduce sugar-sweetened beverage intake in preschool-aged children. Pediatr Obes 2018;13(11):668-676. [doi: 10.1111/ijpo.12258] [Medline: 29119719]
20. Swindle T, Poosala AB, Zeng N, Børsheim E, Andres A, Bellows LL. Digital intervention strategies for increasing physical activity among preschoolers: systematic review. J Med Internet Res 2022;24(1):e28230 [FREE Full text] [doi: 10.2196/28230] [Medline: 35014962]
21. Nyström CD, Sandin S, Henriksson P, Henriksson H, Trolle-Lagerros Y, Larsson C, et al. Mobile-based intervention intended to stop obesity in preschool-aged children: the MINISTOP randomized controlled trial. Am J Clin Nutr 2017;105(6):1327-1335 [FREE Full text] [doi: 10.3945/ajcn.116.150995] [Medline: 28446496]
22. Haines J, McDonald J, O'Brien A, Sherry B, Bottino CJ, Schmidt ME, et al. Healthy habits, happy homes: randomized trial to improve household routines for obesity prevention among preschool-aged children. JAMA Pediatr 2013;167(11):1072-1079. [doi: 10.1001/jamapediatrics.2013.2356] [Medline: 24019074]
23. Marsh S, Taylor R, Galland B, Gerritsen S, Parag V, Maddison R. Results of the 3 Pillars Study (3PS), a relationship-based programme targeting parent-child interactions, healthy lifestyle behaviours, and the home environment in parents of preschool-aged children: a pilot randomised controlled trial. PLoS One 2020;15(9):e0238977 [FREE Full text] [doi: 10.1371/journal.pone.0238977] [Medline: 32941530]
24. McNeill J, Howard SJ, Vella SA, Cliff DP. Compliance with the 24-hour movement guidelines for the early years: cross-sectional and longitudinal associations with executive function and psychosocial health in preschool children. J Sci Med Sport 2020;23(9):846-853. [doi: 10.1016/j.jsams.2020.02.011] [Medline: 32360244]
25. Kracht CL, Katzmarzyk PT, Staiano AE. Household chaos, family routines, and young child movement behaviors in the U.S. during the COVID-19 outbreak: a cross-sectional study. BMC Public Health 2021;21(1):860 [FREE Full text] [doi: 10.1186/s12889-021-10909-3] [Medline: 33947357]
26. Rhodes RE, Guerrero MD, Vanderloo LM, Barbeau K, Birken CS, Chaput J, et al. Development of a consensus statement on the role of the family in the physical activity, sedentary, and sleep behaviours of children and youth. Int J Behav Nutr Phys Act 2020;17(1):74 [FREE Full text] [doi: 10.1186/s12966-020-00973-0] [Medline: 32539730]
27. Sallis J, Owen N, Fisher E. Ecological models of health behavior. In: Health Behavior and Health Education. US: Jossey-Bass/Wiley; 2008:465-485.
28. Liu S, La H, Willms A, Rhodes RE. A "No-Code" app design platform for mobile health research: development and usability study. JMIR Form Res 2022;6(8):e38737 [FREE Full text] [doi: 10.2196/38737] [Medline: 35980740]
29. Marques MM, Wright AJ, Corker E, Johnston M, West R, Hastings J, et al. The behaviour change technique ontology: transforming the behaviour change technique taxonomy v1. Wellcome Open Res 2023;8:308. [doi: 10.12688/wellcomeopenres.19363.1]
30. Connell L, Carey RN, de Bruin M, Rothman AJ, Johnston M, Kelly MP, et al. Links between behavior change techniques and mechanisms of action: an expert consensus study. Ann Behav Med 2019;53(8):708-720 [FREE Full text] [doi: 10.1093/abm/kay082] [Medline: 30452535]
31. Carey R. Behavior change techniques and their mechanisms of action: a synthesis of links described in published intervention literature. Ann Behav Med 2019;53(8):693-707. [doi: 10.31234/osf.io/x5372]
32. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
33. Sadeh A, McGuire JP, Sachs H, Seifer R, Tremblay A, Civita R, et al. Sleep and psychological characteristics of children on a psychiatric inpatient unit. J Am Acad Child Adolesc Psychiatry 1995;34(6):813-819. [doi: 10.1097/00004583-199506000-00023] [Medline: 7608056]

XSL•FO
RenderX

34. Hnatiuk JA, Ridgers ND, Salmon J, Hesketh KD. Maternal correlates of young children's physical activity across periods of the day. J Sci Med Sport 2017;20(2):178-183. [doi: 10.1016/j.jsams.2016.06.014] [Medline: 27460333]
35. Bisson M, Tremblay F, Pronovost E, Julien A, Marc I. Accelerometry to measure physical activity in toddlers: Determination of wear time requirements for a reliable estimate of physical activity. J Sports Sci 2019;37(3):298-305. [doi: 10.1080/02640414.2018.1499391] [Medline: 30009665]
36. Johansson E, Larisch L, Marcus C, Hagströmer M. Calibration and validation of a wrist- and hip-worn actigraph accelerometer in 4-year-old children. PLoS One 2016;11(9):e0162436 [FREE Full text] [doi: 10.1371/journal.pone.0162436] [Medline: 27617962]
37. Chen B, Bernard JY, Padmapriya N, Yao J, Goh C, Tan KH, et al. Socio-demographic and maternal predictors of adherence to 24-hour movement guidelines in Singaporean children. Int J Behav Nutr Phys Act 2019;16(1):70 [FREE Full text] [doi: 10.1186/s12966-019-0834-1] [Medline: 31438965]
38. Trost SG, Pate RR, Freedson PS, Sallis JF, Taylor WC. Using objective physical activity measures with youth: how many days of monitoring are needed? Med Sci Sports Exerc 2000;32(2):426-431. [doi: 10.1097/00005768-200002000-00025] [Medline: 10694127]
39. Cliff DP, Reilly JJ, Okely AD. Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0-5 years. J Sci Med Sport 2009;12(5):557-567. [doi: 10.1016/j.jsams.2008.10.008] [Medline: 19147404]
40. Okely AD, Reilly JJ, Tremblay MS, Kariippanon KE, Draper CE, El Hamdouchi A, et al. Cross-sectional examination of 24-hour movement behaviours among 3- and 4-year-old children in urban and rural settings in low-income, middle-income and high-income countries: the SUNRISE study protocol. BMJ Open 2021;11(10):e049267 [FREE Full text] [doi: 10.1136/bmjopen-2021-049267] [Medline: 34697112]
41. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. International Journal of Human-Computer Interaction 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
42. Brown WH, Pfeiffer KA, McIver KL, Dowda M, Almeida MJCA, Pate RR. Assessing preschool children's physical activity: the observational system for recording physical activity in children-preschool version. Res Quart Exerc Sport 2006;77(2):167-176. [doi: 10.5641/027013606x13080769704361]
43. Zhang Z, Abdeta C, Chelly MS, Cruz JDP, Germana L, Ghofranipour F, et al. Measurement properties of the SUNRISE study parent questionnaire for assessing sleep and sleep-related family practices. Sleep Health 2025;11(5):613-623 [FREE Full text] [doi: 10.1016/j.sleh.2025.04.006] [Medline: 40461398]
44. CDC. National health and nutrition examination survey data. National Center for Health Statistics (NCHS). URL: https://www.cdc.gov/nchs/nhanes/index.html [accessed 2025-12-18]
45. Rhodes RE, Spence JC, Berry T, Faulkner G, Latimer-Cheung AE, O'Reilly N, et al. Parental support of the Canadian 24-hour movement guidelines for children and youth: prevalence and correlates. BMC Public Health 2019;19(1):1385 [FREE Full text] [doi: 10.1186/s12889-019-7744-7] [Medline: 31660925]
46. Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, et al. 2000 CDC growth charts for the United States: methods and development. Vital Health Stat 11 2002;17(246):1-190 [FREE Full text] [Medline: 12043359]
47. Mungas D, Widaman K, Zelazo PD, Tulsky D, Heaton RK, Slotkin J, et al. VII. NIH toolbox cognition battery (CB): factor structure for 3 to 15 year olds. Monogr Soc Res Child Dev 2013;78(4):103-118 [FREE Full text] [doi: 10.1111/mono.12037] [Medline: 23952205]
48. Monteleone M. Updates on the NIH toolbox v3 and introduction to the NIHTB infant and toddler toolbox and mobile toolbox. J Neurosurg Anesthesiol 2025;37(1):122-124. [doi: 10.1097/ana.0000000000001000]
49. Becker L, Condy E, Kaat A, Thurm A. How do 3-year-olds do on the NIH toolbox cognitive battery? Child Neuropsychol 2023;29(4):521-542 [FREE Full text] [doi: 10.1080/09297049.2022.2100337] [Medline: 35876076]
50. Reuben DB, Magasi S, McCreath HE, Bohannon RW, Wang Y, Bubela DJ, et al. Motor assessment using the NIH Toolbox. Neurology 2013;80(11 Suppl 3):S65-S75 [FREE Full text] [doi: 10.1212/WNL.0b013e3182872e01] [Medline: 23479547]
51. Sherman EMS, Brooks BL. Behavior rating inventory of executive function – preschool version (BRIEF-P): test review and clinical guidelines for use. Child Neuropsychology 2010;16(5):503-519. [doi: 10.1080/09297041003679344]
52. Doyle A, Ostrander R, Skare S, Crosby RD, August GJ. Convergent and criterion-related validity of the behavior assessment system for children-parent rating scale. J Clin Child Psychol 1997;26(3):276-284. [doi: 10.1207/s15374424jccp2603_6] [Medline: 9292385]
53. Sirard JR, Nelson MC, Pereira MA, Lytle LA. Validity and reliability of a home environment inventory for physical activity and media equipment. Int J Behav Nutr Phys Act 2008;5:24 [FREE Full text] [doi: 10.1186/1479-5868-5-24] [Medline: 18445280]
54. Ward D, Hales D, Haverly K, Marks J, Benjamin S, Ball S, et al. An instrument to assess the obesogenic environment of child care centers. Am J Health Behav 2008;32(4):380-386. [doi: 10.5555/ajhb.2008.32.4.380] [Medline: 18092898]
55. Hales D, Vaughn AE, Mazzucca S, Bryant MJ, Tabak RG, McWilliams C, et al. Development of HomeSTEAD's physical activity and screen time physical environment inventory. Int J Behav Nutr Phys Act 2013;10:132 [FREE Full text] [doi: 10.1186/1479-5868-10-132] [Medline: 24313962]

XSL•FO
RenderX

56. Matheny AP, Wachs TD, Ludwig JL, Phillips K. Bringing order out of chaos: psychometric characteristics of the confusion, hubbub, and order scale. Journal of Applied Developmental Psychology 1995;16(3):429-444. [doi: 10.1016/0193-3973(95)90028-4]
57. Larouche R, Sarmiento OL, Broyles ST, Denstel KD, Church TS, Barreira TV, ISCOLE Research Group. Are the correlates of active school transport context-specific? Int J Obes Suppl 2015;5(Suppl 2):S89-S99 [FREE Full text] [doi: 10.1038/ijosup.2015.25] [Medline: 27152191]
58. Rhodes RE, Perdew M, Malli S. Correlates of parental support of child and youth physical activity: a systematic review. Int J Behav Med 2020;27(6):636-646. [doi: 10.1007/s12529-020-09909-1] [Medline: 32529629]
59. Vaughn AE, Hales DP, Neshteruk CD, Ward DS. HomeSTEAD's physical activity and screen media practices and beliefs survey: instrument development and integrated conceptual model. PLoS One 2019;14(12):e0226984 [FREE Full text] [doi: 10.1371/journal.pone.0226984] [Medline: 31891610]
60. Rijlaarsdam J, Tiemeier H, Ringoot AP, Ivanova MY, Jaddoe VWV, Verhulst FC, et al. Early family regularity protects against later disruptive behavior. Eur Child Adolesc Psychiatry 2016;25(7):781-789 [FREE Full text] [doi: 10.1007/s00787-015-0797-y] [Medline: 26589300]
61. Verplanken B, Orbell S. Orbell, reflections on past behavior: a self-report index of habit strength. J Applied Social Pyschol 2006;33(6):1313-1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
62. Kracht C, Webster EK, Staiano AE. Sociodemographic differences in young children meeting 24-hour movement guidelines. J Phys Act Health 2019;16(10):908-915 [FREE Full text] [doi: 10.1123/jpah.2019-0018] [Medline: 31491748]
63. Hayes AF. Introduction to mediation, moderation, and conditional process analysis : a regression-based approach. In: Methodology in the Social Sciences. New York: Guilford Press; 2018:692.
64. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Societyries B (Methodological) 2018;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
65. Paudel S, Jancey J, Subedi N, Leavy J. Correlates of mobile screen media use among children aged 0-8: a systematic review. BMJ Open 2017;7(10):e014585 [FREE Full text] [doi: 10.1136/bmjopen-2016-014585] [Medline: 29070636]
66. Jones A, Armstrong B, Weaver RG, Parker H, von Klinggraeff L, Beets MW. Identifying effective intervention strategies to reduce children's screen time: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2021;18(1):126 [FREE Full text] [doi: 10.1186/s12966-021-01189-6] [Medline: 34530867]
67. Feng J, Huang WY, Sit CH, Reilly JJ, Khan A. Effectiveness of a parent-focused intervention targeting 24-hour movement behaviours in preschool-aged children: a randomised controlled trial. Int J Behav Nutr Phys Act 2024;21(1):98. [doi: 10.1186/s12966-024-01650-2] [Medline: 39252110]
68. Šuc A, Einfalt L, Šarabon N, Kastelic K. Validity and reliability of self-reported methods for assessment of 24-h movement behaviours: a systematic review. Int J Behav Nutr Phys Act 2024;21(1):83 [FREE Full text] [doi: 10.1186/s12966-024-01632-4] [Medline: 39095778]
69. Zhang Z, Abdeta C, Chelly MS, Del Pozo-Cruz J, Draper CE, Engberg E, et al. Geocultural differences in preschooler sleep profiles and family practices: an analysis of pooled data from 37 countries. Sleep 2025;48(4) [FREE Full text] [doi: 10.1093/sleep/zsae305] [Medline: 39715306]
70. Carson V, Kuzik N. Demographic correlates of screen time and objectively measured sedentary time and physical activity among toddlers: a cross-sectional study. BMC Public Health 2017;17(1):187 [FREE Full text] [doi: 10.1186/s12889-017-4125-y] [Medline: 28193271]
71. Peña MM, Rifas-Shiman SL, Gillman MW, Redline S, Taveras EM. Racial/ethnic and socio-contextual correlates of chronic sleep curtailment in childhood. Sleep 2016;39(9):1653-1661 [FREE Full text] [doi: 10.5665/sleep.6086] [Medline: 27306269]
72. Dias K, White J, Jago R, Cardon G, Davey R, Janz K, et al. International comparison of the levels and potential correlates of objectively measured sedentary time and physical activity among three-to-four-year-old children. Int J Environ Res Public Health 2019;16(11):1929 [FREE Full text] [doi: 10.3390/ijerph16111929] [Medline: 31159176]

Abbreviations

mHealth: mobile health
HIPAA: Health Insurance Portability and Accountability Act
MoA: mechanism of action
M-PAC: multi-process action control
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

XSL•FO
RenderX

Edited by J Sarvestan; submitted 07.Apr.2025; peer-reviewed by J Arts, WY Huang; comments to author 03.Oct.2025; revised version received 24.Nov.2025; accepted 25.Nov.2025; published 29.Jan.2026.

Please cite as:
Kracht CL, Berge JM, LeBlanc M, Newton Jr RL, Rhodes RE, Bolamperti G, Snodgrass M, Redman LM
A Home-Based Intervention to Improve Adherence to the 24-Hour Movement Guidelines in Young Children: Protocol for a Mobile App–Based Randomized Control Trial
JMIR Res Protoc 2026;15:e75621
URL: https://www.researchprotocols.org/2026/1/e75621
doi:10.2196/75621
PMID:

©Chelsea L Kracht, Jerica M Berge, Monique LeBlanc, Robert L Newton Jr, Ryan E Rhodes, Grace Bolamperti, Madigan Snodgrass, Leanne M Redman. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 29.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Implementation of a Participatory Ergonomics Intervention to Reduce Musculoskeletal and Stress-Related Mental Health Risks in Australian Retail Workers: Protocol for a Randomized Controlled Trial

Elise Condie[1], MAppSc; Victoria Weale[1], PhD; Katrina A Lambert[1], PhD; Jodi Oakman[1], PhD

Centre for Ergonomics and Human Factors, Department of Public Health, La Trobe University, Bundoora, Victoria, Australia

Corresponding Author:
Elise Condie, MAppSc
Centre for Ergonomics and Human Factors
Department of Public Health
La Trobe University
Kingsbury Drive
Bundoora, Victoria, 3083
Australia
Phone: 61 3 9479 3235
Email: e.condie@latrobe.edu.au

Abstract

Background: Worker participation has been identified as important for managing the risks of work-related musculoskeletal disorders (WMSDs) and stress-related mental health problems (MHPs). Previously identified barriers include securing long-term management support to implement risk reduction measures. Few studies evaluate how a manager or decision maker's readiness to act influences the outcomes of a participatory ergonomics program. The Stages of Change (SoC) framework has been suggested for tailoring ergonomics interventions to managers' receptiveness in a workplace setting.

Objective: The main aim is to evaluate the implementation of the "A Participatory Hazard Identification and Risk Management" (APHIRM) toolkit in the online order fulfillment department for a sample of stores in a large retail organization, compared to usual risk management practice.

Methods: This study is a cluster quasi–randomized controlled trial, comparing implementation of the APHIRM toolkit with usual safety risk management practice. As is typical for workplaces, the intervention is facilitated by the organization's safety team. We recruited 9 control and 9 intervention stores to the study through random selection of eligible stores. Quantitative data are collected at baseline and 12-month follow-up. Qualitative data to enable a process evaluation are collected over the duration of the study. Primary outcome measures are physical and psychosocial hazard severity scores. Secondary outcomes are self-rated pain and discomfort scores and action plan implementation measures. Managers' progression through SoC is an additional outcome measure. The primary outcome measures (physical and psychosocial hazard severity ratings) will be analyzed by variance-weighted cluster-level ANCOVA. Ethics approval was granted by the La Trobe University Human Research Ethics Committee (HEC25088).

Results: Funding was provided in June 2025. Recruitment and randomization concluded in early August 2025. The intervention, including data collection, commenced in late August 2025 and is expected to conclude in September 2026. A total of 332 participants have been recruited to the study. Response rates have averaged 46% across control and intervention groups. As of January 2026, no data analyses have been conducted. Primary findings are anticipated to be published in Spring 2028.

Conclusions: This study evaluates the implementation of the APHIRM toolkit survey in a multisite, large retail organization in Australia and describes the use of toolkit resources. It evaluates managers' SoC regarding WMSD and MHP prevention and how this may influence outcomes. Findings from this study should provide additional insight on how to implement the toolkit in large organizations to reduce WMSD and stress-related MHP risk and inform future development of the content of the APHIRM toolkit. This study is anticipated to further inform tailoring of interventions to managers' and decision-makers' SoC.

Trial Registration: OSF Registries 10.17605/OSF.IO/82R9G; https://osf.io/82r9g

International Registered Report Identifier (IRRID): DERR1-10.2196/84864

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e84864) doi:10.2196/84864

KEYWORDS

ergonomics; implementation science; musculoskeletal diseases; occupational stress; randomized controlled trial

Introduction

Background

Work-related musculoskeletal disorders (WMSDs) are one of the most frequent and costly type of compensable injuries in Australia, representing 87% of serious claims for injury or disease during 2019-2020 [1]. "Body stressing" represented 50% of all serious claims lodged in the retail trade sector from 2000 to 2021 [2]. Previous research has indicated that employers and ergonomists remain predominantly focused on controlling the biomechanical hazards associated with WMSDs [3-6]. This is despite consensus that psychosocial hazard exposures also contribute to these injuries and that exposure to multiple hazards simultaneously can have an additive or interactive effect [7-9]. Workplace psychosocial hazard exposure can also lead to a stress response in workers, which can in turn lead to mental health problems (MHPs) [10-13].

Effective WMSD and MHP prevention requires employers to address both the physical and psychosocial sources of injury risk. This requires accounting for the demands of the entire job, rather than evaluating and modifying individual tasks [4,8]. Workforce participation in all steps of the risk management process has also been identified as important to reducing WMSD risk exposure [14-16]. Multifactorial interventions, that is, multiple simultaneous actions to address sources of risk, selected from multiple levels of the hierarchy of controls, have also been supported as effective in reducing WMSD risk [17-20]. Multifactorial interventions are also supported for prevention of MHPs [21-23]. However, employer practices continue to be misaligned with the evidence regarding effective WMSD and MHP prevention [5,6,24,25].

Previous research has identified that employers' approaches to both physical and psychosocial hazard and risk management are typically individualistic rather than system based. These studies have found that implemented changes rarely address the source of injury risk and are therefore unlikely to be effective at preventing WMSDs and MHPs [23]. A review of available WMSD risk management tools [26] identified only 3 comprehensive tools (ie, included both physical and psychosocial risk factors), that also used participative methods to facilitate all stages of the risk management process.

One of the 3 tools identified in the review [26] was "A Participatory Hazard Identification and Risk Management" (APHIRM) toolkit. This is a set of procedures that guides employers and workers through the 5 stages of the risk management cycle to address the sources of WMSD and MHP risk specific to a particular job. These procedures are implemented with guidance from a facilitator who has been trained to implement the toolkit. The procedures focus on implementing changes that address the sources of risks of WMSDs and MHPs. Studies evaluating the toolkit have supported its effectiveness in reducing the severity of exposure to physical and psychosocial hazards that can contribute to WMSDs and MHPs [27-29].

Effective WMSD and MHP risk reduction relies on management support, participation, and commitment of resources to implement identified actions [5,14,30,31]. Aligning approaches with management and workforce readiness for change has been suggested to improve the outcomes resulting from these changes [19,32-34]. The Stages of Change model (SoC) [35] is one approach to adapting the delivery of interventions in the workplace to the recipient's readiness to change [32,36]. This model classifies individuals into one of five SoC: (1) precontemplation (unaware or unconcerned about workplace hazards), (2) contemplation (considering change but not yet ready to act), (3) preparation (intent to change in the near future), (4) action (made changes in the previous 6 months), and (5) maintenance (made changes and are working to consolidate gains and avoid relapse).

An individual's SoC is determined by their current knowledge, attitudes, and beliefs regarding the issue in focus. The key constructs thought to influence movement through the stages are decisional balance, that is, weighing the pros and cons of changing, and habit strength, that is, the ability to sustain a target behavior over time [37,38]. A scoping review [39] identified few intervention studies performed in a workplace setting that used the SoC model. Previous studies on WMSD prevention [36,40,41] have used the SoC model to evaluate whether advice and interventions tailored to the target's SoC are more effective at addressing WMSD hazards in the workplace. In this earlier work, both the managers' and employees' SoC were evaluated simultaneously, using a short questionnaire. Those in earlier SoC made a greater number of successful ergonomics changes when advice was tailored to their SoC. This suggests that such an approach has promise for improving the effectiveness of interventions aimed at making ergonomics changes in the workplace. Less attention has been paid specifically to the relationship between managers' or decision-makers' SoC (as a proxy for readiness to make changes) and the success of ergonomics interventions implemented in their area of control [19,31,36].

There is an opportunity to implement the APHIRM toolkit in a large retail organization and assess its impact on WMSD and MHP hazard exposure. The process evaluation conducted in this study will inform more effective toolkit implementation in organizations of similar scale and/or risk profile, accounting for managers' SoC. Implementation science methodologies have been recommended as an approach to a process evaluation [26,42,43], to improve robustness, while maintaining a "real-world" context. This study intends to build on available evidence relating to effective WMSD and MHP risk management obtained from previous workplace-based studies, particularly those which have implemented the APHIRM toolkit. There are few published studies that describe the implementation of the APHIRM toolkit in large workplaces [44,45].

XSL•FO
RenderX

Study Objectives

To address these gaps, the aims of this research are to evaluate the impact of APHIRM toolkit implementation on hazard exposure levels for workers in the online shopping department (online department), conduct a process evaluation of the toolkit implementation in each participating work group, and evaluate the influence of SoC on the use of APHIRM toolkit resources.

To achieve these aims, this research protocol is designed to answer the following questions: does implementation of the APHIRM toolkit reduce hazard exposure levels for workers in the online department? When implementing the APHIRM toolkit, what are the key requirements to maximize its effectiveness in reducing hazard exposures for workers in the online department? Are APHIRM toolkit resources used differently when senior leaders have different SoC?

Methods

Ethical Considerations

Ethics approval was granted by the La Trobe University Human Research Ethics Committee in July 2025 (approval HEC25088). Survey participants were provided with consent information before commencing the relevant surveys. These included the SoC survey for managers and the APHIRM toolkit survey for online team members at the stores included in the study. Survey response data were anonymized as part of the collection process. Participants in the risk management teams (RMTs) of the intervention group were provided with a participant informed consent survey prior to taking part in the study, as were participants taking part in interviews for the process evaluation component of the study. RMT data were deidentified at the completion of data collection. Participants completed activities related to the study during paid work time but were not provided with additional compensation for their participation.

Registration

The trial has been registered through the Open Science Framework (osf.io/82r9g). Registration was performed after completion of recruitment and following commencement of data collection, but before data were viewed by the research team. Data collection commenced prior to finalization of trial registration to ensure that key activities in the intervention did not take place over the busy Christmas and new year period. Attempting to do so would have likely threatened survey response rates and stores' level of participation in the intervention overall. Although ideally trial registration would have been completed prior to commencement of data collection, this outcome reflects the challenges associated with conducting research in a live workplace, including maintaining project momentum and sustaining organizational stakeholder engagement and motivation.

Trial Design

This is a cluster quasi–randomized controlled trial. Each cluster comprises the employees and managers of an online order fulfilment department (known hereafter as the "online department") in an individual store in a large retail organization. Clusters are described as "work groups" for the purposes of this study. The ratio of intervention work groups to control work groups is 1:1.

The APHIRM toolkit is designed so that nonexperts (ie, individuals other than ergonomists) can implement WMSD and MHP risk reduction measures in their workplaces. On this basis, a facilitator from the organization's health and safety team is assigned to guide toolkit implementation for each work group in the intervention arm. Assignment of facilitators to work groups is based on their geographic location, relative to the location of the store. Facilitators are overseen by the organization's ergonomics team, who monitor adherence to toolkit process and procedure. The organization's ergonomics team are, in turn, overseen by the lead author, who is also employed by the organization in a health and safety role. The lead author is also a member of the research team. The research team monitors the conduct of the study, including compliance with ethics requirements, and guides a robust approach to the study and the subsequent evaluation. This structure is intended to achieve a study design that most closely resembles the real-world conditions in which this intervention would typically occur. A lack of blinding at the cluster level could be considered a threat to the validity of the study; however, discourse of the research-to-practice gap in musculoskeletal disorder (MSD) prevention research [15,33,46] suggests a need for published data on interventions that most resemble a real-world implementation environment. This includes the conduct of research in real workplaces, on real workers, facilitated by those with the typical skill and expertise that would be encountered in a real-world setting [18]. This design reflects a previously identified challenge for workplace safety and ergonomics practitioners, to translate methodologically "correct" research into evidence-based prevention practice in the scope of their roles [23]. The risks of assessment bias are mitigated through review of data collection, analysis, and interpretation by the other members of the research team, who are not part of the intervention.

Participants

Eligible clusters ("work groups") are those stores with an online department, located within 1.5 hours' drive of Sydney or Brisbane, Australia, which are the cities where the facilitators are based. Random sampling within eligible work groups will account for work groups with varying sales volumes and service offerings (eg, delivery to a customer's vehicle parked outside the store, loading orders onto trucks for home delivery, and/or collection of orders by a third-party delivery driver through a service such as Uber). Managers of work groups are recruited to the study if they are the manager of the online department or store associated with the individual work group or are a manager responsible for the group of stores that the individual store or work group sits within.

Facilitators are allocated to work groups in a manner that replicates real-world conditions for toolkit implementation. These facilitators work in the organization's health and safety team in the states where the intervention is conducted and possess a range of qualifications and experience, from a background in store management and operations through to

XSL•FO
RenderX

degree-qualified health and safety professionals with varying years of experience.

Recruitment

A total of 18 work groups (9 intervention and 9 control) will be recruited to the study. All work groups will be located within 1.5 hours' travel time of Sydney or Brisbane, Australia. Settings range from inner urban to semirural locations. The number of work groups required for the study was determined based on the average number of employees in each work group and typical response rates to employee engagement surveys in stores, which are usually 20%-30%.

Stores are recruited to the study through consultation with the resource capacity management unit within the organization. This unit determines which stores are already engaged in other activities which would limit their ability to participate. This unit will then facilitate communication with the management of the groups of stores deemed eligible for participation. This management team then facilitates communication with the stores in their group that are selected for participation via the randomization process. Recruitment of stores continues until the required sample size is achieved.

Individual participants within a work group are eligible for recruitment to the study if they work some or all their work time in the online department, are a manager of an eligible online department or the store, or a manager overseeing an eligible store or "group" of 10-11 stores. These individual employees are involved in the study through completion of the APHIRM toolkit survey and feedback and consultation processes on identified sources of risk and action plans. Recruitment materials and advertising will be distributed through channels only accessible to those employees working in the online department. This is to prevent inadvertent recruitment of employees working in other departments in the store. These channels include the digital discussion board used to communicate with online employees and the posting of physical flyers on the noticeboards in the online room of the store. Informed consent statements and information on the study are provided to participants when they access the online portal used to administer the toolkit.

Recruitment of managers will be conducted through meeting invitations for briefing sessions aimed at communicating the study or providing updates about the study's progress. Informed consent statements are provided to these managers at the commencement of the sessions.

RMT members for each work group are recruited through the same channels as individual participants in the study. Each work group is encouraged to recruit a minimum of 3 RMT members, which is intended to ensure that at least 2 RMT members are available to participate in each of the monthly facilitated sessions in store.

Intervention

Clusters in the experimental group will implement the APHIRM toolkit processes and procedures with the guidance of a facilitator. The APHIRM toolkit guides work groups through the steps of the risk management cycle to address the sources of physical and psychosocial hazards that contribute to WMSDs and MHPs. Each step of toolkit implementation is delivered as a face-to-face session lasting up to 2 hours. Resources used to facilitate these sessions are freely available on the official website [47].

The first step of the process involves identifying members of the work group to act as representatives for their peers during toolkit implementation. This group is known as an RMT.

In the second step, the work group is invited to complete a survey that rates exposure to physical and psychosocial hazards, musculoskeletal pain or discomfort, and stress levels. The survey forms the hazard identification and risk assessment component of the risk management process. This survey is administered through an online portal, hosted as part of the APHIRM toolkit on its official website [47]. This enables the completion of surveys on any computer or mobile device. The survey is anonymous at the individual level, with responses aggregated to the work group (store) level.

In the third step, the results of the survey are provided to the work group, who then provide input on the work-related sources of risk relating to the identified hazards. Over the course of 10 monthly face-to-face sessions, the facilitator guides the RMT through the development and implementation of an action plan to address the sources of risk. It is anticipated that RMTs will identify a combination of actions that are specific to their individual store, as well as actions that are implemented in every store and thus require assistance from personnel above the individual store level to address.

The implementation of actions at the store level is expected to assist in differentiating intervention stores from control stores when measuring hazard severity at follow-up. It is also possible that the act of engaging RMTs from the intervention stores in the development of above-store actions will have an impact on hazard severity scores at follow-up, greater than the impact of the actual implementation of changes to the system of work. The impacts of these changes are expected to be felt by control and intervention stores equally if they are implemented before follow-up data collection. Many changes at the system of work level are unlikely to be implemented before follow-up, further allowing for differentiation between control and intervention groups. If significant changes are implemented in both control and intervention stores during the study, and it is determined that these changes may influence the results of the study, an additional survey will be conducted following the implementation of this change. This will be followed by the planned survey at the 12-month follow-up. This 12-month follow-up comprises the fourth step of the intervention, in which the survey is repeated approximately 12 months after the initial survey. These results are also communicated to the work group, along with an invitation to provide feedback on their perceived effectiveness of the action plan.

In parallel with implementation of the APHIRM toolkit, managers who are responsible for the online department for each intervention work group will be invited to participate in a questionnaire to evaluate their SoC. This will be conducted at commencement of the intervention, then every 3 months during the intervention, and on conclusion of the study, approximately 12 months following commencement.

XSL•FO
RenderX

Clusters in the control arm do not implement the APHIRM toolkit. These work groups only complete the APHIRM toolkit survey at commencement and conclusion of the study, and the managers associated with these work groups complete the SoC assessment every 3 months during the intervention period. These work groups otherwise receive the usual care and safety support from their management team, including the safety professionals who usually support the work group. This includes conducting a monthly safety meeting with the work group's safety committee, performing safety inspections, and participating in other routine activities relating to managing workplace health and safety in the store. Work groups assigned to the control arm are not prevented from implementing changes to address sources of MSD or MHP risk.

Figure 1 illustrates the structure and timing of the intervention and control work groups' involvement in the study for the purposes of providing comparison data.

Figure 1. Structure and timing of intervention activities. APHIRM: A Participatory Hazard Identification and Risk Management; RMT: risk management team.

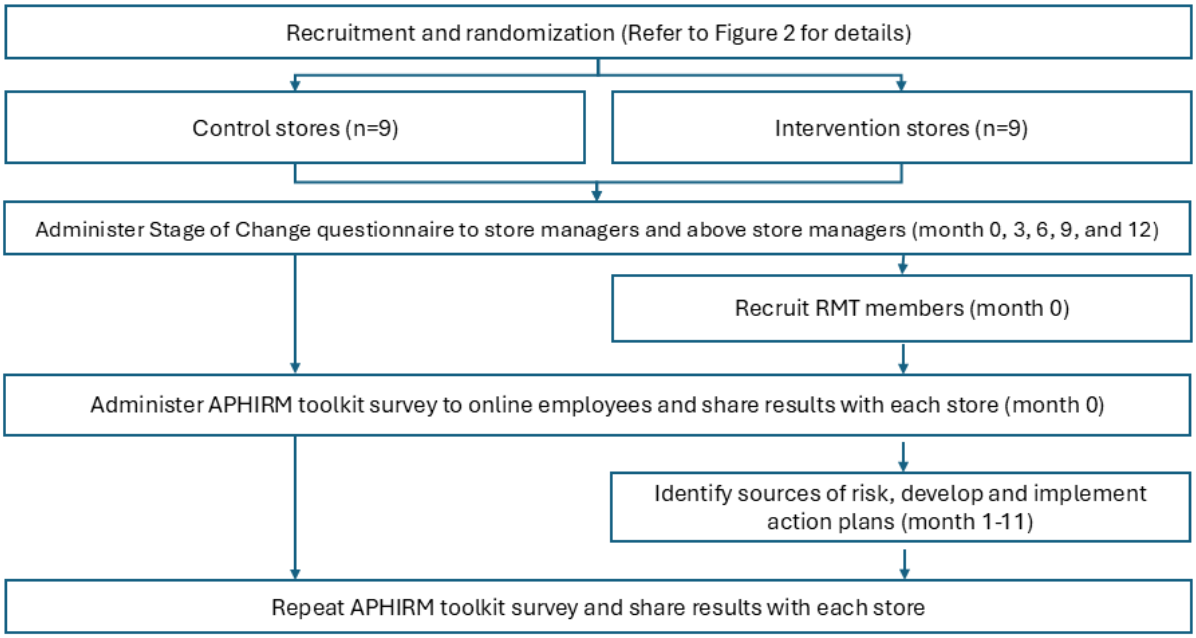

Facilitators from the organization's health and safety team are trained in the administration of the intervention by a member of the research team. This includes training on the procedures to facilitate each of the sessions that comprise the intervention, as well as data collection and ethics compliance procedures. The training is delivered using the resources provided in the APHIRM toolkit. One member of the research team and one member of the health and safety team have completed the 1-day APHIRM training workshop offered by La Trobe University on establishing and delivering the APHIRM toolkit in workplaces.

Adherence to toolkit procedures is monitored by reviewing the outputs of the work group, which are recorded in the work group's toolkit portal on APHIRM's official website [47]. Facilitators use an online notebook to record observations from each session, including how clusters use the resources from the APHIRM toolkit, fidelity to the toolkit procedures, and other observations that are important to the process evaluation completed as part of the study.

If the data provided at the conclusion of a session indicate that procedures may not have been correctly followed, a member of the research team will complete a follow-up session with that facilitator and the work group in question to review and prepare the work group for the next session and step in the intervention.

Outcomes

Quantitative Outcomes

Quantitative data are collected anonymously at the individual level and are therefore aggregated and analyzed at the work group level. Baseline and follow-up data are collected at the same time of the year (August/September), which assists in accounting for seasonal variation and avoids potential impacts on response rates by avoiding busy trading periods.

Primary Outcomes

Previous research has confirmed the validity of the toolkit survey [29] and generated evidence to support the toolkit's effectiveness at reducing exposure to hazards that contribute to WMSDs and MHPs [27,28]. The primary outcome measures for this study are changes in the mean severity ratings for physical and psychosocial hazards, based on the relevant items from the APHIRM toolkit survey. These data are self-reported by members of the work group. Mean severity ratings for each group of hazard items are calculated at the cluster level (ie, each work group), by calculating the difference between baseline and follow-up means. All primary outcome measures will be assessed for mean change by cluster, which will be analyzed using ANCOVA, comparing outcomes for the 2 arms. Outcomes for each work group will also be compared to identify potential outliers. Data will be cleaned and examined for normal distribution prior to analysis. Differences in cluster sizes will

XSL•FO
RenderX

be identified and accounted for using suitable analysis, which will be reported with the results.

Calculation of Severity Ratings for Physical Hazard Measures

A total of 12 survey items relate to physical (biomechanical) hazards that may influence MSD risk. Hazard measures include items related to force, repetition, and awkward and sustained postures, and vibration. Respondents rate their exposure to each hazard, in the last 6 months, on a scale from 1 (almost never) to 5 (almost always). A mean rating is calculated for all 12 survey items relating to physical hazard exposures.

Calculation of Severity Ratings for Psychosocial Hazard Measures

A mean rating for each cluster is calculated for all 44 survey items relating to psychosocial hazards. Examples of psychosocial hazard measures include items relating to control over the pace and content of work, role clarity, and reward and recognition.

Secondary Outcomes: Score for Self-Rated Pain and Discomfort

Using the measures in the APHIRM toolkit survey, participants rate the frequency and severity of their musculoskeletal pain and discomfort (in the last 6 months) in each of 5 body regions: neck/shoulders, arms, hands/fingers, middle/lower back, and hips/bottom/legs/feet. For example, "In the last six months, how often have you felt discomfort or pain in your neck or shoulders?" Response options range from 0 (never) to 4 (almost always). Any response equal to or greater than 1 (occasionally) then generates an additional question, for example "In the last six months, how bad was the discomfort or pain in your neck or shoulders?" Response options range from 1 (mild) to 3 (severe). A score out of 12 (4 × 3) is calculated for each body region and summed to produce a score out of 60 for each participant reporting any pain or discomfort. The change in mean scores between baseline and follow-up is used to evaluate outcomes.

Self-Rated Stress Score

A total of 12 of the APHIRM toolkit survey items are used to calculate a stress score. Eleven of these 12 items are from the Copenhagen Psychosocial Questionnaire [48]. The items relate to stress ("problems relaxing," "been tense," and "been irritable"), cognitive stress ("difficulty with thinking clearly," "concentrating," "making decisions," and "remembering"), and burnout ("worn out," "physically exhausted," "tired," and "emotionally exhausted"). The 12th survey item used to calculate a stress score is "had difficulty in falling or staying asleep." This is from the General Well-Being Questionnaire [49], identified as a valid measure for stress in a workplace context, with more brevity than the equivalent items used in the Copenhagen Psychosocial Questionnaire [29]. Response options range from 0 (never) to 4 (almost always). A score out of 48 is calculated for each respondent by multiplying the ratings for each item (on a scale of 0 to 4). The change in mean scores between baseline and follow-up is used to evaluate outcomes.

Action Plan Implementation

Data on action plan implementation will be collected via inputs and updates to the action plan created and updated for each store and managed through the toolkit portal. Actions will be classified as "completed," "in progress," or "not started" through the review of the action plan during monthly sessions conducted with the facilitator and the work group. Actions will also be classified as "higher-order" actions, involving hazard elimination, substitution, or engineering actions, or "lower-order" actions, involving administrative actions such as process changes.

Actions will be classified as "store-level" or "above-store-level" actions by the facilitation team, partly to determine which actions require escalation above the store's management to be addressed, and partly to evaluate how many store-level actions achieve "completed" status, compared with actions classified as above-store-level. This data will be validated by the research team through an independent review of action plan data collected by the APHIRM toolkit online portal. The count of each action and its classification in each of the above dimensions at the end of the 12-month study for each cluster will be reported.

Additional Outcome Measure

Managers' SoC progression will be measured using the questionnaire included in Multimedia Appendix 1. The questionnaire will be administered every 3 months for the duration of the study via scheduled management meetings. Measurement is conducted for each relevant manager group: all the managers directly responsible for a work group in a store (online department and store managers), and managers with responsibility for groups of stores, known as "above-store managers." Because of this, data for a manager with responsibility above a single store level may be considered when evaluating the implementation process for multiple stores under their direction, while SoC data for a store manager would only be considered in the process evaluation for the individual store under their management.

Movement through the SoC at each timepoint will be reported, as well as the absolute difference in stages, if any, from baseline to follow-up (ie, the total number of stages the work group's managers moved through during the study period and whether movement was progressive or regressive through the SoC). This survey tool was selected as it has been previously validated as an instrument sensitive to detecting both workers' and managers' readiness to implement changes to address physical and psychosocial risk factors in the workplace [36,50,51]. SoC data will be collected and managed using REDCap (Research Electronic Data Capture; Vanderbilt University) electronic data capture tools hosted at La Trobe University [52,53]. REDCap is a secure, web-based software platform designed to support data capture for research studies.

A summary of outcome measures, analysis metrics, methods of aggregation, and measurement timepoints is included in Table 1.

XSL•FO
RenderX

Table 1. Study measures, metrics, methods of aggregation, and frequency of measurement.

Primary or secondary	Outcome measure	Analysis metric	Method of aggregation	Participants providing data	Timepoints
Primary	Physical hazard measures	Change from baseline	Mean	Individual employees from each work group	At commencement and 12 months after commencement
Primary	Psychosocial hazard measures	Change from baseline	Mean	Individual employees from each work group	At commencement and 12 months after commencement
Secondary	Change in self-rated pain and discomfort	Change from baseline	Mean	Individual employees from each work group	At commencement and 12 months after commencement
Secondary	Change in self-rated stress score	Change from baseline	Mean	Individual employees from each work group	At commencement and 12 months after commencement
Secondary	Action plan implementation	Number of actions implemented by action status or type	Count	Risk management teams from intervention stores	At the conclusion of the study
Additional	Stage of change progression	Change from baseline	Difference between initial and final values; changes between each measurement period	Store managers (allocated directly to a store or work group); above-store managers (overseeing multiple work groups)	At commencement, and every 3 months after commencement

Process Evaluation

Data will be collected to enable a process evaluation of the implementation of the toolkit in intervention work groups. Observations and reflections will be recorded by facilitators after each interaction with participating clusters using an online notebook and synthesized to enable evaluation of the intervention using the "Reach, Effectiveness, Adoption, Implementation, and Maintenance" framework [54]. Observation themes will include the extent and nature of engagement of participants; reflections and feedback on participants' comments about the intervention process and their satisfaction with the process; participants' willingness to adopt the toolkit process and identified action plans; adherence to and deviation from toolkit processes; maintenance of implementation of toolkit processes; and actions that were implemented, to the extent possible considering the duration of the study. Observations will be recorded after each face-to-face meeting with each cluster, noting which resources from the APHIRM toolkit were used and how the resources were used. Interactions with members of the RMT or store-level managers for each cluster will be noted. Observations for above-store managers will be either linked to the clusters relevant to the observation being recorded or recorded as a separate record where there is no clear link between the interaction and specific clusters. These observations will be recorded in an online notebook by facilitators and monitored and reviewed by the research team.

This data will also be used to evaluate how different clusters use the resources provided in the APHIRM toolkit and how this relates to the cluster's SoC.

Sample Size

Sample size calculations were based on primary outcome measures of changes in physical and psychosocial hazard severity ratings. Considerations were made to the organization's structure when determining the number of clusters and participant numbers in each cluster. Individual stores are organized into "groups" of 10-11 stores, overseen by a group manager, justifying a 3-level cluster design.

Online department numbers were reviewed by store to determine expected cluster size. Additionally, response rates to previous surveys (not connected with the current project) were reviewed and found to be between 20% and 30%. These values were then used to calculate the sample size within each cluster to account for differences in expected recruitment across clusters. The mean cluster size (number of expected participants in a cluster) was set at 15, with a cluster variation of 0.7. If the mean cluster size deviates significantly from the estimated value of 15, this will be identified and accounted for in the statistical analysis and the interpretation of results.

Table 2 sets out the sample size and power calculations for primary and secondary outcomes evaluated in the randomized controlled trial, taking into account the 3-level cluster design at α=.05, intraclass correlation coefficient level 2=0.10; intraclass correlation coefficient level 3=0.05, mean cluster size=15; and estimated cluster variation=0.7, derived using Monte Carlo simulation. A previous study was used to estimate the effect size for exposure to physical hazards, exposure to psychosocial hazards, and self-rated pain and discomfort scores [55]. Results from Rahimi [56] were used to estimate effect sizes for stress scores.

On this basis, it was determined that 9 control and 9 intervention clusters (18 clusters in total) were required to achieve power of 0.8 or more.

XSL•FO
RenderX

Table 2. Sample size and power calculations.

Outcome	Mean (SD)	Effect size (*d*)	Groups (level 3), count	Stores (level 2), count	Employees (level 1), approximate n	Power (95% CI)
Primary outcomes						
Exposure to physical hazards	3.7 (0.6)	0.5	3	18	270	0.870 (0.816-0.910)
Exposure to psychosocial hazards	2.4 (0.6)	0.5	3	18	270	0.870 (0.816-0.910)
Secondary outcomes						
Pain/discomfort	14.9 (8.3)	4	3	18	270	0.780 (0.753-0.805)
Stress	4.02 (2.56)	0.5	3	18	270	0.742 (0.714-0.768)
Burnout	4.76 (2.54)	0.5	3	18	270	0.734 (0.706-0.760)
Cognitive stress	3.33 (2.57)	0.5	3	18	270	0.745 (0.717-0.771)

Randomization

The exclusion criteria were applied to a complete list of "groups" of stores in the organization prior to randomization. Groups of stores were excluded where the number of eligible stores in the group fell below threshold. Figure 2 illustrates participant flow by cluster for recruitment to the study.

Figure 2. Participant flow for recruitment to the study.

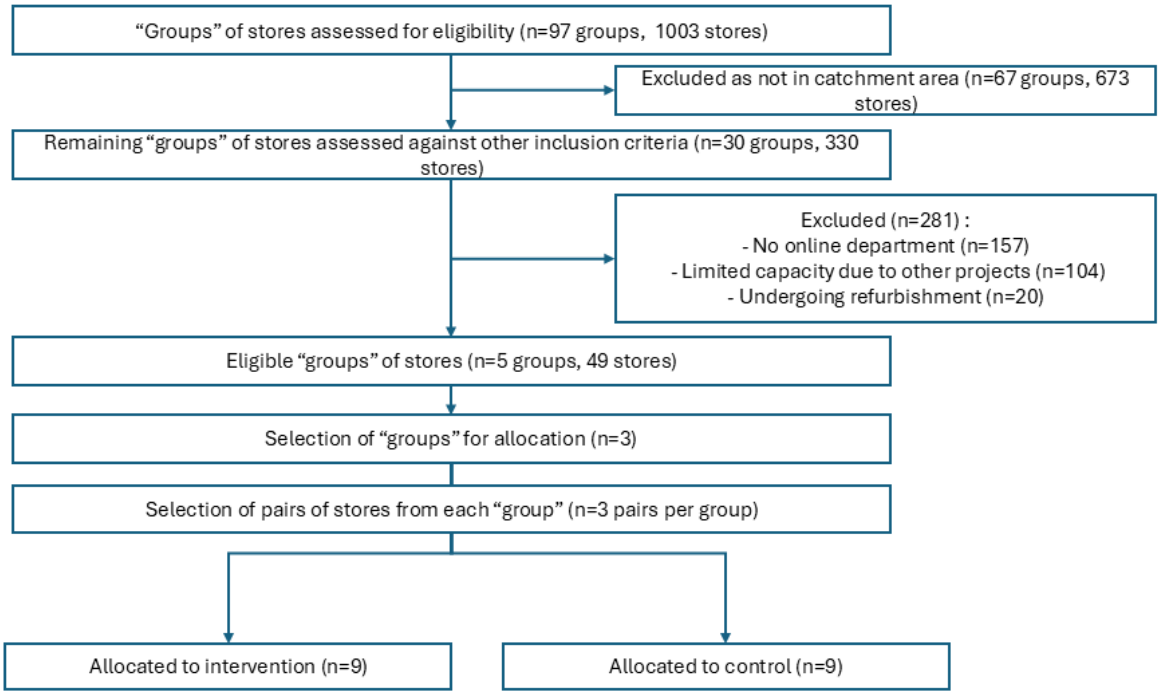

A list of eligible groups of 10-11 stores was provided to an independent researcher to perform randomization using a random number generator based on the stores' numeric identifier and assign them to the control or intervention arm of the study.

A total of 3 groups of stores that met the eligibility criteria were randomly selected first, and then 6 stores were randomly selected from each of those groups. Clusters (stores) were randomly allocated to either the control or intervention arm. Pairs were not matched. Consent was not obtained at the time of recruitment and randomization of stores, as consent procedures will be followed when recruiting individual participants during data collection procedures, such as administration of the APHIRM toolkit survey or the SoC questionnaire.

Blinding

Blinding of cluster allocations was not possible, as each cluster communicates with the other clusters in the study on a regular basis and would thus be able to determine their allocation in the study.

Raw data from the toolkit will be reviewed by a statistician on the research team who is blinded to cluster allocation.

Statistical Methods

Baseline characteristics, including baseline values of the primary outcomes, store size (number of employees), and available aggregate demographic measures, will be summarized at the store level and presented descriptively by treatment arm (mean, SD, median, range, or proportion, as appropriate). Data for each

XSL•FO
RenderX

cluster will be analyzed to identify potential outliers. Data will be cleaned and examined for normal distribution prior to analysis. The primary outcome measures (physical and psychosocial hazard severity ratings) will be analyzed by variance-weighted cluster-level ANCOVA using posttreatment store means regressed on treatment assignment and baseline store mean, with group fixed effects. Weights will be set to the inverse of the estimated theoretical variance of the cluster means to account for potential unequal cluster sizes. Because there are only 18 clusters (9 per arm) and 3 higher-level strata, estimation of higher-level variance components is unstable, and asymptotic cluster-robust SEs perform poorly. However, analyses using generalized estimating equations will be presented as sensitivity checks, with full disclosure of small-sample limitations.

Data on action plan implementation will be analyzed as categorical data. The count of actions by each status category (eg, "completed," "in progress," or "not started") and categorization as "higher-order" or "lower-order" will be performed for each cluster and compared against the other outcome measures for that cluster. Data on SoC progression will be analyzed as ordinal data. The median and mode response for the cluster will be determined as the SoC for that cluster at each measurement point. SoC data for store managers and above-store managers will be analyzed separately for each cluster. Progression through the SoC will be evaluated by assessing the number of SoC each group progresses through at each timepoint and overall, from commencement to the conclusion of the study.

Missing Data

The APHIRM toolkit survey requires completion of each item before the next one is presented, so a missing data strategy is not required. The SoC questionnaire is also managed in the same way. Consequently, a plan for handling missing data with respect to these data is not required.

Facilitators submit a completed worksheet at the end of each session completed with each cluster. This is used to evaluate adherence to protocol and to investigate and document any deviations from the protocol. These will be documented in the records maintained for each cluster, which are used to complete the process evaluation for this study.

Results

Funding for the study was approved in June 2025. Ethics approval for the study was granted in July 2025. Recruitment and randomization commenced in August 2025. This included identification of eligible stores and randomization of the eligible stores to the control or intervention groups until the required sample size was achieved. Data collection commenced in September 2025. These timings resulted from a necessary operational requirement to avoid key study activities coinciding with busy trading periods, such as Christmas and Easter, without delaying the commencement of the study. As of January 2026, a total of 18 stores were recruited to the study, with 332 individuals responding to the APHIRM toolkit survey. Response rates to the APHIRM toolkit survey for each cluster range between 28% and 85%, with an average response rate of 47% of all eligible employees. Data have not been viewed by the research team. The intervention is expected to conclude in September 2026, and results are anticipated to be published in early 2028.

Discussion

Generalizability

WMSDs and MHPs have a significant impact on Australian workers and workplaces, including the retail trade sector [57]. Recommendations from research on WMSD and MHP prevention advocate for a participatory approach, focusing on the whole job rather than discrete tasks, and taking into account both physical and psychosocial hazards [18,58,59]. Previous research has highlighted the evidence-to-practice gap in prevention practices, including those used by safety professionals in the retail sector [5]. This study aims to translate research evidence into practice and address WMSD and MHP hazards through implementation of the APHIRM toolkit. This study seeks to understand what support managers and decision makers require to facilitate effective implementation of these practices in their area of control. This is an important evidence gap in understanding the manager's role in successful workplace interventions [36,41,60]. It is anticipated that these findings could be used to recommend future improvements to the design and content of supporting resources in the APHIRM toolkit.

The retail organization hosting this study is one of Australia's largest employers, with more than 1100 locations across the country. This provides an opportunity to examine the study's research questions at a level of organizational scale and complexity that is often not possible in Australia. The design also provides an opportunity to understand how to support large, multisite employers to bridge the evidence-to-practice gap with respect to WMSD and MHP risk management. This is important for employers, who, due to organizational size, complexity, and geographic spread, must rely on nonexperts to implement and maintain risk management practices with respect to WMSD and MHP prevention. The design of the study endeavors to replicate real-world conditions of toolkit implementation as closely as possible so that findings can be generalized to other work groups in the organization as much as possible. However, the availability of staff in the organization to act as facilitators, and their location—proximal to state capital cities—may present a limitation to the generalizability of findings to regional and remote work group locations.

As part of identifying recommendations to improve these practices, this study provides an opportunity to understand how managers at different levels of a large organization, and in different physical locations, progress through SoC during implementation of the APHIRM toolkit. The study also presents an opportunity to understand how managers' progression through SoC relates to outcomes from toolkit implementation. It is anticipated that this study could also inform how the APHIRM toolkit could be best used, and possibly adapted, for use in large multisite organizations, in ways that support managers of work groups at different SoC. The intervention strategy aims to replicate how the toolkit may be implemented within a large organization, which serves to build the capacity

XSL•FO
RenderX

of individual stores, their leadership teams, and their supporting safety professionals to more effectively manage the risks of WMSDs and MHPs. It is anticipated that the findings of the study could be applied to other large, multisite organizations, particularly in the Australian retail sector.

Limitations

Limitations on facilitators' ability to travel and other time constraints prevent the recruitment of work groups located in regional or remote locations, which typically encounter different operational challenges compared with those located closer to a capital city. This should be considered when translating the results of this study to work groups in these geographies.

All work groups in this study are recruited from the same organization. As such, changes at a whole-of-systems level could possibly impact control and intervention stores alike, thereby influencing the results of this study. This presents a challenge when attempting to attribute changes in outcome measures to any one management level or team within the organization. However, it is anticipated that RMTs in stores will identify a combination of actions to address identified hazards and risks that are specific to their individual store, as well as actions that are applicable at the level of the system of work. Consequently, RMTs will require assistance from personnel above the individual store level to address system-related actions. The implementation of actions at the store level is expected to assist in differentiating intervention stores from control stores when measuring hazard severity at follow-up. The impacts of these systems level changes are expected to be felt by control and intervention stores equally, as it is anticipated these changes will be implemented at all stores. However, it is also highly likely that the majority of systems-level changes will not be implemented prior to the conclusion of the study. It is expected that the majority of any observed changes will be attributable to store-level changes, which can be implemented in a shorter amount of time.

It is also possible that the act of engaging RMTs from the intervention stores in the development of above-store actions will have an impact on hazard severity scores at follow-up, greater than the impact of the implementation of actual changes to the system of work.

Acknowledgments

The authors confirm that no generative artificial intelligence tools were used in any part of the preparation of this manuscript.

Funding

This research is supported by an Australian Government Research Training Program (RTP) scholarship. La Trobe University provided support for the article processing fee associated with this manuscript.

Authors' Contributions

EC contributed to conceptualization, development of the methodology, and writing of the original draft. VW contributed to conceptualization, development of the methodology, supervision, and writing through review and editing. JO contributed to conceptualization, development of the methodology, supervision, and writing through review and editing. KL contributed to development of the methodology and validation.

Conflicts of Interest

EC is employed by the host organization participating in this study.

Multimedia Appendix 1
Stage of change questionnaire for managers.
[DOCX File , 18 KB - resprot_v15i1e84864_app1.docx]

References

1. The Australian workers' compensation statistics (AWCS) 2018-2019. Safe Work Australia. Canberra; 2020. URL: https://www.safeworkaustralia.gov.au/doc/australian-workers-compensation-statistics-2018-19 [accessed 2026-01-10]
2. Interactive data website. Safe Work Australia. 2025. URL: https://data.safeworkaustralia.gov.au/interactive-data [accessed 2026-01-10]
3. Hauke A, Flintrop J, Brun E, Rugulies R. The impact of work-related psychosocial stressors on the onset of musculoskeletal disorders in specific body regions: a review and meta-analysis of 54 longitudinal studies. Work Stress 2011;25(3):243-256. [doi: 10.1080/02678373.2011.614069]
4. O'Reilly H, Van Eerd D. Including the ergonomist's voice in integrating MSD prevention and psychological health and safety: challenges, tools, and considerations. Appl Ergon 2025;122:104405 [FREE Full text] [doi: 10.1016/j.apergo.2024.104405] [Medline: 39454318]
5. Robertson J, Jayne C, Oakman J. Work-related musculoskeletal and mental health disorders: are workplace policies and practices based on contemporary evidence? Saf Sci 2021;138:105098 [FREE Full text] [doi: 10.1016/j.ssci.2020.105098]

XSL•FO
RenderX

6. Weale V, Stuckey R, Kinsman N, Oakman J. Workplace musculoskeletal disorders: a systematic review and key stakeholder interviews on the use of comprehensive risk management approaches. Int J Ind Ergon 2022;91:103338. [doi: 10.1016/j.ergon.2022.103338]
7. Afsharian A, Dollard MF, Glozier N, Morris RW, Bailey TS, Nguyen H. Work-related psychosocial and physical paths to future musculoskeletal disorders (MSDs). Saf Sci 2023;164:106177 [FREE Full text] [doi: 10.1016/j.ssci.2023.106177]
8. Macdonald W, Oakman J. Requirements for more effective prevention of work-related musculoskeletal disorders. BMC Musculoskelet Disord 2015;16(1):293 [FREE Full text] [doi: 10.1186/s12891-015-0750-8] [Medline: 26466897]
9. Stewart S, Rothmore PR, Doda DVD, Hiller JE, Mahmood MA, Pisaniello DL. Musculoskeletal pain and discomfort and associated worker and organizational factors: a cross-sectional study. Work 2014;48(2):261-271. [doi: 10.3233/WOR-131622] [Medline: 23531584]
10. Bakker AB, Demerouti E. Job demands-resources theory: taking stock and looking forward. J Occup Health Psychol 2017;22(3):273-285. [doi: 10.1037/ocp0000056] [Medline: 27732008]
11. Demerouti E, Bakker AB, Nachreiner F, Schaufeli WB. The job demands-resources model of burnout. J Appl Psychol 2001;86(3):499-512. [doi: 10.1037/0021-9010.86.3.499]
12. Schulte PA, Sauter SL, Pandalai SP, Tiesman HM, Chosewood LC, Cunningham TR. An urgent call to address work-related psychosocial hazards and improve worker well-being. Am J Ind Med 2024;67(6):499-514. [doi: 10.1002/ajim.23583] [Medline: 38598122]
13. Stansfeld S, Candy B. Psychosocial work environment and mental health--a meta-analytic review. Scand J Work Environ Health 2006;32(6):443-462 [FREE Full text] [doi: 10.5271/sjweh.1050] [Medline: 17173201]
14. Haines H, Wilson J, Vink P, Koningsveld E. Validating a framework for participatory ergonomics (the PEF). Ergonomics 2002;45(4):309-327. [doi: 10.1080/00140130210123516] [Medline: 12028727]
15. Hansen AF, Hasle P, Caroly S, Reinhold K, Järvis M, Herrig AO. Participatory ergonomics: what works for whom and why? A realist review. Ergonomics 2024;67(1):13-33. [doi: 10.1080/00140139.2023.2202842] [Medline: 37070935]
16. Wilson JR. Participation—A framework and a foundation for ergonomics? J Occup Psychol 2011;64(1):67-80 [FREE Full text] [doi: 10.1111/j.2044-8325.1991.tb00542.x]
17. Caponecchia C, Coman RL, Gopaldasani V, Mayland EC, Campbell L. Musculoskeletal disorders in aged care workers: a systematic review of contributing factors and interventions. Int J Nurs Stud 2020;110:103715. [doi: 10.1016/j.ijnurstu.2020.103715] [Medline: 32758908]
18. Cole DC, Van Eerd D, Bigelow P, Rivilis I. J Occup Rehabil 2006;16(3):359-374. [doi: 10.1007/s10926-006-9032-5] [Medline: 16933147]
19. Oakman J, Rothmore P, Tappin D. Intervention development to reduce musculoskeletal disorders: is the process on target? Appl Ergon 2016;56:179-186. [doi: 10.1016/j.apergo.2016.03.019] [Medline: 27184326]
20. Silverstein B, Clark R. Interventions to reduce work-related musculoskeletal disorders. J Electromyogr Kinesiol 2004;14(1):135-152. [doi: 10.1016/j.jelekin.2003.09.023] [Medline: 14759759]
21. Bailey TS, Dollard MF, McLinton SS, Richards PAM. Psychosocial safety climate, psychosocial and physical factors in the aetiology of musculoskeletal disorder symptoms and workplace injury compensation claims. Work Stress 2015;29(2):190-211. [doi: 10.1080/02678373.2015.1031855]
22. Lang J, Ochsmann E, Kraus T, Lang JW. Psychosocial work stressors as antecedents of musculoskeletal problems: a systematic review and meta-analysis of stability-adjusted longitudinal studies. Soc Sci Med 2012;75(7):1163-1174. [doi: 10.1016/j.socscimed.2012.04.015] [Medline: 22682663]
23. Oakman J, Macdonald W, Bartram T, Keegel T, Kinsman N. Workplace risk management practices to prevent musculoskeletal and mental health disorders: what are the gaps? Safety Science 2018;101:220-230 [FREE Full text] [doi: 10.1016/j.ssci.2017.09.004]
24. Chung AZ, Shorrock ST. The research-practice relationship in ergonomics and human factors--surveying and bridging the gap. Ergonomics 2011;54(5):413-429. [doi: 10.1080/00140139.2011.568636] [Medline: 21547787]
25. Dul J, Bruder R, Buckle P, Carayon P, Falzon P, Marras WS. A strategy for human factors/ergonomics: developing the discipline and profession. Ergonomics 2012;55(4):377-395 [FREE Full text] [doi: 10.1080/00140139.2012.661087] [Medline: 22332611]
26. Oakman J, Kinsman N, Weale V, Stuckey R. A qualitative exploration of tools used by WHS professionals for the prevention of musculoskeletal disorders. Safety Science 2022;149:105685 [FREE Full text] [doi: 10.1016/j.ssci.2022.105685]
27. Oakman J, Macdonald W. The APHIRM toolkit: an evidence-based system for workplace MSD risk management. BMC Musculoskelet Disord 2019;20(1):504 [FREE Full text] [doi: 10.1186/s12891-019-2828-1] [Medline: 31666054]
28. Oakman J, Macdonald WA, McCredie K. Psychosocial hazards play a key role in differentiating MSD risk levels of workers in high-risk occupations. Appl Ergon 2023;112:104053 [FREE Full text] [doi: 10.1016/j.apergo.2023.104053] [Medline: 37270872]
29. Validity of the APHIRM toolkit hazard and risk assessment survey. La Trobe University. 2025. URL: https://www.aphirm.org.au/_files/ugd/4d2769_2e40dea964374e68877fd4b77819d698.pdf [accessed 2025-08-14]

XSL•FO
RenderX

30. van der Molen HF, Sluiter JK, Hulshof CTJ, Vink P, Frings-Dresen MHW. Effectiveness of measures and implementation strategies in reducing physical work demands due to manual handling at work. Scand J Work Environ Health 2005;31 Suppl 2:75-87 [FREE Full text] [Medline: 16363450]
31. Zink KJ, Steimle U, Schröder D. Comprehensive change management concepts. Development of a participatory approach. Appl Ergon 2008;39(4):527-538. [doi: 10.1016/j.apergo.2008.02.015] [Medline: 18395184]
32. Haslam R. Targeting ergonomics interventions--learning from health promotion. Appl Ergon 2002;33(3):241-249. [doi: 10.1016/s0003-6870(02)00016-9] [Medline: 12164508]
33. Rothmore P, Karnon J, Aylward P. Implementation of interventions to prevent musculoskeletal injury at work — lost in translation? Phys Ther Rev 2013;18(5):344-349. [doi: 10.1179/1743288x13y.0000000092]
34. Whysall Z, Haslam C, Haslam R. A stage of change approach to reducing occupational ill health. Prev Med 2006;43(5):422-428. [doi: 10.1016/j.ypmed.2006.07.004] [Medline: 16908057]
35. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983;51(3):390-395. [doi: 10.1037//0022-006x.51.3.390] [Medline: 6863699]
36. Whysall Z, Haslam C, Haslam R. Implementing health and safety interventions in the workplace: An exploratory study. Int J Ind Ergon 2006;36(9):809-818. [doi: 10.1016/j.ergon.2006.06.007]
37. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994;13(1):39-46. [doi: 10.1037//0278-6133.13.1.39] [Medline: 8168470]
38. Whysall ZJ, Haslam C, Haslam R. Developing the stage of change approach for the reduction of work-related musculoskeletal disorders. J Health Psychol 2007;12(1):184-197. [doi: 10.1177/1359105307071753] [Medline: 17158852]
39. Condie E, Weale V, Oakman J. Workplace based interventions which apply the stages of change framework to their design: a scoping review. Appl Ergon 2025;130:104646 [FREE Full text] [doi: 10.1016/j.apergo.2025.104646] [Medline: 40974755]
40. Doda D, Rothmore P, Pisaniello D, Briggs N, Stewart S, Mahmood M, et al. Relative benefit of a stage of change approach for the prevention of musculoskeletal pain and discomfort: a cluster randomised trial. Occup Environ Med 2015;72(11):784-791. [doi: 10.1136/oemed-2015-102916] [Medline: 26304778]
41. Rothmore P, Aylward P, Karnon J. The implementation of ergonomics advice and the stage of change approach. Appl Ergon 2015;51:370-376. [doi: 10.1016/j.apergo.2015.06.013] [Medline: 26154235]
42. Guerin R, Glasgow R, Tyler A, Rabin B, Huebschmann A. Methods to improve the translation of evidence-based interventions: a primer on dissemination and implementation science for occupational safety and health researchers and practitioners. Saf Sci 2022;152:105763. [doi: 10.1016/j.ssci.2022.105763] [Medline: 37854304]
43. Wijk K, Mathiassen SE. Explicit and implicit theories of change when designing and implementing preventive ergonomics interventions--a systematic literature review. Scand J Work Environ Health 2011;37(5):363-375 [FREE Full text] [doi: 10.5271/sjweh.3159] [Medline: 21403982]
44. Macdonald W, Evans O, Armstrong R. Research on the Prevention of Work-Related Musculoskeletal Disorders : Stage 2 : a Study of a Small Sample of Workplaces in High Risk Industries. Bundoora: Centre for Ergonomics & Human Factors, Faculty of Health Sciences, La Trobe University; 2006.
45. APHIRM OHS risk management toolkit. La Trobe University, Centre for Ergonomics and Human Factors. 2025. URL: https://www.aphirm.org.au [accessed 2025-06-15]
46. Shorrock ST, Williams CA. Human factors and ergonomics methods in practice: three fundamental constraints. Theor Issues Ergon Sci 2016;17(5-6):468-482. [doi: 10.1080/1463922x.2016.1155240]
47. APHIRM Toolkit. URL: https://www.aphirm.org.au/ [accessed 2025-01-14]
48. Burr H, Berthelsen H, Moncada S, Nübling M, Dupret E, Demiral Y. The third version of the Copenhagen psychosocial questionnaire. Saf Health Work 2019;10(4):482-503. [doi: 10.1016/j.shaw.2019.10.002] [Medline: 31890332]
49. Wheeler RJ. The theoretical and empirical structure of general well-being. Soc Indic Res 1991;24(1):71-79. [doi: 10.1007/bf00292651]
50. Rothmore P, Aylward P, Gray J, Karnon J. A long-term evaluation of the stage of change approach and compensable injury outcomes - a cluster-randomised trial. Ergonomics 2017;60(5):628-635. [doi: 10.1080/00140139.2016.1199816] [Medline: 27284868]
51. Village J, Ostry A. Assessing attitudes, beliefs and readiness for musculoskeletal injury prevention in the construction industry. Appl Ergon 2010;41(6):771-778. [doi: 10.1016/j.apergo.2010.01.003] [Medline: 20110082]
52. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
53. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
54. Holtrop JS, Estabrooks PA, Gaglio B, Harden SM, Kessler RS, King DK. Understanding and applying the RE-AIM framework: clarifications and resources. J Clin Transl Sci 2021;5(1):e126 [FREE Full text] [doi: 10.1017/cts.2021.789] [Medline: 34367671]

XSL•FO
RenderX

55. Oakman J, Chan S. Risk management: where should we target strategies to reduce work-related musculoskeletal disorders? Safety Science 2015;73:99-105. [doi: 10.1016/j.ssci.2014.11.026]
56. Rahimi M, Arnold B, LaMontagne AD, Riley P. Validation and benchmarks for the Copenhagen psychosocial questionnaire (COPSOQ III) in an Australian working population sample. BMC Public Health 2025;25(1):830 [FREE Full text] [doi: 10.1186/s12889-025-21845-x] [Medline: 40025491]
57. Oakman J, Clune S, Stuckey R. Work-Related Musculoskeletal Disorders in Australia. Australia: La Trobe University; 2019.
58. Eatough EM, Way JD, Chang C. Understanding the link between psychosocial work stressors and work-related musculoskeletal complaints. Appl Ergon 2012;43(3):554-563. [doi: 10.1016/j.apergo.2011.08.009] [Medline: 21944295]
59. Macdonald WA, Oakman J. Changes needed to reduce risk of musculoskeletal disorders. Am J Ind Med 2024;67(7):575-581. [doi: 10.1002/ajim.23613] [Medline: 38752696]
60. Shaw K, Haslam CO, Haslam RA. A staged approach to improving health and safety in the workplace. Contemporary Ergonomics 2007:395-400. [doi: 10.1201/9781315106595-62]

Abbreviations

APHIRM: A Participatory Hazard Identification and Risk Management
MHP: mental health problem
MSD: musculoskeletal disorder
REDCap: Research Electronic Data Capture
RMT: risk management team
SoC: Stages of Change
WMSD: work-related musculoskeletal disorder

Edited by J Sarvestan; submitted 30.Sep.2025; peer-reviewed by S Mooren-van der Meer; comments to author 11.Nov.2025; revised version received 21.Dec.2025; accepted 22.Dec.2025; published 29.Jan.2026.

Please cite as:
Condie E, Weale V, Lambert KA, Oakman J
Implementation of a Participatory Ergonomics Intervention to Reduce Musculoskeletal and Stress-Related Mental Health Risks in Australian Retail Workers: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84864
URL: https://www.researchprotocols.org/2026/1/e84864
doi:10.2196/84864
PMID:

©Elise Condie, Victoria Weale, Katrina A Lambert, Jodi Oakman. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 29.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Brief Intervention for Discontinuing Inappropriate Z-Hypnotic Use Among Older Patients in Primary Care: Protocol for a Cluster Randomized Controlled Trial With a Single Crossover

Tahreem Ghazal Siddiqui[1], PhD; Maria Torheim Bjelkarøy[1], PhD; Tone Breines Simonsen[1], MSci; Maria Lie Selle[1], PhD; Christofer Lundqvist[2], MD, PhD

[1]Health Services Research Unit, Akershus University Hospital, Lørenskog, Norway

[2]Department of Health Services Research, Institute of Clinical Medicine, Campus Akershus University Hospital, University of Oslo, Lørenskog, Norway

Corresponding Author:
Christofer Lundqvist, MD, PhD
Department of Health Services Research
Institute of Clinical Medicine, Campus Akershus University Hospital
University of Oslo
PO Box 1000
Lørenskog, 1478
Norway
Phone: 47 47278881
Email: a.c.lundqvist@medisin.uio.no

Abstract

Background: Older patients are frequent users of Z-hypnotics despite consensus recommendations against extended use. Inappropriate Z-hypnotic use among older patients is frequently reported, posing risks of side effects and dependence. Interventions have been mainly at the population level and through prescription regulations. There are few instruments helping general practitioners (GPs) deal with inappropriate use among individual patients.

Objective: Through a randomized controlled trial (RCT), we aim to test the effectiveness of a behavioral brief intervention (BI) method used by trained GPs for reducing inappropriate Z-hypnotic use among their patients.

Methods: We will conduct a double-blind RCT with a single crossover. Patients (aged >60 years) on participating GPs' lists who are using Z-hypnotics inappropriately, do not have serious mental or physical disorders, and can provide valid informed consent are eligible. GPs randomized to the BI arm will be trained to administer the BI, and those randomized to business as usual (BAU) will not receive training. GPs' patient lists will be screened for inappropriate Z-hypnotic users through an electronic questionnaire. The GP will be informed of patients who should be given an appointment and administered the BI. Untrained GPs will continue BAU. Randomization-blinded outcome evaluation will be conducted at 6 weeks, 6 months, and 1 year in both the study groups.

Results: The main outcome is the proportion of patients with inappropriate Z-hypnotic use, comparing BI versus BAU, after 6 weeks. Secondary outcomes are cognitive function, pain, self-reported sleep evaluation, sleep efficiency (actigraphy) and quality of life, and change compared to baseline. We will also report on the characteristics of the screened GP patient population. Other variables are other medication use or polypharmacy, anxiety and depression, severity of dependence, and mortality.

Conclusions: If RCT-level evidence demonstrates the effectiveness of the BI for reducing inappropriate Z-hypnotic use among older patients without worsening of secondary outcomes, this could be a simple, transferable intervention to implement on a larger scale among GPs, other physicians, and health workers.

International Registered Report Identifier (IRRID): DERR1-10.2196/75670

(JMIR Res Protoc 2026;15:e75670) doi:10.2196/75670

KEYWORDS

Z-hypnotics; inappropriate prescription medication; older people; general practice

XSL•FO
RenderX

Introduction

Background

The use of central nervous system–depressant medications with addictive potential such as opioids, benzodiazepines, and benzodiazepinelike hypnotics (Z-hypnotics) is widespread, and prescription-based use of these has become an increasing problem in many countries. Prescription-based use of such medications may pose an even greater risk among older patients, as is expressed in medication use consensus criteria and guidelines for older persons, such as the Beers Criteria for Potentially Inappropriate Medication Use in Older Adults and the Norwegian General Practice criteria [1,2]. Nevertheless, inappropriate use among older patients is highly prevalent, and despite guidelines, there has been no overall reduction in use [3]. The dominant medication group of those mentioned above among older patients in general practice living at home or in nursery homes is Z-hypnotics, used by 8% to 25% of patients [4-6]. Several studies suggest a clear increase in use over time, with older patients dominating the trend [7].

Although some studies have suggested successful methods for deprescription of Z-hypnotics among older people as well as patients in primary care, there is no consensus on effective deprescription methods [8,9]. This represents a knowledge gap in the area of individual intervention by general practitioners (GPs) against inappropriate Z-hypnotic use, and there is a need to demonstrate the effectiveness of an intervention to do this. Most previous studies have focused on both benzodiazepine and Z-hypnotic use and used different interventions, such as written information brochures and face-to-face interventions by GPs and pharmacists [10,11]. Intervention contents have ranged from abrupt termination of use to tapering over time with support regimens with different types of content [8,9]. Even though there are several similarities between the use of benzodiazepines and of Z-hypnotics among older patients, justifying similar intervention, there are also several differences that may justify the use of different approaches. First, Z-hypnotics in Norway are usually prescribed not on a fixed regimen but on a part pro re nata schedule, with the patients having access to prescriptions but being partly themselves responsible for when to actually take the medication and when not to. This means that, rather than starting with formal deprescription by the GP, the logical first step, also involving patient enablement in our setting, may be to focus on the medication-related behavior of the individual patient without formally removing access to the prescribed medication itself. In addition, it has been shown that, with regard to the use of Z-hypnotics among older patients, dose escalation is rarely observed. Although this does not exclude the possibility of induced tolerance, it has some advantages regarding termination of use in that tapering is usually not necessary from the standard low doses of Z-hypnotics used. This greatly simplifies the communication of the intervention as the aim can usually be to recommend complete termination of intake. However, the other side of the coin is that the risk of rebound insomnia and possible worsening of other aspects of Z-hypnotic drug use must of course be carefully monitored in a study aiming to achieve termination of use. In our previous studies, although they were conducted in a hospital-recruited population, we have demonstrated that cognitive function, pain, and quality of life are associated with use of central nervous system active medication, the dominating one of which in our population was Z-hypnotics, thus justifying that a study focusing on achieving a change in the use of Z-hypnotics should also assess these aspects.

The Brief Intervention and Its Conceptual Framework

The original brief intervention (BI) based on initial screening followed by a short, structured intervention was developed for harmful alcohol use in the last decades of the 20th century. It was based on a validated screening instrument for harmful alcohol use and a structured set of communication skills adapted to be easily usable in any clinical setting, not least in primary care [12]. Similar interventions have since been adapted to many different settings and drugs and, most recently, to inappropriate use of prescribed or nonprescribed medication [9,13-16]. The original framework from the World Health Organization BI manual for alcohol and substance abuse [17] consisted of feedback responsibilities, advice, a menu of options, empathy, and self-efficacy [17-19], and in the preparatory work for this protocol, we adjusted this to our target population: older patients with Z-hypnotic use. The BI is communication based, with some elements from the motivational interviewing counseling method. The original BI manual provided some examples on parts of motivational interviewing in a consultation [17]. However, it does not fully cover the communication strategies and advice for framing information giving and exchanges between physicians and older patients. Thus, the BI was further adapted for the purpose of this protocol based largely on our preparatory work, and the method was piloted in a smaller study focusing on feasibility, logistics, and acceptability. We have previously adapted and successfully tested the BI scheme in interventions by GPs in the context of medication overuse headache (MOH) [20]. The BI for MOH proved to be an effective and acceptable tool for GPs and patients, thus enabling early intervention at the primary care level, where most patients with MOH are [20,21]. Our previous studies on the use of central nervous system active drugs among older patients admitted to hospitals from primary care showed that the use of such drugs is high in this population [22]. We have also shown their association with reduced cognitive function and quality of life, as well as with pain, addiction, and comorbidity [23-26]. The most common of these drugs in our hospital population in these studies were Z-hypnotics [22]. In Norway, health services are built based on the role of the GP as gatekeeper and primary health service contact for most home-living older patients. GPs are also responsible for prescribing the largest share of possibly addictive medication to these patients and, thus, are a key to interventions hoping to reach large groups of patients [27]. Even though GPs in Norway are trained to be restrictive regarding the prescription of such medications to older patients, not least based on the Norwegian General Practice (NORGEP) criteria, such use is increasing, with Z-hypnotics playing a dominant role [28]. On the basis of our previous experience, we have also adapted the BI scheme to users of Z-hypnotics, aiming to give structured advice enabling patients to reduce inappropriate use. The BI setup was piloted in a small case-series feasibility study, with

XSL•FO
RenderX

feedback supporting a focus on the GP setting also in this case [29]. Our previous successful experience with the BI method in other settings with individual inappropriate medication use, as well as both our previous pilot study and our background studies on factors associated with inappropriate medication use, in particular Z-hypnotics, among older patients, suggests that the BI method is a suitable and promising candidate for an intervention study.

In this paper, we describe the protocol for a randomized controlled trial of the BI versus business as usual (BAU). The protocol was developed following the UK Medical Research Council guidelines for developing and evaluating complex interventions [30]. Our primary objective is to evaluate the effectiveness of a BI, conducted by trained GPs, for reducing inappropriate Z-hypnotic use in older patients. The main and secondary effects of the BI will be assessed at the patient level.

Methods

Design

We will conduct a double-blind, randomized, single-crossover controlled trial comparing the BI with BAU. The data will be collected at baseline and 6 weeks, 6 months, and 12 months after baseline.

Trial Status

The first patient was recruited in October 2023. Patient recruitment has been completed as of May 2024. Data collection is ongoing until December 2024.

Setting

In the Norwegian health care system, the main point of entry for patients is via their GP, and it is intended that all patients in Norway be listed with a GP, who then has the primary responsibility for health care contacts, referrals to specialist health care, and prescriptions. The average GP list length is of approximately 1200 patients. GPs are, in most cases, self-employed, but the main funding comes from reimbursement from the national health insurance system, which is based on a fixed sum per patient on the list provided by the municipalities, a fee-for-service payment from the health insurance system, and a small copayment by the patient (approximately NOK 200 [approximately US$ 20]). Most GPs are specialized in general medicine, which requires recertification every 5 years. Recertification is based partly on continuous medical education courses, which are quality approved by the Norwegian Association for General Practice. This study will use such a course, offered for free to GPs and approved and certified by the board of the Norwegian Association of General Practice, to train GPs to administer the BI in their own practices and on their own patients (Figure 1).

Figure 1. General practitioner course package and tasks. BI: brief intervention.

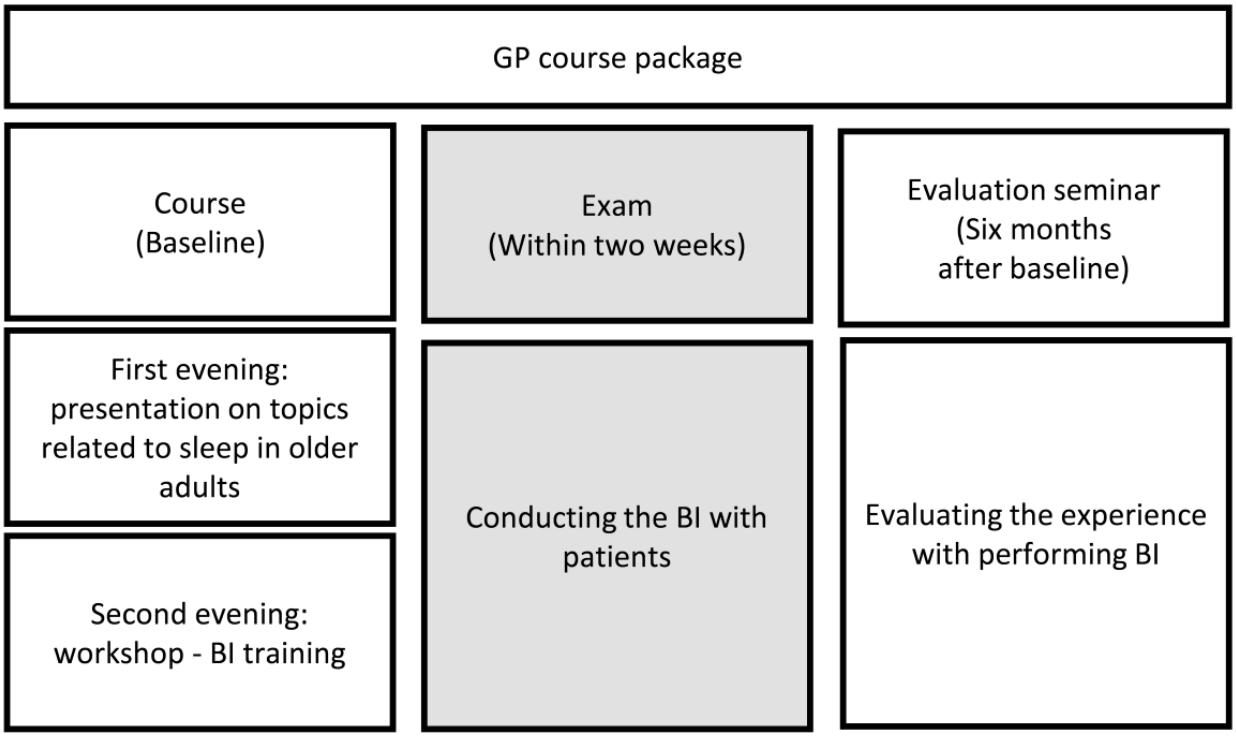

Recruitment, Randomization, and Blinding

GPs will be included by inviting them to a course on the treatment of sleep disorders among older patients. They will be informed that the course is also a part of a research study on an intervention regarding inappropriate use of sleep medications. An external statistician will then randomize the GPs to receive the course at 2 separate time points with an interval of 6 months

XSL•FO
RenderX

between them. This will form the basis for the cluster randomization of the patients on the GPs' lists; thus, patients will be randomized either to a GP participating in an earlier course or to a GP receiving the course later. Each GP and their patients will represent one cluster. We will avoid randomizing GPs belonging to the same GP center to different course groups to avoid cross-contamination of the intervention method. We will attempt to avoid selection bias, which would be expected through GPs selecting their own "suitable" patients, and will avoid collecting prescription patient data before patient consent. Thus, the study board will recruit the patients based on a short screening questionnaire, including written consent, to older patients (aged >60 years) on the patient lists of participating GPs. Self-reported sleeping problems, as well as use of sleeping medication, will be queried. After the course, GPs will receive a list of patients with self-reported Z-hypnotic use whom they should make an appointment with to administer the BI.

We do not expect GPs randomized to the later course to change their handling of their Z-hypnotic users before their course participation, and thus, they will constitute the BAU arm during the first part of the study. Indeed, GPs randomized to the later course will not be informed of their included patients until they have received the BI training course to avoid a change in patient handling in advance, aiming to maintain a true BAU control group. Those randomized to the early course will constitute the active (BI) arm during this period. The second course will enable a single crossover of previous BAU patients who will then also receive the intervention, providing the possibility for enriched before-and-after BI intervention data (Figure 2).

Figure 2. Outline of the study timeline. The upper half shows the tasks of the investigator team, general practitioners (GPs), and patients, and the lower part shows the timeline in weeks. Patients will be initially screened via a questionnaire and then receive self-report electronic questionnaires at specific time points (denoted as Q in the figure). GPs will be randomized to receive the behavioral intervention (BI) course (BIC) in 2 batches (BIC 1 will receive the course early, and BIC 2 will receive the course later) and administer the BI to their patients, who will be screened by the investigators at weeks 0 and 26, respectively (ie, BIC 1 patients at week 0 and BIC 2 patients at week 26). The late BI course will enable a crossover of initial business as usual (BAU) patients after 6 months. The main follow-up assessment (large arrows) will be performed by the investigators as an in-person appointment 6 weeks after the BI (or parallel BAU control).

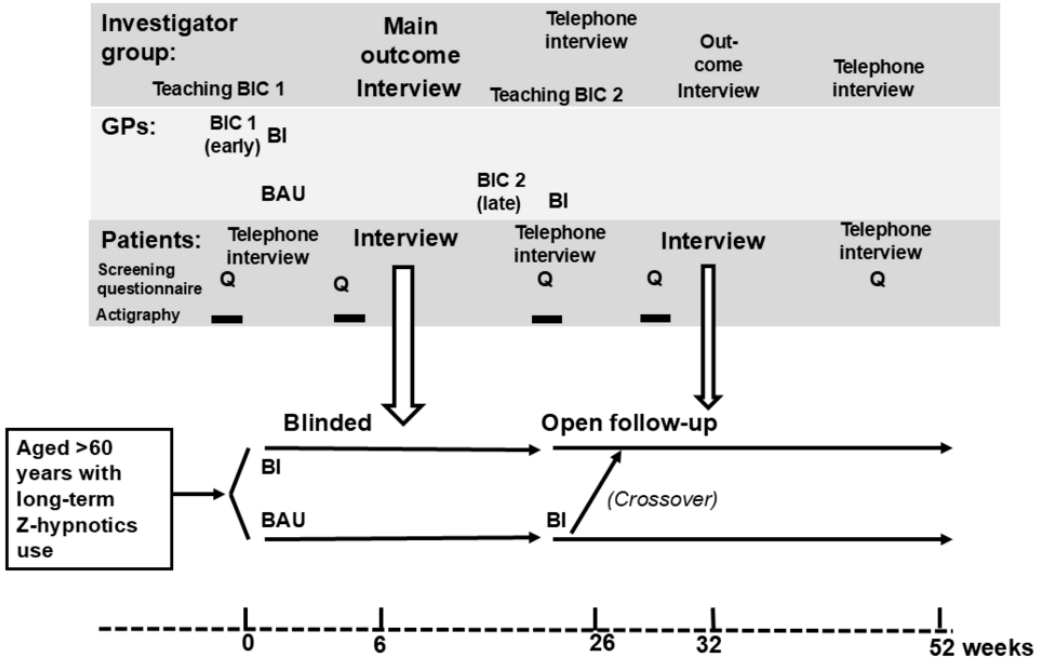

Patient Procedures

Data from patients will be collected as follows. First, GPs who consent to participate will deliver lists of names and contact addresses of all patients on their lists aged >60 years to the study board. The lists will contain no medical information and no personal ID numbers and will not be linked to patient data or hospital records. Second, patients on the list will be sent an email or SMS text message containing an information and consent document and a short screening questionnaire. Third, patients using Z-hypnotics will be recontacted in the same manner with an additional information and consent document inviting them to participate in a study focusing on sleep problems and sleep medication use patterns. Fourth, consenting patients will be contacted for a brief telephone interview before the GPs' BI course (or in parallel to the course for BAU clusters) and administered an online questionnaire at baseline (Figure 2). Fifth, BI GPs will be informed of a maximal number of 3 patients whom they should make an appointment with to administer the BI. The BI will be administered as described below. Sixth, 6 weeks and 6 months after BI patients will receive the follow-up questionnaire, a new telephone interview will be performed, and patients will be invited to a brief face-to face interview either at the hospital or in the patients' homes if they are unable to visit the hospital. Finally, additional telephone interviews and a follow-up online questionnaire will be conducted after 6 and 12 months. Incoming replies will be monitored (without addressing reply contents), and patient

XSL•FO
RenderX

reminders will be given via telephone and SMS text messaging to reduce missing data and dropouts during follow-up. This will also enable patient safety monitoring.

All patient-derived data collection will be conducted via safe links through individual QR codes to safe Data Protection Impact Assessment–approved research servers.

Intervention

The BI for inappropriate Z-hypnotic use taught to and administered by participating GPs consists of (1) an interview-based patient scoring system validated for detection of Z-hypnotic misuse and dependence in this population; (2) evidence-based and individualized information about Z-hypnotics and associated risks in our population, with special attention to the individual patient profile; (3) information about gains and difficulties with reducing inappropriate use (strong emphasis on information regarding expected rebound effects during the first couple of weeks after terminating use and the need to overcome this before improvement can be expected); (4) discussion of possible treatment goals for the patient; (5) formulation of a treatment plan, including goals to aim for, time frame, assumed difficulties and how to overcome them, and a plan for support from and follow-up by the GP office; and (6) summary and take-home message, including giving the patient a copy of the treatment plan.

Communication skills emphasized to GPs during the BI course will include being nonjudgmental and empathetic, using simple language and the teach-back method, formulating a written plan together with the patient, and emphasizing a supportive role.

Inclusion and Exclusion

The inclusion criterion is self-reported use of Z-hypnotics >3 days per week for >4 weeks. Exclusion criteria are diagnosis of dementia, psychosis, major depression, or delirium; inability to provide informed consent; and insufficient Norwegian-language capacity to complete the tests.

Data Collection and Follow-Up Time Points

Table 1 lists the main assessment time points of the study, as well as how data will be collected and which data we will collect for what purpose.

Table 1. Summary of main assessment time points and data to be collected.

Time point	Setting	Data to be collected	Purpose of the data
Before the intervention	• Screening self-report online questionnaire to all patients on the GPs'[a] lists • Self-report online questionnaire • Telephone interview	• Self-report sleep problems and sleep medication use • Basic sociodemographic data and data on anxiety and depression, pain, medication use, sleep, quality of life, and subjective cognitive function • Telephone-administered cognition instruments	• Basic information on GP population; screen for Z-hypnotic users for later intervention by the GP • Baseline predictor variables • Baseline cognitive function
Intervention	• At the GP's office	• None	• Outcome data
6 weeks after the intervention	• Self-report digital questionnaire • Face-to-face interview and cognitive tests	• Anxiety and depression, pain, medication use, sleep, quality of life, and subjective cognitive function • Cognitive function	• Main outcomes: Z-hypnotic use, pain, and cognitive function
6 months after the intervention	• Self-report digital questionnaire • Face-to-face interview and cognitive tests • Qualitative group interviews with GPs	• Anxiety and depression, pain, medication use, sleep, quality of life, and subjective cognitive function • Cognitive function • Recorded group interviews	• Main outcomes: Z-hypnotic use, pain, and cognitive function • Qualitative data on acceptability of the use of the brief intervention by GPs
12 months after the intervention	• Self-report questionnaire • Telephone interview	• Anxiety and depression, pain, medication use, sleep, quality of life, and subjective cognitive function • Telephone-administered cognition instruments	• Long-term outcomes, relapse, and dropout rate

[a]GP: general practitioner.

Follow-up time points will be chosen to enable catching possible rebound insomnia (6-week time point) and long-term relapse (up to 12 months). Missing data will be handled via last observation carried forward, providing at least one follow-up after the intervention.

Actigraphy Subsample

A total of 20 patients will be issued with actigraphy (ActiGraph) wearables without screens, and we will assess their sleep patterns and daily activity patterns over 2 weeks before and after the BI or BAU. The wearables will be delivered to the patients, and the data will be uploaded to a local secure server at the next assessment time point.

XSL•FO
RenderX

Qualitative Interviews

Six months after having tested the intervention with their assigned patients, GPs will be invited to a group interview to assess their experience using the BI method and the perceived acceptability among their patients. The interview will be based on an interview guide and will be video recorded and transcribed verbatim. The material will be qualitatively evaluated.

Outcomes

Main Outcome

The main outcome, on which power has also been based, is the proportion of patients with inappropriate use of Z-hypnotics (defined as use ≥3 days per week for ≥4 weeks without consideration of dosage) in the BI arm compared to BAU.

Secondary Outcomes

Main and additional secondary outcomes are presented in Table 2 with the data collection methods and instruments.

Table 2. Secondary outcomes and the instruments we will use.

Outcome and measures	Instruments
Main secondary outcomes	
Cognitive score	• Montreal Cognitive Assessment (face-to-face and telephone versions) [31,32]
Pain VAS[a]	• 100-mm VAS (numerical; 100 being the worst possible pain) [33]
Sleep self-assessment	• Global Sleep Assessment Questionnaire (11 items with response options of "never," "sometimes," "usually," and "always"; item responses were converted to a common scale from 0 to 100, with a higher score indicating greater likelihood of presence of the disorder) [34]
Additional secondary outcomes	
Sleep effectiveness	• Actigraphy registration 2 weeks before and after the intervention (subgroup only)
Quality of life	• EQ-5D-5L (Norwegian population norm valuation); measured on a scale from 0 to 1, with 1 representing perfect health [35]
Dependence-like behavior	• Severity of Dependence Scale (5 questions, total score range from 0 to 15, with 15 being the most severe) [36]
Additional cognitive assessments	• Cognistat (numerical score from 0 to 84, with 84 being the best) [37] • CFI[b] (Norwegian self-assessed cognitive instrument) and FAS[c] word score
Polypharmacy	• Number of drugs used regularly
Anxiety and depression	• Hospital Anxiety and Depression Scale (0-42, with 42 being the worst) [31,38]
Willingness to change	• Qualitative group interview
Acceptability	• Qualitative group interview

[a]VAS: visual analogue scale.

[b]Cognitive Function Instrument (Norwegian version).

[c]F-A-S word association test.

Statistics

Power

On the basis of previous studies [39], and using a conservative estimate, we assume that the intervention may lead to a 30% detoxification rate in the intervention (BI) arm and that, at most, 2% of patients in the nonintervention (BAU) arm may detoxify over 6 months. With a power of 80% to detect a significant difference with α at .05, we would need to include 52 patients in all (26 per study arm). Power analyses for the 2 tentative main secondary outcomes with α adjusted to .025 (Bonferroni correction) suggest that, altogether, 384 patients are needed to detect cognitive improvement; this sample size is based on the difference between central nervous system–depressant drug users and nonusers in our previous studies [23]. Similarly, over 2000 patients would be needed to detect an improvement in pain [24]. We observed that we were unlikely to attain such recruitment figures and chose to focus on power related to the main outcome, which we suggested as being attainable. In addition, as the number of patients for each GP will be maximized to 3 (in most cases likely just 1-2 patients per GP), we assumed that the clustering effect would not significantly affect power.

Statistical Analyses

The main analyses will follow the intention-to-treat principle, but we will also perform per-protocol analyses based on which

XSL•FO
RenderX

patients actually went through the BI intervention versus BAU. The primary analysis of the main outcome will assess difference in proportion of inappropriate Z-hypnotic users between the BI and BAU groups using the *z* test, which we also based our power calculations on. We will analyze secondary outcomes using linear and logistic regression for numerical and binary outcomes, respectively, as well as using relevant descriptive statistics.

Ethical Considerations

All participation follows the Declaration of Helsinki, with written informed consent collected from both GPs and patients before any study-related procedures and collection of study data. Patients will receive the contact address of the study lead when signing the informed consent form and will be encouraged to make contact if problems or questions arise. In addition, GPs will have a running responsibility toward their own patients. Data will be stored on secure and Data Protection Impact Assessment–approved servers at the research institute. All procedures were assessed and approved by the Research Ethics Committee of South East Norway (REK 556653) and by the Data Inspectorate Office at Akershus University Hospital. Participants received no compensation.

The BI course was approved by the relevant board of the Norwegian Association of General Practice as providing part of the recertification points required for GPs' continuous medical education.

Dissemination of and Feedback on Results

In addition to publishing the results in relevant international peer-reviewed journals, we will ascertain deidentified feedback from the participating GPs and patients. Patients using actigraphy registration will also be able to receive their own registration assessment after completion of the study. Data will be anonymized and stored on safe servers as specified by the General Data Protection Regulation and approved by the research ethics committee and the data protection officer.

We are committed to publishing both positive and negative results We also intend to implement further dissemination of the BI course should the method yield positive results, as well as attempt to disseminate the results through basic teaching for GPs, other physicians, and medical students. Finally, it is a further goal to evaluate and implement parts of the course to give recommendations for treatment through relevant web platforms for health staff. User representatives and health care representatives in our steering committee will actively participate.

Results

We expect to be able to show whether the BI is significantly better at reducing the number of older patients who use Z-hypnotics inappropriately than the nonintervention (ie, BAU), which is the main outcome of this study. In addition, we aim to demonstrate that there is no significant worsening of the main secondary outcomes of sleep (ie, no increase in reported insomnia), cognition, or pain as inappropriate Z-hypnotic use is reduced. Thus, these secondary outcome results may be viewed as assessing possible side effects of the intervention, which we suggest to be an important part of the overall assessment of gains versus losses of the intervention. Should side effects be increased after the intervention, especially if the primary outcome effect is not positive or is small, we will suggest that this is a strong argument against the intervention. However, if the main outcome effect is strong and we observe no downsides to this in terms of increased side effects, this will strengthen the argument for using the BI intervention in this patient population. As the recruitment process is vulnerable and has several steps, and we cannot completely eliminate a carryover effect (from trained GPs to untrained GPs, which would possibly reduce the effect size in the controlled part of the study), a single crossover for the control arm to receive the BI at a later stage was added to increase power and be able to more closely assess the timing of a possible change from the intervention based on when the intervention was delivered. We expect this crossover to also provide additional results on change in medication-related behavior and maintenance of such behavior over different periods up to 12 months. We expect the main results of the study to be published during winter 2025-2026.

Discussion

As there is no established and generally accepted patient-centered intervention method to address inappropriate use of Z-hypnotics among older individuals, despite this being associated with several possible negative effects, there is a need for further evidence-based data on how to approach patients with possible inappropriate use. As this is a "hot topic" that is often debated, our aim is to provide randomized controlled trial–level evidence for the effect of a simple BI-based method that may, if proven effective, help GPs and other physicians address the issue with their patients.

We have chosen to base the intervention in primary care for several reasons. In Norway, GPs represent the most accessible part of the health care system, and almost all Norwegians are listed with a GP. GPs are charged with the responsibility of being gatekeepers for treatment and referral, when required, to the secondary, usually hospital-based, health services, as well as the follow-up of patients after hospital treatment. Therefore, the GP plays a key role in the prescription and termination of prescription, as well as in the assessment and follow-up of possible medication side effects for their patients. GPs are responsible for most prescriptions, whether these are first introduced by other specialists or by the GPs themselves. This is especially noticeable for prescribed drugs with a possible addictive potential [27]. For GPs who often have long-term care and responsibility for their patients, this may be a daunting task. Although simply the awareness of this responsibility and the central rules and recommendations regarding such medications have been suggested as effective means to limit inappropriate use in general, there is a paucity of structured instruments that may help GPs support individual patients in reducing inappropriate use. In addition, effective intervention at the primary health care level may be more cost-effective and follows the recommendations for treatment at the lowest effective treatment level, which is a central focus in the Norwegian decentralized health care system.

XSL•FO
RenderX

The GPs included in the study are from the southeastern health care region in Norway, which has both urban and rural GP settings and may be seen as reasonably representative. Participation of GPs will be via invitation to a course on sleep problems and sleep medication use to GPs in the entire region, although we cannot exclude that this may lead to some recruitment bias and recruit mainly GPs who are positive about a more restrictive use of Z-hypnotics. GPs need to obtain recertification points based on participation in several courses on different themes such as this one. Therefore, we expect to have a reasonably good mixture of representative GPs from our region as participants.

Regarding patient inclusion bias, we will attempt to reduce the risk of those most at risk of Z-hypnotic misuse dropping out and, thus, overrecruit patients who may be easy to "detoxify." Information to patients stating that "detoxification" is a primary aim of a study certainly risks losing patients who are negative about such measures. Although research ethics requires that patients provide informed consent for participation, the way in which this information is formulated may be very important. In close correspondence with the responsible research ethics committee, we landed on a reasonable and ethical compromise regarding the formulation in this regard. We focused on individual evaluation and optimization of the medication to avoid side effects, as well as on individual follow-up of medication effects and side effects and re-evaluation if the effect of the intervention was perceived as negative.

The BI is a simple and short intervention that equips both patients and GPs with evidence-based knowledge and a systematic approach to discontinuing inappropriate Z-hypnotic use. If proven effective, a structured intervention that is simple to use and learn and easy to teach will likely be possible to implement at a larger scale, with the potential to be of benefit to many patients in primary care. Such implementation is a clear aim of our project, and we will use our network and user group to achieve this. We suggest that there is also a potential for implementation of similar individually focused interventions internationally, at least in countries with a similar primary health care system structure to the Norwegian one.

To avoid assessment effect and resulting change in behavior also in the control group, we have limited the amount of data collected in advance of the intervention, with no in-person visits in advance. We have chosen neutrally formulated questions and instruments assessing self-reported sleep, medication use, socioeconomic status, and pain in general. Specific questions focusing on the individuals' use of Z-hypnotics and dependence-like characteristics only come after the intervention or control phase. However, this means that our preintervention data are limited.

Instruments used for data collection are based on our previous experience and have been validated for the age group and setting described and additionally validated by us in previous studies [40,41]. Self-reports on experienced sleep difficulties may be biased by patient expectations. Therefore, we have included a subsample of patients who will be monitored through actigraphy over a 2-week period before the intervention, as well as a similar period after the BI intervention or similarly timed for the BAU controls. Thus, we will be able to supplement self-reports with objective actigraphy measures of sleep efficiency and daytime activity.

Importantly, the design process of the intervention has been extensive and based on the recent guidelines from the UK Medical Research Council for the design of complex intervention studies [30], as well as on our own previous experience. In the process, a special emphasis was made on developing an intervention that was feasible to carry out by GPs in an ordinary GP consultation considering GP workload and available time. The feasibility study suggests that, with some minor logistics changes, we have succeeded in this [29]. However, this will be re-evaluated through qualitative interviews also as part of this study. We have also previously administered a similar BI in general practice for inappropriate medication use in the context of MOH with positive results, which also supports the viability of this study protocol [20,21,42].

Acknowledgments

Jūratė Šaltytė Benth, Division of Health Services Research and Psychiatry, University of Oslo, campus Akershus University Hospital, is gratefully acknowledged for statistical support in the early phases of planning.

Funding

This study was financed through a postdoctoral grant to TGS, with CL as senior project leader, from the South-Eastern Norway Regional Health Authority (grant 2022056).

Data Availability

Though data protection policies preclude us from releasing identifiable data, fully anonymized data are available upon reasonable request to the authors and evaluation by the data inspectorate officer of the hospital.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. American Geriatrics Society 2012 Beers Criteria Update Expert Panel. American Geriatrics Society updated Beers Criteria for potentially inappropriate medication use in older adults. J Am Geriatr Soc 2012 Apr;60(4):616-631 [FREE Full text] [doi: 10.1111/j.1532-5415.2012.03923.x] [Medline: 22376048]
2. Nyborg G, Brekke M, Straand J, Gjelstad S, Romøren M. Potentially inappropriate medication use in nursing homes: an observational study using the NORGEP-NH criteria. BMC Geriatr 2017 Sep 19;17(1):220 [FREE Full text] [doi: 10.1186/s12877-017-0608-z] [Medline: 28927372]
3. O'Mahony D. STOPP/START criteria for potentially inappropriate medications/potential prescribing omissions in older people: origin and progress. Expert Rev Clin Pharmacol 2020 Jan;13(1):15-22. [doi: 10.1080/17512433.2020.1697676] [Medline: 31790317]
4. Johnson CF, Frei C, Downes N, McTaggart SA, Akram G. Benzodiazepine and z-hypnotic prescribing for older people in primary care: a cross-sectional population-based study. Br J Gen Pract 2016 Jun;66(647):e410-e415 [FREE Full text] [doi: 10.3399/bjgp16X685213] [Medline: 27114208]
5. Forthun I, Rørtveit G, Fossum GH, Jørgensen P, Halvorsen PA, Nilsen S, et al. Søvnforstyrrelser og forskrivning av hypnotika i allmennpraksis – en PraksisNett-studie. Helse Bergen. URL: https://www.helse-bergen.no/496524/siteassets/seksjon/sovno/documents/sovnforstyrrelser-og-forskriving-av-hypnotika.pdf [accessed 2025-12-19]
6. Lukačišinová A, Fialová D, Peel NM, Hubbard RE, Brkic J, Onder G, et al. The prevalence and prescribing patterns of benzodiazepines and Z-drugs in older nursing home residents in different European countries and Israel: retrospective results from the EU SHELTER study. BMC Geriatr 2021 Apr 26;21(1):277 [FREE Full text] [doi: 10.1186/s12877-021-02213-x] [Medline: 33902474]
7. Alessi-Severini S, Bolton JM, Enns MW, Dahl M, Collins DM, Chateau D, et al. Use of benzodiazepines and related drugs in Manitoba: a population-based study. CMAJ Open 2014 Oct 01;2(4):E208-E216 [FREE Full text] [doi: 10.9778/cmajo.20130076] [Medline: 25485245]
8. Reeve E, Ong M, Wu A, Jansen J, Petrovic M, Gnjidic D. A systematic review of interventions to deprescribe benzodiazepines and other hypnotics among older people. Eur J Clin Pharmacol 2017 Aug 30;73(8):927-935. [doi: 10.1007/s00228-017-2257-8] [Medline: 28456823]
9. Lynch T, Ryan C, Hughes CM, Presseau J, van Allen ZM, Bradley CP, et al. Brief interventions targeting long-term benzodiazepine and Z-drug use in primary care: a systematic review and meta-analysis. Addiction 2020 Sep 11;115(9):1618-1639. [doi: 10.1111/add.14981] [Medline: 31985127]
10. Vicens C, Leiva A, Bejarano F, Sempere E, Rodríguez-Rincón RM, Fiol F, et al. Intervention to reduce benzodiazepine prescriptions in primary care, study protocol of a hybrid type 1 cluster randomised controlled trial: the BENZORED study. BMJ Open 2019 Jan 30;9(1):e022046 [FREE Full text] [doi: 10.1136/bmjopen-2018-022046] [Medline: 30705235]
11. Vicens C, Leiva A, Bejarano F, Sempere-Verdú E, Rodríguez-Rincón RM, Fiol F, et al. Evaluation of a multicomponent intervention consisting of education and feedback to reduce benzodiazepine prescriptions by general practitioners: the BENZORED hybrid type 1 cluster randomized controlled trial. PLoS Med 2022 May 6;19(5):e1003983 [FREE Full text] [doi: 10.1371/journal.pmed.1003983] [Medline: 35522626]
12. Babor TF, Higgins-Biddle JC. Brief intervention: for hazardous and harmful drinking. World Health Organization. 2001. URL: https://www.drugsandalcohol.ie/14105/1/WHO_Brief_intervention.pdf [accessed 2025-12-21]
13. Henry-Edwards S, Humeniuk R, Ali R, Monteiro M, Poznyak V. Brief intervention for substance use: a manual for use in primary care. World Health Organization. 2003. URL: https://swsphn.com.au/wp-content/uploads/2021/12/Draft_Brief_Intervention_for_Substance_Use.pdf [accessed 2025-12-21]
14. Humeniuk R, Ali R, Babor T, Souza-Formigoni ML, de Lacerda RB, Ling W, et al. A randomized controlled trial of a brief intervention for illicit drugs linked to the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) in clients recruited from primary health-care settings in four countries. Addiction 2012 May;107(5):957-966. [doi: 10.1111/j.1360-0443.2011.03740.x] [Medline: 22126102]
15. Humeniuk R, Dennington V, Ali R. The effectiveness of a brief intervention for illicit drugs linked to the alcohol, smoking and substance involvement screening test (ASSIST) in primary health care settings: a technical report of phase III findings of the WHO ASSIST randomized controlled trial. World Health Organization. 2008. URL: https://www.drugsandalcohol.ie/28120/1/assist_technicalreport_phase3_final.pdf [accessed 2025-12-21]
16. Kaner EF, Dickinson HO, Beyer F, Pienaar E, Schlesinger C, Campbell F, et al. The effectiveness of brief alcohol interventions in primary care settings: a systematic review. Drug Alcohol Rev 2009 May 05;28(3):301-323. [doi: 10.1111/j.1465-3362.2009.00071.x] [Medline: 19489992]
17. Humeniuk R, Henry-Edwards S, Ali R, Poznyak V, Monteiro MG. The ASSIST-linked brief intervention for hazardous and harmful substance use. World Health Organization. 2012 Jun 16. URL: https://www.who.int/publications/i/item/9789241599399 [accessed 2025-12-21]
18. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: applications to addictive behaviors. J Addict Nurs 1993;5(1):2-16 [FREE Full text] [doi: 10.3109/10884609309149692]
19. Prochaska JO. Decision making in the transtheoretical model of behavior change. Med Decis Making 2008;28(6):845-849. [doi: 10.1177/0272989X08327068] [Medline: 19015286]

XSL•FO
RenderX

20. Kristoffersen ES, Straand J, Vetvik KG, Benth J, Russell MB, Lundqvist C. Brief intervention for medication-overuse headache in primary care. The BIMOH study: a double-blind pragmatic cluster randomised parallel controlled trial. J Neurol Neurosurg Psychiatry 2015 May;86(5):505-512. [doi: 10.1136/jnnp-2014-308548] [Medline: 25112307]
21. Frich JC, Kristoffersen ES, Lundqvist C. GPs' experiences with brief intervention for medication-overuse headache: a qualitative study in general practice. Br J Gen Pract 2014 Sep;64(626):e525-e531 [FREE Full text] [doi: 10.3399/bjgp14X681313] [Medline: 25179065]
22. Cheng ST, Siddiqui TG, Gossop M, Kristoffersen ES, Lundqvist C. Sociodemographic, clinical and pharmacological profiles of medication misuse and dependence in hospitalised older patients in Norway: a prospective cross-sectional study. BMJ Open 2019 Sep 05;9(9):e031483 [FREE Full text] [doi: 10.1136/bmjopen-2019-031483] [Medline: 31492795]
23. Siddiqui TG, Cheng S, Gossop M, Kristoffersen ES, Grambaite R, Lundqvist C. Association between prescribed central nervous system depressant drugs, comorbidity and cognition among hospitalised older patients: a cross-sectional study. BMJ Open 2020 Jul 27;10(7):e038432 [FREE Full text] [doi: 10.1136/bmjopen-2020-038432] [Medline: 32718926]
24. Bjelkarøy MT, Cheng S, Siddiqui TG, Benth J, Grambaite R, Kristoffersen ES, et al. The association between pain and central nervous system depressing medication among hospitalised Norwegian older adults. Scand J Pain 2022 Jul 26;22(3):483-493 [FREE Full text] [doi: 10.1515/sjpain-2021-0120] [Medline: 34913326]
25. Bjelkarøy MT, Simonsen TB, Siddiqui TG, Cheng S, Grambaite R, Benth J, et al. Mortality and health-related quality of life in older adults with long-term use of opioids, z-hypnotics or benzodiazepines: a prospective observational study at 5 years follow-up. BMJ Open 2024 Feb 21;14(2):e079347 [FREE Full text] [doi: 10.1136/bmjopen-2023-079347] [Medline: 38387984]
26. Siddiqui TG, Bjelkarøy MT, Cheng S, Kristoffersen ES, Grambaite R, Lundqvist C. The effect of cognitive function and central nervous system depressant use on mortality-a prospective observational study of previously hospitalised older patients. PLoS One 2022 Mar 3;17(3):e0263024 [FREE Full text] [doi: 10.1371/journal.pone.0263024] [Medline: 35239678]
27. Kann IC, Lundqvist C, Lurås H. Prescription of addictive and non-addictive drugs to home-dwelling elderly. Drugs Aging 2014 Jun 20;31(6):453-459. [doi: 10.1007/s40266-014-0169-1] [Medline: 24842594]
28. Kann IC, Biørn E, Lurås H. Competition in general practice: prescriptions to the elderly in a list patient system. J Health Econ 2010 Sep;29(5):751-764. [doi: 10.1016/j.jhealeco.2010.07.004] [Medline: 20708282]
29. Bjelkarøy MT, Simonsen TB, Siddiqui TG, Halset S, Cheng S, Grambaite R, et al. Brief intervention as a method to reduce z-hypnotic use by older adults: feasibility case series. JMIR Form Res 2024 Feb 08;8:e51862 [FREE Full text] [doi: 10.2196/51862] [Medline: 38329779]
30. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
31. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr 30;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
32. Wong A, Nyenhuis D, Black SE, Law LS, Lo ES, Kwan PW, et al. Montreal Cognitive Assessment 5-minute protocol is a brief, valid, reliable, and feasible cognitive screen for telephone administration. Stroke 2015 Apr;46(4):1059-1064 [FREE Full text] [doi: 10.1161/STROKEAHA.114.007253] [Medline: 25700290]
33. Herr KA, Spratt K, Mobily PR, Richardson G. Pain intensity assessment in older adults: use of experimental pain to compare psychometric properties and usability of selected pain scales with younger adults. Clin J Pain 2004;20(4):207-219. [doi: 10.1097/00002508-200407000-00002] [Medline: 15218405]
34. Roth T, Zammit G, Kushida C, Doghramji K, Mathias SD, Wong JM, et al. A new questionnaire to detect sleep disorders. Sleep Med 2002 Mar;3(2):99-108. [doi: 10.1016/s1389-9457(01)00131-9] [Medline: 14592227]
35. Brooks R. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72. [doi: 10.1016/0168-8510(96)00822-6] [Medline: 10158943]
36. Gossop M, Griffiths P, Powis B, Strang J. Severity of dependence and route of administration of heroin, cocaine and amphetamines. Br J Addict 1992 Nov 13;87(11):1527-1536. [doi: 10.1111/j.1360-0443.1992.tb02660.x] [Medline: 1458032]
37. Kiernan RJ, Mueller J, Langston JW, Van Dyke C. The Neurobehavioral Cognitive Status Examination: a brief but quantitative approach to cognitive assessment. Ann Intern Med 1987 Oct 01;107(4):481-485. [doi: 10.7326/0003-4819-107-4-481] [Medline: 3631786]
38. Helvik A, Engedal K, Skancke RH, Selbæk G. A psychometric evaluation of the Hospital Anxiety and Depression Scale for the medically hospitalized elderly. Nord J Psychiatry 2011 Oct 22;65(5):338-344. [doi: 10.3109/08039488.2011.560684] [Medline: 21341979]
39. Llorente MD, David D, Golden AG, Silverman MA. Defining patterns of benzodiazepine use in older adults. J Geriatr Psychiatry Neurol 2000 Oct 01;13(3):150-160. [doi: 10.1177/089198870001300309] [Medline: 11001138]
40. Cheng S, Siddiqui TG, Gossop M, Kristoffersen ES, Lundqvist C. The Severity of Dependence Scale detects medication misuse and dependence among hospitalized older patients. BMC Geriatr 2019 Jun 24;19(1):174 [FREE Full text] [doi: 10.1186/s12877-019-1182-3] [Medline: 31234786]

XSL•FO
RenderX

41. Bjelkarøy MT, Benth J, Simonsen TB, Siddiqui TG, Cheng S, Kristoffersen ES, et al. Measuring pain intensity in older adults. Can the visual analogue scale and the numeric rating scale be used interchangeably? Prog Neuropsychopharmacol Biol Psychiatry 2024 Mar 02;130:110925 [FREE Full text] [doi: 10.1016/j.pnpbp.2023.110925] [Medline: 38143014]
42. Kristoffersen ES, Straand J, Benth JS, Russell MB, Lundqvist C. Study protocol: brief intervention for medication overuse headache--a double-blinded cluster randomised parallel controlled trial in primary care. BMC Neurol 2012 Aug 10;12:70 [FREE Full text] [doi: 10.1186/1471-2377-12-70] [Medline: 22883540]

Abbreviations

BAU: business as usual
BI: brief intervention
GP: general practitioner
MOH: medication overuse headache
RCT: randomized controlled trial

Edited by J Sarvestan; submitted 08.Apr.2025; peer-reviewed by S Khan; comments to author 06.Aug.2025; revised version received 25.Nov.2025; accepted 26.Nov.2025; published 30.Jan.2026.

Please cite as:
Siddiqui TG, Bjelkarøy MT, Simonsen TB, Selle ML, Lundqvist C
Brief Intervention for Discontinuing Inappropriate Z-Hypnotic Use Among Older Patients in Primary Care: Protocol for a Cluster Randomized Controlled Trial With a Single Crossover
JMIR Res Protoc 2026;15:e75670
URL: https://www.researchprotocols.org/2026/1/e75670
doi:10.2196/75670
PMID:

©Tahreem Ghazal Siddiqui, Maria Torheim Bjelkarøy, Tone Breines Simonsen, Maria Lie Selle, Christofer Lundqvist. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Impact of the Population Medicine Multimorbidity Intervention in Xishui County (POPMIX) on People at High Risk for Chronic Obstructive Pulmonary Disease: Protocol for the POPMIX-COPD Cluster Randomized Controlled Trial

Chen Wang[1,2,3,4*], MD; Yuhao Liu[1*], BA; Ke Huang[2,3,4*], MD; Zhoutao Zheng[1], BMed; Shiyu Zhang[5], BSc; Wenjin Chen[1], BMed; Xingyao Tang[2,3], MMSc; Zhong Cao[6], PhD; Xunliang Tong[7,8], MD; Lei Tang[9], PhD; Jinghan Zhao[6], MA; Liu He[1], BMed; Lirui Jiao[10], MPH; Yingping Wang[11], BMed; Tianying Zhao[11], BM; Yingchi Luo[11], BMed; Qiande Lai[12], BMed; Xiangqin Lyu[12], BMed; Qiushi Chen[13], PhD; Aditi Bunker[6], PhD; Sebastian Vollmer[14], PhD; Pascal Geldsetzer[15,16], PhD; Dean Jamison[17], PhD; Till Bärnighausen[1,6,18], MD, ScD; Ting Yang[2,3,4], MD; Simiao Chen[1,4,6], ScD[‡]; POPMIX Group[19]

[1]School of Population Medicine and Public Health, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[2]Department of Pulmonary and Critical Care Medicine, China-Japan Friendship Hospital, Beijing, China

[3]National Center for Respiratory Medicine, Beijing, China

[4]State Key Laboratory of Respiratory Health and Multimorbidity, Beijing, China

[5]School of Health Policy and Management, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

[6]Heidelberg Institute of Global Health, Faculty of Medicine and University Hospital, Heidelberg University, Heidelberg, Germany

[7]Department of Pulmonary and Crtical Care Medicine, Beijing Hospital, Beijing, China

[8]Institute of Geriatric Medicine, Chinese Academy of Medical Sciences, Beijing, China

[9]Guizhou Medical University, Guiyang, China

[10]Department of Health Policy and Management, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[11]Center for Disease Control and Prevention of Xishui County, Zunyi, China

[12]Department of Pulmonary and Critical Care Medicine, People's Hospital of Xishui County, Zunyi, China

[13]The Harold and Inge Marcus Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, PA, United States

[14]Department of Economics and Centre for Modern Indian Studies, University of Göttingen, Göttingen, Germany

[15]Division of Primary Care and Population Health, Department of Medicine, Stanford University, Stanford, CA, United States

[16]Chan Zuckerberg Biohub San Francisco, San Francisco, CA, United States

[17]Department of Epidemiology and Biostatistics and Institute for Global Health Sciences, University of California, San Francisco, San Francisco, CA, United States

[18]Department of Global and Population Health, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, United States

[19]See Acknowledgments

[‡]On behalf of the POPMIX study investigators

[*]these authors contributed equally

Corresponding Author:
Simiao Chen, ScD
School of Population Medicine and Public Health
Chinese Academy of Medical Sciences & Peking Union Medical College
Beijige Santiao 31
Dongcheng District
Beijing, 100730
China
Phone: 49 6221565948
Email: simiao.chen@uni-heidelberg.de

Abstract

XSL•FO
RenderX

Background: Tobacco-related noncommunicable diseases (NCDs) present a major public health challenge in China, requiring population-level management. Chronic obstructive pulmonary disease (COPD) is the most common and prevalent chronic respiratory disease associated with tobacco use. In addition, COPD shares risk factors with other NCDs that frequently co-occur, leading to multimorbidity. This study focuses on the early detection and integrated management of COPD and related multimorbidity among high-risk populations. Population medicine, an emerging and evolving concept aimed at maximizing population health and well-being, provides a promising framework for shifting interventions against COPD from an individual patient focus to a population-level approach.

Objective: This study aims to evaluate the effectiveness of a population medicine–based multimorbidity intervention package among individuals at high risk for COPD.

Methods: We are conducting a 2-arm, population-based, stratified cluster randomized controlled trial (cRCT). The intervention integrates community screening, chronic disease management, patient education, digital follow-up, and team-based care. The trial is being implemented in Xishui County, Guizhou Province, a low-resource county in Southwestern China. Each of the 26 townships in Xishui County was considered a cluster and stratified into large and small townships based on population size. An equal number of residents from each township stratum (large and small) were randomized to undergo the COPD Screening Questionnaire. Individuals identified as being at high risk for COPD were considered study participants and were subsequently enrolled in either the intervention or control arm. The target sample size was approximately 2850 individuals.

Results: Data collection for the POPMIX-COPD trial began in June 2024. Baseline, 3-month, and 6-month assessments have been completed, and 12-month follow-up assessments are planned to be completed in March 2026. All participants in the intervention arm are being followed for 1 year, with 1 telephone follow-up at month 3 and in-person follow-ups at months 6 and 12. Primary outcomes for each participant include the number of chronic conditions controlled, receipt of lung function testing, and forced expiratory volume in 1 second. In addition, secondary outcomes were health-related quality of life, mental and behavioral health status, health care utilization, knowledge of COPD and asthma, and care cascade indicators for chronic conditions.

Conclusions: This cRCT is the first multimorbidity intervention study designed within the population medicine framework to target populations at high risk for COPD. It was featured as a case study in the report of the Lancet Commission on Investing in Health. The results of the trial are expected to inform the next generation of multimorbidity management and population medicine practices among global health authorities and practitioners.

Trial Registration: ClinicalTrials.gov NCT06456996; https://clinicaltrials.gov/ct2/show/NCT06456996

International Registered Report Identifier (IRRID): DERR1-10.2196/85597

(JMIR Res Protoc 2026;15:e85597) doi:10.2196/85597

KEYWORDS

population medicine; multimorbidity; tobacco-related NCDs; high-COPD-risk population; cRCT

Introduction

The concept of population medicine focuses on maximizing health and well-being at the population level—in contrast to clinical medicine's traditional focus on the individual patient—and has been gaining attention in recent literature [1]. Population medicine is a progressive medical discipline that aims to promote health equity through the following mechanisms: (1) integrating clinical medicine and public health; (2) providing comprehensive care that encompasses health promotion, disease prevention, diagnosis, control, treatment, and rehabilitation; (3) integrating and applying the knowledge, academic principles, and technologies of modern medicine with those of related disciplines; (4) coordinating individual health behaviors and collective health actions (eg, community-based smoking cessation campaigns, vaccination programs, and health education activities); and (5) serving as the medical foundation of public health practice [1]. To maximize population health, it is crucial to identify conditions that should be prioritized for population-level management. Targeting the most prevalent, preventable, and impactful conditions enables more efficient allocation of resources and earlier intervention, leading to broad public health gains.

Noncommunicable diseases (NCDs) strongly linked to tobacco use are enumerated among the 7 NCDs and injury-related priority conditions that the Lancet Commission on Investing in Health identified as needing to be addressed to achieve its "50 by 50" goal of halving the probability of premature death by 2050 [2]. Tobacco-related NCDs include chronic obstructive pulmonary disease (COPD) and certain cancers, as well as deaths from atherosclerotic cardiovascular disease and hemorrhagic stroke [2]. In China, tobacco-related NCDs accounted for approximately 1 year of an estimated 3.5-year life expectancy gap attributable to the Lancet Commission's 7 NCDs and injury-related priority conditions, relative to the North Atlantic regions. Furthermore, the portion attributable to tobacco-related NCDs represented approximately 23% of the overall 4.3-year life expectancy gap between the 2 regions [3]. The magnitude of this gap underscores the importance of managing tobacco-related NCDs at the population level, from preventive screening to disease management.

Among all tobacco-related NCDs, COPD is the most common respiratory disease. COPD is a leading cause of mortality and morbidity in China and worldwide [4,5]. It is the fourth leading cause of death and disability-adjusted life years among all diseases and was responsible for over 67.8 million deaths

XSL•FO
RenderX

worldwide and approximately 11.7 million deaths in China in 2021 [6]. Recent estimates suggest that more than 400 million people are affected globally, with this number projected to reach 592 million by 2050 [7]. The immense health burden imposed by COPD also results in a substantial economic burden. It is estimated that COPD will cost China's economy US $1.361 trillion (in 2017 international dollars) over the period 2020-50 [8], amounting to the largest economic burden from COPD faced by any country. The economic value of reducing avoidable mortality from COPD amounts to 1.4% of annual national income in 2019 and 0.9% in 2050; the 2019 figure places COPD as the fourth-ranked cause of death in terms of economic harm, as well as mortality, in China [9]. Moreover, COPD prevalence has steadily increased over the past 2 decades. Among the population aged 40 years and older, prevalence rose from 8% in 2002-04 to 14% in 2014-15, representing an increase of nearly 70% [10].

Two main challenges in COPD care have been widely discussed in the literature. First, awareness and early detection of COPD remain low among both the general and high-risk populations. Second, multimorbidity is frequently observed in patients with COPD and in those at high risk of developing the condition. Recent studies have shown an unmet health care need for COPD screening, formal diagnosis, treatment, and symptom control (ie, poor performance of the COPD care cascade) in China [11,12]. In a large-scale, population-based COPD screening and management program (the "Happy Breathing" program), it was estimated that among all patients with COPD who were spirometry-diagnosed within the program, 41.0% had previously undergone spirometry testing for the condition, 17.6% had been formally diagnosed with COPD, and 8.5% were currently receiving treatment. Among those receiving treatment, 4.6% had mild or no exacerbations in the prior year, and 3.9% had experienced no exacerbations [12]. Given the severe unmet health care need for COPD, increased attention to population-based COPD screening and interventions is warranted.

Community-based screening and management in populations at high risk for COPD represent one potential strategy to enhance awareness and early detection. For example, the TargetCOPD trial showed that screening ever smokers aged 40-79 years can lead to a significant increase in the detection of new COPD cases relative to usual practice; moreover, this approach has proven cost-effective as an active case-finding strategy [13]. Another recent study modeled a community-based COPD screening scenario and demonstrated its cost-effectiveness using a 2-step method [14]. Such research provides a solid scientific basis for conducting population-based COPD screening and management among individuals at high risk for COPD.

Multimorbidity among patients with COPD is extremely common, as these individuals often face behavioral risk factors, such as smoking, and tend to be older (indeed, nearly 50% of all adults suffer from multimorbidity at baseline) [15]. The reported prevalences of other chronic conditions among patients with COPD are high: more than 50% have hypertension [16], roughly 10%-60% experience anxiety and depression [17], 20% have diabetes [18], about 30% are obese and about 17% are underweight [16], 19% have chronic heart failure [18,19], 20% have ischemic heart disease [18,20], and 5% have had a stroke [18]. The high prevalence of coexisting conditions presents significant challenges for health care delivery, particularly with respect to designing care models that address multiple conditions simultaneously. A key challenge in delivering effective care to this population lies in how health care providers are incentivized. In many settings, clinician incentives remain aligned with passive diagnosis and treatment of individual diseases, rather than with providing proactive, inclusive care that holistically addresses multimorbidity.

In response to the substantial burden of COPD and associated multimorbidity, a comprehensive, population-based integrated model of screening and long-term care has been launched. However, there has been limited prior research evaluating the effectiveness of integrated management interventions within a population medicine framework that emphasizes proactive case identification and management among high-risk groups, rather than relying solely on reactive clinical care. Therefore, an important goal of this study is to evaluate the effectiveness of the newly launched multimorbidity care model in addressing the growing challenges of COPD and multimorbidity. In this protocol, we implement an integrated multimorbidity intervention package to evaluate its impact on relevant health outcomes among individuals diagnosed with COPD or at high risk of COPD. We also introduce a performance-linked "pay-for-population" scheme that ties financial incentives for health care workers to concrete population-level indicators, such as spirometry coverage, diagnosis rate, treatment initiation rate, and reduction in acute exacerbations. The design of the integrated package and the decision to evaluate its effects as a whole are based on consensus with the local authority to ensure feasibility in a real-world setting. This combined strategy is intended to shift provider behavior from passive disease response to proactive, integrated care delivery in a manner aligned with the principles of population medicine. This study is part of the Population Medicine Multimorbidity Intervention in Xishui County (POPMIX) project, which aims to develop and evaluate evidence-based, multicomponent strategies for managing chronic diseases and their comorbidities in a resource-limited rural county in China. We aim to investigate whether a population-based integrated multimorbidity intervention package affects our primary outcomes, including the number of chronic diseases controlled, whether participants have received lung function testing, and forced expiratory volume in 1 second (FEV_1), as well as our secondary outcomes, such as health-related quality of life, mental health, behavioral risk factors, health care utilization, productivity loss, knowledge of COPD and asthma, and care cascade indicators for COPD and other chronic diseases.

Methods

Trial Design

This is a parallel, 2-arm, stratified cluster randomized controlled trial (cRCT) conducted in Xishui, Guizhou, China (Figure 1). The protocol was designed in accordance with the Standard Protocol Items: Recommendations for Interventional Trials

XSL•FO
RenderX

(SPIRIT) 2025 Statement. The SPIRIT checklist for this protocol is provided in Multimedia Appendix 1.

Figure 1. Trial flowchart. COPD: chronic obstructive pulmonary disease; COPD-SQ: COPD Screening Questionnaire. See Multimedia Appendix 4 for a higher resolution version of this figure.

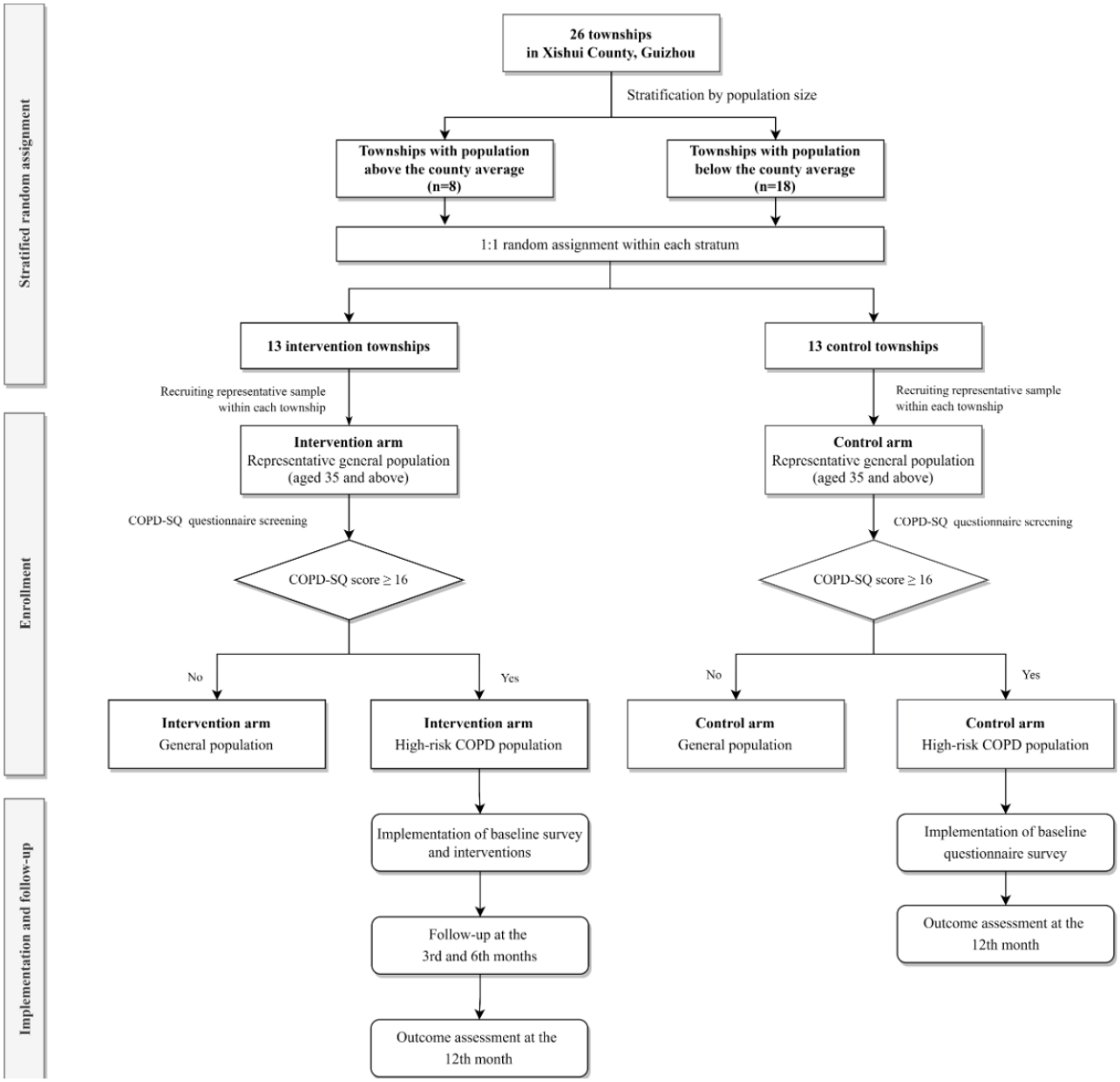

Our clusters are the townships of Xishui County. The demographic characteristics of Xishui County, including the number and size of townships, are summarized in Table 1. The townships were stratified according to whether their population size exceeded the mean size of all townships. Within each stratum, townships were randomly allocated in a 1:1 ratio to either the intervention or control arm using a computer-generated randomization sequence. Participant recruitment was based on a comprehensive roster of permanent residents aged 35 years and above, provided by the local government as of May 10, 2024, with individuals selected from each township for enrollment. Permanent residents were defined as individuals who had resided in Xishui County for at least 3 months and were expected to remain within the same township for the subsequent year at the time the government collected the information.

XSL•FO
RenderX

Table 1. Characteristics of Xishui County, Guizhou.

Characteristics	Xishui County
Total population size (10,000[a]), n	30.20[a]
Rural townships, n	22
Villages within rural townships, n	210
Average rural township size (ie, number of people), n	9518
Urban townships, n	4
Communities within urban townships, n	47
Average urban township size (ie, number of people), n	23,150

[a]The total population size listed in Table 1 was obtained from the Xishui authority. The figures here reflect permanent residents as of May 10, 2024, and are limited to individuals who had already stayed in Xishui County for 3 months in 2024 and would continue to stay within the same township for the next year.

The COPD Screening Questionnaire (COPD-SQ) was used to identify individuals at high risk for COPD. The COPD-SQ includes 7 items: age, cumulative smoking history, BMI, cough, breathlessness, family history of respiratory diseases, and exposure to cooking-related smoke [21]. The total score ranges from 0 to 38, with higher scores indicating a greater risk of developing COPD. This questionnaire has been validated through various epidemiological surveys and community health screenings and has been recommended for COPD screening in primary care settings in China [21]. Participants with a COPD-SQ score of 16 or higher were considered for inclusion in the study, ensuring that the cohort comprised a population at high risk for COPD.

Setting

This study was implemented across 26 township-level clusters in Xishui County, a mountainous area in northern Guizhou Province, China, where the prevalence of smoking was 37.9% in 2018-2019, the second highest among all provinces [22]. A rural area designated as China's National Comprehensive Primary Health Experimental Area, Xishui presents a unique combination of policy-driven innovation and socioeconomic challenges. With a gross domestic product per capita (48%) below the national average in 2024 and limited human capital development, the county exemplifies a resource-constrained rural setting. These conditions, combined with its experimental role in health system transformation, have established Xishui as an ideal setting for evaluating scalable multimorbidity interventions tailored to underserved populations.

Trial Participants (Inclusion and Exclusion Criteria)

All research participants were required to be at least 35 years old. In addition, eligible participants had to meet the following criteria: (1) a COPD-SQ score of 16 or higher, corresponding to individuals at high risk for COPD; (2) residence in a township of Xishui County for the prior 3 months and plans to remain in the same township for the upcoming year; and (3) provision of informed consent.

Individuals with severe cognitive disorders that significantly impair comprehension, decision-making, or adherence to the intervention were excluded. Similarly, those who had completely lost the ability to perform activities of daily living independently were not eligible for participation. These criteria ensure that the study includes individuals who can actively engage in the intervention and follow-up assessments while minimizing confounding variables.

Research Ethics Approval

This study received ethical approval from the Ethics Committee of Peking Union Medical College. To protect participant privacy, all identifying information has been removed from the dataset, and data are anonymized before analysis. Participants have the right to withdraw from the study at any point, without consequences for their standard health care access. The study is committed to upholding the highest standards of ethical research conduct, ensuring that all participants are treated respectfully and that their health and personal information are safeguarded throughout the research process. This study was initially approved by the Ethics Committee of Peking Union Medical College (approval number CAMS&PUMC-IEC-2024-040). A continuing ethics review was completed in June 2025, and updated approval was granted under approval number CAMS&PUMC-IEC-2025-061.

Intervention

Participants assigned to the intervention arm received access to the population medicine multicomponent intervention package, which was developed through iterative prototyping and stakeholder consultations. The specific eligibility criteria for each intervention are summarized in Table 2. Participants in the control arm were notified of their COPD risk status and encouraged to complete a face-to-face interview; they were not provided with any interventions after the interview, but were encouraged to receive usual care.

XSL•FO
RenderX

Table 2. Eligibility criteria for interventions.

Population level and target population	Intervention	Eligibility criteria
General population		
General population	Health education	Permanent residents
General population	Online screening with the COPD[a]-SQ	Permanent residents aged 35 years and above
Patients with COPD or high-COPD-risk population		
High-COPD-risk population	Community-based spirometry pulmonary function tests, interpretation of results, and health education	Individuals with COPD-SQ[b] score ≥16
Smokers within high-COPD-risk population	Health education for smokers to support smoking cessation	Currently smoking or have quit within the last 6 months
Smokers within high-COPD-risk population	Digital health interventions for smoking cessation	Currently smoking or have quit within the last 6 months and own a smartphone
High-COPD-risk population with depression and anxiety symptoms	Mental health digital health interventions	Individuals with Warwick Edinburgh Mental Well-being Scale score <45 who own a smartphone
High-COPD-risk population with mental health symptoms	Health education for smokers with mental health issues	Individuals with Warwick Edinburgh Mental Well-being Scale score <45
High-COPD-risk population with abnormal weight	Weight abnormality interventions	Individuals with BMI <18.5 kg/m^2 or BMI ≥ 24.0 kg/m^2
High-COPD-risk population with hypertension	Hypertension management and education	Three consecutive measurements with average systolic blood pressure ≥140 mmHg or diastolic blood pressure ≥90 mmHg or both
High-COPD-risk population with type 2 diabetes mellitus	Diabetes management and education	Fasting blood glucose ≥7.0 mmol/L or random blood glucose ≥11.1 mmol/L
Patients with COPD or high-COPD-risk population who fail to conduct pulmonary function tests	Encouragement to conduct computed tomography scan and seek professional medical treatment	Individuals with FEV_1[c]/FVC[d] <0.7 after bronchodilation, or high-COPD-risk population who fail to conduct a pulmonary function test for any reason
Health providers		
Health providers	Intrinsic incentive mechanism	Primary care providers who engage with the investigation and intervention
Health providers	Extrinsic incentive mechanism	Primary care providers who engage with the investigation and intervention

[a]COPD: chronic obstructive pulmonary disease.

[b]COPD-SQ: COPD Screening Questionnaire.

[c]FEV_1: forced expiratory volume in 1 second.

[d]FVC: forced vital capacity.

The strategies included in the multicomponent intervention package within this trial have been proven cost-effective and are advocated by the Lancet Commission on Investing in Health [2]. Figure 2 illustrates the full integration pathway of the intervention package implemented in this trial. The structure comprises 4 layers: the screening package, the targeted population, the integrated intervention components, and the long-term management package. The figure shows how the general population is first screened and stratified into risk groups, followed by targeted interventions (eg, spirometry, mental health support, and smoking cessation) and intensive follow-up. It visualizes the conceptual transition from individual treatment to proactive population health management.

XSL•FO
RenderX

Figure 2. Integrated pathway of the multicomponent intervention package. CBT: cognitive behavioral therapy; COPD: chronic obstructive pulmonary disease; COPD-SQ: COPD Screening Questionnaire; CT: computed tomography; WEMWBS: Warwick-Edinburgh Mental Well-Being Scale. See Multimedia Appendix 4 for a higher resolution version of this figure.

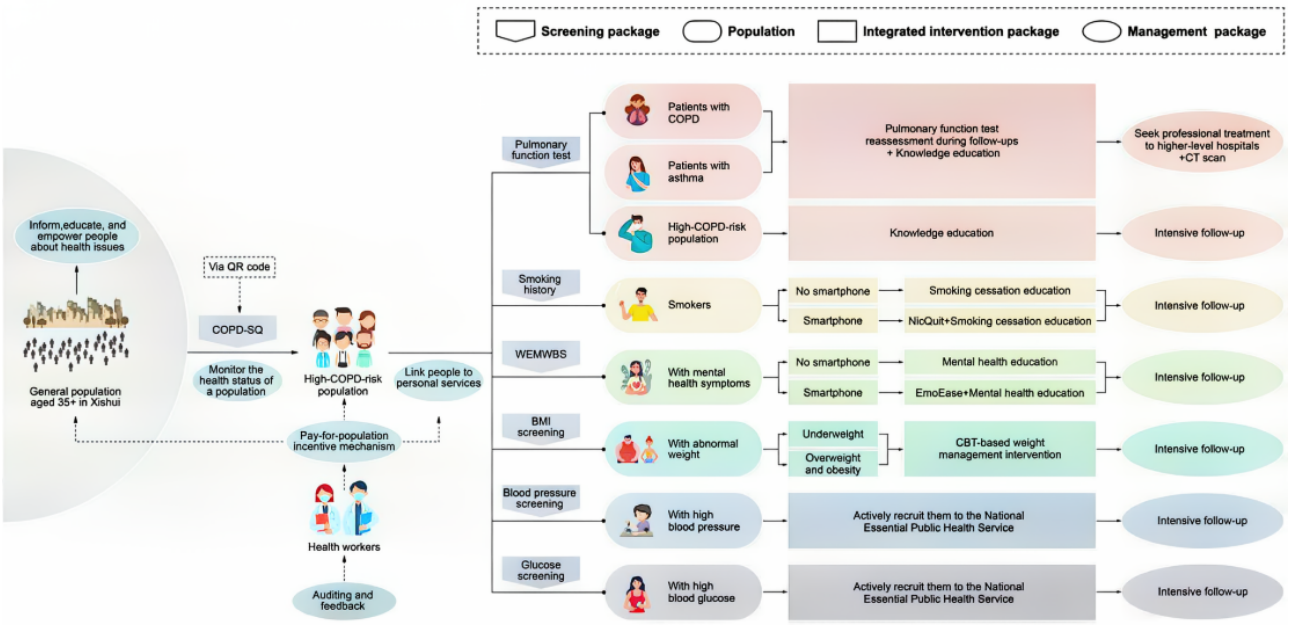

In addition to the intervention components and care pathways described in Figure 2, the implementation of the multicomponent intervention package was supported by a multilevel delivery structure and a performance-linked incentive mechanism. Figure 3 illustrates the organizational structure, population stratification, incentive design, and distribution of responsibilities across administrative levels (county, township, village, household) for delivering the intervention in Xishui. The figure also specifies the eligibility criteria for each target subpopulation, clarifying how risk stratification and service delivery were operationalized in the field.

Figure 3. Implementation structure and execution mechanism of the multicomponent intervention package. CDC: Centers for Disease Control and Prevention; COPD: chronic obstructive pulmonary disease; COPD-SQ: COPD Screening Questionnaire; ECRHS: European Community Respiratory Health Survey; FEV1: forced expiratory volume in 1 second; FVC: forced vital capacity; PCCM: pulmonary and critical care medicine; PFT: pulmonary function test; WEMWBS: Warwick-Edinburgh Mental Well-Being Scale. See Multimedia Appendix 4 for a higher resolution version of this figure.

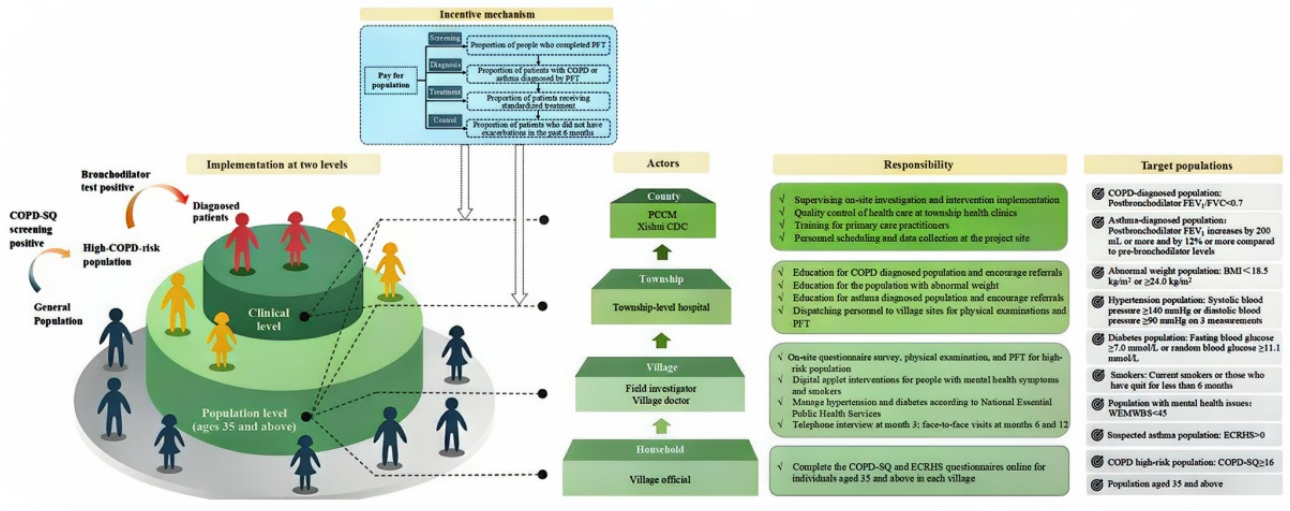

Specific Interventions

Overview

Specific interventions as part of the multicomponent intervention package are described in the following sections.

Health Education

Health education is provided to all permanent residents, regardless of their risk level. This intervention aims to increase population-wide awareness of tobacco-related NCDs, particularly COPD, and to promote early health-seeking behaviors. Health education is delivered through multiple channels, including face-to-face sessions led by trained primary care workers, printed brochures, posters in public areas, and

XSL•FO
RenderX

short videos/messages distributed via WeChat (Tencent Holdings Limited). Educational content covers the importance of lung health, risk factors such as smoking and indoor air pollution, the purpose of spirometry screening, and strategies for chronic disease prevention and lifestyle modification.

Online Screening for COPD

As part of the screening package, all representative permanent residents aged 35 years or older are invited to complete an online screening survey via a QR code–linked digital platform. The survey incorporates the COPD-SQ and is designed to allow broad dissemination and efficient data collection. Upon submission, participants receive automated feedback on their risk level and are guided to appropriate follow-up actions, such as spirometry appointments or health counseling sessions.

Community-Based Spirometry Pulmonary Function Tests, Results Interpretation, and Health Education

The population at high risk of COPD in the intervention arm receives alerts directing them to a community gathering place for spirometry testing, which is conducted using the BH-AX-MAPG spirometry equipment (BreathHome). Participants screened for COPD are referred to the county hospital for computed tomography and formal diagnosis. In addition, they receive further health education on the risks of COPD and strategies for disease prevention and management, delivered verbally by primary health care providers and supplemented with printed materials for distribution.

Health Education for Smokers to Support Smoking Cessation

Participants in the intervention group receive targeted health education to reinforce the importance of smoking cessation. This education emphasizes the health risks associated with smoking and the benefits of quitting, providing participants with evidence-based information and practical guidance. It is delivered through verbal communication by primary health care providers and supplemented with printed posters for distribution.

Digital Health Interventions for Smoking Cessation

NicQuit is a WeChat-based digital smoking cessation intervention that includes cognitive behavioral therapy (CBT) modules focused on smoking cessation strategies, coping with triggers, and reinforcement techniques to maintain abstinence. It is designed for smokers who are currently smoking or have quit within the past 6 months [23]. The intervention specifically targets individuals familiar with smartphone technology, ensuring accessibility and usability. Personalized notifications and reminders are delivered through the WeChat platform to encourage participants to engage regularly with the cessation plan and maintain adherence.

Mental Health Digital Health Interventions

A CBT-based digital mental health intervention, EmoEase [24], is offered to individuals experiencing mental health symptoms (Warwick-Edinburgh Mental Well-Being Scale [WEMWBS] < 45) who have a smartphone. This WeChat-based program includes psychoeducation, mood tracking, guided CBT exercises, and self-regulation techniques. The program emphasizes the link between mental health and respiratory symptoms, aiming to enhance coping, treatment adherence, and sustained behavior change.

Health Education for High-COPD-Risk Individuals With Mental Health Symptoms

Specialized health education is offered to high-COPD-risk individuals with coexisting mental health symptoms, tailored to their specific challenges. This education includes guidance on managing mental health symptoms and is delivered verbally by trained community health workers or general practitioners, supplemented with printed materials for distribution.

Weight Abnormality Interventions

Individuals with a BMI<18.5 kg/m^2 (underweight) or ≥24.0 kg/m^2 (overweight or obese) are considered to have weight abnormalities. CBT-based motivational interviewing is used to help participants self-identify weight-related barriers before providing tailored health education. During the intervention session, participants are asked CBT-informed questions designed to prompt reflection on the disadvantages and advantages of being underweight or overweight.

Hypertension and Diabetes Management

The goal of this intervention is to actively enroll smokers whose blood pressure exceeds 140/90 mmHg [25] or whose random blood glucose exceeds 11.1 mmol/L (or fasting blood glucose ≥ 7.0 mmol/L) into the National Essential Public Health Service in China [26]. These participants also receive health education on hypertension and diabetes through verbal counseling by trained community health workers or general practitioners, supplemented with printed materials for distribution.

Encouragement to Seek Professional Medical Treatment in Superior Hospitals

Patients with a postbronchodilator FEV_1-to-forced vital capacity (FVC) ratio of <0.7 are diagnosed with COPD and are encouraged to seek professional medical care at higher-level hospitals for further diagnosis and management. Some participants at high risk of COPD failed to complete pulmonary function testing due to contraindications, lack of understanding of the procedure, failure to meet spirometry quality-control standards, or unwillingness to undergo the tests. High-risk participants who do not complete pulmonary function testing for any reason are encouraged to undergo a computed tomography scan and seek formal diagnosis and medical treatment at a superior hospital.

Pay-for-Population Mechanism

A novel pay-for-population mechanism is introduced to motivate primary care providers to actively engage with the intervention by aligning compensation with population health goals. This approach ties provider payment to 4 key stages of care: screening, diagnosis, treatment, and control. Health providers in township hospitals are rewarded based on 4 performance indicators. The first is the proportion of individuals completing pulmonary function testing among township residents aged 35 years and above. The second is the proportion of individuals diagnosed with COPD among those identified as high risk in the initial screening (defined as COPD-SQ >16). The third is

XSL•FO
RenderX

the proportion of patients with confirmed COPD who receive standardized inhaled treatment, relative to the total number of patients with confirmed COPD in the township. The fourth is the proportion of patients with confirmed COPD who have not experienced an acute exacerbation in the past 6 months, relative to all patients with confirmed COPD in the township. The county hospital's respiratory department and the county Center for Disease Control and Prevention are assessed using the same 4 performance indicators, with data aggregated across all 13 intervention townships. In addition to providing an extrinsic, results-based financial incentive, providers are offered specialized training and capacity-building opportunities to appeal to intrinsic motivation, support proactive service delivery, and foster a strong sense of responsibility for population health. Details on implementing the pay-for-population mechanism are provided in Multimedia Appendix 2.

Outcomes

We will measure both primary and secondary end points at the end of follow-up for both trial cohorts. End points will be assessed among individuals assigned to the intervention group and compared with those among individuals assigned to the control group (ie, intent-to-treat). This high-COPD-risk population cohort will be followed longitudinally through March 2026 for repeated measurement of end points and covariates, according to the initial intervention versus control assignment. Specifically, all primary and secondary end points will be measured during health worker visits at 6 and 12 months after assignment, except for knowledge of diseases, health service utilization, and the number of chronic diseases controlled, which will be assessed only at the end of the trial (ie, 12 months after assignment). Table 3 defines the trial's primary outcomes, illustrating the functional forms of measured variables and the approach to assessing each end point, while secondary outcomes are detailed in Multimedia Appendix 3.

Table 3. Primary outcomes.

Outcome	Description
Number of chronic conditions controlled	• Definition: Number of controlled conditions among 7 objectively measured chronic health conditions: chronic obstructive pulmonary disease, asthma, depressive symptoms, anxiety symptoms, BMI, hypertension, and type 2 diabetes. • Functional form: Counting data • Measurement: Through objective physical examination or a validated scale
COPD[a] screening	• Definition: Response to the question "Have you ever had a pulmonary function test?" • Functional form: Binary • Measurement: Self-reported response
FEV_1[b] measurement	• Definition: Volume of air forcibly exhaled during the first second of a forced expiratory maneuver • Functional form: Continuous • Measurement: Pulmonary function test and portable spirometry

[a]COPD: chronic obstructive pulmonary disease.

[b]FEV_1: forced expiratory volume in 1 second.

Timeline

Our trial began on June 17, 2024, with recruitment conducted across all townships in Xishui County. Participants in the intervention arm undergo structured assessments at 4 time points: baseline (in-person), 3-month follow-up (telephone), 6-month follow-up (in-person), and 12-month follow-up (in-person). At baseline, following completion of informed consent, participants were invited to undertake a comprehensive in-person assessment. Physical examinations and pulmonary function tests are conducted at multiple points, with protocols tailored by subgroup: prebronchodilator spirometry at baseline and 12 months for the high-COPD-risk population; postbronchodilator spirometry at baseline and 12 months for patients with COPD; and prebronchodilator spirometry at 6 months for high-COPD-risk participants with prebronchodilator $FEV_1/FVC<70\%$ at baseline. All baseline assessments were administered under the supervision of trained site staff to ensure accuracy and adherence to protocol standards. Subsequent follow-ups include a 3-month telephone assessment to monitor intervention adherence and health status. In-person follow-ups at 6 and 12 months repeat elements of the baseline assessment, including the questionnaire, physical examination, and pulmonary function testing. For the control group, the schedule includes a baseline questionnaire and a 12-month on-site survey. Table S2 in Multimedia Appendix 3 presents the timeline of prospective cohort visits and data collection.

Sample Size

We included all 26 townships in Xishui County and allocated them to the intervention or control group using 1:1 stratified randomization. Based on the population size of each township, we recruited a representative sample of 44,000 residents aged 35 years or older. All 44,000 participants underwent COPD-SQ screening via QR code. Considering a 10% attrition rate and an estimated high-risk prevalence of approximately 25% among individuals aged 35 years and above [27], the projected final enrollment is 10,000 individuals in the high-COPD-risk population.

XSL•FO
RenderX

Because of the relatively large number of important secondary end points and the high degree of uncertainty regarding the prevalence and risk factors of COPD in rural and urban communities, we chose FEV_1 as one of the primary end points of interest and conducted the power calculation conservatively. This choice was made because FEV_1 is an objective and clinically relevant measure that helps address uncertainties in previous studies while allowing for a more robust power calculation. Given that 3 outcomes are designated as primary end points, we applied a Bonferroni adjustment for both sample size and power calculations. According to the China Pulmonary Health Study—a cross-sectional study in a nationally representative sample of adults aged 20 years or older—the overall prevalence of spirometry-defined COPD is expected to be 8.6% (95% CI 7.5%-9.9%) [10]. Based on a study conducted by Chen et al [28], the mean and SD of FEV_1 levels among patients with mild COPD and the high-COPD-risk population in China were reported to be 2.62 L and 0.55 L, respectively. Following standard procedure, the intraclass correlation coefficient (ICC) was identified in advance, as it is a key parameter required for accurate sample size estimation in cluster randomized designs. Previous literature reports an ICC of 0.055 for patients with COPD in the PACK Brazil Study and 0.05 in the Danish TeleCare North Study [29]. Thus, we set the ICC at 0.05 for sample size calculation. We followed the formula for population-based stratified cRCTs proposed by Crespi [30] and assumed a 2-sided significance level of 1.67% under the Bonferroni adjustment and a power of 80%. Based on these assumptions, we calculated a conservative sample size estimate of 2850 individuals in the high-COPD-risk population, which is required to detect a minimal difference of 0.1 L in FEV_1 in our study.

Power Calculations

In our power calculations, we estimated the minimum detectable differences (MDDs) in FEV_1 (L) levels under varying sample sizes and ICC values, assuming a 2-sided significance level of 5% and a power of 80%. We further assumed that 10% of participants would refuse to participate in the trial and that 10% of individuals could not be followed up for end point assessment. To our knowledge, this is the first real-world, large-scale multimorbidity implementation trial, and no similar trials are available for reference. Therefore, we conducted a sensitivity analysis to examine how the MDDs in FEV_1 levels change with different proposed population sizes. The results of this sensitivity analysis are presented in Table 4. Our analyses indicate that the population-based cluster randomized trial design can detect differences of 0.052-0.218 L in FEV_1 levels between the treatment and control groups. In the implementation, we applied stratification before randomization to reduce between-cluster variance.

Table 4. Minimum detectable differences in forced expiratory volume in 1 second (liters) across ICC[a] values and proposed sample sizes.

Proposed sample size, n	ICC=0.005	ICC=0.01	ICC=0.025	ICC=0.05
5000	0.089	0.113	0.165	0.226
5500	0.085	0.108	0.157	0.216
6000	0.081	0.103	0.151	0.207
7000	0.075	0.095	0.139	0.191

[a]ICC: intraclass correlation coefficient.

As shown in Table S2 in Multimedia Appendix 3, we listed 3 primary outcomes and a large set of secondary outcomes. For primary outcomes, we will apply the Bonferroni adjustment, setting the significance threshold at 0.0167 (0.05/3). For secondary outcomes, to mitigate concerns of type I error inflation, we plan to use the Benjamini-Hochberg procedure to control the false discovery rate. We will set the false discovery rate at 0.05 and rank all unadjusted *P* values of the secondary outcomes from smallest to largest. Adjusted *P* values will then be reported following the Benjamini-Hochberg procedure, and the significance of each secondary outcome will be determined based on these adjusted *P* values.

In Table 5, we present a sensitivity analysis reporting Cohen *d* for the trial. We set the SD to 1 to explore Cohen *d* (standardized MDDs) under different ICC values and proposed sample sizes.

Table 5. Minimum detectable differences in standardized mean difference (Cohen d) across ICC[a] values and proposed sample sizes.b

Proposed sample size, n	ICC=0.005	ICC=0.01	ICC=0.025	ICC=0.05
5000	0.162	0.205	0.300	0.412
5500	0.155	0.196	0.286	0.393
6000	0.148	0.188	0.274	0.376
7000	0.137	0.174	0.253	0.348

[a]ICC: intraclass correlation coefficient.

[b]Minimum detectable difference values represent standardized mean differences (Cohen *d*), calculated by assuming a pooled SD (σ) of 1 for the outcome.

XSL•FO
RenderX

Assignment of the Intervention

This cRCT assigned participants to either the population medicine multicomponent intervention or the control arm at the township level, with randomization stratified by geographic region to ensure balance across community characteristics. A computer-generated algorithm was used to randomly allocate 13 townships to each arm, with the sequence generated by an independent statistician who was not involved in recruitment or intervention delivery. Given the nature of the intervention, this is an open-label trial, meaning health care providers and research staff are aware of group assignments. However, participants themselves were not explicitly informed of their allocation or that they were part of an "intervention group," to mitigate expectation bias. No additional interventions were provided to the control group, although usual care remained accessible throughout the study period.

To enhance adherence to the intervention protocol, a structured approach was integrated into all follow-up phases. Adherence was reinforced through telephone consultations at the 3-month follow-up and through an in-person visit at the 6-month follow-up. During these sessions, participants received personalized feedback on their health outcomes (eg, spirometry results, smoking status), including comparisons with baseline data to highlight improvements or areas requiring attention. For example, if a participant had not sought professional medical treatment for COPD, a root-cause analysis was conducted by reviewing baseline results and exploring potential barriers, such as lack of motivation or difficulty accessing health care. Adherence was systematically monitored at each follow-up through validated questionnaires assessing compliance with the intervention (eg, whether the participant had sought care, app usage frequency, smoking cessation attempts). In addition, automated reminders and tailored counseling were provided to address gaps in adherence. This iterative process ensured that each follow-up served as an opportunity to motivate participants, troubleshoot challenges, and adapt strategies to sustain engagement throughout the trial.

Blindness

This is an open-label cRCT, in which field workers and participants are not blinded. We consider this scientifically reasonable for large-scale population health intervention trials conducted in real-world community settings, such as the POPMIX trial. However, we made efforts to blind most members of the Data and Safety Monitoring Board (DSMB). Statisticians will remain blinded when conducting data analyses, and DSMB members will not be informed of the exact intervention assignment of each township. Thus, while full blinding was not feasible, we maintained investigator blinding wherever operationally possible, consistent with best practices in open-label population health and behavioral intervention trials.

Data Collection Plan

We are collecting trial data across multiple dimensions to comprehensively assess the impact of the multicomponent intervention (Table 2). At baseline, all participants completed a structured questionnaire capturing demographics and socioeconomic status, quality of life, lifestyle and risk behaviors, disease history, mental health status (9-item Patient Health Questionnaire [PHQ-9], 7-item Generalized Anxiety Disorder [GAD-7], and WEMWBS), and productivity loss. In addition to self-reported data, participants underwent physical examinations, including measurements of height, weight, BMI, waist circumference, heart rate, blood pressure, and blood glucose. Pulmonary function was assessed using spirometry to evaluate respiratory health and detect airflow limitation. For all participants in the high-COPD-risk population, prebronchodilator spirometry was conducted at baseline and again at 12 months to monitor changes in lung function over time. For participants diagnosed with COPD, both pre- and postbronchodilator spirometry were performed at baseline, with prebronchodilator tests repeated at 6 months and both pre- and postbronchodilator tests repeated at 12 months.

Follow-up data are collected at 3 time points: 3 months (via telephone), 6 months (in-person), and 12 months (in-person) for the intervention group and only 12 months for the control group (in-person). At each in-person follow-up, participants complete key sections of the baseline assessments, including self-reported data on COPD symptoms and management (lung function via spirometry, exacerbations, treatment adherence, COPD Assessment Test, and Modified Medical Research Council), asthma symptoms and management, high-risk behaviors (eg, daily cigarette consumption and smoking dependence, alcohol use), and chronic disease status (blood pressure, blood glucose, and BMI). Mental health status is reassessed using the same scales (PHQ-9, GAD-7, and WEMWBS), and health-related quality of life is measured using the EQ-5D-5L and the Saint George Respiratory Questionnaire for patients with COPD. Additionally, lifestyle behaviors, such as physical activity, alcohol consumption, and diet (including sugar, salted vegetable, and general vegetable intake), are monitored through self-report. Health resource utilization, including outpatient and inpatient visits, is also recorded. Data are collected by trained field staff using standardized protocols to ensure accuracy and reliability. All data are entered electronically via an electronic data capture (EDC) system, which includes built-in validation checks to minimize data entry errors.

Data Management

This study uses an EDC system to ensure real-time data entry, automated validation, and secure storage. Data are collected by trained field staff using tablet-based digital surveys that sync directly with a central database. The EDC system incorporates range checks, logical consistency validations, and missing data alerts to minimize errors, with flagged inconsistencies promptly reviewed by the data management team. Data collection forms, covering demographic, clinical, and behavioral assessments, are standardized and precoded, while biometric data (eg, pulmonary function, blood pressure, BMI, and blood glucose) are recorded directly into the system. Spirometry data are transmitted from portable devices and subjected to automated quality control checks to ensure validity.

We store data using numeric identifier codes to maintain participant anonymity in a password-protected database

XSL•FO
RenderX

accessible only to the research team for analysis. Access is restricted to a minimal number of individuals, including the principal investigator, co-investigators, biostatisticians, and data scientists.

Statistical Analysis

Primary analyses will compare outcomes between baseline and the 12-month follow-up using generalized linear mixed models (GLMMs) with township as a random effect under the intention-to-treat principle. To evaluate the robustness of the findings, intervention effects will be estimated using a primary design-adjusted model, which will adjust for stratification factors and baseline values of the outcome measure. This will be complemented by a sensitivity-adjusted model that will further incorporate demographic covariates, such as age, sex, and education level. Adjustment for multiple testing will be applied using the Bonferroni correction and Benjamini-Hochberg procedure.

To evaluate the effects of individual intervention components and explore which components drive specific outcomes, we designed a series of regression discontinuity design analyses, planned for publication after the main trial results. We established 7 clear diagnostic thresholds for the conditions included in our trial. For COPD, the threshold is an FEV_1-to-FVC ratio of 70%. For PHQ-9 and GAD-7, a score of 10 is used to indicate the presence of depressive and anxiety symptoms, respectively. For the WEMWBS questionnaire, a threshold of 45 is used to indicate whether general mental health is adequate. For hypertension, clear thresholds are defined as 140 mmHg for systolic blood pressure or 90 mmHg for diastolic blood pressure. Individuals with a BMI below 18.5 kg/m^2 are considered underweight, while those with a BMI above 24 kg/m^2 are considered overweight. For high blood glucose, thresholds of 7.0 mmol/L for fasting glucose or 11.1 mmol/L for random glucose indicate potential type 2 diabetes mellitus. These well-defined thresholds allow us to conduct regression discontinuity design analyses to evaluate the component-specific effects of the integrated intervention package within the intervention arm.

To better elucidate how the intervention package will reshape the behaviors of both high-COPD-risk populations and health care providers, we also designed a series of qualitative studies involving interviews with high-COPD-risk populations, physicians and community health workers, and policy makers in the intervention arm. These studies will serve as important supplements to the main trial by illustrating the mechanisms of the intervention and the subjective experiences of being intervened under the population medicine framework.

Monitoring

We appointed an independent DSMB to provide scientific rigor and ensure an independent evaluation of safety. The DSMB comprised leading experts in public health, epidemiology, clinical research, and respiratory disease. The DSMB convened 3 meetings in which baseline and follow-up descriptive analyses for the intervention arm were evaluated for process monitoring. The DSMB convened for a formal interim analysis in February 2026. At this meeting, the research team reviewed the conduct of the trial, including adherence to intervention components, logistical issues, and the completeness and quality of the collected data. Our biostatisticians also presented emerging effect estimates for the intervention and control arms to the steering committee. Stopping rules were established regarding safety and study completion:

- If deaths or severe adverse effects were reported, these would be immediately conveyed to the DSMB. The DSMB would determine if the deaths or adverse effects were related to the intervention and advise if the study should be stopped immediately.
- In February 2026, during the planned interim analysis, the trial was approaching its completion as field workers confirmed that the majority of eligible participants had been contacted. At this stage, the results were assessed to determine the study's conclusion based on the following:
  - Null findings, suggesting that the intervention had no detectable effect and indicating that the study should be stopped.
  - Statistically significant findings, indicating that the study had achieved its primary objectives and should be stopped.
  - Emerging but statistically insignificant effects, suggesting that the study might require further observation if applicable. However, the study would be terminated if field workers confirmed that all eligible participants had been contacted and no additional individuals remained to undergo follow-up assessments.

Plans for Communicating Important Protocol Amendments to Relevant Parties

The Ethics Commission of the Medical Faculty of Peking Union Medical College was contacted and notified of any protocol amendments requiring its approval. Following approval by the Ethics Committee and before implementation, all research team members were informed of the protocol amendments.

In this trial, an unavoidable amendment was made to the secondary outcome measures. Although outcome modifications are generally discouraged, this change was necessitated by a nonnegotiable regulatory restriction. Specifically, we initially planned to measure health care utilization (ie, outpatient visits, inpatient visits, and medical expenditure) using electronic health records from medical institutions in Xishui County. However, due to unexpected government regulations, we were unable to reach an agreement to obtain these data. This amendment was approved by the Ethics Committee, and all research personnel were notified before implementation to ensure consistency in data collection and analysis.

Consent and Withdrawal

We provided written study information and the informed consent form to eligible individuals. The field workers explained the study's aims and detailed procedures, in the presence of a witness if required. Participants were given sufficient time to decide whether to participate in the study and the opportunity to ask questions, which were answered accordingly. For illiterate participants, consent was documented before enrollment using a thumbprint signature and a witness's signature.

XSL•FO
RenderX

We informed eligible participants about the risks and benefits of participation in the trial. A decision not to participate in the study did not carry any consequences for the individual. All participants were free to refuse or discontinue data collection at any stage, and participation did not require relinquishing any concomitant care. Although there were no predefined criteria for modifying or discontinuing allocated interventions, reasons for attrition will be documented and reported in future publications.

Confidentiality

Throughout the study, all data will be handled with strict confidentiality, and data collection will conform to the requirements of national data protection legislation. Digital data will be stored in password-protected files at the Center of the Chinese Academy of Medical Sciences and will be accessible exclusively to research team members. Third parties may receive anonymized data solely for research purposes. Informed consent forms, laboratory books, and other participant-related documents will be securely stored during the conduct of the study and will subsequently be stored at the co-principal investigators' (CW, SC, and TY) office premises.

Dissemination

Upon completion of the trial, participant-related information and results will be provided to the participants themselves or to their families. The trial has gained public attention since it was first introduced in the Lancet Commission on Investing in Health and now serves as a critical component of the ongoing Lancet Commission Report on Population Medicine and Public Health. We will disseminate the trial results through international peer-reviewed journals and at other renowned international academic conferences.

Results

Data collection for the POPMIX-COPD trial began in June 2024. Baseline, 3-month, and 6-month assessments have been completed. The trial is planned to finish data collection by March 2026. Data analyses will be conducted after all data have been collected and the dataset has been locked. We plan to publish the POPMIX-COPD paper before mid-2027. At the time of protocol submission, the trial status is active. Recruitment began on June 17, 2024, and was completed on December 31, 2024. Follow-up data collection will conclude on March 31, 2026.

Discussion

To the best of our knowledge, this is the world's first large-scale, community-based population medicine multimorbidity intervention study targeting tobacco-related NCDs among a high-COPD-risk population. As The Lancet Commission on Investing in Health has suggested, tobacco use–related NCDs are among the leading modifiable risk factors contributing to life expectancy gaps in China. However, awareness of tobacco-related NCDs remains low; China is the world's largest producer and consumer of tobacco, and the health and economic burdens of tobacco-related NCDs are substantial, especially in rural China [31]. Against this context, our population medicine trial is designed to improve population health and maximize collective welfare. Through this study, we aim to evaluate whether an integrated multimorbidity intervention based on population medicine theory can enhance health outcomes and awareness of tobacco-related NCDs in a relatively low-resource setting such as Xishui County.

Our population medicine trial emphasizes population health rather than individual outcomes. Unlike most clinical trials that predominantly focus on patients actively seeking care within the health system, our study targets high-COPD-risk individuals in the general population and applies community-level interventions through the use of the COPD-SQ. Early detection and intervention to prevent disease progression among the high-COPD-risk population, including individuals with undiagnosed conditions or unmet care needs, are essential for improving population health.

We also extended the scope of our study to population health outcomes, such as multimorbidity. We specified the "number of chronic diseases controlled" as a primary outcome to capture the intervention's effects not only on COPD but also on common co-occurring conditions, such as hypertension, diabetes, smoking dependence, abnormal BMI, asthma, and symptoms of depression and anxiety. Given that our target population is a high-COPD-risk population, we considered measurement of FEV_1 to be another essential primary outcome. Finally, we included whether participants had ever received lung function testing as a primary outcome to highlight the importance of community-level screening. We additionally listed a series of secondary outcomes, such as care cascade performance for COPD, asthma, hypertension, and diabetes; health care resource utilization; mental health status; knowledge of various conditions; high-risk behaviors and lifestyles; and socioeconomic profile, to facilitate evaluation of the intervention's effects across multiple dimensions of population health, including physical, mental, social, and environmental well-being.

Our trial leverages both intrinsic and extrinsic motivation to encourage health workers to support population health. With respect to intrinsic motivation, we provide professional training focused on actively addressing unmet population health care needs. With respect to extrinsic motivation, we incentivize workers through pay-for-population mechanisms that emphasize care cascade indicators. We acknowledge that implementing a new integrated intervention package entails substantial implementation costs, including financial resources as well as the time and effort of primary care providers, and therefore invested in a 3-month program of courses and training in Xishui County covering population medicine theory, survey skills, pulmonary function testing, and communication. Our experience suggests that population medicine and clinical medicine are complementary rather than contrasting. Population-level activities conducted in the community enable the timely detection of previously undiagnosed individuals at high risk of COPD and facilitate their referral for professional clinical management. By structuring these activities as community care delivered by teams of primary care physicians and village health workers, with many tasks (eg, blood pressure and BMI

XSL•FO
RenderX

measurement, CBT-based interventions, and smoking cessation interventions) not requiring physicians, the integrated package creates a strong linkage between community and clinical care, supporting earlier and more efficient disease management in low-resource settings.

Our trial has strong policy implications. First, we aim to evaluate whether a multicomponent population medicine intervention is suitable for delivery in community and primary care settings in China. The intervention package represents one potential approach to reforming care at both levels. Second, we seek to encourage policy makers to prioritize population-based screening and long-term management at the community level, rather than focusing primarily on hospital-based treatment. Third, we aim to address unmet health care needs related to tobacco-related NCDs among the high-COPD-risk population, as many high-risk individuals lack awareness of both COPD and common co-occurring conditions, such as asthma, hypertension, and mental health symptoms. Through intensive counseling, health education, and encouragement to seek care, we emphasize the importance of early detection and long-term management in low-resource settings. Fourth, this trial encourages primary and community health care workers to shift from a passive to a proactive mode of practice, in line with the principles of population medicine. Population medicine calls on health professionals to provide care beyond the hospital setting, with community-level screening and referral to the health system playing a central role. Accordingly, this trial asks medical professionals to expand the scope of their responsibilities across multiple dimensions. Specifically, it calls for the provision of holistic care spanning health promotion, screening, disease prevention, diagnosis, treatment, control, and rehabilitation. It also encourages a focus on multimorbidity rather than the separate treatment of single diseases, and an expansion of care delivery beyond hospitals to include communities and populations. This integrated approach may provide valuable guidance for policy development and health care planning. Building on these policy implications, our study further envisions a paradigm shift in the role of primary and community health workers within population medicine.

Specifically, we hope that this study will catalyze a paradigm shift in the role of primary and community health workers. In conventional health systems, frontline providers often operate passively, waiting for patients to present at clinics. This trial introduces mechanisms such as pay-for-population, performance-linked process metrics, and home-based outreach to actively engage providers in population screening, health promotion, and continuous care. This transformation, from a patient-centered, reactive model to a proactive, population-centered approach, lies at the core of population medicine. We believe that this shift has the potential to reshape public health governance not only in China but also in other low- and middle-income countries striving for equitable and sustainable health system reform.

This protocol describes the design and implementation of the POPMIX-COPD trial, the first large-scale, community-based, population medicine–oriented multimorbidity intervention targeting a high-COPD-risk population in China. By integrating population-level screening, proactive disease management, patient education, and performance-linked provider incentives, the study aims to address critical gaps in awareness, early detection, and integrated care for tobacco-related NCDs. Conducted in a resource-constrained setting, this trial has the potential to generate robust evidence for scalable, cost-effective models of multimorbidity management. The findings are expected to inform policy development, improve health care delivery, and promote population health not only in China but also in other low- and middle-income countries striving to achieve equitable and sustainable health system reforms.

Acknowledgments

This work was supported by the People’s Government of Xishui County, the Health Bureau of Xishui County, the Youth League Committee of Guizhou Medical University, Medieco Group Co., einPaper Team, the Finance Bureau of Xishui County, the Health Insurance Bureau of Xishui County, the Bureau of Sports and Education of Xishui County, the Bureau of Statistics of Xishui County, the Xishui Sub-Bureau of the Ecology and Environment Bureau of Zunyi City, the Water Authority of Xishui County, the Bureau of Housing and Urban-Rural Development of Xishui County, the Meteorological Bureau of Xishui County, the Centers for Disease Control and Prevention of Xishui County (Health Surveillance Station of Xishui County), the People’s Hospital of Xishui County, the Traditional Chinese Medicine Hospital of Xishui County, the Maternal and Child Health Care Hospital of Xishui County, and the People’s Governments (Offices) of 26 townships (subdistricts) and village committees of Xishui County. Y Liu acknowledges support from the China Scholarship Council Program (Project ID 202406210195). The authors declare the use of generative artificial intelligence in the research and writing process. According to the GAIDeT (Guideline for Reporting Evidence-based Practice Educational Interventions and Teaching) standards (2025), the following tasks were delegated to generative artificial intelligence tools under full human supervision: proofreading and editing, adapting and adjusting emotional tone, reformatting, and quality assessment. The generative artificial intelligence tool used was ChatGPT 5.1 (OpenAI) and DeepSeek (High-Flyer). Responsibility for the final manuscript lies entirely with the authors. Generative artificial intelligence tools are not listed as authors and do not bear responsibility for the final outcomes. The following are the collaborators of this work: The People's Government of Xishui County, Health Bureau of Xishui County, Youth League Committee of Guizhou Medical University, Medieco Group Co. einPaper Team, Finance Bureau of Xishui County, Health Insurance Bureau of Xishui County, Bureau of Sports and Education of Xishui County, Bureau of Statistics of Xishui County, Xishui Sub-Bureau of Ecology and Environment Bureau of Zunyi City, Water Authority of Xishui County, Bureau of Housing and Urban-Rural Development of Xishui County, Meteorological Bureau of Xishui County, Centers for Disease Control and Prevention of Xishui County (Health

XSL•FO
RenderX

Surveillance Station of Xishui County), People's Hospital of Xishui County, Traditional Chinese Medicine Hospital of Xishui County, Maternal and Child Health Care Hospital of Xishui County, The People's Governments (Offices) of 26 Townships (Subdistricts), and Village Committees of Xishui County. The following individuals should also be considered study collaborators: Chun Zhang, Yuzhu Ye, Mingqiang Hou, Shengbo Liao, Yuan Zhou, Dong Liu, Chentao Zhong, Yushuang Zhao, Hongxia Mu, Min Liu, Qiong Lyu, Zhengyu Zhang, Weiwei Li, Jian Wang, Tiansheng Lan, Tingyu Liu, Xin Zhou, Bie Yu, Xianping Wang, Yansong Luo, Wenwu Jiang, Wen Zhang, Dongmei Wu, Qingping Luo, Dajun Rao, Cunkun Yang, Feng Zhao, Qin Wang, Linlin Zhang, Can Li, Min Chen, Yufeng Huang, Guanghai Jian, Li Wang, Yunjiang Zheng, Jianghong Linghu, Li Zhong, Xiaoling Ma, Bin Gui, Yan Cheng, Yu Yang, Jing Yan, Xixi Liao, Mingguo Xu, Zhijun Zhao, Qian Han, Panyan Wu, Dandan Tang, Qingqing Wang, Li Feng, Lilan Wang, Lei Si, Peng Zeng, Jianxiu Tian, Weili Fan, Xin Luo, Sumei Feng, Rong Ding, Huadan Yi, Fengxue Lyu, Ludan Yuan, Yuan Liu, Xueshuang Wang, Aihong Duan, Yang Zhao, Yuping Luo, Yicheng Li, Anchun Liang, Chengmei Li, Hu Tang, Anfeng Liu, Xu Wang, Yi Zhong, Wen Zheng, Benwei Zhao, Qianmei Zhang, Xiaojing Tu, Huanhuan Lu, Xiumei Zhao, Yan Chen, Xuemei Yuan, Yanhong Mu, Yao Zhao, Juan Zhang, Qinhui Li, Fuhui Wang, Xiaomin Mu, Yan Hu, Liuyang Chen, Gang Lei, Zhufei Hu, Xiangyong Tang, Yan Luo, Shunxia Zhou, Ting Pan, Yuanyin Zhou, Xiaoqin Ren, Yaling Zhou, Jing Zeng, Keping Hu, Xiaoyu Liu, Deguo Yang, Qiong Wan, Cai Zhang, Heng He, Chengping Wang, Yuanzhong Wang, Shuyin Wang, Xuhong Lei, Yushen Yuan, Tailiang Luo, Dazhou Zhu, Anrong Luo, Xianmei Hu, Ya Zhang, Ling Yuan, Senlin Luo, Weili Ma, Yonghua Cao, Yunpeng Cao, Taojin Chen, Wen Chen, Zhiqun Chen, Zhongyi Chen, Huagang Cheng, Xianpan Feng, Taimei Gao, Lizhu Jian, Renxian Jiang, Shengmei Kang, Hang Lei, Shanshan Liu, Ruxun Lu, Xiaojiao Luo, Dongshun Lyu, Guanghong Meng, Jianping Mu, Xuerou Mu, Hongxia Qian, Huimin Ren, Juan Wang, Li Wang, Loutao Wang, Qin Wang, Wei Wang, Ya Wang, Yan Wang, Dongjie Wu, Xiaoxia Xu, Jing Yang, Qian Yu, Dan Yuan, Yanlu Yuan, Yu Yuan, Hongying Yue, Yuanmin Yue, Li Zhang, Zhizhong Zhang, Jiaqin Zhao, Yuanxin Zhao, Yue Zhong, Lejie Zhou, Lufang Yang, Xue Chen, Lihui Ren, Guixiang Gong, Ping Shi, Jiangshun Tan, Zhonghang Yu, Lijun Mu, Meilin Chen, Su Yu, Lin Li, Xianbi Mu, Jianxia Wu, Mei Zhao, Jiangmei Chen, Jing Huang, Yanfang Li, Huayan Yang, Jiaheng Yang, Xingqi Yuan, Yunhao Zhou, Mengting Ye, Junlin Zhang, Taijun Luo, Shuai Wang, Ju Hu, Gang Yuan, Shili Mu, Zhongxiang Yuan, Huili He, Jiali Li, Chunlan Liu, Kailiang Cheng, Quan Liu, Huade Yang, Shaojun Zhang, Bo Chen, Guishan Ying, Xumei Zhao, Zhi He, Gan Feng, Huanhuan Guo, Hongling Yuan, Fuyou Yuan, Wei Wang, Lang Chen, Liping Zhao, Jiajun Ning, Jie Feng, Yuxian Chen, Jin Chen, Jing Qi, Yuanting Li, Lingli Zhao, Guo Wang, Ziyi Wang, Huaishu Zhong, Jie Wang, Gang Chen, Xueqin Mu, Yuanfeng Gong, Xinan Lu, Zhengli Chen, Tianwei Chen, Xiaoping Chen, Runqin Mu, Xiaoyi Deng, Yu Duan, Xiujuan Dong, Wangyong Yan, Jiayong Rao, Mingfang Wang, Xiangcai Chen, Zhongquan Wang, Lanfang Zhang, Jianghua Luo, Kaijie Huang, Ye Chen, Jin'e Li, Li Huang, Ju Chen, Hao Huang, Qingqun Long, Yan Chen, Ling Jiang, Man Ding, Xiaolin Yu, Tongjie Wei, Zhengqin Liu, Kaifu Si, Yuxian Mu, Yan Huang, Qian Yang, Tao Xu, Xiaorong Yuan, Cailun Zhao, Minli Chen, Rongfang Luo, Ling Chen, Ying Zhang, Guiwei Yuan, Qianyi Shui, Yong Wu, Qian Huang, Changli Wu, Nan Zhang, Chaojiang Wu, Defei Chen, Wu Xu, Linhui He, Wanlun Zhang, Jinsong Chen, Jin Yuan, Rujun Li, Minghe Wang, Chunmei Zhang, Yueshan Wang, Lang Liao, Anqian Feng, Dengyang Wang, Zelin Deng, Hupiao Yang, Kaimeng You, Lijuan Deng, Qiong Huang, Xiaoqing Zhang, Zhengqian Zhu, Xiaoyan Liao, Yongbo Song, Yushui Yuan, Qinghong Cai, Lu Liu, Hong Yu, Xue Luo, Xiaoshan Huang, Qin Hu, Dian Yuan, Zengye Zhao, Jiemin Zeng, Zhaoyan Huang, Peng Luo, Taibo Luo, Wenwu Wu, Guangni Liu, Gengning Zhang, Mingyuan Ren, Mei Yang, Zhongjin Wang, Rongping Yuan, Yunfeng Yuan, Anquan Yuan, Yunqiang Chen, Hongxian Liu, Xingyu He, Tongxiang Zhang, Mingkai Xu, Minqing Hu, Chunyi Qian, Yunfen Zhang, Yueguang Lei, Jun Wang, Xiaoju Yuan, Changxian Yang, Ai Xiang, Minjing Zhao, Xiaomei Ren, Guimei Li, Huiying Ren, Ju Teng, Yu Chen, Jing Yuan, Shunfeng Mu, Xiangyu Zhang, Ye Li, Yuwei Sun, Shangjing Wang, Fangping He, Guoxiang Wang, Wei Cai, Weijiang Tian, Tingfen Cai, Zhao Chen, Dakai Mu, Yufu Mu, Cai Cheng, Xiaofang Pan, Fang Wang, Mingqian Zhao, Qijun Liu, Bangqi Yang, Ye Yu, Ting Min, Xiaoya Wang, Jun Li, Yang Zhang, Yan Wang, Lihong Zhao, Zhihua Luo, Hao Yuan, Huihua Ruan, Jian Wu, Chengfen Wang, Dezhi Zhang, Yunfang Zhang, Weining Ao, Pan Zhao, Yunxia Liang, Yicheng Zhao, Xiaosong Mu, Lyufang Luo, Ziyi Zhang, Wei Huang, Jun Cao, Huan Li, Jun Liu, Liye Yuan, Na Zhao, Ju Wang, Qiuyan Duan, Weizhong Chen, Lirong Zhao, Jiangbo Wu, Liqin Deng, Kunyu Wang, Zhengguo Yan, Chaolian Wang, Can Liu, Quan Cao, Gang Ma, Youyong Lu, Yuan Zhang, Kaijing Xiang, Yixiang Wang, Xingye Zhu, Jiamei Wang, Demin Hu, Dongmin Li, Peng Chen, Chao Tian, Lu Wang, Mei Wang, Lulu Xiao, Lingling Yu, Min Wei, Chengcheng Yang, Limei Zhang, Rong Zhao, Chongguang Zhong, Yan Xiao, Jian Luo, Hongfa Chen, Pengju Dai, Jin Liu, Yingtao Luo, Guangjie Ma, Shiyou Mu, Zhengli Rao, Deqin Wen, Guowang Wu, Yong Hu, Xiaorong Shi, Shaofen Huang, Ya Chen, Wenmao Zhang, Min Chen, Xin Luo, Hongqiong Ruan, Liming Liu, Lihui Fan, Xiaoyan Zou, Huizhi He, Xinggao Yang, Linfei Wang, Hong Wang, Wenxiang Li, Jialian Wu, Jinyong Yang, Jianying Wu, Guangliang Liu, Xiaolin Xiong, Gang Yang, Ping Feng, Cheng Luo, Qiwei Zhang, Qihui Yang, Taoshan Ni, Rensong Tian, Li Zhang, Yuanmei Zhang, Hui Bi, Bijue Huang, Yu Yan, Shan Chen, Xinghui Chen, Jiangfei Wang, Guanglun Zou, Junting Liu, Rongjiang Zhao, Li Yuan, Ruyin Yuan, Jianbing Li, Aiping Wang, Daijie Wang, Mei Yang, Kaiyan Si, Peng Wang, Yan Linghu, Tuye Yuan, Zhiyuan Yuan, Weiwei Qian, Daimei Wang, Rongqin Ren, Dong Dai, Qianhua Deng, Yan Xu, Jiali Zhang, Qingyan Bai, Yingwen Zhang, Xingke Wang, Qiang Liu, Liyuan Feng, Fancha Wang, Shiwen Wang, Xiangguang Liu, Runping Liu, Xiaoying Wang, Daizhi Wang, Huihai Cao, Jiang Yuan, Qinhui Cao, Min Luo, Hao Hou, Lu Kong, Zhu Ren, Yu Zhou, Xiaoyan Yuan, Hong Liu, Shixian Gong, Yonghong Zhao, Rusheng Wang, Ji Chen, Fuping Bai, Yang Zhou, Wenbi Chen, Yu Sun, Qunying Xu, Gang Zheng, Yong Wang, Juan Chen, Ling Zhong, Ji Ren, Qiling Zhu, Zheng Yang, Yongsong Chen, Zhonghui Huang, Xiaojiang Zhang, Junyan Yu, Panting Yu, Qun Wang, Qimei Yu, Hao Zhou, Yuanmei Lu, Yiyu Wang, Min Huang, Feng Wu, Jiangyu Yuan, Guangjun Xiong, Jian Zhang, Hong Zhou, Fenfang Liu, Yuqin Chen, Xiaoli Ye, Qingli Luo, Qijie Dai, Yunxian Feng, Tianmei

XSL•FO
RenderX

Huang, Yunfeng Xia, Lei Zhao, Anshui Wang, Fuxiu Yang, Minqin Hu, Tingting Yuan, Song Zhou, Lin Yuan, Mingjian Yuan, Maozhao Wen, Hongming Zhao, Hongbo Zhao, Yuanchao Zhao, Binyuan Yuan, Hua Xiong, Qiyong Yuan, Qin Yuan, Xingqin Zhao, Qifu Luo, Qigui Luo, Tulin Yuan, Xianglin Huang, Chun Luo, Dehong Wang, Rongwei Zhang, Xingji Chen, Zhong'en Wang, Deping Yuan, Qihua Luo, Huahua Li, Yueming Wang, Jian Chen, Wei Teng, Jisheng Wang, Mingli He, Shan Wu, Junjun Xia, Tingxian Mao, Tingmin Luo, Sihong Yi, Ruixue Li, Kang Yang, Chengshang Ao, Xu Zhong, Yang Liu, Langsha Chen, Huan Luo, Qingxia Wang, Chen Chen, Wenqiang Wu, Yumei Wu, Limei Feng, Caixian Lei, Kai Li, Youqun Chen, Lulu Zhao, Han Ren, Juhang Yang, Shanshan Luo, Xiaoyi Zhang, Huiqin Yuan, Li Chen, Guishuang Liu, Jiakuan Ren, Dongqiao Jiang, Yi Yuan, Dengmin Ren, Xiaojiang Mu, Xue Yang, Wuyao He, Wanyi Hu, Piaopiao Huang, Guijiang Li, Jinman Li, Suosuo Li, Yiling Lu, Yuqi Lu, Huan Luo, Qingqing Luo, Wenqin Mu, Daolin Qin, Jinlan Quan, Yingying Tang, Junbao Wang, Lingling Wang, Rui Xu, Chenxi Yang, Dengqin Yang, Liqin Yang, Qingting Zhang, Lei Zhao, Lin Zhao, Jingmei Zhou, Lijuan Zuo, Xiaoning Xia, Yujia Xie, Hangyu Chen, Li Yang, Yu Zhao, Sisi Li, Juan Pan, Ju Yang, Minjia Zhao, Meiyuan Zhou, Guimei Hu, Chunqin Huang, Haiyan Long, Zeyuan Luo, Donghua Pan, Xuelang Ying, Yanyan Zhang, Yuting Zhang, Hong Zhao, Yunli Zhou, Bijin Zhu, Kang Chen, Kaidi Ding, Jiacheng Dong, Tai Fu, Qin Li, Zaijing Liu, Siqi Huang, Lingli Ou, Zhuoting Tan, Yan Tang, Qingzhong Wu, Xinyi Xu, Bingyang Xu, Yulin Zhang, Shuangshuang Zhou, Zena Zhu, Jiancao Zuo, Guangxianfeng Chen, Jiahui Luo, Zhiqi Duan, Shangrong Gu, Binyao Ou, Xingtai Zhao, Yanran He, Xingli Liao, Shunxin Shi, Cheng Wang, Jiayu Wang, Rongmei Wang, Ruirui Zhang, Yuhan Zhou, Serena Luo, Jiwei Jin, Xiaoqi Shi, Ziqianqian Yang, Xiaoxue An, Jiwei Jin, Mingli Huang, Yinghua Mu, Xiaozheng Wang, Songlin Zhang, Xue Zhou, Ruirui Ao, Shaohuan Bao, Yating Jiang, Yijun Liu, Zhirui Nie, Dingyuan Rao, Jiangqin Wan, Xue Wang, Die Wei, Lijin Xie, Jianan Xu, Xuerong Zhang, Die Deng, Jingyu Fu, Shunli Guo, Binhan He, Sitong Liu, Lin Liu, Linqu Qin, Teng Tian, Kehao Wang, Qirui Wu, Laitian Ye, Xiao Zhang, Tiantian Zhou, Sitong Zhou, Jiajia Liu, Huali Yin, Yue Luo, Jianxiu Bai, Yuchan Li, Xinglian Zhang, Shenghui Zhu, Xue Chen, Dan Jiang, Xiaoqin Yang, Changnan Zhao, Fangjing Geng, Jia Li, Jie Wang, Wenfan Hu, Xingling Li, Yumei Lu, Zhuying Long, Junxue Qian, Guoxian Shi, Ruoxi Wang, Guangjian Wu, Guangyan Zhang, Rongrong Zhou, Hongyu Hu, Zhaoyi Ren, Lingxuan Shi, Meng Zeng, Shengsheng Li, Ying Qin, Yi Zhou, Tianyu Liu, Jia He, Jinan Li, Mingyuan Mao, Jiayi Wang, Ronghui Ma, and Bin Li.

Data Availability

The datasets used or analyzed in this study are available from the principal investigator (CW) on reasonable request.

Funding

This study was supported by the Noncommunicable Chronic Diseases–National Science and Technology Major Project (2024ZD0522300), AstraZeneca China (2024-HX-22), the CAMS Innovation Fund for Medical Sciences (2023-I2M-2-001 and 2022-I2M-CoV19-003), the Nonprofit Central Research Institute Fund of the Chinese Academy of Medical Sciences (2022-ZHCH330-01), the State Key Laboratory Special Fund (2060204), the Tencent Sustainable Social Value Inclusive Health Lab (SSVPJ202307060001), the EU Horizon Europe Programme (HORIZON-MSCA-2021-SE-01 and 101086139-PoPMeD-SuSDeV), and the China Medical Board (22-469 to SC).

Authors' Contributions

Conceptualization: CW, TY, SC, Y Liu, KH, TB
Methodology: CW, TY, SC, Y Liu, KH, TB, PG, AB
Data curation: Y Liu, KH, ZZ, SZ, YW, TZ, QL, X Tang, Y Luo, X Lyu, ZC, JZ, LH, LJ, CW, TY, SC, LT, X Tong
Formal analysis: Y Liu, KH, ZZ, SZ, CW
Investigation: Y Liu, KH, ZZ, SZ, YW, TZ, QL, X Tang, Y Luo, X Lyu, ZC, JZ, LH, LJ
Supervision: CW, TY, SC
Validation: CW, TY, SC
Visualization: Y Liu, KH, ZZ, SZ
Writing – original draft: CW, Y Liu, KH
Writing – review & editing: All
All authors contributed to revising the manuscript. All authors have contributed to this protocol manuscript according to the International Committee of Medical Journal Editors' guidelines. CW, Y Liu, and KH contributed equally to this work and are designated as cofirst authors. TB, DJ, TY, SC, and CW are jointly listed as cosenior authors. SC, CW, and TY are cocorresponding authors.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) 2025 checklist.

XSL•FO
RenderX

[DOCX File , 32 KB - resprot_v15i1e85597_app1.docx]

Multimedia Appendix 2
The pay-for-population incentive mechanism.
[DOCX File , 22 KB - resprot_v15i1e85597_app2.docx]

Multimedia Appendix 3
Secondary outcomes and trial timeline.
[DOCX File , 44 KB - resprot_v15i1e85597_app3.docx]

Multimedia Appendix 4
Full size versions of Figures 1–3.
[ZIP File (Zip Archive), 3486 KB - resprot_v15i1e85597_app4.zip]

References

1. Wang C, Chen S, Shan G, Leng Z, Bärnighausen T, Yang W. Strengthening population medicine to promote public health. Chinese Medical Journal 2022 May 20;135(10):1135-1137. [doi: 10.1097/cm9.0000000000002221]
2. Jamison DT, Summers LH, Chang AY, Karlsson O, Mao W, Norheim OF, et al. Global health 2050: the path to halving premature death by mid-century. Lancet 2024 Oct 19;404(10462):1561-1614. [doi: 10.1016/S0140-6736(24)01439-9] [Medline: 39419055]
3. Verguet S, Bolongaita S, Chang AY, Cardoso DS, Stevens GA. The economic value of reducing mortality due to noncommunicable diseases and injuries. Nat Med 2024 Nov;30(11):3335-3344. [doi: 10.1038/s41591-024-03248-4] [Medline: 39333314]
4. Kahnert K, Jörres R, Behr J, Welte T. The diagnosis and treatment of COPD and its comorbidities. Deutsches Ärzteblatt International 2023 Jun 23;120(25):434. [doi: 10.3238/arztebl.m2023.0027]
5. GBD 2019 DiseasesInjuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
6. GBD 2021 Causes of Death Collaborators. Global burden of 288 causes of death and life expectancy decomposition in 204 countries and territories and 811 subnational locations, 1990-2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):2100-2132 [FREE Full text] [doi: 10.1016/S0140-6736(24)00367-2] [Medline: 38582094]
7. de Oca MM, Perez-Padilla R, Celli B, Aaron SD, Wehrmeister FC, Amaral AFS, et al. The global burden of COPD: epidemiology and effect of prevention strategies. The Lancet Respiratory Medicine 2025 Aug;13(8):709-724. [doi: 10.1016/s2213-2600(24)00339-4]
8. Chen S, Kuhn M, Prettner K, Yu F, Yang T, Bärnighausen T, et al. The global economic burden of chronic obstructive pulmonary disease for 204 countries and territories in 2020-50: a health-augmented macroeconomic modelling study. Lancet Glob Health 2023 Aug;11(8):e1183-e1193 [FREE Full text] [doi: 10.1016/S2214-109X(23)00217-6] [Medline: 37474226]
9. Chang AY, Stevens GA, Cardoso DS, Cao B, Jamison DT. The economic value of reducing avoidable mortality. Nat Med 2024 Nov;30(11):3327-3334. [doi: 10.1038/s41591-024-03253-7] [Medline: 39333313]
10. Wang C, Xu J, Yang L, Xu Y, Zhang X, Bai C, et al. Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. The Lancet 2018 Apr;391(10131):1706-1717. [doi: 10.1016/s0140-6736(18)30841-9]
11. Schroedl C, Yount S, Szmuilowicz E, Hutchison P, Rosenberg S, Kalhan R. A qualitative study of unmet healthcare needs in chronic obstructive pulmonary disease. A potential role for specialist palliative care? Ann Am Thorac Soc 2014 Nov;11(9):1433-1438 [FREE Full text] [doi: 10.1513/AnnalsATS.201404-155BC] [Medline: 25302521]
12. Wang C, Qi W, Yang T, Jiao L, Chen Q, Huang K, et al. The care cascade of chronic obstructive pulmonary disease in China: a cross-sectional study of individual-level data at enrolment into the national 'Happy Breathing' Programme. EClinicalMedicine 2024 Aug;74:102597 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102597] [Medline: 39114273]
13. Jordan RE, Adab P, Sitch A, Enocson A, Blissett D, Jowett S, et al. Targeted case finding for chronic obstructive pulmonary disease versus routine practice in primary care (TargetCOPD): a cluster-randomised controlled trial. Lancet Respir Med 2016 Sep;4(9):720-730. [doi: 10.1016/S2213-2600(16)30149-7] [Medline: 27444687]
14. Chen Q, Fan Y, Huang K, Li W, Geldsetzer P, Bärnighausen T, et al. Cost-effectiveness of population-based screening for chronic obstructive pulmonary disease in China: a simulation modeling study. Lancet Reg Health West Pac 2024 May;46:101065 [FREE Full text] [doi: 10.1016/j.lanwpc.2024.101065] [Medline: 38721063]
15. Fabbri LM, Celli BR, Agustí A, Criner GJ, Dransfield MT, Divo M, et al. COPD and multimorbidity: recognising and addressing a syndemic occurrence. Lancet Respir Med 2023 Nov;11(11):1020-1034. [doi: 10.1016/S2213-2600(23)00261-8] [Medline: 37696283]

XSL•FO
RenderX

16. Divo M, Cote C, de Torres JP, Casanova C, Marin JM, Pinto-Plata V, BODE Collaborative Group. Comorbidities and risk of mortality in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2012 Jul 15;186(2):155-161. [doi: 10.1164/rccm.201201-0034OC] [Medline: 22561964]
17. Pumar MI, Gray CR, Walsh JR, Yang IA, Rolls TA, Ward DL. Anxiety and depression-important psychological comorbidities of COPD. J Thorac Dis 2014 Nov;6(11):1615-1631 [FREE Full text] [doi: 10.3978/j.issn.2072-1439.2014.09.28] [Medline: 25478202]
18. Groenewegen A, Zwartkruis VW, Smit LJ, de Boer RA, Rienstra M, Hoes AW, et al. Sex-specific and age-specific incidence of ischaemic heart disease, atrial fibrillation and heart failure in community patients with chronic obstructive pulmonary disease. BMJ Open Respir Res 2022 Dec;9(1):e001307 [FREE Full text] [doi: 10.1136/bmjresp-2022-001307] [Medline: 36585036]
19. Kaszuba E, Odeberg H, Råstam L, Halling A. Heart failure and levels of other comorbidities in patients with chronic obstructive pulmonary disease in a Swedish population: a register-based study. BMC Res Notes 2016 Apr 12;9:215 [FREE Full text] [doi: 10.1186/s13104-016-2008-4] [Medline: 27067412]
20. Svendsen CD, Kuiper KKJ, Ostridge K, Larsen TH, Nielsen R, Hodneland V, et al. Factors associated with coronary heart disease in COPD patients and controls. PLoS One 2022;17(4):e0265682 [FREE Full text] [doi: 10.1371/journal.pone.0265682] [Medline: 35476713]
21. Zhou Y, Chen S, Tian J, Cui J, Li X, Hong W, et al. Development and validation of a chronic obstructive pulmonary disease screening questionnaire in China. Int J Tuberc Lung Dis 2013 Dec;17(12):1645-1651. [doi: 10.5588/ijtld.12.0995] [Medline: 24200283]
22. Liu Z, Li Y, Cui Z, Li L, Nie X, Yu C, et al. Prevalence of tobacco dependence and associated factors in China: findings from nationwide China Health Literacy Survey during 2018-19. Lancet Reg Health West Pac 2022 Jul;24:100464 [FREE Full text] [doi: 10.1016/j.lanwpc.2022.100464] [Medline: 35538934]
23. Wadland WC, Holtrop JS, Weismantel D, Pathak PK, Fadel H, Powell J. Practice-based referrals to a tobacco cessation quit line: assessing the impact of comparative feedback vs general reminders. Ann Fam Med 2007 Mar 01;5(2):135-142 [FREE Full text] [doi: 10.1370/afm.650] [Medline: 17389537]
24. Chen S, Chen W, Li Y, Yu Y, Chen Q, Jiao L, et al. Effectiveness and cost-effectiveness of an integrated digital psychological intervention (EmoEase) in Chinese chronic obstructive pulmonary disease patients: study protocol of a randomized controlled trial. Digit Health 2024;10:20552076241277650 [FREE Full text] [doi: 10.1177/20552076241277650] [Medline: 39381816]
25. World Health Organization. WHO Package of Essential Noncommunicable (PEN) Disease Interventions for Primary Health Care in Low-Resource Settings. Geneva, Switzerland: World Health Organization; Sep 07, 2020.
26. World Health Organization. Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycemia. Geneva, Switzerland: World Health Organization; Apr 21, 2006.
27. Huang X, Yu H, Huang E, Wang S, Chen J, Weng J, et al. Feasibility and performance of the chronic obstructive pulmonary disease population screener and chronic obstructive pulmonary disease screening questionnaire in a Chinese physical examination center. J Thorac Dis 2024 Feb 29;16(2):1378-1387 [FREE Full text] [doi: 10.21037/jtd-23-1967] [Medline: 38505045]
28. Chen S, Wang C, Li B, Shi G, Li H, Zhang J, et al. Risk factors for FEV₁ decline in mild COPD and high-risk populations. COPD 2017 Jan;Volume 12:435-442. [doi: 10.2147/copd.s118106]
29. Lilholt PH, Witt Udsen F, Ehlers L, Hejlesen OK. Telehealthcare for patients suffering from chronic obstructive pulmonary disease: effects on health-related quality of life: results from the Danish 'TeleCare North' cluster-randomised trial. BMJ Open 2017 May 09;7(5):e014587 [FREE Full text] [doi: 10.1136/bmjopen-2016-014587] [Medline: 28490555]
30. Crespi CM. Improved designs for cluster randomized trials. Annu Rev Public Health 2016;37:1-16. [doi: 10.1146/annurev-publhealth-032315-021702] [Medline: 26789386]
31. Chen S, Kuhn M, Prettner K, Bloom DE. Noncommunicable diseases attributable to tobacco use in China: macroeconomic burden and tobacco control policies. Health Aff (Millwood) 2019 Nov 01;38(11):1832-1839. [doi: 10.1377/hlthaff.2019.00291] [Medline: 31682483]

Abbreviations

CBT: cognitive behavioral therapy
COPD: chronic obstructive pulmonary disease
COPD-SQ: COPD Screening Questionnaire
cRCT: cluster randomized controlled trial
DSMB: Data and Safety Monitoring Board
EDC: electronic data capture
FEV1: forced expiratory volume in 1 second
FVC: forced vital capacity
ICC: intraclass correlation coefficient
MDD: minimum detectable difference

XSL•FO
RenderX

NCD: noncommunicable disease
PHQ-9: 9-item Patient Health Questionnaire
POPMIX: Population Medicine Multimorbidity Intervention in Xishui County
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
WEMWBS: Warwick-Edinburgh Mental Well-being Scale

Edited by J Sarvestan; submitted 10.Oct.2025; peer-reviewed by Z Jin; comments to author 17.Nov.2025; accepted 18.Dec.2025; published 18.Feb.2026.

Please cite as:
Wang C, Liu Y, Huang K, Zheng Z, Zhang S, Chen W, Tang X, Cao Z, Tong X, Tang L, Zhao J, He L, Jiao L, Wang Y, Zhao T, Luo Y, Lai Q, Lyu X, Chen Q, Bunker A, Vollmer S, Geldsetzer P, Jamison D, Bärnighausen T, Yang T, Chen S, POPMIX Group
Impact of the Population Medicine Multimorbidity Intervention in Xishui County (POPMIX) on People at High Risk for Chronic Obstructive Pulmonary Disease: Protocol for the POPMIX-COPD Cluster Randomized Controlled Trial
JMIR Res Protoc 2026;15:e85597
URL: https://www.researchprotocols.org/2026/1/e85597
doi:10.2196/85597
PMID:

©Chen Wang, Yuhao Liu, Ke Huang, Zhoutao Zheng, Shiyu Zhang, Wenjin Chen, Xingyao Tang, Zhong Cao, Xunliang Tong, Lei Tang, Jinghan Zhao, Liu He, Lirui Jiao, Yingping Wang, Tianying Zhao, Yingchi Luo, Qiande Lai, Xiangqin Lyu, Qiushi Chen, Aditi Bunker, Sebastian Vollmer, Pascal Geldsetzer, Dean Jamison, Till Bärnighausen, Ting Yang, Simiao Chen, POPMIX Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 18.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effects of Allulose vs Aspartame Consumption on Postprandial Glucagon-Like Peptide-1 Profiles and Metabolic Health: Protocol for a Randomized, Crossover, Double-Blind, Placebo-Controlled Trial

Selina Busch[1], BSc; Paola G Ferrario[1], Dr rer nat; Ann-Kathrin Henk[1], Dr med; Ann Katrin Engelbert[1], Dr troph; Oliver Wittek[1], Dr rer nat; Stephanie Seifert[1], Dr rer nat; Achim Bub[1,2], Dr med; Carina I Mack[3], Dr rer nat; Bettina Hieronimus[1], Dr rer nat

[1]Department of Physiology and Biochemistry of Nutrition, Max Rubner-Institut, Karlsruhe, Germany

[2]Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

[3]Department of Safety and Quality of Fruit and Vegetables, Max Rubner-Institut, Karlsruhe, Germany

Corresponding Author:
Bettina Hieronimus, Dr rer nat
Department of Physiology and Biochemistry of Nutrition
Max Rubner-Institut
Haid-und-Neu-Str. 9
Karlsruhe, 76131
Germany
Phone: 49 721 6625 ext 349
Email: bettina.hieronimus@mri.bund.de

Abstract

Background: Excessive sugar consumption is a public health concern. Allulose, a low-calorie sugar with similar functional properties to sucrose, offers potential metabolic benefits. Animal and limited human studies suggest it may stimulate glucagon-like peptide-1 (GLP-1) secretion, improve glucose regulation, and support weight management. However, evidence to substantiate these effects in humans remains scarce.

Objective: The primary aim of this study, the low-calorie sweetener intervention study allulose (LisA), was to assess differences in the postprandial GLP-1 profile (primary outcome) between an acute intake of allulose and aspartame interventions in healthy adults. Secondary goals included exploratively assessing potential subacute adaptation effects over a 4-week consumption period and evaluating a comprehensive set of parameters as hypothesis-generating findings for future large-scale research.

Methods: We conducted a randomized, double-blind, placebo-controlled, crossover trial in healthy adults. Participants daily consumed either 3 allulose-sweetened or aspartame-sweetened beverages for 4 weeks in crossover, with a washout in between. Standardized inpatient procedures were conducted at the study baseline and at the beginning and end of each intervention phase. The primary outcome is the postprandial profile of GLP-1. Secondary outcomes include further parameters of gut hormone secretion, insulin sensitivity (Matsuda Index), body composition (body impedance analysis), subjective satiety (visual analog scales), and gastrointestinal tolerance. We also assess multiomic endpoints, including sugaromics and gut microbiome composition. The primary outcome will be analyzed using the incremental area under the curve with a 2-tailed paired *t* test. All further outcomes (including peak and total area under the curve for GLP-1) will be assessed using linear mixed models.

Results: A total of 10 participants (4 female and 6 male; mean age 31.2, SD 6.8 years; BMI 25.1, SD 2.6 kg/m^2) completed all study procedures. The sample collection phase was successfully concluded in November 2023. Data processing and statistical analysis for the primary outcome are expected to be completed by June 2026.

Conclusions: The comprehensive study protocol, integrating a rigorous crossover design with multiomic analysis, is poised to provide confirmatory evidence for the acute GLP-1 effects of allulose and generate valuable mechanistic hypotheses regarding its subacute metabolic and gut health effects. The findings will contribute to the evidence base required for evaluating allulose's potential role in public health sugar reduction strategies.

Trial Registration: German Clinical Trials Register DRKS00028521; https://drks.de/search/en/trial/DRKS00028521

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/81857

(JMIR Res Protoc 2026;15:e81857) doi:10.2196/81857

KEYWORDS

allulose; aspartame; gut hormones; insulin sensitivity; randomized controlled study; satiety

Introduction

The dietary intake of free sugars consistently exceeds recommended maximum levels across all age groups in Germany [1-3]. This is a growing public health concern, as excessive sugar consumption—especially from sugar-sweetened beverages—is linked to dental caries, overweight, obesity, and associated metabolic diseases [4,5]. To mitigate these risks, the use of low-calorie sweeteners has increased in recent years, with further growth anticipated [6].

One promising low-calorie sweetener is allulose (D-psicose), a rare naturally occurring sugar found in small quantities in dried fruits and high-sugar processed food [7]. Allulose offers significant calorie reduction, providing only 0.2-0.4 kcal/g, or 5%-10% of the energy content of caloric sugars [8], while maintaining similar functional properties, such as browning, texture, and bulk addition. This makes allulose a promising candidate for sugar reduction strategies. Like its epimer fructose, allulose is absorbed primarily via glucose transporter 5 transporters [9]; however, it is believed to be excreted unmetabolized in the urine [10]. Studies suggest it is well-tolerated at moderate doses, with a gastrointestinal tolerance similar to fructose [11].

Beyond its potential to reduce caloric value in sweetened products, animal studies have highlighted further metabolic benefits of allulose. For example, allulose consumption has been shown to stimulate the release of glucagon-like peptide-1 (GLP-1), a gut hormone that plays a critical role in regulating satiety and postprandial glucose metabolism [12,13], in rodents [14,15]. Increased GLP-1 levels have been associated with improved glucose regulation and reduced food intake in rodent models. In humans, fructose induces GLP-1 release, suggesting that allulose may elicit a similar response [16-18]. However, human data are limited to 1 isolated study showing modest GLP-1 increases after administration of allulose by feeding tube without significant effects on satiety or food intake [19], underscoring the need for further investigation. Furthermore, the metabolic and endocrine effects of sweeteners and fermentable carbohydrates can change over time. Studies investigating dietary fibers and specific nutrients suggest that alterations in the GLP-1 response may be subject to physiological accommodation or adaptation following continuous exposure, driven by changes in the gut microbiota or L-cell sensitivity [20]. Crucially, the currently available human data on allulose focus exclusively on acute, single-dose effects [19]. Previous human trials over several weeks have observed no significant change in GLP-1 release following repeated exposure to sweeteners like aspartame or sucralose [21,22], while preclinical animal studies suggest that chronic sweetener intake can modulate GLP-1 dynamics via changes in enteroendocrine signaling or L-cell function [23]. There is a significant gap in knowledge regarding whether the immediate postprandial GLP-1 response is sustained, potentiated, or diminished after a period of regular consumption. Therefore, a subacute intervention period, such as 4 weeks, is essential to adequately capture these potential longitudinal adaptation effects on GLP-1 secretion and other downstream metabolic markers.

Beyond its effects on satiety, allulose may influence glucose metabolism independently of GLP-1. Acute intervention trials in humans have suggested that allulose coingested with glucose or mixed meals can modestly reduce postprandial blood glucose responses, although the effects on insulin levels are inconsistent [24-26]. Proposed mechanisms include inhibition of intestinal α-glucosidase activity and enhanced hepatic glucose usage, potentially mediated by glucokinase activation [27]. Animal studies further support these findings, showing increased hepatic glycogen storage and improved glucose tolerance in response to allulose supplementation [27,28]. However, these effects are not well-characterized in humans, and the potential for allulose to aid in blood glucose regulation warrants further exploration. Allulose could also play a role in body weight regulation. In animal studies, allulose supplementation has been shown to reduce diet-induced weight gain, possibly through enhanced satiety [27-33] and reduced energy intake [32]. Preliminary human trials have also reported weight loss and reductions in body fat with allulose consumption, even in the absence of significant changes in total energy intake or physical activity [34,35]. One proposed mechanism contributing to weight loss is a direct effect of allulose on substrate oxidation within energy metabolism. Studies in both rats [32] and humans [36] have demonstrated that allulose consumption increases postprandial fat oxidation, as measured by indirect calorimetry. These findings suggest that allulose may influence body weight beyond its caloric reduction, potentially via mechanisms related to gut hormone modulation or altered glucose metabolism. However, these effects require confirmation in well-controlled, long-term human studies.

In addition to its metabolic effects, allulose may influence the composition and activity of the gut microbiome. The intestinal absorption capacity for allulose is limited [10]. Following oral intake, a portion of the ingested allulose therefore reaches the lower sections of the intestine, where it becomes available for microbial fermentation. Animal studies have shown that the consumption of allulose leads to changes in the composition of the gut microbiota and short-chain fatty acids [37,38]. The digestion of allulose by bacteria requires specific enzymes encoded by a dedicated gene cluster, which is not ubiquitously present across bacterial species [10,39]. The presence of allulose could therefore offer certain bacteria a selective survival advantage and induce a shift in the microbiota composition [40]. However, whether and to what extent allulose induces changes in the complex human gut microbiome remains unknown.

XSL•FO
RenderX

Despite its promise as a low-calorie sugar alternative, the effects of allulose on human metabolism remain incompletely understood [41]. This study aims to address critical gaps by investigating the impact of allulose consumption on GLP-1 secretion, satiety, glucose regulation, and body weight in humans, as well as potential effects on the microbiome. The findings will contribute to the evidence needed to guide the responsible integration of allulose into the food supply.

Methods

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and the Good Clinical Practice guidelines. The trial was prospectively registered in the German Clinical Trials Register under the identifier DRKS00028521 (Universal Trial Number U1111-1291-0885) on May 3, 2023. Ethical approval was obtained in April 2022 from the Ethics Committee of the State Medical Chamber of Baden-Württemberg (approval number F-2022-029). All potential participants were thoroughly informed, both verbally and in writing, about the aims, procedures, potential risks, and benefits of the study, and all participants provided written informed consent prior to their inclusion. To safeguard participant data, all data were pseudonymized immediately upon collection. All identifying personal information is stored securely, separate from the study data, and access to the identity key is restricted to authorized personnel in compliance with applicable data protection regulations. Finally, participants received financial compensation (US $850) for their time commitment.

Outcomes

The primary endpoint of this randomized crossover trial is the acute postprandial GLP-1 profile comparison between the allulose and aspartame interventions. For the primary analysis, the GLP-1 profile will be quantified using the incremental area under the curve (iAUC), as this metric best reflects the net stimulation effect relevant to appetite modulation and glucose regulation. Peak concentration and total area under the curve (AUC) will also be assessed. No minimal important change for within-participant GLP-1 has been established. The clinical significance of the observed changes will be interpreted in the context of available literature. Secondary outcomes comprise a range of parameters assessed following 4 weeks of daily exposure. A central secondary goal is the investigation of potential subacute adaptation effects on the GLP-1 profile by comparing the acute response with the response at the end of the intervention phase. Additionally, comprehensive parameters such as glucose and insulin homeostasis, insulin sensitivity, satiety (visual analog scales [VAS]), body weight and composition, and gastrointestinal tolerance will be evaluated (secondary outcomes). The study also includes exploratory multiomics analyses, namely sugaromics and gut microbiome composition analysis, to generate mechanistic hypotheses for future, larger-scale studies.

Study Design and Setting

The study was planned as a randomized, crossover human trial. The study included 2 phases, each consisting of a 7-day baseline period and a 4-week intervention period, separated by a 3- to 4-week washout phase (Figure 1). During each intervention period, participants consumed 3 beverages daily consisting of water sweetened either with allulose or aspartame (control). The sequence of interventions was randomized (allulose–aspartame or aspartame–allulose). Before and during inpatient procedure days (baseline and intervention), participants received a standardized low-sugar dinner. Beverages were consumed with inpatient meals on procedure days, while they were consumed in a private setting on days without procedures. This study protocol was developed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines. The completed SPIRIT checklist is provided as Multimedia Appendix 1 [42].

Figure 1. Study schematic: participants were randomly assigned to an intervention sequence. Procedures and study protocol were the same in both intervention phases. SDY: study day.

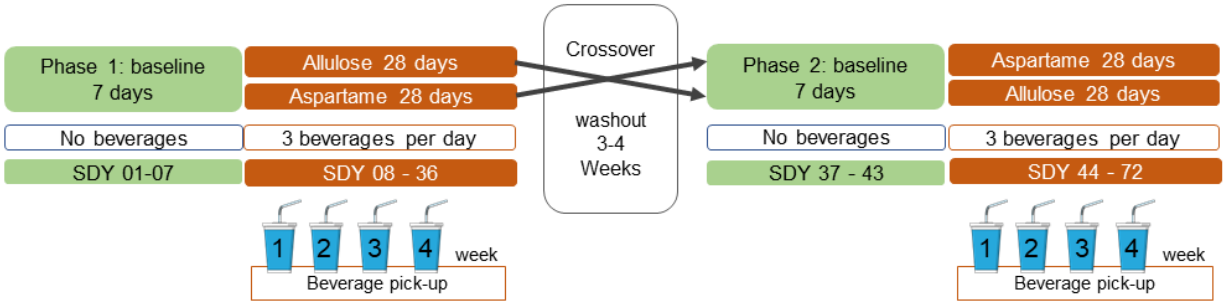

Participants

We aimed to include 12 participants (male or female; aged 18-50 years; BMI 18.5-30 kg/m^2) with self-reported stable body weight during the prior 6 months. Participants were required to be in good metabolic health, as assessed by the study physician, without nutritional sensitivities (food allergies, fructose malabsorption, celiac disease, irritable bowel disease, etc) and should not show signs of restricted eating (restrain score <13; Fragebogen zum Essverhalten [43], which is the German version of the Three-Factor Eating Questionnaire [44]). Further detailed eligibility criteria are listed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Women and men aged 18-50 years
- BMI 18.5-30 kg/m^2
- No desired weight change
- Good health
- Written informed consent to participate voluntarily in the study
- No diseases affecting nutrient absorption, digestive function, metabolism, or excretion, in particular malabsorption or intolerance of fructose
- Nonsmokers
- Parameters of liver, kidney, and thyroid function and carbohydrate and lipid metabolism without pathological values (complete blood count, clinical chemical parameters, and preliminary examination)
- No indication of restrictive eating behavior

Exclusion criteria

- Pregnant or breastfeeding
- Acute or regular use of medications that affect lipid metabolism, blood glucose regulation, or the gastrointestinal tract
- Use of antibiotics in the past 6 months
- Allergy or intolerance to any food or ingredients in the test meal

Participants were recruited through various channels, including flyers, local and regional newspapers, Instagram, and direct outreach through our internal participant database. Initial screening was conducted through a 20-minute telephone interview, during which participants received an overview of the study and their general eligibility was assessed using a structured interview form.

Individuals meeting the formal inclusion criteria were provided with a copy of the consent form by email, detailing the study objectives and procedures (Multimedia Appendix 2). Interested participants were invited to a screening appointment at the Max Rubner-Institut, where the study principal investigator and study physician explained the study, reviewed the consent form, addressed any questions, and obtained written informed consent from those who chose to participate.

During the screening visit, a fasting blood sample was collected to analyze clinical chemical parameters, blood count, and lipids. Participants also completed the Fragebogen zum Essverhalten [43]. Candidates who met all inclusion criteria and were deemed responsible and likely to adhere to the dietary protocol were invited to participate in the study.

Intervention

To be labelled “sugar reduced,” a product marketed in the European Union must contain 30% less sugar and energy than comparable traditionally sweetened products [45]. In this study, the amount of allulose was calculated to reflect the substitution of one-third of the sugar in “sugar-reduced” products. Free sugar intake in German adults aged 19-50 years averages 15.3% of daily energy requirements, equivalent to 81.8 g/day [46]. Based on individual daily energy requirements (calculated using the Mifflin equation, with an adjustment of 1.5 for activity), participants received allulose equivalent to 5.1% of their daily energy requirements (15.3% × 1/3), representing a realistic average exposure if one-third of sugar were replaced by allulose upon market introduction. Beverages were prepared on-site by staff who were not otherwise involved in the study.

Aspartame was used in the control group for blinding, as it does not affect the main outcomes assessed in this study [22,47,48]. Beverages were matched for sweetness, and beverage volumes were equal for both interventions. Aspartame concentration was determined by sensory evaluation of sweetness to match the sweetness of the allulose beverage, ensuring effective blinding. Testing was performed by a small group of panelists with prior experience in sensory testing. Allulose concentration was 50 g/L, and aspartame concentration was 113.75 mg/L. Beverage volumes per bottle were adjusted for each participant to match the individual allulose quantity, bottled and frozen at −20 °C until they were handed to the participants.

Participants were asked to consume 3 beverages daily with main meals. Drinks could be flavored to taste (eg, with lemon slices). Participants were instructed to label empty bottles with the consumption date and accompanying meal (breakfast, lunch, or dinner) and return them at the weekly in-person appointments.

The 4-week duration for the intervention was selected to balance participant burden with the physiological need to observe cumulative metabolic and microbial adaptation effects. Furthermore, the washout and intervention periods were chosen to facilitate the testing of female participants during the same phase of their menstrual cycle, thereby reducing a major source of biological variability in the outcomes of interest.

Randomization and Blinding Procedure

Randomization was performed before study initiation by a staff member who was not involved in participant enrollment or study implementation, ensuring allocation concealment. The 2 intervention sequences—aspartame–allulose and

XSL•FO
RenderX

allulose–aspartame—were each entered into stacked columns of 6 cells in a Microsoft Excel spreadsheet. The RAND() function was applied in an adjacent column to generate random numbers, and the RANK() function was used to assign ranks to these numbers. Based on the resulting ranks, the sequences were allocated to participant IDs.

The staff member performing the randomization was also responsible for preparing the drinks and therefore was the only person knowing the sequence allocation during the active phase of the study. Participants were not informed about the sequence of intervention they were receiving. Additionally, both the study staff interacting with the participants and the technical staff analyzing the samples were blinded to the intervention assignments during and after the study.

Standardized Meals

During the outpatient period, participants were instructed to follow their usual ad libitum diets. On the evenings before study procedures, participants were provided with standardized dinners and were restricted to eating only these meals. The standardized ad libitum meal resembled a typical German cold dinner, consisting of mixed wheat and whole-meal bread, butter, a packet of sliced cheese, cream cheese, a tomato, a cucumber, a carrot, and an apple. Participants were asked to finish eating before 9 PM and to drink only water afterward.

The standardized breakfast consisted of energy-balanced semolina pudding (semolina, milk, cream, and vanilla aroma) providing 25% of the participant's daily energy requirement. Lunch was a vegetable gratin (pasta, mixed vegetables, cheese, and herb sauce with milk) which was provided ad libitum. These meals contained 50% energy from carbohydrates, 15% energy from protein, and 35% energy from fat. For the inpatient meals, daily energy requirements were calculated by the Mifflin equation [49] with an adjustment of 1.3 for activity, as participants had restricted mobility on the respective days.

Participant Retention and Follow-Up

Several strategies were implemented to promote participant retention and ensure complete follow-up. Participants received regular personal contact during weekly appointments to return empty bottles and collect the next week's beverages. This contact provided an opportunity to address any concerns, reinforce adherence, and document any missed doses or adverse effects.

Participants were also reminded of upcoming visits and procedures by email. The study team made every effort to accommodate participants' schedules and offered flexibility for beverage and meal pickup times.

Experimental Procedures

Anthropometric Measurements and Assessment of Body Composition

Participants were weighed on inpatient days in the morning. Body weight was measured to the nearest 0.05 kg in underwear and was determined on the first and last day of each baseline phase (study day [SDY] 1 and SDY 37; SDY 7 and SDY 44) and at the end of each intervention phase (SDY 35/SDY 71). Height was measured to the nearest 0.1 cm using the Seca 285 measuring station (Seca). Waist and hip circumferences were measured with an ergonomic body circumference measuring tape from Seca. Body composition was measured using bioelectrical impedance analysis (Medical Body Composition Analyzer Seca 515; Seca). Bioelectrical impedance analysis measurement was performed at baseline (SDY 7 and SDY 44) and at the end of each intervention phase (SDY 35 and SDY 71).

Satiety Assessment

Satiety was tested and analyzed according to best practice suggestions by experts in the field [50]. Participants were provided with water (baseline) or aspartame- or allulose-sweetened beverages (intervention) as a preload. Fifteen minutes after the preload, participants were provided with a standardized breakfast that had to be eaten within 15 minutes.

Gut hormones will be analyzed in all plasma samples by enzyme-linked immunosorbent assay.

In addition, participants filled out a total of 9 VAS to assess subjective satiety. VAS, 100 mm in length with words anchored at each end, expressing the most positive and the most negative rating, were used to assess hunger, satiation, fullness, and prospective food consumption. The questionnaires were made as small booklets showing 1 question at a time. The first VAS was filled out before the preload and then every 30 minutes until 4 hours after breakfast. Satiety was assessed at 3 timepoints during each study arm: on the first day of the baseline and on the first and last day of receiving intervention beverages in each intervention phase.

Participants were given a standardized lunch on SDY 1 and SDY 37 and SDY 35 and SDY 71 to be eaten in their individual rooms. Participants were served a meal exceeding their calculated lunch energy needs (1.5×), presented in a casserole dish and eaten from dinner plates. They were informed that the portion was too large to be finished. They were asked to fully focus on the meal and not to use any digital device while eating or distract themselves otherwise and continue to eat until they are comfortably satiated. The eating time and amount of food consumed were recorded.

3-Hour Oral Glucose Tolerance Test

Oral glucose tolerance tests (OGTTs) were conducted at baseline (SDY 7 and SDY 43) and on the first day after the 4-week intervention period (SDY 36 and SDY 72). Participants fasted overnight and drank a glucose solution (75 g glucose in 250 mL water) after an initial blood draw. Blood was drawn at 0.5, 1, 1.5, 2, and 3 hours after drinking the glucose solution. Glucose and insulin concentrations will be assessed in all plasma samples. OGTT glucose and insulin AUC and the Matsuda Index for whole-body insulin sensitivity [51] will be calculated.

Postprandial Course of Respiration Quotient

Respiratory gas exchange was determined on the first day of the baseline phase (SDY 1 and SDY 37) and on the penultimate day of the intervention phase (SDY 35 and SDY 71) using a calorimetry system (Cosmed Quark RMR; software Omnia 1.6.10). Measurements were conducted before lunch and then

XSL•FO
RenderX

once per hour for 3 hours. Measurements were conducted in a quiet room with an initial room temperature of 22 °C. The room was ventilated after each session. Participants wore earplugs to minimize sound exposure. Total energy expenditure, carbohydrate energy expenditure, fat energy expenditure (kJ/min/kg body weight), and the respiratory quotient were calculated.

Dietary Intake Assessment (Food Frequency Questionnaires)

Participants filled out food frequency questionnaires (FFQs) at the end of the intervention periods (SDY 35 and SDY 71) after lunch. The FFQ contains 53 food groups and covers the previous 4 weeks and was validated at the Max Rubner-Institut [52]. Given that the FFQ was designed in 2010 and plant-based alternative products have a higher market share today, we added 12 questions in total that asked for the type and amount of plant-based alternatives for milk or meat-based products.

Participants were instructed in advance on the use of the FFQ by trained staff. Mean dietary intake (g) and total energy consumption (kcal) were calculated from the FFQ data. As FFQs are prone to high interindividual differences [53], percent differences between the 2 phases were compared for each participant.

Continuous Glucose Monitoring

A continuous glucose monitoring sensor (FreeStyle Libre Pro iQ Continuous Glucose Monitoring System, Abbott Diabetes Care) was placed on the participants' upper nondominant arm during each baseline week and the last week of each intervention. Participants carried the sensor during 5 outpatient days when food consumption was not controlled and 2 days when they consumed standardized meals. This allowed comparison of glucose responses (24-hour AUC and postprandial responses) under free-living and under controlled conditions, both at baseline and during the intervention. Data were recorded after the OGTT on SDY 8 and SDY 44 and SDY 36 and SDY 72 (FreeStyle Libre Pro iQ Recorder, Abbott Diabetes Care). The participants were blinded to their blood glucose readings. Data were extracted using LibreView Software (Abbott).

Sample Collection

Table 1 provides an overview of the sample collection timepoints and analyses.

Table 1. Sample type and uses.

Sample type and study days	Collection details	Analyses and use
Blood samples		
1, 8, 35, 37, 44, and 71	Satiety assessment: intravenous catheter, samples collected at 0, 0.5, 1, 1.5, 2, 3, 4, and 5 hours	(Gut) hormone analyses (active GLP-1[a], GIP[b], CCK[c], PYY[d], and insulin), GLUT5[e] analysis
7, 36, 43, and 72	OGTT[f]: intravenous catheter, samples collected at 0, 0.5, 1, 1.5, 2, and 3 hours	Glucose, lactate, insulin, metabolomics, and sugaromics
Urine samples		
8, 35, 44, and 71	Spot urine at 0 hours and fractions at 0–2 hours and 2–4 hours	Sugaromics, metabolomics, and microbiome analyses (16S)
7, 36, 43, and 72	Spot urine at 0 hour and fraction at 0–3 hours	Sugaromics and metabolomics
Saliva samples		
7, 14, 36, 43, 50, and 72	Stimulation using cooled dental wax	Sugaromics and microbiome analyses (16S)
Stool samples		
1–7	Fresh sample processed within 4 hours	In vitro fermentation
5–7, 41–43, 8–9, 32–36, and 70–72	Frozen sample	Microbiome analyses (16S) and short-chain fatty acids

[a]GLP-1: glucagon-like peptide-1.

[b]GIP: glucose-dependent insulinotropic polypeptide.

[c]CCK: cholecystokinin.

[d]PYY: peptide tyrosine-tyrosine.

[e]GLUT5: glucose transporter 5.

[f]OGTT: oral glucose tolerance test.

Blood Samples

Satiety Assessment

Satiety assessment was performed on the first day of baseline and on the first and last day of each intervention (SDY 1, SDY 8, SDY 35, SDY 37, SDY 44, and SDY 71). Following insertion of an intravenous catheter, blood samples were collected before the preload (baseline, 0 hour) and then at 0.5, 1, 1.5, 2, 3, and 4 hours. During baseline and on the last day of the interventions (SDY 1, SDY 35, SDY 36, and SDY 71), one additional blood sample was collected after lunch at 5 hours. All blood samples for hormone analysis were collected into chilled K3-ethylenediaminetetraacetic acid blood collection tubes

XSL•FO
RenderX

(Sarstedt AG & Co KG) containing 50 µM dipeptidyl peptidase-4 inhibitor (Sigma Aldrich) and 500 KIU/mL aprotinin (Merck KGaA), gently mixed, and stored on ice until centrifugation at 2500 × g for 10 minutes at 20 °C. Serum samples for sugaromics analysis were collected in serum separation tubes (Sarstedt AG & Co KG) at 0, 0 0.5, 1, 1.5, 2, 3, and 4 hours, allowed to clot for 30 minutes, and centrifuged accordingly. All samples were aliquoted and frozen at −70 °C. At 3 hours, additional lithium-heparin plasma samples were collected (Sarstedt AG & Co KG) for glucose transporter 5 analysis. Whole blood was centrifuged at 2000 × g for 10 minutes at 4 °C. Plasma was collected, aliquoted, and stored at −70 °C, and blood cells were processed immediately.

Oral Glucose Tolerance Test

OGTTs were conducted on the last day of baseline and 1 day after the intervention (SDY 7, SDY 36, SDY 43, and SDY 72). Following insertion of an intravenous catheter, blood samples were collected at 0, 0.5, 1, 1.5, 2, and 3 hours after glucose ingestion. For glucose (and lactate) analysis, blood samples were collected into GlucoEXACT tubes (Sarstedt AG & Co KG) and centrifuged at 2500 × g for 10 minutes at 20 °C to obtain plasma. Additional serum samples were collected in serum separation tubes (Sarstedt AG & Co KG) in the fasting state and at 0.5 hour and 1.5 hours for metabolome and sugaromics analysis. All samples were aliquoted and frozen at −70 °C.

Urine

During satiety assessment (SDY 8, SDY 35, SDY 44, and SDY 71), fasting spot urine samples were collected before ingestion of the allulose drink and as 2-hour fractions after ingestion (0-2 hours and 2-4 hours). Sugaromics will be assessed, and additional metabolome analyses will be conducted if sugaromics results indicate relevant changes.

During OGTT (SDY 7, SDY 36, SDY 43, and SDY 72), fasting spot urine samples were collected before and glucose ingestion, and a 3-hour fraction was collected after glucose ingestion for metabolome and sugaromics analysis. Within 1 hour of collection or end of the collection period, a portion of each sample was centrifuged at 4 °C for 15 minutes at 2500 × g, and 1 mL of supernatant was frozen at −70 °C for subsequent sugar profiling.

For microbiome analysis (16S gene sequencing), participants collected midstream urine after cleaning the area around the ureter opening with a wet cloth on SDY 7, SDY 37, SDY 43, and SDY 71. Samples were put on ice immediately. The remaining urine was incubated at 37 °C for 10 minutes and subsequently centrifuged for 15 minutes at 1500 × g as 2 × 25 or 50 mL depending on the concentration of urine (assessed subjectively according to its color intensity). Supernatants were removed, and the pellets were frozen at −70 °C.

Saliva

On SDY 7, SDY 14, SDY 36, SDY 43, SDY 50, and SDY 72, stimulated saliva samples were collected for microbiome and sugaromics analysis. On these SDYs, participants were asked to neither brush their teeth nor use mouthwash nor apply lip balm or similar products in the morning before saliva collection. Before sample collection, participants rinsed their mouth with demineralized water and waited for 10 minutes. During this time participants were not allowed to drink. For collection of stimulated saliva, participants received cooled dental wax (dental oral care orthodontic wax, neutral taste—ARGOMAX) for chewing. Participants chewed the dental wax to stimulate saliva production and collected saliva in tubes cooled on ice for 2 minutes. The collected saliva was weighed to calculate the salivary flow rate. These first samples are used for sugaromics analysis. Afterward, participants received another cooled dental wax and collected 1 mL of stimulated saliva for microbiome analysis while chewing.

Stool

Sample Collection for In Vitro Fermentation

In vitro fermentation experiments can be used to assess whether bacteria in the samples are capable of metabolizing allulose. Participants were asked to provide 1 fresh stool sample for in vitro fermentation experiments during the course of the first baseline phase (SDY 1-SDY 7). Participants received a prepacked stool collection kit containing a flushable stool sample collection sheet, gloves, a spatula, two 50 mL plastic tubes with perforated caps, an Oxoid AnaeroGen Compact Sachet (Fisher Scientific) to provide an anaerobic atmosphere, and an airtight plastic box. In addition, participants received precise oral instruction and a written manual for the sampling procedure. The stool sample was processed within 4 hours of sampling.

Sample Collection for Microbiome Analysis

Fecal microbiota composition will be assessed using 16S ribosomal RNA gene amplicon sequencing, and short-chain fatty acid concentrations quantified by targeted metabolomics. Study participants were asked to provide a total of 10 stool samples for microbiota analyses during the course of the study. Samples were collected during each baseline phase (2 samples within 72 hours each between SDY 5-SDY 7 and SDY 41-SDY 43), during the first days of the intervention (1 sample each between SDY 8-SDY 9 and SDY 44-SDY 45), and at the end of each intervention (2 samples within 72 hours each between SDY 32-SDY 36 and SDY 70-SDY 72). Participants received a prepacked stool collection kit that contained a flushable stool sample collection sheet, gloves, a spatula, 3 small plastic bags, a small plastic container, and 2 cooling packs. In addition, participants received precise oral instruction and a written manual for the sampling procedure. Participants were asked to collect 3 independent, approximately nut-sized aliquots from the same stool, which were placed in a small plastic bag. The small plastic bag was placed in a plastic container and cooled with prefrozen (−20 °C) cooling packs immediately after collection, stored at −20 °C until delivery to the study center, where storage continued at −70 °C.

Compliance

Participants provided urine samples at weekly visits for beverage pick-up to assess allulose concentrations. As aspartame is efficiently hydrolyzed and absorbed, it cannot be detected in urine samples, therefore lacking the possibility to measure compliance in this intervention directly. Participants were asked

XSL•FO
RenderX

to return the empty beverage containers with the time and date of consumption recorded.

Adverse Events and Safety Monitoring

The participants completed a questionnaire on gastrointestinal symptoms once a week during the intervention phases. The following potential symptoms were recorded: abdominal pain, heartburn, belching, tummy rumbling, bloating, nausea, increased appetite or hunger, decreased appetite or hunger, feeling of fullness, and defecation (diarrhea and constipation). For each of these symptoms, participants were asked whether and to what extent they had experienced it (5-point Likert-scale: none, mild, moderate, intense, and very intense).

In addition to these symptoms, a free text box was provided for any other symptoms to document possible adverse events (AEs). AEs were defined as any health-related occurrences during the intervention phases, regardless of causal relationship to the test substances. All AEs were documented and evaluated by the study physician according to standard criteria (severity, duration, relation to intervention, and outcome). Given the nature of the intervention and the known safety profiles of allulose and aspartame, no serious AEs in connection with their consumption were expected. In case of recurrent or severe AEs, dose adjustment or study withdrawal was predefined.

A blood sample was taken before and after each intervention phase to determine safety parameters. The transaminases gamma-glutamyl transferase and glutamic oxaloacetic transaminase, bilirubin (total), creatinine, uric acid, glucose, hemoglobin A_{1c} (HbA_{1c}) and thyrotropin; the electrolytes sodium, calcium, and potassium (all in serum); and a blood count were measured.

Sample Size Calculation

Sample size was calculated using G*Power 3.1.9.4 with postprandial GLP-1 profiles as the primary outcome. Based on a paired-design analysis, our working hypothesis is that allulose produces a higher postprandial GLP-1 response compared with aspartame. With a significance level of 5% and a power of 90%, a total of 9 participants need to be analyzed. Anticipating a dropout rate of 20%, a total of 12 participants needed to be initially included. The calculation was based on the following assumptions: (1) a comparable effect as seen in Table 1 in Teysseire et al [19] of 1.27 pmol/L, given as Cohen delta after an invention with allulose or water (mean paired difference divided by the SD of the paired differences); (2) a dropout rate of 20%, based on prior studies at the Max Rubner-Institut; and (3) sample size calculation via a 2-tailed paired *t* test due to the crossover design.

Statistical Analysis

Baseline characteristics of the study population will be summarized using descriptive statistics. Data will be reported as mean (SD) for normally distributed variables and as median (IQR) for ordinal variables.

The primary analysis focuses on the acute difference in the postprandial GLP-1 profile between the allulose and aspartame interventions. For this, the primary outcome, the postprandial GLP-1 profile (as iAUC), will be analyzed using within-subject comparisons inherent to the crossover design. For each participant, paired treatment differences between allulose and aspartame will be calculated from the Δ between the acute postprandial GLP-1 and the corresponding period baseline (Δ[end of intervention − baseline]), adjusted for the corresponding period baseline. The primary treatment comparison will be performed using a paired *t* test. In addition, the magnitude of the within-subject treatment effect will be quantified using Cohen delta. Effect estimates will be presented with 95% CIs. Peak concentration and total AUC for GLP-1, together with the secondary outcomes, including other gut hormones and OGTT parameters, will be treated as exploratory. Postprandial profile data (GLP-1, gut hormones, and OGTT) will be analyzed using linear mixed models. The linear mixed models will include fixed effects for intervention and time (acute vs subacute), with the necessary intervention × time interaction term to assess subacute adaptation effects over the 4-week period. Participant ID will serve as the random intercept, with crossover-sequence as an adjusting variable. Satiety will be assessed as differences from fasting values and analyzed using a comparable mixed-model approach. Dietary intake data from FFQs (eg, kcal or g) may be included as covariates, where appropriate. Further outcomes will be analyzed as absolute or relative changes from baseline (Δ[end of intervention – baseline] or %Δ), using ANCOVA with baseline values and BMI as covariates. All analyses will be conducted using R (R Foundation for Statistical Computing).

No formal correction for multiplicity is planned; the primary outcome will be interpreted as confirmatory, whereas all other outcomes are secondary and will be treated as exploratory. These exploratory analyses will be fully described in subsequent publications focusing on the respective outcomes.

Data Exclusion

All randomized and compliant participants will be included in the primary and secondary analyses. As there were no major protocol deviations, all participants are considered adherent and will be analyzed accordingly. For continuous outcomes with values below the limit of quantification (LOQ), imputation will be performed using LOQ or LOQ/2. Missing data will otherwise not be imputed, and questionnaire-based outcomes with substantial missingness may be excluded from specific analyses, as needed. Data cleaning procedures will be transparently documented, and any further exclusions will be based on predefined criteria and reported in the respective publications.

Data Management

All study data will be stored on secure institutional servers maintained at the study site. Access is restricted to study investigators and data managers through individual logins with role-specific permissions. Data are pseudonymized using unique participant codes, and the key linking these codes to personal identifiers is stored securely and separately, accessible only to designated personnel.

Personal information about potential and enrolled participants is collected solely for eligibility assessment and study conduct. Data on screened individuals who did not participate were deleted after documentation of necessary variables for reporting

XSL•FO
RenderX

purposes (eg, reasons for exclusion). Throughout the study, all personal information is handled in compliance with applicable data protection regulations to ensure confidentiality before, during, and after trial completion.

For continuous glucose monitoring sensor data, automated digital imports are used to reduce transcription errors. Paper-based records (FFQs, gastrointestinal symptom questionnaires, and VAS) are manually entered using predefined coding schemes. All manually entered data are independently double-checked by a second team member to ensure accuracy and completeness.

Results

Trial Implementation

This study was a randomized, crossover human trial conducted at the Study Centre of Human Nutrition, Department of Physiology and Biochemistry of Nutrition, Max Rubner-Institut, Karlsruhe, Germany, from July 4, 2023, to November 16, 2023.

Recruitment and Retention

An initial telephone interview was conducted with 67 interested individuals, of whom 35 were invited to an in-person screening. We excluded 21 persons, of whom 10 did not meet the inclusion criteria, 4 had difficulties with blood draws (either own discomfort or vein assessment by the study physician), 3 had scheduling issues with our proposed study dates, and 4 had other reasons (did not show up: n=3 and menstrual cycle issues: n=1). Of the 14 invited individuals, 2 dropped out before any study procedures, 1 was lost to follow-up, and 1 person had conflicts with the study dates. Of the 12 participants who started the study, 2 withdrew their participation. One participant discontinued after the first SDY due to discomfort with the intravenous catheters, and 1 participant discontinued after the first SDY of the second study phase due to the time commitment required for study continuation. The unequal loss of participants resulted in 4 participants completing sequence AB (aspartame–allulose) and 6 participants completing sequence BA (allulose–aspartame). No data from these participants were included in the analyses. The participant flow is shown in Figure 2.

Figure 2. Flowchart of participant screening, exclusions, randomization, and study completion.

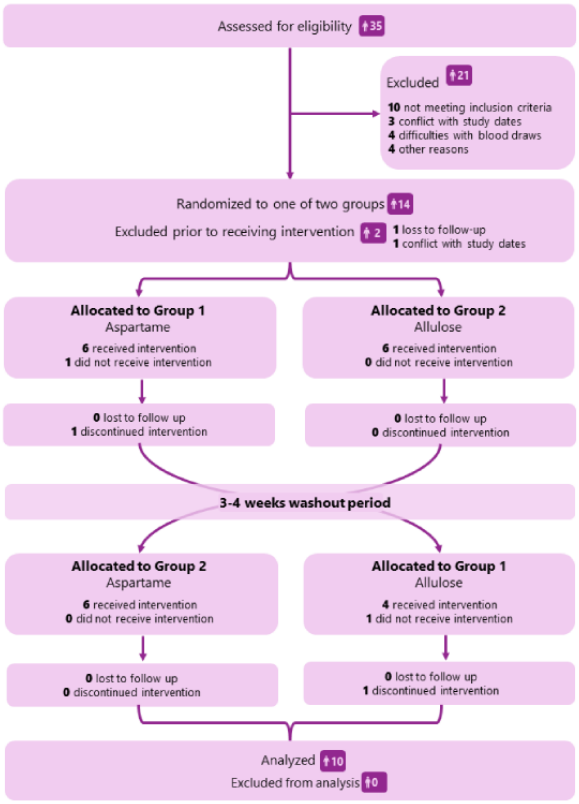

Baseline Characteristics

A total of 10 participants (4 female and 6 male) completed the study. The baseline characteristics of the participants are summarized in Table 2.

XSL•FO
RenderX

Table 2. Baseline characteristics (N=10).

Characteristic	Value
Age (year), mean (SD)	31.2 (6.8)
Sex, n	
Female	4
Male	6
BMI (kg/m^2), mean (SD)	25.1 (2.6)
Waist circumference (cm), mean (SD)	87.5 (8.1)
Hip circumference (cm), mean (SD)	101.9 (8.0)
Waist to hip ratio, mean (SD)	0.86 (0.1)
Energy requirement (kcal/day), mean (SD)	2481 (295)
Fasting glucose (mmol/L), mean (SD)	4.93 (0.4)
Fasting triglyceride (mmol/L), mean (SD)	1.09 (0.6)
Fasting cholesterol (mmol/L), mean (SD)	4.56 (0.44)
HbA_{1c}[a] (mmol/mol), mean (SD)	32.8 (3.5)

[a]HbA_{1c}: hemoglobin A1c.

Data Status

All biological samples and electronic data records were collected as planned by November 16, 2023. The following key milestones describe the progress of data processing and planned analysis.

Raw data, including all questionnaire responses and physiological readings, underwent a thorough cleaning and curation process, which was finalized in June 2025.

Laboratory analysis for the primary and key secondary outcomes (namely gut hormones and OGTT parameters) were performed in the Laboratory for Molecular Physiology at the Max Rubner-Institut and completed in October 2025. Data cleaning and final curation for these key parameters will be finalized in December 2025. The statistical analysis for the primary outcome and key secondary clinical outcomes is scheduled to commence immediately following final data lock in January 2026.

The exploratory multiomics analyses (sugaromics and microbiome sequencing) are expected to be completed and curated in December 2026, given the complexity and external laboratory processes involved.

We anticipate the main results manuscript, focusing on the primary and key secondary clinical outcomes, to be completed and submitted for publication by June 2026. Analysis and publication of the exploratory multiomics data will follow in separate, subsequent publications.

Discussion

Rationale and Significance

This study protocol describes a rigorous, randomized crossover trial designed to evaluate the metabolic and gut health effects of 4 weeks of allulose consumption compared with aspartame. We hypothesize that the consumption of allulose will lead to a measurably higher acute postprandial GLP-1 response compared with the aspartame placebo. We further aim to generate preliminary findings on a potential change in the postprandial GLP-1 response after 4 weeks of consumption and possible effects on insulin sensitivity, body weight and composition, and microbiota composition.

We anticipate that both interventions will be well tolerated, consistent with prior acute studies using similar doses.

This study is poised to provide novel, mechanistic insights by combining continuous glucose monitoring, comprehensive metabolomics (sugaromics), and gut microbiome analysis within a single, controlled, crossover-design pilot. The comprehensive dataset is expected to generate robust parameters for effect size and SDY, which are crucial for designing and powering a definitive, large-scale confirmatory trial on the long-term impact of allulose.

This study addresses the question of whether repeated allulose consumption leads to adaptation of the GLP-1 response, rather than a simple acute effect. Prior human trials have shown no significant changes in GLP-1 after prolonged low-calorie sweetener exposure (aspartame or sucralose) [21]. In contrast, animal models suggest that chronic sweetener intake may dampen the GLP-1 response over time, potentially via alterations in enteroendocrine signaling [23]. By including both acute and 4-week timepoints, our design allows us to explore whether an adaptive incretin response to allulose exists.

The urgency and clinical relevance of these findings have been further underscored by recent regulatory developments. Specifically, the European Food Safety Authority recently indicated that the safety of allulose as a novel food requires additional human data to close identified gaps [54]. Therefore, the results derived from this study will be critical to informing the scientific and regulatory debate surrounding allulose and its potential role in public health sugar reduction strategies.

XSL•FO
RenderX

Strengths and Limitations

The primary strength of this protocol lies in its robust randomized crossover design, which effectively reduces interindividual variability and increases statistical power. The design, coupled with the final analyzed cohort of 10, is statistically powered to test the primary hypothesis regarding postprandial GLP-1 response, thereby providing a strong foundation for the main outcome. Furthermore, the inclusion of multiomic endpoints (sugaromics and microbiome) alongside clinical measures provides a deep, integrated understanding of the underlying mechanisms.

While sufficiently powered for the primary outcome, the modest overall sample size (N=10) confines the interpretation of all secondary endpoints to hypothesis-generating findings and significantly limits the generalizability of the results. This pilot phase will, however, provide essential data to power future confirmatory studies. Despite the robust design, potential limitations in feasibility could arise, such as participant burden leading to potential nonresponse in questionnaires or technical challenges with complex measurements, such as indirect calorimetry, requiring strict environmental control to ensure reliable data. While double-blind, the differing gut effects of active sweeteners may potentially compromise participant blinding, which could influence subjective outcomes.

Conclusions

In summary, this protocol describes a methodologically rigorous study designed to investigate the metabolic and gut health effects of allulose consumption. The established methodology and successful data collection are expected to generate critical feasibility and parameter data. The study's focus on acute postprandial GLP-1 profiles (primary outcome), coupled with comprehensive secondary outcomes such as subacute effects on GLP-1 and other gut hormones, as well as sugaromics and gut microbiome analysis, will provide a novel, integrated understanding of allulose's mechanisms of action.

Acknowledgments

We sincerely thank all study participants for their time, commitment, and valuable contributions to this research. We are also grateful to the study personnel, including nursing staff, laboratory technicians, and administrative support, whose dedication and professionalism were essential to the successful implementation of the study. In particular, we would like to acknowledge Maria Herrmann, Ulrike Rossi, Isabel Albrecht, Eva Hoch, Martina Broßart, Tanja Gadau, Gisela Schultheiss, Julia Renz, and Serap Demirel for their assistance in participant coordination and sample processing, respectively.

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing. The GAI tool used was in-house Llama-3.3 (Meta Platforms, Inc). Responsibility for the final manuscript lies entirely with the authors.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Funding

This study was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft; project number 517694181; principal investigator BH).

Authors' Contributions

Conceptualization: AB, BH, CIM, PGF
Resources: AB, BH, OW, SS
Supervision: AB, BH, SS
Project administration: AB, BH, SS
Writing – Review & Editing: AB, AKE, A-KH, BH, CIM, OW, PGF, SB, SS
Methodology: AKE, BH, CIM, SS
Investigation: A-KH, BH, OW, SB
Formal Analysis: BH, PGF
Data Curation: BH, PGF
Visualization: BH
Funding acquisition: BH
Writing – Original Draft: BH, SB
Software: PGF

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
SPIRIT checklist.
[PDF File (Adobe PDF File), 226 KB - resprot_v15i1e81857_app1.pdf]

Multimedia Appendix 2
Informed consent form in the original study language (German).
[PDF File (Adobe PDF File), 217 KB - resprot_v15i1e81857_app2.pdf]

References

1. Perrar I, Gray N, Kuhnle GG, Remer T, Buyken AE, Alexy U. Sugar intake among German adolescents: trends from 1990 to 2016 based on biomarker excretion in 24-h urine samples. Br J Nutr 2020;124(2):164-172. [doi: 10.1017/S0007114520000665] [Medline: 32102699]
2. Mensink GBM, Haftenberger M, Clarissa LB, Brettschneider AK, Franziska L, Melanie F. EsKiMo II - Die Ernährungsstudie als KiGGS-Modul. Berlin: Robert Koch-Institut; 2020.
3. Nationale Verzehrsstudie II. Ergebnissbericht Karlsruhe. Germany: Max Rubner-Institut; 2008.
4. Huang Y, Chen Z, Chen B, Li J, Yuan X, Li J. Dietary sugar consumption and health: umbrella review. BMJ 2023;381:e071609 [FREE Full text] [doi: 10.1136/bmj-2022-071609] [Medline: 37019448]
5. Lane MM, Travica N, Gamage E, Marshall S, Trakman GL, Young C. Sugar-sweetened beverages and adverse human health outcomes: an umbrella review of meta-analyses of observational studies. Annu Rev Nutr 2024;44(1):383-404 [FREE Full text] [doi: 10.1146/annurev-nutr-062322-020650] [Medline: 39207876]
6. Popkin BM, Ng SW. Sugar-sweetened beverage taxes: lessons to date and the future of taxation. PLoS Med 2021;18(1):e1003412 [FREE Full text] [doi: 10.1371/journal.pmed.1003412] [Medline: 33411708]
7. Oshima H, Kimura I, Izumori K. Psicose contents in various food products and its origin. Food Sci Technol Res 2006;12(2):137-143. [doi: 10.3136/fstr.12.137]
8. Guidance for industry: the declaration of allulose and calories from allulose on nutrition and supplement facts labels. US Food and Drug Administration. 2019. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/guidance-industry-declaration-allulose-and-calories-allulose-nutrition-and-supplement-facts-labels [accessed 2026-01-30]
9. Kishida K, Martinez G, Iida T, Yamada T, Ferraris RP, Toyoda Y. d-Allulose is a substrate of glucose transporter type 5 (GLUT5) in the small intestine. Food Chem 2019;277:604-608. [doi: 10.1016/j.foodchem.2018.11.003] [Medline: 30502192]
10. Iida T, Hayashi N, Yamada T, Yoshikawa Y, Miyazato S, Kishimoto Y. Failure of d-psicose absorbed in the small intestine to metabolize into energy and its low large intestinal fermentability in humans. Metabolism 2010;59(2):206-214. [doi: 10.1016/j.metabol.2009.07.018] [Medline: 19765780]
11. Han Y, Choi BR, Kim SY, Kim SB, Kim YH, Kwon EY. Gastrointestinal tolerance of d-allulose in healthy and young adults. A non-randomized controlled trial. Nutrients 2018;10(12):2010 [FREE Full text] [doi: 10.3390/nu10122010] [Medline: 30572580]
12. Blundell J, Rogers P, Hill A. Evaluating the satiating power of foods: implications for acceptanceconsumption. In: Colms J, Booth DA, Pangborn RM, Raunhardt O, editors. Food Acceptance and Nutrition. London: Academic Press; 1987:205-219.
13. Mela DJ. Eating for pleasure or just wanting to eat? Reconsidering sensory hedonic responses as a driver of obesity. Appetite 2006;47(1):10-17. [doi: 10.1016/j.appet.2006.02.006] [Medline: 16647788]
14. Hayakawa M, Hira T, Nakamura M, Iida T, Kishimoto Y, Hara H. Secretion of GLP-1 but not GIP is potently stimulated by luminal d-allulose (d-psicose) in rats. Biochem Biophys Res Commun 2018;S0006-291X(18):30143-30148. [doi: 10.1016/j.bbrc.2018.01.128] [Medline: 29402406]
15. Iwasaki Y, Sendo M, Dezaki K, Hira T, Sato T, Nakata M. GLP-1 release and vagal afferent activation mediate the beneficial metabolic and chronotherapeutic effects of D-allulose. Nat Commun 2018;9(1):113 [FREE Full text] [doi: 10.1038/s41467-017-02488-y] [Medline: 29317623]
16. Kuhre RE, Gribble FM, Hartmann B, Reimann F, Windeløv JA, Rehfeld JF. Fructose stimulates GLP-1 but not GIP secretion in mice, rats, and humans. Am J Physiol Gastrointest Liver Physiol 2014;306(7):G622-G630 [FREE Full text] [doi: 10.1152/ajpgi.00372.2013] [Medline: 24525020]
17. Teff KL, Elliott SS, Tschöp M, Kieffer TJ, Rader D, Heiman M. Dietary fructose reduces circulating insulin and leptin, attenuates postprandial suppression of ghrelin, and increases triglycerides in women. J Clin Endocrinol Metab 2004;89(6):2963-2972 [FREE Full text] [doi: 10.1210/jc.2003-031855] [Medline: 15181085]
18. Page KA, Chan O, Arora J, Belfort-Deaguiar R, Dzuira J, Roehmholdt B. Effects of fructose vs glucose on regional cerebral blood flow in brain regions involved with appetite and reward pathways. JAMA 2013;309(1):63-70 [FREE Full text] [doi: 10.1001/jama.2012.116975] [Medline: 23280226]
19. Teysseire F, Bordier V, Budzinska A, Weltens N, Rehfeld JF, Holst JJ. The role of d-allulose and erythritol on the activity of the gut sweet taste receptor and gastrointestinal satiation hormone release in humans: a randomized, controlled trial. J Nutr 2022;152(5):1228-1238 [FREE Full text] [doi: 10.1093/jn/nxac026] [Medline: 35135006]

XSL•FO
RenderX

20. Bodnaruc AM, Prud'homme D, Blanchet R, Giroux I. Nutritional modulation of endogenous glucagon-like peptide-1 secretion: a review. Nutr Metab (Lond) 2016;13:92 [FREE Full text] [doi: 10.1186/s12986-016-0153-3] [Medline: 27990172]
21. Orku SE, Suyen G, Bas M. The effect of regular consumption of four low- or no-calorie sweeteners on glycemic response in healthy women: A randomized controlled trial. Nutrition 2023;106:111885. [doi: 10.1016/j.nut.2022.111885] [Medline: 36470113]
22. Higgins KA, Considine RV, Mattes RD. Aspartame consumption for 12 weeks does not affect glycemia, appetite, or body weight of healthy, lean adults in a randomized controlled trial. J Nutr 2018;148(4):650-657 [FREE Full text] [doi: 10.1093/jn/nxy021] [Medline: 29659969]
23. Swithers SE, Laboy AF, Clark K, Cooper S, Davidson TL. Experience with the high-intensity sweetener saccharin impairs glucose homeostasis and GLP-1 release in rats. Behav Brain Res 2012;233(1):1-14 [FREE Full text] [doi: 10.1016/j.bbr.2012.04.024] [Medline: 22561130]
24. Noronha JC, Braunstein CR, Glenn AJ, Khan TA, Viguiliouk E, Noseworthy R. The effect of small doses of fructose and allulose on postprandial glucose metabolism in type 2 diabetes: a double-blind, randomized, controlled, acute feeding, equivalence trial. Diabetes Obes Metab 2018;20(10):2361-2370 [FREE Full text] [doi: 10.1111/dom.13374] [Medline: 29797503]
25. Braunstein CR, Noronha JC, Glenn AJ, Viguiliouk E, Noseworthy R, Khan TA. A double-blind, randomized controlled, acute feeding equivalence trial of small, catalytic doses of fructose and allulose on postprandial blood glucose metabolism in healthy participants: the fructose and allulose catalytic effects (FACE) trial. Nutrients 2018;10(6):750 [FREE Full text] [doi: 10.3390/nu10060750] [Medline: 29890724]
26. Franchi F, Yaranov DM, Rollini F, Rivas A, Rivas Rios J, Been L. Effects of D-allulose on glucose tolerance and insulin response to a standard oral sucrose load: results of a prospective, randomized, crossover study. BMJ Open Diabetes Res Care 2021;9(1):e001939 [FREE Full text] [doi: 10.1136/bmjdrc-2020-001939] [Medline: 33637605]
27. Shintani T, Yamada T, Hayashi N, Iida T, Nagata Y, Ozaki N. Rare sugar syrup containing d-allulose but not high-fructose corn syrup maintains glucose tolerance and insulin sensitivity partly via hepatic glucokinase translocation in wistar rats. J Agric Food Chem 2017;65(13):2888-2894. [doi: 10.1021/acs.jafc.6b05627] [Medline: 28209058]
28. Matsuo T, Izumori K. Effects of dietary D-psicose on diurnal variation in plasma glucose and insulin concentrations of rats. Biosci Biotechnol Biochem 2006;70(9):2081-2085. [doi: 10.1271/bbb.60036] [Medline: 16960391]
29. Han Y, Han HJ, Kim A, Choi J, Cho S, Park YB. d-Allulose supplementation normalized the body weight and fat-pad mass in diet-induced obese mice via the regulation of lipid metabolism under isocaloric fed condition. Mol Nutr Food Res 2016;60(7):1695-1706. [doi: 10.1002/mnfr.201500771] [Medline: 26920079]
30. Chung YM, Hyun Lee J, Youl Kim D, Hwang SH, Hong YH, Kim SB. Dietary D-psicose reduced visceral fat mass in high-fat diet-induced obese rats. J Food Sci 2012;77(2):H53-H58. [doi: 10.1111/j.1750-3841.2011.02571.x] [Medline: 22339545]
31. Ochiai M, Nakanishi Y, Yamada T, Iida T, Matsuo T. Inhibition by dietary D-psicose of body fat accumulation in adult rats fed a high-sucrose diet. Biosci Biotechnol Biochem 2013;77(5):1123-1126 [FREE Full text] [doi: 10.1271/bbb.130019] [Medline: 23649241]
32. Nagata Y, Kanasaki A, Tamaru S, Tanaka K. D-psicose, an epimer of D-fructose, favorably alters lipid metabolism in Sprague-Dawley rats. J Agric Food Chem 2015;63(12):3168-3176. [doi: 10.1021/jf502535p] [Medline: 25754616]
33. Hossain A, Yamaguchi F, Matsunaga T, Hirata Y, Kamitori K, Dong Y. Rare sugar D-psicose protects pancreas β-islets and thus improves insulin resistance in OLETF rats. Biochem Biophys Res Commun 2012;425(4):717-723. [doi: 10.1016/j.bbrc.2012.07.135] [Medline: 22877751]
34. Han Y, Kwon EY, Yu MK, Lee SJ, Kim HJ, Kim SB. A preliminary study for evaluating the dose-dependent effect of d-allulose for fat mass reduction in adult humans: a randomized, double-blind, placebo-controlled trial. Nutrients 2018;10(2):160 [FREE Full text] [doi: 10.3390/nu10020160] [Medline: 29385054]
35. Hayashi N, Yamada T, Takamine S, Iida T, Okuma K, Tokuda M. Weight reducing effect and safety evaluation of rare sugar syrup by a randomized double-blind, parallel-group study in human. J Funct Foods 2014;11:152-159. [doi: 10.1016/j.jff.2014.09.020]
36. Kimura T, Kanasaki A, Hayashi N, Yamada T, Iida T, Nagata Y. d-Allulose enhances postprandial fat oxidation in healthy humans. Nutrition 2017;43-44:16-20. [doi: 10.1016/j.nut.2017.06.007] [Medline: 28935140]
37. Han Y, Kwon EY, Choi MS. Anti-diabetic effects of allulose in diet-induced obese mice via regulation of mrNA expression and alteration of the microbiome composition. Nutrients 2020;12(7):2113 [FREE Full text] [doi: 10.3390/nu12072113] [Medline: 32708827]
38. Han Y, Park H, Choi BR, Ji Y, Kwon EY, Choi MS. Alteration of microbiome profile by D-allulose in amelioration of high-fat-diet-induced obesity in mice. Nutrients 2020;12(2):352 [FREE Full text] [doi: 10.3390/nu12020352] [Medline: 32013116]
39. Martin RM, Cao J, Wu W, Zhao L, Manthei DM, Pirani A. Identification of pathogenicity-associated loci in klebsiella pneumoniae from hospitalized patients. mSystems 2018;3(3):e00015 [FREE Full text] [doi: 10.1128/mSystems.00015-18] [Medline: 29963640]

XSL•FO
RenderX

40. Zuckerersatz Allulose: Für eine gesundheitliche Bewertung als lebensmittelzutat sind weitere daten erforderlich BfR-Stellungnahmen. Bundesinstitut für Risikobewertung. 2010. URL: https://www.bfr.bund.de/stellungnahme/zuckerersatz-allulose-fuer-eine-gesundheitliche-bewertung-als-lebensmittelzutat-sind-weitere-daten-erforderlich/ [accessed 2026-01-30]
41. Daniel H, Hauner H, Hornef MW, Clavel T. Allulose in human diet: the knowns and the unknowns. Br J Nutr 2022;128(2):172-178. [doi: 10.1017/S0007114521003172] [Medline: 34409930]
42. Butcher NJ, Monsour A, Mew EJ, Chan AW, Moher D, Mayo-Wilson E. Guidelines for reporting outcomes in trial protocols: the SPIRIT-outcomes 2022 extension. JAMA 2022;328(23):2345-2356. [doi: 10.1001/jama.2022.21243] [Medline: 36512367]
43. Pudel V, Westenhöfer J. Fragebogen zum Eßverhalten (FEV). 1989. URL: https://www.testzentrale.de/shop/fragebogen-zum-essverhalten.html [accessed 2026-01-30]
44. Stunkard AJ, Messick S. The three-factor eating questionnaire to measure dietary restraint, disinhibition and hunger. J Psychosom Res 1985;29(1):71-83. [doi: 10.1016/0022-3999(85)90010-8] [Medline: 3981480]
45. Regulation (EC) No 1924/2006 of the European Parliament and of the Council of 20 December 2006 on nutrition and health claims made on foods. European Union. 2006. URL: https://eur-lex.europa.eu/eli/reg/2006/1924/oj/eng [accessed 2026-01-30]
46. Heuer T. Zuckerkonsum in Deutschland. Aktuel Ernahrungsmed 2018;43(S 01):S8-S11. [doi: 10.1055/a-0659-8828]
47. Sigala DM, Hieronimus B, Medici V, Lee V, Nunez MV, Bremer AA. The dose-response effects of consuming high fructose corn syrup-sweetened beverages on hepatic lipid content and insulin sensitivity in young adults. Nutrients 2022;14(8):1648 [FREE Full text] [doi: 10.3390/nu14081648] [Medline: 35458210]
48. Boxall LR, Eskandari F, Wallis J, Bielat AD, Appleton KM. The effects of aspartame on glucose, insulin, and appetite-regulating hormone responses in humans: systematic review and meta-analyses. Adv Nutr 2025;16(7):100449 [FREE Full text] [doi: 10.1016/j.advnut.2025.100449] [Medline: 40381807]
49. Mifflin MD, St Jeor ST, Hill LA, Scott BJ, Daugherty SA, Koh YO. A new predictive equation for resting energy expenditure in healthy individuals. Am J Clin Nutr 1990;51(2):241-247. [doi: 10.1093/ajcn/51.2.241] [Medline: 2305711]
50. Blundell J, de Graaf C, Hulshof T, Jebb S, Livingstone B, Lluch A. Appetite control: methodological aspects of the evaluation of foods. Obes Rev 2010;11(3):251-270 [FREE Full text] [doi: 10.1111/j.1467-789X.2010.00714.x] [Medline: 20122136]
51. Matsuda M, DeFronzo RA. Insulin sensitivity indices obtained from oral glucose tolerance testing: comparison with the euglycemic insulin clamp. Diabetes Care 1999;22(9):1462-1470. [doi: 10.2337/diacare.22.9.1462] [Medline: 10480510]
52. Haftenberger M, Heuer T, Heidemann C, Kube F, Krems C, Mensink GBM. Relative validation of a food frequency questionnaire for national health and nutrition monitoring. Nutr J 2010;9:36 [FREE Full text] [doi: 10.1186/1475-2891-9-36] [Medline: 20840739]
53. Naska A, Lagiou A, Lagiou P. Dietary assessment methods in epidemiological research: current state of the art and future prospects. F1000Res 2017;6:926 [FREE Full text] [doi: 10.12688/f1000research.10703.1] [Medline: 28690835]
54. EFSA Panel on Nutrition, Novel FoodsFood Allergens (NDA), Turck D, Cámara M, Bohn T, Castenmiller J, De Henauw S, et al. Safety of D-allulose as a novel food pursuant to regulation (EU) 2015/2283. EFSA J 2025;23(6):e9468. [doi: 10.2903/j.efsa.2025.9468] [Medline: 40487234]

Abbreviations

AE: adverse event
AUC: area under the curve
FFQ: food frequency questionnaire
GLP-1: glucagon-like peptide 1
HbA1c: hemoglobin A1c
iAUC: incremental area under the curve
LisA: low-calorie sweetener intervention study allulose
LOQ: limit of quantification
OGTT: oral glucose tolerance test
SDY: study day
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
VAS: visual analog scales

XSL•FO
RenderX

Edited by A Schwartz; submitted 09.Sep.2025; peer-reviewed by M Anwar Ul Alam; comments to author 13.Nov.2025; accepted 09.Jan.2026; published 19.Feb.2026.

Please cite as:
Busch S, Ferrario PG, Henk AK, Engelbert AK, Wittek O, Seifert S, Bub A, Mack CI, Hieronimus B
Effects of Allulose vs Aspartame Consumption on Postprandial Glucagon-Like Peptide-1 Profiles and Metabolic Health: Protocol for a Randomized, Crossover, Double-Blind, Placebo-Controlled Trial
JMIR Res Protoc 2026;15:e81857
URL: https://www.researchprotocols.org/2026/1/e81857
doi:10.2196/81857
PMID:

©Selina Busch, Paola G Ferrario, Ann-Kathrin Henk, Ann Katrin Engelbert, Oliver Wittek, Stephanie Seifert, Achim Bub, Carina I Mack, Bettina Hieronimus. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Impact of Ocular Massage on Intraocular Pressure and Schlemm Canal Dimensions in Healthy Adults: Protocol for a Randomized Controlled Trial

Mingxuan Zhang[1], MD; Luning Qin[2], MD; Liangzhang Tan[1], MD; Yongtao Li[2], PhD; Qing Zhang[2], PhD; Shuhan Wang[3], MD; Zhihui Zhang[2], PharmB, MSc, PhD; Emmanuel Eric Pazo[4], MD, MSci, PhD; Xinjun Ren[1], MD, PhD

[1]Department of Ocular Trauma, Tianjin Key Laboratory of Retinal Functions and Diseases, Tianjin Branch of National Clinical Research Center for Ocular Disease, Eye Institute and School of Optometry, Tianjin Medical University Eye Hospital, Tianjin, China

[2]Tianjin Key Laboratory of Retinal Functions and Diseases, Tianjin Branch of National Clinical Research Center for Ocular Disease, Eye Institute and School of Optometry, Tianjin Medical University Eye Hospital, Tianjin, China

[3]Independent Consultant, Shanghai, China

[4]Department of Clinical Research, Tianjin Key Laboratory of Retinal Functions and Diseases, Tianjin Branch of National Clinical Research Center for Ocular Disease, Eye Institute and School of Optometry, Tianjin Medical University Eye Hospital, Tianjin, China

Corresponding Author:
Xinjun Ren, MD, PhD
Department of Ocular Trauma
Tianjin Key Laboratory of Retinal Functions and Diseases, Tianjin Branch of National Clinical Research Center for Ocular Disease, Eye Institute and School of Optometry, Tianjin Medical University Eye Hospital
251 Fu Kang Road
Tianjin, 300384
China
Phone: 1 86 022 86428787
Fax: 1 86 022 23346434
Email: xin2023ren@tmu.edu.cn

Abstract

Background: Ocular massage has been reported to lower intraocular pressure (IOP) temporarily. This effect could be related to enhanced aqueous humor outflow; however, the mechanism is unclear.

Objective: This study aims to examine the impact of digital and EyePeace ocular massage on IOP fluctuations and investigate whether the observed reduction in IOP is attributable to morphological changes in the Schlemm canal (SC).

Methods: Participants were randomly assigned in a 1:1:1 ratio to 1 of 3 groups: digital ocular massage, EyePeace ocular massage, or no ocular massage. The primary outcome measure will be IOP. The secondary outcome measures will include anterior segment optical coherence tomography assessment of the SC and trabecular meshwork. In addition, adverse events in the quality of vision will be monitored and documented using a mobile app–based questionnaire. All assessments will be performed at baseline and 10 minutes after the assigned interventions.

Results: Data collection was completed in the summer of 2025. The study results are expected to be published by the end of 2025.

Conclusions: This study will investigate the impact of ocular massage on IOP and SC dimensions by using a randomized controlled design. As preliminary evidence suggests that ocular massage may transiently reduce IOP by enhancing aqueous humor outflow, this study will aim to clarify the underlying mechanism. By comparing digital ocular massage, EyePeace ocular massage, and no ocular massage, the study will assess changes in IOP and SC morphology. The findings are expected to provide insights into the role of mechanical manipulation in modulating aqueous humor outflow, potentially informing nonpharmacological strategies for possible glaucoma management.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2400093512; https://www.chictr.org.cn/showproj.html?proj=250459

International Registered Report Identifier (IRRID): DERR1-10.2196/78864

(JMIR Res Protoc 2026;15:e78864) doi:10.2196/78864

XSL•FO
RenderX

KEYWORDS

ocular massage; Schlemm canal; trabecular meshwork; intraocular pressure; aqueous humor

Introduction

Background

Ocular massage has been recognized as a potential nonpharmacological technique to transiently reduce intraocular pressure (IOP) across various clinical scenarios. It has notably been used after glaucoma surgeries such as trabeculectomy to manage postoperative pressure fluctuations [1,2] presumably through the expansion of filtration blebs, enhancing aqueous humor drainage [3,4]. Moreover, ocular massage has shown efficacy in managing acute elevations of IOP after intravitreal injections and in acute angle-closure scenarios, demonstrating its versatility in ophthalmic practice [5,6].

In the setting of acute angle-closure attacks, patients can be instructed to perform ocular massage as an initial intervention to transiently lower IOP before definitive medical or surgical treatment can be administered [7,8]. Despite these promising clinical applications, previous studies have limitations that hinder drawing definitive conclusions about the underlying mechanisms and the sustained effectiveness of ocular massage [6,7]. Although the prior studies demonstrated temporary reductions in IOP after digital ocular massage, they were limited by the wide variability in the pressure applied by clinicians or patients and in the duration and frequency of massage because even small variations in finger placement or pressure can result in unpredictable IOP responses. These findings collectively highlight the potential clinical applications of ocular massage in various ophthalmic contexts [9]. EyePeace (Eye Comfort Ltd) is a commercially available device for eyelid massage designed to enhance meibum expression for dry eye disease related to meibomian gland dysfunction [10,11]. This flexible, handheld silicone device is used in conjunction with warm compress therapy and applies controlled vertical pressure to the closed eyelids [10,11]. EyePeace addresses the key limitations of manual digital ocular massage by offering a standardized, safe, and reproducible alternative. Unlike digital ocular massage, which is limited by inconsistent finger pressure, variable technique, and anatomical inaccuracy, EyePeace delivers calibrated and uniform mechanical force to the eyelids, ensuring consistent application across users and sessions. Its ergonomic design promotes proper positioning and motion, improving the reliability of therapeutic outcomes. By operating within safe pressure thresholds, EyePeace minimizes the risk of ocular complications such as hypotony, optic nerve stress, or retinal damage that may result from manual overcompression. In addition, its ease of use facilitates home-based therapy, enhancing patient adherence and long-term management. Originally developed for meibomian gland dysfunction, the device may also support aqueous humor outflow through gentle scleral compression, suggesting potential utility in select glaucoma settings. Overall, EyePeace offers a more controlled and patient-friendly approach that overcomes the variability and safety concerns associated with digital ocular massage.

IOP homeostasis is primarily dictated by the dynamic balance between aqueous humor production and outflow. The conventional drainage pathway involves the trabecular meshwork (TM), which directs aqueous humor into the Schlemm canal (SC) and then into aqueous and episcleral veins via collector channels [1,5]. A secondary, unconventional uveoscleral outflow pathway also contributes to aqueous humor clearance, albeit to a lesser extent [2,3,11-13]. Previous studies have provided insights into anatomical variations in aqueous humor drainage structures and their implications for IOP regulation [4,14]. Kagemann et al [15] reported that the cross-sectional area of the SC was significantly reduced in patients with glaucoma compared to healthy individuals, suggesting a structural contribution to elevated IOP [6,7,16,17]. Similarly, Chung et al [18] found that patients with newly diagnosed open-angle glaucoma who had larger baseline SC areas exhibited a more substantial IOP-lowering response to topical hypotensive agents, whereas TM width did not seem to influence the degree of pressure reduction. Advances in high-resolution imaging have now made it possible to visualize the corneoscleral region, including the SC and TM [19,20].

Objectives

This study aims to evaluate changes in IOP after the use of the EyePeace ocular massager. We hypothesize that EyePeace will achieve a more consistent and pronounced IOP reduction compared to digital ocular massage. Furthermore, the observed IOP reduction may be correlated with morphological changes in underlying ocular structures.

Methods

Study Design

This single-masked, randomized controlled trial will enroll 36 healthy Chinese participants (72 eyes) aged 18 years or older. They will be randomly assigned in a 1:1:1 ratio to the digital ocular massage group (24 eyes), EyePeace ocular massage group (24 eyes), or no ocular massage (24 eyes). The trial will be conducted at Tianjin Medical University Eye Hospital (TMUEH) in Tianjin, China, a nationally accredited clinical trial site. Healthy participants who voluntarily consent to undergo ocular massage will be invited to join this study to assess noninvasive clinical parameters. Patients and the public will not be involved in the design, implementation, reporting, or dissemination of this study. This study was registered on the Chinese Clinical Trial Registry (ChiCTR2400093512) on December 6, 2024.

Eligibility Criteria

The inclusion criteria were as follows: (1) age 18 years or older; (2) willingness to provide written informed consent, accept random allocation, and participate in ocular massage interventions, follow-ups, and assessments; and (3) IOP within the target range of 10 to 21 mm Hg. Patients were excluded based on the following criteria: (1) prior diagnosis of an eye disease, (2) history of ocular surgery, (3) current use of medications that might affect anterior segment ocular anatomy,

XSL•FO
RenderX

(4) current contact lens wear, (5) history of ocular allergy, and (6) systemic disorders [21,22].

Interventions

Eligible participants will be randomly assigned in a 1:1:1 ratio to (1) digital ocular massage (group A), (2) EyePeace ocular massage (group B), or (3) no ocular massage (group C). Group A and group B participants will receive ocular massage for 10 minutes [5,10,11]. Any participant experiencing an adverse event (AE) will be withdrawn from the trial, documented, provided appropriate treatment, and referred immediately to the principal investigator (PI). Clinical research staff will receive professional training covering participant enrollment, treatment procedures (including risks and benefits), and data collection processes. Participants will be sufficiently informed about the study and provided with compliance education to enhance participation and minimize noncompliance. Study staff will proactively contact participants in advance by telephone or SMS text message. EyePeace is a flexible, medical-grade silicone device designed to apply gentle, pulsatile vertical pressure to the closed eyelids. Although it does not contain a built-in pressure gauge, manufacturer specifications suggest that the applied force typically ranges from 0.2 to 0.5 N. To ensure consistency, all participants in the EyePeace ocular massage group received standardized instructions and a demonstration before use. In addition, a trained examiner supervised the massage procedure to ensure correct technique and duration. For the digital ocular massage group, clinical staff applied consistent fingertip pressure guided by prior training and feedback from pilot sessions, aiming to replicate the mechanical effect of the EyePeace ocular massage device. Although exact biomechanical equivalence between digital ocular massage and EyePeace ocular massage cannot be assured, the standardization procedures were designed to ensure comparability in massage duration, eyelid area coverage, and application frequency.

Provisions for Posttrial Care

Participants who experience symptoms of conjunctivitis, keratitis, or other adverse reactions during the study or within 6 months after its completion will receive timely, targeted treatment. All such events will be documented in the AE record.

Primary Outcome: IOP

The study's primary outcome is assessing IOP (mm Hg) before and after the intervention. All participants will undergo ophthalmologic examinations on the same day. A slit lamp examination will first be performed on all eyes, followed by IOP measurement in the sitting position, using an iCare IC100 rebound tonometer (iCare Finland Oy) [19]. The instrument will be calibrated regularly according to the manufacturer's instructions. A disposable, single-use probe was loaded into the device and aligned 4 to 8 mm perpendicular to the central cornea. Six consecutive measurements were performed. The software automatically discarded the highest and lowest values, and the IOP was calculated from the remaining 4 values. Only proper measurements (indicated by a green background within acceptable limits) were included.

Secondary Outcomes

Best Corrected Visual Acuity

Best corrected visual acuity, reported in logarithm of the minimum angle of resolution units, will be measured using a standardized visual acuity chart under controlled lighting conditions to assess any changes due to ocular massage [10,20]. Baseline and postintervention measurements will be recorded to determine the impact on visual function.

SC Area and SC Diameter

SC area (μm^2) and SC diameter (μm) will be analyzed using anterior segment optical coherence tomography (AS-OCT) [11,23]. High-resolution images of the temporal corneoscleral region will be captured at baseline and after the intervention, ensuring consistency in scanning alignment. SC area will be delineated manually using ImageJ software, while SC diameter will be measured from the anterior to posterior endpoints of the canal [12,24].

TM Width and TM Thickness

TM parameters, including width (μm) and thickness (μm), will be assessed from AS-OCT images [13]. TM width will be defined as the distance between the scleral spur and the Schwalbe line, whereas TM thickness will be measured at the anterior endpoint and midpoint of the SC [14,25].

Mobile App–Based Quality of Vision Questionnaire

The original Quality of Vision Questionnaire includes 10 visual symptoms, measured in terms of frequency, severity, and impact on daily activities. It was designed to assess patients' subjective visual experience, especially after refractive or cataract surgery [26]. In this study, participants will complete the Quality of Vision Questionnaire [26] before and after ocular massage to assess subjective visual disturbances, visual clarity, and visual discomfort. The questionnaire will be delivered via a web application on the WeChat platform (Tencent Holdings Ltd) [27]. An independent examiner will perform all measurements to minimize bias, and data will be analyzed for statistical significance using appropriate paired comparison tests. In addition to the Quality of Vision Questionnaire, 3 items will be administered to measure overall quality of vision score (Likert scale, 0-10), ocular pain score (Likert scale, 0-10), and ocular discomfort score (Likert scale, 0-10; Multimedia Appendix 1).

Participant Timeline

During the study, participants undergo ocular examinations at baseline and 10 minutes after the intervention. The study procedures for participant recruitment, intervention, assessment, and data follow-up follow the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines (Table 1) [28]. The study process is illustrated in Figure 1.

XSL•FO
RenderX

Table 1. Study procedures conducted in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines.

Time point	Study period			
	Enrollment	Allocation	After allocation	
	Recruitment (2-4 wk)	Random assignment (0 d)	Baseline	Follow-up
Enrollment				
Eligibility screening	✓			
Informed consent	✓			
Random assignment		✓		
Intervention				
Digital ocular massage group			✓	✓
EyePeace ocular massage group			✓	✓
No ocular massage group			✓	✓
Assessments				
Intraocular pressure (mm Hg)			✓	✓
Best corrected visual acuity (logarithm of the minimum angle of resolution)			✓	✓
Schlemm canal area (μm^2)			✓	✓
Schlemm canal diameter (μm)			✓	✓
Trabecular meshwork width (μm)			✓	✓
Trabecular meshwork thickness (μm)			✓	✓
Quality of Vision Questionnaire score			✓	✓

XSL•FO
RenderX

Figure 1. Study design flowchart. AE: adverse event; IOP: intraocular pressure; QoV: Quality of Vision; SC: Schlemm canal; TM: trabecular meshwork; TMUEH: Tianjin Medical University Eye Hospital.

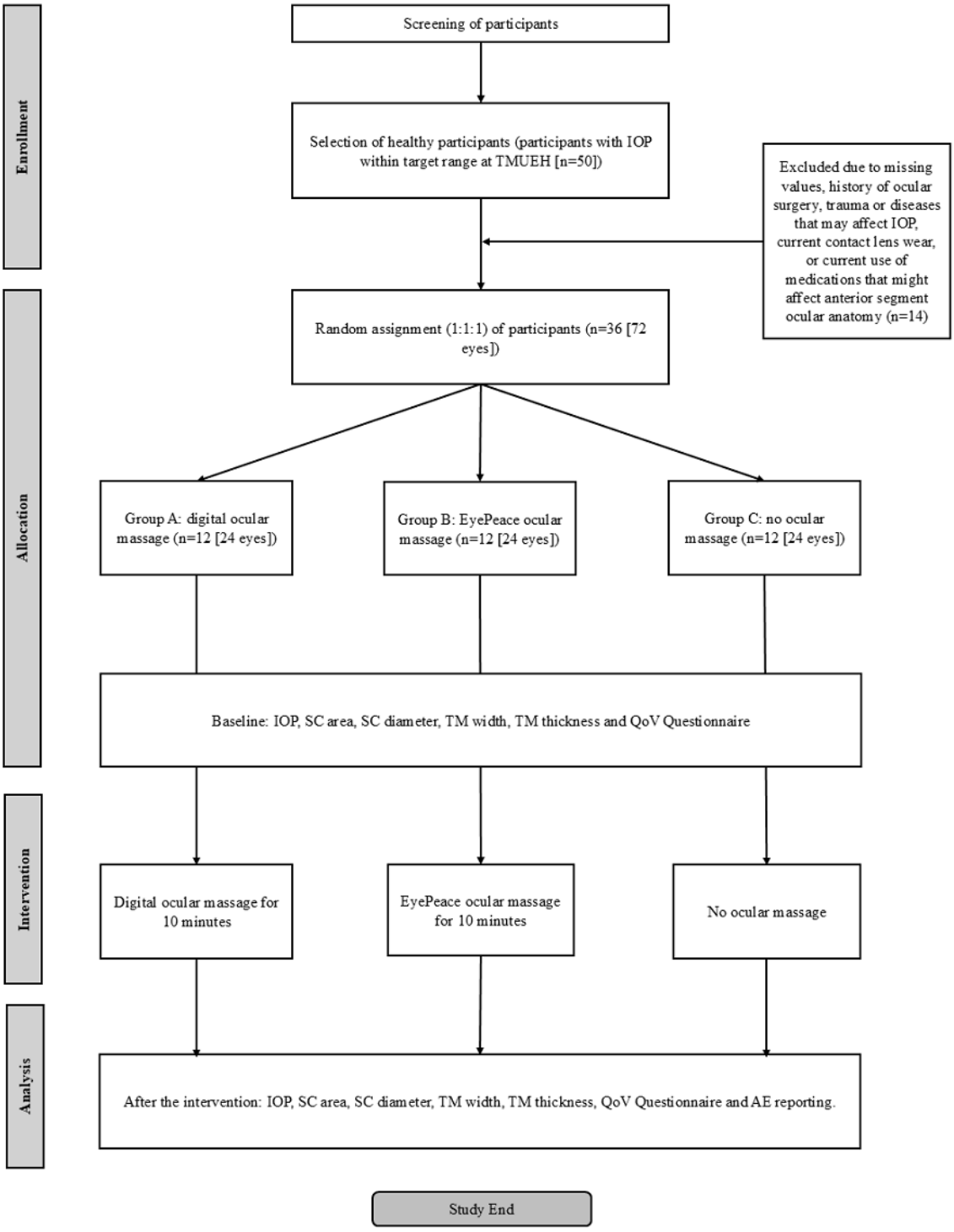

Sample Size

On the basis of the stated assumptions (a common SD of 1.9 and a maximum mean difference in IOP between groups of approximately 2 mmHg), a total of 36 participants (72 eyes [participants will contribute both eyes unless exclusion criteria apply]; 3 groups; 24 eyes per group) are required to achieve 90% power at α=.05 [1,5]. This corresponds to 3 groups of equal size in a randomized controlled trial. The calculations use a 1-way analysis of variance to test overall mean differences and consider a moderate to large effect size in IOP reduction. Allowing for a 10% dropout rate, the recruitment target is 40 participants (80 eyes). If further data refine the expected mean difference, the sample size will be adjusted accordingly (larger samples for smaller expected differences and vice versa).

XSL•FO
RenderX

Recruitment

This study will be conducted from March to June 2025 at TMUEH. Healthy participants in Tianjin, China, will be invited to participate. Recruitment will be carried out through posters, online advertisements, and social media platforms. Specialized professionals will collect demographic information of the participants during the initial screening.

Allocation

The patients will be randomly assigned to 1 of 3 groups: digital ocular massage (group A), EyePeace ocular massage (group B), or no ocular massage (group C). Randomization will be performed by a statistician using a network-based randomization procedure with a fixed block size of 4 to ensure balanced group allocation. Group assignments will be placed in sealed, opaque envelopes held by the PI. Each envelope will include the participant's information and intervention assignment.

This trial will use a single-masked design. Participants will be aware of their assigned intervention (digital ocular massage, EyePeace ocular massage, or no ocular massage). However, the data analysis team and clinical assessment staff will remain masked to group allocation. All clinical assessments will be performed by a masked examiner who is not involved in data collection or group allocation. The conditions and protocols for disclosure of group information, if required, will be established and overseen by the PI.

Data Collection and Management

The clinical assessment staff will collect data on IOP (mm Hg), best corrected visual acuity (logarithm of the minimum angle of resolution), SC area (μm^2), SC diameter (μm), TM width (μm), and TM thickness (μm), as well as responses to the Quality of Vision Questionnaire. Cases that are discontinued will be retained for subanalysis; these cases will be particularly useful for evaluating outcomes such as AEs and adherence. All participants will undergo a comprehensive ophthalmic examination at baseline and after the intervention. When participants enroll for the study, trained clinical researchers and outpatient physicians participating in the trial will enter participant information into electronic case report forms. All study data will be entered into a specialized management program and stored securely in a database to ensure confidentiality. Designated statisticians from the TMUEH research team will analyze the anonymized data according to the trial plan. Participants will be informed in advance and encouraged to actively cooperate with study procedures to minimize attrition. Although AEs are unlikely in this study, the intervention will be suspended upon detection of any ocular symptoms of concern during slit lamp examination. Treatment for the AE, along with regular telephone and SMS text message updates, will be provided to those affected. All AEs are documented and reported to the trial steering committee. Participants unable to complete subsequent follow-up will be withdrawn from the study. Those who experience intolerable side effects will receive treatment and be removed immediately from the study.

Statistical Methods

Following the guidelines for statistical analysis plans in clinical trials [18,29], all statistical analyses will be conducted using SPSS software (version 25.0; IBM Corp). The significance level is set at 5% (2-tailed), with a 95% CI. The normality of data distribution will be assessed using the Shapiro-Wilk test. Data will be compared using 2-way repeated measures analysis of variance. If the hypothesis is unmet ($P>.05$), the Greenhouse-Geisser correction will be applied. Multiple comparisons will be made between groups at different time points using the Bonferroni-Holm correction. If the Shapiro-Wilk test indicates that the data are not normally distributed ($P<.05$), nonparametric tests will be used: paired comparisons will use the Wilcoxon signed-rank test, and between-group comparisons will use the Kruskal-Wallis H test. Statistical significance will be considered at $P<.05$ for 2-tailed tests unless otherwise specified. To ensure participant safety, the distribution of AEs will be described for both digital ocular massage and EyePeace ocular massage groups. The incidence rates of AEs were compared using the chi-square test or the Fisher exact test, as appropriate. An intention-to-treat analysis will be the primary analytic approach. All participants will be included in the analysis based on their original allocation, regardless of protocol adherence. For missing data, multiple imputation will be performed using the fully conditional specification method, assuming data are missing at random. A sensitivity analysis using a per-protocol approach will also be conducted to compare results. Given the short duration of this study, an interim analysis will not be performed.

Ethical Considerations

This study will adhere to the ethical principles outlined in the Declaration of Helsinki and has been approved by the TMUEH Institutional Review Board (2024KY-56). Informed consent will be obtained by clinical physicians who have received specialized training. Written informed consent will be obtained from all participants before their participation. All information and clinical data, including participants' personal medical histories, will remain confidential.

Monitoring

The trial steering committee, comprising project leaders and key decision-makers, is responsible for overseeing and supporting the study. This committee will provide strategic guidance throughout all phases of the research and will be responsible for preparing and submitting the final study report.

Given the low likelihood of AEs in healthy participants, a data monitoring committee will not be established. Data collection and management will be conducted using a Microsoft Excel (version 16) database, which will be regularly updated in accordance with the study protocol.

Any modifications to the inclusion criteria or test procedures will be submitted to the TMUEH Medical Ethics Review Committee for review and approval before implementation.

Classification of AEs

Any unexpected symptoms or signs that may pose a risk to the physical or mental well-being of participants will be classified

XSL•FO
RenderX

as AEs. Localized AEs may include intolerable dry eye symptoms, keratitis, severe corneal ulcers, conjunctivitis, eye irritation, and blurred vision. Mild AEs, such as temporary dry eyes or foreign body sensation, which do not affect daily activities, will be managed by suspending the intervention, administering artificial tears, and re-examining participants after 30 minutes. Moderate AEs, such as conjunctivitis or slightly elevated IOP, will be managed by suspending the treatment, referring to the PI, providing targeted treatment, and following up until the symptoms resolve. In the event of a serious AE, the PI and the TMUEH Medical Ethics Review Committee will be promptly notified, the trial will be immediately suspended, and appropriate medical intervention will be provided to ensure participant safety and minimize further risk.

Auditing Trial Conduct

This study will be subject to weekly evaluations and assessments conducted by independent supervisors to ensure compliance with the research protocol and ethical standards.

Dissemination Plans

The research findings will be disseminated to relevant stakeholders, including participants, their families, physicians, advisory committees, and medical boards. In addition, the results will be shared through high-impact medical journals and presentations at international medical conferences to ensure broad academic and clinical engagement.

Results

As of July 15, 2025, the study recruited 50 participants, and all provided written informed consent. The initial results are anticipated to be reported and published after the completion of data collection and analysis, with publication expected in 2025-2026.

Discussion

Summary

Eye massage is commonly used for relaxation, relief of eye strain, and the management of dry eye syndrome [19,30]. However, improper technique or excessive pressure can lead to ocular complications [20,31]. A case study by Yadgari et al [32] reported the formation of an intracorneal cleft after ocular massage in a patient who had undergone Ahmed valve implantation for glaucoma, highlighting the risk of corneal structural damage when excessive external pressure is applied to the eye. Riede-Pult et al [33] investigated the short-term effects of eyelid massage on corneal shape using topographic analysis. Although the study found no significant short-term changes in corneal curvature or visual acuity, it did not assess potential long-term consequences, suggesting that while brief eyelid massage may be safe, repeated or forceful manipulation could carry risks. A systematic review by Yin et al [34] explored the adverse effects of massage therapy in pain-related conditions. Reported complications included soft tissue trauma, nerve injury, and vascular complications, which could theoretically apply to forceful eye massage as well. The possibility of pressure-induced damage to the delicate ocular and periocular structures warrants caution. Lee et al [30] conducted a randomized controlled trial comparing a thermal eye massager with artificial tears to treat dry eye syndrome. While the thermal massage was found to be effective and generally safe, some patients experienced mild and transient adverse effects such as irritation or temporary discomfort. A broader review assessed the safety of massage therapy and highlighted rare but serious adverse effects such as vascular damage, nerve compression, and eye trauma [24,35]. While these events are uncommon, they emphasize the importance of appropriate technique and caution when applying pressure around the eyes [25,26,36,37]. This study will enroll only healthy adult participants with IOP within the target range of 10 to 21 mm Hg and no ocular comorbidities to minimize confounding factors and ensure a controlled baseline for evaluating the isolated effects of ocular massage on the SC and IOP. However, we acknowledge that this limits the generalizability of the findings, particularly to populations with glaucoma or postsurgical eyes, where ocular massage may be used therapeutically. Future studies in these clinical populations are warranted to validate the safety and efficacy of these interventions under real-world conditions. In this randomized controlled trial, participants were randomly assigned to 1 of 3 groups (digital ocular massage, EyePeace ocular massage, or no ocular massage) and underwent baseline and 10-minute postintervention assessments using AS-OCT and IOP measurements. These short-term findings could inform novel, nonpharmacological approaches for glaucoma risk mitigation. However, the short follow-up window limits conclusions regarding durability. Future work should incorporate longer follow-up periods, include patient populations with glaucoma or ocular hypertension, and explore the correlation between morphological changes and sustained IOP regulation. Safety assessments and comparative efficacy in real-world clinical settings remain crucial.

A more extended follow-up period is needed to determine the long-term effects of digital ocular massage and EyePeace ocular massage on IOP and anterior segment anatomical changes. It is also unclear whether the changes in IOP, anterior segment anatomy, and vision are sustained for long periods and can be considered long-term therapy. This study is a single-masked clinical trial; therefore, potential bias may still be present, despite the researcher's objective evaluation of data. In addition, as this study measures outcomes only 10 minutes after the intervention, it may not capture the longer-term persistence or fluctuation of the IOP-lowering effect. Future studies with extended follow-up periods (eg, 30 minutes, 1 hour, or longer) are warranted to assess the duration and stability of the responses.

Future Directions and Conclusions

Although gentle eye massage can offer therapeutic benefits, excessive or improper pressure may lead to corneal damage, intraocular structural issues, soft tissue trauma, and vascular complications. Patients with preexisting ocular conditions (eg, glaucoma, keratoconus, or recent eye surgery) should exercise caution and seek medical advice before eye massage. More research is needed to fully understand the long-term implications of repeated ocular massage on eye health, including the durability of IOP reduction, cumulative tissue effects, and

XSL•FO
RenderX

optimal massage parameters. Future clinical trials involving diverse patient populations and extended follow-up periods will be essential to determine the safety and efficacy of ocular massage as a nonpharmacological therapeutic strategy.

Acknowledgments

The authors express their gratitude to all individuals who participated in and supported this experiment. This study has been funded by the National Natural Science Foundation of China (82160202), Tianjin Health Science and Technology Project: Special Project for High-Level Talents (TJWJ2024RC005), and Tianjin Education Commission Research Program Project: Key Project in Natural Sciences (2023ZD016).

Data Availability

The datasets generated and analyzed during this study will be available from the corresponding author on reasonable request.

Authors' Contributions

Formal analysis, methodology, software, validation, visualization, writing – original draft, writing – review and editing: MZ
Data curation, investigation, methodology, validation, visualization, writing – review and editing: LQ
Formal analysis, methodology, writing – review and editing: LT
Methodology, validation, writing – review and editing: YL
Data curation, formal analysis, methodology, validation, writing – review and editing: ZZ
Data curation, formal analysis, methodology, project administration, validation, visualization, writing – review and editing: QZ
Formal analysis, methodology, visualization, writing – review and editing: SW
Conceptualization, data curation, formal analysis, methodology, validation, visualization, writing – original draft, writing – review and editing: EEP
Conceptualization, funding acquisition, investigation, project administration, resources, supervision, validation, writing – original draft, writing – review and editing: XR

Conflicts of Interest

None declared.

Multimedia Appendix 1
Modified Quality of Vision Questionnaire.
[DOCX File , 24 KB - resprot_v15i1e78864_app1.docx]

References

1. Wieland M, Spaeth GL. Use of digital compression following glaucoma surgery. Ophthalmic Surg 1988 May;19(5):350-352. [Medline: 3399263]
2. Azuara-Blanco A, Katz LJ. Dysfunctional filtering blebs. Surv Ophthalmol 1998;43(2):93-126. [doi: 10.1016/s0039-6257(98)00025-3] [Medline: 9763136]
3. Agnifili L, Sacchi M, Figus M, Posarelli C, Lizzio RA, Nucci P, et al. Preparing the ocular surface for glaucoma filtration surgery: an unmet clinical need. Acta Ophthalmol 2022 Nov;100(7):740-751. [doi: 10.1111/aos.15098] [Medline: 35088941]
4. Fellman RL, Grover DS. What are surgical pearls for performing trabeculectomy? In: Gedde S, editor. Curbside Consultation in Glaucoma: 49 Clinical Questions. 2nd edition. Boca Raton, FL: CRC Press; 2015:1-6.
5. Wu TH, Lau HK, Lai CH, Wong RW, Wong AK, Lam AK. Effect of digital ocular massage on intraocular pressure and Schlemm's canal dimensions. Sci Rep 2024 Mar 13;14(1):6112 [FREE Full text] [doi: 10.1038/s41598-024-56748-1] [Medline: 38480777]
6. Nuzzi R, Scalabrin S, Becco A. Reduction of intraocular pressure spikes due to intravitreal bevacizumab injections by scleral indentation with cotton swab or digital ocular massage: innovative techniques compared. Clin Ophthalmol 2020 Aug 27;14:2533-2541. [doi: 10.2147/OPTH.S263474] [Medline: 32921980]
7. Duvall-Young J. Emergency, Acute and Rapid Access Ophthalmology: Practical, Clinical and Managerial Aspects. Cham, Switzerland: Springer; Oct 03, 2018.
8. Yu B, Wang K, Zhang X, Xing X. Biometric indicators of anterior segment parameters before and after laser peripheral iridotomy by swept-source optical coherent tomography. BMC Ophthalmol 2022 May 16;22(1):222. [doi: 10.1186/s12886-022-02448-1] [Medline: 35578256]
9. Korenfeld MS, Dueker DK. Review of external ocular compression: clinical applications of the ocular pressure estimator. Clin Ophthalmol 2016 Feb 23;10:343-357. [doi: 10.2147/OPTH.S92957] [Medline: 26966349]

XSL•FO
RenderX

10. Wang MT, Feng J, Wong J, Turnbull PR, Craig JP. Randomised trial of the clinical utility of an eyelid massage device for the management of meibomian gland dysfunction. Cont Lens Anterior Eye 2019 Dec;42(6):620-624. [doi: 10.1016/j.clae.2019.07.008] [Medline: 31358441]
11. Chen J, Yu S, Qin G, Moutari S, Moore JE, Xu L, et al. Immediate sequential changes in the tear film lipid layer following eyelid massage in dry eye syndrome: a comparative control study. Heliyon 2024 Aug 20;10(17):e36590. [doi: 10.1016/j.heliyon.2024.e36590] [Medline: 39286233]
12. Costagliola C, dell'Omo R, Agnifili L, Bartollino S, Fea AM, Uva MG, et al. How many aqueous humor outflow pathways are there? Surv Ophthalmol 2020;65(2):144-170. [doi: 10.1016/j.survophthal.2019.10.002] [Medline: 31622628]
13. Lee JY, Akiyama G, Saraswathy S, Xie X, Pan X, Hong YK, et al. Aqueous humour outflow imaging: seeing is believing. Eye (Lond) 2021 Jan;35(1):202-215. [doi: 10.1038/s41433-020-01215-0] [Medline: 33060830]
14. Mavlyutov TA, Kuhn MS, Bilal SE, De Ieso ML, Chauhan AK, Stamer WD, et al. Decreased outflow facility and Schlemm's canal defects in a mouse model of glaucoma. Exp Eye Res 2022 Dec;225:109249 [FREE Full text] [doi: 10.1016/j.exer.2022.109249] [Medline: 36152913]
15. Kagemann L, Nevins JE, Jan NJ, Wollstein G, Ishikawa H, Kagemann J, et al. Characterisation of Schlemm's canal cross-sectional area. Br J Ophthalmol 2014 Jul;98 Suppl 2(Suppl 2):ii10-ii14 [FREE Full text] [doi: 10.1136/bjophthalmol-2013-304629] [Medline: 24590558]
16. Kagemann L, Wang B, Wollstein G, Ishikawa H, Nevins JE, Nadler Z, et al. IOP elevation reduces Schlemm's canal cross-sectional area. Invest Ophthalmol Vis Sci 2014 Mar 25;55(3):1805-1809 [FREE Full text] [doi: 10.1167/iovs.13-13264] [Medline: 24526436]
17. Kagemann L, Wollstein G, Ishikawa H, Bilonick RA, Brennen PM, Folio LS, et al. Identification and assessment of Schlemm's canal by spectral-domain optical coherence tomography. Invest Ophthalmol Vis Sci 2010 Aug;51(8):4054-4059. [doi: 10.1167/iovs.09-4559] [Medline: 20237244]
18. Chung HW, Park JH, Yoo C, Kim YY. Effects of trabecular meshwork width and Schlemm's canal area on intraocular pressure reduction in glaucoma patients. Korean J Ophthalmol 2021 Aug;35(4):311-317. [doi: 10.3341/kjo.2021.0007] [Medline: 34162183]
19. Chen M, Zhang L, Xu J, Chen X, Gu Y, Ren Y, et al. Comparability of three intraocular pressure measurement: iCare pro rebound, non-contact and Goldmann applanation tonometry in different IOP group. BMC Ophthalmol 2019 Nov 14;19(1):225. [doi: 10.1186/s12886-019-1236-5] [Medline: 31726999]
20. Ng JS, Wong A. Line-by-line visual acuity scoring equivalence with letter-by-letter visual acuity scoring. Clin Exp Optom 2022 May;105(4):414-419. [doi: 10.1080/08164622.2021.1924625] [Medline: 34139959]
21. Marcellán MC, Remón L, Ávila FJ. Corneal hysteresis and intraocular pressure are altered in silicone-hydrogel soft contact lenses wearers. Int Ophthalmol 2022 Sep;42(9):2801-2809. [doi: 10.1007/s10792-022-02270-0] [Medline: 35355168]
22. Anastasopoulos E, Founti P, Topouzis F. Update on pseudoexfoliation syndrome pathogenesis and associations with intraocular pressure, glaucoma and systemic diseases. Curr Opin Ophthalmol 2015 Mar;26(2):82-89. [doi: 10.1097/ICU.0000000000000132] [Medline: 25594764]
23. Ding X, Huang L, Peng C, Xu L, Liu Y, Yang Y, et al. Evaluation of Schlemm's canal with swept-source optical coherence tomography in primary angle-closure disease. BMC Ophthalmol 2023 Jun 07;23(1):256. [doi: 10.1186/s12886-023-03001-4] [Medline: 37286943]
24. Karimi A, Crouch DJ, Razaghi R, Crawford Downs J, Acott TS, Kelley MJ, et al. Morphological and biomechanical analyses of the human healthy and glaucomatous aqueous outflow pathway: imaging-to-modeling. Comput Methods Programs Biomed 2023 Jun;236:107485. [doi: 10.1016/j.cmpb.2023.107485] [Medline: 37149973]
25. Chen L, Chen Z, Deng C, Chen W, Zhang H, Wang J. Changes to outflow structures after pilocarpine in primary open angle glaucoma compared with healthy individuals using optical coherence tomography. J Glaucoma 2023 Jul 01;32(7):593-599. [doi: 10.1097/IJG.0000000000002165] [Medline: 36548467]
26. McAlinden C, Pesudovs K, Moore JE. The development of an instrument to measure quality of vision: the Quality of Vision (QoV) questionnaire. Invest Ophthalmol Vis Sci 2010 Nov;51(11):5537-5545. [doi: 10.1167/iovs.10-5341] [Medline: 20505205]
27. Duan Y, Liang W, Guo L, Zhan H, Xia C, Ma H, et al. Effectiveness of a WeChat mini program-based intervention on promoting multiple health behavior changes among Chinese patients with cardiovascular diseases in home-based rehabilitation: randomized controlled trial. J Med Internet Res 2025 Jun 03;27:e66249. [doi: 10.2196/66249] [Medline: 40460318]
28. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 05;158(3):200-207 [FREE Full text] [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
29. Gamble C, Krishan A, Stocken D, Lewis S, Juszczak E, Doré C, et al. Guidelines for the content of statistical analysis plans in clinical trials. JAMA 2017 Dec 19;318(23):2337-2343. [doi: 10.1001/jama.2017.18556] [Medline: 29260229]
30. Lee JE, Kim NM, Yang JW, Kim SJ, Lee JS, Lee JE. A randomised controlled trial comparing a thermal massager with artificial teardrops for the treatment of dry eye. Br J Ophthalmol 2014 Jan;98(1):46-51. [doi: 10.1136/bjophthalmol-2013-303742] [Medline: 24133024]

XSL•FO
RenderX

31. van den Bosch JJ, Pennisi V, Mansouri K, Weinreb RN, Thieme H, Hoffmann MB, et al. Effect of eyelid muscle action and rubbing on telemetrically obtained intraocular pressure in patients with glaucoma with an IOP sensor implant. Br J Ophthalmol 2023 Oct;107(10):1425-1431 [FREE Full text] [doi: 10.1136/bjophthalmol-2021-320508] [Medline: 35701079]
32. Yadgari M, Noorizadeh F, vafaei A, Sobhani S. Intra corneal cleft secondary to ocular massage after Ahmed valve surgery. J ophthalmic optom sci 2018 Jan 13;2(1):44-46. [doi: 10.22037/joos.v2i1.28555]
33. Riede-Pult BH, Evans K, Pult H. Investigating the short-term effect of eyelid massage on corneal topography. Optom Vis Sci 2017 Jun;94(6):700-706. [doi: 10.1097/OPX.0000000000001076] [Medline: 28471879]
34. Yin P, Gao N, Wu J, Litscher G, Xu S. Adverse events of massage therapy in pain-related conditions: a systematic review. Evid Based Complement Alternat Med 2014;2014:480956 [FREE Full text] [doi: 10.1155/2014/480956] [Medline: 25197310]
35. Liu H, Dai Y, Yu M, Zeng J, Wang C, Tan S, et al. Effects and safety of press-needle therapy for improving visual function and eye blood circulation in patients with glaucoma with controlled intraocular pressure: study protocol for a multicenter randomized controlled trial. JMIR Res Protoc 2025 Apr 01;14:e67737. [doi: 10.2196/67737] [Medline: 40168659]
36. Ernst E. The safety of massage therapy. Rheumatology (Oxford) 2003 Sep;42(9):1101-1106. [doi: 10.1093/rheumatology/keg306] [Medline: 12777645]
37. Liao YF, Lee YC, Lin HJ, Shao YC. Acupuncture as adjuvant therapy for glaucoma: protocol for a randomized controlled trial. JMIR Res Protoc 2024 Oct 08;13:e57888. [doi: 10.2196/57888] [Medline: 39378079]

Abbreviations

AE: adverse event
AS-OCT: anterior segment optical coherence tomography
IOP: intraocular pressure
PI: principal investigator
SC: Schlemm canal
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TM: trabecular meshwork
TMUEH: Tianjin Medical University Eye Hospital

Edited by J Sarvestan; submitted 11.Jun.2025; peer-reviewed by R Razaghi, G Qin; comments to author 14.Jul.2025; revised version received 29.Jul.2025; accepted 09.Sep.2025; published 26.Feb.2026.

Please cite as:
Zhang M, Qin L, Tan L, Li Y, Zhang Q, Wang S, Zhang Z, Pazo EE, Ren X
Impact of Ocular Massage on Intraocular Pressure and Schlemm Canal Dimensions in Healthy Adults: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e78864
URL: https://www.researchprotocols.org/2026/1/e78864
doi:10.2196/78864
PMID:

©Mingxuan Zhang, Luning Qin, Liangzhang Tan, Yongtao Li, Qing Zhang, Shuhan Wang, Zhihui Zhang, Emmanuel Eric Pazo, Xinjun Ren. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Progressive Achilles Loading via Clinician e-Support (PACE): Protocol for a Dual-Site Randomized Controlled Trial

Micah A K M Wong[1], DPT; Mathew Frazier[2], DPT; Ryan Jensen[2], DPT; Timothy Fleagle[1], DPT; Barbara Van Gorp[3], DPT; Nicole Hall[1], MPT; Patrick Ten Eyck[4], PhD; Linder Wendt[4], MS; Molly Pacha[1], MS; Jessica Danielson[4], DDS; Mederic Hall[3], MD; Jason Wilken[1], DPT, PhD; Ruth L Chimenti[1], DPT, PhD

[1]Department of Physical Therapy and Rehabilitation Science, Carver College of Medicine, University of Iowa, 500 Newton Rd., Iowa City, IA, United States

[2]Department of Rehabilitation and Physical Therapy, Carl R. Darnall Army Medical Center, Fort Hood, TX, United States

[3]Department of Orthopedics and Rehabilitation, University of Iowa, Iowa City, IA, United States

[4]Institute for Clinical and Translational Science, University of Iowa, Iowa City, IA, United States

Corresponding Author:
Micah A K M Wong, DPT
Department of Physical Therapy and Rehabilitation Science, Carver College of Medicine, University of Iowa, 500 Newton Rd., Iowa City, IA, United States

Abstract

Background: Achilles tendinopathy (AT) is a common condition that limits activity participation in both active and sedentary populations. While exercise and education are well-established treatments for AT, information is lacking on the amount of intervention required and the prognostic factors associated with response to noninvasive treatments.

Objective: This noninferiority randomized controlled trial will determine the efficacy of a single-visit, physical therapist (PT)–initiated rehabilitation program compared to a multivisit, PT-guided telehealth rehabilitation program for AT and will identify early prognostic factors for individuals who experience the greatest improvement in pain and disability by 4 weeks.

Methods: In this single-blind, 2-arm, parallel-group trial, 160 individuals with AT will be randomized to either a single-visit PT-initiated rehabilitation program or a multivisit, PT-guided telehealth rehabilitation program. Primary outcomes will be the numeric rating scale for movement pain and the Victorian Institute of Sport Assessment–Achilles score for disability. Secondary outcomes will include measures of tendon health, measures within the psychosocial health domain, and military readiness.

Results: This research was funded in July 2024. Participant enrollment began in August 2025 and is expected to conclude in 2028. As of January 2026, 15 participants have been enrolled at the University of Iowa Hospitals and Clinics, meeting the recruitment goals.

Conclusions: The PACE (Progressive Achilles Loading via Clinician e-Support) trial will identify a rehabilitation program focused on education and exercise for AT that expands access to care and will identify factors that predict responses to exercise and education.

Trial Registration: ClinicalTrials.gov NCT06736795; https://clinicaltrials.gov/study/NCT06736795

International Registered Report Identifier (IRRID): DERR1-10.2196/90955

(JMIR Res Protoc 2026;15:e90955) doi:10.2196/90955

KEYWORDS

military medicine; tendinopathy; exercise therapy; telemedicine; delivery of health care

Introduction

Achilles tendinopathy (AT) can severely limit activity participation across the general [1], athletic [2], and military [3] populations. The overall prevalence of AT is between 4% and 7% in the general population [4], with a higher prevalence in middle-aged adults and athletes [5]. Exercise and education interventions have been shown to be effective in decreasing pain and improving function [6] across active and sedentary populations [4] with AT.

While care for AT typically involves an average of 4 visits (range 1-31) with a physical therapist (PT) [7], the evolving civilian health care landscape, marked by an aging population, rising costs, and rural access challenges, demands innovative approaches to care delivery [8,9]. Military health care faces similar barriers to care in high–operational tempo garrison, training, and deployed environments. Large-scale combat operations against near-peer adversaries will necessitate prolonged casualty care, of which telehealth is a component. Telehealth has emerged in the post–COVID-19 pandemic era

XSL•FO
RenderX

as a valid [10,11] and accessible [12] touchpoint for health care access. The US Military Health System made "virtual first" the goal of its restructuring plans, following the trends of many health systems [13]. Telehealth noninferiority has been previously demonstrated for AT [6], total knee replacement rehabilitation [14,15], and general musculoskeletal care [16]. Creating a treatment protocol revolving around the independent management of AT using evidence-based exercise and education interventions, while using telehealth guidance by health care professionals, will expand access to high-quality rehabilitative care.

Recent evidence has shown that overtreatment of musculoskeletal disorders does not lead to improved outcomes. Hopewell et al [17] found that a single session of education and self-management advice with a PT was noninferior to 6 sessions of in-person progressive exercise in improving pain and function among individuals with subacromial shoulder pain. Similarly, Katz et al [18] found that a self-directed home exercise plan alone was noninferior to 14 sessions of PT-guided exercise and manual therapy for pain reduction in individuals with meniscal tear–related knee pain. AT is lacking in trials to gauge the minimal effective dosage of exercise interventions. Pivoting to expand access through novel care delivery pathways is now more feasible, given the evidence that pain and function can improve with focused minimal care [19].

Clinicians currently lack evidence-based prognostic factors to differentiate patients who will benefit from rehabilitation from those who are less likely to respond [20,21]. Given that 26% to 47% of individuals with AT proceed to injections or surgery [22-24], the identification of prognostic factors could inform more efficient and personalized care. For nonresponders to rehabilitation, delays in progressing to invasive treatments prolong Achilles tendon pain, waste patient time and health care resources, and increase the total number of limited duty days as multiple treatments are tried. There is a critical need for a large, prospective study to identify factors that predict which patients with AT are most likely to respond to rehabilitation.

This trial's objective is to identify a rehabilitation program focused on education and exercise for AT that expands access to care and to identify factors that predict responses to rehabilitation. Our aims are twofold: (1) to determine the efficacy of a single-visit, PT-initiated rehabilitation program compared to a multivisit, PT-guided telehealth rehabilitation program for AT and (2) to identify early prognostic factors for individuals who experience the greatest improvement in pain and disability by 4 weeks. We hypothesize that a single-visit, PT-initiated rehabilitation program with a self-guided intervention will be as effective in improving pain and disability as a multivisit, PT-guided telehealth intervention. We also hypothesize that individuals with greater immediate (within 2 weeks) improvements in self-efficacy, pain, and disability will obtain the greatest benefits from rehabilitation.

Methods

Overview

Progressive Achilles Loading via Clinician e-Support (PACE) is a single-blind, noninferiority, 2-arm, parallel-group randomized controlled trial. Individuals with AT will be randomized to a single-visit, PT-initiated rehabilitation program or to a multivisit, PT-guided telehealth rehabilitation program. All participants will initiate care with a single synchronous in-person visit with a PT and have access to information on exercise and education materials via an online learning management system. Primary outcomes will be the numeric rating scale (NRS) for movement pain and the Victorian Institute of Sport Assessment–Achilles (VISA-A) score for disability.

We will enroll 160 individuals with AT across 2 sites: 1 military treatment facility and 1 civilian academic center. Outcomes will be assessed at 2, 4, 8, 12, 26, and 52 weeks, with the primary end point at week 4. Deviations to the protocol will be reported in real time on the Open Science Framework. This protocol is reported in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [25].

Participant Selection

Participants will be recruited and enrolled from 2 sites: the Carl R. Darnall Army Medical Center and the University of Iowa Hospitals & Clinics (UIHC). Active and reserve military personnel, veterans, and civilians will be eligible with recruitment carried out at site-specific clinics (family medicine, physical therapy, and sports medicine); on social media; through email requests; and through the distribution of brochures and flyers in community areas. A full list of the inclusion and exclusion criteria is provided in Textbox 1. To allow for the consideration of sex as a biological variable in the study findings, we will target at least 33% enrollment of female participants across both sites, with an anticipated higher recruitment rate of female participants at the UIHC to compensate for an anticipated lower rate of female recruitment at the CRDAMC (Carl R Darnall Army Medical Center, Fort Hood).

XSL•FO
RenderX

Textbox 1. Eligibility criteria for study enrollment across both sites: the Carl R. Darnall Army Medical Center and the University of Iowa Hospitals and Clinics.

Inclusion criteria

- Pain localized to the Achilles tendon (insertional or midportion)
- Achilles pain ≥3 out of 10 with a tendon-loading exercise

Exclusion criteria

- Age <18 y or >60 y
- BMI >45 kg/m^2
- Pain primarily due to a differential diagnosis, including paratendonitis, bursitis (retrocalcaneal or subcutaneous), posterior ankle impingement or os trigonum, irritation or neuropathy of the sural nerve, or plantaris tendon involvement
- Presence of a partial Achilles tendon tear or rupture on imaging, or a history of Achilles tendon rupture that was verified by surgical or conservative management
- Attendance at physical therapy for Achilles tendinopathy (AT) within the past 3 mo
- History of a steroid injection to lower-extremity tendons or fascia, extracorporeal shock wave therapy, or any injection to the Achilles tendon region within the past 3 mo
- History of fluoroquinolone use within the past 6 mo
- History of surgery or an invasive procedure for AT on the side enrolling for treatment
- Diagnosis of systemic conditions affecting the foot and ankle (eg, rheumatoid arthritis and ankylosing spondylitis), an endocrine disorder with complications (eg, uncontrolled type 1 or type 2 diabetes and diabetic peripheral neuropathy), a connective tissue disorder (eg, Marfan syndrome and Ehlers-Danlos syndrome), peripheral vascular disease, or pregnancy
- High risk for falls (four square step test >15 s)
- Refusal of randomization or inability or uninterestedness in completing virtual visits using a webcam or smartphone, or in completing quizzes and surveys

This study will be stopped prior to its completion if recruitment challenges significantly affect the ability to complete analyses of the primary outcomes. Therefore, we anticipate that it is more likely that we need to add and modify recruitment strategies rather than invoke this stopping rule.

Screening

We will follow a 3-stage screening process at the CRDAMC. First, a member of the study team will review new referrals to CRDAMC physical therapy for lower leg, ankle, and foot conditions. Second, study staff will review the past medical history of potentially eligible participants as part of the referral review process. If eligibility criteria are potentially met, a PT will contact the patient by phone to arrange a Health Insurance Portability and Accountability Act (HIPAA)–compliant video meeting to review the eligibility criteria in more detail and have the participant perform tendon-loading exercises to determine movement-evoked pain intensity localized to the Achilles tendon. The final stage of the screening process includes an in-person clinical examination to confirm the diagnosis of AT and a review of the eligibility criteria.

At the UIHC, we will use a similar 3-stage screening process. First, all recruited potential participants will complete a brief online eligibility survey. Second, if the criteria are potentially met, a PT will meet with them over a video meeting to review the eligibility criteria in more detail and have them perform tendon-loading exercises to determine movement pain intensity localized to the Achilles tendon. The final stage of the screening process includes an in-person clinical examination to confirm the diagnosis of AT and a review of the eligibility criteria.

Randomization and Concealed Allocation

Individuals who meet all eligibility criteria during the in-person visit will be randomized to either a single-visit or multivisit rehabilitation intervention. Randomization will occur at a 1:1 ratio with stratification at each site for each treatment arm using permuted random blocks. Stratification variables will be based on AT type (insertional or midportion) and symptom duration (acute <3 months or chronic >3 months), which will result in 4 strata per site. Variable block sizes of 2, 4, and 6 will be used so that randomization blocks cannot be guessed. To minimize selection bias, group assignments will only be revealed after eligibility is confirmed, all baseline measures have been collected, and the in-person treatment visit has been completed. The PTs enrolling participants will not have access to the allocation sequence, which will be concealed using a computer-based randomization system in REDCap (Research Electronic Data Capture). Unblinding of group allocation will occur in the case of serious adverse events or unanticipated problems related to the intervention, as decided by the safety officer (MH) in conjunction with the steering committee (RLC, JW, and MF).

Exercise and Education Intervention

Both arms of this trial will be using an online learning management system (TalentLMS; Epignosis) for education on the pathophysiology, self-management, and exercise progression for AT. This exercise and education content has been shown to

XSL•FO
RenderX

be equally effective in reducing pain when delivered by telehealth compared to in-person treatment for AT [26,27]. For this study, the learning units (Table 1) are either completed independently at a self-paced rate (single-visit group) or reviewed during weekly sessions with a PT (multivisit group). The learning units include information about the home exercise program and provide guidance on progression. The exercise component has been previously described by Post et al [27] and is shown in Figure 1.

Table . Education and exercise content organized into 6 successive learning units to address common questions about Achilles tendinopathy.

Unit	Unit name	Questions addressed
1	From diagnosis to recovery	• What is Achilles tendinopathy? • How is Achilles tendinopathy diagnosed? • Is stiffness normal in Achilles tendinopathy? • Can exercise reduce Achilles tendinopathy pain and stiffness? • What's the best type of exercise to start with for Achilles tendinopathy?
2	Factors linked with Achilles tendinopathy	• Why is my tendon so swollen? • Would a bone spur or calcification affect my physical therapy? • Does a Haglund's deformity cause Achilles tendinopathy pain? • Should I get imaging of my Achilles tendon? • Am I ready to progress to Phase 2 of the home exercise program?
3	Achilles tendinopathy pain: factors beyond the tendon	• What causes Achilles tendon pain? • How does mental load change your Achilles tendon pain? • How can you use pain as a guide to recovery for Achilles tendinopathy?
4	Progressive loading exercises for Achilles tendinopathy	• How does tendon loading in daily life differ from this exercise program? • How can you use load to increase tendon capacity? • What would happen if you don't treat or do exercise for Achilles tendinopathy? • How do I know when I'm ready to progress to Phase 3 of the home exercise program?
5	Benefits of physical activity	• How does physical activity help your body recover? • What is the recommended amount of exercise? • How do you stay active with Achilles tendinopathy pain?
6	Recognizing recovery	• How long does recovery from Achilles tendinopathy take? • What are common signs of recovery? • What are other treatment options? • When is surgery recommended for Achilles tendinopathy?

XSL•FO
RenderX

Figure 1. Exercise progression performed in both the single-visit and multivisit trial arms. Exercises are divided into 3 phases—isometric, heel raise, and spring—with time points, progression, frequency, dose, and intensity modifications displayed.

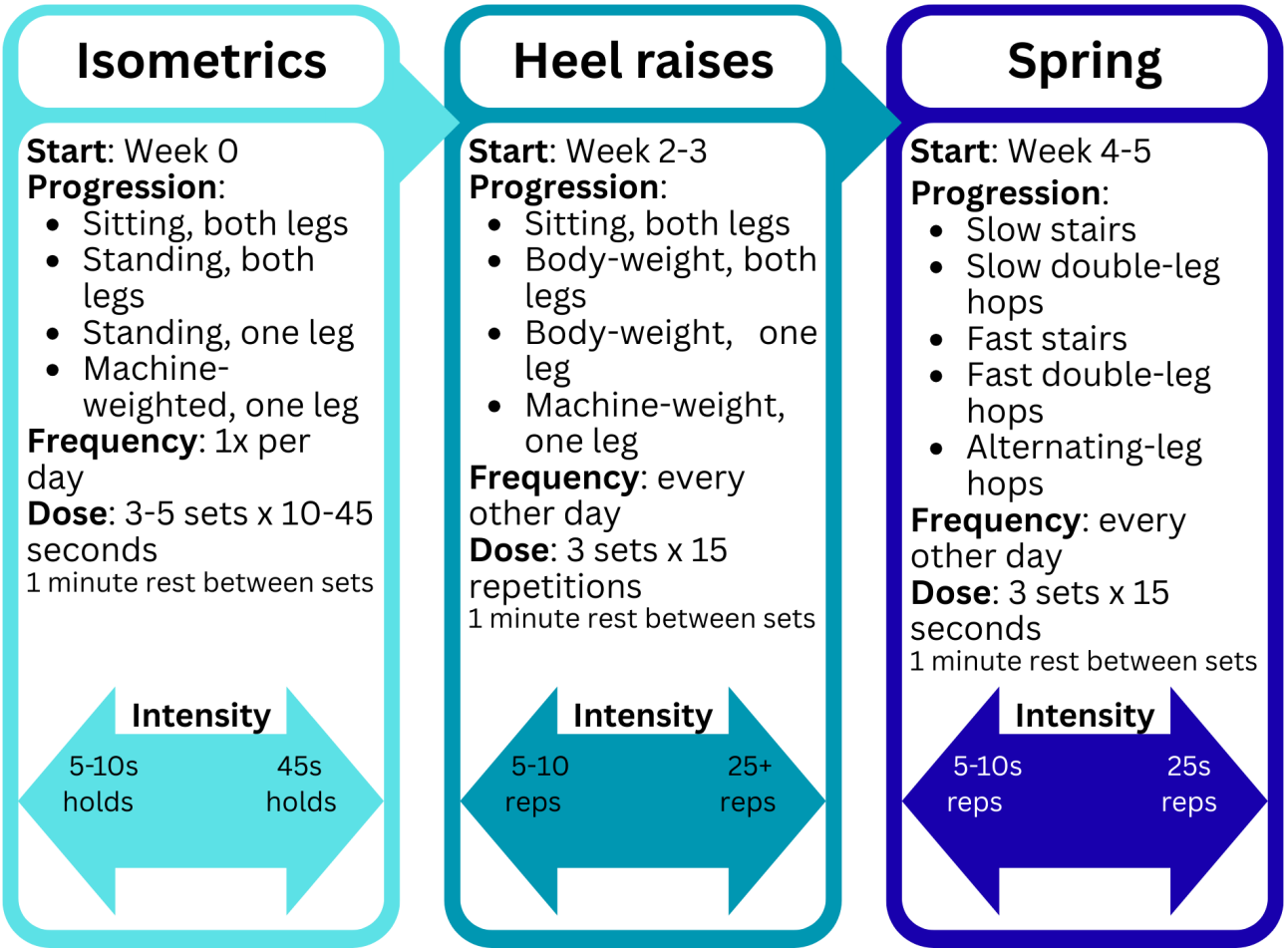

Single-Visit Group

Participants in the single-visit, PT-initiated group will undergo a self-paced rehabilitation program initiated with 1 synchronous, in-person visit with a PT. At the first visit, participants will be instructed to complete the 6 successive online learning units over an 8-week period, with flexibility to choose the timing that works best for them.

Multivisit Group

Participants in the multivisit, PT-guided rehabilitation program will have 7 one-on-one telehealth visits over 8 weeks with a PT. These 30-minute visits will include exercise participation and education. Participants will be expected to complete 1 educational unit per week and will be given an exercise log to track exercises completed at home between treatment sessions. Instruction on exercise performance and progression will be provided by a PT throughout the plan of care. The PT will review the weekly educational content with the participant at each visit.

Outcome Measures

To maintain blinding of outcome assessment, baseline measures will be gathered in person prior to randomization. A research assistant blinded to the treatment groups will capture outcomes after randomization. Participants will complete 3 survey-only evaluation sessions at 2, 12, and 26 weeks. Participants will complete 3 video evaluation sessions, including surveys, at 4 weeks (primary end point), 8 weeks, and 52 weeks. The video evaluation sessions will be used to capture pain with tendon-loading activities (heel raises and hops) in real time, assess heel raise performance (number of heel raises), and screen for adverse events. A visualization of measure administration can be seen in Table 2.

XSL•FO
RenderX

Table . Data collection schedule and format for both the self-guided and clinician-guided intervention groups over the duration of study participation.

Format of data collection	In person	Remote					
	Before the intervention	During the intervention			After the intervention		
	Week 0	Week 2	Week 4[a]	Week 8	Week 12	Week 26	Week 52
Randomization	✓						
Tendon Health Domains (Performance-based)	✓		✓	✓			✓
Tendon Health Domains (Survey-based)	✓	✓	✓	✓	✓	✓	✓
Ultrasound imaging	✓						
Clinical variables	✓		✓	✓	✓	✓	✓

[a]Week 4 indicates primary end point.

Primary Outcomes

The primary outcomes of this study are movement pain and disability (Table 3). Movement pain will be measured using the NRS. Assessing the participant's worst pain during single-limb heel raises allows for comparison of outcomes with previous clinical trials for AT [26,28-30]. Disability will be measured using the VISA-A, which assesses disability specific to AT and allows comparison to other studies, as this is the most commonly used self-reported outcome for AT [31-34].

XSL•FO
RenderX

Table . Measures and psychometric properties used in the PACE (progressive Achilles loading via clinician e-support) study[a].

Health domain	Measure	Psychometric properties
Pain	• Pain with movement during single-limb heel raises [26,28-30] and hops[b,c] • Pain at rest[c]	• MID[d] 1.1 (95% CI 0.9-1.6) [35] on a 0 - 10 NRS[e] • MID for PACE1.7 [36] on a 0 - 10 NRS
Disability	• VISA-A[b,f,g] [37,38] • Tendinopathy severity assessment[c] [39]	• MID 8.2 (SD 16.3) [40] • Cronbach α=0.808 • Reliability: ICC[h] 0.930 (95% CI 0.881-0.959) • Standard error of measurement=6.54 units [41]
Physical function and capacity	• PROMIS[f,i] physical function CAT[j] [42] • Maximum number of single-limb heel raises[c] [43] • Heel raise work[c]	• Internal consistency: person reliability=0.96 • Item reliability=0.99 [44] • Test-retest reliability: ICC 0.83 (95% CI 0.63-0.92) [45] • Interrater reliability: ICC 0.99 (95% CI 0.88-0.99) • Criterion validity: ICC 0.97 [46]
Psychosocial factors	• Tampa Scale of Kinesiophobia (17-item)[f] • PROMIS Self-Efficacy CAT[f] [42] • Optimal Screening for Prediction of Referral and Outcome-Yellow Flag tool (17-item)[f] • PROMIS Sleep Disturbance CAT[f] [47]	• Cronbach α=0.70 - 0.83 • Test-retest reliability: *r*=0.64 - 0.80 • Concurrent validity: *r*=0.33 - 0.59 [48] • Concurrent validity with legacy measures: *r*=0.59 - 0.76 [49] • Test-retest reliability: ICC 0.88; *P*<.001 [50] • Internal consistency, marginal reliabilities=0.92 • Test-retest reliability, *r*=0.80 [47]
Patient rating of condition	• Patient Acceptable State[f] [51] • Patient Global Rating of Change[f] [52]	• Concurrent validity [53] • PROMIS Pain Interference <56.0 (AUC[k] 0.940) • PROMIS Pain Intensity <48.4 (AUC 0.936) • PROMIS Physical Function >44.7 (AUC 0.883) • Test-retest reliability: ICC 0.90 • Face validity: Pearson *r*=0.72 - 0.90
Participation in life activities	• PROMIS Ability to Participate in Social Activities CAT[f] [54] • PROMIS Pain Interference CAT[f] [44]	• Reliability >0.90 • Sufficient unidimensionality root mean square error approximation=0.108 • Cronbach α=0.96 [55] • Concurrent validity with brief pain inventory: *r*=0.76 (95% CI 0.64-0.85); *P*<.01
Quality of life	• World Health Quality of Life–2[f] [56] (subset of the first 2 items on the WHOQOL-BREF[l])	• Cronbach α=0.51-0.82 for entire WHOQOL-BREF

[a]Primary outcomes are movement pain as measured with the NRS and disability measured with the VISA-A.

[b]Primary outcome.

[c]Tendon health domain (performance-based measure).

[d]MCID: minimal important difference.

[e]NRS: numeric rating scale.

[f]Tendon health domain (survey-based measure).

[g]VISA-A: Victorian Institute of Sport Assessment–Achilles.

[h]ICC: intraclass correlation coefficient.

[i]PROMIS: patient-reported outcome measure information system.

[j]CAT: computer adaptive test.

XSL•FO
RenderX

[k]AUC: area under the curve.

[l]WHOQOL-BREF: World Health Organization Quality of Life–Brief.

Secondary Outcomes

At baseline, we will capture demographics, military status, and tendon anterior-posterior diameter using ultrasound imaging [57]. In alignment with the core domain of tendon health, as defined by the International Scientific Tendinopathy Symposium Consensus [58], we will capture additional measures of pain and disability, along with measures of participation in life activities, physical function and capacity, psychological factors, patient rating of the condition, and quality of life over time (Table 3). Activity level will be captured using the Tegner scale [59], minutes per week spent doing moderate and vigorous physical activity, and questions about military readiness (eg, ability to complete a 2-mile run, wear military boots, and wear load-bearing equipment). Potential confounders that may influence the response to rehabilitation will also be captured, including the use of cointerventions or medications, self-reported exercise logs, and completion of educational units as captured in the learning management system. To assess treatment credibility, participants will complete a series of items evaluating the intervention's impact on their perspective, understanding of self-management strategies, clarity of information, and appropriateness of the approach for managing AT, rated on a 5-point Likert scale (0=not at all, 1=a little bit, 2=somewhat, 3=quite a bit, and 4=very much).

Qualitative Measures

To inform the feasibility of translating these rehabilitation strategies to a battlefield setting, we will use purposive sampling to invite participants at the CRDAMC who rotate to the field to participate in focus groups to discuss barriers and facilitators to participation in the rehabilitation programs. Participants will be invited to participate in the qualitative interviews until data saturation is reached. Semistructured interviews will be held via a videoconference platform that allows for HIPAA-compliant audio recording of the interviews. All audio recordings will be transcribed verbatim.

Safety Monitoring and Reporting

Exercise and education interventions are considered minimal risk by the University of Iowa Institutional Review Board (Biomedical IRB-01) and do not exceed the risk associated with standard clinical care; therefore, this study uses a safety officer (MH) rather than a data and safety monitoring board. The data and safety of participants will be monitored regularly by the intervention clinicians and outcome assessors, who will be contacting both the self-paced and PT-guided participants throughout the duration of study participation. All adverse events, serious adverse events, and unanticipated problems will be documented in REDCap by study team members and reviewed regularly by the safety officer. Any potential referrals for appropriate care for participants who report considerable pain or discomfort will be discussed with the safety officer and followed up in a timely manner.

Power Calculation

The trial is powered to evaluate noninferiority between the two treatment groups at 4 weeks for 2 coprimary outcomes: (1) movement pain on a 0 - 10 NRS, using a 1.1-point noninferiority margin, and (2) disability using the VISA-A, with an 8.2-point noninferiority margin. Both margins are consistent with previously published clinically meaningful thresholds [40]. Because the primary analysis will evaluate both outcomes using separate noninferiority tests, we applied a Bonferroni-adjusted, 2-sided α=.025 for each outcome. Power calculations assume comparison of mean group differences at 4 weeks using the same linear mixed-effects models planned for the primary analysis with adjustment for baseline values. A sample of 64 participants per group provides ≥80% power to conclude noninferiority for each outcome at α=.025. To account for an anticipated 20% dropout, we inflated the sample size to 80 participants per group.

Data Analysis

Baseline characteristics will be presented as counts and percentages for categorical measures. Continuous measures will be reported as means and SDs or medians and IQRs, as appropriate. Summary statistics by group will be provided for the primary and secondary outcomes and will be compared using 2-sample *t* tests, Wilcoxon rank-sum tests, or Fisher exact tests, as appropriate. Our intention-to-treat analysis will use linear mixed-effects models to assess the noninferiority of the single-visit intervention compared to the multivisit intervention, adjusting for baseline values of the primary outcomes. Additional intention-to-treat analyses will include any covariates that differed between groups at baseline. Primary outcomes will be the 4-week changes in measures of pain and function, and models will account for site using a random effect. Point estimates, CIs, and *P* values will be obtained for between-group differences in 4-week mean change compared to the margin of noninferiority. Significance for testing pain and function will be assessed at an α level of .025. We plan to report all data disaggregated by sex, AT type, and AT duration.

Secondary analyses will use linear mixed-effects models to examine predictors of improvement in pain and disability separately at 4 weeks. The same analysis will be performed at 12, 26, and 52 weeks as secondary end points. Outcomes will be the changes measured at each follow-up time point, and the models will be controlled for the corresponding baseline values. Model selection will be conducted to identify the optimal predictor set for each change outcome over time. The pool of candidate predictors includes baseline demographic and ultrasound imaging variables, as well as immediate improvement (within 2 weeks) of pain-related psychological factors, knowledge on self-management, pain, and disability. Models for all additional predictor combinations up to order 3 will be fit, and the Akaike information criterion will be calculated to compare model fits with the same outcome. Point estimates, CIs, and *P* values will be reported for predictors in models selected by the Akaike information criterion for each outcome. Significance for model predictors will be assessed at an α level

XSL•FO
RenderX

of .05. Categorization of continuous variables using clinically meaningful cut points, such as obesity categories for the continuous BMI variable, will also be explored.

Missing Data

Reasons for missing data will be recorded and compared qualitatively between participants who drop out and participants who complete the study. Participant characteristics collected prior to dropout will be compared to those of participants who complete the study. For the primary imputation strategy, if comparisons show that participant characteristics do not differ meaningfully between those with and without missing data, we will assume the data are missing at random and proceed accordingly. Under this assumption, we will use multiple imputation of the missing values for variables included in the linear mixed-effects models. Multiple imputation will be performed using SAS PROC MI and PROC MIANALYZE procedures in SAS software (version 9.4; SAS Institute Inc).

All variables in the primary analyses and, if feasible, other variables in the study database predictive of the missing values or influencing the cause of missing data will be included. For the sensitivity analyses addressing missing not at random, if assessment of reasons for missing data suggests that data may be systematically related to unobserved values (eg, inability to attend the next day assessment due to severe AT pain), sensitivity analysis will also be performed to assess the potential impact of including missing not at random data in the analysis. Methods for sensitivity analysis such as marginal delta adjustment, conditional delta adjustment, reference-based controlled imputation, and other pattern-mixture models will be considered. If needed, variables will be transformed to satisfy approximate normality before imputation and then retransformed to the original scale.

Data Management

Data will be captured primarily via REDCap, a secure online data management system developed for building and managing online surveys and databases, and compliant with HIPAA and 21 Code of Federal Regulations Part 11. Electronic records collected will be retained for a minimum of 6 years after the completion of the study.

Ethical Considerations

This study was approved by the University of Iowa Biomedical-01 Institutional Review Board (202405285) on September 19, 2024. During the screening process of this study, potentially eligible participants will be given the study information sheet and electronic consent form, with all questions answered and time provided for review. A copy of the electronically signed consent form will be obtained and kept within REDCap to document compliance, and a copy of the informed consent will be emailed to each participant. Participants will be compensated US $50 for completion of in-person consent and baseline measures, US $15 for each remote survey evaluation (up to US $45 total), US $50 for each remote video evaluation (up to US $150 total), and US $50 for focus group participation. All participant data will be deidentified and assigned a study-specific ID. Electronic records will be stored on University of Iowa network drives and will be accessible only to authorized personnel using password-protected computers.

Results

This research was funded in July 2024. Participant enrollment began in August 2025 and is expected to conclude in 2028. As of January 2026, 15 participants have been enrolled at the UIHC, meeting the recruitment goals.

Discussion

Exercise and education have been shown to be effective for the management of AT [4,23,24,26,28,60,61], but to date, no studies have been conducted to gauge the minimal effective dosage of exercise and education interventions or to identify prognostic factors for success with noninvasive treatments in patients with AT. This randomized controlled trial will identify a rehabilitation program focused on education and exercise for AT that expands access to care and will identify factors that predict responses to rehabilitation. This will be accomplished by determining the efficacy of a single-visit, PT-initiated rehabilitation program compared to a multivisit, PT-guided rehabilitation program for AT and by identifying early prognostic factors for individuals who experience the greatest improvement in pain and disability by 4 weeks.

This trial comprises a robust sample of both midportion and insertional AT. A large heterogeneous sample (N=160) will provide adequate power to detect multiple factors that most strongly predict the success of rehabilitation for AT. Including both subtypes of AT and a range of symptom duration (acute vs chronic) will better represent the natural heterogeneity of the population with AT presenting for care and facilitate the ability to detect potential predictors. Maximizing patient heterogeneity in prognostic studies allows models to account for a wider range of characteristics and factors that may influence outcome prediction, enhancing generalizability to diverse real-world populations and clinical settings [62].

Funding

This trial is funded by the US Department of War Congressionally Directed Medical Research Programs (HT94252410536). The funder had no involvement in the conceptualization of this study or the writing of this manuscript. The views expressed in this paper are those of the authors and do not reflect the official policy or position of the Department of the Army, Department of War, or the US government. Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health (UM1TR004403). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

XSL•FO
RenderX

Data Availability

In compliance with the FAIR (findability, accessibility, interoperability, and reusability) data principles [62], data will be deposited in the University of Iowa open-access institutional repository, Iowa Research Online. The repository is open access and maintained by the University of Iowa Libraries for the preservation and sharing of the intellectual work of faculty, students, and staff.

Individual patient data will be shared within 1 year after completion of the project or upon publication of the data, whichever occurs first. We will share the primary outcome data and research resources, such as the rehabilitation handouts and videos. All data that are necessary to validate the research outcomes and publications will be preserved and shared to allow other researchers to reproduce the findings and generate additional insights that may advance the results.

Conflicts of Interest

RLC has received reimbursement for time and travel related to the presentation of research on Achilles tendinopathy at scientific conferences. All other authors have declared no other conflicts of interest.

Peer Review Report 1
Peer review report by the US Department of War, Congressionally Directed Medical Research Programs (HT94252410536)
[PDF File, 281 KB - resprot_v15i1e90955_app1.pdf]

References

1. Lagas IF, Tol JL, Weir A, et al. One fifth of patients with Achilles tendinopathy have symptoms after 10 years: a prospective cohort study. J Sports Sci 2022 Nov;40(22):2475-2483. [doi: 10.1080/02640414.2022.2163537] [Medline: 36617848]
2. Lavallée-Bourget MH, Roy-Bélanger L, García-Arrabé M, et al. Factors explaining the severity of acute Achilles tendinopathy among runners: a comprehensive cross-sectional analysis. Sports Health 2025 Nov;17(6):1315-1322. [doi: 10.1177/19417381251324929] [Medline: 40114301]
3. Sullivan CK, Janney CF, Fraser JJ. Burden and risk factors for Achilles tendinopathy in the military population, 2006-2015: a retrospective cohort study. J Athl Train 2024 Jun 1;59(6):655-660. [doi: 10.4085/1062-6050-0182.23] [Medline: 37734728]
4. Chimenti RL, Neville C, Houck J, Cuddeford T, Carreira D, Martin RL. Achilles pain, stiffness, and muscle power deficits: midportion Achilles tendinopathy revision - 2024. J Orthop Sports Phys Ther 2024 Dec;54(12):CPG1-CPG32. [doi: 10.2519/jospt.2024.0302] [Medline: 39611662]
5. Wang Y, Zhou H, Nie Z, Cui S. Prevalence of Achilles tendinopathy in physical exercise: a systematic review and meta-analysis. Sports Med Health Sci 2022 Sep;4(3):152-159. [doi: 10.1016/j.smhs.2022.03.003] [Medline: 36090915]
6. Post AA, Rio EK, Sluka KA, et al. Effect of pain education and exercise on pain and function in chronic Achilles tendinopathy: protocol for a double-blind, placebo-controlled randomized trial. JMIR Res Protoc 2020 Nov 3;9(11):e19111. [doi: 10.2196/19111] [Medline: 33141102]
7. Kenyon M, Driver P, Mallows A, et al. Characteristics of patients seeking national health service (NHS) care for Achilles tendinopathy: a service evaluation of 573 patients. Musculoskelet Sci Pract 2024 Nov;74:103156. [doi: 10.1016/j.msksp.2024.103156] [Medline: 39270530]
8. Dall TM, Gallo PD, Chakrabarti R, West T, Semilla AP, Storm MV. An aging population and growing disease burden will require a large and specialized health care workforce by 2025. Health Aff (Millwood) 2013 Nov;32(11):2013-2020. [doi: 10.1377/hlthaff.2013.0714] [Medline: 24191094]
9. Gill TK, Mittinty MM, March LM. Global, regional, and national burden of other musculoskeletal disorders, 1990-2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol 2023 Nov;5(11):e670-e682. [doi: 10.1016/S2665-9913(23)00232-1] [Medline: 37927903]
10. Andino JJ, Eyrich NW, Boxer RJ. Overview of telehealth in the United States since the COVID-19 public health emergency: a narrative review. Mhealth 2023;9(26):26. [doi: 10.21037/mhealth-23-15] [Medline: 37492124]
11. Spaulding EM, Fang M, Commodore-Mensah Y, Himmelfarb CR, Martin SS, Coresh J. Prevalence and disparities in telehealth use among US adults following the COVID-19 pandemic: national cross-sectional survey. J Med Internet Res 2024 May 10;26:e52124. [doi: 10.2196/52124] [Medline: 38728070]
12. Mahdavi S, Fekri M, Mohammadi-Sarab S, Mehmandoost M, Zarei E. The use of telemedicine in family medicine: a scoping review. BMC Health Serv Res 2025 Mar 13;25(1):376. [doi: 10.1186/s12913-025-12449-7] [Medline: 40082924]
13. Hammer R. DHA senior leader: 'virtual first' is the future of military health system. Military Health System. 2023. URL: https://health.mil/News/Dvids-Articles/2023/05/25/news445593 [accessed 2026-03-12]
14. Liu X, Yang G, Xie W, et al. Efficacy of telerehabilitation for total knee arthroplasty: a meta-analysis based on randomized controlled trials combined with a bibliometric study. J Orthop Surg Res 2024 Dec 26;19(1):874. [doi: 10.1186/s13018-024-05381-9] [Medline: 39726029]
15. Klement MR, Rondon AJ, McEntee RM, Greenky MR, Austin MS. Web-based, self-directed physical therapy after total knee arthroplasty is safe and effective for most, but not all, patients. J Arthroplasty 2019 Jul;34(7S):S178-S182. [doi: 10.1016/j.arth.2018.11.040] [Medline: 30591206]

XSL•FO
RenderX

16. Bargeri S, Castellini G, Vitale JA, et al. Effectiveness of telemedicine for musculoskeletal disorders: umbrella review. J Med Internet Res 2024 Feb 2;26:e50090. [doi: 10.2196/50090] [Medline: 38306156]
17. Hopewell S, Keene DJ, Marian IR, et al. Progressive exercise compared with best practice advice, with or without corticosteroid injection, for the treatment of patients with rotator cuff disorders (GRASP): a multicentre, pragmatic, 2 × 2 factorial, randomised controlled trial. Lancet 2021 Jul 31;398(10298):416-428. [doi: 10.1016/S0140-6736(21)00846-1] [Medline: 34265255]
18. Katz JN, Collins JE, Bisson L, et al. A randomized trial of physical therapy for meniscal tear and knee pain. N Engl J Med 2025 Oct 30;393(17):1694-1703. [doi: 10.1056/NEJMoa2503385] [Medline: 41160820]
19. Michaleff ZA, Zadro JR, Traeger AC, O'Keeffe M, Décary S. Overcoming overuse part 2: defining and quantifying health care overuse for musculoskeletal conditions. J Orthop Sports Phys Ther 2020 Nov;50(11):588-591. [doi: 10.2519/jospt.2020.0109] [Medline: 33131396]
20. Hanlon SL, Pohlig RT, Silbernagel KG. Beyond the diagnosis: using patient characteristics and domains of tendon health to identify latent subgroups of Achilles tendinopathy. J Orthop Sports Phys Ther 2021 Sep;51(9):440-448. [doi: 10.2519/jospt.2021.10271] [Medline: 34074130]
21. Chimenti RL, Frey-Law LA, Sluka KA. A mechanism-based approach to physical therapist management of pain. Phys Ther 2018 May 1;98(5):302-314. [doi: 10.1093/ptj/pzy030] [Medline: 29669091]
22. Kvist M. Achilles tendon injuries in athletes. Sports Med 1994 Sep;18(3):173-201. [doi: 10.2165/00007256-199418030-00004] [Medline: 7809555]
23. Nicholson CW, Berlet GC, Lee TH. Prediction of the success of nonoperative treatment of insertional Achilles tendinosis based on MRI. Foot Ankle Int 2007 Apr;28(4):472-477. [doi: 10.3113/FAI.2007.0472] [Medline: 17475142]
24. Maffulli N, Walley G, Sayana MK, Longo UG, Denaro V. Eccentric calf muscle training in athletic patients with Achilles tendinopathy. Disabil Rehabil 2008 Jan;30(20-22):1677-1684. [doi: 10.1080/09638280701786427]
25. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ 2025;389:e081477. [doi: 10.1136/bmj-2024-081477]
26. Chimenti RL, Post AA, Rio EK, et al. The effects of pain science education plus exercise on pain and function in chronic Achilles tendinopathy: a blinded, placebo-controlled, explanatory, randomized trial. Pain 2023 Jan 1;164(1):e47-e65. [doi: 10.1097/j.pain.0000000000002720] [Medline: 36095045]
27. Post AA, Rio EK, Sluka KA, et al. Efficacy of telehealth for movement-evoked pain in people with chronic Achilles tendinopathy: a noninferiority analysis. Phys Ther 2023 Mar 3;103(3):pzac171. [doi: 10.1093/ptj/pzac171] [Medline: 37172125]
28. Beyer R, Kongsgaard M, Hougs Kjær B, Øhlenschlæger T, Kjær M, Magnusson SP. Heavy slow resistance versus eccentric training as treatment for Achilles tendinopathy: a randomized controlled trial. Am J Sports Med 2015 Jul;43(7):1704-1711. [doi: 10.1177/0363546515584760] [Medline: 26018970]
29. Silbernagel KG, Thomeé R, Eriksson BI, Karlsson J. Continued sports activity, using a pain-monitoring model, during rehabilitation in patients with Achilles tendinopathy: a randomized controlled study. Am J Sports Med 2007 Jun;35(6):897-906. [doi: 10.1177/0363546506298279] [Medline: 17307888]
30. Rompe JD, Furia J, Maffulli N. Eccentric loading compared with shock wave treatment for chronic insertional Achilles tendinopathy. A randomized, controlled trial. J Bone Joint Surg Am 2008 Jan;90(1):52-61. [doi: 10.2106/JBJS.F.01494] [Medline: 18171957]
31. Grävare Silbernagel K, Malliaras P, de Vos RJ, et al. ICON 2020-International Scientific Tendinopathy Symposium Consensus: a systematic review of outcome measures reported in clinical trials of Achilles tendinopathy. Sports Med 2022 Mar;52(3):613-641. [doi: 10.1007/s40279-021-01588-6] [Medline: 34797533]
32. Murphy MC, Travers MJ, Chivers P, et al. Efficacy of heavy eccentric calf training for treating mid-portion Achilles tendinopathy: a systematic review and meta-analysis. Br J Sports Med 2019 Sep;53(17):1070-1077. [doi: 10.1136/bjsports-2018-099934]
33. Rhim HC, Kim MS, Choi S, Tenforde AS. Comparative efficacy and tolerability of nonsurgical therapies for the treatment of midportion Achilles tendinopathy: a systematic review with network meta-analysis. Orthop J Sports Med 2020 Jul;8(7):2325967120930567. [doi: 10.1177/2325967120930567] [Medline: 32728589]
34. van der Vlist AC, Winters M, Weir A, et al. Which treatment is most effective for patients with Achilles tendinopathy? A living systematic review with network meta-analysis of 29 randomised controlled trials. Br J Sports Med 2021 Mar;55(5):249-256. [doi: 10.1136/bjsports-2019-101872]
35. Fleagle TR, Post AA, Dailey DL, et al. Minimal clinically important change of movement pain in musculoskeletal pain conditions. J Pain 2024 Aug;25(8):104507. [doi: 10.1016/j.jpain.2024.03.003] [Medline: 38479557]
36. Farrar JT, Pritchett YL, Robinson M, Prakash A, Chappell A. The clinical importance of changes in the 0 to 10 numeric rating scale for worst, least, and average pain intensity: analyses of data from clinical trials of duloxetine in pain disorders. J Pain 2010 Feb;11(2):109-118. [doi: 10.1016/j.jpain.2009.06.007] [Medline: 19665938]
37. Robinson JM, Cook JL, Purdam C, et al. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. Br J Sports Med 2001 Oct;35(5):335-341. [doi: 10.1136/bjsm.35.5.335] [Medline: 11579069]

XSL•FO
RenderX

38. Lagas IF, van der Vlist AC, van Oosterom RF, et al. Victorian Institute of Sport Assessment-Achilles (VISA-A) questionnaire-minimal clinically important difference for active people with midportion Achilles tendinopathy: a prospective cohort study. J Orthop Sports Phys Ther 2021 Oct;51(10):510-516. [doi: 10.2519/jospt.2021.10040] [Medline: 34592827]
39. Murphy MC, Newsham-West R, Cook J, et al. TENDINopathy Severity Assessment – Achilles (TENDINS-A): development and content validity assessment of a new patient-reported outcome measure for Achilles tendinopathy. J Orthop Sports Phys Ther 2024 Jan;54(1):70-85. [doi: 10.2519/jospt.2023.11964]
40. Challoumas D, Zouvani A, Creavin K, et al. Determining minimal important differences for patient-reported outcome measures in shoulder, lateral elbow, patellar and Achilles tendinopathies using distribution-based methods. BMC Musculoskelet Disord 2023 Mar 2;24(1):158. [doi: 10.1186/s12891-023-06261-9] [Medline: 36864412]
41. Murphy MC, McCleary F, Hince D, et al. TENDINopathy Severity assessment–Achilles (TENDINS-A): evaluation of reliability and validity in accordance with COSMIN recommendations. Br J Sports Med 2024 Jun;58(12):665-673. [doi: 10.1136/bjsports-2023-107741]
42. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194. [doi: 10.1016/j.jclinepi.2010.04.011]
43. Lunsford BR, Perry J. The standing heel-rise test for ankle plantar flexion: criterion for normal. Phys Ther 1995 Aug;75(8):694-698. [doi: 10.1093/ptj/75.8.694] [Medline: 7644573]
44. Hung M, Baumhauer JF, Licari FW, Voss MW, Bounsanga J, Saltzman CL. PROMIS and FAAM minimal clinically important differences in foot and ankle orthopedics. Foot Ankle Int 2019 Jan;40(1):65-73. [doi: 10.1177/1071100718800304] [Medline: 30282469]
45. Chimenti RL, Hall MM, Dilger CP, Merriwether EN, Wilken JM, Sluka KA. Local anesthetic injection resolves movement pain, motor dysfunction, and pain catastrophizing in individuals with chronic Achilles tendinopathy: a nonrandomized clinical trial. J Orthop Sports Phys Ther 2020 Jun;50(6):334-343. [doi: 10.2519/jospt.2020.9242] [Medline: 32349638]
46. Dao M, Mosby H, Westphalen E, et al. Reliability and validity of two-dimensional motion capture to assess ankle dorsiflexion motion and heel raise work. Phys Ther Sport 2023 Jul;62(10–6):10-16. [doi: 10.1016/j.ptsp.2023.05.006] [Medline: 37300968]
47. Carlozzi NE, Ianni PA, Tulsky DS, et al. Understanding health-related quality of life in caregivers of civilians and service members/veterans with traumatic brain injury: establishing the reliability and validity of PROMIS fatigue and sleep disturbance item banks. Arch Phys Med Rehabil 2019 Apr;100(4S):S102-S109. [doi: 10.1016/j.apmr.2018.05.020] [Medline: 29932884]
48. Swinkels-Meewisse EJCM, Swinkels RAHM, Verbeek ALM, Vlaeyen JWS, Oostendorp RAB. Psychometric properties of the Tampa Scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. Man Ther 2003 Feb;8(1):29-36. [doi: 10.1054/math.2002.0484] [Medline: 12586559]
49. Gruber-Baldini AL, Velozo C, Romero S, Shulman LM. Validation of the PROMIS® measures of self-efficacy for managing chronic conditions. Qual Life Res 2017 Jul;26(7):1915-1924. [doi: 10.1007/s11136-017-1527-3] [Medline: 28239781]
50. Butera KA, George SZ, Lentz TA. Psychometric Evaluation of the Optimal Screening for Prediction of Referral and Outcome Yellow Flag (OSPRO-YF) tool: factor structure, reliability, and validity. J Pain 2020;21(5-6):557-569. [doi: 10.1016/j.jpain.2019.09.003] [Medline: 31541719]
51. Tashjian RZ, Deloach J, Porucznik CA, Powell AP. Minimal clinically important differences (MCID) and patient acceptable symptomatic state (PASS) for visual analog scales (VAS) measuring pain in patients treated for rotator cuff disease. J Shoulder Elbow Surg 2009;18(6):927-932. [doi: 10.1016/j.jse.2009.03.021] [Medline: 19535272]
52. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther 2009;17(3):163-170. [doi: 10.1179/jmt.2009.17.3.163] [Medline: 20046623]
53. Shaffrey I, Nguyen J, Conti M, et al. Defining the Patient Acceptable Symptom State (PASS) for PROMIS after total ankle replacement. J Bone Joint Surg Am 2024 Aug 7;106(15):1404-1410. [doi: 10.2106/JBJS.23.01133] [Medline: 38809961]
54. Terwee CB, Crins MHP, Boers M, de Vet HCW, Roorda LD. Validation of two PROMIS item banks for measuring social participation in the Dutch general population. Qual Life Res 2019 Jan;28(1):211-220. [doi: 10.1007/s11136-018-1995-0] [Medline: 30203302]
55. Hadlandsmyth K, Dindo LN, St Marie BJ, et al. Patient-Reported Outcomes Measurement Information System (PROMIS) instruments: reliability and validity in veterans following orthopedic surgery. Eval Health Prof 2020 Dec;43(4):207-212. [doi: 10.1177/0163278719856406] [Medline: 31195828]
56. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
57. Thoirs KA, Childs J. Are ultrasound measurements of Achilles tendon size reliable? A systematic review of rater reliability. Ultrasound Med Biol 2018 Dec;44(12):2476-2491. [doi: 10.1016/j.ultrasmedbio.2018.07.011] [Medline: 30154035]
58. Vicenzino B, de Vos RJ, Alfredson H, et al. ICON 2019—International Scientific Tendinopathy Symposium Consensus: There are nine CORE health-related DOMAINS for tendinopathy (CORE DOMAINS): Delphi study of healthcare professionals and patients. Br J Sports Med 2020 Apr;54(8):444-451. [doi: 10.1136/bjsports-2019-100894]

XSL•FO
RenderX

59. Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries. Clin Orthop Relat Res 1985 Sep;Sep(198):43-49. [Medline: 4028566]
60. Malliaras P. Physiotherapy management of Achilles tendinopathy. J Physiother 2022 Oct;68(4):221-237. [doi: 10.1016/j.jphys.2022.09.010] [Medline: 36274038]
61. Silbernagel KG, Brorsson A, Lundberg M. The majority of patients with Achilles tendinopathy recover fully when treated with exercise alone: a 5-year follow-up. Am J Sports Med 2011 Mar;39(3):607-613. [doi: 10.1177/0363546510384789] [Medline: 21084657]
62. Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3:160018. [doi: 10.1038/sdata.2016.18] [Medline: 26978244]

Abbreviations

AT: Achilles tendinopathy
CRDAMC: Carl R. Darnall Army Medical Center, Fort Hood
HIPAA: Health Insurance Portability and Accountability Act
NRS: numeric rating scale
PACE: Progressive Achilles Loading via Clinician e-Support
PT: physical therapist
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
UIHC: University of Iowa Hospitals and Clinics
VISA-A: Victorian Institute of Sport Assessment–Achilles

Edited by J Sarvestan; submitted 14.Jan.2026; this is a non–peer-reviewed article;accepted 25.Feb.2026; published 07.Apr.2026.

Please cite as:
Wong MAKM, Frazier M, Jensen R, Fleagle T, Van Gorp B, Hall N, Ten Eyck P, Wendt L, Pacha M, Danielson J, Hall M, Wilken J, Chimenti RL
Progressive Achilles Loading via Clinician e-Support (PACE): Protocol for a Dual-Site Randomized Controlled Trial
JMIR Res Protoc 2026;15:e90955
URL: https://www.researchprotocols.org/2026/1/e90955
doi:10.2196/90955

© Micah A K M Wong, Mathew Frazier, Ryan Jensen, Timothy Fleagle, Barbara Van Gorp, Nicole Hall, Patrick Ten Eyck, Linder Wendt, Molly Pacha, Jessica Danielson, Mederic Hall, Jason Wilken, Ruth L Chimenti. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 7.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility of Vigorous Extended Reality Tele-Exergaming for Cardiometabolic Health in Youth With Mobility Disabilities: Protocol for a Case Series Study

Byron Lai[1*], MS, PhD; Maggie Logan[1,2*], BS; Raven Young[1*], BS; Ashley Wright[1*], MS; Jordyn Terrell[3*], BS; Larsen Bright[1*]; Drew Davis[1*], MD; Christen J Mendonca[4*], PhD

[1]Division of Pediatric Rehabilitation Medicine, Department of Pediatrics, University of Alabama at Birmingham, 1720 2nd Ave South, Birmingham, AL, United States

[2]Department of Occupational Therapy, University of Alabama at Birmingham, Birmingham, AL, United States

[3]Department of Physical Therapy, University of Alabama at Birmingham, Birmingham, AL, United States

[4]Research Collaborative, School of Health Professions, University of Alabama at Birmingham, Birmingham, AL, United States

[*]all authors contributed equally

Corresponding Author:
Byron Lai, MS, PhD
Division of Pediatric Rehabilitation Medicine, Department of Pediatrics, University of Alabama at Birmingham, 1720 2nd Ave South, Birmingham, AL, United States

Abstract

Background: Young people with mobility disabilities have limited options to maintain their cardiometabolic health and cardiorespiratory fitness. Active video gaming using extended reality head-mounted displays is becoming increasingly common for promoting serious exergaming. However, there is a need to identify dosing protocols that can potentially lead to meaningful improvements in cardiometabolic health and cardiorespiratory fitness.

Objective: This feasibility study aims to explore potential benefits of vigorous-intensity extended reality exergaming, conducted at home with telemonitoring for body composition and cardiometabolic health, in 4 young people with cerebral palsy and overweight or obesity. The secondary aim is to assess the effects of the program on cardiorespiratory fitness. The tertiary aim is to describe the safety and acceptability of the intervention.

Methods: This case study is a phase 1 feasibility trial with a pretest-to-posttest design including 4 participants. Young people with cerebral palsy and overweight or obesity (aged 13-24 years) will be purposively selected to participate based on 2 mobility categories (ambulatory: n=2; nonambulatory: n=2). The intervention prescription will include 240 minutes per week of vigorous-intensity exercise at home for a total of 6 weeks, with telemonitoring of exercise data that will be supplemented with weekly coaching calls. Participants will exercise using a head-mounted display to control an immersive exergame. Caregivers will agree to manage a play schedule and assess their children's safety during play. Body composition will be measured using dual-energy x-ray absorptiometry at the pre- and postintervention time points (weeks 0 and 7, respectively). Blood-related health (fasting insulin, lipids, high-sensitivity C-reactive protein, and glycated hemoglobin) will be measured via a blood spot test, and blood pressure will be measured via a sphygmomanometer. Cardiorespiratory fitness (peak oxygen consumption) will be measured using a portable metabolic cart (COSMED K5) during a graded exercise test on an arm ergometer at the pre- and postintervention time points. Quantitative process metrics and qualitative feedback will be used to assess feasibility. Changes in outcomes over time will be descriptively analyzed.

Results: Recruitment began in October 2025. All data are anticipated to be collected by December 2025. Full results are anticipated to be analyzed and submitted for publication by April 2026.

Conclusions: This feasibility study tests an accessible and intensive program, which leverages high-intensity exercise gaming with telehealth procedures for children with cerebral palsy. Study findings will inform a pilot efficacy trial to improve cardiometabolic health and fitness among young people with mobility disabilities.

Trial Registration: ClinicalTrials.gov NCT07155213; https://clinicaltrials.gov/study/NCT07155213

International Registered Report Identifier (IRRID): PRR1-10.2196/85246

(JMIR Res Protoc 2026;15:e85246) doi:10.2196/85246

KEYWORDS

active video gaming; exercise; disability; physical activity; exergaming; extended reality

XSL•FO
RenderX

Introduction

Physical Activity, Disability, and Early Disease Risk

As children and adolescents with cerebral palsy (CP) transition into adulthood, their risk of cardiometabolic disease increases exponentially compared to their peers without disabilities [1-5]. This is thought to be a result of living a physically inactive lifestyle [6-9] owing to a lack of accessible exercise options and supportive environments [10,11]. While health professionals generally recommend more than 150 minutes of moderate-intensity physical activity per week for adults and more than 60 minutes of moderate to vigorous physical activity per day for children to improve and manage cardiometabolic health, most young people with physical disabilities are not meeting these guidelines [11-15].

Limited Options for Exercise

For people with mobility disabilities, there is substantial evidence supporting the improvement of cardiorespiratory fitness, muscle strength and function, and gross motor skills through exercise [11]. However, there are limited studies demonstrating the positive effects of exercise on cardiometabolic health, particularly among children and adolescents with various disabilities [11]. To achieve improvements in cardiometabolic health, individuals need to engage in moderate to vigorous exercise regularly over a minimum of 1 to 3 months [16,17]. CP is characterized by disorders of movement and posture due to disturbances of a fetal or developing brain, which can include a variety of secondary complications such as epilepsy, pain, or musculoskeletal issues [18]. Therefore, individuals with CP typically encounter difficulty in long-duration participation in common modes of aerobic exercise, such as walking, running, and cycling.

Extended Reality Exergaming

Exergaming is an exercise strategy that blends video games with serious exercise programming to improve health. Extended reality (XR) includes virtual, augmented, and mixed reality applications. For young people with various disabilities, XR has been used in exercise research to enhance motor and executive functions, physical fitness, movement quality, spatial orientation, and mobility, as well as to manage pain [19-24]. XR using a head-mounted display (HMD) is a common and commercially available method of exergaming that enables the user to engage with simulated environments by moving their body to play. HMDs can be used in both sitting and standing positions and have been found to elicit moderate- to vigorous-intensity exercise among children and youth with various disabilities [19,23,25].

Telehealth and XR Exergaming

Telehealth programs include virtual communication and monitoring to provide a sense of accountability, which fosters strong adherence to home-based interventions [26]. Combining telehealth strategies with home-based exergaming has the potential to engage children and adolescents with mobility disabilities in high doses of exercise that are needed to improve cardiometabolic health. Additional advantages of using telehealth strategies with exergaming could include increased social support, cost-effectiveness, access to services, and reduced trainer burden [27]. Given that HMDs can be obtained at any major retailer, telehealth exergaming could address a critical limitation in exercise research and disability: confirmatory sample sizes to achieve generalizable study findings [11,23]. Although studies have examined the effects of play using XR HMDs to elicit moderate-intensity exercise [28], to the best of our knowledge, no published study has explored the effects of a high-volume and vigorous-intensity exercise program on objective biomarkers of cardiometabolic health among youth with CP. Therefore, this study has three purposes:

1. Describe the feasibility of high-volume, vigorous-intensity XR exergaming in 4 youth with CP who are overweight or obese through metrics related to participant safety and adherence to the exercise prescription
2. Explore the potential benefits of high-volume, vigorous-intensity XR exergaming for 6 weeks for cardiometabolic health and cardiorespiratory fitness in 4 youth with mobility disabilities who are overweight or obese
3. Describe the acceptability of the intervention through participants' perceptions of completing the program via qualitative semistructured interviews

Methods

Study Design and Overview

This case study of feasibility uses a pretest-to-posttest trial design with a single group. The intervention will be conducted remotely for 6 weeks. Data collection will take place on-site at a university research laboratory in intervention week 0 and week 7. This study will include a separate (ie, not mixed methods) descriptive analysis of both quantitative and qualitative data.

Participants and Recruitment

Recruitment will take place through word of mouth and patient referrals from the Division of Pediatric Rehabilitation Medicine at Children's of Alabama. The study will purposively select a total of 4 youth with CP from the following 2 mobility groups: 2 who are ambulatory (Gross Motor Function Classification System [GMFCS] level I-II) and 2 who are nonambulatory (GMFCS level III-IV).

Inclusion Criteria

Inclusion criteria are as follows:

- Age of 13 to 25 years (to account for the World Health Organization definition of youth [15-25 years] and the previous minimum age of 13 years that was specified by the HMD manufacturer [now revised to 10 years])
- Overweight or obesity (based on Centers for Disease Control and Prevention BMI normative ranges for age and sex)
- Physical inactivity (<150 minutes per week of moderate exercise)
- GMFCS levels I to IV
- Wi-Fi internet access at home
- Physician clearance to participate

XSL•FO
RenderX

Exclusion Criteria

Exclusion criteria are as follows:

- Inability to communicate in English
- Complete blindness or deafness
- Propensity to photosensitive seizures
- GMFCS level V

Equipment and Monitoring

This study will use the Meta Quest 3S (Reality Labs) for active video gaming. This is a commercially available video gaming system. The system includes an HMD and 2 handheld controllers (Figure 1).

Figure 1. Meta Quest 3 head-mounted display with controllers and a Polar Sense heart rate monitor.



Heart rate data will be recorded and transmitted using a heart rate monitoring device (Polar Verity Sense) and mobile app (Polar Beat; Figure 1). The heart rate device will be placed on an armband and worn on the forearm. Participants will record their heart rate during exercise using the Polar Verity Sense, which will be uploaded to the Polar Flow for Coach cloud platform via a mobile app on the participants' phones. Additionally, participants will record their rating of perceived exertion in a provided logbook after each session using the modified Borg scale from 0 to 10 (0 indicates no effort, and 10 indicates maximal effort) [29].

Overview of Study Procedures

Participants will complete a 6-week exercise intervention with 2 data collection visits to the laboratory at the pre- (week 0) and postintervention (week 7) time points. Before arrival on a data collection visit, participants will be asked to fast for 10 to 12 hours overnight.

The first data collection visit (week 0) will include a study briefing covering the purpose, procedures, roles, risks, and benefits of the study; informed consent and assent documentation; study questionnaires; dual-energy x-ray absorptiometry (DEXA) scans; a graded exercise test; a dried blood spot test; and an intervention briefing (6-week prescription and equipment instructions).

The second data collection visit (week 7) will include study questionnaires; a DEXA scan; a graded exercise test; a dried blood spot test; and a debriefing (reviewing pre- and posttest changes in outcomes).

Home-Based Intervention

For the 6-week intervention, participants will be asked to achieve 4 days of XR exergaming per week for 60 minutes per day, totaling 240 minutes per week. This prescription aims to be consistent with a previous case study in which a child with spina bifida achieved a substantial reduction in weight from 187 minutes per week of exergaming at approximately 76% of

XSL•FO
RenderX

the maximum heart rate [30]. A duration of 240 minutes per week was chosen for this study based on preliminary (nonpublished) findings from a recently completed trial [28]. A prescription of 240 minutes per week will likely result in 187 minutes per week due to time spent in warm-up, cooldown, and in-game level selection. Two participants will play the exergame standing up, and 2 participants will play seated in a chair or their own wheelchair. This will allow us to observe whether vigorous-intensity play can be achieved while sitting or standing. This methodology is an extension of our recently completed pilot trial [28], which demonstrated that moderate-intensity exercise can be achieved while XR exergaming from either a seated or standing position in young people with CP (GMFCS levels I-IV; findings not yet published).

Participants will be asked to play the game Supernatural and perform high-intensity in-game routines to try to achieve a vigorous-intensity heart rate (>75% of the maximum heart rate [31] as determined by the Tanaka equation [32]). Supernatural is a live-service game designed to improve physical fitness through a variety of activities and programs (Figure 2). The initial setup includes accessibility options for seated gameplay and range of motion and vigorous-intensity programs and requires a monthly subscription to access its continually updated content. The game includes virtual, humanlike exercise trainers that provide verbal praise and instruction, as well as exercise routines that follow popular music from various genres (Figure 3).

Figure 2. First-person perspective of Supernatural gameplay through the Meta Quest head-mounted display.



XSL•FO
RenderX

Figure 3. Example of exercise routines and music genres in the Supernatural game.

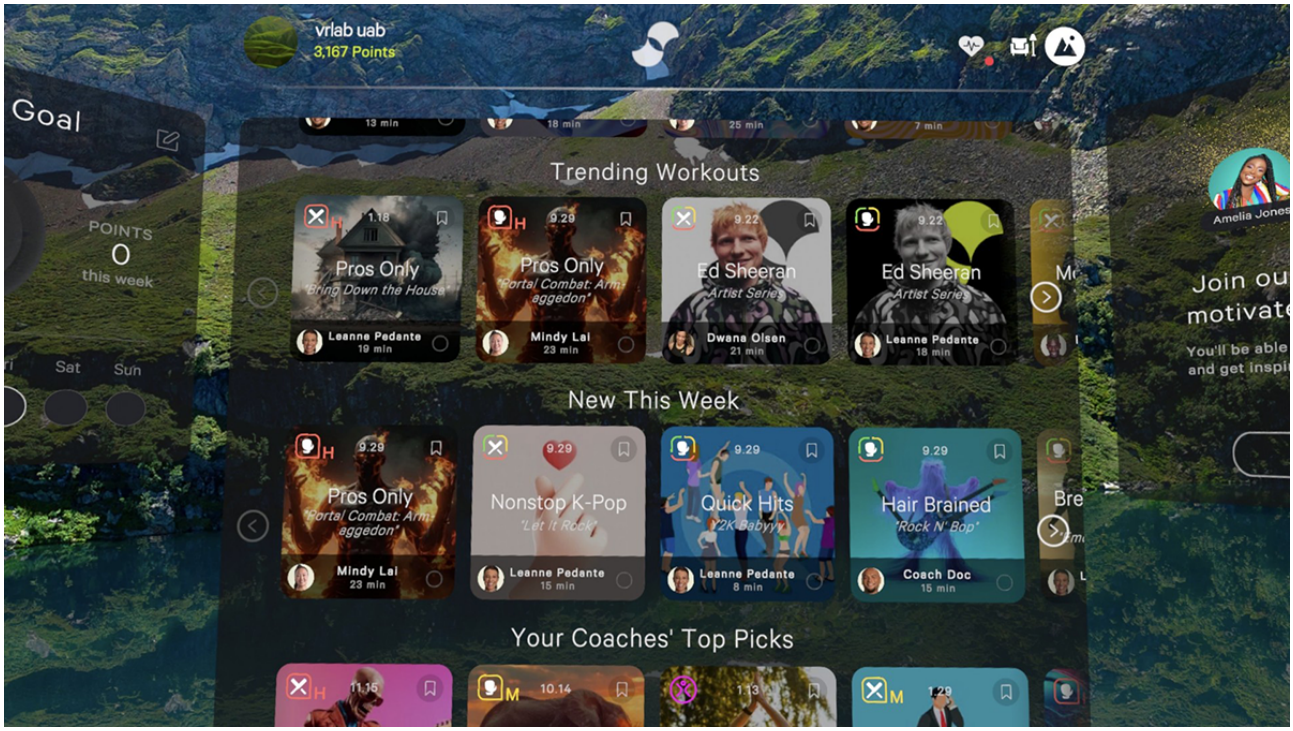

To enhance safety, participants will be instructed on how to set up a play space using the Guardian feature of the headset, a virtual boundary that will halt play should the player exceed the space. Additionally, caregivers will be certified in Red Cross cardiopulmonary resuscitation and automated external defibrillator training by a certified instructor at baseline data collection and be provided with a home automated external defibrillator (Philips HeartStart OnSite Home automated external defibrillator) to use in the case of an emergency [33].

Intervention Coaching

The intervention will incorporate low-dose behavioral physical education coaching via phone call. The purpose of the calls is to enhance adherence to the tele-exergaming program by providing a sense of accountability [26] while enhancing exercise literacy and mastery of the games to promote confidence [28]. Sessions will last 15 minutes and take place weekly. Participants will be accompanied by a caregiver as caregiver knowledge and attitudes significantly influence participation [34-36]. Specific strategies during phone calls will include introducing the importance of aerobic exercise for young people with disabilities; providing verbal praise for meeting exercise goals or smaller steps toward goals; reviewing the benefits of aerobic exercise based on systematic reviews; and solving technological, access, and logistical barriers to XR exercise. These strategies are underpinned by social cognitive theory [37]. Coaching will be delivered by research staff trained by the principal investigator (BL).

Measures and Procedures

Questionnaires

The study will use 3 questionnaires. A questionnaire will be used to assess participant characteristics (eg, age, sex, and BMI). Diet will be recorded using the Adolescent Food Habits Checklist (week 0 and week 7) to account for changes in diet. This short survey includes 23 items in which the person can respond with “true” or “false” to a variety of healthy eating habit questions. The survey is scored using a frequency count of 1 point for each “healthy” response. The final score is calculated using the following equation: number of “healthy” responses × (23/number of items completed) [38]. The Godin Leisure-Time Exercise Questionnaire (GLTEQ) will be used to examine physical activity levels outside of the program (week 0 and week 7) to account for participation in extra physical activity that could affect study outcomes. The GLTEQ [39] is a 3-item self-report questionnaire that is used to measure habitual physical activity. The GLTEQ asks participants to report the number of times that low-, moderate-, and vigorous-intensity physical activity was performed for longer than 15 minutes in a typical week. The numbers reported for moderate and vigorous exercise intensity are then multiplied by 7 and 9, respectively, and summed to obtain a score, which is referred to as the health contribution score. A health contribution score of <24 is classified as physically inactive, whereas a score of ≥24 is considered physically active [40]. There is evidence to support the GLTEQ as a valid and reliable measure of physical activity among adults with multiple sclerosis [41] and adolescents [42].

Feasibility Outcomes (Aim 1)

Components of feasibility will be evaluated through quantitative metrics, namely, intervention safety, participant adherence, and technical issues. Adherence will be measured through minutes of exergaming performed per week throughout the study, which will be objectively recorded by the Polar Verity Sense monitor. Specific measures of exercise minutes will include two categories: (1) vigorous-intensity minutes per week divided by the desired target of 187 minutes and (2) total minutes per week. An arbitrary value of >70% attendance will be used as a cutoff

XSL•FO
RenderX

for classification of “acceptable” adherence, a value of >80% will be used for classification as “good” adherence, and a value of >90% will be used to classify “excellent” adherence. Intervention safety will be measured through both adverse events and problem reports. To monitor safety, participants will be asked during coaching calls whether they experienced any minor problems (eg, joint pain or muscle soreness) or adverse events (eg, falls or injuries). Minor problems and adverse events (major problems) will be recorded and reported appropriately to the institutional review board of the University of Alabama at Birmingham. During weekly coaching calls, study staff will also record technical problems (frequency and type) that participants experience with either the XR or telehealth technology. Adverse events, problems, and technical issues will be descriptively reported, with no a priori criteria for classification as “feasible.”

Cardiometabolic Outcomes (Aim 2)

Aim 2 Primary Outcome

Body composition and weight will be measured using a DEXA scan. The primary outcomes from a whole-body DEXA scan will include body weight, total mass, and total fat mass in pounds, as well as total percentage of lean tissue and percentage of fat tissue. Participants will be instructed to lie on a DEXA scan table (Lunar iDXA; GE HealthCare) for a whole-body scan. Scans will be analyzed using the CoreScan software (GE HealthCare).

Aim 2 Secondary Outcomes

Indicators of cardiometabolic health will be measured via a dried blood spot test and sphygmomanometer. Specific outcomes will include resting blood pressure (systolic and diastolic), high-sensitivity C-reactive protein, hemoglobin A_{1C}, fasting insulin, triglycerides, and cholesterol (total, low-density lipoprotein, and high-density lipoprotein). For the dried blood spot test, a lancet finger prick will be used to obtain 9 drops of blood on a paper card, which will be allowed to dry for several hours. Dried blood spot tests will be stored in an ultralow freezer and mailed in bulk for analysis to the ZRT Laboratory [43]. The ZRT Laboratory blood spot test has demonstrated excellent validity with venous serum samples [43,44].

Cardiorespiratory Fitness Outcomes (Aim 2)

Cardiorespiratory fitness outcomes will include peak oxygen consumption measured using a portable metabolic cart (COSMED K5) during a graded exercise test on an arm ergometer, as well as maximal work output (measured in watts) and resting heart rate. The procedure for the graded exercise test will include a rest period, warm-up, graded exercise to exhaustion, and cooldown [33]. The rest period will last 2 to 3 minutes. Warm-up will include 1 to 2 minutes of cycling at a low, self-selected comfortable pace. For the graded exercise test, participants will be asked to maintain a speed of 60 revolutions per minute (or as close to 60 revolutions per minute as possible if 60 is not tolerable). Resistance will start at 0 watts and increase by 10 watts every minute until volitional fatigue or the participant meets 3 out of 5 of the following indirect criteria: a heart rate greater than 85% of the age-predicted maximum heart rate, self-reporting a score of 8 or higher on the Borg rating of perceived exertion scale from 0 to 10, demonstration of a respiratory exchange ratio of 1:1 or higher, a plateau in oxygen consumption, and subjective signs of exhaustion. Heart rate and oxygen consumption will be measured continuously. Participants will then have a slow, active cooldown until heart rate recovers.

Intervention Acceptability Through Qualitative Inquiry (Aim 3)

Intervention “acceptability” will be evaluated qualitatively. Following completion of the intervention, we will conduct an in-person semistructured interview with the participants led by the study interventionist (BL). The qualitative component of the project will use an intrinsic case study design [45,46] to gain an in-depth understanding of the experiences of a single individual who completes the program. The qualitative component of the study will be underpinned by a critical realism ontological perspective and an interpretivism epistemological perspective [47,48]. These philosophical assumptions will allow the research team to acknowledge that the participant will describe their perceived reality while also recognizing that these accounts may be shaped by interviewer-participant interaction and subsequent interpretive analysis. Consistent with interpretivist traditions, the intention of this component will be to explore how the participant makes sense of engaging in extended reality rather than to quantify code frequencies or identify a singular objective truth.

The interview guide will contain approximately 10 core questions with relevant probes. These prompts will elicit perspectives on overall impressions of the program, perceived strengths and limitations, experiences using VR for upper-extremity exercise, comparison with previous rehabilitation experiences, barriers to and facilitators of meeting the intervention prescription, preferences, and suggestions for refining the protocol in preparation for a larger trial. The interviews will be conducted by a qualitative expert in disability and exercise (BL) who has extensive experience conducting interviews in this area (>500 interviews). All interviews will be audio recorded and transcribed verbatim by a member of the research team.

Data Analyses

Descriptive statistics will be obtained for all study variables, including means, SDs, effect sizes, graphs, and 95% CIs. Quantitative outcomes across time will be graphically reported. Qualitative analysis will include thematic analysis.

Qualitative data analysis will involve 2 analysts and follow the 6-step approach by Braun and Clarke [49,50] to latent thematic analysis. First, the analysts will familiarize themselves with the transcripts through repeated readings and early reflexive noting of tone, context, and emerging ideas. Second, initial codes will be generated across the transcript to capture meaningful units related to actions, perceptions, and emotions. Third, codes will be organized into a single analytic level of themes given the small study sample size and the aim to privilege depth over breadth [51]. Fourth, candidate themes will be reviewed against the full dataset to ensure coherence and alignment with the study aims. Fifth, themes will be refined to clarify their central

XSL•FO
RenderX

concepts, and illustrative quotes will be identified based on conceptual richness rather than frequency. Sixth, findings will be synthesized for an interpretive report and supported by participant quotes. Throughout the analytic process, rigor will be strengthened through reflexive memoing and maintenance of an audit trail documenting coding and theme development. The transcripts will be checked for accuracy by a second member of the research team to ensure fidelity before analysis.

Ethical Considerations

The protocol and informed consent and assent forms were approved by the Institutional Review Board for Human Use of the University of Alabama at Birmingham (IRB-300014468) on September 3, 2025. Prospective participants will provide written informed consent or assent documentation before taking part in the study. Consent and assent forms will be completed on the baseline data collection visit. Participants will receive an electronic gift card that will be loaded with US $590 for completing the data collection visits. The rationale for this value is to account for the commitment and time required of both the child and caregiver to complete the study. This trial has been registered on ClinicalTrials.gov (NCT07155213).

Results

This study was approved by the university institutional review board on September 3, 2025. The study was initiated in October 2025, and the first participant was enrolled in October 2025. All data are anticipated to be collected by December 2025. The study end date is April 4, 2026.

Discussion

Expected Findings

This study will explore the potential impact of high-intensity arm exercise training using XR on both cardiorespiratory fitness and cardiometabolic health in 4 young people with CP. An effective arm exercise program could be extremely beneficial for this population because evidence to support that exercise can reduce body weight or other indicators of cardiometabolic health among young people with CP is limited [11,23,52]. People who have difficulty walking have few options for health-enhancing aerobic exercise [23,24,52,53].

Strengths and Limitations

Considering that this program uses telehealth communication strategies and “off-the-shelf” technology that can be purchased at any major electronics retailer, it has the potential to be carried forward in a large randomized controlled trial for young people with CP. This strength could assist in reaching larger sample sizes that could lead to more confirmatory findings from exercise trials. The average sample size of children and youth with CP for a randomized controlled trial is 27, and one of the largest trials included 159 people [52], sample sizes that are not generalizable to the diverse needs of people with CP [54]. Another strength of this study is that it includes age-appropriate games, themes, and music, potentially leading to high levels of prolonged engagement [55]. This is a crucial benefit given that improvements to cardiometabolic health require higher volumes of exercise over long periods [16,17]. Moreover, the XR exergaming can be done asynchronously at home without supervision of an exercise specialist, which eliminates the need for transportation to an on-site facility and reduces burden on research staff.

A notable limitation of this study is the small sample size. A second major limitation of this study is the requirement for on-site data collection procedures, which will hinder the scale-up potential of the program in a future study. On-site visitations are inconvenient for young people with CP and their caregivers due to a lack of nearby testing facilities and lack of time with busy school schedules and daily activities [10,56]. Thus, there is a need to develop remote data collection procedures with strong psychometric properties that can complement the remote nature of an XR telehealth program. The blood spot test can be performed remotely by shipping blood spot test kits to the home. However, to the best of our knowledge, there is no sound home-based method for measuring cardiorespiratory fitness or body composition imaging among people with mobility disabilities [20].

Conclusions

This case study examines whether intensive XR exergaming produces short-term improvements in key health indicators among a small cohort of young people with CP. If the aims of this study are supported, a larger pilot trial will be warranted.

Acknowledgments

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading, editing, and reformatting. The GenAI tool used was ChatGPT (OpenAI). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

The research reported in this publication was supported by pilot funding from the Center for Engagement in Disability Health and Rehabilitation Sciences at the University of Alabama at Birmingham. The laboratory equipment was provided by the Human Performance Laboratory in the Wellness Health and Research Facility at the University of Alabama at Birmingham.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study will be made publicly available at the completion of the study and are available from the corresponding author on reasonable request.

Authors' Contributions

All authors contributed to the final draft of the manuscript and project design. BL, JT, and ML created the initial manuscript draft. All authors contributed to the study design and procedures.

Conflicts of Interest

None declared.

Peer Review Report 1
Peer review report by the Center for Engagement in Disability Health and Rehabilitation Sciences (CEDHARS) at the University of Alabama at Birmingham.
[PDF File, 97 KB - resprot_v15i1e85246_app1.pdf]

References

1. Nolan MB, Asche SE, Barton K, et al. Cardiometabolic risk in pediatric patients with intellectual and developmental disabilities. Am J Prev Med 2025 Mar;68(3):429-436. [doi: 10.1016/j.amepre.2024.11.013] [Medline: 39615766]
2. Messiah SE, Vidot DC, Somarriba G, et al. Obesity and cardiometabolic disease risk factors among US adolescents with disabilities. World J Diabetes 2015 Feb 15;6(1):200-207. [doi: 10.4239/wjd.v6.i1.200] [Medline: 25685291]
3. Yehuda MB, Barak S, Hutzler Y, et al. Cardiovascular risk profiles clusters among children and adolescents with disabilities. BMC Public Health 2023 May 16;23(1):896. [doi: 10.1186/s12889-023-15796-4] [Medline: 37189074]
4. Rimmer JH, Yamaki K, Lowry BMD, Wang E, Vogel LC. Obesity and obesity-related secondary conditions in adolescents with intellectual/developmental disabilities. J Intellect Disabil Res 2010 Sep;54(9):787-794. [doi: 10.1111/j.1365-2788.2010.01305.x] [Medline: 20630017]
5. Yamaki K, Rimmer JH, Lowry BD, Vogel LC. Prevalence of obesity-related chronic health conditions in overweight adolescents with disabilities. Res Dev Disabil 2011;32(1):280-288. [doi: 10.1016/j.ridd.2010.10.007] [Medline: 21115323]
6. Peterson MD, Kamdar N, Hurvitz EA. Age-related trends in cardiometabolic disease among adults with cerebral palsy. Dev Med Child Neurol 2019 Apr;61(4):484-489. [doi: 10.1111/dmcn.13777] [Medline: 29704244]
7. Heyn PC, Tagawa A, Pan Z, Thomas S, Carollo JJ. Prevalence of metabolic syndrome and cardiovascular disease risk factors in adults with cerebral palsy. Dev Med Child Neurol 2019 Apr;61(4):477-483. [doi: 10.1111/dmcn.14148] [Medline: 30663044]
8. Ryan JM, Crowley VE, Hensey O, Broderick JM, McGahey A, Gormley J. Habitual physical activity and cardiometabolic risk factors in adults with cerebral palsy. Res Dev Disabil 2014 Sep;35(9):1995-2002. [doi: 10.1016/j.ridd.2014.03.051] [Medline: 24864052]
9. Ryan JM, Allen E, Gormley J, Hurvitz EA, Peterson MD. The risk, burden, and management of non-communicable diseases in cerebral palsy: a scoping review. Dev Med Child Neurol 2018 Aug;60(8):753-764. [doi: 10.1111/dmcn.13737] [Medline: 29572812]
10. Martin Ginis KA, Ma JK, Latimer-Cheung AE, Rimmer JH. A systematic review of review articles addressing factors related to physical activity participation among children and adults with physical disabilities. Health Psychol Rev 2016 Dec;10(4):478-494. [doi: 10.1080/17437199.2016.1198240] [Medline: 27265062]
11. Martin Ginis KA, van der Ploeg HP, Foster C, et al. Participation of people living with disabilities in physical activity: a global perspective. Lancet 2021 Jul 31;398(10298):443-455. [doi: 10.1016/S0140-6736(21)01164-8] [Medline: 34302764]
12. Waltersson L, Rodby-Bousquet E. Physical activity in adolescents and young adults with cerebral palsy. Biomed Res Int 2017;2017:8080473. [doi: 10.1155/2017/8080473] [Medline: 29423412]
13. Shkedy Rabani A, Harries N, Namoora I, Al-Jarrah MD, Karniel A, Bar-Haim S. Duration and patterns of habitual physical activity in adolescents and young adults with cerebral palsy. Dev Med Child Neurol 2014 Jul;56(7):673-680. [doi: 10.1111/dmcn.12394] [Medline: 24506509]
14. Carlon SL, Taylor NF, Dodd KJ, Shields N. Differences in habitual physical activity levels of young people with cerebral palsy and their typically developing peers: a systematic review. Disabil Rehabil 2013 Apr;35(8):647-655. [doi: 10.3109/09638288.2012.715721] [Medline: 23072296]
15. Haegele J, Hodge S, Shapiro D, editors. Routledge Handbook of Adapted Physical Education: Routledge; 2020.
16. Chang YH, Yang HY, Shun SC. Effect of exercise intervention dosage on reducing visceral adipose tissue: a systematic review and network meta-analysis of randomized controlled trials. Int J Obes (Lond) 2021 May;45(5):982-997. [doi: 10.1038/s41366-021-00767-9] [Medline: 33558643]

XSL•FO
RenderX

17. Lin X, Zhang X, Guo J, et al. Effects of exercise training on cardiorespiratory fitness and biomarkers of cardiometabolic health: a systematic review and meta-analysis of randomized controlled trials. J Am Heart Assoc 2015 Jun 26;4(7):e002014. [doi: 10.1161/JAHA.115.002014] [Medline: 26116691]
18. The definition and classification of cerebral palsy. Dev Med Child Neurol 2007 Feb;49(s109):1-44. [doi: 10.1111/j.1469-8749.2007.00001.x] [Medline: 17371509]
19. Mat Rosly M, Mat Rosly H, Davis Oam GM, Husain R, Hasnan N. Exergaming for individuals with neurological disability: a systematic review. Disabil Rehabil 2017 Apr;39(8):727-735. [doi: 10.3109/09638288.2016.1161086] [Medline: 27108475]
20. Deutsch JE, Westcott McCoy S. Virtual reality and serious games in neurorehabilitation of children and adults: prevention, plasticity, and participation. Pediatr Phys Ther 2017 Jul;29 Suppl 3:S23-S36. [doi: 10.1097/PEP.0000000000000387] [Medline: 28654475]
21. Sandlund M, McDonough S, Häger-Ross C. Interactive computer play in rehabilitation of children with sensorimotor disorders: a systematic review. Dev Med Child Neurol 2009 Mar;51(3):173-179. [doi: 10.1111/j.1469-8749.2008.03184.x] [Medline: 19191834]
22. Stanmore E, Stubbs B, Vancampfort D, de Bruin ED, Firth J. The effect of active video games on cognitive functioning in clinical and non-clinical populations: a meta-analysis of randomized controlled trials. Neurosci Biobehav Rev 2017 Jul;78:34-43. [doi: 10.1016/j.neubiorev.2017.04.011] [Medline: 28442405]
23. Lai B, Lee E, Kim Y, et al. Leisure-time physical activity interventions for children and adults with cerebral palsy: a scoping review. Dev Med Child Neurol 2021 Feb;63(2):162-171. [doi: 10.1111/dmcn.14751] [Medline: 33241561]
24. Lai B, Young HJ, Bickel CS, Motl RW, Rimmer JH. Current trends in exercise intervention research, technology, and behavioral change strategies for people with disabilities: a scoping review. Am J Phys Med Rehabil 2017 Oct;96(10):748-761. [doi: 10.1097/PHM.0000000000000743] [Medline: 28398967]
25. Lai B, Powell M, Clement AG, Davis D, Swanson-Kimani E, Hayes L. Examining the feasibility of early mobilization with virtual reality gaming using head-mounted display and adaptive software with adolescents in the pediatric intensive care unit: case report. JMIR Rehabil Assist Technol 2021 May 27;8(2):e28210. [doi: 10.2196/28210] [Medline: 34042602]
26. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30. [doi: 10.2196/jmir.1602] [Medline: 21393123]
27. Jennett PA, Affleck Hall L, Hailey D, et al. The socio-economic impact of telehealth: a systematic review. J Telemed Telecare 2003;9(6):311-320. [doi: 10.1258/135763303771005207] [Medline: 14680514]
28. Lai B, Davis D, Young R, et al. The effects of virtual reality tele-exergaming on cardiometabolic indicators of health among youth with cerebral palsy: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2022 Aug 17;11(8):e40708. [doi: 10.2196/40708] [Medline: 35976192]
29. Borg G. Borg's Perceived Exertion And Pain Scales: Human Kinetics; 1998.
30. Lai B, Rimmer J, Barstow B, Jovanov E, Bickel CS. Teleexercise for persons with spinal cord injury: a mixed-methods feasibility case series. JMIR Rehabil Assist Technol 2016 Jul 14;3(2):e8. [doi: 10.2196/rehab.5524] [Medline: 28582252]
31. Liguori G. ACSM's Guidelines for Exercise Testing and Prescription: Wolters Kluwer Health; 2020.
32. Cicone ZS, Holmes CJ, Fedewa MV, MacDonald HV, Esco MR. Age-based prediction of maximal heart rate in children and adolescents: a systematic review and meta-analysis. Res Q Exerc Sport 2019 Sep;90(3):417-428. [doi: 10.1080/02701367.2019.1615605] [Medline: 31157608]
33. Lai B, Oster RA, Davis D, et al. Telehealth movement-to-music with arm-based sprint-intensity interval training to improve cardiometabolic health and cardiorespiratory fitness in children with cerebral palsy: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2024 Mar 5;13:e56499. [doi: 10.2196/56499] [Medline: 38441939]
34. Bloemen MA, Backx FJ, Takken T, et al. Factors associated with physical activity in children and adolescents with a physical disability: a systematic review. Dev Med Child Neurol 2015 Feb;57(2):137-148. [doi: 10.1111/dmcn.12624] [Medline: 25403649]
35. Verschuren O, Wiart L, Hermans D, Ketelaar M. Identification of facilitators and barriers to physical activity in children and adolescents with cerebral palsy. J Pediatr 2012 Sep;161(3):488-494. [doi: 10.1016/j.jpeds.2012.02.042] [Medline: 22494875]
36. Verschuren O, Ketelaar M, Gorter JW, Helders PJ, Uiterwaal CS, Takken T. Exercise training program in children and adolescents with cerebral palsy: a randomized controlled trial. Arch Pediatr Adolesc Med 2007 Nov;161(11):1075-1081. [doi: 10.1001/archpedi.161.11.1075] [Medline: 17984410]
37. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
38. Johnson F, Wardle J, Griffith J. The Adolescent Food Habits Checklist: reliability and validity of a measure of healthy eating behaviour in adolescents. Eur J Clin Nutr 2002 Jul;56(7):644-649. [doi: 10.1038/sj.ejcn.1601371] [Medline: 12080404]
39. Godin G, Shephard RJ. Godin Leisure-Time Exercise Questionnaire. Med Sci Sports Exerc 1997;29:36-38. [doi: 10.1097/00005768-199706001-00009]
40. Amireault S, Godin G. The Godin-Shephard leisure-time physical activity questionnaire: validity evidence supporting its use for classifying healthy adults into active and insufficiently active categories. Percept Mot Skills 2015 Apr;120(2):604-622. [doi: 10.2466/03.27.PMS.120v19x7] [Medline: 25799030]

XSL•FO
RenderX

41. Sikes EM, Richardson EV, Cederberg KJ, Sasaki JE, Sandroff BM, Motl RW. Use of the Godin Leisure-Time Exercise Questionnaire in multiple sclerosis research: a comprehensive narrative review. Disabil Rehabil 2019 Jun;41(11):1243-1267. [doi: 10.1080/09638288.2018.1424956] [Medline: 29343122]
42. Zelener J, Schneider M. Adolescents and self-reported physical activity: an evaluation of the modified Godin Leisure-Time Exercise Questionnaire. Int J Exerc Sci 2016;9(5):587-598. [Medline: 29177023]
43. Kapur S, Kapur S, Zava D. Cardiometabolic risk factors assessed by a finger stick dried blood spot method. J Diabetes Sci Technol 2008 Mar;2(2):236-241. [doi: 10.1177/193229680800200210] [Medline: 19885348]
44. Kapur S, Groves MN, Zava DT, Kapur S. Postprandial insulin and triglycerides after different breakfast meal challenges: use of finger stick capillary dried blood spots to study postprandial dysmetabolism. J Diabetes Sci Technol 2010 Mar 1;4(2):236-243. [doi: 10.1177/193229681000400202] [Medline: 20307382]
45. Tisdell EJ, Merriam SB, Stuckey-Peyrot HL. Qualitative Research: A Guide to Design and Implementation: Jossey-Bass; 2025.
46. Stake RE. The Art of Case Study Research: Sage Publications; 1995. URL: https://link.springer.com/chapter/10.1007/978-3-030-85124-8_7
47. Fletcher AJ. Applying critical realism in qualitative research: methodology meets method. Int J Soc Res Methodol 2017;20(2):181-194. [doi: 10.1080/13645579.2016.1144401]
48. Willis JW. Foundations of Qualitative Research: Interpretive and Critical Approaches: SAGE Publications; 2007.
49. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology: American Psychological Association; 2012:57-71.
50. Braun V, Clarke V, Weate P. Using thematic analysis in sport and exercise research. In: Routledge Handbook of Qualitative Research in Sport and Exercise: Routledge; 2016.
51. O'Reilly M, Parker N. 'Unsatisfactory saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. Qual Res 2013;13(2):190-197. [doi: 10.1177/1468794112446106]
52. Lai B, Lee E, Wagatsuma M, et al. Research trends and recommendations for physical activity interventions among children and youth with disabilities: a review of reviews. Adapt Phys Activ Q 2020 Apr 1;37(2):211-234. [doi: 10.1123/apaq.2019-0081] [Medline: 32131052]
53. Lai B, Cederberg K, Vanderbom KA, Bickel CS, Rimmer JH, Motl RW. Characteristics of adults with neurologic disability recruited for exercise trials: a secondary analysis. Adapt Phys Activ Q 2018 Oct 1;35(4):476-497. [doi: 10.1123/apaq.2017-0109] [Medline: 30382753]
54. McIntyre S, Goldsmith S, Webb A, et al. Global prevalence of cerebral palsy: a systematic analysis. Dev Med Child Neurol 2022 Dec;64(12):1494-1506. [doi: 10.1111/dmcn.15346] [Medline: 35952356]
55. Lai B, Davis D, Narasaki-Jara M, et al. Feasibility of a commercially available virtual reality system to achieve exercise guidelines in youth with spina bifida: mixed methods case study. JMIR Serious Games 2020 Sep 3;8(3):e20667. [doi: 10.2196/20667] [Medline: 32880577]
56. Shields N, Synnot AJ, Barr M. Perceived barriers and facilitators to physical activity for children with disability: a systematic review. Br J Sports Med 2012 Nov;46(14):989-997. [doi: 10.1136/bjsports-2011-090236] [Medline: 21948121]

Abbreviations

CP: cerebral palsy
DEXA: dual-energy x-ray absorptiometry
GLTEQ: Godin Leisure-Time Exercise Questionnaire
GMFCS: Gross Motor Function Classification System
HMD: head-mounted display
XR: extended reality

Edited by J Sarvestan; submitted 03.Oct.2025; this is a non–peer-reviewed article;accepted 05.Dec.2025; published 27.Feb.2026.

Please cite as:
Lai B, Logan M, Young R, Wright A, Terrell J, Bright L, Davis D, Mendonca CJ
Feasibility of Vigorous Extended Reality Tele-Exergaming for Cardiometabolic Health in Youth With Mobility Disabilities: Protocol for a Case Series Study
JMIR Res Protoc 2026;15:e85246
URL: https://www.researchprotocols.org/2026/1/e85246
doi:10.2196/85246

XSL•FO
RenderX

© Byron Lai, Maggie Logan, Raven Young, Ashley Wright, Jordyn Terrell, Larsen Bright, Drew Davis, Christen J Mendonca. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Virtual Reality to Improve Pain Management and Mental Health in Stroke Survivors With Chronic Pain: Study Protocol for a Feasibility Randomized Controlled Trial on Virtual Reality-Acceptance and Commitment Therapy

Sérgio A Carvalho[1], PhD; Paulo Menezes[2], PhD; Catarina Duarte[3], MSc; David Skvarc[4], PhD; Ana Rita Sousa e Silva[1], PhD; Ana Valentim[5], MD; João Emanuel Diogo[6], BA; João Sargento-Freitas[7], MD, PhD; Inês A Trindade[8], PhD; Paula Castilho[1], PhD; Teresa Lapa[9], MD, PhD; Gerhard Andersson[10,11], PhD; Miguel Castelo-Branco[3], MD, PhD

[1]University of Coimbra, Center for Research in Neuropsychology and Cognitive and Behavioral Intervention (CINEICC), Rua do Colégio Novo, s/n, Coimbra, Portugal

[10]Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

[11]Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden

[2]Institute of Systems and Robotics (ISR), Faculty of Sciences and Technology, University of Coimbra, Coimbra, Portugal

[3]Coimbra Institute for Biomedical Imaging and Translational Research (CIBIT), Institute of Nuclear Sciences Applied to Health (ICNAS), University of Coimbra, Coimbra, Portugal

[4]School of Psychology, Faculty of Health, Deakin University, Victoria, Australia

[5]Department of Anesthesiology, Hospitais da Universidade de Coimbra, Coimbra, Portugal

[6]Center for Classical and Humanistic Studies (CECH), Coimbra, Portugal

[7]Department of Neurology, Universidade de Coimbra Faculdade de Medicina, Coimbra, Portugal

[8]EMBRACE Lab, Center for Health and Medical Psychology, School of Behavioural, Social and Legal Sciences, University of Örebro, Orebro, Sweden

[9]Faculty of Health Sciences, Universidade da Beira Interior, Covilhã, Portugal

Corresponding Author:
Sérgio A Carvalho, PhD
University of Coimbra, Center for Research in Neuropsychology and Cognitive and Behavioral Intervention (CINEICC), Rua do Colégio Novo, s/n, Coimbra, Portugal

Abstract

Background: Studies suggest that 40% to 65% of stroke survivors develop chronic poststroke pain (CPSP), which severely affects their quality of life and mental health. Empirical evidence suggests that existing treatments often fall short, underscoring the need for innovative, integrative interventions. Virtual reality (VR) seems to provide valuable tools in stroke rehabilitation. Also, contextual-behavioral psychological approaches, such as acceptance and commitment therapy (ACT), offer promising pain management and mental health resources, which seem to be feasible in VR formats. However, their combined application in CPSP remains unexplored.

Objective: This study protocol describes the VR-ACT study, which will test the feasibility and preliminary efficacy of an 8-week VR-ACT program for CPSP.

Methods: This pilot randomized controlled trial (N=30) will compare a VR-based ACT intervention with a sham VR control. The study will follow a mixed methods approach. Quantitative outcomes include pain intensity, psychological symptoms, and quality of life (via self-report measures), and brain network connectivity of the Triple Network (via functional magnetic resonance imaging). Feasibility will be evaluated through adherence, engagement, and acceptability. Qualitative feedback will be collected postintervention.

Results: This study was funded by the Portuguese Foundation for Science and Technology in February 2025. Data collection is expected to start in December 2025 and end in June 2026. Results are expected to be published in the fall/winter of 2026/2027.

Conclusions: This trial is expected to support the hypothesis that a VR-delivered ACT program is a feasible, acceptable, and potentially effective tool to support pain self-management and mental health in patients with CPSP, thereby laying the groundwork for larger multicenter trials.

Trial Registration: ClinicalTrials.gov NCT06990646; https://clinicaltrials.gov/study/NCT06990646

International Registered Report Identifier (IRRID): PRR1-10.2196/80611

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e80611) doi:10.2196/80611

KEYWORDS

chronic poststroke pain; virtual reality; acceptance and commitment therapy; feasibility; mixed method; study protocol

Introduction

Chronic poststroke pain (CPSP: pain lasting for more than 3 mo after stroke) affects 40% to 65% of stroke survivors [1] and can develop within 6 months after stroke [2]. CPSP substantially impacts patients' quality of life, increases the likelihood of functional dependence and cognitive decline, and increases the risk of depression, anxiety, and fatigue [3]; it is also a predictor of suicidality after stroke [4]. Pain is experienced up to 10 years after stroke [5], prevents optimal participation and gains during rehabilitation, and patients respond poorly to medication [6]. The etiology of CPSP appears to result from an interplay of psychosocial factors (eg, mental health, coping, and social support) [7] and neurophysiological factors (eg, stroke-related damage in corticospinothalamic pathways) [8], suggesting a biopsychosocial etiology. This seems to echo the hyperconnectivity of cognitive emotional brain networks (eg, default mode network [DMN], salience network [SN], and frontal parietal network [FPN]) in chronic pain [9]. Overall, pain after stroke is underresearched [3] and is perceived by patients as an unmet need [10]. Also, health care providers report organizational barriers (eg, insufficient human resources, infrastructure, training) to implementing stroke rehabilitation guidelines [11], which calls for the development of innovative, cost-effective pain management interventions for stroke survivors [12].

Virtual reality (VR) encompasses immersive human-digital interaction technologies [13]. VR seems to be effective in pain management [14] and a viable way to avoid the detrimental effects of opioid-centered analgesia plans [15]. VR technologies can enhance the efficacy of established behavioral interventions by standardizing the presentation of images, instructions, and environmental sounds that accompany behavioral exercises and by increasing the accessibility of practice (eg, patients can use it at home, at their own pace) [16]. VR also seems to be a feasible delivery method of meditation-focused behavioral interventions, which promote a decrease of pain sensitivity by decoupling executive and pain-relevant brain regions (eg, somatosensory cortex) [17]. Additionally, VR poststroke behavioral interventions appear to enhance rehabilitation compliance due to the diverse virtual experiences offered during therapy and the potential for gamification [18]. However, its effectiveness remains inconclusive in terms of clinical validity, as most studies focus on experimental pain [19] and/or fail to adhere to sound randomized controlled trial standards (eg, lack of placebo-controlled studies) [20]. Also, the integration of self-report psychological processes and neurophysiological assessment is still incipient in VR's efficacy studies in pain management [21].

Acceptance and commitment therapy (ACT) is a third-wave cognitive behavioral approach that promotes awareness and acceptance of internal experiences (eg, thoughts, emotions, physical sensations), and engagement with valued activities, through mindfulness and behavioral exercises [22]. ACT is especially helpful in managing chronic health conditions, as its primary focus is not on changing symptoms, but rather on promoting acceptance of difficult internal experiences and making behavioral changes towards a valued and meaningful life. Stroke survivors find ACT helpful in promoting acceptance and adjustment to stroke limitations, in fostering a sense of safety, and in developing self-regulatory skills [23]. ACT seems to lead to reductions in the triple network connectivity involved in pain chronicity, namely brain networks related to self-reflection (DMN), emotion (SN), and cognitive control (FPN) [24]. Although ACT is feasible and effective in reducing psychopathology in stroke survivors [25], its feasibility in CPSP management is yet to be explored. Additionally, although ACT has been found effective in pain management using technological formats (eg, online platforms, mobile apps) [26,27], ACT through VR has not been extensively tested. One recent study suggests ACT through VR is feasible and effective in improving mental health [28]. However, most VR interventions have been almost exclusively focused on exposure techniques or breathing meditation practices [29], rather than full psychological programs that promote emotional self-regulation skills (eg, attentional regulation) and reduction of avoidant-focused behavior (eg, commitment to valued action), and lack methodologically sound designs (eg, placebo-controlled).

Methods

Study Objectives

This project has 2 objectives: (1) we aimed to develop and test the feasibility of an 8-week VR-ACT program in a sample of patients with CPSP. By "feasibility," we mean feasibility in practice, that is, the extent to which the VR-ACT program can be delivered, measured by *retention* and *acceptability* of the recruitment and intervention implementation phases [30]. (2) We aimed to pilot test the efficacy of VR-ACT in improving pain, mental health, and adaptive psychological processes and skills and in reducing the functional connectivity of the triple network (DMN, SN, and FPN).

Design

The current project will conduct a parallel-group randomized controlled trial design and will follow a multimethod approach (qualitative and quantitative assessments) to assess the feasibility of the VR-ACT delivery [31]. The experimental condition (EC): the VR-ACT program: 8 modules of 3D immersive skills training (2 targeting awareness regulation skills; 3 promoting acceptance; 3 focused on behavioral change and commitment to valued action). To assess the preliminary efficacy and control for possible placebo effects, the control condition (CC) will consist of 8 nonimmersive 2D distracting video animations delivered through a VR headset (Sham VR). Participants in both conditions will use the VR devices throughout 8 weeks. Each

XSL•FO
RenderX

module will be available weekly in both conditions and have a duration of 20 minutes. Participants in the CC will be able to use the VR-ACT after the final assessment (see Figure 1).

Figure 1. Flowchart of participants. ACT: acceptance and commitment therapy; CC: control condition; EC: experimental condition; VR: virtual reality.

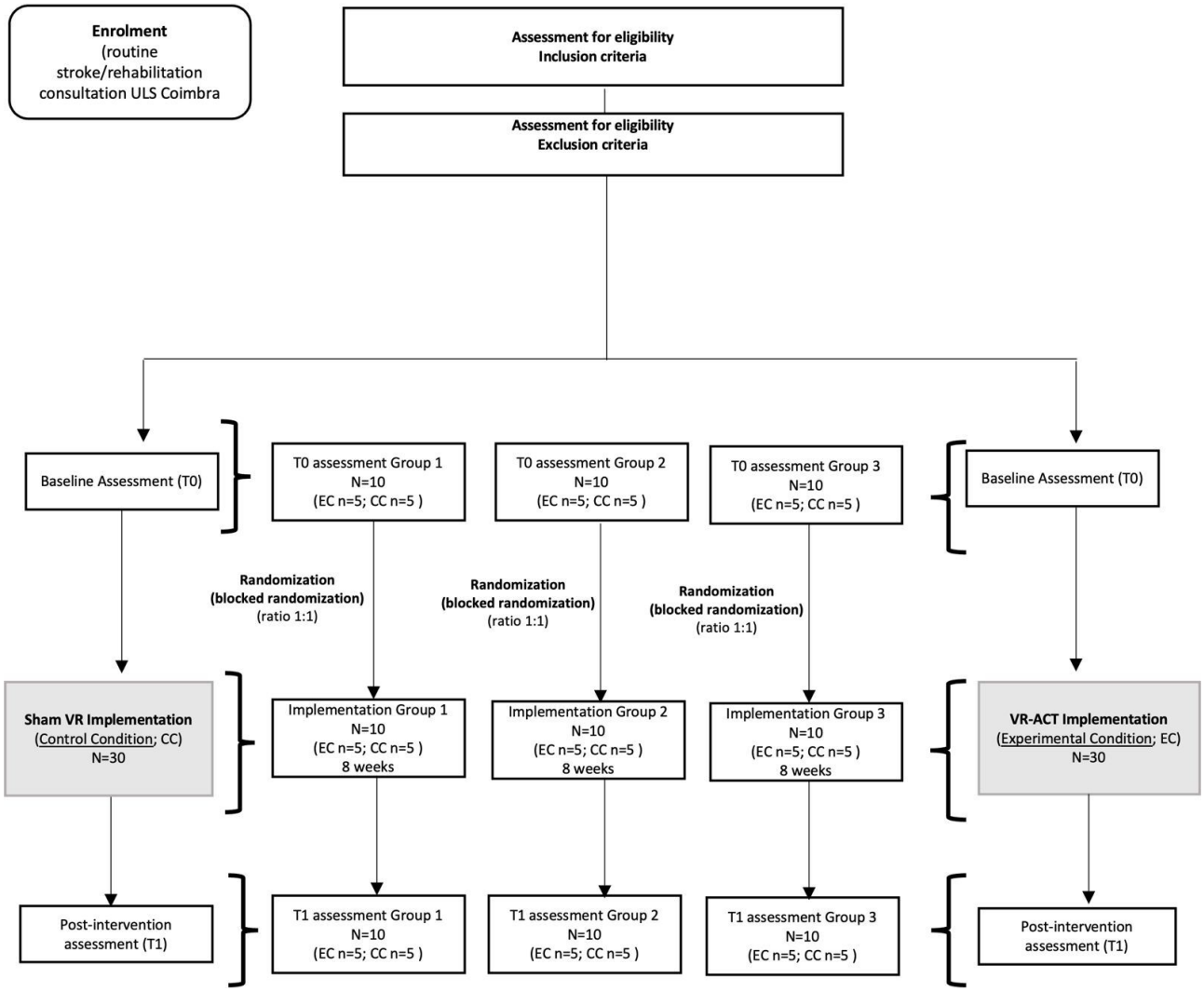

During the development of the VR-ACT intervention, a patient and public involvement group will be formed to monitor the adaptation of the exercises into VR and their tailoring for patients with CPSP. The group (n=5) will consist of 1 medical doctor and 1 nurse working in CPSP rehabilitation, 1 clinical psychologist specializing in contextual behavioral approaches to chronic pain, and 2 patients with CPSP (1 male and 1 female). The research team and the patient and public involvement group will gather monthly during the development of the VR-ACT intervention to ensure the patient-centered adaptation of exercises is successful and acceptable.

Participants

Eligibility Assessment

The study will be conducted in adults diagnosed with CPSP who attend the Neurology Service, Pain Unit, and/or the Rehabilitation Service at the ULS-Coimbra, Portugal. Inclusion criteria include (1) medical diagnosis of CPSP conducted by a medical doctor collaborating with the research team, (2) age between 18 and 80 years, (3) implicit de facto internet and computer literacy, and (4) willingness to comply with the study procedures. Exclusion criteria include (1) currently with active malignancy, (2) severe cognitive impairment, (3) currently undergoing any psychological intervention or VR-delivered pain management program, (4) current severe psychiatric symptoms (ie, psychosis, severe depression, nonsuicidal self-injury, suicide attempt in the last month), (5) language impairment with severe comprehension deficit, (6) history of photosensitive epilepsy or previous experience of severe simulator sickness, and (7) other neurological conditions (eg, dementia or Parkinson disease).

Sample Size Calculation

The study was powered for a large final timepoint between-group effect size (standardized mean difference=0.8), allowing for a variant autoregressive ρ over 2 measurements (preintervention and postintervention; 1 and 0.8) to accommodate the likely instability of the outcome variable(s). Setting the α at .05 and a minimum of 80% power, a minimum sample size of N=30 is required (EC: n=15; CC: n=15). We used the power.mrmm function from the R “longpower” package, through WebPower. A sample size of 15 to 30

XSL•FO
RenderX

participants is also sufficient to attain saturation in qualitative analyses conducted through interviews [32].

To minimize dropout, (1) 2 phone calls will occur during the VR-ACT implementation to address questions on the VR-ACT content. (2) We also expect minor technical difficulties when using the VR headset. At the beginning of each module, an avatar therapist will provide a welcome message, reminding you of the research team's contact information in case of any technical problems. Before the implementation, an online group meeting will be conducted to introduce the VR headset to participants in both conditions and help with navigation. (3) To promote engagement, weekly emails will be sent to all participants reminding them to use the VR headsets. Also, after the VR system registers 3 days of absence, automatic reminders will be sent to the participants' emails.

Procedures

VR Program Development

The VR-ACT intervention content will adapt ACT-based interventions for chronic pain that were previously developed and tested by the research team (eg, the face-to-face COMP.ACT program [33] and the online iACTwithPain platform [34]) into a VR format and tailored to CPSP. The research team will adapt experiential practices (eg, mindfulness meditation) and exercises (eg, values clarification and commitment to valued action) into a VR format through VR headset equipment (MetaQuest). The VR-ACT content will be developed through an "Unreal Engine" application with multiple rooms and different designs/environments, allowing the user-patient/participant to explore those that best suit their current need (eg, pain alleviation; emotional regulation training; establishing valued-based goals for the day) and/or the exercise suggested by the VR system (each module will have a suggested sequence; see Textbox 1).

An embedded registration system will be developed to monitor the progress and number of complete or incomplete sessions and exercises, which will be indicators of usability. This will enable the assessment of preferred time of usage and other user-centered information that can be further utilized in VR-ACT patient-centered tailoring. This registration will be supported by a centralized database and allocated to a server within the University of Coimbra, Portugal. Simultaneously, the research team will develop the content and format of the Sham VR control condition, which will be delivered through the same VR headset equipment, but with different content: 8 distraction-based exercises delivered via 2D nonimmersive video animations.

Textbox 1. The VR-ACT intervention content.

Modules 1 and 2: Awareness Regulation Skills

- Welcome to VR-ACT: control as the problem, and not the solution (Avatar Therapist)
- Introduction to Mindfulness (Avatar Therapist)
- Mindful breathing practice
- Body-scan practice
- Soothing Rhythm Breathing practice
- Mindfulness of thoughts practice (Leaves on a Stream).

Modules 3, 4, and 5: Acceptance of Pain and Internal Experiences

- Introduction to willingness and acceptance: Why accept? (Avatar Therapist)
- Letting go of control—part I: the quicksand metaphor (adapted to immersive VR)
- Letting go of control—part II: Joe the bum metaphor (adapted to immersive VR)
- Mindfulness of difficult emotions practice
- Accepting with self-compassion: soften, soothe, and allow practice.

Modules 6, 7, and 8: Behavior Change and Valued Action

- Introduction to Values (Avatar Therapist)
- The passengers on the bus metaphor (adapted to immersive VR)
- Mindfulness meditation on values: rediscovering what matters to me.
- Interactive game: 4 steps to committed action
- 90th birthday imagery practice (adapted to immersive VR)
- Life after VR-ACT: awareness and acceptance practice as building blocks of valued action.

XSL•FO
RenderX

Data Collection

The research team members, who are physicians at the Neurology Service, the Pain Unit, and/or the Rehabilitation Unit of ULS-Coimbra, Portugal, will invite participants and inform them of the aims and conditions for participation. Those who provide written informed consent will be screened for medical eligibility. Then, psychological eligibility will be individually assessed, online and at a time preferred by participants, by the clinical psychologists of the research team, for severe neurocognitive and psychiatric symptoms (using Mini Mental State Examination and a semistructured clinical interview adapted from Structured Clinical Interview for DSM). Those who meet criteria for severe psychiatric disorders and/or cognitive impairment will be excluded from the study, given feedback, and advised to seek medical/psychological consultation. Eligible participants will receive the link to the online preintervention assessment (T0) self-report protocol. A task-oriented (see Section B in Multimedia Appendix 1) functional magnetic resonance imaging (fMRI) will be scheduled to collect preintervention data on the functional connectivity of the triple network (DMN, SN, and FPN). Participants will then be randomly assigned by a member of the research team to 1 of 2 conditions (EC and CC) using random permuted blocks to ensure balance in age (via an online randomizer, eg, random.org [35]). Participants will be informed of the starting date via telephone (or alternatively via email). The medical/nursing collaborators and the research team members not involved with the implementation monitoring will be blinded to the allocation. The two conditions will be masked to participants by not explicitly informing whether the content of the VR is 3D (VR-ACT) or 2D (Sham VR), with both being delivered through the same model of VR headset (MetaQuest 3.0). Nevertheless, for ethical reasons, the two conditions in the study are described in the patient information sheet. A scheduled meeting in the ULS-Coimbra will be held to provide headset devices to participants. Three groups or waves of recruitment and VR-ACT implementations will be conducted (group 1, n=10; group 2, n=10; group 3, n=10; see Figure 1). At the end of implementation, participants will return the devices and will conduct the post-intervention assessment (T1): online psychosocial self-report data collection; and functional connectivity of the triple network data collection (fMRI). Additionally, participants in the EC will be interviewed online by a research team member who is blind to the allocation to collect qualitative data on the acceptability and usability of the VR-ACT (see Section A in Multimedia Appendix 1). The time points of measurements are described in the SPIRIT figure (See Figure 2).

XSL•FO
RenderX

Figure 2. SPIRIT figure of the VR-ACT study. [a]Conducted by the medical and nursing collaborators of the research team. [b]Conducted by a clinical psychologist member of the research team. [c]See Section A in Multimedia Appendix 1. [d]From 0=not at all appropriate to 4=very much appropriate. [e]Please let us know why you decided not to complete the data collection.

	STUDY PERIOD									
	Enrolment	Allocation	Preintervention			Intervention		Postintervention		
			Wave 1	*Wave 2*	*Wave 3*	*Week 1*	*Week 8*	*Wave 1*	*Wave 2*	*Wave 3*
TIMEPOINTS	*-T0*	*0*	*T0*	*T0*	*T0*			*T1*	*T1*	*T1*
ELIGIBILITY screening: inclusion criteria										
Sociodemographic questionnaire[a]	✓									
Medical screening[a]	✓									
ELIGIBILITY screening: exclusion criteria										
Semistructured clinical interview (SCID)[b]	✓									
ALLOCATION		✓								
INTERVENTIONS										
VR-ACT						◆—	—◆			
Sham-VR						◆—	—◆			
ASSESSMENTS										
Sociodemographic and clinical questionnaire			✓	✓	✓			✓	✓	✓
Self-report questionnaires (primary outcomes)										
Numerical Pain Rating Scale			✓	✓	✓			✓	✓	✓
Pain Disability Index			✓	✓	✓			✓	✓	✓
Depression, Anxiety and Stress Scale - 21			✓	✓	✓			✓	✓	✓
World Health Organization Quality of Life assessment - Brief			✓	✓	✓			✓	✓	✓
Self-report questionnaires (secondary outcomes)										
Comprehensive assessment of acceptance and commitment therapy processes			✓	✓	✓			✓	✓	✓
Mindful Attention Awareness Scale			✓	✓	✓			✓	✓	✓
Difficulties in Emotion Regulation Scale – Short Form			✓	✓	✓			✓	✓	✓
Functional connectivity of the Triple Network (fMRI)			✓	✓	✓			✓	✓	✓
FEASIBILITY ASSESSMENTS										
Systems Usability Scale								✓	✓	✓
Acceptability interview[c]								✓	✓	✓
Acceptability questionnaire[c]								✓	✓	✓
Data collection										
% participants lost from screening to allocation		✓								
Appropriateness of protocol via Likert-like scale[d]			✓	✓	✓					
Reasons for not completing data collection via open question[e]			✓	✓	✓					
Implementation										
% completion								✓	✓	✓
% homework compliance								✓	✓	✓
% skills training between modules								✓	✓	✓
% post-intervention assessment completion								✓	✓	✓

XSL•FO
RenderX

Feasibility Assessment

Assessment of Data Collection and Randomization

The feasibility [36] of the recruitment and data collection will be assessed via (1) the number of participants lost from screening to randomization; (2) *appropriateness* of size, format, and content of the assessment protocol, by responding to a Likert-like scale (from 0=not at all appropriate to 4=very much appropriate) regarding (a) *accessibility* to the protocol, (b) *willingness* to provide the required information, and (c) *size* of the protocol; (3) reasons for not completing data collection (open question "Please, let us know why you decided not to complete the data collection"); (4) support provided by medical/nursing collaborators in referring potential participants measured by average number per week of participants referred by the medical/nursing team.

Assessment of the VR-ACT Delivery

The feasibility of the VR-ACT implementation will be evaluated in terms of (1) *adherence*: (a) percentage of participants who completed the intervention, (b) number of sessions completed, and (c) an open-ended question asking the reasons for not completing the intervention; (2) *engagement*: (a) percentage of participants who completed homework assignments and frequency of skills practice between sessions and (b) participants' average number of VR-ACT device log-ins; (3) *retention/attrition*: number of participants who completed the postintervention assessment; (4) *acceptability*: participants will use a Likert-like scale (eg, 0=not at all; 4=very useful) to rate the VR-ACT modules in terms of its *usefulness*, *comprehensibility*, and *likelihood to continue using learned skills*. Participants will provide suggestions to improve the content of sessions via open-ended questions (see Section A in Multimedia Appendix 1).

Instruments

Demographic and Clinical Data

Demographic data will be collected through a self-report questionnaire (ie, age, marital status, education, sex assigned at birth, gender identity), as well as clinically relevant data (ie, date of stroke, date of pain onset after stroke, poststroke pain diagnosis, current medication, and other chronic pain diagnoses).

Primary Outcomes

Pain Intensity (Numeric Rating Scale)

The Numerical Pain Rating Scale [37,38] is an 11-point scale used to measure pain intensity, ranging from 0 ("No pain") to 10 ("Worst imaginable pain"). It assesses pain through 3 ratings: current pain, highest pain in the last 24 hours, and lowest pain in the last 24 hours, which are then averaged to provide a single score of pain intensity.

Pain Impact (Pain Disability Index)

The pain disability index [39] is an 11-point scale (0="no disability"; 10="worst disability"), a widely used scale that measures pain disability in 7 daily life domains. Higher scores indicate higher pain disability.

Psychological Distress (Depression, Anxiety, and Stress Scale)

The Depression, Anxiety, and Stress Scale (DASS-21) [39,40] is a 21-item questionnaire composed of 3 subscales designed to measure the negative emotional states of depression, anxiety, and stress. Each of the 3 DASS subscales comprises seven items, which respondents rate on a 4-point severity/frequency scale (0 to 3) in relation to the extent to which they experienced each state over the past week.

Quality of Life (World Health Organization Quality of Life-Bref)

The World Health Organization Quality of Life-Bref [41] is a 26-item multidimensional measure of quality of life that assesses subjective quality of life. This instrument measures four dimensions of quality of life (physical, psychological, social relations, and environment). It can also be used as a measure of general quality of life.

Secondary Outcomes

Psychological (In)flexibility (Comprehensive Assessment of ACT)

The Comprehensive Assessment of ACT [42,43] is a measure of psychological flexibility as conceptualized by ACT, with three subscales: openness to experience, valued action, and behavioral awareness. This project will use the short version with 18 items, measured in a Likert-like 7-point scale (0–"Strongly disagree" to 6–"Strongly agree").

Emotional (Dys)regulation (Difficulties in Emotion Regulation Scale-Short Form)

The Difficulties in Emotion Regulation Scale-Short Form (DERS-SF) [44,45] is an 18-item version of the DERS scale that assesses difficulties in regulating emotions in 6 different dimensions, 3 items per dimension, measured in a 5-point Likert scale (1=almost never to 5=almost always): nonacceptance of emotional responses, (2) difficulties engaging in goal-directed behavior, (3) impulse control difficulties, (4) lack of emotional awareness, (5) limited access to emotion regulation strategies, and (6) lack of emotional clarity. This scale can be used unidimensionally to measure difficulties in emotional regulation, where higher scores mean more difficulties in emotional regulation.

Mindful Awareness (Mindful Attention Awareness Scale)

The Mindful Attention Awareness Scale [46,47] is a 15-item scale designed to assess a core characteristic of mindfulness, namely, a receptive state of mind in which attention, informed by a sensitive awareness of what is occurring in the present, simply observes what is taking place. It measures mindful awareness in a 6-point Likert-like scale (1=almost always; 6=almost never).

The Triple Network

The functional connectivity of the Triple Network will be assessed through fMRI, specifically examining the functional connectivity of brain regions involved in chronic pain, including brain networks related to self-reflection (DMN), emotion (SN), and cognitive control (FPN).

XSL•FO
RenderX

System Usability Scale

The System Usability Scale [46,47] is a 10-item scale that measures global view system usability (example item: "I thought the system was easy to use"), through a 5-point response scale ranging from 1=strongly disagree to 5=strongly agree.

Data Analyses: Feasibility and Preliminary Efficacy

The feasibility of VR-ACT will be assessed qualitatively through open-ended questions (eg, "What would you say were the benefits of using the VR-ACT headset?," "What have you learned from the VR-ACT program?") designed to explore adherence, engagement, and acceptability of the intervention (see Section A in Multimedia Appendix 1). Thematic analyses will be conducted with the support of NVivo, and common themes in each question will be categorized. Also, feasibility will be quantitatively assessed through (1) dropout rate, (2) number of usage of the VR headset (embedded software usage monitoring), (3) percentage of intervention completion, (4) a 5-point Likert scale measuring usefulness, comprehensibility, and likelihood to continue using learned skills. The usability of the VR-ACT will be quantitatively assessed through the System Usability Scale.

The preliminary efficacy of VR-ACT will be pilot tested through 2 statistical methods: (1) intention to treat analyses, using linear mixed models to examine statistically significant change over time, moderated by allocation, and post hoc differences, adjusted for multiple comparisons using the Bonferroni technique; (2) reliable change index for each outcome for preintervention to postintervention timepoints, to assess clinically significant individual changes [48]. Network-based statistics [49] will be used to assess significant differences in network connectivity within and between conditions.

Ethical Considerations

All procedures will be conducted in accordance with the 1964 Helsinki Declaration and its subsequent amendments or comparable ethical standards. This study was approved by the Ethics Committee of the Faculty of Psychology and Educational Sciences at the University of Coimbra (ref: CEDI/FPCEUC: 98/R-1). The study will be conducted in accordance with European directives for research involving human participants, such as the General Data Protection Regulation. All participants will provide informed consent to participate in the study. Participants will be provided a patient information sheet with the following information: (1) the aim of the study and the eligibility criteria for participating; (2) the different phases of the study, including assessments and the randomization; (3) the benefits and risks of participating (namely that "We do not anticipate any risks arising from your participation in the study. However, if participating in the study results in a negative emotional experience and you need support, please contact the research team" and other national health resources will be provided). In this case, their participation will be interrupted. (4) Participation is voluntary, participants have the right to drop out without being obligated to provide justification, and access to treatment as usual, as well as their relationships with the medical and nursing team, will not be influenced by this decision; also, the decision to participate will not interfere with their medical treatment; (5) the confidentiality of data will be ensured in all results dissemination activities, and identifiable data in datasets will be password protected, to which only the research team will have access; (6) audio recording of postintervention interviews will be immediately deleted after written transcription. Informed consent will be obtained from each participant by the medical/nursing collaborator who will invite patients to participate before any data collection. (7) The current study will not conduct an assessment of neurological pathology; thus, data collected through fMRI will not be indicative of clinical diagnoses nor associated pathophysiology. The study results will be provided to interested participants in a format that is accessible.

A steering committee composed of 1 participant, 1 physician (anesthesiologist), 1 clinical psychologist, and 1 ethicist will monitor the delivery of the 8-week VR-ACT intervention and ensure the ethical and scientific integrity of the study (eg, adherence to the protocol). The steering committee will gather online at weeks 1, 3, 6, and 8 to discuss the trial progress, methodological issues arising from the implementation, and ethical decision-making. The trial will be interrupted if >50% of participants in the VR-ACT condition report increased psychological distress after session 1, and/or if >50% of the members of the steering committee decide so.

Results

This study was funded by the Portuguese Foundation for Science and Technology in February 2025 (2023.13402.PEX; [50]). Data collection is expected to start in December 2025 and end in June 2026. Data analysis will start in June 2026 and continue until September 2026. Results are expected to be disseminated via a peer-reviewed paper prepared and published in the fall/winter of 2026/2027. The research team has not started recruitment. Recruitment will start on December 1, 2025, and end on January 31, 2026. See Section D in Multimedia Appendix 1 for Trial registration—data set. The VR-ACT project intends to disseminate its results to 3 targeted audiences: (1) the scientific community, via peer-reviewed empirical studies with the results of its feasibility and preliminary efficacy study; (2) the clinical/professional community (eg, health professionals delivering care to patients with CPSP, through oral and poster presentations in scientific/professional conferences, and scientific meetings with health care services; (3) the general community, via social media, mainstream media. The authorship in all scientific outputs will be determined according to the authors' contributions and the recommendations of the International Committee of Medical Journal Editors [51].

Discussion

Expected Scientific Contribution of VR-ACT

This protocol outlines the feasibility and preliminary efficacy study of the VR-ACT project. The project adopts a multidisciplinary approach by integrating scientific knowledge from medicine, neuroscience, psychology, and computer science and engineering to develop an ACT-based VR self-management program that aims to help patients with CPSP improve their pain management skills and mental health. It employs a mixed

XSL•FO
RenderX

methods approach, measuring feasibility and preliminary efficacy through both quantitative and qualitative methodologies. It is expected that (1) the implementation of VR-ACT will be feasible, measured by adherence (ie, attrition rate), engagement (number of VR usage; completion), and acceptability (qualitative postintervention assessment); (2) the VR-ACT program will be significantly more effective in reducing pain intensity and disability, and psychopathological symptoms, and in increasing quality of life, as well as in reducing functional connectivity of the triple network (fMRI), than the Sham VR.

This study is expected to have potential scientific and societal contributions: (1) development of an innovative cost-effective evidence-based pain management and mental health resource for patients with CPSP that integrates recent evidence on protective factors in CP (eg, acceptance-based emotional regulation skills, and attentional redirection) using technologies with potential personal and societal economic sustainability, thus being in alignment with the 2030 agenda for sustainable development of the United Nations (eg, access to physical and mental health and well-being for all); (2) a multimethod approach to feasibility and efficacy, by integrating qualitative and quantitative data, as well as both psychological and neuroimaging data, which will add to the literature a multilevel understanding of (preliminary) efficacy. The project will potentially open a new avenue of research in CPSP (ie, the development of self-regulatory skills through digital solutions in CPSP rehabilitation) and tentatively reinforce the need for biopsychosocial multimodal approaches to CPSP to be included in the training of health professionals working with CPSP. These expected results on the feasibility and preliminary efficacy of VR-ACT in CPSP will serve as a reference point for a future, larger, multicenter, long-term, cross-cultural study.

The current study has some potential limitations that should be considered. Due to its feasibility focus, the study is powered for a small sample size (N=30), which limits the statistical power to detect small effect sizes and also restricts the ability to draw definitive conclusions regarding the efficacy of the VR-ACT intervention. Additionally, although an effort will be made to control for the placebo effect with a sham-VR condition, controlling for expectancy and placebo effects in VR interventions remains challenging, and it can be anticipated that participants may still distinguish between conditions despite our efforts to mask the condition allocated. Finally, the generalizability of the expected results can be hindered by strict eligibility criteria. For example, the inherent technological literacy required to participate, as well as the absence of severe psychological and neurological disorders, may limit the applicability of the results to the broader CPSP population.

Conclusions

This study protocol provides information on the implementation and feasibility testing of the VR-ACT project, specifically the development, feasibility, and preliminary efficacy testing of an ACT intervention delivered via VR, tailored to enhance the pain management skills and mental health of patients with CPSP. This is an early-stage trial that has the potential to serve as a reference point for future larger-scale effectiveness studies of VR-ACT and related resources to improve pain self-management and mental health of patients with CPSP.

Acknowledgments

We would like to thank the stakeholders who were willing to collaborate with the current study, namely the ULS-Coimbra. Additionally, the authors acknowledge the use of an artificial intelligence (AI) software (ie, Grammarly) to assist with minor text edits. No text was written originally by an AI software. Also, a reference management software (Madeley) was used to conduct in-text citations and the reference list.

Funding

This project is funded by the FCT—Fundação para a Ciência e a Tecnologia, I.P. (projetos@fct.pt) via an exploratory R&D funding granted to the principal investigator (SAC) (https://sciproj.ptcris.pt/176452PRJ). Additionally, The first author´s research activity was funded by national funds from FCT—Fundação para a Ciência e a Tecnologia, I.P., under the Individual Call to Scientific Employment Stimulus with ref: 2021.01871.CEECIND (SAC) (https://sciproj.ptcris.pt/157773UID). The funding agency will have no involvement in the scientific process nor the decision to and process of publishing scientific outputs.

Data Availability

Anonymized datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: SAC, PM, CD, DS, MC-B
Funding acquisition: SAC
Methodology: SAC, PM, CD, DS, MC-B
Project administration: SAC
Software: PM, CD, DS, MC-B
Validation: PM, CD, DS, MC-B
Visualization: SAC
Writing – original draft: SAC

XSL•FO
RenderX

Writing – review & editing: SAC, PM, CD, DS, ARSS, AV, JED, JSF, IAT, PC, TL, GA, MC-B

Conflicts of Interest

None declared.

Multimedia Appendix 1
VR-ACT post-intervention semi-structured interview; mindfulness task-based fMRI; information about the study; all items from the World Health Organization Trial Registration Data Set.
[DOCX File, 125 KB - resprot_v15i1e80611_app1.docx]

Peer Review Report 1
Peer review report by Fundação para a Ciência e a Tecnologia (FCT), Portugal.
[PDF File, 945 KB - resprot_v15i1e80611_app2.pdf]

References

1. Hansen AP, Marcussen NS, Klit H, Andersen G, Finnerup NB, Jensen TS. Pain following stroke: a prospective study. Eur J Pain 2012 Sep;16(8):1128-1136. [doi: 10.1002/j.1532-2149.2012.00123.x] [Medline: 22407963]
2. Kumar B, Kalita J, Kumar G, Misra UK. Central poststroke pain: a review of pathophysiology and treatment. Anesth Analg 2009 May;108(5):1645-1657. [doi: 10.1213/ane.0b013e31819d644c] [Medline: 19372350]
3. Plecash AR, Chebini A, Ip A, et al. Updates in the treatment of post-stroke pain. Curr Neurol Neurosci Rep 2019 Nov 13;19(11):86. [doi: 10.1007/s11910-019-1003-2] [Medline: 31720885]
4. Tang WK, Liang H, Mok V, Ungvari GS, Wong KS. Is pain associated with suicidality in stroke? Arch Phys Med Rehabil 2013 May;94(5):863-866. [doi: 10.1016/j.apmr.2012.11.044] [Medline: 23262382]
5. Jönsson AC, Delavaran H, Iwarsson S, Ståhl A, Norrving B, Lindgren A. Functional status and patient-reported outcome 10 years after stroke: the Lund Stroke Register. Stroke 2014 Jun;45(6):1784-1790. [doi: 10.1161/STROKEAHA.114.005164] [Medline: 24803595]
6. Treister AK, Hatch MN, Cramer SC, Chang EY. Demystifying poststroke pain: from etiology to treatment. PM R 2017 Jan;9(1):63-75. [doi: 10.1016/j.pmrj.2016.05.015] [Medline: 27317916]
7. Gandolfi M, Donisi V, Battista S, et al. Health-related quality of life and psychological features in post-stroke patients with chronic pain: a cross-sectional study in the neuro-rehabilitation context of care. Int J Environ Res Public Health 2021 Mar 17;18(6):3089. [doi: 10.3390/ijerph18063089] [Medline: 33802829]
8. Osama A, Abo Hagar A, Elkholy S, Negm M, Abd El-Razek R, Orabi M. Central post-stroke pain: predictors and relationship with magnetic resonance imaging and somatosensory evoked potentials. Egypt J Neurol Psychiatry Neurosurg 2018 Dec;54(1). [doi: 10.1186/s41983-018-0041-z]
9. van Ettinger-Veenstra H, Lundberg P, Alföldi P, et al. Chronic widespread pain patients show disrupted cortical connectivity in default mode and salience networks, modulated by pain sensitivity. J Pain Res 2019;12:1743-1755. [doi: 10.2147/JPR.S189443] [Medline: 31213886]
10. Magaard G, Wester P, Levi R, et al. Identifying unmet rehabilitation needs in patients after stroke with a graphic Rehab-CompassTM. J Stroke Cerebrovasc Dis 2018 Nov;27(11):3224-3235. [doi: 10.1016/j.jstrokecerebrovasdis.2018.07.013] [Medline: 30097401]
11. Cormican A, Hirani SP, McKeown E. Healthcare professionals' perceived barriers and facilitators of implementing clinical practice guidelines for stroke rehabilitation: A systematic review. Clin Rehabil 2023 May;37(5):701-712. [doi: 10.1177/02692155221141036] [Medline: 36475911]
12. Haslam BS, Butler DS, Kim AS, Carey LM. Chronic pain following stroke: current treatment and perceived effect. Disabil Health J 2021 Jan;14(1):100971. [doi: 10.1016/j.dhjo.2020.100971] [Medline: 32830081]
13. Xiong J, Hsiang EL, He Z, Zhan T, Wu ST. Augmented reality and virtual reality displays: emerging technologies and future perspectives. Light Sci Appl 2021 Oct 25;10(1):216. [doi: 10.1038/s41377-021-00658-8] [Medline: 34697292]
14. Goudman L, Jansen J, Billot M, et al. Virtual reality applications in chronic pain management: systematic review and meta-analysis. JMIR Serious Games 2022 May 10;10(2):e34402. [doi: 10.2196/34402] [Medline: 35536641]
15. Gupta A, Scott K, Dukewich M. Innovative technology using virtual reality in the treatment of pain: does it reduce pain via distraction, or is there more to it? Pain Med 2018 Jan 1;19(1):151-159. [doi: 10.1093/pm/pnx109] [Medline: 29025113]
16. Keefe FJ, Huling DA, Coggins MJ, et al. Virtual reality for persistent pain: a new direction for behavioral pain management. Pain 2012 Nov;153(11):2163-2166. [doi: 10.1016/j.pain.2012.05.030] [Medline: 22770840]
17. Hu XS, Beard K, Sherbel MC, et al. Brain mechanisms of virtual reality breathing versus traditional mindful breathing in pain modulation: observational functional near-infrared spectroscopy study. J Med Internet Res 2021 Oct 12;23(10):e27298. [doi: 10.2196/27298] [Medline: 34636731]
18. Demeco A, Zola L, Frizziero A, et al. Immersive virtual reality in post-stroke rehabilitation: a systematic review. Sensors (Basel) 2023 Feb 3;23(3):1712. [doi: 10.3390/s23031712] [Medline: 36772757]

XSL•FO
RenderX

19. Malloy KM, Milling LS. The effectiveness of virtual reality distraction for pain reduction: a systematic review. Clin Psychol Rev 2010 Dec;30(8):1011-1018. [doi: 10.1016/j.cpr.2010.07.001] [Medline: 20691523]
20. Mallari B, Spaeth EK, Goh H, Boyd BS. Virtual reality as an analgesic for acute and chronic pain in adults: a systematic review and meta-analysis. J Pain Res 2019;12:2053-2085. [doi: 10.2147/JPR.S200498] [Medline: 31308733]
21. Ahmadpour N, Randall H, Choksi H, Gao A, Vaughan C, Poronnik P. Virtual reality interventions for acute and chronic pain management. Int J Biochem Cell Biol 2019 Sep;114:105568. [doi: 10.1016/j.biocel.2019.105568] [Medline: 31306747]
22. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006 Jan;44(1):1-25. [doi: 10.1016/j.brat.2005.06.006] [Medline: 16300724]
23. Large R, Samuel V, Morris R. A changed reality: experience of an acceptance and commitment therapy (ACT) group after stroke. Neuropsychol Rehabil 2020 Sep;30(8):1477-1496. [doi: 10.1080/09602011.2019.1589531] [Medline: 30924741]
24. Meier SK, Ray KL, Waller NC, Gendron BC, Aytur SA, Robin DA. Network analysis of induced neural plasticity post-acceptance and commitment therapy for chronic pain. Brain Sci 2020 Dec 23;11(1):10. [doi: 10.3390/brainsci11010010] [Medline: 33374858]
25. Graham CD, Gillanders D, Stuart S, Gouick J. An acceptance and commitment therapy (ACT)–based intervention for an adult experiencing post-stroke anxiety and medically unexplained symptoms. Clin Case Stud 2015 Apr;14(2):83-97. [doi: 10.1177/1534650114539386]
26. Carvalho SA, Skvarc D, Barbosa R, Tavares T, Santos D, Trindade IA. A pilot randomized controlled trial of online acceptance and commitment therapy versus compassion-focused therapy for chronic illness. Clin Psychol Psychother 2022 Mar;29(2):524-541. [doi: 10.1002/cpp.2643] [Medline: 34269493]
27. Trindade IA, Guiomar R, Carvalho SA, et al. Efficacy of online-based acceptance and commitment therapy for chronic pain: a systematic review and meta-analysis. J Pain 2021 Nov;22(11):1328-1342. [doi: 10.1016/j.jpain.2021.04.003] [Medline: 33892153]
28. Gorinelli S, Gallego A, Lappalainen P, Lappalainen R. Virtual reality acceptance and commitment therapy intervention for social and public speaking anxiety: A randomized controlled trial. J Contextual Behav Sci 2023 Apr;28:289-299. [doi: 10.1016/j.jcbs.2023.05.004]
29. Emmelkamp PMG, Meyerbröker K. Virtual reality therapy in mental health. Annu Rev Clin Psychol 2021 May 7;17:495-519. [doi: 10.1146/annurev-clinpsy-081219-115923] [Medline: 33606946]
30. Pearson N, Naylor PJ, Ashe MC, Fernandez M, Yoong SL, Wolfenden L. Guidance for conducting feasibility and pilot studies for implementation trials. Pilot Feasibility Stud 2020 Oct 31;6(1):167. [doi: 10.1186/s40814-020-00634-w] [Medline: 33292770]
31. Czajkowski SM, Powell LH, Adler N, et al. From ideas to efficacy: the ORBIT model for developing behavioral treatments for chronic diseases. Health Psychol 2015 Oct;34(10):971-982. [doi: 10.1037/hea0000161] [Medline: 25642841]
32. Green J, Thorogood N, Melendez-Torres GJ. Qualitative Methods for Health Research: Sage; 2025. URL: https://uk.sagepub.com/en-gb/eur/qualitative-methods-for-health-research/book284045 [accessed 2026-01-14]
33. Carvalho SA, Gillanders D, Forte T, et al. Self-compassion in acceptance and commitment therapy for chronic pain: a pilot study. Scand J Pain 2022 Jul 26;22(3):631-638. [doi: 10.1515/sjpain-2021-0214] [Medline: 34954932]
34. Guiomar R, Trindade IA, Carvalho SA, et al. Usability study of the iACTwithPain platform: an online acceptance and commitment therapy and compassion-based intervention for chronic pain. Front Psychol 2022;13:848590. [doi: 10.3389/fpsyg.2022.848590] [Medline: 35936338]
35. Random.org. URL: https://www.random.org [accessed 2026-01-22]
36. Pfledderer CD, von Klinggraeff L, Burkart S, et al. Consolidated guidance for behavioral intervention pilot and feasibility studies. Pilot Feasibility Stud 2024 Apr 6;10(1):57. [doi: 10.1186/s40814-024-01485-5] [Medline: 38582840]
37. Hartrick CT, Kovan JP, Shapiro S. The numeric rating scale for clinical pain measurement: a ratio measure? Pain Pract 2003 Dec;3(4):310-316. [doi: 10.1111/j.1530-7085.2003.03034.x] [Medline: 17166126]
38. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. Pain 2011 Oct;152(10):2399-2404. [doi: 10.1016/j.pain.2011.07.005] [Medline: 21856077]
39. Lovibond SH, Lovibond PF. Depression anxiety stress scales. APA PsycTests. URL: https://doi.apa.org/doi/10.1037/t01004-000 [accessed 2025-06-05]
40. Pais-Ribeiro JL, Honrado A, Leal I. Contribuição para o estudo da adaptação portuguesa das escalas de ansiedade, depressão e stress (EADS) de 21 itens de lovibond e lovibond [Article in Portuguese]. PSicologia, Saúde &Amp; Doenças 2004;5(1):229-239 [FREE Full text]
41. Development of the World Health Organization WHOQOL-BREF quality of life assessment. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/S0033291798006667]
42. Tyndall I, Vagos P, Cano S, Pinto H, Trindade IA. Assessing psychological flexibility in chronic illness. Psychol Rec 2023 Dec;73(4):565-574. [doi: 10.1007/s40732-023-00572-y]
43. Francis AW, Dawson DL, Golijani-Moghaddam N. The development and validation of the comprehensive assessment of acceptance and commitment therapy processes (CompACT). J Contextual Behav Sci 2016 Jul;5(3):134-145. [doi: 10.1016/j.jcbs.2016.05.003]

XSL•FO
RenderX

44. Kaufman EA, Xia M, Fosco G, Yaptangco M, Skidmore CR, Crowell SE. The difficulties in emotion regulation scale short form (DERS-SF): validation and replication in adolescent and adult samples. J Psychopathol Behav Assess 2016 Sep;38(3):443-455. [doi: 10.1007/s10862-015-9529-3] [Medline: 41522882]
45. Moreira H, Gouveia MJ, Canavarro MC. A bifactor analysis of the difficulties in emotion regulation scale - short form (DERS-SF) in a sample of adolescents and adults. Curr Psychol 2022 Feb;41(2):757-782. [doi: 10.1007/s12144-019-00602-5]
46. Brooke J. SUS: a “quick and dirty” usability scale. In: Usability Evaluation In Industry: CRC Press (Taylor & Francis Group); 1996:207-212 URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke [accessed 2026-01-14]
47. Martins AI, Rosa AF, Queirós A, Silva A, Rocha NP. European Portuguese validation of the system usability scale (SUS). Procedia Comput Sci 2015;67:293-300. [doi: 10.1016/j.procs.2015.09.273]
48. Blampied NM. Reliable change and the reliable change index: still useful after all these years? Cogni Behav Therapist 2022;15:e50. [doi: 10.1017/S1754470X22000484]
49. Zalesky A, Fornito A, Bullmore ET. Network-based statistic: identifying differences in brain networks. Neuroimage 2010 Dec;53(4):1197-1207. [doi: 10.1016/j.neuroimage.2010.06.041] [Medline: 20600983]
50. VR-ACT after stroke: a feasibility RCT of a virtual reality delivery of ACT for pain management and mental health in patients with chronic post-stroke pain. Portal SciPROJ. URL: https://doi.org/10.54499/2023.13402.PEX [accessed 2026-01-22]
51. Recommendations for the conduct, reporting, editing and publication of scholarly work in medical journals. : International Committee of Medical Journal Editors; 2025 URL: https://www.icmje.org/icmje-recommendations.pdf [accessed 2026-01-14]

Abbreviations

ACT: acceptance and commitment therapy
CC: control condition
CPSP: chronic poststroke pain
DASS: Depression, Anxiety and Stress Scale
DERS-SF: Difficulties in Emotion Regulation Scale-Short Form
DMN: default mode network
EC: experimental condition
fMRI: functional magnetic resonance imaging
FPN: frontal parietal network
SN: salience network
VR: virtual reality

Edited by J Sarvestan; submitted 14.Jul.2025; this is a non–peer-reviewed article;accepted 31.Oct.2025; published 06.Feb.2026.

Please cite as:
Carvalho SA, Menezes P, Duarte C, Skvarc D, Sousa e Silva AR, Valentim A, Emanuel Diogo J, Sargento-Freitas J, Trindade IA, Castilho P, Lapa T, Andersson G, Castelo-Branco M
Virtual Reality to Improve Pain Management and Mental Health in Stroke Survivors With Chronic Pain: Study Protocol for a Feasibility Randomized Controlled Trial on Virtual Reality-Acceptance and Commitment Therapy
JMIR Res Protoc 2026;15:e80611
URL: https://www.researchprotocols.org/2026/1/e80611
doi:10.2196/80611

© Sérgio A Carvalho, Paulo Menezes, Catarina Duarte, David Skvarc, Ana Rita Sousa e Silva, Ana Valentim, João Emanuel Diogo, João Sargento Freitas, Inês A Trindade, Paula Castilho, Teresa Lapa, Gerhard Andersson, Miguel Castelo-Branco. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Assessing Associations Between Environmental, Sleep, and Physical Activity Factors and Metabolic Syndrome Risk: Protocol for the FEASible Study

Tania Ramos-Santiago[1], PhD; Elise Huglo[2], MS; Yanrong Li[1], BS; Christian Corral[1], MS; Congyu Wu[3], PhD; Kerry Kinney[2], BS, MS, PhD; Darla Castelli[4], PhD; Andreana Haley[1], PhD

[1]Department of Psychology, University of Texas at Austin, Austin, TX, United States

[2]Maseeh Department of Civil, Architectural, and Environmental Engineering, University of Texas at Austin, Austin, TX, United States

[3]Department of Systems Science and Industrial Engineering, Binghamton University, Greater Binghamton, NY, United States

[4]Department of Physical Therapy, Movement, and Rehabilitation Sciences, Northeastern University, Boston, MA, United States

Corresponding Author:
Tania Ramos-Santiago, PhD
Department of Psychology
University of Texas at Austin
Sarah M & Charles E Seay Bldg
108 E Dean Keeton St
Austin, TX, 78712
United States
Phone: 1 512 471 7926
Email: taniaramos@utexas.edu

Abstract

Background: FEASible is a cross-sectional observational study that explores women's daily living patterns through wearable devices and home environment sensors to validate the use of physical activity and indoor air quality data as indicators of risk for metabolic syndrome (MetS) and cardiovascular disease.

Objective: Leveraging transdisciplinary expertise, we implement a low-cost, dense sampling approach among 800 adult women, 60% Hispanic or Latina, with a subgroup of 225 participants opting for neuroimaging to assess MetS-related brain vulnerability.

Methods: Participants residing in Central Texas use a smartwatch and a custom-built air quality sensor to monitor their activities and environment for two weeks. Other variables, such as social determinants of health, medical history, and lifestyle, are reported through surveys. During their initial visit, we gather blood pressure measurements, body composition, and lipid profile information.

Results: As of August 2025, a total of 805 participants have completed the eligibility survey, of whom 204 have completed 2 weeks of sensor data collection, showing a diverse participant pool comprising 60% (210/348) Hispanic or Latinas, 39% (135/348) non-Hispanic or Latinas, and 1% (3/348) unknown. Participants are aged 18-40 years, with an average age of 27 years (SD 6 years). The study was funded in May 2023 (National Heart, Lung, and Blood Institute; grant R01HL168374), and data collection began in October 2023, with a projected completion date of May 2028. Ongoing analyses and publication of early cohorts are expected to occur throughout 2026 and 2027, with final analyses and dissemination of results anticipated by the end of the funding period in 2028.

Conclusions: We have established a feasible pipeline to collect data that will help yield valuable insights into MetS risk factors among Latina women, including brain scan magnetic resonance imaging data and environmental exposure measurements. This paper aims to provide a rationale for procedures and case examples from the first year of data collection leading to health risk modeling, thus informing future interventions on MetS, heart disease, and all-cause mortality among Latina women.

International Registered Report Identifier (IRRID): DERR1-10.2196/82034

(JMIR Res Protoc 2026;15:e82034) doi:10.2196/82034

KEYWORDS

indoor air quality; Latina women's health; metabolic health; metabolic syndrome; MetS; physical activity; wearable sensors

XSL•FO
RenderX

Introduction

Nearly 1 in 5 people in 2022 identifies as Hispanic or Latino, making this population segment one of the most significant and fastest-growing racial or ethnic minorities in the United States [1]. Texas has been categorized as having one of the nation's largest Hispanic and Latino state populations, accounting for 12.1 million people [2]. Statistics show that Hispanic and Latino communities, a substantial proportion of the population in Texas, face a unique set of health challenges. Among these, cardiovascular diseases (CVDs) are one of the leading causes of death in Hispanic and Latina women in the United States [3,4]. CVD is a broad term covering various heart and blood vessel diseases. Metabolic syndrome (MetS) is a cluster of risk factors specific to CVDs that includes abdominal obesity, elevated fasting glucose, elevated triglycerides, low-density lipoprotein cholesterol, and high blood pressure. MetS has been independently associated with and is considered a precursor to CVDs and mortality [5,6]. While Hispanic and Latina women generally exhibit a higher prevalence of most MetS risk factors compared with non-Hispanic White women, our understanding of what environmental or behavioral factors contribute to these disparities remains limited to prescribed medications, self-report surveys, and risk ratios [7-9]. Social determinants, such as socioeconomic status, access to health care, and lifestyle behaviors, also contribute to the elevated risk of heart disease in this population [10]. Considering that Hispanic and Latino communities are growing fast in the United States, the higher prevalence of MetS in women, and the trends of health disparities after the global pandemic, a critical missing component that needs to be further understood is the correlation of daily behavioral activities with MetS risk.

There is an increasing focus on studying health, environmental, and behavioral patterns through cost-effective and widely available wearable sensors to understand and improve public health. The success of this endeavor is critically dependent on accurate models reflecting unique daily lifestyles, ultimately guiding potential interventions to promote health equity through behavioral interventions. Smartwatches and other wearable devices have garnered considerable attention, as they can help track key health indicators, including sleep patterns, heart rate, and physical activity. Indeed, poor physical activity and sleep time have been consistently associated with an increased risk for MetS in women [11-13]. However, evaluating Hispanic and Latina women's unique environmental exposures and cardiovascular risks is a significant gap in the MetS risk framework.

A second set of essential parameters to monitor environmental health risks is indoor air pollutants and other relevant measures of the indoor environment, such as temperature and humidity. Indeed, adults spend 90% of their time indoors, but this aspect is not generally evaluated as rigorously as it should be [14-16]. Specific sensing technologies now make this possible at a lower cost. For example, particulate matter ≤2.5 μm in diameter ($PM_{2.5}$) is a widely studied outdoor air pollutant in urban areas. It has been found to directly affect the respiratory system, hypertension, physical functioning, and cognitive performance [17-19]. Moreover, chemical exposure affects cognitive function; for example, carbon dioxide (CO_2) is predominantly present in indoor environments and has also been linked to reduced cognitive function and productivity [20]. The development of portable environmental sensing devices offers a promising method for monitoring indoor air quality. However, it is imperative to validate the reliability and accuracy of these low-cost wearables and custom-made air quality sensing devices to better measure attributes of daily living and quantify MetS risk. To establish the utility of these devices in personalized interventions, a thorough calibration and validation of the developed sensing device models are necessary to capture relevant, ecologically valid data in the home environment.

The FEASible study purpose is divided into 3 specific aims to validate a pipeline of low-cost objective measurements and interpret MetS risk and MetS-related brain vulnerability networks. The primary objective is to identify which modifiable peripheral cardiometabolic biomarkers are associated with MetS-related brain vulnerability in Latina women, serving as a ground-truth standard. The hypothesis is that MetS, peripheral inflammation, white and gray matter volume, white matter hyperintensities, and measures of cerebral metabolism will form a network of MetS risk and brain vulnerability. The second aim is to validate the use of low-cost sensors that capture multiple environmental measures inside and around the home to assess exposures that may affect sleep and physical activity. We will use this information to identify potential risk factors for MetS and MetS-related brain vulnerability. The hypothesis states that properly validated, low-cost sensors can capture sleeping habits and home environmental conditions that contribute to more disordered sleep and less participation in physical activity, which in turn mediates MetS outcomes. The third aim is to determine the feasibility and association of mobile sensor measurements of daily behavior, activity, sleep, and environmental attributes with MetS and MetS-related brain vulnerability. The first hypothesis posits that features of daily living detected by mobile sensors will more accurately correlate with MetS risk and MetS-related brain vulnerabilities than participant baseline information. The second hypothesis is that environmental exposures detected by validated, low-cost monitoring devices will mediate the relationships between mobile and wearable sensor measurements of daily behavior and MetS risk and MetS-related brain vulnerabilities. Validating low-cost and objective measurements for MetS risk will aid in preventing MetS development, heart diseases, and all-cause mortality. Given the limited research on Latina women, identifying reliable and accurate customizable devices that target individual behaviors and environments is crucial for designing prevention efforts with greater precision.

Methods

Study Design

FEASible is a cross-sectional study conducted at an academic clinical neuroscience research laboratory. Each month, groups of 25-40 adult women are onboarded in the study. The study is active from October 2023 until May 2028. Figure 1 summarizes the study structure. First, the participant completes an eligibility survey through REDCap (Research Electronic Data Capture;

XSL•FO
RenderX

Vanderbilt University), followed by informed consent and comprehensive surveys. Then, the participant visits the laboratory to perform the health screening, set up the wearable device, and deploy the home environment sensor kit. Participation is required for a minimum of 14 to 30 days, during which dense sensor capture of daily living is conducted. After completing the 2-week data collection, the participant returns to the university to return the toolkit. The participant can schedule a subsequent visit for the in-laboratory blood draw and magnetic resonance imaging (MRI) scanning if eligible.

Figure 1. The FEASible study structure. Participants are asked to complete comprehensive surveys, followed by performing a health screening and a toolkit deployment visit where they get the smartwatch and the home environment sensor kit. The participants complete the 2-week data collection, and after they return the toolkit, they might opt to participate in the magnetic resonance imaging (MRI) scan and blood draw. Created with BioRender by Ramos et al [21], reproduced using license obtained by authors.

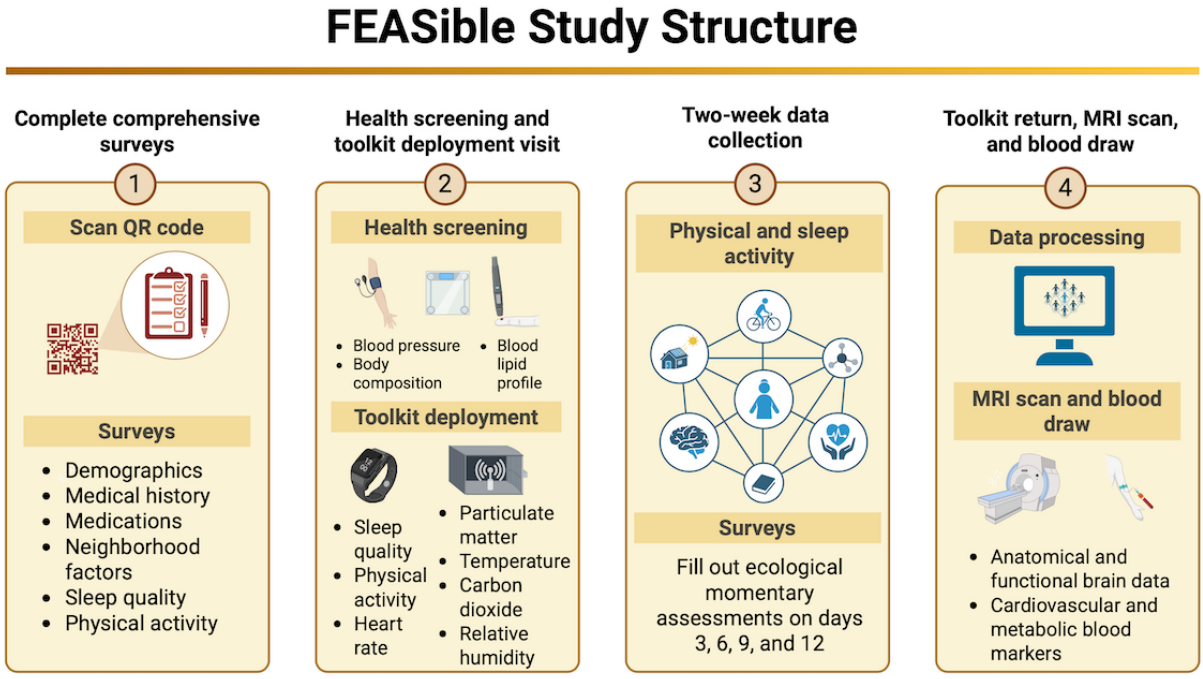

The research participant population of interest consists of adult women, with 60% (210/348) identifying as Latinas. The total sample size comprises 800 participants, with a subgroup of 225 participants opting to participate in neuroimaging. The inclusion criteria are (1) aged 18-40 years during data collection, (2) having no advanced health conditions or illness (eg, brain tumor, psychotic disorder, or history of neurological conditions), (3) not currently pregnant, (4) absence of a physical disability that inhibits mobility, and (5) living in Central Texas (2-hour radius of Austin, Texas). For the neuroimaging subgroup, the additional inclusion criteria are (1) no MRI contraindications, (2) no diagnosed medical diseases or psychological disorders, and (3) no history of head injury with loss of consciousness greater than 5 minutes. All study participants are compensated for completing the comprehensive surveys, participating in the 2-week data collection, and undergoing the MRI scan.

Sample Size and Power Analysis

To demonstrate statistical power, we conducted a simulation analysis for the anticipated experiments to test our hypotheses, aiming to determine the extent to which mobile-sensed behavioral features can correlate high versus low levels of risk for MetS. Datasets were simulated as random draws from a multivariate distribution with a ground-truth covariance based on sensing data collected from the 1012 participants for a 21-day study period during the Fall 2018 deployment of our previous UT1000 study [16,22,23]. We consider a 24-hour day to consist of four 6-hour epochs, namely morning (6-12 hours), afternoon (12-18 hours), evening (18-24 hours), and night (24-6 hours).

Recruitment Strategies and Screening

The recruitment strategies are flyers, business cards, word of mouth, Latina-focused Facebook groups, and Latino community events. To ensure a representative sample, we establish research partnerships with the community, local groups, and nonprofit organizations, while collaborating to offer health and environmental resources, including talks and educational materials. Our primary recruitment strategy is word of mouth, posting flyers in targeted areas, and sending text messages via partnerships with the Special Supplemental Nutrition Program for Women, Infants, and Children partnership. First, participants are screened through a REDCap eligibility survey. REDCap is a secure web application designed to support data capture for research studies. If the participant meets the inclusion criteria, they are invited to complete an informed consent form and a demographics and contact information survey to schedule the initial visit. All materials are available both in English and Spanish.

Instruments and Surveys

There are two types of survey instruments: (1) screening and characterization, and (2) data collection. The screening instruments are used to determine eligibility and confirm living arrangements (Table 1), while the survey instrument provides valid, reliable self-report data (Table 2). All instruments are

XSL•FO
RenderX

hosted on REDCap and distributed by email, through QR code, or by text message.

Table 1. Screening and characterization.

Survey name	Measured constructs and validity
Eligibility	A total of 10 response (yes or no) questions listing the inclusion criteria. Automatically scored. If eligible, the participant is sent to the study dashboard to read the study overview and consent to be contacted by the research team.
Consent	Participants are provided with an overview of the study and confirm their willingness to participate. Consent is again confirmed through scheduling meetings and during the on-site visit.
Demographics	Information about themselves (eg, age) and their living situation (eg, how many people live with them).

Table 2. Data collection.

Survey name	Measured constructs and validity
Medical history (laboratory visit 1)	A total of 25 forced response (yes or no) questions listing general medical problems, conditions, or behaviors. Five questions had follow-up prompts (eg, smoking, when did you quit smoking?).
Current medications (laboratory visit 1)	A total of 20 forced response (yes or no) questions with a follow-up prompt and an open text box when there was a "yes" reply.
Neighborhood Factors Questionnaire (laboratory visit 1)	A total of 7 subscales of walkability, aesthetic quality, safety, violence, availability of healthy foods, social cohesion, and activities with neighbors as factors associated with physical activity [24].
Groningen Sleep Quality Scale (laboratory visit 1)	A total of 15 true or false questions about sleep quality (eg, I slept well last night) [25].
Mini MASQ[a] (laboratory visit 1)	A total of 26 Likert scale questions about perceptions [26,27].
IPAQ[b] (laboratory visit 1)	A total of 27 questions about physical activity duration and intensity enable the calculation of the metabolic equivalent of task-minutes as a measure of physical activity volume [28].
Everyday Discrimination Scale–Everyday Experiences Source (laboratory visit 1)	This brief survey consists of 10 Likert scale questions that identify the frequency of perceived discrimination (eg, never to almost every day) as an indicator of stress and an increased risk for MetS[c]. The following prompt identified the potential reason they thought this occurred (gender, body image, etc) [29].
Home toolkits	The toolkits contain a commercial Fitbit smartwatch (Fitbit Inc) and a custom-built air quality sensor.
EMA[d]	EMA involves the repeated sampling of the participants' behaviors and experiences in real-time in natural environments. EMAs aim to minimize recall bias and maximize ecological validity. Once the toolkits are deployed, EMAs are sent via mobile phones to confirm study compliance and to document whether any unusual circumstances occurred. Participants receive a comprehensive set of surveys during the first days of data collection and an EMA on Days 3, 6, 9, and 12 for a total of 4 EMAs [30].

[a]MASQ: Mood and Anxiety Symptoms Questionnaire.

[b]IPAQ: International Physical Activity Questionnaire.

[c]MetS: metabolic syndrome.

[d]EMA: Ecological Momentary Assessment.

Procedures

Once consent is confirmed, the participants are scheduled for their first in-person health screening. Before arriving, they are asked to complete the survey data collection through REDCap (Table 2). Participants can also volunteer to be screened for eligibility for the MRI phase of the study, which is a second laboratory visit.

Laboratory Visit 1: Health Screening and Toolkit Deployment

During the health screening, we measure anthropometric parameters such as body composition, blood pressure, lipids, and glucose levels. Height, weight, BMI, fat mass percentage, waist circumference, and visceral adipose tissue are measured using a Seca body composition scale (seca mBCA 555/554; seca GmbH & Co KG) through bioelectric impedance. The participant is asked to stand on the platform barefoot for around a minute. For blood pressure measurements, an Omron 5 series upper arm pressure monitor (VP-2000; Omron Healthcare) is used to measure brachial systolic, diastolic, and pulse. It is assessed in the seated position after a 10-minute rest. Three measurements are taken, with a 3-minute interval between each measurement. Additionally, a blood screening using a Cholestech LDX (Alere) is conducted to measure participants' blood lipids and fasting glucose levels. The participant must have fasted for 8 hours before performing the finger prick. A finger prick is done using a lancet on the side of the center finger in the nondominant hand. The sample is collected through a capillary tube of 40 µL (Ref 52193, Abbott) and dispensed with a capillary plunger (Ref 10311, Abbott) into a Cholestech LDX Lipid Profile Glucose Cassette (Ref 97991, Abbott). Values obtained by the Cholestech LDX include blood concentrations of total cholesterol, triglycerides, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, and glucose.

XSL•FO
RenderX

Data Return

In line with best practices in health research and to strengthen rapport with our participants, we provide a concise report of the health screening information [31]. We used the publicly licensed R Markdown (RStudio, PBC) to create personalized, automated reports (Figure 2) [32,33].

Figure 2. Sample health screening data return. Every participant receives a report with their health measurement results.

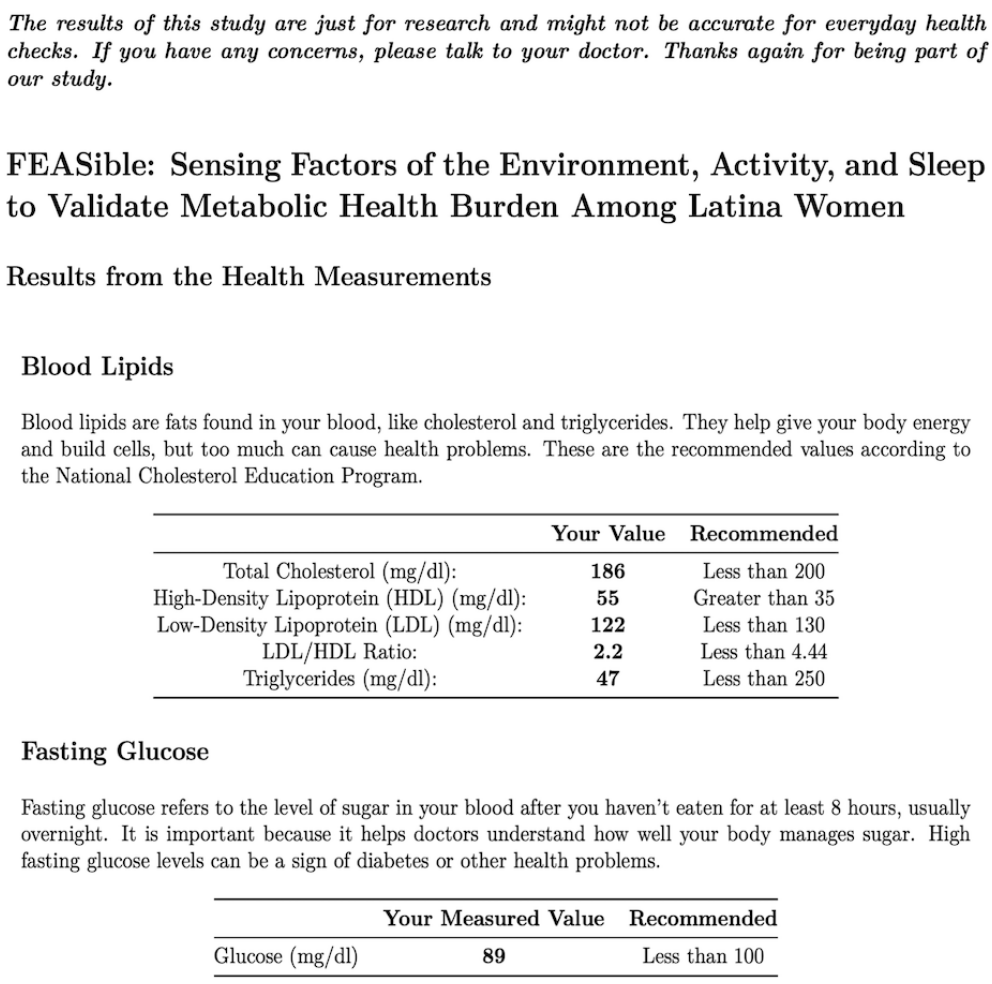

The results of this study are just for research and might not be accurate for everyday health checks. If you have any concerns, please talk to your doctor. Thanks again for being part of our study.

FEASible: Sensing Factors of the Environment, Activity, and Sleep to Validate Metabolic Health Burden Among Latina Women

Results from the Health Measurements

Blood Lipids

Blood lipids are fats found in your blood, like cholesterol and triglycerides. They help give your body energy and build cells, but too much can cause health problems. These are the recommended values according to the National Cholesterol Education Program.

	Your Value	Recommended
Total Cholesterol (mg/dl):	**186**	Less than 200
High-Density Lipoprotein (HDL) (mg/dl):	**55**	Greater than 35
Low-Density Lipoprotein (LDL) (mg/dl):	**122**	Less than 130
LDL/HDL Ratio:	**2.2**	Less than 4.44
Triglycerides (mg/dl):	**47**	Less than 250

Fasting Glucose

Fasting glucose refers to the level of sugar in your blood after you haven't eaten for at least 8 hours, usually overnight. It is important because it helps doctors understand how well your body manages sugar. High fasting glucose levels can be a sign of diabetes or other health problems.

	Your Measured Value	Recommended
Glucose (mg/dl)	**89**	Less than 100

The reports are generated using an R code that references data from the REDCap report function and creates a ready-to-send document for each participant. The report is designed with ease of comprehension in mind, written at an eighth-grade reading level, and includes tables that provide recommended biomarker ranges. They are used as an informational tool for participants to understand their engagement with our research. Health recommendations are compiled from the National Cholesterol Education Program (National Heart, Lung, and Blood Institute) and are publicly available information. Our participants understand that this information is for research purposes only and not intended for clinical diagnosis. If they request, we provide additional health literacy resources from the American Heart Association regarding MetS. We consider this a key component of our engagement with participants and essential in establishing trust and transparency with our participant community.

Toolkit Deployment

After measuring weight and height as part of the health screening, the participant downloads the Fitbit phone app, where the values are used to accurately measure the calories burned throughout the day, step counts, and heart rate. A commercial Fitbit Inspire 3 (Figure 3) is deployed for use by the participant by entering age, sex, and body weight.

XSL•FO
RenderX

Figure 3. The sensing devices used for the FEASible Study in the home environment sensor kit (left) and Fitbit Inspire 3 (right).





The research assistant provides instructions on how to use the Fitbit, including how to navigate the Fitbit App dashboard, charge and maintain the battery, and verify that the device is collecting data accurately. Additionally, participants can try several Fitbit bands to find the one that suits them best in terms of tightness. Participants can track their daily exercise, calorie intake, distance walking or running, and sleep cycles. Sleep cycles are categorized into 4 stages: awake, light, deep, and rapid eye movement.

The participant is expected to wear the Fitbit continuously for 14 days. Participants are instructed to remove the device while bathing and to charge it for 30 minutes. Since the Fitbit Inspire 3 has excellent battery life, it requires only 1 or 2 charging sessions in 2 weeks. Of the 178 participants completed by August 2024, a total of 148 (83%) wore the Fitbit for at least 14 continuous days (mean 17.85 days). The Fitbit device was configured to capture an energy expenditure reading every minute; thus, it would return 24×60=1440 readings per day if worn the entire day without interruption. We calculated a completeness metric for each participant by dividing the actual number of calorie readings by 1440. Out of the 148 participants who wore the Fitbit for at least 14 continuous days, the mean completeness was 93.45%. These numbers indicate that we have collected at least 2 weeks of continuous Fitbit data.

While Fitbit devices are not considered clinical gold standards, they provide accessible and scalable measures that serve as potential proxy features for modeling health risk patterns. To ensure data reliability, we will apply multiple data cleaning and preprocessing steps. Specifically, days with zero variation in minute-by-minute calorie data will be removed, which likely indicate nonwear periods. For each participant, a completeness metric will be calculated, representing the percentage of nonmissing readings out of the total 1440 minutes per day. Only days with at least 50% valid data will be retained for subsequent analyses. Minor gaps within valid days will then be interpolated to maintain temporal continuity. These procedures will help ensure that wearable data used in modeling reflect actual participant activity and sleep behavior rather than device artifacts or missing data.

For the home environment sensor kit, a custom-built environmental sensor (noncommercial device) (Figure 3) is given to the participant to place in their home at the same time as Fitbit usage. The participant is instructed to plug the home environment sensor kit into an AC outlet and to place the device at dresser height (or equivalent) in the bedroom for 1 week and on a countertop (away from the stove) in the kitchen for the second week.

Air Quality Sensors

The home environment sensor kit is an inexpensive, custom-made, real-time sensor and logger developed in the laboratory. It collects continuous temperature, relative humidity, CO_2, and $PM_{2.5}$ time-series data at 1-minute sampling intervals. Version 2 of the environmental sensor is deployed in the FEASible study. Version 1 previously successfully collected data to evaluate the association between sleep and indoor air quality among university students [34].

The home environment sensor kit comprises 5 sensors (Table 3): one Sensirion SCD-30 for CO_2, which has integrated temperature and relative humidity sensors as well, two Sensirion SPS-30 for $PM_{2.5}$, one Honeywell HIH-8120-021 to measure relative humidity, and one TTS-1KC3-BZ thermistor for temperature measurement. Every minute, a measurement is logged on a spreadsheet. Low-cost sensors offer several

XSL•FO
RenderX

advantages, including small size, mobility, lightweight design, affordability, and ease of use, which enable an increase in spatial and temporal resolution [35]. Meanwhile, measurement uncertainty is more significant, and frequent calibration is needed. Particle sensors are typically provided with manufacturer-installed calibrations that depend upon the manufacturer's specific calibration condition; information about these conditions is usually unavailable. Individual calibration of each sensor under controlled conditions is necessary to characterize uncertainty in the data and improve cross-device accuracy. The first calibration round of 20 home environment sensor kits was performed in a controlled environmental chamber against research-grade instruments: a TSI APS 3321 aerodynamic particle sizer and a LI-COR LI-7000 CO_2 analyzer. CO_2 and $PM_{2.5}$ calibration data were collected simultaneously using Arizona Test Dust (12103-1, A3 Medium Test Dust, Powder Technology Inc, PTI ID 14891M) as a standard for $PM_{2.5}$ calibration through a spike and decay process.

Table 3. Home environment sensor kit specifications.

Sensor and type name	Sensing	Range	Characteristics
Sensirion SCD-30 (transmissive NDIR[a])	CO_2[b]	400 ppm to 10,000 ppm	Accuracy: ±(30 ppm +3%)
Sensirion SPS-30 (optical, laser scattering)	$PM_{2.5}$[c]	0.3 µm to 2.5 µm	Precision: ±100 #/cm³ for range 0-1000 #/cm³; +10% mv[d] for range 1000-3000 #/cm³
Honeywell HIH-8120-021 (thermoset-polymer capacitive sensing element)	Humidity	0% to 100% RH[e]	Accuracy: ±2.0% RH
TTS-1KC3-BZ (thermistor with a Maxim MX7705 ADC[f] thermistor)	Temperature	–40 °C to 125 °C	Tolerance (0 °C to 70 °C): ±0.2 °C

[a]NDIR: nondispersive infrared.

[b]CO_2 : carbon dioxide.

[c]$PM_{2.5}$: particulate matter ≤2.5 µm in diameter.

[d]mv: measured value.

[e]RH: relative humidity.

[f]ADC: analog-to-digital converter.

The home environment sensor kit design enables users to verify the validity of the measurements from different sensors. There are 2 particulate matter sensors, 2 temperature sensors, and 2 relative humidity sensors. Therefore, if one drifts unusually and unexpectedly, it can be caught, replaced, and recalibrated before the next deployment. We performed regular calibrations through the first year of deployment.

Laboratory Visit 2: MRI and Blood Draw Visit

After the toolkit deployment visit, participants are asked for interest in an additional visit for MRI scanning and blood draw. Those responding positively are checked for the absence of contraindications to MRI over the phone. To minimize distortion, stricter exclusion criteria are applied to any metal implant in proximity to the head, even if the material is MRI-safe. Eligible participants obtain detailed information about the visit procedures and compensation. Once the participant agrees, an appointment is scheduled for the morning, lasting approximately 1.5 hours. A consenting subgroup of 225 women will participate in the neuroimaging component of the study, which includes a comprehensive MRI scan session and an in-laboratory blood draw.

On the day of the MRI scan and blood draw visit, participants are instructed to fast for 8 hours before the blood draw. Plasma and serum blood samples are collected from the antecubital vein through venipuncture and centrifuged at room temperature for 10 minutes at 1300 rpm in a Centrifuge 5810 (Eppendorf AG). However, the plasma tube is centrifuged immediately after collection, and the serum tube is left at room temperature for 30 minutes before centrifugation. After centrifugation, 0.5 mL of the supernatant from plasma and serum tubes is transferred to three 12 × 75-mm tubes and immediately stored at –80 °C for later analysis of inflammatory markers.

After the blood draw, the participant is escorted to the MRI scan. The MRI scans are conducted using a 3-Tesla PRISMA scanner (Siemens Healthineers). Upon arrival, participants undergo a comprehensive MRI safety screening and receive a detailed explanation of the procedures. Participants are asked to remove any personal clothing and accessories and change into medical scrubs to eliminate any potential sources of metal interference. During the scan, participants are instructed to remain as still as possible and keep their eyes open during the pseudo-continuous arterial spin labeling (pcASL) and blood oxygenation level–dependent functional MRI (BOLD fMRI) sequences. They are encouraged to communicate any discomfort or emergencies to the research staff at any time and are informed of their right to terminate the session if needed.

The protocol includes the following sequences: T1-weighted magnetization prepared rapid gradient echo (MPRAGE; Echo Time [TE]=2.01 ms, repetition time [TR]=2000 ms, inversion time [TI]=900 ms; flip angle=8 degrees; voxel dimension=1.0 × 1.0 × 1.0 mm³; field of view [FOV]=256 × 256 mm; total acquisition time [TA]=4:40); T2-weighted 2D fluid-attenuated inversion recovery (FLAIR; TE=91 ms, TR=9000 ms, TI=2500 ms; voxel dimension=0.9 × 0.9 × 3.0 mm³; FOV=220 × 220 mm; TA=4:05); magnetic resonance spectroscopy (MRS) localizer (TE=1.85 ms, TR=1280 ms, TI=725 ms; voxel dimension=0.4 × 0.4 × 1.1 mm³; FOV=224 × 210 mm; TA=2:02); single voxel spectroscopy (SVS) spin echo (SE)

XSL•FO
RenderX

spectroscopy, water suppressed (TE=30 ms, TR=4750 ms; voxel dimension=31 × 16 × 16 mm³; bandwidth=1600 Hz; TA=5 seconds to 7 seconds; water suppression=weak water suppression); SVS SE spectroscopy, unsuppressed (TE=30 ms, TR=4750 ms; voxel dimension=31 × 16 × 16 mm³; bandwidth=1600 Hz; TA=5 seconds to 7 seconds; water suppression=RF off); and unsuppressed relaxation-resolved advanced multi–spin-echo chemical shift (RRAMSC) sequences (TE=30 ms to 260 ms, TR=1030 ms to 2150 ms voxel dimension=31 × 16 × 16 mm³; bandwidth=1600 Hz; TA=5 seconds to 7 seconds). The total acquisition time was approximately one hour.

Ethical Considerations

This study protocol is approved by the University of Texas's Institutional Review Board (STUDY00003675). To date, 12 Institutional Review Board–approved modifications have been made, primarily to add study personnel, adjust compensation procedures, and clarify protocol details. The informed consent process is obtained in two phases. First, participants complete an eligibility survey; those eligible are prompted to review and electronically sign the informed consent form before proceeding. Second, at the initial laboratory visit, informed consent is reviewed again, and study staff assess comprehension and address any questions. To ensure participant confidentiality, only authorized individuals may access identifiable information for research oversight. However, all participants are assigned a code, not their name. Lastly, participants are compensated by gift cards: US $50 for survey completion, US $200 for contributing sensor data, and an additional US $150 for participants who are eligible and consent to neuroimaging and blood collection. Participants who withdraw early will receive prorated compensation.

Results

Overview

Year one data and selected cases are reported as initial evidence of the procedure used in the FEASible research study. Data collection started in 2023 and will continue until 2028. Some variables in this section are reported as aggregate scores, including eligibility, demographics, and survey summary scores, whereas other variables display results from a single case to illustrate the collected data.

Out of the total of 805 participants screened for eligibility, a total of 225 consented to participate in the study and complete all surveys (Figure 4). Of these, 204 completed the health screening and received sensor toolkits for collecting wearable and home environment data over the 2 weeks. Lastly, a subset of these participants (n=47 and n=44) underwent MRI scans and blood draws, respectively. Table 4 summarizes demographics collected from the initial eligibility survey; these data reflect potential participants, not those who completed the full study.

Figure 4. The FEASible study CONSORT (Consolidated Standards of Reporting Trials) diagram. The study is divided into 3 main categories: online surveys, health screening, 2-week remote sensor data collection, and magnetic resonance imaging (MRI) and blood draw visits. The n represents the number of participants. Created with BioRender by Ramos et al [36], reproduced using license obtained by authors.

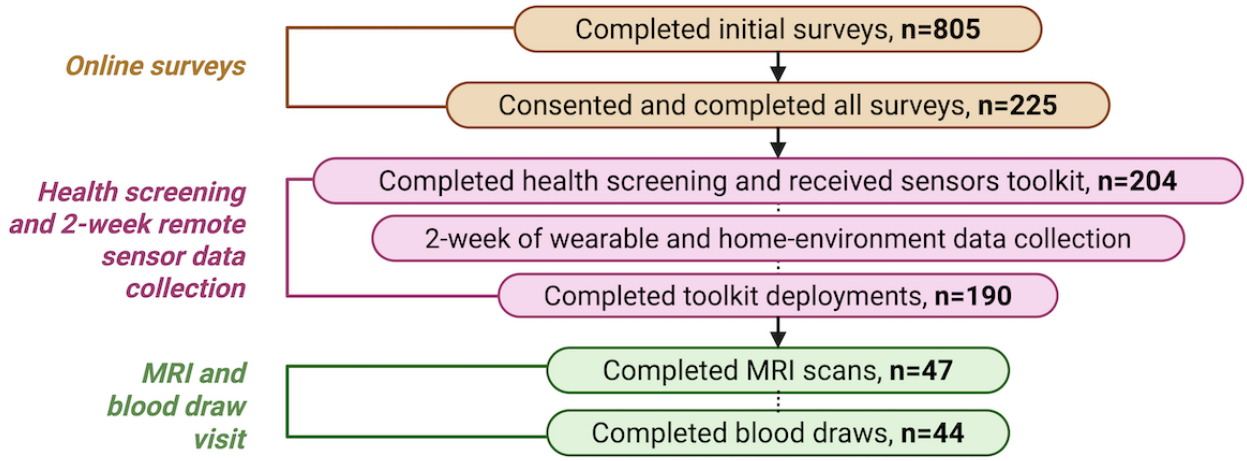

XSL•FO
RenderX

Table 4. Participant demographics from Year one initial eligibility surveys. Percentages are calculated within each demographic factor. Totals within a factor may not equal totals in other factors due to missing responses or participants selecting multiple categories.

Factors	Frequency, n (%)
Age (years)	
18-25	169 (49)
26-33	126 (36)
34-40	52 (15)
Latina heritage	
Yes	210 (60)
No	135 (39)
Do not know	3 (1)
Latina origin	
Central American	23 (11)
Cuban	5 (2)
Dominican	1 (1)
Mexican	138 (63)
Puerto Rican	13 (6)
South American	17 (8)
Spanish or Spaniard	11 (5)
Other	9 (4)
Race	
White	249 (72)
Asian	36 (10)
More than one race	39 (11)
Black or African American	18 (5)
American Indian or Alaska Native	6 (2)
Native Hawaiian or Pacific Islander	0 (0)
Education	
Did not graduate from high school	7 (2)
High school diploma or GED[a]	77 (22)
Some college	104 (30)
College degree	159 (46)

[a]GED: General Educational Development.

Fitbit Data

To show between-subject variance in MetS-indicative daily behavior, we computed 2 example metrics—namely *mean calorie per minute* (calorie.mean) and *proportion of the time spent not moving* (portion.static)—for each of the 148 participants who returned at least 14 continuous days of data, over each person's data collection period. We considered a given minute static if the difference between the current and the previous minute's calorie expense was zero. Figure 5 shows the histogram of these two metrics; vertical, dashed blue lines indicate mean value. The average woman expended mean of 1.42 calories per minute (SD 0.24 calories), which translates to an average of 2048 calories per day (SD 348 calories), and remained static 69% of the time (SD 6.7%).

To illustrate within-person variance in energy expenditure, we plotted the time series of minute-by-minute calorie readings for a sample participant in Figure 6A, where the meaningful fluctuations reflect the daily cycle. Figures 6B and 6C are two different days extracted from the participant's data collection period. Both exhibit an initial period of nonmovement, indicating sleep time, followed by activity spikes at various points throughout the day, which last for different durations.

XSL•FO
RenderX

Figure 5. Histogram of mean calorie (left) and portion static (right).

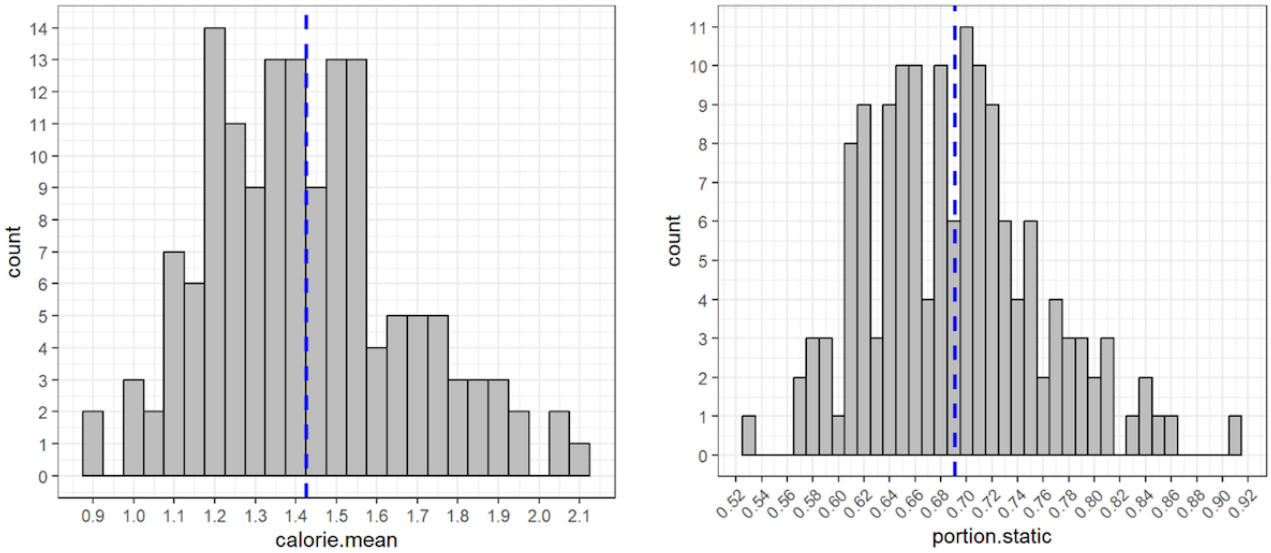

XSL•FO
RenderX

Figure 6. Time series of minute-by-minute energy expenditure of a participant: (A) over the entire data collection period, December 5–19, 2023; (B) December 9, 2023; and (C) December 13, 2023.

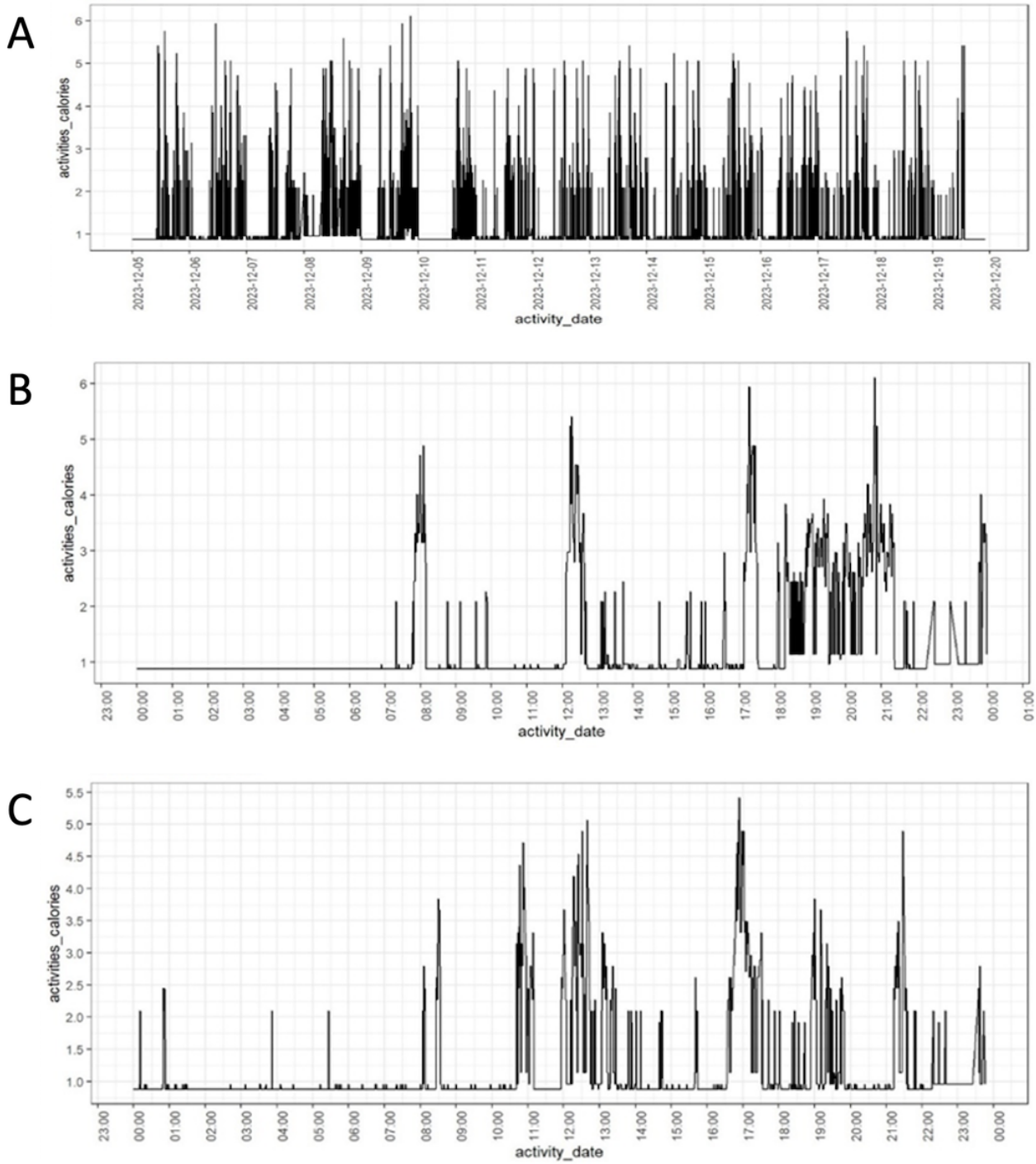

MRI Data

We have demonstrated the feasibility and tolerability of the MRI protocol by completing 47 scans. The representative results of the MRI segmentation and magnetic resonance spectrum are shown below, demonstrating excellent data quality (Figure 7).

XSL•FO
RenderX

Figure 7. Magnetic resonance imaging brain measurements. FreeSurfer gray matter segmentation of T1-weighted magnetization-prepared rapid gradient-echo structural images (left). Linear Combination Model results from one participant's magnetic resonance spectroscopy data (right). Cho: choline+phosphocholine; Cr: creatine+phosphocreatine; Glu: glutamate; mI: myo-inositol; NAA: N-acetylaspartate.

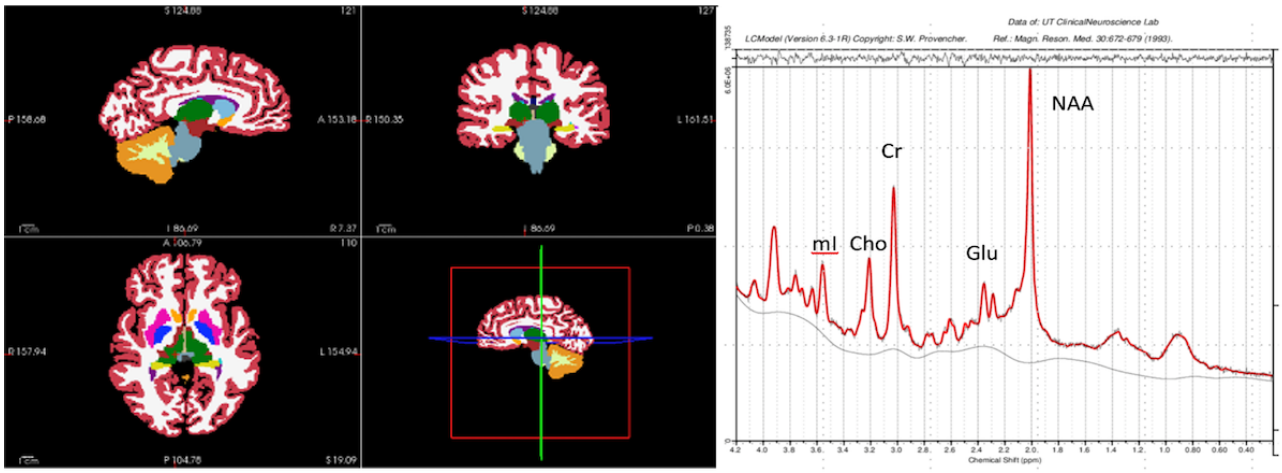

Air Quality Sensor Data

We have successfully gathered CO_2 and $PM_{2.5}$ concentration measurements during the 2-week data collection period, as shown in Figure 8. $PM_{2.5}$ levels increase during dinner time, hypothetically associated with the occurrence of cooking.

Figure 8. The home environment sensor kit measurements. Calibrated carbon dioxide (CO_2; ppm) household indoor air measurements (left) and particulate matter ≤2.5 µm in diameter ($PM_{2.5}$; µg/m³) household indoor air measurements. In the bedroom (week 1: left of the dashed axis), then the kitchen (week 2: right of the dashed axis).

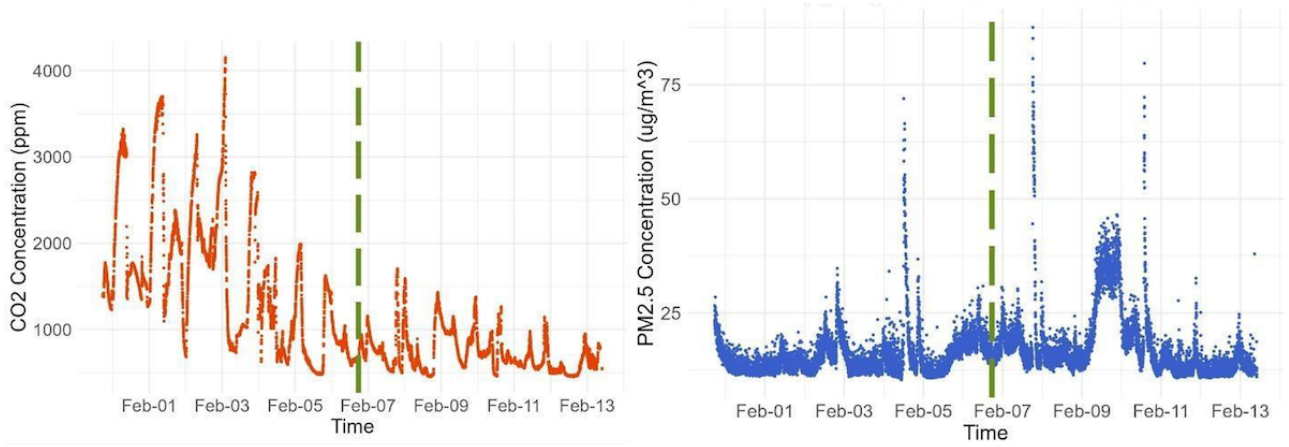

Risk for Metabolic Syndrome

One aim of this study is to provide case examples of data from the first year of data collection, with the goal of promoting the standardization of measurements. This section presents 2 illustrative cases from the FEASible study, which evaluates risk for MetS.

Case 1: A High-Risk 39-Year-Old Non-Hispanic White Woman

At age 39, this participant is a single mother who struggles with sleep (less than 5 hours per night as measured by the Groningen Sleep Scale) because she cares for a teething toddler. She self-identifies as a non-Hispanic White woman with a background in performing arts, fine arts, and high school sports (eg, volleyball). She has a BMI of 44.8 kg/m^2 and a waist circumference of 43 inches. Health screenings suggest she may have MetS and is at risk for CVD. Her common laboratory assessments reveal prehypertensive to hypertensive blood pressure (130/100 mm Hg), borderline high total cholesterol (206 mg/dL), and elevated triglycerides (231 mg/dL). Her self-reported health history revealed diagnosed fatty liver disease, frequent upper respiratory infections, and irregular thyroid function. She lives in a neighborhood with no sidewalks and limited access to healthy food options (Neighborhood Factors Questionnaire). Further, this participant reported experiencing frequent discrimination related to her weight (245 lb) and gender (Everyday Discrimination Scale). Despite working as a nanny, which involves some days of moderate physical activity (self-reported International Physical Activity Questionnaire), she confirmed that she does not meet the weekly or recommended amount of physical activity duration, steps, or intensity. Although she is currently receiving health care, she is at significant risk of a prolonged poor health span. This participant represents our standard practice for addressing health risks: wait for symptomatology, then seek urgent or emergency care. In contrast, our second case reflects the potential for early intervention among young adults. By using wearable devices

XSL•FO
RenderX

and home sensors, we can develop statistical models that link the magnitude of risk and identify points of intervention (eg, home exposures and physical activity).

Case 2: 20-Year-Old Hispanic Woman

The second case involves a 20-year-old Hispanic woman, born in Mexico, who resides in Texas in the same neighborhood as the first case. While she faces similar environmental challenges, including outdoor noise and a lack of access to sidewalks, parks, and healthy foods (as indicated by the Home Air Quality and Neighborhood Factors Questionnaire), her health profile differs significantly. She reports sleeping well, averaging 8.5 hours of sleep per night (as measured by the Groningen Sleep Scale, Ecological Momentary Assessments, and Fitbit data). Her BMI of 26.8 kg/m^2 places her in the overweight category, but all other biomarkers are within healthy ranges. According to the International Physical Activity Questionnaire, she spends most of her awake time sitting, with occasional walking but limited moderate-to-vigorous physical activity. Fitbit data provide further evidence of low activity levels. Over the 17-day observation period (July 8-24, 2024), her average daily energy expenditure was 1882 calories (1.31 calories per minute), compared to the cohort average 2048 calories, indicating lower-than-average activity levels. However, her average daily static percentage was 68% (percentage of time spent not moving during a 24-hour period), no higher than the cohort average of 67%.

By comparing these two cases, which represent mid-life and emerging adult developmental stages, we aim to refine algorithms for estimating MetS risk and brain vulnerability. This approach will help identify behavioral and environmental variables contributing to health disparities, enabling the development of tailored and responsive interventions.

Discussion

Overview

The FEASible study explores women's daily living patterns through wearable devices and home environment sensor kit to validate the use of behavioral and environmental data (eg, physical activity, sleep, and indoor air quality) as indicators of risk for MetS and CVD (National Heart, Lung, and Blood Institute; grant R01HL168374). This investigation is designed to provide a rationale for the standard operating procedures and provide a case example of data from the first year of data collection leading to artificial intelligence and machine learning modeling of health risk, thus informing future interventions focused on MetS, heart disease, and all-cause mortality among Latina women. This study consists of three steps on data collection from participants for one month: (1) self-reported demographics, health, and lifestyle; (2) health screening and toolkit deployment with sensors for 2-week data sensing; and (3) neuroimaging and blood draw. The preliminary results of this study provide valuable insights into the participants' demographics and the progress of data collection. Findings are at an initial stage. Nevertheless, the study has successfully recruited a diverse pool of participants, with a particular focus on Latina women. The primary conclusion of this study is that, upon completion of the first year of data collection, FEASible is feasible, has been implemented as intended, and is ready to scale up**.**

The participants recruited are diverse in terms of age groups, ethnicities, races, and educational backgrounds. Most participants fall within the age range of 18-25 years, which aligns with the study's focus on young adults. However, the lower representation of participants aged 34-40 years warrants targeted recruitment efforts to ensure adequate representation of this age group in the final sample. The most helpful recruitment strategy has been word of mouth among participants. Participants share the study with friends and relatives within their communities, including the University of Texas at Austin. However, consistent with previous studies, recruitment by word of mouth rarely achieves a high sample size, highlighting the need for alternative recruitment strategies [37]. Another recruitment strategy has been posting about the study in Latina-focused Facebook groups, which has helped attract more participants from diverse educational and age backgrounds. Posting on existing Facebook groups and pages has efficiently generated responses from potential participants, particularly since the COVID-19 pandemic [38]. Another recruitment strategy will be to work directly with a community-based health clinic and volunteer at community events.

Community engagement is a valuable asset for this study, as our ultimate goal is to develop potential interventions to address future challenges related to CVD and MetS. Community engagement must be a two-way exchange process between the university and its communities, aiming to facilitate transparent and ongoing communication for study participants [39]. Therefore, the research team will continue to volunteer within the community, offering potential participants talks on CVD and brain health, as well as resources. Our dissemination plan involves providing health screening assessment letters to participants, which they can share with their primary health care practitioner. Moreover, participants can track their physical activity and sleep patterns using wearable devices and obtain preliminary information about the air quality in their homes.

The proportional representation of Latina and non-Latina participants has numerous advantages, as it aims to address the unique challenges and health concerns this population faces. Moreover, the diversity of Latin origin, with participants from various regions, further enriches the findings and provides an opportunity to evaluate the Latina population as disaggregated data, illuminating the community's diversity. This level of granularity is crucial, as previous research has highlighted significant heterogeneity in health outcomes and CVDs and MetS risk factors among different Hispanic and Latino subgroups [40-42]. By capturing this diversity, the study can contribute to a more nuanced understanding of the complex interplay between cultural, environmental, and behavioral factors influencing cardiovascular and metabolic health in Latina women.

Sensing devices address a critical need for cost-effective, accessible, objective, and continuous monitoring of health-related factors. Poor physical activity and sleep patterns have been consistently associated with an increased risk of MetS

XSL•FO
RenderX

in women [43]. However, the association of environmental factors, such as air quality and indoor pollutants, on cardiovascular and cognitive health has been understudied, particularly in Latina women.

It is important to note that the FEASible study is cross-sectional in design, which cannot establish causal relationships between behavioral or environmental factors and MetS. Findings can only identify associations between these variables. Future longitudinal and experimental studies are planned to examine temporal and causal pathways linking lifestyle behaviors, environmental exposures, and cardiometabolic outcomes.

The study's findings have the potential to inform future experimental or clinical studies investigating causal relationships and developing targeted interventions for CVDs and MetS prevention among Latina women. Additionally, validating low-cost and accessible monitoring devices may pave the way for their integration into clinical practices, enabling personalized and continuous monitoring of health risks.

Conclusions

The FEASible study addresses a critical gap in understanding the complex interplay of factors contributing to CVDs and MetS among Latina women, a population disproportionately affected by these conditions. The comprehensive approach to evaluating physical activity, sleep patterns, and environmental factors aligns with the need for tailored prevention strategies. Using low-cost sensing devices is a promising approach to capturing real-world data, but validation is necessary to assess reliability and accuracy to ensure the quality and interpretability of the collected data. Moreover, this study emphasizes the importance of developing affordable and precise devices to accurately monitor individual behaviors and environmental risk factors, thereby enabling modeling approaches that assess health risks. It can aid the development of "sequential multiple assignment randomized trials" interventions to address MetS and CVDs, as well as their cognitive consequences.

Acknowledgments

We thank our team members for their contributions to and review of this study protocol. The authors attest that no generative artificial intelligence tools were used to create any portion of this manuscript. All text, tables, and figures were authored solely by the listed authors.

Funding

This work is supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under award number 1R01HL16837-01 for the project *FEASible: Sensing Factors of Environment, Activity, and Sleep to Validate Metabolic Health Burden Among Latina Women*. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The data collection of the FEASible Study is still in progress. The datasets generated during this study are not publicly available. Interested researchers may contact the Principal Investigators to request access. Data requests will be reviewed on a case-by-case basis following NIH guidelines and participant confidentiality protections.

Authors' Contributions

TR-S led project coordination and administration, methodology, manuscript drafting, review, and supervision. EH, YL, and CC contributed to data collection, data curation, writing, and data visualization. CW, KK, DC, and AH served as principal investigators, providing conceptualization, methodology, funding acquisition, supervision, and guidance on study design and implementation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from the CIDH - Clinical Informatics and Digital Health Study Section, Healthcare Delivery and Methodologies Integrated Review Group (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 148 KB - resprot_v15i1e82034_app1.pdf]

References

1. New Estimates Highlight Differences in Growth Between the U.S. Hispanic and Non-Hispanic Populations. URL: https://www.census.gov/newsroom/press-releases/2024/population-estimates-characteristics.html [accessed 2025-12-06]
2. Krogstad JM, Passel JS, Moslimani M, Bustamante LN. Key Facts About U.S. Latinos for National Hispanic Heritage Month. 2023. URL: https://www.pewresearch.org/short-reads/2023/09/22/key-facts-about-us-latinos-for-national-hispanic-heritage-month/ [accessed 2025-12-06]

XSL•FO
RenderX

3. Health of Hispanic or Latino population. FastStats. URL: https://www.cdc.gov/nchs/fastats/hispanic-health.htm [accessed 2025-12-06]
4. Tsao CW, Aday AW, Almarzooq ZI, Anderson CAM, Arora P, Avery CL, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart disease and stroke statistics-2023 update: a report from the American Heart Association. Circulation 2023;147(8):e93-e621. [doi: 10.1161/CIR.0000000000001123] [Medline: 36695182]
5. Guembe MJ, Fernandez-Lazaro CI, Sayon-Orea C, Toledo E, Moreno-Iribas C, RIVANA Study Investigators. Risk for cardiovascular disease associated with metabolic syndrome and its components: a 13-year prospective study in the RIVANA cohort. Cardiovasc Diabetol 2020;19(1):195 [FREE Full text] [doi: 10.1186/s12933-020-01166-6] [Medline: 33222691]
6. Wilson PWF, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic syndrome as a precursor of cardiovascular disease and type 2 diabetes mellitus. Circulation 2005;112(20):3066-3072. [doi: 10.1161/CIRCULATIONAHA.105.539528] [Medline: 16275870]
7. Martin SS, Aday AW, Almarzooq ZI, Anderson CAM, Arora P, Avery CL, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. 2024 heart disease and stroke statistics: a report of US and global data from the American Heart Association. Circulation 2024;149(8):e347-e913 [FREE Full text] [doi: 10.1161/CIR.0000000000001209] [Medline: 38264914]
8. Campbell B, Aguilar M, Bhuket T, Torres S, Liu B, Wong RJ. Females, Hispanics and older individuals are at greatest risk of developing metabolic syndrome in the U.S. Diabetes Metab Syndr 2016;10(4):230-233. [doi: 10.1016/j.dsx.2016.06.014] [Medline: 27342003]
9. Gomez S, Blumer V, Rodriguez F. Unique cardiovascular disease risk factors in Hispanic individuals. Curr Cardiovasc Risk Rep 2022;16(7):53-61 [FREE Full text] [doi: 10.1007/s12170-022-00692-0] [Medline: 35669678]
10. Suglia SF, Abraido-Lanza AF, Guerrero-Preston RE, Ramos KS. Integration of sociocultural and behavioral factors into the clinical framework of cardiovascular studies in Hispanic/Latino populations: relevance during the SARS-COV-2 pandemic. J Clin Transl Sci 2021;5(1):e101 [FREE Full text] [doi: 10.1017/cts.2021.20] [Medline: 34192056]
11. Peila R, Xue X, Feliciano EMC, Allison M, Sturgeon S, Zaslavsky O, et al. Association of sleep duration and insomnia with metabolic syndrome and its components in the women's health initiative. BMC Endocr Disord 2022;22(1):228 [FREE Full text] [doi: 10.1186/s12902-022-01138-9] [Medline: 36104689]
12. Li X, Xue Q, Wang M, Zhou T, Ma H, Heianza Y, et al. Adherence to a healthy sleep pattern and incident heart failure: a prospective study of 408 802 UK biobank participants. Circulation 2021;143(1):97-99 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.120.050792] [Medline: 33190528]
13. Koniak-Griffin D, Brecht M, Takayanagi S, Villegas J, Melendrez M. Physical activity and cardiometabolic characteristics in overweight Latina women. J Immigr Minor Health 2014;16(5):856-864 [FREE Full text] [doi: 10.1007/s10903-013-9782-z] [Medline: 23355122]
14. Sharkey P. Homebound: the long-term rise in time spent at home among U.S. adults. SocScience 2024;11:553-578. [doi: 10.15195/v11.a20]
15. Kühn S, Mascherek A, Filevich E, Lisofsky N, Becker M, Butler O, et al. Spend time outdoors for your brain - an in-depth longitudinal MRI study. World J Biol Psychiatry 2022;23(3):201-207 [FREE Full text] [doi: 10.1080/15622975.2021.1938670] [Medline: 34231438]
16. Wu C, Fritz H, Bastami S, Maestre JP, Thomaz E, Julien C, et al. Multi-modal data collection for measuring health, behavior, and living environment of large-scale participant cohorts. Gigascience 2021;10(6) [FREE Full text] [doi: 10.1093/gigascience/giab044] [Medline: 34155505]
17. He Q, Ji XJ. The labor productivity consequences of exposure to particulate matters: evidence from a Chinese national panel survey. Int J Environ Res Public Health 2021;18(23) [FREE Full text] [doi: 10.3390/ijerph182312859] [Medline: 34886582]
18. Kim K, Kabir E, Kabir S. A review on the human health impact of airborne particulate matter. Environ Int 2015;74:136-143. [doi: 10.1016/j.envint.2014.10.005] [Medline: 25454230]
19. Huang B, Xiao T, Grekousis G, Zhao H, He J, Dong G, et al. Greenness-air pollution-physical activity-hypertension association among middle-aged and older adults: evidence from urban and rural China. Environ Res 2021;195:110836. [doi: 10.1016/j.envres.2021.110836] [Medline: 33549617]
20. Allen JG, MacNaughton P, Satish U, Santanam S, Vallarino J, Spengler JD. Associations of cognitive function scores with carbon dioxide, ventilation, and volatile organic compound exposures in office workers: a controlled exposure study of green and conventional office environments. Environ Health Perspect 2016;124(6):805-812 [FREE Full text] [doi: 10.1289/ehp.1510037] [Medline: 26502459]
21. Ramos T. The FEASible study structure. Participants are asked to complete comprehensive surveys, followed by performing a health screening and a toolkit deployment visit where they get the smartwatch and the home environment sensor kit. The participants complete the 2-week data collection, and after they return the toolkit, they might opt to participate in the magnetic resonance imaging (MRI) scan and blood draw. BioRender. URL: https://BioRender.com/vngy9wi

XSL•FO
RenderX

22. Wu C, Fritz H, Miller M, Craddock C, Kinney K, Castelli D, et al. Exploring post COVID-19 outbreak intradaily mobility pattern change in college students: a GPS-focused smartphone sensing study. Front Digit Health 2021;3:765972 [FREE Full text] [doi: 10.3389/fdgth.2021.765972] [Medline: 34888544]
23. Wu C, McMahon M, Fritz H, Schnyer DM. Circadian rhythms are not captured equal: exploring circadian metrics extracted by differentcomputational methods from smartphone accelerometer and GPS sensors in daily life tracking. Digit Health 2022;8:20552076221114201 [FREE Full text] [doi: 10.1177/20552076221114201] [Medline: 35874860]
24. Mujahid MS, Diez Roux AV, Morenoff JD, Raghunathan T. Assessing the measurement properties of neighborhood scales: from psychometrics to ecometrics. Am J Epidemiol 2007;165(8):858-867. [doi: 10.1093/aje/kwm040] [Medline: 17329713]
25. Leppämäki S, Meesters Y, Haukka J, Lönnqvist J, Partonen T. Effect of simulated dawn on quality of sleep--a community-based trial. BMC Psychiatry 2003;3:14 [FREE Full text] [doi: 10.1186/1471-244X-3-14] [Medline: 14577838]
26. Casillas A, Clark L. The Mini Mood and Anxiety Symptom Questionnaire (Mini-MASQ). 2000 Presented at: 72nd Annual Meeting of the Midwestern Psychological Association; 2000; Chicago URL: https://scholar.google.com/scholar_lookup?title=The%20mini%20mood%20and%20anxiety%20symptom%20questionnaire%20&publication_year=2000&author=A.%20Casillas&author=L.A.%20Clark
27. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. Psychological Assessment 1995;7(3):309-319. [doi: 10.1037//1040-3590.7.3.309]
28. Booth M. Assessment of physical activity: an international perspective. Res Q Exerc Sport 2000;71 Suppl 2:114-120. [doi: 10.1080/02701367.2000.11082794] [Medline: 25680021]
29. Williams DR, Yu Y, Jackson JS, Anderson NB. Racial differences in physical and mental health: socio-economic status, stress and discrimination. J Health Psychol 1997;2(3):335-351 [FREE Full text] [doi: 10.1177/135910539700200305] [Medline: 22013026]
30. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
31. Mangal S, Niño de Rivera S, Choi J, Reading Turchioe M, Benda N, Sharko M, et al. Returning study results to research participants: data access, format, and sharing preferences. Int J Med Inform 2023;170:104955. [doi: 10.1016/j.ijmedinf.2022.104955] [Medline: 36565546]
32. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://www.R-project.org/ [accessed 2025-08-05]
33. Allaire J, Xie Y, Dervieux C, McPherson J, Luraschi J, Ushey K, et al. rmarkdown: Dynamic Documents for R. R Package Version 228. URL: https://CRAN.R-project.org/package=rmarkdown [accessed 2025-08-05]
34. Fritz H, Kinney K, Wu C, Schnyer D, Nagy Z. Data fusion of mobile and environmental sensing devices to understand the effect of the indoor environment on measured and self-reported sleep quality. Building and Environment 2022 Apr;214:108835. [doi: 10.1016/j.buildenv.2022.108835]
35. Hernández-Gordillo A, Ruiz-Correa S, Robledo-Valero V, Hernández-Rosales C, Arriaga S. Recent advancements in low-cost portable sensors for urban and indoor air quality monitoring. Air Qual Atmos Health 2021;14(12):1931-1951. [doi: 10.1007/s11869-021-01067-x]
36. Ramos T. The FEASible study CONSORT (Consolidated Standards of Reporting Trials) diagram. The study is divided into 3 main categories: online surveys, health screening, 2-week remote sensor data collection, and magnetic resonance imaging (MRI) and blood draw visits. The n represents the number of participants. BioRender. URL: https://BioRender.com/bd6p0gp
37. Botton CE, Santos LP, Moraes BG, Monteiro RB, Gomes MLB, Wilhelm EN, et al. Recruitment methods and yield rates in a clinical trial of physical exercise for older adults with hypertension-HAEL Study: a study within a trial. BMC Med Res Methodol 2022;22(1):42 [FREE Full text] [doi: 10.1186/s12874-022-01535-7] [Medline: 35144532]
38. Wong KHTW, Lau WCY, Man KKC, Bilbow A, Ip P, Wei L. Effectiveness of facebook groups and pages on participant recruitment into a randomized controlled trial during the COVID-19 pandemic: descriptive study. J Med Internet Res 2023;25:e46190 [FREE Full text] [doi: 10.2196/46190] [Medline: 37847536]
39. Weerts DJ, Sandmann LR. Building a two-way street: challenges and opportunities for community engagement at research universities. rhe 2008;32(1):73-106. [doi: 10.1353/rhe.0.0027]
40. Velasco-Mondragon E, Jimenez A, Palladino-Davis AG, Davis D, Escamilla-Cejudo JA. Hispanic health in the USA: a scoping review of the literature. Public Health Rev 2016;37:31. [doi: 10.1186/s40985-016-0043-2] [Medline: 29450072]
41. Giuntella O. The Hispanic health paradox: New evidence from longitudinal data on second and third-generation birth outcomes. SSM Popul Health 2016;2:84-89 [FREE Full text] [doi: 10.1016/j.ssmph.2016.02.013] [Medline: 29349130]
42. Daviglus ML, Pirzada A, Talavera GA. Cardiovascular disease risk factors in the Hispanic/Latino population: lessons from the Hispanic community health study/study of Latinos (HCHS/SOL). Prog Cardiovasc Dis 2014;57(3):230-236. [doi: 10.1016/j.pcad.2014.07.006] [Medline: 25242694]
43. Chaudhry BA, Brian MS, Morrell JS. The relationship between sleep duration and metabolic syndrome severity scores in emerging adults. Nutrients 2023;15(4):1046 [FREE Full text] [doi: 10.3390/nu15041046] [Medline: 36839404]

XSL•FO
RenderX

Abbreviations

BOLD fMRI: blood oxygenation level–dependent functional MRI
CO_2: carbon dioxide
CVD: cardiovascular disease
FLAIR: fluid-attenuated inversion recovery
FOV: field of view
MetS: metabolic syndrome
MPRAGE: magnetization prepared rapid gradient echo
MRI: magnetic resonance imaging
MRS: magnetic resonance spectroscopy
pcASL: pseudo-continuous arterial spin labeling
$PM_{2.5}$: particulate matter ≤2.5 µm in diameter
REDCap: Research Electronic Data Capture
RRAMSC: relaxation-resolved advanced multi–spin-echo chemical shift
SE: spin echo
SVS: single voxel spectroscopy
TA: total acquisition time
TE: echo time
TI: inversion time
TR: repetition time

Edited by J Sarvestan; submitted 07.Aug.2025; this is a non–peer-reviewed article;accepted 17.Nov.2025; published 06.Jan.2026.

Please cite as:
Ramos-Santiago T, Huglo E, Li Y, Corral C, Wu C, Kinney K, Castelli D, Haley A
Assessing Associations Between Environmental, Sleep, and Physical Activity Factors and Metabolic Syndrome Risk: Protocol for the FEASible Study
JMIR Res Protoc 2026;15:e82034
URL: https://www.researchprotocols.org/2026/1/e82034
doi:10.2196/82034
PMID:

©Tania Ramos-Santiago, Elise Huglo, Yanrong Li, Christian Corral, Congyu Wu, Kerry Kinney, Darla Castelli, Andreana Haley. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 06.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Impact of Intranasal Administration of Ayurveda Medicine in Apparently Healthy Individuals on Neurophysiological Variables and Functional Connectivity Using Functional Magnetic Resonance Imaging: Protocol for an Exploratory Randomized Controlled Trial

Devi R Nair[1]; Sandya C J[2]; Sophia Jameela[3]; Remya E[1]; Srikanth Moorthy[2]; Harilal Parasuram[4]; Shruti Khanduri[3]; Bhogavalli Chandrasekhara Rao[3]; Narayanam Srikanth[3]; Rabinarayan Acharya[3]

[1]National Ayurveda Research Institute for Panchakarma, Thrissur, Kerala, India

[2]Department of Radiodiagnosis, Amrita Institute of Medical Sciences, Kochi, Kerala, India

[3]Central Council for Research in Ayurvedic Sciences, New Delhi, India

[4]Amrita Advanced Centre for Epilepsy, Amrita Institute of Medical Sciences, Kochi, Kerala, India

Corresponding Author:
Devi R Nair
National Ayurveda Research Institute for Panchakarma
Cheruthuruthy
Thrissur, Kerala
India
Phone: 91 09074298620
Email: drdevinair87@gmail.com

Abstract

Background: *Nasyakarma*, an Ayurveda nasal drug delivery system, is considered a potent therapeutic modality in *panchakarma* treatment. Uniquely, nasal drug delivery can bypass liver metabolism and the blood-brain barrier for faster drug delivery. Available studies on *nasya karma* focus mainly on its efficacy. This pioneering study aims to explore the mechanisms of *nasya karma* on brain function and neurophysiology, investigating its potential to modulate activity in specific brain regions and affect the functional connectivity between these regions using functional magnetic resonance imaging (fMRI).

Objective: The study aims to map the neurophysiological response of the brain to *nasya karma* using blood oxygen level–dependent fMRI in both rest and task phases and assess the impact of *nasya karma* on quality of life, cognition, sleep, and psychological well-being in healthy volunteers.

Methods: A total of 60 healthy volunteers in the age group of 20 to 40 years who fulfill the selection criteria will be recruited for this randomized controlled trial at the National Ayurveda Institute for Panchakarma, Cheruthuruthy, Kerala, India, and randomized in a 1:1 ratio to either the intervention group (n=30; group 1, receiving *nasya karma* with *anu taila*, an oil-based formulation manufactured using a paste and decoction of herbal medicines with oils as the base material, for a period of 14 days) or the control group (n=30; group 2, not receiving any intervention). The participants will undergo task-based and resting fMRI on day 1 (twice on day 1 before administration and 15 minutes after *nasya karma* for participants in group 1) and day 14 to map the neurophysiological response to *nasya karma* of the brain. A comprehensive neuroimaging protocol using structural magnetic resonance imaging and fMRI will be used in the study. The effect of *nasya karma* on sleep, psychological well-being, cognition, and quality of life will be assessed on the 1st and 30th days in both groups.

Results: After the data collection process of this randomized controlled trial is completed, the study will go into the analysis stage, in which the collected data will be subjected to robust statistical analysis. The final results of the study are expected to be published in 2026.

Conclusions: This study will investigate the neurophysiological mechanisms of *nasya karma* by examining clinical, neuropsychological, and neuroimaging variables to identify associated neural patterns to develop therapeutic protocols for many diseases. The findings will also provide evidence for future research supporting the use of *nasya karma* as a practical and noninvasive therapeutic modality for treating cerebrovascular, behavioral, and neurological disorders as indicated in Ayurveda. The study will propel innovative research focusing on the neural mechanisms responsible for the delivery of central nervous system therapeutics to the brain, thereby bypassing the blood-brain barrier and yielding favorable outcomes in central nervous system diseases.

XSL•FO
RenderX

Trial Registration: Clinical Trials Registry – India CTRI/2023/06/054219; hhttps://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=ODQ1OTU=&Enc=&userName=

International Registered Report Identifier (IRRID): DERR1-10.2196/67132

(JMIR Res Protoc 2026;15:e67132) doi:10.2196/67132

KEYWORDS

nasya; healthy individuals; neurophysiology; functional magnetic resonance imaging; functional MRI; blood oxygen level–dependent; BOLD; intranasal

Introduction

Background

In Ayurveda, *nasya karma* is among the *panchakarma* procedures, a set of 5 therapeutic modalities for body purification involving nasal administration of medicine, and is regarded as a potent approach in health promotive, preventive, and therapeutic applications. The therapeutic applications of *nasya karma* span a wide clinical spectrum, addressing conditions such as chronic sinusitis, cervical spondylosis, and various neurological and psychiatric disorders. *Nasyakarma* is the only procedure that uses the nasal route for medicine administration to deliver therapeutic agents directly to the central nervous system (CNS), bypassing the digestive tract and potentially enhancing the efficacy of treatment owing to its faster absorption.

Intranasal drug administration has gained considerable traction for many decades as a noninvasive route for delivering medications targeting local, systemic, and CNS effects. The leaky intercellular junctions of the nasal mucosa and the extensive vascularity of the lamina propria and nasal mucosa facilitate optimal drug absorption, positioning intranasal delivery as a highly effective and versatile method for a wide range of therapeutic applications [1,2]. In the context of neurological disorders, intranasal drug delivery is advantageous due to its ability to achieve high drug levels in the brain, providing a direct pathway that bypasses the blood-brain barrier (BBB) more effectively than traditional parenteral and oral routes, especially considering that the anatomical features allow for rapid drug absorption, a noninvasive administration method, and the avoidance of hepatic first-pass metabolism, leading to faster therapeutic effects [3]. Direct absorption of molecules through the trigeminal and olfactory pathways provides a direct route to the brain, resulting in a favorable pharmacokinetic and pharmacodynamic profile by bypassing major physiological barriers such as the BBB and the blood–cerebrospinal fluid barrier [4].

The unique relationship that exists between nasal cavity and cranial cavity tissues makes intranasal drug delivery to the brain more feasible and effective. Intranasal delivery provides drugs with short channels to bypass the BBB, greatly enhancing the therapeutic effect on the neurological disease spectrum [5].

Panchakarma encompasses a comprehensive set of Ayurvedic detoxification procedures, including *vamana* (emesis), *virechana* (purgation), *nasya* (nasal administration of herbal oils or powders), *basti* or *vasti* (herbal enema), and *raktamokshana* (bloodletting) [6]. *Panchakarma* is a mainstay of Ayurveda clinical practice where the physician can choose any of the 5 sets of therapeutic modalities to bring about a homeostatic equilibrium in the body tailored to each individual depending on the intended purpose of use (promotive, preventive, therapeutic, or rehabilitative), body constitution, *dosha* (the body humors or functional principles that govern physiological and psychological processes), health or disease status, age, and fitness to undergo the *panchakarma* procedure.

Anu taila [7] is an oil-based formulation manufactured using a paste and decoction of herbal medicines with oils as the base material. The ingredients are combined in specific proportions as highlighted in classic Ayurvedic texts and subjected to a unique heating process to meet certain pharmaceutical parameters, ensuring the retention of both water-soluble and lipid-soluble principles at optimal levels. *Anu taila* can be likened to liposomes, where nanoparticles comprise lipid bilayer membranes surrounding an aqueous interior. The higher viscosity of this formulation increases contact time between the drug and the nasal mucosa, thereby enhancing permeation, and the viscosity may influence mucociliary clearance, potentially reducing it and, thus, prolonging the formulation's retention time in the nasal cavity, which can lead to increased drug absorption [8,9].

Despite its extensive use as a therapeutic modality in clinical practice across a wide range of conditions and in healthy individuals to maintain health and homeostasis, the exact mechanism of action of *nasya karma* remains obscure and mysterious, primarily due to a lack of physiological or mechanistic studies in this area. Given the paucity of research on the neurophysiological mechanisms underlying *nasya*, this study aims to use functional magnetic resonance imaging (fMRI) to investigate its potential to modulate activity in specific brain regions and affect the functional connectivity between these regions. Understanding these neurophysiological correlates is crucial as it will elucidate whether *nasya karma* exerts targeted effects or has more widespread, nonspecific impacts on the brain's intrinsic functioning. This knowledge will benefit stakeholders by providing a scientific basis for conceptualizing future research, potentially leading to optimized therapeutic protocols, improved clinical outcomes, and enhanced integration of *nasya karma* into modern medical practice. By establishing clear spatial and temporal patterns of brain activity associated with *nasya karma*, this study could pave the way for more precise and effective applications of this traditional Ayurvedic practice in contemporary health care. This may also propel innovative research focusing on the neural mechanisms responsible for the delivery of CNS therapeutics to the brain,

XSL•FO
RenderX

thereby bypassing the BBB and yielding favorable outcomes in CNS diseases.

Although *nasya karma* is extensively used for neurological and systemic health benefits, no studies have elucidated its neurophysiological mechanisms of action, specific brain region targets, or the connectivity changes underlying its effects. This knowledge gap prevents scientific validation and optimized therapeutic application of *nasya karma* in clinical neuroscience. This study addresses this gap by using fMRI to characterize the brain's response to *nasya karma*, thereby providing mechanistic insights essential for evidence-based integration of this ancient intervention into modern CNS therapeutics.

Objectives

The primary objective of this study is to investigate the neurophysiological activity of the brain induced by the *nasya karma* procedure using blood oxygen level-dependent (BOLD) fMRI to elucidate the comprehensive neural response to *nasya karma* both in the rest and task phases. The secondary objective is to examine the impact of the *marsha nasya karma* (Ayurveda nasal drug delivery using a medicated or nonmedicated lipid medium) procedure on health parameters in apparently healthy volunteers, such as quality of life; cognition; sleep; and psychological well-being through anxiety, stress, and cognitive health.

The research hypothesis for this study is that the administration of *marsha nasya karma* with *anu taila* will increase the activation of different brain areas and enhance functional connectivity in different neuronal networks compared to baseline.

Methods

Ethical Considerations

The study will strictly adhere to the principles outlined in the Declaration of Helsinki (World Medical Association; 2013). Before enrollment in this trial, all participants will be required to provide informed consent by signing a consent form. The protocol has undergone rigorous ethical review and received approval from the institutional ethics committee of the National Ayurveda Research Institute for Panchakarma (NARIP), Cheruthuruthy, Thrissur, Kerala, Peripheral Institute of CCRAS, Ministry of Ayush, Government of India and Scientific Review Board, Amrita Institute of Medical Sciences, Kochi (F.No.8/16/2023/NARIP/Tech meeting/2508 dated March 31, 2023), and the study is registered with the Clinical Trials Registry – India. The participants will be given incidental support charges of INR 100 (US $1.11) per visit as participant compensation. The trial is covered by clinical trial insurance for any incidents, and the participants will receive appropriate compensation if these arise.

Protocol amendments, if any are needed in the course of the trial, will be communicated to the institutional ethics committees at Amrita Institute of Medical Sciences, Kochi, and NARIP, Cheruthuruthy, India. Personal information about potential and enrolled participants will be collected, shared, and maintained in a confidential manner before, during, and after the trial unless subjected to judiciary (ie, court-related) disputes. Access to trial data will be granted only to the research team involved at NARIP and the Amrita Institute of Medical Sciences (AIMS). There will be no public access to the trial dataset. Dissemination policy includes publication of study findings in peer-reviewed indexed international journals.

Eligibility Criteria

Apparently healthy volunteers aged 20 to 40 years who are physically independent, ambulatory, and right-handed; have a BMI of 18.5 to 24.9 kg/m^2; and are willing to participate and provide written informed consent will be included in the study. Only those who agree to abstain from beverages such as coffee, tea, energy drinks, or other stimulants beginning 12 hours before undergoing fMRI will be considered for recruitment. The healthy status of volunteers will be confirmed through the absence of any active or chronic disease, as determined through a comprehensive medical evaluation that will include a detailed medical history, complete physical examination (including vital signs), 12-lead electrocardiogram, hematology, blood chemistry, serology, psychological assessment, and urinalysis, among other tests.

Individuals with any disease or symptom, with a history of psychiatric disorders as assessed using the Mini-International Neuropsychiatric Interview, with psychosomatic disorders, who used psychotropic medications known to interfere with the dopaminergic system or sedatives within the previous 6 months, with a family history of psychiatric diseases, with past or current abuse of psychoactive drugs, or who use alcohol or smoke will be excluded. Individuals contraindicated for magnetic resonance imaging (MRI), such as those with medical implants, retained wires (eg, temporary pacing), an increased risk of myocardial infarction or other cardiac problems, known pregnancy, nausea (due to aspiration risk), metallic dental work, cochlear implants, body jewelry, and other metallic foreign objects implanted in the body will not be eligible either. Individuals with cognitive impairment, women in the menstrual phase, those with intracranial abnormalities that could adversely affect the interpretation of brain scans, and those who self-identify as claustrophobic are also not eligible for participation in the study. In addition to this, participants deemed unfit for *nasya karma* therapy [10], such as those with fever, rhinitis, cough, or other inflammatory conditions as indicated in the Ayurveda treatises, will be excluded.

Interventions

The participants in the intervention group will be administered *nasya karma* (nasal instillation) with *anu taila* at a dose of 3 mL per nostril for 7 days (on every alternate day for 14 days) as indicated in the Ayurveda treatises for healthy individuals. Before *nasya*, local oleation and sudation of the head region will be conducted for approximately 10 minutes. Following the *nasya karma* procedure, participants will inhale medicated smoke and rinse their mouths with lukewarm water. The control group, consisting also of healthy individuals, will not undergo any interventions to facilitate assessment of the specific effects attributed to *nasya karma* in the intervention group when compared to baseline measures in healthy individuals.

XSL•FO
RenderX

Nasya karma will be administered by the investigator at the hospital's *panchakarma* unit except on assessment days, when it will be conducted at AIMS. Each participant will be closely monitored for 30 minutes following the procedure to observe any immediate effects or reactions. Vital signs such as heart rate, blood pressure, and respiratory rate will be monitored and recorded before and after every session. Participants reporting any adverse symptoms experienced after each session will be recorded through a standardized format designed to capture a wide range of potential side effects, including nasal irritation, headache, dizziness, nausea, or any other discomfort. Any serious adverse event will be reported immediately to the institutional ethics committee, and a causality assessment will also be conducted. Appropriate care will be provided for all adverse events requiring medical intervention, and they will be followed up on until resolution or stabilization.

If a participant experiences significant adverse effects, such as severe nasal irritation, fever, sinusitis, or any other serious discomfort that impacts their well-being, the intervention will be immediately halted, and the participant will be withdrawn. Discontinuation will also occur in cases of noncompliance with the treatment regimen, if the participant wishes to withdraw for any reason, or if the participant meets any of the exclusion criteria during the study. As the trial involves healthy volunteers, administration of concomitant medication is prohibited during the trial period of 30 days.

This study is structured as an exploratory randomized controlled clinical trial, with participants allocated in a 1:1 ratio to the intervention and control groups, matched for age and sex to comprehensively assess and evaluate the neurophysiological and holistic health effects of the *marsha nasya karma* procedure in comparison to a control group (Figure 1). Settings of the study include NARIP and AIMS.

XSL•FO
RenderX

Figure 1. Schematic diagram of study plan.

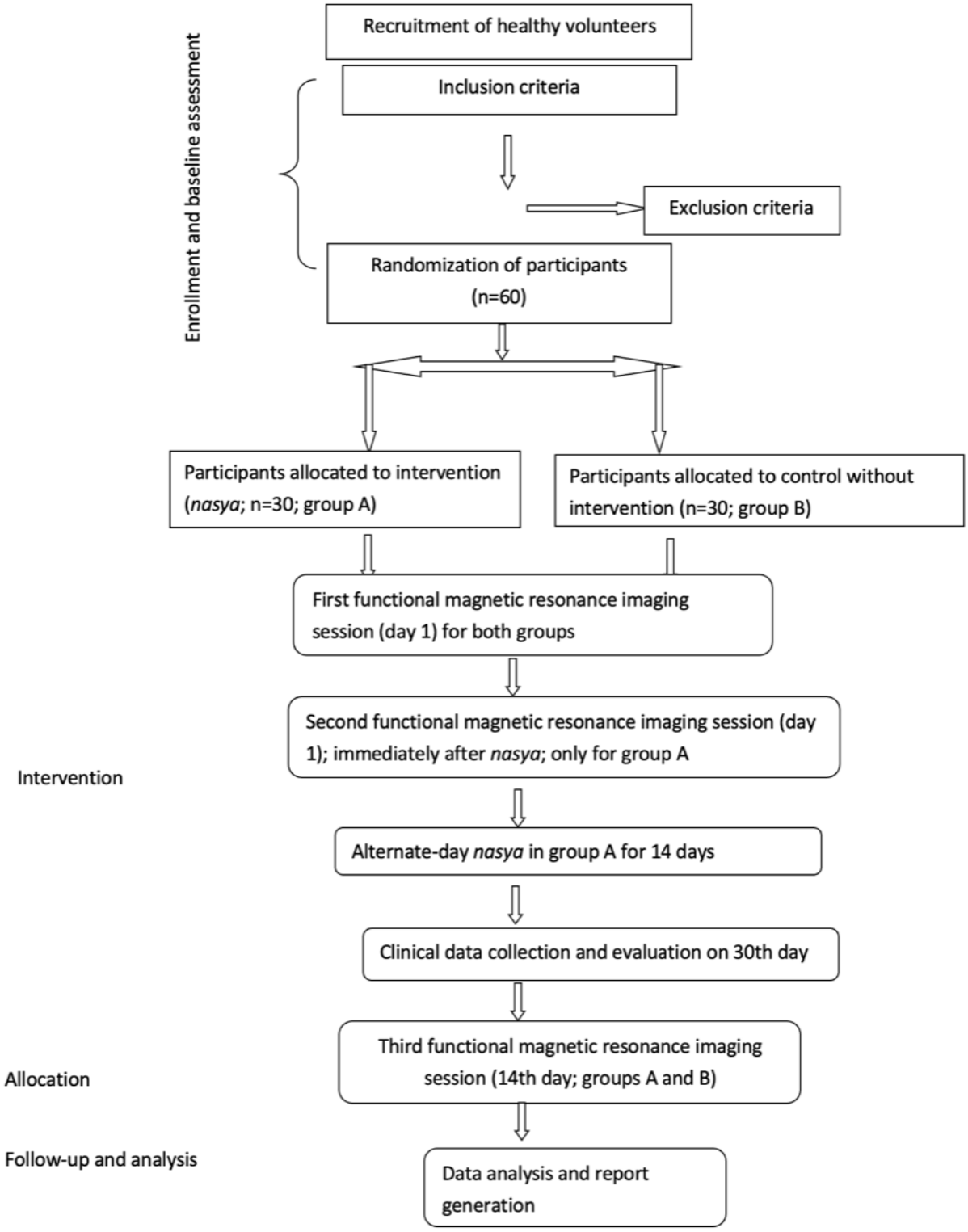

Study Setting

Participants will be screened and enrolled at the NARIP in Cheruthuruthy, Thrissur, an institute under the Central Council for Research in Ayurvedic Sciences, Ministry of Ayush, government of India. The institute is well equipped for comprehensive Ayurvedic assessments and interventions, ensuring a high standard of participant care and protocol adherence during the *nasya karma* procedure. fMRI assessments will be carried out at AIMS, which is equipped with state-of-the-art imaging technology and experienced radiologists, providing the advanced neuroimaging capabilities necessary for detailed BOLD fMRI studies.

Sample Size

As this is designed as an exploratory study aiming to gather preliminary data on the neurophysiological effects of the *nasya karma* procedure and its impact on various health parameters,

XSL•FO
RenderX

a sample size of 60 (30 per group) is considered appropriate to achieve the study objectives and provide valuable insights that will inform the design of future, larger-scale clinical trials. Feasibility is a critical consideration as recruiting a larger sample of participants willing to undergo both the *nasya karma* procedure and twice-repeated fMRI assessments may be challenging, particularly among healthy volunteers. However, the expected variation in response to the interventions is likely to be lower than that in a patient population, which will allow for meaningful conclusions to be drawn from a smaller sample size. Similarly, a sample size of 60 balances the need for preliminary data with the practical constraints of the logistical and financial resources of the study.

Randomization and Allocation

Eligible participants in this study will undergo random allocation to either the intervention or control group using a computer-generated random sequence prepared by an independent statistician, ensuring an unbiased approach to treatment assignment. Each random sequence number will be securely sealed in opaque envelopes, with the enrollment ID on the outside. Before enrollment, participants will open these envelopes to determine their allocated group to minimize bias.

Primary Outcome Measure

The primary outcome measure for this study is the BOLD signal detected via fMRI in specific brain regions that are associated with motor, sensory, memory, and cognitive circuits elicited during both the rest and task phases. A 3-Tesla MRI scanner will be used to assess the brain's response to different stimuli at multiple time points: before the *nasya karma* procedure, immediately before the *nasya karma* procedure, immediately after (within 15 min) the *nasya karma* procedure, and on the 14th day of the *nasya karma* therapy.

Secondary Outcome Measures

These include the therapeutic effects of *marsha nasya karma* on cognitive function assessed via the Montreal Cognitive Assessment scale [11], perceived stress assessed using the Perceived Stress Scale [12], anxiety assessed using the Hamilton Anxiety Rating Scale [13], and health-related quality of life assessed using the 36-Item Short Form Health Survey [14]. The secondary outcomes will be assessed before the intervention and on the 1st and 30th days following the *nasya karma* therapy.

Detailed Methodology of the Primary Outcome Measure

Overview

The primary outcome measure encompasses 2 key components. The first is the assessment of the immediate neurophysiological response elicited by *nasya karma* during both the resting and task phases using MRI. The second component is the evaluation of the neurophysiological response elicited by *nasya karma* after 14 days (following 7 days of nasya) during both the resting and task phases using MRI. Functional neuroimages will be collected using a 3-Tesla scanner (Discovery MR750w; GE HealthCare). Resting state fMRI data will include 3D T1-weighted structural imaging with a fast-spin echo sequence and BOLD imaging with a spin echo, gradient echo, and echo-planar imaging sequence for 40 minutes, covering the entire brain. The parameters for 3D-T1 imaging will include a slice thickness of 1 mm; a repetition time and echo time of 8.5 milliseconds and 3.2 milliseconds, respectively; a field of view of 256 × 256 × 142 mm^3; a flip angle of 12°; a matrix of 256 × 256; an in-plane resolution of 1 × 1 × 1 mm^3; and 142 layers. The parameters for BOLD imaging will include a slice thickness of 3.2 mm; a repetition time and echo time of 3000 milliseconds and 30 milliseconds, respectively; a flip angle of 80°; a matrix of 64 × 64; a field of view of 256 × 256 × 142 mm^3; 35 layers; and 150 time points.

Preparatory Phase

Participants will undergo a robust screening interview to ensure understanding and absence of contraindications for MRI scanning. Before the task phase fMRI, participants will be thoroughly informed and trained. Communication with the operator will be maintained throughout the scan. Incidental findings requiring medical intervention will be disclosed, and participant comfort will be ensured during scanning. All scans will be reviewed by a qualified radiologist, and any clinically significant incidental findings will be communicated confidentially to the participant along with appropriate referral advice for further evaluation and management as per institutional protocols. Verbal communication will be limited due to MRI noise, and the participants will use a 4-button fiber-optic response box positioned for use with their right hand to communicate and record behavioral responses during tasks. To mitigate the effects of scanner noise, participants will wear earplugs. Head motion will be minimized using foam pads and a headband to ensure stable positioning throughout the MRI scans. Metallic items will be removed, and participants will wear MRI-compatible attire. Participants in both the trial and control groups will be directed to avoid heavy meals within 2 hours before and 1 hour after *nasya karma*. Both groups will refrain from consuming coffee, tea, or caffeinated drinks for 12 hours preceding the fMRI. In contrast, the control group will undergo resting state mapping on both the 1st and 14th days, whereas the participants in the intervention group will undergo the *nasya karma* procedure lasting approximately 30 to 45 minutes, after which participants will proceed to the MRI room. Participants in the control group will proceed directly to the MRI without any intervention.

Data Collection

Resting Phase fMRI

In this study, a block design paradigm will be used during task phases before and after *marsha nasya karma*, where stimuli are presented continuously within specific blocks alternated with rest periods to elicit and distinguish brain responses. Resting state fMRI, on the other hand, captures spontaneous brain activity without external stimuli, allowing for the correlation of signal fluctuations among functionally related brain regions. Panbrain activation mapping will be conducted during the resting phase before and immediately after *nasya karma* therapy (within 15 min) in the trial group. fMRI will be conducted before the *nasya karma* procedure and then again after the complete procedure is over. Participants in the control group will proceed

XSL•FO
RenderX

directly to fMRI. During scanning, participants will be instructed to remain still with their eyes closed for 15 minutes to avoid motion artifacts. Any head movement will necessitate repeat mapping. The resting state fMRI paradigm will map spontaneous brain activity across the visual, auditory, motor (left and right), memory, and language domains before and after *marsha nasya karma*.

Task-Based fMRI

Overview

The task paradigms in this study are designed to effectively stimulate relevant brain areas while ensuring that participant performance remains accurate. Rest blocks are strategically interspersed between active blocks to establish baseline fMRI signals and allow participants to rest. Each domain collects sufficient data through a structured design featuring at least 4 blocks per task. Task durations are optimized to maintain participant concentration without inducing fatigue; breaks are included if needed due to head movements, ensuring image consistency.

Visual fMRI

The study will use a checkerboard method with 8-Hz flickering alternating with 30-second rest blocks in a sequence of 1 rest cycle plus 4 flickering cycles (B, AB, ABAB, and AB). Participants receive prior education and training to understand their tasks: viewing a flickering checkerboard followed by a fixation cross during rest periods. The paradigm runs for 4 minutes and 30 seconds, minimizing eye movements while allowing for effective scene categorization.

Auditory fMRI

Auditory tones (1000 and 2000 Hz; ratio of 4:1) alternate with 20-second *veena* sounds in a sequence of ABABABABABAB. Participants, relaxed with their eyes closed, hear alternating tones during rest periods, with more frequent standard tones. Each auditory block is followed by a 20-second *veena* sound. The paradigm duration is 4 minutes.

Motor fMRI (Right and Left)

Right and left motor areas are separately mapped using hand movements. Participants perform finger movements indicated by a flashing green dot on the screen alternating with 30-second rest blocks (1 off plus 4 on cycles [BABABABAB]). They synchronize opening and closing fist movements with the dot, followed by rest periods focusing on a fixation cross. Each motor task runs for 4 minutes and 30 seconds per hand.

Memory fMRI

The working memory attention task lasts 6 minutes and 45 seconds, presenting numbers for 0.5 seconds with 2.5 seconds of a black screen. Blocks of 51 seconds feature numbers with 30-second rest periods in between. Participants press a button when presented with identical numbers, with task initiation and closure including 15-second rest blocks. This paradigm effectively tests working memory under fMRI conditions.

Principles Underlying Assessment of Brain Response to Nasya Karma

The response to *nasya karma* will be assessed considering 3 principles. The first principle is that neural activity correlates with changes in local blood oxygen levels. The second principle is that oxygenated blood exhibits different magnetic properties compared to deoxygenated blood. The third principle is that alterations in the ratio of oxygenated to deoxygenated blood, known as the hemodynamic response function, can be inferred using fMRI through the measurement of the BOLD response.

Blinding

In this study, participant blinding is not feasible as the intervention involves a procedural administration (*nasya karma*) where participants are aware of undergoing the procedure and the study primarily aims to evaluate the procedural effect rather than an isolated pharmacological effect. Assessor blinding will also not be implemented as the primary outcome measure is fMRI, an objective instrumental assessment that is not subject to assessor interpretation bias during data acquisition.

Data Management

Data pertaining to the study, including screening data, recruitment data, baseline information, daily *nasya karma* procedure assessments, laboratory test reports, and secondary outcome measures assessed using scales, will be comprehensively documented in case report forms within 24 hours, and electronic case report forms will be completed within 7 days. Only the study personnel, individuals delegated with the task of data entry, and the statisticians will have access to the data. The imaging data generated from fMRI will be stored in hard disks at AIMS and will undergo further analysis after completion of the study. All study documents will be securely stored for 5 years following study completion.

Statistical Analysis

Overview

The modified intention-to-treat principle will guide the analysis for all outcomes within the full analysis set. Quantitative data will be summarized using mean, SD, minimum, maximum, and median values. Enumeration data will be presented as frequency and the corresponding percentage. Intragroup comparisons will be conducted using the Student paired *t* test or Wilcoxon signed rank test for continuous variables and the chi-square test for categorical variables. Intergroup comparisons will be conducted using the independent-sample *t* test or the Mann-Whitney *U* test. Significance thresholds will be set at $P<.05$ for all 2-tailed tests.

fMRI Data Analysis

The fMRI data will undergo preprocessing using either the Data Processing and Analysis for Brain Imaging or CONN toolboxes, which will encompass essential steps such as slice timing correction, realignment for motion correction, regression to remove confounding effects, normalization to a standard template, smoothing to enhance signal-to-noise ratio, detrending to remove low-frequency drifts, and filtering to focus on relevant frequency bands, aligning with established practices in the field. After data preprocessing, amplitude of low-frequency

XSL•FO
RenderX

fluctuation, functional connectivity, and large-scale functional brain network analysis will be conducted to investigate the difference in neurophysiological responses between the groups. The results will be interpreted in the context of the experimental paradigm to elucidate neural correlates of the studied phenomena, providing insights into the effects of *nasya karma* on brain function. Voxels exhibiting expected time-course correlations (30-second periods) will receive high activation scores, whereas those showing no correlation or deactivation will receive lower or negative scores, respectively. MRI data analysis will use general linear models via statistical parametric mapping (SPM). fMRI data will be processed using MATLAB 2023 (MathWorks Inc).

Resting phase MRI data will be analyzed using software tools such as SPM, the CONN toolbox, and the Functional Magnetic Resonance Imaging of the Brain Software Library (Analysis Group), whereas task phase fMRI will specifically use SPM. Given the focus on group data in this study, complex patterns of brain activation will be interpreted collectively rather than individually. For voxel-based analysis, a significance threshold of $P<.01$ at the voxel level and $P<.05$ corrected for false discovery rate (FDR) at the cluster level will be applied. Connectome-based analysis will use a threshold of $P<.05$ corrected for FDR. Additionally, Pearson correlation analysis will be performed to explore potential associations between clinical outcomes and fMRI findings, examining changes in brain activity following different interventions.

Multiple-Comparison Correction

For voxel-wise analyses, statistical inference will be performed with the following corrections. At the voxel level, uncorrected $P<.001$ will be used as an initial threshold. At the cluster level, familywise error (FWE) correction ($P<.05$) will be applied to control for multiple comparisons across the brain. In parallel, FDR correction will be applied where appropriate to control the expected proportion of false positives. For independent component analysis–based network analyses, dual regression outputs will be thresholded using FWE correction ($P<.05$). This combination ensures robust control of type I error.

Power Analysis

A priori power analysis was conducted using G*Power (version 3.1) and the fMRIPower toolbox. On the basis of pilot data and prior studies of similar paradigms, an expected effect size of Cohen $d=0.6$ (moderate) was used. For a voxel-level threshold of $P<.001$, uncorrected and cluster-level FWE of $P<.05$, and a sample size of 30 per group, the analysis achieves >80% power to detect activation differences in relevant cortical regions. This power analysis informed our sample size and statistical thresholds.

Correlation of Clinical Outcomes and fMRI Data

Associations between clinical measures and fMRI-derived metrics (amplitude of low-frequency fluctuation, functional connectivity, and network strength) will be assessed using partial Pearson correlations controlling for age and sex. Given the multiple regions and measures assessed, the Benjamini-Hochberg FDR correction ($q<0.05$) will be applied to adjust for multiple testing across correlations. In addition, exploratory Spearman correlations will be performed to confirm the robustness of the observed associations.

Protocol Adherence Monitoring

In this study, all *nasya karma* procedures will be administered at the NARIP by trained investigators on alternate days. Participants will not perform *nasya karma* at home. This approach ensures direct procedural supervision, standardized administration, and immediate monitoring for any adverse effects or participant discomfort. Direct assessment by the investigator at each session enhances protocol adherence, fidelity, and the overall robustness of study execution, eliminating the variability associated with home-based administration and, thereby, strengthening the internal validity and reliability of the study findings as in Table 1.

XSL•FO
RenderX

Table 1. Gantt chart.

	Week 0	Baseline		Intervention period (1st-13th day)	Follow-up	
		1st day	1st day (immediately after *nasya*)		14th day	30th day
Participant screening, including laboratory tests	✓					
Informed consent	✓					
Randomization	✓					
Intervention: *nasya karma* (group A)		✓		✓		
Control (group B): without *nasya karma*						
fMRI [a] scan						
Group A		✓	✓		✓	
Group B		✓			✓	
Clinical assessment						
Group A						✓
Group B						✓

[a]fMRI: functional magnetic resonance imaging.

Results

After the data collection process of this randomized controlled trial is completed, the study will go into the analysis stage, in which the collected data will be subjected to robust statistical analysis. The final results of the study are expected to be published in 2026.

Discussion

Anticipated Findings

This study is anticipated to provide preliminary evidence of the neurophysiological responses evoked by the Ayurvedic procedure of *nasya karma* (nasal instillation) on specific brain areas. The findings are expected to demonstrate distinct neuronal activation patterns within visual, auditory, motor, and cognitive domains both immediately (within 15 min) and after cumulative administration (14 days). This will be the first mechanistic exploration mapping the impact of *nasya karma* on brain regions using neuroimaging techniques, bridging traditional Ayurvedic concepts and objective scientific validation.

Comparison to Prior Work

While numerous studies have established the efficacy of *nasya karma* in treating conditions such as sinusitis, headache, and neurological disorders, these works predominantly focus on clinical outcomes rather than mechanistic pathways. Previous pharmacological and anatomical studies have highlighted the naso-brain route as a potential noninvasive drug delivery pathway bypassing the BBB. However, the specific neural activation patterns evoked by *nasya karma* procedures remain unexplored in Ayurveda literature. Thus, this study complements existing clinical research by providing neurophysiological evidence to support Ayurveda's theoretical claims, potentially validating traditional beliefs under modern neuroscientific paradigms.

Strengths and Limitations

The study is a novel approach combining Ayurvedic procedures with neuroscientific imaging methods to generate mechanistic evidence. The mapping of immediate and cumulative neuronal responses that can inform both therapeutic and preventive medicine applications is the key highlight of the study. The study focuses on differential activation across brain domains, providing granular insights rather than generic findings.

Limitations include the small sample size (n=30 per group) due to feasibility constraints in recruiting healthy volunteers for radiological assessments, potentially affecting generalizability. Administration in the outpatient setting limits investigator control over adherence, introducing protocol compliance variability as yet another limitation. As an exploratory study, the design is not powered to evaluate clinical efficacy, focusing solely on mechanistic end points.

Future Directions

Findings from this study will act as preliminary data for larger controlled mechanistic trials evaluating the *nasya karma* procedure for specific neurological conditions such as cognitive impairment, stroke, neurodegenerative disorders, and mood disorders. Optimization of dosage, formulation, and timing parameters for maximum neural impact is another future direction from the study. The study paves the way to translational neuroscience research into Ayurveda therapies as noninvasive approaches to modulate brain function and deliver CNS-targeted treatments via the nasal route.

Dissemination Plan

The results of this research will be disseminated through both national and international peer-reviewed publications and

XSL•FO
RenderX

summary briefs for Ayurveda practitioners and policymakers. This aims to integrate traditional procedures into evidence-based neurological practice and stimulate interdisciplinary research between Ayurveda and modern biomedical sciences.

Conclusions

This study is expected to provide first-of-its-kind neurophysiological evidence supporting the mechanistic action of the *nasya karma* procedure in the brain. By bridging traditional Ayurveda concepts with modern neuroscience, it has the potential to pave the way for innovative interventions in neurological disease management and preventive health care.

Acknowledgments

The authors extend their gratitude to Smitha Mahesh Kushra former data analyst, Amrita Institute of Medical Sciences, who made significant contributions in all stages of drafting the protocol. The authors hereby state that data from artificial intelligence–powered tools were not used in article framing.

Funding

The study is funded by the Central Council for Research in Ayurvedic Sciences, Ministry of Ayush, Government of India.

Authors' Contributions

DRN and SJ drafted the manuscript. Author SCJ provided all technical support in drafting the radiodiagnosis part of the manuscript. Authors SJ, RE and HP reviewed and made necessary corrections to the manuscript. Authors SKM, SK, BCR, NS and RA provided technical advice regarding the framing of the entire protocol.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from the Project Evaluation Monitoring Committee, Central Council for Research in Ayurvedic Sciences, Ministry of AYUSH, Government of India.
[PDF File (Adobe PDF File), 10133 KB - resprot_v15i1e67132_app1.pdf]

References

1. Deli MA. Potential use of tight junction modulators to reversibly open membranous barriers and improve drug delivery. Biochim Biophys Acta 2009 Apr;1788(4):892-910 [FREE Full text] [doi: 10.1016/j.bbamem.2008.09.016] [Medline: 18983815]
2. Lungare S, Bowen J, Badhan R. Development and evaluation of a novel intranasal spray for the delivery of amantadine. J Pharm Sci 2016 Mar;105(3):1209-1220 [FREE Full text] [doi: 10.1016/j.xphs.2015.12.016] [Medline: 26886345]
3. Costa CP, Moreira JN, Sousa Lobo JM, Silva AC. Intranasal delivery of nanostructured lipid carriers, solid lipid nanoparticles and nanoemulsions: a current overview of in vivo studies. Acta Pharm Sin B 2021 Apr;11(4):925-940. [doi: 10.1016/j.apsb.2021.02.012] [Medline: 33996407]
4. Erdő F, Bors LA, Farkas D, Bajza Á, Gizurarson S. Evaluation of intranasal delivery route of drug administration for brain targeting. Brain Res Bull 2018 Oct;143:155-170 [FREE Full text] [doi: 10.1016/j.brainresbull.2018.10.009] [Medline: 30449731]
5. Wu H, Hu K, Jiang X. From nose to brain: understanding transport capacity and transport rate of drugs. Expert Opin Drug Deliv 2008 Oct;5(10):1159-1168. [doi: 10.1517/17425247.5.10.1159] [Medline: 18817519]
6. Vinjamury SP, Vinjamury M, Sucharitakul S, Ziegler I. Panchakarma: Ayurvedic detoxification and allied therapies—is there any evidence? In: Rastogi S, editor. Evidence-Based Practice in Complementary and Alternative Medicine. Berlin, Germany: Springer; 2012.
7. Agnivesa. Matrashiteeya adhyaya. In: Srikantha Murthy KR, editor. Caraka Samhitha. Varanasi, India: Chaukhambha Orientalia; 1995.
8. Singh N, Chaudhary A. A comparative review study of Sneha Kalpana (Paka) vis-a-vis liposome. Ayu 2011 Jan;32(1):103-108 [FREE Full text] [doi: 10.4103/0974-8520.85740] [Medline: 22131767]
9. Singh SK, Rajoria K, Kumar A, Sharma S. A double-blind controlled clinical trial to evaluate the effects of nasal therapy with Vrihatajivakadya oil on different viscosities in patients with migraine. J Ayurveda Integr Med 2023 Mar;14(2):100662 [FREE Full text] [doi: 10.1016/j.jaim.2022.100662] [Medline: 36384709]
10. Vagbhata A. Nasya vidhi adhyaya. In: Srikantha Murthy KR, editor. Astanga Samgraha of Vagbhata. Varanasi, India: Chaukhambha Orientalia; 2012.

XSL•FO
RenderX

11. Wood JL, Weintraub S, Coventry C, Xu J, Zhang H, Rogalski E, et al. Montreal Cognitive Assessment (MoCA) performance and domain-specific index scores in amnestic versus aphasic dementia. J Int Neuropsychol Soc 2020 Oct;26(9):927-931. [doi: 10.1017/S135561772000048X] [Medline: 32423505]
12. Harris KM, Gaffey AE, Schwartz JE, Krantz DS, Burg MM. The Perceived Stress Scale as a measure of stress: decomposing score variance in longitudinal behavioral medicine studies. Ann Behav Med 2023 Sep 13;57(10):846-854 [FREE Full text] [doi: 10.1093/abm/kaad015] [Medline: 37084792]
13. Thompson E. Hamilton Rating Scale for Anxiety (HAM-A). Occup Med (Lond) 2015 Oct 13;65(7):601. [doi: 10.1093/occmed/kqv054] [Medline: 26370845]
14. Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: scoping review. SAGE Open Med 2016 Oct 4;4:2050312116671725 [FREE Full text] [doi: 10.1177/2050312116671725] [Medline: 27757230]

Abbreviations

AIMS: Amrita Institute of Medical Sciences
BBB: blood-brain barrier
BOLD: blood oxygen level–dependent
CNS: central nervous system
FDR: false discovery rate
fMRI: functional magnetic resonance imaging
FWE: familywise error
MRI: magnetic resonance imaging
NARIP: National Ayurveda Research Institute for Panchakarma
SPM: statistical parametric mapping

Submitted 05.Oct.2024; this is a non–peer-reviewed article;accepted 30.Jul.2025; published 09.Jan.2026.

Please cite as:
Nair DR, C J S, Jameela S, E R, Moorthy S, Parasuram H, Khanduri S, Rao BC, Srikanth N, Acharya R
Impact of Intranasal Administration of Ayurveda Medicine in Apparently Healthy Individuals on Neurophysiological Variables and Functional Connectivity Using Functional Magnetic Resonance Imaging: Protocol for an Exploratory Randomized Controlled Trial
JMIR Res Protoc 2026;15:e67132
URL: https://www.researchprotocols.org/2026/1/e67132
doi:10.2196/67132
PMID:

©Devi R Nair, Sandya C J, Sophia Jameela, Remya E, Srikanth Moorthy, Harilal Parasuram, Shruti Khanduri, Bhogavalli Chandrasekhara Rao, Narayanam Srikanth, Rabinarayan Acharya. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 09.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

An App-Based Behavioral Support Intervention Promoting Physical Activity (APPROACH) in Patients Diagnosed With Breast, Prostate, or Colorectal Cancer: Protocol for a Randomized Controlled Trial

Fiona Kennedy[1], BSc, MSc, PhD; Sue L Smith[2], BA, MSc; Rebecca J Beeken[1], BA, MSc, PhD; Caroline Buck[2], BA, MSc, DHealthPsych; Leanne Shearsmith[1], BSc; Hannah Truscott[1], BSc, MSc; Nicholas Counsell[3], MSc; Chloe Thomas[4], BSc, MSc, PhD; Anna Roberts[2], BSc, MSc, PhD; Diana M Greenfield[5], BSc, PhD; Henry WW Potts[6], BA, MSc, PhD; Nicholas Latimer[4], BSc, MSc, PhD; Ed Lowther[7], MA; Lee Smith[8], BSc, MSc, PhD; Rosie Stevens[1], BSc; Amy Creaser[1], BSc, PhD; Jacqui Gath[9], MBCS, CITP, CEng; Lynda Wyld[10], BMedSci, MBChB, PhD, FRCS(Gen Surg), FEBS; Abigail Fisher[2*], BSc, PhD; Phillippa Lally[11*], BSc, MSc, PhD

[1]Leeds Institute of Health Sciences, University of Leeds, Leeds, United Kingdom

[2]Department of Behavioural Science and Health, University College London, London, United Kingdom

[3]Cancer Research UK & UCL Cancer Trials Centre, Cancer Institute, University College London, London, United Kingdom

[4]Sheffield Centre for Health and Related Research, School of Medicine and Population Health, University of Sheffield, Sheffield, United Kingdom

[5]Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, United Kingdom

[6]Institute of Health Informatics, University College London, London, United Kingdom

[7]UCL Advanced Research Computing Centre, University College London, London, United Kingdom

[8]The Centre for Health, Performance and Wellbeing, Anglia Ruskin University, Cambridge, United Kingdom

[9]Independent Cancer Patients' Voice (ICPV), London, United Kingdom

[10]Department of Oncology and Metabolism, University of Sheffield, Sheffield, United Kingdom

[11]University of Surrey, Guildford, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Abigail Fisher, BSc, PhD
Department of Behavioural Science and Health
University College London
1-19 Torrington Place
Room 217
London, WC1E 7HB
United Kingdom
Phone: 44 20 7679 1722
Email: abigail.fisher@ucl.ac.uk

Abstract

Background: Strong evidence highlights that sufficient physical activity (PA) has multiple benefits for people living with and beyond cancer. However, many are not meeting PA recommendations. APPROACH is a trial of a theory-driven, app-based behavioral support intervention to promote brisk walking after breast, prostate, or colorectal cancer.

Objective: The aim of this trial is to evaluate the efficacy and cost-effectiveness of the intervention.

Methods: APPROACH is a multicenter, phase III, 2-armed, individually randomized controlled trial (N=472). We will recruit patients with localized breast, prostate, or colorectal cancer from hospitals in Yorkshire and surrounding areas in the North of England, United Kingdom, and randomize them 1:1 between the intervention and control arm (usual care). The intervention consists of an app designed for the general population to encourage brisk walking (NHS Active 10), supplemented with habit-based behavioral support, including 2 brief telephone or video calls, a leaflet, website, and walking planners. The primary endpoint is the difference between trial arms in the changes from baseline in activPAL-assessed average minutes of brisk walking (≥100 steps per minute) after 3 months. Demographic and medical characteristics will be collected through self-report and hospital records. Secondary outcomes (assessed at 0, 3, and 6 months) will be the other activPAL-assessed outcomes (brisk walking at 6 months, total steps, light PA, standing time, and sitting times, weekly metabolic equivalent of task), self-reported PA, and

XSL•FO
RenderX

self-reported BMI and waist circumference. Patient-reported outcome measures of quality of life, fatigue, sleep, anxiety, depression, self-efficacy, habit strength for walking, and social support will also be collected. Interviews will explore experiences of receiving the intervention. We will use health economic modeling to estimate the cost-effectiveness of the intervention over a lifetime horizon.

Results: The study was funded in June 2019. Trial recruitment commenced in November 2023 and is planned to be completed in 2025. As of December 2025, a total of 473 participants have been randomized. The publication of the main results is expected in autumn 2027 after all follow-up data collection and analysis are complete.

Conclusions: Overall findings will determine the clinical and cost-effectiveness of the intervention for patients diagnosed with breast, prostate, or colorectal cancer. If successful, APPROACH provides a potential model of supportive care to increase PA among people living with and beyond cancer.

Trial Registration: ISRCTN Registry ISRCTN14149329; https://www.isrctn.com/ISRCTN14149329

International Registered Report Identifier (IRRID): DERR1-10.2196/77096

(JMIR Res Protoc 2026;15:e77096) doi:10.2196/77096

KEYWORDS

app; cancer; cost-effectiveness; habit-based behavioral support; physical activity; quality of life; randomized controlled trial

Introduction

The global burden of cancer has more than doubled in the last 20 years [1,2]. Due to advancements in early detection and treatments [3], the number of people surviving cancer is increasing, with 5-year survival rates for some cancers such as breast cancer as high as 86% [4]. As survival rates continue to improve, there is a need to support the ever-growing population of people diagnosed with cancer, enabling them to live well and in good health [5].

The safety and feasibility of physical activity (PA) for oncology patients have been demonstrated by hundreds of trials. Meta-analyses exploring these have shown that PA is safe, feasible, and has beneficial effects on both physical and psychological functioning [6,7]. Benefits include reduced fatigue and pain, improvement in psychological and physical well-being [6,8], improved quality of life [9-13], and evidence of reduced risk of cancer-specific and all-cause mortality [14-19]. Breast, prostate, and colorectal cancer comprise 3 of the most commonly diagnosed cancers in the United Kingdom [20], and the evidence is strongest among these cancer types for the positive effects of PA after a cancer diagnosis [21-23].

Recognizing this evidence, the World Cancer Research Fund published recommendations for people diagnosed with cancer to engage in at least 150 minutes of moderate-to-vigorous physical activity (MVPA) per week, in line with those for the general population [24]. This report also highlighted that PA is not only exercise per se but includes everyday activities such as walking. More recently, these recommendations have been expanded to include efforts to reduce sedentary time. New guidance to “move more, sit less” encompasses the recognized benefits of an overall more active movement profile for people after a cancer diagnosis [25,26]. However, it is estimated that only around 30% of people diagnosed with cancer meet the recommended amounts of PA [27,28], and therefore there is a need to develop interventions that can increase PA levels in this population.

In the United Kingdom, a recommendation for PA advice for people diagnosed with cancer is included within national documents such as the Independent Cancer Taskforce [29], the 2019 National Health Service (NHS) Long Term Plan (albeit health and well-being focused) [5], and cancer charity reports [30]. Despite this and a documented desire among patients to receive advice [31,32], many people diagnosed with cancer do not recall receiving PA advice [33-35]. In one large study, less than a third of 15,254 people diagnosed with colorectal cancer recalled receiving any PA advice as part of their care [34]. In a survey undertaken with 242 people diagnosed with cancer in the United Kingdom, around half recalled receiving PA advice, but only 30% received guidance on the type of PA, and most (85%) were not referred to further information or an exercise specialist [36]. Reasons for this lack of PA advice in clinical practice include health care professionals not having the time and resources to provide advice, insufficient time in appointments, lack of knowledge of what to recommend, and not feeling they are the “right person” to deliver this advice [31,35,37]. These findings demonstrate the need for low-cost, scalable interventions that can be integrated into care and are accessible to a large number of patients.

Digital interventions offer the potential to reach wide populations, with smartphone apps becoming increasingly popular for harnessing behavior change. While a meta-analysis of 15 studies found that digital interventions could increase participation in MVPA by approximately 40 minutes per week in people diagnosed with cancer, only 2 of the included studies included delivery of the intervention by apps [38]. Apps can be used to monitor behaviors, track progress toward goals, and have been used to deliver “in the moment” behavior change techniques (BCTs), such as push notifications to remind people to complete activity [39]. Systematic reviews have also demonstrated that smartphone interventions offer potential to increase PA in people diagnosed with cancer and acknowledged the importance of cost-effectiveness when developing apps for intervention use, particularly focusing on the need to establish long-term changes in behavior that will be maintained [40,41].

XSL•FO
RenderX

To facilitate the maintenance of behaviors, habit theory posits that a habitual behavior is generated through a process where a context-response association is learned through repetition [42,43]. Theory suggests that contextual cues can then trigger the automatic execution of this behavior, even where motivation to perform the behavior in the moment is low [42,43]. Applying this theory in health behavior interventions could help encourage behavior maintenance beyond the intervention period, increasing the chances of long-term enactment of the health-promoting behavior [44,45].

This evidence suggests that an app-based intervention augmented with habit theory and personal contact has the potential to promote PA in people diagnosed with cancer. To explore this, the research group conducted qualitative user experience research with 31 people diagnosed with breast, prostate, or colorectal cancer who used 4 PA apps to identify their preferences [46]. No apps specifically designed to increase PA among people diagnosed with cancer were identified at the time, but the apps included were selected to represent different delivery formats (eg, tracking versus video instruction) and types of PA (eg, walking versus strength training), and were reviewed by a specialist oncology physiotherapist to check their safety for people with cancer [46]. In line with previous research [47], brisk walking was identified as the most achievable form of MVPA for people with cancer [46]. Participants also reported a preference for an app-based intervention that could be tailored to their ability and placed importance on the app being recommended by their health care professional, as well as being "badged" by or linked to the NHS [46]. The publicly available NHS Active 10 app [48] was chosen as it was the app that best captured all of these features, as well as including many of the key BCTs shown to be associated with successful PA change, including the focus on walking, goal setting, and self-monitoring [44,49-51]. Selecting an existing publicly available NHS-hosted app was a more cost-effective option with no development and maintenance costs, and it made it more likely that the app would be sustained and exist throughout our pilot, efficacy trial, and beyond. Qualitative research by our group with oncology clinical nurse specialists demonstrated that they would support an app-based intervention within the care pathway and would be happy to recommend this to patients, provided the app had demonstrated efficacy and was from a reputable source [52], the latter of which was fulfilled by using the NHS-developed and maintained Active 10 app, meaning implementation would be more likely.

This randomized controlled trial was informed by a pilot study [44] that demonstrated the feasibility and acceptability of our trial procedures and intervention content in 90 patients recruited from a single hospital in the United Kingdom. Small changes that were made to the intervention following the pilot are detailed in the "Intervention" section below [53]. The pilot study results also informed the sample size parameters for the trial's primary endpoint. This study aims to investigate the efficacy of an app-based brisk walking intervention with behavioral support to increase PA in people diagnosed with breast, prostate, or colorectal cancer in the United Kingdom.

Methods

Primary Aim

The main objective is to increase brisk walking (activPAL-assessed average minutes ≥100 steps per minute [spm]) by using a theory-driven, app-based intervention that promotes brisk walking in adults living with breast, prostate, or colorectal cancer.

Secondary Aims

The main secondary objectives are to compare the effect of the intervention with usual care on other activPAL-assessed outcomes (including average total steps, light PA, standing time and sitting time, weekly metabolic equivalent of task [MET], self-reported BMI and waist circumference, self-reported PA, habit strength for walking, PA self-efficacy, health status, quality of life outcomes, and health care usage) at 3 and 6 months. The cost-effectiveness of the intervention compared with usual care will be assessed as a secondary outcome. Engagement and satisfaction with the intervention will also be explored.

Design

APPROACH is a multicenter, phase III, individually randomized controlled trial whereby people with a diagnosis of breast, prostate, or colorectal cancer are recruited and randomly allocated (1:1) between the intervention and control (usual care) arms. Study assessments are collected remotely at baseline, 3 months (12-18 weeks), and 6 months (24-30 weeks) after randomization. Figure 1 illustrates the flow of participants through the study. The trial has been designed following the CONSORT (Consolidated Standards of Reporting Trials) statement [54] and has been registered on the ISRCTN Registry [55]. The trial has been conducted in accordance with this trial registry record with no major deviations, only minor changes to trial documents, an extension to the recruitment timescale, and the addition of new recruitment hospital locations to boost recruitment. The reporting of this protocol follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines (Multimedia Appendix 1 contains a completed SPIRIT checklist) [56]. A schedule of enrollment, interventions, and assessments based on the SPIRIT guidelines is shown in Table 1.

XSL•FO
RenderX

Figure 1. Flow chart of patient recruitment and involvement in study.

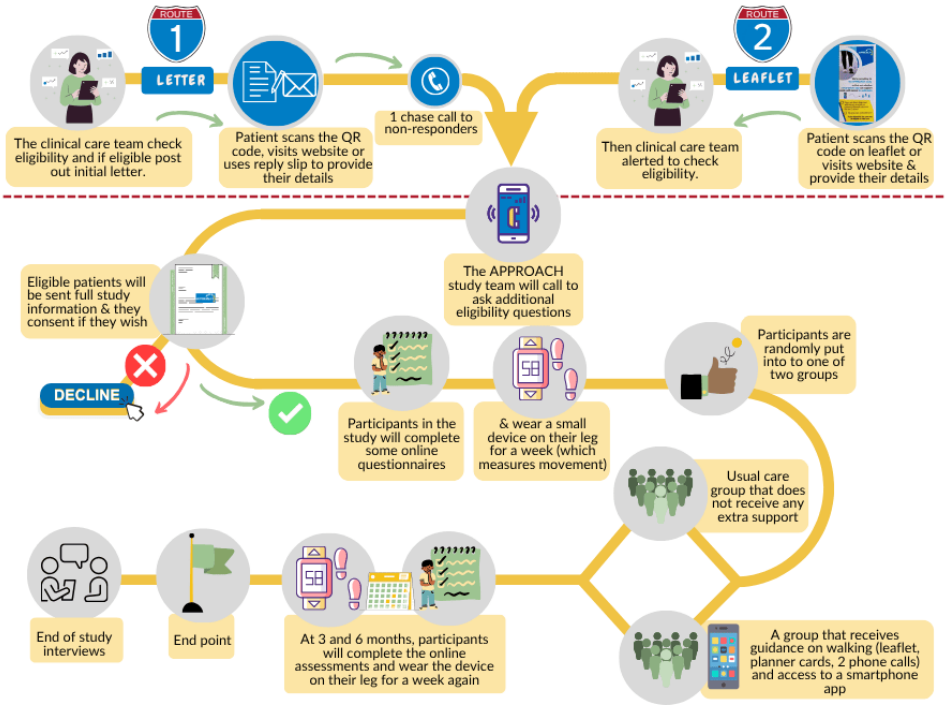

XSL•FO
RenderX

Table 1. SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) schedule of enrollment, interventions, and assessments.

Assessment and activity	Enrollment	Allocation	Post allocation			
	Any time during recruitment window		Weeks 1-12 (intervention)	3-month (12-18 weeks)	6-month (24-30 weeks)	Post study (30+ weeks)
Enrollment						
Screening eligibility	✔					
Eligibility confirmation	✔					
Informed consent	✔					
Allocation		✔				
Intervention						
Intervention group						
Usual care control group			✔			
Assessments						
Demographics and diagnosis	✔					
Individual differences (preference for routine and chronotype)	✔					
Self-reported physical activity (GLTEQ[a])	✔			✔	✔	
Health-related quality of life (EQ-5D-5L)	✔			✔	✔	
Cancer-specific quality of life (FACT-G[b])	✔			✔	✔	
Fatigue (FACIT-F[c])	✔			✔	✔	
Sleep (PSQI[d])	✔			✔	✔	
Anxiety (GAD-7[e])	✔			✔	✔	
Depression (PHQ-9[f])	✔			✔	✔	
Physical activity self-efficacy (PAAI[g])	✔			✔	✔	
Self-efficacy for self-management of cancer (CS-SES[h])	✔			✔	✔	
Habit strength for walking (SR-BAI[i])	✔			✔	✔	
Social support (F-Soz-U K-6[j])	✔			✔	✔	
Health and social care service use (CSRI[k])	✔			✔	✔	
Height, weight, and waist circumference (optional)	✔			✔ (not including height)	✔ (not including height)	
Accelerometry (activPAL) (3-7 days)	✔			✔	✔	
Poststudy qualitative interviews						✔ (intervention group only)
Online Intervention feedback questionnaire				✔ (intervention group only)	✔ (intervention group only)	

[a]GLTEQ: Godin Leisure Time Exercise Questionnaire.

[b]FACT-G: Functional Assessment of Cancer Therapy-General.

[c]FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue.

[d]PSQI: Pittsburgh Sleep Quality Index.

XSL•FO
RenderX

[e]GAD-7: Generalized Anxiety Disorder-7.

[f]PHQ-9: Patient Health Questionnaire-9.

[g]PAAI: Physical Activity Appraisal Inventory.

[h]CS-SES: Cancer Survivorship Self-Efficacy Scale.

[i]SRBAI: Self-Report Behavioral Automaticity Index.

[j]F-Soz-U K-6: 6-item brief form of the Perceived Social Support Questionnaire.

[k]CSRI: Client Service Receipt Inventory.

Eligibility Criteria

Table 2 outlines the inclusion and exclusion criteria across the eligibility screening phases.

Table 2. Eligibility criteria for each screening phase.

Screening phase	Inclusion criteria	Exclusion criteria
Phase 1: medical record (NHS[a] site)	• Confirmed diagnosis of breast, prostate, or colorectal cancer within the past 2 years and registered as a patient at one of the recruiting hospitals • Aged ≥16 years	• Diagnosed with a cognitive impairment (eg, dementia) • Cognitive or physical impairment that prevents participation in brisk walking[b] • Metastatic disease • Incarcerated
Phase 2: participant screening by telephone (APPROACH team)	• Owns a smartphone using either Android or iOS operating systems • Willing and able to provide informed consent • Has access to a computer or tablet and an email address, and is willing to complete online questionnaires • Willing to wear an accelerometer for 1 week at each assessment point	• More than 2.5 years postdiagnosis at consent[c] • Unable to understand spoken or written English • Eastern Cooperative Oncology Group performance status of ≥3 (limited self-care; confined to bed or chair for ≥50% of waking hours) • Cognitive or physical impairment that prevents participation in walking[b] • Scheduled to undergo surgery to remove cancer within the next 5 months at the time of screening • Fewer than 6 weeks after surgery for cancer removal at baseline assessment • Report ≥150 minutes of at least moderate-to-vigorous physical activity per week • Report previous or current use of the intervention app • Currently participating in another health behavior change study (eg, diet, physical activity, alcohol intake, and weight management) • Lives with someone already participating in the trial • Participated in the pilot study (only applies at one site)

[a]NHS: National Health Service.

[b]This criterion is screened twice: first in the medical records and again with participants.

[c]The 6-month gap between medical-record screening (<2 years postdiagnosis) and direct participant screening (2.5 years postdiagnosis) recognizes the potential time between screening and consent.

Recruitment and Setting

Recruitment takes place across multiple hospitals in Yorkshire (our funding stipulates that the majority of recruitment is from the Yorkshire area) and surrounding areas in the North of England, United Kingdom (sites are listed on the ISRCTN trial record [55]). At each site, the clinical care team identifies patients who have been diagnosed with breast, prostate, or colorectal cancer within the last 2 years and examines their medical records for the phase 1 inclusion criteria (Table 2). NHS staff mail or email eligible patients a letter about the study asking them to indicate (by online form, paper reply slip, or telephone) whether the APPROACH team can contact them. If site capacity allows, after at least 2 weeks, nonresponders are telephoned or emailed by hospital staff. Study leaflets and poster advertisements to be distributed in hospital waiting rooms and clinic areas are also available, enabling patients to self-refer (self-referral triggers a check of phase 1 eligibility by the NHS team).

If a patient expresses interest and agrees to a call, the APPROACH team telephones to conduct phase 2 eligibility screening (Table 2) and provide the participant information sheet by email if eligible. The participant information sheet was redesigned following the pilot study to ensure the language was accessible, the format easy to read (including a visual infographic to describe trial activities), and the images culturally sensitive and relevant. These materials were also reviewed and approved by our patient and public involvement members. The email links to a REDCap (Research Electronic Data Capture) online website [57,58] (electronic survey software hosted within the encrypted Data Safe Haven platform at University College London [UCL]) where the full participant information sheet is displayed and the online consent form provided (Multimedia Appendix 2). Patients can request a paper information sheet if desired. Consented participants complete the baseline assessments comprising an activPAL (activPAL4micro accelerometer; PAL Technologies Ltd), an online questionnaire, and optional weight, height, and waist circumference

XSL•FO
RenderX

measurements (provision of weighing scales and tape measure). Participants can also opt into allowing access to and sharing of their NHS number with Hospital Episode Statistics (HES) and the National Disease Registration Service (NDRS) registries to enable exploration of long-term (10-year) health outcomes.

Sample Size

The APPROACH pilot study data [53] were used to inform this power calculation. The target sample size is 472 participants randomized 1:1 to detect a difference between trial arms in the change from baseline in the daily average of 6 minutes per day (ie, 42 minutes per week) of activity at 100 spm or higher. This assumes 90% power, a 2-sided 5% significance level, an SD of 12 minutes per day in the control arm, and a variance ratio of 1:4 (control:intervention). Low attrition (3%) was observed in the pilot, but with multiple recruiting sites and longer follow-up, a 10% attrition rate has been allowed to ensure the required 212 participants per arm is achieved.

Randomization

Post baseline data collection, participants are individually randomized using minimization with a 1:1 allocation ratio by the central trial team at UCL (a researcher not involved with recruitment or data collection) using code developed by a UCL senior research data scientist, based on the paper by O'Callaghan [59], which is available on GitHub [60]. Stratification factors are cancer type (breast, prostate, colorectal) and hospital site. The first participant is randomly allocated; thereafter, subsequent participants are assigned based on the arm with the lowest imbalance score while factoring in a 20% random element. Randomization constitutes full enrollment in the trial.

Intervention

Overview

The intervention consists of the app (NHS Active 10) plus supplementary brief written and telephone behavioral support [44,53] (Multimedia Appendix 3 provides a Template for Intervention Description and Replication [TIDieR] checklist). The intervention has been designed and coded according to the Behavior Change Taxonomy Ontology reported by Michie et al [61] (Table 3 shows the intervention component mapping carried out by PL, FK, CB, SLS, and AF) [62,63].

XSL•FO
RenderX

Table 3. Intervention components and mapping to the Behavior Change Taxonomy Ontology (BCTO) items.

Intervention components and BCTO items	BCTO item keys[a]
Active 10 App	
Introducing the app into participants' environment	007156
The App is hosted by NHS[b] Better Health	007075
Introduction: "Brisk walking is…" "Every minute counts" "Aim for 10 min or more a day"	007058
I am doing Active 10 because…	007005
Set your targets (1, 2, or 3 Active 10s a day)	007300
If regularly achieve 3 Active 10s a day can set a target of 4 or 5	007100
Walking tracker (minutes of walking and minutes of brisk walking)	007023
Rewards	007253, 007254, 007209, 007211
Optional links and functions, including a variety of "tips," which participants may or may not access	
Links to useful websites (eg, NHS, Strava)	N/A[c]
Link to a discussion group for the app	007029
Articles on physical and mental health benefits of physical activity, disabilities, how much physical activity to do, other ways to improve mental health (eg, meditation)	007063, 007065
Ability to set reminders	007080
Pace checker	007017
Tip: How fast is a "brisk" walk	007042
Tip: set a reminder	007080
Tip: keep track (use app to see how you are doing)	007023
Tip: schedule a time for a walk	007010
Tip: walk with a friend	007029
Tip: prepare (get shoes ready)	007168
Tip: Listen to music or podcast	007061
Tip: reward yourself	007187
Leaflet and accompanying letter from clinical team	
Clinical team recommendation to read and use the information provided and to download Active 10	007074, 007075, 007077
Letter and leaflet informing that Active 10 was developed by NHS	007075
Branding: Yorkshire Cancer Research, UCL[d], University of Leeds, University of Sheffield, Anglia Ruskin University, and Sheffield Teaching Hospitals.	007075
Physical activity improves side effects of cancer treatment, recovery and risk of recurrence, mood and confidence. Physical activity reduces risks of other health problems.	007183, 007181
Quotes from patients with cancer: used physical activity to cope with fatigue, chose walking to try to meet guidelines	007074, 007183, 007181
People who have or have had cancer recommended to try to meet the same physical activity guidelines as other adults. Brisk walking 2-3 times every day will meet the activity guidelines (150 minutes); the more, the better	007075, 007076
Recommends brisk walking. This should make you breathe a bit faster…	007058
Recommends start small then build up	007100
Information on downloading Active 10	007058
Recommends planning	007008
Recommends walking at the same time or in the same situation	007094, 007096
Recommends tracking behavior using the walking planners and Active 10	007076
Recommends using the website for more information	N/A

XSL•FO
RenderX

Intervention components and BCTO items	BCTO item keys[a]
Walking planner	
Adding planner to people's environment	007164
Promotes habit formation	007094, 007096
Promotes adapting plans as needed	007299
If your days are different, try to develop a few different habits...	007010
How many Active 10s are you aiming for	007003
Plan: when and how much	007010
Did you complete your plans?	007024
How did you feel after you walked briskly?	007066, 007025
Reminder not to worry if you miss a day and to adjust plans as required (reduce if finding it hard and increase if meeting goals and feel able).	007299
Website	
Website and video branded from Yorkshire Cancer Research	007075
Links to resources about the health benefits of physical activity	007183, 007181
Links to resources to support walking (general walking websites and local information for areas recruiting from)	007042
Promote physical and mental health benefits of brisk walking	007183, 007181
Video: (1) promotes the physical and mental health benefits of brisk walking; (2) states that participants report enjoying using the app, liking having control and doing something for their health, and liking monitoring progress; and (3) shows a participant describing how they built up walking over time, made it part of their routine, and reported less physical pain and feeling emotionally better	007175, 007068, 007120, 007072, 007076
Intervention: phone call or video call (1)	
Introduce self as working with clinical team at the hospital	007075
Ask participants how their cancer and treatment have impacted their lifestyle and activity levels	
Discuss physical and mental health benefits of physical activity	007183, 007181
Discuss motivations to increase activity	007069
Discuss concerns about increasing activity	007069, 007008
Help participant to work out ways to overcome concerns about brisk walking, provide information as appropriate	007008
Discuss why recommending brisk walking, including cancer patients have recommended this	007074
Describe brisk walking	007058, 007051
Provide information on government guidelines (150 minutes of MVPA[e]) as well as WHO[f] Every Move Counts.	007075
Highlight building up over time	007100
Discuss how confident they are and how they can increase their confidence	007008
Suggest trying it to see if that increases their confidence	007054, 007137
If needed tell them that it is possible for them to do this and others have been able to	007137, 007076
Promote habit formation for initiating a walk	007094, 007096
Promote habit formation for planning walks, looking at app or planner	007094, 007096
Make an action plan (when, what, how long)	007010
Promote self-reward during or after walking	007257, 007252
Promote nonspecific self-reward during or after walking	007257, 007252
Promote using the app to track activity	007026
Promote specific cues	007080, 007081
Promote reminders	007080, 007081

XSL•FO
RenderX

Intervention components and BCTO items	BCTO item keys[a]
Promote adjustment of goals if needed	007011
Promote adjustment of plans if needed	007299
Set a target number of Active 10s	007300, 007010
Promote asking friends to support, by encouraging and helping to remember to walk	007030, 007031, 007032, 007033
Promote using the walking planner to track behavior	007024, 007025
Encourage participants to use information provided to overcome their concerns about exercising	007008
Intervention: phone call or video call (2)	
Remind them of their target, and discuss if needs adapting	007011
Ask how they are getting on with their target	007012
Ask participants what is preventing them from walking and what would help them to start (if relevant)	007008
Repeat any of the points from call 1 as appropriate	As relevant

[a]BCTO item key: 007003 set behavior goal behavior change technique (BCT); 007005 set outcome goal; 007008 goal strategizing BCT; 007010 action planning; 007011 review behavior goal BCT; 007012 attend to discrepancy between current behavior and goal BCT; 007017 monitoring BCT; 007023 provide feedback on behavior; 007024 self-monitor behavior BCT; 007025 self-monitor outcome of behavior BCT; 007026 provide biofeedback BCT; 007029 advise to seek support; 007030 advise to seek instrumental support BCT; 007031 advise to seek emotional support BCT; 007032 advise to seek informational support BCT; 007033 advise to seek appraisal support BCT; 007042 deliver informational support; 007051 agree on how to perform behavior BCT; 007054 conduct a behavioral experiment BCT; 007058 instruct how to perform behavior; 007061 plan inclusion of enjoyment; 007063 inform about health consequences; 007065 inform about emotional consequences; 007066 monitor emotional consequences BCT; 007068 increase salience of consequences BCT; 007069 consider pros and cons BCT; 007072 awareness of other people's thoughts, feelings, and actions BCT; 007074 increase awareness of others approval BCT; 007075 present information from credible influence BCT; 007076 suggest to change behavior BCT; 007077 tell to change behavior BCT; 007080 prompt intended action BCT; 007081 cue BCT; 007094 practice behavior BCT; 007096 context-specific repetition of behavior BCT; 007100 set graded tasks BCT; 007120 vicarious reward BCT; 007137 persuade about personal capability BCT; 007156 add objects to the environment BCT; 007164 add objects to the indirectly experienced environment BCT; 007168 advise specific behavior BCT; 007175 increase salience of behavior BCT; 007181 inform about positive emotional consequences BCT; 007183 inform about positive health consequences BCT; 007187 promise consequence for behavior BCT; 007209 promise positive material consequence for behavior BCT; 007211 promise positive material consequence for approximating behavior BCT; 007252 provide positive consequence for behavior BCT; 007253 provide positive consequence for approximating behavior BCT; 007254 provide positive consequence for completion of behavioral sequence BCT; 007257 provide positive material consequence for behavior BCT; 007299 review behavior goal plan BCT; and 007300 set measurable behavior goal BCT.

[b]NHS: National Health Service.

[c]N/A: not applicable.

[d]UCL: University College London.

[e]MVPA: moderate-to-vigorous physical activity.

[f]WHO: World Health Organization.

Several BCTs have shown promise in systematic reviews and meta-analyses focused on increasing PA in inactive adults [49] and specifically after cancer, including adding an object to the environment, self-monitoring, goal setting [64], action planning, social support [65], and rewards [66]. This study focused on BCTs that were considered to have efficacy in this population and those practical to use in the context of an app-based intervention with brief behavioral support.

Intervention participants are mailed a pack containing a clinician endorsed letter, alongside printed materials (leaflet and walking planner cards) and an appointment time for a behavioral support video or telephone call. Since the trial is ongoing at the time of writing this protocol, intervention materials are not included, but the mapping exercise is provided (Table 3) and further detail is available upon request from the corresponding author.

Leaflet and Website

The leaflet covers topics including the benefits of PA during or after cancer treatment (eg, reduces side effects), introduces how walking and particularly brisk walking counts as moderate PA (encouraging individuals to gradually build up), and walking habits (eg, importance of repetition, planning walking into daily routines). The leaflet also introduces the NHS Active 10 app (including how to download it), and has a link to the study website (not used in the pilot as all information was included in the leaflet because the pilot was conducted in one location), which contains copies of all the study resources and also local area-specific resources relating to walking groups or places to walk, which were accumulated by doing internet searches of walking resources in these areas (eg, Get Doncaster Moving council initiative, 5k Your Way). The website used in the randomized controlled trial also includes a brief film of participants from the pilot [53] who used the app and increased their walking.

Walking Planner Cards

The walking planner cards are underpinned by BCT principles and habit theory that have been shown to increase PA, including a focus on helping participants set goals (eg, "How many daily Active 10s are you aiming for?"), plan their walks (action planning, eg, "On Monday I will walk:…"), and self-monitor (eg, "Tick when you complete each plan"). These cards were amended based on feedback from pilot participants [53], many

XSL•FO
RenderX

of whom felt they were too restrictive. The core content remained (eg, prompting the time of day or situation), but the redesign enabled more flexibility for different plans on different days of the week.

App

The NHS Active 10 app is publicly available, developed by Public Health England [48], and currently maintained by the Department of Health and Social Care. This app was selected based on extensive development work [46,52,67], plus patient and public involvement (PPI) and pilot study results [53]. The app incorporates BCTs such as goal setting, self-monitoring behavior, and prompting practice of behaviors [68,69].

The app tracks all walking activity and provides a total for time spent walking and time spent brisk walking. Ten minutes of brisk walking is known as one "Active 10." The app encourages users to initially build up to completing one Active 10. Users can set a goal to complete 1-3 Active 10s per day, with the aim of reaching 30 minutes, and if successful they can set up to 5 Active 10s. Each individual minute of brisk walking is recorded and contributes to the Active 10s (to fit with current UK Chief Medical Officers' PA guidance that updated the previous recommendation that a minimum 10-minute bout of activity was required [70]). Moderate activity has been shown to provide greater health benefits than slower-paced walking [71]. The app uses the mobile phone's inbuilt sensors to measure activity and record the total minutes of walking and, where relevant, brisk walking (100+ spm).

Since the pilot study was completed, a "pace checker" function has been added, enabling users to see timed pulses on the app (in line with 100 spm) while walking for a 1-minute period and, at the end of the 1 minute, shows the number of steps achieved and whether it was classed as brisk. Figure 2 illustrates some screenshots from the app, including the new feature.

Figure 2. Screenshots of Active 10 app.

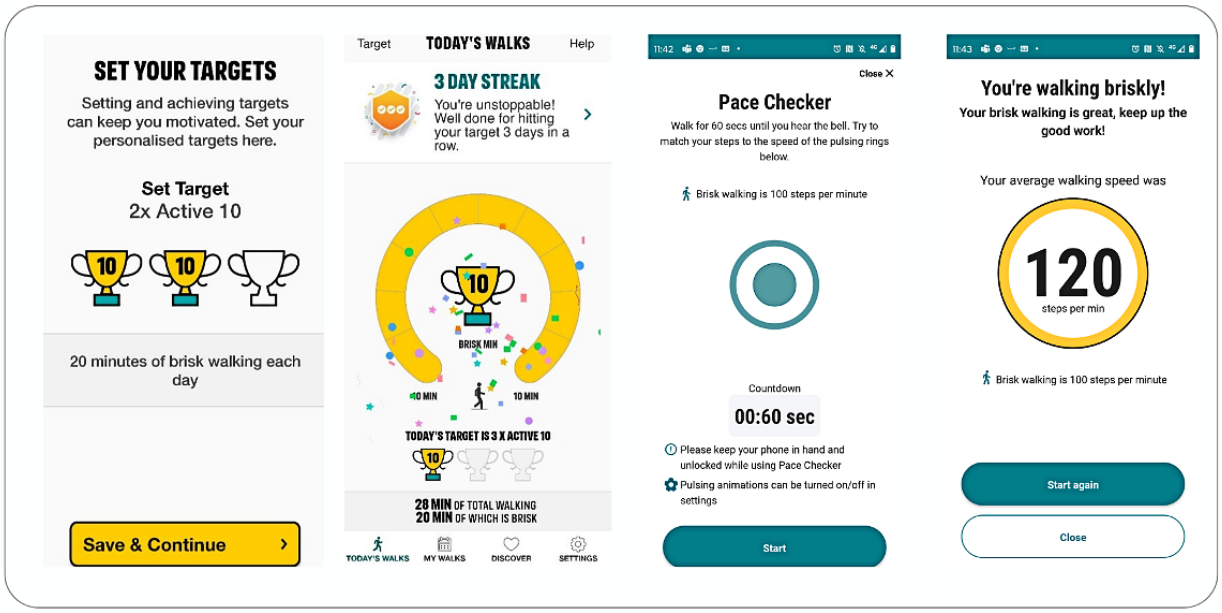

Behavioral Support Calls

Two semiscripted behavioral support video or telephone calls are undertaken by a trained researcher around one month apart. The first call takes place as soon after randomization as possible and introduces the recommendations around PA for people diagnosed with cancer (eg, 150 minutes of MVPA), how brisk walking is a form of moderate PA, the evidence for benefits of meeting these guidelines, and the importance of increasing PA by any amount (eg, highlighting the Every Movement Counts World Health Organization campaign) [72]. The researcher also explains how to form walking habits, how the app works, and encourages use of the app to help set and monitor goals (eg, "What do you want to aim for? One, two or three active tens [so 10, 20 or 30 minutes of brisk walking a day]"). Participants are encouraged to make plans and establish habits for using the app, plan to increase their brisk walking if appropriate, and use the app to track their progress. The call also aims to explore barriers to PA and explore ways of addressing them drawn from evidence (eg, treatment-related fatigue). This call is intended to replicate the type of conversation that a health care professional (eg, clinical nurse specialist) could have with patients as part of routine care. A second call is arranged for around 4 weeks later to see how participants are progressing, troubleshoot (eg, their plans or the app), and remind and recap on any information as required. With the participant's agreement, calls are recorded to allow fidelity checks.

Control Group

Control group participants are notified of their group allocation by telephone or email, then receive "usual care" (no additional support beyond being informed in the information sheet that PA is beneficial for people after a diagnosis of cancer).

Measures

Overview

For all assessments, participants are emailed a link to a questionnaire on REDCap [57,58] and mailed an activPAL. If participants consented to provide body measurements, they are also mailed a set of weighing scales (Seca 803 if they weigh

XSL•FO
RenderX

less than 150 kg; Seca 813 if they weigh over 150 kg) and a tape measure (Seca 201) at baseline. Reminder calls are made if any questionnaires are not completed, and telephone support on completing the questionnaires is provided where necessary. To be included in analysis, 3-month data must be completed between 12-18 weeks postrandomization and 6-month data between 24-30 weeks postrandomization. Researchers involved in collecting follow-up assessment data are blinded to group allocation.

Sociodemographic Characteristics and Disease Characteristics

At baseline, age, sex, employment and education status, marital status, living arrangements, postcode, and ethnicity are self-reported. Participant's postcodes will be used to determine socioeconomic position using the Index of Multiple Deprivation [73]. Participants are asked to self-report their cancer diagnosis and comorbid health conditions [74] at baseline, and clinical details will be verified from medical records (including diagnosis date and stage and treatment received).

Physical Activity

PA is measured objectively [75] using thigh-worn activPAL accelerometer (sent with detailed instructions and materials to apply the device to their thigh, plus a link to a video of a researcher demonstrating how to attach the device). Participants are asked to wear the activPAL for 7 days continuously at each timepoint (ie, baseline, 3 months, and 6 months). ActivPAL data will be used to calculate the primary endpoint. Participants will also complete a log sheet recording when the device was worn and their bedtimes and waketimes. ActivPAL devices have been used successfully in other studies, including those diagnosed with cancer [76], and were used in the pilot [44,53]. The primary endpoint is activPAL-assessed average minutes of brisk walking (≥100 spm) 3 months after randomization (controlling for baseline rates). ActivPAL data will also be used to assess brisk walking at 6 months, average total steps, minutes of light PA, standing time, sitting time, and weekly MET. Participants will also complete the Godin Leisure Time Exercise Questionnaire (GLTEQ) [77], which has demonstrated favorable validity and reliability against objective measures of PA and is widely used in oncology research [78]. This questionnaire was adapted in the pilot [44,53] to include a question about duration of activities to allow the calculation of minutes of MVPA [78].

Anthropometric Outcomes (Optional)

Height, weight (without outer clothing or shoes on), and waist circumference (at the umbilicus) are self-reported by participants into an online questionnaire. Previous studies have illustrated that self-reported weight and waist circumference are adequately reliable and accurate where objective measurement is not possible [79-81]. BMI will be calculated using the standard formula of weight (kg) ÷ height $(m)^2$. Participants who provide these measures at baseline will keep the weighing scales and tape measure for the later follow-ups.

Psychosocial Variables

To assess the broader impact of the intervention, the psychosocial variables outlined in Table 4 are self-reported at each timepoint.

Table 4. Details of the psychosocial variables and measures included in the questionnaires.

Construct	Measure	Reliability and validity references
Health related quality of life (HRQoL)	EQ-5D-5L and visual analogue scale [82]	Reliable and valid among patients with cancer [83]
Cancer-specific quality of life	27-item Functional Assessment of Cancer Therapy-General (FACT-G) scale [84]	Good validity and reliability [84]
Fatigue	13-item fatigue subscale of the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) [85]	A brief reliable and valid measure [85]
Sleep quality	18-item Pittsburgh Sleep Quality Index (PSQI), which assesses sleep quality and disturbances over a 1-month time interval [86]	Used extensively in clinical (including cancer) and nonclinical samples, with good reliability and validity [87]
Anxiety	7-item Generalized Anxiety Disorder-7 (GAD-7) Assessment [88]	Adequate diagnostic accuracy in patients with cancer [89]
Depression	9-item Patient Health Questionnaire (PHQ-9) [90]	Adequate diagnostic accuracy in patients with cancer [91]
Physical activity self-efficacy	13-item Physical Activity Appraisal Inventory (PAAI)	Reliability and validity in women with breast cancer [92]
Self-efficacy to self-manage cancer	11-item Cancer Survivorship Self-Efficacy Scale (CS-SES) [93]	Excellent reliability across mixed patients with cancer [93]
Habit strength for walking	4-item Self-Report Behavioral Automaticity Index (SRBAI) [94], which will be measured focusing on "going for a walk" and "walking briskly"	Good reliability and sensitivity to changes in habit [94]
Social support	6-item brief form of the Perceived Social Support Questionnaire (F-Soz-U K-6) [95]	Good reliability and validity across the lifespan [95]

XSL•FO
RenderX

To conduct exploratory analysis between the activPAL data and habit strength, investigating whether those who walk in the morning have higher habit strength than those who walk in the evening, participants will be asked 4 questions related to their chronotype (individual differences in sleep timing and in preferences for sleeping at given times of day). This will be measured using a subscale of the Morningness-Eveningness Questionnaire [96] used by Randler et al [97], with the English wording taken from Terman et al [98]. Additionally, 2 items (scored on a 5-point Likert scale) from the Creature of Habit scale [99] will be used to assess preference for routine ("I tend to like routine" and "I find comfort in regularity").

Health and Social Care Service Use

Participants are asked to self-report their use of health and social care using items from the Client Service Receipt Inventory (CSRI), which has been validated against objective primary care records [100] and is recommended for usage of hospital and other community health services [101]. As part of the medical records data collection (undertaken by research staff at 6 months postrandomization), details about participant's hospital-based treatment and care received (including the number of hospital contacts) during the 30-week trial period will be recorded.

App Use and Physical Activity Changes

No data collected by the app itself will be obtained by the research team [53] owing to the app data not being stored in a way that would allow linkage with our trial participants. Therefore, participants in both groups are asked to report their use of any PA apps during the study period, whether they have done anything to change their PA since starting the trial, and what prompted them. Additionally, intervention group participants complete brief feedback questions on the intervention content in the follow-up questionnaires, self-report their usage of the app, and complete the Digital Behavior Change Interventions Engagement Scale [102], which was used in the pilot study to capture app engagement [53]. A single item ("How accurately do you think the app recorded the amount of time you spent walking [the total minutes and brisk minutes]?") about perceived accuracy of the app is included at 6 months (5-point scale, 1=not accurate and 5=very accurate).

Qualitative End of Study Interviews

A subsample of intervention participants (approximately n=20-30) will be interviewed after they have completed the 6-month data collection. These telephone or video call interviews will be semistructured (Multimedia Appendix 4 provides the interview guide), audio recorded, and will explore experiences of using the intervention materials, their walking experiences, and their use of the app.

Data Analysis

The primary endpoint, the difference between trial arms in change from baseline in activPAL-assessed average minutes of brisk walking (≥100 spm) 3 months after randomization, will be analyzed by a statistician blind to intervention allocation on an intention-to-treat basis. Estimates of effect sizes will be presented with 95% CIs and 2-sided 5% *P* values. The statistical analysis plan will be finalized prior to completion of data collection. Cross-sectional analysis will be performed at 3 months (primary endpoint) and 6 months (secondary endpoint) using linear regression to compare the two trial arms, with baseline time spent walking at ≥100 spm included as a covariate. Generalized linear mixed models will also be used to model repeated measurements over time, with individual patient random effects, and intervention, time, and intervention-by-time interaction fixed effects. Other continuous secondary outcomes (ie, average daily total steps, minutes of light PA, standing time and sitting time, GLTEQ, EQ-5D-5L, Functional Assessment of Cancer Therapy-General [FACT-G], Functional Assessment of Chronic Illness Therapy-Fatigue [FACIT-F], Pittsburgh Sleep Quality Index [PSQI], Generalized Anxiety Disorder-7 [GAD-7], Patient Health Questionnaire-9 [PHQ-9], Physical Activity Appraisal Inventory [PAAI], Self-Report Behavioral Automaticity Index [SRBAI], BMI, and waist circumference) will be analyzed using a similar approach. Relationships with baseline participant and disease characteristics will be investigated through covariates for main effects as well as interactions with the trial intervention. Subgroup analyses will be performed by stratification factors used in the randomization procedure and presented using forest plots and *P* values from interaction tests. Exploratory analyses to understand the role of habit strength as a mechanism in the relationship between the trial intervention and PA will be carried out through further regression modeling.

Cost-Effectiveness Analysis

Cost-effectiveness will be estimated using a health economic model developed for preliminary analysis of the APPROACH pilot feasibility trial [53]. The model is a cohort Markov model with annual cycles and a lifetime horizon, which links PA to mortality from cancer and other causes. The model will be updated using data from the study, reducing key uncertainties identified as part of the preliminary analysis. The model population will be updated using baseline characteristics from the study. PA changes due to the intervention incorporated in the model will be based on weekly METs rather than on the primary trial outcome, as this measurement aligns better with other data already included in the model defining the relationship between PA and mortality [17,103,104]. Additions to the model will be made to incorporate NHS resource use and quality-of-life data gathered as part of the trial, as these were not included in the previous modeling but are likely to have large impacts on cost-effectiveness. Intervention costs will also be updated. The model will estimate costs and quality-adjusted life years using probabilistic sensitivity analysis to enable uncertainty in the results to be captured. Cost-effectiveness will be estimated across the whole population and in relevant subgroups.

Qualitative Analysis

Qualitative interview data will be analyzed using a framework analysis approach [105]. This approach was chosen because these interviews will inform the process evaluation of the intervention and data will fulfil predetermined information needs about the implementation of the intervention, mechanisms of action, and contextual factors impacting engagement [106]. We also anticipate mapping BCTs against the Behavior Change Taxonomy Ontology [62,63]. The Framework Method involves

XSL•FO
RenderX

both deductive and inductive approaches to analyzing the data, where predefined codes are used alongside open coding [107]. There are 7 stages [107]: transcription, familiarization, coding, developing a working analytical framework, applying the analytical framework, charting the data into the framework matrix, and interpreting the data. This approach enables a systematic procedure for reducing the data while remaining flexible enough to identify unexpected themes and perspectives. Indexing of manuscripts will be undertaken by researchers not involved in the delivery of the intervention to minimize reflexivity biases. A portion of transcripts will be independently double-coded (with any ambiguities or differences discussed and resolved). Final reporting of the qualitative interview data collection and analysis process will be reported according to the COREQ (Consolidated Criteria for Reporting Qualitative Research) Checklist [108].

Data Monitoring

UCL is the trial sponsor. The trial will be overseen by a Trial Management Group consisting of the chief investigators (AF and PL), co-investigators, and research staff employed on the grant. An external Trial Steering Committee, including 2 independent members, all co-investigators, and lay representatives, will be updated regularly on trial progress and will meet as required to provide input. No formal data monitoring committee will be set up, and there are no criteria for stopping the trial as walking is a very low-risk intervention.

Adverse Event Reporting

All adverse events that the research team become aware of from the point of consent will be assessed by severity (mild, moderate, or severe), causality (not related, unlikely, possibly, probably, definitely, or not assessable), and expectedness (expected or unexpected). In this population, various expected adverse events relating to participant's cancer diagnosis and treatment may occur (eg, acute illness, symptoms related to treatment), but as these would not be related to the intervention, they will not be recorded. All other adverse events will be recorded, including "mild skin irritation from the adhesive dressing when wearing the activPAL device" as this has been observed previously in older adults [109] and in the pilot [53].

Ethical Considerations

Favorable ethical opinion for this study was received from the London-Surrey Research Ethics Committee in October 2023 (reference 23/LO/0740). This study describes protocol version 9.0 (dated January 16, 2025), which was approved and implemented in February 2025. All participants will provide written consent.

The trial has been approved by the London-Surrey Research Ethics Committee (reference 23/LO/0740), the Health Research Authority, and the Research and Development department at each recruiting site. Amendments to the protocol will only be implemented after appropriate approval from relevant bodies and communicated where necessary with relevant parties. The study is compliant with the requirements of the General Data Protection Regulation (2016/679) and the UK Data Protection Act (2018). UCL is the data controller, and the study has been registered for Data Protection at UCL (reference Z6364106/2022/07/163 social research). Personal data will only be collected where it is deemed essential to the study. Where possible, personal data are pseudonymized. No identifiable personal data will be used in dissemination of the research. All participants provide informed written consent. Verbally and in the information sheet participants are informed that they are under no obligation to take part and can withdraw at any time without giving a reason and without their medical care being affected. Should any participants ask to withdraw, the data collected up to the point of withdrawal will be analyzed. These participants will also be asked if they allow us to continue to access data from their hospital medical records (collected at 6 months) and linkage with NDRS and HES (if appropriate).

Most study data will be entered directly into, and stored within, the UCL Data Safe Haven secure encrypted platform, but, if on paper, will be stored in locked filing cabinets at UCL. The final dataset will be pseudonymized. All identifiable personal data and research data will be stored for 12 years (allowing time for potential linkage with NHS Digital data 10 years after the trial end date should the research team obtain funding to undertake this) from the trial end date, after which it will be anonymized (removing the study pseudonym and deleting all contact details, including proof of consent). This anonymized dataset (excluding NDRS and HES data) will be entered into the Open Science Framework. The Trial Master File will be archived at UCL, in accordance with the UCL Retentions Schedule, for a minimum of 5 years from the trial end and no longer than 20 years.

Results

The study was funded in June 2019. Trial recruitment began in November 2023. As of December 2025, a total of 473 participants have been randomized. It is anticipated that the recruitment target will be reached during 2025, and the final participant will complete the study in 2026. The publication of the main results is expected in autumn 2027 after all follow-up data collection and analysis are complete.

Discussion

Overview

This study describes the protocol for APPROACH, a multicenter, phase III, randomized control trial, which follows on from the pilot [44,53]. The results will provide robust evidence relating to the clinical and cost-effectiveness of the APPROACH brisk walking intervention among a mixed cohort of individuals diagnosed with breast, colorectal, or prostate cancer, including whether the intervention increases objective, activPal-assessed PA and other outcomes. This study is set within the context of the increasing recognition of the importance of PA in both prehabilitation and rehabilitation settings [6] and growing emphasis on PA advice to be routinely given to individuals diagnosed with cancer [29]. With the introduction of support programs and services being offered to patients locally and nationally [110], the clinical and cost-effectiveness of different types of PA interventions need to be established to support their implementation into routine clinical care.

XSL•FO
RenderX

In addition to the previously outlined small changes to the intervention materials (eg, reformatting the walking planning cards) based on feedback from the pilot [53,111], the following changes were made and incorporated into the original protocol registered in September 2023 (registration number ISRCTN14149329). First, the decision was taken to focus on localized disease. Although systematic reviews have illustrated the safety and potential benefits of PA among metastatic patients [112,113], the limited number of patients recruited to the pilot [53] with metastatic disease (7/90, 7.8%; mainly owing to challenges of obtaining clinical sign-off) meant the feasibility and acceptability of the intervention among this patient group could not be assessed. Second, due to the need to adhere to the research ethics committee requirements, an initial consent-to-contact stage [114] was added. Similarly, to ease the logistics of medical record screening for hospital-based staff, the criteria of time since diagnosis (<2 years) was simplified from the previous pilot criteria (<6 months post end of active treatment). Finally, to give additional flexibility for recruiting hospital sites, a supplementary recruitment route was introduced that allowed the use of leaflets and advertisements (which were reviewed and edits made after following our PPI members' suggestions) to provide patients with the opportunity to self-refer their interest.

Strengths

Compared to some more time- and resource-intensive programs, the theory-informed design of APPROACH provides a potentially low-cost intervention, and it is hoped it could be integrated into usual care or existing support services if shown to be clinical and cost-effective. Participants require no specialist equipment (apart from a mobile phone, ownership of which Office for National Statistics figures illustrate is high and increasing in older adults [115]) and are not required to attend any in-person appointments. Therefore, the intervention has the potential to be flexibly accommodated into daily life (eg, around work/life commitments) and does not discriminate against groups who are not able or do not wish to attend time-intensive time-specific PA sessions. This study also objectively measures PA using thigh-worn accelerometers [75]. The researchers involved in organizing study assessments and the statistician undertaking the primary endpoint data analysis will be blinded to group allocation.

Limitations

This study has several limitations that should be noted. A known limitation of PA research is how study participation (including objective measurement using accelerometers) may influence control group participants to be more active [116]. However, a systematic review by Freene et al [117] suggests these are not significant improvements. Apart from the accelerometer measurement of PA, many of the other outcomes will be self-reported (including use of the app) and therefore have the potential to introduce bias. App engagement will also be explored in the qualitative end-of-study interviews, which, although self-reported, may provide further context on the frequency of use, users' sense of engagement with the app, and reasons for continued or discontinued use. Furthermore, the eligibility screening question relating to current levels of MVPA (ie, those already meeting 150 minutes of MVPA per week are ineligible) is self-reported, which again may introduce bias [118,119] and result in some active participants being recruited [120]. Research by Smith et al [121] indicates that patients diagnosed with prostate cancer tend to overestimate their PA levels, suggesting we may be conservative in our inclusion. Furthermore, the alternative of obtaining objective measures of PA during the screening phase was not practical for this study.

We acknowledge that the nature of this app-based intervention may not appeal to some patients who do not own a smartphone or consider themselves technologically literate. Office for National Statistics figures illustrate smartphone ownership is increasing in older adults [115], and the APPROACH team provide telephone support to all to increase confidence where needed. Future studies could also fund and provide devices for participants to reduce any digital exclusion or partner with charities set up to support digital poverty when delivering the intervention. As the app being explored in this study is publicly available, there is a possibility that control group participants will find and use it independently. This was not seen in the pilot study [44,53] among 90 participants and is being monitored in the trial follow-up. Furthermore, all participants are asked about their use of apps at baseline (to screen out those already using the study app) and at the follow-ups. Furthermore, the intervention itself is a package of support, and use of the app in isolation (without the broader behavioral support) is not considered equivalent to being in the intervention group. Finally, using a publicly available app meant we were unable to obtain and interrogate app diagnostic data for the trial participants as this option was not available at the time of setting up the trial.

Dissemination

Study participants will receive a summary of the study findings after data analysis has taken place. Nonidentifiable study results will be disseminated widely in peer-reviewed publications, conference presentations, social media outlets (eg, LinkedIn), and through our PPI representatives. Direct quotations from interview data will be published anonymously.

Conclusion

To conclude, if found to have a positive effect on PA at 3 months, APPROACH has the potential to provide health professionals with a cost-effective package of support that can be offered to patients diagnosed with cancer. If implemented, this could contribute to improved physical and psychosocial long-term outcomes for those living with and beyond cancer.

Acknowledgments

The authors would like to thank the National Cancer Research Institute Consumer Forum and their members who provided feedback on the intervention materials as they were initially developed, and our patient and public involvement members (Mrs

XSL•FO
RenderX

Jacqui Gath and Mrs Lindy Berkman) who provided further feedback following the pilot study. We would also like to thank all the participants who took part in the pilot study.

The trial sponsor is University College London and can be contacted at uclh.randd@nhs.net. The sponsor has, and will have, no involvement in writing this manuscript, the design of the study, or collection, analysis, and interpretation of data. They will have no role in deciding whether to submit results for publication. We would like to thank the ongoing contributions of our trial steering committee.

Funding

This project is funded by Yorkshire Cancer Research (reference RA/2018/R1/005; joint chief investigators: PL and AF). The funding body has, and will have, no involvement in writing this manuscript, the design of the study, or collection, analysis, and interpretation of data. They will have no role in deciding whether to submit results for publication.

Data Availability

Data sharing is not applicable to this study, as no datasets were generated or analyzed during this study. The full intervention materials are not supplied with the protocol paper since recruitment for the study is ongoing, and we do not want the materials to be publicly available until the end of the trial.

Authors' Contributions

FK, CB, SLS, L Shearsmith, HT, AR, RJB, DMG, HWWP, EL, NC, NL, CT, L Smith, LW, JG, PL, and AF conceptualized the study. FK, SLS, PL, CT, NC, and AF drafted the manuscript. All authors provided critical intellectual input and edits to the study protocol and manuscript. The authors read and approved the final manuscript.

Conflicts of Interest

AF and HWWP have a PhD student employed and funded by BetterPoints Ltd, which has another app to promote physical activity. NL is funded by Yorkshire Cancer Research (award reference number S406NL). The other authors declare they have no conflicts of interest.

Multimedia Appendix 1
SPIRIT 2013 checklist.
[DOCX File , 29 KB - resprot_v15i1e77096_app1.docx]

Multimedia Appendix 2
Online consent form.
[DOCX File , 27 KB - resprot_v15i1e77096_app2.docx]

Multimedia Appendix 3
TIDiER checklist - APPROACH randomized controlled trial.
[DOCX File , 24 KB - resprot_v15i1e77096_app3.docx]

Multimedia Appendix 4
APPROACH randomized controlled trial interview guide.
[DOCX File , 79 KB - resprot_v15i1e77096_app4.docx]

Multimedia Appendix 5
Peer review report from the Yorkshire Cancer Research Advisory Committee.
[PDF File (Adobe PDF File), 639 KB - resprot_v15i1e77096_app5.pdf]

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: Globocan 2000. Int J Cancer 2001;94(2):153-156. [doi: 10.1002/ijc.1440] [Medline: 11668491]
2. Global cancer burden growing, amidst mounting need for services. International Agency for Research on Cancer. 2024. URL: https://www.iarc.who.int/news-events/global-cancer-burden-growing-amidst-mounting-need-for-services/ [accessed 2024-02-14]
3. Miller KD, Nogueira L, Devasia T, Mariotto AB, Yabroff KR, Jemal A, et al. Cancer treatment and survivorship statistics, 2022. CA Cancer J Clin 2022;72(5):409-436 [FREE Full text] [doi: 10.3322/caac.21731] [Medline: 35736631]

XSL•FO
RenderX

4. Cancer survival statistics. Cancer Research UK. 2023. URL: https://www.cancerresearchuk.org/health-professional/cancer-statistics/survival [accessed 2024-12-02]
5. The NHS Long Term Plan. NHS England. 2019. URL: https://www.longtermplan.nhs.uk/ [accessed 2024-10-02]
6. Christensen J, Simonsen C, Hojman P. Exercise training in cancer control and treatment. Compr Physiol 2018;9(1):165-205. [doi: 10.1002/cphy.c180016] [Medline: 30549018]
7. Singh B, Hayes SC, Spence RR, Steele ML, Millet GY, Gergele L. Exercise and colorectal cancer: a systematic review and meta-analysis of exercise safety, feasibility and effectiveness. Int J Behav Nutr Phys Act 2020;17(1):122 [FREE Full text] [doi: 10.1186/s12966-020-01021-7] [Medline: 32972439]
8. Belloni S, Arrigoni C, Caruso R. Effects from physical exercise on reduced cancer-related fatigue: a systematic review of systematic reviews and meta-analysis. Acta Oncol 2021;60(12):1678-1687. [doi: 10.1080/0284186X.2021.1962543] [Medline: 34396915]
9. Mishra S, Scherer RW, Geigle PM, Berlanstein DR, Topaloglu O, Gotay CC, et al. Exercise interventions on health-related quality of life for cancer survivors. Cochrane Database Syst Rev 2012;2012(8):CD007566 [FREE Full text] [doi: 10.1002/14651858.CD007566.pub2] [Medline: 22895961]
10. Mishra S, Scherer RW, Snyder C, Geigle PM, Berlanstein DR, Topaloglu O. Exercise interventions on health-related quality of life for people with cancer during active treatment. Cochrane Database Syst Rev 2012;2012(8):CD008465 [FREE Full text] [doi: 10.1002/14651858.CD008465.pub2] [Medline: 22895974]
11. Aune D, Markozannes G, Abar L, Balducci K, Cariolou M, Nanu N, et al. Physical activity and health-related quality of life in women with breast cancer: a meta-analysis. JNCI Cancer Spectr 2022;6(6):pkac072 [FREE Full text] [doi: 10.1093/jncics/pkac072] [Medline: 36474321]
12. Ramírez-Vélez R, Zambom-Ferraresi F, García-Hermoso A, Kievisiene J, Rauckiene-Michealsson A, Agostinis-Sobrinho C. Evidence-based exercise recommendations to improve mental wellbeing in women with breast cancer during active treatment: a systematic review and meta-analysis. Cancers (Basel) 2021;13(2):264 [FREE Full text] [doi: 10.3390/cancers13020264] [Medline: 33445739]
13. Dennett AM, Peiris CL, Shields N, Prendergast LA, Taylor NF. Moderate-intensity exercise reduces fatigue and improves mobility in cancer survivors: a systematic review and meta-regression. J Physiother 2016;62(2):68-82 [FREE Full text] [doi: 10.1016/j.jphys.2016.02.012] [Medline: 26996098]
14. Lahart IM, Metsios GS, Nevill AM, Carmichael AR. Physical activity, risk of death and recurrence in breast cancer survivors: a systematic review and meta-analysis of epidemiological studies. Acta Oncol 2015;54(5):635-654. [doi: 10.3109/0284186X.2014.998275] [Medline: 25752971]
15. Richman E, Kenfield SA, Stampfer MJ, Paciorek A, Carroll PR, Chan JM. Physical activity after diagnosis and risk of prostate cancer progression: data from the cancer of the prostate strategic urologic research endeavor. Cancer Res 2011;71(11):3889-3895 [FREE Full text] [doi: 10.1158/0008-5472.CAN-10-3932] [Medline: 21610110]
16. Meyerhardt JA, Heseltine D, Niedzwiecki D, Hollis D, Saltz LB, Mayer RJ, et al. Impact of physical activity on cancer recurrence and survival in patients with stage III colon cancer: findings from CALGB 89803. J Clin Oncol 2006;24(22):3535-3541. [doi: 10.1200/JCO.2006.06.0863] [Medline: 16822843]
17. Schmid D, Leitzmann M. Association between physical activity and mortality among breast cancer and colorectal cancer survivors: a systematic review and meta-analysis. Ann Oncol 2014;25(7):1293-1311 [FREE Full text] [doi: 10.1093/annonc/mdu012] [Medline: 24644304]
18. Friedenreich C, Stone CR, Cheung WY, Hayes SC. Physical activity and mortality in cancer survivors: a systematic review and meta-analysis. JNCI Cancer Spectr 2020;4(1):pkz080 [FREE Full text] [doi: 10.1093/jncics/pkz080] [Medline: 32337494]
19. Morishita S, Hamaue Y, Fukushima T, Tanaka T, Fu JB, Nakano J. Effect of exercise on mortality and recurrence in patients with cancer: a systematic review and meta-analysis. Integr Cancer Ther 2020;19:1534735420917462 [FREE Full text] [doi: 10.1177/1534735420917462] [Medline: 32476493]
20. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71(3):209-249 [FREE Full text] [doi: 10.3322/caac.21660] [Medline: 33538338]
21. Campbell K, Winters-Stone KM, Wiskemann J, May AM, Schwartz AL, Courneya KS, et al. Exercise guidelines for cancer survivors: consensus statement from international multidisciplinary roundtable. Med Sci Sports Exerc 2019;51(11):2375-2390 [FREE Full text] [doi: 10.1249/MSS.0000000000002116] [Medline: 31626055]
22. Schmitz KH, Campbell AM, Stuiver MM, Pinto BM, Schwartz AL, Morris GS, et al. Exercise is medicine in oncology: engaging clinicians to help patients move through cancer. CA Cancer J Clin 2019;69(6):468-484 [FREE Full text] [doi: 10.3322/caac.21579] [Medline: 31617590]
23. Patel A, Friedenreich CM, Moore SC, Hayes SC, Silver JK, Campbell KL, et al. American college of sports medicine roundtable report on physical activity, sedentary behavior, and cancer prevention and control. Med Sci Sports Exerc 2019;51(11):2391-2402 [FREE Full text] [doi: 10.1249/MSS.0000000000002117] [Medline: 31626056]
24. Diet, nutrition, physical activity and cancer: a global perspective. World Cancer Research Fund and American Institute for Cancer Research. 2018. URL: https://www.wcrf.org/wp-content/uploads/2024/11/Summary-of-Third-Expert-Report-2018.pdf [accessed 2024-06-13]

XSL•FO
RenderX

25. Cannioto R. "Move more, sit less" is a feasible and impactful guideline for improving cancer survival. JNCI Cancer Spectr 2023;7(2):pkad020 [FREE Full text] [doi: 10.1093/jncics/pkad020] [Medline: 36869677]
26. UK chief medical officers' physical activity guidelines. UK Chief Medical Officers. 2019. URL: https://assets.publishing.service.gov.uk/media/5d839543ed915d52428dc134/uk-chief-medical-officers-physical-activity-guidelines.pdf [accessed 2024-09-01]
27. Garcia DO, Thomson CA. Physical activity and cancer survivorship. Nutr Clin Pract 2014;29(6):768-779 [FREE Full text] [doi: 10.1177/0884533614551969] [Medline: 25335787]
28. Kennedy F, Lally P, Miller NE, Conway RE, Roberts A, Croker H, et al. Fatigue, quality of life and associations with adherence to the world cancer research fund guidelines for health behaviours in 5835 adults living with and beyond breast, prostate and colorectal cancer in England: A cross-sectional study. Cancer Med 2023;12(11):12705-12716 [FREE Full text] [doi: 10.1002/cam4.5899] [Medline: 37021752]
29. Achieving world-class cancer outcomes: a strategy for England 2015-2020. Independent Cancer Taskforce. 2015. URL: https://www.england.nhs.uk/publication/achieving-world-class-cancer-outcomes-a-strategy-for-england-2015-2020/ [accessed 2023-12-02]
30. Principles and guidance for prehabilitation within the management and support of people with cancer. MacMillan Cancer Support. 2019. URL: https://cdn.macmillan.org.uk/dfsmedia/1a6f23537f7f4519bb0cf14c45b2a629/1532-source/prehabilitation-for-people-with-cancer-tcm9-353994 [accessed 2023-12-02]
31. Gokal K, Daley AJ, Madigan CD. "Fear of raising the problem without a solution": a qualitative study of patients' and healthcare professionals' views regarding the integration of routine support for physical activity within breast cancer care. Support Care Cancer 2024;32(1):87 [FREE Full text] [doi: 10.1007/s00520-023-08293-2] [Medline: 38185712]
32. Dennis D, Waring J, Payeur N, Cosby C, Daudt H. Making lifestyle changes after colorectal cancer: insights for program development. Curr Oncol 2013;20(6):e493-e511 [FREE Full text] [doi: 10.3747/co.20.1514] [Medline: 24311950]
33. Fisher A, Wardle J, Beeken RJ, Croker H, Williams K, Grimmett C. Perceived barriers and benefits to physical activity in colorectal cancer patients. Support Care Cancer 2016;24(2):903-910 [FREE Full text] [doi: 10.1007/s00520-015-2860-0] [Medline: 26268781]
34. Fisher A, Williams K, Beeken R, Wardle J. Recall of physical activity advice was associated with higher levels of physical activity in colorectal cancer patients. BMJ Open 2015;5(4):e006853 [FREE Full text] [doi: 10.1136/bmjopen-2014-006853] [Medline: 25922098]
35. Koutoukidis DA, Lopes S, Fisher A, Williams K, Croker H, Beeken RJ. Lifestyle advice to cancer survivors: a qualitative study on the perspectives of health professionals. BMJ Open 2018;8(3):e020313 [FREE Full text] [doi: 10.1136/bmjopen-2017-020313] [Medline: 29593021]
36. Orange ST, Gilbert SE, Brown MC, Saxton JM. Recall, perceptions and determinants of receiving physical activity advice amongst cancer survivors: a mixed-methods survey. Support Care Cancer 2021;29(11):6369-6378 [FREE Full text] [doi: 10.1007/s00520-021-06221-w] [Medline: 33885962]
37. Robinson R, Crank H, Humphreys H, Fisher P, Greenfield DM. Allied health professional's self-reported competences and confidence to deliver physical activity advice to cancer patients at a single centre in England. Disabil Rehabil 2023;45(24):4058-4064. [doi: 10.1080/09638288.2022.2143580] [Medline: 36573404]
38. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HWW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017;11(6):704-719 [FREE Full text] [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
39. Villalobos-Zúñiga G, Cherubini M. Apps that motivate: a taxonomy of app features based on self-determination theory. Int J Hum-Comput Stud 2020;140:102449. [doi: 10.1016/j.ijhcs.2020.102449]
40. Khoo S, Mohbin N, Ansari P, Al-Kitani M, Müller AM. mHealth interventions to address physical activity and sedentary behavior in cancer survivors: a systematic review. Int J Environ Res Public Health 2021;18(11):5798 [FREE Full text] [doi: 10.3390/ijerph18115798] [Medline: 34071342]
41. Ester M, Eisele M, Wurz A, McDonough MH, McNeely M, Culos-Reed SN. Current evidence and directions for future research in ehealth physical activity interventions for adults affected by cancer: systematic review. JMIR Cancer 2021;7(3):e28852 [FREE Full text] [doi: 10.2196/28852] [Medline: 34542415]
42. Gardner B, Lally P. Habit and habitual behaviour. Health Psychol Rev 2023;17(3):490-496. [doi: 10.1080/17437199.2022.2105249] [Medline: 35880387]
43. Gardner B, Rebar AL, de Wit S, Lally P. What is habit and how can it be used to change real-world behaviour? Narrowing the theory-reality gap. Soc Pers Psychol Compass 2024;18(6):e12975. [doi: 10.1111/spc3.12975]
44. Lally P, Miller N, Roberts A, Beeken RJ, Greenfield DM, Potts HWW, et al. An app with brief behavioural support to promote physical activity after a cancer diagnosis (APPROACH): study protocol for a pilot randomised controlled trial. Pilot Feasibility Stud 2022;8(1):74 [FREE Full text] [doi: 10.1186/s40814-022-01028-w] [Medline: 35351187]
45. Gardner B, Rebar A, Lally P. Habit interventions. In: Hagger M, editor. The Handbook of Behavior Change. Cambridge: Cambridge University Press; 2020:599-616.

XSL•FO
RenderX

46. Roberts AL, Potts HW, Koutoukidis DA, Smith L, Fisher A. Breast, prostate, and colorectal cancer survivors' experiences of using publicly available physical activity mobile apps: qualitative study. JMIR Mhealth Uhealth 2019;7(1):e10918 [FREE Full text] [doi: 10.2196/10918] [Medline: 30609982]
47. Elshahat S, Treanor C, Donnelly M. Factors influencing physical activity participation among people living with or beyond cancer: a systematic scoping review. Int J Behav Nutr Phys Act 2021;18(1):50. [doi: 10.1186/s12966-021-01116-9] [Medline: 33823832]
48. Brannan M, Foster CE, Timpson CM, Clarke N, Sunyer E, Amlani A, et al. Active 10 - A new approach to increase physical activity in inactive people in England. Prog Cardiovasc Dis 2019;62(2):135-139. [doi: 10.1016/j.pcad.2019.02.001] [Medline: 30796943]
49. Howlett N, Trivedi D, Troop NA, Chater AM. Are physical activity interventions for healthy inactive adults effective in promoting behavior change and maintenance, and which behavior change techniques are effective? A systematic review and meta-analysis. Transl Behav Med 2019;9(1):147-157 [FREE Full text] [doi: 10.1093/tbm/iby010] [Medline: 29506209]
50. Page EJ, Massey AS, Prado-Romero PN, Albadawi S. The use of self-monitoring and technology to increase physical activity: a review of the literature. Perspect Behav Sci 2020;43(3):501-514 [FREE Full text] [doi: 10.1007/s40614-020-00260-0] [Medline: 33029577]
51. Gildea GC, Spence RR, Jones TL, Turner JC, Macdonald ER, Hayes SC, et al. Barriers, facilitators, perceptions and preferences influencing physical activity participation, and the similarities and differences between cancer types and treatment stages – a systematic rapid review. Prev Med Rep 2023;34:102255 [FREE Full text] [doi: 10.1016/j.pmedr.2023.102255] [Medline: 37273528]
52. Roberts AL, Potts HWW, Stevens C, Lally P, Smith L, Fisher A. Cancer specialist nurses' perspectives of physical activity promotion and the potential role of physical activity apps in cancer care. J Cancer Surviv 2019;13(5):815-828 [FREE Full text] [doi: 10.1007/s11764-019-00801-w] [Medline: 31475306]
53. Lally P, Kennedy F, Smith S, Beeken RJ, Buck C, Thomas C, et al. The feasibility and acceptability of an app-based intervention with brief behavioural support (APPROACH) to promote brisk walking in people diagnosed with breast, prostate and colorectal cancer in the UK. Cancer Med 2024;13(6):e7124 [FREE Full text] [doi: 10.1002/cam4.7124] [Medline: 38529687]
54. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18 [FREE Full text] [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]
55. ISRCTN. 2023. URL: https://doi.org/10.1186/ISRCTN14149329 [accessed 2023-10-31]
56. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013;346:e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
57. Harris P, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
58. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
59. O'Callaghan CA. OxMaR: open source free software for online minimization and randomization for clinical trials. PLoS One 2014;9(10):e110761 [FREE Full text] [doi: 10.1371/journal.pone.0110761] [Medline: 25353169]
60. Lowther E. Minimization. 2024. URL: https://github.com/edlowther/minimization [accessed 2025-03-07]
61. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
62. BCTO Search Website Interface. 2024. URL: https://www.bciosearch.org/ [accessed 2024-01-09]
63. Marques MM, Wright AJ, Corker E, Johnston M, West R, Hastings J, et al. The behaviour change technique ontology: transforming the behaviour change technique taxonomy v1. Wellcome Open Res 2023;8:308. [doi: 10.12688/wellcomeopenres.19363.1]
64. Hailey V, Rojas-Garcia A, Kassianos AP. A systematic review of behaviour change techniques used in interventions to increase physical activity among breast cancer survivors. Breast Cancer 2022;29(2):193-208 [FREE Full text] [doi: 10.1007/s12282-021-01323-z] [Medline: 34989962]
65. Salisbury CE, Hyde MK, Cooper ET, Stennett RC, Gomersall SR, Skinner TL. Physical activity behaviour change in people living with and beyond cancer following an exercise intervention: a systematic review. J Cancer Surviv 2023;17(3):569-594 [FREE Full text] [doi: 10.1007/s11764-023-01377-2] [Medline: 37074621]
66. Finne E, Glausch M, Exner A, Sauzet O, Stölzel F, Seidel N. Behavior change techniques for increasing physical activity in cancer survivors: a systematic review and meta-analysis of randomized controlled trials. Cancer Manag Res 2018;10:5125-5143 [FREE Full text] [doi: 10.2147/CMAR.S170064] [Medline: 30464612]

XSL•FO
RenderX

67. Roberts D, Townsend N, Foster C. Use of new guidance to profile 'equivalent minutes' of aerobic physical activity for adults in England reveals gender, geographical, and socio-economic inequalities in meeting public health guidance: a cross-sectional study. Prev Med Rep 2016;4:50-60 [FREE Full text] [doi: 10.1016/j.pmedr.2016.05.009] [Medline: 27413661]
68. Turner RR, Steed L, Quirk H, Greasley RU, Saxton JM, Taylor SJ, et al. Interventions for promoting habitual exercise in people living with and beyond cancer. Cochrane Database Syst Rev 2018;9(9):CD010192. [doi: 10.1002/14651858.CD010192.pub3] [Medline: 30229557]
69. Turner R, Steed L, Quirk H, Greasley RU, Saxton JM, Taylor SJ, et al. Interventions for promoting habitual exercise in people living with and beyond cancer. Cochrane Database Syst Rev 2018;9(9):CD010192 [FREE Full text] [doi: 10.1002/14651858.CD010192.pub3] [Medline: 30229557]
70. UK chief medical officers' physical activity guidelines. Department of Health & Social Care. 2019. URL: https://assets.publishing.service.gov.uk/media/5d839543ed915d52428dc134/uk-chief-medical-officers-physical-activity-guidelines.pdf [accessed 2025-12-02]
71. Füzéki E, Engeroff T, Banzer W. Health benefits of light-intensity physical activity: a systematic review of accelerometer data of the national health and nutrition examination survey (NHANES). Sports Med 2017;47(9):1769-1793. [doi: 10.1007/s40279-017-0724-0] [Medline: 28393328]
72. Every move counts towards better health - says WHO. World Health Organisation. 2020. URL: https://www.who.int/news/item/25-11-2020-every-move-counts-towards-better-health-says-who [accessed 2024-04-10]
73. English indices of deprivation 2019. Ministry of Housing. 2019. URL: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019 [accessed 2025-12-02]
74. Katz JN, Chang LC, Sangha O, Fossel AH, Bates DW. Can comorbidity be measured by questionnaire rather than medical record review? Med Care 1996;34(1):73-84. [doi: 10.1097/00005650-199601000-00006] [Medline: 8551813]
75. Gresham G, Schrack J, Gresham LM, Shinde AM, Hendifar AE, Tuli R, et al. Wearable activity monitors in oncology trials: current use of an emerging technology. Contemp Clin Trials 2018;64:13-21 [FREE Full text] [doi: 10.1016/j.cct.2017.11.002] [Medline: 29129704]
76. George SM, Alfano CM, Groves J, Karabulut Z, Haman KL, Murphy BA, et al. Objectively measured sedentary time is related to quality of life among cancer survivors. PLoS One 2014;9(2):e87937 [FREE Full text] [doi: 10.1371/journal.pone.0087937] [Medline: 24505335]
77. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. Can J Appl Sport Sci 1985;10(3):141-146. [Medline: 4053261]
78. Amireault S, Godin G, Lacombe J, Sabiston CM. The use of the Godin-Shephard Leisure-Time Physical Activity Questionnaire in oncology research: a systematic review. BMC Med Res Methodol 2015;15:60 [FREE Full text] [doi: 10.1186/s12874-015-0045-7] [Medline: 26264621]
79. Ross KM, Wing RR. Concordance of in-home "smart" scale measurement with body weight measured in-person. Obes Sci Pract 2016;2(2):224-248 [FREE Full text] [doi: 10.1002/osp4.41] [Medline: 27499884]
80. Yorkin M, Spaccarotella K, Martin-Biggers J, Quick V, Byrd-Bredbenner C. Accuracy and consistency of weights provided by home bathroom scales. BMC Public Health 2013;13:1194 [FREE Full text] [doi: 10.1186/1471-2458-13-1194] [Medline: 24341761]
81. Contardo Ayala AM, Nijpels G, Lakerveld J. Validity of self-measured waist circumference in adults at risk of type 2 diabetes and cardiovascular disease. BMC Med 2014;12:170 [FREE Full text] [doi: 10.1186/s12916-014-0170-x] [Medline: 25274418]
82. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
83. Pickard AS, Wilke CT, Lin H, Lloyd A. Health utilities using the EQ-5D in studies of cancer. Pharmacoeconomics 2007;25(5):365-384. [doi: 10.2165/00019053-200725050-00002] [Medline: 17488136]
84. Cella DF, Tulsky DS, Gray G, Sarafian B, Linn E, Bonomi A, et al. The functional assessment of cancer therapy scale: development and validation of the general measure. J Clin Oncol 1993;11(3):570-579. [doi: 10.1200/JCO.1993.11.3.570] [Medline: 8445433]
85. Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the functional assessment of cancer therapy (FACT) measurement system. J Pain Symptom Manage 1997;13(2):63-74 [FREE Full text] [doi: 10.1016/s0885-3924(96)00274-6] [Medline: 9095563]
86. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
87. Mollayeva T, Thurairajah P, Burton K, Mollayeva S, Shapiro CM, Colantonio A. The Pittsburgh Sleep Quality Index as a screening tool for sleep dysfunction in clinical and non-clinical samples: A systematic review and meta-analysis. Sleep Med Rev 2016;25:52-73. [doi: 10.1016/j.smrv.2015.01.009] [Medline: 26163057]
88. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing Generalized Anxiety Disorder: the GAD-7. Arch Intern Med 2006;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]

XSL•FO
RenderX

89. Esser P, Hartung TJ, Friedrich M, Johansen C, Wittchen H, Faller H, et al. The Generalized Anxiety Disorder screener (GAD-7) and the Anxiety module of the Hospital and Depression Scale (HADS-A) as screening tools for generalized anxiety disorder among cancer patients. Psychooncology 2018;27(6):1509-1516. [doi: 10.1002/pon.4681] [Medline: 29473255]
90. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
91. Hartung TJ, Friedrich M, Johansen C, Wittchen H, Faller H, Koch U, et al. The Hospital Anxiety and Depression Scale (HADS) and the 9-item patient health questionnaire (PHQ-9) as screening instruments for depression in patients with cancer. Cancer 2017;123(21):4236-4243 [FREE Full text] [doi: 10.1002/cncr.30846] [Medline: 28654189]
92. Haas BK, Northam S. Measuring self-efficacy: development of the physical activity assessment inventory. Southern Online Journal of Nursing Research 2010;10(4).
93. Foster C, Breckons M, Cotterell P, Barbosa D, Calman L, Corner J, et al. Cancer survivors' self-efficacy to self-manage in the year following primary treatment. J Cancer Surviv 2015;9(1):11-19 [FREE Full text] [doi: 10.1007/s11764-014-0384-0] [Medline: 25028218]
94. Gardner B, Abraham C, Lally P, de Bruijn G. Towards parsimony in habit measurement: testing the convergent and predictive validity of an automaticity subscale of the self-report habit index. Int J Behav Nutr Phys Act 2012;9:102 [FREE Full text] [doi: 10.1186/1479-5868-9-102] [Medline: 22935297]
95. Kliem S, Mößle T, Rehbein F, Hellmann DF, Zenger M, Brähler E. A brief form of the perceived social support questionnaire (F-SozU) was developed, validated, and standardized. J Clin Epidemiol 2015;68(5):551-562. [doi: 10.1016/j.jclinepi.2014.11.003] [Medline: 25499982]
96. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol 1976;4(2):97-110. [Medline: 1027738]
97. Randler C. Chronotype correlates with developmental index, intelligence and academic achievement: A study based on nationwide indicators. Chronobiol Int 2017;34(7):985-992. [doi: 10.1080/07420528.2017.1332070] [Medline: 28594245]
98. Terman MR, Jacobs JB, White TM. Morningness-Eveningnes05s Questionnaire Self-Assessment Version (MEQ-SA). 2001. URL: https://cet.org/wp-content/uploads/2017/11/Morningness-Eveningness-Self-assessment-from-cet.org_.pdf [accessed 2023-12-02]
99. Ersche KD, Lim T, Ward LH, Robbins TW, Stochl J. Creature of habit: a self-report measure of habitual routines and automatic tendencies in everyday life. Pers Individ Dif 2017;116:73-85 [FREE Full text] [doi: 10.1016/j.paid.2017.04.024] [Medline: 28974825]
100. Patel A, Rendu A, Moran P, Leese M, Mann A, Knapp M. A comparison of two methods of collecting economic data in primary care. Fam Pract 2005;22(3):323-327. [doi: 10.1093/fampra/cmi027] [Medline: 15824056]
101. Byford S, Leese M, Knapp M, Seivewright H, Cameron S, Jones V, et al. Comparison of alternative methods of collection of service use data for the economic evaluation of health care interventions. Health Econ 2007;16(5):531-536. [doi: 10.1002/hec.1175] [Medline: 17001749]
102. Perski O, Lumsden J, Garnett C, Blandford A, West R, Michie S. Assessing the psychometric properties of the digital behavior change intervention engagement scale in users of an app for reducing alcohol consumption: evaluation study. J Med Internet Res 2019;21(11):e16197 [FREE Full text] [doi: 10.2196/16197] [Medline: 31746771]
103. Salam A, Woodman A, Chu A, Al-Jamea LH, Islam M, Sagher M, et al. Effect of post-diagnosis exercise on depression symptoms, physical functioning and mortality in breast cancer survivors: a systematic review and meta-analysis of randomized control trials. Cancer Epidemiol 2022;77:102111. [doi: 10.1016/j.canep.2022.102111] [Medline: 35091272]
104. Wang Y, Jacobs EJ, Gapstur SM, Maliniak ML, Gansler T, McCullough ML, et al. Recreational physical activity in relation to prostate cancer-specific mortality among men with nonmetastatic prostate cancer. Eur Urol 2017;72(6):931-939. [doi: 10.1016/j.eururo.2017.06.037] [Medline: 28711382]
105. Hackett A, Strickland K. Using the framework approach to analyse qualitative data: a worked example. Nurse Res 2019;26(2):8-13. [doi: 10.7748/nr.2018.e1580] [Medline: 30215482]
106. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
107. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
108. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
109. Schrack JA, Cooper R, Koster A, Shiroma EJ, Murabito JM, Rejeski WJ, et al. Assessing daily physical activity in older adults: unraveling the complexity of monitors, measures, and methods. J Gerontol A Biol Sci Med Sci 2016;71(8):1039-1048 [FREE Full text] [doi: 10.1093/gerona/glw026] [Medline: 26957472]

XSL•FO
RenderX

110. Keen C, Phillips G, Thelwell M, Humphreys L, Evans L, Myers A, et al. Correction: Keen et al. establishing innovative complex services: learning from the active together cancer prehabilitation and rehabilitation service. Healthcare (Basel) 2024;12(12):3007 [FREE Full text] [doi: 10.3390/healthcare12121175] [Medline: 38921352]
111. Kennedy F, Smith S, Beeken RJ, Buck C, Williams S, Martin C, et al. An app-based intervention with behavioral support to promote brisk walking in people diagnosed with breast, prostate, or colorectal cancer (APPROACH): process evaluation study. JMIR Cancer 2025;11:e64747 [FREE Full text] [doi: 10.2196/64747] [Medline: 39928926]
112. Toohey K, Chapman M, Rushby A, Urban K, Ingham G, Singh B. The effects of physical exercise in the palliative care phase for people with advanced cancer: a systematic review with meta-analysis. J Cancer Surviv 2023;17(2):399-415. [doi: 10.1007/s11764-021-01153-0] [Medline: 35040076]
113. Rodríguez-Cañamero S, Cobo-Cuenca AI, Carmona-Torres JM, Pozuelo-Carrascosa DP, Santacruz-Salas E, Rabanales-Sotos J, et al. Impact of physical exercise in advanced-stage cancer patients: systematic review and meta-analysis. Cancer Med 2022;11(19):3714-3727 [FREE Full text] [doi: 10.1002/cam4.4746] [Medline: 35411694]
114. Callard F, Broadbent M, Denis M, Hotopf M, Soncul M, Wykes T, et al. Developing a new model for patient recruitment in mental health services: a cohort study using Electronic Health Records. BMJ Open 2014;4(12):e005654 [FREE Full text] [doi: 10.1136/bmjopen-2014-005654] [Medline: 25468503]
115. OFCOM TECHNOLOGY TRACKER 2021. OFCOM. 2021. URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0015/219102/technology-tracker-2021-data-tables.pdf [accessed 2022-12-02]
116. Waters L, Reeves M, Fjeldsoe B, Eakin E. Control group improvements in physical activity intervention trials and possible explanatory factors: a systematic review. J Phys Act Health 2012;9(6):884-895. [doi: 10.1123/jpah.9.6.884] [Medline: 22898467]
117. Freene N, Davey R, Sathiyakumar R, McPhail SM. Can physical activity measurement alone improve objectively-measured physical activity in primary care?: A systematic review and meta-analysis. Prev Med Rep 2020;20:101230 [FREE Full text] [doi: 10.1016/j.pmedr.2020.101230] [Medline: 33163332]
118. Milton K, Bull FC, Bauman A. Reliability and validity testing of a single-item physical activity measure. Br J Sports Med 2011;45(3):203-208 [FREE Full text] [doi: 10.1136/bjsm.2009.068395] [Medline: 20484314]
119. Jennings CA, Vandelanotte C, Caperchione CM, Mummery WK. Effectiveness of a web-based physical activity intervention for adults with type 2 diabetes-a randomised controlled trial. Prev Med 2014;60:33-40. [doi: 10.1016/j.ypmed.2013.12.011] [Medline: 24345601]
120. Vandelanotte C, Stanton R, Rebar AL, Van Itallie AK, Caperchione CM, Duncan MJ, et al. Physical activity screening to recruit inactive randomized controlled trial participants: how much is too much? Trials 2015;16:446 [FREE Full text] [doi: 10.1186/s13063-015-0976-7] [Medline: 26449210]
121. Smith L, Lee JA, Mun J, Pakpahan R, Imm KR, Izadi S, et al. Levels and patterns of self-reported and objectively-measured free-living physical activity among prostate cancer survivors: a prospective cohort study. Cancer 2019;125(5):798-806 [FREE Full text] [doi: 10.1002/cncr.31857] [Medline: 30516839]

Abbreviations

BCT: behavior change technique
CONSORT: Consolidated Standards of Reporting Trials
COREQ: Consolidated Criteria for Reporting Qualitative Research
CSRI: Client Service Receipt Inventory
FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue
FACT-G: Functional Assessment of Cancer Therapy-General
GAD-7: Generalized Anxiety Disorder-7
GLTEQ: Godin Leisure Time Exercise Questionnaire
HES: Hospital Episode Statistics
MET: metabolic equivalent of task
MVPA: moderate-to-vigorous physical activity
NDRS: National Disease Registration Service
NHS: National Health Service
PA: physical activity
PAAI: Physical Activity Appraisal Inventory
PHQ-9: Patient Health Questionnaire-9
PPI: patient and public involvement
PSQI: Pittsburgh Sleep Quality Index
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
spm: steps per minute
SRBAI: Self-Report Behavioral Automaticity Index

XSL•FO
RenderX

TIDieR: Template for Intervention Description and Replication
UCL: University College London

Edited by A Schwartz; submitted 07.May.2025; this is a non–peer-reviewed article;accepted 29.Nov.2025; published 13.Jan.2026.

Please cite as:
Kennedy F, Smith SL, Beeken RJ, Buck C, Shearsmith L, Truscott H, Counsell N, Thomas C, Roberts A, Greenfield DM, Potts HWW, Latimer N, Lowther E, Smith L, Stevens R, Creaser A, Gath J, Wyld L, Fisher A, Lally P
An App-Based Behavioral Support Intervention Promoting Physical Activity (APPROACH) in Patients Diagnosed With Breast, Prostate, or Colorectal Cancer: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e77096
URL: https://www.researchprotocols.org/2026/1/e77096
doi:10.2196/77096
PMID:

©Fiona Kennedy, Sue L Smith, Rebecca J Beeken, Caroline Buck, Leanne Shearsmith, Hannah Truscott, Nicholas Counsell, Chloe Thomas, Anna Roberts, Diana M Greenfield, Henry WW Potts, Nicholas Latimer, Ed Lowther, Lee Smith, Rosie Stevens, Amy Creaser, Jacqui Gath, Lynda Wyld, Abigail Fisher, Phillippa Lally. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Enhancing Cognitive Functions in Older Adults With Mild Cognitive Impairment via Virtual Sail 3D: Protocol for a Feasibility Randomized Controlled Trial

Federica Sancassiani[1], PsyD, PhD, Prof Dr; Alessandra Perra[1], PhD; Veronica Vacca[2], MSc; Roberta Montisci[1], MD, Prof Dr Med; Mariarita Monni[2], BSc; Stefano Lorrai[1], MSc; Giulia Cossu[1], PsyD, PhD; Donatella Rita Petretto[3], PsyD, Prof Dr; Lorenzo di Natale[4,5], PsyD; Diego Primavera[1], MD, PhD; Stefania Redolfi[1], MD, PhD, Prof Dr Med; Angelo Scuteri[1], MD, Prof Dr Med; Antonio Egidio Nardi[6], MD, Prof Dr Med; Goce Kalcev[2], MD, PhD; Massimiliano Pau[7], ING, PhD, Prof Dr; Mauro Giovanni Carta[1], MD, Prof Dr Med

[1]Department of Medical Sciences and Public Health, University of Cagliari, Cagliari, Italy

[2]University Hospital of Cagliari (AOU Cagliari), Cagliari, Italy

[3]Department of Pedagogy, Psychology and Philosophy, University of Cagliari, Cagliari, Italy

[4]Cerebrum VR Society, Rome, Italy

[5]IDEGO Digital Psychology Society, Rome, Italy

[6]Laboratory of Panic and Respiration-IPUB, Institute of Psychiatry of the Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

[7]Department of Mechanical, Chemical and Materials Engineering, University of Cagliari, Cagliari, Italy

Corresponding Author:
Goce Kalcev, MD, PhD
University Hospital of Cagliari (AOU Cagliari)
Via Ospedale 54
Cagliari, 09126
Italy
Phone: 39 00393663579333
Email: gocekalcev@yahoo.com

Abstract

Background: Innovative approaches, such as virtual reality (VR)–based cognitive rehabilitation, are becoming essential to healthy aging.

Objective: This study aims to evaluate the feasibility and generate preliminary evidence regarding the effectiveness of an immersive VR-based cognitive remediation intervention for older individuals with mild cognitive impairment (MCI).

Methods: A total of 40 individuals aged 65 years and older, of both genders with MCI, will be recruited and randomly assigned to the experimental or control group. The experimental group will participate in a 6-week VR-based cognitive remediation program, while the control group will receive healthy lifestyle education. Feasibility will be assessed through measures, including dropout rates, side effects, and satisfaction with the program. Preliminary effectiveness will be evaluated using assessments of cognitive functions, quality of life, functional mobility, depression and anxiety symptoms, regulation of social and biological rhythms, body awareness, and physical activity levels.

Results: Recruitment began in May 2024 and is expected to conclude by November 2025; to date, 153 participants have been recruited and screened, with 39 enrolled. Preliminary analyses are planned for January 2026 and follow-up analyses for January 2027.

Conclusions: This trial will provide insights into the feasibility and preliminary effectiveness of a VR-based intervention for older adults with MCI, with implications for its integration into preventive and rehabilitative care.

Trial Registration: ClinicalTrial.gov NCT06579378; https://clinicaltrials.gov/search?term=NCT06579378

International Registered Report Identifier (IRRID): DERR1-10.2196/85089

(JMIR Res Protoc 2026;15:e85089) doi:10.2196/85089

XSL•FO
RenderX

KEYWORDS

aging; cognitive functions; emerging technologies; feasibility; mild cognitive impairment; Sail; virtual reality

Introduction

With the ongoing global increase in life expectancy, mild cognitive impairment (MCI) and dementia have emerged as leading causes of disability, especially among older adults [1]. Italy has the highest proportion of older adults in Europe, with 23% of its population aged 65 years and older [2]. Projections suggest that by 2050, 1 in 5 Italians will be in advanced age [2]. This demographic shift is expected to result in a shrinking working-age population, accompanied by a significant increase in both direct and indirect health care and social costs associated with age-related health issues [3]. MCI is characterized as a transitional phase bridging the gap between typical cognitive decline and the early stages of dementia [4]. Cognitive decline is frequently associated with symptoms of depression and anxiety, both of which contribute to a progressive deterioration in functional capacity during daily activities [5]. These conditions negatively impact quality of life (QoL) and disrupt the regulation of biological and social rhythms [6]. Given these considerations, both researchers and the European Union have increasingly focused attention on the critical objective of advancing research in the field of healthy aging, particularly in the development and implementation of nonpharmacological interventions aimed at preventing or delaying cognitive decline [7,8].

Emerging evidence from recent literature reviews highlights the beneficial impact of cognitive training, cognitive remediation interventions, and physical activity on cognitive performance in older adults, especially when these approaches are embedded within a comprehensive public health framework [9-11]. Research on active aging has been identified as a priority, particularly concerning the feasibility and effectiveness of nonpharmacological interventions aimed at preventing and/or delaying cognitive decline [11-13]. Risk factors such as physical, cognitive, and social inactivity are estimated to account for about 40% of the risk of developing dementia in a lifetime [14]. Contemporary reviews have shown that cognitive remediation interventions are effective in improving cognitive performance in older adults with and without cognitive decline [15,16]. Furthermore, cognitive remediation is a behavior-based training intervention aimed at improving cognitive functions (memory, attention, executive functions, social cognition, and metacognition) to achieve lasting results and their generalization [17]. There is substantial evidence supporting the effectiveness of such interventions in neuropsychological disorders, such as dementia, MCI, and behavioral disorders [18].

In recent years, there has been an increase in the use of immersive virtual reality (VR) software in clinical and rehabilitative settings as tools for rehabilitation and the improvement of various skills [18,19]. To date, there are few studies where immersive VR is used to enhance cognitive functions related to personal and social functioning, especially to prevent cognitive decline in older adults (aged 65 years and older), and the quality of evidence is rated as moderate to low [20,21]. VR serves as a core tool in facilitating cognitive and functional skill acquisition or enhancement, owing to its capacity to deliver immersive, ecologically valid, and highly realistic experiential learning environments [22]. These findings highlight not only improvements in specific cognitive domains, such as memory, language, and executive functioning, but also improvements in overall clinical, personal, and social functioning [20,21,23]. Despite the encouraging evidence, the methodological rigor of existing studies remains limited. Many studies lack comprehensive descriptions of the VR interventions and the specific methods used during the sessions. Additionally, most studies do not clarify the framework used to develop the interventions, which is essential for understanding the alignment between the hypotheses, outcomes, methods, and the VR programs designed to target specific cognitive domains. This lack of detailed information prevents the establishment of a standard operating procedure, which is crucial for ensuring the reproducibility of the intervention and for guiding the development of a future gold standard in the field [20,21,23-27]. Moreover, an integrative, multidisciplinary approach, grounded in individualized recovery-oriented objectives, is imperative for addressing the multifaceted and evolving needs of people with MCI.

The trial aims to assess the feasibility (primary outcome) of a 6-week intervention to improve cognitive functions using the immersive VR software “CEREBRUM” (Cerebrum VR Society) with sailing scenarios among the older adults with MCI. It also aims to evaluate, through a randomized controlled trial (RCT) design, the effectiveness of the experimental intervention in enhancing cognitive functions and/or mitigating further cognitive decline and deterioration (secondary outcome). Finally, it will verify whether the experimental intervention has a positive impact on (1) QoL, (2) functional mobility, (3) depressive and anxiety symptoms, (4) social and biological rhythms, (5) body awareness, and (6) physical activity levels (tertiary outcomes).

Methods

Study Design

This study is designed as a feasibility RCT. Assessments will be conducted at the following time points: T0 (0 weeks - baseline), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), T5 (48 weeks from T1). The study adopted a longitudinal, 2-arm design, with participants randomized to either an active interventional protocol based on the Virtual Sail 3D (VSail) experimental protocol or to a control setting, “WEB-Health” (education on healthy lifestyles to promote active aging). Participants will be independently randomized in a 1:1 ratio to one of the two parallel assignment groups: (1) VSail or (2) WEB-Health. The study protocol is written according to the CONSORT (Consolidated Standards of Reporting Trials) 2010 statement for randomized pilot and feasibility studies [28] (Figure 1).

XSL•FO
RenderX

Figure 1. CONSORT flow diagram.

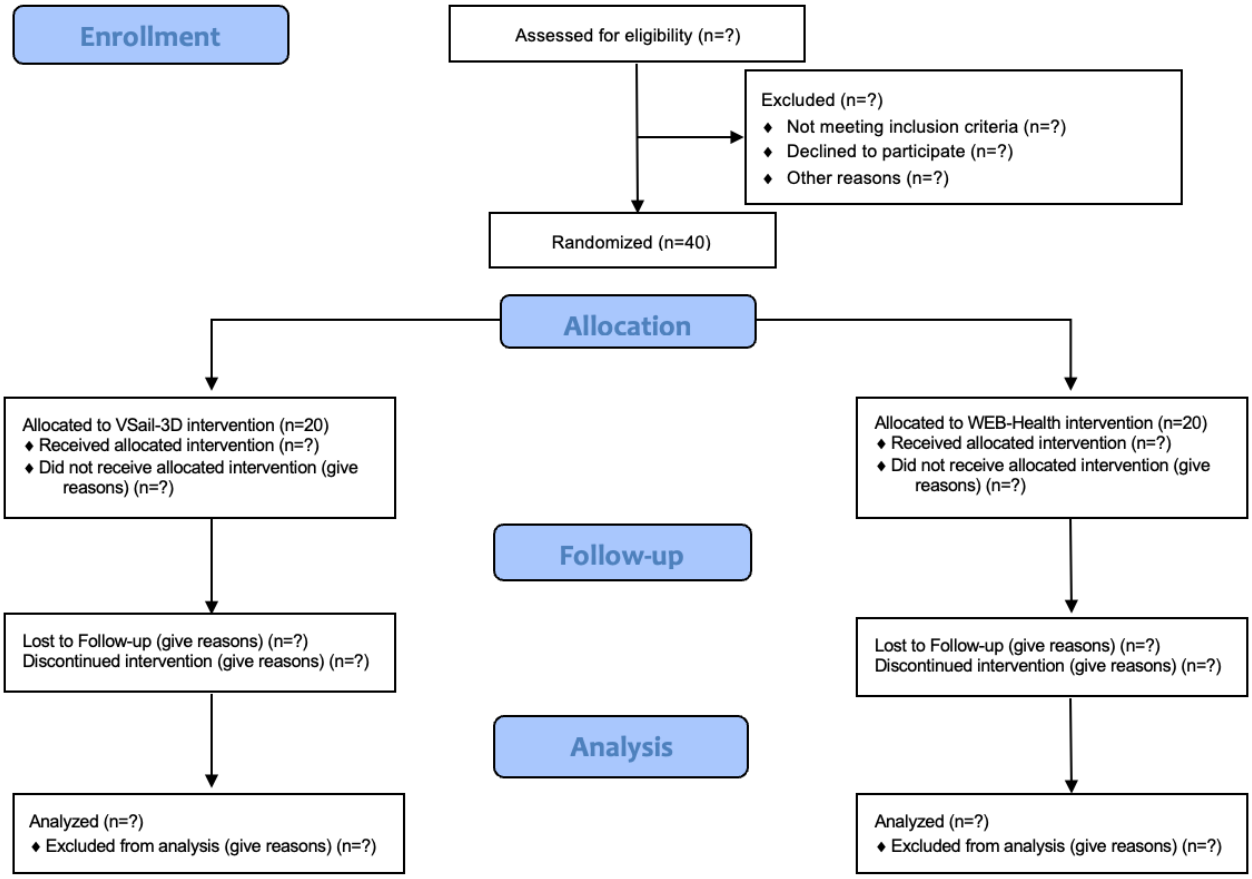

Participants, Recruitment, and Setting

The target population includes older participants with MCI conditions.

Inclusion criteria are being 65 years and older of both genders, without conditions that would prevent them from participating in the activities of the experimental intervention "VSail," without severe issues with autonomous mobility, with MCI according to Addenbrooke's Cognitive Examination-Revised (ACE-R) for cognitive functions, which also includes the Mini-Mental State Examination (MMSE) [29], with the following ranges: ACE=66.93-84.93 and/or MMSE=27.02-27.11, and signing the informed consent.

Exclusion criteria are severe cardiovascular conditions, severe issues with autonomous mobility, severe metabolic disorders not pharmacologically compensated, severe neurological conditions that prevent from participating in the experimental protocol, such as a stroke within the past 2 years, Parkinson disease, epilepsy, or dementia (Alzheimer, vascular, etc), severe ongoing bronchopulmonary disorders, severe ongoing renal disorders, glaucoma, retinal detachment, or other serious vision conditions that do not allow the safe use of VR technology, active malignant neoplasm that do not allow the hinders participation in the intervention.

Recruitment activities will be conducted at the sites of the project's partner units, namely, the Consultation-Liaison Psychiatry and Psychosomatics Center, the Cardiology–CCU–Hemodynamics Unit, and the Multidisciplinary Sleep Medicine Center of the University Hospital of Cagliari (AOU Cagliari), Italy. At each site, infographics will be displayed and informational brochures about the project will be made available so that outpatients attending the relevant clinics can express their interest in participating. In addition, a health care professional will contact by telephone all patients who, based on the study's inclusion and exclusion criteria, will be identified by clinicians of the partner units as potentially eligible. All patients who express interest and meet the eligibility criteria will then be invited to provide written informed consent and to undergo screening with the ACE-R [29].

Participants who screen positive for MCI will be enrolled and subsequently randomized to the experimental or control arm. Those who screen negative will be excluded from study participation.

According to the inclusion and exclusion criteria, recruitment will continue until the target sample size is reached, allowing for randomization and the conduct of interventions.

All assessment activities using the instruments specified in the protocol, as well as the implementation of both the control and experimental interventions, will take place at the Consultation-Liaison Psychiatry and Psychosomatics Center. These activities will be conducted by mental health professionals who will be duly trained in the use of the assessment tools and in the delivery of the VR-based cognitive remediation protocol (experimental intervention) and the web-based healthy lifestyle seminars (control intervention). Motor performance assessment

XSL•FO
RenderX

(balance, gait, and functional mobility) will be carried out with the support of biomechanical engineers who will receive appropriate training for this purpose.

Randomization and Allocation Concealment

Participants who meet the eligibility criteria will be randomly assigned to one of the two groups: an intervention group and a control group. Randomization will be conducted after baseline assessment. Randomization will be computer-based and concealed. The individuals conducting the randomization procedure will be blinded to participants' identities and statuses and will not receive any information about them. Participants will be randomly assigned to experimental or control groups in a 1:1 ratio using block randomization. Allocation codes were concealed.

Blinding

Due to the intervention method, it may not be feasible to effectively blind both the participants and professionals engaged in the study.

Interventions

The control intervention encompasses web-based activities aimed at promoting healthy lifestyles for active aging (WEB-Health). This intervention includes education on healthy lifestyles to promote active aging (topics: physical activity in old age, effective communication, stress management, technologies, and health). It consists of 3 seminars, delivered in the form of webinars, on a biweekly basis over 6 consecutive weeks, lasting 50-60 minutes each. Each webinar will be led by a qualified mental health professional and delivered live in plenary session through the Zoom platform (Zoom Video Communications, Inc).

The experimental intervention (VSail) includes 12 sessions (2 times a week, for 6 consecutive weeks) with scenarios (virtual environments) dedicated to the sport of sailing. The duration of each session is 50-60 minutes. It is based on the immersive VR software "CEREBRUM" developed by the Cerebrum VR Society. The software represents one of the most recent tools in psychiatric rehabilitation, specifically designed for cognitive remediation. Developed by mental health professionals, "CEREBRUM" is compatible with the "Oculus Quest" VR headset, which carries the CE marking, ensuring adherence to European safety, health, and environmental protection standards. Within the immersive virtual environment, participants interacted with scenarios replicating daily activities in both home and urban settings.

For the present project, the "CEREBRUM" software was further developed through the addition of 4 virtual scenarios specifically designed around the sport of sailing. The decision to add these scenarios was driven by the aim of making the virtual experience during cognitive training more interactive and of simultaneously engaging motor functions. The "CEREBRUM" sailing-based scenarios enable users to move within the virtual space to grasp and manipulate objects (eg, the helm) and to perform specific movements required for executing sailing manoeuvres (eg, approaching the sheet, bending down to reach it, and alternating rapid arm movements to haul the line and adjust the sail). The "CEREBRUM" software comprises exercises of varying difficulty designed to train various cognitive functions, including attention, memory, learning, cognitive estimation, working memory, executive functions, motor abilities, and language. The different degrees of difficulty are designed to adapt to the user's functional diagnosis. The operator who conducts the intervention must adapt the difficulty level to the user's residual abilities so that the exercises are neither too easy nor too complex. Each session, after an initial part of welcome, psychoeducation and orientation to the instrument, involves alternating VR exercises for a maximum of 15-20 minutes, positive and corrective feedback, and suggestions of practical homework related to the cognitive function trained during the session that the individual should try during daily life (generalization phase).

Outcomes

The primary outcomes regard feasibility and include (1) dropout rates, (2) the proportion of recruited participants among those considered eligible, (3) the incidence of potential adverse effects associated with the use of immersive VR technology (eg, nausea, dizziness, visual fatigue, and headache), and (4) the level of perceived satisfaction reported by participants following the experimental intervention.

The secondary outcomes regard cognitive functioning and include (1) executive functions, (2) memory, (3) visuospatial abilities, (4) selective attention, and (5) general cognitive functioning.

The tertiary outcomes regard psychosocial and physical functioning and include (1) QoL, (2) depressive symptoms, (3) anxiety symptoms, (4) regulation of social and biological rhythms, (5) body awareness, (6) levels of physical activity, and (7) mobility.

Also, the following sociodemographic and clinical variables will be collected for each participant: gender and age, marital status, educational attainment, past and current occupational status, history of and current organic physical illnesses, history of and current mental health diagnoses, and current medication use.

Evaluation Tools and Timeline

Outcomes will be assessed using validated questionnaires, tests, and performance-based measures at predefined time points (Textbox 1).

XSL•FO
RenderX

Textbox 1. The tools and evaluation procedures.

Primary outcomes (feasibility): the tools and evaluation procedures

- Tolerability, including dropout rates and acceptability, which considers the proportion of recruited participants among those considered eligible. These will be calculated following this timeline: T0 (0 weeks - baseline), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Side and secondary effects through the Simulator Sickness Questionnaire (SSQ) [30], a self-report questionnaire including 16 items that evaluate the frequency of unwanted effects due to VR technologies, such as nausea, dizziness, headaches, eye strain, etc. This questionnaire will be administered following this timeline: T0 (0 weeks - baseline) and T1 (6 weeks).
- The level of satisfaction with the program as a whole, the effect on psychosocial and physical health, assisting and listening abilities demonstrated by the operators, organizational support for activities, and satisfaction with initial expectations by an ad hoc self-report satisfaction questionnaire. This questionnaire will be administered at T1 (6 weeks after T0).

Secondary outcomes (cognitive functioning): the tools and the evaluation procedures

- Trail Making Test [31]: it is a widely administered neuropsychological assessment tool designed to evaluate cognitive flexibility, visual attention, and processing speed. It consists of 2 parts: TMT-A, which requires participants to connect numbered circles sequentially, and TMT-B, which alternates between numbers and letters, assessing more complex executive functions such as task-switching and set-shifting. Reliability studies have reported Cronbach α values ranging from 0.70 to 0.89, depending on the sample characteristics and testing conditions. The questionnaires will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Digit Span [32]: it is a neuropsychological test used to evaluate working memory, attention, and immediate verbal memory. It consists of 2 components: Digit Span Forward, which evaluates simple attention and memory span by requiring participants to repeat a sequence of numbers in the same order, and Digit Span Backward, which assesses working memory and cognitive flexibility by requiring participants to repeat the sequence in reverse order. Reliability studies have reported Cronbach α values ranging from 0.70 to 0.90, indicating good internal consistency across different populations and testing contexts. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Stroop Test [33]: it is a widely used neuropsychological test that measures cognitive control, selective attention, and processing speed. It evaluates the ability to inhibit automatic responses and resolve cognitive interference by requiring participants to identify the ink color of a word that may denote a different color name (eg, the word “red” printed in blue ink). The test typically includes 3 conditions: “reading colour names,” “naming coloured blocks,” and “naming the ink colour of incongruent colour-word pairs.” The Stroop effect, reflected in the increased time or errors during the incongruent condition, provides insights into executive functioning and attentional control. Cronbach α values reported in reliability studies range from 0.72 to 0.91, depending on the version of the test and the characteristics of the population studied. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Rey Figure Test [34]: it is a neuropsychological assessment that measures praxis abilities, visuospatial abilities, working memory, long-term memory, and executive functions. The test involves having a person first copy a complex drawing and then reproduce it from memory after a delay. Studies have reported varying Cronbach α values, but they generally indicate adequate internal consistency for the Rey Figure Test scoring methods. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Frontal Assessment Battery (FAB) [35]: it is a brief neuropsychological assessment instrument designed to evaluate executive functions and cognitive processes associated with frontal lobe activity. It consists of 6 subtests assessing different aspects of executive function: conceptualization, mental flexibility, motor programming, sensitivity to interference, inhibitory control, and environmental autonomy. Studies have reported Cronbach α values ranging from 0.78 to 0.90, indicating good internal consistency across various populations. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Matrix Test [36]: it is a cognitive assessment tool designed to evaluate nonverbal reasoning, abstract thinking, and problem-solving abilities. It typically involves identifying patterns or completing sequences in visual matrices. The reliability of the Matrix Test, as indicated by Cronbach α, generally ranges from 0.80 to 0.90, reflecting good to excellent internal consistency. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Rey Word Test [37]: it is a neuropsychological measure designed to assess verbal memory performance and identify potential malingering through tasks involving the recall and recognition of word lists. The Cronbach α falls within a range of 0.70 to 0.85, depending on the specific population and administration conditions, indicating acceptable to good internal consistency. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Addenbrooke’s Cognitive Examination-Revised (ACE-R) [29]: it is a brief tool used for the assessment of 5 key cognitive domains: attention and orientation, memory, verbal fluency, language, and visuospatial abilities. The test has a maximum score of 100, with a recommended range of 66.93-84.93 for the detection of mild cognitive impairment. Notably, ACE-R incorporates the Mini-Mental State Examination (MMSE) as part of its structure. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).

Tertiary outcomes (psychosocial and physical functioning): the tools and evaluation procedures

-

XSL•FO
RenderX

Short Form Health Survey-12 item (SF-12) [38,39]: it assesses the following dimensions of well-being and quality of life: vitality, physical functioning, bodily pain, general health perception, mental, physical, and emotional health, work functioning, and social role. The total score ranges from 12 to 47, with higher scores indicating better quality of life. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).

- Patient Health Questionnaire-9 (PHQ-9) [40]: it assesses the presence and severity of 9 depressive symptoms experienced over the past 2 weeks, including depressed mood, anhedonia, sleep and appetite disturbances, low energy, difficulty concentrating, psychomotor disturbances, loss of self-esteem, and suicidal ideation. The total score ranges from 0 to 27, with higher scores indicating greater severity of depressive symptoms. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Generalized Anxiety Disorder-7 item (GAD-7) [41]: it assesses the presence and severity of 7 anxiety symptoms experienced over the previous 2 weeks, including nervousness, inability to stop worrying, excessive worry, restlessness, difficulty relaxing, irritability, and fear that something awful might happen. Scores range from 0 to 21, with higher values reflecting more severe anxiety symptoms. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Brief Social Rhythms Scale (BSRS) [42]: it assesses the (ir)regularity of rhythms over the past week using 10 items. It evaluates sleep (wake-up and bedtime), eating habits (across different meals), and social interactions (eg, work and/or leisure activities). Scores range from 10 to 60, with higher scores indicating poorer rhythm regularity. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Physical Body Experiences Questionnaire Simplified for Active Aging (PBE-QAG) [43]: it assesses body awareness over the past month, focusing on 4 dimensions using 12 items: relationship with one's own body, body acceptance, awareness of physical abilities, and awareness of physical limitations. Scores range from 12 to 60, with higher scores indicating lower overall body awareness. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- International Physical Activity Questionnaire (IPAQ) [44]: it assesses the type and amount of physical activity performed over the past 7 days, including both work-related and leisure activities, using 9 items. The total score is based on the conventional unit of measure "MET," where higher MET corresponds to higher levels of physical activity. The test will be administered following this timeline: T0 (0 weeks), T1 (6 weeks), T2 (12 weeks from T1), T3 (24 weeks from T1), T4 (36 weeks from T1), and T5 (48 weeks from T1).
- Sway area in the presence and absence of visual input [45]: it assesses the quantitative balance maintenance abilities under static conditions. The evaluation will be conducted under 2 conditions: eyes open and eyes closed. Participants will be asked to step onto the platform barefoot and remain as still as possible for 30 seconds. The test will be repeated 3 times for each condition. This assessment will be administered according to the following timeline: T0 (baseline, 0 weeks) and T1 (6 weeks).
- Gait speed analysis (GA) [46]: it assesses the analysis of gait, using a miniaturized wearable inertial sensor. The sensor is positioned on the lumbar region of the trunk at the level of vertebrae L5-S1. The participant is asked to walk at a comfortable, self-selected speed along a straight, unobstructed path of at least 2 m. This assessment will be administered according to the following timeline: T0 (baseline, 0 weeks) and T1 (6 weeks).
- Timed-Up-and-Go (TUG) [47]: it assesses the functional mobility in older adults. The test involves standing up from a chair, walking for 3 m, turning around an obstacle, returning to the starting point, performing a 180° turn, and sitting back down. The participants will perform an instrumented version of the TUG using a wearable inertial sensor placed on the lower back at the level of vertebra L2. This assessment will be administered according to the following timeline: T0 (baseline, 0 weeks) and T1 (6 weeks).

Data Management and Security

The main goal of the data management process is to guarantee the timely delivery of high-quality data in order to meet the requirements of appropriate statistical analysis. It includes all phases of data management, such as data collection, processing, and application. Clinical data will be compiled into a structured longitudinal database to ensure consistency and facilitate research activities. Each participant will be assigned a unique identification number at the start of the study. The study will be conducted following the Declaration of Helsinki [48] and the requirements of all applicable local and international standards, according to data protection laws.

Adverse Event Reporting

Each adverse event will be documented with detailed information, including the type of event, onset and resolution dates, severity, expectedness, outcome, and its potential relationship to the study.

Data Analysis, Statistical Power, and Sample Size

The statistical analysis will use SPSS software (version 21; IBM Corp). Descriptive statistics will be calculated for all variables, with nominal variables reported as frequencies and percentages (n; %) and continuous variables as means and SDs (mean, SD).

Given the feasibility-oriented nature of this trial, all analyses of secondary and tertiary outcomes are exploratory and are not intended to provide definitive evidence of effectiveness. The inclusion of multiple cognitive, psychosocial, and physical outcomes is justified by the need to obtain preliminary effect size estimates and variability parameters that will guide the selection and prioritization of outcomes in a future fully powered RCT. Therefore, the analyses of secondary and tertiary outcomes will be interpreted as hypothesis-generating. No formal power calculations for effectiveness end points have been performed, as the current sample size is intended to assess procedural feasibility and parameter estimation rather than hypothesis testing.

XSL•FO
RenderX

To assess baseline homogeneity between the control group and the intervention group, chi-square tests will be used for nominal variables and 1-way ANOVA for continuous variables. Expected differences related to "group," "time," and the "group × time" interaction for all dependent variables will be analyzed using a series of multivariate ANOVA. Specifically, repeated-measures ANOVA and multivariate ANOVA will be applied to continuous variables, while the De Castellan test will be used for nominal variables. Bonferroni correction will be applied where appropriate. The assumption of normality for dependent variables will be examined in relation to sphericity, with the Mauchly test used to assess the equality of variances across related group combinations. Cases with missing data will be treated as incomplete and excluded from the analyses. Secondary analyses may be conducted on specific subgroups (eg, by age) and on selected outcomes (eg, individual items from the administered questionnaires).

Regarding statistical power, it is important to note that no previous RCTs have examined the effects of an immersive VR intervention using virtual sailing scenarios in this population. For this reason, this study is designed as a feasibility and exploratory trial. Although this, the planned sample size is consistent with previous feasibility studies using VR-based paradigms, which have typically enrolled small to moderate cohorts [18,46-48]. On this basis, a total sample of 40 participants (20 per arm) was considered appropriate to assess procedural feasibility, obtain reliable estimates of variability, and capture preliminary effect size signals to inform the design for a fully powered RCT.

Ethical Considerations

In accordance with the guidelines of the Declaration of Helsinki [48], written informed consent will be obtained from all participants prior to enrollment and data collection. Participants will be fully informed about the aims and procedures of the study and about their right to withdraw at any time without consequences. The study protocol was approved by the Ethical Committee of the "Azienda Ospedaliero-Universitaria di Cagliari," Italy (Prot. NP/2023/2557, June 16, 2023). All data will be collected and processed in compliance with the European General Data Protection Regulation (GDPR; EU Regulation 2016/679). Data will be anonymized and stored securely, and only the research team will have access to identifiable information. No financial compensation will be provided to participants for taking part in the study.

Results

This project was initially funded in 2022 for a planned duration of 2 years. A 12-month extension was requested and subsequently granted due to delays in the recruitment and implementation phases. Particularly, recruitment was slowed by post–COVID-19 challenges: given the advanced age and MCI of the target population, regular access to the participating health care facilities and engagement in the study were not always feasible. Furthermore, the intervention was originally designed to be delivered through a sailing navigation simulator based on semi-immersive VR technology, which required a dedicated physical environment. Persistent logistical barriers in securing a suitable setting led the research team to develop and adopt a more advanced immersive VR device, which preserved the original study objectives but required additional time before submission and approval by the Ethics Committee.

Recruitment activities began in May 2024 and are planned to be completed by November 2025. As of manuscript submission, 153 participants were recruited and screened regarding inclusion and exclusion criteria. Among them, 39 participants have been enrolled. Preliminary pre-post analyses are scheduled for January 2026; follow-up analyses will be completed by January 2027.

Discussion

Principal Findings

This pilot RCT is expected to provide critical information regarding the feasibility, acceptability, and preliminary signals of effectiveness of an immersive VR-based cognitive remediation program for older adults with MCI. We anticipate that the intervention will be feasible to deliver within routine clinical settings, well tolerated by participants, and associated with improvements in selected cognitive domains, psychosocial functioning, and motor performance. These anticipated findings will guide the refinement of both the intervention and the study procedures for a future fully powered trial.

Over the past decade, the integration of VR as a rehabilitative intervention for older individuals with MCI has shown promising outcomes [20,21,23-26]. Recent meta-analyses conducted on adult individuals with neuropsychiatric disorders, including MCI, demonstrated a positive effect on clinical cognitive outcomes, particularly in the domains of memory, language, and global cognition [25,26]. However, limited evidence is available concerning improvements in clinical and functional outcomes, primarily due to the heterogeneity of variables and assessment tools used, as well as the need for advanced methodological quality in existing studies [20,21,23-27]. Overall, the use of technological innovations and tools such as VR, supported by the existing evidence, can strengthen the rehabilitation experience by increasing engagement and promoting the generalization of acquired skills to real-world contexts, given that the simulated scenarios offer ecologically valid and contextually relevant experiences [18,19,22,49]. While technological tools offer significant support to rehabilitation, they should complement rather than replace the therapeutic relationship, which remains a cornerstone of effective treatment.

The methodological approach adopted in this study reflects a person-centered and recovery-oriented model of rehabilitation, enabling participants to develop their skills and realize a significant improvement in their health and psychological well-being [19,50,51]. Moreover, it adheres to the contemporary framework for the development of complex interventions, wherein ensuring the reproducibility of such interventions constitutes a critical methodological consideration [27]. This study is primarily designed as a feasibility investigation. The findings of this study will also contribute valuable insights that may be translated into clinical practice, particularly concerning

XSL•FO
RenderX

the practicability of an innovative therapeutic approach aimed at preventing cognitive decline in older adults with MCI. These findings may carry important practical implications for health care systems, particularly in light of the substantial economic burden associated with cognitive disorders and preliminary clinical effectiveness.

Findings of the present RCT will be able to guide clinical decisions in terms of both the choice of treatment and personalization of the same. The study will ultimately favor the application of safe, noninvasive, and nonpharmacological interventions for the population of older adults with MCI. The systematic monitoring of the intervention phase will support informed decision-making by policymakers regarding the integration of this evidence into standard clinical care pathways.

Strengths of this study include the use of a rigorous feasibility RCT design, standardized VR-based cognitive remediation procedures, multimodal assessment (cognitive, psychosocial, and motor domains), and alignment with current frameworks for the development of complex interventions. The integration of both mental health professionals and biomechanical engineers ensures high-quality implementation and fidelity.

Limitations may arise from recruitment challenges in an older population with cognitive impairment, potential variability in participants' technological familiarity, and the single-center nature of the trial, which may affect generalizability. As this is a pilot study, it is not powered to detect robust clinical effects; instead, it is designed to generate preliminary signals and inform the next phase of research.

Next Steps for Scale-Up and Implementation

Should the feasibility indicators and preliminary signals of effectiveness be confirmed, the results of this pilot study will directly inform the design of a fully powered RCT. The subsequent phase of research will focus on optimizing recruitment procedures, refining the intervention dosage and session structure, and further standardizing the delivery of the VR-based cognitive remediation protocol across different clinical settings. Implementation-oriented elements, such as training requirements for health care personnel, logistical needs for VR infrastructure, and preliminary assessments of cost-effectiveness, will also be incorporated to evaluate the scalability and sustainability of the intervention within routine care pathways. Moreover, the findings of this study will support the development of an implementation framework aimed at facilitating the integration of immersive VR-based cognitive remediation into prevention and rehabilitation services. Larger-scale trials will ultimately be required to assess long-term outcomes, generalizability, and health-economic impact.

Risk and Benefits

The use of immersive VR may be associated with mild and transient side effects, including dizziness, nausea, headaches, visual fatigue, impaired limb coordination, diminished postural stability, a decreased sense of presence, and potentially inappropriate reactions to real-world environments. Nonetheless, significant adverse effects are not expected, as VR is considered a safe and well-tolerated approach, with an extremely low risk of side effects [18,19,22,49-51].

Dissemination of Results

Findings from this pilot trial will be disseminated through peer-reviewed publications, conference presentations in translational medicine, neuropsychology, and digital health, and stakeholder meetings with local health care professionals and authorities. Additionally, user-friendly summaries will be provided to participants, caregivers, and clinical staff. If the intervention proves feasible and preliminary outcomes are encouraging, dissemination will also target policymakers to support the potential integration of VR-based cognitive interventions into standard care.

Conclusions

This study will yield clinically translatable insights into the feasibility and preliminary clinical potential of an innovative, VR-based, nonpharmacological cognitive remediation program designed to prevent cognitive decline in older adults with MCI. The findings are expected to inform preventive and rehabilitative practices within health care systems, supporting treatment selection and personalization. By promoting safe, noninvasive strategies, this trial aims to advance the integration of VR interventions into routine care for cognitive decline. Future trials with larger samples will be required to establish efficacy and cost-effectiveness.

Acknowledgments

The authors wish to remember their colleague Antonio Crisafulli for his contribution to the submission proposal to the Fondazione di Sardegna call, which funded this project.

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following task was delegated to GenAI tools under full human supervision: proofreading and editing. The GenAI tool used was ChatGPT-5.1. Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by: collective responsibility. Additional note: we used ChatGPT-5.1 for text editing to correct any typographical errors and to ensure that the English text was grammatically accurate, clear, coherent, and appropriate.

XSL•FO
RenderX

Funding

This study is funded by Fondazione di Sardegna: 3-year agreement between the Fondazione di Sardegna and the University of Cagliari, Italy - 2021 call, CUP: F73C22001260007. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the European Union's individual privacy rules, but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: FS, MGC
Methodology: FS
Software: LDN
Formal analysis: FS, MGC
Investigation: AP, VV, MM, SL, GK, RM, SR, DRP, AS, GC, DP, AEN
Resources: FS
Data curation: AP, VV, MM, SL
Writing – original draft: GK, FS
Writing – review & editing: FS, AP, VV, RM, MM, SL, GC, DRP, LDN, DP, SR, AS, AEN, GK, MP, MGC
Supervision: FS
Project administration: FS and AP
Funding acquisition: FS

Conflicts of Interest

LDN is the legal administrator of the society that has developed the CEREBRUM software.

Multimedia Appendix 1
Peer review report 1 from Fondazione di Sardegna (Foundation of Sardinia, Italy).
[PDF File (Adobe PDF File), 99 KB - resprot_v15i1e85089_app1.pdf]

Multimedia Appendix 2
Peer review report 2 from Fondazione di Sardegna (Foundation of Sardinia, Italy).
[PDF File (Adobe PDF File), 131 KB - resprot_v15i1e85089_app2.pdf]

References

1. Salemme S, Lombardo FL, Lacorte E, Sciancalepore F, Remoli G, Bacigalupo I, et al. The prognosis of mild cognitive impairment: a systematic review and meta-analysis. Alzheimers Dement (Amst) 2025;17(1):e70074. [doi: 10.1002/dad2.70074] [Medline: 40078377]
2. International Monetary Fund. Monetary and Capital Markets Department. Global Financial Stability Report, April 2012: The Quest for Lasting Stability. Washington, DC: International Monetary Fund; Apr 18, 2012.
3. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
4. Valdevila Figueira JA, Valdevila Santiesteban R, Carvajal Parra ID, Benenaula Vargas LP, Ramírez A, Leon-Rojas JE, et al. Multimorbidity patterns in dementia and mild cognitive impairment. Front Psychiatry 2024;15:1432848 [FREE Full text] [doi: 10.3389/fpsyt.2024.1432848] [Medline: 39575196]
5. Pu Z, Huang H, Li M, Li H, Shen X, Du L, et al. Screening tools for subjective cognitive decline and mild cognitive impairment based on task-state prefrontal functional connectivity: a functional near-infrared spectroscopy study. Neuroimage 2025 Apr 15;310:121130 [FREE Full text] [doi: 10.1016/j.neuroimage.2025.121130] [Medline: 40058532]
6. Gopalakrishnan P, Tiwari S, Nagaraja R, Krishnan G. Quality of life in persons with mild cognitive impairment: a systematic review and meta-analysis. Dement Neuropsychol 2024;18:e20230093. [doi: 10.1590/1980-5764-DN-2023-0093] [Medline: 39193465]
7. Horizon 2020 - Work Programme 2018-2020. Health, demographic change and wellbeing. European Commission. URL: https://ec.europa.eu/research/participants/data/ref/h2020/wp/2018-2020/main/h2020-wp1820-health_en.pdf [accessed 2025-12-25]

XSL•FO
RenderX

8. World Health Organization. Risk Reduction of Cognitive Decline and Dementia. Geneva, Switzerland: World Health Organization; 2019:978-992.
9. Gavelin HM, Lampit A, Hallock H, Sabatés J, Bahar-Fuchs A. Cognition-oriented treatments for older adults: a systematic overview of systematic reviews. Neuropsychol Rev 2020;30(2):167-193 [FREE Full text] [doi: 10.1007/s11065-020-09434-8] [Medline: 32266520]
10. Webb SL, Birney DP, Loh V, Walker S, Lampit A, Bahar-Fuchs A. Cognition-oriented treatments for older adults: a systematic review of the influence of depression and self-efficacy individual differences factors. Neuropsychol Rehabil 2022;32(6):1193-1229. [doi: 10.1080/09602011.2020.1869567] [Medline: 33509053]
11. Carta MG, Cossu G, Pintus E, Zaccheddu R, Callia O, Conti G, et al. Moderate exercise improves cognitive function in healthy elderly people: results of a randomized controlled trial. Clin Pract Epidemiol Ment Health 2021;17:75-80 [FREE Full text] [doi: 10.2174/1745017902117010075] [Medline: 34733346]
12. Mura G, Sancassiani F, Migliaccio GM, Collu G, Carta MG. The association between different kinds of exercise and quality of life in the long term. Results of a randomized controlled trial on the elderly. Clin Pract Epidemiol Ment Health 2014;10:36-41. [doi: 10.2174/1745017901410010036] [Medline: 24707314]
13. Hardoy MC, Seruis ML, Floris F, Sancassiani F, Moro MF, Mellino G, et al. Benefits of exercise with mini tennis in intellectual disabilities: effects on body image and psychopathology. Clin Pract Epidemiol Ment Health 2011;7:157-160. [doi: 10.2174/1745017901107010157] [Medline: 22016751]
14. Tang L, Feng Z, Liu W, Zhang W, Liu Y, Xiong N, et al. Risk factors for mortality in individuals with dementia: a systematic review and meta-analysis. Arch Gerontol Geriatr 2025;137:105951. [doi: 10.1016/j.archger.2025.105951] [Medline: 40683148]
15. Yun S, Ryu S. The effects of cognitive-based interventions in older adults: a systematic review and meta-analysis. Iran J Public Health 2022;51(1):1-11 [FREE Full text] [doi: 10.18502/ijph.v51i1.8286] [Medline: 35223620]
16. Mendes L, Oliveira J, Barbosa F, Castelo-Branco M. A conceptual view of cognitive intervention in older adults with and without cognitive decline-a systemic review. Front Aging 2022;3:844725 [FREE Full text] [doi: 10.3389/fragi.2022.844725] [Medline: 35821828]
17. McGurk SR, Mueser KT. Introduction to special issue on cognitive remediation. Psychiatr Rehabil J 2017;40(1):1-3. [doi: 10.1037/prj0000263] [Medline: 28368178]
18. Perra A, Riccardo CL, De Lorenzo V, De Marco E, Di Natale L, Kurotschka PK, et al. Fully immersive virtual reality-based cognitive remediation for adults with psychosocial disabilities: a systematic scoping review of methods intervention gaps and meta-analysis of published effectiveness studies. Int J Environ Res Public Health 2023;20(2) [FREE Full text] [doi: 10.3390/ijerph20021527] [Medline: 36674283]
19. Perra A, De Lorenzo V, Zaccheddu R, Locci A, Piludu F, Preti A, et al. Cognitive remediation virtual reality tool a recovery-oriented project for people with bipolar disorder: protocol of a feasibility randomized clinical trial. Clin Pract Epidemiol Ment Health 2022;18:e174501792208220 [FREE Full text] [doi: 10.2174/17450179-v18-e2208220] [Medline: 37274852]
20. Tortora C, Di Crosta A, La Malva P, Prete G, Ceccato I, Mammarella N, et al. Virtual reality and cognitive rehabilitation for older adults with mild cognitive impairment: a systematic review. Ageing Res Rev 2024;93:102146 [FREE Full text] [doi: 10.1016/j.arr.2023.102146] [Medline: 38036103]
21. Yang Q, Zhang L, Chang F, Yang H, Chen B, Liu Z. Virtual reality interventions for older adults with mild cognitive impairment: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2025;27:e59195 [FREE Full text] [doi: 10.2196/59195] [Medline: 39793970]
22. Riva G, Serino S. Virtual reality in the assessment, understanding and treatment of mental health disorders. J Clin Med 2020;9(11):3434 [FREE Full text] [doi: 10.3390/jcm9113434] [Medline: 33114623]
23. Ren Y, Wang Q, Liu H, Wang G, Lu A. Effects of immersive and non-immersive virtual reality-based rehabilitation training on cognition, motor function, and daily functioning in patients with mild cognitive impairment or dementia: a systematic review and meta-analysis. Clin Rehabil 2024;38(3):305-321. [doi: 10.1177/02692155231213476] [Medline: 38289618]
24. Ramalho A, Duarte-Mendes P, Paulo R, Serrano J, Petrica J. Crossing the digital frontier: are older adults ready for virtual reality workouts? Front Public Health 2024;12:1324004 [FREE Full text] [doi: 10.3389/fpubh.2024.1324004] [Medline: 38389949]
25. Li X, Zhang Y, Tang L, Ye L, Tang M. Effects of virtual reality-based interventions on cognitive function, emotional state, and quality of life in patients with mild cognitive impairment: a meta-analysis. Front Neurol 2025;16 [FREE Full text] [doi: 10.3389/fneur.2025.1496382]
26. Du Q, Wei Y, Ma Y, Liu C, Du S, Zhang Q, et al. Efficacy of virtual reality-based interventions on cognitive function in patients with neuropsychiatric disorders: systematic review and meta-analysis of randomized controlled trials. JMIR Serious Games 2025;13:e67501 [FREE Full text] [doi: 10.2196/67501] [Medline: 40341171]
27. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]

XSL•FO
RenderX

28. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot Feasibility Stud 2016;2:64 [FREE Full text] [doi: 10.1186/s40814-016-0105-8] [Medline: 27965879]
29. Pigliautile M, Ricci M, Mioshi E, Ercolani S, Mangialasche F, Monastero R, et al. Validation study of the Italian Addenbrooke's Cognitive Examination Revised in a young-old and old-old population. Dement Geriatr Cogn Disord 2011;32(5):301-307. [doi: 10.1159/000334657] [Medline: 22262124]
30. Brown P, Spronck P, Powell W. The simulator sickness questionnaire, and the erroneous zero baseline assumption. Front Virtual Real 2022;3. [doi: 10.3389/frvir.2022.945800]
31. Llinàs-Reglà J, Vilalta-Franch J, López-Pousa S, Calvó-Perxas L, Torrents Rodas D, Garre-Olmo J. The Trail Making Test. Assessment 2017;24(2):183-196. [doi: 10.1177/1073191115602552] [Medline: 26318386]
32. Bowden SC, Petrauskas VM, Bardenhagen FJ, Meade CE, Simpson LC. Exploring the dimensionality of Digit Span. Assessment 2012;20(2):188-198. [doi: 10.1177/1073191112457016]
33. Grapperon J, Delage M. [Stroop test and schizophrenia]. Encephale 1999;25(1):50-58. [Medline: 10205734]
34. Caffarra P, Vezzadini G, Dieci F, Zonato F, Venneri A. Rey-Osterrieth complex figure: normative values in an Italian population sample. Neurol Sci 2002;22(6):443-447. [doi: 10.1007/s100720200003] [Medline: 11976975]
35. Iavarone A, Ronga B, Pellegrino L, Loré E, Vitaliano S, Galeone F, et al. The Frontal Assessment Battery (FAB): normative data from an Italian sample and performances of patients with Alzheimer's disease and frontotemporal dementia. Funct Neurol 2004;19(3):191-195. [Medline: 15595714]
36. Brand T, Wagener KC. Characteristics, advantages, and limits of matrix tests. HNO 2016;65(3):182-188 [FREE Full text] [doi: 10.1007/s00106-016-0224-9]
37. Bell-Sprinkel TL, Boone KB, Miora D, Cottingham M, Victor T, Ziegler E, et al. Re-examination of the Rey Word Recognition Test. Clin Neuropsychol 2013;27(3):516-527. [doi: 10.1080/13854046.2012.744853] [Medline: 23216255]
38. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
39. Apolone G, Mosconi P, Quattrociocchi L, Gianicolo E, Groth N, Ware JE. Questionario sullo stato di salute SF-12 Versione Italiana. Milan, Italy: Guerini Associati; 2005.
40. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
41. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
42. Cossu G, Agus M, Atzori L, Aviles Gonzales CI, Minerba L, Ferreli C, et al. Principal component analysis of the Social and Behavioral Rhythms Scale in elderly. J Public Health Res 2021;11(1) [FREE Full text] [doi: 10.4081/jphr.2021.2546] [Medline: 34619859]
43. Cossu G, Loi E, Carta MG, Bramanti A. The Physical Body Experiences Questionnaire Simplified for Active Aging (PBE-QAG). Clin Pract Epidemiol Ment Health 2018;14:70-77 [FREE Full text] [doi: 10.2174/1745017901814010070] [Medline: 29643930]
44. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nutr 2006;9(6):755-762. [doi: 10.1079/phn2005898] [Medline: 16925881]
45. Aghapour M, Affenzeller N, Peham C, Lutonsky C, Tichy A, Bockstahler B. Effect of vision and surface slope on postural sway in healthy adults: a prospective cohort study. Life (Basel) 2024;14(2):227 [FREE Full text] [doi: 10.3390/life14020227] [Medline: 38398736]
46. Paillard T, Pau M, Noé F, González LM. Rehabilitation and improvement of the postural function. Biomed Res Int 2015;2015:703679 [FREE Full text] [doi: 10.1155/2015/703679] [Medline: 26640793]
47. Caronni A, Sterpi I, Antoniotti P, Aristidou E, Nicolaci F, Picardi M, et al. Criterion validity of the instrumented Timed Up and Go test: a partial least square regression study. Gait Posture 2018;61:287-293. [doi: 10.1016/j.gaitpost.2018.01.015] [Medline: 29413799]
48. World Medical Association Declaration of Helsinki ethical principles for medical research involving human subjects. JAMA. 2013. URL: https://doi.org/10.1001/jama.2013.281053 [accessed 2025-12-25]
49. Manzanares A, Camblor ?, Romero-Arenas S, Segado F, Gil-Arias A. Adapted sailing teaching methodology using vsail-trainer simulator as rehabilitation therapy. A feasibility study. J Spinal Cord Med 2024 Nov;47(6):960-967. [doi: 10.1080/10790268.2023.2228585] [Medline: 37534919]
50. Primavera D, Aviles Gonzalez C, Perra A, Kalcev G, Cantone E, Cossu G, et al. Virtual reality cognitive remediation in older adults with bipolar disorder: the effects on cognitive performance and depression in a feasibility randomized controlled trial. Healthcare (Basel) 2024;12(17):1753 [FREE Full text] [doi: 10.3390/healthcare12171753] [Medline: 39273777]
51. Perra A, Carta MG, Primavera D, Cossu G, Locci A, Zaccheddu R, et al. An integrated cognitive remediation and recovery-oriented program for individuals with bipolar disorder using a virtual reality-based intervention: 6- and 12-month cognitive outcomes from a randomized feasibility trial. Behav Sci (Basel) 2025;15(4):452 [FREE Full text] [doi: 10.3390/bs15040452] [Medline: 40282074]

XSL•FO
RenderX

Abbreviations

ACE-R: Addenbrooke's Cognitive Examination-Revised
CONSORT: Consolidated Standards of Reporting Trials
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
RCT: randomized controlled trial
QoL: quality of life
VR: virtual reality
VSail: Virtual Sail 3D

Edited by A Schwartz; submitted 30.Sep.2025; this is a non–peer-reviewed article;accepted 30.Nov.2025; published 15.Jan.2026.

Please cite as:
Sancassiani F, Perra A, Vacca V, Montisci R, Monni M, Lorrai S, Cossu G, Petretto DR, di Natale L, Primavera D, Redolfi S, Scuteri A, Nardi AE, Kalcev G, Pau M, Carta MG
Enhancing Cognitive Functions in Older Adults With Mild Cognitive Impairment via Virtual Sail 3D: Protocol for a Feasibility Randomized Controlled Trial
JMIR Res Protoc 2026;15:e85089
URL: https://www.researchprotocols.org/2026/1/e85089
doi:10.2196/85089
PMID:

©Federica Sancassiani, Alessandra Perra, Veronica Vacca, Roberta Montisci, Mariarita Monni, Stefano Lorrai, Giulia Cossu, Donatella Rita Petretto, Lorenzo di Natale, Diego Primavera, Stefania Redolfi, Angelo Scuteri, Antonio Egidio Nardi, Goce Kalcev, Massimiliano Pau, Mauro Giovanni Carta. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Smartphone-Based Psychological Intervention for Nonsuicidal Self-Injury (Kalmer App): Protocol for a Multicenter Double-Blind Randomized Controlled Trial

Anna Julià[1,2,3], MSc, MPGS; Irene Jaén[4], PhD; Mónica Conesa Giménez[5], MPGS; Azucena García-Palacios[5,6], PhD; Juan C Pascual[7,8], MD, PhD; Anna Sintes Estévez[9], ClinPsy, PhD; Anaís Lara[10], ClinPsy; Iria Méndez[11], MD, PhD; Soledad Romero[8,12,13], MD, PhD; Joaquim Puntí[8,14,15], ClinPsy; Joaquim Soler[2,8,16], ClinPsy, PhD; Jordi Solé-Casals[17,18], PhD; Marina López-Solà[13,19,20], PhD; Daniel Vega[1,3], ClinPsy, PhD

[1]Fundació Sanitària d'Igualada, Igualada, Spain
[2]Department of Psychiatry and Forensic Medicine, School of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain
[3]Department of Psychiatry, Hospital Universitari d'Igualada, Consorci Sanitari de l'Anoia, Igualada, Spain
[4]Department of Developmental, Educational, Social and Methodology Psychology, Universitat Jaume I, Castellón de la Plana, Spain
[5]Department of Basic and Clinical Psychology and Psychobiology, Universitat Jaume I, Castellón de la Plana, Spain
[6]CIBER de Physiopathology of Obesity and Nutrition (CIBEROBN), Instituto de Salud Carlos III, Madrid, Spain
[7]Mental Health and Psychiatry Department, Vic Hospital Consortium, Vic, Spain
[8]Centro de Investigación Biomédica en Red de Salud Mental, Instituto de Salud Carlos III, Madrid, Madrid, Spain
[9]Servicio de Psiquiatría y Psicología, Hospital Sant Joan de Déu Barcelona, Barcelona, Catalonia, Spain
[10]Servicio de Salud Mental Infanto Juvenil, ALTHAIA, Xarxa Assistencial Universitaria de Manresa, Manresa, Spain
[11]Servei de Psiquiatria & Psicologia Infantil Juvenil, Hospital Universitari Mútua de Terrassa, Fundació Recerca Mútua de Terrassa, Terrassa, Spain
[12]Department of Child and Adolescent Psychiatry and Psychology, 2021SGR01319, Institute of Neuroscience, Hospital Clínic de Barcelona, Barcelona, Spain
[13]Neuropsychology and Neuroimaging Group, Clinical and Experimental Neurosciences Area, IDIBAPS, Barcelona, Spain
[14]Salut Mental, Corporació Sanitària Parc Taulí de Sabadell, Sabadell, Spain
[15]Department of Clinical and Health Psychology, Faculty of Psychology and Speech Therapy, Universitat Autònoma de Barcelona, Barcelona, Spain
[16]Department of Psychiatry, Hospital de la Santa Creu i Sant Pau, Institut d'Investigació Biomèdica- Sant Pau (IIB-SANT PAU), Barcelona, Spain
[17]Data and Signal Processing Group, Universitat de Vic - Universitat Central de Catalunya, Vic, Spain
[18]Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom
[19]Department of Medicine, Unit of Psychological Medicine, Universitat de Barcelona, Barcelona, Spain
[20]Institute of Neurosciences, Universitat de Barcelona, Barcelona, Catalonia, Spain

Corresponding Author:
Daniel Vega, ClinPsy, PhD
Department of Psychiatry
Hospital Universitari d'Igualada
Consorci Sanitari de l'Anoia
Av. Catalunya, 11
Igualada
Spain
Phone: 34 938 075 500
Email: daniel.vegamo@gmail.com

Abstract

Background: Nonsuicidal self-injury (NSSI), defined as the deliberate, self-inflicted damage of body tissue without suicidal intent, is increasingly prevalent among adolescents and young adults and poses a major public health concern. Current treatments are often costly, difficult to access, and not tailored to the specific needs of young people. Mobile health (mHealth) interventions represent a promising avenue for scalable, accessible, and cost-effective support for NSSI, especially when combined with real-time assessments and personalized treatment strategies.

XSL•FO
RenderX

Objective: This randomized controlled trial will evaluate the effectiveness of Kalmer, a brief app-based intervention for reducing NSSI and improving emotional well-being. The study has 2 aims: (1) to evaluate a newly developed app-based intervention for adolescents and young adults engaging in NSSI and (2) to assess predictors of treatment outcomes for this app-based intervention. We hypothesize that participants receiving a mobile app–based brief intervention specifically tailored to address NSSI will show a greater reduction in NSSI frequency at the end of treatment and at follow-up than participants receiving a nonspecific app-based intervention. In this paper, we present our study protocol.

Methods: This 2-arm randomized controlled trial, lasting 6 weeks, will include 240 participants aged 14 to 24 years who engage in NSSI. The intervention app, Kalmer, was developed through iterative consultation with clinical and research experts and guided by survey results and evidence-based frameworks such as dialectical behavior therapy and cognitive behavioral therapy. The intervention will include 5 core components: distress tolerance, emotion regulation, mindfulness and self-compassion, interpersonal regulation, and problem-focused coping. The app will deliver multimedia-based ecological momentary interventions triggered by real-time ecological momentary assessments to tailor support to users' current contexts and needs. Participants are randomized to receive either the Kalmer intervention or a psychoeducational control app. The primary outcome will be NSSI frequency, assessed through self-report at baseline, 6 weeks after intervention, and at 1-month and 3-month follow-ups.

Results: Ethics approval was obtained in December 2022. As of January 2026, a total of 145 participants had consented to participate in the study and completed baseline assessments. Preliminary data show high app engagement and acceptability and positive user feedback regarding app usability and content. Recruitment is ongoing.

Conclusions: This study will provide evidence on the effectiveness of a mobile app–based intervention for NSSI and will explore potential mechanisms of change, supporting the development of accessible digital mental health tools for adolescents and young adults.

Trial Registration: International Standard Randomized Controlled Trial Number Registry ISRCTN63093907; https://www.isrctn.com/ISRCTN63093907

International Registered Report Identifier (IRRID): DERR1-10.2196/86413

(JMIR Res Protoc 2026;15:e86413) doi:10.2196/86413

KEYWORDS

nonsuicidal self-injury; adolescents and young people; smartphone; randomized controlled trial; RCT; intervention; ecological momentary intervention; app; cognitive behavioral therapy; CBT; dialectical behavior therapy; DBT

Introduction

Nonsuicidal self-injury (NSSI) is the deliberate destruction of one's own body tissue in the absence of suicidal intent and is a growing public health concern, particularly among adolescents and young adults [1,2]. NSSI imposes a substantial burden not only on individuals and their families but also on health care systems and society at large. Prevalence rates suggest that 17.2% of adolescents and 13.4% of young adults in community samples report at least one episode of NSSI, with rates reaching up to 58% in clinical adolescent populations [3,4]. NSSI is associated with potentially negative consequences, including impaired interpersonal, academic, and daily functioning, and is one of the strongest predictors of suicidal behavior [5]. In recognition of its severity, the World Health Organization has identified NSSI as one of the top 5 major health threats to adolescents, and it has been included in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, as a condition requiring further research.

NSSI often begins during early adolescence and is frequently associated with chronic psychosocial stress, peer rejection, or family-related neglect [6]. Individuals who engage in NSSI typically report greater psychological distress than their peers, regardless of clinical status [4]. In clinical settings, NSSI is commonly observed in the context of broader internalizing and externalizing conditions such as depression, borderline personality disorder (BPD), and eating disorders [7]. Although NSSI is primarily used to regulate negative emotions, it serves a range of functions, both intrapersonal (eg, to relieve distress) and interpersonal (eg, to communicate needs or influence relationships) [8]. According to the 4-factor model of NSSI, these functions can be either positively or negatively reinforced and often co-occur within the same individual [9].

Despite progress in understanding NSSI, significant gaps remain, particularly in the identification of effective, scalable interventions and the mechanisms underlying treatment response. Established treatments such as dialectical behavior therapy (DBT), cognitive behavioral therapy (CBT), and mentalization-based therapy have shown efficacy in reducing NSSI, especially among young adults with BPD traits, but are resource intensive and limited in accessibility [10-12]. Many adolescents and young adults with NSSI face substantial barriers to accessing care, including stigma, a lack of specialized clinicians, and reluctance to seek face-to-face treatment [13].

In this context, digital mental health interventions, particularly mobile apps, represent a promising avenue for improving access to care. Mobile apps can deliver evidence-based support in a self-guided, scalable, and cost-effective format [14]. Young people report high levels of interest in and engagement with digital tools for mental health. However, most app-based interventions currently available either do not specifically target NSSI or have limited empirical support [15]. Furthermore, very few studies have explored their utility in nonclinical populations, despite the growing prevalence of NSSI in these groups.

XSL•FO
RenderX

Importantly, there is also limited knowledge regarding individual-level predictors of treatment response, which hinders the development of personalized digital interventions.

To address these gaps, we have developed Kalmer, a brief mobile app–based intervention designed specifically for adolescents and young adults who engage in NSSI. The app has been developed through a structured process that includes expert consultations and survey-based feedback, ensuring that the intervention is grounded in clinical and empirical knowledge. Kalmer integrates ecological momentary assessments (EMAs) to monitor users' emotional states in real time and trigger tailored ecological momentary interventions (EMIs) delivered through multimedia formats (eg, video, audio, text, and image-based content).

This paper presents the protocol for a randomized controlled trial (RCT) to evaluate the feasibility, acceptability, and preliminary efficacy of the Kalmer intervention. The study will include both clinical and subclinical participants aged 14 to 24 years with a history of NSSI. The primary aim is to develop Kalmer and assess whether it reduces NSSI frequency. The secondary aim is to identify behavioral predictors of treatment response using self-report and real-time app data. This research has the potential to inform the development of accessible, personalized, and evidence-based digital interventions for young people who engage in self-injury.

Methods

Study Design and Intervention Arms

Overview

This study is a multicenter, double-blinded RCT evaluating the effectiveness of a brief mobile app–based intervention for NSSI in adolescents and young adults. The study was registered with the International Standard Randomized Controlled Trial Number registry. Participants are being randomized to either an NSSI-specific intervention app (intervention 1) or a nonspecific control app (intervention 2).

Intervention 1: NSSI-Specific Intervention App

Development and Design of Kalmer

The NSSI-specific intervention app, Kalmer, has been designed as a brief app-based program targeting NSSI in adolescents and young adults. It combines evidence-based components from CBT and DBT, with the aim of expanding users' repertoire of coping strategies to improve well-being, distress tolerance (DT), autonomy, decision-making, problem-solving, sense of meaning, and effectiveness in social relationships.

The intervention content was developed through an iterative process involving both clinical and academic experts in NSSI, digital mental health, and adolescent psychopathology. Experts participated in structured meetings and completed targeted surveys to validate and prioritize the core therapeutic elements. This process led to the definition of 5 therapeutic domains that structured the intervention: distress tolerance (DT); mindfulness and self-compassion (M); interpersonal regulation (IR); and problem-focused coping (PC). A schematic representation of the EMA-EMI sequence is shown in Figure 1 (left panel).

XSL•FO
RenderX

Figure 1. Screenshots of the Kalmer app. (A) Initial ecological momentary assessment (EMA), in this example rating the extent to which the participants feel distracted in the present moment. (B) Based on the EMA response, the app then delivers a tailored ecological momentary intervention (EMI) in a multimedia format, such as audio, video, or an infographic. (C) An EMA is used to collect feedback on the perceived usefulness of the exercise. (D) A screenshot of the main menu of the Kalmer app. T1: time point 1; T2: time point 2; T3: time point 3.

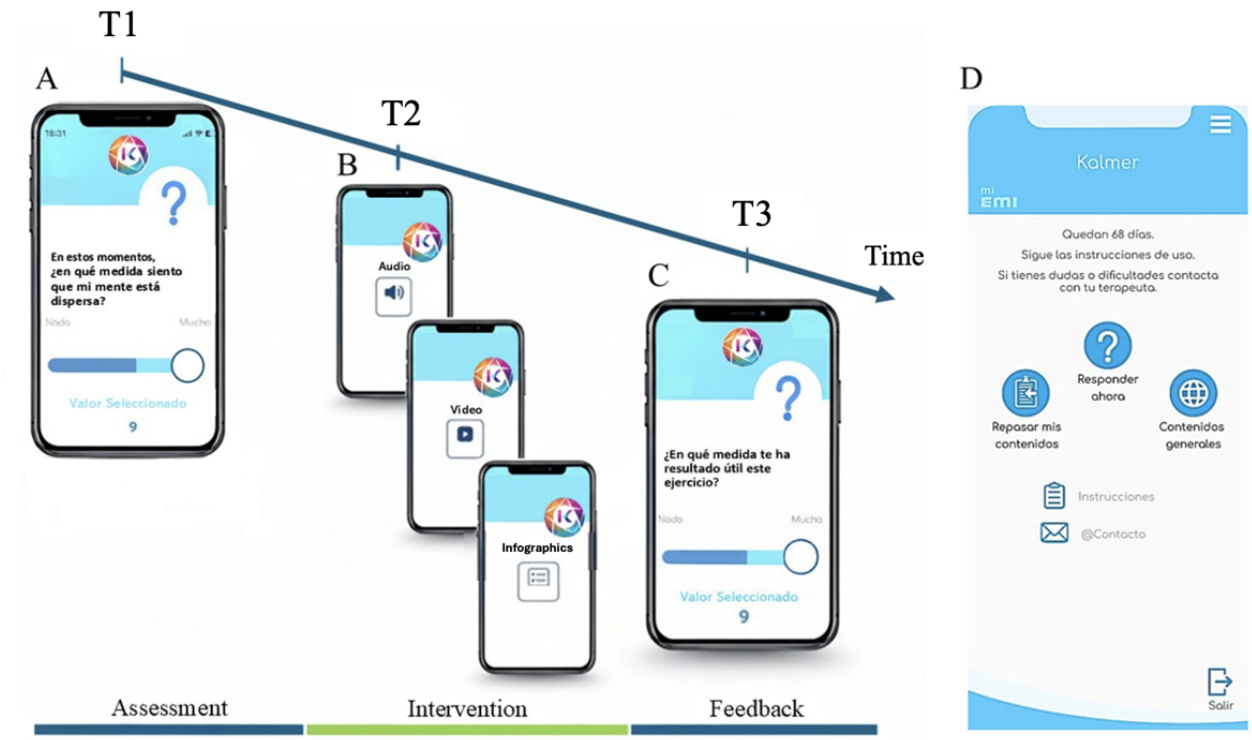

Ecological Momentary Assessments

EMAs were constructed to gather daily, real-time information on participants' emotional states, perceived social support, NSSI behaviors, and coping strategies aligned with the 5 therapeutic domains. The assessments were derived from validated psychometric instruments, including the 15-item Five Facet Mindfulness Questionnaire [16,17], the 28-item Coping Orientation to Problems Experienced questionnaire [18,19], the 12-item Interpersonal Support Evaluation List [20,21], the Difficulties in ER Scale (DERS) [22,23], and the Inventory of Statements About Self-Injury (ISAS) [24].

Table 1 shows 2 types of EMAs that were implemented. On-demand EMAs were accessible at any time and provided participants with an immediate, tailored EMI upon completion. Scheduled EMAs were delivered daily at 6:00 PM, with a reminder at 7:00 PM, to ensure consistent engagement. The question sets varied slightly between the on-demand and the scheduled EMAs. Additional items regarding social media use were included in the scheduled EMAs, given their relevance in this age group and their established association with NSSI behaviors [25,26].

XSL•FO
RenderX

Table 1. Scheduled and on-demand ecological momentary assessment (EMA) questions.

Component	On-demand EMA	Scheduled EMA
Emotion regulation	At this moment, to what extent do my emotions seem overwhelming to me?	At this moment, to what extent do my emotions seem overwhelming to me?
Interpersonal regulation I	During the day today, to what extent am I spending more time than usual on social media because I feel lonely and/or rejected?	At this moment, to what extent do I feel lonely and/or rejected?
Interpersonal regulation II	N/A[a]	During the day today, how many hours have I spent on social media?
Interpersonal regulation III	N/A	During the day today, to what extent have I spent more time than usual on social media because I have felt lonely and/or rejected?
Mindfulness	At this moment, to what extent do I feel that my mind is wandering?	At this moment, to what extent do I feel that my mind is wandering?
Problem-focused coping	At this moment, to what extent do I feel able to cope with my problems?	At this moment, to what extent do I feel able to cope with my problems?
Distress tolerance	At this moment, to what extent do I think about harming myself?	At this moment, to what extent do I think about harming myself?

[a]N/A: not applicable.

Ecological Momentary Interventions

EMAs dynamically trigger EMIs tailored to participants' real-time needs. EMIs are delivered through multimedia formats, including videos, images, infographics, audio clips, and external links. Each EMI module features a short video with psychoeducation, examples, and simulated testimonies, followed by exercises designed for real-world application.

The first week of app use includes daily introductory videos presenting each therapeutic component and encouraging adherence. Both scheduled and on-demand EMAs and EMIs will be available throughout the intervention period.

The DT domain was prioritized because of its strong association with NSSI behaviors [27]. DT-focused EMIs aimed to help users withstand intense emotional states without resorting to NSSI, emphasizing skills for tolerating, but not necessarily reducing, the intensity of distress until it subsides.

The ER component focused on helping participants understand and modulate strong positive and negative emotions, as emotional cascades are a known trigger for self-injury [27]. The mindfulness and self-compassion domain equipped users with skills to remain present and attuned to their emotions, thereby enhancing the application of other coping strategies.

The interpersonal regulation and problem-focused coping domains address relational difficulties and life challenges that are particularly salient during adolescence and young adulthood [28]. Strategies within these domains aim to mitigate feelings of rejection and abandonment, both of which are linked to NSSI [6,29].

Multimedia Content and Prioritization

A total of 32 EMI videos, 50 images, 42 infographics, 8 audio files, and 6 YouTube links were created to explain therapeutic strategies, provide psychoeducation, and encourage skill application. EMIs were prioritized based on expert judgment: when the user reaches the specific scores detailed in Table 2, an EMI will be activated to offer the intervention objects linked with the score and its component. If multiple components warranted intervention, the app delivered the EMI corresponding to the most maladaptive state at that moment. The decision rules for EMI selection and prioritization are detailed in Table 2.

XSL•FO
RenderX

Table 2. Ecological momentary intervention (EMI) component order.

EMA[a] score	EMI component
DT[b] between 8 and 10	DT content related to doing something else to avoid self-injury, as "STOP[c]," "opposite action," and "distraction" techniques, but focused on dealing with NSSI[d] behaviors
DT between 6 and 7	DT content related to activities or strategies that can be used to delay NSSI behaviors
ER[e] between 8 and 10	ER content related to doing something else or different, as "STOP," "opposite action," and "distraction" techniques
IR[f] between 8 and 10	IR content related to not acting impulsively, contact with support network and emotion management of difficult situations.
M[g] between 5 and 10	M content related to practicing mindfulness with different strategies.
ER between 6 and 7	ER content related to reevaluation and relaxation.
IR between 6 and 7	IR content related to how to connect with others
DT between 4 and 5	DT content related to things to do to protect yourself in the event of NSSI behaviors appearing, such as safety plans.
PC[h] between 0-3	PC content related to the steps to solve problems and to make decisions.
ER between 4 and 5	ER content related to normalizing the emotions felt in the moment.
IR between 4 and 5	IR content related to remind the user to connect with people and that is not alone.
PC between 4 and 6	PC content related to short reminders of how to deal with problems.
SM[i] between 7 and 10	SM content to explain how social media do not help with rejection feelings.

[a]EMA: ecological momentary assessment.
[b]DT: distress tolerance.
[c]STOP: stop, take a breath, observe, and proceed mindfully.
[d]NSSI: nonsuicidal self-injury.
[e]ER: emotion regulation.
[f]IR: interpersonal regulation.
[g]M: mindfulness.
[h]PC: problem-focused coping.
[i]SM: social media.

The app also featured 2 content libraries. The "Review My Contents" section served as a repository of all previously received EMIs, allowing users to review and practice the strategies at any time. The "General Contents" section provided crisis hotline numbers and links to reputable websites related to self-injury and mental health. Figure 1D shows the 3 principal options displayed in the app: "Answer now," "Review my content," and "General content."

Youth Advisory Board

A youth advisory board composed of 9 young adults with lived experience of NSSI provided systematic feedback on the app's functionality, design, and overall user experience. Their perspectives informed the development of the final version of the Kalmer app.

Intervention 2: Nonspecific Control App

Participants assigned to intervention 2 group use the same Kalmer app interface, which adheres to the same EMA schedule as those in intervention 1. However, this version delivers static, nonpersonalized content that is not tailored to individual EMA responses. Specifically, the app comprises 2 psychoeducational modules: (1) an introductory module outlining the app-based intervention and providing general information on NSSI and (2) a module focused on emotional well-being, healthy lifestyle behaviors, and strategies for managing anxiety.

For both intervention 1 and intervention 2, app use metrics (eg, session completion rates and time spent within the app) are being collected continuously. In either condition, any participant reporting recurrent self-injury or heightened clinical risk will prompt the research team to notify the participants' clinical care team to initiate an appropriate safety protocol. Additionally, both versions of the app include a "General contents" section containing the national suicidal behavior hotline number and links to validated online resources pertaining to self-injury and mental health.

Participants, Eligibility Criteria, and Recruitment

Recruitment is ongoing, and we aim to recruit a total of 240 participants aged 14 to 24 years. As shown in Table 3, all participants will have engaged in at least 5 episodes of NSSI in the past year and will have access to a smartphone. They must also provide informed consent; parental consent is required for minors. Exclusion criteria for all participants include a current psychiatric crisis or inability to complete study procedures.

XSL•FO
RenderX

Table 3. Inclusion and exclusion criteria for the study groups.

Inclusion and exclusion criteria	Study groups	
	Clinical	Subclinical
Aged 14-24 y	✓	✓
Providing informed consent online	✓	✓
Owning a smartphone	✓	✓
Having engaged in NSSI[a] on ≥5 different days in the last 12 months	✓	✓
Co-occurring with any other diagnosis	✓	
Currently receiving psychological or psychiatric treatment	✓	
Absence of important clinical instability	✓	✓

[a]NSSI: nonsuicidal self-injury.

Participants with NSSI are being recruited from clinical sites (ie, clinical group, n=120) and from the community (ie, subclinical group, n=120). Participants categorized as clinical are receiving treatment from one of the mental health service sites involved in this study, whereas subclinical participants are not receiving treatment from a public mental health unit. Recruitment is being carried out through multiple channels including clinics, schools, social media, and mental health organizations. A dedicated study website provides information, facilitates consent, and guides participants through the screening and assessment process.

Screening, Baseline, and Follow-Up Assessments

Interested individuals are being contacted by the study's research assistant and invited to attend a screening interview, during which they receive detailed information about the study, undergo an eligibility assessment for participation, and provide informed consent in electronic or paper form. For participants younger than 18 years, parental consent is also required. Upon consent, participants complete the online assessment protocol consisting of sociodemographic information and validated self-report questionnaires. Participants are invited to complete online questionnaires at multiple time points throughout the study to assess primary and secondary outcomes and momentary data during the app-based intervention. Figure 2 provides an overview of the instruments administered at each assessment point (T0-T3), including baseline, treatment start, postintervention, and 3-month follow-up evaluations.

XSL•FO
RenderX

Figure 2. Schedule of self-report measures administered across study time points (T0-T3). BEAQ: Brief Experiential Avoidance Questionnaire; BSL-23: Borderline Symptom List–23; DASS-21: Depression Anxiety Stress Scales; DERS-18: brief version of the Difficulties in Emotion Regulation Scale; DoM: Delusion of Me; DTS: Dialectical Thinking Scale; ISAS: Inventory of Statements About Self-Injury; MPSS: Multidimensional Scale of Perceived Social Support; NSSI: nonsuicidal self-injury; NSSIDS: nonsuicidal self-injury disorder scale; PANAS: Positive and Negative Affect Schedule; PHI: Pemberton Happiness Index; SUS: System Usability Scale.

Self-report measures	Time points			
	Screening	Intervention		Follow-up
	T0	T1	T2	T3
Recruitment				
Consent and sociodemographic data	✓			
Screening interview	✓			
Interventions				
Treatment 1		⟼		
Treatment 2		⟼		
Assessment				
ISAS	✓	✓	✓	✓
NSSIDS	✓			
DERS-18	✓		✓	✓
BEAQ	✓		✓	✓
DASS-21	✓		✓	✓
BSL-23	✓		✓	✓
MPSS	✓		✓	✓
DTS	✓		✓	✓
DoM	✓		✓	✓
PHI	✓			
PANAS	✓		✓	✓
SUS			✓	
Momentary				
Engagement in NSSI		⟼		
Social support		⟼		
Emotional state		⟼		

Description of Measures

Screening Interview

Questionnaires at the baseline interview assessment include the validated Spanish version of the Patient Health Questionnaire-9 [30-32]. Selected items from the Columbia-Suicide Severity Rating Scale [33] are also used. Finally, the clinician-rated severity of NSSI, a 5-point scale developed to assess the severity of NSSI behaviors experienced by an individual in the past year, is being completed [34].

Primary Outcome

The primary outcome is the frequency of NSSI behaviors, assessed using the ISAS [24] at the predefined study time points. At baseline (T0), the ISAS assesses the lifetime frequency and functions of NSSI behaviors [24]. At follow-up assessments, a modified version of ISAS is used to inquire about NSSI behavior in the last month. The Spanish version of the ISAS is used in this study, demonstrating strong internal consistency for both the interpersonal and intrapersonal subscales (Cronbach α=0.87 and 0.90, respectively) [35].

Secondary Outcomes

NSSI Features

The Nonsuicidal Self-Injury Disorder Scale [36] is a self-report measure designed to assess the proposed criteria for NSSI disorder (as stated in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*) [37]. The Spanish version of the scale shows good internal reliability (Cronbach α=0.88) [36].

XSL•FO
RenderX

Emotion Regulation

The brief version of the DERS-18 is the shortened form of the original DERS [22]. The Spanish translation of this scale is being used. The DERS-18 preserves the original 6-subscale structure assessing key domains of ER difficulties, including a lack of emotional awareness, a lack of emotional clarity, nonacceptance of emotions, inability to engage in goal-directed behavior when feeling emotional, engagement in impulsive behavior when feeling emotional, and inability to access ER strategies. The scale demonstrates strong internal consistency, with a total score of Cronbach α=0.91 [38].

The Brief Experiential Avoidance Questionnaire is a 15-item measure assessing an individual's tendency to avoid or escape from unwanted internal experiences such as uncomfortable emotions, thoughts, memories, or sensations [39]. The Spanish version is being used and shows good internal consistency (Cronbach α=0.82) [39].

Symptoms

The Depression Anxiety Stress Scales–21 is a self-report measure that evaluates the frequency and severity of 21 negative emotional symptoms experienced over the past week, organized into 3 subscales: depression, anxiety, and stress. The Spanish version of the scale has shown moderate to good internal consistency, with Cronbach α coefficients of 0.84 for depression, 0.70 for anxiety, and 0.82 for stress [40].

The Borderline Symptom List–23 [41] assesses BPD symptom severity. Participants rate each item on a 5-point Likert scale from 0 (not at all) to 4 (very strong). The Spanish version of the scale is being used, which has high reliability (Cronbach α=0.949) [42].

Social Support

Perceived social support is being measured using the Multidimensional Scale of Perceived Social Support [43]. In this study, the validated Spanish version of this scale is being used, which shows good internal reliability (Cronbach α=0.88) [44]. A high score indicates greater social support, and the response format ranges from 1 (completely disagree) to 7 (completely agree).

Process Measures

The Dialectical Thinking Scale [45] is a validated 5-item measure assessing dialectical thinking. It consists of a 2-factor structure: "Both Sides," which reflects the ability to acknowledge and accept opposing viewpoints as equally valid (thereby promoting cognitive flexibility and minimizing polarized thinking) and "Both Sides in Me," which captures the capacity to incorporate and accept internal contradictions. This scale demonstrates strong psychometric properties, with high internal consistency (Cronbach α=0.81) and sound test-retest reliability (intraclass correlation coefficient=0.82 for "Both Sides" and 0.64 for "Both Sides in Me").

The Delusion of Me is a 24-item questionnaire assessing a latent construct proposed to unify 3 core mindfulness-related processes: acceptance, decentering, and nonattachment. The general factor showed good reliability, with a hierarchical omega (ωh) of 0.81, indicating that most of the reliable variance in the composite score is attributable to the general construct, rather than to specific subcomponents [46]. Participants rate each item on a 5-point Likert scale ranging from 1 (not at all true or rarely true) to 5 (always or very frequently true).

Well-Being

The Pemberton Happiness Index [47] integrates hedonic and eudaimonic aspects of well-being. The Spanish version of the index is being used, which shows high internal consistency (Cronbach α=0.91).

The Positive and Negative Affect Schedule is a 20-item self-report measure assessing positive affect (ie, interested, satisfied, and excited) and negative affect (ie, anxious, guilty, and aggressive). Participants rate each item on a scale ranging from 1 (very slightly or not at all) to 5 (extremely), indicating the extent to which each adjective describes their current mood at the time of completing the questionnaire. The result is the sum of the values. This scale is useful to measure emotional fluctuations throughout a specific period, and a Spanish adaptation is being used [48].

User Experience

To evaluate user experience with the app, participants complete the System Usability Scale, a 10-item tool with strong reliability for assessing perceived usability of digital technologies. Items are rated from 1 (strongly disagree) to 5 (strongly agree). The SUS score can be interpreted with a qualitative scale ranging from "worst imaginable" to "best imaginable" usability [49].

Daily Assessment

EMAs delivered via the app collect real-time data throughout the intervention phase corresponding to the 5 different EMI domains as previously described in Table 1.

Randomization and Procedure

Figure 3 depicts the evaluation, randomization, and allocation of participants to intervention 1 or intervention 2 between participant groups. Following screening, eligible participants in the NSSI clinical and NSSI subclinical groups are being randomized to either intervention 1 or intervention 2 using an adaptive randomization procedure stratified by age (14-16, 17-19, and 20-24 years) and sex. This approach ensures a balanced distribution across intervention arms. Each treatment group will include 120 participants, divided evenly between the clinical and subclinical groups. In addition, both intervention conditions continue to receive treatment as usual in the NSSI clinical group, which may include psychological or pharmacological care within a public health unit.

Participants receive access to either version of the app, which has been developed for both Android and iOS platforms. To use the app, the user obtains an access code from the researcher, and the app remains active for 6 weeks.

XSL•FO
RenderX

Figure 3. Study flowchart. TAU: treatment as usual.

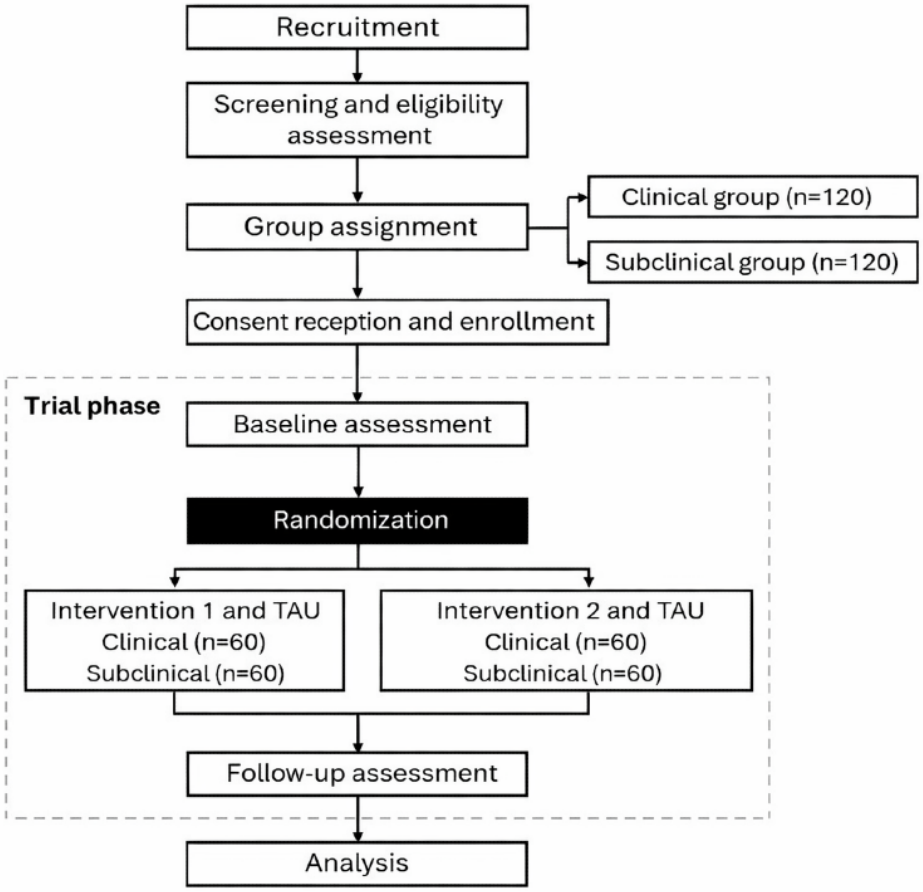

Sample Size Calculation

Power analyses conducted using G*Power (version 3.1) indicated that, to detect a medium effect size (Cohen *f*=0.27) with 90% power and a significance level of .05, 100 participants per group are required. Allowing for up to 20% attrition, 120 participants will be recruited per intervention arm, resulting in a total NSSI sample of 240 participants.

Statistical Analysis

Primary analyses will be conducted using linear mixed-effects models to evaluate the impact of the intervention on both momentary and self-report outcomes. Time (T0, T1, T2, and T3) will be included as a within-subject factor, while intervention (intervention 1 vs intervention 2) and group (clinical vs subclinical) will be modeled as between-subject factors. Random intercepts for participants and random slopes for time will be specified. Secondary analyses will explore potential moderators and mediators of treatment response, including app engagement, baseline severity, and demographic variables. Missing data will be addressed using maximum likelihood estimation under the assumption that data are missing at random.

Ethical Considerations

The study received ethics approval from the Research Ethics Committee of Bellvitge University Hospital on December 12, 2022 (reference PR336/22; CSA PR4/2022). All prospective participants are provided with written information clarifying that their involvement in the study is entirely voluntary. They are informed that choosing not to participate will not result in any negative consequences and that they retain the right to withdraw from the study at any point without penalty. Additionally, participants are given the opportunity to seek clarification or ask any further questions.

Participant privacy and confidentiality are strictly protected. Personal data are pseudonymized and stored on secure, encrypted servers, accessible only to authorized members of the research team. Data handling procedures comply with applicable data protection regulations, including the General Data Protection Regulation (EU 2016/679). No personally identifiable information will be included in study analyses or publications. Participants who complete the full intervention receive financial compensation of €45 (US $52.20). Compensation is provided for time and resources invested and is not contingent on study outcomes.

For ethical reasons, when the group assigned to intervention 2 completes the posttreatment assessment (2 months after the baseline evaluation), the Kalmer app will be offered to participants in the control group if the study shows a significant reduction in the frequency of NSSI.

Results

This study was funded in 2022 and received ethics approval in December 2022. It was also registered in the UK's clinical study registry (International Standard Randomized Controlled Trial Number; ISRCTN63093907) in March 2025. As of January 2026, a total of 145 participants have consented to participate

XSL•FO
RenderX

in our study and completed baseline assessments. Preliminary data show high app engagement and acceptability and positive user feedback regarding app usability and content. Recruitment is ongoing. We expected to complete recruitment by March 2026 and finish data analysis and publish final results in October 2026. We plan to publish the data and summary findings in a peer-reviewed journal and at conferences.

We anticipate that the results of this study will offer valuable insight into the prevalence and patterns of NSSI in both clinical and nonclinical populations, particularly in the postpandemic context. Using EMAs and self-report measures, we expect to detect real-time fluctuations in emotional states, social experiences, and NSSI urges and behaviors.

The Kalmer app–based intervention, delivered via EMIs tailored to real-time affective states, is expected to reduce the frequency and intensity of NSSI urges and behaviors. Improvements in ER, perceived social support, and self-compassion are also anticipated. Furthermore, we expect the intervention to show high acceptability and usability among young adults with a history of NSSI, as assessed through validated self-report scales.

Discussion

NSSI is a significant public health concern that typically emerges during adolescence and young adulthood and is associated with a range of negative mental health outcomes [5]. Despite increasing awareness and prevalence, access to evidence-based treatments remains limited [10-12,50]. Traditional psychotherapies such as DBT and CBT have demonstrated efficacy in reducing NSSI but are resource intensive; often inaccessible due to long waitlists, stigma, or a lack of specialized providers; and frequently underused by the populations most affected [51-54].

Digital mental health interventions, particularly mobile app–based formats, offer a promising avenue for addressing these barriers. They provide scalable, cost-effective, and private access to evidence-informed strategies that can be tailored to users' needs and embedded into daily life [55]. However, few digital tools have been specifically developed or validated for individuals engaging in NSSI, especially in young populations. There is also a lack of evidence regarding their feasibility, acceptability, and the mechanisms through which they might influence self-injurious behavior, although recent studies show promising results, including reductions in NSSI frequency, improvements in ER, and generally positive feasibility and acceptability outcomes [56].

This paper presents the protocol for an RCT to evaluate the feasibility, acceptability, and preliminary efficacy of the NSSI-specific Kalmer intervention. Kalmer is a brief mobile app–based program co-designed with clinical experts and individuals with lived experience of NSSI. It integrates EMAs with real-time EMIs targeting emotional dysregulation through mindfulness-based and self-compassion–based strategies. The app has been designed to engage young users in both clinical and subclinical settings and deliver timely, context-sensitive support.

By assessing usability, app engagement, and changes in NSSI-related thoughts and behaviors, this study aims to generate early evidence on whether such an intervention is viable and beneficial for young people. The inclusion of real-time assessments and personalized content delivery allows for a granular understanding of how emotional states and environmental contexts interact with intervention use and may potentially inform how to optimize timing and modality for future personalization.

Expected results may inform the development of personalized digital care pathways by identifying both stable (eg, personality traits) and dynamic (eg, emotional fluctuations) predictors of intervention response. This approach aligns with contemporary calls for precision mental health frameworks and could help segment users into subgroups most likely to benefit from specific components or delivery modes.

However, some limitations are anticipated. As a pilot trial, the study may be underpowered to detect small-to-moderate clinical effects and may lack a long-term follow-up. Reliance on self-report and EMA data, although informative, introduces potential for missing data and response biases. App engagement may also vary significantly across users, and understanding dropout or nonuse patterns will be critical for informing future iterations. Additionally, although the intervention is informed by established therapeutic models, its brevity and digital delivery format may limit its scope compared with comprehensive face-to-face treatment.

Despite these limitations, this trial represents a meaningful step toward a scalable and accessible intervention for people with NSSI. It leverages real-time data to adapt support to user needs, draws from stakeholder expertise, and targets an underserved but highly affected population. Findings from this feasibility phase will inform iterative improvements to Kalmer and lay the groundwork for future large-scale evaluations of its clinical efficacy and cost-effectiveness.

Acknowledgments

Following the classification of artificial intelligence (AI) activities developed by the STM Association Task and Finish Group [57], the authors disclose the use of generative AI (GAI) for the generation of specific components of images for illustrative purposes only. The GAI tool used was ChatGPT (version 4.0; OpenAI) under full human supervision. Responsibility for the final manuscript rests solely with the authors. The GAI tools are not listed as authors and bear no responsibility for the final content.

XSL•FO
RenderX

Funding

This research protocol was funded by “La Marató de TV3” Foundation (grant 202233-30-31-32-33 [Multimedia Appendix 1]). The study has been funded for a 3-year period starting in May 2023.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AJ: investigation, data curation, visualization, methodology, writing – original draft, writing – review and editing. IJ: data curation, software, resources, writing – review and editing. MCG: resources and visualization. AG-P: conceptualization, investigation, visualization, methodology, resources, software, supervision, and funding acquisition. JCPM: conceptualization, funding acquisition, software, supervision, and resources. ASE: conceptualization, funding acquisition, and resources. AL: conceptualization, funding acquisition, and resources. IM: conceptualization, funding acquisition, and resources. SR: conceptualization, funding acquisition, and resources. JP: conceptualization, funding acquisition, and resources. JS: conceptualization, funding acquisition, and resources. JS-C: conceptualization, investigation, methodology, visualization, resources, supervision, and funding acquisition. ML-S: conceptualization, investigation, methodology, visualization, resources, supervision, and funding acquisition. DV: conceptualization, investigation, methodology, visualization, project administration, resources, supervision, funding acquisition, writing – review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from ‘La Marató de TV3’ Foundation (Catalonia, Spain).
[PDF File (Adobe PDF File), 122 KB - resprot_v15i1e86413_app1.pdf]

Multimedia Appendix 2
SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist.
[DOCX File , 34 KB - resprot_v15i1e86413_app2.docx]

References

1. Vega D, Sintes A, Fernández M, Puntí J, Soler J, Santamarina P, et al. Review and update on non-suicidal self-injury: who, how and why? Actas Esp Psiquiatr 2018 Jul;46(4):146-155 [FREE Full text] [Medline: 30079928]
2. Nock MK. Self-injury. Annu Rev Clin Psychol 2010;6:339-363. [doi: 10.1146/annurev.clinpsy.121208.131258] [Medline: 20192787]
3. Swannell SV, Martin GE, Page A, Hasking P, St John NJ. Prevalence of nonsuicidal self-injury in nonclinical samples: systematic review, meta-analysis and meta-regression. Suicide Life Threat Behav 2014 Jun;44(3):273-303. [doi: 10.1111/sltb.12070] [Medline: 24422986]
4. Mendez I, Sintes A, Pascual JC, Puntí J, Lara A, Briones-Buixassa L, et al. Borderline personality traits mediate the relationship between low perceived social support and non-suicidal self-injury in a clinical sample of adolescents. J Affect Disord 2022 Apr 01;302:204-213 [FREE Full text] [doi: 10.1016/j.jad.2022.01.065] [Medline: 35038480]
5. Mars B, Heron J, Klonsky ED, Moran P, O'Connor RC, Tilling K, et al. Predictors of future suicide attempt among adolescents with suicidal thoughts or non-suicidal self-harm: a population-based birth cohort study. Lancet Psychiatry 2019 Apr;6(4):327-337 [FREE Full text] [doi: 10.1016/S2215-0366(19)30030-6] [Medline: 30879972]
6. Brunner R, Kaess M, Parzer P, Fischer G, Carli V, Hoven CW, et al. Life-time prevalence and psychosocial correlates of adolescent direct self-injurious behavior: a comparative study of findings in 11 European countries. J Child Psychol Psychiatry 2014 Apr 12;55(4):337-348. [doi: 10.1111/jcpp.12166] [Medline: 24215434]
7. Cipriano A, Cella S, Cotrufo P. Nonsuicidal self-injury: a systematic review. Front Psychol 2017;8:1946 [FREE Full text] [doi: 10.3389/fpsyg.2017.01946] [Medline: 29167651]
8. Klonsky ED. The functions of deliberate self-injury: a review of the evidence. Clin Psychol Rev 2007 Mar;27(2):226-239. [doi: 10.1016/j.cpr.2006.08.002] [Medline: 17014942]
9. Nock MK, Prinstein MJ. A functional approach to the assessment of self-mutilative behavior. J Consult Clin Psychol 2004 Oct;72(5):885-890. [doi: 10.1037/0022-006X.72.5.885] [Medline: 15482046]
10. Kothgassner OD, Goreis A, Robinson K, Huscsava MM, Schmahl C, Plener PL. Efficacy of dialectical behavior therapy for adolescent self-harm and suicidal ideation: a systematic review and meta-analysis. Psychol Med 2021 May 20;51(7):1057-1067 [FREE Full text] [doi: 10.1017/S0033291721001355] [Medline: 33875025]

XSL•FO
RenderX

11. Ougrin D, Tranah T, Stahl D, Moran P, Asarnow JR. Therapeutic interventions for suicide attempts and self-harm in adolescents: systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry 2015 Feb;54(2):97-107. [doi: 10.1016/j.jaac.2014.10.009] [Medline: 25617250]
12. Kamody RC, Pluhar E, Burton ET, Lois BH, Martin A. Nonsuicidal self-injury and barriers to accessibility of dialectical behavior therapy among black youth. J Am Acad Child Adolesc Psychiatry 2023 Apr;62(4):389-393. [doi: 10.1016/j.jaac.2022.08.002] [Medline: 35981668]
13. Hasking P, Rees CS, Martin G, Quigley J. What happens when you tell someone you self-injure? The effects of disclosing NSSI to adults and peers. BMC Public Health 2015 Oct 09;15:1039 [FREE Full text] [doi: 10.1186/s12889-015-2383-0] [Medline: 26453187]
14. Simmonds-Buckley M, Bennion MR, Kellett S, Millings A, Hardy GE, Moore RK. Acceptability and effectiveness of NHS-recommended e-therapies for depression, anxiety, and stress: meta-analysis. J Med Internet Res 2020 Oct 28;22(10):e17049 [FREE Full text] [doi: 10.2196/17049] [Medline: 33112238]
15. Arshad U, Farhat-Ul-Ain, Gauntlett J, Husain N, Chaudhry N, Taylor PJ. A systematic review of the evidence supporting mobile- and internet-based psychological interventions for self-harm. Suicide Life Threat Behav 2020 Feb;50(1):151-179 [FREE Full text] [doi: 10.1111/sltb.12583] [Medline: 31448847]
16. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
17. Feliu-Soler A, Pérez-Aranda A, Luciano JV, Demarzo M, Mariño M, Soler J, et al. Psychometric properties of the 15-item five facet mindfulness questionnaire in a large sample of Spanish pilgrims. Mindfulness 2020 Nov 20;12(4):852-862. [doi: 10.1007/S12671-020-01549-6]
18. Carver CS. You want to measure coping but your protocol's too long: consider the brief COPE. Int J Behav Med 1997;4(1):92-100. [doi: 10.1207/s15327558ijbm0401_6] [Medline: 16250744]
19. Morán Astorga C, Landero Hernández R, González Ramírez MT. COPE-28: un análisis psicométrico de la versión en español del brief COPE. Univ Psychol 2010 Jun 16;9(2):543-552. [doi: 10.11144/javeriana.upsy9-2.capv]
20. Cohen S, Mermelstein R, Kamarck T, Hoberman HM. Measuring the functional components of social support. In: Sarason IG, Sarason BR, editors. Social Support: Theory, Research and Applications. Cham, Switzerland: Springer; 1985:73-94.
21. Merz EL, Roesch SC, Malcarne VL, Penedo FJ, Llabre MM, Weitzman OB, et al. Validation of interpersonal support evaluation list-12 (ISEL-12) scores among English- and Spanish-speaking Hispanics/Latinos from the HCHS/SOL Sociocultural Ancillary Study. Psychol Assess 2014 Jun;26(2):384-394 [FREE Full text] [doi: 10.1037/a0035248] [Medline: 24320763]
22. Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. J Psychopathol Behav Assess 2004 Mar;26(1):41-54. [doi: 10.1023/B:JOBA.0000007455.08539.94]
23. Hervás G, Jódar R. Adaptación al castellano de la Escala de Dificultades en la Regulación Emocional. Clin Salud 2008;9(2):139-156 [FREE Full text]
24. Klonsky ED, Glenn CR. Assessing the functions of non-suicidal self-injury: psychometric properties of the inventory of statements about self-injury (ISAS). J Psychopathol Behav Assess 2009 Sep 30;31(3):215-219 [FREE Full text] [doi: 10.1007/s10862-008-9107-z] [Medline: 29269992]
25. Spínola LG, Calaboiça C, Carvalho IP. The use of social networking sites and its association with non-suicidal self-injury among children and adolescents: a systematic review. J Affect Disord Rep 2024 Apr;16:100781. [doi: 10.1016/j.jadr.2024.100781]
26. Reddy N, Rokito L, Whitlock J. What is the link? The relationship between non-suicidal self-injury and social media. Cornell Research Program on Self-Injury and Recovery. 2016. URL: https://www.selfinjury.bctr.cornell.edu/perch/resources/the-relationship-between-non-suicidal-self-injury-and-social-media.pdf [accessed 2025-05-29]
27. Akbari M, Seydavi M, Firoozabadi MA, Babaeifard M. Distress tolerance and lifetime frequency of non-suicidal self-injury (NSSI): a systematic review and meta-analysis. Clin Psychol Psychother 2024 Feb 12;31(1):e2957. [doi: 10.1002/cpp.2957] [Medline: 38343352]
28. Zhou J, Zhang J, Huang Y, Zhao J, Xiao Y, Zhang S, et al. Associations between coping styles, gender, their interaction and non-suicidal self-injury among middle school students in rural west China: a multicentre cross-sectional study. Front Psychiatry 2022 Aug 9;13:861917 [FREE Full text] [doi: 10.3389/fpsyt.2022.861917] [Medline: 36016979]
29. De Luca L, Pastore M, Palladino BE, Reime B, Warth P, Menesini E. The development of Non-Suicidal Self-Injury (NSSI) during adolescence: a systematic review and Bayesian meta-analysis. J Affect Disord 2023 Oct 15;339:648-659 [FREE Full text] [doi: 10.1016/j.jad.2023.07.091] [Medline: 37479039]
30. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
31. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]

XSL•FO
RenderX

32. Diez-Quevedo C, Rangil T, Sanchez-Planell L, Kroenke K, Spitzer RL. Validation and utility of the patient health questionnaire in diagnosing mental disorders in 1003 general hospital Spanish inpatients. Psychosom Med 2001;63(4):679-686. [doi: 10.1097/00006842-200107000-00021] [Medline: 11485122]
33. Al-Halabí S, Sáiz PA, Burón P, Garrido M, Benabarre A, Jiménez E, et al. Validación de la versión en español de la Columbia-Suicide Severity Rating Scale (Escala Columbia para Evaluar el Riesgo de Suicidio). Rev Psiquiatr Salud Ment 2016 Jul;9(3):134-142. [doi: 10.1016/j.rpsmen.2016.06.004]
34. Clinician-rated severity of non-suicidal self-injury. American Psychiatric Association. 2013. URL: https://www.psychiatry.org/File%20Library/Psychiatrists/Practice/DSM/APA_DSM5_Clinician-Rated-Severity-of-Non-Suicidal-Self-Injury.pdf [accessed 2025-06-01]
35. Vega D, Torrubia R, Soto À, Ribas J, Soler J, Pascual JC, et al. Exploring the relationship between non suicidal self-injury and borderline personality traits in young adults. Psychiatry Res 2017 Oct;256:403-411. [doi: 10.1016/j.psychres.2017.07.008] [Medline: 28692921]
36. Victor SE, Davis T, Klonsky ED. Descriptive characteristics and initial psychometric properties of the non-suicidal self-injury disorder scale. Arch Suicide Res 2017 Jun 07;21(2):265-278 [FREE Full text] [doi: 10.1080/13811118.2016.1193078] [Medline: 27267416]
37. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. 5th edition. Arlington, TX: American Psychiatric Publishing; 2013.
38. Victor SE, Klonsky ED. Validation of a brief version of the difficulties in emotion regulation scale (DERS-18) in five samples. J Psychopathol Behav Assess 2016 May 13;38(4):582-589. [doi: 10.1007/S10862-016-9547-9]
39. Vázquez-Morejón R, León Rubio JM, Martín Rodríguez A, Vázquez Morejón AJ. Validation of a Spanish version of the Brief Experiential Avoidance Questionnaire (BEAQ) in clinical population. Psicothema 2019 Aug 01;31(3):335-340. [doi: 10.7334/psicothema2019.60] [Medline: 31292050]
40. Bados A, Solanas A, Andrés R. Psychometric properties of the Spanish version of Depression, Anxiety and Stress Scales (DASS). Psicothema 2005;17(4):679-683 [FREE Full text]
41. Bohus M, Kleindienst N, Limberger MF, Stieglitz R, Domsalla M, Chapman AL, et al. The short version of the Borderline Symptom List (BSL-23): development and initial data on psychometric properties. Psychopathology 2009;42(1):32-39. [doi: 10.1159/000173701] [Medline: 19023232]
42. Soler J, Vega D, Feliu-Soler A, Trujols J, Soto Á, Elices M, et al. Validation of the Spanish version of the borderline symptom list, short form (BSL-23). BMC Psychiatry 2013 May 14;13(1):1317. [doi: 10.1186/1471-244X-13-139]
43. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Pers Assess 1988 Mar;52(1):30-41. [doi: 10.1207/s15327752jpa5201_2]
44. Calderón C, Ferrando PJ, Lorenzo-Seva U, Gómez-Sánchez D, Fernández-Montes A, Palacín-Lois M, et al. Multidimensional Scale of Perceived Social Support (MSPSS) in cancer patients: psychometric properties and measurement invariance. Psicothema 2021 Feb;33(1):131-138. [doi: 10.7334/psicothema2020.263] [Medline: 33453746]
45. Soler J, Arqueros M, Schmidt C, Otero D, Soria-Madrid A, Campins MJ, et al. Scale to measure dialectical thinking from dialectical behavior therapy perspective. Eur J Psychiatry 2025 Apr;39(2):100305. [doi: 10.1016/j.ejpsy.2025.100305]
46. Soler J, Montero-Marin J, Domínguez-Clavé E, González S, Pascual JC, Cebolla A, et al. Decentering, acceptance, and non-attachment: challenging the question "is it me?". Front Psychiatry 2021 Nov 18;12:659835 [FREE Full text] [doi: 10.3389/fpsyt.2021.659835] [Medline: 34867498]
47. Hervás G, Vázquez C. Construction and validation of a measure of integrative well-being in seven languages: the Pemberton Happiness Index. Health Qual Life Outcomes 2013 Apr 22;11(1):66 [FREE Full text] [doi: 10.1186/1477-7525-11-66] [Medline: 23607679]
48. Bonifacio S, Chorot P, Lostao L, Joiner TE, Santed MA, Valiente RM. Escalas PANAS de afecto positivo y negativo: validación factorial y convergencia transcultural. Psicothema 1999;11(1):37-51.
49. Brooke J. SUS - a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. New York, NY: Taylor & Francis; 1996:189-194.
50. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]
51. Kapusta ND, Posch M, Niederkrotenthaler T, Fischer-Kern M, Etzersdorfer E, Sonneck G. Availability of mental health service providers and suicide rates in Austria: a nationwide study. Psychiatr Serv 2010 Dec;61(12):1198-1203. [doi: 10.1176/ps.2010.61.12.1198] [Medline: 21123403]
52. Kawaguchi H, Koike S. Association between the Density of Physicians and Suicide Rates in Japan: nationwide ecological study using a spatial Bayesian model. PLoS One 2016 Feb 3;11(2):e0148288 [FREE Full text] [doi: 10.1371/journal.pone.0148288] [Medline: 26840389]
53. Tondo L, Albert MJ, Baldessarini RJ. Suicide rates in relation to health care access in the United States: an ecological study. J Clin Psychiatry 2006 Apr 15;67(4):517-523. [doi: 10.4088/jcp.v67n0402] [Medline: 16669716]
54. Bruffaerts R, Demyttenaere K, Hwang I, Chiu W, Sampson N, Kessler RC, et al. Treatment of suicidal people around the world. Br J Psychiatry 2011 Jul;199(1):64-70. [doi: 10.1192/bjp.bp.110.084129] [Medline: 21263012]

XSL•FO
RenderX

55. Olff M. Mobile mental health: a challenging research agenda. Eur J Psychotraumatol 2015 May 19;6(1):27882 [FREE Full text] [doi: 10.3402/ejpt.v6.27882] [Medline: 25994025]
56. Julià A, Jaén I, García-Palacios A, Pascual JC, Sintes A, Lara A, et al. Delivering real-time support for self-injury: a systematic review on ecological momentary interventions. Internet Interv 2025 Jun;40:100826 [FREE Full text] [doi: 10.1016/j.invent.2025.100826] [Medline: 40342957]
57. Recommendations for a classification of AI Use in academic manuscript preparation. STM Association. 2022 Sep 22. URL: https://stm-assoc.org/document/recommendations-for-a-classification-of-ai-use-in-academic-manuscript-preparation/ [accessed 2025-05-29]

Abbreviations

BPD: borderline personality disorder
CBT: cognitive behavioral therapy
DBT: dialectical behavior therapy
DERS: Difficulties in Emotion Regulation Scale
DT: distress tolerance
EMA: ecological momentary assessment
EMI: ecological momentary intervention
ISAS: Inventory of Statements About Self-Injury
NSSI: nonsuicidal self-injury
RCT: randomized controlled trial

Edited by J Sarvestan; submitted 27.Oct.2025; this is a non–peer-reviewed article;accepted 12.Dec.2025; published 27.Jan.2026.

Please cite as:
Julià A, Jaén I, Conesa Giménez M, García-Palacios A, Pascual JC, Sintes Estévez A, Lara A, Méndez I, Romero S, Puntí J, Soler J, Solé-Casals J, López-Solà M, Vega D
A Smartphone-Based Psychological Intervention for Nonsuicidal Self-Injury (Kalmer App): Protocol for a Multicenter Double-Blind Randomized Controlled Trial
JMIR Res Protoc 2026;15:e86413
URL: https://www.researchprotocols.org/2026/1/e86413
doi:10.2196/86413
PMID:

©Anna Julià, Irene Jaén, Mónica Conesa Giménez, Azucena García-Palacios, Juan C Pascual, Anna Sintes Estévez, Anaís Lara, Iria Méndez, Soledad Romero, Joaquim Puntí, Joaquim Soler, Jordi Solé-Casals, Marina López-Solà, Daniel Vega. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Evaluating the Effectiveness of the Motivational Interviewing–Based Wellness Coaching for Kids and Families (WC4K) Program in Pediatric Obesity Care: Protocol for a Cluster Randomized Pragmatic Trial

Corinna Koebnick[1,2*], PhD; Margo Sidell[1*], SCD; Jessica D Vallejo[1], PhD; Xia Li[1], MSc; Emerson Delacroix[3], MS; Deborah Young[1,2], PhD; Poornima Kunani[1], MD; Ken Resnicow[4], PhD

[1]Department of Research and Evaluation, Kaiser Permanente, Pasadena, CA, United States

[2]Department of Health Systems Science, Kaiser Permanente Bernard J. Tyson School of Medicine, Pasadena, CA, United States

[3]School of Public Health, University of Michigan, Ann Arbor, MI, United States

[4]School of Public Heatlth, University of Minnesota, Minneapolis, MN, United States

[*]these authors contributed equally

Corresponding Author:
Corinna Koebnick, PhD
Department of Research and Evaluation
Kaiser Permanente
100 S. Los Robles Ave
Pasadena, CA, 91101
United States
Phone: 1 6267730307
Email: Corinna.Koebnick@kp.org

Abstract

Background: The increasing prevalence of severe obesity among youth and the early onset of comorbidities highlight the urgent need for effective solutions to support behavior and lifestyle changes. Motivational interviewing (MI), a client-centered counseling technique, has shown promise in improving obesity-related outcomes and is now recommended by the American Academy of Pediatrics as a key component of behavioral interventions for children with overweight and obesity.

Objective: This study aimed to describe the design and baseline characteristics of the Wellness Coaching for Kids and Families (WC4K) program, an MI-based behavioral health intervention integrated into pediatric primary care within a large integrated health care system. This trial aims to inform broader implementation strategies for other health care systems.

Methods: We are conducting a cluster randomized pragmatic trial across 50 pediatric clinics within Kaiser Permanente Southern California. Clinics were randomized to either the intervention arm (n=24, 48.00%) or the usual care arm (n=26, 52.00%), targeting children aged between 3 and 8 years with overweight or obesity. Pediatricians in intervention clinics received MI training and referred families to centralized health coaches for tailored telephone counseling. Key behavioral targets include increased fruit and vegetable intake, reduced consumption of sugar-sweetened beverages, increased physical activity, and reduced screen time. The primary outcome is a change in BMI-for-age, measured as relative distance from the median using electronic medical record data. Secondary outcomes include parent-reported behavior change indicators.

Results: The source population includes 150,482 children from clinics serving predominantly low-income and racial and ethnic minority populations. After randomization, intervention and control clinics were similar in demographics (standardized mean differences <0.2 for key variables), with 77,481 (51.49%) children in the WC4K intervention clinics and 73,001 (48.51%) children in the control clinics. In the total clinic population, 36.58% (55,052/150,482) of the children were overweight or with obesity. Enrollment started in fall 2022, study results are expected in spring 2027.

Conclusions: If the trial results indicate success in reducing BMI and improving health behaviors, WC4K may offer a scalable and sustainable model for integrating behavioral health coaching into pediatric primary care. By leveraging MI-trained pediatricians and centralized health coaching, WC4K has the potential to facilitate meaningful lifestyle changes in children with overweight or obesity.

Trial Registration: ClinicalTrials.gov NCT05143697; https://clinicaltrials.gov/study/NCT05143697

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/78792

(JMIR Res Protoc 2026;15:e78792) doi:10.2196/78792

KEYWORDS

childhood obesity; motivational interviewing; pediatric primary care; pragmatic trial; lifestyle coaching; health equity

Introduction

Childhood obesity in the United States has escalated substantially over the past 3 decades, with the prevalence increasing from 10.0% in 1988 to 1994 to 19.3% in 2017 to 2020 among children aged between 2 and 11 years [1]. This trend has immediate consequences, including elevated blood pressure, insulin resistance, and psychosocial issues [2-8] as well as long-term risks, such as type 2 diabetes, cardiovascular disease, and premature mortality [9-14].

With approximately 1 in 5 US children living with obesity by the age of 11 years [1], effective interventions are urgently needed. Despite decades of public health efforts [15], the increasing prevalence of severe obesity and the early onset of obesity-related comorbidities in youth underscore the urgent need for scalable, effective treatments [16-18].

The American Academy of Pediatrics (AAP) now recommends comprehensive, intensive behavioral interventions, with motivational interviewing (MI) identified as a key component of treatment [16]. Traditional counseling approaches often rely on directive advice giving, which may not be the most effective method for promoting sustained behavior change [19]. In contrast, MI is a client-centered, collaborative technique that helps individuals resolve ambivalence and align behavior with personal values [19,20]. Originally developed to treat substance use disorders [21], MI has since demonstrated efficacy across a range of health behaviors, including diet, physical activity, and weight management [16,22-24].

In pediatric populations, MI has shown promise in improving obesity-related outcomes when delivered by trained health care providers [25-29]. Pediatricians are uniquely positioned to address childhood obesity through routine screening, anticipatory guidance, and family-centered counseling [30-33]. Although many express a strong desire to play a more active role in prevention and treatment, pediatricians report feeling underprepared to manage obesity effectively, citing limited training, a lack of confidence in achieving results, or being overwhelmed by other responsibilities [33-36]. More than 80% of the pediatricians report that they are disappointed in their ability to treat childhood obesity [36]. However, when equipped with behaviorally informed tools such as MI, pediatricians can successfully engage families and support meaningful change [37,38].

Qualitative studies have shown that parents often underestimate the long-term impact of early weight gain unless explicitly counseled by their pediatrician [39]. To address these needs, the Wellness Coaching for Kids and Families (WC4K) program was developed to adapt an existing primary care obesity intervention previously tested in the AAP Pediatric Research in Office Settings network for real-world implementation in a large, integrated health system [37,40]. Key features of WC4K include the integration of program components into the electronic medical record (EMR) system, the initiation of lifestyle treatment through pediatricians, centralized MI-trained health coaches, flexible MI dosing, telehealth delivery, and mobile-friendly tools for communication and engagement. This pragmatic trial aims to evaluate the effectiveness of WC4K in routine pediatric care and inform broader implementation strategies.

Methods

Study Design

The WC4K trial is a cluster randomized pragmatic trial with a parallel design, evaluating the effectiveness and sustainability of an MI-based behavioral health intervention integrated into pediatric primary care. Clinics were the unit of randomization. This study was guided by the Pragmatic-Explanatory Continuum Indicator Summary tool [41,42], with scores of 4 or 5 across all 9 domains indicating a highly pragmatic approach.

The WC4K program uses MI to promote healthy lifestyle changes aimed at reducing excess weight in children with overweight or obesity. Key behavioral targets include fruit and vegetable intake, consumption of sugar-sweetened beverages, physical activity, and screen time. The intervention is a modified version of the Brief Motivational Interviewing to Reduce BMI (BMI^2) intervention [40]. This study uses an implementation-effectiveness hybrid design (type 2) to allow a simultaneous evaluation of clinical outcomes and implementation processes [43].

Ethical Considerations

The Kaiser Permanente Southern California (KPSC)/Hawaii Institutional Review Board approved this study and granted a waiver of informed consent because the study posed minimal risk, involved secondary use of routinely collected clinical data, was embedded into routine clinical care (including wellness coaching), and could not practicably be carried out with individual consent. The study is registered on ClinicalTrials.gov (NCT05143697). Wellness coaching was offered as a covered clinical service; no additional compensation was provided for research participation. All interactions occurred with parents or caregivers; no direct data collection or assent procedures were conducted with children. Study data were derived from KPSC electronic medical records. All data were stored on secure, password-protected servers behind institutional firewalls. Data access was restricted to study personnel, and all analyses were conducted within the health system's secure computing environment. No individual-level data are publicly available.

XSL•FO
RenderX

Population

KPSC provides care for more than 4.6 million members, including approximately 900,000 children and adolescents as of January 1, 2025. The membership population is generally reflective of the population in the region [44]. The eligible population includes parents and caregivers of children aged between 3 and 8 years with overweight or obesity who receive care at participating clinics. No exclusion criteria were applied, consistent with the pragmatic nature of the trial.

Health Care Setting

KPSC operates under a prepaid, integrated care model where physicians are salaried, enabling a focus on evidence-based care rather than service volume [45]. The infrastructure supports coordinated care across primary, specialty, and hospital services, with a strong emphasis on prevention and chronic disease management. Members typically have access to preventive services with minimal out-of-pocket costs. Lifestyle coaching is a covered benefit.

Patient information is captured in EMRs, which support population health management tools, such as electronic referrals, centralized services, and follow-up reminders. These features facilitate rapid implementation of care adaptations across the region.

Clinic Selection

Of the 88 pediatric clinics in the KPSC region, 50 (56.82%) clinics with the highest proportion of low-income and racial and ethnic minority patients were selected. One clinic served as a pilot site. The 50 clinics were grouped into small (<300 children), medium (300-999 children), and large (≥1000 children) clinics based on their pediatric population at the target age group and block randomized using clinic size, resulting in 24 (48%) clinics in the WC4K intervention arm and 26 (52%) clinics in the usual care arm.

Recruitment

Recruitment occurred during scheduled well-child visits using routine workflows. Pediatricians used EMR-embedded tools to identify and refer eligible families to coaching (Figure 1). Recruitment concluded in intervention clinics in December 2024. For control clinics, a synthetic control cohort will be created to address potential differences in demographic and clinical characteristics (more details are provided in the Statistical Analysis section).

Figure 1. Study flowchart. KPSC: Kaiser Permanente Southern California.

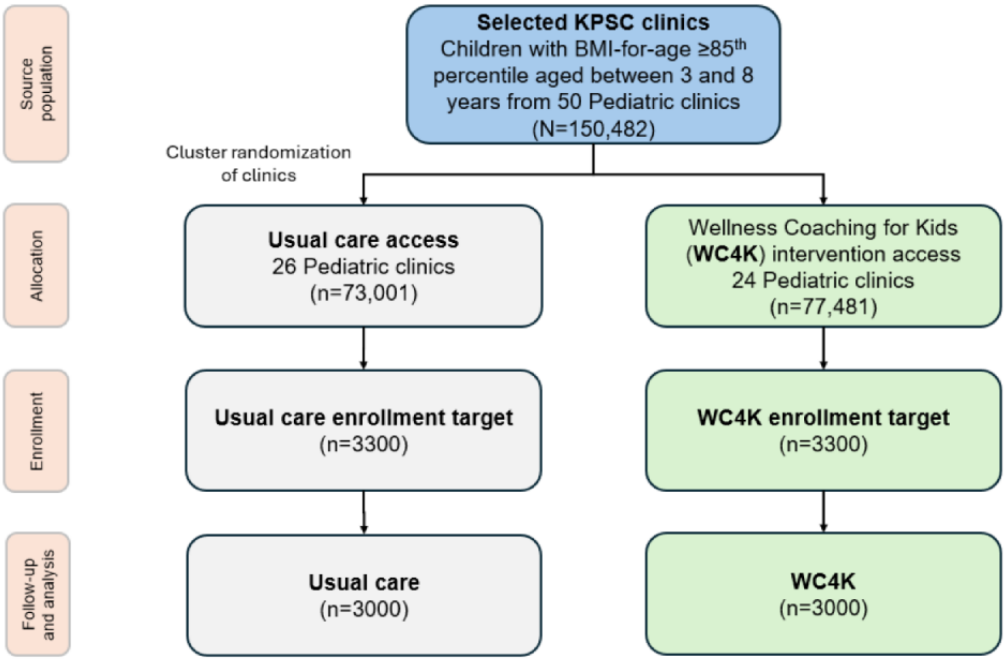

Planning and Piloting Phases

Overview

One pediatric clinic volunteered to serve as the pilot site. Workflows were developed and vetted with regional leadership to ensure seamless integration into usual care. EMR tools were built during this phase to support pediatrician referrals and family engagement.

The pilot led to a streamlined role for pediatricians, focusing on initiating MI-based referrals rather than delivering extensive counseling. Pediatricians used MI strategies to engage families and refer them to centralized health coaches. Table 1 summarizes the differences in workflow and treatment components between WC4K and usual care, and Figure 2 shows the wellness screener.

XSL•FO
RenderX

Table 1. Comparison of lifestyle treatment in Wellness Coaching for Kids and Families (WC4K) program and usual care and the roles of pediatricians and lifestyle coaches.

Roles	WC4K	Usual care
Previsit questionnaires	• AAP[a] Bright Futures–consistent assessment for well-child visits • Wellness screener to assess key target behaviors	• AAP Bright Futures–consistent assessment for well-child visits
Pediatricians	• Initiate discussion about LT[b] during well-child visits • Use key MI[c] strategies (explore, guide, and choose) to obtain buy-ins for referral to MI-based lifestyle coaching • Enter referral to health coaches • Provide educational materials and self-monitoring logs to parents	• Initiate discussion about LT during well-child visits • Usual communication style, mostly directive • Enter referral to LT classes or self-study videos • Provide educational materials and inform families about KP.org materials
Health behavior and lifestyle treatment mode	• MI-based lifestyle coaching • Intake appointment (45 min); 5-9 follow-up visits (20 min) over 24 months	• Online group class with a trainer or coach: 16 sessions (75 min); abbreviated version available (4×60 min) • On-demand self-study videos about a healthy diet and lifestyle (3×20 min)
Target behaviors	• Fruits and vegetable intake, exercise, sedentary time, sugar-sweetened beverages, and snack foods, consistent with AAP Bright Futures guidelines	• Fruit and vegetable intake, exercise, sedentary time, sugar-sweetened beverages, and snack foods, consistent with AAP Bright Futures guidelines

[a]AAP: American Academy of Pediatrics.

[b]LT: lifestyle treatment.

[c]MI: motivational interviewing.

Figure 2. Wellness screener sent at scheduling of well-child visit (intervention clinics only).

1. **Are you concerned about your child's weight?**

1	2	3	4	5	6	7
Not at all concerned			Concerned			Extremely concerned

2. **For each of the behaviors below, please grade how your child is doing.**

A = Great/Healthy to F = Poor/Unhealthy

Behaviors	A Great/ Healthy	B	C	D	F Poor/ Unhealthy
Snack Foods					
Drinking Sweetened Beverages					
Eating Out/Carry Out Dinners					
Eating Fruits					
Eating Vegetables					
Watching TV/Screen Time					
Playing Video Games/Internet Games					
Physical Activity/Exercise					
Having family home prepared meals together					
Portion size/Second Helpings					
Emotional eating					

Usual Pediatric Care With Attention Control

In control clinics, pediatric care followed standard practices aligned with AAP Bright Futures guidelines [46-48]. Families completed a previsit questionnaire, and pediatricians used responses and Centers for Disease Control and Prevention (CDC) BMI growth charts to guide discussions [49,50]. Pediatricians provided general lifestyle advice and referred families to virtual health behavior and lifestyle treatment classes or self-study videos. As part of usual care, target behaviors for general lifestyle advice are fruit and vegetable intake, sugar-sweetened beverages, and energy-dense foods, promoting at least 60 minutes of daily physical activity and limiting screen

XSL•FO
RenderX

time to below 2 hours per day. No MI training was provided in control clinics.

Pediatric Care Following the WC4K Program

The WC4K program builds on the BMI^2 framework and emphasizes MI as its core behavioral strategy [40,51,52]. Pediatricians in intervention clinics received structured MI training, including EMR-integrated tools to support behavior change conversations. Families completed the same previsit questionnaire as in control clinics and a wellness screener to prioritize health behaviors for discussion (agenda setting; Figure 2). Target behaviors are similar to usual care.

During the well-child visits, pediatricians applied MI techniques to engage families. For parents of children identified as being overweight or having obesity, they encouraged participation in the WC4K program and initiated a "warm handoff" by entering a referral to centralized health coaches. Depending on the clinical encounter, pediatricians also provided families with self-monitoring logs and printed educational materials.

MI Training and Counseling Approach

MI differs from traditional directive counseling by focusing on eliciting intrinsic motivation. MI emphasizes autonomy, collaboration, and intrinsic motivation. In pediatric care, especially with young children aged between 3 and 8 years, MI is primarily directed toward parents or caregivers, who play a central role in shaping the child's health behaviors. MI-trained pediatricians and health coaches avoid direct persuasion and instead focus on building high-quality motivation before initiating action planning.

In MI encounters, pediatricians assess and reflect on the patient's and their family's routines, values, and goals and encourage families to consider how current behaviors impact their child's health and development. Ambivalence and resistance are explored respectfully, allowing families to voice their own reasons for and against change. This process helps align behavior change with what matters most to the family, such as supporting their child's growth, confidence, or long-term well-being.

Key MI techniques include reflective listening and eliciting change talk—statements that express desire, ability, reasons, or a need for change. Reflections can be viewed as a form of hypothesis testing, helping caregivers clarify their thoughts and motivations. MI assumes that individuals are more likely to act on ideas they articulate themselves [53], a concept known as change talk [53-56].

The WC4K program trains pediatricians to use the 3-phase MI model: *explore, guide, and choose* [19,20,40,57,58]. Although this model often follows a temporal sequence, starting with exploring concerns and guiding toward options before choosing a plan, not all encounters follow a linear path. Some patients may arrive ready to act, while others may need more time to build motivation. The flexibility of MI allows pediatricians to tailor conversations to each family's readiness, cultural context, and emotional state, making it especially valuable in pediatric obesity care.

To support implementation, pediatricians in intervention clinics had access to a study dashboard displaying referral and uptake metrics. For ongoing training, video modules were provided featuring example encounters and demonstrations of MI techniques tailored to pediatric care. The training was eligible for maintenance of certification part 4 credit, with participation actively encouraged by regional leadership.

WC4K Lifestyle Coaching

Health coaching was delivered by a centralized team of MI-trained coaches with more than 5 years of experience. Coaching included a 45-minute intake session and 5 to 9 follow-up sessions (approximately 20 min each) over 24 months.

Coaches worked with parents to set goals, track progress, and reinforce behavior change. Children were encouraged to attend the coaching sessions, but participation was optional and dependent on family preferences and logistics.

MI Fidelity

To ensure fidelity, coaches participated in ongoing supervision and periodic session reviews using validated fidelity checklists [59]. Pediatricians were offered MI fidelity assessments using standardized patient encounters and could receive up to 3 one-on-one coaching sessions to improve MI skills.

Outcome Measures

The primary outcome is a change in BMI-for-age over 24 months, measured as the relative distance from the median using EMR data and CDC extended BMI percentiles [50,60-62]. No additional visits were scheduled for height and weight collection. Secondary outcomes include parent-reported behavior changes.

At KPSC, height and weight are measured by clinical staff during routine care and automatically recorded in the EMR, allowing for real-time BMI calculation (weight in kg divided by height in m^2).

Statistical Analysis

Identification of the Control Cohort

Children in the control cohort were identified from EMR records based on having a well-child visit during the study period. To construct a comparable control cohort, we will apply inverse probability of treatment weighting using propensity scores to adjust for confounding [63].

Each child's propensity score (defined as the probability of being referred to WC4K) will be estimated using observed covariates, including age, sex, baseline relative distance from the median BMI-for-age, socioeconomic status, and other EMR-derived factors. Children in control clinics will then be weighted by the inverse of their propensity score, allowing a balanced comparison with the intervention group and reducing selection bias [63].

Power

The power analysis was performed for the primary outcome, the difference in the change in the relative distance to the median BMI-for-age, and takes into account the intraclass correlation among children within clinics [64].

XSL•FO
RenderX

Across the 49 selected medical centers (Figure 1), the number of children aged between 3 and 8 years with overweight or obesity ranged from 500 to 6000 per site, with a coefficient of variation of 0.9. On the basis of previous studies [65,66], we assumed that the participating children would average approximately 30% to 40% above the median (which equates to the 97th percentile; SD 20% to 25%). Based on an expected effect size of 5.3% (SD 14.7%), a Cohen *h* of approximately 0.3 was used.

With 24 clinics in the intervention arm and 26 in the control arm and assuming a 2-sided α of .05, an intraclass correlation of 0.01, a coefficient of variation of 0.9, and an SD of 27.0% for change in BMI distance, a sample size of approximately 3300 children per arm provides adequate power to detect a difference in difference (DID) of 1 to 5 units in the primary outcome.

Planned Statistical Analysis

To evaluate the intervention effect, we will use a DID approach comparing the change in BMI-for-age distance between the WC4K and usual care groups. Using repeated measurement data, the DID model will estimate the average within-child change from baseline to follow-up, using all available BMI data points.

A linear mixed-effects regression model will be used to account for clustering. The model will include a random intercept for "child" to account for repeated measures. Baseline covariates with standardized mean differences (SMDs) greater than 0.1 between groups will be included as fixed effects, along with relevant demographic, clinical, and site-level characteristics.

For descriptive comparisons, SMDs will be used to assess the balance between groups.

Results

Overview

A total of 368,570 children aged between 3 and 8 years received care across KPSC pediatric clinics (Table 2). The population was racially and socioeconomically diverse, with the largest subgroup identifying as Hispanic (n=171,400, 46.50%) and 86,715 (23.53%) children identifying as non-Hispanic White. Of the total clinic population, 113,443 (30.24%) children were overweight or with obesity.

Clinic selection based on high proportions of racial and ethnic minority and low-income patients resulted in higher proportions of patients with public health insurance, racial and ethnic minority background, and living in neighborhoods with a high deprivation index (SMD >0.2). However, the selected clinics also had a slightly lower proportion of children living in neighborhoods with incomes more than US $100,000 (59,483/150,482, 39.53% vs 98,062/218,088, 44.97%) and a higher proportion living in neighborhoods with incomes below US $35,000 (30,882/150,482, 20.52% vs 39,045/218,088, 17.90%).

XSL•FO
RenderX

Table 2. Characteristics of children aged between 3 and 8 years receiving care in Kaiser Permanente Southern California (KPSC), the pediatric clinics selected for the pragmatic trial (higher than average proportion of Hispanic and Black patients and those from low-income background), and pediatric clinics that were excluded from the trial (lower than average proportion of Hispanic and Black patients and those from low-income background)[a].

Characteristic	Pediatric clinics			Standardized mean difference
	All KPSC (N=368,570)	Included (n=150,482)	Excluded (n=218,088)	Included versus excluded
Sex, n (%)				0.02
Male	180,177 (48.89)	72,884 (48.43)	107,293 (49.20)	
Female	188,393 (51.11)	77,598 (51.57)	110,795 (50.80)	
Age (y), mean (SD)	5.63 (1.85)	5.81 (1.78)	5.51 (1.88)	0.16
Race or ethnicity, n (%)				0.37
Asian or Pacific Islander	38,891 (10.55)	14,868 (9.88)	24,023 (11.02)	
Black	24,463 (6.64)	12,371 (8.22)	12,092 (5.54)	
Hispanic	171,400 (46.50)	82,963 (55.13)	88,437 (40.55)	
White	86,715 (23.53)	27,560 (18.31)	59,155 (27.12)	
Other or unknown	47,101 (12.78)	12,720 (8.46)	34,381 (15.77)	
Public health plan, n (%)				0.21
No	276,256 (75.95)	104,578 (69.50)	171,678 (78.72)	
Yes	92,314 (25.05)	45,904 (30.50)	46,410 (21.28)	
Neighborhood education[a]				0.36
Less than high school	68,858 (18.68)	32,090 (21.32)	36,768 (16.86)	
High school	85,604 (23.23)	37,114 (24.66)	48,490 (22.23)	
Some college or higher	214,108 (58.09)	81,278 (54.02)	132,830 (60.91)	
Neighborhood income[a] (US $) , n (%)				0.26
<35,000	69,927 (18.97)	30,882 (20.52)	39,045 (17.90)	
35,000-49,999	36,246 (9.83)	15,810 (10.51)	20.436 (9.37)	
50,000-74,999	56,315 (15.28)	24,097 (16.01)	32,218 (14.77)	
75,000-99,999	48,537 (13.17)	20,210 (13.43)	28,327 (12.99)	
≥100,000	157,545 (42.75)	59,483 (39.53)	98,062 (44.97)	
Neighborhood deprivation index quintile[b] , n (%)				0.36
Least deprivation	50,834 (13.79)	13,417 (8.92)	37,417 (17.16)	
Below-average deprivation	75,571 (20.50)	25,675 (17.06)	49,896 (22.88)	
Average deprivation	98,134 (26.63)	39,147 (26.01)	58,987 (27.05)	
Above-average deprivation	82,334 (22.34)	40,170 (26.69)	42,164 (19.33)	
Most deprivation	61,697 (16.74)	32,073 (21.32)	29,624 (13.58)	
BMI-for-age, n (%)				0.08
Missing	49,930 (13.55)	3270 (2.17)	46,660 (21.40)	
Underweight	8224 (2.58)[c]	3715 (2.52)[d]	4509 (2.63)[e]	
Normal weight	196,973 (61.82)[c]	88,445 (60.08)[d]	108,528 (63.31)[e]	
Overweight	53,212 (16.70)[c]	24,950 (16.95)[d]	28,262 (16.49)[e]	
Obese	60,231 (18.90)[c]	30,102 (20.45)[d]	30,129 (17.57)[e]	

[a]Neighborhood income and education were derived from US census tract data and reflect median household income and the percentage of adults aged 25 years or older without a high school diploma.

XSL•FO
RenderX

[b]Least deprivation (–2.15 to –0.68), below-average deprivation (–0.68 to –0.15), average deprivation (–0.15 to 0.42), above-average deprivation (0.42-1.12), and most deprivation (1.12-5.77) calculated relative to the Kaiser Permanente Southern California population.

[c]N=318,640.

[d]N=147,212.

[e]N=171,428.

Intervention Versus Control Clinic Population Characteristics

Clinics were randomized into either the WC4K intervention (n=24 clinics; n=77,481 children) or the usual care (n=25 clinics; n=73,001 children) arm. The demographic characteristics of children in both arms were similar across key variables, including age, sex, race or ethnicity, insurance status, BMI category, and neighborhood deprivation indicators (SMD <0.2 for all comparisons; Table 3).

XSL•FO
RenderX

Table 3. Characteristics of the children aged between 3 and 8 years receiving care in pediatric clinics randomized into the Wellness Coaching for Kids and Families (WC4K) intervention (n=24) or usual care (n=26) arms[a].

Characteristic	WC4K intervention clinics (n=77,481)	Usual care clinics (n=73,001)	Standardized mean difference
Sex, n (%)			0.002
Male	37,483 (48.38)	35,401 (48.49)	
Female	39,998 (51.62)	37,600 (51.51)	
Age (y), mean (SD)	5.83 (1.77)	5.79 (1.78)	0.02
Race or ethnicity			0.09
Asian or Pacific Islander	6711 (8.66)	8157 (11.17)	
Black	6537 (8.44)	5834 (7.99)	
Hispanic	43,003 (55.50)	39,960 (54.74)	
White	14,783 (19.08)	12,777 (17.50)	
Other or unknown	6447 (8.32)	6273 (8.60)	
Public health plan, n (%)			0.05
No	52,938 (68.32)	51,640 (70.74)	
Yes	24,543 (31.68)	21,361 (29.26)	
Neighborhood education[b]			0.04
Less than high school	16,311 (21.05)	15,779 (21.61)	
High school	19,674 (25.39)	17,440 (23.89)	
Some college or higher	41,496 (53.56)	39,782 (54.5)	
Neighborhood income[b] (U S $) , n (%)			0.05
<35,000	16,081 (20.75)	14,800 (20.27)	
35,000-49,999	8252 (10.65)	7559 (10.35)	
50,000-74,999	12,408 (16.01)	11,689 (16.01)	
75,000-99,999	10,446 (13.48)	9765 (13.38)	
100,000	30,295 (39.11)	29,188 (39.99)	
Neighborhood deprivation index quintiles[b] , n (%)			0.11
Least deprivation	6061 (7.82)	7356 (10.08)	
Below-average deprivation	12,354 (15.94)	13,321 (18.25)	
Average deprivation	21,299 (27.49)	17,848 (24.45)	
Above-average deprivation	20,783 (26.82)	19,387 (26.56)	
Most deprivation	16,984 (21.93)	15,089 (20.66)	
BMI-for-age group, n (%)			0.02
Missing	1829 (2.36)	1441 (1.97)	
Underweight	2037 (2.69)[c]	1678 (2.34)[d]	
Normal weight	45,434 (60.06)[c]	43,011 (60.1)[d]	
Overweight	12,685 (16.77)[c]	12,265 (17.14)[d]	
Obese	15,496 (20.48)[c]	14,606 (20.42)[d]	

[a]Neighborhood income and education were derived from US census tract data and reflect median household income and the percentage of adults aged 25 years or older without a high school diploma, respectively.

[b]Least deprivation (–2.15 to –0.68), below-average deprivation (–0.68 to –0.15), average deprivation (–0.15 to 0.42), above-average deprivation (0.42-1.12), and most deprivation (1.12-5.77) calculated relative to the Kaiser Permanente Southern California population.

[c]N=75,652.

XSL•FO
RenderX

[d]N=71,560.

Discussion

Overview and Impact of the WC4K Trial

The WC4K trial addresses a critical gap in pediatric obesity care: the need for scalable, sustainable interventions that can be embedded into routine primary care. Despite decades of clinical guidelines and public health campaigns, the prevalence of childhood obesity remains high, particularly among low-income and racial and ethnic minority populations [1,16,67-69].

WC4K was designed to evaluate a pragmatic, family-centered, and health system–integrated approach to promoting healthy lifestyle behaviors in children aged between 3 and 8 years with overweight or obesity. The intervention builds on the BMI[2] framework [40,52,65,70] and leverages the infrastructure of a large integrated health system to embed MI and centralized telehealth coaching into routine pediatric care. This trial's hybrid type 2 implementation-effectiveness design allows for simultaneous evaluation of clinical outcomes and implementation processes, generating insights that are both scientifically rigorous and operationally actionable.

Key Contributions and Innovations

WC4K introduces several innovations that distinguish it from previous pediatric obesity interventions. The integration of MI into routine pediatric care is supported by structured training, EMR-embedded tools, and a centralized coaching infrastructure. This approach enables consistent delivery of MI across clinics, regardless of local staffing or resources.

The use of telehealth-based coaching with flexible MI dosing allows for tailored support that adapts to family needs, availability, and readiness to change. This flexibility is particularly important for families facing logistical barriers to in-person care, such as transportation, work schedules, or caregiving responsibilities.

Digital tools, including wellness screeners, goal-setting worksheets, and self-monitoring logs, enhance family engagement while minimizing burden on pediatricians. These tools are embedded into existing workflows and accessible through EMR systems, supporting seamless integration into clinical practice.

The centralized coaching model ensures fidelity to MI principles and allows for ongoing supervision, quality assurance, and scalability. Coaches are trained to deliver MI-based counseling across a range of behavioral domains, supported by standardized protocols and fidelity monitoring.

Addressing Gaps in Pediatric Obesity Care

WC4K responds to well-documented barriers in pediatric obesity management. Many families underestimate the long-term risks associated with early weight gain or struggle to make sustained lifestyle changes without structured support [16,34]. At the same time, pediatricians often report limited training, time, and confidence in delivering effective obesity counseling [33-36]. WC4K addresses these challenges by combining brief MI-based engagement by pediatricians with centralized, telehealth-delivered coaching. This dual approach allows pediatricians to initiate behavior change conversations during routine well-child visits and refer families to trained health coaches for ongoing support.

Importantly, this trial targets clinics serving predominantly low-income and racial and ethnic minority population groups facing higher obesity prevalence and greater barriers to care. By embedding the intervention into routine care and minimizing additional burden on families and pediatricians, WC4K offers a more equitable and accessible model of care.

Balancing Intensity and Feasibility

The US Preventive Services Task Force recommends behavioral interventions with at least 26 contact hours to achieve meaningful outcomes [71]. This recommendation was reaffirmed in a 2024 evidence update, which found insufficient evidence supporting the effectiveness of interventions with fewer contact hours [72].

However, such high-intensity programs are often impractical for families and difficult to implement in primary care. Time constraints, competing demands, and limited resources can hinder both health care provider delivery and family participation.

WC4K tests a lower-dose model of MI-based coaching, with the option to extend support based on family needs. This flexible approach balances the need for effectiveness with the realities of clinical practice and family life, particularly in resource-constrained settings [35,38].

Tailoring and Implementation Science

Tailoring, defined here as the adaptation of the intervention through implementation science rather than individual-level customization, is a central feature of WC4K and a key strength of its pragmatic design [73-77].

At the pediatrician level, MI training and tools are adapted to fit within existing workflows, allowing pediatricians to initiate behavior change conversations during routine care without disrupting clinical efficiency.

At the family level, coaching is personalized based on readiness to change, cultural context, and logistical constraints. Coaches work with caregivers to identify goals that align with their values and circumstances, adjusting the intensity and focus of sessions as needed.

At the system level, the intervention is embedded within the infrastructure of a large integrated health system. EMR-integrated tools, centralized coaching, and leadership support facilitate rapid implementation and scalability.

Conclusions and Future Directions

WC4K represents a novel and pragmatic approach to addressing childhood obesity within pediatric primary care. By integrating MI-based coaching into routine care and leveraging telehealth

XSL•FO
RenderX

and EMR systems, the program addresses key barriers to implementation, scalability, and sustainability.

This trial is designed to evaluate both clinical effectiveness and implementation outcomes, with a focus on underserved populations disproportionately affected by pediatric obesity. If successful, WC4K could serve as a model for other health systems seeking to embed behavioral health interventions into pediatric workflows.

The findings from this trial will inform future efforts to scale and sustain lifestyle interventions in real-world settings. The centralized coaching infrastructure, flexible delivery model, and integration with EMR tools position WC4K for potential replication across diverse health care environments. Insights from the reach, effectiveness, adoption, implementation, and maintenance framework will guide adaptation and dissemination strategies, supporting broader public health impact.

Acknowledgments

The authors thank Edgar Ordonez, Jimmy Wannares, Diego Gomez-Porras, and Nancy Daghighian as well as all staff of the Center of Healthy Living, Kaiser Permanente, for their help in developing this protocol, improving the workflows, and assisting with data collection, particularly the data collected during coaching visits.

Funding

This project is funded by the National Institute of Diabetes and Digestive and Kidney Diseases (; principal investigator CK) and by Kaiser Permanente Southern California community funds.

Data Availability

The datasets generated and analyzed during this study are not publicly available due to patient privacy concerns and restrictions under the Health Insurance Portability and Accountability Act (HIPAA) but deidentified data are available from the corresponding author on reasonable request.

Authors' Contributions

CK conceptualized and designed the study, led protocol development, and oversaw all aspects of implementation and manuscript preparation. MS led the statistical design and analysis planning and contributed to data interpretation and manuscript drafting. JDV managed project operations, coordinated clinic workflows, and contributed to data collection and manuscript review. XL programmed the analytic datasets, supported statistical analyses, and contributed to data interpretation. ED developed and delivered motivational interviewing coaching protocols, contributed to fidelity monitoring, assisted with manuscript review, and contributed to data interpretation. DY contributed to study design, implementation strategy, data interpretation, and manuscript editing. PK provided clinical expertise in pediatric workflows and contributed to integrating the protocol into routine care. KR led the motivational interviewing training design and fidelity strategy and contributed to intervention development, data interpretation, and manuscript editing. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017-2018. NCHS Data Brief 2020 Feb(360):1-8 [FREE Full text] [Medline: 32487284]
2. Daniels SR, Arnett DK, Eckel RH, Gidding SS, Hayman LL, Kumanyika S, et al. Overweight in children and adolescents: pathophysiology, consequences, prevention, and treatment. Circulation 2005 Apr 19;111(15):1999-2012. [doi: 10.1161/01.CIR.0000161369.71722.10] [Medline: 15837955]
3. The Surgeon General's Vision for a Healthy and Fit Nation. Rockville, MD: Office of the Surgeon General (US); 2010.
4. Freedman DS, Mei Z, Srinivasan SR, Berenson GS, Dietz WH. Cardiovascular risk factors and excess adiposity among overweight children and adolescents: the Bogalusa Heart Study. J Pediatr 2007 Jan;150(1):12-7.e2. [doi: 10.1016/j.jpeds.2006.08.042] [Medline: 17188605]
5. Li C, Ford ES, Zhao G, Mokdad AH. Prevalence of pre-diabetes and its association with clustering of cardiometabolic risk factors and hyperinsulinemia among U.S. adolescents: National Health and Nutrition Examination Survey 2005-2006. Diabetes Care 2009 Feb;32(2):342-347 [FREE Full text] [doi: 10.2337/dc08-1128] [Medline: 18957533]
6. Dietz WH. Overweight in childhood and adolescence. N Engl J Med 2004 Feb 26;350(9):855-857. [doi: 10.1056/NEJMp048008] [Medline: 14985480]
7. Koskinen J, Magnussen CG, Sinaiko A, Woo J, Urbina E, Jacobs DRJ, et al. Childhood age and associations between childhood metabolic syndrome and adult risk for metabolic syndrome, type 2 diabetes mellitus and carotid intima media

XSL•FO
RenderX

thickness: the International Childhood Cardiovascular Cohort Consortium. J Am Heart Assoc 2017 Aug 16;6(8):e005632 [FREE Full text] [doi: 10.1161/JAHA.117.005632] [Medline: 28862940]
8. Koebnick C, Black MH, Wu J, Martinez MP, Smith N, Kuizon B, et al. High blood pressure in overweight and obese youth: implications for screening. J Clin Hypertens (Greenwich) 2013 Nov;15(11):793-805. [doi: 10.1111/jch.12199] [Medline: 24119024]
9. Guo SS, Chumlea WC. Tracking of body mass index in children in relation to overweight in adulthood. Am J Clin Nutr 1999 Jul;70(1):145S-148S. [doi: 10.1093/ajcn/70.1.145s] [Medline: 10393162]
10. Freedman DS, Patel DA, Srinivasan SR, Chen W, Tang R, Bond MG, et al. The contribution of childhood obesity to adult carotid intima-media thickness: the Bogalusa Heart Study. Int J Obes (Lond) 2008 May;32(5):749-756. [doi: 10.1038/sj.ijo.0803798] [Medline: 18227845]
11. Dunford LJ, Langley-Evans SC, McMullen S. Childhood obesity and adult cardiovascular disease risk: a systematic review. Int J Obes (Lond) 2010 Jan;34(1):18-28. [doi: 10.1038/ijo.2009.61] [Medline: 19434067]
12. Feng C, Osgood ND, Dyck RF. Low birth weight, cumulative obesity dose, and the risk of incident type 2 diabetes. J Diabetes Res 2018;2018:8435762 [FREE Full text] [doi: 10.1155/2018/8435762] [Medline: 29541643]
13. Wang LX, Filipp SL, Urbina EM, Gurka MJ, DeBoer MD. Longitudinal associations of metabolic syndrome severity between childhood and young adulthood: the Bogalusa Heart Study. Metab Syndr Relat Disord 2018 Jun;16(5):208-214 [FREE Full text] [doi: 10.1089/met.2017.0160] [Medline: 29584578]
14. Zhang H, Zhang T, Li S, Guo Y, Shen W, Fernandez C, et al. Long-term excessive body weight and adult left ventricular hypertrophy are linked through later-life body size and blood pressure: the Bogalusa Heart Study. Circ Res 2017 May 12;120(10):1614-1621 [FREE Full text] [doi: 10.1161/CIRCRESAHA.116.310421] [Medline: 28232594]
15. McGuire S. Institute of Medicine. 2012. Accelerating progress in obesity prevention: solving the weight of the nation. Washington, DC: the National Academies Press. Adv Nutr 2012 Sep 01;3(5):708-709 [FREE Full text] [doi: 10.3945/an.112.002733] [Medline: 22983849]
16. Hampl SE, Hassink SG, Skinner AC, Armstrong SC, Barlow SE, Bolling CF, et al. Clinical practice guideline for the evaluation and treatment of children and adolescents with obesity. Pediatrics 2023 Feb 01;151(2):e2022060640. [doi: 10.1542/peds.2022-060640] [Medline: 36622115]
17. Hales CM, Fryar CD, Carroll MD, Freedman DS, Ogden CL. Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007-2008 to 2015-2016. JAMA 2018 Apr 24;319(16):1723-1725 [FREE Full text] [doi: 10.1001/jama.2018.3060] [Medline: 29570750]
18. Stierman B, Ogden CL, Yanovski JA, Martin CB, Sarafrazi N, Hales CM. Changes in adiposity among children and adolescents in the United States, 1999-2006 to 2011-2018. Am J Clin Nutr 2021 Oct 04;114(4):1495-1504 [FREE Full text] [doi: 10.1093/ajcn/nqab237] [Medline: 34291279]
19. Resnicow K, McMaster F, Rollnick S. Action reflections: a client-centered technique to bridge the WHY-HOW transition in Motivational Interviewing. Behav Cogn Psychother 2012 Jul;40(4):474-480 [FREE Full text] [doi: 10.1017/S1352465812000124] [Medline: 22414686]
20. Resnicow K, McMaster F. Motivational Interviewing: moving from why to how with autonomy support. Int J Behav Nutr Phys Act 2012 Mar 02;9:19 [FREE Full text] [doi: 10.1186/1479-5868-9-19] [Medline: 22385702]
21. Schwenker R, Dietrich CE, Hirpa S, Nothacker M, Smedslund G, Frese T, et al. Motivational interviewing for substance use reduction. Cochrane Database Syst Rev 2023 Dec 12;12(12):CD008063. [doi: 10.1002/14651858.CD008063.pub3] [Medline: 38084817]
22. Almansour M, AlQurmalah SI, Abdul Razack HI. Motivational interviewing-an evidence-based, collaborative, goal-oriented communication approach in lifestyle medicine: a comprehensive review of the literature. J Taibah Univ Med Sci 2023 Oct;18(5):1170-1178 [FREE Full text] [doi: 10.1016/j.jtumed.2023.03.011] [Medline: 37206188]
23. Mercado J, Espinosa-Curiel IE, Martínez-Miranda J. Embodied conversational agents providing motivational interviewing to improve health-related behaviors: scoping review. J Med Internet Res 2023 Dec 08;25:e52097 [FREE Full text] [doi: 10.2196/52097] [Medline: 38064707]
24. Resnicow K, Sonneville KR, Naar S. The heterogeneity of MI interventions studies for treatment of obesity. Pediatrics 2018 Nov;142(5):e20182471 [FREE Full text] [doi: 10.1542/peds.2018-2471] [Medline: 30348752]
25. Davis MM, Gance-Cleveland B, Hassink S, Johnson R, Paradis G, Resnicow K. Recommendations for prevention of childhood obesity. Pediatrics 2007 Dec;120 Suppl 4:S229-S253. [doi: 10.1542/peds.2007-2329E] [Medline: 18055653]
26. Moyce SC, Bell JF. Receipt of pediatric weight-related counseling and screening in a national sample after the expert committee recommendations. Clin Pediatr (Phila) 2015 Dec;54(14):1366-1374. [doi: 10.1177/0009922815584216] [Medline: 25926668]
27. Spear BA, Barlow SE, Ervin C, Ludwig DS, Saelens BE, Schetzina KE, et al. Recommendations for treatment of child and adolescent overweight and obesity. Pediatrics 2007 Dec;120 Suppl 4:S254-S288. [doi: 10.1542/peds.2007-2329F] [Medline: 18055654]
28. Borrello M, Pietrabissa G, Ceccarini M, Manzoni GM, Castelnuovo G. Motivational interviewing in childhood obesity treatment. Front Psychol 2015 Nov 12;6:1732 [FREE Full text] [doi: 10.3389/fpsyg.2015.01732] [Medline: 26617555]

XSL•FO
RenderX

29. Lutaud R, Mitilian E, Forte J, Gentile G, Reynaud R, Truffet C, et al. Motivational interviewing for the management of child and adolescent obesity: a systematic literature review. BJGP Open 2023 Dec;7(4):BJGPO.2022.0145. [doi: 10.3399/BJGPO.2022.0145] [Medline: 37402547]
30. Barlow SE, Expert Committee. Expert committee recommendations regarding the prevention, assessment, and treatment of child and adolescent overweight and obesity: summary report. Pediatrics 2007 Dec;120 Suppl 4:S164-S192. [doi: 10.1542/peds.2007-2329C] [Medline: 18055651]
31. Barlow SE, Bobra SR, Elliott MB, Brownson RC, Haire-Joshu D. Recognition of childhood overweight during health supervision visits: does BMI help pediatricians? Obesity (Silver Spring) 2007 Jan;15(1):225-232 [FREE Full text] [doi: 10.1038/oby.2007.535] [Medline: 17228051]
32. Woo Baidal JA, Criss S, Goldman RE, Perkins M, Cunningham C, Taveras EM. Reducing Hispanic children's obesity risk factors in the first 1000 days of life: a qualitative analysis. J Obes 2015;2015:945918 [FREE Full text] [doi: 10.1155/2015/945918] [Medline: 25874127]
33. Klein JD, Sesselberg TS, Johnson MS, O'Connor KG, Cook S, Coon M, et al. Adoption of body mass index guidelines for screening and counseling in pediatric practice. Pediatrics 2010 Feb;125(2):265-272 [FREE Full text] [doi: 10.1542/peds.2008-2985] [Medline: 20083518]
34. Story MT, Neumark-Stzainer DR, Sherwood NE, Holt K, Sofka D, Trowbridge FL, et al. Management of child and adolescent obesity: attitudes, barriers, skills, and training needs among health care professionals. Pediatrics 2002 Jul;110(1 Pt 2):210-214. [Medline: 12093997]
35. Brown J, Weedn A, Gillaspy S. Childhood obesity management in primary care: a needs assessment of pediatricians to guide resource development. J Okla State Med Assoc 2014;107(9-10):518-522. [Medline: 25790601]
36. Jelalian E, Boergers J, Alday CS, Frank R. Survey of physician attitudes and practices related to pediatric obesity. Clin Pediatr (Phila) 2003 Apr;42(3):235-245. [doi: 10.1177/000992280304200307] [Medline: 12739922]
37. Resnicow K, Harris D, Wasserman R, Schwartz RP, Perez-Rosas V, Mihalcea R, et al. Advances in motivational interviewing for pediatric obesity: results of the brief motivational interviewing to reduce body mass index trial and future directions. Pediatr Clin North Am 2016 Jun;63(3):539-562 [FREE Full text] [doi: 10.1016/j.pcl.2016.02.008] [Medline: 27261549]
38. Moore E, Vallejo J, Sidell MA, Li X, Delacroix E, Young DR, et al. Training in motivational interviewing for pediatricians was associated with a higher uptake of health behavior and lifestyle treatment: results from the KP Wellness Coaching for Families and Kids (WC4K) Program. Patient Educ Couns 2025 Aug;137:109173 [FREE Full text] [doi: 10.1016/j.pec.2025.109173] [Medline: 40412364]
39. Taveras EM, Gortmaker SL, Hohman KH, Horan CM, Kleinman KP, Mitchell K, et al. Randomized controlled trial to improve primary care to prevent and manage childhood obesity: the High Five for Kids study. Arch Pediatr Adolesc Med 2011 Aug;165(8):714-722 [FREE Full text] [doi: 10.1001/archpediatrics.2011.44] [Medline: 21464376]
40. Resnicow K, McMaster F, Woolford S, Slora E, Bocian A, Harris D, et al. Study design and baseline description of the BMI2 trial: reducing paediatric obesity in primary care practices. Pediatr Obes 2012 Feb;7(1):3-15 [FREE Full text] [doi: 10.1111/j.2047-6310.2011.00001.x] [Medline: 22434735]
41. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. BMJ 2015 May 08;350:h2147 [FREE Full text] [doi: 10.1136/bmj.h2147] [Medline: 25956159]
42. Zwarenstein M, Treweek S, Loudon K. PRECIS-2 helps researchers design more applicable RCTs while CONSORT Extension for Pragmatic Trials helps knowledge users decide whether to apply them. J Clin Epidemiol 2017 Apr;84:27-29. [doi: 10.1016/j.jclinepi.2016.10.010] [Medline: 28137671]
43. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
44. McGlynn EA, Lieu TA, Durham ML, Bauck A, Laws R, Go AS, et al. Developing a data infrastructure for a learning health system: the PORTAL network. J Am Med Inform Assoc 2014;21(4):596-601. [doi: 10.1136/amiajnl-2014-002746] [Medline: 24821738]
45. Kanter MH, Lindsay G, Bellows J, Chase A. Complete care at Kaiser Permanente: transforming chronic and preventive care. Jt Comm J Qual Patient Saf 2013 Nov;39(11):484-494. [doi: 10.1016/s1553-7250(13)39064-3] [Medline: 24294676]
46. No author. 2015 Recommendations for preventive pediatric health care committee on practice and ambulatory medicine and bright futures periodicity schedule workgroup. Pediatrics 2015 Sep;136(3):e727. [doi: 10.1542/peds.2015-2009] [Medline: 26324870]
47. Committee on Practice and Ambulatory Medicine, Bright Futures Periodicity Schedule Workgroup. 2017 recommendations for preventive pediatric health care. Pediatrics 2017 Apr;139(4):e20170254. [doi: 10.1542/peds.2017-0254] [Medline: 28213605]
48. Hagan JFJ, Shaw JS, Duncan PM. Bright Futures Guidelines for Health Supervision of Infants, Children, and Adolescents, Fourth Edition. Itasca, IL: American Academy of Pediatrics; 2017.
49. Kuczmarski RJ, Ogden CL, Guo SS, Grummer-Strawn LM, Flegal KM, Mei Z, et al. 2000 CDC growth charts for the United States: methods and development. Vital Health Stat 11 2002 May(246):1-190 [FREE Full text] [Medline: 12043359]

XSL•FO
RenderX

50. Ogden CL, Freedman DS, Hales CM. CDC extended BMI-for-age percentiles versus percent of the 95th percentile. Pediatrics 2023 Sep 01;152(3):e2023062285. [doi: 10.1542/peds.2023-062285] [Medline: 37615069]
51. Resnicow K, McMaster F, Bocian A, Harris D, Zhou Y, Snetselaar L, et al. Motivational interviewing and dietary counseling for obesity in primary care: an RCT. Pediatrics 2015 Apr;135(4):649-657 [FREE Full text] [doi: 10.1542/peds.2014-1880] [Medline: 25825539]
52. Resnicow K, Delacroix E, Sonneville KR, Considine S, Grundmeier RW, Shu D, et al. Outcome of BMI2+: motivational interviewing to reduce BMI through primary care AAP PROS practices. Pediatrics 2024 Jan 01;153(2):e2023062462. [doi: 10.1542/peds.2023-062462] [Medline: 38282541]
53. Rollnick S, Butler CC, Stott N. Helping smokers make decisions: the enhancement of brief intervention for general medical practice. Patient Educ Couns 1997 Jul;31(3):191-203. [doi: 10.1016/s0738-3991(97)01004-5] [Medline: 9277242]
54. Rollnick S, Heather N, Gold R, Hall W. Development of a short 'readiness to change' questionnaire for use in brief, opportunistic interventions among excessive drinkers. Br J Addict 1992 May;87(5):743-754. [doi: 10.1111/j.1360-0443.1992.tb02720.x] [Medline: 1591525]
55. Rollnick S, Mason P, Butler C. Health Behavior Change: A Guide for Practitioners. London, UK: Churchill Livingstone; 1999.
56. Resnicow K, Gobat N, Naar S. Intensifying and igniting change talk in motivational interviewing: a theoretical and practical framework. Eur Health Psychol 2015;17(3):102-110 [FREE Full text]
57. Rollnick S, Butler CC, McCambridge J, Kinnersley P, Elwyn G, Resnicow K. Consultations about changing behaviour. BMJ 2005 Oct 22;331(7522):961-963. [doi: 10.1136/bmj.331.7522.961] [Medline: 16239696]
58. Rollnick S, Miller WR, Butler C. Motivational Interviewing in Health Care: Helping Patients Change Behavior. New York, NY: Guilford Press; 2008.
59. McMaster F, Resnicow K. Validation of the one pass measure for motivational interviewing competence. Patient Educ Couns 2015 Apr;98(4):499-505. [doi: 10.1016/j.pec.2014.12.014] [Medline: 25624118]
60. Freedman DS, Berenson GS. Tracking of BMI z scores for severe obesity. Pediatrics 2017 Sep;140(3):e20171072. [doi: 10.1542/peds.2017-1072] [Medline: 28830920]
61. Freedman DS, Butte NF, Taveras EM, Goodman AB, Ogden CL, Blanck HM. The limitations of transforming very high body mass indexes into z-scores among 8.7 million 2- to 4-year-old children. J Pediatr 2017 Sep;188:50-6.e1 [FREE Full text] [doi: 10.1016/j.jpeds.2017.03.039] [Medline: 28433203]
62. Freedman DS, Goodman AB, King RJ, Blanck HM. Tracking of obesity among 2- to 9-year-olds in an electronic heath record database from 2006 to 2018. Obes Sci Pract 2020 Feb 08;6(3):300-306 [FREE Full text] [doi: 10.1002/osp4.407] [Medline: 32523719]
63. Austin PC, Stuart EA. Moving towards best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. Stat Med 2015 Dec 10;34(28):3661-3679. [doi: 10.1002/sim.6607] [Medline: 26238958]
64. Donner A. Sample size requirements for stratified cluster randomization designs. Stat Med 1992 Apr;11(6):743-750. [doi: 10.1002/sim.4780110605] [Medline: 1594813]
65. Wright ME, Delacroix E, Sonneville KR, Considine S, Proctor T, Steffes J, et al. Reducing paediatric overweight and obesity through motivational interviewing: study protocol for a randomised controlled trial in the AAP PROS research network. BMJ Open 2020 Jul 28;10(7):e035720 [FREE Full text] [doi: 10.1136/bmjopen-2019-035720] [Medline: 32723736]
66. Freedman DS, Woo JG, Ogden CL, Xu JH, Cole TJ. Distance and percentage distance from median BMI as alternatives to BMI z score. Br J Nutr 2020 Sep 14;124(5):493-500 [FREE Full text] [doi: 10.1017/S0007114519002046] [Medline: 31439056]
67. Cunningham SA, Hardy ST, Jones R, Ng C, Kramer MR, Narayan KM. Changes in the incidence of childhood obesity. Pediatrics 2022 Aug 01;150(2):e2021053708 [FREE Full text] [doi: 10.1542/peds.2021-053708] [Medline: 35789417]
68. Cunningham SA, Kramer MR, Narayan KM. Incidence of childhood obesity in the United States. N Engl J Med 2014 Apr 24;370(17):1660-1661 [FREE Full text] [doi: 10.1056/NEJMc1402397] [Medline: 24758623]
69. Emmerich SD, Fryar CD, Stierman B, Gu Q, Afful J, Ogden CL. Trends in obesity-related measures among US children, adolescents, and adults. JAMA 2025 Mar 25;333(12):1082-1084. [doi: 10.1001/jama.2024.27676] [Medline: 39946125]
70. Fiks AG, Scheindlin B, Shone L. 30th anniversary of pediatric research in office settings (PROS): an invitation to become engaged. Pediatrics 2016 Sep;138(3):e20161126. [doi: 10.1542/peds.2016-1126] [Medline: 27485697]
71. O'Connor EA, Evans CV, Burda BU, Walsh ES, Eder M, Lozano P. Screening for obesity and intervention for weight management in children and adolescents: evidence report and systematic review for the US Preventive Services Task Force. JAMA 2017 Jun 20;317(23):2427-2444. [doi: 10.1001/jama.2017.0332] [Medline: 28632873]
72. O'Connor EA, Evans CV, Henninger M, Redmond N, Senger CA. Interventions for weight management in children and adolescents: updated evidence report and systematic review for the US Preventive Services Task Force. JAMA 2024 Jul 16;332(3):233-248. [doi: 10.1001/jama.2024.6739] [Medline: 38888913]
73. Albers B, Rapley T, Nilsen P, Clack L. Editorial: tailoring in implementation science. Front Health Serv 2023;3:1233597 [FREE Full text] [doi: 10.3389/frhs.2023.1233597] [Medline: 37389149]

XSL•FO
RenderX

74. Griffith DM, Efird CR, Baskin ML, Webb Hooper M, Davis RE, Resnicow K. Cultural sensitivity and cultural tailoring: lessons learned and refinements after two decades of incorporating culture in health communication research. Annu Rev Public Health 2024 May;45(1):195-212 [FREE Full text] [doi: 10.1146/annurev-publhealth-060722-031158] [Medline: 37931182]
75. McHugh SM, Riordan F, Curran GM, Lewis CC, Wolfenden L, Presseau J, et al. Conceptual tensions and practical trade-offs in tailoring implementation interventions. Front Health Serv 2022 Nov 17;2:974095 [FREE Full text] [doi: 10.3389/frhs.2022.974095] [Medline: 36925816]
76. McHugh SM, Riordan F, Kerins C, Curran G, Lewis CC, Presseau J, et al. Understanding tailoring to support the implementation of evidence-based interventions in healthcare: the CUSTOMISE research programme protocol. HRB Open Res 2023 Jan 25;6:7 [FREE Full text] [doi: 10.12688/hrbopenres.13675.1] [Medline: 37361339]
77. Metz A, Kainz K, Boaz A. Intervening for sustainable change: tailoring strategies to align with values and principles of communities. Front Health Serv 2022 Jan 18;2:959386 [FREE Full text] [doi: 10.3389/frhs.2022.959386] [Medline: 36925846]

Abbreviations

AAP: American Academy of Pediatrics
BMI2: Brief Motivational Interviewing to Reduce BMI
CDC: Centers for Disease Control and Prevention
DID: difference in difference
EMR: electronic medical record
KPSC: Kaiser Permanente Southern California
MI: motivational interviewing
SMD: standardized mean difference
WC4K: Wellness Coaching for Kids and Families

Edited by J Sarvestan; submitted 14.Oct.2025; this is a non–peer-reviewed article;accepted 30.Nov.2025; published 10.Feb.2026.

Please cite as:
Koebnick C, Sidell M, Vallejo JD, Li X, Delacroix E, Young D, Kunani P, Resnicow K
Evaluating the Effectiveness of the Motivational Interviewing–Based Wellness Coaching for Kids and Families (WC4K) Program in Pediatric Obesity Care: Protocol for a Cluster Randomized Pragmatic Trial
JMIR Res Protoc 2026;15:e78792
URL: https://www.researchprotocols.org/2026/1/e78792
doi:10.2196/78792
PMID:

©Corinna Koebnick, Margo Sidell, Jessica D Vallejo, Xia Li, Emerson Delacroix, Deborah Young, Poornima Kunani, Ken Resnicow. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Feasibility and Acceptability of Collaborative Augmented Reality for Older Adults and Companions: Protocol for a Randomized Controlled Trial

Nilanjan Sarkar[1], PhD; Judith A Tate[2], PhD; Mahrukh Tauseef[3], MS ECE, BSEE; Akshith Ullal[4], PhD; Cathy A Maxwell[5], PhD; Lisa A Juckett[6], PhD; Douglas W Scharre[7], MD; Alai Tan[2], MD, PhD; Rendong Zhang[3], MSE; Zalen Ingram[1], BSCS; Lorraine C Mion[2], PhD

[1]Mechanical Engineering and Computer Science, School of Engineering, Vanderbilt University, Nashville, TN, United States

[2]College of Nursing, The Ohio State University, Columbus, OH, United States

[3]Electrical and Computer Engineering, School of Engineering, Vanderbilt University, Nashville, TN, United States

[4]Intelligent Clinical Care Center, University of Florida, Gainesville, FL, United States

[5]College of Nursing, University of Utah, Salt Lake City, UT, United States

[6]School of Health & Rehabilitation Sciences, College of Medicine, The Ohio State University, Columbus, OH, United States

[7]Division of Cognitive Neurology, Department of Neurology, College of Medicine, The Ohio State University, Columbus, OH, United States

Corresponding Author:
Nilanjan Sarkar, PhD
Mechanical Engineering and Computer Science
School of Engineering
Vanderbilt University
2400 Highland Avenue
Nashville, TN, 37212
United States
Phone: 1 615 343 7219
Email: nilanjan.sarkar@vanderbilt.edu

Abstract

Background: Loneliness and social isolation are common among older adults and are associated with dire consequences. Studies using interactive communication technology (ICT) interventions with older adults have yielded mixed results. Advancements in collaborative head-mounted display augmented reality (HMD-AR) can provide older adults who are geographically distant from their families with a more diverse range of interactive activities, thus offering greater potential to enhance social connection.

Objective: We examined the feasibility and acceptability of cocreated collaborative HMD-AR activities and 2D ICT (eg, Zoom) activities among older adults and their family members and close friends.

Methods: In total, 8 pairs of older adults and their designated companion (family or friend) from the greater Nashville area were randomized to the HMD-AR or 2D ICT groups. Eligibility criteria for older adults included being 60 years or older, being able to tolerate HMD-AR, being cognitively and physically able to participate, and having a companion willing to participate. For long-term care settings, participants must have been a resident for at least 3 months. All participants lived within a 1-hour driving distance from the investigators' university. Each older adult–companion pair participated in eight 30-minute sessions over 4 weeks. Participants randomized to the HMD-AR group had photorealistic avatars created; they participated in collaborative activities (ie, fireplace decoration, checkers) that were cocreated in an earlier study. Those randomized to the 2D ICT group had the opportunity to play virtual checkers or house décor games. Engineers remained on-site for all participants to assist as needed. The primary outcome was the feasibility of the study processes (ie, recruitment, retention, and data collection) and the technology (ie, viability, usability, comfort, ease of use, and acceptability). The study used face-to-face questionnaires and observations to collect data.

Results: Funding was received from the National Institute on Aging in August 2022 and from the National Science Foundation in October 2022. This study was approved by the local institutional review board. Following the initial design and testing of the HMD-AR activities, recruitment and data collection for the feasibility and acceptability study began in April 2024 and were completed in May 2025. Of the 8 enrolled older-companion pairs, 8 (100%) completed all sessions. The final data acquisition has been completed, and data cleaning is currently ongoing. Results are intended to be published in 2026.

XSL•FO
RenderX

Conclusions: To our knowledge, this study is the first collaborative augmented reality study using photorealistic avatars between older adults and their family members or friends. Our study will determine whether the use of HMD-AR is feasible, and the results of this pilot study will inform a full-scale randomized controlled trial evaluating the efficacy of the intervention to reduce loneliness among older adults.

Trial Registration: ClinicalTrials.gov NCT06179225; https://clinicaltrials.gov/study/NCT06179225

International Registered Report Identifier (IRRID): DERR1-10.2196/83864

(JMIR Res Protoc 2026;15:e83864) doi:10.2196/83864

KEYWORDS

older adults; loneliness; photorealistic avatars; interactive communication technology; social connection

Introduction

Background

Loneliness and Social Isolation Are Common Health Problems Among Older Adults

Social connection is a critical health determinant in preventing loneliness (the subjective feeling of being alone or isolated) and social isolation (the objective measure of social contact, communication, or social activities), both of which are common health problems among older adults [1-4]. Among community-dwelling older adults, up to one-fourth experience social isolation, and up to 46% report feeling lonely [2-4]. Among older adults residing in long-term care (LTC) facilities, such as assisted living, social isolation occurs in 4% of cases, moderate loneliness in 61%, and severe loneliness in 35% [5-8]. Social isolation and loneliness have significant adverse consequences, including increased mortality, cardiovascular disease, depression, suicide, cognitive and physical decline, reduced quality of life, and increased health care use [4,9-11]. Numerous recommendations have been made to alleviate social isolation and loneliness among older adults, which include technology-based solutions [3,4,12].

Social Presence as the Underlying Mechanism in Effective Interactive Communication Technology

Interactive communication technology (ICT) interventions, such as social media and video-mediated visits, have been examined for their effects on social connection, social isolation, and loneliness among older adults [4,13-19]. Results have been mixed, and the varying study designs, ICT mediums, settings, communication partners (eg, family members, friends, volunteers, or groups), measures, and outcomes make it difficult to draw firm conclusions. One potential explanation for the mixed results on social connection is the lack of attention to social presence within ICT interventions. In a recent systematic review of interactive technology interventions with older adults, few studies had a theoretical basis, and none examined social presence [19].

In information communication technology, social presence is the subjective feeling that other real people are present, involved, and connected within the mediated digital environment, influencing the warmth and personal connection felt during the interaction [20-22]. Satisfaction with ICTs is largely based on the quality of the social presence afforded. With its foundation in theories of interpersonal communication and symbolic interactionism, social presence is strongly correlated with the awareness of others and with a sense of connection with them [23]. Social presence is a transient state that varies with the ICT used, familiarity with the other person, content, environment, and context [21,23,24].

What Is Augmented Reality and Why Use It With Older Adults?

Augmented reality (AR) superimposes digital elements onto the real world, allowing real-time interaction between the user and the digital elements. Unlike virtual reality (VR), which immerses the person into a synthetic environment, AR allows the user to maintain visualization of their physical environment [25]. AR terminology and type vary [26-28]; our focus is on head-mounted display AR (HMD-AR), which projects virtual content onto some portion of the real world, also referred to as mixed reality [29,30].

Advantages of HMD-AR over HMD-VR include the following: (1) HMD-AR overlays synthetic elements onto the real world, thus allowing the user to move about safely because room obstacles are apparent; (2) it is less resource intensive because only virtual objects need to be created and overlaid onto a room setting; (3) it provides greater depth perception, enabling the user to see one's own body, the interacting person's body, and physical objects, resulting in faster task completion; and (4) it is associated with fewer side effects, including reduced motion sickness, digital eye strain, or isolation effects [30-35]. Importantly, older adults generally accept HMDs [35,36].

With increased computing power in newer AR-based HMD devices, AR is seen as a viable alternative for developing interaction-based interventions for older adults, primarily due to its increased immersion and social presence [22]. The latest AR-based HMDs have increased field of vision, seamless hand and eye tracking, and robust workspace localization. Importantly, users are able to wear these HMDs with their eyeglasses on, which is an important requirement when dealing with older adults.

AR Studies Involving Older Adults

Although HMD-AR technology has matured, research examining its applications among adults aged 65 years and older is in its early stages. Existing study designs are primarily nonexperimental, consist of small sample sizes, focus on hardware and software development, and include limited activities, such as park design [37]; exercise, balance training, and fall prevention [38-40]; and cognitive training and e-learning

XSL•FO
RenderX

[27,40,41]. One study used augmented VR (a spectrum combining AR and VR elements) to examine a joint meal among 3 older adults [42]. Few studies have involved older adults with cognitive impairment [41].

All these previous studies reveal the promise for HMD-AR; most older adults accepted the technology and experienced few side effects. Suggestions for improvement included a lighter-weight HMD type that fits over eyeglasses, a more realistic avatar design, larger synthetic objects, enhanced user training, and the incorporation of varying levels of difficulty.

Preliminary Work

As part of the overall grant funding, we conducted several sequential studies before implementing the pilot feasibility and acceptability study.

Older and Younger Adults' Perceptions of AR Photorealistic Avatars as a Viable Medium for Interpersonal Communication

Given that AR as an ICT has been understudied, we examined perceptions of older adults (n=31) and younger adults (n=31) using HMD-AR with photorealistic avatars as a medium for social communication [43]. In addition, we examined the extent to which photorealistic avatars could portray 6 common emotions compared to video clips of a real person.

To assess participants' perceptions, older adults (n=31) and younger adults (n=31) interacted with a volunteer's 3D photorealistic avatar using a structured conversational activity through an HMD. Participants rated the quality of HMD-AR communication based on physical and human realism, comfort while talking to the avatar, and the degree of social presence offered by HMD-AR. Participants then identified 6 basic emotions exhibited by (1) video clips of a real person and (2) an animation of their 3D photorealistic avatars via an HMD. Each participant viewed 36 videos and avatar stimuli. Subgroup analyses were conducted by age group.

Participants reported a positive communication experience with the 3D photorealistic avatar, with older adults rating the quality higher than young adults. Ratings were generally lowest for perceptions of how life-like the avatar appeared (68% among younger adults and 61% among older adults). Most participants were able to accurately identify emotions displayed by the 3D photorealistic avatars, although younger adults outperformed older adults (93% accuracy versus 80% accuracy).

Cocreation of Collaborative HMD-AR Activities

Given the range of physical and cognitive impairments among older adults, especially for those who reside in LTC settings, we used an iterative participatory design methodology involving LTC older adults (n=8), staff (n=5), and family members (n=2) to develop collaborative HMD-AR activities [44]. Several factors were considered, including older adults' ability to tolerate HMD-AR for up to 30 minutes without experiencing adverse effects, preferences for HMD-AR activities, and ability to navigate the HMD-AR environment and manipulate AR objects. All iterative participatory design activities took place at an LTC in sequential visits over 6 months. At each visit, participants interacted with the HMD-AR activities and provided feedback that was used to refine the specific modes of interaction, game or activity logic, and user interface elements. We identified the need to tailor AR activities for older adults who experienced a range of physical limitations, including impaired fine motor control (eg, grasping and pinching), cognitive limitations (eg, reduced ability to verbalize thoughts), and sensory impairments (eg, difficulty seeing or reading text). A number of modifications have enabled easier interaction for older adults. Ultimately, 2 collaborative HMD-AR activities were developed for the pilot randomized controlled trial (RCT): fireplace decoration and checkers.

Specific Aims

The primary aim of this pilot study was to examine the feasibility and acceptability of cocreated collaborative HMD-AR activities using photorealistic avatars. To evaluate the feasibility and acceptability of the study procedures for a future RCT comparing HMD-AR with photorealistic avatars versus 2D ICT (eg, Zoom), we used an RCT design.

Methods

Study Design

This study was a feasibility and acceptability evaluation of older adults' use of HMD-AR compared with 2D ICT (ClinicalTrials.gov NCT06179225). To inform a future RCT, we designed the pilot as a nonblinded 2-arm RCT (Figure 1).

XSL•FO
RenderX

Figure 1. Study flow diagram.

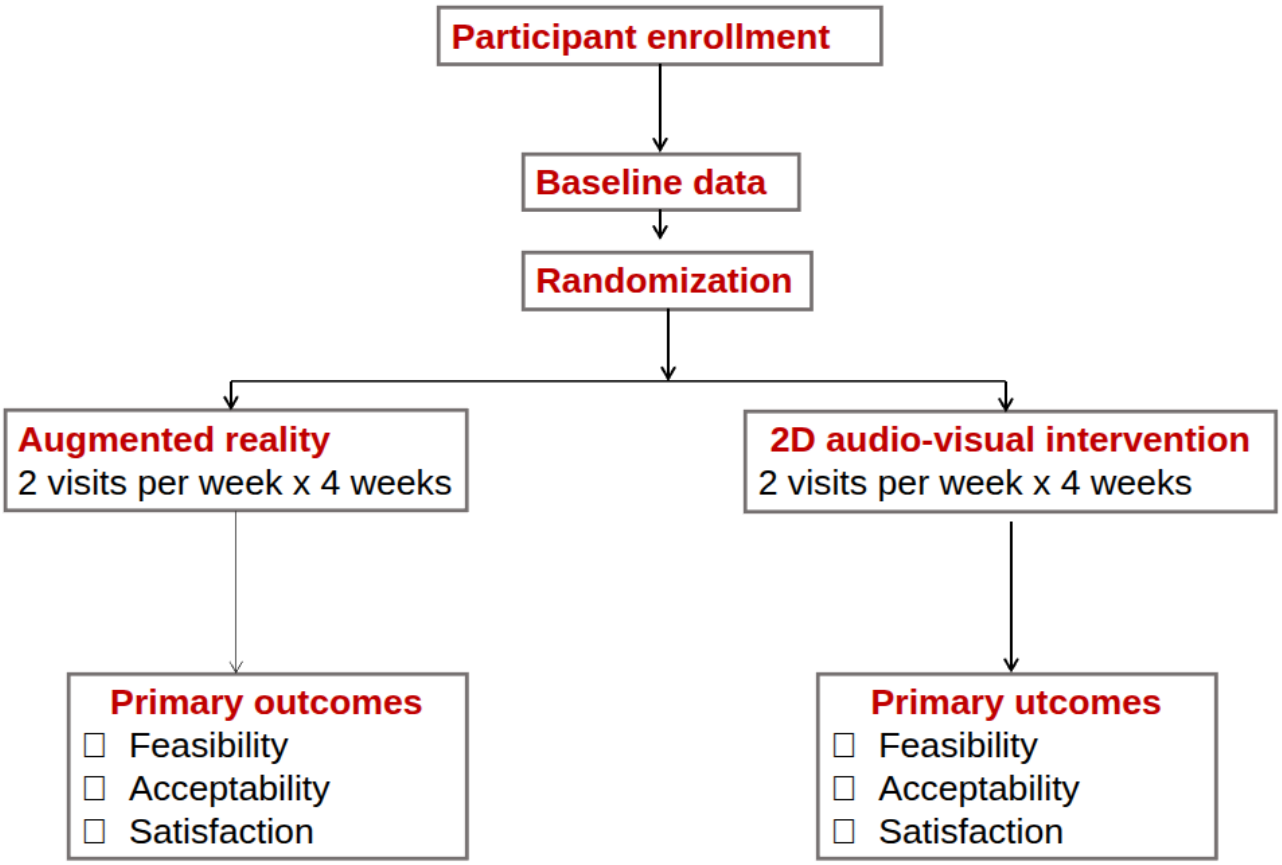

Participants were recruited from April 2024 through May 2025. The primary aim was to examine the feasibility and acceptability of the 2 ICTs. We enrolled 8 pairs of older adults and their designated companions (ie, designated family members or friends). After completing baseline assessments, 2 to 4 pairs were enrolled at a time because of logistical constraints. Randomization was conducted at a 1:1 ratio. The intervention consisted of 8 sessions delivered over 4 weeks (2 sessions per week). Baseline evaluations were conducted at week 0. Social presence was evaluated after the second and eighth sessions. Feasibility and acceptability outcomes were evaluated at completion of the study.

Study Design Rationale

We targeted older adults, with and without cognitive impairment, who resided at home or in LTC settings because of the high prevalence of loneliness or social isolation in these populations, resulting in further cognitive and physical decline [4-7]. Reviews have yielded mixed findings on whether men or women are more likely to report loneliness [4,6,45]. Age does not appear to have an association with levels of social presence in ICT; however, women tend to report higher levels of social presence when using ICT [24].

We used a 2D audio-visual intervention as the comparator because it has been extensively used since the COVID-19 pandemic. Although it provides audio and visual communication, it limits the quality of the interaction and has produced mixed results among older adults residing in the community [4,13-15,46,47]. Our focus is on known family members or close friends because of older adults' desire to maintain meaningful long-term relationships [48].

We examined feasibility and acceptability because these factors are highly salient for adoption and intervention success [49,50]. An RCT design was chosen for its robustness in conducting future intervention studies. We chose 2 visits per week for a 4-week duration to allow sufficient exposure to assess feasibility and acceptability. There are several HMD products; at the time of this study, we chose HoloLens 2 (Microsoft Corporation) because of its robustness, widespread use (eg, US Army with >80,000 hours of soldier feedback) [51], usability among older adults [52], and ability to fit over eyeglasses.

Ethical Considerations

This study protocol was approved by the Vanderbilt University Medical Center Institutional Review Board (IRB; 221150). Written informed consent was obtained in person from all participants before data collection and study procedures. To ensure that potential participants understood the study purpose, risks, and benefits, we administered the University of San Diego Brief Protection of Human Subjects Capacity to Consent instrument [53].

Research assistants (RAs) underwent training to recognize signs of frustration, anxiety, or stress displayed by participants; any sign of discomfort resulted in termination of the session. RAs were trained in procedures to protect privacy and confidentiality. Participants were assigned unique study ID numbers, and all deidentified data were stored in a secure password-protected database.

XSL•FO
RenderX

Participants received financial compensation for study activities, including the interactive sessions and data collection procedures, to acknowledge their time and reduce attrition. An external safety officer, designated by the National Institute on Aging (NIA), reviewed and approved the study protocol and informed consent documents before the start of the study. Biannual reports were provided to both the safety officer and the NIA program officer to review progress and any untoward events.

Conceptual Framework

Several frameworks guided the study (Figure 2). First, social presence theory postulates that the ICT properties and the persons' perceptions, behaviors, or attitudes regarding the physical presence of other people within the ICT environment affect the quality of the interaction [20,21,23]. We compared 2 ICTs, HMD-AR and 2D audio-video communication. Second, we applied the engagement framework proposed by Cohen-Mansfield [54-56], which emphasizes the role of social interactions, activities, and environment in shaping engagement. Third, the social connection framework by Holt-Lunstad et al [1] guided our focus on the functional and qualitative aspects of ICT visits. Finally, we adapted the unified theory of acceptance and use of technology framework [57] to guide the evaluation of users' acceptance and use of the ICTs. Individual characteristics such as age, gender, cognition, relationship with designated companion, and social network can impact acceptance and use of ICT [57], the type and quality of engagement [58], and loneliness—our future condition of interest [4].

Figure 2. Conceptual framework guiding the study design and intervention. ICT: interactive communication technology.

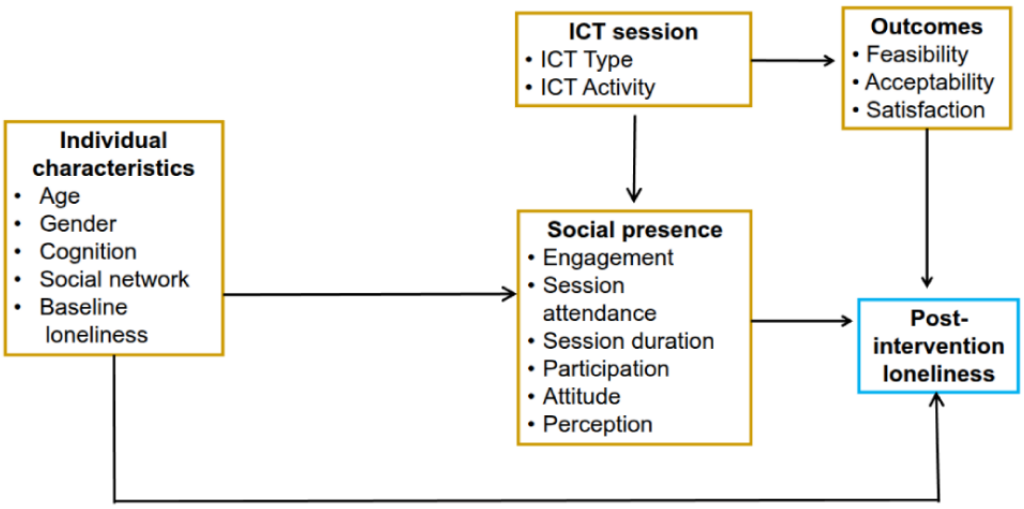

Participants and Setting

Overview

This study took place in the greater Nashville, Tennessee area. Eligibility criteria for older adults included being aged 60 years or older and living within a 1-hour drive from the principal investigator's office. If the older adults resided in LTC settings, they must have been residents for 3 months or more. Exclusion criteria included severe cognitive, sensory, or physical impairments that impeded participation; major psychiatric disorders; inability to provide assent; inability to understand or speak English; and acute or terminal illness.

Each older adult recruited 1 companion (family member or close friend). Companions had to be aged 18 years or older and willing to participate in both HMD-AR and 2D audio-video visits and related study procedures. For both older adults and their companions, adequate internet bandwidth was necessary. Eligible older adult–companion pairs were randomized at a 1:1 ratio to ICT type using a computer-generated randomization scheme after baseline measures were completed.

Recruitment

Using IRB-approved communication scripts, flyers were posted in public areas, including Vanderbilt University parking garages, local community and older adult centers, churches, and libraries. RAs provided several in-person demonstrations at local centers. Emails were sent to several assisted living facilities located near Vanderbilt University. Interested older adults and family members contacted the investigator team directly. RAs followed the IRB-approved informed consent document to explain the study either in person or by telephone. A standardized screening checklist was completed to ensure eligibility.

Sample Size

There is no precise sample size for an early-stage proof-of-concept technology feasibility study. Nielsen and Landauer [59] found that 5 users uncovered 85% of usability problems with the technology. Our target was to enroll 12 pairs of older adults and designated companions for a sample size of 24, as the recommended sample size for pilot studies is 20 to 30 [60,61]. Our final sample of 8 pairs (16 participants) with 8 sessions per pair (64 sessions) was judged sufficient to address the primary aim of feasibility and acceptability of ICT in the home setting.

HMD-AR and 2D Audio-Visual Interventions

For both interventions, the internet bandwidth within the homes of older adults and family members was measured, and a hot spot was used when necessary to facilitate data streaming. A minimum download speed of 80 Mbps was required for data communication. For each session, a trained RA engineer was

XSL•FO
RenderX

physically present in both the older adults' and companions' homes to troubleshoot for technological issues during the sessions. The RA trained the participants in establishing the ICT (HMD-AR or 2D audio-visual intervention) connection and assisted participants as needed. The RA remained in the room to monitor for technological issues and maintain field notes. Session duration was at the discretion of the older adult and family member or close friend, but it did not exceed 60 minutes.

HMD-AR Intervention

Creation of Photorealistic Avatars

We created photorealistic avatars for all participants randomized to the HMD-AR study arm. A mobile app, Polycam, was used to take multiple photos of the participant's face using a 360° view. Polycam constructed a 3D model from these photos that was then used to create a MetaHuman on Unreal Engine, matching the facial geometry and the facial features of the participant. However, MetaHuman Creator is limited in its ability to match the features exactly; therefore, the RAs added manual touch-ups to ensure the avatar looked similar to the participant. After creating the first version of the avatar, a video call was arranged with the participants to share their avatar with them and receive feedback. Participants were able to pick their avatar's clothing, hair, body shape, and accessories, and apply makeup. If a participant was not satisfied with the likeness of their avatar, the process was repeated, and more manual touch-ups were added. A second meeting was arranged with the participants to share the new version. For most participants, it took 1 meeting to finalize their avatars.

Experimental Setup

Participants, in consultation with the RAs, determined the physical space for conducting the sessions. General criteria for physical space were that it was free from clutter, furniture, and objects to allow for clear overlays of virtual objects and photorealistic avatars and located in an area of low traffic to minimize session interruptions. For each participant, equipment included a Kinect camera (Microsoft Corporation), a microphone, an HMD, and a laptop on which the data sharing and the activity ran and were remotely displayed on each participant's HMD. An Alienware (Dell Technologies Inc) laptop with a 3060 Graphics Processing Unit was used due to the high-performance demand of the AR activity. The RAs ensured that both the older adult and their family members or close friends were connected before moving to an unobtrusive viewing area. Upon completion of the session, the RA dismantled all equipment and cleaned the HMD-AR with disinfectant wipes.

Collaborative AR Activity Sessions

Before the start of the experimental sessions, RAs provided one-on-one training and orientation to the HMD-AR to ensure that the participants were able to navigate the system. In addition, a hard-copy, 1-page quick reference guide was provided to each participant. Each older adult–companion pair participated in 2 sessions per week for 4 weeks (total of 8 sessions). Two collaborative AR activities were available to the participants [44]: fireplace decoration and checkers. If the fireplace activity was selected, participants could choose from a variety of virtual objects to decorate the fireplace mantel, hearth, or nearby wall. For the checkers activity, each participant sat at a table, and the checkerboard and checkers appeared as a virtual object in space. Participants were able to see each other's photorealistic avatars during both activities. Upon completion of each session, the participants would let the RAs know who turned off the AR display and stopped the data communication.

2D Audio-Visual Intervention

Experimental Setup

Participants, in consultation with the RAs, determined the physical space for conducting the sessions, preferably in an area of low traffic to minimize session interruptions. If the participants were placed in the 2D ICT group, they could use their PC. In the event that participants did not own a device, a laptop was provided for each session. The RAs ensured that both the older adult and their companion were connected before moving to an unobtrusive viewing area. Upon completion of the session, the RAs recovered any loaned devices.

Collaborative 2D Audio-Video Sessions

For this study, we used the Zoom videoconferencing application for all sessions. Before the start of the experimental sessions, RAs provided 1:1 training and orientation to using the laptop device, if needed, especially for older adults unaccustomed to tablets or computers and logging in to Zoom. Participants were provided with links to an online checkers game and an online room decoration activity that they could engage in during the session. A 1-page quick reference guide was provided to each participant.

Ensuring Intervention Fidelity

We incorporated recommendations from the National Institutes of Health Behavior Change Consortium to maximize treatment fidelity across RAs and over time [62]. In brief, we used theoretical models to guide our study, conducted standardized training for RAs with ongoing reviews, and used standardized protocols. Issues were discussed at the weekly investigator-RA meetings. RAs completed field notes after each session, documenting the session date, start and end times, notes on issues encountered during the session, and quotes from participants, and evaluated the amount of effort the older adults exerted to accomplish tasks.

Data Collection Procedures

Data were collected throughout the experiment from the older adult and their companion, as detailed subsequently. A final interview was conducted at the experiment's conclusion. All data were collected in person using standardized assessment forms and entered into REDCap (Research Electronic Data Capture; Vanderbilt University) [63], a secure, web-based software platform, and double checked for accuracy. During face-to-face interactions, participants read the questionnaire while the research staff read the questions aloud and documented the responses. Instruments were chosen with established validity and reliability. All data collection forms were scanned and uploaded into REDCap. Table 1 displays the study assessment measures and timing.

XSL•FO
RenderX

Table 1. Study assessments and the schedule of evaluations.

Assessment	Enrollment and baseline assessment	Week 1 (sessions 1 and 2)	Week 2 (sessions 3 and 4)	Week 3 (sessions 5 and 6)	Week 4 (sessions 7 and 8)	Week 6 (follow-up: final visit)
Eligibility review and informed consent	✓					
Baseline measures						
Demographics	✓					
Physical function: Barthel index[a]	✓					
Cognition						
AD8 questionnaire	✓					
Dementia Severity Rating Scale	✓					
Self-Administered Gerocognitive Examination Questionnaire	✓					
Social network						
Lubben Social Network Scale for Long-Term Care[b]	✓					
6-item UCLA[c] Loneliness Scale	✓					
Session variables						
Attendance		✓	✓	✓	✓	
Engagement: Observational Measurement of Engagement tool		✓	✓	✓	✓	
3-item UCLA Loneliness Scale		✓	✓	✓	✓	
6-item Mutual Awareness Subscale		✓	✓	✓	✓	
Outcomes						
Feasibility and acceptability						
Feasibility of Intervention Measure questionnaire						✓
Acceptability of Intervention Measure questionnaire						✓
Exploratory variables						
6-item UCLA Loneliness Scale						✓
Network Minds Measure of Social Presence						✓
Semistructured interviews assessing facilitators and barriers						✓

[a]Only for participants residing in assisted living facilities.

[b]For participants residing in assisted living facilities, the Lubben Social Network Scale for Long-Term Care (LSNS-LTC) was used for long-term care settings.

[c]UCLA: University of California, Los Angeles.

XSL•FO
RenderX

Measures

Feasibility and Acceptability Outcomes

Feasibility, defined as the practicality of conducting the study [64], was measured by observation and documentation of recruitment and retention rates, data completion, and session attendance (ie, intervention completion). Logistical, technological, and user issues were observed and documented after each session. We measured intervention implementation rates for both study arms as the proportion of successful connections. Feasibility was also measured by the participants' perceptions of the ease of using the technology using a valid and reliable 4-item questionnaire, the Feasibility Intervention Measure [49]. Each item was rated on a 5-point Likert scale ranging from 1 (completely disagree) to 5 (completely agree). Data were collected during the final 2 weeks after the follow-up visit.

Acceptability, defined as the perception that the intervention (HMD-AR or 2D audio-visual intervention) was agreeable and palatable, was measured using a valid and reliable 4-item questionnaire—the Acceptability Intervention Measure [49]. Each item was rated on a 5-point Likert scale ranging from 1 (completely disagree) to 5 (completely agree). Data were collected 2 weeks after the follow-up visit. Overall satisfaction with the ICT was ascertained at the final session via semistructured interviews.

Baseline Variables

For both older adults and companions, baseline data were collected on demographics (age, gender, and race), social network, and loneliness. Social network was assessed using the 12-item Lubben Social Network Scale [65], a self-report measure of social engagement with family and friends. Each item is rated using a 6-point Likert scale ranging from 0 (no monthly contact) to 5 (daily contact); scores range from 0 to 60, with higher scores indicating a greater social network.

For older adults residing in LTC, the Revised Lubben Social Network Scale–Long-Term Care Scale [66] was used; this scale includes an additional 11 items that assess quality and frequency of interactions with other residents and staff members. Loneliness was assessed using the 6-item revised University of California, Los Angeles (UCLA) Loneliness Scale [67], a self-report measure consisting of 6 items rated using a 4-point Likert scale. The score ranges from 6 to 24, with higher scores indicating greater loneliness.

For older adults, cognition was assessed using the Self-Administered Gerocognitive Examination [68,69]. The Self-Administered Gerocognitive Examination is a brief tool that assesses language, reasoning or computation, visuospatial, executive, memory, and orientation domains. Scores range from 0 to 22, with scores greater than 16 suggesting normal cognition. Physical function was assessed via self-report using the Barthel Index [70]. Participants rate their degree of independence in performing 10 activities of daily living (eg, bathing and dressing); scores range from 0 to 100, with higher scores indicating greater independence.

ICT Sessions

Immediately after each session, RAs completed a modified Observational Measurement of Engagement [54] instrument to note engagement behaviors, type of activity, level of noise, and lighting. Additional field notes provided contextual information for understanding the older adults' engagement with their family members across each type of ICT. Participants were asked to complete the 3-item UCLA Loneliness Scale [71].

The Mutual Awareness Subscale (MAS) of the Networked Minds Measure of Social Presence [72] was completed after the second and eighth sessions. The MAS is an 8-item questionnaire rated using a 5-point Likert scale, ranging from 0 (never) to 4 (a great deal); scores range from 0 to 32, with higher scores indicating a greater sense of social presence with one's partner.

At the last session, participants completed the Networked Minds Measure of Social Presence instrument, consisting of several subscales: MAS (8 items), Perceived Attentional Engagement (6 items), Perceived Emotional Contagion (8 items), Perceived Comprehension (6 items), and Perceived Behavioral Interdependence (6 items). Higher scores indicated greater social presence. Loneliness is the targeted primary outcome for future studies; therefore, participants also completed the 6-item UCLA Loneliness Scale [68]. Finally, the RAs conducted a brief interview using a semistructured questionnaire to elicit participants' perceptions of facilitators and barriers.

Analysis Plan

We will produce a CONSORT (Consolidated Standards of Reporting Trials) flowchart to track the progress through the phases of the trial (enrollment, intervention allocation, follow-up, and data analysis). Reasons for noneligibility, nonparticipation, and attrition will be documented and summarized. Feasibility and acceptability (eg, participant recruitment, enrollment, and retention rates) will be descriptively analyzed using frequency distributions and measures of central tendency and dispersion, as appropriate. For participant-completed instruments (eg, social presence, Feasibility Intervention Measure, and Acceptability Intervention Measure), we will examine the extent of missing data to inform the selection of future instruments.

Results

Funding was received from the NIA in August 2022 and the National Science Foundation in October 2022 (Multimedia Appendices 1 and 2). This study was approved by the local IRB and registered at ClinicalTrials.gov. Following the initial design and testing of the HMD-AR activities, recruitment and data collection for the feasibility and acceptability study began in April 2024 and were completed in May 2025. Of the 8 enrolled older adult and family member pairs, 8 (100%) completed all sessions. Final acquisition of data has been completed, and the data are currently undergoing cleaning. Results are intended to be published in 2026.

XSL•FO
RenderX

Discussion

This study will assess the feasibility of implementing HMD-AR communication between older adults and family members or close friends and identify key facilitators of and barriers to implementation. This project has several innovations. First, collaborative HMD-AR delivers enhanced social connection for older adults and their family members or friends via a 3D hologram using realistic avatars that are transported to each other's environments. Most previous work on older adults in AR focuses only on a single user and does not include collaborative aspects. Second, the proposed technology has the capability to map the remote users' interactions with augmented objects in a chosen environment. Interactions can be mapped efficiently even when the remote and local users' environments are nonisomorphic. Third, assessing the feasibility and acceptability of collaborative AR among older adults residing alone in the community will be novel to the literature. This work contributes to the knowledge on technology-based interventions targeting loneliness by examining the underlying mechanism of social presence in ICT.

Several limitations exist. It will not be possible to blind participants to the study arms. Attrition can significantly impact studies with small sample sizes. On the basis of our previous experiences, we believe attrition will be low but may occur due to illness or disinterest. We will address attrition through regular communication with the participants and by scheduling sessions at times convenient for participants. Cost and complexity may be issues in applying advanced ICT. We chose to use commercially available, nonproprietary equipment to minimize costs. We plan to make computational algorithms freely available. AR technologies and components are becoming cheaper as demand and applications increase. Almost any new technological intervention plan will introduce additional costs and complexity, both in the home and in the existing system of care; our proposed system is no exception. Whether such an additional burden justifies the new technology is a question that needs to be answered on a case-by-case basis.

Our work will expand on previous studies and focus on HMD-AR to address loneliness. We expect a greater understanding of the feasibility, acceptability, and social presence associated with the use of varying collaborative HMD-AR activities and environments for older adults with different levels of cognitive impairment and their family members or close friends. We also anticipate a greater understanding of the logistics and deployment of this technology in homes and in LTCs.

Acknowledgments

REDCap (Research Electronic Data Capture; Vanderbilt University) is supported in part by the National Institutes of Health (NIH/NCATS UL1 TR000445). The authors acknowledge their safety officer, Jeremy Neal, PhD, RN, professor at the Vanderbilt University School of Nursing, who oversaw research procedures to ensure the safety of participants, validity of findings, and quality of data collection.

A generative artificial intelligence tool, Gemini (version 3, 2025; Google DeepMind), was used to (1) assist in identifying key studies in augmented reality and older adults and (2) generate a title meeting the journal guidelines.

Funding

This research protocol was funded by the National Institute on Aging of the National Institutes of Health (R21AG078480) and the National Science Foundation (2225890).

Data Availability

Final research data will be uploaded and made available on the website, Open Science Framework [73]. This will occur as our manuscripts are accepted for publication and no later than the online publication date. The engineering know-how for creating the collaborative augmented reality activities and programming techniques will be shared with other researchers through a no-fee research-use only license to be issued upon written request to the principal investigators.

Authors' Contributions

Conceptualization: NS, JAT, LAJ, CAM, DWS, AT, LCM
Data curation: MT, AT
Formal analysis: AT
Investigation and data collection: MT, ZI, RZ
Methodology: NS, LCM, AT
Resources: NS
Software: AU, MT, NS, ZI, RZ
Supervision: NS, JAT
Validation: MT, NS, JAT
Visualization: MT, LCM, AT

XSL•FO
RenderX

Writing—original draft: LCM, MT
Writing—review and editing: NS, JAT, LAJ, CAM, DWS, LCM, MT, RZ, ZI, LAJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from SPIP - Social Psychology, Personality and Interpersonal Processes Study Section, Risk, Prevention and Health Behavior Integrated Review Group, National Institute on Aging (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 149 KB - resprot_v15i1e83864_app1.pdf]

Multimedia Appendix 2
Investigator's response to the peer review report.
[DOCX File , 23 KB - resprot_v15i1e83864_app2.docx]

References

1. Holt-Lunstad J, Robles TF, Sbarra DA. Advancing social connection as a public health priority in the United States. Am Psychol 2017 Sep;72(6):517-530 [FREE Full text] [doi: 10.1037/amp0000103] [Medline: 28880099]
2. Cudjoe TK, Roth DL, Szanton SL, Wolff JL, Boyd CM, Thorpe RJ. The epidemiology of social isolation: National Health and Aging Trends Study. J Gerontol B Psychol Sci Soc Sci 2020 Jan 01;75(1):107-113 [FREE Full text] [doi: 10.1093/geronb/gby037] [Medline: 29590462]
3. Donovan NJ, Blazer D. Social isolation and loneliness in older adults: review and commentary of a National Academies report. Am J Geriatr Psychiatry 2020 Dec;28(12):1233-1244 [FREE Full text] [doi: 10.1016/j.jagp.2020.08.005] [Medline: 32919873]
4. National Academies of Sciences, Engineering, and Medicine, Division of Behavioral and Social Sciences and Education, Health and Medicine Division, Board on Behavioral, Cognitive, and Sensory Sciences, Board on Health Sciences Policy, Committee on the Health and Medical Dimensions of Social Isolation and Loneliness in Older Adults. Social Isolation and Loneliness in Older Adults: Opportunities for the Health Care System. Washington, DC: National Academies Press; 2020.
5. Chamberlain SA, Duggleby W, Teaster PB, Estabrooks CA. Characteristics of socially isolated residents in long-term care: a retrospective cohort study. Gerontol Geriatr Med 2020;6:2333721420975321 [FREE Full text] [doi: 10.1177/2333721420975321] [Medline: 33283023]
6. Boamah SA, Weldrick R, Lee TS, Taylor N. Social isolation among older adults in long-term care: a scoping review. J Aging Health 2021 Mar 27;33(7-8):618-632 [FREE Full text] [doi: 10.1177/08982643211004174] [Medline: 33779366]
7. Simard J, Volicer L. Loneliness and isolation in long-term care and the COVID-19 pandemic. J Am Med Dir Assoc 2020 Jul;21(7):966-967 [FREE Full text] [doi: 10.1016/j.jamda.2020.05.006] [Medline: 32505516]
8. Smith BJ, Lim MH. How the COVID-19 pandemic is focusing attention on loneliness and social isolation. Public Health Res Pract 2020 Jun 30;30(2):3022008 [FREE Full text] [doi: 10.17061/phrp3022008] [Medline: 32601651]
9. Lapane KL, Lim E, McPhillips E, Barooah A, Yuan Y, Dube CE. Health effects of loneliness and social isolation in older adults living in congregate long term care settings: a systematic review of quantitative and qualitative evidence. Arch Gerontol Geriatr 2022;102:104728 [FREE Full text] [doi: 10.1016/j.archger.2022.104728] [Medline: 35597183]
10. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. PLoS Med 2010 Jul 27;7(7):e1000316 [FREE Full text] [doi: 10.1371/journal.pmed.1000316] [Medline: 20668659]
11. Holt-Lunstad J, Smith TB, Baker M, Harris T, Stephenson D. Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect Psychol Sci 2015 Mar;10(2):227-237. [doi: 10.1177/1745691614568352] [Medline: 25910392]
12. Hoang P, King JA, Moore S, Moore K, Reich K, Sidhu H, et al. Interventions associated with reduced loneliness and social isolation in older adults: a systematic review and meta-analysis. JAMA Netw Open 2022 Oct 03;5(10):e2236676 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.36676] [Medline: 36251294]
13. Chen YR, Schulz PJ. The effect of information communication technology interventions on reducing social isolation in the elderly: a systematic review. J Med Internet Res 2016 Jan 28;18(1):e18 [FREE Full text] [doi: 10.2196/jmir.4596] [Medline: 26822073]
14. Heins P, Boots LM, Koh WQ, Neven A, Verhey FR, de Vugt ME. The effects of technological interventions on social participation of community-dwelling older adults with and without dementia: a systematic review. J Clin Med 2021 May 25;10(11):2308 [FREE Full text] [doi: 10.3390/jcm10112308] [Medline: 34070660]
15. Casanova G, Zaccaria D, Rolandi E, Guaita A. The effect of information and communication technology and social networking site use on older people's well-being in relation to loneliness: review of experimental studies. J Med Internet Res 2021 Mar 01;23(3):e23588 [FREE Full text] [doi: 10.2196/23588] [Medline: 33439127]

XSL•FO
RenderX

16. Monin JK, Ali T, Syed S, Piechota A, Lepore M, Mourgues C, et al. Family communication in long-term care during a pandemic: lessons for enhancing emotional experiences. Am J Geriatr Psychiatry 2020 Dec;28(12):1299-1307 [FREE Full text] [doi: 10.1016/j.jagp.2020.09.008] [Medline: 33004262]
17. Ibarra F, Baez M, Cernuzzi L, Casati F. A systematic review on technology-supported interventions to improve old-age social wellbeing: loneliness, social isolation, and connectedness. J Healthc Eng 2020;2020:2036842 [FREE Full text] [doi: 10.1155/2020/2036842] [Medline: 32765823]
18. Barbosa Neves B, Franz R, Judges R, Beermann C, Baecker R. Can digital technology enhance social connectedness among older adults? A feasibility study. J Appl Gerontol 2019 Jan;38(1):49-72. [doi: 10.1177/0733464817741369] [Medline: 29166818]
19. Ambagtsheer RC, Borg K, Townsin L, Pinero de Plaza MA, O'Brien LM, Kunwar R, et al. The effectiveness of technology interventions in reducing social isolation and loneliness among community-dwelling older people: a mixed methods systematic review. Arch Gerontol Geriatr Plus 2024 Mar;1(1):100008. [doi: 10.1016/j.aggp.2024.100008]
20. Rettie R. Connectedness, awareness and social presence. In: Proceedings of the 6th Annual International Workshop on Presence. 2003 Presented at: AIWP 2003; Oct 6-8, 2003; Aalborg , Denmark URL: https://researchinnovation.kingston.ac.uk/en/publications/connectedness-awareness-and-social-presence-2/
21. Kreijns K, Xu K, Weidlich J. Social presence: conceptualization and measurement. Educ Psychol Rev 2022;34(1):139-170 [FREE Full text] [doi: 10.1007/s10648-021-09623-8] [Medline: 34177204]
22. Osmers N, Prilla M, Blunk O, Brown GG, Janßen M, Kahrl N. The role of social presence for cooperation in augmented reality on head mounted devices: a literature review. In: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021 Presented at: CHI '21; May 8-13, 2021; Yokohama, Japan URL: https://dl.acm.org/doi/10.1145/3411764.3445633 [doi: 10.1145/3411764.3445633]
23. Biocca F, Harms C, Burgoon JK. Toward a more robust theory and measure of social presence: review and suggested criteria. Presence 2003 Oct;12(5):456-480. [doi: 10.1162/105474603322761270]
24. Oh CS, Bailenson JN, Welch GF. A Systematic Review of Social Presence: Definition, Antecedents, and Implications. Front Robot AI 2018;5:114 [FREE Full text] [doi: 10.3389/frobt.2018.00114] [Medline: 33500993]
25. Venkatesan M, Mohan H, Ryan JR, Schürch CM, Nolan GP, Frakes DH, et al. Virtual and augmented reality for biomedical applications. Cell Rep Med 2021 Jul 20;2(7):100348 [FREE Full text] [doi: 10.1016/j.xcrm.2021.100348] [Medline: 34337564]
26. Lee LN, Kim MJ, Hwang WJ. Potential of augmented reality and virtual reality technologies to promote wellbeing in older adults. Appl Sci 2019 Aug 30;9(17):3556-3847 [FREE Full text] [doi: 10.3390/app9173556]
27. Cicconi S, Marchese M. Augmented learning: an e-learning environment in augmented reality for older adults. In: Proceedings of the 13th International Technology, Education and Development Conference. 2019 Presented at: INTED2019 Proceedings; March 11-13, 2019; Valencia, Spain URL: https://library.iated.org/view/CICCONI2019AUG [doi: 10.21125/inted.2019.0937]
28. Nishchyk A, Chen W, Pripp AH, Bergland A. The effect of mixed reality technologies for falls prevention among older adults: systematic review and meta-analysis. JMIR Aging 2021 Jun 30;4(2):e27972 [FREE Full text] [doi: 10.2196/27972] [Medline: 34255643]
29. Piumsomboon T, Day A, Ens B, Lee Y, Lee G, Billinghurst M. Exploring enhancements for remote mixed reality collaboration. In: Proceedings of the SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications. 2017 Presented at: SA '17; November 27-30, 2017; Bangkok, Thailand. [doi: 10.1145/3132787.3139200]
30. Ping J, Liu Y, Weng D. Comparison in depth perception between virtual reality and augmented reality systems. In: Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces. 2019 Presented at: VR 2019; March 23-27, 2019; Osaka, Japan URL: https://ieeexplore.ieee.org/document/8798174 [doi: 10.1109/vr.2019.8798174]
31. Krichenbauer M, Yamamoto G, Taketom T, Sandor C, Kato H. Augmented reality versus virtual reality for 3D object manipulation. IEEE Trans Vis Comput Graph 2018 Feb 1;24(2):1038-1048. [doi: 10.1109/TVCG.2017.2658570] [Medline: 28129181]
32. Xu X, Mangina E, Campbell AG. HMD-based virtual and augmented reality in medical education: a systematic review. Front Virtual Real 2021 Jul 6;2. [doi: 10.3389/frvir.2021.692103]
33. Roettl J, Terlutter R. The same video game in 2D, 3D or virtual reality - how does technology impact game evaluation and brand placements? PLoS One 2018 Jul 20;13(7):e0200724 [FREE Full text] [doi: 10.1371/journal.pone.0200724] [Medline: 30028839]
34. Park S, Lee G. Full-immersion virtual reality: adverse effects related to static balance. Neurosci Lett 2020 Aug 10;733:134974. [doi: 10.1016/j.neulet.2020.134974] [Medline: 32294492]
35. Appel L, Appel E, Bogler O, Wiseman M, Cohen L, Ein N, et al. Older adults with cognitive and/or physical impairments can benefit from immersive virtual reality experiences: a feasibility study. Front Med (Lausanne) 2019;6:329 [FREE Full text] [doi: 10.3389/fmed.2019.00329] [Medline: 32010701]
36. Huygelier H, Schraepen B, van Ee R, Vanden Abeele V, Gillebert CR. Acceptance of immersive head-mounted virtual reality in older adults. Sci Rep 2019 Mar 14;9(1):4519 [FREE Full text] [doi: 10.1038/s41598-019-41200-6] [Medline: 30872760]

XSL•FO
RenderX

37. Vieira ER, Civitella F, Carreno J, Junior MG, Amorim CF, D'Souza N, et al. Using augmented reality with older adults in the community to select design features for an age-friendly park: a pilot study. J Aging Res 2020;2020:8341034 [FREE Full text] [doi: 10.1155/2020/8341034] [Medline: 32953175]
38. Vorwerg S, Stamm O, Müller-Werdan U. Movement Training in Augmented Reality for Older Adults with Essential Hypertension: A Qualitative Study Determining Training Requirements. Games Health J 2020 Dec;9(6):436-445. [doi: 10.1089/g4h.2019.0133] [Medline: 33301387]
39. Blomqvist S, Seipel S, Engström M. Using augmented reality technology for balance training in the older adults: a feasibility pilot study. BMC Geriatr 2021 Feb 26;21(1):144 [FREE Full text] [doi: 10.1186/s12877-021-02061-9] [Medline: 33637043]
40. Mostajeran F, Steinicke F, Nunez OJ, Gatsios D, Fotiadis D. Augmented reality for older adults: exploring acceptability of virtual coaches for home-based balance training in an aging population. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020 Presented at: CHI '20; April 25-30, 2020; Honolulu HI URL: https://dl.acm.org/doi/10.1145/3313831.3376565 [doi: 10.1145/3313831.3376565]
41. Aruanno B, Garzotto F. MemHolo: mixed reality experiences for subjects with Alzheimer's disease. Multimed Tools Appl 2019 Jan 5;78(10):13517-13537. [doi: 10.1007/s11042-018-7089-8]
42. Korsgaard D, Bjorner T, Bruun-Pedersen JR, Sorensen PK, Perez-Cueto FJ. Eating together while being apart: a pilot study on the effects of mixed-reality conversations and virtual environments on older eaters' solitary meal experience and food intake. In: Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops. 2020 Presented at: VRW 2020; March 22-26, 2020; Atlanta, GA URL: https://ieeexplore.ieee.org/document/9090427 [doi: 10.1109/vrw50115.2020.00079]
43. Tauseef M, Ullal A, Watkins A, Dietrich MS, Maxwell C, Tate J, et al. Older and younger adults' perceptions of augmented reality photorealistic avatars as a viable medium for interpersonal communication. Innov Aging 2025;9(9):igaf083 [FREE Full text] [doi: 10.1093/geroni/igaf083] [Medline: 41069994]
44. Ullal A, Tauseef M, Watkins A, Juckett L, Maxwell C, Tate J, et al. An iterative participatory design approach to develop collaborative augmented reality activities for older adults in long-term care facilities. Proc SIGCHI Conf Hum Factor Comput Syst 2024 May;2024:1009. [doi: 10.1145/3613904.3642595] [Medline: 39015130]
45. O'Rourke HM, Sidani S. Definition, determinants, and outcomes of social connectedness for older adults: a scoping review. J Gerontol Nurs 2017 Jul 01;43(7):43-52. [doi: 10.3928/00989134-20170223-03] [Medline: 28399313]
46. Bonsaksen T, Thygesen H, Leung J, Ruffolo M, Schoultz M, Price D, et al. Video-based communication and its association with loneliness, mental health and quality of life among older people during the COVID-19 outbreak. Int J Environ Res Public Health 2021 Jun 10;18(12):6284 [FREE Full text] [doi: 10.3390/ijerph18126284] [Medline: 34200670]
47. Noone C, McSharry J, Smalle M, Burns A, Dwan K, Devane D, et al. Video calls for reducing social isolation and loneliness in older people: a rapid review. Cochrane Database Syst Rev 2020 May 21;5(5):CD013632 [FREE Full text] [doi: 10.1002/14651858.CD013632] [Medline: 32441330]
48. Eriksen S, Helvik AS, Juvet LK, Skovdahl K, Førsund LH, Grov EK. The experience of relations in persons with dementia: a systematic meta-synthesis. Dement Geriatr Cogn Disord 2016;42(5-6):342-368. [doi: 10.1159/000452404] [Medline: 27866199]
49. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
50. Powell BJ, Fernandez ME, Williams NJ, Aarons GA, Beidas RS, Lewis CC, et al. Enhancing the impact of implementation strategies in healthcare: a research agenda. Front Public Health 2019;7:3 [FREE Full text] [doi: 10.3389/fpubh.2019.00003] [Medline: 30723713]
51. Bach D. U.S. Army to use HoloLens technology in high-tech headsets for soldiers. Microsoft Transform. 2021 Jun 08. URL: https://news.microsoft.com/source/features/digital-transformation/u-s-army-to-use-hololens-technology-in-high-tech-headsets-for-soldiers/ [accessed 2021-10-12]
52. Azhar A. Mixed reality storytelling for social engagement with older adults. OCAD University. 2021 May. URL: https://openresearch.ocadu.ca/id/eprint/3450/1/Azhar_Amna_2021_MDes_INCD_MRP.pdf [accessed 2021-10-12]
53. Jeste DV, Palmer BW, Appelbaum PS, Golshan S, Glorioso D, Dunn LB, et al. A new brief instrument for assessing decisional capacity for clinical research. Arch Gen Psychiatry 2007 Aug;64(8):966-974. [doi: 10.1001/archpsyc.64.8.966] [Medline: 17679641]
54. Cohen-Mansfield J, Hai T, Comishen M. Group engagement in persons with dementia: the concept and its measurement. Psychiatry Res 2017 May;251:237-243. [doi: 10.1016/j.psychres.2017.02.013] [Medline: 28214783]
55. Cohen-Mansfield J, Jensen B. Attendance in recreational groups for persons with dementia: the impact of stimulus and environmental factors. Am J Alzheimers Dis Other Demen 2018 Nov;33(7):471-478 [FREE Full text] [doi: 10.1177/1533317518788158] [Medline: 30012002]
56. Cohen-Mansfield J, Thein K, Dakheel-Ali M, Marx MS. Engaging nursing home residents with dementia in activities: the effects of modeling, presentation order, time of day, and setting characteristics. Aging Ment Health 2010 May;14(4):471-480 [FREE Full text] [doi: 10.1080/13607860903586102] [Medline: 20455123]

XSL•FO
RenderX

57. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478 [FREE Full text] [doi: 10.2307/30036540]
58. Cohen-Mansfield J. Activity groups for persons with dementia: personal predictors of participation, engagement and mood. Psychiatry Res 2017 Nov;257:375-380. [doi: 10.1016/j.psychres.2017.07.045] [Medline: 28806713]
59. Nielsen J, Landauer T. A mathematical model of the finding of usability problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 Presented at: INTERCHI 1993; April 24-29, 1993; Amsterdam, The Netherlands URL: https://doi.org/10.1145/1690169166 [doi: 10.1145/169059.169166]
60. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Nov 24;4(4):287-291. [doi: 10.1002/pst.185]
61. Moore CG, Carter RE, Nietert PJ, Stewart PW. Recommendations for planning pilot studies in clinical and translational research. Clin Transl Sci 2011 Oct;4(5):332-337 [FREE Full text] [doi: 10.1111/j.1752-8062.2011.00347.x] [Medline: 22029804]
62. Resnick B, Bellg AJ, Borrelli B, Defrancesco C, Breger R, Hecht J, et al. Examples of implementation and evaluation of treatment fidelity in the BCC studies: where we are and where we need to go. Ann Behav Med 2005 Apr;29 Suppl:46-54. [doi: 10.1207/s15324796abm2902s_8] [Medline: 15921489]
63. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
64. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022 Jan 01;60(1):95-103 [FREE Full text] [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
65. Lubben J. Assessing social networks among elderly populations. Fam Community Health 1988;11(3):42-52 [FREE Full text] [doi: 10.1097/00003727-198811000-00008]
66. Munn J, Radey M, Brown K, Kim H. Revising the Lubben Social Network Scale for use in residential long-term care settings. J Evid Inf Soc Work 2018;15(4):385-402. [doi: 10.1080/23761407.2018.1460734] [Medline: 29672239]
67. Wongpakaran N, Wongpakaran T, Pinyopornpanish M, Simcharoen S, Suradom C, Varnado P, et al. Development and validation of a 6-item Revised UCLA Loneliness Scale (RULS-6) using Rasch analysis. Br J Health Psychol 2020 May 30;25(2):233-256. [doi: 10.1111/bjhp.12404] [Medline: 31999891]
68. Scharre DW, Chang SI, Nagaraja HN, Yager-Schweller J, Murden RA. Community cognitive screening using the self-administered gerocognitive examination (SAGE). J Neuropsychiatry Clin Neurosci 2014;26(4):369-375. [doi: 10.1176/appi.neuropsych.13060145] [Medline: 24419587]
69. Scharre DW, Chang SI, Murden RA, Lamb J, Beversdorf DQ, Kataki M, et al. Self-administered Gerocognitive Examination (SAGE): a brief cognitive assessment instrument for mild cognitive impairment (MCI) and early dementia. Alzheimer Dis Assoc Disord 2010;24(1):64-71. [doi: 10.1097/WAD.0b013e3181b03277] [Medline: 20220323]
70. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
71. Hughes ME, Waite LJ, Hawkley LC, Cacioppo JT. A short scale for measuring loneliness in large surveys: results from two population-based studies. Res Aging 2004;26(6):655-672 [FREE Full text] [doi: 10.1177/0164027504268574] [Medline: 18504506]
72. Harms C, Biocca F. Internal consistency and reliability of the Networked Minds Social Presence measure. In: Proceedings of the Seventh Annual International Workshop: Presence 2004. 2004 Presented at: PRESENCE 2004; October 13-15, 2004; Valencia, Spain URL: https://web-archive.southampton.ac.uk/cogprints.org/7026/1/Harms_04_reliability_validity_social_presence_(Biocca).pdf
73. OSF. URL: https://osf.io/ [accessed 2026-01-25]

Abbreviations

AR: augmented reality
CONSORT: Consolidated Standards of Reporting Trials
HMD: head-mounted display
HMD-AR: head-mounted display augmented reality
ICT: interactive communication technology
IRB: Institutional Review Board
LTC: long-term care
MAS: Mutual Awareness Subscale
NIA: National Institute on Aging
RA: research assistant
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
UCLA: University of California, Los Angeles
VR: virtual reality

XSL•FO
RenderX

Edited by J Sarvestan; submitted 10.Sep.2025; this is a non–peer-reviewed article;accepted 08.Dec.2025; published 11.Feb.2026.

Please cite as:
Sarkar N, Tate JA, Tauseef M, Ullal A, Maxwell CA, Juckett LA, Scharre DW, Tan A, Zhang R, Ingram Z, Mion LC
Feasibility and Acceptability of Collaborative Augmented Reality for Older Adults and Companions: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e83864
URL: https://www.researchprotocols.org/2026/1/e83864
doi:10.2196/83864
PMID:

©Nilanjan Sarkar, Judith A Tate, Mahrukh Tauseef, Akshith Ullal, Cathy A Maxwell, Lisa A Juckett, Douglas W Scharre, Alai Tan, Rendong Zhang, Zalen Ingram, Lorraine C Mion. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

The Third Study of Infectious Intestinal Disease (IID3 Study) in the Community: Protocol for UK-Based Prospective Cohort Studies Investigating the Disease Burden

Ben W Rowland[1*], PhD; Vanashree Sexton[2*], PhD; Aileen Mill[1], PhD; Stephen Rushton[1], PhD; Roy Sanderson[1], PhD; Caroline Grundy[1], BA; Simon de Lusignan[2], MBBS; Nigel A Cunliffe[3], PhD; Daniel Hungerford[3], PhD; Mark Hopkins[4], PhD; Saheer Gharbia[5], PhD; Claire Jenkins[6], PhD; Gauri Godbole[6], PhD; Roberto Vivancos[6,7,8], MD; Alex J Elliot[6,9,10], PhD; Dominic J Mellor[11], PhD; Lesley Larkin[11], MSc; Rachel Chalmers[12], PhD; Sarah O'Brien[1], MBBS; IID3 Consortium[13*]

[1]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, England, United Kingdom

[2]Clinical Informatics and Health Outcomes Research Group, Nuffield Department of Primary Care Health Science, University of Oxford, Oxford, England, United Kingdom

[3]Department of Clinical Infection, Microbiology and Immunology, Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, England, United Kingdom

[4]Liverpool Clinical Laboratories, Liverpool University Hospitals NHS Trust, Liverpool, England, United Kingdom

[5]Wellcome Sanger Institute, Hinxton, England, United Kingdom

[6]UK Health Security Agency, London, England, United Kingdom

[7]Emerging and Zoonotic Infections, National Institute for Health Research, Liverpool, England, United Kingdom

[8]Warwick Medical School, University of Warwick, Coventry, England, United Kingdom

[9]National Institute for Health Research, University of East Anglia, Norwich, United Kingdom

[10]NIHR Health Protection Research Unit in Emergency Preparedness and Response, University of Birmingham, Birmingham, United Kingdom

[11]Public Health Scotland, Edinburgh, Scotland, United Kingdom

[12]Swansea University, Swansea, Wales, United Kingdom

[13]See Acknowledgements

[*]these authors contributed equally

Corresponding Author:
Ben W Rowland, PhD
School of Natural and Environmental Sciences
Newcastle University
Kings Gate
Barras Bridge
Newcastle upon Tyne, England, NE1 7RU
United Kingdom
Phone: 44 01912086098
Email: Ben.Rowland@newcastle.ac.uk

Abstract

Background: There is a significant hidden burden of infectious intestinal disease (IID) in the UK community, which has increased over time. In the late 2000s, the Second Study of Infectious Intestinal Disease (IID2 study) estimated 17 million IID cases annually in the United Kingdom. However, only a small proportion of cases present to health care, and even those are often not tested for causative organisms.

Objective: The Third Study of Infectious Intestinal Disease (IID3 study) aims to determine the IID burden in the UK community, estimate the underreporting level in routine practice and the general population, and recalibrate UK national surveillance based on the new incidence rates.

Methods: We will follow methods of previous studies, along with modern pathogen detection methods and digital platforms for recruitment and follow-up. Participants will be recruited to three population-based prospective cohorts: cohort 1 (the general population), cohort 2 (patients with IID presenting to general practices [GPs]), and cohort 3 (enumeration study of IID cases

XSL•FO
RenderX

presenting to GPs). Microbiological analysis of stool samples in cohorts 1 and 2 will include testing for a wide range of causative organisms using molecular assays, including pathogen targets not routinely sought by National Health Service (NHS) laboratories. Additional characterization of pathogens will be conducted at national reference laboratories. The incidence rates of IID and organisms detected in cohorts 1-3 will be compared to national surveillance systems, both laboratory and syndromic. Descriptive statistics and analysis will allow comparison of IID rates within each cohort, estimate the overall burden of disease caused by different pathogens, and compare findings to earlier IID studies.

Results: A favorable ethical opinion was obtained from the UK Health Research Authority on August 4, 2022. A pilot phase to test the sampling process was conducted from January to August 2023. Participant recruitment commenced on September 1, 2023, for cohort 2 and on March 16, 2024, for cohort 1; recruitment ceased on August 31, 2025. Data collection is complete, and data analysis is to begin. The study is expected to end in September 2026.

Conclusions: Since the first and second IID studies, changes have occurred within national surveillance systems, the NHS structure, and public recommendations about when to consult a GP and where to seek health care advice, which may have altered the extent of IID reporting and the perceived burden in the community, creating greater uncertainty about the representativeness of IID rates. The IID3 study results will provide insight into trends in disease incidence over time and help quantify inequalities in IID in the UK community. Revised estimates can inform policy related to prevention, including food standards and disease management. Furthermore, advances in molecular diagnostics will significantly enhance pathogen detection, increasing our understanding of the causes of IID.

International Registered Report Identifier (IRRID): DERR1-10.2196/88759

(JMIR Res Protoc 2026;15:e88759) doi:10.2196/88759

KEYWORDS

infectious intestinal disease; IID; cohort; general practitioner; sample; pathogen

Introduction

Background

Infectious intestinal disease (IID) is a common illness, with an estimated 17 million cases in the United Kingdom each year [1]. In 2018, the estimated cost of foodborne illness to the UK economy was approximately GBP 9 (USD 12.4) billion per year [2]. Although most cases are self-limiting, these infections can lead to significant complications and mortality, with substantial age and pathogen-specific inequalities [3,4]. IID is caused by a wide range of causative organisms, including bacteria (eg, *Salmonella* spp. and *Campylobacter* spp.), viruses (eg, norovirus and rotavirus), and parasites (eg, *Giardia* spp. and *Cryptosporidium* spp.). Transmission of these pathogens occurs through a variety of routes, including foodborne, waterborne, and person-to-person transmission, as well as contaminated environments [5-7]. Foodborne transmission alone accounts for an estimated 2.4 million cases in the United Kingdom annually. Only a small proportion of people with IID report to primary care, but their illness may prevent them from attending work or school. In the Second Study of Infectious Intestinal Disease (IID2 study; 2007-2009) [1], a major, unexpected result was that this hidden burden of disease had increased between the First Study of Infectious Intestinal Disease (IID1 study; 1993-1996) [8] and the IID2 study, largely due to changes in the delivery of primary care. The societal and economic impact of IID and the resulting health inequalities were greatly underestimated because they could not be measured through routine surveillance, which relies on contact with health services.

Understanding the level of underreporting is important because public health actions by regulators, such as the Food Standards Agency (FSA), need to be based on robust information that accurately identifies hazards and risks. Much of the current information used by policy advisors is based on the results of IID2. However, uncertainty is growing about the reliability of IID2 results to accurately reflect the current status of IID rates and reporting practices in the United Kingdom, as older estimates inevitably become less accurate over time.

Furthermore, four major changes make recalibrating national surveillance data particularly important to the FSA's core work. First, significant changes to primary care delivery are still evolving. Following the COVID-19 pandemic, there has been an increase in digital and non–general practice (GP)–based access to primary care and advice (eg, triage via National Health Service [NHS] 111/NHS 24 and pharmacies) [9]. These changes likely result in fewer cases of disease being directly reported, and in the absence of independent estimates to support revised underreporting figures, the allocation of funding and investment in prevention and disease management is not well informed.

Second, the public health and social measures enacted during the COVID-19 pandemic had a substantial impact on the confirmed cases and reported outbreaks of IID (34% and 52% decrease, respectively) [10]. This was reflected in national surveillance data, with the rates of disease being reported having returned to prepandemic levels [11]. Despite this, it is important to consider the impacts of the COVID-19 pandemic on current reporting patterns of disease, engagement in research, and how and when people access the health care system for routine care.

Third, laboratory diagnostic methods have changed significantly over time. Improved molecular methods have expanded the range and sensitivity of pathogen detection, enabling a more complete description of infections and potentially altering previous estimates. For example, in the IID1 study, traditional bacterial culture methods were used, whereas in the IID2 study, a mix of traditional and molecular methods (polymerase chain reaction [PCR]) were used. The Third Study of Infectious

XSL•FO
RenderX

Intestinal Disease (IID3 study) will use molecular methods as the primary diagnostic approach for detecting bacteria, viruses, and parasites, reflecting the shift by NHS laboratories toward molecular methods for pathogen detection.

Fourth, infant immunization against rotavirus IID was introduced in the United Kingdom in 2013, and vaccine uptake rapidly exceeded 90%. Before vaccine introduction, rotavirus was estimated to account for 80,000 GP consultations and 750,000 episodes of diarrhea [12,13]. Since the introduction, there have been substantial reductions in health care–attended rotavirus-associated gastroenteritis among vaccine-eligible children [14,15]. There is also some evidence of indirect effects in older unvaccinated populations. Therefore, the burden and epidemiology of rotavirus-associated IID is likely to be substantially different since the IID2 study was conducted. It is, therefore, timely to conduct the third UK study of IID to quantify the impact of changes in the epidemiology of infection and disease reporting, and to investigate established and emerging pathogens. The design of the IID3 study will allow for the identification of underreporting rates across primary care to assess whether the societal burden of IID has changed since the late 2000s. There is an opportunity to pivot over time toward lower-cost, continuous sentinel surveillance, but these approaches need to be benchmarked against conventional ones.

Study Overview

The IID3 study comprises three population-based prospective studies (one cohort study and two studies in primary care), supported by diagnostic and reference microbiological investigation of fecal specimens. Routine public health surveillance data, provided by the United Kingdom Health Security Agency (UKHSA), Public Health Scotland (PHS), Public Health Wales (PHW), and the Public Health Agency Northern Ireland (PHA NI), allow calibration of the IID incidence rates reported to national surveillance records.

Overall Study Aims

The IID3 study aims to determine the incidence rates of and identify the microorganisms associated with IID in the community, as well as incidence rates of IID presenting to GPs; compare these data with routine testing in primary care in the United Kingdom; and investigate how this has changed since the previous IID2. The seven study aims are as follows:

- Determine the overall incidence of IID in the UK population.
- Establish the incidence of IID presenting to primary care.
- Clarify the proportion of IID that is UK acquired.
- Describe the pathogens causing IID in the United Kingdom, including levels of antimicrobial resistance (AMR).
- Recalibrate UK surveillance data for IID, overall and by pathogen.
- Determine the number of cases in the community, GP-reported cases, and hospitalizations and deaths due to IID in the United Kingdom.
- Compare results from the IID3 study to previous studies.

Methods

Study Setting and Case Definition

Participants and GPs were recruited from across all four UK nations. We used the same case definition used for the earlier IID studies to ensure direct comparability [16]. Thus, “cases” were defined as “persons with loose stools or clinically significant vomiting lasting less than 2 weeks, in the absence of a known noninfectious cause, preceded by a symptom-free period of 3 weeks.” Vomiting is considered clinically significant if it occurs more than once in a 24-hour period and if it incapacitates the case or is accompanied by other symptoms, such as cramps or fever.

Study Design

The IID3 study comprises three prospective cohorts: cohort 1 (prospective population), cohort 2 (GP presentation), and cohort 3 (GP enumeration). The three cohorts are described in detail later and in Figure 1, which shows the study’s reverse pyramid structure. Recruitment into the cohorts and record-linked access to primary care routine data from GPs were enabled through the Oxford Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) [17].

XSL•FO
RenderX

Figure 1. Overview of study design. GP: general practice.

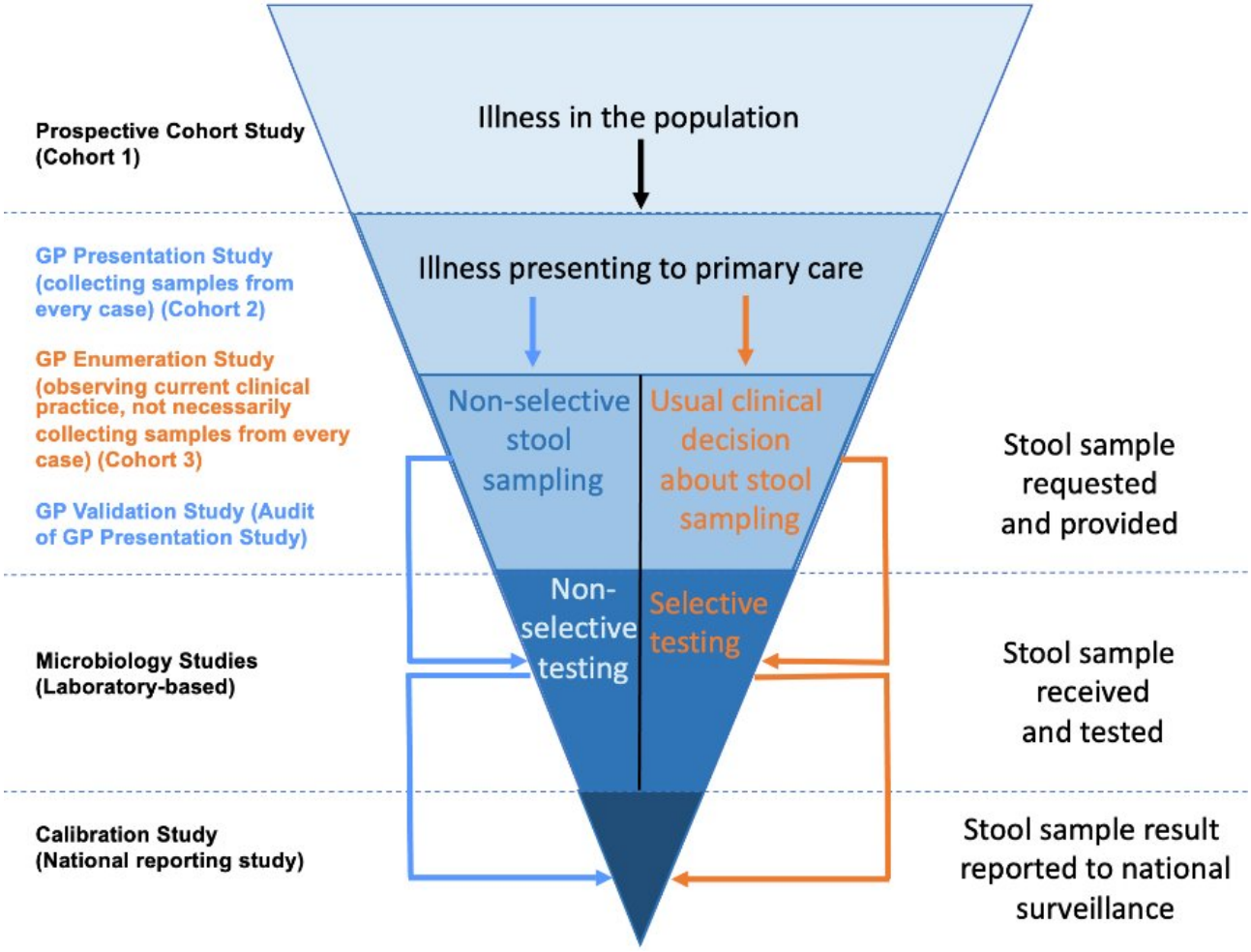

Prospective Cohort Study (Cohort 1)

Cohort 1 is a prospective, population-based household group designed to identify the incidence of IID in the community who may not present to GPs. In addition, cohort 1 data will provide insights into the transmission of IID within households. Participants with no current symptoms of IID were recruited to this cohort via invitations to join the study from their GPs. The invitations were sent in batches via text messages to the majority of the GP-registered population using clinical management software to apply the study's exclusion criteria (see Textbox 1).

Textbox 1. Patient eligibility criteria.

Inclusion criteria:

- Patients with diarrhea or clinically significant vomiting lasting less than 2 weeks, with no known noninfectious cause, preceded by a symptom-free period of 3 weeks
- Any age

Exclusion criteria (used for cohort 1 participant recruitment):

- Patients with terminal illness
- Patients with severe mental incapacity
- Patients who do not speak English and for whom a suitable interpreter is not available
- Surgical obstruction

Participants' IID status (whether they had experienced any symptoms of vomiting and/or diarrhea) was followed up weekly for 12 months through the study mobile app (PeopleWith). If participants reported symptoms, they were asked to complete a symptom questionnaire (see the full questionnaire in Multimedia Appendices 1 and 2) and to submit a stool sample.

XSL•FO
RenderX

GP Presentation Study (Cohort 2)

Cohort 2 is a prospective cohort of patients presenting to primary care with symptoms of IID. GPs were instructed to invite all patients who presented with IID symptoms (as per the case definition). Recruitment was primarily opportunistic, with patients presenting with IID symptoms asked to participate in the study following a face-to-face or telephone consultation. Patients who consented were asked to complete a symptom questionnaire (see the full questionnaire in Multimedia Appendix 3). GPs participating in cohort 2 collected stool samples from recruited patients.

GP Enumeration Study (Cohort 3)

The enumeration study will provide contemporary data on routine practice through a 24-month prospective audit of IID cases presenting to GPs, observing them as they occur in routine clinical practice. We included all RSC practices, excluding participating in cohorts 1 and 2 to prevent any participant from appearing in two cohorts, and compiled data on the clinical severity of IID cases, stool sample submission, and microbiology laboratory results. Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) codes will be used to identify IID cases.

Validation Study

To assess the level of underascertainment of IID cases in cohort 2 (where patients meeting the case definition were not invited or did not consent to the study), the participating practices will be audited by identifying all IID cases using the same SNOMED CT codes as the enumeration study.

Calibration Study

The overall IID incidence rates and pathogen-specific incidence rates will be calculated for each cohort. Cohort 1 data will provide a rate for the general population; cohort 2 data will allow us to estimate the true level of IID illness among those seeking primary care; and cohort 3 data will provide a rate observed in routine primary care. Comparison between these cohorts and with rates of IID reported to national surveillance systems will allow estimation of the degree of underreporting to primary care and national surveillance systems.

Data Sources

Participant and Symptom Information

Participant consent, baseline information, and information about IID symptoms were collected for cohorts 1 and 2 through a bespoke mobile app. The app was linked to a web-based portal that allows GPs to view participants who have provided consent, which was monitored by the study team to track recruitment. For cohort 1, GPs used the portal to view when participants reported illness on their weekly symptom questionnaires and to send an automated email asking participants to collect a stool sample kit.

Participants in cohort 1 who reported IID symptoms or those in cohort 2 who presented to GPs with IID were asked to complete the symptom questionnaire to obtain data on the following items: symptoms experienced and number of days the symptoms lasted; date of symptom onset; prior health care service use for symptoms, including use of the NHS 111 telephone service or other urgent care services; and recent travel outside the United Kingdom. The complete questionnaire is available in Multimedia Appendices 1 and 2.

The RSC provided primary care medical record data for cohorts 1, 2, and 3. These data include participant-associated variables, such as age, ethnicity, index of multiple deprivation (IMD), records of vaccination history, recent changes to medication, antibiotic usage (within the 3 months prior to the start of the study), referrals to gastrointestinal (GI) specialists, a history of GI or respiratory illness, and details on pregnancy.

Stool Microbiology

Participants in cohorts 1 and 2 who reported symptoms of IID were requested to supply a stool sample. Samples were subject to PCR testing at the Liverpool Clinical Laboratories, Liverpool University Hospitals NHS Foundation Trust, to identify the presence of causative organisms (see Table 1 for a list of organisms sought). Commercial assays used for this purpose included Serosep EntericBio DX, Viral and *Clostridioides difficile* assays detecting 15 pathogens, plus additional PCR tests (Genetic Signatures) that target *Clostridium perfringens*, *Aeromonas* spp., *Cyclospora cayetanensis*, and diarrhoeagenic *Escherichia coli* (enterotoxigenic, enteropathogenic, and enteroaggregative strains) [18]. Routine diagnostic laboratory results from cohorts 1 and 2 were reported to the participants' general practitioners.

XSL•FO
RenderX

Table 1. Causative organisms tested in stool samples using PCR[a].

Bacterium	Virus	Parasite
Clostridioides difficile	Adenovirus	*Cryptosporidium parvum/hominis*
Clostridium perfringens	Astrovirus	*Giardia lamblia*
Campylobacter jejuni/coli/lari	Norovirus	*Cyclospora cayetanensis*
Diarrheagenic *Escherichia coli* (DEC)	Rotavirus A	*Entamoeba histolytica*
Enteroaggregative *E. coli* (DEC_EAEC)	Sapovirus	—[b]
Enteropathogenic *E. coli* (DEC_EPEC)	—	—
Enterotoxigenic *E. coli* (DEC_ETEC)	—	—
Salmonella enterica spp.	—	—
Shigella spp.	—	—
Shiga toxin-producing *Escherichia coli* (STEC)	—	—
Yersinia enterocolitica spp.	—	—
Aeromonas spp.	—	—
Vibrio cholerae/parahaemolyticus	—	—

[a]PCR: polymerase chain reaction.

[b]Not applicable.

Further characterization of selected pathogens identified using these primary diagnostic methods will be carried out at the UKHSA's Gastrointestinal Bacteria Reference Unit (bacteria), Enteric Virus Unit (viruses), and Public Health Wales *Cryptosporidium* Reference Unit (parasites).

Cohort 3 data will enable us to report on routine clinical practice, where any stool samples will follow standard diagnostic routes; no additional testing will be performed for these cases. We received data on typing, bacterial multilocus sequence typing (MLST), multiple-locus variable number (MVLA) for parasites, and single-nucleotide polymorphism (SNP) addresses, as well as a variety of AMR data.

Sample Archiving

Stool samples from participants who provided consent for storage are archived at the University of Liverpool's Good Clinical Practice (GCP) Laboratory Biobank for potential use in future research. A sample aliquot is saved in Shield Fluid (Zymo Research) to preserve nucleic acids for future molecular applications, where the material allows.

National Surveillance

Laboratory Pathogen Surveillance

National surveillance data on GI pathogens were provided for each of the four UK nations by the UKHSA, the PHS, the PHW, and the PHA NI. To ensure standardization from these sources, we developed a common script to be used across the four nations to extract and deduplicate the data. Preprocessing and checking of these datasets ensured deduplication based on patient name, NHS number, and date, as well as pathogen genus, species, and serotype. The data provided vary depending on availability by country, but broadly, the datasets include International Organization for Standardization (ISO) week and year, 5-year age group (<1 year, 1-4 years, 5-9 years, and so on), sex, postcode district, whether the case was imported (the UKHSA only), measure of deprivation (eg, the English IMD, Scottish IMD, and Welsh IMD), organism (genus, species, type, and subtype), and antimicrobial sensitivity to different antimicrobials where such tests have been performed.

Syndromic Surveillance

Syndromic surveillance data, in the form of nonidentifiable records of phone calls relating to IID made to the national telephone health services NHS 111 (England) and NHS 24 (Scotland), were also provided by the UKHSA and the PHS, respectively. These data include date of call, syndrome (symptoms reported by the caller, limited to diarrhea or vomiting), 5-year age group, sex, and postcode district.

Data Storage and Security

Data collected for the study were pseudonymized and stored within a secure, trusted research environment (TRE) operated by the University of Oxford. Where available, cohort 1 and cohort 2 sample collection was facilitated through the Lab Links system. The system was intended to minimize the study-related workload of sample processing by enabling practices to follow their usual clinical sample workflow for study samples, allowing them to order tests and receive results. However, as this was not available to most practices, an alternative manual approach, where sample kits were to be returned by post, was provided to participating GPs and patients, as required.

Data Linkage

For cohorts 1 and 2, data from three sources can be linked to an individual participant: the participant's information and symptom data from the study mobile app, results from the study laboratory, and medical history information from the primary care record. The datasets are linked within the secure Oxford Royal College of General Practitioners, Clinical Informatics Digital Hub (ORCHID) TRE using an encrypted key that is consistent across the three data sources. Cohort 3 data will be

XSL•FO
RenderX

cross-checked to ensure no study participant (cohort 1 and cohort 2 participants) appears in the cohort 3 dataset.

Ethical Considerations

Ethical opinion was sought from the UK Health Research Authority using the Integrated Research Applications System (IRAS). Further information governance approvals were obtained from the Public Benefit and Privacy Panel for Health and Social Care (NHS Scotland). A consortia-wide study agreement outlines partner roles and responsibilities, and data-sharing agreements were put in place between organizations for data or material transfer. Approval for the study was obtained from the RCGP, which is a policy-level data controller of the RSC. As mentioned earlier, data were pseudonymized and stored within the secure ORCHID TRE. Access to the data is limited to data analysts on the IID3 study team. The ORCHID TRE meets the requirements of NHS Digital's Data Security and Protection (DSP) Toolkit and is compliant with university policies and relevant legislation. Informed consent was obtained from all adults before entry to the study. For children, consent was sought from a parent or guardian.

Data Analysis

Assessing Representativeness

Representativeness and inequalities will be assessed by comparing the characteristics of the study population with those from routine data from the national census and GP statistics. We will quantify inter-GP variation in patient characteristics, including sociodemographic features associated with the patient population, using mixed-effect modeling. Seasonal variation in the data will be removed through detrending, and factors driving residual nonseasonal disease variation will be determined using count negative binomial (generalized) linear mixed-effects models. Detrending will occur by including harmonic variables (sine and cosine of time) as covariates in the linear mixed-effect models [19]. Model fit will be assessed by comparing Akaike information criterion (AIC) or similar values between models, comparing R^2 values, and examining model diagnostics, such as QQ plots. If any remaining temporal autocorrelation remains in the data, it will be adjusted for using autoregression.

Assessing Compliance

Compliance in cohorts 1 and 2 will be calculated as the proportion of participants submitting a questionnaire or stool sample relative to the total number of participants who had confirmed participation in the presentation study. Differences in compliance across age, sex, or practice groups/areas will be assessed using mixed-effects analyses with GP as a random effect; a Pearson chi-square test and multiple logistic regression will be used.

Assessing Underascertainment

Underascertainment in cohorts 1 and 2 will be investigated by comparing the frequency of IID diagnosis in the GP presentation study data with that ascertained in the enumeration study. This will also be used to assess the likely contribution of unmeasured GP variation on ascertainment. Factors influencing underascertainment will be investigated using multiple logistic regression in a mixed-effects framework (ie, generalized linear mixed-effects models with a binomial error structure). The inverse of the predicted underascertainment from this model will be used as a multiplier for the number of cases reported in each practice, considering practice characteristics and intraclass correlation from the mixed-effects models.

Estimating the IID Incidence

We will exploit the hierarchical structure of patients within a GP as our modeling framework. All analyses of the incidence of individual IID pathogens will be based on PCR results; bacterial PCR-positive samples will be subject to routine culture and identification to allow comparison with the IID1 and IID2 studies.

Incidence Rate of IID

The incidence rate of IID in the community will be calculated using the total number of cases reported in the population (cohort 1) and the follow-up time of cohort members. Follow-up periods will be censored for times when participants were not considered at risk (ie, while traveling). Incidence rates of IID will be calculated as an overall rate and adjusted to account for age and sex.

Incidence Rate of IID Cases Presenting to the GP

The incidence rate of IID cases presenting to GPs (cohort 2) will be calculated using the total number of cases reported in the presentation and enumeration components (cohorts 2 and 3, respectively) and the follow-up time of all cohort members registered in the practice populations. The numerator and denominator of the GP presentation rates will be adjusted to account for underascertainment and list inflation. Information about individual patient symptoms will be incorporated into the analyses, as appropriate.

The Role of Individual Pathogens

Number of Cases by Organism

For cohorts 1 and 2, age and seasonality incidence will be estimated for each target organism using denominators derived from the age/sex registers of the study practices and the sampled population. Organism-specific incidence rates will be calculated from the number of cases in which each target organism was identified, regardless of whether other organisms were present in the same sample. Underascertainment will be assumed to be the same for all organisms. Rates will be adjusted for noncompliance with submitting stool samples based on overall compliance data. We will assume that noncompliant cases have the same identification rate for each organism as compliant cases. Calculation of CIs will be adjusted to ensure that there is no increase in the apparent precision due to this correction for noncompliance.

Comparing the Etiology of IID and Surveillance Pyramids

Incidence Rates and Presentation Rates

Incidence rates and presentation rates (overall for IID and by target organism) will be compared between the IID1 and IID2 studies and the new IID3 study based on the incidence in this research estimated using the laboratory methods used in the

XSL•FO
RenderX

first study; changes in incidence in the community and presentation to GPs will be estimated, together with their associated 95% CIs calculated using a Wald test.

Surveillance Pyramids

Surveillance pyramids will be created using the overall community incidence rates and presented to GPs, with ratios between these rates and their respective CIs included. The ratio of community rates to nationally reported cases will be calculated by projecting the overall cohort and presentation incidence rates to the population and comparing these to ratios reported to the laboratory reporting surveillance system over the same period. When more than one laboratory method is used, the results from both methods will be estimated separately for comparison with the IID1 study and as a baseline for the future.

Three Key Components of the Surveillance Pyramids

Three key components of the surveillance pyramids will be compared across the IID1 and IID2 studies and the new IID3 study: overall incidence rates in the community, presentation to GPs, and reporting to the laboratory surveillance system. A simple comparison will be made between these ratios. Performing the same calculation on the upper and lower CIs for each ratio will derive the upper and lower sensitivity bands.

Molecular Data

Molecular data used to derive some of the surveillance pyramids in the IID1 and IID2 studies were obtained using different molecular methods than those available for the IID3 study. The surveillance pyramid for the IID3 study can be recalculated assuming that the molecular techniques previously used with different targets are now in place. This will allow a more direct comparison between the new IID3 surveillance and the original IID1 and IID2.

Patient and Public Involvement and Engagement

We will have public and patient involvement and engagement (PPIE) at all stages of the research. We received input from patient participation groups in primary care, the PPIE research theme and panel from the NIHR Health Protection Research Unit in Gastrointestinal Infections, and PPIE representatives on the IID3 Study Executive Committee and the External Advisory Panel. We sought and incorporated feedback on how to phrase questions about sex/gender and the material's reading level for children. A primary school teacher gave us feedback, and they said the material should be suitable for a reading age of 7-8 years. We also incorporated feedback from older members of the community to ensure the material is unlikely to cause offence. PPIE involvement will continue throughout the study's life cycle.

Results

A favorable ethical opinion was obtained from the UK Health Research Authority on August 4, 2022 (IRAS ID 314268, REC Reference 22/EM/0130). A pilot phase to test the sampling process was conducted from January to August 2023. Participant recruitment commenced on September 1, 2023, for cohort 2 and on March 16, 2024, for cohort 1; participation in both cohorts ceased on August 31, 2025 [20]. Data collection is complete, and data analysis is to begin. The results will be disseminated to key stakeholders, including the FSA, participating GPs, and the academic community, through reports, publications, and conference presentations. The study is expected to end in September 2026. The study-generated data will be archived after study completion and will be available to the general public.

Discussion

Summary

Our ability to determine the true burden of IID in a population remains a challenge, as only a small proportion of cases appear in national surveillance statistics. Large-scale interdisciplinary prospective cohort studies are critical to maintaining our understanding of disease incidence and causative pathogens. This study is timely as it will provide new estimates of the disease burden postpandemic and also follows the introduction of the new rotavirus vaccine in 2013 in children under 5 years of age. New pathogen-specific incidence rates will inform policy and pathogen management, such as the rollout of the norovirus vaccine currently under development.

All stool samples in the study were subject to sensitive molecular assays, using significant advancements in molecular diagnostics since the IID2 study, to allow detection of a broad range of bacteria, viruses, and parasites. Further characterization of positive samples by national reference laboratories in England and Wales will provide additional information about the strain types circulating in the community. The enhanced microbiological testing will also improve our understanding of pathogen detection through routine surveillance and where this could be improved. The revised incidence rates across the study surveillance pyramids will allow recalibration of the national surveillance of IID, while accounting for improvements in our ability to detect disease since the previous estimates.

The IID3 study methodology will provide valuable insight into changing patterns of primary care reporting practices following the worldwide COVID-19 pandemic. The past 5 years have seen substantial changes to how patients access primary care in the United Kingdom, with more people relying on electronic communications (NHS 111/NHS 24 and GP messaging systems) to report illness and seek advice, when they may previously have had an appointment with their GP. We anticipate this to be reflected by a change in IID incidence rates for the GP presentation cohort.

Strengths and Limitations

Although the study uses the same overall methodology as the previous IID1 and IID2, the digital recruitment approach for cohort 1 is novel, using a direct message system to reach the community population through GP registrations [20]. Participant follow-up was conducted via a mobile app to enable real-time collection of weekly questionnaire data, removing the need for paper-based communications and data entry.

We recognize that digital study recruitment may have limited access to the study to those who are permanently registered with a GP and have access to a phone or email to complete the weekly

XSL•FO
RenderX

questionnaires. Recruitment of children (under 18 years) was challenging due to limitations with direct contact and consenting minors through text message invitation and the mobile app, which may result in underrepresentation of viral causes of IID, as these are likely to be more dominant in young children. However, embedding the study within the Oxford RCGP Research and Surveillance Centre primary care sentinel surveillance network has facilitated recruitment from nationally representative research-ready GPs with experience in undertaking primary care research, including on IID [17].

Conclusion

Here, we describe the protocol of the IID3 study in the United Kingdom, which will provide contemporary estimates of the disease burden in the UK population. Updated IID reporting and incidence rates are urgently needed to reflect changes in primary care interaction and improved diagnostic capability and to recalibrate national surveillance reporting. This prospective cohort study approach will provide comparable rate estimates to previous studies (IID1 and IID2 studies), putting us in the unique position of being able to report trends over three decades. The World Health Organization, national governments, and public health organizations rely on these robust estimates of disease incidence and underreporting as evidence to inform policy agendas, interventions. and decision-making.

Acknowledgments

We acknowledge fundings for this research from the Food Standards Agency (FSA) and Food Standards Scotland. We would also like to thank all the technology suppliers that have enabled participant recruitment: PeopleWith, who developed the patient recruitment platform (Mark Bradley and Christopher Johnston), and the OPTUM recruit team (Hazel Rogers, Amal Qureshi, Paul Davies, and Jamie Cole), who supported the participant invitation process through the Optum clinical system. We are grateful to patients and general practices in the Oxford Royal College of General Practitioners Research and Surveillance Centre who agreed to data sharing for research and surveillance and OPTUM, TPP, In-Practice Systems, and Magentus for collaboration in pseudonymized data extraction. We thank National Health Service (NHS) 111 and NHS 24 for their support in provision of syndromic surveillance data and Danielle McMichael from Public Health Agency Northern Ireland for support with national surveillance records. The full research was implemented by the IID3 study consortium. This study has been funded by the FSA (project code FS301058). The views expressed are those of the authors and not necessarily those of the NIHR, the Department of Health and Social Care, or the UKHSA.

The IID3 consortium as of Feb 2026 is comprised of Aileen C Mill, Simon de Lusignan, Ben W Rowland, Elizabeth Button, Steve Rushton, Roy Sanderson, Dominic J Mellor, Daniel Hungerford, Emma Ackermann, Mark Hopkins, Angela Hughes, Michael McDowell, Farieda Kassim, Gauri S Godbole, Kristin Elwin, Rachel M Chalmers, Anna R Lawson, Claire Jenkins, Nigel A Cunliffe, Caroline Grundy, Sarah O'Brien, Vanashree Sexton, Uy Hoang, Filipa Ferreira, Christopher Williams, Roberto Vivancos, Andrew John Dudley Nelson, Annemarie van Heelsum, Christina Bronowski, Edward Cunningham-Oakes, Alistair Darby, Nichola Love, Danielle McMicheal, Karin André, Alex J Elliot, Melanie Oates, Anna L Donaldson, Tianlai Yang, Jack Macartney, Darius Oraee, Daniel Gallaher, Eva Klamerus, Timea Süli, Jessica Smylie, Gunjan Jiwnani, Rosalind Goudie, and Mili Muraleedharan. We thank all consortium members for their contributions to the IID3 study.

Data Availability

The data collected from the Third Study of Infectious Intestinal Disease (IID3 study) will be archived in the UK Data Archive at the end of the study (September 2026). At this time, the data will be publicly available. National surveillance data will not be archived and will require data-sharing agreements with the relevant public health agency to obtain.

Authors' Contributions

Conceptualization and funding acquisition were handled by SOB, AM, SR, RS, SdL, NAC, DH, MH, SG, CJ, GG, LL, RV, AJE, DJM, and RC; data curation and writing—original draft by BWR and VS; investigation by SOB, AM, SR, RS, SdL, NAC, DH, MH, SG, CJ, GG, LL, RV, AJE, DJM, RC, BWR, VS, and CG; methodology and writing—review and editing by SOB, AM, SR, RS, SdL, NAC, DH, MH, SG, CJ, GG, LL, RV, AJE, DJM, RC, BWR, and VS; project administration by SOB, AM, and CG; software by BWR; and supervision by SOB, AM, SR, BWR, SdL, NAC, and DJM. BWR and VS were joint first authors of the manuscript and contributed equally to the drafting and editing process.

Conflicts of Interest

NAC is a National Institute for Health and Care Research (NIHR) senior investigator (NIHR203756). NAC, CJ, AJE, RV, and DH are affiliated to the NIHR Health Protection Research Unit (HPRU) in Gastrointestinal Infections at the University of Liverpool, a partnership with the United Kingdom Health Security Agency (UKHSA) in collaboration with the University of Warwick. NAC is based at the University of Liverpool. DH is in receipt of grants from Sanofi Pasteur and Merck and Co (Kenilworth, NJ) for rotavirus strain surveillance, has received honorariums for presentation at a Merck Sharp and Dohme (UK) symposium on vaccines, and has consulted on rotavirus strain surveillance. RV, CJ, AJE and BR receive funding from the NIHR HPRU in Gastrointestinal

XSL•FO
RenderX

Infections at the University of East Anglia. RV receives funding from the NIHR HPRU in Emerging and Zoonotic Infections at the University of Liverpool. AJE receives funding from the NIHR HPRU in Emergency Preparedness and Response and the University of Birmingham. SdeL is the Director of the Oxford-Royal College of General Practitioners Research and Surveillance Centre (RSC) sponsored and run in partnership with UKHSA. Its data infrastructure and practice network were used to support IID3.

Multimedia Appendix 1
Cohort symptom questionnaire for adults.
[DOC File , 113 KB - resprot_v15i1e88759_app1.doc]

Multimedia Appendix 2
Cohort symptom questionnaire for children.
[DOC File , 115 KB - resprot_v15i1e88759_app2.doc]

Multimedia Appendix 3
GP symptom questionnaire. GP: general practice.
[DOC File , 459 KB - resprot_v15i1e88759_app3.doc]

References

1. Tam C, Viviani L, Adak B, Bolton E, Dodds J, Cowden J, et al. The second study of infectious intestinal disease in the community (IID2 study). Food Standards Agency. 2012. URL: https://www.food.gov.uk/research/foodborne-disease/the-second-study-of-infectious-intestinal-disease-in-the-community-iid2-study [accessed 2026-01-29]
2. Daniel N, Casadevall N, Sun P, Sugden D, Aldin V. The burden of foodborne disease in the UK 2018. Food Standards Agency. 2020. URL: https://www.food.gov.uk/sites/default/files/media/document/the-burden-of-foodborne-disease-in-the-uk_0.pdf [accessed 2026-01-29]
3. Adams N, Rose T, Hawker J, Violato M, O'Brien SJ, Barr B, et al. Relationship between socioeconomic status and gastrointestinal infections in developed countries: a systematic review and meta-analysis. PLoS One 2018;13(1):e0191633 [FREE Full text] [doi: 10.1371/journal.pone.0191633] [Medline: 29360884]
4. Buczkowska M, Butt S, Jenkins C, Hungerford D, Hawker J, Verlander N, et al. Association between socioeconomic deprivation and incidence of infectious intestinal disease by pathogen and linked transmission route: an ecological analysis in the UK. Epidemiol Infect 2023 Jun 14;151:e109 [FREE Full text] [doi: 10.1017/s0950268823000869]
5. Payment P. Transmission of gastrointestinal diseases: hygiene as the final barrier. Am J Infect Control 2001 Aug;29(4):218-221 [FREE Full text] [doi: 10.1067/mic.2001.115683] [Medline: 11486257]
6. Hughes C, Gillespie I, O'Brien S. Foodborne transmission of infectious intestinal disease in England and Wales, 1992–2003. Food Control 2007 Jul;18(7):766-772 [FREE Full text] [doi: 10.1016/j.foodcont.2006.01.009]
7. Sanderson R, Maas J, Blain A, Gorton R, Ward J, O'Brien SJ, et al. Spatio-temporal models to determine association between Campylobacter cases and environment. Int J Epidemiol 2018 Feb 01;47(1):202-216 [FREE Full text] [doi: 10.1093/ije/dyx217] [Medline: 29069406]
8. Wheeler J, Cowden J, Cumberland P, Roberts J, Roderick P, Rodrigues L, et al. A report of the study of infectious intestinal disease in England. Food Standards Agency. 2000. URL: https://www.food.gov.uk/sites/default/files/media/document/iid1_study_final_report.pdf [accessed 2026-01-29]
9. Peek N, Sujan M, Scott P. Digital health and care in pandemic times: impact of COVID-19. BMJ Health Care Inform 2020 Jun 21;27(1):e100166 [FREE Full text] [doi: 10.1136/bmjhci-2020-100166] [Medline: 32565418]
10. Love NK, Elliot AJ, Chalmers RM, Douglas A, Gharbia S, McCormick J, et al. Impact of the COVID-19 pandemic on gastrointestinal infection trends in England, February-July 2020. BMJ Open 2022 Mar 21;12(3):e050469 [FREE Full text] [doi: 10.1136/bmjopen-2021-050469] [Medline: 35314468]
11. National norovirus and rotavirus report, week 45 report: data to week 44 (data up to 2 November 2025). UK Health Security Agency. 2025. URL: https://www.gov.uk/government/statistics/national-norovirus-and-rotavirus-surveillance-reports-2025-to-2026-season/national-norovirus-and-rotavirus-report-week-45-report-data-to-week-44-data-up-to-2-november-2025 [accessed 2026-01-29]
12. Harris J, Jit M, Cooper D, Edmunds W. Evaluating rotavirus vaccination in England and Wales. Part I. Estimating the burden of disease. Vaccine 2007 May 16;25(20):3962-3970 [FREE Full text] [doi: 10.1016/j.vaccine.2007.02.072] [Medline: 17395343]
13. Tam C, Rodrigues L, Viviani L, Dodds J, Evans M, Hunter P, IID2 Study Executive Committee. Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. Gut 2012 Jan;61(1):69-77 [FREE Full text] [doi: 10.1136/gut.2011.238386] [Medline: 21708822]
14. Atchison CJ, Stowe J, Andrews N, Collins S, Allen DJ, Nawaz S, et al. Rapid declines in age group–specific rotavirus infection and acute gastroenteritis among vaccinated and unvaccinated individuals within 1 year of rotavirus vaccine

XSL•FO
RenderX

introduction in England and Wales. J Infect Dis 2016 Jan 15;213(2):243-249 [FREE Full text] [doi: 10.1093/infdis/jiv398] [Medline: 26232438]
15. Hungerford D, Vivancos R, Read J, Iturriza-Gȯmara M, French N, Cunliffe N. Rotavirus vaccine impact and socioeconomic deprivation: an interrupted time-series analysis of gastrointestinal disease outcomes across primary and secondary care in the UK. BMC Med 2018 Jan 29;16(1):10 [FREE Full text] [doi: 10.1186/s12916-017-0989-z] [Medline: 29375036]
16. O'Brien SJ, Rait G, Hunter P, Gray J, Bolton F, Tompkins D, et al. Methods for determining disease burden and calibrating national surveillance data in the United Kingdom: the second study of infectious intestinal disease in the community (IID2 study). BMC Med Res Methodol 2010 May 05;10:39 [FREE Full text] [doi: 10.1186/1471-2288-10-39] [Medline: 20444246]
17. Leston M, Elson W, Watson C, Lakhani A, Aspden C, Bankhead C, et al. Representativeness, vaccination uptake, and COVID-19 clinical outcomes 2020-2021 in the UK Oxford-Royal College of general practitioners research and surveillance network: cohort profile summary. JMIR Public Health Surveill 2022 Dec 19;8(12):e39141 [FREE Full text] [doi: 10.2196/39141] [Medline: 36534462]
18. Bronowski C, Hughes A, Jenkins C, Godbole G, Chalmers R, Celma C, the IID3 Consortium. The third study of infectious intestinal disease in the community microbiological methods. medRxiv. Preprint posted online 2025. [doi: 10.1101/2025.11.20.25340668] [FREE Full text] [doi: 10.1101/2025.11.20.25340668]
19. Rowland BW, Rushton SP, Shirley MDF, Brown MA, Budge GE. Identifying the climatic drivers of honey bee disease in England and Wales. Sci Rep 2021 Nov 09;11(1):21953 [FREE Full text] [doi: 10.1038/s41598-021-01495-w] [Medline: 34754028]
20. Oraee D, Button E, Sexton V, Andre K, Ferreira F, Mill A, et al. Implementation and evaluation of automated, online study recruitment from computerised medical records in a primary care sentinel surveillance network. Stud Health Technol Inform 2024 Nov 22;321:124-128. [doi: 10.3233/SHTI241076] [Medline: 39575793]

Abbreviations

AMR: antimicrobial resistance
FSA: Food Standards Agency
GI: gastrointestinal
GP: general practice
IID: infectious intestinal disease
IID1 study: First Study of Infectious Intestinal Disease
IID2 study: Second Study of Infectious Intestinal Disease
IID3 study: Third Study of Infectious Intestinal Disease
IMD: index of multiple deprivation
NHS: National Health Service
ORCHID: Oxford Royal College of General Practitioners, Clinical Informatics Digital Hub
PCR: polymerase chain reaction
PHA NI: Public Health Agency Northern Ireland
PHS: Public Health Scotland
PHW: Public Health Wales
PPIE: public and patient involvement and engagement
RCGP: Royal College of General Practitioners
RSC: Research and Surveillance Centre
SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms
TRE: trusted research environment
UKHSA: United Kingdom Health Security Agency

Edited by J Sarvestan; submitted 01.Dec.2025; this is a non–peer-reviewed article;accepted 19.Jan.2026; published 25.Feb.2026.

Please cite as:
Rowland BW, Sexton V, Mill A, Rushton S, Sanderson R, Grundy C, de Lusignan S, Cunliffe NA, Hungerford D, Hopkins M, Gharbia S, Jenkins C, Godbole G, Vivancos R, Elliot AJ, Mellor DJ, Larkin L, Chalmers R, O'Brien S, IID3 Consortium
The Third Study of Infectious Intestinal Disease (IID3 Study) in the Community: Protocol for UK-Based Prospective Cohort Studies Investigating the Disease Burden
JMIR Res Protoc 2026;15:e88759
URL: https://www.researchprotocols.org/2026/1/e88759
doi:10.2196/88759
PMID:

XSL•FO
RenderX

©Ben W Rowland, Vanashree Sexton, Aileen Mill, Stephen Rushton, Roy Sanderson, Caroline Grundy, Simon de Lusignan, Nigel A Cunliffe, Daniel Hungerford, Mark Hopkins, Saheer Gharbia, Claire Jenkins, Gauri Godbole, Roberto Vivancos, Alex J Elliot, Dominic J Mellor, Lesley Larkin, Rachel Chalmers, Sarah O'Brien, IID3 Consortium. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Culturally Tailored Mobile Health Intervention to Improve Quality of Life in Black Survivors With Prostate Cancer: Protocol for a Stratified Randomized Controlled Trial

Gaurav Kumar[1], MBBS, MPH, PhD; Parisa Ghasemi[1], MD; Adam C Alexander[1,2], PhD; Kathleen Dwyer[3], PhD; Jordan M Neil[1,2], PhD; Perry Cole[4], AA; Patrick Beckford[4], BS; Donald Reese[4], MA, CNA; Everett Montgomery[4], BS, MACC, DBA; Roland Odeleye[4], BSCS, MBA, JD; Mark A Harris[4], CES, CPT; Yan D Zhao[5], PhD; Zsolt Nagykaldi[2], BTh, PhD; Andrew G McIntosh[6], MD; Ernie Kaninjing[7], MPH, MBA, CHES, DrPH; Mary E Young[8], PhD; Sabrina Dickey[9], MSN, PhD; Daniel J Morton[10,11], PhD; Opeyemi Bolajoko[12], PhD; Folakemi T Odedina[12], PhD; Lourdes G Planas[13], RPh, PhD; Michael S Businelle[1,2], PhD; Darla E Kendzor[1,2], PhD; Motolani E Ogunsanya[1,2], DPH, PhD

[1]TSET Health Promotion Research Center, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[2]Department of Family and Preventive Medicine, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[3]Fran and Earl Ziegler College of Nursing, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[4]The Multidisciplinary Health Outcomes Research and Economics (MORE) Lab Community Advisory Board, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[5]Department of Biostatistics and Epidemiology, Hudson College of Public Health, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[6]Department of Urology, Stephenson Cancer Center, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[7]School of Health and Human Performance, Georgia College and State University, Milledgeville, GA, United States

[8]Mayo Foundation for Medical Education and Research, Jacksonville, FL, United States

[9]College of Nursing, Florida State University, Tallahassee, FL, United States

[10]Department of Pediatrics, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[11]Stephenson Cancer Center, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[12]Mayo Clinic Comprehensive Cancer Center, Jacksonville, FL, United States

[13]College of Pharmacy, University of Oklahoma Health Campus, Oklahoma City, OK, United States

Corresponding Author:
Gaurav Kumar, MBBS, MPH, PhD
TSET Health Promotion Research Center
University of Oklahoma Health Campus
655 Research Parkway
Oklahoma City, OK, 73104
United States
Phone: 1 4052718001 ext 18001
Email: gaurav-kumar-3@ou.edu

Abstract

Background: Prostate cancer (CaP) disproportionately affects Black men in the United States, leading to significant disparities in incidence, survival, and quality of life (QoL). Treatment-related side effects, including urinary dysfunction, pain, fatigue, and psychological distress, contribute to poor long-term outcomes. There is an urgent need for culturally-tailored, technology-based interventions to support symptom self-management and survivorship care.

Objective: This study aims to develop, refine, and evaluate the Survivorship App for Ethnically Diverse Black Prostate Cancer Survivors (SAFE-CaPS), a tailored mobile health intervention designed to enhance QoL, improve symptom self-management, and provide psychosocial support.

Methods: This 12-month stratified randomized clinical trial will enroll 248 Black survivors with CaP (including native born, African-born, and Caribbean-born men) into the SAFE-CaPS intervention arm and standard care (SC) control group (1:1 allocation). Eligible participants are self-identified Black men (including native born, African-born, and Caribbean-born) aged 40 to 80 years

XSL•FO
RenderX

with a confirmed CaP diagnosis within the past 5 years who own or are willing to use a smartphone. SC will consist solely of routine oncology or primary care follow-up, including clinic visits, referrals, and supportive services as needed, with no digital components. The SAFE-CaPS intervention, developed using the Insight platform, includes daily and weekly ecological momentary assessments of pain, urinary and bowel symptoms, sleep, fatigue, sexual function, mood, physical activity, and diet; adaptive educational modules addressing treatment-related side effects and culturally grounded survivorship concerns; automated symptom alerts prompting nurse follow-up and provider notification; and engagement features such as reminders, goal-setting, and literacy-appropriate content. Intervention content incorporates culturally tailored messaging, patient narratives, and feedback from prior qualitative work with ethnically diverse Black CaP. Stratified randomization will ensure balanced representation of ethnic subgroups. Primary outcomes include overall QoL. Secondary outcomes include specific symptom domains, mental health (depression and anxiety), health care engagement, and app acceptability. Assessments will occur at baseline, 3, 6, 9, and 12 months, and linear mixed models with intent-to-treat principles will evaluate intervention effects over time.

Results: The study was funded in 2024, institutional review board approval was received in mid-2024, and the study began recruitment on August 24, 2025. Overall, 6 participants have been enrolled, with 2 in the SAFE-CaPS intervention and 4 in SC. Recruitment is expected to conclude in July 2027. Data analysis will occur in late 2027, with article preparation and publication planned for mid-2028 and early 2029. We hypothesize that SAFE CaPS participants will experience greater improvements in QoL, reduced symptom burden, and higher engagement in care compared to SC.

Conclusions: This study will provide critical evidence on the feasibility, acceptability, and preliminary effectiveness of a culturally tailored mobile health intervention designed to enhance survivorship outcomes among Black survivors with CaP, informing future large-scale trials and real-world implementation.

Trial Registration: ClinicalTrials.gov NCT06651359; https://clinicaltrials.gov/study/NCT06651359

International Registered Report Identifier (IRRID): PRR1-10.2196/81503

(JMIR Res Protoc 2026;15:e81503) doi:10.2196/81503

KEYWORDS

Black men; prostate cancer; cancer survivorship; symptoms monitoring; quality of life; mobile health; culturally-tailored interventions

Introduction

Prostate cancer (CaP) remains a leading cancer diagnosis among men in the United States, with 313,780 new cases and 35,770 deaths projected in 2025 [1]. Among racial groups, Black men have the highest incidence rates nationwide at 191.5 cases per 100,000 compared to 114.5 per 100,000 among white men, and a higher CaP mortality rate (37.2 vs 18.1 deaths per 100,000, respectively) [2]. Black men have a greater likelihood of more advanced or aggressive disease at diagnosis, and they are frequently diagnosed at later disease stages, which limits treatment options and contributes to higher morbidity and mortality rates [3,4].

Disparities in CaP survival among Black men are influenced by genetic susceptibility, limited access to health care, family history, socioeconomic factors, and delayed detection [5,6]. Despite these disparities, the overall 5-year survival rate for Black men with CaP is high at approximately 97% [2]. However, survival does not equate to optimal posttreatment outcomes, as many Black men experience a disproportionate burden of treatment-related complications, psychosocial distress, and lower quality of life (QoL) [4,7-9]. CaP survivorship presents unique and persistent challenges, particularly for Black men, who often face disproportionate burdens of physical, emotional, and social distress [5,10]. Long-term treatment-related side effects such as sexual dysfunction, bowel and urinary incontinence, and mobility impairments can significantly diminish QoL [11-14]. Additionally, concerns about cancer recurrence, medical mistrust, and limited access to survivorship resources create further barriers to effective coping and reintegration into daily life [15-17]. These issues are often intensified by stigma surrounding sexual dysfunction, cultural beliefs about masculinity, and discomfort discussing intimate symptoms, which can delay help-seeking among Black men [9]. Addressing these multifaceted survivorship needs requires culturally tailored interventions that improve symptom management and psychosocial support. While many survivors with CaP face physical and emotional challenges, Black men often experience lower levels of functioning in these areas compared to White men [18-20]. Studies report slower recovery rates, decreased satisfaction with recovery, and higher levels of regret following prostatectomy among Black men [3,21,22].

The Black population in the United States is increasingly ethnically diverse, encompassing native-born Black men, African-born Black men, and Caribbean-born Black men [23,24]. While these groups share genetic similarities, their cultural beliefs, health care–seeking behaviors, and survivorship experiences vary significantly [25,26]. Across the African diaspora, Black men are at particularly high risk for CaP, but this risk is not uniform: native-born Black men have higher CaP mortality than African-born Black men and Caribbean-born Black men living in the United States, whereas some Caribbean-born Black men populations show an even greater likelihood of aggressive disease [27]. These patterns are likely driven by a combination of biological, environmental, behavioral, dietary, and geographic factors, including differences in risk exposures, screening practices, and access to high-quality care [4,28]. Studies suggest that perceptions of CaP, symptom management strategies, and engagement with health care differ

XSL•FO
RenderX

based on cultural background, highlighting the need for culturally tailored research and interventions to address CaP disparities [8,9,29]. These needs also vary by ethnicity, particularly in areas such as communication about sexual health, stigma surrounding symptom disclosure, preferences for family involvement, and expectations of provider support, further reinforcing the need for tailored survivorship strategies [8,9,30]. Despite the high survivorship burden, Black men often report slower recovery rates, lower satisfaction with treatment outcomes, and higher levels of regret following prostatectomy compared to White men [18,20,31-33]. At the same time, many face systemic barriers to posttreatment support, including financial constraints, geographic challenges, and medical mistrust. These barriers may limit participation in traditional in-person survivorship programs, underscoring the need for scalable, culturally responsive approaches that address ethnic variation in survivorship priorities [3,34].

Mobile health (mHealth) interventions have shown great promise for supporting survivors with cancer by providing easy access to information, symptom tracking, and virtual counseling [35-39]. This technology is highly relevant for Black men with CaP, as more than 85% of Black Americans own smartphones [40], allowing them to access health resources remotely and overcome traditional barriers to care [41,42]. mHealth solutions also provide a convenient and scalable way to deliver tailored survivorship support, reducing reliance on in-person visits, which may be infrequent, inaccessible, or underused among Black men [39]. Additionally, these tools can deliver real-time symptom tracking, personalized resources, and interactive peer support, which have been shown to improve health outcomes in other cancer populations [39,43]. However, there are very limited existing mHealth interventions focusing on CaP [39,44], and none that are specifically designed for Black men or integrate culturally responsive features to enhance engagement and accessibility [39,44-46]. Given the high smartphone adoption rate among Black men and the potential for digital health interventions to reduce disparities, there is an urgent need to develop a culturally tailored mHealth solution that enhances QoL and addresses the unique survivorship challenges faced by Black survivors with CaP.

Given the persistent symptom burden, cultural variation in survivorship experiences, and limited engagement with supportive care among ethnically diverse Black men with prostate cancer, there is a need for theoretical frameworks that can both inform clinical targets and guide user engagement with digital resources. The Revised Wilson and Cleary Model (WCM) of health-related quality of life [47,48] provides a clinically grounded pathway linking biological function, symptom experience, functional status, and overall QoL, offering a clear rationale for monitoring and managing treatment-related side effects that disproportionately affect Black survivors with CaP. Complementing this approach, the Unified Theory of Acceptance and Use of Technology (UTAUT) [49,50] addresses factors that influence technology adoption and sustained use, including usability, cultural relevance, trust, social influence, and access to supportive conditions. Together, these frameworks provide a comprehensive foundation for designing interventions that are both clinically meaningful and acceptable to diverse Black survivor populations with CaP.

Therefore, this study aims to develop, culturally tailor, and evaluate the Survivorship App for Ethnically Diverse Black Prostate Cancer Survivors (SAFE-CaPS), a mHealth intervention designed to improve QoL, enhance symptom self-management, and strengthen psychosocial support among native-born, African-born, and Caribbean-born Black survivors with CaP. Specifically, this protocol describes the design and methodology of a stratified randomized clinical trial comparing SAFE-CaPS to standard care (SC). We hypothesize that participants receiving SAFE-CaPS will experience greater improvements in overall QoL, reductions in treatment-related symptom burden, and higher engagement in survivorship care compared with those receiving SC.

Methods

Study Design

This 12-month, 2-arm, stratified randomized clinical trial will use purposive sampling. A total of 248 Black survivors with CaP will be allocated in a 1:1 ratio to one of two study arms: (1) The SAFE-CaPS arm (n=124) where participants will receive access to the SAFE-CaPS app in addition to SC; or (2) Control arm (n=124) where participants will receive SC as recommended by their oncologist or other health care provider. SC will vary across participants and may include provider visits, referrals to specialists as indicated, and opportunities to receive support as needed. A stratified randomization approach will ensure a balanced representation of participants across diverse ethnic backgrounds. This study was reviewed and funded by the Department of Defense Prostate Cancer Research Program under Award Number HT942525PCRPIDA (Multimedia Appendix 1). This protocol was developed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2025 guideline [51]. The completed SPIRIT checklist is included in Multimedia Appendix 2.

Inclusion and Exclusion Criteria

Participants will be eligible if they (1) self-identify as Black, including native-born Black men, Caribbean-born Black men, and African-born Black men; (2) have a self-reported confirmed diagnosis of CaP within the last 5 years; (3) are between the ages of 40 and 80; and (4) own or will use a smartphone (Android or iOS).

Participants will be excluded if they do not meet the eligibility criteria listed above or if they are (1) unable or unwilling to provide a valid Social Security Number or Individual Taxpayer Identification Number, or do not meet U.S. residency requirements needed to comply with university reporting and compensation policies; or (2) currently enrolled in another structured CaP survivorship or support program.

Study Recruitment and Procedure

Using purposive and snowball sampling, participants will be recruited through multiple channels across the United States via our current registries (more than 250 registered survivors with CaP), community-based approaches including our

XSL•FO
RenderX

Community Advisory Board [52] members [30], fraternities, and the Inclusive Cancer Care Research Equity [48] for Black Men Consortium [53] and other existing research networks and clinics. The goal is to enroll a diverse population of Black survivors with CaP who meet the study criteria.

The first recruitment strategy will involve active collaboration with 1 of the co-investigators (medical oncologist), allowing recruitment through the urologic cancer clinics at the OU Health Stephenson Cancer Center (SCC). Clinical staff will identify eligible patients and assist with data collection during clinic visits. The second recruitment strategy will leverage the Clinical Research Data Warehouse at the SCC, enabling targeted recruitment. An electronic medical record data pull will identify potentially eligible participants. Identified participants will be contacted via mail and or phone. The third strategy will use existing research networks in Florida and Georgia, as well as a database of 108 Black survivors with CaP maintained by the senior author (MO). National social media recruitment will be conducted via ResearchMatch [54], a national health volunteer registry, in collaboration with an in-house marketing team to promote study participation through digital channels. Finally, community-based recruitment will be implemented through partnerships with the SCC Community Outreach and Engagement Core, the SCC African American Cancer Research CAB, and established community relationships, including faith-based organizations, the Prostate Cancer Survivorship CAB [30], the Inclusive Cancer Care Research Equity for Black Men Consortium [53], and the Mayo Clinic Minority Health & Health Equity Research Support initiatives [55]. Community education events will serve as engagement opportunities to enhance recruitment. A 1-page QR code-embedded recruitment flyer containing study details, eligibility criteria, and the study phone number will be distributed at community organizations, churches, and barbershops frequented by Black men. The multipronged recruitment strategy outlined has been successfully used in Black communities and is designed to address potential reluctance to participate in clinical trials [56,57]. Interested individuals will complete an online screening form or call the study phone number.

Participants in the SAFE-CaPS intervention group who do not own a smartphone will be provided with an Android smartphone and basic cell phone service to ensure equitable access. Those who already own an Android or Apple smartphone will use their personal device. All participants randomized to the intervention will receive a unique, single-use code to download the Insight app developed by the SCC mHealth Technology Shared Resource [58]. Upon entering the code, they will complete the baseline assessment and have access to the app for a 12-month period. Participants will receive practical instructions on using the app from the research team and there will be an option for a video tutorial outlining the features, and a video tutorial outlining the app's features will be available during the initial visit. This tutorial video can be saved to the phone's home screen for future reference, if needed by the participants. Participants will also be instructed on how to exclude the Insight app from “Do Not Disturb” mode settings on Android and Apple devices to ensure that ecological momentary assessments (EMAs) are received and completed in a timely manner. This guidance aims to reduce missed notifications due to phone settings and enhance EMA completion rates. The Insight app will serve as the platform for assessment and intervention for study groups. During the baseline assessment, participants will receive instructions on app features and complete prompted EMAs. These instructions are also available on demand via the “App Instructions” button on the app's home screen. Participants will complete a weekly, app-prompted EMA throughout the study. The app will also prompt participants to complete follow-up surveys at months 3,6, 9, and 12. Additionally, a subset of participants in the intervention group (n=30) will be invited to schedule an exit qualitative interview with a research team member to gather feedback on their experiences using the app.

Participants in the control arm will receive SC as directed by their oncology providers, which may include clinic visits, specialist referrals, or support services. No additional educational materials, digital tools, or support will be provided by the study team beyond the scheduled assessments. They will not have access to the Insight app or EMAs. Participants in the control group will complete baseline and follow-up surveys through a secure REDCap (Research Electronic Data Capture; Vanderbilt University) link [59]. QoL assessments will be completed via REDCap at months 3, 6, 9, and 12. The planned flow of participants through screening, enrollment, randomization, intervention, and follow-up assessments is presented in Figure 1.

Figure 1. Participant flow for the Survivorship App for Ethnically Diverse Black Prostate Cancer Survivors stratified randomized clinical trial.

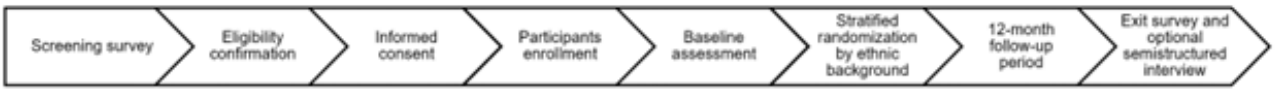

Study Randomization

This study will use a 1:1 stratified randomized controlled trial design and will enroll 248 participants who will be randomly assigned to either the SAFE-CaPS intervention group (n=124) or the control group (n=124) using REDCap’s stratified randomization module. Stratification will be based on ethnic background to ensure a balanced representation across study arms: native-born Black men (n=100), African-born Black men (n=74), and Caribbean-born Black men (n=74). Within each subgroup, participants will be randomized in equal numbers to the intervention and control arms (native-born Black men: 50/50; African-born Black men: 37/37; Caribbean-born Black men: 37/37). This stratified approach ensures balanced subgroup representation and supports exploratory analyses assessing potential differences in survivorship experiences and intervention response across ethnic backgrounds.

Due to the nature of the intervention, blinding of participants, care providers, and recruitment staff is not feasible. However,

XSL•FO
RenderX

the data analyst will be blinded to group allocation to reduce analysis bias.

Mobile Health Platform

The Insight mHealth Platform [58], developed by the SCC mHealth Technology Shared Resource [60], serves as the foundation for the SAFE-CaPS intervention. This state-of-the-art platform supports the rapid design, testing, and implementation of digital health studies, including secure EMAs, adaptive learning modules, automated symptom alerts, and provider notification tools, all within a HIPAA (Health Insurance Portability and Accountability Act)-compliant, encrypted infrastructure. EMA data are stored locally on study smartphones in a sandboxed, encrypted database and securely transmitted via encrypted channels to enterprise-grade Microsoft Azure servers, ensuring data confidentiality, integrity, and protection against unauthorized access. The platform has supported a broad portfolio of National Institutes of Health–funded clinical trials across cancer care, tobacco cessation, HIV prevention, chronic pain, and behavioral health [61-65], including studies specifically tailored for racial and ethnically diverse adults [61,66,67]. This evidence base demonstrates the platform's scalability, cultural adaptability, and suitability for populations experiencing health inequities. Insight was selected for SAFE-CaPS because it provides oncology-specific safety features (eg, distress monitoring and nurse alerts), literacy-appropriate content options, and flexible cultural adaptation capabilities. The research team has collaborated closely with the mHealth Technology Shared Resource to develop and refine the SAFE-CaPS content and usability (Figure 2), ensuring alignment with study goals and adherence to clinical and ethical standards.

XSL•FO
RenderX

Figure 2. Home page of the Survivorship App for Ethnically Diverse Black Prostate Cancer Survivors (SAFE-CaPS) app.

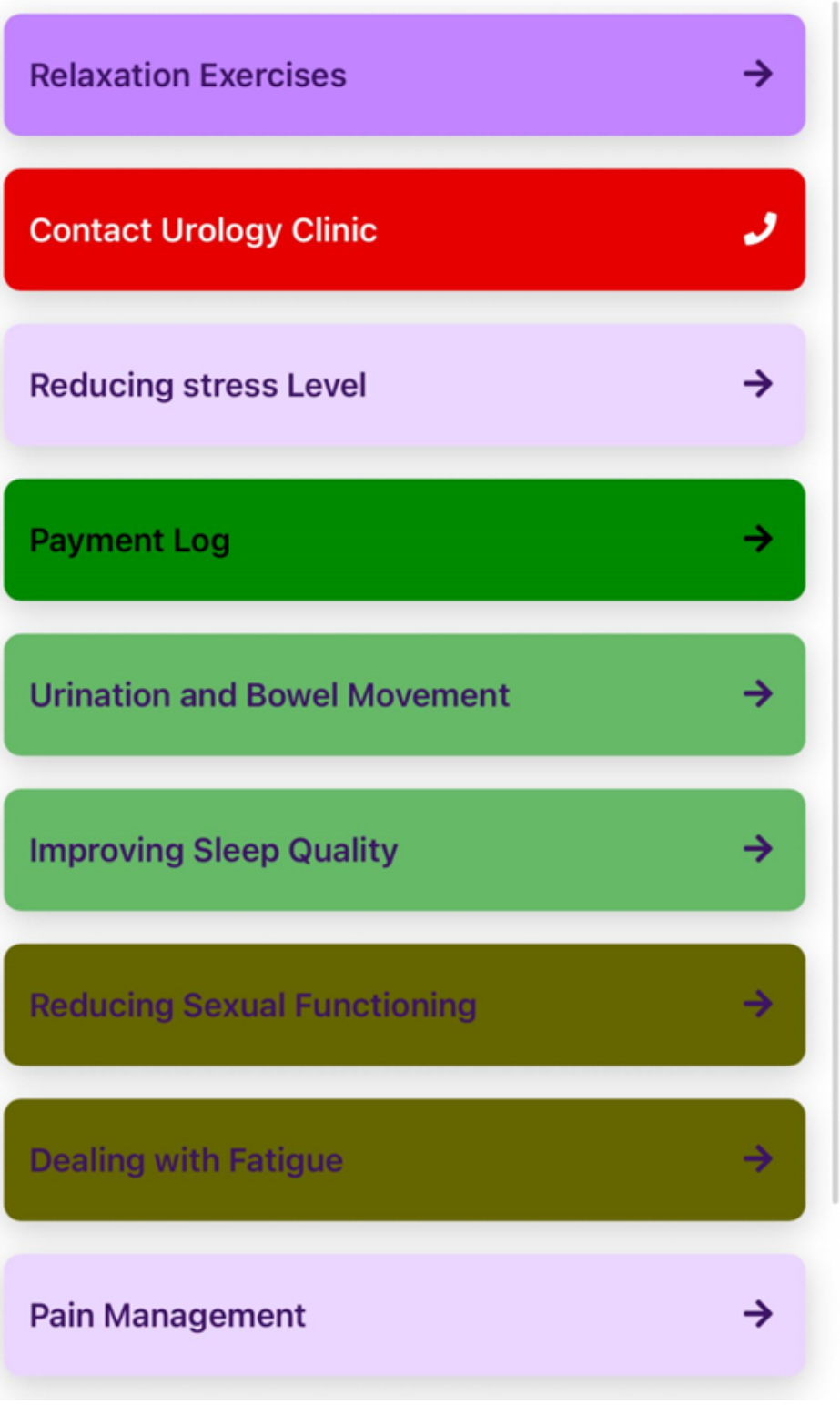

Conceptual Frameworks

The SAFE-CaPS study is guided by an adaptation of the Revised WCM of QoL [47,48], which was informed by findings from our extensive integrative review [8] and prior qualitative studies [9,29,68]. The Revised WCM model provides a structured approach for understanding the impact of biological function, symptom experience, functional status, and general health perceptions on QoL among Black men with CaP [48]. The model posits that biological function (eg, CaP-related pathophysiology) influences symptom status (eg, urinary dysfunction, fatigue, and depression), which in turn affects functional status (eg, ability to perform daily activities) [48,69]. These factors shape an individual's general health perception, ultimately determining their overall QoL, as measured by the Functional Assessment of Cancer Therapy-Prostate (FACT-P) questionnaire. The SAFE-CaPS intervention leverages this framework to enhance

XSL•FO
RenderX

symptom monitoring, self-management, and early intervention to improve functional status and QoL outcomes.

Demographic factors (eg, age, education, ethnicity, income, marital status, and geographic residence) and clinical variables (eg, age at diagnosis, treatment type, disease stage, prostate-specific antigen level, family history, overall health perception, and comorbidities) may moderate the impact of the intervention on QoL. The integration of the WCM framework within the SAFE-CaPS intervention is designed to account for individual differences, ensuring a comprehensive, culturally tailored approach to improve QoL among Black survivors with CaP (Figure 3).

Figure 3. Proposed model of treatment mechanism. PSA: prostate-specific antigen; QoL: quality of life; SAFE-CaPS: Survivorship App for Ethnically Diverse Black Prostate Cancer Survivors.

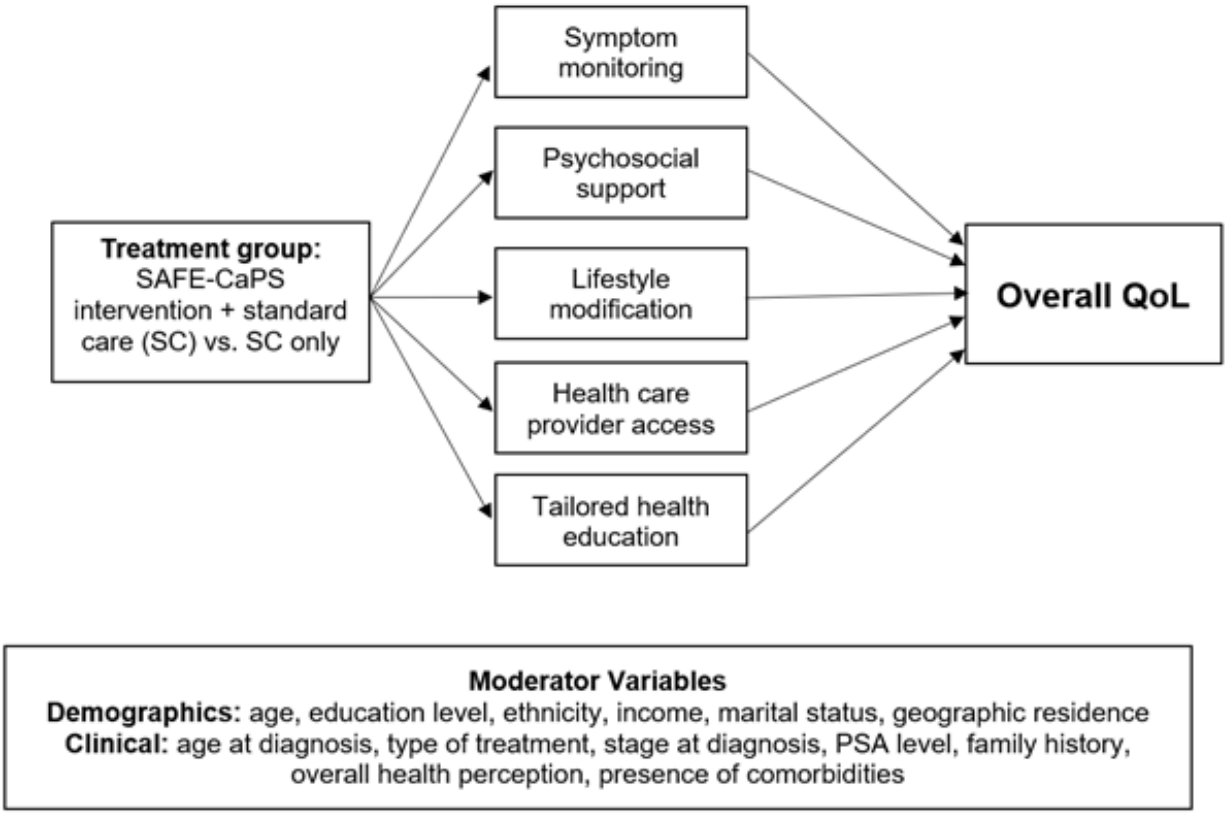

Additionally, the study incorporates the UTAUT framework to assess participant engagement and technology adoption [49]. The UTAUT model evaluates 4 key constructs that have been shown to influence behavioral intent and sustained app use (Figure 4) [49,70]. Performance expectancy refers to whether participants believe the intervention app will help them manage symptoms and improve their health behaviors. Effort expectancy captures perceptions of ease of use and overall accessibility. Social influence considers the extent to which health care providers, caregivers, and peers encourage engagement with the app. Facilitating conditions assess the availability of necessary resources and technical support to enable effective use of the intervention. Additionally, age serves as a moderating variable in the UTAUT model [49,70]. By integrating WCM for health outcomes and UTAUT for technology adoption, the study ensures a comprehensive evaluation of intervention effectiveness, allowing researchers to measure both clinical improvements in survivorship and behavioral factors affecting participant engagement. This dual-framework approach will inform future enhancements and scalability of the SAFE-CaPS intervention.

XSL•FO
RenderX

Figure 4. Proposed Unified Theory of Acceptance and Use of Technology model.

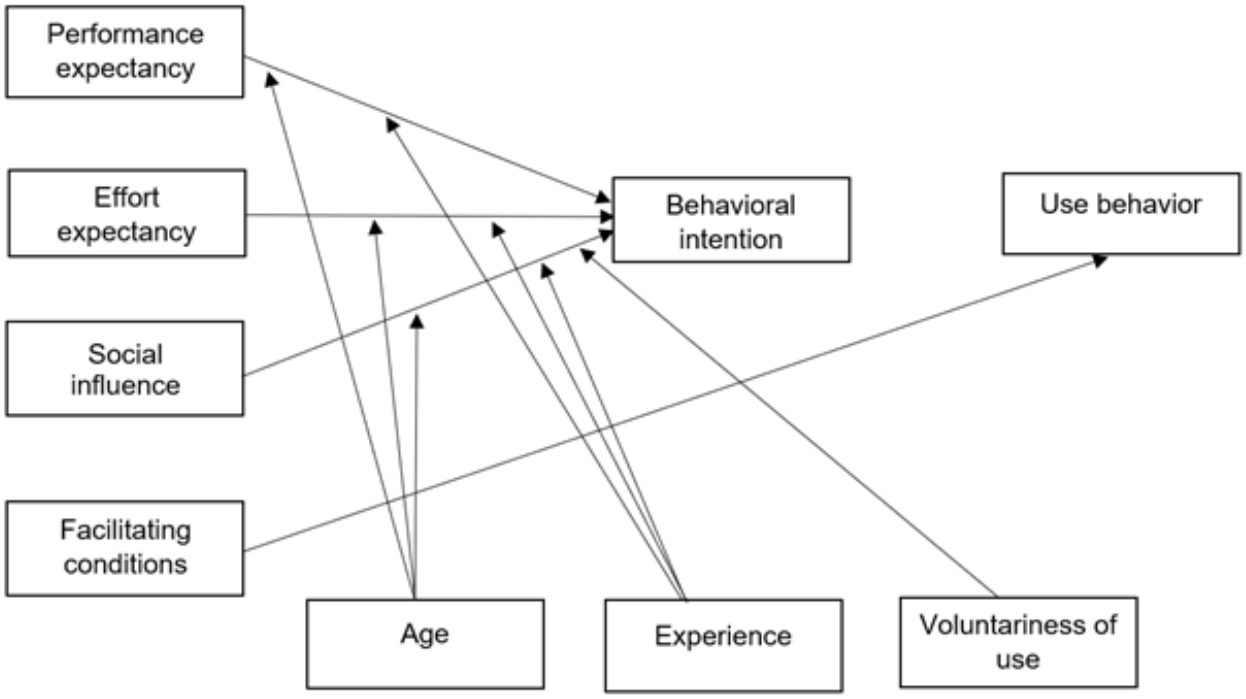

Intervention (SAFE-CaPS)

The SAFE-CaPS intervention is designed to provide real-time symptom monitoring and management of CaP-related symptoms while integrating culturally relevant content and psychosocial support. Acknowledging the importance of emotional well-being, the intervention includes resources for mental health support to improve coping with psychological challenges related to CaP survivorship.

The development of the SAFE-CaPS app was guided by a Community-Based Participatory Research approach, ensuring that the needs and lived experiences of Black men with CaP shaped its design and implementation [8,9,29,30]. The intervention is built upon findings from our prior work exploring QoL among ethnically diverse Black men with CaP [8,9,29]. This research identified key survivorship concerns, including symptom burden, psychosocial distress, and unmet informational needs, highlighting the necessity of a culturally responsive intervention. To increase relevance and usability, the app underwent iterative refinement in collaboration with a CAB composed of ethnically diverse Black men survivors with CaP [30]. Through focus groups, usability testing, and feedback from Black men experiencing CaP survivorship, the app was tailored to reflect their preferences, values, and health care experiences. Unique features include culturally relevant health education and an emphasis on peer support and interactions with trusted health care providers. Additionally, the intervention addresses historical health care mistrust by integrating educational modules that empower users to navigate the health care system and communicate effectively with providers.

A key feature of the SAFE-CaPS app is its incorporation of culturally tailored content, recognizing potential within-group differences among Black men from native-born, African-born, and Caribbean-born backgrounds. The intervention also acknowledges historical mistrust in health care and includes educational modules to enhance users' understanding of the health care system and encourage effective communication with providers. The intervention provides both dynamic and static resources, featuring content that adapts to user inputs and a comprehensive content library. The intervention also enables participants to log and share their symptoms with health care providers or designated proxies, supporting timely symptom reporting. Furthermore, the intervention uses behavioral engagement strategies, including push notifications and reminders, to encourage consistent participation. Using a longitudinal tracking approach, the SAFE-CaPS intervention captures detailed insights into the survivorship journey, including symptom burden, QoL, health behaviors (eg, physical activity, diet, and sleep), app usage patterns, and user-reported outcomes through weekly EMAs and quarterly follow-up surveys.

SAFE-CaPS includes daily assessments of key symptoms, including pain, sleep disturbances, fatigue, weight gain, urinary frequency and difficulty, bowel control issues, depression, anxiety, and worry. Participants also have the option to add additional comments in an open-text section to report any concerns or symptoms not captured in the standard items. Specific symptoms will trigger alerts based on frequency or distress levels, notifying both the study team and the designated health care providers selected by participants (Figure 5).

XSL•FO
RenderX

Figure 5. Example ecological momentary assessment item from risk assessment (left) and adaptive educational material (right).

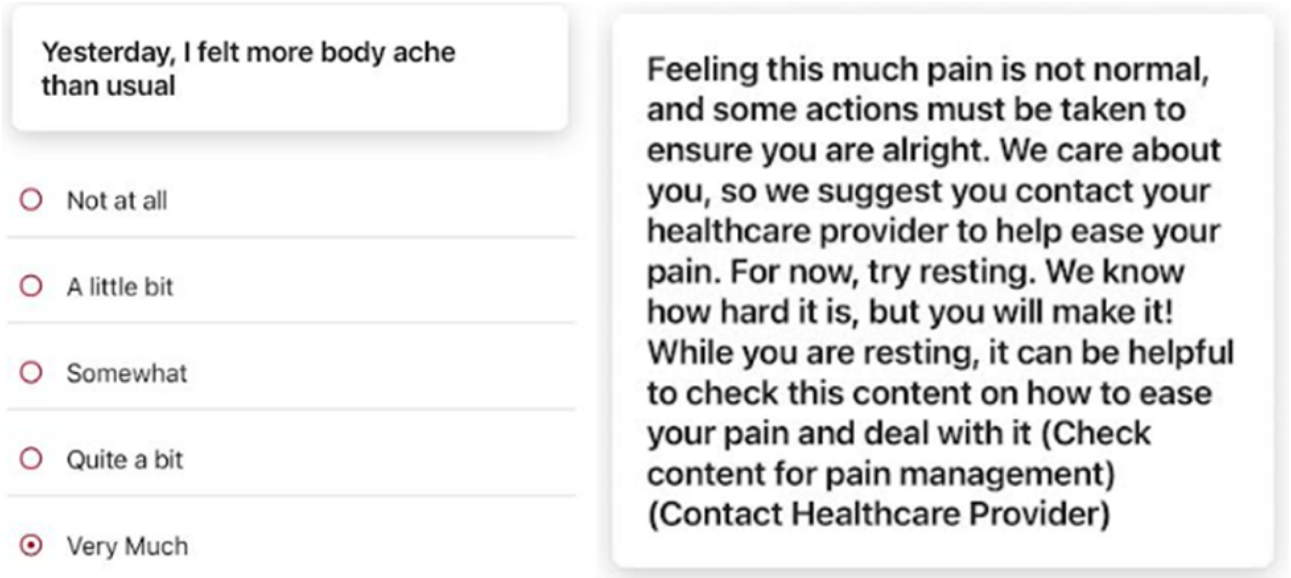

These alerts will prompt adaptive educational materials that provide self-management strategies until medical consultation is obtained (Figure 6). Specific alert triggers include urinary urgency, difficulty urinating, hematuria, constipation, blood in stool, pain, worry, or depression. Through these integrated components, SAFE-CaPS delivers a holistic and culturally responsive approach (developed with input from Black survivors with CaP and community advisors to ensure that language, imagery, and messaging reflect the lived experiences, preferences, and cultural values) to support Black men in navigating the complexities of CaP survivorship.

Figure 6. Samples of educational content.

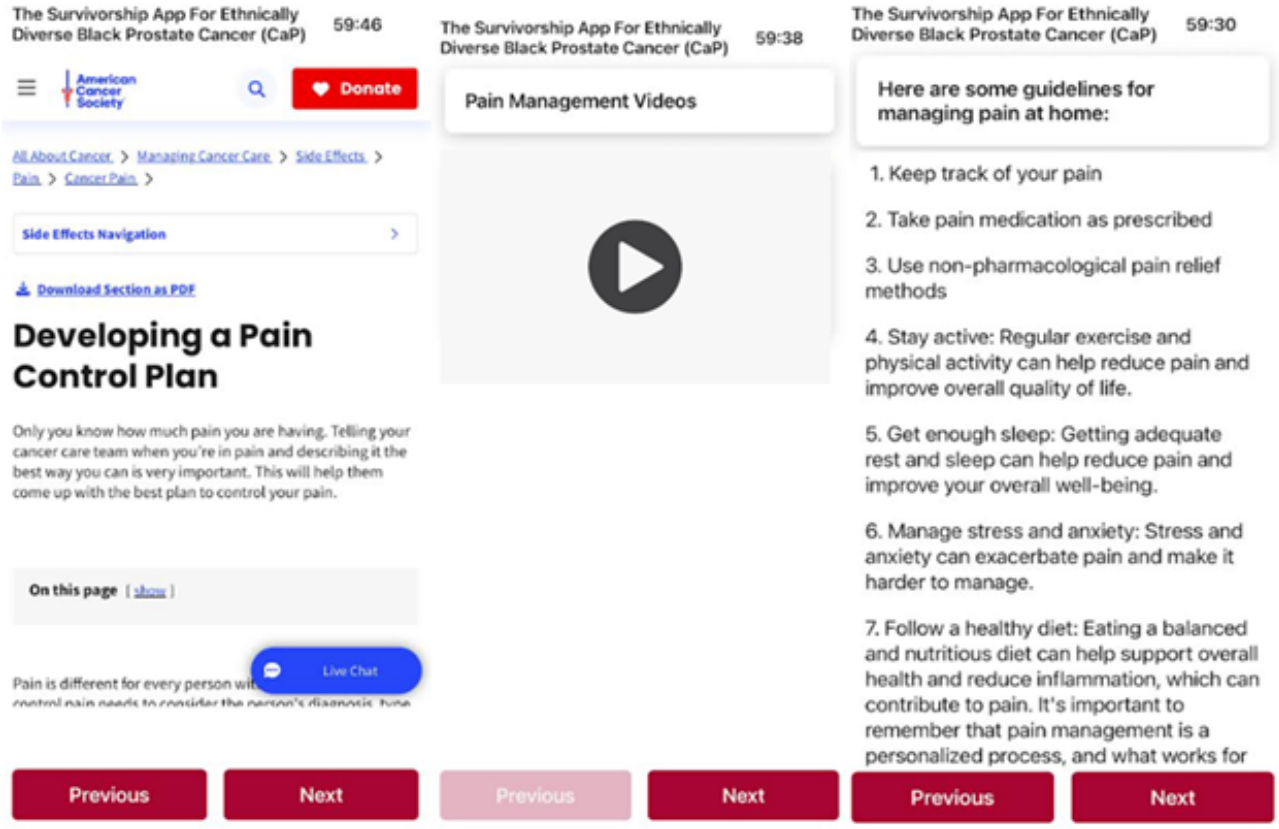

Control Group (SC)

Participants in the control group will not have access to the SAFE-CaPS intervention content. Still, they will receive standard survivorship care, including routine medical follow-ups and referrals as needed through their health care providers.

XSL•FO
RenderX

EMA

The SAFE-CaPS daily EMAs will include brief, real-time assessments of symptom status, functioning, and overall QoL. The goal is to minimize participant burden while capturing the essential data points necessary for the study's analyses.

Data Collection

Assessments will be conducted at baseline, post intervention, and interim time points at 3, 6, and 9 months. The assessment schedule is outlined in Table 1 (details of the assessment measures are provided in the Multimedia Appendix 3).

Table 1. Overview of the questionnaire assessments.

EMA[a] measures	Baseline (week 1)	Assessments (week 1-12)	Follow-up
Weekly[b] assessments (domains)			
Sociodemographic characteristics	✓		
Clinical characteristics	✓		
The positive and negative affect schedule	✓	✓	✓
Daily assessments (domains)			
Pain level and impact	✓	✓	
Sleep quality	✓	✓	
Sexual functioning	✓	✓	
Fatigue level	✓	✓	
Weight gain	✓	✓	
Urine frequency and difficulty	✓	✓	
Bowel control	✓	✓	
Physical activity frequency	✓	✓	
Dietary records	✓	✓	
Monthly[c] assessments			
PHQ-4[d]	✓		✓
FACT-P[e]	✓		✓

[a]EMA: ecological momentary assessment.

[b]For weekly assessments, the follow-up was done for months 6-12.

[c]For monthly assessments, the follow-up was done for months 3-12.

[d]PHQ-4: Patient Health Questionaire-4.

[e]FACT-P: Functional Assessment Cancer Therapy–Prostate.

One-Time Assessments

At baseline, sociodemographic and clinical characteristics will be collected, including age, education level, ethnicity, marital status, geographic residence, and clinical history (eg, age at diagnosis, stage of cancer, treatment type, family history of CaP, treatment status, prostate-specific antigen level, and comorbidities).

Weekly EMAs

Mood will be assessed at baseline and weekly for the first 12 weeks and during the week following each QoL assessment at 3, 6, 9, and 12 months, totaling 15 weeks of evaluation. The Positive and Negative Affect Schedule will be used to measure affect, consisting of two 10-item scales assessing positive and negative mood [71]. Participants will rate the intensity of several dimensions of their affect on a 5-point scale (1=very slightly or not at all; 5=very much), with higher scores indicating greater positive or negative affect [71].

Daily EMAs

The SAFE-CaPS group will have access to the app-based intervention for 12 weeks. During this time, participants will complete daily assessments of various symptoms, including pain intensity and impact, sleep quality, fatigue levels, weight changes, urinary frequency and difficulty, body aches, bowel control, sexual function, and depression or anxiety. Additional measures will assess physical activity frequency, duration, and dietary intake (fruit and vegetable consumption). Participants will also complete daily EMAs for 1 week following their 6, 9, and 12-month QoL assessments. Each daily evaluation is expected to take approximately 5–10 minutes to complete.

Quarterly Assessments

QoL and mental health outcomes, including depression and generalized anxiety disorder, will be evaluated at baseline and 3, 6, 9, and 12 months. These assessments are designed to take approximately 5 minutes to complete.

XSL•FO
RenderX

Post Intervention Survey

Twelve months post-randomization, participants assigned to SAFE-CaPS will complete a structured online survey via REDCap to assess intervention engagement, usability, and changes in health behavior, based on the UTAUT framework. The survey will evaluate participants' responses to the 4 key intervention constructs: performance expectancy, effort expectancy, social influence, and facilitating conditions, along with their final behavioral intent.

Exit Interview

A subset of 30 intervention participants will be invited to participate in semistructured interviews and focus groups conducted via Zoom to collect qualitative data on the usability, engagement, and perceived benefits of the SAFE-CaPS app. The interview guide will explore several key domains related to participants' experiences with the intervention. Specifically, the interviews will delve into their CaP survivorship journey, including interactions with health care providers, the long-term impact of treatment, and motivations for participating in the study. Questions will assess participants' affective attitude (ie, feelings about the intervention), perceived effectiveness (ie, perceived improvements in health and well-being), and self-efficacy (ie, confidence in carrying out intervention-related behaviors). The domain of burden will capture perceived effort, challenges, and time investment associated with the intervention. Intervention coherence will assess how well participants understand the purpose and functioning of the app. Additionally, participants will be invited to offer recommendations for improvement, describe their ideal model of survivorship support, and share whether they would recommend the app to others facing similar experiences.

Sample Size and Power Analysis

The sample size required to detect a significant change in the primary outcome (FACT-P Total Score) was estimated using a 2-sample *t* test. Although FACT-P will be assessed at baseline and at 3, 6, 9, and 12 months, the sample size calculation was based on detecting a statistically significant difference between groups at the 12-month timepoint, which serves as the primary endpoint. The calculation assumed a significance level (α) of 0.05, 80% power, and a 1:1 randomization ratio. Based on prior research by Cella et al [72], which reported an effect size between 0.39 and 0.58, the sample size was conservatively determined using the smaller effect size (0.39). Under these assumptions, a minimum of 105 participants per group is required to detect a meaningful difference in FACT-P scores.

We anticipate that 15% of the full sample will be lost to attrition; therefore, the final target sample size was set at 124 participants per group (N=248 total). A participant sampling plan will be implemented to ensure adequate representation of ethnic groups and balanced allocation across intervention and control conditions, in alignment with the randomized controlled trial design.

Measures

This study will use a range of validated instruments and self-reported measures to assess key constructs related to QoL, symptom burden, psychosocial well-being, lifestyle behaviors, and health care access among Black men with CaP.

Primary Outcome

The primary outcome will be QoL, which will be assessed using the (FACT-P) questionnaire [72] at baseline, 3, 6, 9, and 12 months. FACT-P is a 39-item instrument that evaluates physical, emotional, and functional well-being, as well as CaP-specific concerns such as weight loss, appetite, pain, urinary symptoms, bowel function, and sexual health. Scores will be transformed onto a 0-100 scale, with higher scores indicating better QoL.

Secondary Outcomes

App engagement and usability will be evaluated based on frequency and duration of use, symptom-tracking interaction, and engagement with educational content. Technology acceptance and behavioral intent will also be assessed using the UTAUT framework, which evaluates performance expectancy, effort expectancy, social influence, and facilitating conditions. These outcome variables will be analyzed at multiple time points to determine the effectiveness of the SAFE-CaPS intervention in enhancing QoL, improving mental health, and supporting symptom self-management among Black survivors with CaP. Additionally, semistructured interviews and focus groups will be conducted to explore app usability, engagement, and perceived benefits.

Analytic Plan

The statistical analyses for the SAFE-CaPS study will include descriptive, bivariate, and multivariate methods to assess the intervention's impact on QoL, mental health, symptom management, and app engagement. Each variable will be examined for distribution, range, mode, median, and mean (SD). Skewness and kurtosis will be determined for all ratio-level variables, while frequency distributions will be calculated for nominal- and ordinal-level variables. Descriptive statistics will be used to summarize all study variables, and bivariate statistics will be computed to assess relationships among theoretical constructs.

Differences in demographic and clinical characteristics between the 2 groups will be analyzed using a 2-tailed independent *t* test, Fisher exact test, and the chi-square test. Intervention usage will be analyzed using descriptive and inferential statistics for the number of EMAs completed, alerts triggered, and views of the educational section. An adjusted linear regression mixed-effect model will assess the impact of treatment group assignment on FACT-P scores at 12 months, adjusting for relevant covariates such as age, socioeconomic status, treatment type, ethnicity, and cancer stage. Preliminary analyses will include checks for multicollinearity and score distributions to ensure model validity. The goal is to determine the effect of the intervention on QoL as measured by FACT-P, independent of these baseline characteristics.

Chi-square analyses and *t* tests will examine baseline differences between ethnic groups. One-way ANOVA and correlations will be conducted to determine whether demographic and clinical variables are significantly associated with perceived acceptability, accessibility, and engagement. Due to the

XSL•FO
RenderX

preliminary nature of these analyses, only covariates significant at $P<.05$ will be included in the final models. Cronbach α will assess the reliability of the multi-item instrument to determine the internal consistency of the model used to measure the constructs of the proposed framework, as shown in Figure 3. A Cronbach α value ≥0.70 will be considered acceptable for internal consistency [73]. Data will be analyzed using SPSS (version 29; IBM Corp).

Qualitative Analysis Plan

Qualitative data from postintervention interviews will be analyzed using a reflexive thematic analysis approach to explore participants' experiences with the SAFE-CaPS mHealth app. All interviews will be transcribed verbatim by a professional transcription service with established safeguards to protect Protected Health Information, and transcripts will be coded using ATLAS.ti v24 (Lumivero/ATLAS.ti Scientific Software Development GmbH). Two researchers (MO and GK) will independently review transcripts, develop initial codes, and iteratively refine a shared codebook through discussion and consensus. Codes will then be organized into broader categories and themes that capture perceptions of usability, acceptability, cultural relevance, and factors influencing engagement with the intervention, with attention to variation across ethnic subgroups. Emerging themes will be reviewed collaboratively with the research team to ensure analytic rigor. Qualitative findings will be integrated with quantitative outcomes using an explanatory sequential mixed-methods approach, allowing qualitative insights to contextualize, explain, and complement observed quantitative patterns to provide a comprehensive understanding of intervention usability and impact.

Study Compensation

Participants (N=248) will receive US $50 each for completing both the baseline and postintervention assessments, totaling US $100 per participant. In addition, all participants will receive US $25 for completing each interim survey at 3, 6, and 9 months (a total of US $75). Participants in the SAFE-CaPS arm who complete the semistructured exit interview (n=30) will receive an additional US $50 payment. To enhance EMA compliance, SAFE-CaPS participants (n=124) will receive US $2 per day for completing daily assessments once a day over 15 weeks (up to US $210) and US $20 per completed weekly assessment (up to US $300). Thus, participants in this arm may earn up to US $685, or US $735 if they also complete the exit interview. Participants in the control arm may earn up to US $175 in total. Compensation will be provided weekly via a reloadable study credit card [74]. Participants will not be compensated for on-demand app usage or engagement with intervention components. Although the total potential payment varies by arm, the additional amounts in SAFE-CaPS accurately reflect the increased time and effort required for intensive EMA data collection. This structure was reviewed and approved by the University of Oklahoma Health Campus Institutional Review Board (IRB). To guard against bias, total compensation received will be captured and evaluated as a covariate in sensitivity analyses.

Ethical Considerations

The SAFE-CaPS study will adhere to strict ethical guidelines to protect participants' rights, privacy, and safety. The study protocol has been approved by the University of Oklahoma Health Campus Institutional Review Board (OUHC IRB; protocol number 17116). This trial is registered at ClinicalTrials.gov (NCT06651359). All participants will provide written informed consent after being fully informed of the study purpose, procedures, risks, benefits, and voluntary nature of participation, including the right to withdraw at any time without penalty.

To protect privacy and confidentiality, all participant data will be deidentified, encrypted, and stored on secure, password-protected computers and servers. Audio-recorded interviews will be professionally transcribed, and all transcripts will be deidentified. Only authorized study personnel will have access to identifiable data. Additional protections, including user access controls, password authentication, auto-logout features, and audit trails, will be implemented to prevent unauthorized access to Protected Health Information. Participants who experience distress or report clinically concerning symptoms will be referred to appropriate health care providers, and the app's automated alert system will notify the study team when severe symptoms are reported to facilitate timely follow-up. Ethical oversight will be maintained through ongoing IRB monitoring and adherence to all federal and institutional regulations governing human subjects' research. Participants will receive monetary compensation for completing study assessments across the 12-month period, with payments distributed at each assessment point. Participants in the SAFE-CaPS arm may earn up to US $685, or up to US $735 if they complete the exit interview, whereas participants in the control arm may earn up to US $175 in total.

Results

The SAFE-CaPS study received funding in January 2024, and IRB approval was received in May 2024. Study recruitment began on August 24, 2025. Overall, 6 participants have been enrolled, including 2 in the SAFE-CaPS intervention arm and 4 in the SC group. Recruitment is expected to conclude in July 2027. Data analysis is anticipated to occur toward the end of 2027, with article preparation and publication planned for mid-2028 and early 2029. The SAFE-CaPS study is expected to improve QoL, mental health, and symptom management among ethnically diverse Black men with CaP. It is anticipated that participants in the intervention group will demonstrate significant improvements in FACT-P scores, along with reductions in depression and anxiety symptoms over time. Real-time symptom monitoring is expected to enhance self-management and facilitate timely intervention, resulting in fewer severe symptom episodes and improved health outcomes. High app engagement is anticipated, with frequent use of symptom tracking, educational content, and provider alerts. Increased engagement may be associated with better health behaviors, including increased physical activity and improved dietary intake. Exploratory analyses may reveal differences in app adoption based on sociodemographic factors, enabling the

XSL•FO
RenderX

identification of potential disparities in digital health use. Furthermore, qualitative findings are expected to provide insights into facilitators and barriers to app usability, informing future refinements of the intervention and guiding the scalability and adaptation of the SAFE-CaPS model for broader implementation.

Discussion

Principal Findings

The SAFE-CaPS study is expected to demonstrate that participants receiving the culturally tailored mHealth intervention will experience greater improvements in QoL, reductions in symptom burden, and higher engagement in survivorship care than those receiving SC. We anticipate that real-time symptom monitoring, adaptive educational support, and culturally grounded content will enhance symptom self-management, reduce distress, and facilitate more timely health care follow-up. Qualitative findings are also expected to illuminate user experiences, highlight facilitators and barriers to engagement, and guide iterative refinement of the app.

Comparison With Prior Work

Prior research has consistently shown that Black men with CaP experience a higher symptom burden, lower engagement in survivorship care, reduced QoL, and significant disparities in health outcomes compared to other racial groups [3,8,39,75]. Digital health solutions have demonstrated promise in improving symptom management, mental health, and survivorship outcomes across various cancer populations [76]. However, existing CaP mHealth interventions have rarely been designed specifically for Black men, nor have they accounted for within-group cultural diversity among native-born, African-born, and Caribbean-born survivors with CaP. SAFE-CaPS advances this literature by integrating culturally relevant educational resources and survivorship priorities identified directly from diverse Black survivors with CaP and community advisors, addressing a critical gap in prior mHealth and survivorship studies.

Strengths and Limitations

The SAFE-CaPS study has several key strengths that distinguish it as a pioneering mHealth intervention for Black survivors with CaP. First, its culturally tailored design enables the intervention content, symptom tracking, and psychosocial support features to directly address unique survivorship challenges faced by Black men, particularly considering historical health care mistrust and disparities in symptom management. Second, the app includes real-time symptom monitoring with a tiered alert system: when a participant reports a symptom that exceeds a prespecified frequency or distress threshold, an alert is routed to the research team, who will contact the participant within 24 hours and, with the participant's permission, notify the designated health care provider for clinical follow-up. Third, the study uses a comprehensive evaluation framework that integrates quantitative assessments (FACT-P and Patient Health Questionnaire-4 scales) with qualitative data from semistructured interviews to assess the effectiveness, feasibility, engagement, and ongoing refinement of the app. Additionally, the use of validated theoretical frameworks (WCM and UTAUT) strengthens the study's conceptual foundation, linking clinical outcomes, psychosocial factors, and technology adoption to enhance survivorship care.

Despite its strengths, the study has several limitations. One potential challenge is variability in intervention engagement, which may be influenced by factors such as digital literacy, socioeconomic status, and access to technology, potentially limiting the generalizability of the findings. To address this, participants will receive technical support and training to enhance the usability and engagement of the intervention. A further limitation is the potential for differential retention between study arms, as participants in the SAFE-CaPS group may have greater motivation to remain in the study due to access to the intervention and associated compensation, which could contribute to higher loss to follow-up in the usual care group. Retention will be monitored across study arms, and results will be interpreted with consideration of potential differential attrition. Another limitation is reliance on self-reported data for QoL and symptom tracking, which may introduce recall and response bias. However, the use of validated assessment instruments and objective app engagement metrics will enhance data reliability and mitigate some concerns. Despite these limitations, demonstrating the preliminary efficacy of the SAFE-CaPS app will provide valuable insights into the feasibility and effectiveness of culturally tailored digital interventions for improving CaP survivorship outcomes in Black men.

Future Directions

Findings from this study will guide the refinement of SAFE-CaPS, including enhancing usability features, tailoring educational content, and improving the symptom alert process. Future work may involve adapting the intervention for broader geographic implementation, testing the app in a larger multisite randomized clinical trial, and exploring integration with electronic health records for automated clinical alerts. The study team plans to disseminate findings through peer-reviewed publications, national conferences, community presentations, and partnerships with cancer centers and survivorship networks. Ultimately, SAFE-CaPS has the potential to inform scalable, equitable digital survivorship models for diverse and underserved cancer survivor populations.

Conclusions

The SAFE-CaPS study represents a significant effort to address persistent survivorship disparities among Black prostate cancer survivors through a culturally tailored, mHealth–enabled approach. By combining real-time symptom surveillance, tailored education, and community-informed content, this intervention aims to provide a scalable and accessible tool to support improved survivorship outcomes. The results of this protocol-driven trial will contribute critical evidence to the fields of digital oncology, health equity, and mHealth implementation, and will inform future policy, clinical integration, and larger-scale trials aimed at expanding culturally responsive survivorship care.

XSL•FO
RenderX

Acknowledgments

The authors thank the Mobile Health Shared Resource at the OU Health Stephenson Cancer Center for providing programming and technological support for the development and deployment of the SAFE-CaPS intervention. We also acknowledge the contributions of the research team, community partners, and participating survivors with prostate cancer, whose insights helped guide the cultural tailoring of the SAFE-CaPS intervention.

During the preparation of this manuscript, the lead author used Paperpal Preflight to revise the clarity of a few sentences. After using this tool, the lead author reviewed and edited all content and takes full responsibility for the integrity and accuracy of the final publication.

Funding

This research was supported by an Early Investigator Research Award from the Department of Defense (EIRAW81XWH-20-1-0080), a Health Disparity Research Award from the Department of Defense (HT94252410290), and a National Institute on Minority Health and Health Disparities grant (R25MD011564), all awarded to MO. This work was also supported by the National Institute on Minority Health and Health Disparities of the National Institutes of Health under Award Number K01MD015295 awarded to AA. Additional funding support was provided by the American Cancer Society through grants PF-24-1198256-01-HOPS to NMM and CHERC-24-1160021-01-CHERC to DEK/GK. This resource is supported in part by the National Cancer Institute Cancer Center Support Grant (P30CA225520). Additional support for infrastructure and shared resources was provided by the Oklahoma Tobacco Settlement Endowment Trust (TSET) contract STCST00400_FY25), awarded to the OU Health Stephenson Cancer Center and used to access the Mobile Health Technology Shared Resource.

Data Availability

The datasets generated or analyzed during this study will be available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: MO
Data curation: MO, GK, and PG
Formal analysis: MO, GK, and PG
Funding acquisition: MO
Investigation: GK, PG, ACA, KD, JMN, PC, PB, DR, EM, RO, MAH, YDZ, ZN, AGM, EK, MEY, SD, DJM, OB, FTO, LGP, MSB, DEK, MEO
Methodology: MO
Project administration: PG and GK
Software: MB
Supervision: MO
Validation: MO, GK, and PG
Visualization: MO, GK, and PG
Writing-original draft: GK
Writing-review and editing: GK, PG, ACA, KD, JMN, PC, PB, DR, EM, RO, MAH, YDZ, ZN, AGM, EK, MEY, SD, DJM, OB, FTO, LGP, MSB, DEK, and MEO

Conflicts of Interest

MSB and DEK are inventors of the Insight mHealth Platform, and they receive royalties when investigators outside OUHC use Insight. However, since the study PI is affiliated with OUHC, they will not receive royalties in this case.

Multimedia Appendix 1
Department of Defense peer-review summary report.
[PDF File (Adobe PDF File), 234 KB - resprot_v15i1e81503_app1.pdf]

Multimedia Appendix 2
SPIRIT 2025 checklist.
[DOCX File , 34 KB - resprot_v15i1e81503_app2.docx]

Multimedia Appendix 3

XSL•FO
RenderX

Overview of scheduled assessment.
[DOCX File , 20 KB - resprot_v15i1e81503_app3.docx]

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin 2025;75(1):10-45 [FREE Full text] [doi: 10.3322/caac.21871] [Medline: 39817679]
2. Saka AH, Giaquinto AN, McCullough LE, Tossas KY, Star J, Jemal A, et al. Cancer statistics for African American and Black people, 2025. CA Cancer J Clin 2025;75(2):111-140 [FREE Full text] [doi: 10.3322/caac.21874] [Medline: 39976243]
3. Lillard JW, Moses KA, Mahal BA, George DJ. Racial disparities in Black men with prostate cancer: a literature review. Cancer 2022;128(21):3787-3795 [FREE Full text] [doi: 10.1002/cncr.34433] [Medline: 36066378]
4. Murphy A, Cottrell-Daniels CC, Awasthi S, Katende E, Park JY, Denis J, et al. Understanding and addressing prostate cancer disparities in diagnosis, treatment, and outcomes among Black men. Cancer Control 2024;31:10732748241275389 [FREE Full text] [doi: 10.1177/10732748241275389] [Medline: 39149902]
5. Chowdhury-Paulino IM, Ericsson C, Vince R, Spratt DE, George DJ, Mucci LA. Racial disparities in prostate cancer among black men: epidemiology and outcomes. Prostate Cancer Prostatic Dis 2022;25(3):397-402 [FREE Full text] [doi: 10.1038/s41391-021-00451-z] [Medline: 34475523]
6. Zhang B, Li J, Tang M, Cheng C. Reduced racial disparity as a result of survival improvement in prostate cancer. Cancers (Basel) 2023;15(15):3977 [FREE Full text] [doi: 10.3390/cancers15153977] [Medline: 37568792]
7. Finney JM, Hamilton JB, Hodges EA, Pierre-Louis BJ, Crandell JL, Muss HB. African American cancer survivors: do cultural factors influence symptom distress? J Transcult Nurs 2015;26(3):294-300. [doi: 10.1177/1043659614524251] [Medline: 24797252]
8. Dickey SL, Ogunsanya ME. Quality of life among black prostate cancer survivors: an integrative review. Am J Mens Health 2018;12(5):1648-1664 [FREE Full text] [doi: 10.1177/1557988318780857] [Medline: 29926761]
9. Ogunsanya ME, Kaninjing E, Ellis TN, Morton DJ, McIntosh AG, Zhao J, et al. Health-related quality of life in ethnically diverse Black prostate cancer survivors: a convergent parallel mixed-methods approach. Qual Life Res 2024;33(10):2705-2719. [doi: 10.1007/s11136-024-03718-9] [Medline: 39008143]
10. Baden M, Lu L, Drummond FJ, Gavin A, Sharp L. Pain, fatigue and depression symptom cluster in survivors of prostate cancer. Support Care Cancer 2020;28(10):4813-4824 [FREE Full text] [doi: 10.1007/s00520-019-05268-0] [Medline: 31980895]
11. Nguyen C, Lairson DR, Swartz MD, Du XL. Risks of major long-term side effects associated with androgen-deprivation therapy in men with prostate cancer. Pharmacotherapy 2018;38(10):999-1009. [doi: 10.1002/phar.2168] [Medline: 30080934]
12. Sutton E, Lane JA, Davis M, Walsh EI, Neal DE, Hamdy FC, et al. Men's experiences of radiotherapy treatment for localized prostate cancer and its long-term treatment side effects: a longitudinal qualitative study. Cancer Causes Control 2021;32(3):261-269 [FREE Full text] [doi: 10.1007/s10552-020-01380-3] [Medline: 33394204]
13. Davis KM, Kelly SP, Luta G, Tomko C, Miller AB, Taylor KL. The association of long-term treatment-related side effects with cancer-specific and general quality of life among prostate cancer survivors. Urology 2014;84(2):300-306 [FREE Full text] [doi: 10.1016/j.urology.2014.04.036] [Medline: 24975711]
14. Unger JM, Till C, Tangen CM, Hershman DL, Goodman PJ, LeBlanc M, et al. Long-term adverse effects and complications after prostate cancer treatment. JAMA Oncol 2024;10(12):1654-1662. [doi: 10.1001/jamaoncol.2024.4397] [Medline: 39509091]
15. Spendelow JS, Eli Joubert H, Lee H, Fairhurst BR. Coping and adjustment in men with prostate cancer: a systematic review of qualitative studies. J Cancer Surviv 2018;12(2):155-168 [FREE Full text] [doi: 10.1007/s11764-017-0654-8] [Medline: 29063497]
16. Ogunsanya ME, Kaninjing E, Ellis TN, Morton DJ, McIntosh AG, Zhao J, et al. Health-related quality of life in ethnically diverse Black prostate cancer survivors: a convergent parallel mixed-methods approach. Qual Life Res 2024;33(10):2705-2719. [doi: 10.1007/s11136-024-03718-9] [Medline: 39008143]
17. Okoro FO, Song L, Auten B, Whitaker-Brown C, Cornelius J. African-American survivors of prostate cancer: a meta-synthesis of qualitative studies. J Cancer Surviv 2021;15(1):40-53 [FREE Full text] [doi: 10.1007/s11764-020-00909-4] [Medline: 32666419]
18. Rencsok EM, Slopen N, Autio K, Morgans A, McSwain L, Barata P, et al. Quality of life in the year after new diagnosis with advanced prostate cancer for Black and White individuals living in the US. Qual Life Res 2023 Nov;32(11):3209-3221 [FREE Full text] [doi: 10.1007/s11136-023-03468-0] [Medline: 37410340]
19. Rice K, Hudak J, Peay K, Elsamanoudi S, Travis J, Lockhart R, et al. Comprehensive quality-of-life outcomes in the setting of a multidisciplinary, equal access prostate cancer clinic. Urology 2010;76(5):1231-1238. [doi: 10.1016/j.urology.2010.03.087] [Medline: 20579698]
20. Chhatre S, Wein AJ, Malkowicz SB, Jayadevappa R. Racial differences in well-being and cancer concerns in prostate cancer patients. J Cancer Surviv 2011;5(2):182-190. [doi: 10.1007/s11764-011-0170-1] [Medline: 21274650]

XSL•FO
RenderX

21. Imm KR, Williams F, Housten AJ, Colditz GA, Drake BF, Gilbert KL, et al. African American prostate cancer survivorship: exploring the role of social support in quality of life after radical prostatectomy. J Psychosoc Oncol 2017;35(4):409-423 [FREE Full text] [doi: 10.1080/07347332.2017.1294641] [Medline: 28398149]
22. Koelker M, Labban M, Frego N, Ye J, Lipsitz SR, Hubbell HT, et al. Racial differences in patient-reported outcomes among men treated with radical prostatectomy for prostate cancer. Prostate 2024;84(1):47-55. [doi: 10.1002/pros.24624] [Medline: 37710385]
23. Facts about the U.S. black population. Pew Research Center. 2025. URL: https://www.pewresearch.org/race-and-ethnicity/fact-sheet/facts-about-the-us-black-population/ [accessed 2025-07-18]
24. A rising share of the U.S. black population is foreign born. Pew Research Center. 2015. URL: https://www.pewresearch.org/social-trends/2015/04/09/a-rising-share-of-the-u-s-black-population-is-foreign-born/ [accessed 2025-07-18]
25. Pinheiro PS, Medina H, Callahan KE, Kwon D, Ragin C, Sherman R, et al. Cancer mortality among US blacks: variability between African Americans, Afro-Caribbeans, and Africans. Cancer Epidemiol 2020;66:101709 [FREE Full text] [doi: 10.1016/j.canep.2020.101709] [Medline: 32240959]
26. Ragin C, Mutetwa B, Attong-Rogers A, Roach V, Taioli E. Geographic and outcome variation among black men diagnosed with prostate cancer. Infect Agent Cancer 2011;6 Suppl 2(Suppl 2):S2 [FREE Full text] [doi: 10.1186/1750-9378-6-S2-S2] [Medline: 21992682]
27. Zeigler-Johnson C, McDonald AC, Pinheiro P, Lynch S, Taioli E, Joshi S, et al. Trends in prostate cancer incidence among black men in the Caribbean and the United States. Prostate 2023;83(12):1207-1216. [doi: 10.1002/pros.24580] [Medline: 37244749]
28. Bright M, Grant A, Zhu J, Choudhri A, Dash C, Adams-Campbell L. Preventative cancer screening among African and Afro-Caribbean immigrants. J Natl Med Assoc 2025 [FREE Full text] [doi: 10.1016/j.jnma.2025.11.006] [Medline: 41290476]
29. Ogunsanya M, Kaninjing E, Ellis T, Bamidele O, Morton D, McIntosh A, et al. Quality of life assessment among ethnically diverse black prostate cancer survivors: a constructivist grounded theory approach. Res Sq 2024 [FREE Full text] [doi: 10.21203/rs.3.rs-3941497/v1] [Medline: 38464107]
30. Ogunsanya ME, Kaninjing E, Morton DJ, Dwyer K, Young ME, Odedina FT. Bridging the Gap: a community advisory board promoting community engagement in cancer research for ethnically diverse populations. Am J Mens Health 2024;18(5):15579883241280826 [FREE Full text] [doi: 10.1177/15579883241280826] [Medline: 39340388]
31. Chornokur G, Dalton K, Borysova ME, Kumar NB. Disparities at presentation, diagnosis, treatment, and survival in African American men, affected by prostate cancer. Prostate 2011;71(9):985-997 [FREE Full text] [doi: 10.1002/pros.21314] [Medline: 21541975]
32. Sinfield P, Baker R, Camosso-Stefinovic J, Colman AM, Tarrant C, Mellon JK, et al. Men's and carers' experiences of care for prostate cancer: a narrative literature review. Health Expect 2009;12(3):301-312 [FREE Full text] [doi: 10.1111/j.1369-7625.2009.00546.x] [Medline: 19754693]
33. Burnett AL. Racial disparities in sexual dysfunction outcomes after prostate cancer treatment: myth or reality? J Racial Ethn Health Disparities 2016;3(1):154-159. [doi: 10.1007/s40615-015-0126-7] [Medline: 26896115]
34. Alexis O, Worsley AJ. The impact of ethnicity on decisions and decision making in prostate cancer: an integrative review. Ethn Health 2025;30(2):316-336 [FREE Full text] [doi: 10.1080/13557858.2024.2429416] [Medline: 39560506]
35. Stubbins R, He T, Yu X, Puppala M, Ezeana CF, Chen S, et al. A behavior-modification, clinical-grade mobile application to improve breast cancer survivors' accountability and health outcomes. JCO Clin Cancer Inform 2018;2:1-11 [FREE Full text] [doi: 10.1200/CCI.18.00054] [Medline: 30652617]
36. Howell KE, Baedke JL, Bagherzadeh F, McDonald A, Nathan PC, Ness KK, et al. Using mHealth technology to evaluate daily symptom burden among adult survivors of childhood cancer: a feasibility study. Cancers (Basel) 2024;16(17):2984 [FREE Full text] [doi: 10.3390/cancers16172984] [Medline: 39272842]
37. Hernandez Silva E, Lawler S, Langbecker D. The effectiveness of mHealth for self-management in improving pain, psychological distress, fatigue, and sleep in cancer survivors: a systematic review. J Cancer Surviv 2019;13(1):97-107. [doi: 10.1007/s11764-018-0730-8] [Medline: 30635865]
38. Seiler A, Klaas V, Tröster G, Fagundes CP. eHealth and mHealth interventions in the treatment of fatigued cancer survivors: a systematic review and meta-analysis. Psychooncology 2017;26(9):1239-1253. [doi: 10.1002/pon.4489] [Medline: 28665554]
39. Ogunsanya ME, Sifat M, Bamidele OO, Ezenwankwo EF, Clifton S, Ton C, et al. Mobile health (mHealth) interventions in prostate cancer survivorship: a scoping review. J Cancer Surviv 2023;17(3):557-568. [doi: 10.1007/s11764-022-01328-3] [Medline: 36627464]
40. Mobile fact sheet. Pew Research Center. 2024. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2025-01-09]
41. Smartphones help blacks, hispanics bridge some?but not all?digital gaps with whites. Benton Institute for Broadband & Society. 2017. URL: https://www.benton.org/headlines/smartphones-help-blacks-hispanics-bridge-some-%E2%80%93-not-all-%E2%80%93-digital-gaps-whites [accessed 2025-07-18]

XSL•FO
RenderX

42. Home broadband adoption, computer ownership vary by race, ethnicity in the U.S. Pew Research Center. 2021. URL: https://www.pewresearch.org/short-reads/2021/07/16/home-broadband-adoption-computer-ownership-vary-by-race-ethnicity-in-the-u-s/ [accessed 2025-07-18]
43. Mahmood A, Mahmood A, Kedia S, Chang CF. Rural-urban disparities in mobile health application ownership and utilization among cancer survivors. Med Care 2025;63(2):111-116. [doi: 10.1097/MLR.0000000000002092] [Medline: 39791845]
44. Rincon E, Monteiro-Guerra F, Rivera-Romero O, Dorronzoro-Zubiete E, Sanchez-Bocanegra CL, Gabarron E. Mobile phone apps for quality of life and well-being assessment in breast and prostate cancer patients: systematic review. JMIR Mhealth Uhealth 2017;5(12):e187 [FREE Full text] [doi: 10.2196/mhealth.8741] [Medline: 29203459]
45. Yuan MJ, Hébert ET, Johnson RK, Long J, Vandewater EA, Vickers AJ. A personalized automated messaging system to improve adherence to prostate cancer screening: research protocol. JMIR Res Protoc 2012 Nov 28;1(2):e20 [FREE Full text] [doi: 10.2196/resprot.2398] [Medline: 23612443]
46. Pham Q, Cafazzo JA, Feifer A. Adoption, acceptability, and effectiveness of a mobile health app for personalized prostate cancer survivorship care: protocol for a realist case study of the ned app. JMIR Res Protoc 2017;6(10):e197 [FREE Full text] [doi: 10.2196/resprot.8051] [Medline: 29025699]
47. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. JAMA 1995 Jan 04;273(1):59-65. [Medline: 7996652]
48. Ferrans CE, Zerwic JJ, Wilbur JE, Larson JL. Conceptual model of health-related quality of life. J Nurs Scholarsh 2005;37(4):336-342. [doi: 10.1111/j.1547-5069.2005.00058.x] [Medline: 16396406]
49. Venkatesh V. User acceptance of information technology: Toward a unified view. MIS quarterly 2003:425-478. [doi: 10.2307/30036540]
50. Marikyan M, Papagiannidis P. Unified Theory of Acceptance and Use of Technology. Bristol, UK: TheoryHub book; 2021.
51. Chan A, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. The Lancet 2025;405(10491):e19-e27. [doi: 10.1016/s0140-6736(25)00770-6]
52. Halpern MT, McCabe MS, Burg MA. The cancer survivorship journey: models of care, disparities, barriers, and future directions. Am Soc Clin Oncol Educ Book 2016;35:231-239 [FREE Full text] [doi: 10.1200/EDBK_156039] [Medline: 27249703]
53. Ashing KT. The iCCARE consortium for prostate cancer in black men: creating a survivorship care plan for black prostate cancer survivors. 2022 Presented at: 15th AACR Conference on the Science of Cancer Health Disparities in Racial/Ethnic Minorities and the Medically Underserved; September 16-19, 2022; Philadelpha, PA. [doi: 10.1158/1538-7755.disp22-a033]
54. Harris PA, Scott KW, Lebo L, Hassan N, Lightner C, Pulley J. ResearchMatch: a national registry to recruit volunteers for clinical research. Acad Med 2012;87(1):66-73 [FREE Full text] [doi: 10.1097/ACM.0b013e31823ab7d2] [Medline: 22104055]
55. Minority health and wellness. Mayo Clinic. 2025. URL: https://www.mayoclinic.org/about-mayo-clinic/minority-health-wellness [accessed 2025-07-17]
56. Odedina F, Scrivens JJ, Larose-Pierre M, Emanuel F, Adams AD, Dagne GA, et al. Modifiable prostate cancer risk reduction and early detection behaviors in black men. Am J Health Behav 2011;35(4):470-484. [doi: 10.5993/ajhb.35.4.9] [Medline: 22040593]
57. Cobran EK, Wutoh AK, Lee E, Odedina FT, Ragin C, Aiken W, et al. Perceptions of prostate cancer fatalism and screening behavior between United States-born and caribbean-born black males. J Immigr Minor Health 2014;16(3):394-400 [FREE Full text] [doi: 10.1007/s10903-013-9825-5] [Medline: 23576029]
58. Insight mHealth platform. TSET Health Promotion Research Center. URL: https://healthpromotionresearch.org/Mobile-Health [accessed 2025-07-18]
59. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
60. Mobile health (mHealth) technology. OU Health Stephenson Cancer Center. URL: https://www.ouhealth.com/stephenson-cancer-center/cancer-research/cancer-shared-resources-services/mobile-health-mhealth-technology/ [accessed 2025-02-10]
61. Kendzor DE, Businelle MS, Frank-Pearce SG, Waring JJC, Chen S, Hébert ET, et al. Financial incentives for smoking cessation among socioeconomically disadvantaged adults: a randomized clinical trial. JAMA Netw Open 2024;7(7):e2418821 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.18821] [Medline: 38954415]
62. Keopaseuth P, Xangsayarath P, Kulkarni S, Phandouangsy K, Soulaphy C, Alounlungsy P, et al. Mobile health technology for personalized tobacco cessation support among cancer survivors and caregivers in laos (project survlaos): protocol for a pilot randomized controlled trial. JMIR Res Protoc 2025;14:e66517 [FREE Full text] [doi: 10.2196/66517] [Medline: 40736921]
63. Hébert ET, Kendzor DE, Vidrine DJ, Langford JS, Kezbers KM, Montgomery A, et al. Just-in-time adaptive intervention for smoking cessation in low-income adults: a randomized clinical trial. JAMA Netw Open 2025;8(8):e2526691 [FREE Full text] [doi: 10.1001/jamanetworkopen.2025.26691] [Medline: 40810943]

XSL•FO
RenderX

64. Azizoddin DR, DeForge SM, Edwards RR, Baltazar AR, Schreiber KL, Allsop M, et al. Serious games for serious pain: development and initial testing of a cognitive behavioral therapy game for patients with advanced cancer pain. JCO Clin Cancer Inform 2024;8:e2400111 [FREE Full text] [doi: 10.1200/CCI.24.00111] [Medline: 39546741]
65. Maria DS, Nyamathi A, Lightfoot M, Paul M, Quadri Y, Padhye N, et al. Results of a randomized wait-list controlled trial of CAYA: a nurse case management HIV prevention intervention for youth experiencing homelessness. AIDS Behav 2025;29(2):613-625 [FREE Full text] [doi: 10.1007/s10461-024-04544-3] [Medline: 39531116]
66. Cheney MK, Alexander AC, Garey L, Gallagher MW, Hébert ET, Vujanovic AA, et al. Adapting a mobile health app for smoking cessation in black adults with anxiety through an analysis of the mobile anxiety sensitivity program proof-of-concept trial: qualitative study. JMIR Form Res 2025;9:e53566 [FREE Full text] [doi: 10.2196/53566] [Medline: 39918847]
67. Neil J, Mao B, Shao R, Ogunsanya ME, Frank-Pearce S, Businelle M, et al. A pilot randomized clinical trial of a smartphone-based application to support at-home PSA screening and culturally tailored prostate cancer education for African American men: a study protocol. Contemp Clin Trials 2024;147:107737 [FREE Full text] [doi: 10.1016/j.cct.2024.107737] [Medline: 39532237]
68. Ogunsanya ME, Kaninjing E, DeForge SM, Morton DJ, Cole P, Beckford P, et al. Bridging the gap: enhancing prostate cancer survivorship and advocacy among ethnically diverse black men through community town halls. J Cancer Educ 2025;40(6):865-874. [doi: 10.1007/s13187-025-02587-1] [Medline: 39994169]
69. Ojelabi AO, Graham Y, Haighton C, Ling J. A systematic review of the application of wilson and cleary health-related quality of life model in chronic diseases. Health Qual Life Outcomes 2017;15(1):241 [FREE Full text] [doi: 10.1186/s12955-017-0818-2] [Medline: 29228977]
70. Marikyan M, Papagiannidis P. Unified Theory of Acceptance and Use of Technology. Bristol, UK: Theoryhub Book; 2025. URL: https://open.ncl.ac.uk/ [accessed 2026-01-10]
71. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]
72. Cella D, Nichol MB, Eton D, Nelson JB, Mulani P. Estimating clinically meaningful changes for the functional assessment of cancer therapy--prostate: results from a clinical trial of patients with metastatic hormone-refractory prostate cancer. Value Health 2009;12(1):124-129 [FREE Full text] [doi: 10.1111/j.1524-4733.2008.00409.x] [Medline: 18647260]
73. Hussey I. An aberrant abundance of Cronbach's alpha values at .70. Adv Methods Pract Psychol Sci 2025;8(1):25152459241287123. [doi: 10.31234/osf.io/dm8xn]
74. Patient-centric clinical trial technology. Greenphire. 2025. URL: https://greenphire.com/participant-payments/ [accessed 2025-07-19]
75. Mahal BA, Gerke T, Awasthi S, Soule HR, Simons JW, Miyahira A, et al. Prostate cancer racial disparities: a systematic review by the prostate cancer foundation panel. Eur Urol Oncol 2022;5(1):18-29. [doi: 10.1016/j.euo.2021.07.006] [Medline: 34446369]
76. Pan L, Wu X, Lu Y, Zhang H, Zhou Y, Liu X, et al. Artificial intelligence empowered digital health technologies in cancer survivorship care: a scoping review. Asia Pac J Oncol Nurs 2022;9(12):100127 [FREE Full text] [doi: 10.1016/j.apjon.2022.100127] [Medline: 36176267]

Abbreviations

CaP: prostate cancer
EMA: ecological momentary assessment
FACT-P: Functional Assessment of Cancer Therapy – Prostate
HIPAA: Health Insurance Portability and Accountability Act
IRB: institutional review board
mHealth: mobile health
QoL: quality of life
REDCap: Research Electronic Data Capture
SAFE-CaPS: Survivorship App for Ethnically Diverse Black Prostate Cancer Survivors
SC: standard care
SCC: Stephenson Cancer Center
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
UTAUT: Unified Theory of Acceptance and Use of Technology
WCM: Wilson and Cleary Model

XSL•FO
RenderX

Edited by A Schwartz; submitted 30.Jul.2025; this is a non–peer-reviewed article;accepted 31.Dec.2025; published 24.Mar.2026.

Please cite as:
Kumar G, Ghasemi P, Alexander AC, Dwyer K, Neil JM, Cole P, Beckford P, Reese D, Montgomery E, Odeleye R, Harris MA, Zhao YD, Nagykaldi Z, McIntosh AG, Kaninjing E, Young ME, Dickey S, Morton DJ, Bolajoko O, Odedina FT, Planas LG, Businelle MS, Kendzor DE, Ogunsanya ME
A Culturally Tailored Mobile Health Intervention to Improve Quality of Life in Black Survivors With Prostate Cancer: Protocol for a Stratified Randomized Controlled Trial
JMIR Res Protoc 2026;15:e81503
URL: https://www.researchprotocols.org/2026/1/e81503
doi:10.2196/81503
PMID:

©Gaurav Kumar, Parisa Ghasemi, Adam C Alexander, Kathleen Dwyer, Jordan M Neil, Perry Cole, Patrick Beckford, Donald Reese, Everett Montgomery, Roland Odeleye, Mark A Harris, Yan D Zhao, Zsolt Nagykaldi, Andrew G McIntosh, Ernie Kaninjing, Mary E Young, Sabrina Dickey, Daniel J Morton, Opeyemi Bolajoko, Folakemi T Odedina, Lourdes G Planas, Michael S Businelle, Darla E Kendzor, Motolani E Ogunsanya. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy of Bacopa monnieri (Linn.) on Cognitive Function and Alterations in Blood Metabolites in Patients With Amnestic Mild Cognitive Impairment and Early Alzheimer Disease: Protocol for an Exploratory Double-Blind, Randomized, Placebo-Controlled Trial

Abhilasha Dwivedi[1], MA, MPhil; Anjali Anjali[1], BMLT; Hina Narzari[2], MSc, MPhil; Yashwant Kumar[3], MSc, PhD; Hanuman Prasad Sharma[4], MPharma, PhD; Aditi Dubey[1], MA, MPhil; Nilima Nilima[2], PhD; Roopa Rajan[1], MD, DM; Mamta Bhushan Singh[1], MD, DM; Venugopalan Y Vishnu[1], MD, DM; Rohit Bhatia[1], MD, DM; Gautam Sharma[5], MD, DM; Anu Gupta[1], MD, DM

[1]Department of Neurology, Faculty of Neurology, All India Institute of Medical Sciences, 7th Floor, Room no. 708, New Delhi, India

[2]Department of Biostatistics, All India Institute of Medical Sciences, New Delhi, India

[3]Noncommunicable Diseases, Translational Health Science and Technology Institute, Faridabad, India

[4]Bioanalytics Facility, Centralized Core Research Facility, All India Institute of Medical Sciences, New Delhi, India

[5]Department of Cardiology, Incharge Center for Integrative Medicine and Research, Faculty of Cardiology, All India Institute of Medical Sciences, Delhi, India

Corresponding Author:
Anu Gupta, MD, DM
Department of Neurology, Faculty of Neurology, All India Institute of Medical Sciences, 7th Floor, Room no. 708, New Delhi, India

Abstract

Background: Amnestic mild cognitive impairment (aMCI) represents a transitional stage between normal aging and Alzheimer disease (AD), where early intervention is critical for preserving cognition and delaying or preventing progression to dementia. Due to the limited availability of curative pharmacological treatments, there is growing interest in traditional and indigenous medicinal interventions, such as *Bacopa monnieri* (Linn.) or Brahmi, a widely used Ayurvedic nootropic herb. *Bacopa* consumption is known to enhance cognitive performance in healthy individuals and is associated with alterations in pathways related to branched-chain and aromatic amino acid biosynthesis. These pathways have been implicated in MCI due to AD pathophysiology. Hence, the efficacy of *Bacopa* in aMCI and the mediating metabolic changes need to be systematically evaluated through clinical trials.

Objective: The aim of this study is to assess the effects of *B. monnieri* on cognitive function and plasma metabolites in individuals with aMCI (early or prodromal AD).

Methods: This study will employ a double-blind, randomized, placebo-controlled trial design, with 76 participants (38 per group) of aMCI diagnosed using the clinical National Institute on Aging–Alzheimer's Association 2011 criteria in a tertiary care setting in India. Participants will receive either *B. monnieri* (Linn.) (300 mg standardized plant extract) or a matched placebo daily for 12 weeks. Comprehensive cognitive assessments (0, 12, and 24 wk) and untargeted plasma metabolomic profiling (0 and 12 wk) will be conducted to evaluate both cognitive changes and alterations in plasma metabolites. MetaboAnalyst 6.0 (with integrated features like Kyoto Encyclopedia of Genes and Genomes and the Human Metabolome Database) will be used for a comprehensive statistical and functional analysis pertaining to the metabolomics. The primary outcome will be a change in the composite *z* score of memory between the 2 groups at 12 weeks. The secondary outcomes will include alterations in metabolites and pathways, and adverse events at 12 weeks, and cognitive performance at 12 and 24 weeks.

Results: The study was funded in October 2023, and the first participant was enrolled in April 2024. As of November 2025, a total of 60 participants have been recruited, with a mean (SD) age of 62.2 (8.1) years. The cohort predominantly comprises men (53/60, 88.3%), individuals with at least a high school education (47/60, 78.3%), and participants at elevated risk for cognitive decline. Specifically, 39 of 60 participants (65%) have diabetes, 29 (48.3%) have hypertension, and 13 (21%) report a positive family history of dementia. Data collection will conclude in June 2026, after which data analysis will begin and be completed by September 2026, with primary findings targeted for publication in spring 2027.

XSL•FO
RenderX

Conclusions: This protocol investigates the efficacy of *B. monnieri* (Linn.) in improving cognitive function and altering the blood metabolites in patients with aMCI. If effective, this intervention could provide an accessible and cost-effective approach to manage early AD in resource-limited settings.

Trial Registration: Clinical Trial Registry of India CTRI/2021/12/038731; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=NjI4NDg=&Enc=&userName=

International Registered Report Identifier (IRRID): DERR1-10.2196/82891

(JMIR Res Protoc 2026;15:e82891) doi:10.2196/82891

KEYWORDS

Bacopa monnieri (Linn.); mild cognitive impairment; MCI; amnestic; Alzheimer disease; AD; metabolomics

Introduction

Background

Mild cognitive impairment (MCI), an intermediate clinical state between cognition of normal aging and mild dementia, is characterized by subjective cognitive complaints and objective impairment in one or more cognitive domains, such as memory, executive function, visuospatial skills, and language, while independence in instrumental activities of daily living is preserved [1]. The estimated prevalence of MCI in India is 17.6%, while the prevalence of major neurocognitive disorder such as Alzheimer disease (AD) is 7.2% [2]. Worldwide, MCI affects approximately 15% to 20% of older adults, while AD affects about 5% to 7% of the global population aged 60 years and above [3]. Considered the prodromal stage of dementia, MCI is a critical window period for early therapeutic intervention.

MCI has 2 clinical subtypes: amnestic mild cognitive impairment (aMCI) and nonamnestic MCI. aMCI is characterized by significant memory impairment, while in nonamnestic MCI, individuals face challenges with language, executive functions, or visuospatial skills [4]. aMCI subtypes are often linked to AD and pose a significantly higher risk of progression toward dementia [5]. Recently, antiamyloid therapies have received approval for MCI due to AD, but these drugs are costly, require sophisticated imaging or fluid tests before administration, and will take time to align to routine clinical practice [6]. The other US Food and Drug Administration approved medications for AD, such as donepezil (a cholinesterase inhibitor) and memantine (an N-methyl-D-aspartate receptor antagonist), are not approved for use in MCI. These medications also have numerous adverse effects, such as gastrointestinal effects (nausea, vomiting, diarrhea, loss of appetite), vagotonic (bradycardia, heart block), headache, confusion, and bladder incontinence [7,8]. In this scenario, traditional interventions for tackling cognitive decline are gaining focus.

The Indian system of medicine or Ayurveda (Sushruta Samhita, Charak Samhita, and Atharva Veda) describes plants that have a specific action on the intellect and memory as “Medhya Rasayana.” The word “Medhya” stands for “intellect or retention,” while “Rasayana” means “procedure or preparation” [9]. *Bacopa monnieri* (known as “Brahmi” in India), a herbal nutraceutical native to South and Southeast Asia, is recognized for this role and has been traditionally used for memory loss, anxiety, epilepsy, and insomnia [10]. Its leaves contain important natural compounds such as triterpenoid saponins (bacoside A and B), alkaloids, flavonoids, and phytosterols, with bacosides being essential for its effects on the brain, which are known to be pleiotropic. Numerous experimental studies have shown that *Bacopa* protects cholinergic neurons [11,12], upregulates expression of N-methyl-D-aspartate receptors (which mediate structural changes that occur in neurons during learning and memory formation) in the prefrontal cortex and hippocampus [13], reduces stress-induced hippocampal damage [14], enhances cerebral blood flow [15], increases expression of neuroplasticity markers like brain-derived neurotrophic factor and Arc protein expression [16], and has an antioxidant and free radical scavenging action [17-19]. It also has an anti-amyloidogenic potential, reducing formation of amyloid fibrils and offering protection against neuronal cell death induced by amyloid-β protein [20-22].

Looking at the clinical translation of the above experimental studies, Kongkeaw et al [23] did a meta-analysis of 9 randomized controlled trials (RCTs) where standardized extracts of *B. monnieri* were administered to 437 eligible participants (heterogeneous study population in terms of age range; most were healthy volunteers, except for 79 participants who had memory complaints) for at least 12 weeks. The authors reported a significant improvement in cognitive processing speed after 12 weeks of *B. monnieri* extract. However, the effect of *Bacopa* on the memory domain was inconclusive because of high heterogeneity. Subsequent studies using *Bacopa* in healthy older adults have reported improvement in working memory, a decrease in N100 (assessing selective attention) and P300 latencies (assessing cognitive processing) of event-related potentials, and suppression of plasma acetylcholinesterase activity [24].

To evaluate the role of *Bacopa* in people with cognitive impairment, Basheer et al [25] did a meta-analysis on the effectiveness of *B. monnieri* in mild, moderate, or severe dementia due to AD or MCI-AD. The authors found no difference between *B. monnieri* and the placebo or donepezil. The analysis included 5 eligible studies; 2 compared *Bacopa* with donepezil, and the rest compared Bacopa with placebo. There was heterogeneity in dosage, duration, formulation, follow-up, and outcomes, and the overall quality of evidence was very low due to high risk of bias, small sample size, and wide CIs. However, no major safety issues were reported in the included trials.

XSL•FO
RenderX

Recently, Minale et al [26] used liquid chromatography-mass spectrometry (LC-MS) in healthy individuals (age range, 55 - 80 y) who had consumed *Bacopa* extract for 12 weeks and showed changes in metabolites (aminoacyl-transfer RNA biosynthesis, aromatic and branched-chain amino acid biosynthetic pathways) in plasma, urine, and feces. Prior research findings on blood and cerebrospinal fluid metabolites in AD and MCI have suggested possible alterations in metabolomic pathways associated with cellular energy and biosynthesis, oxidative stress, amino acid utilization, lipid homeostasis, mitochondrial oxidative metabolism, and neurotransmitters [27-31]. Yilmaz et al [32] combined proton nuclear magnetic resonance spectroscopy, plasma LC-MS, and machine learning statistical approaches in a community-based sample cohort acquired from different sites across the United States and reported significant perturbations in amino sugar, alanine, glucose-alanine cycle (energy supply to muscles and nitrogen disposal), cysteine and aspartate metabolism, and urea cycle in patients with MCI compared with healthy controls. Our recent work using untargeted metabolomics in Indian patients with AD (vs cognitively normal controls) showed disruptions in branched-chain amino acid metabolism, nicotinamide metabolism, the tricarboxylic acid cycle, and neurotransmitter pathways [33]. Although the study by Minale et al [26] in healthy subjects suggests that *Bacopa* may affect pathways implicated in MCI and AD, it needs further validation.

Need for the Study

aMCI has a high risk of progression to Alzheimer dementia. Currently, there are few approved treatments for aMCI. Antiamyloid agents are the only approved pharmacological options and need sophisticated tests for accurate diagnosis of amyloid pathology in the brain. These drugs are costly, need monitoring for serious adverse effects (SAEs), and are not yet available in low- and middle-income countries. *B. monnieri*, a herb that has been used as a memory enhancer for centuries in India, is a culturally aligned intervention, cheap, and without SAEs. Most studies in the literature on the effectiveness of *Bacopa* on cognitive functions have focused on healthy individuals. Reports in cognitively impaired older adults are few and provide low-certainty evidence. Furthermore, few studies have explored the underlying biological mechanisms of *Bacopa* using the modern "omics" approaches. This study aims to fill these critical gaps by assessing the efficacy of *B. monnieri* (Linn.) in improving memory and other cognitive functions in individuals with aMCI (early or prodromal AD) through an exploratory double-blind randomized controlled trial design. It will also examine related changes in blood metabolites through untargeted metabolomics. We will use a standard *Bacopa* extract containing 50% bacosides, a matching placebo, and culturally valid cognitive tests. The results may contribute to therapeutic strategies for this entity and promote evidence-based traditional medicine within mainstream health care.

Objectives

The primary objective of the study is to evaluate the effect of *B. monnieri* (Linn.) on memory as compared with placebo among individuals with aMCI after 12 weeks of intervention. The secondary objectives are to identify alterations in metabolites, pathways, and networks after administration of *B. monnieri* (Linn.) and to evaluate the effect on other cognitive domains (processing speed, attention, and executive functions) at 12 and 24 weeks. The study tests the hypothesis that *B. monnieri* (Linn.) (300 mg standardized whole-plant extract) is more effective than placebo in improving the memory in patients with aMCI over 12 weeks.

Methods

Study Design

This study will follow a double-blind, randomized, placebo-controlled trial design with 3 phases of assessment: baseline evaluation (week 0), postintervention assessment at 12 weeks, and follow-up assessment at 24 weeks.

Study Settings

The study will be conducted at the outpatient services of the Department of Neurology, All India Institute of Medical Sciences (AIIMS), New Delhi, a tertiary academic and research institute in India. Participants presenting with memory complaints will be screened and recruited by qualified neurologists and professional clinical psychologists based on predefined inclusion and exclusion criteria. Eligible participants will be enrolled in the study and randomly assigned to 1 of 2 groups using a stratified block randomization method (stratified by age: 45 - 60 and 61 - 75 y) with a fixed block size of four and a 1:1 allocation ratio. Participants will receive either *B. monnieri* (300 mg standardized whole-plant extract) or a matched placebo, taken once daily with water after breakfast for 12 weeks. All clinical assessments, blood sample collections, and follow-up visits will be done at the designated study location.

Study Population (Inclusion Criteria)

The trial will recruit participants aged 45 to 75 years who have a clinical diagnosis of aMCI or early AD, diagnosed using the clinical criteria given by the National Institute of Ageing-Alzheimer's Association workgroup (2011) [34]. Participants should report a decline in memory compared with previous ability, have impaired scores (≤15th percentile on the published norms as abnormal) on tests of verbal or visuospatial memory on the detailed neuropsychological battery, and preserved functional independence [35]. They should be willing to participate in the study, give written informed consent, be able to comply with the study requirements, read and comprehend either English or Hindi for cognitive test assessments, and should have at least 5 years of formal school education (participants with low education may not perform as well on neuropsychological testing) [36].

Exclusion Criteria

Participants with advanced malignancies; significant systemic diseases such as chronic kidney disease, severe hepatic, pulmonary, cardiac disease; other neurologic diseases or drugs that are known to affect cognition (like structural brain lesions, stroke, epilepsy, traumatic brain injury, Parkinson disease, Parkinson-plus syndromes, normal pressure hydrocephalus, autoimmune/paraneoplastic encephalitis); history of a major

XSL•FO
RenderX

surgical procedure in the last 6 months; uncontrolled hypothyroidism (thyroid-stimulating hormone >10 µU/ml at time of enrollment) will be excluded. Participants with a clinically diagnosed major psychiatric disorder (eg, psychosis, major depression, bipolar disorder), those who use psychoactive medications that could affect their ability to reliably perform neurocognitive testing, have visual or auditory impairment that would preclude the participant from participating in or cooperating with the protocol, have a history of intake of *B. monnieri* or any other investigational drug or cholinesterase inhibitors within 1 month prior to screening visit, or with alcohol abuse as defined by the National Institute on Alcohol Abuse and Alcoholism will also be excluded. The National Institute on Alcohol Abuse and Alcoholism defines heavy drinking as follows: for men, consuming more than 4 drinks on any day or more than 14 drinks per week; for women, consuming more than 3 drinks on any day or more than 7 drinks per week [37].

Recruitment and Baseline Assessment

Eligible participants will be identified by neurologists and clinical psychologists. Participants meeting the inclusion and exclusion criteria will undergo cognitive screening using the Indian adaptation of the Addenbrooke's Cognitive Examination III (ACE-III; Hindi version) [38] and functional screening Instrumental Activities of Daily Living in the Elderly scale[39]. Participants with an ACE-III score greater than 62 [40] and a cognitive disability index ≤16 will be invited for a detailed cognitive assessment. This excludes mild dementias and ensures that all the recruited participants are at the MCI stage. A structured questionnaire will be used to gather information on sociodemographic characteristics, medical history, concomitant medications, vital signs, lifestyle factors, and risk factors (Table S1 in Multimedia Appendix 1). Cognitive assessments will be conducted by trained clinical psychologists using the Indian Council of Medical Research-Neurocognitive Tool Box [35], a culturally validated tool for diagnosis of MCI in India (Table 1). The study flow is presented in Figure 1.

Table . Details of cognitive assessment.

Cognitive domain	Test
Attention and executive functioning [35]	
Visuospatial attention, concentration, and set shifting	Trail making test (black and white)
Verbal fluency	Category fluency (animal, food, vegetable); phonemic fluency (Ka, Ma, Pa)
Episodic memory	
Verbal memory	Verbal learning test (total learning, learning over trials, delayed recall, and delayed recognition)
Visuospatial memory	Test des Neuf Images du—93 (TNI-93; immediate recall, free recall, cued recall and spatial recall)
Global cognition	
Screening of all domains	Addenbrooke's Cognitive Examination III
Functional activities	
Instrumental activities of daily living	Instrumental Activities of Daily Living–Elderly

XSL•FO
RenderX

Figure 1. Study flow. Inclusion criteria included: aged 45-75 years, with a clinical diagnosis of amnestic mild cognitive impairment (NIA-AA, 2011 clinical criteria), willing to participate and give written informed consent, able to read and comprehend either English or Hindi, and with at least 5 years of formal school education. Exclusion criteria included: participants with advanced malignancy, significant systemic diseases like chronic kidney disease, severe hepatic, pulmonary, cardiac disease, other neurologic diseases or drugs known to affect cognition, major surgical procedure in the last 6 months; thyroid-stimulating hormone >10 µU/ml at time of enrollment; major psychiatric disorder (eg, psychosis, major depression, bipolar disorder); use of psychoactive medications; visual or auditory impairment; history of intake of *Bacopa monnieri* or any other investigational drug or cholinesterase inhibitors within 1 month prior to screening visit; or presence of alcohol abuse (National Institute on Alcohol Abuse and Alcoholism guidelines). AIIMS: All India Institute of Medical Sciences; NIAAA: National Institute on Alcohol Abuse and Alcoholism; NIA-AA: National institute on Ageing-Alzheimer's Association.

XSL•FO
RenderX

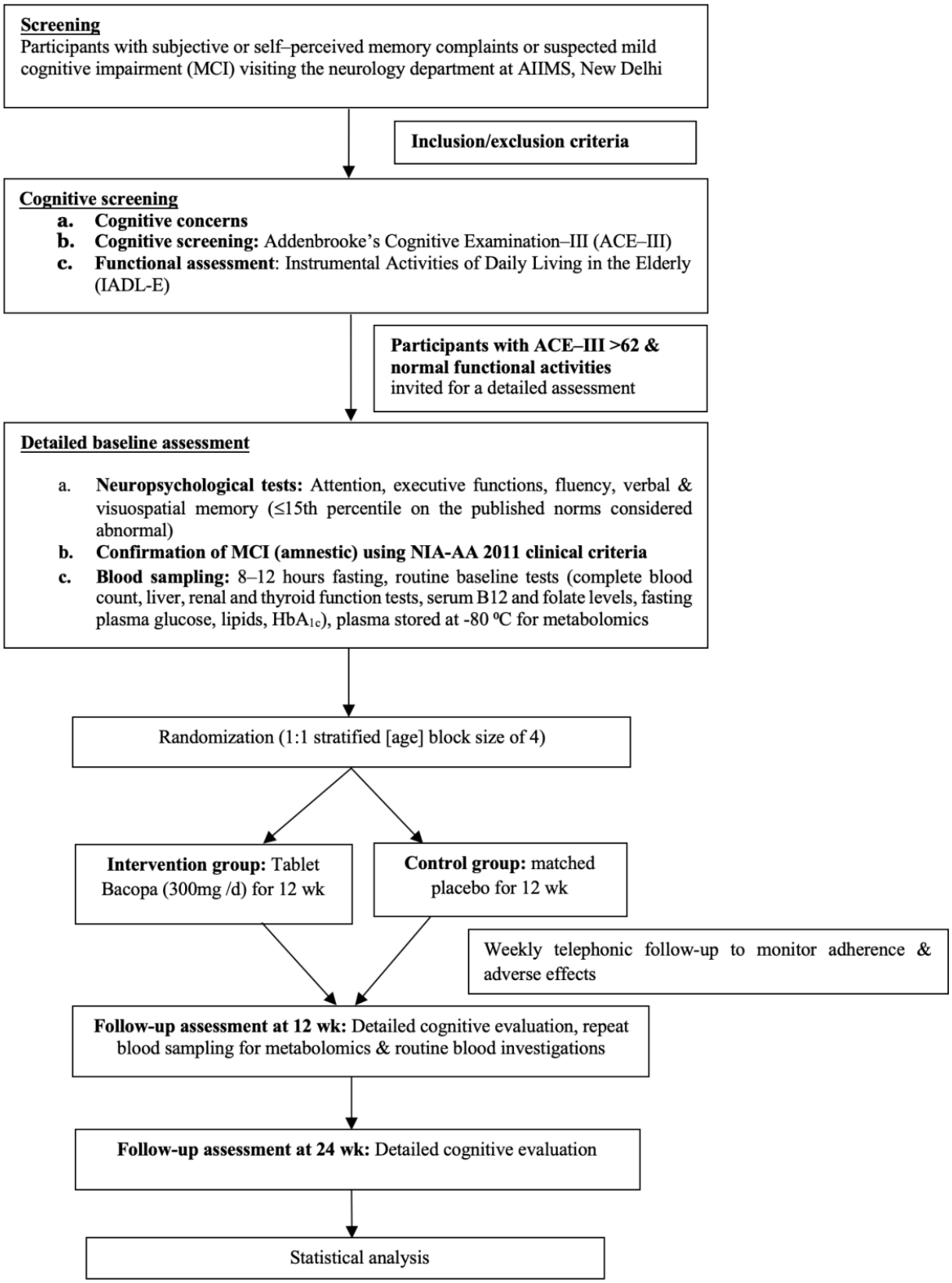

XSL•FO
RenderX

Blood Sample Collection and Storage

Venous blood samples will be collected by certified laboratory technicians. Participants will be instructed to maintain an 8 to 12 hours fast and to abstain from alcohol for at least 72 hours before sample collection. Blood will be drawn aseptically into the following tubes: 2 plain tubes, 3 ethylenediaminetetraacetic acid tubes, and 1 gray-top (fluoride) tube. Samples showing signs of hemolysis will be excluded from further analysis. The collected samples will be used for routine biochemical investigations, including complete blood count, glycated hemoglobin (HbA_{1c}), liver function tests, kidney function tests, lipid profile, thyroid profile, fasting blood glucose, vitamin B12, and folate levels in serum. For metabolomic analysis, 1 plain and 1 ethylenediaminetetraacetic acid tube will be used, and the samples will be transported via cold chain. All samples will be centrifuged at 2500 rpm for 15 minutes at room temperature to separate serum and plasma within 1 hour of drawing the sample. The separated components will be aliquoted into cryovials and stored immediately at –80 °C until further analysis.

Randomization

Participants will be randomly assigned to the intervention or control group using a computer-generated list. As age is the most significant risk factor for cognitive decline, and our study has a relatively small sample size, we will do a stratified randomization by age (45 - 60 and 61 - 75 y) to avoid significant baseline differences pertaining to age in the 2 groups. The sequence will be created using Stata 17 by an independent statistician. Unique IDs starting with “BRY2001-BRY2038” for younger and “BRE3001-BRE3038” for older participants will be assigned prior to allocation to ensure a systematic and unbiased process. To maintain a 1:1 allocation ratio and minimize selection bias, a block size of 4 will be applied within each age stratum. This approach ensures that participants in both age groups are evenly distributed between the intervention and placebo arms. To prevent selection bias, the randomization list will be kept strictly confidential and accessible only to the independent statistician, who will not be involved in recruitment or data collection.

Allocation Concealment

The statistician generating the randomization list will prelabel the medicine bottles (intervention and placebo) with sequential IDs based on the randomization list. These prelabeled bottles will be used during participant enrollment, ensuring that the recruitment team remains unaware of group assignments at all times. This process ensures that the allocation remains concealed before and during assignment, maintaining the integrity of the randomization process.

Blinding

This study will adopt a double-blind design, in which participants, caregivers, and investigators including those administering the intervention, conducting assessments, and handling data are blinded to group allocation throughout the study. This blinding helps reduce both performance and detection bias, supporting the reliability and objectivity of the trial results.

Intervention Arm

Participants will receive either the *B. monnieri* (Linn.) tablets (300 mg) or a matched placebo, identical in appearance, weight, shape, size, color, and packaging. *B. monnieri* (Linn.) will be administered as a standardized whole-plant extract at a dose of 300 mg once daily, taken orally in the morning after meals with water. At baseline, each participant will be provided with a bottle containing 90 tablets, along with 6 additional tablets to account for potential missed doses. The *B. monnieri* (Linn.) formulation (Dabur India Limited) contains 50% bacosides, and the placebo will be devoid of active bacosides while maintaining physical similarity. The intervention period will last for 12 weeks. Concomitant medications for comorbidities will be allowed, and any change in the prescriptions will be documented.

Control Arm

The placebo tablets will be carefully matched with the investigational tablets in terms of color, taste, texture, shape, and size, and will be supplied in similar bottles. Placebo tablets will be prepared using starch and the same excipients that are known to be biologically inactive. Each participant will be given dispensed bottles (containing 90 tablets) at the time of the baseline visit.

Methodology for Metabolomics

This study will use the LC-MS approach for untargeted metabolite analysis.

Sample Preparation for Protein Extraction

Prior to LC-MS/MS analysis, plasma samples will be deproteinized using a protein precipitation technique. Nonaqueous solvents such as chilled methanol will be used to remove proteins. Plasma samples will be thawed on ice and be vortexed using a vortex mixer to ensure uniformity. For each sample, 100 µl of plasma will be transferred into a new 2-ml microcentrifuge tube, followed by the addition of 400 µl of chilled methanol. The mixture will be vortexed for 1 minute and then centrifuged at 10,000 rpm for 10 minutes at 4 °C. After centrifugation, the clear supernatant will be carefully transferred into new 2 ml microcentrifuge tubes in a fixed volume. These samples will then be subjected to vacuum drying using a CentriVap Cold Trap vacuum concentrator (Labconco, Kansas City, MO, USA). Finally, each sample will be resuspended in 25 µl of a methanol-water mixture (3:17, methanol:water), vortexed briefly, and centrifuged at 11,000 rpm for 10 min at 4 °C and injected into LC-MS/MS.

LC-MS/MS Methodology

The Orbitrap Fusion mass spectrometer (Thermo Scientific) coupled with a heated electrospray ion source will be used for data acquisition. Data acquisition methods have been followed as per published protocols [41,42] with minor modifications. Briefly, for MS1 mode, mass resolution will be kept at 120,000, and for MS2 acquisition, mass resolution will be 30,000. The mass range of data acquisition will be 60 to 900 Da. Extracted metabolites will be separated on UPLC Ultimate 3000. Data will be acquired on reverse-phase and hydrophilic interaction liquid chromatography columns in both positive and negative

XSL•FO
RenderX

ionization modes. HSS T3 will serve as a reverse phase column, while XBridge BEH Amide (Waters Corporation) will be used as a hydrophilic interaction liquid chromatography column. For polar compound separation, solvent A will have 20 mM ammonium acetate in the water (pH 9), and mobile phase B will have 100% acetonitrile. The elution gradient will start from 85% B to 10% B over 14 min with a flow rate of 0.35 ml/min. For reverse phase, solvent A will be water, and B will be methanol with 0.1% formic acid added in both solvents. The elution gradient will start with 1% B to 95% B over 10 min with a flow rate of 0.3 ml/min. The sample injection volume will be 5 µl. Pooled quality control samples will run after every five samples to monitor signal variation and drift in mass error.

Data Processing

All LC-MS–acquired data will be processed using the Progenesis QI for Metabolomics (Water Corporation) software using the default settings. The untargeted workflow of Progenesis QI facilitates retention time alignment, feature detection, deconvolution, and prediction of elemental composition. The Metascope plug-in of Progenesis QI will be used for the in-house library, incorporating accurate mass, fragmentation patterns, and retention times for database searches. The study will use an online spectral library for further confirmation of identification. The retention time match cut-off will be set at 0.5 minutes, with a spectral similarity exceeding 30% for fragmentation matches in Progenesis QI. Peaks with a coefficient of variation below 30% in pooled quality control samples will be chosen for subsequent data analysis. Furthermore, manual verification of each identified feature will be conducted to select the appropriate peaks.

Follow-Up

There are 3 in-person primary data collection points in the study: “baseline evaluation (week 0),” “postintervention assessment at 12 weeks,” and “second follow-up assessment at 24 weeks.” Between the in-person visits, participants will receive weekly telephonic follow-up calls to promote participant retention, monitor medication adherence, and collect information on any SAEs. In the case of any SAEs, participants will be instructed to discontinue the study medication immediately and to contact the study team for appropriate medical evaluation. In such scenarios, unblinding will be done by the statistician who has access to the randomization and conveyed to the treating medical team. Following completion of the 12-week intervention period, participants will return to the hospital for a comprehensive follow-up visit, which will include assessment of primary and secondary outcome measures; and blood sample collection for postintervention metabolomic analysis, and repeat laboratory investigations, including hematological, hepatic, and renal function tests. Participants will be allowed to withdraw from the study if they wish to do so, and the reason for withdrawal will be clearly documented. They will be encouraged to return the used bottles at the end of 12 weeks or if they withdraw from the study at any time point, for pill count to confirm adherence. The second follow-up assessment will take place at 24 weeks and will include detailed cognitive evaluation. Blood sampling will not be done at this time point. Participants will continue to receive posttrial standard of care from their treating physicians.

Outcomes

The primary outcome will be change in the composite *z* score of memory between the 2 groups at 12 weeks. The memory domain will include both verbal (Verbal Learning Test: learning over trials, delayed recall, delayed recognition) and visuospatial memory (Test des Neuf Images du 93-Spatial Recall). The secondary outcomes will include alterations in metabolites, pathways, and molecular networks, and frequency of SAEs after 12 weeks of treatment with *B. monnieri*; change in *z* scores for other cognitive domains (attention, executive functioning, and information processing) at 12 and 24 weeks (including *z* score for the memory domain at 24 wk); proportion of participants in both groups who demonstrate clinically meaningful cognitive improvement, defined as either ≥5 point improvement on the ACE-III [43], or ≥0.5 SD improvement in the composite *z* score of memory domain in the *Bacopa* versus placebo group at 12 and 24 weeks. Adherence will be monitored through self-reported medication logbooks and weekly telephonic follow-ups.

Ethical Considerations

This study has received ethical approval from the Institute Ethics Committee of the All-India Institute of Medical Sciences (AIIMS), New Delhi, India (reference IEC–660/03.09.2021, RP–35/2021). Written informed consent will be obtained from all participants before their enrollment in the study by the clinical psychologist under the supervision of the principal investigator. Participants will be told about the study’s goals, methods, risks, and benefits, as well as their right to withdraw and the steps taken to keep their data confidential. Data will be entered into REDCap (Research Electronic Data Capture), which is a secure web application for managing databases. Any data sharing will involve anonymization of personally identifiable information. All procedures will adhere to the ethical standards established by the institutional review board and the principles outlined in the Declaration of Helsinki. This is an academic, investigator-initiated, nonpharma-funded clinical trial; hence, there is no provision of compensation in the study protocol. However, any adverse event related to the research will be treated free of cost in the hospital.

Statistical Methods

Sample Size Estimation

There are no previous studies investigating the effects of *B. monnieri* versus placebo in patients with aMCI, looking at changes in composite memory *z* scores as the primary outcome. By considering a clinically relevant effect size as a SD change of 0.5, a 2-sided α at 5%, and 80% power, a sample size of 128 (64 in each group) was calculated. With a 20% dropout rate, the final sample size reached 160 (80 in each group), which did not seem feasible for the selected patient population in the specified duration. Therefore, a sample size of 60 (30 in each group) is determined for this exploratory study based on feasibility. The sample size was further adjusted for a dropout rate of 20% and 76 participants (38 in each group) will be recruited for this study.

XSL•FO
RenderX

Planned Analyses

As mentioned previously, data will be entered into REDCap, a secure web-based application for managing databases. A data monitoring committee is not needed for this trial, as it is a short-duration study and the known risks of the intervention drug are minimal. No interim analysis is planned. Data will be analyzed using Stata version 17.0 (StataCorp LLC, College Station, TX, USA) and presented data as number (%), mean (SD), or median (IQR), as appropriate. Pearson chi-square test or Fisher exact test will be used (as appropriate) for qualitative variables. For quantitative variables, normality will be assessed using the Shapiro-Wilk test. The Student *t* test for independent samples will be used to compare differences in means for normally distributed quantitative variables. The Mann-Whitney *U* test will be used for quantitative variables with a distribution other than normal. All tests will be 2-tailed, and a *P* value <.05 will be considered statistically significant. Raw scores of each cognitive test will be converted to a *z* score (*z* score=[individual participant raw score – group mean]/group SD) for each participant. The group mean and SD will be derived by pooling the baseline scores of all participants. For all the tests, a higher *z* score will indicate better performance, except for the trail making tests A and B. The signs of these two tests will be inverted for a uniform interpretation. Both intention-to-treat and per protocol analysis will be done. Generalized linear mixed models with a between-subject factor of group (treatment vs placebo) and within-subject factor of time (baseline, mid-point, end point) will be conducted to assess differences in continuous primary and secondary outcome measures. Planned simple contrasts will be conducted to assess all time points (baseline vs. 12 wk, 12 wk vs. 24 wk, and baseline vs 24 wk). To handle any missing data, analysis of last observation carried forward will be undertaken within the intention-to-treat analysis model. Comprehensive statistical and functional analysis pertaining to the metabolomics will be performed using MetaboAnalyst version 6.0. The investigation based on observed peak values will encompass principal component analysis, partial least squares discriminant analysis, ANOVA, and molecular pathway identification. Heat maps will be developed with samples arranged in columns and features in rows, presented in a gradient from cold to hot. Pathway analysis will be conducted utilizing the integrated features of MetaboAnalyst 6.0, which includes the Kyoto Encyclopedia of Genes and Genomes and the Human Metabolome Database.

Reporting Guidelines

The SPIRIT 2025 reporting guidelines have been used while reporting this study protocol. The checklist is provided in Checklist 1. Changes made to protocol after submission to CTRI are summarized in Table S2 in Multimedia Appendix 1.

Dissemination of Results

The results will be presented at scientific meetings and will be published in scientific journals. The results will also be conveyed to the public in simple language during patient awareness programs after study completion.

Results

Study Status and Timeline

The study was funded in October 2023, and the first participant was enrolled in April 2024. Data collection will conclude in June 2026, after which data analysis will begin and be completed by September 2026, with primary findings targeted for publication in spring 2027.

Baseline Characteristics

As of November 2025, a total of 60 participants had been recruited, with a mean age of 62.2 (SD 8.1) years. Most participants were men (53/60, 88.3%), and 47/60 (78.3%) had completed at least high school (Table 2). Among the risk factors, 39/60 (65%) had diabetes, 33/59 (55.9%) had dyslipidemia, 29/60 (48.3%) had hypertension, 29/60 (48.3%) were overweight or obese, 15/60 (25%) had coronary artery disease, and 13/60 (21%) reported a positive family history of dementia (Table 3). The biochemical profile and the cognitive raw scores are provided in Tables S3 and S4 in Multimedia Appendix 1, respectively.

XSL•FO
RenderX

Table . Baseline sociodemographic characteristics of the participants.

Characteristics	Whole cohort (N=60)
Age, y, mean (SD)	62.2 (8.1)
Men, n (%)	53 (88.3)
Married, n (%)	57 (95)
Education (high school and above), n (%)	47 (78.3)
Occupation, n (%)	
Unemployed	3 (5)
Semiskilled and skilled workers	27 (45)
Semiprofessional to professional	16 (26.7)
Arithmetic skill job	14 (23.3)
Monthly family income (INR[a]), mean (SD)	37,462.5 (18,584.7)
Socioeconomic status, n (%)	
Upper class	37 (61.7)
Middle class to upper-middle class	20 (33.3)
Lower class to lower-middle class	3 (5)

[a]1 INR equals US $0.011.

XSL•FO
RenderX

Table . Baseline risk factor profile of the participants.

Characteristics	Whole cohort (N=60)
Height (cm), mean (SD)	163.7 (6.9)
Weight (kg), mean (SD)	68.4 (10.2)
Body mass index (kg/m²), mean (SD)	25.6 (3.8)
Average SBP[a] (mm Hg), mean (SD)[b]	126.7 (16.5)
Average DBP[c] (mm Hg), mean (SD)[b]	75 (10.5)
Pulse rate (bpm), mean (SD)	92.5 (11.8)
Duration of forgetfulness (months), mean (SD)	40.8 (23.6)
Hypertension, n (%)	29 (48.3)
Duration of hypertension (months), median (IQR)	72 (3-138)
Diabetes, n (%)	39 (65)
Duration of diabetes (months), median (IQR)	96 (36-186)
CAD[d], n (%)	15 (25)
Duration of CAD (months), median (IQR)	120 (90-162)
Dyslipidemia, n/N (%)	33/59 (55.9)
Overweight, n (%)	20 (33.3)
Obese, n (%)	9 (15)
Hypothyroidism, n (%)	3 (5)
Hyperthyroidism, n (%)	5 (8.3)
Alcohol use, n (%)[e]	3 (5)
History of smoking, n (%)	0 (0)
Family history of dementia, n (%)	13 (21)

[a]SBP: systolic blood pressure.

[b]Average of 3 readings taken 2 minutes apart.

[c]DBP: diastolic blood pressure.

[d]CAD: coronary artery disease.

[e]Not fulfilling the National Institute on Alcohol Abuse and Alcoholism criteria of alcohol abuse.

Deviations From Original Registered Protocol

We made some amendments to the inclusion and exclusion criteria, allocation concealment procedure, and cognitive test battery before the commencement of participant recruitment (December 2023). These revisions were minor and intended to improve implementation, feasibility, and the precision of cognitive assessments. They did not affect the overall study design, sample size, randomization strategy, or predefined clinical outcomes. These deviations are summarized in Table S2 in Multimedia Appendix 1. We have added “at least 5 years of formal school education” to the inclusion criteria, and “participants who have undergone a major surgical procedure in the last 6 months, or thyroid-stimulating hormone >10 µU/ml at time of enrollment” to the exclusion criteria. We also elaborated on “other neurologic diseases or drugs known to affect cognition” (eg, structural brain lesions, stroke, epilepsy, traumatic brain injury, Parkinson disease, Parkinson-plus syndromes, normal pressure hydrocephalus, autoimmune/paraneoplastic encephalitis) in the exclusion criteria. Prelabeled medicine bottles (intervention and placebo) with sequential IDs based on the randomization list are being used for allocation concealment instead of “sequentially numbered opaque sealed envelopes.” Two changes were made to the cognitive battery: ACE-III (clinically significant improvement will be defined as ≥5-point improvement in score) is being used for cognitive screening and Test des Neuf Images du 93 for visuospatial memory in place of Montreal Cognitive Assessment scale and Modified Taylor Complex Figure Test, respectively.

Discussion

Principal Findings

The current clinical trial seeks to evaluate the potential of *B. monnieri* (Linn.) as a therapeutic intervention for middle-aged and older adults (aged 45 - 75 y) diagnosed with aMCI. Previous literature on the efficacy of *B. monnieri* in enhancing cognitive functions has shown variable results. Most clinical studies that support its efficacy for improving memory and other cognitive functions have been done in healthy adults [23,24]. There is limited data on its efficacy in aMCI. A Cochrane review on the

XSL•FO
RenderX

clinical efficacy and safety of *B. monnieri* in persons with mild, moderate, or severe dementia due to AD or with MCI-AD by Basheer et al [25] found a high risk of bias in the included trials, and imprecise results due to small sample sizes (most studies had 30 - 40 participants in total) and wide CIs. There was heterogeneity in the formulation (some studies have used *Bacopa* in combination with herbal extracts, others have used *Bacopa* extract only), comparator (donepezil or placebo), treatment duration, follow-up, and outcomes. We have summarized the randomized controlled trials in which *B. monnieri* has been used for healthy older adults, or MCI, or AD dementia in Table S5 in Multimedia Appendix 1.

Very few studies have reported the efficacy and safety of *Bacopa* in different dosages and formulations. Most studies have compared a dried extract composed of 300 to 320 mg of *Bacopa*, either once a day or in two divided dosages, against placebo and reported efficacy with no major side effects [24,26,44,45]. Others have used doses ranging from 250 to 450 mg and 600 mg [11,46-48]. Studies have not demonstrated any incremental benefit with a higher dose [24].

From a safety perspective, Pravina et al [49], in a phase I study, evaluated the short-term safety and tolerability of *B. monnieri* in healthy adult volunteers (aged 22 - 42 y) through a randomized, open-label, dose-escalation design. The standardized extract of *B. monnieri* was given to each participant in a dose of 300 mg per day for 15 days, and then 450 mg per day for another 15 days. Pre- and post-treatment examination of clinical, hematological, biochemical, and electrocardiographic parameters of the treated volunteers did not indicate any untoward effects. There were mild adverse events related to the gastrointestinal system in the trial (300 mg dosage: epigastric burning sensation in one volunteer, nausea in one volunteer; 450 mg dosage: fullness and bloating sense of abdomen in one volunteer) which resolved spontaneously without any need to discontinue treatment. There were no dropouts in the study period. At the dose administered for the given duration of the trial period, *B. monnieri* extract was found to be safe in healthy adults. Stough et al [50] have also reported a higher percentage of participants with nausea (18% vs 4%), dry mouth (23% vs 16%) and fatigue (14% vs 4%) in the *B. monnieri* (300 mg for 12 wk) group in comparison to placebo (double-blind placebo-controlled study in 46 healthy volunteers aged between 18 and 60 y). Other studies have reported rash, constipation, drowsiness, and aftertaste following ingestion of *B. monnieri* formulation [25]. Prabhakar et al [51] in their study comparing *Bacopa* to donepezil reported 3 deaths due to myocardial infarction (2 in the donepezil arm and 1 in the *B. monnieri* arm).

Given the complexity of AD pathophysiology, omics-based investigations of molecular and cellular pathway disruptions are gaining momentum, as they can guide biomarker and therapeutic development. From a mechanistic perspective, *B. monnieri* has been postulated to have antiapoptotic, antioxidant actions, repairs damaged neurons, stimulates kinase activity, restores synaptic function, and improves neurotransmission. A systematic review of 22 studies reported that *B. monnieri* reduces the levels of proinflammatory markers, oxidative stress, and also reduces nuclear factor-κB phosphorylation [52]. However, there is a lack of adequate clinical research on the molecular targets and pathways of *Bacopa* in modulating neuroinflammation and neurotrophism. In our previous work using untargeted metabolomics in Indian patients with AD, we found significant disruption in branched-chain amino acid metabolism, and *Bacopa* consumption for 12 weeks has been shown to alter metabolites of this pathway in healthy adults [33]. We intend to gain better mechanistic insights through the rigorous RCT design of this study and also correlate metabolomic changes with cognitive outcomes.

Strengths and Limitations

The strengths of the study include a double-blind, randomized, placebo-controlled trial design, which is considered the gold standard for evaluating intervention efficacy, clearly defined relevant predictor variables and primary and secondary outcomes; a standardized and culturally validated cognitive battery tailored for the Indian population; and the use of standardized *B. monnieri* (with 50% bacosides) formulation and a matching placebo, ensuring consistency in dosage and bioactive content across participants. The second assessment at 24 weeks (12 wk after the intervention is stopped) will inform about the sustainability of cognitive changes and differences, if any, in the 2 groups. Finally, integration of metabolomics with the clinical outcomes will provide mechanistic insights on the actions of *B. monnieri* on cognition. The limitations include the exploratory nature of the study, lack of sufficient power to evaluate the efficacy, lack of biomarkers to confirm the etiological diagnosis of AD, and a relatively short follow-up period to observe meaningful change in cognitive functions.

Future Work

Future studies could look at the efficacy and mechanisms of *B. monnieri* Linn in a larger sample through multicenter studies, using a longer duration of intervention (12 - 18 mo), in a more precisely selected patient population using blood biomarkers, in genetically diverse phenotypes (eg, APOE4 positive and negative), and in the preclinical phase as a neuroprotector.

Conclusion

To conclude, this study aims to evaluate the effectiveness of *B. monnieri* in improving cognitive performance, particularly memory, and resultant changes in blood metabolites over a 12-week intervention period in middle-aged and older Indian adults with aMCI or early AD. The results of this study have the potential to generate evidence on the role of traditional therapies as early intervention strategies for cognitive decline, particularly in conditions for which conventional treatments are unavailable.

We plan to disseminate the results of this study through scientific meetings and publication in scientific journals. The results will also be conveyed to the public in simple language during patient awareness programs after study completion. The SPIRIT 2025 reporting guidelines were used in reporting this study protocol (Checklist 1).

XSL•FO
RenderX

Acknowledgments

The authors acknowledge the Indian Council of Medical Research for getting this protocol reviewed by an expert committee and supporting this study financially. We also thank Dabur India Limited for providing the active drug and placebo. We are grateful to all participants for giving consent and actively participating in the study. Generative artificial intelligence tools were not used for any portion of this manuscript.

Funding

The study is financially supported by the Indian Council of Medical Research (grant 5/4-5/3/7/ Neuro/2022-NCD-1). Dabur India Limited provided the active drug and placebo. The funding agency and the drug provider had no role in the study design, data collection, analysis, interpretation, decision to publish, or manuscript preparation.

Data Availability

Deidentified individual-level data sets will be available upon reasonable request to the corresponding author following publication of the study results.

Authors' Contributions

AG, GS, and RB conceptualized the study. A Dwivedi, A Dubey, AA, HPS, and YK collected and analyzed data. HN and NN helped with methodology, randomization, allocation concealment, and designing the statistical analysis plan. AG supervised the data collection, analysis, and interpretation. A Dwivedi wrote the first draft. All authors critically read and revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary tables.
[DOCX File, 46 KB - resprot_v15i1e82891_app1.docx]

Checklist 1
SPIRIT 2025 checklist.
[DOCX File, 24 KB - resprot_v15i1e82891_app2.docx]

Peer Review Report 1
Peer-review report from the Indian Council of Medical Research.
[PDF File, 1838 KB - resprot_v15i1e82891_app3.pdf]

References

1. Duff K. Mild cognitive impairment: quantifying a qualitative disorder. Neurol Clin 2024 Nov;42(4):781-792. [doi: 10.1016/j.ncl.2024.05.007] [Medline: 39343474]
2. Gross AL, Nichols E, Angrisani M, et al. Prevalence of DSM-5 mild and major neurocognitive disorder in India: results from the LASI-DAD. PLoS ONE 2024;19(2):e0297220. [doi: 10.1371/journal.pone.0297220] [Medline: 38324518]
3. Song WX, Wu WW, Zhao YY, et al. Evidence from a meta-analysis and systematic review reveals the global prevalence of mild cognitive impairment. Front Aging Neurosci 2023;15:1227112. [doi: 10.3389/fnagi.2023.1227112] [Medline: 37965493]
4. Portet F, Ousset PJ, Visser PJ, et al. Mild cognitive impairment (MCI) in medical practice: a critical review of the concept and new diagnostic procedure. Report of the MCI Working Group of the European Consortium on Alzheimer's Disease. J Neurol Neurosurg Psychiatry 2006 Jun;77(6):714-718. [doi: 10.1136/jnnp.2005.085332] [Medline: 16549412]
5. Csukly G, Sirály E, Fodor Z, et al. The differentiation of amnestic type MCI from the non-amnestic types by structural MRI. Front Aging Neurosci 2016;8:52. [doi: 10.3389/fnagi.2016.00052] [Medline: 27065855]
6. Beshir SA, Hussain N, Menon VB, Al Haddad AHI, Al Zeer RAK, Elnour AA. Advancements and challenges in antiamyloid therapy for Alzheimer's disease: a comprehensive review. Int J Alzheimers Dis 2024;2024:2052142. [doi: 10.1155/2024/2052142] [Medline: 39081336]
7. Varadharajan A, Davis AD, Ghosh A, et al. Guidelines for pharmacotherapy in Alzheimer's disease – a primer on FDA-approved drugs. J Neurosci Rural Pract 2023;14(4):566-573. [doi: 10.25259/JNRP_356_2023] [Medline: 38059250]
8. Zhang X, Lian S, Zhang Y, Zhao Q. Efficacy and safety of donepezil for mild cognitive impairment: a systematic review and meta-analysis. Clin Neurol Neurosurg 2022 Feb;213:107134. [doi: 10.1016/j.clineuro.2022.107134] [Medline: 35078087]

XSL•FO
RenderX

9. Chaudhari KS, Tiwari NR, Tiwari RR, Sharma RS. Neurocognitive effect of nootropic drug Brahmi (Bacopa monnieri) in Alzheimer's disease. Ann Neurosci 2017 May;24(2):111-122. [doi: 10.1159/000475900] [Medline: 28588366]
10. Kulkarni R, Girish KJ, Kumar A. Nootropic herbs (Medhya Rasayana) in Ayurveda: an update. Pharmacogn Rev 2012 Jul;6(12):147-153. [doi: 10.4103/0973-7847.99949] [Medline: 23055641]
11. Uabundit N, Wattanathorn J, Mucimapura S, Ingkaninan K. Cognitive enhancement and neuroprotective effects of Bacopa monnieri in Alzheimer's disease model. J Ethnopharmacol 2010 Jan 8;127(1):26-31. [doi: 10.1016/j.jep.2009.09.056] [Medline: 19808086]
12. Le XT, Pham HTN, Do PT, et al. Bacopa monnieri ameliorates memory deficits in olfactory bulbectomized mice: possible involvement of glutamatergic and cholinergic systems. Neurochem Res 2013 Oct;38(10):2201-2215. [doi: 10.1007/s11064-013-1129-6] [Medline: 23949198]
13. Rai R, Singh HK, Prasad S. A special extract of Bacopa monnieri (CDRI-08) restores learning and memory by upregulating expression of the NMDA receptor subunit GluN2B in the brain of scopolamine-induced amnesic mice. Evid Based Complement Alternat Med 2015;2015:254303. [doi: 10.1155/2015/254303] [Medline: 26413117]
14. Kumar SS, Saraswathi P, Vijayaraghavan R. Effect of Bacopa monniera on cold stress induced neurodegeneration in hippocampus of wistar rats: a histomorphometric study. J Clin Diagn Res 2015 Jan;9(1):AF05-AF07. [doi: 10.7860/JCDR/2015/10199.5423] [Medline: 25737970]
15. Kamkaew N, Norman Scholfield C, Ingkaninan K, Taepavarapruk N, Chootip K. Bacopa monnieri increases cerebral blood flow in rat independent of blood pressure. Phytother Res 2013 Jan;27(1):135-138. [doi: 10.1002/ptr.4685] [Medline: 22447676]
16. Konar A, Gautam A, Thakur MK. Bacopa monniera (CDRI-08) upregulates the expression of neuronal and glial plasticity markers in the brain of scopolamine induced amnesic mice. Evid Based Complement Alternat Med 2015;2015:837012. [doi: 10.1155/2015/837012] [Medline: 26413129]
17. Simpson T, Pase M, Stough C. Bacopa monnieri as an antioxidant therapy to reduce oxidative stress in the aging brain. Evid Based Complement Alternat Med 2015;2015:615384. [doi: 10.1155/2015/615384] [Medline: 26413126]
18. Kapoor R, Srivastava S, Kakkar P. Bacopa monnieri modulates antioxidant responses in brain and kidney of diabetic rats. Environ Toxicol Pharmacol 2009 Jan;27(1):62-69. [doi: 10.1016/j.etap.2008.08.007] [Medline: 21783922]
19. Russo A, Izzo AA, Borrelli F, Renis M, Vanella A. Free radical scavenging capacity and protective effect of Bacopa monniera L. on DNA damage. Phytother Res 2003 Sep;17(8):870-875. [doi: 10.1002/ptr.1061] [Medline: 13680815]
20. Mathew M, Subramanian S. Evaluation of the anti-amyloidogenic potential of nootropic herbal extracts in vitro. Int J Pharm Sci Res 2012;3(11):4276-4280. [doi: 10.13040/IJPSR.0975-8232.3(11).4276-80]
21. Holcomb LA, Dhanasekaran M, Hitt AR, Young KA, Riggs M, Manyam BV. Bacopa monniera extract reduces amyloid levels in PSAPP mice. J Alzheimers Dis 2006 Aug;9(3):243-251. [doi: 10.3233/jad-2006-9303] [Medline: 16914834]
22. Limpeanchob N, Jaipan S, Rattanakaruna S, Phrompittayarat W, Ingkaninan K. Neuroprotective effect of Bacopa monnieri on beta-amyloid-induced cell death in primary cortical culture. J Ethnopharmacol 2008 Oct 30;120(1):112-117. [doi: 10.1016/j.jep.2008.07.039] [Medline: 18755259]
23. Kongkeaw C, Dilokthornsakul P, Thanarangsarit P, Limpeanchob N, Norman Scholfield C. Meta-analysis of randomized controlled trials on cognitive effects of Bacopa monnieri extract. J Ethnopharmacol 2014;151(1):528-535. [doi: 10.1016/j.jep.2013.11.008] [Medline: 24252493]
24. Peth-Nui T, Wattanathorn J, Muchimapura S, et al. Effects of 12-week Bacopa monnieri consumption on attention, cognitive processing, working memory, and functions of both cholinergic and monoaminergic systems in healthy elderly volunteers. Evid Based Complement Alternat Med 2012;2012:606424. [doi: 10.1155/2012/606424] [Medline: 23320031]
25. Basheer A, Agarwal A, Mishra B, et al. Use of Bacopa monnieri in the treatment of dementia due to Alzheimer disease: systematic review of randomized controlled trials. Interact J Med Res 2022 Aug 1;11(2):e38542. [doi: 10.2196/38542] [Medline: 35612544]
26. Minale G, Saesong T, Temkitthawon P, et al. Characterization of metabolites in plasma, urine and feces of healthy participants after taking Brahmi essence for twelve weeks using LC-ESI-QTOF-MS metabolomic approach. Molecules 2021 May 15;26(10):2944. [doi: 10.3390/molecules26102944] [Medline: 34063409]
27. Mai Y, Huang F, Mi H, et al. Metabolomics and lipidomics study on serum metabolite signatures in Alzheimer's disease and mild cognitive impairment. Neurotherapeutics 2025 Oct;22(6):e00756. [doi: 10.1016/j.neurot.2025.e00756] [Medline: 41006122]
28. Shao Y, Ouyang Y, Li T, et al. Alteration of metabolic profile and potential biomarkers in the plasma of Alzheimer's disease. Aging Dis 2020 Dec;11(6):1459-1470. [doi: 10.14336/AD.2020.0217] [Medline: 33269100]
29. Hajjar I, Liu C, Jones DP, Uppal K. Untargeted metabolomics reveal dysregulations in sugar, methionine, and tyrosine pathways in the prodromal state of AD. Alzheimers Dement (Amst) 2020;12(1):e12064. [doi: 10.1002/dad2.12064] [Medline: 32793799]
30. Weng WC, Huang WY, Tang HY, Cheng ML, Chen KH. The differences of serum metabolites between patients with early-stage Alzheimer's disease and mild cognitive impairment. Front Neurol 2019;10:1223. [doi: 10.3389/fneur.2019.01223] [Medline: 31824405]

XSL•FO
RenderX

31. François M, Karpe AV, Liu JW, et al. Multi-omics, an integrated approach to identify novel blood biomarkers of Alzheimer's disease. Metabolites 2022 Oct 6;12(10):949. [doi: 10.3390/metabo12100949] [Medline: 36295851]
32. Yilmaz A, Ustun I, Ugur Z, et al. A community-based study identifying metabolic biomarkers of mild cognitive impairment and Alzheimer's disease using artificial intelligence and machine learning. J Alzheimers Dis 2020;78(4):1381-1392. [doi: 10.3233/JAD-200305] [Medline: 33164929]
33. Tiwari P, Gupta A, Kaushik M, et al. Comprehensive metabolomics profiling reveals novel biomarkers and pathways for early detection of Alzheimer's disease. Brain Commun 2025;7(6):fcaf410. [doi: 10.1093/braincomms/fcaf410] [Medline: 41229615]
34. Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279. [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
35. Menon RN, Varghese F, Paplikar A, et al. Validation of Indian Council of Medical Research Neurocognitive Tool Box in diagnosis of mild cognitive impairment in India: lessons from a harmonization process in a linguistically diverse society. Dement Geriatr Cogn Disord 2020;49(4):355-364. [doi: 10.1159/000512393] [Medline: 33412549]
36. Tripathi R, Kumar K, Bharath S, Marimuthu P, Varghese M. Demographic effects on neuropsychological functions in older Indians. Dement Neuropsychol 2014;8(2):148-154. [doi: 10.1590/S1980-57642014DN83000007] [Medline: 29213896]
37. Understanding alcohol drinking patterns. National Institute on Alcohol Abuse and Alcoholism. URL: https://www.niaaa.nih.gov/alcohols-effects-health/alcohol-drinking-patterns [accessed 2026-01-27]
38. Mekala S, Paplikar A, Mioshi E, et al. Dementia diagnosis in seven languages: the Addenbrooke's Cognitive Examination-III in India. Arch Clin Neuropsychol 2020 Jul 24;35(5):528-538. [doi: 10.1093/arclin/acaa013] [Medline: 32188967]
39. Mathuranath PS, George A, Cherian PJ, Mathew R, Sarma PS. Instrumental activities of daily living scale for dementia screening in elderly people. Int Psychogeriatr 2005 Sep;17(3):461-474. [doi: 10.1017/s1041610205001547] [Medline: 16252378]
40. Bajpai S, Upadhyay A, Sati H, Pandey RM, Chaterjee P, Dey AB. Hindi version of Addenbrooke's Cognitive Examination III: distinguishing cognitive impairment among older Indians at the lower cut-offs. Clin Interv Aging 2020;15:329-339. [doi: 10.2147/CIA.S244707] [Medline: 32184582]
41. Naz S, Gallart-Ayala H, Reinke SN, et al. Development of a liquid chromatography-high resolution mass spectrometry metabolomics method with high specificity for metabolite identification using all ion fragmentation acquisition. Anal Chem 2017 Aug 1;89(15):7933-7942. [doi: 10.1021/acs.analchem.7b00925] [Medline: 28641411]
42. Kumar A, Kumar Y, Sevak JK, Kumar S, Kumar N, Gopinath SD. Metabolomic analysis of primary human skeletal muscle cells during myogenic progression. Sci Rep 2020 Jul 16;10(1):11824. [doi: 10.1038/s41598-020-68796-4] [Medline: 32678274]
43. Carrick J, Cheung SC, Foxe D, Piguet O. Interpreting Addenbrooke's Cognitive Examination-III scores in dementia: performance distributions and clinically meaningful change. Eur J Neurol 2025 Jul;32(7):e70257. [doi: 10.1111/ene.70257] [Medline: 40583793]
44. McPhee GM, Downey LA, Wesnes KA, Stough C. The neurocognitive effects of Bacopa monnieri and cognitive training on markers of brain microstructure in healthy older adults. Front Aging Neurosci 2021;13:638109. [doi: 10.3389/fnagi.2021.638109] [Medline: 33692683]
45. Cicero AF, Bove M, Colletti A, et al. Short-term impact of a combined nutraceutical on cognitive function, perceived stress and depression in young elderly with cognitive impairment: a pilot, double-blind, randomized clinical trial. J Prev Alzheimers Dis 2017;4(1):12-15. [doi: 10.14283/jpad.2016.10] [Medline: 29188854]
46. Raghav S, Singh H, Dalal PK, Srivastava JS, Asthana OP. Randomized controlled trial of standardized Bacopa monniera extract in age-associated memory impairment. Indian J Psychiatry 2006 Oct;48(4):238-242. [doi: 10.4103/0019-5545.31555] [Medline: 20703343]
47. Roodenrys S, Booth D, Bulzomi S, Phipps A, Micallef C, Smoker J. Chronic effects of Brahmi (Bacopa monnieri) on human memory. Neuropsychopharmacology 2002 Aug;27(2):279-281. [doi: 10.1016/S0893-133X(01)00419-5] [Medline: 12093601]
48. Sathyanarayanan V, Thomas T, Einöther SJL, Dobriyal R, Joshi MK, Krishnamachari S. Brahmi for the better? New findings challenging cognition and anti-anxiety effects of Brahmi (Bacopa monniera) in healthy adults. Psychopharmacology (Berl) 2013 May;227(2):299-306. [doi: 10.1007/s00213-013-2978-z] [Medline: 23354535]
49. Pravina K, Ravindra KR, Goudar KS, et al. Safety evaluation of BacoMind in healthy volunteers: a phase I study. Phytomedicine 2007 May;14(5):301-308. [doi: 10.1016/j.phymed.2007.03.010] [Medline: 17442556]
50. Stough C, Downey LA, Lloyd J, et al. Examining the nootropic effects of a special extract of Bacopa monniera on human cognitive functioning: 90 day double-blind placebo-controlled randomized trial. Phytother Res 2008 Dec;22(12):1629-1634. [doi: 10.1002/ptr.2537] [Medline: 18683852]
51. Prabhakar S, Vishnu VY, Modi M, et al. Efficacy of Bacopa monnieri (Brahmi) and donepezil in Alzheimer's disease and mild cognitive impairment: a randomized double-blind parallel phase 2b study. Ann Indian Acad Neurol 2020;23(6):767-773. [doi: 10.4103/aian.AIAN_610_19] [Medline: 33688125]

XSL•FO
RenderX

52. Valotto Neto LJ, Reverete de Araujo M, Moretti Junior RC, et al. Investigating the neuroprotective and cognitive-enhancing effects of Bacopa monnieri: a systematic review focused on inflammation, oxidative stress, mitochondrial dysfunction, and apoptosis. Antioxidants (Basel) 2024 Mar 25;13(4):393. [doi: 10.3390/antiox13040393] [Medline: 38671841]

Abbreviations

ACE-III: Addenbrooke's Cognitive Examination III
AD: Alzheimer disease
aMCI: amnestic mild cognitive impairment
LC-MS: liquid chromatography-mass spectrometry
MCI: mild cognitive impairment
REDCap: Research Electronic Data Capture
SAE: serious adverse effect

Edited by A Schwartz; submitted 23.Aug.2025; this is a non–peer-reviewed article;accepted 31.Dec.2025; published 25.Feb.2026.

Please cite as:
Dwivedi A, Anjali A, Narzari H, Kumar Y, Sharma HP, Dubey A, Nilima N, Rajan R, Bhushan Singh M, Vishnu VY, Bhatia R, Sharma G, Gupta A
Efficacy of Bacopa monnieri (Linn.) on Cognitive Function and Alterations in Blood Metabolites in Patients With Amnestic Mild Cognitive Impairment and Early Alzheimer Disease: Protocol for an Exploratory Double-Blind, Randomized, Placebo-Controlled Trial
JMIR Res Protoc 2026;15:e82891
URL: https://www.researchprotocols.org/2026/1/e82891
doi:10.2196/82891

© Abhilasha Dwivedi, Anjali Anjali, Hina Narzari, Yashwant Kumar, Hanuman Prasad Sharma, Aditi Dubey, Nilima Nilima, Roopa Rajan, Mamta Bhushan Singh, Venugopalan Y Vishnu, Rohit Bhatia, Gautam Sharma, Anu Gupta. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effectiveness of Low-Dose Atropine Combined With Bright Light Therapy for Controlling Myopic Eye Growth in Schoolchildren: Study Protocol for a Randomized Controlled Trial

Dennis Yan-Yin Tse[1,2,3], PhD; Martin Ming-Leung Ma[3], PhD; Ying Hon[3,4], PhD; Lai Ming Ho[5], PhD; Wai Ching Lam[4,6], MD, FRCSC; Kendrick Co Shih[6], FCOphth HK; Xiujuan Zhang[6], PhD; Desmond Man Kit Lam[2,3], BSc; Andy Chi Wai Kong[3], MSc; Thomas Chuen Lam[1,3], PhD; Carly Siu Yin Lam[3], PhD; Tina Jinxiao Lian[1,2,3], PhD; Ian Morgan[7], PhD; Chi Ho To[2,3], PhD; Rachel Ka Man Chun[1,2,3], PhD; Christopher Kai Shun Leung[6], MD, MB ChB

[1]Research Centre for SHARP Vision, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]InnoHK, Centre for Eye and Vision Research, 17W Hong Kong Science Park, Hong Kong, China (Hong Kong)

[3]Centre for Myopia Research, School of Optometry, The Hong Kong Polytechnic University, 11 Yuk Choi Rd, Hung Hom, Kowloon, China (Hong Kong)

[4]Vancouver Coastal Health Research Institute, Vancouver, Canada

[5]School of Public Health, The University of Hong Kong, Hong Kong, China (Hong Kong)

[6]Department of Ophthalmology, The University of Hong Kong, Hong Kong, China (Hong Kong)

[7]Research School of Biology, Australian National University, Canberra, Australia

Corresponding Author:
Dennis Yan-Yin Tse, PhD
Research Centre for SHARP Vision, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

Abstract

Background: Myopia is increasingly prevalent worldwide, with projections indicating that nearly 50% of the global population may have myopia by 2050. This surge poses significant concerns due to its impact on vision and quality of life and its link to a range of blinding diseases, including myopic macular degeneration, glaucoma, and retinal detachment. Current pharmacologic and optical interventions offer limited effectiveness in slowing myopia progression, highlighting the urgent need for more effective treatments.

Objective: This study aims to examine the combined effect of bright light therapy and low-dose atropine on myopic progression.

Methods: This is a single-site, 2-arm, single-masked (examiner-masked) randomized controlled trial to compare the effectiveness of low-dose atropine alone and its combination with bright light therapy in retarding myopia progression. The study protocol has been approved by the institutional review boards of Hong Kong Polytechnic University (HSEARS 20180829002-05) and the University of Hong Kong and Hospital Authority Hong Kong West Cluster (UW 20-362). Schoolchildren with myopia aged 7 to 12 years who have not undergone any previous myopic control intervention will be recruited and randomly allocated into 2 groups (n=67 per group) after baseline measurements. Both groups will receive 0.01% atropine twice daily for 24 months. The combination treatment group will also receive a high-intensity lamp for bright light therapy. The primary and secondary outcome measures will be the changes in cycloplegic autorefraction in spherical equivalent refraction and axial length, respectively, measured every 6 months over 2 years from baseline.

Results: The project was funded in January 2019. The recruitment process started on March 21, 2023, and was completed on February 2, 2024. Data collection is expected to be completed in April 2026.

Conclusions: This study will provide new information on whether the combination of bright light therapy and low-dose atropine is more effective than atropine alone in slowing down myopia progression. It will also assess the effectiveness of low-dose atropine used twice daily. Combining bright light therapy and atropine could become a new treatment option if shown to be effective. New data on the effectiveness of using atropine twice daily might also expand available treatment options.

Trial Registration: ClinicalTrials.gov NCT04923841; https://clinicaltrials.gov/study/NCT04923841

International Registered Report Identifier (IRRID): DERR1-10.2196/90893

(JMIR Res Protoc 2026;15:e90893) doi:10.2196/90893

KEYWORDS

myopia control; bright light; atropine; refractive errors; myopia; clinical trial; vision

XSL•FO
RenderX

Introduction

Myopia has become increasingly prevalent worldwide, particularly among children in the Asian region. It has been predicted that almost 50% of the world's population will have myopia by 2050 [1]. This rapid rise in myopia is alarming not only because of its impact on vision and quality of life [2] but also because of its association with a higher risk of developing a range of blinding diseases later in life, including myopic macular degeneration, glaucoma, and retinal detachment [3].

Previous studies have identified pharmacologic and optical methods to retard myopia progression that only have a low to moderate effectiveness [4]. Thus, there is still an urgent need to explore innovative treatment approaches that are more capable than current methods. One promising approach is bright light therapy. Outdoor illumination in direct sunlight can reach up to 100,000 lux, whereas indoor illumination typically ranges from 100 to 500 lux [5]. One hypothesis explaining the rising prevalence of myopia among schoolchildren is that urban populations have embraced a lifestyle that limits exposure to bright outdoor light [6]. Research on animals (chicks, mice, and rhesus monkeys) has demonstrated that exposure to bright light ranging from 15,000 to 40,000 lux can entirely inhibit the development of myopia [7-10]. Observational clinical studies have indicated that an increase in time spent outdoors is associated with a more hyperopic refractive error and a lower incidence of juvenile-onset myopia [11,12]. In addition, a 1-year-long randomized clinical trial by Wu et al [13] has shown that an increase in outdoor activity can lead to less myopic shift and axial elongation in both children with and without myopia. The protective effect against myopia does not appear to be due to physical exercise, as indoor exercise does not inhibit myopia [14].

While it is clear that spending more time outdoors is beneficial for preventing and controlling myopia, the next clinically relevant question is whether artificially applied bright light indoors could achieve the same effect. Therefore, our team has designed a randomized controlled study to investigate the effect of bright light therapy (approximately 10,000 lux) on myopia [15]. This innovative approach will also be examined to explore the effect of its combination with low-dose atropine.

As no current treatment has been found to be able to completely halt myopia progression, another option is the potential use of a combination treatment. It is hypothesized that bright light therapy could provide additive efficacy to slow myopia progression when used in combination with low-dose atropine. Low-dose atropine has demonstrated a dose-dependent effect on slowing myopia with minimal side effects [16]. This combination is promising as there are preliminary data suggesting that these 2 therapies work on different mechanisms. While the exact mechanism for atropine's action is still unknown, it is thought to affect the development of myopia mainly by acting as a reversible competitive antagonist of acetylcholine muscarinic receptors that are found in many parts of the eye [17]. Regarding bright light therapy, its inhibitory effect on myopia is suggested to be mediated by the retinal dopaminergic pathways [18].

On the basis of these rationales, this clinical trial was designed to examine the combined effect of bright light therapy and low-dose atropine on myopic progression. The primary aim of this trial is to investigate whether the combination treatment with bright light therapy and low-dose atropine would be more effective than atropine alone. This trial will also assess the effectiveness of low-dose atropine in slowing myopia progression. Factors associated with the effectiveness of the treatments will also be identified. In addition, compliance and participant acceptability (particularly in terms of visual performance) of the combination therapy will be evaluated.

Methods

Study Design and Setting

This protocol describes a single-site, 2-arm, single-masked (examiner-masked) randomized controlled trial comparing the effectiveness of low-dose atropine alone and in combination with bright light therapy in retarding myopia progression. This manuscript is based on protocol version 1.6 (September 2022). This study will last for 2 years and is an auxiliary study to a previous study (phase 1 study) [15] conducted by the same team sharing the same study design except for the choice of treatment. The phase 1 study [15] will consist of 1 control group and 2 treatment groups, whereas this phase 2 study will consist of 2 treatment groups only. This paper will only describe the design of the phase 2 study, and details of the intervention for each group are shown in Figure 1 [15]. The investigation will be separated into 2 phases due to logistical issues. Participants from the phase 1 study will not continue into the phase 2 study. This study consists of 2 arms: atropine and atropine combined with bright light therapy. Participants will be randomized to the 2 treatment groups at a ratio of 1:1 using an age-stratified simple randomization. Data from the 2 phase 2 treatment groups will be compared with each other, and comparisons with phase 1 data will be conducted for exploratory analysis. Identical inclusion criteria will be used for both studies to minimize potential selection bias. The application of consistent methodologies and statistical approaches will enable meaningful comparisons across interventions.

XSL•FO
RenderX

Figure 1. Intervention groups in the phase 1 and phase 2 studies. The protocol for the phase 1 study has been published elsewhere [15]. This paper only describes the design of the phase 2 study.

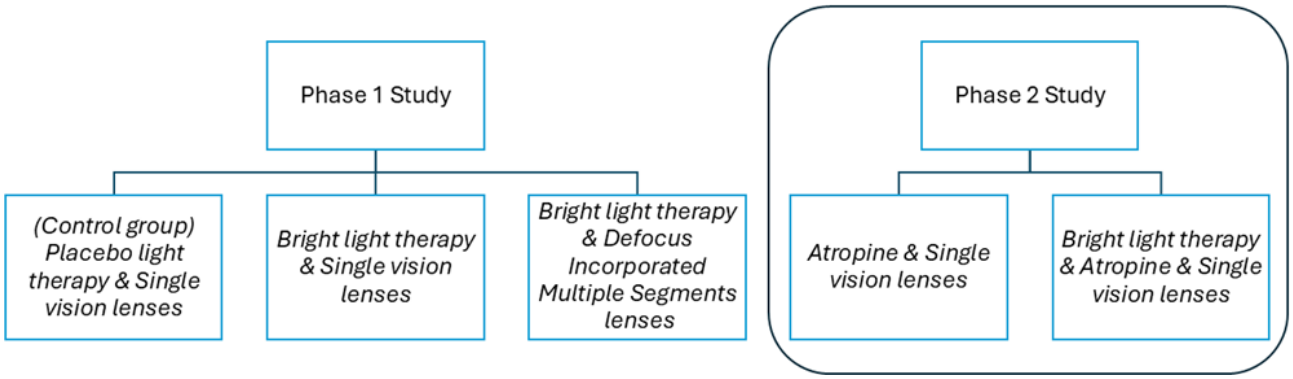

All data collection will take place at the Optometry Research Clinic of Hong Kong Polytechnic University or the HKU Eye Centre, Department of Ophthalmology, University of Hong Kong. This trial protocol is reported in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [19].

Ethical Considerations

This study protocol meets the tenets of the Declaration of Helsinki and has been approved by the institutional review boards of Hong Kong Polytechnic University (HSEARS 20180829002-05) and the University of Hong Kong and Hospital Authority Hong Kong West Cluster (UW 20-362). Written consent from the participants and their parents and assent from the participants will be obtained by examiners before any study procedures are performed. The study was prospectively registered at ClinicalTrials.gov (NCT04923841) on June 11, 2021. All procedures will adhere to the good clinical practice guidelines. Study personnel, including administrative staff and investigators, will complete the good clinical practice certification course. Patients and members of the public were not involved in the design, conduct, reporting, or dissemination planning of this research. After publication, parents or guardians and participants will be informed of the study results via a newsletter. This paper does not contain any individual identifiable data.

Sample Size Calculation

The sample size is estimated based on the 2-year change in refractive error in the 0.01% atropine treatment group (−0.64, –0.44 to +0.44 D) from a previous myopia control study [20]. Assuming that, over 2 years, combination treatment is more effective than 0.01% atropine alone by 0.25 D, which is considered a clinically significant change, 56 participants per group will be required. This sample size will enable the demonstration of a statistically significant difference between the 2 groups at a 5% significance level with 85% power. The sample size for each study arm will be adjusted to 67 per arm assuming a 20% loss to follow-up. As a result, a total sample size of 134 participants will be recruited for this study.

Inclusion Criteria

Children meeting the following criteria will be included:

- Children who are Hong Kong Chinese
- Myopia (spherical equivalent) of −0.75 D to −5.00 D in both eyes
- Age at enrollment of 7 to 12 years
- Astigmatism and anisometropia of 1.50 D or less
- Spectacle-corrected monocular visual acuity of 0.04 logMAR or better in both eyes
- Normal binocular function
- Parents' understanding and acceptance of the random allocation of grouping and masking
- Ability to wear the prescribed spectacles, apply eye drops, and undergo light therapy daily

Exclusion Criteria

Children meeting the following criteria will be excluded:

- Eye disease or binocular vision problems (eg, strabismus, amblyopia, oculomotor nerve palsies, or corneal disease)
- Previous intraocular or corneal surgery
- Ocular or systemic disease that may affect vision function or development (eg, endocrine, cardiac, and respiratory diseases; diabetes; or Down syndrome)
- Allergy to atropine
- Previous or current use of gas-permeable, soft bifocal, or orthokeratology contact lenses or bifocal, progressive addition lenses; specific myopic control spectacle wear; or use of atropine or pirenzepine (longer than 1 month)
- Previous or current participation in myopia control studies

Randomization and Masking

Participants will be randomized to the 2 treatment groups at a ratio of 1:1 using stratified simple randomization. The randomization list will be generated using a random number generator in a Microsoft Excel spreadsheet. Participants will be stratified into 2 age groups (7-9 years and 10-12 years). This is a single-masked study in which examiners are masked to the treatment allocation. Allocation concealment will be ensured using sequentially numbered, opaque, sealed envelopes. The envelopes will be opened by the designated unmasked investigators in numerical order only after the participants are enrolled. To maintain masking, access to the randomization list will be limited to the trial steering committee; study management committee; statistician; and unmasked investigators designated to handle randomization, allocation, and delivery of group assignments. These individuals will not be involved in participant enrollment or data collection.

XSL•FO
RenderX

Procedures

Mass emails to the staff of Hong Kong Polytechnic University, advertising posts on social media (Facebook and Instagram), and referrals from parents or guardians who are already participating in the study will be used to recruit participants. Parents or guardians and their children will be invited to the Optometry Research Clinic for a screening visit to confirm their eligibility. Informed consent and assent will be obtained upon confirmation of fulfillment of the inclusion criteria.

Successfully enrolled participants will be randomly allocated to 1 of the 2 arms after baseline measurement. Delivery of spectacles and a light therapy lamp will be arranged 2 weeks after the baseline visit. Training will be provided to participants and parents or guardians on how to set up and use the light therapy lamp at home. The participants in both treatment groups will be asked to return 1 month after the intervention for aftercare appointments to ensure that there are no allergic reactions to atropine or adverse events associated with the light therapy lamp. Participants will also be asked to return for regular follow-up visits with the study ophthalmologist for dispensing of 0.01% atropine.

Data collection will be conducted every 6 months from the baseline visit up to 24 months. Baseline measurements will include demographic data; medical and medication history; visual acuity; tests for strabismus, autorefraction, and keratometry; pupil size; amplitude of accommodation; accommodative lag; subjective and objective refraction with and without cycloplegia; intraocular pressure; choroidal thickness; external and internal ocular health; peripheral refraction; and axial length via the IOLMaster 500 (ZEISS) and other ocular biometry (eg, corneal curvature, corneal thickness, anterior chamber depth, and lens thickness) via the Lenstar LS 900 (Haag-Streit AG). In addition, questionnaires on visual performance, daily activities, and compliance will be administered at the 6-, 12-, 18-, and 24-month visits. A SPIRIT schedule of the study is shown in Table 1.

Table . Schedule of enrollment, interventions, and assessments.

Timepoint	Baseline	1 month	6 months	12 months	18 months	24 months
Enrollment						
Allocation	✓					
Interventions						
Single treatment: atropine & single vision lenses		✓	✓	✓	✓	✓
Combination treatment: bright light therapy & atropine & single vision lenses		✓	✓	✓	✓	✓
Assessments						
Primary outcome: cycloplegic objective refraction	✓		✓	✓	✓	✓
Secondary outcome: axial length	✓		✓	✓	✓	✓
Other outcomes: intraocular pressure, accommodative functions, pupil size, and peripheral refraction	✓		✓	✓	✓	✓
Safety outcomes: treatment compliance, visual performance, and adverse events		✓	✓	✓	✓	✓

Interventions

Eligible participants will be randomized to 1 of 2 groups: (1) bright light therapy and atropine group (high-intensity lamp for bright light therapy, 0.01% atropine eye drops twice a day [1 drop in the morning and 1 drop at night before bedtime], and single vision spectacle lenses) and (2) atropine group (0.01% atropine eye drops twice a day [1 drop in the morning and 1 drop at night before bedtime] and single vision spectacle lenses). Participants will stay in these groups for 24 months.

Bright Light Therapy

During bright light therapy, participants are exposed to an LED lamp on a daily basis as they perform near visual tasks (eg, homework, reading, or using a computer). Each participant will be provided with a lamp (model DL93011, Carex Health Brands;

XSL•FO
RenderX

Figure 2) commonly used in previous psychiatric studies [21]. It emits 4000-K glare-free white light rated at 10,000 lux at 30 cm from the screen to the cornea. The spectral radiance of the lamp can be found in Figure 3. The polycarbonate diffuser of the unit blocks UV light with a cutoff of approximately 400 nm. Participants will use the lamp at home and be given standardized verbal instructions, including viewing distance, viewing angle, and duration of use. The lamp head should be angled at approximately 15 degrees from vertical, with the center of the lamp head located 30 to 55 cm from the participant's eye. Participants will be asked to use the unit for at least 45 minutes per day (not more than 120 minutes per day) and 7 days per week. This dosage is chosen because clinical trials using similar protocols (30-60 minutes per day) have reported effective remission of depressive symptoms and well-tolerated treatment [22]. Participants may choose when to use the lamp during the day and to skip the treatment on days when they have already had more than 2 hours of outdoor activities under sunlight. To avoid causing insomnia, participants will be advised not to use the lamp within 3 hours prior to bedtime. They will also be trained on how to use the unit properly. There will be a network-enabled device attached to the lamp to monitor and record the duration of treatment, working distance, and ambient light illuminance. To promote compliance, researchers will make phone calls to participants (or their guardians) regarding lamp use during the first month of the trial and to provide guidance when needed. All lamps will be returned to investigators at the end of the study.

XSL•FO
RenderX

Figure 2. Light therapy lamp with monitoring sensor attached underneath.



XSL•FO
RenderX

Figure 3. Spectral radiance of the light therapy lamp measured using a spectroradiometer (model S-R3; Topcon Technohouse Co).

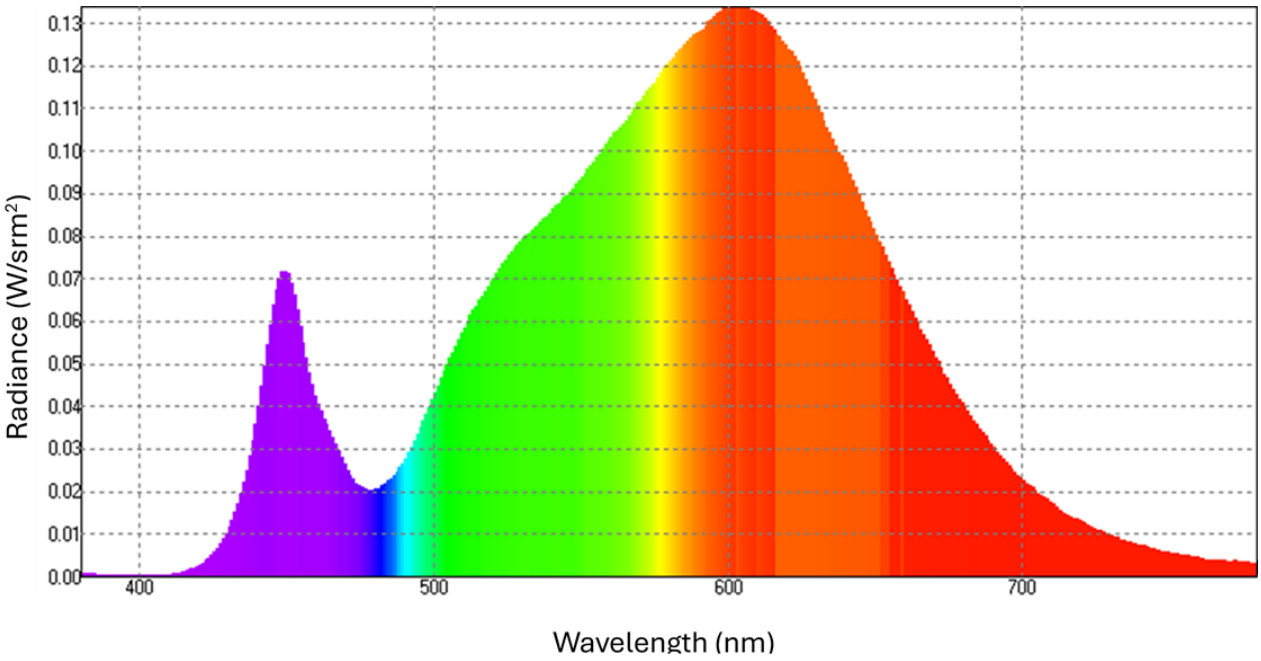

Atropine

Sterile, unpreserved, unidose 0.01% atropine eye drops will be provided to parents or guardians for administering to the participants as 1 drop in the morning and 1 drop at night before bedtime (twice a day) for 24 months. Each vial will be disposed of daily to prevent microbial contamination. An ophthalmologist will perform an ocular health assessment at the atropine dispensing visit, 1 month later, and every 6 months thereafter.

Single Vision Spectacle Lenses

The single vision spectacles, prescribed according to the cycloplegic subjective refraction, will be provided to participants in both treatment groups. The lenses are made of plastic, with a refractive index of 1.6. Participants will be instructed to wear the spectacles full time. At each follow-up visit, an update in prescription will be made if either of these criteria is fulfilled: (1) an increase in myopia of 0.50 D (spherical overrefraction) or more or (2) habitual visual acuity equal to or worse than 0.20 logMAR.

Outcomes and Measurements

The primary and secondary outcome measures will be changes in cycloplegic autorefraction in spherical equivalent refraction and axial length, respectively, measured every 6 months over 2 years from baseline. Other outcomes include peripheral refraction, pupil size, choroidal thickness, amplitude of accommodation, accommodative lag, corneal curvature, noncontact intraocular pressure, visual performance questionnaire, and treatment compliance (assessed via the frequency and/or duration of the use of atropine and bright light therapy).

Compliance

Compliance with spectacle wear will be monitored through questionnaires at 6-month intervals. Compliance with light therapy will be monitored through 2 means. First, participants or their parents or guardians will be instructed to record the treatment duration daily using a mobile app. Second, a cloud-enabled monitoring sensor attached to the bottom of each light therapy lamp (Figure 2) will be used to collect data on background lighting, working distance, and use time for each participant. Compliance with atropine eye drops will be monitored through a self-reporting log through the aforementioned mobile app. Additionally, custom-built medicine boxes with cloud-enabled sensors will be randomly distributed to 20 participants in the atropine group. The sensors will collect data on the time and frequency of the actions of opening and closing the boxes, providing another measure of compliance.

Compliance data will be analyzed monthly, and participants with poor compliance will be reminded to follow the treatment modality via WhatsApp messages.

Safety Considerations

The side effects of bright light therapy have been extensively researched in both psychiatric patients and healthy adults [22,23]. No ocular abnormalities have been found with either short- or long-term bright light therapy [24]. Some commonly reported side effects include headaches, eye strain, blurred vision, and nausea, but these were mild and temporary [23,25]. Bright light therapy is generally considered a safe and well-tolerated treatment [22].

A meta-analysis of the safety of low-dose atropine revealed that no major adverse events were detected in all 4 studies included. Side effects of atropine include photophobia, blurred near vision, lower accommodation amplitude, and a larger pupil size [26]. Overall, low-dose atropine is considered a safe treatment for myopia control [26,27].

The visual functions, ocular health, and visual and nonvisual symptoms of both treatment groups will be strictly monitored at every follow-up visit. Adverse events, causally related or not to the study intervention, will be recorded and classified by

XSL•FO
RenderX

severity. These include but are not limited to reduced visual acuity, visual disturbances, eye strain, eye infection or allergic reaction, photophobia, headache, or nausea. Participants will receive appropriate medical care for any clinically significant adverse events. Any serious adverse events will be reported to the Institutional Review Board of Hong Kong Polytechnic University within 48 hours. All adverse events will be followed up on until adequate resolution. An investigator may withdraw a participant from the study if any clinical adverse event or situation occurs such that continued participation in the study would not be in their best interest.

Data Management Plan

The source documents consist of signed consent and assent forms, printouts of ophthalmic investigations (eg, autorefractor and tonometer), and electronic case report forms (eCRFs). All paper documents will be stored in lockable cabinets. All study data will be recorded on standardized eCRFs and stored in an online REDCap (Research Electronic Data Capture; Vanderbilt University) [28] database, in which the server is built and protected under a secure campus network. Access is restricted to authorized study investigators, each with a unique account, ID number, and password. Only specific forms can be viewed and modified based on the investigators' assigned roles. Anonymized participant identifiers will be used for any data entry and documents uploaded to REDCap. An unmasked investigator must review and approve each eCRF after all data have been entered. An independent monitor will check the presence and correct date of all signed consent forms. The monitor will sample approximately 10% of all randomized participants to check correct data collection and data entry for the key points of this study.

Oversight and Monitoring

The trial steering committee comprises DYYT (chair), CHT (co–principal investigator), LMH (co–principal investigator), and Professor Chea-su Kee (departmental research committee chair of the School of Optometry, Hong Kong Polytechnic University). The study management committee comprises DYYT (principal coordinator), CKSL (co–principal investigator), CSYL (co–principal investigator), and Professor Victor Chi-Pang Woo (external member from Hong Kong Polytechnic University). The day-to-day coordination of the trial will be managed by the study team at the School of Optometry, Hong Kong Polytechnic University, and the Department of Ophthalmology, University of Hong Kong.

Statistical Analyses

The latest version of the SPSS software (IBM Corp) will be used for statistical analyses. This study will use the same statistical approach as that used in the previous phase 1 study [15]. Data from the right eyes will be used for analyses. The normality of the data will be examined using the Kolmogorov-Smirnov test. Demographic and optometric data will be presented as means and SDs or as numbers and percentages as appropriate. The chi-square test (or, if appropriate, the Fisher exact test) and one-way ANOVA will be performed to compare baseline between-group differences in demographics (age and sex). Changes in primary, secondary, and other outcomes at different time points will be analyzed using a mixed-effects model for repeated measures, which includes fixed effects for the treatment group, time, the interaction between group and time, and covariates including age and baseline myopia. If statistically significant differences in baseline demographic characteristics between arms are found, these will also be adjusted for.

The progression of myopia over 2 years will be calculated as the difference in cycloplegic autorefraction in spherical equivalent refraction and axial length between the baseline and the 2-year visit. The relative treatment effect in each arm will be determined by dividing the difference in myopia progression between the treatment group and the control group in the phase 1 study [15] by myopia progression in the control group and then multiplying by 100%. Additionally, the absolute treatment effect [29] will be presented to facilitate comparison with other studies. Multiple regression analysis will be used to identify factors associated with myopia control effectiveness.

All analyses will follow the intention-to-treat approach. Missing values of outcome variables and covariates will be replaced using a multiple imputation procedure with 10 sets of imputations assuming missingness at random. Sensitivity analyses for the primary outcomes include (1) complete-case analysis and (2) per-protocol analysis. All analyses will be 2-tailed, with a significance level of 5%.

Results

The project was funded in January 2019. The recruitment process started on March 21, 2023, and was completed on February 2, 2024. Data collection is expected to be completed in April 2026. The results are expected to be published in winter 2026.

Discussion

This 2-arm randomized controlled trial will be the first to investigate the combined effect of bright light therapy and low-dose atropine on myopia progression in schoolchildren. We hypothesized that bright light therapy would have an additive effect on slowing myopia progression when combined with low-dose atropine. If proven effective, this combination treatment could represent a novel therapeutic strategy. Combination treatment, also known as polytherapy, has emerged as a new approach in recent years. The rationale for its use is that there may be an additive effect on treatment efficacy if the treatment of choice affects myopia pathogenesis through distinct mechanisms. A recent meta-analysis suggests that combination treatment is the most effective intervention for myopia control [30]. Both Tan et al [31] and Kinoshita et al [32] have shown that the combination of 0.01% atropine and orthokeratology is more effective than orthokeratology alone in slowing axial elongation. As the proposed myopia control mechanisms differ between atropine and bright light therapy, this randomized clinical trial is designed to investigate whether combining these 2 promising therapies results in an additive effect.

Additionally, bright light therapy has not yet been studied for its effectiveness in controlling myopia progression. Therefore,

XSL•FO
RenderX

our results will contribute valuable insights to the literature regarding its potential efficacy. Another feature of this study is the administration of low-dose atropine twice daily. At the time of writing, all other randomized clinical trials [20,33-40] using low-dose atropine to control or prevent myopia only administered the treatment once daily. It has been repeatedly shown that 0.01% atropine is a safe and well-tolerated treatment but is limited in its effectiveness [34]. Increasing the concentration of atropine would lead to a stronger myopia control effect [34], but unfortunately, both the risk of adverse effects [34] and the extent of the rebound effect are dose dependent [38]. Administering atropine twice daily provides a more prolonged presence of the medication, which could result in a stronger myopia control effect without significantly increasing the risk of adverse events.

A recent 1-year–long randomized clinical trial by He et al [41] provides supporting evidence for this hypothesis. In their study, the axial elongation and spherical equivalent refraction progression in the twice-daily group (n=36) were 34% and 53% lower, respectively, than in the once-daily group (n=34). There were no clinically significant differences between the 2 groups in safety parameters, including accommodation amplitude, pupil diameter, and intraocular pressure. The study proposed in this paper, with a larger sample size and longer follow-up, will provide additional insights into optimizing the balance between efficacy and safety in low-dose atropine treatment for myopia control.

The strengths of this study include its prospective, randomized design and the use of masked examiners. The primary limitation of this study is the lack of masking of participants. The study duration is only 24 months due to budget constraints, so longer-term results will not be available. In addition, while the effectiveness of bright light therapy is examined in this study, further research is still needed to determine the optimal light level. Whether the benefit, if any, would apply to patients with different characteristics also remains to be assessed.

The results of this study will be disseminated at scientific conferences and in peer-reviewed, indexed journals in the fields of optometry and ophthalmology.

Acknowledgments

The authors acknowledge Shenyang Xingqi Pharmaceutical Co., Ltd, for donating the atropine eye drops used in this trial. The authors acknowledge Mr Billy Cheung, Ms Pauline Cheung, Ms Katrina Cheung, Dr Jason Lau, Ms Sandy Au, and Mr Leo Lee for data collection support; Ms Winky Wong, Ms Sze Wai Fong, and Ms Cathy Cheung for liaison with parents; Mr Anson Lau for technical support with experimental tools and database management; and Dr Maureen Boost for English-language editing. No generative artificial intelligence tools were used in the preparation of this manuscript.

Funding

The study is supported by the Research Grants Council Research Impact Fund (grant R5032-18), University Grants Committee Research Matching Grant Scheme (grants U-ZEYE and U-ZEZ8), and Hong Kong Polytechnic University Grants (grant BDDR), and the InnoHK scheme from the Hong Kong Special Administrative Region Government. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

Data sharing is not applicable to this article as this manuscript is the study protocol without pilot or preliminary data. Deidentified research data will be made publicly available upon completion and publication of the study.

Authors' Contributions

Conceptualization: DYYT, RKMC, CSYL, CHT, LMH, WCL, TCL, TJXL, IM, CKSL
Formal analysis: DYYT, RKMC, CHT, LMH, TJXL, MMLM, YH
Funding acquisition: CHT, DYYT, CKSL, RKMC, LMH, WCL, TCL, CSYL, TJXL, IM
Investigation: MMLM, DMKL, YH, KCS, XJZ
Project administration: YH, RKMC
Software: ACWK, DMKL
Supervision: DYYT, CKSL
Validation: MMLM, YH, LMH
Visualization: RKMC, DYYT
Writing—original draft: MMLM
Writing—review and editing: MMLM, YH, DYYT, RKMC, CYSL, CKSL, and MB
RKMC, and CKSL are co-corresponding authors of this work.

Conflicts of Interest

DYYT, RKMC, CSYL, and CHT have received research funding from Hoya Corporation. DYYT, CSYL and CHT have received royalty payments from Hoya Corporation. DYYT, CSYL, and CHT hold patents on myopia treatment using spectacle lenses, contact lenses, and pharmaceuticals. All other authors declare no other conflicts of interest.

XSL•FO
RenderX

Checklist 1
Peer review report by the Research Impact Fund Committee, University Grants Committee, Research Grants Council (Hong Kong).
[PDF File, 970 KB - resprot_v15i1e90893_app1.pdf]

References

1. Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology 2016 May;123(5):1036-1042. [doi: 10.1016/j.ophtha.2016.01.006] [Medline: 26875007]
2. Sankaridurg P, Tahhan N, Kandel H, et al. IMI impact of myopia. Invest Ophthalmol Vis Sci 2021 Apr 28;62(5):2. [doi: 10.1167/iovs.62.5.2] [Medline: 33909036]
3. Haarman AEG, Enthoven CA, Tideman JWL, Tedja MS, Verhoeven VJM, Klaver CCW. The complications of myopia: a review and meta-analysis. Invest Ophthalmol Vis Sci 2020 Apr 9;61(4):49. [doi: 10.1167/iovs.61.4.49] [Medline: 32347918]
4. Lanca C, Pang CP, Grzybowski A. Effectiveness of myopia control interventions: a systematic review of 12 randomized control trials published between 2019 and 2021. Front Public Health 2023;11:1125000. [doi: 10.3389/fpubh.2023.1125000] [Medline: 37033047]
5. Norton TT, Siegwart JT Jr. Light levels, refractive development, and myopia--a speculative review. Exp Eye Res 2013 Sep;114:48-57. [doi: 10.1016/j.exer.2013.05.004] [Medline: 23680160]
6. Morgan IG, French AN, Ashby RS, et al. The epidemics of myopia: aetiology and prevention. Prog Retin Eye Res 2018 Jan;62:134-149. [doi: 10.1016/j.preteyeres.2017.09.004] [Medline: 28951126]
7. Ashby R, Ohlendorf A, Schaeffel F. The effect of ambient illuminance on the development of deprivation myopia in chicks. Invest Ophthalmol Vis Sci 2009 Nov;50(11):5348-5354. [doi: 10.1167/iovs.09-3419] [Medline: 19516016]
8. Smith EL III, Hung LF, Huang J. Protective effects of high ambient lighting on the development of form-deprivation myopia in rhesus monkeys. Invest Ophthalmol Vis Sci 2012 Jan 25;53(1):421-428. [doi: 10.1167/iovs.11-8652] [Medline: 22169102]
9. Chen S, Zhi Z, Ruan Q, et al. Bright light suppresses form-deprivation myopia development with activation of dopamine D1 receptor signaling in the ON pathway in retina. Invest Ophthalmol Vis Sci 2017 Apr 1;58(4):2306-2316. [doi: 10.1167/iovs.16-20402] [Medline: 28431434]
10. Cohen Y, Peleg E, Belkin M, Polat U, Solomon AS. Ambient illuminance, retinal dopamine release and refractive development in chicks. Exp Eye Res 2012 Oct;103:33-40. [doi: 10.1016/j.exer.2012.08.004] [Medline: 22960317]
11. Dharani R, Lee CF, Theng ZX, et al. Comparison of measurements of time outdoors and light levels as risk factors for myopia in young Singapore children. Eye (Lond) 2012 Jul;26(7):911-918. [doi: 10.1038/eye.2012.49] [Medline: 22562184]
12. Dirani M, Tong L, Gazzard G, et al. Outdoor activity and myopia in Singapore teenage children. Br J Ophthalmol 2009 Aug;93(8):997-1000. [doi: 10.1136/bjo.2008.150979] [Medline: 19211608]
13. Wu PC, Chen CT, Lin KK, et al. Myopia prevention and outdoor light intensity in a school-based cluster randomized trial. Ophthalmology 2018 Aug;125(8):1239-1250. [doi: 10.1016/j.ophtha.2017.12.011] [Medline: 29371008]
14. Guggenheim JA, Northstone K, McMahon G, et al. Time outdoors and physical activity as predictors of incident myopia in childhood: a prospective cohort study. Invest Ophthalmol Vis Sci 2012 May 14;53(6):2856-2865. [doi: 10.1167/iovs.11-9091] [Medline: 22491403]
15. Hon Y, Chun RKM, Cheung BKK, et al. Effectiveness of bright light therapy and combination with myopic defocus for controlling myopic eye growth in schoolchildren: study protocol for a randomised controlled trial (phase 1). BMJ Open Ophthalmol 2025 May 19;10(1):e002044. [doi: 10.1136/bmjophth-2024-002044] [Medline: 40389302]
16. Yam JC, Jiang Y, Tang SM, et al. Low-Concentration Atropine for Myopia Progression (LAMP) study: a randomized, double-blinded, placebo-controlled trial of 0.05%, 0.025%, and 0.01% atropine eye drops in myopia control. Ophthalmology 2019 Jan;126(1):113-124. [doi: 10.1016/j.ophtha.2018.05.029] [Medline: 30514630]
17. Upadhyay A, Beuerman RW. Biological mechanisms of atropine control of myopia. Eye Contact Lens 2020 May;46(3):129-135. [doi: 10.1097/ICL.0000000000000677] [Medline: 31899695]
18. Zhou X, Pardue MT, Iuvone PM, Qu J. Dopamine signaling and myopia development: what are the key challenges. Prog Retin Eye Res 2017 Nov;61:60-71. [doi: 10.1016/j.preteyeres.2017.06.003] [Medline: 28602573]
19. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomized trials. JAMA 2025 Aug 5;334(5):435-443. [doi: 10.1001/jama.2025.4486] [Medline: 40294593]
20. Lee SSY, Lingham G, Blaszkowska M, et al. Low-concentration atropine eyedrops for myopia control in a multi-racial cohort of Australian children: a randomised clinical trial. Clin Exp Ophthalmol 2022 Dec;50(9):1001-1012. [doi: 10.1111/ceo.14148] [Medline: 36054556]
21. Lam RW, Levitt AJ, Levitan RD, et al. Efficacy of bright light treatment, fluoxetine, and the combination in patients with nonseasonal major depressive disorder: a randomized clinical trial. JAMA Psychiatry 2016 Jan;73(1):56-63. [doi: 10.1001/jamapsychiatry.2015.2235] [Medline: 26580307]
22. Terman M, Terman JS. Light therapy for seasonal and nonseasonal depression: efficacy, protocol, safety, and side effects. CNS Spectr 2005 Aug;10(8):647-663. [doi: 10.1017/s1092852900019611] [Medline: 16041296]
23. Botanov Y, Ilardi SS. The acute side effects of bright light therapy: a placebo-controlled investigation. PLoS One 2013;8(9):e75893. [doi: 10.1371/journal.pone.0075893] [Medline: 24086658]

XSL•FO
RenderX

24. Gallin PF, Terman M, Remé CE, Rafferty B, Terman JS, Burde RM. Ophthalmologic examination of patients with seasonal affective disorder, before and after bright light therapy. Am J Ophthalmol 1995 Feb;119(2):202-210. [doi: 10.1016/s0002-9394(14)73874-7] [Medline: 7832227]
25. Kogan AO, Guilford PM. Side effects of short-term 10,000-lux light therapy. Am J Psychiatry 1998 Feb;155(2):293-294. [doi: 10.1176/ajp.155.2.293] [Medline: 9464216]
26. Gong Q, Janowski M, Luo M, et al. Efficacy and adverse effects of atropine in childhood myopia: a meta-analysis. JAMA Ophthalmol 2017 Jun 1;135(6):624-630. [doi: 10.1001/jamaophthalmol.2017.1091] [Medline: 28494063]
27. Yam JC, Jiang Y, Tang SM, et al. Low-Concentration Atropine for Myopia Progression (LAMP) study. Ophthalmology 2019 Jan;126(1):113-124. [doi: 10.1016/j.ophtha.2018.05.029]
28. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
29. Brennan NA, Toubouti YM, Cheng X, Bullimore MA. Efficacy in myopia control. Prog Retin Eye Res 2021 Jul;83:100923. [doi: 10.1016/j.preteyeres.2020.100923] [Medline: 33253901]
30. Zhang G, Jiang J, Qu C. Myopia prevention and control in children: a systematic review and network meta-analysis. Eye (Lond) 2023 Nov;37(16):3461-3469. [doi: 10.1038/s41433-023-02534-8] [Medline: 37106147]
31. Tan Q, Ng AL, Cheng GP, Woo VC, Cho P. Combined 0.01% atropine with orthokeratology in childhood myopia control (AOK) study: a 2-year randomized clinical trial. Cont Lens Anterior Eye 2023 Feb;46(1):101723. [doi: 10.1016/j.clae.2022.101723] [Medline: 35654683]
32. Kinoshita N, Konno Y, Hamada N, et al. Efficacy of combined orthokeratology and 0.01% atropine solution for slowing axial elongation in children with myopia: a 2-year randomised trial. Sci Rep 2020 Jul 29;10(1):12750. [doi: 10.1038/s41598-020-69710-8] [Medline: 32728111]
33. Lee SH, Tseng BY, Wang JH, Chiu CJ. Efficacy and safety of low-dose atropine on myopia prevention in premyopic children: systematic review and meta-analysis. J Clin Med 2024 Mar 5;13(5):1506. [doi: 10.3390/jcm13051506] [Medline: 38592670]
34. Hou P, Wu D, Nie Y, Wei H, Liu L, Yang G. Comparison of the efficacy and safety of different doses of atropine for myopic control in children: a meta-analysis. Front Pharmacol 2023;14:1227787. [doi: 10.3389/fphar.2023.1227787] [Medline: 37767401]
35. Cui C, Li X, Lyu Y, et al. Safety and efficacy of 0.02% and 0.01% atropine on controlling myopia progression: a 2-year clinical trial. Sci Rep 2021 Nov 15;11(1):22267. [doi: 10.1038/s41598-021-01708-2] [Medline: 34782708]
36. Hieda O, Hiraoka T, Fujikado T, et al. Efficacy and safety of 0.01% atropine for prevention of childhood myopia in a 2-year randomized placebo-controlled study. Jpn J Ophthalmol 2021 May;65(3):315-325. [doi: 10.1007/s10384-021-00822-y] [Medline: 33586090]
37. Wei S, Li SM, An W, et al. Safety and efficacy of low-dose atropine eyedrops for the treatment of myopia progression in Chinese children: a randomized clinical trial. JAMA Ophthalmol 2020 Nov 1;138(11):1178-1184. [doi: 10.1001/jamaophthalmol.2020.3820] [Medline: 33001210]
38. Yam JC, Zhang XJ, Zhang Y, et al. Three-Year Clinical Trial of Low-Concentration Atropine for Myopia Progression (LAMP) study: continued versus washout: phase 3 report. Ophthalmology 2022 Mar;129(3):308-321. [doi: 10.1016/j.ophtha.2021.10.002] [Medline: 34627809]
39. Yam JC, Zhang XJ, Zhang Y, et al. Effect of low-concentration atropine eyedrops vs placebo on myopia incidence in children: the LAMP2 randomized clinical trial. JAMA 2023 Feb 14;329(6):472-481. [doi: 10.1001/jama.2022.24162] [Medline: 36786791]
40. Sharma I, Das GK, Rohatgi J, Sahu PK, Chhabra P, Bhatia R. Low dose atropine in preventing the progression of childhood myopia: a randomised controlled trial. Curr Eye Res 2023 Apr;48(4):402-407. [doi: 10.1080/02713683.2022.2162925] [Medline: 36576170]
41. He M, Zhang Z, Tong L, et al. Twice versus once daily application of 0.01% atropine for myopia control in children: a randomised controlled trial. Eye (Lond) 2025 Dec;39(17):3105-3112. [doi: 10.1038/s41433-025-03964-2] [Medline: 41034350]

Abbreviations

eCRF: electronic case report form
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

XSL•FO
RenderX

Edited by J Sarvestan; submitted 08.Jan.2026; this is a non–peer-reviewed article;accepted 25.Feb.2026; published 13.Apr.2026.

Please cite as:
Tse DYY, Ma MML, Hon Y, Ho LM, Lam WC, Shih KC, Zhang X, Lam DMK, Kong ACW, Lam TC, Lam CSY, Lian TJ, Morgan I, To CH, Chun RKM, Leung CKS
Effectiveness of Low-Dose Atropine Combined With Bright Light Therapy for Controlling Myopic Eye Growth in Schoolchildren: Study Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e90893
URL: https://www.researchprotocols.org/2026/1/e90893
doi:10.2196/90893

© Dennis Yan-Yin Tse, Martin Ming-Leung Ma, Ying Hon, Lai Ming Ho, Wai Ching Lam, Kendrick Co Shih, Xiujuan Zhang, Desmond Man Kit Lam, Andy Chi Wai Kong, Thomas Chuen Lam, Carly Siu Yin Lam, Tina Jinxiao Lian, Ian Morgan, Chi Ho To, Rachel Ka Man Chun, Christopher Kai Shun Leung. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effectiveness and Implementation of Adapted Physical Activity Delivery Strategies for Older Adults Living With HIV in Ivory Coast: Protocol for a Type 2 Hybrid Randomized Controlled Trial

Nicolas Diadhiou[1,2,3,4*], MPH, MD; Nelly Assoumou[5], PhD; Joseph Tegbe[5], MD; Herve Etchin Boa[6], MD; Marie-Laura Bao[7], MD; Simon Amador-Paz[8], PharmB; Damien Vitiello[9], PhD; Patrick Coffie[5,10], MD, PhD; Pierre Debeaudrap[1,2,3,4*], MD, PhD

[1]CEPED, French National Research Institute for Sustainable Development, Paris, France

[2]ERL 1244, INSERM, Paris, France

[3]University Paris Sorbonne, Paris, France

[4]Université Paris Cité, Paris, Île-de-France, France

[5]PAC-CI program, Treichville Teaching Hospital, Abidjan, Côte d'Ivoire

[6]FSU Com of Abobo Avocatier-Abidjan, Abidjan, Côte d'Ivoire

[7]Centre Medico-social El Rapha-Abidjan, Abidjan, Côte d'Ivoire

[8]ANRS-MIE (Emerging Infectious Diseases), Paris, France

[9]URP 3625, Institut des Sciences du Sport Santé de Paris (I3SP), UFR STAPS, Université Paris Cité, Paris, Île-de-France, France

[10]Dermatology and infectious diseases unit, Félix Houphouët-Boigny University, Abidjan, Côte d'Ivoire

[*]these authors contributed equally

Corresponding Author:
Pierre Debeaudrap, MD, PhD
CEPED
French National Research Institute for Sustainable Development
45, rue des Saints Pères
Paris, 75006
France
Phone: 33 176533453
Email: pierre.debeaudrap@ird.fr

Abstract

Background: With improved access to antiretroviral treatment, HIV infection has become a chronic disease, and the proportion of people living with HIV aged 50 years or older is increasing. However, the long-term evolution of this disease is associated with an increased risk of comorbidities and functional impairments, which negatively impact the social participation and quality of life of people living with HIV. In resource-limited countries, population aging is a new situation, and significant challenges remain unaddressed to respond to this demographic shift. Strong evidence supports the role of physical activity (PA) in improving health and decreasing functional limitations in many chronic conditions, including HIV. However, there is a lack of information on how to effectively implement this type of nondrug intervention in resource-limited contexts.

Objective: This study aims to examine the effectiveness and implementation of 2 strategies to deliver an adapted PA program to older adults living with HIV.

Methods: This is a prospective, randomized controlled trial following a type 2 hybrid design, with a dual focus on intervention effectiveness and implementation outcomes. It also includes prior formative research that provides information on the context and guides the implementation. Conducted in Ivory Coast, the study aims to randomize 180 people living with HIV aged 50 years or older, receiving antiretroviral treatment and presenting moderate functional limitations or disabilities into the following three arms: (1) a reference arm receiving a group-based PA program supervised by a coach, (2) an exploratory arm receiving a home-based PA program with remote supervision via phone calls and messaging apps, or (3) a control arm receiving health education sessions. The total follow-up period is planned for 12 months, with an initial 6-month active phase and then a 6-month maintenance phase. The primary effectiveness outcome is the increase in the number of steps in the 6-minute step-up test between baseline and 6 months. The secondary outcomes include changes in performance on other functional tests and improvements in

XSL•FO
RenderX

cardiometabolic risk factors. The implementation outcomes include the acceptability, adoption, feasibility, and sustainability of the intervention.

Results: This study is funded by ANRS Emerging Infectious Disease, which is also the study sponsor. It received ethical approval from the National Ethical Committee of Ivory Coast (00231 3124/MSHPCMU/CNESVS-km). As of manuscript submission, the baseline formative research has been completed, participants have been randomized, and they have started the PA activity program. Results dissemination will involve civil society and decision-makers through workshops and policy briefs.

Conclusions: This study builds on previous research on healthy aging while living with HIV. Its hybrid design allows for a comprehensive evaluation of implementation processes and outcomes alongside effectiveness outcomes.

Trial Registration: ClinicalTrial.gov NCT06139497; https://clinicaltrials.gov/study/NCT06139497

International Registered Report Identifier (IRRID): DERR1-10.2196/84677

(JMIR Res Protoc 2026;15:e84677) doi:10.2196/84677

KEYWORDS

Africa; aging; exercise; HIV; implementation science

Introduction

Overview

With the scale-up of antiretroviral therapy (ART), the life expectancy of people living with HIV has improved dramatically over the last few decades, and HIV has become a chronic disease [1]. However, chronic HIV infection is associated with several long-term health challenges [2]. The chronic immune activation caused by HIV persistence, ART-related toxicity, and the occurrence of opportunistic infections increases the risk of cardiovascular diseases and musculoskeletal or neurocognitive disorders, among other conditions [3-8]. In resource-limited countries, the situation is compounded by factors such as malnutrition and opportunistic infections, which remain frequent and can increase the risk and severity of these chronic conditions [9,10]. These chronic HIV-related complications result in significant functional limitations, which translate into restrictions in daily activities and social participation, thereby negatively impacting the quality of life of people living with HIV [11-14]. According to available studies (mostly conducted in high-income countries), between one and three-quarters of people living with HIV on ART experience some form of activity limitation [15-20]. This wide range of reported prevalence likely reflects the variability of the study methodologies and population characteristics. When these limitations are ignored, they may compromise ART adherence and, consequently, treatment outcomes [14,16]. In addition, they are associated with depressive symptoms and financial hardship [19,20]. Addressing disability is therefore necessary to achieve the ambitious objective that most people living with HIV receive ART and maintain good health [21].

However, while more than two-thirds of people living with HIV reside in sub-Saharan Africa, there is a dearth of data documenting the situation of adults aging with HIV in these settings [22]. As life expectancy increases for both people living with HIV and the general population, the demographic profile of the HIV epidemic in sub-Saharan Africa is shifting, and an increasing proportion of people living with HIV are now older than 50 years [23]. Because of the significant differences in the characteristics of population with HIV (eg, sex distribution, treatment characteristics, genetics, and socioeconomic environment), data from high-income contexts cannot be simply extrapolated to African contexts.

Physical Activity

There is now strong evidence that regular physical activity (PA) increases overall life expectancy [24] and enhances health outcomes in many chronic diseases [25], including HIV [26-34]. Various types of exercise are available and can be combined, such as aerobic exercise (low to moderate intensity), resistance exercise (strength and muscular resistance), flexibility exercise, and balance exercise [35]. Two systematic reviews including 30 randomized trials revealed that aerobic activity (with or without resistance) for at least 6 weeks significantly improved cardiorespiratory function, strength, weight, and body composition and improved quality of life [31,32]. When engaged in regularly, exercise training reduces the risk of disability caused by sarcopenia among people living with HIV [36], although studies on the long-term effects of exercise training are lacking. Some studies have reported a positive effect of PA on cognitive performance [37].

In a recent study conducted in Ivory Coast and including 300 adults living with HIV and 200 without infection, we found that, after the age of 50 years, people living with HIV who reported higher levels of PA achieved similar functional performance on the 6-minute walk test (6MWT) compared with controls [38]. In contrast, those who reported low PA had lower performance on the 6MWT [38].

The Need for Research on How to Promote PA Among Aging Adults in Resource-Limited Contexts

Adopting a new lifestyle is a difficult task that requires motivation, opportunities, and capabilities. Therefore, the first step when implementing an intervention promoting PA is to obtain good knowledge of the participants' living environment and of the contextual factors that may influence the adoption of the new behavior. Grounding interventions on one or more behavior change frameworks can increase their effectiveness and improve their transferability [39]. However, few interventions promoting PA have been implemented in African settings, and evidence on methods to achieve behavior change and successfully implement PA remains scarce.

XSL•FO
RenderX

The Health Action Process Approach (HAPA) is well-suited for developing activities directed toward the adoption and maintenance of PA [40-43]. The HAPA distinguishes the motivational and volitional phases during the behavior change process. The motivational phase is characterized by the building of motivation to be engaged in the new behavior, whereas the volitional phase relates to the engagement in and maintenance of the behavior. Furthermore, the HAPA identifies 6 constructs shaping the adoption process: intention, risk perception, outcome expectancies, self-efficacy, planning, and action control [40]. Among these, self-efficacy, planning, and action control are particularly influential in the volitional phase and have been shown to be more strongly associated with PA adoption and maintenance [41,43].

The Need for an Implementation Strategy for PA Intervention

Despite strong evidence showing the effectiveness and cost-effectiveness of PA in preventing and reducing HIV- and age-related disabilities, very few programs have been implemented in sub-Saharan Africa, and none exist in West Africa [44]. Further research is needed to assess the effects of endurance and resistance training on older people living with HIV in African contexts [44]. Indeed, there is a lack of reference data on physical or neurocognitive outcomes in these populations, which are essential for establishing population norms and defining context-specific cutoffs. A significant gap persists between the scientific evidence produced by randomized clinical trials and its translation into actual and effective interventions [45]. Research is needed to explore how effective interventions can be integrated into real-world day-to-day activities of already busy health services, taking cultural differences into account [46].

In this context, this research aims to fill this gap by assessing the effectiveness and implementation of a structured PA program implemented through 2 strategies, group-based and home-based exercises, in improving the quality of life and exercise capacity of individuals aged 50 years and older living with HIV in Ivory Coast. It will primarily focus on the short-term effects of the intervention.

Methods

Study Design

VIRAGE+ (VIH Et Réadaptation Grâce À L'Exercice) is a prospective randomized controlled pilot trial with 2 phases: a 6-month active phase to implement the intervention and a 6-month maintenance phase to evaluate the sustainability of the intervention's effects (Figure 1). This study adopts a type 2 hybrid design to assess both the effectiveness of the intervention on physical, cognitive, and cardiometabolic health outcomes and the implementation outcomes to identify factors influencing its adoption by users, as well as its integration into health care practices [47]. The trial comprises the following three arms: (1) a reference arm receiving a group-based PA program delivered by a certified coach, (2) an exploratory arm receiving a home-based PA program remotely delivered via mobile phone communication, and (3) a control arm receiving general advice on healthy behaviors. Participants in the control arm will be offered the intervention after the trial if it is proven effective. Details on the intervention are provided below.

Figure 1. Study design. M12: month 12; QoL: quality of life.

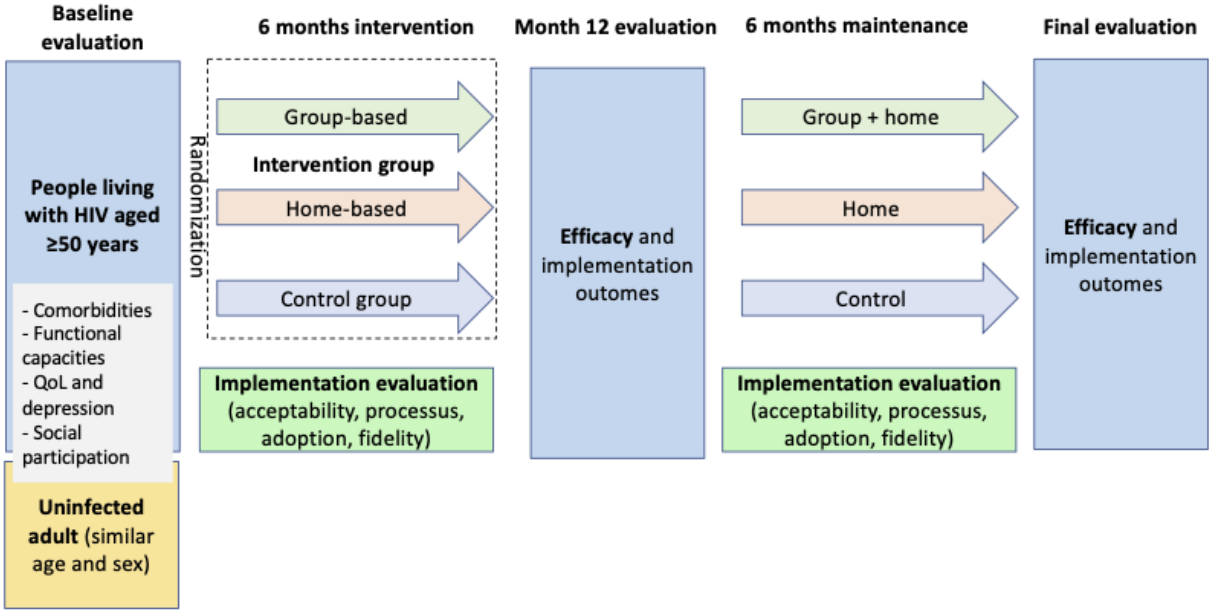

This research is coordinated by the PAC-CI program in Ivory Coast and will be conducted in 2 health centers in Abobo, Ivory Coast.

Study Population and Recruitment

The study population consists of people living with HIV aged ≥50 years, receiving ART for ≥12 months, residing in the catchment area of the study sites, and meeting at least one criterion for functional limitation or disability, as defined below.

XSL•FO
RenderX

The exclusion criterion was any condition that could impede participation in the intervention. Eligibility criteria are listed in Textbox 1.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Age ≥50 years.
- Documented HIV infection.
- Stable antiretroviral therapy for ≥12 months.
- Residing in the Abobo district.
- The presence of at least one functional limitation or disability criterion among the following:
  - Performance in the 6-minute walk test ≤300 m
  - Handgrip strength ≤35 kg for men or ≤24 kg for women.
  - World Health Organization Disability Assessment Schedule (WHODAS) score ≥10.
  - Score at the short physical performance battery (SPPB) between 4 and 9.
  - Signed informed consent.

Exclusion criteria

- Medical contraindications to physical activity.
- Anticipated unavailability.
- Severe functional limitation (SPPB ≤3).
- Another health condition requiring priority care.

Recruitment will take place in 2 health centers in the Abobo district, a popular neighborhood on the outskirts of Abidjan. The first center, "Formation Urbaine Sanitaire Abobo Avocatier," located in eastern Abobo, has an active cohort of approximately 1500 people living with HIV, including 500 individuals aged 50 years and older. The second center, "Centre Médicalisé Spécialisé El Rapha," has an active cohort of 2000 people living with HIV, including approximately 650 individuals aged 50 years and older. Participants who fulfill the inclusion criteria are contacted by their clinicians, who will obtain their informed consent to participate in the study (Multimedia Appendix 1).

Randomization and Blinding

The participants will be randomized using a computer-generated randomization list in a 1:1:1 ratio into one of three arms. Randomization will use fixed-size blocks and be stratified by study site. An independent team, not involved in intervention delivery, recruitment, or participant assessment, will generate and keep the randomization sequences. The arm allocation will be provided to investigators through a web application. An alternative procedure has been prepared in case of difficulties with the internet connection. Because of the nature of the intervention, participants and coaches will be aware of the arm allocated. However, the investigators responsible for the outcome evaluation are not involved in the intervention delivery and will remain blinded to the group assignments.

Patient and Public Involvement

This research received valuable input from people living with HIV and health workers to refine the intervention design. A preliminary formative study was conducted to explore perceptions and social representations of aging and PA among the targeted population and to identify preferences for intervention delivery. Semistructured individual interviews (n=8) and focus groups (n=4) were conducted with potential participants, health workers, and coaches. This preliminary research revealed several key aspects of the intervention, including the composition of exercise groups (mixed- or single-gender), the acceptability and feasibility of home-based sessions, and practical considerations such as preferred schedules and methods for self-monitoring progress. These results were used to refine the design, framework, and materials of the intervention to ensure cultural appropriateness and feasibility.

In addition, a steering committee including people living with HIV, health workers, and representatives of the national AIDS program will follow the study progress and advise on its implementation and conduct, as well as on the result-sharing process.

Study Intervention

The intervention consists of a 48-week exercise program including endurance, resistance, balance, and flexibility exercises performed 2-3 times per week. Endurance training sessions will include exercises conducted at 60%-70% of the maximal heart rate for at least 30 minutes, whereas resistance training sessions will involve exercises involving major muscle groups using resistance (eg, squats and lunges for lower limbs, and push-ups and dips for upper limbs) at 60%-70% of 1 repetition maximum, with 8-12 repetitions per set, performed for 2-4 sets. The Rating of Perceived Exhaustion (RPE) scale will also be used to calibrate the exercise intensity. The number

XSL•FO
RenderX

of sets will gradually increase over 24 weeks to increase strength and endurance while minimizing the risk of injury. A session should last 60 minutes by the end of this period. Together, endurance and resistance training represent 80% of the activities proposed. Exercises are usually combined with circuit training, which involves performing a series of exercises in sequence with minimal rest to maximize efficiency and cardiovascular benefits. Alternatively, exercises can be performed in isolation, such as walking, to accommodate individual preferences and abilities.

Program Initiation

The 2 intervention arms will start with a 3-week initiation phase designed to build trust between coaches and participants, ensure that the participants understand and are comfortable with the exercises planned, and establish a foundation for long-term adherence. This initiation phase includes weekly supervised sessions for both arms (group-based PA and home-based PA). The first supervised session will be devoted to an individual interview with the coach to assess participants' motivation to engage in PA, set personalized goals (frequency, intensity, duration), and discuss the practical aspects of the PA program (ie, activity type, session format, visit schedule, and communication methods). Goals will be tailored on the basis of the initial physical assessments and the participants' preferences. The following supervised sessions will focus on explaining the organization of a typical training session (warm-up, endurance, strength, and cooldown), how to use the RPE scale [48], and measure the heart rate, teaching safety aspects of physical training, and discussing the integration of the PA exercises into daily routines when performed at home.

Delivery Strategies

A total of 2 delivery approaches for the PA program are proposed. In the group-based arm, participants will take part in group sessions directly supervised by a coach. The supervised sessions are planned once a week during the first 24 weeks (active phase) and transition to once a month during the 24-week maintenance phase. This construction in 2 phases is based on behavior change frameworks such as the HAPA or the transtheoretical model of change, which distinguish 3 components (initiation, maintenance, and recovery) in the process of adoption of a new behavior [40]. The other session of the PA program will be performed at home, and the overall weekly PA will be reviewed with the coach at the end of the supervised session.

In the home-based arm, participants will perform the same exercise program at home exclusively. They will be called every week by a research assistant to review the activities performed and collect information on their adherence and perceptions of the program. During this call, the research assistant will also provide positive feedback on the participants' activity, help them address challenges, and reinforce their motivation. The weekly program will be shared via phone calls and sent via SMS text message or WhatsApp according to participants' preferences. Participants can access and download videos demonstrating each exercise, as well as a training book displaying all exercises they will have to perform during the program.

Various components of the HAPA model are targeted in both interventional arms, even though the actual activities will differ between the arms. In each interventional arm, participants will receive motivational interviews with a coach or a research assistant [49,50]. Motivational interviews will initially cover the advantages of PA and the disadvantages of sedentarism and discuss participants' intentions to engage in the adapted PA program. They will avoid a paternalistic or authoritative approach and help to build commitment to adopt this new behavior by discussing an action plan with clear, realistic goals and offering optimistic views. Motivational interviews will continue during the intervention, either through face-to-face discussions with the coach or through phone calls. The focus of these interviews will progressively shift toward self-monitoring and action planning. The participants will be asked to monitor their PA on a weekly basis using a dedicated instrument. During the interviews, they will review their goals as well as the potential problems faced and solutions identified, and they will be able to adapt their program to their goals with the help of the coach or research assistant. When needed, coping strategies will be explored and discussed.

Maintenance Phase

In the maintenance phase, participants will be encouraged to maintain their exercise routine, with 3 sessions per week. The frequency of supervised group sessions will gradually decrease, and similarly, the home-based group will experience a gradual reduction in follow-up call frequency. This gradual reduction is intended to support the development of self-efficacy and long-term adherence to the exercise regimen.

Control Group

Participants in the control group will receive monthly 30- to 60-minute health education sessions. This frequency ensures regular engagement without overburdening participants. These sessions will promote scientifically validated lifestyle habits for healthy aging, including balanced diet recommendations, age-appropriate advice for PA, and strategies for daily PA integration. The control arm will be offered the intervention after the trial if it is proven effective.

Effectiveness Outcome Assessment

In this type 2 hybrid design, both effectiveness and implementation will be evaluated and receive equal attention. For clarity, effectiveness outcomes are presented in this section, and implementation research is presented in the following section. The primary effectiveness outcome is the difference in the number of steps completed during a 6-minute step-up test (6MST) at 6 months. The 6MST is a submaximal aerobic test that provides information on cardiorespiratory and functional capacities. Although less accurate than maximal exercise tests (with ergocycles or treadmills), the 6MST was selected because it does not require sophisticated technology and is easier to implement in resource-limited settings. Compared with the 6MWT, a reference for submaximal exercise tests [51], the 6MWT has similar good reproducibility and reliability [51,52] but is easier to perform. In addition, it is slightly more demanding, which is appealing for the evaluation of adults with relatively good health. The other effectiveness outcomes

XSL•FO
RenderX

included in this study reflect key health outcomes relevant to aging individuals living with HIV and will be measured at 6 and 12 months. These include evaluations of strength, balance, cognitive performance, quality of life, depression, daily autonomy, social participation, blood pressure, and glucose metabolism. Details on the operationalization of all effectiveness outcomes are provided in Multimedia Appendix 2. All effectiveness outcomes will be assessed by investigators blinded to the participants' assigned intervention arm (single-blind).

Implementation Research

The implementation research aims to evaluate the success of program implementation and to determine which factors influence it and how. The implementation outcomes will be assessed using the RE-AIM (Reach, Efficacy, Adoption, Implementation, and Maintenance) [53] evaluation framework and cover the reach, adoption, effectiveness, satisfaction, resource requirements, and sustainability aspects of implementation. The RE-AIM framework was selected for its comprehensive approach, which evaluates both individual- and organizational-level outcomes critical for public health interventions. A description of these outcomes and of the measurement methods is provided in Table 1. The analysis of the implementation determinants will examine different levels of factors that may influence one or several aspects of the implementation process and will be guided by the Consolidated Framework of Implementation Research (CFIR) [54]. It starts with the contextual factors, which include the characteristics of the 2 implementation sites and the organizational aspects of the health units and fitness centers involved. Elements of internal context of interest are the implementation climate, the perception of PA by health workers and users, the attitudes toward change, the values and norms related to the intervention, the attitudes of the coaches toward people living with HIV and older people, and their confidence in their ability to conduct the program.

Table 1. RE-AIM (Reach, Efficacy, Adoption, Implementation, and Maintenance) framework outcomes.

Dimension and variables	Measurement method
Reach	
Number of participants who did not consent to participate and number not included in the trial	Monitoring (registers)
Reasons	Interviews and registers
Number lost to follow-up	Monitoring (baseline to month 6)
Efficacy	
Proportion of participants who have ≥1 session per week	Monitoring (registers and journals)
Proportion with participation of ≥75% of the sessions	Monitoring (registers and journals)
Reason for adherence and nonadherence	Interviews
Adoption	
Acceptability	Questionnaires (weeks 12, 24, and 48) and interviews
Satisfaction	Interviews
Implementation (fidelity)	
Number of modifications implemented	Monitoring and observations
Maintenance (phase 2)	
Proportion of participants with participation of ≥75% of sessions	Monitoring (≥6 months)
Proportion of participants with decreasing participation	Monitoring (≥6 months)
Proportion of participants having ≥1 session per week	Monitoring (≥6 months)
Reasons	Interviews

The acceptability of the PA program will be assessed before and during implementation using the theoretical framework of acceptability [55], which encompasses the following constructs: affective attitude, burden, perceived effectiveness, ethicality, intervention coherence, opportunity costs, and self-efficacy.

The implementation outcomes and determinants will be measured using quantitative methods (process indicators and questionnaires) along with qualitative methods (semistructured interviews with participants and observations). Both approaches will be implemented concurrently with systematic triangulation of data as the project progresses.

Evaluations and Data Collection

Data on the intervention's efficacy and its implementation will be collected at multiple time points during the study to address the 2 main objectives of this research.

Initial (Baseline) Evaluation

This evaluation follows the International Classification of Functioning, Disability, and Health (ICF) framework [56], chosen for its holistic approach to health, encompassing physical, functional, and social dimensions. It provides information on participants' physical capacities (strength, balance, and endurance), functional limitations, and social

XSL•FO
RenderX

participation. In addition, cardiometabolic risk will be assessed via clinical (anthropometric characteristics and blood pressure) and biological (fasting glucose, glycated hemoglobin, cholesterol levels, and urinary branched-chain amino acid detection) evaluations. Urinary branched-chain amino acid is a screening test for insulin resistance, and elevated fasting glucose or glycated hemoglobin indicates impaired glucose metabolism and type 2 diabetes mellitus. Obesity, type 2 diabetes, insulin resistance, high blood pressure, and dyslipidemia are frequent conditions among people aging with HIV; these conditions are associated with major health challenges and can be effectively prevented through PA. Finally, depressive symptoms, dietary quality, alcohol consumption, and usual PA levels will be evaluated using validated questionnaires. The list of the evaluations performed is presented in Table 2, and a description of the physical tests is provided in Multimedia Appendix 2.

Table 2. Summary of the evaluations performed (a detailed description of the tests is provided in Multimedia Appendix 2).

	Initial	Week 12	Week 24	Week 48
Clinical examination	✓		✓	✓
Heart rate, blood pressure, height, weight, body composition	✓		✓	✓
Sociodemographic characteristics and medical history	✓			
Diet quality questionnaire				
Functional test: 6-minute step-up test, grip strength measurement, gait speed test, stand-up test	✓		✓	✓
Balance tests (Y balance test, tandem, semitandem)	✓		✓	✓
Glycated hemoglobin, fasting blood glucose, cholesterol LDL[a] and HDL[b], triglyceride levels	✓			✓
IDIR[c]	✓		✓	✓
Neurocognitive evaluation (Neuroscreen)	✓			✓
Self-reported PA[d] (GPAQ[e] or IPAQ[f])	✓		✓	✓
Social participation and activity limitation: IADL[g], WHODAS-2[h], HDQ[i]	✓		✓	✓
Quality of life (WHOQoL[j]-short)	✓		✓	✓
Alcohol consumption (AUDIT[k])	✓			
Depression (PHQ-9[l])	✓		✓	✓
Satisfaction with PA (questionnaire)		✓	✓	✓
Acceptability of the PA program		✓	✓	✓

[a]LDL: low-density lipoprotein (cholesterol).

[b]HDL: high-density lipoprotein (cholesterol).

[c]IDIR: urinary branched-chain amino acid measurement to detect insulin resistance.

[d]PA: physical activity.

[e]GPAQ: Global Physical Activity Questionnaire.

[f]IPAQ: International Physical Activity Questionnaire.

[g]IADL: Instrumental Activities of Daily Living.

[h]WHODAS-2: World Health Organization Disability Assessment Schedule-2.

[i]HDQ: HIV Disability Questionnaire.

[j]WHOQoL: World Health Organization Quality of Life.

[k]AUDIT: Alcohol Use Disorders Test.

[l]PHQ-9: Patient Health Questionnaire 9.

Follow-Up Visit (Week 12)

The aim of this visit is to identify potential challenges faced by participants in the intervention groups, particularly those in the home-based group. It will involve a structured interview with the coach to discuss the training experiences and potential problems encountered, and questionnaires on self-reported PA (Internation Physical Activity Questionnaire), satisfaction with PA, and acceptability of the intervention. The interviews will explore barriers to adherence, perceptions of exercise intensity, and suggestions for program improvement. The major issues identified will lead to tailored adjustments, which will be documented and monitored for their effects.

Some participants will also have semistructured in-depth interviews with a social scientist to delve deeper into their experiences during the initial months of the intervention.

XSL•FO
RenderX

Intervention (Week 24) and Maintenance (Week 48) Evaluation Visits

These visits will provide data for assessing the primary and secondary effectiveness outcomes (Table 2). The same evaluations as for the baseline will be performed alongside questionnaires on satisfaction, acceptability, and quality of life. In addition, in-depth semistructured interviews will be conducted with a subset of participants.

Data Management

Data will be collected on paper forms before being entered into a web application designed for this study. It includes an application server and database servers hosted at the PAC-CI program in Ivory Coast with regular backup on a second server located in another site and secure access via https, passwords, and multiple firewalls. The data entered into the database are monitored with consistency checks. In addition, a monitoring of the data collected is performed by the clinical research assistant according to a monitoring plan.

Statistical Analysis

Sample Size

In a previous study, participants aged 50 years or older with HIV who reported high PA presented an improvement of approximately 10% in performance compared with those with low activity in the 6MWT (422 meters versus 380 meters) or the 6MST test (115 steps vs 105 steps) [38]. This difference is consistent with an estimated minimal clinically important difference of about 40 meters reported in other studies on older people or people with chronic disease [57,58], which translates into a difference of approximately 10 steps. Using these results, 50 participants per group will ensure 80% power to detect a significant difference between the intervention and control groups, assuming a type I error rate of 0.025 (adjusted for multiple comparisons regarding the primary end point). To accommodate attrition, each planned group size was increased by 20%, resulting in a total planned sample size of 180 participants. Importantly, this pilot study is powered to compare both intervention arms with the control arm but not to test for noninferiority between the exploratory and reference arms.

Data Analysis

Primary analyses will use the intention-to-treat approach, considering all included participants in their initial groups and using all data, regardless of the intervention participation level. If protocol deviations are frequent, a per-protocol analysis will be conducted, with a focus on participants who followed the intervention.

The primary outcome analysis will compare scores between groups with modified covariance analysis (ANCOVA), using a linear regression model to analyze postintervention results in relation to the intervention group while accounting for baseline values, age, and sex. Robust linear regression will be used if necessary. A parallel gatekeeping procedure with the following four families of end points will be used to control the type I error risk for secondary end points: (1) the primary family includes the end points related to physical capacity (6MST, strength, and 30-second stand-up test), (2) comorbidities (high blood pressure, high fasting glycemia, obesity, and insulin resistance), (3) disability (HIV Disability Questionnaire scores), (4) mental health (quality of life and Personal Health Questionnaire 9 scores). End points of comorbidities, disability, and mental health families will be analyzed only if at least one primary end point is significant (with the Hommel method used to control for type I error). Secondary end points will be analyzed using the same approach with the ANCOVA model. Exploratory analyses will analyze repeated data using mixed-effects linear models, and Generalized Additive Mixed Model will be used to model nonlinear relationships. Missing data will be described and imputed using multiple imputations by chained equations, assuming that they are randomly missing, with estimates obtained via Rubin's method.

Ethical Considerations

This study is carried out in compliance with the ethical principles set out in the current revised version of the Declaration of Helsinki and with the International Council for Harmonisation guidelines. It received ethical approval from the National Ethical Committee of Ivory Coast (number 00231 3124/MSHPCMU/CNESVS-km issued on October 11, 2024). It is registered on the ClinicalTrials.gov website (NCT06139497). Changes to the study protocol should be validated by the sponsor and scientific committee and will be promptly communicated to the investigators and ethics committee. The study has been notified to and obtained authorization from the French data protection authority for the data processing carried out as part of this research.

A knowledge transfer plan will be prepared. The study results will be published in journal articles and presented at relevant scientific conferences. Results dissemination will also involve civil society and decision-makers through workshops and policy briefs.

Because of the limited risk of this nontherapeutic trial, there is no data monitoring committee. However, a scientific committee meets at least once a year to provide guidance on the study rationale and progress. Adverse events will be reported via specific forms to the sponsor.

Results

This study is funded by ANRS Emerging Infectious Disease (ANRS-MIE), which is also the study sponsor. The trial has received approval from the National Ethical Committee of Ivory Coast and from the French Data Protection Authority (CNIL, DR-2024-229). As of manuscript submission, the baseline formative research has been completed, and all participants have been recruited and randomized. The PA program has started with an overall participation rate of 85% (102/120) though some participants started the program with several months delay. The 6-month evaluation will start in November 2025, and the study is expected to be completed in June 2026.

XSL•FO
RenderX

Discussion

Overview

This study protocol outlines the design and methodology of the VIRAGE+ trial, a multicenter, randomized controlled pilot study aimed at evaluating the effectiveness and implementation of a structured PA intervention for people living with HIV aged 50 years and older in Ivory Coast. This trial addresses the need to better understand how PA can mitigate age-related and HIV-associated disabilities in resource-limited settings and improve the exercise capacity and quality of life of older people living with HIV.

Contribution to Existing Knowledge and Strengths

Although evidence is accumulating on the beneficial effects of PA on the physical and mental health of people living with HIV, most studies have been conducted in high-income countries with limited applicability to African contexts. Moreover, socioeconomic and cultural factors affect the risk of HIV and age-related functional limitations as well as the risk of chronic disease, as shown by the significant heterogeneity in the prevalence of comorbidities and functional limitations among people living with HIV found in studies originating from sub-Saharan Africa. These contextual factors are therefore important to consider when designing and implementing preventive interventions. The mixed methods approach of this study, along with its hybrid design and preliminary formative research, will help gather information on different aspects of the context (eg, cultural, socioeconomic, and psychological characteristics).

Another important goal of this research is to investigate how a program of adapted PA could be embedded into routine HIV health care in a resource-limited context. To achieve this goal, a 3-arm hybrid trial was designed with a dual purpose. First, the group-based arm (reference strategy) will provide information on the scale of change that might be achieved with a structured PA program in the specific context of urban HIV health care centers in Ivory Coast. The home-based arm will provide evidence on the effectiveness and implementation of a simplified intervention, which is more feasible for adoption by stakeholders. An important strength of this research is the type 2 hybrid design, which will help to identify factors influencing the adoption and maintenance of PA together with evaluating their effectiveness. Various methods will contribute to implementation research, encompassing process evaluation, observations, and interviews with participants and stakeholders. The information collected will be triangulated and discussed with people living with HIV, health care workers, and program managers. Importantly, intervention implementation and implementation evaluation will be guided by frameworks [59]. The use of one or several frameworks should be considered a heuristic approach, facilitating analysis, organization, and knowledge sharing.

The intervention leverages the HAPA framework to enhance adherence and sustain long-term behavioral change. This framework was selected on the basis of preliminary formative research and helped organize the intervention's logic model. By targeting constructs such as self-efficacy, risk perception, and planning, the program is designed to address common challenges faced by aging people living with HIV, including a lack of confidence in their ability to engage in regular PA and competing socioeconomic priorities. Other factors may emerge during program implementation, and alternative frameworks could be identified through empirical data. Additional frameworks were selected to guide the implementation research [60]. The RE-AIM framework was selected to assess the different aspects of implementation success. Because it is a widely used and comprehensive framework, this approach facilitates the dissemination of the study results. The implementation research also aimed to identify factors that influence implementation using the CFIR. This framework was selected because it provides a general and systematic approach to the evaluation of implementation determinants and can easily be translated into interview guides and frameworks for questionnaire and indicator collection.

Anticipated Challenges and Mitigation Strategies

Adherence or Attrition

Adherence to PA interventions is a challenge [61], particularly in resource-limited settings where socioeconomic factors and competing priorities may diminish participation [62]. To address this, this study will incorporate motivational interviews, personalized goal setting, and regular follow-ups, which have been shown to increase self-efficacy and thereby enhance engagement [63]. Additionally, the home-based arm will include flexible scheduling and remote support via phone calls and messaging apps, accommodating participants with time constraints or mobility limitations.

Contextual Factors and Participatory Approach

Cultural perceptions of aging, PA, and HIV may influence participants' willingness to engage in the intervention. For example, group sessions are expected to foster social support, which has been shown to promote PA maintenance among older people [61], whereas home-based exercises can offer privacy for individuals who may feel stigmatized, a critical challenge among people living with HIV [64-66]. The formative research conducted during the design phase of the study provided information on the perceptions of the targeted population of the program, gathered suggestions on its implementation, and helped to identify potential implementation barriers and facilitators. Therefore, it allows for culturally tailored intervention materials and strategies. This iterative feedback from key stakeholders is critical for refining and adapting the intervention and ensuring its relevance [67]. Although this research is not strictly participatory [68], people living with HIV, health workers, and coaches will be consulted before and during the study and will be invited to discuss different aspects of the intervention, study design, and results. We believe that the discussions will be enriched as the study progresses, insofar as the stakeholders will become more familiar with the intervention and the research process, strengthening the confidence bound between them and the researchers.

Broader Implications

The situation of older people in resource-limited countries requires particular attention because of the unique challenges

XSL•FO
RenderX

posed by resource limitations. Access to health care in many resource-limited countries is limited, especially for aging populations, and noncommunicable diseases are rarely screened for. However, HIV care offers a unique opportunity for intervention, as ART programs often provide free or subsidized services. These programs could serve as entry points for integrating PA into routine care. In addition, some people living with HIV may already benefit from preventive health care services, such as cervical cancer screening, highlighting the potential for leveraging existing health care infrastructure to introduce scalable PA interventions.

Limitations and Perspectives

As a pilot study, VIRAGE+ is not powerful enough to detect noninferiority between the 2 intervention arms, which limits the conclusions that can be drawn about their relative effectiveness. However, these findings will inform the design of larger, more definitive trials. Additionally, the study relies on self-reported measures for certain outcomes, such as the PA performed, which may introduce reporting biases. Future research should explore the use of objective monitoring tools, such as wearable activity trackers, to complement self-reported data.

Another limitation is the study's focus on urban health centers. Expanding similar interventions to rural areas in future studies will be critical to ensuring equitable access to PA programs for all people living with HIV.

Conclusion

The VIRAGE+ trial represents a step toward addressing the dual challenges of aging with HIV in resource-limited settings, such as in Ivory Coast. By combining the evaluation of health outcomes with a detailed assessment of implementation processes, this study aims to generate actionable data that can inform the integration of PA into HIV care. The findings are expected to contribute to improving the quality of life of aging people living with HIV in Ivory Coast and serve as a protocol for similar interventions using adapted PA in other low- and middle-income countries.

Acknowledgments

The manuscript was edited by the American Journal of Experts. Generative artificial intelligence was not used to prepare this manuscript.

Funding

This work was supported by the ANRS Emerging Infectious Disease grant number 0396.

Authors' Contributions

PD and PC designed the study with the support of DV. PD, PC, ND, NA, JT, and SAP developed the protocol and the study instruments. ND, PD, HEB, and MLB developed the implementation plan. ND and PD wrote the first version of the manuscript. All authors edited and reviewed the manuscript and approved the final version of the article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Informed consent form.
[PDF File (Adobe PDF File), 181 KB - resprot_v15i1e84677_app1.pdf]

Multimedia Appendix 2
Description of the physical tests performed.
[PDF File (Adobe PDF File), 33 KB - resprot_v15i1e84677_app2.pdf]

Multimedia Appendix 3
Peer review comments from the l'Agence nationale de recherches sur le sida et les hépatites virales - Maladies infectieuses émergentes (National Research Agency for HIV, viral hepatitis and emerging infectious diseases; ANRS-MIE, France).
[PDF File (Adobe PDF File), 188 KB - resprot_v15i1e84677_app3.pdf]

References

1. Antiretroviral Therapy Cohort Collaboration. Life expectancy of individuals on combination antiretroviral therapy in high-income countries: a collaborative analysis of 14 cohort studies. Lancet 2008;372(9635):293-299 [FREE Full text] [doi: 10.1016/S0140-6736(08)61113-7] [Medline: 18657708]
2. Deeks SG, Lewin SR, Havlir DV. The end of AIDS: HIV infection as a chronic disease. Lancet 2013;382(9903):1525-1533 [FREE Full text] [doi: 10.1016/S0140-6736(13)61809-7] [Medline: 24152939]

XSL•FO
RenderX

3. Richert L, Brault M, Mercié P, Dauchy FA, Bruyand M, Greib C, Groupe d'Epidémiologie Clinique du SIDA en Aquitaine (GECSA). Decline in locomotor functions over time in HIV-infected patients. AIDS 2014;28(10):1441-1449. [doi: 10.1097/QAD.0000000000000246] [Medline: 24566096]
4. Habib AG, Yakasai AM, Owolabi LF, Ibrahim A, Habib ZG, Gudaji M, et al. Neurocognitive impairment in HIV-1-infected adults in sub-Saharan Africa: a systematic review and meta-analysis. Int J Infect Dis 2013;17(10):e820-e831 [FREE Full text] [doi: 10.1016/j.ijid.2013.06.011] [Medline: 23953699]
5. Robertson KR, Jiang H, Kumwenda J, Supparatpinyo K, Marra CM, Berzins B, AIDS Clinical Trials Group. Human immunodeficiency virus-associated neurocognitive impairment in diverse resource-limited settings. Clin Infect Dis 2019;68(10):1733-1738 [FREE Full text] [doi: 10.1093/cid/ciy767] [Medline: 30219843]
6. Stringer WW. Mechanisms of exercise limitation in HIV+ individuals. Med Sci Sports Exerc 2000;32(7 Suppl):S412-S421. [doi: 10.1097/00005768-200007001-00006] [Medline: 10910298]
7. Authier FJ, Chariot P, Gherardi RK. Skeletal muscle involvement in human immunodeficiency virus (HIV)-infected patients in the era of highly active antiretroviral therapy (HAART). Muscle Nerve 2005;32(3):247-260. [doi: 10.1002/mus.20338] [Medline: 15902690]
8. Calligaro GL, Gray DM. Lung function abnormalities in HIV-infected adults and children. Respirology 2015;20(1):24-32 [FREE Full text] [doi: 10.1111/resp.12385] [Medline: 25251876]
9. Kerac M, Postels DG, Mallewa M, Alusine Jalloh A, Voskuijl WP, Groce N, et al. The interaction of malnutrition and neurologic disability in Africa. Semin Pediatr Neurol 2014;21(1):42-49 [FREE Full text] [doi: 10.1016/j.spen.2014.01.003] [Medline: 24655404]
10. Hestad KA, Chinyama J, Anitha MJ, Ngoma MS, McCutchan JA, Franklin DR, et al. Cognitive impairment in Zambians with HIV infection and pulmonary tuberculosis. J Acquir Immune Defic Syndr 2019;80(1):110-117 [FREE Full text] [doi: 10.1097/QAI.0000000000001880] [Medline: 30531294]
11. Piggott DA, Erlandson KM, Yarasheski KE. Frailty in HIV: epidemiology, biology, measurement, interventions, and research needs. Curr HIV/AIDS Rep 2016;13(6):340-348 [FREE Full text] [doi: 10.1007/s11904-016-0334-8] [Medline: 27549318]
12. Rusch M, Nixon S, Schilder A, Braitstein P, Chan K, Hogg RS. Prevalence of activity limitation among persons living with HIV/AIDS in British Columbia. Can J Public Health 2004;95(6):437-440 [FREE Full text] [doi: 10.1007/BF03403989] [Medline: 15622793]
13. Banks LM, Zuurmond M, Ferrand R, Kuper H. The relationship between HIV and prevalence of disabilities in sub-Saharan Africa: systematic review (FA). Trop Med Int Health 2015;20(4):411-429 [FREE Full text] [doi: 10.1111/tmi.12449] [Medline: 25495989]
14. Rajasuriar R, Chong ML, Ahmad Bashah NS, Abdul Aziz SA, Mcstea M, Lee ECY, et al. Major health impact of accelerated aging in young HIV-infected individuals on antiretroviral therapy. AIDS 2017;31(10):1393-1403. [doi: 10.1097/QAD.0000000000001475] [Medline: 28358731]
15. Myezwa H, Hanass-Hancock J, Ajidahun AT, Carpenter B. Disability and health outcomes--from a cohort of people on long-term anti-retroviral therapy. SAHARA J 2018;15(1):50-59 [FREE Full text] [doi: 10.1080/17290376.2018.1459813] [Medline: 29635976]
16. Greene M, Covinsky K, Astemborski J, Piggott DA, Brown T, Leng S, et al. The relationship of physical performance with HIV disease and mortality. AIDS 2014;28(18):2711-2719 [FREE Full text] [doi: 10.1097/QAD.0000000000000507] [Medline: 25493597]
17. Laverick R, Haddow L, Daskalopoulou M, Lampe F, Gilson R, Speakman A, Cognitive Impairment in People with HIV in the European Region (CIPHER) Study Group. Self-reported decline in everyday function, cognitive symptoms, and cognitive function in people with HIV. J Acquir Immune Defic Syndr 2017;76(3):e74-e83. [doi: 10.1097/QAI.0000000000001468] [Medline: 29016449]
18. Van As M, Myezwa H, Stewart A, Maleka D, Musenge E. The international classification of function disability and health (ICF) in adults visiting the HIV outpatient clinic at a regional hospital in Johannesburg, South Africa. AIDS Care 2009;21(1):50-58. [doi: 10.1080/09540120802068829] [Medline: 19085220]
19. Hanass-Hancock J, Myezwa H, Carpenter B. Disability and living with HIV: baseline from a cohort of people on long term ART in South Africa. PLoS One 2015;10(12):e0143936 [FREE Full text] [doi: 10.1371/journal.pone.0143936] [Medline: 26625001]
20. Hanass-Hancock J, Misselhorn A, Carpenter B, Myezwa H. Determinants of livelihood in the era of widespread access to ART. AIDS Care 2017;29(1):32-39. [doi: 10.1080/09540121.2016.1201192] [Medline: 27350256]
21. Disability and HIV. UNAIDS. Geneva: UNAIDS; 2017. URL: http://www.unaids.org/en/resources/documents/2017/jc2905_disability-and-hiv [accessed 2025-11-28]
22. Bernard C, Dabis F, de Rekeneire N. Physical function, grip strength and frailty in people living with HIV in sub-Saharan Africa: systematic review. Trop Med Int Health 2017;22(5):516-525 [FREE Full text] [doi: 10.1111/tmi.12852] [Medline: 28170120]

XSL•FO
RenderX

23. Vollmer S, Harttgen K, Alfven T, Padayachy J, Ghys P, Bärnighausen T. The HIV epidemic in sub-Saharan Africa is aging: evidence from the demographic and health surveys in sub-Saharan Africa. AIDS Behav 2017;21(Suppl 1):101-113. [doi: 10.1007/s10461-016-1591-7] [Medline: 27837426]
24. Arem H, Moore SC, Patel A, Hartge P, Berrington de Gonzalez A, Visvanathan K, et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. JAMA Intern Med 2015;175(6):959-967 [FREE Full text] [doi: 10.1001/jamainternmed.2015.0533] [Medline: 25844730]
25. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee I, American College of Sports Medicine. American college of sports medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc 2011;43(7):1334-1359. [doi: 10.1249/MSS.0b013e318213fefb] [Medline: 21694556]
26. Montoya JL, Jankowski CM, O'Brien KK, Webel AR, Oursler KK, Henry BL. Evidence-informed practical recommendations for increasing physical activity among persons living with HIV. Aids Jan 2019:22. [doi: 10.1097/qad.0000000000002137]
27. Nixon S, Forman L, Hanass-Hancock J, Mac-Seing M, Munyanukato N, Myezwa H, et al. Rehabilitation: a crucial component in the future of HIV care and support. South Afr J HIV Med 2011;12(2):12. [doi: 10.4102/sajhivmed.v12i2.191]
28. Gomes Neto M, Conceição CS, Oliveira Carvalho V, Brites C. Effects of combined aerobic and resistance exercise on exercise capacity, muscle strength and quality of life in HIV-Infected Patients: a systematic review and meta-analysis. PLoS One 2015;10(9):e0138066 [FREE Full text] [doi: 10.1371/journal.pone.0138066] [Medline: 26378794]
29. Li A, McCabe T, Silverstein E, Dragan S, Salbach NM, Zobeiry M, et al. Community-based exercise in the context of HIV: factors to consider when developing and implementing community-based exercise programs for people living with HIV. J Int Assoc Provid AIDS Care 2017;16(3):267-275 [FREE Full text] [doi: 10.1177/2325957416686836] [Medline: 28074681]
30. Mabweazara SZ, Leach LL, Ley C. Development of a context-sensitive physical activity intervention for persons living with HIV and AIDS of low socioeconomic status using the behaviour change wheel. BMC Public Health 2019;19(1):774 [FREE Full text] [doi: 10.1186/s12889-019-7091-8] [Medline: 31208375]
31. O'Brien KK, Tynan A, Nixon SA, Glazier RH. Effectiveness of progressive resistive exercise (PRE) in the context of HIV: systematic review and meta-analysis using the Cochrane collaboration protocol. BMC Infect Dis 2017;17(1):268 [FREE Full text] [doi: 10.1186/s12879-017-2342-8] [Medline: 28403830]
32. O'Brien KK, Tynan A, Nixon SA, Glazier RH. Effectiveness of aerobic exercise for adults living with HIV: systematic review and meta-analysis using the Cochrane Collaboration protocol. BMC Infect Dis 2016;16:182 [FREE Full text] [doi: 10.1186/s12879-016-1478-2] [Medline: 27112335]
33. Vancampfort D, Stubbs B, Mugisha J. Physical activity and HIV in sub-Saharan Africa: a systematic review of correlates and levels. Afr Health Sci 2018;18(2):394-406 [FREE Full text] [doi: 10.4314/ahs.v18i2.25] [Medline: 30602967]
34. Voigt N, Cho H, Schnall R. Supervised physical activity and improved functional capacity among adults living with HIV: a systematic review. J Assoc Nurses AIDS Care 2018;29(5):667-680 [FREE Full text] [doi: 10.1016/j.jana.2018.05.001] [Medline: 29861318]
35. Stuart M, Benvenuti F, Macko R, Taviani A, Segenni L, Mayer F, et al. Community-based adaptive physical activity program for chronic stroke: feasibility, safety, and efficacy of the Empoli model. Neurorehabil Neural Repair 2009;23(7):726-734 [FREE Full text] [doi: 10.1177/1545968309332734] [Medline: 19318465]
36. Paes LDS, Borges JP, Dos Santos FM, de Oliveira TP, Dupin JG, Harris EA, et al. Effects of a 2-year supervised exercise program upon the body composition and muscular performance of HIV-infected patients. Open AIDS J 2015;9:80-88 [FREE Full text] [doi: 10.2174/1874613601509010080] [Medline: 26587076]
37. Quigley A, O'Brien K, Parker R, MacKay-Lyons M. Exercise and cognitive function in people living with HIV: a scoping review. Disabil Rehabil 2019;41(12):1384-1395. [doi: 10.1080/09638288.2018.1432079] [Medline: 29376434]
38. Debeaudrap P, Etoundi N, Tegbe J, Assoumou N, Dialo Z, Tanon A, et al. The association between HIV infection, disability and lifestyle activity among middle-aged and older adults: an analytical cross-sectional study in Ivory Coast (the VIRAGE study). BMC Public Health 2024;24(1):1549 [FREE Full text] [doi: 10.1186/s12889-024-19020-9] [Medline: 38851706]
39. Willmott T, Rundle-Thiele S. Are we speaking the same language? Call for action to improve theory application and reporting in behaviour change research. BMC Public Health 2021;21(1):479 [FREE Full text] [doi: 10.1186/s12889-021-10541-1] [Medline: 33691658]
40. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the health action process approach (HAPA). Rehabil Psychol 2011;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]
41. Zhang CQ, Zhang R, Schwarzer R, Hagger MS. A meta-analysis of the health action process approach. Health Psychol 2019;38(7):623-637. [doi: 10.1037/hea0000728] [Medline: 30973747]
42. Silva-Smith AL, Hanson CL, Neubeck L, Rowat A, McHale S. Physical activity interventions framed by the health action process approach for adults with long-term conditions: a scoping review. Int J Behav Med 2024;31(6):987-1017. [doi: 10.1007/s12529-024-10305-2] [Medline: 39009797]
43. Peng S, Khairani AZ, Zhou Z, Yuan F, Liang J. Effectiveness of interventions based on health action process approach in promoting physical activity among rehabilitation patients: a systematic review and meta-analysis. Int J Sport Exer Psychol 2024:1-25. [doi: 10.1080/1612197x.2024.2436066]

XSL•FO
RenderX

44. Chetty L, Cobbing S, Chetty V. Physical activity and exercise for older people living with HIV: a scoping review. HIV AIDS (Auckl) 2021;13:1079-1090 [FREE Full text] [doi: 10.2147/HIV.S336886] [Medline: 34984030]
45. Gottfredson DC, Cook TD, Gardner FEM, Gorman-Smith D, Howe GW, Sandler IN, et al. Standards of evidence for efficacy, effectiveness, and scale-up research in prevention science: next generation. Prev Sci 2015;16(7):893-926 [FREE Full text] [doi: 10.1007/s11121-015-0555-x] [Medline: 25846268]
46. Gautier L, Karambé Y, Dossou J, Samb OM. Rethinking development interventions through the lens of decoloniality in sub-Saharan Africa: the case of global health. Glob Public Health 2022;17(2):180-193 [FREE Full text] [doi: 10.1080/17441692.2020.1858134] [Medline: 33290183]
47. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
48. Haddad M, Stylianides G, Djaoui L, Dellal A, Chamari K. Session-RPE method for training load monitoring: validity, ecological usefulness, and influencing factors. Front Neurosci 2017;11:612 [FREE Full text] [doi: 10.3389/fnins.2017.00612] [Medline: 29163016]
49. Hettema J, Steele J, Miller WR. Motivational interviewing. Annu Rev Clin Psychol 2005;1:91-111. [doi: 10.1146/annurev.clinpsy.1.102803.143833] [Medline: 17716083]
50. O'Halloran PD, Blackstock F, Shields N, Holland A, Iles R, Kingsley M, et al. Motivational interviewing to increase physical activity in people with chronic health conditions: a systematic review and meta-analysis. Clin Rehabil 2014;28(12):1159-1171. [doi: 10.1177/0269215514536210] [Medline: 24942478]
51. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med 2002;166(1):111-117. [doi: 10.1164/ajrccm.166.1.at1102] [Medline: 12091180]
52. da Costa CH, da Silva KM, Maiworm A, Raphael Y, Parnayba J, Da Cal M, et al. Can we use the 6-minute step test instead of the 6-minute walking test? An observational study. Physiotherapy 2017;103(1):48-52. [doi: 10.1016/j.physio.2015.11.003] [Medline: 27012823]
53. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
54. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
55. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
56. International classification of functioning, disability and health (ICF). In: World Health Organisation. Geneva: World Health Organisation; 2011.
57. Redelmeier DA, Bayoumi AM, Goldstein RS, Guyatt GH. Interpreting small differences in functional status: the six minute walk test in chronic lung disease patients. Am J Respir Crit Care Med 1997;155(4):1278-1282. [doi: 10.1164/ajrccm.155.4.9105067] [Medline: 9105067]
58. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. J Am Geriatr Soc 2006;54(5):743-749. [doi: 10.1111/j.1532-5415.2006.00701.x] [Medline: 16696738]
59. Ridde V, Pérez D, Robert E. Using implementation science theories and frameworks in global health. BMJ Glob Health 2020;5(4):e002269 [FREE Full text] [doi: 10.1136/bmjgh-2019-002269] [Medline: 32377405]
60. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015;10:53 [FREE Full text] [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
61. Maula A, LaFond N, Orton E, Iliffe S, Audsley S, Vedhara K, et al. Use it or lose it: a qualitative study of the maintenance of physical activity in older adults. BMC Geriatr 2019;19(1):349 [FREE Full text] [doi: 10.1186/s12877-019-1366-x] [Medline: 31830900]
62. Kitilya B, Sanga E, PrayGod G, Kavishe BB, Ditlevsen K, Peck R, et al. Perceptions, facilitators and barriers of physical activity among people living with HIV: a qualitative study. BMC Public Health 2023;23(1):360 [FREE Full text] [doi: 10.1186/s12889-023-15052-9] [Medline: 36803443]
63. McAuley E, Blissmer B. Self-efficacy determinants and consequences of physical activity. Exerc Sport Sci Rev 2000;28(2):85-88. [Medline: 10902091]
64. Brener L, Callander D, Slavin S, de Wit J. Experiences of HIV stigma: the role of visible symptoms, HIV centrality and community attachment for people living with HIV. AIDS Care 2013;25(9):1166-1173. [doi: 10.1080/09540121.2012.752784] [Medline: 23311451]
65. Sahel-Gozin N, Loutfy M, O'Brien KK. Exploring experiences engaging in exercise from the perspectives of women living with HIV: a qualitative study. PLoS One 2023;18(6):e0286542 [FREE Full text] [doi: 10.1371/journal.pone.0286542] [Medline: 37267270]

XSL•FO
RenderX

66. Grov C, Golub SA, Parsons JT, Brennan M, Karpiak SE. Loneliness and HIV-related stigma explain depression among older HIV-positive adults. AIDS Care 2010;22(5):630-639 [FREE Full text] [doi: 10.1080/09540120903280901] [Medline: 20401765]
67. Peters DH, Tran NT, Adam T. Implementation Research in Health: A Practical Guide. Geneva: World Health Organisation; 2013.
68. Cornwall A, Jewkes R. What is participatory research? Soc Sci Med 1995;41(12):1667-1676. [doi: 10.1016/0277-9536(95)00127-s] [Medline: 8746866]

Abbreviations

6MST: 6-minute step-up test
6MWT: 6-minute walk test
ANCOVA: covariance analysis
ANRS-MIE: ANRS Emerging Infectious Disease
ART: antiretroviral therapy
CFIR: Consolidated Framework of Implementation Research
HAPA: Health Action Process Approach
ICF: International Classification of Functioning, Disability and Health
PA: physical activity
RE-AIM: Reach, Efficacy, Adoption, Implementation, and Maintenance
RPE: Rating of Perceived Exhaustion
VIRAGE+: VIH Et Réadaptation Grâce À L'Exercice

Edited by J Sarvestan; submitted 23.Sep.2025; this is a non–peer-reviewed article;accepted 24.Nov.2025; published 06.Jan.2026.

Please cite as:
Diadhiou N, Assoumou N, Tegbe J, Boa HE, Bao ML, Amador-Paz S, Vitiello D, Coffie P, Debeaudrap P
Effectiveness and Implementation of Adapted Physical Activity Delivery Strategies for Older Adults Living With HIV in Ivory Coast: Protocol for a Type 2 Hybrid Randomized Controlled Trial
JMIR Res Protoc 2026;15:e84677
URL: https://www.researchprotocols.org/2026/1/e84677
doi:10.2196/84677
PMID:

©Nicolas Diadhiou, Nelly Assoumou, Joseph Tegbe, Herve Etchin Boa, Marie-Laura Bao, Simon Amador-Paz, Damien Vitiello, Patrick Coffie, Pierre Debeaudrap. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 06.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effects of Performing Eccentric Contractions to Failure After Concentric Muscle Failure in Resistance Training Sessions: Protocol for a Within-Participant Randomized Trial

Pedro Henrique Alves Campos[1*], BSc; Renan Vieira Barreto[2*], MSc; Gabriel Fontanetti[1,2*], MSc; Leonardo Santos Lopes da Silva[1,3*], BSc; Matheus Machado Gomes[1*], PhD; Leonardo Lima[1*], PhD

[1]School of Physical Education and Sport of Ribeirão Preto, University of São Paulo, Ribeirão Preto, Brazil

[2]Department of Physical Education, Institute of Biosciences, São Paulo State University, Rio Claro, Brazil

[3]Department of Internal Medicine, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil

[*]all authors contributed equally

Corresponding Author:
Leonardo Lima, PhD
School of Physical Education and Sport of Ribeirão Preto
University of São Paulo
Av. Bandeirantes, 3900
Monte Algre
Ribeirão Preto, 14040907
Brazil
Phone: 55 1633158774
Email: leonardocrlima@usp.br

Abstract

Background: Resistance training is a well-established strategy to promote muscle hypertrophy and strength gains. Performing sets to concentric muscle failure (MF_{CON}) is commonly used to maximize neuromuscular adaptations. However, after reaching MF_{CON}, there is a remaining capacity for eccentric contractions that could be used. Increasing eccentric contraction volume may represent a promising and practical alternative to enhance training volume load and optimize adaptations, although its effectiveness in this specific application has not yet been tested.

Objective: This study aims to investigate whether performing additional eccentric contractions to eccentric muscle failure (MF_{EXC}), after the occurrence of MF_{CON}, enhances neuromuscular and morphological adaptations beyond those promoted by a traditional protocol to MF_{CON}.

Methods: In a randomized within-subject design, untrained young adult females will perform 2 upper-limb resistance training protocols over 10 weeks, including traditional (TRAD) training to MF_{CON} and a training protocol (ECC+) consisting of sets to MF_{CON} followed by eccentric-only contractions to MF_{EXC}. Each arm will be assigned to one of the protocols. Sessions (twice per week) will consist of 6 sets of unilateral elbow flexion with a load between 9 and 12 repetition maximum, with 2-minute rest intervals. Muscle function (isometric, concentric, and eccentric strength) and body composition (biceps brachii and brachialis muscle thickness and dual-energy x-ray absorptiometry [DXA]–based analysis) will be assessed pre and post intervention. Comparisons between limbs and across time will be analyzed using 2-way ANOVA. The level of significance will be set at $P<.05$.

Results: As of July 2024, a total of 7 participants have completed the intervention. Data collection was conducted between March and July 2024, with a new phase planned for the first half of 2025. Manuscript submission is expected in the second half of 2025.

Conclusions: If the hypothesis is confirmed, the ECC+ protocol may represent a practical, simple, and low-cost strategy to increase training volume and optimize strength and hypertrophy outcomes. This study may contribute to evidence-based resistance training prescriptions, particularly for women, and support the use of additional eccentric contractions as an effective tool to enhance localized muscle adaptations.

International Registered Report Identifier (IRRID): DERR1-10.2196/67537

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e67537) doi:10.2196/67537

KEYWORDS

eccentric exercise; resistance training; muscle hypertrophy; muscle function; exercise-induced muscle damage

Introduction

Background

Resistance training (RT) with free weights and gym machines is widely accepted as an effective intervention for inducing increases in strength and muscle size (ie, muscle hypertrophy) [1]. Evidence indicates that manipulating RT variables, such as volume (ie, number of sets) and intensity (ie, percentage of one-repetition maximum [RM]), is crucial for optimizing these adaptive responses [2-7]. The improvement in muscle strength in response to RT [1,5,8,9] primarily occurs due to neuromuscular and physiological adaptations that are important for muscle health, such as improved recruitment of motor units during contraction [10,11] and increased coordination resulting from enhanced activation of agonist muscles and reflex inhibition of antagonist muscles, contributing to more precise, coordinated, and efficient movements [12-14]. Additionally, RT allows for improvements in muscle capillary density and vascularization (ie, angiogenesis), which enhances the delivery of oxygen and nutrients to muscle fibers, resulting in greater metabolic efficiency and reduced fatigue during exercise [15]. Increases in strength in response to RT can be measured in various ways, such as the determination of 1RM increases for a given exercise [5], increases in maximal concentric, eccentric, and isometric torque production by the trained muscles, and changes in the optimal angle for force production [16-18], which can be assessed through specific tests on isokinetic equipment [19].

Among the variables that can be manipulated to enhance the effectiveness of RT, one particularly noteworthy strategy is performing sets to concentric muscle failure (MF_{CON}), which has been shown to effectively promote increases in muscle cross-sectional area [20-22]. This approach ensures that the individual completes the maximum number of repetitions possible for a given load. Muscle failure is defined as the point at which an individual can no longer complete a repetition with proper technique, even when exerting maximal effort. Concentric failure occurs when the individual is no longer able to perform the shortening phase of the muscle, for example, during a biceps curl, it is the moment when the person can no longer lift the weight. Taking the set to MF_{CON} can result in the recruitment of a higher number of motor units, especially high-threshold motor units and muscle fibers [10], which are not normally activated during less intense efforts or those that do not bring the muscle close to MF_{CON}. This can result in greater activation and stress on muscle fibers. Moreover, at lower intensities, a greater number of repetitions is required to reach MF_{CON}. In these cases, due to the higher number of repetitions and the high degree of muscle stress, there is greater accumulation of metabolites, such as lactate, hydrogen ions, and inorganic phosphate in muscle cells, leading to muscle swelling, which can also be a trigger for hypertrophy [23].

Another benefit of performing sets close to MF_{CON} is the greater accumulation of time under tension, an important variable for the hypertrophic process due to its role in optimizing mechanotransduction [24]. Although recent evidence shows that reaching MF_{CON} is not a sine qua non condition for muscle hypertrophy [4,25-27], it is well established that protocols incorporating this strategy are effective in promoting muscle mass increase [4,6,21,28].

Equally important, the amount of mechanical work performed in an RT session, typically estimated by the product of resistance (kg) divided by the number of repetitions performed (also referred to as volume load [VL]), is considered one of the most important variables for developing muscle strength and, especially, hypertrophy [6,29]. Therefore, progressive increases in the resistance used or the number of repetitions performed—ideally close to MF_{CON}—in a given exercise are recommended during RT interventions to promote an increase in VL [1,30,31].

There is a large body of research investigating different strategies for manipulating RT variables, especially volume and intensity, and comparing their impacts on hypertrophic responses and improvements in muscle function after chronic interventions. Examples include pyramid training systems [32], drop-set [33], rest-pause [34,35], cluster set [36], and conjugate sets [37-39], among others. However, it is commonly observed that in most studies comparing so-called advanced techniques or training models with the traditional method (ie, concentric failure), regardless of how the factors (ie, volume and intensity) are manipulated, when the VL is kept equal, the responses tend to be similar. Therefore what seems to determine the adaptations to RT is not the technique itself, but its ability to alter the final product, that is, VL [5]. Such techniques should be studied, taking into account their main practical application, increasing the training VL in order to verify whether such an increase would result in a better response. A plausible VL-increasing strategy that has yet to be investigated is performing sets of repetitions to eccentric muscle failure (MF_{EXC}). In the context of resistance training, it is common to distinguish between 2 main types of muscle failure, such as concentric and eccentric. Eccentric failure refers to the inability to control the muscle lengthening phase under load, such as during the lowering portion of a biceps curl. In this case, even with external assistance to initiate the repetition (thus overcoming the concentric phase), the individual reaches such a high level of fatigue that they can no longer resist the load during the return phase, causing the weight to drop uncontrollably. Therefore, while concentric failure represents the limit of the ability to generate force to move a load, eccentric failure indicates the exhaustion of the capacity to control that load during the muscle's lengthening phase.

Each repetition in RT consists of a concentric phase (when the external resistance is moved as a result of muscle shortening)

XSL•FO
RenderX

and an eccentric phase (when the movement of the external resistance is controlled by active muscle lengthening) [40,41]. The distinct mechanisms of force production during eccentric and concentric contractions result in significant differences in force and power generation, energy cost, and fatigue between the 2 types of contraction [41-44].

For example, a recent meta-analysis estimated that maximal eccentric strength is approximately 40% greater than maximal concentric strength [40]. More relevant to this study, it was demonstrated that to achieve the same level of fatigue observed in a concentric exercise (ie, 40% reduction in maximal isometric force), a 28.8% greater volume of contractions (ie, 134 vs 104) was required when the exercise was performed eccentrically [42]. Supporting this, it has also been shown that when the same external resistances are used (ie, 70%-95% of 1RM), protocols of concentric and eccentric contractions to MF_{CON} and MF_{EXC} result in a substantially greater volume (64% vs 152%) of the latter type of contraction compared to the former [44]. Based on this information, it can be assumed that (1) during traditional RT sessions (performed with free weights and gym machines), the eccentric phase of repetitions is performed at a relatively lower intensity (ie, % of maximal eccentric strength) since the determination of external resistance is based on the mass the performer can move during muscle shortening (ie, concentric phase); and (2) in sets performed to MF_{CON}, MF_{EXC} is not reached, resulting in a reserve of eccentric contractions that could contribute to an increase in the VL of the session but is underused.

In addition to having distinct intrinsic characteristics from concentric contractions, eccentric contractions, when emphasized in specific RT protocols, also seem to induce unique adaptations compared to traditional RT protocols. For example, when using eccentric loads greater than maximal concentric strength (ie, >1RM), exclusively eccentric training appears to promote greater overall strength gains (ie, combined concentric, isometric, and eccentric strength) compared to exclusively concentric and traditional training (ie, alternating concentric and eccentric contractions) [45]. Furthermore, high-frequency eccentric training (2.5 repetitions per second) performed on an isokinetic ergometer was associated with greater increases in muscle power (assessed through changes in jump performance) compared with traditional training [46].

An interesting characteristic of eccentric training is that it appears to promote regional hypertrophy, with a tendency to induce greater increases in the cross-sectional area of the distal portion of the vastus lateralis [47,48]. There is also a preferential increase in type II muscle fiber size, as well as an increase in cross-sectional area and tendon stiffness induced by training with rapid, high-intensity eccentric contractions. Furthermore, from a muscle mechanics perspective, training with an emphasis on eccentric contractions has been shown to result in optimization of the stretch-shortening cycle, allowing for improvements in muscle power and inducing an increase in muscle fascicle length (ie, addition of sarcomeres in series), which may lead to increased muscle shortening velocity and force production when the muscle is in a stretched state [49]. Furthermore, when comparing the chronic effects of training with different types of contractions of the elbow flexor muscles, where one group performed only concentric contractions, another performed concentric and eccentric contractions, and a third group performed only eccentric contractions, similar increases in strength and muscle mass were found between the eccentric and concentric–eccentric groups, despite a lower total training volume for the first [50]. These findings indicate a greater contribution of this type of contraction to muscular adaptations.

Eccentric training at different speeds also seems to influence cross-education (ie, increased contralateral eccentric strength) [51]. It has also been demonstrated that unilateral eccentric training promotes attenuation of muscle function loss in the contralateral limb induced by immobilization [52]. This cross-education effect is not exclusive to eccentric training; however, the cited studies demonstrate that the effect is more pronounced with eccentric training compared to traditional training. It is essential to note that investigations identifying contralateral transfer of neuromuscular adaptations induced by eccentric training were conducted with one limb exercising while the contralateral limb (which benefited from cross-education) remained inactive.

Justification

Considering that (1) the VL of RT sessions is positively associated with the magnitude of functional and morphological adaptations induced in skeletal muscle after a long-term RT intervention; (2) traditional RT protocols performed to MF_{CON} do not use a reserve of eccentric contractions that still exists after this phenomenon; (3) when MF_{CON} occurs, skeletal muscle is in an anabolic state due to both the VL and the metabolic state induced by anaerobic metabolism activation; and (4) different RT strategies with some emphasis on eccentric contractions result in unique functional and morphological adaptations in skeletal muscle. There is a gap in the literature regarding the impacts of increasing VL during RT sessions by performing eccentric muscle contractions to MF_{EXC} after the occurrence of MF_{CON}, a situation where hypertrophic signaling in the muscle is already increased. Given the rationale developed above, it is reasonable to consider that increasing VL using this strategy may result in significantly greater increases in both cross-sectional area and muscle function. If confirmed, the proposed protocol could be efficiently and easily implemented in the context of physical education and sports professionals when the goal is to promote increases in muscle mass and function.

Objectives

The objectives of this study are (1) to compare, through a randomized intraparticipant design, the acute and chronic impacts of using sets to MF_{EXC} with using sets to MF_{CON} on the VL produced during training sessions; (2) to compare the chronic impacts of using sets to MF_{EXC} with using sets to MF_{CON} on muscle morphological variables (body composition and muscle cross-sectional area); and (3) to investigate the chronic impacts of applying the 2 mentioned techniques on muscle function (maximal isometric strength, rate of development of

XSL•FO
RenderX

isometric strength, maximal concentric strength, and maximal eccentric strength).

Hypotheses

The hypotheses of this study are associated with the specific objectives presented above and are as follows: (1) the total work produced during training sessions will be greater for the intervention involving sets to MF_{EXC}; (2) the magnitude of changes in muscle morphological variables will be greater after the intervention involving sets to MF_{EXC}; and (3) improvements in muscle function will be more pronounced after the intervention involving sets to MF_{EXC}.

Methods

Participants

Young adult females (18-30 years old) will be invited to participate in the study. As an inclusion criterion, participants must have no recent experience (ie, 6 months) with upper body strength training of any kind (ie, resistance training, functional training, Pilates, and CrossFit). Exclusion criteria will include a history of musculoskeletal or joint injuries in the upper limbs within the past 12 months and use of anabolic steroids. Participants will be instructed to maintain their usual dietary habits and regular physical activity routines.

Experimental Design

The study will be conducted as a within-participant randomized trial. Each participant will undergo 2 different interventions over 10 weeks, each being performed on one of their upper limbs, with 2 training sessions per week separated by a minimum of 48 hours and a maximum of 96 hours. The intervention for each participant's upper limb will be randomly assigned via electronic randomization. Using commonly available online randomization tools, the first randomization draw will determine whether the left or right upper limb will be selected. Next, the intervention to be applied to the previously selected limb (traditional or ECC+) will be randomly selected. Finally, a randomization draw will decide whether the previously selected protocol and limb will be conducted before or after the unselected protocol (ie, both protocols will be performed in the same session). No counterbalancing based on limb dominance (ie, right-handedness) will be implemented in either the applied protocol or the order of performance. The limb assigned to traditional training (TRAD) will follow a protocol with sets of repetitions until concentric failure (MF_{CON}). The limb assigned to the eccentric overload training protocol (ECC+) will follow the same exercise protocol, but with the addition of exclusively eccentric contractions until eccentric failure (MF_{EXC}) after reaching MF_{CON}.

The existence of a widely reported cross-education effect, particularly for variables related to muscle function, can be considered a limitation of this design. However, there is a growing trend in strength training research suggesting that, despite this effect, experimental designs comparing active unilateral strength training interventions (ie, with the control limb remaining nonsedentary) performed on each limb of the same individual are more suitable than designs comparing interventions in 2 different groups of people, even if randomized [53]. This is justified by the greater control in such experimental designs over intrinsic factors (eg, gene expression of protein transcription factors, muscle fiber type distribution, and capillary-to-fiber ratio) and extrinsic factors (eg, sleep quality, training level, and diet), which are known to influence responses to strength training and cannot be controlled in designs involving different individuals in groups. Furthermore, when the primary outcome of the investigation is muscle hypertrophy, the unilateral training design appears even more suitable, as there is evidence that there is no cross-education effect for this type of outcome [54-56].

Neuromuscular function and body composition markers will be collected 1 week before and 1 week after the interventions and compared between limbs. The VL performed in each training session and the session's perceived effort will be recorded for comparison over time and between limbs. The experimental design of the study is illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. Timeline of the experiment including procedures and assessments. RM: repetition maximum.

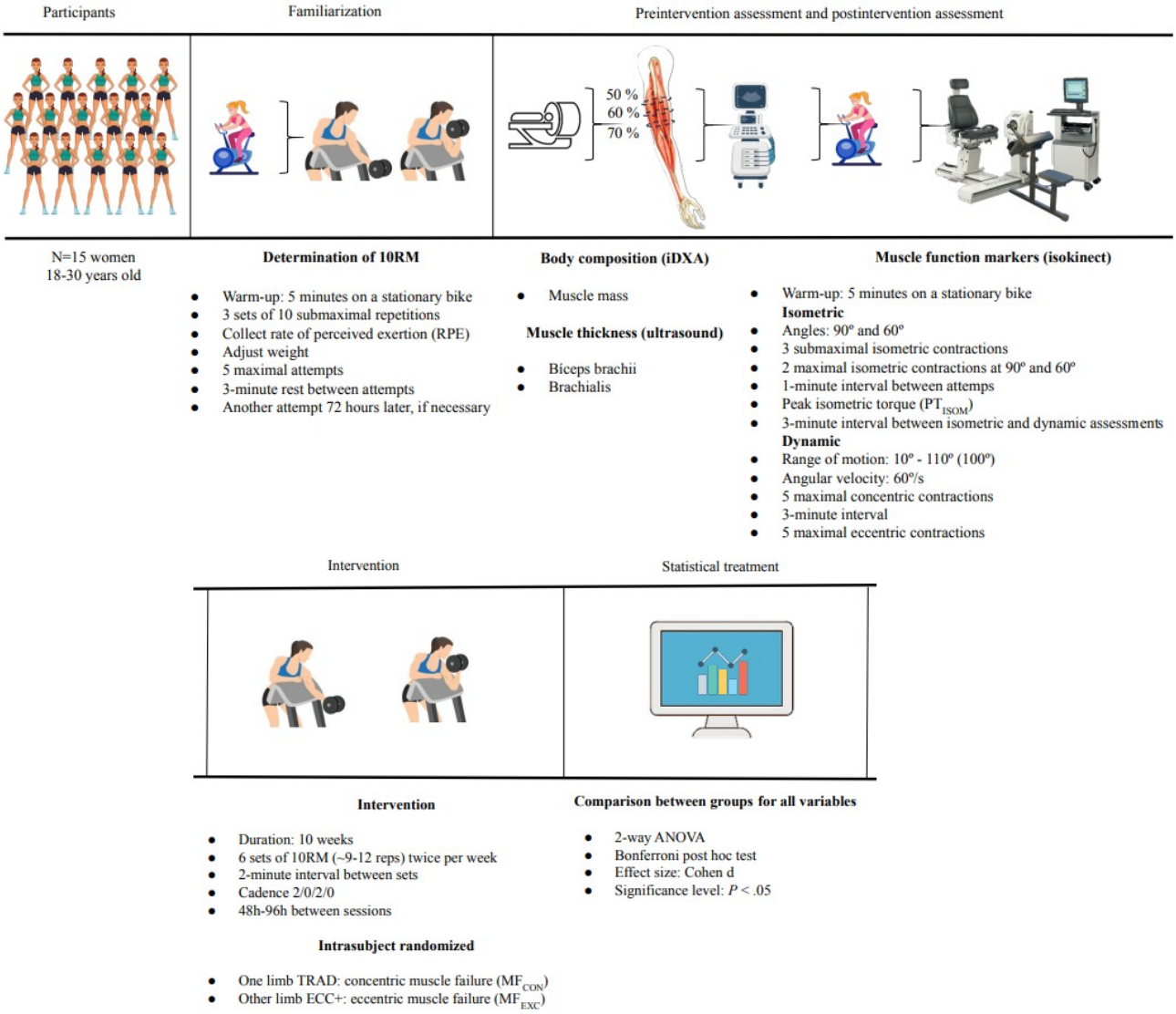

Training Protocol

The exercise to be performed by participants in this study will be unilateral elbow flexion using a preacher curl bench, Scott, OriGym Brasil, with 45º shoulder flexion, carried out at the Laboratory of Kinanthropometry and Human Performance of the School of Physical Education and Sports of Ribeirão Preto. For both protocols, participants will be instructed to perform repetitions at a cadence of 2/0/2/0 seconds (ie, 2 seconds for the concentric phase, 0 seconds of isometry between the concentric and eccentric phases, 2 seconds for the eccentric phase, and 0 seconds of isometry between the eccentric and concentric phases) with the assistance of the Beats Metronome app developed by Stonekick Apps, which will set the rhythm to 60 bpm (ie, one beat per second). This method will help control the cadence, as there will be an auditory reference at the beginning, middle, and end of each contraction. Muscle failures in the concentric and eccentric phases will be considered as the first concentric and eccentric repetitions, respectively, that the participant cannot complete within a 2-second interval. The range of motion adopted will be 100º, limited by a physical barrier during extension (maintaining a slight flexion of 10º at the end of extension and a visual apparatus for reference at the end of flexion, where the participant's forearm should be parallel to the apparatus at the end of the range of motion of the concentric phase and beginning of the eccentric phase), as illustrated in Figure 2.

XSL•FO
RenderX

Figure 2. Illustration of the settings adopted for the training sessions.

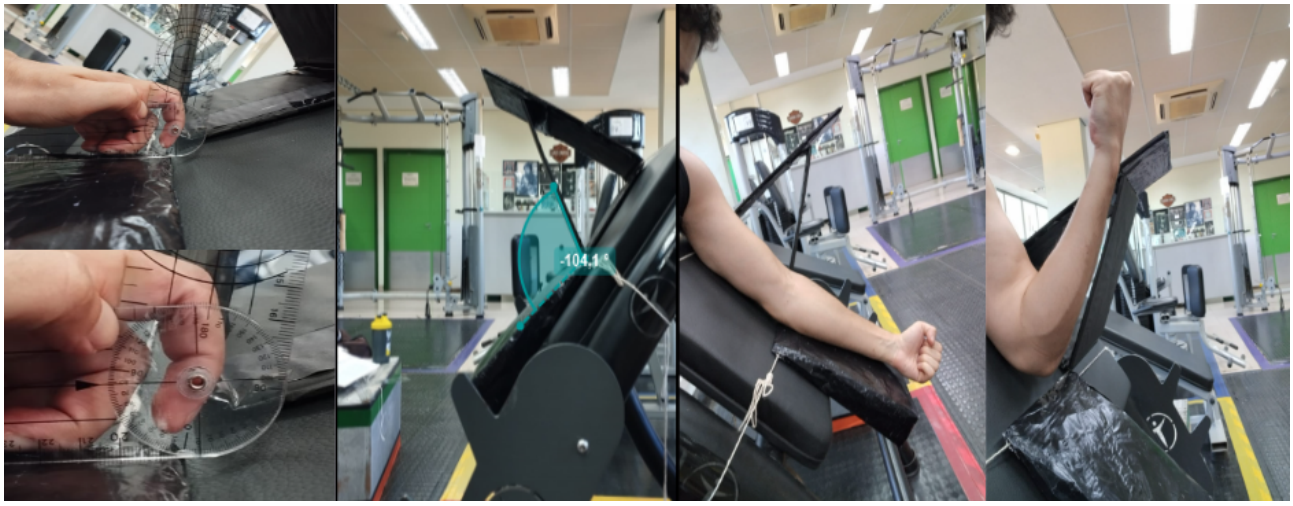

Initial Load and Training Protocol

The initial load adopted will be 10RM, which should be maintained throughout the training in the range of 9-12 maximum concentric repetitions. This means that whenever a participant reaches MF_{CON} with more than 12 or fewer than 9 repetitions in a set, the load will be increased or decreased, respectively, in the subsequent set. Load adjustments will be made in increments of at least 1 kg. For the ECC+ protocol, additional eccentric contractions will be performed with the same load used during repetitions until MF_{EXC}.

Training sessions will consist of 6 sets of elbow flexions with 2-minute recovery intervals. The recovery interval will begin when the participant reaches MF_{CON} or MF_{EXC} (for TRAD and ECC+ protocols, respectively). To allow for additional eccentric contractions after MF_{CON} in the ECC+ protocol, examiners will move the weight to the elbow-flexed position, where participants will start the eccentric phase. This will make the concentric phase of elbow flexion load-free for the participants—except for the weight of the arm itself, which will need to be moved to the flexed position.

Determination of 10RM Load

The determination of the 10RM load for the elbow flexion exercise will be conducted following a warm-up protocol consisting of 3 sets of 10 submaximal unilateral elbow flexions with a 3 kg load on the preacher curl bench, with a 3-minute rest between sets. At the end of the 30 submaximal warm-up contractions, participants will be asked to register their rate of perceived exertion (RPE) using the OMNI scale [57].

The load for the first attempt will be based on the RPE during the warm-up with a 3 kg load. If the reported value is between 8 and 9, the load for the first attempt will be increased by 0.5 kg. If the reported value is between 5 and 7, the load will be increased by 1-2 kg. If the reported value is less than 5, the initial load will be increased by 2.5-3 kg.

All attempts to determine the 10RM load will follow the same protocol described for the training sessions (ie, metronome-guided rhythm, 2/0/2/0 seconds cadence, and 100° range of motion). A maximum of 5 attempts per session will be allowed, with 3-minute intervals between attempts. If participants reach MF_{CON} before the 10th repetition during an attempt, the load for the next attempt will be decreased by 1-2 kg, depending on the number of repetitions required to reach the 10th. In cases where MF_{CON} does not occur by the 10th repetition, and participants successfully complete the 11th repetition, the attempt will be terminated, and the load for the next attempt will be increased by 0.5-1 kg. If the 10RM load is not successfully determined within the 5 attempts made in a single day, the session will be concluded, and a new determination session will be scheduled for another day, allowing at least 72 hours of recovery.

Monitoring of VL

VL will be quantified as the total work performed in each session. Perceived exertion (RPE) will also be considered as a perceptual indicator of internal training load. VL for each session will be calculated as the product of resistance (kg) and the number of repetitions completed. For the ECC+ protocol, additional eccentric contractions will be multiplied by 0.5, as they represent half of the concentric–eccentric contraction cycle. Incomplete contractions that characterize the points of concentric failure (MF_{CON}) and eccentric failure (MF_{EXC}) will not be included in the calculations.

The RPE will be reported by the participants immediately after each set using the OMNI scale [57]. Additionally, the overall session RPE will be assessed 15 minutes after the end of the session using the same scale, with instructions for participants to evaluate the entire session, not just the last set performed.

Markers of Neuromuscular Adaptation to Training Protocols

Body composition of both arms will be determined using dual-energy x-ray absorptiometry (DXA; Lunar, iDXA, GE Healthcare) in a supine position. The variable of interest for this study will be the muscle mass of the arm and the entire upper limb of the participants.

The thickness of the biceps brachii and brachialis muscles will be determined at 3 locations, based on percentages of the distance between the posterior margin of the acromion of the scapula and the olecranon of the ulna, namely proximal (50%),

XSL•FO
RenderX

middle (60%), and distal (70%) [58], using images obtained by a mode-B ultrasound device (Saevo FP 102, Ribeirão Preto) with a 40 mm linear transducer and a frequency of 10 MHz. Participants will remain in a supine position for 5 minutes before recordings to allow for fluid accommodation. The transducer will be positioned perpendicularly to the skin over the muscle and transversely to the limb. Three images of each assessed region of both muscles will be captured and exported. The ImageJ software (National Institutes of Health) will be used to determine the thickness of each muscle by calculating the difference between the superficial and deep fascias, as well as the total muscle compartment thickness, by calculating the difference between the superficial fascia of the muscle compartment and the osteomuscular interface. Reliability of ultrasound measurements will be carried out by means of the intraclass correlation coefficient.

Muscle function will be evaluated using an isokinetic dynamometer (System 4, Biodex Systems) with participants positioned on a preacher curl bench set at a 45º angle adjusted to the dynamometer. Peak isometric torque and maximum rate of torque development (RTD) will be determined from 2 maximal voluntary isometric contractions of 5 seconds at 2 different angles, 60º and 90º of elbow flexion, with a 1-minute interval between each attempt. After proper positioning on the dynamometer, following the manufacturer's instructions, and a standardized warm-up (5 minutes of exercise with a self-selected load on a stationary bike followed by 3 submaximal isometric contractions), participants will be instructed to flex the assessed elbow as quickly and with as much force as possible, sustaining the maximal contraction for 5 seconds. A 1-minute rest will be observed between the 2 contractions at each angle. The torque-time curves will be filtered and analyzed in MatLab (Mathworks, IBM), with the highest torque value recorded in the 2 contractions considered as peak isometric torque and the steepest slope of the torque-time curve as RTD. These procedures will be performed for each investigated angle (ie, 60º and 90º of elbow flexion). The contraction with the highest torque value will be used for analysis. Peak torque will be calculated as the highest value on the torque-time curve. The maximum RTD will be calculated as the steepest slope of the torque-time curve. The onset of muscle contraction will be defined as the point where torque values exceed 2.5% of the difference between baseline and peak torque values [59]. For more details on the procedures for filtering and analyzing isometric torque data, refer to [60]. Three minutes after the assessment of isometric strength, dynamic muscle function will be evaluated using nonintercalated maximal isokinetic concentric and eccentric contractions. First, 5 maximal isokinetic concentric contractions of the elbow flexors will be performed with a range of motion of 100° (10º-110º of elbow flexion) at an angular velocity of 60°/s. This range of motion was chosen to allow for greater joint comfort during movement, avoiding full extension at 0º at the beginning of the movement. Additionally, a semipronated grip at 30º is standardized for greater wrist joint comfort. The interval between contractions will be approximately 1.5 seconds, while the dynamometer arm returns passively to the initial position. After a 3-minute recovery interval, 5 maximal isokinetic eccentric contractions of the elbow flexors will be performed with the same range of motion (100°: 10º-110º), angular velocity (–60°/s), and interval between contractions. The highest torque values produced for each contraction will be recorded, and the average of the highest values from the 5 contractions (for each type of contraction) will be calculated and recorded as concentric torque peak and eccentric torque peak.

All data obtained from isokinetic force analyses and ultrasound muscle thickness measurements were performed by an evaluator who was blinded to which protocol each limb performed.

Statistical Treatment

A formal a priori power analysis was conducted using G*Power software (version 3.1.9.7; Heinrich-Heine-University Düsseldorf). The calculation was based on an effect size of 0.57 [61], for a study that observed significant increases in elbow flexor muscle thickness following resistance training performed to failure, also over a 10-week period. Considering a power (1-β) of .80, a significance α level of .05, and a 2-tailed test for repeated measures (within-subject design), the estimated required sample size was 21 upper limbs. Accordingly, the minimum number of participants needed was 11, assuming each would complete both unilateral protocols. This estimation was based on a 10-week intervention period and informed the recruitment target of this study. The normality of data distribution will be assessed using the Shapiro-Wilk test. Nonparametric data will undergo a logarithmic transformation. The assumptions of homogeneity and sphericity of the data will be checked using Levene and Mauchly tests, respectively. Comparisons over time and between limbs for all dependent variables will be conducted using a 2-way mixed model ANOVA for repeated measures (time effect) and nonrepeated measures (limb effect). If significant effects of limb or time are identified, Bonferroni post hoc adjustments will be applied. Main effect sizes will be calculated and expressed as partial η^2. If significant interaction effects between limb and time are detected, pairwise comparisons between relevant points will be performed. Effect sizes for significant differences in dependent variables between limbs or over time will be calculated and expressed as Cohen *d*. The significance level adopted will be $P<.05$.

Ethical Considerations

All participants will sign an informed consent form prior to their inclusion in the study, ensuring they are fully aware of their rights and of the possibility of withdrawing at any time without any repercussions. The data collected will be fully anonymized to ensure the privacy and confidentiality of the participants. No identifiable information will be retained, and robust protective measures will be implemented to safeguard all collected data. The present project was submitted on July 2, 2024, and approved by the Research Ethics Committee, with the Certificate of Presentation for Ethical Consideration: 76595923.8.0000.5659.

Results

Up until now, a total of 7 participants have been recruited and completed the intervention. The first preintervention data collection was conducted on January 3, 2024, and the final

XSL•FO
RenderX

postintervention test of the last participant was completed on October 7, 2024. Data collection from new participants is planned for the first semester of 2025. The paper is expected to be published in the second semester of 2025. This study was financed, in part, by the São Paulo Research Foundation (Fundação de Amparo à Pesquisa do Estado de São Paulo [FAPESP]), Brasil (process number 2023/13099-7). The workplan for the experiment is presented in Table 1.

Table 1. Baseline participant characteristics after matching.

Period	Month of execution								
	January 2024	February-March 2024	April-June 2024	July-August 2024	October-December 2024	January-March 2025	April-June 2025	July-September 2025	October-December 2025
Ethical assessment of the project (already completed)	✓								
Research team training (already completed)	✓	✓			✓				
Determining the reproducibility of assessments (already completed)	✓								
Data collection (ongoing)		✓	✓	✓	✓	✓	✓	✓	
Analysis of the data obtained					✓		✓	✓	
Discussion of the results obtained								✓	✓
Preparation of the final report								✓	✓
Preparation of abstracts and scientific articles									✓

Discussion

Anticipated Findings

The objectives of this study are to compare, through a randomized intraparticipant trial, the chronic effects of RT sets performed up to MF_{EXC}, after MF_{CON}, in relation to series performed only up to MF_{CON} on the VL produced during the training sessions, muscle mass, muscle thickness of the proximal, medial and distal regions of the arm, and muscle function assessed by maximum isometric, concentric, and eccentric strength. The hypotheses are that (1) the total work performed during the sessions will be greater in the sets performed up to MF_{EXC}, (2) the muscle morphological adaptations will be more pronounced after the intervention with sets up to MF_{EXC}, and (3) the improvements in muscle function will also be more pronounced with the ECC+ protocol. If the hypotheses that a greater training volume induces greater hypertrophy adaptations are confirmed, the results will align with previous findings in the literature [2,6,29,31]. If so, the proposed intervention could be implemented as a new practical method in the personal trainer's professional context, providing session time efficiency (ie, increased training density), as the volume load could be substantially increased with sets taken to MF_{EXC}, thereby inducing greater responses. Regarding regional hypertrophy, the analysis of the 3 regions (50%, 60%, and 70% of the arm length) may indicate that the thickness response is different between the studied protocols. Studies indicate that hypertrophy may occur nonuniformly throughout the muscle, being more pronounced in the regions that undergo greater stretching during the exercise [62]. In the preacher curl bench exercise, used in this study, peak torque occurs at the beginning of the concentric phase and at the end of the eccentric phase, which maximizes stress on the distal portion of the biceps brachii when stretching and suggests that exercise biomechanics plays a crucial role in regional hypertrophy [63]. Some possible limitations could be that the fatigue and muscle damage induced by the ECC+ protocol may not maximize hypertrophic responses, even with a higher volume load [64,65]. Another limitation is the choice to use this method on fusiform muscles, making it impossible to analyze pennation angle, which could be altered with greater emphasis on eccentric contractions, as well as not measuring fascicle length, which could be carried out in future work using the ECC+ method in other pennate muscles. The reason for choosing to investigate the elbow flexor muscles was exclusively related to the implementation logistics and the availability of equipment in our laboratory. However, as the study does not aim to answer questions related to kinesiological variables but rather to the manipulation of training variables, such as hypertrophy and strength, it is likely that this choice will not significantly interfere with the main outcome.

XSL•FO
RenderX

Additionally, the unequal volume load between limbs could be considered a limitation of the protocol. However, this choice was made to ensure greater ecological validity. The participants' uncontrolled diet, particularly in terms of carbohydrate and protein intake, is also considered a limitation. Although the study uses a within-subject design, this factor could still restrict training responses. Finally, another limitation of the present protocol concerns the absence of counterbalancing based on upper limb dominance. Although randomization was applied to the selection of the limb, the intervention (TRAD or ECC+), and the order of execution, no specific control was implemented to ensure that either the dominant or nondominant arm would perform a given protocol or initiate the session. This decision aimed to preserve the simplicity and feasibility of the protocol's practical application. Although lateral dominance may influence muscular performance, such an effect could potentially be minimized if, in this study, 10RM testing reveals similar strength values between participants' right and left upper limbs. This equivalence would suggest that, at least from an initial functional standpoint, dominance may not have a relevant impact on strength capacity. Nevertheless, the literature indicates that dominance can affect neuromuscular performance due to factors such as motor learning, neural recruitment, and habitual use. Therefore, we acknowledge that subtle dominance-related differences may exist and contribute to variability in training responses. As such, researchers applying this protocol are encouraged to consider including dominance as a covariate or to investigate its influence in future analyses, particularly in studies with greater statistical power.

Studies on strength training frequently use samples predominantly composed of men, which results in a significant bias in the application of findings to the female population. This limitation underscores the need for research that explores the impact of different training protocols specifically on women, enabling the development of evidence-based guidelines tailored to this group. Conducting studies that promote the enhancement of strength and muscle health in women also presents a significant social impact, contributing to the demystification of historical prejudices related to women's participation in strength training. Additionally, evidence suggests that women exhibit greater fatigue resistance compared to men, which can influence responses to different training intensities and volumes [66]. This superiority may be associated with the higher proportion of type I muscle fibers in women, conferring them greater resistance capacity to repetitive contractions, particularly in muscles such as elbow flexors, as used in protocols like ECC+ [67,68]. Moreover, studies indicate that women tend to recover faster after training sessions, which can directly impact the optimal periodization and training frequency [69].

Conclusion

This study aims to demonstrate whether the ECC+ protocol, which includes additional eccentric contractions to failure, can be effective in promoting greater adaptations of localized muscle hypertrophy when compared to the TRAD protocol. Both may result in significant morphological changes, such as muscle thickness, mainly in the distal portions. In addition, both may be effective in increasing concentric and eccentric strength, but the ECC+ protocol may maximize overall strength gains. If our hypotheses are confirmed, the results may suggest that additional eccentric training may be an effective strategy to provide strength and hypertrophy gains, especially in specific areas of the muscle. In practice, this means that it may be effective for a trainer to assist in extra eccentric repetitions in order to increase the total work volume of the muscles during a resistance training session to induce better responses in both strength and hypertrophy. Furthermore, it will be another study with a female target audience, which can be used to prescribe exercises for this audience.

Data Availability

While the experiment is underway, there are currently no datasets to be presented. However, the final datasets will be made available by the corresponding author on reasonable request.

Funding

This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brasil (Process Number 2023/13099-7).

Conflicts of Interest

None declared.

References

1. Schoenfeld BJ, Grgic J, Ogborn D, Krieger JW. Strength and hypertrophy adaptations between low- vs high-load resistance training: a systematic review and meta-analysis. J Strength Cond Res Dec 2017;31(12):3508-3523. [doi: 10.1519/jsc.0000000000002200]
2. Wernbom M, Augustsson J, Thomeé R. The influence of frequency, intensity, volume and mode of strength training on whole muscle cross-sectional area in humans. Sports Med 2007;37(3):225-264. [doi: 10.2165/00007256-200737030-00004] [Medline: 17326698]
3. Krzysztofik, Wilk, Wojdała, Gołaś. Maximizing muscle hypertrophy: a systematic review of advanced resistance training techniques and methods. Int J Environ Res Public Health 2019;16(24):4897 [FREE Full text] [doi: 10.3390/ijerph16244897] [Medline: 31817252]

XSL•FO
RenderX

4. Bergamasco JGA, da Silva SD, Bittencourt DF. Low-load resistance training performed to muscle failure or near muscle failure does not promote additional gains on muscle strength, hypertrophy, and functional performance of older adults. J Strength Cond Res 2022;36(5):1209-1215. [doi: 10.1519/jsc.0000000000003632]
5. Carvalho L, Junior RM, Barreira J, Schoenfeld BJ, Orazem J, Barroso R. Muscle hypertrophy and strength gains after resistance training with different volume-matched loads: a systematic review and meta-analysis. Appl Physiol Nutr Metab 2022;47(4):357-368. [doi: 10.1139/apnm-2021-0515] [Medline: 35015560]
6. Nóbrega SR, Scarpelli MC, Barcelos C, Chaves TS, Libardi CA. Muscle hypertrophy is affected by volume load progression models. J Strength Cond Res 2023;37(1):62-67. [doi: 10.1519/jsc.0000000000004225]
7. Schoenfeld BJ, Ogborn D, Krieger JW. Dose-response relationship between weekly resistance training volume and increases in muscle mass: a systematic review and meta-analysis. J Sports Sci 2017;35(11):1073-1082. [doi: 10.1080/02640414.2016.1210197] [Medline: 27433992]
8. Fiatarone MA, Marks EC, Ryan ND, Meredith CN, Lipsitz LA, Evans WJ. High-intensity strength training in nonagenarians. Effects on skeletal muscle. JAMA 1990;263(22):3029-3034. [Medline: 2342214]
9. Mangine GT, Hoffman JR, Gonzalez AM, Townsend JR, Wells AJ, Jajtner AR, et al. The effect of training volume and intensity on improvements in muscular strength and size in resistance-trained men. Physiol Rep 2015;3(8):e12472 [FREE Full text] [doi: 10.14814/phy2.12472] [Medline: 26272733]
10. Sale D. Neural adaptation to strength training. In: Komi PV, editor. Strength and Power in Sport. Hoboken, NJ: Wiley; 2008:281-314.
11. Del Vecchio A, Casolo A, Negro F, Scorcelletti M, Bazzucchi I, Enoka R, et al. The increase in muscle force after 4 weeks of strength training is mediated by adaptations in motor unit recruitment and rate coding. J Physiol 2019;597(7):1873-1887. [doi: 10.1113/jp277250]
12. Frontera WR, Meredith CN, O'Reilly KP, Knuttgen HG, Evans WJ. Strength conditioning in older men: skeletal muscle hypertrophy and improved function. J Appl Physiol (1985) 1988;64(3):1038-1044. [doi: 10.1152/jappl.1988.64.3.1038] [Medline: 3366726]
13. Cormie P, McGuigan MR, Newton RU. Developing maximal neuromuscular power: part 1--biological basis of maximal power production. Sports Med 2011;41(1):17-38. [doi: 10.2165/11537690-000000000-00000] [Medline: 21142282]
14. Suchomel TJ, Nimphius S, Bellon CR, Stone MH. The importance of muscular strength: training considerations. Sports Med 2018;48(4):765-785. [doi: 10.1007/s40279-018-0862-z] [Medline: 29372481]
15. Mang ZA, Ducharme JB, Mermier C, Kravitz L, de Castro Magalhaes F, Amorim F. Aerobic adaptations to resistance training: the role of time under tension. Int J Sports Med 2022;43(10):829-839 [FREE Full text] [doi: 10.1055/a-1664-8701] [Medline: 35088396]
16. Methenitis S, Theodorou AA, Chatzinikolaou PN, Margaritelis NV, Nikolaidis MG, Paschalis V. The effects of chronic concentric and eccentric training on position sense and joint reaction angle of the knee extensors. Eur J Sport Sci 2023;23(7):1164-1174 [FREE Full text] [doi: 10.1080/17461391.2023.2184726] [Medline: 36815692]
17. Welsh MC, Allen DL, Batliner ME, Byrnes WC. Revisiting the force-joint angle relationship after eccentric exercise. J Strength Cond Res Dec 2015;29(12):3284-3291. [doi: 10.1519/jsc.0000000000001003]
18. Chen TC, Nosaka K, Sacco P. Intensity of eccentric exercise, shift of optimum angle, and the magnitude of repeated-bout effect. J Appl Physiol (1985) 2007;102(3):992-999 [FREE Full text] [doi: 10.1152/japplphysiol.00425.2006] [Medline: 17138839]
19. Davies GJ, Wilk KE, Ellenbecker TS. Assessment of strength. In: Malone TR, Mcpoil T, Nitz NJ, editors. Orthopaedic and Sports Physical Therapy. 3rd Edition. Louis: Mosby; 1997:225-257.
20. Lacerda L, Marra-Lopes R, Lanza M, Diniz RCR, Lima FV, Martins-Costa HC, et al. Resistance training with different repetition duration to failure: effect on hypertrophy, strength and muscle activation. PeerJ 2021;9:e10909 [FREE Full text] [doi: 10.7717/peerj.10909] [Medline: 33665031]
21. Lasevicius T, Schoenfeld BJ, Silva-Batista C. Muscle failure promotes greater muscle hypertrophy in low-load but not in high-load resistance training. J Strength Cond Res 2022;36(2):346-351. [doi: 10.1519/jsc.0000000000003454]
22. Schoenfeld BJ, Grgic J. Does training to failure maximize muscle hypertrophy? Strength Cond J 2019;41(5):108-113. [doi: 10.1519/ssc.0000000000000473]
23. Schoenfeld BJ, Contreras B. The muscle pump: potential mechanisms and applications for enhancing hypertrophic adaptations. Strength Cond J 2014;36(3):21-25. [doi: 10.1519/ssc.0000000000000021]
24. Robinson ZP, Pelland JC, Remmert JF, Refalo MC, Jukic I, Steele J, et al. Exploring the dose-response relationship between estimated resistance training proximity to failure, strength gain, and muscle hypertrophy: a series of meta-regressions. Sports Med 2024;54(9):2209-2231. [doi: 10.1007/s40279-024-02069-2] [Medline: 38970765]
25. Grgic J, Schoenfeld BJ, Orazem J, Sabol F. Effects of resistance training performed to repetition failure or non-failure on muscular strength and hypertrophy: a systematic review and meta-analysis. J Sport Health Sci 2022;11(2):202-211 [FREE Full text] [doi: 10.1016/j.jshs.2021.01.007] [Medline: 33497853]
26. Refalo MC, Helms ER, Trexler ET, Hamilton DL, Fyfe JJ. Influence of resistance training proximity-to-failure on skeletal muscle hypertrophy: a systematic review with meta-analysis. Sports Med 2023;53(3):649-665 [FREE Full text] [doi: 10.1007/s40279-022-01784-y] [Medline: 36334240]

XSL•FO
RenderX

27. Santanielo N, Nóbrega SR, Scarpelli M, Alvarez I, Otoboni G, Pintanel L, et al. Effect of resistance training to muscle failure vs non-failure on strength, hypertrophy and muscle architecture in trained individuals. Biol Sport 2020;37(4):333-341 [FREE Full text] [doi: 10.5114/biolsport.2020.96317] [Medline: 33343066]
28. Ruple BA, Plotkin DL, Smith MA, Godwin JS, Sexton CL, McIntosh MC, et al. The effects of resistance training to near failure on strength, hypertrophy, and motor unit adaptations in previously trained adults. Physiol Rep 2023;11(9):e15679 [FREE Full text] [doi: 10.14814/phy2.15679] [Medline: 37144554]
29. Baz-Valle E, Balsalobre-Fernández C, Alix-Fages C, Santos-Concejero J. A systematic review of the effects of different resistance training volumes on muscle hypertrophy. J Hum Kinet 2022;81:199-210 [FREE Full text] [doi: 10.2478/hukin-2022-0017] [Medline: 35291645]
30. Aube D, Wadhi T, Rauch J. Progressive resistance training volumeffects on muscle thickness, mass, and strength adaptations in resistance-trained individuals. J Strength Cond Res 2022;36(3):600-607. [doi: 10.1519/jsc.0000000000003524]
31. Scarpelli MC, Nóbrega SR, Santanielo N. Muscle hypertrophy response is affected by previous resistance training volume in trained individuals. J Strength Cond Res 2022;36(4):1153-1157. [doi: 10.1519/jsc.0000000000003558]
32. Cardozo DC, de Souza Destro D. Pyramidal resistance training: a brief review of acute responses and long-term adaptations. J Bodyw Mov Ther 2023;35:21-27. [doi: 10.1016/j.jbmt.2023.04.070] [Medline: 37330772]
33. Angleri V, Ugrinowitsch C, Libardi CA. Crescent pyramid and drop-set systems do not promote greater strength gains, muscle hypertrophy, and changes on muscle architecture compared with traditional resistance training in well-trained men. Eur J Appl Physiol 2017;117(2):359-369. [doi: 10.1007/s00421-016-3529-1] [Medline: 28130627]
34. Enes A, Alves RC, Schoenfeld BJ, Oneda G, Perin SC, Trindade TB, et al. Rest-pause and drop-set training elicit similar strength and hypertrophy adaptations compared with traditional sets in resistance-trained males. Appl Physiol Nutr Metab 2021;46(11):1417-1424. [doi: 10.1139/apnm-2021-0278] [Medline: 34260860]
35. Prestes J, de Araujo Sousa E. Strength and muscular adaptations after 6 weeks of rest-pause vs. traditional multiple-sets resistance training in trained subjects. J Strength Cond Res 2019;33:S113-S121. [doi: 10.1519/jsc.0000000000001923]
36. Tufano JJ, Brown LE, Haff GG. Theoretical and practical aspects of different cluster set structures: a systematic review. J Strength Cond Res Mar 2017;31(3):848-867. [doi: 10.1519/jsc.0000000000001581]
37. Fleck SJ, Kramer WJ. Fundamentals of Muscle Strength Training, 4th edition. Porto Alegre: Artmed; 2017.
38. Weakley JJS, Till K, Read DB, Roe GAB, Darrall-Jones J, Phibbs PJ, et al. The effects of traditional, superset, and tri-set resistance training structures on perceived intensity and physiological responses. Eur J Appl Physiol 2017;117(9):1877-1889 [FREE Full text] [doi: 10.1007/s00421-017-3680-3] [Medline: 28698987]
39. Belo WR, Øvretveit K, De Salles BF, dos Santos LG, Ribeiro FM, Dias IB, et al. The effects of straight and alternating sets on volume load, training efficiency, and metabolic response in grapplers. J Sports Med Phys Fitness 2020;60(5):713-719. [doi: 10.23736/s0022-4707.20.10490-0]
40. Nuzzo JL, Pinto MD, Nosaka K, Steele J. The eccentric: concentric strength ratio of human skeletal muscle in vivo: meta-analysis of the influences of sex, age, joint action, and velocity. Sports Med 2023;53(6):1125-1136 [FREE Full text] [doi: 10.1007/s40279-023-01851-y] [Medline: 37129779]
41. Tomalka A. Eccentric muscle contractions: from single muscle fibre to whole muscle mechanics. Pflugers Arch 2023;475(4):421-435 [FREE Full text] [doi: 10.1007/s00424-023-02794-z] [Medline: 36790515]
42. Souron R, Nosaka K, Jubeau M. Changes in central and peripheral neuromuscular fatigue indices after concentric versus eccentric contractions of the knee extensors. Eur J Appl Physiol 2018;118(4):805-816. [doi: 10.1007/s00421-018-3816-0] [Medline: 29411127]
43. Ruas C, Latella C, Taylor J, Haff G, Nosaka K. Comparison between eccentric-only and coupled concentric-eccentric contractions for neuromuscular fatigue and muscle damage. Med Sci Sports Exerc 2022;54(10):1635-1646. [doi: 10.1249/mss.0000000000002959]
44. Shibata K, Yamaguchi T, Takizawa K, Nosaka K. Comparison in repetitions to failure between concentric-only and eccentric-only dumbbell arm curl exercise at four different relative intensities. J Strength Cond Res 2023;37(9):1754-1760. [doi: 10.1519/jsc.0000000000004470]
45. Brandenburg JP, Docherty D. The effects of accentuated eccentric loading on strength, muscle hypertrophy, and neural adaptations in trained individuals. J Strength Cond Res 2002;16(1):25-32. [Medline: 11834103]
46. Liu C, Chen C, Ho W, Füle R, Chung P, Shiang T. The effects of passive leg press training on jumping performance, speed, and muscle power. J Strength Cond Res 2013;27(6):1479-1486. [doi: 10.1519/jsc.0b013e31826bde9f]
47. Franchi MV, Atherton PJ, Reeves ND, Flück M, Williams J, Mitchell WK, et al. Architectural, functional and molecular responses to concentric and eccentric loading in human skeletal muscle. Acta Physiol (Oxf) 2014;210(3):642-654 [FREE Full text] [doi: 10.1111/apha.12225] [Medline: 24387247]
48. Seger JY, Arvidsson B, Thorstensson A. Specific effects of eccentric and concentric training on muscle strength and morphology in humans. Eur J Appl Physiol Occup Physiol 1998;79(1):49-57. [doi: 10.1007/s004210050472] [Medline: 10052660]
49. Douglas J, Pearson S, Ross A, McGuigan M. Chronic adaptations to eccentric training: a systematic review. Sports Med 2017;47(5):917-941. [doi: 10.1007/s40279-016-0628-4] [Medline: 27647157]

XSL•FO
RenderX

50. Sato S, Yoshida R, Murakoshi F, Sasaki Y, Yahata K, Kasahara K, et al. Comparison between concentric-only, eccentric-only, and concentric-eccentric resistance training of the elbow flexors for their effects on muscle strength and hypertrophy. Eur J Appl Physiol 2022;122(12):2607-2614. [doi: 10.1007/s00421-022-05035-w] [Medline: 36107233]
51. Farthing JP, Chilibeck PD. The effect of eccentric training at different velocities on cross-education. Eur J Appl Physiol 2003;89(6):570-577. [doi: 10.1007/s00421-003-0841-3] [Medline: 12756570]
52. Valdes O, Ramirez C, Perez F, Garcia-Vicencio S, Nosaka K, Penailillo L. Contralateral effects of eccentric resistance training on immobilized arm. Scand J Med Sci Sports 2021;31(1):76-90. [doi: 10.1111/sms.13821] [Medline: 32897568]
53. MacInnis MJ, McGlory C, Gibala MJ, Phillips SM. Investigating human skeletal muscle physiology with unilateral exercise models: when one limb is more powerful than two. Appl Physiol Nutr Metab 2017;42(6):563-570. [doi: 10.1139/apnm-2016-0645] [Medline: 28177712]
54. Wilkinson SB, Tarnopolsky MA, Grant EJ, Correia CE, Phillips SM. Hypertrophy with unilateral resistance exercise occurs without increases in endogenous anabolic hormone concentration. Eur J Appl Physiol 2006;98(6):546-555. [doi: 10.1007/s00421-006-0300-z] [Medline: 16972050]
55. West DW, Burd NA, Staples AW, Phillips SM. Human exercise-mediated skeletal muscle hypertrophy is an intrinsic process. Int J Biochem Cell Biol 2010;42(9):1371-1375. [doi: 10.1016/j.biocel.2010.05.012] [Medline: 20541030]
56. Bell ZW, Wong V, Spitz RW, Yamada Y, Song JS, Kataoka R, et al. Unilateral high-load resistance training influences strength changes in the contralateral arm undergoing low-load training. J Sci Med Sport 2023;26(8):440-445. [doi: 10.1016/j.jsams.2023.06.011] [Medline: 37423835]
57. Robertson RJ, Goss FL, Rutkowski J, Lenz B, Dixon C, Timmer J, et al. Concurrent validation of the OMNI perceived exertion scale for resistance exercise. Medicine & Science in Sports & Exercise 2003;35(2):333-341. [doi: 10.1249/01.mss.0000048831.15016.2a]
58. Matta T, Simão R, de Salles BE, Spineti J, Oliveira L. Strength training's chronic effects on muscle architecture parameters of different arm sites. J Strength Cond Res Jun 2011;25(6):1711-1717. [doi: 10.1519/jsc.0b013e3181dba162]
59. Andersen LL, Andersen JL, Zebis MK, Aagaard P. Early and late rate of force development: differential adaptive responses to resistance training? Scand J Med Sci Sports 2010;20(1):e162-e169. [doi: 10.1111/j.1600-0838.2009.00933.x] [Medline: 19793220]
60. Lima LCR, Nosaka K, Chen TC, Pinto RS, Greco CC, Denadai BS. Decreased running economy is not associated with decreased force production capacity following downhill running in untrained, young men. Eur J Sport Sci 2021;21(1):84-92 [FREE Full text] [doi: 10.1080/17461391.2020.1727570] [Medline: 32090683]
61. Martorelli S, Cadore EL, Izquierdo M, Celes R, Martorelli A, Cleto VA, et al. Strength training with repetitions to failure does not provide additional strength and muscle hypertrophy gains in young women. Eur J Transl Myol 2017;27(2):6339 [FREE Full text] [doi: 10.4081/ejtm.2017.6339] [Medline: 28713535]
62. Wakahara T, Fukutani A, Kawakami Y, Yanai T. Nonuniform muscle hypertrophy: its relation to muscle activation in training session. Med Sci Sports Exerc 2013;45(11):2158-2165. [doi: 10.1249/mss.0b013e3182995349]
63. Pedrosa GF, Simões MG, Figueiredo MOC, Lacerda LT, Schoenfeld BJ, Lima FV, et al. Training in the initial range of motion promotes greater muscle adaptations than at final in the arm curl. Sports (Basel) 2023;11(2):39 [FREE Full text] [doi: 10.3390/sports11020039] [Medline: 36828324]
64. Damas F, Libardi CA, Ugrinowitsch C. The development of skeletal muscle hypertrophy through resistance training: the role of muscle damage and muscle protein synthesis. Eur J Appl Physiol 2018;118(3):485-500. [doi: 10.1007/s00421-017-3792-9] [Medline: 29282529]
65. Paulsen G, Mikkelsen UR, Raastad T, Peake JM. Leucocytes, cytokines and satellite cells: what role do they play in muscle damage and regeneration following eccentric exercise? Exerc Immunol Rev 2012;18:42-97 [FREE Full text] [Medline: 22876722]
66. Salvador EP, Cyrino ES, Gurjão ALD, Dias RMR, Nakamura FY, Oliveira ARD. Comparison between the motor performance of men and women in multiple sets of weight training exercises. Rev Bras Med Esporte 2005;11(5):257-261. [doi: 10.1590/s1517-86922005000500002]
67. Haizlip KM, Harrison BC, Leinwand LA. Sex-based differences in skeletal muscle kinetics and fiber-type composition. Physiology (Bethesda) 2015;30(1):30-39 [FREE Full text] [doi: 10.1152/physiol.00024.2014] [Medline: 25559153]
68. Hunter SK. Sex differences in fatigability of dynamic contractions. Exp Physiol 2015;101(2):250-255. [doi: 10.1113/ep085370]
69. Stupka N, Tarnopolsky MA, Yardley NJ, Phillips SM. Cellular adaptation to repeated eccentric exercise-induced muscle damage. J Appl Physiol (1985) 2001;91(4):1669-1678 [FREE Full text] [doi: 10.1152/jappl.2001.91.4.1669] [Medline: 11568149]

Abbreviations

DXA: dual-energy X-ray absorptiometry
FAPESP: Fundação de Amparo à Pesquisa do Estado de São Paulo
MFCON: Concentric muscle failure

XSL•FO
RenderX

MFEXC: Eccentric muscle failure
RM: percentage of one-repetition maximum
RPE: Rate of perceived exertion
RT: Resistance training
RTD: Rate of torque development
TRAD: traditional training
VL: volume load

Edited by A Schwartz; submitted 14.Oct.2024; this is a non–peer-reviewed article;accepted 21.Jul.2025; published 08.Jan.2026.

Please cite as:
Campos PHA, Barreto RV, Fontanetti G, da Silva LSL, Gomes MM, Lima L
Effects of Performing Eccentric Contractions to Failure After Concentric Muscle Failure in Resistance Training Sessions: Protocol for a Within-Participant Randomized Trial
JMIR Res Protoc 2026;15:e67537
URL: https://www.researchprotocols.org/2026/1/e67537
doi:10.2196/67537
PMID:

©Pedro Henrique Alves Campos, Renan Vieira Barreto, Gabriel Fontanetti, Leonardo Santos Lopes da Silva, Matheus Machado Gomes, Leonardo Lima. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 08.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effectiveness and Safety of a Supplement Containing a Pharmacologically Active Basidiomycete Mushroom for Chronic Fatigue and Post–COVID-19 Fatigue Syndrome: Protocol for a Randomized Controlled Trial

Yunqing Xun[1], MSc; Tung Leong Fong[1], MSc; Guang Chen[1], PhD; Yibin Feng[1], Prof Dr; Linda Chan[1,2,3,4], Prof Dr; Ning Wang[1], Prof Dr Med

[1]School of Chinese Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Bau Institute of Medical and Health Sciences Education, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Department of Family Medicine and Primary Care, The University of Hong Kong-Shenzhen Hospital, Hong Kong, China (Hong Kong)

[4]School of Clinical Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Ning Wang, Prof Dr Med
School of Chinese Medicine
Li Ka Shing Faculty of Medicine
The University of Hong Kong
School of Chinese Medicine, Li Ka Shing Faculty of Medicine, The University of Hong Kong
Hong Kong
China (Hong Kong)
Phone: 852 3917 6441
Email: ckwang@hku.hk

Abstract

Background: Chronic fatigue syndrome, or myalgic encephalomyelitis, is characterized by persistent, unexplained exhaustion unalleviated by rest, with a pathophysiology distinct from underlying medical conditions. It is diagnostically complex due to symptoms that overlap with other disorders and the absence of definitive biomarkers, contributing to limited therapeutic options in current medicine. *Lingzhi*, a pharmacologically active basidiomycete mushroom, has been empirically used in traditional Chinese medicine for two millennia. This study has a dual focus: (1) systematically evaluating the efficacy and safety of lingzhi in managing chronic fatigue and post–COVID-19 fatigue syndrome and (2) elucidating its clinical associations with inflammatory, immune, and oxidative stress biomarkers to uncover potential therapeutic mechanisms.

Objective: The primary objective is to investigate whether a 6-week intake of CP003, a lingzhi-containing supplement, can reduce physical and mental fatigue in patients with chronic fatigue syndrome (ie, myalgic encephalomyelitis) or post–COVID-19 fatigue. Secondary objectives include assessing its effects on sleep quality, anxiety, depression, and general health status and exploring associations with inflammatory, immune, and oxidative stress biomarkers.

Methods: This randomized, waitlist-controlled trial will enroll 130 participants in Hong Kong, equally allocated (1:1) to either the CP003 intervention group or a waitlist control group. The intervention period spans 6 weeks, followed by a 6-week follow-up phase to assess the sustained effects. The trial data will be managed using REDCap (Research Electronic Data Capture) and analyzed via an intention-to-treat approach with inverse probability weighting for missing data, assessed through generalized linear regression (adjusted for covariates and interaction terms) at 6 and 12 weeks and supplemented by subgroup and sensitivity analyses in R, with results reported as mean differences (with 95% CIs) and $P<.05$ considered significant.

Results: This study was funded in March 2024 (ITF/PRP/029/24FX). As of December 2025, all participants had been enrolled, but data entry had not yet commenced. Data analysis will commence after the completion of the 12-week follow-up for all participants. Results are expected to be submitted for publication in mid-2026. Safety outcomes will also be assessed and reported.

Conclusions: This trial will evaluate the efficacy and safety of CP003 for chronic fatigue and post–COVID-19 fatigue syndrome and explore potential underlying mechanisms of action. The findings will provide experimental data to inform future clinical applications and research on traditional Chinese medicine–based interventions for fatigue-related disorders.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT06739720; https://clinicaltrials.gov/study/NCT06739720
International Registered Report Identifier (IRRID): PRR1-10.2196/82633

(JMIR Res Protoc 2026;15:e82633) doi:10.2196/82633

KEYWORDS

CP003; chronic fatigue; post–COVID-19 fatigue syndrome; randomized controlled trial; protocol

Introduction

Background and Rationale

Chronic fatigue is a term that is often used interchangeably with chronic fatigue syndrome (CFS), also known as myalgic encephalomyelitis (ME), which is a complex, long-term medical disorder characterized by extreme fatigue that cannot be explained by any underlying medical condition. The fatigue is not alleviated by rest and may worsen with physical or mental activity [1]. In addition to fatigue, people with CFS/ME may experience other symptoms, such as muscle pain, joint pain, sleep disturbances, headaches, cognitive difficulties, and postexertional malaise. Estimates of the prevalence of CFS/ME vary widely, ranging from 0.2% to 2.6% of the population depending on the diagnostic criteria used and the population studied. Some studies suggest that approximately 0.4% to 1% of the population may be affected by CFS/ME. CFS/ME can affect people of any age, but it is most commonly diagnosed in people aged between 40 and 60 years, with a female-to-male ratio ranging from 2:1 to 4:1 [2]. Diagnosing CFS/ME can be challenging due to the lack of a specific diagnostic test and the similarity of its symptoms to many other medical conditions. Due to the complex nature and the challenges in diagnosing and studying the condition, there is currently no cure for CFS/ME, but several interventions can help manage symptoms, improve the quality of life, and support daily functioning—including pacing and energy management, cognitive behavioral therapy, graded exercise therapy, sleep management, and symptom medications [3]. The treatment approach is often individualized, as the severity and types of symptoms can vary greatly among patients.

The exact pathological mechanism of CFS is not yet fully understood, but several theories have been proposed. It is likely that a combination of factors contributes to the development and persistence of the condition. Studies have suggested that CFS may be associated with an abnormal immune response, possibly due to a viral infection [4]. People with CFS often have elevated levels of immune system molecules called cytokines that could contribute to the fatigue and influenza-like symptoms. Patients with CFS/ME may have higher levels of inflammatory markers, such as cytokines, or evidence of impaired immune function, such as decreased natural killer cell activity [5]. Oxidative stress is another possible mechanism involved in CFS/ME, as some studies have found that patients with CFS/ME may have reduced levels of antioxidants or impaired antioxidant defenses, which could make them more susceptible to oxidative stress [6]. Increased oxidative stress may lead to a decrease in the production of adenosine triphosphate, the primary energy source for cells. This reduced energy production may contribute to the fatigue experienced by patients with CFS/ME. Oxidative stress can cause damage to various cellular components, including lipids, proteins, and DNA. This damage may impair cellular function and contribute to the symptoms of CFS/ME [7]. Poor sleep quality can exacerbate CFS/ME and other symptoms associated with the condition. One of the key diagnostic criteria for CFS/ME is unrefreshing sleep, where patients do not feel rested or refreshed even after a full night's sleep. This can be due to poor sleep quality, difficulty falling asleep, or frequent awakenings during the night, all of which are characteristic of insomnia [8]. People with CFS/ME often experience various sleep disturbances, such as sleep onset insomnia (difficulty falling asleep), sleep maintenance insomnia (difficulty staying asleep), and early morning awakenings. These sleep disturbances can contribute to the fatigue and cognitive difficulties experienced by patients with CFS/ME. They may also have disruptions in their circadian rhythm or sleep-wake cycle that can lead to insomnia or other sleep disorders. This disruption can result from a combination of factors, including hormonal imbalances, immune system dysfunction, and external factors, such as light exposure or lifestyle habits [9]. Specifically, there is growing evidence that some individuals recovering from COVID-19 may develop symptoms similar to those of CFS/ME, suggesting a potential link between chronic fatigue and post–COVID-19 fatigue syndrome. Although CFS and post–COVID-19 fatigue are heterogeneous in etiology, they exhibit several overlapping biological mechanisms. For example, both CFS/ME and post–COVID-19 fatigue are associated with persistent immune activation and elevated levels of proinflammatory cytokines and autoantibodies. Some studies also reported a downregulation of activities in hypothalamic pituitary axes in patients with CFS/ME and post–COVID-19 fatigue. These mechanisms can contribute to prolonged fatigue and other systemic symptoms. Both CFS/ME and post–COVID-19 syndrome are characterized by persistent fatigue, unrefreshing sleep, cognitive difficulties, and other symptoms that can significantly impact daily functioning. In the meantime, fatigue is a common and significant symptom reported by many individuals experiencing post–COVID-19 syndrome. Understanding the clinically relevant mechanism of CFS/ME will be useful for the development of effective management of CFS/ME [10].

Lingzhi, also known as "*reishi* mushroom" or *Ganoderma lucidum*, is a type of medicinal mushroom that has been used in traditional Chinese medicine (TCM) and other East Asian medicinal practices for more than 2000 years. It is a woody, shelf-like mushroom that is typically reddish-brown in color with a glossy, varnished appearance. Lingzhi is believed to have numerous medicinal properties and health benefits. It has been used as an herbal remedy to promote health, longevity, and overall well-being. Some potential benefits of lingzhi include

XSL•FO
RenderX

boosting the immune system, reducing stress and anxiety, improving liver function, supporting heart health, and providing antiaging effects. Numerous contemporary comprehensive research efforts have conclusively demonstrated the promising medical capabilities of *G lucidum*, a medicinal mushroom, in various areas, such as combating cancer, safeguarding liver health, reducing inflammation, regulating immune system functions, neutralizing harmful free radicals, and providing protection against viral infections. Clinical studies observed that consumption of *G lucidum* does not lead to the disruption of blood clotting processes and the immune system in healthy individuals, indicating its safety and tolerability in human use [11,12]. In patients with lower urinary tract symptoms, the use of lingzhi was well tolerated and provided an improvement in the International Prostate Symptom Score [13,14]. Lingzhi can significantly improve antioxidant enzymes [15], enhance physical fitness [16], and maintain the immune functions of patients who are immunocompromised [17]. However, the effect of lingzhi on diabetes remains controversial and unconfirmed [18,19]. The traditional use of lingzhi and the results of recent clinical trials support its potential application in the management of CFS in patients with or without a history of COVID-19 infections.

Randomized controlled trials are required to determine whether CP003, a lingzhi-containing supplement, is effective and safe for chronic fatigue and post–COVID-19 fatigue syndrome. Because there is no evidence-based pharmacological treatment for these fatigue conditions, the ideal comparator group is a placebo. Considering the feasibility of this trial, we designed a randomized, waitlist-controlled trial for this study.

Objectives

The primary objective of this study is to investigate whether the intake of lingzhi-containing CP003 over 6 weeks can reduce physical and mental fatigue in patients with both CFS/ME and post–COVID-19 fatigue syndrome. We will primarily focus on the effectiveness and safety of CP003-containing lingzhi for addressing chronic fatigue and post–COVID-19 fatigue symptoms as well as its clinical associations with various factors, including inflammation, immunity, and oxidative aging.

Methods

Trial Design

This is a randomized, waitlist-controlled trial on patients stratified by CFS/ME or post–COVID-19 fatigue syndrome who will receive the intervention for 6 weeks, followed by a 6-week follow-up. Because there is no approved pharmaceutical intervention for CFS/ME or post–COVID-19 fatigue, a waitlist control without a pharmacologically active drug is designed in this trial. The study is registered at the ClinicalTrials.gov (NCT06739720).

This study will be conducted in accordance with the ethical principles of the Declaration of Helsinki. The representative work flowchart is shown in Figure 1. This protocol is reported in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 statement.

Figure 1. Study flowchart for this randomized, waitlist-controlled trial evaluating CP003 in participants with chronic fatigue syndrome or myalgic encephalomyelitis (CFS/ME) or post–COVID-19 fatigue in Hong Kong.

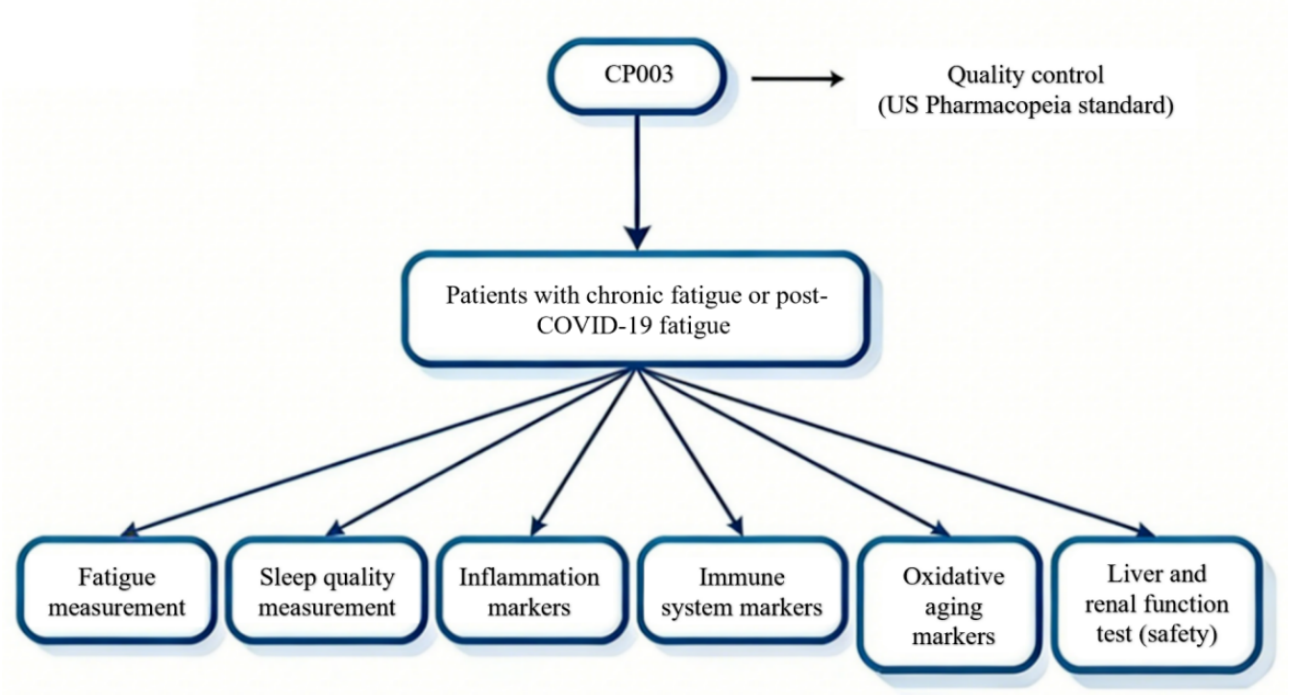

XSL•FO
RenderX

Ethical Considerations

Overview

This trial has been approved by the University of Hong Kong/Hospital Authority Hong Kong West Cluster Institutional Review Board (UW 24-639). All procedures will be conducted in accordance with the ethical principles of the Declaration of Helsinki.

Informed Consent

Written informed consent will be obtained from all participants before any study-related procedures. Research staff will explain the study's purpose, procedures, potential risks and benefits, and the right to withdraw at any time without penalty, using a detailed information sheet. Consent will be obtained electronically via the secure REDCap (Research Electronic Data Capture; Vanderbilt University) platform. Participants will have the opportunity to discuss any questions with a member of the research team before providing consent. The consent process and documentation have been approved by the aforementioned institutional review board.

Participation and Compensation

Participation in this project is entirely voluntary, and patients are not charged for anything associated with it. To assist with patients' travel to the Specialist Clinical Centre for Teaching and Research, and the blood sampling center, they will be reimbursed 800 HKD (US $104) to cover travel expenses. This reimbursement will be provided after the completion of the 12-week follow-up and is intended purely to compensate for travel costs incurred by participating in the study. Patients' participation or decision to withdraw will not affect their access to medical care in this center.

Confidentiality and Privacy

The privacy rights of all participants will be strictly upheld. Personally identifiable information will be collected only when absolutely necessary to conduct the study. All data will be deidentified before analysis, with participant codes replacing any direct identifiers, such as names, initials, hospital numbers, or specific dates. In any publications or presentations, no identifiable details will be disclosed. Should the publication of any potentially identifiable information in case descriptions, images, or other materials be deemed scientifically essential, separate and explicit written consent for publication will be obtained from the affected individuals, and the manuscript will be presented to them for review before submission.

Study Setting

This study will take place at the Specialist Chinese Medicine Clinic, School of Chinese Medicine, The University of Hong Kong, Hong Kong.

Patient and Public Involvement

Patients were not involved in the design or development of this study protocol. Participants meeting the eligibility criteria (ie, aged 40-60 years, diagnosed with CFS/ME per the Fukuda criteria or with post–COVID-19 persistent fatigue, and able to complete the 6-week intervention period and online fatigue assessments) were recruited through advertisement campaigns and outpatient clinics in Hong Kong. Potential participants were provided with detailed study information, including intervention requirements and assessment procedures, before obtaining written informed consent. Exclusion criteria, such as recent use of supplements or herbal medicine, acute inflammation, pregnancy, or severe comorbidities, were clearly communicated during recruitment to ensure appropriate enrollment. Although study results will not be disseminated directly to participants, individuals may request outcome summaries or the published manuscript upon trial completion.

Eligibility Criteria

Inclusion Criteria

Participants are included in this trial if they meet the following criteria: aged between 40 and 60 years, diagnosed with CFS/ME according to the Fukuda criteria or experiencing persistent fatigue following post–COVID-19 condition [20], capable of undergoing a 6-week intervention program, and able to complete online validated fatigue assessments. Recruitment will be conducted through both advertisement campaigns and outpatient clinic enrollments in Hong Kong, targeting individuals with CFS/ME or post–COVID-19 condition who meet the eligibility requirements.

Exclusion Criteria

The exclusion criteria are as follows: (1) patients who have taken any type of supplements or herbal medicine within the past 6 months, (2) patients who have serious medical conditions that might limit their participation in this intervention, (3) patients who are diagnosed with acute inflammation, or (4) patients who are pregnant or planning to become pregnant in the next 3 months.

Randomization and Blinding

After screening and baseline assessments, eligible participants will be randomly assigned to the CP003 intervention group or waitlist control group in a 1:1 ratio stratified by CFS/ME and post–COVID-19 condition. The CP003 intervention group received a therapeutic dose (5 capsules taken orally once per day for 6 weeks). The waitlist control group will wait for 6 weeks after the randomization and then start to receive CP003 capsules (5 capsules taken orally once per day for 6 weeks), followed by a 6-week follow-up. The random sequences generated in the R statistical package will have a block of 8 in each cycle, consisting of 2 CFS/ME control groups, 2 post–COVID-19 condition control groups, 2 CFS/ME intervention groups, and 2 post–COVID-19 condition intervention groups. Assignments were concealed in sealed opaque envelopes and were only opened after eligibility was confirmed and consent was obtained. The research study staff who recruit and screen the patients and the assessors who conduct baseline and follow-up assessments will be blinded to treatment allocation.

Interventions

Participants assigned to the intervention group will receive 5 CP003 capsules daily for 6 weeks, followed by a 6-week follow-up. CP003 is a Chinese medicine nutritional supplement used in Hong Kong, manufactured by Chinese Pharmaceuticals

XSL•FO
RenderX

Co Ltd. Per the manufacturer's guidelines, a complete treatment course consists of 6 weeks of CP003. Quality control will ensure that the similarity of multiple batches exceeds 96%, as tested by ultraperformance liquid chromatography with fingerprint analysis. During the intervention period, patients will be asked to complete a supplement intake diary to record any missed doses and any additional supplements taken. At the 6-week visit, participants are required to return any unused trial capsules.

The control group (ie, the waitlist) will wait for 6 weeks after randomization before receiving CP003 capsules and then start taking 5 capsules daily for 6 weeks, followed by a 6-week follow-up. During the waiting period, participants will be asked to complete a supplement intake diary to record any supplements they take in their daily lives. The overview of the study design is presented in Figure 2.

Figure 2. Schematic overview of the trial design and participant timeline for the CP003 intervention study. Patients with fatigue will be randomly assigned (1:1) to the intervention group and control group.

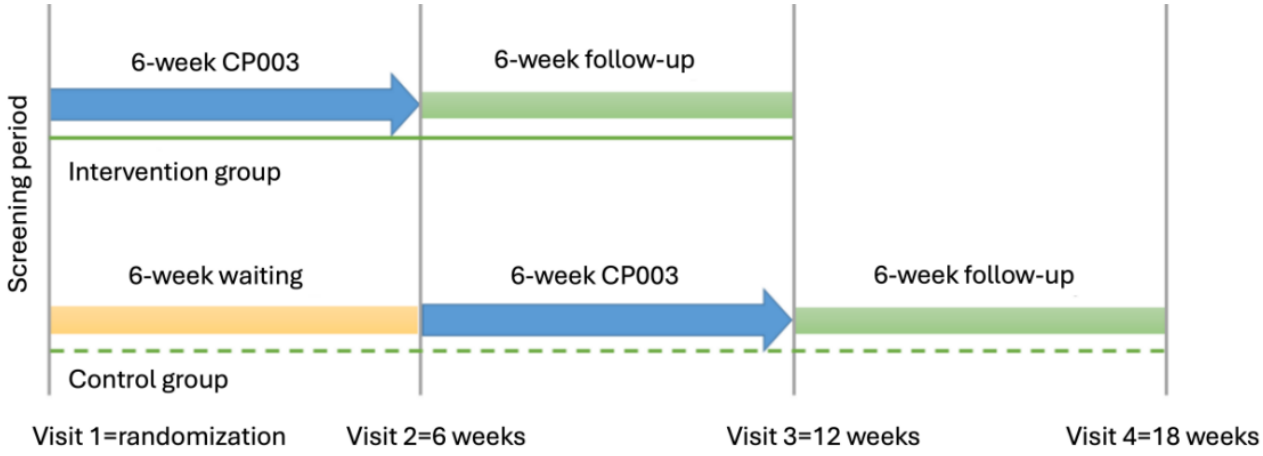

Outcomes

Outcome measures, including the Chalder Fatigue Scale score (Likert score 0-4 for 11 items, with higher scores indicating greater fatigue) [21], Fatigue Severity Score (Likert score 1-7 for 9 items, where higher scores indicate more severe fatigue) [22], Patient-Reported Outcomes Measurement Information System (PROMIS) Short Form Fatigue 7a survey (a 5-point Likert scale of 7 items, with higher scores indicating greater fatigue) [23], 36-Item Short Form Health Survey (SF-36; ranging 0-100, with higher scores indicating better health status) [23], Hospital Anxiety and Depression Scale (HADS; ranging 0-21, with higher scores indicating more severe symptoms) [24], and Pittsburgh Sleep Quality Index (ranging 0-21, with higher scores indicating worse sleep quality) [25]. Additional measures include an outcome expectations scale; physical examinations (assessing tongue and pulse, blood pressure, and body temperature); serum markers for inflammation, immunity, and oxidative aging; as well as laboratory tests for safety assessments, including complete blood count, renal function (blood urea nitrogen and creatinine), and liver function (aspartate aminotransferase and alanine aminotransferase).

Primary Outcome Measures

The primary outcome is the mean difference in the Chalder Fatigue Scale score from baseline to week 6.

Secondary Outcome Measures

The secondary outcomes include the mean difference in the Chalder Fatigue Scale scores from baseline to 12 weeks, the mean difference in the SF-36 scores from baseline to both 6 and 12 weeks, the mean difference in the Fatigue Severity Scores from baseline to both 6 and 12 weeks, the mean difference in the PROMIS Short Form Fatigue 7a survey scores from baseline to both 6 and 12 weeks, the mean difference in the HADS scores from baseline to both 6 and 12 weeks, and the mean difference in the Pittsburgh Sleep Quality Index scores from baseline to both 6 and 12 weeks.

Blood Tests

Patients will be asked to undergo a blood test at baseline and at the end of the intervention program to investigate the inflammation-, immune-, and oxidative aging–related measures. For chronic inflammation and immunity assessment, the BD Cytometric Bead Array system will be used to quantify the level of inflammation panel (interleukin-10, interleukin-12 p70, interleukin-1β, and interleukin-6), T helper 1 cells and T helper 2 cells panel (interleukin-2, interleukin-4, interleukin-5, tumor necrosis factor, and IFN-γ), B cell activation panel (CD79b, B cell linker protein, Btk, Syk, and PLCγ), T cell activation panel (TCRz, SLP-76, ZAP-70, Pyk2, Itk), and the complements panel (C4a, C3a, and C5a). The plasma level of C-reactive protein will be determined by the enzyme-linked immunosorbent assay. The oxidative aging maker in the plasma, including antioxidative enzyme activities of superoxide dismutase, glutathione peroxidase, and catalase as well as GSH (reduced form of glutathione), GSSG (oxidized form of glutathione), and malondialdehyde concentration, will be determined by colorimetric biochemical assays. For cardio-related markers, triglyceride, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and total cholesterol will be tested. Fasting plasma glucose will also be evaluated. Blood vessel elasticity will be evaluated by the arterial velocity-pulse index and arterial pressure-volume index by a noninvasive PASESA detector (Win Horizon).

XSL•FO
RenderX

Sample Size

The sample size for this study is based on the minimal important difference threshold of 2 points, which is the expected mean difference between groups [26]. Using a 2-sided hypothesis test at a significance level of .05 with an allocation ratio of 1:1, a sample size of 110 is calculated to achieve approximately 80% power to detect this effect size difference. Considering a 15% loss to follow-up or dropout rate, the final estimated sample size for this trial is 130.

Feasibility and Process Evaluation

To assess the feasibility of the trial and the acceptability of the CP003 intervention, we will monitor the following process measures: (1) recruitment rate (ie, the number of participants enrolled per month), (2) retention and dropout rates at each study visit, (3) intervention adherence calculated based on the count of the remaining capsules and self-reported intake diaries, and (4) participant satisfaction with the intervention and study procedures assessed via a brief survey at the final visit. These measures will inform the practicality of conducting a larger definitive trial.

Recruitment

Patients in this trial will be recruited through both advertisements and outpatient clinic enrollments in Hong Kong.

Adverse Events

Adverse events will be assessed using the Treatment Emergent Symptom Scale and evaluated during the whole process as well as laboratory tests (eg, whole blood count, renal, and liver function tests), if needed. All clinical adverse events will be recorded in terms of intensity (mild, moderate, or severe), duration, outcome, and relationship to the study.

Handling and Storage of Personal Data and Study Data

Data will be managed by and accessible only to authorized investigators and designated personnel on a password-protected system throughout the study period. In addition, an independent data and safety monitoring committee, separate from sponsors, researchers, and stakeholders, will convene quarterly to oversee the safety and effectiveness of the research in the medium term. The principal investigator will have access to the personal and study data during and after the study.

Our project team will ensure the security of raw data by storing it in a locked cabinet within a secure room with limited access. Data entry will be performed on a password-protected computer accessible solely to team members. Throughout and following the study, only designated team members will have the authority to access raw data or study records. Upon study completion, raw data will be preserved for a period of 5 years and subsequently disposed of, unless an application for extended retention is sanctioned by the institutional review board. Before statistical analysis, any personal identifying information will be expunged from the dataset.

Data Processing and Analysis

The primary analysis will follow the intention-to-treat principle. The between-group difference in the primary outcome, defined as the mean change in the Chalder Fatigue Scale score from baseline to week 6, will be assessed using a generalized linear regression model, adjusted for the baseline score and stratification factors of CFS/ME versus post–COVID-19 fatigue. Results will be presented as the adjusted mean difference with a 95% CI. For participants lost to follow-up, inverse probability weighting will be used to address potential bias from missing data. Analysis of secondary outcomes and biomarker data will use similar generalized linear models at 6 and 12 weeks, adjusted for prespecified covariates, including time since COVID-19 infection, vaccination doses, routine exercise pattern, and baseline sleep quality. Subgroup analyses will be conducted based on gender, age category, and history of COVID-19 infection. Sensitivity analyses, including those using multiple imputation for missing data, will be performed. All analyses will be conducted using R software (version 4.5.2; R Foundation for Statistical Computing), with a 2-sided P value of <.05 considered statistically significant.

Withdrawal and Dropout

If participants use a combination of treatments that are prohibited by the study protocol, withdraw their consent, or cease communication, they will be excluded from the study. Researchers will report the reasons for withdrawal and dropout and acquire the time of last treatment and outcome measures where possible.

Consent or Assent

Research staff will ask interested participants to provide consent via the secure REDCap electronic consent platform. Participants will receive information regarding the trial electronically and have the opportunity to discuss the trial specifics and meet with a research team member, virtually or in person, before deciding on participation.

Dissemination Policy

Upon study completion, the main findings will be disseminated to patients, primary care providers, and physicians through the hospital authority network in Hong Kong. Results will be presented at academic conferences and published in peer-reviewed journals.

Results

This is a study protocol, and no results from the trial are available at this stage. This study received funding in March 2024 (grant ITF/PRP/029/24FX). Participant recruitment concluded in September 2025. Follow-up for each patient will be 6 weeks. The total duration of the study will be approximately 24 months. As of December 2025, all participants had been enrolled, but data entry had not yet commenced. Data collection, including baseline and all follow-up assessments, will proceed according to the timeline presented in Figure 2. Data analysis will commence after all participants complete the 12-week follow-up. Results are expected to be submitted for publication in mid-2026. Upon trial completion, we will report the between-group differences in primary and secondary outcomes, such as the Chalder Fatigue Scale, Fatigue Severity Score, SF-36, PROMIS fatigue 7a, HADS, and the Pittsburgh Sleep Quality Index, at the 6- and 12-week time points. Changes in inflammatory, immune, and oxidative stress biomarkers from

XSL•FO
RenderX

baseline to the end of intervention will also be presented. Safety outcomes will be assessed through adverse event monitoring, and laboratory tests will be reported.

In case of any changes to the protocol, the investigators, clinical research assistants, the ethics committee, and the sponsor will be immediately informed via email.

Discussion

Our protocol uses a randomized, waitlist-controlled trial design to achieve two primary objectives: (1) systematically evaluate the efficacy and safety of CP003 for CFS/ME and post–COVID-19 fatigue management and (2) elucidate clinical correlations between CP003 and key biomarkers of inflammation, immune function, and oxidative stress to reveal potential therapeutic mechanisms. This design was selected based on 3 key considerations. First, the waitlist-controlled approach balances scientific rigor with ethical considerations by ensuring that all participants eventually receive the intervention while avoiding placebo-related ethical concerns in this vulnerable population [27]. Second, our dual clinical-biomarker assessment strategy overcomes limitations of purely subjective measures by combining validated questionnaires (ie, the Chalder Fatigue Scale) with quantitative biomarker analysis (eg, interleukin-6 and superoxide dismutase), as recommended by the International Consensus Criteria [28]. Third, our sample size (N=130) was powered to detect clinically meaningful 20% fatigue score reductions (α=.05 and power=80%) while accounting for 15% attrition, ensuring both statistical robustness and practical feasibility [29].

However, this study is subject to several limitations. The heterogeneity inherent in fatigue-related disorders, particularly the symptomatic overlap with conditions such as depression and fibromyalgia, may influence treatment response and generalizability. Although standardized fatigue scales were used, the primary efficacy evaluation remains dependent on patient-reported outcomes in the absence of definitive biomarkers. In addition, the 12-week study duration, comprising a 6-week intervention and 6-week follow-up periods, may be inadequate to fully assess the long-term benefits or potential risks associated with prolonged lingzhi administration. Future studies with extended observation periods and more rigorously stratified participant cohorts are warranted to corroborate these initial findings.

In conclusion, this trial adds to the expanding evidence base for natural interventions in fatigue-related conditions and provides new insights into the therapeutic potential of lingzhi as a TCM-derived agent. Should efficacy be established, subsequent research ought to prioritize several directions: (1) refining dosage and treatment protocols through dedicated dose-finding studies, (2) conducting comprehensive long-term safety evaluations in extended longitudinal settings, and (3) elucidating the underlying mechanisms in broader and more diverse patient populations. This study emphasizes the value of integrating traditional medicine with contemporary research frameworks to tackle complex, multifactorial syndromes, such as CFS/ME and post–COVID-19 fatigue syndrome, while underscoring the promising yet underexplored role of TCM-based formulations, such as lingzhi.

Acknowledgments

The authors gratefully acknowledge the Innovation and Technology Commission of Hong Kong for the research grant support (ITF/PRP/029/24FX). The authors attest that no generative artificial intelligence technologies were used in the generation of this manuscript.

Funding

This study was funded by the Innovation and Technology Commission of Hong Kong (grant ITF/PRP/029/24FX).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

GC contributed to the study design and development of the proposal. YX contributed to drafting the manuscript. YX and TLF contributed to the recruitment and acquisition of study data. YF contributed to reviewing and editing the manuscript, supervision, resources, data curation, and conceptualization. All authors read and approved the final manuscript.

Conflicts of Interest

NW received funding from Chinese Pharmaceuticals, the manufacturer of CP003, for partial financial support of this trial. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Peer-review report from The Innovation and Technology Commission Grant (Hong Kong).
[PDF File (Adobe PDF File), 65 KB - resprot_v15i1e82633_app1.pdf]

XSL•FO
RenderX

References

1. Avellaneda Fernández A, Pérez Martín A, Izquierdo Martínez M, Arruti Bustillo M, Barbado Hernández FJ, de la Cruz Labrado J, et al. Chronic fatigue syndrome: aetiology, diagnosis and treatment. BMC Psychiatry 2009 Oct 23;9(Suppl 1):1-11 [FREE Full text] [doi: 10.1186/1471-244X-9-S1-S1] [Medline: 19857242]
2. Lim EJ, Ahn YC, Jang ES, Lee SW, Lee SH, Son CG. Systematic review and meta-analysis of the prevalence of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME). J Transl Med 2020 Feb 24;18(1):100 [FREE Full text] [doi: 10.1186/s12967-020-02269-0] [Medline: 32093722]
3. Bateman L, Bested AC, Bonilla HF, Chheda BV, Chu L, Curtin JM, et al. Myalgic encephalomyelitis/chronic fatigue syndrome: essentials of diagnosis and management. Mayo Clin Proc 2021 Nov;96(11):2861-2878 [FREE Full text] [doi: 10.1016/j.mayocp.2021.07.004] [Medline: 34454716]
4. Poenaru S, Abdallah SJ, Corrales-Medina V, Cowan J. COVID-19 and post-infectious myalgic encephalomyelitis/chronic fatigue syndrome: a narrative review. Ther Adv Infect Dis 2021 Apr 20;8:20499361211009385 [FREE Full text] [doi: 10.1177/20499361211009385] [Medline: 33959278]
5. Mensah FK, Bansal AS, Ford B, Cambridge G. Chronic fatigue syndrome and the immune system: where are we now? Neurophysiol Clin 2017 Apr;47(2):131-138. [doi: 10.1016/j.neucli.2017.02.002] [Medline: 28410877]
6. Wood E, Hall KH, Tate W. Role of mitochondria, oxidative stress and the response to antioxidants in myalgic encephalomyelitis/chronic fatigue syndrome: a possible approach to SARS-CoV-2 'long-haulers'? Chronic Dis Transl Med 2021 Mar;7(1):14-26 [FREE Full text] [doi: 10.1016/j.cdtm.2020.11.002] [Medline: 33251031]
7. Bhatti JS, Bhatti GK, Reddy PH. Mitochondrial dysfunction and oxidative stress in metabolic disorders - a step towards mitochondria based therapeutic strategies. Biochim Biophys Acta Mol Basis Dis 2017 May;1863(5):1066-1077 [FREE Full text] [doi: 10.1016/j.bbadis.2016.11.010] [Medline: 27836629]
8. Milrad SF, Hall DL, Jutagir DR, Lattie EG, Ironson GH, Wohlgemuth W, et al. Poor sleep quality is associated with greater circulating pro-inflammatory cytokines and severity and frequency of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) symptoms in women. J Neuroimmunol 2017 Feb 15;303:43-50 [FREE Full text] [doi: 10.1016/j.jneuroim.2016.12.008] [Medline: 28038892]
9. McCarthy MJ. Circadian rhythm disruption in myalgic encephalomyelitis/chronic fatigue syndrome: implications for the post-acute sequelae of COVID-19. Brain Behav Immun Health 2022 Jan 10;20:100412 [FREE Full text] [doi: 10.1016/j.bbih.2022.100412] [Medline: 35465246]
10. Nalbandian A, Sehgal K, Gupta A, Madhavan MV, McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. Nat Med 2021 Apr;27(4):601-615 [FREE Full text] [doi: 10.1038/s41591-021-01283-z] [Medline: 33753937]
11. Kwok Y, Ng KF, Li CC, Lam CC, Man RY. A prospective, randomized, double-blind, placebo-controlled study of the platelet and global hemostatic effects of Ganoderma lucidum (Ling-Zhi) in healthy volunteers. Anesth Analg 2005 Aug;101(2):423-426. [doi: 10.1213/01.ANE.0000155286.20467.28] [Medline: 16037156]
12. Wicks SM, Tong R, Wang CZ, O'Connor M, Karrison T, Li S, et al. Safety and tolerability of Ganoderma lucidum in healthy subjects: a double-blind randomized placebo-controlled trial. Am J Chin Med 2007;35(3):407-414. [doi: 10.1142/S0192415X07004928] [Medline: 17597499]
13. Noguchi M, Kakuma T, Tomiyasu K, Kurita Y, Kukihara H, Konishi F, et al. Effect of an extract of Ganoderma lucidum in men with lower urinary tract symptoms: a double-blind, placebo-controlled randomized and dose-ranging study. Asian J Androl 2008 Jul;10(4):651-658 [FREE Full text] [doi: 10.1111/j.1745-7262.2008.00336.x] [Medline: 18097503]
14. Noguchi M, Kakuma T, Tomiyasu K, Yamada A, Itoh K, Konishi F, et al. Randomized clinical trial of an ethanol extract of Ganoderma lucidum in men with lower urinary tract symptoms. Asian J Androl 2008 Sep;10(5):777-785 [FREE Full text] [doi: 10.1111/j.1745-7262.2008.00361.x] [Medline: 18097505]
15. Sargowo D, Ovianti N, Susilowati E, Ubaidillah N, Widya Nugraha A, Vitriyaturrida, et al. The role of polysaccharide peptide of Ganoderma lucidum as a potent antioxidant against atherosclerosis in high risk and stable angina patients. Indian Heart J 2018;70(5):608-614 [FREE Full text] [doi: 10.1016/j.ihj.2017.12.007] [Medline: 30392496]
16. Collado Mateo D, Pazzi F, Domínguez Muñoz FJ, Martín Martínez JP, Olivares PR, Gusi N, et al. Ganoderma lucidum improves physical fitness in women with fibromyalgia. Nutr Hosp 2015 Nov 01;32(5):2126-2135 [FREE Full text] [doi: 10.3305/nh.2015.32.5.9601] [Medline: 26545669]
17. Gao Y, Zhou S, Jiang W, Huang M, Dai X. Effects of ganopoly (a Ganoderma lucidum polysaccharide extract) on the immune functions in advanced-stage cancer patients. Immunol Invest 2003 Aug;32(3):201-215. [doi: 10.1081/imm-120022979] [Medline: 12916709]
18. Chu TT, Benzie IF, Lam CW, Fok BS, Lee KK, Tomlinson B. Study of potential cardioprotective effects of Ganoderma lucidum (Lingzhi): results of a controlled human intervention trial. Br J Nutr 2012 Apr;107(7):1017-1027. [doi: 10.1017/S0007114511003795] [Medline: 21801467]
19. Klupp NL, Kiat H, Bensoussan A, Steiner GZ, Chang DH. A double-blind, randomised, placebo-controlled trial of Ganoderma lucidum for the treatment of cardiovascular risk factors of metabolic syndrome. Sci Rep 2016 Aug 11;6:29540 [FREE Full text] [doi: 10.1038/srep29540] [Medline: 27511742]

XSL•FO
RenderX

20. Fukuda K, Straus SE, Hickie I, Sharpe MC, Dobbins JG, Komaroff A. The chronic fatigue syndrome: a comprehensive approach to its definition and study. International Chronic Fatigue Syndrome Study Group. Ann Intern Med 1994 Dec 15;121(12):953-959. [doi: 10.7326/0003-4819-121-12-199412150-00009] [Medline: 7978722]
21. Morriss RK, Wearden AJ, Mullis R. Exploring the validity of the Chalder Fatigue scale in chronic fatigue syndrome. J Psychosom Res 1998 Nov;45(5):411-417. [doi: 10.1016/s0022-3999(98)00022-1] [Medline: 9835234]
22. Hewlett S, Dures E, Almeida C. Measures of fatigue: Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire (BRAF MDQ), Bristol Rheumatoid Arthritis Fatigue Numerical Rating Scales (BRAF NRS) for severity, effect, and coping, Chalder Fatigue Questionnaire (CFQ), Checklist Individual Strength (CIS20R and CIS8R), Fatigue Severity Scale (FSS), Functional Assessment Chronic Illness Therapy (Fatigue) (FACIT-F), Multi-Dimensional Assessment of Fatigue (MAF), Multi-Dimensional Fatigue Inventory (MFI), Pediatric Quality Of Life (PedsQL) Multi-Dimensional Fatigue Scale, Profile of Fatigue (ProF), Short Form 36 Vitality Subscale (SF-36 VT), and Visual Analog Scales (VAS). Arthritis Care Res (Hoboken) 2011 Nov;63(S11):S263-S286 [FREE Full text] [doi: 10.1002/acr.20579] [Medline: 22588750]
23. Yang M, Keller S, Lin JM. Psychometric properties of the PROMIS® Fatigue Short Form 7a among adults with myalgic encephalomyelitis/chronic fatigue syndrome. Qual Life Res 2019 Dec;28(12):3375-3384. [doi: 10.1007/s11136-019-02289-4] [Medline: 31506915]
24. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
25. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
26. Nøstdahl T, Bernklev T, Fredheim OM, Paddison JS, Raeder J. Defining the cut-off point of clinically significant postoperative fatigue in three common fatigue scales. Qual Life Res 2019 Apr;28(4):991-1003. [doi: 10.1007/s11136-018-2068-0] [Medline: 30506178]
27. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
28. Carruthers BM, van de Sande MI, De Meirleir KL, Klimas NG, Broderick G, Mitchell T, et al. Myalgic encephalomyelitis: International Consensus Criteria. J Intern Med 2011 Oct;270(4):327-338 [FREE Full text] [doi: 10.1111/j.1365-2796.2011.02428.x] [Medline: 21777306]
29. Sim J, Lewis M. The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. J Clin Epidemiol 2012 Mar;65(3):301-308. [doi: 10.1016/j.jclinepi.2011.07.011] [Medline: 22169081]

Abbreviations

CFS: chronic fatigue syndrome
HADS: Hospital Anxiety and Depression Scale
ME: myalgic encephalomyelitis
PROMIS: Patient-Reported Outcomes Measurement Information System
REDCap: Research Electronic Data Capture
SF-36: 36-Item Short Form Health Survey
TCM: traditional Chinese medicine

Edited by J Sarvestan, A Mavragani; submitted 19.Aug.2025; this is a non–peer-reviewed article;accepted 09.Dec.2025; published 20.Jan.2026.

Please cite as:
Xun Y, Fong TL, Chen G, Feng Y, Chan L, Wang N
Effectiveness and Safety of a Supplement Containing a Pharmacologically Active Basidiomycete Mushroom for Chronic Fatigue and Post–COVID-19 Fatigue Syndrome: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e82633
URL: https://www.researchprotocols.org/2026/1/e82633
doi:10.2196/82633
PMID:

©Yunqing Xun, Tung Leong Fong, Guang Chen, Yibin Feng, Linda Chan, Ning Wang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Combining Noninvasive Brain Stimulation and Physiotherapy to Improve the Management of Chronic Low Back Pain in Veterans: Protocol for a Multi-Arm Randomized Controlled Trial

Frederique Dupuis[1,2], PT, PhD; Yannick Tousignant-Laflamme[3,4], PT, PhD; Pascale Marier Deschênes[1,2], PhD; Philippe Fournier[1,2], PhD; Ildephonse Nduwimana[1,2], PT, PhD; Orlane Ballot[1], PhD; Loris Loisel[1,2], PT; Anne Marie Pinard[1,2,5], MD; Luc Lacombe[6]; Pierre Langevin[1,2], PT, PhD; Alain Gaumond[7], PT, MSc; Jean-Sébastien Roy[1,2], PT, PhD; Hugo Massé-Alarie[1,2], PT, PhD

[1]Centre for Interdisciplinary Research in Rehabilitation and Social Integration, Quebec, QC, Canada
[2]École des Sciences de la Réadapation, Faculté de Médecine, Université Laval, Québec, QC, Canada
[3]School of Rehabilitation, Université de Sherbrooke, Sherbrooke, QC, Canada
[4]Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, QC, Canada
[5]Chronic Pain Service, CHU de Québec-Université Laval, Quebec, Canada
[6]Canadian Armed Forces, Québec, QC, Canada
[7]Espace Pleine Conscience, Québec, QC, Canada

Corresponding Author:
Hugo Massé-Alarie, PT, PhD
Centre for Interdisciplinary Research in Rehabilitation and Social Integration
525, Wilfrid-Hamel Boulevard
Wing H Suite 1300
Quebec, QC, G1M 2S8
Canada
Phone: 1 418 649 3735
Email: hugo.masse-alarie@fmed.ulaval.ca

Abstract

Background: Low back pain (LBP) is the most common chronic pain condition in veterans, but the effectiveness of standard management approaches is modest. Addressing the psychological risk factors of chronic pain that are often observed in this population (eg, anxiety, depression, stress and mood disorders) may be important to enhance outcomes. Psychologically informed physiotherapy (PiP) identifies and mitigates the negative impacts of emotional and cognitive factors alongside the biomedical aspects of chronic LBP to improve physical functioning and has shown promising results in this population. However, residual pain and disability often persist in veterans. The combination of PiP with repetitive transcranial magnetic stimulation (rTMS) to the prefrontal cortex may enhance its effectiveness by modulating cognition, emotion, and pain perception.

Objective: The aim of this study is to compare the effects of (1) combining active rTMS with PiP, (2) combining sham rTMS with PiP, and (3) usual physiotherapy (UP) on physical functioning in veterans with chronic LBP and comorbid psychological risk factors. Secondary objectives include comparing their effect on pain intensity, quality of life, depression symptoms, pain catastrophizing, movement pain–related fear, self-efficacy, medication use, and posttraumatic stress disorder symptoms.

Methods: Ninety-six veterans with chronic LBP and comorbid psychological risks factors of pain will be enrolled in this 3-arm parallel randomized controlled trial. Individuals will be allocated to receive an 8-week intervention of (1) active rTMS + PiP, (2) sham rTMS + PiP, or (3) UP. Online self-administered questionnaires will be completed at baseline, 2, 8, and 26 weeks after the first treatment session. A linear mixed model will be used to assess the treatment effects by using intention-to-treat analyses. We hypothesize that active rTMS + PiP will be more effective than sham rTMS + PiP and that active PiP + rTMS or sham rTMS will be more effective than UP.

Results: Ethics approval was obtained in January 2025, and participating physiotherapists completed the 2-day PiP training in May 2025. Participants have been recruited since June 2025. As of December 2025, 28 participants have been included, and recruitment is expected to continue up to June 2027, targeting the inclusion of approximately 4 new participants per month. Follow-up should be completed by December 2027, and results will be analyzed. The results of this randomized controlled trial should be published and available in June 2028.

XSL•FO
RenderX

Conclusions: It is paramount to identify innovative and effective interventions for the management of chronic LBP in veterans. This study will provide new evidence on the effectiveness of two innovative interventions targeting cognitive and emotional factors of pain (ie, PiP and rTMS). If our hypothesis is confirmed, it could motivate changes in clinical practice and improve the quality of life of veterans living with chronic LBP.

Trial Registration: ClinicalTrials.gov NCT06999772; https://clinicaltrials.gov/study/NCT06999772

International Registered Report Identifier (IRRID): PRR1-10.2196/78952

(JMIR Res Protoc 2026;15:e78952) doi:10.2196/78952

KEYWORDS

low back pain; chronic pain; rehabilitation; transcranial magnetic stimulation; psychosocial; physiotherapy; military

Introduction

Chronic low back pain (CLBP) stands as the leading cause of years lived with disability globally [1,2]. Addressing this widespread chronic condition poses a notable challenge to both the public and military health care systems. The effectiveness of standard management approaches for CLBP has been found to be modest in both general [3] and military/veteran [4,5] populations, further negatively impacting the quality of life of these people.

In military veterans, CLBP is the most prevalent painful condition [6,7], with prevalence rates reaching 45% [7]. Moreover, veterans are more likely to experience more severe pain intensity and pain-related disability than the civilian population [8]. The poor response to current interventions in this highly impacted population exacerbates the current opioid crisis [9], urging the need for new, effective, and noninvasive interventions.

The limited effectiveness of current interventions suggests that they may not adequately target specific mechanisms underlying CLBP. In the last few years, a paradigm shift has been noted in CLBP management, transitioning from a strictly biomedical perspective to a more comprehensive biopsychosocial approach [10]. Moreover, our understanding of the mechanisms of CLBP has evolved. For example, evidence reveals abnormal brain functioning among individuals with CLBP [11], suggesting the involvement of alterations within the central nervous system in certain people [12]. As the painful symptoms persist, a shift from circuits processing nociceptive inputs to circuits processing emotional states may occur in those at the highest risk to develop chronic pain [13,14]. The corticolimbic system (including the prefrontal cortex [PFC] and limbic system), which is critical in the regulation, processing, and control of emotions, plays a significant role in the inhibition of pain through cognitive control mechanisms within the cortex [14]. Enhanced cognitive and emotional valuation of pain may act as an amplifier of nociception [14]. This transition aligns with the identified multidimensional factors contributing to chronicity [15].

Psychological disorders like depression and anxiety and factors like pain catastrophizing, and movement pain–related fear have repeatedly been linked to the persistence of LBP and long-term disability [16,17]. These psychological factors represent higher-order cognitive and emotional processes occurring, *inter alia*, within the PFC [18], and alterations in brain processing may exacerbate nociception and contribute to the persistence of pain. This physiological phenomenon is referred to as nociplastic pain mechanisms [19]. Considering that psychological factors are more prevalent among veterans compared to civilians [7,20] and are associated with the emergence of chronic pain in this population [7,20-22], there might be a critical need for interventions identifying and mitigating these factors [23]. However, they are frequently overlooked in standard health care practices [24,25].

Usual physiotherapy (UP) constitutes a standard health care intervention for veterans experiencing LBP, and although being effective to improve symptoms and disability compared to no intervention, it often fails to consider the psychological prognostic factors of persistent pain. Psychologically informed practice utilizes a biopsychosocial approach to recognize and address emotional and cognitive components associated with the condition, using different therapeutic modalities such as education, gradual exposure to movement, and stress management strategies [26]. When delivered by physiotherapists combined with UP care, the psychologically informed practice is called psychologically informed physiotherapy (PiP) [26]. PiP represents a promising intervention for the management of CLBP, especially in veterans, as it targets the highly prevalent psychological factors present in this population. This approach is currently supported by evidence from civilian populations: meta-analyses have shown that integrating psychological interventions with UP leads to greater improvements in pain intensity and physical functioning in people with CLBP [27], including those presenting comorbid psychological factors [28], when compared to UP alone.

In addition to PiP, noninvasive brain stimulation over the PFC by using repetitive transcranial magnetic stimulation (rTMS) may be useful to modulate the cognitive and emotional processing involved in pain mechanisms [29,30]. This approach is supported by large randomized controlled trials in psychiatric populations [29]. rTMS depolarizes neurons under the stimulating coil, and modulates brain excitability [31], leading to changes in network connectivity [29] that have been linked to reductions in symptoms such as rumination and negative bias [29,31]. Through potential similar mechanisms, rTMS over the PFC has been shown to reduce pain intensity in cases of neuropathic pain [32], a finding that may relate to the high prevalence of psychological comorbidities among people with chronic pain [16,17]. To date, most studies that have investigated the use of rTMS for chronic musculoskeletal pain stimulated the motor cortex (M1), yielding conflicting results [33-35]. Stimulating the PFC could optimize results by targeting brain

XSL•FO
RenderX

areas specifically involved in pain modulation. A pilot randomized controlled trial (RCT) (n=20) studied the efficacy of noninvasive brain stimulation over a prefrontal cortical area in veterans with CLBP, and this stimulation resulted in a large improvement in the physical functioning and depressive symptoms compared to a placebo [36]. PiP and rTMS are two promising interventions specifically targeting psychological factors and cortical mechanisms contributing to CLBP in veterans. Despite these promising pilot results, good quality RCTs investigating the effectiveness of rTMS over the PFC for chronic LBP are still needed.

PiP and rTMS are two promising interventions specifically targeting psychological factors and cortical mechanisms contributing to CLBP in veterans. Combining psychologically informed clinical intervention to rTMS could enhance treatment outcomes, as they could both modulate cognitive and emotional processing involved in pain mechanisms, through different pathways [29]. However, the combination of these two interventions has never been studied in chronic musculoskeletal pain and compared to usual care.

Therefore, the main objective of this study will be to compare the effectiveness of (1) combining PiP + active rTMS, (2) PiP + sham rTMS, and (3) UP alone in improving the physical functioning in veterans experiencing CLBP and comorbid psychological factors. The secondary objectives will be to compare their effectiveness on pain intensity, quality of life, movement pain–related fear, pain catastrophizing, self-efficacy, depressive symptoms, medication use, and posttraumatic stress disorder symptoms. We hypothesize that the combination of active rTMS + PiP will result in a larger effect on primary and secondary outcomes compared to PiP + sham rTMS and UP alone. In addition, we hypothesize that PiP with active rTMS or sham rTMS will be more effective than UP.

Methods

Study Design

This protocol is reported according to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines [37]. Participants included in this 3-arm, parallel-group RCT will be randomized to one of the 3 groups (PiP + active rTMS, PiP + sham rTMS, or UP) and will undergo an 8-week intervention. Outcomes measurements will occur at baseline, 2, 8, and 26 weeks after the beginning of the intervention (Figure 1). The protocol of this RCT has been registered on ClinicalTrials.gov (NCT06999772).

Figure 1. Flowchart of the participants. PiP: psychologically informed physiotherapy; rTMS: repetitive transcranial magnetic stimulation; UP: usual physiotherapy.

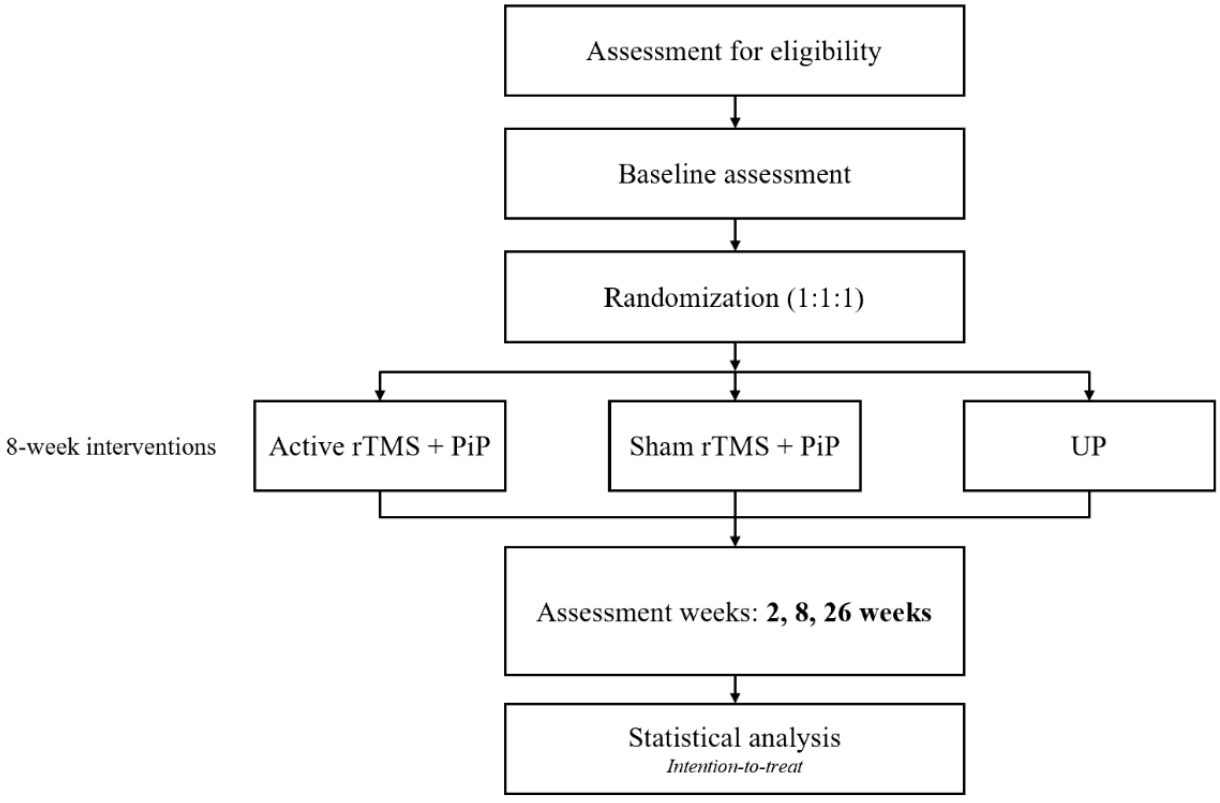

Ethical Considerations

Ethics approval has been granted by the Comité d'éthique de la Recherche Sectoriel en Réadaptation et Intégration Sociale du CIUSSS de la Capitale-Nationale (Rehabilitation and Social Integration section; 2025-3215). All participants will provide a written informed consent prior to their participation in the study and will be free to opt out whenever they wish to.

XSL•FO
RenderX

Participants will receive compensation of CAD $40 (US $29) after the first 4 weeks of treatment and CAD $40 (US $29) after the last 4 weeks to offset the costs incurred for traveling to appointments. In addition, CAD $20 (US $15) will be given to participants who complete their 6-month follow-up questionnaires in order to increase the retention rate. All collected data using REDCap will be deidentified, and any participant information at recruitment will be kept on an online spreadsheet secured with a password. Only members in the research team have access to this information.

Study Participants

Veterans with nonspecific CLBP and comorbid psychological factors will be recruited through social media and recruitment emails. A patient partner (LL) will help the team target relevant groups on social media and organizations for the recruitment. Nonspecific CLBP will be defined as LBP that has persisted for at least 3 months and resulted in pain on at least half the days in the past 6 months [38]. A minimal level of disability of ≥15% on the Oswestry Disability Index (ODI) will also ensure that included participants are limited by LBP in their daily activities. The presence of psychological risk factors will be defined as a score of ≥4 on the Star Back Screening Tool [39]. Potential participants will be excluded if they have (1) a change of drug dosage in the last month (ie, drug for the treatment of pain or mental health disorders), (2) any cause of specific LBP (eg, fracture, cancer), (3) any specific rTMS-related exclusion criteria (eg, previous seizure, metal in the head, history of stroke), (4) any confirmed drug or alcohol abuse problem [40], and (5) they cannot understand written French (for self-administrated questionnaires). A research professional or a PhD student will obtain written consent of all the participants before the start of the study after receiving information on the nature and objectives of the study.

Randomization and Blinding

A randomization list will be generated using a computer random number generator by blocks of 3, 6, or 9 by an independent research assistant. The randomization list will be uploaded on a Research Electronic Data Capture (REDCap, Vanderbilt University) platform by using the randomization module. Once the randomization list will be uploaded, REDCap will automatically retrieve the group allocation. This method digitally replicates the allocation concealment and prevents selection bias without having to manipulate sealed envelopes. Randomization will be stratified by self-reported gender and physical functioning (ODI score ≤ or > 40).

To ensure that no assessor could bias data collection, all patient-reported outcome measures (PROMs) will be assessed using REDCap. Although it is not possible to blind participants to a physical intervention (UP and PiP), they will be blinded to rTMS group allocation (active vs sham). The rTMS intervention will be delivered at the research center, and the physiotherapy interventions will be delivered in local physiotherapy clinics. The 2 interventions (PiP or UP) will be provided in 2 different physiotherapy clinics, reducing the risks of potential group contamination. Furthermore, the physiotherapists providing PiP will not be informed of the type of rTMS (sham vs active), and participants will be asked not to discuss rTMS sessions with their treating physiotherapist. The statistician will be blinded to group allocation. Overall, the RCT will be triple-blind for rTMS (ie, participants, treating physiotherapist, statistician) and single-blind for physiotherapy (ie, statistician). At week 8, participants in the PiP groups will be asked to determine which stimulation they believed they received (sham, active, or don't know) and to provide the reasons for their choice through a blinding evaluation questionnaire. At any time during the trial, if any severe adverse event occurs (ie, seizure), the research team and the participant will be informed of its group allocation to appropriately manage this event.

Intervention

All participants will undergo an 8-week intervention. All participants will receive 6 physiotherapy sessions within 8 weeks, while the rTMS + PiP and sham rTMS + PiP groups will receive 11 additional rTMS sessions in the same period. Figure 2 shows the intervention timeline for the active rTMS + PiP and sham rTMS + PiP groups, while Figure 3 shows the intervention timeline for the UP group. A detailed description of each intervention is provided below.

Figure 2. Intervention timelines for the active rTMS + PiP and sham rTMS + PiP groups. Each bar represents a session of intervention. A black bar represents an rTMS (or sham) session offered at the research center, while a grey bar represents a physiotherapy session offered in one of the local clinics. PiP: psychologically informed physiotherapy; rTMS: repetitive transcranial magnetic stimulation.

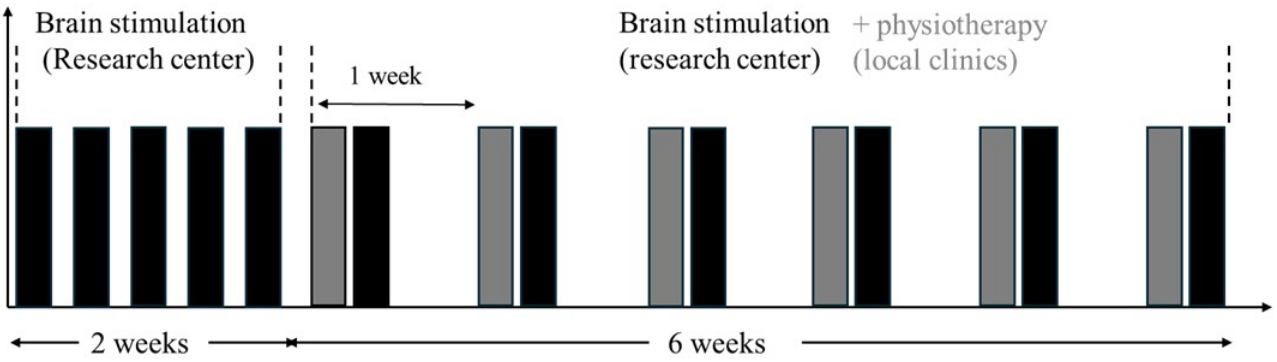

XSL•FO
RenderX

Figure 3. Intervention timelines for the usual care group. Each bar represents a physiotherapy session provided in one of the participating local clinics. In the first 4 weeks, 4 sessions were offered at 1-week interval. In the last 4 weeks of the intervention, 2 additional sessions were offered at 2-weeks interval.

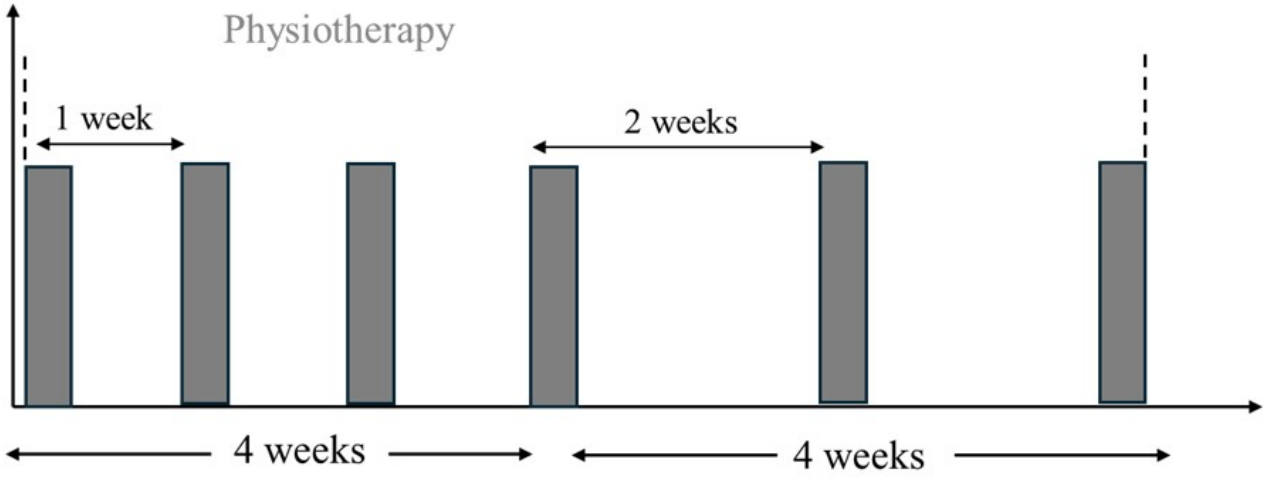

Active rTMS

Participants will receive 11 sessions of rTMS. For the first 2 weeks, participants will attend 5 brain stimulation sessions that will serve as an induction phase to optimize the effects of rTMS [41]. During weeks 3 to 8, 6 rTMS sessions will be provided, once per week. Each rTMS session will last 30-40 minutes (10-20 min of installation + 20 min of stimulation). Stimulation parameters will be obtained by measuring the motor threshold of the right first dorsal interosseous muscle, and then the rTMS intensity will be set at 110% of this motor threshold [32,42]. High-frequency rTMS consisting of 60 trains of 5 seconds at 10 Hz with 15-second intertrain intervals, for a total of 3000 pulses per session, will be used [32]. High-frequency rTMS has been shown to reduce chronic pain [32] and depression symptoms [43]. The coil will be positioned over the left dorsolateral PFC by using the BeamF3 methods [44].

Sham rTMS

For the sham stimulation, a sham coil looking like the regular rTMS coil but with a magnetic shield blocking the magnetic field will be used to produce the auditory effects and vibration similar to the active coil. This type of sham has been shown to be effective to blind participants [32]. All other parameters will be the same as the active rTMS.

PiP Intervention

The objective of PiP is to identify the biopsychosocial factors that may impede recovery and to address these factors within the physiotherapy scope of practice. To standardize PiP delivery, treating physiotherapists providing PiP will receive a 2-day training provided by an expert in PiP who also has expertise with military personnel. They will learn to use strategies such as the establishment of common goals and therapeutic alliance, the use of the behavior change model, motivational interviewing approach, education on pain neuroscience, gradual exposure to movement, and stress management strategies (eg, breathing techniques, meditation) [45]. The training will also include a section on the military culture to foster therapeutic alliance with veterans and optimize the treatment adherence. It provides foundational knowledge about the Canadian Armed Forces, including its branches, core values, and professional expectations. It supports participants in understanding how military culture influences health behaviors and offers guidance on how this awareness can inform clinical assessments and interventions with veterans. These physiotherapists will also be allowed to use standard physiotherapy modalities presented in the UP section, based on their clinical judgement. However, they will be encouraged to assess and address psychosocial factors by using strategies learned during their training. This training has been shown to be effective in leading to changes in the clinical practice of physiotherapists, leading to a greater use of modalities targeting psychological factors compared to UP [45]. Participants in the PiP + active rTMS and PiP + sham rTMS groups will have access to a website offering support information on topics such as pain neuroscience education, common myths and misconceptions about back pain, the psychological factors influencing pain, self-management strategies, and healthy lifestyle habits (eg, sleep and nutrition). This site will be freely available to participants during the study. Physiotherapists will be asked to discuss one topic at each session with the participants to assess and reframe (if necessary) their conceptualization of these concepts.

UP Intervention

This pragmatic group aims to represent real-world clinical practice, including interventions commonly used in physiotherapy to manage CLBP. No training will be provided, and all modalities typically used by physiotherapists in the province of Quebec will be allowed, including but not limited to manual therapy, active exercises, education, stretching, etc. Each physiotherapy session, including PiP and UP, will last between 30 and 45 minutes. Physiotherapists providing UP and PiP will all complete a checklist to indicate the interventions provided at each treatment session.

XSL•FO
RenderX

Outcomes

Overview

All PROMs will be assessed and filled online by using REDCap. Sociodemographic information will first be collected before randomization. PROMs baseline measurements will be measured after randomization, just prior to the beginning of the treatment. This decision was taken because the time before care initiation can vary depending on the groups (rTMS vs no rTMS) and the clinics delivering the different physiotherapy approaches (PiP vs UP). A longer delay before care initiation for a specific group may influence the outcomes trajectories. The baseline primary outcome will be measured 48 hours prior to the beginning of treatment. Nonetheless, the primary outcome will also be collected before randomization (during enrollment for screening purpose) to ensure similar between-group evolution until baseline measurement. However, participants will be informed on their allocated groups only after baseline PROMs to avoid any anticipation bias. PROMs will then be measured again via email to participants at 2, 8, and 26 weeks after the first treatment session. Figure 4 presents the SPIRIT diagram of this study.

Figure 4. SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) diagram. PiP: psychologically informed physiotherapy; rTMS: repetitive transcranial magnetic stimulation.

Timepoint		Enrollment		Postrandomization				
		$-T_i$ to 0	0	T_1 *48 h prior to first treatment session*	*First treatment session*	T_2 *2 weeks*	T_3 *8 weeks*	T_4 *26 weeks*
Enrollment								
	Eligibility criteria	√						
	Informed consent	√						
	Sociodemographic characteristics		√					
	Randomization		√					
Interventions								
	Pip + rtms				◆—	—	—◆	
	Pip + sham rtms				◆—	—	—◆	
	Usual physiotherapy care				◆—	—	—◆	
Assessments								
	Oswestry Disability Index	√		√		√	√	√
	Numeric Pain Rating Scale			√		√	√	√
	EQ-5D-5L			√		√	√	√
	Pain Catastrophizing Scale			√		√	√	√
	Pain Health Questionnaire 9			√		√	√	√
	Pain Self-Efficacy Questionnaire			√		√	√	√
	Medication Quantification Scale 4			√		√	√	√
	Posttraumatic Stress Disorder Checklist			√		√	√	√
	Tampa Scale of Kinesiophobia			√		√	√	√
	Patient Acceptable Symptoms State			√		√	√	√
	Global Rating of Change					√	√	√
	Blinding evaluation questionnaire						√	
	Adverse events					√	√	

Primary Outcomes

The primary outcome will be physical functioning, which will be evaluated using ODI, a 10-item questionnaire (0=no disability; 100=extremely severe disability) that is valid and reliable and recommended to evaluate physical functioning in LBP [46]. The primary time point will be at 8 weeks.

Secondary Outcomes

The average pain intensity over the last week will be assessed with the 11-point numerical pain rating scale (0: “no pain” to 10: “worst pain you can imagine”). The numerical pain rating scale is reliable and responsive to change and recommended in LBP trials [46]. The quality of life will be measured using the EQ-5D-5L, which is a generic questionnaire evaluating five dimensions of health-related quality of life, each rated on a 5-point scale [47]. Depression symptoms will be measured using the Patient Health Questionnaire (9 items), which is reliable and valid [48]. Pain catastrophizing will be measured using the valid and reliable Pain Catastrophizing Scale [49]. Movement pain–related fear will be assessed with the Tampa Scale of Kinesiophobia questionnaire (13 items), which has good reliability and validity [50]. Self-efficacy will be measured using the Pain Self-Efficacy Questionnaire, which is valid, reliable, and responsive to change in a population with CLBP [51]. Medication use will be measured using the Medication

XSL•FO
RenderX

Quantification Scale Version IV that quantifies 22 drug classes used in pain management (eg, opioid) with a detriment weight [52]. Posttraumatic stress disorder symptoms will be assessed with the Posttraumatic Stress Disorder Checklist for Diagnostic and Statistical Manual of Mental Disorders, which is a 20-item self-report questionnaire that is consistent and valid [53]. The Patient Acceptable Symptom State will be used to determine if the participants consider their current functional and symptom states as satisfactory or not. In addition, the Global Rating of Change, an 11-point scale ranging from –5 (a great deal worst) to 5 (a great deal better) will be used to measure the perceived change of health status after the intervention (ie, at 8 weeks and 26 weeks). Finally, all adverse events and reasons for drop-out will be recorded and reported.

Analysis

Sample Size Calculation

The trial is designed to detect a 9-point (SD 10) between-group difference in physical functioning at 8 weeks, as measured using the minimal important difference of the ODI [54]. The standard deviation was estimated from our previous RCTs [45,55]. Considering 20% loss at follow-up, 1-β=.90, and α=.05, 32 participants per group will be recruited (n_{total}=96).

Statistical Analysis

A linear mixed model will be used to compare treatment effects on outcomes, using intention-to-treat analyses. Group (active rTMS + PiP, sham rTMS + PiP, and UP) and time (baseline, 2, 8, 26 weeks) will be entered as fixed factors, participants as random intercept, and stratification factors (gender and baseline ODI score) as covariate, to determine the presence of any significant group × time interaction. Missing values will be handled with multiple imputations if necessary, and sensitivity analyses will be completed to confirm/infirm the results.

Trial Management

The study will be coordinated by the principal investigator. A research professional will be the head coordinator for this project. He and a PhD student will be responsible for the day-to-day management of the RCT by overseeing the recruitment efficacy (eg, rate of recruitment), regularly contacting the participating physiotherapists, and preparing and managing the data collection in REDCap.

Results

Ethics approval was obtained in January 2025, and participating physiotherapists completed the 2-day PiP training in May 2025. Participants have been recruited since June 2025. As of December 2025, 28 participants have been included, and recruitment is expected to continue up to June 2027, targeting the inclusion of approximately 4 new participants per month. Follow-up should be completed by December 2027, and results will be analyzed. The results of this randomized controlled trial should be published and available in June 2028.

Discussion

This study will investigate the effectiveness of combining psychologically informed physiotherapy and rTMS over the dorsolateral PFC for veterans with CLBP. The management of veterans with CLBP is challenging. Although LBP is the most prevalent source of chronic pain, only modest therapeutic effects are observed [4,5]. Furthermore, veterans with CLBP experience more severe pain and limitations than civilians and are more likely to be prescribed opioids, increasing the risk of chronic opioid use, overdose, and death [9,56]. Identifying innovative, effective, and noninvasive interventions for the management of CLBP in this population is paramount.

The presence of psychological factors such as anxiety, depression, and stress disorders has been frequently observed in veterans living with chronic pain, and growing evidence suggest that these factors contribute to the persistence and intensity of symptoms [7,20-22]. Therefore, targeting the emotional and cognitive factors of pain with PiP and rTMS interventions might be a solution to further improve the quality of life of these people.

Evidence for the effectiveness of rTMS over PFC in chronic musculoskeletal pain is sparse [57], with previous studies mostly targeting M1 and yielding conflicting results [34,35]. Targeting the dorsolateral PFC could be more effective than M1, as the left dorsolateral PFC is involved in cognition and emotion processing that may negatively influence recovery in patients with chronic pain [31,39]. This is also supported by previous studies yielding promising results on the effectiveness of rTMS over the left dorsolateral PFC for neuropathic pain [41].

Various types of psychotherapy (eg, cognitive behavioral therapy, acceptance and commitment therapy, mindfulness-based approaches), when combined with physiotherapy, have been associated with greater improvements in pain intensity and physical functioning among individuals with CLBP, compared to usual care [27]. Moreover, psychologically informed interventions have shown particularly promising results in the treatment of CLBP among patients presenting comorbid psychological factors [28]. Thus, combining these two approaches could enhance treatment response in veterans with CLBP, a population with a high prevalence of psychological disorders. One possible explanation would be that they both target cognitive and emotional mechanisms involved in pain processing but through different pathways [29]. The hypothesis for an enhanced response when combining the two approaches is supported by previous research conducted in patients with major depression. For example, the combination of cognitive behavioral therapy and rTMS treatment yields greater improvements in symptoms and remission rates than previous RCTs evaluating these treatment modalities alone [29].

This study will address some of the top priorities in the management of LBP, that is, to prioritize noninvasive and nonpharmacological biopsychosocial intervention, as recommended by both civilian [58,59] and veteran [60] guidelines. More specifically, it will provide new evidence on the effectiveness of two innovative interventions little studied to date in veterans with CLBP (ie, PiP and rTMS). Furthermore,

XSL•FO
RenderX

while the body of evidence supporting the efficacy of rTMS in treating mental health disorders such as major depression or posttraumatic stress disorders is steadily growing [61], important gaps remain in our understanding of its application among patients experiencing chronic musculoskeletal pain, particularly those presenting psychological risk factors. Should our hypotheses be confirmed and PiP demonstrates superior outcomes compared to UP, the findings of this study could yield substantive clinical implications. For example, such results could motivate clinicians to acquire training in administering PiP, as this training can be easily adapted across various health care professions.

Some limitations should, however, be mentioned. First, the optimal parameters for rTMS treatment over the PFC are still unknown. The parameters for this study are based on available evidence, which has mainly been conducted in patients with neuropathic pain [62-64]. It is possible that they are not optimal for the treatment of CLBP. Second, based on evidence from the population with depression, the use of psychotropic medication may attenuate the treatment response to rTMS [65]. We anticipate a high prevalence of psychotropic medication use among our participants, given the elevated rates of psychological disorders observed in veteran populations [7,20]. Finally, several clinicians will be providing care for whom PiP is a new concept. As such, it is possible that the standardization of the PiP intervention may be suboptimal despite the two days of training provided. To optimize standardization and application of the concepts, clinicians will be able to contact the expert clinician who provided the training throughout the study to validate information about the intervention provided. In addition, clinicians' logbooks will allow us to see which interventions were used with participants.

Funding

This work is funded by the Chronic Pain Centre of Excellence for Canadian Veterans and the Foundation for Physical Medicine and Rehabilitation. The funder has no role in the design of this study or the interpretation of the results.

Authors' Contributions

HMA, YTL, JSR, AMP, PL and LL contributed to the elaboration of the protocol. FD and ID developed the RedCap platform for the project, and FD wrote the manuscript. FD, IN, OB and LL are responsible for participants recruitment and follow-ups. AG developed the psychologically informed physiotherapy training adapted for veterans. PMD and LL developed a complementary training on military culture and the impacts of service. LL is a patient partner and helped develop the protocol, and the training. All the authors revised the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from the Chronic Pain Centre of Excellence for Canadian Veterans (CPCoE).
[PDF File (Adobe PDF File), 112 KB - resprot_v15i1e78952_app1.pdf]

References

1. The Lancet Rheumatology. The global epidemic of low back pain. Lancet Rheumatol 2023 Jun;5(6):e305. [doi: 10.1016/S2665-9913(23)00133-9] [Medline: 38251593]
2. GBD 2016 Disease Injury Incidence Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
3. Saragiotto B, G. Maher C, P Yamato T, OP Costa L, Menezes Costa LC, Ostalo RWJG, et al. Motor control exercise for chronic non-specific low-back pain. Cochrane Database Syst Rev 2016 Jan 16;CD012004:1-2. [doi: 10.1002/14651858.cd012004] [Medline: PMC8761501]
4. Bushey MA, Slaven JE, Outcalt SD, Kroenke K, Kempf C, Froman A, et al. Effect of medication optimization vs cognitive behavioral therapy among US veterans with chronic low back pain receiving long-term opioid therapy: a randomized clinical trial. JAMA Netw Open 2022 Nov 01;5(11):e2242533 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.42533] [Medline: 36394874]
5. Haun JN, Fowler CA, Venkatachalam HH, Saenger MS, Alman AC, Smith BM, et al. Empower Veterans Program (EVP): a chronic pain management program demonstrates positive outcomes among veterans. BMC Health Serv Res 2023 May 03;23(1):431 [FREE Full text] [doi: 10.1186/s12913-023-09327-5] [Medline: 37138319]
6. Goulet JL, Kerns RD, Bair M, Becker WC, Brennan P, Burgess DJ, et al. The musculoskeletal diagnosis cohort: examining pain and pain care among veterans. Pain 2016 Aug;157(8):1696-1703 [FREE Full text] [doi: 10.1097/j.pain.0000000000000567] [Medline: 27023420]

XSL•FO
RenderX

7. Talukdar JR, Zeraatkar D, Thomas A, Busse JW. An exploration of the increasing prevalence of chronic pain among Canadian veterans: Life After Service Studies 2016 and 2019. Can J Pain 2025;9(1):2443631 [FREE Full text] [doi: 10.1080/24740527.2024.2443631] [Medline: 39896960]
8. Nahin RL. Severe pain in veterans: the effect of age and sex, and comparisons with the general population. J Pain 2017 Mar;18(3):247-254 [FREE Full text] [doi: 10.1016/j.jpain.2016.10.021] [Medline: 27884688]
9. Schmidt C, Borgia M, Zhang T, Gochyyev P, Shireman TI, Resnik L. Initial treatment approaches and healthcare utilization among veterans with low back pain: a propensity score analysis. BMC Health Serv Res 2023 Mar 21;23(1):275 [FREE Full text] [doi: 10.1186/s12913-023-09207-y] [Medline: 36944926]
10. Jevne J. Stabbed in the back: catalysts for a paradigm shift in back pain care. Br J Sports Med 2016 Feb;50(4):198-199. [doi: 10.1136/bjsports-2015-094713] [Medline: 26130698]
11. Kregel J, Meeus M, Malfliet A, Dolphens M, Danneels L, Nijs J, et al. Structural and functional brain abnormalities in chronic low back pain: A systematic review. Semin Arthritis Rheum 2015 Oct;45(2):229-237. [doi: 10.1016/j.semarthrit.2015.05.002] [Medline: 26092329]
12. Nicholas M, Vlaeyen JWS, Rief W, Barke A, Aziz Q, Benoliel R, et al. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019 Jan;160(1):28-37. [doi: 10.1097/j.pain.0000000000001390] [Medline: 30586068]
13. Apkarian A, Hashmi JA, Baliki MN. Pain and the brain: specificity and plasticity of the brain in clinical chronic pain. Pain 2011 Mar;152(3 Suppl):S49-S64 [FREE Full text] [doi: 10.1016/j.pain.2010.11.010] [Medline: 21146929]
14. Vachon-Presseau E, Centeno M, Ren W, Berger S, Tétreault P, Ghantous M, et al. The emotional brain as a predictor and amplifier of chronic pain. J Dent Res 2016 Jun;95(6):605-612 [FREE Full text] [doi: 10.1177/0022034516638027] [Medline: 26965423]
15. Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. Lancet 2018 Jun 09;391(10137):2356-2367 [FREE Full text] [doi: 10.1016/S0140-6736(18)30480-X] [Medline: 29573870]
16. da C Menezes Costa L, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LOP. The prognosis of acute and persistent low-back pain: a meta-analysis. CMAJ 2012 Aug 07;184(11):E613-E624 [FREE Full text] [doi: 10.1503/cmaj.111271] [Medline: 22586331]
17. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? JAMA 2010 Apr 07;303(13):1295-1302. [doi: 10.1001/jama.2010.344] [Medline: 20371789]
18. Hashmi J, Baliki MN, Huang L, Baria AT, Torbey S, Hermann KM, et al. Shape shifting pain: chronification of back pain shifts brain representation from nociceptive to emotional circuits. Brain 2013 Sep;136(Pt 9):2751-2768 [FREE Full text] [doi: 10.1093/brain/awt211] [Medline: 23983029]
19. Shraim MA, Massé-Alarie H, Hodges PW. Methods to discriminate between mechanism-based categories of pain experienced in the musculoskeletal system: a systematic review. Pain 2021 Apr 01;162(4):1007-1037. [doi: 10.1097/j.pain.0000000000002113] [Medline: 33136983]
20. Hadlandsmyth K, Zhuang C, Driscoll MA, Lund BC. Comorbid chronic pain and post-traumatic stress disorder: current rates and psychiatric comorbidities among U.S. military veterans. Mil Med 2024 Nov 05;189(11-12):2303-2306. [doi: 10.1093/milmed/usae313] [Medline: 38869274]
21. Griffin S, Young JR, Naylor JC, Allen KD, Beckham JC, Calhoun PS. Reciprocal effects between depressive symptoms and pain in veterans over 50 years of age or older. Pain Med 2022 Feb 01;23(2):295-304 [FREE Full text] [doi: 10.1093/pm/pnab294] [Medline: 34643735]
22. Kneeman J, Battalio SL, Korpak A, Cherkin DC, Luo G, Rundell SD, et al. Predicting persistent disabling low back pain in veterans affairs primary care using the STarT Back tool. PM R 2021 Mar;13(3):241-249. [doi: 10.1002/pmrj.12488] [Medline: 32902134]
23. O'Sullivan P, P Caneiro J, O'Keeffe M, Smith A, Dankaerts W, Fersum K, et al. Cognitive functional therapy: an integrated behavioral approach for the targeted management of disabling low back pain. Phys Ther 2018 May 01;98(5):408-423. [doi: 10.1093/ptj/pzy087] [Medline: 29669082]
24. Valjakka A, Salanterä S, Laitila A, Julkunen J, Hagelberg N. The association between physicians' attitudes to psychosocial aspects of low back pain and reported clinical behaviour: a complex issue. Scand J Pain 2013 Jan 01;4(1):25-30 [FREE Full text] [doi: 10.1016/j.sjpain.2012.08.003] [Medline: 29913889]
25. Vanneman M, Jo Larson M, Chen C, Sayko Adams R, V Williams T, Meerwijk E, et al. Treatment of low back pain with opioids and nonpharmacologic treatment modalities for army veterans. Med Care 2018;56(10):855-861. [doi: 10.1097/mlr.0000000000000977] [Medline: 30134347]
26. Ballengee LA, Zullig LL, George SZ. Implementation of psychologically informed physical therapy for low back pain: where do we stand, where do we go? JPR 2021 Dec;14:3747-3757. [doi: 10.2147/jpr.s311973]
27. Ho EK, Chen L, Simic M, Ashton-James CE, Comachio J, Wang DXM, et al. Psychological interventions for chronic, non-specific low back pain: systematic review with network meta-analysis. BMJ 2022 Mar 30;376:e067718 [FREE Full text] [doi: 10.1136/bmj-2021-067718] [Medline: 35354560]
28. Rabiei P, Keough C, Patricio P, Côté-Picard C, Desgagnés A, Massé-Alarie H. Are tailored interventions to modifiable psychosocial risk factors effective in reducing pain intensity and disability in low back pain? a systematic review with

XSL•FO
RenderX

meta-analysis of randomized trials. J Orthop Sports Phys Ther 2025 Feb;55(2):89-108. [doi: 10.2519/jospt.2025.12777] [Medline: 39873474]
29. Donse L, Padberg F, Sack AT, Rush AJ, Arns M. Simultaneous rTMS and psychotherapy in major depressive disorder: clinical outcomes and predictors from a large naturalistic study. Brain Stimul 2018;11(2):337-345 [FREE Full text] [doi: 10.1016/j.brs.2017.11.004] [Medline: 29174304]
30. Loo C, Martin D, Pigot M, Arul-Anandam P, Mitchell P, Sachdev P. Transcranial direct current stimulation priming of therapeutic repetitive transcranial magnetic stimulation: a pilot study. J ECT 2009;25(4):256-260. [doi: 10.1097/yct.0b013e3181a2f87e] [Medline: 19440158]
31. Patricio P, Massé-Alarie H. The use of non-invasive brain stimulation for the management of chronic musculoskeletal pain: fad or future? Brain Sci 2025 Jul 17;15(7):760-1169. [doi: 10.3390/brainsci15070760]
32. Lefaucheur JP, Aleman A, Baeken C, Benninger DH, Brunelin J, Di Lazzaro V, et al. Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS): an update (2014-2018). Clin Neurophysiol 2020 Feb;131(2):474-528 [FREE Full text] [doi: 10.1016/j.clinph.2019.11.002] [Medline: 31901449]
33. Patricio P, Roy JS, Rohel A, Gariépy C, Émond C, Hamel É, et al. The effect of noninvasive brain stimulation to reduce nonspecific low back pain: a systematic review and meta-analysis. Clin J Pain 2021 Jun 01;37(6):475-485. [doi: 10.1097/AJP.0000000000000934] [Medline: 33949359]
34. Johnson S, Summers J, Pridmore S. Changes to somatosensory detection and pain thresholds following high frequency repetitive TMS of the motor cortex in individuals suffering from chronic pain. Pain 2006 Jul;123(1-2):187-192. [doi: 10.1016/j.pain.2006.02.030] [Medline: 16616419]
35. Ambriz-Tututi M, Alvarado-Reynoso B, Drucker-Colín R. Analgesic effect of repetitive transcranial magnetic stimulation (rTMS) in patients with chronic low back pain. Bioelectromagnetics 2016 Dec;37(8):527-535. [doi: 10.1002/bem.22001] [Medline: 27548757]
36. Mariano TY, Burgess FW, Bowker M, Kirschner J, Van't Wout-Frank M, Jones RN, et al. Transcranial direct current stimulation for affective symptoms and functioning in chronic low back pain: a pilot double-blinded, randomized, placebo-controlled trial. Pain Med 2019 Jun 01;20(6):1166-1177 [FREE Full text] [doi: 10.1093/pm/pny188] [Medline: 30358864]
37. Chan AW, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ 2025 Apr 28;389:e081477. [doi: 10.1136/bmj-2024-081477] [Medline: 40294953]
38. Deyo RA, Dworkin SF, Amtmann D, Andersson G, Borenstein D, Carragee E, et al. Report of the NIH Task Force on research standards for chronic low back pain. Pain Med 2014 Aug;15(8):1249-1267. [doi: 10.1111/pme.12538] [Medline: 25132307]
39. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. Arthritis Rheum 2008 May 15;59(5):632-641. [doi: 10.1002/art.23563] [Medline: 18438893]
40. McClintock SM, Reti IM, Carpenter LL, McDonald WM, Dubin M, Taylor SF, et al. Consensus recommendations for the clinical application of repetitive transcranial magnetic stimulation (rTMS) in the treatment of depression. J Clin Psychiatry 2018;79(1):16cs10905 [FREE Full text] [doi: 10.4088/JCP.16cs10905] [Medline: 28541649]
41. Lefaucheur JP, Nguyen JP. A practical algorithm for using rTMS to treat patients with chronic pain. Neurophysiol Clin 2019 Sep;49(4):301-307. [doi: 10.1016/j.neucli.2019.07.014] [Medline: 31375381]
42. Fitzgibbon BM, Hoy KE, Knox LA, Guymer EK, Littlejohn G, Elliot D, et al. Evidence for the improvement of fatigue in fibromyalgia: a 4-week left dorsolateral prefrontal cortex repetitive transcranial magnetic stimulation randomized-controlled trial. Eur J Pain 2018 Aug;22(7):1255-1267. [doi: 10.1002/ejp.1213] [Medline: 29542208]
43. Blumberger DM, Vila-Rodriguez F, Thorpe KE, Feffer K, Noda Y, Giacobbe P, et al. Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): a randomised non-inferiority trial. Lancet 2018 Apr 28;391(10131):1683-1692. [doi: 10.1016/S0140-6736(18)30295-2] [Medline: 29726344]
44. Seminowicz DA, de Martino E, Schabrun SM, Graven-Nielsen T. Left dorsolateral prefrontal cortex repetitive transcranial magnetic stimulation reduces the development of long-term muscle pain. Pain 2018 Dec;159(12):2486-2492. [doi: 10.1097/j.pain.0000000000001350] [Medline: 30431520]
45. Desgagnés A, Côté-Picard C, Gaumond A, Langevin P, Piché M, Pagé G, et al. Efficacy of a psychologically-informed physiotherapy intervention in patients with chronic low back pain at high risk of poor prognosis: a pilot and feasibility randomized controlled trial. Physiother Can 2024 May;76(2):163-174. [doi: 10.3138/ptc-2023-0038] [Medline: 38725600]
46. Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LOP, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. Pain 2018 Mar;159(3):481-495. [doi: 10.1097/j.pain.0000000000001117] [Medline: 29194127]
47. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]

XSL•FO
RenderX

48. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
49. Lopes RA, Dias RC, Queiroz BZD, Rosa NMDB, Pereira LDSM, Dias JMD, et al. Psychometric properties of the Brazilian version of the Pain Catastrophizing Scale for acute low back pain. Arq Neuropsiquiatr 2015 May;73(5):436-444 [FREE Full text] [doi: 10.1590/0004-282X20150026] [Medline: 26017211]
50. Dupuis F, Cherif A, Batcho C, Massé-Alarie H, Roy JS. The Tampa Scale of Kinesiophobia: a systematic review of its psychometric properties in people with musculoskeletal pain. Clin J Pain 2023 May 01;39(5):236-247. [doi: 10.1097/AJP.0000000000001104] [Medline: 36917768]
51. Dubé MO, Langevin P, Roy JS. Measurement properties of the Pain Self-Efficacy Questionnaire in populations with musculoskeletal disorders: a systematic review. Pain Rep 2021;6(4):e972 [FREE Full text] [doi: 10.1097/PR9.0000000000000972] [Medline: 34963996]
52. De Clifford-Faugère G, Nguena Nguefack H, Godbout-Parent M, Diallo M, Guénette L, Gabrielle Pagé M, et al. The Medication Quantification Scale 4.0: an updated index based on prescribers' perceptions of the risk associated with chronic pain medications. J Pain 2024 Feb;25(2):508-521 [FREE Full text] [doi: 10.1016/j.jpain.2023.09.010] [Medline: 37838346]
53. Ashbaugh AR, Houle-Johnson S, Herbert C, El-Hage W, Brunet A. Psychometric validation of the english and french versions of the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5). PLoS One 2016;11(10):e0161645 [FREE Full text] [doi: 10.1371/journal.pone.0161645] [Medline: 27723815]
54. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. BMC Musculoskelet Disord 2006 Oct 25;7:82 [FREE Full text] [doi: 10.1186/1471-2474-7-82] [Medline: 17064410]
55. Patricio P, Mailloux C, Wideman TH, Langevin P, Descarreaux M, Beaulieu LD, et al. Assessment of exercise-induced hypoalgesia in chronic low back pain and potential associations with psychological factors and central sensitization symptoms: A case-control study. Pain Pract 2023 Mar;23(3):264-276. [doi: 10.1111/papr.13189] [Medline: 36461643]
56. Hudson TJ, Painter JT, Martin BC, Austen MA, Williams JS, Fortney JC, et al. Pharmacoepidemiologic analyses of opioid use among OEF/OIF/OND veterans. Pain 2017 Jun;158(6):1039-1045 [FREE Full text] [doi: 10.1097/j.pain.0000000000000874] [Medline: 28195856]
57. Patricio P, Roy JS, Macedo L, Roy M, Léonard G, Hodges P, et al. Repetitive transcranial magnetic stimulation alone and in combination with motor control exercise for the treatment of individuals with chronic non-specific low back pain (ExTraStim trial): study protocol for a randomised controlled trial. BMJ Open 2021 Mar 24;11(3):e045504 [FREE Full text] [doi: 10.1136/bmjopen-2020-045504] [Medline: 33762244]
58. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, et al. Prevention and treatment of low back pain: evidence, challenges, and promising directions. Lancet 2018 Jun 09;391(10137):2368-2383. [doi: 10.1016/S0140-6736(18)30489-6] [Medline: 29573872]
59. Qaseem A, Wilt TJ, McLean RM, Forciea MA, Clinical Guidelines Committee of the American College of Physicians, Denberg TD, et al. Noninvasive treatments for acute, subacute, and chronic low back pain: a clinical practice guideline from the american college of physicians. Ann Intern Med 2017 Apr 04;166(7):514-530 [FREE Full text] [doi: 10.7326/M16-2367] [Medline: 28192789]
60. Pangarkar SS, Kang DG, Sandbrink F, Bevevino A, Tillisch K, Konitzer L, et al. VA/DoD clinical practice guideline: diagnosis and treatment of low back pain. J Gen Intern Med 2019 Nov;34(11):2620-2629 [FREE Full text] [doi: 10.1007/s11606-019-05086-4] [Medline: 31529375]
61. Vida RG, Sághy E, Bella R, Kovács S, Erdősi D, Józwiak-Hagymásy J, et al. Efficacy of repetitive transcranial magnetic stimulation (rTMS) adjunctive therapy for major depressive disorder (MDD) after two antidepressant treatment failures: meta-analysis of randomized sham-controlled trials. BMC Psychiatry 2023 Jul 27;23(1):545 [FREE Full text] [doi: 10.1186/s12888-023-05033-y] [Medline: 37501135]
62. Zhou J, Wang Y, Luo X, Fitzgerald PB, Cash RFH, Fitzgibbon BM, et al. Revisiting the effects of rTMS over the dorsolateral prefrontal cortex on pain: an updated systematic review and meta-analysis. Brain Stimul 2024;17(4):928-937 [FREE Full text] [doi: 10.1016/j.brs.2024.07.011] [Medline: 39089648]
63. Mori N, Hosomi K, Nishi A, Matsugi A, Dong D, Oshino S, et al. Exploratory study of optimal parameters of repetitive transcranial magnetic stimulation for neuropathic pain in the lower extremities. Pain Rep 2021;6(4):e964 [FREE Full text] [doi: 10.1097/PR9.0000000000000964] [Medline: 34667918]
64. Naik A, Bah M, Govande M, Palsgaard P, Dharnipragada R, Shaffer A, et al. Optimal frequency in repetitive transcranial magnetic stimulation for the management of chronic pain: a network meta-analysis of randomized controlled trials. World Neurosurg 2024 Apr;184:e53-e64. [doi: 10.1016/j.wneu.2024.01.010] [Medline: 38185460]
65. Intrator J, Noto J, Vallesteros RD, Peltier M, O'Reardon J, Abbas M. Effects of concomitant use of various psychotropic medications on the treatment response to transcranial magnetic stimulation for depression: a literature review. Cureus 2024 Nov;16(11):e72993. [doi: 10.7759/cureus.72993] [Medline: 39634972]

XSL•FO
RenderX

Abbreviations

CLBP: chronic low back pain
LBP: low back pain
ODI: Oswestry Disability Index
PFC: prefrontal cortex
PiP: psychologically informed physiotherapy
PROM: patient-reported outcome measure
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
rTMS: repetitive transcranial magnetic stimulation
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
UP: usual physiotherapy

Edited by A Schwartz; submitted 13.Jun.2025; this is a non–peer-reviewed article;accepted 07.Nov.2025; published 26.Jan.2026.

Please cite as:
Dupuis F, Tousignant-Laflamme Y, Marier Deschênes P, Fournier P, Nduwimana I, Ballot O, Loisel L, Pinard AM, Lacombe L, Langevin P, Gaumond A, Roy JS, Massé-Alarie H
Combining Noninvasive Brain Stimulation and Physiotherapy to Improve the Management of Chronic Low Back Pain in Veterans: Protocol for a Multi-Arm Randomized Controlled Trial
JMIR Res Protoc 2026;15:e78952
URL: https://www.researchprotocols.org/2026/1/e78952
doi:10.2196/78952
PMID:

©Frederique Dupuis, Yannick Tousignant-Laflamme, Pascale Marier Deschênes, Philippe Fournier, Ildephonse Nduwimana, Orlane Ballot, Loris Loisel, Anne Marie Pinard, Luc Lacombe, Pierre Langevin, Alain Gaumond, Jean-Sébastien Roy, Hugo Massé-Alarie. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effectiveness of an Educational and Counseling Program (the Green Mother Project Phase 2) to Enhance Breastfeeding and Improve Mothers' Diets From an Environmental Perspective: Protocol for a Cluster Randomized Controlled Trial

Rosa Maria Cabedo-Ferreiro[1,2]; Liudmila Liutsko[2,3,4], MSci, PhD; Judit Cos-Busquets[5]; Rosa García-Sierra[3,6], PhD; Margalida Colldeforns-Vidal[1]; Azahara Reyes-Lacalle[2,5]; Pere Torán-Monserrat[3,6,7,8], PhD; M Mercedes Vicente-Hernández[2,9]; Miriam Gómez-Masvidal[2,10]; Concepció Violán[3,7,11,12,13], PhD; Laura Montero-Pons[2,10], PhD; Gemma Falguera-Puig[2,5], PhD; Gemma Cazorla-Ortiz[2,5], PhD

[1]Atenció a la Salut Sexual I Reproductiva (ASSIR) Granollers, Institut Català de la Salut, Granollers, Spain

[2](2021-SGR-793), Consolidated Research Group on Sexual and Reproductive Healthcare (GRASSIR), Barcelona, Spain

[3]Unitat de Suport a la Recerca Metropolitana Nord, Institut Universitari d'Investigació en Atenció Primària Jordi Gol, Mataró, Spain

[4]Dep. Health Sciences; Group PSICONLINE, Universidad Internacional De La Rioja, Logroño, Spain

[5]Atenció a la Salut Sexual I Reproductiva (ASSIR) Sabadell, Institut Català de la Salut, Sabadell, Spain

[6](2021-SGR-0148), Multidisciplinary Research Group in Health and Society (GREMSAS), Barcelona, Spain

[7]Institut d'Investigació Germans Trias i Pujol (IGTP), Badalona, Spain

[8]Department of Medicine, Faculty of Medicine, Universitat de Girona, Girona, Spain

[9]Atenció a la Salut Sexual I Reproductiva (ASSIR) Badalona - Sant Adrià, Institut Català de la Salut, Sant Adrià del Besós, Spain

[10]Atenció a la Salut Sexual I Reproductiva (ASSIR) Mataró, Institut Català de la Salut, Mataró, Spain

[11]Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain

[12]Institut Universitari D'Investigació en Atenció Primària Jordi Gol (IDIAPJGol), (2021 SGR 01537), Grup de Recerca en Impacte de les Malalties Cròniques i les seves Trajectòries (GRIMTra), Barcelona, Spain

[13]Institut Universitari D'Investigació en Atenció Primària Jordi Gol (IDIAPJGol), (RD21/0016/0029), El grupo de investigación en servicios sanitarios en Atención Primaria (GrenSSAP), RICAPPS, Barcelona, Spain

Corresponding Author:
Liudmila Liutsko, MSci, PhD
Unitat de Suport a la Recerca Metropolitana Nord
Institut Universitari d'Investigació en Atenció Primària Jordi Gol
Carrer Mare de Déu de Guadalupe, 2
Mataró, 08303
Spain
Phone: 34 93 741 53 38
Email: lliutskop.mn.ics@gencat.cat

Abstract

Background: Exclusive breastfeeding is recommended as healthier and more sustainable than formula feeding. It produces less waste, requires fewer resources, and has a smaller environmental impact. Breastfeeding has some environmental impact related to increased maternal dietary needs and the use of feeding accessories. In light of the global climate emergency and suboptimal breastfeeding rates, targeted interventions are urgently needed to promote sustainable infant feeding practices. There are few studies that evaluate sustainability interventions in the postpartum period.

Objective: The objective of this study is to evaluate the effectiveness of an educational and counseling intervention on breastfeeding and healthy maternal nutrition from an environmental perspective.

Methods: A multicenter prospective intervention study is being conducted in 2 cohorts in primary care centers and hospitals in the north metropolitan area of Barcelona. The control group received standard obstetric care. The experimental group additionally received an educational intervention and health care support on breastfeeding and healthy and sustainable maternal nutrition. Pregnant women were monitored from 24 weeks of gestation to 6 months post partum. The rates of different types of breastfeeding,

XSL•FO
RenderX

the women's diet, and the associated environmental impacts (climate change and water footprint) will be analyzed to assess the effectiveness of the intervention.

Results: The development of the educational and counseling intervention has been completed, including the creation of the *Guide to Good Practices in Breastfeeding, Nutrition, and Sustainability*. Health care professionals received targeted training. Recruitment of pregnant women was conducted from December 2023 to December 2024. Prenatal education sessions and specialized care pathways were designed and implemented. Breastfeeding-friendly spaces were adapted to support the participating centers. Data collection for monitoring breastfeeding practices, maternal diet, and environmental impact indicators (carbon footprint and water footprint), with the follow-up period of 6 months post partum, was extended until September 2025, with a complementary missing data collection in October 2025. Data cleaning for final analysis is expected to conclude by January 2026. This study hypothesizes that mothers who receive higher levels of education and counseling support will (1) breastfeed for a longer duration, (2) adopt healthier and more sustainable dietary practices, and (3) reduce environmental impacts associated with both infant feeding accessories and dietary choices.

Conclusions: We expect an increase in the incidence and prevalence rates of breastfeeding and a shift toward a healthy and sustainable diet with low environmental impact.

Trial Registration: ClinicalTrials.gov NCT05729581; https://clinicaltrials.gov/study/NCT05729581

International Registered Report Identifier (IRRID): DERR1-10.2196/80358

(JMIR Res Protoc 2026;15:e80358) doi:10.2196/80358

KEYWORDS

public health; planetary health; nutrition; post partum; breastfeeding; bottle feeding; counseling; prenatal education; health education; public health education for professionals

Introduction

The World Health Organization (WHO) recommends exclusive breastfeeding (EBF) for the first 6 months of life, continuing until the child is aged 2 years or as desired by mother and child, as it provides numerous physical and mental health benefits [1-5]. Factors that impact breastfeeding include the type of delivery; the socioeconomic status of the mother; her return to work; support from partners, families, and peer groups; and the level of prenatal breastfeeding education [6]. Breastfeeding is a sustainable and environmentally friendly infant feeding method compared to formula feeding, as it generates less waste and produces minimal greenhouse gas (GHG) emissions; therefore, it has a smaller carbon footprint. However, we cannot affirm that breastfeeding has a 0 cost from an environmental point of view because the impact of the increase in maternal diet and the use of breastfeeding accessories must be calculated [1,7-9]. Maternal diet [10,11] and associated consumption habits (local purchasing; type of energy and cooking used) can also contribute not only to the quality of nutrition [12] but also to sustainability and environmental protection [13,14]. The Paris Agreement has set ambitious targets for Europe to reduce its carbon footprint, aiming for a 55% reduction in GHG emissions by 2030 and climate neutrality by 2050 [15]. Food production accounts for 30% of the total GHG emissions. A healthy diet based on the consumption of vegetables, legumes, and low-impact foods contributes to reducing the carbon footprint [8]. By contrast, the manufacturing and distribution of industrial infant formulas harm the environment, creating polluting waste and requiring substantial use of energy and water [15].

Supporting breastfeeding mothers is crucial to achieving the WHO 2025 target (reaching 50% EBF during the first 6 months of life) [16] and the goals of the United Nations Paris Agreement [15]. However, despite many mothers' desire to breastfeed, obstacles and difficulties often lead to premature cessation of breastfeeding [1,17].

Health professionals play a vital role in providing support and guidance for successful and prolonged breastfeeding. Specific training for professionals followed by advice and education for users has been shown to be effective in improving breastfeeding care [18]. Shafaei et al [19], in 2020, demonstrated that after a bad experience with breastfeeding, prenatal education or counseling solves the main breastfeeding problems and improves breastfeeding self-efficacy. The implementation of supportive interventions in both the prenatal and postpartum periods can improve the rate, exclusivity, and duration of breastfeeding [19-21]. Souza et al [20] confirmed the effectiveness of an intervention also in the postpartum period. Educational programs have demonstrated greater effectiveness than support programs conducted through telephone and promotional materials, such as leaflets [22]. Sattari et al [22], in a systematic review of 30 trials, could not determine whether the combination of education with support was more effective than education alone. If we consider the information on a healthy diet, a Cochrane review with 1090 women evaluated the effectiveness in terms of perinatal outcomes of prenatal dietary education [23]. Several studies have demonstrated the effectiveness of educational interventions and professional support in improving both the rates and duration of breastfeeding [18-20]. However, existing research has predominantly focused on clinical and psychosocial aspects, with limited attention given to the interrelationship among health education, sustainable nutrition, and environmental impact.

The education of health care professionals is a key factor. Previous studies have identified a lack of environmental knowledge among health professionals, which limits their ability to guide families toward more sustainable practices [24]. Jadotte et al [24] emphasized the need to incorporate environmental

XSL•FO
RenderX

content into health care education so that professionals are equipped to transfer this knowledge to the wider population.

Research into educational interventions on the environmental benefits of breastfeeding and mothers' diets is essential to achieve changes in the attitude toward prolonged breastfeeding and a healthy and sustainable diet for the whole family. Currently, there are no studies on the impact of a training intervention for professionals and education for users on breastfeeding and a healthy and sustainable diet. Therefore, conducting research in this area, increasing the knowledge of professionals, and updating standard prenatal courses can contribute to raising awareness and promoting a change in attitude toward the protection of the environment and the promotion of public health as well as ensuring the future well-being of both the population and the planet [25,26].

The main objective of our study is to evaluate the effectiveness of an educational intervention with counseling and material resources for breastfeeding in improving breastfeeding rates, healthy diet, and environmental impacts. At the same time, it is hoped that the rates of different types of breastfeeding up to 6 months of life of the newborn can be calculated; the quality of the maternal diet in terms of calories and macro- and micronutrients can be described; and, finally, the environmental impact of infant and maternal nutrition between the 2 groups can be compared. To sum up, this study aims to evaluate whether a nutritional and environmental educational intervention improves EBF rates and maternal diet quality and reduces environmental impact compared to standard obstetric care.

We hypothesize that women who receive a nutritional and environmental educational intervention will have higher rates of EBF and a more balanced diet than women who receive standard obstetric care. Furthermore, the environmental impact produced by the participants who receive intervention will be smaller than that produced by the women who receive the standard obstetric care.

Methods

This research protocol is part of a broader project, where in phase 1, the environmental impact of infant feeding is calculated, considering the accessories necessary for it and the maternal diet, accounting for food consumption habits [9,12,27]. With these results, it has been possible to design the informative materials and courses that were used in the intervention, that is, phase 2 of the Green Mother project.

Design of This Study

A multicenter intervention cluster study with a randomized controlled trial is being conducted in the north metropolitan area of Barcelona, Spain, between 2023 and 2026. It is implemented in the public universal primary care system, specifically from the reproductive and sexual health care units (ASSIRs; Catalan: Atenció a la Salut Sexual i Reproductiva) [28,29] and in the territory's reference public hospitals for childbirth care. The participating centers were randomly assigned (primary care and their reference hospitals).

Participants and Inclusion and Exclusion Criteria

The study population is composed of pregnant women from the ASSIR centers, where pregnancy monitoring is conducted, with deliveries planned in the reference hospitals. Participants were enrolled in order of care, provided they met the inclusion criteria and had been accepted into the study, until the predetermined sample size was reached. The inclusion criteria were as follows: (1) receiving care in these ASSIRs and hospitals; (2) being aged 16 years or more, with a gestational age between 24 and 32 weeks; and (3) wishing to breastfeed. The exclusion criteria were as follows: (1) linguistic barriers that prevented the understanding of the information and the ability to complete the questionnaires and (2) not wishing to participate. Participants were followed from 24 to 32 weeks of gestation to 6 months post partum. The follow-up period lasted until September 2025.

Ethical Considerations

We followed the guidelines of the Declaration of Helsinki regarding bioethical principles of clinical research. This research protocol was approved by the research ethics committee of the Jordi Gol Primary Care Research Institute as well as the reference hospitals participating in this study: Hospital Parc Taulí, Hospital General de Granollers, Hospital Germans Trias i Pujol, and Hospital de Mataró. The final version of this protocol (22/101-P) was approved and signed on February 22, 2023.

Data collected through questionnaires on REDCap (Research Electronic Data Capture; Vanderbilt University) are encrypted and stored in a secure database. Data for the study were collected directly from project participants based on their prior written consent, in accordance with the provisions of Articles 6.1.a and 9.2.a of the General Data Protection Regulation.

Sample Size Calculation

A total of 126 participants is needed in each group to detect a statistically significant difference between the 2, assuming a confidence level of 95% and a power of 80% in a 2-tailed test, according to the calculations of the GRANMO sample size and power calculator (version 8.0; Datarus). Based on the data published in 2019 on the population studied, the EBF rate at 6 months would be 17% in the control group (CG) and 34% in the experimental group (EG) [21]. The loss to follow-up rate is estimated to be 20%. A total of 252 participants will be studied, with 126 (50%) in each group.

Procedure

Recruitment

Midwives at the participating centers invited pregnant women who met the inclusion criteria and provided consent to take part in this study by signing the informed consent form.

Intervention

Randomization was performed at the cluster level (health care centers), ensuring that all participants within a given center received the same type of care. The intervention study included two groups: (1) CG, which received standard obstetric care (ASSIRs of Badalona, Sant Adrià, and Mataró together with Hospital Germans Trias i Pujol and Hospital de Mataró) and

XSL•FO
RenderX

(2) EG or intervention group, which received an extra educational program and health care support (ASSIRs of Granollers, Sabadell, and Cerdanyola with Hospital Parc Taulí and Hospital General de Granollers).

A multimodal intervention included activities for professionals and the study participants, such as dissemination, training, and counseling support actions.

The dimensions of the intervention (Table 1 and Figure 1) are mentioned subsequently.

Table 1. The educational and support intervention program overview performed in the experimental group centers.

Intervention target and intervention type (dimensions)	Before recruitment (primary care)	Pregnancy		Delivery	Postpartum visits					
		24 gestational weeks	28-36 gestational weeks	Birth	Immediate	Early	Month 1	Month 3	Month 4	Month 6
		Primary care		Hospital		Primary care				
Health professionals										
Creation of the Guide to Good Practices in Breastfeeding, Nutrition, and Sustainability	✓									
Training of health workers (virtual course: 20 h)	✓									
Pregnant and postpartum persons										
Equipping breastfeeding spaces and providing other resources	✓									
Educational leaflets, posters with QR codes, and videos	✓									
The educational program: special sessions on sustainable feeding and nutrition for pregnant persons: (1) breastfeeding, (2) sustainable feeding, and (3) healthy nutrition (1 hour per session[a])			✓							
Individual special care and counseling for postpartum persons in hospitals and primary care centers following the guide		✓		✓	✓	✓	✓	✓		✓

[a]Versus standard educational courses and standard obstetric care in the control group.

XSL•FO
RenderX

Figure 1. Workflow chart.

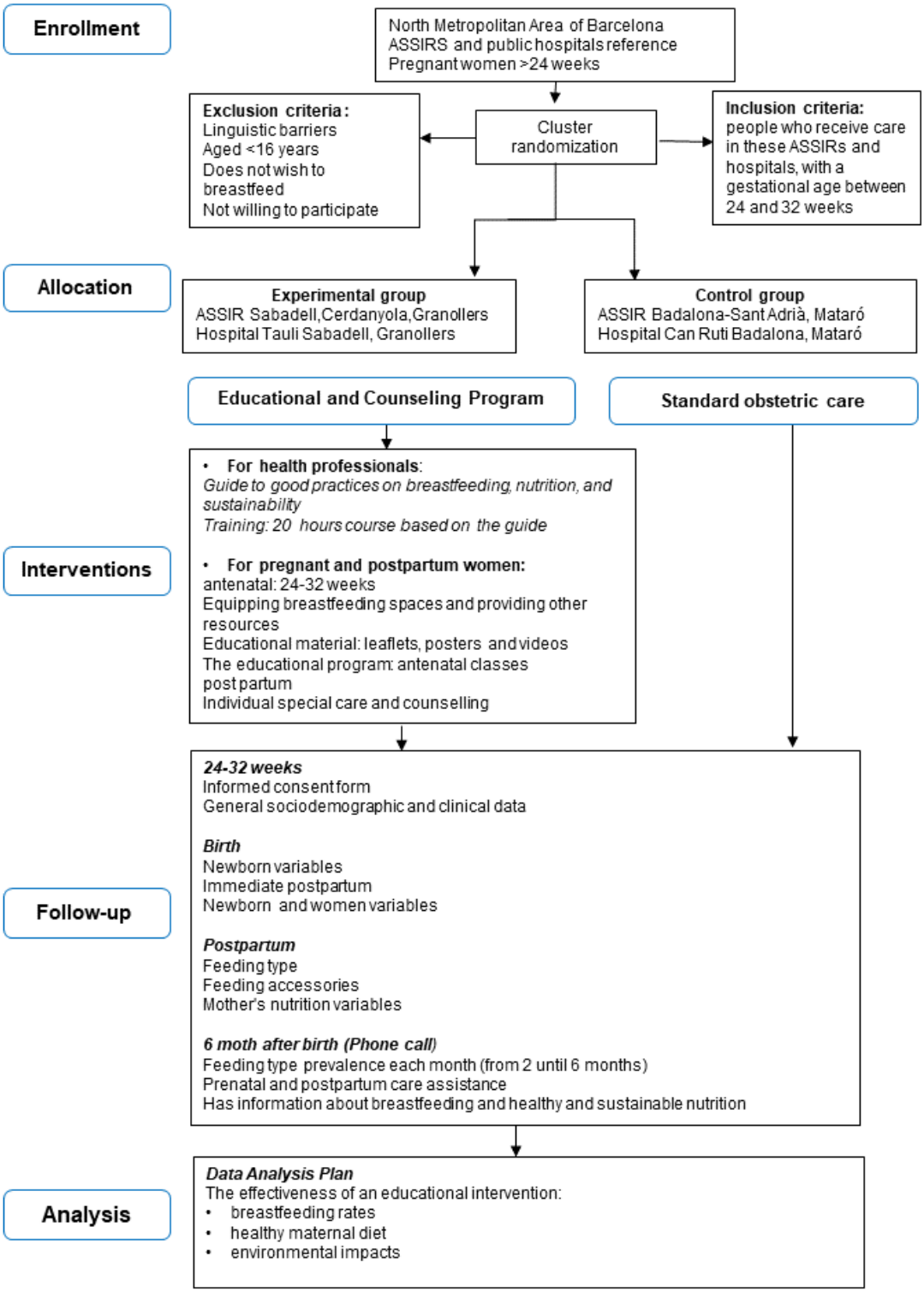

For primary care and hospital health professionals (with the aim of enhancing and homogenizing knowledge), the first dimension was the creation of a *Guide to Good Practices in Breastfeeding, Nutrition, and Sustainability* (based on breastfeeding topics in the *Clinical Practice Guide on Breastfeeding* published by the Spanish Ministry of Health). The content was structured around the following topics: anatomy, physiology, benefits, initiation, monitoring, and challenges of breastfeeding; healthy maternal nutrition; sustainability; and sustainable diets. It was used in both primary care centers and hospitals to achieve excellence in care for mothers and babies and included updated information on breastfeeding and healthy and sustainable nutrition.

The second dimension was training provided to health professionals, such as midwives, pediatric nurses, and pediatricians, in ASSIR centers, primary care centers, and hospitals. This was conducted through a 20-hour course taught by the Green Mother research team using a virtual training platform. The course consisted of 12 topics based on the content

XSL•FO
RenderX

of the guide named in the first dimension and the results of phase 1 of the Green Mother study on breastfeeding, maternal diet, and environmental impact. Viewing of the videos was monitored, and participants completed an assessment consisting of a series of questions on the content to evaluate their knowledge. These courses were scheduled between 2023 and 2024, before the start of participant recruitment for the study.

For pregnant and postpartum women, the third dimension was equipment for waiting rooms, other spaces, and resources. Financial resources were allocated to support the adaptation of primary care centers and hospitals to provide optimal breastfeeding counseling, for example, creating cozy and calm spaces with sofas and supports (ie, nursing pillows, breast pumps, nipple shields, and reusable nursing pads).

The fourth dimension included educational materials. Information brochures were created for participants with advice on effective breastfeeding; tips or information messages for a healthy and sustainable diet; as well as details on the environmental impacts of formula milk, accessories, and food.

Informative videos on breastfeeding and healthy and sustainable nutrition, developed in accordance with the guidelines, were displayed in the waiting areas of primary care centers and hospitals.

Posters for surgeries and waiting rooms with QR codes were provided, where all the project information could be accessed.

The fifth dimension, the prenatal education course taught by trained professionals, was expanded and adapted with specific information on the Green Mother project. It included 3 sessions, namely breastfeeding, healthy maternal diet, and sustainability. Presentations in Microsoft PowerPoint format and recorded classes on healthy eating, breastfeeding, and environmental education were designed for users who cannot attend classes. Each class lasted 1 hour.

The sixth dimension was specific care for participants from recruitment to the study until post partum. Counseling was carried out at an individual level, both in primary care centers and at hospital admission, following the *Guide to Good Practices in Breastfeeding, Nutrition, and Sustainability*(first dimension) and offering support for behavioral changes that facilitate EBF and a healthier and more sustainable diet using the resources provided.

Data Collection and Data Sources

All data were collected through questionnaires completed by the researcher in charge of participant recruitment, using the REDCap secure web application [30] housed in the Catalan Health Institute.

A range of variables were collected to enable the computation of outcome measures used to evaluate the effectiveness of the intervention (refer to the Data Analysis Plan section).

The timing of participants' data collection is presented in Table 2, and the variables (more details are provided in Multimedia Appendix 1) are presented in Textbox 1.

Data collection by time, place, and the instruments used are presented in Table 2.

XSL•FO
RenderX

Table 2. Data collection protocol for the Green Mother phase 2 study.

Validated instrument	Primary care	Hospital		Primary care		Phone call		
	24 to 32 weeks of pregnancy	Birth	Immediate puerperium	Post partum	4 to 6 weeks	Month 3	Month 4	Month 6
Ad hoc questionnaires (Research Electronic Data Capture)								
Informed consent to be signed	✓							
General sociodemographic and clinical data	✓							
Newborn variables		✓	✓	✓	✓			
Breastfeeding observation: Latch, Audible Swallowing, Type of Nipple, Comfort, and Hold scale			✓	✓				
Obstetric pain: visual analog scale			✓	✓				
Tongue tie			✓	✓				
Nipple wounds			✓	✓				
Other puerperium variables		✓	✓	✓				
Feeding type			✓			✓	✓	✓
Mother's nutrition variables								
Purchasing and cooking habits					✓			
The mother's daily diet recall (24 h)					✓			
Feeding accessories				✓				

Textbox 1. Data collection variables.

Prenatal primary care data

- Study recruitment survey: confirmation of informed consent and inclusion criteria
- Sociodemographic and clinical data

Hospital data at the time of delivery

- Delivery room survey: birth and breastfeeding initiation data
- Maternity department survey: data from the visual analog scale (VAS) [31]; assessment of breastfeeding, such as the Latch, Audible Swallowing, Type of Nipple, Comfort, and Hold (LATCH) scale [32]; refer to Multimedia Appendix 1 for ankyloglossia and other variables

Postpartum data (primary care data)

- Early postpartum survey (primary care; 1-2 weeks after birth): data from the VAS [31] and breastfeeding assessment, such as the LATCH scale [32], ankyloglossia, etc (Multimedia Appendix 1)
- Quarantine survey (4-6 weeks after birth): postpartum survey that includes data from the VAS [25] and breastfeeding assessment, such as the LATCH scale [31] and frenulum
- Survey on infant feeding and accessories
- Survey on the nutrition of the mother
- Ferrari 24-hour recall of maternal diet [33,34]
- Eating and cooking habits
- Phone call survey after 6 months: survey on types of infant feeding from 2 to 6 months; survey to evaluate the use of educational and support interventions
- Survey on withdrawal from the study

XSL•FO
RenderX

Data Analysis Plan

The effectiveness of the educational intervention will be evaluated as differences between the EG and CG across 3 principal outcome variables. These outcomes are derived from the data collected during the study and are structured from general descriptive indicators to specific interventional impacts.

Infant Feeding Practices

Breastfeeding practices will be assessed from birth to 6 months of age. The incidence and prevalence of EBF, mixed feeding, and formula feeding will be calculated.

Maternal Nutrition

Maternal dietary intake will be evaluated in comparison with the recommended nutritional standards [35], focusing on caloric intake (total kilocalories consumed), macronutrient distribution (proportions of carbohydrates, proteins, and fats [saturated, monounsaturated, and polyunsaturated]), and micronutrient levels (iron, calcium, magnesium, and vitamin D).

The analysis will determine whether the average maternal diet meets the recommended nutritional guidelines.

Environmental Impact Assessment

Environmental impacts associated with infant feeding practices and maternal diet will be assessed at 6 weeks post partum. This includes consumption of artificial milk, use of feeding accessories, and maternal dietary habits.

A comparative life cycle assessment will be conducted in accordance with ISO 14040:2006, focusing on 3 environmental indicators:

- Climate change—measured in kg CO2 equivalents using the Intergovernmental Panel on Climate Change Global Warming Potential 100-year method [36]
- Water consumption—measured in m3 using the ReCiPE method, accounting for both direct and indirect water use [37]
- Water scarcity—measured in m3 equivalents using the available water remaining method [38]

Statistical Analysis

Categorical variables (such as gender, country of birth, and infant feeding type) will be summarized using absolute and relative frequencies for the total and categories.

Quantitative variables will be described using means and SDs for normally distributed data and medians and 95% CIs for nonnormally distributed data.

The differences will be checked with the chi-square test for categorical variables and ANOVA for continuous variables with normal distribution or the Kruskal-Wallis *H* test for continuous variables with nonnormal distribution. The statistical significance will be set at $P<.05$. The registers with missing data for main outcome variables will not be included in the analysis.

Results

Progress Timing and Flow of Participants Through the Study

The timeline and implementation status of phase 2 of the Green Mother project are mentioned subsequently.

Implementation Status

Following approval by the ethics committee on February 22, 2023, the research team organized a multidisciplinary group of experts to develop the *Guide to Good Practices in Breastfeeding, Nutrition, and Sustainability*. This guide was based on the Clinical Practice Guide on Breastfeeding published by the Spanish Ministry of Health and was developed during the first half of 2023.

Guide Development and Professional Training

A total of 16 health care professionals, each specialized in relevant content areas, contributed to the development of the guide. In addition, a 20-hour training course was organized to educate professionals on best practices in breastfeeding, nutrition, and environmental impact. This training was conducted between June 2023 and July 2023.

Recruitment Status

Recruitment of pregnant women began in September 2023 and finished in December 2024.

Data Collection and Follow-Up

Data collection was conducted through follow-up of participants until 6 months post partum for the last recruited woman. This phase extended until September 2025, with complementary data collection and revision in October 2025.

Data Exploitation and Analysis Status

Between November 2025 and January 2026, the research team will complete data cleaning and afterward conduct the final data analysis. The final results are anticipated to be ready for publication and dissemination in 2026.

The participant flowchart is presented in Figure 2.

XSL•FO
RenderX

Figure 2. Flow chart of study participation.

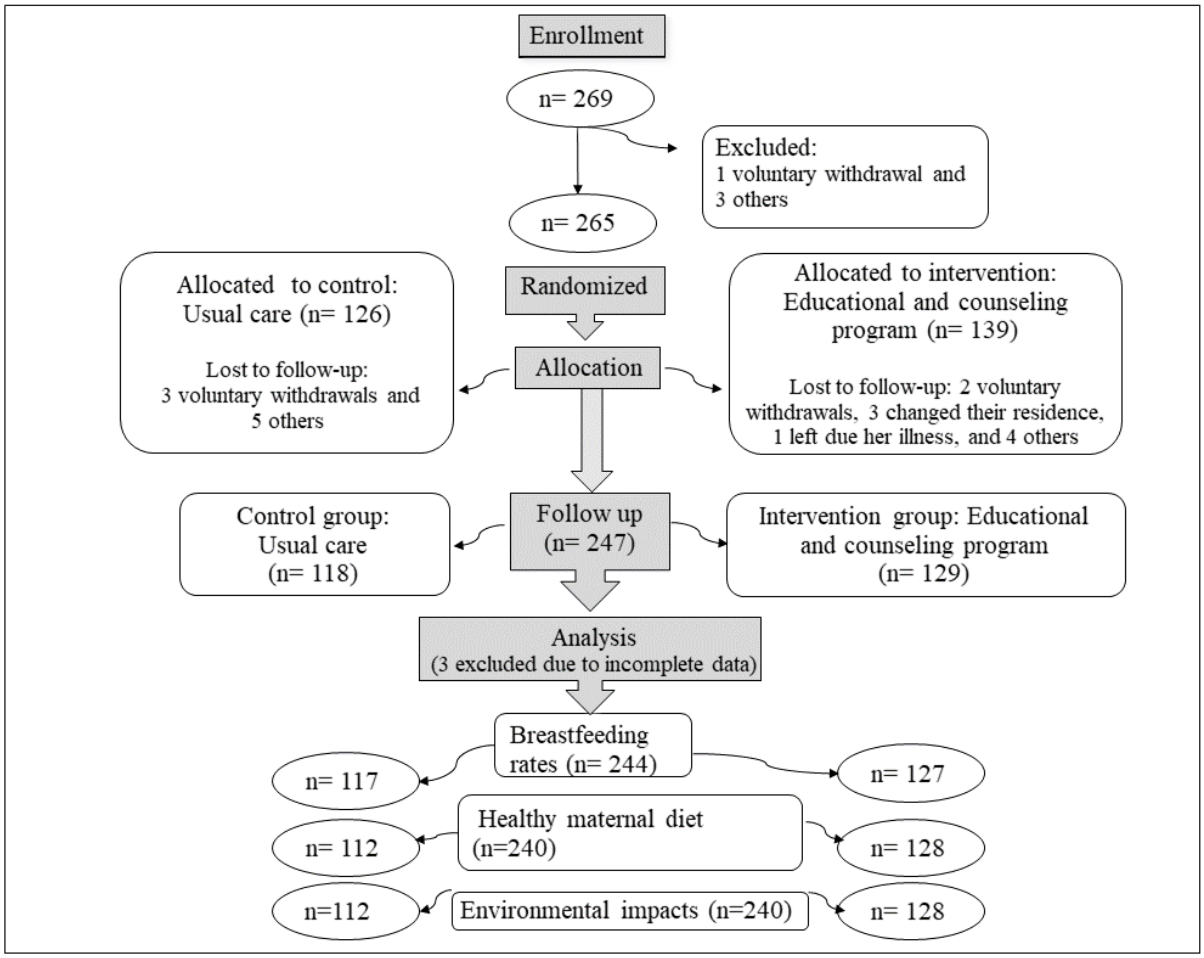

Citizen Participation

Mothers involved in breastfeeding and parenting will collaborate with the research team to evaluate and validate the content of informational pamphlets and other resources of the proposed educational interventions before implementation. Their opinions will also be considered for discussion of the project results and their subsequent dissemination to communities and the public.

Discussion

Expected Results

Starting with the objective of evaluating the effectiveness of this educational and reinforcing intervention, we expect to find the following differences between the 2 groups:

- A higher rate of EBF among EG mothers in the first 6 months of the baby's life
- A lower environmental impact associated with infant feeding, considering the use of accessories, the mothers' diet, and eating habits among mothers in the EG at 1 month of the baby's life
- A more healthful diet and a more adequate proportion of nutrients in the mothers' diets in the EG

It is crucial to adopt evidence-based strategies and new guidelines for complementary education in aspects such as breastfeeding, mothers' nutrition, and environmental information for the health personnel involved in the care of pregnant and breastfeeding women.

Group education, both written and audiovisual dissemination, has proven to be an effective tool used by primary care professionals. The self-determination of women and their partners as providers of food for the entire family and their role in obtaining sustainable food must be reinforced.

Once the project is completed, the guide, different materials, and courses will be used to train all health professionals in the entire territory of Catalonia involved in mother and baby care to focus on the nursing mother, providing the best conditions for a good start to breastfeeding and offering support for any difficulties that may arise. It will also include information on how to follow a healthy and environmentally friendly diet.

Limitations and Strengths

Prenatal counseling and education in primary care centers will be updated with the materials created for this project, and they will be disseminated to all health centers. Potential limitations of the study during data collection are that participants in the CG, without specific intervention, can obtain information through social networks, television, and the internet. However, to control for this fact, we included a final questionnaire to verify the sources of information.

In this study, all participants intending to breastfeed are included, as its objective is not to influence their decision of

XSL•FO
RenderX

which type of infant feeding, they choose but rather to support those who wish to breastfeed for an extended period (6 months), helping them achieve this successfully and without complications. For the infant feeding rates in the general case, we conducted phase 1 of the Green Mother project, and from the previous study, LACTEM [21], from the same geographic area, we found that 96% of the pregnant mothers wanted to breastfeed. The mothers' illness, country of birth, and age were also observed as facilitators of and barriers to breastfeeding. During the recruitment phase 1 of this project, which investigated the environmental impact of different types of infant feeding, we observed a 93% breastfeeding rate at hospital discharge [12].

Furthermore, all users with the intention of breastfeeding are included because the objective of this study is not to influence the type of breastfeeding that the participants will choose but rather to ensure that those who want to breastfeed for a long time can achieve it. This also helps to see the rates of those who wanted but could not initiate or follow the EBF or breastfeeding.

The development of this research will not only enable the updating of standard prenatal programs but also contribute to the achievement of the Sustainable Development Goals (SDGs), particularly those related to health (SDG 3), gender equality (SDG 5), responsible consumption (SDG 12), and climate action (SDG 13). Ultimately, strengthening health education through an environmental lens may foster a shift in attitudes toward healthier and more sustainable practices, thereby ensuring the long-term well-being of both the population and the planet [25,26].

Conclusions

The results of this study will demonstrate the effectiveness of the educational and counseling intervention. This intervention will not only help improve support for prolonged breastfeeding and therefore benefit children's health but also help reduce environmental impact, raising awareness among mothers and families of the importance of breastfeeding their children and having healthy and sustainable diets, as stated in the 2030 SDGs.

The findings of this project will be shared with both the scientific community and the public through various channels. These include continuing education courses, informative talks, community discussions, as well as presentations at conferences and congresses. By disseminating the results through these media, the knowledge gained from the project will reach a broader audience and contribute to the advancement of understanding in the field and the achievement of the SDGs.

Acknowledgments

The authors are thankful to all the researchers, community midwives, and other medical professionals who are contributing to this study. The authors thank the collaborators of the Green Mother group: Glòria Seguranyes-Guillot, Josep Mª Manresa-Dominguez, Alba Silva-Camacho, Alba Rodriguez-Biosca, Paula Amezcua-Latorre, Anna Xufré-Nadal, Carmen Barrionuevo-Ramirez, Susana Sancho-Esteban, Assumpta Prats-Oliveras, Anna Vila-Coriminas, Nuria Risques-Fernábdez, Carolina Alós-Rodríguez, Glòria Miralpeix-Pomar, Ariadna Torres-Franco, Olga Derm-Garriga, Nuria Esteban-Guzman, Ruth Marchena-Miranda, Nuria Buitrago-Torrijos, Gladys Margarita Maldonado-Aubian, Maria Camacho-Donézar, Inés María García-Martín, Cristina Olivas-Meñendez, Maria Inmaculada Rodriguez-Alvarez, Elena Imbernon-Bustamante, Anna Estruch-Riu, Alba Llobera-Sanz, Paloma Amado-Barroso, Soraya Vera-Pérez, Soraya González-Perez, MG-M, Margarita Barluengo Perez-Coussio, Antonia Garcia-López, Raquel Martinez-Mondéjar, Paula Amorós-Ferrer, Esther Cerro-Hernández, Maria Raja-Carcaña, Ana García-López, Núria Nebot-Rodrigo, Salut Puig-Calsina, Montserrat Pujol-Abajo, Mireia Monllau-Ros, Mercè Sesa-Nogueras, Rosa Tamaral-Cepas, Yolanda Tortola-Brocal, Marta Guillen-Vila, Laura Muñoz-Tamajon, Alba García-Muñoz, Mònica Martinez-Terron, Eduard Lobera-Gutierrez De Pando, Lorena Segovia-Navarro, Eva Bueno-López, Antonio López-Ollero, Laura Tarrats-Velasco, Susana Calle del Fresno, and Meritxell Casajoana-Guerrero. The authors would also like to acknowledge the enterprise inèdit, which will support them in calculating the carbon footprint impacts, especially Mariana Filipe and Edilene Pereira. The authors attest that there was no use of generative artificial technology in the generation of text, figures, or tables in this manuscript.

Funding

This project is carried out with the support of the department of health of the Generalitat de Catalunya, as part of the 2021 call for grants of the Strategic Plan for Research and Innovation in Health (Pla Estratègic de Recerca i Innovació en Salut) , for research projects in the area of primary care (SLT021/21/000063).

Data Availability

Data sharing is not applicable to this study as no datasets were generated or analyzed during this study.

Authors' Contributions

Conceptualization: RMC-F, GF-P, JC-B, RG-S
Funding acquisition: GF-P
Investigation and article review: RMC-F, LL, JC-B, RG-S, AR-L, MC-V, MMV-H, MG-M, LM-P, CV, PT-M, GC-O, GF-P

XSL•FO
RenderX

Methodology: RMC-F, JC-B, GF-P, AR-L, MMV-H, MC-V, MG-M, LM-P, LL
Software: RG-S, PT-M, CV
Supervision and validation of the translated version: LL, JC-B, CV, RMC-F, MC-V
Writing—original draft: LL, RMC-F, JC-B, AR-L
Writing—review and editing: LL, RMC-F, JC-B, AR-L

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study surveys.
[DOCX File , 17 KB - resprot_v15i1e80358_app1.docx]

Multimedia Appendix 2
CONSORT 2025 checklist.
[DOCX File , 33 KB - resprot_v15i1e80358_app2.docx]

Multimedia Appendix 3
Peer review report from the PERIS Primary Care Programme, Agency for Health Quality and Assessment of Catalonia (AQuAS; Spain).
[PDF File (Adobe PDF File), 307 KB - resprot_v15i1e80358_app3.pdf]

References

1. Joffe N, Webster F, Shenker N. Support for breastfeeding is an environmental imperative. BMJ 2019 Oct 02;367:l5646 [FREE Full text] [doi: 10.1136/bmj.l5646] [Medline: 31578194]
2. Legendre M, Pirozzi G. Convención Sobre Los Derechos Del Niño. UNICEF. URL: https://www.un.org/es/events/childrenday/pdf/derechos.pdf [accessed 2025-05-29]
3. Pérez-Escamilla R, Tomori C, Hernández-Cordero S, Baker P, Barros A, Bégin F, 2023 Lancet Breastfeeding Series Group. Breastfeeding: crucially important, but increasingly challenged in a market-driven world. Lancet 2023 Feb 11;401(10375):472-485 [FREE Full text] [doi: 10.1016/S0140-6736(22)01932-8] [Medline: 36764313]
4. La alimentación del lactante y del niño pequeño. World Health Organization (WHO). 2010. URL: https://www.who.int/es/publications/i/item/9789241597494 [accessed 2025-05-29]
5. Kramer MS, Aboud F, Mironova E, Vanilovich I, Platt RW, Matush L, Promotion of Breastfeeding Intervention Trial (PROBIT) Study Group. Breastfeeding and child cognitive development: new evidence from a large randomized trial. Arch Gen Psychiatry 2008 May 24;65(5):578-584 [FREE Full text] [doi: 10.1001/archpsyc.65.5.578] [Medline: 18458209]
6. Sayres S, Visentin L. Breastfeeding: uncovering barriers and offering solutions. Curr Opin Pediatr 2018 Aug;30(4):591-596 [FREE Full text] [doi: 10.1097/MOP.0000000000000647] [Medline: 29782384]
7. Karlsson JO, Garnett T, Rollins NC, Röös E. The carbon footprint of breastmilk substitutes in comparison with breastfeeding. J Clean Prod 2019 Jun 10;222:436-445 [FREE Full text] [doi: 10.1016/j.jclepro.2019.03.043] [Medline: 31190697]
8. Andresen EC, Hjelkrem AG, Bakken AK, Andersen LF. Environmental impact of feeding with infant formula in comparison with breastfeeding. Int J Environ Res Public Health 2022 May 24;19(11):6397 [FREE Full text] [doi: 10.3390/ijerph19116397] [Medline: 35681983]
9. Cos-Busquets J, Cabedo-Ferreiro RM, Liutsko L, Reyes-Lacalle A, García-Sierra R, Colldeforns-Vidal M, GREEN MOTHER Group. A comprehensive assessment of the environmental impact of different infant feeding types: the observational study GREEN MOTHER. J Adv Nurs 2025 Dec;81(12):8230-8241 [FREE Full text] [doi: 10.1111/jan.16473] [Medline: 39362795]
10. Serra-Majem L, Tomaino L, Dernini S, Berry EM, Lairon D, Ngo de la Cruz J, et al. Updating the mediterranean diet pyramid towards sustainability: focus on environmental concerns. Int J Environ Res Public Health 2020 Nov 25;17(23):8758 [FREE Full text] [doi: 10.3390/ijerph17238758] [Medline: 33255721]
11. Alimentació durant l'embaràs. Gencat. URL: https://canalsalut.gencat.cat/ca/salut-a-z/e/embaras-part-i-postpart/alimentacio/ [accessed 2025-05-29]
12. Reyes-Lacalle A, Cabedo-Ferreiro RM, Cos-Busquets J, Liutsko L, Colldeforns-Vidal M, García-Sierra R, et al. Characteristics, preventive factors, and barriers to breastfeeding and mixed feeding in the first month of life in Barcelona: the multicenter observational study GREEN MOTHER. Nutrients 2025 Sep 30;17(19):3109 [FREE Full text] [doi: 10.3390/nu17193109] [Medline: 41097185]
13. Green feeding-climate action from birth. International Baby Food Action Network. 2019. URL: https://www.ibfan.org/ [accessed 2025-05-29]

XSL•FO
RenderX

14. Martinelli SS, Cavalli SB. Healthy and sustainable diet: a narrative review of the challenges and perspectives. Cien Saude Colet 2019;24(11):4251-4262 [FREE Full text] [doi: 10.1590/1413-812320182411.30572017] [Medline: 31664397]
15. The Paris agreement. United Nations. URL: https://unfccc.int/process-and-meetings/the-paris-agreement [accessed 2025-05-29]
16. Metas mundiales de nutrición 2025: documento normativo sobre lactancia materna. Organizasión Mundial de la Salud. URL: https://www.who.int/es/publications/i/item/WHO-NMH-NHD-14.7 [accessed 2025-05-29]
17. Brown A. Breastfeeding as a public health responsibility: a review of the evidence. J Hum Nutr Diet 2017 Dec;30(6):759-770 [FREE Full text] [doi: 10.1111/jhn.12496] [Medline: 28744924]
18. Rana R, McGrath M, Sharma E, Gupta P, Kerac M. Effectiveness of breastfeeding support packages in low- and middle-income countries for infants under six months: a systematic review. Nutrients 2021 Feb 20;13(2):681 [FREE Full text] [doi: 10.3390/nu13020681] [Medline: 33672692]
19. Shafaei FS, Mirghafourvand M, Havizari S. The effect of prenatal counseling on breastfeeding self-efficacy and frequency of breastfeeding problems in mothers with previous unsuccessful breastfeeding: a randomized controlled clinical trial. BMC Womens Health 2020 May 05;20(1):94 [FREE Full text] [doi: 10.1186/s12905-020-00947-1] [Medline: 32370804]
20. Souza ED, Pina-Oliveira AA, Shimo AK. Effect of a breastfeeding educational intervention: a randomized controlled trial. Rev Lat Am Enfermagem 2020;28:e3335 [FREE Full text] [doi: 10.1590/1518-8345.3081.3335] [Medline: 33027400]
21. Cabedo R, Manresa JM, Cambredó MV, Montero L, Reyes A, Gol R. Types of breastfeeding and factors that influence its abandonment up to 6 months. LACTEM Study. Matronas Profesión 2019;20(2):54-61 [FREE Full text]
22. Sattari M, Serwint JR, Levine DM. Maternal implications of breastfeeding: a review for the internist. Am J Med 2019 Aug;132(8):912-920 [FREE Full text] [doi: 10.1016/j.amjmed.2019.02.021] [Medline: 30853481]
23. Ota E, Hori H, Mori R, Tobe-Gai R, Farrar D. Antenatal dietary education and supplementation to increase energy and protein intake. Cochrane Database Syst Rev 2015 Jun 02;2015(6):CD000032 [FREE Full text] [doi: 10.1002/14651858.CD000032.pub3] [Medline: 26031211]
24. Jadotte YT, Caron RM, Kearney GD. Ecosystemic theory, practice, and policy: training recommendations for environmental public health. Am J Prev Med 2022 Jan;62(1):135-144 [FREE Full text] [doi: 10.1016/j.amepre.2021.09.004] [Medline: 34774390]
25. Liutsko L. The integrative model of personality and the role of personality in a planetary health context. Pers Individ Differ 2019 Dec;151:109512 [FREE Full text] [doi: 10.1016/j.paid.2019.109512]
26. Lerner H, Berg C. A comparison of three holistic approaches to health: one health, ecohealth, and planetary health. Front Vet Sci 2017;4:163 [FREE Full text] [doi: 10.3389/fvets.2017.00163] [Medline: 29085825]
27. Cabedo-Ferreiro RM, Liutsko L, Cos-Busquets J, García-Sierra R, Colldeforns-Vidal M, Reyes-Lacalle A, GREEN MOTHER Group. Environmental impact of infant feeding type, accessories used and maternal dietary habits: the GREEN MOTHER-I project, a cross-sectional study protocol. Nutr J 2024 Aug 21;23(1):97 [FREE Full text] [doi: 10.1186/s12937-024-01000-9] [Medline: 39164727]
28. Reyes-Lacalle A, Montero-Pons L, Manresa-Domínguez JM, Cabedo-Ferreiro R, Seguranyes G, Falguera-Puig G. Perinatal contraceptive counselling: effectiveness of a reinforcement intervention on top of standard clinical practice. Midwifery 2020 Apr;83:102631 [FREE Full text] [doi: 10.1016/j.midw.2020.102631] [Medline: 32036192]
29. Jané M, Amorós P, Molina MC, Mateum A. Pla de salut afectiva i sexual (PSAS). Agència de Salut Pública de Catalunya. URL: https://scientiasalut.gencat.cat/bitstream/handle/11351/1277/pla_salut_afectiva_sexual_2013.pdf [accessed 2025-05-29]
30. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
31. Scott J, Huskisson EC. Graphic representation of pain. Pain 1976;2(2):175-184. [doi: 10.1016/0304-3959(76)90113-5]
32. Báez León C, Blasco Contreras R, Martín Sequeros E, Pozo Ayuso ML, Sánchez Conde AI, Vargas Hormigos C. Validación al castellano de una escala de evaluación de la lactancia materna: el LATCH. Análisis de fiabilidad. Index Enferm 2008 Sep;17(3):205-209. [doi: 10.4321/S1132-12962008000300012]
33. Cabedo R, Manresa JM, Cambredó MV, Montero L, Reyes A. Tipos de lactancia materna y factores que influyen en su abandono hasta los 6 meses. Matronas Prof 2019;20(2):54-61 [FREE Full text]
34. Ferrari MA. Estimación de la Ingesta por Recordatorio de 24 Horas. Diaeta 2013;31(143):43 [FREE Full text]
35. European Food Safety Authority (EFSA). Dietary reference values for nutrients summary report. EFS3 2017 Dec;14(12):e15121E [FREE Full text] [doi: 10.2903/sp.efsa.2017.e15121]
36. Kirschbaum MU. Climate-change impact potentials as an alternative to global warming potentials. Environ Res Lett 2014;9(1):034014 [FREE Full text] [doi: 10.1088/1748-9326/9/3/034014]
37. Dekker E, Zijp MC, van de Kamp ME, Temme E, van Zelm R. A taste of the new ReCiPe for life cycle assessment: consequences of the updated impact assessment method on food product LCAs. Int J Life Cycle Assess 2019 Jul 09;25(12):2315-2324 [FREE Full text] [doi: 10.1007/S11367-019-01653-3]
38. Boulay AM, Bare J, Benini L, Berger M, Lathuillière M, Manzardo A, et al. The WULCA consensus characterization model for water scarcity footprints: assessing impacts of water consumption based on available water remaining (AWARE). Int J Life Cycle Assess 2017 Jun 8;23(2):368-378 [FREE Full text] [doi: 10.1007/S11367-017-1333-8]

XSL•FO
RenderX

Abbreviations

ASSIR: Atenció a la Salut Sexual i Reproductiva (reproductive and sexual health care unit)
CG: control group
EBF: exclusive breastfeeding
EG: experimental group
GHG: greenhouse gas
REDCap: Research Electronic Data Capture
SDG: Sustainable Development Goal
WHO: World Health Organization

Edited by A Schwartz; submitted 09.Jul.2025; this is a non–peer-reviewed article;accepted 30.Nov.2025; published 26.Jan.2026.

Please cite as:
Cabedo-Ferreiro RM, Liutsko L, Cos-Busquets J, García-Sierra R, Colldeforns-Vidal M, Reyes-Lacalle A, Torán-Monserrat P, Vicente-Hernández MM, Gómez-Masvidal M, Violán C, Montero-Pons L, Falguera-Puig G, Cazorla-Ortiz G
Effectiveness of an Educational and Counseling Program (the Green Mother Project Phase 2) to Enhance Breastfeeding and Improve Mothers' Diets From an Environmental Perspective: Protocol for a Cluster Randomized Controlled Trial
JMIR Res Protoc 2026;15:e80358
URL: https://www.researchprotocols.org/2026/1/e80358
doi:10.2196/80358
PMID:

©Rosa Maria Cabedo-Ferreiro, Liudmila Liutsko, Judit Cos-Busquets, Rosa García-Sierra, Margalida Colldeforns-Vidal, Azahara Reyes-Lacalle, Pere Torán-Monserrat, M Mercedes Vicente-Hernández, Miriam Gómez-Masvidal, Concepció Violán, Laura Montero-Pons, Gemma Falguera-Puig, Gemma Cazorla-Ortiz. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Process Evaluation of a Structured Method for Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems (SIOHPS): Protocol for a Convergent Parallel Mixed Methods Study

Camilla Göras[1,2], PhD; Ann-Sofie Ersson[3,4]; Therese Hellman[5,6], PhD; Petronella Bjurling-Sjöberg[4,7], PhD; Gunnar Bergström[8,9], PhD; Robert Sarkadi Kristiansson[4,6], PhD; Malin Lohela-Karlsson[3,4,10], PhD

[1]Department of Caring Sciences, Faculty of Health and Occupational Sciences, University of Gävle, Gävle, Sweden
[2]Department of Caring Sciences, School of Health and Welfare, Dalarna University, Falun, Sweden
[3]Centre for Clinical Research Västmanland, Uppsala University, Västerås, Uppsala, Sweden
[4]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden
[5]Department of Medical Sciences, Uppsala University, Uppsala, Uppsala, Sweden
[6]Primary Care and Health, Region Uppsala, Uppsala, Sweden
[7]Centre for Clinical Research Sörmland, Uppsala University, Eskilstuna, Sweden
[8]Unit of Intervention and Implementation Research for Worker Health, Karolinska Institutet, Stockholm, Sweden
[9]Department of Occupational Health, Psychology and Sports Sciences, Faculty of Health and Occupational Studies, University of Gävle, Gävle, Sweden
[10]School of Health, Care and Social Welfare, Mälardalen University, Västerås, Sweden

Corresponding Author:
Camilla Göras, PhD
Department of Caring Sciences
Faculty of Health and Occupational Sciences
University of Gävle
Kungsbäcksvägen 47
Gävle, 80176
Sweden
Phone: 46 26 648468
Email: camilla.goras@hig.se

Abstract

Background: Health care systems are increasingly challenged by demographic shifts, rising chronic illnesses, human resource constraints, and growing efficiency demands. Improving both occupational safety and health (OSH) and patient safety (PS) management has been identified as a pivotal strategy to address these challenges. However, there is a paucity of evidence-based methods that support systematic and integrated OSH and PS. In response, the Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems (SIOHPS) intervention was developed, guided by the Safer Culture Framework and the Medical Research Council (MRC) framework for complex interventions.

Objective: This paper outlines the study protocol for a process evaluation embedded in the SIOHPS trial.

Methods: The process evaluation will use a convergent parallel mixed methods design. The SIOHPS trial is conducted in 13 Swedish hospital settings. Guided by the Consolidated Framework for Implementation Research (CFIR), quantitative and qualitative data will be collected before, during, and after the trial through questionnaires, telephone interviews, focus group interviews (FGIs), observations, the SIOHPS digital tool, and other relevant documentation. In line with the convergent parallel mixed methods study design, the quantitative and qualitative data will be analyzed in a stepwise manner, initially independently from each other, followed by iterative triangulation.

Results: Funding began in January 2023. The development phase was completed in early 2024, and the evaluation phase started in June 2024, with completion planned in early February 2026. Quantitative data collection for two of three clusters (baseline, 4- and 8-month follow-up) is complete, and data cleaning is underway. All qualitative data collected to date have been transcribed. Final data collection for cluster III, including the 8-month survey and FGIs, is scheduled for the end of January and early February

XSL•FO
RenderX

2026. Data analysis will begin in early 2026, with results to be disseminated through publications and conference presentations later in 2026.

Conclusions: The process evaluation will integrate quantitative and qualitative data sources to elucidate the mechanisms through which the SIOHPS intervention influences safety culture, health care worker (HCW) health, PS, and quality of care. This comprehensive approach unpacks the "black box" of the implementation process, which will provide an in-depth nuanced picture of the intervention's effectiveness and valuable insights into scalability and transferability across diverse health care contexts.

Trial Registration: ClinicalTrials.gov NCT06398860; https://clinicaltrials.gov/study/NCT06398860

International Registered Report Identifier (IRRID): DERR1-10.2196/89185

(JMIR Res Protoc 2026;15:e89185) doi:10.2196/89185

KEYWORDS

complex intervention; health care; mixed methods; occupational safety and health; patient safety; process evaluation; study protocol

Introduction

Health care systems are increasingly challenged by demographic shifts, rising chronic illness, human resource constraints, and growing efficiency demands [1]. These challenges complicate the delivery of high-quality and safe care [2], place significant strain on the work environment of health care workers (HCWs), and threaten the sustainability of health care. Occupational safety and health (OSH) and patient safety (PS) are closely intertwined fields [3]. The COVID-19 pandemic underscored this interdependence, emphasizing that protecting HCW well-being is essential to maintaining a resilient health care system [4].

Integrated OSH and PS management, to simultaneously prevent adverse outcomes in both fields, has been raised as having potential for addressing the challenges in a resource-effective manner [4,5]. However, evidence-based methods integrating OSH and PS are lacking (M Lohela-Karlsson, Associate Professor, unpublished data, January 2026), highlighting the need for the development and evaluation of such methods. In response, a theoretically grounded project was launched to develop and evaluate a structured method for OSH and PS management, called Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems (SIOHPS). To evaluate the SIOHPS intervention, a large-scale multidisciplinary trial has been performed [6].

From a theoretical perspective [7], effective integration of OSH and PS requires addressing organizational factors that support a culture of safety across multiple system levels. This complexity needs to be considered in intervention development and evaluation, preferably codesigning it with key stakeholders and addressing the needs of different system levels. However, the efficiency of complex interventions, defined by multiple interacting components, diverse outcomes, and system-level impact [8], depends not only on the intervention itself, but also on the implementation process [9,10] and sustained enactment [11].

The viability of an intervention depends on implementation outcomes, that is, stakeholders' satisfaction with the intervention (acceptability); its fit and relevance to the context, problem, and users (appropriateness); the extent to which it can be successfully deployed (feasibility); and the financial impact and resources required for the implementation effort (implementation cost) [8]. Additionally, the extent to which an intervention is implemented as intended (fidelity) by its developers is critical for internal validity and for understanding intervention effectiveness [12]. Context-sensitive adaptations of an intervention's design or delivery (adaptations) may be needed, while preferably preserving its core components [13]. Once viewed as a threat to fidelity, adaptations are currently considered vital for aligning complex interventions with real-world contexts and enhancing implementation outcomes [14]. Similarly, the extent to which an implemented intervention is sustained over time (sustainability) has been seen as a final implementation outcome [15], while being contemporary viewed as a dynamic, adaptive process that ensures continued use of an intervention in response to evolving needs [10]. Implementation outcomes are also shaped by the process through which interventions produce change, that is, enabling factors, enacting behaviors, and eventual unexpected pathways and consequences (mechanisms of impact) [7]. Hence, facilitators and barriers in the organizational, cultural, and structural environment in which an intervention is implemented (contextual conditions) [16,17] can provide valuable insights into intervention effectiveness [18].

Implementation processes of complex interventions are often referred to as the "black box," insinuating a lack of transparency that undermines reliable interpretation of outcomes and eventual scalability across health care settings. Additionally, despite growing interest, few studies address sustained interventions, probably due to varied definitions of sustainability and challenges in adapting interventions across contexts [19]. Addressing questions beyond the scope of effectiveness studies, process evaluations unpack the "black box" of the implementation process, with the purpose to understand how and why interventions succeed or fail. Such understanding can be used to assess whether interventions are viable and what adaptations and implementation strategies might be suitable in eventual future implementation across diverse contexts [18]. This paper outlines the study protocol for a process evaluation embedded in the SIOHPS trial [6].

The overall aim of this process evaluation is to understand how the SIOHPS intervention is delivered and adhered to in practice,

XSL•FO
RenderX

how context influences implementation, and the mechanisms through which it produces change.

The specific aims of the process evaluation of the SIOHPS intervention are to:

- Describe the content, frequency and dose, duration, and coverage (fidelity), as well as adaptations for contextual fit.
- Explore the contextual factors that facilitate or hinder the implementation of the intervention.
- Describe and compare short- and long-term outcomes between settings in relation to the implementation process.
- Identify variables (mechanisms of impact) that may influence or contribute to the development of the intervention outputs.

Methods

Study Design

This study was registered at ClinicalTrials.gov (NCT06398860). The process evaluation follows the Medical Research Council (MRC) framework for complex interventions [18]. A convergent parallel mixed methods design [20] is used within a hybrid type 1 effectiveness-implementation framework.

The following section briefly describes the overall design and setting of the SIOHPS project. Next, the intervention (the SIOHPS method) and its underlying theoretical assumptions are provided, visualized in a logic model (Figure 1). Thereafter, participants, the procedure, and the process evaluation are outlined, including domains of the evaluation and the theoretical rationale for studying them. Finally, the planned analysis is described.

Figure 1. Logic model illustrating the interrelationship between the process evaluation components: contextual factors; the implementation process and mechanisms of impact (coloured boxes); and the theoretical foundation, core components of the intervention, and expected intervention outputs (grey boxes). COM-B: capability, opportunity, motivation, and behavior.

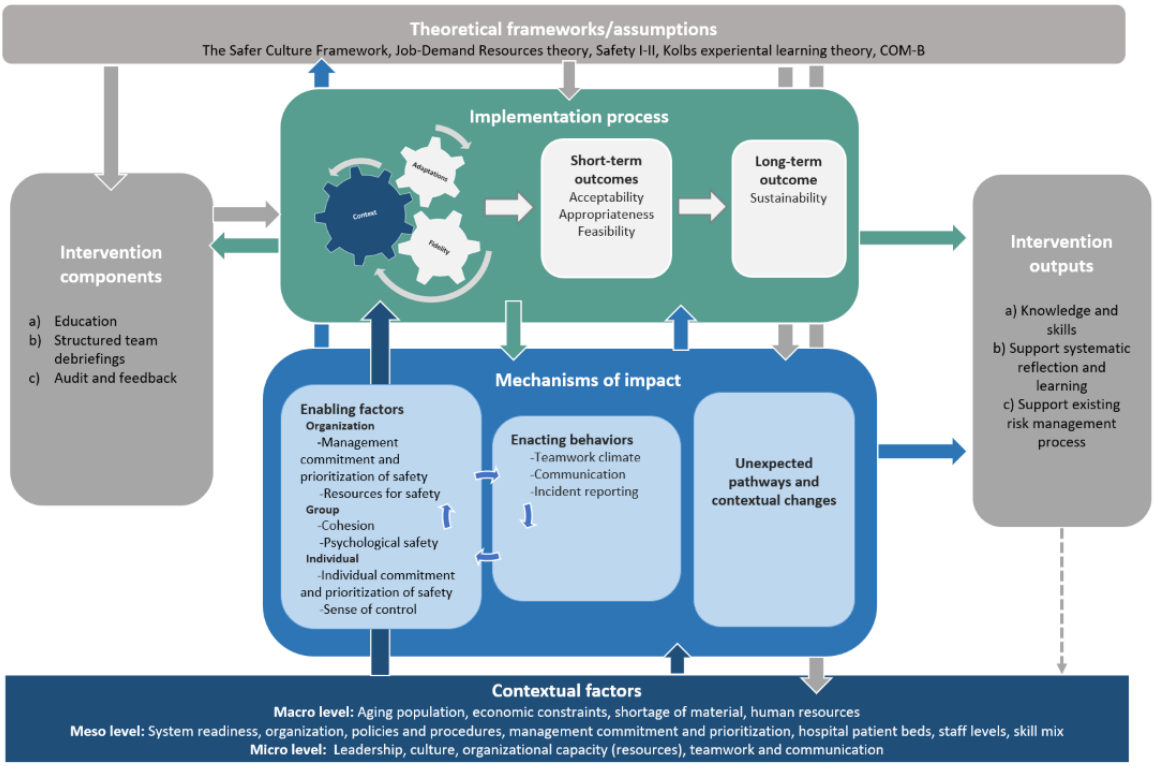

The SIOHPS Project

The SIOHPS project evaluates whether a structured, integrated approach to managing OSH and PS can improve HCW health and quality of care. It follows the MRC framework for complex interventions [8], emphasizing stakeholder codesign and contextual adaptation. The project is led by a core team consisting of the principal investigator (PI); two regional project managers, responsible for coordinating units in respective regions; and an administration assistant. The core team members bring complementary research and clinical expertise from health care fields, such as OSH management, occupational health services, PS, and PS culture. An advisory board (with clinical and academic background and expertise, including OSH, PS, and safety culture; intervention and implementation research; and roles such as chief medical officer, PS coordinator, and statistician) provides ongoing guidance.

The project comprised a development phase (2023-2024); it also comprises an evaluation phase. During the development phase, the program theory, intervention components, and evaluation design were refined in accordance with O'Cathain et al [21]. This work was coproduced with key stakeholders, including managers, HCWs involved in PS tasks, safety representatives, and central organizational functions. The development phase will be reported in a separate paper.

The evaluation uses a large-scale, multidisciplinary hybrid type 1 effectiveness-implementation framework [22], with the

XSL•FO
RenderX

primary focus on effectiveness and cost-effectiveness and the secondary focus on the implementation process. The effectiveness study uses a pragmatic stepped-wedge cluster-controlled trial across two Swedish regions, involving 13 hospital units with round-the-clock care. Units transition from control to intervention in three steps, with at least 4 units per step and ≥360 HCWs. The sample size was determined to ensure sufficient unit-level representation, acknowledging variation in HCW numbers across units. All participating units had to provide round-the-clock care. The study protocol for the trial is outlined in a separate paper [6], while further details of the process evaluation are outlined later. Evaluation of effectiveness and process will be conducted by two separate but collaborative teams, led by the PI and one of the regional project managers, respectively. Integrating the teams supports coherent data collection, prevents duplication of effort, minimizes participant burden, and enables meaningful integration between implementation data and process outcomes [18].

The PI holds overall responsibility for ensuring the evaluation of established study protocols. The evaluation design, including the development of data collection instruments, was carried out in close collaboration with the advisory board during the development phase. The advisory board will continue to be consulted and receive regular updates, providing feedback on interim findings and overall implementation.

The SIOHPS Intervention

The developed intervention (ie, the SIOHPS method) consists of three core components: (1) *targeted education*, delivered in live format or in online format provided by the research team, to build foundational knowledge and practical skills to adhere to the SIOHPS method; (2) *team debriefings* (also known as end-of-shift huddles or after-action reviews), which HCWs are encouraged to perform after every shift; and (3) monthly *audit and feedback* by management in collaboration with the support functions based on data from team debriefings (Figure 1). The components target the needs of different system levels, involving management at both organizational and group levels, support functions (including those involved in PS work and the safety representatives of the work environment), and HCWs. The method encourages an integrated OSH and PS perspective when identifying patterns and root causes and taking measures, supported by a novel digital tool to be used in team debriefings, audit, and feedback.

Theoretical Assumptions

The SIOHPS intervention and its process evaluation is theoretically grounded in the principles of safety culture theory [7] and complemented with four additional theoretical foundations: job demand-resources theory [23], the Safety I-II approach [24], Kolb's experiential learning theory [25], and the COM-B (capability, opportunity, motivation, and behavior) model for behavior change [26]. To help clarify causal assumptions, a logic model was developed by the core team based on the relevant literature and in collaboration with the advisory board (Figure 1).

Briefly, the logic model visualizes that to evaluate the *implementation* and sustainability of SIOHPS, it is crucial to consider fidelity, adaptation, and context, as these domains interact in dynamic ways. The fidelity-adaptation dilemma [14,16,27] highlights the tension between preserving the core components of an intervention and adapting it to the local context. Striking this balance is crucial for sustainability and effectiveness, shaped by contextual factors across system levels (micro, meso, and macro) [16]. The core components of the *intervention* are designed to address the theoretical *mechanisms of impact* (ie, enabling factors and enacting behaviors) that, based on the Safer Culture Framework, influence the safety culture in the workplace [7]. Although the enabling factors do not constitute an integral component of SIOHPS or the chain of effects, they are nevertheless considered a contextual factor that possesses the capacity to exert influence on the outcome of the intervention. Additionally, unexpected pathways and consequences, due to contextual conditions and changes, affect both short- and long-term process outcomes and intervention output.

In the short term, the expected intervention output includes enhanced HCW participation and learning, alongside improved working conditions (eg, job demand and resources) and improved safety culture. These improvements are anticipated to positively affect work engagement [28], HCW health [29,30], productivity and PS [31-33], and perceived quality of care [31], contributing to reduced sick leave and improved health care quality [33,34]. The intermediate- and long-term effects are hypothesized to be mediated by improved safety culture and psychological safety [33]. However, both short- and long-term outcomes are dependent on the implementation process. Depending on whether participants adhere to the intervention (fidelity adaptations), effects are achieved on staff-related outcomes, patient-related outcomes, and safety culture (OSH culture and PS culture). Adherence to the intervention can be explained by the COM-B model [26].

Participants and Procedure

Participants will include nursing assistants, registered nurses, physicians, managers, paramedical services, and support functions, together with internal facilitators. Stakeholders at the meso level, those functions involved in OSH and PS, will also contribute as informants. All HCWs will be encouraged to take part in the intervention, but to participate in the study, they must work at least 50% full-time at the unit. Data will be collected from various participants and stakeholders throughout the project period using a range of quantitative and qualitative methods. The sample size has been determined pragmatically for various parts of the project. The number of units included and HCWs was based on a power calculation for the intervention effect. Each unit was assumed to include an average of 30 HCWs. Based on the sample size calculation, at least 12 units and 360 HCWs were required to be included in the study to achieve sufficient statistical power for further analyses. The estimate was sufficiently conservative to detect the lowest expected intervention effect. The sample size was considered a minimum requirement, with primary emphasis placed on achieving the estimated number of units, as the number of staff members varied across units. A dropout rate of less than 20% for

XSL•FO
RenderX

participating units and an attendance rate of at least 70% of all HCWs have been assumed in this study. For further details, please see the effectiveness study protocol [6]. The data collection process is illustrated in a Gantt chart (Figure 2). Overviews of the explored process domains, data collection methods, and their operationalization in relation to the three core components of SIOHPS are presented in Tables 1-4 for the *implementation process* and Tables 5 and 6 for the *mechanisms of impact*.

Survey data from HCWs will be collected in parallel with the effectiveness evaluation and comprise both quantitative measures (eg, Likert-scale items) and qualitative data from open-ended questions (Tables 1-6). All HCWs who fulfill the inclusion criteria will be invited to answer the questionnaires at each of the three measurement points (baseline, 4 months, and 8 months). Additional quantitative data will be collected from the SIOHPS digital tool. Qualitative data will be collected through individual telephone interviews with managers at each health care unit at the beginning of the study (n=13) and after 4 months (n=13). Focus group interviews (FGIs) with managers, support functions, and internal facilitators will be conducted at each health care unit after 8 months (n=13).

The performance of the SIOHPS method during team debriefings will be examined through nonparticipating observations conducted twice at each unit, after 1 and 4 months, respectively (n=at least 1 observation per data collection, yielding a minimum of 26 observations in total). The observations will be conducted by a single researcher and will not audio- or video-recorded in order to minimize potential inhibitory effects on team behavior. The observations will be documented contemporaneously in a template, and reflexive field notes will be documented after each observation. The observation protocol comprises two parts: intervention fidelity and psychological safety. The intervention fidelity part will be based on the SIOHPS method used during team debriefings. The psychological safety part will be based on O'Donovan et al's work [35], with adjustments to fit debriefings under a short duration with possible interruptions. The observer will be one of the researchers (author ASE), who is well informed about the intervention and the study. Additionally, the observer will be trained in pilot observations, in which the observer and the PI independently will rate the same team debriefings. These will be discussed to ensure consistency in interpretation. The protocol will be refined accordingly, followed by additional pilot observations to further train the observer and to ensure that the protocol functions according to its purpose.

Documentation in logbooks will occur iteratively during the different phases, as described in the MRC framework [18]. Additionally, qualitative data will be gathered through face-to-face interactions, telephone calls, video conferences and informal observations, and continuous process notes documented in connection with every contact with participating units.

All data will be stored pseudonymized, with only the date and time of the interviews documented, and will be analyzed solely at the group level.

Figure 2. Gantt chart illustrating data collection during the process evaluation phase, visualized in relation to the process evaluation components in accordance with the logic model. The boxes represent the type of data collected, the timing, and the month from intervention start when each data collection activity was conducted. C: cluster (C1-C3); HCW: health care worker; m: month (1=1 month, 4=4 months, 8=8 months); Pl: project logbook; SIOHPS: Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems.

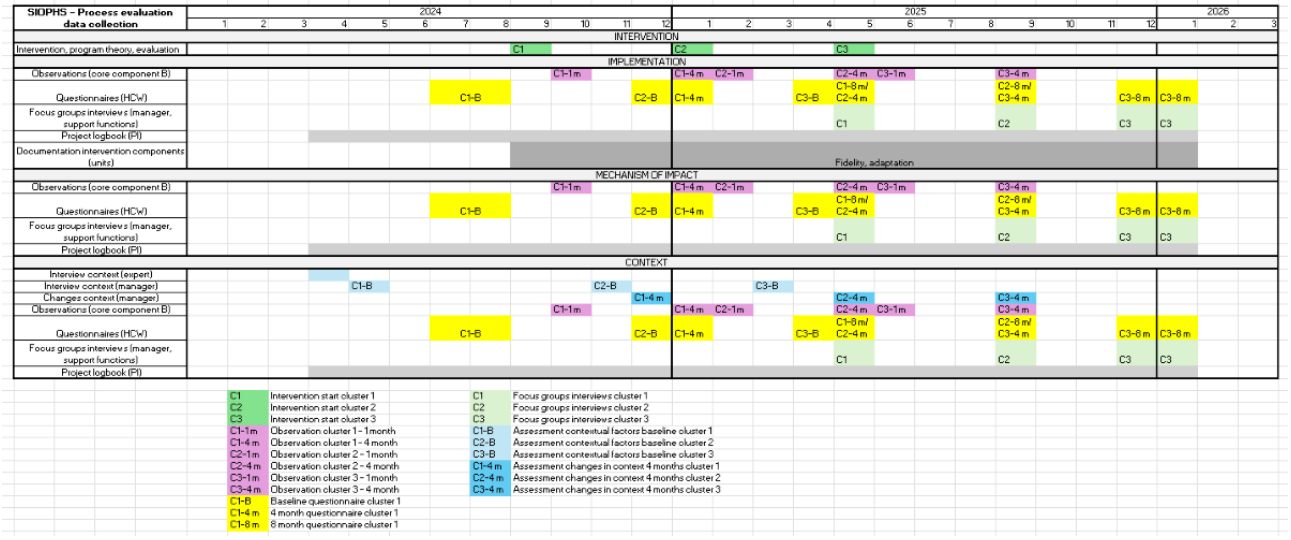

XSL•FO
RenderX

Table 1. Overview of the "fidelity" process domain, data collection, and operationalization in relation to the three core components for exploration of the SIOHPS[a] implementation.

Dimension, core component, and data sources	Data collection methods (number of planned data collections, where relevant)	Operationalization	Data type
Content and targeted education			
First-line managers and support functions	FGIs[b] (n=13)	Experience of the targeted education	Qualitative
Content and team debriefings			
HCWs[c]	Direct observations (≥26)	Adherence to the protocol	Quantitative
HCWs	Direct observations (≥26)	Interruptions and the causes of observed interruptions	Quantitative, qualitative
First-line managers and support functions	FGIs (n=13)	Experience of structured team debriefings	Qualitative
Content, and audit and feedback			
First-line managers and support functions	FGIs (n=13)	Experience of audit and feedback	Qualitative
Duration and team debriefings			
HCWs	Direct observations (≥26)	Duration of structured team debriefings	Quantitative
Duration, and audit and feedback			
First-line managers	Documentation and FGIs (n=13)	Experience of audit and feedback	Qualitative
HCWs	Questionnaires (n=2)	Perception of regular feedback	Quantitative
Frequency and dose, and targeted education			
HCWs	Questionnaires (n=2)	Participation in targeted education and in what format	Quantitative
First-line managers and support functions	Documentation and FGIs	Participation in targeted education and in what format	Quantitative, qualitative
Frequency and dose, and team debriefings			
SIOHPS digital tool	All HCWs	Number of reported estimates, including distribution by work shift	Quantitative
HCWs	Questionnaires (n=2)	Participation in structured team debriefings in the past 4 months	Quantitative
Frequency and dose, and audit and feedback			
First-line managers and support functions	FGIs (n=13)	Experience of conducted audit and feedback	Qualitative
HCWs	Questionnaires (n=2)	Perception of regular feedback on reported data in the SIOHPS digital tool	Quantitative
Coverage (reach) and targeted education			
First-line managers and support functions	Documentation/logbook	Participation in targeted education and in what format	Qualitative
Coverage (reach) and team debriefings			
First-line managers and support functions	FGIs (n=13)	Presence during team debriefings	Qualitative
HCWs	Direct observations (≥26) and questionnaires (n=2)	Participating professionals during structured team debriefings	Quantitative, qualitative
Coverage (reach), and audit and feedback			
First-line managers and support functions	FGIs	Providing audit and feedback	Qualitative

XSL•FO
RenderX

Dimension, core component, and data sources	Data collection methods (number of planned data collections, where relevant)	Operationalization	Data type
HCWs	Questionnaires (n=2)	Perception of regular feedback on reported data in the SIOHPS digital tool	Quantitative

[a]SIOHPS: Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems.
[b]FGI: focus group interview.
[c]HCW: health care worker.

Table 2. Overview of the "adaptations" process domain, data collection, and operationalization in relation to the three core components for exploration of the SIOHPS[a] implementation.

Core components and data source	Data collection methods (number of planned data collections, where relevant)	Operationalization	Data type
Targeted education			
First-line managers	Documentation in a logbook	Contextual adaptations of education sessions	Qualitative
Targeted education and team debriefings			
HCWs[b]	Direct observations with memos (≥26)	Adaptations from the original protocol and as intended by developers	Qualitative
HCW	A follow-up question postobservation (≥26)	Experiences of adaptation needs, the cause, and adaptations implemented	Qualitative

[a]SIOHPS: Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems.
[b]HCW: health care worker.

Table 3. Overview of the "context" process domain, data collection, and operationalization in relation to the three core components for exploration of the SIOHPS[a] implementation.

Core components and data sources	Data collection methods (number of planned data collections, where relevant)	Operationalization	Data type
Targeted education, team debriefings, and audit and feedback			
Central support functions working within the area of PS[b] or the work environment and first-line managers	Telephone interviews (n=13+4)	Mapping of factors at micro and meso levels with a potential impact on the conditions for systematic quality and safety work related to the work environment and PS	Quantitative, qualitative
First-line managers	Telephone interviews (n=13)	Experience of organizational/contextual changes	Qualitative
First-line managers and support functions	FGIs[c] (n=13)	Experience of barriers and facilitators	Qualitative
Logbook	Documentation	Contextual changes over time	Qualitative

[a]SIOHPS: Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems.
[b]PS: patient safety.
[c]FGI: focus group interview.

XSL•FO
RenderX

Table 4. Overview of the "short- and long-term outcomes" process domain, data collection, and operationalization in relation to the three core components for exploration of the SIOHPS[a] implementation.

Dimension, core components, and data sources	Data collection methods (number of planned data collections, where relevant)	Operationalization	Data type
Acceptability, appropriateness, and feasibility (short term); team debriefings			
HCWs[b]	Questionnaires (n=2)	Perception of the SIOHPS digital tool	Quantitative
Acceptability, appropriateness, and feasibility (short term); targeted education, team debriefings, and audit and feedback			
First-line managers and support functions	FGIs[c] (n=13)	Experience of the SIOHPS method	Qualitative
Sustainability (long term); team debriefings			
HCWs	Questionnaires (n=2)	Perception of motivational factors involvement, communication, psychological safety, and consequences	Quantitative
HCWs	Questionnaires (n=2)	Perception of intervention complexity, compatibility, and advantages of the method	Quantitative
Sustainability (long term); audit and feedback			
First-line managers and support functions	FGIs (n=13)	Experiences of intervention complexity and advantages of the method	Qualitative
Sustainability (long term); targeted education, team debriefings, and audit and feedback			
HCWs	Questionnaires (n=2)	Perception of management commitment	Quantitative
First-line managers	FGIs (n=13)	Experience of support from head of department	Qualitative
First-line managers and support functions	FGIs (n=13)	Experience of preconditions and goals	Qualitative
Sustainability (long term); team debriefings, and audit and feedback			
First-line managers and support functions	FGIs (n=13)	Experience of internal and external support	Qualitative

[a]SIOHPS: Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems.

[b]HCW: health care worker.

[c]FGI: focus group interview.

XSL•FO
RenderX

Table 5. Overview of the "enabling factors" process domain, data collection, and operationalization in relation to the three core components for exploration of the mechanisms of impact.

System level and factors	Data source	Data collection methods (number of planned data collections, where relevant)	Operationalization	Data type
Organizational level				
Management commitment and prioritization of safety	First-line managers and support functions	FGIs[a] (n=13)	Experiences of SIOHPS[b]	Qualitative
Resources for safety	First-line managers and support functions	FGIs (n=13)	Experiences of SIOHPS	Qualitative
Group level				
Cohesion	HCWs[c]	Questionnaire	Perceptions of communication and participation	Quantitative
Psychological safety	HCWs	Questionnaire	Perceptions of psychological safety	Quantitative
Individual level				
Individual commitment and prioritization of safety	HCWs	Questionnaire	Perceptions of work engagement	Quantitative
Sense of control	HCWs	Questionnaire	Perceptions of resources	Quantitative

[a]FGI: focus group interview.

[b]SIOHPS: Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems.

[c]HCW: health care worker.

Table 6. Overview of the "enacting factors" process domain, data collection, and operationalization in relation to the three core components for exploration of the mechanisms of impact.

Factors	Data source	Data collection methods (number of planned data collections, where relevant)	Operationalization	Data type
Teamwork climate	HCWs[a]	Questionnaire	Perceptions of quality of collaboration between team members	Quantitative
Communication	HCWs	Questionnaire	Perceptions of communication	Quantitative
Incident reporting	Incident-reporting system	—[b]	Frequency of reported incidents	Quantitative
Unexpected pathways and consequences				
Contextual conditions and changes	First-line managers	Telephone interviews (n=13)	Experienced contextual conditions and changes	Qualitative
Contextual conditions and changes	First-line managers	FGIs[c] (n=13)	Barriers and facilitators	Qualitative

[a]HCW: health care worker.

[b]Not applicable.

[c]FGI: focus group interview.

Process Evaluation

The process evaluation will explore the implementation process, including fidelity, adaptations, context and process outcomes (short and long term), and mechanisms of impact [18], with theoretical assumptions based on the logic model (Figure 1). An overview of explored process domains, data collection, and operationalization in relation to the three core components of SIOHPS is presented in Tables 1-6.

Implementation Process

The implementation process describes how the intervention is delivered and adhered to in practice [18]. During this process, particular attention will be given to *intervention fidelity*, *adaptations*, and *contextual conditions*, as these factors are expected to influence both *short-* and *long-term outcomes*. Short-term outcomes, including acceptability, appropriateness, and feasibility, are primarily assumed to impact the intervention output and also sustainability of the intervention.

XSL•FO
RenderX

Fidelity

Given that SIOHPS is evaluated using an effectiveness-implementation hybrid design, fidelity will be assessed primarily as intervention fidelity, including fidelity of delivery and participant adherence. Fidelity of delivery refers to the interventionists' actions and reflects the quality, consistency, and alignment of delivered activities in relation to the intervention protocol. Participant adherence captures the extent to which participants engage with the intervention as intended [36]. This part will be explored across four dimensions (*content*, *coverage*, *frequency* and *dose*, and *duration* [10]) related to the three core components of SIOHPS using multiple data sources.

Content will be explored through participants' experiences of education, information, and support provided for systematic management, complemented by observations of adherence to protocol, interruptions, and the causes of interruptions during team debriefings. Experience of the content in team debriefings and the audit and feedback process will also be examined.

The duration of team debriefings will be observed, and the approximate time spent preparing presentation materials for audit and feedback sessions during workplace meetings will be obtained from the FGIs.

Frequency and dose will be assessed through participation in the targeted education, including the delivery format. Additional information includes the frequency and distribution of team debriefings across work shifts and patterns of participation over time. The audit and feedback component will be examined by assessing the regularity of feedback provided through the SIOHPS digital tool and through routine workplace meetings or other communication channels.

Coverage will be evaluated by documenting participation in education (live or online), presence during team debriefings, and the professional groups represented. Materials used when presenting audit and feedback results in workplace meetings will be collected as part of the documentation. Participants will also be surveyed regarding the regularity of feedback received based on reports in the SIOHPS digital tool.

Adaptations

To explore adaptations, information on participation in targeted education and the format will be collected as these are essential for contextual fit. In addition, experiences regarding the need for adaptations, their underlying causes, and the specific adaptations implemented will also be explored.

Context

The context domain will be guided by the Consolidated Framework for Implementation Research (CFIR) [37]. Prior to the start of the intervention, a comprehensive mapping of contextual factors in participating regions and units will be conducted through telephone interviews. A follow-up after 4 months will be performed to identify any organizational changes that could influence the implementation of SIOHPS. In addition, barriers and facilitators will be explored following completion of the intervention. Contextual changes during the study period will be continuously documented in the project logbook.

Short- and Long-Term Outcomes

During the implementation process, fidelity, adaptation, and contextual factors are expected to influence short-term outcomes, such as *acceptability*, *appropriateness*, and *feasibility*, which, in turn, are expected to mediate long-term outcomes, such as *sustainability*.

Acceptability, appropriateness, and feasibility of the SIOHPS digital tool will be assessed by examining perceptions of its complexity, compatibility, and perceived advantages. Experiences of the SIOHPS method as a whole will also be explored.

To examine long-term sustainability, motivational factors, such as involvement, will be assessed. Perceptions of management commitment will also be collected, and experiences of organizational support, prerequisites, and goals will be explored.

Mechanisms of Impact

Mechanisms of impact refer to the processes through which interventions produce change. For SIOHPS, these mechanisms are explored from a theoretical perspective, with *enabling factors* from the Safer Culture Framework. The Safer Culture Framework [7] and the COM-B model [26] informed the operationalization of the mechanisms of impact and guided data collection through questionnaires, telephone interviews, and FGIs (Tables 5 and 6).

At the *organizational level*, management commitment and prioritization of safety and investment in safety competencies are expected to foster an environment conducive to effective implementation and sustained impact by enhancing fidelity, promoting safety culture, motivating HCWs, and enabling capacity building. On a *group level*, cohesion and psychological safety facilitate open communication, engagement, and collective learning through trust, key processes for cultural change, and long-term sustainability. At the *individual level*, commitment, prioritization of safety, and perceived control are assumed to be drivers of sustained engagement.

Together, these enabling factors influence *enacting behaviors*, such as communication, teamwork, incident reporting, and ownership of safety practices, thereby activating mechanisms for lasting improvements in OSH culture [7]. SIOHPS is designed to reinforce these processes by promoting effective communication and regular incident reporting.

Contextual factors and organizational changes may also generate unanticipated pathways influencing short- and long-term outcomes, reflecting the dynamic interplay between external demands, internal dynamics, and organizational structures.

Data Management and Analysis

In line with the MRC framework [18], qualitative findings will be integrated with quantitative measures of process variables and short- and long-term outcomes, particularly where these are expected to influence the SIOHPS intervention's effectiveness and functionality. All analyses will be explorative and framed by the underlying logic model (Figure 1). Aligned with a developmental evaluation approach, the analysis will remain open to emergent findings and local contextual

XSL•FO
RenderX

adaptations, supported by continuous learning [38]. This will enable the study to capture how the intervention functions in practice and to identify refinements for future large-scale implementation [8,39].

Quantitative Data

For all quantitative data (Tables 1-6), analyses will include descriptive and comparative statistics to summarize and explore patterns in process and outcome domains across measurement points and settings. Descriptive analyses will be used to explore intervention delivery (eg, coverage, frequency and dose, duration, and adaptations), while comparative analyses will be applied to examine associations and potential relationships between process variables, contextual factors, and short- and long-term outcomes. Depending on data type, distributions, and study-specific requirements, appropriate parametric and nonparametric tests will be conducted using established statistical software, such as R (R Foundation for Statistical Computing), IBM SPSS, or equivalent.

Qualitative Data

For all qualitative data (Tables 1-6), an inductive-deductive approach will be applied guided by a thematic analysis outlined by Braun and Clarke [40], using NVivo software. Interviews will be audio-recorded and transcribed verbatim. Data analysis will follow an iterative process, allowing insights from initial telephone interviews, surveys, nonparticipant observations, and FGIs to inform subsequent data collection. The analysis process will follow a customary process, including repeated reading, coding, and categorizing, with gradually increasing abstraction. The final abstraction level will be determined based on the depth of data and the requirement of the different specific study aims.

Triangulation

In line with the convergent parallel mixed methods study design, all data will be analyzed in a stepwise manner. Initially, quantitative and qualitative data will be analyzed independently from each other, followed by iterative triangulation according to Creswell and Plano-Clark [41]. The breadth and complexity of the data, combined with data collection before, during, and after the intervention phase, means that different parts of the quantitative and qualitative results will be merged in multiple combinations to meet each specific aim. The formulation of the study aims reflects different analytic logics (convergent, exploratory, and explanatory) across aims, which will guide the integration of quantitative and qualitative findings, without altering the overall convergent design. Based on each aim, integration will occur through systematic comparison and complementary quantitative and qualitative results to inform understanding of observed outcomes. The different combinations of quantitative and qualitative data used in triangulation are outlined in Tables 1-6 for each process domain. Integration will be achieved through merging quantitative and qualitative results using complementary strategies, including joint displays, side-by-side comparisons, narrative weaving by aim, and matrix mapping of integrated findings to components of the underlying logic model, following the typology described by Fetters et al [42]. The logic model will further serve as a theoretical framework to guide the interpretation of interrelationships between contextual factors, implementation processes, mechanisms of impact, and intervention outputs (Figure 1). This strategy aims to strengthen the trustworthiness [43] of the process evaluation and to evaluate and refine the underlying logic model in alignment with the empirical results.

Ethical Considerations

The project was approved by the Swedish Ethics Review Authority (IDs 2023-02402-01 and 2024-01407-02). Informed consent was collected from the participants before data collection. Project datasets will be stored in secure storage for research data provided by Region Västmanland. Members of the Systematic and Integrated Occupational Health and Patient Safety (SIOHPS) project core team will be given access to the cleaned datasets. To ensure confidentiality, all data forms and data files will be pseudonymized, and project team members will be blinded to any identifying participant information. The code key will be stored separately. Access to the code key will be restricted to researchers in the core research team.

Results

Research funding for the project began in January 2023. The development phase was completed in the beginning of 2024, including refinement of the program theory, intervention components, and evaluation design in close collaboration with key stakeholders. The evaluation phase started in June 2024, with completion planned in early February 2026. The effectiveness trial protocol has been published, and description of the development phase is currently in preparation. Quantitative data collection for two of three clusters (baseline, 4- and 8-month follow-up) is complete, and data cleaning of these datasets is in progress. Interim feedback has been provided to participating units in clusters I and II. All qualitative material collected to date (contextual mapping and FGIs) has been fully transcribed. Planning for the final data collection for cluster III is underway and scheduled for the end of January and early February 2026, including the 8-month follow-up survey and FGIs. Data analysis will start in February-March 2026. Results will be disseminated through publications and conference presentations starting in 2026.

Discussion

Summary

This protocol outlines the process evaluation embedded in the SIOHPS trial, an effectiveness-implementation hybrid type 1 study assessing an integrated approach to systematic OSH and PS management. Conducting a process evaluation alongside the trial is essential to clarify how the intervention is delivered, how contextual factors influence implementation, and how through them change is generated [18,44]. Because the hybrid type 1 design primarily targets effectiveness [22], fidelity focuses on intervention fidelity, including fidelity of delivery and participant adherence [36] instead of implementation fidelity. The logic model (Figure 1) will guide the analytical process [18] and support examination of how fidelity, adaptations, and contextual conditions interact across system levels.

XSL•FO
RenderX

A mixed methods design [41] is appropriate for capturing the complexity of the SIOHPS intervention. Quantitative process data collected at multiple time points will describe implementation patterns and indicators of sustainability, while qualitative data will be gathered iteratively throughout implementation. This reflects the intervention's complexity and the dynamic nature of its implementation context. Multiple mixed methods will support meaningful data integration and accommodate emerging contextual changes. Coproduction with stakeholders will help ensure that diverse perspectives are presented.

Comparison With Prior Work

Implementation of complex interventions often occurs in organizational environments characterized by imbalances between job demands and available resources [23]. Consistent with previous research, factors such as organizational readiness [45], leadership engagement [46], and traditional separation of systematic OSH management and PS work are likely to influence the implementation of SIOHPS.

Organizational interventions aimed at strengthening PS culture have shown positive effects on HCW working conditions and health [47], yet few are designed to support systematic OSH management or reinforce OSH culture. Despite the interdependence between OSH and PS, a recent scoping review did not identify interventions targeting improvements in both safety cultures simultaneously (M Lohela-Karlsson, Associate Professor, unpublished data, January 2026). This gap likely reflects longstanding organizational silos within health care. Within SIOHPS, the work processes are integrated, which is essential for influencing both PS culture and OSH culture, thereby supporting the interventions primary long-term outcomes. Such organizational integration introduces additional complexity, as the intervention interacts dynamically with local structures, workflows, and contextual conditions. This aligns with contemporary perspectives emphasizing the need to embrace complexity in the design and evaluation of organizational interventions [48], reinforcing the importance of a rigorous process evaluation.

Balancing fidelity and adaptation [49] are particularly relevant in complex interventions. Achieving fidelity, while allowing contextually appropriate adaptations, requires preserving core components while tailoring implementation to local needs. This process evaluation will examine how this balance is negotiated across units and regions. Although specific implementation strategies are not tested, mapping contextual factors at micro and meso levels, together with documentation of barriers and facilitators, will support a nuanced understanding of how SIOHPS is enacted in practice.

With the Safer Culture Framework, encompassing enabling factors and enacting behaviors, this evaluation adopts a theory-informed lens that strengthens understanding of integrated OSH and PS management interventions. This approach positions key factors for safety culture as the key unifying mechanism connecting the two fields and influences how integrated interventions produce change. The SIOHPS evaluation therefore offers a theoretically grounded example of how mechanisms of impact operate across organizational, group, and individual levels, thereby advancing the understanding of mechanisms of change [50].

Strengths and Limitations

A strength of this process evaluation is the use of a logic model to guide data collection and analysis. This facilitates systematic assessment of fidelity, adaptations, contextual conditions, and mechanisms of impact and supports iterative refinement of theoretical assumptions. The mixed methods design and longitudinal data collection allow for the examination of temporal patterns and triangulation of findings across diverse data sources. Coproduction provides valuable insights into multiple perspectives on the implementation process.

The process evaluation also has potential limitations. Achieving adequate response rates in web-based surveys remains challenging [51] and will require close monitoring and active involvement of first-line managers and internal facilitators. The round-the-clock operational context of participating units might limit opportunities to convene support functions and internal facilitators for focus groups. The complexity of the SIOHPS intervention and the substantial volume of data collected also present analytical challenges. To mitigate these issues, analytic procedures will be anchored in the logic model and supported by iterative triangulation across methods.

Conclusion

This process evaluation is designed to illuminate how SIOHPS is implemented across diverse real-world settings and to identify the mechanisms through which it generates both short- and long-term outcomes. Through contextual mapping, exploration of barriers and facilitators, and examination of fidelity and adaptations, the evaluation is expected to generate insights that can inform future refinements of the SIOHPS intervention and guide the selection of appropriate implementation strategies for broader scale-up. The mixed methods approach, combined with iterative engagement with stakeholders, will enhance the robustness and transferability of findings and contribute to advancing knowledge on integrated complex organizational interventions, such as OSH and PS interventions.

Acknowledgments

The Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems (SIOHPS) research group is grateful for the contributions of the research assistants involved in the project. We extend our thanks specifically to Ellen Arbman Hansing, the current research assistant, for her efforts in participating during the focus group interviews and supporting the project during all phases of the process evaluation.

XSL•FO
RenderX

The SIOHPS project has received funding from the Swedish Research Council for Health, Working Life and Welfare (FORTE) for the data collection part of the process evaluation (diary number 2022-00404) and Region Västmanland internal research funds (diary number LTV-1003601). This was peer-reviewed, and now is developed with further details. The funder will play no role in the conceptualization, design, data collection, analysis, decision to publish, or preparation of the manuscript.

Data Availability

The Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems (SIOHPS) project group will be open for collaboration with other researchers, and data will be shared, provided that a research plan is submitted to the project group and aligns with the ethical approval.

Authors' Contributions

CG is the principal investigator (PI) of the process evaluation; is leading the design, analysis, and interpretation of the process evaluation; and has contributed to the development of the intervention. She wrote the first draft of the study protocol for the process evaluation and will participate in the analysis and interpretation of the data. ASE and CG are the project managers of the trial, overseeing all activities and are responsible for the implementation. ASE is responsible for managing participating workplaces at Region Västmanland. She contributed to the development of the intervention and lead the development of the supporting Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems (SIOHPS) digital tool used in the intervention. CG is responsible for managing participating workplaces at Region Dalarna. MLK is the PI of the SIOHPS project. She is also leading the design of the effectiveness and economic evaluation. As a PI for the general project, she holds overall responsibility for the evaluation of the SIOHPS project, ensuring that both effectiveness and process evaluations are conducted in line with the established plan. All data collections are conducted by ASE, CG, and a research assistant, with support from MLK. MLK, ASE, and CG are the core team of the research project. TH, GB, PBS, and RSK are research associates of the project, acting as an advisory board to the core team, contributing to the study design, logic model, and interpretation of data; together with MLK and ASE, they have critically revised the protocol draft for important intellectual content. All authors have approved the final version of the manuscript and have agreed to be accountable for all aspects of the work. The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to Generative AI Delegation Taxonomy (GAIDeT) 2025, the following tasks were delegated to a GAI tool (ChatGPT 5.2) under full human supervision: editing, proofreading, and translation. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. CG used a GAI tool to clarify wording, translate text, and check grammar in sentences and sections. Additionally, CG sought assistance in shortening extended passages, while preserving nuances and the essence of the content and maintaining an academic tone.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer-review report from the Swedish Research Council for Health, Working Life and Welfare (FORTE).
[PDF File (Adobe PDF File), 43 KB - resprot_v15i1e89185_app1.pdf]

References

1. Page B, Irving D, Amalberti R, Vincent C. Health services under pressure: a scoping review and development of a taxonomy of adaptive strategies. BMJ Qual Saf 2024 Oct 18;33(11):738-747 [FREE Full text] [doi: 10.1136/bmjqs-2023-016686] [Medline: 38050158]
2. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Care Services, Board on Global Health, Committee on Improving the Quality of Health Care Globally. Crossing the Global Quality Chasm: Improving Health Care Worldwide. Washington, DC: National Academies Press; 2018.
3. Welp A, Manser T. Integrating teamwork, clinician occupational well-being and patient safety - development of a conceptual framework based on a systematic review. BMC Health Serv Res 2016 Jul 19;16(1):281 [FREE Full text] [doi: 10.1186/s12913-016-1535-y] [Medline: 27430287]
4. Nathan Ezie K, Amekpor F, Scott G, Andigema AS, Musa S, Takoutsing B, et al. Healthcare workers' safety; a necessity for a robust health system. Ann Glob Health 2023;89(1):57 [FREE Full text] [doi: 10.5334/aogh.4167] [Medline: 37692379]
5. Keller S, Jelsma JGM, Tschan F, Sevdalis N, Löllgen RM, Creutzfeldt J, et al. Behavioral sciences applied to acute care teams: a research agenda for the years ahead by a European research network. BMC Health Serv Res 2024 Jan 13;24(1):71 [FREE Full text] [doi: 10.1186/s12913-024-10555-6] [Medline: 38218788]
6. Lohela-Karlsson M, Ersson A, Hellman T, Condén Mellgren E, Bergström G, Bjurling-Sjöberg P, et al. Effectiveness- and cost effectiveness of a structured method for systematic and integrated occupational safety and health and patient safety

XSL•FO
RenderX

management systems (SIOHPS) - a study protocol for a pragmatic stepped wedge cluster randomised controlled trial. BMC Health Serv Res 2025 Oct 22;25(1):1391 [FREE Full text] [doi: 10.1186/s12913-025-13537-4] [Medline: 41126149]
7. Kilcullen MP, Bisbey TM, Ottosen MJ, Tsao K, Salas E, Thomas EJ. The Safer Culture Framework: an application to healthcare based on a multi-industry review of safety culture literature. Hum Factors 2022 Feb 22;64(1):207-227 [FREE Full text] [doi: 10.1177/00187208211060891] [Medline: 35068229]
8. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
9. Nilsen P, Kirk JW, Gunnarsson KU, Thomas K. Tempering implementation optimism: distinguishing between efficacy and effectiveness in implementation research. Implement Sci Commun 2025 Aug 23;6(1):90 [FREE Full text] [doi: 10.1186/s43058-025-00781-2] [Medline: 40849653]
10. Nilsen P, Birken SA. Handbook on Implementation Science. Cheltenham, UK: Edward Elgar; 2020.
11. Nilsson K, Bååthe F, Andersson AE, Wikström E, Sandoff M. Experiences from implementing value-based healthcare at a Swedish University Hospital - a longitudinal interview study. BMC Health Serv Res 2017 Feb 28;17(1):169 [FREE Full text] [doi: 10.1186/s12913-017-2104-8] [Medline: 28241823]
12. Proctor EK, Bunger AC, Lengnick-Hall R, Gerke DR, Martin JK, Phillips RJ, et al. Ten years of implementation outcomes research: a scoping review. Implement Sci 2023 Jul 25;18(1):31 [FREE Full text] [doi: 10.1186/s13012-023-01286-z] [Medline: 37491242]
13. Stirman SW, Gamarra J, Bartlett B, Calloway A, Gutner C. Empirical examinations of modifications and adaptations to evidence-based psychotherapies: methodologies, impact, and future directions. Clin Psychol (New York) 2017 Dec;24(4):396-420 [FREE Full text] [doi: 10.1111/cpsp.12218] [Medline: 29593372]
14. Moore JE, Bumbarger BK, Cooper BR. Examining adaptations of evidence-based programs in natural contexts. J Prim Prev 2013 Jun 19;34(3):147-161 [FREE Full text] [doi: 10.1007/s10935-013-0303-6] [Medline: 23605294]
15. Proctor E, Luke D, Calhoun A, McMillen C, Brownson R, McCrary S, et al. Sustainability of evidence-based healthcare: research agenda, methodological advances, and infrastructure support. Implement Sci 2015 Jun 11;10(1):88 [FREE Full text] [doi: 10.1186/s13012-015-0274-5] [Medline: 26062907]
16. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. Implement Sci 2013 Oct 02;8(1):117 [FREE Full text] [doi: 10.1186/1748-5908-8-117] [Medline: 24088228]
17. Kitson A, Harvey G, McCormack B. Enabling the implementation of evidence based practice: a conceptual framework. Qual Health Care 1998 Sep 01;7(3):149-158 [FREE Full text] [doi: 10.1136/qshc.7.3.149] [Medline: 10185141]
18. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]
19. Wiltsey Stirman S, Kimberly J, Cook N, Calloway A, Castro F, Charns M. The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research. Implement Sci 2012 Mar 14;7(1):17 [FREE Full text] [doi: 10.1186/1748-5908-7-17] [Medline: 22417162]
20. Steinmetz-Wood M, Pluye P, Ross NA. The planning and reporting of mixed methods studies on the built environment and health. Prev Med 2019 Sep;126:105752 [FREE Full text] [doi: 10.1016/j.ypmed.2019.105752] [Medline: 31226344]
21. O'Cathain A, Croot L, Duncan E, Rousseau N, Sworn K, Turner KM, et al. Guidance on how to develop complex interventions to improve health and healthcare. BMJ Open 2019 Aug 15;9(8):e029954 [FREE Full text] [doi: 10.1136/bmjopen-2019-029954] [Medline: 31420394]
22. Landes SJ, McBain SA, Curran GM. Reprint of: an introduction to effectiveness-implementation hybrid designs. Psychiatry Res 2020 Jan;283:112630 [FREE Full text] [doi: 10.1016/j.psychres.2019.112630] [Medline: 31722790]
23. Bakker AB, Demerouti E, Sanz-Vergel A. Job demands–resources theory: ten years later. Annu Rev Organ Psychol Organ Behav 2023 Jan 23;10(1):25-53 [FREE Full text] [doi: 10.1146/annurev-orgpsych-120920-053933]
24. Aven T. A risk science perspective on the discussion concerning Safety I, Safety II and Safety III. Reliab Eng Syst Saf 2022 Jan;217:108077 [FREE Full text] [doi: 10.1016/j.ress.2021.108077]
25. Kolb D. Experiential Learning: Experience as the Source of Learning and Development. NJ: Prentice-Hall; 1984.
26. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
27. Stirman SW, Miller CJ, Toder K, Calloway A. Development of a framework and coding system for modifications and adaptations of evidence-based interventions. Implement Sci 2013 Jun 10;8(1):65 [FREE Full text] [doi: 10.1186/1748-5908-8-65] [Medline: 23758995]
28. Platania S, Morando M, Caruso A, Scuderi VE. The effect of psychosocial safety climate on engagement and psychological distress: a multilevel study on the healthcare sector. Safety 2022 Sep 02;8(3):62 [FREE Full text] [doi: 10.3390/safety8030062]
29. Cheng H, Yang H, Ding Y, Wang B. Nurses' mental health and patient safety: an extension of the Job Demands-Resources model. J Nurs Manag 2020 Apr 25;28(3):653-663 [FREE Full text] [doi: 10.1111/jonm.12971] [Medline: 32052511]

XSL•FO
RenderX

30. Wang J, Wu C, Tsai Y, Weng S, Hsu Y. The effects of safety climate on psychosocial factors: an empirical study in healthcare workplaces. J Patient Saf 2021 May 3;18(2):e528-e533 [FREE Full text] [doi: 10.1097/pts.0000000000000856] [Medline: 34009872]
31. Lake E, Sanders J, Duan R, Riman K, Schoenauer K, Chen Y. A meta-analysis of the associations between the nurse work environment in hospitals and 4 sets of outcomes. Med Care 2019 May;57(5):353-361 [FREE Full text] [doi: 10.1097/MLR.0000000000001109] [Medline: 30908381]
32. Sturm H, Rieger MA, Martus P, Ueding E, Wagner A, Holderried M, WorkSafeMed Consortium. Do perceived working conditions and patient safety culture correlate with objective workload and patient outcomes: a cross-sectional explorative study from a German university hospital. PLoS One 2019 Jan 4;14(1):e0209487 [FREE Full text] [doi: 10.1371/journal.pone.0209487] [Medline: 30608945]
33. O'Connor DB, Hall LH, Johnson J. Job strain, burnout, wellbeing and patient safety in healthcare professionals. In: Montgomery A, van der Doef M, Panagopoulou E, Leiter MP, editors. Connecting Healthcare Worker Well-Being, Patient Safety and Organisational Change. Aligning Perspectives on Health, Safety and Well-Being. Cham: Springer; 2020:11-23.
34. Teoh K, Hassard J. Linking organisational factors and patient care: does healthcare workers' well-being matter? In: Montgomery A, van der Doef M, Panagopoulou E, Leiter MP, editors. Connecting Healthcare Worker Well-Being, Patient Safety and Organisational Change. Aligning Perspectives on Health, Safety and Well-Being. Cham: Springer; 2020:41-57.
35. O'Donovan R, McAuliffe E. Exploring psychological safety in healthcare teams to inform the development of interventions: combining observational, survey and interview data. BMC Health Serv Res 2020 Aug 31;20(1):810 [FREE Full text] [doi: 10.1186/s12913-020-05646-z] [Medline: 32867762]
36. Páez A, Nunan D, McCulloch P, Beard D. The influence of intervention fidelity on treatment effect estimates in clinical trials of complex interventions: a metaepidemiological study. J Clin Epidemiol 2025 Jan;177:111610 [FREE Full text] [doi: 10.1016/j.jclinepi.2024.111610] [Medline: 39528004]
37. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
38. Patton MQ. What is essential in developmental evaluation? On integrity, fidelity, adultery, abstinence, impotence, long-term commitment, integrity, and sensitivity in implementing evaluation models. Am J Eval 2016 Mar 22;37(2):250-265 [FREE Full text] [doi: 10.1177/1098214015626295]
39. Skivington K, Craig N, Craig P, Rycroft-Malone J, Matthews L, Simpson SA, et al. Introducing the revised framework for developing and evaluating complex interventions: a challenge and a resource for nursing research. Int J Nurs Stud 2024 Jun;154:104748 [FREE Full text] [doi: 10.1016/j.ijnurstu.2024.104748] [Medline: 38564983]
40. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Los Angeles: SAGE; 2022.
41. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research. Los Angeles: SAGE; 2017.
42. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
43. Lincoln YS, Guba EG. Naturalistic Inquiry. Beverly Hills, CA: SAGE; Jun 1985.
44. French C, Pinnock H, Forbes G, Skene I, Taylor SJC. Process evaluation within pragmatic randomised controlled trials: what is it, why is it done, and can we find it?-a systematic review. Trials 2020 Nov 09;21(1):916 [FREE Full text] [doi: 10.1186/s13063-020-04762-9] [Medline: 33168067]
45. Weiner BJ, Clary AS, Klaman SL, Turner K, Alishahi-Tabriz A. Organizational readiness for change: what we know, what we think we know, and what we need to know. In: Albers B, Shlonsky A, Mildon R, editors. Implementation Science 3.0. Cham: Springer; 2020:101-144.
46. Karanika-Murray M, Gkiontsi D, Baguley T. Engaging leaders at two hierarchical levels in organizational health interventions: insights from the intervention team. Int J Workplace Health Manag 2018 Sep 25;11(4):210-227 [FREE Full text] [doi: 10.1108/ijwhm-07-2018-0086]
47. Finn M, Walsh A, Rafter N, Mellon L, Chong HY, Naji A, et al. Effect of interventions to improve safety culture on healthcare workers in hospital settings: a systematic review of the international literature. BMJ Open Qual 2024 May 07;13(2):e002506 [FREE Full text] [doi: 10.1136/bmjoq-2023-002506] [Medline: 38719514]
48. von Thiele Schwarz U, Sørensen O, Tafvelin S, Roczniewska M. Complexity embraced: a new perspective on the evaluation of organisational interventions. Work Stress 2024 May 08;38(4):373-379 [FREE Full text] [doi: 10.1080/02678373.2024.2349000]
49. von Thiele Schwarz U, Aarons GA, Hasson H. The value equation: three complementary propositions for reconciling fidelity and adaptation in evidence-based practice implementation. BMC Health Serv Res 2019 Nov 21;19(1):868 [FREE Full text] [doi: 10.1186/s12913-019-4668-y] [Medline: 31752846]
50. Lewis C, Klasnja P, Powell B, Lyon A, Tuzzio L, Jones S, et al. From classification to causality: advancing understanding of mechanisms of change in implementation science. Front Public Health 2018;6:136 [FREE Full text] [doi: 10.3389/fpubh.2018.00136] [Medline: 29868544]
51. Shiyab W, Ferguson C, Rolls K, Halcomb E. Solutions to address low response rates in online surveys. Eur J Cardiovasc Nurs 2023 May 25;22(4):441-444 [FREE Full text] [doi: 10.1093/eurjcn/zvad030] [Medline: 36827086]

XSL•FO
RenderX

Abbreviations

FGI: focus group interview
HCW: health care worker
MRC: Medical Research Council
OSH: occupational safety and health
PI: principal investigator
PS: patient safety
SIOHPS: Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems

Edited by J Sarvestan; submitted 08.Dec.2025; this is a non–peer-reviewed article;accepted 09.Jan.2026; published 27.Jan.2026.

Please cite as:
Göras C, Ersson AS, Hellman T, Bjurling-Sjöberg P, Bergström G, Sarkadi Kristiansson R, Lohela-Karlsson M
Process Evaluation of a Structured Method for Systematic and Integrated Occupational Safety and Health and Patient Safety Management Systems (SIOHPS): Protocol for a Convergent Parallel Mixed Methods Study
JMIR Res Protoc 2026;15:e89185
URL: https://www.researchprotocols.org/2026/1/e89185
doi:10.2196/89185
PMID:

©Camilla Göras, Ann-Sofie Ersson, Therese Hellman, Petronella Bjurling-Sjöberg, Gunnar Bergström, Robert Sarkadi Kristiansson, Malin Lohela-Karlsson. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Combined Acupoint Massage and Abdominal Mirabilite Application for Accelerating Gastrointestinal Recovery in Pediatric Patients After Endoscopic Retrograde Cholangiopancreatography: Protocol for a Randomized Controlled Trial

Xiaowei Pu[1*], MS; Shan Wang[2*], MS; Chen Cui[1], PhD; Libing Liang[2], MS; Qunsong Shen[2], MS; Wei Luan[3*], PhD; Caiqin Wu[2*], PhD

[1]Department of Pediatrics, Shanghai University of Traditional Chinese Medicine Affiliated Shuguang Hospital, Shanghai, Shanghai, China

[2]School of Nursing, Shanghai University of Traditional Chinese Medicine, Shanghai, Shanghai, China

[3]Department of Nursing, Shanghai University of Traditional Chinese Medicine Affiliated Shuguang Hospital, Shanghai, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Caiqin Wu, PhD
School of Nursing
Shanghai University of Traditional Chinese Medicine
No. 1200 Cai Lun Road, Pudong New Area, Shanghai
Shanghai, Shanghai, 201203
China
Phone: 86 15802173609
Fax: 86 21 51322234
Email: yuwuchina@aliyun.com

Abstract

Background: Endoscopic retrograde cholangiopancreatography (ERCP) is crucial for managing pediatric hepatobiliary diseases but frequently results in postoperative gastrointestinal dysfunction such as delayed flatus and defecation, which can prolong recovery and increase the risk of complications. Nonpharmacological interventions such as acupoint massage and mirabilite application offer potential benefits, but evidence for their efficacy and synergy in children after ERCP is lacking.

Objective: This study aims to evaluate the individual and combined effects of acupoint massage and abdominal mirabilite application on accelerating gastrointestinal recovery in pediatric patients following ERCP.

Methods: A single-center, single-blind, randomized controlled trial will be conducted involving 72 children aged between 2 and 12 years who have undergone ERCP. Participants will be randomly allocated to 1 of the 4 groups: acupoint massage group (stimulating ST-36, ST-25, CV-12, and PC-6 for 2 minutes per point; twice daily), mirabilite group (abdominal application for 1 hour; twice daily), combination group (both interventions; at least 1 hour apart), or conventional care group. The primary outcome is the time to first postoperative flatus and defecation. The secondary outcomes include the frequency of postoperative nausea and vomiting and inflammatory markers. Data will be analyzed using intention-to-treat analysis with SPSS, using ANOVA, chi-square, and nonparametric tests as appropriate.

Results: This trial is currently in progress, having received funding in January 2023 and July 2023. Participant recruitment commenced on November 1, 2024, with an anticipated completion date of October 30, 2026. As of January 2026, a total of 38 pediatric patients have been enrolled and randomized, while data collection and intervention delivery are ongoing. No interim analysis has been conducted. The final analysis will begin after recruitment and follow-up are complete, with primary results expected to be published in the second half of 2026.

Conclusions: This study is the first to rigorously assess the efficacy and potential synergistic effects of acupoint massage and mirabilite application for enhancing gastrointestinal recovery after pediatric ERCP. Positive findings could establish a safe, nonpharmacological protocol to improve postoperative outcomes, reduce hospital stays, and minimize complications in this vulnerable population.

Trial Registration: International Traditional Medicine Clinical Trial Registry ITMCTR2025000670; https://itmctr.ccebtcm.org.cn/mgt/project/view/3385640124071694411

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/87961

(JMIR Res Protoc 2026;15:e87961) doi:10.2196/87961

KEYWORDS

acupoint massage; mirabilite; pediatric; endoscopic retrograde cholangiopancreatography; ERCP; gastrointestinal function; postoperative recovery; randomized controlled trial

Introduction

Background

Endoscopic retrograde cholangiopancreatography (ERCP) is a significant minimally invasive diagnostic and therapeutic procedure extensively used in the diagnosis and treatment of pediatric hepatobiliary disorders, including common bile duct stones, biliary strictures, biliary fistulas, and chronic pancreatitis [1-3]. Research indicates that children undergoing ERCP experience varying degrees of postoperative gastrointestinal dysfunction, including delayed passage of flatus and stool, abdominal distension, nausea, vomiting, and diminished bowel sounds [2,4]. If not promptly addressed, these complications may lead to severe outcomes, such as intestinal perforation and pancreatitis [5]. Postoperative gastrointestinal dysfunction not only increases the child's discomfort but may also prolong hospital stays, raise intravenous nutrition requirements, and heighten the risk of complications, ultimately impacting the overall recovery process and exacerbating the economic burden on health care systems [6,7].

The mechanisms contributing to gastrointestinal dysfunction following ERCP are multifactorial, encompassing preoperative anxiety, opioid use, mechanical stimulation of the intestines from surgical procedures, vagal nerve dysfunction, and postoperative inflammatory responses [8-10]. Furthermore, the gastrointestinal development of children is incomplete, and their metabolic and neuroregulatory functions are more sensitive than those of adults, potentially exacerbating the delay in postoperative gastrointestinal recovery [11,12]. Current postoperative management strategies have shown limited effectiveness in the pediatric population. For instance, the administration of low-dose opioids after gastrointestinal surgery fails to alleviate postoperative pain and is associated with an increased incidence of postoperative gastrointestinal dysfunction [9]. Currently, there is no efficient, safe, and adherence-friendly specific protocol specifically designed to enhance gastrointestinal function in children after ERCP, highlighting the pressing need to explore nonpharmacological, low-risk alternative therapies.

Traditional Chinese medicine (TCM) posits that postoperative gastrointestinal dysfunction is categorized as "abdominal distension" or "constipation," with the underlying pathogenesis frequently associated with Qi stagnation and bowel obstruction [13-15]. Acupoint massage stimulates specific acupoints to unblock meridians and harmonize Qi and blood, thereby enhancing gastrointestinal motility [16,17]. Research has demonstrated that massage activates parasympathetic nerve activity, promoting intestinal movement [18]. The topical application of mirabilite (sodium sulfate) aligns with TCM's principle that "external treatment corresponds with internal treatment." This substance possesses a cold nature and salty flavor, allowing it to clear heat, reduce swelling, and soften hardness to disperse masses; its high osmotic properties alleviate edema of the intestinal wall and stimulate local nerve reflexes, thus accelerating the recovery of bowel function [19]. The integration of these 2 methods creates a "comprehensive internal and external treatment" model; massage regulates the overall organ function and Qi, while mirabilite addresses localized swelling and constipation. This approach is advantageous due to its noninvasive nature, the ease of application, and cost-effectiveness, making it particularly suitable for pediatric patients who are intolerant to medication.

Research on adults has confirmed the effectiveness of the topical application of mirabilite in enhancing gastrointestinal function and reducing inflammation following laparoscopic cholecystectomy and laparoscopic appendectomy for mild acute biliary pancreatitis [20,21]. Additionally, acupoint massage has been shown to improve gastrointestinal function in patients undergoing gynecologic laparoscopic procedures and gastrointestinal surgery [17,18]. Despite the potential demonstrated by single therapies in adult studies, there is a significant lack of high-quality evidence supporting the use of acupoint massage in conjunction with the topical application of mirabilite to enhance gastrointestinal function in children following ERCP. Existing research not only lacks evidence for the effectiveness of acupoint massage, topical application of mirabilite, and their combined interventions for improving gastrointestinal function following ERCP in children but also fails to standardize intervention parameters related to gastrointestinal recovery in this population, such as the choice of acupoint for massage, duration of the procedure, and dosage of mirabilite. We hypothesize that all 3 intervention strategies will shorten the recovery time of gastrointestinal function and reduce postoperative adverse reactions compared to standard care, and the combined group is expected to yield the best results.

This protocol uses a 4-group parallel control design, comprising an acupoint massage group, a mirabilite external application group, a combined intervention group, and a routine care group. The primary objective is to compare the independent and synergistic effects of acupoint massage and mirabilite external application on the recovery of gastrointestinal function, specifically the time to first flatulence and defecation following pediatric ERCP. The secondary objective is to clarify the differences in the effects of various interventions on secondary outcomes, including the frequency of postoperative nausea and vomiting (PONV), as well as inflammatory markers, such as serum amylase and white blood cell (WBC) count. Additionally, the operational standards, including massage techniques, the duration of mirabilite application, and the evaluation processes

XSL•FO
RenderX

of interventions, will be standardized to ensure methodological rigor, thereby providing a reliable basis for subsequent clinical translation.

Study Aims

This study aims to comprehensively evaluate the effects of acupoint massage, the external application of mirabilite, and their combined intervention on gastrointestinal function recovery in children following ERCP procedures. The primary objectives include comparing the efficacy of the 3 interventions—acupoint massage, mirabilite application, and the combined approach—against a routine care group, specifically in terms of shortening the time to first postoperative flatus and bowel movement. Additionally, this study aims to clarify whether the combined intervention offers advantages over individual therapies to determine the optimal clinical strategy. The secondary objectives involve assessing the degree of improvement in postoperative symptoms, including PONV; analyzing changes in inflammatory markers, including serum amylase and WBC count; and verifying the safety of the interventions regarding issues such as skin irritation and procedural tolerance. This research not only determine the optimal strategy to address the existing evidence gap in accelerated recovery following pediatric ERCP but also provides a scalable, standardized protocol for rapid postoperative recovery in pediatric patients.

Methods

Study Design

A parallel 4-group, single-blind, randomized controlled trial (RCT) will be conducted from November 1, 2024, to October 30, 2026. The study will take place in the pediatric department of Shuguang Hospital in Shanghai. After baseline assessments, eligible participants will be randomly assigned in a 1:1:1:1 ratio to 4 groups: acupoint massage group, mirabilite group, combination group, and conventional care group (Figure 1; Table 1). Table 2 presents the content and procedures used in each intervention group. The standardized intervention manual can be found in Multimedia Appendix 1. The primary outcome measure will be the time to postoperative gas passage or stool, whereas the secondary outcomes will include the incidence of PONV as well as changes in inflammatory markers. This study will adhere to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines (Multimedia Appendix 2) [22]. This protocol was registered at the International Traditional Medicine Clinical Trial Registry (ITMCTR2025000670) on April 7, 2025.

Figure 1. Flow diagram of the study design. PONV: postoperative nausea and vomiting; WBC: white blood cell.

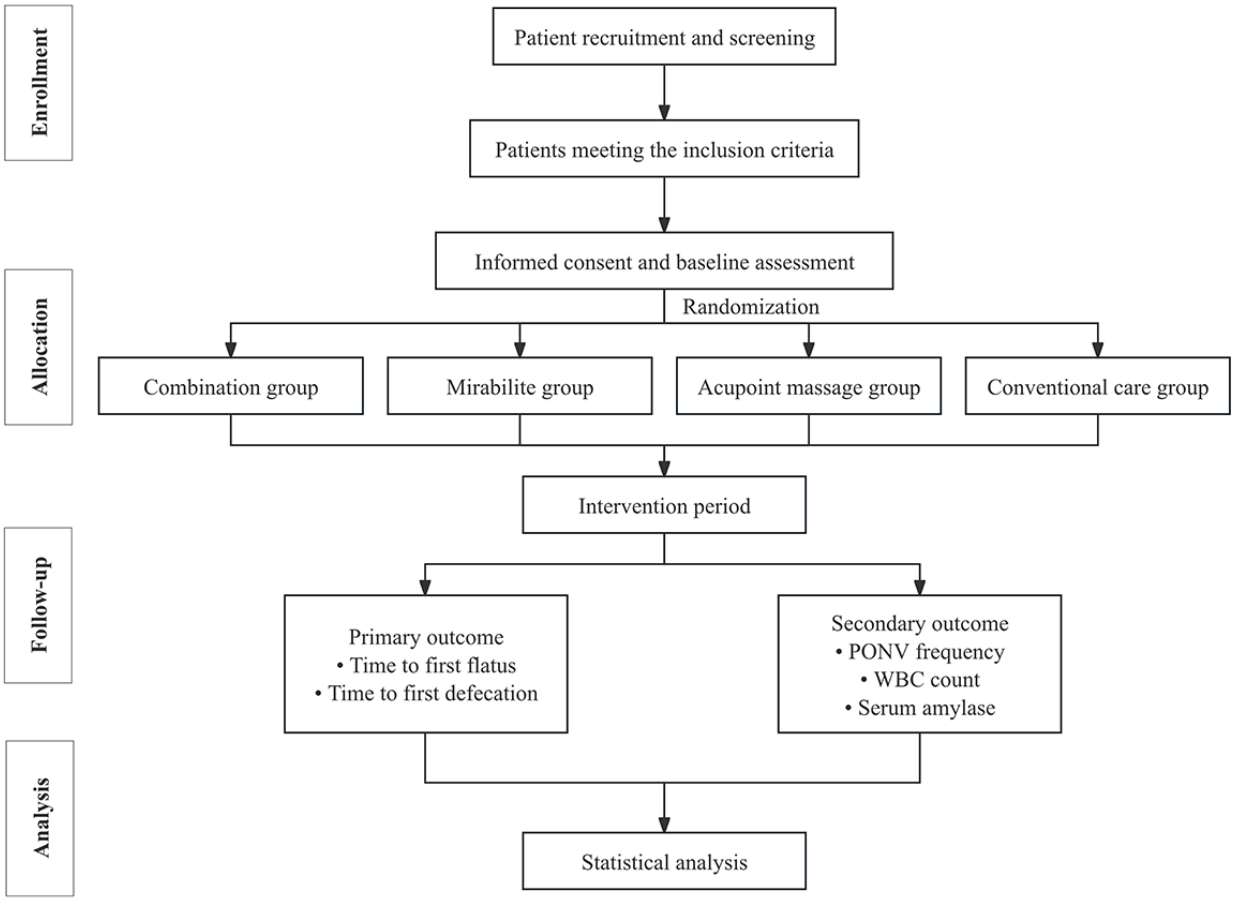

XSL•FO
RenderX

Table 1. Clinical trial procedure.

Study period	Enrollment	Baseline (before surgery)	Intervention and assessment (after surgery)	Closeout (discharge)
Enrollment				
Eligibility screen	✓			
Informed consent	✓			
Randomization			✓	
Interventions				
Combination intervention			✓	
Mirabilite application			✓	
Acupoint massage			✓	
Routine care		✓	✓	
Assessments				
General social demographic data	✓			
Primary outcomes				
Time to first flatus			✓	✓
Time to first defecation			✓	✓
Secondary outcomes				
Postoperative nausea and vomiting frequency			✓	✓
White blood cell count		✓	✓	✓
Serum amylase		✓	✓	✓
Safety monitoring				
Skin irritation			✓	✓
Massage tolerance			✓	✓

Table 2. Intervention program.

Group	Intervention components	Frequency or duration	Key operational parameters	Safety monitoring
Acupoint massage	Stimulation of Zhongwan (CV-12), Tianshu (bilateral; ST-25), Zusanli (bilateral; ST-36), and Neiguan (bilateral; PC-6)	Twice daily (2 minutes per acupoint)	• Technique: rotatory kneading+vertical pressing • Pressure: ≤0.5 kg (aged 2-5 y) and ≤1 kg (aged 6-12 y) • Sequence: CV-12→ST-25→ST-36→PC-6	• Preprocedure wound assessment • Real-time tolerance monitoring • Skin response check • Emergency stop for vomiting or pain
Mirabilite application	Periumbilical paste application (xiphoid process to pubic symphysis)	Twice daily (1 hour, with a 5-minute margin on either side)	• Allergy test • Preparation: 1:1 mirabilite: water (36 °C-38 °C) • Thickness: 0.5 cm uniform layer • Fixation: hypoallergenic tape • Constant temperature heating pack	• Preapplication skin test • Temperature monitoring (36 °C-38 °C; ΔT≤1 °C) • Visual inspection • Discontinuation for erythema or blisters
Combination group	Both interventions	Massage: 2 per day; mirabilite: 2 per day; minimum interval: 60 minutes	• Sequence: massage→mirabilite (preferred) • Prohibited: massage within 1 hour after mirabilite application • Separate documentation for both	• Combined monitoring of both interventions

XSL•FO
RenderX

Participants

Participants will be recruited from children who have undergone ERCP in the pediatric department of Shanghai Shuguang Hospital. Inclusion criteria are as follows: (1) aged between 2 and 12 years, (2) having completed ERCP, (3) deemed suitable for pediatric acupressure and external mirabilite application, (4) having provided written informed consent by legal guardians, and (5) having provided age-appropriate assent from the child. Exclusion criteria are as follows: (1) history of abdominal surgery or severe gastrointestinal dysmotility; (2) concomitant failure of vital organs, such as the heart, lung, or kidney; (3) known allergy to mirabilite or skin lesions contraindicating topical application; (4) cognitive or behavioral impairment that would compromise cooperation with acupressure; and (5) intraoperative conversion to open surgery due to major complications, such as perforation or severe bleeding. Participants will be withdrawn under the following circumstances: (1) voluntary withdrawal by the child or their parent; (2) loss of communication during the intervention that precludes further assessment; or (3) the occurrence of a serious adverse event, including marked clinical deterioration or an acute medical condition.

Eligible children will be identified through an electronic medical record system, with 2 senior physicians independently evaluating their enrollment criteria. In the event of disagreement, a third expert will adjudicate. All participants will have the research protocol thoroughly explained by a dedicated research nurse before surgery, and informed consent will be obtained from the guardians and the children.

Sample Size

G*Power (version 3.1; Heinrich Heine University Düsseldorf) software was used to calculate the sample size. We used the *F* test as the method for statistical analysis, establishing an α of .05 and a power (1–β) of .95. The estimation was based on data from a previous high-quality clinical trial, which used a 3-arm parallel design evaluating combined mirabilite application with massage, massage alone, and mirabilite application alone in patients undergoing laparoscopic hepatectomy [23]. Using the reported group means, a pooled SD of 13.03 hours, and a sample size of 30 per group from that trial, the software estimated an effect size (*f*) of 0.593. On the basis of this effect size, the calculated total sample size required for this 4-group study was 56 participants. To account for an anticipated dropout rate of 20%, the final planned total sample size is 72 participants, with approximately 18 (25%) individuals allocated to each group.

Randomization and Blinding

This study uses a randomized block design with a block length of 4. Random numbers for grouping were generated using Stata (version 18.0; StataCorp). A dedicated researcher, who will not participate in the recruitment of participants or data collection, will oversee the random grouping process, thereby ensuring that the grouping decisions remain independent of actual operations. Before the initiation of the study, personnel not involved in participant recruitment or data collection will create an allocation table based on the random number sequence and maintain this table in strict confidentiality to prevent any potential leaks. Group allocations will be concealed within sealed, opaque envelopes, which the researcher will open before the intervention to access the allocation plan for each patient.

Given the overt nature of acupoint massage and external mirabilite application, blinding of participants, their guardians, and the health care providers administering the interventions was not feasible. This represents an inherent methodological limitation of this study. To mitigate potential measurement bias due to the lack of blinding, we implemented several procedural strategies. A strict and standardized outcome assessment protocol was established, wherein independent assessors, specifically nurses who were blinded to group allocation, evaluated patients at 4-hour intervals. These evaluations were conducted through direct auscultation of bowel sounds combined with patient or parental reporting. Documentation of first flatus or defecation requires both active bowel sounds and clear verbal confirmation, thereby reducing sole reliance on subjective reporting. Researchers responsible for data collection and statistical analysis remained blinded to group assignment and accessed only deidentified datasets to maintain analytic independence. Furthermore, throughout the informed consent process and study duration, families received standardized, neutral communication, emphasizing that the study aimed to compare different postoperative care approaches, avoiding implications of differential efficacy to reduce expectation bias. All outcomes were recorded on structured case report forms (CRFs), mandating concurrent documentation of objective signs, such as auscultation findings, alongside the reporting source.

Interventions

Conventional Care Group

The children will receive standard postoperative care, including monitoring of vital signs, necessary medication, as-needed pain relief, dietary adjustments, and rehabilitation training, but will not undergo acupoint massage or the external application of mirabilite nor will any investigational interventions be administered.

Acupoint Massage Group

On the basis of conventional care, standardized acupoint massage will be performed by certified physicians and nurses to promote the recovery of organ function and alleviate symptoms of gastrointestinal dysfunction. The selected acupoints include Zusanli (bilateral; ST-36), Tianshu (bilateral; ST-25), Zhongwan (CV-12), and Neiguan (PC-6), which are informed by pertinent studies on the recovery of gastrointestinal function following laparoscopic surgery and adhere strictly to the “World Health Organization Standard Acupuncture Point Localization” [24,25]. Each acupuncture point will receive a combination of kneading and pressure techniques at a frequency of 30 repetitions per minute, with a pressure depth ranging from 0.5 to 1 cm, sustained for a duration of 2 minutes, and administered twice daily. The applied pressure must remain within a tolerable range for the child, ensuring that the skin becomes slightly reddened without eliciting pain. The child’s abdominal wound will be assessed before the procedure to avoid the puncture site.

XSL•FO
RenderX

Mirabilite Group

On the basis of conventional care, an external application of mirabilite will be administered to the abdomen, drawing on relevant studies concerning acute pancreatitis [26]. Preparation entails grinding medical-grade mirabilite into a powder, which is then mixed with warm water in a 1:1 ratio to create a paste. This paste will be evenly spread onto gauze to a thickness of approximately 0.5 cm and placed over the area surrounding the navel, extending from above the xiphoid process to below the pubic symphysis. The gauze will be anchored with breathable tape to prevent pressure on the drainage tube, while a constant temperature heating pack will maintain warmth during the application. Each application will last for 1 hour, administered twice daily. Skin reactions will be closely monitored, and the application will be discontinued immediately if redness or allergic reactions occur.

Combination Group

This group will simultaneously receive both acupoint massage and external application of mirabilite, adhering to the same protocols as the single-intervention groups, but spaced apart from each other. An interval of at least 1 hour will be maintained between the 2 interventions to avoid interference. The acupoint massage will promote the smooth flow of Qi and blood within the meridians, while the localized stimulation from the external application of mirabilite will further facilitate the recovery of gastrointestinal function and enhance therapeutic effects.

Outcome Measures

General Sociodemographic Data

To characterize the study population and assess potential confounding factors, we will collect basic demographic and clinical information from pediatric patients before the intervention, including age, gender, height, weight, and BMI, as well as disease diagnosis, specific indications for undergoing ERCP, history of abdominal surgeries, comorbidities, relevant family medical history, the number of previous ERCP procedures, guardians' educational level, annual family income and the type of residence (urban or rural). Annual family income will be categorized as: <CNY ¥50,000 (US $7183), CNY ¥50,000-100,000 (US $7183-14,366), CNY ¥100,000-200,000 (US $14,366-28,731), and >CNY ¥200,000 (US $28,731).

Primary Outcomes

The primary objective of this study is to investigate the speed of gastrointestinal function recovery in pediatric patients following ERCP. The core indicators include recording and comparing the time until the first passage of gas and the time until the first bowel movement after surgery between the 2 patient groups. Assessments will be performed every 4 hours by trained nursing staff through direct inquiry of the child or guardian, combined with auscultation of bowel sounds, to objectively validate the occurrence of flatus or defecation. These 2 time points serve as critical objective markers that reflect the recovery of gastrointestinal motility.

Secondary Outcomes

Overview

To comprehensively evaluate postoperative recovery and investigate potential mechanisms, the subsequently mentioned secondary outcomes will be monitored. The first outcome is the incidence of PONV. The frequency of PONV episodes will be recorded every 4 hours. The second outcome is postoperative inflammatory markers. To establish baseline values and track the perioperative inflammatory trajectory, venous blood samples for quantification of WBC count and serum amylase levels will be collected before surgery, at 3 hours after surgery, and on the morning of postoperative day 1. These parameters will serve to assess pancreatic functional recovery, screen for complications, and gauge the magnitude of the systemic inflammatory response. Furthermore, they will provide indirect indicators of gastrointestinal functional status, offering valuable reference data for monitoring postoperative clinical progression.

Adherence

To ensure strict adherence to the intervention plan, researchers will actively communicate with both the parents and the child to confirm the timing of the intervention, thereby ensuring that participants can attend on time. Additionally, the child will receive a gift upon enrollment and again at the conclusion of the final follow-up to enhance motivation and engagement throughout the study.

Data Collection and Management

Data collection will be conducted by trained, blinded researchers. Baseline data will be collected after the children have completed random allocation and before the initiation of the first intervention. Data will be meticulously recorded during each acupoint massage session and the application of mirabilite, which includes specific time points of the intervention, actual duration of execution, selected acupoints for acupressure, intensity of techniques, site of mirabilite application, and the child's tolerance and compliance with the intervention. Core observational indicators for postoperative recovery will be collected strictly, with the times of first flatus and first bowel movement recorded through scheduled inquiries by medical staff to the child or family, validated through physical examinations, and accurately documented. Instances of PONV will also be recorded, while inflammatory indicators will be documented at 3 hours after surgery and the following morning.

All collected data will be recorded using standardized paper CRFs. All paper CRFs will be securely stored by the principal investigator, and data entry will follow a double-entry system. An electronic database will be established using Microsoft Excel, with regular checks to identify potential errors or outliers. Source data verification will be performed regularly, comparing CRF data with original medical records to ensure transcription accuracy.

Quality Control and Precautions

The research team developed a comprehensive standard operating procedure in advance, clearly delineating the intervention specifics related to acupoint massage and the external application of mirabilite. Before the formal

XSL•FO
RenderX

commencement of the study, all intervention researchers must undergo unified theoretical and practical training, focusing on controlling the intensity of the technique and accurately locating acupoints. Consistency assessments will ensure a high degree of homogeneity and reproducibility in intervention measures among different practitioners.

Safety monitoring will be conducted continuously throughout the intervention process. During each acupoint massage session, the operator must assess and document the child's tolerance to the treatment in real time, with particular attention to controlling the acupressure intensity to minimize discomfort or injury. Before, during, and after each application of mirabilite, skin reactions in the treated area must be closely monitored and documented, emphasizing any allergic or irritative symptoms, such as rashes, itching, blisters, or ulceration, as well as the child's reported discomfort. Any adverse events or suspected adverse reactions, irrespective of their relation to the research intervention, must be recorded promptly and thoroughly according to the protocol, detailing the time of occurrence, manifestations, severity, duration, management measures, and outcomes.

Emergency response plans have been pre-established to address potential safety issues. To manage possible skin allergic reactions resulting from the external application of mirabilite, emergency medications will be readily available on-site for localized treatment; should a clear allergic reaction or severe skin irritation arise, the application of mirabilite will be immediately terminated, symptomatic treatment will be administered, and adverse events will be reported as required. If a child shows discomfort, distress, or verbal or behavioral refusal during an intervention, the procedure will be paused immediately. The researcher will comfort the child and consult with the guardian. The intervention may resume only if the distress is transient and the child and guardian agree. Continued refusal or significant distress will lead to termination of that session or withdrawal from the study. All research personnel must be thoroughly familiar with and strictly adhere to these safety protocols and emergency plans, prioritizing the safety of the children involved and promptly reporting any incidents that exceed the prescribed protocols or severe adverse events to the principal investigator and the ethics committee.

Data Analysis

Statistical analyses will be performed using SPSS (version 25.0; IBM Corp) and RStudio (Posit PBC), incorporating the *survival* package. A 2-sided significance level of .05 will be applied for all tests, and an intention-to-treat analysis will be implemented to handle missing data.

Categorical baseline variables will be summarized as frequencies and percentages, with group comparisons conducted using chi-square tests or Fisher exact tests as appropriate. Continuous variables will be described as mean with SD if normally distributed and compared via 1-way ANOVA; nonnormally distributed variables will be presented as median with IQR and compared using the Kruskal-Wallis *H* test. Time-to-event analysis will be applied to the primary outcomes, namely time to first flatus and time to first defecation. Cumulative event rates will be visualized using Kaplan-Meier curves, and overall, between-group differences will be assessed via the log-rank test. To address multiple comparisons, a prespecified hierarchical contrast approach will be used. Primary contrasts, including each intervention group compared to the routine care group, will be evaluated using a Bonferroni-adjusted significance threshold of .0167. Exploratory pairwise contrasts between the combination group and each monotherapy group will be conducted only if significant differences are observed in the primary contrasts. Additionally, a Cox proportional hazards regression model will be used to estimate adjusted hazard ratios with 95% CIs for each intervention relative to routine care while controlling for baseline covariates identified in univariable analysis with a *P* value of <.10.

The secondary outcome analysis will include comparisons of PONV incidence using chi-square tests, with Bonferroni-adjusted pairwise testing where applicable. Inflammatory markers, including WBC count and serum amylase, will be analyzed using repeated-measures ANOVA to evaluate changes at 3 hours after surgery and on the morning of postoperative day 1. Safety outcomes will be summarized descriptively, and between-group differences in adverse events will be examined using chi-square or Fisher exact tests.

Ethical Considerations

This study adheres to the Declaration of Helsinki and Chinese pediatric research ethics guidelines [27]. A structured, age-appropriate informed consent process will be implemented. Trained research nurses will explain the study purpose, procedures, risks, benefits, privacy measures, and the right to withdraw to the legal guardian. Written consent will be obtained after ensuring full understanding.

We place the highest priority on respecting the child's own will. In accordance with established pediatric ethical guidelines, a tiered assent process will be applied [27]. Children aged between 8 and 12 years will receive a simplified explanation using age-appropriate language and visual aids, and those who demonstrate understanding will be invited to sign a child assent form. For children aged between 6 and 8 years, verbal assent will be sought using simple language, and their agreement will be formally documented. For children aged between 2 and 6 years, participation will be based primarily on guardian consent; however, researchers will continuously monitor the child's emotional and behavioral responses. Persistent distress or clear refusal will be respected as an expression of dissent, leading to immediate pausing or adjustment of the procedures. The child's right to refuse is absolute. Regardless of guardian consent, if a child explicitly refuses or withdraws assent at any point, their decision will be honored without exception, resulting in nonparticipation or withdrawal from the study.

Aas a token of appreciation for participation and to support ongoing engagement, each child will receive a small, age-appropriate gift upon enrollment and upon completion of the final follow-up assessment. This compensation is not contingent on study outcomes and is provided to acknowledge the time and contribution of the participants and their families.

To protect participant privacy, all data will be anonymized using a unique study ID number. The linkage file matching study IDs

XSL•FO
RenderX

to personal identifiers will be encrypted and stored separately, with access restricted to authorized study personnel only. Sensitive information, such as socioeconomic data, will be collected solely for aggregated analysis and will be reported in grouped categories without any identifiable details. All study documents, including original paper records, consent forms, and the electronic database, will be stored securely both during and after the completion of the study. Paper materials will be kept in locked cabinets within designated hospital areas, and electronic files will be maintained on password-protected, access-controlled hospital servers. Data retention will comply with all applicable ethical guidelines and regulatory requirements.

The protocol has been reviewed and approved by the ethics committee of Shanghai Shuguang Hospital (2024-1622-205).

Results

This trial is currently in progress. This study received funding in January 2023 and July 2023. Participant enrollment commenced on November 11, 2024, with an anticipated completion date of October 30, 2026. Data collection and intervention delivery are actively ongoing. As of January 2026, a total of 38 pediatric patients have been successfully enrolled and randomized. No interim data analysis has been performed. The final data analysis is scheduled to commence upon completion of recruitment and follow-up, with primary study results expected to be available for publication in the second half of 2026.

Discussion

ERCP is a minimally invasive procedure that is essential for the diagnosis and treatment of pediatric liver, biliary, and pancreatic diseases [28-30]. However, due to the immature gastrointestinal systems in children, gastrointestinal dysfunction following ERCP is common among pediatric patients, which increases the risk of complications and negatively impacts overall recovery [5,31,32]. Therefore, this study aims to evaluate both the independent and synergistic effects of acupoint massage combined with topical mirabilite on gastrointestinal recovery after ERCP in children using an RCT design. This research addresses the evidence gap in the field of rehabilitation following ERCP in children and provides new intervention options for pediatric enhanced recovery in surgical practice.

Both acupoint massage and the external application of mirabilite are grounded in a solid theoretical framework of TCM and are supported by preliminary evidence from modern medical research. TCM posits that postoperative gastrointestinal dysfunction falls within the categories of “distention” and “constipation,” attributing the core pathology to the obstruction of Qi and impaired digestive flow [33]. By stimulating Zusanli, Tianshu, Zhongwan, and Neiguan, the massage aims to unblock the meridians and harmonize Qi and blood, thereby restoring the function of visceral Qi and alleviating abdominal distention. Research involving patients following gynecologic laparoscopic surgery indicates that acupoint massage can significantly reduce the time to first bowel movement and the time to first gas passage while also regulating gastrointestinal hormone secretion, thereby facilitating the recovery of gastrointestinal function after surgery [16,34]. The external application of mirabilite adheres to the TCM principle that external treatments correspond to internal effects. It is cold in nature and salty in taste, capable of clearing heat, reducing swelling, softening hardness, and dispersing masses. Studies reveal that this application can shorten the recovery time of gastrointestinal function, alleviate postoperative gastrointestinal symptoms, reduce the hospital stay duration, decrease hospitalization costs, and does not increase the incidence of adverse events [21,23]. The efficacy of mirabilite in patients with sepsis combined with gastrointestinal dysfunction is evident, promoting gastrointestinal recovery, accelerating clinical improvement, and enhancing prognosis [35]. Relevant research suggests that mirabilite can exhibit anti-inflammatory and antiedematous effects and enhances gastrointestinal motility by stimulating the release of vasoactive intestinal peptide, substance P, and ghrelin [36]. In summary, acupoint massage and the external application of mirabilite have been implemented in adult postoperative care and have demonstrated efficacy, showing feasibility in both theoretical and clinical applications.

The core innovation of this study is the introduction of a combined intervention model that uses acupoint massage to regulate organ Qi and enhance gastrointestinal motility, alongside the external application of mirabilite to alleviate swelling and promote bowel movements. This integrated approach aims to create a synergistic effect that comprehensively supports the recovery of gastrointestinal function.

This research uses a rigorous RCT design, systematically evaluating the effects of acupoint massage, external application of mirabilite, and their combined approach on the improvement of gastrointestinal function in children after ERCP, which clearly distinguishes the independent contributions of each intervention and their potential synergistic effects, thus providing a reliable basis for determining the optimal treatment strategy. Furthermore, this study aims to establish standardized operational parameters for acupoint massage and the external application of mirabilite for children, including acupoint selection, techniques, treatment frequency, and duration for external application, laying the groundwork for future clinical applications. The research pays attention to clinical end point and biological indicators, aiming to provide preliminary evidence on the mechanisms of action of the interventions, especially concerning their anti-inflammatory effects. Additionally, rigorously recorded safety data that meet expectations will become crucial for supporting clinical promotion. This study has significant advantages, as a noninvasive, nonpharmacological external treatment method; its high safety and ease of operation make it more acceptable to children and parents. Compared to mandatory early ambulation, this method is less painful, with low costs for mirabilite and controllable labor costs, providing good health economic value.

This study has certain limitations that need to be considered. First, the research was conducted in a single center, which may lead to participant characteristics being confined to the patient population, surgical practices, and nursing routines of that

XSL•FO
RenderX

center. Second, the highly observable nature of acupoint massage and mirabilite application made full blinding of participants, nursing staff, and families unachievable. While assessor blinding and a standardized outcome evaluation protocol incorporating objective signs, such as bowel sound auscultation, were used to reduce measurement bias, the awareness of group allocation among nurses could lead to implementation bias, and symptom reporting by families may be influenced by expectation bias. These factors may impact the reliability of key reported outcomes, specifically the timing of first flatus and defecation as well as the incidence of PONV. Additionally, while monitoring inflammatory markers and amylase levels provides some clues, the lack of direct measurement of vagal nerve activity or motilin levels limits a deeper understanding of the synergistic mechanism of the combined intervention.

Future research should concentrate on several crucial directions. First, it is essential to validate the reliability and generalizability of the results through multicenter, large-sample trials, thereby providing stronger evidence for the establishment of postoperative gastrointestinal recovery guidelines for pediatric ERCP. Second, researchers should deepen their exploration of underlying mechanisms by using gastrointestinal hormone monitoring and metagenomic analysis. Finally, it is important to broaden the applications of this strategy by investigating its effectiveness in other pediatric abdominal surgeries, ultimately benefiting a wider population of children.

In summary, this study aims to clarify the independent and synergistic effects of acupoint massage and mirabilite application on shortening gastrointestinal recovery time, alleviating related symptoms, and improving recovery processes in children after ERCP surgery through rigorous and standardized research. It aims to provide a safe, nondrug, child-friendly, and cost-effective option in clinical practice. This approach can not only optimize postoperative comfort for children, reducing hospitalization time and lowering the risk of complications and medical costs, but also provide critical evidence for developing standardized accelerated recovery pathways for pediatric ERCP. Furthermore, it promotes the scientific integration and application of TCM external treatment methods in modern pediatric perioperative care.

Acknowledgments

The authors would like to thank all the researchers who participated in this study.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Funding

This study was supported by the Foundation of Shanghai University of Traditional Chinese Medicine (grants 2023HLXL09 and 23HLZX06).

Authors' Contributions

XP, SW, CC, and WL contributed to the protocol design and writing the manuscript. SW and LL were responsible for the randomization design. XP was responsible for treatment. SW and CW were responsible for guidance and statistics. All authors were instrumental in the conceptualization of the study design, contributed substantially to the development of the manuscript, and endorsed the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Postoperative application of mirabilite and acupoint massage after pediatric endoscopic retrograde cholangiopancreatography.
[DOCX File , 23 KB - resprot_v15i1e87961_app1.docx]

Multimedia Appendix 2
SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist.
[DOCX File , 23 KB - resprot_v15i1e87961_app2.docx]

Multimedia Appendix 3
Peer review report by the Foundation of Shanghai University of Traditional Chinese Medicine (2023HLXL09, 23HLZX06).
[DOCX File , 18 KB - resprot_v15i1e87961_app3.docx]

References

XSL•FO
RenderX

1. AbiMansour JP, Martin JA. Biliary endoscopic retrograde cholangiopancreatography. Gastroenterol Clin North Am 2024 Dec;53(4):627-642. [doi: 10.1016/j.gtc.2024.08.011] [Medline: 39489579]
2. Baiu I, Visser B. Endoscopic retrograde cholangiopancreatography. JAMA 2018 Nov 20;320(19):2050. [doi: 10.1001/jama.2018.14481] [Medline: 30458491]
3. Hosseini A, Sohouli MH, Sharifi E, Sayyari A, Sridharan K, Tajalli S, et al. Indications, success, and adverse event rates of pediatric endoscopic retrograde cholangiopancreatography (ERCP): a systematic review and meta-analysis. BMC Pediatr 2023 Nov 24;23(1):596 [FREE Full text] [doi: 10.1186/s12887-023-04392-5] [Medline: 37996785]
4. Papaefthymiou A, Landi R, Arvanitakis M, Tringali A, Gkolfakis P. Endoscopic retrograde cholangiopancreatography: a comprehensive review as a single diagnostic tool. Best Pract Res Clin Gastroenterol 2025 Feb;74:101976. [doi: 10.1016/j.bpg.2025.101976] [Medline: 40210330]
5. Barakat M, Saumoy M, Forbes N, Elmunzer BJ. Complications of endoscopic retrograde cholangiopancreatography. Gastroenterology 2025 Aug;169(2):230-243. [doi: 10.1053/j.gastro.2025.03.009] [Medline: 40120770]
6. Guelrud M. Endoscopic retrograde cholangiopancreatography. Gastrointest Endosc Clin N Am 2001 Oct;11(4):585-601. [Medline: 11689358]
7. Asenov Y, Akın M, Cantez S, Gün Soysal F, Tekant Y. Endoscopic retrograde cholangiopancreatography in children: retrospective series with a long-term follow-up and literature review. Turk J Gastroenterol 2019 Feb 06;30(2):192-197 [FREE Full text] [doi: 10.5152/tjg.2018.18165] [Medline: 30459128]
8. Chen X, Zhang X, Gao R, Huang Y, Mao S, Wang B, et al. Effect of alleviating preoperative anxiety on gastrointestinal function recovery after laparoscopic high ligation of the hernia sac in children with indirect inguinal hernia. Drug Des Devel Ther 2024;18:3477-3486 [FREE Full text] [doi: 10.2147/DDDT.S461097] [Medline: 39132623]
9. Dai F, Zhang R, Deng R, Wang G, Guo H, Guo C. Regular use of low-dose of opioids after gastrointestinal surgery may lead to postoperative gastrointestinal tract dysfunction in children: a Chinese national regional health center experience sharing. BMC Gastroenterol 2023 Oct 31;23(1):369 [FREE Full text] [doi: 10.1186/s12876-023-02999-7] [Medline: 37907841]
10. Tan S, Wu G, Yu W, Li N. [Research advance in causes of postoperative gastrointestinal dysfunction]. Zhonghua Wei Chang Wai Ke Za Zhi 2016 Mar;19(3):351-355. [Medline: 27003660]
11. Neu J. Gastrointestinal development and meeting the nutritional needs of premature infants. Am J Clin Nutr 2007 Feb;85(2):629S-634S [FREE Full text] [doi: 10.1093/ajcn/85.2.629S] [Medline: 17284768]
12. Harris J, Chorath K, Balar E, Xu K, Naik A, Moreira A, et al. Clinical practice guidelines on pediatric gastroesophageal reflux disease: a systematic quality appraisal of international guidelines. Pediatr Gastroenterol Hepatol Nutr 2022 Mar;25(2):109-120 [FREE Full text] [doi: 10.5223/pghn.2022.25.2.109] [Medline: 35360381]
13. Kong Q, Chen LM, Liu CY, Li W, Yin PH. The effect of acupuncture on gastrointestinal recovery after abdominal surgery: a narrative review from clinical trials. Int J Surg 2024 Sep 01;110(9):5713-5721. [doi: 10.1097/JS9.0000000000001641] [Medline: 38759698]
14. Su H, Wen Y, Kang D. Application of refined nursing combined with comprehensive treatment of traditional Chinese and western medicine in gastrointestinal dysfunction after tumor operation. Evid Based Complement Alternat Med 2022 Jul 20;2022:4957061-4957067 [FREE Full text] [doi: 10.1155/2022/4957061] [Medline: 35911141]
15. Wang Y, Wang L, Ni X, Jiang M, Zhao L. Effect of acupuncture therapy for postoperative gastrointestinal dysfunction in gastric and colorectal cancers: an umbrella review. Front Oncol 2024 Feb 5;14:1291524 [FREE Full text] [doi: 10.3389/fonc.2024.1291524] [Medline: 38375156]
16. Ruan D, Li J, Liu J, Li D, Ji N, Wang C, et al. Acupoint massage can effectively promote the recovery of gastrointestinal function after gynecologic laparoscopy. J Invest Surg 2021 Jan 27;34(1):91-95 [FREE Full text] [doi: 10.1080/08941939.2019.1577515] [Medline: 30917712]
17. Zhang M, Chen J. Effect of ear point burying combined with acupoint massage on recovery of gastrointestinal function after gastrointestinal surgery. Minerva Pediatr (Torino) 2023 Oct;75(5):768-770 [FREE Full text] [doi: 10.23736/S2724-5276.23.07295-6] [Medline: 37284812]
18. Dehghan M, Malakoutikhah A, Ghaedi Heidari F, Zakeri MA. The effect of abdominal massage on gastrointestinal functions: a systematic review. Complement Ther Med 2020 Nov;54:102553. [doi: 10.1016/j.ctim.2020.102553] [Medline: 33183670]
19. Liu P, Qin Z, Wu S, Zeng W, Guo W. Effect of external application of Rhubarb and Mirabilite combined with warm acupuncture on the recovery after laparoscopic appendectomy. Drug Eval 2024 2024;21(01):108-111. [doi: 10.19939/j.cnki.1672-2809.2024.01.27]
20. Cai H, Du J, Luo C, Li S. External application of mirabilite before surgery can reduce the inflammatory response and accelerate recovery in mild acute biliary pancreatitis. BMC Gastroenterol 2023 Aug 02;23(1):264 [FREE Full text] [doi: 10.1186/s12876-023-02901-5] [Medline: 37532999]
21. Gao B, Ding Y, Wang X. The effect of external application of mirabilite on the recovery of gastrointestinal function after laparoscopic appendectomy. Clin Med 2024 2024;44(03):114-116. [doi: 10.19528/j.issn.1003-3548.2024.03.037]
22. Chan AW, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomized trials. Nat Med 2025 Jun;31(6):1784-1792. [doi: 10.1038/s41591-025-03668-w] [Medline: 40295741]

XSL•FO
RenderX

23. Hu L, Liu Y, Li J, Wang L, Zeng Y. Effect of external application of Glauber's salt of umbilical cord and point massage of Zusanli on postoperative gastrointestinal function in patients undergoing laparoscopic hepatectomy. J Nurs (China) 2019;26(01):5-7 [FREE Full text] [doi: 10.16460/j.issn1008-9969.2019.01.005]
24. Hu S, Chen J, Zhou Q, Jiang L. Influence of acupoint massage on gastrointestinal function recovery of patients after gynecologic laparoscopic surgery: a meta-analysis. Chin Evid-Based Nurs 2023 2023;9(05):776-780 [FREE Full text]
25. Huang L. Standard acupuncture point localization by the World Health Organization. World Health Organization (WHO).: People's Health Publishing House URL: https://wap.pmphmall.com/gdsdetail/613244-287842 [accessed 2025-05-25]
26. Zhang Q, Zhu L, Ju Y. Effect of enteroclysis with Dachengqi Decoction combined with external application of mirabilite in early enteral nutrition support of patients with moderate severe acute pancreatitis. Hebei J TCM 2025 2025;47(05):754-758. [doi: 10.3969/j.issn.1002-2619.2025.05.012]
27. Guo C, Zhang Y, Ding Q, Wang Q, Yang Y, Feng X. Ethical guidance for pediatric clinical research. Chin J Clin Pharmacol 2021 2021;37(11):1476-1480. [doi: 10.13699/j.cnki.1001-6821.2021.11.042]
28. Santángelo A, Scarpin A, Imaz F, Marino P, Vargas R, Cardona Bidart L, et al. One-step endoscopic retrograde cholangiopancreatography and laparoscopic cholecystectomy: a safe strategy in pediatrics. Cir Pediatr 2025;38(2):45-49. [doi: 10.54847/cp.2025.02.03]
29. Qin XM, Yu FH, Lv CK, Liu ZM, Wu J. Endoscopic retrograde cholangiopancreatography for diagnosing and treating pediatric biliary and pancreatic diseases. World J Gastrointest Surg 2023 Oct 27;15(10):2272-2279 [FREE Full text] [doi: 10.4240/wjgs.v15.i10.2272] [Medline: 37969723]
30. Sun R, Xu X, Zheng Q, Zhan J. Therapeutic endoscopic retrograde cholangiopancreatography for pediatric hepato-pancreato-biliary diseases: a systematic review and meta-analysis. Front Pediatr 2022 Jun 30;10:915085 [FREE Full text] [doi: 10.3389/fped.2022.915085] [Medline: 35844750]
31. Kheir FA, Moustafa L, Ahmad L, Kamil H, Mahmod J, Sandouk F. Duct-related complications of pediatric post-traumatic pancreatitis: a case report from Syria. Int J Surg Case Rep 2025 Feb;127:110885 [FREE Full text] [doi: 10.1016/j.ijscr.2025.110885] [Medline: 39809054]
32. Lorio E, Moreau C, Hernandez B, Rabbani T, Michaud K, Hachem J, et al. Pediatric ERCP: factors for success and complication-a 17-year, multisite experience. J Pediatr Gastroenterol Nutr 2023 Sep 01;77(3):413-421. [doi: 10.1097/MPG.0000000000003875] [Medline: 37399144]
33. Liu G, Zhang J, Zhao S, Li HJ. Research progress on the treatment of gastrointestinal dysfunction after laparoscopic surgery using traditional Chinese medicine external treatment methods. China Mod Doctor 2025;63(08):117-120. [doi: 10.3969/j.issn.1673-9701.2025.08.032]
34. Zhang Y, Hou C, Tang S, Duan Y, Weng S. The effect of acupuncture manipulation on gastrointestinal function recovery and gastrointestinal hormone levels after gynecological laparoscopic surgery. Mod J Integr Tradit Chin West Med 2020;29(08):93-96. [doi: 10.3969/j.issn.1008-8849.2020.08.011]
35. Wang C, Wang L, Xiang J, Sun W. Protective effect and mechanism of mirabilite external application on gastrointestinal function in patients with sepsis complicated with gastrointestinal dysfunction. Chin Mod Doctor 2022;60(07):0131-0133 [FREE Full text]
36. Tao L, Fu J, Wang F, Song Y, Li Y, Zhang J, et al. The application of mirabilite in traditional Chinese medicine and its chemical constituents, processing methods, pharmacology, toxicology and clinical research. Front Pharmacol 2023 Jan 4;14:1293097 [FREE Full text] [doi: 10.3389/fphar.2023.1293097] [Medline: 38239194]

Abbreviations

CRF: case report form
ERCP: endoscopic retrograde cholangiopancreatography
PONV: postoperative nausea and vomiting
RCT: randomized controlled trial
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TCM: traditional Chinese medicine
WBC: white blood cell

XSL•FO
RenderX

Edited by J Sarvestan; submitted 17.Nov.2025; this is a non–peer-reviewed article;accepted 05.Jan.2026; published 03.Feb.2026.

Please cite as:
Pu X, Wang S, Cui C, Liang L, Shen Q, Luan W, Wu C
Combined Acupoint Massage and Abdominal Mirabilite Application for Accelerating Gastrointestinal Recovery in Pediatric Patients After Endoscopic Retrograde Cholangiopancreatography: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e87961
URL: https://www.researchprotocols.org/2026/1/e87961
doi:10.2196/87961
PMID:

©Xiaowei Pu, Shan Wang, Chen Cui, Libing Liang, Qunsong Shen, Wei Luan, Caiqin Wu. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 03.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Feasibility and Acceptability of Barbershop-Based HIV Prevention Among Heterosexual Men in Kalangala Islands, Uganda: Protocol for a Cluster Randomized Trial (HPTN 111)

Caitlin Scoville[1*], MPH; Brenda Gati Mirembe[2*], MSc Epi, MBChB; Emily Voldal[3*], PhD; Joel Maena[2*], MBChB, MPH; Juliane Etima[2*], MA; Jayla Harris-Wisecarver[1*], MPH; Ivan Rukundo[2], BS, MS; Hassan Ssemere[2], BA, MA; Doreen Kemigisha[2*], BA, MPH; Teopista Nakyanzi[2*], BA; Rita Nakalega[2*], MBChB, MPH, MMED; Donaldson F Conserve[4*], MS, PhD; Maggie Albano[1*], MA; Molly Dyer[1*], MPH; Estelle Piwowar-Manning[5], BS, MLS; Hans Spiegel[6*], MD, PhD, DTMH; Clemensia Nakabiito[2*], MBChB, MMed; Andrew Mujugira[7*], MBChB, PhD; Deborah Donnell[3*], PhD; Zubair Lukyamuzi[2,8*], MBChB, MPH, MSc Pharm

[1]Family Health International 360, Durham, NC, United States

[2]Makerere University - Johns Hopkins University Research Collaboration, Kampala, Central Region, Uganda

[3]Fred Hutch Cancer Center, Seattle, WA, United States

[4]Milken Institute School of Public Health, George Washington University, Washington, DC, United States

[5]Department of Pathology, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[6]Kelly Government Solutions, Rockville, MD, United States

[7]Makerere University, Infectious Diseases Institute, Kampala, Central Region, Uganda

[8]Department of Pharmacology and Therapeutics, King Ceasar University, Kampala, Uganda

[*]these authors contributed equally

Corresponding Author:
Jayla Harris-Wisecarver, MPH
Family Health International 360
359 Blackwell St
Suite 200
Durham, NC, 27701
United States
Phone: 1 9156139830
Email: jharris@fhi360.org

Abstract

Background: Globally, successful strategies to engage high-risk heterosexual men in HIV prevention are scarce, resulting in limited access and uptake. Barbershops offer a potential venue for HIV prevention.

Objective: The primary objective was to evaluate the feasibility and acceptability of a barbershop-based HIV prevention initiative. The secondary objectives were to compare completion of self-initiated HIV testing between intervention and control groups, evaluate the preliminary effectiveness of the intervention on change in behaviors associated with HIV acquisition, compare interest in or use of HIV prevention services between intervention and control groups, and assess interest in long-acting preexposure prophylaxis among all participants and by study arm. The exploratory objective was to evaluate the preliminary effectiveness of the intervention on incident sexually transmitted infections.

Methods: HIV Prevention Trials Network 111 (HPTN 111; Testing a Barbershop-based HIV Prevention Initiative Among Men [TRIM]) is a cluster randomized trial conducted in Kalangala district, Uganda, among barbershop-going men. Approximately 250 men were assigned to either an intervention barbershop (n=12) or a control barbershop to receive the standard of care (n=6). Participants assigned to intervention barbershops received an intervention package that included HIV-status neutral education, distribution of HIV self-test kits, and barber-led peer group sessions. Feasibility and acceptability of the intervention were assessed from participants in the intervention group at week 26 and week 52. The study will also assess the effectiveness of the intervention on changes in HIV testing and use of prevention services. Self-reported sexual behaviors associated with HIV incidence and sexually transmitted infection incidence rates will also be compared to the standard of care.

XSL•FO
RenderX

Results: Data collection began in March 2024 and concluded in June 2025. Participants were followed for 12 months. Data analysis has been completed. The primary manuscript is expected to be submitted for publication by March 30, 2026.

Conclusions: The results of this study will provide crucial information about the feasibility and acceptability of novel interventions, such as barbershops, to impact behavior change, as well as about the engagement of heterosexual men in high HIV transmission settings in HIV prevention and treatment trials in the future.

Trial Registration: ClinicalTrials.gov NCT06148584; https://clinicaltrials.gov/study/NCT06148584

International Registered Report Identifier (IRRID): RR1-10.2196/87612

(JMIR Res Protoc 2026;15:e87612) doi:10.2196/87612

KEYWORDS

barbershop; heterosexual men; HIV prevention; HIV risk reduction counseling; HIV testing; risk factors; sexually transmitted diseases; STI testing; Uganda

Introduction

Background

The Joint United Nations Program on HIV/AIDS (UNAIDS) and partners are focused on ending the HIV/AIDS epidemic using a variety of strategies [1,2]. However, over the past decade, it has become increasingly clear that the global HIV response is inadequately engaging men [3]. This shortfall has been largely due to a lack of emphasis on acceptable and approachable ways to reach men, which has greatly affected clinical and public health outcomes and targets. Global evidence shows that, when compared to women, there are 1 million more men living with HIV who are unaware of their HIV status and 1.8 million more men not on treatment [1,2]. More AIDS-related deaths have occurred among men than among women over the last decade [4,5].

In 2023, the 95-95-95 strategy for HIV testing and treatment was 83-86-94 among men aged ≥15 years versus 91-91-94 among women counterparts [6]. However, in 2022, a meta-analysis of 129 studies from sub-Saharan Africa (SSA) found that 49% (95% CI 0.41-0.58) of men did not know their HIV status, 58% (95% CI 0.51-0.65) were not on treatment, and 21% (95% CI 0.19-0.23) on treatment were not virally suppressed [7]. The observed difference may be attributable to the meta-analysis incorporating studies conducted between 2014 and 2020, whereas the comparison data are derived from 2023 program reports. Notably, gender disparities in HIV prevention and treatment outcomes have persisted through 2024. Recent estimates indicate that 84%, 87%, and 94% of men know their HIV status, are on treatment, and are virally suppressed, respectively, compared with 92%, 91%, and 95% of women, according to UNAIDS. Suboptimal HIV prevention and care services for men disproportionally increase both their risk of death and HIV transmission [4]. Unfortunately, undiagnosed and untreated HIV among men leads to continued transmission of HIV, especially to adolescent girls and young women, the most susceptible and affected age group for new HIV infections [4,8]. Therefore, diagnosing and treating HIV in men is not only essential for promoting men's health but also for breaking the cycle of HIV transmission [4,9].

In Eastern and Southern Africa, the HIV epidemic is largely heterosexually transmitted, and heterosexual men are greatly affected [10,11]. However, HIV services increasingly target key populations [12-15], with limited emphasis on heterosexual men [7,16-18]. Fishing communities are especially affected; for example, the Kalangala Islands have an HIV prevalence of 12%, twice the national average of 5.1%, and an HIV incidence of 231 per 100,000 persons [19-21]. Additionally, men make up four times the population of women in the islands, with 48.4% of them unaware of their HIV status [20,22-24].

Compared to women, men and boys are less likely to test for HIV, initiate antiretroviral therapy, and remain engaged in care [4,25,26]. A significant body of research and experience shows a range of complex, multilevel barriers to male engagement in HIV-related services [4,26-28]. These include fear of knowing one's HIV status and related stigma [29,30], and prevailing cultural, social, and gender norms—such as equating illness with "weakness" and viewing clinical settings as "female" spaces [31,32]. Men in Africa often assume their status matches their partner's [33] and worry about loss of income from taking time off work [30,34-37]. Health systems also create structural barriers by focusing on reproductive-age women who can access care through antenatal care, and maternal and child health programs—entry points that do not exist for men [4]. Limited clinic hours and facility-based service models further exclude men, especially those who work [4,26].

HIV testing is the gateway to prevention and care. The World Health Organization recommends HIV self-testing (HIVST) as an additional HIV testing approach, and it is feasible and acceptable among various populations, including heterosexual men [38-40]. Over 4.8 million test kits have been distributed in SSA, including Uganda [41]. However, a recent systematic review found that, despite high uptake, many questions remain about HIVST's effectiveness in reaching historically underserved men and ensuring linkage to prevention and care [42].

Community-based HIV testing services (CHTS), such as mobile and home testing [4], show high potential for reaching men [4,38]. Men prefer CHTS over facility-based testing [26,38-40,43,44] due to its convenience and lower stigma [40,44-48]. Specifically, CHTS improves testing among low- and middle-income men, young men, and those who would otherwise avoid health facilities [40,44-48]. Still, broader reach and linkage to care mechanisms are lacking [38].

XSL•FO
RenderX

Peer intervention approaches, which use individuals with similar demographic and social backgrounds to the target population, have proven effective in promoting behavior change, sustaining reductions in HIV risk behaviors [49,50]. Peer-led education enhances knowledge exchange, reduces stigma, builds trust, and supports uptake of services such as HIVST [40,44,51-53]. Notably, community-based peer support has improved HIV service uptake and adherence in several settings [54,55].

Despite this progress, affordable and scalable combination programs tailored for heterosexual men in high-burden settings remain scarce [56]. Thus, innovative community-based strategies are urgently needed to improve men's HIV testing and linkage to prevention and care services. Barbershops offer informal, easily accessible (ie, convenient location and hours of operation), and collegial spaces where men often feel safe to talk about personal health issues, and the majority use the same barber most of the time [57]. Studies in the United States on barbershop-based health promotion interventions for Black men showed that barbershop-based interventions achieved satisfactory recruitment and retention of men in health services [58]. Another study in the United States found barbershop business venues appropriate for HIVST promotion and distribution [59].

Therefore, the purpose of HIV Prevention Trials Network (HPTN) 111, also known as TRIM (Testing a Barbershop-based HIV Prevention Initiative Among Men), was to develop and assess the feasibility, acceptability, and preliminary effectiveness of a barbershop-based HIV prevention initiative among heterosexual men in a high–HIV-prevalence setting. The study leverages trusted relationships between men and their barbers to deliver a prevention package consisting of HIV education, HIVST distribution, and barber-led peer group sessions. This protocol responds to a critical need to develop focused interventions to engage men in HIV prevention and care to bring male HIV prevention and care outcomes on par with those of women.

Objectives

The primary objective was to evaluate the feasibility and acceptability of a barbershop-based HIV prevention initiative. The secondary objectives were to compare completion of self-initiated HIV testing between intervention and control groups, evaluate the preliminary effectiveness of the intervention on change in behaviors associated with HIV acquisition, compare interest in or use of HIV prevention services between intervention and control groups, and assess interest in long-acting preexposure prophylaxis (PrEP) among all participants and by study arm. The exploratory objective was to evaluate the preliminary effectiveness of the intervention on incident sexually transmitted infections (STIs).

Methods

Overview

HPTN 111 (TRIM) was a pilot study conducted through the HPTN. HPTN 111 is funded by the National Institute of Allergy and Infectious Diseases (NIAID), under the National Institutes of Health. NIAID provides technical assistance and coordination of the study, including regulatory reviews and study conduct oversight. HPTN 111 is registered on ClinicalTrials.gov (NCT06148584). The study is based on the capability, opportunity, or motivational components of the Capability, Opportunity, and Motivation Model of Behavior Change (COM-B; Figure 1) [60]. The COM-B postulates that changing the behavior of an individual, a group, or a population requires changing either their capability, opportunity, and/or motivation needed to perform the behavior [61]. Therefore, in this study, we used the COM-B to provide a barbershop-based HIV prevention intervention to men living in a high heterosexual HIV incidence setting and who could benefit from HIV prevention services. The initiative aimed to increase men's capability, opportunity, and motivation through the provision of HIV education and prevention measures directly from their barbers and based at their barbershops. This study will provide valuable information about the feasibility and acceptability of a barbershop-based HIV prevention initiative and the potential for scale-up in other settings where barber-client relationships are an important component of men's social structure. Lastly, we followed the SPIRIT (Standardized Protocol Items: Recommendations for Interventional Trials) reporting guidelines during the drafting of this protocol paper.

Figure 1. The Capability, Opportunity, and Motivation Model of Behavior Change used to motivate and design the barbershop-based HIV prevention intervention in the HIV Prevention Trials Network 111 trial.

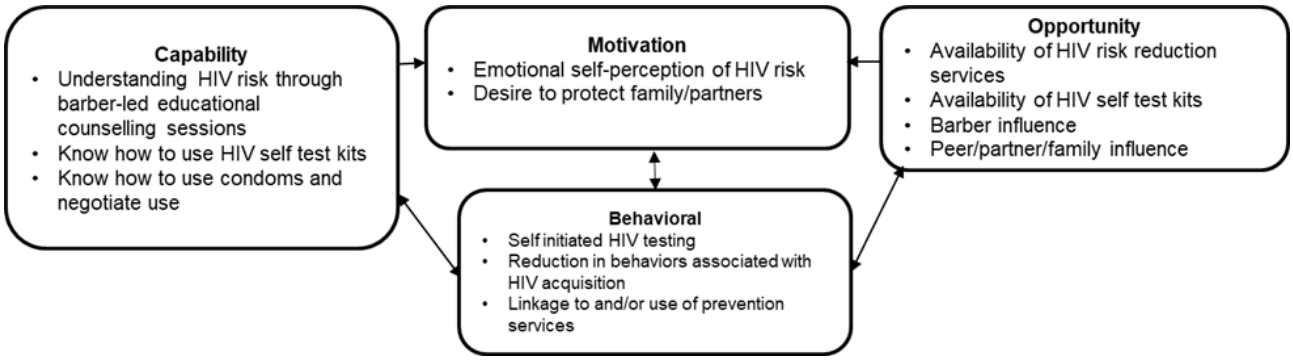

Study Design

HPTN 111 (TRIM) is a cluster randomized trial testing the feasibility and acceptability of a barbershop-based HIV prevention intervention among heterosexual men who have certain risk factors for HIV acquisition. The study was conducted in Kalangala district, Uganda, among barbershop-going men. Barbershops acted as clusters, and 18

XSL•FO
RenderX

purposively selected barbershops were randomized in a 2:1 ratio to the intervention or control group (Figure 2).

Figure 2. Overview schema of the HIV Prevention Trials Network 111 study design conducted in barbershops in the Kalangala Islands. HIVST: HIV self-testing; STI: sexually transmitted infection.

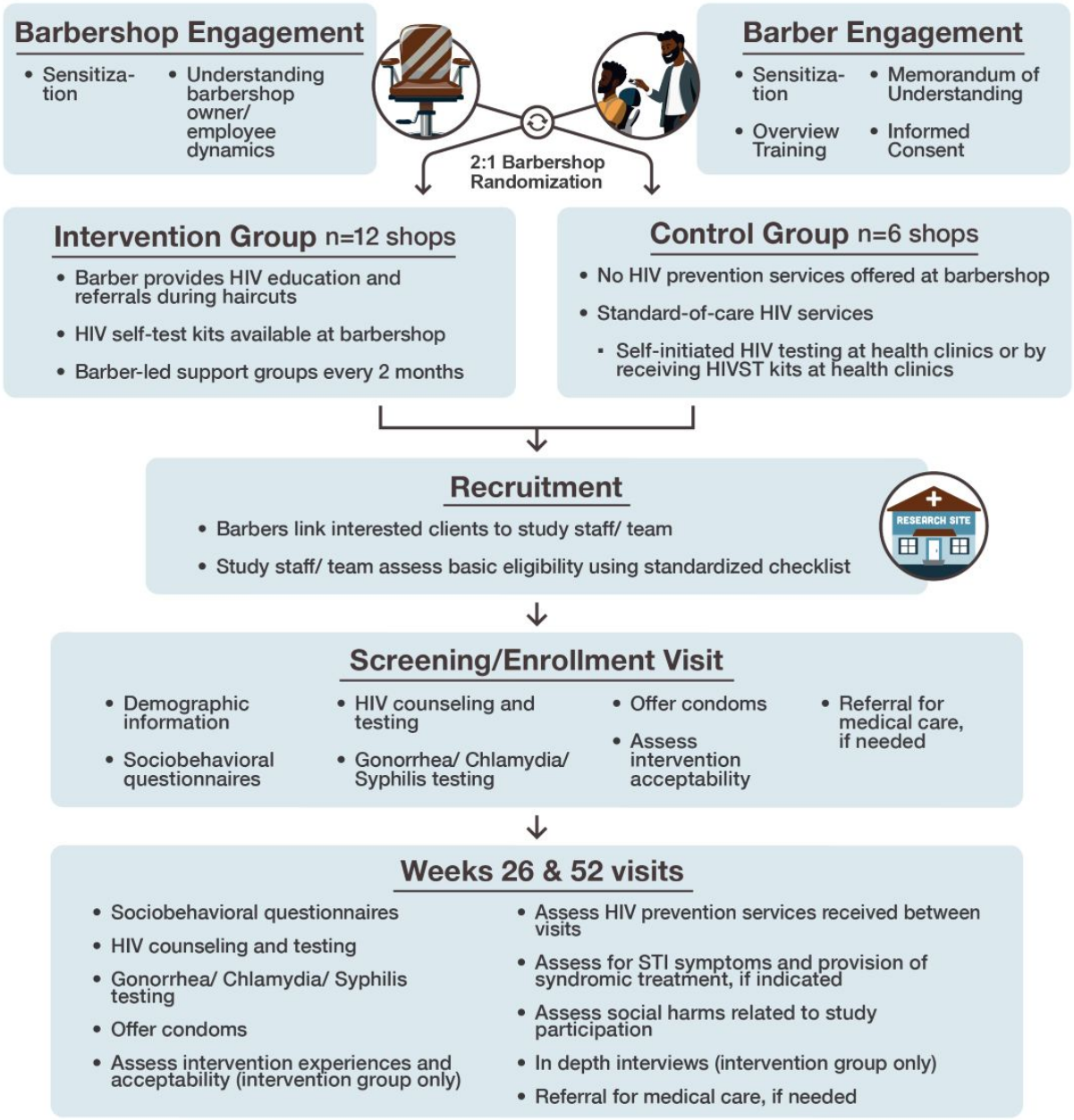

Men enrolled from the intervention barbershops received a barbershop-based HIV prevention intervention from their barber, while those enrolled in control barbershops received standard-of-care (SoC) HIV services. At intervention barbershops, trained barbers provided HIV status-neutral education, distributed HIVST kits at the barbershop, and led peer group sessions. The SoC services included existing HIV prevention services—namely, HIV testing and related services from nearby health facilities. All enrolled participants were followed for 12 months with biannual study visits at a designated study clinic. Intervention barbers also had quarterly follow-up visits to assess social impacts, acceptability, and feasibility of the intervention, and interviews to help better understand their experience within the intervention. Study objectives and corresponding end points are provided in Table 1.

XSL•FO
RenderX

Table 1. HIV Prevention Trials Network 111 study objectives and end points to be evaluated among 18 barbershops randomized 2:1 to implement the barbershop-based HIV prevention services package or receive standard of care.

Objective and topic	End point assessment	Population
Primary		
Evaluate the feasibility and acceptability of a barbershop-based HIV prevention initiative		
Participant survey responses about feasibility/acceptability	• Questions about how participants feel about the proposed intervention before receiving it • Questions about acceptability after receiving the intervention	• Both groups • Intervention group
Barber survey responses about feasibility/acceptability	• Questions about how barbers feel about the proposed intervention before delivering it • Questions about the feasibility/acceptability of talking to clients about the study • Questions about acceptability after delivering the intervention	• Both groups • Both groups • Intervention group
Recruitment and retention	• Participant recruitment in the study • Participant visit completion • Barber visit completion	• Both groups • Both groups • Intervention group
Barber-participant interactions and intervention delivery	• Participant-reported frequency of interactions with the barber and interventions received • Barber-reported frequency of intervention delivery and interventions delivered • Frequency and content of group sessions as observed by study staff	• Both groups • Intervention group • Intervention group
Secondary		
Compare the completion of self-initiated HIV testing between intervention and control groups		
Self-initiated HIV testing by participants	• Whether participants test for HIV between visits	• Both groups
Evaluate the preliminary effectiveness of the intervention on the change in behaviors associated with HIV acquisition		
Self-reported participant risk behaviors	• Participant survey responses about risk behaviors, such as condom use and the number of sexual partners	• Both groups
Compare interest in or use of HIV prevention services between intervention and control groups		
Self-reported participant use of HIV prevention services	• Uptake of VMMC[a] • PEP[b] access and use • PrEP[c] access and use	• Both groups • Both groups • Both groups
Assess interest in long-acting PrEP among all participants and by study arm		
Self-reported participant survey on long-acting PrEP	• Interest in the use of long-acting injectable PrEP	• Both groups
Exploratory		
Evaluate the preliminary effectiveness of the intervention on incident STIs		
Diagnostic testing results	• Incidence of HIV • Incidence of gonorrhea and chlamydia	• Both groups • Both groups

[a]VMMC: voluntary medical male circumcision.

[b]PEP: postexposure prophylaxis.

[c]PrEP: preexposure prophylaxis.

Sample Size

HPTN 111 (TRIM) is a pilot study with a primary objective of assessing feasibility and acceptability, measured in the intervention group only (12 barbershops). The sample size was determined based on the following two key considerations: (1) logistical and implementation feasibility [62] and (2) sufficient power to detect differences in a key secondary outcome of HIV testing uptake. A total of 18 barbershops were to be strategically selected, reflecting the number that could be feasibly supported by the study team and within the available budget, as guided in

XSL•FO
RenderX

feasibility studies [63]. Selection would be guided by the predefined eligibility criteria, with additional consideration for logistical efficiency. Specifically, barbershops should not be located excessively far from one another, yet should also avoid being in very close proximity. A minimum separation radius of approximately 5 km between sites would be considered to prevent contamination between intervention and control barbershops. While logistical feasibility was the primary driver, to inform client recruitment targets, we also assessed the number of participants per shop needed to detect meaningful differences in HIV testing between study arms, defined as a difference between arms of 20% in the proportion of participants self-initiating HIV testing [63]. A 2:1 randomization was applied, which both increased the number of intervention shops contributing data on feasibility and acceptability and allowed comparisons between arms on secondary outcomes. Enrollment targets of at least 10 participants per each of the 18 barbershops were based on the key considerations above, as well as maintaining at least 80% power for the HIV testing analyses under a variety of possible scenarios (Multimedia Appendix 1). Therefore, each shop would recruit at least 10 participants to reduce the variability in sample size between barbershops, and the protocol team would be informed that it is optimal for power if the 250 participants are evenly distributed across the 18 barbershops, although the variability in the barbershop and village sizes might prevent this.

Community Engagement, Barbershop Selection, and Barber Training

The study team obtained administrative approval and permission from local leaders to engage various communities and stakeholders in the study area (Kalangala district). This process involved barbershop assessment and selection, district and community stakeholder engagement, and barber training. The study team conducted stakeholder engagement meetings organized with the support of district and local leaders. Stakeholders included local and religious leaders, community representatives from fishing communities, boda-boda cyclists, palm plantation workers, nongovernmental organizations, and village health teams. Meetings were used to discuss perceptions, attitudes, and challenges surrounding the barbershop-based HIV prevention initiative and possible ongoing mitigation strategies.

Participating barbershops were purposively selected on Bugala Island (the largest island in Kalangala District) based on set criteria (Textbox 1) assessed by the study team. The barbershops were largely selected from busy trading centers/towns and landing sites. The barbers participating in each shop were also required to meet set criteria listed in Textbox 1.

Textbox 1. HIV Prevention Trials Network (HPTN) 111 eligibility criteria for the barbershops, barbers, and participants recruited in the Kalangala Islands.

Barbershop inclusion criteria

- Adequate space to deliver the intervention to clients
- Existence as a shop for at least 6 months
- A customer base of at least 15 clients per week
- A minimum radius of 5 km between other participating shops

Barber inclusion criteria

- At least a primary level of education
- Willing and able to participate in the study, recruit participants, and deliver the intervention if their shop is randomized to the intervention group

Participant

- Inclusion criteria
  - At least 16 years of age
  - Any participants 16-17 years of age must meet the definition of a mature or emancipated minor per the Uganda National Council for Science and Technology guidelines.
  - Identifies as a heterosexual male
  - Able and willing to provide informed consent
  - Has certain risk factors for HIV acquisition, based on self-report of at least one of the following in the last 3 months: (1) had condomless sex with a person of unknown HIV status or a person living with HIV; (2) had more than one sexual partner
  - HIV negative, per the Ugandan Ministry of Health guidelines and the HPTN 111 Study Specific Procedural Manual
  - A regular customer at a participating barbershop, defined as visiting a specific participating barbershop for barber services at least three times prior study participation
- Exclusion criteria
  - Not planning to stay in the study catchment area in the next 12 months
  - Any other condition that, in the opinion of the Investigator of Record, would preclude informed consent, make study participation unsafe, complicate interpretation of study outcome data, or otherwise interfere with achieving the study objectives

XSL•FO
RenderX

On completion of the barbershop assessment and selection, barbers from the 18 selected barbershops were further engaged and trained through consultative and participatory approaches. They were engaged to assess commitment and trained on the study overview, general HIV education, communication skills, and study recruitment requirements. The study team used didactic presentations, illustrations, group discussions, and role plays for these trainings.

Barbers from barbershops assigned to the intervention group received additional training on the implementation of the intervention. Pre- and posttraining assessments were used to ensure that barbers had sufficient knowledge about HIV and the intervention before participant recruitment. Training was ongoing during the study, with study staff providing support to barbers and assessing knowledge and confidence in participant recruitment and intervention delivery. Refresher training and cross-sharing with other barbers were completed if knowledge gaps were identified.

Randomization

The randomization procedures assigned the 18 selected barbershops to either the intervention or control arm in a 2:1 ratio. Due to the relatively small number of units being randomized (n=18), covariate-constrained randomization was used to balance 3 barbershop characteristics between treatment groups: the estimated population of the village or town where the barbershop is located, the average number of customers a barber served per day (as reported by the lead barber), and the estimated travel distance between the barbershop and the nearest health facility (in km). The restrictions would ensure that the difference between groups is within 300 for the average village population, within 4 for the average number of clients per day, and within 0.75 km for the average distance to the nearest health facility.

Randomization was done in 2 stages. In the first stage, the shops were allocated into 3 groups of 6 shops each using covariate-constraint randomization described above. In the second stage, a public randomization event was conducted, and 2 groups (12 shops total) were assigned to the treatment group, and 1 group (6 shops) was assigned to the control group. The first stage of randomization was executed by statisticians at Fred Hutchinson Cancer Center; the second stage involved barbers publicly drawing opaque balls from a transparent bucket to randomly determine treatment group assignment, ensuring transparency and fostering community engagement.

Due to the nature of the intervention, neither barbers nor study participants were blinded. Barbers were asked not to disclose their treatment assignment during recruitment activities; however, because of the public randomization process and visible delivery of the intervention to other clients, clients would become aware of their barbershop's treatment assignment prior to enrollment.

Ethical Considerations

The HPTN Ethics Working Group developed the HPTN Ethics Guidance for Research, a network-wide ethical principles document, which guided the HPTN 111 (TRIM) study conduct. The protocol and the informed consent forms (ICFs) were reviewed and approved by the HPTN Scientific Review Committee with respect to scientific content and compliance with applicable research and human participants' regulations.

The protocol, site-specific ICFs, participant education and recruitment materials, other requested documents, and any subsequent modifications were reviewed and approved by the Joint Clinical Research Center (JCRC 2023-54) and the Johns Hopkins Medicine (IRB00420132) and cleared by the Uganda National Council for Science and Technology (HS3430ES), which are responsible for oversight of research conducted at the study site.

After initial review and approval, the responsible institutional review boards (IRBs)/ethics committees will review the protocol at least annually. The investigator will make safety and progress reports to the IRBs/ethics committees at least annually and provide a study termination or completion report within 3 months after the study is completed. The annual reports include the total number of participants enrolled in the study, the number of participants who completed the study, all changes in the research activity, and all unanticipated problems involving risks to participants or others. The study site submits documentation of continuing review to the Division of AIDS (DAIDS) Protocol Registration Office, in accordance with the current DAIDS Protocol Registration Policy and Procedures Manual.

All participants provided informed consent prior to their involvement in any study procedure (Multimedia Appendix 2). We ensured participant privacy and confidentiality by using anonymized and deidentified data. Participant data, including ICFs, are kept under lock and key and only accessed by the research team or authorized personnel. Participants were reimbursed and compensated for the time spent at the study site at the rate approved by the IRB.

Study Procedures

Recruitment, Screening, and Enrollment

Participating barbers were responsible for recruiting their male clients into the study. Barbers were provided with posters and information brochures to help with recruitment. All barbers were asked to complete a simple, anonymous log of the number of men contacted and linked to the study for screening. Men who were interested in the study following barber-led recruitment were referred directly to study staff who would complete a prescreening assessment. During the prescreening, potential participants were assessed for sexual behaviors associated with HIV acquisition, their residence, barbershop attendance status, and age. Potential participants who remained eligible following this prescreening assessment would proceed to the central study clinic for further eligibility assessment and enrollment. Study procedures are shown in Figure 2.

At the study clinic, potential participants completed a full screening visit, including providing written informed consent, collection of basic demographic and sociobehavioral information, and blood collection for HIV and syphilis testing. Participants who remained eligible after screening procedures would immediately proceed to enrollment or return for enrollment within 14 days. Eligibility criteria for participants

XSL•FO
RenderX

are provided in Textbox 1. Ineligible individuals would be referred to appropriate medical care, including HIV prevention services or HIV care, based on their needs. Participants could be rescreened up to 2 times if found to be ineligible at the initial screening attempt. Participants who would have a reactive HIV test at screening would not be able to rescreen.

Individuals were considered enrolled in the study once study staff verified that all eligibility criteria were met. Subsequent enrollment procedures included gonorrhea and chlamydia testing, administration of additional sociobehavioral questionnaires via audio computer–assisted self-interview (ACASI), assessment of the acceptability of the barbershop-based intervention, and counseling on the recommended quarterly HIV testing schedule. This recommendation aligned with the Uganda Ministry of Health National Implementation for HIV Counselling and Testing Services, which advises that individuals with certain HIV acquisition risk factors undergo HIV testing every 3 months [64]. Participants were encouraged to contact the study team between visits to share results of any HIV tests they would have completed.

Enrolled participants were informed of the study group to which their barbershop was randomized, so that they were aware of whether they would receive the intervention from their barber. For participants from an intervention shop, the respective barber was notified to ensure that the participant received the barbershop-based intervention during their next regular visit to the barbershop.

Follow-Up Study Visits

Following enrollment, follow-up study visits occurred at 26 weeks and 52 weeks post enrollment. HIV testing and counseling, as well as gonorrhea and chlamydia testing, were conducted at both visits. A symptom-directed physical exam was completed, and any STI symptoms were treated as per the SoC syndromic management. Participants were asked to provide information on their HIV testing behaviors between visits, use of other HIV prevention services (eg, HIV PrEP, postexposure prophylaxis, voluntary medical male circumcision), sexual behavior (via ACASI), and social harms experienced due to study participation. All participants were asked about their experiences of receiving barbershop-based HIV information to assess fidelity in the intervention group and cross-contamination in the control shops.

Interim study visits were completed upon participant request or as deemed necessary by the study team. For example, interim visits would occur if a participant contacted study staff with HIV test results, either requiring follow-up testing or referral to care, or if they received a positive STI test result and required treatment.

Participants either received the barbershop-based HIV prevention initiative (intervention group) or SoC HIV prevention services (control group) based on the barbershop from which they were recruited. Participants were expected to continue visiting their respective barbershops for regular services, but these barbershop visits were not considered formal study visits.

Qualitative Interviews

Participants in the intervention group were further asked about the acceptability of receiving the intervention from their barber. A subset of approximately 25 participants from the intervention group were purposively selected to participate in individual in-depth interviews at week 26 and/or week 52. Additional interviews were conducted if saturation was not reached—that is, the point at which no new themes or insights emerge, ensuring a comprehensive understanding of participants' experience within the study—barbers and clients alike. The interviews explored the acceptability of the barbershop-based HIV prevention initiative and use of HIV prevention strategies to gain deeper insight into how the intervention may influence behavioral change (Multimedia Appendix 3).

Sample size was guided by the principle of informational adequacy and thematic saturation [65]. Studies examining relatively homogeneous populations with narrowly defined research objectives have demonstrated that saturation of major themes often occurs within 20-30 interviews [66]. The study population in this protocol, heterosexual men attending selected barbershops in the Kalangala Islands, represents a geographically and socially bounded group exposed to a standardized intervention.

Purposive sampling further strengthens informational yield. Participants were selected based on characteristics such as age, HIV risk behaviors, use of prevention services, experiences during the study, and barbershop attended. This approach ensures variation within key implementation domains while maintaining contextual coherence. A sample of approximately 25 interviews is therefore methodologically appropriate to generate an in-depth understanding of acceptability, behavioral responses, and contextual influences within a feasibility trial framework.

Intervention Group

Participants in the intervention group received the HIV prevention intervention from their barber. The intervention includes (1) status-neutral HIV education and referral information for HIV services, (2) distribution of HIVST kits through the barbershops, and (3) barber-led peer group sessions.

Status-Neutral HIV Education and Referral Information for HIV Services

During the routine barber-participant interaction, barbers would provide general, status-neutral HIV education while providing haircut services. A counseling guide was provided to ensure that barbers follow standardized messaging during participant education. If other clients were present while an enrolled participant was receiving a haircut and study-related services, the barbers were trained to extend the education to everyone in the shop at that time. The provided HIV education and counseling was neutral and general information as guided by the Ugandan Ministry of Health manuals for HIV education in public spaces and, therefore, approved by the IRB to share in spaces with nonparticipants. Participants were educated on the risks of HIV acquisition, strategies to reduce those risks, the recommended HIV testing schedule, and where to find other HIV prevention services. Barbers would also share knowledge

XSL•FO
RenderX

and information about existing HIV prevention services and resources in Kalangala district, including comprehensive HIV education and counseling, PrEP, postexposure prophylaxis, and voluntary medical male circumcision services.

Distribution of HIVST Kits Through the Barbershops

Barbers maintained a stock of study-provided HIVST kits in the shop and offered them to participants to take home and use at their convenience. Participants would take as many test kits as needed, including for use with their partners. Barbers would educate participants on how to properly complete the test and inform them of the recommended testing frequency (ie, every 3 months for individuals with certain risk factors for HIV). HIVST kits were packed in opaque bags to allow participants to carry them discreetly out of the shop. If a nonstudy client was present during HIV education and expressed interest in taking an HIVST kit, they would be permitted to receive one.

Barber-Led Peer-Group Sessions

Each barber led his study clients in a peer group session every 2 months to provide additional HIV education and facilitate discussion. These sessions would foster peer support and the sharing of experiences among participants. Study staff assisted with coordinating the peer groups, recording attendance, documenting the HIV education topics covered, and assessing the quality and accuracy of the information delivered.

The barbers documented the provision of the intervention to their enrolled clients using a study intervention tracking document (Figure 3). They were trained on best practices for documentation, and study staff would provide real-time refresher training as needed.

Figure 3. Template for tracking barber delivery of the intervention to their male clients, used by barbers in the intervention arm during participating clients' haircuts.

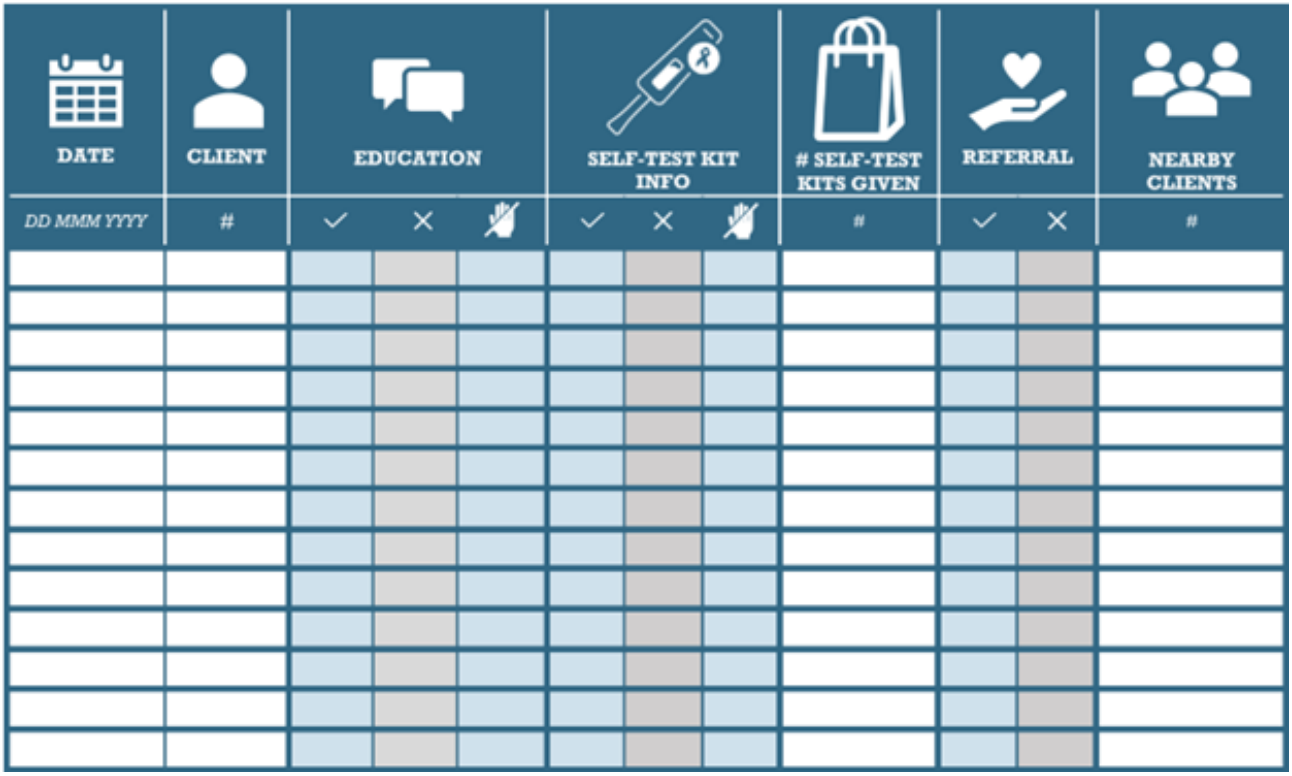

Study staff conducted approximately weekly check-ins with barbershops to review processes, troubleshoot issues, address concerns, and encourage barbers to deliver the intervention to study clients during routine barbering services (see Multimedia Appendix 4 for participant schedule of events).

Control Group

Participants in the control group would not receive any HIV prevention services from their barbershop or barber. During study visits at the clinic, they received counseling on routine HIV services, including HIV testing, HIVST kits, and other HIV prevention services available at nearby health facilities. Barbers in the control group were not notified about participants enrolled from their shops and were withdrawn from the study upon completion of the enrollment phase (see Multimedia Appendix 4 for participant schedule of events).

Retention

The study aimed to achieve participant retention of 90% at week 26 and week 52 visits. Strategies such as the collection of full locator information for participants, flexible study clinic hours, and the option for off-site visits would be used. However, to prevent undue influence in HIV prevention-seeking behaviors, the study team limited contact with participants between study visits.

Barber Study Procedures

Barbers in this study had a unique role as recruiters, implementers of the intervention, and study participants. Therefore, barbers participating in HPTN 111 (TRIM) carried out a series of administrative and study-specific activities. Barbers/barbershop owners and the study team signed a

XSL•FO
RenderX

memorandum of understanding to clearly outline expectations of both parties before the commencement of study activities. All participating barbers also completed written informed consent to ensure they fully understood and voluntarily agreed to participate in the research procedures required throughout the study.

Initially, study staff collected basic information about each participating shop, such as location, total number of clients, hours of operation, cost of services, etc. At baseline, barbers provided demographic information and completed a questionnaire on the acceptability and feasibility of the intervention. At the end of the recruitment period, barbers were asked to complete brief questionnaires about their recruitment experiences. Barbers in the intervention group remained in the study and completed brief questionnaires on acceptability, feasibility, and social impacts of delivering the intervention approximately quarterly during the study follow-up. Intervention barbers also participated in individual in-depth interviews at approximately 6 and 12 months after starting intervention delivery to further understand their views on feasibility, acceptability, facilitators and barriers, and overall experiences of delivering the intervention (see Multimedia Appendix 5 for barber-specific schedule of events and Multimedia Appendix 6 for barber in-depth interview guide). Intervention barbers completed a simple checklist (Figure 3) of the services provided to each client enrolled in the study each time the participant visited the barbershop. The checklist did not contain any direct identifiers of the clients and was stored in a lockable cabinet at the shop, to be collected weekly by study staff.

Data Management

HPTN 111 (TRIM) data were managed in REDCap (Research Electronic Data Capture; Vanderbilt University) Cloud, a platform used by the HPTN Statistical and Data Management Center to receive and manage study data. The study site completed electronic case report forms (CRFs) by entering data into the REDCap Cloud study database. Data would be entered directly into the study database, collected first on paper CRFs and then entered into the study database, entered directly by study participants (ACASI responses only), and/or entered into the study database based on other non-CRF source documents (eg, laboratory reports, testing logs, and chart notes).

Data Analysis

The primary objective of HPTN 111 (TRIM) is to assess the feasibility and acceptability of a barbershop-based HIV prevention intervention. To achieve this, a variety of end points will be assessed (Table 1).

To assess feasibility and acceptability of the intervention, the following were assessed from participants in the intervention group: participant responses from surveys at week 26 and week 52 regarding acceptability of the intervention, barber responses from surveys assessing feasibility and acceptability of delivering the intervention, recruitment rates during enrollment and participant retention in the study at week 26 and week 52, and the frequency of barber-participant interactions (Multimedia Appendix 7). These outcomes were prespecified and will be summarized and described, but specific progression criteria or targets for these outcomes were not prespecified, which is a limitation driven by the novelty of this intervention. The study also assessed the effectiveness of the intervention on changes in HIV testing and use of prevention services. Self-reported sexual behaviors associated with HIV incidence and STI incidence rates will also be compared to SoC.

For end points involving comparisons between groups, as well as those assessed only within the intervention group, analyses will follow intention-to-treat principles. Formal comparisons will account for the clustering by aggregating data at the barbershop level and adjusting for the covariates constrained by randomization [67]. No interim analyses are planned. Secondary end points (Table 1) will be assessed using the same statistical methods. Because feasibility and acceptability outcomes overlap with critical information about study conduct, certain data, such as recruitment and enrollment summaries and information on the frequency and quality of intervention delivery by the barber—would be shared with investigators and study staff. Summaries of other end points not related to study conduct (eg, self-initiated HIV testing) would remain unavailable to investigators and study staff for the duration of the trial. Data and safety monitoring procedures for the trial will involve summaries of reported outcomes and social impacts. Reporting of end points not related to study conduct would be available only to members of the study monitoring committee (SMC) for this study and will be withheld from study investigators and staff members.

Qualitative interviews were digitally recorded, transcribed, translated where necessary, and analyzed systematically. A hybrid thematic analysis approach, combining deductive and inductive coding strategies [68], was used to analyze the data. The intervention itself is guided by COM-B [69]. Therefore, our analytic framework incorporated both deductive coding informed by COM-B constructs and inductive coding.

Emergent themes not predetermined by the framework, such as trust in barbers, social cohesion within fishing communities, mobility between islands, or unintended social impacts, were identified through open coding. Thematic analysis proceeded through the following stages [70]: (1) review of interview debrief reports, (2) familiarization with transcripts, (3) generation of initial codes, (4) development and refinement of a codebook, (5) double coding of selected transcripts to enhance analytic rigor and for determination of interrater reliability among the coders, (6) identification and refinement of themes, and (7) interpretation of themes in relation to feasibility, acceptability, and preliminary effectiveness.

This hybrid approach allowed theoretical alignment with the COM-B while ensuring contextual sensitivity to the lived experiences of participants in Kalangala. Such integration is consistent with implementation science methodologies that combine theory-driven and data-driven analysis [71].

Determination of saturation included descriptive saturation, interpretive saturation, and team-based review. Descriptive saturation was assessed based on no new themes emerging from successive interviews and would be implemented by team-based reviews of interview debrief reports, which were discussed at weekly team meetings. Similarly, at the weekly team-based

XSL•FO
RenderX

meetings, which included the protocol team and the interviewers, interpretive (meaning) saturation was assessed. Meaning saturation occurs when themes are fully elaborated, and additional interviews do not provide new depth, nuance, or variation [72]. The team used the interview debrief reports to document thematic development, review variation across purposively selected subgroups (eg, age groups, different barbershops, and levels of HIV prevention uptake), and determine whether thematic explanations sufficiently captured contextual complexity. This guided the team in determining that saturation had been reached.

Data Safety and Monitoring

An SMC conducted interim reviews of study progress, including participant accrual, retention, social harms, and quality of data collection. Reviews occurred approximately every 6 months. The SMC provided recommendations on study conduct to ensure that the study can meet the primary objectives.

Due to the minimal-risk nature of the study and absence of a biomedical intervention, there was no centralized reporting of adverse events. However, the study site would report all adverse events and social harms according to local IRB and regulatory requirements.

Results

Data collection began in March 2024 and concluded in June 2025. Participants were followed for 12 months. Data analysis has been completed. The primary manuscript is expected to be submitted for publication by March 30, 2026.

Discussion

This protocol is designed to evaluate the feasibility and acceptability of a barbershop-based HIV prevention intervention and test its preliminary effect on male engagement in HIV prevention services. The intervention was piloted among men in Kalangala district, Uganda—a population that may benefit significantly from expanded HIV prevention efforts. Barbershops, which serve as popular gathering places for men in many settings, especially in African communities, provide a promising platform for this intervention. The intervention involved 3 main components delivered by the barbers: HIV education and referral information for HIV services, distribution of HIV self-test kits through barbershops, and barber-led peer group sessions. Given the documented barriers men face while engaging in HIV services [26,73]—such as stigma around testing, services that are often tailored to women, inconvenient service locations and hours [29,30,33-37,74]—this protocol offers an opportunity to explore a novel, community-based strategy to address these gaps. Ultimately, this approach aims to reach a historically underserved population within the HIV prevention and care continuum.

Barbershops may be convenient and ideal venues for health promotion [75]. Earlier studies showed that health initiatives based in barbershops improved condom use [51] and high blood pressure control [76,77]. More recently, a trustworthy relationship between barbers and their clients has been reported as a facilitator for collaboration with barbershop businesses to reach Black men for HIVST distribution [59]. Similarly, salon-based PrEP initiatives also improved PrEP knowledge and uptake among Black women in the United States [78]. These successes suggest that salon or barbershop-based interventions may be transferable across settings, informing the development of similar HIV prevention initiatives targeting barbershop-going men in Uganda and other settings. HPTN 111 was implemented in a rural fishing community, where the lessons learned may be transferable to similar rural contexts, although they may not be generalizable to urban and nonfishing communities. The study area was selected due to the high HIV burden in the Ugandan fishing communities [73] and the demographic reality that more men than women live in these areas. This imbalance has contributed to high rates of transactional sex, with the majority of men being unaware of their HIV status [79-82].

Peer support in this study included barber-led education and barber-led group sessions. Peer support draws on shared personal experience to provide knowledge, social interaction, emotional assistance, or practical help, often in a mutually beneficial way, where the source of support is a peer with relevant lived experience [83-85]. In this study, the barbers provided education, actively listened, built rapport, and offered emotional and social support to promote behavioral change, consistent with recommended peer support practices [51-53]. The intervention used status-neutral education during barber-client interactions, a strategy shown to be both feasible and effective in community-based interventions [86,87]. The barber-delivered peer support was provided face-to-face during routine haircuts, in line with evidence that peer support is most effective when delivered regularly in familiar and comfortable settings [88,89]. Barbers also facilitated group sessions, allowing men to discuss HIV prevention practices among peers. We anticipated that these regular interactions would improve HIV prevention knowledge and enhance men's engagement in HIV prevention practices.

Male-led, community-based interventions can help overcome stigma and masculinity-related barriers that typically deter men from using HIV services [40,44]. Moreover, during and after adolescence, men often rely on peer networks for support, guidance, and advice in the development of a sense of belonging and identity [57,90]. This study leverages this notion since most barbers in this setting are men. Further, qualitative evidence suggests that peer counseling fosters trust and confidence toward HIV testing, including HIVST [51-53].

HIVST is a robust tool that could benefit historically hard-to-reach populations in need of HIV prevention and care services. However, concerns remain about whether individuals who use HIVST are adequately linked to follow-up care. HPTN 111 seeks to address this gap via barbershops to expand access to HIVST among men who may face barriers in obtaining it from health facilities and improve their linkage to HIV prevention services and care.

Data on the feasibility and acceptability of the intervention, collected through quantitative surveys and qualitative interviews, will provide critical insights into barriers and facilitators to implementation. Systematic evaluation and analysis of these data will inform refinement of the intervention package and

XSL•FO
RenderX

guide the development of strategies for broader scale-up. In addition, findings from both predefined quantitative objectives and exploratory analyses will help assess the potential effectiveness of the intervention and identify areas for improvement. Collectively, these data will generate practical insights on the provision of barbershop-based HIV services in high HIV-burden fishing communities and similar settings.

Subsequently, if successful, this intervention will contribute to the community-based decentralization of HIV services by enhancing men's awareness of HIV and risk-reduction strategies, increasing access to and use of HIV testing services, reducing HIV-related stigma associated with testing, and strengthening linkage to HIV prevention and care services. Overall, the intervention will improve male engagement in HIV prevention and care, reduce HIV infections among men, and ultimately decrease HIV transmission to adolescent girls and young women.

This protocol is limited by the absence of prespecified progression criteria, as the novelty of the intervention meant there was no sufficient prior data to serve as a benchmark. Nevertheless, the protocol clearly outlines how feasibility and acceptability will be assessed and how the results will inform subsequent steps, in line with guidance from the literature [41,91].

In conclusion, the HPTN 111 (TRIM) study will answer the question of whether the barbershop-based HIV prevention initiative is feasible, acceptable, and provide insight into preliminary effectiveness in improving HIV testing, promoting behavioral change, and minimizing HIV and STI incidence. This study highlights the potential of barbershops as accessible, trusted, and far-reaching venues for delivering HIV prevention services to men. This barber-led intervention is anticipated to effectively engage and sustain heterosexual men in HIV prevention and care. If found feasible and acceptable, the data obtained will inform modifications and guide the implementation of the intervention in a future larger-scale effectiveness trial.

Acknowledgments

We would like to thank the HIV Prevention Trials Network (HPTN) 111 study team for their tireless efforts in developing this protocol, as well as the HPTN Leadership and Operations Center for administrative and technical support, and overall guidance on this protocol.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the pending publication of the primary manuscript. Protocol and letter of amendment files are available at the HIV Prevention Trials Network [92].

Funding

Research reported in this publication was supported by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health under award number UM1AI068619. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The Overall support for the HIV Prevention Trials Network (HPTN) is provided by the NIAID, Office of the Director, National Institutes of Health, National Institute on Drug Abuse, the National Institute of Mental Health, and the Eunice Kennedy Shriver National Institute of Child Health and Human Development under award numbers UM1AI068619 (HPTN Leadership and Operations Center), UM1AI068617 (HPTN Statistical and Data Management Center), and UM1AI068613 (HPTN Laboratory Center).

Authors' Contributions

Conceptualization: ZL, BGM
Data curation: EV, DD
Formal Analysis: EV, DD
Investigation: ZL, BGM, JM
Methodology: ZL, BGM, CN, RN, EP-M, DD, JM
Project Administration: CS, JM, JH-W, MA, MD, H Ssemere, IR, TN, DK
Supervision: CN, DD
Validation: EV, DD
Visualization: EV
Writing — original draft: CS, ZL, BGM, EV
Writing — review & editing: JH-W, CS, AM, MD, DFC, MA, RN, CN, H Spiegel

Conflicts of Interest

None declared.

Multimedia Appendix 1
HIV Prevention Trials Network 111 protocol V2.0.

XSL•FO
RenderX

[DOCX File , 439 KB - resprot_v15i1e87612_app1.docx]

Multimedia Appendix 2
Sample participant and barber informed consent forms.
[DOCX File , 52 KB - resprot_v15i1e87612_app2.docx]

Multimedia Appendix 3
HIV Prevention Trials Network 111 (Testing a Barbershop-based HIV Prevention Initiative Among Men) participant in-depth interview guide.
[PDF File (Adobe PDF File), 290 KB - resprot_v15i1e87612_app3.pdf]

Multimedia Appendix 4
Participant schedule of events.
[DOCX File , 27 KB - resprot_v15i1e87612_app4.docx]

Multimedia Appendix 5
Barber schedule of events.
[DOCX File , 23 KB - resprot_v15i1e87612_app5.docx]

Multimedia Appendix 6
HIV Prevention Trials Network 111 (Testing a Barbershop-based HIV Prevention Initiative Among Men) barber in-depth interview guide.
[PDF File (Adobe PDF File), 279 KB - resprot_v15i1e87612_app6.pdf]

Multimedia Appendix 7
HIV Prevention Trials Network 111 feasibility and acceptability scoring details.
[PDF File (Adobe PDF File), 151 KB - resprot_v15i1e87612_app7.pdf]

Multimedia Appendix 8
Peer review report by the HIV Prevention Trials Network (HPTN) Science Review Committee (SRC), National Institute of Allergy and Infectious Diseases (NIAID) (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 685 KB - resprot_v15i1e87612_app8.pdf]

References

1. Global AIDS update 2021: confronting inequalities; lessons for pandemic responses from 40 years of AIDS," UNAIDS Joint United Nations Programme on HIV/AIDS, 20 Avenue Appia 1211 Geneva 27 Switzerland. UNAIDs. 2021. URL: https://www.unaids.org/sites/default/files/media_asset/2021-global-aids-update_en.pdf [accessed 2026-03-11]
2. Cornell M, Majola M, Johnson LF, Dubula-Majola V. HIV services in sub-Saharan Africa: the greatest gap is men. Lancet 2021;397(10290):2130-2132. [doi: 10.1016/s0140-6736(21)01163-6]
3. Grimsrud A, Ameyan W, Ayieko J, Shewchuk T. Shifting the narrative: from "the missing men" to "we are missing the men". J Int AIDS Soc 2020;23 Suppl 2(Suppl 2):e25526 [FREE Full text] [doi: 10.1002/jia2.25526] [Medline: 32589325]
4. UNAIDs. Male engagement in HIV testing, treatment and prevention in Eastern and Southern Africa: a framework for action. 2023. URL: https://reliefweb.int/report/world/male-engagement-hiv-testing-treatment-and-prevention-eastern-and-southern-africa [accessed 2023-03-15]
5. UNAIDS. Joint United Nations Programme on HIV/AIDS: global AIDS updates confronting inequalities," in "lessons for pandemic responses from 40 years of AIDS. 2021. URL: https://reliefweb.int/report/world/2021-unaids-global-aids-update-confronting-inequalities-lessons-pandemic-responses-40 [accessed 2026-03-11]
6. The urgency of now: AIDS at a crossroads. UNAIDs. 2024. URL: https://www.unaids.org/sites/default/files/media_asset/2024-unaids-global-aids-update_en.pdf [accessed 2026-03-11]
7. Nardell MF, Adeoti O, Peters C, Kakuhikire B, Govathson-Mandimika C, Long L, et al. Men missing from the HIV care continuum in sub-Saharan Africa: a meta-analysis and meta-synthesis. J Int AIDS Soc 2022;25(3):e25889 [FREE Full text] [doi: 10.1002/jia2.25889] [Medline: 35324089]
8. Boily MC, Baggaley RF, Wang L, Masse B, White RG, Hayes RJ, et al. Heterosexual risk of HIV-1 infection per sexual act: systematic review and meta-analysis of observational studies. Lancet Infect Dis 2009;9(2):118-129. [doi: 10.1016/s1473-3099(09)70021-0]

XSL•FO
RenderX

9. de Oliveira T, Kharsany ABM, Gräf T, Cawood C, Khanyile D, Grobler A, et al. Transmission networks and risk of HIV infection in Kwazulu-Natal, South Africa: a community-wide phylogenetic study. Lancet HIV 2017;4(1):e41-e50. [doi: 10.1016/s2352-3018(16)30186-2]
10. Winskell K, Singleton R, Sabben G, Tiendrébéogo G, Obong'o C, Dia FL, et al. Social representations of the prevention of heterosexual transmission of HIV among young Africans from five countries, 1997-2014. PLoS One 2020;15(3):e0227878 [FREE Full text] [doi: 10.1371/journal.pone.0227878] [Medline: 32142515]
11. Jia KM, Eilerts H, Edun O, Lam K, Howes A, Thomas ML, et al. Risk scores for predicting HIV incidence among adult heterosexual populations in sub-Saharan Africa: a systematic review and meta-analysis. J Int AIDS Soc 2022;25(1):e25861 [FREE Full text] [doi: 10.1002/jia2.25861] [Medline: 35001515]
12. Babel RA, Wang P, Alessi EJ, Raymond HF, Wei C. Stigma, HIV risk, and access to HIV prevention and treatment services among men who have sex with men (MSM) in the United States: a scoping review. AIDS Behav 2021;25(11):3574-3604. [doi: 10.1007/s10461-021-03262-4]
13. Dunbar W, Labat A, Raccurt C, Sohler N, Pape JW, Maulet N, et al. A realist systematic review of stigma reduction interventions for HIV prevention and care continuum outcomes among men who have sex with men. Int J STD AIDS 2020;31(8):712-723. [doi: 10.1177/0956462420924984] [Medline: 32631213]
14. Phillips G, McCuskey D, Ruprecht MM, Curry CW, Felt D. Structural interventions for HIV prevention and care among US men who have sex with men: a systematic review of evidence, gaps, and future priorities. AIDS Behav 2021;25(9):2907-2919 [FREE Full text] [doi: 10.1007/s10461-021-03167-2] [Medline: 33534056]
15. Maulsby C, Millett G, Lindsey K, Kelley R, Johnson K, Montoya D, et al. A systematic review of HIV interventions for black men who have sex with men (MSM). BMC Public Health 2013;13(1):625 [FREE Full text] [doi: 10.1186/1471-2458-13-625] [Medline: 23819660]
16. Pascoe L, Peacock D, Stemple L. Reaching men: addressing the blind spot in the HIV response. Int J Men Soc Commun Health 2018;1(SP1):e57-e70. [doi: 10.22374/ijmsch.v1isp1.3]
17. Kusemererwa S, Akena D, Nakanjako D, Kigozi J, Nanyunja R, Nanfuka M, et al. Strategies for retention of heterosexual men in HIV care in sub-Saharan Africa: a systematic review. PLoS One 2021;16(2):e0246471 [FREE Full text] [doi: 10.1371/journal.pone.0246471] [Medline: 33539424]
18. Male engagement in HIV testing, treatment and prevention in Eastern and Southern Africa: a framework for action. UNAIDs. 2022. URL: https://www.unaids.org/sites/default/files/media_asset/east-south-africa-engaging-men_en.pdf [accessed 2026-03-11]
19. 2023 fact sheet? Facts on HIV and AIDS in Uganda 2023 (based on data ending 31st December 2022). Uganda AIDS Commision. 2023. URL: https://uac.go.ug/media/attachments/2024/01/23/hiv-aids-factsheet-2023.pdf [accessed 2026-03-11]
20. Facts on HIV and AIDS in Uganda 2021, Kampala, Uganda. UA Commission. 2021. URL: https://uac.go.ug/media/attachments/2021/09/13/final-2021-hiv-aids-factsheet.pdf [accessed 2026-03-11]
21. Kalangala district health management information system. USEHMI System. URL: https://hmis2.health.go.ug/hmis2/dhis-web-commons/security/login.action [accessed 2026-03-11]
22. Investing in sexual and reproductive health is a pathway to better health in Kalangala district. Population Reference Bureau, Uganda. 2018. URL: https://www.prb.org/resource/fact-sheet-investing-in-sexual-and-reproductive-health-is-a-pathway-to-better-health-in-kalangala-district/ [accessed 2026-03-11]
23. HIV/AIDS related knowledge, attitudes, and behaviour. Uganda Demographic and Health Survey. 2016. URL: https://www.dhsprogram.com/pubs/pdf/SR194/SR194.pdf [accessed 2026-03-11]
24. Facts on HIV and AIDs in Uganda in 2022 Fact sheets,. Uganda AIDS Commision. 2022. URL: https://uac.go.ug/media/attachments/2023/01/12/hiv-aids-factsheet-2022.pdf [accessed 2026-03-11]
25. UNAIDS. UNAIDS Global AIDS Update 2022: in danger. In: Joint United Nations Programme on HIV/AIDS. Geneva: https://www.unaids.org/sites/default/files/media_asset/2022-global-aids-update_en.pdf; 2023.
26. Sharma M, Barnabas RV, Celum C. Community-based strategies to strengthen men's engagement in the HIV care cascade in sub-Saharan Africa. PLoS Med 2017;14(4):e1002262 [FREE Full text] [doi: 10.1371/journal.pmed.1002262] [Medline: 28399122]
27. Colvin CJ. Strategies for engaging men in HIV services. The Lancet HIV 2019;6(3):e191-e200. [doi: 10.1016/s2352-3018(19)30032-3]
28. Dovel K, Dworkin SL, Cornell M, Coates TJ, Yeatman S. Gendered health institutions: examining the organization of health services and men's use of HIV testing in Malawi. J Int AIDS Soc 2020;23(2):e25517 [FREE Full text] [doi: 10.1002/jia2.25517] [Medline: 32589346]
29. Katz DA, Kiarie JN, John-Stewart GC, Richardson BA, John FN, Farquhar C. Male perspectives on incorporating men into antenatal HIV counseling and testing. PLoS One 2009;4(11):e7602 [FREE Full text] [doi: 10.1371/journal.pone.0007602] [Medline: 19881884]
30. Larsson EC, Thorson A, Nsabagasani X, Namusoko S, Popenoe R, Ekström AM. Mistrust in marriage-reasons why men do not accept couple HIV testing during antenatal care- a qualitative study in eastern Uganda. BMC Public Health 2010;10(1):769 [FREE Full text] [doi: 10.1186/1471-2458-10-769] [Medline: 21167040]

XSL•FO
RenderX

31. Betron M, Barker G, Manuel JC, Peacock D. Men, masculinities and HIV/AIDS: strategies for action. Equimundo. URL: https://www.equimundo.org/wp-content/uploads/2014/12/Men-Masculinities-and-HIV-AIDS-Strategies-for-Action.pdf [accessed 2026-03-27]
32. Dworkin SL, Treves-Kagan S, Lippman SA. Gender-transformative interventions to reduce HIV risks and violence with heterosexually-active men: a review of the global evidence. AIDS Behav 2013;17(9):2845-2863. [doi: 10.1007/s10461-013-0565-2] [Medline: 23934267]
33. Brou H, Djohan G, Becquet R, Allou G, Ekouevi DK, Viho I, et al. When do HIV-infected women disclose their HIV status to their male partner and why? A study in a PMTCT programme, Abidjan. PLoS Med 2007;4(12):e342 [FREE Full text] [doi: 10.1371/journal.pmed.0040342] [Medline: 18052603]
34. Byamugisha R, Tumwine JK, Semiyaga N, Tylleskar T. Determinants of male involvement in the prevention of mother-to-child transmission of HIV programme in Eastern Uganda: a cross-sectional survey. Reprod Health 2010;7(1):12 [FREE Full text] [doi: 10.1186/1742-4755-7-12] [Medline: 20573250]
35. Reece M, Hollub A, Nangami M, Lane K. Assessing male spousal engagement with prevention of mother-to-child transmission (pMTCT) programs in western Kenya. AIDS Care 2010;22(6):743-750. [doi: 10.1080/09540120903431330] [Medline: 20461572]
36. Orne-Gliemann J, Tchendjou PT, Miric M, Gadgil M, Butsashvili M, Eboko F, et al. Couple-oriented prenatal HIV counseling for HIV primary prevention: an acceptability study. BMC Public Health 2010;10(1):197 [FREE Full text] [doi: 10.1186/1471-2458-10-197] [Medline: 20403152]
37. Falnes EF, Moland KM, Tylleskär T, de Paoli MM, Msuya SE, Engebretsen IM. "It is her responsibility": Partner involvement in prevention of mother to child transmission of HIV programmes, northern Tanzania. J Int AIDS Soc 2011;14(1):21-21 [FREE Full text] [doi: 10.1186/1758-2652-14-21] [Medline: 21521511]
38. Hlongwa M, Mashamba-Thompson T, Makhunga S, Hlongwana K. Mapping evidence of intervention strategies to improving men's uptake to HIV testing services in sub-Saharan Africa: a systematic scoping review. BMC Infect Dis 2019;19(1):496. [doi: 10.1186/s12879-019-4124-y]
39. Hensen B, Taoka S, Taoka JJ, Lewis HA, Hargreaves J. Systematic review of strategies to increase men's HIV-testing in sub-Saharan Africa. AIDS (London) 2014;28:2133. [doi: 10.1097/qad.0000000000000395]
40. Ezeanolue EE, Obiefune MC, Yang W, Ezeanolue CO, Pharr J, Osuji A, et al. "What do you need to get male partners of pregnant women tested for HIV in resource limited settings" The baby shower cluster randomized trial. AIDS Behav 2016;21(2):587-596. [doi: 10.1007/s10461-016-1626-0]
41. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
42. Hlongwa M, Hlongwana K, Makhunga S, Choko AT, Dzinamarira T, Conserve D, et al. Linkage to HIV care following HIV self-testing among men: systematic review of quantitative and qualitative studies from six countries in sub-Saharan Africa. AIDS Behav 2023;27(2):651-666 [FREE Full text] [doi: 10.1007/s10461-022-03800-8] [Medline: 36094641]
43. Ashaba S, Cooper-Vince CE, Vořechovská D, Rukundo GZ, Maling S, Akena D, et al. Community beliefs, HIV stigma, and depression among adolescents living with HIV in rural Uganda. Afr J AIDS Res 2019;18(3):169-180 [FREE Full text] [doi: 10.2989/16085906.2019.1637912] [Medline: 31339461]
44. DiCarlo AL, Mantell JE, Remien RH, Zerbe A, Morris D, Pitt B, et al. 'Men usually say that HIV testing is for women': gender dynamics and perceptions of HIV testing in Lesotho. Cult Health Sex 2014;16(8):867-682 [FREE Full text] [doi: 10.1080/13691058.2014.913812] [Medline: 24854495]
45. Hensen B, Lewis JJ, Schaap A, Tembo M, Mutale W, Weiss HA, et al. Factors associated with HIV-testing and acceptance of an offer of home-based testing by men in rural Zambia. AIDS Behav 2015;19(3):492-504. [doi: 10.1007/s10461-014-0866-0] [Medline: 25096893]
46. Huerga H, Van Cutsem G, Ben Farhat J, Puren A, Bouhenia M, Wiesner L, et al. Progress towards the UNAIDS 90-90-90 goals by age and gender in a rural area of KwaZulu-Natal, South Africa: a household-based community cross-sectional survey. BMC Public Health 2018;18(1):1-8 [FREE Full text] [doi: 10.1186/s12889-018-5208-0] [Medline: 29499668]
47. Hamilton A, Thompson N, Choko AT, Hlongwa M, Jolly P, Korte JE, et al. HIV self-testing uptake and intervention strategies among men in sub-Saharan Africa: Aa systematic review. Front Public Health 2021;9:594298. [doi: 10.3389/fpubh.2021.594298]
48. Matovu JK, Wanyenze RK, Wabwire‐Mangen F, Nakubulwa R, Sekamwa R, Masika A, et al. "Men are always scared to test with their partners … It is like taking them to the Police": Motivations for and barriers to couples' HIV counselling and testing in Rakai, Uganda: a qualitative study. J Int AIDS Soc 2014;17(1):19160. [doi: 10.7448/ias.17.1.19160]
49. Simoni JM, Franks JC, Lehavot K, Yard SS. Peer interventions to promote health: conceptual considerations. Am J Orthopsychiatry 2011;81(3):351-359 [FREE Full text] [doi: 10.1111/j.1939-0025.2011.01103.x] [Medline: 21729015]
50. Bradford JB, Coleman S, Cunningham W. HIV System Navigation: an emerging model to improve HIV care access. AIDS Patient Care STDS 2007;21 Suppl 1(s1):S49-S58 [FREE Full text] [doi: 10.1089/apc.2007.9987] [Medline: 17563290]
51. Wilson TE, Gousse Y, Joseph MA, Browne RC, Camilien B, McFarlane D, et al. HIV prevention for black heterosexual men: the barbershop talk with brothers cluster randomized trial. Am J Public Health 2019;109(8):1131-1137. [doi: 10.2105/ajph.2019.305121]

XSL•FO
RenderX

52. Medley A, Kennedy C, O'Reilly K, Sweat M. Effectiveness of peer education interventions for HIV prevention in developing countries: a systematic review and meta-analysis. AIDS Educ Prev 2009;21(3):181-206 [FREE Full text] [doi: 10.1521/aeap.2009.21.3.181] [Medline: 19519235]
53. Shangani S, Escudero D, Kirwa K, Harrison A, Marshall B, Operario D. Effectiveness of peer-led interventions to increase HIV testing among men who have sex with men: a systematic review and meta-analysis. AIDS Care 2017;29(8):1003-1013 [FREE Full text] [doi: 10.1080/09540121.2017.1282105] [Medline: 28150501]
54. Chang LW, Nakigozi G, Billioux VG, Gray RH, Serwadda D, Quinn TC, et al. Effectiveness of peer support on care engagement and preventive care intervention utilization among pre-antiretroviral therapy, HIV-infected adults in Rakai, Uganda: a randomized trial. AIDS Behav 2015;19(10):1742-1751 [FREE Full text] [doi: 10.1007/s10461-015-1159-y] [Medline: 26271815]
55. Adebajo S, Eluwa G, Njab J, Oginni A, Ukwuije F, Ahonsi B, et al. Evaluating the effect of HIV prevention strategies on uptake of HIV counselling and testing among male most-at-risk-populations in Nigeria: a cross-sectional analysis. Sex Transm Infect 2015;91(8):555-560. [doi: 10.1136/sextrans-2014-051659] [Medline: 25921019]
56. The path that ends AIDS; 2023 UNAIDS global AIDS update. UNAIDS. 2023. URL: https://www.unaids.org/en/resources/documents/2023/global-aids-update-2023 [accessed 2026-03-11]
57. Barry KR, Seager M, Sullivan L. The Palgrave Handbook of Male Psychology and Mental Health. London: Palgrave; 2019.
58. Wippold GM, Frary SG, Garcia KA, Wilson DK. Implementing barbershop-based health-promotion interventions for Black men: a systematic scoping review. Health Psychol 2023;42(7):435-447 [FREE Full text] [doi: 10.1037/hea0001294] [Medline: 37227823]
59. Conserve DF, Mathews A, Janson S, Nwaozuru U, Jennings Mayo-Wilson L, Ritchwood TD, et al. Perceptions of HIV self-testing promotion in black barbershop businesses: implications for equitable engagement of black-owned small businesses for public health programs. BMC Public Health 2024;24(1):2440 [FREE Full text] [doi: 10.1186/s12889-024-19973-x] [Medline: 39245719]
60. Mayne J. The COM-B theory of change model. 2018. URL: https://www.researchprotocols.org/citations/parse/87612/1395683/6 [accessed 2019-07-22]
61. Semitala FC, Musinguzi A, Ssemata J, Welishe F, Nabunje J, Kadota JL, et al. Acceptance and completion of rifapentine-based TB preventive therapy (3HP) among people living with HIV (PLHIV) in Kampala, Uganda—patient and health worker perspectives. Implement Sci Commun 2021;2(1):17. [doi: 10.1186/s43058-021-00173-2]
62. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot Feasibility Stud 2016;2(1):64 [FREE Full text] [doi: 10.1186/s40814-016-0105-8] [Medline: 27965879]
63. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022;60(1):95-103. [doi: 10.1097/mlr.0000000000001664]
64. Consolidated guidelines for the prevention and treatment of HIV and AIDS in Uganda. Ministry of Health, Uganda. 2022. URL: http://library.health.go.ug/communicable-disease/hivaids/consolidated-guidelines-prevention-and-treatment-hiv-uganda [accessed 2026-03-11]
65. Guest G, Bunce A, Johnson L. How many interviews are enough? an experiment with data saturation and variability. Field Method 2006;18(1):59-82. [doi: 10.1177/1525822x05279903]
66. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
67. Hayes RJ, Moulton LH. Cluster Randomised Trials. New York: Chapman and Hall/CRC; 2017.
68. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Method 2006;5(1):80-92. [doi: 10.1177/160940690600500107]
69. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6(1):42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
70. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
71. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009;4(1):50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
72. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017;27(4):591-608 [FREE Full text] [doi: 10.1177/1049732316665344] [Medline: 27670770]
73. Musumari PM, Techasrivichien T, Srithanaviboonchai K, Wanyenze RK, Matovu JKB, Poudyal H, et al. HIV epidemic in fishing communities in Uganda: a scoping review. PLoS One 2021;16(4):e0249465 [FREE Full text] [doi: 10.1371/journal.pone.0249465] [Medline: 33793652]
74. Morfaw F, Mbuagbaw L, Thabane L, Rodrigues C, Wunderlich A, Nana P, et al. Male involvement in prevention programs of mother to child transmission of HIV: a systematic review to identify barriers and facilitators. Syst Rev 2013;2(1):5 [FREE Full text] [doi: 10.1186/2046-4053-2-5] [Medline: 23320454]

XSL•FO
RenderX

75. Barry JA, Kingerlee R, Seager M, Sullivan L. The Palgrave Handbook of Male Psychology and Mental Health. Berlin and Heidelberg: Springer; 2019.
76. Victor RG, Lynch K, Li N, Blyler C, Muhammad E, Handler J, et al. A cluster-randomized trial of blood-pressure reduction in black barbershops. N Engl J Med 2018;378(14):1291-1301. [doi: 10.1056/nejmoa1717250]
77. Victor RG, Ravenell JE, Freeman A, Leonard D, Bhat DG, Shafiq M, et al. Effectiveness of a barber-based intervention for improving hypertension control in black men: the BARBER-1 study: a cluster randomized trial. Arch Intern Med 2011;171(4):342-350 [FREE Full text] [doi: 10.1001/archinternmed.2010.390] [Medline: 20975012]
78. Randolph SD, Johnson R. A salon-based intervention to improve PrEP Uptake among black women. N Engl J Med 2024;390(9):776-777. [doi: 10.1056/nejmp2313708]
79. Factsheet-facts on HIV and AIDS in Uganda 2021 (based on Data ending 31st December 2020). UA Commission. 2020. URL: https://library.health.go.ug/communicable-disease/hivaids/facts-hiv-and-aids-uganda-2021-based-data-ending-31st-december-2020 [accessed 2020-12-31]
80. Facts on HIV and AIDS in Uganda 2020. UA Commission. 2020. URL: https://uac.go.ug/index.php/easy-customization/custom-404-page-and-offline-page [accessed 2026-03-11]
81. Investing in sexual and reproductive health is a pathway to better health in Kalangala District. Population Reference Bureau, Uganda. 2018. URL: https://www.prb.org [accessed 2026-03-11]
82. Uganda demographic health survey: HIV/AIDs related knowledge, attitudes, behaviour. UBOS. 2022. URL: https://www.ubos.org/uganda-demographic-and-health-survey-2022-main-report/ [accessed 2026-03-27]
83. Peer support for diabetes, heart disease and HIV/AIDS: a review of the clinical effectiveness, cost-effectiveness, and guidelines. Canadian Agency for Drugs and Technologies in Health. URL: https://pubmed.ncbi.nlm.nih.gov/24716255/ [accessed 2026-03-27]
84. Colella TJ, King KM. Peer support. An under-recognized resource in cardiac recovery. Eur J Cardiovasc Nurs 2004;3(3):211-217. [doi: 10.1016/j.ejcnurse.2004.04.001] [Medline: 15350230]
85. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehabil J 2004;27(4):392-401. [doi: 10.2975/27.2004.392.401] [Medline: 15222150]
86. Gamarel KE, Rebchook G, McCree BM, Jadwin-Cakmak L, Connolly M, Reyes LA, et al. The ethical imperative to reduce HIV stigma through community-engaged, status-neutral interventions designed with and for transgender women of colour in the United States. J Int AIDS Soc 2022;25 Suppl 1(Suppl 1):e25907 [FREE Full text] [doi: 10.1002/jia2.25907] [Medline: 35818894]
87. Hirshfield S, Birnbaum JM, Turner D, Roberson M, Bailey MM, Smith MDR, et al. An innovative adaptation of an HIV status-neutral, community-informed, socioemotional asset-building intervention with the house ball community. Health Promot Pract 2023;24(3):398-400. [doi: 10.1177/15248399221140794] [Medline: 36524579]
88. Nesta. Peer support what is it and does it work in summarising evidence from more than 1000 studies. National Voices. URL: https://www.nationalvoices.org.uk/sites/default/files/public/publications/peer_support_-_what_is_it_and_does_it_work.pdf [accessed 2026-03-11]
89. Berg RC, Page S, Øgård-Repål A. The effectiveness of peer-support for people living with HIV: a systematic review and meta-analysis. PLoS One 2021;16(6):e0252623 [FREE Full text] [doi: 10.1371/journal.pone.0252623] [Medline: 34138897]
90. Dooley BA, Fitzgerald A. My world survey: national study of youth mental health in Ireland. UCD School of Psychology. 2012. URL: https://researchrepository.ucd.ie/entities/publication/f028b522-c3ed-4e3a-8ffe-e1578443f885 [accessed 2026-03-27]
91. Craig P, Martin A, Browne S, Simpson SA, Wight D, Robling M, et al. Development of guidance for feasibility studies to decide whether and how to proceed to full-scale evaluation of complex public health interventions: a systematic review. Lancet 2018;392:S7. [doi: 10.1016/s0140-6736(18)32265-7]
92. HIV Prevention Trials Network. URL: https://www.hptn.org/research/studies/hptn-111 [accessed 2026-03-16]

Abbreviations

ACASI: audio computer–assisted self-interview
CHTS: community-based HIV testing services
COM-B: Capability, Opportunity, and Motivation Model of Behavior Change
CRF: case report form
DAIDS: Division of AIDS
HIVST: HIV self-testing
HPTN: HIV Prevention Trials Network
ICF: informed consent form
IRB: institutional review board
NIAID: National Institute of Allergy and Infectious Diseases
PrEP: preexposure prophylaxis
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

SMC: study monitoring committee
SoC: standard of care
SPIRIT: Standardized Protocol Items: Recommendations for Interventional Trials
SSA: sub-Saharan Africa
STI: sexually transmitted infection
TRIM: Testing a Barbershop-based HIV Prevention Initiative Among Men
UNAIDS: Joint United Nations Program on HIV/AIDS

Edited by A Schwartz; submitted 12.Nov.2025; this is a non–peer-reviewed article;accepted 23.Feb.2026; published 17.Apr.2026.

Please cite as:
Scoville C, Gati Mirembe B, Voldal E, Maena J, Etima J, Harris-Wisecarver J, Rukundo I, Ssemere H, Kemigisha D, Nakyanzi T, Nakalega R, Conserve DF, Albano M, Dyer M, Piwowar-Manning E, Spiegel H, Nakabiito C, Mujugira A, Donnell D, Lukyamuzi Z
Feasibility and Acceptability of Barbershop-Based HIV Prevention Among Heterosexual Men in Kalangala Islands, Uganda: Protocol for a Cluster Randomized Trial (HPTN 111)
JMIR Res Protoc 2026;15:e87612
URL: https://www.researchprotocols.org/2026/1/e87612
doi:10.2196/87612
PMID:

©Caitlin Scoville, Brenda Gati Mirembe, Emily Voldal, Joel Maena, Juliane Etima, Jayla Harris-Wisecarver, Ivan Rukundo, Hassan Ssemere, Doreen Kemigisha, Teopista Nakyanzi, Rita Nakalega, Donaldson F Conserve, Maggie Albano, Molly Dyer, Estelle Piwowar-Manning, Hans Spiegel, Clemensia Nakabiito, Andrew Mujugira, Deborah Donnell, Zubair Lukyamuzi. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Digital Diabetes Prevention Program for Hispanic Adolescents (Fit24+): Protocol for a Feasibility Randomized Controlled Trial

Erica G Soltero[1], PhD; Marbelly Partida[1], BS; Sandra Mihail[1], BS, MPH; Lizette Villanueva[1]; Callie Lopez[1], BS; Salma M Musaad[1], MD, PhD; Teresia M O'Connor[1], MD, MPH; Maria J Redondo[2], MD, MPH, PhD; Matthew Buman[3], PhD; Debbe Thompson[1], PhD

[1]USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, 1100 Bates St, Houston, TX, United States
[2]Division of Diabetes and Endocrinology, Department of Pediatrics, Baylor College of Medicine, Houston, TX, United States
[3]College of Health Solutions, Arizona State University, Phoenix, AZ, United States

Corresponding Author:
Erica G Soltero, PhD
USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, 1100 Bates St, Houston, TX, United States

Abstract

Background: Hispanic youth are disproportionately impacted by obesity and subsequent type 2 diabetes (T2D) yet remain underrepresented in diabetes prevention research. Digital health interventions hold promise for increasing the accessibility to and engagement in disease prevention programming, particularly among high-risk populations. However, there is a significant gap in the literature regarding digital T2D prevention programs for adolescents or Hispanic youth.

Objective: The objective of this article is to describe the protocol for examining the feasibility of a 12-week digital diabetes prevention program among Hispanic adolescents with obesity.

Methods: Participants (N=40; aged 12-16 years) will be randomized (1:1) to a 12-week intervention group or a control group. Youth in the intervention group will receive access to an e-learning platform with 12 nutrition and wellness video content sessions, a Fitbit Charge 5, and daily SMS text messages grounded in the self-determination theory to promote physical activity. Youth in the standard control group will receive information on diet and physical activity guidelines and guidance on setting behavior change goals. The study findings will focus on the evaluation of feasibility criteria: (1) recruitment of 40 Hispanic adolescents aged 12 to 16 years; (2) retention of 80% of the participants for postassessments; (3) integrity of the study protocol, defined as 70% or higher completion of content sessions and Fitbit wear on 5 days per week or more with response to 80% of SMS text messages when prompted; (4) 10% or lower incidence of technical issues; and (5) 80% or higher satisfaction among participants.

Results: This study was funded in August 2022 and intervention implementation is ongoing. To date, 35 participants have been enrolled. Study findings will be available before December 2026 and will focus on an evaluation of the a priori feasibility criteria on participant recruitment, data collection, integrity of the study protocol, technical issues, and satisfaction. Study findings will also focus on secondary outcomes.

Conclusions: The feasibility and process evaluation data obtained from this study will provide novel insights on the use of digital T2D prevention strategies among Hispanic youth and families and will inform the development of future digital health interventions among high-risk pediatric populations.

Trial Registration: ClinicalTrials.gov NCT06943001; https://clinicaltrials.gov/study/NCT06943001

International Registered Report Identifier (IRRID): PRR1-10.2196/75331

(JMIR Res Protoc 2026;15:e75331) doi:10.2196/75331

KEYWORDS

adolescent health; type 2 diabetes; physical activity; sleep; digital health; health disparities

Introduction

Over the past 20 years, Hispanic adolescents have experienced the largest increase in obesity (22.3% to 30.6%) and severe obesity (7.6% to 12.9%) compared with other pediatric and ethnic subgroups [1]. As a result, Hispanic youth have the highest lifetime risk of type 2 diabetes (T2D) and are disproportionately impacted by this disease [2]. When diagnosed with T2D in adulthood, Hispanic adults disproportionately experience diabetes-related complications and hospitalizations, highlighting the downstream consequences of diabetes disparities [3]. Lifestyle interventions focused on improving

XSL•FO
RenderX

diet and physical activity (PA) can reduce diabetes risk by 20% among high-risk adults [4]. However, evidence on the impact of lifestyle-based diabetes prevention interventions among Hispanic youth remains limited [5,6].

Among health promotion and disease prevention studies that do focus on Hispanic youth, accessibility remains an issue for reaching and engaging this population [6]. Traditional in-person interventions are typically delivered in 1 geographic location at a set day and time. However, limited transportation, the need for childcare, and conflicting work schedules can limit participation in in-person lifestyle interventions, especially among Hispanic youth and families, who are more likely to be impacted by some of these challenges [7-9]. Innovations in digital health have been recommended for overcoming barriers to accessibility as some digital devices and services allow for the delivery of disease prevention content directly to participants in their home environment [7,10]. By leveraging digital tools, digital health interventions (DHIs) can assist in overcoming geographic and temporal barriers, increasing the reach of these approaches [11]. DHIs that use devices and services already owned and operated by youth, such as smartphones or SMS text messaging, can increase accessibility by being cost-effective [12,13]. Approximately 95% of Hispanic teenagers in the United States report that they own a smartphone, which is comparable to the prevalence of smartphone ownership among non-Hispanic White youth (94%) [14]. Adolescents are the highest users of SMS text message communication, with a median rate of 60 messages sent per day [15], suggesting that the use of these digital tools is a familiar, accessible, and low-cost strategy. More advanced digital solutions such as chatbots can be deployed via SMS text message to provide personalized PA promotion strategies, which has been shown to increase engagement. Chatbots can also provide continuous, on-demand support for behavior change, further increasing the accessibility of health promotion and disease prevention content. However, there is a significant gap in the literature regarding digital T2D prevention programs for adolescents, particularly among Hispanic youth [16,17].

One limitation of existing DHIs is the lack of evidence-based interventions [18]. In addition to established efficacy, interventions that are culturally grounded and family based are essential for increasing the engagement with and effectiveness of prevention efforts among high-risk minority populations [6,19]. This study leverages previous work conducted on the development and testing of a 12-week culturally grounded, family-focused T2D prevention program [20]. This previously tested prevention program is grounded in social cognitive theory and is designed to foster social support and self-efficacy for improving diet and PA through key behavior change techniques, including goal setting and self-monitoring [21]. This 12-week program is traditionally delivered in a clinic or community setting by trained health educators or registered dietitians to cohorts of 5 to 7 families at a time. Feasibility and acceptability for this prevention program were first established in a pilot study (N=15) among Hispanic adolescents (12-16 years) with obesity (BMI ≥95th percentile) [22]. Participants demonstrated significant increases in insulin sensitivity, assessed through an oral glucose tolerance test [22]. The intervention was then refined and rigorously evaluated in an efficacy trial (N=160), which demonstrated significant short-term (12 weeks) improvements in insulin sensitivity and short- and long-term (12 months) improvements in BMI percentile and body fat percentage [20]. Given the demonstrated effectiveness of this evidence-based prevention program, adapting it into a DHI may serve to increase its reach and accessibility among those who need it the most.

To that end, the evidence-based prevention program described above was adapted for remote, digital delivery using a co-design approach, a process that has been previously described in detail [23]. Hispanic adolescents with obesity were engaged using mixed methods strategies, including in-depth interviews, a youth and expert panel, and several interactive activities to guide the development, design, and usability testing of all study components. The aim of this study is to assess the feasibility of a 12-week digital diabetes prevention program among Hispanic adolescents with obesity. We hypothesize that the program will be feasible and acceptable as demonstrated by a priori criteria for recruitment, retention, the integrity of the study protocol, technical issues, and satisfaction. The feasibility and process evaluation data to be obtained from this study will provide novel insights on the use of digital T2D prevention strategies among Hispanic youth and families and will inform the development of future DHIs among high-risk pediatric populations.

Methods

Study Design

This 2-armed, pretest-posttest design will include 40 Hispanic adolescents aged 12 to 16 years with obesity (BMI ≥95th percentile). Youth will be randomized (1:1 ratio) to the intervention or control group. Youth in the 12-week intervention group will receive access to an e-learning platform, a Fitbit Charge 5 (Google), and SMS text messages grounded in the self-determination theory. Youth randomized to the control group will receive a 1-page handout on dietary and PA guidelines for adolescents and standard counseling on developing specific, measurable, attainable, relevant, and time-bound (SMART) goals to improve these health behaviors. All participants will undergo data collection at baseline and 3 months after the intervention.

Participants

We will recruit equal numbers of male and female participants who meet the following specific inclusion criteria: (1) self-identification as Hispanic individuals; (2) age between 12 and 16 years; (3) obesity, defined as a BMI in the 95th percentile or above; and (4) ownership of a smartphone. Participants will be excluded if they (1) take medication or have been diagnosed with a condition that significantly influences their diet or ability to be active, (2) have experienced a recent hospitalization or injury that would impact participation in diet changes or activity, (3) have been diagnosed with T2D, (4) are pregnant, or (5) are currently enrolled in an exercise program or currently using a PA monitoring device. We will recruit 40 youths anticipating 20% attrition, yielding a final sample size of 32. This sample size is adequate for determining feasibility parameters and generating data on preliminary efficacy [24,25].

XSL•FO
RenderX

Participant Recruitment

Recruitment will be conducted in collaboration with local pediatric clinics, including Texas Children's Pediatrics, Harris Health pediatric clinics, and smaller pediatric practices across the greater Houston metropolitan area, Texas. These community-based clinics offer preventive services to youth in underserved, predominantly minority communities. Most patients treated in these clinics qualify for Medicaid. The research team will work with each clinic to implement strategies for physicians to refer potentially eligible patients to the study. We will also collaborate with community-based organizations (eg, Houston Public Library and BakerRipley Community Developers) for recruitment. Study information will be disseminated by attending community events (eg, health fairs and back-to-school fairs), hosting health education workshops at partnering organizations, and attending service events such as Houston food bank food drives. The research team includes bilingual and bicultural staff members who share study information and answer questions in English and Spanish to facilitate recruitment.

Ethical Considerations

The study protocol and all study-related materials have been approved by the Institutional Review Board at Baylor College of Medicine (H-49195). All study-related documents will be available in English and Spanish, with bilingual and bicultural research staff administering consent procedures, conducting data collection, and answering questions. Prior to any study procedures, trained research team members will obtain written parental consent and child assent. Participants will be informed that their participation is voluntary and they are free to withdraw from the study at any time. Participants will also be informed that nonparticipation will not affect any health or medical services they currently receive and that confidentiality will be maintained using unique identification numbers and a password-protected database. Given the use of third-party digital services, business associate agreements, which are legal contracts required under HIPAA (Health Insurance Portability and Accountability Act) to protect data and personal health information, have been exacted with Fitabase (Fitbit data), Mobile Coach (SMS text messaging), and LearnWorlds (nutrition education platform) to ensure that no data are shared outside of these companies or used in any way outside of the study aims. Participants will receive US $140 in compensation for completion of data collection assessments. This study is registered at ClinicalTrials.gov (NCT04953442).

Procedures

Intervention Overview

For the nutrition education component, we will implement a nutrition and wellness curriculum, which is designed to foster self-efficacy and social support to promote behavior changes. Participants will be asked to watch 1 nutrition and wellness content session (25-35 minutes) per week as a family on a day and time that works best for their schedule. Self-efficacy will be fostered through the educational content, which will include nutrition knowledge regarding topics such as macronutrients, portion sizes, and reading nutrition labels. Textbox 1 provides the titles of all 12 sessions. Each session consists of a prerecorded content video available in Spanish or English that includes didactic teaching delivered in a conversational style by 2 bilingual, bicultural research team members. The content videos include interactive features such as quizzes, animations, and additional links to health information on the content topic. To access the content videos, the adolescent participants and all participating family members will use an e-learning platform hosted by LearnWorlds (LearnWorlds Ltd). This platform is a password-protected website and has an accompanying mobile phone app that mirrors the website, allowing participants to access intervention content using a laptop, desktop computer, tablet, or smartphone.

At the start of the intervention, families will receive an intervention toolkit that will include a workbook. Inside the workbook are content session handouts (for the family and for the individual youth), youth-focused activities to reinforce the concepts and behavior change techniques taught in the content sessions, and a section for listing weekly SMART goals for behavior change that will lead to the mastery of skills through goal attainment. Social support will be fostered within families through the family-focused activities and as families are encouraged to support each other toward their SMART behavior goals. Social support will also be fostered as families will receive a monthly phone call from a research team member to answer questions, provide encouragement, and address any technical issues. Families will be asked to work together as a team to make 4 different healthy, culturally appropriate recipes throughout the intervention. To increase the feasibility of making these recipes, the intervention toolkit will also include a few nonperishable food items. For example, one of the recipes is a lentil soup, so each family will receive a can of lentils and a can of beef broth. Brief (1-3 minutes) demonstration videos for each recipe have been developed and are available through the e-learning platform.

XSL•FO
RenderX

Textbox 1. Nutrition education curriculum topics.

- Session 1: Getting Started
- Session 2: Health Awareness
- Session 3: Roles & Responsibilities
- Session 4: Keep it Moving
- Session 5: How Sweet Are You?
- Session 6: Champions with Breakfast
- Session 7: Slim the Fat
- Session 8: Fast Food
- Session 9: Snack Attack
- Session 10: Stay Strong
- Session 11: Self-Esteem
- Session 12: Balancing Act

Additional components have been developed within the e-learning platform to further foster self-efficacy and social support among and within participating families. The platform includes a discussion board, a recipe tab, and a support tab where participants can ask questions about content or report technical issues (Figure 1). Families will be encouraged to use the discussion board to share successes and challenges regarding SMART goals, their experience in trying new recipes, and other tips and strategies for behavior change. The recipe tab will provide families with additional healthy snack, breakfast, lunch, and dinner recipes. The LearnWorlds platform also allows us to connect a secure, institutional Zoom account (Zoom Video Communications) to the platform, which will be used to host monthly live, synchronous sessions between families and the research team. Live sessions will be hosted separately for parents and youth to enable participants to ask questions about the nutrition education content, share their experiences throughout the program, and further share their success and challenges regarding behavior change.

XSL•FO
RenderX

Figure 1. Screenshot of the LearnWorlds e-learning platform.



PA Component

For the PA component, adolescent participants will receive a Fitbit Charge 5 device. The Fitbit Charge has shown high validity and reliability (intraclass correlation coefficient=0.71-1.00) for assessing steps per day in adolescents [26,27]. Youth will also receive access to Steph, a chatbot that will be deployed via SMS text message. Steph is designed to assist youth in goal setting and intention planning and to provide support for achieving step goals and feedback on goal attainment. The needs support offered by Steph is grounded in the self-determination theory, as shown in Figure 2.

Figure 2. Conceptual map demonstrating how SMS text message components align with self-determination theory constructs (autonomy, competence, and relatedness) to promote autonomous motivation for physical activity.

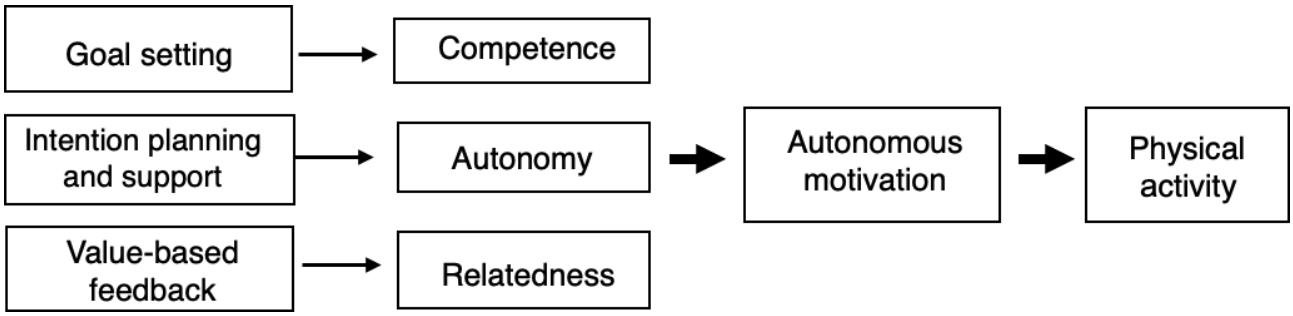

On Mondays, Steph will provide youth with a step goal. In line with the self-determination theory, goal-setting SMS text messages are written in a manner that considers the participants' autonomy by giving them the choice to accept the step goal that was provided by the research team or allowing them to suggest a goal that is more feasible for them that week. Goals will be adaptive and will increase successively by 10% of the previous week's performance throughout the 12-week intervention to guide the participants to meet or maintain the current PA recommendation of 12,000 steps per day [28,29]. For the

XSL•FO
RenderX

intention planning component, Steph will ask the following questions: (1) "What activities will you do to meet your goal?" (2) "On what day and at what time will you do this activity to meet your goal?" (3) "Why is this goal important to you?" These questions are also designed to promote autonomy by providing control and choice over how the participant will achieve their goal. Steph will also provide accountability by following up with the participant on the days and times they intend to be active. On Tuesdays and Thursdays, Steph will provide competence support by providing the participants with a menu of multiple-choice support options: (1) an evidence-based behavior change technique, (2) tips or strategies for overcoming common barriers to PA, or (3) activity suggestions. On the basis of the participant's selection, Steph will send a corresponding support message. On Wednesdays, Steph will ask participants to report on their step progress to further encourage self-monitoring toward goal attainment. Asking youth to report on their own step progress as opposed to sending these data to them further engages them in the self-monitoring process [30]. On Saturdays, Steph will send value-based feedback on goal attainment to promote relatedness. These messages will tie goal attainment to commonly held values among Hispanic adolescents, including being healthy and fit and getting good grades [31]. Connecting goal attainment with a self-endorsed value can promote internalization of the behavior so that it becomes integrated into one's sense of self [32].

Steph is a rule-based chatbot trained on a library of 125 messages that were developed in a previous pilot study conducted by the research team, and she is constrained so that she can only draw from our library of messages [31]. Steph is not capable of 2-way communication and only uses natural language processing to provide youth with PA suggestions when prompted. For example, a participant may ask Steph to recommend a 20-minute aerobic exercise that can be done in the home environment, and Steph will provide the participant with a feasible recommendation.

Control Group

Participants randomized to the control group will receive a 1-page informational handout with diet and PA guidelines for adolescents [33]. They will also receive general behavioral counseling information via SMS text message on setting SMART goals to promote healthy diet and activity behaviors.

Pre- and Postintervention Follow-Up

All youths, regardless of group assignment, will participate in a baseline preintervention assessment and a 12-week postintervention assessment.

Primary Outcomes

The primary outcome of this study is feasibility, which will be evaluated using the a priori criteria outlined in Table 1 [34]. Recruitment and data collection criteria will assess the feasibility of recruiting and retaining high-risk Hispanic youth and families. Process evaluation data on the number of participants screened, enrolled, and assessed will be collected via a REDCap (Research Electronic Data Capture; Vanderbilt University) electronic database. The "integrity of the study protocol" criteria will be used to identify the feasible and acceptable intervention dose [25]. This will be assessed using data extracted from the e-learning platform, including the number of content sessions completed, the frequency of engagement with the e-learning platform measured using the number of log-ins to the website, and the days of Fitbit device wear [35]. Criteria regarding technical issues will assess the feasibility of using an online e-learning platform, a Fitbit device, and an SMS text message–enabled chatbot. Technical issues experienced by the research team and technical issues reported by participants will be logged in REDCap. Satisfaction criteria will assess program acceptability via an exit satisfaction survey. Satisfaction will be further evaluated by conducting exit interviews among all participants randomized to the intervention group. The purpose of these interviews will be to elicit participants' attitudes, perceptions, and suggestions for improving the program. Interviews will be audio recorded and transcribed, and a thematic content or narrative analysis will be used to identify emergent themes on participant satisfaction [36,37].

Table . A priori feasibility criteria.

Category	Criteria
Recruitment	Recruiting and enrolling 40 Hispanic adolescents aged 12-16 y
Data collection	Data collection on 100% of the participants at baseline and 80% of the participants at the postintervention time point
Integrity of the study protocol	≥70% completion of content sessions and Fitbit device wear on ≥5 d per wk, with response to 80% of SMS text messages when prompted
Technical issues	≤10% technical issues with the e-learning platform, Fitbit device, and SMS text message transmissions
Satisfaction	≥80% of participants reporting "excellent" to "good" satisfaction with the intervention

Secondary Outcomes

The secondary outcomes will include BMI percentile and PA. Height and weight will be measured to the nearest 0.1 cm and 0.1 kg, respectively, and used to calculate sex- and age-adjusted BMI percentile as well as the percentage of participants in the 95th percentile, which is more sensitive to detecting clinical changes in BMI in populations with severe obesity [38-40]. We will assess hemoglobin A_{1c} (HbA_{1c}) using a point-of-care

XSL•FO
RenderX

analyzer (DCA Analyzer; Abbott Rapid Diagnostics [41]). Given that Fitbit devices provide less accurate measures of PA compared with research-grade devices, a hip-worn accelerometer (ActiGraph wGT3X-BT; Ametris) will be used to assess PA at the pre- and postintervention time points only [42]. Participants will be asked to wear the accelerometer for 7 days and will receive daily SMS text messages to promote compliance in device wear [43,44]. A valid day of wear will be defined as 10 hours or more of wear on 4 days or more, one of which has to be a weekend day [45,46]. Data will be analyzed using Evenson cutoff points in the ActiLife software (Ametris) to report average minutes of moderate to vigorous PA, light PA, and total PA per day [47,48].

Tertiary Outcomes

Tertiary outcomes will include autonomous motivation for PA and digital health equity. Autonomous motivation for PA will be assessed using the Autonomous Motivation for Physical Activity scale of the Behavioral Regulation in Exercise Questionnaire [49,50], which has been validated among adolescents [51]. We used the framework for digital health equity developed by Richardson et al [52] to develop a digital health equity survey to assess access to digital technology, digital self-efficacy, and digital literacy among participating adults and youth. This information will be used to assess the degree to which the digital devices and services used in this study were implemented in an equitable manner.

Data Analyses

Power Analysis

A formal power calculation was not used as the primary purpose of this study is to examine feasibility [25,53]. We will oversample and recruit 40 youths with the conservative assumption that 20% will be lost to attrition, leaving a proposed final sample size of 32, which is adequate for estimating feasibility parameters needed to inform a fully powered intervention [24,25]. We will strive to recruit equal numbers of male and female participants.

Primary Outcome Analysis

Standard descriptive analyses including mean estimates and their 95% CIs will be performed on process evaluation metrics, including recruitment; data collection; website engagement; device wear; technical issues regarding the e-learning platform, Fitbit devices, and SMS text message transmission; and participant satisfaction as assessed via the exit survey. Descriptive data on process evaluation will be compared with feasibility criteria to determine the overall feasibility of this approach. To reduce bias in the analysis and interpretation, the lead biostatistician is blinded, and the randomization scheme is implemented by a research staff member who is not involved in recruitment or daily interactions with study participants. To assess satisfaction and acceptability of the intervention, in-depth exit interviews will be audio recorded; transcribed verbatim using a professional transcription service; coded by 2 trained, independent coders; and qualitatively analyzed using NVivo (version 9; Lumivero). Thematic content analysis will be used to identify emergent patterns and insights on participant satisfaction and feasibility of the digital health tools used in the intervention [36].

Analysis of Secondary Outcomes

Comparisons of BMI percentile, percentage of participants in the 95th percentile, HbA_{1c}, and PA between (Mann-Whitney *U* test) and within (pretest to posttest; paired 1-tailed *t* tests and signed rank test) treatment groups will be conducted. While clinically significant changes in secondary and tertiary outcomes are not expected, we will examine changes in the outcomes and clinically meaningful effects to determine the preliminary efficacy of this pilot study [54,55].

Results

This study is ongoing. Study recruitment began in September 2025 and is projected to conclude in March 2026. At the time of this publication, we have enrolled 35 participants. Study findings will be available before December 2026 and will focus on an evaluation of the feasibility of this study, as well as secondary outcomes (BMI percentile, HbA_{1c}, and PA). Study outcomes will be reported using the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) reporting guidelines for DHIs. Study results will be communicated to participating pediatric clinics and community partners through in-person presentations and written reports. The results will also be disseminated through scientific publications, as well as through national and international conferences.

Discussion

Anticipated Findings

Only 1 culturally grounded T2D prevention program has been developed and tested among Hispanic youth and families [6]. Despite its efficacy, this in-person program is limited to delivery in 1 location and is time and resource intensive. Because DHIs can leverage commonly owned and operated digital devices and services to deliver intervention content directly to participants, they have the potential for increasing the reach and engagement of evidence-based programs [10]. This paper describes the protocol for examining the feasibility of using an e-learning platform in conjunction with a Fitbit device and chatbot via SMS text messaging for delivering a culturally grounded, evidence-based T2D prevention program among Hispanic youth with obesity and their families.

This study expands upon traditional prevention programs by using digital devices and services to develop a flexible, adaptive intervention that is highly accessible and meets the needs of the focus population. Using an e-learning platform with asynchronous content sessions, a Fitbit device, and a chatbot deployed via SMS text messaging will enable families to access the content on multiple types of digital devices at their convenience, allowing participants to navigate common barriers such as work schedules and other household responsibilities. Use of the chatbot to provide a PA recommendation significantly expands access to behavior change support by providing real-time support that is personalized, contextualized, and

XSL•FO
RenderX

accessible when it is needed the most by the participant. Schedule conflicts and lack of program flexibility are consistently named among the most common reasons why parents withdraw their children from traditional disease prevention interventions [56]. Among adults, digital diabetes prevention programs have demonstrated that having highly accessible and flexible content can increase engagement and long-term outcomes [57]. Specific features within the e-learning platform such as recipes and the discussion board were recommended by youth and parents as desired intervention components [23]. Additionally, interacting with the chatbot through action planning was also desired by youth as they recognized the need for assistance in outlining the steps needed to achieve their behavior goals and their need for accountability. Providing a variety of interactive features has also been shown to increase participant engagement and contribute to study outcomes [58,59]. The feasibility and outcome data from this study will provide novel insights on the ability of this digital strategy to impact reach and engagement among Hispanic youth and families.

This study also expands upon traditional prevention programs by focusing on a population impacted by health disparities. Most DHIs are developed and tested on populations with a higher socioeconomic status, and few have focused on high-risk, minority youth [60,61]. This study fills this gap by focusing on Hispanic youth with obesity during the critical life stage of adolescence. The feasibility criteria and additional measures such as the digital health equity survey will provide important information on the user experience, including ease of use, perspectives on functionality, and the accessibility of the digital and intervention components to be used in this study. A recent review of DHIs implemented among adolescents reported that very few studies included any measure of digital health equity [62]. This information is important given that failure to use digital technology in an equitable manner could exacerbate health disparities and potentially contribute to digital disparities [61,62].

Implications of Anticipated Findings

DHIs that are personalized, adaptive, and accessible have the potential to change the way in which disease prevention programs are delivered in vulnerable communities. According to the US Preventive Services Task Force, improvements in metabolic health requires more than 52 contact hours, which is not feasible in the current health care model [33]. By providing flexible access to nutrition and wellness content through the e-learning platform and the continuous delivery of adaptive, personalized PA content through Fitbit and SMS text messaging, these tools can extend clinical care to meet task force recommendations, possibly leading to greater improvements in metabolic outcomes [63]. Advancements in clinical informatics have demonstrated that data from e-learning platforms and activity trackers can be integrated into electronic health record systems, thus having the potential for DHIs to be integrated into clinical care [64,65]. If found to be feasible, this is a scalable approach that could have significant clinical utility and could make a significant contribution toward addressing T2D disparities within the Hispanic population.

Potential Limitations

This study is limited by its focus on Hispanic youth with obesity. However, this focus is warranted given the disproportionate impact of T2D on Hispanic youth. In light of the digital divide, there is potential for some Hispanic youths to be excluded from the study due to lack of smartphone ownership. However, recent national statistics have demonstrated that smartphone ownership among Hispanic adolescents is very high (90%) and is comparable to smartphone ownership among non-Hispanic White youth. We will carefully document the number of participants who are excluded due to lack of access to a smartphone. Because youth randomized to the intervention group receive a Fitbit device, SMS text messages, and an online nutrition education platform, there is potential for attention bias that may impact the study effects, which is a limitation. This study is also limited by the small sample size. However, the sample size is appropriate for establishing the feasibility and acceptability of the intervention prior to conducting a fully powered trial that is designed to detect statistically and clinically meaningful changes in health outcomes. Finally, an additional limitation of this study is that it is designed to assess feasibility and, therefore, not powered to identify clinically meaningful changes in weight, BMI, or health behaviors. However, assessing feasibility using metrics that are collected both objectively (eg, recruitment, data collection, technical issues, and device wear) and subjectively (eg, satisfaction) will fill a knowledge gap and inform the development of future DHIs for high-risk adolescents.

Conclusions

There is a critical need for T2D prevention research focused on high-risk Hispanic youth and families with obesity. This study will deliver a digital T2D prevention program to address growing T2D disparities among this population using an e-learning platform, a Fitbit activity tracker, and SMS text messages grounded in the self-determination theory. This study will lead to novel insights on the feasibility of using these digital tools and services to deliver disease prevention programming to Hispanic youth and families. If found to be feasible, this study will guide the development and implementation of DHIs among other high-risk pediatric populations.

Funding

This work was supported by the US Department of Agriculture Agricultural Research Service (cooperative agreement 58-3092-0-001). The contents of this work are solely the responsibility of the authors and do not necessarily represent the official views of the US Department of Agriculture. This work was also supported by a grant from the National Institute for Diabetes and Digestive and Kidney Diseases of the National Institutes of Health (K01DK128682) awarded to EGS.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Ogden CL, Fryar CD, Martin CB, et al. Trends in obesity prevalence by race and Hispanic origin-1999-2000 to 2017-2018. JAMA 2020 Sep 22;324(12):1208-1210. [doi: 10.1001/jama.2020.14590] [Medline: 32857101]
2. Imperatore G, Boyle JP, Thompson TJ, et al. Projections of type 1 and type 2 diabetes burden in the U.S. population aged <20 years through 2050: dynamic modeling of incidence, mortality, and population growth. Diabetes Care 2012 Dec;35(12):2515-2520. [doi: 10.2337/dc12-0669] [Medline: 23173134]
3. Aguayo-Mazzucato C, Diaque P, Hernandez S, Rosas S, Kostic A, Caballero AE. Understanding the growing epidemic of type 2 diabetes in the Hispanic population living in the United States. Diabetes Metab Res Rev 2019 Feb;35(2):e3097. [doi: 10.1002/dmrr.3097] [Medline: 30445663]
4. Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. Diabetes Care 2002 Dec;25(12):2165-2171. [doi: 10.2337/diacare.25.12.2165] [Medline: 12453955]
5. McCurley JL, Crawford MA, Gallo LC. Prevention of type 2 diabetes in U.S. Hispanic youth: a systematic review of lifestyle interventions. Am J Prev Med 2017 Oct;53(4):519-532. [doi: 10.1016/j.amepre.2017.05.020] [Medline: 28688727]
6. Cruz P, Granados A. Type 2 diabetes in Latino youth: a clinical update and current challenges. Curr Probl Pediatr Adolesc Health Care 2019 Jan;49(1):16-22. [doi: 10.1016/j.cppeds.2018.11.008] [Medline: 30567629]
7. Soltero EG, O'Connor TM, Thompson D, Shaibi GQ. Opportunities to address obesity disparities among high-risk Latino children and adolescents. Curr Obes Rep 2021 Sep;10(3):332-341. [doi: 10.1007/s13679-021-00445-x] [Medline: 34263434]
8. Skelton JA, Palakshappa D, Moore JB, Irby MB, Montez K, Rhodes SD. Community engagement and pediatric obesity: incorporating social determinants of health into treatment. J Clin Transl Sci 2019 Dec 19;4(4):279-285. [doi: 10.1017/cts.2019.447] [Medline: 33244407]
9. Haire-Joshu D, Hill-Briggs F. The next generation of diabetes translation: a path to health equity. Annu Rev Public Health 2019 Apr 1;40(391-410):391-410. [doi: 10.1146/annurev-publhealth-040218-044158] [Medline: 30601723]
10. López L, Tan-McGrory A, Horner G, Betancourt JR. Eliminating disparities among Latinos with type 2 diabetes: effective eHealth strategies. J Diabetes Complications 2016 Apr;30(3):554-560. [doi: 10.1016/j.jdiacomp.2015.12.003] [Medline: 26774790]
11. Erku D, Khatri R, Endalamaw A, et al. Digital health interventions to improve access to and quality of primary health care services: a scoping review. Int J Environ Res Public Health 2023 Sep 28;20(19):6854. [doi: 10.3390/ijerph20196854] [Medline: 37835125]
12. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36(393-415):393-415. [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
13. Maher CA, Lewis LK, Ferrar K, Marshall S, De Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40. [doi: 10.2196/jmir.2952] [Medline: 24550083]
14. Pew Research Center. Teens, Social Media and Technology 2018. 2018. URL: https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018 [accessed 2026-03-13]
15. Lenhart A, Ling R, Campbell S, Purcell K. Teens and Mobile Phones.: Pew Research Center; 2010. URL: https://www.pewresearch.org/internet/2010/04/20/teens-and-mobile-phones [accessed 2026-03-13]
16. Giovanelli A, Ozer EM, Dahl RE. Leveraging technology to improve health in adolescence: a developmental science perspective. J Adolesc Health 2020 Aug;67(2S):S7-S13. [doi: 10.1016/j.jadohealth.2020.02.020] [Medline: 32718517]
17. Partridge SR, Redfern J. Strategies to engage adolescents in digital health interventions for obesity prevention and management. Healthcare (Basel) 2018 Jun 21;6(3):70. [doi: 10.3390/healthcare6030070] [Medline: 29933550]
18. Mumtaz H, Riaz MH, Wajid H, et al. Current challenges and potential solutions to the use of digital health technologies in evidence generation: a narrative review. Front Digit Health 2023;5:1203945. [doi: 10.3389/fdgth.2023.1203945] [Medline: 37840685]
19. Thornton PL, Kumanyika SK, Gregg EW, et al. New research directions on disparities in obesity and type 2 diabetes. Ann N Y Acad Sci 2020 Feb;1461(1):5-24. [doi: 10.1111/nyas.14270] [Medline: 31793006]
20. Soltero EG, Olson ML, Williams AN, et al. Effects of a community-based diabetes prevention program for Latino youth with obesity: a randomized controlled trial. Obesity (Silver Spring) 2018 Dec;26(12):1856-1865. [doi: 10.1002/oby.22300] [Medline: 30426694]
21. Williams AN, Konopken YP, Keller CS, et al. Culturally-grounded diabetes prevention program for obese Latino youth: Rationale, design, and methods. Contemp Clin Trials 2017 Mar;54(68-76):68-76. [doi: 10.1016/j.cct.2017.01.004] [Medline: 28104469]
22. Shaibi GQ, Konopken Y, Hoppin E, Keller CS, Ortega R, Castro FG. Effects of a culturally grounded community-based diabetes prevention program for obese Latino adolescents. Diabetes Educ 2012;38(4):504-512. [doi: 10.1177/0145721712446635] [Medline: 22585870]

XSL•FO
RenderX

23. Mihail S, Partida M, Villanueva L, et al. Community participatory co-design and development of a digital diabetes prevention education program for Hispanic families with obesity: mixed methods study. JMIR Form Res 2026 Feb 11;10:e67800. [doi: 10.2196/67800] [Medline: 41671476]
24. Billingham SAM, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. BMC Med Res Methodol 2013 Aug 20;13(104):104. [doi: 10.1186/1471-2288-13-104] [Medline: 23961782]
25. Arain M, Campbell MJ, Cooper CL, Lancaster GA. What is a pilot or feasibility study? A review of current practice and editorial policy. BMC Med Res Methodol 2010 Jul 16;10(67):67. [doi: 10.1186/1471-2288-10-67] [Medline: 20637084]
26. Ridgers ND, McNarry MA, Mackintosh KA. Feasibility and effectiveness of using wearable activity trackers in youth: a systematic review. JMIR Mhealth Uhealth 2016 Nov 23;4(4):e129. [doi: 10.2196/mhealth.6540] [Medline: 27881359]
27. Witt D, Kellogg R, Snyder M, Dunn J. Windows into human health through wearables data analytics. Curr Opin Biomed Eng 2019 Mar;9(28-46):28-46. [doi: 10.1016/j.cobme.2019.01.001] [Medline: 31832566]
28. Mendoza JA, Baker KS, Moreno MA, et al. A Fitbit and Facebook mHealth intervention for promoting physical activity among adolescent and young adult childhood cancer survivors: a pilot study. Pediatr Blood Cancer 2017 Dec;64(12). [doi: 10.1002/pbc.26660] [Medline: 28618158]
29. Tudor-Locke C, Craig CL, Aoyagi Y, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act 2011 Jul 28;8(80):80. [doi: 10.1186/1479-5868-8-80] [Medline: 21798044]
30. Morton K, Sutton S, Hardeman W, et al. A text-messaging and pedometer program to promote physical activity in people at high risk of type 2 diabetes: the development of the PROPELS follow-on support program. JMIR Mhealth Uhealth 2015 Dec 15;3(4):e105. [doi: 10.2196/mhealth.5026] [Medline: 26678750]
31. Soltero E, Lopez C, Mihail S, et al. An SMS text message-based type 2 diabetes prevention program for Hispanic adolescents with obesity: qualitative co-design process. JMIR Form Res 2023 Aug 2;7:e46606. [doi: 10.2196/46606] [Medline: 37531191]
32. Thompson D, Cantu D, Bhatt R, et al. Texting to increase physical activity among teenagers (TXT Me!): rationale, design, and methods proposal. JMIR Res Protoc 2014 Mar 12;3(1):e14. [doi: 10.2196/resprot.3074] [Medline: 24622344]
33. US Preventive Services Task Force, Grossman DC, Bibbins-Domingo K, et al. Screening for obesity in children and adolescents: US Preventive Services Task Force recommendation statement. JAMA 2017 Jun 20;317(23):2417-2426. [doi: 10.1001/jama.2017.6803] [Medline: 28632874]
34. O'Connor TM, Beltran A, Musaad S, et al. Feasibility of targeting Hispanic fathers and children in an obesity intervention: Papás Saludables Niños Saludables. Child Obes 2020 Sep;16(6):379-392. [doi: 10.1089/chi.2020.0006] [Medline: 32466678]
35. Cadmus-Bertram L, Marcus BH, Patterson RE, Parker BA, Morey BL. Use of the Fitbit to measure adherence to a physical activity intervention among overweight or obese, postmenopausal women: self-monitoring trajectory during 16 weeks. JMIR Mhealth Uhealth 2015 Nov 19;3(4):e96. [doi: 10.2196/mhealth.4229] [Medline: 26586418]
36. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
37. Bernard HR, Ryan GW. Analyzing Qualitative Data: Systematic Approaches: Sage Publications; 2009.
38. Flegal KM, Wei R, Ogden CL, Freedman DS, Johnson CL, Curtin LR. Characterizing extreme values of body mass index-for-age by using the 2000 Centers for Disease Control and Prevention growth charts. Am J Clin Nutr 2009 Nov;90(5):1314-1320. [doi: 10.3945/ajcn.2009.28335] [Medline: 19776142]
39. Gulati AK, Kaplan DW, Daniels SR. Clinical tracking of severely obese children: a new growth chart. Pediatrics 2012 Dec;130(6):1136-1140. [doi: 10.1542/peds.2012-0596] [Medline: 23129082]
40. National Health and Nutrition Examination Study (NHANES): Anthropometry Procedures Manual: Centers for Disease Control and Prevention; 2020. URL: https://wwwn.cdc.gov/nchs/data/nhanes/2019-2020/manuals/2020-Anthropometry-Procedures-Manual-508.pdf [accessed 2026-03-16]
41. Yonel Z, Kuningas K, Sharma P, et al. Concordance of three point of care testing devices with clinical chemistry laboratory standard assays and patient-reported outcomes of blood sampling methods. BMC Med Inform Decis Mak 2022 Sep 22;22(1):248. [doi: 10.1186/s12911-022-01999-z] [Medline: 36138408]
42. Reddy RK, Pooni R, Zaharieva DP, et al. Accuracy of wrist-worn activity monitors during common daily physical activities and types of structured exercise: evaluation study. JMIR Mhealth Uhealth 2018 Dec 10;6(12):e10338. [doi: 10.2196/10338] [Medline: 30530451]
43. Trost SG, Pate RR, Freedson PS, Sallis JF, Taylor WC. Using objective physical activity measures with youth: how many days of monitoring are needed? Med Sci Sports Exerc 2000 Feb;32(2):426-431. [doi: 10.1097/00005768-200002000-00025] [Medline: 10694127]
44. Sirard JR, Slater ME. Compliance with wearing physical activity accelerometers in high school students. J Phys Act Health 2009;6 Suppl 1(Suppl 1):S148-S155. [doi: 10.1123/jpah.6.s1.s148] [Medline: 19998861]
45. Belton S, O'Brien W, Wickel EE, Issartel J. Patterns of noncompliance in adolescent field-based accelerometer research. J Phys Act Health 2013 Nov;10(8):1181-1185. [doi: 10.1123/jpah.10.8.1181] [Medline: 23223684]
46. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S531-S543. [doi: 10.1249/01.mss.0000185657.86065.98] [Medline: 16294116]

XSL•FO
RenderX

47. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci 2008 Dec;26(14):1557-1565. [doi: 10.1080/02640410802334196] [Medline: 18949660]
48. Evenson KR, Terry JW Jr. Assessment of differing definitions of accelerometer nonwear time. Res Q Exerc Sport 2009 Jun;80(2):355-362. [doi: 10.1080/02701367.2009.10599570] [Medline: 19650401]
49. Wilson PM, Rodgers WM, Loitz CC, Scime G. "It's Who I Am … Really!' the importance of integrated regulation in exercise contexts. J Appl Biobehavioral Res 2006 Apr;11(2):79-104 [FREE Full text] [doi: 10.1111/j.1751-9861.2006.tb00021.x]
50. Mullan E, Markland D, Ingledew DK. A graded conceptualisation of self-determination in the regulation of exercise behaviour: development of a measure using confirmatory factor analytic procedures. Pers Individ Dif 1997 Nov;23(5):745-752. [doi: 10.1016/S0191-8869(97)00107-4]
51. Markland D, Tobin V. A modification to the behavioural regulation in exercise questionnaire to include an assessment of amotivation. J Sport Exerc Psychol 2004;26(2):191-196. [doi: 10.1123/jsep.26.2.191]
52. Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. NPJ Digit Med 2022 Aug 18;5(1):119. [doi: 10.1038/s41746-022-00663-0] [Medline: 35982146]
53. Kraemer HC, Mintz J, Noda A, Tinklenberg J, Yesavage JA. Caution regarding the use of pilot studies to guide power calculations for study proposals. Arch Gen Psychiatry 2006 May;63(5):484-489. [doi: 10.1001/archpsyc.63.5.484] [Medline: 16651505]
54. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011 May;45(5):626-629. [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]
55. Tingey R, Lambert M, Burlingame G, Hansen N. Assessing clinical significance: proposed extensions to method. Psychother Res 1996;6(2):109-123. [doi: 10.1080/10503309612331331638] [Medline: 22242565]
56. Sallinen Gaffka BJ, Frank M, Hampl S, Santos M, Rhodes ET. Parents and pediatric weight management attrition: experiences and recommendations. Child Obes 2013 Oct;9(5):409-417. [doi: 10.1089/chi.2013.0069] [Medline: 24028563]
57. Sepah SC, Jiang L, Ellis RJ, McDermott K, Peters AL. Engagement and outcomes in a digital Diabetes Prevention Program: 3-year update. BMJ Open Diabetes Res Care 2017;5(1):e000422. [doi: 10.1136/bmjdrc-2017-000422] [Medline: 28948027]
58. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
59. Metzendorf MI, Wieland LS, Richter B. Mobile health (m-health) smartphone interventions for adolescents and adults with overweight or obesity. Cochrane Database Syst Rev 2024 Feb 20;2(2):CD013591. [doi: 10.1002/14651858.CD013591.pub2] [Medline: 38375882]
60. Nollen NL, Mayo MS, Carlson SE, Rapoff MA, Goggin KJ, Ellerbeck EF. Mobile technology for obesity prevention: a randomized pilot study in racial- and ethnic-minority girls. Am J Prev Med 2014 Apr;46(4):404-408. [doi: 10.1016/j.amepre.2013.12.011] [Medline: 24650843]
61. Coghlan S, Leins K, Sheldrick S, Cheong M, Gooding P, D'Alfonso S. To chat or bot to chat: ethical issues with using chatbots in mental health. Digit Health 2023;9:20552076231183542. [doi: 10.1177/20552076231183542] [Medline: 37377565]
62. Partridge SR, Knight A, Todd A, et al. Addressing disparities: a systematic review of digital health equity for adolescent obesity prevention and management interventions. Obes Rev 2024 Dec;25(12):e13821. [doi: 10.1111/obr.13821] [Medline: 39256958]
63. Yingling LR, Brooks AT, Wallen GR, et al. Community engagement to optimize the use of web-based and wearable technology in a cardiovascular health and needs assessment study: a mixed methods approach. JMIR Mhealth Uhealth 2016 Apr 25;4(2):e38. [doi: 10.2196/mhealth.4489] [Medline: 27113680]
64. Dinh-Le C, Chuang R, Chokshi S, Mann D. Wearable health technology and electronic health record integration: scoping review and future directions. JMIR Mhealth Uhealth 2019 Sep 11;7(9):e12861. [doi: 10.2196/12861] [Medline: 31512582]
65. Baer HJ, Rozenblum R, De La Cruz BA, et al. Effect of an online weight management program integrated with population health management on weight change: a randomized clinical trial. JAMA 2020 Nov 3;324(17):1737-1746. [doi: 10.1001/jama.2020.18977] [Medline: 33141209]

Abbreviations

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
DHI: digital health intervention
HbA_{1c}: hemoglobin A_{1c}
HIPAA: Health Insurance Portability and Accountability Act
PA: physical activity
REDCap: Research Electronic Data Capture
SMART: specific, measurable, attainable, relevant, and time-bound

XSL•FO
RenderX

T2D: type 2 diabetes

Edited by A Mavragani; submitted 01.Apr.2025; peer-reviewed by E Beaulieu, M Hachem; revised version received 06.Nov.2025; accepted 07.Nov.2025; published 15.Apr.2026.

Please cite as:
Soltero EG, Partida M, Mihail S, Villanueva L, Lopez C, Musaad SM, O'Connor TM, Redondo MJ, Buman M, Thompson D
A Digital Diabetes Prevention Program for Hispanic Adolescents (Fit24+): Protocol for a Feasibility Randomized Controlled Trial
JMIR Res Protoc 2026;15:e75331
URL: https://www.researchprotocols.org/2026/1/e75331
doi:10.2196/75331

© Erica G Soltero, Marbelly Partida, Sandra Mihail, Lizette Villanueva, Callie Lopez, Salma M Musaad, Teresia M O'Connor, Maria J Redondo, Matthew Buman, Debbe Thompson. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Role of Surgical Telementoring in the Acquisition of Surgical Skills in Laparoscopic Cholecystectomy: Protocol for a Pilot Randomized Controlled Trial

Khayam Butt[1], MD; Laurent Olivier Trichet[2], MSc; Gunnar Hartvigsen[3], Prof Dr; Ole Sjo[4], MD, PhD; Bjørn von Gohren Edwin[5], MD, PhD; Knut-Magne Augestad[6], MD, PhD

[1]Department of Surgery, Nordland Hospital Trust, Parkveien 96, Bodø, Norway

[2]Department of Research, Nordland Hospital Trust, Bodø, Norway

[3]Institute of Informatics, UiT The Arctic University of Norway, Tromsø, Norway

[4]Department of Hepatic, Gastrointestinal and Paediatric Surgery, Oslo University Hospital, Oslo, Norway

[5]The Intervention Center, Medical Faculty, Oslo University Hospital, Oslo, Norway

[6]Department of Surgery, Akershus University Hospital, Lillestrøm, Norway

Corresponding Author:
Khayam Butt, MD
Department of Surgery, Nordland Hospital Trust, Parkveien 96, Bodø, Norway

Abstract

Background: Developing surgical skills among residents requires time and resources, often in conflict with the goals of efficiency and hospital economics in surgical practice. Optimizing resident training may prevent a decline in quality. Skilled surgeons are vital to effective and safe surgical performance, yet the training trajectory for surgical residents is often inefficient, particularly in the era of minimally invasive surgery. A framework that efficiently disseminates surgical skills is essential to cultivate skilled future surgeons within an acceptable time frame. Surgical telementoring (ST) aligns well with minimally invasive surgery because the mentee operating in the field shares the telementor's perspective. Despite growing experience with telemedicine, data on educational outcomes remain limited.

Objective: We aim to conduct a pilot randomized controlled trial (RCT). Data collection will inform a future full-scale RCT evaluating whether real-time ST, combined with structured postoperative coaching, is superior to traditional on-site mentoring in developing laparoscopic cholecystectomy skills among surgical residents.

Methods: Eligible residents will be randomly assigned in a 1:1 ratio to either the intervention group (real-time ST and postoperative coaching) or the control group (traditional intraoperative hands-on teaching). This research adheres to the extended SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) statements for pilot RCTs and follows the intention-to-treat principle. Residents in both groups will perform 5 consecutive procedures under standardized intraoperative conditions. The primary outcomes are the Global Operative Assessment of Laparoscopic Skills (GOALS) score and the Non-Technical Skills for Surgeons (NOTSS). We hypothesize that the intervention group will show a 3- to 5-point improvement in the GOALS score compared with the control group. To achieve 0.8 statistical power, a P value <.05, and a 20% dropout rate in a definitive RCT, 12 residents per group are required. We present a pilot RCT with 5 residents per group. Challenges encountered during the trial will be reported, and amendments will be proposed to assess the need for a future definitive RCT.

Results: The primary outcomes will be technical and nontechnical surgical skills, measured by the GOALS and NOTSS scores, respectively. Patient-reported experience measures and satisfaction scores of eligible mentees and mentors will be analyzed as secondary outcomes. The ST system configuration at the 2 inclusion centers was completed in December 2023. Ethics approval and the most recent update report to ClinicalTrials.gov were obtained on October 23, 2023, and December 8, 2024, respectively. Recruitment of residents to the study commenced in spring 2024. Recruitment is expected to be completed in May 2026, with data analysis in November 2026 and manuscript submission of results in January 2027.

Conclusions: The SURGTEACH (Surgical Telementoring in Acquisition of Surgical Skills of Laparoscopic Cholecystectomy) trial is the first pilot RCT evaluating ST for educational outcomes. This research may provide high-quality evidence to improve surgical education, particularly in rural hospitals.

Trial Registration: ClinicalTrial.gov NCT06421584; https://clinicaltrials.gov/study/NCT06421584

International Registered Report Identifier (IRRID): DERR1-10.2196/73159

(JMIR Res Protoc 2026;15:e73159) doi:10.2196/73159

XSL•FO
RenderX

KEYWORDS

surgical telementoring; pilot randomized controlled trial; technical and nontechnical surgical skills; interprofessional education; competency-based education; videoconferencing; surgical education; laparoscopic cholecystectomy

Introduction

Developing surgical skills among residents is a time- and resource-consuming process [1]. Increasing demands for efficiency and hospital economics in surgical practice have contributed to a decrease in the number of surgeons per capita [2]. Concerning reports show increasingly limited access to the operating room (OR) during a surgical residents' educational path [3]. Hence, efficient training trajectories for general surgeons are needed to ensure educational quality and efficiency [4].

Geographical challenges, including long distances and harsh weather conditions in Northern Norway, necessitate complex transportation arrangements to reach the hospital. A healthcare policy has been established that requires rurally located institutions to provide basic surgical care within an acceptable proximity to patients [5]. However, these institutions have limited access to surgical expertise, which may contribute to reduced quality of OR training for residents. The Norwegian National Medical Association has already taken this into account. This is one of the leading reasons for the reduced number of surgical procedures required to meet the Norwegian national requirements for achieving a specialty in general and colorectal surgery.

The centralization of surgical care, while improving outcomes for complex procedures at high-volume centers, has led to a decline in surgical services at smaller hospitals and greater challenges in recruiting and retaining consultant surgeons in rural areas.

This phenomenon is particularly pronounced in Scandinavian countries, where vast geographic distances amplify the impact of centralization policies. Norwegian rural hospitals increasingly rely on rotating surgeons and visiting consultants, creating discontinuity in surgical training programs and limiting residents' exposure to consistent mentorship. A consequence of centralization is a lack of expert mentorship and a reduction in patient volume. Increased use of simulation training, surgical telementoring (ST), and information and communication technologies in many aspects of health care delivery are interventions being adopted to address these gaps [5-7]. These structural challenges underscore the need for innovative educational solutions that can deliver expert surgical guidance regardless of physical location.

Laparoscopic cholecystectomy (LC) is a technically challenging procedure that residents become familiar with early on under supervision. Potential complications can have serious and hazardous consequences for the patient during major bile duct injury. Substantial hands-on supervision during the initial phases of the learning curve, with gradual achievement of independence, is expected.

Laparoscopic procedures have a complex learning curve that affects clinical and financial outcomes [5,8]. However, expertise in laparoscopic surgery is not readily available, posing a challenge for resident training, particularly in smaller hospitals with limited surgical case volume. These factors call for optimizing the training period for surgical residents. There is some evidence supporting the cost-effectiveness and safety of ST systems [9].

Despite significant technological advances and growing experience with telecommunication and telementoring in health care, the evidence base for educational outcomes remains limited [10-12]. A systematic review by Erridge et al [13] found that, although telementoring demonstrates feasibility and safety, most studies lack rigorous comparative designs to evaluate skill acquisition. A subsequent systematic review from December 2023 by Fernandes et al [14] examined the educational outcomes of surgical telementoring (ST) during surgery. Using ST as the intervention in a continuous professional development program, the authors used the Moores Outcome Framework to evaluate educational outcomes [15]. They reported a lack of high-quality research, such as randomized controlled trials (RCTs), to elucidate the causal effect of ST on educational outcomes.

Furthermore, Chepkoech et al [16] reviewed telementoring systems for surgical training in low- and middle-income countries. They called for standardized qualitative and quantitative metrics, embedded within well-defined training frameworks, to assess the long-term effects of ST on surgical competence and to enable reproducibility and cross-study comparisons.

The Laparoscopic Colorectal (Lapco) Training the Trainers (TT) program, developed in the United Kingdom, provides a validated framework for standardized surgical mentoring and communication. The program established structured communication protocols between the mentor (trainer) and mentee (trainee) during Lapco surgery [17,18]. These principles have been adopted internationally, including in the Norwegian adoption (LapcoNor), to ensure consistent, high-quality intraoperative teaching.

Initiated in 2014, LapcoNor was a collaboration between the Norwegian Health Authority (South East) and Lapco International UK to standardize training in Lapco surgery in Norway [19]. The Norwegian program was described in detail in the Norwegian surgical journal, *Kirurgen* [20]. Its standardized communication framework provides an ideal foundation for ST, where a clear, structured dialogue is essential to overcome the physical separation between mentor and trainee. The Lapco TT program, adapted to Norway by the LapcoNor group, was used exclusively to prepare and train the involved mentors prior to participation. LapcoNor principles served as the intraoperative communication framework during resident supervision of surgical procedures.

Multimedia Appendix 1 describes how the LapcoNor group supervised the included telementors and on-site mentors to

XSL•FO
RenderX

integrate the communication principles described by Mackenzie et al [17] into the study.

While Lapco principles were used at the micro level to facilitate concise verbal communication and enhance technical competency by addressing surgical challenges, the GROW (goals, reality, options, wrap-up) model was used for surgical coaching [21,22]. Surgical coaching has emerged as an effective method for developing surgical competence and will be implemented using the GROW model. Unlike mentoring, surgical coaching using the GROW model provides a macrostructure for verbal interactions during the telementoring session. It may facilitate the mentee's ability to make independent decisions, thereby evolving both technical and nontechnical surgical skills. A description of the GROW model of surgical coaching, adapted for telementoring sessions, is provided in Multimedia Appendix 2.

There has been a focus on procedure-specific metrics to structure surgical education and assessment, minimizing subjectivity in the trainee-trainer relationship. van Zwieten et al [23] presented a systematic review of assessment methods in Lapco surgery. They differentiated between procedure-specific and general rating scales for the objective assessment of technical skills, enabling accurate feedback in a training environment.

This study applies Lapco communication principles and standardized coaching using the GROW model in the intervention arm, thereby minimizing communicative bias while evaluating the educational efficacy of the telementoring modality. Using the Global Operative Assessment of Laparoscopic Skills (GOALS) score, the most widely used general rating system for laparoscopic procedures, the study enables comparison with existing and future literature. The GOALS score, including its parameters and scoring systems, is shown in Figure 1.

Given the increasing reliance on remote surgical education following the COVID-19 pandemic, there is a need for RCTs to assess ST as an intervention, enabling the acquisition of technical and nontechnical surgical skills to improve individual surgeons' performance and patient outcomes. An RCT of this caliber will inherently pose strategic and practical challenges that require resources, which might be wasted due to poor study design [24]. Pilot trials are recommended as an essential initial step in assessing the feasibility of a full-scale RCT [25].

ST may offer several advantages over traditional hands-on mentoring: (1) the telementor shares the same visual perspective as the operating resident without competing for the operative field; (2) the structured communication required in telementoring may enhance deliberate feedback; (3) the combination of real-time guidance with systematic postoperative coaching using the GROW model may reinforce learning more effectively than traditional, unstructured mentoring; and (4) the telementor can provide undivided attention to teaching without the distraction of potentially needing to take over the procedure.

This pilot RCT aims to provide foundational data for a future full-scale trial. The study aims to evaluate whether real-time ST, combined with structured postoperative coaching, is superior to traditional on-site mentoring for developing surgical skills among surgical residents. Our primary hypothesis is that residents receiving telementoring will show a 3- to 5-point improvement in GOALS scores compared with the control group.

XSL•FO
RenderX

Figure 1. GOALS (Global Operative Assessment of Laparoscopic Skills). The assessed parameters and the scoring system are shown.

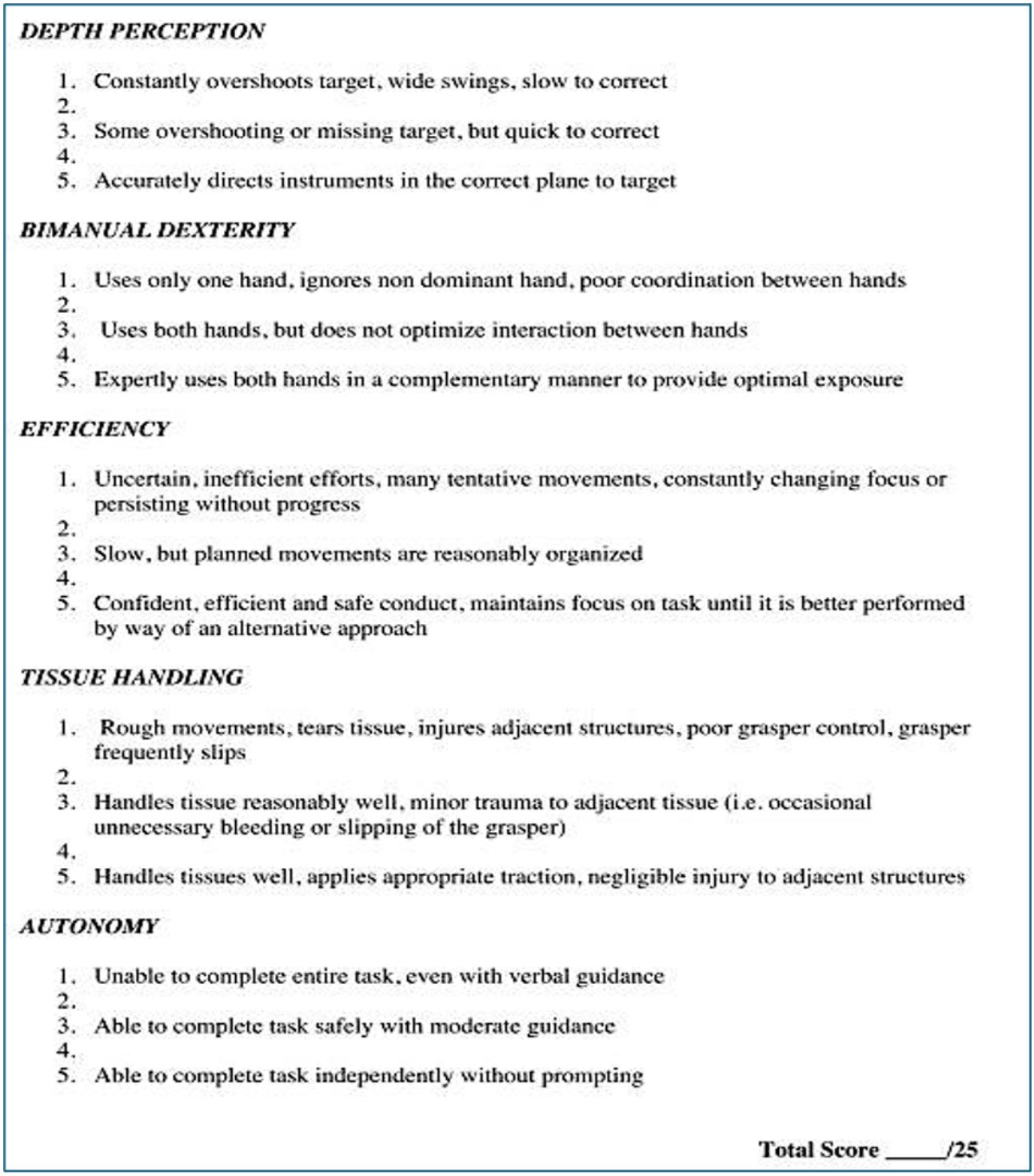

DEPTH PERCEPTION

1. Constantly overshoots target, wide swings, slow to correct
2.
3. Some overshooting or missing target, but quick to correct
4.
5. Accurately directs instruments in the correct plane to target

BIMANUAL DEXTERITY

1. Uses only one hand, ignores non dominant hand, poor coordination between hands
2.
3. Uses both hands, but does not optimize interaction between hands
4.
5. Expertly uses both hands in a complementary manner to provide optimal exposure

EFFICIENCY

1. Uncertain, inefficient efforts, many tentative movements, constantly changing focus or persisting without progress
2.
3. Slow, but planned movements are reasonably organized
4.
5. Confident, efficient and safe conduct, maintains focus on task until it is better performed by way of an alternative approach

TISSUE HANDLING

1. Rough movements, tears tissue, injures adjacent structures, poor grasper control, grasper frequently slips
2.
3. Handles tissue reasonably well, minor trauma to adjacent tissue (i.e. occasional unnecessary bleeding or slipping of the grasper)
4.
5. Handles tissues well, applies appropriate traction, negligible injury to adjacent structures

AUTONOMY

1. Unable to complete entire task, even with verbal guidance
2.
3. Able to complete task safely with moderate guidance
4.
5. Able to complete task independently without prompting

Total Score _____/25

Secondary objectives include assessing Non-Technical Skills for Surgeons (NOTSS) score, participant satisfaction, and feasibility outcomes. Table 1 presents the NOTSS score, including their parameters and scoring system, and Figure 2 presents the scoring systems used to obtain satisfaction scores from the involved mentees and mentors in both the intervention and control arms.

XSL•FO
RenderX

Table . The NOTSS (Non-Technical Skills for Surgeons) score.

Category and elements	Rating[a]
Situation awareness	
Gathering information	1, 2, 3, 4
Understanding information	1, 2, 3, 4
Projecting and anticipating future state	1, 2, 3, 4
Decision-making	
Considering options	1, 2, 3, 4
Selecting and communicating options	1, 2, 3, 4
Implementing and reviewing decisions	1, 2, 3, 4
Communication and teamwork	
Exchanging information	1, 2, 3, 4
Establishing shared understanding	1, 2, 3, 4
Coordinating team activities	1, 2, 3, 4
Leadership	
Setting and maintaining standards	1, 2, 3, 4
Supporting others	1, 2, 3, 4
Coping with pressure	1, 2, 3, 4

[a]1: poor, 2: marginal, 3: acceptable, 4: good.

This pilot study will specifically (1) evaluate the feasibility of recruitment, randomization, and intervention delivery; (2) assess the acceptability of the telementoring system among residents and mentors; (3) identify logistical challenges requiring modification before a definitive RCT; and (4) provide preliminary effect size estimates to refine sample size calculations for a future full-scale trial.

Figure 2. Resident and mentor satisfaction survey to obtain satisfaction scores from both the intervention and control arms of the study. The satisfaction survey used a 5-point Likert scale, with 1=strong disagreement and 5=strong agreement.

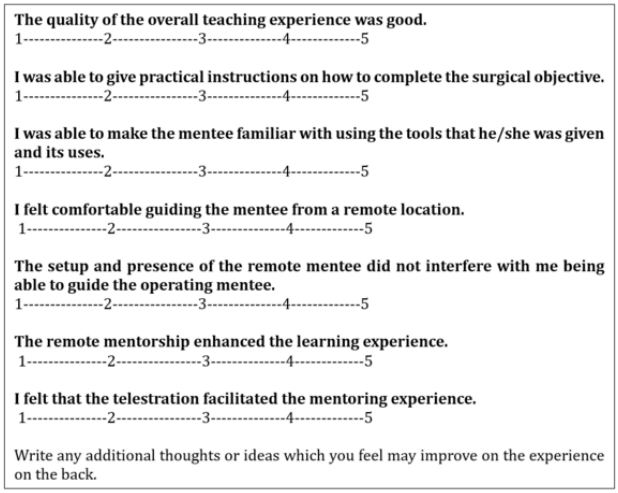

Mentors

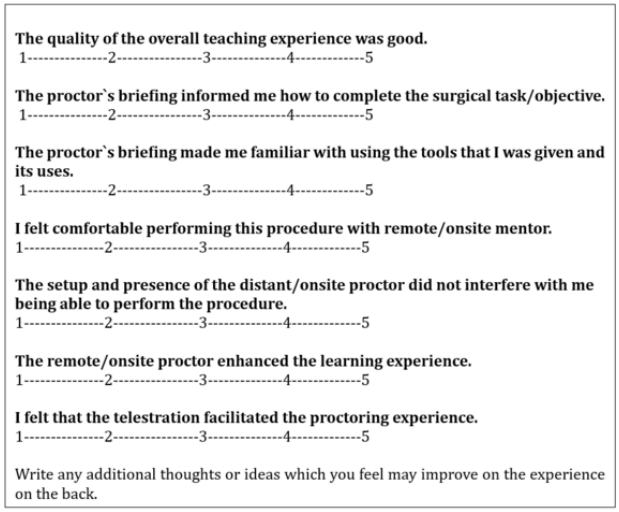

Mentees

Methods

Study Design and Setting

This study is a parallel, 2-group, assessor-blind, multicenter pilot RCT. Participating surgical residents are allocated in a 1:1 ratio to either the intervention group, which receives real-time intraoperative telementoring and postoperative coaching, or the control group, which receives traditional intraoperative mentoring. The intention-to-treat principle will be applied. The study adheres to the latest SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2025 statements recommended for RCT protocols, as shown in Checklist 1 [26].

The Northern Norwegian Health Care Trust comprises 11 hospitals, including 1 university hospital and 10 local hospitals. Four hospitals within the Northern Norwegian Health Care Trust were initially assessed for eligibility as study sites:

XSL•FO
RenderX

Sandnessjøen, Mo i Rana, Stokmarknes, and Bodø. Following a feasibility assessment, 2 hospitals—Sandnessjøen and Mo i Rana—were selected as active study sites based on surgical volume, resident availability, geographic location for ease of patient and resident inclusion in the study, and telementoring infrastructure, including the availability of telementors and on-site mentors. Stokmarknes and Bodø were not included as active recruitment sites due to anticipated pressure from acute cases in these locations, which would affect the elective surgical plan that the telementoring project was part of. Additionally, one of the initially included telementors based in Stokmarknes had to withdraw from the study. Although the telementoring infrastructure was initially installed at these locations to support patient and resident inclusion in the study, we ultimately decided to restrict inclusion facilities to Mo i Rana and Sandnessjøen. This facilitated optimal collaboration among actively participating telementors, residents, and the project administrator, KB, who was located in Bodø. Depending on the results of this pilot study, the expansion to include Stokmarknes and Bodø in a definitive RCT will be considered.

For this pilot RCT, 2 of the local hospitals, Sandnessjøen and Mo i Rana, served as study sites. Both locations offer surgical residency training programs that provide LC training.

Technical Setup of the Telementoring System

Both hospitals, including patients in the study, have implemented ST systems in the ORs. The Olympus telecollaboration platform, MedPresence, provides a telementoring system that uses voice instructions delivered via headphones to assist the mentee throughout laparoscopic surgery. Simultaneously, the telementor observes the procedure in real time via a live stream from the laparoscopic camera. Telestration may be used, at the request of the telementor or the telementee, to provide annotations on a static picture of the ongoing surgical procedure. The telestration session begins when the telementor instructs the mentee to cease surgery and directs their attention to the visual aspect. Telementors may demonstrate surgical planes and strategize dissection.

The telementor may alternate between the external camera, which provides an overview of the laparoscopic trocars and the spatial positions of the OR and anesthesia nurse, and the laparoscopic camera. The system captures audio from the mentor-mentee dialog and footage from the laparoscopic surgery and the external camera. The external camera and its audio recording provide a foundation for assessing interactions between the telementee and OR staff, enabling the evaluation of the telementee's nontechnical skills.

In both study arms, the standard operative team includes the operating resident, a camera-holding assistant (assisting surgeon), a scrub nurse, and a circulating nurse.

In the intervention group, the local assisting surgeon, an experienced laparoscopic surgeon, holds the laparoscopic camera and assists the operating resident solely by moving the camera to meet the resident's demands, providing instructional guidance only when patient safety requires it. The camera-holding surgeon's role is limited to emergency assistance. In the control arm, the on-site mentor, an experienced laparoscopic surgeon, is scrubbed and provides active instruction throughout the procedure, while the same support personnel configuration is maintained in the OR. Figure 3 shows the intraoperative setup during a telementored procedure in the intervention arm.

In the control arm, the Lapco principles of intraoperative communication and the GROW model are not applied during the mentor's preoperative, intraoperative, and postoperative guidance of the mentee.

XSL•FO
RenderX

Figure 3. Configuration of the operating room (OR) setup during telementored sessions in the intervention group. Participating personnel include a resident, a camera-holding surgeon, a scrub nurse, a telementor (remotely located), an anesthesia nurse, and an anesthesiologist. Note the small microphone placed beside the soundbar, beneath the screens displaying the laparoscopic camera images.

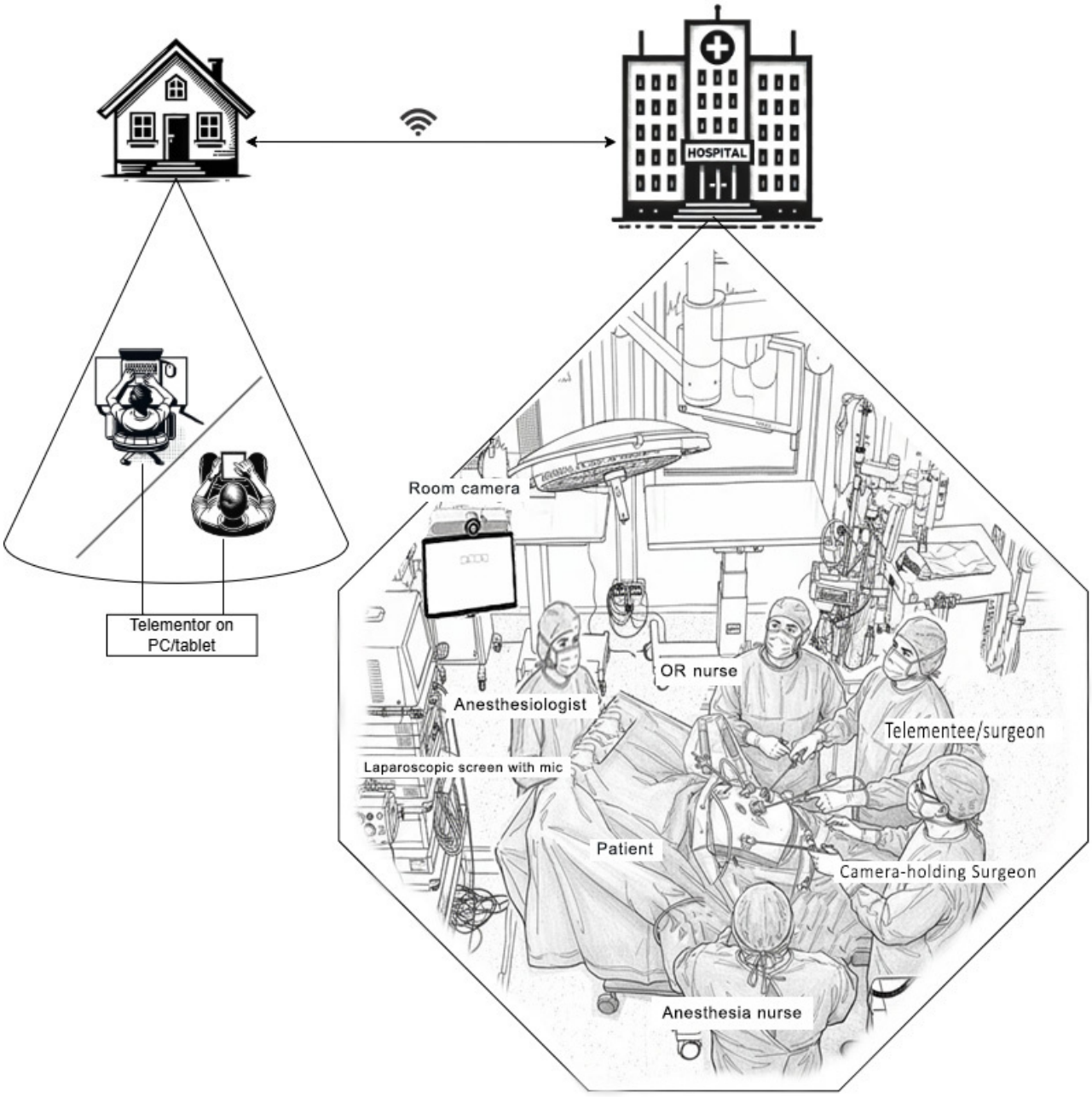

Eligibility Criteria: Residents, Mentors, and Patients

Residents

The eligibility criteria for the residents are as follows:

- General surgery residents in years 1 to 6 of their specialty education who have partially or completely performed more than 10 laparoscopic procedures
- Stratification according to experience
- Having completed the mandatory national course on general laparoscopic principles and the mandatory LC course
- Agreement with the mentor on the communication model during surgery. This model is derived from the LapcoNor principles [17]. Residents in the intervention group received an introduction to communication principles during telementoring and an orientation to the telementoring equipment used in the OR.

Mentors

The eligibility criteria for the mentors are as follows:

- Consultant surgeon with more than 5 years of experience performing LCs independently
- Acquainted with the assessment of videos for the GOALS score [27]
- Both telementors and on-site mentors had to attend the national LapcoNor TT course and follow standardized communication norms with the mentee during surgical mentoring, thus diminishing the bias of communicative difference.

XSL•FO
RenderX

Patients

The eligibility criteria for the patients are as follows:

- Gallstone disease without a clinical history of cholecystitis
- BMI <38
- No previous history of upper abdominal laparotomy
- No previous history of percutaneous gallbladder drainage
- Patient provided informed consent

Each telementoring procedure will be followed by a coaching session in which the mentee describes their performance (an opportunity for self-reflection). The telementor will provide performance-enhancing feedback to formulate a take-home message. Both the GROW model and the Lapco TT model will be used in telementoring sessions as macro- and micromodels for surgical guidance.

Residents will continue their standard surgical training during the study period. No restrictions will be imposed on other educational activities (eg, simulation training and courses). To promote participant retention, procedures will be scheduled within a 3-day block to minimize scheduling conflicts. Residents who cannot complete all 5 procedures will be followed for available outcomes. Reasons for withdrawal will be documented.

Recruitment Procedures

The residents will be randomized into 2 groups. Residents in the control group will receive traditional OR supervision throughout all 5 consecutive procedures, with on-site mentoring by an experienced surgeon. The intervention group will receive guidance from a remotely located expert via live stream during 5 consecutive procedures.

Recruitment of research participants (surgical residents) will primarily occur at their local hospitals. During 2024 and 2025, all patients referred to hospital surgical departments with a diagnosis of gallstone disease and scheduled for LC will be assessed for inclusion. Study residents will be scheduled for 5 LCs within 3 days. If not enough residents are recruited at the local hospitals, recruitment efforts will be extended to include residents from other hospitals.

Allocation and Blinding

The randomization process will use a predetermined statistical sequence generated by an external researcher who was not directly engaged in the study. The allocation sequence will be concealed using a centralized web-based randomization system. Personnel enrolling participants will not have access to the sequence until after enrollment. All individuals who actively participate in the experiment will be granted access to the allocation codes after randomization. Residents are enrolled by KB. After enrollment confirmation, the allocation is revealed by KB.

Computer-generated stratification will be used for randomization to create balanced and unpredictable groups. To ensure balance across groups with respect to significant confounders, the randomization process will be stratified by trainees' experience, explicitly distinguishing between those who had performed 5 to 10 LCs and those who had performed 10 to 20.

Participant Timeline

To ensure transparency and adherence to reporting standards for RCTs, according to the Consolidated Standards of Reporting Trials (CONSORT), the trial design is summarized in the CONSORT flowchart in Figure 4. The timeline of surgical residents' inclusion in both groups is shown. Residents who consented to inclusion will be informed of randomization 1 week before the planned procedures. In addition to the planned stepwise procedure description, standardized information on communication during the telementoring procedure will be provided.

This protocol follows the SPIRIT 2025 guidelines for clinical trial protocols. A complete SPIRIT checklist is provided as Checklist 1.

XSL•FO
RenderX

Figure 4. Design of the SURGTEACH (Surgical Telementoring in Acquisition of Surgical Skills of Laparoscopic Cholecystectomy) trial: CONSORT flow and participant timeline. The intention-to-treat (ITT) principle will be applied. CONSORT: Consolidated Standards of Reporting Trials; GOALS: Global Operative Assessment of Laparoscopic Skills; NOTSS: Non-Technical Skills for Surgeons.

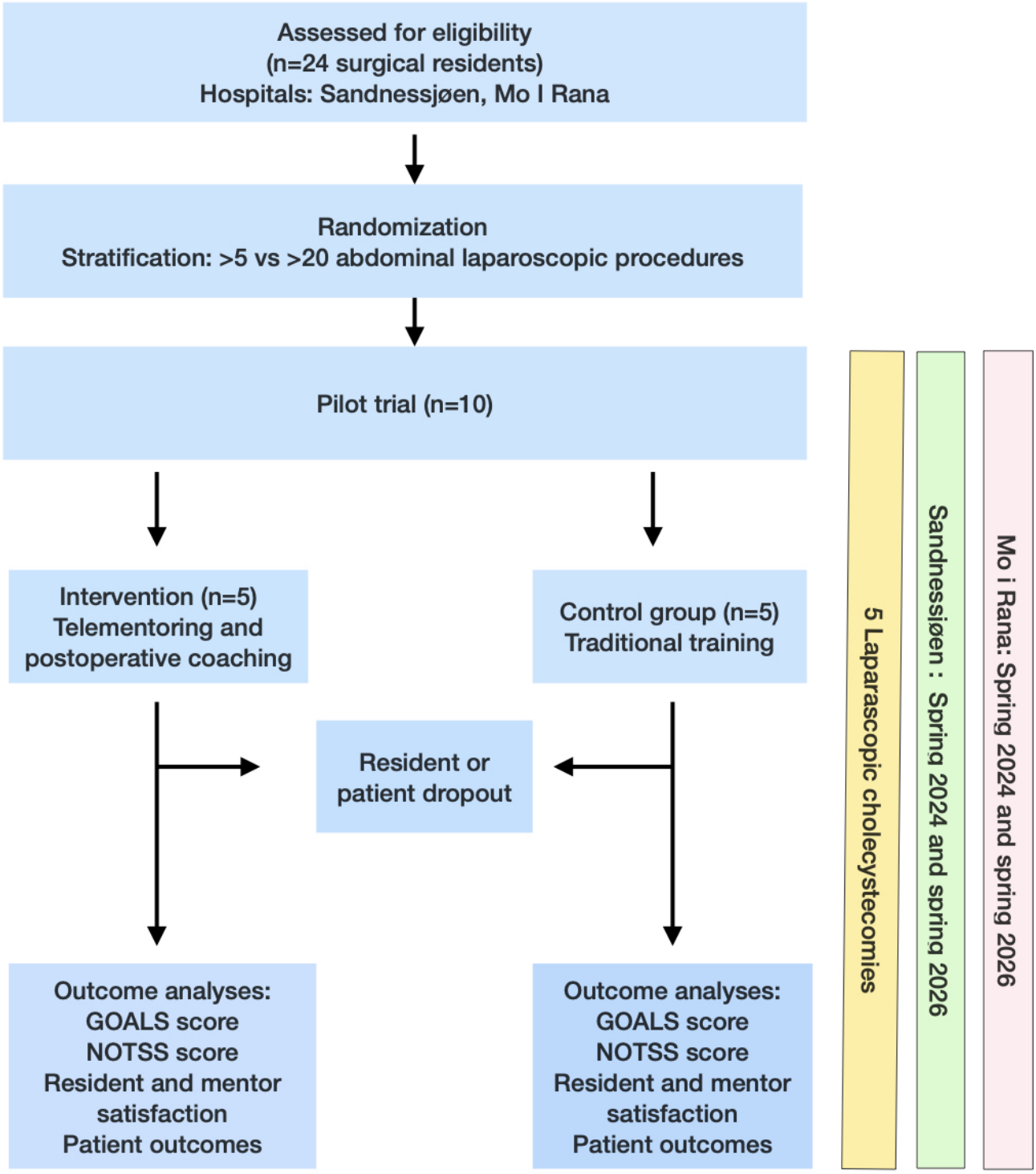

Blinding

Blinding the trainees is not feasible given the characteristics of the intervention. Nevertheless, the evaluators who assessed performance needed to be unaware of the assignment of individuals to specific groups. Given the nature of the intervention, unblinding is not applicable to participants. For outcome assessors, any unintentional unblinding would be documented. No emergency unblinding procedures are required because the intervention poses no direct safety risk.

Interventions

A skilled surgeon will serve as the camera-holding assistant for all procedures in the intervention group. They will intervene only at the time of unexpected events during surgery requiring emergency on-site assistance. Second, residents in the intervention group will receive postoperative coaching grounded in Lapco TT principles and the GROW model. In contrast, the control group will follow standard, unstructured guidance during the procedure. The mentors and mentees will sign a contract

XSL•FO
RenderX

accepting the educational setting and the prerequisites before participating in the trial. Table 2 summarizes the intervention and control conditions.

Table . Comparison of intervention and control groups.

Component	Telementoring group (intervention)	Traditional training group (control)
Mentor location	Remote via secure video link. May use a laptop or other mobile devices	Physically present in operating room
Verbal communication mode	Headphones with a mic during surgery for real-time mentor-mentee communication	Direct verbal communication; physical gestures
Visual guidance	Live stream of procedure. Possibility for static image annotation (telestration); verbal description of dissection planes	Direct pointing; hands-on demonstration capability
Mentor's view	Laparoscopic camera feed + external OR[a] camera.	Direct laparoscopic monitor view + peripheral vision of OR
Physical intervention	Not possible (local consultant available for emergencies)	Can take instruments to demonstrate; tactile guidance
Postoperative coaching	Mandatory structured GROW[b] model session (15 - 20 min)	Standard informal feedback (unstructured, variable).
Documentation	All sessions recorded; coaching documented	Standard operative notes only.

[a]OR: operating room.

[b]GROW: goals, reality, options, wrap-up.

Intervention Group (Telementoring)

The telementoring protocol comprises 3 main phases: preprocedural setup and coaching, intraoperative telementoring, and postoperative coaching.

Preprocedure Setup and Coaching

Before commencing the surgical procedure, a standardized setup phase of approximately 10 to 15 minutes is implemented. Key steps to ensure operational readiness include the following: the resident confirms connectivity to the telementoring system with the remote mentor; a bilateral communication check of the audio headset is performed, and the video feeds from both the laparoscopic and external OR cameras are verified by the telementor. This phase concludes with a brief preoperative discussion to review the surgical plan and anticipate potential challenges.

Intraoperative Telementoring

During the operation, the telementor provides real-time guidance by observing a high-definition video stream and maintaining continuous audio communication via a wireless headset. When anatomical clarification is required, the telementor activates the telestration mode. This process involves the following:

- The procedure is temporarily paused using a specific "stop word" command from the mentor.
- A static image is captured from the laparoscopic feed.
- The mentor annotates the image using lines, circles, or arrows to highlight relevant anatomical structures, dissection planes, or areas of concern.
- The resident confirms their understanding of the annotations before the procedure is resumed.

The Lapco communication protocol, an integral component of the Lapco standardized teaching framework, follows the "6 steps": stop → identify → explain → instruct → check → assess. The steps are defined as follows:

1. Stop. The telementor (trainer) initiates the sequence by pausing the procedure completely with the stop signal, often when the telementee (trainee) encounters difficulty or reaches a critical juncture, to gain the telementee's full attention and prevent errors.
2. Identify. The telementor asks the telementee to clearly identify the specific problem, concern, or relevant anatomical structure to ensure mutual understanding of the issue.
3. Explain. The telementor provides a brief explanation of the underlying principle or rationale behind the next required action, connecting theory to practice.
4. Instruct. Clear, concise instructions are provided to the telementee on how to perform the next step correctly.
5. Check. The telementee verbally confirms that they understand the instructions and are prepared to execute them safely by repeating the instructions provided in step 4.
6. Assess: After the telementee attempts the maneuver, the telementor provides an assessment and determines whether the procedure can proceed safely without further instructions or whether the on-site surgeon should assume control of the procedure. The latter should be avoided to maximize learning.

A local consultant surgeon is scrubbed for all procedures and assists as the camera-holding surgeon during the laparoscopic procedure but does not provide teaching input unless patient safety requires intervention.

XSL•FO
RenderX

Postoperative Coaching (10-15 Minutes Within 1 Hour of Procedure Completion):

A structured coaching session is conducted postoperatively, within 1 hour of procedure completion, and lasts approximately 10 to 15 minutes. This session uses the GROW model and is conducted via a videoconferencing platform.

The process begins with the resident completing a self-assessment of their performance, focusing on their initial goals and the actual surgical outcomes (goals and reality). Subsequently, the mentor provides constructive feedback, referencing specific procedural moments observed during the operation. The session concludes with the joint identification of specific action points for improvement (options and wrap-up). Finally, the key learning points generated during the session are formally documented for future reference and to support the continuity of training.

Control Group (Traditional Training)

The control group operates within the framework of conventional surgical training protocols, characterized by direct, in-person mentorship.

Intraoperative Mentoring

During the surgical procedure, an experienced consultant surgeon is physically scrubbed in alongside the resident. The primary mode of instruction is direct verbal guidance throughout the operation. Mentors retain discretion to physically interact during the case. They may guide the resident's hands, indicate anatomical structures on the surgical screen, or temporarily assume control of instruments to demonstrate a specific technique. While the same standardized procedural steps are typically followed, the mentor provides teaching and feedback at their discretion, in the absence of a formalized, structured teaching framework.

Postoperative Feedback

Postoperative feedback in this group consists of a standard informal debriefing, conducted solely at the mentor's discretion. This approach is defined by the absence of mandated structure, specific timing requirements, or formal documentation requirements. This methodology represents the current standard practice in many surgical training environments.

The Surgical Procedure

The LC procedure was standardized using a stepwise approach, incorporating the principles of creating the critical view of safety before transecting any structure in Calot triangle [28,29]. All included residents were instructed in intraoperative communication in accordance with Lapco and GROW principles. A step-by-step procedure was provided for both mentees and mentors, and adherence was required. A detailed description of the step-by-step surgical procedure is provided in Multimedia Appendix 3.

Outcomes

Primary Outcomes

The primary outcomes are the GOALS and NOTSS scores (Figures 1 and 2). The GOALS score is the only global rating scale developed and validated for assessing laparoscopic surgery. Five subscales are assessed: depth perception, bimanual dexterity, efficiency, tissue handling, and autonomy. Each subscale is scored from 1 to 5, yielding a maximum score of 25 when assessing the resident's laparoscopic technical skills. The NOTSS score assesses the operating surgeon's nontechnical surgical skills by evaluating defined parameters [30].

The NOTSS score will be derived from video recordings of interactions between the operating telementee and the rest of the OR team. A 3-member evaluation group of specialized OR nurses will be established to assess the NOTSS score for each procedure by reviewing video footage from the room camera.

The GOALS score will be assessed by video recording the laparoscopic camera footage of each procedure performed by the included registrars. A 3-member evaluation group of expert laparoscopic surgeons will be established to assess the GOALS score for each procedure by reviewing the laparoscopic video footage, including the verbal communication between the mentee and the mentor.

Both evaluation groups will be presented with video footage of each included mentee in a random order to obtain blinded evaluations of the GOALS and NOTSS scores.

Secondary Outcome Measurements

Satisfaction Scores

Residents in both groups completed a satisfaction questionnaire after each procedure, allowing them to self-assess their satisfaction with the learning outcomes.

The satisfaction survey used a 5-point Likert scale, with 1 indicating strong disagreement and 5 indicating strong agreement. A similar satisfaction survey was administered to the mentors in each group.

Patient-Related Outcomes and Complications

Intraoperative unfavorable incidents will be classified according to the Oslo Classification of Intraoperative Unfavorable Incidents [31].

- Grade 1: minor incidents with no consequences (eg, minor bleeding controlled immediately and small gallbladder perforation with minimal spillage)
- Grade 2: incidents requiring intervention but not affecting the procedure's outcome (eg, moderate bleeding requiring additional hemostatic measures and bile spillage requiring irrigation)
- Grade 3: incidents significantly affecting the procedure (eg, conversion to open surgery, common bile duct injury, and bowel injury).

This classification complements the Clavien-Dindo Classification, which captures postoperative complications [32]. By applying both systems, a comprehensive assessment of intraoperative and postoperative morbidity will be achieved. Interventions by the attending consultant surgeon and conversions to laparotomy will also be assessed.

XSL•FO
RenderX

Sample Size

This pilot RCT is designed to identify challenges and unforeseen pitfalls in conducting a proper RCT. Given the novelty of the topic and the resources required for an RCT, it is necessary to conduct a pilot study to assess the feasibility of a full-scale RCT.

The full RCT is designed as a superiority trial to demonstrate the superiority of telementored resident skill acquisition in the OR compared with standard training methods. Previous investigations have shown that the SD of the GOALS score is 2.5. We hypothesize that the intervention group will demonstrate significant clinical improvement in skills, as evidenced by a 3‐ to 5-point increase in the GOALS score relative to the control group. To achieve a statistical power of 0.8, a significance level (*P* value) of <.05, and to account for an expected dropout rate of 20%, 12 residents per group are required for an RCT.

This pilot study will include 5 residents in both the experimental and control arms (n=10). Residents will be stratified into 2 groups according to the number of prior LCs.

Data Collection and Management

Each video recording will be stored in the OR recording system. The video recordings will be edited by KB (corresponding author) into predefined segments of the surgical procedure, as specified in the stepwise procedure description. The edited video recordings will be annotated. These standardized, edited recordings will be anonymized and randomized for each resident and distributed to the evaluation group for assessment of GOALS and NOTSS scores.

The project secretary will provide each evaluation group with a Questback questionnaire to collect their assessments. The project secretary, who holds the anonymization and randomization key, will then collect and decode the assessments.

Data will be collected and managed using REDCap, hosted at the University Hospital of Northern Norway [33]. Data items (resident satisfaction and mentor satisfaction) will be collected immediately after each procedure. Given the pilot nature of this study, no formal steering committee or data safety monitoring board was established.

GOALS and NOTSS scores will be collected by giving the mentor group 1 month to assess the video records. Patient-related data will be collected retrospectively by reviewing the electronic health record. The timeline of data collection is shown in Figure 5.

XSL•FO
RenderX

Figure 5. Trial data collection timeline. GOALS: Global Operative Assessment of Laparoscopic Skills; NOTSS: Non-Technical Skills for Surgeons.

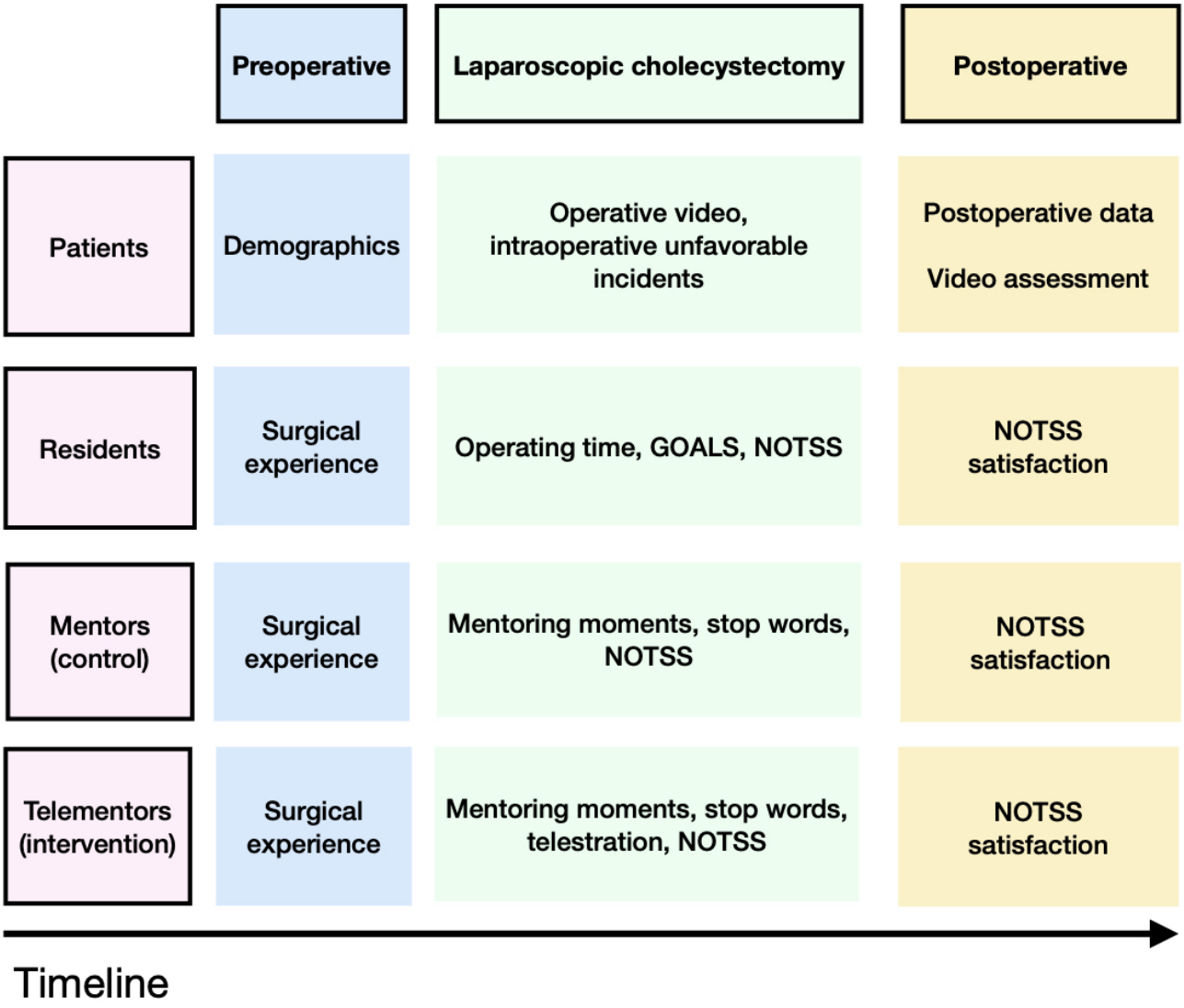

Statistical Analysis

The primary analysis will compare mean GOALS scores between the intervention and control groups. An independent-samples *t* test will be used for this comparison, provided that the assumption of normality is met; if normality is violated, the nonparametric Mann-Whitney *U* test will be applied. Effect sizes will be calculated using Cohen *d* to quantify the magnitude of the difference between groups and to inform sample size estimation for a future definitive RCT.

Secondary analyses will include comparisons of NOTSS scores between the intervention and control groups using the same statistical approach as that used for the primary outcome. Within-group improvement in technical performance across the training period will be assessed by comparing GOALS scores from the first to the fifth procedure using paired *t* tests. Participant satisfaction scores will be compared between groups using the Mann-Whitney *U* test, as these data are collected on an ordinal Likert scale.

Multivariate analyses will be conducted using linear regression models to examine the association between group allocation and GOALS scores while adjusting for potential confounding factors. Group allocation (intervention vs control) will be included as the primary predictor. Covariates will be selected a priori based on theoretical relevance and will include baseline surgical experience (defined as completion of 5 - 10 vs 10 - 20 prior LCs), baseline GOALS score from the first procedure, year of residency training, study site (Sandnessjøen vs Mo i Rana), and indicators of procedure complexity, including intraoperative findings and operative time.

Exploratory analyses will be performed to further characterize performance development and relationships between outcomes. Learning curves across the 5 procedures will be evaluated by examining trajectories of GOALS scores using repeated-measures ANOVA or mixed effects models, depending on the data structure and model assumptions. Subgroup analyses stratified by baseline experience level will be undertaken, and correlations between GOALS and NOTSS scores will be assessed to explore the relationship between technical and nontechnical performance.

As this study is a pilot trial, feasibility outcomes will be summarized descriptively. These outcomes will include the recruitment rate, expressed as the number of participants enrolled per month; balance between randomized groups;

XSL•FO
RenderX

adherence to the study protocol (including completion of all 5 procedures and scheduled coaching sessions); the frequency and nature of technical failures related to the telementoring system; and withdrawal rates, along with reasons for withdrawal. Participants may withdraw from the study at any time. In cases where a resident discontinues the intervention, outcome data will be collected, where feasible, and analyzed using the intention-to-treat principle.

Data Presentation

Continuous variables will be presented as mean (SD) for normally distributed data and as median (IQR) for skewed distributions. Categorical variables will be presented as frequencies and percentages. A *P* value <.05 will be considered statistically significant. Given the pilot nature, emphasis will be placed on effect sizes and CIs rather than *P* values alone. Descriptive statistics and data analysis will be performed using Microsoft Excel - Office 365 and IBM SPSS Statistics software (version 31.0; IBM Corp).

Data Monitoring

Data and study monitoring will be performed once 10 residents have been recruited to the pilot trial. Given the pilot nature and low-risk intervention, a formal data monitoring committee has not been established. The principal investigator will review safety data after every 5 residents. The principal investigator, the telementors, and the on-site consultant surgeon will conduct monitoring in brief meetings (telephone calls or Teams meetings) prior to the inclusion of each resident. Technical issues related to telementoring, the mentoring framework, and coaching will be addressed. Given the study's novelty, adjustments to the telementoring framework may be necessary.

Given the pilot study's intensive follow-up (5 procedures in 3 d), minimal missing data are expected. For the primary analysis, available-case analysis will be used. Descriptive patterns of missing data will be reported to inform the definitive RCT.

Trial monitoring will be conducted by the principal investigator. This includes a quarterly review of recruitment rates, protocol deviations, technical system performance, and adverse events. Given the single-site coordination and pilot scale, on-site monitoring visits are not planned.

Ethical Considerations

The study was approved by the ethical committee (REK HELSE NORD 32592) and the data protection officer at Nordlandssykehuset Hospital Trust, Bodø, Norway.

All collected data were anonymized prior to analysis to ensure participant (patients and residents) confidentiality. Video recordings were deidentified and securely stored on password-protected institutional servers with access restricted to authorized personnel. Informed consent was obtained from all participants, including consent for recording and data use. The study complies with applicable data protection regulations and was approved by the relevant ethics committee and data protection authority as mentioned above.

Surgical residents will be informed that their participation is entirely optional, that there will be no repercussions if they choose to withdraw, and that they may withdraw without providing a reason. Upon signing in, all surgical residents must sign an electronic permission document and a mentee-mentor contract (intervention arm). Protocol adherence will be monitored by reviewing recorded sessions.

This educational intervention does not pose direct risks to participants beyond those inherent in standard surgical training. No specific posttrial care provisions are required. Residents receive no compensation for participation.

Patients will be asked to provide informed consent. The study group does not anticipate any ethical implications for patient participation, as the telementoring (intervention arm) will provide increased surgical expertise in the OR.

Any protocol modifications will be submitted to REK for approval before implementation. Amendments will be documented with date and rationale, communicated to investigators and the trial registry, and reported in the final publication.

Patient and Public Involvement

This pilot study was designed without formal patient or public involvement in the research design, conduct, or reporting plans. Patients are included solely as surgical cases, with the study focusing on surgical trainee education rather than on patient-centered outcomes. Surgical residents, the primary study participants, were consulted informally during protocol development to assess the acceptability and feasibility of the telementoring intervention. For the planned definitive RCT, we will consider establishing a stakeholder advisory group, including patient representatives and surgical trainees, to inform outcome selection, dissemination strategies, and study conduct.

Dissemination Plan

The results of this pilot study will be submitted for publication in an international peer-reviewed journal, regardless of the outcome (positive, negative, or inconclusive). Findings will be presented at relevant surgical education and telemedicine conferences. National and international conferences in 2026, focusing on the intersection of surgical education and telemedicine, which we are looking forward to participating in, include the following: (1) The Norwegian Surgeons' Annual Meeting, October 2026. Oslo, Norway; (2) AMEE Conference, August 2026, Vienna, Austria; and (3) ASiT Annual Conference, 2027.

The trial protocol and results will remain publicly accessible via ClinicalTrials.gov. If successful, we will engage with Norwegian surgical societies and health care policymakers to discuss implementation implications for rural surgical training programs.

Results

Preparations Before Onset of the Study

Before the trial began, all telementors and on-site consulting surgeons participated in a Lapco TT course. The configuration of the telementoring system was finalized in December 2023.

XSL•FO
RenderX

All residents, consulting surgeons, and telementors were instructed on the setup and use of the telementoring system.

Current Study Status and Timeline

Completed Activities

The project initiated its operations by configuring the telementoring system across all study sites by December 2023. Following this, all mentors received Lapco TT training in February 2024. Ethics approval was subsequently secured from the Regional Committee for Medical and Health Research Ethics (REK nord 32592) on October 23, 2023, and the most recent update of the trial was registered in ClinicalTrials.gov (NCT06421584) on December 8, 2024.

Recruitment Status

Recruitment commenced in the Spring of 2024. Initially, the MedPresence platform was used for the first 5 resident inclusions. Due to changes in the Olympus portfolio of telementoring systems and global partnership agreement with Proximie, a platform transition to Proximie occurred in July/August 202. This included training telementors on the new system in August 2025. By December 2025, 3 residents had been included using the new platform. The remaining recruitment is scheduled to start in February 2026.

Anticipated Timeline of the Pilot

The project anticipates completing all recruitment by May 2026, with data collection concluding in September 2026. The data analysis phase is slated for November 2026, and the manuscript submission for the final results is expected in January 2027.

Discussion

Principal Findings and Interpretation

We acknowledge that telementoring is often positioned as an alternative when on-site expertise is unavailable, and the intuitive assumption is that physical presence should be superior. However, our superiority hypothesis is based on several theoretical considerations.

First, telementoring may enhance the quality of verbal instruction. When the mentor cannot physically intervene, communication must be more precise and explicit, potentially improving the teaching dialog. The structured communication required by telementoring (stop-identify-explain-instruct-check-judge) enforces a pedagogical framework that may be applied less rigorously when mentors can simply take over the instruments.

Second, the mandatory structured postoperative coaching in the telementoring arm (using the GROW model) represents an educational enhancement not typically provided in traditional training. This combination of real-time guidance and systematic reflection may amplify learning beyond that achieved through informal feedback.

Third, the telementor maintains the identical visual perspective as the operating resident without competing for the operative field or camera control, potentially allowing for more focused observational teaching.

We acknowledge, however, that these are theoretical advantages that may not materialize in practice. The pilot study will test whether any educational benefit exists and, equally importantly, will provide data to refine or reject this hypothesis before investing in a larger definitive trial. If the pilot demonstrates no advantage—or disadvantage—for telementoring compared to on-site mentoring, this finding would itself be valuable, suggesting that telementoring can serve as an acceptable equivalent (noninferior) alternative when on-site expertise is unavailable.

Anticipated Outcomes

This pilot RCT is designed to determine whether ST combined with structured postoperative coaching can enhance laparoscopic skill acquisition compared with traditional on-site mentoring. We anticipate that the structured nature of telementoring—requiring explicit verbal communication and systematic feedback—may enhance learning through increased deliberate practice and reflection. If our hypothesis is confirmed, we expect to observe a 3- to 5-point improvement in GOALS scores in the telementoring group, which would represent a clinically meaningful enhancement in surgical proficiency.

Equally important, this pilot will identify practical challenges in implementing telementoring for surgical education, including technological barriers, workflow integration issues, and participant acceptance factors that must be addressed before a definitive trial.

Comparison to Prior Work

Previous telementoring studies have primarily focused on demonstrating technical feasibility and safety rather than educational outcomes [13,14]. The SURGTEACH (Surgical Telementoring in Acquisition of Surgical Skills of Laparoscopic Cholecystectomy) trial extends this work using validated educational assessment tools (GOALS and NOTSS) and a randomized design. The author group of this pilot RCT has previously published a prospective clinical trial evaluating the educational role of ST in minimally invasive surgery [34]. Unlike observational studies that compare outcomes before and after telementoring implementation, this parallel-group design with a concurrent control arm will provide more rigorous evidence of any educational benefit.

The integration of the LapCo TT framework distinguishes this study from prior telementoring research, which often lacked standardized communication protocols. By training all mentors—both telementors and on-site mentors—in identical communication principles, we minimize the confounding effects of differing teaching styles and isolate the effects of the mentoring modality itself.

Strengths and Limitations

Strengths

This study represents the first randomized comparison of telementoring versus traditional training with a primary focus on educational outcomes. It employs validated, objective assessment tools, including GOALS and NOTSS, to ensure robust measurement of both technical and nontechnical skills. Outcome assessment is blinded through standardized video

XSL•FO
RenderX

review, reducing the risk of assessor bias. Both study arms use a standardized communication framework, which minimizes variability in mentor–trainee interaction. In addition, postoperative coaching is delivered in a structured manner using the evidence-based GROW model. Randomization is stratified by trainee experience level, further strengthening internal validity and group comparability.

Limitations

As a pilot study, the relatively small sample size limits statistical power and precludes definitive conclusions regarding effectiveness. Blinding of participants is not feasible given the nature of the intervention, which introduces a potential risk of performance bias. The single-country setting may restrict the generalizability of the findings to other health care systems and training environments. The superiority hypothesis is based on an assumed benefit of structured telementoring that may not be fully realized in the final outcome assessment. Furthermore, a change in the technological platform during recruitment, from MedPresence to Proximie, may introduce heterogeneity in the intervention. The control group receives standard care, which may vary between mentors despite Lapco TT training, potentially increasing intermentor variability. Finally, assessing short-term outcomes limited to the first 5 procedures may not adequately capture long-term skill acquisition or retention.

Future Directions

The results of this pilot study will directly inform the design of a definitive multicenter RCT. Specific aspects that may require modification include recruitment strategies, randomization procedures, outcome timing, and telementoring protocols. If substantial technical or logistical barriers are identified, alternative technological solutions will be explored before the definitive trial.

Beyond the immediate next steps, successful demonstration of telementoring efficacy could enable expansion to other laparoscopic procedures, support the development of national telementoring networks for rural surgical training, and inform educational policy for surgical residency programs in geographically challenging regions.

Conclusion

The SURGTEACH trial is the first known randomized pilot trial to evaluate ST as an educational method. Globally and domestically, there are constraints on the number of surgeons available and strain on the surgical education system. The Norwegian health care system requires support to train sufficient surgeons, primarily due to geographical constraints and limitations in the education system. Therefore, there is a need for a fundamental change in surgical education, and ST has the potential to overcome these obstacles. This study has the potential to provide substantial evidence to enhance surgical education, particularly in rural health care facilities.

Acknowledgments

We thank the consultant surgeons participating as mentors and telementors for their contributions to the study. We also thank the Department of Information and e-Health at Nordland Hospital Trust for their contribution in the OR setup. Generative AI (Leonardo AI) was used to correct the grammar. The scientific content of the paper is the sole responsibility of the authors.

Funding

This study is an investigator-initiated trial sponsored by Nordland Hospital Trust. Financial support was provided by Helgelandsskehuset HF. Olympus Norway contributed by installing and configuring the telementoring system at participating study sites, and Proximie provided the cloud-based platform used during the final phase of recruitment. The funders and supporting organizations had no role in the study design, collection, management, analysis, interpretation of data, writing of the manuscript, or decision to submit the manuscript for publication. The authors retained full independence and responsibility for all decisions.

Data Availability

The datasets generated during this study will be available from the corresponding author upon reasonable request following the publication of the primary results. Anonymized data may be deposited in a suitable repository, subject to ethical approval and data protection regulations. The full study protocol is available on ClinicalTrials.gov. The statistical analysis plan will be finalized before database lock and made available upon request. Data collection instruments will be made available as supplementary materials.

Authors' Contributions

Conceptualization: KB, KMA
Funding acquisition: KB, KMA
Formal analysis: LOT
Investigation: KB, OS
Methodology: KB, LOT, KMA
Project administration: KB
Resources: BvGE, GH

XSL•FO
RenderX

Supervision: KMA, BE
Writing—original draft: KB
Writing—review and editing: all authors
The principal investigator (KB) is responsible for the overall conduct of the trial.

Conflicts of Interest

None declared.

Multimedia Appendix 1
(a) Six-step Lapco-based trainer intervention of intraoperative teaching when a mentee is failing to progress. Adopted for telementoring sessions in the intervention arm of the study. (b) A structured, standardized framework for perioperative setup, dialogue, and closure of telementoring sessions, including performance-enhancing instructions and an opportunity for self-reflection for the telementee. The session is complete after the telementee formulates a take-home message in agreement with the telementor.
[PDF File, 152 KB - resprot_v15i1e73159_app1.pdf]

Multimedia Appendix 2
The adoption of the GROW model for telementoring procedures in the intervention arm of the study.
[PDF File, 126 KB - resprot_v15i1e73159_app2.pdf]

Multimedia Appendix 3
The standardized, detailed stepwise technique applied for laparoscopic cholecystectomy in both the intervention and control arms of the study.
[PDF File, 350 KB - resprot_v15i1e73159_app3.pdf]

Checklist 1
SPIRIT checklist 2025 for the SURGTEACH protocol. Note that the checklist is based on a replica of the manuscript presented in the supplementary documents, including line numbers in the document margins.
[PDF File, 217 KB - resprot_v15i1e73159_app4.pdf]

References

1. Vinden C, Malthaner R, McGee J, et al. Teaching surgery takes time: the impact of surgical education on time in the operating room. Can J Surg 2016 Apr;59(2):87-92. [doi: 10.1503/cjs.017515] [Medline: 27007088]
2. Specialist surgical workforce (per 100,000 population) - Norway. World Bank Group. 2025. URL: https://data.worldbank.org/indicator/SH.MED.SAOP.P5?locations=NO [accessed 2026-03-28]
3. Advancing the surgical workforce: 2023 UK surgical workforce census report. : Royal College of Surgeons of England; 2023 URL: https://www.rcseng.ac.uk/-/media/Files/RCS/Standards-and-research/RCS-Advancing-The-Surgical-Workforce-Digital.pdf [accessed 2026-03-19]
4. The future of surgical traning. : American College of Surgeons; 2023 URL: https://www.facs.org/media/0gtbbj2y/august-2023-acs-bulletin.pdf [accessed 2026-03-19]
5. Skjold-Ødegaard B. More simulation in specialty training. Tidsskr Nor Laegeforen 2024 May 14;144(6). [doi: 10.4045/tidsskr.24.0079] [Medline: 38747670]
6. Miller-Hammond K, Anderson D. Barriers to surgical health care access in rural communities. Am Surg 2025 May;91(5):681-684. [doi: 10.1177/00031348251323712] [Medline: 40012103]
7. Saunes IS, Karanikolos M, Sagan A. Norway: health system review. : WHO; 2020 URL: https://iris.who.int/server/api/core/bitstreams/75a2ffa5-e225-47d3-9c77-ee19529f6c3c/content [accessed 2026-03-19]
8. Patil M, Gharde P, Reddy K, Nayak K. Comparative analysis of laparoscopic versus open procedures in specific general surgical interventions. Cureus 2024 Feb;16(2):e54433. [doi: 10.7759/cureus.54433] [Medline: 38510915]
9. Bilgic E, Turkdogan S, Watanabe Y, et al. Effectiveness of telementoring in surgery compared with on-site mentoring: a systematic review. Surg Innov 2017 Aug;24(4):379-385. [doi: 10.1177/1553350617708725] [Medline: 28494684]
10. Augestad KM, Bellika JG, Budrionis A, et al. Surgical telementoring in knowledge translation--clinical outcomes and educational benefits: a comprehensive review. Surg Innov 2013 Jun;20(3):273-281. [doi: 10.1177/1553350612465793] [Medline: 23117447]
11. Wood D. No surgeon should operate alone: how telementoring could change operations. Telemed J E Health 2011 Apr;17(3):150-152. [doi: 10.1089/tmj.2011.9986] [Medline: 21500973]
12. Butt KA, Augestad KM. Educational value of surgical telementoring. J Surg Oncol 2021 Aug;124(2):231-240. [doi: 10.1002/jso.26524] [Medline: 34245572]
13. Erridge S, Yeung DKT, Patel HRH, Purkayastha S. Telementoring of surgeons: a systematic review. Surg Innov 2019 Feb;26(1):95-111. [doi: 10.1177/1553350618813250] [Medline: 30465477]

XSL•FO
RenderX

14. Fernandes RD, Ghasroddashti A, Sorefan-Mangou F, et al. Educational effectiveness of telementoring as a continuing professional development intervention for surgeons in practice: a systematic review. Ann Surg Open 2023 Dec;4(4):e341. [doi: 10.1097/AS9.0000000000000341] [Medline: 38144497]
15. Moore DE, Green JS, Gallis HA. Achieving desired results and improved outcomes: integrating planning and assessment throughout learning activities. J Contin Educ Health Prof 2009;29(1):1-15. [doi: 10.1002/chp.20001] [Medline: 19288562]
16. Chepkoech M, Malila B, Mwangama J. Telementoring for surgical training in low-resource settings: a systematic review of current systems and the emerging role of 5G, AI, and XR. J Robot Surg 2025 Aug 28;19(1):525. [doi: 10.1007/s11701-025-02703-9] [Medline: 40877481]
17. Mackenzie H, Cuming T, Miskovic D, et al. Design, delivery, and validation of a trainer curriculum for the national laparoscopic colorectal training program in England. Ann Surg 2015 Jan;261(1):149-156. [doi: 10.1097/SLA.0000000000000437] [Medline: 24374538]
18. Wyles SM, Miskovic D, Ni Z, et al. Development and implementation of the Structured Training Trainer Assessment Report (STTAR) in the English national training programme for laparoscopic colorectal surgery. Surg Endosc 2016 Mar;30(3):993-1003. [doi: 10.1007/s00464-015-4281-z] [Medline: 26104793]
19. Case studies: how have we supported our partners: Lapconor–the Norwegian experience. Lapco International. 2024. URL: http://web.archive.org/web/20251208193641/https://www.lapco-international.com/consultancy/case-studies/ [accessed 2026-08-03]
20. Sjo OH, Andersen BT, Gaupset R. LAPCONOR - an educational framework for standardized laparoscopic treatment of patients with colon and rectal cancer. Kirurgen. 2016. URL: https://kirurgen.no/fagstoff/lapconor-et-utdanningsprogram-for-standardisering-av-laparoskopisk-behandling-av-i-pasienter-med-kreft-i-tykk-og-endetarm [accessed 2026-03-19]
21. Julien B, Greer D, Bull N, Zevin B, O'Neill C, Gilligan C. Characteristics of surgical coaching interventions: a systematic review. J Surg Educ 2025 Aug;82(8):103543. [doi: 10.1016/j.jsurg.2025.103543] [Medline: 40414177]
22. Whitmore J. Coaching for Performance: Growing People, Performance and Purpose: Nicholas Brealey Publishing; 2002. URL: https://books.google.co.in/books/about/Coaching_for_Performance.html?id=hTnuiKEE0vkC&redir_esc=y [accessed 2026-03-19]
23. van Zwieten T, Okkema S, van Det M, Pereboom I, Veeger N, Pierie JP. Assessment methods in laparoscopic colorectal surgery: a systematic review of available instruments. Int J Colorectal Dis 2023 Apr 19;38(1):105. [doi: 10.1007/s00384-023-04395-9] [Medline: 37074421]
24. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet 2009 Jul 4;374(9683):86-89. [doi: 10.1016/S0140-6736(09)60329-9] [Medline: 19525005]
25. Corley A, Marsh N, Keogh S. Methodological progress note: pilot randomized controlled trials. J Hosp Med 2024 Sep;19(9):821-826. [doi: 10.1002/jhm.13376] [Medline: 38655805]
26. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. The Lancet 2025 May;405(10491):e19-e27. [doi: 10.1016/S0140-6736(25)00770-6] [Medline: 40294521]
27. Vassiliou MC, Feldman LS, Andrew CG, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. Am J Surg 2005 Jul;190(1):107-113. [doi: 10.1016/j.amjsurg.2005.04.004] [Medline: 15972181]
28. Strasberg SM, Brunt LM. The critical view of safety: why it is not the only method of ductal identification within the standard of care in laparoscopic cholecystectomy. Ann Surg 2017 Mar;265(3):464-465. [doi: 10.1097/SLA.0000000000002054] [Medline: 27763898]
29. Manatakis DK, Antonopoulou MI, Tasis N, et al. Critical view of safety in laparoscopic cholecystectomy: a systematic review of current evidence and future perspectives. World J Surg 2023 Mar;47(3):640-648. [doi: 10.1007/s00268-022-06842-0] [Medline: 36474120]
30. Yule S, Flin R, Paterson‐Brown S, Maran N, Rowley D. Development of a rating system for surgeons' non‐technical skills. Med Educ 2006 Nov;40(11):1098-1104. [doi: 10.1111/j.1365-2929.2006.02610.x] [Medline: 17054619]
31. Kazaryan AM, Røsok BI, Edwin B. Morbidity assessment in surgery: refinement proposal based on a concept of perioperative adverse events. ISRN Surg 2013;2013:625093. [doi: 10.1155/2013/625093] [Medline: 23762627]
32. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 2004 Aug;240(2):205-213. [doi: 10.1097/01.sla.0000133083.54934.ae] [Medline: 15273542]
33. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
34. Butt K, Jonas OS, Trichet LO, et al. Surgical telementoring as teaching tool in the operating room: trans-Nordic IDEAL stage 2a telementored series of a robotic ventral mesh rectopexy learning curve. Br J Surg 2024 Jun 12;111(6):znae123. [doi: 10.1093/bjs/znae123] [Medline: 38864757]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials

XSL•FO
RenderX

GOALS: Global Operative Assessment of Laparoscopic Skills
GROW: goals, reality, options, wrap-up
Lapco: laparoscopic colorectal
LapcoNor: laparoscopic colorectal in Norwegian adoption
LC: laparoscopic cholecystectomy
NOTSS: Non-Technical Skills for Surgeons
RCT: randomized controlled trial
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
ST: surgical telementoring
SURGTEACH: Surgical Telementoring in Acquisition of Surgical Skills of Laparoscopic Cholecystectomy
TT: Training the Trainers

Edited by A Schwartz; submitted 19.Mar.2025; peer-reviewed by A Kazaryan, Å Fretland; revised version received 14.Jan.2026; accepted 30.Jan.2026; published 17.Apr.2026.

Please cite as:
Butt K, Trichet LO, Hartvigsen G, Sjo O, Edwin BVG, Augestad KM
Evaluating the Role of Surgical Telementoring in the Acquisition of Surgical Skills in Laparoscopic Cholecystectomy: Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2026;15:e73159
URL: https://www.researchprotocols.org/2026/1/e73159
doi:10.2196/73159

© Khayam Butt, Laurent Olivier Trichet, Gunnar Hartvigsen, Ole Sjo, Bjørn von Gohren Edwin, Knut-Magne Augestad. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Digital Attention Bias in Cancer Survivors Intervention for Adolescent and Young Adult Cancer Survivors: Protocol for a Pilot Randomized Controlled Trial

Nancy Lau[1,2], PhD; Chuan Zhou[2,3], PhD; Shannon JH Hong[4], BS; Homer Aalfs[2], BS; Shannon Higgins[2], BS; Isabel Badillo[2], BA; Elizabeth McCauley[1,2], PhD; Tyler Ketterl[3,5], MD; Eric J Chow[3,6], MD, MPH; Jesse R Fann[1,6], MD, MPH; Lauren C Heathcote[7], PhD; Tonya M Palermo[2,8], PhD

[1]Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, WA, United States

[2]Center for Child Health, Behavior and Development, Seattle Children's Research Institute, Seattle, WA, United States

[3]Department of Pediatrics, University of Washington School of Medicine, Seattle, WA, United States

[4]Department of Psychology, University of California, Berkeley, Berkeley, CA, United States

[5]Department of Internal Medicine, University of Washington School of Medicine, Seattle, WA, United States

[6]Clinical Research Division, Fred Hutchinson Cancer Center, Seattle, WA, United States

[7]Institute of Psychiatry, Psychology, and Neuroscience, Department of Psychology, King's College London, London, United Kingdom

[8]Department of Anesthesiology & Pain Medicine, University of Washington School of Medicine, Seattle, WA, United States

Corresponding Author:
Nancy Lau, PhD
Department of Psychiatry and Behavioral Sciences
University of Washington School of Medicine
1920 Terry Ave
Seattle, WA, 98101
United States
Phone: 1 2068848238
Email: nancy.lau@seattlechildrens.org

Abstract

Background: Adolescent and young adult cancer survivors face a high burden of psychological late effects, with cancer-related anxiety being a prevalent mental health concern. Despite the significant need for care, more than half of adolescent and young adult cancer survivors, who require psychosocial services, remain untreated. Digital health interventions offer a promising solution to bridge this care gap. Attention bias modification (ABM) is an evidence-based digital intervention for anxiety disorders. This intervention targets automatic and unconscious negative attention biases and retrains attention away from threat and toward neutral or positive stimuli. Recent research has successfully adapted ABM interventions for cancer survivors. However, ABM has not yet been adapted or tested for adolescent and young adult cancer survivors.

Objective: This protocol describes a pilot randomized controlled trial designed to evaluate a novel digital anxiety intervention, the Attention Bias in Cancer survivors (ABCs) intervention, for adolescent and young adult cancer survivors.

Methods: This is a single-site, 2-arm, pilot randomized controlled trial enrolling 60 cancer survivors aged 15 to 29 years. Participants will be randomized 1:1 to the ABCs intervention or a sham control condition. The ABCs intervention combines an ABM mobile intervention with daily gratitude and savoring SMS text messages (positive psychology prompts to savor positive emotions) over a 4-week period. The sham condition consists of sham ABM (showing the same cancer-related word stimuli as the active intervention condition but without attention retraining) and daily mood monitoring SMS text messages (prompts to report on current mood and stress levels). The primary objectives are to evaluate intervention feasibility (defined as ≥50% enrollment and ≥70% retention) and acceptability (defined by cutoff scores on the Client Satisfaction Questionnaire and System Usability Scale). Secondary exploratory outcomes include patient-reported measures of attention bias, anxiety, fear of recurrence, pain, resilience, and other psychosocial outcomes.

Results: This study was funded in August 2023, and study recruitment began in November 2024. We have completed data collection as of February 2026. We anticipate that data analyses will be completed by September 2026. Manuscript preparation and submission are anticipated for December 2026.

XSL•FO
RenderX

Conclusions: This pilot trial examines the feasibility and acceptability of a digital positive psychological intervention targeting anxiety in adolescent and young adult cancer survivors. Exploratory outcomes will inform sample size calculations for a future-powered multisite clinical trial. The ABCs intervention may provide scalable and accessible evidence-based psychosocial care and improve health outcomes.

Trial Registration: ClinicalTrials.gov NCT06682039; https://clinicaltrials.gov/study/NCT06682039

International Registered Report Identifier (IRRID): DERR1-10.2196/82665

(JMIR Res Protoc 2026;15:e82665) doi:10.2196/82665

KEYWORDS

adolescent; young adult; cancer survivors; anxiety; attention bias; psychosocial impact of illness; randomized controlled trial; digital health; mobile apps

Introduction

Background

Nearly 2 million adolescent and young adult cancer survivors reside in the United States [1]. In this study, adolescent and young adult cancer survivors are defined by their current developmental age (15-29 years) and include those diagnosed during childhood and adolescence and treated within pediatric cancer centers. This inclusive definition aligns with the Children's Oncology Group framework for cancer survivorship care, which emphasizes that the psychosocial and medical needs of cancer survivors are based on their current developmental stage and navigation of transition-to-adulthood milestones [2]. Cancer profoundly disrupts normal developmental experiences and life transitions [3,4]. Consequently, adolescent and young adult cancer survivors are at elevated risk for long-lasting negative psychological late effects, with poorer health-related quality of life and psychosocial outcomes than their younger pediatric and adult counterparts [5]. Adolescent and young adult cancer survivors commonly struggle with reintegration into daily life after treatment; miss crucial social, educational, and vocational milestones; face disrupted peer and family relationships; navigate financial toxicity; and experience a forestalled transition to independence. All of these issues can be further exacerbated by discontinuation of comprehensive support from their oncology care team [6]. A seminal longitudinal population-based study (the AYA Health Outcomes and Patient Experience study) showed that more than half of adolescent and young adult cancer survivors, who require psychosocial services, remain untreated [7]. Barriers to psychosocial care include inadequate end-of-treatment transition care planning between pediatric and adult cancer centers and primary care settings, insufficient psychosocial screening and referrals to mental health providers, a shortage of psychosocial clinicians, and persistent mental health stigma [8,9].

Cancer-related anxiety is the most prevalent mental health problem, impacting one-third of adolescent and young adult cancer survivors [10-12]. Muzzin et al [13] aptly characterized this as a "living-dying" experience in which "a person never really 'gets over' cancer: it is a sword of Damocles that continues to hang over the individual and [their] family for the rest of the person's life." In addition to general anxiety, up to 85% of adolescent and young adults experience cancer-specific anxiety, most notably scan-associated anxiety and fear of cancer recurrence [11]. Cancer-related anxiety is associated with long-term negative outcomes such as poor quality of life, depression, distress, posttraumatic stress, substance use, sleep problems, fatigue, and pain [14]. Therefore, the identification and management of anxiety are important targets for follow-up care of adolescent and young adult cancer survivors.

Adopting a developmental science framework, "meeting teens where they are," necessitates leveraging technology and a paradigm shift in traditional treatment norms [15]. Increasingly, mental health delivery models are being deployed in asynchronous mobile health formats. For adolescent and young adults, the internet serves as a primary platform for physical health and mental health–related information seeking and communication [16,17]. Adolescent and young adults prefer digital self-help interventions and may be reluctant to seek traditional forms of psychotherapy due to social stigma and discomfort discussing personal problems [16,18-20]. SMS text messaging interventions may be particularly appealing to younger generations who are accustomed to interacting on smartphones and using the internet [18,19,21,22].

Positive psychological interventions have demonstrated improvements in psychosocial outcomes for adolescent and young adult cancer survivors [23]. Such interventions, designed to increase positive emotions, have beneficial long-term impacts on emotion regulation, mental health, mindset broadening, coping resources, and functional outcomes [24]. Interventions that increase the frequency and duration of positive emotions through practices such as savoring (defined as the process of intentionally bringing focused attention to past, present, and future positive experiences, such as reminiscing about a happy memory or anticipating an upcoming vacation) and gratitude (defined as the process of appreciating the good in one's life and ascribing their internal and external sources, such as writing a thank you note to someone who helped you) have been shown to decrease stress, promote psychological well-being, and improve health-related quality of life in the face of adversity (eg, a cancer diagnosis) [25-28]. Background literature suggests that positive psychological interventions involving gratitude and savoring activities increase treatment adherence [29,30] and have demonstrated efficacy and effectiveness of these interventions in systematic reviews [31-34]. Although few digital psychosocial interventions for adolescent and young adult cancer survivors currently exist, this is a burgeoning area of intervention research that is highly appealing and appropriate for adolescent and young adults [35-38]. To our knowledge, there are no existing adolescent and young adult cancer

XSL•FO
RenderX

survivor–specific digital interventions that are designed to target anxiety, which is the most prevalent mental health symptom impacting adolescent and young adults.

Across pediatric and adult clinical populations alike, attention bias modification (ABM) is a widely used digital intervention for anxiety that has been adapted for a wide range of clinical populations and disease groups [39-41]. Systematic maladaptive patterns of information processing play a causal role in anxiety pathology [42,43]. Cognitive models of anxiety have consistently shown that clinically anxious individuals demonstrate an attention bias, meaning that they selectively attend to and have difficulty disengaging from negative information. Similarly, a systematic review has shown that cancer survivors demonstrate a negative attention bias [44]. For example, an individual with social anxiety may selectively attend to threatening stimuli in their environment, such as threatening facial expressions (eg, angry, judgmental, or disapproving), which can lead to sustained and increased anxiety and the avoidance of social situations [39,45-47].

ABM is a brief, self-guided digital intervention in which individuals complete repetitive association reaction-time tasks. These tasks target automatic and unconscious negative attention biases to retrain attention away from perceived threat and toward neutral or positive stimuli (ie, a modified dot-probe task) [48]. For example, ABM for social anxiety may involve a series of image pairs on a computer screen—one of a neutral or positive face and one of a threatening face. By rapidly and repeatedly training anxious individuals to disengage their attention from the threatening face and direct it toward the neutral or positive face, ABM implicitly modifies their negative attention bias.

ABM delivered digitally for the treatment of anxiety disorders (including social anxiety, generalized anxiety, and posttraumatic stress) has demonstrated moderate treatment effects [49-52]. The effectiveness of ABM is comparable to that of the gold-standard evidence-based treatment for anxiety (ie, cognitive behavioral therapy) [53]. A recent systematic review revealed that adult cancer survivors also experience negative attention biases with large effect sizes [44]. Recent research efforts have successfully adapted ABM for adult cancer survivors, demonstrating feasibility in pilot studies [54]. Notably, a pilot randomized controlled trial (RCT) of ABM in adult breast cancer survivors showed significant improvements in health-related anxiety and fear of cancer recurrence [55]. However, ABM has not yet been adapted or tested for adolescent and young adult cancer survivors.

This paper describes the protocol for an ongoing pilot RCT. This study aims to test the feasibility and acceptability of a novel digital positive psychological intervention for adolescent and young adult cancer survivors, the Attention Bias in Cancer survivors (ABCs) intervention, using a pilot RCT. The ABCs intervention consists of an ABM mobile intervention combined with a gratitude and savoring texting intervention. We hypothesize that the ABCs intervention will meet benchmarks for feasibility (enrollment and retention) and acceptability (user satisfaction and system usability).

Study Aims

The primary aim of this study is to evaluate the feasibility and acceptability of a digital health intervention for adolescent and young adult cancer survivors. We hypothesize that the ABCs intervention will be feasible as measured by ≥50% enrollment of eligible patients and ≥70% retention of participants in both arms (with retention defined as completion of primary outcome measures after treatment). These benchmarks are based on our previous research and known challenges in the recruitment and retention of adolescent and young adult cancer survivors and background literature on expected enrollment and retention rates for remote digital health studies [56,57]. We hypothesize that the ABCs intervention will be acceptable, as measured by the Client Satisfaction Questionnaire (established cutoff score of ≥26 indicating satisfaction) [58-61] and the System Usability Scale (SUS; established cutoff score of ≥70 indicating usability) [62]. An exploratory aim is to examine preliminary efficacy assessed via validated patient-reported outcome measures after treatment.

Methods

Ethical Considerations

All study procedures were approved by the Seattle Children's Hospital institutional review board (STUDY00004811). The trial is registered at ClinicalTrials.gov (NCT06682039). Informed consent (for individuals aged 18 to 29 years) or assent with parental consent (for individuals aged 15 to 17 years) will be obtained from all participants. Privacy and confidentiality protections are in accordance with institutional ethics guidelines. Participants will be compensated US $100 in electronic gift cards for completing a standardized survey assessment battery and attention bias assessment at T1 (baseline) and T2 (4 weeks after the intervention), each designed to be approximately 30 minutes long. Refer to Table 1 for the schedule of assessments. This protocol follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines. The completed SPIRIT checklist is provided in Multimedia Appendix 1 [63].

XSL•FO
RenderX

Table 1. Schedule for outcome measures.

Measures	Baseline assessment (T1)	Posttreatment assessment (T2; 4 weeks)
Client Satisfaction Questionnaire-8		✓
System Usability Scale		✓
Treatment Experiences Questionnaire		✓
Attention bias assessment	✓	✓
Attention Bias Questionnaire	✓	✓
Hospital Anxiety and Depression Scale	✓	✓
Short Health Anxiety Inventory	✓	✓
Fear of Cancer Recurrence Inventory-Short Form	✓	✓
Connor-Davidson 10-item Resilience Scale	✓	✓
Kessler-6 Psychological Distress Scale	✓	✓
Pain frequency	✓	✓
PROMIS[a] Pain Intensity	✓	✓
Child Activity Limitations Interview-9	✓	✓
Adolescent Sleep-Wake Scale-Short Form	✓	✓
Adolescent Insomnia Questionnaire	✓	✓
PROMIS NeuroQoL[b] Cognitive Function-Short Form	✓	✓

[a]PROMIS: Patient-Reported Outcomes Measurement Information System.

[b]NeuroQoL: quality of life in neurological disorders.

Trial Design

This pilot RCT uses a parallel, 2-arm design (Figure 1). A total of 60 participants will be enrolled at a single site and then randomized in a 1:1 allocation ratio after completing baseline surveys. The randomization schedule will be stratified by age group (those aged 15 to 18 years and those aged 19 to 29 years). Age was selected as the stratification factor to account for developmental considerations and to ensure the appropriateness of the intervention design across developmental stages. Randomization was seeded based on the exact date and time of randomization for reproducibility. All randomization was conducted by the study coordinator and data manager in R (version 4.4; R Foundation for Statistical Computing) using the *blockrand* package [64]. Randomization was conducted using permuted varying block sizes of 2, 4, and 6. Participants will be assigned to either the active ABCs intervention condition (ABM and daily savoring and gratitude SMS text messages) or a sham control condition (sham ABM and daily mood SMS text messages). Participants will remain blinded to their study condition for the duration of the study. The principal investigator and the biostatistician responsible for data analysis are blinded to the study condition. The study coordinators responsible for study procedures could not be feasibly blinded to study conditions due to study management and logistics. However, the active treatment and sham treatment are entirely remotely and digitally delivered with no personnel involvement in the administration or facilitation of the intervention.

XSL•FO
RenderX

Figure 1. Timeline of study procedures for parallel, 2-arm, pilot randomized controlled trial. ABM: attention bias modification; ABQ: Attention Bias Questionnaire; AIQ: Adolescent Insomnia Questionnaire; ASWS-SF: Adolescent Sleep-Wake Scale–Short Form; CALI-9: child Activity Limitations Interview-9; CD-RISC-10: Connor-Davidson 10-item Resilience Scale; CSQ-8: Client Satisfaction Questionnaire-8; FCRI-SF: Fear of Cancer Recurrence Inventory-Short Form; HADS: Hospital Anxiety and Depression Scale; K6: Kessler-6 Psychological Distress Scale; NeuroQoL: quality of life in neurological disorders; PROMIS: Patient-Reported Outcomes Measurement Information System; SHAI: Short Health Anxiety Inventory; SF: Short Form; SUS: System Usability Scale; TEQ: Treatment Experiences Questionnaire.

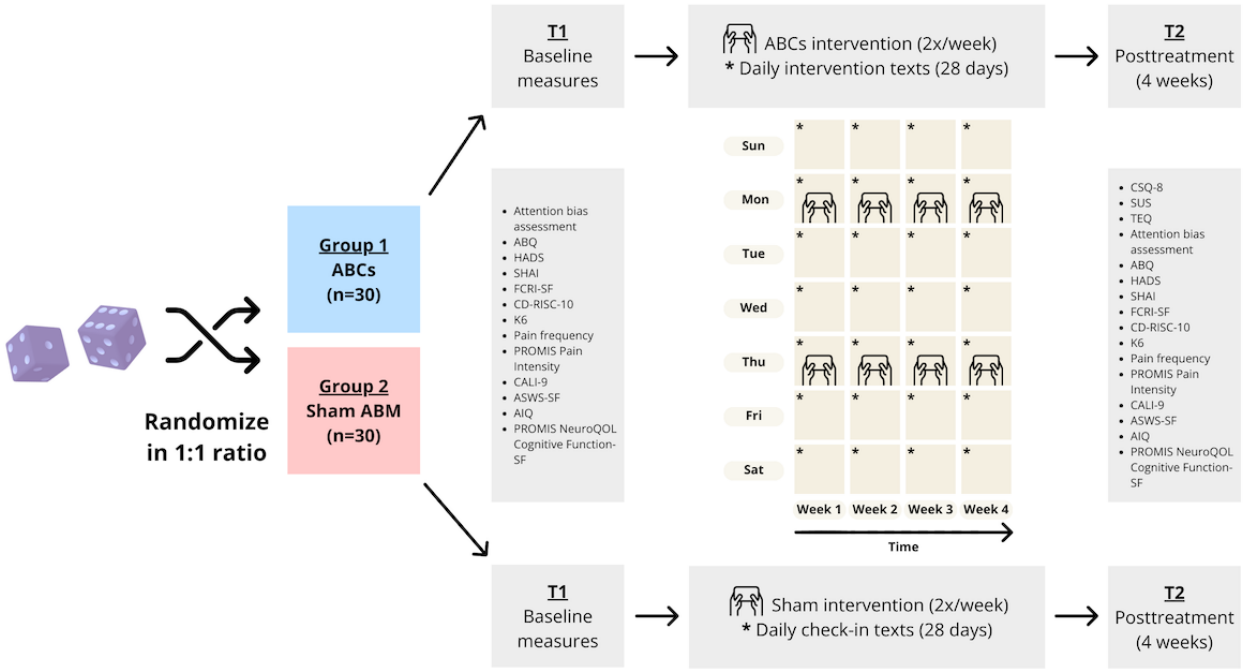

Participants

Individuals are eligible for study participation if they are currently aged 15 to 29 years. Although the National Cancer Institute definition of adolescent and young adults includes ages 15 to 39 years, we targeted a narrower age range (15 to 29 years) to align with the pediatric and adolescent and young adult populations typically treated within pediatric cancer centers, such as Seattle Children's Hospital, where this study is conducted. Eligibility criteria include a pediatric cancer diagnosis (inclusive of those diagnosed before the age of 15 years); completion of cancer treatment and in active follow-up cancer survivorship care at Seattle Children's Hospital; proficiency in English (speaking, reading, and writing); having no cognitive limitations that would make participation difficult; and having access to a smartphone with SMS text messaging capabilities. adolescent and young adults are excluded from study participation if they are receiving active or curative treatment, are not proficient in English (speaking, reading, and writing), or are cognitively or physically unable to participate. Clinical research coordinators are identifying potential participants and conducting their prescreening assessments through clinic rosters and electronic medical records. Recruitment by clinical research coordinators occurs in person at clinic visits and remotely by phone, video, or text.

Dissemination Plan

Study participants who have opted in will receive a layperson-accessible summary of study findings after we have completed primary data analysis. Deidentified study results will be disseminated to the scientific and clinical community through national and international conference presentations, peer-reviewed publications, and social media platforms.

Intervention Conditions

ABCs Intervention (ABM and Daily Savoring and Gratitude SMS Text Messages)

The ABCs intervention integrates 2 distinct components: an ABM mobile intervention and daily gratitude and savoring SMS text messages.

Attention Bias Modification

We created the ABM intervention with word stimuli tailored to cancer-specific anxiety, worries, and fears (Multimedia Appendix 2). Intervention development consisted of an iterative human-centered co-design process to ensure that the stimuli were clinically relevant and developmentally appropriate for adolescent and young adult cancer survivors. Our team first generated a preliminary list of common cancer-specific negative stimuli. The stimuli word list was then formally reviewed by key stakeholders including 6 adolescent and young adult cancer survivors, 6 pediatric oncologists, 6 psychologists, 1 psychiatrist, 1 ethicist, and 2 social workers. Stakeholders evaluated the word stimuli for readability, emotional salience, and relevance to the adolescent and young adult cancer survivorship population. Stakeholders were also encouraged to propose new cancer-related threat stimuli. The stimuli word list was further refined based on stakeholder group consensus and finalized by the study team.

Participants are instructed to complete ABM over a 4-week intervention period. ABM is completed using an app, Inquisit

XSL•FO
RenderX

6 Player (version 6.6.3; Millisecond Software LLC), to enable psychological tests remotely via mobile devices [65]. A description of active ABM intervention sessions is provided in Table 2. A single dot-probe trial is depicted in Figure 2. The total number of trials per ABM session was modeled after a large international multisite study of children and adolescents without cancer [66].

Table 2. Description of active attention bias modification (ABM) and sham ABM sessions.

	Active ABM	Sham ABM
Frequency of sessions	• Twice per week	• Twice per week
Duration of sessions	• 4 weeks	• 4 weeks
Single dot-probe trial (Figure 2)	• A fixation cross appears for 500 ms. • Paired negative and neutral (positive) stimuli appear for 500 ms. • A dot-probe appears in a location *consistent with neutral (positive) stimuli*, disappearing once the participant presses the "left" (letter "E") or "right" (letter "I") key. • If no key is pressed within 2000 ms, the program automatically advances to the next trial.	• A fixation cross appears for 500 ms. • Paired negative and neutral (positive) stimuli appear for 500 ms. • A dot-probe is *equally likely to appear in a location consistent with either negative or neutral (positive) stimuli*, disappearing once the participant presses the "left" (letter "E") or "right" (letter "I") key. • If no key is pressed within 2000 ms, the program automatically advances to the next trial.
Number of trials per session	• 120	• 120

Figure 2. Example of a single dot-probe trial of the attention bias modification intervention.

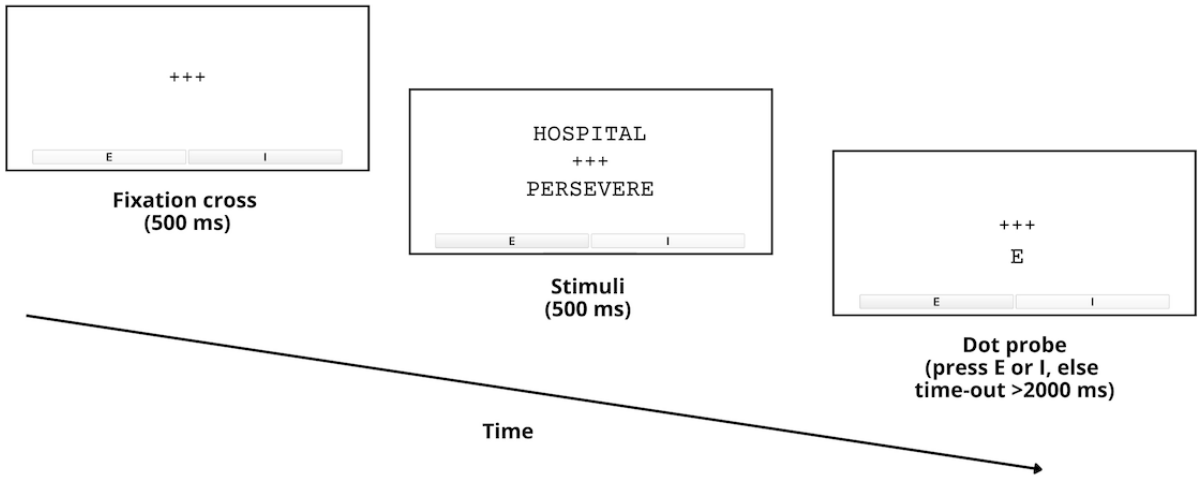

Gratitude and Savoring SMS Text Messages

Automated interactive savoring and gratitude SMS text messages are sent to participants daily at a prespecified time of their preference, soliciting open-ended responses. Gratitude and savoring intervention SMS text messaging prompts encourage participants to engage in past, present, and future–oriented savoring of positive emotions related to people, places, and things for which they are grateful (Figure 3). Intervention prompts aim to extend and increase positive emotions by fostering a meta-awareness of positive experiences, including positive memory building, sensory-perceptual experiences, positive imagination, and reminiscing about past positive memories. Participants also respond to daily mood monitoring prompts: (1) How content, happy, or calm are you feeling right now? (2) How stressed are you feeling right now?

XSL•FO
RenderX

Figure 3. Example of daily savoring and gratitude intervention SMS text messaging prompts.



Sham Control Condition (Sham ABM and Mood SMS Text Messages)

The sham control condition is designed to mirror the active intervention condition, providing a nontherapeutic sham ABM mobile intervention plus daily mood monitoring texts.

Sham ABM

Participants are instructed to complete sham ABM via the Inquisit app over a 4-week intervention period. Consistent with established ABM protocols, the sham condition uses the same negative threat stimuli as the active treatment condition *without* attention retraining toward neutral or positive stimuli, which is necessary to modify attention bias. This design ensures that the sham condition is identical to the active treatment condition in duration, task structure, and appearance. A description of sham ABM intervention sessions has been provided in Table 2.

Mood Monitoring SMS Text Messages

Automated interactive mood monitoring SMS text messages are sent to participants daily at a prespecified time of their preference. SMS text messages consist of the following prompts: (1) How content, happy, or calm are you feeling right now? (2) How stressed are you feeling right now?

Primary Outcomes

Feasibility

Feasibility is defined by recruitment and retention rates. The intervention will be considered feasible if we achieve ≥50% enrollment of eligible patients and ≥70% retention across both study arms. We will report the eligibility rate, defined as the proportion of adolescent and young adults eligible to participate in the study out of all adolescent and young adults we attempted to contact. We will report the participation rate, defined as the proportion of eligible adolescent and young adults who enrolled in the study. Adherence to the 4-week ABCs intervention and the sham control condition will be monitored via the app backend and REDCap (Research Electronic Data Capture) SMS text messaging logs. For both the active treatment and sham conditions, we will track the number of ABM sessions completed and the number of text prompts answered. There are no established benchmarks for optimal digital intervention adherence, and such interventions are typically associated with low engagement. For the purpose of this pilot trial, we have defined intervention exposure as the completion of at least one ABM session and one SMS text message interaction. We will report on data completeness for all primary and exploratory outcome measures, aiming for ≥80% data completeness across patient-reported outcome measures at T1 (baseline) and T2 (after treatment).

Client Satisfaction Questionnaire-8

The Client Satisfaction Questionnaire-8 is an 8-item questionnaire used to assess the level of satisfaction with care. Items are scored on a Likert scale from 1 (low satisfaction) to 4 (high satisfaction) with different descriptors for each response point. Total scores range from 8 to 32, with higher scores indicating greater satisfaction (scores ≥26 indicate satisfaction). The Client Satisfaction Questionnaire-8 has been found to have high internal consistency and concurrent validity in inpatient and outpatient mental health settings [58-61].

System Usability Scale

The SUS is a well-validated and widely used 10-item scale to evaluate the perceived usability of digital interventions [62]. Items are rated on a 5-point Likert scale, reverse-scored as applicable, and converted to a 4-point scale, then multiplied by 2.5 to yield a total score ranging from 0 to 100; scores ≥70 demonstrate adequate usability.

XSL•FO
RenderX

Exploratory Outcomes

Treatment Experiences Questionnaire

The Treatment Experiences Questionnaire is a measure developed by the senior author and has been used in digital health intervention trials conducted by our research team. This is a 2-item measure in which participants in both study arms will indicate whether they experienced any negative effects from the program and how severely the program affected them.

Attention Bias Assessment

Attention bias will be assessed using a dot-probe task administered via the Inquisit 6 Player app. Attention bias scores will be computed by subtracting the mean reaction time from trials in which probes replace negative stimuli from the mean reaction time of trials in which probes replace neutral or positive stimuli. A positive score will indicate an attentional bias toward threatening stimuli, and negative scores will indicate an attentional bias toward nonthreatening stimuli.

Attention Bias Questionnaire

The Attention Bias Questionnaire consists of 9 items, reflecting the subjective experience of attention bias toward threats, with two subscales: engagement with threat and difficulty disengaging from threat. High scores in the Attention Bias Questionnaire are correlated with trait anxiety, social anxiety, posttraumatic stress disorder, and depression. It has been validated in individuals aged 18 years and older, with high internal consistency (Cronbach α=0.90) [67].

Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale is a 14-item measure that assesses symptoms of anxiety and depression, respectively, in patients with serious illness. It has been validated in adolescent and young adults with chronic illness and cancer survivors, with excellent reliability (Cronbach α=0.83) [68-71]. Items are scored from 0 to 3 (subscale range 0-21), with scores ≥8 categorized as borderline abnormal and ≥11 categorized as abnormal.

Short Health Anxiety Inventory

The Short Health Anxiety Inventory contains 14 items that assess health anxiety independent of physical health status. Items assess worry about one's health, awareness of bodily sensations and changes, and the feared consequences of having an illness. It can be used in both healthy and physically ill individuals, including those who were temporarily sick or diagnosed with a serious and chronic illness. The scale has been validated in clinical and nonclinical samples [72,73].

Fear of Cancer Recurrence Inventory-Short Form

The Fear of Cancer Recurrence Inventory-Short Form is a widely used and validated 9-item measure that evaluates the presence and severity of intrusive thoughts associated with fear of cancer recurrence [74-76]. Each item is rated on a Likert scale ranging from 0 ("not at all" or "never") to 4 ("a great deal' or "all the time"). A higher score indicates higher levels of fear of cancer recurrence.

Connor-Davidson 10-Item Resilience Scale

The Connor-Davidson 10-item Resilience Scale is a validated and widely used instrument to measure inherent resilience [77]. Questions revolve around personal problem-solving and approaches to adversity. The 10-item instrument has high internal consistency (Cronbach α=0.85) and has been used in diverse populations including adolescents, parents, and patients with cancer [78]. Each item consists of a 5-point Likert scale (scored from 0 to 4), for a total of 40 points, with higher scores reflecting greater resilience.

Kessler-6 Psychological Distress Scale

This 6-item scale measures the level of psychological distress experienced in the past month. The instrument strongly discriminates between community cases and noncases of the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*) psychiatric disorders such as serious emotional distress or serious mental illness. It has been extensively cross-validated, including among adolescents [79,80]. Responses are scored on a 5-point Likert scale, generating a range of 0 to 24 points. Previous studies have shown that scores ≥7 are consistent with "high" distress, and those ≥13 meet criteria for serious or debilitating psychological distress.

Pain Frequency

Three questions will assess pain frequency (6-item or 7-item Likert scale ranging from "not at all" to "daily"), pain intensity (11-point Numeric Rating Scale, 0-10), and pain interference (11-point Numeric Rating Scale, 0-10).

Patient-Reported Outcomes Measurement Information System Pain Intensity

The Patient-Reported Outcomes Measurement Information System (PROMIS) are publicly available, rigorously tested, and well-validated instrument supported by the National Institutes of Health Common Fund. The PROMIS Pain Intensity instrument assesses how much a person hurts [81].

Child Activity Limitations Interview-9

The Child Activity Limitations Interview-9 is a brief measure for assessing activity limitations in children and adolescents with chronic pain. This measure has demonstrated good internal consistency and high cross-informant reliability [82].

Adolescent Sleep-Wake Scale–Short Form

The Adolescent Sleep-Wake Scale–Short Form is a 10-item measure of behavioral sleep patterns that has been validated in general adolescent and young adult populations [83,84]. The measure provides an overall sleep quality score and 3 subscale scores: going to bed, falling asleep and reinitiating sleep, and returning to wakefulness. Items are rated on a 6-point Likert scale and averaged to compute subscale and total measure scores. Total scores range from 1 to 6, with higher scores indicating better sleep quality.

Adolescent Insomnia Questionnaire

The Adolescent Insomnia Questionnaire is a 13-item screening measure of insomnia in adolescents. The measure has demonstrated strong reliability and high internal consistency in

XSL•FO
RenderX

study samples up to the age of 19 years [85]. Items are rated on a 5-point Likert scale. Total scores range from 0 to 52, with higher scores indicating more insomnia symptoms.

PROMIS Quality of Life in Neurological Disorders Cognitive Function-Short Form

The PROMIS NeuroQoL (quality of life in neurological disorders) Cognitive Function-Short Form measures perceived difficulties in everyday cognitive abilities such as memory, attention, and decision-making [86].

Data Analysis

The primary aim of this study is to determine the feasibility and acceptability of the ABCs intervention. Feasibility will be defined by meeting the following criteria: (1) ≥50% enrollment of eligible patients and (2) ≥70% retention of participants in both arms. Acceptability will be defined as (1) an average score ≥26 on the Client Satisfaction Questionnaire and (2) an average score ≥70 on the SUS.

All participants who complete baseline surveys and are randomized will be included in intent-to-treat analyses to explore the preliminary efficacy of all patient-reported outcome measures in our assessment battery. Preliminary efficacy will be examined using 2-sample comparisons and regression models. Two-sample comparisons will be based on 2-tailed *t* tests or chi-square tests, depending on the type of outcome. The regression models aim to assess the intervention's effect on continuous patient-reported outcomes while controlling for baseline scores and other potential confounders. Potential covariates include age, gender, cancer diagnosis, and other sociodemographic or disease-related or treatment-related characteristics previously shown to influence psychosocial outcomes. All statistical analyses will be performed in the R software [64].

Results

This project was funded in August 2023. The study launch date (date of first participant enrolled) was November 8, 2024. Data collection has been completed as of February 2026. After data collection, analyses are expected to be completed by September 2026. Dissemination of findings and manuscript submission are planned for December 2026.

Discussion

Anticipated Findings

This manuscript describes the protocol for a pilot RCT designed to evaluate the feasibility and acceptability of the ABCs intervention, which is a novel digital intervention tailored for adolescent and young adult cancer survivors. The ABCs intervention is specifically designed to target anxiety, addressing a critical gap in cancer survivorship care. We hypothesize that the ABCs digital intervention will demonstrate feasibility and acceptability within the adolescent and young adult cancer survivor population. The exploratory patient-reported outcomes data collected from this pilot trial will provide effect size estimates to inform power calculations and sample sizes for future multisite efficacy trials.

Comparison to Previous Work

Digital health interventions offer a unique opportunity to address existing access care gaps by using technologies that align with adolescent and young adult communication and learning preferences. The ABCs intervention builds on a robust body of literature establishing ABM as a digital anxiety intervention across the age continuum and in various disease groups [39-41]. A previous systematic review has confirmed that cancer survivors, such as individuals with clinical anxiety disorders, demonstrate a negative attention bias [44]. However, existing digital psychosocial interventions for adolescent and young adult cancer survivors have not been designed to target anxiety symptoms [35-38]. To our knowledge, this protocol represents the first study to design and test a digital anxiety intervention for adolescent and young adult cancer survivors, targeting cognitive mechanisms of anxiety [42,43].

Strengths

A primary strength of this study is the use of a widely used digital anxiety intervention, ABM, that we have adapted for adolescent and young adult cancer survivors. The development of the ABCs intervention followed a rigorous user-centered co-design approach involving a multidisciplinary group of stakeholders. The digital deployment model allows for remote delivery and scalability and helps overcome the access care gap that adolescent and young adult cancer survivors face when seeking traditional mental health services.

Limitations

Our study has several anticipated limitations. First, this is a single-site pilot trial, which limits the generalizability of findings. However, a strength of this approach is the presence of preexisting relationships with the clinic to ensure project completion and the collection of pilot data for future clinical trials. Second, cognitive late effects may present difficulties in engaging with ABM, which consists of repetitive reaction-time association tasks. Although not yet tested in adolescent and young adult cancer survivors, findings from attention bias assessment and ABM studies in pediatric medical populations and adult breast cancer survivors support its potential feasibility. On the basis of the results of this ongoing trial and participant feedback, we may vary stimulus presentation times for future intervention refinement.

Future Directions

Findings from this pilot trial will inform a larger multisite hybrid effectiveness-implementation trial. This is a critical next step in digitizing evidence-based positive psychological interventions and co-designing for engagement and sustainability. This study and future studies building on this work aim to expand the adoption and reach of evidence-based mobile health psychosocial interventions.

XSL•FO
RenderX

Acknowledgments

Dr Lau is supported for unrelated work by the National Cancer Institute at the National Institutes of Health (grant 1K08CA263474). The research reported in this publication was supported by the Fred Hutch/University of Washington/Seattle Children's Cancer Consortium Survivorship Program Pilot Award.

Funding

The Cancer Consortium is supported by funding from the US National Institutes of Health (grant P30 CA015704). The opinions herein represent those of the authors and not necessarily those of the funders.

Data Availability

Deidentified human subjects data from this research study will be banked indefinitely. Data will be made available upon reasonable request to other researchers who seek approval from the research team via email to the principal investigator (NL).

Authors' Contributions

NL, CZ, HA, EM, TK, EJC, JRF, LCH, and TMP conceptualized the study. NL drafted the manuscript. Visualization was conducted by SJHH. CZ, SJHH, HA, SH, IB, EM, TK, EJC, JRF, LCH, and TMP reviewed and edited the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2025 checklist.
[DOCX File , 34 KB - resprot_v15i1e82665_app1.docx]

Multimedia Appendix 2
Attention bias modification word stimuli list for the Attention Bias in Cancer survivors intervention.
[DOCX File , 17 KB - resprot_v15i1e82665_app2.docx]

References

1. GBD 2019 Adolescent Young Adult Cancer Collaborators. The global burden of adolescent and young adult cancer in 2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Oncol 2022 Jan;23(1):27-52 [FREE Full text] [doi: 10.1016/S1470-2045(21)00581-7] [Medline: 34871551]
2. DeVine A, Landier W, Hudson MM, Constine LS, Bhatia S, Armenian SH, et al. The children's oncology group long-term follow-up guidelines for survivors of childhood, adolescent, and young adult cancers: a review. JAMA Oncol 2025 May 01;11(5):544-553. [doi: 10.1001/jamaoncol.2024.6812] [Medline: 39976936]
3. Zebrack BJ, Zeltzer LK, Whitton J, Mertens AC, Odom L, Berkow R, et al. Psychological outcomes in long-term survivors of childhood leukemia, Hodgkin's disease, and non-Hodgkin's lymphoma: a report from the Childhood Cancer Survivor Study. Pediatrics 2002 Jul;110(1 Pt 1):42-52. [doi: 10.1542/peds.110.1.42] [Medline: 12093945]
4. Gurney JG, Krull KR, Kadan-Lottick N, Nicholson HS, Nathan PC, Zebrack B, et al. Social outcomes in the childhood cancer survivor study cohort. J Clin Oncol 2009 May 10;27(14):2390-2395 [FREE Full text] [doi: 10.1200/JCO.2008.21.1458] [Medline: 19224833]
5. Kazak AE, Derosa BW, Schwartz LA, Hobbie W, Carlson C, Ittenbach RF, et al. Psychological outcomes and health beliefs in adolescent and young adult survivors of childhood cancer and controls. J Clin Oncol 2010 Apr 20;28(12):2002-2007 [FREE Full text] [doi: 10.1200/JCO.2009.25.9564] [Medline: 20231679]
6. Patterson P, McDonald FE, Zebrack B, Medlow S. Emerging issues among adolescent and young adult cancer survivors. Semin Oncol Nurs 2015 Feb;31(1):53-59. [doi: 10.1016/j.soncn.2014.11.006] [Medline: 25636395]
7. Keegan TH, Lichtensztajn DY, Kato I, Kent EE, Wu X, West MM, AYA HOPE Study Collaborative Group. Unmet adolescent and young adult cancer survivors information and service needs: a population-based cancer registry study. J Cancer Surviv 2012 Sep 29;6(3):239-250 [FREE Full text] [doi: 10.1007/s11764-012-0219-9] [Medline: 22457219]
8. Perez GK, Salsman JM, Fladeboe K, Kirchhoff AC, Park ER, Rosenberg AR. Taboo topics in adolescent and young adult oncology: strategies for managing challenging but important conversations central to adolescent and young adult cancer survivorship. Am Soc Clin Oncol Educ Book 2020 Mar;40:1-15 [FREE Full text] [doi: 10.1200/EDBK_279787] [Medline: 32324424]
9. Jones JM, Fitch M, Bongard J, Maganti M, Gupta A, D'Agostino N, et al. The needs and experiences of post-treatment adolescent and young adult cancer survivors. J Clin Med 2020 May 13;9(5):1444 [FREE Full text] [doi: 10.3390/jcm9051444] [Medline: 32413981]

XSL•FO
RenderX

10. Use of a grading system in the evaluation of complications in a randomised controlled trial on cataract surgery. Oxford Cataract Treatment and Evaluation Team (OCTET). Br J Ophthalmol 1986 Jun;70(6):411-414 [FREE Full text] [doi: 10.1136/bjo.70.6.411] [Medline: 2872914]
11. Yang Y, Li W, Wen Y, Wang H, Sun H, Liang W, et al. Fear of cancer recurrence in adolescent and young adult cancer survivors: a systematic review of the literature. Psychooncology 2019 Apr 17;28(4):675-686. [doi: 10.1002/pon.5013] [Medline: 30703261]
12. Osmani V, Hörner L, Klug SJ, Tanaka LF. Prevalence and risk of psychological distress, anxiety and depression in adolescent and young adult (AYA) cancer survivors: a systematic review and meta-analysis. Cancer Med 2023 Sep 10;12(17):18354-18367 [FREE Full text] [doi: 10.1002/cam4.6435] [Medline: 37559504]
13. Muzzin LJ, Anderson NJ, Figueredo AT, Gudelis SO. The experience of cancer. Soc Sci Med 1994 May;38(9):1201-1208. [doi: 10.1016/0277-9536(94)90185-6] [Medline: 8016685]
14. Derry-Vick HM, Heathcote LC, Glesby N, Stribling J, Luebke M, Epstein AS, et al. Scanxiety among adults with cancer: a scoping review to guide research and interventions. Cancers (Basel) 2023 Feb 22;15(5):1381 [FREE Full text] [doi: 10.3390/cancers15051381] [Medline: 36900174]
15. Giovanelli A, Ozer EM, Dahl RE. Leveraging technology to improve health in adolescence: a developmental science perspective. J Adolesc Health 2020 Aug;67(2S):S7-13 [FREE Full text] [doi: 10.1016/j.jadohealth.2020.02.020] [Medline: 32718517]
16. Gulliver A, Griffiths KM, Christensen H. Perceived barriers and facilitators to mental health help-seeking in young people: a systematic review. BMC Psychiatry 2010 Dec 30;10(1):113. [doi: 10.1186/1471-244x-10-113]
17. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. Health information-seeking behaviour in adolescence: the place of the internet. Soc Sci Med 2005 Apr;60(7):1467-1478. [doi: 10.1016/j.socscimed.2004.08.010] [Medline: 15652680]
18. Ebert DD, Zarski AC, Christensen H, Stikkelbroek Y, Cuijpers P, Berking M, et al. Internet and computer-based cognitive behavioral therapy for anxiety and depression in youth: a meta-analysis of randomized controlled outcome trials. PLoS One 2015 Mar 18;10(3):e0119895 [FREE Full text] [doi: 10.1371/journal.pone.0119895] [Medline: 25786025]
19. Ybarra ML, Eaton WW. Internet-based mental health interventions. Ment Health Serv Res 2005 Jun;7(2):75-87. [doi: 10.1007/s11020-005-3779-8] [Medline: 15974154]
20. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar 20;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
21. Granger D, Vandelanotte C, Duncan MJ, Alley S, Schoeppe S, Short C, et al. Is preference for mHealth intervention delivery platform associated with delivery platform familiarity? BMC Public Health 2016 Jul 22;16(1):619 [FREE Full text] [doi: 10.1186/s12889-016-3316-2] [Medline: 27450240]
22. Radesky J, Weeks H, Schaller A, Robb M, Mann S, Lenhart A. Constant companion: a week in the life of a young person's smartphone use. Common Sense. URL: https://www.commonsensemedia.org/sites/default/files/research/report/2023-cs-smartphone-research-report_final-for-web.pdf [accessed 2025-05-24]
23. Salsman JM, Rosenberg AR. Fostering resilience in adolescence and young adulthood: considerations for evidence-based, patient-centered oncology care. Cancer 2024 Apr 01;130(7):1031-1040 [FREE Full text] [doi: 10.1002/cncr.35182] [Medline: 38163249]
24. Stifter C, Augustine M, Dollar J. The role of positive emotions in child development: a developmental treatment of the broaden and build theory. J Posit Psychol 2019 Nov 22;15(1):89-94. [doi: 10.1080/17439760.2019.1695877]
25. Corman M, Aubret D, Ghazal J, Berthon M, Chausse P, Lohou C, et al. Attentional bias modification with a new paradigm: the effect of the Detection Engagement and Savoring Positivity (DESP) task on eye-tracking of attention. J Behav Ther Exp Psychiatry 2020 Sep;68:101525. [doi: 10.1016/j.jbtep.2019.101525] [Medline: 31739249]
26. Garland EL, Farb NA, Goldin P, Fredrickson BL. Mindfulness broadens awareness and builds eudaimonic meaning: a process model of mindful positive emotion regulation. Psychol Inq 2015 Oct 01;26(4):293-314 [FREE Full text] [doi: 10.1080/1047840X.2015.1064294] [Medline: 27087765]
27. Garland EL, Thielking P, Thomas EA, Coombs M, White S, Lombardi J, et al. Linking dispositional mindfulness and positive psychological processes in cancer survivorship: a multivariate path analytic test of the mindfulness-to-meaning theory. Psychooncology 2017 May 22;26(5):686-692 [FREE Full text] [doi: 10.1002/pon.4065] [Medline: 26799620]
28. Goldberg SB, Tucker RP, Greene PA, Davidson RJ, Wampold BE, Kearney DJ, et al. Mindfulness-based interventions for psychiatric disorders: a systematic review and meta-analysis. Clin Psychol Rev 2018 Feb;59:52-60 [FREE Full text] [doi: 10.1016/j.cpr.2017.10.011] [Medline: 29126747]
29. Davis DE, Choe E, Meyers J, Wade N, Varjas K, Gifford A, et al. Thankful for the little things: a meta-analysis of gratitude interventions. J Couns Psychol 2016 Jan;63(1):20-31. [doi: 10.1037/cou0000107] [Medline: 26575348]
30. Geraghty AW, Wood AM, Hyland ME. Attrition from self-directed interventions: investigating the relationship between psychological predictors, intervention content and dropout from a body dissatisfaction intervention. Soc Sci Med 2010 Jul;71(1):30-37. [doi: 10.1016/j.socscimed.2010.03.007] [Medline: 20400220]

XSL•FO
RenderX

31. Diniz G, Korkes L, Tristão LS, Pelegrini R, Bellodi P, Bernardo W. The effects of gratitude interventions: a systematic review and meta-analysis. Einstein (Sao Paulo) 2023;21:eRW0371 [FREE Full text] [doi: 10.31744/einstein_journal/2023RW0371] [Medline: 37585888]
32. Cullen K, Murphy M, Di Blasi Z, Bryant FB. The effectiveness of savouring interventions in adult clinical populations: a systematic review. Int J Appl Posit Psychol 2024 Jul 24;9(3):1589-1615. [doi: 10.1007/S41042-024-00182-1]
33. Saboor S, Medina A, Marciano L. Application of positive psychology in digital interventions for children, adolescents, and young adults: systematic review and meta-analysis of controlled trials. JMIR Ment Health 2024 Aug 14;11:e56045 [FREE Full text] [doi: 10.2196/56045] [Medline: 39141906]
34. Zheng J, Zaremohzzabieh Z, Roslan S. Examining the influence of savoring interventions on positive emotions in university students: a systematic review and meta-analysis. Asian J Soc 2025 Jan 31;21(1):103. [doi: 10.5539/ass.v21n1p103]
35. Devine KA, Viola AS, Coups EJ, Wu YP. Digital health interventions for adolescent and young adult cancer survivors. JCO Clin. Cancer Inform 2018 Dec(2):1-15. [doi: 10.1200/cci.17.00138]
36. Bertrand A, Escot N, Girodet M, Congiusta T, Baudry AS, Demoor C, et al. Digital tools used to enhance the quality of life and well-being of adolescent and young adult cancer survivors: a systematic review of the literature. J Adolesc Young Adult Oncol (Forthcoming) 2025 Sep 12. [doi: 10.1177/21565333251377393] [Medline: 40938393]
37. Zhong C, Luo X, Tan M, Chi J, Guo B, Tang J, et al. Digital health interventions to improve mental health in patients with cancer: umbrella review. J Med Internet Res 2025 Feb 21;27:e69621 [FREE Full text] [doi: 10.2196/69621] [Medline: 39984165]
38. Zhu S, Wang Y, Hu Y. Facilitators and barriers to digital mental health interventions for depression, anxiety, and stress in adolescents and young adults: scoping review. J Med Internet Res 2025 Mar 24;27:e62870 [FREE Full text] [doi: 10.2196/62870] [Medline: 40127430]
39. Blankers M, Salemink E, Wiers R. Cognitive behavioral therapy, and cognitive bias modification in internet-based interventions for mood, anxiety, and substance use disorders. In: Mucić D, Hilty DM, editors. Digital Mental Health: The Future is Now. Cham, Switzerland: Springer; 2025:351-377.
40. Fodor LA, Georgescu R, Cuijpers P, Szamoskozi C, David D, Furukawa TA, et al. Efficacy of cognitive bias modification interventions in anxiety and depressive disorders: a systematic review and network meta-analysis. Lancet Psychiatry 2020 Jun;7(6):506-514. [doi: 10.1016/S2215-0366(20)30130-9] [Medline: 32445689]
41. Hang Y, Xu L, Wang C, Zhang G, Zhang N. Can attention bias modification augment the effect of CBT for anxiety disorders? A systematic review and meta-analysis. Psychiatry Res 2021 May;299:113892. [doi: 10.1016/j.psychres.2021.113892] [Medline: 33799125]
42. Vasey M, MacLeod C. Information-processing factors in childhood anxiety: a review and developmental perspective. In: Vasey MW, Dadds MR, editors. The Developmental Psychopathology of Anxiety. Oxford, UK: Oxford University Press; 2001:253-277.
43. Songco A, Hudson JL, Fox E. A cognitive model of pathological worry in children and adolescents: a systematic review. Clin Child Fam Psychol Rev 2020 Jun 28;23(2):229-249 [FREE Full text] [doi: 10.1007/s10567-020-00311-7] [Medline: 31989444]
44. Pradhan P, Sharpe L, Butow P. The role of attentional biases in the context of cancer: a systematic review and meta‐analysis. Psycho-Oncol 2021 Feb 05;30(5):649-658. [doi: 10.1002/pon.5617]
45. Yoon KL, Zinbarg RE. Interpreting neutral faces as threatening is a default mode for socially anxious individuals. J Abnorm Psychol 2008 Aug;117(3):680-685. [doi: 10.1037/0021-843x.117.3.680]
46. Heuer K, Lange WG, Isaac L, Rinck M, Becker ES. Morphed emotional faces: emotion detection and misinterpretation in social anxiety. J Behav Ther Exp Psychiatry 2010 Dec;41(4):418-425. [doi: 10.1016/j.jbtep.2010.04.005] [Medline: 20511123]
47. Mohlman J, Carmin CN, Price RB. Jumping to interpretations: social anxiety disorder and the identification of emotional facial expressions. Behav Res Ther 2007 Mar;45(3):591-599. [doi: 10.1016/j.brat.2006.03.007] [Medline: 16643844]
48. Eldar S, Apter A, Lotan D, Edgar KP, Naim R, Fox NA, et al. Attention bias modification treatment for pediatric anxiety disorders: a randomized controlled trial. Am J Psychiatry 2012 Feb;169(2):213-220. [doi: 10.1176/appi.ajp.2011.11060886] [Medline: 22423353]
49. Amir N, Beard C, Burns M, Bomyea J. Attention modification program in individuals with generalized anxiety disorder. J Abnorm Psychol 2009 Feb;118(1):28-33 [FREE Full text] [doi: 10.1037/a0012589] [Medline: 19222311]
50. Wald I, Fruchter E, Ginat K, Stolin E, Dagan D, Bliese PD, et al. Selective prevention of combat-related post-traumatic stress disorder using attention bias modification training: a randomized controlled trial. Psychol Med 2016 Sep 05;46(12):2627-2636. [doi: 10.1017/S0033291716000945] [Medline: 27377418]
51. MacLeod C, Clarke PJ. The attentional bias modification approach to anxiety intervention. Clin Psychol Sci 2015 Jan 06;3(1):58-78 [FREE Full text] [doi: 10.1177/2167702614560749] [Medline: 22423353]
52. de Rutte J, Myruski S, Davis E, Findley A, Dennis-Tiwary TA. A randomized clinical trial investigating the clinical impact of a game-based digital therapeutic for social anxiety disorder. J Anxiety Disord 2025 Apr;111:103000. [doi: 10.1016/j.janxdis.2025.103000] [Medline: 40068434]

XSL•FO
RenderX

53. Hakamata Y, Lissek S, Bar-Haim Y, Britton JC, Fox NA, Leibenluft E, et al. Attention bias modification treatment: a meta-analysis toward the establishment of novel treatment for anxiety. Biol Psychiatry 2010 Dec 01;68(11):982-990 [FREE Full text] [doi: 10.1016/j.biopsych.2010.07.021] [Medline: 20887977]
54. Tenschert DD, Wilbers J, Eichelsheim P, Delbressine V, Schenkel E, Custers JA, et al. Examining the feasibility of cognitive bias modification for patients with cancer and survivors in 3 independent pilot studies. J Psychosoc Oncol Res Pract 2025;7(1):e153. [doi: 10.1097/OR9.0000000000000153]
55. Lichtenthal WG, Corner GW, Slivjak ET, Roberts KE, Li Y, Breitbart W, et al. A pilot randomized controlled trial of cognitive bias modification to reduce fear of breast cancer recurrence. Cancer 2017 Apr 15;123(8):1424-1433 [FREE Full text] [doi: 10.1002/cncr.30478] [Medline: 28055119]
56. Harlan LC, Lynch CF, Keegan TH, Hamilton AS, Wu X, Kato I, AYA HOPE Study Collaborative Group. Recruitment and follow-up of adolescent and young adult cancer survivors: the AYA HOPE study. J Cancer Surviv 2011 Sep 28;5(3):305-314 [FREE Full text] [doi: 10.1007/s11764-011-0173-y] [Medline: 21274648]
57. Daniore P, Nittas V, von Wyl V. Enrollment and retention of participants in remote digital health studies: scoping review and framework proposal. J Med Internet Res 2022 Sep 09;24(9):e39910 [FREE Full text] [doi: 10.2196/39910] [Medline: 36083626]
58. Larsen DL, Attkisson CC, Hargreaves WA, Nguyen TD. Assessment of client/patient satisfaction: development of a general scale. Eval Program Plann 1979 Jan;2(3):197-207. [doi: 10.1016/0149-7189(79)90094-6] [Medline: 10245370]
59. Attkisson CC, Zwick R. The client satisfaction questionnaire. Psychometric properties and correlations with service utilization and psychotherapy outcome. Eval Program Plann 1982 Jan;5(3):233-237. [doi: 10.1016/0149-7189(82)90074-x] [Medline: 10259963]
60. Miles H, Johnson S, Amponsah-Afuwape S, Finch E, Leese M, Thornicroft G. Characteristics of subgroups of individuals with psychotic illness and a comorbid substance use disorder. Psychiatr Serv 2003 Apr;54(4):554-561. [doi: 10.1176/appi.ps.54.4.554] [Medline: 12663845]
61. De Wilde EF, Hendriks VM. The client satisfaction questionnaire: psychometric properties in a Dutch addict population. Eur Addict Res 2005 Aug 19;11(4):157-162. [doi: 10.1159/000086396] [Medline: 16110221]
62. Peres SC, Pham T, Phillips R. Validation of the System Usability Scale (SUS): SUS in the wild. Proc Hum Factors Ergon Soc Annu Meet 2013 Sep 30;57(1):192-196. [doi: 10.1177/1541931213571043]
63. Chan AW, Boutron I, Hopewell S, Moher D, Schulz KF, Collins G, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. PLoS Med 2025 Apr;22(4):e1004589 [FREE Full text] [doi: 10.1371/journal.pmed.1004589] [Medline: 40294521]
64. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. URL: https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf [accessed 2025-05-29]
65. Inquisit 6. Millisecond Software. URL: https://www.millisecond.com/ [accessed 2025-05-29]
66. Abend R, de Voogd L, Salemink E, Wiers RW, Pérez-Edgar K, Fitzgerald A, et al. Association between attention bias to threat and anxiety symptoms in children and adolescents. Depress Anxiety 2018 Mar 06;35(3):229-238 [FREE Full text] [doi: 10.1002/da.22706] [Medline: 29212134]
67. Azriel O, Britton JC, Gober CD, Pine DS, Bar-Haim Y. Development and validation of the Attention Bias Questionnaire (ABQ). Int J Methods Psychiatr Res 2022 Jun 17;31(2):e1905 [FREE Full text] [doi: 10.1002/mpr.1905] [Medline: 35297127]
68. White D, Leach C, Sims R, Atkinson M, Cottrell D. Validation of the hospital anxiety and depression scale for use with adolescents. Br J Psychiatry 1999 Nov 03;175(5):452-454. [doi: 10.1192/bjp.175.5.452] [Medline: 10789277]
69. Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the hospital anxiety and depression scale: an updated literature review. J Psychosom Res 2002 Feb;52(2):69-77. [doi: 10.1016/s0022-3999(01)00296-3] [Medline: 11832252]
70. Moons P, Kovacs AH, Luyckx K, Thomet C, Budts W, Enomoto J, APPROACH-IS consortiumthe International Society for Adult Congenital Heart Disease (ISACHD). Patient-reported outcomes in adults with congenital heart disease: inter-country variation, standard of living and healthcare system factors. Int J Cardiol 2018 Jan 15;251:34-41. [doi: 10.1016/j.ijcard.2017.10.064] [Medline: 29107358]
71. Lemay KR, Tulloch HE, Pipe AL, Reed JL. Establishing the minimal clinically important difference for the hospital anxiety and depression scale in patients with cardiovascular disease. J Cardiopulm Rehabil Prev 2019 Nov;39(6):E6-11. [doi: 10.1097/HCR.0000000000000379] [Medline: 30489438]
72. Abramowitz JS, Deacon BJ, Valentiner DP. The short health anxiety inventory: psychometric properties and construct validity in a non-clinical sample. Cognit Ther Res 2007 Feb 15;31(6):871-883 [FREE Full text] [doi: 10.1007/s10608-006-9058-1] [Medline: 32214558]
73. Alberts NM, Hadjistavropoulos HD, Jones SL, Sharpe D. The short health anxiety inventory: a systematic review and meta-analysis. J Anxiety Disord 2013 Jan;27(1):68-78. [doi: 10.1016/j.janxdis.2012.10.009] [Medline: 23247202]
74. Simard S, Savard J. Fear of cancer recurrence inventory: development and initial validation of a multidimensional measure of fear of cancer recurrence. Support Care Cancer 2009 Mar 15;17(3):241-251. [doi: 10.1007/s00520-008-0444-y] [Medline: 18414902]

XSL•FO
RenderX

75. Peng L, Huang W, Zhang W, Xu Y, Lu F, Zhong L, et al. Psychometric properties of the short form of the Fear of Cancer Recurrence Inventory (FCRI) in Chinese breast cancer survivors. Front Psychiatry 2019 Aug 7;10:537 [FREE Full text] [doi: 10.3389/fpsyt.2019.00537] [Medline: 31447709]
76. Custers JA, Gielissen MF, de Wilt JH, Honkoop A, Smilde TJ, van Spronsen D, et al. Towards an evidence-based model of fear of cancer recurrence for breast cancer survivors. J Cancer Surviv 2017 Feb 13;11(1):41-47 [FREE Full text] [doi: 10.1007/s11764-016-0558-z] [Medline: 27412726]
77. Connor KM, Davidson JR. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). Depress Anxiety 2003 Sep 02;18(2):76-82. [doi: 10.1002/da.10113] [Medline: 12964174]
78. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the Connor-davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. J Trauma Stress 2007 Dec 21;20(6):1019-1028. [doi: 10.1002/jts.20271] [Medline: 18157881]
79. Mewton L, Kessler RC, Slade T, Hobbs MJ, Brownhill L, Birrell L, et al. The psychometric properties of the Kessler Psychological Distress Scale (K6) in a general population sample of adolescents. Psychol Assess 2016 Oct;28(10):1232-1242. [doi: 10.1037/pas0000239] [Medline: 26619095]
80. Umucu E, Fortuna K, Jung H, Bialunska A, Lee B, Mangadu T, et al. A national study to assess validity and psychometrics of the short Kessler Psychological Distress Scale (K6). Rehabil Couns Bull 2022 Jan 03;65(2):140-149. [doi: 10.1177/00343552211043261] [Medline: 39301105]
81. Stone AA, Broderick JE, Junghaenel DU, Schneider S, Schwartz JE. PROMIS fatigue, pain intensity, pain interference, pain behavior, physical function, depression, anxiety, and anger scales demonstrate ecological validity. J Clin Epidemiol 2016 Jun;74:194-206. [doi: 10.1016/j.jclinepi.2015.08.029] [Medline: 26628334]
82. Holley AL, Zhou C, Wilson AC, Hainsworth K, Palermo TM. The CALI-9: a brief measure for assessing activity limitations in children and adolescents with chronic pain. Pain 2018 Jan 18;159(1):48-56 [FREE Full text] [doi: 10.1097/j.pain.0000000000001063] [Medline: 28957835]
83. Sufrinko AM, Valrie CR, Lanzo L, Bond KE, Trout KL, Ladd RE, et al. Empirical validation of a short version of the adolescent sleep-wake scale using a sample of ethnically diverse adolescents from an economically disadvantage community. Sleep Med 2015 Oct;16(10):1204-1206 [FREE Full text] [doi: 10.1016/j.sleep.2015.07.002] [Medline: 26429746]
84. Huber NL, Nicoletta A, Ellis JM, Everhart DE. Validating the adolescent sleep wake scale for use with young adults. Sleep Med 2020 May;69:217-219. [doi: 10.1016/j.sleep.2020.01.021] [Medline: 32179455]
85. Bromberg MH, de la Vega R, Law EF, Zhou C, Palermo TR. Development and validation of the adolescent insomnia questionnaire. J Pediatr Psychol 2020 Jan 01;45(1):61-71 [FREE Full text] [doi: 10.1093/jpepsy/jsz073] [Medline: 31584648]
86. Iverson GL, Connors EJ, Marsh J, Terry DP. Examining normative reference values and item-level symptom endorsement for the Quality of Life in Neurological Disorders (Neuro-QoL™) v2.0 cognitive function-short form. Arch Clin Neuropsychol 2021 Jan 15;36(1):126-134. [doi: 10.1093/arclin/acaa044] [Medline: 32851403]

Abbreviations

ABM: attention bias modification
ABCs: Attention Bias in Cancer survivor
DSM-5: Diagnostic and Statistical Manual of Mental Disorders-5
NeuroQoL: quality of life in neurological disorders
PROMIS: Patient-Reported Outcomes Measurement Information System
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SUS: System Usability Scale

Edited by A Schwartz; submitted 19.Aug.2025; peer-reviewed by M Jong; comments to author 08.Dec.2025; revised version received 15.Jan.2026; accepted 23.Jan.2026; published 25.Feb.2026.

Please cite as:
Lau N, Zhou C, Hong SJH, Aalfs H, Higgins S, Badillo I, McCauley E, Ketterl T, Chow EJ, Fann JR, Heathcote LC, Palermo TM
Digital Attention Bias in Cancer Survivors Intervention for Adolescent and Young Adult Cancer Survivors: Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2026;15:e82665
URL: https://www.researchprotocols.org/2026/1/e82665
doi:10.2196/82665
PMID:

XSL•FO
RenderX

©Nancy Lau, Chuan Zhou, Shannon JH Hong, Homer Aalfs, Shannon Higgins, Isabel Badillo, Elizabeth McCauley, Tyler Ketterl, Eric J Chow, Jesse R Fann, Lauren C Heathcote, Tonya M Palermo. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Supportive Group Intervention for Caregivers to Patients Diagnosed With Glioblastoma: Protocol for the SUGRI Study

Sara Nordentoft[1], RN, MHSc; Rikke Guldager[2], RN, PhD; Helle Pappot[3], MD; Tiit Mathiesen[4], MD; Anders Tolver[5], PhD; Karin Piil[6,7], RN, PhD

[1]Copenhagen University Hospital, Copenhagen, Denmark and Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark., Copenhagen Ø, Denmark

[2]Copenhagen University Hospital, Copenhagen, Denmark. Dept. of People and Technology, Roskilde University, Roskilde, Denmark, Copenhagen University Hospital, Denmark

[3]Copenhagen University Hospital and University of Copenhagen, Copenhagen, Denmark and Department of clinical medicine, University of Copenhagen, Copenhagen, Denmark, copenhagen, Denmark

[4]Department of Neurosurgery, Copenhagen University Hospital; Department of clinical medicine, University of Copenhagen, Copenhagen, Denmark; Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden, Copenhagen Ø, Denmark

[5]Statistics and Data Analysis, Danish Cancer Instititute, Danish Cancer Institute, Copenhagen, Denmark, Copenhagen Ø, Denmark

[6]Copenhagen University Hospital, Copenhagen and Department of clinical medicine, University of Copenhagen, Denmark, Copenhagen, Denmark

[7]Curtin University, Perth, Western Australia, Australia

Corresponding Author:
Sara Nordentoft, RN, MHSc
Copenhagen University Hospital, Copenhagen, Denmark and Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark.
Inge Lehmanns Vej 8, 2100
Copenhagen
Copenhagen Ø, 2100
Denmark
Phone: 45 29866734
Email: sara.marie.juel.nordentoft@regionh.dk

Abstract

Background: Caregivers to a person diagnosed with a glioblastoma often face significant responsibilities, balancing the demands of care with the complexities of the disease and treatment trajectory, while also coping with concerns and uncertainty of the future. Caregivers report unmet needs of information and support throughout the patient's disease and treatment trajectory, and they may benefit from targeted supportive care interventions.

Objective: The aim of this study is to develop, test, and evaluate the feasibility of a supportive group intervention among caregivers to patients diagnosed with a glioblastoma.

Methods: The study consists of 3 phases with ongoing patient and public involvement (PPI). In the first phase, a systematic review will be carried out exploring the outcome of supportive group interventions for caregivers to patients diagnosed with a primary brain tumor. In the second phase, the design and development of the intervention will be conducted in cooperation with a panel of caregivers using PPI process. The third phase will include the feasibility and evaluation of the intervention. The study will be guided by the British Medical Research Council's framework for developing complex interventions. The feasibility of the intervention will be tested in reference to relevant parameters. Quantitative data in terms of questionnaires will be analyzed using descriptive statistics, and qualitative evaluation interviews will be analyzed using thematic analysis.

Results: A national panel of caregivers (N=10) has been established, and the final design of the intervention is currently developed through an ongoing PPI process. A total of 3 hospitals in Denmark have committed to participating in the recruitment for the feasibility study. Recruitment began in April 2025, and the SUGRI (Supportive Group Intervention for caregivers to patients diagnosed with a glioblastoma) intervention will undergo a feasibility assessment in a multicenter study starting in August 2025, with the final evaluation planned for April 2026.

Conclusions: This study is expected to provide necessary insights to guide caregiver initiatives, ultimately improving support for caregivers within the neuro-oncology field. A supportive group intervention offered to the caregivers has the potential to address specific caregiver needs and strengthen their supportive network. Providing caregivers with support may enhance their

XSL•FO
RenderX

perceived support, strengthen family functioning, and provide them with strategies to manage caregiving challenges, ultimately benefiting both caregivers and patients.

Trial Registration: ClinicalTrials.gov NCT06869577; https://clinicaltrials.gov/study/NCT06869577

International Registered Report Identifier (IRRID): PRR1-10.2196/81948

(JMIR Res Protoc 2026;15:e81948) doi:10.2196/81948

KEYWORDS

caregiver; glioblastoma; group intervention; high-grade glioma; neuro-oncology; supportive care

Introduction

Primary malignant brain tumors are among the most fatal cancers and have an international incidence rate of 6.89 per 100,000, compared with 7.03 per 100,000 in Denmark [1,2]. The most common malignant brain tumor is a glioblastoma (GBM), accounting for 51.5% of all malignant brain tumors. It is associated with a poor prognosis: a 5-year relative survival rate of 5% [1]. The typical disease trajectory of 12 to 15 months includes initial surgical intervention (resection or biopsy), followed by a 9- to 10-month standard treatment plan consisting of concomitant chemotherapy or radiation therapy, with subsequent adjuvant chemotherapy of temozolomide [3]. Despite aggressive treatment, tumor progression occurs almost inevitably [4].

The diagnosis of a GBM is usually preceded by symptoms such as an epileptic seizure, focal neurological deficits, or cognitive changes that may affect the patient's functional status [2]. Patients often develop a complex symptom burden, including physical symptoms such as aphasia and hemiparesis and psychological symptoms such as cognitive deficits, anxiety, distress, and depression. The symptoms may change as tumor progresses [5-7]. The combination of a rapid onset of the disease, the intense disease and treatment trajectories, and the uncertainty of the future affect not only the individual but also those close to them who function as informal caregivers, a role that often demands great caring responsibilities [8,9]. The caregiver burden is described as multidimensional with components affecting caregivers physically, emotionally, socially, and economically [10]. Caregivers express being unprepared for the new role and report symptoms of anxiety and depression [4,9]. Additionally, social isolation is often a consequence of the caregiver role. Furthermore, the *invisibility* of the patients' symptoms can lead to withdrawal of support from family and previous networks [4,11]. Caregivers emphasize the importance of meeting other caregivers in similar situations so that they can share experiences and recognize emotions or concerns in a peer-to-peer environment [12,13]. Peer support may represent an important access to information, advice, and practical help for caregivers and address caregivers' individual needs of support [12,14,15]. It is often a struggle for caregivers to be separated from the patient, as they worry about risk of seizures or cognitive symptoms that necessitate constant supervision, a struggle that can be alleviated with support [16].

Online access to peer support and information may be valuable for caregivers. Such a tool is accessible and has the potential to ensure broad reach while remaining scalable at the national level. Moreover, online support aligns well with the recommendations of the International Council of Nurses, which indicate that digital health can address population needs, strengthen health systems, and remain feasible despite shortages in the health workforce [17].

Several authors address the lack of longitudinal studies and advocate for ongoing supportive interventions to target caregivers of patients diagnosed with a brain tumor [10,18,19]. A need of such interventions is further echoed in international research agendas [20,21]. A systematic review from 2021 concluded feasibility of interventions for people with a brain tumor and their caregivers [14]. Importantly, support groups exclusively for caregivers may fulfill their need for peer support [14,22]. Therefore, this study aims to develop, test, and evaluate the feasibility of a supportive group intervention among caregivers to patients diagnosed with a GBM.

Methods

Study Design

This study titled SUGRI (Supportive Group Intervention for caregivers to patients diagnosed with a glioblastoma) is a 3-phase study. The overall framework is guided by the British Medical Research Council (MRC) framework for developing complex intervention. This methodological framework entails 4 main phases of intervention research: development or identification of the intervention, feasibility, evaluation, and implementation. In this study, we develop, test, and evaluate the feasibility of an intervention; therefore, we are primarily guided by the first 2 phases of the MRC framework: the development phase and the feasibility phase [23-25]. The phases of the MRC framework applied to the SUGRI study are presented in Figure 1. An overview of the intervention logic, including planned activities and expected outcomes, is presented in Figure 2. A patient and public involvement (PPI) plan will be carried out throughout all phases of the SUGRI study, guided by the workbook to guide the development of a Patient Engagement in Research plan [26]. The recommended checklist for research protocols, SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials), has been applied (Multimedia Appendix 1) [27].

XSL•FO
RenderX

Figure 1. Complex intervention for the SUGRI (Supportive Group Intervention for caregivers to patients diagnosed with a glioblastoma) study. Solid lines represent this study's activities; dashed lines represent planned future activities. HCP: health care professional.

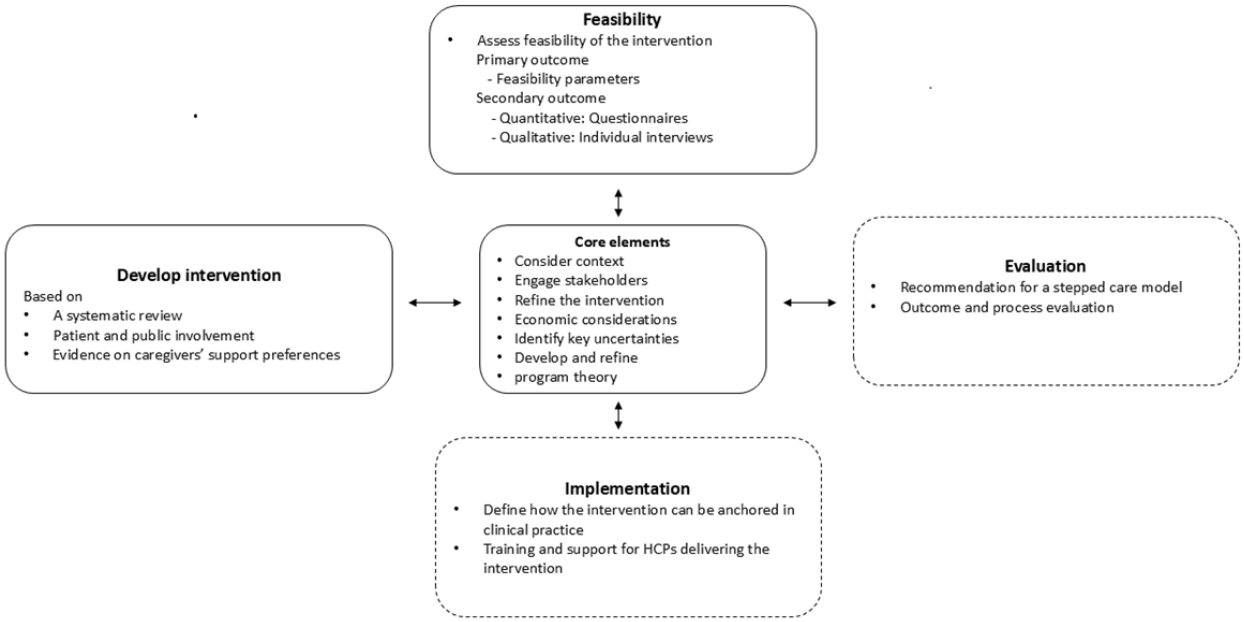

Figure 2. Logic model. Solid lines represent planned work; dashed lines represent intended results. HCP: health care professional; PEIR: Patient Engagement in Research; PPI: patient and public involvement; REDCap: Research Electronic Data Capture; SUGRI: Supportive Group Intervention for caregivers to patients diagnosed with a glioblastoma.

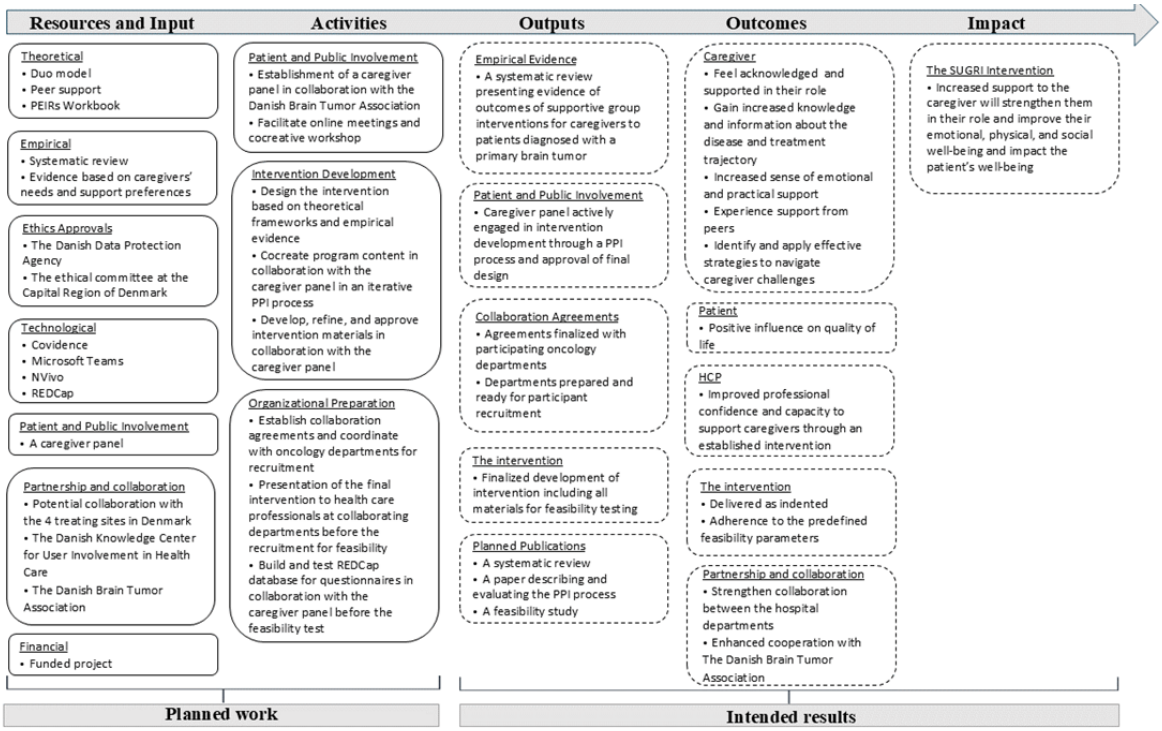

Study Population and Eligibility Criteria

The term "caregivers" is in this study defined according to the criteria provided by The National Cancer Institute [28]: "Informal caregiving is broadly defined as services provided by an unpaid person, such as helping with personal needs and household chores, managing a person's finances, arranging for outside services, or visiting regularly. An informal caregiver is usually a relative or friend who may or may not live in the same household as the person with cancer who requires care." Inclusion criteria are family caregivers (aged ≥18 years) to patients (aged ≥18 years) diagnosed with a GBM grade 4 or diffuse astrocytoma grade 4 receiving chemoradiotherapy treatment. Despite differential terminology of GBM grade 4 and diffuse astrocytoma grade 4 since 2021 [29], the study will

XSL•FO
RenderX

include both categories of grade 4 tumors; both are referred to as "GMB" in the protocol. Participants must be able to understand, read, and speak Danish.

Recruitment Procedure

Participants will be recruited from the Department of Neuro-oncology at Copenhagen University Hospital in Denmark. The feasibility of the intervention is planned to be evaluated nationally, and the 3 other departments of neuro-oncology in Denmark at Odense University Hospital, Aarhus University Hospital, and Aalborg University Hospital will be invited to participate in the recruitment procedure. A purposive sampling strategy will be applied, and participants will be recruited face-to-face.

Study Phases

Exploratory Phase

A systematic review will identify and explore available evidence of outcomes of supportive group interventions for caregivers to patients diagnosed with a primary brain tumor. To define the group element of an intervention, we decided to include "any kind of intervention, that sought to facilitate some element of interaction between a minimum of two caregivers." The search will be conducted in the following databases: PubMed, Embase, Web of Science, Emcare, Cochrane Library, and PsycInfo. Peer-reviewed studies in English, German, or Scandinavian will be included without any time limits. The software program Covidence (Veritas Health Innovation Ltd) will support the process of screening [30]. Guidance on the conduct of narrative synthesis in systematic reviews will be used to interpret findings [31]. To address the quality of the included studies, the Mixed Methods Appraisal Tool will be applied [32]. To ensure transparency, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guideline will be used [33].

Development Phase

According to the MRC framework, the development phase focuses on designing the intervention based on existing evidence and relevant theories, engaging stakeholders, and refining the intervention to ensure contextual fit and relevance to the target population [25]. Therefore, the development of the SUGRI intervention will be grounded in existing evidence from the neuro-oncology field, informed by the findings from the systematic review, and be carried out in close collaboration with stakeholders.

PPI Activities

PPI will be central in the development phase, where a national panel of current and bereaved caregivers to patients with a GBM will be established. A panel consisting of 6 to 12 members will allow for deeper interaction and a more trusting environment. Members will be recruited nationally through the Danish Brain Tumor Association. The intervention will be developed and designed in collaboration with a panel of caregivers, using Microsoft Teams as the meeting platform. Online meetings will be scheduled for every second month, and 1 cocreative workshop requiring physical attendance will be planned and facilitated in cooperation with the Danish Knowledge Center for User Involvement in Health Care. The PPI process will be evaluated through a focus group interview with the panel members, guided by the approach described by Krueger and Casey [34]. For the quantitative evaluation, the 22-item Patient Engagement in Research Scale will be used to measure the level of meaningful caregiver engagement in the research process [35,36].

Intervention

The intervention will be conducted online and will encompass 2 primary components: an informational element and a group-based element. The final design of the intervention and how the 2 central elements will be incorporated will be developed through an iterative PPI process in cooperation with the panel to integrate existing evidence on caregivers' needs and preferences. Results from the systematic review from the exploratory phase will also guide the design of the final intervention. The group element will be designed and inspired by existing knowledge of caregiver preferences of supportive group interventions [11,22].

Feasibility and Evaluation Phase

The feasibility phase in the MRC framework includes assessment of the feasibility and acceptability of the intervention. It also involves evaluation of the design and the predefined feasibility parameters, addressing uncertainties identified during development, refining the intervention, and considering practical and economic aspects [25]. In the third study phase, the online intervention will be tested on a new group of caregivers, and feasibility will be analyzed and evaluated by integration of qualitative and quantitative methods. A sample size of approximately 30 participants is recommended for feasibility trials [37].

Primary Objective and End Points

We hypothesize that offering caregivers a supportive group intervention will empower them in their role by boosting their perceived support, strengthening their family function, and helping them to identify strategies to navigate the challenges of caregiving, ultimately benefiting both caregivers and patients.

The primary objective of this study is to evaluate the feasibility of the SUGRI intervention. The following parameters will be evaluated: recruitment, fidelity, adherence, acceptability, and safety (detailed in Table 1).

XSL•FO
RenderX

Table 1. Progression criteria and feasibility parameters.

Study phase	Study	Aim and primary assessment	Data collection or assessment method	Progression criteria or predefined benchmarks
Exploratory phase	Systematic review	• Aim: to identify and explore available evidence of outcome of supportive group interventions for caregivers to patients with a primary brain tumor • Primary assessment: evidence from peer-reviewed literature	• Data synthesis: guidance on the conduct of narrative synthesis in systematic reviews [31] • Quality assessment: the Mixed Methods Appraisal Tool [32] • PRISMAa guidelines [33]	• Proceed to the development phase if: • Review confirms a gap for group interventions of the target population • Intervention already exists, reconsider scope or adapt intervention • Review identifies sufficient evidence to inform group interventions to be positively evaluated
Development phase	PPI[b] study	• Aim: to involve caregivers in the design and development of a supportive group intervention targeted at caregivers to patients diagnosed with a glioblastoma and evaluate the PPI process • Primary assessment: caregivers' experiences and perceived meaningful impact of participation	• Qualitative data: focus group interview guidelines by Krueger and Casey [34] • Quantitative data: 22-item Patient Engagement in Research Scale [35,36]	• Proceed to feasibility and evaluation phase if: • A caregiver panel can be successfully established • Intervention can be realistically developed in cooperation with the panel, integrating evidence and caregiver input
Feasibility and evaluation phase	Feasibility study	• Aim: to test the feasibility of a supportive group intervention among caregivers to patients diagnosed with a glioblastoma • Primary assessment: predefined feasibility parameters: • Recruitment: is it feasible to recruit participants to the intervention? • Fidelity: is the intervention provided as designed? • Adherence: how feasible is it to ensure participants adhere to the intervention? • Acceptability: what are the experiences and ratings of satisfaction with the intervention? • Safety: is the intervention safe for participants?	• Qualitative data: semistructured individual interviews with caregivers • Quantitative data: patient- and caregiver-reported outcome • Quantitative registration of recruitment rate • Participants are offered the intervention • Quantitative registration of intervention participation and participant feedback via study-specific questionnaires and interviews • Participant feedback via study-specific questionnaires and interviews • Adverse events from validated questionnaires and interviews	• Proceed to development of a stepped care model if: • All 5 feasibility parameters meet their predefined benchmarks • Estimated recruitment ≥60%c • Estimated fidelity 100%c • Estimated adherence ≥70%c • Estimated acceptability mean ≥4d • Safety reported narrativelye

[a]PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

[b]PPI: patient and public involvement.

[c]Recruitment rate, fidelity, and adherence are presented as estimates informed by other caregiver interventions and reported as percentages.

[d]Acceptability is defined as an average score of ≥4 on a 5-point Likert scale in a study-specific questionnaire. The questionnaire includes the item: "Overall, I found the caregiver program to be supportive," rated from 1=not at all to 5=to a very great extent.

[e]Safety will be reported narratively based on individual postintervention evaluation interviews with caregivers, guided by the following question: "Have you experienced any concerns as a result of your participation that you did not have before?"

The proposed benchmarks for the 5 feasibility parameters have been guided by existing literature on caregiver interventions [38-41]. Predefined progression criteria, as proposed by the MRC framework, will be assessed through a feasibility study aligned with the evaluation design [24]. Moreover, progression criteria are used to guide the decision on whether to proceed to the next phase of the project [25]. Progression criteria for the 3 study phases—the exploratory phase, the development phase, and the feasibility and evaluation phase—are also presented in Table 1.

XSL•FO
RenderX

Measurements

Secondary outcomes are patient-reported outcome data completed by caregivers and patients at baseline, postintervention, and follow-up (Table 1). Caregivers will complete the following 3 questionnaires: the Caregiver Roles and Responsibilities Scale to assess broad life impacts for caregivers [42], the Iceland-Family Perceived Support Questionnaire to measure families' perceived support from nurses and other health care professionals [43], and the Hospital Anxiety and Depression Scale to measure symptoms of anxiety and depression [44]. Patients will complete the following 3 questionnaires: the Functional Assessment of Cancer Therapy-Brain to assess the patients' health-related quality of life [45]; the MD Anderson Symptom Inventory-Brain Tumor Module to measure patients' symptom prevalence, intensity, and interference with the daily life [46]; and Hospital Anxiety and Depression Scale [44]. The questionnaires and time assessment are presented in Table 2 and will be collected electronically using Research Electronic Data Capture (REDCap; Vanderbilt University). Demographic data of the participants will be completed at baseline, and information on patient histology will be obtained from the patient's medical journal and stored in REDCap.

Table 2. Measurement tools and assessment time points for the feasibility study.

Outcome measurement tools	Caregiver			Patient		
	T0[a]	T1[b]	T2[c]	T0	T1	T2
Quantitative						
Informed consent	✓[d]			✓		
Demographic data	✓			✓		
CRRS[e] 41 items (0-164 points)	✓	✓	✓			
ICE-FPSQ[f] 14 items (14-70 points)	✓	✓	✓			
HADS[g] 14 items (0-21 points)	✓	✓	✓	✓	✓	✓
FACT-Br[h] 50 items (0-200 points)				✓	✓	✓
MDASI-BT[i] 13 core cancer-related symptoms and 9 brain tumor–specific symptoms (0-220 points)				✓	✓	✓
Qualitative						
Semistructured interview		✓				

[a]T0: baseline.

[b]T1: postintervention week 12.

[c]T2: follow-up week 18.

[d]Time points for data collection.

[e]CRRS: Caregiver Roles and Responsibilities Scale.

[f]ICE-FPSQ: Iceland-Family Perceived Support Questionnaire.

[g]HADS: Hospital Anxiety and Depression Scale.

[h]FACT-Br: Functional Assessment of Cancer Therapy-Brain.

[i]MDASI-BT: MD Anderson Symptom Inventory-Brain Tumor Module.

To explore experiences and evaluate the intervention qualitatively, semistructured individual interviews will be carried out with the participating caregivers postintervention. A semistructured interview guide will be applied, and data will be audiotaped and verbatim transcribed.

Statistical Analysis

Results regarding the feasibility of the intervention and the attainment of predefined feasibility parameters will be presented separately. Feasibility parameters and secondary outcomes, including subscale scores from patient-reported outcomes, will be presented using descriptive statistics. Mean change from baseline to later assessment times will be analyzed using appropriate statistical methods, possibly a linear mixed model to account for repeated measurements on each participant. This will further ensure that no imputation is required to handle missing data under the missing-at-random assumption. No formal adjustment of *P* values to control the number of false positives will be used, as analyses of secondary outcomes should be regarded as exploratory. Patient-reported outcome data will provide valuable insight into the impact of the caring responsibilities, perceived support, and the caregiver role combined with the patient's well-being, with a particular emphasis on their emotional and functional challenges.

Qualitative Analysis

Qualitative interview data will be analyzed using thematic analysis, as described by Braun and Clarke [47,48], using the NVivo software (Lumivero) [49]. The interview data will provide the opportunity to evaluate specific elements of the intervention, contributing to a more comprehensive understanding of how different aspects of the intervention

XSL•FO
RenderX

impact caregivers. The qualitative and quantitative data will be collected and analyzed separately and then merged. In the overall interpretation, the 2 datasets will be compared, and we will search for potential common patterns [50-52].

Ethical Considerations

The study adheres to the ethical standards of the institutional and national research committee and complies with the Declaration of Helsinki. The study was approved by the ethical committee of the Capital Region of Denmark (F-24004933) and registered with the Danish Data Protection Agency (Jr. P-2022-739). No compensation is provided for participation. Potential ethical risks include breaches of privacy or integrity, while other types of harm are considered unlikely. Study participants as well as future patients and caregivers may benefit from the study. Measures will be taken to protect integrity. Participants will be enrolled based on oral and written informed consent, which can be withdrawn at any time. All participants will provide written informed consent before enrollment and will be informed about procedures related to confidentiality, safeguarding, and support throughout the intervention. Clear rules for confidentiality will be established at the beginning of the online group sessions, including reminders to not share identifiable information outside the group. To ensure participant well-being, facilitators will be attentive to recognizing signs of emotional distress during the sessions. If a participant appears overwhelmed, they will be offered a one-on-one follow-up conversation and, if necessary, referred to appropriate support services.

Study Registration

The study is registered on ClinicalTrials.gov (NCT06869577), and the systematic review is registered in PROSPERO (CRD42023392222).

Results

A national panel of caregivers (N=10) has been successfully established, and the final design of the intervention is currently being developed through an ongoing PPI process. A total of 3 hospitals in Denmark have committed to participating in the recruitment for the feasibility study. Recruitment began in April 2025. The multicenter feasibility assessment of the SUGRI intervention is scheduled to start in autumn 2025, and the final evaluation is planned for spring 2026

Discussion

Positioning Supportive Caregiver Care Within the Neuro-Oncology Landscape

Caregiver support within the neuro-oncology field remains suboptimal. This paper describes a study protocol for a supportive group intervention for caregivers to patients diagnosed with a GBM. The intervention is designed to address the challenges faced by caregivers, with particular emphasis on their need for information and peer support. We expect that participation in a group intervention will empower caregivers by enhancing their perceived social support, strengthening family functioning, and helping them identify strategies to navigate caregiving challenges—ultimately benefiting both caregivers and patients.

Caregivers may benefit from continuous access to support throughout the patient's treatment trajectory, tailored to address the specific emotional challenges and stressors that caregivers face in their role as a caregiver [53]. Previous interventions have been carried out using a dyadic approach, where patients and caregivers receive the intervention together, allowing for mutual support [54,55]. However, Mulbury et al [56] examined preferences for intervention delivery, comparing a dyadic format, involving both patient and caregiver, with a caregiver-only format and found that the caregiver-only approach was significantly more beneficial for caregivers. This supports the findings of Mallya et al [22], who reported a clear preference among caregivers for dedicated support groups separate from those involving the patient. Group-based interventions offer caregivers a chance to connect with others in similar situations, creating a safe and supportive environment for sharing emotions, concerns, and experiences [13,22]. Peer support may also provide valuable practical advice, help normalize experiences, and offer new perspectives on personal caregiving experiences [13,14]. Group-based interventions may also offer validation, a sense of belonging, and foster hope [4].

Given the diverse nature of caregiver needs, shaped by demographic, course of the disease, and patient symptom burden, future research may benefit from moving beyond standardized approaches and explore the potential of multicomponent interventions [14,22]. Therefore, the MRC framework may provide an overall structured framework designed to guide this study. An intervention may be described as complex when it includes multiple components interacting with each other and allows for flexibility or tailoring [25]. The SUGRI intervention is planned to include 2 primary components: an informational element and a group-based element. Integration of these 2 components will be developed through a PPI process.

The updated MRC framework from 2021 includes 6 core elements to be addressed throughout the research process, one of which is engaging stakeholders [24]. Engagement of the stakeholders meaningfully at every stage of the research process is essential to develop interventions that positively impact health outcomes for patients and caregivers [24]. In the SUGRI project, we plan to establish a PPI collaboration with a panel of current and bereaved caregivers throughout the project. The MRC framework recommends involving stakeholders early in the research process [25]; however, out of respect for the caregivers' time, we have chosen not to include them at the very beginning during the protocol stage. The protocol, while serving as the foundation for the project, is intentionally described in broad terms, as the detailed design of the intervention will be developed collaboratively with the caregiver panel. Given the poor prognosis of a GBM, we anticipate the panel to be dynamic, with caregivers leaving as patients enter end-of-life phases and new caregivers are invited to join. Most of the panel activities are scheduled to be online, as caregivers may have practical difficulties to schedule physical meetings due to their caregiver role.

XSL•FO
RenderX

The feasibility test of the SUGRI intervention aims to address caregivers' supportive needs throughout the patient's disease and treatment trajectory. Moreover, we anticipate the SUGRI intervention to inform how a tailored intervention can be both feasible and meaningful within the context of caregivers' daily lives.

Strengths and Limitations

A major strength of the project is that it addresses an evident and urgent need for support among caregivers of patients with GBM, a need that has also been expressed by clinicians seeking structured ways to assist caregivers in navigating this profoundly life-changing situation. A limitation of this protocol is that the intervention has not yet been fully developed. Consequently, there is uncertainty regarding its specific content, structure, and timing, which will determine the feasibility and overall evaluation of the intervention.

The study incorporates PPI, which represents a methodological strength by ensuring the intervention is tailored to the target population and enhances its relevance and acceptability. The project also includes an evaluation of the PPI process itself, thereby contributing to the validation of the intervention from the perspective of the involved caregivers.

The guidance of the MRC framework is another strength, as it provides a structured and evidence-informed approach to intervention development. It supports transparency, theoretical grounding, and relevance to the target population, thereby enhancing the quality and potential impact of the final intervention.

Future Directions

Health care resources should be used in an efficient manner to ensure the greatest impact. At the same time, it is essential that support efforts are tailored to the individual needs, particularly in light of the evolving shortages in the health workforce [17]. Moreover, caregiver support is likely to be cost-effective, providing societal benefits by improving health outcomes. Cost-benefit of caregiver intervention is not part of this protocol but would be interesting to evaluate.

Stepped care is a health care delivery model that allocates resources efficiently by matching the level of care to the individual's needs [57,58]. Depending on the design of the SUGRI intervention, a stepped care model may be beneficial for the future direction of the study. In a stepped care approach, the intensity of support provided to caregivers can be adjusted according to their individual needs and preferences, ensuring that they receive appropriate and timely support. Applying a stepped care model to future caregiver interventions may enhance their adaptability to hospital workflows and better address the individual needs of caregivers.

Dissemination Plan

Dissemination of results includes publication of articles in peer-reviewed journals: a systematic review, a PPI article, and an article on the feasibility of the intervention. Moreover, results will be presented at scientific conferences, to clinicians within the neuro-oncology field, and to the panel of caregivers who participated in the PPI process.

Conclusions

The SUGRI study will develop, test, and evaluate the feasibility of a supportive group intervention. The intervention has the potential to address specific caregiver needs and strengthen their supportive network by offering a supportive group intervention to the caregivers. Only few group interventions targeting the neuro-oncology caregiver have been carried out, and results may contribute with new internationally relevant knowledge to improve conditions and care for patients with GBM and their caregivers. Providing caregivers with support may enhance their perceived support, strengthen family functioning, and provide them with strategies to manage caregiving challenges, ultimately benefiting both caregivers and patients.

Clinical Implications

Family members of patients diagnosed with a GBM often assume caregiver roles and take on a great responsibility. The SUGRI study accepts this fact and assumes caregivers not only not only need support but also represent an untapped resource. An online intervention may enhance support for caregivers, offering flexibility and helping to reduce the demand on clinical staff. This approach underscores the need for labor-saving technologies to address the nursing shortage and improve patient care [17]. To the best of our knowledge, a group intervention has not yet been systematically offered to caregivers of patients with a GBM in a Danish context, making its feasibility and outcomes of significant interest to the field of neuro-oncological caregiver research.

Acknowledgments

The authors would like to acknowledge the caregiver panel for their valuable cooperation in the SUGRI (Supportive Group Intervention for caregivers to patients diagnosed with a glioblastoma) project. Moreover, the authors would like to gratefully acknowledge the established collaboration with the national treatment sites. We used the generative AI tools Microsoft Copilot and ChatGPT for language refinement, including grammar checks and readability improvement. All content was subsequently reviewed and revised by the authors.

Funding

This work was supported by the Copenhagen University Hospital, Rigshospitalet (grant E-23565-08); the Danish Cancer Society (grant R344-A19341); and the Novo Nordisk Foundation (grant NNF22OC0080425) and is a part of the Models of Cancer Care research program.

XSL•FO
RenderX

Authors' Contributions

Conceptualization and design of the project: SN, KP
Development of protocol: SN, KP, RG
Review and feedback on protocol: HP, TM
Drafting of the manuscript: SN, KP, RG
Critical revision of the manuscript: all authors
Final approval of the manuscript: all authors

Conflicts of Interest

None declared.

Multimedia Appendix 1
SPIRIT 2025 checklist of items to address in a randomized trial protocol.
[PDF File (Adobe PDF File), 293 KB - resprot_v15i1e81948_app1.pdf]

References

1. Price M, Ballard C, Benedetti J, Neff C, Cioffi G, Waite KA, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2017-2021. Neuro Oncol 2024 Oct 06;26(Supplement_6):vi1-vi85. [doi: 10.1093/neuonc/noae145] [Medline: 39371035]
2. Rasmussen BK, Hansen S, Laursen RJ, Kosteljanetz M, Schultz H, Nørgård BM, et al. Epidemiology of glioma: clinical characteristics, symptoms, and predictors of glioma patients grade I-IV in the the Danish Neuro-Oncology Registry. J Neurooncol 2017 Dec;135(3):571-579. [doi: 10.1007/s11060-017-2607-5] [Medline: 28861666]
3. Stupp R, Hegi ME, Mason WP, van den Bent MJ, Taphoorn MJ, Janzer RC, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol 2009 May;10(5):459-466. [doi: 10.1016/S1470-2045(09)70025-7] [Medline: 19269895]
4. Piil K, Laegaard Skovhus S, Tolver A, Jarden M. Neuro-oncological symptoms: a longitudinal quantitative study of family function, perceived support, and caregiver burden. J Fam Nurs 2022 Feb 21;28(1):43-56. [doi: 10.1177/10748407211029986] [Medline: 34286624]
5. Molassiotis A, Wilson B, Brunton L, Chaudhary H, Gattamaneni R, McBain C. Symptom experience in patients with primary brain tumours: a longitudinal exploratory study. Eur J Oncol Nurs 2010 Dec;14(5):410-416. [doi: 10.1016/j.ejon.2010.03.001] [Medline: 20363189]
6. Noll K, King AL, Dirven L, Armstrong TS, Taphoorn MJ, Wefel JS. Neurocognition and health-related quality of life among patients with brain tumors. Hematol Oncol Clin North Am 2022 Feb;36(1):269-282. [doi: 10.1016/j.hoc.2021.08.011] [Medline: 34711455]
7. Armstrong TS, Dirven L, Arons D, Bates A, Chang SM, Coens C, et al. Glioma patient-reported outcome assessment in clinical care and research: a Response Assessment in Neuro-Oncology collaborative report. Lancet Oncol 2020 Feb;21(2):e97-103. [doi: 10.1016/S1470-2045(19)30796-X] [Medline: 32007210]
8. Schmer C, Ward-Smith P, Latham S, Salacz M. When a family member has a malignant brain tumor: the caregiver perspective. J Neurosci Nurs 2008 Apr;40(2):78-84. [doi: 10.1097/01376517-200804000-00006] [Medline: 18481737]
9. Schubart JR, Kinzie MB, Farace E. Caring for the brain tumor patient: family caregiver burden and unmet needs. Neuro Oncol 2008 Feb;10(1):61-72 [FREE Full text] [doi: 10.1215/15228517-2007-040] [Medline: 17993635]
10. Bayen E, Laigle-Donadey F, Prouté M, Hoang-Xuan K, Joël ME, Delattre JY. The multidimensional burden of informal caregivers in primary malignant brain tumor. Support Care Cancer 2017 Jan 13;25(1):245-253. [doi: 10.1007/s00520-016-3397-6] [Medline: 27624465]
11. Cubis L, Ownsworth T, Pinkham MB, Chambers S. The social trajectory of brain tumor: a qualitative metasynthesis. Disabil Rehabil 2018 Aug 19;40(16):1857-1869. [doi: 10.1080/09638288.2017.1315183] [Medline: 28420297]
12. Nordentoft S, Dieperink KB, Johansson SD, Jarden M, Piil K. Evaluation of a multimodal rehabilitative palliative care programme for patients with high-grade glioma and their family caregivers. Scand J Caring Sci 2022 Sep 23;36(3):815-829. [doi: 10.1111/scs.13019] [Medline: 34296773]
13. Kanter C, D'Agostino NM, Daniels M, Stone A, Edelstein K. Together and apart: providing psychosocial support for patients and families living with brain tumors. Support Care Cancer 2014 Jan 29;22(1):43-52. [doi: 10.1007/s00520-013-1933-1] [Medline: 23989499]
14. Heinsch M, Cootes H, Wells H, Tickner C, Wilson J, Sultani G, et al. Supporting friends and family of adults with a primary brain tumour: a systematic review. Health Soc Care Community 2022 May 11;30(3):869-887. [doi: 10.1111/hsc.13586] [Medline: 34633723]

XSL•FO
RenderX

15. Au TH, Willis C, Reblin M, Peters KB, Nghiemphu PL, Taylor JW, et al. Caregiver burden by treatment and clinical characteristics of patients with glioblastoma. Support Care Cancer 2022 Feb 11;30(2):1365-1375 [FREE Full text] [doi: 10.1007/s00520-021-06514-0] [Medline: 34510238]
16. Paterson C, Roberts C, Li J, Chapman M, Strickland K, Johnston N, et al. What are the experiences of supportive care in people affected by brain cancer and their informal caregivers: a qualitative systematic review. J Cancer Surviv 2024 Oct 31;18(5):1608-1629 [FREE Full text] [doi: 10.1007/s11764-023-01401-5] [Medline: 37256499]
17. The future of nursing and digital health: new ICN position statement highlights opportunities and risks. International Council of Nurses. 2023 Sep 1. URL: https://www.icn.ch/news/future-nursing-and-digital-health-new-icn-position-statement-highlights-opportunities-and [accessed 2026-02-18]
18. Piil K, Juhler M, Jakobsen J, Jarden M. Controlled rehabilitative and supportive care intervention trials in patients with high-grade gliomas and their caregivers: a systematic review. BMJ Support Palliat Care 2016 Mar 02;6(1):27-34 [FREE Full text] [doi: 10.1136/bmjspcare-2013-000593] [Medline: 24890014]
19. Boele FW, Rooney AG, Bulbeck H, Sherwood P. Interventions to help support caregivers of people with a brain or spinal cord tumour. Cochrane Database Syst Rev 2019 Jul 02;7(7):CD012582 [FREE Full text] [doi: 10.1002/14651858.CD012582.pub2] [Medline: 31264707]
20. MacDonald L. Top 10 priorities for clinical research in primary brain and spinal cord tumours. Neuro-Oncology Group. 2015 Jun. URL: https://www.jla.nihr.ac.uk/media/4826/download [accessed 2026-02-18]
21. Piil K, Jarden M, Pii KH. Research agenda for life-threatening cancer. Eur J Cancer Care (Engl) 2019 Jan;28(1):e12935. [doi: 10.1111/ecc.12935] [Medline: 30345723]
22. Mallya S, Daniels M, Kanter C, Stone A, Cipolla A, Edelstein K, et al. A qualitative analysis of the benefits and barriers of support groups for patients with brain tumours and their caregivers. Support Care Cancer 2020 Jun;28(6):2659-2667. [doi: 10.1007/s00520-019-05069-5] [Medline: 31641869]
23. Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, et al. Designing and evaluating complex interventions to improve health care. BMJ 2007 Mar 03;334(7591):455-459 [FREE Full text] [doi: 10.1136/bmj.39108.379965.BE] [Medline: 17332585]
24. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
25. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. Health Technol Assess 2021 Sep;25(57):1-132 [FREE Full text] [doi: 10.3310/hta25570] [Medline: 34590577]
26. Workbook to guide the development of a Patient Engagement In Research (PEIR) Plan. Arthritis Research Canada. 2018. URL: https://www.arthritisresearch.ca/wp-content/uploads/2018/06/PEIR-Plan-Guide.pdf [accessed 2026-02-18]
27. Chan AW, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ 2025 Apr 28;389:e081477. [doi: 10.1136/bmj-2024-081477] [Medline: 40294953]
28. Informal caregivers in cancer: roles, burden, and support (PDQ®): health professional version. In: PDQ Cancer Information Summaries [Internet]. Bethesda, MD: National Cancer Institute (US); 2002.
29. Torp SH, Solheim O, Skjulsvik AJ. The WHO 2021 classification of central nervous system tumours: a practical update on what neurosurgeons need to know-a minireview. Acta Neurochir (Wien) 2022 Sep;164(9):2453-2464 [FREE Full text] [doi: 10.1007/s00701-022-05301-y] [Medline: 35879477]
30. Covidence. URL: https://www.covidence.org/ [accessed 2026-02-18]
31. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC Methods Programme. Lancaster University. 2006 Apr. URL: https://tinyurl.com/5cscchr7 [accessed 2026-02-18]
32. Hong QN, Pluye P, Fàbregues S, Bartlett G, Boardman F, Cargo M, et al. Mixed methods appraisal tool (MMAT) version 2018: user guide. Canadian Intellectual Property Office, Industry Canada. 2018. URL: http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf [accessed 2026-02-18]
33. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021 Mar 29;10(1):89 [FREE Full text] [doi: 10.1186/s13643-021-01626-4] [Medline: 33781348]
34. Krueger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research. Thousand Oaks, CA: SAGE Publications; 2008.
35. Hamilton CB, Hoens AM, McQuitty S, McKinnon AM, English K, Backman CL, et al. Development and pre-testing of the Patient Engagement In Research Scale (PEIRS) to assess the quality of engagement from a patient perspective. PLoS One 2018 Nov 1;13(11):e0206588 [FREE Full text] [doi: 10.1371/journal.pone.0206588] [Medline: 30383823]
36. Christiansen MG, Pappot H, Hamilton C, Lund-Jacobsen T, Kristensen AW, Nordentoft S, et al. Patient Engagement in Research Scale (PEIRS-22): Danish translation, applicability, and user experiences. Res Involv Engagem 2023 Dec 07;9(1):115 [FREE Full text] [doi: 10.1186/s40900-023-00526-2] [Medline: 38062535]

XSL•FO
RenderX

37. Billingham SA, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database. BMC Med Res Methodol 2013 Aug 20;13(1):104 [FREE Full text] [doi: 10.1186/1471-2288-13-104] [Medline: 23961782]
38. Whisenant M, Weathers SP, Li Y, Aldrich E, Ownby K, Thomas J, et al. Simulation-based caregiving skills training for family members of high-grade glioma patients. Neurooncol Pract 2024 Aug;11(4):432-440 [FREE Full text] [doi: 10.1093/nop/npae025] [Medline: 39006525]
39. Rooney AG, Hewins W, Walker A, Mackinnon M, Withington L, Robson S, et al. Lifestyle coaching is feasible in fatigued brain tumor patients: a phase I/feasibility, multi-center, mixed-methods randomized controlled trial. Neurooncol Pract 2023 Jun;10(3):249-260 [FREE Full text] [doi: 10.1093/nop/npac086] [Medline: 37188163]
40. Vranceanu AM, Bannon S, Mace R, Lester E, Meyers E, Gates M, et al. Feasibility and efficacy of a resiliency intervention for the prevention of chronic emotional distress among survivor-caregiver dyads admitted to the neuroscience intensive care unit: a randomized clinical trial. JAMA Netw Open 2020 Oct 01;3(10):e2020807 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.20807] [Medline: 33052404]
41. Halkett GK, Lobb EA, Miller L, Shaw T, Moorin R, Long A, et al. Feasibility testing and refinement of a supportive educational intervention for carers of patients with high-grade glioma - a pilot study. J Cancer Educ 2018 Oct 11;33(5):967-975. [doi: 10.1007/s13187-017-1175-x] [Medline: 28190236]
42. Shilling V, Starkings R, Jenkins V, Cella D, Fallowfield L. Development and validation of the caregiver roles and responsibilities scale in cancer caregivers. Qual Life Res 2019 Jun 18;28(6):1655-1668 [FREE Full text] [doi: 10.1007/s11136-019-02154-4] [Medline: 30887387]
43. Sveinbjarnardottir EK, Svavarsdottir EK, Hrafnkelsson B. Psychometric development of the Iceland-Family Perceived Support Questionnaire (ICE-FPSQ). J Fam Nurs 2012 Aug 20;18(3):328-352. [doi: 10.1177/1074840712449203] [Medline: 22821443]
44. Snaith RP. The Hospital Anxiety And Depression Scale. Health Qual Life Outcomes 2003 Aug 01;1:29 [FREE Full text] [doi: 10.1186/1477-7525-1-29] [Medline: 12914662]
45. Thavarajah N, Bedard G, Zhang L, Cella D, Beaumont JL, Tsao M, et al. Psychometric validation of the functional assessment of cancer therapy--brain (FACT-Br) for assessing quality of life in patients with brain metastases. Support Care Cancer 2014 Apr 28;22(4):1017-1028. [doi: 10.1007/s00520-013-2060-8] [Medline: 24287508]
46. Piil K, Whisenant M, Mendoza T, Armstrong T, Cleeland C, Nordentoft S, et al. Psychometric validity and reliability of the Danish version of the MD Anderson Symptom Inventory Brain Tumor Module. Neurooncol Pract 2021 Apr;8(2):137-147 [FREE Full text] [doi: 10.1093/nop/npaa068] [Medline: 33898047]
47. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
48. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: SAGE Publications; 2021.
49. NVivo 11. QSR International. URL: https://download.qsrinternational.com/Document/NVivo11/11.3.0/en-US/NVivo11-Getting-Started-Guide-Pro-edition.pdf [accessed 2026-02-18]
50. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015 Nov 09;13(6):554-561 [FREE Full text] [doi: 10.1370/afm.1865] [Medline: 26553895]
51. Creswell JW, Inoue M. A process for conducting mixed methods data analysis. J Gen Fam Med 2025 Jan 03;26(1):4-11. [doi: 10.1002/jgf2.736] [Medline: 39776872]
52. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
53. Clarke JJ, Halkett GK, McDougall E, Dhillon HM, Lobb E, Phillips JL, et al. What do carers of people with high-grade glioma perceive could improve their preparedness to care, and what additional support do they require? Neurooncol Pract 2024 Jun;11(3):296-306 [FREE Full text] [doi: 10.1093/nop/npae015] [Medline: 38737602]
54. Milbury K, Li J, Weathers SP, Mallaiah S, Armstrong T, Li Y, et al. Pilot randomized, controlled trial of a dyadic yoga program for glioma patients undergoing radiotherapy and their family caregivers. Neurooncol Pract 2019 Jul;6(4):311-320 [FREE Full text] [doi: 10.1093/nop/npy052] [Medline: 31386042]
55. Ownsworth T, Chambers S, Aitken JF, Foote M, Pinkham MB, Gordon LG, et al. Evaluation of a telehealth psychological support intervention for people with primary brain tumour and their family members: study protocol for a randomised controlled trial. Eur J Cancer Care (Engl) 2019 Jul 10;28(4):e13132. [doi: 10.1111/ecc.13132] [Medline: 31293057]
56. Milbury K, Whisenant M, Weathers SP, Malliaha S, Snyder S, Jackson N, et al. Dyadic versus individual delivery of a yoga program for family caregivers of glioma patients undergoing radiotherapy: results of a 3-arm randomized controlled trial. Cancer Med 2023 Mar;12(6):7567-7579 [FREE Full text] [doi: 10.1002/cam4.5514] [Medline: 36468605]
57. Yang P, Zhang L, Tian X. Application of stepped care model in cancer patients: a scoping review. Support Care Cancer 2024 Dec 17;33(1):34. [doi: 10.1007/s00520-024-09086-x] [Medline: 39688623]
58. Bower P, Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. Br J Psychiatry 2005 Jan 02;186(1):11-17. [doi: 10.1192/bjp.186.1.11] [Medline: 15630118]

XSL•FO
RenderX

Abbreviations

GBM: glioblastoma
MRC: Medical Research Council
PPI: patient and public involvement
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SUGRI: Supportive Group Intervention for caregivers to patients diagnosed with a glioblastoma

Edited by J Sarvestan; submitted 06.Aug.2025; peer-reviewed by LPV Bovim; comments to author 09.Oct.2025; revised version received 19.Dec.2025; accepted 22.Dec.2025; published 11.Mar.2026.

Please cite as:
Nordentoft S, Guldager R, Pappot H, Mathiesen T, Tolver A, Piil K
A Supportive Group Intervention for Caregivers to Patients Diagnosed With Glioblastoma: Protocol for the SUGRI Study
JMIR Res Protoc 2026;15:e81948
URL: https://www.researchprotocols.org/2026/1/e81948
doi:10.2196/81948
PMID:

©Sara Nordentoft, Rikke Guldager, Helle Pappot, Tiit Mathiesen, Anders Tolver, Karin Piil. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 11.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of the Comparative Efficacy of an Ayurvedic Formulation (Nimba-Amalakyadi Powder) vs Metformin in the Management of Type 2 Diabetes Mellitus: Protocol for a Pilot Study

Punam Sawarkar[1*], BAMS, MD, PhD; Gaurav Rajendra Sawarkar[2*], BAMS, MD, PhD

[1]Department of Panchakarma, Mahatma Gandhi Ayurved College, Hospital and Research Centre, Datta Meghe Institute of Higher Education and Research, Salod(H), Wardha, Maharashtra, India

[2]Department of Rachana Sharir, Mahatma Gandhi Ayurved College, Hospital and Research Centre, Datta Meghe Institute of Higher Education and Research, Salod(H), Wardha, Maharashtra, India

[*]all authors contributed equally

Corresponding Author:
Punam Sawarkar, BAMS, MD, PhD
Department of Panchakarma, Mahatma Gandhi Ayurved College, Hospital and Research Centre, Datta Meghe Institute of Higher Education and Research, Salod(H), Wardha, Maharashtra, India

Related Article:

This is a corrected version. See correction statement: https://www.researchprotocols.org/2026/1/e96344

Abstract

Background: In Ayurveda, *Prameha* (diabetes mellitus) is considered *Ashtamahagada Vyadhi* (one among eight diseases difficult to treat and cure). *Tridosha* (three bio-humor in Ayurveda) with certain metabolic factors are involved in its pathogenesis. Frequent urination is a common symptom of this disease. Diabetes mellitus can be confused with type 2 diabetes mellitus in conventional science due to the resemblance in cardinal features.

Objective: The primary objectives of this pilot study are to assess feasibility through recruitment, retention, adherence, data completeness, and participant acceptability. It also aims to estimate variability in outcome measures like fasting blood sugar and glycated hemoglobin (HbA_{1c}). The secondary, exploratory objective is to observe trends in these clinical indicators to inform effect size assumptions for future trials.

Methods: In this pilot study 36 individuals with type 2 diabetes mellitus will be recruited and split into two equal groups at random. In Group A (Control), metformin tablets (500 mg) will be prescribed daily once before lunch with plain water for 45 consecutive days. In Group B (Interventional group), the *Nimba-Amalakyadi* formulation will be prescribed 5 grams twice daily with an empty stomach (ie, 7 AM-5 PM) with lukewarm water for 45 consecutive days. Follow-up will be taken on days 45 and 90 to study the sustained effect of the drug.

Results: The preliminary analysis of participant responses began in January 2025. Final results from this phase are expected to be available by June 2025. Recruitment rate (number recruited/month), eligibility rate (number eligible/number screened), retention rate at 45 days, compliance with intervention (percentage of prescribed doses taken), adherence to follow-up investigations, acceptability of intervention (based on participant-reported feedback), and data completeness rates are variables. Moreover, changes in the objective parameters, ie, reductions in blood sugar level (fasting and postprandial) and HbA_{1c} values of patients will be also observed and recorded.

Conclusions: A conclusion will be drawn according to the clinical findings obtained at baseline and follow-up visits with biochemical reports of the patients. The trial will prove the feasibility and comparative efficacy of *Nimba Amalakyadi* formulation versus metformin tablets for the management of type 2 diabetes mellitus.

Trial Registration: The Clinical Trials Registry- India (CTRI) CTRI 2021/08/036034; https://tinyurl.com/2xwc84p2

International Registered Report Identifier (IRRID): PRR1-10.2196/63574

(JMIR Res Protoc 2026;15:e63574) doi:10.2196/63574

XSL•FO
RenderX

KEYWORDS

type 2 diabetes mellitus; metformin; evaluation; diabetes mellitus; type 2 diabetes; clinical trial protocol; diabetes; treatment; cure; cardinal; assessment; pathogenesis; powder form; blood sugar level; sugar level; biochemical report; clinical findings

Introduction

Diabetes mellitus (DM) is a metabolic syndrome with disturbed carbohydrates, fat, and protein metabolism. It is characterized by prolonged hyperglycemia due to disturbances in insulin secretion [1].

It remains a silent but life-threatening lifestyle disorder due to its irreversible complications. It is associated with certain tissue or vascular damage with the disease's progression, resulting in severe complications, including diabetic retinopathy, neuropathy, nephropathy, stroke, and diabetic foot. The broad spectrum of complications occurs as a result of uncontrolled diabetes [2]. In the modern era, a sedentary lifestyle is attributed to a lot of stress and nutrition, which makes diabetes one of the most prevalent diseases globally. Although there have been many advancements in modern medicine, there is still no clear and proper solution to the utmost satisfaction in redressing diabetes. Thus, it becomes the biggest silent killer in today's world [3].

DM is a complex metabolic clinical condition generated from the disturbed secretion of insulin or insulin receptor insensitivity that induces further resistance to an elevated blood sugar level. In 1980, there were 108 million patients with diabetes; by 2014, there were 422 million. Adults worldwide now have an 8.5% prevalence rate of diabetes [4]. This rate is increasing rapidly, especially in middle- and low-income countries. Uncontrolled DM is the major contributor to death occurring in the age group below 70 years, which is the leading cause of death in India, as per the World Health Organization [5].

There are two main types of DM, that is, types 1 and 2. Type 1 DM is an auto-immune disorder commonly occurring in pediatric age groups and young people, resulting from destroying ß cells in the pancreas due to certain auto-immune responses. The second type, that is, type 2 DM (T2DM), is the most common in adult patients, often associated with modifiable lifestyle factors such as physical inactivity, poor dietary habits, and stress, resulting in impaired insulin resistance (insulin resistance) or insulin insufficiency (ß-cell dysfunction). The increased demand for insulin in the body to control elevated sugar levels induces an excess burden over these pancreatic ß cells, ultimately resulting in the failure of ß-cells that get converted into impaired fasting blood glucose and postprandial hyperglycemia [6].

In the Indian context, increasing urbanization, industrialization, and changing lifestyles have contributed to the increasing prevalence of DM. The common features of DM include frequent micturition, excessive thirst, weight loss, and impairment of vision due to significant hyperglycemia. Such patients also become more susceptible to getting certain infections and growth impairment in the prolonged state of hyperglycemia. It is a complex metabolic disorder with multiple long-term complications despite regular treatment, such as retinopathy, microangiopathy (nephropathy), and neuropathy. These challenging entities are not preventable, even with satisfactory control of blood sugar levels. Moreover, these are quite difficult to treat in clinical practice [7].

Despite the rapid advancement of modern medical science, especially in understanding the disease, there is still no permanent remedy or known complete cure for the management of DM [8]. Moreover, the prognosis of this disease is very poor due to certain irreversible major changes or dysfunction. It is mostly treated with a combination of lifestyle modification and oral hypoglycemic drugs, such as metformin or insulin injections. However, these drugs have numerous side effects due to long-term consumption, such as weight gain, diarrhea, bloating in the abdomen, gastrointestinal problems, liver damage, hypersensitivity reactions, and flatulence [9]. For example, metformin has certain side effects such as a metallic taste in the mouth, fatigue, diarrhea, and stomach ache [10,11]. Therefore, it becomes imperative to search for some safe, alternative solutions for such kinds of chronic and complex diseases to alter the basic pathology of the disease as well as to avoid any complications.

A thorough analysis of the Ayurvedic literature indicates that lifestyle changes, herbal medication, and Panchakarma (putative therapies in Ayurveda) are safe and effective for managing DM [12]. This study is focused exclusively on individuals with T2DM, a condition typically managed through lifestyle modification and pharmacological agents such as metformin. Type 1 DM, which involves absolute insulin deficiency and requires insulin therapy from diagnosis, is not included in this study population. To address the drawbacks and adverse effects of traditional treatments of T2DM, it is imperative to investigate safe and alternative Ayurvedic therapy approaches [10,11].

Much research work has been carried out on various Ayurvedic interventions, including *Shodhana* (putative) and *Shamana Chikitsa* (palliative treatment in Ayurveda) with different herbal and herbal-mineral formulations. Among them, Panchakarma (penta-bio-purificatory) therapies offer promising results for controlling pathogenesis and curing symptoms of DM. However, many patients refuse to undergo such treatments due to the time-consuming nature, tedious nature of pre-procedure like internal oleation therapy, the requirement of a certain number of sittings for procedures, and certain dos and don'ts regarding diet and lifestyle to follow during these procedures. Moreover, there is apprehension in society regarding mineral formulations that can cause adverse effects, such as liver or kidney failure. Therefore, there are many limitations to the benefit of Panchakarma and palliative treatments in the form of mineral or herbo-mineral compounds for managing diabetes. The hypoglycemic effect of the individual drug in the *Nimba-Amalakyadi* formulation has been observed [13].

However, the synergistic effects of such anti-diabetic medicines, which are readily available, have not been studied yet. Moreover, no clinical data exist about their comparative effects with standard oral hypoglycemic agents. Therefore, this study

XSL•FO
RenderX

will make novel efforts to study the synergistic hypoglycemic effects of Ayurvedic herbal formulation, ie, *Nimba–Amalakyadi Yoga*, which consists of seven herbs, ie, *Nimba (Azadaricta Nimba Linn.)*, *Amalaki (Embelica officinalis* Linn.), *Haridra (Curcumin longa* Linn.), *Kutaj (Holarerna antidystentrica* Linn.), *Vijaysara (Pterocarpus marsupium* Linn.), *Pippali (Piper longum* Linn,) and *Marich (Piper nigram* Linn) for DM. Moreover, its efficacy will also be compared with that of metformin tablets. The primary aim of this study is to study the comparative efficacy of *Nimba Amalakyadi* formulation with that of metformin in managing T2DM (*Prameha*), as well as to estimate the variability and completeness of outcome data collection (eg, fasting blood sugar and glycated hemoglobin [HbA_{1c}]); evaluate the acceptability of the treatment and procedures from the participants' perspective; and assess the recruitment rate, retention rate, and adherence to the intervention. In order to assist in creating effect size assumptions for upcoming trials, the secondary, exploratory objective is to track trends in clinical indicators such fasting blood sugar and HbA_{1c}.

Methods

Ethical Considerations

The trial protocol was reviewed and approved by the institutional ethics committee, Mahatma Gandhi Ayurveda College Hospital and Research Centre, Salod (Hirapur), Wardha, Maharashtra, India, IEC Ref No. MGACH/IEC/April-2021/202 dated 09.04.2021. The enrollment of patients was initiated only after receiving the CTRI registration number, CTRI/2021/08/036034 received dated 31/08/2021. Every procedure involving human participants complies with the 1964 Helsinki Declaration and its subsequent amendments, as well as the institutional and/or national research committee's ethical requirements. We ensured that each patient gave their written informed consent before recruiting them. Participants received comprehensive information about the study's goals, interventions, possible risks, and advantages. They are made aware of their freedom to leave at any moment without repercussions. Throughout the study, each patient's privacy will be protected. The confidentiality of participants is rigorously respected. Unique participant numbers are used to anonymize all personal information, and the data are safely stored on password-protected systems with limited access. No publication arising from this research will reveal any personally identifiable information. No financial payment was made to participants for taking part. Because the interventions utilized in this study comprise well-established medications with recognized safety profiles (standard metformin and Ayurvedic formulations), no risk is involved.

Study Design

This study was an interventional study (randomized open-labeled standard controlled equivalence clinical trial). This whole protocol is designed as per the guidelines of the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) Checklist(Checklist 1).

Sample Size

The sample size is limited due to practical constraints such as funding unavailability, study duration, and participant considerations. For this pilot Phase II study, 18 participants per group (total n=36) are proposed, accounting for a 10% dropout. While a formal power calculation is not required, this size offers approximately 90% power to detect an increase in the primary outcome from 32% to 70% at a 5% significance level. This range is appropriate for estimating parameters to inform a future definitive randomized controlled trial while maintaining feasibility and resource efficiency [13-18].

Recruitment

Eligible participants with T2DM will be identified from inpatient and outpatient departments of the institute. Those meeting the inclusion criteria will be approached for written informed consent. Individual patients will be added until the desired sample size is attained.

Randomization

After obtaining consent and performing baseline assessment, participants will be randomly allocated into intervention or control groups via simple randomization using a computer-generated table method.

Setting

Location of the Study

The study will be carried out in the outpatient and inpatient departments of *Panchkarma* and *Kayachikitsa* at the Mahatma Gandhi Ayurveda College Hospital and Research Centre, Salod (Hirapur) Wardha, Maharashtra, India.

Study Setting

Following clearance from the Clinical Trials Registry - India (CTRI), patients will be enrolled, and their treatment will involve a 1.5-month (45 d) exposure period. There will be a follow-up on day 90 of the treatment.

Criteria for Discontinuing or Modifying Treatment

Participants may withdraw from the study at any time. Those who withdraw before starting the intervention will be replaced to maintain the intended sample size. However, participants who withdraw after receiving part of the intervention (eg, after 30 days) will be considered dropouts and included in the intention-to-treat analysis. They will not be replaced, as their data remain valuable for assessing safety and efficacy. Participants who develop acute or life-threatening conditions or show non-compliance will also be excluded from the study.

Participants

The patients diagnosed with ICD-10-CM Diagnosis Code E11.9, that is, newly diagnosed patients with T2DM having a duration of disease of <6 months and having fasting blood sugar levels between 126 and 200 mg/dl and post-meal blood sugar level between 140 and 300 mg/dl will be recruited in the study via simple randomization using the computerized generated table method.

XSL•FO
RenderX

Eligibility Criteria

Inclusion Criteria

The inclusion criteria for the study consists of patients with T2DM of either sex, aged between 20 and 60 years, who were recently diagnosed (<6 months) with uncomplicated T2DM and having a fasting blood sugar level between the range of 126 and 200 mg/dl, post-meal blood sugar level between the range of 140 and ‐300 mg/dl, and HbA_{1c} level of 5.7% ‐ 8%. Along with the aforementioned criteria, patients who were not taking any other anti-hypoglycemic drugs (ie, no allopathic, Ayurvedic, homeopathic, or Unani drugs) were also included. Patients with diabetes who have controlled hypertension (blood pressure of no more than 140 mmHg or an average diastolic blood pressure of no more than 90 mmHg) and patients willing to provide written informed consent were also included.

Exclusion Criteria

Patients with insulin-dependent diabetes mellitus, patients with T2DM who are on insulin therapy, patients with juvenile diabetes or gestational diabetes (ICD-10 criteria 024), and those with impaired glucose tolerance (ICD-10 criteria R73.0) will be excluded from the study. Other exclusion criteria include patients with diabetes who have complications such as retinopathy, nephropathy, neuropathy, or a history of coma; patients with uncontrolled hypertension; and pregnant and lactating women.

Interventions

Overview

For Group A (Control Group), metformin tablets (500 mg) will be prescribed daily once before lunch with plain water for 45 consecutive days. In Group B (Interventional group), 5 grams *Nimba-Amalakyadi* formulation in powder form will be prescribed twice daily on an empty stomach (ie, 7 AM-5 PM) with lukewarm water for 45 consecutive days. Follow-up will be performed on days 45 and 90 to study the efficacy and sustained effect of the drug.

Assessment Criteria

The assessment criteria included the fasting and postprandial blood sugar level, urine sugar level, and HbA_{1c} level.

Variables to Be Measured

As it is a pilot study, ecruitment rate (number recruited/month), ligibility rate (number eligible/number screened), retention rate at 45 days, compliance with intervention (percentage of prescribed doses taken), adherence to follow-up investigations, acceptability of intervention (based on participant-reported feedback), and data completeness rates are variables. Moreover, the reduction in Blood sugar levels (fasting and postprandial), HbA_{1c}, and urine sugar level values of patients with T2DM will be also assessed to calculate the effect size.

Participant Timeline

The duration of intervention for both groups is 45 days. The first and second follow-ups will be on the 45th day and 90th day of study, respectively.

Assignment of Interventions (For Controlled Trials) and Blinding

The computer-generated table will be used for convenient randomization in the recruitment of patients. The main investigator and co-investigator will assign and enroll the patients. The researcher generates random allocation cards using computer-generated random numbers. The researchers will store the initial randomly selected sequences in a secure place, and the patients will use a separate copy. Serial numbers will be written on the outside of the envelopes for the purpose of blinding. The envelope will contain the following necessary information: the date, time, patient ID, postintervention results, and others.

Data Collection and Management

Reminders for appointments and question-and-answer sessions will be created to keep patients engaged in the project. The patient file will include information about the whole follow-up, and if patients are asked to stop or withdraw, the file will be closed and kept on file for documentation purposes.

After the completion of the study, observations will be noted and drawn from case record forms with thorough examinations, history, and assessment sheets; findings of values of biochemical parameters; and a proforma for follow-up evaluation.

Monitoring

The lead investigator and a co-investigator will handle data monitoring and coding. A hard-copy case record form and soft-copy data entry will be completed. The values will be checked twice before being sent in electronic format. A unique patient code protects every patient file and planning for collecting, assessing, and managing auditing adverse events.

The trial conduct will be audited quarterly by an independent monitoring team using a predefined checklist to ensure adherence to the study protocol and regulatory compliance.

Statistical Analysis

The initial analysis of participant responses commenced in January 2025. The outcomes from this phase are anticipated by June 2025. Objective measures, including reductions in fasting and postprandial blood glucose levels and HbA_{1c} values, will be monitored and documented. After CTRI Registration, a delay occurred in the recruitment of patients due to a delay in the availability of raw drugs and the preparation of the Ayurveda medicine as a result of some administrative issues.

Statistical comparisons between the Ayurvedic formulation and metformin groups will be performed using the Mann-Whitney *U* test for continuous variables and χ^2/Fisher exact test for categorical variables.

Only descriptive statistics, confidence intervals, and exploratory calculation of variability—avoidance of *P* values where appropriate—will be reported; no hypothesis testing will be done. All analyses will be conducted using a recent version of the SPSS software (version 25.0, IBM Corp), and results will be interpreted as exploratory due to the small sample size.

XSL•FO
RenderX

Results

Participant recruitment and baseline data collection commenced in January 2025. As of April 2025, 28 out of the planned 36 participants have been enrolled. The intervention and follow-up for all participants are expected to be completed by July 2025. Data analysis will be conducted in August 2025, with manuscript preparation and submission of final results planned for September–October 2025.

Discussion

Comparison With Previous Works

T2DM is a chronic metabolic disorder with a rising global prevalence. Ayurvedic formulations have shown promise in managing hyperglycemia through multiple mechanisms, such as antioxidant and insulin-sensitizing effects. As common people always prefer herbal drugs due to their safe and non-invasive nature, the *Ayurvedic* herbal formulation, that is, *Nimba-Amalakyadi* formulation, selected for this trial consists of seven herbs including *Nimba (Azadaricta nimba Linn.), Amalaki (Embelica officinalis* Linn.), *Haridra (Curcumin longa* Linn.), *Kutaj (Holarerna antidystentrica* Linn.), *Vijaysara (Pterocarpus marsupium* Linn.), *Pippali (Piper longum* Linn.), and *Marich (Piper nigram* Linn). The *Prameghna* (hypoglycemic action) of each herb can be explained based on the classical textual references from Ayurveda and modern science as described below.

Nimba is an important ingredient of *Nimba-Amrutadi Erand*, an effective *Ayurvedic* medicine for DM. The hypoglycemic effect, that is, a significant reduction in blood sugar levels in adrenaline and glucose-induced hyperglycemia due to aqueous extracts of *Neem* leaves, has been discussed [19]. He also quotes that aqueous extracts of Neem leaves are also effective in controlling streptozotocin-induced diabetes. The hypoglycemic effect of the aqueous extracts of Neem leaves may be due to the rich content of the flavonoid quercetin that potentiates insulin secretion from the pancreas. The *Nimba-Amalakyadi* formulation is quite effective due to its anti-hyperglycemic potential to lower peripheral glucose utilization and block the effects of epinephrine on glucose metabolism in diabetic and normal rats [20].

Amalaki is the main ingredient of one of the Ayurvedic formulations, *Dhatri-Nisha Yoga*, whose efficacy has been proven by several clinical studies [14-18]. It is quite a common drug in clinical practice for the management of DM. The large content of vitamin C, tannoids, and polyphenols present in *Phyllanthus Emblica L* is useful to control diabetes by controlling oxidative stress and avoiding hyperglycemia along with hyperlipidemia induced by it. It reduces the blood sugar level in diabetes by enhancing insulin secretion by stimulating the pancreas [21].

Haridra is considered the best hypoglycemic herb in Ayurveda. The rhizome of *Haridra* possesses hypoglycemic action due to its *Tikta, Katu Rasa* (bitter and pungent taste as per Ayurveda), *Ushna Virya* (hot potency), *Katu Vipaka* (pungent effect after digestion as per Ayurveda), *Laghu* (easy for digestion), and *Ruksha* (dry nature) properties. *Haridra* breaks the pathogenesis of the disease by balancing the disturbed *Vata*, *Pitta*, and *of Kapha*, all of which are vitiated in T2DM as per Ayurveda. The anti-diabetic effect of *Haridra* may be due to the curcumin present in its rhizomes that reduce insulin resistance and oxidative stress. It also protects insulin-secreting cells in the pancreas. The three main constituents, that is, curcumin, dimethoxy-curcumin, bisdemethoxycurcumin, and ar-turmerone, reduce hyperglycemia via peroxisome proliferator-activated receptor-gamma activation [22].

The anti-diabetic potential of *Kutaja* (especially its seeds and bark) is explained by Saroya [23]. *Kutaja* pacifies *Kapha Dosha*, the main culprit pathogenesis of *Prameha* due to its *Tikta*, *Kashaya Rasa* (bitter and astringent taste as per Ayurveda), *Katu Vipaka* (pungent effect after digestion as per Ayurveda), *Laghu*, and *Ruksha* properties. The antioxidant and antidiabetic properties of *Kutaja* are based on the rich content of flavonoids and phenolic compounds present in its bark. Such anti-diabetic effects may be induced by enhancing the uptake of liver glycogen and glucose [24].

Flavonoids in *Kutaja* possess an antioxidant property that reduces hyperglycemia induced by oxidative stress. *Vijaysara* is one of the hypoglycemic herbs used to prepare *Bijak* Glass (*Nisha* Herbal company) or *Nisha-Kathakadi Choorna* (powder)and *Qwath* (decoction). The hypoglycemic activity of the wood of *Vijaysara* in dogs was shown by Singh et al [25]. This effect was due to the pterostilbene and tannates present in it [24].

Pippali and *Marich* possess *Katu-Tikta Rasa*, and *Laghu, Snigdha, and Tikshna* properties that have *Kaphaghna* (a phlegmn-alleviating effect) and *Dipana* (digestion enhancers) actions. Both herbs also induce immunomodulation due to their *Rasayana* (rejuvenative) properties useful in *Prameha*. Their *Kapha–Medoghna* (alleviation of phlegmn and fat) property and *Prameghna* action (anti-diabetic activity) were proved by Reddy [26] and Khaliq et al [27], respectively.

In a nutshell, most of the ingredients of *Nimba –Amalakyadi Yoga* (formulation) are primarily *Katu and Tikta Rasa Dravyas*, *Kledahara* (absorbing hydrous content in the body) and *Agnidipaka* (enhancement in digestive fire) in nature. The *Kledaghna* properties of *Tikta Rasa*, useful in the management of *Prameha*, are described by *Acharya Charaka* based on its *Panchbhautik* (made of *Panchamahahabhuta*) composition, that is, the predominance of *Vata* and *Akasha Mahabhuta* that induces *Stroto Vishodhana* (channel-clearing effect) effect due to its *Sukshma Guna* (easy penetration capacity at the micro-cellular level). All the above-mentioned herbs are natural immunomodulators due to their *Rasayana* (rejuvenating) property that reduces oxidative stress and hyperglycemia induced by it. Moreover, it also prevents further *Dhatu Kshaya* (tissue emaciation) present in *Prameha* due to the vitiation of Vata and causes improvement in the clinical features of *Prameha* due to this *Rasayana* property.

The anti-diabetic effect of all these herbs may be induced by increasing insulin secretion or enhancing insulin sensitivity at the receptor level, promoting the regeneration of ß cells, reducing the absorption of glucose from the gastrointestinal

XSL•FO
RenderX

tract, lowering blood glucose level through the improved peripheral utilization of glucose, decreasing their resistance, decreasing oxidative stress by the inhibition of superoxide radicals or hydroxyl radical scavenging activity or converting reactive oxygen species to non-reactive products, lowering cholesterol levels by the inhibition of lipid peroxidation.

Based on the above clinical shreds of evidence, the study was planned with this formulation to assess the synergistic hypoglycemic effects of herbs. In patients with diabetes, this can become a scientific approach for the clinical management of DM, which becomes a ray of hope for such patients. At the same time, the efficacy of *Nimba Amalakyadi Choorna* will be more than that of metformin for the management of T2DM.

Preliminary observations from this pilot study suggest favorable trends in glycemic control with the Ayurvedic formulation, aligning with previous studies demonstrating the hypoglycemic effects of herbs such as *Pterocarpus marsupium* and *Curcuma longa* [28, 29]. Unlike studies that lacked comparator arms [30, 31, 32], our protocol directly compares the Ayurvedic formulation with metformin, providing a more rigorous assessment.

The feasibility outcomes will be revealed in terms of the steady recruitment through outpatient services, participant compliance in percentage, and the tolerability of both interventions along with the observation of any major adverse effects. Some logistical challenges, such as coordinating follow-up visits, will be mitigated through flexible scheduling.

This pilot study will inform a future full-scale RCT by providing estimates of the effect size, participant retention rates, and potential refinements to intervention protocols. However, the limitations include a small sample size, short trial duration (45 days), and the lack of statistical power for definitive efficacy conclusions. The limited sample size due to limited financial assistance is the study's major limitation. Moreover, patients with other types of diabetes, for example, those with type 1 DM, or juvenile or gestational DM, were excluded from this study. Further clinical trials in these populations with similar interventions can be planned in the future.

Upon completion, data will guide protocol adjustments for a larger, adequately powered RCT aimed at validating the clinical efficacy and safety of the Ayurvedic formulation in managing T2DM. If the intervention offers significant results, then safe, cost-effective, time-saving, and alternative treatment in Ayurveda can be identified for patients with T2DM who are not willing to undergo *Panchakarma* Therapy or are contraindicated for conventional existing treatment modalities.

Conclusions

If this trial yields effective and promising results, it could provide clinical evidence supporting the efficacy of a safe, herbal alternative to metformin for managing T2DM. This Ayurvedic formulation may offer an option for individuals who seek Ayurveda treatment but cannot undergo procedures like *Vamana* (therapeutic emesis) or *Virechana* (therapeutic purgation), or those who struggle with strict dietary and lifestyle modifications required for *Panchakarma* therapies. It may also appeal to those wary of mineral-based hypoglycemic drugs. The trial would highlight the synergistic hypoglycemic effects of *Nimba-Amalakyadi Yoga*, a combination of herbs with proven hypoglycemic properties. Future studies could expand on this research with larger sample sizes and compare its effects to Shodhana therapies like Vamana and Virechana in patients with T2DM. Further clinical trials could explore its potential in treating juvenile diabetes, pregnancy-induced diabetes, or drug-induced DM.

Acknowledgments

We would like offer our sincere gratitude for the research cell for valuable guidance and encouragement for the planning of this project.

Artificial intelligence was used for making grammatical corrections and rephrasing statements.

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Funding

The authors declare that no grants were involved in supporting this work.

Data Availability

Protocol-related information is accessible on the CTRI PORTAL and will be updated in accordance with the study's progress. The collected data are only accessible to the principle investigator and authorized research personnel. The information will only be utilized for this study. In compliance with institutional regulations, deidentified data may be made available upon request after publishing. No data are given in this article.

Authors' Contributions

PS: conceptualization, project administration, writing–original draft preparation
GRS: conceptualization, writing–review & editing

XSL•FO
RenderX

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[PDF File, 216 KB - resprot_v15i1e63574_app1.pdf]

References

1. Dilworth L, Facey A, Omoruyi F. Diabetes mellitus and its metabolic complications: the role of adipose tissues. Int J Mol Sci 2021 Jul 16;22(14):7644. [doi: 10.3390/ijms22147644] [Medline: 34299261]
2. Maida CD, Daidone M, Pacinella G, Norrito RL, Pinto A, Tuttolomondo A. Diabetes and ischemic stroke: an old and new relationship an overview of the close interaction between these diseases. Int J Mol Sci 2022 Feb 21;23(4):2397. [doi: 10.3390/ijms23042397] [Medline: 35216512]
3. Ashtankar PV, Patil M, Singh K. Comparative evaluation of the efficacy of Panchatikta Panchaprasutik Niruha Vasti and lifestyle modification in Prediabetes (Prameha Poorvaroopa). Wutan Huatan Jisuan Jishu 2020 Dec;16(XII):661-667 [FREE Full text]
4. Alam S, Hasan M, Neaz S, Hussain N, Hossain M, Rahman T. Diabetes mellitus: insights from epidemiology, biochemistry, risk factors, diagnosis, complications and comprehensive management. Diabetology 2021;2(2):36-50. [doi: 10.3390/diabetology2020004]
5. Jha RP, Shri N, Patel P, Dhamnetiya D, Bhattacharyya K, Singh M. Correction to: trends in the diabetes incidence and mortality in India from 1990 to 2019: a joinpoint and age-period-cohort analysis. J Diabetes Metab Disord 2021 Dec;20(2):1741-1741. [doi: 10.1007/s40200-021-00865-5] [Medline: 34898664]
6. Khawandanah J. Double or hybrid diabetes: a systematic review on disease prevalence, characteristics and risk factors. Nutr Diabetes 2019 Nov 4;9(1):33. [doi: 10.1038/s41387-019-0101-1] [Medline: 31685799]
7. Sorrentino FS, Matteini S, Bonifazzi C, Sebastiani A, Parmeggiani F. Diabetic retinopathy and endothelin system: microangiopathy versus endothelial dysfunction. Eye (Lond) 2018 Jul;32(7):1157-1163. [doi: 10.1038/s41433-018-0032-4] [Medline: 29520046]
8. Aloke C, Egwu CO, Aja PM, et al. Current advances in the management of diabetes mellitus. Biomedicines 2022 Sep 29;10(10):2436. [doi: 10.3390/biomedicines10102436] [Medline: 36289697]
9. Jasim ME, Abdul-Razzaq NE. A comparative study between metformin and insulin on biochemical variables in diabetic patients. Endocrine and Metabolic Science 2024 Sep;16:100206. [doi: 10.1016/j.endmts.2024.100206]
10. Zhou Y, Xue F. Unlocking new potential: the evolving landscape of metformin repurposing trials. J Transl Med 2025;23(1):24. [doi: 10.1186/s12967-024-06061-2]
11. Klein KR. Metformin for glucocorticoid-induced toxicity: why not? Diabetes Care 2025 May 1;48(5):688-690. [doi: 10.2337/dci25-0010] [Medline: 40273353]
12. Iheagwam FN, Iheagwam OT. Diabetes mellitus: the pathophysiology as a canvas for management elucidation and strategies. Medicine in Novel Technology and Devices 2025 Mar;25:100351. [doi: 10.1016/j.medntd.2025.100351]
13. Joo S, Chun H, Lee J, Seo S, Lee J, Leem J. Hypoglycemic effect of an herbal decoction (modified Gangsimtang) in a patient with severe type 2 diabetes mellitus refusing oral anti-diabetic medication: a case report. Med Bogota Colomb 2023 Oct 30;59(11):1919. [doi: 10.3390/medicina59111919]
14. Uchio R, Okuda-Hanafusa C, Saji R, et al. A hot water extract of Curcuma longa L. improves fasting serum glucose levels in participants with low-grade inflammation: reanalysis of data from two randomized, double-blind, placebo-controlled trials. Nutrients 2022 Sep 13;14(18):3763. [doi: 10.3390/nu14183763] [Medline: 36145139]
15. Pathomwichaiwat T, Jinatongthai P, Prommasut N, et al. Effects of turmeric (Curcuma longa) supplementation on glucose metabolism in diabetes mellitus and metabolic syndrome: an umbrella review and updated meta-analysis. PLoS ONE 2023;18(7):e0288997. [doi: 10.1371/journal.pone.0288997] [Medline: 37471428]
16. Majeed M, Nagabhushanam K, Devarajan TV, Saklecha S, Reddy SVK, Mundkur L. A minor metabolite from Curcuma longa effective against metabolic syndrome: results from a randomized, double-blind, placebo-controlled clinical study. Food Funct 2023 May 22;14(10):4722-4733. [doi: 10.1039/d2fo03627d] [Medline: 37114318]
17. Kumar S, Yadav A. Comparative study of hypoglycemic effect of Holarrhena antidysenterica seeds and glibenclamide in experimentally induced diabetes mellitus in albino rats. Biomed Pharmacol J 2015 Jun 30;8(1):477-483. [doi: 10.13005/bpj/637]
18. Patidar A, Tonpay SD, Agrawal N. Hypoglycemic activity of Pterocarpus marsupium in patients with Type 2 diabetes mellitus. Int J Basic Clin Pharmacol 2015;4(1):1189-1193. [doi: 10.18203/2319-2003.ijbcp20151356]
19. Ogbuewu IP, Odoemenam VU, Obikaonu HO, et al. The growing importance of neem (Azadirachta indica A. Juss) in agriculture, industry, medicine and environment: a review. Research J of Medicinal Plant 2011 Mar 1;5(3):230-245. [doi: 10.3923/rjmp.2011.230.245]

XSL•FO
RenderX

20. Cook JA, Julious SA, Sones W, et al. Practical help for specifying the target difference in sample size calculations for RCTs: the DELTA2 five-stage study, including a workshop. Health Technol Assess 2019 Oct;23(60):1-88. [doi: 10.3310/hta23600] [Medline: 31661431]
21. Ying L, Chaudhry MT, Xiao F, et al. The effects and mechanism of quercetin dietary supplementation in streptozotocin-induced hyperglycemic arbor acre broilers. Oxid Med Cell Longev 2020;2020:9585047. [doi: 10.1155/2020/9585047] [Medline: 32104545]
22. Munshi R, Karande-Patil S, Kumbhar D, Deshmukh A, Hingorani L. A randomized, controlled, comparative, proof-of-concept study to evaluate the efficacy and safety of Nisha-Amalaki capsules in prediabetic patients for preventing progression to diabetes. J Ayurveda Integr Med 2023;14(6):100806. [doi: 10.1016/j.jaim.2023.100806] [Medline: 37857033]
23. Saroya AS. Contemporary Phytomedicines: CRC Press; 2017. [doi: 10.1201/9781315367071]
24. Sultana R, Rahman MN, Yusuf M, Karim M. Therapeutic potential of Inderjo Talkh (Holarrhena antidysenterica) in Unani medicine: an in-depth study. J Drug Delivery Ther 2025;15(4):165-171. [doi: 10.22270/jddt.v15i4.7073]
25. Singh K, Alam MI, Phalke E, Lakshman K. Traditional diabetes care with herbal medicine. WJPPS 2023;12(11):648-677 [FREE Full text]
26. Reddy KRC. Remedial merits of Piper longum Linn with astonishing antidiabetic potential. IJGP 2017 Dec 29;11(4). [doi: 10.22377/ijgp.v10i04.1286]
27. Khaliq T, Sarfraz M, Ashraf MA. Recent progress for the Utilization of Curcuma longa, Piper nigrum and Phoenix dactylifera seeds against type 2 diabetes. West Indian Med J 2015 Dec;64(5):527-532. [doi: 10.7727/wimj.2016.176] [Medline: 27399905]
28. Grover JK, Yadav S, Vats V. Medicinal plants of India with anti-diabetic potential. J Ethnopharmacol 2002 Jun;81(1):81-100. [doi: 10.1016/s0378-8741(02)00059-4] [Medline: 12020931]
29. Aggarwal BB, Sundaram C, Malani N, Curcumin IH. The Indian solid gold. Adv Exp Med Biol 2007;595:1-75. [doi: 10.1007/978-0-387-46401-5_1] [Medline: 17569205]
30. Sharma S, Jadhav LL. Bhudhatryadi Yoga in Madhumeha (diabetes mellitus type 2): An open label single arm clinical study. J Ayurveda Integr Med Sci 2022;7(3):01-06. [doi: 10.21760/jaims.7.3.1]
31. Totad M. An open label, single arm prospective clinical study on the effectiveness of Pugaarimeda Kashaya in Madhumeha (Diabetes Mellitus Type 2). J Ayurveda Integr Med Sci 2022;7(8):01-07. [doi: 10.21760/jaims.7.8.1]
32. Benakatti VB. Efficacy of the polyherbal formulation Phalatrikadi Vati in the management of type 2 diabetes mellitus: a single arm open labeled trial. J Nat Remedies 2023. [doi: 10.18311/jnr/2023/29166]

Abbreviations

DM: diabetes mellitus
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
T2DM: type 2 diabetes mellitus

Edited by A Schwartz; submitted 24.Jun.2024; peer-reviewed by AN Ali, D Prasad; revised version received 01.Aug.2025; accepted 10.Sep.2025; published 13.Mar.2026.

Please cite as:
Sawarkar P, Sawarkar GR
Evaluation of the Comparative Efficacy of an Ayurvedic Formulation (Nimba-Amalakyadi Powder) vs Metformin in the Management of Type 2 Diabetes Mellitus: Protocol for a Pilot Study
JMIR Res Protoc 2026;15:e63574
URL: https://www.researchprotocols.org/2026/1/e63574
doi:10.2196/63574

© Punam Sawarkar, Gaurav Rajendra Sawarkar. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating a Social Robot Reception Service in an Italian Geriatric Hospital (HOSPER Study): Protocol for a Usability Study

Giulio Amabili[1], MEng; Arianna Sgolastra[1], MSc; Arianna Margaritini[1], MScPH; Federico Barbarossa[1], MEng; Elvira Maranesi[1], PhD; Giacomo Cucchieri[1], BEng; Lidia Pascucci[2], MD; Letizia Ferrara[1], MD; Roberta Bevilacqua[1], MScPH

[1]IRCCS INRCA (Istituto Nazionale di Ricovero e Cura per Anziani), Via Santa Margherita 5, Ancona, Italy

[2]Local Health Agency of Ancona, Ancona, Italy

Corresponding Author:
Elvira Maranesi, PhD
IRCCS INRCA (Istituto Nazionale di Ricovero e Cura per Anziani), Via Santa Margherita 5, Ancona, Italy

Abstract

Background: Humanoid robots are being used more frequently in health care: in geriatrics and pediatrics, robots such as "Pepper" and "Nao" have been shown to enhance communication, emotional well-being, and patient engagement. During the pandemic, their role expanded, for example, to include remote monitoring and reducing the risk of infection. Given these applications, our study examines the integration of 2 advanced humanoid robots, Pepper and Nao, within the reception area of a geriatric hospital.

Objective: The study aims to develop a reception service directed to enhance the quality of the IRCCS Istituto Nazionale di Riposo e Cura per Anziani (INRCA), Ancona, Italy. The primary end point of the study is the assessment of the perceived quality of an innovative reception service. Second, the study aims at evaluating the usability, human-robot interaction, user experience, and technical performance of the solution.

Methods: The HOSPER (analysis of the quality of reception services using robotic platforms, Nao and Pepper) study is a feasibility pilot project coordinated by IRCCS INRCA. During the trial, 200 older adults using services at IRCCS INRCA in Ancona will be enrolled. Two humanoid robots, Nao and Pepper, which are already available at the hospital, will be used to support reception and orientation services. Data collection will include general demographics, cognitive and functional status, usability measures, perceived service quality, and human-robot interaction indicators.

Results: The recruitment and data collection started in January 2025 and will be concluded by December 2025. Results regarding the feasibility of a robotic-based reception service in a geriatric hospital will be published in 2027.

Conclusions: The HOSPER study aims to assess the feasibility of an innovative robot-mediated reception service in a geriatric hospital, with specific focus on recruitment feasibility, completion rates, user acceptability, and perceived service quality. In fact, the usability and user experience evaluation will demonstrate the feasibility of the approach and provide indications for future developments and improvements.

Trial Registration: ClinicalTrials.gov NCT07002411; https://clinicaltrials.gov/study/NCT07002411

International Registered Report Identifier (IRRID): DERR1-10.2196/81395

(JMIR Res Protoc 2026;15:e81395) doi:10.2196/81395

KEYWORDS

social robots; hospital setting; reception service; human-robot interaction; HRI; older adults; usability; user experience; user-centered design

Introduction

Background

Hospital reception and orientation represent critical first touchpoints in the patient experience [1-3]. In geriatric settings, the process of navigating an unfamiliar hospital environment to reach the correct department or exam location is often associated with anxiety, confusion, and poor first impressions of care [4,5]. These challenges are particularly acute for older adults, who may experience navigation difficulties due to cognitive or sensory changes, physical fatigue, or limited familiarity with institutional settings [6,7]. Current human-staffed reception services, while important, are often resource-constrained and may not fully address the emotional and relational dimensions of a welcoming care experience [3,8]. In recent years, social and socially assistive robots (SAR)—autonomous machines with humanoid features designed to interact empathetically with humans—have demonstrated promise both in health care settings and in interactions with

XSL•FO
RenderX

older people [9,10]. Systematic reviews confirm the effectiveness of SAR in maintaining cognitive and physical functions, promoting sociability and emotional support, and reducing feelings of loneliness and anxiety among older adults [11,12]. These robots have been successfully deployed across diverse health care contexts, from therapeutic interventions with children and individuals with autism to supportive roles in geriatric care facilities [13-15]. Prior work highlights that SAR can engage older adults in meaningful interactions, improve mood and cooperation during care activities, and enhance perceived quality of care experiences [16,17]. A growing body of evidence suggests that robots possess specific affordances for the reception and orientation context: they can provide consistent, patient, and nonjudgmental guidance; they can communicate in multiple languages and adapt their communication style to individual needs; and their embodied, interactive presence may reduce anxiety and foster a sense of welcome that improves the initial hospital experience [18-20]. On this basis, we hypothesize that deploying SAR as reception agents can improve older adults' perceived quality of the reception and orientation experience, as well as the usability and acceptability of the service. However, current research on social robots and SAR in health care focuses on therapeutic, rehabilitative, or patient care activities. The application of SAR specifically to hospital reception and orientation services—particularly in the Italian health care context—remains understudied. Thus, a feasibility pilot is warranted to assess whether this innovation can be successfully integrated into routine hospital operations and whether it meaningfully enhances older adults' experiences during intake.

Goal of the Study

The HOSPER (analysis of the quality of reception services using robotic platforms, Nao and Pepper) study aims to assess the feasibility of assigning receptionist duties in an Italian geriatric hospital to two widely used SAR in research and health care settings, Nao and Pepper (both developed by SoftBank Robotics). These robots were selected because they offer multimodal interaction capabilities (voice, gesture, and visual interface), are highly customizable, and are already available at the IRCCS INRCA (Istituto Nazionale di Riposo e Cura per Anziani) hospital. Thus, this protocol focuses on improving the perceived quality of the reception and orientation experience for older adults, rather than directly measuring navigation performance (eg, time or errors). Therefore, the primary outcome is perceived service quality, with secondary outcomes capturing usability, human-robot interaction (HRI) quality, and emotional responses during robot-mediated guidance.

Methods

Study Design

The HOSPER study is a feasibility pilot study coordinated by IRCCS INRCA. Its aim is to assess the perceived quality of an innovative reception service through the introduction of 2 robotic platforms designated to provide information on how to reach the location where the person will receive the service of interest.

During the trial period, 200 older adults referred to IRCCS INRCA in Ancona will be enrolled. Figure 1 shows the study design, which consists of 2 different phases: recruitment (T0) and interaction (T1).

XSL•FO
RenderX

Figure 1. The study design. SPMSQ: Short Portable Mental Status Questionnaire; SUS: system usability scale.

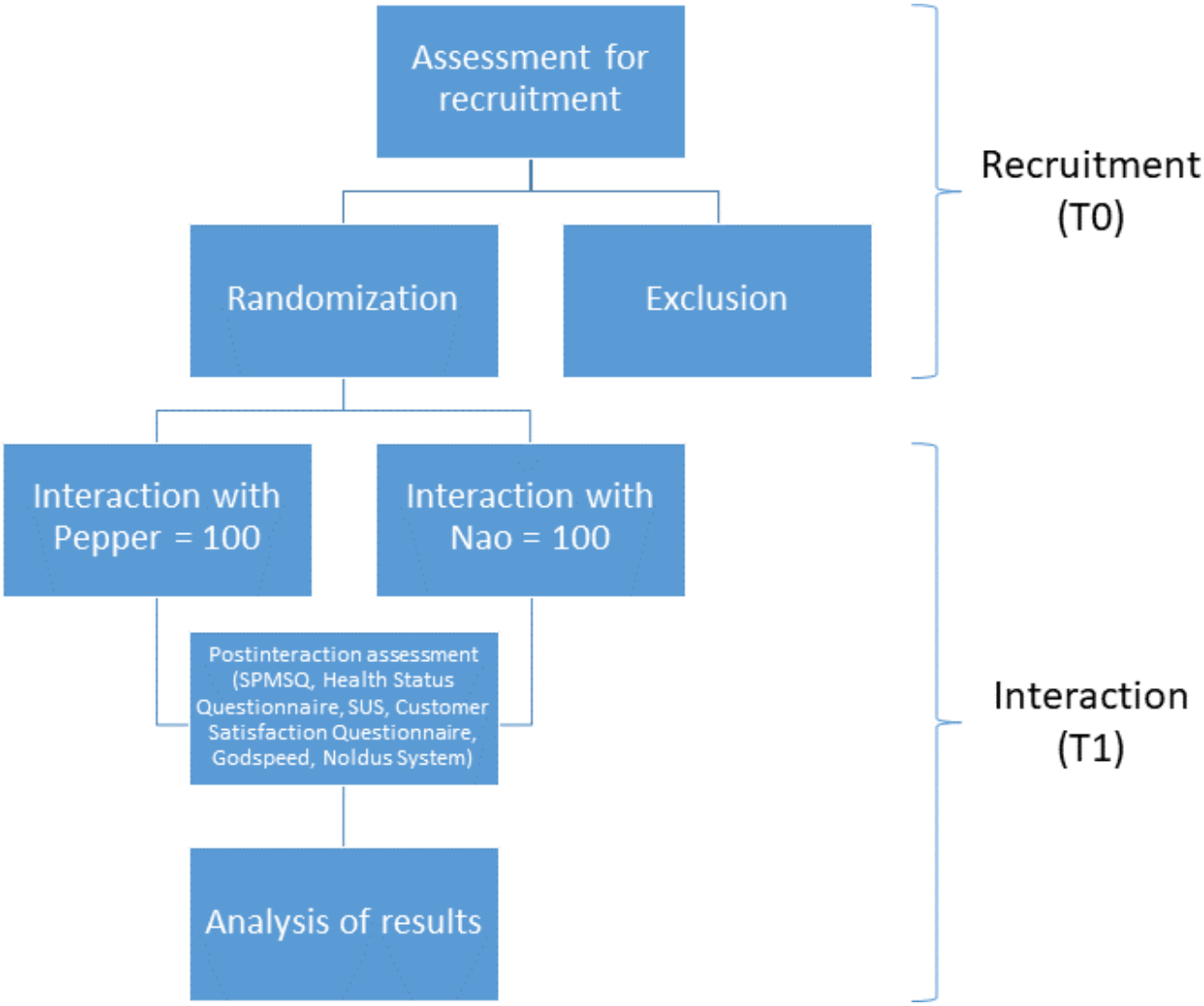

A total of 200 older adults will be recruited for this single-center feasibility pilot with a pragmatic quasi-experimental design. For logistical reasons, entire days will be dedicated to either Nao or Pepper. Participants will be equally allocated to interact with Nao or Pepper based on alternating study days and a predetermined quota system: once 100 participants have interacted with one robot, all subsequent participants will use the other robot until a total of 100 participants are reached for each. The allocation schedule will be predetermined and documented to ensure transparency.

Participants

The inclusion and exclusion criteria are reported in Textbox 1.

Textbox 1. Inclusion and exclusion criteria for patients.

Inclusion criteria

- Aged 65 years and older
- Capacity to consent
- Possession of full cognitive abilities (Short Portable Mental Status Questionnaire=0)

Exclusion criteria

- Failure to meet the inclusion criteria
- Concomitant participation in other interventional studies based on human-robot interaction
- Lack of written informed consent
- Pacemaker wearer
- Significant visual or hearing impairment

XSL•FO
RenderX

Equipment

For this study, 2 humanoid social robots already available at IRCCS INRCA will be used: NAO v6 (58 cm in height) and Pepper (120 cm in height), both developed by SoftBank Robotics. The HOSPER applications for both platforms were developed using the Choregraphe software (Nao v2.8.6.23 and Pepper: v2.5.10.7; SoftBank Robotics), with standardized interaction flows programmed via “choice” objects handling up to 30 ward inputs. Interactions follow identical scripted sequences: they are user-initiated by the keyword pronunciation (HOSPER), followed by a fixed welcome prompt, speech recognition (ie, desired ward to reach), path directions provided by the robot (during this passage Pepper also shows a visual map on the tablet), repetition on request (≤40% understanding), and standardized closure. Detailed technical development is reported in Amabili et al [21]. Figure 2 shows the dedicated area and one of the robots. This is a closed box in which the robots will be located. Figure 2A shows the experimental setting, while Figure 2B focuses on the experimental setting with the high-definition camera that will record HRI.

XSL•FO
RenderX

Figure 2. The experimental setting. (A) Dedicated area for the study with 1 robot (Pepper). (B) Focus on the experimental setting with a high-definition camera to record the human-robot interaction.



Outcomes

Overview

The primary end point of the study is the assessment of the perceived quality of an innovative reception service, as measured by the Customer Satisfaction Questionnaire.

The secondary end points are (1) assessment of the usability of 2 robotic platforms by older adults for hospital reception, as measured by the system usability scale (SUS) [22]; (2) analysis of HRI using the Godspeed Questionnaire [23]; (3) dynamic analysis of the interaction and user experience using the Noldus system [24]; and (4) analysis of robot logs to assess technical performance.

XSL•FO
RenderX

As this is a feasibility study, these outcomes will be interpreted descriptively. For the primary end point, we will examine mean scores on the Customer Satisfaction Questionnaire and the proportion of participants reporting medium-to-high satisfaction levels to judge whether the robot-mediated reception service is acceptable to older adults. For the secondary end points, the distributions of SUS, Godspeed Questionnaire scores, and Noldus-derived emotional indicators will be used to evaluate whether usability, HRI quality, and emotional responses are compatible with moving towards a larger-scale trial. In addition to these measures, the study will use standard feasibility indicators, including recruitment rate, the proportion of participants who complete the interaction and postinteraction assessments, and completeness of data collected at T0 and T1. These indicators will be summarized descriptively to judge whether the protocol is feasible in routine practice. Finally, qualitative notes on technical performance (eg, malfunctions, interruptions, or need for operator intervention) will be collected throughout the study both during the sessions and through the analysis of video recordings to inform the system-level feasibility of the robot-mediated reception service and guide future updates of the service.

While outcomes focused on usability and HRI emphasize validated patient-reported measures, triangulation is provided through existing data sources: Noldus video analysis for objective emotional and behavioral indicators, and robot logs capturing interaction duration, comprehension rates, and repetition requests as objective efficiency metrics. In fact, these complementary data sources mitigate self-report limitations while remaining within the approved protocol scope.

Finally, as this is a feasibility pilot, a priori success criteria are defined. The study will be considered feasible if: (1) at least 70% of eligible older adults approached agree to participate; (2) at least 70% of enrolled participants complete the interaction with the assigned robot and the postinteraction assessment; and (3) questionnaire data are complete for at least 80% of enrolled participants. These thresholds are intended to inform the design and sample size of a subsequent definitive trial rather than to support hypothesis testing.

Sociodemographic Questionnaire

This is a simple tool used to collect key personal data from study participants. The information collected includes date of birth (dd/mm/yy), sex (male, female, or intersex), current marital status (married, cohabiting, separated, divorced, single, widowed, and do not know or refuse to answer), level of education (no education, primary education, secondary education, tertiary education, and do not know or refuse), total years of education and current employment status (retired, full-time employment, part-time employment, unemployed, working from home, do not know, and refuse), details of last job, current source of income (work, pension, annuity, help from relatives, social security benefits, other, do not know, and refuse), and possible cohabitants (none, spouse or partner, children, grandchildren, sons-in-law or daughters-in-law, brothers or sisters, mother or father, assistant, others, and refuse to answer).

The Short Portable Mental Status Questionnaire or the Pfeiffer Test

This is a tool designed to assess cognitive impairment in older patients [25]. This tool was designed to provide health care professionals working with older patients with an effective and simple method of assessing cognitive impairment. The Short Portable Mental Status Questionnaire is a short questionnaire consisting of 10 questions investigating certain aspects of cognitive ability: 7 on orientation (spatial, temporal, personal, and environmental), 2 on long-term memory (telephone number or address and mother's maiden name), and 1 on concentration (serial subtraction). Questions should be asked in the order they appear in the protocol, and a score of 0 is assigned for each unanswered or incorrect question, or 1 for each correct answer. The maximum possible score ranges from 0 (indicating maximum cognitive impairment) to 10 (indicating no cognitive impairment). A correction must be applied to the obtained score based on the participant's level of education: if the number of errors is 0, no correction is made for the patient's level of education; if the patient's level of education is less than 8 years, subtract 1 from the raw score; if the patient's level of education is 8 years, the raw score is not modified; and (4) if the patient's level of education is more than 8 years, add 1 point to the raw score.

The results are interpreted based on the number of errors made: 0 to 2 errors (cognitive functions intact), 3 to 4 errors (indicative of mild cognitive decline), 5 to 7 errors (indicative of moderate cognitive decline), and 8 to 10 errors (severe cognitive decline).

This quick tool helps to define the presence and intensity (mild, moderate, and severe) of cognitive disorders of organic origin in older patients. It can be used by medical staff and other healthcare professionals in any setting. It can also be used with blind people.

Health Status Questionnaire

As well as asking for the participant's age, the questionnaire uses dichotomous questions and scales ranging from 1 to 5 to assess their visual and hearing abilities. For example, it asks whether they can read a book or watch television with adequate lighting, with or without lenses, and whether they can follow a conversation or use the telephone, even with a hearing aid. The user will assess their own sense of sight and hearing, rather than the researcher.

SUS Tool

The SUS is a reliable tool for measuring usability. It consists of a 10-item questionnaire in which respondents are asked to indicate their level of agreement or disagreement with a series of statements using a 5-point scale ranging from "Strongly Agree" to "Strongly Disagree." The SUS allows for the evaluation of a wide variety of products and services, including hardware, software, mobile devices, websites, and applications. The SUS is easy to administer and can produce reliable results in small samples, effectively distinguishing between usable and unusable systems.

XSL•FO
RenderX

Customer Satisfaction Questionnaire

This tool is designed to assess user satisfaction after using the 2 robotic platforms. Satisfaction is assessed by quantifying various variables, such as: the attention received from the robot in terms of accuracy and courtesy (scoring from 1="not at all satisfied" to 7="very satisfied"), the clarity and completeness of the information received (scoring from 1="not at all satisfied" to 7="very satisfied"), respect for personal privacy (scoring from 1="not at all satisfied" to 7="very satisfied"), overall satisfaction with the service offered (scoring from 1="not at all satisfied" to 7="very satisfied"), and whether they would recommend our service to others (scoring from 1="absolutely not" to 7="definitely yes").

Godspeed

It allows users to quantify certain characteristics of the platform perceived during use, using a score ranging from 1 to 5 for each item. In particular, it refers to anthropomorphism, animation, friendliness, intelligence, and perceived safety.

Procedure

Recruitment (T0)

The robots will be placed in a dedicated area (a closed box) at the hospital entrance. Only one user will be allowed in at a time. A researcher will stop users at the hospital entrance and invite them to participate in the study. Participants must confirm their willingness to participate in writing. After signing the consent forms, the technology will be assigned to them in a randomized alternating manner (ABAB) according to the order of inclusion of participants. Then, the researcher will collect general information about the patient and administer the cognitive scale to verify that the user meets the inclusion criteria.

Interaction With the Robot (T1)

The user, who has been deemed suitable and agreed to participate, starts the interaction with the robot assigned. In both cases, the researcher will suggest touching the robot's head to initiate the interaction. From this point onwards, the researcher will not intervene in the interaction between the user and the robot, and the user will be free to interact with the robot and to stop the interaction at any time. When an explicit support request is made by the user, the researcher can intervene a maximum of 3 times, in line with good practice in experiments involving HRI analysis. At this point, the robot greets the user and asks how they can be helped, using the phrase "Good morning and welcome to INRCA. I'm (robot's name) and I'm here to provide useful information to help you reach your desired destination within the hospital. Which department do you need to reach, or which exam do you need to take? Please, answer by saying only the name of the department or the exam as soon as my eyes turn blue." At this point, the user responds, and the robot replies in one of two ways:

1. If it does not understand the request (or if the request is outside the robot's knowledge), it invites the user to rephrase it and they return to the previous point; and
2. If it understands the request, it explains the path that the user should follow. In the case of Pepper, this path will also be shown on a display with photographic materials integrated with directional arrows.

After 1 successful interaction or 2 failed attempts, the robot asks the user if it has provided all the necessary information. If so, the robot says goodbye to the user as follows: "Thank you for participating. Have a nice day." Otherwise, the interaction resumes from the point where the robot asks for the department to be reached or the test to be performed.

The researcher suggests that the user should touch the robot's head to stop the session only in the case of an interaction loop or if the user is willing to stop the interaction.

All interactions will be recorded on video using a GoPro camera placed behind the robot. The robot will be placed in a dedicated area, and only one user will be permitted at a time to ensure that no data is collected about "unaware individuals." These videos will be used for postanalysis to assess the emotional aspect of the interaction using the FaceReader system of Noldus in the "Youse" Usability Lab, which is located within the library of the Institute.

Postinteraction Assessment

A brief interview will be administered after the interaction to assess satisfaction with the service. The SUS scale will be used to measure usability grade and the HRI scale will be used to assess HRI.

Statistical Analysis

Sample size determination followed feasibility study conventions rather than power calculations for confirmatory testing. Previous studies evaluating patient satisfaction with novel health care service innovations enrolled 124 to 177 participants [26,27]. To provide robust descriptive data across both feasibility indicators (recruitment, completion, and acceptability) and the balanced quasi-experimental design (n=100 per robot), the HOSPER study targets 200 participants (100 Nao and 100 Pepper). This sample size supports reliable estimation of prevalence estimates (eg, satisfaction proportions) while remaining practical for a single-center pilot. In the case that the statistical analysis reveals a positive trend, a new protocol will be proposed with the aim of conducting a trial.

Continuous variables will be reported as mean and SD, or median and IQR, depending on their distribution (as assessed using the Kolmogorov-Smirnov). Categorical variables will be expressed as absolute numbers and percentages.

Descriptive statistical analysis will be conducted on quantitative data with SPSS (Statistical Package for the Social Sciences; IBM Corp) and RStudio (Posit PBC). The robot's logs (computer records) will be extracted in order to make a technical assessment of the objective parameters of HRI: dialogue comprehension ability, questions asked, answers given, dialogue duration, and so on.

Moreover, HRI will be analyzed using the Noldus system, considered the gold standard for behavioral analysis. Specifically, FaceReader and Observer (Noldus Information Technology) will be used to analyze audio-video recording of interactions. FaceReader uses facial analysis to identify the

XSL•FO
RenderX

Ekman 7 emotions (neutrality, happiness, anger, sadness, disgust, fear, and surprise) in real-time [28] and summarize emotional valence distributions. Instead, the Observer enables the analysis of the interaction quantitatively and qualitatively by manually inserting labels to code the user's emotional state and behaviors. Results will describe emotional and behavioral patterns during robot interactions to complement self-reported questionnaires. These analyses can be integrated and stacked with the Noldus software, enabling correlations to be found between the robot's actions and the user's behavior. This observational approach is useful for describing the dynamics of interaction and assessing the user's experience.

Finally, the comparison between groups (Nao vs Pepper) will be carried out with a 2-tailed *t* test or the Mann-Whitney *U* test and the Fisher exact test considering the variables typology and distribution. As this is a feasibility study, these comparisons are descriptive only and not powered for formal inference. Linear or logistic regression models may be used exploratorily to examine potential confounders, but results will not support causal conclusions.

Due to the enrollment of the older population, despite the short duration of the testing and the presence of a researcher, it could happen to have missing data. Missing data will be reported by variable and time point. Moreover, available case analysis will be used for descriptive summaries, with complete-case denominators for primary outcome and feasibility indicators.

Risks and Benefits Analysis

This study involves risks and benefits for participants. Expected benefits include the possibility that participants may feel more supported during reception, thanks to standardized information and dedicated guidance from the robots. Risks are minimal, but not absent. There is a risk that older people will stop interacting with the robots (eg, because they dislike Pepper and Nao or because they are afraid to interact with 2 humanoid robots). If this were to happen, the session would be interrupted and concluded immediately.

As the presence of the 2 robots could cause distress, participants will interact with them in an environment that guarantees an adequate level of privacy during the interaction.

To protect the security of participants, the following measures have been adopted:

1. Placement of robots within a dedicated structure that allows private data collection without external interference and ensures a human-robot interaction that respects user privacy and security;
2. The robots will function in plug-in mode to avoid service interruption and to limit their free movements. They will remain static, although they are free to gesticulate and move their heads, as these are essential aspects of empathetic interaction;
3. Clear and complete information about the appropriate use of the robots (eg, they must not lean on the devices or make any movements that could destabilize them);
4. The researcher will be present in the study area throughout the interactive session to adequately train participants on the use of robots and provide assistance in case of difficulties or problems, without impacting interaction;
5. All phases will also be illustrated with easily interpretable iconographic material to guide the researcher.

Despite the precautions described above, adverse events may still occur: (1) participants will be asked not to touch or disconnect the robot in case of problems or malfunctions; (2) participants should call the researcher if they experience any difficulties; and (3) when called by participants, the researcher will ensure that no damage has been caused to anyone or to the robots.

Although risks are expected to be low, they are not negligible. Some older adults may experience temporary fear, embarrassment, frustration, or discomfort when interacting with humanoid robots, or may decide to interrupt the interaction. To manage these risks, a researcher (either a psychologist or a physician) will be present throughout the sessions, participants will be reminded that they can stop at any time without consequences for their care, and any signs of distress will lead to immediate interruption of the interaction. A further potential risk is overattachment to or over-reliance on the robots. This will be mitigated by providing clear information that the robots are tools to support reception only, by limiting the duration of interactions, and by ensuring that human staff remain available at all times for conventional reception and orientation. Overall, the controlled environment, explicit right to withdraw, and continuous monitoring by trained researchers are expected to keep psychological risks within acceptable limits for a feasibility pilot.

Finally, if participants request longer interaction with these devices at the end of the study, they will be asked to participate in similar studies to ensure continuity in the use of the technological devices used.

Ethical Considerations

Ethical Approval

The study was approved by the Marche Territorial Ethics Committee (CET MARCHE) and registered under the protocol 42021 at INRCA on December 16, 2024. The study protocol was recorded in ClinicalTrials.gov (NCT07002411). The principles of the Declaration of Helsinki and Good Clinical Practice guidelines were adhered to.

Data Management

All participants in this study will provide written informed consent. Throughout all procedures, including screening, recruitment, testing, evaluation, and dissemination, the project is committed to maintaining the anonymity and confidentiality of participants. Procedures for data collection, use, and storage comply with national legislation, the General Data Protection Regulation (GDPR) and the European Union's Data Protection Impact Assessment (DPIA), including participants' rights to access their data, to be informed, to withdraw consent, and to request data deletion. Data will be collected in accordance with the principle of data minimization, that is, only the personal information that is directly relevant and necessary for achieving the specific objectives of the testing and evaluation work

XSL•FO
RenderX

packages will be collected. Data entry will be performed using dedicated software that includes checks and validations to reduce input errors. All screening data will be deleted upon project completion. During testing procedures, all visual, auditory, and sensory data collected and processed by the robot to function as intended will be discarded once the procedures have been completed. The only exception is the recording of the number of interactions the robot has with each participant. However, these interactions will remain anonymous. All research data will be made openly available for secondary analysis 3 years after the project's completion.

The data collected and processed by the robot is only stored locally (ie, within the robot itself) and is not transmitted over the internet or cloud services. Only the relevant data mentioned in the "Statistical analysis" section will be extracted from the log.

These data will be anonymous because it will be impossible to trace the identity of those involved in the interaction. The logs only contain information about the functions activated by the robot (eg, spoken phrases and gestures) and the phrases spoken by the user, along with the percentage of understanding achieved by the robot. The robots do not save any audio-video data. In the Noldus system, data are also saved and analyzed locally to avoid any internet or cloud sharing. Data imported into the system will be MP4 files from GoPro camera recordings. These files will be named using a code that identifies the user without revealing their identity (eg, "User_01" or "User_01_ddmmyyyy" to link to the date of interaction). Both the robot logs and the Noldus analyses will be saved on INRCA computers assigned to the Youse lab and protected by passwords, making them accessible only to the study's research team.

Research data will remain available for secondary analyses for up to 3 years after study completion, as declared in the approved protocol.

Legal and Technical Aspects

The study will be conducted in accordance with all relevant regulatory and legal requirements and received approval from the Marche Territorial Ethics Committee (CET MARCHE).

Additionally, all potentially eligible participants must receive full information about the study and provide their consent to participate. Participants must also consent to the processing of their personal data in an anonymous and aggregated form, in accordance with EU Regulation 2016/679 (GDPR), the Italian Data Protection Code (APPI), and Legislative Decree No. 101/2018, which adapts national regulations to the European Regulation 2016/679. Participants must be informed that authorized personnel, members of the relevant ethics committee, and officials from the competent regulatory authorities may review their data. They will also be asked to provide specific informed consent to participate in the study, including consent for data retention for up to 3 years after the study's completion.

All study information and consent forms will be explained verbally in plain, nontechnical language by a trained researcher. Also, flyers that present the project in an understandable language will be available at the pilot site. Moreover, participants may have a family member or caregiver present during the explanation if they wish. Only individuals who demonstrate understanding of the study and voluntarily agree will provide written informed consent.

Each signature must be dated by the signatory in person, and the informed consent form and any additional patient information must be retained by the investigator. Each patient will receive a signed copy of the informed consent form and the information sheet.

Participant information and consent forms are included in the documentation submitted to the local ethics committee for approval. No compensation was provided to participants.

Participants may opt to consent to the storage and use of their data long after the end of the project as part of open access to scientific publications and open research data, as required by the European Commission.

Results

The recruitment started in January 2025 and will continue until December 2025. The recruited end users are reported in Table 1, grouped by phase of study. The assessment of the perceived quality and usability of the reception service and the analysis of HRI aim to provide an overall overview of the experience of older users with robots in a hospital setting.

XSL•FO
RenderX

Table . Results and clinical assessment.

Dimension	Scale	T0[a]	T1[b]
General information	Sociodemographic Questionnaire	✓	
Cognitive status	SPMSQ[c]	✓	
Functional status	Health Status Questionnaire	✓	
Usability	SUS[d]		✓
Perceived quality	Customer Satisfaction Questionnaire		✓
HRI[e]	Godspeed		✓
HRI	Noldus system		✓

[a]T0: recruitment.

[b]T1: assessment at the end of the interaction.

[c]SPMSQ: Short Portable Mental Status Questionnaire.

[d]SUS: system usability scale.

[e]HRI: human-robot interaction.

The SUS will be used to quantify the level of perceived usability for each of the 2 robotic platforms tested. The score obtained will also allow a direct comparison between the 2 systems, to explore potential differences in usability and user experience between Nao and Pepper, without being powered for definitive conclusions.

The Godspeed Questionnaire will be used to identify aspects related to the perception of robotic systems, such as the degree of anthropomorphism, animacy, likeability, perceived intelligence, and safety. These data may suggest useful ideas for optimizing the behavioral design of social robots.

Finally, the Noldus system will enable a dynamic analysis of the interaction to be conducted, based on objective and coded observation of the user behavior during interaction with the social robot. In addition to the temporal analysis of action sequences (eg, response times), the Noldus will also detect the emotions elicited during the interaction, thus providing additional qualitative elements useful for understanding the user's subjective experience and the emotional impact of the interaction.

The feasibility of the HOSPER study will be evaluated and discussed in a separate publication expected to be published by 2027.

Discussion

The HOSPER study addresses the feasibility of providing an innovative reception service within the IRCCS INRCA hospital in Ancona through social robots. This feasibility pilot is expected to show that a social robot–based reception service is acceptable to older adults accessing a geriatric hospital, with adequate recruitment, interaction completion, and questionnaire completion rates according to the predefined feasibility thresholds. It is hypothesized that the service will be associated with favorable satisfaction and usability profiles, suggesting potential to improve older adults' reception experience, while also revealing technical and organizational constraints that must be addressed before routine implementation. These anticipated findings will primarily inform refinement of the service and the design of a subsequent, adequately powered trial, rather than demonstrating the effectiveness of the solution. A distinctive feature of the protocol is its person-centered design, which integrates standard usability measurements (ie, SUS) with robots' logs and measures of participants' subjective experience, including dynamic behavioral observation through the Noldus system.

Previous studies have shown that social assistive robots such as Pepper and Nao can support communication, emotional well-being, and engagement among older adults and other vulnerable groups, including in hospital and long-term care settings [13,27,29]. However, most prior work has focused on ward-based activities, companionship, or therapy, whereas HOSPER concentrates on the reception and orientation phase, which represents a critical aspect in the patient journey and has been less systematically explored [2,30]. Thus, this study extends existing evidence on social robotics in healthcare towards front-desk services in a geriatric hospital context.

The target population in this pilot was required to show full cognitive capacity to ensure participant safety and data quality during the first implementation of the robot-mediated reception service. This choice was intended to avoid both excessive frustration among participants and the introduction of major confounding factors in the interpretation of feasibility outcomes (eg, completion rates, usability, and satisfaction scores), but it also limits generalizability to older adults with cognitive impairment, who represent a substantial proportion of the geriatric population accessing the hospital. Once feasibility, usability, and basic reliability of the solution are established in this cognitively intact group, extending the interaction to older people with mild cognitive impairment will therefore be an important next step for future studies. Similarly, another limitation of the present study is the continuous presence of a dedicated researcher during robot interactions, which does not represent a real-world scenario. However, this controlled set-up is appropriate for early-stage feasibility and will inform how responsibilities could be progressively transferred to regular reception staff in later implementation phases, including training

XSL•FO
RenderX

needs and procedures in case of technical failures. Further constraints arise from the need to operate the robots offline, which limits vocabulary size and dynamic updates, and from the controlled environment that does not fully reproduce the ambient noise and congestion of the hospital hall. In addition, potential resistance or skepticism from reception staff and clinicians, as well as unexpected breakdowns in robot functioning in a busy lobby, are anticipated challenges that will need to be monitored and addressed in future implementation phases through targeted training and clear repair procedures.

From an implementation science perspective, this feasibility pilot primarily focuses on intervention-level characteristics (usability, acceptability, and basic technical reliability) rather than a full evaluation of organizational context and economic impact. Therefore, only preliminary insights on the inner setting (eg, compatibility with current reception workflows and staff workload) will be obtained through the researcher's observations, and no formal cost or cost-effectiveness analysis will be conducted at this stage. Future phases, informed by implementation frameworks, will need to explicitly examine organizational readiness, long-term integration into hospital processes, and the economic implications of scaling a robot-based reception service for administrators and policymakers [31-33].

Data Availability

The datasets generated during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The Nao and Pepper robots used in this study are part of the standard equipment of IRCCS INRCA and were not provided or funded by SoftBank Robotics, but were purchased as part of other research projects prior to this activity. The authors declare that they have no financial or commercial relationship with SoftBank Robotics or any other robot manufacturer and/or company that could be construed as a potential conflict of interest.

References

1. Gueye B, Bop MC, Diagne N, et al. Assessment of satisfaction in the reception and emergency department of a health center. Open J Epidemiol 2024;14(2):284-296. [doi: 10.4236/ojepi.2024.142020]
2. Gualandi R, Masella C, Viglione D, Tartaglini D. Exploring the hospital patient journey: what does the patient experience? PLoS One 2019;14(12):e0224899. [doi: 10.1371/journal.pone.0224899] [Medline: 31805061]
3. Ferreira DC, Vieira I, Pedro MI, Caldas P, Varela M. Patient satisfaction with healthcare services and the techniques used for its assessment: a systematic literature review and a bibliometric analysis. Healthcare (Basel) 2023 Feb 21;11(5):639. [doi: 10.3390/healthcare11050639] [Medline: 36900644]
4. Pouyan AE, Ghanbaran AH, Hosseinzadeh A, Shakibamanesh A. The elderly wayfinding performance in an informative healthcare design indoors. J Build Eng 2024 Jun;87:108843. [doi: 10.1016/j.jobe.2024.108843]
5. Davis R, Veltkamp A. Wayfinding strategies and spatial anxiety in older adults with and without Alzheimer's disease. Res Gerontol Nurs 2020 Mar 1;13(2):91-101. [doi: 10.3928/19404921-20191022-03] [Medline: 31697391]
6. Wallace LG, Bradway CK, Cacchione PZ. The relationship between sensory loss and health literacy in older adults: a systematic review. Geriatr Nurs 2022;47:1-12. [doi: 10.1016/j.gerinurse.2022.06.003] [Medline: 35779376]
7. Reed N, Yeh C. Sensory loss and the healthcare system: outcomes and navigation. Innov Aging 2020 Dec 16;4(Suppl 1):797. [doi: 10.1093/geroni/igaa057.2888]
8. Litchfield I, Gale N, Burrows M, Greenfield S. The future role of receptionists in primary care. Br J Gen Pract 2017 Nov;67(664):523-524. [doi: 10.3399/bjgp17X693401] [Medline: 29074695]
9. Maranesi E, Amabili G, Cucchieri G, Bolognini S, Margaritini A, Bevilacqua R. Understanding the acceptance of IoT and social assistive robotics for the healthcare sector: a review of the current user-centred applications for the older users. In: Scataglini S, Imbesi S, Marques G, editors. Internet of Things for Human-Centered Design: Springer Nature Singapore; 2022:331-351. [doi: 10.1007/978-981-16-8488-3_16]
10. Ghafurian M, Chandra S, Hutchinson R, et al. Systematic review of social robots for health and wellbeing: a personal healthcare journey lens. ACM Trans Hum-Robot Interact 2025 Mar 31;14(1):1-48. [doi: 10.1145/3700446]
11. Pu L, Moyle W, Jones C, Todorovic M. The effectiveness of social robots for older adults: a systematic review and meta-analysis of randomized controlled studies. Gerontologist 2019 Jan 9;59(1):e37-e51. [doi: 10.1093/geront/gny046] [Medline: 29897445]
12. Lim J. Effects of a cognitive-based intervention program using social robot PIO on cognitive function, depression, loneliness, and quality of life of older adults living alone. Front Public Health 2023;11:1097485. [doi: 10.3389/fpubh.2023.1097485] [Medline: 36815168]
13. Lu LC, Lan SH, Hsieh YP, Lin LY, Lan SJ, Chen JC. Effectiveness of companion robot care for dementia: a systematic review and meta-analysis. Innov Aging 2021;5(2):igab013. [doi: 10.1093/geroni/igab013] [Medline: 34316517]

XSL•FO
RenderX

14. Amabili G, Cucchieri G, Margaritini A, et al. Social robotics and dementia: results from the eWare Project in supporting older people and their informal caregivers. Int J Environ Res Public Health 2022 Oct 16;19(20):13334. [doi: 10.3390/ijerph192013334] [Medline: 36293915]
15. Rakhymbayeva N, Amirova A, Sandygulova A. A long-term engagement with a social robot for autism therapy. Front Robot AI 2021;8:669972. [doi: 10.3389/frobt.2021.669972] [Medline: 34222353]
16. Figueroa D, Yamazaki R, Nishio S, et al. Social robot for older adults with cognitive decline: a preliminary trial. Front Robot AI 2023;10:1213705. [doi: 10.3389/frobt.2023.1213705] [Medline: 38077459]
17. Murayama H, Takase M. Evaluating the effectiveness of digital social robots in reducing loneliness among community-dwelling older adults in Japan: randomized controlled trial and qualitative analysis. JMIR Aging 2025 Oct 22;8:e74422. [doi: 10.2196/74422] [Medline: 41124645]
18. Salins PL, Anandan G, Ananda BD, Reshmi B, Jathanna RD. Internet of things–based wayfinding for hospital visitors: a digital solution for complex health care infrastructures. Mayo Clin Proc Digit Health 2025 Dec;3(4):100293. [doi: 10.1016/j.mcpdig.2025.100293] [Medline: 41245354]
19. Getson C, Nejat G. The adoption of socially assistive robots for long-term care: during COVID-19 and in a post-pandemic society. Healthc Manage Forum 2022 Sep;35(5):301-309. [doi: 10.1177/08404704221106406] [Medline: 35714374]
20. Blavette L, Dacunha S, Alameda-Pineda X, et al. Acceptability and usability of a socially assistive robot integrated with a large language model for enhanced human-robot interaction in a geriatric care institution: mixed methods evaluation. JMIR Hum Factors 2025 Aug 1;12:e76496. [doi: 10.2196/76496] [Medline: 40750072]
21. Amabili G, Barbarossa F, Maranesi E, et al. Design and implementation of a social robot reception service in an Italian geriatric hospital. Presented at: 2025 IEEE Cyber Science and Technology Congress (CyberSciTech); Oct 21-24, 2025; Hakodate, Japan p. 777-781. [doi: 10.1109/CyberSciTech68397.2025.00118]
22. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry: CRC Press; 1996:189-194.
23. Bartneck C, Kulić D, Croft E, Zoghbi S. Godspeed questionnaire series [Database record]. APA PsycTests. 2019 Apr 8. URL: https://doi.org/10.1037/t70855-000 [accessed 2026-02-16]
24. Skiendziel T, Rösch AG, Schultheiss OC. Assessing the convergent validity between the automated emotion recognition software Noldus FaceReader 7 and Facial Action Coding System scoring. PLoS One 2019;14(10):e0223905. [doi: 10.1371/journal.pone.0223905] [Medline: 31622426]
25. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. J Am Geriatr Soc 1975 Oct;23(10):433-441. [doi: 10.1111/j.1532-5415.1975.tb00927.x] [Medline: 1159263]
26. Pitsch K, Wrede S, Seele JC, Süssenbach L. Attitude of German museum visitors towards an interactive art guide robot. Presented at: 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI); Mar 8-11, 2011. [doi: 10.1145/1957656.1957744]
27. Blavette L, Rigaud AS, Anzalone SM, et al. A robot-mediated activity using the Nao robot to promote COVID-19 precautionary measures among older adults in geriatric facilities. Int J Environ Res Public Health 2022 Apr 25;19(9):5222. [doi: 10.3390/ijerph19095222] [Medline: 35564615]
28. Ekman P. An argument for basic emotions. Cognition and Emotion 1992 May;6(3-4):169-200. [doi: 10.1080/02699939208411068]
29. Liao YJ, Jao YL, Boltz M, et al. Use of a humanoid robot in supporting dementia care: a qualitative analysis. SAGE Open Nurs 2023;9:23779608231179528. [doi: 10.1177/23779608231179528] [Medline: 37324571]
30. Tanner A, Urech A, Schulze H, Manser T. Older adults' engagement and mood during robot-assisted group activities in nursing homes: development and observational pilot study. JMIR Rehabil Assist Technol 2023 Dec 25;10:e48031. [doi: 10.2196/48031] [Medline: 38145484]
31. Gasteiger N, Zhu T, Broadbent E, Lim J, MacDonald BA, Ahn HS. Social robots in hospital settings: an initial exploration of the services provided, interaction style and in the field evaluation. Presented at: 19th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI 2024); Mar 11-14, 2024. [doi: 10.1145/3610978.3640714]
32. Kirk MA, Kelley C, Yankey N, Birken SA, Abadie B, Damschroder L. A systematic review of the use of the consolidated framework for implementation research. Implement Sci 2016 May 17;11:72. [doi: 10.1186/s13012-016-0437-z] [Medline: 27189233]
33. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]

Abbreviations

GDPR: General Data Protection Regulation
HOSPER: analysis of the quality of reception services using robotic platforms, Nao and Pepper
HRI: human-robot interaction
IRCCS INRCA: Istituto Nazionale di Riposo e Cura per Anziani
SAR: socially assistive robot

XSL•FO
RenderX

SPSS: Statistical Package for the Social Sciences
SUS: system usability scale

Edited by J Sarvestan; submitted 28.Jul.2025; peer-reviewed by R Marshall; revised version received 31.Dec.2025; accepted 31.Dec.2025; published 26.Feb.2026.

Please cite as:
Amabili G, Sgolastra A, Margaritini A, Barbarossa F, Maranesi E, Cucchieri G, Pascucci L, Ferrara L, Bevilacqua R
Evaluating a Social Robot Reception Service in an Italian Geriatric Hospital (HOSPER Study): Protocol for a Usability Study
JMIR Res Protoc 2026;15:e81395
URL: https://www.researchprotocols.org/2026/1/e81395
doi:10.2196/81395

© Giulio Amabili, Arianna Sgolastra, Arianna Margaritini, Federico Barbarossa, Elvira Maranesi, Giacomo Cucchieri, Lidia Pascucci, Letizia Ferrara, Roberta Bevilacqua. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Transcutaneous Electrical Nerve Stimulation vs Transcutaneous Electrical Nerve Stimulation With Music in Persistent Low-Back Pain: Protocol for a Pilot Feasibility Trial

Seema Radhakrishnan[1,2,3,4], PhD; Michael Glynn[1], MSc; Anne-Marie Robertson[1], MBBS; Anna Fedyczkowski[1], BN; Lieu Trinh[1], PhD; Sabrina Naz[5], PhD

[1]Pain Management Centre, Westmead Hospital, Cnr Darcy Road and Hawkesbury Road, Westmead, Australia

[2]Westmead Clinical School, University of Sydney, Sydney, Australia

[3]School of Population Health, University of New South Wales, Sydney, Australia

[4]Department of Geriatrics and Rehabilitation, Fairfield Hospital, Prairiewood, Australia

[5]Pain Management Centre, Western Sydney Local Health District, Research and Education Network (REN), Westmead, Australia

Corresponding Author:
Seema Radhakrishnan, PhD
Pain Management Centre, Westmead Hospital, Cnr Darcy Road and Hawkesbury Road, Westmead, Australia

Abstract

Background: Transcutaneous electrical nerve stimulation (TENS) works on the principle of the gate control theory of pain and is used as a nonpharmacological pain management intervention. Music therapy or listening to self-selected music has also been shown to reduce pain intensity, anxiety, and depression. There has been no published literature that has explored whether these modalities can work synergistically.

Objective: This study aims to test the feasibility of a clinical trial comparing TENS alone to TENS combined with self-selected music in participants with persistent low-back pain lasting more than 3 months.

Methods: This will be a prospective, randomized controlled pilot study to compare TENS alone vs TENS combined with participant self-selected music. We hope to enroll 20 participants. The participants will be their own controls, and the order in which the intervention is delivered will be determined by randomization.

Results: The study is open for recruitment from August 2025. The results are expected in 18 months from the start of recruitment. The primary outcome will be the feasibility of the intervention, monitored by adherence rate, dropout rate, and the percentage of the eligible population that consented. Secondary outcomes collected will include pain intensity; mood using the 21-item Depression, Anxiety, and Stress Scale; medication consumption; participant satisfaction; and adverse reactions.

Conclusions: If the pilot trial is feasible and indicates a positive synergistic trend with TENS and music, the trial will help to inform the feasibility for a larger trial with more varied pain diagnoses, including acute and chronic pain.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12625000552482; https://anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12625000552482

International Registered Report Identifier (IRRID): PRR1-10.2196/82382

(JMIR Res Protoc 2026;15:e82382) doi:10.2196/82382

KEYWORDS

transcutaneous electrical nerve stimulation; TENS; music therapy; chronic pain; pilot trial; spinal pain

Introduction

Background

Low-back pain is the leading cause of disability across the globe [1]. Pharmacological interventions such as opioids, when used as a long-term management strategy for persistent spinal pain, can result in significant adverse effects [2]. Nonpharmacological noninvasive interventions that facilitate endogenous pain modulatory pathways have the potential to provide improvement in pain and function while minimizing adverse events [3].

Transcutaneous electrical nerve stimulation (TENS) is a nonpharmacological method to manage both acute and persistent pain that works on the principle of gate theory. TENS uses electrical stimulation delivered through surface electrodes encased in a patch to activate large afferent fibers, which causes presynaptic inhibition of afferent nociceptive input at the dorsal horn of the spinal cord [3]. TENS also activates central

XSL•FO
RenderX

inhibitory pathways in the brain and spinal cord and acts through multiple neurotransmitters involving opioid, gamma-aminobutyric acid, serotonin, muscarinic, and cannabinoid receptors [4]. Analgesia following TENS can last beyond the period of application for up to 24 hours [5]. There is some evidence that the type of TENS used can impact the duration of analgesia, with higher frequency, lower intensity, or conventional TENS providing faster analgesia but less poststimulation analgesia and low frequency, high intensity, or acupuncture such as TENS providing longer poststimulation analgesia, which can last several hours after the stimulation is terminated [6]. Burst mode TENS, which delivers pulses of electrical stimulation in bursts, is thought to combine the effects of the abovementioned 2 modes and provide quicker onset analgesia that lasts longer than the duration of the stimulation [7]. TENS has been shown to improve functional disability in persistent low-back pain [8].

Music listening interventions, such as music therapy delivered by trained music therapists or listening to self-selected music, have been shown to reduce pain intensity, anxiety, and depression [9]. Functional neuroimaging indicates that this could be due to increasing the descending inhibition as well as changes to the salient stimuli detection circuitry involving somatosensory and prefrontal areas, amygdala, and cingulate cortex, as well as the insula. There is also evidence that reward pathways involving dopamine play a part in the analgesic and mood-altering effects of music listening interventions [10]. Cognitive mechanisms involved in analgesia and mood-altering effects of music listening interventions may include sustained attention on the music rather than pain, associations of self-selected music to pleasant memories, and agency related to selection of the type of music and the volume at which it is played [9]. Listening to self-selected music as well as engaging with the content of the music, or active listening, may reduce pain intensity and improve emotional distress more than music therapy [11]. Listening to self-selected music could have the potential advantage of improved self-efficacy and motivation, increased engagement in daily activities, and incorporation as a cost-effective noninvasive component to the overall pain management strategy [9].

There are no studies in the published literature that combine the abovementioned 2 nonpharmacological interventions in pain management. There is potential for the 2 interventions to act synergistically, providing superior pain relief and improved function compared with either intervention alone, through engagement of central pain modulation pathways and activation of descending inhibitory mechanisms.

Aims

The primary aim of this study is to test the feasibility of a clinical trial comparing TENS alone to TENS combined with listening to self-selected music in participants with low-back pain with or without leg pain lasting more than 3 months. The results of this feasibility study will be used to inform a larger trial of these interventions in participants with persistent neck and low-back pain lasting more than 3 months. A larger study will be considered feasible if 80% or more of the eligible population consents to the study and 80% or more of the participants are able to adhere to the protocol. The secondary aims are to explore if combining the 2 modalities has a synergistic effect on analgesia and function. A difference of 2 points in the numerical rating scale for pain intensity will be considered significant.

Hypotheses

First, we hypothesize that administration of TENS combined with music will be equally feasible compared to TENS alone and will be well tolerated by participants. Second, we hypothesize that TENS combined with listening to self-selected music will have improved analgesia compared to using TENS alone.

Methods

Overview

TENS with music study will be a prospective, randomized controlled pilot study to compare TENS alone to TENS combined with participant self-selected music. The participants will be randomized to receive 2 weeks of TENS therapy and 2 weeks of TENS with self-selected music using a streaming service, CD player, Long Play, or radio. The order in which this is delivered will be determined by randomization. Therefore, 1 group will receive standard TENS for 2 weeks, followed by TENS with self-selected music for 2 weeks. This order will be reversed in the other group with no washout period. The participants will be compared against themselves for the outcomes, thus reducing the number of participants to power the study. We are hoping to recruit 20 participants for this pilot trial (Figure 1). We will use TENS and TENS with music as part of a holistic pain management plan involving pain neuroscience education, behavior activation, and coping skills training that is individually tailored and not as a stand-alone intervention for pragmatic reasons.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) diagram: TENS with music pilot study. TENS: transcutaneous electrical nerve stimulation.

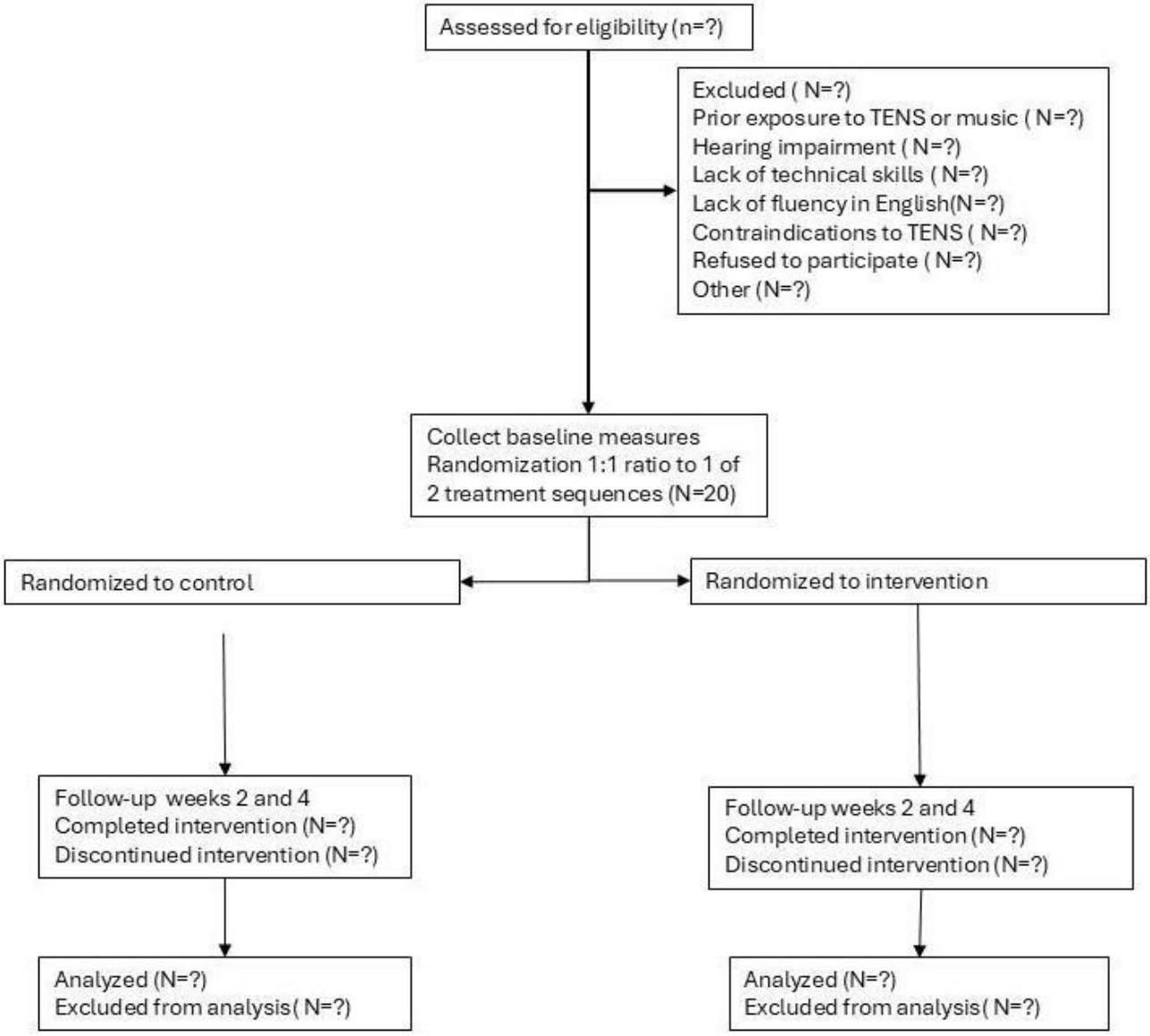

Eligibility

All participants aged 18 years or older, presenting with low-back pain lasting more than 3 months with or without referred or radicular limb pain, and attending a pain clinic at a metropolitan tertiary hospital in Australia will be eligible to participate in the study. Participants who have used TENS or music therapy in the past to manage their pain and those participants with a significant hearing impairment, the lack of technical ability to play music through a streaming platform or operate other music listening devices such as an LP or CD player, or an inability to communicate in English will be excluded from participation. Individuals with contraindications to the use of a TENS machine, such as having a permanent cardiac pacemaker, implantable cardioverter defibrillator, or neurostimulators, will also be excluded [12].

Treatment

TENS will be administered for 1 hour at a time, up to a maximum of 6 times every 24 hours. Participants are free to select from available modes, which include burst, continuous, or modulate. Other parameters, such as pulse width, pulse rate, and pulse intensity, can also be modified by the participant. The pulse intensity of the TENS machine can be adjusted from 0 to 80 mA, and the pulse frequency from 2 to 150 Hz. The aim is to use the frequencies, amplitude, and mode to provide a comfortable paresthesia [13]. The participants can vary the electrode placement to provide the best possible analgesia.

The TENS will be self-administered. Music from a self-selected playlist from a set of genres such as classical, easy listening, pop, rhythm and blues, vocal, and musical will be used. We will use a participant diary to record the duration of TENS each day and the type of music listened to. Music can be listened to with or without headphones. Participants will be provided with a 1-month Spotify (Spotify AB) subscription voucher to enable ad-free music listening.

Randomization

Participants will be randomized in a 1:1 ratio to 1 of 2 treatment sequences:

XSL•FO
RenderX

- Sequence AB: 2 weeks of standard TENS followed by 2 weeks of TENS combined with self-selected music
- Sequence BA: 2 weeks of TENS combined with self-selected music followed by 2 weeks of standard TENS

Randomization will be performed using a computer-generated allocation list created in Stata (StataCorp LLC) before study commencement. A permuted block design with variable block sizes will be used to ensure balance between sequences throughout recruitment. The randomization seed and code will be archived to allow reproducibility. Allocation concealment will be maintained by storing the randomization list in the Research Electronic Data Capture (REDCap; Vanderbilt University) randomization module, which will reveal the assigned sequence only after a participant has provided informed consent and completed baseline assessments.

The randomization result will be communicated to participants by telephone within 24 hours of their initial assessment. Investigators responsible for outcome collection and statistical analysis will remain blinded to treatment sequence. Participants cannot be blinded due to the pragmatic nature of the intervention.

Outcomes

The primary outcome will be the feasibility of the intervention, measured as adherence rate, dropout rate, and the percentage of the eligible population that consented. Adherence rates will include the number of participants using TENS alone or TENS with music regularly more than once per week, adherence to filling out the pain diaries and other patient-reported outcome measures, and attendance at follow-up clinic visits.

Secondary outcomes include the following:

- Pain intensity as a numerical rating scale (NRS) score collected before and after each session using a pain diary. The baseline pain will be collected for 1 week before the intervention when they initially attend the clinic.
- Opioid and other analgesic consumption over a 24-hour period collected in the pain diary
- Participant satisfaction using a 5-point Likert scale
- Mood measured through the 21-item Depression, Anxiety, and Stress Scale
- Function measured through the Patient-Specific Functional Scale [14]
- Compliance with other interventions, such as participation in interdisciplinary pain programs
- Adverse outcomes, including worsening pain and skin reaction to the electrodes used in TENS

All reported adverse reactions will also be collected through the study period when using the TENS machines through entries in the pain diary as well as during interviews with the participants. The self-reported outcome measures, such as the 21-item Depression, Anxiety, and Stress Scale and the Patient-Specific Functional Scale, will be recorded through REDCap forms during the clinic appointments. Figure 2 summarizes the process.

Figure 2. Workflow for transcutaneous electrical nerve stimulation with music pilot study. DASS-21: 21-item Depression, Anxiety, and Stress Scale; PSFS: Patient-Specific Functional Scale.

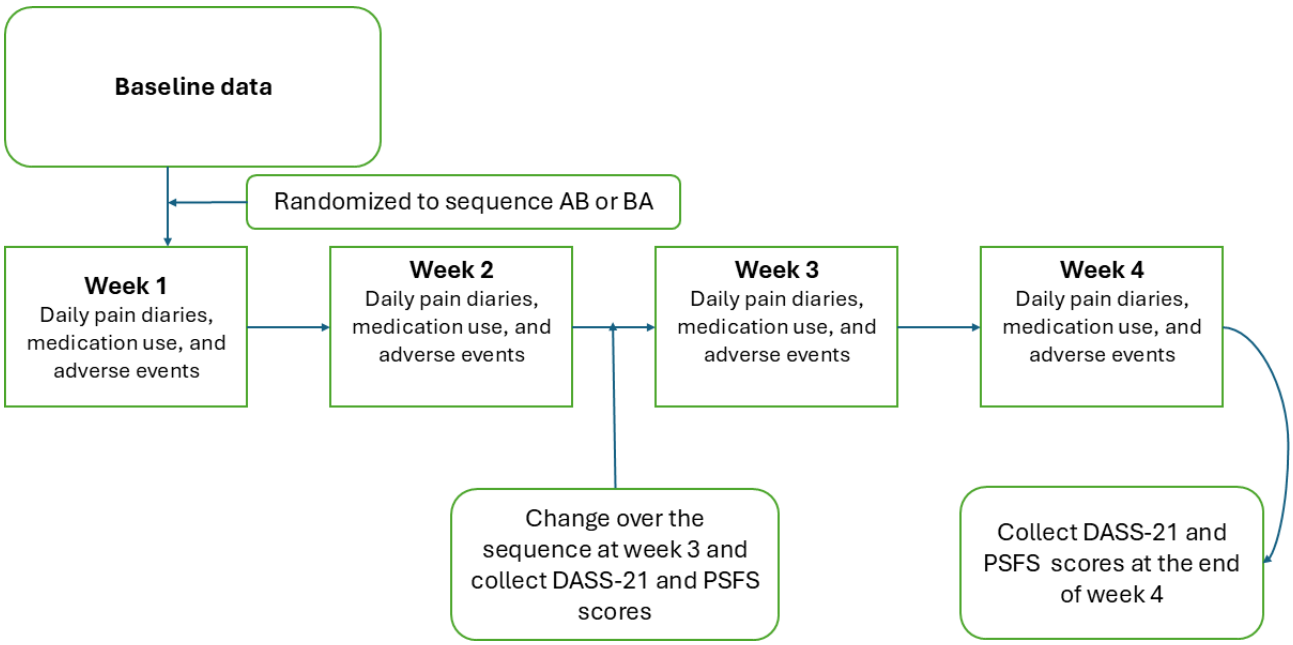

Blinding

The observer who collates the data and the statistician who analyzes the results will be blinded. The participants will not be blinded, as blinding is not pragmatic for this intervention. The equipment required includes a TENS machine, a pain diary, and the participants' own device for listening to music on Spotify, CD, LP, radio, or other music listening devices.

Data Collection

A standardized pain diary (Multimedia Appendix 1) will be used for collecting the primary outcome and some of the secondary outcomes, such as medication taken over 24 hours. Specific forms will be used for other secondary outcomes, which are available through REDCap. The data will be stored in a password-protected, secure database via REDCap. The data

XSL•FO
RenderX

collected will be deidentified and assigned a study code. The pain management unit will be the custodian of the data stored in the password-protected drive within the local health district computer, with access restricted to the investigators only. The data will be stored for 7 years after the completion of the study. Table 1 summarizes the data collection points.

Table . Data collection points.

Measure	Frequency	Collection points
Pain diaries	Daily	Baseline, week 1, week 2, week 3, and week 4
Average pain over 24 hours (numerical rating scale)	Weekly	Baseline, week 1, week 2, week 3, and week 4
Medications	Weekly	Baseline, week 1, week 2, week 3, and week 4
Patient-Specific Functional Scale	Fortnightly	Baseline, week 2, and week 4
21-item Depression, Anxiety, and Stress Scale	Fortnightly	Baseline, week 2, and week 4
Pain sites update	1 time	Baseline
Pain duration	1 time	Baseline
Demographics	1 time	Baseline

Power Calculation

Power was planned for the primary within-subject comparison (paired difference between TENS with self-selected music and TENS), with a 2-sided α=.05 and 80% power. In the absence of previous variance data, sensitivity analyses were conducted across standardized paired effect sizes (dz). The required numbers of completed pairs (participants who complete both study periods) are presented in Table 2.

Table . Sample size calculation.

Effect size (dz)	Completed pairs needed, n	Recruits with 10% dropout, n	Recruits with 20% dropouts, n	Recruits with 30% dropouts, n
0.4	52	57	62	68
0.5	34	37	41	44
0.6	24	26	29	31
0.7	19	21	23	25
0.8	15	17	18	20

As a pilot study (target recruitment N=20), the trial is adequately powered only for larger effects (approximately dz≥0.70). Therefore, the primary purpose of this pilot study is to assess the feasibility, estimate variance, and provide effect size estimates to inform the definitive sample size for the main trial.

All calculations were performed using G*Power version 3.1 (Heinrich Heine University), which is widely used for sample size and power analysis in clinical research.

Statistical Methods

This study is a 2-period, 2-sequence crossover pilot trial without a washout period between treatments. Analyses will primarily focus on feasibility outcomes (recruitment, adherence, data completeness, and dropout) using descriptive statistics (means, medians, and proportions) with 95% CIs rather than formal hypothesis testing. These feasibility metrics will inform the practicality of conducting a larger definitive trial.

For the primary clinical outcome of pain intensity, exploratory analyses will first be conducted using paired *t* tests to compare mean scores between TENS and TENS with self-selected music within participants. Normality of paired differences will be assessed using graphical methods and the Shapiro-Wilk test. If assumptions are violated, the Wilcoxon signed-rank test will be applied as a nonparametric alternative. Categorical variables (eg, adverse events and compliance categories) will be analyzed initially using chi-square tests, or Fisher exact test if expected cell counts are small. All tests will be 2-sided with a significance level of α=.05.

Following these preliminary analyses, formal modeling tailored to the crossover design will be undertaken. Linear mixed-effects models will be specified with fixed effects for treatment (TENS vs TENS with self-selected music), period, and sequence, and a random intercept for participant to account for within-participant correlations. Because no washout period is present, carryover effects will be explicitly assessed by including a prior-treatment term and by examining period×sequence interactions. If evidence of carryover is observed, analyses will be restricted to first-period data as a sensitivity check, treating it as a parallel-group comparison.

Secondary outcomes (mood, function, satisfaction, medication use, and adverse events) will be analyzed using appropriate mixed-effects regression models: linear for continuous outcomes, cumulative logit for ordinal outcomes, negative binomial for counts, and logistic for binary outcomes. These analyses are exploratory in nature. Given the pilot design and small sample size, no formal correction for multiple testing will

XSL•FO
RenderX

be applied; instead, effect sizes and CIs will be reported to guide future sample size calculations.

Intention-to-treat principles will be followed, analyzing participants according to their randomized sequence regardless of adherence or protocol deviations. Mixed-effects models allow inclusion of all available data, including partial data from participants who drop out after the first period, thereby minimizing bias. Per-protocol analyses will be conducted as sensitivity checks. Missing data will be handled descriptively, with analyses restricted to available cases. No multiple imputation will be attempted due to the small sample size.

As a feasibility study with 20 participants, the trial is powered only to detect large effects (approximately $dz \geq 0.70$). The observed variance and effect sizes will be used to refine sample size calculations for the definitive trial.

All analyses will be conducted using Stata and SPSS Statistics (IBM Corp).

Ethical Considerations

The trial has been approved by the local health district ethics committee (2024/PID02885 - 2024/ETH02492) and is currently screening participants for recruitment. Written informed consent will be obtained from all participants. The participants will be informed that a withdrawal of consent is possible at any point and that their clinical care will not be compromised. The participants will not receive any remuneration.

Results

This manuscript describes the protocol for an ongoing pilot trial, with results to be reported upon trial completion. Recruitment for the trial began in August 2025. As of manuscript submission, 4 participants have been enrolled, with a target of 20 participants by the end of August 2026. Data analysis is expected to be completed by December 2026. The results are planned to be published as an open-access article in 2027 and used to inform the design of a larger trial.

A summary of the results will be disseminated to the participants.

Discussion

Anticipated Findings

This pilot study primarily aims to assess the feasibility of combining TENS with self-selected music. The anticipated results include approximately 80% adherence to daily use of TENS or TENS combined with self-selected music and completion of daily diaries on 80% of the study days. An adherence rate of 80% will be considered sufficient for the current data collection methods to be deemed feasible for a larger trial. It is hoped that adherence to data collection procedures, including real-time documentation of pain scores and medications used, will be evaluated among participants [15]. The trial is not likely to reveal any additive analgesic benefit for TENS and self-selected music due to the small participant numbers, but it is anticipated that a trend toward improved pain scores, better function, and reduced analgesia requirements by combining the 2 modalities may be revealed. A change of 2 points in the 11-point scale on the NRS will be most clinically relevant, but even a point difference in the NRS can be considered as the minimal clinically significant change to the participants [16]. A change in pain intensity of 2 points in the NRS is accepted as best correlated with general patient satisfaction and other meaningful outcome measures in chronic low-back pain [17]. The clinically meaningful change will differ by baseline pain; however, for pragmatic reasons in this feasibility trial, it was decided to not stratify participants based on baseline pain and adopt the general threshold. It is hoped that the acceptability of combining TENS with self-selected music as an intervention for pain management will be determined, and whether the outcome measurements and methods are perceived as burdensome and time-consuming will be assessed in this small sample. Additional insights are expected regarding noncompliance with the treatment, which may include lack of perceived improvement in pain or adverse events associated with use, such as worsening pain or skin irritation from the electrodes, as well as other factors such as lack of motivation [18].

Chronic pain is a complex condition, and the intensity of pain and the functional limitations experienced by people with chronic pain are best managed using a biopsychosocial framework [19]. The use of TENS and self-selected music will fit within this framework, along with education and an interdisciplinary pain self-management program, to achieve person-centered functional goals by addressing the multiple contributors to a person's pain experience [20]. The 2-week time frame for the intervention, rather than an in-clinic comparison, will provide a more pragmatic approach to test the real-world efficacy of combining the interventions [21]. Focusing not only on improvement of pain scores but also on other aspects of pain management, including mood, function, and medication use, will provide a comprehensive picture of the interaction between the 2 modalities.

The self-selection of the pulse amplitude, frequency, and mode is recommended due to the complexity of the pain experience. TENS provides optimal analgesia when the delivered sensation is comfortable for the person experiencing pain [22]. Self-selection of the TENS parameters and music also has the potential to increase the sense of control of the participants, thus improving self-efficacy, which can also reduce pain intensity and improve compliance [23]. Due to constraints related to resources and time, a smaller number of participants were selected, as this is a feasibility trial. A crossover trial design was chosen where participants were their own controls to reduce the sample size [24]. We chose not to have a washout period due to the short-acting nature of the intervention, with evidence that postuse analgesia usually wears off by 24 hours for TENS [22].

The duration of analgesia from music listening interventions in chronic pain is unclear, but available evidence indicates that it is short-lived and does not last hours or days beyond the session [25]. The clinical trials using TENS do not consistently report adherence rates and can vary from 50% to 80% [18]. We chose the higher adherence rate, as this is a feasibility study. Those with previous exposure to TENS or music listening interventions

XSL•FO
RenderX

were excluded due to the potential for placebo effects arising from previous experience [26]. Stratification of participants based on their previous use of TENS or music listening interventions could have reduced this bias in recruitment. However, given the small sample size and the pilot feasibility nature of this study, stratification would have introduced unnecessary complexity and was not aligned with the primary objective, which was to assess the feasibility of data collection methods [27].

Limitations

The study requires the participants to fill out a significant amount of information daily, which may contribute to participant burden. It also increases the risk of noncompliance with regular entries and introduces recall bias if participants try to record the information before review. One of the main limitations of the study will be the lack of concurrent electronic monitoring to complement the participant-reported data. The study methodology also does not allow for blinding of the participants due to the crossover design. We have excluded participants who cannot communicate in English, those who cannot use the TENS machine due to functional limitations, and those who are unable to use streaming technology or other technology that plays music for pragmatic reasons in this feasibility trial. The research team acknowledges that this will affect the generalizability of the results by excluding culturally diverse cohorts and those who are less tech savvy. However, such individuals are expected to represent only a small proportion of otherwise eligible participants. This point will be addressed if a larger trial is considered feasible. Participants have considerable flexibility in choosing the frequency, mode, and electrode placement when using TENS, as well as the type of music they listen to. While this variability makes it challenging to determine which parameters were truly effective, predict underlying mechanisms, and replicate results, these limitations are acknowledged. As this study primarily focuses on feasibility, such factors will be carefully addressed when designing a larger trial. The pilot study is designed to detect only large effect sizes and to inform power calculations for future trials, and it may fail to detect small effect sizes. It is anticipated that approximately 90% of participants will need to demonstrate a positive trend, defined as at least a 1-point improvement on the NRS, with the combination of TENS and self-selected music, to calculate the participant numbers for a larger trial that is adequately powered to detect a moderate effect size [28]. However, the authors anticipate that the effect sizes and the CIs from the pilot study will help to calculate a conservative estimate to detect smaller effect sizes from a larger definitive trial [29].

Future Directions

If the trial is feasible and indicates a positive synergistic trend with TENS and self-selected music, it will help to plan for a larger trial with more varied pain diagnoses, including acute and chronic pain other than spinal pain. The aim of a larger study will be to show if combining the 2 modalities of music and TENS will improve pain intensity and function more than TENS alone. The pilot will test the feasibility of using pain diaries as a reliable indicator of gathering information in a larger trial. Future directions could include automated methods for data collection, which could provide more reliable indicators of adherence, function, and medication use.

Acknowledgments

The authors would like to thank associate professors Stefan Dieleman and Ms Louise Cope for reviewing the protocol.

Funding

No external funding was received for this trial.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

Conceptualization: SR, MG
Formal analysis: SN, LT
Investigation: SR, MG
Methodology: SR, MG
Project administration: SR, MG, AF
Resources: SR, MG, AF
Software: SN, LT, AF, AMR
Supervision: SR
Validation: SR, MG
Visualization: SR, SN, LT
Writing – original draft: SR, SN, LT, MG
Writing – review & editing: SR, SN, LT, MG, AF, AMR

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data collection template.
[DOCX File, 30 KB - resprot_v15i1e82382_app1.docx]

References

1. Ferreira ML, de Luca K, Haile L, et al. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol 2023 May 22;5(6):e316-e329. [doi: 10.1016/S2665-9913(23)00098-X] [Medline: 37273833]
2. Thomas KH, Dalili MN, Cheng HY, et al. Prevalence of problematic pharmaceutical opioid use in patients with chronic non-cancer pain: a systematic review and meta-analysis. Addiction 2024 Nov;119(11):1904-1922. [doi: 10.1111/add.16616] [Medline: 39111346]
3. Shi Y, Wu W. Multimodal non-invasive non-pharmacological therapies for chronic pain: mechanisms and progress. BMC Med 2023 Sep 29;21(1):372. [doi: 10.1186/s12916-023-03076-2] [Medline: 37775758]
4. Choi JC, Kim J, Kang E, et al. Brain mechanisms of pain relief by transcutaneous electrical nerve stimulation: a functional magnetic resonance imaging study. Eur J Pain 2016 Jan;20(1):92-105. [doi: 10.1002/ejp.696] [Medline: 25847717]
5. Ezema CI, Onyeso OK, Nna EO, et al. Transcutaneous electrical nerve stimulation effects on pain-intensity and endogenous opioids levels among chronic low-back pain patients: a randomised controlled trial. J Back Musculoskelet Rehabil 2022;35(5):1053-1064. [doi: 10.3233/BMR-210146] [Medline: 35253730]
6. Vance CG, Dailey DL, Chimenti RL, Van Gorp BJ, Crofford LJ, Sluka KA. Using TENS for pain control: update on the state of the evidence. Medicina (Kaunas) 2022 Sep 22;58(10):1332. [doi: 10.3390/medicina58101332] [Medline: 36295493]
7. Patel P, Green M, Tram J, et al. Latest advancements in transcutaneous electrical nerve stimulation (TENS) and electronic muscle stimulation (EMS): revisiting an established therapy with new possibilities. J Pain Res 2025 Jan 9;18:137-153. [doi: 10.2147/JPR.S493162] [Medline: 39816205]
8. Wu LC, Weng PW, Chen CH, Huang YY, Tsuang YH, Chiang CJ. Literature review and meta-analysis of transcutaneous electrical nerve stimulation in treating chronic back pain. Reg Anesth Pain Med 2018 May;43(4):425-433. [doi: 10.1097/AAP.0000000000000740] [Medline: 29394211]
9. Howlin C, Rooney B. The cognitive mechanisms in music listening interventions for pain: a scoping review. J Music Ther 2020 May 2;57(2):127-167. [doi: 10.1093/jmt/thaa003] [Medline: 32249313]
10. Sihvonen AJ, Pitkäniemi A, Särkämö T, Soinila S. Isn't there room for music in chronic pain management? J Pain 2022 Jul;23(7):1143-1150. [doi: 10.1016/j.jpain.2022.01.003] [Medline: 35124251]
11. Lee JH. The effects of music on pain: a meta-analysis. J Music Ther 2016;53(4):430-477. [doi: 10.1093/jmt/thw012] [Medline: 27760797]
12. Guidance for the clinical use of electrophysical agents. : Chartered Society of Physiotherapy; 2006 URL: https://cdm16198.contentdm.oclc.org/digital/collection/p16198coll2/id/767/ [accessed 2026-01-28]
13. Pallett EJ, Rentowl P, Johnson MI, Watson PJ. Implementation fidelity of self-administered transcutaneous electrical nerve stimulation (TENS) in patients with chronic back pain: an observational study. Clin J Pain 2014 Mar;30(3):224-231. [doi: 10.1097/AJP.0b013e31828dc828] [Medline: 24503978]
14. Horn KK, Jennings S, Richardson G, Vliet DV, Hefford C, Abbott JH. The patient-specific functional scale: psychometrics, clinimetrics, and application as a clinical outcome measure. J Orthop Sports Phys Ther 2012 Jan;42(1):30-42. [doi: 10.2519/jospt.2012.3727] [Medline: 22031594]
15. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022 Jan 1;60(1):95-103. [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
16. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004 Aug;8(4):283-291. [doi: 10.1016/j.ejpain.2003.09.004] [Medline: 15207508]
17. Suzuki H, Aono S, Inoue S, et al. Clinically significant changes in pain along the Pain Intensity Numerical Rating Scale in patients with chronic low back pain. PLoS ONE 2020 Mar 3;15(3):e0229228. [doi: 10.1371/journal.pone.0229228] [Medline: 32126108]
18. Johnson MI, Paley CA, Jones G, Mulvey MR, Wittkopf PG. Efficacy and safety of transcutaneous electrical nerve stimulation (TENS) for acute and chronic pain in adults: a systematic review and meta-analysis of 381 studies (the meta-TENS study). BMJ Open 2022 Feb 10;12(2):e051073. [doi: 10.1136/bmjopen-2021-051073] [Medline: 35144946]
19. Nielson WR, Weir R. Biopsychosocial approaches to the treatment of chronic pain. Clin J Pain 2001 Dec;17(4 Suppl):S114-S127. [doi: 10.1097/00002508-200112001-00020] [Medline: 11783824]
20. Nicholas MK. The biopsychosocial model of pain 40 years on: time for a reappraisal? Pain 2022 Nov 1;163(Suppl 1):S3-S14. [doi: 10.1097/j.pain.0000000000002654] [Medline: 36252231]

XSL•FO
RenderX

21. Johnson MI. Resolving long-standing uncertainty about the clinical efficacy of transcutaneous electrical nerve stimulation (TENS) to relieve pain: a comprehensive review of factors influencing outcome. Medicina (Kaunas) 2021 Apr 14;57(4):378. [doi: 10.3390/medicina57040378] [Medline: 33919821]
22. Paley CA, Wittkopf PG, Jones G, Johnson MI. Does TENS reduce the intensity of acute and chronic pain? A comprehensive appraisal of the characteristics and outcomes of 169 reviews and 49 meta-analyses. Medicina (Kaunas) 2021 Oct 4;57(10):1060. [doi: 10.3390/medicina57101060] [Medline: 34684097]
23. Köke AJ, Schouten JS, Lamerichs-Geelen MJ, et al. Pain reducing effect of three types of transcutaneous electrical nerve stimulation in patients with chronic pain: a randomized crossover trial. Pain 2004 Mar;108(1-2):36-42. [doi: 10.1016/j.pain.2003.10.013] [Medline: 15109505]
24. Louis TA, Philip WL, Bailar JC, Polansky M. Crossover and self-controlled designs in clinical research. In: Bailar JC, Hoaglin DC, editors. Medical Uses of Statistics, 3rd edition: CRC Press; 2009:83-103.
25. Chen S, Yuan Q, Wang C, Ye J, Yang L. The effect of music therapy for patients with chronic pain: systematic review and meta-analysis. BMC Psychol 2025 Apr 30;13(1):455. [doi: 10.1186/s40359-025-02643-x] [Medline: 40307940]
26. Colloca L, Akintola T, Haycock NR, et al. Prior therapeutic experiences, not expectation ratings, predict placebo effects: an experimental study in chronic pain and healthy participants. Psychother Psychosom 2020;89(6):371-378. [doi: 10.1159/000507400] [Medline: 32492688]
27. Ahmed SK. How to choose a sampling technique and determine sample size for research: a simplified guide for researchers. Oral Oncology Reports 2024 Dec;12:100662. [doi: 10.1016/j.oor.2024.100662]
28. Sim J. Should treatment effects be estimated in pilot and feasibility studies? Pilot Feasibility Stud 2019 Aug 28;5:107. [doi: 10.1186/s40814-019-0493-7] [Medline: 31485336]
29. Ying X, Freedland KE, Powell LH, Stuart EA, Ehrhardt S, Mayo-Wilson E. Determining sample size for pilot trials: a tutorial. BMJ 2025 Aug 8;390:e083405. [doi: 10.1136/bmj-2024-083405] [Medline: 40780848]

Abbreviations

NRS: numerical rating scale
REDCap: Research Electronic Data Capture
TENS: transcutaneous electrical nerve stimulation

Edited by J Sarvestan; submitted 15.Aug.2025; peer-reviewed by E Güleç; revised version received 02.Jan.2026; accepted 05.Jan.2026; published 24.Feb.2026.

Please cite as:
Radhakrishnan S, Glynn M, Robertson AM, Fedyczkowski A, Trinh L, Naz S
Transcutaneous Electrical Nerve Stimulation vs Transcutaneous Electrical Nerve Stimulation With Music in Persistent Low-Back Pain: Protocol for a Pilot Feasibility Trial
JMIR Res Protoc 2026;15:e82382
URL: https://www.researchprotocols.org/2026/1/e82382
doi:10.2196/82382

© Seema Radhakrishnan, Michael Glynn, Anne-Marie Robertson, Anna Fedyczkowski, Lieu Trinh, Sabrina Naz. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Adaptation of a Protocol to Evaluate Online Recruitment Strategies for Mailing HIV Self-Tests and Pilot an Intervention to Improve Linkage to HIV Prevention and Care Among Transgender Women: Protocol for a Randomized Controlled Trial

Joanna A Caldwell[1], MPH; Ruth Dana[1], MPH; Irah L Lucas[1], MPH; Patrick Sean Sullivan[1], DVM, PhD; Evelyn Olansky[2], MPH; Stephen Sullivan[3], MPH; Marissa Hannah[1], MPH; Travis Sanchez[1], MPH, DVM; Lisa Hightow-Weidman[1], MPH, MD; Akshay Sharma[4], MBBS, MPH, PhD

[1]Rollins School of Public Health, Department of Epidemiology, Emory University, Address 1518 Clifton Rd NE, Atlanta, GA, United States

[2]a Subsidiary of DLH Corp, Social and Scientific Systems, Inc, Atlanta, GA, United States

[3]Medical School, University of Michigan – Ann Arbor, 1109 Geddes Ave, Ann Arbor, MI, United States

[4]Department of Health Behavior and Clinical Sciences, University of Michigan, School of Nursing, 1109 Geddes Ave, Ann Arbor, MI, United States

Corresponding Author:
Joanna A Caldwell, MPH
Rollins School of Public Health, Department of Epidemiology, Emory University, Address 1518 Clifton Rd NE, Atlanta, GA, United States

Abstract

Background: Transgender women are disproportionately affected by HIV in the United States. Evidence shows that HIV self-testing increases awareness of HIV status, preventing transmission. Culturally appropriate strategies are needed to provide HIV self-tests (HIVSTs) to transgender women.

Objective: This paper aims to describe an adaptation for transgender women of a comparative effectiveness trial of an intervention (Know@Home website) on the use of HIVSTs, counseling, and linkage of participants to appropriate services initially designed to reach Black or African American and Hispanic or Latino men who have sex with men (Implementation of Rapid HIV Self-Testing Among MSM Project).

Methods: The ADAPT-ITT (Assessment, Decision, Adaptation, Production, Topical Experts, Integration, Training, and Testing) framework guided the adaptation. Project staff worked with consultants and a community advisory board to revise all study materials (ie, intervention web content, marketing text and images, and language used in the surveys) to be tailored to, and inclusive of, health topics relevant to transgender women. We conducted a pilot study with transgender women recruited through online venues. Participants were randomly assigned 2:1 to the intervention arm or control arm. All participants were mailed 2 HIVSTs, the OraQuick In-Home HIV Test, and had access to a standard resource list including information on how to find a nearby testing location, HIV prevention, and sexual health. Only intervention arm participants had access to the study website, "Know@Home.org." The Know@Home website allowed participants in the intervention arm to order sexually transmitted infection test kits, condoms, and lubricant and provided HIV prevention information, service locators, and risk assessment tools. Adaptations to the intervention included modifying language, images, and links to outside resources to be transgender specific or gender neutral. Online surveys were adapted by removing inappropriate gender terms and replacing them with culturally appropriate terms for gender identity, genitalia, and condom use. Four months after completing the baseline (enrollment) survey, all participants were asked to complete a final follow-up survey. Upon completing the 4-month survey, participants were mailed a dried blood spot collection kit to be returned by mail for the laboratory testing of the sample. The results of all HIVSTs could be reported in real time in an online survey. Video health counseling was available, upon request, for all participants during study participation.

Results: Enrollment into the pilot study began on April 22, 2021, and concluded on July 7, 2021, yielding a total of 102 transgender women participants, with data collection completed in January 2022.

Conclusions: By using the ADAPT-ITT model, we produced materials to be culturally and linguistically appropriate for transgender women. The findings of this study have the potential to inform future research studies among transgender women and underscore the importance of involving subject matter experts and community members in the development of tailored materials.

Trial Registration: ClinicalTrials.gov NCT04219878; https://clinicaltrials.gov/study/NCT04219878

XSL•FO
RenderX

International Registered Report Identifier (IRRID): PRR1-10.2196/75345

(JMIR Res Protoc 2026;15:e75345) doi:10.2196/75345

KEYWORDS

HIV; HIV prevention; transgender women; linkage to care; HIV self-testing; protocol

Introduction

The Ending the HIV Epidemic in the United States initiative is built on 4 science-based strategies. The first is diagnosing all individuals with HIV as early as possible. HIV testing allows people to learn their HIV status and initiate pre-exposure prophylaxis (PrEP) or antiretroviral treatment, preventing new infections. HIV self-testing (HIVST) is one component of this testing strategy [1-3].

In 2019, approximately 2% of HIV infection diagnoses were among transgender persons. The majority (93%) of these were among transgender women [4]. Transgender women in the United States are disproportionately affected by HIV with a prevalence of 14.1%, compared with an estimate of less than 0.5% for all US adults [5-7].

The underutilization of HIV testing may contribute to the disproportionate impact of HIV among transgender women [7], as HIV testing is the entry to engagement with HIV prevention and treatment [8]. A study in the United Kingdom found that transgender people are less likely to have ever tested for HIV than cisgender people (49% vs 64%, respectively) [9]. Barriers to HIV testing reported by transgender women include low risk perception, lack of access to HIV testing, and concerns of HIV-related stigma and transphobia in health care settings [10-12]. Providing HIVSTs to transgender women may reduce some barriers they experience when seeking gender-affirming HIV testing services by expanding routine testing options and partially mitigating the barrier of medical mistrust [10,13-15]. Self-testing allows users the flexibility of not having to present to clinical or community-based testing sites where providers may not be culturally competent. This testing approach provides an alternative to those who are distrustful of medical settings, who want to avoid settings that could trigger gender dysphoria, or who do not wish to be seen attending testing sites, due to stigma, discrimination, homophobia, transphobia, and racism [16,17].

Many studies conflate cisgender men who have sex with men (MSM) with transgender women, and the studies that disaggregate data on transgender women are limited [18,19]. Tailoring existing interventions is an important strategy to expedite the creation and deployment of effective interventions and improve public health outcomes for targeted populations [18,20,21]. Preliminary evidence indicates that technology-based behavioral health interventions, such as using social media, have the potential to promote HIV testing and decrease prevention disparities [22]. Social media–based interventions can increase engagement in care [23], and computerized counseling and text-based interventions can improve viral suppression among transgender women [24-27], although some studies did not show statistically significant results, and more research is needed [28].

This paper presents the adaptation of the study protocol and materials from the Implementation of Rapid HIV Self-Testing Among MSM Project (iSTAMP) [29] for use with transgender women. The iSTAMP study evaluated strategies to reach participants for distributing HIVST and incorporated a pilot evaluation of an HIV prevention phone app to improve linkage to HIV prevention and care designed for Black and Hispanic MSM [30]. Mailing HIVSTs to gay, bisexual, and other MSM increased rates of testing and awareness of infection status. An adaptation of the original protocol for transgender women was necessary to consider social and behavioral characteristics unique to this population.

Through an existing cooperative agreement with the Centers for Disease Control and Prevention, we obtained funds for conducting pilot activities for transgender women. This pilot project was funded separately from the original iSTAMP study at the end of year 3 of the 4-year project period.

The study aimed to evaluate (1) the comparative effectiveness of the intervention (Know@Home website) on the use of HIVSTs, counseling, and linkage of participants to appropriate services (ie, HIV treatment, HIV PrEP, sexually transmitted infection [STI] testing, additional prevention, and social services) and (2) the cost of implementing recruitment strategies and the intervention. We describe a community-engaged process to adapt the iSTAMP protocol to be culturally appropriate for transgender women, which included adapting recruitment materials, survey tools, and the interactive intervention website.

Methods

Overview

This pilot study assessed whether the iSTAMP protocol, online advertisements, and recruitment materials [30], originally designed for recruiting and engaging Black and Hispanic MSM, could be adapted for and implemented for transgender women. Most study procedures were the same as those described in the iSTAMP protocol paper [30]. Briefly, the iSTAMP randomized control trials (both the original study and this adaptation) were designed to evaluate a program that distributed HIVST kits. The iSTAMP included a nested randomized component assessing the utility of comprehensive prevention content (Know@Home [31]) to support related HIV prevention services for those assigned only to the intervention arm. The adaptation substituted the Know@Home mobile app for a website named Know@Home with similar materials tailored to transgender women. The intervention content included HIV prevention information; a way to order condoms and STI self-collection kits; risk assessment tools; and locators for HIV testing, HIV treatment, and PrEP services. Four months after completing the

XSL•FO
RenderX

baseline (enrollment) survey, all participants were asked to complete a final follow-up survey, which collected information on sexual behavior, HIV testing and status, and linkage to care. Upon completing the 4-month survey, participants were mailed a dried blood spot collection kit to be returned by mail for laboratory testing. The results of all HIVSTs could be reported in real time into an online survey. Video health counseling was available, upon request, for all participants during study participation. Further details about study design, participant flow, data collection, and planned analyses can be found in the original iSTAMP protocol paper [30].

This paper focuses on materials affecting the transgender participant, rather than internal staff documents, such as standard operating procedures and email, text, or phone scripts. Electronic case report forms captured enrollment and retention, HIVST result follow-up communications, referrals, and video counseling.

Our adaptation was guided by the 8 phases of the ADAPT-ITT (Assessment, Decision, Adaptation, Production, Topical Experts, Integration, Training, and Testing) model [32] (Table 1). We chose the iSTAMP study for assessment, reviewed literature, identified domains of the study for adaptation, and developed a list of needed adaptations. Topical experts provided input on marketing and recruitment materials and recommended changes to intervention materials and survey assessments. Suggested changes were integrated into all study materials. The testing phase of the pilot study and associated data analyses occurred as the final phase.

Table . Use of ADAPT-ITT (Assessment, Decision, Adaptation, Production, Topical Experts, Integration, Training, and Testing) framework for the development of a protocol for distribution of HIV self-test kits for transgender women in the United States (2020).

Phase	Staff roles	Activity	Outcomes
Assessment	• Researchers	• Literature review	• Determined need for increased HIV testing options
Decision	• Researchers	• Literature review	• Selected the iSTAMP[a] protocol
Adaptation	• Community consultants	• Review of iSTAMP materials	• Identified questions, images, and content for revision
Production	• Researchers	• Review of input from community consultants	• Created required changes in recruitment materials, images, intervention materials, and survey assessments
Topical experts	• Subject matter experts • Community advisory board	• Recommend revisions • Formative discussion to provide feedback	• Provided feedback on intervention materials and survey assessment • Provided feedback on marketing, images, and recruitment materials
Integration	• Research staff	• Create new website and modify study assessments	• Revised marketing materials, recruitment materials, intervention materials, and assessment
Training or testing	• Researchers, research staff	• Implement pilot study	• Analyzed preliminary data on feasibility, acceptability, recruitment, retention, and HIV testing outcomes

[a]iSTAMP: Implementation of Rapid HIV Self-Testing Among MSM Project.

Assessment

In the assessment phase, we reviewed relevant literature and decided to adapt an intervention to promote HIV self-testing to increase HIV testing and awareness of HIV status for transgender women. An informal literature review indicated that transgender women were disproportionately affected by HIV [6,7,33]. Transgender women's unique experiences and their desire for HIV interventions and services necessitated specialized research to assess risks that differentiate transgender women from MSM [33,34].

The literature review indicated that additional behavioral health survey items, which are not typically included in HIV research with MSM, for example, gender identity and hormone replacement therapy, would be relevant for transgender women. These items were chosen and included in the survey tool. Research staff determined their capacity to adapt and implement the HIV counseling and referral component of the study and assessed the availability of transgender-specific referral resources for HIV testing and prevention services.

XSL•FO
RenderX

Decision and Adaptation

In the decision phase, we selected the iSTAMP study because the distribution of HIV self-test kits has shown to increase HIV testing in MSM [35], and there were specifically designed materials and processes available that could be adapted.

In the adaptation phase, although many aspects of the original iSTAMP study protocol remained the same (eg, OraQuick In-Home HIV test, the recruitment, and mailing systems), several domains required revision, including the recruitment materials (images and text), survey questions, and an intervention website. In March 2020, 5 members of the community and several culturally competent iSTAMP study staff identified language in existing survey materials that was not culturally appropriate for transgender women. They provided alternative appropriate language. A US $50 gift card was provided for the community members' time.

Production

During the production phase, staff reviewed the relevant literature from the assessment phase and created a list of required changes. This was a preliminary list of potential adaptations required to tailor the iSTAMP study for transgender women.

Topical Experts: Recruitment and Advertising Materials

Study staff convened an online meeting of consultants to identify effective recruitment strategies and tailor advertising materials specifically for transgender women. Consultants were either transgender women, transfeminine-identifying individuals, or nonbinary-identifying individuals. We recruited the consultants through social media and flyers sent to lesbian, gay, bisexual, transgender, and queer organizations located in the 11 states from which we planned to recruit participants. Each consultant was provided a US $50 e-gift card for their time. Four consultants participated in a recorded, online videoconference held in November 2020.

The consultants discussed the suitability of recruiting transgender women from online dating websites and social media platforms and explored potential platforms that may have had a greater transgender women user base than those already under consideration, such as TikTok. The facilitator presented the examples of advertising images and text messages to assess cultural acceptability.

Consultants identified that the pop cultural references used in the original intervention were more relevant to MSM and the drag community than to transgender women. The examples included language such as "Demon time" that references horniness and "Faster than beating your mug" that references applying makeup. The use of drag references with transgender women was deemed by the consultants to be offensive and inappropriate.

The language that alluded to using sex-seeking apps (eg, "You host, we're mobile" appearing to reference the dating app Grindr, or "Plan for your fans" appearing to reference the website OnlyFans, which is often used by sex workers) was viewed negatively by consultants. The consultants stated that the transgender women's community has often been portrayed as hypersexualized within mainstream media and that these text messages and images reinforced the stereotype. Advertisement text that focused on respecting transgender identities and intimated themes of empowerment was preferred.

Consultants emphasized the importance of using photographs with transgender women. Images where the individual's transgender or gender nonconforming identity was more prominent were rated higher. Consultants stated that text specifically used to provide assurances of data security or confidentiality (eg, "Your identity is safe here") increased the skepticism of data protections and how participant information may be used. Another identity-related concern the consultants discussed was that the study should consider developing procedures to allow participants to use different names in different contexts. For example, some participants may prefer to use their legal name or "deadname" at home because they are not "out." They reported that participants might want to use a chosen name on email and text communications but use their legal name or "deadname" on postal packages to avoid disclosing their transgender identity to the people with whom they live.

There was a consensus among consultants that obtaining free HIVSTs and STI home collection kits would be viewed as beneficial by transgender women. Consultants also agreed that the financial token of appreciation for the participant's time should be included in the advertisement because it may be viewed by many potential participants as immediately beneficial. Figures 1 and 2 provide examples of recruitment advertisements that resulted in greater response from potential participants during the study. All advertisements were created using licensed stock photos.

XSL•FO
RenderX

Figure 1. High-performing ad copy (sample 1).



Figure 2. High-performing ad copy (sample 2).



Topical Experts: Intervention Website

In this pilot study, we substituted the Know@Home mobile app, provided for the MSM intervention arm of the original iSTAMP trial, for a website named Know@Home, which included the same materials adapted for transgender women. The website contained HIV prevention information and allowed participants to order free materials (eg, condoms and STI self-collection kits) and access online risk assessment tools and service locators for HIV testing, HIV treatment, and PrEP services.

A subject matter expert (SME) on transgender health reviewed the Know@Home website intervention content and supporting materials. The expert proposed revisions to make the content suitable for transgender women based on recommendations from the Provincial Health Services Authority of British Columbia [36], which was a reliable source for recommendations at the time of study design. The reviewed intervention sections included written and video messaging components, risk-assessment quizzes to help assess one's risk for HIV infection, embedded links to external pages, and the static informational pages consisting of the "About Us" page and STI and HIV information.

Materials were revised with the gender-neutral terms "person" and "people." Materials that directly referred to sexual anatomy were revised to use gender-neutral anatomical terms [36]. Video

XSL•FO
RenderX

messages and external links to study materials were reviewed, and those appropriate for MSM were replaced with gender-neutral alternatives. Content that featured only MSM was replaced with content that featured transgender women or images that would be appropriate for all audiences regardless of gender identity. Table 2 lists the 4 core themes identified as requiring revisions and examples of appropriate terminology replacements. The four core themes included: (1) gendered terminology referring to the participant, (2) gendered terminology referring to the participant's sex partners, (3) terminology describing genitalia and sexual acts, and (4) condom use.

Table . Thematic content from the Implementation of Rapid HIV Self-Testing Among MSM Project (iSTAMP) study identified as not suitable for HIV self-test distribution for transgender women (2020).

Theme identified	iSTAMP terms and phrases	Revised terms and phrases	Reasons text modified	Study materials updated
Participant gender identity	• Men, man, MSM[a], guys • Gay and bisexual men	• People, person • Transfeminine people, transgender women	• Avoids misgendering transgender women	• Website • Surveys • Advertisements • Videos • Images
Sex partner or partners gender identity	• Men, man, MSM, guys	• Partner, partners, sex partner(s)	• Avoids potentially misgendering a participant's sex partner or partners • Avoids assumption that a participant has sex with only cisgender men	• Website • Surveys
Terms describing genitalia and sex acts	• Penis, scrotum, testicles • Penis, vulva • Vagina • Vaginal fluids, semen	• Outer parts • External pelvic area • Internal pelvic area • Genital fluids	• Avoids biological terms associated with a specific gender that may cause discomfort or gender dysphoria for a participant or their partner or partners • Avoids biological terms that may not be the preferred term of the participant or participant's partner or partners	• Website • Surveys

[a]MSM: men who have sex with men.

Topical Experts: Survey Assessments

An SME on transgender health reviewed the online surveys and identified items that needed to be rephrased with culturally appropriate terminology based on current best practices [37]. The SME identified and recommended survey domains that were relevant to transgender health that were not included in the original surveys designed for MSM.

Similar to the intervention materials, baseline and follow-up surveys were revised to ensure clarity and appropriateness of questions pertaining to the gender, sex, and genitalia of respondents and their sexual partner or partners. Questions were removed if they fell within the 4 core themes identified as inappropriate for transgender women. Supplemental questions related to the experiences of transgender women were developed using appropriate terminology for transgender women. These questions were provided to participants directly after they completed the baseline and follow-up surveys.

The supplemental questions obtained data on the following domains: gender identity, sexual orientation, gender presentation, mental health and suicidality, sexual behavior (using affirming language for genitalia), gender confirmation surgeries, hormone replacement therapy, electrolysis, hormonal and medical gender affirmation, gender-based stigma, and history of intimate partner violence. For transgender-specific intimate partner violence, we used the Peitzmeier Transgender-Specific Intimate Partner Violence scale, which is tailored for transgender women and consists of four domains: (1) coercive control of gender transition or gender presentation, (2) emphasizing the undesirability of transgender individuals as intimate partners, (3) blackmail via "outing," and (4) sabotaging transition [38]. This tool has demonstrated adequate internal consistency (KR-20=0.56) [38]. To measure gender-based stigma, we used a 13-question scale developed by Maksut et al [39] that assessed perceived, anticipated, and enacted stigmas due to gender identity. This gender identity stigma scale has demonstrated acceptable internal consistency (Cronbach =0.79) [39].

Integration and Testing

Project staff, the principal investigators, and SMEs reviewed and approved each adaptation. With the deployment of the redesigned website and the implementation of the revised

XSL•FO
RenderX

materials for transgender women, we conducted a randomized control trial in which we aimed to recruit up to 200 participants. The eligibility criteria are described in Textbox 1.

Textbox 1. Eligibility and exclusion criteria for an HIV self-test pilot study of transgender women living in 11 US states (2020-2021).

Eligibility criteria

- Provided informed consent online
- At least 18 years of age
- Identifies as a transgender woman or transgender female
- Assigned male sex at birth
- Has an Android or iOS mobile smartphone with active service
- Resides in Alabama, California, Florida, Georgia, Louisiana, Mississippi, North Carolina, Nevada, New York, South Carolina, and Texas
- Provides valid contact information
- Successfully completed the baseline survey

Exclusion criteria

- Declined consent to participate
- Currently participating in another HIV prevention research study or program
- Has a bleeding disorder
- Participant in an HIV vaccine study
- Currently taking pre-exposure prophylaxis for HIV prevention
- Reported a prior positive HIV test result
- Plans to move out of a study state during the 4 months study period

The goal of the pilot study was to provide preliminary data on recruiting transgender women online and the feasibility of enrolling transgender women into an HIV self-testing study. The data collected provide estimates for the length of time needed to recruit participants, the effectiveness of online recruitment strategies, the cost of recruitment, retention rates, the use and acceptability of video counseling services, and the proportion of transgender women who reported using the HIVSTs and STI self-collection kits. These data will be used in conducting power calculations for future HIV self-testing research studies with transgender women.

We planned to use the data to identify key considerations for programmatic implementation of a mail-out HIV self-testing program for transgender women that could be implemented by local health departments or community-based service providers. Factors that are relevant for implementation by these providers include online recruitment and costs to implement.

For this pilot study, potential participants were recruited through online venues, such as social networking sites (eg, Facebook) and dating apps (eg, Jack'd), which were chosen using advertising data from the previous recruitment of transgender women. Participants completed an initial eligibility screening survey, provided electronic informed consent, and completed a baseline survey. An initial validation process was conducted to determine enrollment status. Participants were considered fully enrolled in the study and mailed 2 HIVSTs after they had provided the required contact information, and duplicate accounts were excluded.

We used a two-stage stratified randomization procedure [40]. First, we grouped participants based on race or ethnicity (Black, White, Other, Hispanic or Latina) for a total of 4 strata. Next, we randomly assigned participants from each strata to the intervention arm (mailed self-test kit+provided link to the Know@Home website) or control arm (mailed self-test kit only) with a 2:1 ratio.

Ethical Considerations

This pilot study was approved by the Emory University Institutional Review Board (IRB00099710) and is registered at ClinicalTrials.gov (NCT04219878). Personal identifying information was only stored at Emory University. Confidential information was stored off site on secure web servers. All data files were encrypted with strong password protection. Access to personally identifying information was restricted to Emory study staff. Informed consent was acquired online. Incentives in the form of e-gift cards were provided to participants for completing various study activities totaling up to US $70 over the 4-month study period. Additional optional surveys allowed participants to earn an additional US $60.

Results

The main research outcomes from this pilot study are summarized in Table 3. During the enrollment period, from April 22, 2021, to July 7, 2021, a total of 2014 people consented to be screened for the study, most of whom were ineligible for multiple reasons. Not identifying as female or a transgender woman was the most common reason for ineligibility. Of 217 people who were eligible based on screening criteria, 169

XSL•FO
RenderX

provided contact information. Of these, 108 participants passed the initial validation process, responded to a text message to complete the validation process, and completed the baseline survey. Subsequently, we determined that an additional 6 people had provided false information (eg, they revealed that they did not identify as transgender women) and were excluded from the analyses. A total of 102 transgender women constituted the sample for this pilot study, and 81 provided follow-up information. In total, 20 participants (18%) were non-Hispanic Black, 49 (48%) were non-Hispanic White, 18 (17%) were Hispanic or Latina of any race, and 15 (14%) were of another race.

Table . Research outcomes, data types, and relevance to study design implementation for HIV self-test kit distribution for transgender women, 11 US states, 2020.

Research outcomes	Types of data	Relevance to study design
Feedback on sample advertisements	• Qualitative report from consultant's video conference	• Allows for more effective and culturally appropriate recruitment
Proposed recruitment strategies for transgender women (specific online sites, influencers, advertisement images, language)	• Cost data • Advertisement click-through rates • Eligibility and enrollment counts based on venue and advertisement copy	• Allows for cost analysis for recruitment venues
Feasibility of online enrollment of transgender women	• Characteristics of transgender women enrolled • Ability to obtain target enrollment	• Allows for assessment of potential biases in recruitment of transgender women (eg, by age, region, race, or ethnicity) and development of plans to mitigate selection biases and allows ability to determine if sample size will be met
Time to recruit	• Time required to recruit transgender women	• Allows for planning of appropriate recruitment period and funds for recruitment
Retention rates	• Stratified retention estimates by age • Region • Race or ethnicity	• Allows for planning of sample size of a large-scale trial, accounting for retention
Percent of participants using study HIVST[a]	• Reported use of HIVST kits • Distribution of HIVST kits to others • Reasons for using or not using kits	• Allows for sample size calculations and budgeting and number of HIVST kits required for distribution
HIV prevalence among pilot study participants	• Reported results of HIVST	• Allows for estimation of initial sample size to develop a cohort of transgender women who have a negative HIV status for prospective research; allows planning for needs for video counseling resources
Percent of participants using video counseling services	• Number of counseling sessions requested, scheduled, and attended • Client preferences for time of day and day of week	• Allows for planning of the number of counselors needed and meeting the capacity for counseling sessions each week based on preferences for time of day and day of week

[a]HIVST: HIV self-testing.

Discussion

Principal Findings

This paper documents the process of adapting the iSTAMP protocol developed for Black and Hispanic MSM to be appropriate for engaging transgender women, using the 8-phase ADAPT-ITT model. The purpose of this adaptation is to ensure that study procedures and materials presented to participants are culturally and linguistically appropriate and do not cause discomfort or harm. In our adaptation process, we strove to ensure that study materials would not misgender participants and that questions and responses would be culturally relevant and inclusive [41,42]. These adaptations were crucial to ensure the data obtained were relevant for transgender women who participated in the study and would allow us to gain a comprehensive understanding of the experiences of transgender women. Other adaptation studies reflected the importance of including content related to gender dysphoria, police or law enforcement interactions, HIV testing, PrEP, and antiretroviral treatment as well as reflections of transgender women in the study materials [43]. Several studies used focus groups or community advisory boards, including transgender members [43-45]. Some studies noted the importance of including transgender women at any stage of transition as well as nonbinary individuals [43,44]. The resources developed in this protocol can be used to provide preliminary data and tools for further development and evaluation of a robust trial HIVST

XSL•FO
RenderX

distribution and linkage to prevention services for transgender women in the United States.

Limitations

We encountered several challenges during the development and implementation of this pilot study. This activity was conducted after the completion of a trial for MSM with supplemental funding awarded in the final year of the original study. Consequently, the achieved sample size was half of what was originally planned because there were time constraints in which all activities had to be completed, leading to potential limitations around the ability to assess feasibility. Sample representativeness is limited by sampling technique and sample size.

Only 4 of the 9 consultants were able to participate in the online videoconference due to scheduling conflicts, limiting the input on the adaptation process; however, written feedback was elicited from all 9 consultants, and we had coauthors who are SMEs who contributed to the adaptation. The study was conducted in 11 US states, predominantly located in jurisdictions with a high prevalence of HIV in the United States. The sample does not represent all transgender women in the United States.

Due to time constraints, we created a web-based version of the original intervention developed for use on a mobile phone. We used the same eligibility screener questions from the MSM trial due to timeline constraints, and participants were required to identify as “male to female transgender.” This language may have discouraged some transgender women from participating in the study if they did not identify with this terminology. To earn trust among the community, it is necessary to use appropriate language [41], and future studies can incorporate an expanded eligibility criterion to include transgender women who may identify with alternative terminologies, such as transgender woman, transfeminine, or other terms that are commonly used at the time of study implementation.

This study examines a mobile intervention, which is accessible to people with smartphones. Additionally, nearly half of the study population is non-Hispanic White. Although over 91% of the US population owns a smartphone [46] and over half the US population is non-Hispanic White, these factors could have led to an underrepresentation of groups bearing a high HIV burden among transgender women.

Future Adaptations

During the adaptation process, we identified several key principles that may be useful for future adaptations of protocols or interventions for transgender women. When collecting names and pronouns, researchers can inform participants that legal names may not be required to participate, and when possible, collect both the legal name and the name the participant wishes to use during study participation. Engaging members and representatives of the community in planning the research design and materials can mitigate cultural competency issues when tailoring materials. Involving SMEs on transgender women, particularly when the experts are a part of the intended study population, in planning the research design and materials for future studies for transgender women could allow studies to be culturally appropriate [47-50].

Mixed methods approaches provide an opportunity for increasing cultural competency of materials and survey instruments. Qualitative methods can be used to allow transgender women to provide detailed feedback on materials and terminology before implementation. Validating survey instruments that were designed for other populations as well as formal A or B testing would improve recruitment optimization and study rigor.

Linguistic choices that pathologize or marginalize people are potentially harmful to participants during the research and publication of findings [41,51]. As languages and cultures evolve, it is necessary to adapt terminology so that terms are culturally appropriate for the proposed populations. This may require modifying demographic terminology (Table 2) used to collect data to reflect the culture of the population. Finally, while this pilot study engaged transgender women, future studies could widen inclusion criteria to include other populations who may also benefit from HIV self-test programs, such as nonbinary people, gender-nonconforming people, transgender men, and transmasculine people.

Conclusions

We adapted study materials for an HIV prevention intervention for transgender women. By using the ADAPT-ITT model, we systematically produced materials that are culturally and linguistically appropriate for transgender women. The research process and resulting materials have the potential to inform future research studies among transgender women and underscore the importance of involving SMEs and community members in the development of tailored materials.

Acknowledgments

We thank the study consultants and participants for providing valuable information and participation and Susie P Danner for her review of the paper.

Funding

This study was funded through the CDC Cooperative Agreement #U01PS005181 to Emory University and with supplemental funds from the Minority HIV/AIDS Fund. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

XSL•FO
RenderX

Data Availability

After the primary outcome analyses are completed, deidentified datasets can be made available upon written request to the investigator PSS. Interested parties should propose a specific analytic hypothesis. The data use agreement would be developed to ensure that (1) data are used solely for research purposes without attempting to identify individual participants, (2) there is a clear plan for maintaining the safety and security of the data, and (3) commitment to the destruction or return of the data after analysis.

Authors' Contributions

JAC, RD, PSS, and SS conceptualized, wrote, reviewed, and edited the paper. EO and ILL conceptualized, reviewed, and edited the paper. MH contributed to data curation. TS, LH-W, and AS conceptualized and reviewed the paper. All authors contributed to the investigation.

Conflicts of Interest

TS is the editor in chief of *JMIR Public Health and Surveillance*. The other authors declare no conflicts of interest.

References

1. EHE overview. HIV.gov. URL: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview/ [accessed 2026-01-31]
2. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. JAMA 2019 Mar 5;321(9):844-845. [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
3. Issue brief: the role of HIV self-testing in ending the HIV epidemic. : Centers for Disease Control and Prevention; 2022 URL: https://stacks.cdc.gov/view/cdc/129025/cdc_129025_DS1.pdf [accessed 2026-02-5]
4. HIV surveillance special report: 2019; no. 28. : Centers for Disease Control and Prevention; 2021 Aug 1 URL: https://stacks.cdc.gov/view/cdc/149116 [accessed 2026-01-31]
5. Issue brief: HIV and transgender communities. : Centers for Disease Control and Prevention; 2024 URL: https://stacks.cdc.gov/view/cdc/129026/cdc_129026_DS1.pdf [accessed 2026-01-31]
6. Estimated HIV incidence and prevalence in the United States, 2010–2016. : Centers for Disease Control and Prevention; 2019 URL: https://stacks.cdc.gov/view/cdc/129019/cdc_129019_DS2.pdf [accessed 2026-01-31]
7. HIV surveillance supplemental report: estimated HIV incidence and prevalence in the United States, 2018–2022. : Centers for Disease Control and Prevention; 2024 May 21 URL: https://stacks.cdc.gov/view/cdc/156513 [accessed 2026-01-31]
8. Myers JE, Braunstein SL, Xia Q, et al. Redefining prevention and care: a status-neutral approach to HIV. Open Forum Infect Dis 2018 Jun;5(6):ofy097. [doi: 10.1093/ofid/ofy097] [Medline: 29977957]
9. Hibbert MP, Wolton A, Weeks H, et al. Psychosocial and sexual factors associated with recent sexual health clinic attendance and HIV testing among trans people in the UK. BMJ Sex Reprod Health 2020 Apr;46(2):116-125. [doi: 10.1136/bmjsrh-2019-200375] [Medline: 31666303]
10. Olakunde BO, Pharr JR, Adeyinka DA, Conserve DF. Nonuptake of HIV testing among transgender populations in the United States: results from the 2015 U.S. Transgender Survey. Transgend Health 2022 Oct;7(5):430-439. [doi: 10.1089/trgh.2020.0141] [Medline: 36644483]
11. Poteat T, Wirtz AL, Reisner S. Strategies for engaging transgender populations in HIV prevention and care. Curr Opin HIV AIDS 2019 Sep;14(5):393-400. [doi: 10.1097/COH.0000000000000563] [Medline: 31219887]
12. Sevelius JM, Patouhas E, Keatley JG, Johnson MO. Barriers and facilitators to engagement and retention in care among transgender women living with human immunodeficiency virus. Ann Behav Med 2014 Feb;47(1):5-16. [doi: 10.1007/s12160-013-9565-8] [Medline: 24317955]
13. Frye V, Wilton L, Hirshfied S, et al. “Just because it’s out there, people aren’t going to use it.” HIV self-testing among young, Black MSM, and transgender women. AIDS Patient Care STDS 2015 Nov;29(11):617-624. [doi: 10.1089/apc.2015.0100] [Medline: 26376029]
14. Adams LM, Simoni JM. The need for multi-level mitigation of medical mistrust among social network members contributing to antiretroviral treatment nonadherence in African Americans living with HIV: comment on Bogart et al. (2016). Soc Sci Med 2016 Jun;159:58-60. [doi: 10.1016/j.socscimed.2016.04.025] [Medline: 27173741]
15. Witzel TC, Wright T, McCabe L, et al. Impact and acceptability of HIV self-testing for trans men and trans women: a mixed-methods subgroup analysis of the SELPHI randomised controlled trial and process evaluation in England and Wales. EClinicalMedicine 2021 Feb;32:100700. [doi: 10.1016/j.eclinm.2020.100700] [Medline: 33681732]
16. Van Gerwen OT, Austin EL, Bethune CW, Sullivan PS, Muzny CA. Sexual healthcare and at-home STI test collection: attitudes and preferences of transgender women in the Southeastern United States. Front Public Health 2023;11:1187206. [doi: 10.3389/fpubh.2023.1187206] [Medline: 37325308]
17. Arnold EA, Rebchook GM, Kegeles SM. “Triply cursed”: racism, homophobia and HIV-related stigma are barriers to regular HIV testing, treatment adherence and disclosure among young Black gay men. Cult Health Sex 2014 Jun;16(6):710-722. [doi: 10.1080/13691058.2014.905706] [Medline: 24784224]

XSL•FO
RenderX

18. Poteat T, German D, Flynn C. The conflation of gender and sex: gaps and opportunities in HIV data among transgender women and MSM. Glob Public Health 2016;11(7-8):835-848. [doi: 10.1080/17441692.2015.1134615] [Medline: 26785751]
19. Rocha GM, Cândido RCF, de Carvalho NP, et al. Strategies to increase HIV testing among men who have sex with men and transgender women: an integrative review. BMC Infect Dis 2023 Apr 18;23(1):240. [doi: 10.1186/s12879-023-08124-z] [Medline: 37072705]
20. McKleroy VS, Galbraith JS, Cummings B, et al. Adapting evidence-based behavioral interventions for new settings and target populations. AIDS Educ Prev 2006 Aug;18(4 Suppl A):59-73. [doi: 10.1521/aeap.2006.18.supp.59] [Medline: 16987089]
21. Baumann AA, Stirman SW, Cabassa LJ. Adaptation in dissemination and implementation science. In: Brownson RC, Colditz GA, Proctor EF, editors. Dissemination and Implementation Research in Health: Translating Science to Practice: Oxford University Press; 2017:172-191. [doi: 10.1093/oso/9780197660690.003.0008]
22. Horvath KJ, Walker T, Mireles L, Bauermeister JA, Hightow-Weidman L, Stephenson R. A systematic review of technology-assisted HIV testing interventions. Curr HIV/AIDS Rep 2020 Aug;17(4):269-280. [doi: 10.1007/s11904-020-00506-1] [Medline: 32507984]
23. Tanner AE, Song EY, Mann-Jackson L, et al. Preliminary impact of the weCare social media intervention to support health for young men who have sex with men and transgender women with HIV. AIDS Patient Care STDS 2018 Nov;32(11):450-458. [doi: 10.1089/apc.2018.0060] [Medline: 30398955]
24. Reback CJ, Fletcher JB, Kisler KA. Text messaging improves HIV care continuum outcomes among young adult trans women living with HIV: text me, girl!. AIDS Behav 2021 Sep;25(9):3011-3023. [doi: 10.1007/s10461-021-03352-3] [Medline: 34164763]
25. Reback CJ, Blue T, Jalali A, et al. Comparative- and cost-effectiveness research determining the optimal intervention for advancing transgender women with HIV to full viral suppression (text me, Alexis!): protocol for a randomized controlled trial. JMIR Res Protoc 2025 Jan 23;14:e65313. [doi: 10.2196/65313] [Medline: 39847410]
26. Kuo I, Liu T, Patrick R, et al. Use of an mHealth intervention to improve engagement in HIV community-based care among persons recently released from a correctional facility in Washington, DC: a pilot study. AIDS Behav 2019 Apr;23(4):1016-1031. [doi: 10.1007/s10461-018-02389-1] [Medline: 30627850]
27. Holloway IW, Jordan SP, Dunlap SL, Ritterbusch A, Reback CJ. Leveraging social networks and technology for HIV prevention and treatment with transgender women. AIDS Educ Prev 2020 Apr;32(2):83-101. [doi: 10.1521/aeap.2020.32.2.83] [Medline: 32539480]
28. Arnold EM, Rotheram-Borus MJ, Christodoulou J, et al. Increasing sustained viral suppression among youth living with HIV: a randomized controlled trial of stepped care intervention. AIDS Behav 2025 Jul;29(7):2287-2298. [doi: 10.1007/s10461-025-04692-0] [Medline: 40185959]
29. MacGowan RJ, Chavez PR, Borkowf CB, et al. Effect of internet-distributed HIV self-tests on HIV diagnosis and behavioral outcomes in men who have sex with men: a randomized clinical trial. JAMA Intern Med 2020 Jan 1;180(1):117-125. [doi: 10.1001/jamainternmed.2019.5222] [Medline: 31738378]
30. Dana R, Sullivan S, MacGowan RJ, et al. Engaging Black or African American and Hispanic or Latino men who have sex with men for HIV testing and prevention services through technology: protocol for the iSTAMP comparative effectiveness trial. JMIR Res Protoc 2023 Jan 6;12:e43414. [doi: 10.2196/43414] [Medline: 36607707]
31. Sullivan PS, Stephenson R, Hirshfield S, et al. Behavioral efficacy of a sexual health mobile app for men who have sex with men: randomized controlled trial of mobile messaging for men. J Med Internet Res 2022 Feb 2;24(2):e34574. [doi: 10.2196/34574] [Medline: 35025755]
32. Wingood GM, DiClemente RJ. The ADAPT-ITT model: a novel method of adapting evidence-based HIV Interventions. J Acquir Immune Defic Syndr 2008 Mar 1;47 Suppl 1:S40-S46. [doi: 10.1097/QAI.0b013e3181605df1] [Medline: 18301133]
33. Everhart AR, Boska H, Sinai-Glazer H, et al. “I’m not interested in research; I’m interested in services”: how to better health and social services for transgender women living with and affected by HIV. Soc Sci Med 2022 Jan;292:114610. [doi: 10.1016/j.socscimed.2021.114610] [Medline: 34923191]
34. Sevelius JM, Keatley J, Calma N, Arnold E. ‘I am not a man’: trans-specific barriers and facilitators to prep acceptability among transgender women. Glob Public Health 2016;11(7-8):1060-1075. [doi: 10.1080/17441692.2016.1154085] [Medline: 26963756]
35. MacGowan RJ, Chavez PR, Gravens L, et al. Pilot evaluation of the ability of men who have sex with men to self-administer rapid HIV tests, prepare dried blood spot cards, and interpret test results, Atlanta, Georgia, 2013. AIDS Behav 2018 Jan;22(1):117-126. [doi: 10.1007/s10461-017-1932-1] [Medline: 29058163]
36. Gender inclusive language building relationships with new clients. : TransCare BC; 2022 URL: https://www.transcarebc.ca/sites/default/files/2024-03/Gender_Inclusive_Language_General.pdf [accessed 2026-01-31]
37. Bauer GR, Braimoh J, Scheim AI, Dharma C. Transgender-inclusive measures of sex/gender for population surveys: mixed-methods evaluation and recommendations. PLoS One 2017;12(5):e0178043. [doi: 10.1371/journal.pone.0178043] [Medline: 28542498]

XSL•FO
RenderX

38. Peitzmeier SM, Hughto JMW, Potter J, Deutsch MB, Reisner SL. Development of a novel tool to assess intimate partner violence against transgender individuals. J Interpers Violence 2019 Jun;34(11):2376-2397. [doi: 10.1177/0886260519827660] [Medline: 30735080]
39. Maksut JL, Sanchez TH, Wiginton JM, et al. Gender identity and sexual behavior stigmas, severe psychological distress, and suicidality in an online sample of transgender women in the United States. Ann Epidemiol 2020 Dec;52:15-22. [doi: 10.1016/j.annepidem.2020.07.020] [Medline: 32768521]
40. Kernan WN, Viscoli CM, Makuch RW, Brass LM, Horwitz RI. Stratified randomization for clinical trials. J Clin Epidemiol 1999 Jan;52(1):19-26. [doi: 10.1016/s0895-4356(98)00138-3] [Medline: 9973070]
41. Adams N, Pearce R, Veale J, et al. Guidance and ethical considerations for undertaking transgender health research and institutional review boards adjudicating this research. Transgend Health 2017;2(1):165-175. [doi: 10.1089/trgh.2017.0012] [Medline: 29098202]
42. Tebbe EA, Budge SL. Research with trans communities: applying a process-oriented approach to methodological considerations and research recommendations. Couns Psychol 2016 Oct;44(7):996-1024. [doi: 10.1177/0011000015609045]
43. Lee JJ, Cruz Y, Aguirre J, et al. Adaptation of an evidence-based HIV prevention intervention to address the needs of transgender Latina women. Transgend Health 2024 Sep 12. [doi: 10.1089/trgh.2024.0054] [Medline: 41585992]
44. Kuhns LM, Hereth J, Garofalo R, et al. A uniquely targeted, mobile app-based HIV prevention intervention for young transgender women: adaptation and usability study. J Med Internet Res 2021 Mar 31;23(3):e21839. [doi: 10.2196/21839] [Medline: 33787503]
45. MacDonell KK, Wang B, Phanuphak N, et al. Optimizing an mHealth intervention to improve uptake and adherence to HIV pre-exposure prophylaxis in young transgender women: protocol for a multi-phase trial. JMIR Res Protoc 2022 May 19;11(5):e37659. [doi: 10.2196/37659] [Medline: 35587370]
46. Mobile fact sheet. Pew Research Center. 2025. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2026-01-31]
47. Katz-Wise SL, Sansfaçon AP, Bogart LM, et al. Lessons from a community-based participatory research study with transgender and gender nonconforming youth and their families. Action Res (Lond) 2019 Jun;17(2):186-207. [doi: 10.1177/1476750318818875] [Medline: 40303848]
48. Singh AA, Richmond K, Burnes TR. Feminist participatory action research with transgender communities: fostering the practice of ethical and empowering research designs. Int J Transgend 2013 Jul;14(3):93-104. [doi: 10.1080/15532739.2013.818516]
49. Andrasik MP, Yoon R, Mooney J, et al. Exploring barriers and facilitators to participation of male-to-female transgender persons in preventive HIV vaccine clinical trials. Prev Sci 2014 Jun;15(3):268-276. [doi: 10.1007/s11121-013-0371-0] [Medline: 23446435]
50. Mann-Jackson L, Alonzo J, Garcia M, et al. Using community-based participatory research to address STI/HIV disparities and social determinants of health among young GBMSM and transgender women of colour in North Carolina, USA. Health Soc Care Community 2021 Sep;29(5):e192-e203. [doi: 10.1111/hsc.13268] [Medline: 33369811]
51. Martin JI, Meezan W. Applying ethical standards to research and evaluations involving lesbian, gay, bisexual, and transgender populations. J Gay Lesbian Soc Serv 2003 Apr;15(1-2):181-201. [doi: 10.1300/J041v15n01_12]

Abbreviations

ADAPT-ITT: Assessment, Decision, Adaptation, Production, Topical Experts, Integration, Training, and Testing
HIVST: HIV self-test
iSTAMP: Implementation of Rapid HIV Self-Testing Among MSM Project
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
SME: subject matter expert
STI: sexually transmitted infection

Edited by J Sarvestan; submitted 02.Apr.2025; peer-reviewed by D Westmoreland; revised version received 16.Dec.2025; accepted 17.Dec.2025; published 20.Feb.2026.

Please cite as:
Caldwell JA, Dana R, Lucas IL, Sullivan PS, Olansky E, Sullivan S, Hannah M, Sanchez T, Hightow-Weidman L, Sharma A Adaptation of a Protocol to Evaluate Online Recruitment Strategies for Mailing HIV Self-Tests and Pilot an Intervention to Improve Linkage to HIV Prevention and Care Among Transgender Women: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e75345
URL: https://www.researchprotocols.org/2026/1/e75345
doi:10.2196/75345

XSL•FO
RenderX

© Joanna A Caldwell, Ruth Dana, Irah L Lucas, Patrick Sean Sullivan, Evelyn Olansky, Stephen Sullivan, Marissa Hannah, Travis Sanchez, Lisa Hightow-Weidman, Akshay Sharma. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Therapeutic Improvement in People With Schizophrenia Undergoing tACS/CBTp (Transcranial Alternating Current Stimulation/Cognitive Behavioral Therapy for Psychosis) Associated With Usual Medication Regimen: Protocol for a Pilot, Randomized, Single-Blind Trial

Felicia Iftene[1,2*], MD, PhD; Adriana Farcas[1*], PhD; Simon O'Brien[1,2*], MD; Christopher Bowie[3], Prof Dr; Michael Best[4], PhD; Oyedeji Ayonrinde[1], Prof Dr Med; Terry Landry[2], PhD; Jennifer Carlson[5], MSc; Scott John Davidson[5], MSc; Ellie Rodgerson[6], MA; Ashley Theis[6], MSc

[1]Division of Adult Psychoses, Department of Psychiatry, Queen's University, Kingston, ON, Canada

[2]Providence Care Mental Health Services, Kingston, ON, Canada

[3]Department of Psychology, Queen's University, Kingston, ON, Canada

[4]Department of Psychology, University of Toronto, Toronto, ON, Canada

[5]Center for Neuroscience Studies, Queen's University, Kingston, ON, Canada

[6]Research Compliance and Training- Vice-Principal Research Portfolio, Queens University, Kingston, ON, Canada

[*]these authors contributed equally

Corresponding Author:
Felicia Iftene, MD, PhD
Division of Adult Psychoses
Department of Psychiatry
Queen's University
24 Kingsgate Pl
Kingston, ON, K7M 7K8
Canada
Phone: 1 6134847330
Email: iftenef@providencecare.ca

Abstract

Background: Transcranial alternating current stimulation (tACS) applies a low-intensity sinusoidal electrical current through electrodes placed on the scalp to boost the brain's own oscillation by way of entrainment. When a single frequency is applied, this exogenous oscillation synchronizes with the brain's endogenous frequency. Gamma frequency synchrony stands out as a binding mechanism for integrating disparate brain networks, mediating perception, cognition, and memory, typically disturbed in schizophrenia. The treatment of schizophrenia includes medication and cognitive behavioral therapy for psychosis (CBTp). We are adding tACS to these usual treatments, targeting gamma oscillation stimulation, to augment the CBTp efficacy in people living with schizophrenia.

Objective: This study aims to elicit cognitive readiness and therapeutic engagement by adding tACS to each CBTp session in individuals with schizophrenia taking their usual medication, to evaluate the possible improvement of the level of functioning, and to determine if the response to intervention is gender specific.

Methods: This is a pilot, prospective, randomized, repeated-measures, single-blind study design. We expect to enroll 28 participants randomly assigned to one of two treatment arms: arm 1 (tACS/CBTp, n=14) or arm 2 (sham tACS/CBTp, n=14; tACS is sham, but CBTp is active). The intervention with active or sham tACS/CBTp will take place weekly for 16 weeks. Primary outcome measures—electroencephalogram, Positive and Negative Syndrome Scale, 16-item Negative Symptom Assessment, and Cognitive Flexibility Scale—will evaluate the efficacy of treatment at the end of the intervention and at the two follow-ups. We will use SPSS (version 29, IBM Corp); the main tests will include repeated measures and mixed design ANOVA.

Results: The timeline for recruitment, treatment, and follow-ups is 18 months, followed by 6 months for data analysis, writing manuscripts, and dissemination activities. By November 1, 2025, we have enrolled 15 participants: 12 are following the intervention

XSL•FO
RenderX

protocol (8 active and 4 sham tACS). Two participants were screening failures, and one participant withdrew after intervention 2.

Conclusions: Our expectations are as follows. CBTp will improve the scores of psychological and psychosocial tests at the end of therapy for both groups, but it will be superior for the group with tACS intervention. Considering that cognitive and emotional status is gender dependent (hormonal differences, brain structure, and sociocultural influences), we expect that the therapeutic response could be gender specific. CBTp will enhance electroencephalogram activity and the heart in clients with schizophrenia at the end of therapy for both groups, but it will be superior for the group with tACS preintervention. The baseline heart rate variability will predict symptom improvement and will increase over the course of therapy. We hope our research will improve the treatment of people living with schizophrenia, thereby enhancing their quality of life, reducing the rate of relapse, and lowering the costs of care.

Trial Registration: ClinicalTrials.gov NCT06889025; https://clinicaltrials.gov/study/NCT06889025

International Registered Report Identifier (IRRID): PRR1-10.2196/80593

(JMIR Res Protoc 2026;15:e80593) doi:10.2196/80593

KEYWORDS

schizophrenia; tACS; transcranial alternating current stimulation; CBTp; cognitive behavioral therapy for psychosis; EEG; electroencephalogram; heart rate variability; gender; clinical or functional improvement

Introduction

Background and Scientific Rationale

Schizophrenia is a debilitating neurodevelopmental disorder with tremendous consequences for the individual and family, resulting in an important socioeconomic burden. There are approximately 360,000 people affected by schizophrenia in Canada. The illness usually begins between the ages of 15 and 35 years, during key developmental milestones of life. Ontario has the largest number of individuals living with schizophrenia; 84,000 Ontarians (and their families) are affected by treatment-resistant schizophrenia.

The mainstay treatment of schizophrenia is antipsychotic medication. However, antipsychotics do not work adequately for up to 60% of individuals [1]. Although evidence shows that medication has beneficial effects on positive symptoms (hallucinations, delusions, etc) and reduces relapses, a significant number of patients do not fully respond to it [2]. The persistence of symptoms throughout its evolution, compounded by the added side effects of medication, leads to significant impairments in social and relational aspects of life, contributing to the significant distress and disability associated with the disorder. Combined, early medical and psychosocial interventions are consistently sought, while research explores new avenues to enhance current approaches or attempt stand-alone, successful, cost-effective, and noninvasive treatments [3]. When cognitive behavioral therapy for psychosis (CBTp) is used alongside medication, some improvements have been observed. Still, the debilitating negative symptoms remain (such as reduction of emotional expression, lack of social involvement, inability to initiate goal-directed activities, etc), which are key factors in long-term disability [4,5].

Schizophrenia is characterized by abnormalities in neural circuitries resulting in dysfunctional cognition and behavior [6,7]. A new treatment, transcranial alternating current stimulation (tACS), has therapeutic results in Alzheimer disease [8-10]. Recent randomized double-blind clinical trials of tACS in Alzheimer disease showed a significant improvement in memory performance, along with restoration of intracortical connectivity, as compared to sham tACS [11-13]. Recent systematic reviews discuss the general optimism within the scientific community regarding the use of electrical currents in enhancing cognitive performance in healthy people and ameliorating symptoms in a wide variety of conditions, including neuropsychiatric ones [6,14-16]. The rationale behind such an approach rests on the understanding that neural oscillations are a fundamental mechanism that organizes the temporal relationship of activity patterns in brain networks [2,17,18]. Certain psychiatric disorders were identified over the last decades as showing imbalances in network connectivity, showing high or slow oscillatory activity by comparison with the healthy population. If tACS improves memory in Alzheimer disease by modulating brain oscillations, then it may also enhance cognitive function in schizophrenia by targeting similar neural mechanisms, particularly those involving working memory and gamma oscillations. Of note, gamma synchrony stands out as a binding mechanism for integrating disparate brain networks, mediating perception, cognition, and memory [3], which is typically disturbed in schizophrenia. Manifesting as synchronous high-frequency oscillations of brain electrical activity, it occurs across several brain regions whose functions are integrated this way [19,20]. Global delays and decreased temporal connectivity in neural activity stemming from impairments in frontal lobe synchrony may be relevant to the disordered cognitive control and flexibility in schizophrenia [21-23].

tACS, the focus of our study, applies a low-intensity sinusoidal electrical current to the brain through electrodes placed on the scalp, and it is thought to boost the brain's own oscillation by way of entrainment. When a single frequency is applied, this exogenous oscillation will synchronize with the brain's endogenous frequency. Conversely, multiple frequencies lead to desynchronization of cortical oscillations [24]. The value of the frequency applied thus becomes a determining factor in the effects sought and identified in tACS studies.

XSL•FO
RenderX

Research focusing on tACS mechanisms has shown that effects result from manipulating the membrane potential of neurons that are aligned with the introduced electric field, mostly pyramidal cells in layer V [25,26]. Upon stimulation by tACS, which will cause an alternating change in membrane potential, these cells demonstrate a resonance specific to the frequency of stimulation, with long-term aftereffects—70 minutes after one stimulation session for 20 minutes [19]—and implications on oscillatory cortical connectivity between different cortical regions [27]. In auditory hallucinations, research has demonstrated interhemispheric miscommunication, where tACS can manipulate auditory perception in healthy participants by decoupling this interhemispheric connectivity [28]. The advantage of using tACS over other noninvasive brain stimulation modalities lies in its effectiveness in entraining endogenous brain oscillations, as it mimics the alternating nature of these oscillations [29,30].

From a practical standpoint, tACS exhibits superior cost, portability, tolerability, and safety profiles [17]. tACS is a feasible tool that reshapes or resynchronizes intrinsic brain rhythms, manipulating the associated brain functions without adding extra excitatory or inhibitory burden.

We hypothesize that tACS targeting gamma oscillations, added to each CBTp session, will improve cognitive performance in individuals living with schizophrenia under their usual antipsychotic medication by restoring neural synchrony and enhancing cortical connectivity. This will subsequently improve clinical and functional status and medication adherence.

Proposed Therapeutic Protocol

The investigators are proposing a new, noninvasive therapeutic model using tACS to augment CBTp efficacy in individuals with schizophrenia. By selecting electroencephalogram (EEG) brain oscillation activity as a biomarker for the progression of cognitive deficits in schizophrenia, this research seeks to determine whether addressing oscillation perturbations can mitigate cognitive deficits. The heart rate variability (HRV) was selected as a biomarker of improvement of somatic and mental health. The investigators are also aiming for an analysis through a Gender-Based Analysis Plus lens by using the Bem Sex-Role Inventory (BSRI), along with specific tests for psychosis (Positive and Negative Syndrome Scale [PANSS], 16-item Negative Symptom Assessment [NSA-16], etc).

Objectives

This study compares the effects of a harmless, low-voltage electrical stimulation (tACS) with placebo stimulation (sham tACS) applied for 20 minutes at the beginning of the CBTp session to increase the efficacy of psychotherapy in people living with schizophrenia.

Primary objectives are to determine whether (1) the new intervention protocol using tACS will enhance the EEG activity, expecting that CBTp will enhance EEG activity in clients with schizophrenia at the end of therapy for both groups, with a better outcome for the group with active tACS stimulation; and (2) tACS applied at the beginning of the CBTp sessions in people with schizophrenia will predict better responses to therapy (compared with the sham tACS/CBTp). CBTp is expected to improve the scores of the PANSS, NSA-16, and Cognitive Flexibility Scale (CFS).

The secondary objectives are to determine whether (1) there will be an improvement of level of functioning and quality of life in both intervention groups, with a superior outcome for the group receiving tACS intervention (psychological and psychosocial tests); (2) higher in-session HRV predicts a better response to CBTp (the expectation is that baseline HRV will predict symptom improvement and increase over the course of therapy); and (3) the response to tACS/CBTp intervention is gender specific; considering that cognitive and emotional status is gender dependent, we expect that the therapeutic response could be gender specific.

Methods

Trial Design

Description

This is a prospective, randomized, repeated-measures, single-blind study design. Preintervention, eligible participants will be randomly assigned to one of two treatment arms: arm 1 (tACS/CBTp, n=14) or arm 2 (sham tACS/CBTp, n=14; tACS is sham, but CBTp is active). In the intervention (16 weeks), participants in arms 1 and 2 will receive once-weekly tACS/CBTp or sham tACS/CBTp. Individuals with schizophrenia (*DSM-5* [*Diagnostic and Statistical Manual for Mental Disorders* {Fifth Edition} criteria) stratified by age and sex. The investigators expect 150 potentially eligible patients from Providence Care Mental Health Services, with 33-40 participants consenting to participate and 28 participants undergoing the research activities (see Figure 1).

Participants will be randomized 1:1 to one of two arms and stratified by age (18-30, 31-55, and >60 years) and by sex (male or female).

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.

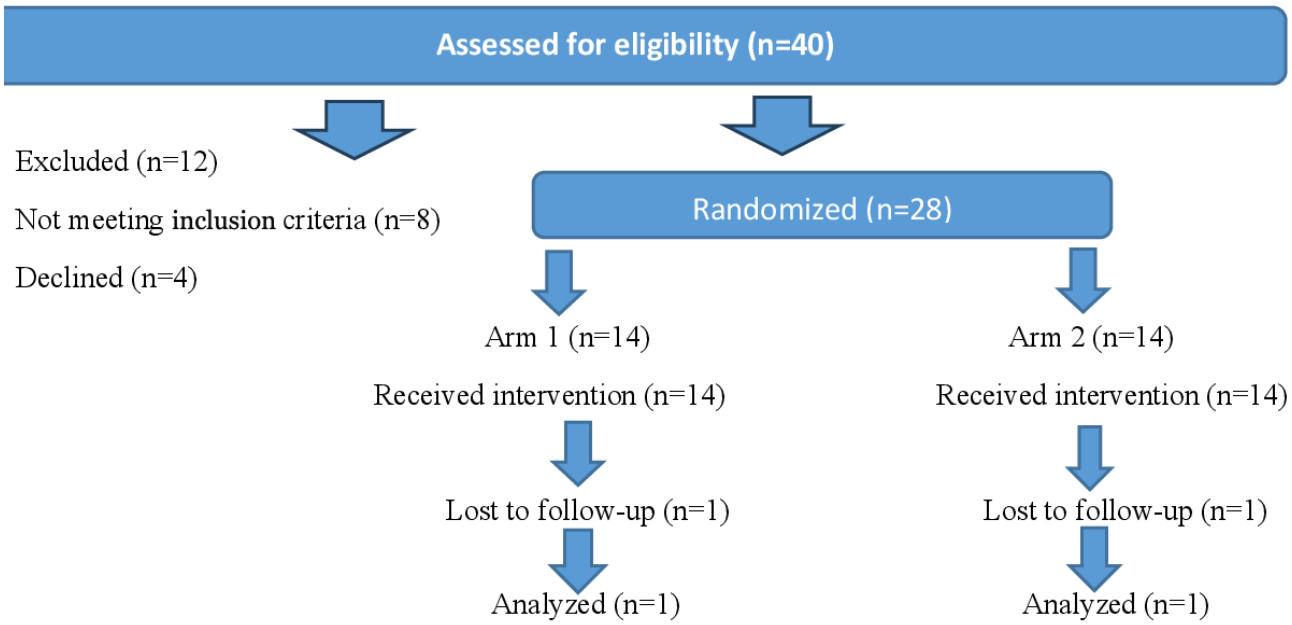

CBTp Intervention

Participants on both arms will receive CBTp in 16 individual-based, 50-minute-long, weekly sessions with 2 booster CBTp sessions (with active or sham tACS) at 1-month and 3-month follow-up visits (after intervention).

tACS/Sham tACS

Participants randomized to the tACS arm will receive gamma band electric stimulation for 20 minutes at the beginning of each CBTp session. After the first 20 minutes of brain stimulation, the tACS device automatically stops, and the device remains in place while the CBTp session continues to run for 30 more minutes.

Participants randomized to the sham tACS arm will receive sham tACS for 20 minutes at the beginning of each CBTp session. After the first 20 minutes of sham tACS, the device automatically stops, and it remains in place, while the CBTp session continues to run for 30 more minutes.

Selection and Withdrawal of Participants

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria for the study participants are listed in Textbox 1.

Changes in medication/new hospitalization for worsening symptoms and/or presence of suicidal ideation are no more exclusion criteria for participants that have passed visit 5 (end of intervention, during follow-up stage); however, the results of the follow-up visits will count at the final results, depending on the severity/imminent danger of symptoms and medication changes, at the best judgment of the principal investigator (FI).

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Individuals, with at least 5 years' duration of illness, who meet diagnostic criteria for schizophrenia according to the *Diagnostic and Statistical Manual for Mental Disorders* (Fifth Edition) and at least one residual positive symptom (as determined by the referring physician)
- No change in medication regimen for at least 1 month, preferably 3 months (minor dose adjustments and/or change in medication involving symptoms, such as sleep, anxiety, or medical symptoms such as fever and pain, are permitted)
- All genders between ages 18-65 years; participants >65 years may be eligible if cognition assessed by the Montreal Cognitive Assessment score is greater than or equal to 26
- Ability to understand English with a reading level at or above grade 6
- Able to understand and comply with the requirements of the study
- Provision of written informed consent

Exclusion criteria

- Current drug substance use
- Current suicidal ideation/plan
- Current enrollment in cognitive behavioral therapy for psychosis or other formalized psychosocial interventions
- Undergone vagotomy or surgery on the vagus nerve
- Comorbid neurological condition, including seizures
- Current fibrillation or pacemaker
- Severe or moderate intellectual disability
- Currently undergoing hormone therapy
- Under 18 years of age
- Changes in medication/new hospitalization for worsening symptoms and/or presence of suicidal ideation

Criteria for Withdrawal

Premature withdrawal or suspension may occur if imminent safety concerns are disclosed to a study team member or a member of the CBTp Clinic by the participant. Withdrawal would be considered with the best interests and safety of the participant considered foremost. An example of this would be if a participant indicated suicidal or homicidal ideation with a plan to carry out such an act at any point in time. After safety considerations are met, the participant would be eligible to return or be included in the study. The research team may temporarily or permanently halt the study if there is any reason to believe that a participant has a medical reason to discontinue participation. Some of the reasons why this might happen are listed below:

- Their condition worsens or does not improve, and their doctor thinks they need a different treatment.
- The study treatment or procedures are found to be unsafe or ineffective.
- They are unable to follow instructions given to them about the study, or they otherwise cannot do or continue to do what they need to participate in the study.
- They develop another serious disease.
- They become pregnant.
- Cancellation by the sponsor or for other unforeseen reasons that make it necessary to stop their study participation.

If a participant is removed from the study, the study doctor will explain to them why the participant was removed. A participant may also be temporarily suspended or stop their participation in the study if they require attention due to a risk of harm to themselves or others. This may be indicated by a member in their circle of care or by the clinician-researchers on the team. In the event there is a risk of harm to themselves or others, after the risk is no longer deemed significant, they may be able to continue the study.

If participants who have passed visit 5 (end of intervention, during the follow-up stage) and do not withdraw their consent to participate, their data can be considered for use.

Participant withdrawal forms have been created to accurately document any withdrawal.

Inclusion of Women and Minorities

Our project is gender and minority inclusive. All participants included will have a diagnosed mental health disorder of schizophrenia. The investigators plan to include individuals who are both inpatients and outpatients at Providence Care Hospital. This includes participants who may be in long-term care.

Disability

Individuals will be excluded if they have severe or moderate intellectual disability or dual diagnosis due to the nature of CBTp, which requires the learning of coping strategies, adequate attention and memory, engagement using speech, and appreciation of consequences of action, thoughts, behaviors,

XSL•FO
RenderX

emotions, and beliefs, which are limited in severe intellectual disability.

Language and Linguistic Proficiency

Individuals will be excluded if they cannot understand English with a reading level at or above grade 6. The ability for an individual to meaningfully engage in a CBTp talk therapy would be limited if their English level is limited to a grade 6 level and is not thought to be substantially beneficial to the individual.

Age

Participants recruited will be included if they are aged 18-65 years. Individuals older than 65 years of age will be required to be assessed by the Montreal Cognitive Assessment (MoCA) and will be included in the study if results validate adequate cognitive performance. The MoCA will be administered at the screening stage (a score under 26 will be the exclusion criterion).

Recruitment

A letter and a poster will be distributed to psychiatrists at Providence Care Mental Health Services, informing them that they can refer clients to the CBTp Clinic, and they are invited to refer potential participants who meet an inclusion or exclusion criterion. Word of mouth and brochures or flyers will be prepared. The clinic staff or the participant will then contact our research coordinator or assistant to provide the potential participant with more information. The investigators plan to include individuals who are both inpatients and outpatients at Providence Care Hospital. This includes participants who may be in long-term care. Potential participants in the study will be referred to the CBTp Clinic at Providence Care Hospital by the participant's primary treating psychiatrist.

During the first meeting at the CBTp Clinic, the potential participant will be informed about the study. The potential participant will be informed that their participation in the study is entirely optional, alongside the standard clinic disclosures about privacy/confidentiality. They will have the opportunity to consider further involvement in a study, with a break provided if they choose to meet with clinic personnel and provide written consent.

Express written consent to participate in the study will be obtained by a member of the research team who is not in the circle of care of the participant (at arm's length).

Participant Information and Informed Consent Procedures

Before data collection, potential participants will be fully informed of the observational nature of the study, in that the sponsor intends to collect information and follow the course of therapy in clinical practice. They will be told that their consent to allow the collection of information within the context of this noninterventional study is voluntary and may be withdrawn at any time. Only participants who are fully able to understand the nature of the research and provide their consent voluntarily will be enrolled. Participants must provide their informed consent after a comprehensive and understandable explanation of the clinical study's nature, scope, and potential consequences. After reviewing the document, participants must give written consent. The consent process must be validated by the personally dated signatures of both the participant and the person facilitating the informed consent discussion. Participants will receive a copy of the signed consent document, while the investigator will keep the original. No study-specific procedures will be carried out until valid consent is obtained. If the participant agrees, their primary physician will be informed about their participation in the study. Where a participant has difficulty reading, a member of the research team will verbally communicate the information on a form or questionnaire if a participant requires clarification. Every effort will be made to help eligible participants with communication difficulties. In this study, the investigators do not enroll participants who are incapable of signing informed consent. To proactively monitor changes in the status of capacity, midway through the study, the research team will seek confirmation of the capacity status of the participant from health care personnel. As is typical ethical practice outside the scope of this study, if members of the CBTp Clinic (by way of observation or professional judgment) suspect a change in capacity in the participant, the participant's primary treating psychiatrist will be notified, and a reassessment will take place.

Reimbursement

The investigators will provide each participant with a US $10 reimbursement to cover out-of-pocket expenses, such as a snack and bus ticket, for each of the 6 required visits as part of the study. Payments will be provided after each task is complete. Participants will be encouraged to bring a lunch with them, as indicated on the informed consent form.

Enrollment and Randomization

After the screening visit, participants will be randomized in a 1:1 ratio with stratification according to age and sex.

Maintaining the blind is covered by the device, which induces a discrete tickling sensation at the starting point of the intervention for participants on sham tACS. This study is single blind, being blinded only for participants. In case of emergency, the research team would be aware of the interventional (or sham) nature of the problem. Communication with participants regarding this aspect will be possible only at the end of the study, at the end of visit 6 (second follow-up), at the patient's request, or after they withdraw from the study. At the study closure, each participant will receive a debriefing letter containing unblinding disclosure.

Trial Intervention

tACS

This trial uses a Soterix Medical 1×1 Mini-CT tES Stimulator, licensed for use in Canada. It is a small and portable device, with 3 rechargeable 800 mAh batteries and the necessary accessories and calibration to begin tACS right out of the box. The device does not allow initiation of stimulation if the contact quality is outside the desired range. If the session is already underway, the device will alert the user with continuous beeping to prompt immediate corrective action. Stimulation is also recorded by the device, which provides a detailed log of the stimulation session, including pause events, critical events such as poor contact quality, and the total time of the session.

XSL•FO
RenderX

Frequency has a 1 Hz resolution and can be adjusted from 1 to 200 Hz. Current intensity has a 0.01 mA resolution and can range from +0.10 to 4.00 mA. Session duration and ramp duration are also adjustable, with session times ranging from 5 to 60 minutes and ramp times from 10 to 30 seconds. The electronic unit is intended to generate the electrical pulses of a specified waveform, duration, and repetition rate. The pulses are delivered to a patient via stimulating electrodes positioned at the patient's head according to the required technique. The mini-CT is European Union Conformity marked. The SNAP Headset can be used to position Soterix Medical SNAP pads at designated transcranial direct current stimulation and transcranial electric stimulation locations, maximizing reproducibility and participant comfort. The custom headgear with fixed electrode sites and a universal, one-size-fits-all design makes for a simplistic setup for 1×1 stimulation.

CBTp

The intervention consists of 16 individual-based weekly sessions, with follow-up assessments (1 month and 3 months after the intervention). The trial intervention is summarized in Table 1.

Table 1. Trial intervention.

Arms	Assigned intervention
Experimental: tACS[a]/CBTp[b] (both active)	• tACS: gamma band electric stimulation for 20 minutes at the beginning of each CBTp session. After 20 minutes, the device remains in place, but the electric stimulation will be automatically stopped. The CBTp will continue up to a total of 50 minutes/session. • CBTp: 16 individual-based, weekly sessions each of 50 minutes in duration and 2× booster sessions at months 1 and 3 after intervention.
Sham comparator: sham tACS/CBTp (only CBTp active)	• For the sham group, the investigators will use the same tES[c] device; the sham stimulation will be applied only for a few seconds in the sham group. • The CBTp will follow the same protocol for both arms.

[a]tACS: transcranial alternating current stimulation.

[b]CBTp: cognitive behavioral therapy for psychosis.

[c]tES: transcranial electric stimulation.

Concomitant Medications/Interventions

Permitted/Supportive

Participants will be on their usual medication regimen, with no significant changes for at least 1 month (preferably 3 months) before enrollment. Small, symptomatic medication changes are permitted at the principal investigator's judgment. These changes should not interfere with the assessments. For example, the use of benzodiazepines will be on hold from midnight of the day before the evaluations. Aside from antipsychotic medication, any pharmacological products are permitted as long as they are part of the usual medication regimen of the participant. Minor adjustments or symptom-treating medication (eg, insomnia and anxiety) are permitted at the principal investigator's best judgment, provided they do not interfere with project activities.

Other interventions are permitted, but only if they are continuous and started 3 months before enrollment. For example, the participants would not be allowed to start a keto diet during the study, but they will be allowed to follow that diet if they started at least 3 months before enrollment, and it will be ongoing until the end of the study.

Not Permitted Concomitant Medications/Interventions

Participants must withhold benzodiazepines and other medications for sleep or anxiety for at least 8 hours before assessments, avoid using illicit drugs during research activities, refrain from hormone therapy throughout trial participation, and not engage in other CBTp or formalized psychosocial interventions concurrently with the study intervention.

Participant Compliance

Being administered by professionals (and not self-administered), it is easy to appreciate the compliance with the intervention. The Personal Evaluation of Transitions in Treatment (PETiT) questionnaire assesses compliance with medication. Even if the participant agrees to participate in this study, they must continue to take their current pharmacological treatment.

Participants who are unable to follow instructions given to them about the study may be temporarily suspended or withdrawn from the study.

Other Participant Care Considerations

The primary standard of care is medication management for adults with schizophrenia. When an individual requires other services or care, generally a referral is made to obtain these services. A proportion of adults with schizophrenia in the community may also have social or skill support via a community treatment team. The research laboratory is situated in a hospital setting, which is well resourced to support care in case of an emergency on-site. Participants will be receiving their standard, usual care throughout the study.

Efficacy Assessments/Measurement of Effect

Therapeutic improvement, in our study, was defined as expected by statistically significant changes in some of our primary as well as secondary outcomes. To capture the progress in clinical improvement, we chose 2 validated scales that show high reliability and sensitivity to schizophrenia symptomatology—the PANSS and the NSA. The CFS was added as an additional measure for the executive function. We considered participants to be good responders to the intervention if there was an increase

XSL•FO
RenderX

in PANSS General Scores higher than or equal to 20% of their baseline scores at the end of the intervention, with an improvement in scores of the PANSS Positive and PANSS Negative scales of at least 15% compared with the baseline. The therapeutic improvement also included a 20% decrease in scores for the NSA scale and a 20% increase in CFS scores at the end of the intervention, compared to baseline.

Safety Considerations

An adverse event (AE) is any untoward medical occurrence in a patient administered an intervention (tACS in our specific project). An AE does not necessarily have a causal relationship with the treatment. An AE can be any unfavorable and unintended sign (including an abnormal finding or lack of expected interventional action), symptom, or disease temporally associated with the use of a tACS, whether or not related to tACS (definition based on the International Council for Harmonisation [ICH]). This includes any occurrence that is new in onset or aggravated in severity from the baseline condition, or abnormal results of any diagnostic procedures that are conducted within clinical practice.

An adverse intervention reaction (AIR) is defined as a response to an intervention (eg, tACS application) that is noxious and unintended. The phrase “response to an intervention” means that a causal relationship between the intervention and an AE is at least a reasonable possibility. The phrase “a reasonable possibility” means that there are facts, evidence, or arguments to support a causal association with the intervention. An AIR, in contrast to an AE, is characterized by the fact that a causal relationship between the intervention and the occurrence is suspected. All AEs judged by either the reporting physician or the sponsor as having a reasonable causal relationship to an intervention qualify as AIRs.

The AEs expected from tACS are a very mild tingling sensation, fatigue, a light itching skin sensation under the stimulus, a mild headache, and dizziness, all of which disappear just after stimulation. With tACS, no persistent AEs have been reported; however, the AEs associated with repetitive tACS application remain unclear. In this type of stimulation technique, no serious adverse effects were reported. Some questionnaires may ask about stressful events in their life, which may cause them to feel discomfort or recall past experiences that upset them. If they find the questionnaire content distressing, they are encouraged to inform the study coordinator immediately or at a later time.

A serious adverse event (SAE), based on ICH and EU Guidelines on Pharmacovigilance for Medicinal Products for Human Use, is any untoward medical occurrence that at any dose (with extending the meaning to the tACS intervention):

- Results in death if life-threatening (the patient was at risk of death at the time of the event; it does not refer to an event that hypothetically might have caused death if it were more severe);
- Requires inpatient hospitalization or prolongation of existing hospitalization;
- Results in persistent or significant disability/incapacity if a congenital anomaly/birth defect, if a suspected transmission of any infectious agent via a therapeutic intervention, or if medically important.

Medical and scientific judgment should be exercised in determining whether other situations warrant consideration as serious, such as significant medical events that, although not immediately life-threatening or resulting in death or hospitalization, could jeopardize the patient or necessitate intervention to prevent one of the aforementioned outcomes.

A product quality complaint (PQC) is defined as any suspicion of a product defect related to manufacturing, labeling, or packaging, that is, any dissatisfaction relative to the identity, quality, durability, reliability, or performance of a distributed interventional product, including its labeling, delivery system, or package integrity. A PQC may have an impact on the safety and efficacy of the product. In addition, it includes any technical complaints, defined as any complaint that indicates a potential quality issue during manufacturing, packaging, release testing, stability monitoring, dose preparation, storage, or distribution of the product or the drug delivery system. Upon identification of a PQC, investigators are encouraged to report it to the marketing authorization holder.

SAEs and AEs of interest (skin irritation and local burning sensation) are to be recorded in the participant’s source records. The collection of these events should start after the informed consent form is signed and will continue until the end of the observation period of the study. Safety evaluations will also include the collection of concomitant medications, as well as the collection of vitals. In addition, the suicidal ideation will be assessed using the Columbia Suicide Severity Rating Scale (C-SSRS) at baseline (exclusion criteria).

All SAEs assessed by the participating physician should be recorded in the participant’s source documents. For reports of hospitalization, it is the sign, symptom, or diagnosis which led to hospitalization that is the serious event for which details must be provided. Any event requiring hospitalization (or prolongation of hospitalization) that occurs during the study must be reported as an SAE, except hospitalizations for the following: (1) hospitalizations not intended to treat an acute illness or AE (eg, social reasons such as pending placement in a long-term care facility); (2) surgery or procedures planned before entry into the study (should be documented in the case report form [CRF]) [note: hospitalizations that were planned before the start of data collection, and where the underlying condition for which the hospitalization was planned has not worsened, will not be considered as SAEs]; (3) any AE that results in a prolongation of the originally planned hospitalization is to be reported as a new SAE; and (4) the cause of death of a participant in a study, whether the event is expected or associated with the intervention, is considered as an SAE.

Pregnancy

Pregnancy here is no contraindication of tACS or CBTp during pregnancy. However, a participant in this special circumstance could have a different response to the intervention, influencing our final results. If a participant becomes pregnant during the study, data specified in the “data collection schedule” that is collected as part of the participant’s standard of care will

XSL•FO
RenderX

continue to be recorded in the CRF for the applicable time points, but the participant will be withdrawn from the study.

Statistical Considerations: Sample Size

This is a pilot randomized study with 28 participants (14 per arm). The primary goals are to evaluate recruitment feasibility, acceptability, and safety, and to produce unbiased estimates of the standard deviation and preliminary effect sizes of the primary outcome to inform sample-size calculation for a future definitive trial [31]. A target of 28 participants is consistent with common pilot rules-of-thumb (approximately 12-20 per arm) and provides sufficient precision to estimate recruitment and retention rates and the outcome variance. This pilot is not powered for hypothesis testing of efficacy (with n=14 per group, the minimal detectable standardized effect is large, ~d=1.1); outcome analyses will therefore be descriptive and will report effect estimates with 95% CIs as inputs for future sample-size calculations. To allow for expected attrition (approximately 15%-20%), we will recruit up to 33-40 participants. The generalizability of our findings is limited by the small sample size (n=28) and recruitment from a single treatment site. Patients in this cohort may differ from the broader schizophrenia population with respect to illness chronicity, symptom severity, comorbid psychiatric or medical conditions, and sociodemographic characteristics. In addition, the specialized treatment context may not reflect the experiences of individuals receiving care in community or primary care settings, where clinical resources and treatment approaches may vary. These factors restrict the extent to which the present results can be applied to other patient populations. Replication in larger, multisite, and more heterogeneous cohorts will be essential to establish the robustness and external validity of the findings.

The statistical analysis plan will use SPSS (version 29, IBM Corp) for the main tests, including effect size (Cohen d) and a 95% CI: repeated measures, mixed design analysis of variance, and t tests for two-group comparisons. Sex and gender differences will be examined using regression models that include biological sex, BSRI masculinity and femininity scores, and their interaction terms, with exploratory subgroup analyses based on BSRI classifications.

Statistical Outcomes

Primary Outcome Measures

EEG gamma oscillations are fast neural rhythms linked to higher-order cognitive functions (attention, working memory, executive control, and consciousness). In the prefrontal cortex, gamma EEG activity is associated with working memory maintenance, cognitive flexibility, top-down attention control, and integrating information across cortical areas. Altered prefrontal gamma observed in schizophrenia. EEG measurement of gamma activity in the prefrontal cortex refers to detecting brain oscillations in the gamma frequency band (typically 30-100 Hz, often subdivided into low gamma: 30-50 Hz, and high gamma: 60-100 Hz) from scalp electrodes positioned over prefrontal areas. In a within-subject design, we will compare gamma power across time windows (pre- and postintervention). In a between-subject design, we use mixed models or t tests on summary metrics (mean power in time×freq window).

PANSS is a standardized clinical interview that rates the presence and severity of positive and negative symptoms and general psychopathology for people with schizophrenia within the past week [32]. The scale consists of 30 items: 7 positive, 7 negative, and 16 general psychopathology symptoms. The symptom severity for each item is rated according to a 7-point scale (1=absent to 7=extreme) to best describe the symptom's presentation. The clinical interview takes approximately 45 minutes. The patient is rated from 1 to 7 on 30 different symptoms based on the discussion and reports of family members or primary care hospital workers. As 1 is the lowest score for each item, a patient cannot score lower than 30 for the total PANSS score. Scores are often given separately for positive items, negative items, and general psychopathology. [Time frame: PANSS will be applied at visit 1, screening (week 0); visit 4, end of intervention (week 17); visit 5, follow-up 1 (week 21); and visit 6, follow-up 2 (week 29).]

The NSA-16 assesses the presence, severity, and range of negative symptoms associated with schizophrenia [33]. For rapid clinical assessment or screening of patients for negative symptoms, the NSA-4 was developed as a simplified version, retaining only 4 of the 16 items: restricted speech quantity, reduced emotion, reduced social drive, and reduced interests. Each of the 4 items and the overall global negative symptoms is rated on a 1- to 6-point scale, where "1" represents no reduction from normal behaviors associated with the item and "6" represents a severe reduction in or absence of the behavior, with markedly impaired functionality. With respect to overall accuracy and predictive validity, the NSA-4 is comparable to the NSA-16. [Time frame: NSA-16 will be applied at visit 2, baseline (week 1); visit 4, end of intervention (week 17); visit 5, follow-up 1 (week 21); and visit 6, follow-up 2 (week 29).]

The CFS assesses cognitive flexibility as a factor that facilitates the individual's adaptability toward events [34]. Cognitive flexibility refers to a person's awareness of the existence of different communication styles in every situation, as well as their willingness and self-efficacy to employ these styles. The scale has 12 items pertaining to the 3 elements of cognitive flexibility: a person's awareness of alternative communicative styles, willingness to be communicatively flexible, and self-efficacy in being communicatively flexible. The scale uses a 6-point Likert-type scale for each item (with 6 as "strongly agree," 5 as "agree," 4 as "slightly agree," 3 as "slightly disagree," 2 as "disagree," and 1 as "strongly disagree"). The total score was used to calculate a global score ranging from 12 to 72, with questions 2, 9, 11, and 12 being reverse-coded. A higher score indicates higher cognitive flexibility. [Time frame: CFS will be applied at visit 2, baseline (week 1); visit 3, midpoint of intervention (week 9); visit 4, end of intervention (week 17); visit 5, follow-up 1 (week 21); and visit 6, follow-up 2 (week 29).]

Secondary Outcome Measures

HRV refers to the variation in time between consecutive heartbeats, a normal physiological phenomenon indicating the body's ability to adapt to stress and demands. There is no single "ideal" HRV; a higher HRV generally indicates better cardiovascular health and fitness, with a normal range

XSL•FO
RenderX

considered to be between 40 and 100 milliseconds. HRV provides insights into the body's autonomic nervous system balance and overall psychological health. It reflects the variations in time intervals between heartbeats, indicating how well the body adapts to stress and recovers. However, individual HRV levels vary based on age, fitness level, and other factors, and tracking baseline HRV and trends for an individual is more important than comparing to others. [Time frame: HRV will be determined at visit 2, week 1, before and immediately after the first intervention (tACS/CBTp); it will also be determined pre- and post-16th intervention, visit 4, week 17.]

The World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0) measures the level of functioning in 6 domains of life [35,36]: cognition—understanding and communicating; mobility—moving and getting around; self-care—attending to one's hygiene, dressing, eating, and staying alone; getting along—interacting with other people; life activities—domestic responsibilities, leisure, work, and school; participation—joining in community activities and participating in society. The 36-item WHODAS 2.0 is the most detailed version (questions about the difficulties experienced in the past 30 days). It allows users to generate scores for the 6 functioning domains and calculate an overall functioning score. The average interview time for the interviewer-administered 36-item version is 20 minutes. The simple sum of the scores of the items across all domains constitutes a statistic that describes the degree of functional limitations. The scores range from 0 to 100, where higher scores indicate greater disability. [Time frame: WHODAS 2.0 will be performed at visit 2, baseline (week 1); visit 4, end of intervention (week 17); visit 5, follow-up 1 (week 21); and visit 6, follow-up 2 (week 29).]

The Quality of Life Enjoyment and Satisfaction Questionnaire—Short Form (Q-LES-Q-SF) is a 16-item, participant-scored survey used to quantify changes in quality of life [37]. The Q-LES-Q-SF evaluates general activities that are assessed in the longer form: physical health, feelings, work, household duties, school/course work, leisure time activities, and social relations. Participants rate their satisfaction using a 5-point scale ranging from 1 (very poor) to 5 (very good). A total score is derived from 14 items, with a maximum score of 70, higher scores indicating greater life satisfaction and enjoyment. [Time frame: Q-LES-Q-SF will be performed at visit 2, baseline (week 1); visit 3, midway intervention (week 9); visit 4, end of intervention (week 17); visit 5, follow-up 1 (week 21); and visit 6, follow-up 2 (week 29).]

The Brief Betrayal Trauma Survey is a semistructured interview that measures 14 potentially traumatic events involving mistreatment by someone close, not so close, and noninterpersonal events [38]. Respondents indicate how frequently they experienced an event before and after 18 years of age. It categorizes items into high, medium, and low betrayal. (Time frame: The Brief Betrayal Trauma Survey will be performed at visit 2, baseline.)

The Perceived Stress Scale (PSS) is one of the most widely used psychological instruments for measuring stress perception [39]. It measures the degree to which situations in one's life are perceived as stressful. The 10 items inquire about how unpredictable, uncontrollable, and overloaded respondents find their lives, as well as the current levels of stress they experience. The questions in the PSS ask about feelings and thoughts concerning experiences during the last month. In each item, respondents are asked how often they felt a certain way. Individual scores on the PSS can range from 0 to 40, with higher scores indicating higher perceived stress. [Time frame: The PSS will be performed at visit 2, baseline (week 1); visit 4, end of intervention (week 17); visit 5, follow-up 1 (week 21); and visit 6, follow-up 2 (week 29).]

PETiT is a self-administered scale, which is user-friendly and sensitive to changes associated with treatment over time [26]. Consisting of 30 items, it assesses the respondent's health over the last week and provides 1 of 3 responses ("often," "sometimes," or "never"). The scale assesses 2 highly relevant domains for schizophrenia: adherence-related attitude (includes 6 items reflecting adherence and feelings toward medication) and psychosocial functioning (24 items describing patient characteristics such as clarity, energy, concentration, functioning, sex drive, and memory). The PETiT total score ranges from 0 to 60, with higher scores indicating better patient health-related quality of life. [Time frame: PETiT will be performed at visit 2, baseline (week 1); visit 4, end of intervention (week 17); visit 5, follow-up 1 (week 21); and visit 6, follow-up 2 (week 29).]

The Working Alliance Inventory—Short Revised (WAI-SR), both forms, that is, for client and therapist, measures 3 critical aspects of the therapeutic alliance: agreement on the tasks of therapy, agreement on the goals of therapy, and development of an affective bond [40]. The 12 items are rated on a 5-point Likert scale ranging from "1=never" to "5=always." The scores range from 5 to 20, with higher scores indicating a better therapeutic alliance. Completing the WAI-SR takes about 5 minutes. [Time frame: WAI-SR will be tested at visit 3, midterm through intervention (week 9); and visit 4, end of intervention (week 17).]

The BSRI is a psychological tool used to measure an individual's perception of their own masculinity and femininity, and it is used to research gender roles, with the understanding that individuals can express both masculine and feminine traits. The interview has 12 items, with scores on a 1 (never or almost never true) to 7 (always or always true) scale, with an average calculated for masculine and feminine traits, and then categorizes them as masculine, feminine, androgynous, or undifferentiated based on those scores. Each item is assigned a category as follows: M for masculine, F for feminine, and N for neutral. [Time frame: BSRI is performed at visit 1, screening (week 0).]

The Stroop Color and Word Test (hereafter just Stroop test), developed by John Ridley Stroop, is a neuropsychological test that measures cognitive interference by asking participants to name the ink color of words, which are printed in a color that conflicts with the word itself. The Stroop scores are based on 3 factors—color naming, interference, and speed—and converted into a *T*-score for interpretation. A *T*-score of 40 or less is considered low, while a score above 40 is considered normal. [Time frame: The Stroop test will be applied at visit 2,

XSL•FO
RenderX

baseline (week 0); at visit 4, end of intervention (week 17); and at the 2 follow-ups (week 21 and week 29).]

Other Prespecified Outcome Measures: Eligibility Screening Measures

The C-SSRS is a questionnaire developed by multiple institutions, including Columbia University, to help establish a person's immediate risk of suicide [41]. Its questions are phrased for use in an interview format but can be completed as a self-report measure if necessary. The C-SSRS measures 4 constructs: the severity of ideation, the intensity of ideation, behavior, and lethality. It includes "stem questions," which, if endorsed, prompt additional follow-up questions to obtain more information. The C-SSRS comprises 10 categories, all of which maintain binary responses (yes/no) to indicate the presence or absence of the behavior. The outcome of the C-SSRS is a numerical score with no specified clinical cutoffs due to the binary nature of the responses to items. Ultimately, interpretation is derived from a thorough clinical assessment, client history, and clinical expertise. [Time frame: at screening, visit 1 (week 0).]

The MoCA is a popular tool used in clinical practice when assessing for mild cognitive dysfunction. This instrument evaluates various cognitive domains, including attention, executive functions, memory, language, visuoconstructional skills, conceptual thinking, calculations, and orientation. Administering the MoCA takes around 10 minutes, and the total possible score is 30 points. A score of 26 or higher is considered normal, while a final score below 26 suggests mild cognitive impairment. [Time frame: MoCA will be performed at visit 1, screening (week 0), for participants over the age of 65 years (to exclude a possible cognitive decline).]

Study Oversight and Data Management

Participant data will be collected as follows: electronic medical record review, wearable device data (EEG and HRV), electronic patient-reported outcomes (Stroop scores and BSRI) and patient-reported outcomes (in papers), paper clinician-administered scales and assessments, and study visit worksheets. The study will use a Microsoft Excel clinical database for electronic data capture. A change log will be maintained as an audit trail for this database.

Even with data security protections in place, there is a risk that participant information could be released by accident. In case of a confidentiality breach, despite the safeguards in place, the investigators will immediately take steps to contain it, notify the participant, investigate, and remediate. Only personnel with designated hospital appointments (Departmental Assistant Status or Departmental Research Assistant Status at Providence Care Hospital) or physicians at Providence Care Hospital who are part of the research team will have access to any personal health information. The access will be necessary to corroborate information given during interviewer-administered interviews linked with questionnaires. Access to medical records and/or study data will be limited to authorized personnel. Electronic data will be stored on a hospital or other institutional network with firewalls and other security and backup measures in place. Data stored on laptops or mobile devices will be encrypted. Paper copies of study data will be stored in locked filing cabinets in a secure location. The principal investigator and their delegate will have access to data. Data will be stored for 15 years under participant ID numbers (anonymized) on the electronic data, which will be stored on a hospital or other institutional network with firewalls and other security and backup measures in place. Any data that exists in an identifiable form (name) will be converted and replaced with participant ID immediately, at the first opportunity. A master list of code-linking identifiers will only be held by the principal investigator or delegate for 15 years after the end of the study. Data collected on paper, except the consent form, will be anonymized. Data collected on paper will be stored in the archive, in locked filing cabinets at Providence Care Hospital. The data will be kept for 15 years in a deidentified format in the archive. Data will be deposited on the Queen's University Dataverse Collection in Borealis (a library-managed collection on a national platform).

Direct access to records will be provided to sponsor monitors as required and outlined in the study monitoring plan.

Monitoring Plan

Monitoring will be performed on-site by the sponsor's monitor in accordance with the study monitoring plan.

Trial monitoring is one of the principal quality control activities of the sponsor-investigator. Monitoring aims to ensure the participants' rights, safety, and well-being, and the reliability of trial results as the trial progresses. Monitoring also verifies that the trial is conducted in compliance with the current research ethics board (REB)–approved protocol, TCPS 2 (Tri-Council Policy Statement 2), and ICH Guideline for Good Clinical Practice E6 and determines appropriate corrective and preventive actions when noncompliance is identified.

The sponsor-investigator has assigned a monitor that is not involved in the clinical conduct of the trial at the site and has completed training on this monitoring plan, the current versions of ICH Guideline for Good Clinical Practice E6 and TCPS 2, and the REB-approved protocol. All monitoring of this trial will be done on-site at Providence Care Hospital.

The monitor will assess investigator and study team qualification, training, and delegation of duties; trial master file and essential records (including standard of care); participant consent; appropriate and timely reporting of AEs, unanticipated problems, and SAEs; randomization process and maintenance of the blind; participant recruitment and retention rates (screening, enrollment logs, and withdrawals); data and device storage and security; device accountability, maintenance, and calibration; clinical trial data, including eligibility; protocol compliance, including endpoints, procedures, assessments, and interventions; source data verification; deviation or violation recording and reporting; and noncompliance and data trends.

The monitor will use standardized templates to review the above and document observations and findings. All observations will be discussed at the end of the on-site review with the sponsor-investigator and study team members. The monitor will review the observations, and all verified and confirmed findings are documented in a modular monitoring report. These reports will be provided to the sponsor and to the Queen's

XSL•FO
RenderX

Vice-Principal of Research Compliance and Training team to assist with addressing noncompliance and identification of trends. Sites will receive a letter outlining findings that require corrective and preventive action (CAPA). Monitors will work with the site to conduct a root cause analysis and formulate the CAPA.

The monitor will perform the first on-site review after the first participant has completed 2 weeks of treatment (study week 4). The second visit will be scheduled 5 weeks later, when the first patient has undergone the first set of on-treatment assessments. If no critical/significant or major findings or deviations have been identified, thereafter the monitor will conduct on-site reviews every 2 months while patients remain on active treatment (assuming a recruitment rate of >5 at the beginning of the study and lower thereafter). More frequent visits may be indicated for higher recruitment rates, rates of protocol deviations, or safety reporting.

As this study was assessed as being medium risk and will use monitoring tools from Network of Networks (N2), 100% source data verification of the following will be done for the first two participants randomized that receive intervention: informed consent process, eligibility, safety (AEs and SAEs), intervention data, endpoints (primary and secondary outcome measures), and protocol compliance. Propose thereafter an additional 25% (7/28) of participants and review for each baseline/eligibility at least one treatment period, end of study, one follow-up, and all endpoints or events (SAEs, death, etc). Additional sources of data verification and visits may be added based on significant error rates or noncompliance. The study monitoring plan will be reviewed annually to ensure that the processes and procedures outlined herein remain applicable to the conduct of the study. Additionally, if amendments are made to the protocol that affect the study procedures and/or participant safety, a review of the monitoring plan will be performed to revise the plan as needed.

Ethical Considerations

This study adheres to the principles outlined in the Tri-Council Policy Statement (2022) and ICH Guideline for Good Clinical Practice E6. It complies with the REB guidelines, involving human participants, medical records, patient information, and observations of public behaviors. The study protocol, along with required updated templates and the letter of consent and information, was reviewed and approved by Queen's University, Health Sciences Research Ethics Board (HSREB), file number PSYC-222-23/6038935/6044118 TRAQ DSS 6036290 (last approval from July 2, 2025), before any research activities commenced. All study participants receive the HSREB-approved detailed consent form, which explains the study and provides enough information for informed decision-making regarding their participation; consent for the involvement is obtained before any study procedures are conducted and is signed by the participant, the person who conducted the informed consent discussion, and the principal investigator. For secondary analyses using existing data with primary consent, the original consent HSREB approval covers secondary analysis without additional consent.

Risks of Participation

There are risks to taking part in any clinical study. Risks associated with tACS are generally mild and disappear after the stimulation. With tACS, no persistent AEs have been reported; however, the AEs related to repetitive tACS application remain unclear. In this type of stimulation technique, no serious adverse effects were reported. Some of the side effects that may occur during this study can be treated. The AEs are a very mild tingling sensation, fatigue, a light itching sensation under the stimulus, mild headache, and dizziness, all of which disappear just after stimulation.

Participants in one of the two intervention groups will not receive the current stimulation but a placebo/sham tACS (the tACS is a placebo, but the CBTp is the intervention). Without receiving the low electric stimulation that may augment their response to psychotherapy (CBTp), their condition might not improve. The previous studies using only CBTp as an intervention reported improvement of the illness (they will receive CBTp anyway). However, these study participants will continue their usual treatments, including antipsychotic treatment.

While generally safe, CBTp can have some risks, primarily during initial sessions or specific techniques such as exposure therapy. These risks can include short-term stress, anxiety, or emotional distress while exploring painful feelings and experiences. There is limited scientific research on the possible side effects of psychotherapy, including CBTp. The risks of CBTp are generally low and can be managed with open communication and the support of a skilled therapist.

Some questionnaires used in this study may ask about stressful events in a participant's life, which may cause them to feel discomfort or recall past experiences that may have upset them. If they find the questionnaire content distressing, they are encouraged to inform the study coordinator immediately or at a later time. A member of the research team can discuss this with them and refer this to a member of their care team whom they identify and consent to (the investigators will ask them for written consent to disclose this information if they decide). The participant's name will never be placed on a questionnaire to maintain their privacy and protect their confidentiality. If they write their name by accident, a sticker containing their participant ID will be placed over it to preserve their privacy. The participants may also feel frustrated or confused by the language in the questionnaires. If investigators need clarification on any points in the written material or explanations, the research team will be happy to assist and provide further details to their satisfaction.

The research team will manage the plan to mitigate psychological and emotional risks. To achieve this, interviewers will be trained to administer questionnaires thoughtfully, sensitively, and with awareness of the types of questions being asked, as well as the potential effects and thought processes participants may experience. These are essential strategies to ensure sensitive and polite communication, helping researchers learn to interpret verbal and nonverbal cues that may convey discomfort.

XSL•FO
RenderX

Privacy and Confidentiality

Participant names will remain confidential and not be disclosed to those not involved in the study. In the CRF, only the participant number will be recorded. If any identifiers appear on any document, they will be redacted in copies stored in the site master file or presented for audit. Data stored electronically will comply with local data protection laws. Participants will be informed that ethics committees or research ethics boards may inspect their records and that all personal information will be treated with utmost confidentiality in line with local data protection laws. The investigator will keep a list linking participant numbers to names for record identification and retrieval.

EEG sample recordings will be stored in a file on the designated laptop, where they will remain anonymous, identified only by a code number. Postrecording, the samples will be transferred to a secure server at Queen's University, anonymized, and then deleted from the original recorder. These anonymized files will be stored on the server for 15 years after data collection ends, after which they will be permanently deleted.

Any chart reviews will be conducted by hospital-appointed individuals who have the privilege to access them.

Results

Duration of Participation

The timeline for recruitment, treatment, and follow-up is 18 months, followed by 6 months for data analysis, writing manuscripts, and dissemination activities.

Total duration of participation in this research will be 29 weeks; the visits, procedures, and timelines are represented in Table 2. In addition to the study visits, participants will be seen weekly for 16 weeks to receive CBTp with either tACS or sham tACS, and there will be a booster session intervention during each of the two follow-up visits.

XSL•FO
RenderX

Table 2. Study calendar, visits, procedures, and timelines[a].

Study phases	Phase I: Preintervention		Phase II: Intervention		Phase III: Postintervention	
Timeline	Visit 1	Visit 2	Visit 3	Visit 4	Visit 5	Visit 6
Procedure	Screening	Baseline	Session 8	Session 16	Follow-up	Follow-up
Duration	Week 0	Week 1	Week 9	Week 17	Week 21	Week 29
Informed consent, 60 minutes	✓					
Chart review, 60 minutes	✓					
Height and weight, 10 minutes	✓					
Vital signs, 15 minutes	✓	✓	✓	✓	✓	✓
Physical exam, 20 minutes	✓					
MoCA[b]/age >65 years, 20 minutes	✓					
C-SSRS[c], 10 minutes	✓					
Bem Sex-Role Inventory, 15 minutes	✓					
Chart review, 30 minutes	✓					
Physical exam, 20 minutes	✓					
Vitals, 10 minutes	✓	✓	✓	✓	✓	✓
PANSS[d], 40 minutes	✓			✓	✓	✓
NSA-16[e], 15 minutes		✓		✓	✓	✓
WHODAS 2.0[f], 10 minutes		✓		✓	✓	✓
CFS[g] and PSS[h], 10 minutes		✓	✓	✓	✓	✓
PETiT[i], 10 minutes		✓	✓	✓	✓	✓
Q-LES-Q-SF[j], 5 minutes		✓	✓	✓	✓	✓
BBTS[k], 10 minutes		✓				
WAI-SR[l], 5 minutes			✓	✓		
Stroop[m], 10 minutes		✓		✓	✓	✓
EEG[n], 30 minutes including prep.	✓			✓		
HRV[o] (beginning and end of session), 5 minutes		✓		✓		

[a]Benzodiazepines and other medications for sleep or anxiety must be held for at least 8 hours before assessments. This questionnaire/assessment must be done at approximately the same time of day each time. If these cannot be done during a scheduled visit, the participant may be rescheduled for another day. Participants will be allowed to take short breaks during these assessments if they need them.

[b]MoCA: Montreal Cognitive Assessment.

[c]C-SSRS: Columbia Suicide Severity Rating Scale.

[d]PANSS: Positive and Negative Syndrome Scale.

[e]NSA-16: 16-item Negative Symptom Assessment.

[f]WHODAS 2.0: World Health Organization Disability Assessment Schedule 2.0.

[g]CFS: Cognitive Flexibility Scale.

[h]PSS: Perceived Stress Scale.

[i]PETiT: Personal Evaluation of Transitions in Treatment.

[j]Q-LES-Q-SF: Quality of Life Enjoyment and Satisfaction Questionnaire—Short Form.

[k]BBTS: Brief Betrayal Trauma Survey.

[l]WAI-SR: Working Alliance Inventory—Short Revised.

[m]Stroop: Stroop Color and Word Test.

XSL•FO
RenderX

[n]EEG: electroencephalogram.

[o]HRV: heart rate variability.

Current State of the Research

The enrollments for this pilot clinical trial are in September 2025. By November 1, 2025, we have enrolled 15 participants. Two of them were screening failures, and one withdrew from the study after intervention session 2—the reasons for withdrawal were not related to our research (they started a full-time job and were no longer able to attend the weekly sessions). The 12 participants who remained in the study are following the intervention protocol, with 8 receiving both active tACS/CBTp and 4 receiving sham tACS/CBTp. The gender distribution of the screening process so far is 7 females, 4 males, and 1 transgender female, with 7 of them under the age of 40 years and 7 older than that. Two male participants were screened for failure, and one female left the study.

Discussion

Expectations

Based on well-established evidence of cognitive deficits in schizophrenia, as well as our own observations in the CBTp Lab, an additional weakness in current schizophrenia treatment protocols was identified. Aside from challenges in therapeutic engagement, patients struggle to retain progress made in individual sessions. This is because the cognitive model-based approach requires implementing actual cognitive and behavioral changes—the “action plan/homework”—without which the therapeutic efforts of both the therapist and patient become futile. Current guidelines, suggesting an hour-long session of CBTp weekly for at least 16 weeks, become insufficient in the context of a complex symptomatology imbued with negative symptoms. A need for a much more concerted effort to obtain significant improvement becomes clear if we were to close the gap of disadvantages that schizophrenia patients are confronted with.

This study hypothesizes that tACS applied at the beginning of the CBTp intervention in people with schizophrenia will predict better clinical and functional responses to therapy (compared with the sham tACS/CBTp).

The expectations are that CBTp will be more efficient in the participants undergoing active tACS compared with sham tACS. Considering that cognitive and emotional status are gender dependent (hormonal differences, brain structure, and sociocultural influences), we expect that the therapeutic response could be gender specific [42,43].

The impact of this research is represented by clinical and social benefits, contributions to the scientific literature, and support for new research applications.

Potential Clinical Benefit

A total of 28 clients will have the opportunity to benefit from our CBTp intervention with or without tACS.

Social Benefits

The new type of intervention will allow for personalized and highly impactful treatment. This will lead to quality-of-life improvement for participants as well as for support members (family and caregivers) and will reduce health care costs.

Disseminations

Aside from the current published protocol and the first scoping review on this topic [18], the investigators anticipate publishing up to 5 new peer-reviewed publications and presenting the results at different scientific events. Two graduate students are team members of this research, and by the end of 2026, they will write their master's theses from the preliminary data of this clinical trial. This pilot study aims to test and redefine our research methods, procedures, and feasibility of the main study. Based on the preliminary/interim analysis, we intend to apply for funding to further expand on the correlation between specific frequency brain stimulation and clinical outcomes in schizophrenia. Adding brain imaging evidence will help us to prove the efficacy of our new intervention protocol.

Limitations

Limitations to the pilot studies include small sample sizes, resulting in less generalizable results. To our knowledge, this intervention protocol, tACS/CBTp, has never been used before in people living with schizophrenia. This study would not only be the first of its kind to highlight the potential capabilities of this therapy (associated with regular antipsychotics) in treating cognitive and negative symptoms, but also to expand the scope of pre-existing literature surrounding the augmentation of the CBTp efficacy.

Further Directions

Replication of this research protocol with a larger sample (N≥30) would allow for greater specificity and rigor in distinguishing predictors, suggesting clear selection criteria for tACS/CBTp candidates.

Conclusions

This study will help us develop a new interventional protocol that could impact the outcome of schizophrenia at personal, familial, and social levels, reducing the rate of relapse and the costs of hospitalization.

Funding

This research was awarded US $20,000 through the Internal Grant Competition 2022 by the Department of Psychiatry at Queen’s University.

XSL•FO
RenderX

Authors' Contributions

FI and AF developed the original protocol for this study. ER developed the monitoring plan. AT supported the entire research work, advising, rewriting, polishing, designing templates, and teaching. All authors helped to finalize and update this version of the protocol. All authors have read and approved this manuscript.

Conflicts of Interest

None declared.

References

1. Lee KH, Williams LM, Haig A, Gordon E. "Gamma (40 Hz) phase synchronicity" and symptom dimensions in schizophrenia. Cogn Neuropsychiatry 2003;8(1):57-71. [doi: 10.1080/713752240] [Medline: 16571550]
2. Higgins E, George MS. The Neuroscience of Clinical Psychiatry. 3rd ed. New York, NY: Wolters Kluwer; 2018.
3. Turner DT, Burger S, Smit F, Valmaggia LR, van der Gaag M. What constitutes sufficient evidence for case formulation-driven CBT for psychosis? Cumulative meta-analysis of the effect on Hallucinations and delusions. Schizophr Bull 2020;46(5):1072-1085 [FREE Full text] [doi: 10.1093/schbul/sbaa045] [Medline: 32221536]
4. Best MW, Milanovic M, Iftene F, Bowie CR. A randomized controlled trial of executive functioning training compared with perceptual training for schizophrenia spectrum disorders: effects on neurophysiology, neurocognition, and functioning. Am J Psychiatry 2019;176(4):297-306. [doi: 10.1176/appi.ajp.2018.18070849] [Medline: 30845819]
5. Farcas A, Iftene F. Findings, limitations and new directions in tACS studies in schizophrenia research: a scoping review. J Psychiatr Res 2022;151:291-298 [FREE Full text] [doi: 10.1016/j.jpsychires.2022.04.036] [Medline: 35525231]
6. Symond MB, Harris AWF, Gordon E, Williams LM. "Gamma synchrony" in first-episode schizophrenia: a disorder of temporal connectivity? Am J Psychiatry 2005;162(3):459-465. [doi: 10.1176/appi.ajp.162.3.459] [Medline: 15741462]
7. Uhlhaas PJ, Haenschel C, Nikolić D, Singer W. The role of oscillations and synchrony in cortical networks and their putative relevance for the pathophysiology of schizophrenia. Schizophr Bull 2008;34(5):927-943 [FREE Full text] [doi: 10.1093/schbul/sbn062] [Medline: 18562344]
8. Jones M, McDermott B, Oliveira BL, O'Brien A, Coogan D, Lang M, et al. Gamma band light stimulation in human case Studies: groundwork for potential Alzheimer's disease treatment. J Alzheimers Dis 2019;70(1):171-185. [doi: 10.3233/jad-190299]
9. Santarnecchi E, Sprugnoli G, Tatti E, Mencarelli L, Neri F, Momi D, et al. Brain functional connectivity correlates of coping styles. Cogn Affect Behav Neurosci 2018;18(3):495-508. [doi: 10.3758/s13415-018-0583-7] [Medline: 29572771]
10. Santarnecchi E, Brem AK, Levenbaum E, Thompson T, Kadosh RC, Pascual-Leone A. Enhancing cognition using transcranial electrical stimulation. Curr Opin Behav Sci 2015;4:171-178. [doi: 10.1016/j.cobeha.2015.06.003]
11. Benussi A, Cantoni V, Cotelli MS, Cotelli M, Brattini C, Datta A, et al. Exposure to gamma tACS in Alzheimer's disease: a randomized, double-blind, sham-controlled, crossover, pilot study. Brain Stimul 2021;14(3):531-540 [FREE Full text] [doi: 10.1016/j.brs.2021.03.007] [Medline: 33762220]
12. Diedrich L, Kolhoff HI, Bergmann C, Bähr M, Antal A. Boosting working memory in the elderly: driving prefrontal theta-gamma coupling via repeated neuromodulation. Geroscience 2025;47(2):1425-1440. [doi: 10.1007/s11357-024-01272-3] [Medline: 38992335]
13. Diedrich L, Kolhoff HI, Bergmann C, Chakraborty S, Antal A. Theta-gamma tACS modulates attention network synchronization, not isolated network performance. Brain Res 2025;1855:149550 [FREE Full text] [doi: 10.1016/j.brainres.2025.149550] [Medline: 40086742]
14. Fertonani A, Miniussi C. Transcranial electrical stimulation: what we know and do not know about mechanisms. Neuroscientist 2017;23(2):109-123 [FREE Full text] [doi: 10.1177/1073858416631966] [Medline: 26873962]
15. Radman T, Ramos RL, Brumberg JC, Bikson M. Role of cortical cell type and morphology in subthreshold and suprathreshold uniform electric field stimulation in vitro. Brain Stimul 2009;2(4):215-228.e3 [FREE Full text] [doi: 10.1016/j.brs.2009.03.007] [Medline: 20161507]
16. Wang L, Chen Y, Piao Z, Gu X, Liu H, Wang D, et al. Medial parietal alpha-frequency transcranial alternating current stimulation for chronic insomnia: a randomized sham-controlled trial. Psychol Med 2025;55:e102. [doi: 10.1017/s0033291725000625]
17. Bland NS, Sale MV. Current challenges: the ups and downs of tACS. Exp Brain Res 2019;237(12):3071-3088. [doi: 10.1007/s00221-019-05666-0] [Medline: 31620829]
18. Doig GS, Simpson F. Randomization and allocation concealment: a practical guide for researchers. J Crit Care 2005;20(2):187-193. [doi: 10.1016/j.jcrc.2005.04.005] [Medline: 16139163]
19. Kasten FH, Dowsett J, Herrmann CS. Sustained aftereffect of α-tACS lasts up to 70 min after stimulation. Front Hum Neurosci 2016;10:1-12 [FREE Full text] [doi: 10.3389/fnhum.2016.00245] [Medline: 27252642]
20. Polese D, Fornaro M, Palermo M, De Luca V, de Bartolomeis A. Treatment-resistant to antipsychotics: a resistance to everything? Psychotherapy in treatment-resistant schizophrenia and nonaffective psychosis: a 25-year systematic review

XSL•FO
RenderX

and exploratory meta-analysis. Front Psychiatry 2019;10:210-218 [FREE Full text] [doi: 10.3389/fpsyt.2019.00210] [Medline: 31057434]
21. Bressler SL. Cortical coordination dynamics and the disorganization syndrome in schizophrenia. Neuropsychopharmacology 2003;28(Suppl 1):S35-S39. [doi: 10.1038/sj.npp.1300145] [Medline: 12827142]
22. Cho RY, Konecky RO, Carter CS. Impairments in frontal cortical gamma synchrony and cognitive control in schizophrenia. Proc Natl Acad Sci USA 2006;103(52):19878-19883 [FREE Full text] [doi: 10.1073/pnas.0609440103] [Medline: 17170134]
23. Tsuchimoto R, Kanba S, Hirano S, Oribe N, Ueno T, Hirano Y, et al. Reduced high and low frequency gamma synchronization in patients with chronic schizophrenia. Schizophr Res 2011;133(1-3):99-105. [doi: 10.1016/j.schres.2011.07.020] [Medline: 21849245]
24. Steinmann S, Leicht G, Andreou C, Polomac N, Mulert C. Auditory verbal hallucinations related to altered long-range synchrony of gamma-band oscillations. Sci Rep 2017;7(1):8401. [doi: 10.1038/s41598-017-09253-7] [Medline: 28827744]
25. Reato D, Rahman A, Bikson M, Parra LC. Effects of weak transcranial alternating current stimulation on brain activity—a review of known mechanisms from animal studies. Front Hum Neurosci 2013;7:687 [FREE Full text] [doi: 10.3389/fnhum.2013.00687] [Medline: 24167483]
26. Voruganti LN, Awad AG. Personal evaluation of transitions in treatment (PETiT): a scale to measure subjective aspects of antipsychotic drug therapy in schizophrenia. Schizophr Res 2002;56(1-2):37-46. [doi: 10.1016/s0920-9964(01)00161-x] [Medline: 12084418]
27. Elyamany O, Leicht G, Herrmann CS, Mulert C. Transcranial alternating current stimulation (tACS): from basic mechanisms towards first applications in psychiatry. Eur Arch Psychiatry Clin Neurosci 2021;271(1):135-156 [FREE Full text] [doi: 10.1007/s00406-020-01209-9] [Medline: 33211157]
28. Strüber D, Herrmann CS. Modulation of gamma oscillations as a possible therapeutic tool for neuropsychiatric diseases: a review and perspective. Int J Psychophysiol 2020;152:15-25. [doi: 10.1016/j.ijpsycho.2020.03.003] [Medline: 32240665]
29. Antal A, Paulus W. Transcranial alternating current stimulation (tACS). Front Hum Neurosci 2013;7:317 [FREE Full text] [doi: 10.3389/fnhum.2013.00317] [Medline: 23825454]
30. Vosskuhl J, Huster RJ, Herrmann CS. Increase in short-term memory capacity induced by down-regulating individual theta frequency via transcranial alternating current stimulation. Front Hum Neurosci 2015;9:257 [FREE Full text] [doi: 10.3389/fnhum.2015.00257] [Medline: 26005411]
31. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005;4(4):287-291. [doi: 10.1002/pst.185]
32. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull 1987;13(2):261-276. [doi: 10.1093/schbul/13.2.261] [Medline: 3616518]
33. Axelrod BN, Goldman RS, Alphs LD. Validation of the 16-item negative symptom assessment. J Psychiatr Res 1993;27(3):253-258. [doi: 10.1016/0022-3956(93)90036-2] [Medline: 7905033]
34. Martin MM, Rubin RB. A new measure of cognitive flexibility. Psychol Rep 1995;76(2):623-626. [doi: 10.2466/pr0.1995.76.2.623]
35. Gold LH. DSM-5 and the assessment of functioning: the World Health Organization Disability Assessment Schedule 2.0 (WHODAS 2.0). J Am Acad Psychiatry Law 2014;42(2):173-181. [Medline: 24986344]
36. Ustun TB, Kostanjsek N, Chatterji S, Rehm J, editors. Measuring Health and Disability: Manual for WHO Disability Assessment Schedule WHODAS 2.0. New York, NY: World Health Organization; 2010:88.
37. Pukrop R, Möller HJ, Steinmeyer EM. Quality of life in psychiatry: a systematic contribution to construct validation and the development of the integrative assessment tool "modular system for quality of life". Eur Arch Psychiatry Clin Neurosci 2000;250(3):120-132. [doi: 10.1007/s004060070028] [Medline: 10941986]
38. Goldberg LR, Freyd JJ. Self-reports of potentially traumatic experiences in an adult community sample: gender differences and test-retest stabilities of the items in a brief betrayal-trauma survey. J Trauma Dissociation 2006;7(3):39-63. [doi: 10.1300/j229v07n03_04]
39. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24(4):385-396. [doi: 10.2307/2136404]
40. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. J Couns Psychol 1989;36(2):223-233. [doi: 10.1037//0022-0167.36.2.223]
41. Posner K, Brown GK, Stanley B, Brent DA, Yershova KV, Oquendo MA, et al. The Columbia-Suicide Severity Rating Scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. Am J Psychiatry 2011;168(12):1266-1277 [FREE Full text] [doi: 10.1176/appi.ajp.2011.10111704] [Medline: 22193671]
42. Chaplin TM. Gender and emotion expression: a developmental contextual perspective. Emot Rev 2015;7(1):14-21 [FREE Full text] [doi: 10.1177/1754073914544408] [Medline: 26089983]
43. Howard LM, Ehrlich AM, Gamlen F, Oram S. Gender-neutral mental health research is sex and gender biased. Lancet Psychiatry 2017;4(1):9-11. [doi: 10.1016/s2215-0366(16)30209-7]

Abbreviations

AE: adverse event

XSL•FO
RenderX

AIR: adverse intervention reaction
BSRI: Bem Sex-Role Inventory
CAPA: corrective and preventive action
CBTp: cognitive behavioral therapy for psychosis
CFS: Cognitive Flexibility Scale
CRF: case report form
C-SSRS: Columbia Suicide Severity Rating Scale
DSM-5: Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition)
EEG: electroencephalogram
HRV: heart rate variability
HSREB: Health Sciences Research Ethics Board
ICH: International Council for Harmonisation
MoCA: Montreal Cognitive Assessment
NSA-16: 16-item Negative Symptom Assessment
PANSS: Positive and Negative Syndrome Scale
PETiT: Personal Evaluation of Transitions in Treatment
PQC: product quality complaint
PSS: Perceived Stress Scale
Q-LES-Q-SF: Quality of Life Enjoyment and Satisfaction Questionnaire—Short Form
REB: research ethics board
SAE: serious adverse event
tACS: transcranial alternating current stimulation
TCPS 2: Tri-Council Policy Statement 2
WAI-SR: Working Alliance Inventory—Short Revised
WHODAS 2.0: World Health Organization Disability Assessment Schedule 2.0

Edited by A Schwartz; submitted 13.Jul.2025; peer-reviewed by E Predescu, R Sipos; comments to author 08.Oct.2025; revised version received 09.Nov.2025; accepted 21.Nov.2025; published 10.Feb.2026.

Please cite as:
Iftene F, Farcas A, O'Brien S, Bowie C, Best M, Ayonrinde O, Landry T, Carlson J, Davidson SJ, Rodgerson E, Theis A
Therapeutic Improvement in People With Schizophrenia Undergoing tACS/CBTp (Transcranial Alternating Current Stimulation/Cognitive Behavioral Therapy for Psychosis) Associated With Usual Medication Regimen: Protocol for a Pilot, Randomized, Single-Blind Trial
JMIR Res Protoc 2026;15:e80593
URL: https://www.researchprotocols.org/2026/1/e80593
doi:10.2196/80593
PMID:41273718

©Felicia Iftene, Adriana Farcas, Simon O'Brien, Christopher Bowie, Michael Best, Oyedeji Ayonrinde, Terry Landry, Jennifer Carlson, Scott John Davidson, Ellie Rodgerson, Ashley Theis. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Flopropione, a Cysteine Conjugate β-Lyase 1 Inhibitor, for Prevention of Cisplatin-Induced Nephrotoxicity: Protocol for a Randomized, Open-Label, Proof-of-Concept Phase 1 and 2a Trial

Takenao Koseki[1,2,3*], PhD; Masashi Kondo[4,5*], MD, PhD; Hidetsugu Fujigaki[6], PhD; Kayoko Kikuchi[7], PhD; Yuko Oya[4], MD, PhD; Hiroshi Kato[5], PhD; Tomohiro Mizuno[1,3], PhD; Naotake Tsuboi[8], MD, PhD; Kenji Kawada[9], MD, PhD; Yasuhiro Goto[4], MD, PhD; Naozumi Hashimoto[4], MD, PhD; Kazuyoshi Imaizumi[4], MD, PhD; Akiko Kada[7], MPH; Hikaru Yabuuchi[10], PhD; Kuniaki Saito[6,10], PhD; Hideyuki Saya[11], MD, PhD

[1]Department of Pharmacy, Fujita Health University Hospital, Toyoake, Aichi, Japan

[2]Department of Regulatory Science for Evaluation and Development of Pharmaceuticals and Devices, Graduate School of Medical Sciences, Fujita Health University, Toyoake, Aichi, Japan

[3]Department of Pharmacotherapeutics and Informatics, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

[4]Department of Respiratory Medicine, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

[5]Department of Development and Education of Clinical Research, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

[6]Department of Advanced Diagnostic System Development, Graduate School of Medical Sciences, Fujita Health University, Toyoake, Aichi, Japan

[7]Center for Translational Research, Fujita Health University, Toyoake, Aichi, Japan

[8]Department of Nephrology, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

[9]Department of Medical Oncology, Fujita Health University School of Medicine, Toyoake, Aichi, Japan

[10]Center for Society-Academia Collaboration, Fujita Health University, Toyoake, Aichi, Japan

[11]Oncology Innovation Center, Fujita Health University, Toyoake, Aichi, Japan

[*]these authors contributed equally

Corresponding Author:
Takenao Koseki, PhD
Department of Regulatory Science for Evaluation and Development of Pharmaceuticals and Devices
Graduate School of Medical Sciences
Fujita Health University
1-98 Dengakugakubo, Kutsukake-cho
Toyoake, Aichi, 470-1192
Japan
Phone: 81 562 93 9396
Email: tkoseki@fujita-hu.ac.jp

Abstract

Background: Cisplatin-induced nephrotoxicity (CIN) is a major dose-limiting adverse event that can lead to both acute and chronic kidney injury. The formation of thiol-cisplatin conjugates within renal tubular cells has been implicated as a key mechanism underlying CIN. Flopropione is an inhibitor of cysteine conjugate β-lyase 1, an enzyme that catalyzes the formation of the thiol-cisplatin conjugate, which might prevent CIN.

Objective: We designed a clinical trial to evaluate the safety of flopropione in patients receiving cisplatin-based chemotherapy and explore its efficacy in preventing CIN.

Methods: This is a phase 1 and 2a, single-center, randomized, open-label trial conducted in patients undergoing cisplatin therapy. Participants are randomized in a 5:2 ratio per cohort to receive either flopropione or no treatment. On the day of cisplatin administration, the flopropione group receives oral flopropione twice daily (80 mg in cohort 1, 160 mg in cohort 2, and 240 mg in cohort 3). On the following day, all cohorts receive 3 doses of 80 mg of oral flopropione. A step-up dose escalation design is adopted, progressing from cohort 1 to 3 after confirming safety at each level. The primary end point is the safety of flopropione use in combination with cisplatin; the secondary end points include changes in the levels of urinary biomarkers of nephrotoxicity such as neutrophil gelatinase-associated lipocalin, liver-type fatty acid-binding protein, and kidney injury molecule-1. Blood and

XSL•FO
RenderX

urine samples are collected within 48 hours before cisplatin administration and at 24 hours, 48 hours, and 1 week after its initiation for safety and efficacy assessments.

Results: The first participant was registered in July 2024. As of January 2026, participant registration is ongoing. The final participant will complete the study by March 2026. Publication of results is expected by March 2027.

Conclusions: This study is expected to contribute to advances in preventive strategies for CIN by providing evidence that inhibition of cysteine conjugate β-lyase 1 by flopropione may attenuate CIN.

Trial Registration: Japan Registry of Clinical Trials jRCTs041220021; https://jrct.mhlw.go.jp/en-latest-detail/jRCTs041220021

International Registered Report Identifier (IRRID): DERR1-10.2196/87907

(JMIR Res Protoc 2026;15:e87907) doi:10.2196/87907

KEYWORDS

cisplatin; cysteine conjugate-β lyase 1; flopropione; kidney injury molecule-1; KIM-1; liver-type fatty acid-binding protein; L-FABP; nephrotoxicity; neutrophil gelatinase-associated lipocalin; N-GAL

Introduction

Background

Nearly 5 decades have passed since cisplatin was approved in the United States in 1978, and it remains one of the most widely used and effective anticancer drugs worldwide [1]. Cisplatin exerts potent antitumor effects primarily through the formation of DNA cross-links, thereby inhibiting DNA replication and transcription. However, because its cytotoxic activity is not selective for cancer cells, cisplatin damages normal tissues and causes adverse effects, such as myelosuppression, ototoxicity, neurotoxicity, and nephrotoxicity, which often limit its clinical use [2]. Among these toxicities, nephrotoxicity is particularly significant as it can result in both acute and chronic kidney injury [3]. Cisplatin-induced nephrotoxicity (CIN) reportedly occurs in approximately 30% of the patients receiving cisplatin-based chemotherapy [4,5].

Hydration therapy can reduce the intratubular concentration of cisplatin, thereby lowering renal toxicity, and it is recognized as the globally established preventive method for CIN [6,7]. Particularly, recently developed short hydration protocols have offered comparable protective effects to conventional hydration while reducing patient burden and improving safety [8,9]. In addition, magnesium supplementation is used for preventing hypomagnesemia and mitigating the subsequent aggravation of renal dysfunction [10,11]. Use of osmotic diuretics, such as mannitol, has also been reported as a preventive strategy; however, their use might increase risk of nausea, vomiting, and potential deterioration in renal function, particularly in patients with comorbid diabetes mellitus or hypertension [12]. Although emerging preclinical evidence in animal and in vitro studies suggests that antioxidant-based agents, including riociguat, JP4-039, and amentoflavone, exert renoprotective effects against CIN, robust clinical data supporting their efficacy and safety are still lacking [13-15]. Thus, in addition to preventive strategies such as hydration therapy and magnesium supplementation, the development of novel prophylactic approaches for CIN is warranted.

One proposed mechanism underlying CIN involves the formation of a highly reactive thiol-cisplatin complex catalyzed by cysteine conjugate-β lyase 1 (CCBL1) in renal tubular cells [16,17]. We previously reported that 2′,4′,6′-trihydroxyacetophenone, a CCBL1 inhibitor, dose dependently suppressed cisplatin-induced increases in blood urea nitrogen (BUN) and serum creatinine levels and tubular cell damage and apoptosis in a CIN mouse model [18]. Flopropione (2′,4′,6′-trihydroxypropiofenone), identified concurrently with 2′,4′,6′-trihydroxyacetophenone as a CCBL1 inhibitor, exhibits comparable inhibitory activity in renal tubular cells and animal models (Koseki, T, unpublished data, April 2022). Flopropione was approved in Japan in 1978 for its antispasmodic effects in hepatobiliary and pancreatic diseases and urinary calculi [19]. Given its known safety profile and potential to inhibit CCBL1, flopropione may also confer protective effects against CIN.

Therefore, for the first time, a phase 1 and 2a clinical trial has been designed for patients initiating cisplatin therapy to obtain proof of concept for the efficacy and safety of flopropione in preventing CIN.

Objective

This phase 1 and 2a clinical trial aims to primarily evaluate the safety of flopropione in patients receiving cisplatin-based chemotherapy while exploring preliminary efficacy and generating proof-of-concept evidence for its potential protective effect against CIN.

Methods

Study Design and Setting

This is a single-center, randomized, open-label trial conducted at Fujita Health University Hospital (Toyoake, Japan) in patients undergoing cisplatin therapy.

Eligibility Criteria

Patients who meet all inclusion criteria and none of the exclusion criteria are eligible for enrollment (Textbox 1).

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Provision of written informed consent before participation after receiving an adequate explanation
- Aged 18 years or older at the time of consent acquisition
- No previous exposure to cisplatin-containing chemotherapy
- Scheduled to receive cisplatin monotherapy or cisplatin in combination with up to 2 other anticancer agents (concurrent radiotherapy permitted) under inpatient care
- Diagnosed with lung cancer, head and neck cancer, or other malignancies scheduled to receive a cisplatin dose of 75 mg/m^2 or higher

Exclusion criteria

- Renal impairment (estimated glomerular filtration rate <60 mL/min/1.73 m^2)
- Pregnancy or possibility of pregnancy
- Breastfeeding
- Known hypersensitivity to flopropione
- Any other condition deemed inappropriate for study participation by the principal investigator or subinvestigator

No predefined criteria are established for using preventive measures against CIN, including hydration, magnesium supplementation, or mannitol administration. However, at the participating institution, hydration is routinely implemented for managing patients undergoing cisplatin regimens at doses of 75 mg/m^2 or higher. In addition, magnesium supplementation and mannitol administration are implemented, involving many of the applicable regimens.

Intervention

Intervention Procedure

Eligible participants are randomized in a 5:2 ratio per cohort to receive either flopropione (flopropione group) or no treatment (control group). Participants in the flopropione group receive oral flopropione twice on day 1 (30 minutes before cisplatin administration and at least 4 hours after) and 3 times on day 2. To ensure participant safety, a step-up dosing approach is applied. The single-dose levels are 80 mg in cohort 1, 160 mg in cohort 2, and 240 mg in cohort 3 on day 1. Progression to the next cohort will occur only after confirming safety in the preceding cohort. On day 2, all participants in the flopropione group will receive 80 mg per dose 3 times daily at intervals of at least 4 hours. Blood and urine samples will be collected within 48 hours before cisplatin initiation and at 24 hours, 48 hours, and 1 week after initiation. Participants in the control group undergo the same procedures as those in the flopropione group and receive standard cisplatin-based chemotherapy but without receiving flopropione. Regarding the analysis, participants assigned to the control group among all cohorts will be pooled for evaluation, whereas participants in the flopropione group will be evaluated according to their assigned dose level. An overview of the study procedure is shown in Figure 1; the step-up dosing approach diagram is shown in Figure 2.

Figure 1. Overview of the study procedure.

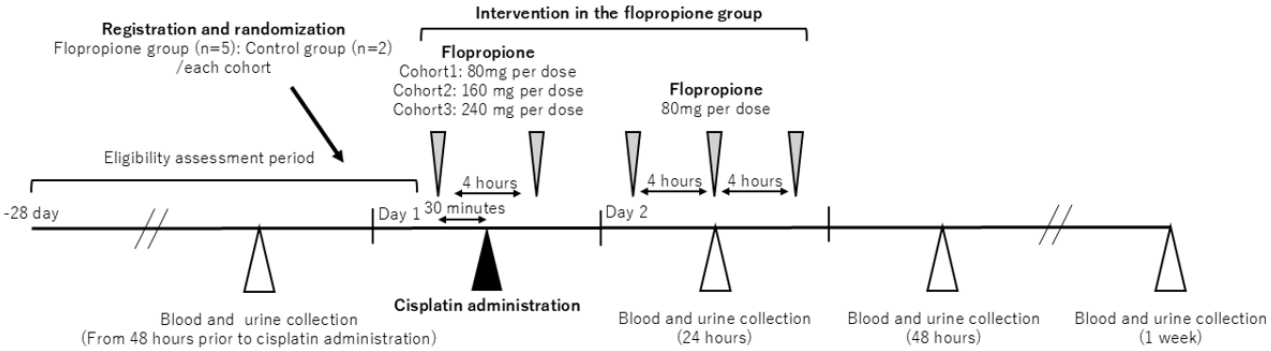

XSL•FO
RenderX

Figure 2. Step-up dosing approach diagram.

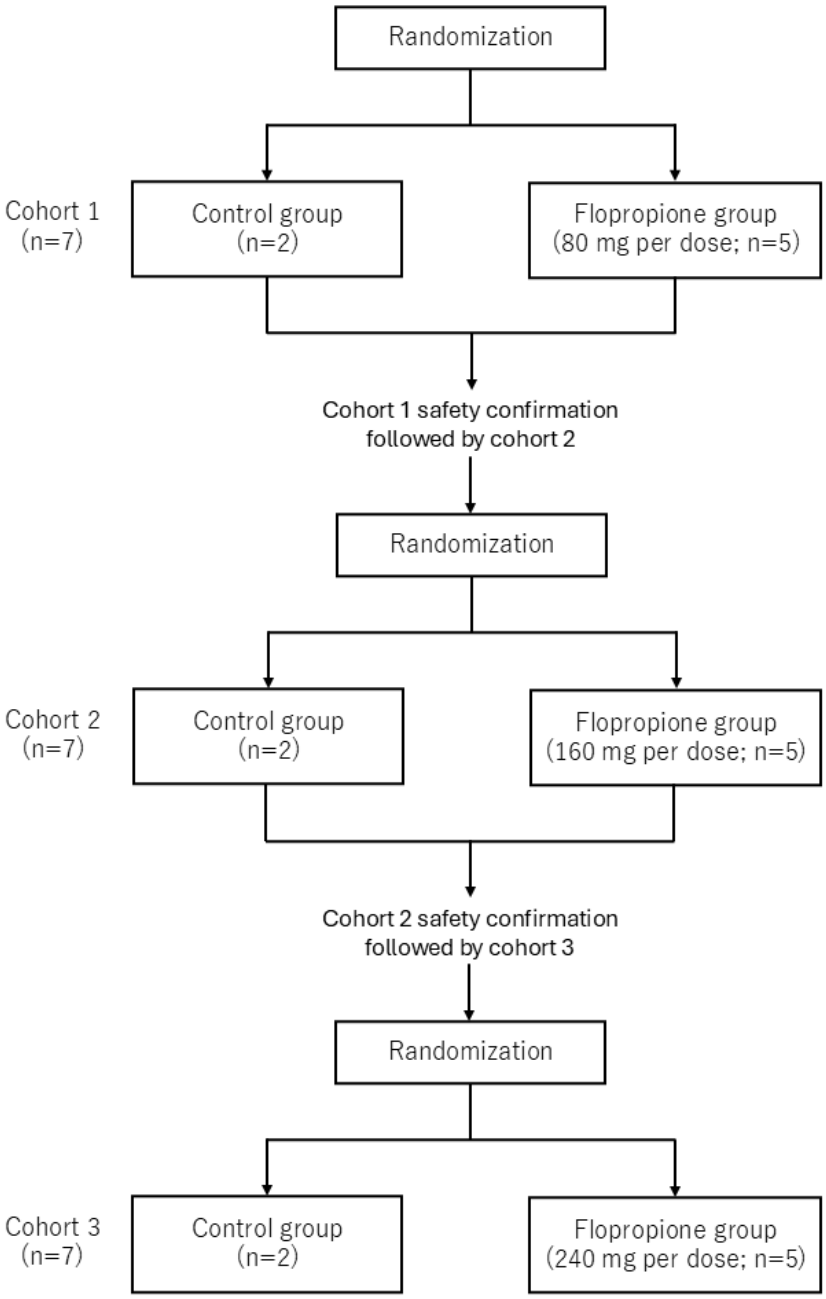

Criteria for Discontinuation of the Protocol Treatment

Treatment will be discontinued if any of the following occur: (1) withdrawal of consent by the participant, (2) inability to continue follow-up, and (3) investigator's judgment that discontinuation is necessary for safety or other reasons.

Prohibited Concomitant Drugs

The concomitant use of nonstudy flopropione is prohibited from 1 day before cisplatin initiation until the 1-week postadministration evaluation.

End Points

Primary End Point

Because this is the first clinical use of flopropione in patients treated with cisplatin, the primary end point is safety. Safety will be assessed by the incidence of adverse events, including abnormal laboratory test results, occurring from study treatment initiation (or 30 minutes before cisplatin initiation in the control group) until 1 week after cisplatin administration. Adverse events will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE; version 4.0), while serious adverse events will be defined based on predefined seriousness criteria.

Secondary End Points

Secondary end points include changes in urinary biomarkers of nephrotoxicity—neutrophil gelatinase-associated lipocalin (N-GAL), liver-type fatty acid-binding protein (L-FABP), kidney injury molecule-1 (KIM-1), *N*-acetyl-β-d-glucosaminidase (NAG), β2-microglobulin, and

XSL•FO
RenderX

α1-microglobulin—from baseline to 24 hours, 48 hours, and 1 week after cisplatin administration.

Exploratory End Points

Exploratory end points include the following:

- Changes in serum creatinine, BUN, and estimated glomerular filtration rate from baseline to 24 hours, 48 hours, and 1 week after cisplatin initiation
- Proportion of participants meeting the CTCAE grade 1 or higher acute kidney injury criteria (increase in serum creatinine of >0.3 mg/dL or 1.5 times above baseline) at 24 hours, 48 hours, and 1 week
- Nontargeted metabolomic analyses of serum and urine samples at each time point

Study Schedule

The schedule for informed consent, registration, randomization, observations, and assessments is presented in Table 1.

Table 1. Study schedule.

Items	Before cisplatin administration		After cisplatin administration		
	From 28 d before cisplatin administration (before visit)	Cisplatin administration day (visit 1)	24 h later (visit 2)	48 h later (visit 3)	1 wk later (visit 4)
Allowable visit windows	—[a]	—	–2 h to +2 h	–2 h to +2 h	–36 h to +36 h
Consent acquisition and registration	✓[b]	✓[b]			
Patient background	✓[b]	✓[b]			
Cisplatin administration		✓			
Flopropione treatment		✓[c]	✓[c]		
Blood and urine collection	✓[b,d]	✓[b,d]	✓	✓	✓
Regimen name and concomitant therapy	✓[e]	✓[e]	✓[e]	✓[e]	✓[e]
Adverse events		✓[f]	✓[f]	✓[f]	✓[f]

[a]Not applicable.

[b]Conduct at either the before visit or visit 1.

[c]The first dose is taken 30 minutes before starting cisplatin administration. On the day of administration, a second dose is taken at least 4 hours apart. On the second day, 3 doses of 80 mg are taken at least 4 hours apart.

[d]Data are collected from 48 hours before cisplatin administration until before starting the regimen.

[e]Data collection period is from registration to 1 week after cisplatin administration.

[f]Data are collected from the start of flopropione treatment (or 30 minutes before starting cisplatin administration in the control group) until the 1-week evaluation point or until discontinuation. As much data as possible are collected at the discontinuation point. If adverse events persist, observation is continued until recovery or until the principal investigator or equivalent determines follow-up is unnecessary.

Eligibility Assessment

Eligibility will be confirmed within 28 days before cisplatin initiation. Baseline characteristics, including age, sex, height, weight, and comorbidities, will be recorded at registration.

Registration and Random Allocation

After confirming eligibility, registration and randomization will be performed through an electronic data capture system, assigning participants in a 5:2 ratio (flopropione:control) within each cohort.

Flopropione Treatment

Flopropione (80 mg per tablet) will be administered orally 30 minutes before cisplatin initiation. Doses are as follows: 80 mg in cohort 1, 160 mg in cohort 2, and 240 mg in cohort 3. A second dose of the same amount will be given at least 4 hours later on day 1, followed by 80 mg 3 times daily on day 2 at intervals of at least 4 hours.

Blood and Urine Collection

Blood and urine samples will be collected at the time points defined as the secondary and exploratory end points. The details of each test item are summarized in Table 2.

XSL•FO
RenderX

Table 2. Laboratory test items.

Classification	Items
Hematological tests	White blood cell count, neutrophil count, lymphocyte count, monocyte count, eosinophil count, basophil count, red blood cell count, hemoglobin, hematocrit, and platelet count
Blood biochemical tests	Total cholesterol, triglycerides, total protein, albumin, albumin/globulin ratio, blood urea nitrogen, uric acid, creatinine, total bilirubin, aspartate aminotransferase, alanine aminotransferase, γ-glutamyl transpeptidase, alkaline phosphatase, lactate dehydrogenase, creatinine kinase (creatine phosphokinase), calcium, phosphorus, sodium, potassium, chlorine, and C-reactive protein
Urinalysis	Hematuria, glucose, protein, creatinine, neutrophil gelatinase-associated lipocalin, liver-type fatty acid-binding protein, kidney injury molecule-1, *N*-acetyl-β-d-glucosaminidase, β2-microglobulin, and α1-microglobulin

Biomarker Measurements

The collected urine samples will be centrifuged at 400×*g* or higher for 5 minutes. Thereafter, the supernatants will be stored at –80 °C until analysis. The samples will be sent to LSI Medience Corporation for N-GAL, L-FABP, KIM-1, NAG, β2-microglobulin, and α1-microglobulin measurement. N-GAL will be measured by a chemiluminescent immunoassay; L-FABP, β2-microglobulin, and α1-microglobulin will be measured by latex agglutination turbidimetry; KIM-1 will be measured by an enzyme-linked immunosorbent assay; NAG will be measured by the artificial substrate MPT method.

Adverse Events

All adverse events, including abnormal laboratory findings, will be collected and recorded in the electronic medical records with information on severity, seriousness, outcome, and the presence or absence of a causal relationship. The seriousness criteria for serious adverse events are (1) death, (2) life-threatening complication, (3) patient requires inpatient hospitalization or prolongation of hospitalization, (4) persistent or significant disability or incapacity, (5) risk of persistent or significant disability or incapacity, (6) congenital anomaly or birth defect, (7) and important medical event.

Sample Size

The planned sample size includes 21 participants: 5 (24%) each in the 80 mg, 160 mg, and 240 mg flopropione cohorts and 6 (29%) in the control group. On the basis of the number of eligible patients in the previous year, anticipated consent rate, and screening dropouts, 21 participants are considered feasible to complete enrollment within 1 to 1.5 years.

Given the early-phase pilot nature of this study, the sample size is selected to balance between exploratory evaluation of potential dose-response trends, safety assessment, ethical considerations, and feasibility. Although not powered for statistical significance, 5 participants per dose level is considered appropriate for an initial exploration of a proof-of-concept study. Participants assigned to the control group will be enrolled among all cohorts and pooled for assessment, providing a common reference group for comparisons with each dose level. This approach allows for efficient use of control data while minimizing unnecessary patient exposure to the investigational intervention.

Statistics

Analysis Set

Full Analysis Set

Patients in the control group are those who receive cisplatin, and participants in the flopropione group are those who receive flopropione. Data of participants in the flopropione group who do not receive flopropione are not included in the analysis.

Per Protocol Set

This is the subset of the full analysis set that excludes participants who have deviations in drug administration or those who use prohibited concomitant medications.

Safety Analysis Set

This set is the same as the full analysis set. The control group is evaluated as 1 group, and the flopropione group is stratified by dose.

Analysis Method

As this is an exploratory pilot study, no formal hypothesis testing is planned. Descriptive and exploratory analyses will be conducted as mentioned subsequently.

The changes in urinary N-GAL, L-FABP, KIM-1, NAG, β2-microglobulin, and α1-microglobulin as well as serum creatinine, BUN, and estimated glomerular filtration rate from baseline to 24 hours, 48 hours, and 1 week after cisplatin administration will be assessed. Summary statistics will be calculated for the change amount, change rate, and observed values at each time point (24 hours, 48 hours, and 1 week after administration) relative to the baseline for each marker. Additionally, between-group differences will be estimated using mixed-effects models for repeated measures if feasible, following an assessment of the data distribution before database lock.

The proportion of participants meeting the CTCAE grade 1 or higher acute kidney injury criteria (creatinine level increase of 0.3 mg/dL or higher; creatinine 1.5 times above baseline) at 24 hours, 48 hours, and 1 week after cisplatin administration will be calculated. The number of participants meeting this criterion at each time point will be analyzed using Fisher exact test.

Nontargeted metabolome analysis will be conducted on serum and urine samples collected at each time point, following previously described protocols [20,21]. Briefly, metabolites will be extracted from the samples and analyzed using capillary electrophoresis coupled with time-of-flight mass spectrometry

XSL•FO
RenderX

(Agilent 6210, Agilent Technologies). Metabolite identification will be performed by matching accurate mass-to-charge ratios (m/z) and migration times to the reference annotation table provided by Human Metabolome Technologies Inc. Peak areas obtained from the capillary electrophoresis coupled with time-of-flight mass spectrometry analysis will be normalized and expressed as relative concentrations to enable comparative evaluation across samples. Statistical analyses will include both univariable approaches and multivariable methods to identify significant temporal changes in the metabolome.

Differences in cancer type, cisplatin dose, renal function, and comorbidities may potentially influence the study end points. Therefore, to explore the impact of these baseline characteristics and clinical variables, subgroup analyses may be considered when deemed necessary.

Data Management

The data will be collected using an electronic data capture system. All data and materials obtained in this study will be managed under the responsibility of the principal investigator. Digital data will be encrypted with passwords, and paper-based materials will be securely stored in locked cabinets. Data will be retained for 5 years following the completion of this clinical study. Upon disposal, all materials will be destroyed in an irreversible manner through methods such as shredding or incineration to ensure destruction.

Ethical Considerations

This study will be conducted in accordance with Japan's Clinical Trials Act and the Declaration of Helsinki. The protocol was reviewed and approved by the clinical research review board of Fujita Health University (CR22-002). At Fujita Health University Hospital, the principal investigator or subinvestigator will screen patients under their care to identify those who may meet the eligibility criteria for participation in this study. Written informed consent will be obtained from eligible patients by the principal investigator or subinvestigator using an explanatory document approved by the clinical research review board, within 28 days before cisplatin administration. All participant data will be deidentified and maintained confidentially. Data will be stored securely with restricted access. Data will be retained for 5 years after study completion. For adverse events, the principal investigator or subinvestigator will provide appropriate medical treatment. For severe adverse events, clinical research insurance will provide appropriate compensation to participants.

Dissemination

The results will be disseminated through peer-reviewed publications, conference presentations, and the Japan Registry of Clinical Trials (jRCTs041220021).

Monitoring and Auditing

On-site monitoring will be performed to ensure compliance with the study protocol, participant safety, and data accuracy. No formal audit is planned.

Trial Registration

This trial was registered in the Japan Registry of Clinical Trials (jRCTs041220021).

Results

This study was approved by the clinical research review board of Fujita Health University in April 2022. This study was subsequently funded by the Japan Agency for Medical Research and Development in April 2024, with the first case registered in July 2024. As of January 2026, participant registration is ongoing. The final participant will complete the study by March 2026. Publication of results is expected by March 2027.

Discussion

Anticipated Findings

CIN is a critical concern about initiating or continuing cisplatin-based chemotherapy. Although hydration therapy is widely implemented, its preventive effects against CIN remain limited. Various pharmacological agents have been investigated for potential protective effects; however, none have yet been successfully translated into clinical practice, and an unmet medical need persists [22]. In addition to currently recommended preventive strategies, such as hydration, the development of mechanism-based pharmacological interventions targeting the pathogenesis of CIN not only reduces the incidence and severity of CIN but also enables cisplatin therapy continuation, potentially leading to improved overall treatment outcomes. Therefore, we designed a clinical trial to evaluate the safety of flopropione, a CCBL1 inhibitor, and obtain proof of concept for its preventive effect against CIN. If proof of concept for the preventive effect of flopropione can be demonstrated in this study, subsequent clinical trials would be able to further verify its efficacy. Such evidence would support the potential of flopropione as an adjunctive therapy to existing preventive strategies, including hydration, with the expectation of achieving greater suppression of CIN incidence compared with those currently available. Furthermore, if its clinical benefit is confirmed, future studies should also consider comparative trials against hydration-based protocols to evaluate whether flopropione could serve as a simpler and more convenient alternative preventive approach to CIN.

In designing this study, several measures are implemented to ensure participant safety while obtaining proof of concept. No previous studies have evaluated the concomitant use of flopropione with cisplatin; therefore, a step-up dosing approach was adopted in the participants who were hospitalized. This design allows for sequential confirmation of safety across cohorts while gradually increasing the flopropione dose. Given that hydration may reduce plasma concentrations of orally administered drugs, we have set the maximum single dose of flopropione on cisplatin administration day 1 at 240 mg, consistent with the previously reported dose administered to healthy adults [19].

N-GAL, L-FABP, and KIM-1 were selected as urinary biomarkers for detecting early renal injury. All have been reported as useful indicators for early detection of CIN [23-25]. Furthermore, urinary KIM-1 levels reportedly continue to rise for up to 10 days after the initial administration of cisplatin [26]. Therefore, urinary biomarkers will be measured at 24 hours, 48 hours, and 1 week after cisplatin initiation.

XSL•FO
RenderX

Limitations

This study has several limitations. First, it is an open-label study using a nontreated control group rather than a placebo control. Although the primary focus is on safety and biomarker evaluation, the potential for participant- or investigator-related bias cannot be completely eliminated. Therefore, a future pivotal study should adopt a randomized, double-blind, placebo-controlled design. Second, this study does not assess the impact of flopropione on cisplatin efficacy. Because this is the first clinical investigation of flopropione used in combination with cisplatin, it is limited to the first cisplatin administration (cycle 1) with an observation period of 1 week. Evaluation of potential effects on antitumor efficacy will require a longer observation period. Third, several studies have suggested that nephrotoxicity develops in a cumulative, dose-dependent manner [27,28]. Therefore, CIN may not necessarily occur after a single cisplatin administration. Accordingly, future studies involving multiple cycles of cisplatin and repeated treatment of flopropione will be necessary for comprehensively assessing its safety and preventive efficacy.

Conclusions

Cisplatin continues to be one of the most important anticancer agents. However, prevention of CIN is crucial for expanding therapeutic options in patients with cancer and enhancing treatment sustainability. This study is a proof-of-concept trial designed to evaluate flopropione, a CCBL1 inhibitor, with a focus on the pathophysiological mechanisms underlying CIN. It occupies an important position in generating evidence that will inform subsequent clinical trials. Through this study and the following clinical trials, we aim to demonstrate the preventive efficacy of flopropione against CIN and establish it as a prophylactic agent capable of further reducing the incidence of CIN.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to participant confidentiality and ongoing regulatory requirements, but are available from the corresponding author on reasonable request, subject to approval by the institutional review board.

Funding

This clinical trial was entirely supported by the Japan Agency for Medical Research and Development (grants 24lk0221187h0001 and 24lk0221187h0002).

Authors' Contributions

Investigation: MK (principal investigator), YO, NT, K Kawada, YG, NH, KI
Project administration: TK, MK, K Kikuchi, HK
Supervision: KS, HS
Writing—original draft: TK, HF
Writing—review and editing: MK, HF, K Kikuchi, YO, HK, TM, NT, K Kawada, YG, NH, KI, AK, HY, KS, HS

Conflicts of Interest

None declared.

References

1. Rottenberg S, Disler C, Perego P. The rediscovery of platinum-based cancer therapy. Nat Rev Cancer 2021 Jan;21(1):37-50 [FREE Full text] [doi: 10.1038/s41568-020-00308-y] [Medline: 33128031]
2. Elmorsy EA, Saber S, Hamad RS, Abdel-Reheim MA, El-Kott AF, AlShehri MA, et al. Advances in understanding cisplatin-induced toxicity: molecular mechanisms and protective strategies. Eur J Pharm Sci 2024 Dec 01;203:106939 [FREE Full text] [doi: 10.1016/j.ejps.2024.106939] [Medline: 39423903]
3. Tang C, Livingston MJ, Safirstein R, Dong Z. Cisplatin nephrotoxicity: new insights and therapeutic implications. Nat Rev Nephrol 2023 Jan;19(1):53-72 [FREE Full text] [doi: 10.1038/s41581-022-00631-7] [Medline: 36229672]
4. Latcha S, Jaimes EA, Patil S, Glezerman IG, Mehta S, Flombaum CD. Long-term renal outcomes after cisplatin treatment. Clin J Am Soc Nephrol 2016 Jul 07;11(7):1173-1179 [FREE Full text] [doi: 10.2215/CJN.08070715] [Medline: 27073199]
5. Madias NE, Harrington JT. Platinum nephrotoxicity. Am J Med 1978 Aug;65(2):307-314 [FREE Full text] [doi: 10.1016/0002-9343(78)90825-2] [Medline: 99034]
6. Jantarat A, Thamlikitkul L, Thephamongkhol K, Setakornnukul J, Phisalprapa P, Kositamongkol C, et al. Efficacy and safety of short intravenous hydration for preventing nephrotoxicity from high-dose cisplatin: a randomized, open-label, phase II trial. JCO Glob Oncol 2025 Jan 11;11:e2400515. [doi: 10.1200/GO-24-00515] [Medline: 39819128]
7. Sikking C, Niggebrugge-Mentink KL, van der Sman AS, Smit RH, Bouman-Wammes EW, Beex-Oosterhuis MM, et al. Hydration methods for cisplatin containing chemotherapy: a systematic review. Oncologist 2024 Mar 02;29(2):e173-e186 [FREE Full text] [doi: 10.1093/oncolo/oyad297] [Medline: 37995306]

XSL•FO
RenderX

8. Tiseo M, Martelli O, Mancuso A, Sormani MP, Bruzzi P, Di Salvia R, et al. Short hydration regimen and nephrotoxicity of intermediate to high-dose cisplatin-based chemotherapy for outpatient treatment in lung cancer and mesothelioma. Tumori 2007;93(2):138-144 [FREE Full text] [doi: 10.1177/030089160709300205] [Medline: 17557559]
9. Hase T, Miyazaki M, Ichikawa K, Yogo N, Ozawa N, Hatta T, et al. Short hydration with 20 mEq of magnesium supplementation for lung cancer patients receiving cisplatin-based chemotherapy: a prospective study. Int J Clin Oncol 2020 Nov;25(11):1928-1935 [FREE Full text] [doi: 10.1007/s10147-020-01755-1] [Medline: 32740717]
10. Ilhan Y, Onder AH, Ozbay MF, Karakaya G, Sezgin Goksu S, Ozturk B, et al. Protective effect of magnesium preloading on cisplatin-induced nephrotoxicity: which dose is more suitable? Eur Rev Med Pharmacol Sci 2024 May;28(9):3403-3413 [FREE Full text] [doi: 10.26355/eurrev_202405_36185] [Medline: 38766796]
11. Li J, Wu Y, Chen C, Zhang W, Yue L, Liu T. A systematic review for prevention of cisplatin-induced nephrotoxicity using different hydration protocols and meta-analysis for magnesium hydrate supplementation. Clin Exp Nephrol 2024 Jan 02;28(1):1-12. [doi: 10.1007/s10157-023-02386-2] [Medline: 37530867]
12. Li S, He X, Ruan L, Ye T, Wen Y, Song Z, et al. Protective effect of mannitol on cisplatin-induced nephrotoxicity: a systematic review and meta-analysis. Front Oncol 2021 Dec 16;11:804685 [FREE Full text] [doi: 10.3389/fonc.2021.804685] [Medline: 34976843]
13. Al Suleimani YM, Nomeir Y, Al Maskari R, Ali H, Manoj P, Abdelrahman AM. Riociguat alleviates cisplatin-caused kidney injury by suppressing oxidative stress and inflammation. Biology (Basel) 2025 Oct 02;14(10):1346 [FREE Full text] [doi: 10.3390/biology14101346] [Medline: 41154749]
14. Airik M, Clayton K, Wipf P, Airik R. JP4-039 mitigates cisplatin-induced acute kidney injury by inhibiting oxidative stress and blocking apoptosis and ferroptosis in mice. Antioxidants (Basel) 2024 Dec 15;13(12):1534 [FREE Full text] [doi: 10.3390/antiox13121534] [Medline: 39765862]
15. Zhang Y, Hu J, Zhang Y, Ci X. Amentoflavone protects against cisplatin-induced acute kidney injury by modulating Nrf2-mediated oxidative stress and ferroptosis and partially by activating Nrf2-dependent PANoptosis. Front Pharmacol 2025 Mar 5;16:1508047. [doi: 10.3389/fphar.2025.1508047] [Medline: 40110131]
16. Bushau-Sprinkle A, Barati M, Zheng Y, Watson W, Gagnon K, Khundmiri S, et al. Na/H exchange regulatory factor 1 deficient mice show evidence of oxidative stress and altered cisplatin pharmacokinetics. Antioxidants (Basel) 2021 Jun 28;10(7):1036 [FREE Full text] [doi: 10.3390/antiox10071036] [Medline: 34203453]
17. McSweeney KR, Gadanec LK, Qaradakhi T, Ali BA, Zulli A, Apostolopoulos V. Mechanisms of cisplatin-induced acute kidney injury: pathological mechanisms, pharmacological interventions, and genetic mitigations. Cancers (Basel) 2021 Mar 29;13(7):1572 [FREE Full text] [doi: 10.3390/cancers13071572] [Medline: 33805488]
18. Sukeda N, Fujigaki H, Ando T, Ando H, Yamamoto Y, Saito K. Identification of 2',4',6'-trihydroxyacetophenone as promising cysteine conjugate beta-lyase inhibitor for preventing cisplatin-induced nephrotoxicity. Mol Cancer Ther 2023 Jul 05;22(7):873-881. [doi: 10.1158/1535-7163.MCT-22-0564] [Medline: 37163384]
19. Cospanon tablets. Daiichi Sankyo. URL: https://www.pmda.go.jp/PmdaSearch/iyakuDetail/ResultDataSetPDF/170033_1249007F1032_1_09 [accessed 2025-11-16]
20. Umehara H, Numata S, Watanabe SY, Hatakeyama Y, Kinoshita M, Tomioka Y, et al. Altered KYN/TRP, Gln/Glu, and Met/methionine sulfoxide ratios in the blood plasma of medication-free patients with major depressive disorder. Sci Rep 2017 Jul 07;7(1):4855 [FREE Full text] [doi: 10.1038/s41598-017-05121-6] [Medline: 28687801]
21. Watanabe Y, Kasuga K, Tokutake T, Kitamura K, Ikeuchi T, Nakamura K. Alterations in glycerolipid and fatty acid metabolic pathways in Alzheimer's disease identified by urinary metabolic profiling: a pilot study. Front Neurol 2021 Oct 27;12:719159 [FREE Full text] [doi: 10.3389/fneur.2021.719159] [Medline: 34777195]
22. Fang CY, Lou DY, Zhou LQ, Wang JC, Yang B, He QJ, et al. Natural products: potential treatments for cisplatin-induced nephrotoxicity. Acta Pharmacol Sin 2021 Dec;42(12):1951-1969 [FREE Full text] [doi: 10.1038/s41401-021-00620-9] [Medline: 33750909]
23. Szumilas D, Owczarek AJ, Brzozowska A, Niemir ZI, Olszanecka-Glinianowicz M, Chudek J. The value of urinary NGAL, KIM-1, and IL-18 measurements in the early detection of kidney injury in oncologic patients treated with cisplatin-based chemotherapy. Int J Mol Sci 2024 Jan 16;25(2):1074 [FREE Full text] [doi: 10.3390/ijms25021074] [Medline: 38256147]
24. Yanishi M, Kinoshita H. Urinary L-type fatty acid-binding protein is a predictor of cisplatin-induced acute kidney injury. BMC Nephrol 2022 Mar 31;23(1):125 [FREE Full text] [doi: 10.1186/s12882-022-02760-4] [Medline: 35361160]
25. Ichimura T, Hung CC, Yang SA, Stevens JL, Bonventre JV. Kidney injury molecule-1: a tissue and urinary biomarker for nephrotoxicant-induced renal injury. Am J Physiol Renal Physiol 2004 Mar;286(3):F552-F563 [FREE Full text] [doi: 10.1152/ajprenal.00285.2002] [Medline: 14600030]
26. George B, Wen X, Mercke N, Gomez M, O'Bryant C, Bowles DW, et al. Time-dependent changes in kidney injury biomarkers in patients receiving multiple cycles of cisplatin chemotherapy. Toxicol Rep 2020;7:571-576 [FREE Full text] [doi: 10.1016/j.toxrep.2020.04.003] [Medline: 32382514]
27. Miller R, Tadagavadi R, Ramesh G, Reeves W. Mechanisms of cisplatin nephrotoxicity. Toxins (Basel) 2010 Nov;2(11):2490-2518 [FREE Full text] [doi: 10.3390/toxins2112490] [Medline: 22069563]

XSL•FO
RenderX

28. Chovanec M, Abu Zaid M, Hanna N, El-Kouri N, Einhorn LH, Albany C. Long-term toxicity of cisplatin in germ-cell tumor survivors. Ann Oncol 2017 Nov 01;28(11):2670-2679 [FREE Full text] [doi: 10.1093/annonc/mdx360] [Medline: 29045502]

Abbreviations

BUN: blood urea nitrogen
CCBL1: cysteine conjugate-β lyase 1
CIN: cisplatin-induced nephrotoxicity
CTCAE: Common Terminology Criteria for Adverse Events
KIM-1: kidney injury molecule-1
L-FABP: liver-type fatty acid-binding protein
NAG: N-acetyl-β-d-glucosaminidase
N-GAL: neutrophil gelatinase-associated lipocalin

Edited by A Schwartz; submitted 17.Nov.2025; peer-reviewed by MS Oba; comments to author 15.Dec.2025; accepted 20.Jan.2026; published 12.Feb.2026.

Please cite as:
Koseki T, Kondo M, Fujigaki H, Kikuchi K, Oya Y, Kato H, Mizuno T, Tsuboi N, Kawada K, Goto Y, Hashimoto N, Imaizumi K, Kada A, Yabuuchi H, Saito K, Saya H
Flopropione, a Cysteine Conjugate β-Lyase 1 Inhibitor, for Prevention of Cisplatin-Induced Nephrotoxicity: Protocol for a Randomized, Open-Label, Proof-of-Concept Phase 1 and 2a Trial
JMIR Res Protoc 2026;15:e87907
URL: https://www.researchprotocols.org/2026/1/e87907
doi:10.2196/87907
PMID:

©Takenao Koseki, Masashi Kondo, Hidetsugu Fujigaki, Kayoko Kikuchi, Yuko Oya, Hiroshi Kato, Tomohiro Mizuno, Naotake Tsuboi, Kenji Kawada, Yasuhiro Goto, Naozumi Hashimoto, Kazuyoshi Imaizumi, Akiko Kada, Hikaru Yabuuchi, Kuniaki Saito, Hideyuki Saya. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Acupuncture Therapies for Chemotherapy-Induced Nausea and Vomiting in Patients With Breast Cancer: Protocol for a Systematic Review and Network Meta-Analysis

Yutao Luo[1], MBBS; Zhujun Bian[1], MBBS; Qi Wang[2], MM; Tiantian Lei[1], MBBS; Hong Zhao[1,3], MD

[1]The First School of Clinical Medicine, Zhejiang Chinese Medical University, No. 548 Binwen Road, Binjiang District, Hangzhou, Zhejiang, China

[2]Department of Breast Surgery, Jiaxing Hospital of Traditional Chinese Medicine, Jiaxing, Zhejiang, China

[3]Department of Breast Surgery, The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), Hangzhou, China

Corresponding Author:
Hong Zhao, MD
The First School of Clinical Medicine, Zhejiang Chinese Medical University, No. 548 Binwen Road, Binjiang District, Hangzhou, Zhejiang, China

Abstract

Background: As a prevalent side effect, chemotherapy-induced nausea and vomiting (CINV) imposes a burden on the daily lives of patients with breast cancer. Multiple clinical trials have suggested the validity of acupuncture in alleviating CINV; however, the optimal acupuncture modality remains unclear.

Objective: This protocol describes a systematic review and network meta-analysis to investigate the efficacy and safety of distinct acupuncture interventions for treating CINV in patients with breast cancer.

Methods: Eight databases (PubMed, Embase, Web of Science, the Cochrane Library, China National Knowledge Infrastructure, VIP Database, Wanfang Database, and the Chinese Biomedical Literature Database) will be searched for eligible studies from their respective inception to July 31, 2025. The language of published studies is limited to English and Chinese. Primary outcomes are CINV intensity and clinical effectiveness rates. Secondary outcomes include recurrence rates, safety outcomes, and quality of life. Risk of bias will be assessed using the Cochrane risk of bias tool. Pairwise meta-analysis will be conducted in Stata using random-effects models with Hartung-Knapp-Sidik-Jonkman CIs. Network meta-analysis will be conducted using Bayesian Markov chain Monte Carlo methods in R software. Convergence will be assessed using Gelman-Rubin statistics and trace plots. Heterogeneity will be summarized using τ^2 and τ, along with prediction intervals, when applicable. Consistency between direct and indirect evidence will be evaluated using the node-splitting method and design-by-treatment interaction test. Small-study effects will be assessed via comparison-adjusted funnel plots and the Egger test in Stata. Where feasible, subgroup analyses and meta-regression analyses will be performed. The certainty of evidence will be evaluated using the Grading of Recommendations Assessment, Development, and Evaluation framework.

Results: A preliminary scoping search was completed in August 2025, and this protocol was finalized in October 2025. The comprehensive literature search and study selection are expected to be completed by June 2026, followed by data extraction by August 2026. Data synthesis and final manuscript preparation are scheduled to be completed by December 2026.

Conclusions: This analysis will expand the range of evidence-based acupuncture options available to clinicians for treating CINV in patients with breast cancer.

Trial Registration: PROSPERO CRD420251030300; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251030300

International Registered Report Identifier (IRRID): DERR1-10.2196/86384

(JMIR Res Protoc 2026;15:e86384) doi:10.2196/86384

KEYWORDS

acupuncture; chemotherapy-induced nausea and vomiting; breast cancer; network meta-analysis; protocol

Introduction

Breast cancer is the most widespread malignant tumor in women and a major contributor to cancer-related deaths in female patients worldwide [1,2]. In the past few decades, the prevalence of breast cancer has risen dramatically, increasing the burden of the disease and making it a key public health problem worldwide. According to the Global Cancer Observatory statistics, the standardized incidence and mortality rates of female patients with breast carcinoma worldwide in 2022 were

XSL•FO
RenderX

47.8 per 100,000 and 13.6 per 100,000, respectively. Breast cancer leads globally in both incidence and fatality across most countries [2]. Currently, surgery, chemotherapy, targeted therapy, endocrine therapy, and radiotherapy are the leading treatments for breast cancer [3]. Chemotherapy can effectively prolong the life of patients [4] and plays an essential role in preoperative neoadjuvant therapy, postoperative adjuvant therapy, and advanced-stage treatment.

However, chemotherapy also causes a series of side effects, such as myelosuppression, cancer-related fatigue, nausea and vomiting, and liver and kidney toxicity [5]. Among these, chemotherapy-induced nausea and vomiting (CINV) remains a frequent adverse effect, occurring in about 70% to 80% of patients undergoing chemotherapy [6] and causing metabolic disorders and malnutrition. It greatly affects the patients' daily lives and may even lead to fear of chemotherapy and a decline in compliance.

The mechanisms by which chemotherapy drugs cause nausea and vomiting are complex and not completely defined. It is generally accepted that most chemotherapeutic drugs irritate the mucosa of the gastrointestinal tract, causing mucosal damage and leading to the release of 5-hydroxytryptamine (5-HT) from enterochromaffin cells in the mucosa between the stomach and the ileum, which binds to appropriate receptors to generate nerve impulses that stimulate the vomiting centers, leading to the occurrence of vomiting [7]. In contrast, the chemoreceptor trigger zone, located at the base of the fourth ventricle of the brain, has a weak blood-brain barrier and is therefore susceptible to stimulation by chemotherapy drugs and their metabolites in the blood. In addition, substance P, which stimulates vomiting by binding to the neurokinin-1 (NK-1) receptor, is a critical neurotransmitter that causes CINV [8].

Antiemetic drugs in clinical use can be broadly categorized into 5-HT3 receptor antagonists (RAs), NK-1 RAs, glucocorticoids, phenothiazines, and others. Among these, 5-HT3 RAs are the primary medications used for the treatment of CINV in clinical practice, including ondansetron, granisetron, dolasetron, and tropisetron [9]. Unlike other 5-HT3 RAs, palonosetron is a safer and more efficient second-generation 5-HT3 RA [10]. It has an extended half-life of 40 hours and a significantly higher receptor affinity. It is especially useful for the prevention and mitigation of acute CINV. In contrast, NK-1 RAs and substance P are closely associated with delayed CINV [11]. Representative drugs include aprepitant, rolapitant, and netupitant, which have a broader range of antiemetic effects and fewer adverse effects. In addition, dexamethasone is often used in combination with other drugs in CINV treatment. Although there is no unanimous consensus on its antiemetic mechanism, it has been clinically shown that a combination of dexamethasone with RAs significantly relieves nausea and vomiting in patients undergoing chemotherapy [12]. Furthermore, for chemotherapy agents with different emetic risk classes, various combination regimens are formulated [13].

Although there is no concept of chemotherapy in traditional Chinese medicine (TCM), TCM possesses extensive and well-established therapeutic experience in effectively addressing nausea and vomiting. According to TCM theory, the pathogenesis of these symptoms is closely linked to the spleen, stomach, and *qi* (the fundamental substance that constitutes the body and sustains vital activities). The spleen-stomach system serves as the origin of *qi* and blood and is responsible for transforming food and transporting nutrients to nourish all *zang-fu* organs and body tissues. Pathogenic factors disrupt normal physiological functions, while chemotherapy-induced cytotoxicity damages vital *qi*. This leads to spleen-stomach deficiency, impaired transformation-transportation functions, dysregulation of *qi* ascending-descending movements, and stomach-*qi* ascending, which ultimately manifests as nausea and vomiting [14].

TCM explores various therapeutic approaches for nausea and vomiting, broadly categorized into internal and external therapies based on administration routes. Internal therapies primarily involve oral herbal decoctions, while external modalities focus on acupuncture and moxibustion. Compared to Western pharmaceuticals, TCM demonstrates advantages such as cost-effectiveness, minimal adverse effects, and procedural simplicity, making it widely adopted in CINV management globally [15-19].

Acupuncture achieves therapeutic effects through acupoint stimulation; therefore, the broader concept of acupuncture encompasses multiple stimulation techniques, including but not limited to electroacupuncture, auricular acupressure, fire needling, catgut embedding, intradermal needling, abdominal acupuncture, warm needling, conventional acupuncture, mild moxibustion, ginger-partitioned moxibustion, and acupressure. In clinical practice, practitioners may use single modalities or combinations of 2 to 3 techniques.

Although these interventions are extensively used clinically and investigated in trials, most comparative studies focus on acupuncture vs Western medications or placebo rather than direct comparisons between different acupuncture modalities. This creates an evidence gap for clinicians selecting optimal acupuncture protocols. The inherent diversity of acupuncture interventions poses challenges for conventional meta-analyses and systematic reviews in addressing these clinical questions, highlighting the need for more head-to-head comparative studies among acupuncture techniques. Network meta-analysis (NMA) can help address this dilemma by comparing multiple interventions when there is insufficient direct evidence to compare all therapies. By combining both direct comparisons from randomized controlled trials (RCTs) and indirect comparisons through a common comparator, NMA allows the simultaneous estimation of relative treatment effects among multiple competing interventions within a single analytical framework. This approach has been increasingly adopted to provide a coherent synthesis of evidence and to facilitate the comparative evaluation and ranking of treatment options in complex clinical decision-making contexts [20]. Its analytical framework extends beyond the scope of pairwise head-to-head comparative meta-analyses [21]. By systematically comparing multiple competing interventions across different but sufficiently comparable study populations and clinical contexts, this approach enables the evaluation of broader clinical questions within an expanded evidence synthesis framework [22].

XSL•FO
RenderX

In this study, an NMA protocol is designed to determine the most efficacious acupuncture interventions (individual or combined therapies) for managing CINV in patients with breast cancer, using evidence derived from RCTs. Additionally, their safety profiles and subgroup therapeutic effects will be systematically evaluated.

Methods

This NMA protocol rigorously follows the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) 2015 statement (Checklist 1) [23,24]. This protocol is registered in the PROSPERO database (CRD420251030300).

Ethical Considerations

Ethics approval is not required as this study is based on publicly available data and does not involve animal testing.

Eligibility Criteria

The literature screening process is conducted according to the population, intervention, comparison, outcome, and study (PICOS) framework.

Population

Eligible participants include adult female patients (aged ≥18 years) with a pathologically confirmed diagnosis of breast cancer who are undergoing moderately or highly emetogenic chemotherapy and consent to receive acupuncture-related interventions. Patients receiving concurrent radiotherapy or with pre-existing gastrointestinal comorbidities, immunodeficiencies, or physical or cognitive impairments will be excluded. No restrictions are applied regarding nationality, ethnicity, education level, marital status, or socioeconomic background.

Interventions

Considering the heterogeneity of TCM acupuncture-related therapies, this review will exclusively include modalities frequently used in CINV management. These include hand needling, electroacupuncture, auricular acupressure, fire needling, warm needling, catgut embedding, intradermal needling, abdominal acupuncture, acupressure, or a combination of these methods. The interventions of plum-blossom needling, dry needling, wrist-ankle acupuncture, acupoint patching, acupoint injection, and moxibustion are outside the scope of this study. We do not set limitations on the specific acupoint selection or treatment duration.

Comparators

Control interventions include placebo acupuncture, sham acupuncture, conventional antiemetic pharmacotherapy (eg, various 5-HT3 RAs), and other acupuncture methods as described previously. Studies investigating variations within the same acupuncture technique, such as comparisons between alternative acupoint choices or different treatment durations, are systematically eliminated.

Outcomes

Primary Outcomes

CINV intensity and clinical effectiveness rates are the primary outcomes of this review.

CINV intensity (continuous outcome) may be quantified using multiple validated assessment tools [16,25-27] (eg, visual analog scale, Multinational Association of Supportive Care in Cancer Antiemetic Tool, Simulator Sickness Questionnaire, Rhodes Index of Nausea, Vomiting and Retching, and the Index of Nausea, Vomiting, and Retching). To enable quantitative synthesis across studies using different instruments, we will treat CINV intensity as a continuous outcome and standardize it using a common effect size. If a study reports multiple time points, we will extract the time point closest to the end of the intervention.

Clinical effectiveness (dichotomous outcomes) is assessed based on the World Health Organization (WHO) criteria for grading CINV and established acupuncture efficacy standards. Clinical efficacy is categorized into 4 levels [28]: complete response (CR; no nausea or vomiting), partial response (PR; mild nausea not affecting appetite and vomiting 1 to 2 times per day), mild response (moderate nausea impairing appetite and vomiting 3 to 5 times per day), and failure (severe nausea with >5 episodes of vomiting per day). The overall effectiveness rate is computed as follows: [(number of patients with CR+PR)/total number of patients]×100%. Additionally, CR has been set as an outcome indicator in some studies. While the National Cancer Institute Common Toxicity Criteria exhibits minor classification discrepancies compared with WHO standards, their inclusion is permitted due to negligible impact on efficacy calculations [29,30]. Studies lacking explicit definitions of CR or PR will be excluded to ensure outcome reliability.

Secondary Outcomes

Secondary outcomes include CINV recurrence rates after the intervention; patients' quality of life, evaluated using related tools (eg, Quality of Life questionnaire-core 30 and Beck Anxiety Inventory); and safety profiles covering the type and incidence of adverse events.

Studies

Only parallel-designed RCTs are included, while crossover trials, case-control studies, cohort studies, and animal or cellular tests are not considered. The language of the studies is confined to English or Chinese. We also excluded conference abstracts, case reports, reviews, guidelines, letters, and studies with missing content.

Search Strategy

A thorough electronic search of 8 major biomedical databases (PubMed, Embase, Web of Science, the Cochrane Library, China National Knowledge Infrastructure, Wanfang Database, VIP Database, and the Chinese Biomedical Literature Database) will be implemented from their respective inception dates to July 31, 2025. The search terms will consist of standardized Medical Subject Headings (MeSH), TCM-related terms, and free keywords. As an illustration, the PubMed search strategy is outlined in Multimedia Appendix 1, with full details provided

XSL•FO
RenderX

therein. The search formula will be subject to targeted minor modifications according to the search requirements for different databases. Non-English or non-Chinese records will be excluded during screening, and such exclusions will be documented in the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flow diagram.

Study Selection

Study selection is independently conducted by 2 reviewers (YL and ZB), who will screen all retrieved studies according to the eligibility criteria. All conflicts in selection are resolved by a third reviewer (HZ), who specializes in diagnosing and treating breast cancer. The search and screening process is depicted in Figure 1. All retrieved research will be imported into EndNote (version X8.2; Clarivate). Initially, duplicate literature will be eliminated using this software. After that, the titles and abstracts of the records will be read for preliminary filtering to remove irrelevant and ineligible studies, followed by a thorough review of the remaining studies to select the final included studies. If the data are incomplete in a study, we will contact the corresponding author to request additional details; the study will be excluded if the author cannot be reached or the missing information cannot be obtained.

XSL•FO
RenderX

Figure 1. The process for records selection. RCT: randomized controlled trial.

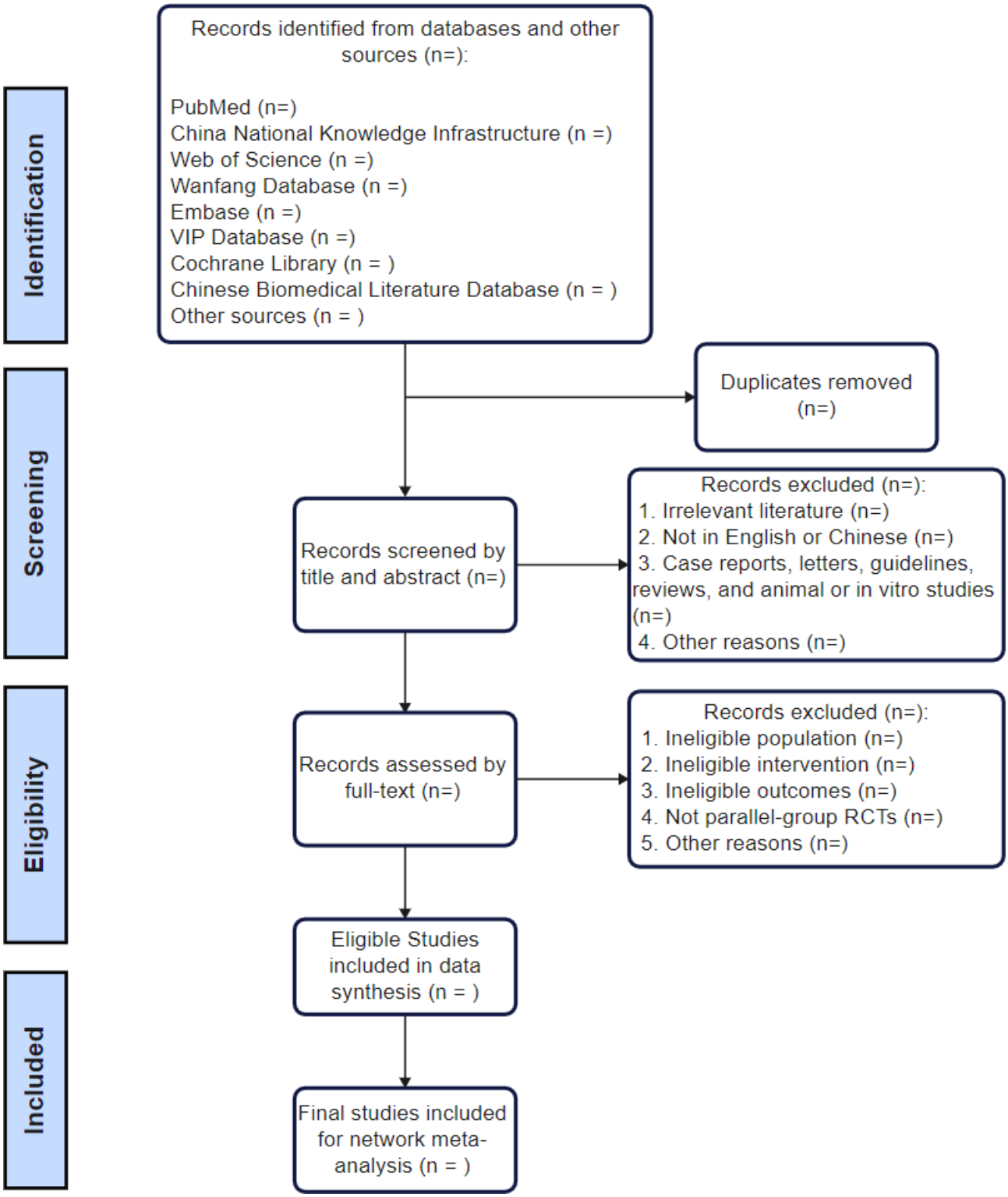

Data Extraction

Two researchers (YL and TL) will systematically extract data from eligible studies into a standardized Microsoft Excel 2019 template, capturing the following elements: literature details (authors, publication year, journal, and country), trial characteristics (blinding methods, randomization processes, number of participants, and chemotherapy scheme), intervention (acupuncture method and acupoint selection), comparator (control measures and Western medicine names), and outcome (preintervention or postintervention values, outcome criteria, and their authoritative sources). Particular attention is given to distinguishing chemotherapy cycle phases and recording acute CINV (≤24 hours) and delayed CINV (24 hours to 7 days) separately. To minimize extraction bias, standardized conversion methods will be applied for studies reporting data in unconventional formats. For continuous outcomes, means and SDs will be estimated from medians, ranges, or IQRs using validated formulas. SEs and CIs will also be converted according to the Cochrane Handbook. For dichotomous outcomes, if only

XSL•FO
RenderX

event counts are reported, the corresponding percentages will be calculated to ensure data consistency. For 3-armed trials, comparisons between the intervention and control groups will be appropriately structured—such as by splitting the shared control group—to enable inclusion as pairwise comparisons in the analysis. If disagreements arise, a third reviewer (HZ) will be invited to reach a consensus when necessary.

Risk of Bias Assessment

The Cochrane risk of bias tool [31] will be used to evaluate the quality of the included literature, which mainly includes 6 aspects of the risk of bias. Selection bias mainly evaluates the randomness of the grouping method (eg, random number table and computer-generated sequences) and the allocation concealment (eg, sealed envelope methods). Blinding during implementation and blinding of outcome assessment are considered performance bias and detection bias, respectively, and are categorized as single-blind, double-blind, or unblinded. Attrition bias focuses on study participant attrition and aims to evaluate whether participant absences and their reasons are truthfully recorded. Reporting bias is assessed to ensure the authenticity and integrity of the study data. Any additional elements that could cause bias are categorized as other biases. All studies are evaluated by 2 independent assessors, and each bias will be categorized into 3 grades: low risk (meets the criteria), high risk (criteria deviant), and unclear risk (inadequate evidence to make a judgment), with adjudication by a third reviewer if disagreements persist. The results will be presented as a bias risk diagram by the Review Manager (Cochrane Collaboration).

Statistical Analysis

Data Integration

Initially, pairwise meta-analysis will be carried out to directly compare the effect of various acupuncture modalities for CINV. Given potential clinical and methodological variations across studies, pairwise syntheses will primarily use random-effects models. For CINV intensity measured by different instruments, we will use the standardized mean difference with Hedges *g* small-sample correction as the common comparative effect size and report the 95% CIs. All intensity scales will be oriented such that negative values represent symptom improvement. For clinical effectiveness outcomes, we will use the risk ratio as the effect size with 95% CIs. The 95% CIs will be computed using the Hartung-Knapp-Sidik-Jonkman method to enhance inference reliability and reduce inflated type I error rates, particularly when the number of studies is small [32]. These steps will be performed using Stata (version 17.0; StataCorp).

A Bayesian probabilistic NMA will be implemented in R (version 4.4.1; R Foundation for Statistical Computing) "*gemtc*" package, if the transitivity assumption is valid. Separate NMAs will be performed for continuous and dichotomous outcomes, specifying appropriate likelihoods and link functions for each outcome type (normal likelihood with an identity link for standardized mean difference; binomial likelihood with a log link for risk ratio). We will conduct data analysis with the Markov chain Monte Carlo model and the random-effects model. Four independent chains will be set up, and 50,000 iterations will be run for each chain. Visual assessment of the convergence will be performed using Gelman-Rubin statistics and trace plots [33,34]. An R-hat value approaching 1 indicates good convergence, trace plots should exhibit stable fluctuations, and autocorrelation should be low. To offer reliable data support for clinical decisions, we will rank the efficacy of varying acupuncture methods. Treatment ranking will be demonstrated by the surface under the cumulative ranking curve; the larger the surface under the curve, the greater the acupuncture therapy effectiveness.

Network Geometry

Network graphs will be plotted to demonstrate the evidence relationships between interventions. Given that separate NMAs will be conducted for different primary outcomes, network plots will be generated separately for each outcome type using studies that report the corresponding outcome. In the graphs, nodes will represent each treatment modality (eligible acupuncture treatments) and control interventions (sham acupuncture, placebo needling, and conventional drugs), with their diameters proportional to the total sample size of the corresponding interventions. The baseline will be set at 6 mm for 200 patients, with an increase of 1 mm for each additional 50 patients. The connecting lines between nodes indicate direct comparative evidence, and the thickness of lines is positively affected by the number of comparative studies. One study will be set to a line thickness of 1 point, with each additional 2 studies increasing the thickness by 0.3 point, up to a maximum width of 2.5 points to prevent graphical distortion. The absence of a line denotes missing head-to-head comparisons. We will assess network density by calculating the ratio of actual direct comparisons to the theoretical maximum possible number to quantify network completeness.

Heterogeneity

The *Q* test and I^2 statistics will be used to evaluate heterogeneity across the included studies. Acknowledging the limitations of I^2, we will emphasize the between-study variance (τ^2) and its square root (τ) to quantify the extent to which true effects vary across studies. Furthermore, to assess the practical implications of this variation in real-world clinical settings, 95% prediction intervals will be calculated when at least 10 studies are available for a given comparison, thereby complementing the 95% CI for the pooled average effect [35]. In cases where substantial heterogeneity is detected, we will perform subgroup analyses and meta-regression to identify possible contributing factors [36]. Furthermore, node-splitting forest plots will be generated using Stata to visualize variation among direct and indirect evidence for each intervention pair.

Consistency

Consistency analyses enhance the credibility of clinical efficacy rankings. The following strategy will be implemented to ensure consistency among direct and indirect evidence. The node-splitting method will be applied to calculate discrepancies between direct and indirect evidence for each pair of interventions [37]. Local inconsistency will be claimed if the node-splitting *P* value is <.10. Moreover, global network consistency will be estimated using the design-by-treatment

XSL•FO
RenderX

interaction model [38], with the consistency assumption accepted if the global *P* value is >.10. All analyses will follow a frequentist framework and report inconsistency factors with 95% CIs.

Subgroup Analysis and Meta-Regression

The following variables will be considered in the subgroup analyses: participant characteristics (eg, chemotherapy regimens, cycle sequence, and stage of breast cancer), outcome measures (eg, scale type and clinical treatment rate [effectiveness rate vs CR rate]), and study region (China vs Europe and America).

Covariates involved in the meta-regression will include study sample size, average age of patients, number of acupuncture treatments, and others.

Small-Study Effects and Publication Bias

In complex intervention networks, small-study effects and publication bias may lead to distorted efficacy rankings. We will assess potential small-study effects using comparison-adjusted funnel plots and statistical tests. When at least 10 studies are included, effect sizes will be calculated and comparison-adjusted funnel plots will be generated using Stata [39]. Funnel plot asymmetry will be interpreted as an indication of potential small-study effects rather than definitive evidence of publication bias, as publication bias is only one of several possible explanations, including clinical heterogeneity and methodological limitations [40,41]. To further investigate these effects, the Egger test will be performed. If small-study effects are detected, we will conduct sensitivity analyses to evaluate the robustness of our findings [42].

Quality of Evidence

The rigor of included evidence will be evaluated using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) framework [43]. We will systematically evaluate 5 critical areas (risk of bias, indirectness, inconsistency, imprecision, and publication bias) through the GRADEpro GDT (Evidence Prime Inc) [44]. Each study will be classified into 4 grades of quality: high, moderate, low, or very low. Any discrepancies in grading will be settled through iterative discussions until mutual agreement is achieved. The final evidence profiles will be synthesized into summary of findings tables, explicitly documenting downgrading rationales for each assessed domain to ensure transparency in quality judgments. The final evidence will be reported in a format that specifies the reasons for downgrading in each area to ensure the reliability of the quality assessment.

Results

A preliminary scoping search was concluded in August 2025 to refine the research question. This protocol was finalized in October 2025. The comprehensive literature search and study selection are expected to be completed by June 2026, followed by data extraction by August 2026. Data synthesis and final manuscript preparation are scheduled to be completed by December 2026. The results will be disseminated through open-access publications.

Discussion

Overview

As one of the most prevalent malignancies globally, breast cancer requires focused attention on the management of chemotherapy-related side effects, particularly CINV, a common complication that significantly affects quality of life. Acupuncture's therapeutic value has been extensively demonstrated in international clinical practice and has gained increasing patient acceptance. A growing number of clinicians, particularly TCM physicians and acupuncturists, have focused on this specialized therapy for CINV in patients with breast cancer.

We note that several published meta-analyses have investigated acupuncture for CINV in breast cancer. Issac et al [45] demonstrated the superiority of acupressure in controlling acute nausea, delayed nausea, and delayed vomiting. Jang et al [46] examined acupuncture's effects on a variety of chemotherapy-related complications in breast cancer, with 2 trials addressing nausea and vomiting management. However, the limitation of these 2 studies lies in their narrow focus on the efficacy of isolated acupoint therapies. The analysis by Wang et al [47] broadly compared nonpharmacological interventions for CINV in all tumors with some acupuncture modalities. In contrast, our study exclusively evaluates multiple acupuncture therapies in patients with breast cancer and incorporates more comprehensive outcomes. This deliberate exclusion of herb-assisted techniques (eg, acupoint patch) eliminates potential confounding from pharmaceutical effects, allowing for a clearer interpretation of traditional acupuncture's intrinsic efficacy in this specific population.

While various acupuncture methods have been widely implemented in China for managing CINV in patients with breast cancer, international clinical practice, especially in Europe and the United States, remains dominated by electroacupuncture, auricular acupuncture, and acupressure, with limited use of some techniques such as fire needling, abdominal acupuncture, and intradermal needling. Emerging evidence from Chinese clinical trials has demonstrated the effects of these techniques [48,49].

Strengths and Limitations

The primary strength of our research lies in the systematic comparison of multiple acupuncture modalities for mitigating CINV in patients with breast cancer using NMA, with the aim of providing clinicians with a broader range of evidence-based therapeutic selection. This protocol might also facilitate the use of acupuncture and moxibustion for CINV in other malignancies or for other side effects generated by chemotherapy.

Some potential limitations and difficulties may arise from the conduct of this research. First, the 9 acupuncture modalities included exhibit clinical variability (eg, differences in stimulation intensity), compounded by diverging outcome assessment criteria between Chinese and Western studies, which may cause significant heterogeneity. Second, despite using unrestricted geographic or ethnic search strategies, we observed that the literature search reveals a predominance of Asian

XSL•FO
RenderX

studies, likely attributable to regional variations in acupuncture acceptance. This phenomenon will raise concerns about potential biases stemming from ethnicity-related genetic and cultural factors. Finally, nowadays, most RCTs tend to adopt the design of "acupuncture combined drug vs drug" [50,51], which may create insufficient direct evidence comparing acupuncture monotherapy with pharmacological interventions and limit assessment of acupuncture's independent efficacy. Future trials are suggested to adopt 3-arm designs (acupuncture alone, drug alone, and combination therapy) to clarify absolute therapeutic benefits.

Conclusions

This protocol outlines the methodology for a systematic review and NMA that will comprehensively compare the efficacy of multiple acupuncture modalities for managing CINV in patients with breast cancer. By integrating both conventional Western and TCM outcome indicators, this study is positioned to generate nuanced, clinically relevant evidence. The findings will not only expand the range of evidence-based nonpharmacological options available to clinicians but also facilitate the application of acupuncture in supportive oncology care.

Funding

This study was supported by the Medical Health Science and Technology Project of Zhejiang Provincial Health Commission (grant 2022KY232), the Natural Science Foundation of Zhejiang Province (grant LY21H280002), and the Zhejiang Province Traditional Chinese Medicine Science and Technology Project (grant 2024ZF072).

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Authors' Contributions

YL and QW conceptualized the protocol, designed the methodology, and developed the analytical framework. TL, ZB, and HZ conducted the literature search. YL, ZB, TL, and HZ are responsible for study selection and data extraction. HZ, QW, and YL supervised the protocol development. YL and HZ drafted the initial manuscript. All authors critically revised the manuscript for intellectual content and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PubMed search strategy.
[DOCX File, 17 KB - resprot_v15i1e86384_app1.docx]

Checklist 1
PRISMA-P 2015 checklist.
[PDF File, 231 KB - resprot_v15i1e86384_app2.pdf]

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin 2025;75(1):10-45. [doi: 10.3322/caac.21871] [Medline: 39817679]
2. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74(3):229-263. [doi: 10.3322/caac.21834] [Medline: 38572751]
3. Wang J, Wu SG. Breast cancer: an overview of current therapeutic strategies, challenge, and perspectives. Breast Cancer (Dove Med Press) 2023;15:721-730. [doi: 10.2147/BCTT.S432526] [Medline: 37881514]
4. Akram M, Iqbal M, Daniyal M, Khan AU. Awareness and current knowledge of breast cancer. Biol Res 2017 Oct 2;50(1):33. [doi: 10.1186/s40659-017-0140-9] [Medline: 28969709]
5. Kavyasudha K, Afrin S, Chandra YP. Pharmacological management of chemotherapy of breast cancer - includes side effects. Int J Appl Res 2025;11(10):15-25. [doi: 10.22271/allresearch.2025.v11.i10a.12911]
6. Wiser W, Berger A. Practical management of chemotherapy-induced nausea and vomiting. Oncology (Williston Park) 2005 Apr;19(5):637-645. [Medline: 15945344]
7. Tian SC, Yang J, Li X, Huang RX, Chen J. Bibliometric and visual analysis of chemotherapy-induced nausea and vomiting (2004-2023). Front Oncol 2024;14:1377486. [doi: 10.3389/fonc.2024.1377486] [Medline: 38720800]
8. Chen W, Zhao Y, Dai Y, Nie K. Gastrointestinal inflammation plays a critical role in chemotherapy-induced nausea and vomiting. Eur J Pharmacol 2022 Dec 5;936:175379. [doi: 10.1016/j.ejphar.2022.175379] [Medline: 36356927]

XSL•FO
RenderX

9. Gupta K, Walton R, Kataria SP. Chemotherapy-induced nausea and vomiting: pathogenesis, recommendations, and new trends. Cancer Treat Res Commun 2021;26:100278. [doi: 10.1016/j.ctarc.2020.100278] [Medline: 33360668]
10. Popovic M, Warr DG, Deangelis C, et al. Efficacy and safety of palonosetron for the prophylaxis of chemotherapy-induced nausea and vomiting (CINV): a systematic review and meta-analysis of randomized controlled trials. Support Care Cancer 2014 Jun;22(6):1685-1697. [doi: 10.1007/s00520-014-2175-6] [Medline: 24590374]
11. Hesketh PJ, Grunberg SM, Gralla RJ, et al. The oral neurokinin-1 antagonist aprepitant for the prevention of chemotherapy-induced nausea and vomiting: a multinational, randomized, double-blind, placebo-controlled trial in patients receiving high-dose cisplatin--the Aprepitant Protocol 052 Study Group. J Clin Oncol 2003 Nov 15;21(22):4112-4119. [doi: 10.1200/JCO.2003.01.095] [Medline: 14559886]
12. Ioannidis JP, Hesketh PJ, Lau J. Contribution of dexamethasone to control of chemotherapy-induced nausea and vomiting: a meta-analysis of randomized evidence. J Clin Oncol 2000 Oct 1;18(19):3409-3422. [doi: 10.1200/JCO.2000.18.19.3409] [Medline: 11013282]
13. Jordan K, Gralla R, Jahn F, Molassiotis A. International antiemetic guidelines on chemotherapy induced nausea and vomiting (CINV): content and implementation in daily routine practice. Eur J Pharmacol 2014 Jan 5;722:197-202. [doi: 10.1016/j.ejphar.2013.09.073] [Medline: 24157984]
14. Zhao R, Li H, Li X. Overview of traditional Chinese medicine in prevention and treatment of gastrointestinal reactions after chemotherapy. Chin J Integr Trad West Med Dig 2021;29(10):749-752. [doi: 10.3969/j.issn.1671-038X.2021.10.15]
15. Tan JY, Molassiotis A, Suen LK, Liu J, Wang T, Huang HR. Effects of auricular acupressure on chemotherapy-induced nausea and vomiting in breast cancer patients: a preliminary randomized controlled trial. BMC Complement Med Ther 2022 Mar 24;22(1):87. [doi: 10.1186/s12906-022-03543-y] [Medline: 35331208]
16. Dibble SL, Luce J, Cooper BA, et al. Acupressure for chemotherapy-induced nausea and vomiting: a randomized clinical trial. Oncol Nurs Forum 2007 Jul;34(4):813-820. [doi: 10.1188/07.ONF.xxx-xxx] [Medline: 17723973]
17. Suh EE. The effects of P6 acupressure and nurse-provided counseling on chemotherapy-induced nausea and vomiting in patients with breast cancer. Oncol Nurs Forum 2012 Jan;39(1):E1-E9. [doi: 10.1188/12.ONF.E1-E9] [Medline: 22201663]
18. Genç F, Tan M. The effect of acupressure application on chemotherapy-induced nausea, vomiting, and anxiety in patients with breast cancer. Palliat Support Care 2015 Apr;13(2):275-284. [doi: 10.1017/S1478951514000248] [Medline: 24787745]
19. Vallim ET, Marques AC, Coelho RC, Guimarães PR, Felix JV, Kalinke LP. Auricular acupressure in the quality of life of women with breast cancer: a randomized clinical trial [Article in Portuguese, English, Spanish]. Rev Esc Enferm USP 2019;53:e03525. [doi: 10.1590/S1980-220X2018043603525] [Medline: 31800817]
20. Ades AE, Welton NJ, Dias S, Phillippo DM, Caldwell DM. Twenty years of network meta-analysis: continuing controversies and recent developments. Res Synth Methods 2024 Sep;15(5):702-727. [doi: 10.1002/jrsm.1700] [Medline: 38234221]
21. Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. BMJ 2013 May 14;346:f2914. [doi: 10.1136/bmj.f2914] [Medline: 23674332]
22. Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. Res Synth Methods 2012 Jun;3(2):80-97. [doi: 10.1002/jrsm.1037] [Medline: 26062083]
23. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015 Jan 2;350:g7647. [doi: 10.1136/bmj.g7647] [Medline: 25555855]
24. Stewart LA, Clarke M, Rovers M, et al. Preferred reporting items for systematic review and meta-analyses of individual participant data: the PRISMA-IPD statement. JAMA 2015 Apr 28;313(16):1657-1665. [doi: 10.1001/jama.2015.3656] [Medline: 25919529]
25. Beith JM, Oh B, Chatfield MD, Davis E, Venkateswaran R. Electroacupuncture for nausea, vomiting, and myelosuppression in women receiving adjuvant chemotherapy for early breast cancer: a randomized controlled pilot trial. Med Acupunct 2012;24(4):241-248. [doi: 10.1089/acu.2012.0876]
26. Wicaksono B, Hadian FI, Firdaus AD, Ulhaq ZS. The effect of digital acupressure on chemotherapy-induced nausea in Indonesian patients with stages III-IV breast cancer. J Ayurveda Integr Med 2023;14(5):100794. [doi: 10.1016/j.jaim.2023.100794] [Medline: 37688948]
27. Shen G, Ren D, Zhao F, et al. Effect of adding electroacupuncture to standard triple antiemetic therapy on chemotherapy-induced nausea and vomiting: a randomized controlled clinical trial. J Clin Oncol 2024 Dec;42(34):4051-4059. [doi: 10.1200/JCO.24.00099] [Medline: 39241208]
28. Guo Q, Yao C, Guo YF, Wang M. Clinical efficacy and action mechanism of mild moxibustion combined with salt-separated moxibustion for gastrointestinal response in breast-cancer chemotherapy patients. Zhongguo Zhen Jiu 2020 Jan 12;40(1):8-12 [FREE Full text] [doi: 10.13703/j.0255-2930.20190116-k0003] [Medline: 31930891]
29. Dueck AC, Mendoza TR, Mitchell SA, et al. Validity and reliability of the US National Cancer Institute's patient-reported outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). JAMA Oncol 2015 Nov;1(8):1051-1059. [doi: 10.1001/jamaoncol.2015.2639] [Medline: 26270597]
30. He P, Pan G, Wang X, et al. A prospective randomized controlled study of "Old Ten Needles" for the prevention and treatment of chemotherapy-related nausea and vomiting in breast cancer [Article in Chinese]. Chin J Trad Chin Med 2017(6) [FREE Full text]

XSL•FO
RenderX

31. Sterne JA, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
32. IntHout J, Ioannidis JP, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol 2014 Feb 18;14:25. [doi: 10.1186/1471-2288-14-25] [Medline: 24548571]
33. Li Y, Zhang L, Wang Q, Deng Y. Acupuncture for gastrointestinal urticaria: a protocol for systematic review and network meta-analysis. Front Med (Lausanne) 2022;9:998655. [doi: 10.3389/fmed.2022.998655] [Medline: 36314007]
34. Shi J, Li P, Wu Y, Li J, Zhang Y, Xiao B. Efficacy of electroacupuncture in the treatment of mild to moderate female stress urinary incontinence: protocol for a systematic review and network meta-analysis. JMIR Res Protoc 2024 Nov 4;13:e55870. [doi: 10.2196/55870] [Medline: 39496309]
35. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010 Apr;1(2):97-111. [doi: 10.1002/jrsm.12] [Medline: 26061376]
36. Xu T, Deng B, Lin S, Wang H, Xu Y, Chen G. Comparative efficacy of acupuncture-related therapy for postmenopausal osteoporosis: protocol for Bayesian network meta-analysis. BMJ Open 2023 Dec 30;13(12):e074740. [doi: 10.1136/bmjopen-2023-074740] [Medline: 38159952]
37. Belavy DL, Diwan AD, Ford J, et al. Network meta-analysis for comparative effectiveness of treatments for chronic low back pain disorders: systematic review protocol. BMJ Open 2021 Nov 29;11(11):e057112. [doi: 10.1136/bmjopen-2021-057112] [Medline: 34845083]
38. Efthimiou O, Debray TP, van Valkenhoef G, et al. GetReal in network meta-analysis: a review of the methodology. Res Synth Methods 2016 Sep;7(3):236-263. [doi: 10.1002/jrsm.1195] [Medline: 26754852]
39. Chen H, Wang Y, Ge S, Li W, Li J, Chen W. The effects of major dietary patterns on patients with type 2 diabetes: protocol for a systematic review and network meta-analysis. PLoS One 2024;19(6):e0306336. [doi: 10.1371/journal.pone.0306336] [Medline: 38941329]
40. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011 Jul 22;343:d4002. [doi: 10.1136/bmj.d4002] [Medline: 21784880]
41. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. BMJ 2006 Sep 16;333(7568):597-600. [doi: 10.1136/bmj.333.7568.597] [Medline: 16974018]
42. Xiong J, Li H, Hua F, et al. Acupuncture for mammary gland hyperplasia: a Bayesian network meta-analysis protocol. Medicine (Baltimore) 2020 Sep 4;99(36):e22055. [doi: 10.1097/MD.0000000000022055] [Medline: 32899068]
43. Puhan MA, Schünemann HJ, Murad MH, et al. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. BMJ 2014 Sep 24;349:g5630. [doi: 10.1136/bmj.g5630] [Medline: 25252733]
44. Izcovich A, Chu DK, Mustafa RA, Guyatt G, Brignardello-Petersen R. A guide and pragmatic considerations for applying GRADE to network meta-analysis. BMJ 2023 Jun 27;381:e074495. [doi: 10.1136/bmj-2022-074495] [Medline: 37369385]
45. Issac A, Nayak SG, Halemani K, Mishra P, Chand G. Acupressure on chemotherapy-induced nausea and vomiting among breast cancer patients: a systematic review and meta-analysis. Asian Pac J Cancer Prev 2024 Oct 1;25(10):3421-3428. [doi: 10.31557/APJCP.2024.25.10.3421] [Medline: 39471007]
46. Jang S, Ko Y, Sasaki Y, et al. Acupuncture as an adjuvant therapy for management of treatment-related symptoms in breast cancer patients: systematic review and meta-analysis (PRISMA-compliant). Medicine (Baltimore) 2020 Dec 11;99(50):e21820. [doi: 10.1097/MD.0000000000021820] [Medline: 33327222]
47. Wang Y, Wang F, Hu F, Long F. Effect of non-pharmacological interventions on chemotherapy induced delayed nausea and vomiting for tumors: a systematic review and Bayesian network meta-analysis. Complement Ther Med 2025 Mar;88:103124. [doi: 10.1016/j.ctim.2024.103124] [Medline: 39798818]
48. Li M, Ma Q, Zhu H, Li X. Influence of intradermal needle buied in Neiguan acupoint on nausea and vomiting of breast cancer patients undergoing chemotherapy. Chin J Integr Tradit West Med 2018;38(8):921-926. [doi: 10.7661/j.cjim.20180616.026]
49. Zheng W. Influence of intradermal needle buied in Neiguan acupoint on nausea and vomiting of breast cancer patients undergoing chemotherapy. Chin Nurs Res 2012;26(28):2641-2642. [doi: 10.3969/j.issn.1009-6493.2012.28.021]
50. Molassiotis A, Helin AM, Dabbour R, Hummerston S. The effects of P6 acupressure in the prophylaxis of chemotherapy-related nausea and vomiting in breast cancer patients. Complement Ther Med 2007 Mar;15(1):3-12. [doi: 10.1016/j.ctim.2006.07.005] [Medline: 17352966]
51. Shen J, Wenger N, Glaspy J, et al. Electroacupuncture for control of myeloablative chemotherapy-induced emesis: a randomized controlled trial. JAMA 2000 Dec 6;284(21):2755-2761. [doi: 10.1001/jama.284.21.2755] [Medline: 11105182]

Abbreviations

5-HT: 5-hydroxytryptamine
CINV: chemotherapy-induced nausea and vomiting
CR: complete response
GRADE: Grading of Recommendations, Assessment, Development, and Evaluation

XSL•FO
RenderX

MeSH: Medical Subject Headings
NK-1: neurokinin-1
NMA: network meta-analysis
PICOS: population, intervention, comparison, outcome, and study
PR: partial response
PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
RA: receptor antagonist
RCT: randomized controlled trial
TCM: traditional Chinese medicine
WHO: World Health Organization

Edited by A Schwartz; submitted 23.Oct.2025; peer-reviewed by T Liu; revised version received 14.Mar.2026; accepted 17.Mar.2026; published 21.Apr.2026.

Please cite as:
Luo Y, Bian Z, Wang Q, Lei T, Zhao H
Acupuncture Therapies for Chemotherapy-Induced Nausea and Vomiting in Patients With Breast Cancer: Protocol for a Systematic Review and Network Meta-Analysis
JMIR Res Protoc 2026;15:e86384
URL: https://www.researchprotocols.org/2026/1/e86384
doi:10.2196/86384

© Yutao Luo, Zhujun Bian, Qi Wang, Tiantian Lei, Hong Zhao. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy of Nutritional Interventions in Adult Poststroke Rehabilitation: Protocol for a Systematic Review and Trial Sequential Meta-Analysis

Chen Chen*, MD; Su-xiang Zhang*, MD; Shao-hua Chen*, MD; Yan Cao, MD; Ji-ming Tao, PhD

Department of Rehabilitation Medicine, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, 185 Pu 'an Road, Huangpu District, Shanghai, China

*these authors contributed equally

Corresponding Author:
Ji-ming Tao, PhD
Department of Rehabilitation Medicine, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, 185 Pu 'an Road, Huangpu District, Shanghai, China

Abstract

Background: Malnutrition is a prevalent and serious concern in adult survivors of stroke, often worsening their clinical status and leading to a poor prognosis. However, the effectiveness of nutritional interventions in this population remains uncertain. Furthermore, the impact of nutritional support on functional recovery, especially in relation to rehabilitation outcomes, has not been sufficiently investigated.

Objective: This meta-analysis aims to synthesize evidence from existing randomized controlled trials (RCTs) to determine whether nutrient supplements confer measurable benefits in poststroke rehabilitation, with overall mortality as the primary outcome and multiple relevant indicators as secondary outcomes.

Methods: RCTs that compare nutrient supplements and a standard diet or placebo supplements in adult survivors of stroke will be included. Literature searches will be conducted in the PubMed, Web of Science, Embase, China National Knowledge Infrastructure, Wanfang, and Cochrane Library databases. Two reviewers will independently perform the processes of literature retrieval, screening, data extraction, and assessment of risk of bias. Risk of bias in included studies will be evaluated using the revised Cochrane risk-of-bias tool for RCTs. Review Manager will be used for data pooling. Subgroup analysis, trial sequential analysis, and sensitivity analysis will be conducted.

Results: This study was funded in 2023 (2023ZDFC0301) and registered in PROSPERO (CRD420251028426) prior to initiation. As of November 25, 2025, literature screening has not yet started. The planned timeline is as follows: complete literature screening by February 28, 2026; finalize full-text assessment and data extraction by May 30, 2026; conclude trial sequential analysis by June 30, 2026; and complete manuscript preparation for publication by July 30, 2026.

Conclusions: This systematic review and meta-analysis will provide evidence for the efficacy of nutritional interventions in adult poststroke rehabilitation, which may inform clinical guidelines and optimize rehabilitation strategies for this vulnerable population.

Trial Registration: PROSPERO CRD420251028426; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251028426

International Registered Report Identifier (IRRID): DERR1-10.2196/79478

(JMIR Res Protoc 2026;15:e79478) doi:10.2196/79478

KEYWORDS

stroke; nutrition; systematic review; rehabilitation; randomized controlled trial

Introduction

Stroke remains a leading cause of long-term disability and mortality worldwide, with survivors often facing significant physical and cognitive impairments that hinder functional recovery. Poststroke rehabilitation is crucial for improving patient outcomes, yet recovery can be complicated by malnutrition, muscle wasting, and secondary infections, which further exacerbate disability and increase mortality risk [1-6]. Nutritional deficiencies, particularly in protein and essential micronutrients, are common among survivors of stroke due to dysphagia, reduced appetite, and metabolic stress, potentially delaying recovery and worsening neurological deficits [7-12]. Given these challenges, nutrient supplementation has emerged as a promising adjuvant therapy to enhance rehabilitation by addressing malnutrition, supporting muscle synthesis, and improving immune function.

XSL•FO
RenderX

Current evidence suggests that nutrient supplements—including proteins, amino acids, vitamins, and minerals—may positively influence stroke recovery by mitigating catabolism, enhancing neurological repair, and reducing infection rates [13,14]. However, findings across randomized controlled trials (RCTs) remain inconsistent, with variations in supplement type, dosage, and patient populations leading to conflicting results [15,16]. While some studies report significant improvements in muscle strength, mobility, and biochemical markers, others show negligible effects on mortality or functional outcomes [17-22]. These discrepancies highlight the need for a comprehensive meta-analysis to evaluate the efficacy of nutrient supplementation in adult survivors of stroke, particularly in comparison with standard dietary care or placebo.

This meta-analysis aims to synthesize evidence from existing RCTs to determine whether nutrient supplements confer measurable benefits in poststroke rehabilitation. The primary outcomes include overall mortality, while the secondary outcomes encompass infection rates; neurological function (National Institutes of Health Stroke Scale [NIHSS] score); nutritional status (albumin, hemoglobin, and protein intake); and physical performance (handgrip strength and 6-minute walking distance). By clarifying the role of nutritional interventions in stroke recovery, this study may inform clinical guidelines and optimize rehabilitation strategies for this vulnerable population.

Methods

Protocol and Registration

In adherence to the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) guidelines, we meticulously reported this systematic review and meta-analysis. Prior to initiation, the protocol for this comprehensive study was formally registered in the PROSPERO database, ensuring transparency and rigor in our research methodology [23] (registration ID CRD420251028426). This protocol describes a future systematic review. As of November 25, 2025, no screening or data extraction has been initiated and no results exist. The team will complete literature screening by February 28, 2026; finalize full-text assessment and data extraction by May 30, 2026; conclude trial sequential analysis by June 30, 2026; and complete manuscript preparation by July 30, 2026.

Inclusion Criteria

Participants

Participants include adults (aged ≥18 years) with confirmed stroke (including ischemic and hemorrhagic stroke diagnosed via computed tomography or magnetic resonance imaging). Stroke severity is not restricted (mild to severe; NIHSS score 0‐42). Patients with or without poststroke complications (eg, dysphagia, muscle weakness, and infections) are included, except for those meeting the exclusion criteria.

Intervention

The intervention of interest is defined as nutritional supplements, which refer to oral, enteral, or parenteral preparations containing one or more nutrients (eg, protein, amino acids, vitamins, minerals, and fatty acids) designed to supplement dietary intake. To be included, these supplements must be administered for a minimum of 2 weeks and a maximum of 6 months. It is important to note that this definition specifically encompasses nutrient supplements, while broader dietary intervention types (eg, high-protein diets and the Mediterranean diet) are excluded. For comparison, the control group comprises patients who receive only a standard diet or placebo supplements.

Primary Outcome

The primary outcome is overall mortality, defined as all-cause mortality during the study follow-up period. The primary time point for assessment is 3 months after the initiation of the nutritional intervention; if multiple follow-up time points are available (eg, 6 months or 1 year), data at the longest available follow-up will be prioritized.

Secondary Outcomes

Secondary outcomes are described below.

The incidence of overall infection is defined as the occurrence of any infection (eg, pneumonia, urinary tract infection, and sepsis) during the intervention period. The assessment time point will be throughout the nutritional intervention period (minimum 2 weeks and maximum 6 months).

The NIHSS is a validated tool for measuring neurological deficit severity (score range: 0‐42, higher scores indicate more severe deficits). Acceptable measurement tools are the original NIHSS or standardized translated versions. A clinically meaningful change will be a reduction of ≥4 points. The assessment time point will be at baseline and after the intervention (3 months).

Albumin (g/L) will be measured via venous blood sampling. A clinically meaningful change will be an increase of ≥3 g/L from baseline. The assessment time point will be at baseline and after the intervention (3 months).

Hemoglobin (g/L) will be measured via venous blood sampling. A clinically meaningful change will be an increase of ≥10 g/L from baseline. The assessment time point will be at baseline and after the intervention (3 months).

Daily protein intake (g/d) will be assessed via 24-hour dietary recall, food frequency questionnaire, or dietary records. Acceptable measurement tools are validated dietary assessment instruments. The assessment time point will be at baseline and after the intervention (3 months).

Handgrip strength (kg) will be measured using a dynamometer (eg, Jamar dynamometer) with the dominant hand, following standard operating procedures (patient seated, elbow flexed at 90°, and forearm neutral). A clinically meaningful change is an increase of ≥3 kg. The assessment time point will be at baseline and after the intervention (3 months).

The 6-minute walking test (m) is measured as the maximum distance walked in 6 minutes on a flat surface, following American Thoracic Society guidelines. A clinically meaningful change is an increase of ≥54 m. The assessment time point will be at baseline and after the intervention (3 months).

XSL•FO
RenderX

Study Design

We will include RCTs. Any other literature, including nonrandomized clinical controlled trials, retrospective research studies, conference abstracts, case reports, duplicate publications, and literature without data will not be considered.

Eligibility Criteria

Inclusion and exclusion criteria are described in Textbox 1. Crossover studies with an adequate washout period (≥2 weeks) and separable data will be included and analyzed using the generic inverse variance model.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Adults (aged ≥18 years) with confirmed stroke (ischemic or hemorrhagic) diagnosed via computed tomography or magnetic resonance imaging
- Stroke severity of any grade (National Institutes of Health Stroke Scale score 0-42)
- Patients with or without poststroke complications (eg, dysphagia, muscle weakness, and infections)
- Nutritional supplement intervention (oral, enteral, or parenteral) containing one or more nutrients, administered for 2 weeks to 6 months
- Control group receiving standard diet or placebo supplements
- Randomized controlled trial study design
- Outcome data including at least one predefined primary or secondary outcome

Exclusion criteria

- Pregnant or lactating women
- Patients with cancer (due to metabolic disturbances and confounding from anticancer treatments)
- Participation in other clinical studies within 3 months before the study
- Complete aphasia with inability to complete outcome assessments and persistent unconsciousness (Glasgow Coma Scale score <8 for >72 hours after stroke)
- Patients who refuse nutritional intervention or follow-up
- Crossover studies with washout period <2 weeks (insufficient to eliminate residual effects of the first intervention) or for which outcome data cannot be separated by the intervention period

Literature Sources and Retrieval Strategy

To ensure accuracy and completeness, 2 reviewers (CC and YC) will independently conduct a comprehensive search of articles across 7 databases and 1 clinical trial registry: PubMed, Web of Science, Embase, China National Knowledge Infrastructure, Wanfang, Chinese Biomedical Literature Database (SinoMed), the Cochrane Library, and the International Clinical Trials Registry Platform. The International Clinical Trials Registry Platform search will focus on completed but unpublished RCTs related to nutritional interventions in poststroke rehabilitation. The detailed retrieval strategy (combining Medical Subject Headings terms and free text keywords) is outlined in Table 1.

Table . The detailed retrieval strategy (combining Medical Subject Headings [MeSH] terms and free text keywords).

Step	Search term (MeSH+free text)	Boolean operation
1	Stroke [MeSH] OR cerebrovascular apoplexy OR intracranial arteriosclerosis OR intracranial embolism OR brain infarction OR cerebral infarction OR cerebral hemorrhage OR ischaemic stroke OR haemorrhagic stroke	OR
2	Nutritional supplements [MeSH] OR nutrition OR energy OR calorie OR carbohydrate OR protein OR amino acid OR fat OR fatty acid OR vitamin OR mineral OR branched-chain amino acid OR whey protein OR omega-3 fatty acid	OR
3	Rehabilitation [MeSH] OR post-stroke rehabilitation OR functional recovery OR muscle strength	OR
4	Step 1 AND Step 2 AND Step 3	AND
5	Filter: randomized controlled trial, humans, and adults (aged ≥18 years)	—[a]

[a]Not applicable.

XSL•FO
RenderX

Literature Screening and Data Extraction

To ensure meticulous accuracy and comprehensive coverage, 2 reviewers (CC and YC) will adhere to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines for rigorous literature screening and data extraction. The screening process will proceed in 2 stages. First, reviewers will independently screen titles and abstracts of retrieved records against the inclusion and exclusion criteria. Records deemed potentially eligible or uncertain will be retained for full-text assessment. Second, reviewers will independently assess the full text of retained records to confirm eligibility. Any discrepancies at either stage will be resolved through discussion; if consensus cannot be reached, a third senior reviewer (JT) will be consulted. The extracted data will include study design, baseline patient information, and statistics on overall mortality, incidence of overall infection, NIHSS score, albumin (g/L), hemoglobin (g/L), daily protein intake (g/d), handgrip strength (kg), and 6-minute walking test (m).

Assessment of Risk of Bias

To ensure accuracy and completeness, 2 reviewers (CC and SC) will independently evaluate the risk of bias in the included studies. They will use the Revised Cochrane risk-of-bias tool (ROB 2) to assess the potential bias in RCTs. Their evaluation will encompass 5 critical dimensions that could impact study quality: bias related to the randomization process, deviations from the intended interventions, missing outcome data, measurement of outcomes, and the selection of reported results. Any discrepancies will be resolved through established protocols.

Ethical Considerations

No ethics approval is required for this review. Our findings will be submitted to peer-reviewed journals.

Statistical Analysis

The Review Manager software (version 5.4; Cochrane Informatics & Technology Services) will be used to aggregate the data and produce forest plots. For dichotomous variables, risk ratios and 95% CIs will be calculated and analyzed using the Mantel-Haenszel method. Continuous variables will be pooled and analyzed as the mean difference with 95% CIs, using the inverse variance method. The statistical significance level (α) is set at .05. Statistical differences are deemed significant if the *P* value <.05. A random effects model will be adopted a priori for all meta-analyses, considering expected clinical heterogeneity (eg, variations in supplement types, doses, durations, and patient stroke types) regardless of I^2 values. The I^2 statistic will be used to quantify statistical heterogeneity: <25% indicates low heterogeneity, 25% to 50% indicates moderate heterogeneity, and >50% indicates high heterogeneity. Publication bias will be assessed only when the number of included studies exceeds 10, as a small number of studies could compromise the robustness of the tests. For missing outcome data, we will first contact the original study authors; if unavailable, we will use multiple imputation or conduct a sensitivity analysis excluding studies with excessive missing data (>20%). To ensure the reproducibility of statistical analyses, we specify the following operational rules: MD will be the primary effect measure for continuous outcomes where all included studies report the same outcome using identical, clinically interpretable units (eg, albumin in g/L, hemoglobin in g/L, handgrip strength in kg, and 6-minute walking distance in m), while standardized mean difference (SMD; Hedges *g* and corrected for small sample size bias) will be used only when studies report the same core outcome using different measurement units or noncomparable validated scales, with SMD effect sizes interpreted per Cohen guidelines (0.2=small effect, 0.5=moderate effect, and 0.8=large effect); for continuous outcomes with convertible nonidentical units, raw data will be uniformly converted to the International System of Units before mean difference calculation, with conversion factors reported and verified by 2 independent reviewers, and SMD will be the sole effect measure for outcomes with nonconvertible units or different measurement scales. For cluster randomized trials, only those with available individual participant data or an adjusted effective sample size accounting for the intracluster correlation coefficient (ICC) will be included, and a conservative ICC of 0.05 (per Cochrane Handbook recommendations) will be used to calculate the effective sample size if the ICC is not reported in the original study, while cluster randomized trials with no available individual participant data, ICC, or raw group-level data will be excluded from the primary meta-analysis and included in a separate sensitivity analysis if raw group-level data are available. For crossover trials meeting the inclusion criteria (washout period ≥2 weeks, separable outcome data by intervention period), only data from the first treatment period will be used for the primary meta-analysis to eliminate carryover effects and avoid double counting of participants, and separable outcome data from the second treatment period (if available) will be used in a prespecified sensitivity analysis to verify the robustness of the primary results. All statistical processing steps, including unit conversion, effect measure selection, and trial design adjustments, will be fully and transparently reported in the final manuscript to ensure complete reproducibility.

Subgroup Analysis and Sensitivity Analysis

Subgroup analyses will be performed based on the following variables to explore potential effect modifiers: participant characteristics, including age (<65 years vs ≥65 years), stroke type (ischemic vs hemorrhagic), diabetes status (with vs without diabetes mellitus), and stroke severity (mild to moderate [NIHSS ≤16] vs severe [NIHSS >16]). Sensitivity analyses for the primary outcome (overall mortality) will include the following methods: (1) excluding studies with a high risk of bias (ROB 2 assessment), (2) excluding studies with a small sample size (<50 participants), (3) excluding studies with excessive missing data (>20%), (4) using a fixed effects model instead of a random effects model to assess the impact of model selection on results, and (5) excluding studies with follow-up duration of less than 6 months to assess the impact of follow-up length on results.

Trial Sequential Analysis

Random errors can lead to spurious conclusions when a meta-analysis involves a limited number of trials and a small patient population. Trial sequential analysis (TSA) will be used to mitigate the risks associated with random errors due to

XSL•FO
RenderX

insufficient sample size or repeated testing and to estimate the required information size for the meta-analysis. We will conduct TSA using the TSA software (version 0.9.5.10 beta). The type I error rate and power will be set at 5% and 80%, respectively.

Certainty of Evidence

Certainty of the evidence will be assessed using GRADEpro, a tool developed by the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working Group [24]. The assessment criteria for certainty include, but are not limited to, the initial study design, risk of bias, imprecision, indirectness, and inconsistency. Following the guidelines, the certainty of the evidence will be categorized as "high," "moderate," "low," or "very low" with the assistance of GRADEpro.

Patient and Public Involvement

Patients and/or the public were not involved in the design, conduct, reporting, or dissemination plans of this research.

Results

This study was funded in 2023 (2023ZDFC0301) and registered in PROSPERO (CRD420251028426) prior to initiation. As of November 25, 2025, literature screening has not yet started. The planned timeline is as follows: complete literature screening by February 28, 2026; finalize full-text assessment and data extraction by May 30, 2026; conclude TSA by June 30, 2026; and complete manuscript preparation for publication by July 30, 2026.

The flowchart of the study selection process is illustrated in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the study inclusion process.

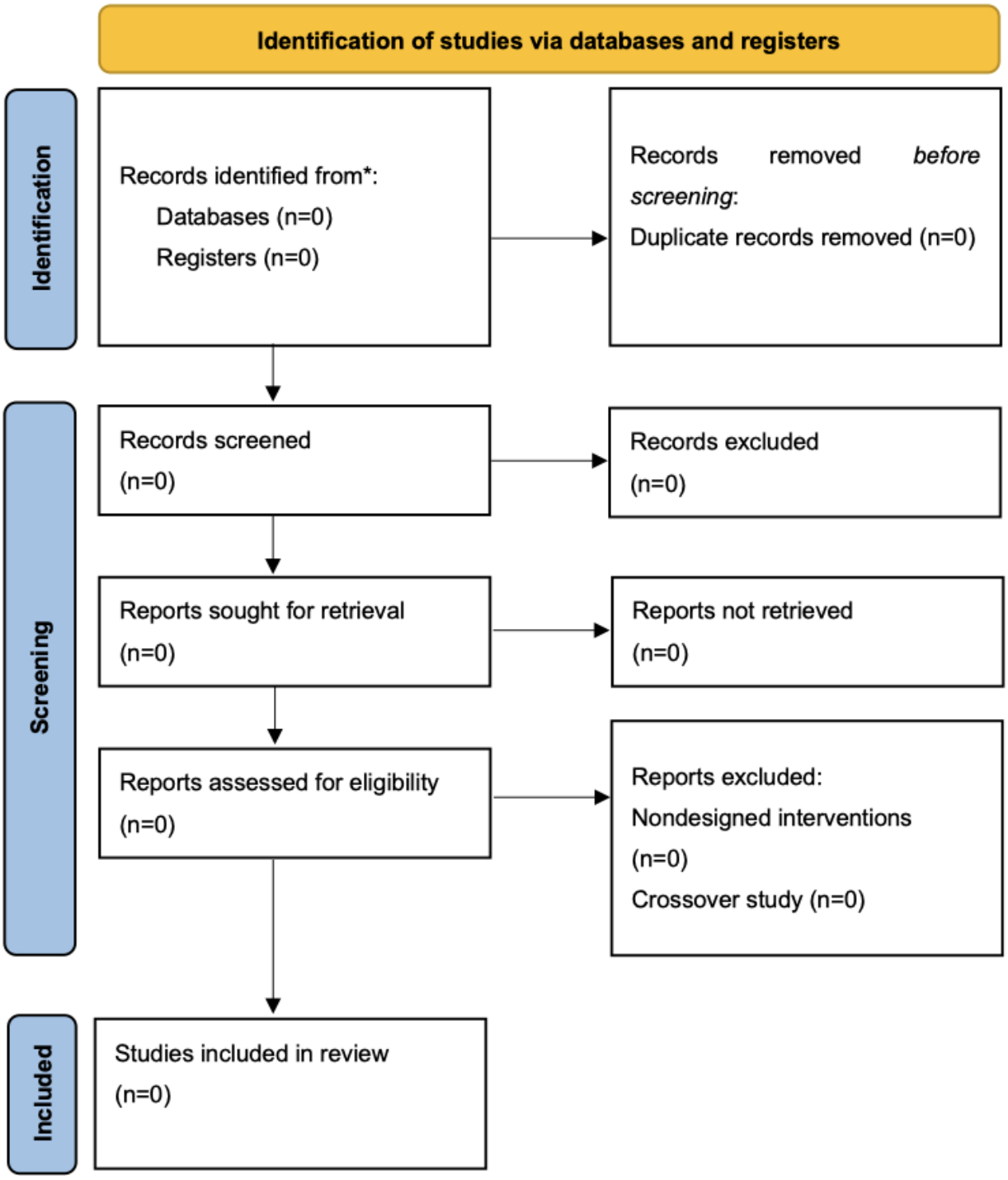

Discussion

Anticipated Findings

This meta-analysis adheres to PRISMA-P guidelines and a preregistered protocol, using rigorous methods to assess nutrient supplementation's role in stroke rehabilitation. A comprehensive search across multiple databases and robust statistical analyses, including trial sequential analysis, will be implemented to ensure reliability. The protocol is specifically designed to minimize bias while maximizing the clinical relevance of our findings for stroke rehabilitation practice.

The study design presents several notable strengths that merit discussion. Our exclusive focus on RCTs provides the highest level of evidence for evaluating intervention efficacy, while the broad inclusion criteria encompassing various nutrient supplements increase the generalizability of findings across different clinical settings. The comprehensive outcome assessment, ranging from clinical end points (mortality) to functional measures (6-minute walk test) and biochemical markers (albumin levels), allows for a multidimensional evaluation of nutritional interventions. Furthermore, the planned subgroup analyses by age and sensitivity analyses enhance our ability to detect potential effect modifiers and verify the

XSL•FO
RenderX

robustness of the results. The incorporation of GRADE methodology for evidence certainty assessment adds further credibility to our conclusions.

Strengths and Limitations of This Study

This systematic review and meta-analysis follows a rigorous methodology with strict inclusion criteria. The use of ROB 2 ensures a rigorous evaluation of study quality. The certainty of the evidence may be limited by heterogeneity in study quality and the small sample sizes of the included trials.

Several methodological considerations also warrant attention. First, while our inclusion of only RCTs enhances internal validity, it may limit generalizability to real-world medical conditions where patient populations are more heterogeneous. The variability in supplement types (eg, proteins and amino acids) and administration protocols across studies introduces clinical heterogeneity that could obscure true treatment effects. Our protocol attempts to address this through subgroup analyses, but these may be limited by the available data. Second, the focus on short-term outcomes (eg, 6-minute walk test and albumin levels) leaves open questions about sustained benefits—a critical gap given stroke's chronic nature. The nutritional field particularly faces measurement challenges; for instance, protein intake reporting often relies on recall methods vulnerable to bias. Third, while trial sequential analysis will be performed to guard against random errors, its utility depends on sufficient sample sizes that may be unrealistic in this research area. These limitations highlight an urgent need for standardized, large-scale trials with longer follow-up periods and objective nutritional biomarkers.

Conclusions and Future Research

Future research should prioritize 3 key directions. Mechanistic studies are needed to clarify how specific nutrients (eg, branched-chain amino acids) influence neuroplasticity and muscle synthesis at the molecular levels, which could explain current inconsistent findings. More trials that compare different supplementation strategies (eg, timing relative to stroke onset and combination with exercise) would help optimize clinical protocols. Most importantly, the field requires the development of validated, stroke-specific nutritional assessment tools—current measures such as handgrip strength and serum markers may not fully capture rehabilitation progress. Emerging technologies such as continuous glucose monitoring or wearable sensors could provide more dynamic nutritional data. There is also growing recognition that stroke recovery may benefit from personalized nutrition approaches based on genetic polymorphisms affecting nutrient metabolism. Finally, integrating nutritional interventions with other rehabilitation modalities represents a promising but underexplored avenue that could yield greater benefits for functional recovery. Addressing these gaps through collaborative, multidisciplinary research will be essential for advancing evidence-based nutritional strategies in stroke care.

Acknowledgments

All content was developed independently by the authors, and artificial intelligence was not used for data analysis or interpretation. The authors attest to the originality and integrity of the manuscript.

Funding

This work was supported by Key Supported Discipline of Shanghai Health and Health System - Rehabilitation Medicine (grant 2023ZDFC0301 awarded to JT).

Data Availability

All data and materials related to this systematic review and meta-analysis are available from the corresponding author upon reasonable request.

Authors' Contributions

CC and JT conceptualized and designed the study. CC drafted the manuscript. CC and YC will retrieve relevant databases. CC and YC will conduct literature screening and data extraction. CC and SC will assess the risk of bias. CC, SC, and JT proofread the manuscript and improved the English language. During the revision stage (in response to peer review comments), SZ (co–first author) made substantial contributions to optimizing the literature search strategy, validating the data extraction framework, and critically revising the manuscript for important intellectual content—including refining the statistical analysis plan and clarifying the PICOTS (population, intervention, comparison, outcome, timing, and setting) framework. CC is responsible for the overall content as guarantor. All authors read and approved the manuscript.

Conflicts of Interest

None declared.

References

1. Labeit B, Michou E, Hamdy S, et al. The assessment of dysphagia after stroke: state of the art and future directions. Lancet Neurol 2023 Sep;22(9):858-870. [doi: 10.1016/S1474-4422(23)00153-9] [Medline: 37596008]

XSL•FO
RenderX

2. Carin-Levy G, Mead GE, Nicol K, Rush R, van Wijck F. Delirium in acute stroke: screening tools, incidence rates and predictors: a systematic review. J Neurol 2012 Aug;259(8):1590-1599. [doi: 10.1007/s00415-011-6383-4] [Medline: 22234842]
3. Balami JS, Chen RL, Buchan AM. Stroke syndromes and clinical management. QJM 2013 Jul;106(7):607-615. [doi: 10.1093/qjmed/hct057] [Medline: 23483140]
4. Huggins HE, Brady M, Emma JP, Thaler DE, Leung LY. Differences in presenting symptoms of acute stroke among young and older adults. J Stroke Cerebrovasc Dis 2020 Aug;29(8):104871. [doi: 10.1016/j.jstrokecerebrovasdis.2020.104871] [Medline: 32689591]
5. Zhu H. A large-scale prospective study on functional ability and mortality following stroke based on HRS data set: implications for rehabilitation. Front Psychol 2022 Jul 11;13:915653. [doi: 10.3389/fpsyg.2022.915653] [Medline: 35899010]
6. Negasa BW, Wotale TW, Lelisho ME, Debusho LK, Sisay K, Gezimu W. Modeling survival time to death among stroke patients at Jimma University Medical Center, Southwest Ethiopia: a retrospective cohort study. Stroke Res Treat 2023 Nov 29;2023:1557133. [doi: 10.1155/2023/1557133] [Medline: 38130889]
7. Hurford R, Sekhar A, Hughes TAT, Muir KW. Diagnosis and management of acute ischaemic stroke. Pract Neurol 2020 Aug;20(4):304-316. [doi: 10.1136/practneurol-2020-002557] [Medline: 32507747]
8. Hollist M, Morgan L, Cabatbat R, Au K, Kirmani MF, Kirmani BF. Acute stroke management: overview and recent updates. Aging Dis 2021 Jul 1;12(4):1000-1009. [doi: 10.14336/AD.2021.0311] [Medline: 34221544]
9. Hasan TF, Hasan H, Kelley RE. Overview of acute ischemic stroke evaluation and management. Biomedicines 2021 Oct 16;9(10):1486. [doi: 10.3390/biomedicines9101486] [Medline: 34680603]
10. Benesch C, Glance LG, Derdeyn CP, et al. Perioperative neurological evaluation and management to lower the risk of acute stroke in patients undergoing noncardiac, nonneurological surgery: a scientific statement from the American Heart Association/American Stroke Association. Circulation 2021 May 11;143(19):e923-e946. [doi: 10.1161/CIR.0000000000000968] [Medline: 33827230]
11. Dávalos A, Ricart W, Gonzalez-Huix F, et al. Effect of malnutrition after acute stroke on clinical outcome. Stroke 1996 Jun;27(6):1028-1032. [doi: 10.1161/01.str.27.6.1028] [Medline: 8650709]
12. Mehta A, De Paola L, Pana TA, et al. The relationship between nutritional status at the time of stroke on adverse outcomes: a systematic review and meta-analysis of prospective cohort studies. Nutr Rev 2022 Nov 7;80(12):2275-2287. [doi: 10.1093/nutrit/nuac034] [Medline: 35640017]
13. Aquilani R, Scocchi M, Boschi F, et al. Effect of calorie-protein supplementation on the cognitive recovery of patients with subacute stroke. Nutr Neurosci 2008 Oct;11(5):235-240. [doi: 10.1179/147683008X301586] [Medline: 18782484]
14. Aquilani R, Scocchi M, Iadarola P, et al. Protein supplementation may enhance the spontaneous recovery of neurological alterations in patients with ischaemic stroke. Clin Rehabil 2008 Dec;22(12):1042-1050. [doi: 10.1177/0269215508094244] [Medline: 19052243]
15. Poppitt SD, Howe CA, Lithander FE, et al. Effects of moderate-dose omega-3 fish oil on cardiovascular risk factors and mood after ischemic stroke: a randomized, controlled trial. Stroke 2009 Nov;40(11):3485-3492. [doi: 10.1161/STROKEAHA.109.555136] [Medline: 19745175]
16. Ha L, Hauge T, Spenning AB, Iversen PO. Individual, nutritional support prevents undernutrition, increases muscle strength and improves QoL among elderly at nutritional risk hospitalized for acute stroke: a randomized, controlled trial. Clin Nutr 2010 Oct;29(5):567-573. [doi: 10.1016/j.clnu.2010.01.011] [Medline: 20176418]
17. Saito G, Zapata R, Rivera R, et al. Long-chain omega-3 fatty acids in aneurysmal subarachnoid hemorrhage: a randomized pilot trial of pharmaconutrition. Surg Neurol Int 2017 Dec 27;8:304. [doi: 10.4103/sni.sni_266_17] [Medline: 29404191]
18. Yoshimura Y, Bise T, Shimazu S, et al. Effects of a leucine-enriched amino acid supplement on muscle mass, muscle strength, and physical function in post-stroke patients with sarcopenia: a randomized controlled trial. Nutrition 2019 Feb;58:1-6. [doi: 10.1016/j.nut.2018.05.028] [Medline: 30273819]
19. Otsuki I, Himuro N, Tatsumi H, et al. Individualized nutritional treatment for acute stroke patients with malnutrition risk improves functional independence measurement: a randomized controlled trial. Geriatr Gerontol Int 2020 Mar;20(3):176-182. [doi: 10.1111/ggi.13854] [Medline: 31854054]
20. Shaafi S, Ebrahimpour-Koujan S, Khalili M, et al. Effects of alpha lipoic acid supplementation on serum levels of oxidative stress, inflammatory markers and clinical prognosis among acute ischemic stroke patients: a randomized, double blind, TNS trial. Adv Pharm Bull 2020 Jun;10(2):284-289. [doi: 10.34172/apb.2020.034] [Medline: 32373498]
21. Honaga K, Mori N, Akimoto T, et al. Investigation of the effect of nutritional supplementation with whey protein and vitamin D on muscle mass and muscle quality in subacute post-stroke rehabilitation patients: a randomized, single-blinded, placebo-controlled trial. Nutrients 2022 Feb 6;14(3):685. [doi: 10.3390/nu14030685] [Medline: 35277045]
22. Wang J, Chang E, Jiang Y. Effects of vitamin C stimulation on rehabilitation of dysphagia after stroke: a randomized trial. Eur J Phys Rehabil Med 2022 Aug;58(4):558-564. [doi: 10.23736/S1973-9087.22.07337-3] [Medline: 35362717]
23. What is PROSPERO? National Institute for Health and Care Research. URL: https://www.crd.york.ac.uk/PROSPERO [accessed 2025-03-25]
24. Streamline evidence-based guidelines with GRADEpro. GRADEpro. URL: https://www.gradepro.org [accessed 2025-03-25]

XSL•FO
RenderX

Abbreviations

GRADE: grading of recommendations assessment, development, and evaluation
ICC: intracluster correlation coefficient
NIHSS: National Institutes of Health Stroke Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
RCT: randomized controlled trial
ROB 2: Revised Cochrane risk-of-bias tool
SMD: standardized mean difference
TSA: trial sequential analysis

Edited by J Sarvestan; submitted 22.Jun.2025; peer-reviewed by KEE Mills; revised version received 12.Feb.2026; accepted 17.Feb.2026; published 24.Mar.2026.

Please cite as:
Chen C, Zhang SX, Chen SH, Cao Y, Tao JM
Efficacy of Nutritional Interventions in Adult Poststroke Rehabilitation: Protocol for a Systematic Review and Trial Sequential Meta-Analysis
JMIR Res Protoc 2026;15:e79478
URL: https://www.researchprotocols.org/2026/1/e79478
doi:10.2196/79478

© Chen Chen, Su-xiang Zhang, Shao-hua Chen, Yan Cao, Ji-ming Tao. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Antenatal Corticosteroids on Brain Function and Underlying Mechanisms in Preclinical Studies: Protocol for a Systematic Review

Mauritz F Herselman[1,2*], BLabMed (Hons), PhD; Tara Hislop[1,2*], BHMS (Hons); Sebastian D McBride[3], BSc, PhD; Mitchell C Lock[1,2], BMedSc, BMedPharmSc (Hons), PhD; Janna L Morrison[1,2], BSc (Hons), MSc, PhD; Kathryn L Gatford[1,2], BAgrSc (Hons), PhD

[1]School of Pharmacy and Biomedical Science, University of Adelaide, Adelaide Health & Medical Sciences building, corner George St and North Terrace, Adelaide, SA, Australia

[2]Robinson Research Institute, College of Health, University of Adelaide, Adelaide, SA, Australia

[3]Department of Life Sciences, Aberystwyth University, Aberyswyth, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Kathryn L Gatford, BAgrSc (Hons), PhD
School of Pharmacy and Biomedical Science, University of Adelaide, Adelaide Health & Medical Sciences building, corner George St and North Terrace, Adelaide, SA, Australia

Abstract

Background: Antenatal corticosteroid (ACS) treatment matures the fetal lung and reduces risks of neonatal morbidity and mortality in babies born preterm. However, ACS treatment also impacts the brain and stress regulatory systems, with increasing clinical evidence for adverse long-term impacts. Preclinical studies are important to investigate the mechanisms for these impacts.

Objective: This systematic review aims to synthesize the best available evidence describing how in utero exposure to ACS affects brain function, the underlying mechanisms, and the hypothalamic-pituitary-adrenal axis in nonhuman mammalian species.

Methods: This review will include peer-reviewed, primary studies that report measures of brain function (eg, learning, behavior) and the underlying mechanisms (eg, brain size, neuron number, myelination, hypothalamic-pituitary-adrenal axis function) in nonhuman mammals exposed in utero to ACS in the second half of pregnancy, in comparison to unexposed individuals. Initial search terms include (corticosteroid* OR glucocorticoid*) AND (antenatal OR fetal* OR pregnan*) AND (brain OR neurodevelopment*). We searched PubMed, Embase, MEDLINE (Ovid), Web of Science, Scopus, and ProQuest for English language publications without date restrictions. Two independent reviewers will perform abstract screening, data extraction, and quality assessment. Antenatal corticosteroid treatment information, study design, methods, and outcomes will be reported for each study.

Results: A narrative synthesis will be presented following standard guidelines. A dose-response meta-analysis will be performed where at least three studies report the same outcome following in utero exposure to the same steroid. As of February 21, 2026, title and abstract screening were completed for 42,024 of 56,493 records, with 40,795 records excluded. Searches will be updated in June-July 2026 to include sources published to the end of 2025. Publication is planned for 2027.

Conclusions: This review will inform future research including intervention studies to reduce the adverse effects of antenatal corticosteroids on the brain.

Trial Registration: PROSPERO CRD420251116681; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251116681

International Registered Report Identifier (IRRID): PRR1-10.2196/84522

(JMIR Res Protoc 2026;15:e84522) doi:10.2196/84522

KEYWORDS

corticosteroid; glucocorticoids; pregnancy; brain; pituitary-adrenal system; gestation; neurodevelopment; animal models

Introduction

Antenatal treatment with corticosteroids (ACS)—discovered in sheep [1], translated rapidly to clinical trials [2], and now standard care for women at risk of preterm delivery [3]—is the only effective antenatal intervention to date for preventing neonatal respiratory distress and improving survival of preterm babies. Maternally delivered ACS rapidly cross the placenta to the fetal circulation, where they activate the glucocorticoid

XSL•FO
RenderX

receptor (GR) in fetal lungs to promote lung maturation, thereby reducing the risk of respiratory distress syndrome by 29% and neonatal death by 22% [3]. The two most commonly used synthetic corticosteroids, dexamethasone and betamethasone, have similar efficacy in preventing respiratory distress syndrome or death of offspring when given to women at risk of preterm birth [4]. However, ACS treatment lacks specificity, resulting in off-target effects, particularly to the developing brain, which also expresses the GR [5-7] and is equally exposed to ACS in the fetal circulation. Clinical concerns center on growing evidence for adverse neurodevelopmental effects of ACS, including greater stress responses and behavioral problems [8,9]. In the largest childhood follow-up to date, mental and behavioral disorders occurred in 12.01% of ACS-exposed individuals compared with 6.45% of nonexposed individuals within the entire cohort of preterm and term-born children [10]. However, as recently reviewed, there is evidence that the effects of ACS exposure vary with gestational age at birth, with benefit in those born preterm and evidence of harm mostly in those born late preterm or at term [11]. For example, exposure to a single course of ACS versus nonexposure was associated with lower odds of neurodevelopmental impairment at 18 - 22 months of age in infants born preterm (adjusted odds ratio 0.69, 95% CI 0.57 - 0.84, low certainty of evidence) [12]. A recent systematic review of outcomes in children born at term identified that early exposure to ACS was associated with greater risks of adverse short-term (neonatal intensive care admission, reduced head circumference, intubation) and long-term outcomes (diagnosis of neurodevelopmental or behavioral disorder), albeit with low or very low certainty, reflecting limited availability of evidence [13]. This group is important, because 40% of ACS-treated mothers do not deliver preterm [13], putting their babies at risk of side effects without respiratory benefit. It is also not clear whether outcomes differ between steroid types, despite no difference in risks of neurosensory disability between children exposed to dexamethasone or betamethasone before birth [14]. Clinical evidence for the mechanistic basis of ACS-mediated changes in neurodevelopment is limited. Although there is moderate-quality evidence that ACS exposure reduces the risk of intraventricular hemorrhage [3], there is limited clinical data on the impacts of ACS exposure on structural or molecular determinants of brain function. Many, but not all, studies identified in a recent scoping review reported reduced head circumference in newborns and children who were exposed to ACS but born late preterm or at term [15]. In two small magnetic resonance imaging studies within this analysis of outcomes in late preterm and term-born offspring, the ACS-exposed group also had a thinner cortex at 6 - 10 years of age and indications of reduced brain gyrification at birth, including lower surface area and cortex convolution index [15]. These structural deficits likely occur along with reprogramming of the hypothalamic-pituitary-adrenal (HPA) axis since exogenous corticosteroids can interfere with negative feedback of this system [16,17]. Indeed, the HPA axis is suppressed in neonates after ACS exposure, which may persist for several months after birth [18]. Longer-term reprogramming of the HPA axis involves alterations in GR expression in the hippocampus and hypothalamus, which may exacerbate stress reactivity and contribute to behavioral disorders in children and adults [17]. For example, ACS exposure increased cortisol responses to psychosocial stress in 6 - 11 year olds born at term, particularly in girls [19].

Preclinically, ACS induces impaired cognition and altered stress responses, as in humans. Neurobehavioral assessments in rodents identified lower activity in open field tests, indications of increased anxiety-like behavior in the elevated plus maze test, more time immobile in forced swim tests, and increased startle responses in those exposed to ACS [20]. Functional impacts of ACS have also been observed in primates, evaluated using the Cambridge Neuropsychological Test Automated Battery system in term-born male and female baboons exposed to repeat courses of clinical doses of betamethasone, compared to offspring of saline-treated mothers [21]. In this study, performance on the moving stimulus task used during touchscreen training, as well as in a simple reversal task, was poorer in betamethasone than control offspring, but only in females [21]. Mechanistically, neurological changes that might underlie altered behavior and cognition after ACS include greater GR expression and lower neuron densities, particularly serotonergic and dopaminergic neurons, with most studies conducted in rodents exposed to high concentrations of dexamethasone [20]. ACS exposure also induces immune and vascular changes in the fetal brain. For example, in preterm-born lambs delivered at gestational day 125 (term ~150 days), exposure to betamethasone was associated with greater cerebral white matter inflammation and vascular permeability [22].

In contrast to the multiple systematic reviews of neurodevelopmental impacts of in utero exposure to ACS in humans, the systematic or scoping reviews describing these outcomes in nonhuman species are of narrower scope and do not fully reflect the clinical context. These include systematic reviews of the impacts of repeat corticosteroid exposure [23], of long-term impacts of ACS, excluding brain changes observed before and shortly after birth [20], and a more recent review of ACS in preclinical models of fetal growth restriction [24]. We therefore plan to conduct a comprehensive and contemporary systematic evaluation of the evidence for impacts of in utero exposure to ACS on the brains of nonhuman mammals. Due to the depth and range of evidence, we will synthesize our results into three reviews, one describing the impacts of ACS on brain function, another discussing the impacts of ACS on underlying mechanisms that may contribute to altered function, and the third describing the impacts of ACS on the HPA axis, including basal and stress-induced function.

Methods

Review Question

What are the impacts of in utero ACS administration on brain function, the underlying mechanisms, and the HPA axis in nonhuman mammalian species?

Search Strategy

An initial search of PubMed and PROSPERO (International Prospective Register of Systematic Reviews) was undertaken to identify articles on the topic and to assess knowledge gaps. A full search strategy was developed using indexed terms and

XSL•FO
RenderX

text words within the titles and abstract (Table 1). This search string was then modified to include syntax specific to each individual database. This search string was validated by checking for the inclusion of a set of 11 key references reporting outcomes spanning brain function, determinants, and HPA axis measures in the output of each search [25-35]. We searched the following electronic databases for published evidence: PubMed, Embase, MEDLINE (Ovid), Web of Science, Scopus, and ProQuest. Searches will be updated in June-July 2026 to identify and include all sources published or accepted for publication to the end of 2025. Publication is planned for 2027.

XSL•FO
RenderX

Table . Search strategy.

Database	Search string
MEDLINE (PubMed)	(Corticosteroid* OR Glucocorticoid* OR Glucocorticoids [mh] OR Betamethasone OR Betamethasone [mh] OR Dexamethasone OR Dexamethasone [mh] OR Cortisol* OR Hydrocortisone* OR Hydrocortisone [mh] OR Corticosterone* OR Corticosterone [mh]) AND (Brain* OR Brain [mh] OR "Brain development" OR Neurodevelopment* OR Neuron* OR Neurons [mh] OR Mental Disorders [mh] OR Intelligence OR Learning [mh] OR Memory [mh] OR Cogniti* OR Cognition [mh] OR Behaviour* OR Behavior* OR Psycholog* OR Anxiety OR Depress* OR Stress* OR "hypothalamo-pituitary-adrenal axis" OR "HPA axis" OR Adrenal glands [mh] OR Adrenocorticotropic hormone [mh] OR Corticotropin-releasing hormone [mh] OR Cortisol OR Corticosterone* OR Corticosterone [mh] OR Hippocamp* OR Hippocampus [mh] OR "Corpus Callosum" OR Corpus Callosum [mh] OR Myelin* OR Neuroglia[mh] OR Oligodendrocyte* OR White matter [mh] OR Gray matter [mh] OR Cerebrum [mh] OR Cerebell* OR Neuroanatom* OR Amygdala* OR Astrocyte*) AND (Antenatal* OR Prenatal* OR Pregnancy Complications [mh] OR Preterm OR Fetus* OR Fetus [mh] OR Foetus* OR Fetal* OR Foetal* OR "Embryonic and Fetal Development" [mh] OR Gestat* OR Pregnan* OR Pregnancy [mh])
MEDLINE (Ovid)	(Corticosteroid*.mp. OR exp Glucocorticoids/ OR Glucocorticoid*.mp. OR Betamethasone.mp. OR exp Betamethasone/ OR Dexamethasone.mp. OR exp Dexamethasone/ OR Cortisol*.mp. OR Hydrocortisone*.mp. OR exp Hydrocortisone/ OR Corticosterone*.mp. OR exp Corticosterone/) AND (Brain*.mp. OR exp Brain/ OR "Brain development".mp. OR Neurodevelopment*.mp. OR Neuron*.mp. OR exp Neurons/ OR exp Mental Disorders/ OR Intelligence.mp. OR exp Learning/ OR exp Memory/ OR Cogniti*.mp. OR exp Cognition/ OR Behavio$r*.mp. OR Psycholog*.mp. OR Anxiety.mp. OR Depress*.mp. OR Stress*.mp. OR hypothalamo-pituitary-adrenal axis.mp. OR HPA axis.mp. OR exp Adrenal glands/ OR exp Adrenocorticotropic hormone/ OR exp Corticotropin-releasing hormone/ OR Cortisol.mp. OR Corticosterone*.mp. OR exp Corticosterone/ OR Hippocamp*.mp. OR exp Hippocampus/ OR Corpus Callosum.mp. OR exp Corpus Callosum/ OR Myelin*.mp. OR exp Neuroglia/ OR Oligodendrocyte*.mp. OR exp White matter/ OR exp Gray matter/ OR exp Cerebrum/ OR Cerebell*.mp. OR Neuroanatom*.mp. OR Amygdala*.mp. OR Astrocyte*.mp.) AND (Antenatal*.mp. OR exp Pregnancy Complications/ OR Prenatal*.mp. OR Preterm.mp. OR F$etus*.mp. OR exp Fetus/ OR F$etal*.mp. OR exp "Embryonic and Fetal Development"/ OR Gestat*.mp. OR Pregnan*.mp. OR exp Pregnancy/)
ProQuest Dissertations & Theses Global	((TI,AB,DISKW(Steroid* OR Corticosteroid* OR cortisol* OR Betamethasone OR Dexamethasone OR Glucocorticoid* OR Corticosterone* OR Hydrocortisone) OR SU(Corticosteroids OR Steroids)) AND (TI,AB,DISKW(Brain OR Neurodevelopment* OR Neuron* OR "Neurological disorder*" OR "Mental disorder*" OR Intelligence OR Learning OR Memory OR Cogniti* OR Recognition OR Behavio$r* OR Psycholog* OR Anxiety OR Depress* OR "Mood disorders" OR Stress* OR "HPA axis" OR "hypothalamo-pituitary-adrenal axis" OR Hypothalamus OR "Adrenal gland*" OR Adrenocorticotropi* OR Corticotropin OR "Corticotropin-releasing" OR Cortisol OR Corticosterone OR Hippocamp* OR "Corpus Callosum" OR Myelin* OR Neuroglia* OR Glia* OR Oligodendro* OR "White matter" OR "Gr?y matter" OR Cerebr* OR Cerebell* OR Neuroanatom* OR Amygdala* OR Astrocyt* OR Neurotransm*) OR SU("Brain development" OR "Neurological disorders" OR Neurotoxicity OR Neurosciences OR Neurons OR Learning OR Memory OR "Cognition & reasoning" OR Neurotransmitters OR Behavior OR "Psychosocial development" OR "Mood disorders" OR Anxiety)) AND (TI,AB,DISKW(Prenatal* OR Antenatal* OR Pregnan* OR Preterm OR F$etal* OR F$etus* OR Gestat*) OR SU("Prenatal care" OR "Prenatal development" OR Fetuses))) AND la.exact("ENG")

XSL•FO
RenderX

Database	Search string
Embase	(exp Corticosteroid/ or exp *Betamethasone/ or exp *Dexamethasone/ or exp *Hydrocortisone/ or exp *Corticosterone/ or (Corticosteroid* or cortisol* or Betamethasone or Dexamethasone or Glucocorticoid* or Corticosterone* or Hydrocortisone).ti,ab,kf,tn.) and (Brain*.ti,ab,kf. or exp *Brain/ or "Brain development".ti,ab,kf. or Neurodevelopment*.ti,ab,kf. or Neuron*.ti,ab,kf. or exp *Nerve cell/ or exp *Mental disease/ or Intelligence.ti,ab,kf. or exp *Learning/ or exp *Memory/ or Cogniti*.ti,ab,kf. or exp *Cognition/ or Behavio$r*.ti,ab,kf. or Psycholog*.ti,ab,kf. or Anxiety.ti,ab,kf. or Depress*.ti,ab,kf. or Stress*.ti,ab,kf. or "hypothalamo-pituitary-adrenal axis".ti,ab,kf. or "HPA axis".ti,ab,kf. or exp *Adrenal gland/ or exp *Corticotropin/ or Cortisol.ti,ab,kf. or Corticosterone*.ti,ab,kf. or exp *Corticosterone/ or Hippocamp*.ti,ab,kf. or exp *Hippocampus/ or "Corpus Callosum".ti,ab,kf. or exp *Corpus Callosum/ or Myelin*.ti,ab,kf. or exp *Myelin/ or exp *Glia cell/ or Oligodendrocyte*.ti,ab,kf. or "White matter".ti,ab,kf. or "Gray matter".ti,ab,kf. or Cerebell*.ti,ab,kf. or Neuroanatom*.ti,ab,kf. or Amygdala*.ti,ab,kf. or Astrocyte*.ti,ab,kf.) and (Antenatal*.ti,ab,kf. or exp *Pregnancy complication/ or Prenatal*.ti,ab,kf. or exp *Prenatal growth/ or Preterm.ti,ab,kf. or F$etus*.ti,ab,kf. or exp *Fetus/ or exp *Fetus growth/ or F$etal*.ti,ab,kf. or exp *Prenatal development/ or Gestat*.ti,ab,kf. or Pregnan*.ti,ab,kf. or exp *Pregnancy/)
Web of Science	(Corticosteroid* OR cortisol* OR Betamethasone OR Dexamethasone OR Glucocorticoid* OR Corticosterone* OR Hydrocortisone) AND (Brain* OR “Brain development” OR Neurodevelopment* OR Neuron* OR “Mental Disorders” OR Intelligence OR Learning OR Memory OR Cogniti* OR Behavio$r* OR Psycholog* OR Anxiety OR Depress* OR Stress* OR “hypothalamo-pituitary-adrenal axis” OR “HPA axis” OR Adrenal* OR “Adrenocorticotropic hormone” OR “Corticotropin-releasing hormone” OR Cortisol OR Corticosterone* OR Hippocamp* OR “Corpus Callosum” OR Myelin* OR Neuroglia* OR Oligodendrocyte* OR “White matter” OR “Gr?y matter” OR Cerebr* OR Cerebell* OR Neuroanatom* OR Amygdala* OR Astrocyte*) AND (Antenatal* OR Prenatal* OR Preterm OR F$etus* OR F$etal* OR Gestat* OR Pregnan*)
Scopus	TITLE-ABS-KEY ((Corticosteroid* OR cortisol* OR Betamethasone OR Dexamethasone OR Glucocorticoid* OR Corticosterone* OR Hydrocortisone) AND (Brain* OR “Brain development” OR Neurodevelopment* OR Neuron* OR “Mental Disorders” OR Intelligence OR Learning OR Memory OR Cogniti* OR Behavior* OR Psycholog* OR Anxiety OR Depress* OR Stress* OR “hypothalamo-pituitary-adrenal axis” OR “HPA axis” OR Adrenal* OR “Adrenocorticotropic hormone” OR “Corticotropin-releasing hormone” OR Cortisol OR Corticosterone* OR Hippocamp* OR “Corpus Callosum” OR Myelin* OR Neuroglia* OR Oligodendrocyte* OR “White matter” OR “Gr?y matter” OR Cerebr* OR Cerebell* OR Neuroanatom* OR Amygdala* OR Astrocyte*) AND (Antenatal* OR Prenatal* OR Preterm OR Fetus* OR Fetal* OR Gestat* OR Pregnan*))

Study Selection

Database search results will be imported to Covidence (Covidence systematic review software, Veritas Health Innovation), where reviewers can collaboratively screen sources for inclusion. Two assessors (any two of the review authors) will independently review sources using eligibility criteria (Table 2) at title, abstract, and full-text levels. Assessors will be blinded during screening. Uncertainties or conflicting decisions at each stage will be resolved through discussion and consensus, with at least one other member of the review team.

XSL•FO
RenderX

Table . Eligibility criteria for abstract screening.

	Inclusion criteria	Exclusion criteria
Population	• Nonhuman mammals • Healthy pregnancies and offspring	• Humans • Nonmammalian species • Genetically modified strains • Unhealthy pregnancies and offspring or models exposed to additional experimental manipulations
Exposure of interest	• In utero exposure to exogenous corticosteroids during the second half of gestation • Specific corticosteroids: dexamethasone, cortisol, corticosterone, or hydroxycortisol with no restrictions on the corticosteroid formulation, dose, or route of administration	• Corticosteroid administration not continuing after 0.5 gestation • Administration of postnatal steroids • Manipulations of endogenous steroids (eg, maternal restraint stress)
Comparator	• Concurrently studied placebo or untreated comparison group assessed at the same fetal or postnatal age	• Studies lacking a control group evaluated at the same fetal or postnatal age
Outcomes	• Functional brain outcomes including measures of cognition, intelligence, learning, memory, attention, social and other behavior, psychological disorders, and executive function • Structural determinants of brain function including gross measures of brain size and shape such as cortical folding, regional volumes, and brain pathology; measures of brain connectivity; measures of numbers, density, proportions, proliferation, and apoptosis for cell types within the brain such as neurons, glia, and astrocytes; measures of cellular morphology such as pruning, dendrites, axons, branching, volumes, and myelination; synaptic neurotransmitter abundance and receptors; measures of blood-brain barrier function, oxidative stress, and inflammation • Molecular determinants of brain function including gene and protein expression in pathways related to glucocorticoid signaling, neurodevelopment, myelination, apoptosis/proliferation, neurotransmitters, and oxidative stress, and epigenetic measures relating to determinants of brain function • Basal and stress-induced measures of hypothalamic-pituitary-adrenal axis function including circulating adrenocorticotropin, cortisol-releasing hormone, and corticosteroid concentrations • Structural and molecular determinants of hypothalamic-pituitary-adrenal axis function including adrenal size, glucocorticoid receptor, and steroidogenic enzyme expression	• No relevant outcomes
Types of studies	• Primary (original) studies • Peer-reviewed literature (published or in press, including theses) • Available in English	• Conference abstracts, reviews, editorials, trial protocols • Literature that has not undergone peer review (eg, preprints) • Full text not available in English

Data Extraction

Data extraction will be performed by two independent reviewers using author-designed fields in Covidence that will reflect the standardized JBI data extraction [36]. Uncertainties or conflicting decisions at each stage will be resolved through discussion and consensus with at least one other member of the review team. Study authors will be contacted to request any

XSL•FO
RenderX

missing or additional data. The extracted data will include citation information, species, strain, and normal term gestation of experimental animals used and information regarding study design including sample size (including number of mothers and number of offspring), method of allocation to treatments, timing of delivery, and assessment blinding. Extracted data relating to the ACS exposure will include the type of corticosteroid used, the dose including the formulation and route of administration, and the timing and duration of administration expressed in days of gestation and as a fraction of term. Extracted data related to outcomes will include sample size, sex, and postnatal age at assessment, details of the functional outcomes or brain regions analyzed, and information regarding the methodology used to measure outcomes. Where possible, the means, standard deviations, and percentage differences will be extracted for each treatment group and outcome. Where outcomes are only presented in figures, the numeric data (means and variance) will be extracted using WebPlotDigitizer v5 (Automeris LLC). Where further data or clarification are required for a study, the authors will be contacted via email. A minimum of three attempts at monthly intervals will be made to contact authors. Following extraction, we will cross-reference author names, experimental design, and study dates to identify multiple publications originating from the same experiment to prevent redundant data extractions arising from multiple publications. Where studies report overlapping data, such as multiple reports of the same outcome from an experimental cohort, only data from the largest study or more relevant exposures will be included. Where the same outcomes are reported in identical cohorts, the earliest publication will be retained. In studies where multiple treatment groups are compared against a single shared control group, the control group sample size will be divided equally across the comparison arms.

Assessment of Methodological Quality

Methodological quality will be assessed through the use of the Systematic Review Center for Laboratory Animal Experimentation Risk of Bias (SYRCLE RoB) tool [37]. Critical appraisal will be conducted by two independent reviewers, and judgment disagreements will be resolved through discussion and consensus or a third review if necessary. Appraisal results will be reported in table format.

Data Synthesis

A narrative synthesis will be presented following standard guidelines. A dose-response meta-analysis will be performed where at least 3 studies report the same outcome following in utero exposure to the same corticosteroid with comparable timing and data are made available by authors. The litter will be treated as the primary unit of analysis. For studies that report data from multiple offspring per litter without adjustment, the litter mean will be used as a single data point for each measured outcome. To ensure that these dose-response analyses are biologically meaningful across different mammalian species, doses will be standardized using allometric scaling (based on body surface area) to calculate human-equivalent dosing where possible. Where sufficient data are available, subgroup analyses will be performed within different species and biological sexes. Studies with missing data where authors do not respond to data requests will be included in the narrative synthesis but will be excluded from meta-analyses. Studies that report mean values for groups without measures of variance will be excluded from the meta-analysis but will be included in the narrative synthesis. For instance, where studies in the functional review report overlapping cohorts or outcomes with the underlying mechanisms and stress response reviews, we will cross-reference these findings to highlight consistent patterns of association across the different domains.

Assessing Certainty in the Findings

For any outcomes where meta-analysis is able to be performed, the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach will be used to grade the certainty of evidence [38]. Following this, a Summary of Findings will be created using GRADEpro GDT (McMaster University). Independent, blinded assessment of the certainty of evidence will be carried out by two reviewers, with any potential disagreements resolved through discussion or the involvement of a third reviewer. The Summary of Findings will present, where appropriate, mean differences, a ranking of the quality of the evidence based on the risk of bias, heterogeneity, precision, and risk of publication bias of the review results.

Dissemination Plan

The results of this systematic review will be disseminated through peer-reviewed publications.

Results

Initial searches were performed between August 12, 2025, and October 28, 2025, yielding 105,457 records. Search results from each database were downloaded to EndNote, duplicates and retractions were removed, and the remaining 105,308 sources were uploaded into Covidence (Figure 1). Records retrieved from more than one database were removed as duplicates by Covidence (n=48,698) or manually removed as duplicates during screening (n=117), resulting in 56,493 unique records (Figure 1). As of February 21, 2026, a total of 42,024 records have undergone title and abstract screening, with 40,795 records excluded (Figure 1). Publication is planned for 2027.

XSL•FO
RenderX

Figure 1. PRIMSA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of study selection as of February 21, 2026. CS: corticosteroid; GMO: genetically modified organisms; TBC: to be confirmed.

XSL•FO
RenderX

Discussion

Principal Findings

We hypothesize that this systematic review will reveal a dose-dependent relationship between ACS exposure and alterations in neurodevelopmental outcomes. Specifically, ACS exposure may be associated with increased anxiety-like behaviors and deficits in spatial learning and memory across various mammalian species. We expect to find evidence of reduced neuronal density and altered myelination patterns in brain regions relevant to behavior and altered HPA axis sensitivity. We hypothesize that the impacts of ACS will differ by biological sex, potentially suggestive of greater vulnerability in one sex regarding cognitive tasks or stress-reactivity. Finally, we anticipate that, while basic mechanisms of glucocorticoid action are conserved across species, the timing of exposure relative to brain maturation may contribute to differences in impacts in different species.

Comparison to Prior Work

Previous scoping and systematic reviews of the effects of ACS on neurodevelopment and stress responses have primarily focused on clinical outcomes in human cohorts [12,15,18]. The systematic or scoping reviews describing these outcomes in nonhuman species are of narrower scope, do not fully reflect the clinical context, and/or require updating [20,23,24]. Unlike clinical studies that may limited by confounding environmental variables, synthesis of preclinical evidence allows for a more controlled evaluation of dose-response relationships and tissue-specific molecular changes.

Limitations

This systematic review will have several limitations. Given that we are excluding publications not in English at screening, there is a possibility that we will miss relevant studies published in other languages, although these will be listed in the exclusion reasons for transparency [39]. It is not clear whether meta-analyses of outcomes will be possible given the heterogeneity of timings of steroid administration and outcome evaluation and the variety of species, antenatal steroid regimens, and outcome assessments evident in previous reviews [20]. Effect sizes may also be overestimated due to the potential nonpublication of null results in preclinical research [40].

Funding

This work was supported in part by a National Health and Medical Research Council (NHMRC) Ideas grant (KLG, JLM, SDM, MCL, grant number 2038332). TH was a recipient of the Faculty of Health and Medical Sciences Honours Scholarship from the University of Adelaide. MCL is funded by a fellowship from the Lloyd Cox O&G Research Fund of Adelaide University. JLM is funded by a NHMRC Investigator Grant (Leadership 2; GNT2041967).

Authors' Contributions

MFH, TH, JLM, and KLG conceived the study and performed initial searches of PubMed and PROSPERO, screening existing systematic reviews to identify knowledge gaps. MFH, TH, JLM, and KLG defined the scope of the study and eligibility criteria and developed the initial search string. TH adapted the search syntax for additional databases, performed systematic searches, and cross-checked results against key sources. MFH registered the protocol in PROSPERO. The initial manuscript draft was written by MFH, TH, JLM, and KLG. All authors contributed to revisions. All authors approved the final submission and agreed to be accountable for all aspects of work, verifying accuracy and integrity.

Conflicts of Interest

None declared.

References

1. Liggins GC. Premature delivery of foetal lambs infused with glucocorticoids. J Endocrinol 1969 Dec;45(4):515-523. [doi: 10.1677/joe.0.0450515] [Medline: 5366112]
2. Liggins GC, Howie RN. A controlled trial of antepartum glucocorticoid treatment for prevention of the respiratory distress syndrome in premature infants. Pediatrics 1972 Oct;50(4):515-525. [Medline: 4561295]
3. McGoldrick E, Stewart F, Parker R, Dalziel SR. Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. Cochrane Database Syst Rev 2020 Dec 25;12(12):CD004454. [doi: 10.1002/14651858.CD004454.pub4] [Medline: 33368142]
4. Williams MJ, Ramson JA, Brownfoot FC. Different corticosteroids and regimens for accelerating fetal lung maturation for babies at risk of preterm birth. Cochrane Database Syst Rev 2022 Aug 9;8(8):CD006764. [doi: 10.1002/14651858.CD006764.pub4] [Medline: 35943347]
5. Cao-Lei L, Suwansirikul S, Jutavijittum P, Mériaux SB, Turner JD, Muller CP. Glucocorticoid receptor gene expression and promoter CpG modifications throughout the human brain. J Psychiatr Res 2013 Nov;47(11):1597-1607. [doi: 10.1016/j.jpsychires.2013.07.022] [Medline: 23948638]
6. Lee CT, Lehrmann E, Hayashi T, et al. Gene expression profiling reveals distinct cocaine-responsive genes in human fetal CNS cell types. J Addict Med 2009 Dec;3(4):218-226. [doi: 10.1097/ADM.0b013e318199d863] [Medline: 20948987]

XSL•FO
RenderX

7. Noorlander CW, De Graan PNE, Middeldorp J, Van Beers J, Visser GHA. Ontogeny of hippocampal corticosteroid receptors: effects of antenatal glucocorticoids in human and mouse. J Comp Neurol 2006 Dec 20;499(6):924-932. [doi: 10.1002/cne.21162] [Medline: 17072842]
8. Asztalos EV, Murphy KE, Matthews SG. A growing dilemma: antenatal corticosteroids and long-term consequences. Am J Perinatol 2022 Apr;39(6):592-600. [doi: 10.1055/s-0040-1718573] [Medline: 33053595]
9. Jobe AH, Goldenberg RL. Antenatal corticosteroids: an assessment of anticipated benefits and potential risks. Am J Obstet Gynecol 2018 Jul;219(1):62-74. [doi: 10.1016/j.ajog.2018.04.007] [Medline: 29630886]
10. Räikkönen K, Gissler M, Kajantie E. Associations between maternal antenatal corticosteroid treatment and mental and behavioral disorders in children. JAMA 2020 May 19;323(19):1924-1933. [doi: 10.1001/jama.2020.3937] [Medline: 32427304]
11. Sultana S, Bruinsma F, Fisher Z, Homer CSE, Vogel JP. Long-term outcomes of antenatal corticosteroids for preterm birth: an overview of systematic reviews. PLOS Glob Public Health 2025;5(5):e0004575. [doi: 10.1371/journal.pgph.0004575] [Medline: 40333673]
12. Ninan K, Liyanage SK, Murphy KE, Asztalos EV, McDonald SD. Evaluation of long-term outcomes associated with preterm exposure to antenatal corticosteroids: a systematic review and meta-analysis. JAMA Pediatr 2022 Jun 1;176(6):e220483. [doi: 10.1001/jamapediatrics.2022.0483] [Medline: 35404395]
13. Ninan K, Gojic A, Wang Y, et al. The proportions of term or late preterm births after exposure to early antenatal corticosteroids, and outcomes: systematic review and meta-analysis of 1.6 million infants. BMJ 2023 Aug 2;382:e076035. [doi: 10.1136/bmj-2023-076035] [Medline: 37532269]
14. Ciapponi A, Klein K, Colaci D, et al. Dexamethasone versus betamethasone for preterm birth: a systematic review and network meta-analysis. Am J Obstet Gynecol MFM 2021 May;3(3):100312. [doi: 10.1016/j.ajogmf.2021.100312] [Medline: 33482400]
15. Sarid EB, Stoopler ML, Morency AM, Garfinkle J. Neurological implications of antenatal corticosteroids on late preterm and term infants: a scoping review. Pediatr Res 2022 Nov;92(5):1225-1239. [doi: 10.1038/s41390-022-02135-3] [Medline: 35681094]
16. Waffarn F, Davis EP. Effects of antenatal corticosteroids on the hypothalamic-pituitary-adrenocortical axis of the fetus and newborn: experimental findings and clinical considerations. Am J Obstet Gynecol 2012 Dec;207(6):446-454. [doi: 10.1016/j.ajog.2012.06.012] [Medline: 22840973]
17. Carson R, Monaghan-Nichols AP, DeFranco DB, Rudine AC. Effects of antenatal glucocorticoids on the developing brain. Steroids 2016 Oct;114(25-32):25-32. [doi: 10.1016/j.steroids.2016.05.012] [Medline: 27343976]
18. Tegethoff M, Pryce C, Meinlschmidt G. Effects of intrauterine exposure to synthetic glucocorticoids on fetal, newborn, and infant hypothalamic-pituitary-adrenal axis function in humans: a systematic review. Endocr Rev 2009 Dec;30(7):753-789. [doi: 10.1210/er.2008-0014] [Medline: 19837868]
19. Alexander N, Rosenlöcher F, Stalder T, et al. Impact of antenatal synthetic glucocorticoid exposure on endocrine stress reactivity in term-born children. J Clin Endocrinol Metab 2012 Oct;97(10):3538-3544. [doi: 10.1210/jc.2012-1970] [Medline: 22869608]
20. van der Merwe JL, Sacco A, Toelen J, Deprest J. Long-term neuropathological and/or neurobehavioral effects of antenatal corticosteroid therapy in animal models: a systematic review. Pediatr Res 2020 Jun;87(7):1157-1170. [doi: 10.1038/s41390-019-0712-1] [Medline: 31822018]
21. Rodriguez JS, Zürcher NR, Keenan KE, Bartlett TQ, Nathanielsz PW, Nijland MJ. Prenatal betamethasone exposure has sex specific effects in reversal learning and attention in juvenile baboons. Am J Obstet Gynecol 2011 Jun;204(6):545. [doi: 10.1016/j.ajog.2011.01.063] [Medline: 21411054]
22. Stojanovska V, Barton SK, Tolcos M, et al. The effect of antenatal betamethasone on white matter inflammation and injury in fetal sheep and ventilated preterm lambs. Dev Neurosci 2018;40(5-6):497-507. [doi: 10.1159/000496466] [Medline: 30840951]
23. Aghajafari F, Murphy K, Matthews S, Ohlsson A, Amankwah K, Hannah M. Repeated doses of antenatal corticosteroids in animals: a systematic review. Am J Obstet Gynecol 2002 Apr;186(4):843-849. [doi: 10.1067/mob.2002.121624] [Medline: 11967518]
24. van de Meent M, Kleuskens DG, Derks JB, et al. Effects of antenatal corticosteroid therapy in animal models of fetal growth restriction: a systematic review and meta-analysis. BMC Pregnancy Childbirth 2025 Mar 13;25(1):281. [doi: 10.1186/s12884-025-07359-9] [Medline: 40082797]
25. Antonow-Schlorke I, Helgert A, Gey C, et al. Adverse effects of antenatal glucocorticoids on cerebral myelination in sheep. Obstet Gynecol 2009 Jan;113(1):142-151. [doi: 10.1097/AOG.0b013e3181924d3b] [Medline: 19104370]
26. Brabham T, Phelka A, Zimmer C, Nash A, López JF, Vázquez DM. Effects of prenatal dexamethasone on spatial learning and response to stress is influenced by maternal factors. Am J Physiol Regul Integr Comp Physiol 2000 Nov;279(5):R1899-R1909. [doi: 10.1152/ajpregu.2000.279.5.R1899] [Medline: 11049876]
27. Hauser J, Knapman A, Zürcher NR, et al. Effects of prenatal dexamethasone treatment on physical growth, pituitary-adrenal hormones, and performance of motor, motivational, and cognitive tasks in juvenile and adolescent common marmoset monkeys. Endocrinology 2008 Dec;149(12):6343-6355. [doi: 10.1210/en.2008-0615] [Medline: 18755792]

XSL•FO
RenderX

28. Huang WL, Harper CG, Evans SF, Newnham JP, Dunlop SA. Repeated prenatal corticosteroid administration delays astrocyte and capillary tight junction maturation in fetal sheep. Int J Dev Neurosci 2001 Aug;19(5):487-493. [doi: 10.1016/s0736-5748(01)00035-1] [Medline: 11470378]
29. Noorlander CW, Tijsseling D, Hessel EVS, et al. Antenatal glucocorticoid treatment affects hippocampal development in mice. PLoS One 2014;9(1):e85671. [doi: 10.1371/journal.pone.0085671] [Medline: 24465645]
30. Oliveira M, Rodrigues AJ, Leão P, Cardona D, Pêgo JM, Sousa N. The bed nucleus of stria terminalis and the amygdala as targets of antenatal glucocorticoids: implications for fear and anxiety responses. Psychopharmacology (Berl) 2012 Apr;220(3):443-453. [doi: 10.1007/s00213-011-2494-y] [Medline: 21935638]
31. Rim C, Park HS, You MJ, et al. Microglia involvement in sex-dependent behaviors and schizophrenia occurrence in offspring with maternal dexamethasone exposure. Schizophrenia (Heidelb) 2022 Sep 8;8(1):71. [doi: 10.1038/s41537-022-00280-6] [Medline: 36075925]
32. Shields A, Thomson M, Winter V, Coalson J, Rees S. Repeated courses of antenatal corticosteroids have adverse effects on aspects of brain development in naturally delivered baboon infants. Pediatr Res 2012 Jun;71(6):661-667. [doi: 10.1038/pr.2012.18] [Medline: 22436975]
33. Uno H, Lohmiller L, Thieme C, et al. Brain damage induced by prenatal exposure to dexamethasone in fetal rhesus macaques. I. Hippocampus. Brain Res Dev Brain Res 1990 May 1;53(2):157-167. [doi: 10.1016/0165-3806(90)90002-g] [Medline: 2357788]
34. van der Merwe J, van der Veeken L, Inversetti A, et al. Neurocognitive sequelae of antenatal corticosteroids in a late preterm rabbit model. Am J Obstet Gynecol 2022 Jun;226(6):850. [doi: 10.1016/j.ajog.2021.11.1370] [Medline: 34875248]
35. Zuloaga DG, Carbone DL, Hiroi R, Chong DL, Handa RJ. Dexamethasone induces apoptosis in the developing rat amygdala in an age-, region-, and sex-specific manner. Neuroscience 2011 Dec 29;199:535-547. [doi: 10.1016/j.neuroscience.2011.09.052] [Medline: 22008524]
36. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, et al. Chapter 7: systematic reviews of etiology and risk. In: Aromataris E, Lockwood C, Poritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis 2024. [doi: 10.46658/JBIMES-24-06]
37. Hooijmans CR, Rovers MM, de Vries RBM, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC Med Res Methodol 2014 Mar 26;14(43):43. [doi: 10.1186/1471-2288-14-43] [Medline: 24667063]
38. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008 Apr 26;336(7650):924-926. [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]
39. Pieper D, Puljak L. Language restrictions in systematic reviews should not be imposed in the search strategy but in the eligibility criteria if necessary. J Clin Epidemiol 2021 Apr;132:146-147. [doi: 10.1016/j.jclinepi.2020.12.027] [Medline: 33383129]
40. van der Naald M, Wenker S, Doevendans PA, Wever KE, Chamuleau SAJ. Publication rate in preclinical research: a plea for preregistration. BMJ Open Sci 2020;4(1):e100051. [doi: 10.1136/bmjos-2019-100051] [Medline: 35047690]

Abbreviations

ACS: antenatal corticosteroids
GR: glucocorticoid receptor
GRADE: Grading of Recommendations, Assessment, Development and Evaluation
HPA: hypothalamic-pituitary-adrenal
SYRCLE RoB: Systematic Review Centre for Laboratory Animal Experimentation Risk of Bias

Edited by J Sarvestan; submitted 21.Sep.2025; peer-reviewed by E Reyna-Villasmil; revised version received 26.Feb.2026; accepted 26.Feb.2026; published 27.Mar.2026.

Please cite as:
Herselman MF, Hislop T, McBride SD, Lock MC, Morrison JL, Gatford KL
Impact of Antenatal Corticosteroids on Brain Function and Underlying Mechanisms in Preclinical Studies: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e84522
URL: https://www.researchprotocols.org/2026/1/e84522
doi:10.2196/84522

© Mauritz F Herselman, Tara Hislop, Sebastian D McBride, Mitchell C Lock, Janna L Morrison, Kathryn L Gatford. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed

XSL•FO
RenderX

under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Determining the Impact of Nonionizing Electromagnetic Waves on Human Pregnancy and Teratogenicity: Protocol for Systematic Review and Meta-Analysis

Desy Armalina[1,2*], MD; Neni Susilaningsih[2*], Dr MSi; Heri Sutanto[3*], Prof Dr; Sunarno Sunarno[4*], Dr MSi SSi

[1]Doctoral Study Program of Medical and Health Science, Universitas Diponegoro, Jl. Prof. Soedarto, Tembalang, Kec. Tembalang, Kota Semarang, Jawa Tengah, Semarang, Central Java, Indonesia

[2]Anatomy-Histology Department, Faculty of Medicine, Universitas Diponegoro, Semarang, Indonesia

[3]Department of Physics, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

[4]Department of Biology, Faculty of Science and Mathematics, Universitas Diponegoro, Semarang, Indonesia

[*]all authors contributed equally

Corresponding Author:
Desy Armalina, MD
Doctoral Study Program of Medical and Health Science, Universitas Diponegoro, Jl. Prof. Soedarto, Tembalang, Kec. Tembalang, Kota Semarang, Jawa Tengah, Semarang, Central Java, Indonesia

Abstract

Background: The widespread use of mobile devices has markedly increased global exposure to nonionizing electromagnetic waves (EMWs). Emerging evidence indicates potential biological effects of EMW exposure in susceptible populations, particularly pregnant women; however, findings remain inconsistent.

Objective: This protocol delineates a systematic review aimed at synthesizing and critically evaluating the teratogenic and pregnancy-related effects of nonionizing EMW exposure in pregnant women.

Methods: This protocol adheres to the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) 2020 guidelines and has been registered with the International Prospective Register of Systematic Reviews (PROSPERO; CRD42023475665). A comprehensive literature search will be conducted in PubMed/MEDLINE, Scopus, Web of Science, Embase, ScienceDirect, SpringerLink, Wiley Online Library, and Google Scholar, with supplementary searches of the World Health Organization International Clinical Trials Registry Platform and ClinicalTrials.gov. Eligible studies will include pregnant women exposed to nonionizing EMWs from mobile phones and related wireless devices. The primary outcomes will be pregnancy complications and fetal anomalies, with secondary outcomes assessed as previously reported. Study selection, data extraction, and risk of bias assessment will be performed independently by 2 reviewers. Where appropriate, a random-effects meta-analysis will be conducted.

Results: Funding for this study was secured in March 2026. The literature search and study screening are planned for April to July 2026, with data extraction, risk of bias assessment, and synthesis expected to be completed by September 2026. The final results are anticipated to be submitted for publication in late 2026.

Conclusions: This systematic review is expected to provide consolidated evidence on the potential teratogenic and pregnancy-related effects of nonionizing EMW exposure, thereby supporting future research and evidence-based recommendations for public health.

Trial Registration: PROSPERO CRD42023475665; https://www.crd.york.ac.uk/PROSPERO/view/CRD42023475665

International Registered Report Identifier (IRRID): PRR1-10.2196/86479

(JMIR Res Protoc 2026;15:e86479) doi:10.2196/86479

KEYWORDS

electromagnetic waves; nonionizing radiation; pregnancy; teratogenicity; systematic review; mobile phone

Introduction

Global digital transformation has fundamentally altered human interactions with electromagnetic energy. Modern lifestyles are now inseparable from mobile phones, Wi-Fi routers, base stations, and wireless devices, all of which emit nonionizing electromagnetic waves (EMWs) in the radiofrequency (RF) and microwave spectra. The World Health Organization has noted that electromagnetic fields (EMFs) are among the most common and fastest-growing environmental exposures worldwide, with

XSL•FO
RenderX

population-level exposure continuing to increase as wireless technologies expand. Although nonionizing radiation lacks the quantum energy necessary to break chemical bonds or ionize atoms, experimental and dosimetric studies have suggested that exposure may induce biological effects through thermal mechanisms related to tissue heating and nonthermal interactions observed under specific exposure conditions, including oxidative stress responses and alterations in cellular signaling pathways [1,2].

Nonionizing EMWs are typically defined as waves with frequencies ranging from 300 Hz to 300 GHz. At these frequencies, the absorbed energy can result in localized tissue heating, which represents the most well-characterized biological effect of EMW exposure. Nonthermal biological interactions, including perturbations in redox balance, cellular signaling, and gene expression, have also been reported in experimental models; however, these findings remain inconsistent, and their relevance to human pregnancy outcomes has not been conclusively established [1,2].

Mobile phones, as ubiquitous sources of EMW exposure, have become integral to our daily life. Global smartphone use has increased from a few hundred million users in 2008 to more than 6.5 billion active users by 2023. In parallel, the average daily usage time has increased substantially, particularly during and after the COVID-19 pandemic [3].

Recent epidemiological studies have explored the potential associations between maternal mobile phone use and adverse pregnancy outcomes, including miscarriage, preterm birth, and abnormal birth weight; however, the results remain heterogeneous across populations and study methodologies [4].

Consequently, the intensity, proximity, and duration of EMW exposure in the general population, particularly among pregnant women, has increased over time. Pregnancy represents a critical window of biological vulnerability due to the complex endocrine, immunological, and vascular adaptations required to sustain fetal growth. The developing embryo and fetus may be especially sensitive to environmental exposures because of rapid cellular proliferation, immature antioxidant defenses, and limited detoxification capacity. As a result, EMW exposure during pregnancy has been investigated for its potential effects on placental function, uteroplacental blood flow, embryonic development, and fetal morphogenesis [5-7].

Importantly, the presence of a potential biological mechanism does not imply a clinically relevant risk. The current international consensus indicates that adverse health effects from nonionizing EMW exposure are strongly dependent on exposure intensity, duration, and proximity to the source. At high power densities, such as in occupational or accidental settings near strong RF emitters, thermal tissue injury may occur; however, under typical environmental and consumer-use conditions, exposure levels are substantially lower and generally remain below established safety limits [8,9]. Consequently, while EMWs represent a potential hazard under specific conditions, the actual health risk during normal daily use is determined by the dose and duration of exposure, which are typically far below the thresholds known to cause harm. This distinction between hazard and risk is central to the interpretation of existing evidence.

Epidemiological investigations have reported mixed findings regarding the association between EMF exposure and adverse pregnancy outcomes. A meta-analysis by Ghazanfarpour et al [3] reported an increased risk of miscarriage associated with higher levels of extremely low-frequency electromagnetic exposure. Similarly, Li et al [10] observed an association between magnetic field exposure and spontaneous abortion in a prospective cohort study. In addition, computational dosimetric modeling by Takei et al [11] demonstrated that mobile phone radiation may increase localized placental and fetal temperatures under specific exposure scenarios, suggesting a potential thermal pathway warranting further investigation.

In contrast, other large-scale epidemiological studies have not identified significant associations between EMW exposure and adverse pregnancy or birth outcomes, highlighting substantial heterogeneity in exposure assessment, study design, and confounder control [5,12]. Many earlier studies were also limited by small sample sizes, indirect exposure metrics, and incomplete adjustment for potential confounders, including maternal stress and environmental coexposure.

From a mechanistic perspective, experimental studies have reported that EMW exposure may induce oxidative stress responses, alter antioxidant enzyme activity, and influence calcium signaling and apoptotic pathways in animal and in vitro models [13-16]. However, the extent to which these experimental findings translate to clinically meaningful effects during human pregnancy remains unclear.

Given the complex and often contradictory findings in the existing literature, a rigorous systematic review is necessary. This systematic review protocol aims to systematically synthesize and critically evaluate the available evidence on the teratogenic and pregnancy-related effects of nonionizing EMW exposure, particularly from mobile phone use, in pregnant women.

Methods

Protocol and Registration

This review adheres to the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) 2020 standards [17] and is registered with International Prospective Register of Systematic Reviews (PROSPERO; ID: CRD42023475665), which ensures transparency and reproducibility. The completed PRISMA 2020 checklist is provided in Checklist 1. All protocol amendments will be documented in PROSPERO.

Eligibility Criteria

Studies will be selected according to the Population Intervention (or Exposure, for observational studies) Comparator Outcomes Study design (PICOS) framework as follows.

Population

Included were pregnant women (any maternal age) and their offspring.

XSL•FO
RenderX

Outcome time window: Outcomes will be eligible if measured during pregnancy (in utero) and/or at birth (perinatal/neonatal periods). If studies report outcomes beyond birth, outcomes will be included up to 28 days postpartum (neonatal period) and analyzed separately as postnatal outcomes. Outcomes measured beyond 28 days will not be included in the primary synthesis but may be described narratively if directly relevant.

Intervention: Exposure (EMW/EMF)

The included sources were nonionizing EMW exposure related to mobile phone use and related consumer wireless devices (eg, smartphones and Wi-Fi–enabled devices) and/or ambient RF sources, where exposure was quantified or reasonably categorized.

Exposure metrics: Studies will be eligible if exposure is reported using at least one of the following: specific absorption rate (SAR), power density (W/m^2 or mW/cm^2), field strength (V/m), magnetic flux density (μT/mG), or a validated exposure classification method (eg, job-exposure matrix or personal dosimetry).

Dose range limits: We will not impose a single universal lower or upper cutoff for power density a priori because consumer and environmental studies use heterogeneous exposure metrics and thresholds. Instead, we will extract the exposure intensity and categorize it into low or typical consumer use versus higher occupational or accidental exposure levels as defined by each study, and we will interpret the findings in relation to established safety guidance (eg, ICNIRP/WHO) in the Discussion. Where data permit, dose–response analyses will be conducted according to the exposure strata.

Comparators: Lower exposure group or nonexposed group, depending on study design.

Outcomes

The primary outcomes of this review will be preterm birth and major congenital anomalies, assessed at birth. These outcomes were selected a priori based on their clinical relevance and biological plausibility in relation to nonionizing EMW exposure. Secondary outcomes will include miscarriage, stillbirth, low birth weight, small for gestational age, placental outcomes (as reported), and neonatal outcomes up to 28 days postpartum. Outcomes reported beyond the neonatal period will be described narratively when relevant but will not be included in the primary quantitative synthesis.

Timing

Timing will be a consideration.

- Minimum exposure or observation duration: Studies must report exposure during pregnancy with an identifiable exposure window (eg, trimester-specific or pregnancy-level exposure).
- Timing categories: We will extract and categorize exposure timing by trimester (if available) and duration of exposure (eg, daily use duration or cumulative exposure).
- Follow-up window: In utero, at birth, and up to 28 days after birth.

Study Design

Included: Human primary research only—randomized trials (if any), cohort studies, case-control studies, cross-sectional studies, and case-crossover designs that reported pregnancy outcomes.

Excluded from quantitative synthesis (meta-analysis): Systematic reviews, narrative reviews, editorials, letters, case reports, and all animal or in vivo studies.

Use of systematic reviews: Systematic reviews will be used only for citation tracking to identify additional eligible primary human studies but will not be included as studies in the meta-analysis.

Comparator

The comparator groups will include pregnant women with no reported exposure to nonionizing EMWs, background environmental exposure only, or the lowest exposure category defined within each individual study. The comparator definitions will be extracted as reported and assessed for comparability across studies.

Exclusion Criteria

Insufficient Exposure Information

Studies will be excluded if EMW exposure is not defined in sufficient detail to allow meaningful interpretation or comparison. Specifically, studies will be considered to have insufficient exposure data if none of the following are reported or inferred:

- Exposure frequency band (eg, extremely low-frequency, RF, microwave, or device-specific frequency range)
- Timing or duration of exposure during pregnancy (eg, trimester-specific exposure, daily use duration, cumulative exposure, or occupational exposure period)

Studies that report qualitative or proxy exposure measures (eg, categorized phone use and job-exposure matrices) will not be excluded solely on this basis, provided that the exposure classification is clearly described and consistently applied.

Language

No language restrictions were applied during the screening stage. Studies published in any language will be considered. Non-English studies will be screened using English abstracts where available, and, when necessary, machine-assisted translation (eg, Google Translate or equivalent tools) will be used. For studies deemed potentially eligible after full-text screening, translations will be used to extract relevant data, with attention to methodological clarity and outcome definition.

Search Strategy and Information Sources

The search strategy will be developed and refined in consultation with an experienced university health sciences librarian to ensure methodological rigor and comprehensive coverage of the literature. The final search strategy will combine controlled vocabulary (eg, Medical Subject Headings terms) and free-text keywords related to EMWs, pregnancy, and reproductive outcomes and will be adapted for each database.

XSL•FO
RenderX

The electronic databases to be searched include PubMed/MEDLINE, Scopus, Web of Science, Embase, ScienceDirect, SpringerLink, Wiley Online Library, and Google Scholar.

In addition, clinical trial registries will be searched, including the WHO International Clinical Trials Registry Platform. Where relevant, other registries, such as ClinicalTrials.gov, will be screened to identify completed or ongoing studies with unpublished results.

In addition to bibliographic databases, gray literature sources will be searched to minimize publication bias. These include ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform. Embase and the Cochrane Library are included to ensure comprehensive coverage of biomedical, epidemiological, and clinical trial literature not fully indexed in PubMed.

Study Selection and Data Extraction

Two reviewers will independently assess titles and abstracts for relevance. Full texts will then be reviewed for eligibility. Discrepancies will be resolved through discussion or by a third reviewer.

The selection process will be documented using a PRISMA 2020 flowchart (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) 2020 flow diagram for new systematic reviews, which included searches of databases and registers only. The asterisk (*) refers to records identified through database searching (PubMed/MEDLINE, Scopus, Web of Science, Embase, ScienceDirect, SpringerLink, Wiley Online Library, and Google Scholar). The double asterisk (**) refers to records excluded during title/abstract screening.

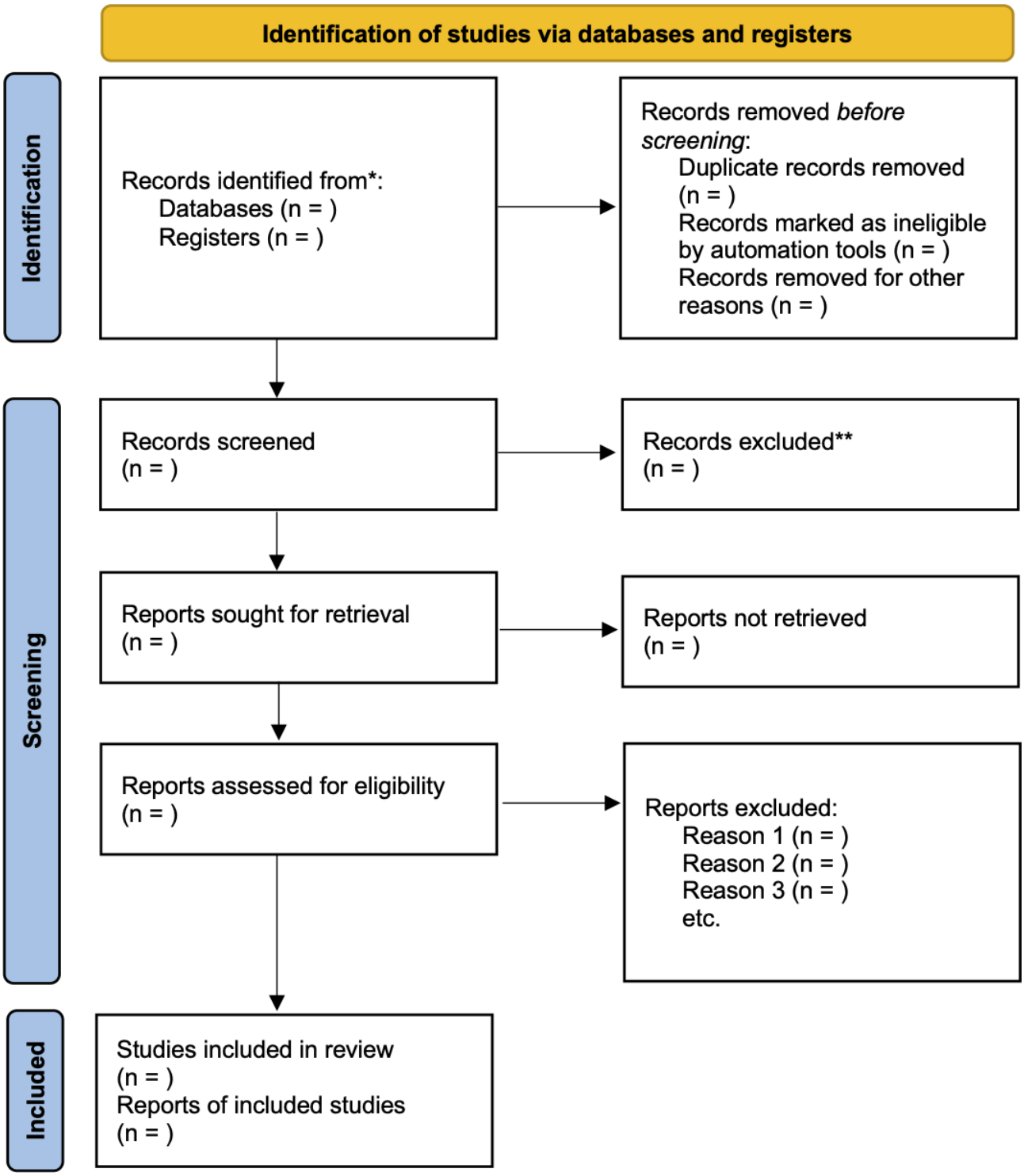

Data will be extracted into Microsoft Excel, capturing the following:

1. Study characteristics (author, year, country)
2. Study design and sample size
3. Exposure parameters (frequency, intensity, duration)
4. Participant details (age, gestational week)
5. Outcomes and key findings
6. Reported statistical measures (odds ratio, risk ratio, CI)

When information is incomplete, the authors will be contacted for clarification.

Risk of Bias and Quality Assessment

Only primary human studies will be included in the risk of bias assessment and quantitative synthesis. The risk of bias in nonrandomized observational studies will be assessed using the Risk Of Bias In Nonrandomized Studies of Interventions tool, which evaluates bias across the domains of confounding,

XSL•FO
RenderX

participant selection, exposure classification, deviations from intended exposure, missing data, outcome measurement, and selective reporting. Where randomized controlled trials are identified, the Cochrane Risk of Bias 2 tool will be applied. Each study will be categorized as having low, moderate, serious, or critical risk of bias according to the tool's guidance. Studies judged to be at critical risk of bias will be excluded from quantitative synthesis but described narratively. Reporting quality will be evaluated descriptively using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist, without influencing risk of bias judgments.

Data Synthesis and Assessment of Heterogeneity

Statistical heterogeneity will be assessed using the I^2 statistic and chi-square test. In accordance with the Cochrane Handbook, I^2 values will be interpreted cautiously and in context, recognizing that fixed thresholds can be misleading and that the importance of heterogeneity depends on multiple factors, including the magnitude and direction of effects and the strength of evidence for heterogeneity (eg, CIs around I^2 and the number of included studies).

Meta-analyses will be conducted using random-effects models when appropriate. Rather than relying on rigid I^2 cutoffs to define "substantial" heterogeneity, heterogeneity will be primarily addressed through clinical and methodological assessments, including differences in study populations, exposure metrics, outcome definitions, and confounder adjustment.

To better reflect the between-study variability and the expected range of true effects across different settings, prediction intervals will be calculated and reported for random-effects meta-analyses, in line with contemporary methodological guidance [18].

Statistical Analysis

When quantitative synthesis is feasible, a random-effects meta-analysis will be performed using the DerSimonian–Laird or restricted maximum likelihood approach, depending on the data structure. Effect estimates will be reported with 95% CIs and, where applicable, with prediction intervals. Sensitivity analyses will be conducted to assess the robustness of the findings to study quality and exposure definitions.

Observational Data and Minimum Requirements for Confounding Control

Given that the available evidence is predominantly observational, included studies must demonstrate a priori adjustment for key confounders to be eligible for quantitative synthesis. At a minimum, studies must adjust for the following:

- maternal age
- gestational age
- socioeconomic status or education
- smoking and/or alcohol use
- at least one indicator of overall health or pregnancy risk (eg, parity or comorbidities)

Studies failing to account for these minimum confounders will be excluded from meta-analysis and described narratively only. This minimum adjustment set will be applied a priori as a criterion for inclusion in quantitative synthesis. Residual confounding will be explicitly considered in the interpretation of findings, and causal language will be avoided in accordance with the limitations of observational evidence.

Certainty of Evidence

The overall quality and certainty of evidence for each outcome will be graded using the GRADE (Grading of Recommendations Assessment, Development, and Evaluation) system as high, moderate, low, or very low [19].

Ethical Considerations

This study is a systematic review protocol based exclusively on the published literature and does not involve the collection of primary data or interaction with human participants. Therefore, ethical approval from an institutional review board or ethics committee is not required, in accordance with institutional and international guidelines for research involving publicly available data.

Results

This protocol was registered with PROSPERO in 2023 and finalized in March 2026 following funding approval. The literature search and study selection are planned for April to July 2026, followed by data extraction, risk of bias assessment, and data synthesis by September 2026. Submission of the final systematic review manuscript is anticipated in late 2026.

Discussion

Expected Findings

This systematic review protocol provides a transparent methodological framework for evaluating the potential effects of nonionizing EMWs on human pregnancy and fetal development. Given the rapid expansion of mobile phone use and wireless technology, understanding the biological implications of prolonged exposure to EMWs has become an increasingly important public health concern [1,2]. Despite extensive research, the existing literature remains heterogeneous, with findings ranging from no observable adverse effects to possible reproductive or developmental impacts [13,17,20]. This protocol is designed to systematically assess and contextualize these inconsistencies through a rigorous methodological appraisal.

Comparison With Prior Work

Biological tissues can absorb EMWs, with thermal effects related to localized tissue heating representing the most well-characterized interaction [21]. In addition, nonthermal biological interactions have been described primarily in experimental and in vitro models, including oxidative stress responses, calcium channel modulation, and alterations in gene expression; however, these findings remain inconsistent, and their relevance to human pregnancy outcomes is not firmly established [1,2].

Previous human studies have reported mixed associations between electromagnetic exposure and adverse pregnancy

XSL•FO
RenderX

outcomes, including miscarriage and fetal growth restriction [3,12]. A systematic review by Wdowiak et al [14] reported potential links between chronic EMF exposure and lower birth weight, placental dysfunction, and developmental delay. Similarly, computational modeling by Takei et al [11] suggested that smartphone radiation may increase localized placental temperature, potentially affecting placental perfusion. Advanced modeling studies have further explored how RF radiation from commonly used mobile phones may influence electromagnetic exposure profiles in pregnant women, highlighting potential heat distribution effects that could have biological relevance; however, these findings require empirical validation in human populations [22].

Conversely, other large-scale epidemiological studies have not confirmed significant teratogenic or adverse pregnancy risks, underscoring the substantial heterogeneity in exposure assessment methods, study design, and confounder control [5,23]. Reviews of nonionizing radiation effects have also noted that while EMW exposure may induce cellular stress responses, such as oxidative stress or altered signaling pathways, the direct clinical relevance of these mechanisms to pregnancy outcomes remains insufficiently established, reinforcing the need for rigorous and balanced systematic synthesis [24].

Strengths and Limitations

These inconsistencies highlight the need for a comprehensive and methodologically rigorous systematic review. Many existing studies rely on self-reported device use or indirect exposure proxies, which may introduce recall bias and exposure misclassification [25]. Additionally, limited adjustment for confounding factors such as maternal stress, occupational radiation exposure, environmental pollutants, and thermal stress further complicates the interpretation of the available evidence [26].

The strengths of this protocol include adherence to PRISMA-P guidelines, a comprehensive search strategy, and the use of contemporary, domain-based risk of bias assessment tools appropriate for observational evidence, with explicit consideration of confounding and exposure misclassification. Nonetheless, heterogeneity in exposure metrics, outcome definitions, and study populations may limit the certainty of any conclusions that can be drawn; therefore, the findings will be interpreted with appropriate caution.

Implications and Future Directions

By synthesizing evidence from human observational studies, this systematic review aims to improve understanding of potential associations between nonionizing EMW exposure and pregnancy outcomes. Evidence from experimental and modeling studies will be discussed contextually in the Discussion to support biological plausibility, without being included in the quantitative synthesis. Experimental studies have suggested that RF exposure may influence oxidative balance, mitochondrial integrity, and calcium signaling in placental and fetal tissues under specific experimental conditions [20,27].

The anticipated contributions of this review are threefold: (1) to synthesize fragmented evidence regarding possible associations between EMW exposure and pregnancy outcomes while carefully accounting for study quality and bias; (2) to critically evaluate exposure assessment methodologies and identify research gaps, particularly regarding real-time dosimetry and gestational timing of exposure; and (3) to support evidence-informed discussion and future research directions rather than definitive risk conclusions, consistent with precautionary principles proposed by international organizations such as the WHO and ICNIRP.

Dissemination Plan

The findings of this systematic review will be disseminated through publications in peer-reviewed open-access journals and presentations at relevant scientific conferences. In addition, the results may be communicated to academic and public health stakeholders to facilitate transparent, evidence-based interpretation and inform future research priorities.

Conclusion

This protocol provides a structured and reproducible framework for systematically synthesizing evidence of nonionizing EMW exposure during pregnancy. By applying rigorous methodological standards and transparent quality assessments, the forthcoming review aims to clarify the current state of evidence while explicitly acknowledging its limitations. The conclusions will be framed cautiously, reflecting the strength and consistency of the available data and are intended to support responsible scientific interpretation and future research rather than draw definitive clinical or causal inferences.

Acknowledgments

The authors express sincere gratitude to the Faculty of Medicine, Diponegoro University, for the provision of research facilities and academic guidance, and to the Ministry of Education, Culture, Research, and Technology (Kemdikbudristek) of Indonesia for continued support in advancing scientific publication. The authors also thank the Systematic Review Methodology Unit at Diponegoro University for consultation and assistance with protocol validation.

Funding

This study was supported by the Endowment Fund for Education (LPDP), Indonesia.

Data Availability

This study is a systematic review protocol and does not generate or analyze new datasets. All data to be analyzed will be derived from published studies identified through a systematic literature search. Therefore, no original datasets are available for this study.

XSL•FO
RenderX

Authors' Contributions

Conceptualization: DA, NS, HS, Sunarno
Methodology: DA, NS, HS, Sunarno
Protocol design: DA (lead)
Search strategy design and PRESS peer review: NS (lead)
Reference management and deduplication: NS
Screening (titles/abstracts and full text): DA, NS
Data extraction: DA (lead), HS
Data curation: DA
Formal analysis: HS, Sunarno
Quantitative synthesis and meta-analysis plan: HS (lead)
Risk of bias assessment: HS (lead), NS, Sunarno
Validation (data extraction verification): NS
Visualization: DA
Project administration: DA
Supervision: HS, Sunarno (lead)
Arbitration for screening disagreements: Sunarno
Guarantor: Sunarno
Writing – original draft: DA (lead)
Writing – review & editing: DA, NS, HS, Sunarno

Conflicts of Interest

None declared.

Checklist 1
PRISMA 2020 checklist.
[PDF File, 196 KB - resprot_v15i1e86479_app1.pdf]

References

1. Pall ML. Electromagnetic fields act similarly in plants as in animals: probable activation of calcium channels via their voltage sensor. Curr Chem Biol 2016 Jul 19;10(1):74-82. [doi: 10.2174/2212796810666160419160433]
2. Yakymenko I, Sidorik E, Kyrylenko S, Chekhun V. Long-term exposure to microwave radiation provokes cancer growth: evidences from radars and mobile communication systems. Exp Oncol 2011 Jun;33(2):62-70. [Medline: 21716201]
3. Ghazanfarpour M, Kashani ZA, Pakzad R, et al. Effect of electromagnetic field on abortion: a systematic review and meta-analysis. Open Med 2021 May 14;16(1):1628-1641. [doi: 10.1515/med-2021-0384]
4. Razavimoghadam M, Sefidkar R, Ehrampoush MH, Teimouri F, Hassanabadi MHZ, Nokhostin F. The association of widely used electromagnetic waves exposure and pregnancy and birth outcomes in Yazd women: a cohort study. BMC Pregnancy Childbirth 2025 Apr 11;25(1):427. [doi: 10.1186/s12884-025-07512-4] [Medline: 40217134]
5. Zhou F, Ma C, Li YJ, Zhang M, Liu W. Effect of extremely low-frequency electromagnetic radiation on pregnancy outcome: a meta-analysis. Afr J Reprod Health 2023 May;27(5):95-104. [doi: 10.29063/ajrh2023/v27i5.9] [Medline: 37584934]
6. Irani M, Aradmehr M, Ghorbani M, Baghani R. Electromagnetic field exposure and abortion in pregnant women: a systematic review and meta-analysis. Malays J Med Sci 2023 Oct;30(5):70-80. [doi: 10.21315/mjms2023.30.5.6] [Medline: 37928787]
7. Cordelli E, Ardoino L, Benassi B, et al. Effects of radiofrequency electromagnetic field (RF-EMF) exposure on pregnancy and birth outcomes: a systematic review of experimental studies on non-human mammals. Environ Int 2023 Oct;180:108178. [doi: 10.1016/j.envint.2023.108178] [Medline: 37729852]
8. Electromagnetic fields. World Health Organization. URL: https://www.who.int/health-topics/electromagnetic-fields#tab=tab_1 [accessed 2026-02-06]
9. International Commission on Non-Ionizing Radiation Protection (ICNIRP). Guidelines for limiting exposure to electromagnetic fields (100 kHz to 300 GHz). Health Phys 2020;118(5):483-524. [doi: 10.1097/HP.0000000000001210]
10. Li DK, Chen H, Ferber JR, Odouli R, Quesenberry C. Exposure to magnetic field non-ionizing radiation and the risk of miscarriage: a prospective cohort study. Sci Rep 2017 Dec 13;7(1):17541. [doi: 10.1038/s41598-017-16623-8] [Medline: 29235463]
11. Takei R, Nagaoka T, Nishino K, Saito K, Watanabe S, Takahashi M. Specific absorption rate and temperature increase in pregnant women at 13, 18, and 26 weeks of gestation due to electromagnetic wave radiation from a smartphone. IEICE Commun Express 2018;7(6):212-217. [doi: 10.1587/comex.2018XBL0026]
12. Liu D, Li C, Kang Y, Zhou Z, Xie Y, Wu T. Numerical analysis for infant's unintentional exposure to 3.5 GHz plane wave radiofrequency electromagnetic fields by field test of fifth generation wireless technologies. Radio Sci 2017 Sep;52(9):1140-1148. [doi: 10.1002/2017RS006382]

XSL•FO
RenderX

13. Davis D, Birnbaum L, Ben-Ishai P, et al. Wireless technologies, non-ionizing electromagnetic fields and children: identifying and reducing health risks. Curr Probl Pediatr Adolesc Health Care 2023 Feb;53(2):101374. [doi: 10.1016/j.cppeds.2023.101374] [Medline: 36935315]
14. Wdowiak A, Mazurek PA, Wdowiak A, Bojar I. Effect of electromagnetic waves on human reproduction. Ann Agric Environ Med 2017 Mar 31;24(1):13-18. [doi: 10.5604/12321966.1228394] [Medline: 28378967]
15. Cervellati F, Franceschetti G, Lunghi L, et al. Effect of high-frequency electromagnetic fields on trophoblastic connexins. Reprod Toxicol 2009 Jul;28(1):59-65. [doi: 10.1016/j.reprotox.2009.03.010] [Medline: 19490996]
16. Abtin S, Seyedaghamiri F, Aalidaeijavadi Z, et al. A review on the consequences of molecular and genomic alterations following exposure to electromagnetic fields: remodeling of neuronal network and cognitive changes. Brain Res Bull 2024 Oct 15;217:111090. [doi: 10.1016/j.brainresbull.2024.111090] [Medline: 39349259]
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
18. Borenstein M, Higgins JPT, Hedges LV, Rothstein HR. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. Res Synth Methods 2017 Mar;8(1):5-18. [doi: 10.1002/jrsm.1230] [Medline: 28058794]
19. Guyatt G, Oxman AD, Akl EA, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011 Apr;64(4):383-394. [doi: 10.1016/j.jclinepi.2010.04.026] [Medline: 21195583]
20. Amandokht Saghezchi S, Azad N, Heidari R, et al. The effect of prenatal exposure to 2.4 GHz radio frequency on the histology and expression of the osteocalcin and RUNX2 gene of the forelimb in an NMRI mouse. J Lasers Med Sci 2019;10(4):283-289. [doi: 10.15171/jlms.2019.46] [Medline: 31875120]
21. International Commission on Non-Ionizing Radiation Protection (ICNIRP). Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields. Health Phys 1998;74:494-522. [Medline: 9525427]
22. Karatsi I, Bakogianni S, Koulouridis S. Assessment of electromagnetic exposure to a child and a pregnant woman inside an elevator in mobile frequencies. Telecom 2025;6(3):52. [doi: 10.3390/telecom6030052]
23. Tsarna E, Reedijk M, Birks LE, et al. Associations of maternal cell-phone use during pregnancy with pregnancy duration and fetal growth in 4 birth cohorts. Am J Epidemiol 2019 Jul 1;188(7):1270-1280. [doi: 10.1093/aje/kwz092] [Medline: 30995291]
24. Wang H, Tong J, Cao Y. Non-ionizing radiation-induced cellular senescence and age-related diseases. Radiat Med Prot 2024 Jun;5(2):68-74. [doi: 10.1016/j.radmp.2024.04.002]
25. Kömürcü Karuserci ̇ Ö, Çöl N, Demi ̇rel C. May electromagnetic field exposure during pregnancy have a negative effect on anthropometric measurements of the newborn? Cukurova Med J 2019;44:290-295. [doi: 10.17826/cumj.568315]
26. Ahlbom IC, Cardis E, Green A, et al. Review of the epidemiologic literature on EMF and health. Environ Health Perspect 2001 Dec;109(Suppl 6):911-933. [doi: 10.1289/ehp.109-1240626] [Medline: 11744509]
27. Kıvrak EG, Yurt KK, Kaplan AA, Alkan I, Altun G. Effects of electromagnetic fields exposure on the antioxidant defense system. J Microsc Ultrastruct 2017;5(4):167-176. [doi: 10.1016/j.jmau.2017.07.003] [Medline: 30023251]

Abbreviations

EMF: electromagnetic field
EMW: electromagnetic wave
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
PICOS: Population Intervention (or Exposure, for observational studies) Comparator Outcomes Study design
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
PROSPERO: International Prospective Register of Systematic Reviews
RF: radiofrequency
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by J Sarvestan; submitted 31.Oct.2025; peer-reviewed by AA Vireque, KEE Mills; revised version received 08.Jan.2026; accepted 09.Jan.2026; published 30.Mar.2026.

Please cite as:
Armalina D, Susilaningsih N, Sutanto H, Sunarno S
Determining the Impact of Nonionizing Electromagnetic Waves on Human Pregnancy and Teratogenicity: Protocol for Systematic Review and Meta-Analysis
JMIR Res Protoc 2026;15:e86479
URL: https://www.researchprotocols.org/2026/1/e86479
doi:10.2196/86479

XSL•FO
RenderX

© Desy Armalina, Neni Susilaningsih, Heri Sutanto, Sunarno Sunarno. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Components and Feasibility of Health Coaching Interventions for Self-Management of Type 2 Diabetes: Protocol for a Systematic Review

Rija Mir[1], BSc, MSc; Shoba Poduval[2], BMBS, MPH, PhD; Jessica Sheringham[1], MA, MSc, PhD; Fiona Louise Hamilton[1], MBBS, MSc, PhD

[1]Research Department of Primary Care and Population Health, University College London, Upper Third Floor, UCL Medical School (Royal Free Campus), Rowland Hill Street, London, United Kingdom

[2]Centre for Primary Care, Wolfson Institute of Population Health, Queen Mary University of London, London, United Kingdom

Corresponding Author:
Fiona Louise Hamilton, MBBS, MSc, PhD
Research Department of Primary Care and Population Health, University College London, Upper Third Floor, UCL Medical School (Royal Free Campus), Rowland Hill Street, London, United Kingdom

Abstract

Background: Type 2 diabetes (T2D) currently has no cure. However, extensive evidence suggests that addressing key risk factors through lifestyle changes can help individuals effectively self-manage their condition. Diabetes self-management primarily involves patients engaging in self-monitoring behaviors and adopting coping strategies to manage their long-term illness. In recent years, health coaching interventions have gained recognition as a valuable approach for providing personalized support, enabling patients to take an active role in managing their health. By fostering behavior change through goal setting, active learning, accountability, and empowerment, health coaching equips patients with the tools to proactively manage their condition over time. This approach is especially important for researchers and policymakers, as it underscores the need for acceptable, engaging, and personalized care interventions that have a lasting positive impact, promoting self-sufficiency and improved quality of life for individuals with T2D.

Objective: The objective of this study is to systematically review published, peer-reviewed, primary research studies to evaluate the effectiveness of health coaching interventions for T2D self-management, with glycated hemoglobin (HbA_{1c}) as the primary outcome. It aims to examine how intervention components, delivery modalities, and health coach characteristics relate to effectiveness, acceptability, and engagement.

Methods: This study protocol outlines the procedure for a systematic review that follows the methodology recommended by PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) guidelines. The review will include randomized controlled trials, quasi-experimental studies, and qualitative studies published in English up to December 1, 2024. The following databases will be used to conduct searches: MEDLINE, PsycINFO, Embase, CINAHL, Web of Science, and the Cochrane Central Register of Controlled Trials, with additional studies identified through citation searching and reference list screening. Included studies will consist of health coaching interventions delivered one-to-one through face-to-face, remote or digital, or artificial intelligence–supported formats. The primary outcome of interest is HbA_{1c}, alongside additional clinical outcomes and qualitative data related to acceptability and engagement. Study selection will involve independent screening by reviewers, with disagreements resolved by consensus. Quantitative data will be synthesized using narrative synthesis, and qualitative findings will be analyzed using thematic and framework analysis within a staged, segregated synthesis approach.

Results: As of September 2025, database searches have been conducted, and decisions have been made related to the included publications. Data extraction and analysis are currently ongoing. Results are expected to be published in summer 2026. The review is underway and is anticipated to be completed by May 2026.

Conclusions: This review will synthesize evidence on the effectiveness and implementation of health coaching interventions for T2D and is expected to inform the design of effective and sustainable coaching models for diabetes self-management.

Trial Registration: PROSPERO CRD42025637862; https://www.crd.york.ac.uk/PROSPERO/view/CRD42025637862

International Registered Report Identifier (IRRID): PRR1-10.2196/71383

(JMIR Res Protoc 2026;15:e71383) doi:10.2196/71383

XSL•FO
RenderX

KEYWORDS

health coaching; digital health interventions; eHealth; type 2 diabetes; T2D; self-management; digital divide; remote coaching; RC; virtual coaching; mobile health

Introduction

Type 2 Diabetes

Recent statistics from the World Health Organization (WHO) show that approximately 41 million people globally die of a chronic illness each year, of which 2 million are attributed to diabetes mellitus alone [1]. Currently, while more than half a billion people live with this condition, with 96% of these cases identified as type 2 diabetes (T2D), these numbers are projected to double in the next 30 years, estimating an increase in global diabetes-related health expenditure to US $1054 billion by 2045 [2,3]. T2D occurs when the body is unable to regulate blood glucose levels because it is either resistant to the insulin being produced or the pancreas does not make enough insulin. Living with this lifelong chronic condition not only affects individuals and their families, but associated risks, such as morbidity, premature mortality, mental health problems, hospitalizations, medications, and other medical costs, also significantly contribute to the overall global burden of T2D on health care systems [4]. It has also been predicted that T2D will be a serious clinical and financial challenge to the UK National Health Service for the next 6 decades, with a significant increase in patient caseloads and the population burden for diabetes-related complications [5].

There is no cure for T2D yet; however, extensive evidence indicates that minimizing several risk factors through lifestyle changes can help people self-manage their condition [6-8]. Self-management in diabetes primarily consists of patients undertaking self-monitoring behaviors and activities or using strategies to cope with their long-term illness. These include frequent monitoring of blood glucose levels, maintaining a healthy daily intake of nutrition, adherence to prescribed medication, insulin dose adjustments, and regular physical activity [9]. Self-management not only empowers patients to make informed decisions and ultimately achieve their health goals but also helps prevent further complications, reducing the strain on health care services and the overall burden of disease [10]. Prior study findings have significantly established that self-management of T2D improves clinical health outcomes and well-being through an overall positive impact on patient health behaviors, such as engaging in moderate physical activity, consuming a low glycemic index diet, and adhering to medication, among others [11-14]. Thus, lifelong proactive diabetes self-management is crucial for T2D care.

Health Coaching

More recently, lifestyle interventions, such as health coaching, have gained increasing recognition as an effective strategy for providing personalized support to enable patients to actively self-manage their health conditions [15]. Rooted in health education and counseling psychology, health coaching adopts a one-on-one, personalized approach that prioritizes patient-centered care [16-19]. It supports patients in learning how to proactively self-manage their condition throughout their lives, by promoting behavior change through goal setting, active learning processes, accountability, and patient empowerment [20,21]. Central to health coaching is the relationship cultivated between the coach and patient through multiple interactions, which significantly influences the patient's understanding, learning, behavior, and progress [22]. Substantial study findings confirm that health coaching interventions have shown a significant positive impact on clinical and behavioral outcomes in patients with T2D [23-27]. However, some systematic reviews have found mixed results, with little to no significant improvement in some clinical outcomes and other health outcomes, such as quality of life, due to self-management [28,29]. Several reasons have been attributed to the mixed evidence in current literature, such as the heterogeneity of the interventions, imprecise descriptions of interventions, and lack of consensus or guidance on the components of a health coaching practice, resulting in variations in the reporting of outcomes [13]. Inconsistencies in the current literature regarding the effectiveness of health coaching interventions justify the need for further in-depth review to aid in our understanding and to identify factors contributing to their impact.

Components of Health Coaching

There is a lack of standardization of how "health coach" or "health coaching" is defined [28]. Numerous definitions of health coaching are used in the current literature and there is an absence of a standardized strategy for coaching practice. Evidence from many studies highlights that health coaches are described as "educator," "navigator," "facilitator," or "partner," indicating that there is no clear definition of the role of a health coach, thus warranting critical analysis to establish key characteristics of a health coach. A health coach is central to a health coaching intervention because they are responsible for providing an effective coaching context where they influence a patient's thoughts and behaviors to subsequently achieve their health goals. There is a strong need to systematically review key ingredients of a standard health coaching practice, such as theoretical underpinnings, the type of content to be included in the intervention, and strategies used to deliver the intervention, to determine the effectiveness of a health coaching intervention. At present, there is no consensus in the current literature on how to design an effective health coaching intervention that includes the most appropriate active components for behavior change [16]. Therefore, a focused systematic review will help identify which coaching practices are more appropriate and effective to replicate and use; this includes critically analyzing the components that make up the intervention (duration, frequency, and mode of delivery of sessions). Consequently, it will help identify factors contributing to effective, personalized, and engaging interactions with patients within a coaching context.

In addition, previous reviews have identified a paucity of research investigating specific health coaching program components in-depth [26,30]. For example, 1 aspect that remains underexplored but holds potential in evaluating the effectiveness

XSL•FO
RenderX

of a health coach for T2D self-management is the delivery modality used to convey information to the patient [22]. This refers to the methods used for interacting with the patient, including (1) "face-to-face" (FF) coaching sessions, where both the coach and the individual are physically present in the same location during the session; (2) "remote" coaching (RC; also known as digital or tele coaching), conducted online without in-person contact through phone calls, video calls, SMS text messaging, emails, mobile apps, and similar means that rely on technology like mobile phones, computers, tablets, and other smart devices; and (3) "virtual coaching or artificial intelligence (AI)–supported coaching," which uses autonomous systems or software to coach patients, such as AI coaches, e-coaches, virtual assistants, chatbots, and so on [31]. Previous literature has primarily examined and compared FF with RC health coaching, either exclusively or in combination, highlighting potential research gaps in fully examining and comparing interventions based on all 3 delivery modalities, as well as their combinations, for patients with T2D [32,33]. Accordingly, the literature recommends examining individual coaching practices and intervention components to clarify their role in intervention effectiveness and to inform future intervention design [34]. These include coaching techniques, intervention delivery modalities, frequency, and duration to determine how these features influence clinical and patient outcomes. Given that glycated hemoglobin (HbA_{1c}) levels in patients with T2D are a strong indicator of disease management, understanding how each intervention component affects these levels is also integral and warranted.

Technology-Enabled Health Coaching

Health coaching interventions for T2D are increasingly delivered through technology-enabled formats, including remote digital platforms and emerging AI-supported systems. The rapid expansion in the use of digital health technologies, such as eHealth and mobile health, has been largely accelerated by the COVID-19 pandemic, resulting in the widespread adoption of remote models of care within primary health care settings [35-37]. These developments have facilitated greater flexibility in service delivery and supported ongoing self-management for people with chronic conditions, including T2D [12,36,38]. As a result, technology-enabled health coaching has become an important modality for delivering personalized behavioral support.

Evidence suggests that remote health coaching interventions may improve access and convenience and have the potential to support behavior change; however, findings regarding their effectiveness remain inconsistent [39-41]. While some studies report positive impacts on self-management and clinical outcomes, others indicate limited benefits when compared with FF coaching, particularly with respect to sustaining engagement and behavior change over time [40-42]. These mixed findings highlight the importance of examining clinical effectiveness alongside intervention components and delivery characteristics when synthesizing outcomes such as HbA_{1c} across studies.

The expansion of digital health has also drawn attention to contextual factors that may influence patient experiences, particularly inequalities in access to and use of digital technologies, commonly described as the digital divide [43-47]. Evidence indicates that older adults, individuals from minority ethnic groups, and those from lower socioeconomic backgrounds may be disproportionately affected by limited digital access and literacy, which may shape engagement with remote or technology-based coaching interventions [48-52]. Although these factors will not be considered as primary outcomes when reviewed, they provide important context for interpreting patient-reported findings related to acceptability and engagement.

In parallel, growing scientific and policy interest in the application of AI within health care has led to the emergence of AI-supported coaching tools, including chatbots and virtual assistants, designed to support self-management behaviors [53]. However, existing systematic reviews have largely focused on FF and RC models, with limited examination of AI-based approaches [32,34]. Thus, this review will examine FF, remote or digital, and AI-supported health coaching interventions and synthesize HbA_{1c} outcomes alongside intervention components and contextual influences, enabling a comparative assessment of delivery modalities relevant to T2D self-management.

Acceptability and Engagement

Although clinical outcomes, such as HbA_{1c} levels, have commonly been used to assess the effectiveness of digital health interventions, their real-world success largely depends on patients finding these interventions acceptable and maintaining engagement over time. Acceptability refers to an individual's willingness to initiate and continue using an intervention [54]. This willingness is shaped by personal factors, such as beliefs, prior knowledge, and attitudes, as well as broader social and contextual influences [55]. Consequently, acceptability cannot be assumed to be uniform across populations, as it varies according to individual needs, experiences, and circumstances.

Engagement, often operationalized through quantitative metrics such as frequency or duration of use, is similarly multifaceted[56]. As Nicholas et al [57] argue, such measures may fail to capture key cognitive, emotional, and social determinants of engagement, including motivation, perceived relevance, and the availability of support. Qualitative evidence further demonstrates that the interplay between intervention content, user experience, and environmental context can either facilitate or hinder sustained participation. Thus, it is crucial to assess the acceptability and engagement of an intervention to determine its effectiveness and long-term sustainability.

Theoretical frameworks offer additional insight into mechanisms assessing acceptability and engagement as key indicators to determine the effectiveness of an intervention. Drawing on the Unified Theory of Acceptance and Use of Technology [58], there are 4 central determinants of technology uptake: perceived benefit, ease of use, social influence, and availability of support with moderators such as age, gender, and prior technology experience further influencing user behavior [59]. Similarly, behavior change interventions are more likely to be accepted when they enable users to establish personally meaningful goals, receive tailored feedback, and understand the rationale for recommended behaviors [60]. Empirical evidence supporting

XSL•FO
RenderX

these propositions remains limited, indicating the need for further research to determine whether theory-driven designs consistently enhance engagement and clinical outcomes.

Perski and Short [55] emphasized the interdependence of acceptability, engagement, and effectiveness, positing that initial willingness to adopt an intervention influences ongoing engagement, which subsequently affects clinical results. Within diabetes care, engagement has been identified as a critical component of effective self-management [61], with features such as access to a health coach, digital monitoring tools, and personalized feedback contributing to sustained participation [62,63]. However, findings across studies remain mixed; some suggest that supplementing digital programs with FF support may reduce subsequent digital engagement [64], whereas others report minimal effects of RC on long-term use [65].

Given the increasing integration of digital and AI-enabled coaching within diabetes self-management, reliance solely on clinical outcomes provides an incomplete assessment of intervention success. Interventions that achieve improvements in HbA_{1c} but are perceived as unacceptable or fail to engage users are unlikely to be effective in routine practice. Furthermore, structural barriers associated with the digital divide, including socioeconomic constraints, cultural factors, and disparities in digital literacy, may limit uptake among certain populations [48-52]. Accordingly, a systematic review that examines not only effectiveness but also patient acceptability and engagement is essential. Such an approach will enable the identification of intervention components and delivery modalities most conducive to promoting long-term, equitable self-management for individuals with T2D.

Moreover, this holds significant importance for researchers and policymakers, ensuring that personalized care and health coaching interventions introduced to patients are sustainable and have a lasting positive impact on their lives, fostering self-sufficiency in managing their health condition. To the best of our knowledge, no prior systematic reviews have examined the acceptability, engagement, and effectiveness of health coaching interventions exclusively for T2D self-management. Therefore, this systematic review will aim to explore in-depth the contextual factors associated with the acceptability and engagement of health coaching interventions, which will aid in better understanding of the most effective intervention components. The findings obtained from studying acceptability and engagement will also help identify barriers and facilitators, offering a set of recommendations aimed at enhancing the effectiveness of health coaching interventions.

Aims and Objectives

Aims

The primary aim of this systematic review is to evaluate the effectiveness of health coaching interventions for T2D self-management, with glycemic control (HbA_{1c}) as the primary outcome. The review also aims to identify and compare intervention components, health coach characteristics, and delivery modalities associated with improved HbA_{1c} outcomes and to draw inferences regarding the feasibility and sustainability of different coaching approaches based on reported outcomes and intervention characteristics. Findings from this review will be used to inform recommendations for the design of effective and sustainable health coaching interventions for adults with T2D. This review seeks to address key research questions: (1) To what extent do health coaching interventions improve HbA_{1c} levels among adults with T2D compared with standard care or control conditions? (2) Which intervention components and health coach characteristics are associated with greater improvements in HbA_{1c} outcomes across included studies? (3) Which intervention components are most frequently present in effective health coaching interventions, and how do these component patterns vary across delivery modalities? (4) What insights do the included studies provide regarding the feasibility and sustainability of different delivery modalities, as inferred from HbA_{1c} outcomes and intervention characteristics?

Objectives

Drawing on existing evidence, this systematic review aims to (1) evaluate the effectiveness of health coaching interventions on glycemic control, with HbA_{1c} as the primary outcome, compared with standard care or control conditions; (2) identify intervention components and health coach characteristics associated with greater improvements in HbA_{1c} outcomes; and (3) compare delivery modalities (FF, remote or digital, and AI-supported coaching) in relation to effectiveness, feasibility, and sustainability, based on HbA_{1c} outcomes and intervention characteristics.

Methods

This protocol follows the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) statement (Checklist 1) [66]. This protocol has been registered with PROSPERO (CRD42025637862).

Eligibility Criteria

Types of Studies

We will include experimental (randomized controlled trial), quasi-experimental, and qualitative study designs, which focus on in-depth interviews and focus group discussions, published until December 1, 2024.

The systematic review will include primary evidence from studies published in English that have high psychometric properties and methodological approaches used to minimize bias. Thus, the review will exclude observational studies, case reports, case series, letters to the editor, preprint papers, expert opinions, editorials, and narrative reviews (if a narrative review is considered relevant to the study, the list of included studies for that paper will be examined to identify studies that may have been missed during the search), abstract-only studies, conference proceedings, and study protocols.

Types of Participants

Studies with adult participants aged 18 years and older, having a confirmed diagnosis of type 2 diabetes or maturity-onset diabetes of the young, will be included. Studies with participants having type 1 diabetes or gestational diabetes will be excluded.

XSL•FO
RenderX

Types of Interventions

The review will include studies of health coaching interventions that provide one-on-one coaching delivered FF, remotely via digital technology, and virtually through autonomous systems or AI to individuals within a coaching context comprising a health coach.

Types of Comparators

The review will include studies with comparator interventions, such as (1) treatment as usual (standard care, leaflets, and health education programs) and (2) health interventions without one-on-one coaching, such as in groups. Studies with pretest or posttest study designs with no comparator at all will also be included.

Types of Outcome Measures

The review will include studies that report HbA_{1c} as an outcome measure related to T2D self-management. The study evaluations must include HbA_{1c} outcomes, and any additional clinical outcomes reported in these studies, such as fasting plasma glucose, lipid profiles, body weight, BMI, waist circumference, blood pressure, or urinary albumin measurements, will also be included where available. The review will also consider any additionally reported qualitative or patient-reported outcomes in these studies, such as facilitators, barriers, acceptability, uptake, or engagement relating to health coaching interventions.

Search Strategy

The search strategy for the systematic review will aim to focus on both published and unpublished literature. An initial search of the databases will be conducted to identify research papers on the topic area. The titles and abstracts of relevant studies, as well as the keywords or index terms used to describe the studies, will be reviewed, and the most relevant terms will be selected to perform a full search using the following databases:

- MEDLINE (via Ovid, 1946 to December 1, 2024)
- PsycINFO (via Ovid, 1806 to December 1, 2024)
- Embase (via Ovid, 1947 to December 1, 2024)
- CINAHL (via EBSCOhost)
- Web of Science (core collection)
- Cochrane Central Library

These databases were selected to ensure comprehensive coverage of biomedical, nursing, behavioral, and digital health literature relevant to health coaching interventions for T2D. Search terms will include various synonyms, spellings, and acronyms for each identified keyword and index term. For example, for the MEDLINE database, the search will include MEDLINE thesaurus MeSH (Medical Subject Headings) terms as well as appropriate keywords related to “type 2 diabetes,” “health coaching interventions,” “remote coaching,” “e-coaching,” “virtual coaching,” “self-care,” “self-management,” “modalities,” “clinical outcomes,” “acceptability,” “engagement,” and “patient reported outcomes” (Multimedia Appendix 1). Similarly, the authors will perform all database searches in collaboration with a research librarian, and the search strategy will be adapted for each included information source. To focus and structure the search for optimal results, various search functions such as thesaurus, Boolean operators, abstract or title or keywords, phrase, free text, and advanced search will be used.

To identify any additional studies that did not appear during database searches, citation searches, and manual searches of the reference lists of all studies selected for critical appraisal will be performed. Moreover, citation searches will be conducted in Google Scholar and Web of Science. If it is difficult to access a study or if any questions arise during the selection or extraction processes, the authors of identified articles may be contacted. All bibliographic data from the search results will be transferred to EndNote (Clarivate), a reference information system software, and the finalized results will be collated and extracted to Microsoft Excel.

Study Selection Process

One author (RM) will be responsible for conducting the database search and managing retrieved records. Following the removal of duplicates, all titles and abstracts identified through the search will be independently screened by at least 2 reviewers against the predefined inclusion criteria. Any records that do not meet the eligibility criteria will be excluded, with reasons for exclusion documented. To ensure consistency in the application of the inclusion criteria, reviewers will first pilot the screening process on an initial subset of studies and discuss discrepancies prior to full screening. All remaining titles and abstracts will then undergo independent double screening, and disagreements will be resolved through discussion and consensus.

Full-text studies of potentially eligible studies will be retrieved and independently assessed by at least 2 reviewers to confirm eligibility. Any disagreements at the full-text stage will be resolved through discussion, with the involvement of a third reviewer if necessary. The study selection process will be documented using a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram, detailing the number of records identified, screened, excluded, and included in the review.

Data Collection Process and Data Extraction

The studies selected for full-text review will be independently examined, and the relevant data will be collected and recorded in a Microsoft Excel file. The data extracted will include specific details about the population, study methods, interventions, and outcomes of significance based on the review objective. More specifically, categories that are considered relevant include characteristics of the health coaching intervention and key results, characteristics of the health coach, and baseline characteristics of the population (Multimedia Appendix 1). The first author will conduct a full-text review of any 5 randomly selected studies and share them with the other authors to independently review and assess whether all the relevant information has been accurately extracted and meets the inclusion criteria. The feedback will be discussed, and any disagreements that arise between the reviewers will be resolved through discussion or with a third or additional reviewer. The author will then continue to review and extract data from the remaining studies based on their feedback.

XSL•FO
RenderX

Outcomes and Prioritization

The primary outcome of interest for this review is HbA_{1c}, reported as a validated clinical measure of glycemic control in adults with T2D. HbA_{1c} outcomes will be extracted as reported in the included studies, including the timing of outcome assessment. Where available, additional clinical outcomes will also be extracted, including fasting plasma glucose, lipid profiles, body weight, BMI, waist circumference, blood pressure, and urinary albumin measurements.

In addition to clinical outcomes, the review will include studies reporting qualitative or patient-reported outcomes related to the use of health coaching interventions. These outcomes will focus on acceptability and engagement, examined using qualitative data and patient-reported findings as defined in the original studies. Acceptability will be interpreted as patients' reported willingness to use, perceived suitability of, and satisfaction with the intervention. Engagement will be interpreted based on reported patterns of use, interaction with the intervention or coach, and sustained participation over time. Data relating to facilitators, barriers, uptake, and user experiences will be synthesized narratively to contextualize and interpret variation in intervention effectiveness across delivery modalities and intervention components.

Risk of Bias and Quality Assessment

Risk of bias and methodological quality of the included studies will be independently assessed by at least 2 reviewers using design-appropriate appraisal tools. Randomized controlled trials will be evaluated using the Cochrane Collaboration's Risk of Bias tool, nonrandomized quantitative studies will be assessed using the ROBINS-I (Risk of Bias in Nonrandomized Studies of Interventions) tool, and qualitative studies will be appraised using the CASP (Critical Appraisal Skills Program) checklist (Checklist 2) [67-69].

To ensure the consistent application of bias assessment criteria, reviewers will pilot the appraisal tools on a subset of included studies and discuss discrepancies prior to completing assessments for all studies. Any disagreements arising during the risk-of-bias assessment will be resolved through discussion and consensus.

All studies will be included in the data extraction and synthesis regardless of their assessed risk of bias. However, risk-of-bias judgments will be taken into account during the interpretation of the findings, with greater caution applied when interpreting results from studies assessed as having a high risk of bias. The potential influence of study quality on observed patterns of effectiveness will be explored narratively as part of the synthesis. Where necessary, authors of the included studies may be contacted to clarify methodological details or to obtain missing information relevant to the risk-of-bias assessment. In addition, the overall quality and certainty of the body of evidence may be assessed using the grading of recommendations, assessment, development, and evaluations (GRADE) approach, where appropriate [70].

Data Synthesis, Narrative Synthesis, or Meta-Analysis

All included studies will be reviewed in depth, and relevant data will be extracted and synthesized in accordance with the review objectives and research questions. Given the inclusion of both quantitative and qualitative study designs, a segregated, staged synthesis approach will be used, whereby quantitative and qualitative evidence will first be analyzed separately using methods appropriate to each data type, followed by an integrative interpretive stage.

For quantitative studies, changes in HbA_{1c} levels will be synthesized using narrative synthesis, structured around study design, intervention characteristics, delivery modality, and effectiveness relative to comparator conditions. Meta-analysis will be considered where sufficient homogeneity exists in study design, intervention characteristics, and outcome reporting. However, substantial heterogeneity across interventions and outcome measures may preclude statistical pooling.

For qualitative studies, data relating to acceptability, engagement, barriers, and facilitators will be analyzed using thematic analysis, followed by framework analysis. An initial analytical framework will be developed based on existing theory and prior research, including the Bandura [71] reciprocal determinism theory, and will be sufficiently flexible to accommodate inductively derived themes. Line-by-line coding will be conducted to generate initial codes, which will be organized into higher-order themes and subthemes [72]. The framework will be iteratively refined by integrating emerging themes and restructuring categories as required to ensure comprehensive representation of the data. Qualitative analyses will be reviewed by multiple authors, with discrepancies resolved through discussion and consensus. Formal double-coding, use of qualitative analysis software, and calculation of inter-rater reliability statistics (eg, κ) are not planned, given the interpretive nature of the thematic and framework analysis; instead, credibility will be ensured through reviewer discussion and consensus.

In the final stage, findings from the quantitative and qualitative syntheses will be integrated interpretively. Qualitative themes will be used to contextualize and explain patterns observed in quantitative outcomes, particularly variations in HbA_{1c} effectiveness across intervention components and delivery modalities. This staged integration will support the development of a conceptual framework identifying how intervention characteristics, delivery strategies, and contextual factors interact to influence effectiveness, acceptability, and engagement with health coaching interventions for T2D self-management.

Heterogeneity in intervention design, delivery modality, follow-up duration, and outcome measurement is anticipated. HbA_{1c} outcomes will be extracted and synthesized as reported in the included studies, with the timing of measurement documented to support interpretation. To accommodate variability across studies, quantitative findings will be synthesized using narrative synthesis, and meta-analysis will only be undertaken where outcome measures and time points are sufficiently comparable. Where sufficient data are available, descriptive subgroup analyses will be conducted by delivery

XSL•FO
RenderX

modality (FF, remote or digital, and AI-supported interventions) and by key intervention components to explore potential sources of heterogeneity. Formal sensitivity analyses will not be undertaken due to anticipated variability in study design, intervention characteristics, and outcome measurement; instead, the robustness of findings will be assessed through structured comparative narrative synthesis across subgroups.

Results

The systematic review is in progress, and the final draft is anticipated to be completed by May 2026. We expect to report a summary of the findings of included published and unpublished studies that involve the delivery and use of health coaching interventions for patients to self-manage their T2D within a coaching context, including changes in HbA_{1c} levels and other clinical outcomes, as well as the results of qualitative studies that identify the barriers, facilitators, and challenges associated with the use and feasibility of health coaching interventions by patients with T2D. Database searches have been conducted, and decisions have been made related to the included publications. Data extraction and analysis are currently ongoing. Results are expected to be published in summer 2026.

Discussion

Principal Findings

The systematic review seeks to synthesize from existing evidence the impacts of using health coaching interventions on glycemic control in adults with T2D. It is anticipated that interventions incorporating structured goal setting, personalized feedback, regular coach-patient interaction, and clearly defined coaching roles will be more consistently associated with favorable HbA_{1c} outcomes. The review also aims to identify differences in effectiveness and feasibility across FF, remote or digital, and AI-supported coaching formats, reflecting the diversity of delivery approaches currently used in practice.

This review is expected to extend existing evidence by synthesizing findings across multiple delivery modalities and explicitly linking HbA_{1c} outcomes to intervention components and health coach characteristics. By adopting a component-focused synthesis, the review aims to provide a more nuanced understanding of how and why certain coaching interventions may be more effective than others.

In addition, by integrating qualitative findings on barriers, facilitators, acceptability, and engagement, this review will contextualize quantitative effectiveness data and support the interpretation of variation in outcomes. This approach is particularly relevant given the growing reliance on remote and technology-enabled coaching interventions, where patient engagement and contextual factors may influence real-world effectiveness.

A key strength of this review is its inclusive and structured synthesis approach, combining quantitative and qualitative evidence using a staged, segregated design. This enables clinical effectiveness (HbA_{1c}) to be prioritized while also incorporating patient-reported experiences relevant to acceptability, engagement, feasibility, and sustainability. The inclusion of FF, remote or digital, and AI-supported interventions allows for a comprehensive comparison across delivery modalities that have not been examined in prior reviews [32-34].

However, several limitations are anticipated. Heterogeneity in intervention design, outcome measurement, follow-up duration, and reporting practices may restrict the feasibility of a meta-analysis, thereby requiring reliance on narrative synthesis. Additionally, acceptability and engagement are expected to be variably defined and reported across studies, which may constrain direct comparisons. Risk-of-bias assessments will, therefore, be used to inform interpretation, and conclusions will be drawn cautiously where evidence quality is limited.

In addition to these methodological limitations, several practical constraints of the planned review may introduce potential bias. The restriction to English-language publications may result in language bias and the exclusion of relevant studies conducted in non–English-speaking settings. Furthermore, the exclusion of gray literature, such as reports, theses, and unpublished studies, may increase the risk of publication bias by preferentially capturing studies with positive or statistically significant findings. These constraints, though common, may limit the comprehensiveness and generalizability of the evidence base. Findings from this review will, therefore, be interpreted with these considerations in mind, and gaps arising from these exclusions will be highlighted to inform future research.

The findings of this review are expected to inform future intervention design by identifying which health coaching components and delivery approaches are most consistently associated with improved glycemic control. For researchers, the synthesis will highlight gaps in reporting and areas where standardization of intervention descriptions and outcome measures is needed. For practitioners and policymakers, the review may offer guidance on selecting or designing health coaching models that are not only effective but also feasible and acceptable across diverse care settings.

Findings from this systematic review will be disseminated through publication in a peer-reviewed journal and presentation at relevant academic and clinical conferences. Results will also be shared with stakeholders involved in diabetes care and digital health implementation to support evidence-informed decision-making regarding the development and scaling of health coaching interventions for T2D.

Conclusions

To conclude, based on current literature, our understanding is that no prior systematic reviews have fully examined the feasibility and active components of existing health coaching interventions exclusively for T2D self-management. Thus, this review will synthesize evidence on the effectiveness of health coaching interventions for T2D, with HbA_{1c} as the primary outcome. By examining intervention components and delivery modalities, it will inform the design of effective and sustainable health coaching interventions.

XSL•FO
RenderX

Funding

The authors declared no financial support was received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
OVID search strategy.
[DOCX File, 16 KB - resprot_v15i1e71383_app1.docx]

Checklist 1
PRISMA-P checklist.
[DOCX File, 34 KB - resprot_v15i1e71383_app2.docx]

Checklist 2
CASP checklist.
[PDF File, 277 KB - resprot_v15i1e71383_app3.pdf]

References

1. Noncommunicable diseases. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases#:~:text=Key%20facts,%2D%20and%20middle%2Dincome%20countries [accessed 2026-03-12]
2. Ahmad E, Lim S, Lamptey R, Webb DR, Davies MJ. Type 2 diabetes. Lancet 2022 Nov 19;400(10365):1803-1820. [doi: 10.1016/S0140-6736(22)01655-5] [Medline: 36332637]
3. The Lancet. Diabetes: a defining disease of the 21st century. Lancet 2023 Jun 24;401(10394):2087. [doi: 10.1016/S0140-6736(23)01296-5] [Medline: 37355279]
4. Lin X, Xu Y, Pan X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. Sci Rep 2020 Sep 8;10(1):14790. [doi: 10.1038/s41598-020-71908-9] [Medline: 32901098]
5. Bagust A, Hopkinson PK, Maslove L, Currie CJ. The projected health care burden of type 2 diabetes in the UK from 2000 to 2060. Diabet Med 2002 Jul;19 Suppl 4:1-5. [doi: 10.1046/j.1464-5491.19.s4.2.x] [Medline: 12121330]
6. Type 2 diabetes. Mayo Clinic. URL: https://www.mayoclinic.org/diseases-conditions/type-2-diabetes/symptoms-causes/syc-20351193 [accessed 2026-03-12]
7. Welch A, Vella A. What is type 2 diabetes? Medicine 2022 Oct;50(10):625-631. [doi: 10.1016/j.mpmed.2022.07.003]
8. Shrivastava SR, Shrivastava PS, Ramasamy J. Role of self-care in management of diabetes mellitus. J Diabetes Metab Disord 2013 Mar 5;12(1):14. [doi: 10.1186/2251-6581-12-14] [Medline: 23497559]
9. Sigurdardóttir AK. Self-care in diabetes: model of factors affecting self-care. J Clin Nurs 2005 Mar;14(3):301-314. [doi: 10.1111/j.1365-2702.2004.01043.x] [Medline: 15707440]
10. Holman HR. The relation of the chronic disease epidemic to the health care crisis. ACR Open Rheumatol 2020 Mar;2(3):167-173. [doi: 10.1002/acr2.11114] [Medline: 32073759]
11. Amin M, Kerr D, Atiase Y, Aldwikat RK, Driscoll A. Effect of physical activity on metabolic syndrome markers in adults with type 2 diabetes: a systematic review and meta-analysis. Sports (Basel) 2023 May 9;11(5):101. [doi: 10.3390/sports11050101] [Medline: 37234057]
12. Adu MD, Malabu UH, Malau-Aduli AEO, Malau-Aduli BS. Enablers and barriers to effective diabetes self-management: a multi-national investigation. PLoS One 2019;14(6):e0217771. [doi: 10.1371/journal.pone.0217771] [Medline: 31166971]
13. Carpenter R, DiChiacchio T, Barker K. Interventions for self-management of type 2 diabetes: an integrative review. Int J Nurs Sci 2019 Jan 10;6(1):70-91. [doi: 10.1016/j.ijnss.2018.12.002] [Medline: 31406872]
14. Health & wellness: You can manage and thrive with diabetes. American Diabetes Association. URL: https://diabetes.org/health-wellness [accessed 2026-03-12]
15. Supported self-management: health coaching guide. National Health Service England. 2023. URL: https://www.england.nhs.uk/long-read/health-coaching/ [accessed 2026-03-12]
16. Almulhim AN, Hartley H, Norman P, Caton SJ, Doğru OC, Goyder E. Behavioural change techniques in health coaching-based interventions for type 2 diabetes: a systematic review and meta-analysis. BMC Public Health 2023 Jan 13;23(1):95. [doi: 10.1186/s12889-022-14874-3] [Medline: 36639632]
17. Todorova I. Health coaching research. In: International Handbook of Evidence-Based Coaching Theory, Research and Practice: Springer International Publishing; 2022:429-445. [doi: 10.1007/978-3-030-81938-5_35]

XSL•FO
RenderX

18. Lindner H, Menzies D, Kelly J, Taylor S, Shearer M. Coaching for behaviour change in chronic disease: a review of the literature and the implications for coaching as a self-management intervention. Aust J Prim Health 2003 Aug 1;9(3):177-185. [doi: 10.1071/PY03044]
19. Butterworth SW, Linden A, McClay W. Health coaching as an intervention in health management programs. Dis Manag Health Outcomes 2007;15(5):299-307. [doi: 10.2165/00115677-200715050-00004]
20. Lambrinou E, Hansen TB, Beulens JW. Lifestyle factors, self-management and patient empowerment in diabetes care. Eur J Prev Cardiol 2019 Dec;26(2_suppl):55-63. [doi: 10.1177/2047487319885455] [Medline: 31766913]
21. Wolever RQ, Simmons LA, Sforzo GA, et al. A systematic review of the literature on health and wellness coaching: defining a key behavioral intervention in healthcare. Glob Adv Health Med 2013 Jul;2(4):38-57. [doi: 10.7453/gahmj.2013.042] [Medline: 24416684]
22. Weimann TG, Schlieter H, Brendel AB. Virtual coaches. Bus Inf Syst Eng 2022 Aug;64(4):515-528. [doi: 10.1007/s12599-022-00757-9]
23. Racey M, Jovkovic M, Alliston P, Ali MU, Sherifali D. Diabetes health coach in individuals with type 2 diabetes: a systematic review and meta analysis of quadruple aim outcomes. Front Endocrinol (Lausanne) 2022;13:1069401. [doi: 10.3389/fendo.2022.1069401] [Medline: 36589795]
24. Ernawati U, Wihastuti TA, Utami YW. Effectiveness of diabetes self-management education (DSME) in type 2 diabetes mellitus (T2DM) patients: systematic literature review. J Public Health Res 2021 Apr 14;10(2):jphr-2021. [doi: 10.4081/jphr.2021.2240] [Medline: 33855427]
25. Lin CL, Huang LC, Chang YT, Chen RY, Yang SH. Effectiveness of health coaching in diabetes control and lifestyle improvement: a randomized-controlled trial. Nutrients 2021 Oct 29;13(11):3878. [doi: 10.3390/nu13113878] [Medline: 34836135]
26. Pirbaglou M, Katz J, Motamed M, Pludwinski S, Walker K, Ritvo P. Personal health coaching as a type 2 diabetes mellitus self-management strategy: a systematic review and meta-analysis of randomized controlled trials. Am J Health Promot 2018 Sep;32(7):1613-1626. [doi: 10.1177/0890117118758234] [Medline: 29658286]
27. Sherifali D, Viscardi V, Bai JW, Ali RMU. Evaluating the effect of a diabetes health coach in individuals with type 2 diabetes. Can J Diabetes 2016 Feb;40(1):84-94. [doi: 10.1016/j.jcjd.2015.10.006] [Medline: 26827684]
28. Mohamed N, Al H, Alamri F, Tofek A. Effectiveness of health coaching on diabetic patients: a systematic review and meta-analysis. Tradit Med Res 2019;4(6):314-325. [doi: 10.53388/TMR20191024143]
29. Veazie S, Winchell K, Gilbert J, et al. Rapid evidence review of mobile applications for self-management of diabetes. J Gen Intern Med 2018 Jul;33(7):1167-1176. [doi: 10.1007/s11606-018-4410-1] [Medline: 29740786]
30. Zimbudzi E, Okada H, Funnell MM, Hamaguchi M. Editorial: Innovation in diabetes self–care management and interventions. Front Endocrinol 2023;14. [doi: 10.3389/fendo.2023.1269437] [Medline: 37670874]
31. Fadhil A, Wang Y, Reiterer H. Assistive conversational agent for health coaching: a validation study. Methods Inf Med 2019 Jun;58(1):9-23. [doi: 10.1055/s-0039-1688757] [Medline: 31117129]
32. Singh H, Kennedy GA, Stupans I. Does the modality used in health coaching matter? A systematic review of health coaching outcomes. Patient Prefer Adherence 2020;14:1477-1492. [doi: 10.2147/PPA.S265958] [Medline: 32904668]
33. Bus K, Peyer KL, Bai Y, Ellingson LD, Welk GJ. Comparison of in-person and online motivational interviewing–based health coaching. Health Promot Pract 2018 Jul;19(4):513-521. [doi: 10.1177/1524839917746634] [Medline: 29212349]
34. Gershkowitz BD, Hillert CJ, Crotty BH. Digital coaching strategies to facilitate behavioral change in type 2 diabetes: a systematic review. J Clin Endocrinol Metab 2021 Mar 25;106(4):e1513-e1520. [doi: 10.1210/clinem/dgaa850] [Medline: 33206975]
35. Khalil-Khan A, Khan MA. The impact of COVID-19 on primary care: a scoping review. Cureus 2023 Jan;15(1):e33241. [doi: 10.7759/cureus.33241] [Medline: 36618499]
36. Ndayishimiye C, Lopes H, Middleton J. A systematic scoping review of digital health technologies during COVID-19: a new normal in primary health care delivery. Health Technol (Berl) 2023;13(2):273-284. [doi: 10.1007/s12553-023-00725-7] [Medline: 36628261]
37. Hutchings R. The impact of COVID-19 on the use of digital technology in the NHS. : Nuffield Trust; 2020 URL: https://www.nuffieldtrust.org.uk/sites/default/files/2020-08/the-impact-of-covid-19-on-the-use-of-digital-technology-in-the-nhs-web-2.pdf [accessed 2026-03-12]
38. Madanian S, Nakarada-Kordic I, Reay S, Chetty T. Patients' perspectives on digital health tools. PEC Innov 2023 Dec;2:100171. [doi: 10.1016/j.pecinn.2023.100171] [Medline: 37384154]
39. Kadir KA, Aziz AA, Bakeri NA, Selamat NFS, Pirus MSM. Online coaching as a strategy for health intervention: a scoping review. Asian J Appl Commun 2021;11(1):17-36 [FREE Full text]
40. Doolittle JS. Virtual coaching is inevitable and effective. Presented at: Regent Research Roundtables – Professional Coaching Proceedings; 2022; Virginia Beach, VA, USA p. 24-31 URL: https://www.regent.edu/wp-content/uploads/2022/10/Regent-Research-Roundtables-2022-Professional-Coaching-Doolittle.pdf [accessed 2026-03-12]
41. Kim Y, Lee S. A systematic review and meta-analysis of the effectiveness of non-face-to-face coaching. Sustainability 2023;15(12):9727. [doi: 10.3390/su15129727]

XSL•FO
RenderX

42. Dodds S, Russell–Bennett R, Chen T, Oertzen AS, Salvador-Carulla L, Hung YC. Blended human-technology service realities in healthcare. J Serv Theory Pract 2022 Jan 13;32(1):75-99. [doi: 10.1108/JSTP-12-2020-0285]
43. Liu K, Tao D. The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services. Comput Human Behav 2022 Feb;127:107026. [doi: 10.1016/j.chb.2021.107026]
44. Lopez de Coca T, Moreno L, Alacreu M, Sebastian-Morello M. Bridging the generational digital divide in the healthcare environment. J Pers Med 2022 Jul 26;12(8):1214. [doi: 10.3390/jpm12081214] [Medline: 35893307]
45. Sanders CK, Scanlon E. The digital divide is a human rights issue: advancing social inclusion through social work advocacy. J Hum Rights Soc Work 2021;6(2):130-143. [doi: 10.1007/s41134-020-00147-9] [Medline: 33758780]
46. Pick J, Sarkar A. Theories of the digital divide: critical comparison. Presented at: 2016 49th Hawaii International Conference on System Sciences (HICSS); Jan 5-8, 2016; Koloa, HI, USA p. 3888-3897. [doi: 10.1109/HICSS.2016.484]
47. van Dijk J. Digital divide research, achievements and shortcomings. Poetics (Amst) 2006;34(4-5):221-235. [doi: 10.1016/j.poetic.2006.05.004]
48. Song Z, Johnston RM, Ng CP. Equitable healthcare access during the pandemic: the impact of digital divide and other sociodemographic and systemic factors. Appl Res Artif Intell Cloud Comput 2021;4(1):19-33 [FREE Full text]
49. Saeed SA, Masters RM. Disparities in health care and the digital divide. Curr Psychiatry Rep 2021 Jul 23;23(9):61. [doi: 10.1007/s11920-021-01274-4] [Medline: 34297202]
50. Litchfield I, Shukla D, Greenfield S. Impact of COVID-19 on the digital divide: a rapid review. BMJ Open 2021 Oct 12;11(10):e053440. [doi: 10.1136/bmjopen-2021-053440] [Medline: 34642200]
51. Harvey M, Hastings DP, Chowdhury G. Understanding the costs and challenges of the digital divide through UK council services. J Inf Sci 2023;49(5):1153-1167. [doi: 10.1177/01655515211040664]
52. Stephens M, Mankee-Williams A. The sickening truth of the digital divide: digital health reforms and digital inequality. J Soc Incl 2021;12(2):57-66. [doi: 10.36251/josi227]
53. Thomason J. Metahealth—how will the metaverse change health care? J Metaverse 2021;1(1):13-16 [FREE Full text]
54. Park SY, Nicksic Sigmon C, Boeldt D. A framework for the implementation of digital mental health interventions: the importance of feasibility and acceptability research. Cureus 2022 Sep;14(9):e29329. [doi: 10.7759/cureus.29329] [Medline: 36277565]
55. Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021 Jul 29;11(7):1473-1480. [doi: 10.1093/tbm/ibab048] [Medline: 33963864]
56. O'Brien HL, Morton E, Kampen A, Barnes SJ, Michalak EE. Beyond clicks and downloads: a call for a more comprehensive approach to measuring mobile-health app engagement. BJPsych Open 2020 Aug 11;6(5):e86. [doi: 10.1192/bjo.2020.72] [Medline: 32778200]
57. Nicholas J, Fogarty AS, Boydell K, Christensen H. The reviews are in: a qualitative content analysis of consumer perspectives on apps for bipolar disorder. J Med Internet Res 2017 Apr 7;19(4):e105. [doi: 10.2196/jmir.7273] [Medline: 28389420]
58. Philippi P, Baumeister H, Apolinário-Hagen J, et al. Acceptance towards digital health interventions—model validation and further development of the Unified Theory of Acceptance and Use of Technology. Internet Interv 2021 Dec;26:100459. [doi: 10.1016/j.invent.2021.100459] [Medline: 34603973]
59. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
60. Morrison LG. Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: a review. Digit Health 2015;1:2055207615595335. [doi: 10.1177/2055207615595335] [Medline: 29942544]
61. Park S, Burford S, Nolan C, Hanlen L. The role of digital engagement in the self-management of type 2 diabetes. Health Commun 2016 Dec;31(12):1557-1565. [doi: 10.1080/10410236.2015.1089468] [Medline: 27124817]
62. Ross J, Cotterill S, Bower P, Murray E. Influences on patient uptake of and engagement with the national health service digital diabetes prevention programme: qualitative interview study. J Med Internet Res 2023 Feb 28;25:e40961. [doi: 10.2196/40961] [Medline: 36853751]
63. Baptista S, Wadley G, Bird D, Oldenburg B, Speight J, My Diabetes Coach Research Group. User experiences with a type 2 diabetes coaching app: qualitative study. JMIR Diabetes 2020 Jul 17;5(3):e16692. [doi: 10.2196/16692] [Medline: 32706649]
64. Lavikainen P, Mattila E, Absetz P, et al. Digitally supported lifestyle intervention to prevent type 2 diabetes through healthy habits: secondary analysis of long-term user engagement trajectories in a randomized controlled trial. J Med Internet Res 2022 Feb 24;24(2):e31530. [doi: 10.2196/31530] [Medline: 35200147]
65. Nelson LA, Spieker A, Greevy R, LeStourgeon LM, Wallston KA, Mayberry LS. User engagement among diverse adults in a 12-month text message–delivered diabetes support intervention: results from a randomized controlled trial. JMIR Mhealth Uhealth 2020 Jul 21;8(7):e17534. [doi: 10.2196/17534] [Medline: 32706738]
66. Moher D, Shamseer L, Clarke M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
67. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]

XSL•FO
RenderX

68. Singh J. Critical appraisal skills programme. J Pharmacol Pharmacother 2013 Mar;4(1):76-77. [doi: 10.4103/0976-500X.107697]
69. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
70. Prasad M. Introduction to the GRADE tool for rating certainty in evidence and recommendations. Clin Epidemiol Glob Health 2024;25:101484. [doi: 10.1016/j.cegh.2023.101484]
71. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: Prentice Hall; 1986.
72. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]

Abbreviations

AI: artificial intelligence
CASP: Critical Appraisal Skills Program
FF: face-to-face
GRADE: grading of recommendations, assessment, development, and evaluations
HbA_{1c}: glycated hemoglobin
MeSH: medical subject headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
RC: remote coaching
ROBINS-I: Risk of Bias in Nonrandomized Studies of Interventions
T2D: type 2 diabetes
WHO: World Health Organization

Edited by J Sarvestan; submitted 16.Jan.2025; peer-reviewed by A DiGiammarino; revised version received 31.Dec.2025; accepted 25.Feb.2026; published 31.Mar.2026.

Please cite as:
Mir R, Poduval S, Sheringham J, Hamilton FL
Exploring Components and Feasibility of Health Coaching Interventions for Self-Management of Type 2 Diabetes: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e71383
URL: https://www.researchprotocols.org/2026/1/e71383
doi:10.2196/71383

© Rija Mir, Shoba Poduval, Jessica Sheringham, Fiona Louise Hamilton. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy and Safety of Massage Therapy for Cognitive Impairment in Patients With Autism Spectrum Disorder: Protocol for a Systematic Review and Meta-Analysis

Yi Zhong, BM; Sijie Dang, BM; Xiaoqiu Wang, BM; Xin Yun Chia, BM; Jing Zhang, PhD; Weijiang Li, MMed; Chongjie Yao, PhD; Zhengyu Li, MD, PhD; Bin Xiao, MD, PhD

School of Acupuncture-Moxibustion and Tuina, Shanghai University of Traditional Chinese Medicine, No. 1200 Cailun Road, Shanghai, China

Corresponding Author:
Bin Xiao, MD, PhD
School of Acupuncture-Moxibustion and Tuina, Shanghai University of Traditional Chinese Medicine, No. 1200 Cailun Road, Shanghai, China

Abstract

Background: Autism spectrum disorder (ASD) is a neurodevelopmental condition marked by social communication deficits and cognitive impairment. As pharmacological and behavioral treatments often show inconsistent efficacy, massage therapy—a low-risk complementary approach including *tuina* and acupressure—is used to improve cognitive symptoms, yet its systematic efficacy and safety remain unevaluated.

Objective: This systematic review and meta-analysis aims to evaluate the efficacy and safety of massage therapy (categorized into traditional Chinese medicine–based and Western-based approaches) for improving specific cognitive domains (eg, executive function, attention, and memory) in individuals with ASD.

Methods: Randomized controlled trials (RCTs) and cluster RCTs will be identified across 9 English and Chinese databases (including PubMed, Cochrane Library, and China National Knowledge Infrastructure) through September 2025. Two reviewers will independently perform study selection, data extraction, and risk-of-bias assessment using version 2 of the Cochrane risk-of-bias tool. A random-effects model will be primarily applied due to anticipated clinical diversity, with fixed-effects models reserved for sensitivity analyses of homogeneous data (I^2=0%). Subgroup analyses will explore age, intervention context, and massage type. Evidence for primary outcomes—including checklist scores and specialized cognitive scales—will be assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach.

Results: Selection results will be summarized in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. This review will synthesize evidence on massage therapy's effects on cognitive outcomes measured using scales such as the Autism Behavior Checklist, Childhood Autism Rating Scale, and Autism Treatment Evaluation Checklist, as well as safety profiles. This study was funded in 2024. Following the September 2025 search, analysis will be completed by June 2026, with final results expected for submission in December 2026.

Conclusions: This study will evaluate massage therapy's role in managing cognitive impairment in ASD. By synthesizing available evidence from RCTs, the findings will provide an evidence-based foundation for clinical decision-making and clarify whether manual therapies serve as safe, effective nonpharmacological adjuncts to existing ASD interventions.

Trial Registration: PROSPERO CRD420251038194; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251038194

International Registered Report Identifier (IRRID): PRR1-10.2196/88284

(JMIR Res Protoc 2026;15:e88284) doi:10.2196/88284

KEYWORDS

autism spectrum disorder; massage therapy; cognitive impairment; systematic review; meta-analysis; protocol

Introduction

Background and Significance

Autism spectrum disorder (ASD), commonly known as autism, is a neurodevelopmental disorder characterized by early onset during childhood. Its core clinical features include impairments in social communication, repetitive stereotyped behaviors, and restricted interests [1]. According to the latest 2025 data from the China Disabled Persons' Federation, the total number of individuals with ASD in China ranges from 13 to 14 million, among whom 3 to 5 million are children aged 0 to 14 years, accounting for one-fifth of the global pediatric ASD population. As stated in the 2022 annual report of the Autism and

XSL•FO
RenderX

Developmental Disabilities Monitoring Network released by the US Centers for Disease Control and Prevention on April 15, 2025, the prevalence of ASD among children in the United States reached 3.23% (1 in 31) in 2022, representing a 15% increase compared to 2020 (1 in 36) [2].

ASD typically manifests in early childhood and is associated with significant cognitive dysfunction, behavioral disturbances, and language developmental delays [3]. This situation not only severely impacts patients' normal social interactions and physical and mental well-being but also imposes a substantial caregiving burden and economic strain on families and society [4-9].

The etiological mechanisms of ASDs remain incompletely understood and may be associated with genetic factors and environmental risks [10-12]. Clinically, treatment primarily relies on symptom-targeted individualized regimens, and there are still no specific curative interventions. Current rehabilitative interventions are mainly categorized into pharmacological and nonpharmacological approaches [13,14]. While Western medication and behavioral training are widely used, they are limited by inconsistent efficacy and significant adverse effects [15-19]. Notably, an increasing number of clinical studies have focused on traditional Chinese medicine (TCM) for ASD treatment in recent years, offering a novel direction for ASD rehabilitation [20-23]. Compared to other complementary therapies in TCM, massage therapy has greater acceptance [24].

Recent clinical and mechanistic studies provide compelling evidence on massage therapy in ASD. First, the *tuina* intervention can significantly reduce core symptom scores (Autism Behavior Checklist [ABC] and Childhood Autism Rating Scale [CARS]) and improve overall clinical efficacy [25]. The therapeutic mechanisms appear multifaceted, involving the modulation of key neuroplasticity pathways such as the PI3K/AKT and BDNF-CREB in animal models [26,27], as well as the activation of social brain regions and oxytocin release in humans [28]. Furthermore, manual techniques may enhance the central nervous system environment by promoting fluid clearance [29]. A systematic review found that massage can improve self-regulation difficulties and emotional distress in pediatric patients. Through rhythmic tactile stimulation, massage may exert autonomic nervous system regulation and psychological relaxation effects on children, thereby contributing to symptom relief to a certain extent [30].

Although a considerable number of clinical studies have explored massage therapy for ASD, there is a lack of systematic reviews or meta-analyses synthesizing evidence on its efficacy and safety, hindering the establishment of reliable evidence-based foundations. Previous meta-analyses on massage therapy for ASD have emphasized behavioral symptoms and stress [30]. However, cognitive impairments—which are a major cause of disability in ASD—remain underinvestigated in the context of manual therapies. This study is novel in its exclusive focus on cognitive impairment as the primary outcome, aiming to provide the first systematic synthesis of evidence regarding the effects of massage therapy on specific cognitive functions and systematically consolidating evidence to address the research gap in nonpharmacological cognitive interventions for ASD. Therefore, we aim to address this issue through this systematic review and meta-analysis.

Research Objectives

This study protocol describes a systematic review and meta-analysis of published randomized controlled trials (RCTs) investigating massage interventions for ASD, aiming to objectively evaluate the clinical efficacy of massage therapy in improving cognitive impairment in ASD through data synthesis and collation. This review protocol is registered in the PROSPERO platform with the registration number CRD420251038194 [31].

Methods

This protocol follows the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) guidelines [31].

Eligibility Criteria (Population, Intervention, Comparator, Outcome, and Study Design Framework)

Participants

The review will focus on patients of any age diagnosed with ASD using standardized criteria (eg, the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*; *International Classification of Diseases, 10th Revision*; or their latest revisions, the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision* and *International Classification of Diseases, 11th Revision*). Given that the critical window for cognitive development in ASD is childhood and intervention responses vary significantly across developmental stages (eg, infants vs adolescents vs adults), age will be treated as a primary factor in our analytical framework.

To minimize confounding effects on cognitive assessment, studies involving participants with specific conditions will be excluded. These exclusions comprise individuals with severe neurological comorbidities, such as active epilepsy defined as 2 or more seizures in the previous 3 months, cerebral palsy, or known structural brain abnormalities. We will also exclude participants with severe intellectual disability, characterized by an IQ below 50 as measured using standardized scales such as the Wechsler Intelligence Scales.

Furthermore, individuals with major psychiatric disorders, including schizophrenia or bipolar disorder, and those with significant sensory impairments, such as severe visual or hearing loss that would preclude valid cognitive testing, will be ineligible.

Finally, the review will exclude patients with severe physical illnesses or unstable health conditions, such as uncontrolled cardiovascular or pulmonary disease, liver or kidney failure, malignant tumors, or the active phase of severe autoimmune diseases.

Textbox 1 summarizes the specific inclusion and exclusion criteria for participants designed to ensure a focused assessment of massage therapy's effects on cognitive function.

XSL•FO
RenderX

Textbox 1. Summary of population, intervention, comparator, outcome, and study design–based inclusion and exclusion criteria for evaluating the effects of massage therapy on cognitive function in individuals with autism spectrum disorder (ASD).

Inclusion criteria

- Patients of any age diagnosed with ASD using standardized criteria (eg, *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*); *International Classification of Diseases, 10th Revision*; or their latest revisions, *DSM-5 Text Revision* and *International Classification of Diseases, 11th Revision*)
- Focus on developmental stages (childhood, adolescence, or adulthood) as a primary analytical factor

Exclusion criteria

- Severe neurological comorbidity: active epilepsy (≥2 seizures in the previous 3 months), cerebral palsy, or known structural brain abnormalities
- Severe intellectual disability: defined as an IQ below 50 measured using standardized scales (eg, Wechsler Intelligence Scales)
- Major psychiatric disorders: such as schizophrenia or bipolar disorder
- Significant sensory impairments: severe visual or hearing loss that would preclude valid cognitive testing
- Severe physical illness or unstable health conditions: uncontrolled cardiovascular or pulmonary disease, liver or kidney failure, malignant tumors, or active severe autoimmune disease

The presence of these excluded conditions, such as severe neurological comorbidities and intellectual disability, may independently and profoundly impact cognitive function, thereby obscuring the specific intervention effect of *tuina* and massage therapy.

Intervention

For the experimental group, massage therapy is defined as any systematic manual manipulation of soft tissues, including *tuina*, acupressure, or Swedish massage, performed by a trained professional for therapeutic purposes to provide a comprehensive evidence map of manual therapies for ASD cognitive symptoms. To ensure consistency and clinical relevance, studies must meet *minimum operational criteria* for *tuina* as recommended by the technical specifications for pediatric *tuina* in medical institutions [32].

Specifically, the intervention must be described as "*tuina*" or "pediatric *tuina*" and incorporate fundamental maneuvers defined in the specifications, such as kneading, pressing, pushing, and rubbing performed by a trained practitioner (eg, *tuina* therapist or physician) on specific body regions (eg, head, back, and limbs) or acupoints (eg, Baihui DU20 and Sishencong EX-HN1).

Furthermore, each treatment session should typically last between 15 and 30 minutes to align with standard clinical practice, and studies must clearly report the treatment frequency and total intervention period to facilitate subgroup analysis of the dose-effect relationship.

Other massage therapies (eg, Swedish massage, Thai massage, shiatsu massage, and acupressure) will also be included provided they offer clear descriptions of the techniques applied. These interventions will be categorized into TCM based (eg, *tuina* and acupressure) and Western based (eg, Swedish and myofascial release) for subsequent analysis. Studies lacking a detailed description of the manual procedures will be categorized as insufficiently described *tuina* for the purpose of subgroup analysis.

Massage therapy could be administered as a stand-alone treatment or as an adjunct to conventional care (eg, behavioral therapy and special education). Cointerventions must be reported and will be accounted for in the analysis.

Comparators

The control groups will include any of the following: no treatment, waitlist, placebo or sham interventions (eg, light touch or simulated massage without therapeutic techniques), or usual care or standard of care (eg, routine behavioral therapy and special education).

Studies using other active interventions (eg, pharmacological treatments) as a comparator will also be included provided the experimental group receives massage therapy as an adjunct to the same standard care.

Studies in which massage therapy is administered to the control group will be excluded.

Outcomes

Primary Outcomes

The primary outcomes of this study involve core ASD symptom severity and cognitive impairment measured using standardized scales, with a focus on extracting domain or item scores that most directly reflect cognitive and language functions. For the ABC, the language and cognitive domain scores will be extracted. For the CARS, the analysis will extract the score from the specific item assessing cognitive function. For the Autism Treatment Evaluation Checklist (ATEC), the total score of the "cognitive/language" module will be used. Furthermore, if a study reports only the total score of these scales but not the specific domain scores, the total score will be recorded and included in a separate sensitivity analysis to assess its consistency with the domain-specific findings.

Secondary Outcomes

Secondary outcomes for this review will include direct assessments of specific cognitive domains using specialized tools to improve the targetedness and accuracy of the evaluation. These assessments will encompass executive function measured

XSL•FO
RenderX

using instruments such as the Behavior Rating Inventory of Executive Function; attention measured via standardized continuous performance tests; and memory or general cognitive ability measured through IQ scores, Wechsler Intelligence Scales, or the Leiter International Performance Scale–Revised.

The clinical efficacy rate will also be evaluated, with the decision to perform a meta-analysis on clinical efficacy rates depending primarily on the clinical comparability of the definitions used across the studies. Rather than applying arbitrary numerical cutoffs for efficacy definitions, we will qualitatively evaluate the criteria for response or marked improvement, conducting quantitative synthesis only if these definitions are deemed clinically similar enough to yield a meaningful pooled effect. If substantial clinical variation exists, such as significantly different diagnostic thresholds, we will perform subgroup analyses based on the stringency of the definition or provide a descriptive qualitative synthesis.

Additionally, we will record physiological markers such as serum levels of tissue plasminogen activator and plasminogen activator inhibitor-1 alongside any reported undesirable effects such as discomfort, crying, or skin irritation during or after the massage sessions.

Finally, long-term outcomes, including cognitive scores or quality of life measures assessed at follow-up time points (eg, 1 or 3 months after treatment) will be systematically analyzed.

Study Design

We will include RCTs, including both parallel-group and cluster RCT designs. For cluster RCTs, we will extract data adjusted for clustering if reported. If not, we will approximate the effective sample size using an intracluster correlation coefficient from similar studies or perform sensitivity analyses using a range of plausible intracluster correlation coefficient values. Nonrandomized studies, case reports, case series, reviews, protocols, and studies with incomplete or unavailable outcome data will be excluded.

Information Sources and Search Strategy

Databases

The comprehensive literature search was conducted from database inception to September 30, 2025, across 5 English-language databases and 4 Chinese-language databases. English-language resources included PubMed, Web of Science, Scopus, Cochrane Library, and Embase, whereas Chinese-language resources encompassed China National Knowledge Infrastructure (CNKI), Wanfang Data, VIP Chinese Scientific Journal Database, and the Chinese Biomedical Database.

To identify ongoing or recently completed trials, we systematically searched clinical trial registries, specifically ClinicalTrials.gov and the Chinese Clinical Trial Registry.

The search identified varying numbers of records across these platforms, including 8 results from PubMed, 37 from Web of Science, 24 from Embase, 161 from the Cochrane Library, 49 from ClinicalTrials.gov, 100 from the Chinese Biomedical Database, 51 from CNKI, 78 from Wanfang Data, and 86 from VIP Chinese Scientific Journal Database. A detailed representative search strategy for PubMed is presented in Table 1 for reference.

XSL•FO
RenderX

Table . Detailed search strategy for PubMed (executed in September 2025).

Search number	Query	Results, n
1	"Autism Spectrum Disorder"[Mesh]	49,279
2	((((((((((((((((((((((((Autistic Spectrum Disorder[Title/Abstract]) OR (Autistic Spectrum Disorders[Title/Abstract])) OR (Disorder, Autistic Spectrum[Title/Abstract])) OR (Autism Spectrum Disorders[Title/Abstract])) OR (ASD(Autism Spectrum Disorder[Title/Abstract]))) OR (Autistic Disorder[Title/Abstract])) OR (Disorder, Autistic[Title/Abstract])) OR (Disorders, Autistic[Title/Abstract])) OR (Autism[Title/Abstract])) OR (Autism, Early Infantile[Title/Abstract])) OR (Early Infantile Autism[Title/Abstract])) OR (Infantile Autism, Early[Title/Abstract])) OR (Autism, Infantile[Title/Abstract])) OR (Infantile Autism[Title/Abstract])) OR (Kanner Syndrome[Title/Abstract])) OR (Pervasive Developmental Disorder[Title/Abstract])) OR (PPD(Pervasive Developmental Disorder[Title/Abstract]))) OR (Asperger Syndrome[Title/Abstract])) OR (Childhood Disintegrative Disorder[Title/Abstract])) OR (Atypical Autism[Title/Abstract])) OR (Child Development Disorders, Pervasive[Title/Abstract])) OR (Pervasive Child Development Disorders[Title/Abstract])) OR (Pervasive Development Disorders[Title/Abstract])) OR (Autism Spectrum Disorder[Title/Abstract])	76,574
3	"Massage"[Mesh]	7355
4	((((((((((((((((((((((Massage Therapy[Title/Abstract]) OR (Massage Therapies[Title/Abstract])) OR (Therapies, Massage[Title/Abstract])) OR (Therapy, Massage[Title/Abstract])) OR (Myofascial Release Therapy[Title/Abstract])) OR (Myofascial Treatment[Title/Abstract])) OR (Myofascial Release[Title/Abstract])) OR (Myofascial Release Treatment[Title/Abstract])) OR (Acupressure[Title/Abstract])) OR (tuina[Title/Abstract])) OR (Musculoskeletal Manipulations[Title/Abstract])) OR (Craniosacral Massage[Title/Abstract])) OR (Manipulation Therapy[Title/Abstract])) OR (Manipulation Therapies[Title/Abstract])) OR (Manipulative Therapies[Title/Abstract])) OR (Manipulative Therapy[Title/Abstract])) OR (Manual Therapies[Title/Abstract])) OR (Manual Therapy[Title/Abstract])) OR (Chiropractic Manipulation[Title/Abstract])) OR (Chiropractic Spinal Adjustment[Title/Abstract])) OR (Osteopathic Manipulation[Title/Abstract])) OR (Osteopathic Manipulative Treatment[Title/Abstract])) OR (Massage[Title/Abstract])	22,163
5	((Randomized Controlled Trial[Title/Abstract]) OR (Randomized Controlled Trials[Title/Abstract])) OR (RCT[Title/Abstract])	167,976
6	Search 1 OR 2	82,335
7	Search 3 OR 4	24,629
8	Search 5 AND 6 AND 7	8

XSL•FO
RenderX

Search Terms

The search terms integrated 3 primary concepts, including ASD-related terms, massage-related interventions such as *tuina* and manual therapy, and cognitive-related outcomes, combined with specific filters for RCTs.

Supplementary Search

Furthermore, to minimize publication bias and identify ongoing or recently completed studies, we conducted supplementary searches for gray literature, including conference proceedings from major international and domestic conferences in the fields of autism research and complementary medicine, such as the International Meeting for Autism Research and key Chinese academic conferences on *tuina* or manual therapy. Academic dissertations were searched within the thesis repositories of the aforementioned Chinese and English databases, including CNKI, Wanfang Data, and ProQuest Dissertations and Theses Global. The reference lists of all included studies and relevant systematic reviews were manually scanned for additional eligible studies.

Study Selection and Data Extraction

Study Selection

The study selection process will begin with a pilot-testing phase conducted to ensure consistent interpretation of the eligibility criteria among reviewers and refine the screening criteria prior to formal screening. During this phase, a random sample of 20 citations from the search results will be independently screened by 2 reviewers based on titles and abstracts. The interrater agreement will be calculated using the Cohen κ statistic.

A κ value of 0.75 or higher will be considered indicative of substantial agreement and sufficient for proceeding to formal screening. If the κ value is below 0.75, the reviewers will discuss the discrepancies, receive additional training on the eligibility criteria, and refine the screening guide if necessary. This pilot process will be repeated with a new sample until the satisfactory agreement level is achieved.

Following the successful pilot test, the study selection process will be performed independently by 2 reviewers. They will first screen the titles and abstracts of all retrieved records against the eligibility criteria using the Rayyan (Qatar Computing Research Institute) or EndNote (Clarivate Analytics) software. The full texts of potentially relevant studies will then be retrieved and assessed for final inclusion. Any discrepancies at any stage will be resolved through discussion or, if necessary, adjudication by a third senior reviewer.

Data Extraction

Data extraction will be performed using a standardized, prepiloted data extraction form that will be developed in Microsoft Excel. Two reviewers will independently extract data from the included studies, with any discrepancies resolved through consensus or consultation with a third reviewer. The extracted information will encompass study characteristics such as the first author, publication year, country, sample size, and study design, as well as participant details (age, gender, diagnostic criteria, baseline severity, and comorbidities).

Regarding intervention and comparator details, the reviewers will record the type of massage or *tuina*, specific techniques, acupoints, treatment duration, session frequency, total sessions, and cointerventions (including the systematic recording of the type, frequency, and intensity of all concurrent interventions reported in each study).

Outcome data extraction will involve collecting means, SDs, or change scores from baseline for all primary and secondary outcomes at all reported time points, as well as the number of events for dichotomous outcomes. Furthermore, the specific cognitive assessment tool or specialized scale used for each relevant outcome in each study will be systematically recorded to ensure a nuanced evaluation of cognitive domains.

Finally, safety data concerning the type and incidence of adverse events will be documented, and in instances in which critical data are missing or unclear, the corresponding authors will be contacted via email twice over a 2-week period to request the information.

Risk-of-Bias Assessment

The risk of bias for each included study will be independently assessed by 2 reviewers using version 2 of the Cochrane risk-of-bias tool for randomized trials (RoB 2) [33]. Any disagreements will be resolved through discussion or consulting a third reviewer. The overall risk of bias for each study and for each domain across studies will be rated as “low risk,” “some concerns,” or “high risk.” The assessment will cover 5 domains.

The first domain is bias arising from the randomization process. We will evaluate whether the allocation sequence was random, whether it was concealed until participants were enrolled, and whether any baseline imbalances suggest a problem with the randomization process. A study will be rated as having “some concerns” if important baseline imbalances in age, gender, or ASD severity are present but not adjusted for in the analysis.

The second is bias due to deviations from the intended interventions. This domain assesses the effect of assignment to the intervention, including whether participants, personnel, and trial contexts were sufficiently blinded.

The third is bias due to missing outcome data. We will evaluate whether the data for the cognitive outcomes (eg, ABC and CARS scores) were complete and whether any missing data could have biased the results.

The fourth is bias in measurement of the outcome. This domain focuses on whether the assessment of cognitive domains (eg, memory and attention) was influenced by knowledge of the intervention, particularly by the outcome assessors.

The fifth is bias in the selection of the reported results. We will examine whether the reported results were selected based on the findings from multiple outcome measurements or different analyses, assessing for potential selective reporting.

Separate from the RoB 2 tool, we will systematically examine the source of funding and the authors’ declarations of interest. Studies funded by organizations with potential commercial or ideological interests in the results, such as massage device manufacturers or specific therapy associations, will be

XSL•FO
RenderX

highlighted. A study will be categorized as having significant concerns if there is evidence that the funder influenced the study design, analysis, or reporting.

Data Synthesis and Analysis

Software

Statistical analyses will be performed using RevMan (version 5.4; The Cochrane Collaboration) and Stata (version 17; StataCorp).

Meta-Analysis

For continuous outcomes (eg, ABC and CARS scores), the mean difference or standardized mean difference with 95% CIs will be calculated [33]. We prefer using change-from-baseline scores (with SDs of the change) for continuous outcomes; if unavailable, posttreatment scores will be used, and a sensitivity analysis will be performed to assess the impact of this choice.

For dichotomous outcomes (eg, efficacy rates), the odds ratio with 95% CIs will be used.

The decision to perform a meta-analysis will depend on both clinical and statistical heterogeneity. We will first evaluate the clinical similarity of the included studies. A quantitative synthesis (meta-analysis) will only be conducted if the clinical heterogeneity is deemed acceptable. If the interventions are considered too diverse (eg, vastly different techniques or contexts) or if the I^2 value remains high despite subgroup investigations, a descriptive qualitative synthesis will be provided instead of a pooled effect size. The fixed-effects model will be considered only if there is minimal statistical heterogeneity.

Subgroup Analyses

If substantial heterogeneity (I^2≥50%) is detected, we will perform subgroup analyses based on several prespecified factors to explore the potential sources of variability. These factors will include participant age tiers, specifically early childhood (<6 years), school age (6-11 years), and adolescence or adulthood (≥12 years) [34]. Additionally, we will examine the type of massage therapy by comparing TCM-based vs Western-based manual therapies and categorize outcomes by specific treatment durations. Outcomes will also be analyzed separately according to time point intervals, specifically immediately after treatment, short-term follow-up (≤3 months after treatment), and medium- or long-term follow-up (>3 months after treatment). Finally, the analyses will account for the comparator type, such as usual care vs placebo, and the intervention context, identifying whether massage therapy was administered as a stand-alone treatment or as an add-on therapy to usual care.

Meta-Regression

If significant heterogeneity (I^2≥50%) persists after the aforementioned subgroup analyses, we will conduct univariate random-effects meta-regression to quantitatively explore the sources of heterogeneity. The following categorical variables, derived from the subgroup analyses, will be tested as potential moderators: patient age, categorized as less than 6 years vs 6 years and above; the total treatment course, categorized as less than 8 weeks vs 8 weeks and above; and the type of control intervention, such as usual care vs placebo or sham therapy. Additionally, the proportion of participants receiving a specific concurrent therapy will be evaluated as a moderator to further investigate the robustness of the findings.

Heterogeneity

Statistical heterogeneity will be assessed using the I^2 statistic and the Cochran Q test (with a significance level of P<.10). Given the inherent clinical diversity in massage techniques (eg, TCM based vs Western based), treatment durations, and participant characteristics across the included trials, a random-effects model will be applied as the primary analytical approach for all outcomes. This model is preferred as it accounts for both within-study and between-study variability, providing more conservative and generalizable estimates. The fixed-effects model will only be used as a sensitivity analysis to test the robustness of the results when statistical heterogeneity is absent (I^2=0%) and the studies are deemed sufficiently similar in their population, intervention, comparator, and outcome elements.

Subgroup, sensitivity, and meta-regression analyses will be performed to explore heterogeneity. Following Cochrane recommendations, subgroup analyses and meta-regression will only be performed if at least 10 studies are available for the specific outcome to avoid spurious findings. Results from analyses with fewer than 10 studies will be interpreted with extreme caution, focusing primarily on descriptive presentation.

Sensitivity Analyses

To test the robustness of the primary meta-analysis results, sensitivity analyses will be conducted by sequentially excluding studies based on specific criteria and comparing the pooled effect sizes. Specifically, we will exclude studies judged to have a high risk of overall bias, those with a single-group sample size of fewer than 20 participants, and those with a follow-up loss rate greater than 15%. Additionally, we will examine the robustness of the overall findings by including only studies that use the most frequently reported massage modality, such as *tuina* only or Swedish massage only, depending on the identified evidence base. Finally, studies involving adult populations (aged ≥18 years) will be excluded to assess whether the findings remain robust specifically for the pediatric and adolescent ASD population.

Publication Bias

If 10 or more studies are included in a meta-analysis, publication bias will be assessed visually using funnel plots and statistically using the Egger linear regression test.

If funnel plot asymmetry is suggested and the Egger test indicates statistical significance (P<.05), the nonparametric trim-and-fill method will be used to estimate the number of potentially missing studies and compute an adjusted effect size. The original and adjusted effect estimates will be compared to evaluate the potential impact of publication bias on the results.

Certainty of Evidence

The certainty of the evidence for the primary outcomes—cognitive and language domain scores from the

XSL•FO
RenderX

ABC, CARS, and ATEC scales—will be assessed using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. The evidence will start as "high certainty" for randomized trials and may be rated down by one or more levels based on the following 5 domains.

First, regarding the risk of bias, the certainty will be downgraded if 50% or more of the participants contributing to the outcome are from studies judged to be at a high risk of bias according to the RoB 2 tool.

Second, indirectness will be considered, with the certainty rated down if significant concerns arise regarding the population, intervention, or outcomes. Specifically, the certainty will be rated down for indirectness of the population if most (>50%) of the included evidence is derived from adult ASD populations despite the target population of the review encompassing all ages.

Third, the presence of substantial unexplained heterogeneity, such as an I^2 value of 50% or more, a low *P* value from the chi-square test, or nonoverlapping CIs, will lead to a downgrade for inconsistency.

Fourth, imprecision will be evaluated, where a downgrade will occur if the total sample size is smaller than the optimal information size (typically considered as <400 participants for continuous outcomes) or if the 95% CI around the pooled effect estimate includes both a potentially important benefit and no effect (the null value of zero for mean differences).

Fifth, the certainty will be downgraded if there is significant evidence of publication bias indicated via funnel plot inspection and a *P* value of .05 or lower in the Egger test.

Two reviewers will independently assess the certainty for each key outcome. The final assessments (high, moderate, low, or very low) will be presented in a "summary of findings" table.

Ethical Considerations

As this study is based exclusively on analysis of previously published literature, ethics approval is not required. The final results will be submitted for publication in a peer-reviewed journal and presented at relevant academic conferences. Any modifications to the study protocol will be documented and updated in the PROSPERO registry.

Results

Study Selection

The results of the literature search and the process of study selection (including screening, eligibility assessment, and final inclusion) will be documented and presented in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram, as shown in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram of the study selection process.

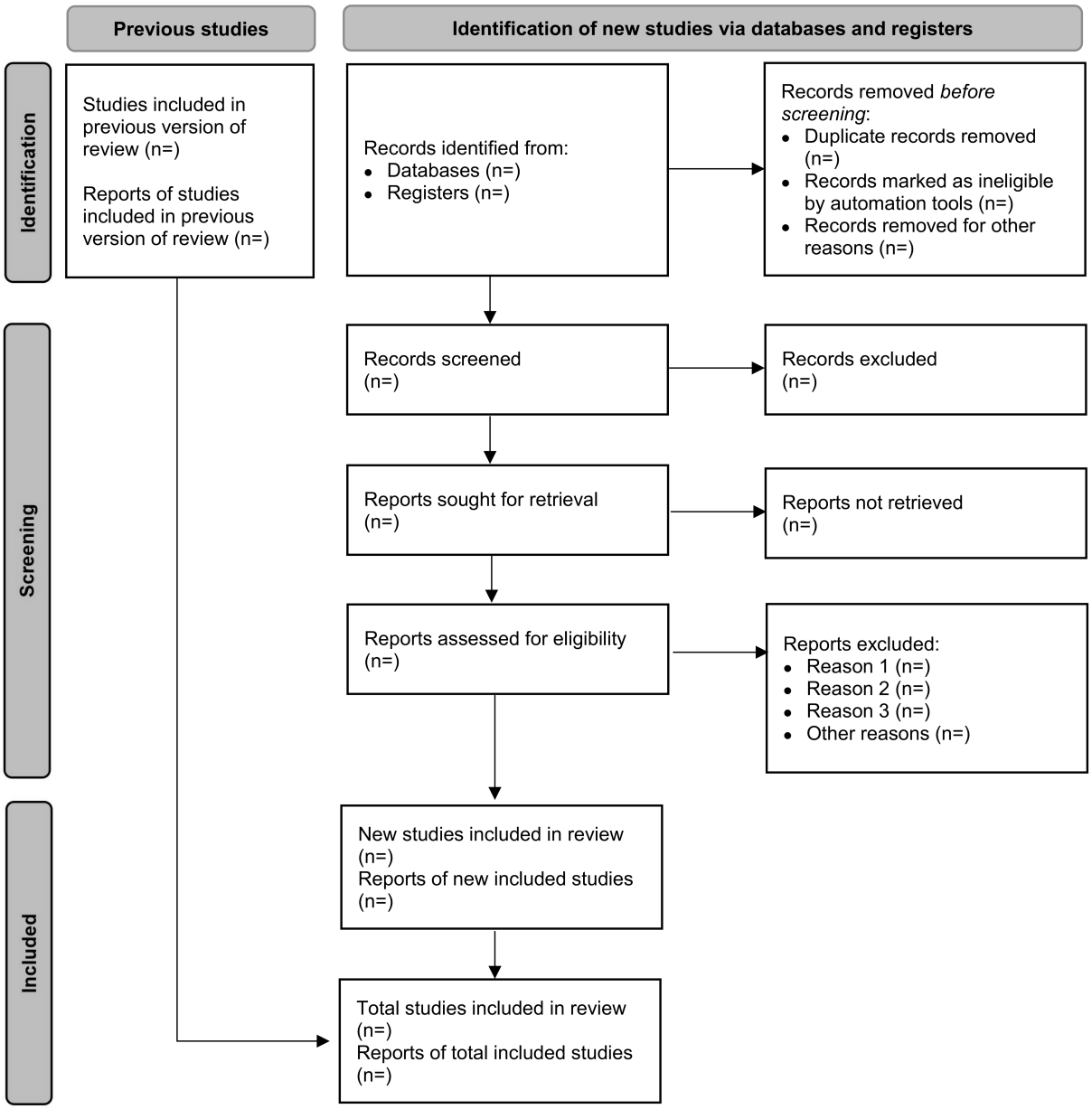

Current Status

This study was funded in 2024. The literature search was initiated in early 2025 and concluded in September 2025. As of September 2025, the comprehensive literature search across eight electronic databases yielded a total of 994 records. Among the English-language resources, the search identified 40 results from PubMed, 104 from Web of Science, 149 from Scopus, and 240 from the Cochrane Library. Regarding the Chinese-language databases, 113 records were retrieved from the CNKI, 159 from Wanfang Data, 97 from the VIP Chinese Scientific Journal Database, and 92 from the Chinese Biomedical Database. These results provide the foundation for the subsequent study selection process as illustrated in the PRISMA flow diagram. Study selection, data extraction, and analysis will be completed by June 2026. The final systematic review is expected to be submitted for publication in December 2026.

Discussion

Expected Findings

This systematic review is expected to provide a comprehensive synthesis of the current evidence regarding massage therapy for cognitive impairment in individuals with ASD. We hypothesize that massage therapy, whether applied as a stand-alone treatment or as an adjunct to conventional care, will demonstrate a positive effect on cognitive and language outcomes as measured using standardized scales such as the ABC, CARS, and ATEC [35,36]. Furthermore, we anticipate that the safety profile will be favorable, with few and minor adverse events reported. By targeting specific cognitive domains, this study aims to fill a

XSL•FO
RenderX

critical gap in the existing literature regarding nonpharmacological cognitive interventions for ASD.

Previous meta-analyses, such as the study by Fadlalmola et al [30], have focused on behavioral symptoms and parenting stress. However, cognitive impairments—which are a major cause of disability and long-term functional challenges in ASD—remain underinvestigated in the context of manual therapies. This review is distinguished by its exclusive focus on cognitive impairment as the primary outcome, aiming to provide the first systematic synthesis of evidence regarding its effects on specific cognitive functions such as memory and executive function.

A major strength of this study is that it will be the first systematic review specifically targeting cognitive outcomes in ASD. The protocol adheres to the PRISMA-P guidelines and includes a rigorous plan for investigating heterogeneity through subgroup analysis and meta-regression. Potential heterogeneity may arise from the diversity of massage techniques and varying degrees of ASD severity [37]. Additionally, the reliance on published data may introduce publication bias, and the availability of long-term follow-up data may be limited.

The findings of this review will identify specific gaps in the current literature, such as the need for more large-scale, high-quality RCTs with standardized protocols. If efficacy is demonstrated, it may encourage the integration of massage therapy into standardized rehabilitation protocols and support clinical decision-making for health care providers and families seeking complementary therapies. Future research should also explore the underlying neurobiological mechanisms through which tactile stimulation influences cognitive development in ASD.

The final results of this systematic review and meta-analysis will be submitted for publication in a high-impact, peer-reviewed journal specializing in pediatrics, integrative medicine, or neurodevelopmental disorders. Additionally, the findings will be presented at relevant international and domestic academic conferences to reach a broad audience of practitioners and researchers.

Strengths and Limitations of This Study

This will be the first systematic review and meta-analysis specifically targeting the efficacy of massage therapy on cognitive impairment in ASD, addressing a clearly defined evidence gap by focusing on specific cognitive domains such as executive function, attention, and memory. The analytical framework is strengthened by a predefined and detailed plan for investigating heterogeneity through subgroup analysis and meta-regression, alongside testing robustness via sensitivity analysis. Despite these strengths, potential limitations include high heterogeneity resulting from varied massage techniques and ASD severity; the risk of publication bias, such as the underreporting of negative findings; and the likely scarcity of long-term data on cognitive outcomes.

Acknowledgments

The authors attest that no generative artificial intelligence tools were used in the writing or revision of this manuscript. All work is original and conducted by the human authors. Patients or the public were not involved in the design, implementation, or reporting of this systematic review.

Funding

This work was supported by the Shanghai Program Innovation and Training Undergraduate on Entrepreneurship (202410268202).

Authors' Contributions

Conceptualization: SD
Data curation: SD, XW
Formal analysis: XYC
Funding acquisition: BX
Investigation: JZ
Methodology: YZ, XYC, WL
Project administration: CY, ZL, BX
Resources: XW
Software: YZ, XYC
Supervision: BX
Validation: SD, XW, XYC, WL
Visualization: SD, CY, ZL
Writing—original draft: YZ
Writing—review and editing: YZ, BX
BX is the guarantor for this study.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Checklist 1
PRISMA-P checklist.
[DOCX File, 30 KB - resprot_v15i1e88284_app1.docx]

References

1. Constantino JN, Charman T. Diagnosis of autism spectrum disorder: reconciling the syndrome, its diverse origins, and variation in expression. Lancet Neurol 2016 Mar;15(3):279-291. [doi: 10.1016/S1474-4422(15)00151-9] [Medline: 26497771]
2. Shaw KA, Williams S, Patrick ME, et al. Prevalence and early identification of autism spectrum disorder among children aged 4 and 8 years - Autism and Developmental Disabilities Monitoring Network, 16 Sites, United States, 2022. MMWR Surveill Summ 2025 Apr 17;74(2):1-22. [doi: 10.15585/mmwr.ss7402a1] [Medline: 40232988]
3. Baranek GT, Watson LR, Boyd BA, Poe MD, David FJ, McGuire L. Hyporesponsiveness to social and nonsocial sensory stimuli in children with autism, children with developmental delays, and typically developing children. Dev Psychopathol 2013 May;25(2):307-320. [doi: 10.1017/S0954579412001071] [Medline: 23627946]
4. Brian AJ, Roncadin C, Duku E, et al. Emerging cognitive profiles in high-risk infants with and without autism spectrum disorder. Res Autism Spectr Disord 2014 Nov;8(11):1557-1566. [doi: 10.1016/j.rasd.2014.07.021]
5. Ozonoff S, Iosif AM, Baguio F, et al. A prospective study of the emergence of early behavioral signs of autism. J Am Acad Child Adolesc Psychiatry 2010 Mar;49(3):256-266. [Medline: 20410715]
6. Soke GN, Maenner MJ, Christensen D, Kurzius-Spencer M, Schieve LA. Prevalence of co-occurring medical and behavioral conditions/symptoms among 4- and 8-year-old children with autism spectrum disorder in selected areas of the United States in 2010. J Autism Dev Disord 2018 Aug;48(8):2663-2676. [doi: 10.1007/s10803-018-3521-1] [Medline: 29524016]
7. Buescher AVS, Cidav Z, Knapp M, Mandell DS. Costs of autism spectrum disorders in the United Kingdom and the United States. JAMA Pediatr 2014 Aug;168(8):721-728. [doi: 10.1001/jamapediatrics.2014.210] [Medline: 24911948]
8. Guerriero V, Fioravanti I, Petrillo M, Di Renzo M, Zavattini GC. Couple satisfaction and parenting stress in parents of children with ASD. Eur Psychiatr 2017 Apr;41(S1):S685-S685. [doi: 10.1016/j.eurpsy.2017.01.1193]
9. Rogge N, Janssen J. The economic costs of autism spectrum disorder: a literature review. J Autism Dev Disord 2019 Jul;49(7):2873-2900. [doi: 10.1007/s10803-019-04014-z] [Medline: 30976961]
10. Lord C, Brugha TS, Charman T, et al. Autism spectrum disorder. Nat Rev Dis Primers 2020 Jan 16;6(1):5. [doi: 10.1038/s41572-019-0138-4] [Medline: 31949163]
11. Bai D, Yip BH, Windham GC, et al. Association of genetic and environmental factors with autism in a 5-country cohort. JAMA Psychiatry 2019 Oct 1;76(10):1035-1043. [doi: 10.1001/jamapsychiatry.2019.1411] [Medline: 31314057]
12. Newschaffer CJ, Fallin D, Lee NL. Heritable and nonheritable risk factors for autism spectrum disorders. Epidemiol Rev 2002;24(2):137-153. [doi: 10.1093/epirev/mxf010] [Medline: 12762089]
13. Sharma SR, Gonda X, Tarazi FI. Autism spectrum disorder: classification, diagnosis and therapy. Pharmacol Ther 2018 Oct;190:91-104. [doi: 10.1016/j.pharmthera.2018.05.007] [Medline: 29763648]
14. Zhang B, Wu H, Zhang C, Wan L, Yang G. Prevalence trends and treatment patterns of autism spectrum disorder among children and adolescents in the United States from 2017 to 2020. Neurol Ther 2024 Dec;13(6):1685-1700. [doi: 10.1007/s40120-024-00665-y] [Medline: 39369113]
15. Maneeton N, Maneeton B, Putthisri S, Woottiluk P, Narkpongphun A, Srisurapanont M. Risperidone for children and adolescents with autism spectrum disorder: a systematic review. Neuropsychiatr Dis Treat 2018;14:1811-1820. [doi: 10.2147/NDT.S151802] [Medline: 30022830]
16. Pandina GJ, Bossie CA, Youssef E, Zhu Y, Dunbar F. Risperidone improves behavioral symptoms in children with autism in a randomized, double-blind, placebo-controlled trial. J Autism Dev Disord 2007 Feb;37(2):367-373. [doi: 10.1007/s10803-006-0234-7] [Medline: 17019624]
17. Owen R, Sikich L, Marcus RN, et al. Aripiprazole in the treatment of irritability in children and adolescents with autistic disorder. Pediatrics 2009 Dec;124(6):1533-1540. [doi: 10.1542/peds.2008-3782] [Medline: 19948625]
18. Sandbank M, Bottema-Beutel K, Crowley S, et al. Project AIM: autism intervention meta-analysis for studies of young children. Psychol Bull 2020 Jan;146(1):1-29. [doi: 10.1037/bul0000215] [Medline: 31763860]
19. White DM, Baeza-Hernandez K, Isotalo C, Gengoux GW. Treatment of autism spectrum disorders. In: Matson JL, editor. Handbook of Clinical Child Psychology: Integrating Theory and Research Into Practice: Springer; 2023:661-686. [doi: 10.1007/978-3-031-24926-6_31]
20. Bang M, Lee SH, Cho SH, et al. Herbal medicine treatment for children with autism spectrum disorder: a systematic review. Evid Based Complement Alternat Med 2017;2017:8614680. [doi: 10.1155/2017/8614680] [Medline: 28592982]
21. Rodrigues JM, Mestre M, Fredes LI. Qigong in the treatment of children with autism spectrum disorder: a systematic review. J Integr Med 2019 Jul;17(4):250-260. [doi: 10.1016/j.joim.2019.04.003] [Medline: 31053555]
22. Li X, Li JC, Lu QQ, Zhang F, Zhang SQ. Research status and prospects of acupuncture for autism spectrum disorders. Front Psychiatry 2023;14:942069. [doi: 10.3389/fpsyt.2023.942069] [Medline: 37304438]
23. Zittoun C, Garbous W, Raffin H, Gusso G. Shiatsu: a complementary medicine for the relief and verbalization of trauma [Article in French]. Encephale 2022 Sep;48 Suppl 1:S56-S60. [doi: 10.1016/j.encep.2022.08.008] [Medline: 36068089]

XSL•FO
RenderX

24. Mursyid F, Chia L, Anwar K, Putri I, Alma B. The impact of tuina massage therapy on sensory processing, sleep disturbances, and maladaptive behaviors in young children with autism spectrum disorder in Jakarta, Indonesia: a controlled clinical study. Scientia Psychiatr 2025;6(4):262-279. [doi: 10.37275/scipsy.v6i2.191]
25. Tang Q, Su LH, Lin LL. Effects of different stimulation doses of chiropractic pinching (niejie) on hippocampal inflammation-related factors and behavioral disorders in autistic rats. J Tradit Chin Med Univ Hunan 2025;45(7):1264-1270. [doi: 10.3969/j.issn.1674-070X.2025.07.009]
26. Feng X, Zhang Y, Tang L, et al. Mechanism of governor vessel pushing manipulation activating the PI3K/AKT signaling pathway to improve behavioral outcomes in rats with autism spectrum disorder. J Beijing Univ Tradit Chin Med 2025;48(6):1. [doi: 10.3969/j.issn.1006-2157.2025.06.020]
27. Sheng ZM, Zhang LN. Mechanisms of external treatment of traditional Chinese medicine in the treatment of childhood autism based on BDNF-CREB related signaling pathways. J Tradit Chin Med 2025;39(8):136-140. [doi: 10.13729/j.issn.1671-7813.Z25022409]
28. Li Q, Becker B, Wernicke J, et al. Foot massage evokes oxytocin release and activation of orbitofrontal cortex and superior temporal sulcus. Psychoneuroendocrinology 2019 Mar;101:193-203. [doi: 10.1016/j.psyneuen.2018.11.016] [Medline: 30469087]
29. Jin H, Yoon JH, Hong SP, et al. Increased CSF drainage by non-invasive manipulation of cervical lymphatics. Nature 2025 Jul;643(8072):755-767. [doi: 10.1038/s41586-025-09052-5] [Medline: 40468071]
30. Fadlalmola HA, Abedelwahed HH, Hamid HI, et al. Effects of massage on self-regulatory difficulties, tactile and oral abnormalities, and parenting stress in children with autism spectrum disorder (ASD): a systematic review and meta-analysis. Int J Ther Massage Bodywork 2024 Dec;17(4):4-19. [doi: 10.3822/ijtmb.v17i4.1079] [Medline: 39669546]
31. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015 Jan 2;350:g7647. [doi: 10.1136/bmj.g7647] [Medline: 25555855]
32. Lan XX, Zhang Y, Wang Y, et al. Technical specifications for pediatric tuina in medical institutions. Int J Tradit Chin Med 2022;44(12):1327-1346. [doi: 10.3760/cma.j.cn115398-20220615-00190]
33. Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions, 2nd edition: John Wiley & Sons; 2019. URL: https://www.cochrane.org/authors/handbooks-and-manuals/handbook [accessed 2026-04-03]
34. Lord C, Charman T, Havdahl A, et al. The Lancet Commission on the future of care and clinical research in autism. Lancet 2022 Jan 15;399(10321):271-334. [doi: 10.1016/S0140-6736(21)01541-5] [Medline: 34883054]
35. Ruan H, Eungpinichpong W, Wu H, Aonsri C. Impact of parent-delivered traditional Thai massage on children with autism spectrum disorder and parenting stress: a randomized controlled trial. Int J Dev Disabil 2024:1-12. [doi: 10.1080/20473869.2024.2404754]
36. Ekradi S, Shahrbanian S, Todd T, Hassani F. Effects of physical activity and massage therapy on sleep disorders among children with autism spectrum disorder: a systematic review. Rev J Autism Dev Disord 2026 Mar;13:185-201. [doi: 10.1007/s40489-024-00433-6]
37. Ruan H, Eungpinichpong W, Wu H, Shen M, Zhang A. Medicine insufficient evidence for the efficacy of massage as intervention for autism spectrum disorder: a systematic review. Evid Based Complement Alternat Med 2022;2022:5328320. [doi: 10.1155/2022/5328320] [Medline: 36193150]

Abbreviations

ABC: Autism Behavior Checklist
ASD: autism spectrum disorder
ATEC: Autism Treatment Evaluation Checklist
CARS: Childhood Autism Rating Scale
CNKI: China National Knowledge Infrastructure
GRADE: Grading of Recommendations Assessment, Development and Evaluation
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
RCT: randomized controlled trial
RoB 2: version 2 of the Cochrane risk-of-bias tool for randomized trials
TCM: traditional Chinese medicine

XSL•FO
RenderX

Edited by J Sarvestan; submitted 22.Nov.2025; peer-reviewed by B Jiao, X Zhang; revised version received 23.Mar.2026; accepted 23.Mar.2026; published 13.Apr.2026.

Please cite as:
Zhong Y, Dang S, Wang X, Chia XY, Zhang J, Li W, Yao C, Li Z, Xiao B
Efficacy and Safety of Massage Therapy for Cognitive Impairment in Patients With Autism Spectrum Disorder: Protocol for a Systematic Review and Meta-Analysis
JMIR Res Protoc 2026;15:e88284
URL: https://www.researchprotocols.org/2026/1/e88284
doi:10.2196/88284

© Yi Zhong, Sijie Dang, Xiaoqiu Wang, Xin Yun Chia, Jing Zhang, Weijiang Li, Chongjie Yao, Zhengyu Li, Bin Xiao. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of Physiotherapy Intervention on Pain, Quality of Life, and Function in Low Back Pain Associated With Piriformis Syndrome: Protocol for Systematic Review

Nikita Deshmukh, PT, MPT

Ravi Nair Physiotherapy College, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Corresponding Author:
Nikita Deshmukh, PT, MPT
Ravi Nair Physiotherapy College, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Abstract

Background: Piriformis syndrome is a neuromuscular condition with hip and buttock pain and other symptoms, including referred pain towards the lower back and leg and radiating towards the foot's medial aspect. Similarly, low back pain caused by piriformis syndrome is undetected or difficult to diagnose because of similar symptoms of lumbar disc herniation, lumbar stenosis, or radiculopathy, as well as neurogenic pain. A study conducted in 2013 found 2910 patients experienced low back pain with sciatica, which is the most common cause of low back pain, because of piriformis muscle stiffness. The prevalence of low back pain in piriformis syndrome is 5% - 36%. It is more commonly seen in women than men.

Objective: This systematic review protocol seeks to identify evidence whether physiotherapy interventions effectively relieve pain, improve functional outcomes, and enhance quality of life among individuals experiencing low back pain associated with piriformis syndrome.

Methods: This review will analyze randomized controlled trials (RCTs) that include physiotherapy for patients with low back pain linked to piriformis syndrome. The included studies must report on pain levels or improvements in function related to quality-of-life outcomes. Searches will take place using Google Scholar, Pubmed/MEDLINE, the Cochrane Library, and PEDro for articles published from January 2014 to January 2025. Two reviewers will individually check the studies, choose relevant ones, and collect data while assessing quality using the Cochrane risk of bias tool. We will provide a narrative summary of the findings, concentrating on data about pain management, functional improvement, and quality of life enhancements.

Results: This review will synthesize knowledge focusing on pain, quality of life and functions in low back pain, which is associated with piriformis syndrome. A synthesis of the findings will be conducted to determine which components of the interventions identified were the most advantageous to the patient population.

Conclusions: The systematic review protocol is designed to identify the effectiveness of physiotherapy interventions for managing low back pain in piriformis syndrome. This analysis will review RCTs with evidence-based recommendations on reducing pain, improving function, and enhancing quality of life.

Trial Registration: PROSPERO CRD42025644124; https://www.crd.york.ac.uk/PROSPERO/view/CRD42025644124

International Registered Report Identifier (IRRID): DERR1-10.2196/72350

(JMIR Res Protoc 2026;15:e72350) doi:10.2196/72350

KEYWORDS

neuromuscular condition; back discomfort; misdiagnosed; affected quality of life; low back pain; piriformis; sciatica; buttock pain; physiotherapy intervention

Introduction

Piriformis syndrome (PS) is a neuromuscular condition characterized by the cardinal symptoms. Sometimes, the syndrome includes hip and buttock pain. The pain radiates towards the back of the leg, sometimes referred down to the medial aspect of the foot. Although this peripheral neuritis is clinically indistinguishable from L5 or S1 radiculopathy, it secondarily leads to abnormal piriformis muscle actions. This abnormal muscle activity results in the compression and irritation of the sciatic nerve as it courses under or through the piriformis muscle [1]. The associated symptoms are challenging to diagnose due to their similar presentation to lumbar disc herniation with radiculopathy, neurogenic pain, and stenosis with radiculopathy. A study showed that in 2013, approximately 2910 patients suffered from low back or buttock pain with sciatica, and 183 (6.25%) patients were diagnosed with PS. If

XSL•FO
RenderX

the diagnosis of PS is delayed, it may lead to pathological complications such as chronic somatic dysfunction and sciatic nerve compression; additionally, compensatory changes may occur, leading to muscle weakness, paraesthesia, hyperesthesia and pain [1,2].

Patients experience discomfort in the back of the thigh, hip area, and lower back, spreading to the leg and foot. Key signs of this condition include numbness, a burning feeling, trouble walking, and difficulty in carrying out everyday tasks. The incidence of this condition ranges from 5% to 36% among individuals with low back pain (LBP). Piriformis syndrome is often misdiagnosed or overlooked in clinical settings due to its symptom overlap with additional ailments, including sacroiliac joint dysfunction, intervertebral disc herniation, and lumbar radiculopathy. Without appropriate treatment, the condition can lead to significant difficulties with prolonged sitting, walking, and standing [3,4]. Piriformis syndrome is a chronic mechanical dysfunction that arises when prolonged sitting places continuous stress on the belly of the piriformis muscle. This leads to muscle spasm and inflammation, resulting in compression of the sciatic nerve. Activities like walking or sitting with one leg crossed can worsen symptoms. Chronic overload, overuse fatigue, radiating limb pain, and severe injuries that may activate skeletal muscle trigger points are additional contributing causes. Sacroiliac joint dysfunction can also result in trigger points in the piriformis muscle, further hindering proper muscle activity. Sacroiliitis, LBP symptoms, or other hip pathologies can also exacerbate the condition [5-7].

Low back pain can often be prevented by encouraging patients to adopt proper sitting postures and avoid prolonged sitting. Various therapeutic approaches are effective in managing different types of LBP. These include manual treatment, electrical stimulation, traction therapy, bed rest, the use of assistive equipment, and heat. Usually, these techniques are regarded as the initial course of treatment for LBP [8]. People who have cerebral palsy, trauma, overexertion, a constricted sciatic foramen, and biomechanical abnormalities may be at risk for PS. These factors can subject the piriformis muscle to excessive stretching or shortening, increasing the risk of developing the condition [9,10]. LBP is among the most prevalent conditions in dancers, accounting for up to 25% of all dance-related injuries. Another study indicated that PS is most commonly observed in individuals in their 40s and 50s. It is more frequent in women than men due to biomechanical differences, such as the wider angle of the quadriceps femoris muscle associated with the female hip structure [11,12].

Tightness in the piriformis limits the range of motion and decreases flexibility, which affects a person's ability to perform both physically and socially. If not identified early, it can progress to PS. Because its symptoms frequently resemble those of disorders including radiculopathy, sacroiliac joint dysfunction, intervertebral disc herniation, or sacroiliitis, misdiagnosis is frequent. However, piriformis tightness is also common in people with low back discomfort. Stiffness in the piriformis muscle is among the most common causes of LBP. Women have piriformis-related illnesses more frequently than men, most likely as a result of anatomical and biomechanical variations. Common signs and symptoms of PS include pain on sitting, muscle tenderness on palpation, muscle spasms, and pain during rectal examination. Piriformis syndrome is also a potential cause of sciatica-related low back and buttock pain. Piriformis syndrome should be considered in the differential diagnosis when pain travels from the back to the leg. Research indicates that most individuals with PS have a history of microtrauma, such as sitting on hard surfaces (36.9%), sitting cross-legged (10.8%), or prolonged walking or running (18.5%) [5,13].

Degenerative disc disease and spondylolysis are the most common causes of LBP in athletes. Diagnosing and managing LBP in athletes can be challenging, as pinpointing the exact source of pain often requires a thorough approach. Physicians must also consider fewer common causes of LBP in this population. While treatment typically targets identifiable causes, over 80% of LBP cases remain unexplained. It is challenging to detect unexplained LBP, which is frequently linked to lower-extremity nerve-related symptoms [14]. Low back pain is a multifactorial condition, requiring a thorough approach when conventional treatments fail to provide relief. A minimum diagnostic workup for athletic LBP should include a comprehensive history to rule out red-flag conditions, alongside specific, evidence-based imaging techniques. Because its definition and mechanism are still up for debate, PS, which is linked to LBP or buttock pain, is still a contentious diagnosis. PS is regarded as a trigger point-based fascial pain condition. Sitting for extended periods usually aggravates symptoms, such as sustained lower back flexion, simultaneous adduction, and internal rotation of the hip joint. These movements, particularly hip adduction and internal rotation in a low posture, are integral to athletic performance and can exacerbate symptoms in affected individuals [6,15]. Which physiotherapy interventions have the most substantial evidence for effectively managing LBP caused by PS? How do these interventions compare in their ability to reduce pain, improve functional limitations, and enhance quality of life for individuals with the LBP associated with PS?

Methods

The eligibility criteria for this systematic review are based on the PICO principle and include the following:

- Exclusion criteria: Individuals who have prolapse intervertebral disc disease; sacro-iliac joint fracture, or posture-related LBP, or any postural impairment associated with LBP.
- Inclusion criteria:Population involves diagnosed cases of PS associated with LBP; dndividuals aged 26 - 50 years, and male and female participants are involved; pain at the intersection of the sciatic nerve and the piriformis muscle; positive modified FAIR test and Lasegue's manoeuvre test positive.
- Intervention: The review includes studies for PS-associated LBP in physical therapy intervention, such as: exercise therapy (ie, strengthening, stretching) [9]; manual therapy (ie, sacro-iliac joint mobilisations, active release technique, neuromuscular inhibition technique, post facilitation stretch technique) [16]; nerve flossing technique (NFT) [10];

XSL•FO
RenderX

postural corrections; appropriate foot support (ie, orthosis); and electrotherapy modalities (ie, shockwave therapy).
- Comparison: Studies must include a comparator group, which could be: positional release technique [1], myofascial release [3], post-isometric relaxation [16], tissue mobilization [16], or conventional physical therapy.
- Outcome measures: The following primary outcomes are used in the study at the beginning (preintervention and after the intervention): pain (using scales such as the Visual Analog Scale (VAS) or Numeric Rating Scale) [17]; functional improvement (Lower Extremity Functional Scale) [18]; and quality of life [19].
- Study design: The focus will be randomized controlled trials (RCTs).
- Language: Studies include only publications published in English because they are easy to understand.
- Publication date: Studies published from January 2014 to January 2025 will be considered.
- Publication type: Published studies like RCTs and relevant content.

Information Sources

The literature search will encompass PubMed/MEDLINE, the Cochrane Library, and PEDro, focusing on English-language articles published between January 2014 and January 2025. Google Scholar will be used as a supplementary search engine.

Search Strategy

A comprehensive literature search will be performed using PubMed/MEDLINE, the Cochrane Library, PEDro, and Google Scholar. The search strategy will include MeSH terms and relevant keywords shown in Textbox 1, such as PS.

The search will be updated before the final analysis to ensure the incorporation of the latest evidence.

Textbox 1. MeSH terms and relevant keywords used in the search strategy.

PS: “piriformis muscle syndrome”[MeSH Terms] OR (“piriformis”[All Fields] AND “muscle”[All Fields] AND “syndrome”[All Fields]) OR “piriformis muscle syndrome”[All Fields] OR (“piriformis”[All Fields] AND “syndrome”[All Fields]) OR “PS”[All Fields]. **PS:** “piriformis muscle syndrome”[MeSH Terms] OR (“piriformis”[All Fields] AND “muscle”[All Fields] AND “syndrome”[All Fields]) OR “piriformis muscle syndrome”[All Fields] OR (“piriformis”[All Fields] AND “syndrome”[All Fields]) OR “PS”[All Fields] **physical therapy**: “physical therapy modalities”[MeSH Terms] OR (“physical”[All Fields] AND “therapy”[All Fields] AND “modalities”[All Fields]) OR “physical therapy modalities”[All Fields] OR (“physical”[All Fields] AND “therapy”[All Fields]) OR “physical therapy”[All Fields]. **PS:** “piriformis muscle syndrome”[MeSH Terms] OR (“piriformis”[All Fields] AND “muscle”[All Fields] AND “syndrome”[All Fields]) OR “piriformis muscle syndrome”[All Fields] OR (“piriformis”[All Fields] AND “syndrome”[All Fields]) OR “PS”[All Fields] **Low back pain:** “low back pain”[MeSH Terms] OR (“low”[All Fields] AND “back”[All Fields] AND “pain”[All Fields]) OR “low back pain”[All Fields].

Selection Process

The reviewers will independently extract data from the included studies using the established extraction form. Information on the study (purpose, sampling strategy, data source, recruitment period, intervention duration, and study duration), participant characteristics, interventions, comparisons, and results is included. The abstracts and titles will be screened for relevancy by two separate reviewers. The eligibility of pertinent full-text research publications will be evaluated next. Discussions or, if required, communication with a third reviewer will be used to resolve any disputes regarding research inclusion.

Data Collection Process

A standardized extraction form, pilot-tested before full implementation, will be used by two independent reviewers to extract data. Essential data, such as research characteristics, participant demographics, intervention and comparator details, outcome measures, and findings, will be gathered via the form. Discussions or, if required, communication with a third reviewer will be used to resolve any disagreements or ambiguities in the data extraction [20].

Data Items

Participant characteristics will include age (26 - 50 y), gender distribution, inclusion and exclusion criteria, and baseline information on health, PS length, physical activity level, and other relevant clinical or demographic data. The sort of physiotherapy intervention employed will be covered in the intervention specifics (eg, exercise therapy (ie, strengthening, stretching); manual therapy (ie, sacroiliac joint mobilizations, active release technique, neuromuscular inhibition technique, post facilitation stretch technique; nerve flossing technique (NFT), postural corrections, appropriate foot support ie, orthosis), electrotherapy modalities (shockwave therapy), the duration, frequency, and intensity of the intervention and the mode of delivery (eg, home-based programs, group sessions, individual sessions). The kind of control or comparator interventions will be included in the comparator group (positional release technique, myofascial release, post isometric relaxation, tissue mobilization, conventional physical therapy), their duration, frequency, and intensity. Outcome measures will assess pain level (using methods likethe VAS or numeric rating Scale), improvement in function (measured with tools like the Lower Extremity Functional Scale), quality of life (36-Item Short Form Health Survey (SF-36)), and any side effects reported. Study details will cover study type (like randomized controlled trial), number of participants, funding sources to check for bias, and any conflicts of interest stated by authors. Planned data assumptions and simplifications will keep data extraction and analysis consistent and accurate across all studies.

Outcomes

Ain intensity will be assessed using continuous scales such as the Visual Analogue Scale (VAS) or the Numeric Pain Rating Scale, with pre- and postintervention scores compared. Functional improvements will be assessed using lower-extremity function scale, which evaluates functional enhancements after treatment by comparing pre- and posttreatment scores. Standardized questionnaires or measures will be used to assess quality of life, such as the SF-36, which focuses on overall

XSL•FO
RenderX

well-being, activities of daily living, and the social and emotional impact of well-being, with pre- and post-treatment assessments conducted (4 wks).

Risk of Bias Assessment

The Cochrane Risk of Bias tool will assess the methodological quality and risk of bias in the included RCTs. The ROBINS-2 tool (Risk of Bias in Non-randomized Studies—of Interventions) will evaluate bias in observational research studies—new CASSTLE.

Data Synthesis

The effectiveness of therapy for LBP linked to PS will be evaluated by calculating standardized mean differences with 95% CIs to measure primary outcomes such as pain intensity, quality of life, and functional improvement. A narrative approach will be the primary method used for data synthesis due to the anticipated variation among studies. A comprehensive summary of the results from the included studies will be provided, emphasizing the main conclusions, different kinds of therapies, and how they affect pain management, functional recovery, and quality of life. Subgroup analyses will be carried out to examine any differences in the effects of the intervention depending on participant demographics or intervention features.

Confidence in Cumulative Evidence

The GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) method will be applied in this systematic review to evaluate the quality of evidence regarding the effectiveness of physiotherapy interventions for LBP associated with PS. The GRADE framework will specifically be applied to the following outcomes: quality of life (measured with the SF-36), functional improvement (rated using the Lower Extremity Functional Scale), and pain severity (measured with the Visual Analog Scale or Numeric Rating Scale), based on comparisons of pre- and post-intervention scores.

Results

The purpose of this research is to determine if rehabilitative therapies are offered for LBP associated PS. The identified studies will be reported in a PRISMA flowchart (Figure 1) [20].

A synthesis of the results will take place to ascertain which elements of the indicated therapies were most beneficial to the patient group. The findings will be used to develop an evidence-based treatment plan to improve quality of life and aid future treatments. The review’s findings will be considered with the panel of experts and decision-makers to create a new evidence-based therapy plan for treating LBP in people with PS .

XSL•FO
RenderX

Figure 1. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) flowchart [20].

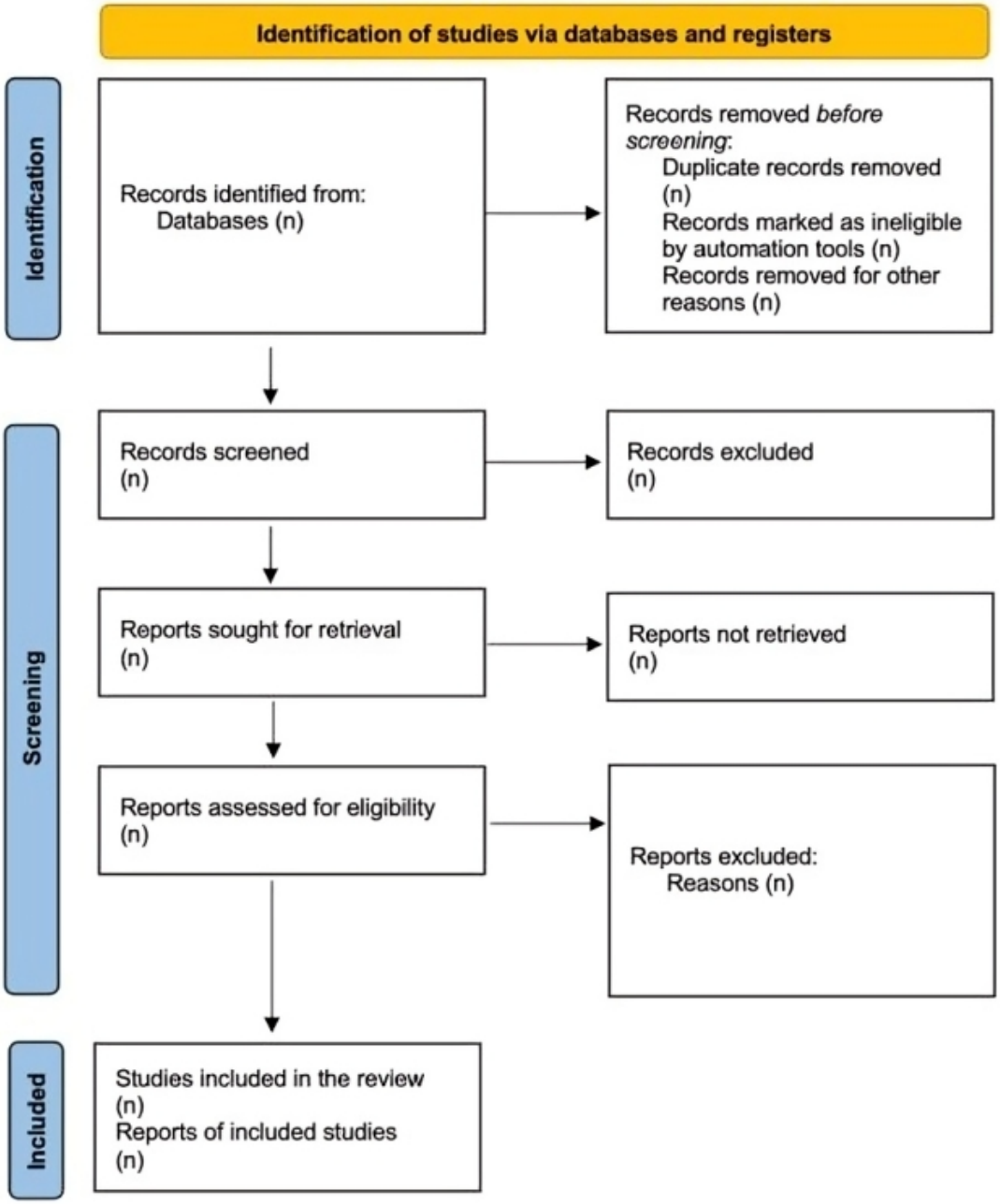

Discussion

Principal Findings

This systematic review protocol outlines a comprehensive approach to evaluating the effectiveness of various physiotherapy interventions for managing LBP associated with PS. PS, with a prevalence of 5% - 36%, significantly impacts quality of life and functional ability due to pain and limitations in daily activities. Underdiagnosis and inadequate treatment are common due to factors such as limited clinical knowledge, difficulty in differentiating PS from other causes of LBP, and a lack of appropriate physiotherapy management. Physical therapy interventions and other holistic approaches and evidence-based practices aim to address underlying factors such as muscular imbalances, biomechanical alterations, and pain to improve patient outcomes and prevent complications. Some evidence suggests promising effects of specific treatments for LBP associated with PS. The results of this systematic review will advance our knowledge and treatment of this illness by offering evidence-based suggestions to direct clinical practice, improve patient outcomes, influence evidence-based guidelines, and pinpoint areas for further study.

Strengths and Limitations

The study addresses LBP associated with PS, a specific condition often overlooked in research, providing valuable insights. The review covers a holistic set of outcomes relevant to patient care by examining pain, quality of life, and function. The systematic review protocol ensures a structured and rigorous methodology, reducing bias and increasing the reliability of findings. The results can guide physiotherapists in developing targeted interventions and improving clinical practice. This

XSL•FO
RenderX

review protocol can consolidate existing evidence on physiotherapy's role in managing PS and address inconsistencies in the current literature. The limitations of physiotherapy interventions can vary widely, making comparisons and synthesis of results challenging. The availability of high-quality RCTs on LBP in PS may be limited, potentially affecting the robustness of the conclusions. Pain and quality of life are subjective and can vary between individuals, possibly introducing variability in the results. Studies with negative or inconclusive results may not be published, skewing the data toward positive outcomes. While focusing on PS is a strength, it may limit generalizability to other causes of LBP.

Conclusion

This systematic review aims to evaluate the effectiveness of physiotherapy techniques for treating LBP associated with PS. The review will offer evidence-based suggestions to alleviate the pain, reduce functional limitations, and improve the quality of life for people with LBP in PS by examining RCTs. In addition to providing opportunities for future study to enhance the quality of life for those with LBP associated with PS, the findings will help inform clinical practice and enrich evidence-based physiotherapy.

Acknowledgments

The author acknowledges the databases and resources required for the data synthesis, as well as the assistance of the Ravi Nair Physiotherapy College and Datta Meghe Institute of Higher Education and Research.

This study is solely the work of the author, and no external reviewers contributed to the development of the manuscript before submission.

The author declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided APF support for the publication of this article.

Funding

The author declared no financial support was received for this work.

Data Availability

Data sharing does not apply to this article as no datasets were generated or analyzed during this study.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist.
[DOCX File, 34 KB - resprot_v15i1e72350_app1.docx]

References

1. Danazumi MS, Yakasai AM, Ibrahim AA, Shehu UT, Ibrahim SU. Effect of integrated neuromuscular inhibition technique compared with positional release technique in the management of piriformis syndrome. J Osteopath Med 2021 May 31;121(8):693-703. [doi: 10.1515/jom-2020-0327] [Medline: 34049428]
2. Rajendran S, Sundaram S. The effectiveness of myofascial release over stretching on pain and range of motion among female college students with piriformis syndrome. Malays J Mov Health Exerc 2020;9(2):45. [doi: 10.4103/2600-9404.323212]
3. Ahmad N, Akram MJ, Hussain H, Imtiaz I, Khan N, Qi L. Effectiveness of post facilitation stretch technique versus myofascial release in piriformis syndrome: a randomised controlled trial. RMJ 2022;47(4):936-939 [FREE Full text]
4. Nazlıkul H, Ural FG, Öztürk GT, Öztürk ADT. Evaluation of neural therapy effect in patients with piriformis syndrome. J Back Musculoskelet Rehabil 2018;31(6):1105-1110. [doi: 10.3233/BMR-170980] [Medline: 30010101]
5. Khan RS, Shafiq Z, Mehar R, et al. A cross-sectional study concerning piriformis tightness: identifying predisposing factors and its association with low back pain in office workers. IHR 2023;3(2):19-31. [doi: 10.32350/ihr.32.02]
6. Othman IK, Raj NB, Siew Kuan C, et al. Association of piriformis thickness, hip muscle strength, and low back pain patients with and without piriformis syndrome in Malaysia. Life (Basel) 2023 May 18;13(5):1208. [doi: 10.3390/life13051208] [Medline: 37240853]
7. Shahzadi B, Taj S, Nawaz S, et al. Sacroiliac joint manipulation helps to improve pain pressure threshold in chronic piriformis syndrome: a 6-week randomized controlled trial US. T Rehabili J 2023 Mar 31;07(1):476-482. [doi: 10.52567/trj.v7i01.202]

XSL•FO
RenderX

8. Elgendy MH, Mohamed M, Hussein HM. A single-blind randomized controlled trial investigating changes in electrical muscle activity, pain, and function after shockwave therapy in chronic non-specific low back pain: pilot study. Ortop Traumatol Rehabil 2022 Apr 30;24(2):87-94. [doi: 10.5604/01.3001.0015.8266] [Medline: 35550356]
9. Alarab A, Unver F. Stretching exercise versus tissue mobilization technique in piriformis syndrome. EJMED 2021;2(6). [doi: 10.24018/ejmed.2020.2.6.610]
10. Saluja J, Parmar R, Patel G. A Comparative study between elongation longitudinaux avec decoaption osteo articulaire (ELDOA) and nerve flossing technique (NFT) to improve pain and flexibility in piriformis syndrome – an experimental study. Archives of Physiotherapy and Global Researches 2023 Nov 19;26(2):31-36. [doi: 10.5604/01.3001.0054.3960]
11. Islam F, Mansha H, Gulzar K, Raza A, Raffique A, Haider S. Prevalence of piriformis muscle syndrome among individuals with low back pain: piriformis muscle syndrome among individuals with low back pain. Pak J Health Sci 2022:48-52. [doi: 10.54393/pjhs.v3i04.98]
12. Rehman AU, Khan LG, Khalid M, et al. Comparison of active release technique and post isometric relaxation in patients with piriformis syndrome. PJMHS 2022;16(5):683-685. [doi: 10.53350/pjmhs22165683]
13. Nadeem N, Yaseen A, Shaheen A, et al. Prevalence of piriformis tightness in pregnant women. IJHR 2024;2(2 (Health& Allied):280-285. [doi: 10.71000/ijhr120]
14. Hashimoto H, Shibahara M. Measurement of piriformis muscle stiffness in athletes with lower back pain and piriformis syndrome: a shear-wave elastography study with consideration of the onset of piriformis syndrome. J Phys Educ Sport 2022;22(7):1748-1760. [doi: 10.7752/jpes.2022.07219]
15. Hashimoto H, Shibahara M, Osaki M. Investigation of the effect of passive and active exercise therapy in athletes suspected of low back pain with piriformis muscle contracture. J Phys Educ Sport 2023;23(8):2060-2072. [doi: 10.7752/jpes.2023.08237]
16. Bhosale P, Kolke(PT) S. Effectiveness of instrument assisted soft tissue mobilization (IASTM) and muscle energy technique (MET) on post-operative elbow stiffness: a randomized clinical trial. Journal of Manual & Manipulative Therapy 2023 Sep 3;31(5):340-348. [doi: 10.1080/10669817.2022.2122372]
17. Koo M, Yang SW. Visual analogue scale (VAS) for pain. Encyclopedia 2025;5(4):190. [doi: 10.3390/encyclopedia5040190]
18. Lower extremity functional scale (LEFS). Emory Healthcare. URL: https://www.emoryhealthcare.org/-/media/Project/EH/Emory/ui/pdfs/msk-pt-forms/hip-lefs.pdf [accessed 2025-01-27]
19. Ruževičius J. Quality of life and working life: conceptions and research. In: 17th Toulon-Verona International Conference Excellence in Services Conference Proceedings: Liverpool John Moores University; 2014.
20. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]

Abbreviations

GRADE: Grading of Recommendations, Assessment, Development, and Evaluation
LBP: low back pain
PS: piriformis syndrome
RCT: randomized controlled trial
SF-36: 36-Item Short Form Health Survey

Edited by J Sarvestan; submitted 08.Feb.2025; peer-reviewed by B Thompson; revised version received 30.Nov.2025; accepted 02.Dec.2025; published 24.Mar.2026.

Please cite as:
Deshmukh N
Impact of Physiotherapy Intervention on Pain, Quality of Life, and Function in Low Back Pain Associated With Piriformis Syndrome: Protocol for Systematic Review
JMIR Res Protoc 2026;15:e72350
URL: https://www.researchprotocols.org/2026/1/e72350
doi:10.2196/72350

© Nikita Deshmukh. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Role of Metabolomics in Understanding Hypertension: Protocol for a Systematic Review and Meta-Analysis

Mohd Naeem Mohd Nawi, MSc, PhD; Puteri Sofia Nadira Megat Kamaruddin, MBBS, MPH, DrPH

Nutrition, Metabolism and Cardiovascular Research Centre, Institute for Medical Research, National Institutes of Health, Ministry of Health Malaysia, No.1, Jalan Setia Murni U13/52, Seksyen U13, Setia Alam, Shah Alam, Malaysia

Corresponding Author:
Mohd Naeem Mohd Nawi, MSc, PhD
Nutrition, Metabolism and Cardiovascular Research Centre, Institute for Medical Research, National Institutes of Health, Ministry of Health Malaysia, No.1, Jalan Setia Murni U13/52, Seksyen U13, Setia Alam, Shah Alam, Malaysia

Abstract

Background: Hypertension is a chronic condition and a leading risk factor for cardiovascular disease, stroke, and premature mortality worldwide. While blood pressure (BP) monitoring—via clinical, home, or ambulatory measurements—remains the primary diagnostic tool, each method is limited by variability, device inaccuracy, and difficulties in detecting atypical BP patterns such as masked or white-coat hypertension. These challenges underscore the need for innovative, complementary diagnostic approaches.

Objective: This systematic review and meta-analysis aims to synthesize current evidence on the role of metabolomic profiling in the detection and understanding of hypertension. Specifically, it seeks to evaluate whether metabolomics can identify preclinical metabolic signatures, improve diagnostic accuracy, and elucidate pathophysiological mechanisms underlying hypertension, thereby complementing traditional BP monitoring.

Methods: A systematic search will be conducted in PubMed, Embase, Web of Science, Scopus, and CENTRAL from inception to the present supplemented with manual screening of metabolomics repositories (MetaboLights and Metabolomics Workbench) to ensure comprehensive coverage. This systematic review will include studies involving adults (aged ≥18 years) that investigate metabolomic biomarkers in hypertension using validated analytical platforms such as nuclear magnetic resonance spectroscopy, gas chromatography–mass spectrometry, or liquid chromatography–mass spectrometry. Eligible studies must stratify participants into hypertensive, prehypertensive, or normotensive groups and report associations between metabolites and BP measurements. Two independent reviewers will conduct study selection, data extraction, and risk-of-bias assessment using version 2 of the Cochrane risk-of-bias tool for randomized trials and the Risk of Bias in Nonrandomized Studies of Interventions and Risk of Bias in Nonrandomized Studies of Exposures tools, with evidence certainty evaluated via the Grading of Recommendations Assessment, Development, and Evaluation framework. A meta-analysis will be performed where possible using random-effects models and subgroup analyses to address heterogeneity.

Results: Database search, screening, and data extraction are ongoing as of January 2026. The results will summarize metabolomic profiles (such as stearidonate and hexadecadienoate), diagnostic accuracy metrics (eg, area under the curve and sensitivity), and metabolic pathways associated with hypertension once data synthesis is completed. The systematic review and meta-analysis is expected to be published in August 2026.

Conclusions: This review will identify metabolites showing consistent hypertension associations and document methodological requirements for clinical translation, including standardization of sample collection, analytical platforms, and data processing. The findings will inform longitudinal validation studies and establish evidence-based pathways for potential applications in hypertension risk assessment and clinical trial design. Validation of consistently identified metabolite-hypertension associations in independent, prospective cohorts will be critical for establishing causality and assessing the temporal stability of metabolomic signatures. The review will also identify methodological gaps requiring standardization, such as sample collection protocols, analytical platform harmonization, and data processing procedures, which must be addressed before metabolomic findings can be translated into clinical applications.

Trial Registration: PROSPERO CRD420251016289; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251016289

International Registered Report Identifier (IRRID): DERR1-10.2196/77536

(JMIR Res Protoc 2026;15:e77536) doi:10.2196/77536

KEYWORDS

hypertension; metabolomics; biomarkers; profiling; diagnosis

XSL•FO
RenderX

Introduction

High blood pressure (BP), or hypertension, is defined by sustained elevations in arterial pressure and is recognized as a leading risk factor for cardiovascular disease, stroke, and premature mortality worldwide [1]. BP monitoring remains the primary method for detecting hypertension using office-based (measurements taken in a clinical setting during a scheduled visit), home, and ambulatory measurements [1]. However, each approach presents notable challenges. Office BP measurement, while widely used, is susceptible to variability due to factors such as improper technique, patient anxiety, and time constraints, which can lead to inaccurate readings and misclassification of hypertension status [2,3]. Additionally, office measurements often fail to detect masked hypertension, in which BP is normal in the office but increased elsewhere, and white-coat hypertension, in which BP is elevated in the clinical setting but normal elsewhere, both of which are prevalent and clinically significant [2,4,5]. Home BP monitoring can help identify these phenomena but depends heavily on patient adherence to proper measurement protocols and the accuracy of devices, many of which are not validated [6,7]. Ambulatory BP monitoring is widely regarded as the most reliable method for diagnosing hypertension, especially for identifying cases of white-coat and masked hypertension, as well as for assessing nocturnal BP patterns and BP variability, which are strong predictors of cardiovascular risk [5,8]. Nevertheless, ambulatory BP monitoring is less accessible, can be poorly tolerated, and is subjected to measurement artifacts that affect reproducibility [3,5]. In summary, while BP monitoring is essential for hypertension detection, its inherent limitations, including measurement variability, device inaccuracy, and challenges in identifying atypical BP patterns, highlight the need for continued innovation and complementary diagnostic strategies to improve detection and management.

Metabolomics, the study of small-molecule metabolites (<1500 Da) in biological systems, provides dynamic insights into physiological states by analyzing pathways such as lipid metabolism and amino acid turnover [9]. Unlike genomic or proteomic approaches, it captures real-time metabolic changes influenced by genetics, environment, and lifestyle, making it pivotal for understanding noncommunicable diseases (NCDs) [10,11]. NCDs, responsible for approximately 74% of global mortality, are linked by shared metabolic disruptions identified through metabolomics, such as altered glycerophospholipid metabolism and mitochondrial dysfunction markers, which underlie conditions such as diabetes and obesity [11,12]. For instance, branched-chain amino acids and ceramides are elevated across metabolic and cardiovascular disorders, serving as cross-disease biomarkers linked to insulin resistance and inflammation [11,13]. Metabolomics additionally uncovers metabolites produced from the gut microbiota, including trimethylamine N-oxide, that contribute to systemic inflammation and renal or hepatic complications in NCDs [14]. Clinically, metabolomic profiling enables early risk prediction—detecting metabolic signatures years before diagnosis—and guides personalized interventions such as micronutrient supplementation or mitochondrial-targeted therapies, improving outcomes such as quality-adjusted life years [10,13]. By mapping these pathways, metabolomics complements traditional diagnostics, offering a proactive approach to managing NCD multimorbidity.

While metabolomics holds considerable promise for advancing hypertension research, existing literature reveals critical gaps that underscore the need for rigorous synthesis of available evidence. The most recent comprehensive narrative review examined studies published between 2015 and 2025, identifying key disrupted metabolic pathways, including amino acid metabolism (particularly branched-chain and aromatic amino acids), fatty acid metabolism (ceramides, oleic acid, and short-chain fatty acids), and pathways related to inflammation and oxidative stress [15,16]. However, this review did not use systematic methodology or conduct meta-analyses to quantify the magnitude of metabolite alterations or evaluate their diagnostic performance characteristics such as sensitivity, specificity, or area under the curve (AUC).

This systematic review and meta-analysis aims to synthesize evidence on the utility of metabolomic profiling in detecting hypertension and complementing BP monitoring by identifying preclinical metabolic signatures, improving diagnostic accuracy, and elucidating pathophysiological mechanisms. This systematic review will evaluate metabolomics' capacity to address limitations of traditional BP measurement, such as episodic variability, white-coat hypertension, and delayed diagnosis, through objective biomarker discovery.

Methods

Eligibility Criteria

Overview

This review focuses on studies that investigate metabolomic biomarkers specifically in relation to hypertension detection and metabolic mechanisms underlying BP regulation, including any linkage such as correlations or diagnostic accuracy metrics (AUC and sensitivity) and excluding broad exploratory or non–hypertension-specific metabolomic analyses. This systematic review will include studies that investigate human participants aged 18 years or older regardless of sex or ethnicity who have been diagnosed with hypertension (systolic BP [SBP]≥130 mm Hg or diastolic BP≥80 mm Hg), prehypertension (SBP of 120 - 129 mm Hg), or normal BP levels. Eligible studies must use metabolomic profiling techniques such as nuclear magnetic resonance (NMR) spectroscopy, gas chromatography–mass spectrometry (GC-MS), or liquid chromatography–mass spectrometry (LC-MS) to identify metabolites or metabolic pathways associated with hypertension. Only original research articles with clearly defined study designs (eg, case-control, cohort, or cross-sectional) will be considered. Exclusion criteria include animal studies, nonoriginal research (reviews and editorials), studies lacking primary metabolomic or BP data, populations with comorbidities (eg, diabetes or chronic kidney disease), pregnancy, or use of confounding medications.

Studies will be selected based on predefined inclusion and exclusion criteria to ensure methodological consistency and

XSL•FO
RenderX

relevance to the research objectives. The criteria were developed using the population, intervention or exposure, comparator, and outcomes framework and aligned with the overall aim of synthesizing metabolomic evidence in hypertension. To enhance readability, the inclusion and exclusion criteria are summarized in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Human studies involving adults aged ≥18 years
- Studies investigating metabolomic biomarkers in hypertension
- Use of validated analytical platforms (eg, nuclear magnetic resonance spectroscopy, gas chromatography–mass spectrometry, or liquid chromatography–mass spectrometry)
- Stratification of participants into hypertensive, prehypertensive, or normotensive groups
- Reporting of associations between metabolites and blood pressure or diagnostic metrics (eg, area under the curve, sensitivity, and specificity)

Exclusion criteria

- Animal or in vitro studies
- Reviews, editorials, conference abstracts without full data, or nonoriginal research
- Studies lacking primary metabolomic or blood pressure data
- Populations with major comorbidities (eg, diabetes or chronic kidney disease), pregnancy, or use of confounding medications
- Studies not meeting the aforementioned inclusion criteria

Population

The target population includes adults aged 18 years and older stratified into hypertensive, prehypertensive, and normotensive groups. Studies involving secondary hypertension, acute illness, or metabolic disorders will be excluded to minimize confounding effects on metabolomic profiles that may arise from conditions such as diabetes or chronic kidney disease. Where eligible studies include mixed populations, data pertaining to primary hypertensive subgroups or sensitivity analyses will be extracted and analyzed separately where possible to maintain relevance and improve generalizability.

Intervention

The intervention or exposure involves conducting metabolomic profiling through techniques such as NMR spectroscopy, GC-MS, or LC-MS, with the aim of identifying biomarkers that are associated with hypertension.

Comparison

Comparisons will be made between individuals with hypertension and normal BP, prehypertensive and normotensive groups, and individuals before or after antihypertensive treatment.

Outcome Measures

Primary outcomes include diagnostic accuracy of metabolomic biomarkers (eg, AUC and sensitivity) and identification of key metabolites linked to hypertension (eg, stearidonate and hexadecadienoate). Secondary outcomes include pathway dysregulation (eg, lipid peroxidation and mitochondrial dysfunction) and correlations between metabolite levels and BP changes.

Data Sources, Search Terms, and Search Strategy

This systematic review will use 5 electronic databases—PubMed, Embase, Web of Science, Scopus, and CENTRAL—from inception to the present as primary data sources. These databases were selected to ensure broad coverage of biomedical and omics-related literature. In addition, specialized metabolomics repositories such as MetaboLights and Metabolomics Workbench will be manually searched to identify relevant studies and supplementary datasets not indexed in the aforementioned databases. Both authors collaborated to determine relevant search terms, including keywords and MeSH (Medical Subject Headings) terms, and search queries were constructed using Boolean operators (“AND” and “OR”). Searches will be conducted from the inception of each database to the most recent available studies without applying search filters. No language restrictions will be applied during the search. In cases in which potentially relevant articles are published in languages other than English, professional and credible translation services will be used to translate the texts into English for inclusion and assessment. Unpublished studies identified at the time of the final search will be excluded. The detailed search strategy is shown in Multimedia Appendix 1. All retrieved references will be downloaded in RIS format and managed using EndNote (version 19; Clarivate Analytics), which will be used to combine records from all databases and remove duplicates.

Study Screening and Selection

The screening of eligible studies will be undertaken after the database searches are complete. All identified references will be managed using Microsoft Excel, where duplicates will be identified and removed. To ensure consistency and familiarity with the screening process, at least 5 randomly selected studies will be pilot-tested by the reviewers. The evaluation will occur in 2 stages. Initially, the 2 reviewers (MNMN and PSNMK)

XSL•FO
RenderX

will independently evaluate the titles and abstracts of all references based on the established eligibility criteria. During the second stage, the same 2 reviewers will evaluate the complete texts of possibly suitable articles for inclusion. Disagreements or differences between the reviewers will be resolved through discussion to achieve consensus. This methodology aims to reduce bias and guarantee a transparent and methodical selection of studies. We will document the selection process in adequate detail to finalize a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [16].

Data Extraction

Data extraction will use a standardized Microsoft Excel template to record analytical platforms (NMR spectroscopy, GC-MS, and LC-MS); quantified metabolites (such as amino acids and lipids); associated metabolic pathways; and statistical associations with BP parameters, including SBP or diastolic BP values and diagnostic accuracy metrics. The template will systematically capture study characteristics, including design, population demographics, interventions, comparators, and outcomes related to metabolite-BP linkages. The 2 reviewers (MNMN and PSNMK) will independently extract data from the selected papers, and any disagreements will be resolved through consensus. If further clarification or information is required, the corresponding authors of the original studies will be contacted.

Risk-of-Bias and Quality Assessment

For randomized controlled trials, the risk-of-bias assessment will use version 2 of the Cochrane risk-of-bias tool for randomized trials; for nonrandomized studies, the Risk of Bias in Nonrandomized Studies of Interventions tool will be applied. For observational studies examining metabolomic exposures and hypertension outcomes, which represent most metabolomics research designs, we will apply the Risk of Bias in Nonrandomized Studies of Exposures (ROBINS-E) tool [17]. The ROBINS-E tool is specifically designed for cohort and case-control studies evaluating biological exposures such as metabolite concentrations, assessing 7 bias domains: confounding, participant selection, exposure classification, deviations from intended exposures, missing data, outcome measurement, and selective reporting. To address metabolomics-specific methodological challenges within the ROBINS-E framework, we will systematically evaluate (1) preanalytical factors, including sample collection protocols, storage conditions, freeze-thaw cycles, and metabolite stability considerations; (2) analytical quality control procedures, including use of pooled quality control samples, internal standards, coefficient of variation reporting, and batch effect correction methods; (3) metabolite identification confidence according to Metabolomics Standards Initiative criteria (levels 1-4); and (4) data processing rigor, including normalization, missing value handling, and multiple testing correction. The 2 reviewers (MNMN and PSNMK) will independently assess each study, and any differences in their assessments will be resolved through discussion. If essential information is unclear or unavailable in the published reports, additional details will be sought by contacting the study investigators. This rigorous approach ensures a transparent and systematic evaluation of study quality and potential sources of bias. The certainty of evidence for each outcome will be assessed using the Grading of Recommendations Assessment, Development, and Evaluation approach, and a summary table will be presented, including justifications for each assessment.

Data Synthesis

A meta-analysis will be conducted if 2 or more studies provide identical outcomes; otherwise, the results will be described narratively. For metabolomics studies, “identical outcomes” will be defined as the same metabolite confirmed through International Chemical Identifier Key [18] matching, ensuring consistent identification across studies regardless of naming conventions, with cross-referencing through PubChem [19], the Human Metabolome Database [20], the Kyoto Encyclopedia of Genes and Genomes [21], and the Reference Set of Metabolite Names [22] to resolve nomenclature discrepancies. Data analysis will use The Cochrane Collaboration’s Review Manager software (version 5.4) for aggregating results. Continuous outcomes will be computed as mean differences or standardized mean differences with 95% CIs, depending on the consistency of the measuring instruments used across studies. Dichotomous outcomes will be represented as odds ratios accompanied by 95% CIs. The I^2 statistic will assess heterogeneity, with values of 0% to 40% indicating low heterogeneity, 30% to 60% indicating moderate heterogeneity, 50% to 90% indicating substantial heterogeneity, and 75% to 100% indicating considerable heterogeneity. Analyses will use random-effects models to account for heterogeneity, particularly when pooling across different metabolomics methodologies.

Given the inherent variability in metabolomics platforms (NMR spectroscopy vs LC-MS vs GC-MS) and metabolite panels, meta-analyses will be stratified by analytical platform type, with platform-specific pooled estimates calculated separately before assessing cross-platform comparability using subgroup heterogeneity tests (Cochran Q statistic; interaction $P<.10$). When substantial between-platform heterogeneity exists ($I^2>75\%$ or significant platform interaction), results will be presented separately by platform rather than pooled across methodologies. Studies will be documented for their analytical platform specifications (instrument type, column chemistry for LC-MS, and field strength for NMR spectroscopy), and these characteristics will inform quality weighting in meta-analyses. Sensitivity analyses will be performed by (1) excluding each analytical platform sequentially to assess platform-specific influence on pooled estimates; (2) comparing targeted vs untargeted metabolomics approaches; (3) restricting analyses to studies using validated, standardized protocols; (4) excluding studies with high risk of bias; and (5) restricting analyses to studies with Metabolomics Standards Initiative level 1 and 2 metabolite identifications. Meta-regression will be performed when at least 10 studies are available, with analytical platform type, sample matrix (serum vs plasma vs urine), and study quality as moderator variables to quantify their contribution to heterogeneity. Predefined subgroups will include age, sex, geographic region, analytical platform (NMR spectroscopy vs LC-MS vs GC-MS), sample type (serum vs plasma vs urine),

XSL•FO
RenderX

metabolite identification confidence level (levels 1-2 vs 3-4), study design (cross-sectional vs prospective cohort), and risk-of-bias categories. Assessment of publication bias will be conducted using funnel plot symmetry and the Egger regression approach when a significant number of papers (generally ≥10) are present. For outcomes characterized by inadequate data, an I^2 value of more than 75%, significant platform-related heterogeneity, or fewer than 2 studies reporting identical metabolites, or when cross-platform differences preclude reliable pooling, findings will be synthesized narratively with results stratified by analytical platform and metabolite class. This comprehensive approach ensures that metabolomics-specific methodological variability is systematically addressed through predefined strategies, enhancing the reliability and interpretability of pooled estimates.

Registration and Reporting

This protocol has been registered with PROSPERO (CRD420251016289). Any protocol amendments will be documented on the PROSPERO record, including the date, description, and rationale for each change. The review will adhere to the PRISMA 2020 statement [23] and PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) [24] guidelines and will be reported according to the *Cochrane Handbook for Systematic Reviews of Interventions* [25]. The PRISMA 2020 checklist has been provided (Checklist 1).

Results

The database search, screening, and data extraction and synthesis are ongoing as of January 2026. The results of this systematic review and meta-analysis will summarize diagnostic accuracy metrics, key metabolites, and metabolic pathways associated with hypertension once the analysis is completed. The systematic review and meta-analysis is expected to be published in August 2026. Figure 1 shows the PRISMA flow diagram.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

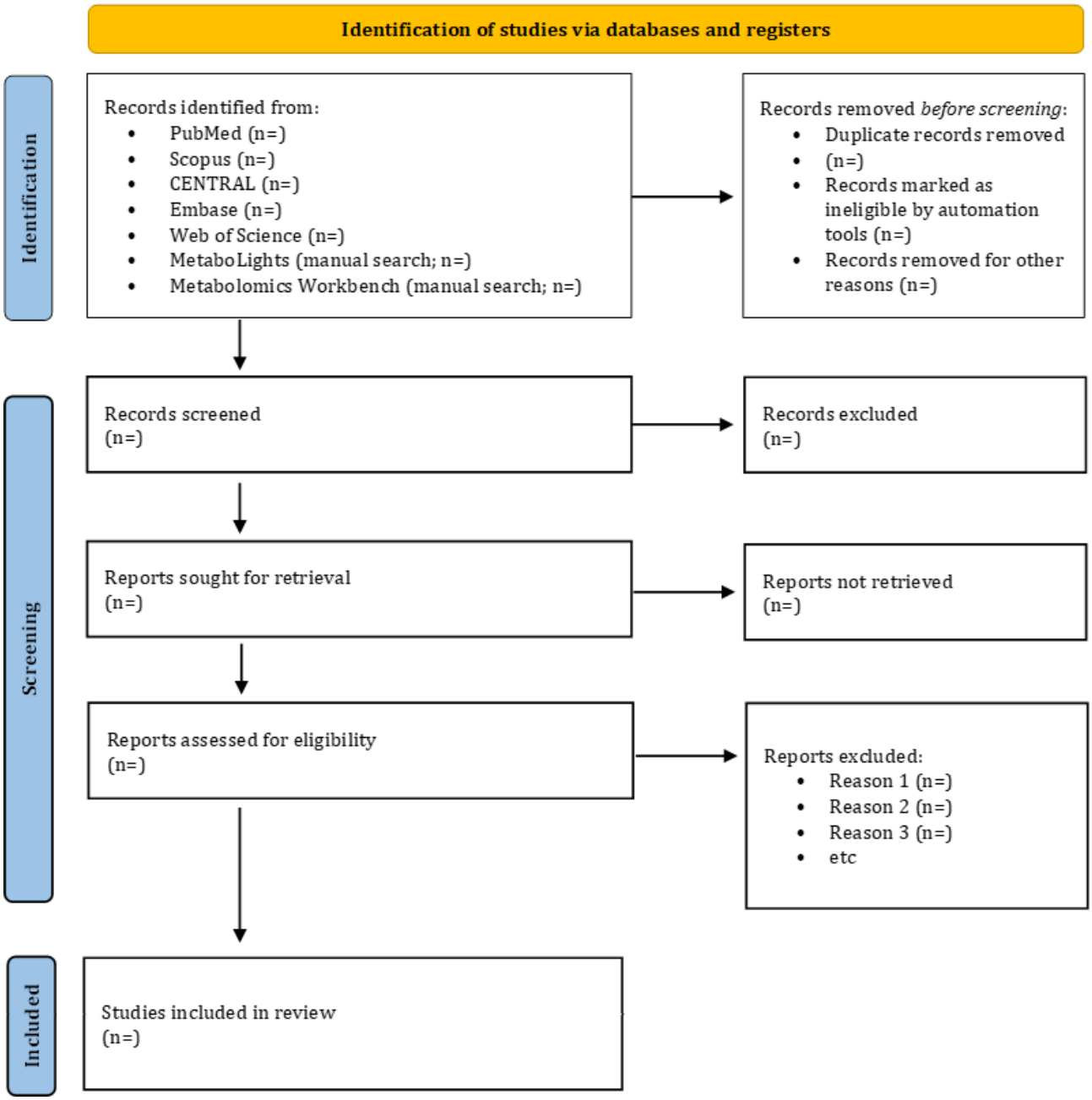

Discussion

Overview

There is increasing recognition of the important role that metabolomics can play in advancing the understanding, detection, and management of hypertension, yet large-scale, systematic data on its application in hypertensive populations remain limited. Comprehensive studies investigating metabolomic signatures and their association with BP are critical for elucidating the underlying mechanisms of hypertension and informing the development of more precise diagnostic and therapeutic strategies.

A central hypothesis in this field is that a combination of genetic, environmental, and lifestyle factors, particularly in vulnerable populations such as those with lower socioeconomic status or coexisting metabolic disorders, may contribute to distinct metabolomic profiles associated with increased hypertension risk [26]. Additionally, variability in the accuracy and validity of different metabolomic techniques may result in inconsistent detection of relevant biomarkers, potentially leading to underestimation or overestimation of hypertension risk. The interplay of multiple contextual risk factors and methodological differences could jointly influence the observed associations between metabolomic markers and hypertension [27].

This systematic review and meta-analysis will synthesize the available literature to clarify the relationship between metabolomic profiles and hypertension, including the prevalence and distribution of key biomarkers; the influence of risk factors such as age, sex, ethnicity, and comorbidities; and the impact of different detection methods on study outcomes. By integrating data from diverse populations and analytical platforms, this review aims to provide a more nuanced and comprehensive

XSL•FO
RenderX

understanding of the role of metabolomics in hypertension. The findings are expected to highlight that factors such as socioeconomic status, lifestyle, and methodological limitations in metabolite detection are key barriers to the effective clinical translation of metabolomics in hypertension. Furthermore, the review will provide a transferable framework for the application of metabolomics in hypertension research and practice.

Limitations

A limitation of this review is that data on metabolomic profiles in hypertension are often derived from secondary analyses of clinical or epidemiological datasets, which may be subjected to incomplete reporting or variability in sample handling and analytical protocols. While such data reflect real-world clinical practice, they may also introduce bias or limit the generalizability of the findings. Implementing a model that requires prospective, standardized metabolomic assessments may not be feasible in all health care settings, particularly where resources or expertise are limited. It is important to recognize that the proposed framework is preliminary and should not be directly applied in clinical practice without further validation. Another limitation is the potential for selection bias, as studies may vary in the timing of sample collection relative to disease onset or treatment and may exclude individuals with acute or secondary forms of hypertension, thereby restricting the applicability of the findings to the broader hypertensive population. This review is also limited by the exclusion of participants with major comorbidities or secondary hypertension, potentially reducing generalizability to real-world hypertensive populations. This criterion was applied to isolate metabolomic signatures specific to primary hypertension, and future research including comorbid groups will be needed to assess external validity.

Conclusions

The findings from this systematic review will inform future research directions by identifying priority areas for investigation, including the need for longitudinal studies that incorporate serial metabolomic assessments to capture temporal changes across different stages of hypertension development and progression. Validation of consistently identified metabolite-hypertension associations in independent, prospective cohorts will be critical for establishing causality and assessing the temporal stability of metabolomic signatures. The review will also identify methodological gaps requiring standardization, such as sample collection protocols, analytical platform harmonization, and data processing procedures, which must be addressed before metabolomic findings can be translated into clinical applications.

The clinical integration of metabolomics into hypertension management will require a staged approach informed by the evidence synthesized in this review. On the basis of the strength and consistency of the findings, several translational pathways can be envisioned. If the review identifies metabolites with robust, replicated associations with hypertension across multiple studies and platforms, these could be incorporated into research-focused risk assessment tools or used to stratify participants in clinical trials of antihypertensive interventions. This would not require changes to clinical guidelines but would enhance research efficiency and mechanistic understanding. Metabolite panels showing consistent associations with hypertension could be evaluated in targeted validation studies comparing their additive predictive value beyond conventional risk factors. If validated metabolite panels demonstrate clinically meaningful improvements in risk prediction (eg, reclassification of 10%-15% of individuals from intermediate to high- or low-risk categories), they could be considered for incorporation into hypertension screening algorithms for high-risk populations, similar to how lipid panels are used in cardiovascular risk assessment.

Acknowledgments

The authors thank the Director General of Health of Malaysia and the Director of the Institute for Medical Research of Malaysia for approval to publish this research. The research has been registered with the Ministry of Health Malaysia (NMRR ID-25-00732-FFP).

Funding

The authors declared no financial support was received for this work.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

MNMN and PSNMK contributed to conceptualization, data collection, data analysis, data interpretation, manuscript writing, and critical revision.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Search strategy.
[DOCX File, 15 KB - resprot_v15i1e77536_app1.docx]

Checklist 1
PRISMA 2020 checklist.
[DOCX File, 275 KB - resprot_v15i1e77536_app2.docx]

References

1. Lee EM. When and how to use ambulatory blood pressure monitoring and home blood pressure monitoring for managing hypertension. Clin Hypertens 2024 Apr 1;30(1):10. [doi: 10.1186/s40885-024-00265-w] [Medline: 38556887]
2. Mizuno H, Choi E, Kario K, et al. Diagnostic accuracy of office blood pressure measurement and home blood pressure monitoring for hypertension screening among adults: Results from the IDH study. J Am Heart Assoc 2023 Dec 19;12(24):e030150. [doi: 10.1161/JAHA.123.030150] [Medline: 38084733]
3. Pilz N, Picone DS, Patzak A, et al. Cuff-based blood pressure measurement: challenges and solutions. Blood Press 2024 Dec;33(1):2402368. [doi: 10.1080/08037051.2024.2402368] [Medline: 39291896]
4. Cai P, Zhong W, Wang Y, Wang X. Effects of white-coat, masked and sustained hypertension on coronary artery stenosis and cardiac arrhythmia. Hypertens Res 2020 Feb;43(2):121-131. [doi: 10.1038/s41440-019-0342-3] [Medline: 31624353]
5. Huang QF, Yang WY, Asayama K, et al. Ambulatory blood pressure monitoring to diagnose and manage hypertension. Hypertension 2021 Feb;77(2):254-264. [doi: 10.1161/HYPERTENSIONAHA.120.14591] [Medline: 33390042]
6. George J, MacDonald T. Home blood pressure monitoring. Eur Cardiol 2015 Dec;10(2):95-101. [doi: 10.15420/ecr.2015.10.2.95] [Medline: 30310433]
7. Parati G, Lombardi C, Pengo M, Bilo G, Ochoa JE. Current challenges for hypertension management: From better hypertension diagnosis to improved patients' adherence and blood pressure control. Int J Cardiol 2021 May 15;331:262-269. [doi: 10.1016/j.ijcard.2021.01.070] [Medline: 33548384]
8. Townsend RR. Out-of-office blood pressure monitoring: A comparison of ambulatory blood pressure monitoring and home (self) monitoring of blood pressure. Hypertension 2020 Dec;76(6):1667-1673. [doi: 10.1161/HYPERTENSIONAHA.120.14650] [Medline: 33012202]
9. Gonzalez-Covarrubias V, Martínez-Martínez E, Del Bosque-Plata L. The potential of metabolomics in biomedical applications. Metabolites 2022 Feb 19;12(2):194. [doi: 10.3390/metabo12020194] [Medline: 35208267]
10. Aderemi AV, Ayeleso AO, Oyedapo OO, Mukwevho E. Metabolomics: A scoping review of its role as a tool for disease biomarker discovery in selected non-communicable diseases. Metabolites 2021 Jun 25;11(7):418. [doi: 10.3390/metabo11070418] [Medline: 34201929]
11. Mastrangelo A, Barbas C. Chronic diseases and lifestyle biomarkers identification by metabolomics. Adv Exp Med Biol 2017;965:235-263. [doi: 10.1007/978-3-319-47656-8_10] [Medline: 28132183]
12. Ding Y, Lin H, Chen X, et al. Comprehensive metabolomics profiling reveals common metabolic alterations underlying the four major non-communicable diseases in treated HIV infection. EBioMedicine 2021 Sep;71:103548. [doi: 10.1016/j.ebiom.2021.103548] [Medline: 34419928]
13. Oliveira ND, Balikian Júnior P, Júnior ADC, et al. Environmental planning and non-communicable diseases: a systematic review on the role of the metabolomic profile. Int J Environ Res Public Health 2023 Jul 23;20(14):6433. [doi: 10.3390/ijerph20146433] [Medline: 37510665]
14. Boopathi S, Priya PS, Haridevamuthu B, et al. Expanding germ-organ theory: understanding non-communicable diseases through enterobacterial translocation. Pharmacol Res 2023 Aug;194:106856. [doi: 10.1016/j.phrs.2023.106856] [Medline: 37460001]
15. Graça ICR, Martins C, Ribeiro F, Nunes A. Understanding hypertension: a metabolomic perspective. Biology (Basel) 2025 Apr 11;14(4):403. [doi: 10.3390/biology14040403] [Medline: 40282268]
16. Page MJ, McKenzie JE, Bossuyt PM, et al. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. J Clin Epidemiol 2021 Jun;134:103-112. [doi: 10.1016/j.jclinepi.2021.02.003] [Medline: 33577987]
17. Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). Environ Int 2024 Apr;186:108602. [doi: 10.1016/j.envint.2024.108602] [Medline: 38555664]
18. Heller SR, McNaught A, Pletnev I, Stein S, Tchekhovskoi D. InChI, the IUPAC International Chemical Identifier. J Cheminform 2015;7(1):23. [doi: 10.1186/s13321-015-0068-4] [Medline: 26136848]
19. Kim S, Chen J, Cheng T, et al. PubChem in 2021: new data content and improved web interfaces. Nucleic Acids Res 2021 Jan 8;49(D1):D1388-D1395. [doi: 10.1093/nar/gkaa971] [Medline: 33151290]
20. Wishart DS, Guo A, Oler E, et al. HMDB 5.0: the Human Metabolome Database for 2022. Nucleic Acids Res 2022 Jan 7;50(D1):D622-D631. [doi: 10.1093/nar/gkab1062] [Medline: 34986597]
21. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res 2016 Jan 4;44(D1):D457-D462. [doi: 10.1093/nar/gkv1070] [Medline: 26476454]

XSL•FO
RenderX

22. Fahy E, Subramaniam S. RefMet: a reference nomenclature for metabolomics. Nat Methods 2020 Dec;17(12):1173-1174. [doi: 10.1038/s41592-020-01009-y] [Medline: 33199890]
23. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
24. PRISMA-protocols. PRISMA. 2025. URL: https://www.prisma-statement.org/protocols [accessed 2026-03-08]
25. Cochrane handbook for systematic reviews of interventions. Cochrane. 2024. URL: https://training.cochrane.org/handbook/current [accessed 2025-05-02]
26. Hernandez-Baixauli J, Quesada-Vázquez S, Mariné-Casadó R, et al. Detection of early disease risk factors associated with metabolic syndrome: a new era with the NMR metabolomics assessment. Nutrients 2020 Mar 18;12(3):806. [doi: 10.3390/nu12030806] [Medline: 32197513]
27. Al Ashmar S, Anwardeen NR, Anlar GG, Pedersen S, Elrayess MA, Zeidan A. Metabolomic profiling reveals key metabolites associated with hypertension progression. Front Cardiovasc Med 2024;11:1284114. [doi: 10.3389/fcvm.2024.1284114] [Medline: 38390445]

Abbreviations

AUC: area under the curve
BP: blood pressure
GC-MS: gas chromatography–mass spectrometry
LC-MS: liquid chromatography–mass spectrometry
MeSH: Medical Subject Headings
NCD: noncommunicable disease
NMR: nuclear magnetic resonance
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
ROBINS-E: Risk of Bias in Nonrandomized Studies of Exposures
SBP: systolic blood pressure

Edited by J Sarvestan; submitted 14.May.2025; peer-reviewed by C Exeter, S Avdeev; revised version received 22.Jan.2026; accepted 23.Jan.2026; published 17.Mar.2026.

Please cite as:
Mohd Nawi MN, Megat Kamaruddin PSN
The Role of Metabolomics in Understanding Hypertension: Protocol for a Systematic Review and Meta-Analysis
JMIR Res Protoc 2026;15:e77536
URL: https://www.researchprotocols.org/2026/1/e77536
doi:10.2196/77536

© Mohd Naeem Mohd Nawi, Puteri Sofia Nadira Megat Kamaruddin. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effects of Cash Transfer and Incentive Programs on Service Utilization and Treatment Outcomes Related to Neglected Tropical Diseases and Their Impact on Health and Nutrition in Low- and Middle-Income Countries: Protocol for a Systematic Review

S M Tafsir Hasan[1*], MBBS, MS; Radhika Dayal[2*], MA; Amena Al Nishan[1], MBBS; Sunetra Ghatak[2,3], PhD; Puja Chakraborty[1], MBBS; Sudeshna Maitra[2], PhD; Sumaiya Tasneem Raisa[1], MBBS; Nizamuddin Khan[2], PhD; Dinesh Mondal[1], MBBS, PhD; Tahmeed Ahmed[1*], MBBS, PhD; Avishek Hazra[2*], PhD

[1]International Centre for Diarrhoeal Disease Research Bangladesh (icddr,b), 68, Shaheed Tajuddin Ahmed Sarani, Mohakhali, Dhaka, Bangladesh

[2]PopulationCouncil Consulting Private Limited, Noida, Uttar Pradesh, India

[3]O. P. Jindal Global University, Sonipat, Haryana, India

[*]these authors contributed equally

Corresponding Author:
S M Tafsir Hasan, MBBS, MS
International Centre for Diarrhoeal Disease Research Bangladesh (icddr,b), 68, Shaheed Tajuddin Ahmed Sarani, Mohakhali, Dhaka, Bangladesh

Abstract

Background: Despite a global decreasing trend in neglected tropical diseases (NTDs), progress is not on track to meet the Sustainable Development Goals by 2030. Evidence suggests a positive effect of conditional cash transfer programs on controlling NTDs. However, it remains to be evaluated whether other financial incentive programs exert similar effects on NTD-related service utilization and treatment outcomes in low- and middle-income countries (LMICs).

Objective: This systematic review aims to evaluate the effects of cash transfer and incentive programs on NTD-related service utilization and treatment outcomes in LMICs; to examine how covariates, program design, and implementation plan influence these outcomes; and to examine how NTD-related service utilization and treatment outcomes influence nutritional status and health in LMICs.

Methods: This review will follow the PICOS (Population, Interventions, Comparators, Outcomes, Study Design) framework. The population of this review will be restricted to adults and children of all ages in LMICs. For intervention selection, we will include any program or policies addressing socioeconomic disadvantage through the provision of cash transfers or incentive programs, including but not limited to in-kind transfers, food vouchers, free medicine vouchers, discount coupons, and microcredit to households or individuals. Eligible comparators will be the population who did not receive any intervention, recipients of interventions from before the intervention, and populations or areas exposed to different levels of intervention coverage. The primary outcomes will be service utilization and treatment outcomes related to any of the 21 NTDs prioritized by the World Health Organization. Studies without any financial incentive intervention will be excluded. Studies available in English between 2000 and 2024 will be searched in databases, including PubMed, Google Scholar, Scopus, Cochrane, Embase, and CINAHL. Screening, data extraction, and risk of bias assessment will be conducted independently by two reviewers, and discrepancies, if any, would be resolved by a third reviewer. Quality assessment will use the Revised Cochrane Risk of Bias for randomized trials (RoB 2), Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I), the Joanna Briggs Institute (JBI) checklist for cross-sectional studies, and the Critical Appraisal Skills Programme (CASP) tool for qualitative studies. Data synthesis will involve narrative summaries by tables and figures, a meta-analysis will be undertaken, and subgroup analyses will be performed where feasible.

Results: This project is supported by the Children's Investment Fund Foundation and was funded in March 2024. As of January 15, 2026, article search and title-abstract screening have been completed, and full-text screening is expected to be completed by February 2026. Findings will then be synthesized and reported by the end of April 2026.

Conclusions: The limitations of the study may include potential language bias and methodological heterogeneity, which may limit the feasibility of meta-analysis. The findings of this review will inform policymakers in developing more effective strategic plans for controlling NTDs in LMICs.

Trial Registration: PROSPERO CRD42024627804; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024627804

XSL•FO
RenderX

International Registered Report Identifier (IRRID): PRR1-10.2196/76450

(JMIR Res Protoc 2026;15:e76450) doi:10.2196/76450

KEYWORDS

neglected tropical disease; cash transfer; incentive program; service utilization; treatment outcome; nutrition; health; low- and middle-income countries

Introduction

Background

Neglected tropical diseases (NTDs) remain a major global public health concern, affecting an estimated 1.62 billion people in tropical and subtropical climatic zones [1,2] and accounting for approximately 200,000 deaths and 14.5 million disability-adjusted life years annually [2,3]. Although there has been a general decline in the global burden of NTDs [2], the progress is not yet on track to meet the ambitious Sustainable Development Goals 2030 target of a 90% reduction in the number of people requiring interventions for NTDs [3]. NTDs continue to pose a big challenge for low- and middle-income countries (LMICs) due to their intricate links with poverty, inadequate hygiene, poor sanitation, and limited health care access [4,5].

Poverty is a key factor exacerbating the burden of NTDs in LMICs. Financial deprivation and disease are closely intertwined, which influence the social and environmental determinants of NTDs, including water, sanitation, and hygiene, education, housing, and health care accessibility [5-7]. Poverty highly modulates the accessibility and utilization of the services related to the 5 important NTD control strategies recommended by the World Health Organization (WHO): innovative and intensified disease management; preventive chemotherapy; vector control; veterinary public health measures; and provision of safe water, sanitation, and hygiene practices [8,9]. NTDs disproportionately affect impoverished and marginalized communities, resulting in increased catastrophic health spending and out-of-pocket expenses [10]. Already living in poverty before the onset of these diseases, affected households often face deeper financial hardship as rising out-of-pocket costs push them further below the poverty line [11-13]. To manage these expenses, they often resort to coping strategies such as borrowing money or selling assets, which further exacerbates their poverty. This cycle undermines access to health care services and negatively affects treatment outcomes [13]. Physical disabilities due to NTDs, such as blindness, lesions, and swelling, may also contribute to decreased income ability [14]. This financial burden perpetuates a cycle of poverty and NTD burden, hindering service uptake and treatment [15]. Consequently, the detrimental effects of NTDs are amplified, including poor NTD-related service utilization and adherence, treatment failure, disability, and deterioration of general health and well-being [15].

Moreover, infection with NTDs increases the nutritional burden of the hosts and impedes overall health. NTDs have been reported to be associated with stunting, anemia, iron deficiency, micronutrient deficiencies, and malnutrition [16-19]. Simultaneously, nutrient deficiency may also increase the susceptibility and vulnerability of NTD-affected individuals and lead to NTD treatment failure [18,20,21]. Actions taken to control NTDs may eventually improve nutritional outcomes, which, in turn, will also influence overall health and well-being, particularly in resource-limited settings.

Despite various programs and policies implementing innovative strategies to improve NTD campaign coverage, multiple layers of poverty-associated and financial obstacles continue to hinder appropriate health-seeking behaviors in affected populations. These barriers include attending screening clinics, receiving adequate care, and treatment adherence [22-25]. Cash transfer programs have demonstrated potential for improving NTD outcomes [26-29]. These programs aim to provide small-scale monetary assistance to individuals or households to break the intergenerational cycle of poverty. Conditional cash transfers (CCTs) impart benefits through a 2-way strategy: direct cash assistance and enforced conditionalities [30]. It is also possible that conditionalities related to the control of NTDs may also impart positive effects on nutrition and overall health outcomes in LMICs. Large-scale, national-level cash transfer programs in Latin America have not only proven effective in improving social protection, economic stability, and sustained developmental outcomes [31] but have also generated interest in using cash transfers to target context-specific health outcomes, including NTD treatment outcomes, maternal health, child health, cognitive development, nutrition, and health care utilization [26,32-34].

A recent systematic review highlighted the positive impact of CCT programs in controlling NTDs. It indicated that the targeted benefits of these programs for vulnerable populations can help reduce health inequalities associated with NTDs. For example, cash transfers have been shown to reduce leprosy incidence and increase the uptake of deworming treatments in certain contexts [35]. However, the review had a major limitation in that it included and synthesized evidence for only 3 of the 21 NTDs and disease groups prioritized by WHO [2,3,35]. While other incentive programs, such as in-kind transfers, microcredits, vouchers for transport, medicines, or food, have shown the potential to promote health knowledge, reduce health inequalities, and improve health-seeking behaviors, their direct or indirect effects on NTD-related indicators have not been comprehensively studied in LMICs [36,37].

Given this context, there is a need for an evidence synthesis on the effects of cash transfers and other incentive programs on NTD-related service utilization and treatment outcomes in LMICs. This proposed systematic review and meta-analysis (depending on information availability) will conduct a comprehensive synthesis and unbiased assessment of the existing studies in this area, critically assessing the strengths and limitations of existing evidence.

XSL•FO
RenderX

Objectives

The systematic review aims to answer the following three questions: (1) What are the effects of cash transfer and incentive programs on NTD-related service utilization and treatment outcomes in LMICs? (2) How do incentive program design (type, amount, and frequency), implementation plan, and other covariates influence NTD-related service utilization and treatment outcomes in LMICs? and (3) How do NTD-related service utilization and treatment outcomes influence nutritional status, overall health, and well-being in LMICs?

Methods

Overview

This protocol was developed in accordance with the guidelines of the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) [38]. A PRISMA-P checklist for this protocol is provided in Checklist 1. The protocol has been registered in PROSPERO (CRD42024627804).

Eligibility Criteria

The eligibility criteria for studies to be included in this review are adapted based on the PICOS (Population, Interventions, Comparators, Outcomes, Study Design) framework [39], as described below.

Population

The population of interest in this review will be adults and children of all ages residing in LMICs.

Interventions

For intervention, this review will include programs or policies addressing socioeconomic disadvantages through the provision of cash transfers or financial incentive programs, including but not limited to in-kind transfers, food vouchers, free medicine vouchers, discount coupons, and microcredit to households or individuals.

CCT programs require beneficiaries to comply with certain conditionalities (eg, regular health check-ups) [31], while unconditional cash transfer programs do not set such requirements [40]. Microcredit is the extension of very small loans (microloans) to impoverished borrowers who typically lack collateral, steady employment, and a verifiable credit history [41]. It is designed to support entrepreneurship and alleviate poverty [41]. An in-kind transfer is a transfer of goods or services rather than cash [42]. Other incentive programs may also include giving free medicine vouchers, food vouchers, discount coupons, social protection, social programs, scholarship programs, family allowances, and microfinance.

Comparators

This review will include studies with suitable comparators; studies without comparators will be excluded. Comparators will include individuals or households who did not receive any cash transfers or incentive programs, recipients of cash transfers or incentive programs from the preintervention period, and populations or areas with varying levels of intervention coverage.

Outcomes

The primary outcomes of interest will be service utilization and treatment outcomes related to any of the 21 NTDs and disease groups prioritized by WHO [2]. These include Buruli ulcer; Chagas disease; dengue and chikungunya; dracunculiasis; echinococcosis; foodborne trematodiases; human African trypanosomiasis; leishmaniasis; leprosy; lymphatic filariasis; mycetoma, chromoblastomycosis, and other deep mycoses; noma; onchocerciasis; rabies; scabies and other ectoparasitoses; schistosomiasis; soil-transmitted helminthiases; snakebite envenoming; taeniasis/cysticercosis; trachoma; and yaws.

Service utilization related to WHO-recommended core strategic interventions for each of the 21 NTDs and disease groups will be evaluated. The aspects of service utilization evaluated will include, but not be limited to, health-seeking behavior, attendance at screening clinics, treatment adherence, and NTD campaign coverage. Treatment outcomes related to WHO-recommended core strategic interventions for each of the 21 NTDs will include, but not be limited to, prevalence, incidence, detection rate, cure rate, disability rate, death rate, and components of morbidity management, including surgery, wound management, physiotherapy, and long-term rehabilitation. A comprehensive list of core strategic interventions and treatment outcomes for each of the 21 NTDs and disease groups is provided in Table 1 [1,3,43-48].

XSL•FO
RenderX

Table . List of core strategic interventions and treatment outcomes for each of the 21 neglected tropical diseases (NTDs) and disease groups.

Serial	NTD name	Core strategic interventions					Treatment outcomes
		Innovative and intensified disease management	Preventive chemotherapy	Vector control	Veterinary public health measures	Provision of safe WASH[a] practices	
1	Buruli ulcer	Early detection, case management (provision of antibiotics: rifampicin, clarithromycin, moxifloxacin, streptomycin, and hygienic wound care)	Bacillus Calmette–Guérin (BCG) vaccination	Aquatic insects, adult mosquitoes, possums, and biting arthropods	Not applicable	Safe water storage practices	Disability, morbidity management including surgery, wound, and lymphedema management, physiotherapy and long-term rehabilitation, prevalence, incidence, detection rate, cure rate, disability rate, death rate
2	Chagas disease	Case management (benznidazole and nifurtimox), MDA[b], surgical treatment	Blood screening before blood transfusion and organ transplantation	Spraying with residual insecticides to remove triatomine bugs, home cleanliness, and housing improvements (eg, crack-free walls, bed nets), wear protective clothing	Not applicable	Good hygiene practices	Prevalence, incidence, detection rate, cure rate, surgery rate, disability rate, death rate
3	Dengue and chikungunya	Case management (fluid management, proactive treatment of hemorrhage, such as platelet transfusion or whole blood transfusion, paracetamol)	Vaccination (dengue)	Reducing mosquito habitats (eg, environmental modification, use of lids), wire mesh, mosquito net, mosquito trap, mosquito insecticide, mosquito repellents	Not applicable	Safe water storage practices, safe water disposal, waste management	Prevalence, incidence, detection rate, cure rate, disability rate, death rate
4	Dracunculiasis	Surveillance, early detection of cases, case containment, cash reward scheme for voluntary reporting, metronidazole or thiabendazole, health education	Not applicable	Temephos to control water fleas/cyclops	Proactive tethering of dogs	Safe drinking water supply and sanitation, water filtration, preventing water contamination	Prevalence, incidence, detection rate, cure rate, disability rate, death rate
5	Echinococcosis	Albendazole prophylaxis, surgery, early detection by ultrasound	Not applicable	Not applicable	Deworming of dogs, foxes, vaccination of sheep, safe slaughtering and disposal	Handwash, avoiding contaminated food and water	Prevalence, incidence, detection rate, cure rate, disability rate, surgery rate, total or partial cystopericystectomy, liver damage, death rate

XSL•FO
RenderX

Serial	NTD name	Core strategic interventions					Treatment outcomes
		Innovative and intensified disease management	Preventive chemotherapy	Vector control	Veterinary public health measures	Provision of safe WASH[a] practices	
6	Foodborne trematodiases	Anthelminthic medicines (praziquantel, triclabendazole), PAIR (puncture, aspiration, instillation, and respiration) methods, surgery (partial hepatectomy), Provision of care and rehabilitative services	MDA with triclabendazole (*Fasciola* spp.) and praziquantel (small liver flukes and *Paragonimus* spp.)	Not applicable	Livestock and domestic animal treatment, safe fish farming, one health intervention	Safe disposal of fecal waste, safe water, safe food practices, safe WASH practices	Prevalence, incidence, detection rate, cure rate, disability rate, surgery rate, liver damage, death rate
7	Human African trypanosomiasis	Screening clinic attendance, active case detection, treatment of gHAT[c]: pentamidine, eflornithine and nifurtimox, fexinidazole, melarsoprol, suramin	Not applicable	Baits and traps, impregnated screens, insecticide spraying, reducing the disease reservoir, controlling the tsetse fly vector	Treatment of animals (cattle, pigs)	Safe water, safe WASH practices	Prevalence, incidence, detection rate, cure rate, disability rate, death rate
8	Leishmaniasis	Early diagnosis, cryotherapy, thermotherapy, medication (liposomal amphotericin B, pentavalent antimonials, miltefosine), rapid diagnostic test rk-39	Not applicable	Insecticide spraying, insecticide-treated nets, environmental management, protection from sandfly bites	Rodent control, vaccination for dogs	Safe WASH practices	Prevalence, incidence, detection rate, cure rate, disability rate, death rate
9	Leprosy	Early detection, MDA, physical therapy, multidrug therapy (rifampicin, clofazimine, dapsone, ofloxacin, and minocycline)	Postexposure prophylaxis of rifampicin, vaccination with BCG, Mycobacterium w (Mw)	Not applicable	Not applicable	WASH in health facility	Prevalence, incidence, detection rate, cure rate, disability rate, death rate

XSL•FO
RenderX

Serial	NTD name	Core strategic interventions					Treatment outcomes
		Innovative and intensified disease management	Preventive chemotherapy	Vector control	Veterinary public health measures	Provision of safe WASH[a] practices	
10	Lymphatic filariasis	Essential package of care (skincare and hygiene, exercises for lymphedema, treatment of adenolymphangitis), surgery to cure hydrocele, antifilarial medicines (albendazole, diethylcarbamazine citrate), ivermectin for lymphatic filariasis co-endemic countries	MDA, annual MDA with diethylcarbamazine citrate and albendazole, home-based disability alleviation and prevention	Vector control to reduce transmission, avoid mosquito bites	Not applicable	Hygiene, sanitation improvement, personal hygiene and self-care of the affected limbs to avoid secondary infections, agriculture, livestock, wildlife, environment (One Health), avoid mosquito bites	Prevalence, incidence, detection rate, cure rate, disability rate, hydrocele management, surgery rate, mental health, death rate
11	Mycetoma, chromoblastomycosis and other deep mycoses (paracoccidioidomycosis [PCM], caused by *Paracoccidioides* spp. and sporotrichosis [ST], caused by *Sporothrix* spp.)	Long-term itraconazole+ antibiotic administration, physical therapy, wound care, surgery (excision, debridement, amputation)	Self-protection by wearing long clothes, closed-toed shoes	Not applicable	Pet control for sporotrichosis, feral cat management	Personal hygiene and self-care of the affected limbs to avoid secondary infections	Prevalence, incidence, detection rate, cure rate, disability rate, local excision, debridement, amputation, surgery rate, death rate
12	Noma	Case management with antibiotics, mouthwash, health education	Nutrition supplements	Not applicable	Not applicable	Personal hygiene	Prevalence, incidence, detection rate, cure rate, disability rate (disfiguration), mental health, death rate
13	Onchocerciasis	Ivermectin, doxycycline, eye care, MDA	MDA of ivermectin for community-based prevention	Insecticides at black fly habitats and larval breeding sites, avoid blackfly bites	Not applicable	Not applicable	Prevalence, incidence, detection rate, cure rate, disability rate (ocular morbidity rate), death rate
14	Rabies	Postexposure prophylaxis (PEP) with rabies vaccine and rabies immunoglobulin	Pre-exposure vaccination for people at high risk (eg, laboratory staff working with rabies virus, veterinarians and animal handlers)	Wild mammals (dogs, bats) control	Mass vaccination of dogs, mammal pet vaccination, agriculture, livestock, wildlife, environment (One Health)	Not applicable	Prevalence, incidence, detection rate, death rate

XSL•FO
RenderX

Serial	NTD name	Core strategic interventions					Treatment outcomes
		Innovative and intensified disease management	Preventive chemotherapy	Vector control	Veterinary public health measures	Provision of safe WASH[a] practices	
15	Scabies and other ectoparasitoses	Topical scabicides, permethrin, benzyl benzoate, malathion, and sulfur ointment, oral ivermectin	MDA using oral ivermectin	Not applicable	Not applicable	WASH practices, hygiene practices	Prevalence, incidence, detection rate, cure rate, co-morbidities (kidney disease, rheumatic heart disease, systemic infection)
16	Schistosomiasis	Case management and surgery, education	MDA with praziquantel	Snail control	Keeping animals away from transmission sites	Safe water, improved sanitation, excreta management, hygiene	Prevalence, incidence, detection rate, cure rate, disability rate, surgery rate, death rate
17	Soil-transmitted helminthiases	Case management (with albendazole/mebendazole/ivermectin), education of hygiene	MDA with albendazole/mebendazole/ivermectin, antenatal care preventive treatment	Not applicable	Not applicable	WASH practices, safe water, improved sanitation, excreta management, hygiene	Prevalence, incidence, detection rate, cure rate, disability rate, death rate
18	Snakebite envenoming	High-quality, safe, and effective snake antivenoms; wound care; ventilation; surgery; WHO[d] Regional Action Plan for prevention and control of snakebite envenoming 2022 - 2030	Protective footwear, sealing house wall, roof and door gaps, bed nets, household rodent controls to keep away the snakes	Not applicable	Not applicable	Not applicable	Prevalence, incidence, detection rate, cure rate, disability rate, surgery rate, intensive care unit (ICU) admission rate, death rate
19	Taeniasis/cysticercosis	Case management (praziquantel, niclosamide, albendazole), neurological treatment, surgery	MDA with niclosamide/ praziquantel/ albendazole, community awareness	Not applicable	Improved pig husbandry, pig vaccination	WASH practices, proper cooking	Prevalence, incidence, detection rate, cure rate, disability rate (neurological symptoms), surgery rate, death rate
20	Trachoma	Eye surgery to treat trachomatous trichiasis, TT (the S of the SAFE strategy) Azithromycin	MDA of azithromycin and tetracycline eye ointment (WHO-recommended SAFE strategy)	Reducing breeding sites for muscid flies	Not applicable	Access to water and improved sanitation	Prevalence, incidence, detection rate, cure rate, disability rate (blindness), surgery rate

XSL•FO
RenderX

Serial	NTD name	Core strategic interventions					Treatment outcomes
		Innovative and intensified disease management	Preventive chemotherapy	Vector control	Veterinary public health measures	Provision of safe WASH[a] practices	
21	Yaws	Case management (azithromycin medication, intramuscular benzathine benzylpenicillin), Morges strategy, health education	Mass treatment, also known as total community treatment (TCT), involves giving oral azithromycin (single dose)	Not applicable	Not applicable	Personal hygiene	Prevalence, incidence, detection rate, cure rate, disability rate (disfiguration)

[a]WASH: water, sanitation, and hygiene.

[b]MDA: mass drug administration.

[c]gHAT: gambiense human African trypanosomiasis.

[d]WHO: World Health Organization.

The review will include only studies reporting primary outcomes. However, in the included studies, exploratory outcomes related to the nutrition, health, and well-being of adults and children will be analyzed to examine how NTD-related service utilization and treatment outcomes may influence these outcomes in LMICs. Exploratory outcomes will include, but not be limited to, nutritional status, various types of malnutrition, micro- and macronutrient deficiency, anemia, cognitive development, school attendance and performance, mortality, and morbidity.

Study Design

The review will include randomized controlled trials, as well as nonrandomized studies with comparator groups, such as cohort, case-control, longitudinal, cross-sectional, ecological, and quasi-experimental studies. Additionally, impact evaluation results, program reports, qualitative studies, and studies employing methodologies like difference-in-difference, pre-post design, interrupted time series, regression discontinuity, instrumental variable estimation, and propensity score matching will be considered. Studies reported in the English language will be considered for inclusion.

Search Strategy

A scoping search will be conducted to identify key Medical Subject Headings terms and other relevant search terms, which will then be combined to develop a comprehensive search strategy. Boolean operators (AND, OR, and NOT), truncation/stemming (*), and phrase searching (“”) will be used. The search will focus on NTDs, using genus names along with key terms, such as cash transfer, unconditional cash transfer, CCT, microcredit, food voucher, and incentive. A preliminary search strategy is provided in Table 2.

XSL•FO
RenderX

Table . Preliminary search strategy.

Domain	Description	Search terms or keywords
Population	Participants	• Adults and children living in LMICs[a] • Adults, young adults, young women, young men, child, infant, adolescents, teen
Interventions	Programs	• Money, monetary, cash, financial transfer, cash transfer, payment, conditional cash transfer (CCT), unconditional cash transfer (UCT), incentive. • Financial, nonfinancial incentive programs, eg, social protection, social program, scholarship programs, family allowance, in-kind transfer, food/medicine voucher, and microfinance/microcredit.

XSL•FO
RenderX

Domain	Description	Search terms or keywords
Outcomes	Neglected tropical diseases and disease groups; core strategic interventions; NTD[b]-related service utilization; and treatment outcomes	

XSL•FO
RenderX

Domain	Description	Search terms or keywords
		• Neglected disease, neglected tropical disease, tropical medicine, tropical disease, Buruli, buruli ulcer, Bairnsdale ulcer, Daintree ulcer, Mossman ulcer, mycobacterium, *Mycobacterium ulcerans*, Chagas, Chagas disease, trypanosome, *Trypanosoma cruzi*, American trypanosomiasis, dengue, dengue virus, break-bone disease, break-bone fever, chikungunya, dracuncul, dracunculus nematode, dracunculiasis, guinea worm, guinea-worm, antihelminthic, antiparasitic, deworm, *Echinococcus*, echinococc, foodborne, trematodiases, trematodias, clonorchiasis, opisthorchiasis, fascioliasis, paragonimiasis, human African trypanosomiasis, sleeping sickness, Leishmaniasis, kala-azar, leprosy, Hansen, trypanosome, lymphatic filariasis, filaria, elephantiasis, *Wuchereria bancrofti*, brugia, onchocerciasis, rabies, river blindness, schistosome, bilharzia, helminth, roundworm, nematode, soil transmitted helminthiasis, ascaris, whipworm, trichuris, hookworm, ancylostom, taenia, cysticercos, chlamydia, tapeworm, hookworm, *Necator americanus*, *Ancylostoma duodenale*, taenia, cysticercosis, cestod, trachoma, treponem, noma, snake bite envenoming, yaws, frambesia, bejel, pinta, parangi, mycetoma, chloroblastomysis, chromomyc, mycosis, mycos, scabies, ectoparasite, mosquito habitat, Temephos, water fleas, small liver flukes, fecal waste, feces, lymphedema, adenolymphangitis, hydrocele, secondary infection, agriculture, wildlife, environment, sporotrichos, feral cats, black fly habitat, larval breeding, wild mammals, veterinarians, animal handlers, rabies virus, pig husbandry, breeding sites, etc • Early detection, antibiotic, Rifampicin, Clarithromycin, hygiene, wound care, BCG[c] vaccine, Benznidazole, Nifurtimox, MDA[d], mass drug administration, surgery, blood screening, blood transfusion, organ transplantation, insecticide, bed nets, spraying, fluid management, wire mesh, repellents, safe water practices, waste management, surveillance, cash reward scheme, voluntary reporting, Cyclops, dogs, water filtration, Albendazole, early detection by USG, deworming, deworming of dogs, sheep vaccination, handwash practices, anthelminthic medicines, Praziquantel, PAIR[e] methods, partial hepatectomy, rehabilitation, Triclabendazole, Praziquantel, one health, WASH[i] practices, screening, Pentamidine, Eflornithine, Fexinidazole, baits, cryotherapy, thermotherapy, liposomal Amphotericin B, pentavalent antimonials, Miltefosine, rodent control, dog vaccination, physical therapy, postexposure prophylaxis, PEP[f], anti-filarial medicines, vector control, sanitation, limb care, excision, debridement, amputation, long clothes, closed-toed shoes, pet control, mouthwash, nutrition, nutrition supplement, Ivermectin, Doxycycline, eye

XSL•FO
RenderX

Domain	Description	Search terms or keywords
		care, pre-exposure vaccination, scabicides, snail control, antenatal care, ANC, excreta management, antivenom, ventilation, footwear, Niclosamide, neurological treatment, pig vaccination, eye surgery, Azithromycin, Tetracycline, eye ointment, SAFE strategy, muscid flies, personal hygiene. • Attendance at screening clinics, treatment adherence, health-seeking behavior, NTD campaign coverage, etc • Prevalence, incidence, detection rate, cure rate, disability, disability rate, morbidity, morbidity rate, death rate, lymphedema management, wound management, total cystopericystectomy, partial cystopericystectomy, surgery rate, liver damage, mental health, disfiguration, ocular morbidity, comorbidities, kidney disease, rheumatic heart disease, systemic infection, ICU[g], intensive care unit, admission rate, blindness,
Study	Design	• RCTs[h], nonrandomized studies, cohort studies, longitudinal studies, case-control studies, cross-sectional studies, ecological studies, quasi-experimental studies, impact evaluation results, and program reports. • Difference-in-difference, pre-post design, time series, regression discontinuity, instrumental variable estimation, propensity score matching, and qualitative study.

[a]LMICs: low- and middle-income countries.

[b]NTD: neglected tropical diseases.

[c]BCG: Bacillus Calmette–Guérin.

[d]MDA: mass drug administration.

[e]PAIR: puncture, aspiration, injection, and re-aspiration.

[f]PEP: postexposure prophylaxis.

[g]ICU: intensive care unit.

[h]RCT: randomized controlled trial.

[i]WASH: water, sanitation, and hygiene.

Article Screening Process

The following databases will be searched: MEDLINE, PubMed, Scopus, Embase, Web of Science, POPLINE, Google Scholar, Cochrane Collection, PsycINFO, Global Health, EconLit, Social Sciences Citation Index, International Bibliography of the Social Sciences, Knowledge Commons of Population Council, and 3ie database. Gray literature databases, including WHO websites, AlignMNH, World Bank, and UNICEF websites, will also be searched for relevant reports. Websites and online resources of universities and research centers in LMICs will be reviewed to identify working papers, dissertations, and theses. Reference lists or bibliographies of included studies will also be hand-searched to include any additional relevant studies. For those studies not available in the public domain, a librarian will be contacted. Studies in English, published or available from 2000 to 2024, will be included. A search log will be maintained throughout the search process to document the database platforms or search interface used, the date the search was conducted, the search strings used, and the number of records or counts identified. The log will be updated as necessary to ensure transparency, robustness, and ease of updating during the study period.

Preliminary search documents will be stored in reference management software (EndNote or Mendeley). All the screened studies will be transferred into Microsoft Excel or Rayyan tool, a systematic review management platform [49].

Title, Abstract, and Full-Text Relevancy

Two reviewers will independently screen studies based on titles and abstracts. During the initial screening phase, decisions made by each reviewer will be blinded. After the initial screening is done, the studies screened by the two reviewers will be compared and checked for discrepancies. Any discrepancies in the categorization will be resolved through consultation between the reviewers or with another reviewer. A second-level independent review will then be conducted by a separate

XSL•FO
RenderX

reviewer to perform quality checks on studies with discrepancies and those deemed ineligible. Studies deemed relevant based on titles and abstracts will undergo full-text screening for relevancy, following the same procedures used during the title and abstract screening phase. Reasons for the exclusion of papers will be documented. The studies will be equally divided among the reviewers for relevancy screening. Data will be extracted individually by each reviewer. Any questions arising during the process will be addressed through consultation among the reviewers to reach a consensus. Eligible studies from the full-text review will be entered directly into a master data extraction spreadsheet.

Data Extraction

After the article screening process is completed, all eligible studies will be extracted into a master data extraction spreadsheet. The data extraction sheet will include the following categories: title, author(s), date of publication, type of publication (eg, summary, synthesis, single study), number and type of included studies (if a summary or synthesis), settings and population studied, interventions implemented, outcomes measured and results, and if relevant, whether results differed among subgroups, such as gender, socioeconomic status, or ethnicity. For studies with multiple outcomes or variables, all eligible outcomes will be extracted. The data extraction form will be pilot-tested and refined based on the information extracted from a few initial studies meeting the inclusion criteria. This process will help minimize bias and improve the validity and reliability of the systematic review. Authors will be contacted through email if key information is missing or if further information or clarification is required.

Risk of Bias Assessment for Individual Studies

The quality assessment will be done by using the Revised Cochrane Risk of Bias for randomized trials (RoB 2) [50], the Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I) [51], the Joanna Briggs Institute (JBI) checklist for analytical cross-sectional studies, and an adapted version of the Critical Appraisal Skills Programme (CASP) [52] for qualitative studies; ecological studies will be evaluated narratively. The quality of each included study will be independently assessed by one reviewer, with the assessment results checked by a second reviewer. Any discrepancies or disagreements will be resolved through consultation with a third reviewer to reach a consensus.

Anticipated Number of Studies to Be Included in the Review

Based on the initial search to develop the protocol, about 20 to 30 studies are anticipated to be included in the review [34,53,54]. This number might change during the actual work of the systematic review based on the search using electronic databases and gray literature.

Data Synthesis

A narrative synthesis will be produced by describing the studies. A geographical information system map will be generated to illustrate the frequency distribution of impact evaluations of cash transfer and incentive programs on NTD-related service utilization, treatment outcomes, nutritional status, and health and well-being, subject to data availability. Interventions of cash transfers and incentive programs will be classified into different categories for different types of NTDs, and their effects will be synthesized for (1) NTD-related service utilization; (2) treatment outcomes; and (3) nutritional status, health, and well-being outcomes. For this review, improvement in nutritional status and health and well-being are considered potential downstream or contextual outcomes that may occur through multiple pathways, in addition to through changes in NTD-related service utilization and treatment outcomes. These pathways may include independent effects of incentives on household resources, food security, health-seeking behavior, and overall living conditions. Given that most primary studies may not be designed to test the full causal pathway, causal inference in this review will be restricted to proximal outcomes (service utilization and treatment outcomes), while nutritional and broader health outcomes will be interpreted as associative and contextual. All the studies will be tabulated by quantitative and qualitative type, and key findings will be described. Primary and exploratory outcomes will be listed along with differences in effectiveness between different cash transfer and incentive programs for NTD control.

We will assess the certainty of evidence for quantitative outcomes using the GRADE (Grading of Recommendations, Assessment, Development and Evaluation) approach. For qualitative or narrative evidence, we will use an appropriate structured approach (eg, GRADE-CERQual) to evaluate confidence in the findings.

The possibility of performing a meta-analysis will be explored depending on the availability of a sufficient number of relevant studies. The ability to undertake a meta-analysis will depend on the heterogeneity of the measures, such as similarities or differences in population, intervention, outcomes, and methodologies. If a meta-analysis is deemed appropriate, statistical heterogeneity of effects will be assessed using χ^2 and I^2 statistics. Efforts will be made to analyze the factors explaining heterogeneity through moderator analysis, including subgroup meta-analysis. All analyses will be contingent on sufficient study density.

For qualitative studies, the themes will be identified using a deductive approach, and the findings will be used to interpret or complement the quantitative findings as part of the narrative synthesis. We will also consider stratifying synthesis by study design strength.

The synthesis will present the factors such as program design, implementation, context, and challenges that influence the impact of cash transfer or incentive programs on NTD-related service utilization, treatment outcomes, nutrition status, and health and well-being. Absolute evidence gaps and synthesis gaps will also be identified that can inform the development of new studies or programs.

Analysis of Subgroups or Subsets

The possibility of performing subgroup analyses will be explored. It will, however, depend on the evidence strength to summarize effects in special populations, such as pregnant women and geriatric populations, or to distinguish effect

XSL•FO
RenderX

differences across different levels of groups, such as female/male, child/adolescent/adult, different geographical regions, malnourished/well-nourished, and type of disease agents (viruses, bacteria, protozoa, fungi, helminths, ectoparasites, snakebites). The effects will also be examined to determine how they vary by incentive program design (type, amount, and frequency), implementation plan, study design, and relevant covariates.

Ethical Considerations

The systematic review protocol has been reviewed and approved by two independent institutional review boards: icddr,b (PR-24112) in Bangladesh and Population Council Institute (IORG0011824) in India. Although systematic reviews do not require institutional review board approvals because no human subjects were involved, these approvals were taken as an institutional requirement.

Results

This project is supported by the Children's Investment Fund Foundation and was funded in March 2024. As of January 15, 2026, an exhaustive article search and title-abstract screening have been completed, and the full-text article screening process is expected to be completed by February 2026. Findings will then be synthesized and reported by the end of April 2026. Based on our initial search, we anticipate that 20 to 30 articles will be included, although the final number may vary. The screening process will be reported using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart (Figure 1) [55].

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.

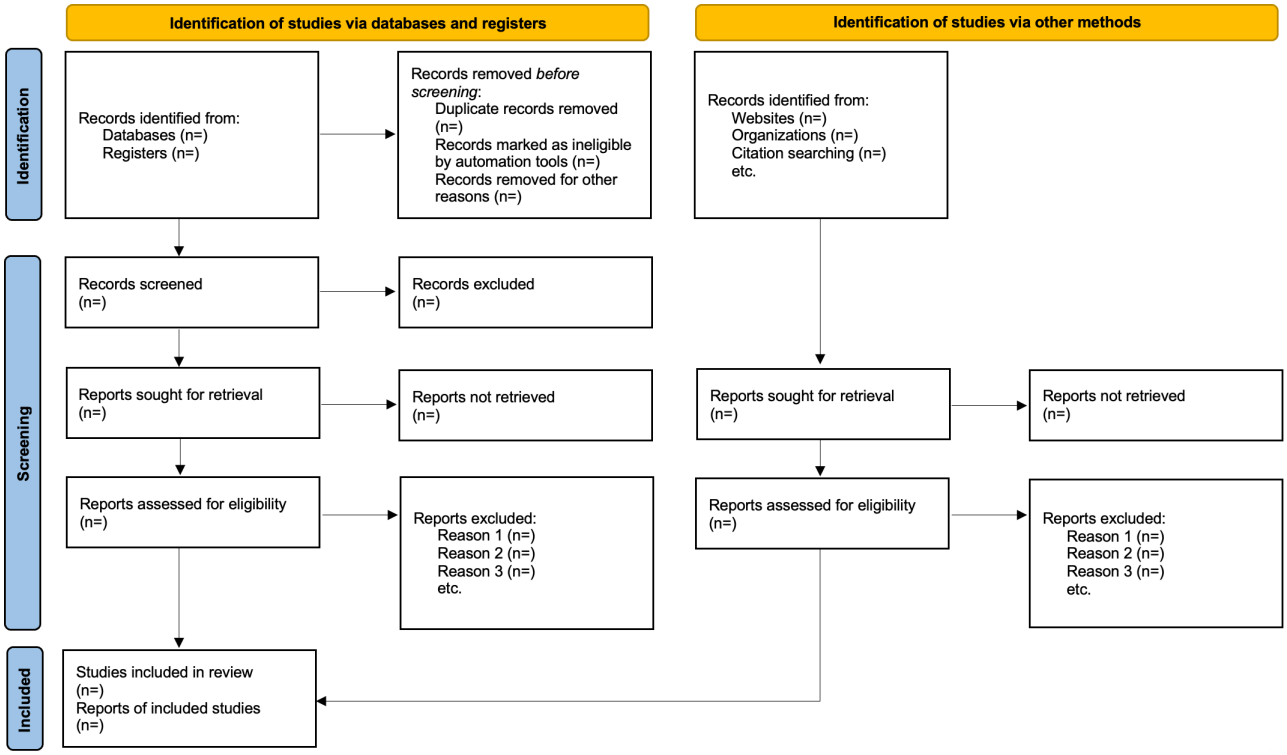

Discussion

Principal Findings

We expect that this study will be completed by April 2026. This systematic review will critically examine and provide a comprehensive assessment of the effects of cash transfer and other financial incentive programs on NTD-related service utilization and treatment outcomes in LMICs. A recent systematic review of 11 studies published in *The Lancet* found that CCTs improve NTD outcomes through both enhanced living standards from the cash benefits and direct health effects linked to health care–related conditionalities [35]. However, this review had a major limitation in that it included and synthesized evidence for only 3 of the 21 NTD groups prioritized by WHO. In contrast, this review will seek evidence across all 21 WHO-prioritized NTDs.

The financial impact of different categories of financial aid for mitigating NTDs will be assessed to inform policymakers about more effective intervention programs and prevention or treatment strategies. Additionally, the potential link between NTD-related service utilization and health and nutrition outcomes in LMICs will be evaluated. A methodological assessment of the published literature will be performed, and the findings will be compared with those of similar reviews to examine the strength of evidence.

The proposed systematic review will synthesize evidence generated over a 25-year period and will critically assess the strengths and limitations of existing studies on NTDs. The use of rigorous methodology and a robust article screening process is expected to ensure a comprehensive, up-to-date, and unbiased synthesis of evidence on the effectiveness of cash transfer and incentive programs in improving NTD-related service utilization,

XSL•FO
RenderX

treatment outcomes, and associated health and nutrition outcomes in LMICs.

One of the major anticipated limitations of this systematic review is that the available evidence may include multiple outcome measures and substantial methodological heterogeneity, potentially limiting the feasibility of meta-analysis. Additional limitations may include potential risk of bias in the included studies, possible publication bias, language bias due to the inclusion of English-language articles only, and limitations of the search strategy. We may face limited sensitivity in the electronic database searches because of the extensive NTD-related keywords and the indexing constraints of certain databases. However, supplementary strategies, including bibliographic citation searching and snowballing from the reference lists of included articles and relevant systematic reviews, are expected to help mitigate these limitations.

Conclusion

This systematic review will provide policymakers with evidence to support more effective strategic planning and targeted interventions for controlling NTDs in LMICs. The findings from this systematic review will also be valuable for researchers, funding agencies, and governmental bodies seeking to design and implement effective NTD control programs.

Dissemination Plan

The findings of this systematic review will be presented at national and international conferences and reported in a peer-reviewed journal in accordance with the PRISMA 2020 guideline for reporting systematic reviews [55]. Any amendments to this protocol during the review process will be documented in PROSPERO and reported in the final manuscript.

Acknowledgments

No artificial intelligence (AI)–based platform was used to develop the protocol paper.

Funding

This study was funded by the Children's Investment Fund Foundation (CIFF; 2312-09106). icddr,b acknowledges with gratitude the commitment of CIFF to its research efforts. icddr,b is also grateful to the Governments of Bangladesh and Canada for providing core/unrestricted support. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Authors' Contributions

SMTH, RD, AAN, and AH initiated the protocol and conceptualized the research plan for the proposed systematic review. SMTH, RD, AAN, and SM finalized the methodology. SMTH, RD, AAN, SG, and PC wrote the manuscript. AH, SM, STR, NK, DM, and TA critically reviewed it for important intellectual content. All authors read and approved the final manuscript. SMTH and RD contributed equally as first authors; TA and AH contributed equally as senior authors.

Conflicts of Interest

None declared.

Checklist 1
PRISMA-P checklist.
[PDF File, 146 KB - resprot_v15i1e76450_app1.pdf]

References

1. Hudu SA, Jimoh AO, Adeshina KA, Otalike EG, Tahir A, Hegazy AA. An insight into the success, challenges, and future perspectives of eliminating neglected tropical disease. Scientific African 2024 Jun;24:e02165. [doi: 10.1016/j.sciaf.2024.e02165]
2. Global report on neglected tropical diseases 2024. : World Health Organization; 2024 May 5 URL: https://iris.who.int/server/api/core/bitstreams/cb44ed69-02bd-4273-b599-26661646f5d3/content [accessed 2026-02-26]
3. Ending the neglect to attain the sustainable development goals: a road map for neglected tropical diseases 2021–2030. : World Health Organization; 2020 URL: https://www.who.int/publications/i/item/9789240010352 [accessed 2026-03-02]
4. Engels D, Zhou XN. Neglected tropical diseases: an effective global response to local poverty-related disease priorities. Infect Dis Poverty 2020 Jan 28;9(1):10. [doi: 10.1186/s40249-020-0630-9] [Medline: 31987053]
5. Hotez PJ, Fenwick A, Savioli L, Molyneux DH. Rescuing the bottom billion through control of neglected tropical diseases. Lancet 2009 May 2;373(9674):1570-1575. [doi: 10.1016/S0140-6736(09)60233-6] [Medline: 19410718]
6. The Causes and Impacts of Neglected Tropical and Zoonotic Diseases: Opportunities for Integrated Intervention Strategies: National Academies Press; 2011. [doi: 10.17226/13087] [Medline: 21977543]

XSL•FO
RenderX

7. Coady D. Alleviating structural poverty in developing countries: the approach of PROGRESA in Mexico. : World Bank Group; 2003 URL: https://documents1.worldbank.org/curated/en/435991468757217796/pdf/269430Coady.pdf [accessed 2026-02-26]
8. Integrating neglected tropical diseases into global health and development: fourth WHO report on neglected tropical diseases. : World Health Organization; 2017 URL: https://iris.who.int/server/api/core/bitstreams/903594da-27f0-44f1-a1e8-f4261ed810a5/content [accessed 2026-02-26]
9. Fürst T, Salari P, Llamas LM, Steinmann P, Fitzpatrick C, Tediosi F. Global health policy and neglected tropical diseases: then, now, and in the years to come. PLoS Negl Trop Dis 2017 Sep;11(9):e0005759. [doi: 10.1371/journal.pntd.0005759] [Medline: 28910281]
10. Patikorn C, Cho JY, Higashi J, Huang XX, Chaiyakunapruk N. Financial hardship among patients suffering from neglected tropical diseases: a systematic review and meta-analysis of global literature. PLOS Negl Trop Dis 2024 May;18(5):e0012086. [doi: 10.1371/journal.pntd.0012086] [Medline: 38739636]
11. Wagstaff A. Measuring financial protection in health. : World Bank Publications; 2008 URL: https://hdl.handle.net/10986/6570 [accessed 2026-02-26] [doi: 10.1596/1813-9450-4554]
12. Nguyen HA, Ahmed S, Turner HC. Overview of the main methods used for estimating catastrophic health expenditure. Cost Eff Resour Alloc 2023 Aug 8;21(1):50. [doi: 10.1186/s12962-023-00457-5] [Medline: 37553675]
13. Adhikari SR, Maskay NM, Sharma BP. Paying for hospital-based care of Kala-azar in Nepal: assessing catastrophic, impoverishment and economic consequences. Health Policy Plan 2009 Mar;24(2):129-139. [doi: 10.1093/heapol/czn052] [Medline: 19181674]
14. Immurana M, Kisseih KG, Abdullahi I, et al. The effects of selected neglected tropical diseases on economic performance at the macrolevel in Africa. BMC Infect Dis 2024 May 2;24(1):462. [doi: 10.1186/s12879-024-09302-3] [Medline: 38698313]
15. Magalhães AR, Codeço CT, Svenning JC, Escobar LE, Van de Vuurst P, Gonçalves-Souza T. Neglected tropical diseases risk correlates with poverty and early ecosystem destruction. Infect Dis Poverty 2023 Apr 10;12(1):32. [doi: 10.1186/s40249-023-01084-1] [Medline: 37038199]
16. Tickell KD, Walson JL. Nutritional enteric failure: neglected tropical diseases and childhood stunting. PLoS Negl Trop Dis 2016 Apr;10(4):e0004523. [doi: 10.1371/journal.pntd.0004523] [Medline: 27124296]
17. Ngui R, Lim YAL, Chong Kin L, Sek Chuen C, Jaffar S. Association between anaemia, iron deficiency anaemia, neglected parasitic infections and socioeconomic factors in rural children of West Malaysia. PLoS Negl Trop Dis 2012;6(3):e1550. [doi: 10.1371/journal.pntd.0001550] [Medline: 22413027]
18. Hall A, Zhang Y, Macarthur C, Baker S. The role of nutrition in integrated programs to control neglected tropical diseases. BMC Med 2012 Apr 25;10:41. [doi: 10.1186/1741-7015-10-41] [Medline: 22533927]
19. Quihui-Cota L, Valencia ME, Crompton DWT, et al. Prevalence and intensity of intestinal parasitic infections in relation to nutritional status in Mexican schoolchildren. Trans R Soc Trop Med Hyg 2004 Nov;98(11):653-659. [doi: 10.1016/j.trstmh.2003.12.017] [Medline: 15363645]
20. Pereira PCM. Interaction between infection, nutrition and immunity in tropical medicine. J Venom Anim Toxins incl Trop Dis 2003;9(2):163-173. [doi: 10.1590/S1678-91992003000200003]
21. Yap P, Utzinger J, Hattendorf J, Steinmann P. Influence of nutrition on infection and re-infection with soil-transmitted helminths: a systematic review. Parasit Vectors 2014 May 19;7(1):229. [doi: 10.1186/1756-3305-7-229] [Medline: 24885622]
22. Anyolitho MK, Nyakato VN, Huyse T, Poels K, Masquillier C. Health-seeking behaviour regarding schistosomiasis treatment in the absence of a mass drug administration (MDA) program: the case of endemic communities along Lake Albert in Western Uganda. BMC Public Health 2023 Jun 5;23(1):1072. [doi: 10.1186/s12889-023-16020-z] [Medline: 37277773]
23. Ackley C, Elsheikh M, Zaman S. Scoping review of neglected tropical disease interventions and health promotion: a framework for successful NTD interventions as evidenced by the literature. PLoS Negl Trop Dis 2021 Jul;15(7):e0009278. [doi: 10.1371/journal.pntd.0009278] [Medline: 34228729]
24. Johnston EA, Teague J, Graham JP. Challenges and opportunities associated with neglected tropical disease and water, sanitation and hygiene intersectoral integration programs. BMC Public Health 2015 Jun 11;15:547. [doi: 10.1186/s12889-015-1838-7] [Medline: 26062691]
25. Aamar H, Siddiqui JA, Siddiqui A, Essar MY. Neglected tropical diseases in Pakistan: challenges, efforts, and recommendations. Int J Surg 2023 Feb 1;109(2):143-144. [doi: 10.1097/JS9.0000000000000186] [Medline: 36799830]
26. Engels D, Elphick-Pooley T. The potential of conditional cash transfers for the control of neglected tropical diseases. Lancet Glob Health 2022 May;10(5):e588-e589. [doi: 10.1016/S2214-109X(22)00115-2] [Medline: 35427503]
27. Ramos AN Jr, Heukelbach J, Oliveira M. A conditional cash transfer programme in Brazil improves leprosy treatment outcomes. Lancet Infect Dis 2020 May;20(5):522-523. [doi: 10.1016/S1473-3099(19)30750-9] [Medline: 32066528]
28. Novignon J, Prencipe L, Molotsky A, et al. The impact of unconditional cash transfers on morbidity and health-seeking behaviour in Africa: evidence from Ghana, Malawi, Zambia and Zimbabwe. Health Policy Plan 2022 May 12;37(5):607-623. [doi: 10.1093/heapol/czac014] [Medline: 35157775]
29. Evans DK, Holtemeyer B, Kosec K. Cash transfers and health: evidence from Tanzania. World Bank Econ Rev 2019 Jun 1;33(2):394-412. [doi: 10.1093/wber/lhx001]

XSL•FO
RenderX

30. Gaarder MM, Glassman A, Todd JE. Conditional cash transfers and health: unpacking the causal chain. J Dev Effect 2010 Apr 13;2(1):6-50. [doi: 10.1080/19439341003646188]
31. Fiszbein A, Schady N, Ferreira FHG, et al. Conditional cash transfers: reducing present and future poverty. : The World Bank Group; 2009 URL: https://documents1.worldbank.org/curated/en/914561468314712643/pdf/476030PUB0Cond101Official0Use0Only1.pdf [accessed 2026-02-26]
32. Álvarez-Hernández DA, Rivero-Zambrano L, Martínez-Juárez LA, García-Rodríguez-Arana R. Overcoming the global burden of neglected tropical diseases. Ther Adv Infect Dis 2020;7:2049936120966449. [doi: 10.1177/2049936120966449] [Medline: 33178435]
33. Fernald LCH, Gertler PJ, Neufeld LM. Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades. Lancet 2008 Mar 8;371(9615):828-837. [doi: 10.1016/S0140-6736(08)60382-7] [Medline: 18328930]
34. Glassman A, Duran D, Fleisher L, et al. Impact of conditional cash transfers on maternal and newborn health. J Health Popul Nutr 2013 Dec;31(4 Suppl 2):48-66. [Medline: 24992803]
35. Ahmed A, Aune D, Vineis P, Pescarini JM, Millett C, Hone T. The effect of conditional cash transfers on the control of neglected tropical disease: a systematic review. Lancet Glob Health 2022 May;10(5):e640-e648. [doi: 10.1016/S2214-109X(22)00065-1] [Medline: 35427521]
36. Hadi A. Promoting health knowledge through micro-credit programmes: experience of BRAC in Bangladesh. Health Promot Int 2001 Sep;16(3):219-227. [doi: 10.1093/heapro/16.3.219] [Medline: 11509457]
37. McFadden A, Green JM, Williams V, et al. Can food vouchers improve nutrition and reduce health inequalities in low-income mothers and young children: a multi-method evaluation of the experiences of beneficiaries and practitioners of the Healthy Start programme in England. BMC Public Health 2014 Feb 11;14:148. [doi: 10.1186/1471-2458-14-148] [Medline: 24517737]
38. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
39. Tacconelli E. Systematic reviews: CRD's guidance for undertaking reviews in health care. Lancet Infect Dis 2010 Apr;10(4):226. [doi: 10.1016/S1473-3099(10)70065-7]
40. Baird S, Ferreira FHG, Özler B, Woolcock M. Relative effectiveness of conditional and unconditional cash transfers for schooling outcomes in developing countries: a systematic review. Campbell Syst Rev 2013 Jan;9(1):1-124. [doi: 10.4073/csr.2013.8]
41. Banerjee A, Karlan D, Zinman J. Six randomized evaluations of microcredit: introduction and further steps. Am Econ J Appl Econ 2015 Jan 1;7(1):1-21. [doi: 10.1257/app.20140287]
42. Cunha JM, De Giorgi G, Jayachandran S. The price effects of cash versus in-kind transfers. Rev Econ Stud 2019 Jan 1;86(1):240-281. [doi: 10.1093/restud/rdy018]
43. Hoefle-Bénard J, Salloch S. Mass drug administration for neglected tropical disease control and elimination: a systematic review of ethical reasons. BMJ Glob Health 2024 Mar 13;9(3):e013439. [doi: 10.1136/bmjgh-2023-013439] [Medline: 38485140]
44. Mitra AK, Mawson AR. Neglected tropical diseases: epidemiology and global burden. Trop Med Infect Dis 2017 Aug 5;2(3):36. [doi: 10.3390/tropicalmed2030036] [Medline: 30270893]
45. WHO to publish first official guidelines on leprosy diagnosis, treatment and prevention [internet]. World Health Organization. 2018. URL: https://www.who.int/news/item/28-06-2018-who-to-publish-first-official-guidelines-on-leprosy-diagnosis-treatment-and-prevention [accessed 2024-07-29]
46. Sharma P, Mukherjee R, Talwar GP, et al. Immunoprophylactic effects of the anti-leprosy Mw vaccine in household contacts of leprosy patients: clinical field trials with a follow up of 8-10 years. Lepr Rev 2005 Jun;76(2):127-143. [Medline: 16038246]
47. Lwin K, Sundaresan T, Gyi MM, et al. BCG vaccination of children against leprosy: fourteen-year findings of the trial in Burma. Bull World Health Organ 1985;63(6):1069-1078. [Medline: 2940028]
48. Macfarlane CL, Budhathoki SS, Johnson S, Richardson M, Garner P. Albendazole alone or in combination with microfilaricidal drugs for lymphatic filariasis. Cochrane Database Syst Rev 2019 Jan 8;1(1):CD003753. [doi: 10.1002/14651858.CD003753.pub4] [Medline: 30620051]
49. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
50. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
51. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
52. CASP checklist: CASP qualitative studies checklist. CASP. 2018. URL: https://casp-uk.net/casp-tools-checklists/qualitative-studies-checklist/ [accessed 2026-02-26]
53. Jacobs W, Downey LE. Impact of conditional cash transfer programmes on antenatal care service uptake in low and middle-income countries: a systematic review. BMJ Open 2022 Nov 25;12(11):e064673. [doi: 10.1136/bmjopen-2022-064673] [Medline: 36428017]

XSL•FO
RenderX

54. Hunter BM, Harrison S, Portela A, Bick D. The effects of cash transfers and vouchers on the use and quality of maternity care services: a systematic review. PLoS One 2017;12(3):e0173068. [doi: 10.1371/journal.pone.0173068] [Medline: 28328940]
55. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021 Mar 29;10(1):89. [doi: 10.1186/s13643-021-01626-4] [Medline: 33781348]

Abbreviations

CASP: Critical Appraisal Skills Programme
CCT: conditional cash transfer
GRADE: Grading of Recommendations, Assessment, Development and Evaluation
JBI: Joanna Briggs Institute
LMICs: low- and middle-income countries
NTD: neglected tropical disease
PICOS: Population, Interventions, Comparators, Outcomes, Study Design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
RoB 2: Revised Cochrane Risk of Bias for randomized trials
ROBINS-I: Risk of Bias in Non-Randomized Studies of Interventions
WHO: World Health Organization

Edited by J Sarvestan; submitted 04.May.2025; peer-reviewed by A Olanlesi-Aliu; revised version received 25.Jan.2026; accepted 30.Jan.2026; published 19.Mar.2026.

Please cite as:
Hasan SMT, Dayal R, Nishan AA, Ghatak S, Chakraborty P, Maitra S, Raisa ST, Khan N, Mondal D, Ahmed T, Hazra A
Effects of Cash Transfer and Incentive Programs on Service Utilization and Treatment Outcomes Related to Neglected Tropical Diseases and Their Impact on Health and Nutrition in Low- and Middle-Income Countries: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e76450
URL: https://www.researchprotocols.org/2026/1/e76450
doi:10.2196/76450

© S M Tafsir Hasan, Radhika Dayal, Amena Al Nishan, Sunetra Ghatak, Puja Chakraborty, Sudeshna Maitra, Sumaiya Tasneem Raisa, Nizamuddin Khan, Dinesh Mondal, Tahmeed Ahmed, Avishek Hazra. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Enhancing Fracture Resistance by Customizing Glass Fiber Posts in Endodontically Treated Teeth: Protocol for a Systematic Review

Sanika Damle[1], MDS; Manoj Chandak[1], MDS; Unnati Shirbhate[2], MDS; Pavan Bajaj[2], MDS

[1]Department of Conservative Dentistry and Endodontics, Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

[2]Department of Periodontics and Oral Implantology, Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, Wardha, Maharashtra, India

Corresponding Author:
Sanika Damle, MDS
Department of Conservative Dentistry and Endodontics, Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Abstract

Background: The loss of tooth structure in endodontically treated teeth compromises their structural integrity and increases their vulnerability to fractures. To strengthen these teeth, post and core systems must be used. Among post materials, glass fiber has become more and more common because of it advantageous mechanical and aesthetic qualities. However, the effectiveness of customized glass fiber posts in enhancing fracture resistance compared to prefabricated posts remains a subject of ongoing debate.

Objective: This systematic review primarily focuses on the fracture resistance of endodontically treated teeth reinforced with customized glass fiber posts. It aims to assess how well they function in terms of resilience, flexibility, and failure patterns in comparison to prefabricated fiber posts and other post materials.

Methods: To locate relevant studies published up to 2025, a comprehensive literature search will be conducted through various materials available electronically on websites such as Google Scholar, Web of Science, PubMed, Semantic Scholar, and Scopus. The selection criteria will include randomized controlled clinical trials that evaluate fracture resistance of customized glass fiber posts in teeth that have had endodontic treatment. The research will be screened using predetermined standards for inclusion and exclusion. Quality assessment and collection of data will be performed strictly, in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines. To examine the results quantitatively, a meta-analysis will be conducted, if at all possible, with a random effects model for heterogeneity. Statistical methods will be used for evaluating the effectiveness of custom glass fiber posts, alternative post materials, and prefabricated posts.

Results: By comparing the mean differences in fracture resistance of teeth treated endodontically using various post systems, the effectiveness of customized glass fiber posts will be assessed. Other aspects to be examined include general biomechanical performance, failure mechanisms, and stress distribution. This study will use a random effects model to estimate the combined effect size measurements and the corresponding 95% CIs. We anticipate that the data synthesis will be conducted between February and March 2026 for this systematic review and will be finished by April or July 2026.

Conclusions: Customized glass fiber posts may show promise in strengthening teeth that have undergone endodontic treatment by lowering the risk of failure and increasing fracture resistance. Even if there is current evidence that they are more effective than prefabricated posts, we need to synthesize the quality of the evidence. Therefore, this review aims to evaluate whether customized glass fiber posts effectively enhance fracture resistance in endodontically treated teeth.

Trial Registration: PROSPERO CRD420250651870; https://www.crd.york.ac.uk/PROSPERO/view/CRD420250651870

International Registered Report Identifier (IRRID): PRR1-10.2196/76027

(JMIR Res Protoc 2026;15:e76027) doi:10.2196/76027

KEYWORDS

root canal; fracture resistance; cemented posts; custom made fiber posts; endodontic therapy

XSL•FO
RenderX

Introduction

Rationale

Endodontics, one of the major fields in dentistry, focuses on the causes, prevention, diagnosis, and management of conditions and traumas affecting the dental pulp and periapical tissues [1]. Today, the emphasis of dental therapy has switched to a more conservative approach, although in the past, extraction was the recommended treatment for the majority of teeth that were highly carious. When the vital contents of the canal are removed by endodontic therapy, the tooth becomes pulpless and has calcified tissues with much lower moisture content than living teeth [2]. Because of these alterations, the tooth becomes weaker and more prone to fracture. To restore the strength and integrity of such a tooth, post and core systems are commonly used [3].

Typically, a post is used to increase core retention, which increases crown retention [4]. Over the past 10 years, the most frequently used post types in dentistry have been cast posts and cores. These typically need an extra laboratory step where the impressions taken from the prepared post area are used to create a custom post [3]. Despite the widespread use of cast or prefabricated metallic posts, there are certain issues that have been identified with these systems, such as corrosion, root fractures, loss of retention, the need to remove considerable root structure, and concentration of stress [5].

Fiber-reinforced composite (FRC) posts are a popular option for already endodontically treated teeth. A milestone was reached with the introduction of carbon fiber–reinforced resin posts, and this changed some of the underlying concepts used to restore previously endodontically treated teeth. Glass fiber posts are now most often used to improve the distribution of masticatory stresses and guarantee greater retention and support for restorative material, leading to better recovery of severely injured teeth. Prefabricated posts and specially customized chair-side posts are the two main types of FRC posts [6].

Prefabricated FRC posts are made by inserting a great volume of continuous, 1-directional, reinforcing fibers into a polymer matrix. Usually composed of materials like carbon, glass, or quartz fibers, these posts require little preparation before being placed. Conversely, non-preimpregnated fibers, usually made of glass or polyethylene, are shaped into posts during the restoration process to make chair-side customized FRC posts [7]. To treat teeth severely affected by cavities or trauma, custom-made posts and cores are sometimes advised. Because they fit precisely in the prepared post space, they are especially helpful for teeth with curved or elliptical canals [8].

Review Question

Is there any role for customized glass fiber posts in fracture resistance for endodontically treated teeth?

Objective

We aimed to evaluate the fracture resistance, mechanical stability, and long-term durability of customized glass fiber posts in endodontically treated teeth and to compare their effectiveness with traditional post systems.

Methods

Eligibility Criteria

Studies will be selected based on predefined inclusion and exclusion criteria to ensure consistency and reproducibility in the review process (Textbox 1).

Textbox 1. Inclusion and exclusion criteria

Inclusion criteria

- Study design: randomized controlled trials, controlled clinical trials, and in vitro experimental studies evaluating fracture resistance
- Participants: patients with endodontically treated teeth
- Intervention: customized glass fiber posts for postendodontic restoration
- Comparison: any conventional post system or no intervention
- Outcome: fracture resistance of the restored teeth
- Language: publications in English

Exclusion criteria

- Study design: case reports, reviews, editorials, conference abstracts, and letters to the editor
- Participants: studies involving teeth that were not endodontically treated

Information Sources

To find papers that meet the review's eligibility criteria, electronically available databases will be searched, including Web of Science, PubMed, Cochrane Library, Google Scholar, Semantic Scholar, and Scopus.

Search Strategy

To locate relevant papers, a limited search of MEDLINE will be conducted. Using terms from the abstracts, titles, and keywords, a thorough search strategy will be created. The following search strategy will be followed: "endodontically treated teeth"[Title/Abstract] OR "endodontically treated tooth"[Title/Abstract] OR "root canal treated teeth"[Title/Abstract] OR "root canal treated tooth"[Title/Abstract] OR "endodontic

XSL•FO
RenderX

treatment*"[Title/Abstract] OR "RCT"[Title/Abstract] OR "root canal treatment*"[Title/Abstract] OR "root canal therap*"[Title/Abstract] OR "endodontic therap*"[Title/Abstract] OR "root canal therapy"[MeSH Terms] AND "customized glass fiber post*"[Title/Abstract] OR "custom made fiber post*"[Title/Abstract] OR "custom-made fiber post*"[Title/Abstract] OR "glass fiber post*"[Title/Abstract] OR "glass fiber post*"[Title/Abstract] OR "endodontic post*"[Title/Abstract].

Study Selection, Data Management, and Data Collection Process

The three reviewers (SD, US, PB) will consult the eligibility criteria when deciding which research papers to include. Recording the selection will help with duplicate removal and record management. Studies will be screened blindly, and disagreements will be resolved through discussion. After loading the retrieved studies, 2 researchers will screen the abstracts and titles. A full-text review screening will be conducted by 2 authors, and any disagreements will be solved by a fourth reviewer (MC).

A data extraction tool used on the included papers will serve as a guide for the reviewers. The randomized controlled trials will include individuals who are in good systemic health as well as those undergoing endodontic treatment. Pregnant women and those with a history of systemic sickness will be excluded. The findings obtained will be narratively summarized. Methodological, statistical, and clinical heterogeneity will be examined and evaluated. If all of the incorporated studies and their findings are found to be sufficiently homogeneous, a quantitative synthesis, also known as a meta-analysis, will be conducted.

Data Items

The extracted data will be duplicated for each of the 2 blinded reviewers (US and SD). A Microsoft Excel spreadsheet will be used to store the extracted data, which will contain publication information, the setting of the research study, the fiber post, push-out bond strength, and luting cement. Customized glass fiber posts, fracture resistance, and endodontic clinical parameters will all be considered as clinical determinants.

Outcomes and Prioritization

The primary outcome is the fracture resistance of teeth that have previously received endodontic treatment and restoration with customized glass fiber posts, measured by fracture load or strength. Secondary outcomes include the failure mode (adhesive, cohesive, or mixed failure), the strength of the bond between the post and dentin, and long-term durability under functional stress. These outcomes will help assess the mechanical stability and clinical efficacy of customized posts compared to conventional systems.

Risk of Bias

To assess the methodological quality and risk of bias of the included studies, standardized and validated tools will be use. For randomized controlled trials, the Risk of Bias 2 (RoB 2) tool, developed by the Cochrane Collaboration, will be used. This tool evaluates five domains: (1) bias arising from the randomization process, (2) bias due to deviations from intended interventions, (3) bias due to missing outcome data, (4) bias in measurement of the outcome, and (5) bias in selection of the reported results. Each domain will be rated as "low risk," "some concerns," or "high risk," and an overall risk of bias judgment will be made accordingly.

If any in vitro studies are included in the review, they will be assessed using the QUIN (Quality Assessment Tool for In Vitro Studies), which considers methodological rigor, sample size justification, standardization of procedures, and statistical analysis, among other relevant criteria.

The risk of bias assessments will be independently conducted by 2 reviewers (US and SD). In case of disagreement, third and fourth reviewers (PB and MC) will act as arbiters to reach consensus. To minimize bias, all reviewers will remain blinded to each other's assessments during the evaluation process.

Data Synthesis and Assessment of Heterogeneity

A random effects model will be used to account for potential heterogeneity across studies. Statistical heterogeneity will be assessed using the I^2 statistic and the chi-square test. If the studies are deemed sufficiently homogeneous in terms of design, population, and outcome measures, a quantitative synthesis (meta-analysis) will be performed.

Effect sizes will be reported as standardized mean differences for continuous outcomes (eg, fracture resistance values) with 95% CIs. If binary outcomes are present, odds ratios will be calculated.

All analyses will be conducted using appropriate statistical software, such as RevMan (Review Manager; version 5.4) or Stata (version 17.0, StataCorp LLC). This approach adheres to the guidelines outlined in the *Cochrane Handbook for Systematic Reviews of Interventions* [9] to ensure methodological transparency and reproducibility.

Results

Identification, screening, and inclusion of studies is expected to be completed in May of 2026. At every stage of the process, the 2 reviewers were blinded. Team meetings were conducted on a regular basis to settle disputes and resolve conflicts between reviewers, increasing the transparency of the evaluation procedure. Every discussion was documented and recorded. When the systematic review is complete, it will be submitted for publication in accordance with the recommendations in the Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) checklist. Figure 1 shows the PRISMA flowchart for this review.

As the literature search and screening processes were completed in May 2025, preliminary data on the number of studies identified, screened, and included have been documented and are presented in the PRISMA flowchart in Figure 1. These numbers are subject to change as full-text screening and risk of bias assessment progress.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) flowchart.

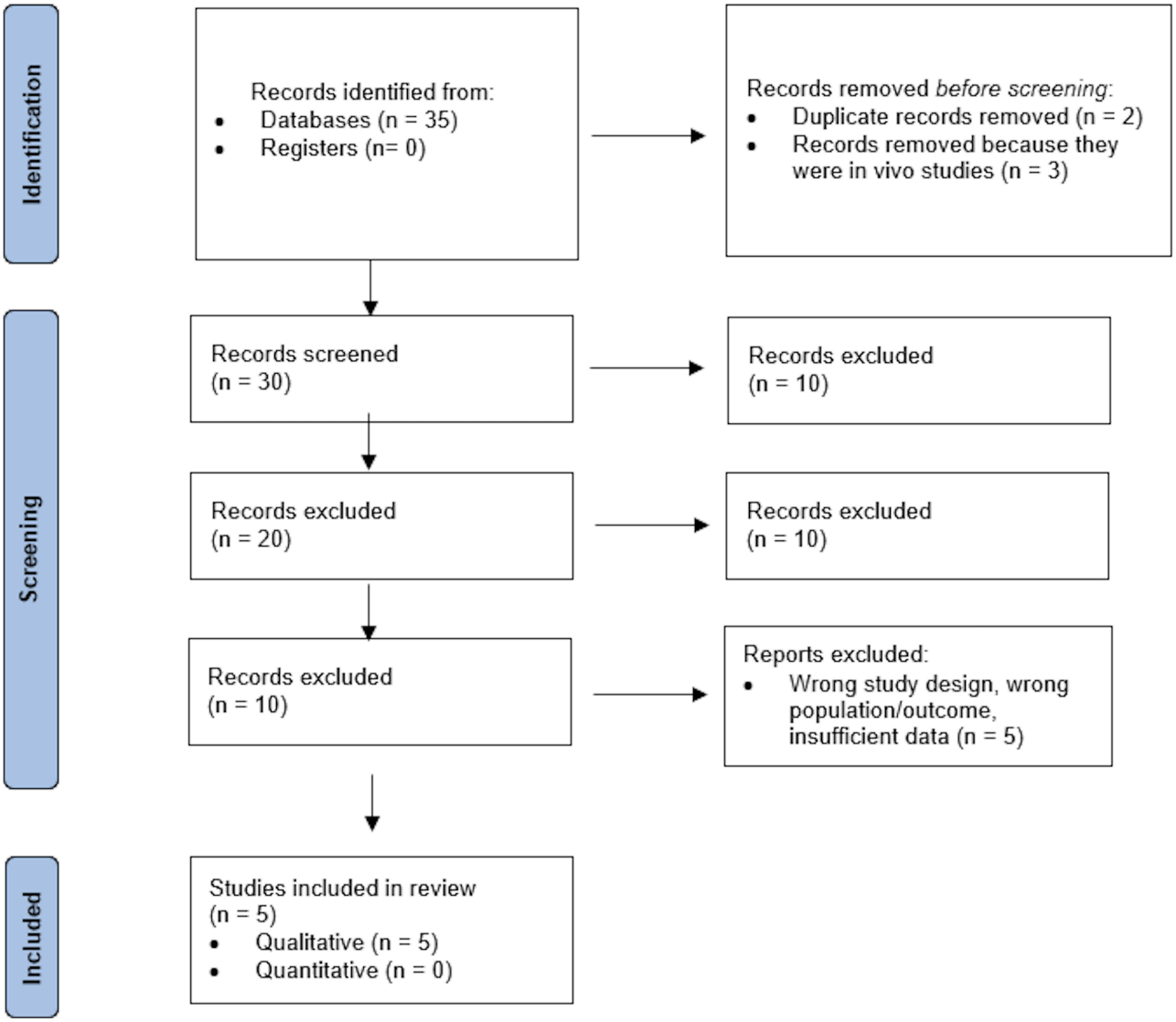

Discussion

Anticipated Findings

This systematic review will assess the effectiveness of customized glass fiber posts for fracture resistance in endodontically treated teeth. Because of their improved stress distribution and increased adaptation to the root canal shape, we anticipate that customized glass fiber posts will offer better fracture resistance in comparison with prefabricated ones [10] due to greater retention and less stress concentration, resulting from the customized posts' superior fit to the canal walls. Customized posts enable a more conservative approach, maintaining the tooth's structural integrity, in contrast to prefabricated fiber posts, which frequently necessitate significant dentin removal to fit the canal [11].

The intrinsic properties of glass fiber posts, such as their elasticity modulus, which is the same as dentin, contribute to improved stress distribution and a reduced incidence of catastrophic root fractures. Customization further enhances these properties by eliminating gaps between the post and canal walls, minimizing microleakage and improving the overall bonding interface with the adhesive cement [5]. We anticipate that the findings of this review will show the importance of post customization in achieving optimal fracture resistance in endodontically treated teeth. Clinicians should consider factors such as canal morphology, residual dentin thickness, and adhesive bonding techniques when selecting and customizing posts. The use of CAD/CAM technology or direct fiber post buildup techniques can further enhance the fit and mechanical performance of customized posts [7].

While the current evidence supports the superiority of customized glass fiber posts, some limitations should be acknowledged. The heterogeneity of study designs, sample sizes, and testing methodologies may impact the generalizability of the results [12]. Additionally, a sufficient number of clinical studies are required for assessing the outcomes of customized posts in functional conditions. Customized glass fiber posts offer significant advantages over prefabricated posts in terms of fracture resistance, biomechanical compatibility, and adhesive bonding. Their ability to preserve dentin and enhance stress distribution makes them a favorable choice for restoring endodontically treated teeth. Further clinical investigations are

XSL•FO
RenderX

required to establish standardized protocols for their fabrication and long-term effectiveness [13,14].

Strengths and Limitations

Strengths of this systematic review include the clinical relevance of an evaluation of customized glass fiber posts for enhancing fracture resistance and mechanical stability in endodontically treated teeth, as well as comprehensive outcome measures, such as mechanical performance, fracture patterns, fracture resistance, and failure modes. The use of multiple high-quality databases and standardized tools for risk of bias assessment (RoB 2 and QUIN) to improve methodological rigor is also a strength.

Limitations include the inclusion of in vitro studies, which may not accurately replicate clinical conditions; potential heterogeneity in study design, materials, and testing protocols across included studies; and the limited availability of long-term clinical data. Additionally, while gray literature sources like Google Scholar will be used (with caution), there remains a risk of publication bias and variable study quality.

Conclusion and Directions for Dissemination

In summary, this systematic review aims to ascertain the effectiveness of customized glass fiber posts for fracture resistance in teeth that have been treated endodontically. However, study variability as well as a lack of long-term clinical data may limit the findings, emphasizing the need for more research to confirm the findings and their clinical relevance. This review's findings will be published in scholarly journals both domestically and abroad. The PRISMA-P criteria will be followed in this review.

Acknowledgments

The databases and resources required for the data synthesis, as well as the assistance of the Datta Meghe Institute of Higher Education and Research, are acknowledged by all authors.

Funding

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Authors' Contributions

The manuscript was drafted and written with input from all contributors. The work was conceptualized and registered by SD and MC. The study design and preliminary paper preparation were completed by US. SD completed the final study design and manuscript writing. Each author made an equal contribution to the effort and reviewed the completed paper.

Conflicts of Interest

None declared.

References

1. Levine M. Root-canal therapy: a means of treating oral pain and infection. Can Fam Physician 1988 Jun;34:1357-1365. [Medline: 21253195]
2. Darvishi F, Oskoee PA, Chaharom MEE, Ajami AA, Kimyai S. Fracture resistance of over flared endodontically treated central incisors restored with multiple prefabricated and custom-made glass fiber posts. Front Dent 2020;17:37. [doi: 10.18502/fid.v17i37.5290] [Medline: 36042794]
3. Adanir N, Belli S. Evaluation of different post lengths' effect on fracture resistance of a glass fiber post system. Eur J Dent 2008 Jan;2(1):23-28. [Medline: 19212505]
4. Heydecke G, Peters MC. The restoration of endodontically treated, single-rooted teeth with cast or direct posts and cores: a systematic review. J Prosthet Dent 2002 Apr;87(4):380-386. [doi: 10.1067/mpr.2002.123848] [Medline: 12011847]
5. Fontana PE, Bohrer TC, Wandscher VF, Valandro LF, Limberger IF, Kaizer OB. Effect of ferrule thickness on fracture resistance of teeth restored with a glass fiber post or cast post. Oper Dent 2019;44(6):E299-E308. [doi: 10.2341/18-241-L] [Medline: 31283420]
6. Silva CF, Martins VM, Melo ADP, Martins LC, Santos-Filho PCF. The use of bulk-fill flow in the customization of glass fiber post. Eur J Dent 2021 Feb;15(1):139-144. [doi: 10.1055/s-0040-1716599] [Medline: 32932532]
7. Schmitter M, Rammelsberg P, Gabbert O, Ohlmann B. Influence of clinical baseline findings on the survival of 2 post systems: a randomized clinical trial. Int J Prosthodont 2007;20(2):173-178. [Medline: 17455439]
8. Godas AGL, Suzuki TYU, Oliveira-Reis B, Briso ALF, Assunção WG, Dos Santos PH. Effect of glass fiber post customization on the mechanical properties of resin cement and underlying dentin. Gen Dent 2020;68(1):72-77. [Medline: 31859667]
9. Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Collaboration. URL: https://www.cochrane.org/authors/handbooks-and-manuals/handbook [accessed 2026-01-12]

XSL•FO
RenderX

10. Purton DG, Love RM. Rigidity and retention of carbon fibre versus stainless steel root canal posts. Int Endod J 1996 Jul;29(4):262-265. [doi: 10.1111/j.1365-2591.1996.tb01379.x] [Medline: 9206443]
11. Wang X, Shu X, Zhang Y, Yang B, Jian Y, Zhao K. Evaluation of fiber posts vs metal posts for restoring severely damaged endodontically treated teeth: a systematic review and meta-analysis. Quintessence Int 2019;50(1):8-20. [doi: 10.3290/j.qi.a41499] [Medline: 30600326]
12. Ruttonji ZR, Kusugal PB, Nayak A, Mahajan D, Sushma KM, Patil VN. Comparative evaluation of the effect of surface treatment of fiber-reinforced posts and prefabricated metal posts on adhesion of a resin-based luting cement: an in vitro study. J Conserv Dent 2019;22(3):245-248. [doi: 10.4103/JCD.JCD_402_18] [Medline: 31367107]
13. Mohamed A, Ismail E, Adam RZ. A comparison of conventional root canal sealers with ones that use green synthesized nanoparticles for antimicrobial activity: protocol for a systematic review. JMIR Res Protoc 2024 Oct 11;13:e51351. [doi: 10.2196/51351] [Medline: 39393053]
14. Nelson S, Albert JM, Milgrom P. Comparative effectiveness of two nonsurgical treatments to reduce oral health disparities from untreated tooth decay in older adults: protocol for a cluster randomized trial. JMIR Res Protoc 2020 Sep 8;9(9):e17840. [doi: 10.2196/17840] [Medline: 32897236]

Abbreviations

FRC: fiber-reinforced composite
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
QUIN: Quality Assessment Tool for In Vitro Studies
RoB 2: Risk of Bias 2

Edited by J Sarvestan; submitted 15.Apr.2025; peer-reviewed by P Dhawan; revised version received 25.Jul.2025; accepted 29.Jul.2025; published 17.Feb.2026.

Please cite as:
Damle S, Chandak M, Shirbhate U, Bajaj P
Enhancing Fracture Resistance by Customizing Glass Fiber Posts in Endodontically Treated Teeth: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e76027
URL: https://www.researchprotocols.org/2026/1/e76027
doi:10.2196/76027

© Sanika Damle, Manoj Chandak, Unnati Shirbhate, Pavan Bajaj. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Diabetic Dyslipidemia and Its Determinants Among People With Diabetes in South Africa: Protocol for a Systematic Review and Meta-Analysis

Mashudu Nemukula[1,2*], MSc; Fentahun Adane[3*], PhD; Arun Kumar Malaisamy[2*], MSc; Neel Sarovar Bhavesh[2*], PhD; Olebogeng Harold Majane[4], PhD; Sechene Stanley Gololo[1*], PhD

[1]Department of Biochemistry and Biotechnology, School of Science and Technology, Sefako Makgatho Health Sciences University, 51 Molotlegi Street, Pretoria, South Africa

[2]Transcription Regulation Group, Integrative Biology, International Centre for Genetic Engineering and Biotechnology, New Delhi, India

[3]Department of Biomedical Sciences, School of Medicine, Debre Markos University, Debre Markos, Ethiopia

[4]Department of Human Physiology, School of Medicine, Sefako Makgatho Health Scienes University, Pretoria, South Africa

[*]these authors contributed equally

Corresponding Author:
Mashudu Nemukula, MSc
Department of Biochemistry and Biotechnology, School of Science and Technology, Sefako Makgatho Health Sciences University, 51 Molotlegi Street, Pretoria, South Africa

Abstract

Background: Diabetic dyslipidemia (DD), characterized by a classical triad of abnormal lipid profiles among the diabetic population, presents a major public health concern in South Africa, particularly among Black South Africans. The increasing prevalence of DD significantly contributes to the development of atherosclerotic cardiovascular disease. With the incidence of diabetes rising from 4.5% in 2010 to 12.7% in 2021, urgent preventive measures and effective treatments are crucial to tackle the risk of premature mortality.

Objective: This systematic review and meta-analysis protocol aims to examine the existing literature on DD, providing an understanding of its prevalence and associated predictors among the diabetic population in South Africa, with the intention of informing more effective clinical and public health interventions.

Methods: The protocol is registered in PROSPERO (International Prospective Register of Systematic Reviews) and will adhere to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The available literature on DD will be systematically searched in common scholarly databases and reviewed accordingly. All published and unpublished studies conducted in South Africa prior to 2024 and written in English will be included. Two members (MN and FA) of the review team will independently screen the studies identified through the database search and assess risk of bias using the revised JBI critical appraisal tools. The review will integrate both quantitative and qualitative data synthesis. Results from both qualitative and quantitative data synthesis will be presented through forest plots, subgroup forest plots, and summary tables, which will present findings on pooled prevalence, odds ratios for predictors, heterogeneity statistics, and sensitivity analyses.

Results: The protocol was finalized in January 2025. The literature search was conducted between October 2024 and March 2025. Title and abstract screening began in April 2025, and full-text review was completed by July 2025, with data extraction scheduled for completion by September 2025. The completion of statistical analyses is expected by October 2025. We anticipate submission of the completed systematic review and meta-analysis for publication in December 2025.

Conclusions: The findings of the study protocol will inform the design of targeted interventions and policies aimed at advancing the management of DD and subsequently reducing the increased risk of atherosclerotic cardiovascular disease among the diabetic population.

Trial Registration: PROSPERO CRD42024614221; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024614221

International Registered Report Identifier (IRRID): PRR1-10.2196/82716

(JMIR Res Protoc 2026;15:e82716) doi:10.2196/82716

KEYWORDS

diabetic dyslipidemia; systematic review; meta-analysis; diabetes; PRISMA

XSL•FO
RenderX

Introduction

Diabetic dyslipidemia (DD) is one of the major public health concerns, particularly in patients with diabetes. It is an established determinant for the development of atherosclerotic cardiovascular disease (ACVD) and contributes to increased rates of morbidity and mortality. Despite international efforts to improve the management of diabetes, DD remains a significant contributor to the burden of cardiovascular morbidity, especially in developing countries [1,2].

The prevalence of DD is high in developing countries, where the burden of diabetes is on the rise [3]. In South Africa, DD is a growing concern, with several studies indicating that a significant proportion of individuals with diabetes have lipid abnormalities, namely hypertriglyceridemia, decreased levels of high-density lipoprotein cholesterol (HDL-C), and elevated levels of low-density lipoprotein cholesterol (LDL-C). This lipid profile aggravates the risk of developing ACVD, which is emerging as the leading cause of preventable mortality in the country [4-6].

In South Africa, diabetes and its associated metabolic disorders, including lipid abnormalities, are increasingly prevalent among certain ethnic groups, with a notably higher prevalence seen among Black South Africans [7,8]. According to the International Diabetes Federation, the prevalence of diabetes in South Africa has notably increased from 4.5% in 2010 to 12.7% in 2019, with an estimated 4.58 million people aged 20 - 70 years affected, of whom 52.4% are undiagnosed [9,10]. Moreover, South Africa has the second-highest number of people living with type 2 diabetes mellitus in sub-Saharan Africa [11]. These metabolic disorders are significant determinants of the national burden of ACVD and premature mortality, with diabetes-related cardiovascular death predicted to increase significantly in the coming years if preventive measures and effective treatment strategies are not implemented [12-14].

Despite the ongoing public health initiatives and policies aimed at addressing diabetes and the determinants of cardiovascular disease, the prevalence of DD continues to rise, particularly among young individuals and those with poor glycemic control [15,16]. Existing studies have shown that DD is prevalent among diabetic patients who have suboptimal glycemic control and engage in unhealthy lifestyle choices, such as poor dietary habits and sedentary lifestyles [17,18]. Therefore, there is a need for updated studies to better understand the epidemiology of DD, particularly in the context of South Africa, to inform targeted intervention strategies and improve patient outcomes.

Predictors of DD include sociodemographic factors such as age, gender, ethnicity, education level, and household income, as well as the presence of other comorbid conditions, including hypertension, obesity, and chronic kidney disease [19,20]. Moreover, psychological factors such as stress, as well as sedentary lifestyle factors, contribute to the development and progression of DD among individuals with diabetes mellitus [21]. Understanding these predictors is important for developing more effective interventions to manage lipid abnormalities in diabetes mellitus.

The aim of this systematic review and meta-analysis is to estimate the pooled prevalence of DD among the diabetic population of South Africa and to assess the association between DD and its predictors. The review intends to provide an understanding of how certain predictors influence lipid profiles in diabetes mellitus. In addition, the review also seeks to inform targeted interventions aimed at reducing the incidence of DD and its associated cardiovascular risks in South Africa.

Methods

Overall Approach

The systematic review will be conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [22] on studies reporting on DD among people with diabetes in South Africa. Following this approach will ensure a transparent and systematic process for identifying, selecting, and including relevant studies. The protocol is registered with PROSPERO (International Prospective Register of Systematic Reviews; registration number CRD42024614221).

Population of Interest

The population of interest will include South Africans diagnosed with diabetes mellitus, including both type 1 and type 2 diabetes. The emphasis is on the diabetic population presenting with DD, characterized by lipid abnormalities such as hypertriglyceridemia, reduced HDL-C, and elevated LDL-C. Nondiabetic individuals will not be included in the review.

Eligibility Criteria

Studies reporting data on DD (hypertriglyceridemia, low HDL-C, and elevated LDL-C) among South Africans with diabetes mellitus will be considered. Observational studies such as cross-sectional, cohort, and case-control studies will be included. Studies reporting data on the prevalence or predictors of diabetic dyslipidemia will be included. Studies conducted among non-South Africans, including those involving nondiabetic populations, will be excluded. Reviews, case reports, editorials, and conference abstracts reporting on diabetic dyslipidemia will be disregarded. Studies without relevant lipid profile data will also be excluded.

Search Strategy

Studies for this systematic review and meta-analysis will be accessed through electronic and other relevant sources. We intend to use common database searches, such as PubMed/MEDLINE, Embase, Scopus, Web of Science, ScienceDirect, African Index Medicus, and Google Scholar, as well as local databases like African Journals Online and SABINET, to identify relevant articles. In addition, we will search publicly accessible institutional repositories such as open-access libraries from the university community, which include publicly available theses and dissertations, for identification of eligible studies. Restricted-access repositories and private databases will not be used for identifying eligible studies. Only publicly accessible sources will be consulted, ensuring full compliance with institutional privacy and data protection. Our search will include all published and unpublished

XSL•FO
RenderX

studies conducted in South Africa before December 2024 and written in the English language. The searches will be conducted using Medical Subject Headings (MeSH) terms, combined key terms, text words, and search strings taken from the review questions. To access the records, the following combination key terms will be used: ("Diabetic dyslipidemia" AND "lipid profile" OR "HDL-C" OR "LDL-C" OR "Triglycerides") AND ("Diabetes mellitus" OR "T1DM" OR "T2DM") AND ("predictors" OR "risk factors" OR "determinants" OR "contributing factors") AND ("South Africa"). After identifying the key relevant studies, their references will also be searched (backward reference searching). Similarly, other studies that cited them will be retrieved (forward reference searching).

Article Screening and Data Extraction

Two members of the review team (MN and FA) will independently screen the titles and abstracts of all studies identified through the database search. Prior to screening, all identified citations will be exported to Microsoft Excel (Microsoft Corp) for duplicate identification and removal. Duplicate records will be identified using sorting functions, conditional formatting, and unique reference identifiers to highlight repeated entries. A manual verification step will be performed to ensure complete and accurate removal of duplicates before title and abstract screening. Prior to the initiation of full screening, MN and FA, the two reviewers, will conduct a pilot assessment phase using a random sample of approximately 10% - 20% of the retrieved studies to enhance and harmonize the application of the eligibility criteria. Interrater agreement will be evaluated using the Cohen κ statistic, and a threshold of $\kappa \geq 0.70$ will be considered acceptable for proceeding with full screening. Any disagreements arising during the pilot phase will be reviewed and discussed between the two reviewers to ensure a consistent interpretation of the criteria. Following the pilot assessment phase, a full-text review will be conducted for studies that appear to meet the inclusion criteria based on their title and abstract. Full-text articles will then be evaluated to determine whether they meet the predefined inclusion and exclusion criteria. Articles reporting on the South African population with diabetes mellitus and DD markers and outcomes will be included. Studies not involving South African populations, as well as nondiabetic populations or those with gestational diabetes or prediabetes, will be excluded. Nonoriginal research and studies that lack adequate data on lipid profiles or cardiovascular disease outcomes will be excluded. All essential data will be extracted and recorded in Microsoft Excel using a checklist data extraction format made by the review team. The same members of the review team will independently extract the data using a standardized form. Study characteristics, population characteristics, exposure and intervention, outcomes, and effect estimates data will be collected and recorded. The third reviewer, AKM, will adjudicate any disagreements between the two reviewers and provide the final decision if consensus cannot be reached. Additionally, any discrepancies will be reviewed and discussed with the third reviewer. All data will be securely stored, with regular backups in place to safeguard against any potential loss. In cases where critical data are missing, the corresponding authors of the studies will be contacted for clarification or additional information. If the missing data cannot be retrieved, the impact of these gaps in the review will be discussed, and sensitivity analyses will be performed to assess the robustness of our review findings.

Risk of Bias Assessment

The review will incorporate a rigorous assessment of risk of bias to evaluate the quality of the selected studies, using tailored tools that depend on the study design. For observational studies such as case-control, cross-sectional, and cohort studies, the Newcastle-Ottawa Scale will be used to evaluate bias in terms of selection methods, comparability, and outcomes [23]. Studies will be classified as high, moderate, or low quality. In the case of randomized controlled trials, the Cochrane Risk of Bias Tool will assess the bias, rating them as low, high, or unclear risk. In addition, an adapted JBI checklist will be used to assess cross-sectional studies, focusing on inclusion criteria, exposure and outcome measurement reliability, and statistical appropriateness [24]. Two independent reviewers will evaluate each study to determine and resolve any discrepancies, with a pilot assessment ensuring consistency. High-risk studies will be disqualified from quantitative data synthesis and will either be included in the qualitative data synthesis or form part of the sensitivity analysis. The overall risk of bias results will be summarized in a table format, and the findings will be interpreted considering the quality of the study and the level of bias.

Data Synthesis and Statistical Analysis

The review will integrate both quantitative and qualitative data synthesis, depending on the availability of data and consistency across studies. For qualitative data synthesis, the descriptive synthesis method will be used to provide a detailed overview of the included studies, highlighting key aspects such as study design, population demographics, sample sizes, and regional distribution across South Africa. A table will be developed, noting the author, year, study location, type of diabetes mellitus, prevalence rates of DD, and reported predictors. Trends and patterns, including regional prevalences, will also be documented. A specified set of predictors of DD will be extracted from all the eligible studies. To avoid selection bias and ensure consistent data extraction, predictors will be categorized into the following groups: sociodemographic factors (age, gender, level of education, and socioeconomic status), lifestyle factors (smoking, alcohol consumption, physical inactivity), clinical and cardiometabolic factors (BMI, blood pressure, hypertension, obesity, duration of diabetes, and glycemic control), and laboratory markers (glycated hemoglobin, total cholesterol, triglycerides, HDL-C, and LDL-C). A quantitative meta-analysis of predictors will be carried out if a minimum of 3 studies report the same predictors with comparable definitions and effect measures; otherwise, they will be reported qualitatively. For quantitative data synthesis, a meta-analysis will be conducted to estimate the overall prevalence of DD and assess the strength of relationships between DD and predictors among South Africans with diabetes mellitus. A random effects model will be used due to the expected heterogeneity in characteristics of the population, including the definition of dyslipidemia across studies. The

XSL•FO
RenderX

pooled prevalence will be calculated, and a meta-analysis of logistic regression will be conducted to assess predictors. Heterogeneity will be examined using the Cochran Q test and I^2 statistics, along with subgroup and sensitivity analyses, to identify potential causes of variance. For articles that cannot be meta-analyzed, qualitative data synthesis will be used to identify trends and variations across populations and outcomes. Publication bias will be assessed using funnel plots (with at least 10 papers included) and Egger or Begg tests. The analysis will be conducted using the R (version 4.3.0 or higher) package for flexibility in handling effect sizes, heterogeneity analysis, and visualization. Stata (version 18) will also be used for meta-regression. Power analysis will also be conducted to ensure a sufficient sample size for detecting statistically significant effects. Receiver operating characteristic curve analysis will be conducted to evaluate the diagnostic accuracy of specific predictors and predictive models [25,26]. Results from both qualitative and quantitative data synthesis will be presented through forest plots, subgroup forest plots, and summary tables, presenting findings on pooled prevalence, odds ratios for predictors, heterogeneity statistics, and sensitivity analyses [27-30].

Subgroup and Sensitivity Analysis

To minimize random discrepancies, a pilot data extraction will be conducted on subgroups and subsets of studies to ensure consistency and accuracy in the data extraction process. This will be done based on the areas in which the studies were conducted. In addition, all the extracted data will be safely deposited, and all files will be backed up to prevent any data loss.

Ethical Considerations

The study is a systematic review that does not involve direct interaction with human participants or the use of identifiable personal data; therefore, ethical approval will not be required.

Dissemination

The review findings will be submitted for publication in a peer-reviewed scientific journal specializing in diabetes, cardiovascular health, or public health. The communication plan details include preparing a manuscript for submission to a reputable journal, presenting the review findings at conferences, and engaging with institutions and the community.

Results

The protocol was finalized in January 2025. The literature search was conducted between October 2024 and March 2025. Title and abstract screening began in April 2025, and full-text review was completed by July 2025, with data extraction scheduled for completion by September 2025. The completion of statistical analyses is expected by October 2025. We anticipate submission of the completed systematic review and meta-analysis for publication in December 2025.

The results of the review will be presented using a PRISMA flow diagram, which will show the number of studies identified, screened, excluded, and included (Figure 1). A table displaying key characteristics of included studies will be presented. The quality assessment of the included studies will be tabulated and summarized using the Newcastle-Ottawa Scale. The included studies will be categorized as high, moderate, or low quality. A pooled prevalence estimate will be presented using a forest plot. This will be followed by subgroup analyses, which will explore differences by geographic setting. Predictors of DD will be synthesized either quantitatively or qualitatively. A detailed summary of the key findings and their implications will be reported.

XSL•FO
RenderX

Figure 1. PRISMA flowchart illustrating the systematic search process. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

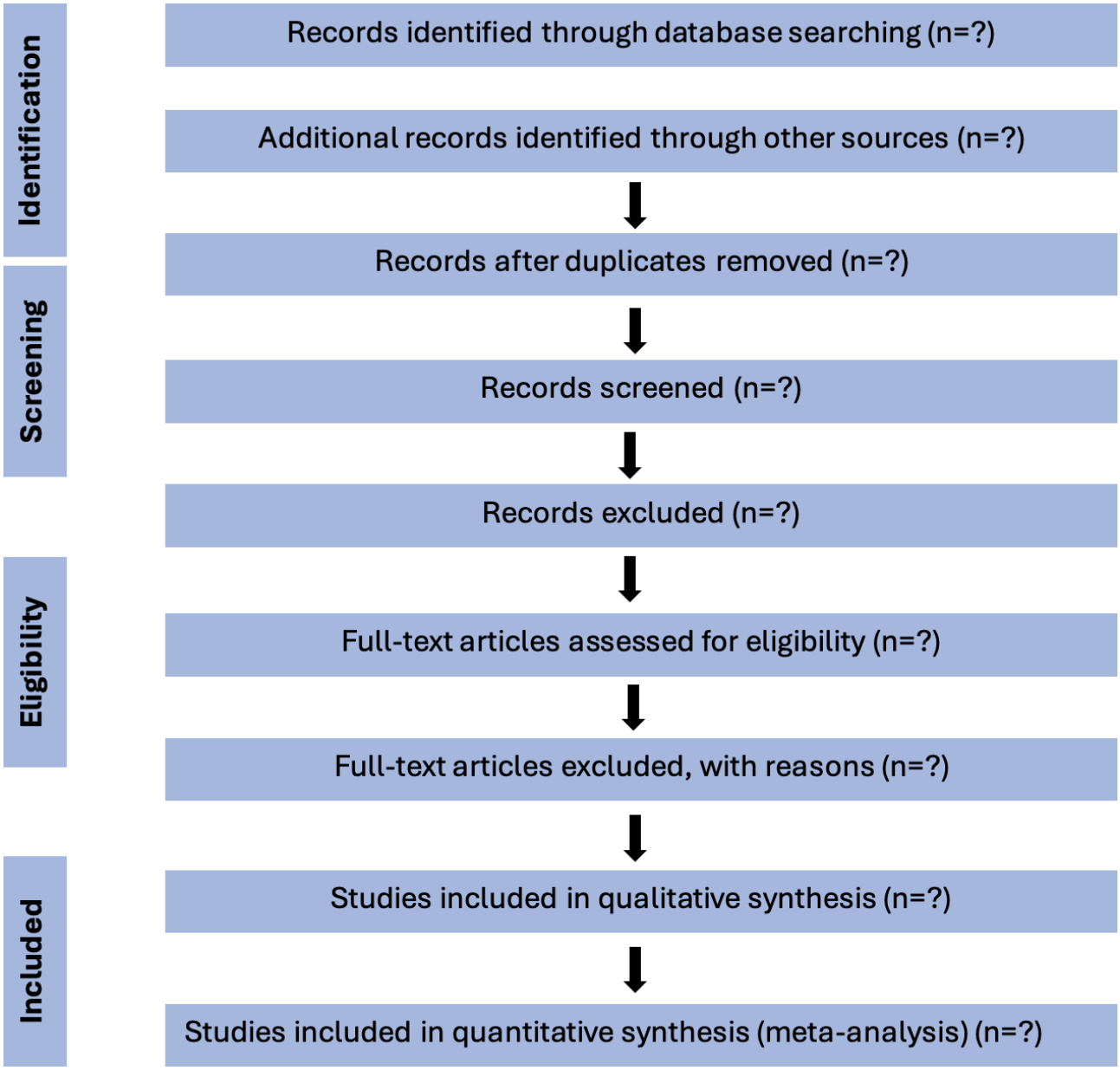

Discussion

Principal Findings

The systematic review and meta-analysis will be conducted to determine the pooled prevalence of DD among South Africans with diabetes mellitus. The review will also evaluate the relationship between DD and its determinants. The review will provide an overview of how certain sociodemographic and clinical factors influence the lipid profile in the diabetic population. The findings will be compared with existing data from other African countries and global trends. Notwithstanding the public health burden of DD, there is a lack of systematic reviews reporting on the prevalence of DD in South Africa. This review will reveal significant insight into the public health challenges associated with DD, including its cardiovascular complications. It will assess the various determinants of DD, including sociodemographic factors and health disparities. Providing a comprehensive understanding of these factors will subsequently inform the development of targeted interventions and strategies for the effective management and prevention of DD among the diabetic population, ultimately enhancing health outcomes.

This meta-analysis will yield significant findings that will be used in crafting better tailored inventions for managing DD in South Africa. Given the country's diverse socioeconomic and cultural landscape, the findings will have the potential to apprise the global strategies for tackling DD. While efforts by the South African health care system to manage complications related to diabetes have shown some advancements, the findings of this review will highlight the need for a more multifaceted approach. This approach will emphasize not only clinical management and treatment guidelines but also public health education, patient support, and community-based programs. In addition, the findings obtained can contribute to the development of future policies, which will offer a beneficial framework for other countries facing similar challenges in mitigating the health burden of DD and its complications.

XSL•FO
RenderX

The implementation of effective policies to manage DD and associated determinants among South Africans with diabetes is significant to the country's public health strategy. Such strategies could include enhancing screening programs for lipid abnormalities, encouraging adherence to lipid-lowering therapies, incorporating nutritional counseling, and addressing socioeconomic barriers to the effective management of DD. The findings of this systematic review and meta-analysis can inform the development of future policies, thereby enhancing their effectiveness. These can lead to a significant reduction in the prevalence and associated complications of DD, a decrease in health disparities among diabetic populations, and subsequently improve the overall public health outcomes in South Africa. Moreover, the findings may have broader implications for global health strategies, as these successful interventions have been implemented across different populations.

Strengths and Limitations of This Study

The study protocol outlines a rigorous and transparent methodology for a systematic review and meta-analysis to provide a more detailed understanding of the prevalence and predictors of DD in South Africa. The study protocol offers significant strengths, including adherence to the PRISMA guidelines, a comprehensive and reproducible search strategy, duplicate article screening, and data extraction by independent reviewers (MN and FA), as well as predefined procedures. The study protocol also details quality appraisal tools and statistical methods to improve the reliability of the proposed synthesis. Several limitations of the study protocol should be taken into consideration. First, restricting the search to exclude non-English electronic databases or studies may introduce selection bias, as it excludes relevant published studies in languages other than English. Second, the review is dependent on the methodological quality and completeness of reporting in the available studies, which may limit comparability and synthesis. Third, potential heterogeneity may arise from variations in study designs, populations, and methods for evaluating lipid profiles, which could hypothetically hinder data synthesis and constrain the interpretability of pooled estimates. Finally, the study protocol does not include access to nonindexed regional reports, restricted-access repositories, and private databases, which may influence the comprehensiveness of the evidence base.

Acknowledgments

The authors would like to acknowledge the late Professor Mototelo Alfred Mogale for his invaluable guidance. His expertise and support were instrumental in shaping the work.

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025) [31], the following tasks were delegated to GAI tools under full human supervision: proofreading, editing, and language refinement. The GAI tools used were ChatGPT and Grammarly. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by MN.

Funding

The authors did not apply for any financial support from public, private, or nonprofit funding agencies for the development of this protocol.

Authors' Contributions

MN, FA, and AKM contributed to protocol writing, conceptualization, and design. NSB, OHM, and SSG contributed to the review and editing of the protocol. All authors read and approved the final protocol. NSB, OHM, and SSG are supervisors of MN's doctoral research. All the authors have read and agreed to the final published version of the manuscript.

Conflicts of Interest

None declared.

References

1. Warraich HJ, Rana JS. Dyslipidemia in diabetes mellitus and cardiovascular disease. Cardiovasc Endocrinol 2017 Mar;6(1):27-32. [doi: 10.1097/XCE.0000000000000120] [Medline: 31646116]
2. Li J, Nie Z, Ge Z, Shi L, Gao B, Yang Y. Prevalence of dyslipidemia, treatment rate and its control among patients with type 2 diabetes mellitus in Northwest China: a cross-sectional study. Lipids Health Dis 2022 Aug 25;21(1):77. [doi: 10.1186/s12944-022-01691-1] [Medline: 36002855]
3. Abdissa D, Hirpa D. Dyslipidemia and its associated factors among adult diabetes outpatients in West Shewa zone public hospitals, Ethiopia. BMC Cardiovasc Disord 2022 Feb 11;22(1):39. [doi: 10.1186/s12872-022-02489-w] [Medline: 35148683]
4. Pitso L, Mofokeng TRP, Nel R. Dyslipidaemia pattern and prevalence among type 2 diabetes mellitus patients on lipid-lowering therapy at a tertiary hospital in central South Africa. BMC Endocr Disord 2021 Aug 8;21(1):159. [doi: 10.1186/s12902-021-00813-7] [Medline: 34365977]

XSL•FO
RenderX

5. Joffe BI, Distiller LA, Kalk WJ. Lipid levels in Black South Africans with type 2 diabetes. Diabetes Care 2004 Jul 1;27(7):1839-1840. [doi: 10.2337/diacare.27.7.1839] [Medline: 15220279]
6. Omodanisi EI, Tomose Y, Okeleye BI, Ntwampe SKO, Aboua YG. Prevalence of dyslipidaemia among type 2 diabetes mellitus patients in the Western Cape, South Africa. Int J Environ Res Public Health 2020 Nov 24;17(23):8735. [doi: 10.3390/ijerph17238735] [Medline: 33255455]
7. Peer N, Lombard C, Steyn K, Levitt N. High prevalence of metabolic syndrome in the Black population of Cape Town: the Cardiovascular Risk in Black South Africans (CRIBSA) study. Eur J Prev Cardiol 2015 Aug;22(8):1036-1042. [doi: 10.1177/2047487314549744] [Medline: 25208906]
8. Motala AA, Esterhuizen T, Pirie FJ, Omar MAK. The prevalence of metabolic syndrome and determination of the optimal waist circumference cutoff points in a rural South African community. Diabetes Care 2011 Apr;34(4):1032-1037. [doi: 10.2337/dc10-1921] [Medline: 21330644]
9. IDF diabetes atlas, 10th edition. : International Diabetes Federation; 2021 URL: https://diabetesatlas.org/resources/previous-editions/ [accessed 2026-02-06]
10. Sifunda S, Mbewu AD, Mabaso M, et al. Prevalence and psychosocial correlates of diabetes mellitus in South Africa: results from the South African National Health and Nutrition Examination Survey (SANHANES-1). Int J Environ Res Public Health 2023 May 12;20(10):5798. [doi: 10.3390/ijerph20105798] [Medline: 37239526]
11. Global report on diabetes. : World Health Organization; 2016 URL: https://www.who.int/publications/i/item/9789241565257 [accessed 2026-02-06]
12. Jain V, Minhas AMK, Ariss RW, et al. Demographic and regional trends of cardiovascular diseases and diabetes mellitus-related mortality in the United States from 1999 to 2019. Am J Med 2023 Jul;136(7):659-668. [doi: 10.1016/j.amjmed.2023.03.002] [Medline: 37183138]
13. Raghavan S, Vassy JL, Ho YL, et al. Diabetes mellitus-related all-cause and cardiovascular mortality in a national cohort of adults. J Am Heart Assoc 2019 Feb 19;8(4):e011295. [doi: 10.1161/JAHA.118.011295] [Medline: 30776949]
14. Mendis S, Graham I, Narula J. Addressing the global burden of cardiovascular diseases; need for scalable and sustainable frameworks. Glob Heart 2022;17(1):48. [doi: 10.5334/gh.1139] [Medline: 36051329]
15. Sunil B, Ashraf AP. Dyslipidemia in pediatric type 2 diabetes mellitus. Curr Diab Rep 2020 Sep 9;20(10):53. [doi: 10.1007/s11892-020-01336-6] [Medline: 32909078]
16. Maahs DM, Wadwa RP, Bishop F, Daniels SR, Rewers M, Klingensmith GJ. Dyslipidemia in youth with diabetes: to treat or not to treat? J Pediatr 2008 Oct;153(4):458-465. [doi: 10.1016/j.jpeds.2008.05.062] [Medline: 18847618]
17. Farzana N, Islam MS, Selim S, et al. The pattern of diabetic care and glycemic control among the ambulatory diabetic patients in tertiary care settings in Bangladesh. Sci Rep 2024 Nov 25;14(1):29220. [doi: 10.1038/s41598-024-67036-3] [Medline: 39587113]
18. Ibrahim AO, Agboola SM, Elegbede OT, Ismail WO, Agbesanwa TA, Omolayo TA. Glycemic control and its association with sociodemographics, comorbid conditions, and medication adherence among patients with type 2 diabetes in southwestern Nigeria. J Int Med Res 2021 Oct;49(10):3000605211044040. [doi: 10.1177/03000605211044040] [Medline: 34632841]
19. Hill-Briggs F, Adler NE, Berkowitz SA, et al. Social determinants of health and diabetes: a scientific review. Diabetes Care 2021 Jan 1;44(1):258-279. [doi: 10.2337/dci20-0053]
20. Duan J, Wang C, Liu D, et al. Prevalence and risk factors of chronic kidney disease and diabetic kidney disease in Chinese rural residents: a cross-sectional survey. Sci Rep 2019 Jul 18;9(1):10408. [doi: 10.1038/s41598-019-46857-7] [Medline: 31320683]
21. Garcia-Ramirez MA, Mora-Cuevas G, Castellanos-Alvarado EA, Lopez-Ortega ME, Ruvalcaba-Jacobo OG, Guillen-Bonilla JT. Stress as a trigger for fatty liver and dyslipidemia. IBRO Neuroscience Reports 2023 Oct;15:S613. [doi: 10.1016/j.ibneur.2023.08.1219]
22. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
23. Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol 2010 Sep;25(9):603-605. [doi: 10.1007/s10654-010-9491-z] [Medline: 20652370]
24. Munn Z, Moola S, Lisy K, Riitano D, Tufanaru C. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. Int J Evid Based Healthc 2015 Sep;13(3):147-153. [doi: 10.1097/XEB.0000000000000054] [Medline: 26317388]
25. Hajian-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J Intern Med 2013;4(2):627-635. [Medline: 24009950]
26. Serdar CC, Cihan M, Yücel D, Serdar MA. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. Biochem Med (Zagreb) 2021 Feb 15;31(1):010502. [doi: 10.11613/BM.2021.010502] [Medline: 33380887]
27. The R Project for Statistical Computing. 2023. URL: https://www.R-project.org [accessed 2026-02-03]
28. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634. [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]

XSL•FO
RenderX

29. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994 Dec;50(4):1088-1101. [Medline: 7786990]
30. What's new in Stata 18. Stata. 2023. URL: https://www.stata.com/stata18/ [accessed 2026-02-03]
31. Suchikova Y, Tsybuliak N, Teixeira da Silva JA, Nazarovets S. GAIDeT (Generative AI Delegation Taxonomy): a taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. Account Res 2025 Aug 8;1–27:1-27. [doi: 10.1080/08989621.2025.2544331] [Medline: 40781729]

Abbreviations

ACVD: atherosclerotic cardiovascular disease
DD: diabetic dyslipidemia
HDL-C: high-density lipoprotein cholesterol
LDL-C: low-density lipoprotein cholesterol
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: International Prospective Register of Systematic Reviews

Edited by J Sarvestan; submitted 20.Aug.2025; peer-reviewed by RSG Mahmoud, Y Kong; revised version received 10.Dec.2025; accepted 12.Dec.2025; published 19.Feb.2026.

Please cite as:
Nemukula M, Adane F, Malaisamy AK, Bhavesh NS, Majane OH, Gololo SS
Diabetic Dyslipidemia and Its Determinants Among People With Diabetes in South Africa: Protocol for a Systematic Review and Meta-Analysis
JMIR Res Protoc 2026;15:e82716
URL: https://www.researchprotocols.org/2026/1/e82716
doi:10.2196/82716

© Mashudu Nemukula, Fentahun Adane, Arun Kumar Malaisamy, Neel Sarovar Bhavesh, Olebogeng Harold Majane, Sechene Stanley Gololo. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effectiveness of Rehabilitation Interventions for Improving Leisure Participation Following Stroke: Protocol for a Systematic Review

Serena Alves-Stein[1,2], BOccThy (Hons), MClinRehab; Natasha A Lannin[1,2], BSc, PhD; Kylie Wales[1,3], BAPSC, PhD; Sharon Kramer[1,2,4], BPhysio, PhD; Laura Jolliffe[5,6,7], BOccThy (Hons), PhD

[1]Department of Neuroscience, School of Translational Medicine, Monash University, Level 6, 99 Commercial Rd, Melbourne, Australia

[2]Department of Occupational Therapy, Alfred Care Group, Bayside Health, Level 6, 99 Commercial Rd, Melbourne, Australia

[3]Faculty of Health Sciences, School of Allied Health, Australian Catholic University, North Sydney, Australia

[4]MedTechVic Hub, Swinburne University, Hawthorn, Australia

[5]Department of Occupational Therapy, School of Primary and Allied Health Care, Monash University, Australia

[6]Department of Occupational Therapy, Peninsula University Hospital, Peninsula Care Group, Bayside Health, Australia

[7]National Centre for Healthy Ageing, Melbourne, Australia

Corresponding Author:
Natasha A Lannin, BSc, PhD
Department of Neuroscience, School of Translational Medicine, Monash University, Level 6, 99 Commercial Rd, Melbourne, Australia

Abstract

Background: Leisure participation is an important rehabilitation goal for survivors of stroke. Following stroke, there is a reduction in leisure participation; however, the focus of rehabilitation is typically on remediation of personal care activities and mobility. Furthermore, previous systematic reviews and current clinical practice guidelines provide inconsistent recommendations for rehabilitation interventions to improve leisure participation. This highlights the need for a comprehensive and targeted review of the literature to help inform clinical practice.

Objective: We propose a systematic review to synthesize data on the effectiveness of rehabilitation interventions to increase leisure participation in adult survivors of stroke, taking into account time since stroke and intervention context.

Methods: Searches will be conducted in MEDLINE, CINAHL, Embase, and CENTRAL. We will include randomized controlled trials and nonrandomized controlled trials that include adult survivors of stroke and test the effectiveness of rehabilitation interventions for leisure participation. Eligible interventions will be those that aim to improve leisure participation or where leisure participation is an outcome of interest. Two reviewers will independently screen full-text articles, and one reviewer will extract data, with a second reviewer providing confirmation. The Physiotherapy Evidence Database Scale will be used to assess the methodological quality of the studies. A random-effects meta-analysis will be performed, and a Cochran *Q* test will assess heterogeneity among studies. Outcome measures of leisure participation may include measures of amount, satisfaction and confidence, and performance. Secondary outcomes will include quality of life measures, adverse events, and resource use.

Results: Results will be discussed based on subgroup analyses where possible, including (1) time since stroke (early subacute, late subacute, and chronic), (2) delivery of the intervention (group or individualized), and (3) type of intervention (functional impairment, leisure education, and recreation participation). At the time of this protocol publication, the systematic review has progressed to data analysis, with publication of results expected in early 2026.

Conclusions: The findings of this review will increase understanding of effective rehabilitation practices to increase leisure participation after stroke and may contribute to updates of existing clinical practice guidelines for stroke rehabilitation.

Trial Registration: PROSPERO CRD42024547133; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024547133

International Registered Report Identifier (IRRID): DERR1-10.2196/71353

(JMIR Res Protoc 2026;15:e71353) doi:10.2196/71353

KEYWORDS

stroke; leisure activities; rehabilitation; community integration; occupational therapy; stroke rehabilitation; social engagement

XSL•FO
RenderX

Introduction

Stroke is a leading cause of disability worldwide, with data suggesting that the increase in incidence in younger age groups has led to an increase in years lived with disability [1,2]. Stroke-related impairments, such as motor, cognitive, and communication impairments, impact activity participation [3]. Typically, rehabilitation programs tend to focus on remediating impairments and improving independence in basic activities of daily living such as feeding and bathing [4,5], all seen as key to supporting discharge from hospital to home [6]. However, it is important that social and community-related needs and goals are also addressed during stroke recovery [5]. With an increased number of survivors of stroke returning to living in the community, there is an increasing need for community-based rehabilitation programs to enable community engagement [7], which includes leisure participation.

Worldwide, survivors of stroke participate less in leisure activities [8-11]. Furthermore, there is a change in the type of leisure activities they engage in, with an increase in sedentary, home-based activities such as watching television [4,7]. Engaging in leisure following stroke has benefits for adjustment, resilience [12], and quality of life [11]. Leisure participation has been found to promote positive mental health outcomes, reduce the risk of cognitive decline, and reduce the risk of physical disability and disease related to aging [13-16].

Despite the importance of leisure participation, clinical practice guidelines provide limited direction on how to address reduced leisure participation. The evidence underpinning guideline recommendations is considered weak [17]. This is likely due to a lack of research and variability in study findings. For example, the Australian guidelines recommend targeted occupational therapy programs to address leisure participation [17], whereas the American guidelines recommend education and support to overcome barriers to physical activity [18]. In comparing the guideline recommendations, it is also notable that some of the cited studies are specific to leisure interventions [19-23], whereas others test the effect of non–leisure-specific interventions on broader participation outcomes [24-26].

A more comprehensive and targeted search of the literature can provide greater clarity on clinical recommendations for poststroke leisure rehabilitation. There is a need to better understand not only the effectiveness of leisure interventions but also the context in which these interventions are delivered. This includes considering time since stroke, the intervention delivery format (eg, one-on-one or in a group), and the content of leisure interventions.

Given these needs, this systematic review will synthesize data from leisure intervention studies that measure leisure participation for community-dwelling survivors of stroke. The review aims to address the following research question: what is the effectiveness of rehabilitation interventions in increasing leisure participation in adult survivors of stroke, taking into account time since stroke and intervention context?

Methods

Ethical Considerations

This systematic review will analyze publicly available data from previously published peer-reviewed articles and will not recruit new human participants. Therefore, approval from a research ethics committee is not required.

Study Design

This systematic review protocol has been developed according to the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses–Protocols) statement [27]; the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement will guide the systematic review reporting. The protocol has been registered in PROSPERO (CRD42024547133).

Eligibility Criteria

Overview

The primary objective of this review is to synthesize effectiveness research on the use of leisure rehabilitation interventions after stroke. We will include publications from any date in our search. We acknowledge that this may influence key definitions of “leisure” used in the context of research [28-30]. Challenges noted to date arise from whether leisure must always be about how one uses discretionary time as opposed to the view that an activity may be considered leisure provided there is a pleasurable mental state that occurs during the activity [28]. Cultural context may also influence whether an activity would be considered leisure [30], and we acknowledge that this may shift over time. Consistent with definitions reported in publications over the past 2 decades, we will consider leisure activities those activities that are not necessary or obligatory to survival or to necessities of life, are activities of choice, and are done for the purpose of enjoyment or pleasure [28,30].

Types of Studies

Studies will be included if they are full publications of pretest-posttest randomized or nonrandomized controlled trials that assess the effects of the interventions (including those with waitlist control groups).

Population and Setting

Studies that include adults (aged ≥18 years) diagnosed with stroke who are community dwelling will be included. Studies involving mixed diagnoses will be included in cases in which the data from those with a diagnosis of stroke can be extracted or where ≥75% of the participants have a diagnosis of stroke. Studies involving mixed age groups will be included in cases in which the data from those aged ≥18 years can be extracted or where ≥75% of the participants are aged ≥18 years. Studies will be excluded if they relate to children and young people aged <18 years (including those with childhood-acquired stroke) or include participants who reside in a care-based facility (eg, residential aged care or group home programs).

XSL•FO
RenderX

Intervention

Studies that provide an intervention aimed at improving leisure participation will be included. The leisure ability model [31] will be used to provide a conceptual framework for describing interventions targeting leisure participation. The leisure ability model describes 3 "categories of service" for providing leisure interventions: functional impairment, leisure education, and recreation participation [31]. Leisure interventions will be provided by rehabilitation professionals, including but not limited to occupational therapists, recreational therapists, physiotherapists, speech and language therapists, nurses, social workers, and allied health assistants. Mode of delivery can include individual or group-based interventions. We will exclude studies in which leisure participation was not a target of intervention for all participants or was not part of the intervention protocol.

Comparison Interventions

Included studies may have either an active comparison, such as those in which participants receive usual care or another nonleisure intervention, or an inactive (no-treatment) comparison. Usual care in the included studies refers to rehabilitation that addresses impairment and/or participation goals but not specifically leisure (eg, addressing self-care and walking ability). A vs B–type intervention studies will be included in cases in which there is a comparison of 2 or more interventions where at least 1 addresses leisure participation.

Outcomes

To be eligible for inclusion, studies must use a leisure-specific outcome measure. Leisure participation may be assessed as amount, satisfaction or confidence, or performance; where possible, we will analyze each separately. Examples of how these domains could be measured include the following:

- Amount—participant diary or Nottingham Leisure Questionnaire [32]
- Satisfaction or confidence—Leisure Satisfaction Scale [33]
- Performance—participant self-reported performance rating (eg, goal attainment scaling of a leisure performance goal [34])

Additional outcomes that will be reviewed include quality of life, adverse events, and resource use, where reported. Where available, economic evaluation reports will be assessed. We will exclude measures primarily assessing engagement in paid employment, volunteering, or homemaking.

Search Strategy

Electronic searches will be conducted in MEDLINE, CINAHL, The Cochrane Library, and Embase. Hand searches of the reference lists of included studies and any related systematic reviews will be conducted to identify further potentially eligible studies. The search strategy will include terms related to stroke, rehabilitation, and leisure. Table 1 provides the search terms that will be used to search for and identify relevant studies. The complete search strategy for each database is available in Multimedia Appendix 1.

Table . Systematic review search terms structured for MEDLINE.[a]

	Population (stroke)	Intervention (rehabilitation)	Outcome (leisure)
Text words	"(Stroke or cerebrovascular accident or cerebr* vasc* or cva* or brain infarct* or cerebral infarct* or brain vascular accident or poststroke or post-stroke or hemipleg* or hemipare*).ti,ab,kw."	"(rehabilitation or community health services or allied health).ti,ab,kw."	"(leisure or hobb* or recreation* or recreational activit* or meaningful) adj3 activit*) or community) adj3 engagement) or community) adj3 integration) or leisure) adj3 (activit* or exploration or participation or education or counselling)).ti,ab,kw."
MeSH[b] terms	"stroke/ or brain infarction/ or hemorrhagic stroke/ or ischemic stroke/"	"rehabilitation/ or neurological rehabilitation/ or stroke rehabilitation/ or occupational therapy/ or recreation therapy/exp Rehabilitation/"	"exp Leisure Activities/ leisure activities/ or recreation/"

[a]Search terms and Boolean operators will be tailored for each database.

[b]MeSH: Medical Subject Headings.

Procedures

All studies identified via the search strategy will be uploaded to Covidence (Veritas Health Innovation) [35], and duplicates will be removed. One author will independently screen studies for eligibility based on title and abstract, and all excluded articles at the title and abstract stage will be reviewed by a senior researcher to ensure that no articles are inadvertently excluded. Full-text articles will then be independently reviewed by 2 authors to determine eligibility. Disagreements will be resolved first through discussion; however, if consensus is not reached, a third research team member will complete an independent review. Study authors may be contacted to clarify details and/or data. The results of the screening process will be provided in detail using a PRISMA study flowchart.

One reviewer will extract data using a predetermined form; all data extraction will be independently verified by the senior author (LJ). If discrepancies occur, these will be resolved through discussion between the authors. If not resolved, a third author will independently extract the data. The data extraction form will include variables necessary for describing the studies and conducting data analysis, including characteristics of the participants (including age, gender, time since stroke, diagnostic details, baseline leisure participation, and comorbidities), study methods (design, randomization, trial duration, follow-up time points, and information required for risk-of-bias assessment),

XSL•FO
RenderX

sample size, country of study, intervention details (according to the TIDieR [Template for Intervention Description and Replication] checklist items [36]), outcomes, and results (number of participants in the analysis, summary data for each intervention group, and estimate of effect along with measure of variance [eg, mean difference and 95% CI]).

The Physiotherapy Evidence Database Scale [37] will be used to examine the risk of bias by one reviewer, and this will then be independently verified by a second reviewer. Any disagreements in appraisal score will be addressed through consensus. A higher score indicates a higher methodological quality of the trial. A score of <4 will indicate poor quality, and these studies will be excluded from meta-analysis. Studies that are excluded from the meta-analysis due to the risk of bias will still be reported descriptively in the text of the review.

Statistical Analysis

Overview

For all outcomes, summary data will be presented by group and reported as estimate of effect size and CI as feasible. If sufficiently homogeneous in terms of clinical (eg, patient characteristics), methodological (eg, study design), and statistical (eg, forest plot consistency) characteristics, statistical analyses will be conducted to determine the treatment effect using a random-effects model. A random-effects model will be used as it is expected that the eligible studies will estimate different intervention effects, contributing to between-study heterogeneity [38]. Statistical heterogeneity will be assessed using the I^2 test and Cochran Q test, which is based on a chi-square distribution. If the I^2 is >80%, heterogeneity will be explored through sensitivity analyses, which will include analyses of studies with low or unclear risk of bias (excluding studies with a high risk of bias) and subgroup analyses. Clinical and methodological heterogeneity will be discussed by the reviewer team (ie, interventions, populations, and assessment methods) to inform subgroup analyses. If meta-analysis is not appropriate as a result of substantial heterogeneity (I^2>80%), a narrative synthesis of the findings from the included studies will be provided.

For continuous data, the treatment effect will be calculated using standardized mean differences and 95% CIs in cases in which different studies use different scales to assess the same outcome. Mean differences and 95% CIs will be calculated in cases where studies have used the same measurement tool. For dichotomous outcomes, a relative risk ratio will compare the risk of an event occurring between the study intervention groups. In cases in which studies report more than 1 leisure participation outcome construct, we will use their reported primary outcome in the main analysis. Other reported outcome constructs will be included in subgroup analyses.

Subgroup Analyses

As described in the eligibility criteria, leisure participation outcome measurement domains are amount, satisfaction or confidence, and performance. Where possible, these outcomes will be analyzed separately. To investigate intervention context, further subgroup analyses may include time since stroke, delivery of the intervention, and type of intervention where possible. In accordance with the Stroke Recovery and Rehabilitation Roundtable recommendations, poststroke time frames will be classified as early subacute (7 days to 3 months), late subacute (3-6 months), and chronic (>6 months) [39]. Noting the Stroke Recovery and Rehabilitation Roundtable point regarding understanding the efficacy of commencing recovery trials in the early subacute phase, subgroup analysis of time since stroke will compare <3 months (early subacute) and >3 months (late subacute and chronic) after stroke. Subgroup analyses on delivery of the intervention will compare individually provided therapies with group programs. Intervention content will be categorized using the leisure ability model categories of service [31].

Results

At the time of publication of this protocol, this systematic review has progressed with searches completed in September 2025 and data analysis completed in October 2025. Results will be discussed based on subgroup analyses where possible, including:

1. time since stroke (early subacute, late subacute, and chronic);
2. delivery of the intervention (group or individualized); and
3. type of intervention (functional impairment, leisure education, and recreation participation).

A manuscript for the systematic review results is being finalized and is anticipated to be published in the first half of 2026.

Discussion

Overview

The findings of this review will benefit researchers, clinicians, and people living in the community who have experienced stroke. For academic researchers, the review will have the potential to identify gaps in available evidence, thereby highlighting key areas for future research. For clinicians, the findings will inform clinical practice guidelines and, thereby, influence intervention selection. For people living in the community who have experienced stroke, the findings may shed light on how to improve leisure participation in the longer term, something that has to date remained unknown. Specifically, this review aims to provide greater clarity and detail on how to deliver effective leisure rehabilitation by considering the content of the interventions in the included studies and exploring other factors such as when leisure rehabilitation should be provided and what aspects of leisure participation are improved.

Comparison With Prior Work

Leisure participation following stroke is an important participation goal in stroke rehabilitation; however, systematic reviews to date have not systematically located or comprehensively analyzed the body of research, limiting the ability to draw reliable conclusions. Two previous systematic reviews [19,40] that included the topic of leisure in stroke recovery have methodological limitations influencing the interpretation of their results. Neither review required studies to include a leisure-specific outcome measure, and the 2 reviews used different working definitions of leisure interventions. The

XSL•FO
RenderX

studies had a broad scope, which limited their ability to provide a targeted and in-depth analysis of the efficacy of leisure rehabilitation. These limitations have been taken into consideration in the design of this systematic review, which is reflected in the decisions to use subgroup analyses to explore outcome measure constructs, intervention delivery, and intervention content.

Strengths and Limitations

A key strength of this study is that it addresses the identified limitations of previous systematic reviews by focusing specifically on leisure rehabilitation and leisure-specific outcomes in a stroke population. This allows for a focused analysis and synthesis of results to inform future research and clinical practice. As with all studies, there are potential limitations. By design, our search is limited to studies published in English. We have also used a single reviewer for screening titles and abstracts, which may increase the risk of missed studies, and had a single reviewer complete data extraction, which may result in errors. We have managed these potential errors with verification of these steps by a second author, as described in the Methods section. With regard to statistical methods, we will report I^2 values to quantify heterogeneity, which may not indicate the extent to which true effects vary across different populations or settings. In addition, we anticipate the following challenges. During the search phase, we anticipate that locating and identifying relevant studies may be a challenge if studies include leisure outcomes among other outcomes of interest or there is variation in how leisure interventions are defined or described. We also acknowledge that how leisure is measured may vary across studies in addition to the outcome measures used. Such variability may limit the ability to meta-analyze data from the included studies.

Conclusions

Summarizing trials completed to date will allow for the analysis of effectiveness alongside discussion of contextual factors that may influence leisure rehabilitation intervention delivery. Thus, this systematic review will help in the decision-making of researchers and clinicians, ultimately advancing the field of rehabilitation.

Funding

This research received no specific grant from any funding agency in the commercial or not-for-profit sectors. SA-S is supported by an Australian government Research Training Program scholarship (Monash University). NAL is supported by a Heart Foundation (Australia) Future Leader Fellowship (106762).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

SA-S, NAL, and LJ conceived the study and developed the methodology. SA-S, NAL, KW, SK, and LJ were involved in all investigation procedures. SA-S managed the project (administration). SA-S, NAL, and LJ completed the initial manuscript draft, and all authors reviewed and edited the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy for each database.
[DOCX File, 22 KB - resprot_v15i1e71353_app1.docx]

References

1. Feigin VL, Stark BA, Johnson CO. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021 Oct;20(10):795-820. [doi: 10.1016/S1474-4422(21)00252-0] [Medline: 34487721]
2. Lannin NA, Anderson CS, Kim J, et al. Treatment and outcomes of working aged adults with stroke: results from a national prospective registry. Neuroepidemiology 2017;49(3-4):113-120. [doi: 10.1159/000484141] [Medline: 29136634]
3. Gittins M, Lugo-Palacios D, Vail A, et al. Stroke impairment categories: a new way to classify the effects of stroke based on stroke-related impairments. Clin Rehabil 2021 Mar;35(3):446-458. [doi: 10.1177/0269215520966473] [Medline: 33131321]
4. Reeves MJ, Thetford C, McMahon N, et al. Life and leisure activities following stroke or transient ischaemic attack (TIA): an observational, multi-centre, 6-month follow-up study. Int J Environ Res Public Health 2022 Oct 25;19(21):13848. [doi: 10.3390/ijerph192113848] [Medline: 36360725]
5. Elphick S, Price C, Tang EY. Shifting the attention to long- term community stroke care. Br J Gen Pract 2024 Dec;74(749):536-537. [doi: 10.3399/bjgp24X739953] [Medline: 39609065]

XSL•FO
RenderX

6. Gonçalves-Bradley DC, Lannin NA, Clemson L, Cameron ID, Shepperd S. Discharge planning from hospital. Cochrane Database Syst Rev 2022 Feb 24;2(2):CD000313. [doi: 10.1002/14651858.CD000313.pub6] [Medline: 35199849]
7. Mayo NE, Wood-Dauphinee S, Côté R, Durcan L, Carlton J. Activity, participation, and quality of life 6 months poststroke. Arch Phys Med Rehabil 2002 Aug;83(8):1035-1042. [doi: 10.1053/apmr.2002.33984] [Medline: 12161823]
8. Yi TI, Han JS, Lee KE, Ha SA. Participation in leisure activity and exercise of chronic stroke survivors using community-based rehabilitation services in Seongnam city. Ann Rehabil Med 2015 Apr;39(2):234-242. [doi: 10.5535/arm.2015.39.2.234] [Medline: 25932420]
9. Northcott S, Moss B, Harrison K, Hilari K. A systematic review of the impact of stroke on social support and social networks: associated factors and patterns of change. Clin Rehabil 2016 Aug;30(8):811-831. [doi: 10.1177/0269215515602136] [Medline: 26330297]
10. Woodman P, Riazi A, Pereira C, Jones F. Social participation post stroke: a meta-ethnographic review of the experiences and views of community-dwelling stroke survivors. Disabil Rehabil 2014;36(24):2031-2043. [doi: 10.3109/09638288.2014.887796] [Medline: 24597937]
11. Schnitzler A, Jourdan C, Josseran L, Azouvi P, Jacob L, Genêt F. Participation in work and leisure activities after stroke: a national study. Ann Phys Rehabil Med 2019 Sep;62(5):351-355. [doi: 10.1016/j.rehab.2019.04.005] [Medline: 31096014]
12. Caldwell LL. Leisure and health: why is leisure therapeutic? Br J Guid Counc 2005 Feb;33(1):7-26. [doi: 10.1080/03069880412331335939]
13. Passmore A, French D. A model of leisure and mental health in Australian adolescents. Behav change 2000 Sep 1;17(3):208-220. [doi: 10.1375/bech.17.3.208]
14. Crowe M, Andel R, Pedersen NL, Johansson B, Gatz M. Does participation in leisure activities lead to reduced risk of Alzheimer's disease? A prospective study of Swedish twins. J Gerontol B Psychol Sci Soc Sci 2003 Sep;58(5):249-255. [doi: 10.1093/geronb/58.5.p249] [Medline: 14507930]
15. Singh-Manoux A, Richards M, Marmot M. Leisure activities and cognitive function in middle age: evidence from the Whitehall II study. J Epidemiol Community Health 2003 Nov;57(11):907-913. [doi: 10.1136/jech.57.11.907] [Medline: 14600119]
16. Paggi ME, Jopp D, Hertzog C. The importance of leisure activities in the relationship between physical health and well-being in a life span sample. Gerontology 2016;62(4):450-458. [doi: 10.1159/000444415] [Medline: 26974682]
17. Stroke Foundation. Community participation and long-term care. Australian and New Zealand Living Clinical Guidelines for Stroke Management. 2024. URL: https://app.magicapp.org/#/guideline/6nYJxE/section/nY7Vgj [accessed 2025-05-25]
18. Winstein CJ, Stein J, Arena R, et al. Guidelines for adult stroke rehabilitation and recovery: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke 2016 Jun;47(6):e98-e169. [doi: 10.1161/STR.0000000000000098] [Medline: 27145936]
19. Dorstyn D, Roberts R, Kneebone I, Kennedy P, Lieu C. Systematic review of leisure therapy and its effectiveness in managing functional outcomes in stroke rehabilitation. Top Stroke Rehabil 2014;21(1):40-51. [doi: 10.1310/tsr2101-40] [Medline: 24521839]
20. Desrosiers J, Noreau L, Rochette A, et al. Effect of a home leisure education program after stroke: a randomized controlled trial. Arch Phys Med Rehabil 2007 Sep;88(9):1095-1100. [doi: 10.1016/j.apmr.2007.06.017] [Medline: 17826452]
21. Walker MF, Leonardi-Bee J, Bath P, et al. Individual patient data meta-analysis of randomized controlled trials of community occupational therapy for stroke patients. Stroke 2004 Sep;35(9):2226-2232. [doi: 10.1161/01.STR.0000137766.17092.fb] [Medline: 15272129]
22. Drummond A, Walker MF. A randomized controlled trial of leisure rehabilitation after stroke. Clin Rehabil 1995 Nov;9(4):283-290. [doi: 10.1177/026921559500900402]
23. Jongbloed L, Morgan D. An investigation of involvement in leisure activities after a stroke. Am J Occup Ther 1991 May;45(5):420-427. [doi: 10.5014/ajot.45.5.420] [Medline: 2048623]
24. Barclay RE, Stevenson TJ, Poluha W, Ripat J, Nett C, Srikesavan CS. Interventions for improving community ambulation in individuals with stroke. Cochrane Database Syst Rev 2015 Mar 13;2015(3):CD010200. [doi: 10.1002/14651858.CD010200.pub2] [Medline: 25767912]
25. Kim M, Cho K, Lee W. Community walking training program improves walking function and social participation in chronic stroke patients. Tohoku J Exp Med 2014 Dec;234(4):281-286. [doi: 10.1620/tjem.234.281] [Medline: 25483170]
26. Harrington R, Taylor G, Hollinghurst S, Reed M, Kay H, Wood VA. A community-based exercise and education scheme for stroke survivors: a randomized controlled trial and economic evaluation. Clin Rehabil 2010 Jan;24(1):3-15. [doi: 10.1177/0269215509347437] [Medline: 20026571]
27. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
28. Raz-Yurovich L. Leisure: definitions, trends, and policy implications. Popul Res Policy Rev 2022 Jun;41(3):981-1019. [doi: 10.1007/s11113-021-09675-2]
29. Aitchison C. From leisure and disability to disability leisure: developing data, definitions and discourses. Disabil Soc 2003 Dec;18(7):955-969. [doi: 10.1080/0968759032000127353]

XSL•FO
RenderX

30. Purrington A, Hickerson B. Leisure as a cross-cultural concept. World Leis J 2013 Jun;55(2):125-137. [doi: 10.1080/04419057.2013.782564]
31. Stumbo NJ, Peterson CA. The leisure ability model. Ther Recreation J 1998;32(2):82-96 [FREE Full text]
32. Drummond AE, Parker CJ, Gladman JR, Logan PA, TOTAL Study Group. Development and validation of the Nottingham Leisure Questionnaire (NLQ). Clin Rehabil 2001 Dec;15(6):647-656. [doi: 10.1191/0269215501cr438oa] [Medline: 11777095]
33. Beard JG, Ragheb MG. Measuring leisure satisfaction. J Leis Res 1980 Jan;12(1):20-33. [doi: 10.1080/00222216.1980.11969416]
34. Cardillo JE, Kiresuk TJ, Smith A. Goal Attainment Scaling: Applications, Theory, and Measurement: Psychology Press; 2014. [doi: 10.4324/9781315801933]
35. Home page. Covidence Systematic Review Software. URL: https://www.covidence.org/ [accessed 2025-05-25]
36. Hoffmann TC, Walker MF. "TIDieR-ing up" the reporting of interventions in stroke research: the importance of knowing what is in the "black box". Int J Stroke 2015 Jul;10(5):657-658. [doi: 10.1111/ijs.12524] [Medline: 26094670]
37. Maher CG, Sherrington C, Herbert RD, Moseley AM, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther 2003 Aug;83(8):713-721. [doi: 10.1093/ptj/83.8.713] [Medline: 12882612]
38. Maitra S. Fixed-effect versus random-effect model in meta-analysis: how to decide? Indian J Anaesth 2025 Jan;69(1):143-146. [doi: 10.4103/ija.ija_1203_24] [Medline: 40046706]
39. Bernhardt J, Hayward KS, Kwakkel G, et al. Agreed definitions and a shared vision for new standards in stroke recovery research: the stroke recovery and rehabilitation roundtable taskforce. Neurorehabil Neural Repair 2017 Sep;31(9):793-799. [doi: 10.1177/1545968317732668] [Medline: 28934920]
40. Graven C, Brock K, Hill K, Joubert L. Are rehabilitation and/or care co-ordination interventions delivered in the community effective in reducing depression, facilitating participation and improving quality of life after stroke? Disabil Rehabil 2011;33(17-18):1501-1520. [doi: 10.3109/09638288.2010.542874] [Medline: 21204742]

Abbreviations

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses–Protocols
TIDieR: Template for Intervention Description and Replication

Edited by A Schwartz; submitted 16.Jan.2025; peer-reviewed by L Shuranova; accepted 31.Dec.2025; published 26.Feb.2026.

Please cite as:
Alves-Stein S, Lannin NA, Wales K, Kramer S, Jolliffe L
The Effectiveness of Rehabilitation Interventions for Improving Leisure Participation Following Stroke: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e71353
URL: https://www.researchprotocols.org/2026/1/e71353
doi:10.2196/71353

© Serena Alves-Stein, Natasha A Lannin, Kylie Wales, Sharon Kramer, Laura Jolliffe. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Prevalence of Problematic Digital Media Use Among Young Adults: Protocol for a Systematic Review

Rachelle Marie Chanmany Pastor[1*], PhD; Fayette Nguyen Truax[1*], PhD; Safiye Sahin[1*], PhD; Janet Donnelly[1,2*], PhD; Jan M Nick[1,3*], PhD; Shanalee Tamares[4*], MLIS

[1]School of Nursing, Loma Linda University, 11262 Campus St., Loma Linda, CA, United States

[2]Sharp Health Care, San Diego, CA, United States

[3]JBI Center for Evidence Synthesis Loma Linda University Health, Loma Linda, CA, United States

[4]Loma Linda University Libraries, Loma Linda, CA, United States

[*]all authors contributed equally

Corresponding Author:
Rachelle Marie Chanmany Pastor, PhD
School of Nursing, Loma Linda University, 11262 Campus St., Loma Linda, CA, United States

Abstract

Background: Problematic digital media use (PDMU) among young people has been on the rise. PDMU is defined as excessive use of digital media, the internet, or electronic communication leading to user dysfunction and harm to other individuals. Evidence links excessive use of media with various mental health disorders, behavioral problems, substance abuse, poor sleep hygiene, and social dysfunction. This maladaptive behavior is pervasive among young people, yet there is a paucity of studies that comprehensively examine the phenomenon in this specific population.

Objective: This systematic review seeks to examine the current global prevalence of PDMU among young adults aged 18 to 24 years, explore the extent of the issue across different regions of the world, and identify key factors contributing to its occurrence.

Methods: The proposed systematic review will use the Prevalence Estimates Reviews – Systematic Review Methodology Group (PERSyst) methodology for systematic reviews of prevalence and incidence. A 3-step search strategy will be used covering 2020 to 2025, with no language restrictions. Nine sources will be searched: Embase (Elsevier), PubMed, PsycInfo (EBSCO), Web of Science Core Collection (Clarivate), Communication & Mass Media Complete (EBSCO), LILACS (VHL Search Portal), China National Knowledge Infrastructure, ProQuest Dissertations and Theses Global, and Google Scholar. Study selection will follow a 3-step process, including critical appraisal for methodological quality. Standardized data extraction tools will be used. Two reviewers will make decisions independently; conflicts will be resolved through consensus. Narrative synthesis will be conducted, and where possible, a meta-analysis will estimate PDMU prevalence with 95% CIs using a random-effects model. Heterogeneity will be assessed using the chi-square and I^2 statistics. The Grading of Recommendations Assessment, Development, and Evaluation approach will be used to assess the certainty of the evidence.

Results: The protocol was completed in July 2025. The comprehensive search of electronic bibliographic databases began in September 2025, followed by deduplication, screening, and selection of eligible studies. Title and abstract screening were completed in January 2026, and full-text review is ongoing. Data extraction and synthesis will be conducted between March and April 2026. The manuscript will be prepared and completed by August 2026, with plans to submit the manuscript for publication in September 2026.

Conclusions: The planned review contributes to the growing body of evidence on digital media use among young adults by highlighting its potential impact on overall well-being. To our knowledge, this is the first comprehensive synthesis focused specifically on this population. The findings are expected to highlight the need for routine screening and early intervention strategies to address the social, mental, and physical health risks associated with digital media overuse in young adulthood.

Trial Registration: PROSPERO CRD420251075912; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251075912

International Registered Report Identifier (IRRID): PRR1-10.2196/82245

(JMIR Res Protoc 2026;15:e82245) doi:10.2196/82245

KEYWORDS

behavioral addiction; technology use disorder; screen time; digital addiction; internet overuse

XSL•FO
RenderX

Introduction

Background

Problematic digital media use (PDMU) among young people has been on the rise [1]. There is evidence suggesting that risks of maladaptive technological use may result in overuse and addiction [2]. PDMU is defined as excessive use of digital media, the internet, or electronic communication leading to user dysfunction and harm to other individuals [2,3]. While PDMU is primarily characterized by harm to the user, emerging literature indicates that its impacts may extend beyond the individual through relational disruption, impaired role functioning, and reduced interpersonal presence (eg, phubbing) [3-5]. The PDMU umbrella term encompasses several conceptualizations, such as internet gaming disorder (IGD) or problematic gaming, social media addiction (SMA) or problematic social media use, and problematic internet use (PIU) or internet addiction. Among youth, problematic use of media is linked to various mental health disorders, behavioral problems, substance abuse, poor sleep hygiene, and social dysfunction [2,6,7]. These health outcomes can negatively impact this population's growth, development, and well-being, as well as their potential to lead active and productive adult lives.

While the body of evidence on PDMU is still emerging, this multifaceted issue is correlated with a multitude of psychological and behavioral conditions [2,3,7] affecting children and young people [8]. Given the ubiquity of digital media, specifically the internet, social media, and interactive gaming, concerns of overstimulation, distraction, and addiction are relevant as these media are easily accessed through handheld electronic devices, such as smartphones or tablets. Among young adults, such pervasive and continuous exposure (eg, gaming disorder) has been associated with difficulties in sustained attention, impaired prefrontal cortex functioning, diminished cognitive control, and potential disruptions in working memory processes [4,5]. However, the amount of time spent on digital media does not necessarily indicate dysfunctional use [9]. Instead, it is how the media is used—and its impact on well-being and daily life—that distinguishes acceptable, normal use from problematic use.

Researchers continue to examine related problematic technological issues, such as addiction to internet gaming [10,11], addiction to social media [12,13], and problematic use of the internet [2,14]. Addiction to gaming, or IGD, is described in the American Psychiatric Association's *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision*. The World Health Organization (WHO) included IGD in the *International Classification of Diseases, 11th Revision*, and defines it as a "pattern of gaming behavior ['digital-gaming' or 'video-gaming'] characterized by impaired control over gaming, increasing priority given to gaming over other activities to the extent that gaming takes precedence over other interests and daily activities, and continuation or escalation of gaming despite the occurrence of negative consequences" [15]. The latest systematic review on IGD, by Gao et al [16], showed that its prevalence was significantly higher among young adults than adolescents (10.4%, 95% CI 8.8%-11.9% vs 8.8%, 95% CI 7.5%-10%); however, the authors cautioned that their systematic review reflected post–COVID-19 activities, when there was a reduction in outdoor activities and increased risk of developing IGD. Indeed, several studies found significant increased incidence of IGD during the pandemic [2,15].

Another area of study in PDMU literature is SMA, described as individuals who experience addiction-like symptoms resulting from social media use, manifesting as extreme preoccupation and compulsion in social media platforms regardless of negative consequences [7]. Social media is highly popular and heavily used by young adults because this population is the first to grow up in an immensely digitalized society. Still, sustained engagement with social media is not without consequences. A recent study on college-aged students (N=571; mean age 23.61, SD 5.00 years) revealed that 22.7% were addicted to social media, scoring significantly higher on depressiveness, anxiety, and somatization than the group who was not addicted to social media [13]. Those addicted to social media were also found to exhibit higher use frequency, time spent, and perceived importance. These findings echo a recent systematic review's findings that showed moderate but statistically significant correlations between SMA and depression, anxiety, and stress among young adults and adolescents [7]. Unlike IGD, SMA is currently not officially recognized and classified as an addiction or mental disorder in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*, or *International Classification of Diseases, 11th Revision*, although research on this condition continues to expand.

PIU is a further subtype related to PDMU. PIU is an addictive behavioral condition characterized by repeated, uncontrollable online behavior that is risky, excessive, or impulsive, resulting in significant harm to the user [2,14]. Consequently, this behavior has been linked to physical and emotional functional impairments, relationship conflicts, compulsive internet use (ie, internet use that is difficult to control and continues despite negative effects on daily life), and mental health problems, such as depression. In a cross-cultural study (N=5130) examining PIU in 15 countries across Europe, South America, North America, and Asia, Lopez-Fernandez et al [14] found that emailing, social networking, shopping, and video streaming were the most common online activities prevalent among young adults with an average age of 24.71 (SD 8.70) years. The study participants tended to be single university students, and most used smartphones and computers or laptops, spending approximately 2 to 3 hours per day on online entertainment via computers and between 1 and 6 hours per day using smartphones. PIU frequency rates were high in English-speaking countries (ie, the United Kingdom and the United States); however, Asian countries were found to have the highest rates of PIU. Despite this, the researchers concluded that further research is needed with more representative samples for better generalization.

As research continues to expand on the divergent subtypes of PDMU [5], combining these subtypes into one comprehensive review provides a macrolevel view of the gravity of PDMU worldwide. Understanding the measured burden of this condition can inform the planning and development of targeted interventions aimed at mitigating risks associated with digital

XSL•FO
RenderX

addiction. This maladaptive behavior is pervasive among young people, yet there is a paucity of studies that comprehensively synthesize the phenomenon in this specific population. The young adulthood period is characterized by high levels of digital media engagement and vulnerability to problematic use behaviors [17]. Limited studies on sociodemographic factors, such as age, gender, and socioeconomic status, show variable associations with PDMU [12-14]. Research on problematic smartphone and internet use has identified a broad range of psychological and behavioral predictors (eg, negative affectivity, fear of missing out, and psychological well-being), whereas sociodemographic factors alone appear less frequently as primary predictors in the literature on young adults [18]. Due to this paucity, a comprehensive systematic review and meta-analysis will be conducted with the aim of (1) examining the global prevalence of PDMU among young adults and (2) exploring the extent of the issue across different regions of the world while identifying key factors contributing to its occurrence. In this study, the term "young adult" is used for individuals aged 18 to 24 years, consistent with classifications used by the Centers for Disease Control and Prevention [19] and the WHO [20].

Review Question

This review will address the following research questions:

1. What is the global prevalence of PDMU among young adults aged 18 to 24 years?
2. How does PDMU vary by age category, gender, country or region, educational level, and employment status?

Methods

Study Design

The proposed systematic review will be conducted in accordance with the methodology for systematic reviews of prevalence and incidence by the Prevalence Estimates Reviews – Systematic Review Methodology Group (PERSyst) [21]. The team used the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) guidelines to prepare this protocol [22] and will use the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) extension for abstracts [22] and the PRISMA 2020 flow diagram to show the results of the search and screening phases [23]. These actions are planned to ensure completeness in meeting the standards for systematic reviews.

Eligibility Criteria

This proposed systematic review uses the elements of the condition, context, and population framework as outlined in the Joanna Briggs Institute (JBI) Manual for Evidence Synthesis [24]. Summarized points are shown in Table 1.

XSL•FO
RenderX

Table . Inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Participants	• Young adults aged 18 to 24 years • Male, female, and nonbinary participants • Any sociodemographic characteristics	• Age <18 or >24 years • Electronic screen dependence due to speech or motor impairments
Condition	• Problematic use of smartphones, social media, online gaming, and the internet • Self-report instruments (ie, IAT[a], PIUQ[b], SMD[c], or GAS[d], as well as other tools) • Clinician diagnostic documentation (ie, *DSM-5*[e] [eg, for internet gaming disorder]; *ICD-11*[f] [eg, for gaming disorder]; clinical interviews that explore patterns of functional impairment, withdrawal, and loss of control associated with digital media use; and the SCID[g] or similar validated diagnostic interviews)	• Studies reporting general or positive digital media use
Context	• Global • Diverse settings, such as the home, the clinical setting, health care, the educational setting, the workplace, and community-based settings • All populations (the general population and marginalized populations [ie, homeless individuals, refugees, or those in transitional housing])	• Studies that focus on correctional or forensic institutions (eg, prisons or juvenile detention centers) and rehabilitation or detoxification centers
Type of study	• Analytical observational studies • Descriptive observational studies • Experimental designs that report prevalence at baseline or the screening or eligibility phase • Conference abstracts	• Qualitative studies • Case reports • Case series • Editorials and commentaries • Letters to the editor • Protocols • Opinion pieces • Synthesis reviews • Literature reviews
Search limits	• Any language • Last 5 years (2020-2025)	• None

[a]IAT: Internet Addiction Test.

[b]PIUQ: Problematic Internet Use Questionnaire.

[c]SMD: Social Media Disorder Scale.

[d]GAS: Gaming Addiction Scale.

[e]*DSM-5*: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*.

[f]*ICD-11*: *International Classification of Diseases, 11th Revision*.

[g]SCID: Structured Clinical Interview for DSM-5.

Search Strategy

The search strategy will aim to locate both published and unpublished studies. A 3-step search strategy will be used in this review. First, an initial search of Embase (Elsevier) was undertaken by a library scientist to identify articles on the topic. The text words contained in the titles and abstracts of relevant articles, the index terms used to describe the articles, and the synonyms suggested by this database will be used to develop the full search strategy (Multimedia Appendix 1). Studies published in any language and from 2020 to 2025 will be included to determine the currency of prevalence rates and trends of prevalence over time. Recent data account for the rapidly evolving technology ecosystems and reflect the current social media landscapes necessary for the recommendation of up-to-date and evidence-based interventions. Another library scientist will independently peer review the search strategy for completeness and rigor, adhering to the PRESS (Peer Review of Electronic Search Strategies) guidelines [25].

The second step of the search strategy will adapt the identified keywords and index terms from the initial search to suit the specific indexing and search functionalities of each included database and/or information source, which will comprise Embase (Elsevier), PubMed, PsycInfo (EBSCO), Web of

XSL•FO
RenderX

Science Core Collection (Clarivate), Communication & Mass Media Complete (EBSCO), LILACS (VHL Search Portal), China National Knowledge Infrastructure, and ProQuest Dissertations and Theses Global. Additionally, Google Scholar will be searched, and only the first 200 results will be considered as recommended [26].

The third stage of the search strategy will involve both backward and forward citation checking. Specifically, the reference lists of all included sources of evidence will be systematically screened to identify additional relevant studies. For forward citation checking, citation-tracking functionalities within search platforms, such as Web of Science and Google Scholar will be used to identify more recent articles that have cited our included studies. Comprehensive documentation of all search processes and reporting of results will adhere to the PRISMA-S (PRISMA extension for reporting literature searches in systematic reviews) statement, thereby ensuring the transparency, reproducibility, and comprehensiveness of the reporting details of the search methodology [27].

Study Screening and Selection

Following the 3-step search, all identified citations will be collated and uploaded into EndNote (version 21; Clarivate Analytics), and duplicates will be removed. After deduplication, potentially relevant studies will be retrieved, and their citation details will be imported into the JBI System for the Unified Management, Assessment, and Review of Information. Study screening and selection will follow a 3-step process to increase specificity and sensitivity.

First, we will carry out two actions to minimize title and abstract screening errors: (1) creating a 1-page summary of inclusion and exclusion criteria to use while voting and (2) conducting a pilot test of 10 titles and abstracts to calibrate clarity of understanding of the inclusion and exclusion of participants, concept, context, and types of studies. Any combination of 2 team members will review titles and abstracts for assessment against the inclusion criteria for the review.

In the second step, the full texts of reports of titles included from the previous step will be collected and reviewed. In the case of any studies published in languages other than English, they will be translated using the large language model DeepL Translate (DeepL SE); this artificial intelligence program has been shown to be highly accurate in translating scientific reports [28]. The full texts of selected reports will be assessed in detail against the inclusion criteria by 2 independent reviewers. The full citations and reasons for exclusion of papers that do not meet the inclusion criteria at the full-text step will be recorded and reported in the systematic review. Reasons for exclusion will include ineligible participants, condition, or context; ineligible study type; duplicate study; unavailability through interlibrary loan; or unavailability due to high costs, among others.

For the third step of screening, included studies from the previous step will undergo assessment of methodological quality using the JBI's critical appraisal tools [29]. Two independent reviewers will conduct the critical appraisal to determine the final studies included.

In each stage of the selection process, conflicts in voting between assessors will be reviewed and resolved through discussion or with one or more additional reviewers. The results of the search and the study inclusion process will be reported in full in the final systematic review and presented in a PRISMA flow diagram [23].

Assessment of Methodological Quality

Two or more independent reviewers will conduct a critical appraisal of the methodological quality of the included studies in this proposed systematic review using the standardized critical appraisal instrument from the JBI for prevalence studies. This critical appraisal tool is specifically designed to evaluate a study's methodological merit regardless of whether the design is cross-sectional or cohort based [24].

During critical appraisal, studies will be assessed using 3 grades: high risk (0%-33% of the criteria met), medium risk (34%-66% of the criteria met), and low risk (>67% of the criteria met). Studies that do not meet a certain quality threshold will be considered as having a high risk of bias and will be excluded from the extraction and synthesis phases. Studies with medium or low risk of bias will be included. Disagreements will be resolved through discussion or with a third reviewer. Results will be reported in a narrative format and risk-of-bias tables for each study design.

Data Extraction

To standardize the data extraction process, the team will use the JBI System for the Unified Management, Assessment, and Review of Information standardized tool to extract data [30]. The extraction template will capture key information on each study, such as (1) bibliometric characteristics, including country, year of publication, time frame for data collection, and language of publication; (2) participant characteristics, such as age, gender, ethnicity, level of education, population density (urban or rural), and other socioeconomic factors (additional characteristics may include mental health status, sleep patterns, personality traits, substance use, living situation, relationship status, digital media use habits [eg, screen time and type of media], and academic or employment status when available); (3) conditions and measurement methods used, such as study design, name of validated tools used to collect data, and type of digital device; and (4) description of the main results, including sample size, number of cases, prevalence, and CIs.

During extraction, any modifications to the tools will be detailed and justified in the narrative of the full review. To ensure consistency, the team will complete in-person training by extracting data from 2 studies together and comparing results. Once training has validated the skill level of the team, any 2 trained team members will extract the data. Any disagreements that arise between the reviewers will be resolved through discussion or with a third reviewer. In cases in which key information is missing, unclear, or contradictory, study authors will be contacted for clarification. If there is no response, data that remain missing or unclearly reported will be noted as "NR" (not reported) in the data extraction form.

XSL•FO
RenderX

Data Synthesis

Study data will, where possible, be pooled in a proportional meta-analysis following the PERSyst guidelines for processing, reporting, and interpretation [21]. The aim of this specific type of meta-analysis will be to pool study results, improve statistical power, and provide guidance on the global prevalence of PDMU. Results can provide evidence for clinicians, researchers, educators, and policymakers. The meta-analysis will be conducted using the PERSyst MA software (version 1.0) based on the R *meta* package (version 7.00; R Foundation for Statistical Computing) [31], and it will be organized following a 3-step approach.

First, outcome data will be standardized to a common effect estimate to ensure consistency. Two reviewers will cross-check the extracted data on prevalence for accuracy and consistency; variations in population labels and definitions of prevalence will be coordinated across studies to improve comparability. As our intention with the meta-analysis is to generalize the pooled prevalence estimate, we will use a random-effects model if there are 5 or more studies included. This is based on guidance by Tufanaru et al [32]. Statistical parameters will be input into the model using the Freeman-Tukey double arcsine transformation to stabilize variance in prevalence estimates. For fewer than 5 studies, a fixed-effects model will be used, and statistical heterogeneity will be checked. If heterogeneity is lower than 50%, the fixed-effects model will be appropriate. However, when heterogeneity is higher than 50%, a narrative approach would be preferred [33,34].

Second, we will conduct the meta-analyses using the selected model and use forest plots to display the distribution of individual and pooled prevalence estimates, corresponding 95% CIs, and measures of heterogeneity. To aid in interpretation, we will express results using percentages [35]. Outliers from the study results will be rechecked for coding errors. If the number of studies is small, we will prioritize point estimates, CIs, and the pooled estimate rather than relying on the chi-square test (Cochran *Q* statistic) for heterogeneity. The I^2 statistic will quantify heterogeneity, categorized as low (<25%), moderate (25% - 75%), or high (>75%) [33,34].

Third, the team will conduct post hoc analyses for sensitivity, subgroup performance, and publication bias. Sensitivity analysis will use the 1-study removal method to assess individual studies' impact on the pooled estimate. Subgroup analyses will explore heterogeneity when possible by stratifying data across different subtypes of PDMU, WHO's 6 geographic regions [36], settings, sample size, diagnostic tools used, and other important bibliometric and participant characteristics. If at least 10 studies are included, publication bias will be assessed using funnel plots in the Comprehensive Meta-Analysis software (version 3.3.070; Biostat, Inc), with asymmetry tested via the Egger test.

If meta-analysis is not possible due to too much variability in the extracted data or too few included studies, the findings will be presented in narrative format, tables, figures, and graphical displays to summarize the evidence and study characteristics descriptively.

Results

The study aims to examine the global prevalence and key factors of PDMU among young adults as this population is at risk of maladaptive behaviors and poor psychosocial health outcomes and underrepresented in the literature. The protocol has been registered in PROSPERO (registration CRD420251075912). The following databases were initially searched in April 2025: PROSPERO, MEDLINE, the Cochrane Database of Systematic Reviews, and *JBI Evidence Synthesis*, and no ongoing systematic reviews of prevalence on the topic were identified. The preliminary search yielded 1189 primary studies, suggesting sufficient primary research for the planned synthesis research.

The comprehensive search began in September 2025, with screening and selection of eligible studies following a 3-step search process. The process involved searching electronic bibliographic databases or information sources: Embase (Elsevier), PubMed, PsycInfo (EBSCO), Web of Science Core Collection (Clarivate), Communication & Mass Media Complete (EBSCO), LILACS (VHL Search Portal), China National Knowledge Infrastructure, ProQuest Dissertations and Theses Global, and Google Scholar. Once the search was completed, and after deduplication, a total of 2113 results were retrieved from the databases for title and abstract screening. The team began title and abstract screening in October 2025 using a 1-page summary of inclusion and exclusion criteria for voting. Conflicts were reviewed and resolved through discussion or with one or more additional reviewers. This phase of screening lasted until the beginning of January 2026. Subsequently, the team began reviewing the full texts of 1096 reports of titles included from the previous step. The full texts of selected reports are being assessed in detail against the inclusion criteria by 2 independent reviewers. We will provide a list of excluded studies from full-text screening in the form of a supplementary file. No deviations from the registered protocol have occurred.

To illustrate the inclusion process, a PRISMA flow diagram will be included. Additionally, results from the comprehensive review will be presented in tables to show general findings, risk of bias, and quality of the studies. Data extraction and synthesis will be conducted between March and April 2026. Subsequently, the manuscript will be prepared and is expected to be completed by the end of August 2026. We plan to submit the manuscript for publication in September 2026. Finally, to bring awareness of this pervasive issue, findings from the study will be disseminated at relevant conferences and shared with stakeholders, such as health care providers, educators, researchers, and public health professionals.

Discussion

Expected Findings

This protocol presents a plan for the synthesis of the existing literature on PDMU among young adults, encompassing related constructs such as PIU, IGD, and SMA. The review will identify a wide range of social media platforms and online games specifically designed to maximize user engagement, which may contribute to adverse health outcomes. Reported consequences of excessive digital media use include increased alcohol

XSL•FO
RenderX

consumption, sleep disturbances, and mental health issues [2]. Additionally, overlapping demographic and health-related patterns have been observed across PIU, IGD, and SMA. While substantial evidence exists linking digital media overuse to negative physical and mental health outcomes in school-aged populations, comparatively little is known about its effects on young adults. This gap is likely due to limited research attention directed toward this age group. Consequently, the young adult population was selected for this review due to its underrepresentation in the literature and the need for a more comprehensive understanding of how digital media use affects their health and well-being.

PDMU among young adults has reached increasingly concerning levels in recent years. However, research on screen time behaviors in this population remains limited, potentially due to their relatively infrequent engagement with health care providers. Despite this gap, a growing body of literature has documented the adverse physical and mental health outcomes associated with excessive media use in young adults [37]. For example, a recent systematic review by Brautsch et al [38] identified significant associations between digital media use and sleep onset latency, sleep duration, early awakening, sleep disturbances, daytime fatigue and dysfunction, sleep deficits, and overall sleep quality. These findings demonstrate that increased digital media use is consistently linked to poor sleep outcomes. Despite the professional guidelines and mounting evidence highlighting the risks of PDMU, many young adults continue to disregard recommended practices for healthy media consumption in their daily routines.

A recent survey conducted as part of the Millenium Cohort Study in the United Kingdom found that 48% of young adults (aged 18-24 years) self-identified as being “addicted” to social media. This finding is concerning as perceived addiction to digital platforms may signal long-term risks to the user’s mental and physical well-being. Additionally, an online cross-sectional survey conducted during the COVID-19 pandemic examined screen time behaviors among young adults (aged 18-28 years). The study found that time spent on specific screen-based activities, particularly social media and entertainment, was associated with increased symptoms of depression, anxiety, psychological distress, and substance abuse [39]. These findings raise significant public health concerns as the long-term health effects of PDMU are not yet well established given that research on this condition is still emerging. A comprehensive investigation into this issue is timely and necessary given the rapid advancement of digital technologies and their pervasive integration into the daily lives of young adults. Although prior studies have explored aspects of this phenomenon, the existing literature remains limited, particularly with respect to this age group.

Strengths and Limitations

The strengths of this protocol methodology include a rigorous approach to creating a comprehensive, unbiased, reproducible, and transparent product while identifying research gaps and providing evidence-based recommendations. First, quantitative designs enable objective measurement of digital media behaviors and associated psychological outcomes using validated scales and rigorous statistical analyses. This approach enhances the reliability and generalizability of the findings across diverse populations. However, quantitative methods also present limitations. They are typically less effective in capturing the contextual and subjective nuances of user experiences, have a limited ability to explore a topic, and often face challenges in establishing causality due to their cross-sectional or observational nature. Additionally, these studies may be subject to recall bias and social desirability bias, particularly when relying on self-reported data. Finally, the possibility of reporting bias and selective publication bias may lead to conflated findings of prevalence.

While a comprehensive search strategy will be used, relevant studies could be overlooked. Furthermore, despite using predefined criteria, judgments about the relevance and quality of the articles involve some level of subjectivity from the study team members. The completion of this systematic review protocol on PDMU is essential to ensure transparency, uphold methodological rigor, and maintain research integrity. Given the rapid evolution of digital media use and its psychosocial impacts, publishing a protocol ensures that the review follows best practices and meets the standards required for evidence-based policymaking and future research.

Conclusions and Recommendations

This review will add to the growing body of evidence on PDMU among young adults by emphasizing its potential impact on their overall well-being. To the best of our knowledge, this is the first comprehensive synthesis focused specifically on this age group. The findings will underscore the importance of implementing routine screening and early intervention strategies to mitigate the social, psychological, and physical health risks associated with digital media overuse in young adulthood.

Future research should prioritize interventional studies aimed at reducing digital media use among young adults while also providing evidence-based recommendations for achieving a balanced integration of screen time and green time in their daily routines.

Acknowledgments

The authors acknowledge using artificial intelligence (ChatGPT-4; OpenAI) for rephrasing, tutoring, and grammar refinement. However, all intellectual contributions, including idea development, analyses, and interpretations, are solely those of the authors. Artificial intelligence served as a tool for enhancing clarity, whereas all critical thinking and conclusions remain the authors’ original work.

XSL•FO
RenderX

Funding

The authors declare that no funding was provided for this study.

Data Availability

The corresponding author will provide access to the database upon request after the complete study has been published.

Authors' Contributions

The initial conceptual framework was developed by RMCP. The content experts are RMCP and FNT. Methodological expertise was provided by JMN and SS, with additional contributions by JD. ST designed the database search strategy. All authors actively participated in shaping the narrative of the protocol. RMCP serves as the guarantor responsible for the review's integrity.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy by database and platform, date searched, and results.
[DOCX File, 21 KB - resprot_v15i1e82245_app1.docx]

Checklist 1
PRISMA-P 2015 checklist.
[DOCX File, 37 KB - resprot_v15i1e82245_app2.docx]

References

1. Teens, screens and mental health. World Health Organization. 2024 Sep 25. URL: https://www.who.int/europe/news/item/25-09-2024-teens--screens-and-mental-health [accessed 2025-06-17]
2. Moreno M, Riddle K, Jenkins MC, Singh AP, Zhao Q, Eickhoff J. Measuring problematic internet use, internet gaming disorder, and social media addiction in young adults: cross-sectional survey study. JMIR Public Health Surveill 2022 Jan 27;8(1):e27719. [doi: 10.2196/27719] [Medline: 34081596]
3. Domoff SE, Borgen AL, Rye B, Barajas GR, Avery K. 12 problematic digital media use and addiction. In: Nesi J, Telzer EH, Prinstein MJ, editors. Handbook of Adolescent Digital Media Use and Mental Health: Cambridge University Press; 2022:300-316. [doi: 10.1017/9781108976237.016]
4. Kuss DJ. Online harms: problematic technology use is a public health concern and requires a multistakeholder approach. Addict Behav Rep 2025 Mar 29;21:100602. [doi: 10.1016/j.abrep.2025.100602] [Medline: 40248689]
5. Werling AM, Kuzhippallil S, Emery S, Walitza S, Drechsler R. Problematic use of digital media in children and adolescents with a diagnosis of attention-deficit/hyperactivity disorder compared to controls. A meta-analysis. J Behav Addict 2022 Jul 13;11(2):305-325. [doi: 10.1556/2006.2022.00007] [Medline: 35567763]
6. Austermann MI, Thomasius R, Paschke K. Assessing problematic social media use in adolescents by parental ratings: development and validation of the Social Media Disorder Scale for Parents (SMDS-P). J Clin Med 2021 Feb 6;10(4):617. [doi: 10.3390/jcm10040617] [Medline: 33561980]
7. Shannon H, Bush K, Villeneuve PJ, Hellemans KG, Guimond S. Problematic social media use in adolescents and young adults: systematic review and meta-analysis. JMIR Ment Health 2022 Apr 14;9(4):e33450. [doi: 10.2196/33450] [Medline: 35436240]
8. Colder Carras M, Aljuboori D, Shi J, et al. Prevention and health promotion interventions for young people in the context of digital well-being: rapid systematic review. J Med Internet Res 2024 Dec 18;26:e59968. [doi: 10.2196/59968] [Medline: 39693138]
9. Domoff SE, Harrison K, Gearhardt AN, Gentile DA, Lumeng JC, Miller AL. Development and validation of the problematic media use measure: a parent report measure of screen media "addiction" in children. Psychol Pop Media Cult 2019 Jan;8(1):2-11. [doi: 10.1037/ppm0000163] [Medline: 30873299]
10. Stevens MW, Dorstyn D, Delfabbro PH, King DL. Global prevalence of gaming disorder: a systematic review and meta-analysis. Aust N Z J Psychiatry 2021 Jun;55(6):553-568. [doi: 10.1177/0004867420962851] [Medline: 33028074]
11. Gentile DA, Bailey K, Bavelier D, et al. Internet gaming disorder in children and adolescents. Pediatrics 2017 Nov;140(Suppl 2):S81-S85. [doi: 10.1542/peds.2016-1758H] [Medline: 29093038]
12. Cheng C, Lau YC, Chan L, Luk JW. Prevalence of social media addiction across 32 nations: meta-analysis with subgroup analysis of classification schemes and cultural values. Addict Behav 2021 Jun;117:106845. [doi: 10.1016/j.addbeh.2021.106845] [Medline: 33550200]
13. Eichenberg C, Schneider R, Rumpl H. Social media addiction: associations with attachment style, mental distress, and personality. BMC Psychiatry 2024 Apr 15;24(1):278. [doi: 10.1186/s12888-024-05709-z] [Medline: 38622677]

XSL•FO
RenderX

14. Lopez-Fernandez O, Romo L, Kern L, et al. Problematic internet use among adults: a cross-cultural study in 15 countries. J Clin Med 2023 Jan 29;12(3):1027. [doi: 10.3390/jcm12031027] [Medline: 36769675]
15. Sherer J. Internet gaming. : American Psychiatric Association; 2023. URL: https://www.psychiatry.org/patients-families/internet-gaming [accessed 2026-02-18]
16. Gao YX, Wang JY, Dong GH. The prevalence and possible risk factors of internet gaming disorder among adolescents and young adults: systematic reviews and meta-analyses. J Psychiatr Res 2022 Oct;154:35-43. [doi: 10.1016/j.jpsychires.2022.06.049] [Medline: 35926424]
17. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. Am Psychol 2000 May;55(5):469-480. [Medline: 10842426]
18. Sánchez-Fernández M, Borda-Mas M. Problematic smartphone use and specific problematic internet uses among university students and associated predictive factors: a systematic review. Educ Inf Technol (Dordr) 2023;28(6):7111-7204. [doi: 10.1007/s10639-022-11437-2] [Medline: 36465425]
19. Leidman E, Duca LM, Omura JD, Proia K, Stephens JW, Sauber-Schatz EK. COVID-19 trends among persons aged 0-24 years - United States, March 1-December 12, 2020. MMWR Morb Mortal Wkly Rep 2021 Jan 22;70(3):88-94. [doi: 10.15585/mmwr.mm7003e1] [Medline: 33476314]
20. Adolescent health. World Health Organization. URL: https://www.who.int/southeastasia/health-topics/adolescent-health [accessed 2025-05-06]
21. PERSystMA app. PERSyst Prevalence Estimates Reviews. URL: https://persyst.group/persystma/ [accessed 2026-02-18]
22. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
23. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
24. Munn Z, Moola S, Lisy K, Riitano D, Tufanaru C. Methodological guidance for systematic reviews of observational epidemiological studies reporting prevalence and cumulative incidence data. Int J Evid Based Healthc 2015 Sep;13(3):147-153. [doi: 10.1097/XEB.0000000000000054] [Medline: 26317388]
25. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75:40-46. [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
26. Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Syst Rev 2017 Dec 6;6(1):245. [doi: 10.1186/s13643-017-0644-y] [Medline: 29208034]
27. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev 2021 Jan 26;10(1):39. [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
28. Kamaluddin MI, Rasyid MW, Abqoriyyah FH, Saehu A. Accuracy analysis of DeepL: breakthroughs in machine translation technology. J Engl Educ Forum 2024;4(2):122-126. [doi: 10.29303/jeef.v4i2.681]
29. Tufanaru C, Munn Z, Aromataris E, Campbell JM, Hopp LJ. Systematic reviews of effectiveness. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: Joanna Briggs Institute; 2020.
30. Munn Z, Aromataris E, Tufanaru C, et al. The development of software to support multiple systematic review types: the Joanna Briggs Institute System for the Unified Management, Assessment and Review of Information (JBI SUMARI). Int J Evid Based Healthc 2019 Mar;17(1):36-43. [doi: 10.1097/XEB.0000000000000152] [Medline: 30239357]
31. PERSystMA app. PERSyst Prevalence Estimates Reviews. URL: https://persyst.group/persystma/ [accessed 2025-06-15]
32. Tufanaru C, Munn Z, Stephenson M, Aromataris E. Fixed or random effects meta-analysis? Common methodological issues in systematic reviews of effectiveness. Int J Evid Based Healthc 2015 Sep;13(3):196-207. [doi: 10.1097/XEB.0000000000000065] [Medline: 26355603]
33. Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: Joanna Briggs Institute; 2020. [doi: 10.46658/JBIMES-20-01]
34. Higgins JP, Thomas J, Chandler J, et al. Cochrane handbook for systematic reviews of interventions version 6.5. Cochrane. 2024. URL: https://www.cochrane.org/authors/handbooks-and-manuals/handbook [accessed 2025-06-05]
35. Barker TH, Migliavaca CB, Stein C, et al. Conducting proportional meta-analysis in different types of systematic reviews: a guide for synthesisers of evidence. BMC Med Res Methodol 2021 Sep 20;21(1):189. [doi: 10.1186/s12874-021-01381-z] [Medline: 34544368]
36. Regional offices. World Health Organization. URL: https://www.who.int/about/who-we-are/regional-offices [accessed 2026-02-18]
37. Gupta C, Jogdand S, Kumar M. Reviewing the impact of social media on the mental health of adolescents and young adults. Cureus 2022 Oct 10;14(10):e30143. [doi: 10.7759/cureus.30143] [Medline: 36381882]
38. Brautsch LA, Lund L, Andersen MM, Jennum PJ, Folker AP, Andersen S. Digital media use and sleep in late adolescence and young adulthood: a systematic review. Sleep Med Rev 2023 Apr;68:101742. [doi: 10.1016/j.smrv.2022.101742] [Medline: 36638702]

XSL•FO
RenderX

39. Wiciak MT, Shazley O, Santhosh D. An observational report of screen time use among young adults (ages 18-28 years) during the COVID-19 pandemic and correlations with mental health and wellness: international, online, cross-sectional study. JMIR Form Res 2022 Aug 24;6(8):e38370. [doi: 10.2196/38370] [Medline: 35878157]

Abbreviations

IGD: internet gaming disorder
JBI: Joanna Briggs Institute
PDMU: problematic digital media use
PERSyst: Prevalence Estimates Reviews – Systematic Review Methodology Group
PIU: problematic internet use
PRESS: Peer Review of Electronic Search Strategies
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for reporting literature searches in systematic reviews
SMA: social media addiction
WHO: World Health Organization

Edited by A Schwartz; submitted 12.Aug.2025; peer-reviewed by J Golding, L Reid-Chassiakos; revised version received 30.Jan.2026; accepted 04.Feb.2026; published 05.Mar.2026.

Please cite as:
Chanmany Pastor RM, Nguyen Truax F, Sahin S, Donnelly J, Nick JM, Tamares S
Prevalence of Problematic Digital Media Use Among Young Adults: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e82245
URL: https://www.researchprotocols.org/2026/1/e82245
doi:10.2196/82245

© Rachelle Marie Chanmany Pastor, Fayette Nguyen Truax, Safiye Sahin, Janet Donnelly, Jan M Nick, Shanalee Tamares. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Determinants Associated With Pesticide Exposure in Patients With Head and Neck Cancer: Protocol for a Systematic Review

Gustavo Adolfo Sánchez-Ramírez[1*], DDS; Manuel López Cabanillas Lomelí[2*], PhD; Blanca Edelia González-Martínez[2*], PhD; Ana Elisa Castro-Sánchez[2*], PhD; Myriam Angélica De la Garza-Ramos[1*], PhD; Guillermo Cano-Verdugo[2*], PhD

[1]Universidad Autónoma de Nuevo León, Facultad de Odontología, Monterrey, Nuevo León, Mexico

[2]Universidad Autónoma de Nuevo León, Facultad de Salud Pública y Nutrición, Calle Dr. Eduardo Aguirre Pequeño No. 905, Col. Mitras Centro, CP., Monterrey, Nuevo León, México

[*]all authors contributed equally

Corresponding Author:
Blanca Edelia González-Martínez, PhD
Universidad Autónoma de Nuevo León, Facultad de Salud Pública y Nutrición, Calle Dr. Eduardo Aguirre Pequeño No. 905, Col. Mitras Centro, CP., Monterrey, Nuevo León, México

Abstract

Background: Currently, head and neck cancer (HNC) associated with pesticide exposure represents a global public health concern. However, there is no consensus regarding the specific determinants involved in this association. Moreover, there is a lack of scientific evidence to support the development of systematic reviews on this topic.

Objective: This study aims to synthesize the methodology for conducting a systematic review to explore the current scientific evidence on the determinants of HNC associated with pesticides.

Methods: The review will follow the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) guidelines to ensure methodological rigor. The protocol includes detailed steps for constructing a robust search strategy using relevant databases such as PubMed, Embase, Scopus, Web of Science, CINAHL, LILACS, and AGRICOLA. Keywords and Medical Subject Headings terms related to “pesticides,” “exposure,” “head and neck neoplasms,” and related concepts will be used to capture the most relevant studies. Eligibility criteria will be clearly defined, including study design (eg, cohort, case-control, and cross-sectional), population characteristics, exposure assessment, and cancer outcomes. Studies published in any language that involve human participants will be included. The studies will be screened in 2 phases: first by title and abstract and then by full-text review. Two independent reviewers will assess the quality of each study and extract key data, such as exposure levels, cancer subtypes, and effect sizes. Any disagreements will be resolved through discussion or by a third reviewer.

Results: This systematic review was initiated in February 2025 after protocol registration. The literature search and review process is ongoing at the time of submission of this paper. Data extraction and quality assessment have not yet been completed. Final results of the review, including a synthesis of determinants associated with pesticide exposure in patients with HNC, are expected to be completed and submitted for publication in June 2026.

Conclusions: The necessary steps to conduct a systematic review must be concise and publicly available to ensure replicability within the scientific community.

International Registered Report Identifier (IRRID): PRR1-10.2196/81312

(JMIR Res Protoc 2026;15:e81312) doi:10.2196/81312

KEYWORDS

occupational exposure; environmental risk factors; agrochemicals; rural health; cancer prevention; vulnerable populations

Introduction

Head and neck cancer (HNC) comprises a heterogeneous group of malignancies arising from the upper aerodigestive tract and represents a significant global public health concern, with more than 930,000 new cases and over 460,000 deaths reported worldwide in 2020 [1]. Within this group, oral cancer is one of the most common subtypes, accounting for nearly 377,000 cases annually. The increasing incidence of HNC and its subtypes, particularly in low- and middle-income countries, underscores the urgent need to understand the role of environmental and occupational determinants in their development [2].

According to the World Health Organization’s (WHO) Commission on Social Determinants of Health framework, health inequities arise from the unequal distribution of structural and intermediary determinants within societies. Structural

XSL•FO
RenderX

determinants—such as socioeconomic status, education, occupation, and governance—shape the conditions of daily life, whereas intermediary determinants include material circumstances, psychosocial factors, and biological pathways through which exposures affect health outcomes. In the context of HNC, these determinants intersect with environmental and occupational exposures, such as pesticides, to modulate vulnerability and disease risk, particularly among rural and agricultural populations. Health outcomes are shaped by a complex interaction of biological, behavioral, social, and environmental determinants that influence individual and population well-being [3,4]. Among these, environmental exposures, especially to toxic or carcinogenic compounds, are increasingly recognized as major contributors to cancer etiology [5].

Pesticides, widely used in agricultural activities and vector-control programs in public health, represent one of the most relevant environmental exposures due to their potential mutagenic and carcinogenic effects. Chronic exposure to certain organophosphates, organochlorines, and carbamates has been associated with an increased risk of several malignancies, including leukemia, prostate cancer, and breast cancer. However, their specific contribution to oral carcinogenesis remains less explored. Mechanistically, pesticide exposure may promote HNC development through oxidative stress, DNA damage, epigenetic alterations, and chronic inflammatory pathways. Lipophilic compounds can accumulate in mucosal tissues, inducing cytotoxic and genotoxic effects that contribute to malignant transformation [6,7].

Despite emerging evidence suggesting a link between pesticide exposure and HNC risk, findings remain fragmented and inconsistent across studies, often due to methodological heterogeneity and limited population-based analyses. Therefore, this document presents a protocol for a systematic review aiming to synthesize the existing scientific evidence on the determinants associated with pesticide exposure in patients with HNC. This review seeks to clarify the potential causal mechanisms and to guide preventive strategies in public and occupational health contexts.

Methods

Protocol Registration

The review will follow the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) [8] guidelines (Checklist 1) and is registered in PROSPERO (ID: CRD42024618939). This protocol was registered on December 6, 2024, and has 2 current versions (1 and 1.1). Through this protocol, we established the following focus question: What are the determinants associated with pesticide exposure in patients with HNC?

Eligibility Criteria

Eligibility criteria will be based on the PECOS (population, exposure, comparator, outcome, and study) framework [9], considering the population as individuals who will have a confirmed diagnosis of HNC. Populations diagnosed exclusively with cancers other than HNC will be excluded. The primary clinical condition under investigation will be head and neck malignancies.

Exposure will be defined as populations with direct or indirect exposure to pesticides, regardless of the type or context of use. Studies where no form of pesticide exposure will be identified will not be considered. This decision aligns with previously reported challenges in evaluating low-level or undocumented exposure risks. Comparators will be considered as an unexposed or low-exposure population. Studies without a comparator will be included via consensus and only if they provide sufficient data to support synthesis.

Outcome will be HNC and will include demographic and exposure-related variables, such as age, sex, geographic location, ethnic background, occupational history, and substance use. Studies that will focus on nonrelevant variables, such as marital status, religious affiliation, number of offspring, sexual orientation, or personal hobbies, will be excluded. These demographic and behavioral variables will be considered essential for understanding the context of exposure and risk, consistent with frameworks for environmental epidemiology and cancer risk profiling.

Studies published within the last 20 years will be included to ensure the relevance and currency of the scientific evidence. This time frame is considered sufficient to capture the most significant developments in the field while excluding potentially outdated data that may no longer reflect current practices, technologies, or knowledge. Language restriction will not be considered, as we will work with a translation workflow if required. This will allow for a broad temporal scope and the inclusion of both historical and recent data.

Study design will include original research studies, such as cohort, case-control, and cross-sectional. Literature such as review articles, news media pieces, and academic theses will be excluded. The selection will be limited to sources that will offer original data and contribute to a robust and credible evidence base, as recommended in systematic review methodology standards.

Search Strategy and Databases

A structured and reproducible search strategy will be developed and implemented across MEDLINE (via PubMed or Ovid), Embase, Scopus, Web of Science, CINAHL, LILACS, and AGRICOLA to identify relevant literature assessing the determinants in patients with HNC associated with pesticides, with a particular focus on cancers of the oral cavity. The search terms will be combined using Boolean operators to maximize sensitivity. It is anticipated that the search will cover literature from the 2000s to the present. The full search strategy and syntax are detailed in Table 1. Gray literature strategy will not be considered as consensus between authors.

XSL•FO
RenderX

Table . Proposed search strategy per database.

Database	Search strategy
PubMed	(“Head and Neck Neoplasms”[Mesh] OR “Oral Cancer”[tiab] OR “Laryngeal Cancer”[tiab] OR “Pharyngeal Cancer”[tiab] OR “Mouth Neoplasms”[tiab]) AND (“Pesticides”[Mesh] OR pesticide*[tiab] OR herbicide*[tiab] OR insecticide*[tiab] OR fungicide*[tiab] OR agrochemical*[tiab]) AND (“Environmental Exposure”[Mesh] OR “Occupational Exposure”[Mesh] OR “Agricultural Workers’ Diseases”[Mesh] OR determinant*[tiab] OR “environmental determinant*"[tiab])
Embase	('head and neck tumor’/exp OR ’oral cancer’/exp OR ’larynx cancer’/exp OR ’pharynx cancer’/exp) AND ('pesticide’/exp OR ’herbicide’/exp OR ’insecticide’/exp OR agrochemical*:ti,ab) AND ('environmental exposure’/exp OR ’occupational exposure’/exp OR determinant*:ti,ab)
Scopus	(TITLE-ABS-KEY(“HNC” OR “oral cancer” OR “pharyngeal cancer” OR “laryngeal cancer”)) AND (TITLE-ABS-KEY(pesticide* OR herbicide* OR insecticide* OR fungicide* OR agrochemical*)) AND (TITLE-ABS KEY(“environmental determinant*” OR “occupational exposure” OR “environmental exposure” OR determinant*))
Web of Science	TS=(“HNC” OR “oral cancer” OR “pharyngeal cancer” OR “laryngeal cancer”) AND TS=(pesticide* OR herbicide* OR insecticide* OR fungicide* OR agrochemical*) AND TS=(“environmental determinant*” OR “occupational exposure” OR “environmental exposure” OR determinant*)
CINAHL	(“Head and Neck Neoplasms” OR “Oral Cancer” OR “Pharyngeal Cancer” OR “Laryngeal Cancer”) AND (pesticide* OR herbicide* OR insecticide* OR fungicide* OR agrochemical*) AND (“Environmental Exposure” OR “Occupational Exposure” OR determinant*)
LILACS	((“neoplasias de cabeza y cuello” OR “cáncer oral” OR “cáncer de laringe” OR “cáncer de faringe”) AND (plaguicida* OR pesticida* OR herbicida* OR insecticida* OR agroquímico*) AND (“exposición ambiental” OR “exposición ocupacional” OR determinante*)
AGRICOLA	(“pesticides” OR “herbicides” OR “insecticides” OR “agrochemicals”) AND (“HNC” OR “oral cancer” OR “pharyngeal cancer” OR “laryngeal cancer”) AND (“occupational exposure” OR “environmental exposure” OR determinant*)

Planned Outcomes and Variables to Be Extracted

In this systematic review, a set of critical and important outcomes will be extracted from the included studies to ensure a comprehensive analysis of determinants in patients with HNC associated with pesticides. Each variable will be clearly defined according to its analytical role in the review. Sociodemographic and occupational characteristics (eg, age, sex, occupation, place of residence, marital status, and ethnicity) will be treated as descriptive variables, whereas pesticide exposure time, type of pesticide, and exposure level will be considered primary determinants (exposures). Potential confounders such as smoking habits, alcohol consumption, and human papillomavirus infection status will be included where available to allow for adjusted interpretation of associations. Family medical history will also be considered as a relevant covariate.

Additionally, a set of contextual and methodological variables will be collected to support the interpretation and synthesis of findings. These include the database, author, and country of the study, and the participants’ access to health care as well as the objective of the study, the outcomes reported, and the key findings described in each publication.

Data Collection and Selection Process

The selection process will follow a manual approach comprising title screening, abstract review, and full-text analysis. Data will be extracted into Microsoft Excel using the prespecified variables (eg, database, country, age, sex, occupation, exposure time, and pesticide type). Unavailable data will be requested from the study authors. An external researcher will serve as an impartial referee to resolve discrepancies.

Primary variables will include age, sex, occupation, duration of pesticide exposure, family history, and pesticide type. Structural stratifiers (eg, place of residence, occupation, ethnicity, and access to care) and intermediate determinants (eg, participant demographics and exposure specifics) will guide the narrative synthesis. Methodological and clinical heterogeneity will be assessed.

Planned Presentation of Results

The results of the included studies will be systematically organized and synthesized in a comprehensive evidence table. This table will summarize key characteristics and findings extracted from each study, allowing for comparisons across study designs, populations, and exposure patterns. For each study, the following data will be reported to ensure a comprehensive and structured synthesis of the evidence:

XSL•FO
RenderX

database and citation information, study design, participant demographics, geographic and occupational context, exposure details, ethnicity and lifestyle factors, health history and access to care, and study objectives and outcomes. These elements will facilitate consistent comparison across studies and support the identification of patterns, determinants, and gaps in the existing literature.

In addition, a narrative synthesis of the data will be conducted to complement the quantitative and qualitative findings. This synthesis will be structured according to the WHO conceptual framework of social determinants of health [10]: (1) structural determinants, which include broader sociopolitical and economic factors, such as income, education, occupation, social class, gender norms, and policies that influence health equity; and (2) intermediate determinants, which refer to the material circumstances, psychosocial factors, behavioral and biological factors, and access to health services that are shaped by structural determinants and directly impact individual health outcomes. This process will be performed by 2 authors and will have a peer-review process.

This thematic framework will guide the interpretation of the findings, helping to identify patterns and relationships between social determinants and the health outcomes reported in the included studies. The use of the WHO framework ensures a comprehensive and globally recognized approach to analyzing the complex social factors that influence health.

Heterogeneity, Risk of Bias, and Certainty Assessment

The methodological heterogeneity of the included studies will be assessed based on the type of study design and the specific inclusion and exclusion criteria applied. Clinical heterogeneity will be examined by evaluating participant characteristics, such as age, sex, and comorbidities [11].

To assess the risk of bias in nonrandomized studies, we will use the ROBINS-I (Risk of Bias in Non-Randomized Studies of Interventions) tool, developed by the Cochrane Bias Methods Group [12]. ROBINS-I evaluates bias across 7 domains: confounding, participant selection, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selection of the reported result. Each domain is rated as low, moderate, severe, or critical risk of bias, or as having no information. This tool is particularly appropriate when synthesizing evidence from observational or quasi-experimental studies.

The certainty of evidence for each outcome will be assessed using the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) approach [13].[13] GRADE considers 5 domains that may lower the certainty of the evidence: risk of bias, inconsistency, indirectness, imprecision, and publication bias. The quality of evidence will be categorized as high, moderate, low, or very low, providing a transparent and structured assessment of confidence in the effect estimates. This process will guide the strength and reliability of the conclusions drawn in the review.

Ethical Considerations

This study did not require ethics committee review or approval because it is a systematic review based only on an analysis of previously published studies. No primary data were collected and no direct human participation took place. Therefore, ethical approval and informed consent were not applicable.

Results

This systematic review was initiated in February 2025 after protocol registration. The literature search and review process is ongoing at the time of submission of this paper. Data extraction and quality assessment have not yet been completed. Final results of the review, including a synthesis of determinants associated with pesticide exposure in patients with HNC, are expected to be completed and submitted for publication in June 2026.

The findings of this systematic review will be presented through a structured synthesis of the included studies. A comprehensive evidence table will summarize essential study characteristics, such as database and citation information, study design, sample size, participant demographics (eg, age, sex, and ethnicity), geographic and occupational context, pesticide exposure type and duration, and cancer subtype diagnosed. Outcomes of interest, including critical and contextual variables, will be extracted and reported systematically.

Patterns in pesticide exposure and their potential determinants with HNC will be analyzed. The review will seek to identify common determinants among affected populations, such as occupational exposure, geographic distribution, and sociodemographic characteristics. Stratified analyses will be conducted, when possible, based on key variables such as pesticide type or exposure duration. Discrepancies between studies and data gaps will be noted, and findings will be synthesized narratively and, if feasible, quantitatively through meta-analysis.

A flow diagram following PRISMA-P guidelines will be used to report the selection process, including the number of records identified, screened, included, and excluded, along with reasons for exclusion (Figure 1). Methodological and clinical heterogeneity will be described, and the risk of bias and certainty of evidence will be reported for each study included. The results are intended to be published in December 2026.

XSL•FO
RenderX

Figure 1. PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) selection flowchart.

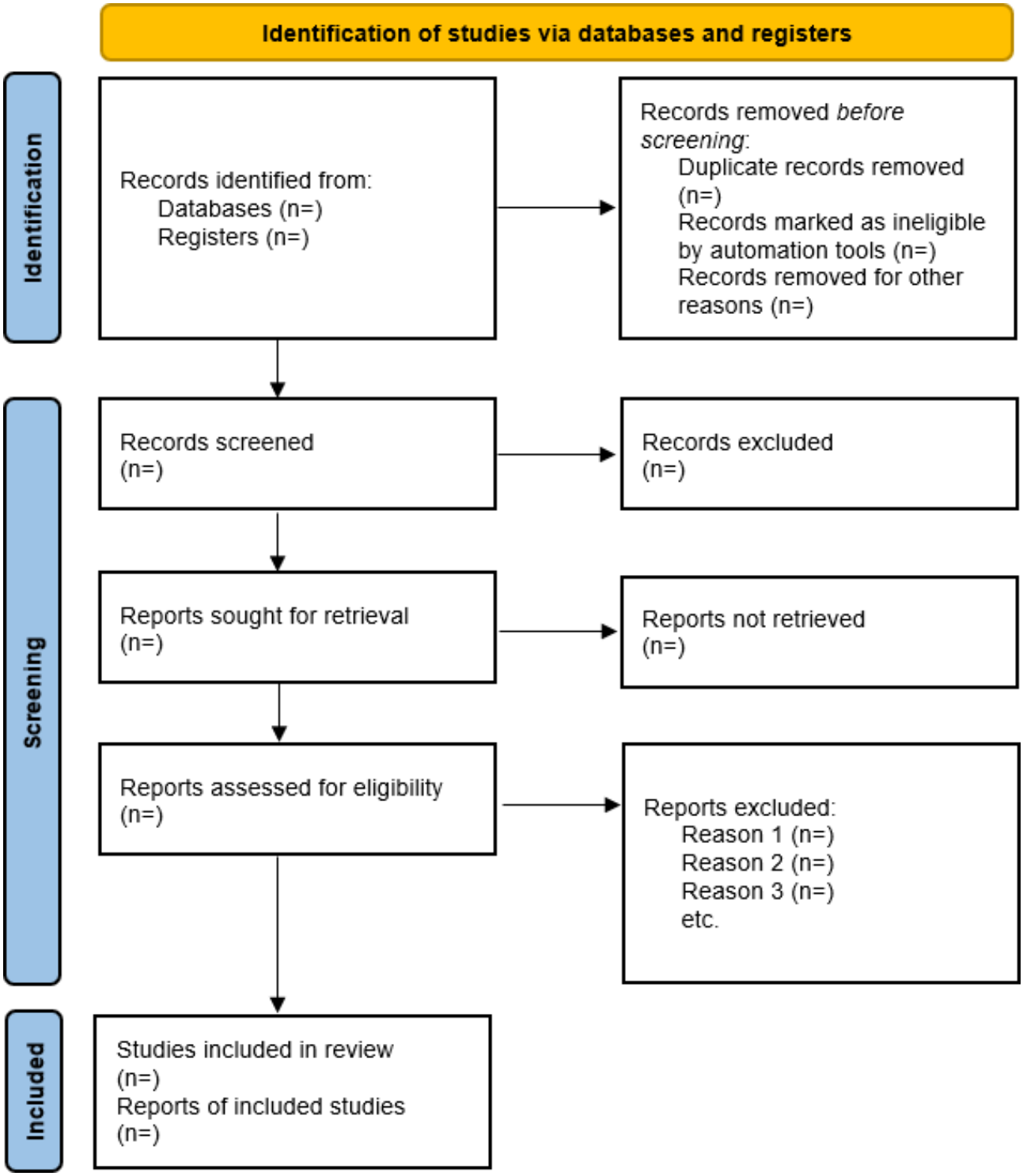

Discussion

Anticipated Findings

The systematic review will provide a detailed synthesis of the current scientific evidence regarding the association between pesticide exposure and HNC. It will examine the role of environmental determinants—particularly those linked to agricultural or occupational settings—in the development of malignancies of the oral cavity, pharynx, and larynx. Through an in-depth evaluation of the literature, the review will explore whether consistent patterns emerge in terms of risk factors, populations affected, and exposure characteristics.

Previous studies have suggested that prolonged exposure to organophosphates, carbamates, and other agricultural pesticides may increase the risk of developing head and neck malignancies. For instance, epidemiological investigations conducted in agricultural communities have reported higher incidence rates of laryngeal and oral cancers among workers exposed to these compounds compared to nonexposed populations. Similarly, systematic reviews and meta-analyses have identified a potential dose-response relationship between cumulative pesticide

XSL•FO
RenderX

exposure and cancer risk, although the strength of this association remains inconsistent across studies.

Findings will be interpreted considering the methodological quality and clinical diversity of the included studies. Particular attention will be given to how structural determinants—such as rural residency, type of employment, or limited access to health care—may modulate risk in vulnerable populations. The review will also discuss the potential influence of behavioral and intermediate factors such as tobacco or alcohol use, which may interact with pesticide exposure and contribute to cancer risk.

The discussion will critically assess limitations of the available evidence, including possible publication bias, insufficient exposure assessment methods, or lack of adjustment for confounders. Additional limitations may include heterogeneity in the definition and measurement of pesticide exposure across studies, retrospective study designs that limit causal inference, and underreporting of environmental or occupational data in low- and middle-income countries. Furthermore, the variability in diagnostic criteria for HNC subtypes could contribute to inconsistencies in observed associations. These limitations will make recommendations for future research, highlighting the need for more robust, longitudinal studies that integrate detailed environmental and occupational exposure data. Future investigations should also aim to standardize exposure metrics, incorporate biomonitoring approaches, and control for behavioral and genetic modifiers of susceptibility to better elucidate the causal pathways involved.

Ultimately, the results of this systematic review will contribute to a better understanding of the environmental determinants of HNC, informing the development of targeted public health strategies and occupational safety policies. The synthesis of current knowledge will also support clinicians and researchers in identifying at-risk populations and formulating hypotheses for future investigations.

Conclusions

This systematic review will aim to fill a critical gap in the scientific literature by synthesizing existing evidence on determinants in patients with HNC associated with pesticides. Through a rigorous and transparent methodology, the review will identify key environmental and demographic determinants that may contribute to cancer risk in exposed populations. The findings are expected to provide a clearer understanding of how pesticide exposure, particularly in occupational and rural settings, may influence the development of malignancies in the oral cavity, pharynx, and larynx. The review will also highlight the importance of integrating environmental health perspectives into cancer prevention strategies and public health policy.

Ultimately, this work will support the identification of high-risk groups, inform regulatory decision-making regarding pesticide use, and guide future research aimed at mitigating environmental carcinogenic exposures. By establishing a structured and replicable approach, the review will contribute to the scientific community's ability to address complex interactions between environmental hazards and cancer epidemiology.

Acknowledgments

The authors would like to thank the academic and research teams from the Universidad Autónoma de Nuevo León for their institutional support during the development of this protocol. No artificial intelligence was used during the development of this manuscript.

Funding

This research protocol was not supported by any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The study was entirely self-funded by the authors.

Data Availability

This is a protocol for a systematic review; therefore, no datasets were generated or analyzed during this study. All data used will be publicly available from previously published sources upon completion of the review.

Authors' Contributions

GAS-R and GC-V conceptualized the study and led the drafting of the protocol. MLCL and BEG-M contributed to the development of the search strategy and inclusion criteria. AEC-S and MADlG-R assisted in the methodological design and critical revision of the manuscript. All authors reviewed and approved the final version of the manuscript.

Guillermo Cano-Verdugo is the co-corresponding author and can be reached at Universidad Autónoma de Nuevo León Facultad de Salud Pública y Nutrición, Calle Dr. Eduardo Aguirre Pequeño No. 905, Col. Mitras Centro Monterrey, Nuevo León, 64460, México Phone: 52 81 8329 4000 ext 3046 Email: guillermo.canovrd@uanl.edu.mx

Conflicts of Interest

None declared.

Checklist 1
PRISMA-P checklist.

XSL•FO
RenderX

[PDF File, 261 KB - resprot_v15i1e81312_app1.pdf]

References

1. Sun H, Yu M, An Z, et al. Global burden of head and neck cancer: epidemiological transitions, inequities, and projections to 2050. Front Oncol 2025;15:1665019. [doi: 10.3389/fonc.2025.1665019] [Medline: 41079071]
2. Barsouk A, Aluru JS, Rawla P, Saginala K, Barsouk A. Epidemiology, risk factors, and prevention of head and neck squamous cell carcinoma. Med Sci (Basel) 2023 Jun 13;11(2):42. [doi: 10.3390/medsci11020042] [Medline: 37367741]
3. Goldstein SP, Tovar A, Espel-Huynh HM, Stowers KC. Applying a social determinants of health framework to guide digital innovations that reduce disparities in chronic disease. Psychosom Med 2023 Sep 1;85(7):659-669. [doi: 10.1097/PSY.0000000000001176] [Medline: 36800264]
4. Bolster-Foucault C, Ho Mi Fane B, Blair A. Structural determinants of stigma across health and social conditions: a rapid review and conceptual framework to guide future research and intervention. Health Promot Chronic Dis Prev Can 2021 Mar;41(3):85-115. [doi: 10.24095/hpcdp.41.3.03] [Medline: 33688694]
5. Zhu Y, Costa M. Metals and molecular carcinogenesis. Carcinogenesis 2020 Sep 24;41(9):1161-1172. [doi: 10.1093/carcin/bgaa076] [Medline: 32674145]
6. Salazar-Flores J, Lomelí-Martínez SM, Ceja-Gálvez HR, Torres-Jasso JH, Torres-Reyes LA, Torres-Sánchez ED. Impacts of pesticides on oral cavity health and ecosystems: a review. Int J Environ Res Public Health 2022 Sep 7;19(18):11257. [doi: 10.3390/ijerph191811257] [Medline: 36141526]
7. Ojha A, Shekhar S, Gupta P, Jaiswal S, Mishra SK. Comparative study of oxidative stress in cancer patients occupationally exposed to the mixture of pesticides. Discov Oncol 2024 Oct 5;15(1):526. [doi: 10.1007/s12672-024-01235-2] [Medline: 39367924]
8. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160. [doi: 10.1136/bmj.n160] [Medline: 33781993]
9. Morgan RL, Whaley P, Thayer KA, Schünemann HJ. Identifying the PECO: a framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. Environ Int 2018 Dec;121(Pt 1):1027-1031. [doi: 10.1016/j.envint.2018.07.015] [Medline: 30166065]
10. Solar O, Irwin A. A Conceptual Framework for Action on the Social Determinants of Health: World Health Organization; 2010. URL: https://www.who.int/publications/i/item/9789241500852
11. Gagnier JJ, Kienle G, Altman DG, et al. The CARE guidelines: consensus-based clinical case reporting guideline development. Glob Adv Health Med 2013 Sep;2(5):38-43. [doi: 10.7453/gahmj.2013.008] [Medline: 24416692]
12. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016 Oct 12;355:i4919. [doi: 10.1136/bmj.i4919] [Medline: 27733354]
13. Prasad M. Introduction to the GRADE tool for rating certainty in evidence and recommendations. Clin Epidemiol Glob Health 2024 Jan;25:101484. [doi: 10.1016/j.cegh.2023.101484]

Abbreviations

GRADE: Grading of Recommendations, Assessment, Development, and Evaluations
HNC: head and neck cancer
PECOS: population, exposure, comparator, outcome, and study
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
ROBINS-I: Risk of Bias in Non-Randomized Studies of Interventions I
WHO: World Health Organization

Edited by J Sarvestan; submitted 25.Jul.2025; peer-reviewed by H Amqam, T Li; revised version received 26.Nov.2025; accepted 27.Nov.2025; published 22.Jan.2026.

Please cite as:
Sánchez-Ramírez GA, López Cabanillas Lomelí M, González-Martínez BE, Castro-Sánchez AE, De la Garza-Ramos MA, Cano-Verdugo G
Determinants Associated With Pesticide Exposure in Patients With Head and Neck Cancer: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e81312
URL: https://www.researchprotocols.org/2026/1/e81312
doi:10.2196/81312

© Gustavo Adolfo Sánchez-Ramírez, Manuel López Cabanillas-Lomelí, Blanca Edelia González-Martínez, Ana Elisa Castro-Sánchez, Myriam Angélica De la Garza-Ramos, Guillermo Cano-Verdugo. Originally published in JMIR Research

XSL•FO
RenderX

Protocols (https://www.researchprotocols.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Frameworks for Guiding the Selection of Digital Data Collection Tools Used in Clinical Trials: Protocol for a Systematic Review

Bavatharani Ananthesayanan[1], MSc; Kayode Philip Fadahunsi[2], PhD; Huanhuan Xiong[1], PhD; John O'Donoghue[1,3,4], PhD

[1]Department of Business Information Systems, University College Cork, College Road, Cork, Ireland

[2]Department of Primary Care and Public Health, The School of Public Health, Imperial College London, London, England, United Kingdom

[3]Malawi eHealth Research Centre, University College Cork, Cork, Ireland

[4]School of Medicine, University College Dublin, Dublin, Ireland

Corresponding Author:
Bavatharani Ananthesayanan, MSc
Department of Business Information Systems, University College Cork, College Road, Cork, Ireland

Abstract

Background: Data collection is an essential aspect of clinical trials because it forms the basis of the scientific analysis that evaluates the performance and safety of interventions. With the wide variety of digital data collection tools available, decision-makers responsible for choosing the appropriate tools for clinical trials must exercise caution. There are numerous challenges that could impact data collection, and a careful selection of tools is necessary to ensure that they effectively support the trial. Therefore, an evidence-based framework is needed to support the selection of an appropriate digital data collection tool in clinical trials.

Objective: This systematic review aims to develop an evidence-based framework for the selection of digital data collection tools for clinical trials.

Methods: Bibliographic databases including IEEE Xplore, eAIS, PubMed, CINAHL, MEDLINE, Embase, ClinicalTrials.gov, Scopus, and Web of Science will be searched for published articles. Additionally, searches will be performed for publicly available gray literature from reputable institutions such as the United States Food and Drug Administration and World Health Organization. Studies should include a framework that is relevant to selecting digital data collection tools for clinical trials. Two reviewers will independently use Covidence to screen and review the articles to be included. Data related to the selection of digital data collection tools will be extracted. Thematic synthesis will be conducted to develop a new evidence-based framework to select digital data collection tools for clinical trials.

Results: The review started in May 2025 and is expected to be completed in December 2025. The searches yielded 9151 studies, which were reduced to 4333 after the removal of duplicates using Covidence.

Conclusions: There is a dearth of established frameworks to guide the selection of digital data collection tools for clinical trials. This review aims to develop an evidence-based framework to support technology decision-makers in identifying and selecting tools that are fit-for-purpose, ensuring they meet the specific needs of clinical research settings.

Trial Registration: PROSPERO CRD420250612895; https://www.crd.york.ac.uk/PROSPERO/view/CRD420250612895

International Registered Report Identifier (IRRID): DERR1-10.2196/78529

(JMIR Res Protoc 2026;15:e78529) doi:10.2196/78529

KEYWORDS

clinical trials; data collection tool; framework; digital health technology; health technology selection; digital data collection

Introduction

Selection of Digital Data Collection Tools in Clinical Trials

There is a growing momentum toward digitizing human clinical trials as they become more complex [1,2] Nowadays, pharmaceutical and biotech companies turn to technological solutions to lower costs, streamline study management, keep study information current, and deliver new interventions [3,4]. Digital health technologies (DHTs) are used to accelerate trial startup and execution [5]. DHTs allow for the rapid collection of efficacy and safety data on interventions [6]. Data collection is a key component of clinical trials, as data form the basis of the scientific analysis that informs researchers regarding the performance of their interventions. Digital data collection tools are designed to provide a structured and consistent way of gathering information to reduce errors and bias, which can affect

XSL•FO
RenderX

the validity and reliability of research findings [7]. When data collection tools are not properly selected in clinical trials, they can cause substantial disruptions to trial management and execution and hence delay the approval of the intervention by the regulatory bodies [8].

With the wide variety of digital data collection tools available, technology decision-makers responsible for selecting the appropriate tools for clinical trials must exercise caution. There are numerous challenges that may impact data collection, and careful selection is necessary to ensure the tools effectively support the trial [1]. The technology decision-makers need to consider the use of DHTs from a multistakeholder perspective to ensure the selected technology is fit-for-purpose. Clinical trials differ from other research settings due to their rigorous requirements for data quality, patient safety, and regulatory compliance. Credible patient data must be collected in accordance with the Good Clinical Practice (GCP) guidelines [9-11]. Additionally, clinical trials progress through distinct phases, each designed to address specific research questions, from early safety and dosing to large-scale efficacy and postmarketing surveillance [12]. Costs typically increase substantially in phase 2 and phase 3 trials due to the larger number of participants and the global scope of these studies, often reaching several million dollars [2]. These highlight the need for a dedicated framework to guide decision-makers in selecting the right technology to ensure both data integrity and patient safety are in full compliance with the GCP guidelines.

Although different frameworks have been developed for DHTs, none are specifically designed to support the selection of DHTs for data collection in clinical trials. For example, the American Psychiatric Association's (APA) app hierarchical framework [13] has been adapted into a checklist to support the selection of DHTs for research purposes, as shown in Table 1 [8].

Table . Decision-making checklist for the selection of digital health technology based on the APA's app hierarchical framework.

Domain	Items
Ethical principles	Informed consent process
Autonomy	How does the technology work? What personal data are produced? How is privacy managed? How is data stored and shared? Is bystander data captured?
Beneficence	Favorable risk/benefit analysis? Summary of known risks? Potential of unknown risks? Sound risk management plan? Rigorous study design?
Justice	Sample selection is appropriate?
Privacy and confidentiality	Are data encrypted? Who has data access? Is data security sound? Does HIPAA[a] apply? What data are shared? Is there a privacy agreement?
Efficacy	Is there evidence that the product works? Has it been used with the study population?
Usability	Is the technology customizable, sustainable, and accessible? What is needed to use the technology?
Interoperability / data sharing	Is data available for postresearch access and sharing? Is there interoperability into electronic health record?

[a]HIPAA: Health Insurance Portability and Accountability Act.

However, the scope of the checklist is narrow as the APA's hierarchical framework solely focuses on the evaluation of smartphone apps in clinical care [13] and does not address critical trial considerations, such as end point validity. Similarly, the NHS digital technology assessment criteria [10] focus on the use of DHTs in health and social care, with limited guidance on whether a DHT can reliably capture trial end points, align with clinical trial procedures and data collection plans, or maintain research-grade data integrity. In the same vein, the National Institute for Health and Care Excellence evidence standards framework delivers a high-level guidance on effectiveness and economic impact of DHTs according to risk [14], but it lacks concrete criteria for evaluating DHTs as instruments for research data collection. Nevertheless, these frameworks offer essential evaluation criteria and provide valuable insights into foundational elements that can inform the development of a trial-specific framework. Thus, there is a need to map the existing evaluation criteria and identify gaps related to trial-specific needs. This will inform the development of a robust, evidence-based framework to support technology decision-makers in making informed choices while upholding scientific rigor and regulatory compliance. Such a framework will ensure that the selected digital data collection tools are capable of accurately capturing trial end points, maintaining data integrity, and integrating seamlessly with study workflows.

Objective

This systematic review aims to develop an evidence-based framework for the selection of digital data collection tools for clinical trials.

Review Questions

This systematic review aims to address the following questions:

1. What frameworks are relevant to the selection of digital data collection tools in clinical trials?
2. What criteria in existing frameworks are relevant to selecting digital data collection tools in clinical trials?

XSL•FO
RenderX

Methods

Trial Registration

The systematic review will be presented in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocol (PRISMA-P) checklist [15] and provided in Checklist 1. This protocol is registered with the PROSPERO database (CRD420250612895) [16]. Although this manuscript provides detailed information than the registry entry, there are no significant deviations.

Ethical Considerations

Formal ethical approval is not required, as primary data will not be collected in this study. The findings from this systematic review will be published in a peer-reviewed journal and public gray literature.

Eligibility Criteria

Inclusion Criteria

The inclusion criteria in this systematic review are as follows:

- Presence of a framework: in this systematic review, a framework is defined as a structured, evidence-informed set of principles, models, or guidelines that provides explicit guidance [17] for the evaluation and selection of DHT for data collection in clinical trials. A framework categorizes relevant characteristics and provides a structured way to evaluate them [7,18]. Ideally, these characteristics would be accompanied by a set of guiding questions specific to each criterion, as shown in Table 1, prompting technology decision-makers to consider critical factors in the selection process [18].
- Presence of criteria relevant to digital data collection tools
- Published studies and gray literature
- Studies written in the English language
- Availability of full-text articles

Exclusion Criteria

We have excluded publications based on the following criteria:

- Studies that do not have a framework for digital data collection tools
- Studies that are not available in full text
- Studies that are published in a non-English language

Information Sources

The search will be conducted using the following bibliographic databases through the online library portal of University College Cork: IEEE Xplore, eAIS, PubMed, CINAHL, MEDLINE, Embase, ClinicalTrials.Gov, Scopus, and Web of Science. These databases were selected to provide a comprehensive coverage of digital and health care research information sources. Additionally, searches will be performed for publicly available gray literature from reputable institutions including the National Institute for Health and Care Excellence, Medicines & Healthcare products Regulatory Agency (MHRA), the Food and Drug Administration (FDA), the American Psychological Association, and the World Health Organization (WHO). Citation tracking of included studies using Google Scholar and a manual search of the references will also be conducted to identify other relevant studies. No date restrictions will be applied to ensure that no potentially relevant literature is inadvertently excluded. The search will include all qualitative, mixed methods, review, and consensus statement designs, including expert consensus papers and regulatory agency guidelines. This approach allows for a comprehensive synthesis of available evidence.

Search Strategy

The search will be conducted using terms related to 3 key concepts: digital health, data collection, and framework. Terms related to each concept will be combined with “OR” and the search results for all the concepts will be combined using “AND” in the databases, as outlined in Textbox 1. Truncations and wildcards will be used to increase the sensitivity of the search. For institutional websites, searches will be conducted using the search terms such as “digital health framework,” “data collection framework,” “digital technology framework,” “digital health guideline,” “data collection guideline,” and “digital technology guideline.” Platform filters will be used to narrow the scope of the search. The full search strategy for all databases is presented in Multimedia Appendix 1.

Textbox 1. Search strategy terms used for each concept.

- System or technology terms**:** "eClinical," "e-clinical," "mHealth," "m-health," "eHealth," "e-health," "telehealth," "tele-health," "health information system*," "health information technolog*," "mobile health," "digital health," "digital technology*," "digital device*," "digital tool*," "digital system*"
- Data collection terms: "data acquisition," "data collection," "data capture," "data gathering," "information acquisition," "information collection," "information capture," "information gathering"
- Method or evaluation terms: "framework*," "checklist*," "guideline*," "model*," "criteria," "taxonomy"

Data Management and Study Selection

Deduplication, selection of studies, and data extraction will be performed using Covidence [19]. The selection of studies will be performed in two stages: (1) screening of titles and abstracts and (2) full-text review. Two reviewers (BA and JO) will independently screen and review papers based on the eligibility criteria. Any disagreement during the study selection will be resolved by discussion between the two reviewers and by adjudication by a third reviewer (KPF). Where a framework or different versions of the same framework appear in different studies, all the studies will be included if they meet the eligibility criteria. The inclusion of multiple studies with the same framework will not affect the result, as thematic synthesis will address redundant and overlapping criteria. A PRISMA

XSL•FO
RenderX

(Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart will be used to illustrate the study selection process.

Data Extraction

Study details and information related to frameworks for selection of digital data collection tools will be extracted. The study details that will be extracted include authors, affiliation, country, year of publication, objective of the study, and study design. Information related to each framework to be extracted will include the following:

- How the framework was developed
- Criteria in the framework and their definition
- Type of digital data collection tools
- Stakeholders associated with the framework

Thus, the main outcomes from the review will include the following:

- Existing frameworks that are relevant to the selection of digital data collection tools for clinical trials
- The criteria that are used in the frameworks
- A new evidence-based framework, identifying criteria that are relevant to selecting digital data collection tools for clinical trials

Quality Assessment

Given the qualitative nature of this systematic review, a formal risk of bias assessment will not be conducted. However, a quality assessment will be conducted using the Mixed Methods Appraisal Tool (MMAT) that allows the evaluation of different study designs [20]. The MMAT was selected for this review due to the anticipated methodological diversity of the included sources, such as qualitative studies, Delphi studies, and quantitative surveys on framework development. The MMAT provides a consistent approach for appraising qualitative, quantitative, and mixed methods designs, thereby ensuring systematic evaluation of all evidence.

Data Synthesis

Thematic synthesis [21] will be used to create a new evidence-based framework for the selection of data collection tools for clinical trials. This process will be performed in three stages: (1) coding of the definitions of the criteria in existing frameworks; (2) development of descriptive themes; and (3) formation of analytical themes. The extracted data will be manually coded using the highlight function of Microsoft Word (Version 2511), as we envisaged that the volume of texts would be manageable. The purpose of this coding is to identify unique concepts within the definitions of the criteria from the selected frameworks. Codes will then be grouped on the basis of their similarities and differences, with a descriptive theme assigned to capture the meaning of each group of similar codes. Coding and generation of descriptive themes will be performed by 2 independent reviewers (BA and KPF), with disagreement resolved by mutual discussion or adjudication by a third reviewer (JO) when necessary to ensure reliable and rigorous synthesis. The analytical themes will be derived from the descriptive themes based on the interpretation of the reviewers. To minimize bias, each reviewer will first develop the analytical themes independently, after which all the reviewers will meet to finalize the analytical synthesis [18]. A new evidence-based framework will be developed based on the descriptive and analytical themes.

Results

The review started in May 2025 and is expected to be completed in December 2025. The searches yielded 9151 studies, which were reduced to 4333 after the removal of duplicates using Covidence. Screening and full-text review of these studies will be performed to identify the final set that will be included for thematic synthesis.

Discussion

Anticipated Findings

The primary objective of this systematic review is to develop an evidence-based framework to guide the selection of digital data collection tools for clinical trials. The resulting framework is expected to highlight core considerations such as usability, interoperability, data quality, regulatory compliance, security, and adaptability to trial-specific needs. By synthesizing findings from a wide range of sources, the review aims to provide structured guidance for selecting tools that are fit-for-purpose and capable of meeting the unique requirements of clinical research settings. Ultimately, this framework will support more consistent, transparent, and evidence-informed decision-making in clinical trial data collection.

Comparison to Prior Work

Existing frameworks have focused on specific technologies [8], health care delivery [10], or regulatory aspects of digital tools [14]. This systematic review will develop a comprehensive, structured framework specifically tailored for the selection of digital data collection tools in clinical trials. This framework will also provide actionable guidance that can be applied in a clinical trial setting.

Strengths and Limitations

A strength of this review is its systematic and comprehensive search strategy across multiple bibliographic databases, combined with the inclusion of gray literature from reputable regulatory and health organizations. This approach increases the likelihood of capturing a broad range of perspectives relevant to tool selection. Another strength is the involvement of 2 independent reviewers in study selection and data synthesis to ensure a transparent and reproducible process.

However, restricting the review to only studies published in English may exclude relevant frameworks in other languages [22,23]. Translating non-English sources could alter the precise meaning of key definitions and concepts. Including only English language materials helps in ensuring that the original intent and nuance of terms are accurately captured. Furthermore, the reliance on published and publicly available literature may overlook valuable insights contained in private industry reports or unpublished evaluations [24].

XSL•FO
RenderX

Implications for Policy, Practice, and Future Research

The anticipated framework has implications for the conduct of clinical trials. By offering structured guidance for tool selection, it may reduce inefficiencies, enhance data quality, and support regulatory compliance, ultimately improving the reliability and success of clinical research. Beyond its immediate utility, the framework could serve as a foundation for the development of best practice guidelines or policies to standardize the selection of digital data collection tools. Future research should focus on getting inputs from relevant stakeholders and validating the framework in clinical trial environments. Thus, an e-Delphi study is being planned with a group of stakeholders involved in clinical trials, including principal investigators, project leads, and technology vendors, to reach consensus on the relevance, completeness, and applicability of the framework criteria.

Conclusions

This systematic review will develop an evidence-based framework to guide the selection of digital data collection tools. Thus, the framework will provide practical guidance to technology decision-makers to improve consistency, transparency, and effectiveness in the adoption of digital technologies in clinical trials, ultimately contributing to improved trial quality and efficiency.

Acknowledgments

The authors thank the library assistants at the Boole Library, University College Cork, Ireland, for their valuable advice on literature search strategies.

Funding

Author BA's PhD is funded by Parexel International, United Kingdom.

Data Availability

All data generated or analyzed during this study are included in this published article.

Authors' Contributions

BA wrote the first draft of the manuscript. JO, KPF, and HX revised the manuscript for intellectual content and contributed to the methodology including search strategy, study selection, data extraction, and data analysis. All authors approved the final draft of the protocol. All authors meet the International Committee of Medical Journal Editors criteria.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 39 KB - resprot_v15i1e78529_app1.docx]

Checklist 1
PRISMA-P checklist.
[DOC File, 82 KB - resprot_v15i1e78529_app2.doc]

References

1. Rosa C, Marsch LA, Winstanley EL, Brunner M, Campbell AN. Using digital technologies in clinical trials: current and future applications. Contemp Clin Trials 2021 Jan;100:106219. [doi: 10.1016/j.cct.2020.106219] [Medline: 33212293]
2. Inan OT, Tenaerts P, Prindiville SA, et al. Digitizing clinical trials. NPJ Digit Med 2020 Jul 31;3:101. [doi: 10.1038/s41746-020-0302-y] [Medline: 32821856]
3. Hird N, Ghosh S, Kitano H. Digital health revolution: perfect storm or perfect opportunity for pharmaceutical R&D? Drug Discov Today 2016 Jun;21(6):900-911. [doi: 10.1016/j.drudis.2016.01.010] [Medline: 26821131]
4. Lu Z, Su J. Clinical data management: current status, challenges, and future directions from industry perspectives. Open Access J Clin Trials 2010;2:93-105. [doi: 10.2147/OAJCT.S8172]
5. Mittermaier M, Venkatesh KP, Kvedar JC. Digital health technology in clinical trials. NPJ Digit Med 2023 May 18;6(1):88. [doi: 10.1038/s41746-023-00841-8] [Medline: 37202443]
6. Colloud S, Metcalfe T, Askin S, et al. Evolving regulatory perspectives on digital health technologies for medicinal product development. NPJ Digit Med 2023 Mar 29;6(1):56. [doi: 10.1038/s41746-023-00790-2] [Medline: 36991116]
7. Karthikeyan R, Al-Shamaa N, Kelly EJ, et al. Investigating the characteristics of health-related data collection tools used in randomised controlled trials in low-income and middle-income countries: protocol for a systematic review. BMJ Open 2024 Jan 29;14(1):e077148. [doi: 10.1136/bmjopen-2023-077148] [Medline: 38286709]

XSL•FO
RenderX

8. Nebeker C, Bartlett Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med 2020 Oct 8;10(4):1004-1015. [doi: 10.1093/tbm/ibz074] [Medline: 31120511]
9. Ohmann C, Kuchinke W, Canham S, et al. Standard requirements for GCP-compliant data management in multinational clinical trials. Trials 2011 Mar 22;12(1):85. [doi: 10.1186/1745-6215-12-85] [Medline: 21426576]
10. Vijayananthan A, Nawawi O. The importance of Good Clinical Practice guidelines and its role in clinical trials. Biomed Imaging Interv J 2008 Jan;4(1):e5. [doi: 10.2349/biij.4.1.e5] [Medline: 21614316]
11. Ehidiamen AJ, Oladapo OO. The role of electronic data capture systems in clinical trials: streamlining data integrity and improving compliance with FDA and ICH/GCP guidelines. World J Biol Pharm Health Sci 2024 Oct;20(1):321-334. [doi: 10.30574/wjbphs.2024.20.1.0789]
12. Kandi V, Vadakedath S. Clinical trials and clinical research: a comprehensive review. Cureus 2023 Feb 16;15(2):e35077. [doi: 10.7759/cureus.35077] [Medline: 36938261]
13. Torous JB, Chan SR, Gipson SY, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. Psychiatr Serv 2018 May 1;69(5):498-500. [doi: 10.1176/appi.ps.201700423] [Medline: 29446337]
14. Unsworth H, Dillon B, Collinson L, et al. The NICE evidence standards framework for digital health and care technologies - developing and maintaining an innovative evidence framework with global impact. Digit Health 2021 Jun 24;7:20552076211018617. [doi: 10.1177/20552076211018617] [Medline: 34249371]
15. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015 Jan 2;350:g7647. [doi: 10.1136/bmj.g7647] [Medline: 25555855]
16. Ananthesayanan B, Fadahunsi KP, Xiong HH, O'Donoghue J. A systematic review of frameworks for guiding the selection of health-related digital data collection tools used in clinical trials. PROSPERO. URL: https://www.crd.york.ac.uk/PROSPERO/view/CRD420250612895 [accessed 2025-05-25]
17. Partelow S. What is a framework? Understanding their purpose, value, development and use. J Environ Stud Sci 2023 Sep;13(3):510-519. [doi: 10.1007/s13412-023-00833-w]
18. Fadahunsi KP, Akinlua JT, O'Connor S, et al. Protocol for a systematic review and qualitative synthesis of information quality frameworks in eHealth. BMJ Open 2019 Mar 5;9(3):e024722. [doi: 10.1136/bmjopen-2018-024722] [Medline: 30842114]
19. Covidence. URL: https://www.covidence.org/ [accessed 2025-05-05]
20. Hong QN, Pluye P, Fàbregues S, et al. Mixed Methods Appraisal Tool (MMAT) version 2018 user guide. BMJ Open 2021;11(1):e039246. [doi: 10.1136/bmjopen-2020-039246]
21. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
22. Jackson JL, Kuriyama A. How often do systematic reviews exclude articles not published in English? J Gen Intern Med 2019 Aug;34(8):1388-1389. [doi: 10.1007/s11606-019-04976-x] [Medline: 31011958]
23. Morrison A, Polisena J, Husereau D, et al. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. Int J Technol Assess Health Care 2012 Apr;28(2):138-144. [doi: 10.1017/S0266462312000086] [Medline: 22559755]
24. Paez A. Gray literature: an important resource in systematic reviews. J Evid Based Med 2017 Aug;10(3):233-240. [doi: 10.1111/jebm.12266] [Medline: 28857505]

Abbreviations

DHT: digital health technology
FDA: Food and Drug Administration
GCP: Good Clinical Practice
MHRA: Medicines & Healthcare products Regulatory Agency
MMAT: Mixed Methods Appraisal Tool
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocol
WHO: World Health Organization

XSL•FO
RenderX

Edited by S Khan; submitted 04.Jun.2025; peer-reviewed by H Anandan, M Ghorbani; revised version received 08.Oct.2025; accepted 19.Nov.2025; published 22.Jan.2026.

Please cite as:
Ananthesayanan B, Fadahunsi KP, Xiong H, O'Donoghue J
Frameworks for Guiding the Selection of Digital Data Collection Tools Used in Clinical Trials: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e78529
URL: https://www.researchprotocols.org/2026/1/e78529
doi:10.2196/78529

© Bavatharani Ananthesayanan, Kayode Philip Fadahunsi, Huanhuan Xiong, John O'Donoghue. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Association and Agreement of Contact-Based Smartphone Photoplethysmography Compared With Electrocardiography: Protocol for a Systematic Review and Meta-Analysis

James M Mather[1], MSc; Nicholas Sculthorpe[1], PhD; Ethan Berry[1], MSc; Jacqueline Mair[2,3], PhD; Nilihan Sanal-Hayes[4], PhD; Lawrence D Hayes[5], PhD

[1]Sport and Physical Activity Research Institute, School of Health and Life Sciences, University of the West of Scotland, Glasgow, United Kingdom

[2]Future Health Technologies, Campus for Research Excellence and Technological Enterprise (CREATE), Singapore-ETH Centre, Singapore

[3]Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

[4]School of Health and Society, University of Salford, Salford, United Kingdom

[5]Lancaster University Medical School, Lancaster University, Health Innovation Campus, Lancaster, United Kingdom

Corresponding Author:
Lawrence D Hayes, PhD
Lancaster University Medical School, Lancaster University, Health Innovation Campus, Lancaster, United Kingdom

Abstract

Background: Mobile health (mHealth), leveraging mobile devices for health measurement and promotion, is rapidly growing. Smartphone cameras can perform photoplethysmography (PPG) to estimate pulse rate (PR) and other features of the cardiac cycle. However, establishing the validity of PR-PPG is essential before it can be adopted for health care applications. There is a pervasive belief that PR-PPG is analogous to heart rate (HR) derived from electrocardiograms (ECGs), and we will conduct a systematic review and meta-analysis to support or challenge this supposition.

Objective: This study aims to synthesize quantitative evidence on the validity of PPG derived from mobile devices (ie, smartphones) for the assessment of HR compared with the gold standard ECG assessment.

Methods: A comprehensive literature search will be performed on CINAHL Ultimate, MEDLINE, ScienceDirect, and Scopus using a predefined search strategy. All retrieved citations will be imported into Rayyan for screening and data management. A minimum of 2 independent reviewers will conduct the title and abstract screening, followed by 2 independent reviewers who will perform full-text screening and data extraction. All stages will be guided by predefined inclusion and exclusion criteria, which will be pilot-tested to ensure consistency and reliability. Any discrepancies will be resolved through discussion with a third reviewer or during a research team meeting. Intrarater reliability will be quantified at the title and abstract stage and the full-text review stage using Cohen κ. To ensure clarity and consistency in the presentation of study characteristics and findings, both narrative synthesis and tabular formats will be used. This review will include studies that report the association and agreement between resting HR and PR from PPG using contact-based smartphone devices versus ECG as the gold standard. PPG signals will be obtained using a contact-based approach, defined as finger-on-camera measurements with the smartphone's built-in camera and flash. Studies will be excluded if they (1) do not use PPG using contact-based smartphone devices, (2) compare PPG to another collection method other than ECG, or (3) are review articles or case studies.

Results: To inform clinical procedures and future studies, the results will contain data on PR-PPG and HR-ECG association (correlations) and agreement (Bland-Altman analysis), sampling devices, and operating systems. This project is unfunded, and the initial screening is expected to start in the first quarter of 2026, with results anticipated to be published in the first quarter of 2027. The projected timeline for the study includes title and abstract screening from the first quarter of 2026, followed by full-text screening in the second quarter of 2026. Results are anticipated by the third quarter of 2026, with publication expected in the first quarter of 2027. Throughout this period, database searches will be regularly updated to capture any newly published studies meeting the inclusion criteria.

Conclusions: This review will provide a comprehensive understanding of the association and agreement between PR-PPG and HR-ECG. The findings may inform future adoption of PR-PPG and HR-ECG with insights into device or setting characteristics for best agreement or association.

Trial Registration: Open Science Framework; https://doi.org/10.17605/OSF.IO/83V7A

International Registered Report Identifier (IRRID): PRR1-10.2196/84837

(JMIR Res Protoc 2026;15:e84837) doi:10.2196/84837

XSL•FO
RenderX

KEYWORDS

photoplethysmography; PPG; mobile; heart rate; validity; artificial intelligence; AI

Introduction

Photoplethysmography (PPG) has been widely used over several decades for the diagnosis, monitoring, and screening of various diseases and disorders, offering clinically relevant physiological insights [1-4]. The term “photoplethysmography” derives from its functional components: “photo” (light), “plethysmo” (volume), and “graphy” (recording) [5]. Initially introduced by Hertzman [4,6] in 1937 to detect blood volume changes, PPG operates by measuring either transmitted (transmissive PPG) or reflected (reflective PPG) light as it interacts with biological tissues [7]. This technique relies on the optical properties of tissue, including absorption, scattering, and transmission [8]. Transmissive PPG detects light that has passed through relatively thin tissue regions, such as the fingers, toes, or earlobes. In contrast, reflective PPG captures light that is scattered back from the skin, which results in a reduction in detected light intensity [9]. While transmissive PPG generally provides more stable signal quality [10], reflective PPG offers greater versatility in terms of measurement site, enabling its application in anatomical regions such as the forehead, wrist, carotid artery, and esophagus—locations where transmissive PPG is less feasible [11-14].

PPG operates based on the Beer-Lambert law, which describes the attenuation of light intensity as a function of the extinction coefficient, concentration of the absorbing medium, and the optical path length through which light travels [15]. Leveraging this principle, a range of PPG devices are used in clinical settings to measure physiological parameters such as pulse rate (PR), a key vital sign [7]. Clinically, PPG is commonly used to monitor cardiac-induced fluctuations in blood volume within microvascular beds at peripheral anatomical sites including the finger, forehead, earlobe, and toe [16,17].

Since the introduction of the first iPhone in 2007, smartphones have become ubiquitous worldwide and are increasingly recognized as practical tools for data collection, addressing several limitations inherent in traditional methods [18]. Conventional health monitoring typically involves periodic, scheduled clinical visits, which may fail to capture dynamic physiological changes that occur longitudinally or during routine daily activities [3,19,20]. In this context, smartphones equipped with integrated cameras offer a cost-effective alternative for PPG acquisition, eliminating the need for additional external devices such as wearables [21].

Consequently, smartphone-based PPG has the potential to extend access to underserved populations, particularly those facing demographic, geographic, or socioeconomic barriers to health care access and delivery [22-24]. The adoption of mobile health (mHealth) technologies has accelerated in recent years, particularly following the COVID-19 pandemic, which underscored the utility of remote and prospective health and symptom monitoring [21,25-27]. As such, the increasing prevalence of smartphone-enabled telemedicine is likely to persist and may play a significant role in advancing global health equity, aligning with targets outlined in the United Nations Sustainable Development Goals, specifically Sustainable Development Goal 3: Good Health and Well-being [28,29].

Smartphone-based PPG can estimate resting PR through the measurement of distal pulse signals at rest, as well as during other activities such as exercise or cognitive tasks [7]. While smartphone-based PPG can measure resting PR through peripheral pulse detection, accuracy hinges on proprietary algorithms that are often undisclosed in the literature. The lack of algorithmic transparency can hinder reproducibility and trust in the results. This is particularly problematic given the proliferation of mHealth technologies and the resulting necessity that such tools be reliable and valid when compared with gold standard measurements before widespread adoption [30]. An earlier meta-analysis by De Ridder et al [31] showed that smartphone PPG could yield results consistent with electrocardiogram (ECG), pulse oximetry, and radial pulse measurements. However, they identified significant variability due to sample characteristics, environmental conditions, and the diversity of smartphone hardware. Their review also reflected on the outdated technology assessed at the time, such as iPhone 5 and Galaxy S4 models. Subsequent technological advancements, including higher-resolution cameras and improved sensors, warrant an updated synthesis of the evidence [21]. Our recent scoping review [3] identified 10 studies directly comparing PR derived from smartphone-based PPG (PR-PPG) and heart rate (HR) derived from ECG (HR-ECG) but did not quantify their agreement or association. This protocol aims to address that gap by conducting a rigorous meta-analysis of studies evaluating the association and agreement between smartphone-based PR-PPG and HR-ECG.

The primary research question is as follows: what is the level of association and agreement between PR-PPG and HR-ECG at rest?

Secondary research questions include the following:

- Does device type or operating system (eg, iOS vs Android) influence the association and agreement of PR-PPG?
- What is the methodological quality and risk of bias among included studies?

We hypothesize that smartphone-derived PR-PPG may provide a valid alternative to HR-ECG at rest. Nonetheless, the current evidence base is likely to be limited and methodologically heterogeneous due to variations in device specifications, study design, and participant characteristics.

Methods

Eligibility Criteria

The eligibility criteria will be assessed using the PICO (population, intervention, comparison, and outcome) framework. For the population component, we will only accept studies that involve humans. For the intervention component, as this is an analysis of association and agreement, rather than magnitude

XSL•FO
RenderX

of change, there will be no intervention per se. However, this systematic review will investigate studies that measure PR-PPG via front- or rear-facing camera of a smartphone using a contact-based approach and HR-ECG. Contact-based smartphone PPG is defined as a finger-on-camera measurement using the device's built-in camera and flash. Participants place the fingertip over the camera lens while the flash illuminates the skin to capture PPG signals. Noncontact methods, such as facial video PPG, are not included in this study. We will restrict inclusion to studies that report agreement or correlation, and we will not compute association or agreement from summary data. For the comparison component, the inclusion of a control group is not necessary in this analysis. Regarding outcome measures, resting pulse rate (beats per min [bpm]) from PR-PPG and resting HR (bpm) from HR-ECG will be reported in the original studies. Moreover, to quantify association and agreement, studies must report a correlation coefficient (Pearson correlation coefficient, concordance correlation coefficient, or interclass correlation coefficient) or a Bland-Altman analysis between PR-PPG and HR-ECG.

The following exclusion criteria will apply: studies using external devices (eg, medical sensors or wearables) connected to smartphones for data acquisition and papers that do not assess the validity of PR-PPG against HR-ECG as an outcome measure. Only original research articles presenting at least preliminary quantitative findings will be included. Qualitative studies, case reports, and literature reviews will not be eligible for inclusion in this review.

Search Strategy

This meta-analysis will be performed in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines. The study protocol has been registered on the Open Science Framework (OSF). Any deviations from this protocol during the review will be documented with justifications and time stamps to ensure full transparency, and database searches will be updated close to the time of analysis.

A comprehensive search strategy has been developed to strike a balance between thorough coverage and practical scope, ensuring methodological rigor. Five electronic databases will be searched: CINAHL Ultimate, MEDLINE, ScienceDirect, Scopus, and EMBASE. These databases were selected for their relevance across health sciences, biomedical research, signal processing, and behavioral science. CINAHL Ultimate is included for its strong representation of allied health literature. MEDLINE, accessed via EBSCOhost, provides authoritative sources in biomedical and cardiovascular research. ScienceDirect is chosen for its wide range of journals covering computational and physiological aspects of PPG. Scopus is used for its comprehensive indexing and citation tracking capabilities. Embase is searched because it offers broader, deeper, and more clinically oriented coverage than many other databases, especially for medical, pharmaceutical, and health-related research. In addition to peer-reviewed literature, gray literature will be searched. This includes preprint servers (eg, MedRxiv and arXiv), theses and dissertations (ProQuest), conference proceedings (eg, Web of Science), and institutional repositories such as DSpace. Citation chaining will also be used by examining reference lists and citing articles (via Google Scholar and Crossref), along with CoCites to identify related work through citation patterns. The search strategy will be refined iteratively if it produces excessive irrelevant records, with adjustments made to keyword combinations, Boolean logic, and subject headings. Known relevant studies will be used to test the sensitivity and specificity of the search. If required data are missing or unclear, the corresponding authors will be contacted using a standardized template. Follow-up emails will be sent 2 and 4 weeks after the initial message if no reply is received. All correspondence will be logged and summarized in the manuscript and supplementary materials. A final search update will be conducted before manuscript submission if more than 6 months have passed since the initial search.

Search Terms

The following is an example Boolean search string used for Scopus: TITLE-ABS-KEY ("heart rate" OR hr) AND TITLE-ABS-KEY (photoplethysmography OR ppg OR "remote ppg" OR "camera-based ppg" OR "remote photoplethysmography" OR "camera based photoplethysmography") AND TITLE-ABS-KEY (smartphone* OR "mobile phone" OR "mobile device" OR camera* OR "smartphone photoplethysmography" OR "phone-based ppg" OR "smartphone ppg") AND TITLE-ABS-KEY (electrocardiogram OR ecg OR ekg OR electrocardiography) AND TITLE-ABS-KEY (validity OR valid OR accuracy OR agreement OR correlation OR reliability).

The example search string includes Scopus syntax; these terms and syntax will be adapted for all databases planned for inclusion in the search. The search period will begin in 2007 to ensure inclusion of early studies related to the first smartphone models (eg, iPhone released in 2007) and will extend through the end of 2025. The review will focus on correlation coefficients comparing HR estimates from PPG to ECG. Supplementary data will include study setting, participant demographics, smartphone model, app characteristics, sampling frequency, camera specifications, ECG details, and environmental conditions during data collection.

Study and Source of Evidence Selection

All references retrieved from the database search will be imported into Rayyan (Rayyan Systems Inc) [32], where duplicate entries will be automatically identified and removed. The screening process will take place in two stages: (1) title and abstract screening and (2) full-text screening.

In both stages, 2 independent reviewers will apply *a priori* inclusion and exclusion criteria. Reviewers will work in a blinded manner, with Rayyan's platform enabling independent decisions without revealing each other's judgments. Disagreements will be flagged automatically and resolved through discussion. If consensus cannot be achieved, a third reviewer will serve as an adjudicator.

To improve consistency, a calibration phase will be conducted before formal screening. Reviewers will test the inclusion and exclusion criteria on a sample of records to align interpretation and application. Interrater reliability will be evaluated using

XSL•FO
RenderX

Cohen κ statistic, which measures agreement beyond chance. This metric will be calculated at both the abstract and full-text screening phases. Screening decisions and bibliographic metadata will be exported in RIS and CSV formats and archived in an open-access repository (eg, OSF) at the time of submission or acceptance.

Data Extraction

Before full-scale data extraction, a pilot phase will be conducted where all reviewers independently extract data from a small subset of studies (5% - 10%). This exercise will be used to refine the extraction template and standardize the approach.

In the main extraction phase, 2 reviewers will independently extract data using a structured and pretested form. Extracted variables will include correlation coefficients and mean bias and limits of agreement between PR-PPG and HR-ECG, participant demographics, smartphone specifications, ECG characteristics, and data collection procedures. As we will include correlation coefficients and Bland-Altman analysis, these tests will report both *r* values (−1 to 1) and agreement values (in bpm or Hz). Therefore, it is important to outline our harmonization strategy. Where a study reports HR or PR in Hz, we will apply the following formula:

$$Hz = \frac{bpm}{60}$$

During the risk of bias assessment, the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) will be used independently by 2 reviewers. QUADAS-2 evaluates 4 domains: patient selection, index test, reference standard, and flow and timing. Each domain will be rated as low, high, or unclear risk of bias, and applicability concerns will be assessed for patient selection, index test, and reference standard.

Disagreements will be resolved through consensus or by involving a third adjudicator. Any ambiguities in reported data will be flagged for discussion. Missing values will be marked as “NR” (not reported), and no estimations will be made unless explicitly stated. All data entries will be double-checked by a second reviewer for accuracy. Where appropriate, artificial intelligence–assisted tools may be used to support metadata extraction or to highlight relevant text, but all final data entries will be manually reviewed and verified. Reconciliation logs and resolution steps will be documented throughout the process. Data collection will adhere to PRISMA and Cochrane Handbook guidelines. Finalized data will be formatted for meta-analysis and saved using a standardized file naming convention. These files will also be uploaded to the OSF repository.

Details on each included study will be systematically documented, beginning with key bibliographic information, such as the study title, authors, publication year, journal name, and study setting (eg, academic or clinical environments). Further extracted information will include total sample size; participant demographics (age, sex, health status, and skin pigmentation); country of study; smartphone model and specifications (eg, camera resolution and flash use); application name and whether it is a proprietary commercial application, open-source software, or custom-developed code; sampling rate of the PPG signal; camera orientation used (front- vs rear-facing); channel used in data computation; ECG equipment used; electrode placement details; ECG processing methods, environmental conditions, and participant preparation (eg, posture, breathing instructions, and dietary control); duration of recording; and number of measurement trials.

In addition to the extracted information above, the extraction form will capture relevant statistical outcomes, such as correlation coefficients, CIs, mean bias, limits of agreement, and any additional reported metrics that relate to association or agreement. Risk of bias for each study will be evaluated using the QUADAS-2 tool, with domain-specific judgments recorded and justified.

To ensure data accuracy, each extraction will be independently performed by 2 reviewers. Their entries will be compared side by side, with discrepancies resolved through consensus. If a resolution cannot be reached, a third reviewer will arbitrate. All final decisions and justifications will be recorded in a reconciliation log. Each data extraction file will follow a standardized naming convention (eg, StudyID_Initials_Extraction.xlsx) and be uploaded to the OSF repository upon completion. Once finalized, the cleaned and verified dataset will be formatted for meta-analysis and made available for public access alongside the manuscript, unless restricted by publisher guidelines.

Data Analysis and Presentation

The analysis will use the Fisher r-to-z transformed correlation coefficient as the primary outcome measure of association. The analysis will use the pooled mean difference and limits of agreement as the primary outcome measure of agreement. A random-effects model will be applied to the data, with all analyses performed in RStudio (Posit PBC) using the *metafor* package. Welz et al [33] proposed that when conducting a meta-analysis involving Fisher r-to-z transformed correlation coefficients, especially in the presence of study heterogeneity, using a random-effects model with appropriate variance estimation techniques yields more reliable and generalizable results compared with fixed-effects models. Therefore, random-effects models will be used. Subgroup analyses will only be performed when there are at least 5 studies per subgroup. Metaregression will require a minimum of 10 studies per covariate to reduce risk of overfitting. Tests for funnel plot asymmetry will only be conducted when the meta-analysis includes 10 or more studies, as recommended by the Cochrane Handbook. Sensitivity analyses will be interpreted cautiously when the evidence base is small.

To identify potential outliers and influential studies, studentized residuals and Cook distances will be examined [34]. Studies with studentized residuals exceeding the $100\times(1-0.05/(2\times k))$th percentile of the standard normal distribution (applying a Bonferroni correction with 2-sided $\alpha=.05$ for k studies) will be flagged as potential outliers. Studies with Cook distances greater than the median plus 6 times the IQR of Cook distances will be deemed influential. Outliers and influential studies will be retained in the primary analysis but flagged for sensitivity analyses to assess their impact on pooled estimates. Funnel plot asymmetry will be assessed using the rank correlation test [35] and regression test [36], where the SE of observed outcomes serves as the predictor. Heterogeneity will be evaluated through

XSL•FO
RenderX

I^2, Cochran Q, and τ^2 statistics. Where possible, subgroup analyses will explore variability in PR-PPG validity based on factors such as operating system, participant characteristics (eg, age and sex), and study design.

Moderator analyses will investigate the influence of PPG parameters and participant traits on outcomes, while sensitivity analyses will test the effects of risk of bias. Additional sensitivity checks will consider the impact of excluding high-risk studies, approaches to missing data, and methodological quality. Provided sufficient data, metaregression will examine relationships between continuous variables and effect sizes. Results will be displayed using forest and funnel plots. Conclusions concerning association will be based on effect sizes, with overall effect sizes of $r>0.9$ (approximately $z>1.5$) considered valid. For context, $r=0.70$ is typically regarded as very large and approximately corresponds to $z=0.8$. Conclusions concerning agreement will be based on bias and limits of agreement (LoA). Bias <1 bpm and LoA <5 bpm will be considered excellent agreement and acceptable for medical-grade use, bias <3 bpm and LoA <10 bpm will be considered acceptable for general PR monitoring and may be borderline for medical-grade use, and bias >5 bpm or LoA >10 bpm will be considered poor agreement and not clinically interchangeable. High heterogeneity ($I^2>50\%$) will warrant cautious interpretation and further exploration through subgroup or moderator analyses. Confidence in conclusions will be guided by the QUADAS-2 framework depending on the quality of accuracy.

Results

To inform clinical procedures and future studies, the results will contain data on PR-PPG and HR-ECG association (correlations) and agreement (Bland-Altman analysis), sampling devices, and operating systems. This project is unfunded, and the initial screening is expected to start in the first quarter of 2026, with results anticipated to be published in the first quarter of 2027. Throughout this period, database searches will be regularly updated to capture any newly published studies meeting the inclusion criteria.

A summary of the study selection procedure is presented in Figure 1 to ensure transparency during study selection. Reporting will follow the PRISMA checklist (Checklist 1).

XSL•FO
RenderX

Figure 1. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram visually summarizes the screening process.

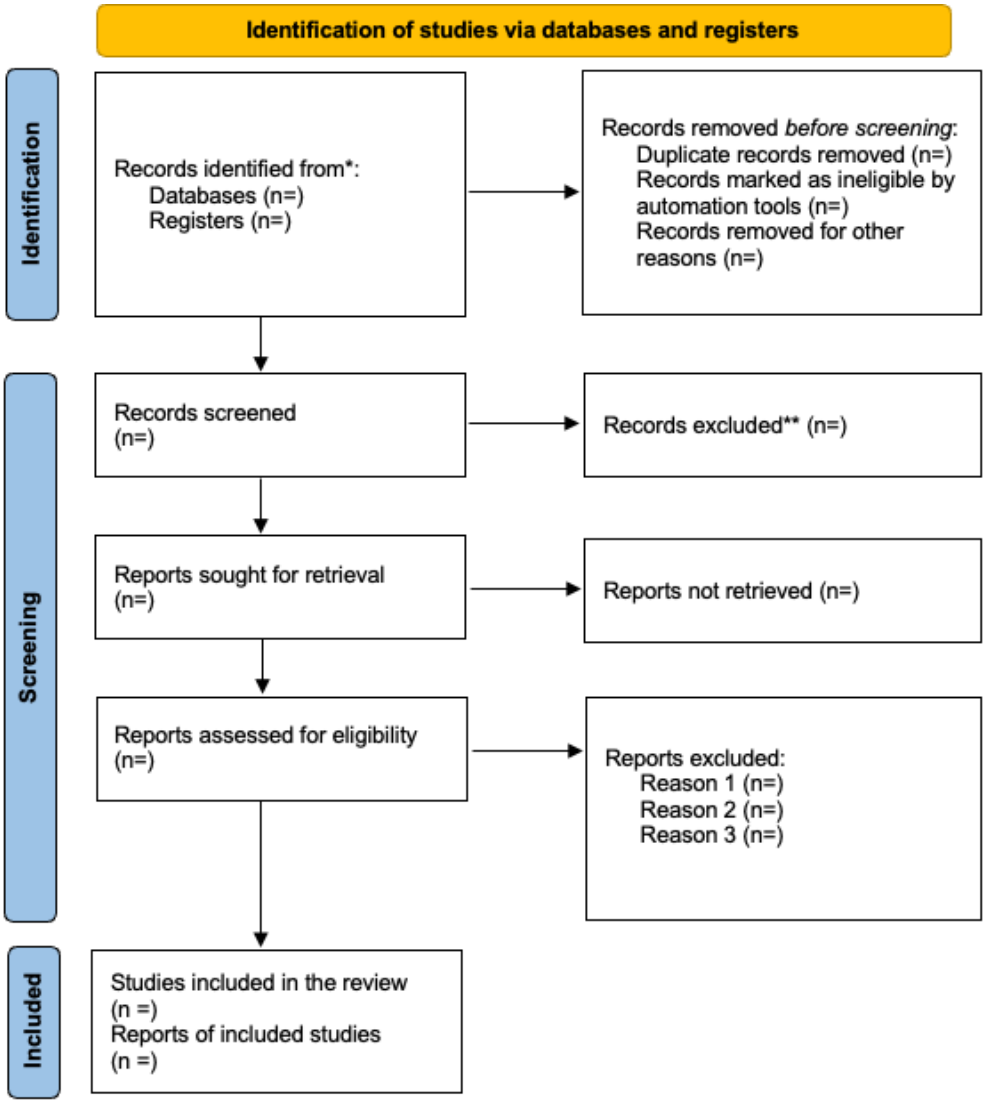

*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers).

**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

Discussion

This systematic review and meta-analysis aims to be the first to provide an updated quantification of the association and agreement of PR-PPG compared with HR-ECG at rest. The rapid development in technology requires continuous assessment of validity, and this meta-analysis will provide important findings, indicating whether PR-PPG can be used with confidence. There is a pervasive belief that PR-PPG is analogous to HR-ECG, and herein we will provide an updated systematic review and meta-analysis to support or challenge this supposition. We hypothesize that smartphone-based PR-PPG will demonstrate strong association and acceptable agreement with HR-ECG at rest. These findings would support the potential of PR-PPG as a valid alternative for remote monitoring in clinical and nonclinical settings.

The decision to undertake a meta-analysis was driven by the need to evaluate the level of agreement between PR-PPG and HR-ECG, including modifying parameters such as frame rate, camera location, skin pigmentation, and operating system. Moreover, if there is a lack of knowledge regarding modifying variables, this systematic review will provide opportunities for future research. By providing context to current findings, this research can guide future studies concerning PR-PPG and its application. Previous reviews, such as the one by De Ridder et al [31], indicated that smartphone-based PR-PPG could approximate HR-ECG. However, these analyses were constrained by outdated technology and lacked detailed agreement metrics, limiting their clinical applicability. The meta-analysis by De Ridder et al [31] included studies published between 2009 and 2016, which means smartphone models assessed were from that era. These typically included

XSL•FO
RenderX

early-generation devices such as the Apple iPhone 4 and iPhone 5 series (released in 2010 and 2012), the Samsung Galaxy S3 and S4 (released in 2012 and 2013), and other similar Android models from that period. These devices are now 10 to 15 years old, and their hardware (camera resolution, LED flash quality, and processing power) is significantly outdated compared to current smartphones. This is why the findings from the study by De Ridder et al [31], while useful at the time, may not fully reflect the performance of modern devices with advanced sensors and algorithms. Our meta-analysis aims to address these gaps by including studies that use modern smartphones and operating systems, reflecting current technological capabilities. By quantifying agreement using Bland-Altman analysis alongside correlation measures, we will provide a more robust assessment of interchangeability. Moreover, by exploring potential moderators, such as device type, operating system, and participant demographics, we will be able to better understand sources of variability.

This review has several strengths, including a rigorous methodology following PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) guidelines and preregistration, clear eligibility criteria using the PICO framework, a comprehensive search strategy of multiple databases, robust screening and data extraction with dual independent reviewers with calibration and Cohen κ for reliability, and the use of structured extraction forms and QUADAS-2 for bias assessment. Our analysis plan includes random-effects meta-analysis using Fisher r-to-z transformation, outlier diagnostics, funnel plot asymmetry tests, and moderator analyses. The main limitations are those of the original studies. These include potential ambiguity in reporting, heterogeneity in study design and reporting, selective reporting, and lack of algorithm transparency due to proprietary software. The results from this systematic review and meta-analysis could identify areas for future research, such as lack of diversity in participants, potential for integration with telehealth and electronic health records, and device-specific or operating system–specific optimization.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Authors' Contributions

All authors contributed equally to the conceptualization, preliminary search, search strategy, writing – original draft, and review and editing.

Conflicts of Interest

None declared.

Checklist 1
PRISMA checklist.
[DOCX File, 18 KB - resprot_v15i1e84837_app1.docx]

References

1. Humphreys K, Ward T, Markham C. Noncontact simultaneous dual wavelength photoplethysmography: a further step toward noncontact pulse oximetry. Rev Sci Instrum 2007 Apr;78(4):044304. [doi: 10.1063/1.2724789] [Medline: 17477684]
2. Nilsson L, Goscinski T, Kalman S, Lindberg LG, Johansson A. Combined photoplethysmographic monitoring of respiration rate and pulse: a comparison between different measurement sites in spontaneously breathing subjects. Acta Anaesthesiol Scand 2007 Oct;51(9):1250-1257. [doi: 10.1111/j.1399-6576.2007.01375.x] [Medline: 17711563]
3. Mather JD, Hayes LD, Mair JL, Sculthorpe NF. Validity of resting heart rate derived from contact-based smartphone photoplethysmography compared with electrocardiography: a scoping review and checklist for optimal acquisition and reporting. Front Digit Health 2024 Feb 29;6:1326511. [doi: 10.3389/fdgth.2024.1326511] [Medline: 38486919]
4. Hertzman AB. Photoelectric plethysmography of the fingers and toes in man. Proc Soc Exp Biol Med 1937 Dec 1;37(3):529-534. [doi: 10.3181/00379727-37-9630]
5. Alnaeb ME, Crabtree VP, Boutin A, Mikhailidis DP, Seifalian AM, Hamilton G. Prospective assessment of lower-extremity peripheral arterial disease in diabetic patients using a novel automated optical device. Angiol Open Access 2007;58(5):579-585. [doi: 10.1177/0003319707305685] [Medline: 18024941]
6. Hertzman AB. The blood supply of various skin areas as estimated by the photoelectric plethysmograph. Am J Physiol Legacy Content 1938;124(2):328-340. [doi: 10.1152/ajplegacy.1938.124.2.328]

XSL•FO
RenderX

7. Almarshad MA, Islam MS, Al-Ahmadi S, BaHammam AS. Diagnostic features and potential applications of PPG signal in healthcare: a systematic review. Healthcare (Basel) 2022 Mar 16;10(3):547. [doi: 10.3390/healthcare10030547] [Medline: 35327025]
8. Challoner AV, Ramsay CA. A photoelectric plethysmograph for the measurement of cutaneous blood flow. Phys Med Biol 1974 May;19(3):317-328. [doi: 10.1088/0031-9155/19/3/003] [Medline: 4445210]
9. Park J, Seok HS, Kim SS, Shin H. Photoplethysmogram analysis and applications: an integrative review. Front Physiol 2022 Mar 1;12:808451. [doi: 10.3389/fphys.2021.808451] [Medline: 35300400]
10. Li K, Zhang S, Yang L, et al. Changes of arterial pulse waveform characteristics with gestational age during normal pregnancy. Sci Rep 2018 Oct 22;8(1):15571. [doi: 10.1038/s41598-018-33890-1] [Medline: 30349022]
11. Venema B, Schiefer J, Blazek V, Blanik N, Leonhardt S. Evaluating innovative in-ear pulse oximetry for unobtrusive cardiovascular and pulmonary monitoring during sleep. IEEE J Transl Eng Health Med 2013 Aug 8;1:2700208. [doi: 10.1109/JTEHM.2013.2277870] [Medline: 27170855]
12. Venema B, Blanik N, Blazek V, Gehring H, Opp A, Leonhardt S. Advances in reflective oxygen saturation monitoring with a novel in-ear sensor system: results of a human hypoxia study. IEEE Trans Biomed Eng 2012 Jul;59(7):2003-2010. [doi: 10.1109/TBME.2012.2196276] [Medline: 22547451]
13. Blanik N, Heimann K, Pereira C, et al. Remote vital parameter monitoring in neonatology - robust, unobtrusive heart rate detection in a realistic clinical scenario. Biomed Tech (Berl) 2016 Dec 1;61(6):631-643. [doi: 10.1515/bmt-2016-0025] [Medline: 27743509]
14. Wannenburg J, Malekian R. Body sensor network for mobile health monitoring, a diagnosis and anticipating system. IEEE Sensors J 2015 Dec;15(12):6839-6852. [doi: 10.1109/JSEN.2015.2464773]
15. Mayerhöfer TG, Pahlow S, Popp J. The Bouguer-Beer-Lambert law: shining light on the obscure. Chemphyschem 2020 Sep 15;21(18):2029-2046. [doi: 10.1002/cphc.202000464] [Medline: 32662939]
16. Elgendi M. On the analysis of fingertip photoplethysmogram signals. Curr Cardiol Rev 2012 Feb;8(1):14-25. [doi: 10.2174/157340312801215782] [Medline: 22845812]
17. Elgendi M, Fletcher R, Liang Y, et al. The use of photoplethysmography for assessing hypertension. NPJ Digit Med 2019 Jun 26;2:60. [doi: 10.1038/s41746-019-0136-7] [Medline: 31388564]
18. Drijkoningen L, Lenaerts F, Vandervoort P, Grieten L, Van der Auwera J, Leysen K. Validation of a smartphone based photoplethysmographic beat detection algorithm for normal and ectopic complexes. Presented at: Computing in Cardiology 2014; Sep 7-10, 2014.
19. Goërtz YM, Braamse AM, Spruit MA, et al. Fatigue in patients with chronic disease: results from the population-based Lifelines Cohort Study. Sci Rep 2021 Oct 25;11(1):20977. [doi: 10.1038/s41598-021-00337-z] [Medline: 34697347]
20. Mair JL, Hayes LD, Campbell AK, Sculthorpe N. Should we use activity tracker data from smartphones and wearables to understand population physical activity patterns? J Meas Phys Behav 2021;5(1):1-5. [doi: 10.1123/jmpb.2021-0012]
21. Raposo A, da Silva HP, Sanches J. Camera-based photoplethysmography (cbPPG) using smartphone rear and frontal cameras: an experimental study. Annu Int Conf IEEE Eng Med Biol Soc 2021 Nov;2021:7091-7094. [doi: 10.1109/EMBC46164.2021.9630847] [Medline: 34892735]
22. Kim BH, Glanz K. Text messaging to motivate walking in older African Americans: a randomized controlled trial. Am J Prev Med 2013 Jan;44(1):71-75. [doi: 10.1016/j.amepre.2012.09.050] [Medline: 23253653]
23. Blumenthal D, Mort E, Edwards J. The efficacy of primary care for vulnerable population groups. Health Serv Res 1995 Apr;30(1 Pt 2):253-273. [Medline: 7721596]
24. Weitz TA, Freund KM, Wright L. Identifying and caring for underserved populations: experience of the National Centers of Excellence in Women's Health. J Womens Health Gend Based Med 2001 Dec;10(10):937-952. [doi: 10.1089/152460901317193521] [Medline: 11788105]
25. Sculthorpe NF, McLaughlin M, Cerexhe L, et al. Tracking Persistent Symptoms in Scotland (TraPSS): a longitudinal prospective cohort study of COVID-19 recovery after mild acute infection. BMJ Open 2025 Jan 15;15(1):e086646. [doi: 10.1136/bmjopen-2024-086646] [Medline: 39819953]
26. Hayes LD, Sanal-Hayes NE, Ellam M, Mclaughlin M, Swainson MG, Sculthorpe NF. A method for determination of hematocrit using the mobile app "HaemoCalc": validity, reliability, and effect of user expertise. Physiol Rep 2025 Apr;13(8):e70314. [doi: 10.14814/phy2.70314] [Medline: 40232943]
27. Hayes LD, Ingram J, Sculthorpe NF. More than 100 persistent symptoms of SARS-CoV-2 (long COVID): a scoping review. Front Med (Lausanne) 2021 Nov 1;8:750378. [doi: 10.3389/fmed.2021.750378] [Medline: 34790680]
28. Goal 3. United Nations Department of Economic and Social Affairs. URL: https://sdgs.un.org/goals/goal3 [accessed 2025-06-06]
29. Asi YM, Williams C. The role of digital health in making progress toward Sustainable Development Goal (SDG) 3 in conflict-affected populations. Int J Med Inform 2018 Jun;114:114-120. [doi: 10.1016/j.ijmedinf.2017.11.003] [Medline: 29126701]
30. Vandenberk T, Stans J, Mortelmans C, et al. Clinical validation of heart rate apps: mixed-methods evaluation study. JMIR Mhealth Uhealth 2017 Aug 25;5(8):e129. [doi: 10.2196/mhealth.7254] [Medline: 28842392]

XSL•FO
RenderX

31. De Ridder B, Van Rompaey B, Kampen JK, Haine S, Dilles T. Smartphone apps using photoplethysmography for heart rate monitoring: meta-analysis. JMIR Cardio 2018 Feb 27;2(1):e4. [doi: 10.2196/cardio.8802] [Medline: 31758768]
32. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
33. Welz T, Doebler P, Pauly M. Fisher transformation based confidence intervals of correlations in fixed- and random-effects meta-analysis. Br J Math Stat Psychol 2022 Feb;75(1):1-22. [doi: 10.1111/bmsp.12242] [Medline: 33934346]
34. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. Res Synth Methods 2010 Apr;1(2):112-125. [doi: 10.1002/jrsm.11] [Medline: 26061377]
35. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994 Dec;50(4):1088-1101. [Medline: 7786990]
36. Sterne JA, Egger M. Regression methods to detect publication and other bias in meta-analysis. In: Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments: John Wiley & Sons; 2006:99-110 URL: https://research-information.bris.ac.uk/en/publications/regression-methods-to-detect-publication-and-other-bias-in-meta-a/ [accessed 2025-12-15]

Abbreviations

AI: artificial intelligence
ECG: electrocardiogram
HR: heart rate
mHealth: mobile health
OSF: Open Science Framework
PICO: population, intervention, comparison, and outcome
PPG: photoplethysmography
PR: pulse rate
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies

Edited by J Sarvestan; submitted 25.Sep.2025; peer-reviewed by RN Alaslani; revised version received 03.Dec.2025; accepted 05.Dec.2025; published 03.Feb.2026.

Please cite as:
Mather JM, Sculthorpe N, Berry E, Mair J, Sanal-Hayes N, Hayes LD
Association and Agreement of Contact-Based Smartphone Photoplethysmography Compared With Electrocardiography: Protocol for a Systematic Review and Meta-Analysis
JMIR Res Protoc 2026;15:e84837
URL: https://www.researchprotocols.org/2026/1/e84837
doi:10.2196/84837

© James M Mather, Nicholas Sculthorpe, Ethan Berry, Jacqueline Mair, Nilihan Sanal-Hayes, Lawrence D Hayes. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 3.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Challenges Faced by Parents in Preventing Online Child Sexual Exploitation and Abuse: Protocol for a Systematic Review

Neelam Punjani[1]; Wenting Yan[1]; Lisa Hartling[2]; Amber Hussain[1]; Farah Elgaweesh[1]; Shannon Scott[1]

[1]College of Health Sciences, Faculty of Nursing, University of Alberta, 5-172 ECHA, 11405 87 Avenue NW, Edmonton, AB, Canada

[2]Pediatrics Department, Faculty of Medicine and Dentistry, University of Alberta, 2J2.00 Walter C. Mackenzie Health Sciences Centre, 8440 112 St. NW, Edmonton, AB, Canada

Corresponding Author:
Amber Hussain
College of Health Sciences, Faculty of Nursing, University of Alberta, 5-172 ECHA, 11405 87 Avenue NW, Edmonton, AB, Canada

Abstract

Background: Online child sexual abuse and exploitation have become a global issue affecting the health and well-being of children and adolescents. Unlike traditional forms of sexual abuse and exploitation, online sexual abuse presents unique challenges due to the complexity of the situations involved. Parents play a crucial role as the first line of defense in preventing these incidents, yet the barriers they encounter have not been thoroughly explored.

Objective: The objective of the systematic review is to identify the challenges parents face in preventing online child sexual abuse, to explore the factors that contribute to those challenges, and to highlight gaps in knowledge and practice regarding how to best support them.

Methods: A comprehensive search will be conducted in MEDLINE (Ovid), Embase (Ovid), PsycINFO (Ovid), HealthSTAR (Ovid), Sociological Abstracts (ProQuest), CINAHL (EBSCOhost), Scopus (Elsevier), and the Cochrane Library (Wiley). The strategy will be developed with a research librarian. Regarding inclusion criteria, this review will include primary research studies (qualitative, quantitative, or mixed methods) that examine the perspectives of parents or primary caregivers of children and adolescents on online child sexual exploitation and abuse (OCSEA) prevention. Studies must focus on parental attitudes, fears, perceptions, or challenges related to OCSEA. Only studies published in English from 2018 onward will be considered. Studies that do not report on parental perspectives will be excluded. Two reviewers will independently screen studies and extract data, with conflicts resolved by a third reviewer. Findings will be narratively synthesized, following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework.

Results: The search strategy has been developed in collaboration with a health sciences librarian and followed PRISMA-Search guidelines. Database searches, screening, and analysis are planned to commence in August, 2025. The findings of the review are anticipated to be ready by Febuary 2026 and will be presented following the PRISMA reporting standards.

Conclusions: This protocol aims to structure a comprehensive and transparent process for a systematic review, which is expected to contribute to a better understanding of parental challenges in preventing OCSEA and provide insights into the types of supportive resources they need to combat OCSEA. This information will contribute to the development of comprehensive strategies that equip parents with the knowledge and tools necessary to effectively safeguard their children in the digital landscape.

Trial Registration: OSF Registries osf.io/fru5k; https://osf.io/fru5k/

International Registered Report Identifier (IRRID): PRR1-10.2196/80968

(JMIR Res Protoc 2026;15:e80968) doi:10.2196/80968

KEYWORDS

digital safety; child online sexual exploitation; parental perspectives; prevention barriers; systematic review

Introduction

Child sexual abuse (CSA) occurs when a child engages in sexual activity that they do not comprehend, cannot consent to, are unaware of, or are not prepared for [1]. It also includes illegal acts; for example, in Canada, 16 sexual offenses in the Criminal Code could apply to CSA [2]. CSA includes different types of abuse, such as sexual touching and sexual harassment, which can involve unwanted words, gestures, or physical contact of a sexual nature [3].

Child sexual exploitation (CSE) is a type of sexual abuse, with a distinguishing feature being the element of exchange [4,5]. This includes coercive situations in which a child or young person is pressured, manipulated, or tricked into sexual acts in

XSL•FO
RenderX

exchange for something they may require or desire [6]. Many children are led to believe that they are in loving, consensual relationships, which can normalize these exploitative situations for them. This grooming process can obscure the reality of the exploitation taking place.

Both CSA and CSE represent severe forms of abuse. When they occur online or through information and communication technologies (ICTs), these abuses can collectively be referred to as online child sexual exploitation and abuse (OCSEA). As digital technology increasingly integrates into everyday life, OCSEA introduces new challenges for prevention and intervention in contemporary society.

OCSEA includes different types of sexual abuse or exploitation that occur online, including online solicitation, grooming, watching pornographic materials, forced or unwanted sexting, image-based abuse, sexual extortion, and exploitation [7-9]. Through ICTs, children may be exposed to cyberbullying, harmful or abusive content, and online grooming by potential predators, placing them at risk of exploitation and abuse [10]. Therefore, the development of technology has reshaped how CSA and CSE are perpetrated and examined, putting children at greater risk and increasing the extent of possible harm. Evidence suggests that when online abuse is categorized as sexual abuse, its incidence rates surge dramatically [11]. Thus, it is crucial to recognize OCSEA as a distinct concern. The breadth of this issue is extensive, and its impact on survivors is more severe than ever. The survivors of OCSEA may experience long-term effects, including problems in building relationships with other individuals in daily life, alienation from family and friends, mental illness, suicide attempts, self-harm, or even participating in criminal behaviors [12].

Although scholars acknowledge the seriousness of OCSEA, it remains an underreported problem for various reasons [13]. Currently, OCSEA is emerging as a significant global public health crisis. Statistics reveal that approximately 20% of children worldwide are affected by OCSEA [3]. Different countries have reported their prevalence. For example, Finkelhor et al [7] analyzed the prevalence of online CSA in the United States and determined that online CSA prevalence was 15.6%. Walsh et al [14] reported that in Australia, the percentage of adults soliciting sex online was 17.7%, and the percentage of nonconsensual dissemination of sexual images of children was 7.6%. In Canada, the percentage of CSE online rose by 18% from 2021 to 2022, with child pornography being the most prevalent offense [15]. The COVID-19 pandemic exacerbated the situation, as lockdowns confined children at home, leading to a shift to online learning environments, thereby increasing exposure to OCSEA. Meanwhile, the issue is also latent; studies show that approximately 1 in every 3 children does not disclose their abuse to anyone, indicating that sexual offenses against children are likely underreported and suggesting that actual incidence rates are much higher than currently reflected [15,16].

To protect future generations, it is essential to proactively address and prevent OCSEA. However, a significant number of children or even their parents are unaware of what OCSEA entails. Therefore, we must educate and empower children and their parents on this issue, enabling them to recognize and report incidents of OCSEA, thus breaking the cycle of abuse [13].

Research has consistently highlighted the critical role that parents play in safeguarding their children from CSA [17-19]. According to Rudolph et al [18], parents serve as a protective factor both directly and indirectly. Directly, they provide supervision, monitor their children's activities, and engage in their daily lives, thereby offering a shield against potential risks. Indirectly, parents foster their children's general self-efficacy, competence, well-being, and self-esteem, which are vital for resilience. As the primary caregivers, parents know their children best and are attuned to their emotional changes. Their involvement is the most effective way to create a secure environment where children feel safe and protected [18]. However, challenges persist regarding effective parental involvement [20]. Key questions arise: How can parents convey a comprehensive prevention message? How can they bridge knowledge and communication effectively with their children? The context of OCSEA, compounded by the rise of ICTs, introduces new barriers for parents as protectors. Many parents may lack the necessary knowledge or skills to guide their children through the growing array of online risks and supervise their internet use effectively. Additionally, detecting OCSEA has become increasingly complex; thus, traditional methods of warning children about threats may no longer suffice. A broader approach is essential—one that empowers parents to guide their children in managing risks and enhancing their capacity to protect themselves from harm.

Several studies have examined parental preventative practices regarding traditional forms of CSA and CSE [17,21,22]. However, there has been less focus on the unique challenges parents face in preventing OCSEA. Novianti [23] investigated and identified several vulnerability factors contributing to OCSEA, highlighting the lack of parental supervision in children's internet use as a significant concern. Despite this, few studies have explored the barriers and challenges parents encounter when trying to mitigate this vulnerability. Addressing these issues is crucial for developing effective programs and strategies that meet the needs of the digital age. Therefore, a systematic review will be undertaken to identify previous research that has examined the challenges parents face and the factors that contribute to those challenges in preventing OCSEA. This review will illuminate gaps in current knowledge and practice regarding how to best equip parents with the necessary information and skills. Ultimately, this research will provide evidence-informed guidance for interventions, policymaking, and future research to enhance the prevention of OCSEA.

This systematic review will be guided by the following research questions:

- What types of evidence exist regarding the challenges parents face in preventing OCSEA?
- How have parents' perceptions related to OCSEA been described in the existing literature?
- What gaps exist in the current evidence base regarding parental involvement in preventing OCSEA?

XSL•FO
RenderX

Methods

This review will be conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement [24].

Identifying Relevant Studies

To identify the relevant studies, a systematic literature search will be conducted by an experienced health sciences librarian. Multimedia Appendix 1 shows the proposed search strategy to be used to search MEDLINE. Searches will be performed in the following databases: MEDLINE, Embase, PsycINFO, HealthSTAR, Sociological Abstracts via ProQuest, CINAHL via EBSCOhost, Scopus via Elsevier, and the Cochrane Library via Wiley. These databases will be searched using a combination of natural language vocabulary and controlled terms (eg, Medical Subject Headings [MeSH]), where available. Natural language terms will be derived from three main concepts: (1) online CSE, online CSA, or technology-based CSA; (2) children and adolescents; and (3) parents or caregivers. Other search methods will also be used to retrieve additional evidence. These methods will include hand searches of the reference lists and forward citation searches.

Inclusion and exclusion criteria were established using the population, concept, context (PCC) framework (Textbox 1). The PCC framework is recommended by the Joanna Briggs Institute to identify the main concepts in the primary review questions [25]. Articles will be included if the study examines parents' perceptions, attitudes, fears, and challenges related to the prevention of OCSEA among their children and adolescents who are younger than 18 years. Articles published in English from 2018 to date will be included to capture the most current evidence following key global developments in online child protection, including the adoption of the WeProtect Global Alliance's Model National Response and United Nations Children's Fund's (UNICEF) call for coordinated action in child online protection [26,27]. Meanwhile, since 2018, the digital environment has experienced substantial qualitative change, with platforms such as Instagram and TikTok adopting algorithm-driven technologies that enable more complex engagement patterns and richer interactive features, while reaching over one billion users globally [28,29]. In addition, this time frame also includes the COVID-19 pandemic period, capturing both prepandemic- and pandemic-related shifts in related challenges. All primary research studies, irrespective of study design, will be considered. In terms of gray literature, we do not intend to include a targeted search of gray literature sources. However, student theses or dissertations identified during the database search and full-text screening process will be screened and included using the same eligibility criteria and will be documented alongside other included records.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Parents, caregivers, and guardians of adolescents and children younger than 18 years
- Studies that explore challenges faced by parents in preventing online child sexual exploitation and abuse
- Online or digital environments across various platforms (eg, social media and messaging apps)
- Any country or cultural context
- All primary research studies, including qualitative, quantitative, and mixed methods study designs
- Published research, including peer-reviewed journal articles and gray literature such as student dissertations and theses, provided they report original research findings relevant to the review questions
- Articles published from 2018 onward
- Publications in English

Exclusion criteria

- Studies that do not involve parents, caregivers, and guardians of adolescents and children younger than 18 years
- Studies that do not mention parental perceptions, attitudes, fears, or challenges related to online child sexual exploitation and abuse
- Studies focused on law enforcement, online safety, educators, or other professionals without parental involvement
- Studies involving adult survivors of online sexual exploitation
- Studies that do not focus on the prevention aspect of online child sexual exploitation and abuse
- Studies that focus exclusively on children or adolescents younger than 18 years without any exploration of parental challenges or roles
- Studies mentioning parental perceptions, attitudes, fears, or challenges related to child sexual exploitation or child sexual abuse or any other nononline format but not online child sexual exploitation and abuse
- Secondary research, such as reviews and content analysis, commentary, and other
- Opinion pieces, blog posts, or editorials without empirical evidence
- Published studies or gray literature that do not report original research findings
- Articles published before 2018
- Non-English publications

Study Selection

Covidence is a web-based tool for systematic review management that offers a platform for importing article records and conducting independent screening. In this study, after exporting all articles from databases, Covidence will be used to import all articles. Duplicates will be detected and removed automatically in Covidence, facilitating the title and abstract screening and full-text screening phases.

A total of 3 reviewers will be involved in the screening process. Two reviewers will independently screen titles and abstracts, categorizing them as included, excluded, or unclear. Studies marked as included or unclear by either reviewer will proceed to full-text screening. Disagreements at any stage will be resolved through discussion or by a third reviewer. This approach ensures a high level of reliability and consistency in the selection of eligible studies. Additionally, eligible studies will also be identified by manually screening the reference lists of related studies. After determining all eligible studies, a diagram demonstrating the flow of the screening process will be produced and reported in a PRISMA flow diagram (see Multimedia Appendix 2).

Data Extraction and Critical Appraisal

A data extraction sheet has been created by researchers on the team and will be used to extract data from all studies that meet the inclusion criteria. After the initial version of the data extraction sheet is created, 2 reviewers will select 10% (2/20) of the total number of included studies to pilot the sheet [30]. They will independently complete data extraction for these 2 studies using the draft sheet to assess its comprehensiveness. If any issues arise during the piloting process, all members of the research team will discuss how to revise and improve the data extraction sheet accordingly before proceeding with the full data extraction. Following the pilot extraction, the research team will compare the results, identify any inconsistencies, and revise the extraction sheet to produce the final version.

Data extracted from each study will include bibliographic details, study design, study population, identified challenges, factors contributing to the challenges, conclusions, research implications, limitations, and identified research gaps. The bibliographic details will include the record number in Covidence, authors, titles, publication year, and the country or region where the study was conducted. The study design will specify whether the study is qualitative (interviews, focus groups, etc), quantitative (surveys, questionnaires, etc), or mixed methods. It will also indicate the type of OCSEA discussed and

XSL•FO
RenderX

the study's research question or hypothesis. The study population will include participant information, such as sample size, age distribution, gender distribution, marital status, ethnic background, and religion. The findings will highlight the challenges identified in the selected studies, including the influencing factors that contributed to these challenges. Two reviewers will independently extract the data, which will be compared for consistency; discrepancies will be identified and resolved through discussion. If there is any uncertainty about the included papers or missing data, the authors will be contacted for clarification.

The methodological quality and potential risk of bias of included studies will be assessed using the Joanna Briggs Institute (JBI) Critical Appraisal Tools by two independent reviewers [31]. Any discrepancies will be resolved through discussion as well. As a single JBI checklist cannot cover all study designs, we will combine relevant JBI checklists for different study designs into a single appraisal framework, with minor adaptations as needed. Study quality will be categorized as high (≥75% of criteria met), moderate (50%-74%), or low (<50%).

Data Analysis and Presentation

The findings of the study will be presented narratively using the PRISMA checklist. A tailored approach, a descriptive numerical summary, and a thematic analysis will be used to synthesize the data and create an evidence table that will summarize key findings from each study [32]. This will include each study's information, including goals or objectives, population, methods, and outcomes of interest. Next, the data will be aggregated and analyzed thematically according to the study's outcomes. A descriptive (narrative) analysis will be conducted to identify common themes in the challenges faced by parents in providing information about OCSEA to their children.

Results

The search strategy has been finalized with input from a health sciences librarian. Database searches and gray literature screening are scheduled to begin in August 2025. Results from the review are expected by Febuary 2026 and will be reported in accordance with the PRISMA guidelines, and search methods will be reported following PRISMA-Search guidance. This protocol is structured to address PRISMA-recommended protocol elements, with detailed search strategies and supporting materials provided in the appendices.

Discussion

Overview

OCSEA is a growing global concern, with children increasingly at risk in the digital environment. While the role of parents as protectors has been acknowledged in the literature addressing traditional forms of CSA and CSE, less is known about the distinct challenges they face in addressing OCSEA. The unique nature of online abuse, such as often invisible, technology-mediated, and rapidly evolving, creates specific barriers for parental involvement in prevention efforts. Despite the critical importance of equipping parents to safeguard their children online, research on their lived experiences, perceived barriers, and informational needs remains fragmented.

We anticipate that this review will map different categories and characteristics of parental challenges related to OCSEA using an established methodological framework and a comprehensive search strategy. We expect that the findings will highlight the complexity of parental roles within contemporary digital environments and provide valuable insight into the areas where parents feel unsupported or underprepared, as well as gaps in the available evidence that could inform future targeted interventions.

By highlighting the nature and extent of these challenges, this review has the potential to guide the development of educational resources, public health messaging, and policy reforms aimed at empowering parents. It also aims to support interdisciplinary collaboration among researchers, educators, child protection specialists, and policymakers to foster safer digital environments for children. The findings will contribute meaningfully to the understanding of OCSEA prevention and inform future research and interventions designed to enhance parental capacity in online safety and abuse prevention.

Limitations

The restriction to studies published from 2018 onward and the absence of a targeted gray literature search may result in the omission of earlier foundational work or unpublished evidence.

Dissemination Plan

Results from this review will be disseminated through publication in a peer-reviewed journal and presentation at relevant academic and professional conferences. Findings may also be shared with stakeholders through knowledge translation activities.

Acknowledgments

The authors would like to acknowledge the contribution of health sciences librarian Megan Kennedy for her efforts in developing the search strategies for this systematic review and identifying suitable databases.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

XSL•FO
RenderX

Data Availability

All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Authors' Contributions

NP conceptualized the review protocol. All authors (NP, WY, LH, SS, AH, and FE) collaborated on drafting the manuscript, contributed through multiple revisions, and provided substantial feedback and critical evaluation. All authors have reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
MEDLINE search strategy.
[DOCX File, 16 KB - resprot_v15i1e80968_app1.docx]

Multimedia Appendix 2
PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.
[DOCX File, 1721 KB - resprot_v15i1e80968_app2.docx]

References

1. Report of the consultation on child abuse prevention, 29-31 March 1999. : World Health Organization; 1999 URL: https://iris.who.int/handle/10665/65900 [accessed 2026-01-16]
2. Wells M. Canada's law on child sexual abuse. U.S. Department of Justice. 1990. URL: https://www.ojp.gov/ncjrs/virtual-library/abstracts/canadas-law-child-sexual-abuse [accessed 2026-01-16]
3. Shuster A, Mason H, Erwood S. Online child sexual exploitation and abuse (OCSEA). : Child Helpline International; 2023 URL: https://childhelplineinternational.org/wp-content/uploads/2023/11/2023-OCSEA-Report.pdf [accessed 2026-01-16]
4. Kelly L, Karsna K. Measuring the scale and changing nature of sexual abuse and child sexual exploitation: scoping report. : London Metropolitan University; 2017 URL: https://www.csacentre.org.uk/app/uploads/2023/10/CSA-Scale-and-Nature-full-report-2018.pdf [accessed 2026-01-16]
5. Rimer JR. Child sexual exploitation. In: Oxford Research Encyclopedia of Criminology and Criminal Justice: Oxford University Press; 2024.
6. Laird JJ, Klettke B, Hall K, Hallford D. Toward a global definition and understanding of child sexual exploitation: the development of a conceptual model. Trauma Violence Abuse 2023 Oct;24(4):2243-2264. [doi: 10.1177/15248380221090980] [Medline: 35603675]
7. Finkelhor D, Turner H, Colburn D. Prevalence of online sexual offenses against children in the US. JAMA Netw Open 2022 Oct 3;5(10):e2234471. [doi: 10.1001/jamanetworkopen.2022.34471] [Medline: 36239942]
8. Álvarez-Guerrero G, Fry D, Lu M, Gaitis KK. Online child sexual exploitation and abuse of children and adolescents with disabilities: a systematic review. Disabilities 2024;4(2):264-276. [doi: 10.3390/disabilities4020017]
9. Online child sexual exploitation. Public Safety Canada. 2024. URL: https://www.canada.ca/en/public-safety-canada/campaigns/online-child-sexual-exploitation.html [accessed 2026-01-16]
10. Kahimise J, Shava FB. An analysis of children's online activities and behaviours that expose them to cybercrimes. Presented at: Proceedings of the 27th Telecommunications Forum; Nov 26-27, 2019. [doi: 10.1109/TELFOR48224.2019.8971089]
11. Finkelhor D, Turner H, Colburn D. The prevalence of child sexual abuse with online sexual abuse added. Child Abuse Negl 2024 Mar;149:106634. [doi: 10.1016/j.chiabu.2024.106634] [Medline: 38227986]
12. Schmidt F, Varese F, Bucci S. Understanding the prolonged impact of online sexual abuse occurring in childhood. Front Psychol 2023 Oct 24;14:1281996. [doi: 10.3389/fpsyg.2023.1281996] [Medline: 37941760]
13. Kloess JA, Beech AR, Harkins L. Online child sexual exploitation: prevalence, process, and offender characteristics. Trauma Violence Abuse 2014 Apr;15(2):126-139. [doi: 10.1177/1524838013511543] [Medline: 24608540]
14. Walsh K, Mathews B, Parvin K, et al. Prevalence and characteristics of online child sexual victimization: findings from the Australian Child Maltreatment Study. Child Abuse Negl 2025 Feb;160(107186):107186. [doi: 10.1016/j.chiabu.2024.107186] [Medline: 39644728]
15. Savage L. Online child sexual exploitation: a statistical profile of police-reported incidents in Canada, 2014 to 2022. Statistics Canada. 2024. URL: https://www150.statcan.gc.ca/n1/pub/85-002-x/2024001/article/00003-eng.htm [accessed 2026-01-16]
16. Alaggia R, Collin-Vézina D, Lateef R. Facilitators and barriers to child sexual abuse (CSA) disclosures: a research update (2000–2016). Trauma Violence Abuse 2019 Apr;20(2):260-283. [doi: 10.1177/1524838017697312] [Medline: 29333973]

XSL•FO
RenderX

17. Mendelson T, Letourneau EJ. Parent-focused prevention of child sexual abuse. Prev Sci 2015 Aug;16(6):844-852. [doi: 10.1007/s11121-015-0553-z] [Medline: 25757528]
18. Rudolph J, Zimmer-Gembeck MJ, Shanley DC, Hawkins R. Child sexual abuse prevention opportunities: parenting, programs, and the reduction of risk. Child Maltreat 2018 Feb;23(1):96-106. [doi: 10.1177/1077559517729479] [Medline: 28920456]
19. Rudolph JI, Walsh K, Shanley DC, Zimmer-Gembeck MJ. Child sexual abuse prevention: parental discussion, protective practices and attitudes. J Interpers Violence 2022 Dec;37(23-24):NP22375-NP22400. [doi: 10.1177/08862605211072258] [Medline: 35098765]
20. Rudolph J, Zimmer-Gembeck MJ. Parents as protectors: a qualitative study of parents' views on child sexual abuse prevention. Child Abuse Negl 2018 Nov;85:28-38. [doi: 10.1016/j.chiabu.2018.08.016] [Medline: 30180986]
21. Babatsikos G. Parents' knowledge, attitudes and practices about preventing child sexual abuse: a literature review. Child Abuse Rev 2010;19(2):107-129. [doi: 10.1002/car.1102]
22. Russell DH, Trew S, Harris L, et al. Engaging parents in child-focused child sexual abuse prevention education strategies: a systematic review. Trauma Violence Abuse 2024 Oct;25(4):3082-3098. [doi: 10.1177/15248380241235895] [Medline: 38477488]
23. Novianti N. Online child sexual exploitation and abuse (OCSEA) of children and adolescents: a systematic literature review. Psikologika J Pemikir dan Penelit Psikol 2024;29(2):281-306. [doi: 10.20885/jstl.vol29.iss2.art7]
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
25. Pollock D, Peters MD, Khalil H, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth 2023 Mar 1;21(3):520-532. [doi: 10.11124/JBIES-22-00123] [Medline: 36081365]
26. Model national response to end child sexual exploitation & abuse online. WeProtect Global Alliance. URL: https://www.weprotect.org/resources/frameworks/model-national-response [accessed 2026-01-16]
27. The state of the world's children 2017: children in a digital world. UNICEF. 2017. URL: https://www.unicef.org/reports/state-worlds-children-2017 [accessed 2026-01-16]
28. Bhandari A, Bimo S. Why's everyone on TikTok now? The algorithmized self and the future of self-making on social media. Social Media + Society 2022 Jan;8(1). [doi: 10.1177/20563051221086241]
29. Timeline of Instagram. Wikipedia. 2025. URL: https://en.wikipedia.org/wiki/Timeline_of_Instagram [accessed 2026-01-01]
30. Brink PJ, Wood MJ, Lackey NR, Wingate AL. The pilot study: one key to research success. In: Brink PJ, Wood MJ, editors. Advanced Design in Nursing Research: SAGE Publications; 1998.
31. Critical appraisal tools. JBI. URL: https://jbi.global/critical-appraisal-tools?utm_source=chatgpt.com [accessed 2026-01-01]
32. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. J Health Serv Res Policy 2005 Jan;10(1):45-53. [doi: 10.1177/135581960501000110] [Medline: 15667704]

Abbreviations

CSA: child sexual abuse
CSE: child sexual exploitation
ICT: information and communication technology
MeSH: Medical Subject Headings
OCSEA: online child sexual exploitation and abuse
PCC: population, concept, context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
UNICEF: United Nations Children's Fund

Edited by J Sarvestan; submitted 20.Jul.2025; peer-reviewed by E Walker; revised version received 02.Jan.2026; accepted 05.Jan.2026; published 09.Feb.2026.

Please cite as:
Punjani N, Yan W, Hartling L, Hussain A, Elgaweesh F, Scott S
Challenges Faced by Parents in Preventing Online Child Sexual Exploitation and Abuse: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e80968
URL: https://www.researchprotocols.org/2026/1/e80968
doi:10.2196/80968

XSL•FO
RenderX

© Neelam Punjani, Wenting Yan, Lisa Hartling, Amber Hussain, Farah Elgaweesh, Shannon Scott. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Donor Risk Factors Affecting Graft Survival in Pediatric Kidney Transplants: Protocol for a Systematic Review and Meta-Analysis

Cahyani Gita Ambarsari[1,2,3], MD; Jon Jin Kim[4,5], PhD; Nur Hasnah Maamor[6], PhD; Izzah Athirah Rosli[6], BSc; Nor Asiah Muhamad[6], MPH

[1]School of Medicine University of Nottingham, University Park, Nottingham, United Kingdom

[2]Department of Child Health, Faculty of Medicine Universitas Indonesia-Cipto Mangunkusumo Hospital, Central Jakarta, Jakarta, Indonesia

[3]Medical Technology Cluster, Indonesian Medical Education and Research Institute, Faculty of Medicine Universitas Indonesia, Central Jakarta, Jakarta, Indonesia

[4]Department of Paediatric Nephrology, Nottingham University Hospitals, Nottingham, United Kingdom

[5]Department of Surgery, University of Cambridge, Cambridge, United Kingdom

[6]Sector for Evidence-based Healthcare, National Institutes of Health, Ministry of Health, Shah Alam, Selangor, Malaysia

Corresponding Author:
Cahyani Gita Ambarsari, MD
School of Medicine University of Nottingham, University Park, Nottingham, United Kingdom

Abstract

Background: Pediatric patients with end-stage kidney disease require kidney transplants (KTs) throughout their lifetime. Long-term graft survival is dependent on multiple factors, which are broadly categorized as donor- and recipient-related factors. Advances in transplant care and changes in donor population demographics necessitate an updated analysis on donor risk factors to guide clinical decision-making.

Objective: In this systematic review and meta-analysis, we will focus on the impact of donor factors on graft survival in pediatric KT, excluding transplants from donation after circulatory death as the latter are less common in children.

Methods: This review encompasses studies reporting donor-related risk factors for graft survival in pediatric KT, including age, size, comorbidities, and ethnicity for living and deceased donors, as well as the cause of death and length of hospitalization for deceased donors. The literature search will use the following databases: PubMed, Scopus, Web of Science, Embase, and Cochrane. Two independent reviewers will select studies and assess their quality. Pooled estimates of relevant factors will be computed via a random-effects model using the Stata/BE (version 19) software. Depending on data availability, subgroup analyses will be conducted based on donor type (living vs deceased). The reporting of findings will adhere to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines.

Results: The search and screening for the systematic literature review are anticipated to be finished in June 2026. Data extraction, quality appraisal, and subsequent data synthesis will begin in July 2026. The review is expected to be completed by October 2026, and the study results will be published in 2027.

Conclusions: Our review will provide a comprehensive synthesis of the available evidence on kidney donor risk factors impacting graft survival in pediatric KT. The results of this review could provide valuable insights for clinical decisions, policy development, and ongoing efforts to improve outcomes for children with end-stage kidney disease requiring KT.

Trial Registration: PROSPERO CRD42024500442; https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42024500442

International Registered Report Identifier (IRRID): PRR1-10.2196/71620

(JMIR Res Protoc 2026;15:e71620) doi:10.2196/71620

KEYWORDS

graft failure; human leukocyte antigen; HLA; hypertension; increased-risk donor; kidney transplant; mismatch; pediatric; weight

Introduction

Kidney disease, including chronic kidney disease and end-stage kidney disease (ESKD), is a major global health issue, with ESKD posing profound challenges to the health and well-being of pediatric patients and their families [1-9]. Among the treatment options for ESKD, kidney transplant (KT) is the most preferred, demonstrating better outcomes with reduced mortality and morbidity [10-13]. Additionally, pediatric KT creates opportunities for improved growth, development, education, and social interaction, leading to an overall enhanced quality

XSL•FO
RenderX

of life for young recipients [1]. Pediatric transplant recipients are projected to live over 50 years and, therefore, will require more than one transplant during their lifetime. Therefore, minimizing the need for multiple transplants and maximizing the lifetime of each transplant is crucial. Acknowledging the multiple risk factors, including both medical and psychosocial factors, we aim to examine donor-related factors that affect long-term allograft survival. The impact of immunological risk factors, including human leukocyte antigen (HLA) mismatching [14-16] and sensitization reflected by panel-reactive antibodies [17], along with recipient-related risk factors such as prior kidney replacement therapy modality [18,19] and its duration [20-22], has been previously reviewed.

Prognostication is important to guide clinical decision-making both at the population level, as facilitated by kidney allocation schemes (KASs), and at the patient level, as informed decision-making between the patient and clinical team. When a kidney offer is presented to a pediatric patient, nephrologists and transplant surgeons must rapidly make intricate decisions regarding the organ's suitability for transplantation. This process entails evaluating various clinical, psychological, and social factors relying on often incomplete information about the donor and existing knowledge about the patient from the organ procurement organization [23,24]. Factors known to influence decisions include donor quality, recipient risk factors, and physicians' decision-making behaviors [14,18,25-29]. Nonetheless, trying to weigh each factor is difficult. Clinical risk indexes that combine multiple factors have been developed, but they have mainly used adult population data, and their applicability in pediatric patients has not been proven [26,30-32].

The increasing demand for KT among pediatric patients far surpasses the limited supply of organs from deceased donors, highlighting the necessity to expand the donor pool [33,34]. The scarcity of kidney allografts poses a significant obstacle for patients with ESKD seeking transplants, emphasizing the crucial need to expand the donor pool and optimize the use of available allografts [34]. Additionally, the demographics of the donor population have changed, and donors are becoming increasingly older with more comorbidities [26,33]. Therefore, health care professionals need to be able to make careful considerations on accepting donors with risk factors rather than wait for an ideal donor, which might prolong waiting times [26].

This review aims to examine donor risk factors and their impact on graft survival in pediatric patients undergoing KT. By combining recent evidence, it seeks to offer valuable insights for clinical decisions, policy development, and ongoing efforts to improve outcomes for this vulnerable patient group. Our review aims to address the following question: in pediatric KT, what donor risk factors influence allograft survival? To the best of our knowledge, no systematic review or meta-analysis has previously examined this question.

Methods

This review will be conducted and reported following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol guidelines. We will also integrate this review with the MOOSE (Meta-Analyses of Observational Studies in Epidemiology) guidelines as the outcome will be a meta-analysis of selected observational studies.

Eligibility Criteria

We will include all original articles that meet the following criteria:

- Studies involving pediatric patients as KT recipients (aged <18 years at the time of transplant)
- Studies reporting the first kidney-only transplants, defined as the patient's first transplant and excluding simultaneous liver-kidney or other combined organ transplants, from living donors and deceased donation after brain death
- Studies that investigate donor-related risk factors for graft survival and report graft survival as the primary or secondary outcome
- Peer-reviewed original articles, including randomized controlled trials and observational studies (cohort, case-control, and cross-sectional studies)
- Articles published or accepted for publication beginning from 2000

The exclusion criteria are as follows:

- Studies involving adult recipients in which the pediatric group data cannot be extracted
- Studies involving transplants from deceased donors following circulatory death in which the donation after brain death group data cannot be extracted
- Studies involving multiorgan transplants in which the kidney-only group data cannot be extracted
- Studies involving kidney retransplantations in which the first KT group data cannot be extracted
- Reviews, case studies, commentaries, qualitative studies, or editorials

Outcome Measures

The following are the outcomes of this review:

- Death-censored graft survival: graft survival, considering only cases in which graft failure occurred due to factors other than patient death. Graft survival is defined as the time from transplant to graft failure. Graft failure is defined as the start of dialysis or retransplantation. Our primary aim is to evaluate 5-year graft survival. However, we will adhere to the availability of the data identified through our search strategy and anticipate that outcomes may be reported at varying time points, including 1-, 5-, and 10-year graft survival.
- All-cause graft survival: combined outcome of graft survival and patient survival.

Information Sources

Electronic Search

We will systematically conduct a comprehensive literature search using various literature databases (MEDLINE [PubMed], Web of Science, Scopus, Embase, and Cochrane Central Register of Controlled Trials (CENTRAL) to identify eligible studies. Secondary searches will be conducted on Google Scholar and other websites, and the reference section of the

XSL•FO
RenderX

included studies will also be hand searched for additional relevant studies. We will include both English- and non–English-language studies in this review. For studies published in languages other than English, translation support will be sought to ensure accurate screening, data extraction, and interpretation. This approach minimizes language bias and ensures that potentially relevant evidence is not excluded based on language of publication. The search will be performed from database inception to May 2026.

Search Strategy

The proposed search terms will cover 3 key themes. The first theme will be "pediatric kidney transplantation," including keywords for children and adolescents. The second theme is "risk factor," including unrelated, living, or tissue donors and donor selection. The third theme is "graft survival." The exploded versions of the MeSH (Medical Subject Headings) terms for each theme will be included. All 3 search themes will be combined using the Boolean operators "OR" and "AND." The detailed search terms for each database are presented in Multimedia Appendix 1.

Study Selection

We will identify and remove duplicates and collate multiple reports. Two review authors will independently screen all the titles and abstracts to examine the potential studies for inclusion and exclude studies that are clearly irrelevant. We will identify the studies and code them as "retrieve" (eligible or potentially eligible or unclear) or "do not retrieve." We will retrieve the full-text study reports and publications, and the review authors will independently screen the full texts to identify studies for inclusion as well as identify and record reasons for the exclusion of ineligible studies. We will resolve any disagreements through discussion. If no consensus is reached, the other 2 authors will act as arbiters.

We will document the study selection process in sufficient detail to complete a PRISMA flow diagram and construct a table describing the characteristics of the excluded studies [35,36]. In accordance with PRISMA standards, the flow diagram summarizing records identified, screened, excluded, and included will be incorporated into the final review. Any deviations from or amendments to this protocol will be transparently documented, together with their rationale, and reported in the final review. The EndNote reference management software (Clarivate Analytics) will be used to store, organize, and manage all the articles identified from the databases.

Data Extraction and Management

We will use a standardized data extraction form created in the Microsoft Excel spreadsheet software for study characteristics and outcome data. Two review authors will independently extract outcome data from the included studies. In the "characteristics of the included studies" table, we will note whether outcome data are reported in a usable way. We will resolve any disagreements through consensus or by involving the other 2 authors. We will double-check that data are entered correctly by comparing the data present in the systematic review with the data in the study reports. The following characteristics from the included studies will be extracted:

- Title, authors, study country, region, and publication year
- Methods: study design, source of data, total duration of the study, and method of analysis
- Participants: number of patients, age range, and sex
- Outcome: graft survival
- Exposures or risks: donor age, sex, diabetes, hypertension, smoking, body height, body weight, ethnicity, cause of death, days in hospital, cytomegalovirus, and hepatitis C virus (however, where applicable, we will extract data on other exposure risks or modifiable risk factors, control conditions, and adjustment variables)

Quality Assessment

Two review authors will independently assess the studies' quality based on the criteria in the Newcastle-Ottawa Scale (NOS) for observational studies. The NOS is a widely used tool for assessing the quality of nonrandomized studies, including cohort studies and case-control studies. The NOS applies a star system, in which the study is assessed based on three broad perspectives: (1) the selection of the study groups, (2) the comparability of the groups, and (3) the ascertainment of exposure and outcome [37]. The maximum score is 9 points, and the studies can be classified as being of good, fair, or poor quality according to the following standard thresholds:

- Good quality: 3 or 4 stars in the selection domain AND 1 or 2 stars in the comparability domain AND 2 or 3 stars in the outcome or exposure domain
- Fair quality: 2 stars in the selection domain AND 1 or 2 stars in the comparability domain AND 2 or 3 stars in the outcome or exposure domain
- Poor quality: 0 or 1 star in the selection domain OR 0 stars in the comparability domain OR 0 or 1 star in the outcome or exposure domain

To assess the risk of bias in the included intervention studies, we will use a Cochrane Risk of Bias Assessment Tool for Nonrandomized Studies of Interventions [38]. This assessment tool consists of 5 domains of bias (selection bias, confounding bias, performance bias, detection bias, and attrition bias), which will be assessed using a set of signaling questions. Each domain will be rated as having a low, high, or unclear risk of bias, and an overall risk-of-bias rating will be assigned for each study. The use of a Cochrane Risk of Bias Assessment Tool for Nonrandomized Studies of Interventions will allow us to provide a standardized assessment of the quality of evidence from nonrandomized studies of interventions and determine the overall risk of bias across studies.

To assess the risk of bias in randomized controlled trials, we will use version 2 of the Cochrane risk-of-bias tool for randomized trials, which evaluates bias across domains including the randomization process, deviations from intended interventions, missing outcome data, measurement of outcomes, and selection of reported results [38].

Statistical Analysis

Data Analysis and Statistical Analysis

Statistical analyses will be conducted using the Stata software (StataCorp). The main R packages *meta* and *metafor* (R

XSL•FO
RenderX

Foundation for Statistical Computing) will be used for meta-analysis. We will calculate the pooled estimates of graft survival using a random-effects model to allow for heterogeneity across studies. Where appropriate, effect estimates will be stratified according to donor type (living vs deceased) to explore potential differences in outcomes.

Assessment of Heterogeneity

Assessing heterogeneity is a critical step in conducting a meta-analysis as it allows for an evaluation of the degree of inconsistency or variability among the results of individual studies. For this review, we will evaluate heterogeneity using both the I^2 and Q statistics. The I^2 statistic will be used to quantify the impact of heterogeneity, with percentages of 25%, 50%, and 75% representing a low, moderate, and high degree of heterogeneity, respectively [39]. The Q test is a statistical test that determines whether there is significant heterogeneity among the studies. The significance level for the Q test will be set at .01 in this review [40]. If significant heterogeneity is detected using the Q test or I^2 index (>50%), we will explore potential sources of heterogeneity using subgroup analyses and meta-regression. We will also explore the possible causes (eg, differences in study quality, participants, or outcome assessments) and evaluate the studies in terms of their methodological characteristics to determine whether the degree of heterogeneity can be explained by differences in those characteristics and whether a meta-analysis is appropriate.

Assessment of Publication Bias

We will create and examine a funnel plot to explore possible small-study biases if we can pool more than 10 studies in a single meta-analysis. The number of studies missing from the funnel plot will be estimated. The effect size after the inputting of these missing studies will be estimated using the trim-and-fill method. The trim-and-fill method is a simple estimation approach proposed by Duval and Tweedie [41] in which the asymmetric outlying part of the funnel is trimmed, the symmetric remainder is then used to estimate the true center of the funnel, and the trimmed studies and their counterparts are then replaced around the center. Other methods to assess publication bias, including the Begg rank correlation and the Egger weighted regression method test, will also be performed [42,43].

Ethical Considerations

We registered this systematic review with PROSPERO (CRD42024500442). As this review will use published data, ethics approval has been waived.

Dissemination

The systematic review will focus on donor risk factors in pediatric KT, the results of which will be made publicly available and disseminated through publication in a peer-reviewed journal after completion and presentation at conferences.

Results

The search and screening for the systematic literature review are anticipated to be finished in June 2026. Data extraction, quality appraisal, and subsequent data synthesis will begin in July 2026. The review is expected to be completed by October 2026, and the study results will be published in 2027.

Discussion

We will conduct a systematic review on donor risk factors that affect kidney allograft and patient survival in pediatric patients. This is an important clinical question as it influences the decisions of clinicians and families on the choice of donor to accept, for example, whether there is an upper limit of donor age to accept or whether donor comorbidities (eg, diabetes, hypertension, or obesity) translate to poorer outcomes in the recipient.

Since the first successful KT in 1954, surgical transplantation techniques have evolved considerably. New approaches to transplant protocols have been revised over the past 2 decades to overcome organ donor shortage, for example, careful selection of older living donors, donors deceased from cardiac causes, and donors with a history of increased infection risk [44]. Concurrently, the management of infants on dialysis has improved, and the number of transplants in smaller-sized recipients has increased [45-47]. Immunosuppressive treatment regimens have also evolved, such as the introduction of tacrolimus and mycophenolate mofetil in 1994, which replaced azathioprine and cyclosporine [48]. In parallel, clinical factors such as posttransplant proteinuria have emerged as important markers of graft injury, with strong associations with impaired graft survival and adverse outcomes in pediatric recipients [49-52]. Therefore, a contemporary analysis of the factors affecting long-term transplant survival is required. Our emphasis on uniformity in the inclusion criteria and consideration of the evolution of different practices will help ensure sample representativeness. The design, use of standardized study rating instruments, and adherence to systematic review and meta-analysis guidelines further enhance the study's robustness.

In view of the advances in KT, our review will focus on studies from the year 2000 onward. During this time, the systems for allocating deceased donor KT have also evolved considerably [53]. In the United Kingdom, a national KAS was first introduced in 1989. HLA matching was introduced in 2006, and utility matching using donor and recipient risk indexes was introduced in 2019 [54]. The United States implemented a national sharing and donor service area system with their KAS in 1987. In 2014, a revised US KAS scheme was implemented. Currently, this system uses the Kidney Donor Profile Index to assess donor kidney quality, with pediatric patients given priority [53,55]. Australia and New Zealand's allocation system, established in 2008 and revised in 2011, emphasizes national prioritization for well-matched grafts using state-based algorithms for local allocations [56,57]. European countries have collaborated since 1967 on a registry of KT candidates to optimize HLA matching and introduced the Eurotransplant Kidney Allocation System in 1996. This Eurotransplant Kidney

XSL•FO
RenderX

Allocation System point scoring system considers factors such as HLA match grade, mismatch probability, waiting time, geographic distance, national balance, medical urgency status, and pediatric age [58]. Therefore, there are variations in KASs internationally reflecting local priorities and variations in the way in which pediatric patients are prioritized [59]. Nonetheless, the donor risk indexes, which assess donor quality, have not been shown to be as effective in prognosticating outcomes in pediatric patients [26,30-32].

We hope that a large meta-analysis that combines results worldwide will overcome the problem of small sample sizes in pediatric KT and improve the validity of the analysis. However, there is a need to acknowledge the potential limitations arising from the diverse designs of the included studies. Our review will use subgroup analyses and meta-regression to address high heterogeneity and enhance the overall quality of evidence. Despite potential biases in observational study designs, the use of standardized study rating instruments, compliance with all relevant guidelines for systematic reviews and meta-analyses, and synthesis of evidence from multiple studies can strengthen the validity of our conclusions.

By providing solid and updated reports on donor risk factors in pediatric KT, this review aims to contribute substantially to the refinement of transplant strategies, policy development, and clinical practice in this critical area of pediatric nephrology.

Acknowledgments

The authors would like to thank Chika Carnation Tandri, MD, for general assistance with the references for this protocol.

Funding

This work was supported by a PhD studentship awarded to CGA from the Indonesia Endowment Fund for Education (LPDP), Ministry of Finance, Republic of Indonesia.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

CGA, NAM, and JJK conceptualized the study. CGA, NHM, and IAR conducted the investigation. CGA, NAM, JJK, NHM, and IAR developed the methodology. CGA and NAM were selected for project administration. CGA, NAM, and JJK managed the resources. NAM and JJK were in charge of supervision and validation. CGA was in charge of visualization and wrote the original draft. CGA, NAM, NHM, IAR, and JJK were responsible for review and editing. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Proposed search terms.
[DOCX File, 19 KB - resprot_v15i1e71620_app1.docx]

Checklist 1
PRISMA-P checklist.
[DOC File, 92 KB - resprot_v15i1e71620_app2.doc]

References

1. Sypek MP, Davies CE, Le Page AK, et al. Paediatric deceased donor kidney transplant in Australia: a 30-year review-what have paediatric bonuses achieved and where to from here? Pediatr Transplant 2021 Sep;25(6):e14019. [doi: 10.1111/petr.14019] [Medline: 33942949]
2. Palupi-Baroto R, Hermawan K, Murni IK, et al. High fibroblast growth factor 23 as a biomarker for severe cardiac impairment in children with chronic kidney disease: a single tertiary center study. Int J Nephrol Renovasc Dis 2021;14:165-171. [doi: 10.2147/IJNRD.S304143] [Medline: 34135617]
3. Ambarsari CG, Trihono PP, Kadaristiana A, et al. Low-dose maintenance intravenous iron therapy can prevent anemia in children with end-stage renal disease undergoing chronic hemodialysis. Int J Nephrol 2020;2020:3067453. [doi: 10.1155/2020/3067453] [Medline: 32566294]
4. Palupi-Baroto R, Hermawan K, Murni IK, et al. Carotid intima-media thickness, fibroblast growth factor 23, and mineral bone disorder in children with chronic kidney disease. BMC Nephrol 2024 Oct 21;25(1):369. [doi: 10.1186/s12882-024-03771-z] [Medline: 39433982]

XSL•FO
RenderX

5. Ambarsari CG, Agianda HA, Saraswati M, Kim JJ. Dent disease 1 presented early with Bartter-like syndrome features and rickets: a case report. Case Rep Nephrol Dial 2025;15(1):15-25. [doi: 10.1159/000543719] [Medline: 40881168]
6. Ambarsari CG, Nadhifah N, Lestari HI. Perioperative blood pressure management recommendations in pediatric pheochromocytoma: a 10-year narrative review. Kidney Blood Press Res 2025;50(1):61-82. [doi: 10.1159/000542897] [Medline: 39626645]
7. Santoso DN, Sinuraya FA, Ambarsari CG. Distal renal tubular acidosis presenting with an acute hypokalemic paralysis in an older child with severe vesicoureteral reflux and syringomyelia: a case report. BMC Nephrol 2022 Jul 14;23(1):248. [doi: 10.1186/s12882-022-02874-9] [Medline: 35836135]
8. Ambarsari CG, Palupi-Baroto R, Sinuraya FA, Suryati E, Widyastuti E, Widhiati S. Nephropathy in a child with severe recessive dystrophic epidermolysis bullosa treated with cyclophosphamide: a case report. Case Rep Nephrol Dial 2023;13(1):75-83. [doi: 10.1159/000530875] [Medline: 37484797]
9. Ambarsari CG, Bermanshah EK, Putra MA, Rahman FH, Pardede SO. Effective management of peritoneal dialysis-associated hydrothorax in a child: a case report. Case Rep Nephrol Dial 2020;10(1):18-25. [doi: 10.1159/000506119] [Medline: 32232056]
10. Wyld ML, De La Mata NL, Hedley J, Kim S, Kelly PJ, Webster AC. Life years lost in children with kidney failure: a binational cohort study with multistate probabilities of death and life expectancy. J Am Soc Nephrol 2023 Jun 1;34(6):1057-1068. [doi: 10.1681/ASN.0000000000000118] [Medline: 36918386]
11. Chandar J, Chen L, Defreitas M, Ciancio G, Burke G III. Donor considerations in pediatric kidney transplantation. Pediatr Nephrol 2021 Feb;36(2):245-257. [doi: 10.1007/s00467-019-04362-z]
12. Lepeytre F, Delmas-Frenette C, Zhang X, et al. Donor age, donor-recipient size mismatch, and kidney graft survival. Clin J Am Soc Nephrol 2020 Oct 7;15(10):1455-1463. [doi: 10.2215/CJN.02310220] [Medline: 32843375]
13. Ambarsari CG, Hidayati EL, Hasan I, Grace A, Oswari H. Successful treatment of Hepatitis C virus infection using direct-acting antiviral agents (DAAs) in adolescents with kidney transplantation: a case series. Int J Nephrol Renovasc Dis 2020;13:139-146. [doi: 10.2147/IJNRD.S248632] [Medline: 32606888]
14. Kim JJ, Fuggle SV, Marks SD. Does HLA matching matter in the modern era of renal transplantation? Pediatr Nephrol 2021 Jan;36(1):31-40. [doi: 10.1007/s00467-019-04393-6] [Medline: 31820146]
15. Fichtner A, Gauché L, Süsal C, et al. Incidence, risk factors, management strategies, and outcomes of antibody-mediated rejection in pediatric kidney transplant recipients-a multicenter analysis of the Cooperative European Paediatric Renal Transplant Initiative (CERTAIN). Pediatr Nephrol 2025 Feb;40(2):491-503. [doi: 10.1007/s00467-024-06487-2] [Medline: 39283519]
16. Shi X, Liu R, Xie X, et al. Effect of human leukocyte antigen mismatching on the outcomes of pediatric kidney transplantation: a systematic review and meta-analysis. Nephrol Dial Transplant 2017 Nov 1;32(11):1939-1948. [doi: 10.1093/ndt/gfx259] [Medline: 28992320]
17. Engen RM, Smith JM, Bartosh SM. The kidney allocation system and pediatric transplantation at 5 years. Pediatr Transplant 2022 Nov;26(7):e14369. [doi: 10.1111/petr.14369] [Medline: 35919967]
18. Ambarsari CG, Cho Y, Milanzi E, et al. Epidemiology and outcomes of children with kidney failure receiving kidney replacement therapy in Australia and New Zealand. Kidney Int Rep 2023 Oct;8(10):1951-1964. [doi: 10.1016/j.ekir.2023.07.006] [Medline: 37850025]
19. Hou YF, Wang XX, Yang HJ, Zhong S. Impact of pre-transplant dialysis modality on kidney transplant outcomes: a systematic review and meta-analysis. Eur Rev Med Pharmacol Sci 2022 Apr;26(7):2292-2304. [doi: 10.26355/eurrev_202204_28459] [Medline: 35442484]
20. Amaral S, Sayed BA, Kutner N, Patzer RE. Preemptive kidney transplantation is associated with survival benefits among pediatric patients with end-stage renal disease. Kidney Int 2016 Nov;90(5):1100-1108. [doi: 10.1016/j.kint.2016.07.028] [Medline: 27653837]
21. Marcou M, Galiano M, Tzschoppe A, et al. Risk factor analysis for long-term graft survival following pediatric kidney transplantation: the importance of pretransplantation time on dialysis and donor/recipient age difference. J Clin Med 2023 Nov 9;12(22):7014. [doi: 10.3390/jcm12227014] [Medline: 38002629]
22. Cransberg K, Smits JM, Offner G, Nauta J, Persijn GG. Kidney transplantation without prior dialysis in children: the Eurotransplant experience. Am J Transplant 2006 Aug;6(8):1858-1864. [doi: 10.1111/j.1600-6143.2006.01405.x] [Medline: 16771812]
23. Chaudhuri A, Gallo A, Grimm P. Pediatric deceased donor renal transplantation: an approach to decision making I. Pediatric kidney allocation in the USA: the old and the new. Pediatr Transplant 2015 Nov;19(7):776-784. [doi: 10.1111/petr.12569] [Medline: 26426316]
24. Chaudhuri A, James G, Grimm P. Whether or not to accept a deceased donor kidney offer for a pediatric patient. Pediatr Nephrol 2015 Sep;30(9):1529-1536. [doi: 10.1007/s00467-015-3139-x] [Medline: 26130248]
25. Le Page AK, Johnstone LM, Kennedy SE. Australian deceased donor kidney allocation protocols: transplant waiting and graft quality for children and adolescents. Pediatr Transplant 2015 Sep;19(6):588-594. [doi: 10.1111/petr.12544] [Medline: 26103210]

XSL•FO
RenderX

26. Kim JJ, Curtis RM, Reynolds B, et al. The UK kidney donor risk index poorly predicts long-term transplant survival in paediatric kidney transplant recipients. Front Immunol 2023;14:1207145. [doi: 10.3389/fimmu.2023.1207145] [Medline: 37334377]
27. Annual report on kidney transplantation report for 2021/2022 2022. NHS Blood and Transplant. URL: https://nhsbtdbe.blob.core.windows.net/umbraco-assets-corp/27956/kidney-annual-report-2021-22.pdf [accessed 2025-05-25]
28. Marlais M, Pankhurst L, Martin K, Mumford L, Tizard EJ, Marks SD. Renal allograft survival rates in kidneys initially declined for paediatric transplantation. Pediatr Nephrol 2018 Sep;33(9):1609-1616. [doi: 10.1007/s00467-018-3969-4] [Medline: 29808263]
29. Ambarsari CG, Hidayati EL, Trihono PP, et al. Experience of the first 6 years of pediatric kidney transplantation in Indonesia: a multicenter retrospective study. Pediatr Transplant 2020 Dec;24(8):e13812. [doi: 10.1111/petr.13812] [Medline: 32794281]
30. Montgomery A, Goff C, Adeyeri B, et al. A modified kidney donor risk index for pediatric kidney transplant recipients. Pediatr Nephrol 2023 Apr;38(4):1309-1317. [doi: 10.1007/s00467-022-05722-y] [Medline: 36066770]
31. Van Arendonk KJ, Chow EK, James NT, et al. Choosing the order of deceased donor and living donor kidney transplantation in pediatric recipients. Transplantation 2015 Feb;99(2):360-366. [doi: 10.1097/TP.0000000000000588] [Medline: 25594552]
32. Bae S, Massie AB, Thomas AG, et al. Who can tolerate a marginal kidney? Predicting survival after deceased donor kidney transplant by donor–recipient combination. Am J Transplant 2019 Feb;19(2):425-433. [doi: 10.1111/ajt.14978]
33. Mumford L, Maxwell H, Ahmad N, Marks SD, Tizard J. The impact of changing practice on improved outcomes of paediatric renal transplantation in the United Kingdom: a 25 years review. Transpl Int 2019 Jul;32(7):751-761. [doi: 10.1111/tri.13418] [Medline: 30801866]
34. Naesens M, Kambham N, Concepcion W, Salvatierra O, Sarwal M. The evolution of nonimmune histological injury and its clinical relevance in adult-sized kidney grafts in pediatric recipients. Am J Transplant 2007 Nov;7(11):2504-2514. [doi: 10.1111/j.1600-6143.2007.01949.x] [Medline: 17725681]
35. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341. [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]
36. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009 Oct;62(10):e1-34. [doi: 10.1016/j.jclinepi.2009.06.006] [Medline: 19631507]
37. Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Ottawa Hospital Research Institute. 2021. URL: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp [accessed 2025-05-25]
38. A cochrane risk of bias assessment tool: for non-randomized studies of interventions (ACROBAT- NRSI). University of Bristol. URL: https://www.bristol.ac.uk/media-library/sites/social-community-medicine/images/centres/cresyda/ACROBAT-NRSI%20Version%201_0_0.pdf [accessed 2026-02-05]
39. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-1558. [doi: 10.1002/sim.1186] [Medline: 12111919]
40. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010 Apr;1(2):97-111. [doi: 10.1002/jrsm.12] [Medline: 26061376]
41. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000 Jun;56(2):455-463. [doi: 10.1111/j.0006-341x.2000.00455.x] [Medline: 10877304]
42. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994 Dec;50(4):1088-1101 [FREE Full text] [Medline: 7786990]
43. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634. [doi: 10.1136/bmj.315.7109.629] [Medline: 9310563]
44. Tantisattamo E, Maggiore U, Piccoli GB. History of kidney transplantation: a journey of progression and evolution for success. J Nephrol 2022 Sep;35(7):1783-1786. [doi: 10.1007/s40620-022-01453-3] [Medline: 36040564]
45. Pickles CW, Brown C, Marks SD, Reynolds BC, Kessaris N, Dudley J. Long term outcomes following kidney transplantation in children who weighed less than 15 kg - report from the UK Transplant Registry. Pediatr Nephrol 2023 Nov;38(11):3803-3810. [doi: 10.1007/s00467-023-06024-7] [Medline: 37209174]
46. de Feltran LS, Genzani CP, Fonseca M, et al. Strategy to enable and accelerate kidney transplant in small children and results of the first 130 transplants in children ≤15 kg in a single center. Transplantation 2020 Aug;104(8):e236-e242. [doi: 10.1097/TP.0000000000003300] [Medline: 32732842]
47. Kizilbash SJ, Evans MD, Vock DM, Chinnakotla S, Chavers BM. Should children receive a kidney transplant before 2 years of age? Pediatr Transplant 2024 Feb;28(1):e14631. [doi: 10.1111/petr.14631] [Medline: 37937507]
48. Verghese PS. Pediatric kidney transplantation: a historical review. Pediatr Res 2017 Jan;81(1-2):259-264. [doi: 10.1038/pr.2016.207] [Medline: 27732587]
49. Rosík T, Chadimová M, Dušek J, et al. Proteinuria 1 year after renal transplantation is associated with impaired graft survival in children. Pediatr Nephrol 2015 Oct;30(10):1853-1860. [doi: 10.1007/s00467-015-3114-6] [Medline: 25925040]

XSL•FO
RenderX

50. Ambarsari CG, Utami DA, Tandri CC, Satari HI. Comparison of three spot proteinuria measurements for pediatric nephrotic syndrome: based on the International Pediatric Nephrology Association 2022 guidelines. Ren Fail 2023;45(2):2253324. [doi: 10.1080/0886022X.2023.2253324] [Medline: 37724557]
51. Sypek M, Francis A, Chadban S. Predicting graft survival in pediatric kidney transplantation: does the box fit? Am J Transplant 2023 Oct;23(10):1481-1482. [doi: 10.1016/j.ajt.2023.08.021] [Medline: 37652175]
52. Ambarsari CG, Tambunan T, Pardede SO, Fazlur Rahman FH, Kadaristiana A. Role of dipstick albuminuria in progression of paediatric chronic kidney disease. J Pak Med Assoc 2021 Feb;71(Suppl 2)(2):S103-S106. [Medline: 33785952]
53. Wu DA, Watson CJ, Bradley JA, Johnson RJ, Forsythe JL, Oniscu GC. Global trends and challenges in deceased donor kidney allocation. Kidney Int 2017 Jun;91(6):1287-1299. [doi: 10.1016/j.kint.2016.09.054] [Medline: 28320531]
54. Johnson RJ, Fuggle SV, Mumford L, Bradley JA, Forsythe JL, Rudge CJ, et al. A new UK 2006 National Kidney Allocation scheme for deceased heart-beating donor kidneys. Transplantation 2010 Feb 27;89(4):387-394. [doi: 10.1097/TP.0b013e3181c9029d] [Medline: 20177339]
55. Chopra B, Sureshkumar KK. Changing organ allocation policy for kidney transplantation in the United States. WJT 2015;5(2):38. [doi: 10.5500/wjt.v5.i2.38]
56. Faire B, Dittmer I. Improving equity of access to deceased donor kidneys in New Zealand. Prog Transplant 2008 Mar;18(1):10-12. [doi: 10.1177/152692480801800103] [Medline: 18429575]
57. Clinical guidelines for organ transplantation from deceased donors: organ and tissue authority, version 111. The Transplantation Society of Australia and New Zealand. 2023. URL: https://tsanz.com.au/storage/documents/TSANZ_Clinical_Guidelines_Version-111_13062023Final-Version.pdf [accessed 2025-05-25]
58. Mayer G, Persijn GG. Eurotransplant kidney allocation system (ETKAS): rationale and implementation. Nephrol Dial Transplant 2006 Jan 1;21(1):2-3. [doi: 10.1093/ndt/gfi269]
59. Hernández Benabe S, Batsis I, Dipchand AI, Marks SD, McCulloch MI, Hsu EK. Allocation to pediatric recipients around the world: an IPTA global survey of current pediatric solid organ transplantation deceased donation allocation practices. Pediatr Transplant 2023 Feb;27 Suppl 1:e14317. [doi: 10.1111/petr.14317] [Medline: 36468320]

Abbreviations

CENTRAL: Cochrane Central Register of Controlled Trials
ESKD: end-stage kidney disease
HLA: human leukocyte antigen
KAS: kidney allocation scheme
KT: kidney transplant
MeSH: Medical Subject Headings
MOOSE: Meta-Analyses of Observational Studies in Epidemiology
NOS: Newcastle-Ottawa Scale
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses–Protocols

Edited by A Schwartz; submitted 22.Jan.2025; peer-reviewed by AN Ali; revised version received 14.Sep.2025; accepted 18.Sep.2025; published 12.Feb.2026.

Please cite as:
Ambarsari CG, Kim JJ, Maamor NH, Rosli IA, Muhamad NA
Donor Risk Factors Affecting Graft Survival in Pediatric Kidney Transplants: Protocol for a Systematic Review and Meta-Analysis
JMIR Res Protoc 2026;15:e71620
URL: https://www.researchprotocols.org/2026/1/e71620
doi:10.2196/71620

© Cahyani Gita Ambarsari, Jon Jin Kim, Nur Hasnah Maamor, Izzah Athirah Rosli, Nor Asiah Muhamad. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evidence for Digital Mental Health Assessment Tools in the Post–COVID-19 Era: Protocol for a Systematic Review on Diagnostic Accuracy Across Age Groups

Kathryn Babbitt[1], BA; Erin Lucy Funnell[1], BSc; Nayra Martin-Key[1], PhD; Eleanor Barker[2], BSc (Hons), MSc, MCLIP; Sabine Bahn[1], MD, PhD

[1]Cambridge Centre for Neuropsychiatric Research, Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge, United Kingdom

[2]Engineering Library, Department of Engineering, University of Cambridge, Cambridge, United Kingdom

Corresponding Author:
Sabine Bahn, MD, PhD
Cambridge Centre for Neuropsychiatric Research, Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge, United Kingdom

Abstract

Background: Digital assessment tools in health care are increasingly used to aid clinicians in diagnosing mental health conditions. Particularly since the quarantine and isolation guidelines of the COVID-19 pandemic moved much of health care online, there has been an accelerated adoption of digital assessment tools. The diagnostic accuracy of digital mental health assessment tools for a range of psychiatric conditions has yet to be fully explored, especially for their use in populations of older adults and children.

Objective: This systematic review aims to (1) summarize recent studies on digital self-report question-and-answer–based mental health assessment tools for use in all ages across a range of psychiatric conditions (eg, the type and number of questions, if available; reference tests; timing; and blinding procedures), (2) present their validity (ie, diagnostic accuracy), and (3) assess study quality and applicability.

Methods: The PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) guided the development of this protocol. The protocol has been registered with PROSPERO. The searches were guided by the PICO (population, intervention, comparator, and outcome) framework. A systematic search was conducted of the following databases of literature published since 2021: MEDLINE, Embase, Cochrane Library, ASSIA, Web of Science Core Collection, CINAHL, and PsycINFO. Searches of clinical trial databases and hand searching of reference lists will be completed. Two authors have independently screened titles and abstracts of identified papers and selected studies according to eligibility criteria, resolving inconsistencies through discussion. Full texts were screened following the same process. The authors extracted data using the Covidence data extraction tool (Veritas Health Innovation Ltd; eg, sensitivity and specificity). Two authors will use the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool to assess risk of bias for each full-text inclusion.

Results: Scoping for this review began in December 2024. Searches of databases were completed in January 2025. Full-text screening and identification of the relevant gray literature were completed by the end of August 2025, and the final review is expected to be completed by December 2025.

Conclusions: The review aims to present the validity and quality of the diagnostic accuracy of digital mental health assessment tools across different ages (including children and older adults), particularly following the COVID-19 pandemic due to the exponential increase in development and use of such tools. This review will provide evidence for the wider deployment of digital mental health assessment tools across a wide age range. There will also be a discussion about future research for digital tools and avenues for policy around digital mental health assessments.

Trial Registration: PROSPERO CRD420250654734; https://www.crd.york.ac.uk/PROSPERO/view/CRD420250654734

International Registered Report Identifier (IRRID): DERR1-10.2196/73492

(JMIR Res Protoc 2026;15:e73492) doi:10.2196/73492

KEYWORDS

systematic review; digital assessment tool; digital mental health; depression; generalized anxiety disorder; validity

XSL•FO
RenderX

Introduction

The implementation of digital technologies in health care has rapidly increased since the COVID-19 pandemic, as health care services were forced to operate in the unprecedented circumstances of quarantine. This posed great challenges as global health care systems had to adapt quickly to continue to serve patients, but there was also a unique opportunity for growth and improvements when systems were challenged to transition to online platforms [1,2]. Simultaneously, the surge in the global burden of mental health symptoms since the pandemic [3] resulted in an increased demand for evidence-based interventions that are accessible and can reduce the strain of mental health symptoms and conditions on the wider health care system. Currently, mental health care systems both in the United Kingdom and globally are struggling to address the clinical needs of patients, leaving help-seeking individuals on long waiting lists that span months and even years for certain diagnoses and presentations [4-6].

Digital mental health interventions pose a scalable solution to various system-level problems that may contribute to poor mental health care outcomes [7,8]. By their nature, digital tools are accessible and cost-effective, allowing anyone with access to a device and an internet connection to retrieve resources [9,10]. Individuals unable or less willing to seek in-person care now have the option to be seen online, potentially reducing stress associated with visits and the cost of travel [11]. Additionally, digital interventions can be easily deployed by health care professionals with adequate training [12], while also allowing for the results of the intervention to be accessed anywhere. One specific type of digital mental health intervention is assessments, which can range from computerized versions of traditional pen and paper assessments to sophisticated algorithms that incorporate artificial intelligence and machine learning to assess patients' symptoms and clinical needs. Digital assessment tools vary greatly in their processes and may be used outside or alongside clinical care. Their outcomes vary as well, from providing diagnostic aid to serving as symptom severity checkers and risk assessments.

Digital assessment tools are part of an ever-changing landscape of health technologies, and while the evidence for the accuracy of digital assessment tools for mental health care has become more robust [13], continued updated research about the safety and validity of digital assessment tools is needed due to the potential for rapid development and deployment, which may result in the availability of low-quality tools [14]. It is also important to investigate the use and accuracy of digital assessment tools in the wider population, especially for younger and older people.

This study is intended to summarize recently published literature on digital assessment tools for mental health, with a particular focus on the years following the initial outbreak of the COVID-19 pandemic, which shifted many aspects of health care online. A previous systematic review from 2021 focused on digital mental health tools provided a basis to establish the quality of these assessments [13]. However, as outlined earlier, there is an opportunity for a more comprehensive summary for ages outside the range of 18 to 65 years, which was the population of interest in the previous review. As digital interventions become more widespread and more people rely on digital assessment tools, including young people [15] and older adults [16], it is important to understand the evidence for the validity of such tools for a wider age range, especially given the specific needs of these populations. Additionally, as the previous review ended in 2021, there is a need to examine literature in the years following the first wave of the pandemic. Summarizing the accuracy of these assessments in diverse age populations is imperative due to the exponential increase in their development and deployment since 2021 [17].

Therefore, the following research question is left unanswered in this literature: What is the evidence on the validity of digital assessment tools for mental health in children and young people (ie, under the age of 18 years, working-age adults (ie, between the ages of 18 and 65 years), and older adults (ie, aged >65 years) since the increased digitization of health care following the COVID-19 pandemic? In this study, we aim to present validity through evidence of diagnostic accuracy using typically reported metrics (ie, area under the curve [AUC], specificity, sensitivity, positive predictive value [PPV], negative predictive value [NPV], and accuracy). These metrics were chosen to be in line with STARD (Standards for Reporting of Diagnostic Accuracy Studies) 2015 guidelines, which suggest all diagnostic accuracy studies should mention a metric of accuracy such as sensitivity, specificity, predictive value (ie, PPV and NPV), or accuracy (ie, AUC) [18].

This systematic review aims to (1) summarize recent studies on digital self-report question-and-answer–based mental health assessment tools for use in all ages across a range of psychiatric conditions (eg, the type and number of questions [if available], reference tests, timing, and blinding procedures), (2) present their validity (ie, diagnostic accuracy), and (3) assess study quality and applicability. By presenting key evidence related to the diagnostic accuracy of digital mental health assessment tools, this review will inform our understanding of the quality and safety of currently available tools, as well as potentially support the wider acceptance and uptake of such tools.

Methods

Overview

The authors decided that the research questions would best be answered using a systematic review approach due to its applications for questions around international evidence, current practices, and intention to inform future research [19]. The primary objectives of this systematic review are to provide a comprehensive summary of existing digital mental health assessment tools and their validity in studies that use a gold standard reference. First, we will categorize the types of digital assessment tools available, distinguishing, for example, between digitized versions of traditional pen-and-paper psychiatric questionnaires and more advanced digital platforms such as those using machine learning approaches. Next, we will present the validity of these tools for each mental health condition under consideration. Finally, we will appraise the risk of bias and overall applicability of all included studies.

XSL•FO
RenderX

We have used the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) [20] guidelines for the development of this protocol (Checklist 1). The protocol for this systematic review was registered with PROSPERO (CRD420250654734). The searches were guided by the PICO (population, intervention, comparator, and outcome) framework. For the full report, the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [21] will be used as a guide (Checklist 2). All documentation will be submitted on our adherence to all the above guidelines in the final report, and any changes to the protocol will be noted.

Inclusion Criteria

Overview

To summarize the current evidence for the validity of digital mental health assessment tools, the inclusion and exclusion criteria were developed according to the PICO framework (Textbox 1). There is a rapidly growing number of digital assessment tools and substantial improvements in these technologies [22]. As digital tools in health care have been widely implemented, particularly since the pandemic [17], this review focuses on studies published since October 2021 to capture the current evidence base of the accuracy of these assessments across different age groups and to contribute to existing evidence for digital mental health assessment tools. October was chosen as the starting month to be included as it aligns with the end cutoff in a previous similar systematic review [13], therefore ensuring that the review builds on existing evidence and considers all published studies since then. Finally, studies were included regardless of study design.

Textbox 1. PICO (population, intervention, comparator, and outcome) framework (see Multimedia Appendix 1 for full search strategies).

Population

- Individuals assessed for symptoms of major depressive disorder, bipolar disorder, generalized anxiety disorder, panic disorder, social anxiety disorder, attention-deficit/hyperactivity disorder, autism spectrum disorders, insomnia, anorexia nervosa, bulimia nervosa, obsessive-compulsive disorder, psychosis, alcohol use disorder, substance use disorder, posttraumatic stress disorder, acute stress disorder, adjustment disorder, borderline personality disorder, emotionally unstable personality disorder, self-harm, and suicidality

Intervention

- Patient-completed digital question-and-answer–based mental health tools for assessment or similar assessments completed by a caregiver in the case of children or older adults

Comparator

- Reference standard including assessment by a clinician (eg, clinician interview), clinical diagnosis, self-reported diagnosis (including parent-reported diagnosis for children or young people, polysomnography for insomnia, or standardized structured or semistructured interview based on the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision* [*DSM-5-TR*] and *International Classification of Diseases, 11th Revision* [*ICD-11*] criteria [or *DSM*-5 and *ICD-10* for older publications])

Outcomes

- Validity (eg, accuracy, area under the curve, sensitivity, specificity, positive predictive value, and negative predictive value)

Population

All study locations, mental health condition severity, and populations were included regardless of gender or ethnicity. While the previously referenced systematic review included working-age adults (aged 18-65 years) [13], the population of interest for this review has been expanded to include individuals of any age. The inclusion of a wider age range provides more evidence for the validity of digital mental health assessment tools in both children and older adults.

Intervention

Research that includes a digital question-and-answer–based mental health assessment tool intended for the assessment of any of the mental health conditions of interest specified in the population was included. The intervention required for inclusion in this review was defined as digital question-and-answer–based mental health tools for assessments that were completed by a patient or user. Because this review includes children and older adults, studies with tools aimed to investigate this population that require additional input from a parent or caregiver were included when applicable.

Comparator

To be included, studies must have had a “gold standard” comparator (ie, a clinician assessment, a structured clinical interview, self- or parent-reported diagnosis, or a clinical diagnosis). We also considered polysomnography to be a “gold standard” comparator for insomnia. Identification of this comparator was assessed via screening of the title, abstract, or full text, rather than defined through searches.

Outcome

The primary outcome measures of the study are metrics of validity and accuracy for question-and-answer–based digital assessment tools. This includes metrics such as AUC, specificity, sensitivity, PPV, and NPV.

XSL•FO
RenderX

Exclusion Criteria

Overview

Exclusion criteria were developed to ensure the inclusion of the most relevant studies. To prevent including papers multiple times in the analysis, previous systematic reviews were excluded, although their reference lists were hand-searched for relevant papers for inclusion. While excluded, systematic reviews of similar topics were saved to compare the results of this review with previous results.

Other exclusion criteria were literature published before October 2021, papers with fewer than 500 words, ongoing clinical trials without results, papers published in non-English languages, editorials, opinion pieces, newspaper articles, and other various forms of popular media. Non–peer-reviewed literature (eg, dissertations and theses) was also excluded.

Population

Studies that specifically focus on military personnel and veterans, athletes, incarcerated or imprisoned individuals, detained individuals, police officers or other emergency personnel, or pregnant people were excluded. The only exceptions were studies that provide evidence for an assessment tool for post-traumatic stress disorder, where the experiences of these populations (eg, war, detention, and traumatic birth) may be relevant due to the diagnostic criteria. Studies that evaluate assessment tools designed for use in currently psychotic individuals, for assessment within the context of mental health crises or emergency mental health presentations within emergency department settings (eg, accident and emergency and emergency room), or use in groups solely due to a diagnosis of physical illnesses (eg, patients with cancer) were also excluded. Studies that consist of nonhuman studies were excluded.

Intervention

Studies with interventions that are designed to predict future mental illnesses and high-risk behaviors (eg, suicidality, suicide attempts, suicidal acts, and self-harm) were excluded. These tools were considered out of scope for this study because they fall outside the clinical use cases investigated in this review and have been identified as possibly high-risk for use in health care settings [23]. We aimed to look at tools that assess current mental health conditions rather than tools that predict the likelihood of future mental illnesses and high-risk behaviors. Interventions that are nondigital (ie, pen and paper) question-and-answer–based tools were excluded. Studies that evaluate digital tools that use other assessment methods outside of question-and-answer–based methods such as blood tests, imaging, genome analyses, wearables, speech biomarkers, and other digital biomarkers were excluded.

Comparator

Literature that does not address the outlined research question or include the specified reference standards was also excluded.

Outcomes

Studies that do not include one of the primary outcome measures of validity and accuracy for question-and-answer–based digital assessment tools were also excluded.

Search Strategies

A medical librarian completed formal searches on January 14, 2025, of the following databases: MEDLINE (via Ovid), Embase (via Ovid), Cochrane Library, ASSIA (via ProQuest), Web of Science (Core Collection), CINAHL (via EBSCOhost), and PsycINFO (via EBSCOhost). Searches of the ClinicalTrials.gov database [24] were also completed by the medical librarian on January 28, 2025. Authors KB and EF completed searches of the World Health Organization clinical trial database [25] on February 10, 2025. To build on the existing literature [13] and to review digital tools for mental health assessment after the COVID-19 pandemic, searches were conducted for literature published in and after 2021. As this systematic review aims to provide a comprehensive view of the current state of digital assessment tools, the authors conducted additional gray literature searches. Gray literature searches through Google were completed on February 10, 2025, by authors KB and EF in collaboration with the medical librarian.

While previous systematic reviews without meta-analyses were excluded from this analysis, hand searches were performed of the reference lists of relevant systematic reviews captured by database searches to identify studies that were then screened against the eligibility criteria.

The included search terms were adapted from the previously published systematic review [13], which demonstrated an effective search strategy for identifying relevant studies. The full search strategy can be found here (Multimedia Appendix 1).

Screening and Study Selection

Following the formal searches done by the medical librarian and gray literature searches, all of the identified literature was stored in Covidence (Veritas Health Innovation Ltd), a systematic review software. The literature was screened to remove duplicates before it was uploaded to Covidence.

Two independent reviewers (KB and EF) worked in Covidence to screen titles and abstracts against predetermined eligibility and exclusion criteria. Papers were flagged by the authors to include or exclude or to label if they were unsure. Any disagreements about whether to include or exclude papers were resolved by discussion. Following the initial screening, both reviewers screened the full texts of all included papers to determine their final eligibility based on the predetermined inclusion and exclusion criteria. Any disagreements were discussed, and when unable to come to a conclusion, a third reviewer (SB) was consulted.

Reasons for exclusion at the full-text screening stage were recorded in the Covidence software. The authors documented the number of papers included and excluded in every stage of review in the PRISMA 2020 flow diagram.

XSL•FO
RenderX

Results from gray literature searches in Google and the World Health Organization clinical trials database were screened against the same inclusion and exclusion criteria in a spreadsheet (Multimedia Appendix 2). Due to the comprehensive nature of the searches and the feasibility of screening such a large number of results, the decision was made not to conduct hand searching of references when systematic reviews were pulled from the gray literature searches.

Data Extraction

Data are currently being extracted via the standardized Covidence data extraction form. Two reviewers (KB and ELF) will complete each form for every included paper. The following data are being extracted: general publication information (author, date, title, and location of study), study design, sample characteristics (size, mean age or age range, gender, and ethnicity breakdowns), mental health conditions addressed by the study, assessment type (including the index test and reference standard used in the study), and the relevant outcome measures addressed by the study (eg, PPV, NPV, AUC, sensitivity, specificity, and accuracy). Where reported, the authors will extract data on cutoff scores for the index tests and the performance at each cutoff. In the case that multiple cutoffs are presented, if there are no suggestions of the optimal cutoffs or suggested cutoffs based on previous literature in the paper, then no associated data will be extracted.

The reviewers will continue to compare data extraction forms for each study and resolve any discrepancies. Anything that remains unresolved will be decided by a third reviewer (SB). Following this, all extracted data will then be transferred to a shared online spreadsheet.

Quality Appraisal

To evaluate the quality of the selected studies, the Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) is being used [26]. Two independent reviewers (KB and EF) are completing this assessment for all included studies using a shared online spreadsheet while blinded to the other reviewer's responses. The QUADAS-2 is intended to evaluate possible biases and applicability of the study in relation to the research question for the review. The assessment consists of 4 domains, each of which addresses different aspects of the study. The domains are patient selection, index tests, reference standard, and flow and timing. Each domain consists of a subdomain about risk of bias, and three have a subdomain about applicability. The subdomains consist of questions that the rater evaluates as either "low," "unclear," or "high." The subdomains are then evaluated together for the larger domain to reach a final decision about the risk of bias and concern about applicability. This method will be completed for all included studies. The results of the 2 reviewers will be compared. There will be a discussion about reaching a consensus when the reviewers have reached different conclusions, and a third reviewer will be consulted as needed (SB). A table will be created with the results.

Data Synthesis

Due to the anticipated heterogeneity of studies, a descriptive approach will be used for results synthesis. A meta-analysis will not be pursued because of the potential range of findings, conditions included, and extended inclusion of all ages, and the potential for a large number of different screening tools to be pulled. Tables will also be used to present findings of extracted data (ie, study name, authors, year of publication, demographics, index test, reference standard, and outcome data of the index test's performance). Results will be summarized based on the mental health condition and population (ie, children and young people, working-age adults, and older adults), and the current guidelines on the validity of digital tools will be described.

Results

Scoping of this review began in December 2024. All database searches were completed on February 10, 2025. Manual searching of reference lists was completed on August 25, 2025, and the final review with analysis is expected to be completed by December 2025. The final results will be submitted to a peer-reviewed journal for dissemination.

Gray literature searches from Google included 55 URLs pulled from using an advanced search string (Multimedia Appendix 2).

Currently, at the data extraction stage of the review, there are 40 studies that investigate 50 tools and screen for 17 conditions. The number of studies included at each stage is recorded in a PRISMA flowchart (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram for study selection, which included searches of databases, registers, and other sources.

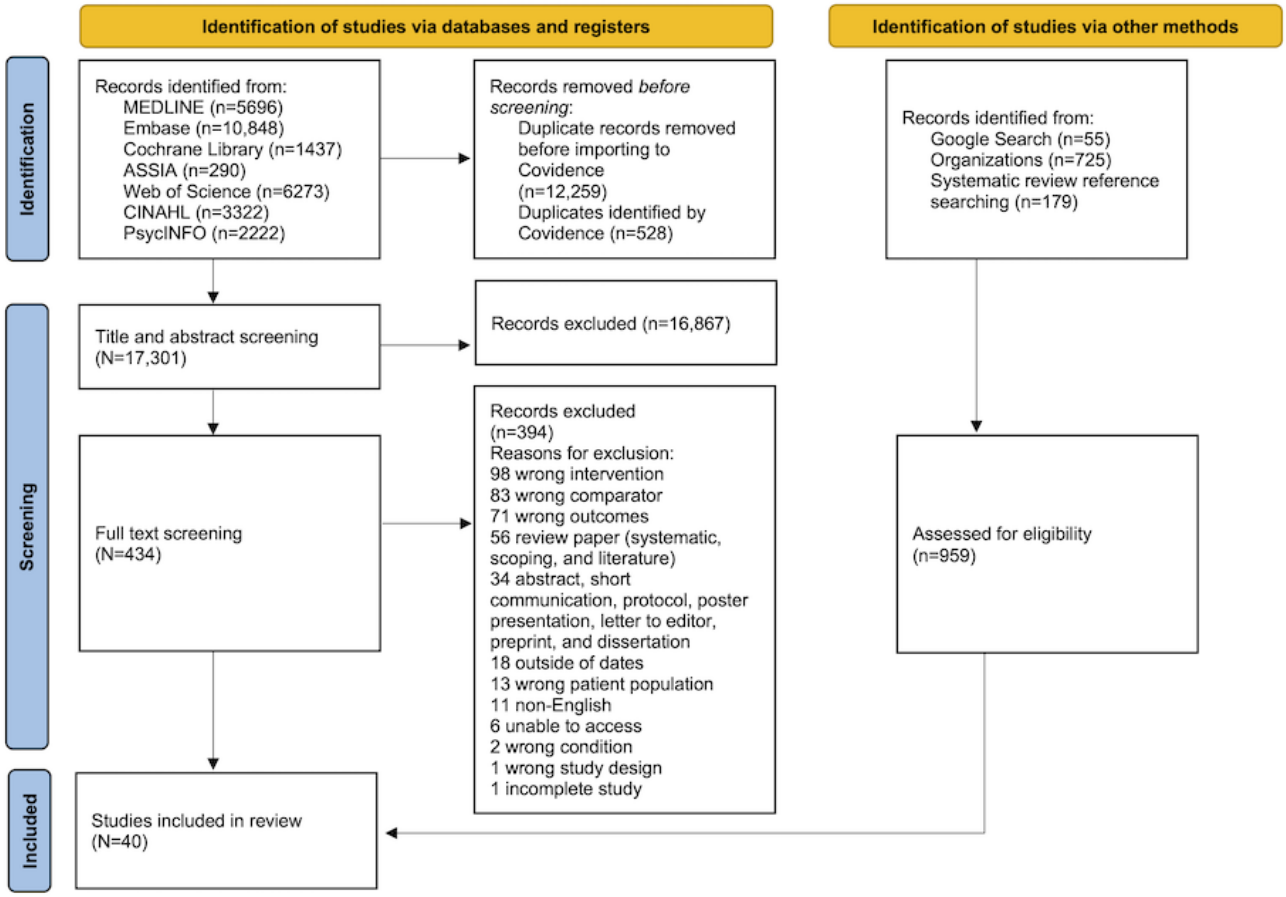

Discussion

Principal Findings

We hypothesize that, in line with previous systematic reviews [13], the evidence for the accuracy of digital assessment tools will be mixed. We anticipate that a large number of studies will fit the inclusion criteria due to the increased need for and implementation of digital interventions in the years after the COVID-19 pandemic. We also anticipate that proportionally, there will be fewer studies validating the digitalization of paper-and-pen questionnaires and a high number of studies that use machine learning techniques due to the rise in machine learning methods over the past 5 years [27].

Comparison With Prior Work

Previous similar reviews on digital assessment tools have focused on adult populations only [13] or the landscape of tools for specific conditions [28,29]. This study aims to complement and further these previous reviews by searching for an expanded age range, including both older adults and children, as well as by including multiple conditions within one review.

Strengths and Limitations

The main strengths of this study are the range of findings possible due to the breadth of conditions included and the expanded age range, which will allow for novel insights on the use of digital mental health assessment tools beyond just focusing on the use in the adult population. However, the broadened scope of this study limits the possible conclusions drawn due to likely heterogeneity in the pulled literature. Additionally, it is important to note that while this study addresses a range of conditions, it is not exhaustive, and there are several psychiatric conditions that are not examined. Finally, this review focuses only on question-and-answer–based assessments, which do not cover the recent technological advances in other forms of assessment.

Future Directions

Future work should consider summarizing the landscape of digital mental health assessment tools for conditions not examined in this study that have seen a significant increase in the years following the COVID-19 pandemic and pandemic-related lockdowns, such as health anxiety, specific phobias, and specific addictions, especially those related to technology. We also did not examine other conditions that were present with physical and mental health symptoms (eg, irritable bowel syndrome and chronic fatigue).

Finally, future work should consider a broader search strategy to include other digital tools for mental health assessments, which are not question-and-answer–based such as digital biomarkers and wearable technologies that have also seen a rise in uptake in recent years [30]. Given their currently limited use in clinical settings, they were not included in this study, but this analysis could contribute to a more comprehensive understanding of the current landscape for digital assessments in the field beyond what will be presented in the review.

XSL•FO
RenderX

Conclusions

This review aims to expand on previous literature reviews regarding digital mental health assessment tools, particularly in light of the increase in deployment of digital technologies in health care following the COVID-19 pandemic. By doing this, the current validity and quality of digital mental health assessments in adult populations will be presented while also expanding to include children and young people as well as older people to address this gap in the literature. This review will potentially provide evidence for the safe and effective deployment of digital assessment tools in a wider patient population and explore the potential opportunities for digital assessment tools to support safe and cost-effective health care delivery. There will also be a discussion about the implications and potential avenues for policy changes and future research directions.

Acknowledgments

The authors would like to thank Veronica Phillips at the University of Cambridge Medical Library for her assistance in developing search terms and completing database searches. The authors declare that no generative artificial intelligence was used at any stage for this manuscript.

Funding

This review has been funded by Stanley Medical Research Institute (grant 07R-1888).

Data Availability

All data generated or analyzed as a part of this study will be included in the full published version of the article or its supplementary materials.

Authors' Contributions

Conceptualization: KB (lead), ELF (equal), NM-K (equal), SB (equal)
Data curation: KB (lead), ELF (equal), EB (supporting)
Formal analysis: KB (lead), ELF (equal)
Funding acquisition: SB
Investigation: EB (lead), KB (supporting), ELF (supporting)
Methodology: EB (lead), KB (supporting), ELF (supporting)
Project administration: KB (lead), ELF (equal)
Resources: EB
Supervision: SB
Validation: KB (lead), ELF (equal), NM-K (supporting), SB (supporting)
Visualization: KB
Writing – original draft: KB (lead), ELF (equal), NM-K (supporting), EB (supporting), SB (supporting)
Writing – review & editing: KB (lead), ELF (equal), NM-K (supporting), EB (supporting), SB (supporting)

Conflicts of Interest

SB is a director of Psynova Neurotech Ltd and Psyomics Ltd. KB and ELF are consultants with Psyomics Ltd. KB, ELF, and SB have financial interests in Psyomics Ltd. The other authors have no conflicts to declare.

Multimedia Appendix 1
Search strategies.
[DOCX File, 28 KB - resprot_v15i1e73492_app1.docx]

Multimedia Appendix 2
World Health Organization (WHO) clinical trials and advanced Google searches.
[XLSX File, 1938 KB - resprot_v15i1e73492_app2.xlsx]

Checklist 1
PRISMA-P checklist.
[DOCX File, 35 KB - resprot_v15i1e73492_app3.docx]

Checklist 2
PRISMA checklist.
[DOCX File, 274 KB - resprot_v15i1e73492_app4.docx]

XSL•FO
RenderX

References

1. Budd J, Miller BS, Manning EM, et al. Digital technologies in the public-health response to COVID-19. Nat Med 2020 Aug;26(8):1183-1192. [doi: 10.1038/s41591-020-1011-4] [Medline: 32770165]
2. Getachew E, Adebeta T, Muzazu SGY, et al. Digital health in the era of COVID-19: reshaping the next generation of healthcare. Front Public Health 2023;11:942703. [doi: 10.3389/fpubh.2023.942703] [Medline: 36875401]
3. COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide. World Health Organization. URL: https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide [accessed 2022-03-02]
4. Punton G, Dodd AL, McNeill A. "You're on the waiting list": an interpretive phenomenological analysis of young adults' experiences of waiting lists within mental health services in the UK. PLoS ONE 2022;17(3):e0265542. [doi: 10.1371/journal.pone.0265542] [Medline: 35303040]
5. Thomas KA, Schroder AM, Rickwood DJ. A systematic review of current approaches to managing demand and waitlists for mental health services. Mental Health Rev J 2021 Feb 17;26(1):1-17. [doi: 10.1108/MHRJ-05-2020-0025]
6. Edbrooke-Childs J, Deighton J. Problem severity and waiting times for young people accessing mental health services. BJPsych Open 2020 Oct 12;6(6):e118. [doi: 10.1192/bjo.2020.103] [Medline: 33040766]
7. Spadaro B, Martin-Key NA, Funnell E, Benáček J, Bahn S. Opportunities for the implementation of a digital mental health assessment tool in the United Kingdom: exploratory survey study. JMIR Form Res 2023 Aug 7;7:e43271. [doi: 10.2196/43271] [Medline: 37549003]
8. Buck B, Kadakia A, Larsen A, Tauscher J, Guler J, Ben-Zeev D. Digital interventions for people waitlisted for mental health services: a needs assessment and preference survey. Pract Innov (Wash D C) 2025 Mar;10(1):32-42. [doi: 10.1037/pri0000250] [Medline: 40256356]
9. Gentili A, Failla G, Melnyk A, et al. The cost-effectiveness of digital health interventions: a systematic review of the literature. Front Public Health 2022;10:787135. [doi: 10.3389/fpubh.2022.787135] [Medline: 36033812]
10. Richards D, Enrique A, Eilert N, et al. A pragmatic randomized waitlist-controlled effectiveness and cost-effectiveness trial of digital interventions for depression and anxiety. NPJ Digit Med 2020;3(1):85. [doi: 10.1038/s41746-020-0293-8] [Medline: 32566763]
11. Sweeney GM, Donovan CL, March S, Forbes Y. Logging into therapy: adolescent perceptions of online therapies for mental health problems. Internet Interv 2019 Mar;15:93-99. [doi: 10.1016/j.invent.2016.12.001] [Medline: 30792959]
12. Wosny M, Strasser LM, Hastings J. Experience of health care professionals using digital tools in the hospital: qualitative systematic review. JMIR Hum Factors 2023 Oct 17;10:e50357. [doi: 10.2196/50357] [Medline: 37847535]
13. Martin-Key NA, Spadaro B, Funnell E, et al. The current state and validity of digital assessment tools for psychiatry: systematic review. JMIR Ment Health 2022 Mar 30;9(3):e32824. [doi: 10.2196/32824] [Medline: 35353053]
14. Neary M, Schueller SM. State of the field of mental health apps. Cogn Behav Pract 2018 Nov;25(4):531-537. [doi: 10.1016/j.cbpra.2018.01.002] [Medline: 33100810]
15. Boydell KM, Hodgins M, Pignatiello A, Teshima J, Edwards H, Willis D. Using technology to deliver mental health services to children and youth: a scoping review. J Can Acad Child Adolesc Psychiatry 2014 May;23(2):87-99. [Medline: 24872824]
16. Harerimana B, Forchuk C, O'Regan T. The use of technology for mental healthcare delivery among older adults with depressive symptoms: a systematic literature review. Int J Ment Health Nurs 2019 Jun;28(3):657-670. [doi: 10.1111/inm.12571] [Medline: 30666762]
17. Li J. Digital technologies for mental health improvements in the COVID-19 pandemic: a scoping review. BMC Public Health 2023 Mar 1;23(1):413. [doi: 10.1186/s12889-023-15302-w] [Medline: 36859184]
18. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open 2016 Nov 14;6(11):e012799. [doi: 10.1136/bmjopen-2016-012799] [Medline: 28137831]
19. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
20. Moher D, Shamseer L, Clarke M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
21. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100. [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
22. Chang JE, Lai AY, Gupta A, Nguyen AM, Berry CA, Shelley DR. Rapid transition to telehealth and the digital divide: implications for primary care access and equity in a post-COVID era. Milbank Q 2021 Jun;99(2):340-368. [doi: 10.1111/1468-0009.12509] [Medline: 34075622]
23. Alderman J, Riley R, Parekh D, Summers C, Liu X, Denniston A. Hidden risks of predictive models in healthcare. BMJ Evid Based Med 2025 Sep 17:bmjebm-bm2025. [doi: 10.1136/bmjebm-2025-113730] [Medline: 40962433]
24. ClinicalTrials.gov. URL: https://clinicaltrials.gov [accessed 2025-01-28]

XSL•FO
RenderX

25. International Clinical Trials Registry Platform. World Health Organization. URL: https://www.who.int/clinical-trials-registry-platform [accessed 2025-02-10]
26. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011 Oct 18;155(8):529-536. [doi: 10.7326/0003-4819-155-8-201110180-00009] [Medline: 22007046]
27. Madububambachu U, Ukpebor A, Ihezue U. Machine learning techniques to predict mental health diagnoses: a systematic literature review. Clin Pract Epidemiol Ment Health 2024;20:e17450179315688. [doi: 10.2174/0117450179315688240607052117] [Medline: 39355197]
28. De Angel V, Lewis S, White K, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. NPJ Digit Med 2022 Jan 11;5(1):3. [doi: 10.1038/s41746-021-00548-8] [Medline: 35017634]
29. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry 2016;39:24-31. [doi: 10.1016/j.genhosppsych.2015.11.005] [Medline: 26719105]
30. Şahin Tokatlıoğlu T, Oflaz F, Semiz B. Wearable technologies and psychiatry: strengths, weaknesses, opportunities, and threats analysis. Cyprus J Med Sci 2025. [doi: 10.4274/cjms.2025.2024-60]

Abbreviations

AUC: area under the curve
NPV: negative predictive value
PPV: positive predictive value
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies 2
STARD: Standards for Reporting of Diagnostic Accuracy Studies

Edited by A Schwartz; submitted 05.Mar.2025; peer-reviewed by M Gottschalk; revised version received 14.Nov.2025; accepted 29.Nov.2025; published 05.Jan.2026.

Please cite as:
Babbitt K, Funnell EL, Martin-Key N, Barker E, Bahn S
Evidence for Digital Mental Health Assessment Tools in the Post–COVID-19 Era: Protocol for a Systematic Review on Diagnostic Accuracy Across Age Groups
JMIR Res Protoc 2026;15:e73492
URL: https://www.researchprotocols.org/2026/1/e73492
doi:10.2196/73492

© Kathryn Babbitt, Erin Lucy Funnell, Nayra Martin-Key, Eleanor Barker, Sabine Bahn. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Self-Guided Mental Health Interventions to Prevent Workplace Burnout and Enhance Psychological Wellness: Protocol for a Systematic Review

Ehsan Etezad[1], MSc; John Fiset[2], PhD; Raghid Al Hajj[3], PhD

[1]Applied Organizational Psychology, Applied Science Department, Saint Mary's University, 923 Robie Street, Halifax, Canada

[2]Management Department, Sobey School of Business, Saint Mary's University, Halifax, Canada

[3]Business Administration Department, Gulf University for Science & Technology, Hawally, Kuwait

Corresponding Author:
Ehsan Etezad, MSc
Applied Organizational Psychology, Applied Science Department, Saint Mary's University, 923 Robie Street, Halifax, Canada

Abstract

Background: Employee burnout has reached critical levels, with nearly 40% of workers reporting symptoms driven by excessive workloads, inadequate managerial support, and toxic organizational cultures. Research has consistently indicated that workplace stress significantly impacts employees' physical health. Consequently, a variety of self-guided mental health programs have been developed to address these challenges. Despite numerous evidence-based self-guided mental health interventions, there remains a lack of understanding and clarity regarding which specific content modules, intervention features, active components, and theoretical frameworks most effectively drive meaningful improvements in workplace mental health.

Objective: This systematic review aims to synthesize evidence on self-guided digital mental health interventions in workplace settings. Specifically, it seeks to identify which content, design features, activities, and assignments are most effective for preventing burnout and enhancing psychological well-being. This review will also examine underlying theoretical mechanisms and assess the methodological rigor of included studies to provide actionable recommendations for intervention developers and organizational stakeholders.

Methods: A systematic search will be executed across PsycINFO, PubMed, Web of Science, and Cochrane Library for relevant studies published since 2000, using comprehensive search strings targeting working adults, digital delivery modes, outcomes related to burnout prevention, stress reduction, well-being, and controlled experimental designs. Titles, abstracts, and keywords will be screened, with additional records identified through manual searches of reference lists. Following the removal of duplicates, a 2-step screening process will be applied to studies based on the defined inclusion and exclusion criteria. Data from eligible studies will be extracted into a standardized Excel template covering authors, sample characteristics, study design, intervention content and theoretical framework, outcome measures, intervention effectiveness, implementation fidelity, and risk-of-bias assessment.

Results: Preliminary searches were conducted in early 2025. The review is anticipated to be completed by May 2026.

Conclusions: This review will identify theoretical mechanisms and core components driving the effectiveness of self-guided digital interventions for workplace burnout, stress, and psychological well-being. By identifying theoretical background, rationale, content, activities, characteristics, and implementation factors behind evidence-based digital interventions, the findings will guide the development of scalable and accessible programs that enhance employee wellness, boost productivity, and inform future organizational mental health and workplace wellness initiatives.

Trial Registration: PROSPERO International Prospective Register of Systematic Reviews CRD420251035459; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251035459

International Registered Report Identifier (IRRID): DERR1-10.2196/80417

(JMIR Res Protoc 2026;15:e80417) doi:10.2196/80417

KEYWORDS

burnout; job stress; mental health; psychological well-being; self-guided intervention; workplace wellness

XSL•FO
RenderX

Introduction

Overview

In North America, approximately 40% of employees experience burnout, driven primarily by heavy workloads, insufficient managerial support, and toxic work cultures [1,2]. This increased level of burnout has tangible consequences: 7.5% of the employed population has taken time off due to stress or mental health concerns, leading to an average loss of 2.4 workdays per employee [3]. Moreover, about 77% of employees reported that work stress had a negative impact on their physical health [4]. Therefore, among a wide range of mental health concerns, burnout has emerged as a critical driver of productivity loss. Its prevalence signals a deeper systemic issue, with costs extending beyond absenteeism to increased health care expenses, diminished employee engagement, and higher turnover rates [5-11]. In a recent meta-analysis study reviewing data from 5022 participants, nearly half of the participants reported elevated levels of burnout [12]. A survey conducted by the Canadian Medical Association on 4000 physicians and medical learners in 2021 showed that 53% of the participants indicated elevated levels of burnout during the pandemic. This number was 30% in 2017, pre-pandemic. Additionally, nearly half of the participants expressed intentions to reduce their clinical workload within the next 2 years [13]. Given the wide-ranging consequences, burnout should not be viewed solely as an individual health issue but as a strategic organizational challenge requiring timely and practical action to safeguard both employee well-being and overall productivity.

Despite growing awareness, a substantial gap remains between the demand for and access to mental health services. Even prior to the COVID-19 pandemic, mental health challenges were identified as a leading cause of loss of productivity and disability [14]. The World Health Organization has similarly announced that depression symptoms are the primary contributors to loss of productivity globally [15]. In the wake of the pandemic, the urgency of addressing mental health in the workplace was even more pronounced. Employers increasingly recognized the need to support psychological well-being by promoting mental health and providing resources for employees experiencing distress [16]. The prevalence of mental health issues has risen sharply since the onset of the COVID-19 pandemic, with anxiety rates increasing 4-fold (from 5% to 20%) and depression rates more than doubling (from 4% to 10%) [17]. Insurance data further reflect this trend, showing a 25% increase in claims for mental health support in 2020 [18]. Moreover, a recent survey revealed a 6% rise in the number and 12% in the duration of short-term mental health−related disability claims in 2021 [19]. Overall, mental health issues, including burnout, now account for approximately 30%-40% of short-term disability claims [20].

Burnout is defined as a multidimensional occupational phenomenon characterized by chronic work-related stress that has not been successfully managed [21]. It primarily encompasses 3 core dimensions: emotional exhaustion, cynicism, and reduced personal accomplishment [22]. Emotional exhaustion refers to feelings of being overextended and overstretched and depleted of one's emotional resources, often resulting from excessive job demands and prolonged exposure to stressors [23]. Cynicism, or depersonalization, involves a detached or negative attitude toward one's work and the people served, as it signifies an individual's attempt to cope with overwhelming emotional demands by distancing themselves from their professional role [24]. The third dimension, reduced personal accomplishment, captures a decline in one's sense of competence and achievement at work, often leading to a sense of inefficacy and a diminished capacity to perform job responsibilities effectively [25-28].

In addition to examining the alarming prevalence of workplace stress and burnout, it is equally important to evaluate the current solutions being used to address these challenges. Burnout interventions are generally classified into 2 main categories: organizational-level and employee-level support. Organizational-level support refers to the strategies and policies that companies implement that are aimed at fostering a work environment where employees can thrive and reduce exposure to excessive stress and burnout [29,30]. At its core, these initiatives are typically embedded in the company's culture and leadership practices [31]. Employee-level support primarily consists of extended health benefits, including access to counseling and psychological therapies that are designed to improve the psychological health and well-being of individual employees. These resources are generally made available for employees and working adults to access at their own discretion. Numerous studies have demonstrated the effectiveness of these programs [32-34].

Despite the effectiveness and availability of these traditional evidence-based burnout interventions [35-38], several key obstacles impede employees' access and engagement. First, financial constraints, both at the organizational level (budget caps on program funding) and the individual level (copays or out-of-pocket costs), can prevent participation or engagement with these programs [39]. Second, time demands present a significant hurdle: structured programs often require multisession commitments during standard business hours, creating scheduling conflicts or necessitating employees and working adults to sacrifice personal time for attendance [40]. Third, systemic wait lists for in-person services can delay support by weeks or months, undercutting the immediacy of the intervention when stress is most acute [41]. Finally, stigma surrounding mental health in the workplace, whether fear of being perceived as weak by colleagues or concerns about confidentiality, further depresses uptake, even when services are fully funded and readily available [39]. Collectively, these barriers contribute to persistently low participation rates and call into question the scalability of traditional support models.

To overcome these limitations and improve the scalability of mental health support programs, there has been growing interest in digital self-guided mental health programs [42]. Self-guided programs offer a flexible, accessible, and cost-effective complement to traditional burnout interventions. These programs empower working adults to engage with content at their own pace or convenience through an online platform that incorporates evidence-based techniques such as mindfulness, cognitive restructuring, and stress management [38]. By eliminating barriers like high cost, long wait times, scheduling conflicts,

XSL•FO
RenderX

and the stigma sometimes associated with seeking professional help, self-guided approaches enable employees and working adults to take proactive control of their psychological wellness. Moreover, the self-paced nature of these interventions allows users to revisit material as needed and tailor practices to fit their unique circumstances.

Although prior reviews have provided valuable insights into the effectiveness of digital interventions targeting various psychological outcomes [43-48], they have largely overlooked the theoretical foundations, rationale, and specific content or modules that contribute to these positive effects. The present systematic review differs by focusing on the underlying theories, rationales, and intervention components that characterize effective self-guided digital interventions in the workplace and their effects on burnout, stress, and psychological well-being. While a recent systematic review and meta-analysis examined the mechanisms of change in digital interventions for clinical depression [49], that work reviewed 11 digital interventions targeting a clinical population. In contrast, the current review aims to identify the theoretical bases and content elements of self-guided digital interventions designed to prevent burnout and enhance psychological well-being among generally healthy working adults.

This systematic review is primarily focused on self-guided digital interventions since other intervention options, such as hybrid or guided digital interventions, still present several practical limitations that restrict their scalability and accessibility in workplace settings. Moreover, guided formats typically require ongoing facilitator involvement, which reintroduces cost, scheduling, and staffing challenges similar to those of traditional in-person interventions [50]. In contrast, fully self-guided programs remove the dependency on human facilitation, allowing organizations to reach a larger employee populations [51,52]. Furthermore, self-guided interventions provide a higher degree of privacy and autonomy, which may be particularly valuable in work cultures where mental health stigma remains prevalent [53]. Focusing on self-guided interventions thus allows this review to address an emerging yet underexplored domain in occupational health programs.

This systematic review aims to offer a comprehensive, transparent, and replicable synthesis of literature on the content and theoretical explanations of these digital self-guided mental health interventions in the workplace. Specifically, the review will focus on identifying the most effective content, activities, and assignments within these digital interventions that contribute to preventing burnout and improving psychological wellness. To ensure a rigorous and methodologically sound process, the review will follow the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [54,55].

Review Aim

The aim of this systematic literature review is to identify the most effective content, design features, activities, and assignments included in brief digital self-guided mental health interventions for the workplace. Specifically, the review will assess which components are most effective in preventing burnout and improving psychological wellness, as well as their theoretical background and the rationale behind their effectiveness.

Review Objectives

This review has 2 objectives: (1) to catalog and describe the content and components of each self-guided digital intervention: one of the key goals of this systematic review is to identify and detail the content, design features, activities, and assignments used in brief self-guided mental health interventions within workplace settings, focusing on those aimed at preventing burnout and enhancing psychological wellness. (2) To examine the theoretical underpinnings and rationale of each self-guided digital intervention: this systematic review will investigate the theoretical frameworks informing these digital interventions and determine how the underlying theory or mechanisms are associated with greater effectiveness in achieving mental health outcomes.

Methods

Study Design

This systematic review aims to offer a comprehensive, transparent, and replicable synthesis of the literature on the content and theoretical explanations of brief digital mental health interventions in the workplace. This systematic review protocol has been developed according to the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols) checklist [56] (Checklist 1)(). This protocol has been registered on PROSPERO International Prospective Register of Systematic Reviews with the ID CRD420251035459.

Search Methods

A comprehensive review of the literature will be conducted using 4 major relevant academic databases, including PsycINFO, PubMed, Web of Science, and Cochrane Library. These databases were selected to ensure a broad coverage of psychological research (PsycINFO), biomedical and clinical studies (PubMed), multidisciplinary citation indexing (Web of Science), and high-quality evidence syntheses (Cochrane Library). The search was conducted to identify records that match our inclusion criteria published between January 2000 and the date of commencement. This timeline was chosen as this review focuses on rapidly evolving digital mental health interventions. Given the fast-paced development of new technologies, studies prior to 2000 are unlikely to remain relevant to the aim of this systematic review. A systematic search will be conducted across multiple databases using a combination of search terms related to the target population (working adults), the mode of intervention delivery (digital, web-based, or mobile interventions), the intended outcomes (burnout prevention, stress reduction, and mental well-being), and the study design (randomized controlled trials and other experimental studies). The full list of search strings used for each database is provided in Table 1. Search terms will be applied to titles, abstracts, and keywords. Additionally, reference lists of included studies and relevant systematic reviews will be manually screened to identify any additional potentially eligible studies.

XSL•FO
RenderX

Table . Complete list of terms used as a search method of the systematic review.

Cluster	Search string
Population	adult* OR worker* OR employee* OR "working population" OR workplace* OR occupational OR job OR staff OR personnel OR "work* place" OR occupation*
Intervention	"digital intervention*" OR "web-based intervention*" OR "online intervention*" OR "internet-based intervention*" OR "e-health" OR "ehealth" OR "m-health" OR "mhealth" OR "smartphone app*" OR "mobile app*" OR "computer-based" OR "email-based"
Outcome	burnout OR "burn-out" OR "occupational stress" OR "work-related stress" OR "emotional exhaustion" OR cynicism OR "reduced efficacy" OR "stress" OR "work* stress" OR "exhaustion" OR "fatigue" OR "mental health" OR "depression" OR "anxiety" OR "wellness" OR "well-being" OR "wellbeing"
Design	"randomized controlled trial*" OR RCT OR "controlled clinical trial*" OR "quasi-experiment*" OR "experimental stud*" OR "control group"

Screening and Selection Process

The search results from the selected databases will be imported into a reference management software (version 7; Zotero, developed by Digital Scholar) to facilitate duplicate detection and streamline the screening process. Papers will be managed using the web platform Rayyan [57], which is a web solution that expedites the initial screening of abstracts, followed by a 2-stage screening process. First, all duplicates will be removed. Second, the titles and abstracts will be reviewed for eligibility, and full texts will be retrieved for studies where eligibility remains unclear. Then, the full texts will be assessed based on the inclusion and exclusion criteria, considering factors such as intervention type, study population, and reported outcomes. Reasons for exclusion will be recorded at each stage of the selection process. As this study is part of a PhD student's thesis, the initial search and eligibility assessment of retrieved articles will be conducted by the PhD student (EE). Any articles with uncertain eligibility will be discussed with my coauthor (JF). In instances where discrepancies or disagreements arise, RAH will be consulted to provide additional input and facilitate consensus.

Eligibility Criteria

This systematic review will include studies that focus on digital, web-based health interventions aimed at addressing burnout or workplace wellness in working adults (aged 19 years and older) who are generally healthy and have no diagnosed clinical conditions, in order to evaluate preventive strategies and ensure that findings are broadly applicable to the wider workforce rather than being driven by treatment-seeking clinical samples. Eligible studies must examine workplace-related populations, including part-time and full-time employees. To ensure methodological rigor, only randomized controlled trials, experiments, or quasi-experiments with a control group (eg, no treatment, treatment as usual, waitlist control, or active control) will be included. Studies must assess intervention effectiveness using at least 1 pre- and postintervention measurement and report quantitative outcome data. The primary outcomes of this review include burnout, perceived stress, and psychological well-being, and the secondary outcomes include depression and anxiety symptoms. Interventions must be delivered through digital platforms such as websites, emails, computer programs, or smartphone apps, with active components included. Additionally, studies must describe the intervention's content, modules, or activities in sufficient detail. Only peer-reviewed journal articles or dissertations published after 2000 will be considered to ensure studies included have undergone a rigorous editorial and methodological vetting process, thereby maximizing the quality and reliability of our findings [58]. Finally, both the abstract and full text must be available in English to allow for comprehensive review and analysis. The full list of eligibility and exclusion criteria is shown in Table 2.

XSL•FO
RenderX

Table . Inclusion and exclusion criteria of the systematic review.

Element	Eligibility criteria	Exclusion criteria
Population	• Age: Adults (19 y old or over) • Working population (part-time, full-time) and related to workplace • Participants are generally healthy	• Individuals who are unemployed, retired, or on disability leave • Participants involved in young or senior care mental health contexts • Individuals with psychiatric or complex mental health conditions (eg, PTSD[a], schizophrenia, or comorbid substance misuse)
Intervention	• Digital or web-based health interventions, where active elements are delivered via website, email, computer program, or smartphone app • Intervention content or modules or activities are described	• Interventions using tracking devices or wearable technologies (eg, biofeedback) • Interventions incorporating virtual reality or augmented reality • Telehealth or videoconferencing interventions where technology is only a mode of communication (eg, Skype, e-counseling) • Text or SMS-based health interventions • Interventions where homework is digital, but the main intervention is delivered in-person
Comparison	• Studies must include a control group (eg, nontreatment, treatment-as-usual, waitlist control, or active control)	• Studies without a control group
Outcome	• · Burnout measured as an outcome (directly or indirectly) • Reported quantitative effectiveness outcomes • Included at least one pre- and 1 post-intervention assessment	• Studies focusing on risk factors, protective factors, motivation, or willingness to participate • Studies focused on participant satisfaction, sources of stress, or burnout prevention strategies • Studies assessing organizational- or system-level strategies or initiatives only • Protocol papers or cost-analysis studies
Study design	• Randomized controlled trials (RCTs), experiments, or quasi-experimental designs • Minimum of 5 participants	• Observational, qualitative, or correlational studies • Case studies with fewer than 5 participants
Publication type	• Published in peer-reviewed journals or dissertations • Published post-2000 • Abstract and full text available in English	• Gray literature, opinion pieces, protocols, and reviews

[a]PTSD: posttraumatic stress disorder.

Data Extraction Process

For each eligible study, we will extract key data into a structured Microsoft Excel template to ensure consistency and accuracy [59]. This will include (1) Authors and References: Names of the authors and publication year; (2) Sample: Sample size, target population (eg, employees, industry type), and setting (eg, online, workplace); (3) Methodology or Study Design: Research design used in the study (eg, randomized controlled trial, quasi-experimental), including details about the control group (active control, passive waitlist group, or treatment as usual); (4) Objective or Goal: The primary aim of the intervention, such as reducing burnout or improving psychological wellness; (5) Outcomes or Measures: Instruments used to assess effectiveness; (6) Type of Intervention: Focus and delivery method (eg, digital self-guided mental health program, web-based training, mobile application); (7) Theoretical Framework: The underlying model or framework guiding the intervention (eg, Cognitive Behavioral Therapy, Mindfulness-Based Stress Reduction); (8) Content: A brief description of the intervention's structure, components, and activities; (9) Effectiveness: A brief summary of the quantitative results indicating its effectiveness and effect sizes; and (10) Fidelity: A brief summary of whether the intervention was implemented as intended and adhered to the planned structure. This systematic extraction process will ensure a comprehensive and standardized analysis of all relevant aspects of the interventions.

Risk of Bias Assessment

The risk of bias (ROB) will be assessed using the revised Cochrane Risk-of-Bias (RoB 2; version 2) tool for randomized trials [60]. Each study will be evaluated across the 5 RoB 2

XSL•FO
RenderX

domains: (1) bias resulting from the randomization process, (2) bias resulting from deviations from intended interventions, (3) bias due to missing outcome data, (4) bias in outcome measurement, and (5) bias in the selection of reported results. The primary author will perform the initial ROB assessment. Included studies will be classified as having "low risk," "some risk," "high risk," or "unclear risk."

Data Analysis

An integrated narrative synthesis will be conducted to investigate and explore the connection between the findings within and across the included interventions as recommended by Popay et al [61]. This structured iterative approach consists of 6 key steps: (1) developing an overarching theoretical understanding, (2) conducting an initial analysis, (3) examining relationships between findings, (4) evaluating the strength and consistency of the evidence, and (5) drawing conclusions and (6) practical recommendations.

To facilitate this synthesis, various analytical techniques will be employed, including content analysis, tabulation, concept mapping, and critical reflection. The preliminary synthesis will involve summarizing study findings and categorizing them into meaningful themes through content analysis and tabulation. Relationships within and across studies will then be explored using concept mapping and visual representations to enhance conceptual and methodological integration. Finally, a critical reflection and quality appraisal process will be conducted. The GRADE (Grading of Recommendations Assessment, Development, and Evaluation) framework will be used to evaluate the quality of evidence within each study and its recommendations [62].

More specifically, objective 1 (to catalog and describe the content and components of interventions) will be primarily analyzed through content analysis and tabulation by including the characteristics of selected studies, such as intervention goals, delivery format, duration, activities, and behavioral assignments. These entries will then be grouped into broader categories that capture recurring design features and intervention components. Additionally, objective 2 (to examine the theoretical underpinnings and rationale of each intervention) will be analyzed by linking each intervention's theoretical framework or psychological mechanism (eg, cognitive-behavioral theory, mindfulness, self-determination theory) to its reported outcomes. Quantitative findings, such as effect sizes and direction of effects, will be compared across studies sharing similar theoretical foundations and levels of evidence quality (including ROB) to determine which frameworks and mechanisms are most strongly associated with significant improvements in burnout and psychological wellness outcomes. Finally, a critical reflection will be applied to interpret how contextual factors (such as setting, population, and intervention format) may influence the effectiveness of each theoretical approach.

The synthesis of the findings will lead to evidence-based recommendations on the content and theories behind digital interventions for psychological wellness in the workplace.

Results

In April 2025, the study protocol was submitted to PROSPERO. Shortly after, we conducted preliminary searches on the academic databases. As of October 2025, the title and abstract screening and full-text screening have been completed, and data analysis and reporting have commenced. This systematic review is anticipated to be completed by May 2026.

Discussion

Principal Findings

This systematic review is positioned to make a meaningful contribution to the field of digital mental health interventions in workplace contexts. This study will identify the "active ingredients" and underlying theoretical mechanisms that underpin the effectiveness of self-guided digital interventions that can be applied in organizational settings.

Although the results are pending, this proposed systematic review is expected to provide the most comprehensive analysis of how self-guided digital mental health interventions function for employees and working adults. We also anticipate identifying substantial growth in publications, particularly in the past decade, which signals the field's rapid maturation and growing recognition of digital interventions as a legitimate therapeutic modality to improve employee mental health outcomes beyond experimental approaches [63].

The review is expected to reveal core intervention components that characterize effective programs, including psychoeducational modules, skill-building exercises, multimedia integration, and personalization features. We anticipate that effective interventions will demonstrate sophisticated adaptation of evidence-based psychological frameworks, such as cognitive behavioral therapy, acceptance and commitment therapy, social cognitive theory, and positive psychology, to digital formats while maintaining theoretical fidelity. Additionally, the review is also expected to uncover and capture implementation challenges such as engagement barriers, time constraints, privacy concerns, and variations in organizational support that may influence intervention effectiveness [64]. By focusing exclusively on self-guided formats, this review will shift attention from resource-intensive guided interventions to scalable solutions that could overcome traditional barriers to mental health support in organizational settings [65].

Comparison to Previous Work

Existing literature on workplace mental health interventions has established the efficacy of various psychological frameworks, including cognitive behavioral therapy and mindfulness-based approaches, in reducing occupational stress and burnout [66-68]. However, the translation of these frameworks to fully self-guided digital formats remains understudied. Our review builds upon foundational work in digital mental health while extending the focus to implementation characteristics, theoretical mechanisms, and contextual factors unique to self-guided workplace interventions. This targeted approach is expected to reveal insights into the minimum viable components necessary for effective self-guided

XSL•FO
RenderX

digital interventions, the role of personalization and adaptation, and the organizational conditions that facilitate or hinder intervention success.

Strengths, Limitations, and Future Directions

This systematic review protocol possesses several methodological strengths that will enhance the reliability and comprehensiveness of the findings. The multidatabase search strategy (PubMed, Web of Science, PsycINFO, Cochrane) combined with additional source identification through existing systematic reviews will minimize publication bias and ensure broad literature coverage. The temporal scope spanning from 2000 will capture the evolution of digital mental health technologies and their application in workplace contexts.

The review's exclusive focus on self-guided interventions addresses a critical research gap, as these formats offer unique scalability advantages essential for widespread implementation in organizational settings. The planned detailed analysis of intervention components, delivery mechanisms, and theoretical foundations will provide valuable insights for both researchers and practitioners. Additionally, the comprehensive data extraction protocol will enable a nuanced understanding of how various design features and implementation contexts influence intervention effectiveness.

However, several anticipated limitations warrant consideration. The exclusion of guided or blended interventions, while maintaining methodological focus, may overlook important insights about the continuum of support needed in digital workplace mental health programs. Self-guided interventions may face unique challenges, including lower engagement rates compared to guided formats, and the review may reveal limited information about implementation contexts, organizational support structures, and workplace-specific adaptations. The predominance of studies from high-income countries may also limit the generalizability of the findings to diverse global workplace contexts.

The anticipated contribution of this review is multifaceted, advancing both academic understanding and practical application of self-guided digital mental health interventions in workplace settings. For the research community, this review will provide a comprehensive synthesis of the current evidence base, identify critical knowledge gaps, and establish a foundation for future comparative effectiveness studies. The detailed analysis of theoretical frameworks and intervention components is expected to inform the development of next-generation digital interventions that leverage emerging technologies while maintaining evidence-based foundations.

For organizational stakeholders and practitioners, this review will offer practical guidance on selecting, implementing, and evaluating self-guided digital mental health programs, enabling evidence-informed decision-making about digital health investments through the identification of core intervention components and implementation facilitators. The findings will be particularly valuable for small- and medium-sized enterprises that may lack resources for traditional employee assistance programs yet seek scalable, accessible mental health solutions. From a policy perspective, the review will provide essential evidence for developing standards and guidelines that ensure workplace digital mental health programs balance innovation with ethical safeguards and appropriate oversight. By synthesizing effectiveness data, implementation considerations, and quality indicators, the review will inform policy frameworks that promote equitable access to evidence-based digital interventions across diverse workplace populations. Future research directions will likely include investigations into optimal intervention dosage and duration, comparative analyses of theoretical frameworks across similar populations, and evaluations of implementation strategies that enhance engagement and effectiveness. Moreover, the anticipated findings will underscore the need for standardized outcome measures and reporting protocols to facilitate meta-analytic synthesis, while recognizing that emerging technologies—such as artificial intelligence and virtual reality—require ongoing, rigorous evaluation to ensure that innovation remains evidence-based and continues to expand scalable, effective workplace mental health support.

Conclusion

The findings of this systematic review will inform the design of innovative, scalable, and accessible interventions aimed at enhancing psychological well-being and reducing burnout among working adults.

The significance of this review extends beyond documenting intervention effectiveness to addressing fundamental questions about the role of self-guided digital support within comprehensive workplace mental health strategies. As the global workforce continues to face unprecedented stressors and organizations seek scalable solutions to support employee well-being, understanding the capabilities and limitations of self-guided interventions becomes increasingly critical. The review's findings are expected to challenge traditional assumptions about the necessity of human guidance in digital mental health interventions while identifying the conditions under which self-guided support can achieve meaningful clinical and organizational outcomes.

Ultimately, this systematic review will contribute to the transformation of workplace mental health support from a reactive, resource-intensive model to a proactive, accessible, and scalable approach that can reach diverse employee populations. This proposed review has the potential to advance our understanding of how digital tools can create psychologically healthy workplaces that not only prevent burnout and distress but also actively promote employee thriving and organizational success. The knowledge generated through this systematic review will be instrumental in realizing the potential of digital mental health interventions to address the growing mental health crisis in workplace settings worldwide.

XSL•FO
RenderX

Acknowledgments

This study protocol is part of the EE's PhD Thesis in Applied Organizational Psychology at Saint Mary's University, Canada, and portions of the text may appear in the final thesis. The study protocol was created by EE, JF, and RAH.

Funding

The authors received no direct funding for this study.

Authors' Contributions

Conceptualization: EE (lead), JF (supporting), RAH (supporting) Methodology: EE (lead), JF (supporting), RAH (supporting) Writing – original draft: EE (lead) Writing – review & editing: EE (equal), JF (equal), RAH (supporting) Supervision: JF (lead), RAH (supporting) All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Checklist 1
PRISMA-P checklist.
[DOCX File, 39 KB - resprot_v15i1e80417_app1.docx]

References

1. Nearly 4 in 10 professionals report rising burnout. Robert Half Canada. 2023. URL: https://press.roberthalf.ca/2023-05-31-Nearly-4-in-10-Professionals-Report-Rising-Burnout [accessed 2026-01-06]
2. State of the global workplace report. Gallup. 2024. URL: https://www.gallup.com/workplace/349484/state-of-the-global-workplace.aspx [accessed 2026-01-07]
3. Work-related stress most often caused by heavy workloads and work-life balance. Statistics Canada. 2023 Jun 19. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/230619/dq230619c-eng.htm [accessed 2025-03-24]
4. 2024 workforce state of mind. Headspace. 2024. URL: https://get.headspace.com/2024-workforce-state-of-mind [accessed 2025-03-24]
5. Amanullah S, Ramesh Shankar R. The impact of COVID-19 on physician burnout globally: a review. Healthcare (Basel) 2020 Oct 22;8(4):421. [doi: 10.3390/healthcare8040421] [Medline: 33105757]
6. Boudreau RA, Grieco RL, Cahoon SL, Robertson RC, Wedel RJ. The pandemic from within: two surveys of physician burnout in Canada. Can J Community Ment Health 2006 Sep 1;25(2):71-88. [doi: 10.7870/cjcmh-2006-0014]
7. Etezad S. Secure your own mask first before assisting others: investigating the health of frontline care workers during the COVID-19 pandemic [Master's Thesis]. : Saint Mary's University; 2021 URL: http://library2.smu.ca/xmlui/handle/01/29948 [accessed 2023-01-27]
8. Etezad S, Fleming M, Weigand HA, et al. Exploring the well-being of community pharmacy professionals, turnover intention and patient safety: time to include operational responsibility. Can Pharm J (Ott) 2023;156(2):71-84. [doi: 10.1177/17151635231152170] [Medline: 36969306]
9. Nahrgang JD, Morgeson FP, Hofmann DA. Safety at work: a meta-analytic investigation of the link between job demands, job resources, burnout, engagement, and safety outcomes. J Appl Psychol 2011 Jan;96(1):71-94. [doi: 10.1037/a0021484] [Medline: 21171732]
10. Van Bogaert P, Meulemans H, Clarke S, Vermeyen K, Van de Heyning P. Hospital nurse practice environment, burnout, job outcomes and quality of care: test of a structural equation model. J Adv Nurs 2009 Oct;65(10):2175-2185. [doi: 10.1111/j.1365-2648.2009.05082.x] [Medline: 20568322]
11. West CP, Dyrbye LN, Erwin PJ, Shanafelt TD. Interventions to prevent and reduce physician burnout: a systematic review and meta-analysis. Lancet 2016 Nov 5;388(10057):2272-2281. [doi: 10.1016/S0140-6736(16)31279-X] [Medline: 27692469]
12. Parandeh A, Ashtari S, Rahimi-Bashar F, Gohari-Moghadam K, Vahedian-Azimi A. Prevalence of burnout among health care workers during coronavirus disease (COVID-19) pandemic: a systematic review and meta-analysis. Prof Psychol Res Pract 2022;53(6):564-573. [doi: 10.1037/pro0000483]
13. A profession under pressure: results from the CMA's 2021 national physician health survey. Canadian Medical Association. 2021. URL: https://www.cma.ca/news/profession-under-pressure-results-cmas-2021-national-physician-health-survey [accessed 2025-12-12]
14. Better understanding of youth mental health. Lancet 2017 Apr 29;389(10080):1670. [doi: 10.1016/S0140-6736(17)31140-6] [Medline: 28463125]
15. Friedrich MJ. Depression is the leading cause of disability around the world. JAMA 2017 Apr 18;317(15):1517. [doi: 10.1001/jama.2017.3826] [Medline: 28418490]

XSL•FO
RenderX

16. Psychological health and safety in the workplace: employer practices in response to COVID-19. CSA Group. 2022. URL: https://www.csagroup.org/wp-content/uploads/CSA-Group-Research-Psychological-Health-and-Safety-in-the-Workplace-Employer-Practices-COVID-19.pdf [accessed 2025-03-28]
17. Dozois DJA, Mental Health Research Canada. Anxiety and depression in Canada during the COVID-19 pandemic: a national survey. Canadian Psychology / Psychologie canadienne 2021;62(1):136-142. [doi: 10.1037/cap0000251]
18. Rolfe K. "A pandemic after the pandemic": insurers brace for disability claims "deluge" from mental, physical strain of crisis. Financial Post. 2021. URL: https://financialpost.com/fp-work/a-pandemic-after-the-pandemic-insurers-brace-for-disability-claims-deluge-from-mental-physical-strain-of-crisis [accessed 2021-08-17]
19. Telus Health. Four ways to prepare for the disability deluge. Health Benefits Hub May. 2021 May 7. URL: https://plus.telushealth.co/blogs/health-benefits/en/four-ways-to-prepare-for-the-disability-deluge/ [accessed 2022-01-23]
20. The ROI in workplace mental health programs: good for people, good for business. Deloitte Insights. 2019. URL: https://www2.deloitte.com/us/en/insights/topics/talent/workplace-mental-health-programs-worker-productivity.html [accessed 2022-01-24]
21. Burn-out an "occupational phenomenon": international classification of diseases. World Health Organization. 2019. URL: https://www.who.int/news/item/28-05-2019-burn-out-an-occupational-phenomenon-international-classification-of-diseases [accessed 2021-07-11]
22. Maslach C, Jackson SE, Leiter MP, Schaufeli WB, Schwab RL. Maslach burnout inventory instruments and scoring guides. Mind Garden Inc. 1996. URL: http://mindgarden.com/117-maslach-burnout-inventory-mbi [accessed 2024-12-09]
23. Bell C, Roberts GLR, Millear PM, et al. The emotionally exhausted treating the mentally unwell? A systematic review of burnout and stress interventions for psychologists. Clin Psychol Psychother 2024 Jan;31(1):e2909. [doi: 10.1002/cpp.2909] [Medline: 39084599]
24. Ng IK, Goh WG, Thong C, Teo KS. "Quiet quitting" among medical practitioners: a hallmark of burnout, disillusionment and cynicism. J R Soc Med 2025 Mar;118(3):73-77. [doi: 10.1177/01410768241311059] [Medline: 39946171]
25. Leiter MP, Maslach C. Areas of worklife: a structured approach to organizational predictors of job burnout. In: Emotional and Physiological Processes and Positive Intervention Strategies: Elsevier Science/JAI Press; 2004:91-134. [doi: 10.1016/S1479-3555(03)03003-8]
26. Leiter MP, Maslach C. Latent burnout profiles: a new approach to understanding the burnout experience. Burn Res 2016 Dec;3(4):89-100. [doi: 10.1016/j.burn.2016.09.001]
27. Maslach C, Schaufeli WB, Leiter MP. Job burnout. Annu Rev Psychol 2001;52(1):397-422. [doi: 10.1146/annurev.psych.52.1.397] [Medline: 11148311]
28. Maslach C, Leiter MP. Understanding the burnout experience: recent research and its implications for psychiatry. World Psychiatry 2016 Jun;15(2):103-111. [doi: 10.1002/wps.20311] [Medline: 27265691]
29. Eisenberger R, Huntington R, Hutchison S, Sowa D. Perceived organizational support. J Appl Psychol 1986;71(3):500-507. [doi: 10.1037//0021-9010.71.3.500]
30. Eisenberger R, Cummings J, Armeli S, Lynch P. Perceived organizational support, discretionary treatment, and job satisfaction. J Appl Psychol 1997 Oct;82(5):812-820. [doi: 10.1037/0021-9010.82.5.812] [Medline: 9337610]
31. Shanafelt TD, Gorringe G, Menaker R, et al. Impact of organizational leadership on physician burnout and satisfaction. Mayo Clin Proc 2015 Apr;90(4):432-440. [doi: 10.1016/j.mayocp.2015.01.012] [Medline: 25796117]
32. Layard R, Clark D, Knapp M, Mayraz G. Cost-benefit analysis of psychological therapy. Natl Inst econ rev 2007 Oct;202(1):90-98. [doi: 10.1177/0027950107086171]
33. Vasiliadis HM, Tempier R, Lesage A, Kates N. General practice and mental health care: determinants of outpatient service use. Can J Psychiatry 2009 Jul;54(7):468-476. [doi: 10.1177/070674370905400708] [Medline: 19660169]
34. Global burden of mental disorders and the need for a comprehensive, coordinated response from health and social sectors at the country level: report by the Secretariat. : World Health Organization; 2012 A65/10 URL: https://apps.who.int/iris/handle/10665/78898 [accessed 2025-03-24]
35. Gilin DA, Anderson GG, Etezad S, Lee-Baggley D, Cooper AM, Preston RJ. Impact of a wellness leadership intervention on the empathy, burnout, and resting heart rate of medical faculty. Mayo Clin Proc Innov Qual Outcomes 2023 Dec;7(6):545-555. [doi: 10.1016/j.mayocpiqo.2023.09.005] [Medline: 38075429]
36. Geraedts AS, Kleiboer AM, Twisk J, Wiezer NM, van Mechelen W, Cuijpers P. Long-term results of a web-based guided self-help intervention for employees with depressive symptoms: randomized controlled trial. J Med Internet Res 2014 Jul 9;16(7):e168. [doi: 10.2196/jmir.3539] [Medline: 25008127]
37. Gold KJ, Dobson ML, Sen A. "Three good things" digital intervention among health care workers: a randomized controlled trial. Ann Fam Med 2023;21(3):220-226. [doi: 10.1370/afm.2963] [Medline: 37217328]
38. Grime PR. Computerized cognitive behavioural therapy at work: a randomized controlled trial in employees with recent stress-related absenteeism. Occup Med 2004 Aug 1;54(5):353-359. [doi: 10.1093/occmed/kqh077] [Medline: 15289593]
39. Knaak S, Mantler E, Szeto A. Mental illness-related stigma in healthcare: barriers to access and care and evidence-based solutions. Healthc Manage Forum 2017 Mar;30(2):111-116. [doi: 10.1177/0840470416679413] [Medline: 28929889]

XSL•FO
RenderX

40. Petrie K, Crawford J, Baker STE, et al. Interventions to reduce symptoms of common mental disorders and suicidal ideation in physicians: a systematic review and meta-analysis. Lancet Psychiatry 2019 Mar;6(3):225-234. [doi: 10.1016/S2215-0366(18)30509-1] [Medline: 30744997]
41. McFarling L, D'Angelo M, Drain M, Gibbs DA, Rae Olmsted KL. Stigma as a barrier to substance abuse and mental health treatment. Mil Psychol 2011 Jan 1;23(1):1-5. [doi: 10.1080/08995605.2011.534397]
42. Almeda N, Díaz-Milanés D, Guiterrez-Colosia MR, García-Alonso CR. A systematic review of the international evolution of online mental health strategies and recommendations during the COVID-19 pandemic. BMC Psychiatry 2022 Sep 20;22(1):621. [doi: 10.1186/s12888-022-04257-8] [Medline: 36127666]
43. Adam D, Berschick J, Schiele JK, et al. Interventions to reduce stress and prevent burnout in healthcare professionals supported by digital applications: a scoping review. Front Public Health 2023;11:1231266. [doi: 10.3389/fpubh.2023.1231266] [Medline: 38026413]
44. Cameron G, Mulvenna M, Ennis E, et al. Effectiveness of digital mental health interventions in the workplace: umbrella review of systematic reviews. JMIR Ment Health 2025 Jan 24;12:e67785. [doi: 10.2196/67785] [Medline: 39854722]
45. Henshall C, Ostinelli E, Harvey J, et al. Examining the effectiveness of web-based interventions to enhance resilience in health care professionals: systematic review. JMIR Med Educ 2022 Sep 6;8(3):e34230. [doi: 10.2196/34230] [Medline: 36066962]
46. Linardon J, Torous J, Firth J, Cuijpers P, Messer M, Fuller-Tyszkiewicz M. Current evidence on the efficacy of mental health smartphone apps for symptoms of depression and anxiety. A meta-analysis of 176 randomized controlled trials. World Psychiatry 2024 Feb;23(1):139-149. [doi: 10.1002/wps.21183] [Medline: 38214614]
47. Schäfer SK, von Boros L, Schaubruch LM, et al. Digital interventions to promote psychological resilience: a systematic review and meta-analysis. NPJ Digit Med 2024 Feb 8;7(1):30. [doi: 10.1038/s41746-024-01017-8] [Medline: 38332030]
48. Paleri V, Valmaggia L, Kaleva I, Martland R, Riches S. Digital staff support interventions for the psychological wellbeing of healthcare professionals: a systematic review. J technol Behav Sci 2024;10(2):250-282. [doi: 10.1007/s41347-024-00434-6]
49. Angerer F, Mennel V, Grund S, et al. Mechanisms of change in digital interventions for depression: a systematic review and meta-analysis of six mediator domains. J Affect Disord 2025 Jan 1;368:615-632. [doi: 10.1016/j.jad.2024.09.055] [Medline: 39284530]
50. Jonas B, Leuschner F, Tossmann P. Efficacy of an internet-based intervention for burnout: a randomized controlled trial in the German working population. Anxiety Stress Coping 2017 Mar;30(2):133-144. [doi: 10.1080/10615806.2016.1233324] [Medline: 27602992]
51. Prasek A. Randomized controlled trial to evaluate a self-guided, web-based mindfulness program for stress reduction and wellbeing promotion [Dissertation]. : University of Minnesota; 2015 URL: https://hdl.handle.net/11299/175304 [accessed 2025-12-09]
52. Super A, Yarker J, Lewis R, Keightley S, Summers D, Munir F. Developing self-compassion in healthcare professionals utilising a brief online intervention: a randomised waitlist control trial. Int J Environ Res Public Health 2024 Oct 11;21(10):1346. [doi: 10.3390/ijerph21101346] [Medline: 39457319]
53. Brindal E, Kakoschke N, Golley S, Rebuli M, Baird D. Effectiveness and feasibility of a self-guided mobile app targeting emotional well-being in healthy adults: 4-week randomized controlled trial. JMIR Ment Health 2023 Apr 3;10:e44925. [doi: 10.2196/44925] [Medline: 37010898]
54. Rojon C, McDowall A, Saunders MNK. On the experience of conducting a systematic review in industrial, work, and organizational psychology. J Pers Psychol 2011 Jan;10(3):133-138. [doi: 10.1027/1866-5888/a000041]
55. Siddaway AP, Wood AM, Hedges LV. How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. Annu Rev Psychol 2019 Jan 4;70(1):747-770. [doi: 10.1146/annurev-psych-010418-102803] [Medline: 30089228]
56. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
57. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
58. Chen X, Zou D, Xie H, Cheng G. Twenty years of personalized language learning: topic modeling and knowledge mapping. Educ Technol Soc 2021;24(1):205-222. [doi: 10.30191/ETS.202101_24(1).0016]
59. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series, 1st edition: Wiley; 2008. [doi: 10.1002/9780470712184]
60. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898. [doi: 10.1136/bmj.l4898] [Medline: 31462531]
61. Popay J, Roberts H, Sowden A, et al. Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC methods programme. : Lancaster University; 2006. [doi: 10.13140/2.1.1018.4643]
62. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008 Apr 26;336(7650):924-926. [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]

XSL•FO
RenderX

63. Armeni P, Polat I, De Rossi LM, Diaferia L, Meregalli S, Gatti A. Exploring the potential of digital therapeutics: an assessment of progress and promise. Digit Health 2024;10:20552076241277441. [doi: 10.1177/20552076241277441] [Medline: 39291152]
64. Underdahl L, Ditri M, Duthely LM. Physician burnout: evidence-based roadmaps to prioritizing and supporting personal wellbeing. J Healthc Leadersh 2024;16:15-27. [doi: 10.2147/JHL.S389245] [Medline: 38192639]
65. Aye LM, Tan MM, Schaefer A, et al. Self-help digital mental health intervention in improving burnout and mental health outcomes among healthcare workers: a narrative review. Digit Health 2024;10:20552076241278313. [doi: 10.1177/20552076241278313] [Medline: 39257871]
66. Varela-Agra N, Rueda-Extremera M, Dorta-Afonso D, Gómez-Martínez S, Cantero-García M. Psychological interventions to reduce burnout and depression in healthcare professionals: a systematic review. Curr Psychol 2025 Apr;44(8):6963-6977. [doi: 10.1007/s12144-025-07655-9]
67. Musker M, Othman S. Effective interventions to reduce burnout in nurses: a meta-analysis. Complement Ther Clin Pract 2024 Feb;54:101827. [doi: 10.1016/j.ctcp.2023.101827] [Medline: 38181650]
68. Bagasi A, Al Harbi EK, Alabbasi SM, Alqaedi RO, Alharbi BA, Alhomaid TA. Effectiveness of workplace mental health programs in reducing occupational burnout: a systematic review. Cureus 2025 Jul;17(7):e88715. [doi: 10.7759/cureus.88715] [Medline: 40861700]

Abbreviations:

GRADE: Grading of Recommendations Assessment, Development and Evaluation
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols
ROB: risk of bias

Edited by A Schwartz; submitted 10.Jul.2025; peer-reviewed by A Adepoju, G Cameron, S Mohmad; accepted 25.Nov.2025; published 12.Jan.2026.

Please cite as:
Etezad E, Fiset J, Al Hajj R
Digital Self-Guided Mental Health Interventions to Prevent Workplace Burnout and Enhance Psychological Wellness: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e80417
URL: https://www.researchprotocols.org/2026/1/e80417
doi:10.2196/80417

© Ehsan Etezad, John Fiset, Raghid Al Hajj. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Effect of Multimodal Interventions on Pain, Stress, and Neurodevelopmental Outcomes in Neonates With Malignant and Benign Tumors: Protocol for a Systematic Review

Sharath Hullumani[*], MPT; Shrutika Sharad[*], BPT; Aishik Ghosh, BPT; Rutuja Sawalkar[*], BPT; Irshad Qureshi[*], BPT, MPT

Faculty of Physiotherapy, Datta Meghe Institute of Higher Education and Research (DU), Sawangi Meghe, Wardha, Maharashtra, India

[*]these authors contributed equally

Corresponding Author:
Sharath Hullumani, MPT
Faculty of Physiotherapy, Datta Meghe Institute of Higher Education and Research (DU), Sawangi Meghe, Wardha, Maharashtra, India

Abstract

Background: Both benign and malignant tumors in neonates frequently necessitate invasive diagnostic and therapeutic procedures, exposing infants to significant pain and stress during critical periods of brain development. Procedural stress can disrupt the hypothalamic-pituitary-adrenal axis, affect synaptic pruning, and alter myelination, potentially leading to long-term cognitive and neurodevelopmental impairments. While standard medical care remains the cornerstone of management, there is growing interest in multimodal interventions—including integrative therapies, physiotherapy, and non-pharmacological approaches such as massage, music therapy, kangaroo care, and sensory stimulation—to mitigate discomfort and support neurodevelopmental outcomes.

Objective: This study aimed to systematically evaluate how well multimodal therapies reduce stress and suffering while enhancing neurodevelopmental outcomes in infants with both benign and malignant tumors.

Methods: The PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) standards will be followed in the conduct of this systematic review. Studies that use multimodal therapies (eg, physiotherapy, massage, music therapy, kangaroo care, or other non-pharmacologic techniques) for newborns with proven malignant or benign tumors will be considered. A thorough literature search will be conducted using databases such as the PubMed, Embase, and CINAHL, Cochrane, and CancerLi. Cohort studies, case-control studies, quasi-experimental studies, and randomized controlled trials will all be considered. Measures of pain (such as validated neonatal pain scales), stress (such as cortisol levels and behavioral markers), and neurodevelopment (such as the Bayley Scales of Infant Development) will be the main objectives.

Results: The anticipated timeline for completing this systematic review is June 2025 to December 2027. Two independent reviewers will conduct each stage of the review process in a blinded manner to minimize bias.

Conclusions: This systematic review aims to provide evidence-based insights into the role of multimodal interventions in enhancing standard care for neonates undergoing tumor-related procedures. Conclusions regarding clinical impact will be conditional on the quality and consistency of the evidence, as assessed using GRADE (Grading of Recommendations, Assessment, Development, and Evaluation). The findings are expected to guide future research, inform cautious clinical recommendations, and support the development of integrative strategies to optimize neurodevelopmental outcomes in this vulnerable population.

Trial Registration: PROSPERO CRD420251053012; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251053012

International Registered Report Identifier (IRRID): PRR1-10.2196/77969

(JMIR Res Protoc 2026;15:e77969) doi:10.2196/77969

KEYWORDS

stress reduction; non-pharmacological therapy; integrative therapy; pain; newborns; neonate

Introduction

Rationale

Newborns with both benign and malignant tumors frequently undergo invasive diagnostic and therapeutic procedures—including surgery, chemotherapy, and prolonged hospital stays—subjecting them to substantial physical and psychological stress. Early-life stress can adversely affect behavior, brain development, and long-term health outcomes by disrupting the hypothalamic-pituitary-adrenal axis, impairing

XSL•FO
RenderX

synaptic pruning, and altering myelination. Non-pharmacological and physiotherapeutic interventions, such as music therapy, massage, kangaroo care, and sensory stimulation, are increasingly being investigated as strategies to mitigate these negative effects. However, the efficacy of such multimodal approaches in neonatal oncology remains poorly supported by aggregated evidence [1-3].

The rarity and heterogeneity of neonatal tumors make it challenging to define their true prevalence, sites of origin, and clinical behavior. Neonatal tumors occur in approximately 1 per 12,500 - 27,500 live births and account for roughly 2% of pediatric malignancies. Despite their low incidence, these tumors offer a unique opportunity to study tumor biology with minimal confounding from environmental factors. Common tumor types include teratomas, neuroblastomas, mesoblastic nephromas, and fibromatoses, which often present as a mass at birth and are frequently detected through prenatal ultrasonography [4].

Although benign tumors may occasionally undergo malignant transformation, true malignancies in the neonatal period are rare. Histological features do not always correlate with clinical behavior, complicating accurate classification. Benign tumors can be life-threatening depending on their size or location, whereas some histologically malignant tumors may exhibit indolent or ambiguous behavior [5-11]. Genetic predispositions, such as constitutional chromosomal abnormalities, can favor tumor development during the fetal and neonatal stages. Environmental exposures—including ionizing radiation, maternal medications, infections, and maternal tumors—may also contribute to tumor risk, although evidence varies in strength [6]. Given these complexities, a systematic synthesis of evidence on multimodal, non-pharmacological interventions is critical to inform supportive care strategies aimed at minimizing procedural stress and optimizing neurodevelopment in this vulnerable population.

Objective

This study aimed to systematically evaluate the efficacy of multimodal, non-pharmacological interventions—including physiotherapy, integrative therapies, and sensory-based approaches—in reducing procedural pain and stress, and enhancing short- and long-term neurodevelopmental outcomes in neonates with benign and malignant tumors. The review will also identify intervention timing, delivery methods, and gaps in current evidence to guide future research and clinical practice.

Methods

Overview

The PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) guidelines will be followed in conducting this review. Evidence from observational and interventional trials that evaluate the use of multimodal (non-pharmacological) treatments in newborns with tumors will be identified and synthesized in this review. A narrative synthesis and meta-analysis (if applicable) will be performed to evaluate findings and guide future research and clinical practices.

Eligibility Criteria

The eligibility criteria were determined using the PICOT (population, intervention, comparison, outcome measures, and timing) framework

The population (P) included neonates (≤28 days) diagnosed with benign or malignant tumors who were treated with non-pharmacologic interventions.

The intervention (I) was multimodal, non-pharmacological interventions, including but not limited to physiotherapy (eg, chest physiotherapy or gentle handling), sensory stimulation (visual, auditory, or tactile), music therapy, massage therapy, Kangaroo mother care, non-nutritive sucking, and additional integrative or supportive non-pharmacological techniques.

The comparison (C) was between standard medical care without adjunct non-pharmacologic intervention and the absence of intervention or placebo controls, where applicable.

The outcome measure (O) was pain alleviation; this was assessed using validated neonatal pain scales, such as the Premature Infant Pain Profile (PIPP) orPremature Infant Pain Profile-Revised (PIPP-R) and Neonatal Infant Pain Scale (NIPS).

The main aim was stress reduction that was measured via physiological markers (eg, salivary or serum cortisol) and behavioral indicators (eg, facial expressions, crying, and agitation); this assessment was performed immediately or short-term, as indicated above.

Neurodevelopmental outcomes were assessed using standardized tools such as the Bayley Scale of Infant and Toddler Development, electroencephalography, and relevant imaging markers. Short-term (up to 1 mo post-intervention) and long-term follow-up (≥1 mo to assess ongoing neurodevelopment) assessments were performed.

Subgroup analyses are pre-specified for the (1) intervention modality, (2) tumor type (benign vs malignant), and (3) outcome category. This structure minimizes uninformative averaging.

Study Design

The study will be focused on observational studies, experimental studies, and randomized controlled trials (RCTs).

Languages of the Included Studies

Only studies published in English or with available translations will be included. This restriction is due to resource limitations for accurate translation and verification. To minimize potential bias, non-English abstracts will be screened, and studies deemed highly relevant will be noted, with attempts to access translations when feasible.

Publication Date

Studies published from January 2000 to May 2025 will be included. No restrictions were applied on the year of publication to ensure the inclusivity of all relevant evidence.

Publication Type Included

The types of studies included were both published articles and gray literature.

XSL•FO
RenderX

Information Sources

A comprehensive search will be conducted using both electronic databases and manual methods to identify studies relevant to multimodal, non-pharmacological interventions in neonates with tumors. The electronic databases searched included PubMed, Embase, Cochrane Library, CancerLit (an oncology-specific database), and Google Scholar (to supplement formal database searches). Search strategies will combine controlled vocabulary terms (eg, MeSH and Emtree) and free-text keywords, customized for each database. Tumor-specific terms (eg, neonatal tumor, neoplasm, neuroblastoma, teratoma, and germ cell tumor), intervention terms (eg, physiotherapy, massage, music therapy, and kangaroo care), and validated neonatal outcome measures (eg, PIPP-R and NIPS) will be included. The full database-specific search strings are provided in Checklist 1.

Eligibility of Evidence

Only studies available in full text will be included. Case reports, case series, and expert opinions will be excluded from data extraction due to limited generalizability. Both published literature and gray literature (theses, conference abstracts, clinical trial registries) will be screened, provided full text is accessible.

Search Strategy

The search strategy will combine controlled vocabulary terms (eg, MeSH and Emtree) with free-text keywords, tailored for each database. Key terms will include “physiotherapy,” “stress reduction,” “non-pharmacological therapy,” “integrative therapy,” “pain,” and “newborn,” along with tumor-specific terms (ie, “neonatal tumor,” “neoplasm,” “neuroblastoma,” and “teratoma”). Truncation (*) and Boolean operators (AND, OR) will be used to ensure comprehensive retrieval. Medline and the Cochrane Library will particularly utilize MeSH terms for additional coverage. The full database-specific search strings are provided in Checklist 1. To ensure the inclusion of the most recent studies, all searches will be updated prior to the final analysis. In addition, trial registries, investigator contacts, and conference proceedings will be searched for unpublished studies. Both published and gray literature (eg, conference abstracts, theses, and clinical trial registries) will be included if the full text is available.

Study Selection and Synthesis

Observational studies without control groups will be included in narrative synthesis but excluded from quantitative meta-analysis. Only studies in English or with available translations will be included, with non-English abstracts screened for potential relevance.

Selection Process and Data Collection Process

Two calibrated and independent reviewers (SH and SS) will perform the study in two stages. The first step involves screening abstracts and titles for alignment with the inclusion criteria and research question. To decide on the final inclusion, the full texts of possibly pertinent research are examined in the second stage. This procedure will be guided by the inclusion and exclusion criteria. In phase 1, the titles and abstracts of the chosen papers will be read by the two reviewers, SH and SS. The entire texts of the previously included papers will be read by the same reviewers (SH and SQ) in phase two. The selection criteria will be the same as those from the first round. A third reviewer (IQ) will make the decision if there is any dispute.

Data Items

The data synthesis will concentrate on assessing how well multimodal therapies work to improve neurodevelopmental outcomes, pain, and stress in newborns with benign and malignant tumors. Pre-established outcome measures will be used to aggregate the quantitative and qualitative data from qualified research. A narrative summary will be presented for each of the included studies, organized according to the demographic characteristics, intervention type, and outcomes that were measured.

A meta-analysis will be carried out using R software (version 4.1.2; R Foundation for Statistical Computing), including the tidyverse package (version 1.3.2) when the outcome data are sufficiently consistent across trials in terms of the design, intervention, and measurement instruments. For continuous outcomes, such as pain scores, cortisol levels, and neurodevelopmental test scores, effect sizes will be presented as 95% CI or the mean difference or standardized mean difference. The risk ratio or odds ratio with 95% CI will be used to express dichotomous outcomes, such as the presence or absence of adverse occurrences. Given the expected clinical and methodological heterogeneity, a random-effects model will be used.

Assessment of Risk of Bias in Included Studies

The risk of bias (RoB) in included studies will be assessed independently by two reviewers (SH and SS), with disagreements resolved by a third reviewer (AG). The choice of the RoB tool will depend on the study design: for RCTs, the RoB 2.0 will be used; for non-randomized intervention studies, ROBINS-I will be used; and for observational studies, ROBINS-E will be used [12].

Before formal assessment, reviewers will discuss each tool and standardize the evaluation criteria, followed by a calibration exercise on a sample of included studies to ensure consistency. Inter-rater agreement will be calculated using Cohen κ, with $\kappa \geq 0.80$ considered acceptable. The RoB results will be presented according to each tool’s recommended format and visualized using Robvis [13]. Publication bias will be assessed if ≥10 studies are available for a given outcome. Funnel plots will be generated, and asymmetry will be tested using the Egger or Begg tests, as appropriate. This structured approach ensures reproducibility, transparency, and reliable evaluation of the methodological quality across diverse study designs.

Data Synthesis

Data synthesis will be conducted according to study design, with RCTs pooled using RoB 2 assessments, non-randomized intervention studies evaluated with ROBINS-I, and observational exposure-based studies assessed using ROBINS-E. Where sufficient homogeneous RCT data are available, meta-analyses will be performed using standardized mean differences for

XSL•FO
RenderX

continuous outcomes and relative risks for dichotomous outcomes, each reported with 95% CIs. Funnel plots and Egger tests will be applied only to homogeneous RCT subsets to assess publication bias. Given the anticipated heterogeneity across interventions, tumor types, and outcome measures, a narrative synthesis will serve as the primary approach. This narrative summary will describe key outcomes, intervention types, and their effects on pain, stress, and neurodevelopment; highlight patterns and gaps in the evidence; and include pictorial representations such as summary tables and figures to improve clarity. Subgroup analyses will be conducted, where possible, to explore variations in intervention effects according to participant characteristics (eg, tumor type, gestational age, and baseline clinical status) and intervention attributes (eg, modality, frequency, and timing). This combined quantitative and qualitative approach ensures a rigorous and comprehensive synthesis of available evidence while accounting for heterogeneity and supporting meaningful conclusions regarding multimodal interventions in neonates with tumors.

Confidence in Cumulative Evidence

Using the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) method, this systematic review will be assessed. When it comes to handling neonates with respiratory distress syndrome, the GRADE approach may be able to perform a thorough evaluation of the presence of evidence that supports the efficacy of multimodal treatments.

Results

The anticipated timeline for completing this systematic review is June 2025 to December 2027. Two independent reviewers will conduct each stage of the review process in a blinded manner to minimize bias. Regular team meetings will be held to resolve any discrepancies or conflicts, with all discussions documented to ensure transparency and reproducibility. The review will be carried out in strict adherence to the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) guidelines, and the protocol has been prospectively registered with PROSPERO (CRD420251053012). Upon completion, the full review and its findings will be prepared and submitted for peer-reviewed publication.

Discussion

Anticipated Findings

It is anticipated that this systematic review will identify a spectrum of non-pharmacological interventions capable of reducing pain and physiological stress while promoting early neurodevelopment in neonates diagnosed with benign or malignant tumors. The evidence may highlight specific multimodal strategies—such as physiotherapy, sensory stimulation, and integrative therapies—that are both effective and feasible within neonatal oncology and critical care contexts. These findings are expected to contribute valuable insights into how supportive, non-pharmacological approaches can complement standard medical treatments in this highly vulnerable population.

Strengths and Limitations

This protocol has been developed in alignment with the MOOSE (Meta-Analyses and Systematic Reviews of Observational Studies) and PRISMA-P guidelines, ensuring methodological rigor and transparency. Two independent reviewers will screen and extract data, with a third reviewer mediating any disagreements, thereby minimizing selection and extraction bias. However, certain limitations are anticipated. Restricting inclusion to English-language studies may reduce the comprehensiveness of the evidence base due to translation constraints. Additionally, the inclusion of non-randomized studies may introduce bias; however, this will be systematically assessed and discussed as part of the review's limitations. Despite these constraints, the review will help identify knowledge gaps and inform priorities for future empirical and interventional research in neonatal oncology care.

Conclusion and Dissemination

By synthesizing available evidence on multimodal non-pharmacological therapies, this review aims to clarify their role in enhancing both clinical and developmental outcomes in neonates with tumors. The findings will contribute to the formulation of evidence-based recommendations and integrated care strategies that prioritize comfort, neuroprotection, and developmental support. Dissemination of the results will occur through peer-reviewed publication, presentations at neonatal and pediatric conferences, and sharing within professional clinical and academic networks to ensure the translation of findings into practice and guide future research directions.

Acknowledgments

No generative AI tools were used.

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Authors' Contributions

Conceptualization: SH, SS
Data curation: SH, SS
Funding acquisition: SH, RR

XSL•FO
RenderX

Investigation: SH, IQ, AG
Methodology: SH, IQ, SS
Project administration: SH, RR
Software: SH, RS
Supervision: IQ, RR
Validation: SH, IQ, RR
Visualization: SH, IQ, RR
Writing – original draft: SH
Writing – review & editing: SH

Conflicts of Interest

None declared.

Checklist 1
PRISMA checklist.
[PDF File, 89 KB - resprot_v15i1e77969_app1.pdf]

References

1. Othman AH, Mohamad MF, Sayed HAR. Transdermal fentanyl for cancer pain management in opioid-naive pediatric cancer patients. Pain Med 2016 Jul 1;17(7):1329-1336. [doi: 10.1093/pm/pnw004] [Medline: 26921071]
2. Sposito NPB, Rossato LM, Bueno M, Kimura AF, Costa T, Guedes DMB. Assessment and management of pain in newborns hospitalized in a neonatal intensive care unit: a cross-sectional study. Rev Lat Am Enfermagem 2017 Sep 12;25:e2931. [doi: 10.1590/1518-8345.1665.2931] [Medline: 28902934]
3. Anbazhagan S, Priyadarshini GN, Subramanian VV, Parthasarathy S. Addressing pain in neonatal cancers: a systematic review. J Neonatal Surg 2025;14(4):320-327. [doi: 10.52783/jns.v14.2782]
4. Solanki S, Menon P, Samujh R, Gupta K, Rao KLN. Clinical presentation and surgical management of neonatal tumors: retrospective analysis. J Indian Assoc Pediatr Surg 2020;25(2):85-90. [doi: 10.4103/jiaps.JIAPS_241_18] [Medline: 32139986]
5. De Ioris MA, Fabozzi F, Lodi M, et al. The fight just born-neonatal cancer: rare occurrence with a favorable outcome but challenging management. Cancers (Basel) 2022 Apr 29;14(9):2244. [doi: 10.3390/cancers14092244] [Medline: 35565372]
6. Yeh JM, Stout NK, Chaudhry A, et al. Universal newborn genetic screening for pediatric cancer predisposition syndromes: model-based insights. Genet Med 2021 Jul;23(7):1366-1371. [doi: 10.1038/s41436-021-01124-x] [Medline: 33767345]
7. Williams MD, Lascelles BDX. Early neonatal pain-a review of clinical and experimental implications on painful conditions later in life. Front Pediatr 2020;8:30. [doi: 10.3389/fped.2020.00030] [Medline: 32117835]
8. Neshat H, Hassankhani H, Jabraeili M, Negarandeh R. Organisational challenges of pain management in neonatal intensive care unit: a qualitative study. BMJ Open 2023 Sep 5;13(9):e072695. [doi: 10.1136/bmjopen-2023-072695] [Medline: 37669843]
9. Gaspari S, Di Ruscio V, Stocchi F, Carta R, Becilli M, De Ioris MA. Case report: early association of vemurafenib to standard chemotherapy in multisystem langerhans cell histiocytosis in a newborn: taking a chance for a better outcome? Front Oncol 2021;11:794498. [doi: 10.3389/fonc.2021.794498] [Medline: 34966688]
10. Gow KW, Dellinger M. Neonatal tumors. In: Pediatric and Neonatal Surgery 2017. [doi: 10.5772/68070]
11. Cooper TE, Heathcote LC, Anderson B, Grégoire MC, Ljungman G, Eccleston C. Non-steroidal anti-inflammatory drugs (NSAIDs) for cancer-related pain in children and adolescents. Cochrane Database Syst Rev 2017 Jul 24;7(7):CD012563. [doi: 10.1002/14651858.CD012563.pub2] [Medline: 28737843]
12. Risk of bias tools. riskofbias.info. URL: https://www.riskofbias.info/ [accessed 2025-12-12]
13. Risk of bias tools. robvis (visualization tool). URL: https://www.riskofbias.info/welcome/robvis-visualization-tool [accessed 2025-12-12]

Abbreviations

GRADE: Grading of Recommendations, Assessment, Development, and Evaluation
MOOSE: Meta-Analyses and Systematic Reviews of Observational Studies
NIPS: Neonatal Infant Pain Scale
PICOT: population, intervention, comparison, outcome measures, and timing
PIPP/PIPP-R: Premature Infant Pain Profile
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
RCT: randomized controlled trial
RoB: risk of bias

XSL•FO
RenderX

Edited by J Sarvestan; submitted 23.May.2025; peer-reviewed by D Poddighe, R Nooripour; revised version received 16.Oct.2025; accepted 17.Oct.2025; published 13.Jan.2026.

Please cite as:
Hullumani S, Sharad S, Ghosh A, Sawalkar R, Qureshi I
Effect of Multimodal Interventions on Pain, Stress, and Neurodevelopmental Outcomes in Neonates With Malignant and Benign Tumors: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e77969
URL: https://www.researchprotocols.org/2026/1/e77969
doi:10.2196/77969

© Sharath Hullumani, Shrutika Sharad, Aishik Ghosh, Rutuja Sawalkar, Irshad Qureshi. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Lived Experiences of Children Who Have Undergone Kidney Replacement Therapy and Their Families: Protocol for a Qualitative Systematic Review

Kenji Takao[1*], PhD; Maki Fujitsuka[2*], MNS; Shingo Ueki[2*], PhD; Tomoyuki Sakai[3*], MD, PhD

[1]Department of Nursing, Faculty of Nursing, Osaka Dental University, 11-8 Kuzuha Hanazono-cho, Hirakata, Japan

[2]Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan

[3]Department of Pediatrics, Shiga University of Medical Science, Otsu, Shiga, Japan

[*]all authors contributed equally

Corresponding Author:
Kenji Takao, PhD
Department of Nursing, Faculty of Nursing, Osaka Dental University, 11-8 Kuzuha Hanazono-cho, Hirakata, Japan

Abstract

Background: In many high-income countries, 5 to 10 children per million of the age-related population start kidney replacement therapy (KRT), which includes both dialysis (peritoneal dialysis and hemodialysis) and kidney transplantation (KT) for end-stage kidney disease. After peritoneal dialysis is introduced, or after KT, self-care at home typically becomes the main focus. Providing support for each developmental stage and transition period in the treatment process from dialysis to KT is an urgent issue.

Objective: This review aims to synthesize the lived experiences of children undergoing KRT and their families. We will use identified themes to develop a self-care program aimed at solving the life tasks of children and their families.

Methods: A search strategy will follow the Joanna Briggs Institute methodology and will be conducted in 3 steps: an initial limited search, a comprehensive database search, and a reference search of the included articles. MEDLINE (EBSCO), CINAHL Plus, and PsycINFO will be searched with no restriction on language or publication dates. The study selection, critical appraisal, data extraction, and data synthesis will be performed according to Joanna Briggs Institute guidelines for systematic reviews of qualitative research. Final synthesis will be assessed using the ConQual (confidence in the output of qualitative research synthesis) approach.

Results: The review will include studies focusing on the experiences of children with KRT and their families. These experiences include physical, mental, and social issues, hygiene care, diet, fluid intake, medication, strict infection prevention, delays in growth and development, restrictions on social life, and a lack of social resources. This is an important issue because after starting dialysis, or after KT, many children are treated at home and experience these issues in their local communities and homes. As of May 2025, the authors have conducted 2 pilot searches to test and refine keywords of results with the help of the librarian and have identified 1003 studies for screening in MEDLINE. This systematic review is scheduled to be completed by April 2026.

Conclusions: This systematic review synthesizes qualitative evidence regarding the daily life experiences of children and their families after initiating KRT, contributing to the development of a self-care program that enhances their quality of life.

Trial Registration: PROSPERO CRD42025644504; https://www.crd.york.ac.uk/PROSPERO/view/CRD42025644504

International Registered Report Identifier (IRRID): PRR1-10.2196/77725

(JMIR Res Protoc 2026;15:e77725) doi:10.2196/77725

KEYWORDS

end-stage kidney disease in children; kidney replacement therapy; patient and family experience; qualitative research; quality of life

Introduction

Background

End-stage kidney disease (ESKD) in children is a rare and irreversible condition that develops as a result of chronic kidney disease. Previous studies in many high-income countries have reported that approximately 5-10 children per million of the age-related population experience ESKD and undergo kidney replacement therapy (KRT) each year, although definitions and selection criteria vary from country to country [1-3]. The most common causes of ESKD in children are congenital anomalies of the kidney and urinary tract and hereditary nephropathy [4-6]. However, because ESKD is also associated with preterm birth

XSL•FO
RenderX

and low birth weight, its incidence is increasing worldwide with recent advances in neonatal care [7]. Treatment for ESKD in children involves various KRT methods, such as hemodialysis (HD), peritoneal dialysis (PD), and kidney transplantation (KT), although treatment may be limited by financial, social, or religious constraints [8]. PD is considered the first choice for dialysis in children because it does not require painful vascular access, has less of an effect on circulation, and involves less strict dietary restrictions than HD. Additionally, it is easier for children to go to school or kindergarten while receiving home health care [9]. In most high-income countries, national registries for KT (including dialysis) have been implemented to describe the global incidence of pediatric KRT and to capture variations in treatment and outcomes [2]. As a result of advances in KRT, children and their families can now receive treatment not only in hospitals but also in their own communities. KRT cannot be maintained without being positioned within the lives of children and their families, including their growth and development, so self-care to improve quality of life (QOL) becomes a necessity.

Previous studies on QOL have reported that children on dialysis and children who have received a KT exhibit significantly impaired health-related QOL [10,11]. In addition, children with ESKD have been found to exhibit a significant decrease in health-related QOL at all stages [12]. In addition, KT has been found to improve QOL compared with dialysis, and children who have received KT exhibit better growth, neurocognitive development, academic performance, and QOL than children who have received dialysis [13]. KT represents an essential treatment for children in terms of physical and psychomotor development, and the benefits of successful KT are immense. According to the North America Pediatric Renal Transplant Cooperative Study of 2495 children under 18 years of age at the time of transplant, the rate of graft survival in the preemptive renal transplant (PRT) group was clearly superior to that in the non-PRT group [14]. Additionally, the estimated life expectancy of pediatric transplant patients is 30 to 40 years longer than that of pediatric dialysis patients [14]. The importance of early KT has also been demonstrated in terms of life prognosis [15]. However, there are also issues with donors and medical facilities in PRT, and children with ESKD often cannot avoid dialysis. Therefore, the transition to dialysis or kidney transplants is a continuous process that children and their families follow as KRT, and they face many physical, mental, and social difficulties in their daily lives, either for short periods of time or across developmental stages. For example, with regard to oral medication, infants and preschool children may find it difficult to take medication due to the bitter taste, discomfort, or shape of the medicine, and even if they do manage to take it, it may induce vomiting, which can lead to distress when they have to take it again. For school-age children and adolescents, the factors that hinder medication adherence are said to be at the individual, family, health care system, and community levels [16]. Difficulties in daily life not only affect children's self-care but also the QOL of children and their families. It is necessary to consider ways to improve the issues in daily life in accordance with the developmental stage of the child.

We conducted a preliminary search of the following databases in April 2025: PROSPERO, MEDLINE (EBSCO), CINAHL Plus (EBSCOhost), and Joanna Briggs Institute (JBI) Evidence Synthesis. Few systematic reviews have examined KRT. Previous reviews can be divided into reviews focusing on children who have experienced PD/HD or their families [17,18] and those focusing on children who have experienced KT or their families [19]. However, children with ESKD often undergo repeated KRT throughout their lives. Thus, no previous systematic reviews have attempted to synthesize the qualitative evidence regarding the lived experiences of PD/HD and KT in children with ESKD and their families. Although substantial progress has been made in the development of KRT methods, the issues facing children and their families remain unclear. It is necessary to clarify these difficulties and to ensure that holistic medical care is provided for children and their families by the medical staff involved [20].

Objectives

This systematic review will synthesize evidence regarding the life experiences of children undergoing KRT and their families and to clarify the life issues. We aim to develop a life-support system focused on the perspectives of children undergoing KRT and their families. The primary and secondary objectives of this study are as follows:

- The primary objective is the needs and dissatisfaction experienced by children and their families undergoing KRT.
- The secondary objective is the contextual and cultural factors influencing these family experiences, as well as the interactions between children and families.

Review Questions

The review question for this study is "What needs and dissatisfaction are experienced by children and their families in daily life during KRT, including dialysis (PD/HD) and KT?" The subquestions are as follows:

- What are the contextual and cultural factors that influence the experiences of these children and families?
- What is the type of interaction between the children and their families?

Methods

Methodology and Protocol Registration

The proposed systematic review will be conducted in accordance with the JBI methodology for systematic reviews of qualitative evidence [21]. This protocol was registered in PROSPERO (CRD 42025644504).

Eligibility Criteria

Eligible studies will include qualitative research designs (eg, phenomenology, grounded theory, ethnography, and qualitative descriptive studies) focusing on patient or family experiences after the initiation of KRT.

Quantitative studies, mixed method studies without extractable qualitative data, editorials, and commentaries will be excluded.

XSL•FO
RenderX

Population

The proposed review will include studies of children who have undergone KRT for ESKD (glomerular filtration rate <15 ml/min/1.73 m^2) when they were 18 years or younger and their families. Although no age limit at the time of study participation will be applied, it is likely that young adults (under 39 y) will be selected because the use of PD began in the 1980s, followed by the development of automated PD and continuous hemodiafiltration in the 1990s, as well as the advent of superior immunosuppressants [9]. We will include cases of families that meet the inclusion criteria even if their child is no longer alive when the study recruitment takes place. KRT includes 3 types of treatment: PD, HD, and KT. KT includes both living-donor and deceased-donor KT. Additionally, KT also includes PRT, in which a kidney is transplanted before dialysis is needed, as well as including dialysis after PRT and second transplants. Data extraction will be conducted independently by two reviewers using a predeveloped standardized form. Information will be extracted separately for each treatment modality—HD, PD, and KT. For each included study, the following population characteristics will be extracted (where available): age range or mean age of participants, sex distribution (male/female), underlying kidney disease (including proportions by diagnostic category, if reported), comorbidities or complications, and family relationship to the patient (eg, parent, child, caregiver).

These details will help contextualize the diversity of participant backgrounds and ensure comparability across studies.

In this review, “family” is defined as a person identified as the primary caregiver for the child or as holding a significant role within the household. Therefore, biological parents, siblings, and grandparents are included as family members. Nonbiological relatives such as adoptive parents, parents of children born through surrogacy, and stepparents are also included. During screening, both reviewers will independently confirm whether the study explicitly describes participants as family members or primary caregivers. Any ambiguity will be clarified by contacting the authors to clarify or consulting with a third reviewer (SU).

Phenomena of Interest

This review will consider research that explores the lived experiences of children who have undergone KRT (ie, PD, HD, and KT) and their families. These experiences will include individuals’ demands and needs regarding KRT in daily life.

Context

This review will consider studies that are conducted in any context of the children’s lives to reflect the diverse voices of children and families with KRT. This includes the community, school, family settings, leisure activities, and visits to hospitals. Moreover, this review will not be restricted to any country or region.

Types of Studies

We will consider all studies in which qualitative data were collected, including, but not limited to, designs such as phenomenology, grounded theory, ethnography, action research, discourse analysis, and qualitative descriptions. Qualitative components of mixed methods studies that contain qualitative data will also be included if the qualitative results are reported separately. We will review academic papers, theses and dissertations, and conference papers. Moreover, the research will only include studies in which informed consent was obtained from the parents and informed assent was obtained from the children to ensure thorough consideration of ethical aspects in our review because the research involves minors.

Search Strategy

The search strategy will aim to locate both published and unpublished studies. An initial limited search of MEDLINE (EBSCO) and CINAHL was undertaken to identify articles on the topic. The text words contained in the titles and abstracts of relevant articles and the index terms used to describe the articles were used to develop a full search strategy (Table 1). The search strategy, including all identified keywords and index terms, will be adapted for each included information source. The reference list of all studies selected for critical appraisal will be screened for additional studies. There will be no restrictions in terms of language or year of publication. For studies published in languages other than English and Japanese, we will use translation software. The databases to be searched will include MEDLINE (EBSCO), CINAHL, PsycINFO, and Igaku Chuo Zasshi. The unpublished studies will be searched using Mednar. Gray literature will be searched using ProQuest Dissertations and Theses and Google Scholar.

XSL•FO
RenderX

Table . Search conducted on MEDLINE (EBSCO) on April 23, 2025.

Search	Query	Records retrieved, n
1[a]	MH[b] "Renal Replacement Therapy+" OR MH "Renal Dialysis+" OR MH "Peritoneal Dialysis+" OR (MH "Kidney Transplantation" OR MH "Organ Transplantation+")	380,767
2[a]	MH "Child+" OR MH "Infant+" OR MH "Adolescent" OR MH "Minors+" OR MH "Family+" OR MH "Parents+" OR MH "Caregivers"	4,346,908
3[c]	MH "Adverse Childhood Experiences" OR MH "health services needs and demand" OR MH Stress, Psychological OR MH Quality of Life OR MM[d] "Patient Preference" OR MH "Decision Making+" OR MH "Patient-Centered Care+"	756,711
4[c]	XB[e] (experience* OR demand* OR need* OR stress* OR feel* OR anxiety* OR burden* OR preference* OR dissatisf* OR life OR live*)	7,886,771
5	3 OR 4	8,137,015
6[f]	(MH "Interviews as Topic" OR MH "Interview") OR MH "Focus Groups" OR MH "Qualitative Research+" OR MH "Grounded Theory" OR MH "Anthropology, Cultural+"	390,534
7[f]	XB (qualitative* OR Narrative OR ethno* OR Phenomenology OR "action research" OR "content analysis" OR "thematic analysis" OR interview* OR "mixed method*")	1,158,783
8	6 OR 7	1,320,861
9	1 AND 2 AND 5 AND 8	1003

[a]1 and 2: Population.

[b]MH: MeSH heading.

[c]3 and 4: Phenomena of interest.

[d]MM: major MeSH heading

[e]XB: exploded MeSH heading.

[f]6 and 7: Methods and data collection approaches related to qualitative research.

Study Selection

Following the search, all identified citations will be collated and uploaded into EndNote 21 (Clarivate Analytics), and duplicates will be removed. A pilot screening involving 30 studies will then be conducted by 2 reviewers (KT and MF) to independently review the titles and abstracts. We will continue to perform pilot tests and engage in discussion until agreement is reached. The titles and abstracts will then be screened against the review inclusion criteria. Potentially relevant studies will be retrieved in full and their citation details imported into the JBI System for the Unified Management, Assessment, and Review of Information (JBI SUMARI; JBI) [22].

The full text of selected citations will be assessed in detail against the inclusion criteria by 2 independent reviewers. Reasons for exclusion of full-text studies that do not meet the inclusion criteria will be recorded and reported in the systematic review. Any disagreements that arise between the reviewers at each stage of the study selection process will be resolved through discussion or with a third reviewer (SU).

Assessment of Methodological Quality

Eligible studies will be critically appraised by 2 independent reviewers (KT and MF) for methodological quality using the standard JBI critical appraisal checklist [21] for qualitative research. Authors of studies will be contacted to request missing or additional data for clarification, where required. The evaluation content of the checklist includes congruity between the stated philosophical perspective and the research methodology; congruity between the research methodology and the research question or objectives; congruity between the research methodology and the methods used to collect data; congruity between the research methodology and the representation and analysis of data; congruity between the research methodology and the interpretation of results; a statement locating the researcher culturally or theoretically; the influence of the researcher on the research, and vice versa, addressed; are participants, and their voices, adequately represented; evidence of ethical approval by an appropriate body; and the conclusions drawn in the research report flow from the analysis, or interpretation, of the data. All items were evaluated by using "yes," "no," "unclear," and "not applicable"

XSL•FO
RenderX

to appraisal. The appraisal items are listed in Table 2. Any disagreements that arise between the reviewers will be resolved through discussion or with a third reviewer (SU). The results of critical appraisal will be reported in narrative form and in formats. Following critical appraisal, only studies rated "no" for Q8 (Are participants, and their voices, adequately represented?) will be excluded. This item is treated as a methodological requirement specific to JBI meta-aggregation because findings must be supported by participant quotations (illustrations). When participant voices are not adequately represented, the necessary illustrations for extraction and credibility assessment are unavailable, making synthesis technically impossible. Therefore, excluding studies with a "no" rating on Q8 is not a quality-based threshold but a requirement for data extractability. For all other items on the checklist, studies rated "no" or "unclear" will be included as these limitations do not prevent extraction of findings and illustrations. If there is insufficient information to assess methodological quality, it will be rated as "unclear" without contacting the study authors.

Table . Joanna Briggs Institute quality appraisal tool.

	Yes	No	Unclear	N/A[a]
1. Is there congruity between the stated philosophical perspective and the research methodology?	☐	☐	☐	☐
2. Is there congruity between the research methodology and the research question or objectives?	☐	☐	☐	☐
3. Is there congruity between the research methodology and the methods used to collect data?	☐	☐	☐	☐
4. Is there congruity between the research methodology and the representation and analysis of data?	☐	☐	☐	☐
5. Is there congruity between the research methodology and the interpretation of results?	☐	☐	☐	☐
6. Is there a statement locating the researcher culturally or theoretically?	☐	☐	☐	☐
7. Is the influence of the researcher on the research, and vice versa, addressed?	☐	☐	☐	☐
8. Are participants, and their voices, adequately represented?	☐	☐	☐	☐
9. Is the research ethical according to current criteria or, for recent studies, and is there evidence of ethical approval by an appropriate body?	☐	☐	☐	☐
10. Do the conclusions drawn in the research report flow from the analysis, or interpretation, of the data?	☐	☐	☐	☐
Overall appraisal: Include ☐ Exclude ☐ Seek further info ☐				
Comments (Including reason for exclusion)				

[a]Not Applicable.

XSL•FO
RenderX

Data Extraction

Data will be extracted from studies included in the review by 2 independent reviewers (KT and MF) using the standardized JBI data extraction tool in JBI SUMARI [22]. The extracted data will include specific details about the populations, context, culture, geographical location, study methods, and the phenomena of interest relevant to the review objective. In this review, context refers to family composition, the setting where treatment occurs (hospital, home, etc), health care and insurance systems, and schooling. Culture refers to lifestyle habits such as attitudes toward treatment (views on dialysis and transplantation), family roles, and dietary customs. Focusing on and exploring the daily lives of children and their families, their findings and illustrations will be extracted and assigned a level of credibility. When required, authors of specific studies will be contacted to request missing or additional data.

Data Synthesis

Qualitative research findings will, where possible, be pooled using JBI SUMARI with the meta-aggregation approach. This will involve the aggregation or synthesis of findings to generate a set of statements that represent that aggregation, through assembling the findings and categorizing these findings on the basis of similarity in meaning. These categories will then be subjected to a synthesis in order to produce a single comprehensive set of synthesized findings that can be used as a basis for evidence-based practice. Where textual pooling is not possible, the findings will be presented in narrative form. Only unequivocal and credible findings will be included in the synthesis. The terms “unequivocal” and “credible” will follow the JBI’s explanation:

- Unequivocal relates to evidence beyond reasonable doubt, which may include conclusions that are matter of fact directly reported/observed and not open to challenge.
- Credible relates to those conclusions that are, albeit interpretations, plausible in light of the textual data and theoretical framework. As the conclusions are interpretive, they can be challenged.

Only findings classified as unequivocal or credible will be included in the synthesis; unsupported statements will be excluded.

Assessing Confidence in the Findings

The methodological quality of included studies will be appraised using the JBI Critical Appraisal Checklist for Qualitative Research (Table 2).

The ConQual (confidence in the output of qualitative research synthesis) approach will be used to establish the confidence level in each synthesized finding [23] (Textbox 1).

Each finding will receive an initial high score that is downgraded based on 2 domains: dependability (reflecting methodological rigor) and credibility (reflecting the strength of evidence support).

A summary table will present the final ConQual scores for all synthesized findings.

The summary of findings will include the major elements of the review and detail how the ConQual score has been developed and the title, population, phenomena of interest, and context for the specific review.

Textbox 1. ConQual (confidence in the output of qualitative research synthesis) summary of findings

Review title: The lived experiences of children who have undergone kidney replacement therapy (KRT) and their families: a qualitative systematic review

Population: children who have undergone KRT for end-stage kidney disease when they were 18 years old or younger and their families.

Phenomena of interest: the exposition of lived experiences of children who have undergone KRT and their families.

Context: the lived experiences of children who have undergone KRT and their families.

Summary of findings will include:

- Synthesized finding
- Type of research
- Dependability
- Credibility
- ConQual score

Results

The results will report on all studies concerning children and families’ experiences with daily life focusing on KRT, including cleanliness care, diet, fluid intake, medication, strict infection prevention, delay in growth and development, restriction of social life, and insufficient social resources. The results of the search will be reported in full in the final systematic review and presented in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 1) [24].

As of May 2025, the authors have conducted 2 pilot searches to test and refine keywords of results with the help of the librarian. We have identified 1003 studies for screening in MEDLINE. We will apply the number of records identified, screened, excluded, and included at each stage in Figure 1, after the full search and study selection are performed. This qualitative synthesis will determine the current studies exploring the lived experiences of children undergoing KRT and their

XSL•FO
RenderX

families. The results will presumably report on 3 treatment types such as PD, HD, and KT.

We will summarize our findings in a table and a conceptual diagram and discuss them in a descriptive and narrative review form. We will discuss the implications for future research, policies, and practices. Our findings will be presented at a relevant conference and submitted in a peer-reviewed journal. The review team is currently engaged in data extraction, and this systematic review is expected to be completed by April 2026.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram for new systematic reviews which included searches of databases, registers, and other sources. If feasible to do so, we will report the number of records identified from each database or register searched (rather than the total number across all databases/registers). If automation tools were used, we will indicate how many records were excluded by a human and how many were excluded by automation tools. Source: Page et al [24].

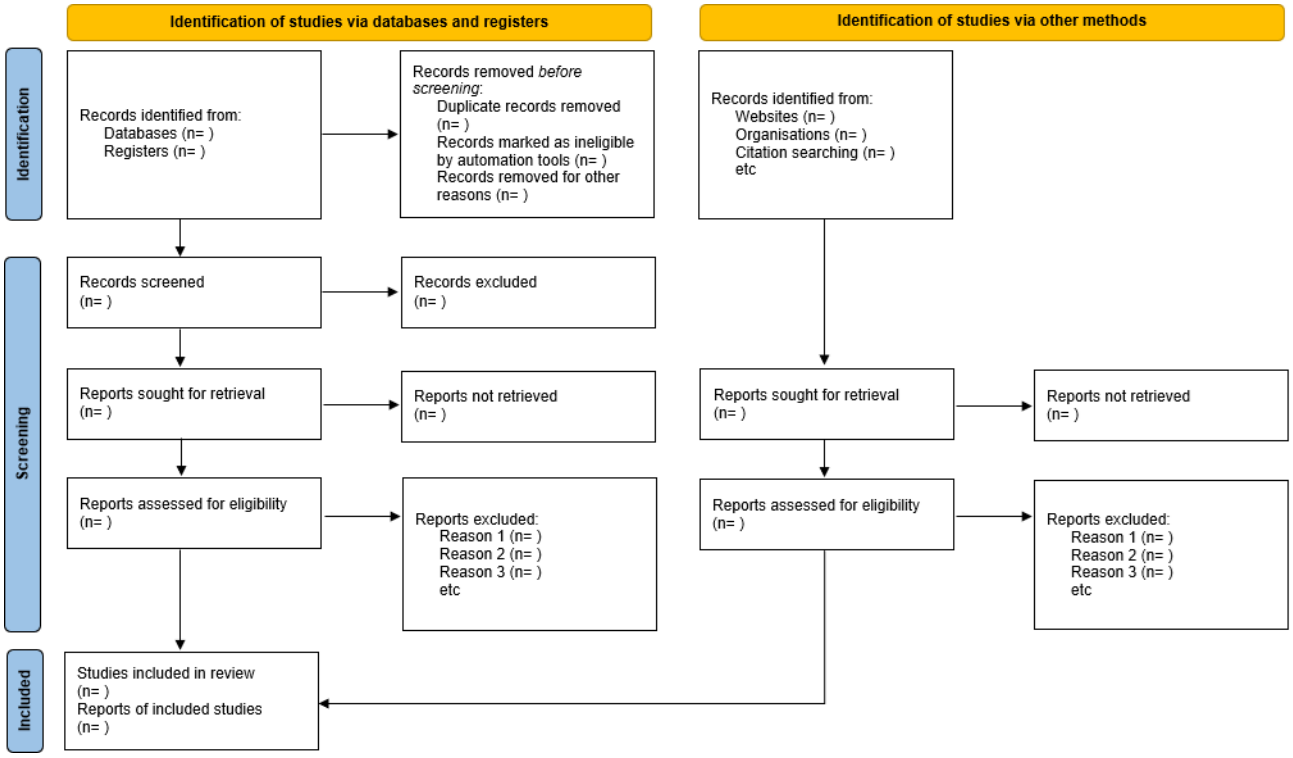

Discussion

Principal Findings

Children with ESKD must continue to undergo some form of KRT throughout their lives. To our knowledge, there is no other qualitative synthesis investigating the experiences of children undergoing KRT (particularly from dialysis to KT) and their families. This qualitative synthesis will be unique in identifying issues for seamless support for children and their families. Globally, there is a need for health care that emphasizes the perspectives of patients and their families. It is essential to investigate the real voices of patients and their families. By conducting a systematic review from the perspective of patients and their families, it will be possible to bridge the gap between the care provided by health care professionals and the care sought by patients and their families.

Strengths and Limitations of This Study

The qualitative systematic review will evaluate the lived experiences from the perspectives of both children undergoing KRT and their families and may reveal comprehensive demands and mutually beneficial opportunities for qualitative improvement for the whole family. Dialysis and kidney transplants are integrated into a comprehensive process of living experience as KRT. This study included the literature from multidisciplinary databases, gray literature, and all languages to maximize data diversity and comprehensiveness. Human errors are minimized by having two independent researchers carry out literature selection, quality assessment, and data extraction according to the JBI methodology. Meta-integration will be performed to gain a comprehensive understanding of the topic and identify gaps for further research.

Conclusions

This systematic review protocol aims to comprehensively synthesize qualitative research findings regarding the daily life experiences—such as difficulties, needs, and dissatisfactions—faced by children and their families after initiating KRT.

We will use identified themes to develop a self-care program aimed at solving the life tasks of children and their families. This research will contribute to the realization of a society in which the QOL of children and their families is ensured to the greatest extent possible.

XSL•FO
RenderX

Acknowledgments

We would like to express our sincere gratitude to Osaka Dental University Library for creating the search strategy. We thank Benjamin Knight, MSc, from Edanz for editing a draft of this manuscript.

Funding

This work was supported by JSPS KAKENHI (grant-in-aid for young scientists: JP25K20764), which assisted with the collection of articles and Open Access fees, and by Osaka Dental University Personal Research funding, which assisted with translation from other languages into English and English-language editing of manuscripts.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed in this study protocol.

Authors' Contributions

Conceptualization: KT, MF, TS
Data curation: KT
Funding acquisition: KT.
Methodology: KT, MF, SU
Validation: MF
Writing – original draft: KT
Writing – review & editing: MF, SU, TS
All authors read and approved the final protocol.

Conflicts of Interest

None declared.

References

1. Harambat J, van Stralen KJ, Kim JJ, Tizard EJ. Epidemiology of chronic kidney disease in children. Pediatr Nephrol 2012 Mar;27(3):363-373. [doi: 10.1007/s00467-011-1939-1] [Medline: 21713524]
2. Ploos van Amstel S, Noordzij M, Warady BA, et al. Renal replacement therapy for children throughout the world: the need for a global registry. Pediatr Nephrol 2018 May;33(5):863-871. [doi: 10.1007/s00467-017-3863-5] [Medline: 29273970]
3. Maurer E, Neuhaus TJ, Weitz M, Kuehni CE, Laube GF. Paediatric end-stage renal disease and renal replacement therapy in Switzerland: survival and treatment trends over four decades. Swiss Med Wkly 2020 Jul 27;150:w20300. [doi: 10.4414/smw.2020.20300] [Medline: 32920790]
4. Mong Hiep TT, Ismaili K, Collart F, et al. Clinical characteristics and outcomes of children with stage 3-5 chronic kidney disease. Pediatr Nephrol 2010 May;25(5):935-940. [doi: 10.1007/s00467-009-1424-2] [Medline: 20148340]
5. Ardissino G, Daccò V, Testa S, et al. Epidemiology of chronic renal failure in children: data from the ItalKid project. Pediatrics 2003 Apr;111(4 Pt 1):e382-e387. [doi: 10.1542/peds.111.4.e382] [Medline: 12671156]
6. Ishikura K, Uemura O, Ito S, et al. Pre-dialysis chronic kidney disease in children: results of a nationwide survey in Japan. Nephrol Dial Transplant 2013 Sep;28(9):2345-2355. [doi: 10.1093/ndt/gfs611] [Medline: 23825101]
7. Sanderson KR, Shih WV, Warady BA, Claes DJ. Severe fetal CAKUT (Congenital Anomalies of the Kidneys and Urinary Tract), prenatal consultations, and initiation of neonatal dialysis. Am J Perinatol 2024 May;41(S 01):e156-e162. [doi: 10.1055/a-1850-4429] [Medline: 35554891]
8. Lalji R, Francis A, Wong G, et al. Disparities in end-stage kidney disease care for children: a global survey. Kidney Int 2020 Sep;98(3):527-532. [doi: 10.1016/j.kint.2020.04.058] [Medline: 32828230]
9. Hattori M. The history and current state of the treatment of end-stage chronic renal failure in children: from the 53rd Japanese Society for Dialysis Therapy Educational Lecture. J Jpn Soc Dial Ther 2009;42(2):137-144. [doi: 10.4009/jsdt.42.137]
10. Splinter A, Tjaden LA, Haverman L, et al. Children on dialysis as well as renal transplanted children report severely impaired health-related quality of life. Qual Life Res 2018 Jun;27(6):1445-1454. [doi: 10.1007/s11136-018-1789-4] [Medline: 29374855]
11. Doshi K, Raina R, Ng KH, et al. Health-related quality of life for pediatric patients with end-stage kidney disease: a systematic review and meta-analysis of the Pediatric Quality of Life Inventory (PedsQL). Hemodial Int 2024 Apr;28(2):198-215. [doi: 10.1111/hdi.13138] [Medline: 38468403]
12. Francis A, Didsbury MS, van Zwieten A, et al. Quality of life of children and adolescents with chronic kidney disease: a cross-sectional study. Arch Dis Child 2019 Feb;104(2):134-140. [doi: 10.1136/archdischild-2018-314934] [Medline: 30018070]

XSL•FO
RenderX

13. Francis A, Johnson DW, Melk A, et al. Survival after kidney transplantation during childhood and adolescence. Clin J Am Soc Nephrol 2020 Mar 6;15(3):392-400. [doi: 10.2215/CJN.07070619] [Medline: 32075809]
14. Vats AN, Donaldson L, Fine RN, Chavers BM. Pretransplant dialysis status and outcome of renal transplantation in North American children: a NAPRTCS Study12. Transplantation 2000 Apr;69(7):1414-1419. [doi: 10.1097/00007890-200004150-00035]
15. McDonald SP, Craig JC, Australian and New Zealand Paediatric Nephrology Association. Long-term survival of children with end-stage renal disease. N Engl J Med 2004 Jun 24;350(26):2654-2662. [doi: 10.1056/NEJMoa031643] [Medline: 15215481]
16. Varnell CD Jr, Rich KL, Nichols M, et al. Assessing barriers to adherence in routine clinical care for pediatric kidney transplant patients. Pediatr Transplant 2017 Nov;21(7). [doi: 10.1111/petr.13027] [Medline: 28762577]
17. Tjaden L, Tong A, Henning P, Groothoff J, Craig JC. Children's experiences of dialysis: a systematic review of qualitative studies. Arch Dis Child 2012 May;97(5):395-402. [doi: 10.1136/archdischild-2011-300639] [Medline: 22399673]
18. Tong A, Lowe A, Sainsbury P, Craig JC. Experiences of parents who have children with chronic kidney disease: a systematic review of qualitative studies. Pediatrics 2008 Feb;121(2):349-360. [doi: 10.1542/peds.2006-3470] [Medline: 18245427]
19. Araújo NSS, Pereira RRF, Fram D, Hino P, Longo MCB, Taminato M. Quality of life in children with kidney transplant: systematic review. Rev Bras Enferm 2018;71(suppl 6):2818-2823. [doi: 10.1590/0034-7167-2018-0464] [Medline: 30540061]
20. Gotoh Y, Nagai T, Yamakawa S, Uemura O. The present conditions in the pediatric renal transplant and problems: based on a pediatric renal transplant patients of 55 this hospital. Jpn J Pediatr Nephrol 2008;21(1):36-41. [doi: 10.3165/jjpn.21.36]
21. Porritt K, Evans C, Bennett C, et al. Systematic reviews of qualitative evidence. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI; 2024. [doi: 10.46658/jbimes-24-02]
22. Munn Z, Aromataris E, Tufanaru C, et al. The development of software to support multiple systematic review types: the Joanna Briggs Institute System for the Unified Management, Assessment and Review of Information (JBI SUMARI). Int J Evid Based Healthc 2019 Mar;17(1):36-43. [doi: 10.1097/XEB.0000000000000152] [Medline: 30239357]
23. Munn Z, Porritt K, Lockwood C, Aromataris E, Pearson A. Establishing confidence in the output of qualitative research synthesis: the ConQual approach. BMC Med Res Methodol 2014 Sep 20;14:108. [doi: 10.1186/1471-2288-14-108] [Medline: 25927294]
24. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]

Abbreviations

ConQual: confidence in the output of qualitative research synthesis
ESKD: end-stage kidney disease
HD: hemodialysis
JBI: Joanna Briggs Institute
JBI SUMARI: JBI System for the Unified Management, Assessment, and Review of Information
KRT: kidney replacement therapy
KT: kidney transplantation
PD: peritoneal dialysis
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRT: preemptive renal transplant
QOL: quality of life

Edited by J Sarvestan; submitted 19.May.2025; peer-reviewed by CG Ambarsari; revised version received 28.Nov.2025; accepted 30.Nov.2025; published 12.Jan.2026.

Please cite as:
Takao K, Fujitsuka M, Ueki S, Sakai T
The Lived Experiences of Children Who Have Undergone Kidney Replacement Therapy and Their Families: Protocol for a Qualitative Systematic Review
JMIR Res Protoc 2026;15:e77725
URL: https://www.researchprotocols.org/2026/1/e77725
doi:10.2196/77725

© Kenji Takao, Maki Fujitsuka, Shingo Ueki, Tomoyuki Sakai. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Quality in Use in Connected Mental Health: Protocol for a Systematic Mapping Study

Shweta Premanandan[1], PhD; Sofia Ouhbi[1], PhD; Magdalena Ramstedt Stadin[2], PhD; Charlotte Blease[3], PhD; Åsa Cajander[2], PhD; Maria Hägglund[3], PhD

[1]Division of Computing Science, Department of Information Technology, Uppsala University, Uppsala, Sweden

[2]Division of Vi3, Department of Information Technology, Uppsala University, Uppsala, Sweden

[3]Participatory eHealth and Health Data Research Group, Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden

Corresponding Author:
Shweta Premanandan, PhD
Division of Computing Science
Department of Information Technology
Uppsala University
Box 337
Hus 10, Regementsvägen 10
Uppsala, 75105
Sweden
Phone: 46 0707689331
Email: shweta.premanandan@it.uu.se

Abstract

Background: Quality in use (QiU), a stakeholder-centered dimension of software quality encompassing effectiveness, efficiency, satisfaction, and freedom from risk, is essential in evaluating digital systems, particularly in health-related domains. Although QiU has been explored in various fields, its application within connected mental health (CMH) systems remains fragmented and understudied. Given the rapid rise in CMH technologies, ranging from mobile apps to teletherapy platforms, understanding how QiU is conceptualized, evaluated, and reported in this domain has become increasingly urgent.

Objective: This study aims to systematically map and synthesize existing research on QiU in CMH applications. It seeks to identify current trends, research gaps, evaluation methods, and the range of technologies examined concerning QiU.

Methods: A systematic mapping methodology following the guidelines by Petersen et al will be used. The process includes defining mapping questions, developing a classification scheme, and systematically searching and analyzing peer-reviewed literature from databases—Scopus, PubMed, IEEE Xplore, and ACM Digital Library. Eight mapping questions will guide the analysis, focusing on publication trends, research types, empirical evaluations, QiU characteristics and subcharacteristics, and technologies studied.

Results: As this paper presents the protocol for an ongoing mapping study, results are not yet available. The literature search and data analysis are scheduled for completion in 2026. Preliminary screening suggests variability in how QiU is defined and evaluated across CMH technologies, highlighting the need for systematic synthesis.

Conclusions: This systematic mapping study will fill a critical gap by providing a comprehensive overview of QiU research in the context of CMH. By organizing and classifying the existing literature, the study will inform future research, support the development of more user-centered CMH tools, and contribute to establishing more consistent evaluation practices in this growing field.

International Registered Report Identifier (IRRID): DERR1-10.2196/79611

(JMIR Res Protoc 2026;15:e79611) doi:10.2196/79611

KEYWORDS

connected mental health; quality in use; systematic mapping study; digital mental health evaluation; user-centered design; eHealth; mental health technologies; usability; human-centered evaluation

XSL•FO
RenderX

Introduction

Background

Quality in use (QiU) is an important aspect of software evaluation, focusing on the stakeholders' perspective and experience. The ISO/IEC 25019:2023 standard defines QiU as "extent to which the system or product when it is used in a specified context of use satisfies or exceeds stakeholders needs to achieve specified beneficial goals or outcomes" [1], which has replaced the ISO/IEC 25010:2011 standard, defines QiU as the "extent to which the system or product when it is used in a specified context of use satisfies or exceeds stakeholders needs to achieve specified beneficial goals or outcomes" [2]. Measuring QiU presents challenges due to complex standard models and limitations of customized quality models [3]. Researchers have proposed various approaches to assess QiU, including latent semantic analysis of user reviews [4], sentiment analysis [3], and systematic evaluation frameworks [5]. QiU models have been developed for specific applications such as web portals [6] and e-learning systems [5]. The concept of QiU extends beyond usability and user experience (UX; Figure 1), integrating human factors into the software engineering life cycle [2].

Figure 1. Relationship between usability, user experience, and quality in use.



To establish a clear scope for this mapping study, we explicitly define QiU as outlined in the widely used quality framework in software engineering in the ISO/IEC 25010:2011 standard and its replacement in terms of QiU, the ISO/IEC 25019:2023 standard. Therefore, QiU goes beyond traditional constructs such as usability (focused primarily on effectiveness, efficiency, and satisfaction during interaction), UX (the broader set of emotional, cognitive, and contextual responses), and freedom from risk (minimizing potential harm). While these concepts overlap, QiU represents an integrative, outcome-oriented perspective that links interaction, performance, and user-perceived value in real contexts of use.

In this study, we will map QiU dimensions, methods used to evaluate QiU, and conceptual or analytical frameworks related to QiU in connected mental health (CMH). Pure usability studies or UX-focused evaluations will be included only when they explicitly address or report on one or more QiU characteristics. To support transparency and avoid conceptual drift, we have added a visual conceptual model (Figure 1) that illustrates how QiU relates to but remains distinct from usability, UX, and safety in CMH technologies. This model anchors the boundaries of what is being mapped and ensures clear differentiation from prior CMH usability reviews.

CMH has emerged as a crucial tool for addressing the global mental health crisis, offering innovative solutions to improve access to and quality of care [7,8]. These technologies range from smartphone apps and virtual reality to artificial intelligence–powered chatbots, providing support, prevention, and treatment options [7]. CMH has shown promise in improving mental health outcomes, such as in the case of underserved populations such as lesbian, gay, bisexual, transgender, and queer individuals [9]. However, challenges remain, including ensuring efficacy and user safety and addressing the digital divide [10]. Ethical considerations such as data privacy and the need for evidence-based practices are crucial [11]. Despite these challenges, investing in CMH presents significant economic and societal benefits, especially in low-resource settings [12]. As the field evolves, focus should be placed on meaningful community engagement, personalization, and integration with traditional health care models [7,13].

As CMH technologies become increasingly embedded in mental health care delivery, evaluating these tools from a user-centered perspective is crucial. Unlike traditional clinical interventions, CMH applications are often accessed independently by users without direct supervision from mental health professionals (eg, psychologists or psychiatrists). This autonomy underscores the importance of QiU as a critical determinant of the effectiveness, engagement, and long-term adoption of CMH applications. Evaluating QiU allows developers and researchers to better understand how users experience these technologies in real-world settings, including aspects such as trust, perceived safety, ease of use, and usefulness, factors that directly influence outcomes such as adherence and therapeutic benefit [14].

Moreover, CMH interventions must be particularly sensitive to user vulnerabilities, ethical risks, and contextual appropriateness. Poorly designed systems may not only fail to provide support but can also actively cause harm, for instance, by triggering distress or mishandling sensitive data [15]. Therefore, QiU is not just a technical consideration but an ethical imperative in the design and deployment of CMH technologies.

XSL•FO
RenderX

While the broader literature on CMH has grown rapidly [8], few reviews have specifically addressed how QiU is conceptualized, evaluated, or reported in the CMH domain. Existing systematic reviews have tended to focus on clinical effectiveness [16], usability testing [14], or specific populations and conditions [15], often treating UX as a secondary concern. Furthermore, evaluations often rely on ad hoc instruments rather than established standards such as the ISO/IEC 25010, leading to inconsistent reporting and a lack of comparability across studies [1].

There is also considerable variability in how QiU subcharacteristics are interpreted and applied in CMH research. For example, usability is frequently assessed, but other critical dimensions such as safety, learnability, and risk mitigation receive little attention [1]. Additionally, there is a lack of clarity on which types of technologies, user groups, and mental health contexts have been studied in relation to QiU. Without a comprehensive overview of these patterns, it remains difficult to build a coherent body of knowledge to inform both theory and practice.

This systematic mapping study addresses these gaps by organizing and analyzing the existing literature on QiU in CMH applications. By identifying trends, challenges, and underexplored areas, it aims to support more consistent, rigorous, and user-centered evaluation practices in CMH.

Study Objectives

This systematic mapping study aims to identify, categorize, and synthesize research evidence on how QiU is evaluated in CMH applications with the purpose of informing more robust, user-centered design and evaluation practices.

Methods

Overview

This study will adopt the mapping process outlined by Petersen et al [17], which involves selecting relevant publications, developing a classification scheme, and systematically mapping the literature. The primary goal of a systematic mapping study is to organize a research field and provide a comprehensive overview of existing literature, focusing on exploring topics and categorizing the available contributions [18].

Our research team brings together international and interdisciplinary perspectives from medical informatics, human-computer interaction, information systems, and software engineering and includes researchers at different career stages (from early-career researchers to professors). This heterogeneity provides a broad understanding of digital health, sociotechnical systems, user-centered evaluation, and software quality, enriching the analysis and reducing the dominance of any single disciplinary lens. All coders have prior experience conducting systematic literature reviews [19,20], scoping reviews [21,22], or mapping studies [8,23], and several have expertise in usability testing, digital health evaluation, participatory design, or software quality.

We acknowledge that our positionalities as academic researchers working within digital health, sociotechnical, and design-oriented communities may influence how we interpret constructs related to QiU, such as usability, satisfaction, trust, or freedom from risk. While we do not have clinical roles or direct relationships with the study populations, our collective orientation toward user-centered and standard-driven evaluation may predispose us to notice certain QiU subcharacteristics more readily than others. To mitigate these potential influences, we anchor our classification scheme in the ISO/IEC 25019:2023 and ISO/IEC 25010:2011 models and use systematic intercoder reliability procedures: 2 independent reviewers will screen and code each study, discrepancies will be resolved through consensus, and a senior member will review unresolved conflicts.

Mapping Questions

Textbox 1 presents the 8 mapping questions. The mapping questions were defined to provide an overview and a structured understanding of the existing literature on QiU in CMH.

Textbox 1. Mapping questions.

- What is the temporal distribution of publications in this field? (Mapping question 1)
- What are the publication avenues of quality in use (QiU) in connected mental health (CMH) research? (Mapping question 2)
- What are the research types in which QiU is used in CMH? (Mapping question 3)
- What types of mental health problems are addressed in the QiU literature? (Mapping question 4)
- What are the types of empirical evaluation, if any (eg, case studies or interviews), of QiU in CMH research? (Mapping question 5)
- What QiU characteristics or subcharacteristics are reported in the CMH literature? (Mapping question 6)
- What are the key challenges or limitations identified in QiU in CMH? (Mapping question 7)
- What technologies (eg, mobile apps, wearable devices, or telehealth) are studied for QiU in CMH? (Mapping question 8)

Search Strategy

The review will analyze studies that discuss QiU in CMH applications. The search strategy for this systematic mapping study was developed in collaboration with Görel Sundström, a librarian from Uppsala University, and the research team (SP, SO, MH, CB, ÅC, and MRS). The population, intervention, comparator, and outcome framework was used to construct the search strategy, as shown in Table 1.

XSL•FO
RenderX

Table 1. Preliminary search strategy.

Search number	Keyword	Database search algorithm
1	"Quality in use"	TITLE-ABS-KEY ("IEC 25010" OR "ISO 25010" OR "ISO/IEC 25010" OR "ISO 25019" OR "IEC 25019" OR "ISO/IEC 25019" OR "quality-in-use" OR "quality in use" OR accessib* OR appropriateness OR beneficialness OR effectiveness OR efficiency OR flexibility OR "freedom from risk" OR "interaction capability" OR "interface design" OR learnability OR operability OR quality OR recognizability OR "risk mitigation" OR safe* OR satisfaction OR suitab* OR trust* OR usability OR usefulness OR "user experience*" OR UX OR "user need*")
2	"Connected mental health"	TITLE-ABS-KEY ("e-mental health" OR "m-mental health" OR (("mental health" OR psychotherapy OR "behavi* therapy") W/2 (connected OR digital OR mobile OR online OR smart OR tele OR video OR web)))
3	"Systems"	TITLE-ABS-KEY (app OR apps OR application* OR software OR system OR systems OR device*)
4	__[a]	Search 1 AND search 2 AND search 3

[a]Not applicable.

Scopus was selected as the first database for this research due to its extensive collection of multidisciplinary academic content, covering fields such as engineering, medicine, business, and computer science [24]. The selection process aimed to identify the articles most relevant to this mapping study's objectives.

The literature search will be conducted using multiple electronic databases: PubMed, IEEE Xplore, Scopus, and ACM Digital Library. These databases have been selected based on their relevance to the research topic and their widespread use in academic and scientific research. This systematic literature review will not involve the collection of any sensitive personal data.

Study Selection Criteria

Overview

To ensure broad inclusion consistent with the goals of a mapping study, we defined the eligibility criteria based on the characteristics of the studies rather than on a clinical evidence framework. The criteria reflect the types of technologies, participants, evaluation focuses, and publication formats relevant to understanding QiU in CMH. This is detailed in Textbox 2.

Textbox 2. Criteria for inclusion and exclusion.

Inclusion criteria

- Study focus or participants: studies involving end users (eg, patients, clinicians, or caregivers), stakeholders, or evaluators (eg, researchers or designers) who interact with or assess connected mental health (CMH) technologies; no restrictions on demographics or setting
- Technology or system type: information and communications technology tools explicitly related to CMH, including mobile apps, teletherapy platforms, chatbots, wearable-supported mental health interventions, and web-based self-help systems
- Evaluation focus: studies that report on aspects of quality in use (QiU; eg, usability, effectiveness, efficiency, safety, satisfaction, trust, or freedom from risk) in CMH applications, whether measured formally or described qualitatively
- Study type or methodology: qualitative methods, quantitative methods, mixed methods, solution proposals, evaluation research, and experience papers
- Publication type: formally published peer-reviewed journal articles and conference papers

Exclusion criteria

- Study focus or participants: studies not involving any users, stakeholders, or evaluators of CMH applications (eg, theoretical papers or technology descriptions without user focus)
- Technology or system type: interventions unrelated to mental health (eg, general wellness or physical health monitoring) or papers that discuss CMH without evaluating QiU
- Evaluation focus: studies that do not report on any user experience or QiU-related outcome or that focus solely on technical performance or clinical efficacy without user-centered evaluation
- Study type or methodology: reviews (eg, systematic reviews, scoping reviews, or meta-analyses)
- Publication type: gray literature, opinion pieces, protocols, and reviews

Types of Studies

Qualitative Studies

These encompass studies using methods such as interviews, focus groups, case studies, usability evaluations, participatory design, thematic analysis, grounded theory, and ethnographic observation that explore user or evaluator perspectives on CMH technologies.

XSL•FO
RenderX

Quantitative Studies

These encompass studies using experimental designs, surveys, user testing, task performance metrics, or statistical analysis to assess QiU aspects such as usability, efficiency, satisfaction, or effectiveness in CMH contexts.

Mixed Methods Studies

These encompass studies combining qualitative and quantitative approaches in a single research design to investigate user-centered evaluation of CMH applications. Examples include convergent, explanatory, or exploratory designs.

Language

Due to resource limitations, only studies published in English will be considered. The research team acknowledges that this may limit the inclusion of relevant studies published in other languages. This may introduce geographic and cultural bias, potentially limiting the diversity of perspectives and contexts represented in the findings.

Time Frame

No publication date limits will be applied to capture the full scope of research addressing QiU in CMH technologies.

Study Screening

Librarian Görel Sundström initiated the search process using a set of predefined keywords that reflect the study's inclusion and exclusion criteria. Duplicate records were removed using EndNote (version 21; Clarivate Analytics) according to the procedures outlined by Bramer et al [25]. The deduplicated results were then uploaded to Rayyan (Qatar Computing Research Institute) [26] for further processing. Screening will be conducted in 2 stages: first by evaluating titles and abstracts followed by full-text assessment. During the initial phase, at least 2 reviewers will independently screen the titles and abstracts, remaining blinded to each other's decisions [27]. Any discrepancies during the full-text review will be resolved through discussion, and a third, independent reviewer will be involved when consensus cannot be reached.

Data Extraction

The extraction of data from the selected studies will be guided by the mapping questions defined in this study. A standardized data extraction form will be developed collaboratively by the research team to ensure consistency. The extracted information will be used to categorize and classify the literature following the classification criteria derived from the mapping questions.

Specifically, data extraction will focus on the following dimensions:

- Bibliographic information (mapping questions 1 and 2)—includes author names, year of publication (temporal distribution), publication type (journal or conference), and country of study.
- Research characteristics (mapping questions 3 and 5)—captures the research type (qualitative, quantitative, or mixed methods), contribution type, and empirical method used (eg, usability study, interviews, surveys, or case studies).
- Technology and context (mapping questions 4 and 8)—documents the type of CMH technology studied (eg, mobile app, wearable, or telehealth platform), the mental health condition addressed (eg, anxiety, depression, or general well-being), and the user group (eg, patients, clinicians, or caregivers).
- QiU dimensions (mapping question 6)—extracts the subcharacteristics of QiU reported (eg, usability, satisfaction, effectiveness, or safety), along with any frameworks or tools used to assess them.
- Reported challenges (mapping question 7)—identifies limitations, risks, and barriers related to the evaluation or achievement of QiU in CMH systems.

All selected studies will be reviewed in full text to extract the required data. Two reviewers will carry out the initial data extraction, and a third reviewer will verify the results for accuracy and completeness. To ensure consistency between reviewers, we will assess interrater agreement using the Cohen κ coefficient [28] during the evaluation phase, along with the blind review feature in Rayyan, which provides an agreement percentage. Any discrepancies will be discussed with the third reviewer until consensus is reached. The extracted data will then be entered into a shared Microsoft Excel sheet for organization and analysis.

Quality Assessment

Although formal quality appraisal is not always required in mapping studies, QiU research varies substantially in methodological rigor, from controlled laboratory usability tests to real-world clinical or field evaluations. To contextualize these differences, we will include a descriptive methodological quality coding for all included studies. This coding is based on the study by Ouhbi et al [29] and will capture key methodological attributes such as study design, evaluation setting, participant characteristics, data collection methods, and reported limitations.

Data Analysis and Synthesis

Data from the included studies will be analyzed using quantitative and qualitative approaches guided by the 8 predefined mapping questions. Descriptive statistics will be used to address mapping questions 1 to 5 and 8, focusing on the temporal distribution of publications (mapping question 1), publication venues (mapping question 2), research types (mapping question 3), types of mental health and CMH technologies (mapping question 4), empirical evaluation methods (mapping question 5), and the technologies studied (mapping question 8). These patterns will be visualized through tables and charts generated in Microsoft Excel.

To address mapping questions 6 and 7, which relate to the reporting of QiU subcharacteristics and the challenges identified in CMH research, qualitative data will be analyzed thematically using Microsoft Excel. At least 2 reviewers will independently conduct the initial coding, and a third reviewer will independently review the data to ensure consistency.

Findings will be synthesized narratively combining quantitative trends with qualitative insights. Visual representations will be used to enhance clarity and support interpretation. The PRISMA (Preferred Reporting Items for Systematic Reviews and

XSL•FO
RenderX

Meta-Analyses) 2020 flow diagram will be used to document and present the study selection process.

Dissemination Strategy

The findings of this study will be published in a peer-reviewed scientific journal and presented at relevant academic conferences. In addition, the results will be shared through blog posts hosted by research groups affiliated with the authors.

Ethical Considerations

According to the Ethical Review Act (2003:460) established by the Swedish Ethical Review Authority, this research is not subject to ethics approval requirements.

Results

The study selection procedure will be illustrated using the PRISMA 2020 flow diagram [30], as demonstrated in Figure 2. The authors have so far conducted 2 pilot searches in collaboration with a librarian to refine the search terms and assess the initial relevance of the results. The outcomes are intended to be reported in a systematic mapping study and submitted for publication in 2026.

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

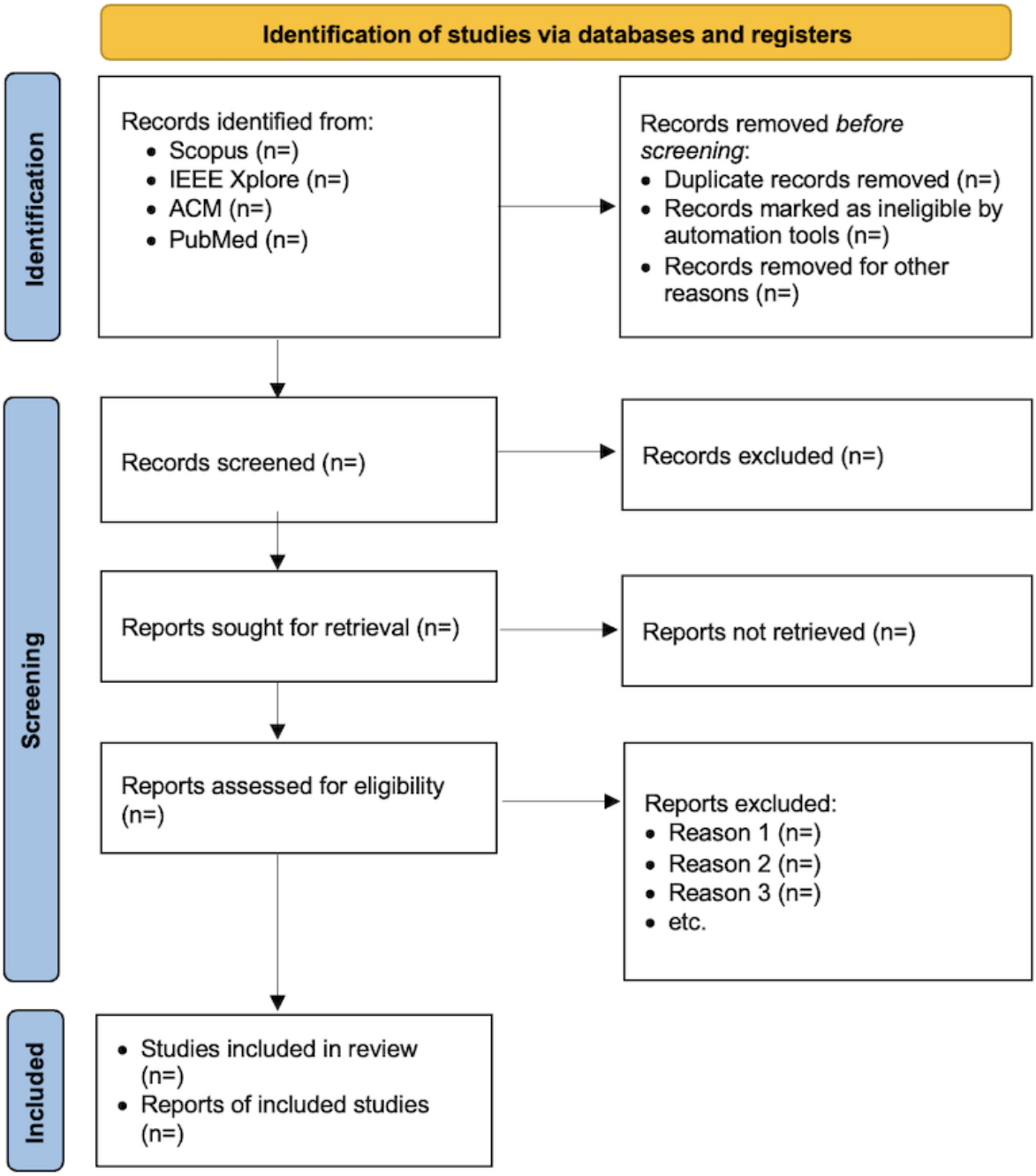

XSL•FO
RenderX

Discussion

Anticipated Principal Findings

On the basis of the scope of this mapping study, we anticipate 3 overarching patterns in how QiU is currently addressed in CMH research. First, QiU evaluations are expected to be highly heterogeneous, with studies variably emphasizing usability, UX, or safety rather than the integrated, standard-based definition provided by the ISO/IEC 25019:2023. Second, we expect to find a dominance of usability-focused assessments (eg, task success and satisfaction), with comparatively fewer studies addressing broader QiU dimensions such as context-of-use coverage, freedom from risk, or sustained goal achievement in real-world settings. Third, methodological diversity, ranging from laboratory usability tests to field deployments, is likely to reveal a fragmented evaluation landscape, suggesting the absence of a unified framework for QiU in CMH technologies. Together, these patterns will highlight both what the field currently measures and where systematic gaps limit our understanding of user-centered quality in mental health technologies.

Comparison to Prior Work

Prior reviews in CMH have primarily focused on clinical effectiveness, usability, feasibility, or ethical considerations rather than on a comprehensive assessment of QiU. For example, usability-oriented reviews highlight methodological inconsistencies, short-term evaluations, and limited attention to real-world contexts [14,16]. Broader CMH reviews have mapped technology types, use cases, and research trends [8,16], whereas others emphasize challenges related to privacy, safety, and the ethical implications of digital mental health tools [11,15]. Although these studies provide valuable insights, they typically address individual QiU-related constructs such as usability, acceptability, engagement, trust, or safety in isolation. None adopt an integrated, standard-based conceptualization anchored in the ISO/IEC 25010 or ISO/IEC 25019 standard, and few explicitly analyze how QiU subcharacteristics are explored or assessed in CMH research.

In contrast, the software engineering literature offers conceptual and methodological foundations for QiU, including theoretical models [2], application-specific frameworks [5,6], and computational approaches based on user-generated data [3,4]. However, these works rarely intersect with CMH and have not been synthesized in a mental health context. As a result, existing CMH evaluations often rely on ad hoc measures or usability-centric assessments rather than standardized QiU frameworks, contributing to fragmented reporting and limited comparability across studies [14,15]. Therefore, this mapping study is distinct from prior work as it is the first to integrate CMH and software engineering perspectives by systematically examining how QiU is conceptualized, operationalized, and evaluated in CMH technologies using the ISO/IEC 25019:2023 standard as an explicit conceptual anchor.

Strengths and Limitations

A major strength of this mapping study is the adoption of a clear conceptual boundary for QiU based on the ISO/IEC 25019:2023 and ISO/IEC 25010:2011 standards, which reduces the ambiguity that could limit comparability across CMH evaluations. Additionally, the eligibility framework was intentionally designed to be broad, enabling the inclusion of design studies, exploratory investigations, mixed methods evaluations, and experience reports from diverse mental health contexts.

However, some limitations are anticipated. First, QiU terminology is inconsistently used across the CMH literature, which may result in implicit QiU evaluations being categorized under related constructs (eg, usability or acceptability). However, this risk is mitigated because the search string includes terms related to QiU. Second, because QiU encompasses broad real-world outcomes, studies using short-term or laboratory-based assessments may contribute unevenly to the mapping. Third, like any mapping study, the methodological approach based on the established framework by Petersen et al [17] represents one of several possible ways to structure and categorize a broad field. Alternative strategies for data gathering could yield different groupings or emphases, and therefore, our findings should be interpreted as one evidence-based representation among multiple possible perspectives. Fourth, the study focuses exclusively on peer-reviewed journals and conference publications. Although this approach enhances methodological rigor and transparency, it excludes gray literature, dissertations, practitioner reports, and informal sources that may contain emerging ideas or practical insights. Fifth, while we use multiple independent reviewers during screening, coding, and verification to minimize individual bias, the interpretive nature of qualitative synthesis means that researcher perspectives may still influence decisions. We acknowledge that factors such as researcher background, expertise, and disciplinary orientation can shape categorization, and we have attempted to mitigate this through independent coding, verification by a third reviewer, and transparent reporting. Finally, the scope of this study is necessarily bounded by practical considerations, including available time, funding, and methodological feasibility. Broader analyses, such as the use of artificial intelligence–assisted synthesis techniques, the integration of sociocultural or political dimensions, multilingual considerations, or deeper engagement with stakeholder and developer perspectives, fall outside the remit of this protocol but represent promising areas for future research. By making these boundaries explicit, we aim to enhance the transparency of our approach and encourage subsequent studies to build on and extend the insights generated by this mapping study.

Future Directions

The anticipated findings of this mapping study point to several promising directions for advancing QiU research in CMH. First, future studies would benefit from real-world, longitudinal evaluation designs that capture sustained UX, safety, and effectiveness beyond short-term laboratory interactions. Second, the adoption of standardized reporting guidelines is essential to enhance reproducibility and enable meaningful cross-study comparisons. Third, expanding QiU research to underrepresented populations and contexts, such as caregivers, adolescents, culturally diverse groups, and low-resource environments, could provide a more inclusive evidence base.

XSL•FO
RenderX

Relatedly, expanding searches to include gray literature could strengthen and complement the peer-reviewed evidence base. Fourth, closer integration between design and evaluation processes, particularly by embedding QiU principles early in the development life cycle, can help ensure that CMH technologies align with users' needs, risks, and real-world contexts. Overall, by illuminating how QiU is currently conceptualized and assessed, this mapping study establishes a foundation for more rigorous, comprehensive, and contextually grounded evaluation practices in CMH research.

Conclusions

This mapping study will provide the first structured overview of how QiU has been conceptualized, operationalized, and evaluated in CMH technologies. The contribution of this work lies in establishing the conceptual boundaries of QiU, identifying the diversity of evaluation practices, and revealing where the field lacks coherence or methodological grounding. By synthesizing QiU dimensions, methods, and frameworks, the study will clarify how current research captures user-centered quality and where critical gaps remain, particularly in areas such as freedom from risk, context-of-use coverage, and real-world evaluation. The findings will support researchers and designers in selecting or developing QiU-aligned evaluation strategies and will help the CMH community move toward more consistent, standard-informed assessment practices. Ultimately, the mapping will serve as a foundation for future empirical work, methodological advancement, and the refinement of QiU models tailored to mental health technologies.

Acknowledgments

The authors would like to thank Uppsala University librarian Görel Sundström for assisting with the search strategy.

Funding

This research is funded by FORTE (Swedish Research Council for Health, Working Life, and Welfare; STY-2024/0004).

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

SP and SO contributed to conceptualization. SP wrote the initial manuscript draft. Manuscript revision was performed by SP, SO, MH, CB, ÅC, and MRS. Visualization was conducted by SP. All authors reviewed and edited the manuscript and approved the final version.

Conflicts of Interest

None declared.

References

1. ISO/IEC 25019:2023: systems and software engineering — systems and software quality requirements and evaluation (SQuaRE) — quality-in-use model. International Organization for Standardization. URL: https://www.iso.org/standard/78177.html [accessed 2025-02-25]
2. Bevan N, Azuma M. Quality in use: incorporating human factors into the software engineering lifecycle. In: Proceedings of the 1997 IEEE International Symposium on Software Engineering Standards. 1997 Presented at: SESS '97; June 1-6, 1997; Walnut Creek, CA URL: https://ieeexplore.ieee.org/document/595963 [doi: 10.1109/sess.1997.595963]
3. Atoum I, Bong CH. A framework to predict software "quality in use" from software reviews. In: Proceedings of the 1st International Conference on Advanced Data and Information Engineering. https://link.springer.com/chapter/10.1007/978-981-4585-18-7_48; 2013 Presented at: DaEng '13; December 16-18, 2013; Kuala Lumpur, Malaysia. [doi: 10.1007/978-981-4585-18-7_48]
4. Wendy S, How BC, Atoum I. Using latent semantic analysis to identify quality in use (QU) indicators from user reviews. arXiv Preprint posted online March 25, 2015 [FREE Full text] [doi: 10.48550/arXiv.1503.07294]
5. Covella GJ, Olsina LA. Assessing quality in use in a consistent way. In: Proceedings of the 6th international conference on Web engineering. 2006 Presented at: ICWE '06; July 11-14, 2006; Palo Alto, CA URL: https://dl.acm.org/doi/10.1145/1145581.1145583 [doi: 10.1145/1145581.1145583]
6. Herrera M, Moraga M, Caballero I, Calero C. Quality in use model for web portals (QiUWeP). In: Proceedings of the 10th International Conference on Current Trends in Web Engineering. 2010 Presented at: ICWE '10; July 5-6, 2010; Vienna, Austria URL: https://link.springer.com/chapter/10.1007/978-3-642-16985-4_9 [doi: 10.1007/978-3-642-16985-4_9]
7. Bond RR, Mulvenna MD, Potts C, O'Neill S, Ennis E, Torous J. Digital transformation of mental health services. Npj Ment Health Res 2023 Aug 22;2(1):13 [FREE Full text] [doi: 10.1038/s44184-023-00033-y] [Medline: 38609479]

XSL•FO
RenderX

8. Drissi N, Ouhbi S, Janati Idrissi MA, Fernandez-Luque L, Ghogho M. Connected mental health: systematic mapping study. J Med Internet Res 2020 Aug 28;22(8):e19950 [FREE Full text] [doi: 10.2196/19950] [Medline: 32857055]
9. Fowler JA, Buckley L, Muir M, Viskovich S, Paradisis C, Zanganeh P, et al. Digital mental health interventions: a narrative review of what is important from the perspective of LGBTQIA+ people. J Clin Psychol 2023 Nov 02;79(11):2685-2713. [doi: 10.1002/jclp.23571] [Medline: 37528773]
10. Ogugua JO, Okongwu CC, Akomolafe OO, Anyanwu EC, Daraojimba OD. Mental health and digital technology: a public health review of current trends and responses. Int Med Sci Res J 2024 Feb 02;4(2):108-125. [doi: 10.51594/imsrj.v4i2.754]
11. Wies B, Landers C, Ienca M. Digital mental health for young people: a scoping review of ethical promises and challenges. Front Digit Health 2021 Sep 6;3:697072 [FREE Full text] [doi: 10.3389/fdgth.2021.697072] [Medline: 34713173]
12. Faria M, Zin ST, Chestnov R, Novak AM, Lev-Ari S, Snyder M. Mental health for all: the case for investing in digital mental health to improve global outcomes, access, and innovation in low-resource settings. J Clin Med 2023 Oct 25;12(21):6735 [FREE Full text] [doi: 10.3390/jcm12216735] [Medline: 37959201]
13. Prescott J. Digital technology to support mental health: a brief introduction to what it is and why it is important? Ment Health Soc Incl 2022 Apr 06;26(2):103-106. [doi: 10.1108/mhsi-02-2022-0010]
14. Inal Y, Wake JD, Guribye F, Nordgreen T. Usability evaluations of mobile mental health technologies: systematic review. J Med Internet Res 2020 Jan 06;22(1):e15337 [FREE Full text] [doi: 10.2196/15337] [Medline: 31904579]
15. Catania J, Beaver S, Kamath RS, Worthington E, Lu M, Gandhi H, et al. Evaluation of digital mental health technologies in the United States: systematic literature review and framework synthesis. JMIR Ment Health 2024 Aug 30;11:e57401 [FREE Full text] [doi: 10.2196/57401] [Medline: 39213023]
16. Zhong S, Yang X, Pan Z, Fan Y, Chen Y, Yu X, et al. The usability, feasibility, acceptability, and efficacy of digital mental health services in the COVID-19 pandemic: scoping review, systematic review, and meta-analysis. JMIR Public Health Surveill 2023 Feb 13;9:e43730 [FREE Full text] [doi: 10.2196/43730] [Medline: 36634261]
17. Petersen K, Feldt R, Mujtaba S, Mattsson M. Systematic mapping studies in software engineering. In: Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering. 2008 Presented at: EASE '08; June 26-27, 2008; Rome, Italy URL: https://dl.acm.org/doi/10.5555/2227115.2227123 [doi: 10.14236/ewic/ease2008.8]
18. Petersen K, Vakkalanka S, Kuzniarz L. Guidelines for conducting systematic mapping studies in software engineering: an update. Inf Softw Technol 2015 Aug;64:1-18. [doi: 10.1016/j.infsof.2015.03.007]
19. Drissi N, Ouhbi S, Marques G, de la Torre Díez I, Ghogho M, Janati Idrissi MA. A systematic literature review on e-mental health solutions to assist health care workers during COVID-19. Telemed J E Health 2021 Jun 22;27(6):594-602. [doi: 10.1089/tmj.2020.0287] [Medline: 32970532]
20. Alneyadi M, Drissi N, Almeqbaali M, Ouhbi S. Biofeedback-based connected mental health interventions for anxiety: systematic literature review. JMIR Mhealth Uhealth 2021 Apr 22;9(4):e26038 [FREE Full text] [doi: 10.2196/26038] [Medline: 33792548]
21. Premanandan S, Ahmad A, Cajander Å, Ågerfalk P, van Gemert-Pijnen L. Designing and evaluating IT applications for informal caregivers: scoping review. J Med Internet Res 2024 Oct 23;26:e57393 [FREE Full text] [doi: 10.2196/57393] [Medline: 39442167]
22. Meier-Diedrich E, Lyckblad C, Davidge G, Hägglund M, Kharko A, McMillan B, et al. Impact of patient online record access on documentation: scoping review. J Med Internet Res 2025 Feb 20;27:e64762 [FREE Full text] [doi: 10.2196/64762] [Medline: 39977853]
23. Karampela M, Isomursu M, Porat T, Maramis C, Mountford N, Giunti G, et al. The extent and coverage of current knowledge of connected health: systematic mapping study. J Med Internet Res 2019 Sep 25;21(9):e14394 [FREE Full text] [doi: 10.2196/14394] [Medline: 31573915]
24. Montoya FG, Alcayde A, Baños R, Manzano-Agugliaro F. A fast method for identifying worldwide scientific collaborations using the Scopus database. Telemat Inform 2018 Apr;35(1):168-185. [doi: 10.1016/j.tele.2017.10.010]
25. Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Libr Assoc 2016 Jul;104(3):240-243 [FREE Full text] [doi: 10.3163/1536-5050.104.3.014] [Medline: 27366130]
26. AI powered tool for systematic literature reviews. Rayyan. URL: https://www.rayyan.ai/ [accessed 2023-05-11]
27. Harrison H, Griffin SJ, Kuhn I, Usher-Smith JA. Software tools to support title and abstract screening for systematic reviews in healthcare: an evaluation. BMC Med Res Methodol 2020 Jan 13;20(1):7 [FREE Full text] [doi: 10.1186/s12874-020-0897-3] [Medline: 31931747]
28. Gisev N, Bell JS, Chen TF. Interrater agreement and interrater reliability: key concepts, approaches, and applications. Res Social Adm Pharm 2013;9(3):330-338. [doi: 10.1016/j.sapharm.2012.04.004] [Medline: 22695215]
29. Ouhbi S, Idri A, Fernández-Alemán JL, Toval A. Requirements engineering education: a systematic mapping study. Requir Eng 2013 Nov 21;20(2):119-138. [doi: 10.1007/S00766-013-0192-5]
30. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]

XSL•FO
RenderX

Abbreviations

CMH: connected mental health
IEC: International Electrotechnical Commission
ISO: International Organization for Standardization
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QiU: quality in use
UX: user experience

Edited by J Sarvestan; submitted 24.Jun.2025; peer-reviewed by M Pulier; comments to author 28.Oct.2025; revised version received 12.Dec.2025; accepted 17.Dec.2025; published 20.Jan.2026.

Please cite as:
Premanandan S, Ouhbi S, Stadin MR, Blease C, Cajander Å, Hägglund M
Quality in Use in Connected Mental Health: Protocol for a Systematic Mapping Study
JMIR Res Protoc 2026;15:e79611
URL: https://www.researchprotocols.org/2026/1/e79611
doi:10.2196/79611
PMID:

©Shweta Premanandan, Sofia Ouhbi, Magdalena Ramstedt Stadin, Charlotte Blease, Åsa Cajander, Maria Hägglund. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effectiveness and Theoretical Foundations of mHealth Apps for Physical Activity, Healthy Eating, and Weight Loss: Protocol for a Systematic Review and Meta-Analysis

Victoria Riccalton[1], BA, MSc; Madison Milne-Ives[1,2], BAS, MSc, PhD; Rosiered Brownson-Smith[1], BSc, MSc, PhD; Ananya Ananthakrishnan[1], BTech, BA, MSc; Cen Cong[1], BEng, MA, PhD; Emily Oliver[3], Bsc, PhD; Edward Meinert[1,4,5], MA, MSc, MBA, MPA, MPH, PhD, CEng, FBCS, EUR, ING

[1]Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom

[2]Centre for Health Technology, University of Plymouth, Plymouth, United Kingdom

[3]Population Health Sciences Institute, Newcastle University, Newcastle upon Tyne, United Kingdom

[4]School of Public Health, Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

[5]Gnosis Health Limited, Newcastle upon Tyne, England, United Kingdom

Corresponding Author:
Edward Meinert, MA, MSc, MBA, MPA, MPH, PhD, CEng, FBCS, EUR, ING
School of Public Health
Department of Primary Care and Public Health
Imperial College London
Exhibition Rd, South Kensington
London, SW7 2AZ
United Kingdom
Phone: 44 191 208 6000
Email: e.meinert14@imperial.ac.uk

Abstract

Background: Obesity is a significant global public health concern. Primary prevention and health promotion to encourage positive health behavior to address obesity could be delivered via mobile health (mHealth), but evidence of apps improving health outcomes over sufficient time frames to be clinically meaningful is limited. mHealth interventions for physical activity, healthy eating, and weight loss typically prioritize intention as the primary driver of behavior. This may limit their impact, as intention does not consistently translate into behavior.

Objective: This review updates a previous systematic review on the effectiveness of mobile apps for health behavior change while narrowing the scope to weight management interventions to enable a more focused analysis. Our primary objective is to investigate the effectiveness of mHealth apps in improving health behaviors with respect to physical activity and healthy eating and to explore the inclusion of behavioral theories and behavior change techniques and the evidence for their effectiveness.

Methods: This protocol follows the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols) checklist, and the review will be structured using the PRISMA 2020 statement. Nine databases (PubMed, EMBASE, CINAHL, APA PsycINFO, Cochrane Library, SPORTDiscus, SCOPUS, Web of Science, and Science Direct) will be searched for studies reporting evaluation of the impact of mHealth interventions on weight loss, healthy eating, or physical activity outcomes. EndNote 21 software will be used for deduplication and initial screening, followed by manual title and abstract screening, and then full-text screening by 2 independent reviewers. Data regarding the studies, intervention characteristics, their theoretical basis (eg, use of behavior change frameworks such as the COM-B [Capability, Opportunity, Motivation-Behavior] model or the social cognitive theory), evaluation methods, and outcomes will be extracted into a predetermined form. A meta-analysis will be conducted on eligible studies (reporting control group comparisons) to synthesize evidence of their effectiveness, and the remaining quantitative data will be descriptively analyzed.

Results: The review is expected to start in December 2025 and to be submitted for publication by the end of 2026.

Conclusions: This review will synthesize evidence on the theoretical basis underpinning mHealth interventions for enhancing physical activity, healthy eating, and weight loss and generate new insights into how particular behavior change techniques can

XSL•FO
RenderX

best support intended outcomes. This will help guide the development of more impactful mobile-based interventions to support healthy behaviors that are better able to reduce the risk factors for chronic health conditions.

Trial Registration: PROSPERO CRD42024602819; https://www.crd.york.ac.uk/PROSPERO/view/CRD42024602819

International Registered Report Identifier (IRRID): PRR1-10.2196/72664

(JMIR Res Protoc 2026;15:e72664) doi:10.2196/72664

KEYWORDS

digital health; mHealth; behavioral change; review; intention

Introduction

Background

The global rise of noncommunicable diseases is increasing pressure on health care systems and economies, alongside the impact on individuals' quality and length of life [1,2]. Obesity is a risk factor in many noncommunicable diseases, and weight management through healthy behaviors, including healthy eating and physical activity [3,4], is key to prevention [5-7]. Digital health interventions (DHIs) offer opportunities to support greater numbers of at-risk individuals than traditional, offline weight management interventions [8], but evidence of DHI effectiveness, particularly over sufficient durations to be clinically meaningful, is limited [9,10]. Key barriers to digital health impact include lack of engagement [11-16] and lack of digital access and/or literacy, which can also widen health inequities [8]. Increasing evidence suggests that theory-based mobile-based interventions for health behavior changes could enhance their effectiveness in various aspects such as user engagement [17], self-efficacy [18], patient self-management [19], and attitudes toward physical activity [20]. The influence of specific theories and behavior change techniques (BCTs) on the effectiveness of mobile-based interventions is still unclear. This review will provide an update to previous work [21] exploring the theoretical basis and BCTs (or "active ingredients"[22]) included in the interventions, as well as their effectiveness through a meta-analysis. Given their important roles in chronic disease prevention, we will focus on physical activity, healthy eating, and weight loss [23,24].

Incorporating behavioral theory into the development of DHIs can improve their impact on health behaviors [25], but many DHIs are developed without a theoretical basis [26,27]. Potential reasons for this lack include theory complexity, with some designers citing that they are inaccessible to nonexperts of behavior change theory [26] and the diversity of behavioral theories available. For example, Michie et al's (2011) [28] COM-B (Capability, Opportunity, Motivation-Behavior) model, which posits that behavior change requires a combination of considerations of individuals' capability, opportunity, and motivation, is associated with the taxonomy of BCTs [22]. BCTs are increasingly being deliberately incorporated in digital behavior change interventions [29-31], but the diversity of behavioral theories can lead to confusion about which BCTs may be most useful to include in different health contexts [32]. For example, the social cognitive theory, commonly used as a theoretical basis in DHIs [21], states that goal setting is key to behavior change [33], which would suggest prioritizing the goal-setting BCT within interventions [34]. However, there is substantial evidence that people's intentions do not necessarily align with their actual behavior, particularly in interventions targeting long-term healthy lifestyle behaviors such as physical activity and healthy eating [35-38].

Behavioral theories such as the dual process theory [39] have been applied to attempt to explain this gap, with researchers positing that conscious and automatic processes (ie, habits) must be activated for sustained behavior change. Evidence suggests that web-based physical activity interventions targeting habit formation effectively support sustained behavior change [40], supporting this theory in the context of mobile health (mHealth) intervention development. Other theories that emphasize the role of cognition (eg, the cognitive continuum theory, a theory of decision-making suggesting that decision-making involves both intuitive, ie, automatic and analytical thinking, influenced by task complexity and familiarity [41] and the cognitive load theory, which posits that the learning is impacted by the mental effort required) are also potentially applicable to the behavior change intervention design [42]. Interventions aiming to minimize the cognitive load on users may be more successful at facilitating behavior change by reducing barriers to engagement. Digital-specific models such as the Technology Acceptance Model and the Unified Theory of Acceptance and Use of Technology emphasize the importance of the solution's ease of use and perceived usefulness in facilitating engagement [43,44]. While hardly exhaustive, and drawn predominantly from the cognitive rather than wider social domains of behavioral sciences, these examples highlight the wide variety of theories that could be used to support DHI design for sustainable behavior change.

The effectiveness of mobile app–based interventions in promoting health behaviors has been explored in systematic reviews and meta-analyses, with mixed findings. For example, significant effectiveness was not found in mobile app–based interventions for physical activity in healthy adults aged 18 to 35 years [45]. In contrast, recent meta-analysis studies revealed that mobile apps can lead to significant improvements in behavioral and health outcomes in various population groups, such as increasing total physical activity and reducing BMI in children and adolescents [46], as well as various weight-related outcomes in adults with overweight and obesity (reduced weight, BMI, waist circumference, fat mass, and diastolic blood pressure) [47]. An umbrella review on mHealth interventions examined the impact of mHealth across various health outcomes, some of which involved weight loss and physical activity [48]. There is a lack of exploration of the behavioral frameworks underlying the interventions, and although the use of specific BCTs in mHealth interventions was examined in one review

XSL•FO
RenderX

[49], it focused on frequency and effectiveness in interventions that did not use mobile technologies. This highlights the need for further research to explore the association between these outcomes and the use of specific BCTs. Knowing the effects of BCTs on behavior change outcomes could provide valuable insights to guide the development of more effective mHealth apps. The impact of mobile-based interventions on healthy eating was also not explored in previous reviews, leaving a gap in the literature. A healthy diet has a significant positive impact on weight loss [50]. Exploring these aspects together may yield a better understanding of the effectiveness of mobile app interventions in promoting health behavior change.

We previously conducted a systematic review in 2019 (literature searched between 2014-2019), aiming to evaluate the effectiveness of mHealth apps in improving physical activity, diet, drug and alcohol use, and mental health and to explore the use of BCTs in these apps and found limited evidence supporting their effectiveness [21]. However, while the broad scope provided a comprehensive overview, its variety of outcomes meant we could not conduct a meta-analysis, which was a key limitation of the review. In addition, app development accelerated exponentially during and after COVID-19, which occurred after our review was conducted [51,52]. This highlights a need to update the previous review with a more specific focus to enable a meta-analysis and the inclusion of more recent literature to synthesize the developments in the field.

Objectives

The primary aim of this review is to update a previous systematic review that evaluated the effectiveness of mobile apps for health behavior change [21], with the addition of a meta-analysis to quantitatively synthesize effectiveness. To ensure a more focused approach, we will narrow the scope only to include studies on weight management interventions. Secondary objectives include exploring behavioral theories and BCTs incorporated in the apps and synthesizing the apps' feasibility, usability, and user experience. Three main research questions will guide this review. First, what is the effectiveness of mHealth apps at improving and maintaining health behavior change with respect to healthy eating and physical activity? Second, what behavioral theories and BCTs are being incorporated in mHealth app design to support engagement with these behaviors and which are the most effective at promoting behavior change? Third, what is the feasibility, usability, and user experience associated with these apps? By providing a comprehensive synthesis of these aspects, we aim to inform the development of mHealth apps for weight management, particularly concerning behavior change theories, models, and frameworks.

Methods

Study Design

This protocol has been prepared in line with the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols) guidelines (see Multimedia Appendix 1) and the search strategy structured with reference to the PICOS (Population, Intervention, Comparator, Outcome, and Study) framework [53]. This review was registered on PROSPERO (CRD42024602819). A systematic review design was chosen because the primary objective of the review will be to synthesize the effectiveness of mHealth apps. Although the secondary objectives are more exploratory, they will primarily be used to contextualize the evidence and synthesized using systematic methods [54].

Search Strategy

We will search 9 databases: PubMed, Embase, CINAHL, APA PsycINFO, Cochrane Library, SPORTDiscus, SCOPUS, Web of Science, and Science Direct. These have been chosen, as they comprehensively cover the fields of medicine, sport/exercise, nutrition, and computing/artificial intelligence.

The eligibility criteria and search terms are shown in Table 1. A literature review identified keywords and MeSH (Medical Subject Headings) terms relating to evaluations of mHealth apps targeting weight loss, healthy eating, and physical activity outcomes. Keywords and MeSH terms were expanded through database searching and grouped into themes. The 3 themes resulted in the following search structure: mHealth (MeSH OR keywords) AND weight-related behaviors (MeSH OR keywords) AND evaluation (keywords).

The full search string used for an initial search of PubMed and the number of references returned are included in Multimedia Appendix 2.

XSL•FO
RenderX

Table 1. Eligibility criteria, search themes, MeSHa terms, and keywords structured according to the PICOS (Population, Intervention, Comparator, Outcome, and Study) framework.

Element	Eligibility criteria	MeSH terms	Keywords
Population	The population will include people of any age (children, adolescents, or adults).	N/A[b]	N/A
Intervention	mHealth[c] interventions targeting physical activity, sedentary and healthy eating behaviors, and weight loss will be included.	Cell Phone OR Telemedicine OR Mobile Applications	Smartphone OR "mobile phone" OR "mHealth" OR "mobile health" OR App OR apps OR "mobile app*" OR "smartphone app*" OR "digital health"
Comparator	Not required for inclusion in the systematic review Studies with control group comparisons will form a subset for meta-analysis	N/A	N/A
Outcomes	Outcomes include weight loss, healthy eating, physical activity, health behavior change, and engagement.	Health Behavior OR Exercise OR Weight Loss OR Obesity (diet therapy, prevention & control, rehabilitation, therapy) OR Nutrition Therapy OR Diet OR Sedentary Behavior	overweight OR obesity OR obese OR "weight loss" OR "weight management" OR "weight reduction" OR "weight maintenance" OR "maintaining weight" OR "ideal weight" OR "body fat" OR "waist circumference" OR "body mass index" OR BMI OR "healthy eating" OR diet OR nutrition OR calori* OR "meal plan" OR "eating behav*" OR cook* OR food* OR fruit* OR veg OR vegetable* OR "physical activit*" OR "moderate to vigorous" OR "active living" OR exercise OR fitness OR "active transport" OR sedentary OR physical inactivity OR walk* OR "behavior change" OR "behavior change" OR Exercise ADJ3 (increase or start or maintain*)) OR "physical activity" OR "Weight loss" OR "healthy weight" OR "five a day" OR "diet" OR "nutrition" OR ((Maintenance OR maintain* OR achiev* Or retain*) ADJ4 (weight goal OR "weight loss" OR "goal weight" OR BMI))
Study types	Studies in English, published since 2019 evaluating the impact of mHealth interventions supporting intention or habit on the specified outcomes will be included. Protocols, conference abstracts, reviews, and other studies that do not contain an evaluative component will be excluded.	N/A	effective* OR efficacy OR impact* OR evaluat* OR study OR test OR trial OR Feasibility OR usability OR "outcome*" OR acceptability OR adherence OR "effectiv*" OR "adoption" OR "assess*"

[a]MeSH: Medical Subject Headings.

[b]N/A: not applicable.

[c]mHealth: mobile health.

Inclusion Criteria

All types of studies evaluating the impact of an mHealth intervention on weight-related, healthy eating, or physical activity outcomes will be eligible for inclusion. Given that digital health is a rapidly advancing field, particularly after COVID-19 [55,56], and this review is related to the current state of mHealth interventions, it will include studies published since 2019. This time frame allows for a comprehensive update to Milne-Ives et al's [21] systematic review, which included papers from 2014 to 2019. The subset of studies reporting control group comparisons will additionally be eligible for inclusion in the meta-analysis.

Exclusion Criteria

References that describe the development or implementation of interventions without reporting the impact on weight loss, healthy eating, or physical activity outcomes will be excluded, as this study focuses on the interventions' impact on outcomes. Studies of apps that were designed for use by health care professionals or those that focus on mobile phones or other digital technology but not apps (eg, email, SMS text messaging interventions) will also be excluded. Studies not reported in English will be excluded due to the language capabilities of the research team. Where full text is unavailable, references will be excluded. Reviews, conference abstracts, surveys, and case studies will also be excluded, as our focus is on original, evaluative research concerning digital behavior change interventions.

XSL•FO
RenderX

Screening and Study Selection

References will be compiled and duplicates removed in EndNote 21 software. A series of keyword search passes will be conducted to exclude ineligible references. The remaining references will be uploaded to Rayyan to facilitate manual screening of titles and abstracts by 2 independent reviewers, after which full-text screening by 2 independent reviewers will be completed to arrive at a final set of included references. Disagreements between reviewers will be resolved by discussion and consultation with a third reviewer if required. EndNote-automated screening passes will be recorded for reference, and the outcomes of each stage of the process will be recorded in a PRISMA flow diagram; all details will be available in the review report.

Data Extraction

Table 2 shows the study data to be extracted for analysis. As with the screening, 2 reviewers will extract the data independently, with disagreements discussed until consensus and a third reviewer involved if necessary.

Table 2. Study data to be extracted.

Information area	Data to be extracted
Study characteristics	• Publication year • Country of study • Study type • Study duration, including intervention duration and follow-up timepoint or timepoints • Population demographics • Sample size • Comparator
Intervention characteristics	• Target health behavior or behaviors • Intervention focus (ie, healthy eating, physical activity, or both) • Whether BCTsa explicitly described in paper (yes/no) • BCTs employed (list - coded based on behavior change ontology [57]) • Behavioral theory or theories' basis • Name of the app • Developers • Platform • Design features (eg, feedback provision, notifications, monitoring)
Evaluation	• What outcomes were reported • Impact on health outcomes (eg, change in BMI, body weight, body fat, waist circumference, blood pressure, cholesterol levels)b • Impact on behavioral outcomes (eg, change in calorie intake, intake of fruit/vegetables, intake of processed foods, step count, metabolic equivalent of tasks, minutes of exercise per week, minutes of physical activity by level, sedentary periods)b • Engagement outcomes (eg, engagement self-report scales, app usage data) • Feasibility • Usability • User experience/satisfaction • Other outcomes • Stated strengths and limitations

[a]BCT: behavior change technique.

[b]Quantitative outcomes potentially eligible for meta-analysis if sufficient data are available.

Quality and Risk of Bias Assessments

Quality and risk of bias will be assessed by 2 reviewers, any disagreements discussed, and a third reviewer consulted if a consensus cannot be reached. Randomized controlled trials will be assessed using the revised Cochrane Risk of Bias tool [58]. Nonrandomized studies will be assessed using the Newcastle-Ottawa Scale [59].

Data Analysis and Synthesis

Based on initial literature searches, it is anticipated that a subset of the studies will be suitable for inclusion in a meta-analysis and that these studies will be heterogeneous in their outcome measures. A random-effects meta-analysis model for overall effect size, using Hedge g to calculate the standardized mean difference (with a 95% CI), and τ^2 to measure heterogeneity will be employed. SPSS software (version 29; IBM Corp) will be used for the statistical analysis. The meta-analysis will aim to synthesize the effectiveness of the included apps, and separate analyses will be conducted for the different outcome measures identified (eg, change in BMI or body weight, change in waist circumference, change in calorie intake). Missing data will be handled on a case-by-case basis, following the Cochrane Handbook for Systematic Reviews of Interventions [60]. Where possible, original investigators will be contacted to request missing data (particularly for missing studies, to mitigate publication bias). Based on the data found, assumptions as to whether they are missing at random will be made and clearly outlined, and risk of bias assessments will take this into account. Additionally, sensitivity analyses will be performed based on the data found to assess the impact of reasonable changes in the assumptions made. A descriptive analysis will be conducted to

XSL•FO
RenderX

summarize the quantitative results of studies not suitable for inclusion in the meta-analysis. Data about the use of BCTs and the feasibility, usability, and user experience of the apps will be summarized following Popay et al's (2006) [61] narrative synthesis guidance, chosen as it would enable us to integrate a variety of qualitative and quantitative research findings. The strengths and limitations of the identified studies and areas for future research and development will be discussed.

Results

Work on this review has not yet commenced as of October 3, 2025. Data collection will be completed by December 2025, and findings will be submitted for publication by the end of 2026. Data used for analysis in this review will be deposited in the Open Science Framework [59] upon completion of the study for transparency.

Discussion

Summary of the Anticipated Findings

This review will provide a synthesis of the effectiveness of mobile apps for health behavior change in terms of healthy eating and physical activity. We will also examine the apps' feasibility, usability, and user experience, and explore the behavioral theories and BCTs used in developing them. By synthesizing all this information in one place, we hope to explore how the use of theory in DHI development might best contribute to their effectiveness, particularly in the long term. The findings of this review will update existing syntheses of similar apps, providing a focused summary of apps for physical activity and diet change. Although the mixed findings in previous reviews preclude a hypothesis, we anticipate that apps incorporating specific BCTs may show increased evidence of effectiveness.

Dissemination and Stakeholder Engagement

The review will be disseminated via publication in a peer-reviewed journal and via an online newsletter distributed by the research group. We will also distribute a summary of the research through Newcastle University's networks, including the Behavioral Sciences and Psychology Theme.

Comparison to Previous Work and Future Directions

By synthesizing the effectiveness of DHIs for behavior change through a meta-analysis and contextualizing this via an exploration of user experience, usability, feasibility, and the use of BCT, we aim to enable the development of apps grounded in theory that enable long-term engagement and sustained change. Although previous research has synthesized the effectiveness of similar apps [46,47], none have conducted a comprehensive assessment of the qualitative and quantitative outcomes as well as the theoretical grounding, and few have focused on diet in addition to physical activity. Our previous review attempted to analyze this comprehensively, its broad scope meant a meta-analysis was infeasible, and an update is needed as this was conducted in 2019. We expect that this review will provide better insights into such apps, particularly adding to the knowledge of their impact on healthy diet, and their use of BCTs to enable these changes. This review will highlight for health care professionals and app developers the characteristics of mHealth interventions that might lead to sustained change, inform future development, and allow individuals to choose effective DHIs to use and recommend, if any are identified.

Limitations

Due to the language capability of the research team, this study will not include studies in languages other than English, which may miss relevant studies in other languages. Studies published before 2019 will also be excluded to avoid duplicated efforts in previous studies and to provide a more recent update on the evidence. Conference abstracts and grey literature will be excluded to focus the review on studies that are likely to be higher quality, but there is a risk that relevant information may be missed.

Conclusions

This review will contribute to the body of knowledge around the theoretical foundations for mHealth interventions, particularly those related to physical activity, healthy eating, and weight loss, by synthesizing recent mHealth evaluations. To address the gap in previous reviews, we will explore the associations between these BCTs and the effectiveness of mHealth interventions through a meta-analysis. This will offer new insights about which technique is more effective for specific outcomes. Such understanding can contribute to the design of more impactful mHealth interventions in the future.

Funding

This manuscript is independent research funded by Parkinson's UK (grant H-2101). The research was also supported by the National Institute for Health and Care Research (NIHR) Newcastle Biomedical Research Centre (BRC) based at the Newcastle-upon-Tyne Hospitals NHS Foundation Trust, Newcastle University, and the Cumbria, Northumberland, and Tyne and Wear NHS Foundation Trust. The views expressed in this publication are those of the authors and not necessarily those of NIHR or BRC or any of the authors' affiliated universities. The funding body was not involved in the study design, data collection or analysis, or the writing and decision to submit the paper for publication.

Authors' Contributions

Conceptualization: VR
Funding acquisition: EM

XSL•FO
RenderX

Methodology: VR
Supervision: EM (lead), EO (supporting)
Writing – original draft: VR
Writing – review & editing: MM-I (lead), RB-S (supporting), AA (supporting), CC (supporting), EO (supporting), EM (supporting)

Conflicts of Interest

EM is the Editor-in-Chief of JMIRx Med. EM is the Co-Founder and Chief Executive Officer of Gnosis Health Limited, a company specializing in the design and development of digital tools for chronic disease management. Newcastle University and the University of Plymouth are shareholders in Gnosis Health Limited.

Multimedia Appendix 1
PRISMA-P checklist.
[PDF File (Adobe PDF File), 214 KB - resprot_v15i1e72664_app1.pdf]

Multimedia Appendix 2
Sample search on PubMed.
[DOCX File , 17 KB - resprot_v15i1e72664_app2.docx]

References

1. Santos AC, Willumsen J, Meheus F, Ilbawi A, Bull FC. The cost of inaction on physical inactivity to public health-care systems: a population-attributable fraction analysis. Lancet Glob Health 2023 Jan;11(1):e32-e39 [FREE Full text] [doi: 10.1016/S2214-109X(22)00464-8] [Medline: 36480931]
2. Muka T, Imo D, Jaspers L, Colpani V, Chaker L, van der Lee SJ, et al. The global impact of non-communicable diseases on healthcare spending and national income: a systematic review. Eur J Epidemiol 2015 Apr;30(4):251-277. [doi: 10.1007/s10654-014-9984-2] [Medline: 25595318]
3. Donnelly JE, Blair SN, Jakicic JM, Manore MM, Rankin JW, Smith BK, American College of Sports Medicine. American College of Sports Medicine Position Stand. Appropriate physical activity intervention strategies for weight loss and prevention of weight regain for adults. Med Sci Sports Exerc 2009 Feb;41(2):459-471. [doi: 10.1249/MSS.0b013e3181949333] [Medline: 19127177]
4. Johnston BC, Kanters S, Bandayrel K, Wu P, Naji F, Siemieniuk RA, et al. Comparison of weight loss among named diet programs in overweight and obese adults: a meta-analysis. JAMA 2014 Sep 03;312(9):923-933. [doi: 10.1001/jama.2014.10397] [Medline: 25182101]
5. Budreviciute A, Damiati S, Sabir DK, Onder K, Schuller-Goetzburg P, Plakys G, et al. Management and prevention strategies for non-communicable diseases (NCDs) and their risk factors. Front Public Health 2020;8:574111 [FREE Full text] [doi: 10.3389/fpubh.2020.574111] [Medline: 33324597]
6. Noncommunicable diseases: risk factors and conditions. World Health Organization. URL: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/ncd-risk-factors [accessed 2024-10-04]
7. Pi-Sunyer FX. The medical risks of obesity. Obes Surg 2002 Apr 01;12(2):6-11. [doi: 10.1381/096089202762552601]
8. Milne-Ives M, Burns L, Swancutt D, Calitri R, Ananthakrishnan A, Davis H, et al. The effectiveness and usability of online, group-based interventions for people with severe obesity: a systematic review and meta-analysis. Int J Obes (Lond) 2025 Apr;49(4):564-577. [doi: 10.1038/s41366-024-01669-2] [Medline: 39558074]
9. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med 2020;3:4 [FREE Full text] [doi: 10.1038/s41746-019-0206-x] [Medline: 31970289]
10. Covolo L, Ceretti E, Moneda M, Castaldi S, Gelatti U. Does evidence support the use of mobile phone apps as a driver for promoting healthy lifestyles from a public health perspective? A systematic review of Randomized Control Trials. Patient Educ Couns 2017 Dec;100(12):2231-2243. [doi: 10.1016/j.pec.2017.07.032] [Medline: 28855063]
11. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283 [FREE Full text] [doi: 10.2196/20283] [Medline: 32990635]
12. Torous J, Lipschitz J, Ng M, Firth J. Dropout rates in clinical trials of smartphone apps for depressive symptoms: a systematic review and meta-analysis. J Affect Disord 2020 Feb 15;263:413-419. [doi: 10.1016/j.jad.2019.11.167] [Medline: 31969272]
13. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]
14. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. Acad Emerg Med 2015 Jun;22(6):754-756 [FREE Full text] [doi: 10.1111/acem.12692] [Medline: 25997375]
15. Yeager CM, Benight CC. If we build it, will they come? Issues of engagement with digital health interventions for trauma recovery. Mhealth 2018;4:37 [FREE Full text] [doi: 10.21037/mhealth.2018.08.04] [Medline: 30363749]

XSL•FO
RenderX

16. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. J Med Internet Res 2019 Sep 25;21(9):e14567 [FREE Full text] [doi: 10.2196/14567] [Medline: 31573916]
17. Yang Q. Theory-based social and non-social engagement features in smoking cessation mobile apps: a content analysis. Int J Environ Res Public Health 2021 Aug 29;18(17):9106 [FREE Full text] [doi: 10.3390/ijerph18179106] [Medline: 34501696]
18. Chao DY, Lin TM, Ma W. Enhanced self-efficacy and behavioral changes among patients with diabetes: cloud-based mobile health platform and mobile app service. JMIR Diabetes 2019 May 10;4(2):e11017 [FREE Full text] [doi: 10.2196/11017] [Medline: 31094324]
19. Song T, Qian S, Cui T, Yu P. The use of theory in mobile health interventions for patient self-management of chronic diseases. Stud Health Technol Inform 2019 Aug 21;264:1982-1983. [doi: 10.3233/SHTI190745] [Medline: 31438439]
20. Besoain F, Gallardo I. Think about it: promoting physical activity with a mobile app with a theory-based approach. IEEE Access 2023;11:37021-37030. [doi: 10.1109/access.2023.3265891]
21. Milne-Ives M, Lam C, De Cock C, Van Velthoven MH, Meinert E. Mobile apps for health behavior change in physical activity, diet, drug and alcohol use, and mental health: systematic review. JMIR Mhealth Uhealth 2020 Mar 18;8(3):e17046 [FREE Full text] [doi: 10.2196/17046] [Medline: 32186518]
22. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95 [FREE Full text] [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
23. Anderson E, Durstine JL. Physical activity, exercise, and chronic diseases: a brief review. Sports Med Health Sci 2019 Dec;1(1):3-10 [FREE Full text] [doi: 10.1016/j.smhs.2019.08.006] [Medline: 35782456]
24. Elliot CA, Hamlin MJ. Combined diet and physical activity is better than diet or physical activity alone at improving health outcomes for patients in New Zealand's primary care intervention. BMC Public Health 2018 Feb 08;18(1):230 [FREE Full text] [doi: 10.1186/s12889-018-5152-z] [Medline: 29422040]
25. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
26. Voorheis P, Bhuiya AR, Kuluski K, Pham Q, Petch J. Making sense of theories, models, and frameworks in digital health behavior change design: qualitative descriptive study. J Med Internet Res 2023 Mar 15;25:e45095 [FREE Full text] [doi: 10.2196/45095] [Medline: 36920442]
27. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. Health Psychol Rev 2015;9(3):323-344 [FREE Full text] [doi: 10.1080/17437199.2014.941722] [Medline: 25104107]
28. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
29. Thomas PC, Curtis K, Potts HWW, Bark P, Perowne R, Rookes T, et al. Behavior change techniques within digital interventions for the treatment of eating disorders: systematic review and meta-analysis. JMIR Ment Health 2024 Aug 01;11:e57577 [FREE Full text] [doi: 10.2196/57577] [Medline: 39088817]
30. Van Rhoon L, Byrne M, Morrissey E, Murphy J, McSharry J. A systematic review of the behaviour change techniques and digital features in technology-driven type 2 diabetes prevention interventions. Digit Health 2020;6:2055207620914427 [FREE Full text] [doi: 10.1177/2055207620914427] [Medline: 32269830]
31. Martín-Martín J, Roldán-Jiménez C, De-Torres I, Muro-Culebras A, Escriche-Escuder A, Gonzalez-Sanchez M, et al. Behavior change techniques and the effects associated with digital behavior change interventions in sedentary behavior in the clinical population: a systematic review. Front Digit Health 2021;3:620383 [FREE Full text] [doi: 10.3389/fdgth.2021.620383] [Medline: 34713097]
32. Michie S, West R, Sheals K, Godinho CA. Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used. Transl Behav Med 2018 Mar 01;8(2):212-224 [FREE Full text] [doi: 10.1093/tbm/ibx019] [Medline: 29381786]
33. Bandura A, Freeman W, Lightsey R. Self-efficacy: the exercise of control. J Cogn Psychother 1999 Jan 01;13(2):158-166. [doi: 10.1891/0889-8391.13.2.158]
34. Taj F, Klein MCA, van Halteren A. Digital health behavior change technology: bibliometric and scoping review of two decades of research. JMIR Mhealth Uhealth 2019 Dec 13;7(12):e13311 [FREE Full text] [doi: 10.2196/13311] [Medline: 31833836]
35. Feil K, Fritsch J, Rhodes RE. The intention-behaviour gap in physical activity: a systematic review and meta-analysis of the action control framework. Br J Sports Med 2023 Oct;57(19):1265-1271. [doi: 10.1136/bjsports-2022-106640] [Medline: 37460164]

XSL•FO
RenderX

36. McDermott MS, Oliver M, Iverson D, Sharma R. Effective techniques for changing physical activity and healthy eating intentions and behaviour: a systematic review and meta-analysis. Br J Health Psychol 2016 Nov;21(4):827-841. [doi: 10.1111/bjhp.12199] [Medline: 27193530]
37. Adriaanse MA, Vinkers CDW, De Ridder DTD, Hox JJ, De Wit JBF. Do implementation intentions help to eat a healthy diet? A systematic review and meta-analysis of the empirical evidence. Appetite 2011 Feb;56(1):183-193 [FREE Full text] [doi: 10.1016/j.appet.2010.10.012] [Medline: 21056605]
38. Rhodes RE, de Bruijn GJ. How big is the physical activity intention-behaviour gap? A meta-analysis using the action control framework. Br J Health Psychol 2013 May;18(2):296-309. [doi: 10.1111/bjhp.12032] [Medline: 23480428]
39. Evans J. Dual Process Theory 2.0. London: Routledge; Nov 09, 2017:137-155.
40. Ma H, Wang A, Pei R, Piao M. Effects of habit formation interventions on physical activity habit strength: meta-analysis and meta-regression. Int J Behav Nutr Phys Act 2023 Sep 12;20(1):109 [FREE Full text] [doi: 10.1186/s12966-023-01493-3] [Medline: 37700303]
41. Dunwoody P, Haarbauer E, Mahan R, Marino C. Cognitive adaptation and its consequences: a test of cognitive continuum theory. Journal of Behavioral Decision Making. 2000 Jan 10. URL: https://www.researchgate.net/publication/285649903_Cognitive_adaptation_and_it's_consequences_A_test_of_cognitive_continuum_theory [accessed 2025-12-23]
42. Bannert M. Managing cognitive load—recent trends in cognitive load theory. Learning and Instruction 2002 Feb;12(1):139-146. [doi: 10.1016/s0959-4752(01)00021-4]
43. Williams M, Rana N, Dwivedi Y. The unified theory of acceptance and use of technology (UTAUT): a literature review. Journal of Enterprise Information Management 2015 Apr 13:443-488. [doi: 10.1108/JEIM-09-2014-0088]
44. Marangunić N, Granić A. Technology acceptance model: a literature review from 1986 to 2013. Universal Access in the Information Society 2014 Feb 16:81-95. [doi: 10.1007/s10209-014-0348-1]
45. He Z, Hassan MA, Saiz-González P, Ryu S, Wang R, Gao Z. Smartphone app-based interventions on physical activity behaviors and psychological correlates in healthy young adults: a systematic review. PLoS One 2024;19(4):e0301088 [FREE Full text] [doi: 10.1371/journal.pone.0301088] [Medline: 38578729]
46. Wang J, Zhu Z, Shuling Z, Fan J, Jin Y, Gao Z, et al. Effectiveness of mHealth app-based interventions for increasing physical activity and improving physical fitness in children and adolescents: systematic review and meta-analysis. JMIR Mhealth Uhealth 2024 Apr 30;12:e51478 [FREE Full text] [doi: 10.2196/51478] [Medline: 38687568]
47. Li S, Zhou Y, Ma HM, Wang AQ, Tang XY, Pei RY, et al. Effects of mobile applications on weight-related, behavior and metabolic outcomes for adults with overweight and obesity: a systematic review and meta-analysis. Stud Health Technol Inform 2024 Jul 24;315:386-391. [doi: 10.3233/SHTI240175] [Medline: 39049288]
48. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
49. Dugas M, Gao GG, Agarwal R. Unpacking mHealth interventions: a systematic review of behavior change techniques used in randomized controlled trials assessing mHealth effectiveness. Digit Health 2020;6:2055207620905411 [FREE Full text] [doi: 10.1177/2055207620905411] [Medline: 32128233]
50. Institute of Medicine (US) Subcommittee on Military Weight Management. Weight Management: State of the Science and Opportunities for Military Programs. Washington (DC): National Academies Press (US); 2004.
51. Kalgotra P, Raja U, Sharda R. Growth in the development of health and fitness mobile apps amid COVID-19 pandemic. Digit Health 2022;8:20552076221129070 [FREE Full text] [doi: 10.1177/20552076221129070] [Medline: 36211794]
52. Dang A, Dang D, Rane P. The expanding role of digital therapeutics in the post-COVID-19 era. TOCOVIDJ 2021 May 21;1(1):32-37. [doi: 10.2174/2666958702101010032]
53. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Healthcare. York: CRD, University of York; Jan 2009.
54. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143 [FREE Full text] [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
55. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3 [FREE Full text] [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
56. Paul M, Maglaras L, Ferrag M, Almomani I. Digitization of healthcare sector: a study on privacy and security concerns. ICT Express 2023 Aug;9(4):571-588. [doi: 10.1016/j.icte.2023.02.007]
57. Marques M, Wright A, Corker E, Johnston M, West R, Hastings J, et al. The behaviour change technique ontology: transforming the behaviour change technique taxonomy v1. Wellcome Open Res 2023;8:308 [FREE Full text] [doi: 10.12688/wellcomeopenres.19363.2] [Medline: 37593567]
58. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ 2019 Aug 28;366:l4898 [FREE Full text] [doi: 10.1136/bmj.l4898] [Medline: 31462531]
59. Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. URL: https://ohri.ca/en/who-we-are/core-facilities-and-platforms/ottawa-methods-centre/newcastle-ottawa-scale [accessed 2024-10-04]

XSL•FO
RenderX

60. Deeks J, Higgins J, Altman D. Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions. United States: Wiley Blackwell; Sep 20, 2019:241-284.
61. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A Product from the ESRC Methods Programme. England: Lancaster University; 2006.

Abbreviations

BCT: behavior change technique
COM-B: Capability, Opportunity, Motivation-Behavior
DHI: digital health intervention
MeSH: Medical Subject Headings
mHealth: mobile health
PICOS: Population, Intervention, Comparator, Outcome, and Study
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols

Edited by A Schwartz; submitted 14.Feb.2025; peer-reviewed by H Kirk; comments to author 16.May.2025; revised version received 30.Jun.2025; accepted 19.Aug.2025; published 17.Feb.2026.

Please cite as:
Riccalton V, Milne-Ives M, Brownson-Smith R, Ananthakrishnan A, Cong C, Oliver E, Meinert E
Effectiveness and Theoretical Foundations of mHealth Apps for Physical Activity, Healthy Eating, and Weight Loss: Protocol for a Systematic Review and Meta-Analysis
JMIR Res Protoc 2026;15:e72664
URL: https://www.researchprotocols.org/2026/1/e72664
doi:10.2196/72664
PMID:

©Victoria Riccalton, Madison Milne-Ives, Rosiered Brownson-Smith, Ananya Ananthakrishnan, Cen Cong, Emily Oliver, Edward Meinert. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Investigating Placebos and Controls Used in Large Language Model–Based Chatbot Intervention Trials: Protocol for a Methodological Review

Leo Druart[1,2], PT, PhD; Vanda Faria[1,3,4], PhD; Marco Annoni[5], PhD; John Torous[6], MD; Moa Pontén[7,8], PhD; Charlotte Blease[1,6], PhD

[1]Participatory eHealth and Health Data Research Group, Department of Women's and Children's Health, Uppsala University, Uppsala, Uppsala, Sweden

[2]Department of Radiology, Warren Alpert Medical School, Brown University, Providence, RI, United States

[3]Brain and Eye Pain Imaging Lab, Pain and Affective Neuroscience Center, Department of Anesthesiology, Critical Care and Pain Medicine, Boston Children's Hospital and Harvard Medical School, Harvard University, Boston, MA, United States

[4]Comprehensive Pain Center, Faculty of Medicine, University Hospital Carl Gustav Carus, Dresden, Saxony, Germany

[5]Interdepartmental Center for Research Ethics and Integrity, National Research Council, Milan, Lazio, Italy

[6]Department of Psychiatry, Beth Israel Deaconess Medical Center, Division of Digital Psychiatry, Harvard Medical School, Boston, MA, United States

[7]Department of Clinical Neuroscience, Centre for Psychiatry Research, Karolinska Institutet, Stockholm, Stockholm, Sweden

[8]Stockholm Health Care Services, Stockholm, Stockholm, Sweden

Corresponding Author:
Leo Druart, PT, PhD
Participatory eHealth and Health Data Research Group, Department of Women's and Children's Health
Uppsala University
MTC-huset (MTC Building), Dag Hammarskjölds väg 14B, 1 tr
Uppsala, 75237
Sweden
Phone: 46 18 471 00 00
Email: leo.druart@uu.se

Abstract

Background: Large language model (LLM)–based chatbots are rapidly being repurposed as patient-facing digital health tools. Their interactive, adaptive, and seemingly empathic behavior can heighten engagement and expectancy—nonspecific factors that complicate causal inference. Yet, comparator strategies in LLM trials are inconsistently defined and often undermatched (eg, minimal education vs highly engaging chatbots), risking biased effect estimates and poor reproducibility.

Objective: The aim of this study was to systematically identify and categorize the control conditions used in interventional studies of LLM-based, patient-facing digital health interventions and to evaluate their methodological appropriateness. Secondary aims are to describe variability by health domain and study design and to explore whether control type/quality relates to the direction of reported effects.

Methods: This protocol follows PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) and is registered in PROSPERO. Eligible studies are interventional designs that evaluate LLM-based, patient-facing digital health interventions; any control condition is eligible (including no control, waitlist, treatment-as-usual, attention/education, active comparator, or sham digital control). We will search PubMed, PsycINFO, CENTRAL, CINAHL, and Scopus for records from January 1, 2023, onward. All records will be managed and screened in Rayyan by 2 independent reviewers. Dual, independent data extraction will target study context, intervention details, and control-arm characteristics (typology, rationale, matching to nonspecifics, blinding, reporting). No formal risk-of-bias assessments are planned, as the focus is on meta-research.

Results: At submission, the protocol is registered in PROSPERO and has received no specific funding. Scoping searches are complete; full screening and extraction have not yet commenced.

Conclusions: This review will provide an empirical map of control practices in LLM chatbot trials and guidance for designing better-matched comparators, supporting more valid and interpretable evaluations as LLMs diffuse into patient care.

Trial Registration: PROSPERO CRD420251246148; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251246148

International Registered Report Identifier (IRRID): PRR1-10.2196/90507

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e90507) doi:10.2196/90507

KEYWORDS

large language models; chatbots; digital health; control conditions; methodological review

Introduction

Background

Since ChatGPT's launch in November 2022, interest in patient-facing conversational agents whose replies are generated by large language models (LLM), in other words, LLM-based chatbots, has surged, with growing attention to their potential clinical applications. Just a year after its launch, already 1 in 5 primary care physicians in the United Kingdom reported using it in their professional activities [1]. While the application of LLMs in health care presents significant opportunities to improve health care, for example, by diminishing the administrative burden for clinicians while also helping patients through its accessibility and availability, they also invite significant challenges that warrant consideration [2,3].

Now, patients are using these chatbots for their health care needs. For example, by early 2025, 10% of Australians had already used ChatGPT for health-related questions, with 61% asking questions about taking action that typically require clinical advice [4]. Of the participants who had not yet used ChatGPT for a health-related question, 40% reported that they would do so within the next 6 months. More striking, in October 2025, OpenAI announced that, every week, 1.2 million users disclosed thinking about taking their own lives with ChatGPT [5]. It is clear that these tools are being repurposed for lifestyle counselling or mental health support. Studies demonstrate that these chatbots can simulate empathy, personalize responses, and sustain meaningful dialogue, potentially amplifying expectancy and engagement effects beyond those seen in rule-based apps [6].

One key challenge, however, lies in measuring their efficacy and determining whether current methodological tools are adapted to assess these new interventions. While rigorous trials are a prerequisite for introducing new drugs to patients, LLM tools have not been subjected to similar scrutiny. Ideally, an LLM should only be used as a health intervention after safety and effectiveness have been demonstrated [7].

Such rapid dissemination raises concerns that deployment may have outpaced the development of appropriate evaluation practices. Like other nonpharmacological interventions such as surgery, physiotherapy, or psychotherapy [8], LLMs, and more generally digital interventions, pose unique challenges for evaluation when compared to drugs. Their effects often rely on multiple components, and they are, as interventions, highly adaptive. Moreover, they inherently depend on user engagement, interface design, and personalization, all of which could potentially influence outcomes independently of any hypothesized active ingredient. As the field evolves, these methodological specificities will need greater methodological attention [9]. Specifically, methodological attention should focus on designing appropriate control groups for evaluating these tools. Controls are necessary to sift out any confounding factors and accurately assess the efficacy of LLM interventions [10]. Yet, adequate controls are challenging to design and implement [11]. These considerations are especially salient for unsupervised LLMs whose conversational, adaptive, and apparently empathic responses may heighten engagement and expectancy effects [6,12,13]. In this context, it is important to distinguish between placebo responses—any observed change following a placebo intervention, and placebo effects, proper, which refer specifically to the component of that response attributable to placebo mechanisms such as expectancy or conditioning [11,14].

This is why the choice of control condition is fundamental to assess the efficacy of these tools—a consideration that has not only scientific but also ethical implications [12]. Foundational guidance for nonpharmacological trials emphasizes that controls should be selected to answer a specific scientific question while accounting for nonspecific factors such as time, attention, therapeutic alliance, and expectancy effects [15,16]. In LLM trials, an adequate control may need to match not only time and attention but also the look-and-feel and level of interactive engagement (eg, notifications, chat, interface aesthetics) to approximate a placebo control that isolates the hypothesized mechanism of action [17]. Without such methodological matching, apparent benefits may be inflated (if controls are minimally engaging) or genuine effects obscured (if controls inadvertently contain active elements), limiting the validity of the findings.

Recent scholarship has begun to map the challenges in digital health more generally. Goldberg and colleagues [17] proposed a typology of 11 control condition types used in mobile health randomized trials, along with a decision framework and a gradation of comparison strength, illuminating how different control choices afford different inferences. Methodological work has urged for comparators that better approximate what some researchers have dubbed "active placebo" characteristics in digital trials—mirroring the active arm's interactivity and engagement while withholding the hypothesized mechanism [17,18]. A systematic review of digital shams in digital therapeutics highlighted practical and conceptual challenges—among them, inconsistently defined controls, inadequate matching on engagement, and limited reporting [18].

These contributions underscore that controls are heterogeneous, variably justified, and unevenly reported in digital intervention trials and that stronger standards are needed to ensure valid inference and comparability across studies. However, this has yet to be done for LLM-based interventions, which are increasingly gaining popularity.

Scoping reviews show that evaluation methods for LLMs are lacking in high-quality effectiveness research. The evaluation of LLMs in coaching and lifestyle has been suggested to present a lack of methodological standardization and a general trend toward less rigorous evaluation designs [19]. Similar findings

XSL•FO
RenderX

were reported for LLMs used in mental health [20]. Moreover, effectiveness has been defined subjectively in over two-thirds of studies evaluating LLMs to provide health advice [21]. These findings highlight that the need for improved control designs is particularly acute for LLM, given their widespread use has advanced faster than their systematic evaluation.

In summary, the exponential growth of LLMs used as patient-facing digital health tools and the unique methodological challenges they present create a critical empirical gap: we lack a consolidated mapping of the control conditions actually used across these interventional studies of LLM-based chatbots. We also need to better understand how consistently these controls are justified and described and the extent to which they match nonspecific factors that are especially salient for conversational/artificial intelligence (AI) systems. Moreover, while guidance documents (eg, CONSORT-EHEALTH [Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth]) have improved reporting expectations, they do not extend to control adaptations such as which control to use for a given scientific aim, and evidence suggests that control selection and labeling (eg, treatment as usual, education, waitlist) remain inconsistent across the literature [17,22]. A methodological synthesis focused specifically on LLM-based interventions, where conversationality and perceived intelligence may intensify nonspecific effects, can provide the field with a clearer description of practices, highlight reporting deficits, and inform design decisions in upcoming trials.

Accordingly, this protocol describes a methodological review that aims to answer the question: what type of control conditions are used in interventional studies of LLM-based, patient-facing digital health interventions, and how methodologically appropriate are these controls?

Objectives

This methodological review aims to systematically identify and categorize the control conditions used in interventional studies of LLM-based chatbot interventions and to evaluate their methodological appropriateness—specifically, whether a clear rationale is provided, whether placebo response components are adequately matched, and whether risks of blinding failure or contamination are addressed and reported transparently. We will also describe how control design and reporting vary across 3 health domains (mental health, chronic disease, lifestyle changes) and study designs and explore whether the type or quality of control relates to the direction of reported intervention effects (positive, null, or negative findings).

The population consists of individuals in any health domain who are intended users of an AI-based digital health tool. The intervention is an LLM-based, patient-facing digital modality, and the comparator may be any control condition, including no control, a waitlist, treatment-as-usual, attention/education, an active comparator, or another type of intervention control. The outcomes of the review include the methodological description, classification, and appraisal of control arms rather than clinical effectiveness; notwithstanding, the quality of the former will be relevant to measuring the latter.

Methods

Study Design

This is a protocol for a methodological (meta-research) review reported in accordance with PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) [23]. Where relevant, general methods are informed by the Cochrane Handbook [24]. Methodological reviews identify inconsistencies, gaps, and sources of bias in the design, conduct, and reporting of research. They are particularly useful when an area of research is rapidly evolving and lacks standardization, allowing researchers to map current practices and evaluate their adequacy [25]. Given the heterogeneity and emerging nature of control conditions in LLM-based digital health interventions, this approach enables a rigorous assessment of whether existing trials use appropriate comparators and adhere to methodological best practices.

Eligibility Criteria

We will include interventional studies evaluating LLM-based health interventions, including randomized controlled trials, nonrandomized trials, quasi-experimental studies, and single-group trials (pilot/feasibility studies). Published study protocols may be considered. We will exclude observational designs (cohort, cross-sectional, case-control), case reports, editorials, reviews and commentaries, and non–peer-reviewed publications.

Regarding the population, we will include studies with individuals of any age/sex/health status who are intended users of the intervention with no restriction by clinical domain. Subpopulation analyses will be conducted for mental health, lifestyle change, or chronic disease management interventions.

The interventions concerned will be LLM-based chatbots used as patient-facing digital health interventions in which LLM constitutes the active ingredient (eg, mobile apps, web platforms, chatbots/virtual agents, generative-AI tools). We will exclude studies where the digital component serves only as a delivery vehicle (eg, videoconferencing for standard therapy) and interventions that are not directly patient-facing. An LLM will be considered the active ingredient when it generates patient-facing linguistic output intended to contribute to the intervention experience, including dialogue, explanation, summarization, or rephrasing. This reflects theoretical and empirical work suggesting that AI-generated language itself may influence expectancy, meaning-making, or engagement [11,26]. By contrast, LLM uses limited to back-end processing, or clinician-facing functions will be excluded.

Regarding controls, any control condition is eligible, including none (eg, no-treatment), waitlist, treatment-as-usual, active comparators, attention controls, placebo/sham digital controls, or standard care. Lastly, clinical outcomes will not be the eligibility criteria in this review. As a methodological review, our primary outcomes concern the characteristics of the control arms used in the included studies (see “Data items”). Included publications will be in English and published from January 2023 onward to reflect the adoption and investigation of LLMs. This cutoff was selected because it marks the point at which

XSL•FO
RenderX

patient-facing, conversational LLM-based chatbots became both technically plausible and widely deployed at scale. Earlier generative or GPT-based systems were typically experimental, constrained, or not deployed as adaptive patient-facing interventions, and are therefore methodologically distinct from the intervention class examined in this review. An additional sense of urgency into the effectiveness of these tools is required because LLM-based chatbots are reported being adopted at scale by consumers and patients. For example, OpenAI reported in 2025 that 1.2 million adults every month disclosed suicidal thoughts to ChatGPT [5] and a recent UK study found that 37% of adults had used the tools for emotional or mental health support [27].

Search Strategy

We will search 5 major bibliographic databases: PubMed, PsycINFO via EBSCO, CENTRAL via Cochrane Library, CINAHL via EBSCO, and Scopus. We will screen references of included studies and relevant reviews and consider backward and forward citation tracking. The coverage period will begin on January 1, 2023, and the search will be performed by the Uppsala University librarian team on the project (Görel Sundström and Emma Kristina Carlsson). Given the rapid pace of development in this field, we will conduct an updated search and incorporate any newly identified studies before submitting the manuscript or during the peer-review process.

The final database-specific strategies will be designed in collaboration with study librarians. The search strategies will be peer-reviewed in accordance with the PRESS 2015 framework [28]. The search strategy will be designed for PubMed using Medical Subject Headings (MeSH) and then adapted to the other databases and their thesauri. Terms will also be searched for in the title/abstract of publications. The PubMed search string is available in Multimedia Appendix 1.

Data Management and Selection Process

Records will be managed in Rayyan for deduplication and screening [29]. Screening will proceed in two stages, that is, (1) titles/abstracts and (2) full texts, conducted independently by two reviewers through the Rayyan software. Disagreements will be resolved by discussion or by a third reviewer. We will report a PRISMA 2020 flow diagram for identification, screening, eligibility, and inclusion [30]. Prior to formal screening, we will pilot test on a smaller sample of studies. Any disagreements at this stage will be discussed and help refine the screening process to ensure consistent interpretation.

Data Extraction

At least two reviewers will independently extract data by using an extraction form. Disagreements will be resolved through discussion or by a third reviewer. Where intervention or control details are unclear, we will attempt to contact the author (via one initial email and one reminder). We will record when information is not reported or unclear. Extracted data are presented in Table 1.

XSL•FO
RenderX

Table 1. Planned extracted data.

Type of data	Data
Study identification	• First author name • Year of publication • Country of study
Study design and sample	• Study design (options: randomized controlled trial, nonrandomized trial, single group) • Sample size • Participant characteristics (age, gender, clinical condition, self-report or clinical diagnosis, verification of self-reported data) • Recruitment (online, in-person) • Follow-up characteristics (length of study, compensation or incentives) • Research aim (copy research aim) • Requires a control group? (deduced from research aim, possible options: yes, no, unsure)
Intervention characteristics	• Name/description of large language model intervention • Mode of delivery (eg, mobile app, web-based content) • Target condition or domain • Duration, frequency, and intensity of the intervention • Supposed mechanism of action explained (yes/no) • Supposed mechanism of action (detail in free text)
Comparator characteristics	• Type of control condition • Type of control condition (describe in free text) • Justification for choice of control explained (yes/no) • Justification for choice of control (detail in free text) • Blinding status (participants, providers, assessors, analysts)
Methodological appropriateness	• Given the study design, research aim, need for control, intervention description, supposed mechanism of action and type of control condition, is the control group appropriate? (Yes/No) • Comment on the previous answer • Are placebo responses controlled for adequately? • Are placebo effects controlled for adequately?
Outcomes and results	• Conclusion on primary outcome (positive finding, null finding, unclear)

Outcomes

The primary outcome of this review is the type and frequency of control conditions used. The absence of a comparator will be coded as a distinct control category (no control). Control-arm appropriateness will be evaluated relative to the stated research aim; where a comparator is required to support causal inference but absent, the study will be classified as not appropriately controlled.

Secondary outcomes will include the adequacy/appropriateness of controls with respect to the placebo response and theoretical justification; the reporting quality of control descriptions and justifications. Control-arm appropriateness will be evaluated using a predefined set of extraction items (Table 1: items in methodological appropriateness), which are considered jointly as a qualitative decision framework. These items assess whether a control condition is present, justified, aligned with the study aim, and designed to address relevant placebo response components. Appropriateness will be classified categorically rather than scored numerically, as different control choices afford different inferential aims rather than representing a linear quality continuum. Where judgments require interpretation, disagreements will be resolved through collegial discussion among authors with complementary expertise in placebo research, clinical trial methodology, and digital health interventions [11,31]. For Table 1 items regarding the methodological appropriateness of accounting for placebo responses and effects, we will consider that placebo responses refer to observed changes following intervention exposure, whereas placebo effects proper denote changes attributable to mechanisms such as expectancy or conditioning [11,14].

An exploratory outcome is to assess whether the type and quality of the control condition are associated with differences in the reported effectiveness of digital interventions, by comparing success proportions (see Table 1) across control categories.

Risk of Bias in Individual Studies

We will not conduct a formal risk-of-bias assessment because the review does not synthesize treatment effects; instead, it characterizes control-arm design and reporting (ie, a meta-research focus). We will, however, extract bias-relevant features (eg, blinding, contamination, missing data statements) to inform the appraisal of control adequacy.

Assessment of publication bias or selective reporting at the meta-analytic level is not planned, as we do not pool effect

XSL•FO
RenderX

sizes. GRADE (Grading of Recommendations Assessment, Development and Evaluation) is not applicable because we are not grading the certainty of an effect estimate [32].

Data Synthesis

We will use descriptive statistics (counts, proportions, medians/IQRs where relevant) to summarize control-type frequencies, presence/absence of justifications, and reporting completeness. Stratified summaries will be presented by health domain (mental health, lifestyle changes, chronic disease management). For studies with more than two arms, we will treat each LLM-vs-control comparison in a multi-arm study as a separate comparison unit for exploratory analyses of outcome direction by control type.

Amendments

Any protocol amendments (eg, eligibility refinements, added subgroup) will be documented, dated, and justified within the PROSPERO record and reported in the final manuscript [23].

Results

As of submission, the protocol has been registered in PROSPERO (CRD420251246148). The database-specific search strategies have been pretested, and preliminary scoping searches have been conducted to estimate the volume and refine the eligibility criteria. No study data have been synthesized yet; this manuscript details the a priori methods that will guide the review as shown in Figure 1.

XSL•FO
RenderX

Figure 1. Anticipated PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis) 2020 flow diagram for the planned study selection process.

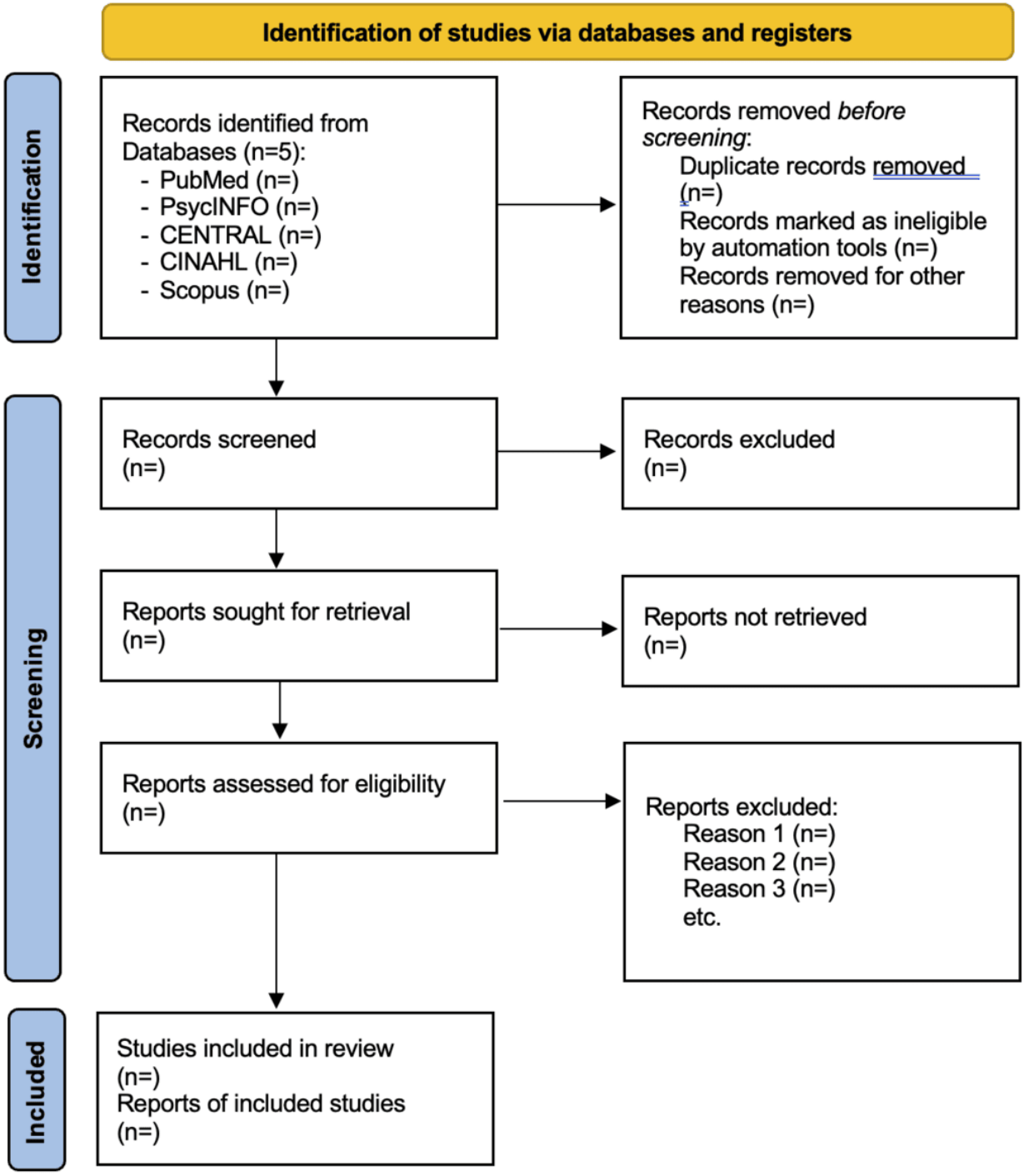

Discussion

Principal Results

This review will provide researchers, clinicians, and policymakers with a systematic assessment of the control conditions used to measure the efficacy of LLM-based interventions and their adequacy. This will enable a more comprehensive evaluation of the current evidence regarding the effectiveness of LLM interventions, allowing health care providers, policymakers, and, most importantly, patients to make informed decisions when considering the use of LLMs in health care.

We hypothesize that many LLM-based health trials employ control conditions that are variably defined, underjustified, and insufficiently matched for nonspecific effects, reflecting the specific challenges that evaluating digital interventions pose, yet potentially biasing inferences about the efficacy of the intervention. The outcome of this research will be highly informative for all stakeholders and will allow researchers to design more optimal control conditions in the future.

XSL•FO
RenderX

Limitations

A key limitation of this review is that the evidence base for LLM-based digital health interventions is rapidly evolving, which restricts the strength of any conclusions about optimal control conditions. This challenge is further accentuated by the inherent delay between the emergence of new technologies, the conduct of research, and the publication of findings. Nonetheless, adoption of these systems is already occurring at scale, and the ethical risks associated with inadequate or poorly designed controls are substantial. In this context, we argue that providing an early methodological assessment is preferable to delaying synthesis until the field matures, as timely guidance may help prevent avoidable methodological weaknesses from becoming entrenched in future research.

Comparison With Prior Work

There has yet to be any systematic mapping of the design and use of control conditions in LLM trials. However, research has been performed for digital interventions more broadly, allowing researchers to adapt reporting guidelines to the specifics of eHealth interventions through the CONSORT-EHEALTH, emphasizing the need to specifically describe components such as usability, dose/exposure, and usage metrics [22].

A previous review provided an initial examination of how control conditions were used in digital therapeutics; however, this work preceded the advent of LLMs in health care [18]. Goldberg and colleagues [17] offered a precise mapping of the typology of control conditions that were used in mobile health interventions, showing a clear need for improvement.

Conclusions

This protocol outlines a timely and rigorous plan for a methodological synthesis of control conditions used in LLM-based health trials. By systematically classifying control arms, assessing their appropriateness, and evaluating the quality of reporting, the review will provide an empirical foundation for improved comparator design in future LLM health research. The findings are expected to inform investigators and reviewers in the rapidly evolving field of LLM-based chatbots and generative AI, where standards for comparators are still being developed.

Acknowledgments

The authors would like to thank Emma Kristina Carlsson and Görel Sundström of Uppsala University Library for their assistance in developing the search strategy for this research protocol. The authors declare the use of generative artificial intelligence (GAI) in the manuscript preparation process. According to the GAIDeT (GAI Delegation Taxonomy, 2025), the following tasks were, in part, delegated to GAI tools under full human supervision: (1) proofreading and editing and (2) publication support (fitting word counts). The GAI tool used was ChatGPT5.2. The responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

This article is a research protocol and does not include any data. Data resulting from this research will be made available upon publication of the results.

Conflicts of Interest

CB is an Associate Editor of *JMIR Mental Health* and JT is Editor-in-Chief of *JMIR Mental Health*. This manuscript was handled by an academic editor with no conflicts of interest. CB received funding from Google DeepMind in 2026 for independent ethics work, which did not influence the content of this manuscript.

Multimedia Appendix 1
PubMed search strategy.
[DOCX File , 14 KB - resprot_v15i1e90507_app1.docx]

References

1. Blease CR, Locher C, Gaab J, Hägglund M, Mandl KD. Generative artificial intelligence in primary care: an online survey of UK general practitioners. BMJ Health Care Inform 2024 Sep 17;31(1):e101102 [FREE Full text] [doi: 10.1136/bmjhci-2024-101102] [Medline: 39288998]
2. Blease C, Torous J. ChatGPT and mental healthcare: balancing benefits with risks of harms. BMJ Ment Health 2023 Nov;26(1):e300884 [FREE Full text] [doi: 10.1136/bmjment-2023-300884] [Medline: 37949485]

XSL•FO
RenderX

3. Faria V, Goturi N, Dynak A, Talbert C, Pondelis N, Annoni M, et al. Triadic relations in healthcare: surveying physicians' perspectives on generative AI integration and its role on empathy, the placebo effect and patient care. Front Psychol 2025;16:1612215 [FREE Full text] [doi: 10.3389/fpsyg.2025.1612215] [Medline: 41312268]
4. Ayre J, Cvejic E, McCaffery K. Use of ChatGPT to obtain health information in Australia, 2024: insights from a nationally representative survey. Med J Aust 2025 Mar 03;222(4):210-212 [FREE Full text] [doi: 10.5694/mja2.52598] [Medline: 39901778]
5. 2025. Amélioration des réponses de ChatGPT lors de conversations sensibles. URL: https://openai.com/fr-FR/index/strengthening-chatgpt-responses-in-sensitive-conversations/ [accessed 2025-12-29]
6. Ayers JW, Poliak A, Dredze M, Leas EC, Zhu Z, Kelley JB, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023 Jun 01;183(6):589-596 [FREE Full text] [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]
7. Kohane IS. Injecting artificial intelligence into medicine. NEJM AI 2024 Jan;1(1):1. [doi: 10.1056/aie2300197]
8. Boutron I, Altman DG, Moher D, Schulz KF, Ravaud P, CONSORT NPT Group. CONSORT statement for randomized trials of nonpharmacologic treatments: a 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. Ann Intern Med 2017 Jul 04;167(1):40-47 [FREE Full text] [doi: 10.7326/M17-0046] [Medline: 28630973]
9. van Mierlo T. From efficacy to scale: addressing digital health's original sin. J Med Internet Res 2025 Oct 27;27:e85878 [FREE Full text] [doi: 10.2196/85878] [Medline: 41143679]
10. Torous J, Firth J. The digital placebo effect: mobile mental health meets clinical psychiatry. Lancet Psychiatry 2016 Feb;3(2):100-102. [doi: 10.1016/S2215-0366(15)00565-9] [Medline: 26851322]
11. Druart A, Blease C. Should digital health trials be placebo controlled? JME Practical Bioethics Institute of Medical Ethics 2025 Dec 10;1:1. [doi: 10.1136/jmepb-2025-000027]
12. Blease C. Placebo, nocebo, and machine learning: how generative AI could shape patient perception in mental health care. JMIR Ment Health 2025 Aug 15;12:e78663 [FREE Full text] [doi: 10.2196/78663] [Medline: 40815809]
13. Kharko A, McMillan B, Hagström J, Muli I, Davidge G, Hägglund M, et al. Generative artificial intelligence writing open notes: a mixed methods assessment of the functionality of GPT 3.5 and GPT 4.0. Digit Health 2024;10:20552076241291384 [FREE Full text] [doi: 10.1177/20552076241291384] [Medline: 39493632]
14. Evers AWM, Colloca L, Blease C, Annoni M, Atlas LY, Benedetti F, et al. Implications of placebo and nocebo effects for clinical practice: expert consensus. Psychother Psychosom 2018;87(4):204-210 [FREE Full text] [doi: 10.1159/000490354] [Medline: 29895014]
15. Locher C, Gaab J, Blease C. When a placebo is not a placebo: problems and solutions to the gold standard in psychotherapy research. Front Psychol 2018;9:2317 [FREE Full text] [doi: 10.3389/fpsyg.2018.02317] [Medline: 30542310]
16. Blease C. Psychotherapy and placebos: manifesto for conceptual clarity. Front Psychiatry 2018;9:379 [FREE Full text] [doi: 10.3389/fpsyt.2018.00379] [Medline: 30177892]
17. Goldberg SB, Sun S, Carlbring P, Torous J. Selecting and describing control conditions in mobile health randomized controlled trials: a proposed typology. NPJ Digit Med 2023 Sep 30;6(1):181 [FREE Full text] [doi: 10.1038/s41746-023-00923-7] [Medline: 37775522]
18. Bos E, Preller KH, Kaur G, Malhotra P, Kharawala S, Motti D. Challenges with the use of digital sham: systematic review and recommendations. J Med Internet Res 2023 Oct 24;25:e44764 [FREE Full text] [doi: 10.2196/44764] [Medline: 37874638]
19. Lai X, Lai Y, Chen J, Huang S, Gao Q, Huang C. Evaluation strategies for large language model-based models in exercise and health coaching: scoping review. J Med Internet Res 2025 Oct 14;27:e79217 [FREE Full text] [doi: 10.2196/79217] [Medline: 41086432]
20. Hua Y, Na H, Li Z, Liu F, Fang X, Clifton D, et al. A scoping review of large language models for generative tasks in mental health care. NPJ Digit Med 2025 Apr 30;8(1):230 [FREE Full text] [doi: 10.1038/s41746-025-01611-4] [Medline: 40307331]
21. Huo B, Boyle A, Marfo N, Tangamornsuksan W, Steen JP, McKechnie T, et al. Large language models for chatbot health advice studies: a systematic review. JAMA Netw Open 2025 Feb 03;8(2):e2457879 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.57879] [Medline: 39903463]
22. Eysenbach G. CONSORT-EHEALTH: implementation of a checklist for authors and editors to improve reporting of web-based and mobile randomized controlled trials. Stud Health Technol Inform 2013;192:657-661. [Medline: 23920638]
23. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015 Jan 02;350:g7647 [FREE Full text] [doi: 10.1136/bmj.g7647] [Medline: 25555855]
24. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. Cochrane Handbook for Systematic Reviews of Interventions. 1st ed. Wiley. USA: Wiley; 2019.
25. Mbuagbaw L, Lawson DO, Puljak L, Allison DB, Thabane L. A tutorial on methodological studies: the what, when, how and why. BMC Med Res Methodol 2020 Sep 07;20(1):226 [FREE Full text] [doi: 10.1186/s12874-020-01107-7] [Medline: 32894052]

XSL•FO
RenderX

26. Song E, Bernstein M, Lay P, Druart L, Dibble E, Lourenco A, et al. Accessing AI mammography reports impacts patient follow-up behaviors: the unintended consequences of including AI in patient portals. NPJ Digit Med 2025 Dec 03;8(1):761 [FREE Full text] [doi: 10.1038/s41746-025-02201-0] [Medline: 41339742]
27. Orpwood G. Over one in three using AI Chatbots for mental health support, as charity calls for urgent safeguards. Mental Health UK. URL: https://mentalhealth-uk.org/blog/over-one-in-three-using-ai-chatbots-for-mental-health-support-as-charity-calls-for-urgent-safeguards/ [accessed 2026-02-24]
28. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75:40-46 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
29. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
30. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021 Mar 29;372:n160 [FREE Full text] [doi: 10.1136/bmj.n160] [Medline: 33781993]
31. Blease C, Druart L. Difficulty is not an ethical exemption: placebo controls and epistemic responsibility in digital therapeutics. JME Practical Bioethics 2026 Feb 12;2(1):e000129-e000141. [doi: 10.1136/jmepb-2026-000129]
32. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008 Apr 26;336(7650):924-926 [FREE Full text] [doi: 10.1136/bmj.39489.470347.AD] [Medline: 18436948]

Abbreviations

AI: artificial intelligence
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
GRADE: Grading of Recommendations Assessment, Development and Evaluation
LLM: large language model
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Review and Meta-Analysis
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols

Edited by J Sarvestan; submitted 29.Dec.2025; peer-reviewed by L He, H Su; comments to author 04.Feb.2026; revised version received 24.Feb.2026; accepted 03.Mar.2026; published 17.Mar.2026.

Please cite as:
Druart L, Faria V, Annoni M, Torous J, Pontén M, Blease C
Investigating Placebos and Controls Used in Large Language Model–Based Chatbot Intervention Trials: Protocol for a Methodological Review
JMIR Res Protoc 2026;15:e90507
URL: https://www.researchprotocols.org/2026/1/e90507
doi:10.2196/90507
PMID:

©Leo Druart, Vanda Faria, Marco Annoni, John Torous, Moa Pontén, Charlotte Blease. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Mathematical Prediction Models for Sentinel Node Status in Early-Stage Breast Cancer: Protocol for a Systematic Review

Justin James[1,2*]; Kirti Mehta[1*]; Mohammadali Ahmadipour[3*]; Tegan Ormston[1*]; Emily Schembri[1*]; Michael Law[1*]; Shomik Sengupta[1,2*]; Christobel Saunders[4*]

[1]Eastern Health, Melbourne, Victoria, Australia

[2]Eastern Health Clinical School, Faculty of medicine, nursing and health sciences, Monash University, Melbourne, Victoria, Australia

[3]Deakin University, Melbourne, Australia

[4]Universtiy of Melbourne, Melbourne, Australia

[*]all authors contributed equally

Corresponding Author:
Justin James
Eastern Health
Breast and Cancer Centre, Maroondah hospital,
Davey Drive, Ringwood East
Melbourne, Victoria, 3135
Australia
Phone: 61 0439709752
Email: justin.james@monash.edu

Abstract

Background: The status of the axilla remains a significant prognostic factor and influences adjuvant systemic and locoregional treatment choices in early-stage breast cancer (EBC). Sentinel node (SN) biopsy continues to be the preferred technique for establishing axillary nodal status in clinically node-negative EBC. A multivariable prediction model with adequate accuracy and generalizability has been explored as a potential alternative to SN.

Objective: This systematic review aims to evaluate the predictive performance, methodological quality, and risk of bias associated with the available mathematical models (MMs), excluding artificial intelligence (AI)–based models, for predicting SN status in patients with EBC.

Methods: A systematic search will be conducted across PubMed, Cochrane CENTRAL, and Embase to identify studies reporting the development of an SN status prediction model. Only studies that report SN status using mathematical modeling techniques will be included. Two independent reviewers will screen the search results and extract data from the included articles. The primary outcome of this systematic review is to evaluate the methodological adequacy and generalizability of individual MMs and compare the reported predictive performances of methodologically robust MMs. The secondary objective is to identify key predictive factors contributing to SN status prediction in MMs. A narrative synthesis of all the included studies will be undertaken. The details of this protocol are accessible on PROSPERO, where it was registered on January 23, 2025. Ethics approval is not required for this study because only published data will be analyzed.

Results: Funding for this review was obtained in 2022. The literature search was completed on December 15, 2023, and screening began in December 2023. Data extraction and assessment using the Prediction Model Risk of Bias Assessment Tool was completed by December 2025, with synthesis planned for March 2026. Of the 3458 screened records, 122 (3.5%) were selected for data extraction. Results will be prepared for submission for a peer review and publication in mid-2026.

Conclusions: This review will provide a consolidated evaluation of non–machine learning MMs for predicting SN status in EBC. By clarifying the predictive performance and methodological quality of traditional statistical approaches, the findings will serve as a benchmark against which emerging AI-based tools can be compared. This review is also expected to identify predictors that consistently contribute to accurate modeling, informing the development of future statistical and AI-enhanced prediction tools.

Trial Registration: PROSPERO CRD42025637632; https://www.crd.york.ac.uk/PROSPERO/view/CRD42025637632

International Registered Report Identifier (IRRID): DERR1-10.2196/82523

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e82523) doi:10.2196/82523

KEYWORDS

breast cancer; sentinel node; nomograms; systematic review; artificial intelligence; AI

Introduction

Background

Sentinel node (SN) biopsy (SNB) is the procedure of choice for assessing the axilla in all patients presenting with clinically node-negative early-stage breast cancer (EBC) [1]. Information on nodal status obtained through SNB contributes to prognostication and can influence the selection of adjuvant systemic and locoregional treatment options [2]. However, as more systemic treatment decisions rely on tumor biology, SN status information is becoming increasingly redundant for many patient subgroups [3]. Importantly, SNB has not demonstrated any proven therapeutic benefit [4]. Moreover, SNB has been shown to adversely affect the quality of life after EBC treatment [5]. Therefore, excluding SNB from the EBC treatment protocol could lead to substantial improvements in workflow efficiency and cost-effectiveness. The accuracy of radiological evaluations of the axilla, combined with the significant impact of tumor biology on systemic and locoregional therapy decisions, has prompted many practitioners to explore alternatives to SNB [4,6-8].

The overall false-negative rate (FNR) of SNB is 8.4% (range 0%-29%), according to the American Society of Clinical Oncology expert panel [9]. It was hypothesized that a reliable prediction model could replace SNB if its accuracy (specifically its FNR and negative predictive value) was comparable to that of SNB [10]. Consequently, many researchers have attempted to predict SN status using preoperative data, and some of these SN status prediction models are well known and have been validated [11,12].

Recent trials and international consensus statements have further accelerated interest in axillary de-escalation for selected patients with clinically node-negative EBC. Randomized studies, such as the Sentinel Node versus Observation After Axillary Ultrasound (SOUND) and Intergroup-Sentinel-Mamma (INSEMA) trials, have demonstrated that omitting SNB in carefully selected low-risk populations is oncologically safe and associated with reduced morbidity [13,14]. Reflecting this evolving evidence base, axillary management recommendations from professional societies, including the American Society of Clinical Oncology, have progressively incorporated principles of de-escalation for selected patients [1,15].

All de-escalation approaches inherently accept low-volume axillary disease that will not be confirmed or excluded in de-escalated patients. Importantly, the St. Gallen International Breast Cancer Consensus has acknowledged that contemporary de-escalation strategies implicitly accept an FNR of approximately 10% to 15% [16].

Historically, nomogram-based mathematical models (MMs) for predicting SN status were not adopted for routine clinical decision-making because their predictive performance was considered insufficient to replace SNB [17]. However, as the paradigm shifts toward accepting low-volume axillary disease, nomograms may assume a new role in aiding patient selection for axillary de-escalation [18,19].

Additionally, more recent studies using advanced modeling approaches—particularly artificial intelligence (AI), radiomics, and genomic data—have reported substantially higher discriminatory performance, with area under the receiver operating characteristic curve values frequently exceeding 0.80 [20-22]. The number of these models has surged recently and has led to optimism toward the development of a noninvasive alternative to SNB [23].

Ideally, a prediction model should adequately report all components of the study so that the results can be interpreted correctly and the model's applicability in a real-world clinical setting can be inferred. Structured and complete reporting, in accordance with the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) statement and the PROBAST (Prediction Model Risk of Bias Assessment Tool), is desirable for all prediction model studies [24,25]. Such comprehensive reporting will help replicate and validate the reported models and combine the results in a meta-analysis. Given the significant differences in the reporting parameters, modified standards (TRIPOD+AI and PROBAST+AI) for reporting studies of AI-based models (AIM) have been proposed [26,27].

Therefore, a comprehensive review of currently available models attempting to predict SN status is critical for assessing their quality, transparency, usability, methodological rigor, and risk of bias. This will aid clinicians looking to use a model-based selection strategy for axillary de-escalation. As the parameters used to determine the robustness of prediction models differ between the 2 types—MMs and AIMs—it is challenging to review them together and compare their results within a single systematic review. Although newer AIMs are claimed to be more effective at predicting SN status, MMs remain widely available and could be adopted as decision aids. Moreover, understanding the strengths and limitations of existing MMs will provide a benchmark against which emerging AI-based or hybrid prediction tools can be compared. Identifying critical variables that consistently contribute to SN status prediction in these traditional MMs will also be helpful for future research. These variables may then be selected for the development of more advanced models based on newer modeling techniques.

Objectives

The primary aim of this study is to systematically review available MMs for predicting SN status in patients with clinically node-negative EBC undergoing primary surgery. We will evaluate the methodological adequacy, suitability, and generalizability of individual MMs and compare the reported predictive performances of methodologically robust MMs. We will conduct a narrative synthesis of robust MMs suitable for

XSL•FO
RenderX

clinical use and identify key predictive factors influencing SN status prediction in these MMs.

Methods

Overview

This systematic review was designed to meet the reporting standards recommended by TRIPOD Systematic Reviews and Meta-Analyses [28]. The protocol was prepared in accordance with the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) 2015 statement [29] and was registered with PROSPERO (CRD42025637632) on January 23, 2025.

Articles describing primary research aimed at developing multivariable models to predict SN status in patients with EBC undergoing primary surgery were eligible for inclusion. Eligible studies were required to include at least 2 predictors, use any statistical or mathematical modeling methodology, and be published up to December 15, 2023. MMs are defined as statistical, regression-based, or probabilistic modeling techniques that estimate the association between preoperative clinicopathological variables and SN status using prespecified mathematical functions. Included model types are logistic regression models, generalized linear models, generalized additive models, Bayesian statistical models, point-based or rule-based scoring systems derived from statistical modeling, and nomograms derived from these techniques. These models share the following characteristics: (1) prespecified functional form, (2) estimation via statistical inference, and (3) interpretability based on coefficients or explicit equations.

This time frame for including studies was selected to achieve the most comprehensive range of articles that reflect various strategies for predicting SN status. We will include studies of any design and data source, if the predicted outcome is SN status in a population with a clinically node-negative EBC. Further details about the inclusion criteria are provided in Textbox 1.

Textbox 1. Inclusion and exclusion criteria for the study.

Inclusion criteria

- Study design: primary report on development of mathematical models, with or without validation, based on data from experimental (randomized controlled trials) or observational (cohort, case control, and case cohort) studies
- Variables: clinical, radiological, histopathological, or any other measured variables
- Outcome: axillary node status (micrometastasis or macrometastasis) identified using any of the axillary assessment methods (sentinel node biopsy, axillary dissection, axillary sampling, or a combination of these methods) and any of the standard histopathological or immunohistochemical methods
- Population: all or a subgroup of patients with clinically node-negative early-stage breast cancer undergoing surgical primary axillary assessment before any systemic treatment
- Modeling: mathematical modeling techniques including but not restricted to regression-based approaches
- Data source: registry data, hospital and other databases, or clinical trial data

Exclusion criteria

- Predictive factor studies: articles describing predictive factors without developing a prediction model
- Validation studies: articles describing the validation of an existing model on new data
- Outcome: models predicting further nodes after a positive sentinel node biopsy and N2 status
- Population: models predicting nodal status after neoadjuvant systemic therapy or in patients with in situ breast cancer
- Modeling: models based on artificial intelligence and machine learning techniques
- Article type and quality: secondary research, conference abstracts, non-English publications, and papers without available full text

Articles will be excluded if they report prediction models developed using machine learning (ML) or AI techniques. These exclusion approaches include but are not limited to (1) decision tree–based methods and ensemble algorithms (eg, classification and regression trees, random forests, Adaptive Boosting, and Extreme Gradient Boosting), (2) support vector machines, (3) neural networks and deep learning models, (4) k-nearest neighbor classifiers, (5) naïve Bayes classifiers, and (6) any nonparametric or high-dimensional modeling approaches in which the functional form is learned algorithmically rather than prespecified. Hybrid mathematical ML approaches will also be excluded unless the ML component is used solely for variable preselection and the final prediction model is purely mathematical. These approaches rely on automated pattern recognition or feature learning and therefore fall within the scope of AI and ML rather than traditional mathematical modeling. We will exclude studies focused on predictive factors, secondary research, conference abstracts, or studies lacking full-text availability in English. Further details on the exclusion criteria are provided in Textbox 1.

A systematic literature search will be conducted in 3 major publicly available electronic medical databases—PubMed, Embase, and Cochrane CENTRAL—to identify all eligible articles published until December 15, 2023. To maximize completeness and minimize the risk of missing eligible studies, we plan to supplement database searches with both backward and forward citation screening. The combination of targeted database searching with citation tracking is expected to retrieve

XSL•FO
RenderX

all relevant, high-quality mathematical prediction models; therefore, no additional gray literature sources or trial registries will be searched.

The search strategy was built using keywords related to breast cancer (ie, “breast malignancy” and “breast neoplasm”), SN status (ie, “sentinel node” and “lymph node”), and prediction modeling (ie, “nomograms” and “predictive models”). We selected 10 articles that met our eligibility criteria. A trial run of our search confirmed that it retrieved all 10 chosen studies. The final search strategy is presented in the Multimedia Appendix 1.

Study records, including titles and abstracts, will be imported into Covidence (Veritas Health Innovation) [30], a web-based collaboration platform that streamlines the production of systematic and other literature reviews. It tracks and archives all activities throughout the reviewing process. Once eligible studies are identified, full-text articles will be uploaded into the program for screening and data extraction using a customized template. Two authors (JJ and MA) will independently screen titles and abstracts to identify eligible studies according to the criteria and subsequently review the full texts of potentially relevant studies. If necessary, any disagreements between the reviewers will be resolved by consensus or through consultation with additional investigators (CS and KM). The study flow will be illustrated in a PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flowchart.

We will perform double data extraction for all included articles. One of the 2 reviewers (MA and TO) will extract data from each article using a predefined data extraction template, while the lead investigator (JJ) will extract data from all included articles and compare the extracted data. If necessary, a third reviewer (CS) will adjudicate any disagreements in data extraction. The data extraction form will be tested on the first 10 studies and revised as needed. If critical data are required, the authors of the articles will be contacted for additional information and clarification.

The lead investigator (JJ) will maintain all data and records, which will be stored on a shared, secure Covidence platform accessible to all investigators. Data extraction will be guided by TRIPOD adherence criteria, PROBAST, and CHARMS (Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies). Table 1 presents the main categories of the extracted data.

Table 1. Extracted data categories and their definitions.

Data category	Definition
Study design	Type of study (eg, model development, temporal validation, and geographic validation), methodological design (retrospective, prospective, and registry based), and year of publication
Setting	Geographic location, clinical setting (hospital, multicenter, and registry), and period of data collection
Study population and data source	Eligibility criteria, recruitment method, source of data (hospital records, registry, and cohort study), and population characteristics relevant to model applicability
Patient characteristics	Baseline clinicopathological features (eg, age distribution; tumor size; histology; grade; estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 status; and lymphovascular invasion)
Sample size	Number of participants included for model development, (if applicable) internal or external validation, and number of outcome events
Outcome	Definition of sentinel lymph node positivity (eg, macrometastasis >2 mm), method of outcome assessment (SNB[a] with ALND[b] and SNB without ALND), and event rate
Predictors	Candidate predictors examined, final predictors retained in the model, definitions of predictors, and handling of missing data
Internal validation	Method and sample used for internal validation (eg, bootstrap, cross validation, and train-test split), and internal performance metrics
External validation	Presence and type of external validation (temporal, geographic, and independent cohort) and characteristics of the validation dataset
Predictive performance measures	Measures of model discrimination (eg, area under the receiver operating characteristic curve), calibration (eg, calibration slope, intercept, and calibration plot), and classification metrics (eg, sensitivity, specificity, positive predictive value, and negative predictive value)

[a]SNB: sentinel node biopsy.

[b]ALND: axillary lymph node dissection.

The form will include instructions for reviewers on assessing the methodological and analytic steps used to develop the model and extracting the relevant statistical parameters presented in the articles. Predictors will be identified from the main article and its associated supplemental files. We will list all the predictors (1) examined as candidates, (2) tested in the model, and (3) included in the final model. The final data extraction form is presented in Multimedia Appendix 2.

The risk of bias in individual studies will be assessed using the PROBAST tool. Before beginning the formal risk-of-bias assessment, all reviewers will complete a calibration exercise using the PROBAST tool on a sample of 3 to 5 included studies. Differences in the interpretation of signaling questions and

XSL•FO
RenderX

domain-level judgments will be discussed, and decision rules will be refined to ensure consistency across reviewers. Two reviewers will independently assess the risk of bias for each included study using the PROBAST tool (Multimedia Appendix 3). Disagreements at the signaling question, domain, or overall judgment level will be resolved through discussion. If consensus cannot be reached, a third senior reviewer will adjudicate. Agreement rates will be monitored, and additional calibration will be undertaken if needed. Results from the PROBAST assessment will be fully integrated into the narrative synthesis. Studies judged to be at high overall risk of bias will be highlighted, and their findings will be interpreted with caution. Patterns of bias across domains (participants, predictors, outcomes, and analysis) will be summarized descriptively. Where appropriate, studies at low risk of bias will be prioritized when interpreting the evidence base. No statistical pooling is planned; therefore, PROBAST will primarily guide the qualitative interpretation of study validity and the confidence placed in individual models.

We will conduct a narrative synthesis of the extracted data. Descriptive statistics and visual plots will be used to summarize the data. Categorical data on reporting, methodological practices, and risk of bias will be presented as counts and percentages. The distribution of continuous data, such as sample size and the number of predictors, will be assessed and reported using the mean and SD for normally distributed data and the median with the IQRs for nonnormally distributed data.

The risk-of-bias assessment will be summarized and graphically illustrated for each PROBAST domain, along with an overall risk-of-bias judgment. Results will be stratified by study type (model development only, model development with internal validation, and model development with external validation). Meta-bias will not be investigated, and the strength of the body of evidence will not be assessed in this study.

Protocol amendments will be listed and made available on the PROSPERO registration. Each amendment will include the date, description, and rationale.

Ethical Considerations

Ethics approval is not required for this study as only deidentified data will be analyzed.

Dissemination of Findings

The findings of this systematic review will be published in an open-access journal and shared at various scientific conferences.

Results

Funding for this review was obtained in 2022. The literature search was completed on December 15, 2023, and screening began in December 2023. Data extraction and PROBAST assessment were completed by December 2025, with synthesis planned in March 2026. As of manuscript submission, 3458 records have been screened, and 122 (3.5%) full-text articles were selected for data extraction (Figure 1). Data extraction is ongoing, and the results are expected to be prepared for submission by mid-2026.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart showing the progress of data extraction. AI: artificial intelligence.

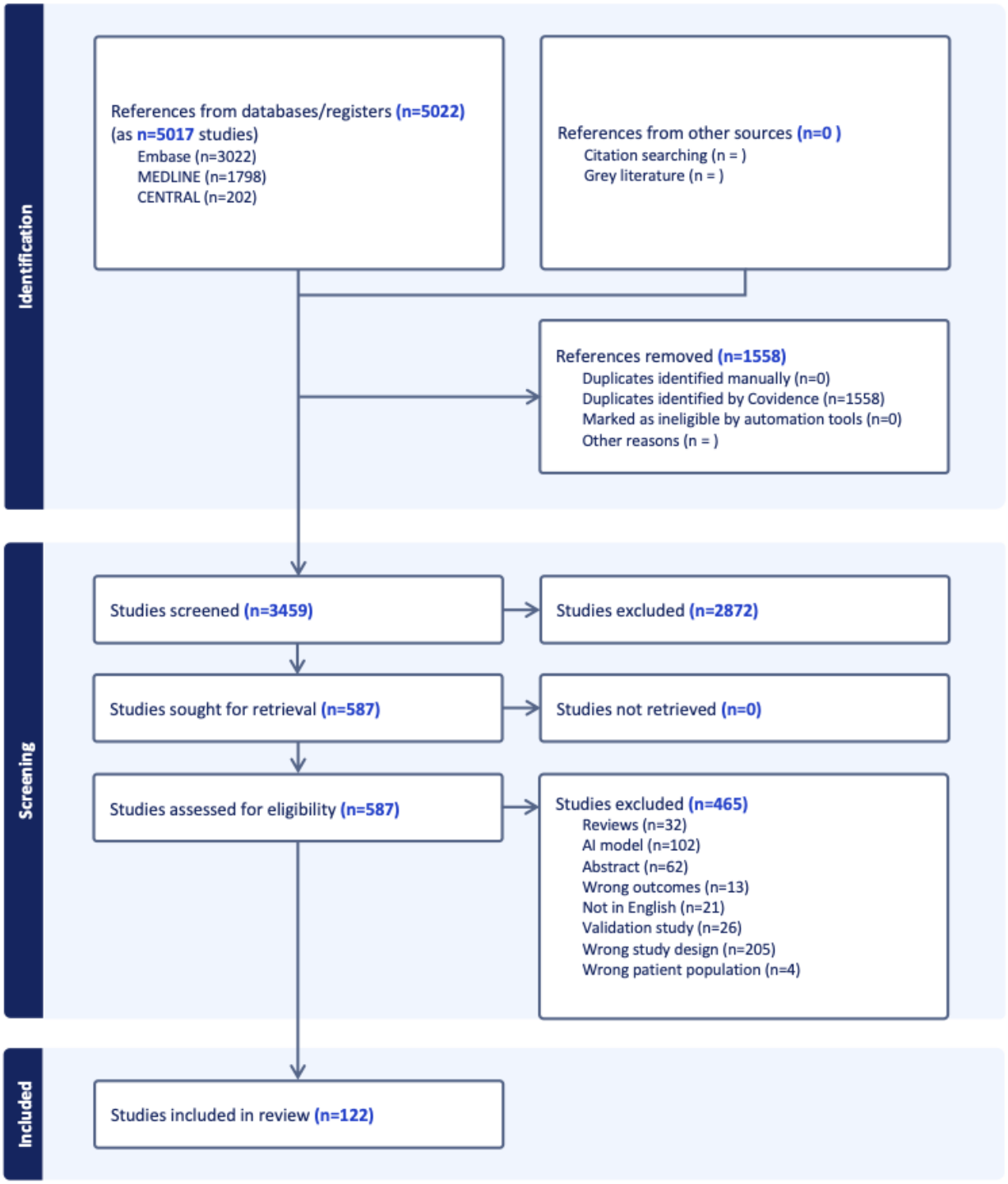

Discussion

Anticipated Findings

This systematic review is expected to provide the first consolidated synthesis of MMs developed to predict SN status in patients with EBCs. By focusing on MMs, this review will clarify the performance, methodological quality, and potential clinical utility of these traditional statistical models, which rely on routinely collected clinicopathological variables. Therefore, this review fills a gap by examining the evidence base for interpretable, equation-driven prediction tools that promise to aid critical decision-making in the management of EBCs.

A major strength of this protocol is the use of clear operational definitions that distinguish MMs from AIMs, thereby enabling reproducible screening and reducing misclassification. The structured data extraction framework, predefined eligibility criteria and decision rules, and planned use of PROBAST to assess risk of bias across all relevant domains enhance methodological rigor. Supplementing the electronic database searches with forward and backward citation screening increases

XSL•FO
RenderX

the likelihood of capturing older or less consistently indexed regression-based models.

The findings of this review will consolidate the strengths and benefits of traditional statistical modeling based on routinely measured clinicopathological variables. By establishing a comprehensive benchmark of model performance and methodological quality, this review will provide a standard against which the performance of emerging AI-based prediction tools can be compared. The synthesis is expected to help determine whether statistical models remain adequate for predicting SN status or whether AI approaches offer meaningful improvements. It will inform clinicians about appropriate models for use with patient groups of interest and in clinical scenarios in which they intend to apply the model. Furthermore, by identifying predictors that consistently contribute to accurate statistical models, this review may guide the selection of key variables for inclusion in future AIMs or hybrid prediction models, supporting the development of efficient, interpretable, and clinically practical tools.

This systematic review is intended to provide a methodological and performance benchmark for MMs of SN status rather than to evaluate their clinical implementation or long-term oncological impact. By focusing exclusively on model development studies, this review will summarize predictive performance, reporting quality, and risk of bias and describe the clinical settings and patient subpopulations (eg, biologically defined subgroups or specific diagnostic pathways) for which individual models were developed. These contextual details will be incorporated into the narrative synthesis to inform the interpretation of model applicability and generalizability.

The findings are expected to support comparison with emerging AIMs or hybrid prediction models, all of which should undergo transparent development, internal validation, rigorous external validation, and prospective testing. As axillary de-escalation strategies ultimately aim to demonstrate oncological safety rather than acceptable FNRs alone, any prediction model intended to guide omission of SNB would require evaluation in randomized controlled trials to establish effects on recurrence and survival. However, such downstream clinical evaluation lies beyond the remit of this review.

By design, this review excludes AIMs that can limit direct comparison across all contemporary modeling techniques. Restricting the search to 3 major databases may omit models published in nontraditional biomedical journals, although citation tracking mitigates this risk. Heterogeneity in study populations, predictor definitions, modeling strategies, and reporting standards may limit comparability between included studies and preclude meta-analysis.

Conclusions

This systematic review will synthesize and critically appraise MMs predicting SN status in EBCs. By applying rigorous eligibility criteria, structured data extraction methods, and formal risk-of-bias assessment using PROBAST, this review will provide a clear and comprehensive evaluation of the current landscape of statistical prediction tools. The findings will help define the role of mathematical modeling in contemporary axillary decision-making, establish a benchmark for comparison with AI-based approaches, and identify essential predictors that consistently contribute to nodal status prediction, thereby informing future model development and validation.

Acknowledgments

The authors acknowledge the use of an artificial intelligence language model, ChatGPT (version 5.1; OpenAI), to assist with language editing during manuscript preparation. All outputs were critically reviewed, edited, and verified by the authors to ensure accuracy, integrity, and appropriateness for publication.

Funding

JJ received a research grant from the Eastern Health Foundation through the Linda Williams Memorial Research and Innovation Grant for the project “A Model to Select Early Breast Cancers Where Sentinel Node Biopsy Can Be Avoided.” The sponsor had no role in the study design; the collection, analysis, or interpretation of data; the writing of the manuscript; or the decision to submit the manuscript for publication.

Data Availability

The extraction framework and the extracted data may be made available on request to support transparency and future methodological research.

Authors' Contributions

JJ contributed to conceptualization and acted as the guarantor for this review.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.

XSL•FO
RenderX

[DOCX File , 16 KB - resprot_v15i1e82523_app1.docx]

Multimedia Appendix 2
Data extraction template.
[DOCX File , 41 KB - resprot_v15i1e82523_app2.docx]

Multimedia Appendix 3
Quality assessment template.
[DOCX File , 51 KB - resprot_v15i1e82523_app3.docx]

References

1. Brackstone M, Baldassarre FG, Perera FE, Cil T, Chavez Mac Gregor M, Dayes IS, et al. Management of the axilla in early-stage breast cancer: Ontario Health (Cancer Care Ontario) and ASCO guideline. J Clin Oncol 2021 Sep 20;39(27):3056-3082 [FREE Full text] [doi: 10.1200/JCO.21.00934] [Medline: 34279999]
2. Marrazzo A, Taormina P, Gebbiab V, David M, Riili I, Lo Gerfo D, et al. Is sentinel lymph node biopsy more accurate than axillary dissection for staging nodal involvement in breast cancer patients? Chir Ital 2007;59(5):693-699. [Medline: 18019642]
3. Goldhirsch A, Ingle JN, Gelber RD, Coates AS, Thürlimann B, Senn HJ, Panel members. Thresholds for therapies: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2009. Ann Oncol 2009 Aug;20(8):1319-1329 [FREE Full text] [doi: 10.1093/annonc/mdp322] [Medline: 19535820]
4. Gentilini O, Veronesi U. Abandoning sentinel lymph node biopsy in early breast cancer? A new trial in progress at the European Institute of Oncology of Milan (SOUND: Sentinel node vs Observation after axillary UltraSouND). Breast 2012 Oct;21(5):678-681 [FREE Full text] [doi: 10.1016/j.breast.2012.06.013] [Medline: 22835916]
5. James J, Law M, Sengupta S, Saunders C. Assessment of the axilla in women with early-stage breast cancer undergoing primary surgery: a review. World J Surg Onc 2024 May 09;22(1):127 [FREE Full text] [doi: 10.1186/S12957-024-03394-6]
6. Reimer T, Stachs A, Nekljudova V, Loibl S, Hartmann S, Wolter K, et al. Restricted axillary staging in clinically and sonographically node-negative early invasive breast cancer (c/iT1-2) in the context of breast conserving therapy: first results following commencement of the Intergroup-Sentinel-Mamma (INSEMA) trial. Geburtshilfe Frauenheilkd 2017 Feb;77(2):149-157 [FREE Full text] [doi: 10.1055/s-0042-122853] [Medline: 28331237]
7. van Roozendaal LM, Vane M, van Dalen T, van der Hage JA, Strobbe LJ, Boersma LJ, et al. Clinically node negative breast cancer patients undergoing breast conserving therapy, sentinel lymph node procedure versus follow-up: a Dutch randomized controlled multicentre trial (BOOG 2013-08). BMC Cancer 2017 Jul 01;17(1):459 [FREE Full text] [doi: 10.1186/s12885-017-3443-x] [Medline: 28668073]
8. Tucker NS, Gillanders WE, Eberlein T, Aft R, Margenthaler J. Abstract OT3-5-01: a prospective, randomized trial of sentinel lymph node biopsy versus no additional staging in patients with T1-T2 invasive breast cancer and negative axillary ultrasound. Ongoing Clin Trials 2015;3:3-5 [FREE Full text] [doi: 10.1158/1538-7445.sabcs14-ot3-5-01]
9. Lyman GH, Giuliano AE, Somerfield MR, Benson AB, Bodurka DC, Burstein HJ, et al. American Society of Clinical Oncology guideline recommendations for sentinel lymph node biopsy in early-stage breast cancer. J Clin Oncol 2005 Oct 20;23(30):7703-7720 [FREE Full text] [doi: 10.1200/jco.2005.08.001]
10. Chen J, Chen J, Yang BL, Liu ZB, Huang XY, Liu GY, et al. Predicting sentinel lymph node metastasis in a Chinese breast cancer population: assessment of an existing nomogram and a new predictive nomogram. Breast Cancer Res Treat 2012 Oct;135(3):839-848 [FREE Full text] [doi: 10.1007/s10549-012-2219-x] [Medline: 22941537]
11. Bevilacqua J, Kattan M, Fey JV, Cody HS, Borgen PI, Van Zee KJ. Doctor, what are my chances of having a positive sentinel node? A validated nomogram for risk estimation. J Clin Oncol 2007 Aug 20;25(24):3670-3679 [FREE Full text] [doi: 10.1200/jco.2006.08.8013]
12. Breast cancer nomogram to predict positive sentinel lymph nodes, without neoadjuvant chemotherapy. The University of Texas MD Anderson Cancer Center. URL: https://www3.mdanderson.org/app/medcalc/bc_nomogram3/index.cfm?pagename=sln [accessed 2025-01-24]
13. Gentilini OD, Botteri E, Sangalli C, Galimberti V, Porpiglia M, Agresti R, SOUND Trial Group. Sentinel lymph node biopsy vs no axillary surgery in patients with small breast cancer and negative results on ultrasonography of axillary lymph nodes: the SOUND randomized clinical trial. JAMA Oncol 2023 Nov 01;9(11):1557-1564 [FREE Full text] [doi: 10.1001/jamaoncol.2023.3759] [Medline: 37733364]
14. Reimer T, Stachs A, Veselinovic K, Kühn T, Heil J, Polata S, et al. Axillary surgery in breast cancer - primary results of the INSEMA trial. N Engl J Med 2025 Mar 13;392(11):1051-1064. [doi: 10.1056/NEJMoa2412063] [Medline: 39665649]
15. Park KU, Vega RB, Shams S, Hawley J, Somerfield MR, Thompson AM, et al. sentinel lymph node biopsy in early-stage breast cancer: ASCO guideline clinical insights. JCO Oncol Pract (Forthcoming) 2025 Jun 18 [FREE Full text] [doi: 10.1200/OP-25-00447] [Medline: 40532132]

XSL•FO
RenderX

16. Burstein HJ, Curigliano G, Gnant M, Loibl S, Regan MM, Loi S, et al. Tailoring treatment to cancer risk and patient preference: the 2025 St Gallen International Breast Cancer Consensus Statement on individualizing therapy for patients with early breast cancer. Ann Oncol 2025 Dec;36(12):1433-1446 [FREE Full text] [doi: 10.1016/j.annonc.2025.09.007] [Medline: 41072918]
17. Patani N, Dwek M, Douek M. Predictors of axillary lymph node metastasis in breast cancer: a systematic review. Eur J Surg Oncol 2007 May;33(4):409-419 [FREE Full text] [doi: 10.1016/j.ejso.2006.09.003] [Medline: 17125963]
18. Tran HT, Pack D, Mylander C, Martino L, Rosman M, Tafra L, et al. Ultrasound-based nomogram identifies breast cancer patients unlikely to harbor axillary metastasis: towards selective omission of sentinel lymph node biopsy. Ann Surg Oncol 2020 Aug;27(8):2679-2686 [FREE Full text] [doi: 10.1245/s10434-019-08164-3] [Medline: 32026063]
19. Tran HT, Jackson RS. ASO author reflections: preoperative nomogram identifies patients for selective omission of SLNB in breast cancer-to promote implementation of the choosing wisely campaign recommendations. Ann Surg Oncol 2020 Aug;27(8):2687-2688 [FREE Full text] [doi: 10.1245/s10434-020-08288-x] [Medline: 32350715]
20. Soh CL, Shah V, Arjomandi Rad A, Vardanyan R, Zubarevich A, Torabi S, et al. Present and future of machine learning in breast surgery: systematic review. Br J Surg 2022 Oct 14;109(11):1053-1062 [FREE Full text] [doi: 10.1093/bjs/znac224] [Medline: 35945894]
21. Eldaly AS, Avila FR, Torres-Guzman RA, Maita K, Garcia JP, Serrano LP, et al. Radiomics and artificial intelligence in predicting axillary lymph node metastasis in breast cancer: a systematic review. Curr Med Imaging 2023;19(6):564-578 [FREE Full text] [doi: 10.2174/1573405618666220822093226] [Medline: 35996255]
22. Calabrese A, Santucci D, Landi R, Beomonte Zobel B, Faiella E, de Felice C. Radiomics MRI for lymph node status prediction in breast cancer patients: the state of art. J Cancer Res Clin Oncol 2021 Jun;147(6):1587-1597 [FREE Full text] [doi: 10.1007/s00432-021-03606-6] [Medline: 33758997]
23. Pesapane F, Mariano L, Magnoni F, Rotili A, Pupo D, Nicosia L, et al. Future directions in the assessment of axillary lymph nodes in patients with breast cancer. Medicina (Kaunas) 2023 Aug 25;59(9):1544 [FREE Full text] [doi: 10.3390/medicina59091544] [Medline: 37763661]
24. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med 2015 Jan 06;162(1):55-63 [FREE Full text] [doi: 10.7326/m14-0697]
25. Wolff RF, Moons KG, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019 Jan 01;170(1):51-58 [FREE Full text] [doi: 10.7326/M18-1376] [Medline: 30596875]
26. Collins GS, Moons KG, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378 [FREE Full text] [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
27. PROBAST+AI: assessing quality, risk of bias and applicability of diagnostic and prognostic prediction models based on AI or ML techniques. The Cochrane Collaboration. 2023. URL: https://abstracts.cochrane.org/2023-london/probastai-assessing-quality-risk-bias-and-applicability-diagnostic-and-prognostic [accessed 2025-01-24]
28. Snell KI, Levis B, Damen JA, Dhiman P, Debray TP, Hooft L, et al. Transparent reporting of multivariable prediction models for individual prognosis or diagnosis: checklist for systematic reviews and meta-analyses (TRIPOD-SRMA). BMJ 2023 May 03;381:e073538 [FREE Full text] [doi: 10.1136/bmj-2022-073538] [Medline: 37137496]
29. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 01;4(1):1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
30. Covidence. URL: https://covidence.org [accessed 2025-01-24]

Abbreviations

AI: artificial intelligence
AIM: artificial intelligence–based model
CHARMS: Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies
EBC: early-stage breast cancer
FNR: false-negative rate
INSEMA: Intergroup-Sentinel-Mamma
ML: machine learning
MM: mathematical model
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
PROBAST: Prediction Model Risk of Bias Assessment Tool
SOUND: Sentinel Node versus Observation After Axillary Ultrasound

XSL•FO
RenderX

SN: sentinel node
SNB: sentinel node biopsy
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by J Sarvestan; submitted 16.Aug.2025; peer-reviewed by L Huss, H Anandan; comments to author 24.Nov.2025; revised version received 23.Dec.2025; accepted 24.Dec.2025; published 23.Mar.2026.

Please cite as:
James J, Mehta K, Ahmadipour M, Ormston T, Schembri E, Law M, Sengupta S, Saunders C
Mathematical Prediction Models for Sentinel Node Status in Early-Stage Breast Cancer: Protocol for a Systematic Review
JMIR Res Protoc 2026;15:e82523
URL: https://www.researchprotocols.org/2026/1/e82523
doi:10.2196/82523
PMID:

©Justin James, Kirti Mehta, Mohammadali Ahmadipour, Tegan Ormston, Emily Schembri, Michael Law, Shomik Sengupta, Christobel Saunders. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Longitudinal Associations Between Biomarkers and Frailty in Older Adults: Protocol for a Scoping Review

Jo Woon Seok[1], BSN, PhD; Kyusol Shim[2], BSN; Yihan Yang[2]; Soomin Hong[3], BSN, OCN, PhD

[1]Department of Basic Nursing Science, College of Nursing, Korea University, Seongbuk-gu, Seoul, Republic of Korea

[2]College of Nursing, Ewha Womans University, Seodaemun-gu, Seoul, Republic of Korea

[3]Red Cross College of Nursing, Chung-Ang University, Dongjak-gu, Seoul, Republic of Korea

Corresponding Author:
Soomin Hong, BSN, OCN, PhD
Red Cross College of Nursing
Chung-Ang University
84, Heukseok-ro, Building 102, Room 712
Dongjak-gu, Seoul, 06974
Republic of Korea
Phone: 82 2 820 5645
Fax: 82 2 820 7961
Email: soominsnow@naver.com

Abstract

Background: Frailty is a progressive, dynamic clinical syndrome characterized by reduced physiological reserves and increased vulnerability to stressors, leading to adverse outcomes in older adults. Despite its clinical significance, routine detection of frailty remains challenging owing to subtle presentations and time constraints during assessments. Longitudinal studies are essential for capturing its biological trajectory and identifying early biomarkers.

Objective: This scoping review aims to systematically identify and map biomarkers—including inflammatory, hormonal, metabolic, genetic, and imaging-based markers—investigated for their longitudinal associations with frailty in older adults. Studies involving older adults with comorbid conditions (eg, cancer, infectious diseases, cardiovascular diseases) will also be included when frailty is assessed as an outcome.

Methods: We will conduct a comprehensive literature search in 6 databases (PubMed, CINAHL, Cochrane Library, Scopus, Web of Science, and Embase). We will include longitudinal studies that examine the association between objectively measured biomarkers and frailty, as assessed with validated tools, in older adults. Studies published in English, Korean, or Chinese will be considered, with no restriction on publication date.

Results: We identified 3315 records after duplicate removal in an exploratory search conducted in July 2024, indicating a substantial body of literature for review. We plan to complete this review by December 2026.

Conclusions: This review will identify and describe longitudinal biomarkers that precede frailty in older adults, highlighting their potential for personalized clinical application. Anticipated limitations include reliance on observational study designs, methodological heterogeneity, and limited causal inference.

International Registered Report Identifier (IRRID): PRR1-10.2196/83312

(JMIR Res Protoc 2026;15:e83312) doi:10.2196/83312

KEYWORDS

frailty; biomarkers; aged; longitudinal studies; systematic review

Introduction

Background

The rapid growth of the global older adult population has made frailty an emerging challenge in geriatric medicine and public health [1]. Frailty is a clinical syndrome, distinct from disability or comorbidity, characterized by reduced physiological reserves. This reduction increases vulnerability to stressors and results in functional decline in daily activities [2,3].

Frailty in community-dwelling older adults has been associated with adverse clinical outcomes, including falls, disability,

XSL•FO
RenderX

hospitalization, and death [4,5]. It also negatively affects the health of individuals with cardiovascular disease, infectious diseases (eg, COVID-19, pneumonia), or cancer because it strongly predicts treatment intolerance [6], more frequent complications [7], and increased mortality [8]. These outcomes emphasize the clinical importance of timely frailty identification in older adults and suggest that biomarkers may improve early detection strategies [9,10].

Although frailty is clearly associated with adverse outcomes, routine detection remains challenging because of subtle or nonspecific presentations and time constraints during assessments, particularly in older adult care [8,11-13]. The multisystem and dynamic nature of frailty reduces the effectiveness of single-indicator approaches. This highlights the need for objective, scalable biomarkers that enable rapid and reproducible identification. In recent years, the unmet need has led to increased interest in biological markers that may offer a more accurate and less subjective approach to predicting, diagnosing, and monitoring frailty in older adults [14-16]. Several systematic reviews and meta-analyses have shown that the levels of cytokines, including interleukin-6, tumor necrosis factor-α, and C-reactive protein, are consistently elevated in frail older adults [10,17]. Beyond cytokines and inflammation, interest has increased in hormonal, metabolic, and molecular markers, as well as multimarker panels. Omics-based approaches (proteomics and transcriptomics) have been proposed to more accurately capture the heterogeneous and dynamic trajectories of frailty than any single biomarker. European consortia such as the FRAILOMIC initiative are conducting comprehensive profiling of genomic, proteomic, and clinical biomarkers to improve the prediction of frailty onset and progression, as well as the potential therapeutic response [18,19].

Despite growing interest in frailty biomarkers, most previous studies have been cross-sectional, primarily comparing biomarker levels between frail and robust individuals at a single point in time [17,20]. These explorations have highlighted inflammatory and metabolic dysregulation; however, because they rely on static measurements, they cannot capture biomarker dynamics or clarify whether these changes precede or result from frailty progression. This limitation reinforces the need for longitudinal evidence to delineate causal pathways and identify biomarkers that predict frailty trajectories.

Clinical inconsistency in frailty definitions, biomarker assays, and analytical techniques has further increased heterogeneity across studies [14,19]. For example, frailty has been operationalized using different instruments such as the Fried phenotype (which focuses on physical manifestations) and the Frailty Index (which aggregates deficits across multiple domains), leading to discordant classifications of the same individuals [14]. Similarly, inflammatory biomarkers such as interleukin-6 and C-reactive protein have been measured using assays with varying sensitivities (eg, standard enzyme-linked immunosorbent assay vs high-sensitivity assays) and inconsistent cutoff values, complicating cross-study comparisons [19]. These discrepancies have impeded the comparability and replication of the findings. The performance of biomarkers has been erratic because sensitivity and specificity have been affected by complicating factors, including comorbidity, drug use, and population variation. Large-scale longitudinal studies remain limited, which restricts our understanding of dynamic biomarker trajectories and their predictive value [21,22]. Consequently, no single biomarker panel has demonstrated sufficient robustness for routine clinical use in identifying frailty in older adults [23]. These gaps indicate the need for a systematic synthesis of longitudinal evidence to clarify consistencies, identify weaknesses, and guide future research.

Objectives

This scoping review will systematically map longitudinal evidence on biomarkers associated with frailty in older adults. Existing evidence has identified the relevance of inflammatory and molecular markers; however, most studies have focused on single biological pathways or relied on cross-sectional associations. In contrast, we aim to describe findings across a broader range of biomarker types to reflect the biological complexity of frailty. Rather than following a hypothesis-driven approach, we will present a structured framework for interpreting and classifying diverse biomarkers (eg, inflammatory markers, hormonal mediators, genetic signatures, imaging-based measures) evaluated across diverse longitudinal study designs (eg, prospective cohort studies, follow-up assessment, and nested case-control analyses) and various demographic groups (eg, sex/gender, age strata, and clinical subpopulations such as heart failure or chronic kidney disease cohorts). This synthesis aims to identify clinically promising biomarkers, describe consistent trends, reveal key knowledge gaps to inform future research, and provide evidence-based suggestions for developing personalized frailty care strategies.

Methods

Study Design

We will conduct a scoping review to identify and review biomarkers with longitudinal associations with frailty, following the Joanna Briggs Institute (JBI) Manual for Evidence Synthesis guidelines [24]. We will report the review in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guideline to ensure transparency and methodological rigor [25]. The review questions are as follows:

1. What types of biomarkers have been reported to be longitudinally associated with frailty in older adults?
2. What are the follow-up durations used in studies examining the association between biomarkers and frailty?
3. How do variations in longitudinal study design, frailty definitions and assessment tools, and key demographic factors (eg, age and sex) differ across studies?
4. What statistical methods have been employed to analyze the relationship between biomarkers and frailty over time?

Eligibility Criteria

The review will include studies that meet the following criteria, structured according to population, concept, and context formats.

Population

Individuals aged 60 years or older with frailty were evaluated using defined tools, including the Frailty Index, Fried Frailty

XSL•FO
RenderX

Phenotype, or laboratory-based frailty indices. Frailty was measured as a dependent variable rather than as a risk factor for other health outcomes such as mortality, hospitalization, and disability.

Concept

Biomarkers are defined as objective measures that represent cellular and organismal phenomena in the human body at a specific time [26]. However, the review will include studies that measured biomarkers directly reflecting the physiological mechanisms of frailty or its risk factors based on the pathogenesis of frailty [27,28]. Biomarkers of interest include hormones, inflammatory proteins, gene expression, and imaging markers. Studies that assessed biomarkers using methods not based on biological fluids (eg, body composition data from InBody or bioelectrical impedance analysis) were excluded to maintain consistency in biomarker definition and measurement approaches; however, imaging biomarkers such as functional magnetic resonance imaging (fMRI), resting-state fMRI, or diffusion tensor imaging will be included when the data capture neural function related to brain activity. Acceptable biological fluids include blood (eg, serum, plasma, whole blood for DNA/RNA analysis), urine, and other accessible fluids such as saliva.

Context

A study will be classified as longitudinal if it includes at least one follow-up frailty assessment in addition to the baseline evaluation and if biomarker measurements were obtained at baseline with frailty assessed at one or more later time points. This temporal ordering ensured prospective assessment of frailty incidence or progression.

Search Methods

Sources of Evidence

This scoping review will include only peer-reviewed original research articles. We will also exclude reviews, qualitative studies, case studies, posters, abstracts, editorials, books, and commentaries because they lack original data or sufficient methodological details required to evaluate longitudinal associations between biomarkers and frailty.

Search Strategy

We will search 8 databases (PubMed, CINAHL, Cochrane Library, Scopus, Web of Science, Embase, CNKI, and RISS) by using key concept words such as “frailty,” “biomarkers,” and “longitudinal studies.” The search will have no date restrictions and will include literature published in Korean, English, or Chinese. We will not impose limitations on the study setting or geographic location.

The search strategy was initially developed using core controlled vocabularies from major databases, including Medical Subject Headings (MeSH) in PubMed, Subject Headings in CINAHL, and Emtree terms in Embase. Based on these controlled terms, we supplemented additional keywords relating to specific biomarkers and frailty phenotypes to construct the final search trails.

Two reviewers (KS and YY) independently drafted and refined the initial list of search terms, resolving any discrepancies through discussion. A third reviewer (JWS) subsequently evaluated the proposed terms and adjudicated remaining disagreements regarding the inclusion or exclusion of specific keywords. Following this adjudication process, a fourth reviewer (SH) performed a final review to ensure the completeness, accuracy, and coherence of the search terms before finalization. The details of the preliminary search trails for each database are presented in the Multimedia Appendix 1.

Study Selection

We will import all retrieved citations into Covidence (Veritas Health Innovation Ltd, Melbourne, Australia) for reference management, with duplicates automatically removed. Two independent reviewers (KS and YY) will screen titles and abstracts to assess eligibility according to predefined inclusion and exclusion criteria (Textbox 1). The same reviewers will then independently review the full texts of potentially relevant articles. We will resolve any disagreements through discussion; if consensus cannot be reached, a third reviewer (JWS or SH) will make the final decision. We will record reasons for excluding studies at the full-text stage. We will document and report the entire process using a PRISMA flow diagram [29].

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for study selection.

Inclusion criteria

- Adults aged ≥60 years
- Frailty assessed using validated tools (Fried Frailty Phenotype, Frailty Index, Laboratory-based Frailty Index, Tilburg Frailty Indicator, Edmonton Frailty Scale)
- Biomarkers measured at baseline prior to frailty assessment
- Biomarkers from biological fluids (blood, urine, saliva, cerebrospinal fluid)
- Biomarker types: molecular, genetic, transcriptomic, functional imaging biomarkers
- Longitudinal study design with ≥2 time points
- Frailty incidence or progression assessed over time
- Peer-reviewed, full-text, published in English/Korean/Chinese

Exclusion criteria

- Animal studies or nonhuman data
- Frailty not defined or not measured as an outcome
- Studies assessing only frailty risk factors without frailty outcome
- Nonbiological measures (InBody, Bioelectrical Impedance Analysis)
- Imaging not related to physiological or neural function
- Longitudinal design but cross-sectional analysis only
- Preprints, abstracts, dissertations, or unavailable full-text

Data Extraction

Two independent reviewers will extract data using a standardized form developed for this review. Extracted information will include study characteristics (author, year of publication, country, and study design), participant details (age and frailty assessment tools used), types of biomarkers measured at baseline, specimen used for measurement, follow-up duration, statistical or analytical methods, and key findings on the longitudinal association between biomarkers and frailty. When annual (or other interval) assessments are available, each follow-up period will be explicitly recorded to capture how biomarker trajectories and frailty status change over time, enabling clearer interpretation of the temporal relationships reported across studies.

We will refine the data extraction form as needed during the extraction process to ensure comprehensive capture of information relevant to the review questions. Discrepancies between the two reviewers (KS and YY) will be resolved through discussion. If consensus cannot be reached, a third reviewer (JWS) will make the final decision. For unclear or missing data, we will record “not reported” in the data extraction sheet. We will not apply imputation or substitution. These gaps will be acknowledged in the narrative synthesis.

Data Synthesis

We will provide a narrative synthesis to describe how the findings align with the objectives and research questions of the review, following [30]. In the results section, we will synthesize the identified biomarkers and their longitudinal associations with frailty. We will stratify results by study population, including community-dwelling older adults and patients with metabolic disorders, musculoskeletal conditions, neurological diseases, cancer, infectious disease, and cardiovascular disease. Biomarkers will be categorized by specific lineage: inflammatory (eg, C-reactive protein, interleukin-6, tumor necrosis factor-α), metabolic (eg, glucose, hemoglobin A_{1c}, lipid profiles), hormonal (eg, testosterone, sex hormone–binding globulin, insulin-like growth factor I), and other groups. We will create a summary figure to present the initial and follow-up periods for each study, enabling comparison of the longitudinal effects over time.

Critical Appraisal

We will conduct this review using the JBI critical appraisal tools for cohort studies [31]. In this review, we will consider biomarkers as the exposure and frailty as the outcome. We will present the results of the critical appraisal in a summary table, which will show the total number and percentage of “yes” responses for each study. As this study is a scoping review, all relevant studies will be included regardless of methodological quality. However, we will provide a summary of the number of high-quality studies in the current evidence base.

Ethical Considerations

Because this study relied exclusively on publicly available published and gray literature, ethics committee approval was not required.

Results

We initiated an exploratory search in July 2024 to obtain preliminary results. After removing duplicates, the preliminary

XSL•FO
RenderX

search identified 3315 records. We plan to complete this review by December 2026 (Figure 1).

Figure 1. Flow diagram of identification, screening, and inclusion of studies for the scoping review.

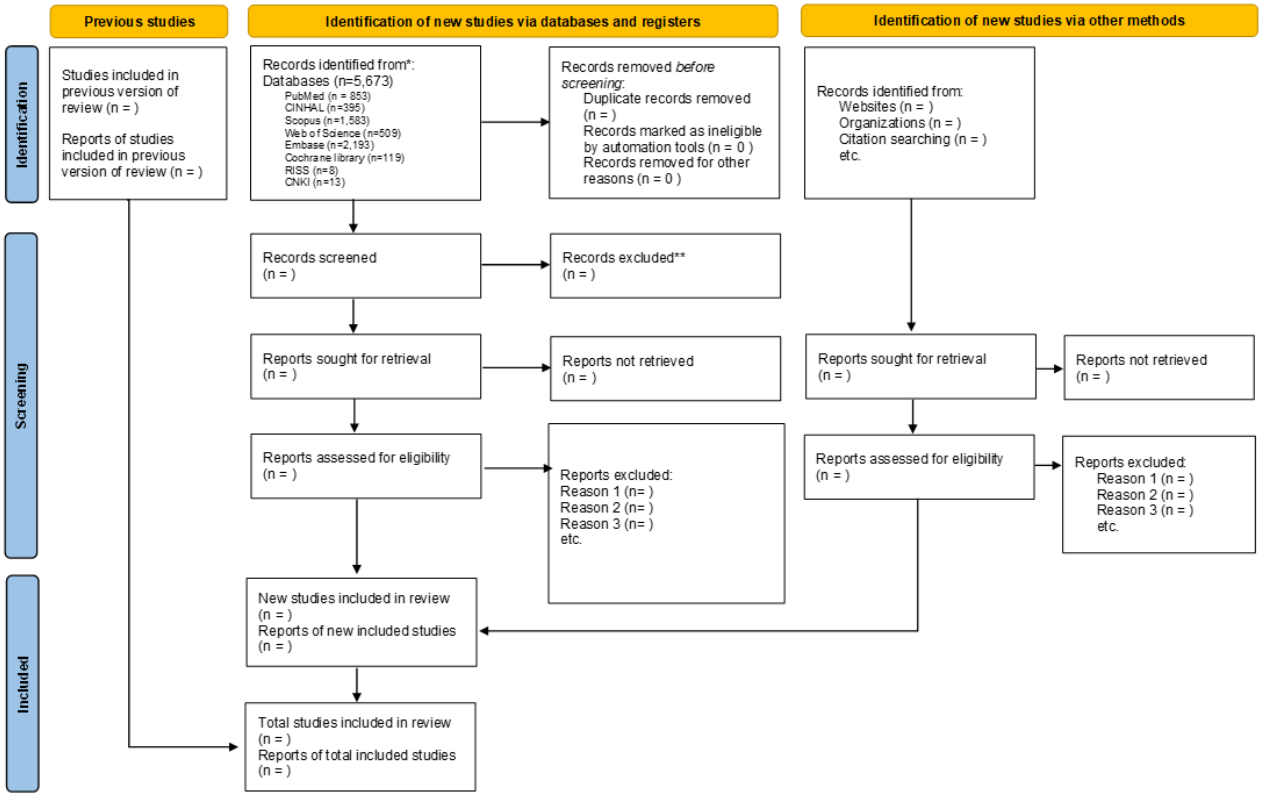

Discussion

Anticipated Findings

Through this scoping review, we expect to identify and synthesize biomarkers of frailty that reflect the sequential influence of biological factors on frailty progression in older adults. We hypothesize that multiple heterogeneous biomarkers, including inflammatory, metabolic, cardiovascular, endocrine, and neurologic markers, exist. However, we also expect substantial variability in biomarker types, frailty definitions, follow-up durations, and statistical and analytic approaches across studies.

Comparison to Prior Work

Previous reviews suggested albumin, hemoglobin, interlukin-6, and tumor necrosis factor-α as potential biomarkers for frailty and sarcopenia and reported inflammatory markers such as leukocyte, lymphocytes, C-reactive protein, interleukin-6, interleukin-10, and tumor necrosis factor-α as a biomarker for frailty in a cross-sectional study [10,32]; however, few longitudinal studies have been conducted, leaving a knowledge gap about the potential use of biomarkers of frailty for early detection and prevention. This review will address the range of biomarkers of frailty that precede frailty incidence or progression and overview existing findings by biological process and biomarker source, as well as follow-up duration and analytical and statistical methods.

Strengths and Limitations

This review addresses the challenge of systematically examining the complex biological changes that precede frailty and evaluating their potential clinical relevance. By integrating current evidence without restricting the scope to predefined biomarker categories, this work captures the multifaceted biological processes underlying frailty and reflects its dynamic and systemic nature. Importantly, this approach allows for a more comprehensive understanding of frailty-related biological alterations across diverse physiological systems.

A key strength of this review lies in its inclusive exploration of heterogeneous biomarker types and analytical strategies. We examined a broad spectrum of biomarkers and analytical methodologies rather than focusing on a single class, thereby providing insight into how different approaches contribute to biomarker discovery and interpretation. This approach also enabled a critical discussion of methodological considerations, including study design, analytic frameworks, and appropriate follow-up durations needed to meaningfully capture longitudinal changes in frailty status.

Furthermore, by incorporating not only blood-based biomarkers but also neurologic markers such as fMRI, this review highlights the multidimensional nature of frailty. This broader biological perspective is particularly relevant for understanding cognitive frailty and other neurobiological dimensions of frailty that extend beyond traditional physical or inflammatory markers, thereby supporting a more holistic and personalized framework for frailty prediction and management.

XSL•FO
RenderX

Despite these strengths, several limitations should be acknowledged. As noted, most included studies were observational or cohort-based, limiting causal inference. In addition, substantial heterogeneity in biomarker assessment methods, statistical analyses, and outcome definitions precluded meta-analytic synthesis and the identification of statistically robust biomarkers. The lack of standardized protocols further limited direct comparisons across studies. Moreover, this review did not include body composition or functional biomarkers derived from measures such as bioelectrical impedance analysis (eg, InBody), which may restrict the scope of biological domains represented. Excluding such markers may underestimate the contribution of musculoskeletal and body composition-related pathways to frailty and sarcopenia, highlighting an important area for future integrative research.

Future Directions

Future biomarker research should adopt a more integrative approach by combining molecular biomarkers with functional and body-composition measures such as bioelectrical impedance–derived indices, physical performance metrics, and wearable-based life-log data, while integrating multi-omics data (eg, genomics, transcriptomics, proteomics, and metabolomics) within a systems biology framework to better capture the complex biological mechanisms underlying frailty.

In addition, future studies should also identify and describe biomarker evidence specific to distinct frailty subgroups, including physical, cognitive, and affective frailty, to better reflect the heterogeneous biological pathways underlying frailty. For biomarkers that are consistently reported and show strong associations with frailty across studies, meta-analytic synthesis is needed to strengthen the evidence base and support clinical translation.

Dissemination Plans

The findings of this scoping review will be disseminated through publication in a peer-reviewed journal, with manuscript submission anticipated in December 2026.

Conclusions

This scoping review will systematically identify and examine longitudinal biomarkers that precede frailty in older adults, providing an integrated overview of the complex biological changes underlying frailty progression. By synthesizing evidence across diverse biomarker domains and analytical approaches, this review will emphasize the potential for advancing personalized frailty prediction and management while highlighting the multifaceted and dynamic nature of frailty biology.

Acknowledgments

Generative artificial intelligence tools were used only for language refinement and not for scientific content development, data interpretation, conceptualization, or analysis. All scientific ideas, methods, and conclusions were developed by the authors, and the final manuscript was thoroughly reviewed and verified by the authors.

Funding

This study is supported by the Basic Science Research Program through the National Research Foundation of Korea, funded by the Ministry of Science and ICT (RS-2024-00338005).

Authors' Contributions

JWS contributed to conceptualization, methodology, writing the original draft, supervision, and funding acquisition. YY and KS contributed to methodology and writing the original draft. SH contributed to conceptualization, methodology, writing the original draft, and supervision. All authors contributed to writing, review, and editing and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search trails.
[DOCX File , 31 KB - resprot_v15i1e83312_app1.docx]

References

1. Dlima SD, Hall A, Aminu AQ, Akpan A, Todd C, Vardy ERLC. Frailty: a global health challenge in need of local action. BMJ Glob Health 2024 Aug 09;9(8):e015173 [FREE Full text] [doi: 10.1136/bmjgh-2024-015173] [Medline: 39122463]
2. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]
3. Gobbens RJ, van Assen MA, Luijkx KG, Schols JM. Testing an integral conceptual model of frailty. J Adv Nurs 2012 Sep;68(9):2047-2060. [doi: 10.1111/j.1365-2648.2011.05896.x] [Medline: 22150238]

XSL•FO
RenderX

4. Sánchez-García S, García-Peña C, Salvà A, Sánchez-Arenas R, Granados-García V, Cuadros-Moreno J, et al. Frailty in community-dwelling older adults: association with adverse outcomes. Clin Interv Aging 2017;12:1003-1011 [FREE Full text] [doi: 10.2147/CIA.S139860] [Medline: 28721028]
5. Chittrakul J, Siviroj P, Sungkarat S, Sapbamrer R. Physical frailty and fall risk in community-dwelling older adults: a cross-sectional study. J Aging Res 2020;2020:3964973 [FREE Full text] [doi: 10.1155/2020/3964973] [Medline: 32714612]
6. Pearce J, Martin S, Heritage S, Khoury EG, Kucharczak J, Nuamek T, et al. Frailty and outcomes in adults undergoing systemic anticancer treatment: a systematic review and meta-analysis. J Natl Cancer Inst 2025 Jul 01;117(7):1316-1339. [doi: 10.1093/jnci/djaf017] [Medline: 39886936]
7. Wleklik M, Denfeld Q, Czapla M, Jankowska EA, Piepoli MF, Uchmanowicz I. A patient with heart failure, who is frail: how does this affect therapeutic decisions? Cardiol J 2023;30(5):825-831 [FREE Full text] [doi: 10.5603/CJ.a2023.0027] [Medline: 37067336]
8. Liu C, Yu J, Li X, Wei H, Liu X, Zhang W, et al. Progress in lung cancer study coupled with cognitive frailty in elderly individuals. Geriatr Nurs 2025;61:423-428. [doi: 10.1016/j.gerinurse.2024.12.001] [Medline: 39693687]
9. Faller JW, Pereira DDN, de Souza S, Nampo FK, Orlandi FDS, Matumoto S. Instruments for the detection of frailty syndrome in older adults: A systematic review. PLoS One 2019;14(4):e0216166 [FREE Full text] [doi: 10.1371/journal.pone.0216166] [Medline: 31034516]
10. Picca A, Coelho-Junior HJ, Calvani R, Marzetti E, Vetrano DL. Biomarkers shared by frailty and sarcopenia in older adults: A systematic review and meta-analysis. Ageing Res Rev 2022 Jan;73:101530. [doi: 10.1016/j.arr.2021.101530] [Medline: 34839041]
11. Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013 Mar 02;381(9868):752-762 [FREE Full text] [doi: 10.1016/S0140-6736(12)62167-9] [Medline: 23395245]
12. Avram L, Ungureanu MI, Crişan D, Donca V. Assessment of frailty scores among geriatric patients hospitalized in the north-western region of Romania: a cross-sectional study. Medicina (Kaunas) 2024 Nov 26;60(12):1947 [FREE Full text] [doi: 10.3390/medicina60121947] [Medline: 39768829]
13. Lin KP, Li HY, Chen JH, Lu FP, Wen CJ, Chou YC, et al. Prediction of adverse health outcomes using an electronic frailty index among nonfrail and prefrail community elders. BMC Geriatr 2023 Aug 07;23(1):474 [FREE Full text] [doi: 10.1186/s12877-023-04160-1] [Medline: 37550602]
14. Saedi AA, Feehan J, Phu S, Duque G. Current and emerging biomarkers of frailty in the elderly. Clin Interv Aging 2019;14:389-398 [FREE Full text] [doi: 10.2147/CIA.S168687] [Medline: 30863033]
15. Wang X, Ji J. Explainable machine learning framework for biomarker discovery by combining biological age and frailty prediction. Sci Rep 2025 Apr 22;15(1):13924 [FREE Full text] [doi: 10.1038/s41598-025-98948-3] [Medline: 40263505]
16. Yamasaki T, Tanaka M, Kumagai S. Editorial: recent advances in research on cognitive frailty and related conditions. Front Aging Neurosci 2025;17:1595526 [FREE Full text] [doi: 10.3389/fnagi.2025.1595526] [Medline: 40259947]
17. Sepúlveda M, Arauna D, García F, Albala C, Palomo I, Fuentes E. Frailty in aging and the search for the optimal biomarker: a review. Biomedicines 2022 Jun 16;10(6):1426 [FREE Full text] [doi: 10.3390/biomedicines10061426] [Medline: 35740447]
18. Erusalimsky JD, Grillari J, Grune T, Jansen-Duerr P, Lippi G, Sinclair AJ, FRAILOMIC Consortium. In search of 'omics'-based biomarkers to predict risk of frailty and its consequences in older individuals: the FRAILOMIC initiative. Gerontology 2016;62(2):182-190. [doi: 10.1159/000435853] [Medline: 26227153]
19. Kane AE, Sinclair DA. Frailty biomarkers in humans and rodents: current approaches and future advances. Mech Ageing Dev 2019 Jun;180:117-128 [FREE Full text] [doi: 10.1016/j.mad.2019.03.007] [Medline: 31002925]
20. Cardoso AL, Fernandes A, Aguilar-Pimentel JA, de Angelis MH, Guedes JR, Brito MA, et al. Towards frailty biomarkers: candidates from genes and pathways regulated in aging and age-related diseases. Ageing Res Rev 2018 Nov;47:214-277 [FREE Full text] [doi: 10.1016/j.arr.2018.07.004] [Medline: 30071357]
21. Gonçalves RSDSA, Maciel ÁCC, Rolland Y, Vellas B, de Souto Barreto P. Frailty biomarkers under the perspective of geroscience: a narrative review. Ageing Res Rev 2022 Nov;81:101737. [doi: 10.1016/j.arr.2022.101737] [Medline: 36162706]
22. Shi X, Yang H, Wang S, Yang Y, Li Y, Dou G, et al. Development and validation of a predictive nomogram for frailty based on thyroid function in older adults. Eur Geriatr Med 2025 Oct;16(5):1861-1870. [doi: 10.1007/s41999-025-01247-3] [Medline: 40526339]
23. Rodriguez-Mañas L, Araujo de Carvalho I, Bhasin S, Bischoff-Ferrari HA, Cesari M, Evans W, et al. ICFSR task force perspective on biomarkers for sarcopenia and frailty. J Frailty Aging 2020;9(1):4-8 [FREE Full text] [doi: 10.14283/jfa.2019.32] [Medline: 32150207]
24. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis. Australia: JBI; 2024.
25. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
26. Biomarkers. Sciences NIoEH. URL: https://www.niehs.nih.gov/health/topics/science/biomarkers [accessed 2025-04-23]
27. Leng S, Chen X, Mao G. Frailty syndrome: an overview. CIA 2014 Mar;9:433. [doi: 10.2147/cia.s45300]

XSL•FO
RenderX

28. Fried LP, Cohen AA, Xue Q, Walston J, Bandeen-Roche K, Varadhan R. The physical frailty syndrome as a transition from homeostatic symphony to cacophony. Nat Aging 2021 Jan;1(1):36-46 [FREE Full text] [doi: 10.1038/s43587-020-00017-z] [Medline: 34476409]
29. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
30. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368(12):l6890 [FREE Full text] [doi: 10.1136/bmj.l6890] [Medline: 31948937]
31. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, et al. Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis. Australia: JBI; 2020.
32. Xu Y, Wang M, Chen D, Jiang X, Xiong Z. Inflammatory biomarkers in older adults with frailty: a systematic review and meta-analysis of cross-sectional studies. Aging Clin Exp Res 2022 May;34(5):971-987. [doi: 10.1007/s40520-021-02022-7] [Medline: 34981430]

Abbreviations

fMRI: functional magnetic resonance imaging
JBI: Joanna Briggs Institute
MeSH: Medical Subject Headings
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by J Sarvestan; submitted 31.Aug.2025; peer-reviewed by CY Chen, Z Xia; comments to author 15.Nov.2025; revised version received 08.Jan.2026; accepted 23.Jan.2026; published 24.Mar.2026.

Please cite as:
Seok JW, Shim K, Yang Y, Hong S
Longitudinal Associations Between Biomarkers and Frailty in Older Adults: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e83312
URL: https://www.researchprotocols.org/2026/1/e83312
doi:10.2196/83312
PMID:

©Jo Woon Seok, Kyusol Shim, Yihan Yang, Soomin Hong. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Inside-Out Mechanism of Referred Visceral Pain: Protocol for a Systematic Review of Cutaneous Neurogenic Inflammation in Preclinical Animal Models of Visceral Diseases

EunMee Yang[1], MPH, PhD; Dennis Muñoz-Vergara[1], MPH, DVM; Andrew C Ahn[1], MD; Weidong Lu[2], MB, PhD; Vitaly Napadow[3,4], PhD; Alexandria E Cronin[5], BA, MLIS; Peter M Wayne[1], PhD

[1]Osher Center for Integrative Health, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States

[2]Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, MA, United States

[3]Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital and Harvard Medical School, Boston, MA, United States

[4]Department of Radiology, Athinoula A Martinos Center for Biomedical Imaging and Harvard Medical School, Charlestown, MA, United States

[5]Countway Library of Medicine, Harvard Medical School, Boston, MA, United States

Corresponding Author:
EunMee Yang, MPH, PhD
Osher Center for Integrative Health, Brigham and Women's Hospital and Harvard Medical School
900 Commonwealth Ave, 3rd floor
Boston, MA, 02215
United States
Phone: 1 617 582 9291
Email: eyang13@bwh.harvard.edu

Abstract

Background: Visceral diseases often lead to referred pain, but clinical management is challenging due to poor understanding of the underlying mechanisms. Prior preclinical studies have shown that experimental induction of visceral diseases in animal models can cause cutaneous neurogenic inflammation, as assessed by Evans Blue dye extravasation. Neurogenic inflammation is associated with increased pain sensitivity. Thus, these preclinical findings may provide important mechanistic insights into the relationship between visceral diseases and referred somatic pain.

Objective: The objective of this systematic review is to synthesize evidence on cutaneous neurogenic inflammation and corresponding somatic pain sensitivity in preclinical animal models of visceral diseases, to lay out the current state of knowledge, and to identify any gaps or limitations to be addressed in future research on this topic.

Methods: This systematic review will be conducted, following the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol) guidelines. The literature search will be conducted in MEDLINE (Ovid), Embase, and Web of Science, using search terms related to cutaneous neurogenic inflammation and animal models of visceral diseases. Two investigators will independently screen and select studies meeting the eligibility criteria. For each included study, a predefined template will be used to extract relevant information, including the details of animal species and strains, visceral disease models, outcomes related to the distribution and extent of Evans Blue dye extravasation, and other pertinent findings and methodologies. Quality assessments will be independently performed by two investigators using the Risk of Bias tool from the Systematic Review Centre for Laboratory Animal Experimentation (SYRCLE). A qualitative summary of the findings will be provided for outcomes related to the extent and distribution of Evans Blue dye extravasation.

Results: As of August 2025, a total of 536 records were identified from database search, and 13 studies met the eligibility criteria to be included in the review. Results from data extraction and quality assessments are expected to be published in fall 2026.

Conclusions: Prior preclinical animal studies have investigated cutaneous neurogenic inflammation as a potential mechanistic outcome of referred visceral pain. The proposed review will be the first to summarize research findings and study quality on this topic. This will be an important step toward gaining insights into one of the potential mechanisms underlying referred visceral pain and informing future research in this area.

International Registered Report Identifier (IRRID): DERR1-10.2196/67852

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e67852) doi:10.2196/67852

KEYWORDS

referred visceral pain; cutaneous neurogenic inflammation; Evans Blue dye; somatic hypersensitivity; systematic review protocol; preclinical studies; neurogenic; inflammation; visceral pain; pain; hypersensitivity; animal models; systematic review; review; visceral diseases; somatic pain; sensitivity; sensitive; animal

Introduction

Referred visceral pain is a common yet poorly understood phenomenon, whereby pain originating from internal organs is perceived as pain in other areas of the body, often in somatic structures such as the skin [1-3]. Examples include pain in the chest, left arm, neck, jaw, and upper back during myocardial infarction, as well as abdominal or back pain in pancreatitis, gastrointestinal disorders, and interstitial cystitis [4-12]. In clinical practice, management of referred visceral pain remains challenging because currently available treatments offer limited relief, and its underlying mechanisms are not fully understood [5,13].

To date, explanations for referred visceral pain have focused on central mechanisms, particularly the concept of viscerosomatic convergence. In this model, visceral and somatic afferents converge onto shared spinal neurons, causing the central nervous system to misinterpret visceral pain as originating from somatic regions [13]. While this framework explains the central integration of visceral and somatic signals, it does not account for any physiological changes observed in the somatic structures.

Emerging evidence from preclinical animal studies suggests an additional mechanism involving cutaneous neurogenic inflammation [14-20]. This process involves crosstalk between visceral and somatic afferents, whereby noxious input from the viscera activates the convergent somatic afferents, triggering antidromic signaling and peripheral release of vasoactive neuropeptides such as substance P and calcitonin gene-related peptide. These neuropeptides, in turn, cause localized inflammatory changes, including vasodilation and plasma extravasation and are associated with increased pain sensitivity [21,22]. Therefore, such cutaneous manifestations may represent an “inside-out” extension of referred visceral pain that contribute to somatic hyperalgesia.

In their 1997 study, Wesselmann and Lai [14] provided the first key experimental evidence for this mechanism by demonstrating that uterine inflammation in rats induced neurogenic inflammation on the body surface along the referred area of somatotopic overlap. Upon intravenous injection of Evans Blue (EB) dye, plasma extravasation was visualized as blue-stained spots on the skin. Subsequent studies have applied this approach (ie, intravenous injection of EB dye) to models of other visceral diseases, reporting segmentally organized patterns of EB dye extravasation [15-20]. Despite these important findings, to date, no systematic or narrative review has been conducted to summarize the preclinical evidence investigating cutaneous neurogenic inflammation in visceral disease models.

Understanding the relationship between cutaneous neurogenic inflammation and visceral diseases has important translational implications for pain research. Unlike central models that primarily explain altered pain perception, evidence of peripheral physiological changes—such as increased microvascular permeability and inflammation in the skin—suggests a measurable biological correlate of referred visceral pain. Characterizing these changes may provide objective biomarkers of viscerosomatic interactions and improve understanding of how visceral disease may give rise to referred pain. Such insights could inform the development of targeted diagnostic and therapeutic strategies for managing visceral pain and its somatic manifestations.

Therefore, the purpose of this systematic review is to synthesize preclinical evidence of cutaneous neurogenic inflammation in animal models of visceral disease. Specifically, we aim to examine whether experimentally induced models of visceral disease manifest signs of cutaneous neurogenic inflammation, as assessed by EB dye extravasation, and whether there are corresponding changes in pain sensitivity. Secondary aims include assessing the methodological quality and risk of bias of the included studies, summarizing complementary laboratory findings, and identifying gaps in the existing literature to inform future research.

Methods

Study Design and Registration

We will conduct a systematic review of preclinical animal studies, according to the guidelines recommended in PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) [23]. A completed PRISMA-P checklist can be found as Multimedia Appendix 1. In preparing this protocol, we have also adopted an animal systematic review protocol format developed by de Vries et al [24], which was designed specifically for review of animal studies. Any changes or amendments made to the protocol will be recorded in PROSPERO as minor amendments or major revisions, following their procedures, with each update including the date, version number, and a summary of the changes. Any deviations from the registered protocol will also be justified and documented in the final manuscript.

Search Strategy

A literature search will be conducted in MEDLINE (Ovid), Embase, and Web of Science by using terms related to the research question, including but limited to “cutaneous neurogenic inflammation” and “Evans Blue dye extravasation.” The initial search strategy was jointly developed by two investigators (EY and PW) and a medical librarian/information scientist (AC). It was then reviewed and refined based on suggestions and feedback from other co-investigators. The search will be limited to English-language publications. No limitation will be placed on the publication date range. In

XSL•FO
RenderX

addition to electronic database search, we will hand search references of the included studies to identify additional relevant studies. The full search strategies for all databases can be found in Table 1. The search will be rerun prior to data synthesis to capture newly published studies, and updated search dates and strategies will be documented in the supplementary materials and reported in the final manuscript to ensure transparency and reproducibility.

Table 1. Database search strategy.

Database	Search strategy
MEDLINE	1. exp “Extravasation of Diagnostic and Therapeutic Materials”/ OR exp “Evans Blue”/ OR exp “Acupuncture Points”/
	2. ((Acupuncture OR neurogenic) adj2 (spot* OR point*)).ab,kf,kw,ti.
	3. (acupoint* OR Neuro-Sp* OR “Evans blue” OR “EB dye” OR “dye extravasation” OR “plasma extravasation”).ab,kf,kw,ti.
	4. 1 OR 2 OR 3
	5. exp “Neurogenic Inflammation”/
	6. (“neurogenic inflammation” OR (visceral AND (disease OR pain OR inflam*) AND neuro*) OR (referred AND pain AND neuro*)).ab,kf,kw,ti.
	7. 5 OR 6
	8. 4 AND 7
Embase	1. ('contrast medium extravasation'/exp OR 'evans blue'/exp OR 'acupuncture point'/exp) AND [embase]/lim
	2. (((acupuncture OR neurogenic) NEAR/2 (spot* OR point*)):ab,kw,ti) AND [embase]/lim
	3. ((acupoint* OR Neuro-Sp* OR 'Evans blue' OR 'EB dye' OR 'dye extravasation' OR 'plasma extravasation'):ab,kw,ti) AND [embase]/lim
	4. 1 OR 2 OR 3
	5. ('neurogenic inflammation'/exp) AND [embase]/lim
	6. (('neurogenic inflammation' OR (visceral AND (disease OR pain OR inflam*) AND neuro*) OR (referred AND pain AND neuro*)):ab,kw,ti) AND [embase]/lim
	7. 5 OR 6
	8. 4 AND 7
Web of Science	1. TI=((acupuncture OR neurogenic) NEAR/2 (spot* OR point*))
	2. AB=((acupuncture OR neurogenic) NEAR/2 (spot* OR point*))
	3. AK=((acupuncture OR neurogenic) NEAR/2 (spot* OR point*))
	4. TI=(acupoint* OR Neuro-Sp* OR “Evans blue” OR “EB dye” OR “dye extravasation” OR “plasma extravasation”)
	5. AB=(acupoint* OR Neuro-Sp* OR “Evans blue” OR “EB dye” OR “dye extravasation” OR “plasma extravasation”)
	6. AK=(acupoint* OR Neuro-Sp* OR “Evans blue” OR “EB dye” OR “dye extravasation” OR “plasma extravasation”)
	7. #1 OR #2 OR #3 OR #4 OR #5 OR #6
	8. TI=(“neurogenic inflammation” OR (visceral AND (disease OR pain OR inflam*) AND neuro*) OR (referred AND pain AND neuro*))
	9. AB=(“neurogenic inflammation” OR (visceral AND (disease OR pain OR inflam*) AND neuro*) OR (referred AND pain AND neuro*))
	10. AK=(“neurogenic inflammation” OR (visceral AND (disease OR pain OR inflam*) AND neuro*) OR (referred AND pain AND neuro*))
	11. #8 OR #9 OR #10
	12. #7 AND #11

Of note, we found that some of the preclinical studies on this topic were conducted by acupuncture researchers aiming to compare the locations and distributions of cutaneous neurogenic inflammatory spots with that of acupuncture points, the specific points on the body surface that are stimulated during acupuncture treatment [15-20]. Thus, we also included terms related to “acupuncture points” to be more comprehensive in our search strategy (Table 1).

Eligibility

A summary of the inclusion and exclusion criteria is provided in Textbox 1. Briefly, this review will include preclinical animal studies, investigating experimentally induced models of acute or chronic visceral diseases. Studies involving humans, in vitro, or ex vivo experiments will be excluded. Studies using EB dye or other validated markers of plasma extravasation will be considered for inclusion. EB is an azo dye that binds strongly to blood albumin, and its leakage into the surrounding tissue can be visualized to assess the extent and distribution of cutaneous neurogenic inflammation [25,26]. Only studies with a control or baseline comparison and sufficient methodological detail for data extraction will be included.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Preclinical animal studies using acute or chronic visceral disease models
- Inclusion of a comparison condition
- Use of Evans Blue dye or equivalent validated measures of plasma extravasation to assess cutaneous neurogenic inflammation
- Studies reporting original data in peer-reviewed journals
- Published in English with sufficient methodological detail for data extraction

Exclusion criteria

- Human studies, clinical trials, or in vitro/ex vivo experiments
- Studies lacking a control or baseline condition for comparison with the disease model
- Studies not using Evans Blue dye or other equivalent validated measures of plasma extravasation
- Reviews, commentaries, editorials, conference abstracts, theses, preprints, or duplicated datasets
- Non-English publications or reports with incomplete methodological information preventing reliable data extraction

Study Selection

Two investigators (EY and DMV) will independently screen all the studies resulting from the literature search to include relevant studies that meet the predefined inclusion and exclusion criteria. Screening will first be conducted based on titles and abstracts. Studies that do not meet the eligibility criteria will be excluded at this stage. The remaining studies will be screened by retrieving and examining the full texts for eligibility. Interrater agreement will be assessed using Cohen κ. Any disagreements between reviewers will first be discussed to reach a consensus; if unresolved, they will be adjudicated by a third investigator (PW).

Duplicate publications and overlapping datasets will be identified by cross-checking study details (eg, author names, animal models, experimental methods). When multiple reports describe the same dataset, the most comprehensive or recent publication will be included, and additional information from other reports will be extracted as appropriate to avoid duplication.

Data Extraction

Two investigators (EY and DMV) will independently extract data from all included studies, using a predefined template (Multimedia Appendix 2). Data extraction will first be piloted on a subset of studies to ensure consistency and completeness between reviewers. Any discrepancies during data extraction will be resolved through discussion and consensus; if unresolved, they will be adjudicated by a third investigator (PW).

Briefly, the data extraction table includes the following: (1) publication details (eg, title, authors, publication year); (2) study population (eg, details on animal species and strains, sample size, and other characteristics of the animals such as sex and age); (3) visceral disease model, including the details of the methodology used to induce the disease model (eg, name and dosage of chemical irritant used to induce the disease model); (4) comparison or control group (eg, use of saline); (5) outcomes measures (eg, comparisons of EB dye extravasation in disease and control groups in terms of distribution and numbers of blue dots), including the details of the methodology (eg, dosage and route of EB dye administration); (6) outcomes and methods relating to pain sensitivity assessment of blue dots; (7) any other relevant methods and results from included studies, which provide additional insights about the neurocircuitry of referred visceral pain as well as other pathophysiological features of the skin in visceral diseases; and (8) if applicable, pertinent information about acupuncture point(s) (eg, names of acupuncture points in close proximity to blue dots).

The primary outcome of this review will be the extent of cutaneous neurogenic inflammation in response to experimental induction of visceral disease, as assessed by the distribution and extent of EB dye extravasation. Secondary outcomes will include pain sensitivity assessments as well as additional measures of neurogenic inflammation such as immunohistochemistry for neuropeptides (eg, substance P, calcitonin gene-related peptide), if available. Acupuncture point-related data will also be considered secondary and extracted only when directly relevant to the main mechanistic question.

Risk of Bias Assessment

The included studies will be independently assessed for their methodological quality and risk of bias by two investigators (EY and DM-V). Given that this review involves preclinical animal studies, we will adopt the Risk of Bias (RoB) tool from the Systematic Review Centre for Laboratory Animal Experimentation (SYRCLE), which has been developed to address the specific challenges and complexities associated with assessing the rigor of evidence derived from animal research [24,27]. Briefly, SYRCLE's RoB tool includes 10 entries related to 6 types of bias: selection bias, performance bias, detection bias, attrition bias, reporting bias, and other biases. This tool is designed to be applicable across different animal species and experimental models. Using this RoB tool, we will identify any potential limitations in the design, conduct, and reporting of included studies and determine the overall strength of evidence regarding cutaneous neurogenic inflammation in visceral diseases. Both reviewers have previously been trained in the

XSL•FO
RenderX

use of the SYRCLE risk of bias tool and will independently assess the included studies and then discuss their evaluations to reach consensus. Any disagreements will be resolved through discussion and consensus or review; if unresolved, they will be adjudicated by a third investigator (PW).

Data Synthesis

The findings of individual studies included in this review will be synthesized to identify overall patterns and draw conclusions relevant to the research question, taking into account the methodological quality of the included studies. For outcomes related to the distribution and extent of EB dye extravasation, a descriptive synthesis will be performed to summarize key trends across models, including species, visceral organ involved, and induction method. Given the anticipated heterogeneity in study design, disease induction, and outcome quantification, a meta-analysis will not be planned at this stage. Instead, findings will be reported narratively, highlighting consistent observations and methodological differences among studies. If sufficient homogeneity and quantitative data become available, a meta-analysis may be considered in future analyses.

Results

As of August 2025, a total of 536 records were identified from database search, and 13 studies met the eligibility criteria to be included in the review (Figure 1). Results from data extraction and quality assessment are expected to be published in fall 2026.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

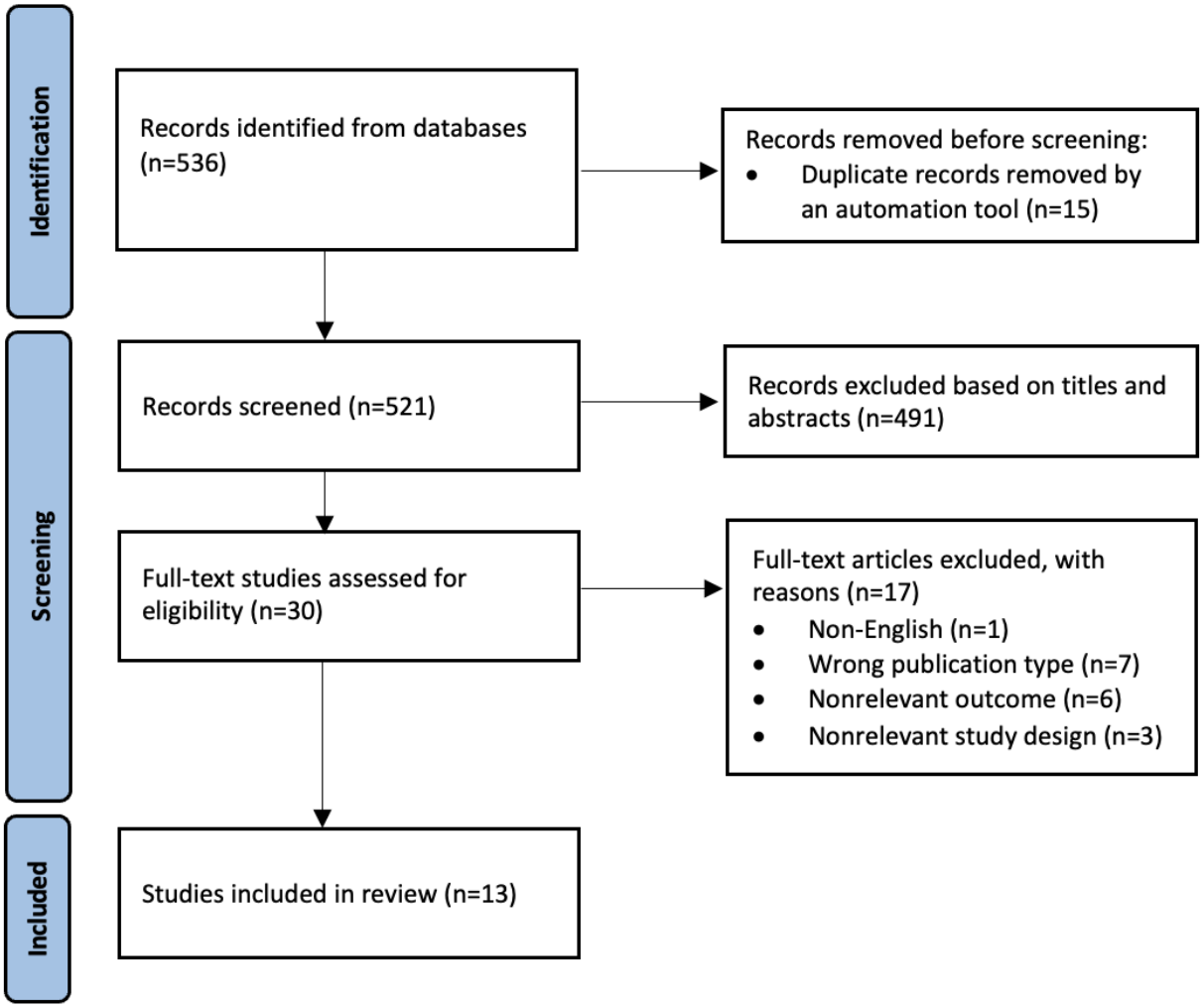

Discussion

To date, several preclinical studies have investigated cutaneous neurogenic inflammation as a potential mechanistic correlate of referred visceral pain, yet no synthesis of the collective evidence has been performed. To our knowledge, this proposed review will be the first to systematically search, summarize, and evaluate the quality of relevant studies on this topic.

This protocol was designed to ensure methodological rigor, transparency, and reproducibility. It follows the PRISMA-P guidelines, with predefined eligibility criteria, data extraction procedures, and a risk-of-bias assessment using the SYRCLE tool. These measures are intended to minimize bias and strengthen the validity of the findings. Our literature search will be conducted across multiple databases to ensure the inclusion of relevant studies from diverse sources. We will also include studies with diverse animal models, with no restriction on the specific species or strains of the animals used nor on the specific

XSL•FO
RenderX

visceral disease induced, to ensure a comprehensive overview of the topic. Independent screening and data extraction by two investigators with arbitration procedures for unresolved discrepancies will ensure consistency and reliability.

There are, however, limitations in the proposed review. Our literature search will be limited to the English language due to resource and translation constraints, which may introduce language bias and limit the comprehensiveness of the review. Grey literature (eg, conference abstracts, theses, preprints) will be also excluded, as these sources often lack sufficient detail for accurate data extraction and quality assessment; however, this exclusion may introduce publication bias. The review will focus primarily on EB dye and equivalent markers of plasma extravasation; although this ensures methodological consistency, it may not provide a comprehensive overview of all relevant evidence relating to mechanisms of referred visceral pain. In addition, while animal studies provide essential mechanistic insights, there are fundamental differences between animal models and humans, including variations in anatomy, physiology, neurocircuitry, and disease characteristics, which may limit direct translation of findings [28,29].

In summary, referred visceral pain remains a complex and incompletely understood phenomenon. While several preclinical studies have identified cutaneous neurogenic inflammation as a potential peripheral mechanism contributing to somatic manifestations of visceral disease, no systematic review has been conducted to summarize and evaluate the rigor of relevant studies. The proposed review will address this gap by providing the first comprehensive synthesis of preclinical findings on this topic, serving as an important step toward elucidating mechanisms underlying referred visceral pain and informing future research in this area.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Authors' Contributions

EY and PW conceived of the study. EY, PW, and AC developed the initial search strategy. EY drafted the initial manuscript. All authors read, contributed to the revision, and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-P checklist.
[PDF File (Adobe PDF File), 26 KB - resprot_v15i1e67852_app1.pdf]

Multimedia Appendix 2
Data extraction template.
[PDF File (Adobe PDF File), 24 KB - resprot_v15i1e67852_app2.pdf]

References

1. Lai HH, Gardner V, Ness TJ, Gereau RW. Segmental hyperalgesia to mechanical stimulus in interstitial cystitis/bladder pain syndrome: evidence of central sensitization. J Urol 2014 May;191(5):1294-1299 [FREE Full text] [doi: 10.1016/j.juro.2013.11.099] [Medline: 24316091]
2. Farrell KE, Callister RJ, Keely S. Understanding and targeting centrally mediated visceral pain in inflammatory bowel disease. Front Pharmacol 2014;5:27 [FREE Full text] [doi: 10.3389/fphar.2014.00027] [Medline: 24634658]
3. Gornet ME, Chandawarkar A, Herati A, Dellon AL. Interstitial cystitis/bladder pain syndrome as referred pain from injured T12/L1 nerves: symptomatic improvement with resection of ilioinguinal and iliohypogastric nerves. Urology 2021 Mar;149:24-29. [doi: 10.1016/j.urology.2020.11.034] [Medline: 33279610]
4. Foreman R, Garrett K, Blair R. Mechanisms of cardiac pain. Compr Physiol 2015 Apr;5(2):929-960. [doi: 10.1002/cphy.c140032] [Medline: 25880519]
5. van Oosterhout REM, de Boer AR, Maas AHEM, Rutten FH, Bots ML, Peters SAE. Sex differences in symptom presentation in acute coronary syndromes: a systematic review and meta‐analysis. JAHA 2020 May 05;9(9):e014733. [doi: 10.1161/jaha.119.014733]
6. Sikandar S, Dickenson A. Visceral pain: the ins and outs, the ups and downs. Curr Opin Support Palliat Care 2012 Mar;6(1):17-26 [FREE Full text] [doi: 10.1097/SPC.0b013e32834f6ec9] [Medline: 22246042]
7. Luz L, Fernandes E, Sivado M, Kokai E, Szucs P, Safronov B. Monosynaptic convergence of somatic and visceral C-fiber afferents on projection and local circuit neurons in lamina I: a substrate for referred pain. Pain 2015;156(10):a. [doi: 10.1097/j.pain.0000000000000267] [Medline: 26098437]

XSL•FO
RenderX

8. Beissner F, Henke C, Unschuld PU. Forgotten features of head zones and their relation to diagnostically relevant acupuncture points. Evid Based Complement Alternat Med 2011;2011:240653 [FREE Full text] [doi: 10.1093/ecam/nen088] [Medline: 19141490]
9. Arendt-Nielsen L, Schipper K, Dimcevski G, Sumikura H, Krarup AL, Giamberardino MA, et al. Viscero-somatic reflexes in referred pain areas evoked by capsaicin stimulation of the human gut. Eur J Pain 2008 Jul;12(5):544-551. [doi: 10.1016/j.ejpain.2007.08.010] [Medline: 17977761]
10. Van Oudenhove L, Kragel PA, Dupont P, Ly HG, Pazmany E, Enzlin P, et al. Common and distinct neural representations of aversive somatic and visceral stimulation in healthy individuals. Nat Commun 2020 Nov 23;11(1):5939 [FREE Full text] [doi: 10.1038/s41467-020-19688-8] [Medline: 33230131]
11. Johns E, Tracey I. Neuroimaging of visceral pain. Rev Pain 2009 Oct;3(2):2-5 [FREE Full text] [doi: 10.1177/204946370900300202] [Medline: 26527574]
12. Chang L. Brain responses to visceral and somatic stimuli in irritable bowel syndrome: a central nervous system disorder? Gastroenterol Clin North Am 2005 Jun;34(2):271-279. [doi: 10.1016/j.gtc.2005.02.003] [Medline: 15862935]
13. Jin Q, Chang Y, Lu C, Chen L, Wang Y. Referred pain: characteristics, possible mechanisms, and clinical management. Front Neurol 2023;14:1104817 [FREE Full text] [doi: 10.3389/fneur.2023.1104817] [Medline: 37448749]
14. Wesselmann U, Lai J. Mechanisms of referred visceral pain: uterine inflammation in the adult virgin rat results in neurogenic plasma extravasation in the skin. Pain 1997;73(3):309-317. [doi: 10.1016/s0304-3959(97)00112-7] [Medline: 9469520]
15. Fang Y, Han S, Li X, Xie Y, Zhu B, Gao X, et al. Cutaneous hypersensitivity as an indicator of visceral inflammation via c-nociceptor axon bifurcation. Neurosci Bull 2021 Jan;37(1):45-54 [FREE Full text] [doi: 10.1007/s12264-020-00577-5] [Medline: 32902804]
16. Fan Y, Ryu Y, Zhao R, Bills KB, Steffensen SC, Yang CH, et al. Enhanced spinal neuronal responses as a mechanism for increased number and size of active acupoints in visceral hyperalgesia. Sci Rep 2020 Jun 25;10(1):10312 [FREE Full text] [doi: 10.1038/s41598-020-67242-9] [Medline: 32587303]
17. He W, Wang X, Shi H, Bai W, Cheng B, Su Y, et al. Cutaneous neurogenic inflammation in the sensitized acupoints induced by gastric mucosal injury in rats. BMC Complement Altern Med 2017 Mar 07;17(1):141 [FREE Full text] [doi: 10.1186/s12906-017-1580-z] [Medline: 28270193]
18. Rong P, Li S, Ben H, Li L, Yu L, Cui C, et al. Peripheral and spinal mechanisms of acupoint sensitization phenomenon. Evid Based Complement Alternat Med 2013;2013:742195 [FREE Full text] [doi: 10.1155/2013/742195] [Medline: 24191171]
19. Kim D, Ryu Y, Hahm DH, Sohn BY, Shim I, Kwon OS, et al. Acupuncture points can be identified as cutaneous neurogenic inflammatory spots. Sci Rep 2017 Nov 09;7(1):15214 [FREE Full text] [doi: 10.1038/s41598-017-14359-z] [Medline: 29123119]
20. Kim H, Hahm D, Sohn B, Choi Y, Pyun K, Lee H, et al. Skin on GV01 acupoint in colonic inflammatory states: tenderness and neurogenic inflammation. J Physiol Sci 2006;56(4):317-320. [doi: 10.2170/physiolsci.sc001606]
21. Gee MD, Lynn B, Cotsell B. The relationship between cutaneous C fibre type and antidromic vasodilatation in the rabbit and the rat. J Physiol 2004 Sep 30;503(1):31-44. [doi: 10.1111/j.1469-7793.1997.031bi.x] [Medline: 9288672]
22. Sorkin LS, Eddinger KA, Woller SA, Yaksh TL. Origins of antidromic activity in sensory afferent fibers and neurogenic inflammation. Semin Immunopathol 2018 May;40(3):237-247 [FREE Full text] [doi: 10.1007/s00281-017-0669-2] [Medline: 29423889]
23. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015 Jan 02;350:g7647 [FREE Full text] [doi: 10.1136/bmj.g7647] [Medline: 25555855]
24. de Vries RBM, Hooijmans CR, Langendam MW, van Luijk J, Leenaars M, Ritskes‐Hoitinga M, et al. A protocol format for the preparation, registration and publication of systematic reviews of animal intervention studies. Evid Based Preclin Med 2015 Feb 24;2(1):1-9. [doi: 10.1002/ebm2.7]
25. Wick MJ, Harral JW, Loomis ZL, Dempsey EC. An optimized Evans blue protocol to assess vascular leak in the mouse. JoVE 2018 Sep 12(139):57037. [doi: 10.3791/57037-v]
26. Radu M, Chernoff J. An in vivo assay to test blood vessel permeability. J Vis Exp 2013 Mar 16(73):e50062 [FREE Full text] [doi: 10.3791/50062] [Medline: 23524912]
27. Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC Med Res Methodol 2014 Mar 26;14(1):43 [FREE Full text] [doi: 10.1186/1471-2288-14-43] [Medline: 24667063]
28. Van Norman GA. Limitations of animal studies for predicting toxicity in clinical trials: is it time to rethink our current approach? JACC Basic Transl Sci 2019 Nov;4(7):845-854 [FREE Full text] [doi: 10.1016/j.jacbts.2019.10.008] [Medline: 31998852]
29. Domínguez-Oliva A, Hernández-Ávalos I, Martínez-Burnes J, Olmos-Hernández A, Verduzco-Mendoza A, Mota-Rojas D. The importance of animal models in biomedical research: current insights and applications. Animals (Basel) 2023 Mar 31;13(7):1223 [FREE Full text] [doi: 10.3390/ani13071223] [Medline: 37048478]

XSL•FO
RenderX

Abbreviations

EB: Evans Blue
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols
RoB: Risk of Bias
SYRCLE: Systematic Review Centre for Laboratory Animal Experimentation

Edited by J Sarvestan; submitted 22.Oct.2024; peer-reviewed by I Karunarathna; comments to author 04.Sep.2025; revised version received 18.Nov.2025; accepted 18.Nov.2025; published 17.Apr.2026.

Please cite as:
Yang E, Muñoz-Vergara D, Ahn AC, Lu W, Napadow V, Cronin AE, Wayne PM
Inside-Out Mechanism of Referred Visceral Pain: Protocol for a Systematic Review of Cutaneous Neurogenic Inflammation in Preclinical Animal Models of Visceral Diseases
JMIR Res Protoc 2026;15:e67852
URL: https://www.researchprotocols.org/2026/1/e67852
doi:10.2196/67852
PMID:34515028

©EunMee Yang, Dennis Muñoz-Vergara, Andrew C Ahn, Weidong Lu, Vitaly Napadow, Alexandria E Cronin, Peter M Wayne. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Role of Digital Media in Chronic Disease Self-Management: Protocol for a Multimethod Study of the DISELMA Research Consortium

Constanze Rossmann[1], MA, PhD, Prof Dr; Veronika Karnowski[2], MA, PhD, Prof Dr; Julia Metag[3], MA, PhD, Prof Dr; Juliana Raupp[4], MA, PhD, Prof Dr; Doreen Reifegerste[5], MA, PhD, Prof Dr; Claudia Riesmeyer[1], MA, PhD, PD; Nariman Sawalha[1], MA; Alexandra Lux[1], MA; Anna-Lena Esser[5], MA; Rebecca Kammerer[1], MA; Franca Singh[3], MA; Natalie Rödel[2], MA; Janine Brill[2], MA; Elisabeth Gerling[4], MA; Annemarie Wiedicke[2], MA, PhD

[1]Department of Media and Communication, LMU Munich, Oettingenstr. 67, Munich, Germany

[2]Institute for Media Research, Chemnitz University of Technology, Chemnitz, Germany

[3]Department of Communication, University of Münster, Münster, Germany

[4]Institute for Media and Communication Studies, Freie Universität Berlin, Berlin, Germany

[5]School of Public Health, Bielefeld University, Bielefeld, Germany

Corresponding Author:
Constanze Rossmann, MA, PhD, Prof Dr
Department of Media and Communication, LMU Munich, Oettingenstr. 67, Munich, Germany

Abstract

Background: Chronic diseases, such as type 1 and type 2 diabetes, asthma, and chronic obstructive pulmonary disease , demand long-term treatment and permanent adaptation. One important pillar in coping with these diseases is individuals' self-management, including support from digital media. Research on their effects confirms their potential. However, it is flawed by theoretical underdevelopment and methodological weaknesses, such as a focus on short-term effects, single digital features, and microlevel studies.

Objective: The research unit (RU) DISELMA ("Digital Media in Chronic Disease Self-Management") aims to examine the continued use patterns and effects of the digital self-management of chronic diseases, as well as the role of the interpersonal, organizational, and societal levels to gain a comprehensive picture of the individual processes, their contextual embeddedness, and cross-level interactions.

Methods: To fully capture the manifold multilevel influences, the RU comprises 6 individual projects (IPs), each of which conducts several studies. Two projects at the individual level analyze determinants of use, usage patterns, and effects of digital media, combining systematic reviews, experience sampling method studies, focus groups, panel surveys, and content analysis of apps used. Two projects examine the interpersonal context by analyzing the role of health care providers and the diffusion of digital media in informal networks, conducting a scoping review, online surveys with physicians, semistructured interviews, and participant observations of physician-patient dyads, patient focus groups, and interviews with peers. One project aims to analyze the role of organizations within the mobile health market by conducting a content analysis of organizational messages and a survey. Finally, one project analyzes journalistic and social media to gain insight into the discourses about digital chronic disease self-management on the societal level.

Results: The RU received funding approval from the Deutsche Forschungsgemeinschaft (German Research Foundation; grant 456132969) in July 2023, and the 4-year funding period ranges from December 2023 to November 2027. IP1 is currently conducting its systematic reviews and experience sampling method studies, both to be finalized in 2026. IP2 is conducting its systematic review and meta-analysis alongside panel surveys until June 2026. IP3 has completed its online survey with physicians and is currently conducting observations until August 2026. IP4 is conducting its scoping review and peer interviews through 2026, while IP5 is working on its content analysis and survey, and IP6 on its manual content analysis. First publications of the results are expected in 2026.

Conclusions: The results will contribute to the existing research through a theoretically and methodologically comprehensive approach that improves our understanding of the processes within and between all levels. These insights will inform providers of digital health solutions and health care practitioners about users' needs, advance evidence-based disease self-management programs, and contribute to better coping with chronic diseases, improved well-being of affected individuals, and reduced health care costs.

International Registered Report Identifier (IRRID): DERR1-10.2196/77811

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e77811) doi:10.2196/77811

KEYWORDS

digital media; mHealth; self-management; diabetes; COPD; asthma; ecological model; continued use; network approach; framing approach; mobile health; chronic obstructive pulmonary disease

Introduction

Background

As a consequence of both lifestyle changes and a growing number of older people, chronic diseases are the most common and economically significant health problems in industrialized countries [1-3]. Among these, type 1 and 2 diabetes, asthma, and chronic obstructive pulmonary disease (COPD) belong to those chronic diseases with the highest prevalence rates, ranging between 6% and 10% in Germany [4-7]. Given that chronic diseases are prolonged, do not resolve on their own, and are rarely cured completely, they are characterized as lifelong challenges that demand long-term treatment and permanent adaptation [8]. Despite numerous therapeutic options, disease control remains inadequate in most of these conditions [9]. In addition, most chronic diseases are heavily influenced by a patient's daily life habits, such as dietary behavior, physical activity, smoking, and medication adherence [10], requiring individuals to be responsible for their health. Consequently, self-management is becoming an important pillar in coping with chronic diseases [11].

Self-Management of Chronic Diseases

The self-management of chronic diseases embraces three different domains: (1) the acquisition of knowledge about one's condition and its treatment; (2) behavioral management, including both medical adherence and everyday life management (condition management); and (3) emotional management for maintaining adequate psychosocial functioning [12,13]. To facilitate the self-management of chronic diseases, several disease management programs (DMPs) have been developed. They offer a structured and systematized schedule and provide patients with an understanding of their disease and the behavioral skills that they need to master, monitor, and respond to their current health status [14]. In Germany, the statutory health insurance funds offer DMPs for, for example, asthma, COPD, and type 1 and 2 diabetes [15,16]. Ultimately, self-management programs aid in improving health and quality of life by helping patients gain a sense of control over their disease and preparing them for disease management at home [17,18].

However, meta-analyses examining the effects of such traditional chronic disease self-management programs across various chronic diseases show that interventions often have small to medium, unsustainable effects on clinical outcomes and behavior change [19,20]. In the context of diabetes self-management, several reviews and meta-analyses have found small improvements in disease-specific knowledge and clinical parameters [21,22]. Regarding asthma and COPD, some evidence suggests that the frequency of dyspnea is reduced and that adherence to written action plans is enhanced after following self-management programs [23]. However, almost no effects were found for hospital admissions [24,25] and health-related quality of life [24]. Furthermore, studies on both diabetes and respiratory diseases have observed a high chance of relapse, especially after interventions on weight loss, exercising, and smoking cessation [26-28], high dropout rates, and overall low participation [29].

The observation of small benefits, relapse, and low uptake of traditional self-management programs emphasizes the need for improvement, especially regarding the inclusion of emotional management in addition to mere medical management strategies [30,31], tailoring to individual needs and obstacles [32], and offering continuous day-to-day support to combat lack of self-efficacy and motivation, as well as to facilitate engagement in self-management behaviors [26]. In comparison to one-time or short-term self-management programs, interventions with regular and reinforced follow-ups have been substantially more promising [21]. However, traditional face-to-face programs are naturally limited by accessibility constraints and time restrictions [33]. These circumstances underscore the urgent need for innovative interventions that accompany patients through their everyday lives and are tailored to their specific needs—demands that can be achieved using digital media for chronic disease self-management.

Digital Media in Chronic Disease Self-Management

Over the past decades, digital media have become deeply embedded in people's everyday lives. Considering the great variety of devices, services, and platforms, digital media allow a broad range of uses, building upon the core assets of multimediality and interactivity, granting communication on the continua from synchronous to asynchronous and interpersonal to mass communication [34,35]. This digitalization has transformed virtually all aspects of our lives and every area of society [36,37]. Even more so, mobile media, which now encompasses a growing proportion of digital media, has enabled ubiquitous access to digital content across all aspects of our lives. Hence, in current media ecologies, users are most often provided with access to mobile, digital, and internet-enabled devices (such as smartphones) throughout their daily lives. These devices allow constant access to a wealth of information, communication, and other services, which users can continuously arrange and rearrange according to their specific needs [38].

Against this background, digital media technologies have become predestined to improve chronic disease self-management and facilitate long-term adherence. This potential is further enhanced through the dedicated use of digital media for health information and services (eHealth) or the use of mobile media for health purposes, such as health care or health promotion (mobile health [mHealth]) [39-41]. Indeed, recent use data have shown that information procurement for health matters is increasingly occurring online and on mobile. According to the

XSL•FO
RenderX

Health Information National Trends Survey Germany, the internet is the first port of call for 1 in 5 Germans when seeking health-related information (21%), second only to physicians (69%) [42]. In addition, digital offers are gaining momentum in the diagnosis and therapy of diseases. Almost one-third of patients already use apps to manage their disease or monitor health data [43,44].

Key use cases of eHealth and mHealth include the collection, management, processing, sharing, and storage of patient-reported data and clinical parameters [45,46], also through self-tracking systems. By affording convenient and continuous access to timely personal data and health-related information, individuals living with chronic diseases are encouraged to—quite literally—take matters into their own hands [39,47,48]. Even if this places more responsibility on individuals living with a chronic disease, they are not necessarily on their own. In fact, eHealth and mHealth can facilitate and even increase the frequency of communication between patients and health care providers [39,49,50]. In addition, by integrating technology, physicians can remotely adjust treatment plans, provide the necessary feedback, and conduct regularly scheduled online consultations [51,52]. Increasingly, informal caregivers (ie, relatives) are involved, either by accessing the same digital health system as patients or by using a stand-alone caregiver system. To fulfill their caregiving responsibilities, digital media provide informal caregivers with a platform to continuously learn about the patient's disease, help manage medical care, and provide practical and emotional social support, both mobile and online [53]. Furthermore, traditional health care providers and informal caregivers have been joined by new players, such as technology companies. These are also involved in chronic disease management through the proliferation of digital health technologies [54]. Thus, disease management becomes a collaborative process between actors on various levels in which individuals with chronic diseases are embedded and actively involved.

With these possibilities, digital health technologies can foster adherence to disease self-management more effectively than conventional health care models. While traditional self-management programs remain anchored in didactic structures, which only teach patients why, how, and when to perform certain self-management behaviors and rely on them to take them into practice [55,56], digital systems accompany individuals with chronic conditions throughout the day. Thus, digital systems can hardly be bypassed, thereby encouraging long-term adherence [57,58]. Ultimately, digital health technologies may present an avenue to strengthen patient empowerment, overcome communication barriers, and enhance treatment adherence. Against this background, health-related digital media have also unsurprisingly become a main priority of political stakeholders in Germany [59,60].

Despite this promising picture, digital technologies not only provide numerous benefits for the self-management of chronic diseases but also bear individual and societal risks associated with their use. Depending on legal standards, the increasing datafication of patients' health inevitably poses concerns about privacy and misuse of data arising from the conscious or unconscious sharing of data with physicians, service providers, and other third parties [61]. Another vein of skepticism toward digital health technologies lies in the individualization of health. As eHealth and mHealth systems allow remote adjustment of treatment plans, physician visits may occur less frequently, thus actually impairing the development and preservation of a strong and trusting physician-patient relationship [62,63]. Relating thereto, critics argue that the replacement of real-life interactions with digital ones may exacerbate the degree of social isolation and loneliness, which older people living with chronic conditions already regularly experience [64]. Finally, instead of bridging the digital divide and social gaps in health care, digital media carry the risk of widening existing health disparities. Hence, creating equal access to technology and promoting digital literacy among users are priority tasks for society to truly leverage the benefits of digital media in disease self-management [65,66].

Evidence on the Use and Effects of Digital Media in Chronic Disease Self-Management

Considering that the use of digital media for chronic disease self-management bears both potential and risks, we require profound empirical research on their effects and their conditions of use. Thus far, evidence syntheses provide an inconsistent picture, ranging from positive to limited or no effects on clinical health outcomes (refer to systematic reviews and meta-analyses for various chronic diseases [33,67-72]). For behavioral (eg, dietary behavior or medical adherence) or cognitive outcomes (eg, self-efficacy), the results are also inconsistent [73]. Holmen et al [74] concluded that the effects of research in this context are still too heterogeneous and low in quality to "provide reliable evidence of effects for stakeholders," which still held true 5 years later [69]. Frequent criticism also mentions a lack of theory and a limitation to single digital media features or apps and short-term effects [67,69,75-78].

Consequently, another branch of research started to examine how and why individuals use digital media for chronic disease self-management to explain the inconsistent evidence for its effects [79,80]. However, again, the explanatory value of most of these studies (although often theory-based) is limited because they are constrained by adoption factors, a rather narrow view of specific tools used, and critical gaps and variability in the measurement of use [81,82]. Only a few studies analyzed continued use [82,83]; however, the multifaceted patterns of everyday life integration are neglected. In a qualitative study with patients with diabetes, Rossmann et al [84] showed that those patients do not use just one specific self-management app; rather, they make use of their entire mobile media ecosystem; for example, they use WhatsApp [Meta] for communication with their caregivers, social media for exchange with other patients, or health websites to seek diabetes-related information. Thus, to fully understand the effects of using digital media in chronic disease self-management, we must look beyond the adoption of single apps and rather consider the continued use of users' entire digital media ecosystem.

Need for a Multilevel Perspective

At present, research on the use of digital media for chronic disease self-management has mostly dealt with effects from a microlevel perspective. However, considering the individual

XSL•FO
RenderX

level only unduly limits the holistic view necessary to fully grasp digital chronic disease management for two reasons:

(1) a holistic view is necessary to fully understand individual decisions, since individuals living with chronic diseases do not live in a vacuum but are surrounded and potentially influenced by their interpersonal, organizational, and societal contexts. Accordingly, studies provide evidence that the use of digital media for disease self-management highly depends on the recommendations of health care providers [84,85], the support of family members [86], the organizational implications of health data markets [87], and the public discourse in public and social media [88,89]. Thus, these influences must be integrated into research on the use and effects of digital media for chronic disease self-management. (2) Given the contradictory view of the potential and risks of digital media for chronic disease self-management, we will only draw the full picture if we gain an understanding of how professional and personal networks discuss digital chronic disease self-management, how organizational structures evolve in the context of mHealth for chronic disease self-management, and how media discourse develops around this issue.

Hence, the overarching research questions of this research unit (RU) are:

- RQ1: Which role do digital media play in the self-management of chronic diseases?
- RQ2: How is the digital self-management of chronic diseases negotiated on an interpersonal, organizational, and societal level?

Ecological Model and Objectives

To systemize the various perspectives within our multilevel approach, we adapted the idea of ecological models to digital chronic disease self-management as a basis to derive the specific objectives of the RU. Ecological models have been developed in the behavioral sciences and public health to explain people's interactions beyond behavioral models that only investigate individual characteristics and their proximal social influences [90]. In addition, an ecological model integrates multiple theories and serves "as a meta-model to ensure that environmental and policy factors are considered in developing comprehensive approaches to studying and intervening on health behaviors" [91]. To structure and integrate these different perspectives into a common framework, we developed an ecological model of digital chronic disease self-management (Figure 1).

The model illustrates interrelations between individual digital media use and its indirect (ie, individual evaluations of behavioral consequences, norm evaluations, and restriction evaluations) as well as its direct context influences at various levels (eg, privacy regulations, cost regulations, and technical developments). In the following, we will describe the factors and interrelations at various levels relevant to the understanding of digital chronic disease management that will be jointly examined in the RU.

XSL•FO
RenderX

Figure 1. Ecological model of digital chronic disease self-management.

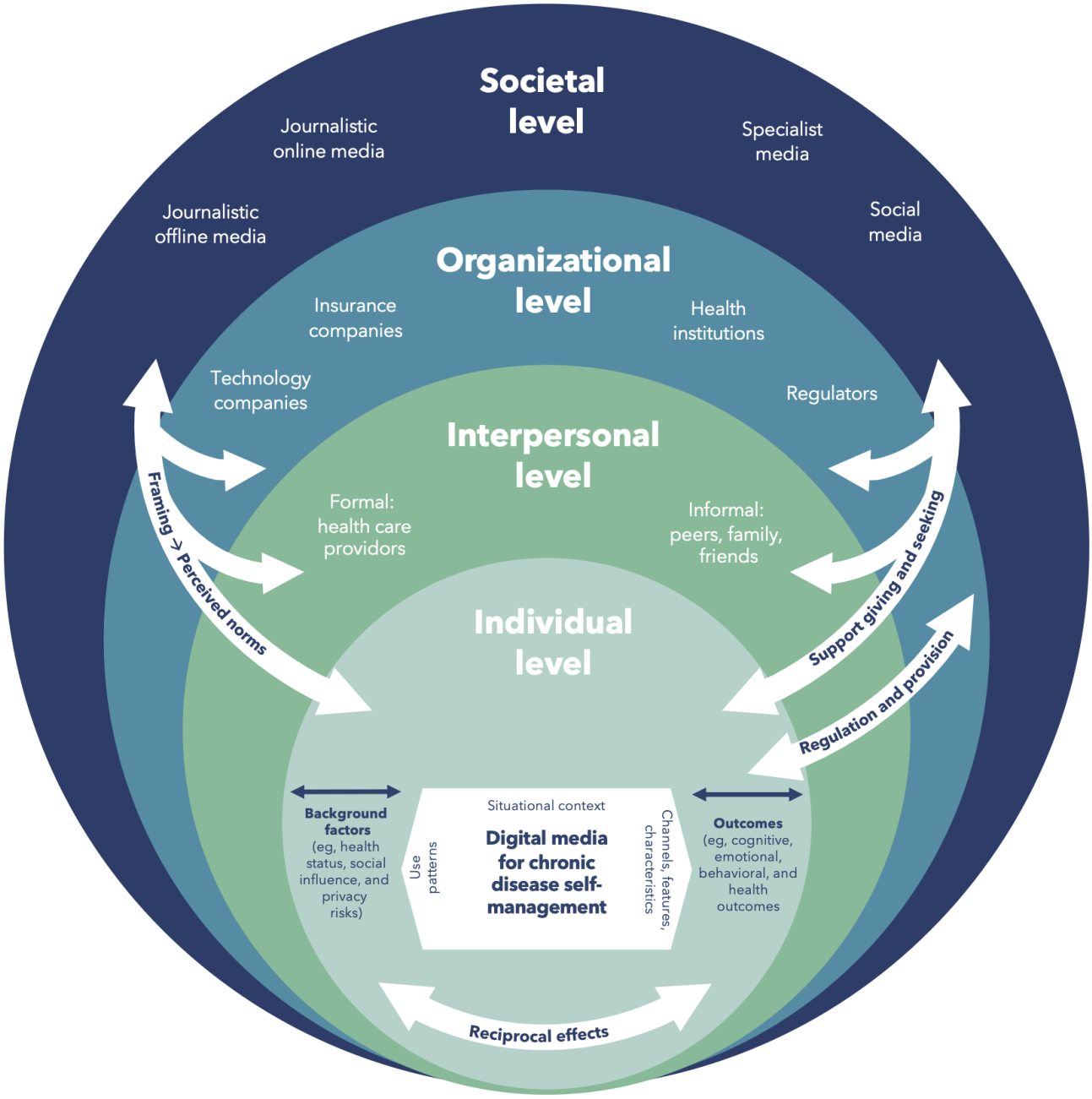

Individual Level

At the basis of the model are individuals and their manifold digital media use patterns. These patterns are determined by several factors, such as psychological determinants, media characteristics, or situational contexts [85,92-95]. The continued use leads to specific psychological, behavioral, or health-related outcomes evolving over time. Interactions between uses and effects will reinforce each other directly and indirectly via experiences, imposing an impact on future determinants of use and use patterns.

Interpersonal Level

The interpersonal level embraces individuals' social contacts. This includes both formal contacts with health care professionals and informal social contacts, such as peers, friends, and family members [96]. Both groups are related to individuals' media- and health-related behaviors by mediating knowledge and negotiating how digital media can be used, providing support, or setting norms and values, as postulated in theories of social networks [97,98] and social norms as expectations [99].

Organizational Level

Furthermore, communication about eHealth and mHealth has an important organizational dimension. Advertising and organizational messaging are central components of strategic public health communication [100]. How pharmaceutical companies, app developers, health insurance companies, and public health authorities, among others, communicate about digital self-management of chronic conditions influences the perception of digital self-management technologies at both the individual and societal levels.

XSL•FO
RenderX

Societal level

Individuals with chronic diseases and their social surroundings are embedded in and influenced by the societal context embracing the political system, economy, cultural norms, and public discourse. Since digitalization, individualization, and the occurrence of specific stressors are (meta-)trends in contemporary societies, they are also topics of public discourse. This means that public discourses about digital chronic disease self-management, as shared in (journalistic) online and offline mass and specialist media, as well as social media, offer frames of an issue associated with evaluations and expectations that are influential for organizations, regulators, health care providers, and individuals [101-103].

This comprehensive perspective not only allows the integration of different levels but also considers cross-level interactions, as apparent in support seeking and giving; formal and informal communication around digital media for chronic disease self-management; the provision and regulation of the app market in relation to individual needs; and framing effects across all levels.

The following objectives emerge from the background and the ecological model of digital chronic disease self-management:

1. To gain an understanding of the continued use patterns of digital media for chronic disease self-management resulting from determinants of use, media characteristics, and situational influences, and the effects of continued use of digital media for chronic disease self-management on cognitive, emotional, behavioral, and health outcomes as well as reciprocal effects.

2. To gain insight into the influences of interpersonal networks on digital chronic disease self-management as imposed by formal (health care providers) and informal (family, friends, and peers) support.

3. To understand the public discourse and framing of digital chronic disease self-management in strategic public health communication (including organizational messaging and corporate media) and in journalistic online and offline media, specialist media, and social media.

4. To gain a conclusive picture of the interrelations among individual, interpersonal, organizational, and societal influences on digital chronic disease self-management, as apparent in framing and perceived norms, support seeking and giving, and regulation and provision.

5. To extend the ecological model of digital chronic disease self-management by theoretically and empirically specifying cross-level relationships.

Methods

Overview of Contributions to the Objectives of the Individual Projects

To achieve these objectives, the RU will assemble multimethod studies on the individual, interpersonal, organizational, and societal levels, which are structured into 6 individual projects (IPs). These 6 IPs are conducted by 6 different teams, located at 5 different universities in Germany. Each of the IPs again combines several studies and methods to address the specific research objectives.

Two IPs are located at the individual level and analyze the determinants of continued use, usage patterns, and effects of digital media for the self-management of chronic diseases (objective 1). Specifically, these projects aim to understand the continued use of mobile media in chronic disease self-management by mapping usage situations and their situational context (IP1: use). Furthermore, they aim to understand the effects of use by integrating both a long-term perspective and heterogeneous use patterns (IP2: effects).

Three IPs aim to examine digital chronic disease self-management at the interpersonal and organizational levels (objective 2) by analyzing the role of health care providers within their institutional contexts, the informal social network, and the organizational communication context. Specifically, by transferring the actor-structure-dynamics framework [97] to the health sector, IP3 (health care) aims to understand the relationship between health care providers and patients when adopting and integrating digital media into everyday life. While this project focuses on the professional context, IP4 (peers) studies the influence of informal social contacts and related interpersonal communication on the acceptance of chronic disease–related digital media use. The third project, located at the organizational level, examines organizations' communication strategies about the use of digital media for the self-management of chronic diseases, contributing to an understanding of the organizational and institutional contexts of digital media use (IP5: organizations).

At the societal level (objective 3), IP6 (reporting) analyzes how digital media use for chronic disease self-management is represented in journalistic and social media discourses. Drawing on framing theory and news value theory, this project focuses on how the topic is constructed, which actors and frames dominate, and how scientific evidence is portrayed and evaluated in the public sphere.

Overview of IPs

The RU is funded for 48 months. In accordance with the broad spectrum of theoretical backgrounds, the IPs apply various qualitative and quantitative methods, as well as mixed methods approaches. The following paragraphs describe each IP in more detail. For an overview of the IPs and their studies and timelines, refer to Table 1.

XSL•FO
RenderX

Table . Overview of studies in each individual project (IP).

Studies and methods	Main objectives	Analysis methods	Timeline (start-end)
IP1[a]: use			
Two systematic reviews of current state of health app use (each for diabetes and asthma/COPD[b])	To gain a comprehensive overview of the measurement of "use" of mHealth[c] in the self-management of diabetes and asthma/COPD, respectively, in extant research.	Inductive-deductive content analyses	12/2023-12/2025
Two mixed-methods in-situ studies integrating quantitative surveys, mobile experience sampling studies, and automated observations for diabetes and asthma/COPD	To uncover both person- and situation-level types of mHealth use for diabetes self-management and asthma/COPD self-management.	Multilevel latent class using Latent-GOLD 6.0 (Informer Technologies) and multivariate analysis methods in R (The R Foundation).	04/2024-12/2026
Two focus groups to contextualize the findings of our 2 in-situ studies based on participants' individual and subjective experiences	To contextualize our findings with participants' subjective experiences and beliefs, and brainstorm open questions in a participatory approach.	Inductive-deductive content analyses	01/2027-09/2027
IP2: effects			
Systematic review and meta-analysis of effect studies on mobile chronic disease self-management	Develop and apply a structured framework to synthesize the status quo of mHealth system components and measured self-management outcomes and to meta-analytically assess effectiveness and moderating factors.	Narrative synthesis techniques, 3-level meta-analytical model, and meta-regressions in R.	12/2023-12/2025
Two 4-wave panel surveys to identify patterns of continued mHealth use, behavioral determinants, and long-term effects on behavioral and health outcomes, as well as reciprocal effects for diabetes and asthma/COPD	Identify patterns of continued mHealth use and explore the transactional effects in self-management.	Latent transition analysis, random-intercept cross-lagged panel model, and multigroup analyses in R.	03/2025-12/2026
Content analysis of the mHealth systems used (based on Study 2) using the developed framework (based on Study 1)	Examine the design and assess the quality of commonly used mHealth systems for self-management in Germany.	Descriptive and comparative analyses in R and integrative analysis with Study 2.	06/2026-05/2027
IP3: health care			
Online survey of physicians about their professional health-related media use, recommendation of digital media for self-management, and perception of the physician-patient relationship	Examine the self-perception of actor constellations, preferred physician-patient communication, mediation strategies, digital media within therapy, and participant recruitment for Studies 2, 3, and 4.	Multivariate analysis methods in R.	06/2024-10/2024
Semistructured interviews with 40 physician-patient dyads	Exploring self and mutual perception of actor constellation, social structures, influencing factors, and dynamics and typology of physician-patient relationship.	Theory-driven approach per MAXQDA (VERBI GmbH).	Wave 1: 05/2025-11/2025 and Wave 2: 05/2026-08/2026
Participant observations of consulting hours between the physician and the patient (Studies 2 and 3 were designed as qualitative panels with 2 waves)	Exploring self and mutual perception of actor constellation, social structures, influencing factors, and dynamics and typology of physicians-patient relationship.	Theory-driven approach per MAXQDA.	Wave 1: 05/2025-11/2025 and Wave 2: 05/2026-08/2026
Two focus groups, each with 4 patients and physicians	Discussion, interpretation, and contextualization of the results to identify further open questions.	Theory-driven approach per MAXQDA.	03/2027-06/2027
IP4: peers			

XSL•FO
RenderX

Studies and methods	Main objectives	Analysis methods	Timeline (start-end)
Scoping review on structure, content, and effects of metacommunication in digital chronic disease self-management	Examine the influence of metacommunication on knowledge, persuasion, adoption, and continued use of digital health apps.	Narrative synthesis techniques.	12/2024-12/2025
Qualitative interviews with ego-centered network maps of people with type 1 and 2 diabetes, or asthma/COPD	Exploration of the functions, content, and processes of metacommunication with informal contacts.	Inductive-deductive content analyses.	03/2025-06/2026
Online survey to identify relevant alternatives, their attributes, and tie attributes and gather knowledge about the content and perceived effects	Structural investigation into health-related ego-centered social networks of people with type 1 and 2 diabetes, or asthma/COPD.	Descriptive and multivariate techniques.	06/2026-06/2027
IP5: organizations			
Organizational analysis: typology of central organizational actors based on their characteristics, relationships, and field position	Development of a comprehensive database for further substudies and insights into the institutional structures of the digital health services sector.	Desktop research, documentary analysis, and organizational network analysis.	03/2024-06/2026
Survey of communication practitioners across sectors on attitudes toward digital media for chronic-disease self-management, with follow-up qualitative interviews	Examine organizational communication practices, networks, and affective and normative orientations toward digital health technologies.	Descriptive statistics and regressions in R and qualitative interview analysis for contextualization.	08/2025-08/2026
Automated and manual content analyses of organizational documents, including websites, press releases, and CSR[d] reports, on digital self-management of chronic diseases	Map value orientations, patient images, normative expectations, and emotional framing in organizational communication.	Website analysis using Python and LDA[e] packages in R, CSR reports analyzed in MAXQDA with cluster analysis, and descriptive statistics to describe content analyzed press releases.	03/2024-04/2026
Population survey examining how organizational communication affects perceptions of responsibility for chronic-disease self-management	Assess how organizational messages shape responsibility attributions, emotional responses, and normative expectations in the public.	Descriptive statistics and regressions in R.	02/2026-02/2027
Network analysis of central organizational actors and their social media presence, examining characteristics, relationships, and positions within the digital health field	Develop a comprehensive database and analyze how organizations communicate via social media, including how their interactions and positions shape the digital health communication landscape.	Organizational network analysis using R and descriptive analysis of desktop and documentary research.	04/2026-05/2027
IP6: reporting			
Automated content analyses of the entire coverage of chronic diseases	Investigate how the topic of digital media use for the self-management of chronic diseases is constructed within the coverage and to assess its share.	Automated content analysis using the STM[f] package in R.	12/2023-02/2026
Manual content analysis to tease out the detailed aspects of the coverage of digital media use for the self-management of chronic diseases	Provide a more detailed understanding of how the use of digital media for chronic disease self-management is portrayed in journalistic and social media coverage.	Descriptive and comparative analyses.	04/2025-09/2027

[a]IP: individual project.

[b]COPD: chronic obstructive pulmonary disease.

[c]mHealth: mobile health.

[d]CSR: corporate social responsibility.

[e]LDA: Linear Discriminant Analysis.

[f]STM: structural topic modeling.

XSL•FO
RenderX

IP1: Use

IP1 maps the specific situations of mobile media use in chronic disease self-management by building on 2 systematic reviews (Studies 1 and 3), 2 experience sampling studies (Studies 2 and 4), and reflective focus groups (Study 5).

Studies 1 and 3 map the measurement of continued use of mobile media in the self-management of diabetes and asthma/COPD, respectively. Following the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) guidelines [104], both systematic reviews are preregistered [105,106]. The selection of databases covered and development of the search string are handled in parallel with IP2, to allow for comparability. Records are stored in the reference management software Zotero (Digital Scholar) and from there transferred to Rayyan (Rayyan Systems Inc) [107]. Coding follows an inductive-deductive approach.

Studies 2 and 4 investigate both person- and situation-level types of mHealth use for diabetes self-management and asthma/COPD self-management. Both Study 2 [108] and Study 4 [109] have been preregistered. Both studies received ethical approval from Chemnitz University of Technology's Institutional Review Board (IRB). Study 2 and Study 4 each consist of a prequestionnaire, 6 months with 2 daily in-situ measurements per week for 1 week each month, and a postquestionnaire. This process will lead to an ideal maximum of 84 data points per participant (plus pre- and post-questionnaire), exceeding the recommendations given by [110]. Again, following Park and Yu [110], and adding a healthy amount of oversampling, we aim for a sample of 100 individuals per study. The studies are carried out using SoSci Survey, with the signals for the in-situ surveys being delivered via text message. A semirandom sampling scheme with predefined intervals was used for signaling: the first signal of the day was sent between 8 AM and 3 PM, and the second signal between 3 PM and 10 PM. Field time for Study 2 covers June 2025 to November 2025, with 103 diabetics taking part in the study; field time for Study 4 will start in February 2026. Both studies will be analyzed using multilevel latent class analysis in LatentGOLD 6.0 (Informer Technologies) [111,112], thus uncovering both person- and situation-level types of mHealth use for diabetes self-management and asthma/COPD self-management.

Study 5 will contextualize the findings of Studies 2 and 4 with participants' subjective experiences and beliefs, as well as brainstorm open questions for future research in a participatory approach. We will conduct a final qualitative study consisting of 2 focus groups. Interview guidelines will be developed based on the outcomes of Studies 2 and 4. Eight participants from the in-situ study on asthma/COPD and diabetes will be invited to take part in each group. Both focus groups will be audio- and video-recorded and transcribed. Data analysis will be conducted based on Kuckartz [113] using MAXQDA (VERBI GmbH).

IP2: Effects

IP2 investigates long-term use patterns and effects of mobile media in chronic disease self-management through 3 complementary studies.

Study 1 (systematic review and meta-analysis) synthesizes mHealth effect studies in type 1 and type 2 diabetes and asthma, and COPD. Following the *Cochrane Handbook for Systematic Reviews of Interventions* [114] and PRISMA statement [104], the preregistered protocol [115] specifies search, eligibility, and analysis procedures. Record screening is managed in Rayyan [107]. Drawing on scoping searches and concept mapping, we develop an organizing framework that decomposes mHealth systems into core components (devices, channels, targeted behavioral contents, features, characteristics, and theoretical grounding) and classifies addressed self-management outcomes into first- (determinants), second- (behaviors), and third-order (health-related outcomes) levels. We then empirically apply this framework by (1) narratively synthesizing mHealth systems and outcomes with frequency distributions and (2) conducting 3-level meta-analyses [116] to estimate effect sizes (Hedges *g*, 95% CI; significance will be set at $P<.05$) and heterogeneity between and within studies (Q test and multilevel I^2 statistics). Meta-regressions with omnibus tests will be used to assess moderating variables (eg, mHealth components and study and sample characteristics). Analysis will be carried out in R (The R Foundation; metafor [117] and dmetar [118]). Study quality will be assessed with the mixed methods appraisal tool [119].

Study 2 (two 4-wave panel surveys) examines how individuals with diabetes (panel 1) or asthma/COPD (panel 2) in Germany use mHealth and how this influences self-management outcomes over time. Two 4-wave online surveys (3-month lags) will be conducted via SoSci Survey with German-speaking adults (≥18 years and smartphone owners). The preregistration [120] details theoretical considerations and the developed questionnaire and specifies the 3 core objectives and analytical steps: (1) identify continued mHealth use patterns through latent class and latent transition analyses, classified by used devices, channels, and system features and characteristics, as well as intended purposes for supporting different self-management behaviors (as per Study 1). Multinomial logistic regression examines predictors of class membership and transitions (eg, determinants of use, privacy perceptions, and demographics); (2) model paths and transactional effects with cross-sectional (Wave 1; structural equation modeling) and longitudinal data (Wave 1 - 4; random intercept cross-lagged panel models) to test relations between self-management determinants, behaviors, and health-related outcomes (as per Study 1); (3) multigroup structural equation modelings and random intercept cross-lagged panel models test whether (reciprocal) relations differ across latent usage patterns. Analyses will be performed in R (poLCA [121], depmixS4 [122], and lavaan [123,124]).

Study 3 (content analysis of mHealth systems) examines the design and quality of mHealth systems most used for self-management, based on usage data from Study 2 (names of used apps and websites). All systems used by ≥5 % of participants will be included. The Study 1 framework guides codebook development. Analyses will be primarily descriptive and comparative (eg, chi-square tests). A quality index derived from meta-analytic evidence will be computed and linked to Study 2 data to test the moderating effects of system effectiveness.

XSL•FO
RenderX

IP3: Health Care

IP3 expands the overall project by adopting a multimethod design that takes a process perspective on the physician-patient relationship (formal network) and examines the negotiation and mediation processes involved in the use of health-related digital media for chronic disease self-management in their entirety.

Study 1 conducts an online survey with 203 physicians (diabetologists, nephrologists, and pulmonologists). The questionnaire was developed and programmed using SoSci Survey (field time April 2024 to October 2024). Study 1 provides information on physicians' use of and attitudes toward professionally relevant health-related digital media, as well as their stance on digital media for the self-management of chronic diseases and the perception of the physician-patient relationship. For this purpose, it includes the technology acceptance model [125], which assesses the perceived usefulness and ease of use of information technologies. To gain insights into the theoretical model of the physician-patient relationship [126] and its potential implementation in practice, it uses a validated scale on the physician-patient relationship (adapted versions of PRA and PRA-D) [127,128]. To capture the challenges and recommendations for digital media in self-management, items were developed based on a qualitative systematic review and a meta-synthesis, identifying key themes that facilitate or hinder the use of mHealth or digital health interventions [129,130], complemented by expert interviews with physicians. The questionnaire is available at Open Science Framework [131]. In addition, the surveyed physicians were recruited for Studies 2-4.

Studies 2 and 3 (semistructured interviews and participant observations, 2 waves) use a qualitative approach to examine actor constellations and social structures within the physician-patient relationship. The overarching goal is to identify patterns that characterize the physician-patient relationship during the process of adopting and integrating health-related digital media for chronic disease self-management into everyday life. Semistructured interviews with 20 physicians and 1-2 chronically ill patients (ie, diabetes, asthma/COPD), designed as physician-patient dyads as the smallest form of a formal network, are conducted (Wave 1: field time from May 2025 to November 2025 and Wave 2: field time from May 2026 to August 2026). The interviews take place separately, allowing physicians and patients to open up. The guide includes questions according to the theoretically driven categories on the use of digital media in a private and professional context, the usability and usefulness of digital technologies, as well as the opportunities and challenges of digital self-management in the health care sector. In addition, it covers aspects such as the confidentiality of information sources, social structures, and the physician-patient relationship, including knowledge transfer and shared decision-making. The guide is designed for both physicians and patients, in corresponding versions. In addition, participant observations of the consultation hours between physician and patient (Study 3) are conducted. Due to the complexity and multilayered nature of the multimethod research design, on-site data collection in the physicians' offices is necessary. Only in this way are interactions between physician and patient observable, and face-to-face interviews can be conducted. All interviews will be transcribed verbatim, anonymized, and pseudonymized. Data analysis will be conducted using a theory-driven approach with MAXQDA [132].

In Study 4 (focus groups), findings from the previous studies will be discussed, interpreted, and contextualized with physicians and patients. Two focus groups, 1 with 4 physicians and 1 with 4 patients, are planned between March 2027 and June 2027. Recruitment will take place during Studies 2 and 3. We will strive for maximum heterogeneity among the invited participants. The focus groups will be transcribed verbatim, anonymized, and pseudonymized. Data analysis will be conducted using a theory-driven approach with MAXQDA [132].

IP4: Peers

IP4 investigates the influence of informal network contacts and the related metacommunication on the everyday life integration of digital media for the self-management of the chronic conditions type 1 and 2 diabetes and asthma/COPD.

Study 1 (scoping review) synthesizes results on the structure, content, and effects of metacommunication with informal network contacts in the context of digital media used for illness self-management, based on existing literature on general health topics. The review will systematically examine the influence of metacommunication (alteri, content, and effects), determining the knowledge, persuasion, adoption, and continued use of health apps. Relevant papers will be searched in the databases Communication and Mass Media Complete, PsycInfo, PubMed, and Google Scholar, the journal JMIR (mHealth and uHealth), and the literature of included studies. Specific eligibility criteria will be based on the population, intervention, comparator, and outcome framework and adjusted to the material (eg, considering specifics such as study design, health topic, and methodology). The included studies will be coded based on quality assessment (using the mixed methods appraisal tool [119]), theories used, and results considering data sources, content, and consequences. Data will be analyzed using narrative synthesis.

In Study 2, qualitative interviews with ego-centered network maps are conducted with individuals with (1) asthma/COPD and (2) diabetes to explore the underlying mechanisms, content of metacommunication, and its effect in more detail (eg, who or what initiated the metacommunication about health-related digital media, such as patient, alteri, or a specific event). Thus, the guiding questions of the qualitative interviews are derived from adoption and appropriation models, as well as the scoping review (refer to "IP4: Peers, Study 1"). Both the content and the (perceived) effects are thereby openly queried and considered in the relationship structures. Supported by structuring questions, the interviewees are asked to report on the course of their vicarious information search (occasion, search, selection, transmission, and effects), taking their relationship to the person concerned into account. In addition, the respondents are asked about the sources of information they used for health information search beyond the informal contacts.

In Study 3, a quantitative survey with ego-centered network maps will be conducted with N=900 individuals living with

XSL•FO
RenderX

type 1 or type 2 diabetes, or asthma/COPD, to identify relevant social contacts, their attributes, tie attributes, and to gather knowledge about the content and perceived effects. The ego-centered network analysis comprises 3 steps [133]. In the first step, a so-called name generator is used to identify persons (alteri) from whom the respondent received health information in the last 12 months. In the second step, the characteristics of the alteri (so-called alteri attributes) are assessed, such as their health status, their age, and gender. In the third step, details about the relationships between the ego and the alteri (so-called tie attributes) are recorded. Participants will first be asked to name informal individuals or groups (the alteri) who are related to their health-related digital media knowledge, adoption, and appropriation on a network map. This network map will be implemented with a visualization program, such as GENSI [134], and will provide insights into the social network structure of metacommunication about health-related digital media in the self-management of chronic illnesses.

IP5: Organizations

IP5 expands the overall project by examining how organizations across sectors communicate about digital health technologies and chronic disease self-management. The project adopts a multimethod design that traces responsibilities, normative expectations, and affective evaluations across organizational actors, public interfaces, and social media environments.

Study 1 has created a comprehensive database of organizations and categorized the actors according to their functions in the organizational field. In parallel, a systematic analysis of the key organizational messages of the most important organizations in this database is in progress. Based on this, an organizational network analysis is carried out, which maps the positions of the various actor groups and the semantic network between selected actors in the institutional field of the digital health market. Furthermore, a theoretical sample of contacts for the survey is being drawn.

Study 2 comprises an online survey of communications practitioners working in the organizations identified in Study 1. The questionnaire assesses attitudes toward digital media for chronic disease self-management, perceived organizational responsibilities, emotional and normative orientations, and assessments of interorganizational communication networks. Recruitment is based on the contact information gathered in the desk research. Participants willing to engage further are invited to expert interviews exploring communicative practices, organizational strategies, and sector-specific perspectives on mHealth technologies. Survey data are analyzed descriptively and through multivariate models, while interview data are examined qualitatively to contextualize the quantitative findings.

Study 3 combines several content analyses. In-depth content analyses of app providers' websites and organizational reports (eg, corporate social responsibility and diversity, equality, inclusion reports) of major providers of digital health solutions are conducted with regard to the communication of central values, the portrayal of patients and users, and the attribution of responsibilities. Furthermore, automated and manual content analyses of organizational documents, including websites, press releases, and promotional materials, will be conducted to examine how digital media and chronic disease self-management are communicated across various social sectors. The analysis is guided by established approaches to organizational framing and strategic communication [135], which inform the identification of value orientations, patient images, normative expectations, and affective elements. Automated procedures such as dictionary-based emotion detection, semantic network analysis, and indices of empowerment and responsibilization are combined with qualitative coding. This allows us to capture framing patterns and to map how organizations construct digital self-management in their public communication in terms of responsibility, support, risk, and emotion.

Study 4 consists of a population-level online survey that examines how organizational communication affects public perceptions of chronic-disease self-management and digital health technologies. Respondents view selected organizational materials from Study 3 and provide information on their own exposure to organizational campaigns. The questionnaire assesses perceived responsibility for self-management, emotional reactions to digital health tools, and normative expectations associated with chronic conditions. Items also cover evaluations of empowerment, burden, legitimacy, and trust. Data will be analyzed with multivariate statistical techniques to identify how organizational communication influences responsibility attributions, emotional responses, and perceived norms in the general population.

Study 5 investigates how organizational communication circulates and transforms within social media environments, with a focus on platforms such as TikTok (Beijing Bytedance Technology Ltd) and Instagram (Meta). The analysis includes posts from organizational accounts such as health insurers, governmental institutions, NGOs, and pharmaceutical companies, as well as influencer content that engages with or integrates organizational messages. In addition, user comments beneath posts are collected to capture audience reactions, interpretations, and contestations of organizational communication. Sampling includes platform-specific practices such as remixes, collaborations, and responses to organizational campaigns. Automated extraction of textual and metadata features is combined with qualitative coding to identify framing strategies, normative cues, emotional expressions, and responsibility dynamics. The study examines how organizational communication is amplified, reframed, or challenged through the interaction of institutional actors, creators, platform logics, and user engagement.

IP6: Reporting

IP6 comprises 2 complementary studies: an automated content analysis (Study 1) and a manual content analysis (Study 2).

Study 1 (automated content analysis) analyzes media coverage of chronic diseases (ie, diabetes, asthma, and COPD) in leading German media outlets from 2010 to the end of June 2025 (N=27,423). The media articles stem from German print and online newspapers, trade and customer magazines, television programs, Instagram post captions from newspapers, Instagram post captions from health influencers, online news sites, and from 4 health forums dedicated to diabetes, asthma, and COPD. The selection of specific media outlets was based on their reach

XSL•FO
RenderX

and relevance in the German media landscape. Thus, newspapers and magazines with high circulation figures, online news platforms and health forums with large audiences and members, and health influencers with high follower counts were included. This diverse selection of media types allows for the analysis of different communication logics, ranging from traditional journalistic reporting to personalized storytelling on social media. To identify topics, we conducted a structural topic modeling in R [136]. From this underlying corpus of public discourse on chronic diseases, a subset of articles explicitly addressing the use of digital media for chronic disease self-management will be extracted.

Study 2 (manual content analysis) analyzes how the use of digital media for chronic disease self-management is portrayed in journalistic and social media coverage in more detail. Key categories of the codebook include covered topics, actors, evaluations of health-related digital media, and news factors such as relevance, controversy, personalization, prominence of actors, reach, and negativity [137]. Frames are coded following Dan and Raupp's [135] typology of frames in health communication (eg, health severity, medical, consequence, gain, and loss). Furthermore, we code how scientific evidence on digital self-management is represented, both explicitly and implicitly, including references to the certainty and uncertainty of research findings [138]. Before starting the manual coding, coders undergo extensive training, and pretests will be run. Krippendorff alpha is used to assess intercoder reliability. Results from both automated and manual analyses will be integrated to provide a thorough interpretation of how digital self-management of chronic diseases is constructed and contextualized in public discourse.

Joint Research Activities

The global aim of the RU is the integration of IPs within different levels, as well as across different levels, to gain a conclusive picture of the interrelations among the individual, interpersonal, organizational, and societal levels in the context of digital chronic disease self-management (objective 4) and to extend the ecological model of digital chronic disease self-management by specifying interrelations between the different levels (objective 5). These objectives will be achieved by sharing common concepts, jointly developing measurement instruments, and joint data analysis and integration of results from all IPs within the RU, as will be described in the sections on collaborations within concept groups, bi- and multilateral collaborations, and overall integration.

Collaboration Within Concept Groups (CGs)

The RU shares several key concepts guiding the IPs and joint activities. Beyond self-management, these pertain to continued use, the network approach, and framing:

Continued Use

A prerequisite for the effective implementation of digital media in chronic disease self-management is that individuals not only adopt this technology for a short time, but also integrate the use of digital media for chronic disease self-management into their everyday lives. Nevertheless, previous research on the effects of digital media in chronic disease self-management is dominated by short-term randomized trials, neglecting aspects of use, and research on the use of digital media for chronic disease self-management is mostly bound to adoption decisions, neglecting the crucial role of continued use and everyday life integration. Furthermore, previous research often misses the fact that users differ in their drivers of use and usage patterns, especially when considering not only single apps but also the entire digital media ecosystem. Therefore, this concept group (CG) focuses on the drivers and patterns of the continued use of digital media ecosystems. This common understanding is especially prevalent in IP1, IP2, IP3, and IP4, which work together in the CG (continued use). The IPs in this group examine, with varying foci, the individual use of digital media for chronic disease self-management and assess continued use as part of their IPs. On a theoretical level, the IPs reflect on the patterns of continued use and strive to overcome the limitation of focusing on simple adoption decisions and highlight the circumstantiality of such once-off decisions. To optimize the measurement and produce comparable data, this CG jointly develops and implements measures to assess use. Finally, the CG will jointly interpret (and publish) their data, thus contributing to objectives 1, 4, and 5.

Network Approach

Previous research indicates that the role of digital media in chronic disease self-management cannot be fully assessed by considering only individuals living with a chronic disease. In contrast, the individual is embedded in a network of health care providers, family, friends, peers, and organizations that may differently influence the use and effects of digital media for chronic disease self-management, which is evidenced by recommendations to use specific apps, the provision of apps to individuals and health care providers, and seeking and giving social support in interpersonal contacts or via social media. Network theory and empirical network analyses are well established in communication science and have become increasingly important in recent years. Several projects apply a network approach and thus broaden the methodological scope of the RU. Network approaches are especially relevant in IP3, IP4, IP5, and IP6, which consequently collaborate within the CG (network approach). Even if these IPs use different specific theories (eg, action theories, theories of adoption and appropriation of media technologies, and social and semantic network theory) and different multimethodological approaches (eg, online surveys, semistructured interviews, participant observations, focus groups, automated content analyses, and ego-centered network maps), the IPs share a common conceptual ground regarding the role of networks, which shall be fostered through joint discussions of various relevant theories and methodological approaches. In this way, the CG will further integrate findings from network approaches into communication research. Finally, integrated analyses will be conducted to contribute to the role of networks and both theoretical and methodological learnings, thus contributing to objectives 2, 4, and 5.

Framing Approach

Beyond personal and organizational networks, public and organizational discourses in journalistic and social media and organizational outlets may affect how individuals deal with

XSL•FO
RenderX

digital media in managing their diseases as well. The pertaining IPs (IP5: organizations and IP6: reporting) work together in the CG framing approach, which commonly draws on the framing approach to identify and describe these discourses. With framing as a key concept within communication research, this CG outlines its specific relevance for the RU and allows for a critical interrogation of its meaning for this specific field. Given the manifold definitions and applications, as well as disciplinary traditions of framing, the IPs exchange their perspectives on framing, find a common conceptual ground, and exchange measures of frames for the identification of frames in the respective IPs. To provide a broad picture of the framing of digital chronic disease self-management from different perspectives, we will integrate the findings to commonly contribute to objectives 2, 3, 4, and 5.

Bi- and Multilateral Collaborations Between the IPs

Beyond the key concepts and pertaining CGs, the IPs are interlaced in further bi- and multilateral collaborations resulting from common theoretical or methodological approaches, shared measurements, and objects of investigation. For example, IP1 and IP2 share a common theoretical meaning regarding continued use, use patterns, and the conceptualization of mobile media ecosystems. In addition, methodological connections exist between the systematic reviews conducted in both IPs and the common preparation of survey constructs pertaining to use. On this basis, the results of both IPs shall be integrated through method triangulations, contributing to the improvement of theory on use and effects.

Another example is IP3, which has strong ties with IP4, IP5, and IP6. Being located at the interpersonal level, both IP3 and IP4 refer to network approaches as part of the theoretical background and will integrate their results to provide a full picture of the role of formal and informal networks in digital chronic disease self-management. IP3 and IP5 also share the network approach as their common key concept, the notion of embedded agency, as well as their interest in the relevance of various health organizations for physicians and patients. They collaborate in their preparation of interview guidelines and in joint analyses of the mutual perceptions and expectations of market actors, physicians, and patients. In addition to bilateral collaborations, IP3 will also engage in a multilateral collaboration with IP4 and IP6 to jointly develop specific categories in the codebook of IP6 and questions in the interview guides of IP3 and IP4. Through this, the 3 IPs will be able to compare the content of exchanges with content in media coverage and jointly analyze mentions of media content in the exchanges.

Overall Integration

The global aim of the RU is the integration of the IPs within the different levels, as well as between the different levels, to gain a conclusive picture of the interrelations among the individual, interpersonal, organizational, and societal levels in the context of digital chronic disease self-management (objectives 4 and 5). At this stage, the interrelations between the levels are mostly examined by assessing perceptions; that is, IP2 integrates perceived collective norms of digital chronic disease self-management that are understood as a result of framing in public discourse (IP6: reporting) and by organizations (IP5: organizations). Similarly, IP3 and IP5 assess, among other things, how providers of mHealth services anticipate the expectations of their users (IP1: use and IP2: effects). This enables us to specify potential interdependencies. Accordingly, the theoretical goal of these integrated analyses is to improve the ecological model of digital chronic disease self-management over the course of the 4 years of the RU by theoretically specifying (causal) cross-level relationships (objective 5). Consequently, the results from all IPs will be jointly interpreted in the final phase of Year 4.

Ethical Considerations

All studies collecting data from humans will adhere to the principle of voluntariness and informed consent to participate in the study and to data storage and data publication provided in advance. Studies working with external access panel providers will ensure that the panel provider fulfills the quality criteria with regard to adherence to data protection and ethical rules. In any case, data containing personal information (eg, qualitative interviews) will be stored only as long as necessary and anonymized as soon as possible during the research process to ensure data protection. No datasets with personal information will be shared in open science repositories. Within the framework of a data management plan, binding agreements on the rights of use of the data and materials, their storage locations, and naming conventions will be made.

In addition, each IP will seek IRB approval for each of their studies with human participants. Since each IP conducts multiple studies, many of which build on the results of previously conducted studies, ethical approvals will be obtained in a stepwise process, shortly before the pertaining studies are about to start their fieldwork. Up to now, ethical approvals have been obtained for the following subprojects.

IP1: Use

Study 2 (experience sampling method [ESM] study on diabetes), ethical approval granted by the IRB of Chemnitz University of Technology on May 28, 2025 (approval no 101751518); Study 4 (ESM study on asthma/COPD), ethical approval granted by the IRB of Chemnitz University of Technology on November 14, 2025 (extension to approval no 101751518).

IP2: Effects

Study 1 (systematic review and meta-analysis), ethical approval granted by the IRB of the Department of Media and Communication, LMU Munich on May 22, 2024 (approval no 2024 - 11); Study 2 (panel survey), ethical approval granted by the IRB of the Department of Media and Communication, LMU Munich on May 20, 2025 (approval no 2025 - 11).

IP3: Health Care

Studies 1-4 (online survey, semistructured interviews, participant observations, and focus groups), approval granted by the IRB of the Faculty of Social Sciences, LMU Munich, on December 6, 2022 (approval no GZ22-05).

XSL•FO
RenderX

IP4: Peers

Study 2 (qualitative ego-centered interviews with network maps), approval granted by the IRB at Bielefeld University on November 27, 2024 (approval no EUB-2024 - 328).

IP5: Organizations

Study 2 (interviews with communication practitioners), approval for quantitative interviews granted on September 29, 2025 (approval no ZEA24041), approval for qualitative interviews granted on February 10, 2025 (approval no ZEA25027), both by the IRB at Freie Universität Berlin.

IP6: Reporting

Study 1 (automatic content analysis) and Study 2 (manual content analysis), approval granted by the IRB of the Department of Communication, University of Münster, on May 2, 2024.

For details on ethical aspects per study, refer to Table 2.

XSL•FO
RenderX

Table . Detailed description of ethical aspects per study.

Study	Human participants	Informed consent, opt-out option	Anonymization	Compensation	No identification in supplements	Ethical approval obtained (approval no)
IP1:[a] Study 1	—[b]	—	—	—	—	—
IP1: Study 2	Yes	Yes	Yes	Participants taking part in at least 50% of the in-situ surveys receive 75 EUR (a currency exchange rate of 1 EUR=US $1.04 to US $1.17 is applicable).	Yes	101751518
IP1: Study 3	—	—	—	—	—	—
IP1: Study 4	Yes	Yes	Yes	Participants taking part in at least 50% of the in-situ surveys receive 75 EUR	Yes	Extension to 101751518
IP1: Study 5	Yes	Yes	Yes	25 EUR	Yes	Not yet submitted
IP2: Study 1	—	—	—	—	Yes	November 2024
IP2: Study 2	Yes	Yes	Yes	Wave 1: 2.50 EUR, Waves 2 and 3: 4 EUR, and Wave 4: 7 EUR	Yes	November 2025
IP2: Study 3	—	—	—	—	Yes	TBD[c]
IP3: Study 1	Yes	Yes	Yes	100 EUR (cash, voucher, or donation)	Yes	GZ22-05
IP3: Study 2 and 3	Yes	Yes	Yes	Physicians: 400 EUR for both waves and patients: 300 EUR for both waves	Yes	GZ22-05
IP3: Study 4	Yes	Yes	Yes	100 EUR	Yes	GZ22-05
IP4: Study 1	—	—	—	—	—	—
IP4: Study 2	Yes	Yes	Yes	100 EUR	Yes	2024 - 328
IP4: Study 3	Yes	Yes	Yes	TBD[c]	Yes	Not yet submitted
IP5: Study 1	—	—	—	—	—	—
IP5: Study 2	Yes	Yes	Yes	None[d]	Yes	Survey: ZEA25027 and interview: ZEA24041
IP5: Study 3	—	—	—	—	—	—
IP5: Study 4	Yes	Yes	Yes	None[d]	Yes	Not yet submitted
IP5: Study 5	—	—	Yes	None[d]	Yes	—
IP6: Studies 1 and 2	—	—	Names of health influencers will be anonymized in datasets	—	Yes	—

[a]IP: individual project.

[b]Not applicable.

[c]TBD: to be determined.

[d]No compensation provided.

XSL•FO
RenderX

Results

General Status of the RU

The project was submitted to the Deutsche Forschungsgemeinschaft in a 2-stage process: the first-stage proposal was submitted in July 2020, reviews and the call to submit the full proposal were obtained in January 2022. Thereafter, we submitted the full proposal in December 2022, orally presented it to the reviewer committee in April 2023, and finally received the funding acknowledgment in July 2023, together with reviews for each of the IPs (refer to Peer Review Report 1 for the reviews in German). After a preparation process (implementation of the projects at the departments, hiring of PhD candidates), the RU started its 4-year funding period in December 2023.

Status of the IPs

IP1: Use

IP1 has finalized the systematic review on continued use of digital media in diabetes self-management (Study 1 of the IP) and is currently working on finishing the systematic review on continued use of digital media in asthma/COPD self-management (Study 3 of the IP). In addition, Study 2, the ESM study on continued use of mobile media in diabetes self-management, is currently being conducted. The last of 6 waves of data collection will be finished by November 2025, and data analysis using multilevel latent class analysis, following the 3-step procedure proposed by Lukočienė et al [139], using LatentGOLD 6.0 [111,112] will start. At the same time, preparations for Study 4 (mirroring the procedures for Study 2) have started, and field time for this study will start in February 2026. Study 5, two focus groups with participants from Studies 2 and 4 to contextualize the results, is planned for early 2027.

IP2: Effects

IP2 is currently conducting a systematic review and meta-analysis of studies on mobile chronic disease self-management (Study 1). The planning phase and literature search took place from December 2023 to January 2024, yielding 4852 records. Abstract and full-text screening was conducted from February 2024 to August 2024, narrowing the selection to 202 studies for review. A codebook has been developed, and data extraction is underway. R scripts for data analysis are being prepared. Completion of Study 1 is planned for December 2025.

As of March 2025 to June 2025, we prepared the questionnaire, preregistration, and IRB application for the 4-wave panel surveys (Study 2). Field time for both surveys is scheduled to take place from June 2025 to March 2026, with a time lag of 3 months between each wave. Data collection for Waves 1 and 2 has been completed, resulting in a sample size of n=2171 for diabetes and n=2438 for asthma/COPD in Wave 1, and n=1446 for diabetes and n=1531 for asthma/COPD in Wave 2. The next immediate steps include data cleaning and preparation for Waves 3 and 4, followed by the implementation of cross-sectional and longitudinal analytical models (October 2025 to December 2026).

In parallel, the content analysis (Study 3) will be prepared and conducted over a period from June 2025 to March 2027. This will involve extracting and persisting data on commonly used mHealth systems identified over the course of Study 2. Following the completion of these studies, integrative data analysis will be performed from March to May 2027.

IP3: Health Care

IP3 conducted an online survey with physicians from June to October 2024 to analyze their professional use of, perceptions of, and attitudes toward digital media for the self-management of chronic diseases (Study 1), as well as their perception of the physician-patient relationship and their recommendation behavior regarding digital media for chronic disease self-management [131]. Diabetologists, nephrologists, and pulmonologists (N=2808) were recruited via the National Association of Statutory Health Insurance Physicians (KBV) database; n=203 participated in the survey. Physicians were invited by mail, followed by a reminder letter and email. As an incentive, participants could either donate €100 (approximately US $110) to a charity or receive it as a voucher. Data were analyzed using R.

Furthermore, IP3 is conducting Studies 2 and 3. Semistructured interviews and participant observations of 40 physician-patient dyads will be conducted from May to November 2025 (Wave 1) and May to August 2026 (Wave 2) to examine the physician-patient relationship (including comparisons of physicians' and patients' self- and mutual perceptions), as well as the recommendation, adoption, and integration of digital media for chronic disease self-management into everyday life. The data will be transcribed verbatim, anonymized, and pseudonymized. The data analysis is based on a theory-driven approach and uses MAXQDA. The coding system is currently being pretested. Coding will start after the data collection of Wave 1. Two focus groups (Study 4) will be conducted between March and June 2027 to contextualize the results.

IP4: Peers

First, IP4 conducted a scoping review to analyze the structure of informal network contacts and their role in the use of digital media for the self-management of chronic diseases (Study 1). The preparation and literature search resulted in n=975 hits. After the removal of duplicates, a title and abstract screening was carried out and completed in April 2024. A full-text screening of n=62 identified sources, and the preparation of data analysis is currently ongoing and planned to be completed by December 2025. In addition, IP4 conducted qualitative interviews with peers (Study 2). The corresponding ethics application for the study was approved in November 2024, and the study was carried out from March to September 2025. Study 3, a quantitative survey of peers, is scheduled for 2026. Integrative analysis of qualitative and quantitative data will be carried out in 2027.

IP5: Organizations

After completion of Study 1, Study 2 has entered the data collection phase; around 1800 communications practitioners from a broad range of health-sector organizations were identified and contacted, and the survey will remain in the field until

XSL•FO
RenderX

February 2026, with first descriptive examinations planned for the end of 2025. Study 3 has completed the compilation and analysis of organizational communication materials. A corpus of approximately 250 organizational websites was crawled, from which keyword and length filters produced around 1000 relevant documents. Complementing this, a focused press release corpus of n=146 items from 38 organizations has been analyzed, with reliability checks completed. Across both datasets, recurring patterns appear in responsibility attributions, value orientations, emotional expressions, and representations of patients in communication on digital self-management. These materials form the empirical foundation for the stimulus development in Study 4, for which questionnaire construction is currently in progress, and fieldwork is scheduled for the middle of 2026. Study 5 has not yet commenced; sampling strategies for organizational posts and user comments on major social media platforms, as well as the technical framework for extraction and coding, are currently being discussed.

IP6: Reporting

IP6 has created and validated 2 search strings: 1 to identify journalistic articles covering chronic diseases and another to capture articles specifically addressing the use of digital media for chronic disease self-management. Using these search strings, the main corpus has been compiled, consisting of n=27,423 media articles from German print and online newspapers, trade and customer magazines, television programs, Instagram post captions from newspapers, Instagram post captions from health influencers, online news sites, and 4 health forums dedicated to diabetes, asthma, and COPD. Media articles in this corpus were published between January 2010 and June 2025. Compiling media articles from these various sources was carried out between April 2024 and May 2025. In June 2025, the structural topic modeling was calculated. In total, researchers identified 14 topics within the overall media coverage on chronic diseases: forum discussions on glucose monitoring, lifestyle, cardiovascular diseases, scientific research, health insurance companies, asthma and COPD, diabetes management among adolescents, comorbidities and complications of diabetes, diet and weight loss, food and nutrition, digital diabetes self-management, diabetes, obesity and weight-loss injections, and family. Notably, digital self-management only appears as a distinct topic in the context of diabetes (n=1222), suggesting that discussions of digital health technologies are currently concentrated within diabetes-related reporting, while asthma and COPD receive little attention in this regard. As of November 2025, IP6 has begun pretesting the codebook for the manual content analysis, which will further explore the framing of digital self-management, the representation of scientific evidence, and the role of news factors. Relevant articles will be extracted from the main corpus using the second validated search string, focusing on digital self-management.

Discussion

Principal Findings

By bringing together the results from the IPs and collaborating in bi- and multilateral constellations and within the CGs, the RU will be able to jointly develop best practices in measurement, share data and results, and interpret them jointly. Moreover, this collaborative venture will contribute not only to the objectives formulated in each IP but also to common objectives beyond the individual scopes (as outlined in the “Ecological Model and Objectives” section). Finally, the results of all IPs will be integrated to contribute to the overarching objectives of the RU. The expected results will have both scientific and practical implications.

Regarding scientific contributions, four aspects are worth mentioning: (1) both the independent and the integrated results will contribute to improving the theoretical multilevel understanding of digital chronic disease self-management at the individual, interpersonal, organizational, and societal levels and between the levels [84-89]. The IPs mainly operate on one level but are closely related via perceived influences from other levels. By integrating the results from different levels, the results from this stage of the RU will allow for specification of the interrelations between the different levels; (2) the relevance and application spectrum of our results for digital chronic disease self-management have implications beyond this field of research. Disease self-management is closely related to coping, which, in a broader sense and as described in the transactional stress model [140], embraces problem- and emotion-focused strategies to overcome stressors beyond disease-related stressors. Thus, by analyzing coping in this context and applying a multilevel perspective, our results will also contribute to the field of media-related coping research, especially with regard to a closer understanding of problem-focused strategies and the recognition of a situational perspective that has been mostly neglected in coping research [141,142]; (3) beyond coping, many communication phenomena, which are usually examined solely with an individual focus or a societal focus will benefit from a multilevel perspective. Sallis and Owen [91] indicated that despite their strengths, ecological models generally are characterized by a “lack of specificity about the most important hypothesized influences” and “the lack of information about how the broader levels of influence operate or how variables interact across levels.” Thus, this project will contribute to previous research in this field by proposing theory-based interrelations and specifying these interrelations relevant to other fields of communication research; (4) the results of this RU will also contribute to methodological development. This refers not only to commonly developed and tested measurement scales (eg, measurement of use within the mobile media ecosystem) but also to insights into experience sampling research, computational methods, and long-term panel research, as well as qualitative and quantitative network analyses.

Moreover, the results of this RU will have several practical implications: (1) the results will provide insights into the existing app market, the continued use of the entire digital media ecosystem by individuals living with chronic diseases [82-84], their effects [33,67-72], and the relevant communication strategies of health care providers, peers, and organizations [96-100], as well as the public discourse on this issue [101-103]. With this, the results will be highly relevant for app developers and health care providers by providing insights into users’ needs and expectations, effective strategies, and the surrounding discourse [100]; (2) furthermore, the results can inform the

XSL•FO
RenderX

improvement and advancement of evidence-based disease self-management programs [15,16]. As outlined in the "Introduction" section, traditional disease self-management is often ineffective [19-29]. Thus, by providing evidence on the mechanisms of effective digital disease self-management programs and their surrounding influences, DMPs can be improved; (3) in the long term, this will potentially contribute to better coping with chronic diseases, thus improving the well-being of many people living with chronic diseases; and (4) on a larger scale, this may also be a small but important piece of the puzzle to reduce health care costs in the future.

Limitations

Our ecological model is by no means exhaustive; it does not include all factors potentially relevant to the interplay of digital media and chronic disease self-management. For example, on the societal level, cultural and political factors may also play a role. However, their influences emerge as part of media discourse or within institutional and organizational structures, and the structural influences on health care providers become evident in formal interpersonal contexts. Thus, our proposed ecological model and the work of the RU will still provide a concise framework to structure and assess the communication processes most relevant to an integrative understanding of digital chronic disease self-management.

Comparison With Previous Work

Previous research on the effects of digital chronic disease self-management reveals some evidence for its potential [33,67-72] but is sometimes flawed by theoretical underdevelopment and methodological weaknesses, such as the focus on short-term effects, single digital features, and microlevel studies lacking external validity [67,69,75-78]. In addition, it lacks a comprehensive perspective reaching beyond the individual level [74].

Therefore, this RU contributes to existing research by proposing an ecological model of digital chronic disease self-management, integrating different communication perspectives and levels of analysis as well as their interactions, and examining the continued use patterns and effects of digital chronic disease self-management, as well as the role of the interpersonal, organizational, and societal levels, to gain a comprehensive picture of the individual processes, surrounding mechanisms, and cross-level interactions.

Conclusions

By the end of the project, the results will contribute to the existing research on digital self-management through a theoretically and methodologically comprehensive approach, uniquely situated in communication studies, to improve the understanding of the processes at and between all levels. The results in this context will also allow for conclusions to be drawn about the broader interplay of digital media and coping with modern stressors.

Acknowledgments

We used DeepL and ChatGPT to support translation and proofreading. No further AI support was used.

Funding

The preparation of the project proposal was funded by the University of Erfurt and LMU Munich; the RU is funded by the Deutsche Forschungsgemeinschaft (DFG; funding no 456132969). We thank all colleagues, critical friends, advisory board members, and reviewers who provided their valuable inputs to the conceptualization of this research proposal, thus contributing to the improvement of this work.

Data Availability

Studies have been or will be preregistered on suitable platforms (eg, PROSPERO for systematic reviews and meta-analyses and Open Science Framework (OSF) for experiments and surveys). All measurement instruments and stimuli will be made publicly available on open online repositories (eg, OSF). In addition, collected data will be published online if permitted under data protection regulations (exceptions include qualitative data and data on the situational contexts of use). Please see the IPs' method descriptions above for details.

Authors' Contributions

CR, VK, ClR, DR, JM, and JR contributed to the conceptualization of the RU. CR, DR, NS, VK, ClR, JM, and JR drafted the research proposal on which this manuscript is based. CR, VK, JM, ClR, DR, JR, NS, AL, ALE, RK, FS, and NR drafted the first version of this manuscript. CR, VK, JM, ClR, DR, JR, NS, AL, ALE, RK, FS, JB, EG, and AW contributed to the revised version.

Conflicts of Interest

None declared.

Peer Review Report 1

Peer review report by the Deutsche Forschungsgemeinschaft (DFG) / German Research Foundation.

XSL•FO
RenderX

[PDF File, 1270 KB - resprot_v15i1e77811_app1.pdf]

References

1. World health statistics 2023: monitoring health for the SDGs, sustainable development goals. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240074323 [accessed 2025-10-29]
2. Choi M, Sempungu JK, Lee EH, Lee YH. Living longer but in poor health: healthcare system responses to ageing populations in industrialised countries based on the findings from the Global Burden of Disease Study 2019. BMC Public Health 2024 Feb 22;24(1):576. [doi: 10.1186/s12889-024-18049-0] [Medline: 38388412]
3. Cordova R, Viallon V, Fontvieille E, et al. Consumption of ultra-processed foods and risk of multimorbidity of cancer and cardiometabolic diseases: a multinational cohort study. Lancet Reg Health Eur 2023 Dec;35:100771. [doi: 10.1016/j.lanepe.2023.100771] [Medline: 38115963]
4. Eder W, Ege MJ, von Mutius E. The asthma epidemic. N Engl J Med 2006 Nov 23;355(21):2226-2235. [doi: 10.1056/NEJMra054308] [Medline: 17124020]
5. Heidemann C, Scheidt-Nave C. Prävalenz, Inzidenz und Mortalität von Diabetes mellitus bei Erwachsenen in Deutschland – Bestandsaufnahme zur Diabetes-Surveillance (prevalence, incidence and mortality of diabetes mellitus in adults in Germany - inventory of diabetes surveillance). J Health Monit 2017;2(3). [doi: 10.17886/RKI-GBE-2017-050]
6. Steppuhn H, Kuhnert R, Scheidt-Nave C. 12-Monats-Prävalenz von Asthma bronchiale bei Erwachsenen in Deutschland (12-month prevalence of bronchial asthma in adults in Germany). J Health Monit 2017;2(3). [doi: 10.17886/RKI-GBE-2017-052]
7. Steppuhn H, Kuhnert R, Scheid-Nave C. 12-Monats-Prävalenz der bekannten chronisch obstruktiven Lungenerkrankungen (COPD) in Deutschland (12-month prevalence of known chronic obstructive pulmonary disease [COPD] in Germany). J Health Monit 2017;2(3). [doi: 10.17886/RKI-GBE-2017-053]
8. Sinclair C. Effectiveness and user acceptance of online chronic disease management interventions in rural and remote settings: systematic review and narrative synthesis. Clin Med Insights Ther 2015 Jan;7. [doi: 10.4137/CMT.S18553]
9. Alkandari A, Vaucher J, Marques-Vidal P. Trends in glycemic, blood pressure, and lipid control in adults with diabetes in Switzerland: the CoLaus|PsyCoLaus Study. BMJ Open Diabetes Res Care 2023 May;11(3):e003377. [doi: 10.1136/bmjdrc-2023-003377] [Medline: 37188394]
10. Noncommunicable diseases. World Health Organization. URL: https://www.who.int/health-topics/noncommunicable-diseases#tab=tab_1 [accessed 2025-11-28]
11. Huygens MWJ, Vermeulen J, Swinkels ICS, Friele RD, van Schayck OCP, de Witte LP. Expectations and needs of patients with a chronic disease toward self-management and eHealth for self-management purposes. BMC Health Serv Res 2016 Jul 8;16:232. [doi: 10.1186/s12913-016-1484-5] [Medline: 27391471]
12. Clark NM, Becker MH, Janz NK, Lorig K, Rakowski W, Anderson L. Self-management of chronic disease by older adults. J Aging Health 1991 Feb;3(1):3-27. [doi: 10.1177/089826439100300101]
13. Corbin J, Strauss A. Managing chronic illness at home: three lines of work. Qual Sociol 1985;8(3):224-247. [doi: 10.1007/BF00989485]
14. Lennon S, McKenna S, Jones F. Self-management programmes for people post stroke: a systematic review. Clin Rehabil 2013 Oct;27(10):867-878. [doi: 10.1177/0269215513481045] [Medline: 23543340]
15. Disease-management-programme (DMP). Gemeinsamer Bundesausschuss. 2019. URL: https://www.g-ba.de/themen/disease-management-programme/ [accessed 2025-12-22]
16. Burns J, Kurz C, Laxy M. Effectiveness of the German disease management programs: quasi-experimental analyses assessing the population-level health impact. BMC Public Health 2021 Dec;21(1). [doi: 10.1186/s12889-021-12050-7]
17. Udlis KA. Self-management in chronic illness: concept and dimensional analysis. J Nurs Healthc Chronic Illn 2011 Jun;3(2):130-139. [doi: 10.1111/j.1752-9824.2011.01085.x]
18. Pal S, Panduragan SL, Said FM, Poddar S. Effect of self-management intervention on improvement of quality of life in chronic kidney disease patients: a scoping review. TONURSJ 2024 Oct 23;18(1). [doi: 10.2174/0118744346323309241009074442]
19. Hermanns N, Ehrmann D, Finke-Groene K, Kulzer B. Trends in diabetes self-management education: where are we coming from and where are we going? A narrative review. Diabet Med 2020 Mar;37(3):436-447. [doi: 10.1111/dme.14256] [Medline: 32017188]
20. Nolte S, Osborne RH. A systematic review of outcomes of chronic disease self-management interventions. Qual Life Res 2013 Sep;22(7):1805-1816. [doi: 10.1007/s11136-012-0302-8]
21. Norris SL, Engelgau MM, Venkat Narayan KM. Effectiveness of self-management training in type 2 diabetes. Diabetes Care 2001 Mar 1;24(3):561-587. [doi: 10.2337/diacare.24.3.561]
22. Vas A, Devi ES, Vidyasagar S, et al. Effectiveness of self-management programmes in diabetes management: a systematic review. Int J Nurs Pract 2017 Oct;23(5). [doi: 10.1111/ijn.12571] [Medline: 28758701]
23. Andrews KL, Jones SC, Mullan J. Asthma self management in adults: a review of current literature. Collegian 2014;21(1):33-41. [doi: 10.1016/j.colegn.2012.12.005] [Medline: 24772988]

XSL•FO
RenderX

24. Jolly K, Majothi S, Sitch A, et al. Self-management of health care behaviors for COPD: a systematic review and meta-analysis. COPD 2016:305. [doi: 10.2147/COPD.S90812]
25. Monninkhof E, van der Valk P, van der Palen J, van Herwaarden C, Partridge MR, Zielhuis G. Self-management education for patients with chronic obstructive pulmonary disease: a systematic review. Thorax 2003 May;58(5):394-398. [doi: 10.1136/thorax.58.5.394] [Medline: 12728158]
26. Fisher EB, Brownson CA, O'Toole ML, Shetty G, Anwuri VV, Glasgow RE. Ecological approaches to self-management: the case of diabetes. Am J Public Health 2005 Sep;95(9):1523-1535. [doi: 10.2105/AJPH.2005.066084] [Medline: 16051929]
27. Kaptein A, Fischer M, Scharloo M. Self-management in patients with COPD: theoretical context, content, outcomes, and integration into clinical care. COPD 2014:907. [doi: 10.2147/COPD.S49622]
28. Kanbour S, Ageeb RA, Malik RA, Abu-Raddad LJ. Impact of bodyweight loss on type 2 diabetes remission: a systematic review and meta-regression analysis of randomised controlled trials. Lancet Diabetes Endocrinol 2025 Apr;13(4):294-306. [doi: 10.1016/S2213-8587(24)00346-2] [Medline: 40023186]
29. Verevkina N, Shi Y, Fuentes-Caceres VA, Scanlon DP. Attrition in chronic disease self-management programs and self-efficacy at enrollment. Health Educ Behav 2014 Dec;41(6):590-598. [doi: 10.1177/1090198114529590] [Medline: 24771128]
30. Disler RT, Gallagher RD, Davidson PM. Factors influencing self-management in chronic obstructive pulmonary disease: an integrative review. Int J Nurs Stud 2012 Feb;49(2):230-242. [doi: 10.1016/j.ijnurstu.2011.11.005] [Medline: 22154095]
31. Koch T, Jenkin P, Kralik D. Chronic illness self-management: locating the "self". J Adv Nurs 2004 Dec;48(5):484-492. [doi: 10.1111/j.1365-2648.2004.03237.x] [Medline: 15533086]
32. Maier B. Dtsch gesundheitsbericht diabetes 2012. : diabetesDE – Deutsche Diabetes-Hilfe; 2012. [Medline: 24360366]
33. Song X, Hallensleben C, Zhang W, et al. Blended self-management interventions to reduce disease burden in patients with chronic obstructive pulmonary disease and asthma: systematic review and meta-analysis. J Med Internet Res 2021 Mar 31;23(3):e24602. [doi: 10.2196/24602] [Medline: 33788700]
34. Schweiger W. Online media. In: Donsbach W, editor. The International Encyclopedia of Communication, 1st edition: Wiley; 2009. [doi: 10.1002/9781405186407]
35. Pius Nedumkallel J. Interactivity of digital media: literature review and future research agenda. Int J Interact Commun Syst Technol 2020;10(1):13-30. [doi: 10.4018/IJICST.2020010102]
36. Brennen JS, Kreiss D. Digitalization. In: Jensen KB, Craig RT, Pooley J, Rothenbuhler EW, editors. The International Encyclopedia of Communication Theory and Philosophy: John Wiley & Sons; 2016:1-11. [doi: 10.1002/9781118766804]
37. Farkaš I. Transforming cognition and human society in the digital age. Biol Theory 2024. [doi: 10.1007/s13752-024-00483-3]
38. Humphreys L, Karnowski V, Pape T. Smartphones as metamedia: a framework for identifying the niches structuring smartphone use. Int J Commun 2018;12:2793-2809 [FREE Full text]
39. Ghose A, Guo X, Li B, Dang Y. Empowering patients using smart mobile health platforms: evidence from a randomized field experiment. arXiv. Preprint posted online on Feb 10, 2021. [doi: 10.48550/ARXIV.2102.05506]
40. Rossmann C, Karnowski V. EHealth und mhealth: gesundheitskommunikation im internet (ehealth and mhealth: health communication on the internet). In: Hurrelmann K, Baumann E, editors. Handbuch Gesundheitskommunikation (Handbook of Health Communication): Verlag Hans Huber; 2014:271-285 URL: https://www.researchgate.net/publication/267325850_eHealth_und_mHealth_Gesundheitskommunikation_im_Internet [accessed 2025-12-24]
41. Farao J, Malila B, Conrad N, Mutsvangwa T, Rangaka MX, Douglas TS. A user-centred design framework for mHealth. In: Schnall R, editor. PLOS ONE 2020;15(8):e0237910. [doi: 10.1371/journal.pone.0237910] [Medline: 32813711]
42. Baumann E, Czerwinski F, Rosset M, Seelig M, Suhr R. Wie informieren sich die Menschen in Deutschland zum Thema Gesundheit? Erkenntnisse aus der ersten Welle von HINTS Germany. Bundesgesundheitsbl 2020 Sep;63(9):1151-1160. [doi: 10.1007/s00103-020-03192-x]
43. Baumann E, Czerwinski F, Großmann U, Calhoun K. Ist Gesundheit schon digitaler Alltag Teilergebnisse der Studie „HINTS Germany" zur Nutzung digitaler Gesundheitsangebote (is health already digital everyday life? Partial results of the study "HINTS Germany). 2021. URL: https://www.stiftung-gesundheitswissen.de/sites/default/files/pdf/2021_07_08_trendmonitor_Digitaliserung_AktFass_vf_0.pdf [accessed 2025-10-29]
44. Eco. Gesundheits-Apps nutzen ein Drittel der Deutschen (31,6 Prozent) (one third of the Germans use health-apps). 2023. URL: https://www.eco.de/presse/gesundheits-apps-nutzen-ein-drittel-der-deutschen-316-prozent/ [accessed 2025-10-29]
45. Daruwalla Z, Thakkar V, Aggarwal M, Kiasatdolatabadi A, Guergachi A, Keshavjee K. Patient empowerment: the role of technology. In: Lau F, editor. Stud Health Technol Inform 2019;257(70-74):70-74. [doi: 10.3233/978-1-61499-951-5-70] [Medline: 30741175]
46. Rossmann C, Krömer N. mHealth in der medizinischen Versorgung, Prävention und Gesundheitsforderung (mHealth in medical care, prevention and health promotion). In: Fischer F, Kramer A, editors. eHealth in Deutschland. Anforderungen und Potenziale innovativer Versorgungsstrukturen (eHealth in Germany Requirements and Potentials of Innovative Care Structures): Springer; 2016:441-456. [doi: 10.1007/978-3-662-49504-9_24]
47. Akeel AU, Mundy D. Re-thinking technology and its growing role in enabling patient empowerment. Health Informatics J 2019 Dec;25(4):1278-1289. [doi: 10.1177/1460458217751013] [Medline: 29366367]

XSL•FO
RenderX

48. Arsad FS, Syed Soffian SS, Megat Kamaruddin PSN, et al. The impact of eHealth applications in healthcare intervention: a systematic review. J Health Res 2022;37(3):178-189. [doi: 10.56808/2586-940X.1020]
49. Woldeyohannes HO, Ngwenyama OK. Factors influencing acceptance and continued use of mhealth apps. In: Nah FFH, Tan CH, editors. HCI in Business, Government and Organizations Interacting with Information Systems: Springer International Publishing; 2017:239-256. [doi: 10.1007/978-3-319-58481-2_19]
50. Agnihothri S, Cui L, Delasay M, Rajan B. The value of mHealth for managing chronic conditions. Health Care Manag Sci 2020 Jun;23(2):185-202. [doi: 10.1007/s10729-018-9458-2] [Medline: 30382448]
51. Varshney U. Mobile health: four emerging themes of research. Decis Support Syst 2014 Oct;66:20-35. [doi: 10.1016/j.dss.2014.06.001]
52. Mata-Lima A, Paquete AR, Serrano-Olmedo JJ. Remote patient monitoring and management in nephrology: a systematic review. Nefrología 2024 Sep;44(5):639-667. [doi: 10.1016/j.nefro.2024.01.005]
53. Ghahramani F, Wang J. Intention to adopt mHealth apps among informal caregivers: cross-sectional study. JMIR Mhealth Uhealth 2021 Mar 17;9(3):e24755. [doi: 10.2196/24755] [Medline: 33729166]
54. Herrmann M, Boehme P, Mondritzki T, Ehlers JP, Kavadias S, Truebel H. Digital Transformation and Disruption of the Health Care Sector: Internet-Based Observational Study. J Med Internet Res 2018 Mar 27;20(3):e104. [doi: 10.2196/jmir.9498] [Medline: 29588274]
55. Boland L, Bennett K, Connolly D. Self-management interventions for cancer survivors: a systematic review. Support Care Cancer 2018 May;26(5):1585-1595. [doi: 10.1007/s00520-017-3999-7] [Medline: 29199362]
56. de Souza Ferreira E, de Oliveira AHM, Dias MA, et al. Mobile solution and chronic diseases: development and implementation of a mobile application and digital platform for collecting, analyzing data, monitoring and managing health care. BMC Health Serv Res 2024 Aug 31;24(1):1009. [doi: 10.1186/s12913-024-11505-y] [Medline: 39217352]
57. Mathur S, Janaudis-Ferreira T, Hemphill J, et al. User-centered design features for digital health applications to support physical activity behaviors in solid organ transplant recipients: a qualitative study. Clin Transplant 2021 Dec;35(12):e14472. [doi: 10.1111/ctr.14472] [Medline: 34510558]
58. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52. [doi: 10.2196/jmir.3951] [Medline: 25803266]
59. Digital health and care. European Commission. URL: https://health.ec.europa.eu/ehealth-digital-health-and-care/digital-health-and-care_en [accessed 2025-10-29]
60. Digitalisation in healthcare. Federal Ministry of Health. 2024. URL: https://www.bundesgesundheitsministerium.de/en/topics/digitalisation/digitalisation-in-healthcare.html [accessed 2025-10-29]
61. Sainz-de-Abajo B, Torre-Díez I, Góngora-Alonso S, López-Coronado M. Privacy issues in ehealth and mhealth apps. In: Tzanou M, editor. Health Data Privacy under the GDPR: Big Data Challenges and Regulatory Responses: Routledge; 2021:71-82. [doi: 10.4324/9780429022241-6]
62. Bork U, Weitz J, Penter V. Apps and mobile health: many potentials still untapped. Deutsches Ärzteblatt. 2018. URL: https://www.aerzteblatt.de/archiv/apps-und-mobile-health-viele-potenziale-noch-nicht-ausgeschoepft-85d2810d-e649-492d-8c77-1aa32106f566
63. Sze KP, Fong QW, De Roza JG, Lee ES, Tan SY. Exploring physicians' perceptions of digital health's impact on the patient-physician relationship in the primary health care setting: qualitative descriptive study. J Med Internet Res 2024 Oct 15;26:e53705. [doi: 10.2196/53705] [Medline: 39405515]
64. Locsin RC, Soriano GP, Juntasopeepun P, Kunaviktikul W, Evangelista LS. Social transformation and social isolation of older adults: digital technologies, nursing, healthcare. Collegian 2021 Oct;28(5):551-558. [doi: 10.1016/j.colegn.2021.01.005] [Medline: 34720647]
65. Kuhn B, Amelung V. Gesundheits-Apps und besondere Herausforderungen (health apps and special challenges). In: Albrecht UV, editor. Chancen Und Risiken von Gesundheits-Apps (CHARISMHA) (Opportunities and Risks of Health Apps [CHARISMHA]) 2016:100-114 URL: http://www.digibib.tu-bs.de/?docid=60009 [accessed 2025-10-29]
66. Badr J, Motulsky A, Denis JL. Digital health technologies and inequalities: a scoping review of potential impacts and policy recommendations. Health Policy 2024 Aug;146:105122. [doi: 10.1016/j.healthpol.2024.105122] [Medline: 38986333]
67. Cucciniello M, Petracca F, Ciani O, Tarricone R. Development features and study characteristics of mobile health apps in the management of chronic conditions: a systematic review of randomised trials. NPJ Digit Med 2021 Oct 5;4(1):144. [doi: 10.1038/s41746-021-00517-1] [Medline: 34611287]
68. Fan K, Zhao Y. Mobile health technology: a novel tool in chronic disease management. Intelligent Medicine 2022 Feb;2(1):41-47. [doi: 10.1016/j.imed.2021.06.003]
69. Sawalha N, Rossmann C. Mobile health for the self-management of chronic diseases: a systematic review of tool characteristics, usage and health-related effects. Presented at: 73rd International Conference of the International Communication Association; May 25-29, 2023 URL: https://www.diselma.de/wp-content/uploads/2026/01/mHealth-Systematic-Review_ICA.pdf

XSL•FO
RenderX

70. Schulte MHJ, Aardoom JJ, Loheide-Niesmann L, Verstraete LLL, Ossebaard HC, Riper H. Effectiveness of eHealth interventions in improving medication adherence for patients with chronic obstructive pulmonary disease or asthma: systematic review. J Med Internet Res 2021 Jul 27;23(7):e29475. [doi: 10.2196/29475] [Medline: 34313593]
71. Whittemore R, Siverly L, Wischik DL, Whitehouse CR. An umbrella review of text message programs for adults with type 2 diabetes. Diabetes Educ 2020 Dec;46(6):514-526. [doi: 10.1177/0145721720965464] [Medline: 33353511]
72. Song T, Yu P, Zhang Z. Design features and health outcomes of mhealth applications for patient self-management of asthma: a systematic review. 2022 Presented at: ACSW 2022: Australasian Comput Science Week 2022; Feb 14-18, 2022; Brisbane, Australia p. 153-160 URL: https://dl.acm.org/doi/proceedings/10.1145/3511616 [accessed 2025-12-24] [doi: 10.1145/3511616.3513110]
73. Yi JY, Kim Y, Cho YM, Kim H. Self-management of chronic conditions using mHealth interventions in Korea: a systematic review. Healthc Inform Res 2018 Jul;24(3):187-197. [doi: 10.4258/hir.2018.24.3.187] [Medline: 30109152]
74. Holmen H, Wahl AK, Cvancarova Småstuen M, Ribu L. Tailored communication within mobile apps for diabetes self-management: a systematic review. J Med Internet Res 2017 Jun 23;19(6):e227. [doi: 10.2196/jmir.7045] [Medline: 28645890]
75. Bashi N, Fatehi F, Fallah M, Walters D, Karunanithi M. Self-management education through mHealth: review of strategies and structures. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e10771. [doi: 10.2196/10771] [Medline: 30341042]
76. Chib A, Lin SH. Theoretical advancements in mHealth: a systematic review of mobile apps. J Health Commun 2018;23(10-11):909-955. [doi: 10.1080/10810730.2018.1544676] [Medline: 30449261]
77. Duplaga M, Tubek A. mHealth - areas of application and the effectiveness of interventions. Zdrowie Publiczne i Zarządzanie 2018;16(3):155-166. [doi: 10.4467/20842627OZ.18.018.10431]
78. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where Is the evidence? PLoS Med 2013;10(2):e1001382. [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]
79. Chib A. The promise and peril of mHealth in developing countries. Mobile Media & Communication 2013 Jan;1(1):69-75. [doi: 10.1177/2050157912459502]
80. Morton K, Dennison L, May C, et al. Using digital interventions for self-management of chronic physical health conditions: a meta-ethnography review of published studies. Patient Educ Couns 2017 Apr;100(4):616-635. [doi: 10.1016/j.pec.2016.10.019] [Medline: 28029572]
81. Dockweiler C, Boketta R, Schenke JH. Nutzungsverhalten und akzeptanz von mhealth-applikationen bei jungen erwachsenen in deutschland (use behavior and acceptance of mhealth applications among young adults in germany). Presented at: TELEMED 2015 - Erfahrungen, Stand und Perspektiven von Gesundheitstelematik und Telemedizin in Deutschland; Jun 23-24, 2015 URL: https://www.researchgate.net/publication/295205869_Use_and_acceptance_of_mHealth_applications_among_young_adults_in_Germany_Nutzungsverhalten_und_Akzeptanz_von_mHealth-Applikationen_bei_jungen_Erwachsenen_in_Deutschland [accessed 2026-01-07]
82. Eaton C, Vallejo N, McDonald X, et al. User engagement with mHealth interventions to promote treatment adherence and self-management in people with chronic health conditions: systematic review. J Med Internet Res 2024 Sep 24;26:e50508. [doi: 10.2196/50508] [Medline: 39316431]
83. Cho J. The impact of post-adoption beliefs on the continued use of health apps. Int J Med Inform 2016 Mar;87:75-83. [doi: 10.1016/j.ijmedinf.2015.12.016]
84. Rossmann C, Riesmeyer C, Brew-Sam N, et al. Appropriation of mobile health for diabetes self-management: lessons from two qualitative studies. JMIR Diabetes 2019 Mar 29;4(1):e10271. [doi: 10.2196/10271] [Medline: 30924786]
85. Brew-Sam N, Chib A, Rossmann C. Differential influences of social support on app use for diabetes self-management - a mixed methods approach. BMC Med Inform Decis Mak 2020 Jul 7;20(1):151. [doi: 10.1186/s12911-020-01173-3] [Medline: 32635919]
86. Whitehead L, Jacob E, Towell A, Abu-Qamar M, Cole-Heath A. The role of the family in supporting the self-management of chronic conditions: a qualitative systematic review. J Clin Nurs 2018 Jan;27(1-2):22-30. [doi: 10.1111/jocn.13775] [Medline: 28231630]
87. Van Dijck J, Poell T. Understanding the promises and premises of online health platforms. Big Data Soc 2016 Jun 1;3(1):2053951716654173. [doi: 10.1177/2053951716654173]
88. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292. [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
89. Rogers EM. Diffusion of Innovations, 5th edition: The Free Press; 2003. URL: https://teddykw2.wordpress.com/wp-content/uploads/2012/07/everett-m-rogers-diffusion-of-innovations.pdf [accessed 2025-12-24]
90. Moran MB, Frank LB, Zhao N, et al. An argument for ecological research and intervention in health communication. J Health Commun 2016;21(2):135-138. [doi: 10.1080/10810730.2015.1128021] [Medline: 26788682]
91. Sallis JF, Owen N. Ecological models of health behavior. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior: Theory, Research, and Practice: Jossey-Bass; 2015:43-65 URL: https://content.e-bookshelf.de/media/reading/L-3717019-50e4fa74c0.pdf [accessed 2025-12-24]

XSL•FO
RenderX

92. Schnauber-Stockmann A, Scharkow M, Karnowski V, Naab TK, Schlütz D, Pressmann P. Distinguishing person-specific from situation-specific variation in media use: a meta-analysis. Communic Res 2025 Mar;52(2):143-172. [doi: 10.1177/00936502241262664]
93. Schnauber-Stockmann A, Bayer JB, Harari GM, Karnowski V. The situation in media and communication research. Commun Theory 2025 Feb 1;35(1):25-36. [doi: 10.1093/ct/qtae021]
94. Schretzlmaier P, Hecker A, Ammenwerth E. Suitability of the unified theory of acceptance and use of technology 2 model for predicting mHealth acceptance using diabetes as an example: qualitative methods triangulation study. JMIR Hum Factors 2022 Mar 9;9(1):e34918. [doi: 10.2196/34918] [Medline: 35262493]
95. Zhang Y, Liu C, Luo S, et al. Factors influencing patients' intentions to use diabetes management apps based on an extended unified theory of acceptance and use of technology model: web-based survey. J Med Internet Res 2019 Aug 13;21(8):e15023. [doi: 10.2196/15023] [Medline: 31411146]
96. Levonian Z, Dow M, Erikson D, et al. Patterns of patient and caregiver mutual support connections in an online health community. arXiv. Preprint posted online on Jul 31, 2020. [doi: 10.48550/ARXIV.2007.16172]
97. Schimank U. Funktionale differenzierung und systemintegration der modernen gesellschaft. In: Differenzierung Und Integration Der Modernen Gesellschaft: VS Verlag für Sozialwissenschaften; 2005:185-197. [doi: 10.1007/978-3-322-80766-3_11]
98. Valente TW. Social networks and health behaviors. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education Theory, Research, and Practice, 5th edition: Jossey-Bass; 2015:205-222 URL: https://content.e-bookshelf.de/media/reading/L-3717019-50e4fa74c0.pdf [accessed 2025-12-24]
99. Rimal RN, Real K. How behaviors are influenced by perceived norms. Communic Res 2005 Jun;32(3):389-414. [doi: 10.1177/0093650205275385]
100. Niederdeppe J, Boyd AD, King AJ, Rimal RN. Strategies for effective public health communication in a complex information environment. Annu Rev Public Health 2025 Apr;46(1):411-431. [doi: 10.1146/annurev-publhealth-071723-120721] [Medline: 39656948]
101. Entman RM, Matthes J, Pellicano L. Nature, sources, and effects of news framing. In: Wahl-Jorgensen K, Hanitzsch T, editors. The Handbook of Journalism Studies: Routledge; 2009:175-190. [doi: 10.4324/9780203877685-22]
102. Foley K, McNaughton D, Ward P. Monitoring the "diabetes epidemic": a framing analysis of United Kingdom print news 1993-2013. PLoS ONE 2020;15(1):e0225794. [doi: 10.1371/journal.pone.0225794] [Medline: 31951616]
103. Reveilhac M, Blanchard A. The framing of health technologies on social media by major actors: prominent health issues and COVID-related public concerns. Int J Inf Manag Data Insights 2022 Apr;2(1):100068. [doi: 10.1016/j.jjimei.2022.100068]
104. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. [doi: 10.1136/bmj.n71]
105. Mobile media use in diabetes self-management. OSF. 2024. URL: https://osf.io/xvmgj/?view_only=2b713ac07037487db959c5fa6c2df01d [accessed 2025-12-07]
106. Review asthma/COPD. OSF. 2025. URL: https://osf.io/dznjs/overview?view_only=b4dc6439d7f64108a9bd498db6f6e1fb [accessed 2025-12-07]
107. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
108. TP1 use. OSF. 2025. URL: https://osf.io/gsfpw/overview?view_only=31a63dccfd8d467fb50dcf365af9b22d [accessed 2025-12-07]
109. Niches structuring mobile media use in chronic disease self-management (ESM asthma/COPD). OSF. 2025. URL: https://osf.io/yg96r/overview?view_only=edbb4a5f5655419591dbb557f7e54604 [accessed 2025-12-07]
110. Park J, Yu HT. Recommendations on the sample sizes for multilevel latent class models. Educ Psychol Meas 2018 Oct;78(5):737-761. [doi: 10.1177/0013164417719111] [Medline: 32655168]
111. Vermunt JK. Latent class and finite mixture models for multilevel data sets. Stat Methods Med Res 2008 Feb;17(1):33-51. [doi: 10.1177/0962280207081238]
112. Vermunt JK, Magidson J. How to perform three-step latent class analysis in the presence of measurement non-invariance or differential item functioning. Struct Equ Modeling 2021 May 4;28(3):356-364. [doi: 10.1080/10705511.2020.1818084]
113. Kuckartz U. Qualitative Inhaltsanalyse Methoden, Praxis, Computerunterstützung [Qualitative Content Analysis Methods, Practice, Computer Support]: Beltz; 2018. URL: https://www.abebooks.com/9783779936824/Qualitative-Inhaltsanalyse-Methoden-Praxis-Computerunterst%C3%BCtzung-3779936828/plp [accessed 2025-12-27]
114. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane; 2024. URL: https://www.cochrane.org/handbook [accessed 2025-12-27]
115. Sawalha N, Lux A, Rossmann C. Mobile health in the self-management of chronic diseases: a systematic meta-analytical review of effects across the mobile media ecosystem. PROSPERO. 2024. URL: https://www.crd.york.ac.uk/PROSPERO/view/CRD42024521002 [accessed 2025-12-07]
116. Assink M, Wibbelink CJM. Fitting three-level meta-analytic models in R: a step-by-step tutorial. TQMP 2016;12(3):154-174. [doi: 10.20982/tqmp.12.3.p154]

XSL•FO
RenderX

117. Viechtbauer W. metafor: meta-analysis package for R. CRAN. 2015. URL: https://cran.r-project.org/package=metafor [accessed 2025-10-29]
118. Harrer M, Cuijpers P, Furukawa T, dmetar EDD. dmetar: companion R package for the guide “Doing Meta-Analysis in R”. 2019. URL: https://dmetar.protectlab.org [accessed 2025-10-29]
119. Hong QN, Pluye P, Fàbregues S, Bartlett G, Boardman F, Cargo M, et al. Mixed methods appraisal tool (MMAT) version 2018. : McGill University, Department of Family Medicine; 2018 URL: http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf [accessed 2025-12-27]
120. Mobile health use patterns and transactional effects in chronic disease self-management: a four-wave investigation. OSF. 2025. URL: https://osf.io/fzu96/overview?view_only=d8ad7769cd0d47a19b808bcfe8de61c6 [accessed 2025-12-07]
121. Linzer DA, Lewis JB. PoLCA: an R package for polytomous variable latent class analysis. J Stat Softw 2011;42(10):1-29. [doi: 10.18637/jss.v042.i10]
122. Visser I, Speekenbrink M. DepmixS4: an R package for hidden Markov models. J Stat Softw 2010;36(7):1-21. [doi: 10.18637/jss.v036.i07]
123. Rosseel Y, Jorgensen T, Wilde L. lavaan: latent variable analysis. CRAN. 2025. URL: https://CRAN.R-project.org/package=lavaan [accessed 2025-12-27]
124. Mulder JD, Hamaker EL. Three extensions of the random intercept cross-lagged panel model. Struct Equ Modeling 2021 Jul 4;28(4):638-648. [doi: 10.1080/10705511.2020.1784738]
125. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
126. Roter DL, Hall JA. Doctors Talking with Patients/Patients Talking with Doctors: Improving Communication in Medical Visits, 2nd edition: Greenwood Publishing Group, Inc; 2006. [doi: 10.5040/9798400641794]
127. Brenk-Franz K, Hunold G, Galassi JP, Tiesler F, Herrmann W, Freund F, et al. Quality of the physician-patient relationship: evaluation of the German version of the Patient Reactions Assessment (PRA-D). Z Allg Med 2016;92(3):103-108. [doi: 10.3238/zfa.2016.0103-0108]
128. Galassi JP, Schanberg R, Ware WB. The patient reactions assessment: a brief measure of the quality of the patient-provider medical relationship. Psychol Assess 1992;4(3):346-351. [doi: 10.1037//1040-3590.4.3.346]
129. Gu Y, Guan Y, Meng Z. Health providers’ perceptions and experiences of using mHealth for chronic noncommunicable diseases: qualitative systematic review and meta-synthesis. J Med Internet Res 2023 Sep 12;25:e45437. [doi: 10.2196/45437] [Medline: 37698902]
130. Posselt J, Klawunn R, Dierks M. Digital health intervention prescriptions for people with depressive disorders: results of a qualitative study. Z Allg Med 2023;99(3):145-150. [doi: 10.1007/s44266-023-00034-2]
131. Digital self-management for chronic diseases – curse or blessing? Decision factors for or against recommendations by physicians. OSF. 2024. URL: https://osf.io/z7kpj/overview?view_only=6de45045e4994935a5571088427f39d7 [accessed 2025-12-07]
132. Riesmeyer C. Das leitfadeninterview. königsweg der qualitativen journalismusforschung? In: Jandura O, Quandt T, Vogelgesang J, editors. Methoden Der Journalismusforschung 2011:223-236. [doi: 10.1007/978-3-531-93131-9_13]
133. Herz A. Erhebung und analyse ego-zentrierter netzwerke. In: Kulin S, Frank K, Fickermann D, Schwippert K, editors. Soziale Netzwerkanalyse: Theorie, Methoden, Praxis: Waxmann; 2012:133-150 URL: https://www.researchgate.net/publication/268521346_Erhebung_und_Analyse_ego-zentrierter_Netzwerke [accessed 2025-12-27]
134. Stark TH, Krosnick JA. GENSI: a new graphical tool to collect ego-centered network data. Soc Networks 2017 Jan;48:36-45. [doi: 10.1016/j.socnet.2016.07.007]
135. Dan V, Raupp J. A systematic review of frames in news reporting of health risks: characteristics, construct consistency vs. name diversity, and the relationship of frames to framing functions. Health Risk Soc 2018 Aug 18;20(5-6):203-226. [doi: 10.1080/13698575.2018.1522422]
136. Roberts ME, Stewart BM, Tingley D. stm: an R package for structural topic models. J Stat Softw 2019;91(2):1-40. [doi: 10.18637/jss.v091.i02]
137. Guenther L, Ruhrmann G, Zaremba MC, Weigelt N. The newsworthiness of the “March for Science” in Germany: comparing news factors in journalistic media and on Twitter. JCOM 2021;20(02):A03. [doi: 10.22323/2.20020203]
138. Kessler SH. Scientific evidence/uncertainty (science and health communication). DOCA 2021. [doi: 10.34778/2h]
139. Lukočienė O, Varriale R, Vermunt JK. 6. The simultaneous decision(s) about the number of lower- and higher-level classes in multilevel latent class analysis. Sociol Methodol 2010 Aug;40(1):247-283. [doi: 10.1111/j.1467-9531.2010.01231.x]
140. Lazarus RS, Stress FS. Appraisal and Coping: Springer; 1984.
141. Wolfers LN, Utz S. Social media use, stress, and coping. Curr Opin Psychol 2022 Jun;45:101305. [doi: 10.1016/j.copsyc.2022.101305] [Medline: 35184027]
142. Wolfers LN, Schneider FM. Using media for coping: a scoping review. Communic Res 2021 Dec;48(8):1210-1234. [doi: 10.1177/0093650220939778]

XSL•FO
RenderX

Abbreviations

CG: concept group
COPD: chronic obstructive pulmonary disease
DMP: disease management program
ESM: experience sampling method
IP: individual project
IRB: Institutional Review Board
mHealth: mobile health
PRISMA : Preferred Reporting Items for Systematic reviews and Meta-Analyses
RU: research unit

Edited by A Schwartz; submitted 21.May.2025; this is a non–peer-reviewed article;accepted 05.Dec.2025; published 28.Jan.2026.

Please cite as:
Rossmann C, Karnowski V, Metag J, Raupp J, Reifegerste D, Riesmeyer C, Sawalha N, Lux A, Esser AL, Kammerer R, Singh F, Rödel N, Brill J, Gerling E, Wiedicke A
The Role of Digital Media in Chronic Disease Self-Management: Protocol for a Multimethod Study of the DISELMA Research Consortium
JMIR Res Protoc 2026;15:e77811
URL: https://www.researchprotocols.org/2026/1/e77811
doi:10.2196/77811

© Constanze Rossmann, Veronika Karnowski, Julia Metag, Juliana Raupp, Doreen Reifegerste, Claudia Riesmeyer, Nariman Sawalha, Alexandra Lux, Anna-Lena Esser, Rebecca Kammerer, Franca Singh, Natalie Rödel, Janine Brill, Elisabeth Gerling, Annemarie Wiedicke. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 28.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Usability and Cultural Relevance of an mHealth App for Hispanic/Latino Individuals Living With Rheumatoid Arthritis: Protocol for a Mixed Methods Study

Thais F Alves[1], PhD; Ronnie Horner[1], PhD; Marie Chantel Montas[2], MPH; Melanie Cozad[1], PhD

[1]Department of Health Services Research and Administration, College of Public Health, University of Nebraska Medical Center, 42nd and Emile, Omaha, NE, United States

[2]Department of Global Health and Population, Harvard T. H. Chan School of Public Health, Boston, MA, United States

Corresponding Author:
Thais F Alves, PhD
Department of Health Services Research and Administration, College of Public Health, University of Nebraska Medical Center, 42nd and Emile, Omaha, NE, United States

Abstract

Background: Hispanic and Latino individuals represent 14.6% of rheumatoid arthritis (RA) cases in the United States and experience significant disparities in access to rheumatologic care, disease management, and health outcomes. Mobile health (mHealth) apps are promising tools to improve patient-provider communication and self-management among populations with language and literacy barriers. However, few RA-focused digital health interventions (DHIs) have been culturally adapted for Spanish-speaking Hispanic and Latino individuals.

Objective: This study aims to assess the health literacy, eHealth literacy, technology trust, and digital self-efficacy of Hispanic and Latino individuals with RA, and to evaluate the cultural relevance, usability, and patient satisfaction of the Spanish-language RunRA app. Additionally, it will explore health care providers' perceptions of the app's usefulness for clinical decision-making and communication with Hispanic and Latino patients.

Methods: A prospective, iterative convergent mixed methods design integrated with the Framework for Reporting Adaptations and Modifications–Expanded (FRAME) will be used. We will recruit 25 Hispanic and Latino patients with RA and 7 Spanish-speaking health care professionals. Quantitative data will include standardized questionnaires (SAHL-S, eHEALS, Human-Computer Trust Scale, Digital Self-Efficacy Scale) and app analytics. Qualitative data will be collected via interviews and focus groups using the Cultural Relevance Questionnaire (CRQ), System Usability Scale (SUS), and Mobile Application Rating Scale (uMARS). Data will be analyzed using an independent intramethod strategy, with integration guided by FRAME to inform culturally relevant app modifications.

Results: We anticipate enrolling 32 participants (25 patients and 7 providers). This study will be the first to evaluate the cultural relevance and usability of an mHealth app specifically designed for Spanish-speaking Hispanic and Latino individuals living with RA.

Conclusions: Our long-term goal is to assess the potential for the mHealth app to act as a vehicle for the dissemination of accurate, useful, usable, and understandable health information to populations that experience health disparities and their health care providers. Findings will inform iterative refinements to RunRA and contribute to the development of culturally responsive DHIs aimed at improving communication, shared decision-making, and health outcomes in underserved populations.

International Registered Report Identifier (IRRID): PRR1-10.2196/88401

(JMIR Res Protoc 2026;15:e88401) doi:10.2196/88401

KEYWORDS

digital health intervention; mobile health; arthritis, rheumatoid; Hispanic or Latino; cultural relevance; usability

Introduction

Approximately 1.5 million people in the United States have rheumatoid arthritis (RA), with racial and ethnic minority groups, women, and elderly people experiencing a disproportionately greater burden. Hispanic and Latino individuals represent 14.6% of RA cases and paradoxically have lower risk of disease-related and treatment-related complications but they experience greater disparities in pain, fatigue, functional limitation, and risk of depression compared to non-Hispanic populations [1]. Disparities in outcomes for Hispanic and Latino individuals are associated with significant delays in obtaining initial rheumatological care related to challenges in health care access [2,3]. Challenges include low income, lower

XSL•FO
RenderX

socioeconomic status, lack of health insurance, and rurality in addition to lower English language proficiency and health literacy levels [4-7]. Low ELP and health literacy for Hispanic and Latino individuals also contribute to lower levels of shared decision-making regarding treatment and symptom management and trust in providers, which may also contribute to worse outcomes [8].

Digital health interventions (DHI) are a highly promising mechanism to improve access to health care for patients with diverse language backgrounds or communication barriers, even for those in rural areas [9-11]. Appropriately designed mHealth apps represent one type of DHI with the potential to improve the patient-provider communication gap among Hispanic and Latino individuals, empowering them to obtain more patient-centric, timely, and effective care [12]. Current studies exploring the potential of culturally relevant DHIs for Hispanic and Latino individuals in the United States predominantly focus on creating behavioral change for the prevention and management of obesity, diabetes, and hypertension [10-17]. To date, only one mobile app for RA has been translated into Spanish for Hispanic and Latino patients, and limited information exists regarding the cultural relevance of RA-related DHIs for this population. To improve goal setting and shared treatment decisions among patients, our team developed 'RunRA,' a mHealth app designed for patients with rheumatoid arthritis. This app contains disease activity tracking, goal setting, and note-taking features that create instantaneous graphics the user can pull up on their smartphone to support shared decision-making with their provider about treatment. RunRA was recently translated into Spanish. Further refinement of this mHealth app fills an important knowledge gap by demonstrating the potential of a culturally relevant RA-related DHI to improve the experience with care of Hispanic and Latino patients with RA.

Effective use of mHealth apps as DHIs to improve share decision-making regarding treatment for RA depends on user's engagement with the app and their willingness to use it as a frequency that provides clinical valuable information on disease activity and goal setting. As an initial step towards app refinement, this study's objective is to assess Hispanic and Latino RA patients' preferences for using DHIs as well as the cultural relevance, usability, and satisfaction the Spanish version of the RunRA app. It also seeks to understand providers' perceptions on the app's use to improve provider communication with Hispanic and Latino patients with RA.

Our specific aims are to:

1. Assess the health literacy, eHealth literacy, technology trust, and digital self-efficacy levels of Hispanic and Latino individuals living with rheumatoid arthritis;
2. Determine the cultural relevance, usability, and patient satisfaction of the mHealth app RunRA, modified for Spanish speakers, among Hispanic and Latino individuals living with rheumatoid arthritis; and
3. Assess health care providers' perspectives on the challenges and potentialities of the RunRA app on improving patient-provider communication, and perceived usefulness of the app's capabilities for clinical decision-making.

The results that emerge from the qualitative and quantitative outcomes' data analysis will allow us to track and plan the adaptations and modifications needed to improve the technology. Our findings will inform refinements to the RunRA mHealth app before its implementation and help support a future study to test its efficacy.

Methods

Overall Study Design

This study will apply a prospective, iterative convergent mixed methods design (ICMM) [18], integrated with the Framework for Reporting Adaptations and Modifications-Expanded (FRAME) [19]. Following ICCM, we will collect qualitative and quantitative data from Hispanic and Latino individuals living with RA and their health care providers. We will conduct an independent intra-method strategy [18] for data analyses to synthesize qualitative and quantitative data collected, and interpret those findings by applying the Integration in the Data Interpretation Dimension [18] to FRAME to determine appropriate, culturally relevant app adaptations and modifications (Figure 1). FRAME comprises eight aspects: (1) when the modification was made; (2) if it was proactive or reactive; (3) who determined that the modification should be made; (4) what is modified; (5) the level of delivery; (6) nature of content or context modifications; (7) fidelity-consistency; and (8) reasons for the modification [18]. This methodological strategy will allow us to (1) understand the app's potential, (2) improve the technology, and (3) prepare for future evaluation of its efficacy.

XSL•FO
RenderX

Figure 1. An iterative convergent mixed-method design for the Spanish RunRA evaluation. FRAME: Framework for Reporting Adaptations and Modifications-Expanded. RA: rheumatoid arthritis.

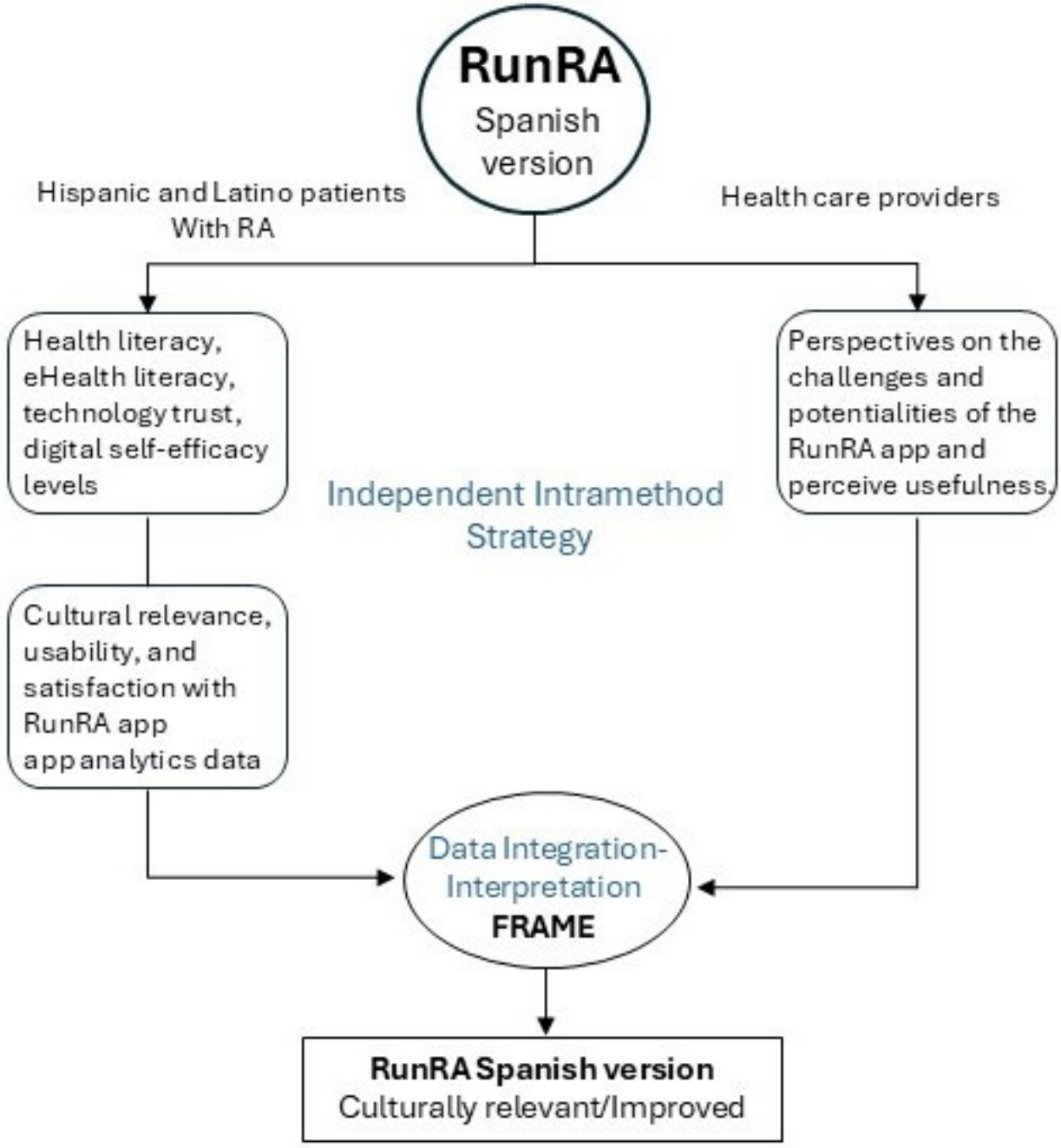

Specific Aim 1 Design

To assess the health literacy, eHealth literacy, technology trust, and digital self-efficacy levels of Hispanic and Latino individuals living with RA, we will conduct a survey applying the Short Assessment of Health Literacy–Spanish (SAHL-S) [20], eHealth Literacy Scale (eHEALS) [21,22], the Human-computer Trust Scale [23], and the Digital Self-Efficacy Scale [24]. We will use these standardized questionnaires in their Spanish versions. This will inform us of the preferences of Hispanic or Latino patients with RA and their readiness to incorporate evidence-based health-related information delivered through technology in their health care routine.

Specific Aim 2 Design

To determine the cultural relevance, usability, and patient satisfaction of the mHealth app RunRA, modified for Spanish speakers, among Hispanic or Latino individuals living with RA, we will use a mixed methods approach to collect: (1) quantitative utilization data through mHealth analytics collected within the app, and (2) qualitative data through personal interviews with participants using the Cultural Relevance Questionnaire (CRQ) [25], System Usability Scale (SUS) [26], the Mobile Application Rating Scale user version (uMARS) [27], and open questions regarding app's content.

Specific Aim 3 Design

For assessing health care providers' perspectives on the challenges and potentialities of the RunRA app on improving

XSL•FO
RenderX

patient-provider communication, and perceived usefulness of the app's capabilities for clinical decision-making, we will conduct a Focus Group (FG) [28] with Spanish-speaking health professionals that assist Hispanic Latino individuals living with RA, at any level of health care (ie, nurses, community health workers, physicians). The FG script will be based on the three main categories of the CRQ: functional equivalence, conceptual equivalence, and linguistic equivalence, and on the eight aspects presented in the FRAME [19].

Participants and Recruitment

We will recruit 25 patients for specific Aims 1 and 2, and 7 health care professionals for specific Aim 3. Patients living with RA, of Hispanic and Latino ethnicity, who are 19 years of age or older, either sex, who are Spanish-literate, and have access to a mobile device in which the RunRA app can be installed (Android or iOS operating systems) will be eligible to participate in the study. Spanish-speaking health professionals who assist Hispanic and Latino individuals living with rheumatoid arthritis in any level of health care are eligible.

To recruit participants for Aims 1 and 2, we will work with the Research, Education, Administration, and Development of Biomedical Informatics (READi) Core, which is the health informatics operations core for a large Medical Center in the Midwest. It provides cutting-edge resources to support public health research and clinical care. According to a READi query requested by our team, there are currently 2854 patients living with RA in Nebraska who opted in to participate in further research through the electronic medical record. This query included patients with a diagnosis of rheumatoid arthritis (codes M05 - M06.9) who were 19 years of age or older, either sex, any race or ethnicity, and without regard for comorbidities. For this study, we anticipate contacting up to 400 participants who are Hispanic and Latino patients with RA. This number was derived based on the national prevalence of disease among this population of 14.6%. The Center for Clinical & Translation Research Opt-In process data pull allows researchers to obtain contact information for Nebraska Medical Center patients who agreed to receive information about research opportunities when they completed their consent for treatment form. Potential participants on the READi query list will receive personalized digital fliers explaining the study and providing information to contact the researcher coordinator with their interest.

If our recruitment efforts fall behind, we will also pursue social media recruitment of participants through the University of Nebraska Medical Center for Clinical & Translation Research social media posts as a secondary strategy. Social media posts will contain information about the study and contact information for the research coordinator.

For Aim 3, we will recruit a convenience sample for the recruitment of health care professionals. When first meeting with the patients, we will ask them about their local health care providers , most often visit to help manage their condition, and based on that information, we will contact the clinics, hospitals, or health departments to invite the professionals to participate in this study.

Data Collection of Outcomes

For Aims 1 and 2, we will collect the data in two ways – through virtual encounters and via participants' use of the app. During our first virtual meeting, participants will answer the following questionnaires: the (SAHL-S) [20], eHEALS [21,22], the Human-computer Trust Scale [23], the Digital Self-Efficacy Scale [24]. Following completion of the survey, we will guide them on how to download and use the Spanish version of the RunRA app (Aim 1). Participants will use the app for 2 to 4 weeks [29], and our team will collect quantitative data related to mHealth app analytics (such as functions accessed, and time spent on the page) (Aim 2). We selected this time frame as it allowed sufficient time to garner appropriate user analytics following other usability studies. We will schedule a second virtual encounter with the same participants for a qualitative interview, using CRQ [25], SUS [26], the uMARS [27], and open questions regarding the app's content.

For Aim 3, we will conduct a focus group with health care professionals who assist these patients to understand their perspectives regarding the potentialities and challenges of using the technology, and their perception of the app's usefulness for clinical decision-making. The script will comprise three main categories of the CRQ, and the eight aspects presented at the FRAME framework (Aim 3).

All virtual encounters and the focus group will be recorded and transcribed using NVivo software for further analysis. A bilingual facilitator will conduct the interviews and focus groups, respecting the participants' preferred language, English or Spanish. Table 1 shows the study outcomes we will explore in Aims 1, 2. and 3, including the source of data, specific measures, participants, and data collection specificity.

XSL•FO
RenderX

Table . Spanish RunRA app evaluation study outcomes.

Outcome	Source of data	Measures	Participants	Outcome data collection
Aim 1: Access the health literacy, eHealth literacy, technology trust, and digital self-efficacy levels of Hispanic/Latino individuals living with RA[a]				
Health literacy	standardized questionnaire	Short Assessment of Health Literacy–Spanish (SAHL-S)	Hispanic and Latino individuals living with RA	During the first meeting for app download
eHealth literacy	Standardized questionnaire	The eHealth Literacy Scale (eHEALS)	Hispanic and Latino individuals living with RA	During the first meeting for app download
Technology trust	Standardized questionnaire	The Human Computer Trust Scale	Hispanic and Latino individuals living with RA	During the first meeting for app download
Digital self-efficacy	Standardized questionnaire	Digital Self-Efficacy Scale	Hispanic and Latino individuals living with RA	During the first meeting for app download
App's ability to help with RA management	Questionnaire	Survey developed by the research team	Hispanic and Latino individuals living with RA	During the first meeting for app download
Aim 2: Determine the cultural relevance, usability, and satisfaction of RunRA				
Identify the app functions used most often and to what extent	Quantitative data on mHealth app analytics	Number of times a page (ie, screen) within the app was visited, amount of time participants remained there, average views on a page per user.	Hispanic and Latino individuals living with RA	As participants use the app after initial download
Usability	Standardized questionnaire	System Usability Scale	Hispanic and Latino individuals living with RA	Within 2 to 4 weeks after app download, during the interview
Satisfaction	Standardized questionnaire	Mobile Application rating Scale user version	Hispanic and Latino individuals living with RA	Within 2 to 4 weeks after app download, during the interview
Cultural relevance, and feedback on app's content	Qualitative, in-person interview	Semi-structured interview script based on the Cultural Relevance Questionnaire (CRQ)	Hispanic and Latino individuals living with RA	Within 2 to 4 weeks after app, during the interview
Aim 3: Identify the challenges and potentialities of the RunRA mHealth App on improving patient-provider communication				
Challenges and potentialities	Qualitative focus group	Focus Group with a script based on CRQ and FRAME	Health care professionals that assist Hispanic and Latino individuals living with RA	Following interactions with patient and app after download
Perceived usefulness of the app's capability for clinical decision-making	Questionnaire	Survey developed by the research team	Health care professionals that assist Hispanic and Latino individuals living with RA	Following interactions with patient and app after download

[a]RA: rheumatoid arthritis.

Data Analysis

For data analysis, we will apply the independent intramethod strategy [18], which implies that the qualitative data is analyzed separately from the quantitative data, followed by the integration in the data interpretation dimension [18] using the FRAME framework [19].

Independent Intramethod Strategy

For the qualitative data outcomes collected during the interviews and the focus group (Aims 2 and 3, respectively), two researchers will independently conduct content analysis [30] in 4 stages – decontextualization, recontextualization, categorization, and compilation – using the software NVivo. To avoid personal bias and to increase validity, the two researchers will use a codebook previously established based on the constructs of CRQ and FRAME, and will meet to garner consensus after finishing the independent coding process. The quantitative outcomes collected through surveys and apps' use (Aim 2) will be analyzed using descriptive statistics such as means, standard deviations, and tabulations. Regarding the standardized questionnaires (Aims 1 and 2), we will follow the developers' recommendations for scoring each instrument.

Integration in the Data Interpretation Dimension

Results that emerged from the qualitative and quantitative outcomes data analysis will be confronted and interpreted based on the eight aspects of FRAME to help us track and plan the adaptations and modifications needed to improve the technology (Textbox 1). In the ICMM strategy, this is the step that precedes the developer updates to the mHealth intervention and, consequently, the development of the app's new version.

XSL•FO
RenderX

Textbox 1. Example of integration in the data interpretation dimension [18] using FRAME [19].

Independent intramethod strategy results

- Quantitative data: SAHL-S scores between 0 and 14 dimension: health literacy level
- Qualitative data: Participants' quotes from interviews and focus group dimension: linguistic equivalence perceptions

Merging

Interpretation: SAHL-S (Short Assessment of Health Literacy–Spanish) score results showed that the target-population has a low literacy level, and the interviews and focus group content analysis indicates that they find app's written information too complicated.

Adaptation and modification tracking report

- RunRA app improvement: Research team will revise app's written information content and adapt grammar until achieve a lower readability level (5th grade or under)
- Adaptation and Modification FRAME (Framework for Reporting Adaptations and Modifications-Expanded) coding:
- When: pilot planned; reactive
- What: content level: target intervention group
- Nature: tailoring/ tweaking/ refining
- Goal: increase engagement and retention; improve feasibility and fit with recipients; address cultural factors
- Reason: literacy and education level
- Fidelity-consistency/core elements: preserved

Ethical Considerations

Data collection for this study will commence after the approval of the University of Nebraska Medical Center Institutional Review Board and the University of Nebraska Medical Center Information Technology security team. The research coordinator will meet with interested participants virtually, explain the study to them, and then have them complete the narrative consent process. Participants will receive a copy of the narrative consent term. We will offer the participants a $50 gift card for each study data collection step, totaling $100 per patient and $50 per health care provider. Regarding data confidentiality, the data elements recorded through the mobile application will be stored on a secure, cloud-based server. The mobile application will pass the deidentifiable data to a secure cloud-based storage system that is accessed only by study leads. All information in the secure-cloud storage system will be deidentifiable as this study requires no names or personal identifiers of the participants be collected. All the qualitative interviews' data will be stored in a deidentified manner with a numerical patient identifier. A list of the participants' names and addresses will be kept only to provide compensation for the study following the qualitative interview. This information will only be collected if the participant wishes to be compensated for the study. This list will not be linked to any of the data collected from the participants and will be kept in a secure shared folder accessible only to the study leads.

Results

We anticipate the enrollment of 32 participants: 25 Hispanic and Latino patients living with RA, and 7 health care professionals. We expect to recruit the first participants by April 2026 and have 100% of planned enrollment recruited by August 2026. The completion of our primary endpoint data analysis is February 2027, and the results will be reported by June 2027.

Discussion

Study Overview

Communication is not just about language and grammar; it is a social phenomenon that involves elements related to a person's life context and history, relationships, how they receive and attribute meaning to information, the cultural relevance of it, literacy levels, and language proficiency. In Latin America, there are approximately ten different Spanish dialects [31], and Hispanic and Latino individuals living in the United States come from these various backgrounds, making this population culturally diverse. Therefore, while developing a DHI for Hispanic and Latino individuals, language translation is essential, but much more is needed to ensure usability, satisfaction, and adoption by these patients. Our long-term goal is to assess the potential for the mHealth app to act as a vehicle for the dissemination of accurate, useful, usable, and understandable health information to populations that experience health disparities and their health care providers. This study will be the first to determine the cultural relevance, usability, and satisfaction of a DHI designed to improve patient-provider communication for the Hispanic and Latino population living with RA. Our team's methodological approach is novel because it integrates an ICMM [20] study design with an implementation science framework—FRAME (Figure 1).

Potential Challenges and Mitigating Strategies

The READi is a health informatics operations core that provides cutting-edge resources to support public health research and clinical care, and one of our strengths for this study recruitment is their ability to directly contact patients who want to participate in research. For our secondary strategy, literature shows that

XSL•FO
RenderX

social media recruitment is a feasible and low-cost option when working with rural communities [32,33]; however, recruiting minority groups, such as Hispanic and Latino individuals, has other challenges, including communication and language, fear, and mistrust [34]. To help mitigate these barriers, all patients' recruitment materials will be available in Spanish, adequate for lower readability levels, and will be revised by bilingual professionals. Also, we understand that sometimes the health care workload can be challenging and impose some difficulties in reuniting all participants simultaneously for a collective discussion. So, if we cannot schedule an FG with the health care professionals, we will conduct personal interviews to address the same content.

Methodological Strengths

We will use the ICMM steps to collect and analyze the data and then apply the Framework for Reporting Adaptations and Modifications-Expanded (FRAME) [19] for integration. Previously, FRAME has been used by researchers in public health program evaluation. It helps researchers plan for intervention scale-up through the understanding of participants' context and improve delivery and effectiveness by tracking modifications and adaptations. To our knowledge, this is the first study to blend these approaches (ie, ICCM and FRAME) for DHI (mobile app) evaluation. The successful use of these methods has the potential to be applied more broadly for all conditions.

Acknowledgments

This project is supported by the College of Public Health Innovation Fund, University of Nebraska Medical Center.

Funding

This study was funded by the UNMC College of Public Health Innovation Fund. The study has been funded for a 1-year period, for $15,000.

Authors' Contributions

TFA, MJC, RDH, were involved in the study design and planning. TFA and MCM were responsible for manuscript drafting and formation, including construction of the tables and figures, with oversight from MJC and RDH. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Peer Review Report 1
Peer review report by the Innovation Fund Committee, College of Public Health, University of Nebraska Medical Center.
[PDF File, 277 KB - resprot_v15i1e88401_app1.pdf]

References

1. Cawley A, Bhalla S, Floresca R, Alvarado L, Dihowm F. Disease profile of rheumatoid arthritis and its complications in Hispanic population. Am J Med Sci 2023 Apr;365(4):337-344. [doi: 10.1016/j.amjms.2022.12.029]
2. Riad M, Dunham DP, Chua JR, et al. Health disparities among Hispanic with rheumatoid arthritis: delay in presentation to rheumatologists contributes to later diagnosis and treatment. J Clin Rheumatol 2020 Oct;26(7):279-284. [doi: 10.1097/RHU.0000000000001085] [Medline: 31220051]
3. Shah A, Macauley C, Ni L, Bay AA, Hackney ME. The relationship between attitudes about research and health literacy among African American and White (non-Hispanic) community dwelling older adults. J Racial Ethn Health Disparities 2022 Feb;9(1):93-102. [doi: 10.1007/s40615-020-00932-2] [Medline: 33415701]
4. Urban and rural differences in community health worker workforce in Nebraska. Nebraska: Department of Health and Human Services. 2021. URL: https://dhhs.ne.gov/MCAH/Urban%20and%20Rural%20Differences%20in%20CHW%20Workforce%20in%20NE%20January%202022%20final.pdf [accessed 2026-01-29]
5. Lahr M, Henning-Smith C, Rahman A, Hernandez AM. Barriers to health care access for rural medicare beneficiaries: recommendations from rural health clinics. 2021 URL: https://rhrc.umn.edu/wp-content/uploads/2025/05/Access-to-Care-Rural-Health-Clinics-PB_revised-2025.pdf [accessed 2026-02-03]
6. The Hispanic population in the United States: 2019. United States Census Bureau. 2021 Oct 28. URL: https://www.census.gov/data/tables/2019/demo/hispanic-origin/2019-cps.html [accessed 2026-01-29]
7. Ainsworth J, Bolic M, Ismail I, Mohammadpour Z, Wood J. Sociodemographic and economic barriers to initial specialist care for patients with rheumatoid arthritis: a scoping review. BMC Rheumatol 2025 Jul 1;9(1):76. [doi: 10.1186/s41927-025-00501-2] [Medline: 40597338]

XSL•FO
RenderX

8. Yip K, Navarro-Millán I. Racial, ethnic, and healthcare disparities in rheumatoid arthritis. Curr Opin Rheumatol 2021 Mar 1;33(2):117-121. [doi: 10.1097/BOR.0000000000000782] [Medline: 33394602]
9. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. 2021. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2026-01-29]
10. Yin Z, Errisuriz VL, Evans M, et al. A digital health intervention for weight management for Latino families living in rural communities: perspectives and lessons learned during development. JMIR Form Res 2020 Aug 20;4(8):e20679. [doi: 10.2196/20679] [Medline: 32726748]
11. Rodriguez JA, Seiglie JA, Caballero AE. What does a new health technology landscape mean for Hispanic/Latino people with diabetes? Clin Diabetes 2023 Apr 1;41(2):289-291. [doi: 10.2337/cd22-0087]
12. Petros De Guex K, Flickinger TE, Mayevsky L, et al. Optimizing usability of a mobile health intervention for Spanish-speaking Latinx people with HIV through user-centered design: a post-implementation study. JAMIA Open 2023 Oct;6(3):ooad083. [doi: 10.1093/jamiaopen/ooad083] [Medline: 37732327]
13. Culhane-Pera KA, Vang KB, Ortega LM, et al. Mobile health technology for hypertension management with Hmong and Latino adults: mixed-methods community-based participatory research. Ethn Health 2023 Apr;28(3):413-430. [doi: 10.1080/13557858.2022.2059451] [Medline: 35387531]
14. Brito FA, Alves TF, Santos N, et al. Feasibility of a culturally adapted technology-delivered, family-based childhood obesity intervention for Latino/Hispanic families in rural Nebraska: the Hispanic Family Connections study protocol. BMJ Open 2024 Oct;14(10):e089186. [doi: 10.1136/bmjopen-2024-089186]
15. Ad G, Dc G, Nj J, et al. A hybrid mobile phone feasibility study focusing on Latino mothers, fathers, and grandmothers to prevent obesity in preschoolers. Matern Child Health J 2023 Sep;27(9):1621-1631. [doi: 10.1007/s10995-023-03700-w] [Medline: 37347374]
16. Rosas LG, Lv N, Lewis MA, et al. A Latino patient-centered, evidence-based approach to diabetes prevention. J Am Board Fam Med 2018;31(3):364-374. [doi: 10.3122/jabfm.2018.03.170280] [Medline: 29743220]
17. Berger Z, Peled Y. Language and health (in)equity in US Latinx communities. AMA J Ethics 2022 Apr 1;24(4):E313-E318. [doi: 10.1001/amajethics.2022.313] [Medline: 35405058]
18. Alwashmi MF, Hawboldt J, Davis E, Fetters MD. The iterative convergent design for mobile health usability testing: mixed methods approach. JMIR Mhealth Uhealth 2019 Apr 26;7(4):e11656. [doi: 10.2196/11656] [Medline: 31025951]
19. Wiltsey Stirman S, Baumann AA, Miller CJ. The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. Implementation Sci 2019 Dec;14(1):58. [doi: 10.1186/s13012-019-0898-y]
20. Lee SYD, Stucky BD, Lee JY, Rozier RG, Bender DE. Short Assessment of Health Literacy-Spanish and English: a comparable test of health literacy for Spanish and English speakers. Health Serv Res 2010 Aug;45(4):1105-1120. [doi: 10.1111/j.1475-6773.2010.01119.x] [Medline: 20500222]
21. Paramio Pérez G, Almagro BJ, Hernando Gómez Á, Aguaded Gómez JI. Validación de la escala eHealth Literacy (eHEALS) en población Universitaria Española. Rev Esp Salud Publica 2015;89(3):329-338. [doi: 10.4321/S1135-57272015000300010]
22. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006;8(4):e27. [doi: 10.2196/jmir.8.4.e27]
23. Gulati S, Sousa S, Lamas D. Design, development and evaluation of a human-computer trust scale. Behav Inf Technol 2019 Oct 3;38(10):1004-1015. [doi: 10.1080/0144929X.2019.1656779]
24. Paredes-Aguirre M, Campoverde Aguirre R, Hernandez-Pozas O, Ayala Y, Barriga Medina H. The Digital Self-Efficacy Scale: adaptation and validation of its Spanish Version. Human Behavior and Emerging Technologies 2024 Jan 5;2024:1-11. [doi: 10.1155/2024/3952946]
25. Salamanca-Sanabria A, Richards D, Timulak L. Adapting an internet-delivered intervention for depression for a Colombian college student population: an illustration of an integrative empirical approach. Internet Interv 2019 Mar;15:76-86. [doi: 10.1016/j.invent.2018.11.005] [Medline: 30740313]
26. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. JMIR Mhealth Uhealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
27. Martin-Payo R, Carrasco-Santos S, Cuesta M, Stoyan S, Gonzalez-Mendez X, Fernandez-Alvarez MDM. Spanish adaptation and validation of the User Version of the Mobile Application Rating Scale (uMARS). J Am Med Inform Assoc 2021 Nov 25;28(12):2681-2686. [doi: 10.1093/jamia/ocab216] [Medline: 34613400]
28. Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995 Jul 29;311(7000):299-302. [doi: 10.1136/bmj.311.7000.299] [Medline: 7633241]
29. Cho H, Yen PY, Dowding D, Merrill JA, Schnall R. A multi-level usability evaluation of mobile health applications: a case study. J Biomed Inform 2018 Oct;86:79-89. [doi: 10.1016/j.jbi.2018.08.012] [Medline: 30145317]
30. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]
31. Lipski JM. Geographical and social varieties of Spanish: an overview. In: The Handbook of Hispanic Linguistics 2012:1-26. [doi: 10.1002/9781118228098]
32. Vos SC, Adatorwovor R, Roberts MK, et al. Community engagement through social media: a promising low-cost strategy for rural recruitment? J Rural Health 2024 Jun;40(3):467-475. [doi: 10.1111/jrh.12809] [Medline: 37985592]

XSL•FO
RenderX

33. Rodríguez MD, Rodríguez J, Davis M. Recruitment of first-generation Latinos in a rural community: the essential nature of personal contact. Fam Process 2006 Mar;45(1):87-100. [doi: 10.1111/j.1545-5300.2006.00082.x] [Medline: 16615255]
34. Rodriguez DK, Hewage SA, Periyakoil VS. Factors affecting the recruitment of Hispanic/Latinx American older adults in clinical trials in the United States: a scoping review. J Am Geriatr Soc 2023 Jun;71(6):1974-1991. [doi: 10.1111/jgs.18264] [Medline: 37013348]

Abbreviations

CRQ: Cultural Relevance Questionnaire
DHI: digital health interventions
eHEALS: eHealth Literacy Scale
FG: Focus Group
FRAME: Framework for Reporting Adaptations and Modifications-Expanded
ICCM: Iterative Convergent mixed-method design
RA: rheumatoid arthritis
READi: Research, Education, Administration, and Development of Biomedical Informatics
SAHL-S: Short Assessment of Health Literacy–Spanish
SUS: System Usability Scale
uMARS: Mobile Application rating Scale user version

Edited by J Sarvestan; submitted 24.Nov.2025; this is a non–peer-reviewed article;accepted 22.Dec.2025; published 13.Feb.2026.

Please cite as:
Alves TF, Horner R, Montas MC, Cozad M
Usability and Cultural Relevance of an mHealth App for Hispanic/Latino Individuals Living With Rheumatoid Arthritis: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e88401
URL: https://www.researchprotocols.org/2026/1/e88401
doi:10.2196/88401

© Thais F Alves, Ronnie Horner, Marie Chantel Montas, Melanie Cozad. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cocreation of a Digital Tool for Proactive End-of-Life Communication: Protocol for a Participatory Action Research Project

Malin Eneslätt[1,2], PhD; Charlèss Dupont[3], PhD; Eva Savolainen[1], PhD; Petter Fjällström[4], PhD; Ida Goliath[5,6], PhD

[1]Department of Health, Education and Technology, Luleå University of Technology, Universitetsvägen 16, Luleå, Norrbotten, Sweden

[2]Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

[3]End-of-Life Care Research Group, Vrije Universiteit Brussel, Brussels, Belgium

[4]Department of Community Medicine and Rehabilitation, Umeå University, Umeå, Sweden

[5]Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Stockholm, Sweden

[6]Stockholm Gerontology Research Center Foundation, Stockholm, Sweden

Corresponding Author:
Malin Eneslätt, PhD
Department of Health, Education and Technology, Luleå University of Technology, Universitetsvägen 16, Luleå, Norrbotten, Sweden

Abstract

Background: Proactive end-of-life (EoL) conversations can help individuals, their significant others, and health care professionals be better prepared to confront dying and future EoL decision-making. Talking about EoL issues may be unfamiliar and difficult; tools are increasingly used to support such conversations. While using digital tools presents many advantages, the development processes of such are seldom robustly reported. The project outlined here has the overall aim of further developing and investigating promotion of early, proactive EoL conversations by cocreating and testing, together with potential community-based end users, a digital version of an existing tool, the DöBra cards.

Objective: The aim of this paper is to outline the protocol for a funded participatory action research project, cocreating a digital tool for proactive EoL conversations, as well as reporting initial steps taken in the project.

Methods: Project design is overall inspired by participatory action research and contains 2 work packages (WPs). WP A encompasses the iterative cocreation process of adapting the initial prototype into a relevant digital tool for the public in Sweden. WP B explores if and how the digital tool can support potential end users and significant others in proactive conversations about future EoL values and preferences. Digital tool development is inspired by a process map and will be conducted in collaborative groups. Data collection for both WPs includes repeated interviews with cocreation partners and meeting documentation. Data from field testing of the tool will encompass audiotaped think-aloud exercises, researchers' observations, usability data, and event log data. Analyses will be qualitative, quantitative, and to some extent mixed methods, and may include inductive thematic analyses, longitudinal qualitative analysis, and descriptive and inferential analyses.

Results: Following funding and project start in January 2024, ethical approval was granted by the Swedish Ethical Review Authority. A stepwise recruitment strategy focusing on heterogeneity yielded an advisory group with 16 members from 13 organizations, representing broad segments of society despite some drop-off. The cocreation process, together with the advisory group, has produced a pilot version of the digital tool, which is currently being field tested. Since May 2025, 42 testers have been recruited. Data analysis is pending. The digital tool will be publicly launched in 2026, readily available, free of charge, and for anyone to use. Results will be disseminated at scientific conferences, in peer-reviewed journals, and through popular science communication.

Conclusions: The cocreative process outlined in this protocol has the potential to develop a digital tool for proactive EoL conversations that is broadly used in the public by varied end users. The digital tool can thus reach new groups in society, potentially highlighting death as a natural part of life and reinforcing normalization of EoL conversations.

International Registered Report Identifier (IRRID): DERR1-10.2196/88452

(JMIR Res Protoc 2026;15:e88452) doi:10.2196/88452

KEYWORDS

advance care planning; digital technology; participatory methods; codevelopment; Go Wish; DöBra

XSL•FO
RenderX

Introduction

Digital tools are increasingly being developed as they may have advantages over analog tools. However, digitalization also brings forth hinders and challenges in usability and availability to the public, especially in people with cognitive hinders. Furthermore, development processes of web-based social health care tools are seldom being fully reported. This protocol paper, therefore, brings a positive addition to robust reporting of digitalization processes.

Proactive reflection and discussion of end-of-life (EoL) values and preferences before death is imminent, a process referred to as advance care planning (ACP) [1,2], can help individuals, their significant others (SOs), and health care professionals be better prepared to confront dying and EoL decision-making when that time comes [3]. While such EoL communication has traditionally been seen as the realm of professionals, which is often still the case in Sweden [4-11], there is increasing international recognition of a need for early, proactive, community-based ACP processes [12-16], not least in the wake of the COVID-19 pandemic [17-19].

The body of knowledge on benefits of ACP practice is vast. Systematic reviews [20-22] show that ACP leads to increased quality of life, health care satisfaction, and use of palliative care. ACP has been associated with reduction in futile life-support treatment at the EoL [20,23] and decreased undesirable hospital admissions [20-23]. International studies [24-27] show that large parts of populations, approximately 60% to 90%, are willing to participate in ACP conversations.

Document-based ACP practices have proven controversial [28-30]. Instead, early and proactive ACP approaches—focusing on preparation of individuals, families, and professionals to be able to make the best possible in-the-moment decisions when they later become necessary—are gaining interest [3,13,21,31]. This focus matches the approach to ACP taken in this project, focusing on reflection and conversations on broad EoL issues in the public.

As talking about EoL issues and dying may be unfamiliar and difficult, tools are increasingly used to support such conversations [32,33]. These tools generally aim to facilitate engagement in ACP discussions by supporting individuals and their SOs to reflect on and/or make decisions for future care and treatment with health care professionals [34]. Initiatives to support community-based, non-professionally led ACP conversations on broader EoL topics are less common, but increasingly being developed [14,35,36].

The Swedish DöBra card is a tool to support EoL conversation that was translated and adapted from the US Go Wish card game, developed by palliative care physician Elizabeth Menkin and the nonprofit organization Coda Alliance [37]. Each card contains a statement, based on seminal research [38] on factors considered important at EoL by patients, family, and caregivers. The process of reflecting on, prioritizing, and ranking card statements serves as the basis for a focused discussion on future EoL care goals [39].

Translating and adapting the Go Wish into the DöBra cards was initiated in 2012, in response to the lack of proactive EoL conversations in the Swedish context at the time [40]. This involved an extensive, cocreative process within a project group of community partners. The DöBra cards have been well-received when tested in various groups in Sweden, that is, community-dwelling older adults [40-42]; staff, residents, and family members in residential elder care [43-45]; and among the Indigenous Sámi [46]. The DöBra cards have been publicly available for purchase since 2018 (without any revenues for researchers); our research indicates an appreciated conversation tool that has had an impact in the public [47].

While the original US Go Wish cards are available online on a basic webpage, scientific studies on the digitalization of the Go Wish cards remain scarce. A recent publication reported on the digitalization of a Go Wish-inspired ACP tool [48]. This project takes inspiration from previous work in Belgium, where a digital version of the Flemish Go Wish cards was developed in a larger project that cocreated a dementia-friendly ACP website [49-53].

Web-based tools can have advantages over paper-based or physical tools as they can be accessed online at any time and place, be used at one's own pace, alone or together with professional or family caregiver over geographical distances, and have potential to reach a wider and more diverse population [54]. Furthermore, web-based tools can include interactive features that allow tailoring to individual needs and preferences [55]. However, limitations highlighted in a systematic review [14] include web-based ACP tools rarely being developed in collaboration with potential end users, that content was not always evidence-based, and tool development was seldom scientifically evaluated. Furthermore, the level of detail in reporting on the development processes varies, and sometimes a clear description of the digital transition is even left out (Dupont et al, unpublished data, February 2026). In this project, the digital ACP tool will be a further adaptation of the robustly researched Swedish DöBra cards [40-47,56], based on the internationally well-studied Go Wish cards [37,39,57-63], to be developed in a cocreative process and evaluated in manners that allow comparison with previous studies [48,49,64].

In this paper, we outline the protocol for a funded participatory action research (PAR) project, as well as report on initial steps taken in the project. The overall aim of the project is to further develop and investigate the promotion of early, proactive EoL conversations by cocreating and testing, together with potential community-based end users, a digital version of an existing ACP conversation tool, the DöBra cards.

Methods

Study Design

Overall project design is inspired by PAR, involving (1) a dynamic and cyclic process of problem identification, planning, action, and evaluation which incrementally changes and builds further on lessons learned, based on (2) partnerships between researchers and community-based stakeholders [65].

This project consists of 2 work packages (WPs). WP A focuses on the iterative cocreation process of adapting the initial

XSL•FO
RenderX

prototype into a relevant digital tool for the public in Sweden. WP B explores if and how the digital tool can support potential end users and SOs in proactive conversations about future EoL values and preferences. Figure 1 presents an outline of the tool development process and WPs.

Figure 1. Timeline of tool development and work packages (WPs).

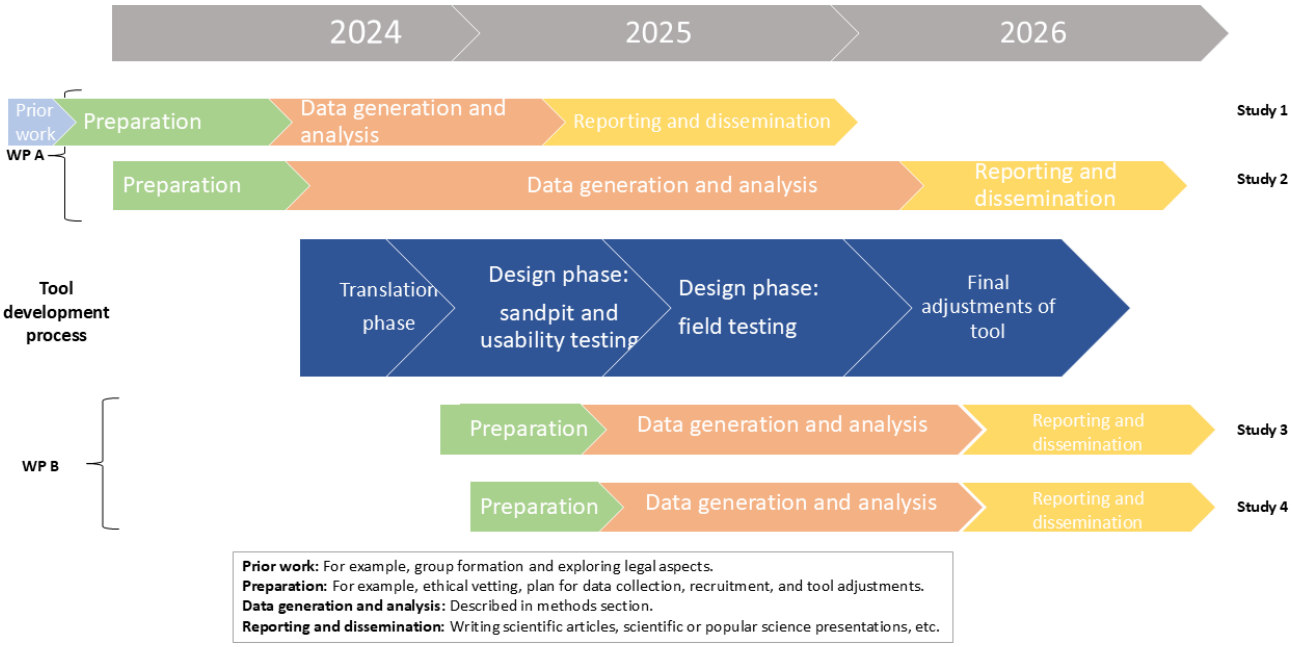

Tool Development Process

The process by Elwyn et al [66] for development of web-based decision support interventions inspires the adaptation process with modifications made to enhance our focus on PAR and cocreation. Four intertwined collaborative groups [66] facilitate the process (Textbox 1):

Textbox 1. Overview of the collaborative groups.

Project management group: A core researcher group, involved in all project phases, responsible for general project management and setting up other groups. The project management group will process feedback from the other groups and make final decisions on content, design, and testing.

Advisory group: Various community partners representing potential end users of the tool, recruited via existing and new community contacts, with diverse ages, backgrounds, experience, and competencies. The advisory group gives repeated feedback on the tool's content, design, usability, and accessibility, primarily in the early design phase.

Scientific reference group: Diverse researchers, with expertise from, for example, health sciences, palliative care, health informatics, computer sciences, design, ethics; the scientific reference group will both meet with the advisory group for feedback on adaptations and inform and contribute to research studies.

Technical production group: This group will design and build a web-based tool based on input coordinated and communicated by the project management group. The project management group will have a contractual agreement with this group, directly supervising the technical production group.

Based on Belgian experiences [49] and the process map [66], tool development is in phases (Textbox 2). Given the dynamic and cyclic PAR approach, the following is a guiding outline—open for adaptation.

Textbox 2. Overview of tool development phases.

Storyboard step: All collaborative groups meet for discussions of storyboard version, feedback looped through technical production group and project management group who deliver the next version.

Sandpit step: In this phase of exploration and testing new ideas, the advisory group provides ideas and feedback to be looped through the technical production group and project management group before sandpit-testing with approximately 10 potential end users.

Usability step: Feedback from sandpit testing is addressed by the project management group and technical production group before discussing the usability version with the advisory group and scientific reference group, resulting in a version for usability testing with approximately 10 new individuals.

Field testing: When the prototype is stable with major issues of navigation and structure processed, field testing commences. The prototype will be tested in a heterogeneous sample of approximately 50 individuals, including approximately 30 dyads of target persons and their significant others who will test the digital tool on their own for approximately 6 wk.

Final tool: Following adaptations and discussions in all groups after field testing, a final digital DöBra card tool will be launched.

XSL•FO
RenderX

Recruitment and Participants

In line with the tool development process, recruitment will be conducted in phases. Initially, members of the advisory group were recruited from various community settings, with an emphasis on heterogeneity in experience, background, age, and gender. This recruitment phase is described in detail in the results section. Forming a multiperspective advisory group was important to facilitate cocreation of a user-friendly digital tool, relevant for the broad public in Sweden today. Potential advisory group members were informed of the prolonged engagement as well as the data collection involved with participating in the group.

Subsequent recruitment of participants for the different stages of digital tool testing will be conducted together with the community partners engaged in the advisory group. Given the potentially sensitive nature of reflecting on future EoL preferences, recruitment will be based on active volunteering to minimize risk of perceived intrusion. Researchers will supply oral or written information about the study in various fora, and potential participants will make the active choice of contacting researchers to sign up for the study. Inclusion criteria for testing are willingness and ability to engage in reflection on EoL preferences by trying the digital tool for approximately 6 weeks; having access to and ability to use any kind of screen, for example, smartphone, tablet, or computer; and self-assessed language proficiency in Swedish on a level efficient for use of the tool and engagement in interviews. Individuals deemed by researchers unable to independently supply an informed consent will be excluded from data collection but may test the tool outside of the research study if they wish.

Recruitment for testing will focus on heterogeneity in participants' age, gender, digital skills, EoL experiences, and educational, cultural, and language backgrounds. While no minimum quotas for subgroups have been decided a priori, in order to maximize the potential to reach different groups, recruitment will be conducted stepwise, mainly through organizations in the advisory group. The stepwise procedure is also helpful in steering recruitment toward characteristics that are not externally visible but can be observed in data collection activities, for example, digital skills and language background.

Data Collection

Work Package A

WP A aims to explore the cocreation and adaptation process of transitioning the paper-based DöBra cards into a digital tool in 2 initial substudies (Table 1). Data will consist of meeting documentations in all phases of cocreation and adaptation, including meeting notes and participant observations of group interaction. All adaptations, including rejected suggestions (from group members and testers), will be carefully documented.

Table . Overview of substudies, data collection, and analyses.

Substudy: tentative focus	Data sources	Tentative data analysis
1: Exploration of the cocreation process	• Meeting documentation • Interviews with advisory, project management, and technical production group members	• Longitudinal qualitative analysis
2: Description of the adaptations made when transferring from a paper-based to a digital tool	• Meeting documentation from cocreation process • Testing: Think-aloud exercises, observations, interviews, System Usability Scale	• Thematic analysis • Descriptive statistics
3: Exploration of test person's opinions and experiences on if/how the digital tool can facilitate reflection on EoL[a] preferences	• Testing: interviews, event log data, System Usability Scale	• Thematic analysis • Descriptive statistics, inferential statistical analyses
4: Longitudinal exploration of alignment of test persons' and significant others' rankings of EoL preferences	• Testing: interviews, rankings of EoL preferences at both time points	• Interpretive description • Observational, longitudinal comparison of rankings

[a]EoL: end-of-life.

Advisory, project management, and technical production group members will be interviewed individually and in groups at multiple time points during the process to elicit their experiences and views on the cocreative work. Interviews will probe into, for example, members' reflections on the current state of the process and pressing issues, collaboration in the group, personal learnings, potential conflicts, and problem-solving.

Data from testing will also be used to describe the adaptation process, for example, audiotaped think-aloud exercises and researchers' observations. Participants will navigate through each prototype version using a think-aloud method [67], for example, verbalizing their thoughts, impressions, and feelings while engaging with the tool. A researcher will observe and note, for example, participants' physical cues, task management, difficulties, and comments with minimal interference [67].

Usability data generated using the System Usability Scale (SUS) [68,69] will also be included. The SUS is a widely used, simple, reliable, and validated 10-item scale that measures subjective usability in terms of effectiveness, efficiency, and satisfaction.

During a 6-week field testing period, testers' usage of the digital tool will be analyzed by event log data. Event logging involves storing 3 variables: (1) unique user ID (who interacted), (2) date

XSL•FO
RenderX

and time stamp (when the interaction occurred), and (3) an event name (describing the interaction) [70].

Work Package B

WP B aims to explore whether and how the digital tool can support potential end users and SOs in ACP conversations about future EoL values and preferences. To this end, interviews with testers will be conducted before the extended field-testing period and after 6 weeks of testing. In addition to sharing their views of how the digital tool technically worked during testing, testers will also be asked to reflect on the tool's potential to support reflection and conversations about EoL preferences as well as EoL decision-making.

Field testers can, if preferred, test the digital tool together with a significant other. In those cases, we will further explore how the tool might facilitate their communication on the subject and help align target persons' (TPs) and SOs perceptions of TPs' EoL preferences. Data consist of TPs and SOs rankings of TPs' EoL values and preferences in the digital tool at the first and the follow-up interview, as well as their recorded reflections about their choices. Based on our previous findings [41,42], congruence between TPs and SOs will be assessed by placement in top 10 DöBra card rankings in the digital tool and reasoning about preferences. Both interviews explore participants' perceptions of how the use of the tool and conversations support preparedness for EoL decision-making.

Event log data may be used to study TPs and SOs patterns of tool use during the trial period to explore, for example, if frequent use of the tool leads to more agreement between TPs and SOs (perception of TPs) EoL preferences.

Data Analysis

Given the iterative and dynamic PAR process used, interactive forms of data collection, as well as the multitude of data sources, data analyses described here are preliminary and will therefore be described briefly.

Analyses of audiotaped and transcribed interview data and think-aloud exercises, field notes, and participant observations may apply inductive thematic analyses [71], interpretive description [72], and longitudinal qualitative analysis [73,74].

Usability as measured by SUS will be analyzed with descriptive statistics and inferential analyses, in combination with above-mentioned qualitative analyses. While a mean score of greater than 68 is a commonly used benchmark to evaluate usability by SUS [75], these measures will be complemented also by qualitative data to assess usability. Combining multiple data sources in evaluating usability is supported by a recent systematic review [76].

In WP B, event log data [70] will be used to analyze frequency and longevity in participants' use of the digital tool, to inform mixed methods analysis on participants' experiences of using the digital tool. By combining event log data and qualitative interview data, we can analyze possible links between stated experiences of using the tool with logged frequency and longevity of tool usage, for example, to see if patterns of usage are related to experience. Mixed methods analysis also includes longitudinal analysis of DöBra card rankings and clustering of event log data, with qualitative data on reasoning about EoL preferences and sense of preparedness in the fourth substudy (Table 1).

In general, analyses will be performed by the first author of the study in close collaboration with the last author but always discussed with the full author team throughout the analysis processes. Reliability of data analyses will be strengthened by triangulation of data sources, methods, and investigators [77], as data from various sources, analyses procedures, and researchers' competences will be combined in the analyses.

Ethical Considerations

This project and its WPs have been ethically vetted and approved by the Swedish Ethical Review Authority (2024-03538-01). This project adheres to the Declaration of Helsinki as it involves human research and written informed consent has been obtained from all participants. Furthermore, participants were informed that their privacy and confidentiality are protected by deidentification (pseudonymization) of their data. They were also informed that their data will be stored in safe servers in accordance with Luleå University of Technology's policy. Participants received no monetary compensation.

Although we are not primarily conducting research with people who are themselves nearing EoL in this project, the substantial literature on and experiences of conducting EoL research suggest that people are often positive to participation, even when nearing the EoL [78]. We are committed to individuals being able to make their own decisions about care and research participation as long as they are able [79]. In order to not be perceived as confrontational for individuals who are not comfortable with the topic of EoL, recruitment will be based on active volunteering.

Furthermore, we are aware of the need for ethical reflection in collaboration with community partners. Discussions and considerations of ethical issues will therefore be integrated intrinsically throughout the project. Community partners may have varied considerations regarding ethical behaviors and issues and may not always agree. Negotiating issues of power, confidentiality, and credibility also demand attention in PAR research, requiring proactive and collective reflection [80,81].

Results

Following major grant funding, the project commenced in January 2024 and has overall been carried out according to the protocol, even though some amendments have been made given the participatory nature. In this section, we will report on the initial steps conducted in the project.

Recruitment of community partners involved in cocreation of the tool in the advisory group began in early 2024. A step-wise recruitment strategy was used to avoid recruiting "more of the same." This strategy resulted in a heterogeneous group (see Figure 2 for an overview of participating organizations). Community partners involved in the previous participatory process translating and adapting the paper-version DöBra cards were first approached, that is, the Swedish Dementia Association

XSL•FO
RenderX

and the national retiree organization Pensionärernas riksorganisation; however, involving new representatives from these organizations as the geographical base of the project shifted from the capital area to the northernmost part of the country and as leadership positions had shifted within the organizations. Given that the geographical base of the project constitutes the realm of 2 national minority groups, the Sàmi and the Tornedalians, these groups were approached via their representative organs, that is, the Sàmi parliament and the National Association of Swedish Tornedalians, Tornionlaaksolaiset. In order to gain health care professionals' input on the project, a registered nurse specialized in dementia care and a nursing aide specialized in education of nursing staff were recruited from nearby municipalities. Striving for heterogeneity in professional input led to recruitment of a death doula, a person offering nonmedical services to dying individuals and their SOs in the final stages of life. Given the ambition to develop a digital tool which would be user-friendly in broad segments of the Swedish public, contact was made with a number of organizations, namely, the Swedish Disability Rights Federation, an organization committed to everybody's right to self-determination and full participation in society. This is an umbrella organization that wanted to participate with the organizational coordinator as well as 2 representatives from patient organizations: the Swedish Heart and Lung Association and the Swedish Association of the Visually Impaired. These local representatives, in turn, made us aware of the support for older people office in the municipality which works, among other things, with supporting older individuals in managing and using everyday digital platforms. Also, to have perspectives of younger people affected by serious illness, the Young Cancer organization, which works to improve living conditions for young adults living with cancer, was asked to participate. In order to gain perspectives of people not born in Sweden and/or not having Swedish as their first language, the local International Women's Association Esperanza was approached and asked to participate. Again, with a focus on inclusion and broad perspectives on EoL, a local lesbian, gay, bisexual, transgender/transsexual, queer, and other minority sexual orientations and gender identities organization was approached. Finally, the publishing company selling the paper-version DöBra cards and thus having important insights into how the paper-version cards are being distributed was asked for their interest to participate. All approached parties and individuals initially agreed to participate in the project.

Figure 2. Community partners participating in cocreation in the advisory group. PRO: Pensionärernas Riksorganisation.

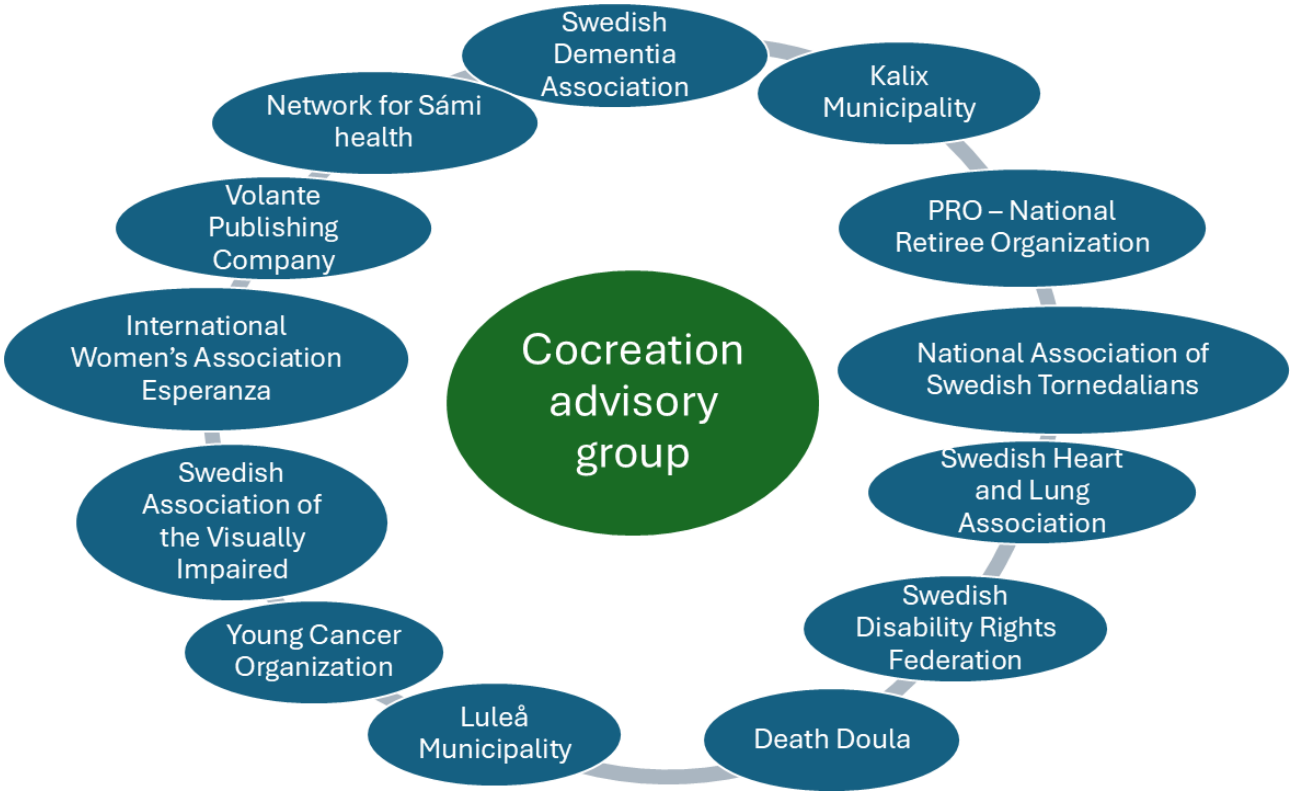

However, during the startup phase, the representative of the lesbian, gay, bisexual, transgender/transsexual, queer, and other minority sexual orientations and gender identities organization and the nursing aide left the advisory group for different personal or organizational reasons. Furthermore, the representatives of the Sàmi Parliament attended the first 2 meetings, but in the second meeting, they communicated a strong wish for the DöBra card statements to be translated and culturally adapted to Sámi settings. This was discussed as a deviating objective from the main project; therefore, the project leader communicated that she envisioned this as a side project that could seek additional funding but not be part of the funded main project. Following this, the Sàmi Parliament representatives left the advisory group. After discussions in subsequent advisory group meetings, the group jointly decided that participation of Sámi representatives was important, and the project leader was given a mandate to try to find other representatives. *The Network for Sámi Health*, an organization

XSL•FO
RenderX

focused on promoting health and equitable care for the Sámi people, was therefore approached and agreed to participate.

Some organizations had double representatives in order to cover for each other, making the total number of advisory group members 16: 13 women and 3 men. Not all members have joined every meeting, which has been held in hybrid form after the first physical meeting in Luleå in April 2024. A deviation from the original plan is that the scientific reference group has been less involved in the cocreation of the digital tool. This is both due to the considerable size of the advisory group which already brings very varied perspectives as well as logistical difficulties in scheduling meetings. Instead, the project leader has met scientific reference group members individually on demand, as well as circulated updates on the project and tool development progress approximately twice a year.

Field testing of the digital tool was initiated in May 2025, and at the submission date of this paper, 42 individuals had started testing. While data from the cocreation process is currently under analysis, data from field testings remains to be analyzed, scientific results publications are expected during 2026-27, and public tool dissemination is planned during fall 2026. From the discussions in the advisory group, a vision of the final digital tool being readily available, free of charge and for anyone to use, has crystalized. Discussions of which platform to use for dissemination are still ongoing, with a focus on already existing websites or platforms that can offer a stable host for the tool and managing necessary updates. The long-term governance of the tool is another important question that is under discussion, in order to facilitate broad public dissemination and ensure that the tool will be readily available.

Dissemination of research results will be conducted at international and national scientific conferences and by publishing peer-reviewed scientific articles in journals. Given that this is a PAR project, much attention will also be paid to popular-science communication, for example, public presentations and (written/oral) communication in the community organizations' preferred channels.

Discussion

Initial Findings

This paper outlines the protocol of a PAR project while also reporting initial steps taken in the project. The project has thus far been successful in recruiting a heterogeneous advisory group, representing various perspectives of importance for EoL care and decision-making. The importance of engaging potential end users in designing digital tools is increasingly recognized [76]. Working with PAR stimulates change processes in community partners, while generating new knowledge researching the process. In line with previous findings [43,44,47,82], this project is thus likely to have impact during the cocreative development process and not only at its conclusion, which is 1 motivation for the process-oriented design in WP A. When researching and developing the paper-version DöBra cards, several ripple effects appeared, for example, a strong demand for the cards, leading to them being made publicly available for purchase soon after initial testing and organizational engagement to inform about EoL and APC within their community [47]. Spin-off effects are hoped for also from this project, which is why collaborating organizations have been strategically chosen to represent a broad range of potential end users, facilitating impact in varied community settings. Furthermore, the disseminated digital tool itself has the potential to reach new groups and even broader societal segments. Dissemination and use in varied community settings can highlight death as a natural part of life and reinforce normalization of EoL conversations by reframing ACP as a health-promoting activity involving public education and engagement [49,83].

This project feeds into both active community collaboration in research and making individuals and their SOs active partners in discussing future EoL care, thus potentially strengthening alignment of future EoL care with individuals' preferences. In Sweden, a major reform of the national health care system is ongoing, aiming at reorganizing and delivering health care close to home based on patients' needs and conditions. This transition includes designing and managing health care services through discussions with patients to strengthen patient participation [84]. Furthermore, the Swedish National Board of Health and Welfare suggests that eHealth solutions can be used to facilitate and encourage patient participation in care and treatment [85]. The broad DöBra statements, covering physical, social, existential, and practical concerns, may facilitate reflection in individuals and their SOs about support needed from formal caregivers, and which needs can be addressed within their social and family circles. Furthermore, the approaches in this project align with the focus of the new Swedish Social Services Act (SFS 2025:400) on preventive measures and proactiveness. Early and proactive engagement in ACP, as addressed in this project, is one means for individuals' needs to be known to others, and for individuals to be prepared to act as discussion partners when meeting health care providers.

From our previous research with the paper-version DöBra cards, we conclude that they have worked well to elicit individuals' future EoL values and preferences and stimulate ACP conversations in a range of contexts [40-47,56]. However, 1 issue that community-based ACP initiatives have yet to address is how to document and communicate the preferences discussed in nonprofessionally led ACP conversations. The digital DöBra tool may be a key in both documenting and saving information derived from these conversations for individuals to keep, as well as digitally communicating one's preferences to someone else, even bridging geographical distance. This project, therefore, also contributes critical new knowledge through addressing the option for individuals to document and communicate EoL preferences resulting from an ACP conversation using the digital DöBra tool.

Strengths and Limitations

This project responds to limited transparent reporting on digital adaptations. Our detailed documentation of prototype changes, think-aloud testing, and continuous input from a heterogeneous advisory group directly addresses this gap. Moreover, from previous research, we know that challenges commonly arise during such transitions, ranging from difficulties in navigation to issues linked to digital literacy [86-89]. Since these challenges

XSL•FO
RenderX

are rarely detailed in published work, transparent reporting becomes crucial for ensuring that others can understand what was done, what worked, and what required adjustment. Transparent reporting is thus essential so that future researchers and developers can build on this work and further strengthen the development of accessible and effective web-based tools.

Conclusions

The cocreative process outlined in this protocol has potential to develop a digital tool for proactive EoL conversations that is broadly used in the public by varied end users. The project's focus on inclusion and broad perspectives on EoL issues has strengthened the cocreative process but also brought challenges in managing expectations and demands. The digital tool can reach new groups in society, potentially highlighting death as a natural part of life and reinforcing normalization of EoL conversations.

Acknowledgments

The authors wish to acknowledge Professor Gert Helgesson, Dr Max Kleijberg, Professor Sabine Koch, Professor Jessica Körning-Ljungberg, Professor Carol Tishelman, and Professor Lieve Van den Block for their valuable contributions to the grant application and continued engagement in the scientific reference group. No parts of this manuscript were generated through the use of artificial intelligence.

Funding

This project was funded by the Swedish Research Council for Health, Working life and Welfare (dnr: 2023-01235).

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

Conceptualization: ME (lead), CD, IG (supporting).
Funding acquisition: ME (lead), CD, IG (supporting).
Methodology: ME (lead), IG (supporting).
Project administration: ME.
Supervision: ME (lead), IG (supporting).
Validation and visualization: ME (lead), CD, ES, PF (supporting).
Writing – original draft: ME.
Writing – review & editing: CD, ES, PF, IG (supporting).

Conflicts of Interest

None declared.

Peer Review Report 1 [PDF File, 47 KB - resprot_v15i1e88452_app1.pdf]

References

1. Rietjens JAC, Sudore RL, Connolly M, et al. Definition and recommendations for advance care planning: an international consensus supported by the European Association for Palliative Care. Lancet Oncol 2017 Sep;18(9):e543-e551. [doi: 10.1016/S1470-2045(17)30582-X] [Medline: 28884703]
2. The society. ACP-i - Advanced Care Planning international. URL: https://www.acp-i.org/society/ [accessed 2026-01-28]
3. Sudore RL, Fried TR. Redefining the "planning" in advance care planning: preparing for end-of-life decision making. Ann Intern Med 2010 Aug 17;153(4):256-261. [doi: 10.7326/0003-4819-153-4-201008170-00008] [Medline: 20713793]
4. Tranberg M, Jacobsen J, Fürst CJ, Engellau J, Schelin MEC. Patterns of communication about serious illness in the years, months, and days before death. Palliat Med Rep 2022;3(1):116-122. [doi: 10.1089/pmr.2022.0024] [Medline: 36059906]
5. Kastbom L, Falk M, Karlsson M, Tengblad A, Milberg A. Prevalence, content and significance of advance care planning in nursing home patients. Scand J Prim Health Care 2022 Mar;40(1):29-38. [doi: 10.1080/02813432.2022.2036429] [Medline: 35170393]
6. Kastbom L, Karlsson M, Falk M, Milberg A. Elephant in the room: family members´ perspectives on advance care planning. Scand J Prim Health Care 2020 Dec;38(4):421-429. [doi: 10.1080/02813432.2020.1842966] [Medline: 33174807]
7. Alftberg Å, Ahlström G, Nilsen P, et al. Conversations about death and dying with older people: an ethnographic study in nursing homes. Healthcare (Basel) 2018 Jun 14;6(2):63. [doi: 10.3390/healthcare6020063] [Medline: 29899220]
8. Andersson S, Granat L, Baxter R, et al. Translation, adaptation, and validation of the Swedish serious illness conversation guide. J Palliat Care 2024 Jan;39(1):21-28. [doi: 10.1177/08258597231210136] [Medline: 37886797]

XSL•FO
RenderX

9. Myrhøj CB, Bernacki R, Bjerre-Bertelsen S, et al. Serious illness conversations and quality of end-of-life care in patients with hematological malignancies: a retrospective quality improvement study. Support Care Cancer 2025 Aug 27;33(9):814. [doi: 10.1007/s00520-025-09855-2] [Medline: 40858896]
10. Morberg Jämterud S, Sandgren A. Health care professionals' perceptions of factors influencing the process of identifying patients for serious illness conversations: a qualitative study. Palliat Med 2022 Jul;36(7):1072-1079. [doi: 10.1177/02692163221102266] [Medline: 35729752]
11. Beck S, Lundblad L, Göras C, Eneslätt M. Implementing advance care planning in Swedish healthcare settings: a qualitative study of professionals' experiences. Scand J Prim Health Care 2023 Mar;41(1):23-32. [doi: 10.1080/02813432.2022.2155456] [Medline: 36519794]
12. Park EJ, Jo M, Park M, Kang SJ. Advance care planning for older adults in community-based settings: an umbrella review. Int J Older People Nurs 2021 Sep;16(5):e12397. [doi: 10.1111/opn.12397] [Medline: 34216191]
13. Howard M, Bernard C, Tan A, Slaven M, Klein D, Heyland DK. Advance care planning: let's start sooner. Can Fam Physician 2015 Aug;61(8):663-665. [Medline: 26273075]
14. Dupont C, Smets T, Monnet F, et al. Publicly available, interactive web-based tools to support advance care planning: systematic review. J Med Internet Res 2022 Apr 20;24(4):e33320. [doi: 10.2196/33320] [Medline: 35442207]
15. Grant MS, Back AL, Dettmar NS. Public perceptions of advance care planning, palliative care, and hospice: a scoping review. J Palliat Med 2021 Jan;24(1):46-52. [doi: 10.1089/jpm.2020.0111] [Medline: 32614634]
16. Sellars M, Simpson J, Kelly H, et al. Volunteer involvement in advance care planning: a scoping review. J Pain Symptom Manage 2019 Jun;57(6):1166-1175. [doi: 10.1016/j.jpainsymman.2019.02.031] [Medline: 30853554]
17. Dujardin J, Schuurmans J, Westerduin D, Wichmann AB, Engels Y. The COVID-19 pandemic: a tipping point for advance care planning? Experiences of general practitioners. Palliat Med 2021 Jul;35(7):1238-1248. [doi: 10.1177/02692163211016979] [Medline: 34041987]
18. Smith GM, Hui FA, Bleymaier CR, Bragg AR, Harman SM. What if i get seriously ill? A virtual workshop for advance care planning during COVID-19. J Pain Symptom Manage 2020 Nov;60(5):e21-e24. [doi: 10.1016/j.jpainsymman.2020.08.022] [Medline: 32835831]
19. McAfee CA, Jordan TR, Cegelka D, et al. COVID-19 brings a new urgency for advance care planning: implications of death education. Death Stud 2022;46(1):91-96. [doi: 10.1080/07481187.2020.1821262] [Medline: 32941112]
20. Martin RS, Hayes B, Gregorevic K, Lim WK. The effects of advance care planning interventions on nursing home residents: a systematic review. J Am Med Dir Assoc 2016 Apr 1;17(4):284-293. [doi: 10.1016/j.jamda.2015.12.017] [Medline: 26861748]
21. Houben CHM, Spruit MA, Groenen MTJ, Wouters EFM, Janssen DJA. Efficacy of advance care planning: a systematic review and meta-analysis. J Am Med Dir Assoc 2014 Jul;15(7):477-489. [doi: 10.1016/j.jamda.2014.01.008] [Medline: 24598477]
22. Brinkman-Stoppelenburg A, Rietjens JAC, van der Heide A. The effects of advance care planning on end-of-life care: a systematic review. Palliat Med 2014 Sep;28(8):1000-1025. [doi: 10.1177/0269216314526272] [Medline: 24651708]
23. Jimenez G, Tan WS, Virk AK, Low CK, Car J, Ho AHY. Overview of systematic reviews of advance care planning: summary of evidence and global lessons. J Pain Symptom Manage 2018 Sep;56(3):436-459. [doi: 10.1016/j.jpainsymman.2018.05.016] [Medline: 29807158]
24. Ng R, Chan S, Ng TW, Chiam AL, Lim S. An exploratory study of the knowledge, attitudes and perceptions of advance care planning in family caregivers of patients with advanced illness in Singapore. BMJ Support Palliat Care 2013 Sep;3(3):343-348. [doi: 10.1136/bmjspcare-2012-000243] [Medline: 24644754]
25. Park HY, Kim YA, Sim JA, et al. Attitudes of the general public, cancer patients, family caregivers, and physicians toward advance care planning: a nationwide survey before the enforcement of the Life-Sustaining Treatment Decision-Making Act. J Pain Symptom Manage 2019 Apr;57(4):774-782. [doi: 10.1016/j.jpainsymman.2018.12.332] [Medline: 30593911]
26. Sævareid TJL, Pedersen R, Magelssen M. Positive attitudes to advance care planning: a Norwegian general population survey. BMC Health Serv Res 2021 Aug 2;21(1):762. [doi: 10.1186/s12913-021-06773-x] [Medline: 34334131]
27. Raijmakers NJH, Rietjens JAC, Kouwenhoven PSC, et al. Involvement of the Dutch general population in advance care planning: a cross-sectional survey. J Palliat Med 2013 Sep;16(9):1055-1061. [doi: 10.1089/jpm.2012.0555] [Medline: 23767949]
28. Montgomery C, Hickman SE, Wilkins C, Fromme EK, Anderson S. Montgomery et al's response to Morrison: advance directives/care planning: clear, simple, and wrong (doi: 10.1089/jpm.2020.0272). J Palliat Med 2021 Jan;24(1):12-13. [doi: 10.1089/jpm.2020.0523] [Medline: 32881594]
29. Sean Morrison R. Advance directives/care planning: clear, simple, and wrong. J Palliat Med 2020 Jul;23(7):878-879. [doi: 10.1089/jpm.2020.0272] [Medline: 32453620]
30. Tishelman C, Eneslätt M, Menkin ES, Van Den Block L. Tishelman et al's response to Morrison: advance directives/care planning: clear, simple, and wrong (doi: 10.1089/jpm.2020.0272). J Palliat Med 2021 Jan;24(1):16-17. [doi: 10.1089/jpm.2020.0540] [Medline: 33393888]

XSL•FO
RenderX

31. Combes S, Nicholson CJ, Gillett K, Norton C. Implementing advance care planning with community-dwelling frail elders requires a system-wide approach: an integrative review applying a behaviour change model. Palliat Med 2019 Jul;33(7):743-756. [doi: 10.1177/0269216319845804] [Medline: 31057042]
32. Myers J, Cosby R, Gzik D, et al. Provider tools for advance care planning and goals of care discussions: a systematic review. Am J Hosp Palliat Care 2018 Aug;35(8):1123-1132. [doi: 10.1177/1049909118760303] [Medline: 29529884]
33. Liu L, Zhao YY, Yang C, Chan HYL. Gamification for promoting advance care planning: a mixed-method systematic review and meta-analysis. Palliat Med 2021 Jun;35(6):1005-1019. [doi: 10.1177/02692163211005343] [Medline: 33775174]
34. Sloan DH, Hannum SM, DeGroot L, et al. Advance care planning shared decision-making tools for non-cancer chronic serious illness: a mixed method systematic review. Am J Hosp Palliat Care 2021 Dec;38(12):1526-1535. [doi: 10.1177/1049909121995416] [Medline: 33583195]
35. Biondo PD, King S, Minhas B, Fassbender K, Simon JE, Advance Care Planning Collaborative Research and Innovation Opportunities Program (ACP CRIO). How to increase public participation in advance care planning: findings from a world café to elicit community group perspectives. BMC Public Health 2019 Jun 3;19(1):679. [doi: 10.1186/s12889-019-7034-4] [Medline: 31159829]
36. Carter R, Hassan E, Barwich D, et al. Peer-facilitated public sessions as a strategy to increase engagement in advance care planning in British Columbia. J Pain Symptom Manage 2018 Dec;56(6):e20. [doi: 10.1016/j.jpainsymman.2018.10.004]
37. Menkin ES. Go Wish: a tool for end-of-life care conversations. J Palliat Med 2007 Apr;10(2):297-303. [doi: 10.1089/jpm.2006.9983] [Medline: 17472498]
38. Steinhauser KE, Christakis NA, Clipp EC, McNeilly M, McIntyre L, Tulsky JA. Factors considered important at the end of life by patients, family, physicians, and other care providers. JAMA 2000 Nov 15;284(19):2476-2482. [doi: 10.1001/jama.284.19.2476] [Medline: 11074777]
39. Lankarani-Fard A, Knapp H, Lorenz KA, et al. Feasibility of discussing end-of-life care goals with inpatients using a structured, conversational approach: the Go Wish card game. J Pain Symptom Manage 2010 Apr;39(4):637-643. [doi: 10.1016/j.jpainsymman.2009.08.011] [Medline: 20413053]
40. Tishelman C, Eneslätt M, Menkin E, Lindqvist O. Developing and using a structured, conversation-based intervention for clarifying values and preferences for end-of-life in the advance care planning-naïve Swedish context: action research within the DöBra research program. Death Stud 2022;46(4):803-815. [doi: 10.1080/07481187.2019.1701145] [Medline: 31858889]
41. Eneslätt M, Helgesson G, Tishelman C. Exploring community-dwelling older adults' considerations about values and preferences for future end-of-life care: a study from Sweden. Gerontologist 2020 Sep 15;60(7):1332-1342. [doi: 10.1093/geront/gnaa012] [Medline: 32201894]
42. Eneslätt M, Helgesson G, Tishelman C. Same, same, but different? A longitudinal, mixed-methods study of stability in values and preferences for future end-of-life care among community-dwelling, older adults. BMC Palliat Care 2021 Sep 22;20(1):148. [doi: 10.1186/s12904-021-00839-7] [Medline: 34551749]
43. Johansson T, Tishelman C, Cohen J, Eriksson LE, Goliath I. Continuums of change in a competence-building initiative addressing end-of-life communication in Swedish elder care. Qual Health Res 2021 Aug;31(10):1904-1917. [doi: 10.1177/10497323211012986] [Medline: 33980082]
44. Johansson T, Tishelman C, Eriksson LE, Cohen J, Goliath I. Use, usability, and impact of a card-based conversation tool to support communication about end-of-life preferences in residential elder care: a qualitative study of staff experiences. BMC Geriatr 2022 Apr 2;22(1):274. [doi: 10.1186/s12877-022-02915-w] [Medline: 35366816]
45. Mikaelsson Å, Eriksson LE, Stenfors T, Goliath I. Proactive end-of-life conversations in residential care homes: a qualitative interview study exploring residents' and family members' experiences. BMC Geriatr 2025 Apr 25;25(1):279. [doi: 10.1186/s12877-025-05916-7] [Medline: 40281509]
46. Kroik L, Eneslätt M, Tishelman C, Stoor K, Edin-Liljegren A. Values and preferences for future end-of-life care among the indigenous Sámi. Scand J Caring Sci 2022 Jun;36(2):504-514. [doi: 10.1111/scs.13047] [Medline: 34841544]
47. Eneslätt M, Helgesson G, Tishelman C. Dissemination, use, and impact of a community-based, conversational advance care planning intervention: ripple effects of the Swedish DöBra cards. Palliat Care Soc Pract 2021;15(1-17):26323524211032983. [doi: 10.1177/26323524211032983] [Medline: 34409297]
48. Campioni D, Ehrler F, Berger A, Clavien C. A digital serious game (Anticip'action) to support advance care planning discussions in the general population: usability study. JMIR Aging 2025 Aug 21;8:e73378. [doi: 10.2196/73378] [Medline: 40840871]
49. Monnet F, Dupont C, Smets T, et al. Advance care planning website for people with dementia and their family caregivers: protocol for a development and usability study. JMIR Res Protoc 2023 Jul 26;12:e46935. [doi: 10.2196/46935] [Medline: 37494084]
50. Monnet F, Dupont C, Pivodic L, et al. A user-centred website to support advance care planning for people with dementia and their family caregivers: development and usability study. Digit Health 2025;11:20552076241304018. [doi: 10.1177/20552076241304018] [Medline: 39991741]
51. Monnet F, Pivodic L, Dupont C, et al. Evaluation of interactive web-based tools to stimulate reflection and communication about advance care planning with people with dementia and their family caregivers. BMC Palliat Care 2024 Jun 28;23(1):162. [doi: 10.1186/s12904-024-01486-4] [Medline: 38943119]

XSL•FO
RenderX

52. Dupont C, Smets T, Potts C, et al. Uncovering specific navigation patterns by assessing user engagement of people with dementia and family caregivers with an advance care planning website: quantitative analysis of web log data. JMIR Aging 2025 Feb 11;8:e60652. [doi: 10.2196/60652] [Medline: 39935007]
53. Dupont C, Smets T, Monnet F, et al. A website to support people with dementia and their family caregivers in advance care planning: results of a mixed-method evaluation study. Patient Educ Couns 2024 Oct;127:108357. [doi: 10.1016/j.pec.2024.108357] [Medline: 38943816]
54. Murray E. Web-based interventions for behavior change and self-management: potential, pitfalls, and progress. Med 2 0 2012;1(2):e3. [doi: 10.2196/med20.1741] [Medline: 25075231]
55. Lustria MLA, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
56. Eneslätt M. Dösnack: metoder och verktyg för samtal om döende, död och framtida vård i livets slutskede [Dissertation in Swedish]. : Karolinska Institutet; 2021 URL: https://www.avhandlingar.se/avhandling/663a55af11/ [accessed 2026-01-28]
57. Lee MC, Hinderer KA, Alexander CS. What matters most at the end-of-life for Chinese Americans? Gerontol Geriatr Med 2018;4:2333721418778195. [doi: 10.1177/2333721418778195] [Medline: 30035200]
58. Li T, Pei X, Chen X, Zhang S. Identifying end-of-life preferences among Chinese patients with cancer using the Heart to Heart card game. Am J Hosp Palliat Care 2021 Jan;38(1):62-67. [doi: 10.1177/1049909120917361] [Medline: 32270684]
59. Litzelman DK, Inui TS, Schmitt-Wendholt KM, et al. Clarifying values and preferences for care near the end of life: the role of a new lay workforce. J Community Health 2017 Oct;42(5):926-934. [doi: 10.1007/s10900-017-0336-5] [Medline: 28353007]
60. Möller UO, Pranter C, Hagelin CL, et al. Using cards to facilitate conversations about wishes and priorities of patients in palliative care. J Hosp Palliat Nurs 2020 Feb;22(1):33-39. [doi: 10.1097/NJH.0000000000000607] [Medline: 31770159]
61. Osman H, El Jurdi K, Sabra R, Arawi T. Respecting patient choices: using the “Go Wish” cards as a teaching tool. BMJ Support Palliat Care 2018 Jun;8(2):194-197. [doi: 10.1136/bmjspcare-2017-001342] [Medline: 28634216]
62. Perin M, Tanzi S, Botrugno C, et al. Translation and cultural adaptation of the Go Wish Game: thinking about personal values to promote advance care planning. J Palliat Med 2022 Oct;25(10):1540-1550. [doi: 10.1089/jpm.2022.0083] [Medline: 35862002]
63. Thompson S, Lyford M, Papertalk L, Holloway M. Passing on wisdom: exploring the end-of-life wishes of Aboriginal people from the midwest of western Australia. Rural Remote Health 2019 Nov;19(4):5444. [doi: 10.22605/RRH5444] [Medline: 31782929]
64. Dupont C, Smets T, Monnet F, Eneslätt M, Tishelman C, Van den Block L. The cultural adaptation of the Go Wish card game for use in Flanders, Belgium: a public health tool to identify and discuss end-of-life preferences. BMC Public Health 2022 Nov 17;22(1):2110. [doi: 10.1186/s12889-022-14523-9] [Medline: 36397020]
65. Bradbury H. Introduction: how to situate and define action research. In: The Sage Handbook of Action Research, 3rd edition: SAGE Publications; 2015. URL: https://study.sagepub.com/sites/default/files/front_tab_request_v3.pdf [accessed 2026-01-28]
66. Elwyn G, Kreuwel I, Durand MA, et al. How to develop web-based decision support interventions for patients: a process map. Patient Educ Couns 2011 Feb;82(2):260-265. [doi: 10.1016/j.pec.2010.04.034] [Medline: 20627644]
67. Eccles DW, Arsal G. The think aloud method: what is it and how do I use it? Qual Res Sport Exerc Health 2017 Aug 8;9(4):514-531. [doi: 10.1080/2159676X.2017.1331501]
68. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum Comput Interact 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
69. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland AL, Weerdmeester BA, editors. Usability Evaluation in Industry: CRC Press; 1996. [doi: 10.1201/9781498710411]
70. McCauley CO, Bond RB, Ryan A, et al. Evaluating user engagement with a reminiscence app using cross-comparative analysis of user event logs and qualitative data. Cyberpsychol Behav Soc Netw 2019 Aug;22(8):543-551. [doi: 10.1089/cyber.2019.0076] [Medline: 31403855]
71. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
72. Thorne S. Interpretive Description: Taylor & Francis; 2016. [doi: 10.4324/9781315545196]
73. Saldaña J. Longitudinal Qualitative Research: Analyzing Change through Time: Rowman & Littlefield; 2003. URL: https://www.biblio.com/book/longitudinal-qualitative-research-analyzing-change-through/d/1687874027 [accessed 2026-01-28]
74. Audulv Å, Hall EOC, Kneck Å, et al. Qualitative longitudinal research in health research: a method study. BMC Med Res Methodol 2022 Oct 1;22(1):255. [doi: 10.1186/s12874-022-01732-4] [Medline: 36182899]
75. Hyzy M, Bond R, Mulvenna M, et al. System Usability Scale benchmarking for digital health apps: meta-analysis. JMIR mHealth uHealth 2022 Aug 18;10(8):e37290. [doi: 10.2196/37290] [Medline: 35980732]
76. Maqbool B, Herold S. Potential effectiveness and efficiency issues in usability evaluation within digital health: a systematic literature review. J Softw Syst 2024 Feb;208:111881. [doi: 10.1016/j.jss.2023.111881]

XSL•FO
RenderX

77. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
78. Bloomer MJ, Hutchinson AM, Brooks L, Botti M. Dying persons' perspectives on, or experiences of, participating in research: an integrative review. Palliat Med 2018 Apr;32(4):851-860. [doi: 10.1177/0269216317744503] [Medline: 29235418]
79. Kellehear A, Leget C. Questioning risk and vulnerability in end-of-life research. Asia Pac J Soc Work Dev 2025:1-13. [doi: 10.1080/29949769.2025.2555001]
80. Kleijberg M. Studio döbra—creating spaces for engaging with end-of-life issues and for interacting across generations through community-based arts activities [Dissertation]. : Karolinska Institutet; 2021 URL: https://www.proquest.com/openview/8d2933067c20a77b22d75ceb8a9bf3a8/1?pq-origsite=gscholar&cbl=2026366&diss=y [accessed 2026-01-28]
81. Kleijberg M, Ahlberg BM, Macdonald A, Lindqvist O, Tishelman C. Navigating power dynamics in engaging communities in end-of-life issues: lessons learned from developing community-based intergenerational arts initiatives about death and loss. Death Stud 2021;45(8):651-664. [doi: 10.1080/07481187.2019.1671547] [Medline: 31604379]
82. Kleijberg M, Hilton R, Ahlberg BM, Tishelman C. Conceptualizing impact in community-based participatory action research to engage communities in end-of-life issues. Palliat Care Soc Pract 2022;16(16):26323524221095107. [doi: 10.1177/26323524221095107] [Medline: 35573090]
83. Prince-Paul M, DiFranco E. Upstreaming and normalizing advance care planning conversations-a public health approach. Behav Sci (Basel) 2017 Apr 12;7(2):18. [doi: 10.3390/bs7020018] [Medline: 28417931]
84. God och nära vård: en gemensam färdplan och målbild [Report in Swedish]. : Statens Offentliga Utredningar; 2017 SOU 2017:53 URL: https://data.riksdagen.se/fil/5E16E1B5-2D0B-4F28-AB50-F83A04B779D4 [accessed 2026-01-28]
85. Insatser under 2020 för att främja omställningen till en god och nära vård [Report in Swedish]. : Socialstyrelsen; 2021 URL: https://www.socialstyrelsen.se/contentassets/cbc8e284105e4e0b863c236dbe15ba3c/2021-3-7305.pdf [accessed 2026-01-28]
86. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: an integrative systematic review. BMC Public Health 2024 Feb 16;24(1):506. [doi: 10.1186/s12889-024-18036-5] [Medline: 38365698]
87. Mumtaz H, Riaz MH, Wajid H, et al. Current challenges and potential solutions to the use of digital health technologies in evidence generation: a narrative review. Front Digit Health 2023;5:1203945. [doi: 10.3389/fdgth.2023.1203945] [Medline: 37840685]
88. Wilson J, Heinsch M, Betts D, Booth D, Kay-Lambkin F. Barriers and facilitators to the use of e-health by older adults: a scoping review. BMC Public Health 2021 Aug 17;21(1):1556. [doi: 10.1186/s12889-021-11623-w] [Medline: 34399716]
89. Iyanna S, Kaur P, Ractham P, Talwar S, Najmul Islam AKM. Digital transformation of healthcare sector. What is impeding adoption and continued usage of technology-driven innovations by end-users? J Bus Res 2022 Dec;153:150-161. [doi: 10.1016/j.jbusres.2022.08.007]

Abbreviations

ACP: advance care planning
EoL: end-of-life
PAR: participatory action research
SO: significant other
SUS: System Usability Scale
TP: target person
WP: work package

Edited by J Sarvestan; submitted 02.Dec.2025; this is a non–peer-reviewed article;accepted 19.Jan.2026; published 25.Feb.2026.

Please cite as:
Eneslätt M, Dupont C, Savolainen E, Fjällström P, Goliath I
Cocreation of a Digital Tool for Proactive End-of-Life Communication: Protocol for a Participatory Action Research Project
JMIR Res Protoc 2026;15:e88452
URL: https://www.researchprotocols.org/2026/1/e88452
doi:10.2196/88452

© Malin Eneslätt, Charlèss Dupont, Eva Savolainen, Petter Fjällström, Ida Goliath. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution,

XSL•FO
RenderX

and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Food Insecurity Determinants, Coping Strategies, and Association With Nutritional Status Among Hemodialysis Patients in Pahang, Malaysia: Protocol for a Mixed Methods Study

Nor Azwani Mohd Shukri[1,2], PhD; Sarah Muneera Karami[1], MHSc; Wan Azdie Mohd Abu Bakar[1,2], PhD; Roselawati Mat Ya[3], PhD; Norhasmah Sulaiman[4], PhD

[1]Department of Nutrition Sciences, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Kuantan, Pahang, Malaysia

[2]Food Security and Public Health Nutrition Research Group (FOSTER), Kulliyyah of Allied Health Sciences, International Islamic University Malaysia, Kuantan, Pahang, Malaysia

[3]Department of Community Medicine, Kulliyyah of Medicine, International Islamic University Malaysia, Kuantan, Pahang, Malaysia

[4]Department of Nutrition, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Serdang, Selangor, Malaysia

Corresponding Author:
Nor Azwani Mohd Shukri, PhD
Department of Nutrition Sciences
Kulliyyah of Allied Health Sciences
International Islamic University Malaysia
Jalan Sultan Ahmad Shah
Kuantan, Pahang, 25200
Malaysia
Phone: 60 95705307
Email: norazwani@iium.edu.my

Abstract

Background: Food insecurity involves the lack of physical and economic access to sufficient, safe, and nutritious food that meets an individual's dietary needs. Patients with end-stage renal disease (ESRD) undergoing hemodialysis have specific dietary needs, but food insecurity may hinder them from adhering to prescribed guidelines and preserving their nutritional status. However, no research has been conducted to elucidate food insecurity among patients undergoing hemodialysis in Malaysia.

Objective: This study aimed to assess the prevalence of food insecurity, its determinants, and its association with nutritional status and explore the coping strategies used among patients undergoing hemodialysis in Pahang, Malaysia.

Methods: This was a cross-sectional study that followed a mixed methods approach and was conducted with patients undergoing hemodialysis for ESRD at the Pahang Islamic Religious Council and Malay Customs (Majlis Ugama Islam dan Adat Resam Melayu Pahang [MUIP]) dialysis centers. The inclusion criteria were patients who were aged 18 years or older, were generally healthy, and had been undergoing hemodialysis regularly for at least 3 months. The food security status of the participants was determined using the Malay version of the Food Insecurity Experience Scale (M-FIES). The nutritional status included anthropometric measurements (height, weight, BMI, triceps skinfold [TSF] thickness, mid-upper arm circumference (MUAC), midarm muscle circumference (MAMC), and body fat percentage), biochemical parameters (serum urea, creatinine, albumin, phosphate, potassium, hemoglobin, and total iron-binding capacity [TIBC]), clinical assessments (Malnutrition Inflammation Score [MIS] and protein energy wasting [PEW]), and dietary intake (adherence to total calorie, protein, sodium, potassium, and phosphorus intake and diet monotony index [DMI]). The determinants were identified using logistic regression, while the association between food security status and nutritional status was analyzed using the chi-square or Fisher's exact test and the independent *t* test. Semistructured interviews involved participants who were categorized as mildly, moderately, or severely food insecure in order to explore the contributing factors of food insecurity and how they coped with it based on their lived experience. The interviews were carried out until the data reached saturation, and then the data were analyzed thematically.

Results: The study was funded by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2023/SS10/UIAM/02/1) starting in September 2023. Data collection was conducted from December 2023 until August 2024, involving 287 participants, and data analysis has also been completed. As of January 2026, quantitative findings are under review and qualitative findings are being prepared.

XSL•FO
RenderX

Conclusions: The implementation of this study protocol will provide new evidence to improve the understanding of food insecurity in this population. Elucidation of its key contributing factors, coping strategies, and potential connections to nutritional status in this population can help guide more informed policymaking and effective interventions.

International Registered Report Identifier (IRRID): RR1-10.2196/84575

(JMIR Res Protoc 2026;15:e84575) doi:10.2196/84575

KEYWORDS

coping strategies; food insecurity determinants; end-stage renal disease; food insecurity; hemodialysis; nutritional status

Introduction

The burden of food insecurity continues to rise worldwide, impacting individuals in diverse socioeconomic and geographic settings. Foor insecurity is increasingly recognized as a growing public health concern in both high- and middle-income countries. The 1996 World Food Summit defined food insecurity as a condition in which individuals lack physical, social, and economic access to sufficient, safe, and nutritious food that meets their dietary needs and food preferences for an active and healthy life [1]. It is associated with cardiometabolic risks, such as obesity, diabetes, hypertension, dyslipidemia, and stress [2]. According to the *2025 Global Report on Food Crises* (GRFC), acute hunger has risen for the sixth consecutive year, with 295.3 million people across 53 countries in the world experiencing acute food insecurity in 2024—threefold the number in 2016 and double that in 2020 [3].

Food insecurity can result in poor nutrition, an elevated risk of chronic diseases, and a deterioration of the health outcomes of patients with chronic kidney disease (CKD), as well as contributing toward the progression of CKD to end-stage renal disease (ESRD) [4,5]. This study is informed by Campbell's conceptual framework [6], which hypothesizes food insecurity as multidimensional, influenced by the researcher's conceptualization of its social context, such as socioeconomic characteristics, food access, and health-related factors, and linked to downstream nutritional and clinical outcomes. Previous studies among patients with CKD have identified demographic, socioeconomic, and social factors, including age, marital status, income, educational level, employment status, financial constraints, limited access to transportation, lack of social support, and a high financial burden, as key determinants of food insecurity [7-9].

Food insecurity has been reported to affect 36% [7] and 26.3% [10] of patients undergoing hemodialysis and is a critical issue among these patients since they have special dietary needs for preserving health. Adequate calorie and protein intake and restricted sodium, potassium, and phosphorus intake, with controlled fluid, are required to prevent malnutrition [11]. Patients with CKD facing food insecurity report greater difficulties in adhering to a costly renal diet compared to those without food insecurity [10]. Exacerbated by financial difficulties and inadequate transportation, which are associated with restricted access to nutritious food, the presence of food insecurity may pose challenges for patients undergoing hemodialysis to comply with these dietary guidelines and maintain their nutritional status [9].

Coping mechanisms are crucial for individuals or households experiencing food insecurity to facilitate managing limited resources and mitigating the immediate impact of inadequate food or resources to buy food. Individuals with chronic conditions (eg, diabetes, HIV/AIDS, and cancer) and food insecurity have adopted coping strategies that fall into two categories: food related and nonfood related. These strategies include consuming less preferred food [12], reducing the meal frequency [12], skipping meals [13], prioritizing food expenses over medical treatments [13], relying on low-cost food to save money for medical supplies [14], and receiving food assistance [13,15].

Although food insecurity among patients with CKD has been widely studied in high-income countries, such as the United States [7,9,10], evidence from Southeast Asia, including Malaysia, remains limited. Malaysia presents a unique context due to its mixed public-private health care system, the rising cost of living, and the increasing prevalence of CKD requiring long-term hemodialysis. Understanding food insecurity within this local context is essential, as socioeconomic conditions, food accessibility, and health care–financing systems may vary substantially from those reported from other regions.

Given the limited local evidence and the unique clinical and socioeconomic challenges faced by patients undergoing hemodialysis in Malaysia, this mixed methods study aims to examine food insecurity among patients undergoing hemodialysis using a quantitative approach and to explore the phenomenon in depth through qualitative inquiry based on the lived experiences of these patients. Specifically, the study aims to (1) assess the prevalence of food insecurity, (2) identify its determinants, (3) examine the association between food security status and nutritional status, and (4) explore the contributing factors and coping strategies related to food insecurity among patients undergoing hemodialysis in Pahang, Malaysia.

Methods

Study Design

This cross-sectional study used a mixed methods approach, which includes quantitative and qualitative data collection. This approach was followed to complement and expand the findings from the quantitative research by seeking elaboration and clarification through qualitative research. Data collection was conducted sequentially. It commenced with a broad survey involving a large number of people to generalize the findings to a population, followed by a second phase that included detailed qualitative, semistructured interviews for a deeper exploration of the experiences of selected cases or individuals

XSL•FO
RenderX

with food insecurity [16]. The timing, setting, and sequence of data collection activities for both the quantitative and qualitative phases are summarized in Table 1.

Table 1. Overview of study procedures, measures, and data collection schedule.

Data collection activity and variables/measures	Timing	Setting
Quantitative questionnaire		
Food security status	During dialysis sessions	Dialysis center
Sociodemographic and other risk factors of food insecurity	During dialysis sessions	Dialysis center
Quantitative nutritional status assessments		
Anthropometric measurements	During and after dialysis sessions	Dialysis center
Biochemical parameters	From medical records	Dialysis center
Clinical status assessments	During dialysis sessions	Dialysis center
Dietary intake assessments	During dialysis sessions	Dialysis center
Qualitative semistructured interviews		
Contributing factors and coping strategies related to food insecurity	Outside dialysis hours, scheduled separately	Telephone call

Study Population

This study was conducted with patients undergoing hemodialysis at four dialysis centers managed by the Pahang Islamic Religious Council and Malay Customs (*Majlis Ugama Islam dan Adat Resam Melayu Pahang* [MUIP]). The council conducts social development programs as part of its *zakat* administration, which is an Islamic charitable contribution aimed at supporting those in need, through the establishment of centers that provide hemodialysis services to patients with ESRD who have a low socioeconomic status or who are classified as *asnaf* (individuals deemed eligible to receive *zakat* funds). Each of these dialysis centers is in the Kuantan, Gambang, Pekan, and Pusat Bandar Jengka areas, covering three different districts in Pahang, one of the east coast states of Malaysia.

Ethical Considerations

This study was approved by the International Islamic University Malaysia Research Ethics Committee (ID IREC 2024-001). Administrative permission from the MUIP dialysis centers was obtained before data collection. Before the study commenced, its details were explained to all participants, and written informed consent was obtained from them. Their personal information was anonymized, and confidentiality was maintained throughout the study.

Quantitative Phase

Research Framework

The quantitative phase of this study was guided by a research framework that focuses on the determinants of food insecurity among patients undergoing hemodialysis and the association of food insecurity with nutritional status (Figure 1). Based on previous studies, demographic and socioeconomic data [8,17,18] as well as physical and economic access to food [19,20] are the known determinants of food security status. However, since this study was conducted with patients undergoing hemodialysis, disease-/treatment-related factors [8,14,21,22] could be one of the variables affecting their food security status. Nutrition literacy and dietary adherence levels of participants [23] were also assessed as potential risk factors to investigate whether the study population is experiencing food insecurity due to a lack of knowledge about a healthy diet for hemodialysis or compliance with dietary recommendations. Furthermore, Campbell's conceptual framework of risk factors and consequences of food insecurity [6] depicts food insecurity as having an impact on nutritional status.

Figure 1. Research framework of the quantitative phase of the study.

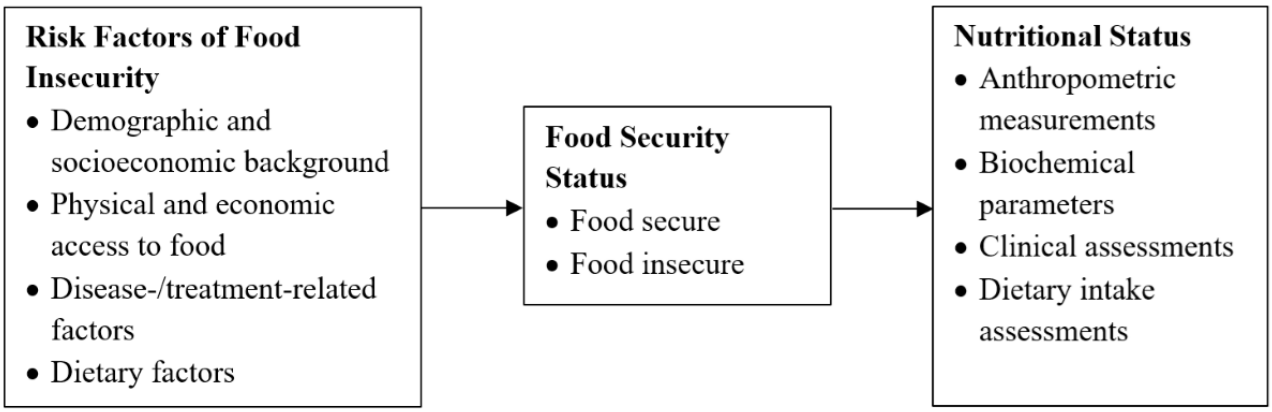

XSL•FO
RenderX

Participant Recruitment and Sampling

The inclusion criteria for the study were as follows:

- Age at least 18 years
- Undergoing regular hemodialysis for 4 hours per session, three times a week
- Undergoing dialysis for at least 3 months before enrolling in the study
- No major acute diseases or psychological disorders

The exclusion criteria were as follows:

- Inability to hear, understand, or speak well
- Not mentally and physically fit to complete all the assessments

All patients undergoing hemodialysis who fulfilled the inclusion criteria were included in this study, except those who were unresponsive after being approached at least three times. Universal sampling was applied in the quantitative phase, whereby all eligible patients undergoing hemodialysis at the MUIP dialysis centers during the study period were invited to participate. As this approach represents a census of the accessible population, a fixed sample size was not predetermined. Instead, the precision of the prevalence estimate was evaluated using the margin of error (d) derived from a single-proportion formula. Assuming a food insecurity prevalence of 26.3% among patients undergoing hemodialysis, as reported previously [9], a sample size of approximately 250 participants would correspond to a margin of error of 5% at a 95% confidence level. Recruitment beyond this number was considered sufficient to ensure adequate precision of the study estimates.

Data Collection and Instruments

Data collection was conducted in the Malay language using interviewer-administered questionnaires, during which the researcher asked the participants questions and recorded their responses.

Objective 1: Assessing the Prevalence of Food Insecurity Among Patients Undergoing Hemodialysis

The food security status of participants was ascertained using the Malay version of the Food Insecurity Experience Scale (M-FIES) [24], which consists of eight questions that assess the severity of food insecurity at the household or individual level, as determined by the individuals' self-reported behaviors and experiences with respect to one of the dimensions of food insecurity, food accessibility [1]. Participants were required to respond with either "yes" or "no" to each question. An affirmative response was assigned a score of 1, while a negative response was assigned a score of 0. The sum of scores from the eight questions was further categorized into the following severity levels: food security (0), mild food insecurity (1-3), moderate food insecurity (4-6), and severe food insecurity (7-8). For data analysis, the M-FIES categories of mild, moderate, and severe food insecurity were combined and classified as "food insecurity."

Objective 2: Identifying the Determinants of Food Insecurity Among Patients Undergoing Hemodialysis

To identify the determinants of food insecurity, four main risk factors of food insecurity were included in the questionnaire. For demographic and socioeconomic data, questions about the participants' age, gender, ethnicity, marital status, educational level, employment status, total monthly income, financial assistance status, living area, and social support status were asked. Social support status was determined using the adapted Malay version of the Multidimensional Perceived Social Support Scale (MSPSS-M) [25,26]. Only 8 of 12 questions to measure an individual's perception of support from family and friends were included. To calculate the mean subscale score for the family subscale, the sum of scores for items 1, 2, 5, and 7 was divided by 4, while for the friend subscale, the sum of scores for items 3, 4, 6, and 8 was divided by 4. A mean scale score between 1 and 2.9 was classified as low social support, a score between 3 and 5 was classified as moderate social support, and a score between 5.1 and 7 was classified as having high social support.

Questions regarding physical and economic access to food included shop availability, the distance from the house, food variety, the affordability of purchasing food, transport ownership, food delivery status, and weekly food expenditure. This section also inquired about the disease- and treatment-related data of participants, such as the presence of comorbidities, duration of the disease, dialysis vintage, and self-reported appetite rating. For the dietary part, two validated questionnaires, the Malay version of the Dialysis-Specific Nutrition Literacy Scale (DSNLS) [27] and the dietary adherence component of the modified End-Stage Renal Disease Adherence Questionnaire (ESRD-AQ) [28], were used to assess the participants' nutrition literacy and dietary adherence, respectively.

Objective 3: Determining the Association Between Food Security Status and Nutritional Status of Patients Undergoing Hemodialysis

To determine how food security status affects the nutritional status of patients undergoing hemodialysis, this component of the study assessed participants' anthropometric measurements, biochemical parameters, clinical characteristics, and dietary intake.

Anthropometric Measurements

The participants' height, (postdialysis) dry weight, BMI, triceps skinfold (TSF) thickness, mid-upper arm circumference (MUAC), midarm muscle circumference (MAMC), and total fat percentage were measured. The anthropometry status assessment was only conducted after the participants completed the hemodialysis session.

Height

The participants' height was measured using a stadiometer placed upright against a straight wall surface to ensure accurate measurement. The procedure was as follows:

- The participant was requested to remove their shoes and hat, if any.

XSL•FO
RenderX

- The head plate of stadiometer was raised to make enough space for the participant to stand underneath it.
- The participant was instructed to step onto the stadiometer; stand upright against the rod, with their heels together and arms on the face outward; and face forward.
- The participant's head was ensured to be in Frankfort position.
- The headboard was lowered onto the highest point of the participant's head with sufficient pressure to compress the hair.
- The height was read at eye level and recorded to the nearest 0.5 cm. The height was measured three times, and the median value was recorded as the final measurement.

For participants who were unable to stand, height was measured by knee height as follows:

- The participant was seated comfortably with feet flat on the floor and knees positioned at a 90° angle.
- Any footwear was removed to ensure accuracy, and a measuring tape was used for measurement.
- The participant's thigh was positioned parallel to the floor.
- The measuring tape was positioned with its upper end placed on the top of the participant's knee.
- The knee height was measured from the bottom of the heel (or the sole if the heel was not accessible) to the tope edge of the measuring tape.
- The measurement was recorded to the nearest centimeter.
- The same process was repeated on the other leg to ensure consistency, and the average of the two readings was used, if appropriate.

Postdialysis Dry Weight

A digital weighing scale (Seca 677) was used to measure the postdialysis dry weight of the participants as follows:

- The participant was asked to remove their shoes and all objects that could affect weight measurement, if any.
- The participant was instructed to step onto the scale and face forward.
- The weight was measured three times, and the median value was recorded as the final measurement.

Body Mass Index

The participants' BMI was calculated as weight (kg)/height squared (m^2). The value was interpreted using the BMI classification [29] shown in Table 2.

Table 2. BMI classification.

BMI range (kg/m²)	Classification
<18.5	Underweight
18.5-24.9	Normal
25.0-29.9	Overweight
≥30.0	Obese

Triceps Skinfold Thickness

The participants' TSF thickness was measured on the nonfistula arm using the Harpenden skinfold caliper (HSK-BI, British Indicators) as follows:

- The acromion and olecranon processes were marked. A flexible, nonstretchable measuring tape (Crescent Lufkin W606PD, Apex Tool Group) was used to mark the midpoint between the two processes on the lateral aspect of the arm.
- The measuring tape was wrapped around the arm horizontally.
- The midpoint mark was drawn on the back of the arm to indicate the location for positioning the caliper.
- The caliper was placed 1 cm below the finger. The TSF was measured three times, and the median value was recorded as the final measurement.

Mid-Upper Arm Circumference

The participants' MUAC was measured using a flexible, nonstretchable measuring tape (Crescent Lufkin W606PD, Apex Tool Group). The measurement was obtained at the midpoint of the upper arm between the acromion and olecranon processes.

Midarm Muscle Circumference

Subsequently, the participants' MAMC was calculated using the following formula:

$$\text{MAMC (cm)} = \text{MUAC (cm)} - [\text{TSF (cm)} \times \pi]$$

The value was interpreted according to the MAMC Standard Reference [30] shown in Table 3.

Table 3. MAMC[a] standard reference.

Gender	Standard reference (cm)	90% of standard reference: moderately malnourished (cm)	60% of standard reference: severely malnourished (cm)
Men	25.3	22.8	15.2
Women	23.5	20.9	13.9

[a]MAMC: midarm muscle circumference.

XSL•FO
RenderX

Body Fat Percentage

The participants' fat mass was measured using a body composition analyzer (Omron Body Composition Monitor HBF-702T) according to the manufacturer's guidelines [31] and interpreted as per Table 4. The procedure was as follows:

- The participant was asked to remove their stockings.
- The participant was instructed to step onto the body composition analyzer and lift the handle of the instrument using both hands.
- The value displayed on the monitor was recorded.

Table 4. Healthy body fat percentage by age.

Age range (years)	Body fat in men (%)	Body fat in women (%)
20-39	8-19	21-32
40-59	11-21	23-33
60-79	13-24	24-35

Biochemical Parameters

Serum urea, creatinine, albumin, phosphate, potassium, hemoglobin, and total iron-binding capacity (TIBC) values were retrieved from the patients' medical records of routine three-monthly blood assessments. The status of each parameter (normal vs abnormal) was classified based on the reference range provided in the laboratory reports.

Clinical Assessments

Malnutrition Inflammation Score

The participants' Malnutrition Inflammation Score (MIS) was determined during anthropometric assessment. The tool is divided into four sections: (12) nutritional history (change in dry weight, dietary intake, gastrointestinal symptoms, functional capacity, and comorbidities with dialysis vintage), (2) physical examination (decrease in fat stores or loss of subcutaneous fat and signs of muscle wasting), (3) BMI, and (4) laboratory values (serum albumin and TIBC). Each component has four severity levels, scored from 0 (normal) to 1 (mild), 2 (moderate), and 3 (very severe). The sum of all 10 MIS components ranges from 0 to 30, with a higher score reflecting a more severe degree of malnutrition and inflammation [32]. Scores ≥8 indicated malnourishment, while scores <8 indicated that participants were well-nourished, as validated in a local study [33].

Protein Energy Wasting

Protein energy wasting (PEW) indicates a state of decreased protein and energy stores and is associated with reduced functional capacity, impaired quality of life, and increased morbidity and mortality in patients with CKD [34]. The diagnostic criteria proposed by the International Society of Renal Nutrition and Metabolism Expert Group [34] was adopted to identify the presence of PEW among the participants. The PEW diagnosis comprises four main criteria: biochemical assessment; low body weight, reduced total body fat, or weight loss; decreased muscle mass; and low energy or protein intake. A diagnosis of PEW was made if any three of the four criteria were met: serum albumin<3.8 g/dL, BMI<23 kg/m^2, >10% reduction in MAMC in relation to the 50th percentile of a reference population, and daily dietary energy intake<25 kcal/kg ideal body weight (IBW) [35].

Dietary Intake Assessments

A multiple-pass 24-hour diet recall technique [36] was used to evaluate participants' dietary intake of total energy, protein, and micronutrients, such as sodium, potassium, and phosphorus, covering 3 days (one dialysis day, one nondialysis day, and one weekend day). Face-to-face interviews were conducted by a trained dietitian (the researcher). The dietary data were analyzed using Nutritionist Pro diet analysis software with the USDA Food Database and Malaysian Food Composition Tables (Axxya System LLC). Each participant's usual dietary intake was calculated from the average of the three 24-hour dietary recalls. The residual method was used to adjust protein intake for total energy intake to reduce the effect of measurement errors, such as misreporting of energy intake, which could potentially impact the overall nutritional values [37]. The protein intake that corresponded with the mean total energy intake of the study population was adjusted by adding the residual (the difference between observed nutrient values for each participant and values predicted by the regression equation) to the intake. As such, protein intake was adjusted for total energy intake using the residual method calculated from regression, with total energy serving as the independent variable and protein intake as the dependent variable. Therefore, the residuals were then used as energy-adjusted protein intake in further analyses.

The adequacy of energy and protein intake was determined through comparison with the Kidney Disease Outcomes Quality Initiative (KDOQI) clinical practice guidelines for nutrition in chronic renal failure [10]. The daily recommended energy intake for patients undergoing hemodialysis is 30-35 kcal/kg IBW, while the recommended protein intake is 1.0-1.2 g/kg IBW per day. The patient's body weight is determined using the dry weight that is taken following the completion of the dialysis session. For micronutrient intake, patients with CKD who are undergoing dialysis are advised to limit their sodium intake to less than 2300 mg/day [10,38]. Their daily potassium intake should not surpass 2730 mg/day [38]. It is also recommended that the daily intake of phosphorus be less than 1000 mg/day [10,38]. Participants who adhered to the recommended dietary intake values were classified as "adhere," while those who exceeded the values were classified as "not adhere."

Dietary intake assessment also included the evaluation of dietary variety using the diet monotony index (DMI), which was calculated based on the serving sizes of 27 food groups, as assessed through the 24-hour dietary recall [39,40]. The

XSL•FO
RenderX

proportion of total servings per week accounted for by each specific food was quantified. The DMI was calculated as follows:

$$[(P_1)^2 + (P_2)^2 + (P_3)^2 + \ldots + (P_n)^2] \times 100,$$

where P_1, P_2, . . ., P_n are the proportion of total servings from each specific food. The resulting score mainly reflected food groups that represent the majority of total consumption. Consequently, higher DMI scores indicated a more monotonous diet, while lower scores suggested a greater variety of diets [39].

Statistical Analysis

IBM SPSS was used to conduct statistical analyses. Descriptive statistics was used to explore the participants' demographic and socioeconomic characteristics, physical and economic access to food, disease-/treatment-related factors, dietary factors, and food security status. Bivariate logistic regression was used to identify the association between each possible determinant and food security status. Variables with $P<.25$ in the bivariate logistic regression were included in the multivariable logistic regression model to determine independent associations with adjusted odds ratios (aORs) [41,42]. The area under the receiver operating characteristic (ROC) curve (AUC) was also tested on the regression model to assess model discrimination [43]. The values ranged from 0 to 1, where 0.5 indicated that the model is useless for discrimination, while values near 1 indicated that the model has a high predictive power and can discriminate between the two outcome groups. The statistical tests that were used to determine the association between food security status and nutritional status are presented in Table 5. The statistical significance value was set at $P<.05$.

XSL•FO
RenderX

Table 5. Statistical tests to determine the association between food security status (independent variable) and nutritional status.

Dependent variables	Statistical test
Anthropometric measurements	
Height	Independent *t* test
Weight	Independent *t* test
BMI status	Chi-square test (or Fisher's exact test)
TSF[a]	Independent *t* test
MUAC[b]	Independent *t* test
MAMC[c]	Independent *t* test
Total fat	Independent *t* test
Biochemical parameters	
Serum urea	Chi-square test (or Fisher's exact test)
Serum creatinine	Chi-square test (or Fisher's exact test)
Serum albumin	Chi-square test (or Fisher's exact test)
Serum phosphate	Chi-square test (or Fisher's exact test)
Serum potassium	Chi-square test (or Fisher's exact test)
Hemoglobin	Chi-square test (or Fisher's exact test)
TIBC[d]	Chi-square test (or Fisher's exact test)
Clinical assessments	
MIS[e]	Chi-square test (or Fisher's exact test)
PEW[f]	Chi-square test (or Fisher's exact test)
Dietary intake assessments	
Adherence to total calorie intake	Chi-square test (or Fisher's exact test)
Adherence to protein intake	Chi-square test (or Fisher's exact test)
Adherence to sodium intake	Chi-square test (or Fisher's exact test)
Adherence to potassium intake	Chi-square test (or Fisher's exact test)
Adherence to phosphate intake	Chi-square test (or Fisher's exact test)
DMI[g]	Independent *t* test

[a]TSF: triceps skinfold.
[b]MUAC: mid-upper arm circumference.
[c]MAMC: midarm muscle circumference.
[d]TIBC: total iron-binding capacity.
[e]MIS: Malnutrition Inflammation Score.
[f]PEW: protein energy wasting.
[g]DMI: diet monotony index.

Qualitative Phase

Participant Recruitment and Sampling

Participants for the qualitative component of this study included patients undergoing hemodialysis with different levels of food insecurity. They were a subgroup of the main study population. Upon completion of the quantitative phase of the study, some of the participants deemed suitable for interviews were approached by the researcher. Using purposive sampling, those who were classified as having mild, moderate, or severe food insecurity based on their M-FIES score were invited to participate in semistructured interviews. They were purposively selected to achieve approximately equal representation across the four hemodialysis centers and the three categories of food insecurity. The interview's purpose and process were explained to them, and if they agreed, their phone numbers were recorded.

XSL•FO
RenderX

Data Collection and Instruments

Objective 4: Exploring the Contributing Factors and Coping Strategies Related to Food Insecurity Among Patients Undergoing Hemodialysis

A semistructured interview guide adopted and adapted by a previous study [44] was used to assist the interviewer and to maintain consistency during the interview session. The guidelines include a brief welcome of the interviewee, an explanation about the procedure, and questions and prompts related to the topic (Multimedia Appendix 1). Before conducting the interview, the researcher familiarized themselves with the interview guide to ensure a smooth session. A pilot test was also conducted to ascertain the anticipated duration of the session.

A one-to-one, semistructured in-depth interview format was used since this study aims to investigate food insecurity, which is considered a sensitive issue because it is closely related to the fundamental aspects of health and well-being and carries emotional and social consequences. Additionally, a semistructured interview is useful in a mixed methods study, as it discovers participants' motivations, attitudes, beliefs, and the impact of specific events in their lives, in addition to supplementing and adding depth to the quantitative approach [45,46].

Prior to the interview session, the participants were approached by the researcher via WhatsApp. A message was sent to each participant, outlining the interview process and the expected duration of the session. If they agreed, the researcher inquired about their preferred date and time for the interview. A one-to-one interview session was subsequently conducted over the phone. In the case of unanswered messages, the researcher made three attempts to reach the participant over the phone during working days. If they were still unreachable, the participant was excluded from the interview list. The interview did not take long, as the guide contained only two main questions that needed to be answered by the participants. To ensure consistency and comparability of data collection, all interviews were conducted over the phone.

Before the interview began, consent from each participant was sought to audio-record the session using a digital voice recorder (TX660, Sony Corporation) as the main recorder and an audio-recording app (macOS, Apple Inc) as a backup recorder. Participants were informed that they could refuse to answer questions and withdraw from the study at any time. The interviews were conducted in Malay until data saturation was achieved, indicating the point when no new information was obtained from the participants and when subsequent interview transcripts yielded no new codes or themes [47]. This was assessed through ongoing comparative analysis using NVivo software, where redundancy in codes and conceptual categories was observed across interviews. Probing techniques, such as repeating the interviewee's statements, summarizing the main idea, and demonstrating engagement through verbal agreement, were used to encourage participants to persist in the conversation, enabling them to provide additional information that could be useful for understanding the topic [46,48]. In addition, the key points and main ideas highlighted during the interviews were noted to guide follow-up questions, ensure that all interview questions were answered, and assist in transcription [45].

Qualitative Analysis

Assisted by the field notes, a word-by-word, verbatim transcription of the audio-recorded interviews was manually composed by the researcher, while ensuring an accurate and detailed record of spoken dialogue. To ensure that the recorded audio was accurately transcribed, the researcher listened to it multiple times during the transcription process. Thematic analysis (systematically identifying, organizing, and delivering information about patterns of meaning [themes] in a dataset [49]) of the qualitative data was performed using NVivo software.

Trustworthiness was ensured during each phase of thematic analysis through several methods [50]. Credibility was enhanced through familiarity with data. The researcher familiarized themselves and thoroughly engaged with the data to ensure comprehension, given the data's extensive range and profound nature. Thus, before beginning coding, the complete dataset was thoroughly reviewed; this process may influence the development of ideas and the identification of potential patterns as researchers gain a comprehensive understanding of the data [51]. In addition, transferability was ensured through the inclusion of direct participant quotations, which will provide conceptual depth and allow readers to access the applicability of the findings to other settings. Initial codes or subthemes from the potentially relevant data in the transcript were generated, which will enable the researcher to simplify and concentrate on particular characteristics of the data.

An inductive approach was followed in defining themes and subthemes, where the codes arose from the data themselves and did not rely on any existing theories or frameworks [52]. Dependability was applied through iterative coding, where the development of codes continued; eventually, similar groups of coded data extracts were sorted into several main themes. The coded data extracts for each theme were then reviewed to determine whether they exhibited a cohesive pattern. Lastly, these themes were labeled and defined, and their significance and rationale were determined [51]. Confirmability was supported by refining themes to ensure they were grounded in the participants' narratives. To enhance analytical rigor, initial coding and theme development were independently reviewed by other members of the research team, and any discrepancies were discussed until consensus was achieved. Following the thematic analysis, the tabulated themes and supporting quotes were clarified with the research team to ensure accuracy, consistency, and credibility of the findings.

Results

The study was funded by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2023/SS10/UIAM/02/1) starting in September 2023., involving 287 participants, and data analysis has also been completed. As of January 2026, quantitative findings are under review and qualitative findings in preparation. A total of 287

XSL•FO
RenderX

participants were enrolled, which exceeded the minimum number required to achieve the targeted margin of error. For the quantitative phase of the study, data collection was conducted from December 2023 until August 2024. For the qualitative phase, 19 (6.6%) participants, drawn from the main study population (N=287), participated in the semistructured interviews. Data analysis was also completed. The manuscript reporting findings from the quantitative phase (prevalence and determinants of food security status and its association with dietary intake) was submitted for publication in December 2025 and is currently under review. Meanwhile, the manuscript reporting findings from the qualitative phase (contributing factors and coping strategies related to food insecurity) is currently being prepared for publication.

In addition, two questionnaires, the Malay version of the DSNLS [27] and the dietary adherence component of the modified ESRD-AQ [28], were translated and validated in this study to assess the participants' nutrition literacy level and dietary adherence, respectively.

Discussion

Summary

This study is expected to provide insight into the prevalence of food insecurity among patients undergoing hemodialysis in Pahang, Malaysia, and to examine the hypothesized determinants of food insecurity, related to demographic and socioeconomic characteristics, physical and economic access to food, disease-/treatment-related factors, and dietary factors. Specifically, the study will evaluate the relationship between food security status and nutritional status indicators, including anthropometric measurements, biochemical parameters, clinical assessments, and dietary intake. The qualitative component will further explore, through semistructured interviews, contributing factors and coping strategies used by patients with food insecurity undergoing hemodialysis.

Strengths and Limitations

The main strength of this study is that it is the first to assess food insecurity among patients undergoing hemodialysis in Malaysia. This study is also the first to explore coping strategies for food insecurity among patients undergoing hemodialysis. In terms of the study design, it used a mixed methods approach that ensures richness of data and provides a comprehensive understanding of data. The findings from the semistructured in-depth interviews are expected to complement the quantitative component of the study by adding nuance to the questionnaire data, providing a more profound understanding of the experiences of patients with food insecurity undergoing hemodialysis and thus strengthening the knowledge about this issue. Another strength is the use of a universal sampling method that eliminated sampling bias by including all patients undergoing hemodialysis at the selected dialysis centers who met the inclusion criteria, ensuring complete coverage of the population. However, it may not be generalizable to other settings and may lack external validity since the characteristics of the group being studied may not represent patients undergoing hemodialysis in other settings.

However, the protocol has a few limitations. Several variables included in the questionnaires were based on participants' self-reports, such as the duration of disease, dialysis vintage, frequency of dietary education, and adherence to diet, as well as dietary assessments, which require participants to recall the food consumed in the past 24 hours and the past few days (including weekends and the nondialysis day). This method may introduce recall bias if the participants do not accurately remember their past food intake. Over- or underestimation of food intake may also occur, particularly in participants' responses, either unintentionally or to seem compliant with dietary recommendations. However, the effect of reporting inaccuracies was minimized by using a statistical adjustment method that adjusted protein intake based on total energy intake, thereby reducing measurement errors. Additionally, using phone call interviews may have overlooked the participants' nonverbal communications, including body language, gestures, and facial expressions. This method would therefore limit additional information useful for the interviews. Hence, face-to-face settings may be advantageous for implementation in further research to improve participants' attention and offer more detailed responses during the interviews.

Conclusion

This protocol outlines a mixed methods study that will generate new evidence and valuable resources for understanding the prevalence and determinants of food insecurity among patients undergoing hemodialysis, its association with nutritional status, and the coping strategies used employed by individuals with food insecurity. The qualitative findings are anticipated to provide useful complementary insights derived from the participants' lived experiences, in addition to the quantitative data obtained through statistical analysis. Collectively, these findings are expected to inform clinical management by supporting early identification of nutritionally vulnerable patients undergoing hemodialysis and guiding targeted nutrition interventions. In addition, evidence generated may support policy-level initiatives aimed at integrating food security screening and assistance into renal care services. By addressing food insecurity through both clinical and policy perspectives, this study has the potential to contribute to improved nutritional care, health outcomes, and quality of life of patients undergoing hemodialysis.

Acknowledgments

The authors gratefully acknowledge the *Majlis Ugama Islam dan Adat Resam Melayu Pahang*, the operator of the hemodialysis centers involved in this study, as well as the staff and all study participants for their cooperation and valuable contributions. This study was funded by the Ministry of Higher Education, Malaysia, under the Fundamental Research Grant Scheme (FRGS/1/2023/SS10/UIAM/02/1).

XSL•FO
RenderX

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author (NAMS) upon reasonable request.

Authors' Contributions

WAMAB, RMY, NS, and NAMS contributed to conceptualization and methodology. SMK, WAMAB, RMY, and NAMS contributed to formal analysis and investigation. SMK contributed to writing the original draft. SMK and NAMS contributed to reviewing and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guidelines.
[PDF File (Adobe PDF File), 66 KB - resprot_v15i1e84575_app1.pdf]

Multimedia Appendix 2
Peer-review report from the Fundamental Research Grant Scheme (FRGS), Ministry of Higher Education, Malaysia.
[PDF File (Adobe PDF File), 271 KB - resprot_v15i1e84575_app2.pdf]

References

1. Food and Agriculture Organization of the United Nations (FAO), International Fund for Agricultural Development (IFAD), United Nations Children's Fund (UNICEF), World Food Programme (WFP), World Health Organization (WHO). The State of Food Security and Nutrition in the World 2024 - Financing to End Hunger, Food Insecurity and Malnutrition in All Its Forms. Rome, Italy: FAO; 2024.
2. Miguel EDS, Lopes SO, Araújo SP, Priore SE, Alfenas RDCG, Hermsdorff HHM. Association between food insecurity and cardiometabolic risk in adults and the elderly: a systematic review. J Glob Health 2020 Dec;10(2):020402 [FREE Full text] [doi: 10.7189/jogh.10.020402] [Medline: 33110569]
3. Food security update (May 16, 2025). The World Bank. 2025. URL: https://thedocs.worldbank.org/en/doc/40ebbf38f5a6b68bfc11e5273e1405d4-0090012022/related/Food-Security-Update-116-May-16-2025.pdf [accessed 2025-07-03]
4. Banerjee T, Crews DC, Wesson DE, Dharmarajan S, Saran R, Ríos Burrows N, CDC CKD Surveillance Team. Food insecurity, CKD, and subsequent ESRD in US adults. Am J Kidney Dis 2017 Jul;70(1):38-47 [FREE Full text] [doi: 10.1053/j.ajkd.2016.10.035] [Medline: 28215947]
5. Puchulu MB, Garcia-Fernandez N, Landry MJ. Food insecurity and chronic kidney disease: considerations for practitioners. J Ren Nutr 2023 Sep;33(5):691-697 [FREE Full text] [doi: 10.1053/j.jrn.2023.06.001] [Medline: 37331455]
6. Campbell CC. Food insecurity: a nutritional outcome or a predictor variable? J Nutr 1991 Mar;121(3):408-415 [FREE Full text] [doi: 10.1093/jn/121.3.408] [Medline: 2002411]
7. Taylor KS, Novick TK, Santos SR, Chen Y, Smith OW, Perrin NA, et al. Material need insecurities among people on hemodialysis: burden, sociodemographic risk factors, and associations with substance use. Kidney360 2023 Nov 9;4(11):1590-1597 [FREE Full text] [doi: 10.34067/kid.0000000000000279]
8. Vargas-Vázquez C, González-Ortíz A, Bertrán-Vilà M, Espinosa-Cuevas A. Impact of SARS-CoV-2 pandemic on food security in patients with chronic kidney disease. J Ren Nutr 2023 Jan;33(1):78-87 [FREE Full text] [doi: 10.1053/j.jrn.2022.07.004] [Medline: 35863603]
9. Zhu M, Arinze N, Buitron de la Vega P, Alonso A, Levin S, Farber A, et al. High prevalence of adverse social determinants of health in dialysis access creation patients in a safety-net setting. Ann Vasc Surg 2024 Mar;100:31-38 [FREE Full text] [doi: 10.1016/j.avsg.2023.10.026] [Medline: 38110081]
10. Perez L, Gendelman S, Kendrick J. Home-delivered medically tailored meal engagement among hemodialysis patients and providers. J Ren Nutr 2024 Jan;34(1):40-46 [FREE Full text] [doi: 10.1053/j.jrn.2023.08.011] [Medline: 37640277]
11. Ikizler TA, Burrowes JD, Byham-Gray LD, Campbell KL, Carrero J, Chan W, et al. KDOQI clinical practice guideline for nutrition in CKD: 2020 update. Am J Kidney Dis 2020 Sep;76(3 Suppl 1):S1-S107 [FREE Full text] [doi: 10.1053/j.ajkd.2020.05.006] [Medline: 32829751]
12. Demu A, Tamire A, Baraki N, Negash A, Dechasa M, Dereje J, et al. Magnitude, determinants, and coping strategies of food insecurity among people living with HIV/AIDS in Eastern Ethiopia. AIDS Res Treat 2025 Apr 19;2025(1):9970515 [FREE Full text] [doi: 10.1155/arat/9970515] [Medline: 40291812]

XSL•FO
RenderX

13. McDougall JA, Jaffe SA, Guest DD, Sussman AL. The balance between food and medical care: experiences of food insecurity among cancer survivors and informal caregivers. J Hunger Environ Nutr 2022 Feb 25;17(3):380-396 [FREE Full text] [doi: 10.1080/19320248.2021.1892295] [Medline: 35757157]
14. Marjerrison S, Cummings EA, Glanville NT, Kirk SF, Ledwell M. Prevalance and associations of food insecurity in children with diabetes mellitus. J Pediatr 2011 Apr;158(4):607-611 [FREE Full text] [doi: 10.1016/j.jpeds.2010.10.003] [Medline: 21126743]
15. Mayer VL, McDonough K, Seligman H, Mitra N, Long JA. Food insecurity, coping strategies and glucose control in low-income patients with diabetes. Public Health Nutr 2015 Sep 02;19(6):1103-1111 [FREE Full text] [doi: 10.1017/s1368980015002323]
16. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2nd ed). Thousand Oaks, CA: Sage; 2003:0761924418.
17. Sulaiman N, Shariff ZM, Jalil RA, Taib MNM, Kandiah M, Samah AA. Validation of the Malaysian coping strategy instrument to measure household food insecurity in Kelantan, Malaysia. Food Nutr Bull 2011 Dec;32(4):354-364 [FREE Full text] [doi: 10.1177/156482651103200407] [Medline: 22590969]
18. Russell J, Flood V, Yeatman H, Mitchell P. Prevalence and risk factors of food insecurity among a cohort of older Australians. J Nutr Health Aging 2014 Jan;18(1):3-8 [FREE Full text] [doi: 10.1007/s12603-013-0339-6] [Medline: 24402381]
19. Gutiérrez OM. Contextual poverty, nutrition, and chronic kidney disease. Adv Chronic Kidney Dis 2015 Jan;22(1):31-38 [FREE Full text] [doi: 10.1053/j.ackd.2014.05.005] [Medline: 25573510]
20. Roselawati MY. The Translation and Validation of Malay Food Insecurity Experience Scale (FIES) and Determinants of Food Insecurity in Kuantan, Pahangdoctoral dissertation. Kuantan, Pahang: International Islamic University Malaysia; 2022.
21. Lee JE, Kim H, Lee MJ, Kwon YE, Kyung M, Park J, et al. Comparison of dietary intake patterns in hemodialysis patients by nutritional status: a cross-sectional analysis. Kidney Res Clin Pract 2020 Jun 30;39(2):202-212 [FREE Full text] [doi: 10.23876/j.krcp.20.037] [Medline: 32576714]
22. Sualeheen A, Khor B, Balasubramanian GV, Sahathevan S, Ali MSM, Narayanan SS, PaTCH Investigators (Malaysia). Habitual dietary patterns of patients on hemodialysis indicate nutritional risk. J Ren Nutr 2020 Jul;30(4):322-332 [FREE Full text] [doi: 10.1053/j.jrn.2019.09.010] [Medline: 31767516]
23. Lim J, Chinna K, Khosla P, Karupaiah T, Daud ZAM. Understanding how nutrition literacy links to dietary adherence in patients undergoing maintenance hemodialysis: a theoretical exploration using partial least squares structural equation modeling. Int J Environ Res Public Health 2020 Oct 14;17(20):7479 [FREE Full text] [doi: 10.3390/ijerph17207479] [Medline: 33066603]
24. Mat Ya R, Mohamed Nor N, Jaafar NH, Sidek S, Ab Rahman J, Sulaiman N, et al. Validation of the Malay version of Food Insecurity Experience Scale (M-FIES) using Rasch analysis. Mal J Nutr 2023 Dec 21;29(3):529-537 [FREE Full text] [doi: 10.31246/mjn-2022-0031]
25. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Pers Assess 1988 Mar;52(1):30-41 [FREE Full text] [doi: 10.1207/s15327752jpa5201_2]
26. Ng C, Amer Siddiq A, Aida S, Zainal N, Koh O. Validation of the Malay version of the Multidimensional Scale of Perceived Social Support (MSPSS-M) among a group of medical students in Faculty of Medicine, University Malaya. Asian J Psychiatr 2010 Mar;3(1):3-6 [FREE Full text] [doi: 10.1016/j.ajp.2009.12.001] [Medline: 23051129]
27. Karami SM, Mohd Abu Bakar WA, Mat Ya R, Sulaiman N, Mohd Shukri NA. The Malay version of Dialysis Specific Nutrition Literacy Scale (DSNLS): translation and validation. Asian J Med Health Sci 2024 Nov;7(2):175-183 [FREE Full text] [doi: 10.70672/4zsbgm58]
28. Karami SM, Mohd Abu Bakar WA, Mat Ya R, Sulaiman N, Mohd Shukri NA. Dietary adherence component of End-Stage Renal Disease Adherence Questionnaire (ESRD-AQ): translation and validation. Int J Allied Health Sci 2024 Feb 14;8(1):2997-3004 [FREE Full text] [doi: 10.31436/ijahs.v8i1.889]
29. Body mass index (BMI). Global Health Observatory. World Health Organization. URL: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/body-mass-index [accessed 2025-09-01]
30. Kadhum I, Mohammed W. Nutritional status of adult hemodialysis patients in Al-Najaf Al-Ashraf Governorate. INJNS 2012;25(1):64-78 [FREE Full text] [doi: 10.58897/injns.v25i1.119]
31. Gallagher D, Heymsfield SB, Heo M, Jebb SA, Murgatroyd PR, Sakamoto Y. Healthy percentage body fat ranges: an approach for developing guidelines based on body mass index. Am J Clin Nutr 2000 Sep;72(3):694-701 [FREE Full text] [doi: 10.1093/ajcn/72.3.694] [Medline: 10966886]
32. Kalantar-Zadeh K, Kopple JD, Block G, Humphreys MH. A malnutrition-inflammation score is correlated with morbidity and mortality in maintenance hemodialysis patients. Am J Kidney Dis 2001 Dec;38(6):1251-1263 [FREE Full text] [doi: 10.1053/ajkd.2001.29222] [Medline: 11728958]
33. Chan YM, Shariff ZM, Thiam C, Lim S, Goh BL. Factors associated with poor nutritional status among hemodialysis patients in Malaysia. Mal J Med Health Sci 2019 Apr;15(SP1):77-83 [FREE Full text]

XSL•FO
RenderX

34. Fouque D, Kalantar-Zadeh K, Kopple J, Cano N, Chauveau P, Cuppari L, et al. A proposed nomenclature and diagnostic criteria for protein-energy wasting in acute and chronic kidney disease. Kidney Int 2008 Feb;73(4):391-398 [FREE Full text] [doi: 10.1038/sj.ki.5002585] [Medline: 18094682]
35. Sahathevan S, Se CH, Ng SH, Chinna K, Harvinder GS, Chee WSS, et al. Assessing protein energy wasting in a Malaysian haemodialysis population using self-reported appetite rating: a cross-sectional study. BMC Nephrol 2015 Jul 07;16:99 [FREE Full text] [doi: 10.1186/s12882-015-0073-x] [Medline: 26149396]
36. Rachida M, Mohamed MF, Mohammed EA, Rekia B. Assessment of nutritional status, dietary intake and adherence to dietary recommendations in hemodialysis patients. GSC Adv Res Rev 2020 May 30;3(2):009-019. [doi: 10.30574/gscarr.2020.3.2.0030]
37. Willett W, Stampfer MJ. Total energy intake: implications for epidemiologic analyses. Am J Epidemiol 1986 Jul;124(1):17-27 [FREE Full text] [doi: 10.1093/oxfordjournals.aje.a114366] [Medline: 3521261]
38. Saglimbene VM, Su G, Wong G, Natale P, Ruospo M, Palmer SC, et al. Dietary intake in adults on hemodialysis compared with guideline recommendations. J Nephrol 2021 Dec 16;34(6):1999-2007 [FREE Full text] [doi: 10.1007/s40620-020-00962-3] [Medline: 33591554]
39. Zimmerer JL, Leon JB, Covinsky KE, Desai U, Sehgal AR. Diet monotony as a correlate of poor nutritional intake among hemodialysis patients. J Ren Nutr 2003 Apr;13(2):72-77 [FREE Full text] [doi: 10.1053/jren.2003.50025] [Medline: 12671828]
40. Khor B, Sualeheen A, Sahathevan S, Chinna K, Gafor AHA, Bavanandan S, PaTCH Investigators. Association of dietary patterns with serum phosphorus in maintenance haemodialysis patients: a cross-sectional study. Sci Rep 2020 Jul 23;10(1):12278 [FREE Full text] [doi: 10.1038/s41598-020-68893-4] [Medline: 32704087]
41. Stoltzfus JC. Logistic regression: a brief primer. Acad Emerg Med 2011 Oct;18(10):1099-1104 [FREE Full text] [doi: 10.1111/j.1553-2712.2011.01185.x] [Medline: 21996075]
42. Malhotra R. Errors in the use of multivariable logistic regression analysis: an empirical analysis. Indian J Community Med 2020;45(4):560. [doi: 10.4103/ijcm.ijcm_16_20]
43. Janssens ACJW, Martens FK. Reflection on modern methods: revisiting the area under the ROC curve. Int J Epidemiol 2020 Aug 01;49(4):1397-1403. [doi: 10.1093/ije/dyz274] [Medline: 31967640]
44. Norhasmah S, Zalilah MS, Mohd Nasir MT, Kandiah M, Asnarulkhadi AS. A qualitative study on coping strategies among women from food insecurity households in Selangor and Negeri Sembilan. Malays J Nutr 2010 Apr;16(1):39-54. [Medline: 22691852]
45. Adams WC. Conducting semi-structured interviews. In: Wholey JS, Harty HP, Newcomer KE, editors. Handbook of Practical Program Evaluation. San Francisco, CA: Jossey-Bass; 2015:492-505.
46. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health 2019;7(2):e000057 [FREE Full text] [doi: 10.1136/fmch-2018-000057] [Medline: 32148704]
47. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
48. Whiting LS. Semi-structured interviews: guidance for novice researchers. Nurs Stand 2008 Feb 13;22(23):35-40 [FREE Full text] [doi: 10.7748/ns2008.02.22.23.35.c6420] [Medline: 18323051]
49. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology, Vol. 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Washington, DC: American Psychological Association; 2012:57-71.
50. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017 Oct 02;16(1):1-13 [FREE Full text] [doi: 10.1177/1609406917733847]
51. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
52. Themelis C, Sime J, Thornberg R. Informed grounded theory: a symbiosis of philosophy, methodology, and art. Scand J Educ Res 2022 Sep 12;67(7):1086-1099 [FREE Full text] [doi: 10.1080/00313831.2022.2115135]

Abbreviations

CKD: chronic kidney disease
DMI: diet monotony index
DSNLS: Dialysis-Specific Nutrition Literacy Scale
ESRD: end-stage renal disease
ESRD-AQ: End-Stage Renal Disease Adherence Questionnaire
IBW: ideal body weight
MAMC: midarm muscle circumference
M-FIES: Malay Food Insecurity Experience Scale
MIS: Malnutrition Inflammation Score

XSL•FO
RenderX

MUAC: mid-upper arm circumference
MUIP: Majlis Ugama Islam dan Adat Resam Melayu Pahang
PEW: protein energy wasting
TIBC: total iron-binding capacity
TSF: triceps skinfold

Edited by A Schwartz; submitted 22.Sep.2025; this is a non–peer-reviewed article;accepted 31.Dec.2025; published 19.Jan.2026.

Please cite as:
Mohd Shukri NA, Karami SM, Mohd Abu Bakar WA, Mat Ya R, Sulaiman N
Food Insecurity Determinants, Coping Strategies, and Association With Nutritional Status Among Hemodialysis Patients in Pahang, Malaysia: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e84575
URL: https://www.researchprotocols.org/2026/1/e84575
doi:10.2196/84575
PMID:

©Nor Azwani Mohd Shukri, Sarah Muneera Karami, Wan Azdie Mohd Abu Bakar, Roselawati Mat Ya, Norhasmah Sulaiman. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Environmental Exposure and Mental Health in Hong Kong: Protocol for a GPS- and Biosensor-Based Observational Study

Yanjia Cao[1], PhD; Tianyu Li[1], MS; Calvin P Tribby[2], PhD; Dorita H F Chang[3], PhD

[1]Department of Geography, The University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Department of Population Health, City of Hope National Comprehensive Cancer Center, Duarte, CA, United States

[3]Department of Psychology, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Yanjia Cao, PhD
Department of Geography, The University of Hong Kong
Rm 1017, The Jockey Club Tower, Pok Fu Lam
Hong Kong, 00000
China (Hong Kong)
Phone: 852 39177451
Email: yanjiac@hku.hk

Abstract

Background: Environmental exposure, including green-blue space and walkability, may influence mental health through various physiological pathways. Traditional methods have limitations in capturing dynamic environmental exposure effects on mental health.

Objective: This study aims to investigate the associations between minute-level environmental exposure and stress-related biomarkers using GPS-based measurements and wrist-worn biosensors in an urban population.

Methods: In this prospective observational study, we will recruit 750 participants (aged 18-80 years) in Hong Kong from September 2024 to December 2026. This includes a pilot study supported by internal funds starting in September 2024 and external funding from the Health and Medical Research Fund by the Hong Kong Health Bureau in July 2025. Participants will undergo a 1-week experiment wearing biosensors and carrying GPS devices while completing daily surveys and entry and exit questionnaires. Primary outcomes include digital biomarkers (electrodermal activity, blood volume pulse, and skin temperature) from wrist-worn sensors and associations with the environmental exposures of air pollution, urban density, and green-blue space access. Secondary outcomes comprise questionnaire responses, physical activity metrics from accelerometer data, and environmental exposure assessments, including mobility barriers, land use mix, road accessibility, and thermal conditions. Focus groups will be conducted to refine study protocol and assess device-wearing experience.

Results: The study received ethical approval from the Human Ethic Research Committee (award numbers: EA240125 and EA240514). As of June 30, 2025, a total of 150 participants completed the pilot study (September 2024 onwards), achieving a 94% (141/150) completion rate with strong device compliance (141/150, 94% valid biosensor days; 137/150, 91% daily questionnaires). Protocol refinements based on participant feedback were implemented without altering core instruments. The main study supported by the Health and Medical Research Fund, commenced in July 2025, with 30 participants enrolled as of September 2025. Data collection for 750 participants will continue until December 2026. We are at an early stage of data cleaning and preprocessing for the pilot study. Data analysis is expected to be completed by June 2027, with results anticipated to be published in 2027 and 2028.

Conclusions: This study's innovative integration of continuous environmental monitoring and stress-related biomarker may provide new insights into environment–health relationships. Findings will inform public health initiatives for mental well-being promotion and urban planning interventions.

International Registered Report Identifier (IRRID): DERR1-10.2196/84919

(JMIR Res Protoc 2026;15:e84919) doi:10.2196/84919

KEYWORDS

digital biomarker; dynamic environmental exposure; electrodermal activity; mental health outcome; physical activity; stress response; wrist-worn biosensor

XSL•FO
RenderX

Introduction

The World Health Organization reported that more than one billion people worldwide experience mental health issues affecting daily lives [1]. High-rise residential buildings, particularly clustered in Hong Kong, were reported to increase the burdens of an already stressful life and threaten the well-being of residents living on the upper floors [2]. Population health surveillance data from the Hong Kong Department of Health for the years 2020 to 2022 indicated that approximately 1.2% of the surveyed population had been diagnosed with depression, while 0.9% of the respondents reported anxiety [3]. These rates were double the numbers reported from 2014 to 2015. Based on this, depression and anxiety are recognized as the most common mental health issues in Hong Kong. Furthermore, the prevalence of severe depression among Hong Kong residents increased by more than 30% between 2020 and 2023 [3]. Mental health problems present a younger trend, as a territory-wide study in 2015 showed 11.3% of children and adolescents in Hong Kong experienced common mental disorders [4]. A separate study reported an increased prevalence of 16.6% between 2019 and 2022, during social unrest and the COVID-19 pandemic [5]. However, very few people are likely to seek professional help. It has been estimated that 74% of mental health sufferers do not seek any form of professional assistance in Hong Kong [4].

Prior research has reported that the status of mental health is associated with genetic factors and wider social determinants encompassing the life span [6,7]. These determinants include socioeconomic conditions, intergenerational relations, living community social and built environment, and the affordability and accessibility of medical care [8,9]. The environment shapes many aspects of daily life. Evidence shows embracing the natural environment, especially exposure to green and blue space, contributes to improving human health [10,11]. However, existing research about environmental exposure related to the status of mental health usually focuses on the residential neighborhoods, leading to a “stationary bias” [12], while dynamic exposure assessments that concentrate on people’s mobility patterns (ie, changes in locations over time) [13] are essential to estimating total environmental exposures [14].

Traditional measurement using questionnaire surveys has been found time- and labor-consuming [15,16], and self-reported instruments are subject to individual characteristics and preferences [17,18]. Recent studies have increasingly adopted wearable devices, allowing for continuous monitoring of physiological parameters and real-time assessment of physical and mental health conditions [19-22].

In this study, we aim to investigate the association between changes in environmental exposure, including green and blue space and walkability, and changes in biomarkers associated with mental health by using minute-level GPS-based environmental measurements and wrist-worn biosensors (Figure 1). Primary analyses will examine associations between 3 environmental exposures (air pollution, urban density, and green-blue space access) and 3 digital biomarkers (electrodermal activity, pulse rate, and skin temperature). Secondary analyses will investigate physical activity metrics and explore questionnaire responses and additional environmental assessments, including walkability, mobility barriers, land use mix, road accessibility, and thermal conditions.

Figure 1. Conceptual framework of the study. GIS: geographic information system; NDVI: normalized difference vegetation index; MET: metabolic equivalent of task.

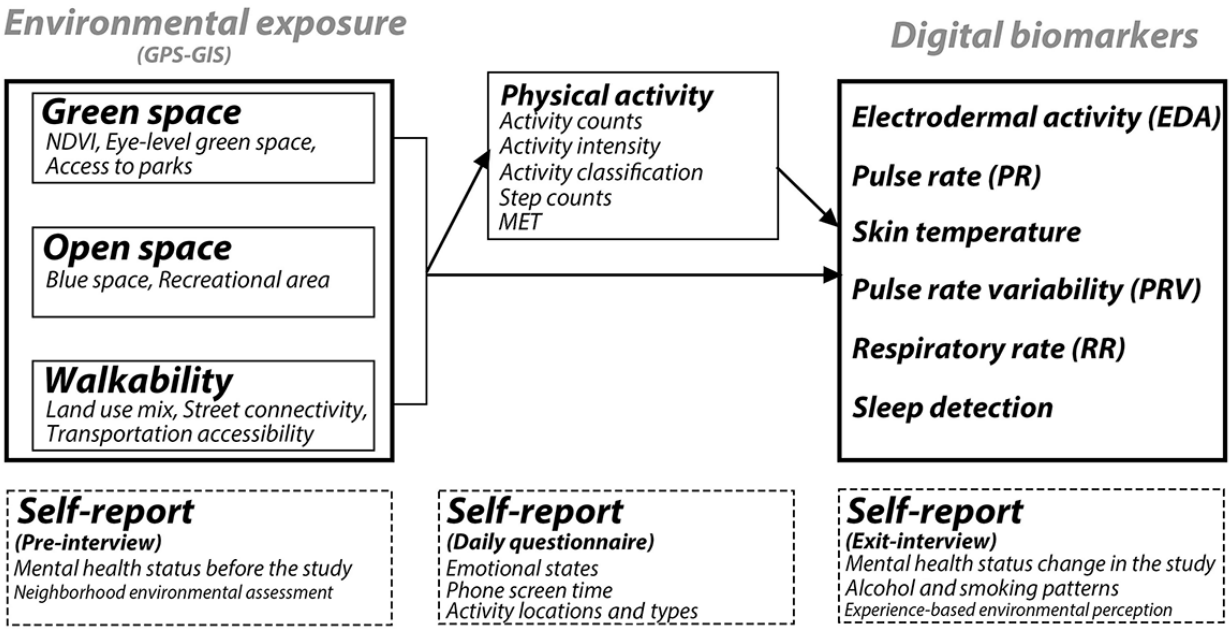

We hypothesize that environmental exposure–biomarker relationships exhibit spatial and temporal heterogeneity across activity space, characterized by both acute physiological responses and cumulative effects. Physical activities and daily

XSL•FO
RenderX

behavior are expected to mediate these relationships, through activity types and intensities. Moreover, we propose that the stress responses to environmental features will demonstrate individual-specific patterns, reflecting the complex interaction between personal characteristics and environmental factors. This protocol addresses current methodological gaps by implementing dynamic geospatial exposure assessment combined with continuous monitoring of digital biomarkers.

Methods

Study Design and Objectives

This study adopts a mixed methods longitudinal design to investigate the pathways through which environmental exposures influence mental health by physical activity mechanisms in Hong Kong adults.

Primary outcomes include digital biomarkers (electrodermal activity, pulse rate, and skin temperature) from wrist-worn sensors and their associations with the 3 environmental exposures of air pollution, urban density, and green-blue space access. Secondary outcomes comprise questionnaire responses, physical activity metrics from accelerometer data, and environmental exposure assessments, including mobility barriers, land use mix, road accessibility, and thermal conditions.

To address concerns regarding multiple testing, we will report the primary associations between environmental exposures and biomarkers defined above with α=.01 to minimize false positives (3 primary outcomes × 3 primary exposure measures = 9 comparisons). For secondary comparisons, we set α=.05, recognizing that this may include some false positives. However, these are exploratory analyses, and significant associations would be interpreted as suggestive to be examined with further study.

Participants' Recruitment

This study adopts a mixed methods approach, combining quantitative and qualitative methods with a cross-disciplinary approach. For sampling strategy, we apply a predefined approach by integrating stratified random sampling with multistage sampling methodologies. Initially, we categorize small territorial planning units (STPUs) in Hong Kong based on two indicators: (1) percentage of older adults (aged ≥65 years) using Hong Kong Census data: this serves as a proxy for age-specific demographic composition in each STPU, indicating areas with aging population densities and their corresponding service demands (eg, health care facilities and community support infrastructure) [23]; and (2) proximity to the nearest green space using OpenStreetMap: this indicator calculates the walking time from an STPU centroid to the nearest urban or country park, representing a crucial environmental factor for physical activities and recreation across all age groups. We further categorize STPUs into low, median, and high groups using terciles for each indicator: (1) Percentage of older adults: low (3.58%-15.78%), medium (15.78%-20.79%), and high (20.79%-65.76%); and (2) proximity to nearest park: low (8.72-56.00 minutes), medium (3.86-8.71 minutes), and high (0.10-3.85 minutes). Finally, we select STPUs based on indicator combination (percentage of older adults–proximity to green space in minutes): high–high (n=32), high–low (n=14), low–high (n=11), and low–low (n=40).

The target sample size of 750 is based on the current Hong Kong population size of 7.4 million. The sample size represents approximately 1 out of 10,000 of the population, which is a common patient representation in a mid-sized region [24,25].

The recruitment strategy includes community engagement through promotional materials in community centers and snowball recruitment through initial participants. Promotional materials include a QR code for participants to provide basic information (ie, contact, age, gender, and residential STPU). During enrollment, we perform prescreening and monitor recruitment rates across STPU strata weekly to ensure balanced distribution across demographics and geographic areas. To mitigate sampling bias toward healthier or smartphone-familiar adults, we have (1) increased targeted outreach in underrepresented areas, (2) provided device-wearing tutorials in multiple languages and formats to accommodate varying digital literacy levels, and (3) provided backup smartphones for participants without compatible devices.

Participant Eligibility

Study participants include adults aged 18-80 years, of any ethnicity or race, who possess a valid Hong Kong identity card and have been living locally for at least 6 months. This age range is selected to provide increased exposure variability while ensuring participants maintain sufficient mobility. Adults aged >80 years are not included because of potential mobility limitations.

To be eligible, participants must be able to walk without human assistance, travel to a study visit, have a mobile phone, read and write fluently in Mandarin, Cantonese, or English, provide informed consent and comply with the protocol, and be willing and able to complete all assessments. Ineligibility criteria include (1) pregnant or nursing, (2) neuropsychological disorders or neurological illnesses that affect understanding or compliance with study protocols, (3) medical conditions that significantly limit outdoor activities or daily mobility, (4) skin conditions that may be irritated by wearing sensors, (5) inability to maintain regular device charging routines, and (6) plans to travel outside Hong Kong during the 7-day study period, or (6) conditions that prevent smartphone use for daily questionnaires.

During the enrollment process, we perform prescreening on the residential location, gender, and age groups. The number of participants from each district and the proportions of different demographic characteristics (gender and age groups) will be regularly reviewed to ensure balanced distribution and avoid oversampling from specific groups or areas.

Pilot Data Collection Protocol Adherence

We defined feasibility thresholds to assess protocol adherence. Study completion is defined as attending both study visits (preinterview and exit interview) and providing data during the 1-week experiment period without withdrawing. We set a threshold of ≥92% completion rate. Among participants who complete the study, we set secondary compliance thresholds: (1) ≥90% of participants provide ≥5 valid biosensor days

XSL•FO
RenderX

(defined as ≥18 hours/day wear time); and (2) ≥90% of participants complete ≥5 of 7 daily questionnaires. Participants who complete study visits but fail to meet these minimum data requirements will be provided 1 additional week to collect data.

Study Procedure

The overview of the data collection process is shown in Figure 2. Each participant is first invited to attend a mandatory prestudy session. During this session, participants are required to sign an informed consent form, which articulates the purpose of the study, the data collection procedure, and data confidentiality protocols. Subsequently, members of the research team provide device training for the wrist-worn biosensor (Empatica Embrace Plus) and GPS tracker (Qstarz BL 1000ST) carried in a bag or pocket. Training includes hands-on practice with proper device use, charging procedures, and a review of common issues. Participants should demonstrate competency in device operation before leaving the session. Following device training, participants are asked to complete a Short Form-12 (SF-12) Health Survey assessing their mental health over the past 4 weeks, along with their environmental perceptions and sociodemographic information. Upon completion, participants receive initial incentives of HKD $100 (US $12).

Figure 2. Overview of data collection process.

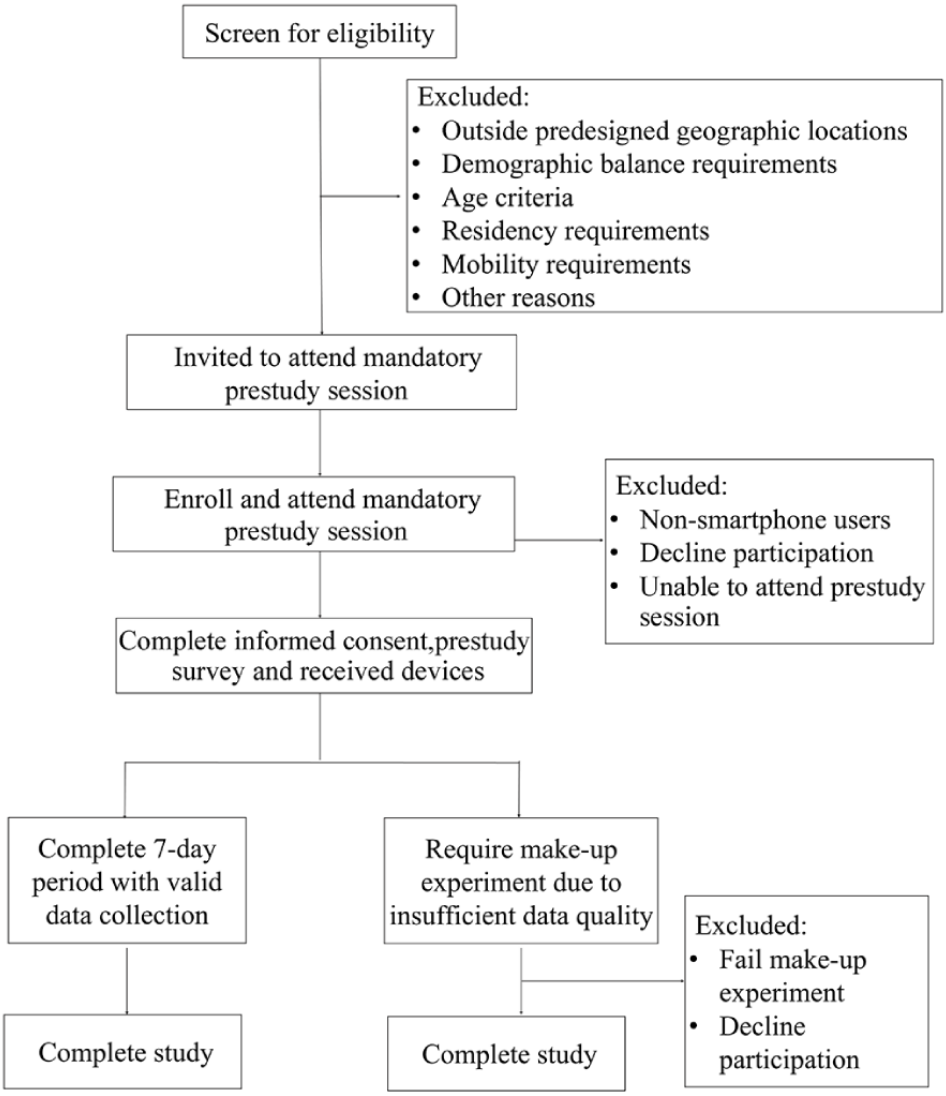

During the 7-day experimental period, participants are required to wear the biosensor for a minimum of 18 hours daily and carry the GPS tracker during all outdoor activities for at least 1 hour per day. This duration is based on prior theory indicating that valid accelerometer data requires a minimum of 3 days with at least 10 hours of daily wear time [26]. Wear time is defined as the period when the device is properly worn and collecting valid data. The biosensor uses on-skin detection algorithms to automatically identify and filter the data recorded while not being used as intended, reducing the likelihood of including unreliable observations [27]. After each day of activities, participants are asked to complete a questionnaire. Daily reminder text messages are sent to ensure questionnaire completion. In case of technical issues or device-related problems, participants are advised to contact the research team by phone for immediate assistance, including device charging support.

Upon completion of the 7-day experiment, participants are asked to complete an exit interview including: (1) the SF-12

XSL•FO
RenderX

questionnaire to assess their mental health during the study week, (2) an exit survey covering study feedback and retrospective behavior, and (3) a questionnaire assessing their perceived built environment based on their daily activity spaces. Participants who can successfully provide at least 5 days of data from both biosensor and GPS devices could receive full incentives of HKD $400 (US $50). If data quality does not meet the criteria, a one-time make-up experiment is offered.

Digital Biomarkers Associated With Mental Health

Participants' mental health is measured by a wrist-worn biosensor, a physiological monitoring device that captures an individual's motion-based activity, blood volume pulse (BVP), and skin temperature and links them to physiological signals (Table 1). Skin conductance is measured in microsiemens (μS) and ranges from 2-100 μS. The sampling rate is 4 Hz. Electrodermal activity (EDA) explorer will be used to detect peaks in 5-second periods of the clean electrodermal 5-minute signal. The biosensor also assesses acceleration and rotation (actigraphy) and BVP by photoplethysmography (PPG) at sampling rates of 64 Hz, as well as peripheral (skin) temperature at sampling rates of 1 Hz.

Table 1. Measurements for raw sensor data used in this study and their relationships with mental health.

High-frequency sensor data	Measurement method	Sampling rate	Pathway to mental health outcomes	Connection to mental health outcomes
EDA[a]	Skin conductance	4 Hz	Sympathetic nervous system activation	Higher EDA indicates increased stress and emotional arousal
BVP[b]	PPG[c]	64 Hz	Cardiovascular regulation	Irregular patterns reflect anxiety and stress states
Skin Temperature	Infrared sensor	1 Hz	Peripheral vasoconstriction	Decreased peripheral temperature indicates sympathetic nervous system activation during stress
Activity (actigraphy)	3-axis accelerometer	64 Hz	Physical activity patterns	Regular physical activity promotes positive mental health through various physiological and psychological pathways

[a]EDA: electrodermal activity.

[b]BVP: blood volume pulse.

[c]PPG: photoplethysmography.

In subsequent analysis, these raw sensor data will be processed into minute-level digital biomarkers through Empatica's proprietary algorithms, which have been validated through prospective clinical studies sponsored by Empatica [28,29]. The digital biomarkers are aggregated into 1-minute windows, including EDA (0.01-100 μS), pulse rate (24-240 beats per minute [bpm]), skin temperature (28 °C-35 °C) as primary outcomes, as well as pulse rate variability (PRV, root-mean-square of successive difference [RMSSD] up to 300 ms), and respiratory rate (6-60 breaths per minute) as secondary outcomes. Sleep detection is also monitored, with periods categorized as awake, sleep, or interruption. All these digital biomarkers will be computed and aligned at 1-minute intervals using specific aggregation methods for each indicator.

Individual baseline values are measured in two phases. First, during the preinterview session, participants wear the biosensor device while completing preinterview survey, which typically takes more than 10 minutes in a quiet environment. These measurements during the survey completion period serve as their initial baseline. Second, during the study period, we identify the daily baseline for each participant by selecting the 15-minute continuous period when their EDA levels are at their lowest, given that EDA is a primary indicator of sympathetic nervous system activation and stress response [30]. During these periods, we also record corresponding pulse rate, respiratory rate, and skin temperature values. These daily baseline values will be averaged across multiple days to establish a personalized baseline.

Physical Activity and Mobility Measurements

Physical activities are measured using accelerometer data from both biosensor and GPS devices. The biosensor device captures raw accelerometry data through its integrated 3-axis accelerometer at 64 Hz sampling rate. Using proprietary machine learning algorithms, the accelerometry signals are processed into minute-level digital biomarkers, including activity counts (0-200 counts), activity intensity levels (categorized as sedentary, low, moderate, or vigorous), movement classification (still, generic, walking, or running), step counts (up to 192 steps), and metabolic equivalent of task (MET, 0.95-8.53).

Raw GPS data are collected with 1-second sampling intervals. Table 2 listed each measurement, sample resolution, and aggregation process. The GPS device logs geographic coordinates, distance, speed, elevation, and time. In subsequent analysis, GPS data (1-second intervals) will be aligned with the biosensor data at 1-minute intervals. Within each 1-minute window, all GPS location points will be retained rather than aggregating to a single point. This approach maintains detailed movement patterns and allows precise assessment of environmental exposures.

XSL•FO
RenderX

Table 2. Measurement of digital biomarkers and location indicators at 1-minute resolution.

Indicator	Raw sampling	Aggregation method	Output interval	Range	Resolution
Activity counts	64 Hz (accelerometer)	Zeros-crossing count in filtered data	1 minute	0-200 counts	1 count
Activity intensity	Derived from MET[a]	MET threshold categorization	1 minute	4 categories (sedentary, low physical activity, moderate physical activity, or vigorous physical activity)	Categorical
Activity classification	64 Hz (accelerometer)	3-second nonoverlapping windows aggregation	1 minute	4 categories (still, generic, walking, or running)	Categorical
Step counts	64 Hz (accelerometer)	Step detection from accelerometer data	1 minute	0-192 steps	1 step
EDA[b]	4 Hz	Average over window	1 minute	0.01-100 µS	~55 pS
MET	64 Hz (accelerometer)	Energy expenditure calculation	1 minute	0.95-8.53 MET	—[c]
Pulse rate	64 Hz (PPG[d])	Aggregation of 10-second windows	1 minute	24-240 bpm[e]	1 bpm
PRV[f] (RMSSD[g])	64 Hz (PPG)	RMSSD calculation from interbeat intervals	1 minute	0-300+ ms	0.000001 ms
Respiratory rate	64 Hz (PPG)	Combining 30-second windows	1 minute	6-60 brpm[h]	1 brpm
Skin temperature	1 Hz	Average over window	1 minute	0 °C-85 °C	0.01 °C
Sleep detection	Derived from activity counts	Activity analysis for sleep/wake periods	1 minute	0 (awake), 101 (sleep), 102 (wake), 300 (interruption)	Categorical
GPS location data	1 second	Total area covered by multiradius buffers around sequential points within 1-minute window	1 minute	0-1 km²	0.0001 km²

[a]MET: metabolic equivalent of task.

[b]EDA: electrodermal activity.

[c]Not applicable/Not available.

[d]PPG: photoplethysmography.

[e]bpm: beats per minute.

[f]PRV: pulse rate variability.

[g]RMSSD: root-mean-square of successive difference.

[h]brpm: breaths per minute.

Environmental Exposure Measurements

We will measure the built environment from multiple perspectives. Primary environmental exposures include air pollution (particulate matter ≤2.5 µm [PM2.5], nitrogen dioxide [NO_2], and air quality index [AQI]), urban density (ground coverage ratio [GCR] and street coverage ratio [SCR]), and green-blue space access (normalized difference vegetation index [NDVI] and eye-level exposure metrics). Secondary environmental assessments comprise mobility barriers, land use mix, road accessibility, and thermal environment.

Air pollution exposure will be evaluated by PM2.5, NO_2, and AQI from Hong Kong's 18 air quality monitoring stations operated by the Environmental Protection Department. Kriging or inverse distance weighting methods will be applied to estimate air pollution levels across the study area. Urban density will be quantified using GCR and SCR for each trajectory buffer area. GCR represents the ratio of total building footprint to buffer area, while SCR indicates the percentage of vehicular street coverage. Green-blue space exposure will be quantified using multiple methods. Greenspace availability will be measured by NDVI from Sentinel-2 satellite imagery (10-meter resolution). Eye-level exposure to green space, recreational areas, blue spaces, and public spaces will be evaluated using Google Maps Street view images [31,32], which will be newly acquired during the analytical process in 2026.

For secondary environmental assessments, mobility barriers will be assessed by analyzing street intersection density, obstacles, slopes, surface conditions, traffic black spots (ie, high-risk areas for accidents), and barrier-free facilities (eg, ramps, staircases, footways, lifts). These factors will be derived from government infrastructure data and 3D pedestrian network datasets within a defined radius of each GPS position. Land use mix will be quantified by the Shannon entropy index [33] from government land use raster files. Point of interest (POI) density will be computed through kernel density estimation of geocoded points from government POI databases. Road accessibility will be calculated using degree, closeness, and betweenness centrality measures will be derived from graph theory algorithms [34]. Furthermore, the thermal environment will be assessed through

XSL•FO
RenderX

land surface temperature, which will be retrieved from Landsat 8 thermal infrared sensor bands using the split-window algorithm and atmospheric correction techniques.

Dynamic environmental exposure will be calculated using kernel density estimation (KDE) at multiple temporal scales based on GPS trajectories [35,36]. Specifically, we will analyze exposure at the minute level, daily level, and weekly level. This multiscale approach allows us to examine both acute and cumulative environmental exposures in relation to physical activity patterns and stress responses. It distinguishes between acute high-intensity exposures at the minute level and sustained low-intensity exposures over longer periods, which may produce equivalent cumulative values but represent distinct exposure patterns across the study period. The KDE-derived exposure surface will be integrated with environmental variables to create composite exposure metrics, which will then be temporally aligned with accelerometer-derived activity data and physiological measurements. For example, minute-level environmental exposure can be directly linked to concurrent physical activity intensity and MET values, while daily and weekly exposure metrics can be associated with aggregate activity patterns and cumulative physiological responses.

For the overall measurement of built environment characteristics and exposure metrics, we will perform index normalization across all environmental indicators, as these measures are originally on different scales (eg, NDVI ranges from –1 to 1, while AQI ranges from 0 to 500). Each environmental indicator will be transformed to a standardized scale (0-1) while maintaining their relative distributions, supporting both within-subject comparisons of different environmental exposures and between-subject comparisons of exposure patterns.

Self-Report Measures and Covariates

In addition to objective measurements (eg, GPS-recorded location and biosensor-derived biomarkers), subjective measurements were collected through structured questionnaires at 3 time points: pre-experiment, during-experiment, and postexperiment. The questionnaires are provided in both English and Cantonese, allowing participants to choose their preferred language.

The preinterview questionnaire contained health-related quality of life assessment using SF-12, which evaluates both physical and mental health components, health behaviors including smoking and alcohol consumption patterns, and perceived built environment quality. The built environment perception was assessed through 3 domains regarding public service accessibility, public open space and green space quality, and transportation convenience. These items were measured on a 5-point Likert scale from 1 (strongly disagree) to 5 (strongly agree). Weather perception and comfort assessment were also included.

During the experiment, participants completed daily surveys recording their travel modes, activity locations and types, estimated phone screen time, and clothing combinations. Emotional states were assessed using a modified version of the Positive and Negative Affect Schedule, comprising 10 mood items rated on Likert scales [37].

At the end of the experiment, participants completed an exit interview reflecting on their experiences over the 7-day experimental period. The interview assessed mental health using the Patient Health Questionnaire for depression symptoms, and health behaviors through detailed alcohol consumption and smoking patterns (frequency, quantity, and motivations). Environmental perceptions and mood responses to different environments were evaluated, including emotional responses to various urban settings such as public gardens, waterfront promenades, urban streets, and mountain trails. Overall device experience and weather comfort preferences were also assessed. These subjective measurements complement our objective data by capturing perceptions and experiences that cannot be recorded by GPS devices or biosensors.

Demographic covariates include age, gender, height (cm), weight (kg), education level, marital status, occupation status, housing type, length of residence in Hong Kong (years), living conditions, monthly household income (excluding welfare or disability allowances), household size (including the participant but excluding foreign domestic workers), residential area in Hong Kong, and building of residence.

Data Processing and Analysis

GPS data (recorded at 1-second intervals) and physiological data from the biosensor will be temporally aligned at 1-minute intervals. To explore the relationship between environmental exposure and mental health, descriptive statistics will be computed for all variables, including biomarkers and demographic characteristics collected from biosensors and survey data. Diagnostic statistics will be applied to assess model adequacy. Outcome variables (eg, biomarkers) will be log-transformed when necessary to better approximate Gaussian distribution for residuals.

Environmental exposure will be quantified by applying kernel weights to the GPS data. This method accurately measures the precise exposure at the GPS point level, with a proximity-weighted raster surface at fine granularity (50 m × 50 m) will be constructed for the activity space of each individual [38]. For each grid cell, exposure values will be calculated based on kernel density estimation, with weights proportional to the time spent in each location.

To analyze the relationship between these environmental exposures and physiological responses, we will use multiple analytical approaches. Statistical testing will follow a hierarchical approach with α=.01 for primary analyses examining the 9 prespecified associations (3 environmental exposures × 3 digital biomarkers) and α=.05 for secondary analyses. Generalized linear mixed models will serve as our primary method to examine the direct relationships between GPS-derived exposure measures and biomarker variations, accounting for individual differences through random effects. To capture potential nonlinear relationships, we will use generalized additive mixed models. These models will be enhanced with time-series analysis techniques, including autoregressive integrated moving average models to examine temporal patterns and cross-correlation functions to identify potential lag effects between environmental exposure and physiological responses. This analytical framework allows us

XSL•FO
RenderX

to understand both the immediate and delayed effects of environmental exposure on physiological responses, while accounting for individual variations and temporal dynamics.

Structural equation modeling will be used to investigate the mediating role of physical activities between environmental exposure and mental health. It accounts for both observable and latent variables, particularly mental health as measured by the SF-12 questionnaire. To further explore the spatiotemporal variation in the relationship between environmental exposures and mental health changes, geographically and temporally weighted regression will be used, which extends traditional geographical weighted regression by incorporating both spatial and temporal dimensions, revealing the varying relationships across space and time. This approach captures dynamic patterns of environmental exposure influences on mental health status across different locations and temporal periods.

Ethical Considerations

The Human Ethics Research Committee at The University of Hong Kong approved the study procedures (award numbers: EA240125 and EA240514). Each participant provided written informed consent before the start of the experiment. Participants received an initial incentive of HKD $100 (US $12) and, upon completing the experiment, the remaining incentive of HKD $400 (US $50).

As we are collecting data with GPS locations, biomarkers, and questionnaires with personal demographic information, the sensitivity of this dataset requires high protection. The study was approved by the Human Research Ethic Committee at the University of Hong Kong. Only staff and student members listed on this ethics approval will have temporary access to study participants' names. This information will be pseudo-anonymized for subsequent data analysis. Student helpers for data collection were all required to sign a data protection and confidentiality agreement. The data will be stored in a password-protected computer and folder managed by the corresponding author throughout the project period. Due to the large quantity of data, the corresponding author has secured access to a high-performance computing account hosted by the University of Hong Kong. All raw data will be backed up daily. All other data and codes for analytical methods can be obtained by contacting the corresponding author. The primary disseminations of our research results will comprise journal publications and presentations at academic conferences. We will also generate reports and provide evidence-based materials for policymakers to facilitate the development of more effective mental disorder intervention strategies.

Results

This study received internal funding in May 2024 and external funding from the Health and Medical Research Fund of Hong Kong Health Bureau in July 2025. The Human Ethics Research Committee at The University of Hong Kong approved the study procedures (award numbers: EA240125 and EA240514). Each participant provides written informed consent before participating in the study.

This pilot study supported by internal funds began in September 2024. The initial setup and testing of GPS devices and biosensors were conducted from October to December 2024. As of June 30, 2025, a total of 150 participants from this pilot study had completed the 1-week experiment period, including participants from diverse population groups such as older adults and foreign domestic workers. This exploratory study aimed to validate the data collection protocol and refine study procedures, particularly focusing on device-wearing compliance and data quality assurance. As a pilot feasibility study, we monitored these feasibility metrics during data collection. We achieved a 94% (141/150) study completion rate, with participants providing ≥5 valid biosensor days and 91% (137/150) completing ≥5 of 7 daily questionnaires. These observations demonstrate the feasibility of our protocols for the larger-scale investigation.

Participant feedback sessions for protocol refinement were conducted during exit interviews throughout the pilot study period, helping identify key challenges in data collection procedures and device usage. Some participants reported difficulty with data collection during the summer months due to high temperatures and extreme weather conditions, which resulted in limited outdoor activity time and consequently reduced GPS recording coverage. A portion of participants found it challenging to maintain consistent device-wearing time, particularly during sleep or water-related activities such as swimming or showering. Several participants also noted that the limited battery life of the biosensors required frequent charging, occasionally leading to data gaps. Based on these findings, the research team has implemented enhanced support protocols with regular check-ins and reminders, detailed device management guidelines, and flexible data collection scheduling. These refinements did not change the core data collection instruments or measures, ensuring comparability between pilot and main study data.

Following the successful pilot phase, the main study, supported by the Health and Medical Research Fund, commenced in July 2025, with 30 participants enrolled and the protocol implemented as of September 2025. Full data collection for 750 participants will continue until December 2026, followed by comprehensive data analysis focusing on the associations between environmental exposure, physical activity, and mental health biomarkers. Environmental exposure data and physiological measurements are being continuously collected and processed, with specific attention to seasonal variations and weather conditions that may affect data quality. We are at an early stage of data cleaning and pre-processing for the pilot study. Data analysis is expected to be completed by June 2027, with results anticipated to be published in 2027 and 2028.

Discussion

Anticipated Findings

We expect to observe improved PRV and reduced EDA when exposed to green and blue spaces, with greater exposure duration and intensity associated with greater effects on these biomarkers. Conversely, we anticipate that higher air pollution levels (PM2.5, NO_2, and AQI) and greater urban density (higher GCR

XSL•FO
RenderX

and SCR) will be associated with elevated EDA and reduced PRV. This study will examine the temporal dynamics of these effects through continuous monitoring and identify the optimal duration and intensity of environmental exposures for stress reduction. We predict cumulative effects of environmental exposure on digital biomarkers associated with mental health over the study period. Participants with higher cumulative exposure to green spaces may exhibit lower baseline stress levels and better stress recovery patterns, while those with long-term exposure to air pollution and high urban density may show elevated baseline stress and impaired recovery.

Additionally, our minute-level assessment examines whether physical activity mediates the relationship between environmental exposure and stress reduction. We anticipate that physical activity, operationalized through MET, will partially mediate this relationship. Specifically, we expect moderate-to-vigorous physical activity during green-blue space exposure to strengthen the associations with lower stress levels.

Furthermore, we expect the associations between all 3 primary environmental exposure and stress biomarkers to vary across demographic subgroups and baseline mental health status. In particular, older adults and individuals with elevated baseline stress may exhibit distinct biomarker responses to environmental exposure compared to younger adults and those with lower baseline stress. As part of our secondary analyses, we will also explore associations between secondary parameters (mobility barriers, land use mix, road accessibility, and thermal environment) and additional stress-related biomarkers such as respiratory rate and sleep quality.

Comparison to Prior Work

This study addresses critical gaps in existing literature through several methodological innovations. Existing studies on mental health have adopted the activity and travel logs to record daily activities [39,40]. However, the locational accuracy of these records is relatively low [41]. Compared to travel logs, GPS tracking, with continuous location information, captures environmental exposure at a more precise spatial and temporal granularity and has been widely used in public health [42,43]. For example, using ecological momentary assessment, researchers have examined the association between momentary depressive symptoms and mobility-based microurban environmental exposure by integrating GPS and accelerometer data to track daily routines [44-47]. While GPS has been widely used in public health research, there remains limited work that combines GPS data with objective measures of urban environments, including urban density, land use mix, and accessibility to facilities, to investigate their impacts on mental health [48,49]. Moreover, measurement on mental health is limited at low temporal resolution with self-reports [15,16]. Our minute-level assessment fills this gap by allowing for precise evaluation of associations between environment, physical activities, and biomarkers associated with mental health.

Previous literature has indicated that exposure to green space is able to mitigate depressive symptoms, alleviate anxiety [50], and stress, and improve cognitive abilities [51]. In addition to the visual characteristics of the environment, practical aspects—particularly the environment's walkability—have been linked to improved mental health. Walkability includes different aspects of the built environment, such as land use mix, intersection density, and public transportation accessibility [52,53]. Walkability has also been linked to higher intensity of physical activity and more social interaction [54]. Therefore, our subsequent research papers will have a focus on the impact from built environment on fluctuations in mental health status.

Wearable technology has the potential to revolutionize mental health care systems by continuously monitoring vital physiological parameters, allowing real-time assessment [20,21]. Physiological parameters, including heart rate variability (HRV), EDA, and behavioral parameters captured by the wearable biosensors, have already demonstrated encouraging results in the identification of stress, anxiety, and depression [55]. For example, using such wearables, it has been shown that short-term exposure to urban green spaces can lead to lower blood pressure, improved HRV indices, and potentially reduced heart rate, indicating a decrease in stress levels [56]. In terms of stress mechanisms, EDA reflects sympathetic nervous system activity, indicating emotional arousal, while skin temperature fluctuations, mediated by vasomotor responses, reveal stress-induced autonomic changes [57]. Respiratory rate, HRV, along with heart rate variations can all reflect stress-related physiological responses.

However, existing studies using biomarkers to link environmental exposure and health outcomes have been limited by insufficient sample size, with participant numbers typically below 20 [36]. These studies often measured immediate physiological responses and acute stress markers during outdoor environmental exposure [58,59], without investigating cumulative effects. The absence of baseline measurements before environmental exposure in many studies prevents accurate assessment of stress-related changes. [60]. This study overcomes these limitations by implementing continuous monitoring over extended periods, establishing baseline measurements, and recruiting a larger and more diverse sample. We investigate both acute and cumulative environmental exposures and their contributions to mental health outcomes through routine residential exposure and daily activity patterns over time. This approach is particularly important since environmental stress in disadvantaged neighborhoods may serve as a mechanism underlying health disparities [61,62].

Strengths and Limitations

This study has several strengths. First, the integration of minute-level GPS tracking with continuous biosensor monitoring represents a significant methodological advancement, allowing precise spatiotemporal assessment of environmental exposure and stress-related responses. Second, we are collecting data from diverse population groups, including older adults and foreign domestic workers, allowing us to compare demographic groups (eg, younger adults vs older adults) and explore differences in stress impact from physical-social-built environments. This diversity enhances the generalizability of our findings across different population segments in Hong Kong. Third, the combination of objective biomarkers, questionnaires, and GPS records provides a comprehensive assessment of the relationship between environmental exposure and mental health.

XSL•FO
RenderX

Fourth, the extended monitoring period allows for examination of both acute and cumulative effects of environmental exposure, addressing a significant gap in existing literature. Lastly, this study design includes baseline measurements before environmental exposure, enabling accurate assessment of stress-related changes.

However, several limitations should be acknowledged. First, our exclusion criteria (no neuropsychological disorders, no severe mobility limitations, able to maintain device charging) may systematically exclude individuals at high mental health risk, who may be less likely to enroll in the study. This limits the generalizability of findings to the broader population with mental health concerns. Second, the study protocol cannot be applied for children and teenagers, who have become a major concern in the region regarding mental health problems. As teenagers and children require a guardian when conducting the experiment, different study protocols are suggested for future research. Third, participant reactivity to wearing monitoring devices may influence behavior and stress-related responses, although this effect typically diminishes over time. Lastly, this protocol is not designed as a randomized controlled trial, and therefore is not applicable for causal inference on environmental exposure.

Feasibility Challenges

Several challenges were encountered during pilot data collection. First, data collection was difficult during the summer time due to the high temperatures or extreme weather conditions, resulting in limited outdoor activity time and thus limited GPS recording. Although subsequent studies could investigate seasonal difference or thermal influence on mental health status, recommendation for data collection shall be considered under normal weather conditions. Second, some participants had difficulty maintaining consistent device-wearing time, particularly during sleep or water-related activities (such as swimming or showering). Third, the battery life of biosensor required regular charging, occasionally leading to data gaps.

In addition, different participant subgroups presented distinct challenges. Older adults (aged >65 years) required longer onboarding sessions (average 45 minutes vs 30 minutes) and reported concerns about device loss when wearing the device outdoors. Foreign domestic workers had limited private time for device charging and questionnaire completion, typically available only late evening, and experienced scheduling constraints due to work commitments. Initial contact-to-enrollment conversion rate was 68%, with time commitment cited as the primary barrier. These population-specific barriers informed adjusted support strategies, including flexible appointment scheduling and charging solutions.

The research team implemented adjusted support strategies. For general operational issues, regular check-ins and detailed guidelines were provided throughout the study procedure, and potential support from backup devices was made available to resolve charging-related data gaps. For population-specific barriers, tailored solutions included flexible appointment scheduling, instructional videos for older adults, and portable charging options.

Next Steps and Future Directions

Full data collection with our target sample of 750 participants is expected to be completed by December 2026. Our subsequent analysis will explore the impact of daily activities and environmental exposure on both sleep quality and stress recovery patterns in urban environments. We will examine the mediating role of physical activity in the relationship between environmental exposure and mental health biomarkers.

Building on this protocol, future research will leverage the Convolutional Neural Networks–Long Short-Term Memory hybrid model to develop real-time stress prediction systems that forecast biomarker responses based on individuals' environmental exposures and physical activities. Random Forests and Gradient Boosting Machines will be used to identify critical environmental features and exposure thresholds that trigger stress reduction, allowing personalized exposure recommendations. We will develop and test targeted interventions based on identified environmental features, to transform findings into actionable mental health promotion strategies. We anticipate contributing to evidence-based urban design and public health policies that promote mental well-being in dense urban environments.

Dissemination Plan

We will disseminate findings through multiple channels to ensure broad impact. At the academic level, we will publish in open-access and high-impact journals, present at conferences, and share findings through research seminars and a quarterly research exchange platform. For policy and community impact, we will collaborate with Hong Kong Health Bureau to incorporate evidence-based recommendations into official mental health guidelines and share findings with therapists to inform personalized intervention strategies. Analytical code will be made available in public repositories upon publication to facilitate reproducibility of similar research in the future.

Acknowledgments

The authors thank Angus Chi-Yan Leung, Edrian Yijun Liu, Amy Lok Yi Zhong, Katrina Yuet Ying Leung, Agatha Wing Yan Szeto, and Vincent Man To Li for their assistance with data collection and for providing feedback from participants on the experiment. We also thank Jiahui Liu for contributions to the initial proposal, including the introduction, literature review, and preliminary project design.

Grammar checking was delegated to artificial intelligence (AI; Claude 3.5 Sonnet) under full human supervision. Responsibility for the final manuscript lies entirely with the authors. AI tools are not listed as authors and do not bear responsibility for the final outcomes.

XSL•FO
RenderX

Data Availability

The datasets generated from this research are not publicly available due to the need to protect participant privacy. But they are available from the corresponding author upon request with justification.

Funding

This research is supported by the Internal Seed Grant Scheme for interdisciplinary research at The University of Hong Kong and the Health Medical Research Fund (No 22232431) from the Health Bureau of Hong Kong Special Administrative Region.

Authors' Contributions

YC contributed to the study design, monitored the data collection, wrote the first draft of the paper, visualization, revised the paper, acquired the funding and supervised the project. TL contributed to visualization, tables, literature review, data analytical plan and revised the paper. CT and DC contributed to the design of the study and revised the paper. All authors reviewed the draft and approved the submission.

Conflicts of Interest

None declared.

References

1. Cuijpers P, Javed A, Bhui K. The WHO World Mental Health Report: a call for action. Br J Psychiatry 2023;222(6):227-229. [doi: 10.1192/bjp.2023.9] [Medline: 36794529]
2. Barros P, Ng Fat L, Garcia LM, Slovic AD, Thomopoulos N, de Sá TH, et al. Social consequences and mental health outcomes of living in high-rise residential buildings and the influence of planning, urban design and architectural decisions: a systematic review. Cities 2019;93:263-272. [doi: 10.1016/j.cities.2019.05.015]
3. Center for Health Protection. Health Behaviour Survey 2023. 2023. URL: https://www.chp.gov.hk/en/static/101105.html [accessed 2025-12-15]
4. Lam LC, Wong CS, Wang M, Chan W, Chen EY, Ng RM, et al. Prevalence, psychosocial correlates and service utilization of depressive and anxiety disorders in Hong Kong: the Hong Kong Mental Morbidity Survey (HKMMS). Soc Psychiatry Psychiatr Epidemiol 2015;50(9):1379-1388. [doi: 10.1007/s00127-015-1014-5] [Medline: 25660760]
5. Wong SM, Chen EY, Suen Y, Wong CS, Chang W, Chan SK, et al. Prevalence, time trends, and correlates of major depressive episode and other psychiatric conditions among young people amid major social unrest and COVID-19 in Hong Kong: a representative epidemiological study from 2019 to 2022. Lancet Reg Health West Pac 2023;40:100881 [FREE Full text] [doi: 10.1016/j.lanwpc.2023.100881] [Medline: 37654623]
6. Kirkbride JB, Anglin DM, Colman I, Dykxhoorn J, Jones PB, Patalay P, et al. The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry 2024;23(1):58-90 [FREE Full text] [doi: 10.1002/wps.21160] [Medline: 38214615]
7. Wang X, Cao Y, DeJohn A. Spatial disparities in access to public and private long-term care facilities in Hong Kong. The Professional Geographer 2025;77(5):611-627. [doi: 10.1080/00330124.2025.2543540]
8. Arango C, Dragioti E, Solmi M, Cortese S, Domschke K, Murray RM, et al. Risk and protective factors for mental disorders beyond genetics: an evidence-based atlas. World Psychiatry 2021;20(3):417-436 [FREE Full text] [doi: 10.1002/wps.20894] [Medline: 34505386]
9. Cao Y, Stewart K, Kalil R. Geographic patterns of end-stage renal disease and kidney transplants in the Midwestern United States. Applied Geography 2016;71:133-143. [doi: 10.1016/j.apgeog.2016.05.001]
10. Labib S, Lindley S, Huck JJ. Estimating multiple greenspace exposure types and their associations with neighbourhood premature mortality: a socioecological study. Sci Total Environ 2021;789:147919 [FREE Full text] [doi: 10.1016/j.scitotenv.2021.147919] [Medline: 34062470]
11. Gascon M, Sánchez-Benavides G, Dadvand P, Martínez D, Gramunt N, Gotsens X, et al. Long-term exposure to residential green and blue spaces and anxiety and depression in adults: a cross-sectional study. Environ Res 2018;162:231-239. [doi: 10.1016/j.envres.2018.01.012] [Medline: 29358115]
12. Helbich M. Toward dynamic urban environmental exposure assessments in mental health research. Environ Res 2018;161:129-135 [FREE Full text] [doi: 10.1016/j.envres.2017.11.006] [Medline: 29136521]
13. Jung PH, Song J. Multivariate neighborhood trajectory analysis: an exploration of the functional data analysis approach. Geographical Analysis 2021;54(4):789-819. [doi: 10.1111/gean.12298]
14. Marquet O, Tello-Barsocchini J, Couto-Trigo D, Gómez-Varo I, Maciejewska M. Comparison of static and dynamic exposures to air pollution, noise, and greenness among seniors living in compact-city environments. Int J Health Geogr 2023;22(1):3 [FREE Full text] [doi: 10.1186/s12942-023-00325-8] [Medline: 36709304]
15. Kruize H, van Kamp I, van den Berg M, van Kempen E, Wendel-Vos W, Ruijsbroek A, et al. Exploring mechanisms underlying the relationship between the natural outdoor environment and health and well-being - results from the

XSL•FO
RenderX

PHENOTYPE project. Environ Int 2020;134:105173 [FREE Full text] [doi: 10.1016/j.envint.2019.105173] [Medline: 31677803]
16. Moore T, Kesten J, López-López JA, Ijaz S, McAleenan A, Richards A, et al. The effects of changes to the built environment on the mental health and well-being of adults: systematic review. Health Place 2018;53:237-257 [FREE Full text] [doi: 10.1016/j.healthplace.2018.07.012] [Medline: 30196042]
17. Latkin CA, German D, Hua W, Curry AD. Individual-level influences on perceptions of neighborhood disorder: a multilevel analysis. J Community Psychol 2009;37(1):122-133 [FREE Full text] [doi: 10.1002/jcop.20284] [Medline: 20027234]
18. Doherty ST, Lemieux CJ, Canally C. Tracking human activity and well-being in natural environments using wearable sensors and experience sampling. Soc Sci Med 2014;106:83-92. [doi: 10.1016/j.socscimed.2014.01.048] [Medline: 24549253]
19. Davis-Martin RE, Alessi SM, Boudreaux ED. Alcohol use disorder in the age of technology: a review of wearable biosensors in alcohol use disorder treatment. Front Psychiatry 2021;12:642813 [FREE Full text] [doi: 10.3389/fpsyt.2021.642813] [Medline: 33828497]
20. Raj Theeng Tamang M, Sharif MS, Al-Bayatti AH, Alfakeeh AS, Omar Alsayed A. A machine-learning-based approach to predict the health impacts of commuting in large cities: case study of London. Symmetry 2020;12(5):866. [doi: 10.3390/sym12050866]
21. Babu M, Lautman Z, Lin X, Sobota MHB, Snyder MP. Wearable devices: implications for precision medicine and the future of health care. Annu Rev Med 2024;75:401-415 [FREE Full text] [doi: 10.1146/annurev-med-052422-020437] [Medline: 37983384]
22. Huckins JF, daSilva AW, Wang R, Wang W, Hedlund EL, Murphy EI, et al. Fusing mobile phone sensing and brain imaging to assess depression in college students. Front Neurosci 2019;13:248 [FREE Full text] [doi: 10.3389/fnins.2019.00248] [Medline: 30949024]
23. Ewen HH, Emerson KG, Washington TR, Carswell AT, Smith ML. Aging in place: community-based services and resources in residential settings among older adults. Housing and Society 2017;44(1-2):114-126. [doi: 10.1080/08882746.2017.1389577]
24. Charan J, Biswas T. How to calculate sample size for different study designs in medical research? Indian J Psychol Med 2013;35(2):121-126 [FREE Full text] [doi: 10.4103/0253-7176.116232] [Medline: 24049221]
25. Jankowska MM, Sears DD, Natarajan L, Martinez E, Anderson CAM, Sallis JF, et al. Protocol for a cross sectional study of cancer risk, environmental exposures and lifestyle behaviors in a diverse community sample: the community of mine study. BMC Public Health 2019;19(1):186 [FREE Full text] [doi: 10.1186/s12889-019-6501-2] [Medline: 30760246]
26. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc 2005;37(11 Suppl):S531-S543. [doi: 10.1249/01.mss.0000185657.86065.98] [Medline: 16294116]
27. de Zambotti M, Goldstein C, Cook J, Menghini L, Altini M, Cheng P, et al. State of the science and recommendations for using wearable technology in sleep and circadian research. Sleep 2024;47(4). [doi: 10.1093/sleep/zsad325] [Medline: 38149978]
28. Albrecht BM, Flaßkamp FT, Koster A, Eskofier BM, Bammann K. Cross-sectional survey on researchers' experience in using accelerometers in health-related studies. BMJ Open Sport Exerc Med 2022;8(2):e001286 [FREE Full text] [doi: 10.1136/bmjsem-2021-001286] [Medline: 35601138]
29. Gerboni G, Comunale G, Chen W, Lever Taylor J, Migliorini M, Picard R, et al. Prospective clinical validation of the Empatica EmbracePlus wristband as a reflective pulse oximeter. Front Digit Health 2023;5:1258915 [FREE Full text] [doi: 10.3389/fdgth.2023.1258915] [Medline: 38111608]
30. Winter M, Pryss R, Probst T, Reichert M. Towards the applicability of measuring the electrodermal activity in the context of process model comprehension: feasibility study. Sensors (Basel) 2020;20(16):4561. [doi: 10.3390/s20164561] [Medline: 32823891]
31. Lu Y. Using Google Street View to investigate the association between street greenery and physical activity. Landscape and Urban Planning 2019;191:103435. [doi: 10.1016/j.landurbplan.2018.08.029]
32. Cheng Y, Browning MH, Zhao B, Qiu B, Wang H, Zhang J. How can urban green space be planned for a ‘happy city’? Evidence from overhead- to eye-level green exposure metrics. Landscape Urban Plann 2024;249:105131. [doi: 10.1016/j.landurbplan.2024.105131]
33. Song Y, Merlin L, Rodriguez D. Comparing measures of urban land use mix. Comput Environ Urban Syst 2013;42:1-13. [doi: 10.1016/j.compenvurbsys.2013.08.001]
34. Sarlas G, Páez A, Axhausen KW. Betweenness-accessibility: estimating impacts of accessibility on networks. J Transp Geogr 2020;84:102680. [doi: 10.1016/j.jtrangeo.2020.102680]
35. Jankowska MM, Yang J, Luo N, Spoon C, Benmarhnia T. Accounting for space, time, and behavior using GPS derived dynamic measures of environmental exposure. Health Place 2023;79:102706 [FREE Full text] [doi: 10.1016/j.healthplace.2021.102706] [Medline: 34801405]
36. Kondo MC, Jacoby SF, South EC. Does spending time outdoors reduce stress? A review of real-time stress response to outdoor environments. Health Place 2018;51:136-150. [doi: 10.1016/j.healthplace.2018.03.001] [Medline: 29604546]
37. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037//0022-3514.54.6.1063] [Medline: 3397865]

XSL•FO
RenderX

38. Liu Y, Kwan M, Wang J, Cai J. Confounding associations between green space and outdoor artificial light at night: systematic investigations and implications for urban health. Environ Sci Ecotechnol 2024;21:100436 [FREE Full text] [doi: 10.1016/j.ese.2024.100436] [Medline: 39027466]
39. Zhang L, Zhou S, Kwan M, Shen M. Assessing individual environmental exposure derived from the spatiotemporal behavior context and its impacts on mental health. Health Place 2021;71:102655. [doi: 10.1016/j.healthplace.2021.102655] [Medline: 34482159]
40. Axhausen KW, Zimmermann A, Schönfelder S, Rindsfüser G, Haupt T. Observing the rhythms of daily life: a six-week travel diary. Transportation 2002;29(2):95-124. [doi: 10.1023/a:1014247822322]
41. Yoo EH, Rudra C, Glasgow M, Mu L. Geospatial estimation of individual exposure to air pollutants: moving from static monitoring to activity-based dynamic exposure assessment. Ann Assoc Am Geogr 2015;105(5):915-926. [doi: 10.1080/00045608.2015.1054253]
42. Pearson AL, Tribby C, Brown CD, Yang J, Pfeiffer K, Jankowska MM. Systematic review of best practices for GPS data usage, processing, and linkage in health, exposure science and environmental context research. BMJ Open 2024;14(2):e077036 [FREE Full text] [doi: 10.1136/bmjopen-2023-077036] [Medline: 38307539]
43. Cao Y, Renschler C, Jacquez G. A data model to capture spatial and temporal exposure. 2014 Presented at: Proceedings of GIScience; September 12-15, 2014; Vienna p. 258-263.
44. Kingsbury C, Buzzi M, Chaix B, Kanning M, Khezri S, Kiani B, et al. STROBE-GEMA: a STROBE extension for reporting of geographically explicit ecological momentary assessment studies. Arch Public Health 2024;82(1):84 [FREE Full text] [doi: 10.1186/s13690-024-01310-8] [Medline: 38867286]
45. Kestens Y, Kingsbury C. In the right place at the right time: leveraging geographic ecological momentary assessment to analyze contexts and health. In: Environmental Neuroscience. Cham: Springer Nature Switzerland; 2024:517-532.
46. Yim SJ, Lui LM, Lee Y, Rosenblat JD, Ragguett R, Park C, et al. The utility of smartphone-based, ecological momentary assessment for depressive symptoms. J Affect Disord 2020;274:602-609. [doi: 10.1016/j.jad.2020.05.116] [Medline: 32663993]
47. Fancello G, Vallée J, Sueur C, van Lenthe FJ, Kestens Y, Montanari A, et al. Micro urban spaces and mental well-being: measuring the exposure to urban landscapes along daily mobility paths and their effects on momentary depressive symptomatology among older population. Environ Int 2023;178:108095 [FREE Full text] [doi: 10.1016/j.envint.2023.108095] [Medline: 37487375]
48. Hu M, Li W, Li L, Houston D, Wu J. Refining time-activity classification of human subjects using the global positioning system. PLoS One 2016;11(2):e0148875 [FREE Full text] [doi: 10.1371/journal.pone.0148875] [Medline: 26919723]
49. Jacobson NC, Bhattacharya S. Digital biomarkers of anxiety disorder symptom changes: Personalized deep learning models using smartphone sensors accurately predict anxiety symptoms from ecological momentary assessments. Behav Res Ther 2022;149:104013 [FREE Full text] [doi: 10.1016/j.brat.2021.104013] [Medline: 35030442]
50. Paul LA, Hystad P, Burnett RT, Kwong JC, Crouse DL, van Donkelaar A, et al. Urban green space and the risks of dementia and stroke. Environ Res 2020;186:109520. [doi: 10.1016/j.envres.2020.109520] [Medline: 32344208]
51. Dadvand P, Nieuwenhuijsen MJ, Esnaola M, Forns J, Basagaña X, Alvarez-Pedrerol M, et al. Green spaces and cognitive development in primary schoolchildren. Proc Natl Acad Sci U S A 2015;112(26):7937-7942 [FREE Full text] [doi: 10.1073/pnas.1503402112] [Medline: 26080420]
52. Renalds A, Smith T, Hale P. A systematic review of built environment and health. Fam Community Health 2010;33(1):68-78. [doi: 10.1097/FCH.0b013e3181c4e2e5] [Medline: 20010006]
53. Yang X, Li T, Cao Y, Zheng X, Tang L. A novel framework integrating GeoAI and human perceptions to estimate walkability in Wuhan, China. Sci Rep 2025;15(1):25377 [FREE Full text] [doi: 10.1038/s41598-025-09779-1] [Medline: 40659718]
54. Wang R, Lu Y, Zhang J, Liu P, Yao Y, Liu Y. The relationship between visual enclosure for neighbourhood street walkability and elders' mental health in China: using street view images. J Transp Health 2019;13:90-102. [doi: 10.1016/j.jth.2019.02.009]
55. Can YS, Chalabianloo N, Ekiz D, Ersoy C. Continuous stress detection using wearable sensors in real life: algorithmic programming contest case study. Sensors (Basel) 2019;19(8):1849 [FREE Full text] [doi: 10.3390/s19081849] [Medline: 31003456]
56. Lanki T, Siponen T, Ojala A, Korpela K, Pennanen A, Tiittanen P, et al. Acute effects of visits to urban green environments on cardiovascular physiology in women: a field experiment. Environ Res 2017;159:176-185 [FREE Full text] [doi: 10.1016/j.envres.2017.07.039] [Medline: 28802208]
57. Acerbi G, Rovini E, Betti S, Tirri A, Flóra Rónai J, Sirianni A, et al. A wearable system for stress detection through physiological data analysis. 2017 Presented at: The 7th Italian Forum on Ambient Assisted Living (ForItAAL 2016); June 20-23, 2016; Pisa p. 31-50. [doi: 10.1007/978-3-319-54283-6_3]
58. Mavros P, J Wälti M, Nazemi M, Ong CH, Hölscher C. A mobile EEG study on the psychophysiological effects of walking and crowding in indoor and outdoor urban environments. Sci Rep 2022;12(1):18476 [FREE Full text] [doi: 10.1038/s41598-022-20649-y] [Medline: 36323718]
59. Korpilo S, Nyberg E, Vierikko K, Ojala A, Kaseva J, Lehtimäki J, et al. Landscape and soundscape quality promote stress recovery in nearby urban nature: a multisensory field experiment. Urban For Urban Green 2024;95:128286. [doi: 10.1016/j.ufug.2024.128286]

XSL•FO
RenderX

60. Yao W, Zhang X, Gong Q. The effect of exposure to the natural environment on stress reduction: a meta-analysis. Urban For Urban Green 2021;57:126932. [doi: 10.1016/j.ufug.2020.126932]
61. Warr D, Feldman P, Tacticos T, Kelaher M. Sources of stress in impoverished neighbourhoods: insights into links between neighbourhood environments and health. Aust N Z J Public Health 2009;33(1):25-33 [FREE Full text] [doi: 10.1111/j.1753-6405.2009.00334.x] [Medline: 19236355]
62. Cao Y, Jankowska MM, Yang J, Shi Y. Spatial and temporal pattern of cannabis use disorder in California 2010-2019. Spat Spatiotemporal Epidemiol 2022;42:100520. [doi: 10.1016/j.sste.2022.100520] [Medline: 35934327]

Abbreviations

AQI: air quality index
bpm: beats per minute
BVP: blood volume pulse
EDA: electrodermal activity
GCR: ground coverage ratio
HRV: heart rate variability
KDE: kernel density estimation
MET: metabolic equivalent of task
NDVI: normalized difference vegetation index
NO2: nitrogen dioxide
PM2.5: particulate matter ≤2.5 μm
POI: point of interest
PPG: photoplethysmography
PRV: pulse rate variability
RMSSD: root-mean-square of successive difference
SCR: street coverage ratio
SF-12: Short Form-12
STPU: small territorial planning unit

Edited by J Sarvestan; submitted 27.Sep.2025; this is a non–peer-reviewed article;accepted 12.Dec.2025; published 08.Jan.2026.

Please cite as:
Cao Y, Li T, Tribby CP, Chang DHF
Environmental Exposure and Mental Health in Hong Kong: Protocol for a GPS- and Biosensor-Based Observational Study
JMIR Res Protoc 2026;15:e84919
URL: https://www.researchprotocols.org/2026/1/e84919
doi:10.2196/84919
PMID:

©Yanjia Cao, Tianyu Li, Calvin P Tribby, Dorita H F Chang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 08.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Promoting Mental Health, Physical Activity, and Healthy and Sustainable Dietary Behavior Among Practical Education Students in the Netherlands: Protocol for a Multiphase Participatory Research Study

Daniëlla van Uden[1], MSc; Annemarie Wagemakers[1], PhD; Margot Peeters[2], PhD; Marloes Kleinjan[3], PhD; Kirsten Verkooijen[1], PhD

[1]Health and Society Group, Wageningen University & Research, Wageningen, The Netherlands

[2]Youth Studies, Department of Interdisciplinary Social Sciences, Utrecht University, Utrecht, The Netherlands

[3]Organisation of Healthcare and Services, HAN University of Applied Sciences, Nijmegen, The Netherlands

Corresponding Author:
Daniëlla van Uden, MSc
Health and Society Group
Wageningen University & Research
Hollandseweg 1
Wageningen, 6706 KN
The Netherlands
Phone: 31 0317488345
Email: daniella.vanuden@wur.nl

Abstract

Background: While a healthy lifestyle at a young age benefits youth now and later in life, not all youth have equal access to resources and support for adopting a healthy lifestyle. Most youth health promotion programs target the general adolescent population without addressing underlying equity issues. Similarly, participatory research, a promising methodology for the development of health promotion programs and addressing health equity, leaves youth in more vulnerable positions, often underrepresented. This research addresses these gaps by focusing on participatory research for health promotion program design with youth in practical education (praktijkonderwijs).

Objective: This research has two objectives: (1) to gain insight into how to meaningfully involve youth in vulnerable positions in participatory research for health promotion by focusing on practical education students as a case study, and (2) to gain insight into possible outputs and outcomes of the developed health promotion programs.

Methods: The research in this protocol is part of the LIFTS (Healthy Lifestyle for Low-Literate Teenagers) project, which aims to promote a healthy lifestyle for practical education students using a participatory approach. The current research uses a multiphase mixed methods research design consisting of 3 studies. Regarding the first objective, we conduct a systematic literature review (study 1) and an empirical qualitative study at practical education schools (study 2). The systematic literature review examines current knowledge on empowering approaches to engage practical education students and similar youth groups as coresearchers in health promotion program design and implementation. The qualitative study explores if and how practical education students can meaningfully be engaged as coresearchers in participatory research for health promotion program design. In line with participatory research, we developed research methods in collaboration with practical education schools and relevant LIFTS stakeholders. Regarding the second objective, we will conduct a realist evaluation of the newly designed health promotion programs within LIFTS (study 3).

Results: On April 11, 2025, the protocol for the systematic literature review (study 1) was ready and submitted to PROSPERO. In September 2025, we finished the data collection and analysis of the empirical qualitative study (study 2) and wrote the results. The realist evaluation (study 3) is foreseen for 2026.

Conclusions: This research contributes to the advancement of academic knowledge in the field of health promotion and participatory research with youth in vulnerable positions. We expect to deliver practical recommendations and lessons learned on how to actualize youth empowerment in participatory research for health promotion, which may inspire future researchers. Additionally, we seek to promote health equity in the Netherlands by contributing to the development and implementation of tailored health promotion programs for practical education students.

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/84723

(JMIR Res Protoc 2026;15:e84723) doi:10.2196/84723

KEYWORDS

community-based participatory research; empowerment; adolescent health; health promotion; health inequities; student; secondary schools; vocational education

Introduction

Background

A healthy lifestyle—comprising regular physical activity, a healthy and sustainable diet, and positive mental health—is widely recognized as a cornerstone of good health across the lifespan. Stimulating a healthy lifestyle at an early age can not only benefit adolescents in the present but also promote the quality of their health in adult life [1-3]. However, not all youth have equal access to a healthy lifestyle, as the ease of adopting and maintaining a healthy lifestyle is strongly affected by social determinants, such as educational level [4]. Secondary education students in vocational tracks generally experience greater challenges to adopt a healthy lifestyle than their peers in academic tracks, both during adolescence and later in life [5]. However, most adolescent health interventions target the general adolescent population, overlooking the role of social determinants shaping health disparities. It is therefore essential to develop and implement youth health promotion programs that cater to underrepresented youth groups to stimulate health equity [6].

Participatory research with youth is a promising methodology for the development of health promotion programs for health equity [7]. In participatory research, scholars strongly collaborate with people with lived experiences by involving them as coresearchers, rather than passive respondents, in various research phases [8]. As a methodology, it prioritizes power-sharing, conducting research with, instead of on, communities, social justice, and locally rooted research approaches [9]. Scholars increasingly value and acknowledge the pivotal impact of participatory research with youth, and a growing number of research projects involve youth as coresearchers [7,10-13]. This trend is commendable, as participatory research with youth can contribute to youth empowerment [13-15], the relevance of interventions for youth [13], the validity of scientific research [13], and health equity [7].

However, it is crucial to recognize the limitations regarding the diversity of youth groups included in participatory research. Current research overrepresents more privileged youth and, therefore, often leaves youth in more vulnerable positions underrepresented [7,16-18]. Vulnerability can be defined based on many characteristics, such as, but not limited to, disability, migration status, and socioeconomic status [19]. Challenges, such as complicated research proposals, fear of rejection by the institutional review board, or fear of inaccessibility of potential participants, can demotivate researchers from conducting research with people in vulnerable positions [20]. Currently, there are still a few examples of studies describing the participatory research process with youth in vulnerable positions to inspire future researchers [21].

Addressing this gap is important, as possibilities for youth empowerment and more equal power relations can be especially valuable for adolescents experiencing greater vulnerability and marginalization [16]. Additionally, participatory approaches provide possibilities to adapt interventions to the needs, wishes, and lived experiences of youth [13]. Therefore, it is precisely the perspective of youth in a vulnerable position that can be crucial for translating general findings to risk populations, for developing appropriate health promotion programs for these groups, and ultimately, for promoting health equity. Furthermore, we lack insight into the outcomes of health promotion programs developed using a participatory approach.

In this research, we aim to address these gaps by focusing on practical education (praktijkonderwijs) students in the Netherlands. In the Dutch context, students enrolled in practical education are often regarded as both vulnerable and underrepresented in participatory research. practical education is a type of Dutch secondary education intended for students aged 12-18 years who are being trained in practical subjects [22]. To our knowledge, practical education is an educational pathway found only in the Netherlands and not easily comparable internationally. Students unable to obtain a diploma in preparatory vocational secondary education (Voorbereidend Middelbaar Beroepsonderwijs [VMBO], EQF level 1-2 [23]) can enroll in practical education [24]. practical education holds two official admission criteria: (1) an IQ between 55 and 80 and (2) a learning delay of 3 years or more in at least two of the following domains: conceptual mathematics, reading comprehension, technical reading, or spelling. The curriculum focuses on personal development and on fostering students' ability to participate independently in society. After graduation, practical education students generally go on to vocational education or directly enter the labor market [22]. Currently, there are 177 practical education schools in the Netherlands and more than 29,000 practical education students, amounting to about 3% of all students in Dutch secondary education [24]. According to the Sectorraad Praktijkonderwijs [25]—the national body representing and advocating for practical education schools— practical education students are generally in a more vulnerable position due to their limited cognitive capacities in an increasingly complex society, and youth are more often from families with lower socioeconomic status. Because of these vulnerabilities, practical education students may experience more health challenges compared to their peers in other types of secondary schools. Health surveys show, for instance, that practical education students are less physically active [26]. However, there are limited health promotion programs specifically designed for this target group. Tailored

XSL•FO
RenderX

programs developed using a participatory approach together with practical education students, their teachers, and caregivers might result in positive changes in the healthy lifestyle of practical education students and ultimately a reduction of health inequities. According to the World Health Organization, health promotion is defined as "the process of enabling people to increase control over and to improve their health" [27]. Actions may focus on individual behavior as well as social and environmental interventions that address the root causes of ill health [27]. As we focus on the school context in this research, we focus on all actions a school takes to promote the health of its students.

Research Objective and Questions

There is still a limited understanding of how to meaningfully involve practical education students in participatory research for health promotion program design and implementation, as well as of the possible outputs and outcomes of the programs that result from a participatory approach. Therefore, this research aims to shed light on these important issues. The research takes place in a practical education setting with practical education students as the main stakeholders. The two main research questions are as follows:

1. How can practical education students be engaged as coresearchers in health promotion program design and implementation in ways that work empowering?
2. What are the outputs and outcomes of a health promotion program resulting from a participatory research approach in the practical education context, and how are they realized?

Methods

Theoretical Framework

The theoretical framework guiding this research is the six-element framework, developed by Chrifou et al [14], and is applied to better understand how to actualize and embed youth empowerment within a participatory research process for health promotion. The authors identify 4 short-term and 2 long-term goals for engaging young people as coresearchers. The short-term goals, which serve to create an enabling environment, include resources, adult facilitation, sense-making, and capacity building. The long-term goals focus on creating opportunities for empowerment and collective participation and include positive child and adolescent development and participatory competence. According to the framework, and in line with Bronfenbrenner's ecological systems theory [14,28,29], the impact of contextual factors on actualizing youth empowerment is situated within an ecological context, including the macrosystem (society and physical environment), the exosystem (community organizations and institutions), and the microsystem (youth, friends, peers, teachers, and family) [14].

The framework is used in the proposed research to operationalize the concept of empowerment and to synthesize the study results. Specifically, it helps to answer the "how-questions," namely "How can practical education students be engaged as coresearchers in ways that work empowering?" and "How are the outputs and outcomes of a health promotion program resulting from a participatory research approach in the practical education context realized?"

Study Design

This research uses a multiphase mixed methods research design consisting of 3 studies. The first study addresses research question 1 through a systematic literature review, while the second study addresses the same research question through an empirical qualitative study. The third study addresses research question 2 and consists of a realist evaluation. During phase 1, we conduct studies 1 and 2 simultaneously to enable an iterative process, which is conducive to answering research question 1. Additionally, using the same analytical framework in these two studies helps us compare the results of both studies and support the iterative process. Phase 2 will consist of study 3.

Study Setting

Overview

The proposed research in this study protocol is part of the LIFTS (Healthy Lifestyle for Low-Literate Teenagers) project (Multimedia Appendix 1), which is funded by the Netherlands Organization for Scientific Research (NWO [Nederlandse Organisatie voor Wetenschappelijk Onderzoek]; grant number KICH1.GZ03.21.011). The main research question of LIFTS is "How to promote, in a participatory way, sustainable, healthy living in practical education students with the use of accessible, acceptable, and engaging technology?" It focuses on 3 health topics, such as mental health, dietary behavior, and physical activity, and covers 4 research phases, such as needs assessment, co-design, implementation, and evaluation. Four PhD researchers are employed at LIFTS, one for each of the three health topics (ie, mental health, dietary behavior, and physical activity), and one PhD researcher to focus on the participatory process across the health topics and research phases. This study protocol involves the aims and activities of the latter PhD researcher.

The LIFTS consortium consists of 3 universities (Wageningen University & Research [WUR], Utrecht University (UU), and Eindhoven University of Technology [TU/e]), practical education schools, industry partners, and societal stakeholders (eg, Knowledge Centre for Sport & Physical Activity Netherlands, Special Heroes the Netherlands, and Trimbos Institute). Three practical education schools host the research activities. To ensure logistical feasibility, we focused on building relationships with the partners to foster collaboration and to ensure logistical feasibility in the preparation phase of LIFTS. In addition, the researchers, their supervision teams, and the consortium as a whole frequently meet to discuss progress and to ensure good alignment between the activities. Furthermore, LIFTS has appointed a program coordinator who acts as a contact person between the partners.

Study 1: Systematic Literature Review

We will conduct a systematic literature review to determine the current knowledge on empowering ways to engage practical education students and similar youth groups as coresearchers in health promotion program design and implementation. As

XSL•FO
RenderX

practical education is difficult to compare internationally, we will first look into participatory research approaches for health promotion program design and implementation for youth in secondary education in general and how these approaches can be empowering. Next, we will investigate how these empowering participatory approaches differ between educational pathways in secondary education. By distinguishing between educational pathways, we hope to get more insight into empowering participatory research approaches specifically suitable for students in practically oriented or vocational tracks, which can be translatable to the practical education context.

The electronic bibliographic databases that will be used are PubMed, ERIC, and Web of Science. The search language will include Dutch and English, and no date restrictions will be used. The study has been preregistered in PROSPERO (registration number CRD420251010102), where the full search strategy can be found.

Studies will be included in the review if (1) the research takes place in secondary school(s) (including vocational, general, or academic tracks) for youth typically aged 12-18 years; (2) the research activities are participatory in nature for youth involved; and 3) the research activities have the goal to promote health for youth, such as (but not limited to) physical activity, mental well-being, and dietary behavior.

The studies resulting from the search will be screened and selected in 2 steps. In the first step, the reviewers will select the studies by screening the title and abstract against the inclusion and exclusion criteria. In the second step, the full text of the selected studies will be screened against the inclusion and exclusion criteria. Studies found eligible at step 2 will be included in the literature review. We will use the QualSyst checklist developed by Kmet et al [30] to assess the quality of both qualitative and quantitative research included.

To synthesize our data, we will first provide an overview of the different participatory research approaches found in the selected literature. We will do this by creating a table with information about the name of the participatory approach, methods, duration process, ethical considerations, stages of youth involvement, and other outcomes than youth empowerment. Second, we will extract information from the selected studies on how they contribute to youth empowerment based on the elements of the framework of Chrifou et al [14]. Finally, we will examine how the participatory research approaches differ by type of secondary education.

Study 2: Qualitative Research

Our second method to explore "how practical education students can be engaged as co-researchers in participatory research for health promotion program design in ways that work empowering" involves an empirical qualitative study. Qualitative research aims to understand and describe people's lived experiences and worldviews and is therefore most suitable to explore the experiences of youth [31].

Data Collection

For the second study, we have collected qualitative data among 3 types of stakeholder groups (practical education students, practical education staff members [teachers and management], and researchers) during the needs assessment phase of LIFTS. We gathered information on the preferred, most meaningful, and feasible ways of practical education students' involvement in the LIFTS research process, as well as the actual experiences of practical education students' and staff members' participation in research activities. In line with participatory research, we developed the methods for data collection in collaboration with the relevant stakeholders. We conducted group interviews with practical education students using visual methods, as they are considered suitable for youth [12,32]. We conducted (group and individual) interviews with practical education teachers. The researchers involved in LIFTS maintained a reflection diary during their research activities and took part in a group interview at the end of data collection. Additionally, participant observation was conducted during the participatory research activities taking place at the schools to get more familiar with the practical education school context and the LIFTS research activities and to get to know students and teachers.

Sampling and Recruitment

The research activities have been designed, planned, and implemented in close collaboration with the staff members of the 3 participating schools. Depending on their expertise, needs, and wishes, the data collection with the practical education students took place (1) at the end of a research activity with the entire class, (2) at the end of a research activity with the class divided into small groups, and (3) at a different moment with a small group of students. Staff members who have been involved in the planning and implementation of research activities were asked to participate in an interview to reflect on their participation. To obtain a balanced and comprehensive answer to the research question, the data collection has covered the health topics and the 3 participating schools as evenly as possible.

Data Analysis

Overview

The interviews with the staff members and researchers were recorded and transcribed, and the data from interviews with students were collected on posters and in field notes. We conducted the data analysis based on the methodology for qualitative data analysis by Beuving and de Vries [33], using open and axial coding. The six-element framework of Chrifou et al [14] serves as an analytical lens to explore whether and how the participatory research process works, empowering in the practical education context.

Study 3: Realist Evaluation

To provide future researchers with insights into the working mechanisms of health promotion programs in the practical education context, we will conduct a realist evaluation [34]. Specifically, we will explore the outputs and outcomes of the health promotion program that results from a participatory research approach in the practical education context and understand how these are realized [34]. In line with studies 1 and 2, we will use the framework of Chrifou et al [14] to operationalize empowerment. Realist evaluation is a theory-driven approach to evaluating programs or interventions,

XSL•FO
RenderX

focusing on understanding how and why they work (or do not) in specific contexts. Rather than simply measuring outcomes, it explores the underlying mechanisms that produce those outcomes and the conditions that trigger them. It acknowledges that programs may work differently for different people or settings and seeks to identify patterns in these variations. Ultimately, realist evaluation aims to generate transferable lessons by explaining what works, for whom, in what circumstances, and why [35].

Corresponding to a realist approach, we will first develop a program theory. A program theory is a conceptual framework that explains how and why a program is expected to work. It outlines the causal pathways between activities, outputs, and desired outcomes, often including assumptions and contextual factors that influence success. It generally includes so-called context-mechanism-outcome configurations [35], which we also plan to create. Due to the participatory nature of LIFTS, we will keep an open approach toward the data collection. The exact methods for data collection during the test and refinement step depend, among others, on the health promotion program that will be developed. However, we foresee that the program theory and context-mechanism-outcome configurations will be shaped with input from the needs assessment and co-design activities by conducting brainstorming sessions with the 3 LIFTS PhD researchers. The following step in the realist evaluation involves the testing and refinement of the program theory. To do so, we plan to conduct realist interviews [36] with relevant stakeholders (staff, LIFTS partners, and caregivers). When refining the program theory with students, we expect, based on study 2, that a variety of practical, visual, interactive, clearly structured research activities with immediate results work best with practical education students, while quantitative methods (eg, questionnaires) are less suitable.

Data Triangulation and Integration

Our study promotes data triangulation and integration in several ways. Both studies 1 and 2 serve to answer research question 1 (How can practical education students be engaged as coresearchers in health promotion program design and implementation in ways that are work empowering?). While the qualitative research seeks to gain insight into the practical education context and provide empirical data on the participatory process, the systematic literature study analyzes participatory research for health promotion in secondary education in the literature, which helps to generalize the findings of the empirical qualitative study. Using the framework of Chrifou et al [14], both studies ensure that findings are comparable. The results of studies 1 and 2 also provide input for the participatory research process of the other three LIFTS PhDs. Additionally, the results of studies 1 and 2 inform the methodological design of the realist evaluation, as described in study 3.

Ethical Considerations

Human Subject Ethics Review Approvals or Exemptions

Study 1, the systematic literature study, does not require ethical approval since it does not involve humans. Study 2, the qualitative study, has been approved by the Wageningen University & Research Research Ethics Committee (WUR-REC; approval number 2024-011). Ethical approval for study 3, the realist evaluation, will be sought in due course.

Informed Consent

All study participants, teachers, and students are fully informed about the aims of the research prior to participation. Informed consent is obtained from all participants in interviews. Before starting the interview, the researcher goes through the information in the information sheet with participants to make sure that participants understand the aims of the study, know what is expected from them, and know what their rights are. With regard to practical education students, who do not have a high literacy and may not have a signature, we believe an oral informed consent is more suitable for this research. Additionally, passive informed consent is asked from one caregiver by email, in which they are informed about the research, and they have the possibility to opt out by responding to this email. Passive consent has been chosen after consultation with the participating practical education schools and the WUR-REC. We believe passive consent is suitable because there is minimal burden and no risk involved in the interviews with their child. Also, when asking for active informed consent by email, there might be difficulties for caregivers to digitally sign an informed consent form, and they might not respond, which leads to a small pool of students able to participate in this research. Furthermore, active consent could lead to nonparticipation—not because of disagreement with the study—but rather because of forgetting to fill out forms. This may also lead to the exclusion of students who would like to participate and share their voices. After carefully weighing the advantages and disadvantages, the WUR-REC approved this option.

Privacy and Confidentiality

LIFTS involves a collaboration between three universities. Four PhD students, who are hosted at these universities (two at WUR, one at UU, and one at TU/e), are the principal researchers of LIFTS. Each PhD student has access to her own data collection, which is stored at a safe (password-protected) place that is also accessible to the supervisors who work in the same research group. The anonymized data are also accessible upon reasonable request by the other PhD students and supervisors in LIFTS. A data management plan, describing how LIFTS data are shared and stored, is in place.

Results from LIFTS will be shared and discussed with the other LIFTS partners (practical education schools, industry partners, and societal stakeholders). However, these partners do not have access to the collected data. A consortium agreement in which the roles and responsibilities of the LIFTS parties are described is signed by all LIFTS partners.

Data are collected, processed, and stored according to the LIFTS data management plan that has been developed according to the WUR privacy policy and regulations. The research described in this protocol neither aims to collect data on confidential or sensitive issues nor to collect special categories of personal data. However, confidential or sensitive issues may be mentioned by the participants. To guarantee confidentiality, names and other information that can be used to trace back to individuals will not be used in the transcription of interviews

XSL•FO
RenderX

or observation notes, and audio recordings will be permanently deleted once transcription is complete. Data will be saved in a file only accessible by the researchers involved in LIFTS.

Compensation Details

It is important to properly value the youth's participation and time investment in this research project. However, it is still contested what a suitable (monetary) remuneration is for youth participation [15]. To determine a suitable remuneration for practical education students involved, we will consult the recommendations of the Netherlands Organization for Health Research and Development (ZonMw; ZorgOnderzoek Nederland / Medische Wetenschappen [37]) and determine it together with practical education staff, practical education students, and other LIFTS researchers.

Results

In March 2023, funding approval for LIFTS was received, and in October 2023, the 5-year project officially started. In January 2024, the author DvU of the studies in this protocol was employed and started to familiarize herself with the practical education context, the project's aims, and the procedures for ethical approval. In March 2024, ethical approval was obtained for the qualitative research (study 2). Data collection for study 2 was started in September 2024 (the beginning of the school year) and ended in January 2025. In total, 179 practical education students participated in the data collection (23 at School A, 91 at School B, and 65 at School C) as part of the needs assessment. Additionally, 17 practical education staff members have been interviewed (3 at School A, 8 at School B, and 6 at School C). Researchers kept a reflection diary during their data collection (from September to January), and in December 2024, the interview was held among the 4 PhD researchers involved in LIFTS. To get more familiar with the practical education school context and the LIFTS research activities and to get to know students and teachers, the author DvU conducted 11.5 hours of participant observation during the participatory research activities taking place at the schools. We have finalized the analysis and are currently, that is November 2025, writing an academic paper about this study. By the end of 2025, the researchers expect to wrap up the results of the needs assessment. The results of this study will also be shared with the LIFTS stakeholders in meetings and the newsletter and will inform the co-design of the health promotion programs of the other three PhDs.

On April 11, 2025, the protocol for the systematic literature review (study 1) was ready and submitted to PROSPERO. The search and selection of studies has started, and it is expected that the systematic literature review will be finished halfway through 2026.

The realist evaluation (study 3) is planned for the school year 2026-2027, when the other PhDs have co-designed an initial design of a health promotion program. At the moment, the other three PhDs are conducting the first cocreation sessions with the practical education schools on the design of the health promotion programs. We expect that they will have the first results in early 2026, and we will design the best evaluation strategy with their input.

Discussion

Anticipated Results

This research has a two-fold aim: (1) to gain insight into how to meaningfully involve youth in vulnerable positions in participatory research for health promotion by focusing on practical education students as a case study and (2) to gain insight into possible outputs and outcomes that result from a participatory approach. Given the limited understanding of these issues to date, this research contributes to the advancement of academic knowledge in the field of health promotion and participatory research with youth in vulnerable positions. We expect to deliver practical recommendations and lessons learned on how to actualize youth empowerment in participatory research for health promotion, which should inspire future researchers.

Additionally, we use the six-element framework for actualizing child and adolescent empowerment in participatory action research for health promotion, developed by Chrifou et al [14], as an analytical framework. This framework seems suitable as a theoretical framework for this research because it (1) allows us to operationalize and analyze the ambiguous concept of youth empowerment in the field of participatory research with youth for health promotion, which is the focus of this research; (2) takes context into account, including the social environment of youth; and (3) builds upon previous models conceptualizing participatory research with youth [17,38,39]. The framework has been published quite recently (2024) and, to our best knowledge, has not been practically assessed yet. This research could potentially add to this theoretical framework.

Next to its academic relevance, this research seeks to have a strong societal impact. First, this research aims to contribute to the health and well-being of practical education students. While practical education students show less healthy behavior than their peers in other types of Dutch secondary education [26], there are limited health promotion programs specifically designed for this target group. LIFTS aims to promote a healthy lifestyle among practical education students by developing and implementing health promotion programs on healthy and sustainable dietary behavior, physical activity, and mental health using a participatory approach. This research will contribute to this development by gaining insight into an empowering and meaningful participatory research approach in the practical education context. It explicitly allows moments for reflection and colearning—important aspects of participatory research [8]—that can guide further steps in the participatory process in LIFTS and, consequently, contribute to better health promotion programs for practical education students. In the long term, this research will hopefully contribute to a healthier lifestyle for practical education students and close the gap between practical education students and their peers in terms of health inequities. Second, young people are considered to have the right to influence decision-making and make possible contributions to change in processes that influence their lives [12,40,41]. One important possible benefit of youth participation in research is

XSL•FO
RenderX

youth empowerment [14]. As practical education students are less represented than their peers in other types of secondary education in the Netherlands, this research is expected to contribute to youth empowerment among practical education students through participatory research. Finally, practical education is still relatively unknown as a type of secondary education in the Netherlands. This is a missed opportunity, as the labor market needs many practically trained people [42]. Practical education trains students for professions in sectors with high demand in the Netherlands, such as construction, metalwork, landscaping, retail, hospitality, logistics, and health care [22]. As this research focuses specifically on practical education, it contributes to awareness raising and improved image about practical education, not only among the LIFTS stakeholders with whom we work directly but also among an academic audience when presenting the results of this study in papers and at conferences.

Strengths and Limitations

The first limitation of this research is its context-specific setting. This research is contextualized in an educational pathway in the Dutch education system. As education systems differ considerably between countries, it might be difficult to generalize results directly to different contexts. However, participatory research inherently prioritizes bottom-up and locally rooted research approaches and close collaboration with the community in a colearning process [9]. Generalization is therefore not the main goal of this research. A strength is that we strive to provide an in-depth case study of the practical education context by collaborating with multiple practical education schools, knowledge institutes, and societal and industrial partners. The lessons learned from this research can inspire other researchers and stakeholders working in similar contexts. Nonetheless, the systematic literature review (study 1) helps us to link the empirical results of studies 2 and 3 to the academic literature.

The second limitation is the anticipated low participation of caregivers in the research. As students with learning difficulties seem to experience more social challenges [43], indicating that a good social network seems to be especially important for this group, it is crucial to involve practical education students' caregivers and teachers in the participatory research process. Until now, we have noticed that it is difficult to involve caregivers, and we expect this might be a challenge throughout LIFTS. To encourage caregivers' participation in participatory research, the literature emphasizes the need for offering practical opportunities to do so, such as evening focus groups for working caregivers and the provision of childcare or meals [44]. Regarding informed consent, research highlights the importance of providing sufficient information to caregivers, as adults can prevent youth from participating [12]. The strong collaboration with schools and societal partners is a strength, as it can foster recruitment of caregivers in LIFTS and limit the risk of missing an important social environment for practical education students in the design of health promotion programs. Specifically, collaborating with the schools' community liaisons (brugfunctionaris in Dutch) is an opportunity, as they bridge the gap between schools and caregivers.

Comparison With Prior Work

To our knowledge, there has been no prior research focusing on participatory research for health promotion in the practical education context. However, participatory research for health promotion has been conducted with somewhat similar youth groups. Prior research exists with VMBO, which, similar to practical education, has a more practical-oriented focus than other educational pathways. For instance, Boonekamp and colleagues [45] have conducted participatory research for physical activity promotion among VMBO students. One of their conclusions is that the development of school-based physical activity programs drawing on active student involvement allows students to shape their physical activity, and it enhances their motivation to engage in physical activity. Another example is the research project VMBO Sportlab [46], which studies how to sustainably enhance the physical activity promotion of VMBO students together with researchers, societal partners, and youth themselves.

An example of a similar methodology to the present research is described by Maenhout and colleagues [21]. The authors share lessons learned from the cocreation of an intervention for physical activity promotion with youth with intellectual disabilities in Belgium. One of the main differences is that the authors rely on researcher reflections, while we also include youth, staff member, and researcher perspectives in study 2. The authors argue that there are still few examples of lessons learned from the participatory research process with youth for health promotion to inspire future researchers. This research starts from the same premise.

Conclusions

No prior participatory research has been conducted with practical education students for health promotion program design and implementation. The development of tailored health promotion programs for this target group is important to contribute to health equity in the Netherlands. Participatory research is a promising methodology, but we lack knowledge on how to do so in a meaningful and empowering way, as well as of the possible outputs and outcomes of the programs that result from a participatory approach. This research aims to shed light on these issues, contributing to both the academic literature on health promotion and participatory research with youth in vulnerable positions as well as promoting health equity in the Netherlands.

Acknowledgments

We would like to thank all practical education students, practical education staff members, LIFTS researchers, and partners for their participation. This research is not possible without them. The authors used the generative artificial intelligence (AI) tool

XSL•FO
RenderX

ChatGPT (OpenAI) for language editing to improve the readability of the paper. The produced suggestions were always judged and checked by the authors. The authors take full responsibility for the content and interpretation of this paper.

Funding

This project is funded by the Netherlands Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek; NWO) grant number KICH1.GZ03.21.011).

Authors' Contributions

Conceptualization: DvU, AW, MP, MK, KV
Formal analysis: KV and DvU
Funding acquisition: KV, AW, MK, MP
Investigation: DvU
Methodology: DvU, AW, MP, MK, KV
Project administration: KV and DvU
Supervision: KV, AW, MK, MP
Writing–original draft: KV and DvU
Writing–review & editing: DvU, AW, MP, MK, KV

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from the Netherlands Organization for Scientific Research (NWO).
[PDF File (Adobe PDF File), 211 KB - resprot_v15i1e84723_app1.pdf]

References

1. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile Health Interventions for Improving Health Outcomes in Youth: A Meta-analysis. JAMA Pediatr 2017 May 01;171(5):461-469 [FREE Full text] [doi: 10.1001/jamapediatrics.2017.0042] [Medline: 28319239]
2. Ferretti A, Hubbs S, Vayena E. Global youth perspectives on digital health promotion: a scoping review. BMC Digit Health 2023 Jul 25;1(1):461-469 [FREE Full text] [doi: 10.1186/S44247-023-00025-0]
3. The National Prevention Agreement. Ministry of Health, Welfare and Sport. 2019. URL: https://www.government.nl/ministries/ministry-of-health-welfare-and-sport/documents/reports/2019/06/30/the-national-prevention-agreement [accessed 2025-12-17]
4. World report on social determinants of health equity. World Health Organization. URL: https://www.who.int/teams/social-determinants-of-health/equity-and-health/world-report-on-social-determinants-of-health-equity [accessed 2025-08-05]
5. Delaruelle K, Van Houtte M, Bracke P. Educational inequalities in general health: does the curricular tracking system matter? Acta Sociologica 2019;63(1):63-81. [doi: 10.1177/00016993188171O1]
6. Macintyre AK, Torrens C, Campbell P, Maxwell M, Pollock A, Biggs H, et al. Socioeconomic inequalities and the equity impact of population-level interventions for adolescent health: an overview of systematic reviews. Public Health 2020 Mar;180:154-162 [FREE Full text] [doi: 10.1016/j.puhe.2019.11.008] [Medline: 31923881]
7. Ozer EJ, Abraczinskas M, Duarte C, Mathur R, Ballard PJ, Gibbs L, et al. Youth Participatory Approaches and Health Equity: Conceptualization and Integrative Review. Am J Community Psychol 2020 Dec;66(3-4):267-278. [doi: 10.1002/ajcp.12451] [Medline: 32969506]
8. Lenette C. Participatory Action Research: Ethics and Decolonization. New York, NY: Oxford University Press; 2022.
9. Yesudhas R, Nikku B. Reimagining Community and Knowledge: Emerging Directions in Participatory Research. International Journal of Qualitative Methods 2025 Jun 19;24:1-2. [doi: 10.1177/16094069251354847]
10. Geurts EMA, Reijs RP, Leenders HHM, Jansen MWJ, Hoebe CJPA. Co-creation and decision-making with students about teaching and learning: a systematic literature review. J Educ Change 2023;25(1):103-125. [doi: 10.1007/s10833-023-09481-x]
11. Kennedy H, DeChants J, Bender K, Anyon Y. More than Data Collectors: A Systematic Review of the Environmental Outcomes of Youth Inquiry Approaches in the United States. Am J Community Psychol 2019 Mar;63(1-2):208-226. [doi: 10.1002/ajcp.12321] [Medline: 30843254]
12. Kim J. Youth Involvement in Participatory Action Research (PAR). Critical Social Work 2019 May 30;17(1):37-53. [doi: 10.22329/csw.v17i1.5891]
13. Ozer EJ. Youth-Led participatory action research. In: Jason JA, Glenwick DS, editors. Handbook of Methodological Approaches to Community-Based Research: Qualitative, Quantitative, and Mixed Methods. New York, NY: Oxford University Press; 2016:263-272.

XSL•FO
RenderX

14. Chrifou R, Anselma M, Christens BD, Israel BA, Jurkowski JM, Perkins DD, et al. Actualizing child and adolescent empowerment in participatory action research for health promotion: a six-element framework. International Journal of Adolescence and Youth 2024 May 16;29(1):2354907. [doi: 10.1080/02673843.2024.2354907]
15. Cullen O, Walsh CA. A Narrative Review of Ethical Issues in Participatory Research with Young People. YOUNG 2019 Dec 22;28(4):363-386. [doi: 10.1177/1103308819886470]
16. Bradbury-Jones C, Isham L, Taylor J. The complexities and contradictions in participatory research with vulnerable children and young people: A qualitative systematic review. Soc Sci Med 2018 Oct;215:80-91. [doi: 10.1016/j.socscimed.2018.08.038] [Medline: 30218806]
17. Cahill H, Dadvand B. Re-conceptualising youth participation: A framework to inform action. Children and Youth Services Review 2018 Dec;95:243-253. [doi: 10.1016/j.childyouth.2018.11.001]
18. Freire K, Pope R, Jeffrey K, Andrews K, Nott M, Bowman T. Engaging with Children and Adolescents: A Systematic Review of Participatory Methods and Approaches in Research Informing the Development of Health Resources and Interventions. Adolescent Res Rev 2022 Feb 03;7(3):335-354 [FREE Full text] [doi: 10.1007/s40894-022-00181-w] [Medline: 38845155]
19. Seidl K, Wahl A, Rojatz D, Affengruber L, Sommer I, Klerings I, et al. How Best to Perform Co-Research With Children and Adolescents: A Rapid Qualitative Evidence Synthesis. International Journal of Qualitative Methods 2025 Jun 21;24:36. [doi: 10.1177/16094069251346879]
20. Moore LW, Miller M. Initiating research with doubly vulnerable populations. Journal of Advanced Nursing 1999 Nov;30(5):1034-1040. [doi: 10.1046/j.1365-2648.1999.01205.x] [Medline: 10564401]
21. Maenhout L, Verloigne M, Cairns D, Cardon G, Crombez G, Melville C, et al. Co-creating an intervention to promote physical activity in adolescents with intellectual disabilities: lessons learned within the Move it, Move ID!-project. Res Involv Engagem 2023 Mar 19;9(1):10 [FREE Full text] [doi: 10.1186/s40900-023-00420-x] [Medline: 36935503]
22. Praktijkonderwijs [Practical education]. Sectorraad Praktijkonderwijs. URL: https://www.praktijkonderwijs.nl/praktijkonderwijs/ [accessed 2025-08-05]
23. EQF and NLQF, the European Qualifications Framework. Europass. URL: https://www.europass.nl/eqfnlqf/ [accessed 2025-09-05]
24. Kengetallen praktijkonderwijs [Key figures for practical education]. Sectorraad Praktijkonderwijs. URL: https://www.praktijkonderwijs.nl/kengetallen-praktijkonderwijs/ [accessed 2025-08-05]
25. De kracht van praktijkonderwijs [The power of practical education]. Sectorraad Praktijkonderwijs. URL: https://www.praktijkonderwijs.nl/wp-content/uploads/2023/06/De-kracht-van-praktijkonderwijs.pdf [accessed 2023-08-05]
26. Regionale resultaten van de gezondheidsmonitors van GGD Noord- en Oost-Gelderland [Regional results of the health monitors of the Municipal Health Service Noord- en Oost Gelderland]. GGD Noord- en Oost-Gelderland. URL: https://kompas.ggdnog.nl/mosaic/dashboard/gezondheidsmonitor-jeugd-praktijkonderwijs [accessed 2025-09-05]
27. Health promotion. World Health Organization. URL: https://www.who.int/health-topics/health-promotion#tab=tab_1 [accessed 2025-11-10]
28. Bronfenbrenner U. Toward an experimental ecology of human development. American Psychologist 1977 Jul;32(7):513-531. [doi: 10.1037/0003-066x.32.7.513]
29. Guy-Evans O. Bronfenbrenner's Ecological Systems Theory. Simply Psychology. 2024. URL: https://www.simplypsychology.org/bronfenbrenner.html [accessed 2025-08-05]
30. Kmet LM, Cook LS, Lee RC. Standard quality assessment criteria for evaluating primary research papers from a variety of fields. University of Alberta. 2004. URL: https://era.library.ualberta.ca/items/48b9b989-c221-4df6-9e35-af782082280e/download/a1cffdde-243e-41c3-be98-885f6d4dcb29 [accessed 2025-08-05]
31. Schelbe L, Chanmugam A, Moses T, Saltzburg S, Williams LR, Letendre J. Youth participation in qualitative research: Challenges and possibilities. Qualitative Social Work 2014 Nov 10;14(4):504-521. [doi: 10.1177/1473325014556792]
32. Aldridge J. The participation of vulnerable children in photographic research. Visual Studies 2012 Mar;27(1):48-58. [doi: 10.1080/1472586x.2012.642957]
33. Beuving J, de Vries G. Doing Qualitative Research – The Craft of Naturalistic Inquiry. Amsterdam: Amsterdam University Press; 2015.
34. Pawson R, Tilley N. Realistic Evaluation. London: Sage; 1997.
35. van Belle S, Westhorp G, Marchal B. Realist evaluation. Better Evaluation. URL: https://www.betterevaluation.org/methods-approaches/approaches/realist-evaluation [accessed 2025-08-05]
36. Verkooijen KT, Super S, Mulderij LS, De Jager D, Wagemakers A. Using realist interviews to improve theory on the mechanisms and outcomes of sport for development programmes. Soc Incl 2020;8(3):152-161. [doi: 10.17645/si.v8i3.2747]
37. Tips voor betekenisvolle jongerenparticipatie in onderzoek [Tips for meaningful youth participation in research]. ZonMw. 2023. URL: https://www.zonmw.nl/sites/zonmw/files/2023-12/019_093_JEUGDPARTICIPATIE_2023_WEB_DEF.pdf [accessed 2025-08-05]
38. Shier H. Pathways to participation: openings, opportunities and obligations. Children & Society 2001;15(2):107-117. [doi: 10.1002/chi.617]

XSL•FO
RenderX

39. Wong NT, Zimmerman MA, Parker EA. A typology of youth participation and empowerment for child and adolescent health promotion. Am J Community Psychol 2010 Sep;46(1-2):100-114 [FREE Full text] [doi: 10.1007/s10464-010-9330-0] [Medline: 20549334]
40. Simmons C, Robinson S. Strengthening participation of children and young people with disability in advocacy. Children and Young People With Disability Australia. 2014. URL: https://cyda.org.au/wp-content/uploads/2023/08/strengthening_participation.pdf [accessed 2025-08-05]
41. Convention on the Rights of the Child. UNICEF. 2017. URL: https://www.unicef.org/child-rights-convention/convention-text [accessed 2025-12-17]
42. Naamsbekendheid [Brand awareness] 2/7. Sectorraad Praktijkonderwijs. URL: https://www.praktijkonderwijs.nl/speerpunten/naamsbekendheid-praktijkonderwijs/ [accessed 2025-08-05]
43. Pearl R, Donahue ML. Learning about learning disabilities. In: Wong B, editor. Peer Relationships and Learning Disabilities. London: Academic Press; 2004:133-165.
44. Bermúdez Parsai M, Castro FG, Marsiglia FF, Harthun ML, Valdez H. Using community based participatory research to create a culturally grounded intervention for parents and youth to prevent risky behaviors. Prev Sci 2011;12(1):34-47 [FREE Full text] [doi: 10.1007/s11121-010-0188-z] [Medline: 21107693]
45. Boonekamp GMM, Dierx JAJ, Jansen E. Motivating students for physical activity: What can we learn from student perspectives? European Physical Education Review 2020 Nov 23;27(3):512-528. [doi: 10.1177/1356336x20970215]
46. Wij brengen vmbo'ers aan het bewegen [we get VMBO students moving]. Vmbo Sportlab. URL: https://www.vmbosportlab.nl/ [accessed 2025-08-05]

Abbreviations

LIFTS: Healthy Lifestyle for Low-Literate teenagers
NWO: Nederlandse Organisatie voor Wetenschappelijk Onderzoek
TU/e: Eindhoven University of Technology
VMBO: Voorbereidend Middelbaar Beroepsonderwijs
UU: Utrecht University
WUR: Wageningen University and Research
WUR-REC: Wageningen University & Research Research Ethics Committee
ZonMw: ZorgOnderzoek Nederland / Medische Wetenschappen

Edited by J Sarvestan; submitted 24.Sep.2025; this is a non–peer-reviewed article;accepted 11.Nov.2025; published 14.Jan.2026.

Please cite as:
van Uden D, Wagemakers A, Peeters M, Kleinjan M, Verkooijen K
Promoting Mental Health, Physical Activity, and Healthy and Sustainable Dietary Behavior Among Practical Education Students in the Netherlands: Protocol for a Multiphase Participatory Research Study
JMIR Res Protoc 2026;15:e84723
URL: https://www.researchprotocols.org/2026/1/e84723
doi:10.2196/84723
PMID:

©Daniëlla van Uden, Annemarie Wagemakers, Margot Peeters, Marloes Kleinjan, Kirsten Verkooijen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 14.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Integrating Physical Health Dimensions in Middle-Aged and Older Adults to Predict Drug Misuse and Relapse Through a Comprehensive Needs Model: Protocol for a Mixed Methods Study

Ema Madyaningrum[1]; Eny Purwandary[2]; Supriyati Supriyati[3]; Ariani Arista Putri Pertiwi[4]; Made Satya Nugraha Gautama[5]; Nurlaela Widyarini[6]

[1]Department of Mental Health and Community Nursing, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Sleman, Yogyakarta, Indonesia

[2]Faculty of Psychology, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia, Surakarta, Center of Java, Indonesia

[3]Department of Health Behavior, Environment, and Social Medicine, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah, Yogyakarta, Indonesia, Sleman, Yogyakarta, Indonesia

[4]Department of Basic Nursing and Emergency; Faculty of Medicine, Public Health, and Nursing; Universitas Gadjah Mada, Yogyakarta, Indonesia, Sleman, Yogayakrta, Indonesia

[5]Department of Nursing, Faculty of Medicine, Universitas Pendidikan Ganesha, Singaraja, Indonesia, Singaraja, Bali, Indonesia

[6]Faculty of Psychology, Universitas Muhammadiyah Jember, Jember, Indonesia, Jember, East Java, Indonesia

Corresponding Author:
Ema Madyaningrum
Department of Mental Health and Community Nursing
Faculty of Medicine, Public Health, and Nursing
Universitas Gadjah Mada
Jl. Farmako, Sekip Utara, Kec. Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta 55281
Sleman, Yogyakarta, 55281
Indonesia
Phone: 62 877378907554
Fax: 62 274545674
Email: ema_nursing@ugm.ac.id

Abstract

Background: Drug misuse poses significant health risks, with rising prevalence globally, particularly affecting middle-aged and older adults. In Indonesia, drug misuse has complex impacts across physical, psychological, social, and spiritual dimensions. Addressing these issues is crucial as Indonesia transitions to an aging population.

Objective: This study protocol aims to develop a comprehensive model incorporating physical health dimensions to predict drug misuse and relapse among middle-aged and older adults, grounded in the health belief model.

Methods: We will use a sequential mixed methods design, combining quantitative data from national databases and medical records with qualitative insights from stakeholders. The respondents will consist of three groups: (1) individuals with substance use disorders recruited from the Badan Narkotika Nasional (National Narcotics Agency), (2) participants from *lembaga pemasyarakatan* (penitentiary institutions) and *balai pemasyarakatan* (community correction centers), and (3) stakeholders involved in drug rehabilitation services and experts in the field of education. Data analysis will be conducted in sequential stages, including descriptive statistics, bivariate analysis, and logistic regression to identify predictors of relapse. Qualitative data will be analyzed using thematic analysis to explore stakeholders' perspectives.

Results: This study was funded by the Ministry of Education, Culture, Research, and Technology of Indonesia in 2024. Data collection started in July 2024 and ended in December 2025. The study is structured into 3 phases. Upon completion, it is expected to support the development of a comprehensive model incorporating physical health dimensions to predict drug misuse and relapse among middle-aged and older adults, grounded in the health belief model. The findings are expected to be published by the end of 2026.

Conclusions: The study is expected to identify key health status indicators and levels of readiness to change for the prevention of drug misuse relapse, as well as to develop a preventive model that addresses the specific needs of vulnerable populations, thereby enhancing the effectiveness of intervention strategies. By using a holistic approach to address the multifaceted impacts of drug misuse, this study will aim to improve health outcomes for middle-aged and older adults.

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/88987

(JMIR Res Protoc 2026;15:e88987) doi:10.2196/88987

KEYWORDS

substance abuse; relapse; middle age; older adults; health belief model

Introduction

People who misuse drugs are among the most vulnerable populations for health issues, with their numbers continuing to rise. Globally, from 2011 to 2021, there was an estimated 5.85% increase in drug misuse, reaching approximately 296 million people [1]. In Indonesia, the prevalence of people who misuse drugs was 1.95% as of 2021 [1]. The types of drugs used are influenced by age, socioeconomic status, and prevailing drug trends at a given time. Illicit drugs commonly used by individuals in Indonesia include cannabis, amphetamines, and nitrazepam (known locally as *nipam* or *koplo*) [2].

Drug misuse leads to complex physical, psychological, social, and spiritual impacts. Physically, it can cause overdoses, drug poisoning, sexually transmitted diseases, liver cirrhosis, cancer, hypertension, and respiratory diseases. In Indonesia, 65.8% of users experience severe addiction [3]. Psychologically, it may result in anxiety, depression, suicidal thoughts, sleep disorders, and posttraumatic stress disorder [4-7]. Social impacts include strained family relationships, divorce, social stigma, and legal issues [8,9]. Spiritually, there is often a decline in religious engagement [4].

Drug misuse is prevalent among individuals aged 15 to 64 years, with adolescents traditionally representing the highest-use demographic [1]. However, there is a growing number of users in older age groups [10-12]. In Indonesia, adults aged 25 to 49 years are the most affected [2], whereas in the United States, opioid use is high among women aged 68 to 77 years [13]. In a previous study in Indonesia, 11.27% of middle-aged and older adults from rehabilitation centers reported drug misuse [3]. Such trends highlight the need to focus on different at-risk groups, especially as Indonesia transitions to an older population structure, with more than 10% of its population comprising older adults since 2021. This shift necessitates addressing the health challenges of an aging population, including the rising number of older people who misuse drugs, which can degrade their overall health status. To improve life quality for older adults, it is crucial to prevent drug misuse. A holistic prevention strategy addressing physical, psychological, social, and spiritual needs is essential [14-16]. Since 2016, Indonesia's National Narcotics Agency (Badan Narkotika Nasional; BNN) has developed the Drugs Rehabilitation Information System (Sistem Rehabilitasi Narkoba; SIRENA) to provide real-time updates on drug misuse nationally, aiding in intervention development.

Using the extensive data available from this system offers a significant opportunity to advance our understanding of drug misuse relapse prevention and to better identify the health care needs of vulnerable older adult populations. As individuals age, they encounter a multitude of health challenges, ranging from chronic illnesses to mobility issues, which can be further complicated by drug misuse. This interrelation underscores the critical need for focused and detailed research aimed specifically at these groups. This study aims to explore several key areas regarding drug misuse among middle-aged and older adult populations. First, it seeks to identify the physical health profile of people who misuse drugs within these groups. Second, the study will examine the risk and protective factors related to physical and biological health that may influence drug misuse relapse. Furthermore, it aims to determine the specific physical and biological health needs that can prevent relapse and anticipate drug misuse among middle-aged and older adults. Finally, the study will intend to develop a health model based on the health belief model to effectively prevent relapse in these vulnerable populations.

Methods

Study Design

This research will use a sequential mixed methods design, combining quantitative and qualitative research approaches [17] using the health belief model approach (Figure 1). The results from the quantitative phase will inform and guide the implementation of the qualitative phase in the next stage of the study [16].

XSL•FO
RenderX

Figure 1. Approach based on the health belief model.

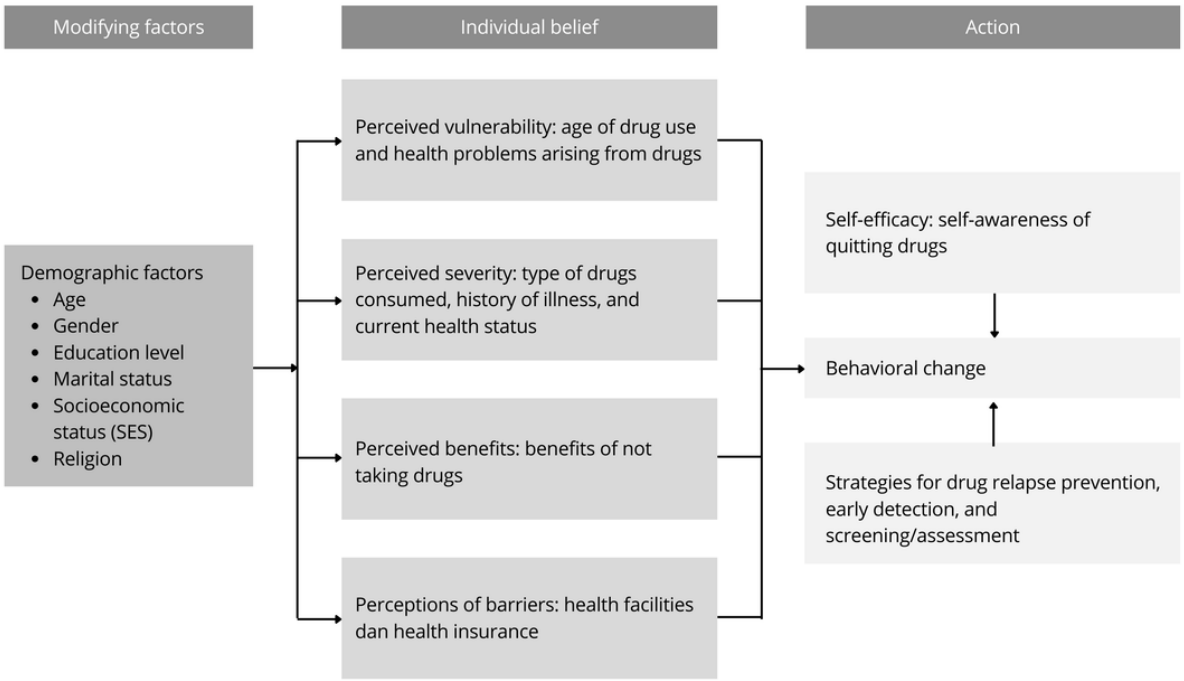

Participants and Recruitment

Phase 1 and 2 (Quantitative Phase)

A total of 355 records will be assessed from the BNN. Furthermore, 381 records will be assessed from penitentiary institutions or *lembaga pemasyarakatan* (LPs) and community correction centers or *balai pemasyarakatan* (BPs). For the quantitative analysis, we used secondary data from the Indonesian government's drug rehabilitation records. These resources originate from the BNN, which serves as the central agency for narcotics regulation; LPs, which provide drug rehabilitation services to prisoners who have been criminalized for drug-related offenses; and BPs, which offer rehabilitation for individuals preparing to reintegrate into their communities.

The database for the years 2022 to 2024 contains detailed information on individuals involved in drug misuse, the services they received, and their health status as recorded in medical records. The medical records will focus on physical health assessments, treatment interventions, and instances of relapse.

Inclusion and Exclusion Criteria

Participant eligibility criteria are presented in Textbox 1.

Textbox 1. Participant eligibility criteria.

Inclusion criteria

- People with a history of drug misuse aged ≥40 years who have previously visited polyclinics or been hospitalized for drug-related problems
- People with a record of health checks
- People with documentation of any illnesses that occurred during their rehabilitation or recovery process

Exclusion criteria

- Incomplete medical record data

The sampling technique used in this study is total sampling, whereby all data that meet the inclusion and exclusion criteria will be included in the analysis.

Phase 3 (Qualitative Phase)

For the qualitative data collection, participants will be selected using purposive sampling techniques. This selection will focus on individuals who have expertise or experience in dealing with drug misuse–related issues. Participants will include stakeholders such as health workers (doctors and nurses) and rehabilitation officers. Their involvement will be determined based on recommendations from the rehabilitation leader of the Provincial National Narcotics Board, ensuring that the participants have relevant knowledge and insights related to the study's focus. The interviews will take place either in person or via videoconferencing and will last approximately 60 to 90 minutes. With participants' consent, each interview will be audio-recorded and then transcribed for detailed analysis.

XSL•FO
RenderX

Variables and Measurements

Phase 1

The first phase aims to explore several key aspects of drug misuse among middle-aged and older adult populations. Specifically, it seeks to identify the physical health profiles of individuals with substance use disorders within these age groups. On the basis of this objective, the study further examines participants' perceptions of their own health status.

To understand the influence of various factors on self-health perception, we apply the health belief model. The modifying factors include demographic variables such as age, gender, education, occupation, ethnicity, marital status, living arrangements, and religion. The perceived susceptibility factors include being a drug dealer, types of drugs used, routine drug misuse, excessive drug misuse, and the impact of these behaviors on occupational functioning and interpersonal relationships. Examining these variables enables the identification of behavioral patterns and their impact on self-assessment of health. The perceived severity variables include length of conviction, recidivism status, and excessive drug misuse.

Phase 2

The second phase aims to examine risk and protective factors related to physical and biological health that may influence drug misuse relapse. Prediction of relapse will be assessed through readiness to change for discontinuing drug misuse. The readiness to change is determined by the application of the transtheoretical model, which explains the principle of change, specifically behavioral change [18,19]. This variable is measured by the University of Rhode Island Change Assessment Scale [20]. The University of Rhode Island Change Assessment Scale has been used to assess people who participated in the rehabilitation program at the BNN [21]. From the total score, the variable will be categorized to precontemplation (score <8), contemplation (8-11), action (11-14), and maintenance (>14).

The independent variables include the demographics of respondents, such as gender, age, education background, marital status, employment status, and religion. Other variables include drug choice; the Addiction Severity Index, Fifth Edition [22]; risk for substance use as measured by the Alcohol, Smoking, and Substance Involvement Screening Test [23]; and rehabilitation monitoring, which encompasses health status, psychological status, social status, and environmental status.

Phase 3

Finally, the study aims to develop a health model grounded in the health belief model to effectively prevent relapse among these vulnerable populations.

Statistical Analysis

Phases 1 and 2

Descriptive statistics will be used to summarize demographic characteristics, substance use history, patterns of drug misuse, and health perceptions among individuals with substance use disorders, and bivariate analyses will be conducted to examine associations between independent and dependent variables. Furthermore, inferential statistical analyses, specifically logistic regression, will be performed to identify predictors of health outcomes and drug misuse relapse while controlling for demographic variables. A P value <.05 will be considered the threshold for statistical significance throughout the analysis.

Phase 3

Thematic analysis will be conducted to identify and analyze patterns within the qualitative data gathered from interviews. Transcripts will be coded inductively to develop themes that reflect participants' experiences and insights. NVivo software (Lumivero) will be used to assist in the organization and analysis of qualitative data.

Ethical Considerations

Ethics approval was obtained from the institutional review board of Medical Health Research Ethics Committee of Faculty of Medicine, Public Health, and Nursing Universitas Gadjah Mada—Dr Sardjito General Hospital (ref KE/FK/1652/EC/2024). Informed consent will be secured from all interview participants, emphasizing the voluntary nature of participation and the right to withdraw at any time. All data will be treated confidentially to ensure participant anonymity.

Results

This study was funded by the Ministry of Education, Culture, Research and Technology of Indonesia in 2024. Data collection started in July 2024 and ended in December 2025. The study is structured into 3 phases. In the last phase of the study, findings from phases 1 and 2 will be used to develop a comprehensive model incorporating physical health dimensions to predict drug misuse and relapse among middle-aged and older adults, grounded in the health belief model. Currently, the research is in the data analysis phase, and the publication of findings is anticipated by the end of 2026.

Discussion

Anticipated Findings

This study aims to develop a comprehensive model incorporating physical health dimensions to predict drug misuse and relapse among middle-aged and older adults, grounded in the health belief model. In the initial phase, the study hypothesizes that illicit drug misuse is associated with perceptions of health status. This phase will also identify factors that influence the health status of individuals who use illicit drugs. In the second phase, the study will assess the readiness of people who misuse drugs to undergo rehabilitation and discontinue illicit drug use. In the final phase, the findings from the previous stages will be integrated to develop a comprehensive model incorporating physical health dimensions to predict drug misuse and relapse among middle-aged and older adults.

This study protocol outlines the approach to developing a comprehensive strategy for drug misuse prevention and treatment among middle-aged and older adults. Recognizing the rising prevalence of drug misuse among these age groups, particularly in Indonesia, this study will target specific interventions in response to demographic shifts.

XSL•FO
RenderX

The Indonesian government has 4 formal drug rehabilitation institutions: government hospitals, LPs, BPs, and the BNN. Although they collaborate in providing drug rehabilitation to people with substance use disorders, they each have different responsibilities. LPs are responsible for drug rehabilitation and social rehabilitation of prisoners who have legal problems related to narcotic use, such as narcotic distributors, sellers, and dealers, after which they will be returned to society. BPs accompany them during the preparation process for returning to society. The next institution is BNN, whose drug misuse rehabilitation service includes inpatient and outpatient services for people who misuse drugs with mild health problems. The last institution is hospitals. people who misuse drugs with health-related problems or drug addition receive more services from hospitals [24,25].

In our study, we will use a mixed methods approach to thoroughly investigate the factors influencing drug misuse, including demographic, behavioral, and psychosocial elements. By incorporating physical health dimensions into our model, we aim to provide a detailed understanding of self-health perception and its impact on drug misuse patterns. The logistic regression analysis will help identify significant predictors of relapse, offering a basis for tailored intervention strategies. Additionally, by involving stakeholders such as health care providers and rehabilitation officers, the study will ground its findings in practical experience, ensuring the proposed health model addresses real-world needs and challenges.

As Indonesia's population ages, addressing health issues related to drug misuse among middle-aged and older adults will become increasingly vital. This study will inform public health policies and intervention programs, aiming to improve their effectiveness and relevance. By prioritizing comprehensive needs and a holistic approach, this protocol aims to pave the way for improving the quality of life for those at risk of drug misuse and relapse.

Strengths and Limitations

A major strength of this study is the involvement of multiple institutions engaged in the rehabilitation of individuals with drug misuse disorders. This multi-institutional approach enables a more comprehensive understanding of behavioral changes across different rehabilitation settings.

The limitation of this study is the restricted availability of variables related to the health status of individuals with substance use disorders, particularly data obtained from LPs and BPs. Additionally, a substantial proportion of health data from individuals undergoing rehabilitation at the BNN are incomplete because the implementation of standardized and comprehensive rehabilitation health assessments was still in its early stages in 2024.

Conclusions

The study is expected to identify key health status indicators and levels of readiness to change for the prevention of drug misuse relapse, as well as to develop a preventive model that addresses the specific needs of vulnerable populations, thereby enhancing the effectiveness of intervention strategies. By using a holistic approach to address the multifaceted impacts of drug misuse, this study will aim to improve health outcomes for middle-aged and older adults with a history of drug misuse.

Funding

This study was supported by research grants from Ministry of Education, Culture, Research, and Technology of Indonesia.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Acknowledgments

This research is a consortium study involving 3 universities in Indonesia: the Department of Mental Health and Community Nursing, Faculty of Medicine, Public Health, and Nursing Universitas Gadjah Mada; the Faculty of Physiology, Universitas Muhammadiyah Surakarta; and Universitas Muhammadiyah Jember. The authors would like to express their sincere gratitude to the leadership of these institutions for their support and facilitation of this research. They would also like to extend their appreciation to the leadership and staff of the Badan Narkotika Nasional, *lembaga pemasyarakatan*, and *balai pemasyarakatan* for their willingness to participate in and support this research. The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing, translation, and reformatting. The GAI tool used was GPT-4.5. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Authors' Contributions

EM, EP, and NW developed the research concept and study design. EM was responsible for project administration, data collection, and writing the research report. EP and NW reviewed the manuscript. SS contributed to data collection and manuscript review. AAPP contributed to writing and reviewing the manuscript. EM and MSNG drafted the manuscript. SS contributed to writing and reviewing the manuscript.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. World Drug Report 2023. United Nations Office on Drugs and Crime. URL: https://www.unodc.org/res/WDR-2023/WDR23_Exsum_fin_DP.pdf [accessed 2026-03-06]
2. Health Research Of Drugs Abuse Effects 2019. Badan Narkotika Nasional. 2019. URL: https://ppid.bnn.go.id/konten/unggahan/2020/10/Health-Research-On-Drug-Abuse-2019.pdf [accessed 2026-03-06]
3. Dewabhrata W, Ahsan A, Bella A, Amalia N, Kusuma D, Pertiwi YBA. Mental health, environmental, and socioeconomic geographic factors of severe drug addiction: analysis of rehabilitation center data in Indonesia. Subst Abuse 2023;17:11782218231203687 [FREE Full text] [doi: 10.1177/11782218231203687] [Medline: 37822734]
4. Piacentine LB. Spirituality, religiosity, depression, anxiety, and drug-use consequences during methadone maintenance therapy. West J Nurs Res 2013 Jul;35(6):795-814. [doi: 10.1177/0193945913479452] [Medline: 23446494]
5. Schulte MT, Hser Y. Substance use and associated health conditions throughout the lifespan. Public Health Rev 2013 Dec 13;35(2):1-2. [doi: 10.1007/bf03391702]
6. Belfrage A, Mjølhus Njå AL, Lunde S, Årstad J, Fodstad EC, Lid TG, et al. Traumatic experiences and PTSD symptoms in substance use disorder: a comparison of recovered versus current users. Nordisk Alkohol Nark 2023 Feb 04;40(1):61-75 [FREE Full text] [doi: 10.1177/14550725221122222] [Medline: 36793481]
7. Dell'Aquila A, Berle D. Predictors of alcohol and substance use among people with post-traumatic stress disorder (PTSD): findings from the NESARC-III study. Soc Psychiatry Psychiatr Epidemiol 2023 Oct 03;58(10):1509-1522 [FREE Full text] [doi: 10.1007/s00127-023-02472-6] [Medline: 37133523]
8. Daley DC. Family and social aspects of substance use disorders and treatment. J Food Drug Anal 2013 Dec;21(4):S73-S76 [FREE Full text] [doi: 10.1016/j.jfda.2013.09.038] [Medline: 25214748]
9. Hawlader MDH, Nabi MH, Hussain A, Al Amin SU, Zaman S, Masud I. Legal and social consequences of substance use: results from a nationwide study in Bangladesh. J Ethn Subst Abuse 2020 Dec 20;21(4):1-11. [doi: 10.1080/15332640.2020.1846008] [Medline: 33342393]
10. Lin J, Arnovitz M, Kotbi N, Francois D. Substance use disorders in the geriatric population: a review and synthesis of the literature of a growing problem in a growing population. Curr Treat Options Psychiatry 2023 Jun 05;10(3):1-20 [FREE Full text] [doi: 10.1007/s40501-023-00291-9] [Medline: 37360959]
11. Yarnell S, Li L, MacGrory B, Trevisan L, Kirwin P. Substance use disorders in later life: a review and synthesis of the literature of an emerging public health concern. Am J Geriatr Psychiatry 2020 Feb;28(2):226-236. [doi: 10.1016/j.jagp.2019.06.005] [Medline: 31340887]
12. Drugs and associated issues among young people and older people. In: Drugs and Age. Austria: The United Nations Office on Drugs and Crime (UNODC); 2018.
13. Vasilenko SA, Evans-Polce RJ, Lanza ST. Age trends in rates of substance use disorders across ages 18-90: differences by gender and race/ethnicity. Drug Alcohol Depend 2017 Nov 01;180:260-264 [FREE Full text] [doi: 10.1016/j.drugalcdep.2017.08.027] [Medline: 28938183]
14. Høyland SA, Schuchert A, Mamen A. A holistic perspective on continuing care for substance use and dependence: results and implications from an in-depth study of a Norwegian continuing care establishment. Nordisk Alkohol Nark 2022 Oct 23;39(5):503-520 [FREE Full text] [doi: 10.1177/14550725221099702] [Medline: 36284745]
15. Fauzia A, Rahayu AYS, Harahap AA. Analisis sistem informasi rehabilitasi narkoba dengan metode hot fit pada klinik BNN guna menunjang ketahanan nasional. J Natl Resil Strateg Stud 2021 Jun 25;4(1):1-2 [FREE Full text] [doi: 10.7454/jkskn.v4i1.10044]
16. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Kolner Z Soz Sozpsychol 2017 Jul 5;69(Suppl 2):107-131 [FREE Full text] [doi: 10.1007/s11577-017-0454-1] [Medline: 28989188]
17. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Thousand Oaks, CA: Sage Publications; 2009.
18. Hashemzadeh M, Rahimi A, Zare-Farashbandi F, Alavi-Naeini A, Daei A. Transtheoretical model of health behavioral change: a systematic review. Iranian J Nursing Midwifery Res 2019;24(2):83. [doi: 10.4103/ijnmr.ijnmr_94_17]
19. Hunter H. The transtheoretical model of behavior change: implications for social work practice. J Hum Behav Soc Environ 2023 Sep 12;34(2):300-314. [doi: 10.1080/10911359.2023.2253307]
20. Richards DK, Schwebel FJ, Cowie KD, Pearson MR, Witkiewitz K. Measurement invariance of the University of Rhode Island Change Assessment Scale in Project MATCH: an exploratory structural equation modeling approach. Alcohol Clin Exp Res 2022 Jul 25;46(7):1258-1267 [FREE Full text] [doi: 10.1111/acer.14849] [Medline: 35488452]
21. Badan Narkotika Nasional. Asesmen Dan Rencana Intervensi Paba Lembaga Rehabilitasi Yang Diselenggaraiian Oleh Masyaraiiat Bagi Pecandu Dan Korban Penyalahgunaan Narkotika. Jakarta, Indonesia: Deputi Bidang Rehabilitasi; 2017. URL: https://perpustakaan.bnn.go.id/sites/default/files/Buku_Digital_2020-10/BK0114_Asesmen_dan_Rencana_Intervensi_pada_Lembaga_Rehabilitasi_yan.pdf [accessed 2026-03-06]
22. Denis CM, Cacciola JS, Alterman AI. Addiction Severity Index (ASI) summary scores: comparison of the Recent Status Scores of the ASI-6 and the composite scores of the ASI-5. J Subst Abuse Treat 2013 Nov;45(5):444-450 [FREE Full text] [doi: 10.1016/j.jsat.2013.06.003] [Medline: 23886822]

XSL•FO
RenderX

23. The Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST) Manual for Use in Primary Care. World Health Organization. URL: https://iris.who.int/bitstream/handle/10665/44320/9789241599382_eng.pdf [accessed 2026-03-06]
24. Martin M, Snyder HR, Otway G, Holpit L, Day LW, Seidman D. In-hospital substance use policies: an opportunity to advance equity, reduce stigma, and offer evidence-based addiction care. J Addict Med 2022 Aug 2;17(1):10-12. [doi: 10.1097/adm.0000000000001046]
25. Hoover K, Lockhart S, Callister C, Holtrop JS, Calcaterra SL. Experiences of stigma in hospitals with addiction consultation services: a qualitative analysis of patients' and hospital-based providers' perspectives. J Subst Abuse Treat 2022 Jul;138:108708 [FREE Full text] [doi: 10.1016/j.jsat.2021.108708] [Medline: 34991950]

Abbreviations

BP: balai pemasyarakata
BNN: Badan Narkotika Nasional
LP: lembaga pemasyarakatan
SIRENA: Sistem Rehabilitasi Narkoba (Drugs Rehabilitation Information System)

Edited by J Sarvestan; submitted 07.Dec.2025; this is a non–peer-reviewed article;accepted 03.Mar.2026; published 19.Mar.2026.

Please cite as:
Madyaningrum E, Purwandary E, Supriyati S, Pertiwi AAP, Gautama MSN, Widyarini N
Integrating Physical Health Dimensions in Middle-Aged and Older Adults to Predict Drug Misuse and Relapse Through a Comprehensive Needs Model: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e88987
URL: https://www.researchprotocols.org/2026/1/e88987
doi:10.2196/88987
PMID:

©Ema Madyaningrum, Eny Purwandary, Supriyati Supriyati, Ariani Arista Putri Pertiwi, Made Satya Nugraha Gautama, Nurlaela Widyarini. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Understanding the Factors Explaining the Growing Use of Medical Assistance in Dying in Québec: Protocol for an Interdisciplinary Mixed Methods and Multimethods Study

Marie-Ève Bouthillier[1,2*], PhD; Isabelle Marcoux[3*], PhD; Catherine Perron[4*], PhD; Bruno Gagnon[5,6], MD, MSc; David Lussier[7], MD; Ghislaine Rouly[2]; Mathieu Moreau[1], MD, MSc; Michel Dorval[8,9], PhD; Sabrina Lessard[10,11], PhD; Dominique Girard[12], PhD; Gina Bravo[13], PhD; Maude Hébert[14], PhD; Michelle Giroux[15], LLL, MA; Simon Lemyre[1], MA; Alexandra Beaudin[3], MSc; Claude Julie Bourque[1,16], PhD; Maryse Soulières[17], PhD; Maude Lévesque[11,18], PhD; Mona Gupta[19], MD, PhD; Valérie Bourgeois-Guérin[20], PhD; David Lavoie[7], PhD; Bertrand Lavoie[21], LLD, MA; Ariane Plaisance[22], PhD; Louise Bernier[21], LLM, LLD; Jacinthe Dupuis[23], MCS

[1]Department of Family Medicine and Emergency Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada
[2]Research Centre, Centre Hospitalier de l'Université de Montréal, Montréal, QC, Canada
[3]Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada
[4]School of Social Work, Faculty of Arts and Humanities, Université de Sherbrooke, Sherbrooke, QC, Canada
[5]Department of Family Medicine and Emergency Medicine, Centre de recherche du CHU de Québec-Université Laval, Québec, QC, Canada
[6]VITAM Sustainable Health Research Center, Université Laval, Québec, QC, Canada
[7]Department of Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada
[8]Faculty of Pharmacy, Université Laval, Québec, QC, Canada
[9]Centre de recherche du CHU de Québec-Université Laval, Québec, QC, Canada
[10]Department of Anthropology, Faculty of Arts and Sciences, Université de Montréal, Montréal, QC, Canada
[11]Centre for Research and Expertise in Social Gerontology, Montréal, QC, Canada
[12]Department of Health Sciences, Université du Québec à Rimouski, Rimouski, QC, Canada
[13]Department of Community Health Sciences, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada
[14]Department of Nursing, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada
[15]Civil Law Section, Law Faculty, University of Ottawa, Ottawa, ON, Canada
[16]Centre de recherche en pédagogie de la santé, Faculté de médecine, Université de Montréal, Montréal, QC, Canada
[17]School of Social Work, Université de Montréal, Montréal, QC, Canada
[18]School of Social Work, Université du Québec à Montréal, Montréal, QC, Canada
[19]Department of Psychiatry and Addiction Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada
[20]Department of Psychology, Université du Québec à Montréal, Montréal, QC, Canada
[21]Law Faculty, Université de Sherbrooke, Sherbrooke, QC, Canada
[22]Qulysis, Québec, QC, Canada
[23]Centre Integre de Sante et de Services Sociaux de Laval, Laval, QC, Canada
[*]these authors contributed equally

Corresponding Author:
Marie-Ève Bouthillier, PhD
Department of Family Medicine and Emergency Medicine
Faculty of Medicine
Université de Montréal
2900 Boulevard Édouard-Montpetit
H3T 1J4
Montréal, QC
Canada
Phone: 1 (514) 343 5931
Email: marie-eve.bouthillier@umontreal.ca
Research Centre
Centre Hospitalier de l'Université de Montréal
900 R. Saint-Denis
Montréal, QC

XSL•FO
RenderX

Canada
Phone: 1 514 890 8000
Email: marie-eve.bouthillier@umontreal.ca

Abstract

Background: Medical assistance in dying (MAiD) became a legal end-of-life option on December 10, 2015, in Québec, and on June 17, 2016, in the rest of Canada. Since its legalization, there has been a steady increase in the number of MAiD requests and provisions. Across permissive jurisdictions, Québec now has the highest rate of assisted death. Despite the growing use of MAiD, research examining the factors driving this increase remains limited and fragmented. Existing studies offer partial and sometimes contradictory explanations, with little integration of legal, institutional, societal, and individual dimensions. Further research is needed to better understand the determinants of MAiD requests and practices, particularly in the Canadian and Québec contexts.

Objective: This research aims to understand the factors influencing changes in MAiD requests and administrations in Québec by examining laws, practices, societal perspectives, organization of care and services, and individual characteristics of those requesting MAiD, as well as their interrelationships. We present the protocol developed by the Consortium interdisciplinaire de recherche sur l'aide médicale à mourir, an interdisciplinary research consortium, including an international advisory committee, set up for this research.

Methods: The design of this protocol is multimethods and convergent mixed methods, including (1) an international cross-thematical approach with 4 main research methods (a scoping review, key informant interviews, focus groups with health care professionals, and a population-based survey) chosen to partially answer research questions across the entire study and to compare with other jurisdictions and (2) 11 theme-specific methods (including community forums, media coverage analysis, comparative legal analyses, case studies of triads, individual interviews, and system mapping) to enrich and complement findings from the cross-thematical approach.

Results: When this 3-year funded study started in July 2024, several research methods not requiring ethics committee approval (because no human participants were involved) were initiated, including scoping and systematic reviews, media coverage analysis, and comparative legal analyses. By August 2025, interviews with key informants were completed, and analyses took place in September. Concurrently, other subteams started data collection (focus groups December 2025) or are getting ready to seek ethics approval for their protocols and data collection processes involving human participants: case studies of triads, individual interviews, and community forums.

Conclusions: Findings from the international cross-thematical approach and theme-specific methods will provide a comprehensive understanding of the factors influencing the use of MAiD in Québec. This study has strengths, including the use of a specific theoretical framework, a variety of complementary methods, and an integrated knowledge mobilization strategy. As for its limitations, we foresee challenges with the comparison of jurisdictions in terms of language, culture, and legal systems, as well as access to data about MAiD cases, since reporting systems may differ between jurisdictions.

International Registered Report Identifier (IRRID): DERR1-10.2196/83549

(JMIR Res Protoc 2026;15:e83549) doi:10.2196/83549

KEYWORDS

assisted dying related factors; Canada; end-of-life care; euthanasia; evolution of assisted dying practices; factors explaining use of MAiD; growing use of MAiD; medical assistance in dying; mixed and multimethod study; protocol; Québec

Introduction

Overview

In the province of Québec, medical assistance in dying (MAiD) has been permitted, as of December 2015, for individuals who meet the eligibility criteria specified in the Act respecting end-of-life care [1]. This legislation sets out a framework for "end-of-life care," which includes palliative care, continuous palliative sedation, as well as euthanasia (referred to in the Act and in this paper as MAiD) [1]. In Québec, MAiD consists "in the administration by a competent professional of medications or substances to a patient, at the patient's request, in order to relieve their suffering by hastening death." [2]. Unlike the Canadian law, Québec legislation excludes assisted suicide or self-administered MAiD [3].

Since 2015, there has been a steady year-over-year increase in the global number of requests and provisions of assisted dying [4]. In fact, Québec now ranks first in the world in terms of the proportion of MAiD deaths, ahead of other Canadian provinces, and even of countries where this practice has been permitted for decades [5,6]. In the September of 2023, at the request of the Québec Ministry of Health and Social Services, a call for proposals was launched by Québec's main funding agency (Fonds de recherche du Québec) to fund a research team willing to investigate the factors explaining this rise in MAiD incidence [7]. The study proposed by our consortium (Consortium interdisciplinaire de recherche sur l'aide médicale à mourir

XSL•FO
RenderX

[CIRAMM]; in English: Interdisciplinary Research Consortium on Medical Assistance in Dying) was successful in securing 3-year financial support [8]. The objectives of this paper are the following: (1) to briefly summarize current knowledge on the factors explaining the use of MAiD in Québec, (2) to specify the research objectives that led to the development of the CIRAMM to carry out this vast study, (3) to present the study's theoretical framework, (4) to describe its methodology, and (5) to discuss its strengths and limitations.

Factors Explaining the Use of MAiD

Studies on the evolution of the use of MAiD and potential explanatory factors remain sparse. Few authors propose hypotheses concerning the increase in MAiD practices at the international level. Boer [9] describes the development of medically assisted death in the Netherlands from 1968 onward, and reports that the number of euthanasia tripled between 2007 and 2021; a period during which eligibility criteria would have been interpreted more broadly, and mobile teams performing euthanasia without a prior doctor-patient relationship were set up. Still in the Netherlands, a study conducted by Groenewoud et al [10] revealed that age, church attendance, political orientation, income, health, and volunteer availability were associated with the geographical variation in the incidence of euthanasia over time (2013-2017).

Byram and Reiner [11] tested 10 hypotheses that could explain why the rates of deaths by MAiD are 15 times higher in Canada than those by assisted suicide in California, despite having adopted their respective laws in 2016. Three of them were confirmed: (1) limited public awareness, (2) fewer MAiD practitioners per capita, and (3) sparse support by health care institutions in California in comparison to Canada. On the other hand, Pullman [12] suggests, as possible hypotheses for the higher rate of MAiD in Canada, that the eligibility criteria are more ambiguous, their interpretation more flexible, and the method used (mainly injection in Canada rather than prescription of lethal drugs in California) implies more active participation of providers, who (even if that's not their intention), have tacit ascendancy over patients in vulnerable positions. According to Khakzedeh [13], other factors associated with the legislation and its implementation should be taken into account. For example, in jurisdictions like the United States, Switzerland, and Austria, assisted suicide is neither a care nor a right, and involves greater mobilization on the part of individuals regarding medical consultations, the purchase of medication, or notarial deeds [13].

In the Canadian context, recent analyses and opinion pieces on what triggers requests for MAiD bring divergent and controversial elements to the conversation. Coelho et al [14] link MAiD requests to suffering and lack of medical, psychosocial, and rehabilitation support (disability support) due in part to structural problems. The lack of access to palliative care and housing alternatives is often cited in the media to explain the use of MAiD [15-19]. Alternatively, according to Downar et al [20], available evidence showed that people who died by MAiD had higher socioeconomic status, and most of them benefited from the involvement of palliative care providers. In Québec, the idealization of "a good death," social secularization, and the widespread emphasis on self-determination have been proposed to explain the increase in the use of MAiD [21,22]. The "quest for autonomy" and the reappropriation of power in the intimate sphere could be explained by the retreat (or rejection) of the argument of authority [22] and, in Québec in particular, the decline of religion [23]. Even with "increasing social acceptance of this care" [24] in Québec, as elsewhere, studies highlight the lack of public knowledge of MAiD in relation to other end-of-life practices [25,26]. Overall, the practice of euthanasia and assisted suicide is still relatively recent in Canada. Studies are needed to better understand the use of these means to end one's life.

Research Objectives

The call for proposals issued by the Fonds de recherche du Québec aimed at better understanding the use of MAiD in the Québec context. It included a series of stated objectives reproduced in Textbox 1.

XSL•FO
RenderX

Textbox 1. Research themes and objectives.

Personal characteristics

- Identify the characteristics of people who request medical assistance in dying (MAiD)
- Identify how social determinants of health influence the personal experience of MAiD
- Investigate the influence of vulnerability factors on personal experience
- Investigate the relationship between the nature of suffering and one's desire to obtain MAiD
- Investigate how personal experiences of support throughout the MAiD process influence one's motivation to consider MAiD

Organization and delivery of health care in Québec

- Demonstrate how Québec's end-of-life care legislative provisions that pertain to the organization of health care are associated with the increased use of MAiD
- Explain how differences in the availability of end-of-life care in different regions and institutions influence the use of MAiD
- Describe the relationship between professional practice in MAiD and the evolution of requests for MAiD in Québec

Societal aspects

- Determine whether sociolegal factors explain the growing use of MAiD
- Examine whether the patient (family)/professional relationship has an impact on MAiD requests and administration
- Explore the factors influencing the social acceptability of MAiD
- Understand local and regional disparities in MAiD with respect to their sociodemographic characteristics

Practices, legislation, and public policies

- Compare and contrast Québec's MAiD legislation and public policies with those of other jurisdictions
- Examine how MAiD numbers in Québec compare to those in other jurisdictions
- Explore how Québec's context can explain differences in MAiD use

Regulation of advance requests for MAiD

- Compare the provisions for advance MAiD requests set out in Québec's Act respecting end-of-life-care with those in effect in jurisdictions that already allow MAiD for incompetent people
- Document experiences and practices related to advance MAiD requests in these other jurisdictions

Theoretical Framework

The Bronfenbrenner ecological systems theory [27] serves as the theoretical framework for this protocol (Figure 1). This theory aims to better understand a phenomenon (namely, the use of MAiD) by situating the individual in relation to the various proximal and distal environments that exert a direct or indirect influence on them. This will enable us to take into account all the specific objectives related to the 5 themes (Textbox 1), each being linked to one or more systems. Figure 1 illustrates how the Bronfenbrenner ecological systems theory will be applied to better understand the use of MAiD in the Québec context.

First, the Ontosystem refers to a person's internal characteristics, states, strengths, and vulnerabilities. A better understanding of what characterizes people who use MAiD will enable us to address the objectives of Theme 1 relating to personal characteristics (eg, health status, social determinants of health, vulnerability factors, personal values, conception of quality of life, and relationship to suffering).

Second, the Microsystem is characterized by the person's immediate environment, and includes the individuals with whom the person has a direct relationship (eg, relatives and health care professionals), as well as their immediate living environment, including the care setting (private dwelling, nursing home, hospital, or palliative care home). The contribution of this system to the explanation of the phenomenon will be addressed by exploring, for example, the influence of social determinants of health (eg, social support; Theme 1) and the relationship between patients and health care professionals on MAiD requests or administration (Theme 3).

Third, the Exosystem encompasses all the environments that have an indirect influence on the person. This will enable us to better understand the disparities in the use of MAiD, while considering factors associated with the organization of care and services, such as characteristics of the end-of-life care offer, depending on facilities and regions (Theme 2).

Fourth, the Macrosystem corresponds to society's cultural matrix (eg, beliefs, values, norms, and ideologies) and includes the laws, rules, and policies that govern a society. A better understanding of the use of MAiD as a function of the Macrosystem will be addressed through several themes' objectives, such as Theme 3 (eg, Québec societal and legal factors, social acceptability, and evolution of roles in

XSL•FO
RenderX

decision-making), Theme 4 (eg, comparative analysis of laws and public policies and factors associated with the evolution of the use of MAiD in different jurisdictions), and Theme 5 (international experiences with advance requests for MAiD [AR-MAiD]).

Fifth, the Chronosystem allows for considering the temporal elements that mark the evolution of MAiD use, depending on different factors, such as the changes in the course of the disease, the changes in institutional policies, the date of coming into force of MAiD legislation, and relevant amendments.

Finally, the dotted line running through the systems represents the bidirectional influence—upward or downward—between society, its systems, and its individuals, thus demonstrating their interdependence.

Figure 1. Bronfenbrenner ecological systems theory to better understand the use of medical assistance in dying in the Québec context.



Methods

Study Design

The diversity of research objectives calls for the use of a multimethods and convergent mixed methods design. Two complementary approaches were adopted to address the research objectives: (1) an international cross-thematical approach comprising 4 research methods and (2) 11 theme-specific methods devoted to each theme. An overview of the research methods developed for this study (n=15) is provided in Textbox 2. The different methods will be deployed in such a way that the preliminary results will inform the development of data collection tools for subsequent theme-specific methods.

XSL•FO
RenderX

Textbox 2. Overview of the methods according to each theme (n=15).

All themes

- Cross-thematical methods
- Jurisdictions: Canada, Belgium, Switzerland, the Netherlands, the United States (California), and Australia (Victoria).
- Scoping review
- Population-based survey
- Key informant interviews
- Focus groups with health care professionals

Personal characteristics

- Theme-specific methods
- Individual interviews
- Case studies of triads
- Retrospective qualitative chart review

Organization and delivery of health care in Québec

- Theme-specific methods
- Document analysis
- System mapping

Societal dimensions

- Theme-specific methods
- Media coverage analysis
- Community forums

Practices, legislation, and public policies

- Theme-specific methods
- Comparative legal analysis on assisted dying
- Realist synthesis

Regulation of advance requests for medical assistance in dying (MAiD)

- Theme-specific methods
- Comparative legal analysis of advance requests for MAiD
- Mixed-literature systematic review

Cross-Thematical Methods

Scoping Review

The scoping review will be conducted following the Joanna Briggs Institute methodology for scoping reviews and reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [28,29]. The scoping review will follow these steps: (1) translating the overarching objective of the study, which is to better understand the use of MAiD in Québec, into an operationalized research question, formulated to make the population, concept, and context explicit; (2) specifying inclusion and exclusion criteria; (3) in collaboration with a librarian, developing the strategy to identify relevant literature, with adaptation for 10 databases (eg, MEDLINE, CINAHL, Embase, PsycINFO, Web of Science, and gray literature); (4) conducting the selection and extraction through Covidence (Veritas Health Innovation) [30] by 2 independent reviewers; (5) presenting the findings in a narrative format, and mapping them in tables and graphs to address the review objectives; and (6) drafting recommendations and implications for research and practice in a language adapted to different target audiences (general population, decision-makers, and practitioners), and identifying future research needs.

Population-Based Survey

Study Design

A cross-sectional population-based survey on the social acceptability of MAiD in Québec, 2 other Canadian provinces, and 5 international jurisdictions (the Netherlands, Belgium, Switzerland, California [US], and Victoria [Australia]) will be conducted. In Canada, British Columbia and Ontario were

XSL•FO
RenderX

selected for their demographic significance, as together with Québec, these provinces accounted for 75% of the national population in 2024 [31]. In British Columbia, the MAiD death rate ranked second after Québec, while Ontario's rate was below the national average [6]. For their part, international jurisdictions were selected for having (1) at least 5 years of assisted dying legal permissibility (as of 2024) since a sufficiently long period is required to analyze potential explanatory factors in its growing use, and (2) different forms of assisted dying (euthanasia and physician-assisted suicide) for comparability purposes. The survey will be conducted among the adult population (18 years of age or older), in the official languages of the selected jurisdictions (either French, English, Dutch, German, or Italian). The target sample size is calculated according to the following parameters: a 95% CI, a maximum proportion of 0.50, a margin of error of ±5%, multiplied by the number of jurisdictions (n=8). A larger sample of respondents will be selected for Québec to allow regional comparisons (expected total n=3580).

Data Collection

The questionnaire, based on and adapted from questionnaires designed by CIRAMM members [25,32-34], will include acceptability and knowledge of MAiD, palliative care, and continuous palliative sedation, previous experiences with MAiD and with other end-of-life care, social determinants of health, and sociodemographic characteristics. Special consideration will be given to tailoring the level of language and MAiD-related terminology to the specific context of each jurisdiction, while preserving thematic consistency across all settings. The survey will be conducted online via a web panel designed by the chosen polling firm (to be determined).

Data Analysis

Data will be weighted according to sociodemographic characteristics to adjust results to represent the target population. Usual descriptive statistics will be used to summarize the factors associated with the acceptability of MAiD and with the use of MAiD for the selected jurisdictions. Multivariable regressions will be conducted to identify which factors (knowledge, previous experiences, social determinants of health, or sociodemographic characteristics) are associated with the acceptability of MAiD and contribute to the differences in use of MAiD across jurisdictions.

Key Informant Interviews

Study Design

To complement the results of the scoping review with data specific to Québec and each abovementioned jurisdictions, a qualitative descriptive study will be carried out on the factors associated with the growing use of MAiD. A total of 50 key informants will be selected for their expertise and experience in the field of MAiD in the abovementioned jurisdictions. We will include all MAiD program managers in Québec, a population of approximately 30 individuals, in order to gain a comprehensive understanding of the situation across the province. In addition, we will include members of MAiD regulatory authorities (eg, members of an oversight body or professional college) or people actively involved in MAiD practice (eg, MAiD program manager or experienced provider). Selection will be done by a nonprobabilistic maximum variation purposive sampling method to obtain rich and meaningful information, as well as to capture the breadth of different experiences and perspectives [35].

Data Collection

We will conduct semistructured interviews with participants in order to gain insight into potential explanatory factors for the use of MAiD. The interview guide, available in French and English, will be pretested and adapted in an evolving and iterative process throughout the data collection process [36]. The interview guide is available in Multimedia Appendix 1. Recruitment in Québec will take place through networks already established by CIRAMM members (ie, communities of practice of Québec MAiD providers and support structures for MAiD, among others) [37-42]. In other provinces of Canada and internationally, we will use a snowball approach [43] and mobilize our many Canadian and international networks: medical associations and federations, professional regulatory bodies, etc [44-47]. Interviews will be conducted by trained interviewers via the online Microsoft Teams platform. They will be recorded, and research notes will be taken by the interviewers to synthesize important ideas.

Data Analysis

The transcription process will leverage Teams' automatic transcription system, followed by editing by the interviewers. A descriptive thematic analysis will be conducted on the transcripts and research notes [48,49], with an iterative review process to refine codes based on the interview guide. Coding disagreements will be discussed between coders until a consensus is reached. The QDA Miner software (Provalis Research) [50] will support the coding process, allowing for the identification of recurring themes across participants' discourse.

Our study is guided by the quality criteria articulated in Tracy's [51] qualitative research framework. Specifically, our approach is informed by the 8 key markers: worthy topic, rich rigor, sincerity, credibility, resonance, significant contribution, ethics, and meaningful coherence, while allowing flexibility, reflexivity, and interpretive dialogue throughout the research process (as for all the qualitative analyses included in this protocol).

Focus Groups With Professionals Involved in MAiD

Study Design

We will explore, with focus groups, the viewpoints of professionals involved in MAiD in the same jurisdictions as the population-based survey and interviews with key informants.

Participants

Physicians and nurse practitioners involved in assessing, providing, or supporting families accessing MAiD will be recruited (nurses, social workers, spiritual care workers, ethicists, etc). Professionals must have participated in at least 10 MAiD cases or have been involved in MAiD for at least 2 years. We will recruit the professionals using a nonprobabilistic purposive sampling to obtain maximum variation (gender, age, experience, location, and type of practice, etc) [35].

XSL•FO
RenderX

Data Collection

A total of 18 online focus groups, each comprising of 6-8 participants, and lasting about 2 hours, will be organized as follows: 4 groups in Québec (physicians and nurse practitioners: n=2; other professionals: n=2); 4 groups in the rest of Canada (physicians and nurse practitioners: n=2; other professionals: n=2); 2 groups in each international jurisdiction (Netherlands: n=2, Belgium: n=2, Switzerland: n=2, California: n=2, and Victoria: n=2). The international focus groups will consist of only physicians because of the limited involvement of other health care professionals in MAiD. We will construct the facilitation guide (with the help of the international advisory committee for jurisdictions other than Québec) based on information extracted from the key informant interviews. The facilitation guide will be pretested with a group of 4 people familiar with the reality of MAiD, but not approached as potential participants (Multimedia Appendix 2). A preliminary version of the facilitation guide can be found in Multimedia Appendix 1. For recruitment, we will use the same methods as for the key informant interviews. Focus groups will be conducted virtually, recorded, and transcribed via Microsoft Teams.

Data Analysis

Analysis of the qualitative data from the focus groups will follow the same procedure as for the key informant interviews.

Theme-Specific Methods

Overview

In this section, we will present the 5 themes that structure the call for proposals and their associated specific methods. For the sake of consistency, we have chosen to present them in accordance with the Bronfenbrenner ecological systems theory, from the microsystem (eg, personal characteristics) to the macrosystem (eg, practices, legislation, and public policies) as presented in Figure 1.

Theme 1: Personal Characteristics

This theme focuses on the personal characteristics of people who have requested MAiD, using qualitative methods to help understand the process that led to their request.

Study Design

The following three complementary methods will be used: (1) individual interviews with people requesting MAiD; (2) case studies of triads involving interviews with people requesting MAiD, one of their close-ones, and a member of their caregiving staff; and (3) a retrospective chart review.

Individual Interviews

Participants

We will conduct interviews inspired by narrative approaches of the "life story" [52-55] with approximately 30 people who have requested MAiD, diversifying the choice of participants according to their personal characteristics (age, gender, identity, cultural or ethnic origin, diagnoses, disability status, and socioeconomic situation).

Data Collection

A semidirective approach will be used to focus on the stories collected, while directing them toward the central themes, encouraging the emergence of new reflections. One or two interviews will be conducted with each participant (varying in length from 30 to 60 minutes, depending on the participant's health and ability to participate). The interviews will be recorded and transcribed integrally. Interviewers will keep a diary in order to note their impressions following each interview and use this information to enrich their analysis (emotional state of the participant, nonverbal cues during the interview, general context of the interview, etc). Recruitment will be based on the CIRAMM's contacts across the province (MAiD interdisciplinary support groups [ISGs], hospice care organizations, patient associations, etc) with the aim of reaching participants in different regions of Québec (urban, semiurban, and rural). During recruitment, we will explore the possibility of involving one of the participant's close-ones, as well as one of their care providers, in order to conduct a triadic case study (see "Case Studies of Triads" section below).

Data Analysis

A thematic analysis of the data [56,57] will be conducted according to the following steps: (1) interview verbatims and diaries will be read, (2) a sample of 3 transcripts will be randomly chosen and analyzed to produce a list of themes, and (3) the content of the remaining interviews and diaries will be classified from this list of themes using NVivo software (version 14; Lumivero; released in 2023). The quality of the codification process will be ensured by an interjudge agreement process in accordance with validity and quality criteria in qualitative research [58].

Case Studies of Triads

Overview

Many studies [59,60] focus on the individual perspectives of people who want to receive assistance in dying, their relatives, or caregiving staff. This approach does not reveal the relational dynamics at play within a request, the contrasts between different individual perspectives, and the way these factors influence each individual's perspective.

Participants and Data Collection

To overcome the shortcomings identified in current literature, we will carry out 10 triadic case studies, each comprising "contextualized narrative" interviews with the person requesting MAiD, one of their close-ones, and one of their care providers (doctor, nurse, allied health clinicians, care assistant, etc). This method was tested in another study conducted in the Netherlands by a coleader of this theme and her colleagues [61]. A different interview guide will be used for each subgroup, aimed at gaining a better understanding of the personal factors that motivate people to request MAiD from the perspective of the 3 subgroups. Each person will be interviewed individually.

Data Analysis

An analysis anchored in phenomenology, combining a longitudinal and multiperspective framework [62], as well as different aspects of reflective lifeworld research [63], will be

XSL•FO
RenderX

used [62,64,65]. The whole process will be iterative and guided by multiple discussions between the members of our research team [66].

Retrospective Qualitative Chart Review

Overview

The researchers will also undertake a retrospective chart review of MAiD requesters in a single institution in Montréal. The purpose of this study is to better understand the role of economic disadvantage (which is defined as dependence on social assistance as one's sole form of revenue) on these individuals' MAiD requests, particularly for those requesters whose natural death is not reasonably foreseeable (Track 2), as it is this group where public debate about the impact of economic disadvantage on requests has arisen. This part of the protocol draws upon similar methods used in 2 studies currently being carried out by CIRAMM researchers, the first involving a population of requesters in which the assessment required a psychiatric consultation [67], the second involving people with amyotrophic lateral sclerosis [68].

Data Collection and Analysis

We aim to review the charts of all Track 2 requesters from 2021 to 2025 who are experiencing economic disadvantage. We will obtain a list of all requesters at the identified institution. We will then do a preliminary review to determine the economic situation of each requester, identifying those in a situation of economic disadvantage. The charts of this subgroup will be subject to a complete chart review. First, we will extract data relating to each individual's sociodemographic profile: age, gender, social circumstance, diagnoses, stated reason for requesting MAiD, and whether MAiD was ultimately received. Second, we will review all notes pertaining to the person's request for an assisted death. We will extract all segments that refer to the person's statements about assisted dying, as well as the writer's views about those statements, if any. We will be particularly attentive to any notes relating to the person's economic situation and its potential impact on the experience of their disease, access to care, and suffering. In our analysis, we will focus on the manifest content of the extracted text. We will identify meaning units from the extracted segments, grouping them together in categories. Ultimately, we will seek patterns that link the categories, which describe and explain the role of a person's economic circumstances in their request for an assisted death, with their sociodemographic profiles, recognizing that a plurality of such narratives may emerge from these records [69].

Theme 2: Health Care Organization and Delivery in Québec

Overview

For this theme, we adopt a systematic comparison approach between Québec regions to examine the relationships between gaps in the use of MAiD and differences in sociodemographic profiles, in health services structures, practices, and perceptions. Québec regions will be further broken into territories corresponding to each of the 30 MAiD ISGs.

Study Design

A mixed methods approach will be implemented, including a document analysis and system mapping. Qualitative discourse analysis will be used to summarize government reports on MAiD (provincial and regional) and information documents available to the public, to explain how MAiD is presented and how resources are promoted in each region. Correlational analysis from MAiD reports and regional sociodemographics will be used to explore the strength and direction of links between regional MAiD rates and the social profile of the population. Both qualitative and quantitative results will be mobilized for a system mapping analysis to show possible causal loops between different categories of items.

Document Analysis

Document analysis is a systematic procedure for reviewing and synthesizing various sources (printed or digital), including their format dimensions (institutional authors, aims, and availability), and their discourse (themes, perspective, posture, and legitimation).

Data Collection

We will first create a mixed database of organizational documents, records, and reports produced by the ISGs on cases of refused, accepted, abandoned, and administered MAiD requests for the period 2019-2025 (case profiles including age of individuals, status of request, trajectory characteristics, access to alternative curative, palliative, and psychosocial care and services). Data gathered by the Québec End-of-Life Commission (Control Commission) will be considered, as well as sociodemographic data by region (population, age, socioeconomic profile, disease incidence, deaths, and causes of death, including rates of unassisted suicide). Other documents from gray literature will also be analyzed (local reports and registers of care and services from the ISGs, the End-of-Life Commission, and the Ministry of Health and Social Services, information and training documents, etc).

Data Analysis

In order to produce a general portrait of the evolution and current state of structures, resources, and practices in each region, the qualitative data will be subject to a documentary analysis using the read, extract, analyze, and distill method [70]. Quantitative data will be the subject of descriptive and comparative statistical analyses of regional realities and their evolution (averages, frequencies, and SDs), population representativeness of cases, as well as correlation analyses between the explanatory variables and the rate of increase in MAiD use. The analysis of the proportionality of the variables for each region will be used, among other things, to determine the associations and the relative weight of population aging in the use of MAiD, which is the main assumption made in the CSFV report [4].

System Mapping

Data Collection

Qualitative and quantitative analyses will be synthesized into a system mapping model, an approach for integrating systems analysis into the implementation of public and community health programs, care, and services using visual representations [71,72].

XSL•FO
RenderX

Data Analysis

This method has the advantage of facilitating the detailed description of complex systems by interested parties and producing a rigorous evaluation of the concrete impacts of an intervention, policy, or program and their local variations [73]. System mapping allows the description and comparison of systems and organization models to produce conceptual maps, diagrams, and causal loops reflecting complex decision-making and accessibility processes [74]. All these analyses will enable us to meet our objectives by modeling the organization of care and services for each region, creating detailed statistical reports as well as typologies of practice models and individual pathways [75].

Theme 3: Societal Aspects

Study Design

This theme encompasses 2 methods. First, media coverage analysis can provide access to the discursive and symbolic predominance of MAiD and, in doing so, shed light on how content contributes to ways of thinking, feeling, and acting about MAiD, and how we integrate it into our "own social value systems" [76]. Second, community forums (CFs) favor an open, participatory approach with citizens and take the pulse of the community's concerns, knowledge, and resources [77,78]. In addition to stimulating the exchange of ideas, this dynamic method brings together people from various backgrounds in a given region to hear and gather their opinions on a subject that concerns them, such as the use of MAiD.

Media Coverage Analysis

Data Collection

The main French- and English-language media published in each of the studied regions of Québec will be analyzed (Journal de Québec, Journal de Montréal, The Gazette, La Presse, Le Devoir, Hebdo Rive-Nord, Journal de Joliette, le Laurentien, le Soleil, and Le Nouvelliste et le Nord-Côtier) from 2014 (the year the Act respecting end-of-life care was adopted in Québec) to the present. A systematic search on Eureka will be carried out with the help of a librarian, by combining corollary terms referring to MAiD. All journalistic articles on MAiD will be included. Two people will select the texts and extract the data using Covidence [30].

Data Analysis

The qualitative discourse will be triangulated through an inductive thematic analysis using NVivo 14 software [79], and a lexicographic analysis [80] conducted with IRaMuTeQ (Pierre Ratinaud) [81]. Media texts, in the form of a textual corpus, will be statistically contrasted in the light of meaningful lexical categories derived from the software's statistical calculations (ie, correspondence factor analysis and top-down hierarchical classification method). This quantitative input will validate and/or complement the discourse analysis initially carried out.

Community Forums

Participants

Adult citizens of a given region will be invited to participate in CFs. The team will use a variety of means to encourage diversity in the participants' sociodemographic profile (gender, sex, age, sexual orientation, ethnic group, socioeconomic status, level of education, religious practices, beliefs, spirituality, and ability and disability). For each CF, we set no maximum number of participants, but anticipate a minimum of 20 representatives per region. The 6 CFs will be organized virtually or in person in the 3 regions with the highest annual proportion of MAiD deaths (Lanaudière, Bas-Saint-Laurent, and Capitale-Nationale), and the 3 regions with the lowest proportion of MAiD deaths (Montréal, Mauricie-Centre-du-Québec, and Côte-Nord) [4]. These regions also represent urban, semiurban, and rural environments. In order to represent and understand the Indigenous perspective on palliative care and MAiD, we plan to consult Nunavik citizens, in coconstruction with community representatives.

Data Collection

The facilitation guide will be elaborated on the basis of the results of the cross-thematical methods and in coconstruction with our partner organizations, patients, and relatives. Recruitment will be based on a call for applications via social media, health care associations, distribution lists of community, associative, and organizational partners, etc. Individuals will be able to register using an online form. Questions about their sociodemographic profile will enable us to assess the number and diversity of people attending, and to adjust accordingly. CFs will be conducted online via Microsoft Teams or face-to-face, and moderated by the researchers. Discussions will be recorded and automatically transcribed by Microsoft Teams, with subsequent verification. Observation notes will be taken by research assistants, and a logbook kept by each of the facilitating researchers will help contextualize the data.

Data Analysis

The data collected will be analyzed using an inductive thematic approach for each region. This qualitative analysis will then be compared with a quantitative content analysis under the same methodology as the media coverage analysis (ie, triangulated through IRaMuTeQ with a lexicographic analysis of the discussion transcriptions).

Theme 4: Comparison of Practices, Legislation, and Public Policies

Comparative Legal Analysis on Assisted Dying

Legal databases (eg, Société québécoise d'information juridique, Centre d'accès à l'information juridique, and La Référence for Québec; Canadian Legal Information Institute for the rest of Canada; and Lexis, Westlaw, and HeinOnline for selected international jurisdictions) will be searched to identify legal documents associated with MAiD legislation for the jurisdictions selected. The comparative positivist legal analysis (legislation, case law, and doctrine) carried out by our team of legal experts will enable us to identify the differences and similarities between laws and the scope of legislative criteria, and to grasp the appropriateness of their interpretation by contrasting them with the findings emerging from the scoping review. This work will be complemented by a comparative analysis of Québec public policies promoting the quality of life of sick people and their families (eg, home care [82] and informal care [83]), as well as

XSL•FO
RenderX

of the various legal mechanisms contained in the Québec legislation (MAiD, advance directives, palliative care, and continuous palliative sedation), while considering recent updates to the law. In addition, by compiling (and critically comparing) data from 2 complementary sources [84], we will be able to address "Examine how MAiD numbers in Québec compare to those in other jurisdictions" objective, which is to examine how MAiD numbers in Québec compare to those in other jurisdictions: (1) the annual reports of the Control Commissions of the selected jurisdictions based on official MAiD declarations, and (2) independent surveys conducted in certain jurisdictions and aimed at documenting the frequency of MAiD practices, in addition to other types of end-of-life medical practices [85,86].

Realist Synthesis

The final objective of Theme 4 (Textbox 1 [Explore how Québec's context can explain differences in MAiD use]) represents one of the overarching goals of the project, requiring the consideration of all results generated across themes, through both cross-thematic and theme-specific methods. To this end, we will draw on the realist synthesis method developed by Pawson et al [87]. The realist synthesis method addresses questions concerning what works for whom, in what circumstances, in what respects, and how a complex social intervention operates, such as a public policy, a health care system, or even a law [87-89].

Realistic synthesis is characterized by (1) the use of a conceptual framework and (2) the involvement of an advisory committee at various stages of the process. The conceptual framework chosen will combine the Bronfenbrenner ecological systems theory and that of the Health System Performance Measurement [90]. It will be adapted to the phenomenon under study (use of MAiD) and will provide a global understanding of the explanatory factors and the underlying mechanisms. In particular, the framework will consider (1) the characteristics and inputs of the legislation (macrosystem) and its operational functioning (2) the characteristics and social determinants of health related to individuals (microsystem) who use MAiD, (3) the outputs of the system (exosystem), which include access to and provision of MAiD and associated services, and (4) the results obtained (the use of MAiD in terms of requests and provisions). We will add the historical context (chronosystem) as a potential explanatory factor for the differences observed in the evolution of the use of MAiD (eg, legislative change initiated by whom and how).

A bipartite advisory committee will be set up, made up of (1) members of an international advisory committee and (2) representatives of the stakeholders (patients and their families, physicians, nurse practitioners, and other health care professionals, managers, and decision-makers) involved in MAiD in Québec. The involvement of the bipartite advisory committee will be particularly decisive in adapting the conceptual framework to the specific characteristics of the MAiD study, as well as in interpreting the results. The comparison of our data, based on the conceptual framework, will enable us to identify the contributing factors for which data already exist, as well as those for which further studies are required.

Theme 5: Regulation of AR-MAiD

Overview

On June 7, 2023, Québec's Act respecting end-of-life care was modified to allow AR-MAiD from people diagnosed with a disease leading to decisional incapacity [2]. Legal dispositions related to AR-MAiD came into effect on October 30, 2024. AR-MAiD may lead to further increases in the use of MAiD, given its widespread acceptability in Québec [26,33,91]. On the other hand, providing MAiD to patients who are no longer competent, based on their advance requests, can be challenging to implement. To inform the implementation of the AR-MAiD regime in Québec, a comparative analysis of the legal regimes of jurisdictions that already allow such requests and a systematic review of implementation challenges will be conducted.

Comparative Legal Analysis on AR-MAiD

A comparative portrait specific to AR-MAiD will be drawn from the legal analysis conducted under Theme 4. Results will highlight processes that led to AR-MAiD legalization in permissive jurisdictions and reasons underlying different legislative choices pertaining to eligibility criteria, procedural safeguards, and the assessment of suffering. Professional standards established by medical associations, learned societies, and regulatory bodies will complement our legal interpretation work [92]. Legal analysis specific to this theme will be anchored in a recent review of Dutch and Belgian AR-MAiD legislation [92], which we will extend to include Luxembourg, Colombia, Spain, and Québec.

Mixed-Literature Systematic Review

We aim to provide a comprehensive overview of international literature on the barriers and facilitators (both anticipated and experienced) related to AR-MAiD. Perspectives from patients, their relatives, health care professionals, legal experts, and scholars will be sought. Primary studies (qualitative, quantitative, and mixed methods studies), theoretical work, as well as relevant documents from gray literature (eg, theses and Control Commissions' reports) will be considered for inclusion. Peer-reviewed studies published after 2002 will be searched by a librarian in 10 electronic databases, with no restriction on language or origin of the authors. Screening of retrieved citations, assessment of the methodological quality of included studies, and data extraction will be conducted independently by pairs of researchers, supported by the software Covidence [30]. A preliminary search for publications on AR-MAiD revealed a predominance of qualitative studies, with a few surveys and mixed methods studies [93-95]. Hence, to enable synthesis and integration, quantitative data will be transformed into themes and subthemes in a convergent integrated approach. Identified barriers and facilitators to implementing AR-MAiD will be summarized following a content analysis of included documents, examined in light of the Québec regime, and contrasted according to (1) whether they are anticipated or truly experienced, (2) the perspective from which they originate (eg, that of health care professionals or relatives), (3) when they might arise along the MAiD process (request drafting, eligibility

XSL•FO
RenderX

assessment, or MAiD provision), and (4) prevailing legislation and codes of practice. We expect our legal analysis and systematic review to highlight significant challenges to implementing AR-MAiD in Québec and also possible ways to overcome them.

Governance and Methods Sequence

To meet the challenge of this mixed methods and multimethods study, CIRAMM rallied leading forces in the field of end-of-life care: 46 coresearchers (Québec, Canada, and international), 14 clinicians, 5 partner organizations, and 10 postdoctoral fellows and graduate students. Researchers are drawn from over 15 disciplines and work in various university departments, programs, and faculties: medicine (palliative care, family medicine, geriatrics, neurology, psychiatry, emergency medicine, and pediatrics), nursing, law, psychology, clinical ethics, epidemiology, social work, occupational therapy, pharmacy, philosophy, anthropology, religious sciences, and theology, sociology, and population health. We developed a governance structure to support CIRAMM members, enabling us to successfully coordinate the various research subteams (Figure 2). We set up an executive committee responsible for the operationalization and general coordination of CIRAMM. A steering committee made up of clinicians, researchers, and patient partners supports the executive committee in general decisions and orientations. A scientific committee brings together all researchers in charge of the specific themes, whose subteams will meet as projects progress. We also set up an international advisory committee made up of 10 collaborators (Canadian and international researchers) with expertise in euthanasia, assisted suicide, and end-of-life care.

Figure 2. Consortium interdisciplinaire de recherche sur l'aide médicale à mourir governance structure. MAiD: medical assistance in dying.

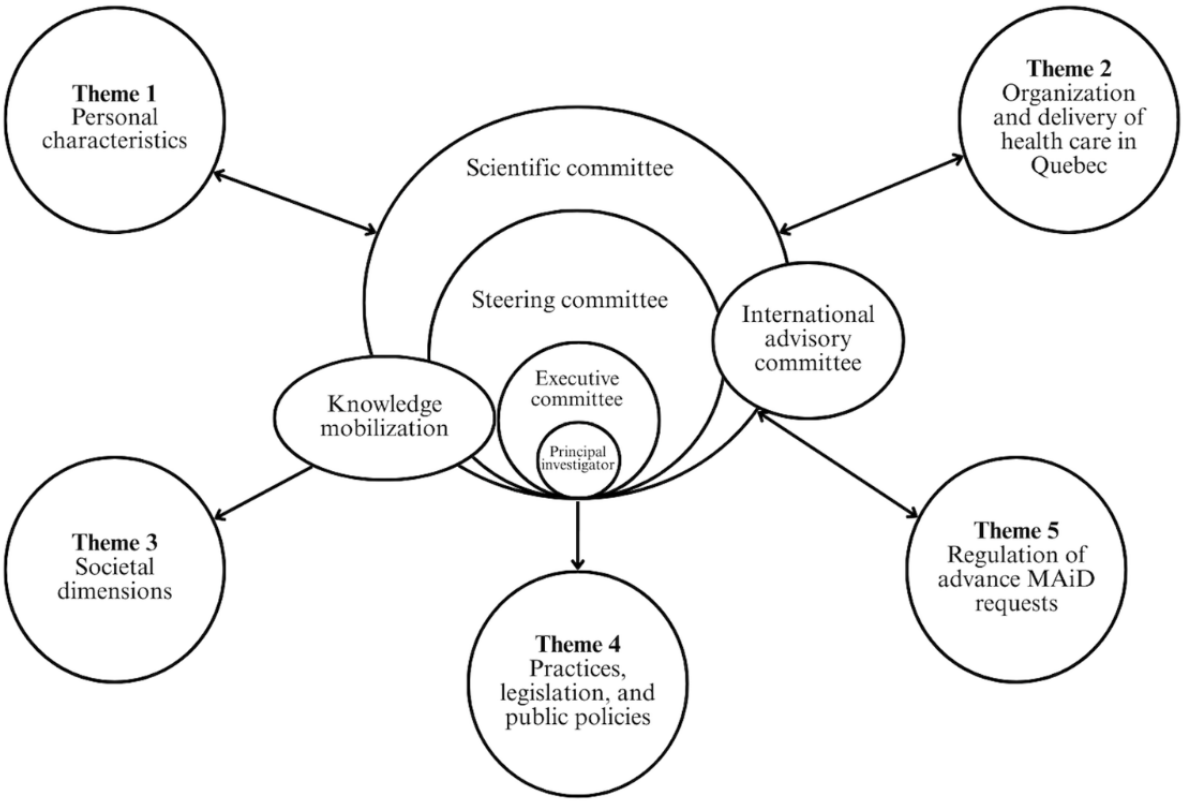

Each theme has a subteam responsible for its theme-specific methods. Coordination tools for the deployment sequence of methods are put in place to structure and adapt the activities of the study as it evolves.

Knowledge Mobilization

Inspired by the Knowledge-to-Action framework of Graham et al [96], we aim not only to produce appropriate knowledge adapted to the target audiences, but also to produce this knowledge by, for, and with them. The strategy put forward for this study is intended to be inclusive and collaborative at all stages of the research. To carry out this mission, we put together a knowledge mobilization committee, including patients, their close-ones, and key stakeholders, to play an anchoring role and to (1) support the subteam in the production of knowledge transfer (KT) tools, (2) plan the project's mobilization and KT activities with partners and users, (3) coordinate deliverables with preidentified milestones, (4) coordinate a unified plan for disseminating and promoting KT activities among partners, and (5) provide feedback on the activities developed, assessing whether the strategy's objectives have been met, while making adjustments where necessary.

Ethical Considerations

In Canada, ethical review by a research ethics board is a mandatory requirement prior to the commencement of any research involving human participants or human biological materials [97]. The ethics approval of the study will be done in stages, given the planned sequence of the various methods. So far, we have received ethical approval from the Scientific and

XSL•FO
RenderX

Research Ethics Committee of the Centre intégré de santé et de services sociaux de Laval (MP-35-2025-1112) and the Research Ethics Committee on Health Science at the University of Montréal (MP-35-2025-1112) for the cross-sectional methods of interviews with key informants and focus groups with professionals. As each methodological component involves distinct procedures, details regarding informed consent, data protection and confidentiality measures, and participant compensation will be described in the corresponding articles reporting on each method.

Results

The study was funded for 3 years and started in July 2024. We initiated projects that do not require the approval of an ethics committee because they do not involve human participants: scoping review, realist synthesis, both comparative legal analyses, mixed-literature systematic review, media coverage analysis, and document analysis. Detailed protocols for the scoping review and the mixed-literature systematic review have been drafted and submitted for publication.

The only method for which we have results to report to date is the interviews with key informants. We completed the interviews in August 2025 and the qualitative analysis in September 2025. The study sample included 63 participants, representing a diverse yet Québec-centered group of stakeholders involved in MAiD. Most participants were based in Québec (n=48), reflecting the primary focus of the study, while a smaller number came from other Canadian provinces (n=5) and international jurisdictions where assisted dying is permitted, including Belgium, the Netherlands, Switzerland, Australia, and the United States (n=2 each). In terms of roles related to MAiD, participants primarily included individuals responsible for MAiD support or coordination structures (n=38), followed by MAiD providers (n=15), experts in the field (n=14), and representatives of regulatory bodies (n=7). This composition ensured the inclusion of complementary perspectives spanning clinical practice, organizational coordination, expertise, and regulatory oversight. We are currently drafting a manuscript reporting the results for publication.

The focus groups are currently underway (December 2025). Ethical approval was obtained in August 2025. All the other subteams are preparing for ethics approval (protocol, data collection tools, consent forms, etc). This includes individual interviews, case studies of triads, retrospective qualitative chart review, CFs, and population-based survey methods.

Discussion

A Better Understanding of the Use of MAiD

This study is the first to comprehensively examine the use of MAiD in Québec. It builds on existing evidence and attempts to fill knowledge gaps. Empirical work aimed at explaining the use of MAiD remains uneven and conceptually dispersed. The literature tends to address isolated aspects of MAiD, often yielding divergent findings and rarely bridging legal frameworks, organizational contexts, societal dynamics, and patient-level factors. A more comprehensive and context-sensitive analysis is therefore required, particularly within the Canadian and Québec settings.

The results will provide a global understanding of the factors explaining the evolution of the use of MAiD, both by cross-thematical methods and by theme-specific methods. Our choice of the Bronfenbrenner ecological systems theory [27] will provide a better understanding of the use of MAiD by situating the individual in relation to the various proximal and distal environments that exert a direct or indirect influence on them.

Accordingly, our multiple methods will allow us to document each level—from the individual to the collective—while accounting for internal and external factors as well as the temporal dimension. We will be able to compare these data with those from other jurisdictions and to triangulate them in order to obtain a richer perspective on this phenomenon.

While focusing on explaining the high uptake of MAiD in Québec, findings could also provide insight into reasons for a decline in demand (if any) or lower rates of deaths by MAiD observed in other jurisdictions [84]. Several hypotheses exist to explain the increase in the use of MAiD, but much still needs to be done to understand this choice made by thousands of people every year.

Strengths and Limitations

The strengths of this study include (1) the use of a variety of methodologies, allowing for methodological triangulation that promotes reliability of the results; (2) the breadth of this interdisciplinary team with its combined set of skills and an international scientific committee ensuring the feasibility of the project; (3) the creation of the knowledge mobilization committee disseminating the results to various target audiences; and (4) the diversity of activities proposed in the study to train the next generation of students in end-of-life research.

Among the limitations, 2 of the selected jurisdictions (Belgium and Switzerland) are multilingual (other than French and English), and, within their respective linguistic communities, the rates of assisted suicide or euthanasia are quite different. It will be very challenging to compare the data, given the linguistic and cultural differences. Also, the evolution of MAiD rates is perceptible in some countries, while not in others. Alternative explanations would need to be taken into consideration. For example, the number of cases reported to the authorities may differ from the actual number of cases. It will also be necessary to consider the evolution of alternatives to MAiD, particularly continuous palliative sedation, which is experiencing an important growth in some countries, notably in the Netherlands and Switzerland [98]. Finally, large and comparative population surveys in different jurisdictions present several scientific and practical challenges that will require rigor and attention to detail from our research team.

XSL•FO
RenderX

Acknowledgments

On behalf of Consortium interdisciplinaire de recherche sur l'aide médicale à mourir (the authors of this paper and other members: Omid Hosseini, Émilie Allard, Elaine Champagne, Josée Chénard, Marc Dumas, Isabelle Dumont, Nicolas Dupré, Sylvie Fortin, Nathalie Gaucher, Claire Godard-Sebillotte, Diane Guay, Manon Guay, Johanne Hébert, Philippe Laperle, Sylvie Lapierre, Jérôme Leclerc-Loiselle, Marianne Olivier-d'Avignon, Antoine Payot, Chantale Simard, Diane Tapp, Deborah Ummel, Bilkis Vissandjee, Carl Bromwich, Andréanne Côté, Thania Crnich-Côté, Dominique Dion, Geneviève Lamothe, Marie-Hélène Marchand, Guy Morissette, Olivia Nguyen, Élodie Petit, Justin Sanders, Jacinthe Dupuis, Patrick Durivage, Sandra Smele, Emilie Richard-Frève, Jean-Thomas Chouinard, Marie-Alexandra Gagné, France Lacharité, Isabelle Martineau, Marta Martisella, Antoine Bé dard, Francine Carmichael, Eliana Close, Luc Deliens, James Downar, Jocelyn Downie, Andrea Frolic, Ralf Jox, Thaddeus Pope, and Agnes Van der Heide), we acknowledge the contribution of Vanessa Finley-Roy, clinical ethics manager at the Centre Hospitalier Universitaire Sainte-Justine, in Montréal, and Gabrielle Lalonde-LeBlond, research coordinator at the Research Centre of the Centre Hospitalier de l'Université de Montréal, for revising the English version of this manuscript.

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: translation and reformatting. The GenAI tool used was ChatGPT 5.2 and DeepL. Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by Marie-Eve Bouthillier.

Data Availability

Some anonymized data may be shared upon request, within the limits authorized by the research ethics committee.

Funding

The study is funded by the Fonds de recherche du Québec, Action concertée Mieux comprendre le recours à l'aide médicale à mourir en contexte québécois (grant 2025-0AMMR-355830). This study is also possible thanks to additional funds from *Réseau québécois de recherche en soins palliatifs et soins de fin de vie* (RQSPAL), *Pôle de recherche sur la transformations des pratiques cliniques et organisationnelles du Centre intégré de santé et de services sociaux de Laval*, and *Chaire de la Famille Blanchard*. We are also grateful to our partners who support the Consortium interdisciplinaire de recherche sur l'aide médicale à mourir with in-kind contributions: Centre de recherche et d'expertise en gérontologie sociale, the Centre intégré de santé et de services sociaux de Laval, and the RQSPAL.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide, key informants.
[PDF File (Adobe PDF File), 222 KB - resprot_v15i1e83549_app1.pdf]

Multimedia Appendix 2
Discussion group facilitation guide.
[PDF File (Adobe PDF File), 182 KB - resprot_v15i1e83549_app2.pdf]

Multimedia Appendix 3
Peer review report from Fonds de recherche du Québec - Société et culture (FRQSC).
[PDF File (Adobe PDF File), 221 KB - resprot_v15i1e83549_app3.pdf]

References

1. Loi concernant les soins de fin de vie. LégisQuébec. 2025. URL: https://www.legisquebec.gouv.qc.ca/fr/document/lc/s-32.0001 [accessed 2025-08-27]
2. Act respecting end-of-life care. LégisQuébec. 2025. URL: https://www.legisquebec.gouv.qc.ca/en/document/cs/s-32.0001 [accessed 2025-08-27]
3. An Act to amend the criminal code (medical assistance in dying). Government of Canada. 2021. URL: https://laws-lois.justice.gc.ca/PDF/2021_2.pdf [accessed 2025-08-27]
4. Rapport annuel d'activités: du 1er avril 2022 au 31 mars 2023. Commission sur les soins de fin de vie. 2023. URL: https://csfv.gouv.qc.ca/fileadmin/docs/rapports_annuels/rapport_annuel_dactivites_2022-2023.pdf [accessed 2025-08-27]

XSL•FO
RenderX

5. "Les soins de fin de vie au Québec: où en sommes-nous". Collège des médecins du Québec. 2023. URL: https://www.cmq.org/fr/actualites/les-soins-de-fin-de-vie-au-quebec-ou-en-sommes-nous [accessed 2025-08-26]
6. Fourth annual report on medical assistance in dying in Canada 2022. Health Canada. 2023. URL: https://www.canada.ca/en/health-canada/services/publications/health-system-services/annual-report-medical-assistance-dying-2022.html#table_3.1 [accessed 2025-08-26]
7. Action concertée: mieux comprendre le recours à l'aide médicale à mourir en contexte québécois. Fonds de recherche du Québec. URL: https://frq.gouv.qc.ca/programme/action-concertee-mieux-comprendre-le-recours-a-laide-medicale-a-mourir-en-contexte-quebecois/ [accessed 2025-09-03]
8. The Canadian Press. Quebec funds a project to better understand the use of medical assistance in dying. CityNews. 2024. URL: https://montreal.citynews.ca/2024/03/27/quebec-funds-project-medical-assistance-dying/ [accessed 2025-08-26]
9. Boer TA. Assisted dying in the Netherlands: recent legal and religious considerations. Rev Derecho Relig 2022;17:185-198 [FREE Full text]
10. Groenewoud AS, Atsma F, Arvin M, Westert GP, Boer TA. Euthanasia in the Netherlands: a claims data cross-sectional study of geographical variation. BMJ Support Palliat Care 2021:bmjspcare-2020-002573. [doi: 10.1136/bmjspcare-2020-002573] [Medline: 33446488]
11. Byram AC, Reiner PB. Disparities in public awareness, practitioner availability, and institutional support contribute to differential rates of MAiD utilization: a natural experiment comparing California and Canada. Mortality 2024;30(1):252-272. [doi: 10.1080/13576275.2024.2328627]
12. Pullman D. Slowing the slide down the slippery slope of medical assistance in dying: mutual learnings for Canada and the US. Am J Bioeth 2023;23(11):64-72 [FREE Full text] [doi: 10.1080/15265161.2023.2201190] [Medline: 37166283]
13. Khakzedeh L. Assisted suicide in Austria - the new legal framework. BioLaw J 2022;1(22) [FREE Full text]
14. Coelho R, Maher J, Gaind KS, Lemmens T. The realities of medical assistance in dying in Canada. Palliat Support Care 2023;21(5):871-878. [doi: 10.1017/S1478951523001025] [Medline: 37462416]
15. Gentile D. Choisir l'aide médicale à mourir faute de soins palliatifs appropriés? Radio-Canada. 2018. URL: https://ici.radio-canada.ca/nouvelle/1104152/aide-medicale-mourir-soins-palliatifs-college-medecins-penurie [accessed 2025-08-26]
16. Noël, M. Mourir plutôt que de vivre en CHSLD. laTribune. 2019. URL: https://www.latribune.ca/2019/12/04/mourir-plutot-que-de-vivre-en-chsld-3a78fb227bc5f02e96f6ab60a664ad5f/ [accessed 2026-08-25]
17. Morais JA, Durand PJ, Pageau F. Notre inquiétude face à l?aide médicale à mourir. Le Devoir. 2021. URL: https://www.ledevoir.com/opinion/idees/595951/notre-inquietude-face-a-l-aide-medicale-a-mourir [accessed 2025-08-26]
18. Vigneault L. Nous voulons une aide à vivre, pas à mourir. leSoleil. 2020. URL: https://www.lesoleil.com/2020/01/28/nous-voulons-une-aide-a-vivre-pas-a-mourir-faeffda3ba4fe8dbbc9fec2c67a25081/ [accessed 2025-08-26]
19. Lessard D. Mort d'andrée simard: Dans la douleur et la détresse. La Presse. 2023. URL: https://www.lapresse.ca/actualites/sante/2023-01-20/mort-d-andree-simard/dans-la-douleur-et-la-detresse.php [accessed 2025-08-26]
20. Downar J, MacDonald S, Buchman S. What drives requests for MAiD? CMAJ 2023;195(40):E1385-E1387 [FREE Full text] [doi: 10.1503/cmaj.230259] [Medline: 37844931]
21. Des Aulniers L, Lapointe BJ. Le choix de l'heure: ruser avec la mort. Somme Toute. 2018. URL: https://editionssommetoute.com/livre/le-choix-de-lheure/ [accessed 2026-02-25]
22. Roy J, Stylios A, Pilote É. Tendances sociétales sur le plan de l'évolution des valeurs et aide médicale à mourir. Des enjeux pour la pratique. Intervention 2023(156):69-83. [doi: 10.7202/1097407ar]
23. Fournier PJ. Les Québécois peu portés sur le religieux, les Canadiens un peu plus. L'actualité. 2022. URL: https://lactualite.com/politique/les-quebecois-peu-portes-sur-le-religieux-les-canadiens-un-peu-plus/ [accessed 2025-08-26]
24. Rapport annuel d'activités 2021-2022. Commission sur les soins de fin de vie. 2022. URL: https://csfv.gouv.qc.ca/fileadmin/docs/rapports_annuels/csfv_rapport_activites_2021-2022.pdf [accessed 2025-08-26]
25. Bérubé A, Tapp D, Dupéré S, Plaisance A, Bravo G, Downar J, et al. Do socioeconomic factors influence knowledge, attitudes, and representations of end-of-life practices? A cross-sectional study. J Palliat Care 2025;40(2):152-161 [FREE Full text] [doi: 10.1177/08258597221131658] [Medline: 36237145]
26. Plaisance A, Mishara B, Masella J, Bravo G, Couture V, Tapp D. Quebec population highly supportive of extending medical aid in dying to incapacitated persons and people suffering only from a mental illness: Content analysis of attitudes and representations. Ethics Med Public Health 2022;21:100759. [doi: 10.1016/j.jemep.2022.100759]
27. Bronfenbrenner U. The Ecology of Human Development. Cambridge: Harvard University Press; 1981.
28. Peters MD, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews. In: JBI Manual for Evidence Synthesis. South Australia: JBI; 2024.
29. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
30. Veritas health innovation. Covidence Systematic Review Software. 2025. URL: https://www.covidence.org/ [accessed 2025-08-26]
31. Population estimates, quarterly. Statistics Canada. 2025. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901&request_locale=en [accessed 2025-08-27]

XSL•FO
RenderX

32. Marcoux I, Mishara BL, Durand C. Confusion between euthanasia and other end-of-life decisions. Can J Public Health 2007;98(3):235-239. [doi: 10.1007/bf03403719]
33. Bravo G, Trottier L, Rodrigue C, Arcand M, Downie J, Dubois M, et al. Comparing the attitudes of four groups of stakeholders from Quebec, Canada, toward extending medical aid in dying to incompetent patients with dementia. Int J Geriatr Psychiatry 2019;34(7):1078-1086. [doi: 10.1002/gps.5111] [Medline: 31034669]
34. Plourde L, Chang SL, Farzin H, Gagnon P, Hébert J, Foxman R, P3A Study Group. Social acceptability of psilocybin-assisted therapy for existential distress at the end of life: a population-based survey. Palliat Med 2024;38(2):272-278 [FREE Full text] [doi: 10.1177/02692163231222430] [Medline: 38253521]
35. Fortin MF, Gagnon J. Fondements et étapes du processus de recherche. Chenelière Éducation 2017;43(1):264-265. [doi: 10.7202/1042088ar]
36. Driscoll DL, Appiah-Yeboah A, Salib P, Rupert DJ. Merging qualitative and quantitative data in mixed methods research: how to and why not. Ecol Environ Anthropol 2007;3(1):19-28 [FREE Full text]
37. À propos. Alliance des maisons de soins palliatifs du Québec. URL: https://www.alliancemspq.com [accessed 2025-08-28]
38. Aide médicale à mourir. Gouvernement du Québec. URL: https://www.quebec.ca/sante/systeme-et-services-de-sante/soins-de-fin-de-vie/aide-medicale-a-mourir [accessed 2025-08-28]
39. Specialists for you. Fédération des médecins spécialistes du Québec. URL: https://fmsq.org/en [accessed 2025-08-28]
40. Médecins de famille. Prendre soin de votre vie donne un sens à la nôtre. Fédération des médecins omnipracticiens du Québec. URL: https://www.fmoq.org/ [accessed 2025-08-28]
41. Protecting the public by ensuring quality medical care. Collège des médecins du Québec. URL: https://www.cmq.org/en [accessed 2025-08-28]
42. La Commission sur les soins de fin de vie. Commission sur les soins de fin de vie. URL: https://csfv.gouv.qc.ca/en/ [accessed 2025-08-28]
43. Geddes A, Parker C, Scott S. When the snowball fails to roll and the use of 'horizontal' networking in qualitative social research. Int J Soc Res Methodol 2017;21(3):347-358. [doi: 10.1080/13645579.2017.1406219]
44. Canadian Association of MAiD Assessors and Providers. URL: https://camapcanada.ca/ [accessed 2025-08-25]
45. Over SCEN. Artsenfederatie KNMG. URL: https://www.knmg.nl/ik-ben-arts/scen/over-scen [accessed 2025-08-25]
46. LevensEinde InformatieForum. URL: https://leif.be/home/ [accessed 2025-08-25]
47. Academy of Aid-in-Dying Medicine. URL: https://www.aadm.org/ [accessed 2026-08-26]
48. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
49. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a researcher. Int J Transgend Health 2023;24(1):1-6 [FREE Full text] [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
50. QDA miner. Provalis Research. URL: https://provalisresearch.com/products/qualitative-data-analysis-software/ [accessed 2025-08-18]
51. Tracy SJ. Qualitative quality: eight "Big-Tent" criteria for excellent qualitative research. Qual Inq 2010;16(10):837-851. [doi: 10.1177/1077800410383121]
52. Dawson P, Mäkelä M. The Routledge Companion to Narrative Theory. New York: Routledge; 2022.
53. Puckett K. Narrative Theory. United Kingdom: Cambridge University Press; 2016.
54. Webster L, Mertova P. Using Narrative Inquiry as a Research Method. New York: Routledge; 2007.
55. Trahar S. Contextualising Narrative Inquiry. New York: Routledge; 2013.
56. Paillé P, Mucchielli A. L'analyse Qualitative En Sciences Humaines et Sociales. Malakoff, France: Armand Colin; 2012.
57. Braun V, Clarke V. Thematic analysis. In: APA Handbook of Research Methods in Psychology, Vol 2: Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Washington: American Psychological Association; 2012:57-71.
58. Morrow SL. Qualitative research in counseling psychology. Counsel Psychol 2007;35(2):209-235. [doi: 10.1177/0011000006286990]
59. Goldberg R, Nissim R, An E, Hales S. Impact of medical assistance in dying (MAiD) on family caregivers. BMJ Support Palliat Care 2021;11(1):107-114. [doi: 10.1136/bmjspcare-2018-001686] [Medline: 30826737]
60. Jackson VA, Mack J, Matsuyama R, Lakoma MD, Sullivan AM, Arnold RM, et al. A qualitative study of oncologists' approaches to end-of-life care. J Palliat Med 2008;11(6):893-906 [FREE Full text] [doi: 10.1089/jpm.2007.2480] [Medline: 18715182]
61. Girard D. Vers une meilleure compréhension de l'expérience de la souffrance dans le contexte de l'aide médicale à mourir: une étude longitudinale alliant la perspective des patients, des proches et des médecins. Fonds de recherche Québec 2022 [FREE Full text] [doi: 10.69777/296311]
62. Vogl S, Zartler U, Schmidt E, Rieder I. Developing an analytical framework for multiple perspective, qualitative longitudinal interviews (MPQLI). Int J Soc Res Methodol 2017;21(2):177-190. [doi: 10.1080/13645579.2017.1345149]
63. Dahlberg K, Dahlberg H, Moodley P. Reflective Life-World Research. Sweden: Studentlitteratur; 2008.
64. Antonides MF, van Wijngaarden E. "It's like crystal gazing": The lived experience of anticipating end-of-life choices in older adults and their close ones. Gerontologist 2024;64(7):gnae061. [doi: 10.1093/geront/gnae061] [Medline: 38832567]

XSL•FO
RenderX

65. Antonides MF, Tholking T, Girard D, Van Wijngaarden E. Bridging multiperspective and longitudinal research with a phenomenological approach: a novel analytical framework. Int J Qual Method 2026.
66. Merriam SB. Qualitative Research: A Guide to Design and Implementation. New Jersey: John Wiley & Sons; 2009.
67. Gupta M. Le service de consultation psychiatrique pour l'AMM et les troubles mentaux: un projet pilote. Fonds de soutien à la pratique académique des psychiatres du CHUM. 2022 Mar. URL: https:// [accessed 2026-03-16]
68. Dupré N. SLA/AMM: Les patients atteints de sclérose latérale amyotrophique qui demandent l'aide médicale à mourir diffèrent-ils des patients qui décèdent directement de la maladie? Fondation du CHU de Québec – Université Laval. 2024 Jan. URL: https:// [accessed 2026-03-16]
69. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
70. Dalglish SL, Khalid H, McMahon SA. Document analysis in health policy research: the READ approach. Health Policy Plan 2021;35(10):1424-1431 [FREE Full text] [doi: 10.1093/heapol/czaa064] [Medline: 33175972]
71. Luke DA, Powell BJ, Paniagua-Avila A. Bridges and mechanisms: integrating systems science thinking into implementation research. Annu Rev Public Health 2024;45(1):7-25 [FREE Full text] [doi: 10.1146/annurev-publhealth-060922-040205] [Medline: 38100647]
72. Tran M, Honarmand K, Sibbald R, Priestap F, Oczkowski S, Ball IM. Socioeconomic status and medical assistance in dying: a regional descriptive study. J Palliat Care 2022;37(3):359-365 [FREE Full text] [doi: 10.1177/08258597211053088] [Medline: 34747239]
73. McGill E, Er V, Penney T, Egan M, White M, Meier P, et al. Evaluation of public health interventions from a complex systems perspective: a research methods review. Soc Sci Med 2021;272:113697 [FREE Full text] [doi: 10.1016/j.socscimed.2021.113697] [Medline: 33508655]
74. Brennan LK, Sabounchi NS, Kemner AL, Hovmand P. Systems thinking in 49 communities related to healthy eating, active living, and childhood obesity. J Public Health Manag Pract 2015;21 Suppl 3:S55-S69. [doi: 10.1097/PHH.0000000000000248] [Medline: 25828223]
75. Stapley E, O'Keeffe S, Midgley N. Developing typologies in qualitative research: the use of ideal-type analysis. Int J Qual Method 2022;21. [doi: 10.1177/16094069221100633]
76. Ginsburg FD, Abu-Lughod L, Larkin B. Media Worlds: Anthropology on New Terrain. Oakland: University of California Press; 2002.
77. Tétreault S, Beaupré P, Deschênes PM, Rajotte T, Gascon H, Boucher N, et al. Comment mobiliser la communauté grâce au forum communautaire ? Serv Soc Que 2013;59(2):65-75. [doi: 10.7202/1019110ar]
78. Bolduc É, Jean E. Organisation Efficiente D'un Forum Citoyen Un Guide Pratique Appuyé Sur Les Meilleures Connaissances. Consortium InterS4. 2023. URL: https://consortiuminters4.uqar.ca/wp-content/uploads/2023/08/20230803ForumCitoyen_GuideDePratique-InterS4-VF.pdf [accessed 2025-08-27]
79. NVivo 14. Lumivero. URL: https://lumivero.com/products/nvivo/ [accessed 2026-02-25]
80. Mucchielli A, Paillé P. L'Analyse Qualitative En Sciences Humaines et Sociales. Paris, France: Armand Colin; 2021.
81. Salone JJ. Analyse textuelle avec IRaMuTeQ et interprétations référentielles des programmes officiels de mathématiques en quatrième. Sciences-Croisées. 2014. URL: http://jjsan.free.fr/JJSanMath/Recherche/2013%20SCIENCES%20CROISEES%20AnalyseStatRefBOS6Math4.pdf [accessed 2025-08-27]
82. Orientations en soutien à domicile - actualisation de la Politique de soutien à domicile «Chez soi: le premier choix». Ministère de la Santé et des Services sociaux. 2023. URL: https://publications.msss.gouv.qc.ca/msss/fichiers/2023/23-704-01W.pdf [accessed 2025-08-27]
83. Politique nationale pour les personnes proches aidantes - reconnaître et soutenir dans le respect des volontés et des capacités d'engagement. Ministère de la Santé et des Services sociaux. 2021. URL: https://publications.msss.gouv.qc.ca/msss/fichiers/2021/21-835-01W.pdf [accessed 2025-08-27]
84. Mroz S, Dierickx S, Deliens L, Cohen J, Chambaere K. Assisted dying around the world: a status quaestionis. Ann Palliat Med 2021;10(3):3540-3553 [FREE Full text] [doi: 10.21037/apm-20-637] [Medline: 32921084]
85. Boivin A, Marcoux I, Garnon G, Lehoux P, Mays N, Prémont MC, et al. Comparing end-of-life practices in different policy contexts: a scoping review. J Health Serv Res Policy 2015;20(2):115-123. [doi: 10.1177/1355819614567743] [Medline: 25592141]
86. Chao Y, Boivin A, Marcoux I, Garnon G, Mays N, Lehoux P, Advisory Committee, Canadian Medical Association, College of Family Physicians of Canada, Canadian Bar Association, Ministère de la santé et des services sociaux du Québec, Réseau des soins palliatifs du Québec, Commissaire à la santé et au bien-être du Québec. International changes in end-of-life practices over time: a systematic review. BMC Health Serv Res 2016;16(1):539 [FREE Full text] [doi: 10.1186/s12913-016-1749-z] [Medline: 27716238]
87. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy 2005;10 Suppl 1:21-34. [doi: 10.1258/1355819054308530] [Medline: 16053581]
88. Wong G, Pawson R, Owen L. Policy guidance on threats to legislative interventions in public health: a realist synthesis. BMC Public Health 2011;11:222 [FREE Full text] [doi: 10.1186/1471-2458-11-222] [Medline: 21477347]

XSL•FO
RenderX

89. Langlois É, Daniels K, Akl EA. Evidence Synthesis for Health Policy and Systems: A Methods Guide. Geneva: World Health Organization; 2018.
90. Cadre de mesure de la performance du système de santé. Institut canadien d'information sur la santé. 2013. URL: https://www.cihi.ca/fr/acceder-aux-donnees-et-aux-rapports/mesure-de-la-performance-du-systeme-de-sante/cadre-de-mesure-de-la-performance-du-systeme-de-sante [accessed 2025-08-27]
91. Bravo G, Delli-Colli N, Dumont I, Bouthillier M, Rochette M, Trottier L. Social workers' attitudes toward medical assistance in dying for persons with dementia: findings from a survey conducted in Quebec, Canada. J Soc Work End Life Palliat Care 2022;18(3):273-292 [FREE Full text] [doi: 10.1080/15524256.2022.2093314] [Medline: 35793419]
92. Marijnissen RM, Chambaere K, Oude Voshaar RC. Euthanasia in dementia: a narrative review of legislation and practices in the Netherlands and Belgium. Front Psychiatry 2022;13:857131 [FREE Full text] [doi: 10.3389/fpsyt.2022.857131] [Medline: 35722561]
93. Schuurmans J, Crol C, Olde Rikkert M, Engels Y. Dutch GPs' experience of burden by euthanasia requests from people with dementia: a quantitative survey. BJGP Open 2021;5(1):bjgpopen20X10112 [FREE Full text] [doi: 10.3399/bjgpopen20X101123] [Medline: 33172849]
94. Schuurmans J, Crol C, Chabot B, Olde Rikkert M, Engels Y. Euthanasia in advanced dementia; the view of the general practitioners in the Netherlands on a vignette case along the juridical and ethical dispute. BMC Fam Pract 2021;22(1):232 [FREE Full text] [doi: 10.1186/s12875-021-01580-z] [Medline: 34789166]
95. Kouwenhoven PSC, Raijmakers NJH, van Delden JJM, Rietjens JAC, Schermer MHN, van Thiel GJMW, et al. Opinions of health care professionals and the public after eight years of euthanasia legislation in the Netherlands: a mixed methods approach. Palliat Med 2013;27(3):273-280. [doi: 10.1177/0269216312448507] [Medline: 22695742]
96. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
97. Natural sciences and engineering research council of Canada. Canadian Institutes of Health Research. 2022. URL: https://nserc-crsng.canada.ca/en [accessed 2026-02-25]
98. Heijltjes MT, van Thiel GJMW, Rietjens JAC, van der Heide A, de Graeff A, van Delden JJM. Changing practices in the use of continuous sedation at the end of life: a systematic review of the literature. J Pain Symptom Manage 2020;60(4):828-846.e3 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.06.019] [Medline: 32599152]

Abbreviations

AR-MAiD: advance requests for medical assistance in dying
CF: community forum
CIRAMM: Consortium interdisciplinaire de recherche sur l'aide médicale à mourir
ISG: interdisciplinary support group
KT: knowledge transfer
MAiD: medical assistance in dying
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by A Schwartz; submitted 05.Sep.2025; this is a non–peer-reviewed article;accepted 28.Jan.2026; published 20.Apr.2026.

Please cite as:
Bouthillier MÈ, Marcoux I, Perron C, Gagnon B, Lussier D, Rouly G, Moreau M, Dorval M, Lessard S, Girard D, Bravo G, Hébert M, Giroux M, Lemyre S, Beaudin A, Bourque CJ, Soulières M, Lévesque M, Gupta M, Bourgeois-Guérin V, Lavoie D, Lavoie B, Plaisance A, Bernier L, Dupuis J
Understanding the Factors Explaining the Growing Use of Medical Assistance in Dying in Québec: Protocol for an Interdisciplinary Mixed Methods and Multimethods Study
JMIR Res Protoc 2026;15:e83549
URL: https://www.researchprotocols.org/2026/1/e83549
doi:10.2196/83549
PMID:15812204

©Marie-Ève Bouthillier, Isabelle Marcoux, Catherine Perron, Bruno Gagnon, David Lussier, Ghislaine Rouly, Mathieu Moreau, Michel Dorval, Sabrina Lessard, Dominique Girard, Gina Bravo, Maude Hébert, Michelle Giroux, Simon Lemyre, Alexandra Beaudin, Claude Julie Bourque, Maryse Soulières, Maude Lévesque, Mona Gupta, Valérie Bourgeois-Guérin, David Lavoie, Bertrand Lavoie, Ariane Plaisance, Louise Bernier, Jacinthe Dupuis. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative

XSL•FO
RenderX

Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Designing a Digital Twin for the Management of Noncommunicable Diseases: Protocol for a Pilot Study and Methodology Validation

Edgar Ross[1*], MD; Jason Ross[2*], MD; Patricia Beschler[3*], MD; David Guydan[4*], MBA; Robert Jamison[5], PhD

[1]Atrius Healthcare, 20 Wall Street, Burlington, MA, United States

[2]Department of Anesthesia, Northwestern Memorial Hospital, Chicago, IL, United States

[3]Department of Medicine, University of Basel, Basel, Switzerland

[4]TheraNetrix, Inc., Wellesley, MA, United States

[5]Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Edgar Ross, MD
Atrius Healthcare, 20 Wall Street, Burlington, MA, United States

Abstract

Background: Noncommunicable diseases (NCDs) have become the leading cause of mortality worldwide. NCDs account for 89% of all deaths in the United States and cost the US economy more than US $47 trillion in direct and indirect expenses. NCDs also account for the main cause of disability worldwide, and the incidence is increasing. The leading NCDs include diabetes, cancer, cardiovascular disease, chronic respiratory disease, and mental health conditions. Outside of aging, NCDs are caused by modifiable behavioral risk factors that include smoking, drug and alcohol abuse, unhealthy diet, obesity, and inadequate physical activity, and treatment must be directed to all of these domains. We hypothesize that a digital twin concept can be used to personalize treatment regimens through analysis of data that allows for artificial intelligence–based decision making.

Objective: This study aims to present a methodology to validate this concept, which would provide a new clinical approach toward addressing the leading cause of disability and mortality worldwide today.

Methods: This study will use delta scores between treatment arms to ascertain whether that distribution was normal for each of the study variables. Parametric (eg, analysis of covariance) or nonparametric analyses will be used to examine the variables to determine the impact of digital twin efficacy over normal treatment paradigms.

Results: Recruitment of participants is expected to begin 6 months after study funding has been awarded and the needed approvals have been obtained. The expected results will show that digital twin modeling using the biopsychosocial characteristics of each participant will be statistically significant, supporting using this approach for personalized medical care.

Conclusions: This study can help to identify significant clinical characteristics to help mitigate the impact of NCDs through biopsychosocial treatment paradigms. This paper proposes a statistical framework to evaluate the validity of the platform's modeling in support of clinical decision making.

International Registered Report Identifier (IRRID): PRR1-10.2196/75934

(JMIR Res Protoc 2026;15:e75934) doi:10.2196/75934

KEYWORDS

digital twins; NCD; personalized medicine; health risk reduction; biopsychosocial treatment paradigm; noncommunicable diseases

Introduction

Noncommunicable Diseases

Since the beginning of the 1900s, worldwide life expectancy has more than doubled. Initially, the gains in life expectancy came from improved sanitation and treatment of infectious diseases [1-5]. Further increases in life expectancy are attributed to improved nutrition, neonatal health care, vaccines, and economic growth, resulting in improvements in living standards and coordinated public health efforts [6]. Now the leading cause of mortality and disability in the world is noncommunicable diseases (NCDs). The most common NCDs include cardiovascular diseases, stroke, type II diabetes, certain cancers, and obstructive lung disease. Unfortunately, longer life expectancy that has come from advances in medicine and public health has not been followed by improvement in quality of life because of NCDs [4,5]. Unhealthy behavior and inactivity have led to an increased risk of myriad health-related problems, with increased incidence of dementia, obesity, arthritis, and psychological comorbidities. The incidence of one key risk factor in developing an NCD, obesity, has only gotten worse.

XSL•FO
RenderX

This is despite a deeper understanding of its cause and the introduction of new, highly effective drug therapies [7].

Longer life expectancy has led to increased years of disability, and a reduction in quality of life stems from an increase in the incidence of NCDs. These NCDs have led to a substantial increase in health care costs and economic and public health burden [8]. NCDs are lifestyle diseases from deeply ingrained behaviors that can be difficult to change. Modifiable risk factors include poor diet and nutrition, obesity, alcohol abuse, cigarette smoking, lack of regular exercise, and environmental and socioeconomic factors [9-12]. Although a few NCDs account for the highest impact on public health and have many commonalities and similar treatments, the prevention and management of these NCDs on the individual level is actually very complex. This complexity stems from the differences in genetics, backgrounds, resources, access, and attitudes, and when taken together, have been labeled biopsychosocial medicine [13,14]. Personalized medicine applies evidence-based treatment guidelines in a framework that factors individual patients' differences, making treatment success more likely. Biopsychosocial medicine recognizes genetic predispositions as risk factors for the development of NCDs, but even these risks can be overcome by behavior change and effective treatments [15].

NCDs are a general category that has been categorized into different diagnoses. The individual health risks of these NCD diagnoses can vary dramatically. The initial diagnosis of an NCD requires an assessment of risk and treatment recommendations that can vary from self-management to the prescribing of medications or even urgent interventions. Although treatment can vary, lifestyle modification should still be part of a comprehensive approach that should include an understanding of the biopsychosocial background that is unique to the patient. Ideally, active management of NCDs begins with prevention through lifestyle modification. Self-directed management using a biopsychosocial approach includes personal recognition of concerning trends that lead to increased health risks. Both preventative primary care and societal public health initiatives are part of this response [16-18]. Once the presence of an NCD has been diagnosed, treatment options usually begin with pharmacological treatment. If an NCD is recognized late, urgent care could be required; this can include procedures or even surgery [19]. All evidence-based guidelines include an emphasis on lifestyle modifications along with needed interventions. Beyond the acute stage, effective treatment and health risk reduction for these chronic conditions require a therapeutic alliance between the primary caregiver and the patient and an appreciation of the biopsychosocial background of the patient. Personalized medicine has been advocated to develop and support this alliance [20-22].

Treatment of NCDs

The treatment of NCDs is guided by the consensus guidelines developed by international and national public health agencies and specialty groups through the analysis of existing literature and expert opinion. These diagnostic-specific guidelines have been validated through randomized controlled trials (RCTs), comparison studies, meta-analysis consensus meetings, and authoritative expert opinions [23-28]. Although guidelines are very helpful in standardizing evidence-based treatment, their application and adoption by a patient can be limited because of differences in biopsychosocial backgrounds. The goal of personalized medicine is to consider these differences and tailor treatment strategies for optimal outcomes. An example of this complexity is illustrated by the barriers to overcoming physical inactivity.

The cause of NCDs is multifactorial, but one of the most important risk factors for developing an NCD is physical inactivity. Physical inactivity is associated with obesity and developmental musculoskeletal pain. The comorbidities of chronic pain can make the restoration of a healthy lifestyle much more difficult [29]. Yet, the health benefits of sustained activity can be evident within 4 weeks, the largest benefit seen among people with the most sedentary lifestyles [30,31]. A goal of 150 minutes of moderate aerobic activity per week has been shown to be the amount that provides the most benefit. This is information that has been well documented since the early 1950s. Most recently, inadequate exercise has even been associated with an elevated cancer risk [32]. Yet nearly one-third of all adults worldwide do not reach recommended weekly activity levels. Most medical treatments do not include recommendations for needed behavioral changes to reduce the risks of increased NCDs. When behavior change recommendations are made, only a small minority of patients are self-motivated enough to follow through with provider recommendations. The vast majority of patients are not sufficiently motivated to make lasting changes in behavior that could significantly improve their health status when recommendations alone are made. This includes increasing and maintaining higher activity levels [33]. Certain psychological traits, such as catastrophizing when combined with chronic pain, are an even more difficult barrier to compliance. Catastrophizing is known to respond to psychological therapy [34]. Other psychological comorbidities, such as depression and anxiety, can also impact compliance with treatment recommendations [35,36]. These examples illustrate the reason a clinician needs to have a holistic perspective on their patients.

Patient Tracking

Tracking a patient's progress with a chronic condition is key to providing feedback to the patient about their progress and supporting treatment decision-making. Examples of tracking patient progress include objective measures such as laboratory tests, serial blood pressures, activity levels, diet, and respiratory reserve, along with regular subjective measures of patient well-being; all are useful longitudinal measures that can be used to measure compliance, identify barriers that are impeding progress, and personalize care. The data collected should also be continuously screened for acute changes that suggest concerning health risks that require urgent reevaluation. The data collected can be used for disease-specific risk assessment calculators to track treatment effectiveness and longitudinal progress [37-39]. Successful treatment of NCDs often requires lifestyle changes that are deeply ingrained. Contextual and regular feedback with personalized messaging is significantly more effective in reinforcing beneficial behaviors than periodic brief conversations typically associated with health care

XSL•FO
RenderX

appointments [40,41]. Using a personalized medicine approach that is facilitated by a DT specifically developed for NCDs has the potential to significantly improve treatment efficiency, outcomes, and decrease the enormous public health burden that NCDs cause. Constructing a DT requires a system that combines multiple different technologies and a methodology that facilitates adoption by a clinician and their patients. A DT is able to combine the biopsychosocial background of a patient, apply the relevant literature, provide decision support, and continuously adapt to a changing clinical picture. A personalized DT will be constructed using the unique characteristics of a patient's biopsychosocial background, risk of presenting with an NCD, and will use the relevant evidence-based literature that contains large numbers of participants to develop a personalized treatment plan for the N-of-1 patient.

Adopting the DT Concept to Treating NCDs

Since the introduction of DTs, the management of increasingly more complex manufacturing processes has evolved from a product focused on the entire process, including design of the product, building, use, and disposal, which collectively is known as product lifecycle management (PLM). The concept of PLM has been pivotal in rethinking manufacturing from a product focused on the management of the entire life cycle [42]. The key advantage of this approach is the ability to simulate performance, identify, and optimize key variables through continuous measurement and the evaluation throughout the lifecycle of a product. Health care has a similar PLM concept known as an episode of care (EOC). This concept is well-established and has been used to profile and improve care and support cost containment programs [43,44]. EOC begins with the evaluation and assessment of a patient, the development of an active treatment plan, followed by tracking of outcomes when a treatment goal has been reached. Risk adjustment measures have been added to EOC management to better understand reasons for treatment variability and differences in clinical efficiency and outcomes [45,46]. These risk adjustment measures typically include physical, psychological, and demographic backgrounds of patients [47]. Medical care has frequently looked to industry and manufacturing as a model for clinical process redesign [48]. However, evaluation and redesign of clinical processes and pathways use EOC data that are retrospective and aggregated data to optimize processes [49,50]. The similarities between PLM and EOC are noted in Table 1 below.

Table . PLM[a]-EOC[b] comparison.

Stages	Description
PLM stage	
Create	Evaluate a need, develop, and design a product specifically for an opportunity
Build	Build prototype, needed manufacturing infrastructure, and supply chain
Use or sustain	Product behavior, performance, prognostics, service, and repair
Dispose	Project archiving, regulatory compliance, and completion of product life cycle
EOC stage	
Evaluation and assessment	Comprehensive patient evaluation, including biopsychosocial profile
Active treatment	Comprehensive treatment plan formulated and implemented
Follow-up	Track treatment plan progress with revisions as indicated
Treatment course completed and outcomes assessment	Patient tracking, change in health risks, and process efficiency

[a]PLM: product lifecycle management.

[b]EOC: episode of care.

Elements Needed to Construct a Personalized NCD DT Program

The process of incorporating a DT to manage a PLM could also be used to manage an EOC. A DT of a patient would be developed with historical and real-time patient data, diagnosis-specific risk assessment using simulation, machine learning, computational modeling, and large language modeling (LLM) review of the relevant evidence-based literature to obtain a unique biopsychosocial profile of the patient. Instead of the typical health care process that consists of intermittent patient-provider contact, the DT would be continuously updated by bidirectional data streams between the actual patient and their personal simulation. At the end of the EOC, the analytics collected would be used to improve and enhance the treatment algorithms that can be used to further clinical efficiency and improve outcomes. Figure 1 illustrates how a DT can be used to manage an EOC.

XSL•FO
RenderX

Figure 1. Integrating a digital twin into the episode-of-care structure. AI: artificial intelligence; BioPsySoc: biopsychosocial; EOC: episode of care; NCD: noncommunicable disease.

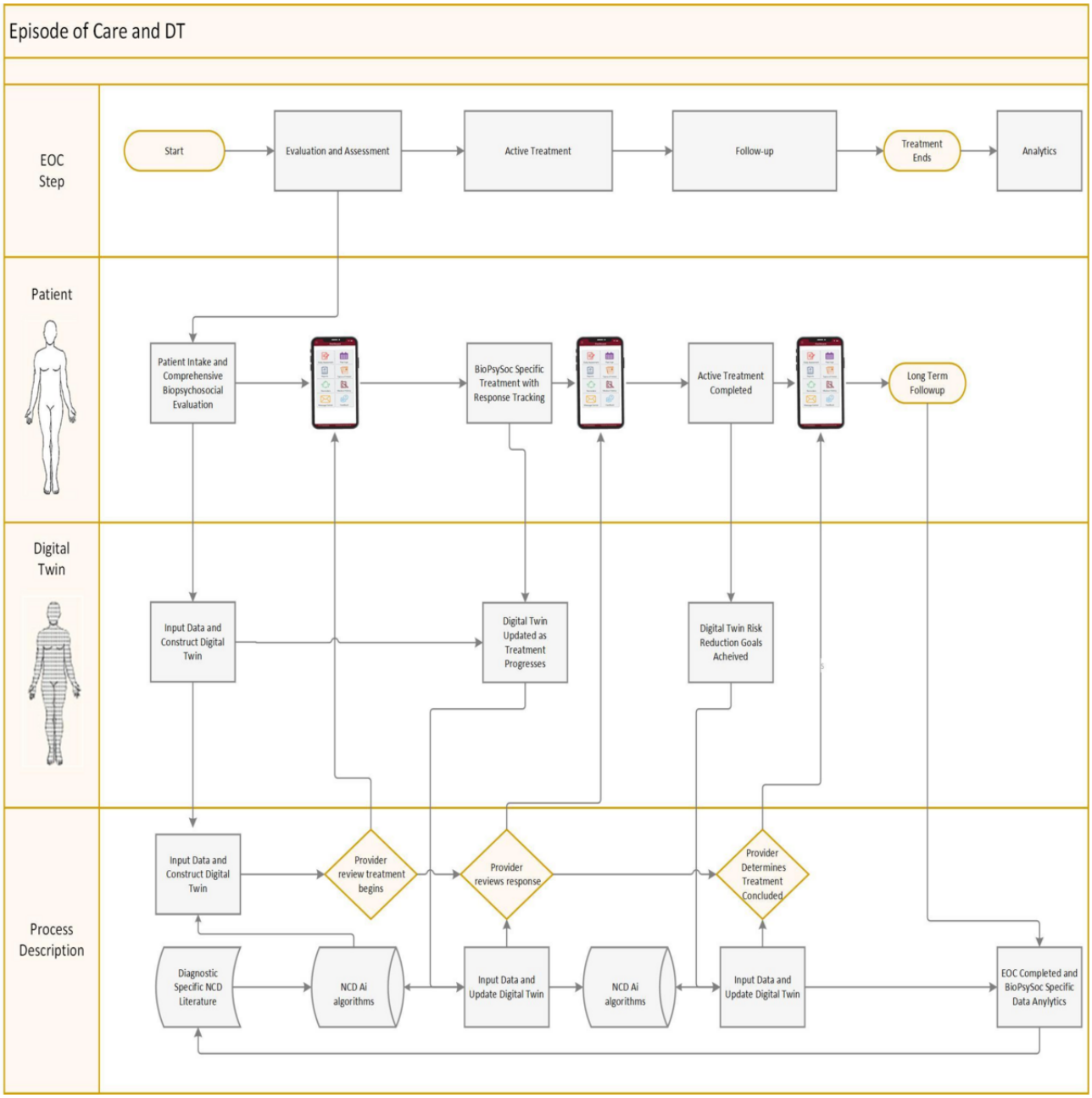

A comprehensive structure that can be used to develop a personalized DT that has the ability to assess and provide decision support in health care requires the following elements:

1. Evidence-based literature is appropriately structured that will inform the background for algorithm decision-making that assesses health risk and provides treatment options for patients with varied backgrounds.
2. The actual patient is evaluated using a biopsychosocial analytic protocol.
3. The virtual patient is constructed as a synthesis of the results of actual patient evaluation and evidence-based literature-based artificial intelligence (AI) algorithms.
4. Secure continuous 2-way data connection between an actual patient and their DT.
5. Data supported DT recalibration through the active treatment phase of the EOC.
6. Tracking treatment outcomes, data analytics and integration, modeling, and refinement of treatment decision support and effectiveness.

The European Virtual Human Twin Initiative (EDITH) project was a 2-year project sponsored by the European Commission to develop a roadmap toward the development of a virtual human twin. In what has been referred to as a moonshot, the EDITH creates the infrastructure and processes with a goal of describing the entire known human pathophysiology [51]. To complete this project, immense data sets, computing power, understanding of both normal and abnormal physiology, and the mechanisms of interventions are needed. This scope of work will require considerable development in many areas that puts it out of reach because of the numerous technical limitations and overall cost. Narrowing the focus of a DT to certain diseases that have smaller numbers of variables and treatment pathways

XSL•FO
RenderX

has been used successfully to overcome these technical limitations. The EDITH project's report "Vision for the VHT and Roadmap Outline" categorizes DTs into three categories. They include (1) generic DT: the expected accuracy is within the range of the studied population, (2) population-specific DT: the expected accuracy is sufficiently close to a mean or median range of values of the referenced population, and (3) participant-specific DT: the accuracy is close to each individual in the reference population with similar characteristics.

Although the above 3 categories can be viewed as distinctive and separate, the categories could also be viewed as a series of stages in a development pathway toward the goal of a DT that will be able to deliver personalized medicine and early disease detection to an N-of-1 (Table 2). All 3 of these stages represent initial development approaches. It is expected that ongoing comprehensive outcome data collection, including sensor data, providing sufficient experience for statistical analysis, will improve the predictive modeling of DTs.

Table . Overview of digital twin development options.

DT[a] type/stage	Input data	Attributes/deficiencies	Required data	Clinical decision support
Generic	Studied populations, accuracy within bounds of referenced population	Least number of variables, easier to develop, may not accurately reflect individuals with important differences outside bounds of referenced studies	Large data sets with similar methodologies, outcomes, and diverse biopsychosocial backgrounds	Generalizable DTs might not be accurate for unstudied key individual differences
Population-specific	Relevant participant-specific study means, medians, or equivalent	Larger number of variables, improved modeling, may not accurately reflect individuals outside SDs	Need studies with sufficient data detail, and subgroup analysis for negative studies	Improved virtual DT representation for patients who fit clinical trial profiles
Participant-specific	Sufficient individual treatment goal–specific accuracy	Largest number of variables, significant evidence-based data mining for AI[b], and in silico modeling will improve DT accuracy	Comprehensive literature review and analysis to find connections that could impact patient care	Most accurate DT, with ability to predict barriers that can impede reaching treatment goals

[a]DT: digital twin.

[b]AI: artificial intelligence.

Medical Literature and Evidence-Based Medicine

Access to large data sets is key to training the AI algorithms needed to deliver the promise of DTs and personalized medicine. Data sets needed for training include well-designed clinical trials for decision support, risk assessment tools that can provide feedback on treatment interventions and large observational cohorts for phenotyping of diverse populations, and prediction of response to interventions. The "All of Us" research program funded by the National Institutes of Health is one example of a data set that could be used to develop a DT infrastructure. With a stated goal of collecting data from a million highly diverse groups of participants, the project will provide considerable biopsychosocial data that advance the goal of personalized medicine [52-54].

The last few decades have seen rapid advances in understanding the pathophysiology of the majority of diseases and the development of new treatment options. The "bench to bedside" process has relied on RCTs of varying designs to obtain needed approval for any recommended clinical treatment. RCTs typically rely on the recruitment of homogenous populations of potential responders to determine the efficacy of a medication or an intervention. RCTs have become the gold standard for evidence-based medicine; however, clinicians treat heterogenous patient populations who often do not share the characteristics of participants found in RCTs. Diseases do not often present with identical symptoms from one patient to another. Treatment response and side effects can also vary greatly from one person to another. Because of this, personalized medicine has been embraced as a way to improve the care of the individual patient. This approach seeks to tailor treatment to each person to account for genetic, physiological, biochemical, behavioral, and environmental differences [24,55]. To be most effective, personalized medicine requires the consideration of a large number of variables. The results from traditional RCTs do not easily lend themselves to providing the needed information to practice evidence-based personalized medicine [17,56].

A consensus panel review of clinical trial design for pain (see the study by Edwards et al [19]) made suggestions to expand and modify existing methodologies to incorporate the needed information that allows for a more accurate understanding of how individual factors predict and influence clinical treatment decision-making. Although the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials panel recommendations are specific to pain management studies, their recommendations also have general application for other conditions [57]. However, for the busy clinician, implementing personalized medicine can become very complicated and time-consuming and not easily adapted for the average provider without support.

Although AI in medicine has received considerable attention because of its promise to improve health care, its true potential in clinical practice has not yet been realized. AI is a broad term that refers to computer-based algorithms that can mirror tasks and abilities of human intelligence. There are many different forms of AI, but all share a common goal—using

XSL•FO
RenderX

computer-based algorithms that review data to uncover diagnoses, assist clinical decision-making, and uncover new therapies [58,59].

AI possesses advantages over human intelligence through its ability to assimilate vast amounts of data, to assist with mundane repetitive tasks, and to consider numerous variables at once to assist with decision-making. This process exceeds the human capability to assimilate information and can lead to greater consistency and speed in decision-making, to remove bias based on anecdotal experience, and to quickly incorporate new advanced treatments [60,61]. The validity of evidence-based literature that would be used by AI decision-making carries with it a key risk factor, the underrepresentation of minorities and certain demographic groups [62-64].

The disadvantages of AI include biased algorithms introduced by racial, gender, socioeconomic, and age prejudice found in the medical literature. Patient safety, ethical, and privacy concerns covering data collection can be ignored. Algorithms can also lack transparency, compromise clinical implementation, and cause distance between the health care provider and the patient [65]. Appropriate governance and transparency of AI are needed to ensure that bias is not introduced into clinical decision-making. Recommendations made by the DT program will be continuously monitored, and clinician feedback on decision-making will be reviewed and used to ascertain that DT algorithmic decision-making is appropriate and effective. Adjustments when needed will be transmitted to the study personnel and clinicians.

The most compelling use of AI in medicine has been in the role of supporting resources that a clinician can easily and effectively use in daily practice. Data supporting this type of collaboration have the potential to offer compelling synergism [66]. The combination of clinician support and AI-enhanced clinical care could greatly improve the goal of implementing personalized medicine [67].

Health Care and DTs

A DT is a virtual representation of a physical object that is used to simulate a process. This simulation can be used to predict the impact of a new process or to modify an existing process to improve risk identification and clinical efficiency. It has the potential to contribute to real-time decision-making to effectively predict outcomes and to make cost-reduction improvements. Because of these attributes, a DT with virtual representations has been shown to have widespread use in industry to simulate the impact of proposed modifications and run “what-if” scenarios before implementation. When a modification is introduced, DTs can be continuously updated with performance data that can be used to manage the process in real time. This real-time monitoring in parallel with the updating of a DT leads to improved management and system stability while enhancing system security [68,69].

Health Care Simulation

Simulation in health care has long been used in training and in improving outcomes in areas such as diagnosing cardiorespiratory arrest, designing mental health treatments, and, most recently, planning treatment interventions such as surgery. The major drawback of simulations is that they are fixed; they do not receive new data, and they are generally not easily modifiable when clinical situations change [70-72]. DTs differ significantly from simulations because of the continuous flow of data from the patient to the DT, which can lead to modification of the DT in real time by simulating the patient’s clinical course [73]. When used in health care, a DT is a virtual representation of a patient. A health care DT has essentially the same attributes as an industrial DT. A patient DT is constructed by analyzing relevant characteristics and problems of that individual through AI to simulate the potential treatment options and provide the best clinical course. DTs have the potential to improve treatment compliance, address individual risks, reduce barriers to improvement, intervene before side effects become problematic, and enhance a therapeutic alliance between the patient and provider. This approach of combining individual characteristics and AI data from a DT that is unique to a patient helps to support the goal of personalized medicine. The health care DT represents a key technology that brings together the advantages of a comprehensive picture of a patient, leading to a precise diagnosis and a personalized treatment plan that is evidence-based yet still personalized to that individual [74,75]. As treatment progresses, the DT is continuously updated by monitoring the patient’s progress through data collection and providing feedback to the patient and treating provider. The changing clinical picture leads to a modification of that patient’s DT.

Treatment Trajectory

Treatment trajectories leading to clinical improvement rarely follow straight linear paths toward a goal [76]. Figure 2 illustrates this for a hypothetical patient. As noted in the figure, the delta of change over a course of treatment can vary. These changes could lead to different approaches or consideration of different treatment options. Ongoing tracking data from the patient and revision of the patient’s DT, as these data become available, modify the DT and can help identify these deltas. This process can provide new insights and a change in clinical decision-making. This back-and-forth data exchange in real time between the patient and that person’s DT can be key to providing personalized medicine.

XSL•FO
RenderX

Figure 2. Hypothetical treatment course.

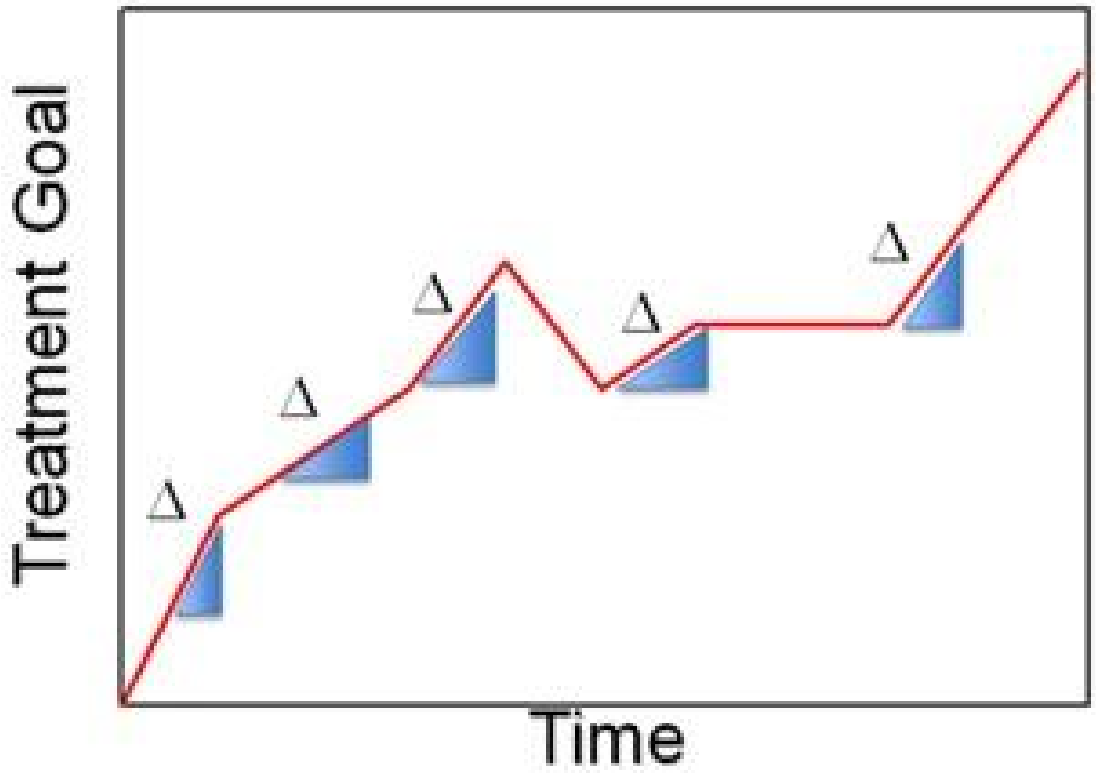

Designing and Developing a DT for NCDs

To our knowledge, a protocol designed to assess the use of a DT to treat NCDs has not been conducted or reported in scientific literature. As noted above, NCDs are lifestyle diseases that can require a multidisciplinary approach toward effective diagnosis and behavior change to impact the long-term reversal of health risk factors using a biopsychosocial approach. The interdisciplinary approach needed to effectively treat a patient with an NCD requires the following:

1. Comprehensive medical review and treatment of the consequences of the NCD.
2. Biopsychosocial evaluation with risk assessment that can impact treatment goals.
3. Psychological assessment and comprehensive treatment when needed.
4. Modification of behaviors that have led to the development of the NCD.
5. Evaluation of current activity, presence, and/or treatment of pain, and integrated rehabilitation when identified.

The design and implementation of a DT for an NCD is as follows:

1. Conduct a literature review of appropriately designed studies related to any particular NCD. These studies will provide an analysis of health risks and therapeutic clinical trials that include both health risks and responder analyses. This information will be used through supportive AI decision-making to form a DT and to offer clinical support.
2. Obtain a patient-centered interface that connects with an electronic medical record to collect existing data. This will allow for the assessment of the patient's baseline key health indicators, to track key variables to be used to update the patient's DT while providing feedback, including contextually specific motivational messaging, and to support the provider. Identification of patient engagement issues and possible measures of treatment outcomes would be generated. Tracking of key patient measures will be facilitated by peripherals enabled to gather objective data such as blood glucose, blood pressure, daily activity, oximetry, and weight [77].
3. The computational process will begin and continue throughout the study process as data are collected, and the patient status is defined by analyzing key health data to form both the initial and updated individual DT using a biopsychosocial framework.
4. Collect deidentified data that will further refine the algorithms used to create the DT and future DTs.

The NCDs that account for the most common cause of death in the United States are cardiovascular disease, cancer, chronic respiratory diseases, and type II diabetes [78]. All of these NCDs are directly influenced by maladaptive behavioral lifestyles such as poor diet, insufficient levels of activity, cigarette smoking, metabolic disorders, and medical conditions, including hypertension and hypercholesterolemia [24]. Biopsychosocial risk factors for the NCDs can include the following items listed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Biopsychosocial risk factors for noncommunicable diseases.

- **Genetics**
- Metabolic risk factors
- **Behavior**
- Smoking
- Excessive alcohol intake
- Unhealthy diet
- Low levels of activity
- **Environment**
- Air pollution
- Exposure to carcinogens and toxic wastes
- Temperature extremes
- Civic infrastructure
- **Demographics**
- Access to health care
- Income
- Age/gender
- Ethnicity

Barriers to Progress and DT Treatment Planning

Effective treatment requires an understanding of barriers to goals, including environmental and demographic factors. Genetics is known to play a role in NCDs, but with personal feedback and targeted treatment, the genetic impact can be mitigated.

We have previously reported on our patient engagement platform that has been used in detecting and treating risk factors for NCDs, including diabetes and chronic pain (MobileNetrix) [79-81,82]. We have also reported on the ability of our patient engagement platform to identify psychological traits that can impact compliance with treatment recommendations and comprehensive care [83]. The patient-facing app is customizable for conditions being studied and is easily adapted to the treatment of all types of NCDs. The app is designed to reinforce patient engagement with their provider, as well as promote behavioral change through automated motivational messaging and psychological principles of change. It can also track patient progress and easily incorporate peripherals and wearables that will provide the needed data to construct a patient-specific DT [84,85]. The programming flowchart for an NCD DT is illustrated in Figure 3. A more detailed explanation of each element is presented in Figure 4.

XSL•FO
RenderX

Figure 3. Digital twin overview. DT: digital twin.

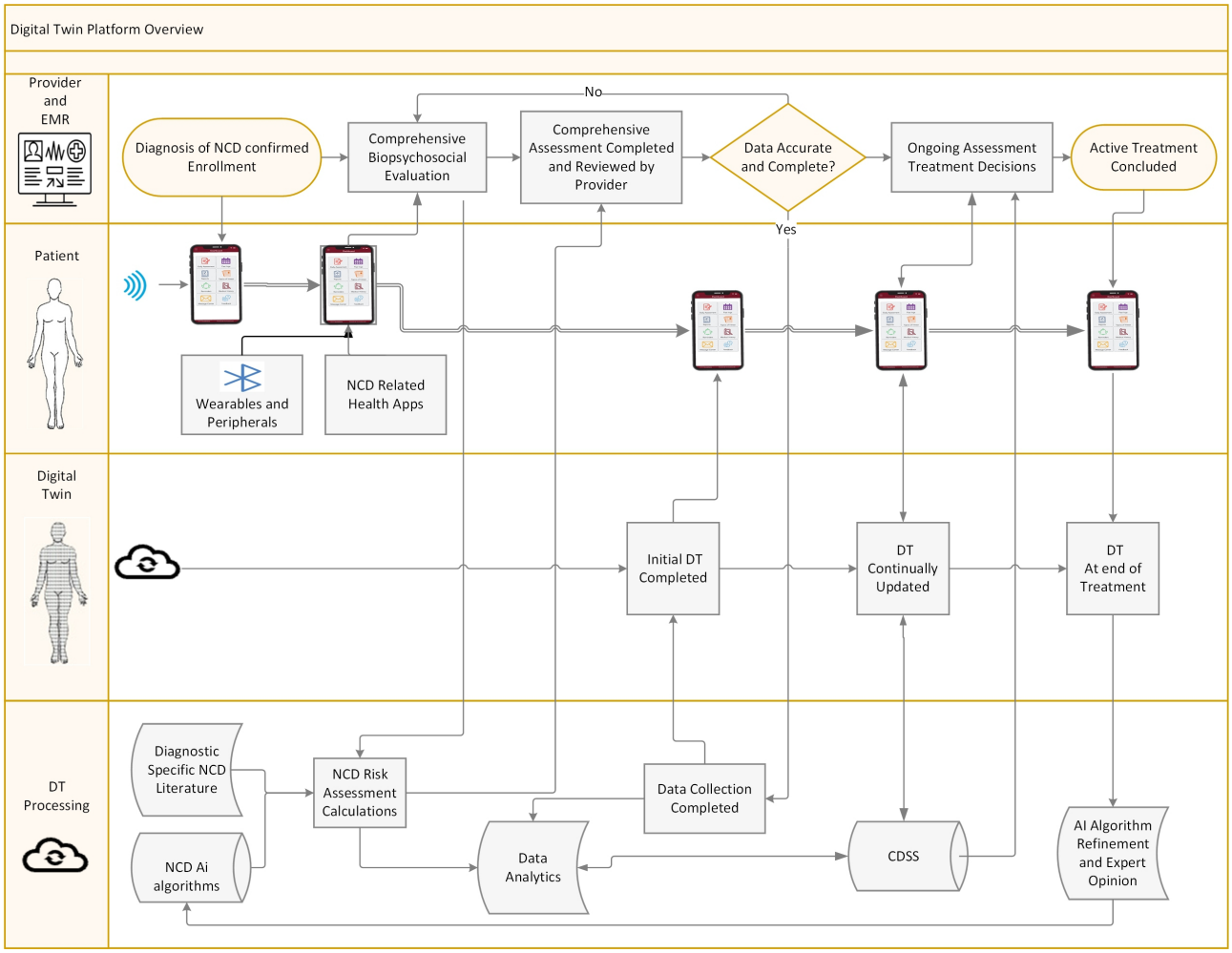

XSL•FO
RenderX

Figure 4. Overview of inputs and information flow needed to construct a DT for NCDs. AI: artificial intelligence; EMR: electronic medical record; *ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*; NCD: noncommunicable disease.

XSL•FO
RenderX

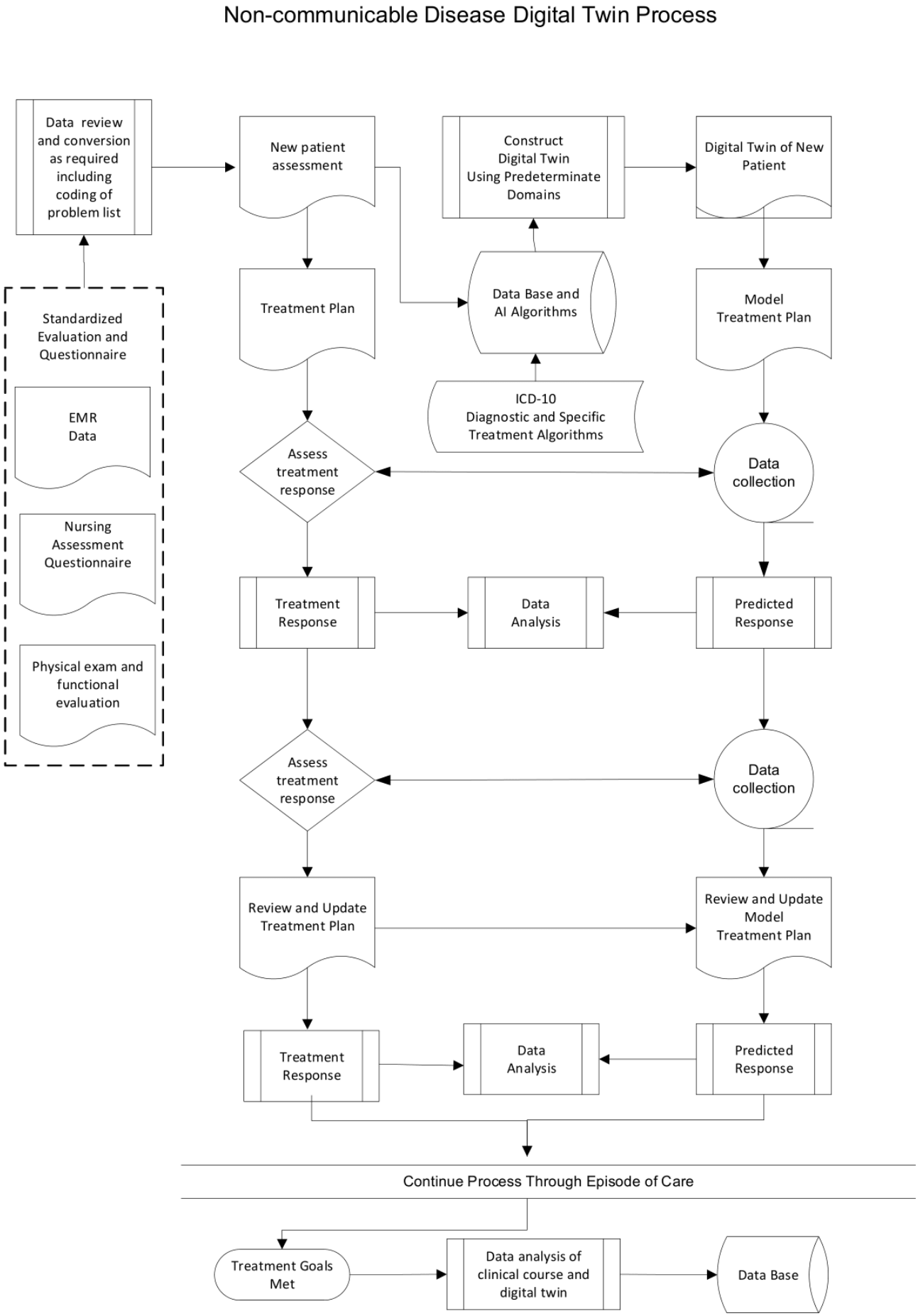

XSL•FO
RenderX

Biological systems are nonlinear. As noted above, the clinical course of an individual is also usually nonlinear, characterized by many deltas in an EOC toward achieving a treatment goal. Any DT must allow for this variability and continually track the individual through efficient data collection and updating of the DT. Our model for a DT will have 4 essential components. These four components will include (1) the identified patient, (2) the unique model of the patient or DT, (3) rapid and timely synchronized data exchange between the individual and the DT, and (4) ongoing analysis of potential variables that are predictive at a macro level. Using a back-and-forth process and a large number of variables, updates of DTs can include an analysis of all the original variables, including their potential relationships and impact. The DT platform consists of the use of the nearly ubiquitous smartphone that will provide patient-facing data collection and biometric and health data from wearables (eg, Google Fit and Apple Health) and a communication portal with 2-way messaging. A DT Health Management App for both iOS and Android platforms that will be HIPAA (Health Insurance Portability and Accountability Act)-compliant will offer electronic medical record integration, enhanced data entry features, Bluetooth-enabled device syncing (eg, glucose monitors), and emergency alerts. This app will be designed to be robust, scalable, and secure, ensuring seamless integration of various components to deliver a comprehensive health care solution. A custom application programming interface (API) will facilitate secure and efficient data exchange between the mobile app, backend infrastructure, and web portals. An administrative portal will be used by the treating provider to access data with electronic medical record integration and offer decision support, and a super-admin portal will monitor system performance and the computational analysis that supports DT construction and updates. A collaborative information model has been proposed that will be able to evaluate complex systems with multiple inputs interacting together to define a specific outcome [45]. When data to construct or update a DT is incomplete, expert knowledge–based evidence will be used. Figure 3 details DT functioning and data processing, flow, and provider supervision. Table S1 in Multimedia Appendix 1 illustrates communication processes and their APIs.

Study Objectives

The goal of this study and the primary objective is validation of our automated patient engagement programs to guide participants toward activity gains, sustainability, opinions of the situational relevance of the engagement tools, and other key performance indicators using a DT. The research plan has several objectives. It will examine the impact that personalized health care messaging and treatment recommendations can have on individuals. Secondary objectives include (1) testing published guidelines on a heterogeneous population, (2) developing a methodology that combines multiple different biopsychosocial characteristics into a personalized treatment plan that improves on and is responsive to an individual's unique background, and (3) refining algorithms that support clinical decision-making of a clinician.

Methods

Validation Methodology for a DT NCD

This proposed study will use the smartphone (iOS and Android) and MobileNetrix as the health management systems platform for a patient DT NCD study. A HIPAA-compliant DT Health Management App will be created, supported by AI and computational analysis. The administrative portal will automatically poll each participant after being identified as a candidate based on validated criteria of having one of the NCDs studied in this protocol. The 3 NCDs that will be included are (1) cardiovascular disease, (2) type II diabetes (non–insulin dependent), and (3) chronic obstructive pulmonary disease.

Candidates will be screened through a 2-step evaluation process. After informed consent is received, the participants will be invited to download the study app. Following a successful app enrollment, a screening questionnaire specific to the provisional diagnosis will be sent through the study app that will contain a link to a designated questionnaire. Once complete, if the individual is found to qualify for the study through the screening questionnaire, an appointment will be made for a comprehensive evaluation to confirm the diagnosis and related risk factors. Table 3 details the screening process for each NCD.

XSL•FO
RenderX

Table . Screening process for specific noncommunicable diseases.

NCD[a] diagnosis	Screening questionnaire	Confirmatory evaluation
Cardiovascular disease	AHA PREVENT[b] risk calculator [86]	Laboratory and physical exam: diabetes (%), current smoking (%), mean total cholesterol (mg/dL), mean systolic blood pressure (mm Hg) [87]
Type II diabetes	Diabetes Risk Score Questionnaire [88]	Fasting blood glucose >126 mg/dL, 2-hour OTT[c] blood glucose >200 mg/dL, HbA_{1C}[d]>6.5% [89,90]
COPD[e]	CDQ[f] questionnaire [91]	Postbronchodilator spirometry ratio of FEV_1/FVC[g] of less than 0.70 confirms the presence of persistent airway obstruction and COPD [92]

[a]Noncommunicable disease.

[b]American Heart Association Predicting Risk of Cardiovascular Disease Events

[c]Oral tolerance test.

[d]Hemoglobin A1C.

[e]Chronic obstructive pulmonary disease.

[f]COPD diagnostic questionnaire.

[g]FEV_1/FVC: forced expiratory volume in 1 second to forced vital capacity.

If the risk factors are confirmed, an additional evaluation will be used to determine other specific risk factors associated with that condition, including level of physical activity, psychological traits or maladaptive behaviors, and barriers to improvement, including the presence of pain, sleep disruption, smoking, and other comorbidities related to worsening health. For an overview of these inputs, refer to Table S2 in Multimedia Appendix 1. These secondary domains will become the individual's study-specific targets. All of these factors will be incorporated into the participant's DT. The DT will be shared with the patient through the study app and reviewed with the patient during an in-person follow-up visit. The study platform will provide daily questionnaires that are specific to each NCD category and problem list. Objective tracking of the patient's activity level and physiological measures (such as blood pressure, glucose levels, and pulse oximetry) will continue throughout the treatment period. All these daily data points will be used to update the DT and will appear in the participant's study portal that is part of the patient-facing app. The updated DT primary endpoint will reflect the change in health risk. Participants found not to be compliant with the program will be contacted and interviewed regarding barriers and their impressions of the program. Clinician decision-making will also be tracked. Variance from DT treatment recommendations and the impact on NCD risk factors will be reviewed and statistically compared to expected outcomes from relevant evidence-based literature.

Automated contextual daily motivational messaging that aligns with the participant's behavioral traits will continue throughout the study period. Daily compliance reminders for medication regimens will be incorporated into the study app. The program goal will be a 37% reduction in long-term health risk from baseline as determined by the NCD-specific risk questionnaire (Tables S1 in Multimedia Appendix 1). This is considered to be a meaningful reduction in health risk using the financial modeling included in this reference that will sustain a DT program [93]. The number of participants in each arm reaching this goal will be used to evaluate the primary objective of this study as noted above. Each candidate will be followed for 6 months, and the study length will be determined by enrollment rate.

The program's daily questionnaire options include (1) the participant's impression of the program and global impressions of change, (2) Global Activity Limitation Index, (3) Brief Pain Inventory, (4) medication adherence and compliance, (5) daily diet summary, (6) quality of sleep, and (7) smoking and alcohol consumption.

Validation Study Material and Methods

Overview

All eligible participants will be provided sufficient details on the study so that they can provide informed consent prior to enrollment. This process will include a complete discussion of the study's purpose, the study procedures, and the risks and benefits of participation. The risk discussion will include the risks and benefits of the use of AI for health care. All study data collected will be kept confidential and securely stored. All personal identification will be removed from the study data. Personal identifiers will be available to treating providers. Participation in this study is completely voluntary, and participants can withdraw at any time. The decision to participate in this study will not affect access to care. All treatment decisions will be made by the patient's provider only. This study will conform to all international standards covering human subjects for this research project.

This study will be approved through the Institutional Review Board (IRB) of Brigham and Women's Hospital in Boston and registered on ClinicalTrials.gov. The health care providers participating in this validation study will identify and sequentially include sufficient patients to validate this process.

XSL•FO
RenderX

An effort will be made to include a diverse group of enrollees to allow for a subgroup analysis of the impact on decision-making in the biopsychosocial and personalized medicine. The potential participants will be new patients in their practice identified as having one of the NCDs that are being targeted in this study. The potential participants will be screened for both inclusion and exclusion criteria listed below. If found to be eligible, the personal smartphone of the identified participants will be polled using SMS text messages to determine interest. Interested individuals will be provided the study app and questionnaires associated with their provisional diagnosis. If the screening questionnaires are compatible with an NCD diagnosis, secondary questionnaires will be transmitted through the study app and an appointment will be made for confirmation of the diagnosis using the criteria listed in Textbox 2. After the confirmation of the diagnosis, each participant will be evaluated by a multidisciplinary team that will include the treating provider, a psychologist, and a physical therapist. After informed consent and sufficient instruction, the participants will be formally enrolled and treatment will begin. Relevant outcomes such as improvement in activity duration and frequency, as well as response to automated messaging, will be tabulated. Figures 4 and 5 illustrate the clinical pathways that will be implemented in this clinical trial of an NCD-focused DT, along with the feedback loops that will be used to refine the validity and increased accuracy of the DT toward being able to support a biopsychosocial personalized medicine at the level of N-of-1. The eligibility criteria are listed in Textbox 2.

Textbox 2. Inclusion and exclusion criteria.

Inclusion criteria

- Personal smartphone (either Android or iPhone iOS).
- Clearance from the primary care provider.
- At-risk noncommunicable disease and confirmed by in-person evaluation.
- Willingness to participate.
- No contraindications to participating in treatment recommendations.

Exclusion criteria

- Diagnosis of stage IV cancer or any other advanced malignant disease.
- Acute osteomyelitis or acute bone disease.
- Present or past DSM-5 (*Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]) diagnosis of schizophrenia, delusional disorder, psychotic disorder, or dissociative disorder that would be judged to interfere with study participation.
- Pregnancy.
- Any clinically unstable systemic illness judged to interfere with treatment.
- Any condition requiring urgent surgery or a procedure.
- An active substance use disorder, such as cocaine or IV heroin use that would interfere with study participation.

Study Risks

The risks of participation in this study are minimal, mainly because the clinician will make all the treatment decisions. All treatment that the participants receive will be standard-of-care treatment that any patient with the same diagnosis would receive. Although the risk is low, the participants' data and communications could be breached. As both platforms are online, the risk of unauthorized viewing or use of online data is higher than that of physical records. The entire platform will be hosted within a HIPAA-compliant firewall, and the smartphone patient-facing app will be compliant with all relevant standards to maintain patient confidentiality. The MobileNetrix app has been used previously in clinical trials conducted at Mass General Brigham and is Veracode tested. Finally, participants may become uncomfortable and emotional while completing some of the questionnaires. Risks of AI include unrecognized biases and hallucinations, leading to the potential of erroneous decision support recommendations.

Enrollment Procedures

The smartphone will be used to communicate with the eligible participants. They will be sent a poll using an SMS text messaging–based survey app for eligibility for in-person evaluation. The administrative portal will automatically categorize participants into a defined level of risk of their NCD. The participants with confirmed NCDs will be evaluated by the interdisciplinary team noted above, and a personal treatment regimen will be designed for them to reduce their health risks, along with a DT that reflects their state of health at enrollment. There will also be a comparison-matched group of participants who will be randomized to a treatment-as-usual condition, where they will not have access to the DT technology. After sufficient instruction, the participants in the experimental treatment condition will be tracked daily and be provided with reinforcing messaging and automated problem-solving based on their DT. Any changes in their health-related behaviors will automatically be reflected in an updated DT. Efficacy of the DT and personalized medicine approach will be measured by the number of individuals reaching a 37% risk reduction, with comparison between enhanced care using the DT and treatment as usual

XSL•FO
RenderX

(TAU). TAU is defined as the treatment of one of the study's NCDs using published guidelines and literature as a guide, but without using any of the DT study interventions. Secondary outcome measures that will be tracked will include each of the diagnosed NCD risk factors and other outcome measures, including compliance and helpfulness ratings of the participants and treating providers. In addition to studying the 3 NCDs noted above, we will select patients diagnosed with type II diabetes with pain for further in-depth analysis, as detailed in the Statistical Methods section.

Statistical Methods

We expect that both groups (DT Experimental and TAU Control) will have missing data. We will assess whether missing data are more prevalent in one group or the other, and if there are differences among those who drop out early and those who complete the 6-month trial. We will also use corrections for multiple comparisons to reduce the likelihood of type I error (Benjamini-Hochberg correction). A calculation will be produced to determine whether the significance value to interpret statistical significance needs to be different from the standard $P<.05$. Power calculations, as outlined by Cohen [94], will be performed to determine the probability of detecting clinically significant differences between treatment conditions in the primary area of measurement. These calculations will assume a 2-tailed test and α level of .05 that use of a DT will improve activity interference on the activity subscale of the Brief Pain Inventory. The Brief Pain Inventory was chosen as a primary outcome measure. This is because persistent pain and other symptoms are often associated with NCDs that can adversely affect quality of life. The Brief Pain Inventory includes 7 items that assess the degree to which pain and other symptoms interfere with daily living and activity, including general activity, walking ability, normal work (both outside the home and housework), relations with others, sleep, mood, and enjoyment in life (0=does not interfere; 10=completely interferes). These items can be collapsed and averaged into one measure of interference.

In our previous work [85,95-97], we had a high rate of consent to participate among recruited patients, with minimal attrition. We assessed power based on our preliminary data and previously published studies. The power analysis revealed that a sample size of 100 participants (50 per treatment arm) using a 2-sample *t* test for a difference in mean pre/post change score gives the study a >85% probability of detecting a 1.5-point group difference on a 0 - 10 rating scale (assuming an SD of 1.8). Conventional metrics suggest that a 30% change or difference score is clinically meaningful [98], and this sample size would provide adequate power to detect a group difference of approximately 20%. Most of the analyses in the proposed study will involve delta scores reflecting changes within participants (both uncorrected and adjusted least squares). We will examine the distribution of the delta scores between treatment arms and ascertain whether that distribution was normal for each of the study variables. Parametric (eg, analysis of covariance) or nonparametric analyses will be used on the variables based on whether the variables were numeric or categorical and showed normal distributions. Collectively, this sample size (N=100), together with our previously observed very high retention rates and the substantial efficacy of the platform, will provide more than adequate power to detect moderate-size effects. Participants will be randomized to the DT and TAU groups with equal allocation. The primary outcome measure will be the Patient Global Impression of Change (PGIC), which represents the participant's overall belief about the efficacy of treatment on a 7-point categorical verbal rating scale. The scale ranges from (1) "no change or condition has gotten worse" to (7) "a great deal better and a considerable improvement that has made all the difference." There will be prespecified secondary efficacy measures. Pain severity and pain interference with function will be evaluated with the Brief Pain Inventory. Neuropathic pain was assessed with the 7-item painDETECT questionnaire. Pain-related disability was evaluated with the Pain Disability Index. Psychological outcomes included the Hospital Anxiety and Depression Scale and the Pain Catastrophizing Scale. These outcome measures have been widely used in prior treatment trials, and their psychometric properties have been validated in US patients with chronic pain. All instruments were delivered via REDCap (Research Electronic Data Capture; Vanderbilt University). All efficacy measures, except for PGIC, will be taken at baseline and 3 months. PGIC will be assessed at 3 months.

The sample size calculation will be designed to confirm the hypothesis that the treatment assigned to the DT group will be associated with a greater PGIC score compared to the TAU control group at 3 months of 0.6 points with an SD of 1.0. The calculation assumes 85% power and a 2-sided type I error rate of 0.05. The estimated sample size of 100 will be increased to a target recruitment of 115 to account for 15% drop-out. The primary analysis of treatment effects will be conducted in the intention-to-treat (ITT) population, which will include all randomized participants. In addition, a prespecified subgroup analysis will be carried out. The study protocol predicts that participants in the DT group will demonstrate the greatest treatment effects.

The mean PGIC score at 3 months and the mean baseline to 3-month change scores for the secondary efficacy measures will be compared between the DT and TAU treatment groups by a mixed model for repeated measures (MMRM) analysis. In the presence of missing data, an MMRM analysis typically has more power than a 2-sample *t* test or analysis of covariance model. The ITT model will include fixed effects for treatment. A subgroup model will include all the parameters in the ITT model. This model will first be used to test for treatment heterogeneity by a significant interaction term at a 2-sided *P* value less than .15. A value of 1 ("no change") will be assigned as the 3-month PGIC score if there is 3-month data for a participant. The outcome vector of the secondary efficacy measures will include a change score of zero for the initial (baseline) visit to account for participants with no treatment data. Missing covariates will be addressed with the missing-indicator approach. Correlations among measurements taken on the same participant will first be modeled with an unstructured covariance assumption. If the model fails to converge, then a first-order autoregressive covariance structure will be used. Marginal effects of treatment will be determined at 3 months, along with corresponding 2-sided *P* values.

XSL•FO
RenderX

Comparisons will be deemed significant if the 2-sided *P* value is less than .05. Adjustments for multiple comparisons of secondary efficacy measures will be performed to account for the risk of elevated type I errors. Multiplicity corrections (eg, Bonferroni) will generally assume that outcomes are independent and will overcompensate for correlated measures, leading to increased type II errors. In this study, moderate correlations among many of the efficacy measures are expected. An MMRM analysis using all available data will impute missing data under a missing at random assumption. A sensitivity analysis will be performed using reference-based multiple imputation with the jump-to-reference method. The multiple imputation model will include the same covariates as the primary MMRM analysis. The imputed datasets will be analyzed with the original MMRM model and combined using the Rubin rule. The impact of missing data will be assessed by comparing the resulting estimates with the primary MMRM estimates.

Responder rates will be compared between treatment groups using logistic regression. Missing outcomes due to study withdrawal will be treated as nonresponders. The dependent variable in the model will be a binary variable indicating whether the participant was a responder or nonresponder. The model includes treatment assignment as an independent variable and baseline pain severity (Brief Pain Inventory average pain item) as a covariate. An interaction term between treatment and baseline pain sensitivity will be added for subgroup analyses of responder rates.

Results

Training of the study personnel will begin in March 2026 after funding has been received and IRB approval. Recruitment of participants is expected to begin by June 2026. The recruitment phase of the study is expected to be completed in 6 months. The active phase of patient tracking is expected to last 1 year. Because participant enrollment will be staggered, the active phase of the study would be expected to last 18 months. It is expected that finalizing of study data and analysis, along with manuscript preparation, will begin 30 months from study initiation and last 6 months. Our hypothesis is that results will show that DT modeling using the biopsychosocial characteristics of each patient will be statistically significant, supporting the use of this approach for personalized medical care.

Discussion

Principal Findings

This paper offers a review of the literature about digital twin (DT) modeling and describes a study protocol for the use of DT methods to help persons with noncommunicable medical conditions of cardiovascular disease, type 2 diabetes, and chronic obstructive pulmonary disease. This study will use a smartphone app to monitor individuals with these conditions to help tailor treatment recommendations for each user. The HIPAA-compliant app will be supported by AI and computational analysis. Through this project, we intend to develop and validate a novel computational framework integrating AI-driven DT modeling to help personalize treatment.

While RCTs remain the gold standard for evidence-based medicine, clinicians treat heterogeneous patient populations whose disease presentation, treatment responses, and side effects vary greatly between individuals. Traditional RCTs rarely provide the multivariable data needed for evidence-based personalized medicine [56,99]. For clinicians, implementing personalized medicine remains complicated and time-consuming, and often these providers must make medical decisions without adequate support. Personalized medicine using DT methodology addresses individual variability by tailoring treatment to each person with the incorporation and benefit of genetic, physiological, biochemical, behavioral, and environmental data [24,55].

It is anticipated that this study will provide a methodology to validate the use of DTs for NCDs. NCDs are the leading cause of death and negative health-related impact in the world today. A body of evidence-based medicine has identified high-quality, effective treatment regimens that have contributed to internationally recognized treatment guidelines. Although evidence-based, authoritative, and widely accepted, the application of these guidelines has not yielded the expected results. Without a personalized medicine approach, application of these guidelines into clinical practice comes with significant shortcomings, as they rarely consider individual differences and the diversity of each individual, which are identified as biopsychosocial characteristics. Personalized medicine has been shown to help overcome these shortcomings. A key technology to deliver personalized medicine is the use of the DT.

We intend, as part of this proposed study, to develop a specific NCD DT Health Management App using AI and machine learning for the management of NCDs. Pilot data collection will include clinical measures and laboratory results, lifestyle and behavioral data, environmental factors, patient-reported outcomes, and wearable device metrics. Reflecting concerns about the lack of diversity found in evidence-based medicine, training of the algorithms needed to formulate the specific NCD DT will be monitored and corrected, as it is identified through a feedback loop incorporated within this study. Weekly feedback to users and providers will serve as input for decision-making and behavior change. We believe that this proposed preliminary study will offer insight into the development and implementation of DTs designed to reduce risk factors associated with the onset of NCDs. Ideally, the protocol described within this paper, incorporating a biopsychosocial approach, will help to provide the necessary data for future use in the development of AI algorithms. If the DT modeling is found to be effective and, as these algorithms are refined, personalized medicine using contextual messaging and predictive risk assessment will become a more viable treatment option to improve treatment outcomes and reduce health care use and expenditures.

A strength of this study resides in the ability of AI to construct a more complete DT of each unique patient. However, we anticipate several challenges with the conduct of this study. First, designing a scalable data architecture that securely integrates retrospective clinical data and real-time patient input will be an undertaking requiring input from many different sources. We plan to use a means for data extraction by integrating with the hospital medical record (HMR) system

XSL•FO
RenderX

through the use of medical record numbers without jeopardizing security. We will incorporate the expertise of specialty clinicians, such as specialized medicine physicians, behavioral health clinicians, primary care providers, neuroscience researchers, IT technicians, and persons with lived experiences in dealing with NCDs. In preparation for this investigation, we have identified collaborators with the needed expertise to assist with this study.

Second, developing learning algorithms through the use of AI that update mechanistic parameters from longitudinal data will be challenging. Demonstrating predictive validity of the DT as it relates to real-world patient outcomes will be important to determine. We acknowledge a potential weakness of this study that may include unrecognized biases within an AI algorithm and its reliance on the current literature and published guidelines to generate treatment recommendations.

Finally, the study eligibility criteria have intentionally been designed to enable enrollment of participants resembling a real-life population of patients with NCDs. Result generalizability can be limited, however, by factors inherent to research that are not typical of clinical settings. Adherence to using the app, engaging in close follow-up with the study participants by the research personnel, and offering incentives for participation will be important. Studies have identified many patient factors that contribute to poor response to treatments. These include age, sensitivity to medications, emotional distress and catastrophizing, physical disability, early childhood trauma, history of substance use disorder, smoking cigarettes, and multiple medical comorbidities [84,96], which will be taken into consideration when evaluating use of the DT app. Also, some re-engineering to address security vulnerabilities and to integrate advanced AI capabilities for personalized predictive analytics may likely be required. The literature has noted the vital involvement of clinicians in eventually embracing the use of software in clinical practice [100-102]. We expect that the results of this trial will be useful in understanding how individuals who experience NCDs and those providers who treat them might incorporate the use of a DT app in daily practice. If the use of a DT app is found to be an effective treatment option for adults with NCDs, it will hopefully become a valuable tool to improve quality of life, increase function, reduce suffering, and address the large global burden of these conditions.

Conclusions

NCDs are the leading cause of death in the world today. This is despite evidence-based, effective self-care and treatment options being widely available. NCDs are recognized as behavior-based diseases. We propose the use of a DT that will personalize care by including biopsychosocial characteristics in treatment plans, along with ongoing feedback to the patient that will lead to increased compliance with treatment recommendations and improvement in this high-impact public health crisis.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Communication methods and DT process elements for NCD care.
[DOCX File, 21 KB - resprot_v15i1e75934_app1.docx]

References

1. Vallée A. Digital twin for healthcare systems. Front Digit Health 2023;5:1253050. [doi: 10.3389/fdgth.2023.1253050] [Medline: 37744683]
2. Turner BE, Steinberg JR, Weeks BT, Rodriguez F, Cullen MR. Race/ethnicity reporting and representation in US clinical trials: a cohort study. Lancet Reg Health Am 2022 Jul;11:11. [doi: 10.1016/j.lana.2022.100252] [Medline: 35875251]
3. Boyle P, Writer S. Clinical trials seek to fix their lack of racial mix. AAMC NEWS. 2021. URL: https://www.aamc.org/news/clinical-trials-seek-fix-their-lack-racial-mix [accessed 2026-01-15]
4. Mehta GU, Pazdur R, Bumpus NN, Califf RM. When diversity goals meet multiregional trials. N Engl J Med 2024 Dec 28. [doi: 10.1056/NEJMp2409392] [Medline: 39752557]
5. Ramamoorthy A, Araojo R, Vasisht KP, Fienkeng M, Green DJ, Madabushi R. Promoting clinical trial diversity: a highlight of select US FDA initiatives. Clin Pharma and Therapeutics 2023 Mar;113(3):528-535 [FREE Full text] [doi: 10.1002/cpt.2822]
6. Dattani S, Rodes-Giurao M, Ritchie H, Ortiz-Ospina E, Roser M. Life expectancy. : Our World in Data; 2013 URL: https://ourworldindata.org/life-expectancy [accessed 2026-01-15]

XSL•FO
RenderX

7. Ng M, Dai X, Cogen RM, et al. National-level and state-level prevalence of overweight and obesity among children, adolescents, and adults in the USA, 1990–2021, and forecasts up to 2050. The Lancet 2024 Dec;404(10469):2278-2298. [doi: 10.1016/S0140-6736(24)01548-4]
8. Chen S, Kuhn M, Prettner K, Bloom DE. The macroeconomic burden of noncommunicable diseases in the United States: Estimates and projections. PLoS ONE 2018;13(11):e0206702. [doi: 10.1371/journal.pone.0206702]
9. Li Y, Schoufour J, Wang DD, et al. Healthy lifestyle and life expectancy free of cancer, cardiovascular disease, and type 2 diabetes: prospective cohort study. BMJ 2020:l6669. [doi: 10.1136/bmj.l6669]
10. Kohl HW 3rd, Craig CL, Lambert EV, et al. The pandemic of physical inactivity: global action for public health. The Lancet 2012 Jul;380(9838):294-305. [doi: 10.1016/S0140-6736(12)60898-8] [Medline: 23041199]
11. Preventing noncommunicable diseases (NCDs) by reducing environmental risk factors. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/WHO-FWC-EPE-17.01 [accessed 2025-12-30]
12. Pate RR, Pratt M, Blair SN, et al. Physical activity and public health. A recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995 Feb 1;273(5):402-407. [doi: 10.1001/jama.273.5.402] [Medline: 7823386]
13. Horwitz RI, Lobitz G, Mawn M, et al. Biosocial medicine: biology, biography, and the tailored care of the patient. SSM Popul Health 2021 Sep;15:100863. [doi: 10.1016/j.ssmph.2021.100863] [Medline: 34430699]
14. Tramonti F, Giorgi F, Fanali A. Systems thinking and the biopsychosocial approach: a multilevel framework for patient‐centred care. Syst Res Behav Sci 2021 Mar;38(2):215-230 [FREE Full text] [doi: 10.1002/sres.2725]
15. Franzago M, Santurbano D, Vitacolonna E, Stuppia L. Genes and diet in the prevention of chronic diseases in future generations. Int J Mol Sci 2020 Apr 10;21(7):2633. [doi: 10.3390/ijms21072633] [Medline: 32290086]
16. McKee M, Haines A, Ebrahim S, et al. Towards a comprehensive global approach to prevention and control of NCDs. Global Health 2014 Dec;10(1):1-7. [doi: 10.1186/s12992-014-0074-8]
17. Ezzati M, Riboli E. Can noncommunicable diseases be prevented? Lessons from studies of populations and individuals. Science 2012 Sep 21;337(6101):1482-1487. [doi: 10.1126/science.1227001] [Medline: 22997325]
18. Mendis S. The policy agenda for prevention and control of non-communicable diseases. Br Med Bull 2010;96(1):23-43. [doi: 10.1093/bmb/ldq037] [Medline: 21059733]
19. NCDs in emergencies. World Health Organization. 2016. URL: https://www.emro.who.int/noncommunicable-diseases/publications/questions-and-answers-on-prevention-and-control-of-noncommunicable-diseases-in-emergencies.html [accessed 2025-12-30]
20. Novelli G, Biancolella M, Latini A, Spallone A, Borgiani P, Papaluca M. Precision medicine in non-communicable diseases. High Throughput 2020 Feb 7;9(1):3. [doi: 10.3390/ht9010003] [Medline: 32046063]
21. Iso H, Kokubo Y. Global strategies for non-communicable disease in terms of predictive, preventive and personalised medicine. EPMA Journal 2014 Feb;5(S1):A4. [doi: 10.1186/1878-5085-5-S1-A4]
22. Bousquet J, Anto JM, Sterk PJ, et al. Systems medicine and integrated care to combat chronic noncommunicable diseases. Genome Med 2011 Jul 6;3(7):1-12. [doi: 10.1186/gm259] [Medline: 21745417]
23. Tackling NCDs: best buys and other recommended interventions for the prevention and control of noncommunicable diseases, 2nd ed. : World Health Organization; 2024 URL: https://www.who.int/publications/i/item/9789240091078 [accessed 2026-01-15]
24. Budreviciute A, Damiati S, Sabir DK, et al. Management and prevention strategies for non-communicable diseases (NCDs) and their risk factors. Front Public Health 2020;8:574111. [doi: 10.3389/fpubh.2020.574111] [Medline: 33324597]
25. Thakur J, S K, Paika R, et al. World NCD Federation guidelines for prevention, surveillance and management of noncommunicable diseases at primary and secondary health-care for low resource settings. Int J Non-Commun Dis 2020;5(5):1. [doi: 10.4103/2468-8827.291731]
26. Mirza S, Clay RD, Koslow MA, Scanlon PD. COPD Guidelines: a review of the 2018 GOLD report. Mayo Clin Proc 2018 Oct;93(10):1488-1502. [doi: 10.1016/j.mayocp.2018.05.026] [Medline: 30286833]
27. Flack JM, Calhoun D, Schiffrin EL. The new ACC/AHA hypertension guidelines for the prevention, detection, evaluation, and management of high blood pressure in adults. Am J Hypertens 2018 Jan 12;31(2):133-135. [doi: 10.1093/ajh/hpx207] [Medline: 29228096]
28. Chamberlain JJ, Rhinehart AS, Shaefer CF Jr, Neuman A. Diagnosis and management of diabetes: synopsis of the 2016 American Diabetes Association standards of medical care in diabetes. Ann Intern Med 2016 Apr 19;164(8):542-552. [doi: 10.7326/M15-3016] [Medline: 26928912]
29. Dahan A, van Velzen M, Niesters M. Comorbidities and the complexities of chronic pain. Anesthesiology 2014 Oct;121(4):675-677. [doi: 10.1097/ALN.0000000000000402] [Medline: 25099749]
30. Lee M, Lee H, Song K, Lee YM. Benefits of sustained physical activity from middle age to early old age on quality of life in early old age. Sci Rep 2022;12(1):16455. [doi: 10.1038/s41598-022-20431-0]
31. Posadzki P, Pieper D, Bajpai R, et al. Exercise/physical activity and health outcomes: an overview of Cochrane systematic reviews. BMC Public Health 2020 Nov 16;20(1):1724. [doi: 10.1186/s12889-020-09855-3] [Medline: 33198717]

XSL•FO
RenderX

32. Shreves AH, Small SR, Walmsley R, et al. Amount and intensity of daily total physical activity, step count and risk of incident cancer in the UK Biobank. Br J Sports Med 2025 Jun 3;59(12):839-847. [doi: 10.1136/bjsports-2024-109360] [Medline: 40139674]
33. Prochaska JJ, Spring B, Nigg CR. Multiple health behavior change research: an introduction and overview. Prev Med 2008 Mar;46(3):181-188. [doi: 10.1016/j.ypmed.2008.02.001] [Medline: 18319098]
34. Siu KC, Hao J. Psychological aspects of virtual reality in chronic non-communicable diseases. In: Handbook of the Behavior and Psychology of Disease: Springer; 2024:1-20. [doi: 10.1007/978-3-031-32046-0_128-1]
35. Mukeshimana M, Mchunu G. The co-morbidity of depression and other chronic non-communicable diseases: a review of literature on the epidemiology, diagnosis and health effects. RWJour 2016;3(1):44. [doi: 10.4314/rj.v3i1.8F]
36. Oldenburg B, O Neil A, Cocker F. Public health perspectives on the co-occurrence of non-communicable diseases and common mental disorders. In: Comorbidity of Mental and Physical Disorders: Karger Publishers; 2015:15-22. [doi: 10.1159/000365524]
37. Schueck A, von Wolff M, Bitterlich N, Stute P. Awareness, use and applicability of online risk calculators for non-communicable diseases-A cross-sectional study. Maturitas 2018 Jul;113:1-6. [doi: 10.1016/j.maturitas.2018.03.015] [Medline: 29903642]
38. Juchli F, Zangger M, Schueck A, von Wolff M, Stute P. Chronic non-communicable disease risk calculators - An overview, part I. Maturitas 2021 Jan;143:25-35. [doi: 10.1016/j.maturitas.2020.07.009] [Medline: 33308633]
39. Juchli F, Zangger M, Schueck A, von Wolff M, Stute P. Chronic non-communicable disease risk calculators - an overview, part II. Maturitas 2021 Jan;143:132-144. [doi: 10.1016/j.maturitas.2020.10.003] [Medline: 33308619]
40. Duan Y, Shang B, Liang W, Du G, Yang M, Rhodes RE. Effects of eHealth-based multiple health behavior change interventions on physical activity, healthy diet, and weight in people with noncommunicable diseases: systematic review and meta-analysis. J Med Internet Res 2021 Feb 22;23(2):e23786. [doi: 10.2196/23786] [Medline: 33616534]
41. Mair JL, Salamanca-Sanabria A, Augsburger M, et al. Effective behavior change techniques in digital health interventions for the prevention or management of noncommunicable diseases: an umbrella review. Ann Behav Med 2023 Sep 13;57(10):817-835. [doi: 10.1093/abm/kaad041] [Medline: 37625030]
42. Grieves M. Intelligent digital twins and the development and management of complex systems. Digital Twin 2022;2:8. [doi: 10.12688/digitaltwin.17574.1]
43. Hornbrook MC, Hurtado AV, Johnson RE. Health care episodes: definition, measurement and use. Med Care Rev 1985;42(2):163-218. [doi: 10.1177/107755878504200202] [Medline: 10274864]
44. Hussey PS, Friedberg MW, Anhang Price R, Lovejoy SL, Damberg CL. Episode-based approaches to measuring health care quality. Med Care Res Rev 2017 Apr;74(2):127-147. [doi: 10.1177/1077558716630173] [Medline: 26896470]
45. Brailer DJ, Kroch EA. Member risk adjustment for ambulatory episodes of care. Health Care Manag Sci 1999 Jul;2(3):125-136. [doi: 10.1023/a:1019091929397] [Medline: 10934537]
46. Kim J, Ock M, Oh IH, et al. Comparison of diagnosis-based risk adjustment methods for episode-based costs to apply in efficiency measurement. BMC Health Serv Res 2023;23(1):1334. [doi: 10.1186/s12913-023-10282-4]
47. MaCurdy T, Shafrin J, Hartmann E, et al. Challenges in the risk adjustment of episode costs. Research (Acumen, LLC). 2010 Feb. URL: www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Reports/downloads/MaCurdy_ERA_2010.pdf [accessed 2026-01-22]
48. Young T, Brailsford S, Connell C, Davies R, Harper P, Klein JH. Using industrial processes to improve patient care. BMJ 2004 Jan 17;328(7432):162-164. [doi: 10.1136/bmj.328.7432.162] [Medline: 14726351]
49. Sambare TD, Bozic KJ. Preparing for an era of episode-based care in total joint arthroplasty. J Arthroplasty 2021 Mar;36(3):810-815. [doi: 10.1016/j.arth.2020.09.028] [Medline: 33069550]
50. Grocott MP, Levett DZ. Redesign of perioperative care pathways. In: Perioperative Care of the Cancer Patient: Elsevier:185-192. [doi: 10.1016/B978-0-323-69584-8.00016-5]
51. Viceconti M, De Vos M, Mellone S, Geris L. Position paper from the digital twins in healthcare to the virtual human twin: a moon-shot project for digital health research. IEEE J Biomed Health Inform 2023 Oct 11;PP(1):491-501. [doi: 10.1109/JBHI.2023.3323688] [Medline: 37819828]
52. Sankar PL, Parker LS. The Precision Medicine initiative's All of Us Research Program: an agenda for research on its ethical, legal, and social issues. Genet Med 2017 Jul;19(7):743-750. [doi: 10.1038/gim.2016.183] [Medline: 27929525]
53. Investigators A. The "All of Us" Research Program. N Engl J Med 2019 Aug 15;381(7):668-676. [doi: 10.1056/NEJMsr1809937]
54. Ramirez AH, Sulieman L, Schlueter DJ, et al. The All of Us Research Program: data quality, utility, and diversity. Patterns (N Y) 2022 Aug 12;3(8):100570. [doi: 10.1016/j.patter.2022.100570] [Medline: 36033590]
55. Goetz LH, Schork NJ. Personalized medicine: motivation, challenges, and progress. Fertil Steril 2018 Jun;109(6):952-963. [doi: 10.1016/j.fertnstert.2018.05.006] [Medline: 29935653]
56. Laska E, Siegel C, Lin Z. A likely responder approach for the analysis of randomized controlled trials. Contemp Clin Trials 2022 Mar;114:106688. [doi: 10.1016/j.cct.2022.106688] [Medline: 35085831]
57. Edwards RR, Dworkin RH, Turk DC, et al. Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. PR9 2021;6(1):e896. [doi: 10.1097/PR9.0000000000000896]

XSL•FO
RenderX

58. Rajpurkar P, Chen E, Banerjee O, Topol EJ. AI in health and medicine. Nat Med 2022 Jan;28(1):31-38. [doi: 10.1038/s41591-021-01614-0]
59. He J, Baxter SL, Xu J, Xu J, Zhou X, Zhang K. The practical implementation of artificial intelligence technologies in medicine. Nat Med 2019 Jan;25(1):30-36. [doi: 10.1038/s41591-018-0307-0]
60. Meskó B, Görög M. A short guide for medical professionals in the era of artificial intelligence. NPJ Digit Med 2020;3(1):126. [doi: 10.1038/s41746-020-00333-z]
61. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med Educ 2023 Sep 22;23(1):689. [doi: 10.1186/s12909-023-04698-z] [Medline: 37740191]
62. Hussain-Gambles M, Atkin K, Leese B. Why ethnic minority groups are under-represented in clinical trials: a review of the literature. Health Soc Care Community 2004 Sep;12(5):382-388. [doi: 10.1111/j.1365-2524.2004.00507.x] [Medline: 15373816]
63. De las Nueces D, Hacker K, DiGirolamo A, Hicks LS. A systematic review of community-based participatory research to enhance clinical trials in racial and ethnic minority groups. Health Serv Res 2012 Jun;47(3 Pt 2):1363-1386. [doi: 10.1111/j.1475-6773.2012.01386.x] [Medline: 22353031]
64. Hsieh N, Shuster SM. Health and health care of sexual and gender minorities. J Health Soc Behav 2021 Sep;62(3):318-333. [doi: 10.1177/00221465211016436] [Medline: 34528481]
65. Khan B, Fatima H, Qureshi A, et al. Drawbacks of artificial intelligence and their potential solutions in the healthcare sector. Biomedical Materials & Devices 2023 Sep;1(2):731-738. [doi: 10.1007/s44174-023-00063-2]
66. Syed K, et al. Artificial intelligence methods in computer-aided diagnostic tools and decision support analytics for clinical informatics. In: Artificial Intelligence in Precision Health: Elsevier:31-59. [doi: 10.1016/B978-0-12-817133-2.00002-1]
67. Schork NJ. Artificial intelligence and personalized medicine. Cancer Treat Res 2019;178:265-283. [doi: 10.1007/978-3-030-16391-4_11] [Medline: 31209850]
68. Botín-Sanabria DM, Mihaita AS, Peimbert-García RE, Ramírez-Moreno MA, Ramírez-Mendoza RA, Lozoya-Santos JDJ. Digital twin technology challenges and applications: a comprehensive review. Remote Sens (Basel) 2022;14(6):1335. [doi: 10.3390/rs14061335]
69. Singh M, Srivastava R, Fuenmayor E, et al. Applications of Digital Twin across Industries: A Review. Appl Sci (Basel) 2022;12(11):5727. [doi: 10.3390/app12115727] [Medline: 35983461]
70. Spitzer M, Dattner I, Zilcha-Mano S. Digital twins and the future of precision mental health. Front Psychiatry 2023;14:1082598. [doi: 10.3389/fpsyt.2023.1082598] [Medline: 36993921]
71. Shu H, Liang R, Li Z, et al. Twin-S: a digital twin for skull base surgery. Int J CARS 2023;18(6):1077-1084. [doi: 10.1007/s11548-023-02863-9]
72. Hagmann K, Hellings-Kuß A, Klodmann J, Richter R, Stulp F, Leidner D. A digital twin approach for contextual assistance for surgeons during surgical robotics training. Front Robot AI 2021;8:735566. [doi: 10.3389/frobt.2021.735566] [Medline: 34621791]
73. Wooley A, Silva DF, Bitencourt J. When is a simulation a digital twin? A systematic literature review. Manuf Lett 2023 Aug;35:940-951. [doi: 10.1016/j.mfglet.2023.08.014]
74. Sun T, He X, Song X, Shu L, Li Z. The digital twin in medicine: a key to the future of healthcare? Front Med 2022;9:907066. [doi: 10.3389/fmed.2022.907066]
75. Critchley LA, Yang XX, Lee A. Assessment of trending ability of cardiac output monitors by polar plot methodology. J Cardiothorac Vasc Anesth 2011 Jun;25(3):536-546. [doi: 10.1053/j.jvca.2011.01.003] [Medline: 21419654]
76. Perkins BA, Nelson RG, Ostrander BEP, et al. Detection of renal function decline in patients with diabetes and normal or elevated GFR by serial measurements of serum cystatin C concentration. J Am Soc Nephrol 2005 May;16(5):1404-1412. [doi: 10.1681/ASN.2004100854] [Medline: 15788478]
77. Patrone C. The role of internet of things and digital twin in healthcare digitalization process. In: Transactions on Engineering Technologies: World Congress on Engineering and Computer Science 2018: Springer; 2020. [doi: 10.1007/978-981-15-6848-0_3]
78. Noncommunicable diseases country profiles 2018. World Health Organization. URL: https://www.who.int/teams/noncommunicable-diseases/surveillance/data/profiles-ncd [accessed 2025-12-30]
79. Jamison R, Deeney M, Al Hamdan E, Ross E, Edwards R. (150) Prevalance of chronic pain and associated risk factors: intake data from a rehabilitation center in Kuwait. J Pain 2015 Apr;16(4):S13. [doi: 10.1016/j.jpain.2015.01.064]
80. Zghoul N, Ross EL, Edwards RR, Ahmed A, Jamison RN. Prevalence of chronic pain with neuropathic characteristics: a randomized telephone survey among medical center patients in Kuwait. JPR 2017. [doi: 10.2147/JPR.S123966]
81. Deeny MC, Al Hamdan E, Ross EL, Edwards RR, Huang CC, Jamison RN. Chronic pain, comorbid medical conditions, and associated risk factors in Kuwait: gender and nationality differences. Pain Med 2015 Nov;16(11):2204-2211. [doi: 10.1111/pme.12840] [Medline: 26118589]
82. MobileNetrix. What is MobileNetrix?. URL: https://mobilenetrix.com/ [accessed 2026-01-12]
83. Jamison RN, Xu X, Wan L, Edwards RR, Ross EL. Determining pain catastrophizing from daily pain app assessment data: role of computer-based classification. J Pain 2019 Mar;20(3):278-287. [doi: 10.1016/j.jpain.2018.09.005] [Medline: 30273687]

XSL•FO
RenderX

84. Jamison RN, Jurcik DC, Edwards RR, Huang CC, Ross EL. A pilot comparison of a smartphone app with or without 2-way messaging among chronic pain patients: who benefits from a pain app? Clin J Pain 2017 Aug;33(8):676-686. [doi: 10.1097/AJP.0000000000000455] [Medline: 27898460]
85. Mariano TY, Wan L, Edwards RR, Lazaridou A, Ross EL, Jamison RN. Online group pain management for chronic pain: preliminary results of a novel treatment approach to teletherapy. J Telemed Telecare 2021 May;27(4):209-216. [doi: 10.1177/1357633X19870369] [Medline: 31431133]
86. Khan SS, Matsushita K, Sang Y, et al. Development and validation of the American Heart Association's PREVENT equations. Circulation 2024 Feb 6;149(6):430-449. [doi: 10.1161/CIRCULATIONAHA.123.067626] [Medline: 37947085]
87. Berry JD, Dyer A, Cai X, et al. Lifetime risks of cardiovascular disease. N Engl J Med 2012 Jan 26;366(4):321-329. [doi: 10.1056/NEJMoa1012848] [Medline: 22276822]
88. Bang H, Edwards AM, Bomback AS, et al. Development and validation of a patient self-assessment score for diabetes risk. Ann Intern Med 2009 Dec 1;151(11):775-783. [doi: 10.7326/0003-4819-151-11-200912010-00005] [Medline: 19949143]
89. Kramer CK, Araneta MRG, Barrett-Connor E. A1C and diabetes diagnosis: The Rancho Bernardo Study. Diabetes Care 2010 Jan;33(1):101-103. [doi: 10.2337/dc09-1366] [Medline: 19837792]
90. American Diabetes Association. 2. Classification and diagnosis of diabetes. Diabetes Care 2016 Jan 1;39(Supplement_1):S13-S22. [doi: 10.2337/dc16-S005]
91. Stanley AJ, Hasan I, Crockett AJ, van Schayck OCP, Zwar NA. COPD Diagnostic Questionnaire (CDQ) for selecting at-risk patients for spirometry: a cross-sectional study in Australian general practice. NPJ Prim Care Respir Med 2014 Jul 10;24(1):14024. [doi: 10.1038/npjpcrm.2014.24] [Medline: 25010522]
92. Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 Report. GOLD Executive Summary. Am J Respir Crit Care Med 2017 Mar 1;195(5):557-582. [doi: 10.1164/rccm.201701-0218PP] [Medline: 28128970]
93. Watkins DA, Msemburi WT, Pickersgill SJ, et al. NCD Countdown 2030: efficient pathways and strategic investments to accelerate progress towards the Sustainable Development Goal target 3.4 in low-income and middle-income countries. The Lancet 2022 Mar;399(10331):1266-1278. [doi: 10.1016/S0140-6736(21)02347-3]
94. Cohen J. Statistical Power Analysis for the Behavioral Sciences: Routledge; 2013. [doi: 10.4324/9780203771587]
95. Jamison RN, Wan L, Edwards RR, Mei A, Ross EL. Outcome of a high-frequency transcutaneous electrical nerve stimulator (hfTENS) device for low back pain: a randomized controlled trial. Pain Pract 2019 Jun;19(5):466-475. [doi: 10.1111/papr.12764] [Medline: 30636101]
96. Jamison RN, Edwards RR, Curran S, et al. Effects of wearable transcutaneous electrical nerve stimulation on fibromyalgia: a randomized controlled trial. J Pain Res 2021;14:2265-2282. [doi: 10.2147/JPR.S316371] [Medline: 34335055]
97. Jamison RN, Curran S, Wan L, Ross EL, Gilligan CJ, Edwards RR. Higher pain sensitivity predicts efficacy of a wearable transcutaneous electrical nerve stimulation device for persons with fibromyalgia: a randomized double-blind sham-controlled trial. Neuromodulation 2022 Dec;25(8):1410-1420. [doi: 10.1111/ner.13463] [Medline: 34056781]
98. Dworkin RH, Evans SR, Mbowe O, McDermott MP. Essential statistical principles of clinical trials of pain treatments. PR9 2021;6(1):e863. [doi: 10.1097/PR9.0000000000000863]
99. McKee M, Haines A, Ebrahim S, et al. Towards a comprehensive global approach to prevention and control of NCDs. Global Health 2014 Dec;10(1):74. [doi: 10.1186/s12992-014-0074-8]
100. Choudhury A, Asan O, Mansouri M. Role of artificial intelligence, clinicians & policymakers in clinical decision making: a systems viewpoint. Presented at: 2019 International Symposium on Systems Engineering (ISSE); Oct 1-3, 2019. [doi: 10.1109/ISSE46696.2019.8984573]
101. Hah H, Goldin DS. How clinicians perceive artificial intelligence–assisted technologies in diagnostic decision making: mixed methods approach. J Med Internet Res 2021 Dec 16;23(12):e33540. [doi: 10.2196/33540] [Medline: 34924356]
102. Wysocki O, Davies JK, Vigo M, et al. Assessing the communication gap between AI models and healthcare professionals: Explainability, utility and trust in AI-driven clinical decision-making. Artif Intell 2023 Mar;316:103839. [doi: 10.1016/j.artint.2022.103839]

Abbreviations

AI: artificial intelligence
API: application programming interface
DT: digital twin
EDITH: European Virtual Human Twin Initiative
EOC: episode of care
HIPAA: Health Insurance Portability and Accountability Act
IRB: Institutional Review Board
ITT: intention-to-treat
LLM: large language model
MMRM: mixed model for repeated measures

XSL•FO
RenderX

NCD: noncommunicable disease
PGIC: Patient Global Impression of Change
PLM: product lifecycle management
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
TAU: treatment as usual

Edited by J Sarvestan; submitted 13.Apr.2025; peer-reviewed by M Collier, R Buschman, S Boodai; revised version received 14.Dec.2025; accepted 15.Dec.2025; published 09.Feb.2026.

Please cite as:
Ross E, Ross J, Beschler P, Guydan D, Jamison R
Designing a Digital Twin for the Management of Noncommunicable Diseases: Protocol for a Pilot Study and Methodology Validation
JMIR Res Protoc 2026;15:e75934
URL: https://www.researchprotocols.org/2026/1/e75934
doi:10.2196/75934

© Edgar Ross, Jason Ross, Patricia Beschler, David Guydan, Robert Jamison. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using AI Algorithms and Machine Learning in the Analysis of a Bio-Purification Method (Therapeutic Emesis, Known as "Vamana Karma"): Protocol for a Mixed Methods Study

Pooja Rani[1*], BAMS, MD; Sumit Kalra[2*], B Tech, M Tech, PhD; Sachin Singh[2], B Tech CE-SE; Richard David[2], BS (EECS); Ashutosh Ravi Gupta[2], B Tech, M Tech; Anandaraman P V[1*], BAMS, MD, PhD

[1]Department of Panchakarma, All India Institute of Ayurveda, Mathura Rd, Gautampuri Awas, Sarita Vihar, New Delhi, Delhi, India

[2]Department of Computer Science and Engineering, Indian Institute of Technology Jodhpur, Jodhpur, Rajasthan, India

[*]these authors contributed equally

Corresponding Author:
Pooja Rani, BAMS, MD
Department of Panchakarma, All India Institute of Ayurveda, Mathura Rd, Gautampuri Awas, Sarita Vihar, New Delhi, Delhi, India

Abstract

Background: Therapeutic emesis (TE), known as *vamana karma*, is a classical method of detoxification performed to eliminate vitiated *kapha* (bio-humor governing fluid regulation and structural cohesion of the body in normalcy) ailments from the body. The assessment of this complete process depends on physicians' visual assessments of vomitus features and patient responses, introducing subjectivity and interobserver variability. Moreover, this method requires more than continuous monitoring; thus, it can sometimes lead to human error, resulting in missed expelled content or complications. We propose an artificial intelligence (AI) model to monitor TE to observe visual changes (ie, patient vomitus content and gestures) to provide better clinical outcomes. This approach is being explored for the first time in the traditional system of medicine.

Objective: This study aims to develop and validate an AI-assisted digital framework for the objective evaluation of TE via (1) automatic vomitus detection, (2) content classification, (3) number of bouts expelled, (4) facial expressions and individual gestures, (5) determination of detoxification type, and (6) provision of a postpurificatory dietary regimen after completion.

Methods: The study will be conducted in 3 phases. The first is the preparation of standard operating procedure for TE data collection. The second is data annotation of detected vomiting events. All analyses will be conducted using Python libraries, including *scikit-learn* (version 1.3.2; developed by the scikit-learn contributors, Python Software Foundation), *TensorFlow* (version 2.14.0; Google Brain Team, Google LLC), and tools supported under Google Summer of Code 2023 (Google LLC), along with SPSS Statistics (version 26.0; IBM Corp) for statistical analysis. In the third phase, model performance will be evaluated using standard machine learning metrics, and agreement with expert assessments will be measured using the Fleiss κ statistic. This study is exploratory in nature. Thus, 50 volunteers will be targeted.

Results: This is the first study of its kind, so to create the dataset, we prepared a standard operating procedure for TE event data collection. Data collection was completed in December 2025. Data annotation and preliminary model preparation are ongoing, with final testing and validation expected to be completed by December 2025. External testing in the health care setting is expected to be completed by February 2026.

Conclusions: This work presents one of the first attempts to apply deep learning for objective analysis of the TE process in Ayurveda. By combining YOLOv9 for vomit detection and residual neural network for classification, the framework achieves promising accuracy in automated vomit identification. The results will demonstrate the potential of AI-assisted analysis in traditional medicine, although further clinical validation and expansion across multiple centers will be necessary before deployment in real-world settings.

Trial Registration: Clinical Trials Registry- India CTRI/2023/03/051087; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=ODI4NTA=&Enc=&userName=

International Registered Report Identifier (IRRID): DERR1-10.2196/79875

(JMIR Res Protoc 2026;15:e79875) doi:10.2196/79875

KEYWORDS

validation; therapeutic emesis; TE; vamana karma; traditional medicine; protocol; artificial intelligence; AI

XSL•FO
RenderX

Introduction

Background

The traditional Indian system of medicine promotes overall health and wellness through personalized therapies and a detoxifying process. Therapeutic emesis (TE), or *vamana karma*, is one of the bio-pentavalent purification procedures used to treat deranged *kapha* ailments that include metabolic, respiratory, and dermatological conditions such as psoriasis and eczema [1]. TE is conducted in 3 phases (ie, preparatory phase, main phase, and postpurificatory phase). In the main phase, volunteers are subjected to induced vomiting through the provision of emetic herbal medicines. After initiation of the process, the individual expels the vitiated content, concurrently evident through physical changes (concurrently evident include the onset of nausea, increased salivation, abdominal discomfort, sweating, piloerection, lacrimation, and the progressive expulsion of gastric contents, reflecting effective mobilization and elimination of morbid bio-humor) [2]. This procedure is carried out early in the morning and takes approximately 1 hour, although it sometimes takes more than an hour to complete.

Previously, TE has been assessed subjectively by trained physicians through visual inspection–based emesis indicators, such as patient gestures that include nausea, discomfort, and distress and vomitus characteristics such as color, consistency, volume, and odor. Nevertheless, dependence on human knowledge may introduce observer bias and interindividual variability, affecting reliability in clinical outputs [3]. Furthermore, extended therapy is often time-consuming and poses challenges for physicians in observing and documenting continuously, potentially creating inconsistency and reducing standardization. Studies suggest that human error is one of the most significant contributors to decreased productivity and quality. Multifactorial interaction among stress, repeated tasks, fatigue, and workplace circumstances impairs cognitive performance, attention, and decision-making accuracy [4]. Similarly, the variable rapid nature of emesis creates difficulties in manual tracking due to the rapid onset of the process, thus limiting real-time monitoring.

Research highlights that the integration of artificial intelligence (AI) techniques in image analysis for clinical decision support showcases their potential in diverse areas. Today's health care practices have changed significantly with the help of AI and machine learning (ML), which help with more accurate diagnoses, custom interventions, and informed decisions [5,6]. Previous studies have substantiated that the use of AI-assisted imaging yields superior accuracy and specificity compared to traditional methods [7]. Further cutting-edge developments using YOLOv9, TensorFlow, and residual neural network (ResNet) models have yielded newer insights in scope for such assessment methods.

There have been many studies conducted on *vamana karma* focusing on its efficacy in diseases and biophysical parameters [8-10]. Presently, no study has been proposed to digitally analyze the procedure. This study has been designed to integrate for the first time an AI component for TE analysis. The primary objective is to support physicians in accurate interpretations of purification type and guiding patients to an appropriate postpurificatory diet. The primary aim is to assess appropriate signs and symptoms of TE through content analysis of vomitus by developing and validating an AI model using digital image processing and an AI algorithm. Additional objectives focus on developing a dataset of TE for training and validating an AI model, training the TE model for real-time assessment, and validating the performance of the developed AI model by comparing its assessments with those of expert Ayurvedic physicians.

Research Imperative

In the traditional system of medicine, TE is a key intervention for body detoxification, which requires documentation of all evident changes during the process. Recent studies have underscored the adoption of the YOLOv9 approach for detecting abnormal crowd behavior to promote crowd safety [11]. Correspondingly, another study used the same methodology in traffic hazards for public safety using the GC-YOLOv9 algorithm [12]. Additionally, there are studies in which image processing was used for microscopic sputum analysis and the velocity distance support algorithm was used for sputum monitoring to detect phlegm stagnation [13,14]. Contemporarily, one study was conducted to automate emesis in *Suncus murinus* using a convolutional neural network and a self-attention mechanism [15]. Such studies draw attention to developing a model that can automate emetic episodes in traditional sciences. Moreover, the convergence of such approaches remains unexplored. Leveraging existing technologies could enable real-time monitoring for automation, enhance reliability, and improve patient outcomes.

Methods

Study Design

This is a prospective, observational, single-center study aimed at developing and validating an AI-based framework for the objective assessment of *vamana karma* (TE), a classical purification therapy in Ayurveda. The study adheres to the Standards for Reporting Diagnostic Accuracy–Artificial Intelligence reporting guidelines [16]. It will ensure data handling with model development, validation, and performance. A checklist has been provided in Multimedia Appendix 1.

Study Setting

The study will be conducted at a tertiary Ayurveda research and teaching institute with an established *panchakarma* (bio-purification methods) center. A dedicated clinical setup will be designed to ensure uniform data collection across participants. The study workflow is depicted in Figure 1A and Figure 1B.

XSL•FO
RenderX

Figure 1. (A) Schematic representation of the data collection setup during *vamana karma* showing patient positioning, camera placement, vomit collection, and green backdrop for optimal video capture. (B) Study workflow and timeline illustrating the sequential phases of the research process, including data collection, manual annotation, machine learning model development, system optimization using trained datasets, and user testing for performance validation.

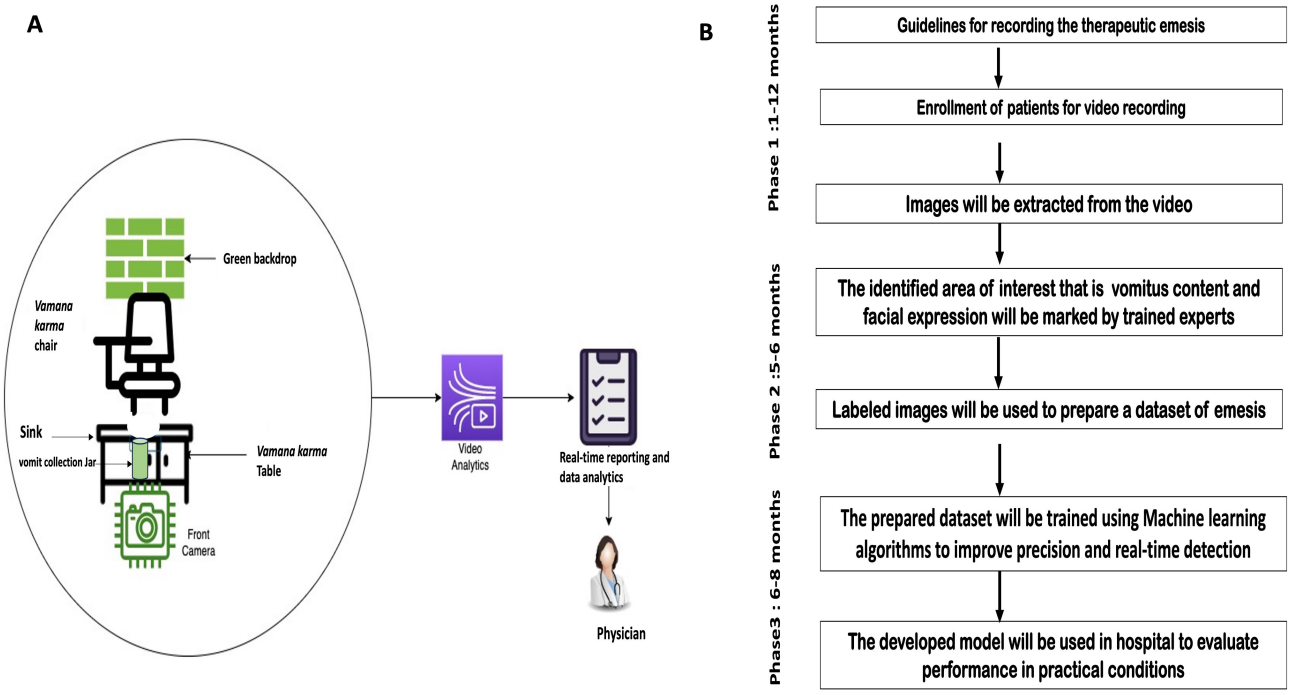

To achieve uniformity in data collection, we will adopt the following guidelines throughout the process:

- A monochromatic green backdrop will be used to optimize background segmentation for video analytics.
- Participants who undergo TE will be dressed in a green gown for the same purpose.
- Standardized artificial lighting will ensure uniform illumination and minimize the effect of varying light conditions.
- High-definition cameras (minimum 1080p resolution) will be positioned at forehead level facing the participant to record both the vomitus events and the patients' facial expressions.
- Vomitus will be collected in transparent jars placed beneath the *vamana karma* chair setup.

Participants

After providing informed consent, adult participants aged between 18 and 65 years will be enrolled in the study. Inclusion and exclusion criteria will be strictly adhered to for participant safety and data reliability. Inclusion criteria are adults clinically recommended for TE as per classic Ayurvedic guidelines for *kapha*-dominant conditions such as asthma, skin diseases, and metabolic disorders. Participants must be willing to consent to video recording during the procedure. Exclusion criteria include individuals with contraindications for *vamana karma*, including pregnancy, lactation, cardiac conditions, anemia, or known hypersensitivity to emetic medicine. Individuals unable or unwilling to provide informed consent will also be excluded.

Sample Size

A total of 40 high-resolution videos (approximately 70 minutes each) were used, yielding over 10,000 frames annotated for vomit detection. This dataset size was considered adequate for training a deep learning model given the high temporal density of labeled events and the pretrained backbone used (YOLOv9 and ResNet). A preliminary learning curve analysis indicated performance saturation beyond this dataset size. Thus, for this study, we targeted 50 participants.

Clinical Phases of TE

The complete process will be conducted in 3 phases, as shown in Figure 2. Phase 1, also known as the preparatory phase (*purva karma*), starts with the administration of *deepana* (digestive stimulants) and *pachana* (digestive agents), followed by *snehapana* (internal oleation) for 5 to 7 days accompanied by *abhyanga* (massage) and *svedana* (sudation) for 1 day. In this 1-day period, a *kapha*-aggravating diet will be administered. Thereafter, the second or main phase (*pradhana karma*) starts on the *vamana karma* day with *abhyanga* (massage) and *svedana* (sudation). Participants will be allowed to take the following emetic formulations early in the morning: (1) rice gruel (consistency, color, and texture), (2) milk (2 L), (3) 8 g of medicine (madanphala powder; *Randia dumetorum* Linn.)+5 g of madhuyashti powder (*Glycyrrhiza glabra* Linn.)+2 g of vacha powder (*Acorus calamus* Linn.)+5 g of salt+honey (q.s.), (4) 30 g of madhuyashti kwath (decoction of *G glabra* Linn.) powder dissolved in 1 L of hot water, and (5) salt water (20 g of salt dissolved in 1 L of water; 1 glass=380 mL). The number, content, and force of vomitus bouts will be recorded. Vomit will be collected in calibrated jars for visual analysis and recorded via video.

XSL•FO
RenderX

Figure 2. Procedural phases of *vamana karma*, including preparatory (*deepana-pachana*), internal oleation (*snehapana*), and rest period showing diet and therapeutic interventions.

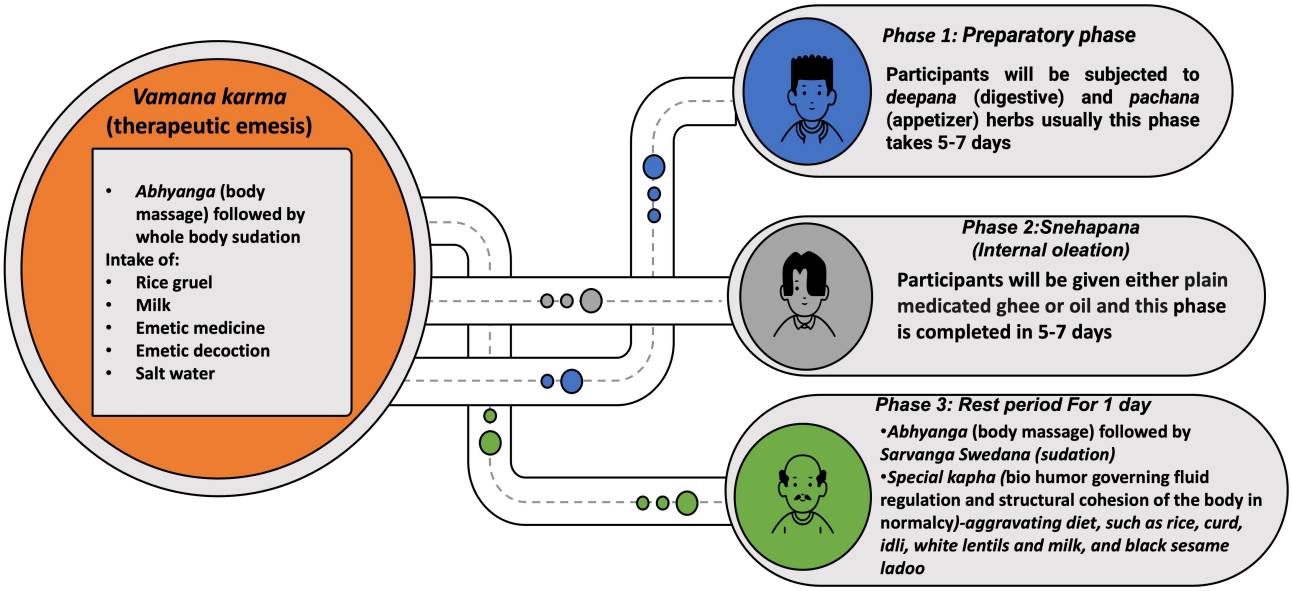

In the third or postprocedure phase (*paschat karma*), after completion of *vamana karma*, depending on the type of purification, the participants will be gradually shifted to *samsarjana karma* (postdetoxification dietary regimen) to normalize the weak digestive fire.

Video data will be captured throughout the procedure to record vomitus events, facial expressions, and physical gestures. The data will form the primary source for model training and validation.

Data Annotation

Video data will be processed through frame extraction and manual annotation to create labeled datasets for ML. More than 11,000 frames will be annotated for vomitus event detection, whereas approximately 700 to 800 cropped images will be classified as per classic parameters. All these annotations will serve as the ground truth for supervised learning and will be validated by experienced Ayurveda physicians to ensure accuracy. Data labeling details are listed in Table 1.

XSL•FO
RenderX

Table . Outcome parameters used for analysis and manual annotation. The outcome parameters constitute the core analytical content of the study and were derived from predefined clinical variables. All parameters were manually annotated and labeled by trained reviewers to ensure consistency and accuracy of the dataset used for analysis.

Outcome parameters	Labeled and description
Milk	Liquid, thin consistency and white color
Yavagu (rice milk, jaggery, and ghee)	Thick consistency; rice particles of brown color mixed with milk
Phanta	Water flow, thin consistency, and light yellow-brown color
Medicine	Thick pastelike consistency, less viscous, dark brown color, and less quantity with decreased flow
Salt water	Normal flow, watery consistency, and transparent
Pitta	Yellow color, thick consistency, normal flow, and much lower quantity
Facial expression analysis	The patients' facial gestures will be analyzed by mapping various features of the face, such as the eyebrows, eyes, and mouth, to the emotions of anger, fear, surprise, sadness, and happiness
Variations in the labeling of parameters	
Phanta+yavagu	Light yellowish-brown color and rice particles
Salt water+milk+medicine	Transparent; brown particles with a slight white color
Pitta+salt water	Water of a yellowish color
Pitta+medicine+rice	Yellow color, brown particle–like consistency, and rice particles
Pitta	Light yellow color
Rice particles	Thick consistency—particles the size of a small dot
Salt water+medicine	Water of a light brown color
Rice particles+medicine particles	Light brown color; medicine particles the size of minute dots
Milk mixed with medicine	Light brownish color
Milk+phanta	White color with light brown texture
Phanta+milk	Light yellowish-brown color; small milk particles
Phanta+pitta	Yellow color with a light yellowish-brown–colored liquid
Phanta+medicine	Light yellowish-brown color
Medicine+milk	White brown color; liquid, slightly thick consistency; and small particles with a white structure
Milk mixed with medicine	White appearance with a brown color texture
Milk+salt water	Transparent with white appearance
Phanta+medicine	Light yellowish-brown color, dark color, and more particles also present
Salt water+rice particles	Transparent water; rice particles appear dotted
Blood+milk	Red color mixed with white particles

ML Pipeline

Overview

The AI framework will integrate several ML components, as shown in Figure 3.

XSL•FO
RenderX

Figure 3. Flow diagram of input-output mapping for vomitus classification based on image analysis and machine learning data training pipeline.

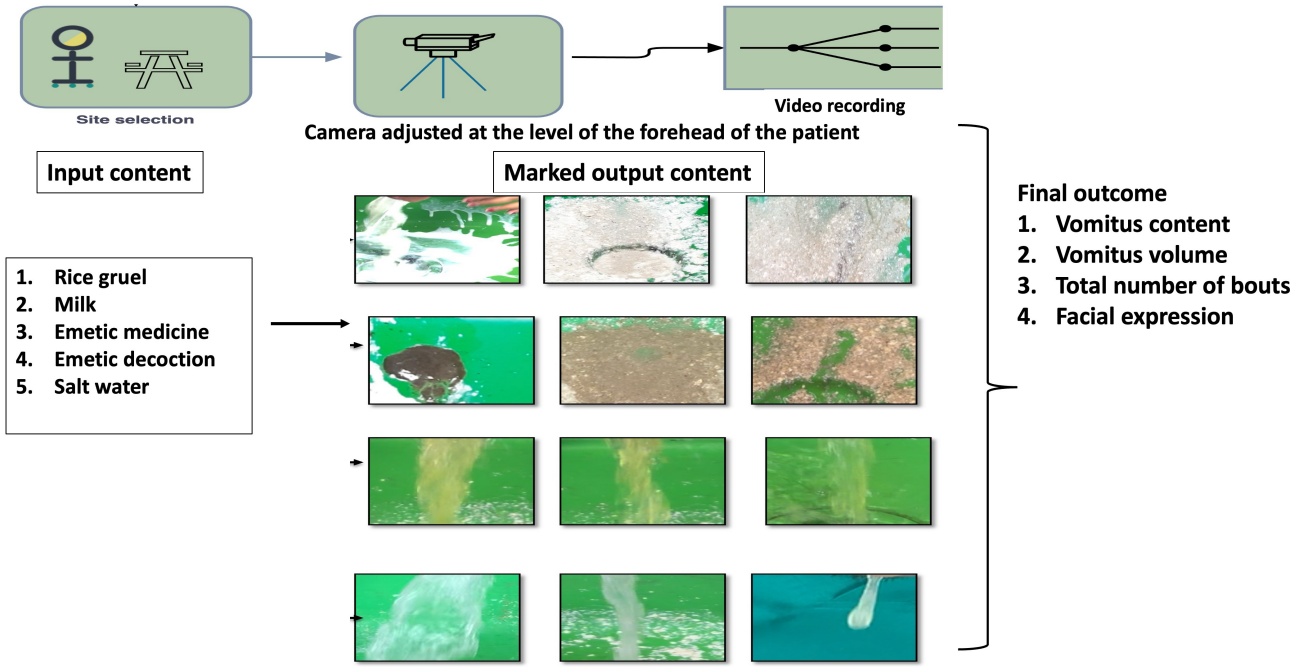

Object Detection

A YOLOv9 deep learning model will be trained to detect and localize vomitus streams in real-time video frames. The model will be trained on the annotated images and validated using k-fold cross-validation. Key performance metrics will include detection accuracy and precision-recall scores.

Content Classification

A convolutional neural network based on ResNet architecture will be developed for vomitus content classification. The model will categorize each vomitus image into one of the predefined classes. Data augmentation techniques will be applied to increase dataset variability and improve generalization.

Facial Expression and Gesture Analysis

Facial expression analysis will be conducted using DeepFace (Meta Platforms), an open-source facial recognition and analysis framework. Expressions such as disgust, sadness, tiredness, and anger will be detected. Physical gestures, including hand and eye movements, will also be analyzed using pose estimation models.

Temporal Analysis

Detected events, classified vomitus types, and facial expressions will be time-stamped and mapped across the procedure timeline. This will enable the correlation of emesis patterns and patient responses, facilitating a structured understanding of the *vamana karma* process.

Statistical Analysis

Descriptive statistics will summarize participant demographics and baseline characteristics. ML model performance will be reported using accuracy, precision, recall, F_1-score, and area under the receiver operating characteristic curve.

Model validation will use 5-fold cross-validation techniques to ensure robustness. Agreement between AI outputs and physician assessments will be evaluated using the Fleiss κ with standard interpretation thresholds of <0.40 for poor agreement, 0.40 to 0.75 for fair to good agreement, and 0.75 for excellent agreement.

All analyses will be conducted using Python libraries (scikit-learn [Google Summer of Code] and TensorFlow [Google Brain Team]) and the SPSS software (IBM Corp) for statistical analysis.

After training both the YOLOv9 and ResNet models, testing will be conducted in the following structured way. The dataset will be divided into 70% for training, 15% for validation, and 15% for testing. The test set will consist of unseen video samples containing vomiting and nonvomiting events to ensure unbiased evaluation. The YOLOv9 model will be tested on the test dataset to detect vomit regions frame by frame, output bounding boxes, and confidence scores. To test the complete pipeline, the full video will be passed through YOLOv9 followed by ResNet. The pipeline will be evaluated on frame-level accuracy (correct vomit identification across frames), event-level accuracy (correct identification of full vomiting events), and processing time per frame (frames per second) to measure real-time feasibility.

K-fold cross-validation (eg, K=5) will be performed to verify model robustness and reduce overfitting. Additionally, to prevent data leakage, participant-level splitting will be used during model evaluation. All frames from a given participant will be kept within a single fold to ensure that no individual's data appear in both the training and test sets. This ensures realistic generalization performance.

To enhance clinical interpretability, gradient-weighted class activation mapping visualization will be explored on the ResNet classifier outputs to highlight the image regions contributing to vomit type classification. This step supports explainability and builds clinician trust in the AI's decisions.

XSL•FO
RenderX

Outcome Measures

The primary outcome is accuracy (or F_1-score) of vomit event detection and classification at the participant level. Secondary outcomes are inference time per frame and classification consistency across participants.

Ethical Considerations

Ethics approval for the study was obtained from the Institutional Ethics Committee of All India Institute of Ayurveda, New Delhi, India before initiation (325/19.12.2022/PhD-04/2022 [reissued in 2023]). All participants will provide written informed consent before recording after being informed about the study objectives, the voluntary nature of participation, and their right to withdraw at any time without consequences. Videos will be anonymized by removing identifiable facial features and stored in encrypted format (Advanced Encryption Standard–256) on a secure local server with access restricted to authorized researchers.

Cultural Considerations

While the potential of AI integration in Ayurveda is significant, it also raises ethical and cultural challenges. AI-based evaluation of deeply traditional and personalized procedures may be viewed with skepticism by practitioners who value individualized clinical judgment. Hence, AI tools should be positioned as decision support systems rather than replacements for practitioner expertise. Transparent algorithms, patient consent, and data privacy must remain central to any implementation.

Results

The study was initiated in January 2024 and is currently ongoing. Participant enrollment and video data collection began in February 2024, with a target of enrolling 50 patients by March 2025. Procedures for annotating data and training ML models are being carried out. Data analysis is scheduled to commence in May 2025, and the final results are expected to be available by December 2025.

Discussion

Expected Findings

The study seeks to test a novel way to assess TE by using an AI model. The proposed framework will evaluate the outcomes more accurately and ease for the physician by automating the vomiting event. Once validated, this framework could serve as a model for the application of AI in other traditional medicine procedures, thereby contributing to the broader field of integrative and digital health. Previous studies in medical imaging and diagnostic AI have established the capacity of ML algorithms to identify subtle patterns in biomedical data—such as radiographs, endoscopic images, and pathology slides—that often escape human perception. In traditional systems of medicine, similar AI-based models have been used for pulse diagnosis, tongue image interpretation, and facial feature analysis, suggesting the need for deployment [17-21]. However, no prior work has attempted this in the domain of vomitus analysis, leaving a critical research gap that needs to be addressed. The framework can be generalized to other *panchakarma* procedures or Ayurvedic practices involving visual inspection provided that adequate domain-specific annotations are available. Its modular design allows for retraining with site-specific data for broader clinical adoption. The study establishes the foundation for image-based pattern recognition models that could one day assist practitioners in evaluating the *shuddhi lakshana* (signs of purification) objectively.

Strengths of the Study

The study will help enhance diagnostic precision and documentation. Moreover, it enables longitudinal data collection that may reveal new correlations among patient profiles, therapy outcomes, and procedural parameters, which were otherwise difficult to quantify manually.

Limitations

The study's shortcomings include a small sample size, possible observer bias in vomitus labeling, and limited generalizability due to single-center data. Data will be collected in a controlled clinical setting, limiting real-world diversity. TE is a seasonal therapy traditionally administered during specific periods of the year, which naturally limits the availability of eligible participants and, consequently, the sample size for ML model development and validation. Furthermore, the digital interpretability of Ayurvedic features such as *pitta* (bio humor responsible for metabolic activities) and *aushadhi* (medicine) remains challenging. Despite these limitations, the study is an important step toward developing a validated framework for combining AI with traditional assessments. More comprehensive multicentric research is necessary to validate the proposed concept across diverse populations and practice settings.

Future Directions

Future work will focus on validating the framework across larger, more diverse patient populations and exploring its application in real-time clinical decision-making. Although this work focuses on model development and internal validation, future studies will involve clinical validation across multiple *panchakarma* centers to assess model robustness under real-world variations in lighting, camera positioning, and practitioner styles.

Conclusions

This work presents one of the first attempts to apply deep learning for objective analysis of the *vamanakarma* process in Ayurveda. By combining YOLOv9 for vomit detection and ResNet for classification, the framework achieves promising accuracy in automated vomit identification. The results will demonstrate the potential of AI-assisted analysis in traditional medicine, although further clinical validation and expansion across multiple centers will be necessary before deployment in real-world settings.

XSL•FO
RenderX

Acknowledgments

The authors would like to express their sincere gratitude to the All India Institute of Ayurveda for providing the infrastructure and funding support required to carry out this research. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this article. The authors acknowledge Dr Neha Chauhan for her involvement in the initial phase of the study. The authors declare that no generative artificial intelligence or artificial intelligence–assisted technologies were used in the preparation of this manuscript.

Funding

This research was supported by the All India Institute of Ayurveda, New Delhi, India.

Authors' Contributions

Conceptualization: APV, PR
Formal analysis: ARG, SS, RD
Methodology: PR
Project administration: APV, SK
Resources: APV, PR, SK
Visualization: ARG, RD, SK, SS
Writing—original draft: PR
Writing—review and editing: APV, RD, SS

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary table S1.
[DOCX File, 20 KB - resprot_v15i1e79875_app1.docx]

References

1. Acharya VY. Caraka Samhita (Cakapani Ayurveda Dipika): Chaukhambha Orientalia; 2015.
2. Charak Samhita New Edition. URL: https://www.carakasamhitaonline.com/index.php?title=Main_Page [accessed 2026-01-31]
3. Acharya VJT, editor. The Charakasamhita by Agnives'a: Nirnaya Sagar Press; 1941. URL: https://archive.org/details/jTZt_charaka-samhita-sanskrit-of-agnivesha-with-ayurveda-dipika-commentary-ed.-by-vai [accessed 2026-01-31]
4. Yeow JA, Ng PK, Tan KS, Chin TS, Lim WY. Effects of stress, repetition, fatigue and work environment on human error in manufacturing industries. J Appl Sci 2014 Dec 1;14(24):3464-3471. [doi: 10.3923/jas.2014.3464.3471]
5. Khalifa M, Albadawy M. AI in diagnostic imaging: revolutionising accuracy and efficiency. Comput Methods Programs Biomed Update 2024;5:100146. [doi: 10.1016/j.cmpbup.2024.100146]
6. Al-Obeidat F, Hafez W, Gador M, et al. Diagnostic performance of AI-based models versus physicians among patients with hepatocellular carcinoma: a systematic review and meta-analysis. Front Artif Intell 2024;7:1398205. [doi: 10.3389/frai.2024.1398205] [Medline: 39224209]
7. Nagareddy B, Vadlamani R, Venkannagari NR, Jain S, Basheer SN, Murugesan S. Comparison of the artificial intelligence versus traditional radiographic interpretation in detecting periapical periodontitis: a diagnostic accuracy study. J Pharm Bioallied Sci 2024 Dec;16(Suppl 4):S3676-S3678. [doi: 10.4103/jpbs.jpbs_1096_24] [Medline: 39926764]
8. N S, D P. A practical approach towards understanding vamana karma: a review on unique detoxification therapy. Int J Res Ayurveda Pharm 2024;15(5):91-96 [FREE Full text] [doi: 10.7897/2277-4343.155162]
9. Kumari K, Bhatted SK. Evaluation of safety and efficacy of vamana karma in healthy and diseased individuals: a systemic review. J Indian Syst Med 2024;12(3):129-134. [doi: 10.4103/jism.jism_5_23]
10. Supriya C, Thakur N, Kaushik P, et al. A clinical study to assess the role of vasantika vamana on lipid profile and weight in healthy individuals. Int J Ayu Pharm Res 2023:104-112. [doi: 10.47070/ijapr.v11i10.2939]
11. Alsabei AA, Alsubait TM, Alhakami HH. Enhancing crowd safety at Hajj: real-time detection of abnormal behavior using YOLOv9. IEEE Access 2025;13:37748-37761. [doi: 10.1109/ACCESS.2025.3545256]
12. An R, Zhang X, Sun M, Wang G. GC-YOLOv9: innovative smart city traffic monitoring solution. Alex Eng J 2024 Nov;106:277-287. [doi: 10.1016/j.aej.2024.07.004]
13. Azman FI, Ghazali KH, Mohamed Z, Hamid R. Detection of sputum smear cell based on image processing analysis. ARPN J Eng Appl Sci 2015;10:9880-9884 [FREE Full text]
14. Gao Z, Yu X. A phlegm stagnation monitoring based on VDS algorithm. J Healthc Eng 2020;2020:8714070. [doi: 10.1155/2020/8714070] [Medline: 32399167]

XSL•FO
RenderX

15. Lu Z, Qiao Y, Huang X, et al. A deep learning-based system for automatic detection of emesis with high accuracy in suncus murinus. Commun Biol 2025 Feb 10;8(1):209. [doi: 10.1038/s42003-025-07479-0] [Medline: 39930110]
16. Sounderajah V, Ashrafian H, Golub RM, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. BMJ Open 2021 Jun 28;11(6):e047709. [doi: 10.1136/bmjopen-2020-047709] [Medline: 34183345]
17. Kumar PV, Deshpande S, Nagendra HR. Traditional practices and recent advances in nadi pariksha: a comprehensive review. J Ayurveda Integr Med 2019;10(4):308-315. [doi: 10.1016/j.jaim.2017.10.007] [Medline: 30100236]
18. Kurande V, Waagepetersen R, Toft E, Prasad R, Raturi L. Repeatability of pulse diagnosis and body constitution diagnosis in traditional Indian ayurveda medicine. Glob Adv Health Med 2012 Nov;1(5):36-42. [doi: 10.7453/gahmj.2012.1.5.011] [Medline: 27257530]
19. Jiang T, Lu Z, Hu X, et al. Deep learning multi-label tongue image analysis and its application in a population undergoing routine medical checkup. Evid Based Complement Alternat Med 2022;2022:3384209. [doi: 10.1155/2022/3384209] [Medline: 36212950]
20. Liu Q, Li Y, Yang P, et al. A survey of artificial intelligence in tongue image for disease diagnosis and syndrome differentiation. Digit Health 2023;9:20552076231191044. [doi: 10.1177/20552076231191044] [Medline: 37559828]
21. Tiwari P, Kutum R, Sethi T, et al. Recapitulation of ayurveda constitution types by machine learning of phenotypic traits. PLoS ONE 2017;12(10):e0185380. [doi: 10.1371/journal.pone.0185380] [Medline: 28981546]

Abbreviations

AI: artificial intelligence
ML: machine learning
ResNet: residual neural network
TE: therapeutic emesis

Edited by J Sarvestan; submitted 01.Jul.2025; peer-reviewed by SM Shaffi; revised version received 18.Nov.2025; accepted 18.Nov.2025; published 03.Feb.2026.

Please cite as:
Rani P, Kalra S, Singh S, David R, Gupta AR, P V A
Using AI Algorithms and Machine Learning in the Analysis of a Bio-Purification Method (Therapeutic Emesis, Known as "Vamana Karma"): Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e79875
URL: https://www.researchprotocols.org/2026/1/e79875
doi:10.2196/79875

© Pooja Rani, Sumit Kalra, Sachin Singh, Richard David, Ashutosh Ravi Gupta, Anandaraman P V. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 3.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Strengthening Health Care Professionals' Collaborative Responses to Women Experiencing Intimate Partner Violence in Pregnancy: Protocol for an Exploratory Mixed Methods Study

Carly Jones[1], BMid, MClinEd; Belinda Lovell[1], BNurs, MMid, PhD; Tracy Humphrey[2], BM, MSc, PhD; Angela Brown[1], BNurs, BMid, MMid, MBA, PhD

[1]College of Health, Adelaide University, Corner of North Terrace and Frome Road, Adelaide, SA, Australia

[2]College of Nursing and Health Sciences, Flinders University, Adelaide, SA, Australia

Corresponding Author:
Carly Jones, BMid, MClinEd
College of Health, Adelaide University, Corner of North Terrace and Frome Road, Adelaide, SA, Australia

Abstract

Background: Effective collaborative practice among health care professionals is crucial for addressing intimate partner violence (IPV) during pregnancy. Therefore, the development and evaluation of an evidence-based intervention for health care professionals is required to work toward meeting the key priorities of the National Plan to End Violence Against Women and Children 2022 - 2032. The consistency, modality, and effectiveness of IPV-focused education vary, and some midwives lack the confidence to respond effectively to disclosures, often due to limited knowledge, education, and skills. This issue is further amplified in interdisciplinary settings, where a lack of cohesiveness and collaboration can negatively impact the experience for pregnant women seeking or needing support.

Objective: The objective of this study is to design, implement, and evaluate an evidence-informed, simulation-based, interprofessional education intervention to improve health care professionals' collaborative response to IPV disclosures during the perinatal period in the South Australian maternity health care sector.

Methods: This study adopts an exploratory mixed methods research design with 3 key phases. Phase 1 explores the experiences of women with lived experience of IPV in pregnancy (a minimum of 9 participants), and health care professionals providing perinatal care (4 - 8 participants in approximately 4 - 8 focus group offerings). The data will inform the development of the educational program. Phase 3 focuses on the evaluation of the intervention to determine the perceived impact. There will be several offerings of the workshops to capture a minimum of 42 health professional participants.

Results: Qualitative data will be collected first from interviews with women with lived experience of IPV in the perinatal period and focus groups with health professionals who have provided care across that context. Phase 3 comprises quantitative survey data through a single-group quasi-experimental pretest-posttest design to determine the perceived impact of the workshop.

Conclusions: This study aims to provide insights into the effectiveness of educational interventions designed to enhance collaboration within interdisciplinary teams in the context of IPV.

International Registered Report Identifier (IRRID): PRR1-10.2196/86289

(JMIR Res Protoc 2026;15:e86289) doi:10.2196/86289

KEYWORDS

intimate partner violence; domestic violence; collaboration; health professional education; interdisciplinary; simulation-based learning

Introduction

Background

Intimate partner violence (IPV) is a long-standing global phenomenon that transcends cultural, social, and economic boundaries, affecting millions of women and their families [1]. IPV is defined as behaviors within an intimate relationship (current or previous) whereby there is physical, sexual, or psychological harm [2]. Coercive control, however, is the main context in which IPV presents [3]. The difference between individual and societal perspectives on how IPV is defined ultimately shapes how the prevalence, patterns, and health consequences of IPV are understood and measured [4]. In fact, the true extent of IPV may not be accurately reflected in current statistics due to the structural and systemic factors that contribute to underreporting [5,6].

In Australia, current statistics report that IPV has accounted for 57 domestic homicide deaths of women between 2023 and 2024

XSL•FO
RenderX

and has affected approximately 1 in 4 women and girls from the age of 16 [7-10]. Pregnancy does not serve as a protective factor for IPV as it has been shown to begin or persist throughout this time, often acknowledged as a risk factor for escalating to death across this period [11,12]. Exposure to IPV while pregnant is associated with poorer maternal and neonatal health outcomes, such as increased likelihood of suicide, emotional distress, miscarriage, stillbirth, preterm labor, and low birth weight [13]. The profound health implications of IPV during pregnancy extend beyond immediate physical harm, contributing to long-term maternal and neonatal morbidity, increased health care costs, and greater strain on health care systems; this highlights the urgent need for targeted intervention and support [13].

The perception of what constitutes IPV is shaped by individual, relational, and structural factors [1]. The complexity of IPV is further exacerbated by health disciplines having varying perceptions of IPV due to the differences in their key responsibilities, organizational structures, and institutional practices [14]. Limited knowledge and understanding of the cross-disciplinary roles and responsibilities is a well-known factor that hinders collaborative responses to support women [15,16]. While institutional and organizational barriers often influence collaboration, by focusing on improving the understanding of cross-discipline roles and individual responsibilities, it can have a considerable impact on collaboration [14,17].

Several programs and interventions have been implemented and evaluated that focused on addressing the challenges and deficits of health professionals in providing care to women experiencing IPV in pregnancy [18-20]. While these programs have had a positive effect on increasing health professionals' knowledge, the effectiveness could be further enhanced through an interprofessional approach [18-20]. While there are existing programs that have an interprofessional approach, they did not target health professionals actively working with women experiencing IPV; instead, they focused on undergraduate students [19,21].

Workforce training and education were emphasized as a key priority in the National Plan to End Violence Against Women and Children 2022 - 2032 report [1]. While a deficit in available IPV-focused education for health professionals has been acknowledged, the problem is further exacerbated by the continual siloed approach of education [21]. When education is offered in a discipline-specific approach, health professionals remain underprepared to work in a multidisciplinary team to address the complex social issues associated with IPV presentations in pregnancy [21]. To achieve meaningful improvements in maternal and neonatal outcomes, it is essential to implement and evaluate IPV educational programs that focus on collaboration among health care professionals [13].

This study aims to develop a workshop that includes the voices of women with lived experience of IPV across the perinatal period and the health professionals who provide the care. The purpose is to integrate their needs, perspectives, and experiences to capture the current context of perinatal care in South Australia and inform the development of a bespoke workshop. This paper describes the study protocol for the development and evaluation of an interprofessional, IPV-focused simulation workshop for health professionals to develop the necessary skills to collaboratively care for women experiencing IPV and accessing health care in the perinatal period.

Aim

The aim of this study is to design, implement, and evaluate an evidence-informed, simulation-based, interprofessional education intervention to improve health care professionals' collaborative response to IPV disclosures and ongoing management of care during the perinatal period in the South Australian maternity health care sector.

Objectives

The objectives of the study are as follows:

1. To design and develop a simulation workshop for health professionals that captures the experiences of women with lived experience of IPV in pregnancy in addition to addressing the needs of health professionals providing that care.
2. To evaluate the perceived impact of the simulation workshop to determine the changes in the health professional's collaboration and behavioral response to IPV disclosures and ongoing management of care throughout the perinatal context.
3. To evaluate changes in participants' objective knowledge related to recognizing and responding to IPV disclosures during maternity care following a simulation-based interprofessional education intervention, using standardized pre- and postintervention knowledge measures.
4. To evaluate changes in participants' interprofessional responses to simulated IPV disclosure scenarios, using standardized assessment that is scored with predefined rubrics before and after the intervention.

Research Questions

The research questions of the study are as follows:

1. What are the key components of an evidence-informed simulation-based interprofessional education intervention?
2. What interprofessional IPV education and training is accessible for health care providers working in the maternity sector within their workplaces and how do they perceive its relevance and effectiveness?
3. What are the experiences and perceptions of engaging with maternity care among women with lived experience of IPV in South Australia?
4. How does simulation-based interprofessional education intervention influence health care professionals' collaboration, knowledge, and behavioral response to IPV disclosures in the perinatal context?

Methods

Study Design

This proposed study will use an exploratory mixed methods research design to collect and analyze qualitative data. Phase 1 explores the perspectives and experiences of victim-survivors

XSL•FO
RenderX

and health care providers in 2 separate data collection modalities. This data will guide the development of the educational workshop, created during phase 2. Phase 3 will therefore involve the evaluation of the simulation education workshop, using a pretest-posttest design, through the integration of the 16-question Interprofessional Education Collaborative (IPEC) competency self-assessment which has been tested for validity, reliability, and usability [22].

Philosophical Framework

The pragmatic philosophy allows for methodological flexibility in approach and provides a comprehensive insight into the problem [23] (Figure 1). Pragmatic research often integrates a mixed methods approach to first identify a problem and collect multiple forms of data from varying cohorts of participants and then apply the findings in a practical method. This approach encourages researchers to remain open to a variety of assumptions and worldviews, allowing for a holistic and deeper understanding of the phenomenon [23]. The main aim is to develop an experience-driven, action-oriented program [24,25].

Figure 1. Phases of the proposed exploratory mixed methods research design. IPV: intimate partner violence; REDCap: Research Electronic Data Capture.

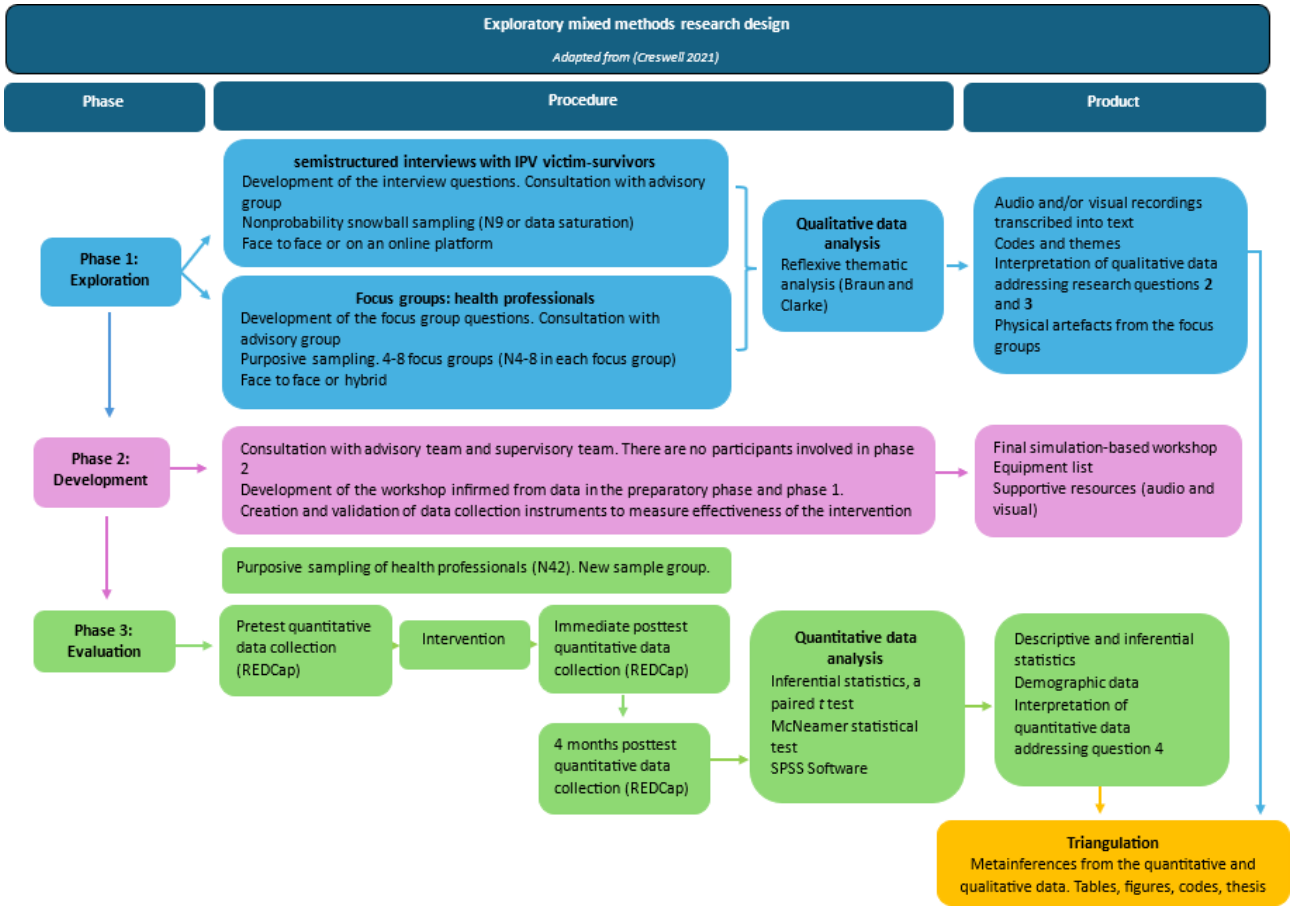

Exploratory Mixed Methods

The project integrates an exploratory mixed methods research design, which involves the collection and integration of qualitative data, followed by quantitative data [23,26,27]. The aim of doing so is to develop a deeper understanding of the phenomena by integrating multiple perspectives through differing data subsets [23,26]. The phase 1 qualitative findings inform the phase 3 quantitative instruments. Dawadi et al [27] note that an exploratory process is important when the research aims to focus on transferability and generalizability. The development of the quantitative instruments in phase 3 of this research is developed following the phase 1 findings, thus consistent with exploratory mixed methods research. The application of mixed methods research in health is an appropriate methodology due to the improved applicability and greater insight generated (depth and breadth) [23,26,27].

Setting

Women with lived experience of IPV will be sought from regional and metropolitan South Australia. The health professionals will be accessed through the same locations; however, they must be practicing in the public or private maternity sector and providing perinatal health care to women and their families. The physical environment, particularly the lived experience interviews, will need to be safe for both the first author and participant and will be taken into consideration. The evaluation of the project will be set at Adelaide University campus in the simulation spaces.

Sampling

The interviews will integrate a nonprobability snowball sampling technique to capture women with lived experience; a useful sampling technique to engage with populations that present

XSL•FO
RenderX

considerable recruitment challenges [28]. This process will be initiated through initial contact with experts and domestic violence organizations.

The focus group participants (phase 1) and evaluation of the workshop (phase 3) will use purposive sampling techniques to ensure a broad representation of disciplines [29].

Participants

Phase 1 has 2 different focuses of participants: women with lived experience of IPV and a minimum of 5 different health profession disciplines. Phase 3 will integrate a new cohort of health professionals.

Phase 1: Lived Experience Participants

Inclusion criteria require people with lived experience of IPV, including any classification such as physical, sexual, or psychological [30,31]. They must have had a pregnancy and accessed maternity care in the past 5 years and be 18 years or older. Participants who are currently pregnant at the time of the interview will be excluded.

Phase 1: Health Professional Participants

For eligibility, participants must be currently practicing in a health profession that is likely to have managed care for pregnant women who have experienced IPV. The health professionals that are most likely to meet this criterion are midwives, social workers, medical staff, perinatal mental health nurse midwives, and Aboriginal Maternal and Infant Care workers (refer to Table 1 for inclusion and exclusion criteria).

Table . Inclusion and exclusion criteria for health professionals.

Discipline	Inclusion	Exclusion
Midwives	Registered with the Ahpra[a] as a midwife and currently working in the clinical environment and has provided care to women who have experienced and disclosed IPV[b] in pregnancy in the past 5 years.	Does not hold registration with Ahpra as a midwife or has conditions as part of their registration and/or has not provided care to women who have experienced and disclosed IPV in pregnancy in the past 5 years.
Social workers	Registered with Ahpra as a social worker and currently practicing clinically and has experience with working with women with lived experience of intimate partner violence in pregnancy in the past 5 years.	No longer practicing clinically or not registered with Ahpra as a social worker and/or does not have experience working with women with lived experience of intimate partner violence in pregnancy in the past 5 years.
Medical staff	Registered with Ahpra as a medical practitioner and undertaken an obstetric rotation in the past 5 years. Additionally, has provided care to women who have experienced and disclosed IPV in pregnancy.	Is not registered with Ahpra as a medical practitioner or has never practiced in obstetrics or has not practiced in the past 5 years. Additionally, if they have not provided care to women who have experienced and disclosed IPV in pregnancy.
Perinatal mental health nurse midwives	Registered with Ahpra as a nurse and/or midwife and currently practicing clinically. Has provided care to women who have experienced and disclosed IPV in pregnancy in the past 5 years. A postgraduate qualification in perinatal mental health has been obtained.	Does not hold registration with Ahpra as a nurse and/or midwife or has conditions as part of their registration and/or has not provided care to women who have experienced and disclosed IPV in the past 5 years. A graduate qualification in perinatal mental health has not been obtained.
AMIC[c] Workers	Have experience working with women with lived experience of intimate partner violence in pregnancy in the past 5 years.	Does not have experience working with women with lived experience of intimate partner violence in pregnancy in the past 5 years.

[a]Ahpra: Australian Health Practitioner Regulation Agency.

[b]IPV: intimate partner violence.

[c]AMIC: Aboriginal Maternal and Infant Care.

Phase 3: Participants

The disciplines will remain the same as phase 1 health professionals; however, participants do not need to have had experience working with women who experienced IPV.

The disciplines included in phase 3 will remain iterative, as the earlier phases may capture additional disciplines who care for women with lived experience of IPV and may benefit from the proposed research.

Recruitment

Phase 1

Self-referral rates among IPV victim-survivors for research participation are often low [32]. Recruitment of lived experience participants will therefore occur through domestic violence organizations and networks with health professionals who are working with women impacted by IPV [33,34]. The chief investigator will email experts expressing the purpose of the research [2]. Those experts who then agree to assist with recruitment will be asked to identify and refer other potential participants who meet the eligibility criteria, thereby extending

XSL•FO
RenderX

recruitment of victim-survivors through their established networks [35].

For inclusion in the focus groups, health professionals will be recruited through the Local Health Networks in South Australia and connections with potential end users of the program [29]. Social media platforms such as Facebook (Meta Platforms, Inc) and LinkedIn (Microsoft Corp) will be used for recruitment, given their cost-effectiveness and ability to reach a broad audience [36]. Furthermore, flyers will be disseminated to management personnel at the Local Health Networks to support recruitment and dissemination.

Phase 3

Phase 3 recruitment of health professionals will be conducted using the same method as phase 1.

Ethical Considerations

Approval

For this research, several ethics applications are required. Phase 1 lived experience interviews received ethics approval from the University of South Australia Human Research Ethics Committee (207340) in November 2025. The research to be undertaken for the focus groups and workshop evaluation (phase 1 and phase 3) will require ethics approval from the Department for Health and Wellbeing Human Research Ethics Committee, and site-specific approval for the Local Health Networks in South Australia [37,38]. Following approval, ratification through the Adelaide University Human Research Ethics Committee (formerly known as the University of South Australia Human Research Ethics Committee) will be sought. Phase 2 is the development of the project and has no engagement with participants, thus not requiring ethics approval.

Informed Consent

All participants across each phase will be provided with a detailed explanation of the participation requirements and expected commitment during the recruitment period [39]. This participant information sheet (PIS) will provide details of the study, requirements of participation, data management, reimbursement, and the consent process. All participants can opt to access the PIS in either a digital or hard copy version. Consent will be specific and limited to the phase of engagement in the proposed project. Participants will also be informed through the PIS that in the event of disclosures indicating a current risk of harm to the participant or child, the researchers have a legal obligation to take appropriate action. Where there is immediate danger to a child and/or participant, the interview will be ceased, and emergency services (000 or 131 444) will be contacted. Where there is a potential risk, however, no immediate danger to a child, the researcher will notify the relevant child protection authority. The consent form requires participants to acknowledge that they have read the PIS as part of the consent process.

Prior to signing the consent form, participants will be asked to complete the self-reflection checklist developed by Lamb et al (refer to Multimedia Appendix 1) [40]. This checklist is designed to ensure that participants engage in self-reflection, are well-informed, and feel adequately prepared to participate in the study [40]. If participants determine they are not prepared to undertake the research after completing the reflective questions, they will be excluded from participation and offered referrals and access to support guided by the study distress protocol [40]. Participants will have the opportunity to ask questions that arise from the lived experience self-reflection checklist to gain clarity and support or can be referred to the support services outlined in the distress protocol (refer to Multimedia Appendix 2) [40].

Gratuity

The inclusion of lived experience participants without payment is often considered exploitative [41,42]; yet, there remains inconsistency in the policy regarding ethical payment in research for lived experience [41]. Interview participants will be offered reimbursement for the costs incurred during participation (a $50 VISA giftcard), in alignment with the recommendations outlined in the National Statement on Ethical Conduct in Human Research (National Health and Medical Research Council [NHMRC]) 2025 policy [43]. Participants in the focus groups are voluntary.

Participant Safety and Withdrawal

All participants will have the right to withdraw at any time during the study, up until the analysis of the data. The health professionals will be informed that after completion of the focus groups, their data will be unable to be removed due to the nature of focus group research.

Women with lived experience of IPV will not be directly recruited through the research team. At no stage will the chief investigator directly reach out to potential participants. Instead, to protect women's safety, snowball sampling techniques through agencies and experts in this field will occur to facilitate recruitment. This indirect way of recruiting aims to prevent coercion. Participants may contact the primary investigator via their preferred method to express interest in the study. By initiating contact with the primary investigator through a participant-initiated method, they are assumed to consider this mode of communication safe. Upon providing the PIS, the primary investigator will verify with each participant that their preferred contact method is suitable and secure, using direct questioning to ensure safe communication.

The process of member checking for women with lived experience of IPV will provide the opportunity to review the full verbatim transcript and redact a statement if they wish. This process will occur via the participants' personal email or another method of communication of their choice. In the email, a summary of the main themes will be included if they choose not to read the entire transcript. Following the interview, they will also receive a follow-up phone call to check in on their well-being. The health professional participants from phase 1 will receive a phone call the following day to check on their well-being. They will also receive an email with the opportunity to clarify the themes and topics from a summary of the focus groups and interview findings. Prior to sending this information, the primary investigator will confirm with the participants that they consent to receive this information through the previously agreed communication method.

XSL•FO
RenderX

Data Management

All data obtained throughout the study will be stored in accordance with the NHMRC management of data guide which was developed in response to the Australian code of conduct of research [43]. All electronic recordings and data will be securely stored on a password-protected laptop (provided by the Adelaide University), OneDrive system, and USB. This system does not allow for external access and sharing, thus it meets the Australian data sovereignty standards [44]. Any physical documents obtained will be securely stored in a locked cabinet within a locked office at the University. Data will be stored in this manner for 7 years, adhering to the minimum 5 years from the first date of publication [43]. After this date, data will be securely destroyed.

Engagement With Participants With Lived Experience of IPV

Overview

Women who have experienced IPV have an increased risk of adverse outcomes when participating in research [43]. The NHMRC recognize people who have experienced social and economic disadvantage, such as trauma and violence, may have potential limitations with their decision-making ability [43]. To mitigate and minimize these risks, consultation with end users, such as social workers from Women's Safety Services SA, will occur prior to the interviews with phase 1 participants to ensure the questions, activities, and discussion is a trauma-informed and healing approach [40]. To ensure victim-survivors are central to the design and empowered by the process, the research will align with best practice principles for coproduced research with victim-survivors and use a framework developed by Lamb et al [40] to support ethical coproduction of research with vulnerable populations of women. These best practice principles align with the NHMRC principles of justice, respect, and beneficence [43].

Distress Protocol

A distress protocol has been developed to ensure the research team is able to respond to participant distress. The flowchart was developed in alignment with recommendations from Whitney and Evered [45], Draucker et al [46], and Labott et al [47] and includes resources and support services (refer to Multimedia Appendix 2). The development of the distress protocol was in collaboration with the advisory group and supervisory team. At the completion of engagement, all participants of each phase will receive a pamphlet with detailed immediate and ongoing psychological support services that are available to access in person and/or online (refer to Multimedia Appendix 3).

Advisory Group

An advisory group, formed through connections with potential end users, will provide guidance throughout the research with an aim to reduce the risk of causing harm for those with lived experience. Six monthly meetings are planned to discuss the proposed research methodologies and intervention. The purpose is to ensure the program supports the needs of women with lived experience in addition to promoting a safe space that protects the well-being of those with lived experience and health professionals throughout the research.

Confidentiality

During data collection, analysis, and reporting, the participants will be labeled using predefined codes (Midwife 1, Social Worker 1, and Woman 1). This process is known as the dominant approach to addressing confidentiality and will be applied during the 3 phases of the process: data collection, data cleaning, and dissemination [48]. The women with lived experience will be informed that generalized, descriptive characteristics will be collected, and information will not be identifiable in the reporting process. This is similar to the health professionals whereby their generalized characteristics, for example, occupation and experience will be included, but not identifiable to maintain confidentiality and anonymity [48]. Identifying information will be stored separately from the transcripts. A deidentified dataset will be analyzed and disseminated, aiming to protect the confidentiality and anonymity of participants [48]. The PIS form provided to focus group participants will have a statement that while confidentiality will be maintained from the researcher's perspective, participants are also obligated to maintain confidentiality across the focus group discussions given the public nature of the forum.

Vicarious Trauma

Due to the nature of the research, there is potential for the researchers involved to experience vicarious trauma. Vicarious trauma can occur from burnout, traumatic stress, and compassion fatigue [49]. The research team will have access to the Employee Assistance Program in preparation for potential discomfort or in response to trauma experienced. All research members will be asked to reflect on their involvement and are aware of the potential challenges of working in this field and agree to embrace differences in perspectives. The supervisory team will have the flexibility to opt into the phases of the research to ensure their comfort, and feedback will be sought iteratively throughout the project to ensure nonmaleficence.

Procedure

Phase 1: Exploration

Overview

The objective of the interviews with women with lived experience of IPV is to understand their perspectives of engaging with maternity care services in South Australia (research question 3). The themes emerging from the interviews will offer valuable insight and inform the development of the workshop, ensuring it reflects and addresses the needs of women with lived experience in South Australia. From there, the importance is shifted to exploring the current interprofessional education and training on IPV received by maternity care providers in South Australia, to then examine their perceptions of its relevance and effectiveness (addressing research question 2). The focus groups aim to cultivate an understanding of each discipline's roles, responsibilities, and values in relation to collaboration with other disciplines when navigating care for women with lived experience or IPV in pregnancy.

XSL•FO
RenderX

Lived Experience Interviews

Semistructured interviews will be used to explore the experiences and perceptions of maternity care among women with lived experience of IPV. The semistructured interviews allow for deeper exploration and the provision of participant autonomy to deviate and explore emerging concepts that may arise throughout the interview and further enhance the understanding of the issues [34,50]. The interview guide will be developed following the data extraction of the scoping review and reviewed by an advisory group and the supervisory team.

Data Saturation: Interviews

It is expected that data saturation will occur after a minimum of 9‐10 interviews [51]. Braun and Clarke [52], however, suggest researchers avoid limiting data collection to a specific number, and therefore in this study, the continuation of interviews will occur until data richness is achieved, and there is notable repetition in the presenting themes and codes.

Health Professional Focus Groups

Focus groups are incorporated due to the known benefits of exploring a group of health professionals' opinions, perspectives, and views to ensure the richness of data [53]. Focus groups are particularly important when exploring new insights and understanding poorly explored topics as communication can be enhanced through a process defined as sharing and comparing [54]. The questions and activities will be structured and planned, commencing with an introduction, transition, and key questions followed by a final activity or question [55]. A range of group activities will be conducted, such as asking the group to arrange a series of statements in a particular order acknowledging that a consensus may not always be reached. The purpose is to observe the insights that arise [56]. The focus groups will be facilitated by 2 members of the research team to reduce bias [55].

Data Saturation: Focus Groups

A total of 4‐8 focus groups is recommended to achieve data saturation, with approximately 4‐8 participants in each group [51,54].

Data Analysis

Data will be analyzed by the primary investigator following the 6 steps to reflexive thematic analysis by Braun and Clarke [52] to determine the patterns and themes from the data. These iterative steps will ensure the data is interpreted and represented meaningfully [52]. The qualitative data will be collected in phase 1 and recorded and transcribed verbatim (by the principal researcher for the victim-survivor interviews and potentially an external third party for the focus groups) for the purpose of identifying, analyzing, and acknowledging patterns that come from the data [52,57]. Analysis will begin with familiarization of the data through transcription and rereading the transcripts to develop a deeper understanding [52,58]. Initial codes will be generated from identified patterns in the data [52]. Reflexive thematic analysis does not require a team of researchers to review codes; instead, the coding process is an "organic" process and does not adhere to specific frameworks [59]. One team member will independently code all data and share the findings with the research team. To improve trustworthiness and confirmability of the data, small portions of the transcripts will be allocated to differing authors, to confer, reflect, provide feedback, and therefore ensure a consensus is reached [60]. Trustworthiness in qualitative research can be strengthened further through reflexive processes [60]. Credibility in this study will be maintained through reflexive journaling to acknowledge personal biases to maintain an objective stance in the collection and data analysis processes [60].

After coding, the initial codes are reevaluated due to the reflexive, iterative nature of the 6 steps described by Braun and Clarke [52] and overarching themes of the data will be sought and determined. The continual reflection and review of the codes and themes will ensure the data represents the whole picture [52]. Once refined, the themes are then clearly defined and titled to articulate the meaning of the data [52]. Finally, the analysis is reported in a manuscript to present to the wider community [52].

The reflexive thematic analysis of focus groups with health care professionals and the interviews with victim-survivors will occur separately. While both qualitative data collection occurs in the same phase, each has a differing purpose and thus the research and analysis will be undertaken and reported separately.

Phase 2: Development of the Advancing Learning in Gender-Based Violence Through Interprofessional Networks Workshop

The simulation-based workshop, referred to as "Advancing Learning in Gender-Based Violence through Interprofessional Networks" (ALIGN), will be developed following the phase 1 qualitative findings, as part of phase 2 of the broader project. The proposed structure of the day will incorporate 4 IPV scenarios, each with a differing focus. The structure will follow best practice standards, outlined by The International Nursing Association for Clinical Simulation and Learning Standards Committee, through the incorporation of structured learner activities, including a prebrief, simulation, and debrief format to ensure transferability and consistency of the program [61].

Simulation education through role play and facilitated discussions is known to improve the participants' knowledge in the clinical environment [58,62]. The proposed simulation-based education program will follow recommendations from the World Health Organization (WHO) through the development of learning objectives that are in alignment with their key recommendations for IPV-focused education for health professionals (clinical skills, resources, behaviors, and values) [58,63]. However, as the program aims to also influence health care professionals' collaboration, the overall program objectives will be combined with the interprofessional competencies as defined by IPEC: values and ethics, roles and responsibilities, communication, teams, and teamwork [64]. This will support the effective delivery of the simulation workshops intended objectives, which are as follows:

1. Apply their own professional role and responsibilities in the clinical scenario, as well as describe the professional role and responsibility of other health professionals when responding to IPV.

XSL•FO
RenderX

2. To improve generalized knowledge of violence against women in the perinatal context and recognizing it as a significant public health concern.
3. Demonstrate ethical and women-centered care when responding, while acknowledging their own personal values and biases to actively collaborate with the interprofessional team.
4. Communicate in a trauma-informed, respectful manner with victim-survivors and interprofessional team members, sharing relevant information and providing feedback to support coordinated, safe, and supportive care.

Phase 3: Evaluation

Overview

The evaluation of the ALIGN workshop will involve a single-group quasi-experimental pretest-posttest design (phase 3). The aim of phase 3 is to evaluate the perceived impact of the simulation-based interprofessional education intervention on health care professionals' collaboration, knowledge, and behavioral responses to IPV disclosures during the perinatal period, addressing research question 4.

Data Collection Tools

A pre- and postquestionnaire will be used to assess the knowledge, attitude, and practice changes of participants. The preworkshop questionnaire will establish a baseline; the postworkshop questionnaire will therefore measure immediate effects on participant knowledge, attitude, practice, and behavior [65]. There will be another follow-up questionnaire at the 4-month postintervention, that is designed to capture practice changes [66].

The questionnaire will be structured to evaluate the Knowledge (multiple choice and binary questions [true or false and yes or no]), Attitudes (Likert scale), and Practices (simulated video scenarios with rubric), referred to as the Knowledge, Attitudes, and Practices (KAP) framework. The KAP framework allows for the exploration of a population to assess knowledge (what is known), attitudes (what is believed), and practices (what is done) [65,67]. Additionally, these questions may identify knowledge gaps and behavioral patterns, aligning with the program's goal of assessing the impact and informing future improvements [67]. The knowledge-based questions will be formulated at the conclusion of the phase 1 qualitative data collection.

To evaluate a change in collaborative practices, the evaluation of the program will integrate the validated 16-question IPEC competency self-assessment tool into the questionnaire [22,68]. This instrument provides a validated means of assessing interprofessional competencies by measuring the participants' self-efficacy, in relation to the IPEC core competency framework [68]. The IPEC competency self-assessment tool is reliable and validated and can therefore be used to assess the development of the participants' collaborative practice skills [68].

Changes in interprofessional practice will also be objectively assessed using predefined scoring rubrics, with standardized criteria applied to minimize interrater variability. This process will involve participants watching a simulated video, aligned with the program's learning outcomes. The participants will be asked to provide a written response that will be graded using a standardized rubric before and after the intervention.

Content Validity

The construction and validation of the knowledge, attitudes, and practice aspect of the questionnaire will be developed in alignment with the instrument development model [69]. The IPEC questions are validated and therefore do not require further validation.

Content validity will be achieved through review of the questionnaire by experts across varying health professionals from the targeted 5 discipline groups [70]. This is not a statistical test; yet, it aims to ensure the questionnaire explores the contextual issues [70]. This process will remain iterative, and the questionnaires will be adapted until each expert agrees on the final set of questions.

The scoring rubrics for the assessed scenarios in the pre- and postquestionnaire, will be developed in alignment with the WHO curriculum guidelines for training health care providers in responding to IPV [63]. Interrater consistency will be supported through assessor collaboration and the use of standardized scoring criteria. Interrater reliability will be measured through the kappa statistic to ensure that the data collected correctly measure the variable prior to the implementation of the project [71]. There will be 2 independent assessors who will grade all data requiring scoring through the rubrics. This process will be refined by piloting the rubric before implementation.

Test-Retest Reliability

The test-retest reliability method will be integrated into the implementation of the project to measure the reliability of the validated questionnaire [70,72]. A randomized subgroup of 20 - 30 health professionals will be sought to volunteer and answer the questionnaire, and 2 weeks later, repeat the same questionnaire to determine the stability and replicability of the questions [73].

Sample Size

A minimum of 42 participants in total, from the 5 or more health professional groups will be recruited (midwives, social workers, medical staff, perinatal mental health nurse midwives, and Aboriginal Maternal and Infant Care workers). There will be several offerings of the simulation workshops conducted, due to the capacity of 10 participants in each session. To avoid bias, the program sessions will continue until all 5 health professional groups have been represented in the data collection.

Data Analysis

Descriptive and inferential statistics will be acquired through the pretest-posttest data collection questionnaires [74]. To determine the significance of perceived change post education, inferential statistics (paired *t* test and McNemar) will be used through the application of SPSS software (IBM SPSS Statistics; version 27).

Standard practice for reporting quantitative data includes detailing the baseline characteristics of participants; therefore,

XSL•FO
RenderX

the protocol will report participant baseline characteristics descriptively (descriptive statistics) [75]. As attrition at the 4-month postintervention follow-up is anticipated, a table of baseline participant characteristics will be presented to allow comparison between participants retained in the study and those lost to follow-up [75]. This process thereby supports the assessment and analysis of potential attrition-related bias. To minimize missing data, questionnaires administered through REDCap (Research Electronic Data Capture; Vanderbilt University) will have the mandatory response settings enabled to reduce the risk of nonresponse bias [76]. Automated reminders regarding outstanding follow-up requirements will be sent to participants through the REDCap platform as a means of retention and tracking procedure recommendations [76].

The KAP survey will be structured to evaluate the knowledge, attitudes, and practices of participants. While the attitudes and practices are subjective self-reported changes, the knowledge-based questions will objectively measure the effect. Data that measure participant knowledge will be analyzed using the McNemar test, which is appropriate for paired categorical (dichotomous) data collected from the same participants pre- and postintervention. The McNemar test focuses on the discordant pairs to determine change for an individual participant over time [77]. Binary outcomes, with one categorical dependent variable (agree or disagree or true or false responses), will be compared to determine the effect, with *P* values calculated to assess the statistical significance [78]. The multiple-choice questions that are correct will receive a score of 1; incorrect and unsure will be scored 0 and will also be assessed at an item level using the McNemar test.

The IPEC tool and attitudes and practice Likert scale responses will be summed to generate a total score and domain-specific scores; higher scores indicate greater self-reported competency. The participant self-reported changes for each domain, and overall scores will be evaluated using a 2-tailed, paired *t* test to determine the mean and SD over time [78]. The Wilcoxon signed-rank test, however, will compare individual ordinal data for each Likert question to compare and determine the statistical significance.

The assessment of the video scenario questions will be scored in ordinal domains based on the rubrics and compared pre- and postinterventions using the Wilcoxon signed-rank test. The total rubric score will be analyzed through a paired *t* test.

Multiple comparisons arise from testing more than one outcome through different methods. Given the exploratory nature of the project, there will be no formal adjustment to multiple comparisons applied; instead, the results will be interpreted cautiously with consideration of the effect sizes and CIs. All data will be displayed in tables.

Triangulation of Findings

This project has both quantitative and qualitative data due to the nature of mixed methods research [23]. Each subset of data from the phases will be analyzed separately to draw meta-inferences in the analysis phase. Following the analysis, these meta-inferences will be triangulated in the interpretation stage of data integration. The triangulation process will adhere to the "triangulation protocol" outlined in O'Cathain et al [79]. The data will be displayed in a convergence coding matrix. Data will be presented in a table to demonstrate where an agreement has been identified and the level of convergence will be presented as either agreement, partial agreement, silence, or dissonance [79]. The triangulation protocol is of particular importance to determine which meta-inference the data are categorized in and therefore allowing for analysis of the strengths and weaknesses of the different methodologies [79].

Results

This research incorporates both qualitative and quantitative approaches through an exploratory research design to guide the development of a simulation-based interprofessional workshop for health professionals. Phase 1 data are qualitative; however, these are drawn from separate cohorts of participants. Interviews with women with lived experience of IPV throughout the perinatal context will be undertaken. Concurrently, focus groups with health professionals who provide care to women in pregnancy will occur. These findings will be analyzed, and recommendations drawn, to integrate into the development of the ALIGN simulation workshop, which comprises phase 2 of the study. There will be no participant involvement in phase 2. Phase 3 is therefore where the quantitative data will occur through a single-group quasi-experimental pretest-posttest design. As the survey will be disseminated online through REDCap, an online web-based platform, the CHERRIES (Checklist for Reporting Results of Internet E-Surveys) checklist will be integrated into the reporting of the process. This will guide the chief investigator to report essential information about the survey design, development, recruitment, and data analysis processes, aiming to improve the project's quality and transparency [80]. The focus of phase 3 is to determine the perceived impact, through statistical analysis, of health professionals' collaboration, knowledge, and behavioral responses to IPV disclosures in the perinatal context.

Discussion

Principal Findings

The paper outlines a proposed mixed methods research project that aims to design, implement, and evaluate an evidence-informed, simulation-based, interprofessional education intervention. Through the integration of input from women with lived experience of IPV in the perinatal context and the health care providers regarding available education and care provision, it is aimed that the workshop will improve health care professionals' collaborative response to disclosures and ongoing management of care in the perinatal period. Phase 1 data are primarily qualitative and inform the workshop development. Phase 2 does not have participants or data collection methods. Quantitative data are collected in phase 3 with the aim of evaluating the perceived impact of the workshop. The integration of qualitative and quantitative data, therefore, seeks to provide a deeper understanding of the issues through the triangulation of data in a final publication.

XSL•FO
RenderX

Strengths and Limitations

It is reasonable to suggest that the use of a mixed methods approach is a strength of the study. In addition to this, while the project is not co-designed, there has been careful consideration of how to ethically integrate the experiences, perspectives, and voices of women impacted by IPV. The project is supported by an advisory group to guide the researchers to approach the research in a trauma-informed manner. While the study has not commenced, it could be argued that the generalizability of the study may be limited through the initial evaluation process as the cohort is situated in South Australia.

Comparison to Prior Work

IPV programs have been evaluated and published, yet there tends to be a focus on the observation of practice and knowledge changes, from a siloed profession [81,82]. This project leverages the insights from previous works and places an emphasis on exploring the perceived impact and changes in the health professional's collaborative response.

There have been several studies undertaken in undergraduate education, with a siloed discipline approach [83-85] or alternatively, incorporating multiple health disciplines [18,19,21]. For example, a simulation workshop integrated into nursing and social work students' curricula explored the impact of the simulation on students' knowledge, perceptions of IPV, and their confidence in acting on disclosures during an interprofessional scenario [19]. The study demonstrated increased confidence in working with people with lived experience of IPV, attributing this finding to the interprofessional education [19]. The data, however, lacked measurable indicators to draw conclusions on the impact on collaboration between disciplines [19]. This highlights the need for an extension of interprofessional IPV programs to evaluate how interprofessional education influences collaboration between disciplines when caring for women with lived experience of IPV across the maternity sector.

Dissemination Plan

The project is part of the chief investigator's doctoral studies and therefore dissemination will occur through a thesis. Planned dissemination of the project also includes the publication of 5 manuscripts in academic journals. At the conclusion of the project, it is aimed that the workshop could potentially be implemented across broader settings throughout Australia.

Conclusion

The development of interprofessional IPV education is integral to improving health care professionals' behavior and responses. The outcomes of this project will provide valuable insights into the current context of health care for victim-survivors accessing health care in South Australia, in contrast to the perspective of health professionals. The findings then develop a simulation workshop. We anticipate that the insights from the evaluation of the simulation workshop will improve the collaborative response of health professionals, to therefore improve care received by victim-survivors.

Acknowledgments

The authors would like to thank and acknowledge the contributions of Ms Kelly Barrett in providing valuable guidance and advice in the design of the study from an industry perspective. The authors also acknowledge Paul Cooper, Kate Toone, and Matthew Aquilla, members of the study's advisory group, for their guidance and support. Drawing on their extensive experience with people with lived experience and best practices, they helped guide the structure and development of the research to meet participants' needs. No generative AI tools were used at any stage in the preparation of this manuscript. The authors take full responsibility for the accuracy, originality, and integrity of all content presented.

Funding

No financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Authors' Contributions

CJ conceptualized and designed the study, drafted the manuscript, and led all aspects of writing and revision. BL and TH provided supervision, significant conceptual input, and contributed to review and editing. AB provided primary supervision, contributed significant conceptual input, and assisted with review and editing of the manuscript. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
An exemplar of the self-reflective questionnaire that all participants are asked to complete prior to engaging with the research.
[DOCX File, 26 KB - resprot_v15i1e86289_app1.docx]

Multimedia Appendix 2

XSL•FO
RenderX

Distress protocol: this is a guideline for management of participants if emotional distress arises throughout the interview, focus groups, or survey.
[PDF File, 193 KB - resprot_v15i1e86289_app2.pdf]

Multimedia Appendix 3
Pamphlet: a “thank you” for engagement pamphlet provided to all participants at the conclusion of the workshop. This pamphlet provides an overview of immediate and ongoing psychological support services.
[PDF File, 192 KB - resprot_v15i1e86289_app3.pdf]

References

1. National plan to end violence against women and children 2022–2032. Department of Social Services. 2022. URL: https://www.dss.gov.au/national-plan-end-gender-based-violence [accessed 2025-02-05]
2. Intimate partner violence. World Health Organization. 2022. URL: https://apps.who.int/violence-info/intimate-partner-violence [accessed 2025-12-17]
3. Defining and responding to coercive control: policy brief. Australia’s National Research Organisation for Women’s Safety. 2021. URL: https://anrows-2019.s3.ap-southeast-2.amazonaws.com/wp-content/uploads/2021/07/30132421/Coercive-Control-Policy-Brief-ANROWS-Insights-1.1.pdf [accessed 2025-11-12]
4. Nicolaidis C, Paranjape A. In: Mitchell C, Anglin D, editors. Intimate Partner Violence: A Health-Based Perspective: Oxford Academic; 2009. [doi: 10.1093/oso/9780195179323.001.0001]
5. Muzingili T, Muzondo E, Chidyausiku W, Kahomwe L. The structural and systemic factors influencing underreporting of intimate partner violence among rural married women: a qualitative research. J Asian Afr Stud 2025 Dec;60(8):4931-4949. [doi: 10.1177/00219096241270676]
6. Loxton D, Powers J, Townsend N, Harris ML, Forder P. Longitudinal inconsistency in responses to survey items that ask women about intimate partner violence. BMC Med Res Methodol 2019 Oct 29;19(1):201. [doi: 10.1186/s12874-019-0835-4] [Medline: 31664906]
7. Domestic homicide. Australian Institute of Health and Welfare. 2025 Feb 28. URL: https://www.aihw.gov.au/family-domestic-and-sexual-violence/responses-and-outcomes/domestic-homicide [accessed 2025-11-10]
8. Intimate partner violence. Australian Institute of Health and Welfare. 2024. URL: https://www.aihw.gov.au/family-domestic-and-sexual-violence/types-of-violence/intimate-partner-violence [accessed 2025-11-10]
9. Change the story. : Our Watch; 2021 URL: https://assets.ourwatch.org.au/assets/Key-frameworks/Change-the-story-Our-Watch-AA.pdf [accessed 2026-03-17]
10. Serpell B, Sullivan T, Doherty L. Homicide in Australia 2019–20. Australian Institute of Criminology. 2022. URL: https://www.aic.gov.au/sites/default/files/2022-03/sr39_homicide_in_australia_2019-20.pdf [accessed 2025-11-15]
11. Chen XY, Lo CKM, Chen Q, et al. Intimate partner violence against women before, during, and after pregnancy: a meta-analysis. Trauma Violence Abuse 2024 Oct;25(4):2768-2780. [doi: 10.1177/15248380241226631]
12. Toivonen C, Backhouse C. National risk assessment principles for domestic and family violence: companion resource. Australia’s National Research Organisation for Women Safety (ANROWS). 2018. URL: https://anrows-2019.s3.ap-southeast-2.amazonaws.com/wp-content/uploads/2018/07/19030421/ANROWS_NRAP_National-Risk-Assessment-Principles.1.pdf [accessed 2025-11-12]
13. Alhusen JL, Ray E, Sharps P, Bullock L. Intimate partner violence during pregnancy: maternal and neonatal outcomes. J Womens Health (Larchmt) 2015 Jan;24(1):100-106. [doi: 10.1089/jwh.2014.4872] [Medline: 25265285]
14. Notko M, Husso M, Piippo S, Fagerlund M, Houtsonen J. Intervening in domestic violence: interprofessional collaboration among social and health care professionals and the police. J Interprof Care 2022;36(1):15-23. [doi: 10.1080/13561820.2021.1876645] [Medline: 33657958]
15. Lalande C, Gauthier S, Damant D, Lessard G, Dubé M. Professional representations of collaboration in the response to intimate partner violence. J Fam Viol 2024 Aug;39(6):1027-1038. [doi: 10.1007/s10896-023-00504-y]
16. Larsson AKL, Olsson H, Strand SJM. Challenges and opportunities in collaborative approaches to responding to intimate partner violence: insights from social workers and police in Sweden. J Soc Work Pract 2024 Apr 2;38(2):143-157. [doi: 10.1080/02650533.2024.2327038]
17. Baird KM, Saito AS, Eustace J, Creedy DK. Effectiveness of training to promote routine enquiry for domestic violence by midwives and nurses: a pre-post evaluation study. Women Birth 2018 Aug;31(4):285-291. [doi: 10.1016/j.wombi.2017.10.014] [Medline: 29102526]
18. Baird KM, Saito AS, Eustace J, Creedy DK. An exploration of Australian midwives’ knowledge of intimate partner violence against women during pregnancy. Women Birth 2015 Sep;28(3):215-220. [doi: 10.1016/j.wombi.2015.01.009] [Medline: 25684254]
19. Childress S, Mammah R, Schrag RV, et al. Preparing to intervene in intimate partner violence: an interprofessional safety planning and assessment simulation. J Soc Work Educ 2024;60(3):448-462. [doi: 10.1080/10437797.2023.2298392] [Medline: 39749145]

XSL•FO
RenderX

20. Childress S, Highfill C, Mammah R, et al. Interprofessional education for domestic violence safety planning in nursing and social work: a scoping review. J Soc Work Educ 2024;60(4):551-564. [doi: 10.1080/10437797.2023.2298389] [Medline: 39749041]
21. Keedle H, Stulz V, Conti J, et al. Psychosocial interprofessional perinatal education: Design and evaluation of an interprofessional learning experience to improve students' collaboration skills in perinatal mental health. Women Birth 2023 Jul;36(4):e379-e387. [doi: 10.1016/j.wombi.2023.01.001] [Medline: 36697285]
22. Lockeman KS, Dow AW, DiazGranados D, et al. Refinement of the IPEC competency self-assessment survey: results from a multi-institutional study. J Interprof Care 2016 Nov;30(6):726-731. [doi: 10.1080/13561820.2016.1220928] [Medline: 27797633]
23. Creswell JW. A Concise Introduction to Mixed Methods Research, 2nd edition: Sage Publications; 2021.
24. Biesta G. In: Tashakkori A, Teddlie C, editors. Pragmatism and the Philosophical Foundations of Mixed Methods Research, 3rd edition: Sage Publications; 2010. [doi: 10.4135/9781506335193.n4]
25. Hothersall SJ. Epistemology and social work: enhancing the integration of theory, practice and research through philosophical pragmatism. Eur J Soc Work 2019 Sep 3;22(5):860-870. [doi: 10.1080/13691457.2018.1499613]
26. Taherdoost H. What are different research approaches? Comprehensive review of qualitative, quantitative, and mixed method research, their applications, types, and limitations. J Manag Sci Eng Res 2022;5(1):53-63. [doi: 10.30564/jmser.v5i1.4538]
27. Dawadi S, Shrestha S, Giri RA. Mixed-methods research: a discussion on its types, challenges, and criticisms. J Pract Stud Educ 2021;2(2):25-36. [doi: 10.46809/jpse.v2i2.20]
28. Sadler GR, Lee HC, Lim RSH, Fullerton J. Recruitment of hard-to-reach population subgroups via adaptations of the snowball sampling strategy. Nurs Health Sci 2010 Sep 1;12(3):369-374. [doi: 10.1111/j.1442-2018.2010.00541.x] [Medline: 20727089]
29. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544. [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
30. Román-Gálvez RM, Martín-Peláez S, Fernández-Félix BM, Zamora J, Khan KS, Bueno-Cavanillas A. Worldwide prevalence of intimate partner violence in pregnancy. A systematic review and meta-analysis. Front Public Health 2021;9:738459. [doi: 10.3389/fpubh.2021.738459] [Medline: 34527656]
31. Korkmaz S, Överlien C, Lagerlöf H. Youth intimate partner violence: prevalence, characteristics, associated factors and arenas of violence. Nord Soc Work Res 2022 Oct 2;12(4):536-551. [doi: 10.1080/2156857X.2020.1848908]
32. Dichter ME, Sorrentino AE, Haywood TN, et al. Women's participation in research on intimate partner violence: findings on recruitment, retention, and participants' experiences. Womens Health Issues 2019;29(5):440-446. [doi: 10.1016/j.whi.2019.03.007] [Medline: 31072755]
33. Sabri B, Lee J, Saha J. Conducting intervention research with immigrant survivors of intimate partner violence: barriers and facilitators of recruitment and retention. J Interpers Violence 2022 Oct;37(19-20):NP18060-NP18084. [doi: 10.1177/08862605211035866] [Medline: 34344214]
34. Adeoye‐Olatunde OA, Olenik NL. Research and scholarly methods: semi‐structured interviews. J Am Coll Clin Pharm 2021 Oct;4(10):1358-1367 [FREE Full text] [doi: 10.1002/jac5.1441]
35. Parker C, Scott S, Geddes A. Snowball Sampling: SAGE Research Methods Foundations; 2019. [doi: 10.4135/9781526421036831710]
36. Darko EM, Kleib M, Olson J. Social media use for research participant recruitment: integrative literature review. J Med Internet Res 2022 Aug 4;24(8):e38015. [doi: 10.2196/38015] [Medline: 35925655]
37. SA Health human research ethics committees. SA Health. 2024. URL: https://www.sahealth.sa.gov.au/wps/wcm/connect/public+content/sa+health+internet/about+us/health+and+medical+research/ethics+and+governance+in+south+australia/sa+health+research+ethics/sa+health+research+ethics [accessed 2025-08-20]
38. Human research ethics application. University of South Australia. URL: https://i.unisa.edu.au/staff/research/research-ethics/human-research-ethics/getting-started/ethics-application-info [accessed 2025-07-20]
39. Guckian J, Thampy H. Gaining ethics approval for health professions education research. Clin Exp Dermatol 2022 Dec;47(12):2081-2084. [doi: 10.1111/ced.15185] [Medline: 35315117]
40. Lamb K, Dembele L, Nina F, Hegart K. An Australian Framework for the Ethical Co-Production of Research and Evaluation with Victim Survivors of Domestic, Family, and Sexual Violence: University of Melbourne; 2023. URL: https://www.researchgate.net/publication/373449435_An_Australian_Framework_for_the_ethical_coproduction_of_research_and_evaluation_with_victim_survivors_of_domestic_family_and_sexual_violence [accessed 2026-02-26]
41. McGregor K, Taylor B, Oakley L. Power, participation, payment and platform: ethical and methodological issues in recruitment in qualitative domestic abuse research. J Fam Viol 2023 Aug;38(6):1029-1041. [doi: 10.1007/s10896-023-00590-y]

XSL•FO
RenderX

42. Taylor J, Clarence B. Beneficial but triggering: experiences and support of survivor speakers in the UK. : Victim Focus Publications; 2021 URL: https://irp.cdn-website.com/4700d0ac/files/uploaded/Beneficial+but+Triggering+Report+2021+FINAL.pdf [accessed 2025-04-30]
43. National statement on ethical conduct in human research 2025. National Health and Medical Research Council, Australian Government. 2025. URL: https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2025#block-views-block-file-attachments-content-block-1 [accessed 2025-04-20]
44. Research data storage. University of South Australia Research. URL: https://i.unisa.edu.au/askresearch/data-management/data-storage [accessed 2025-02-24]
45. Whitney C, Evered JA. The qualitative research distress protocol: a participant-centered tool for navigating distress during data collection. Int J Qual Methods 2022 Apr;21:21. [doi: 10.1177/16094069221110317]
46. Draucker CB, Martsolf DS, Poole C. Developing distress protocols for research on sensitive topics. Arch Psychiatr Nurs 2009 Oct;23(5):343-350. [doi: 10.1016/j.apnu.2008.10.008] [Medline: 19766925]
47. Labott SM, Johnson TP, Feeny NC, Fendrich M. Evaluating and addressing emotional risks in survey research. Surv Pract 2016 Feb 1;9(1):1-9. [doi: 10.29115/SP-2016-0006]
48. Kaiser K. Protecting respondent confidentiality in qualitative research. Qual Health Res 2009 Nov;19(11):1632-1641. [doi: 10.1177/1049732309350879] [Medline: 19843971]
49. Dunkley J, Whelan TA. Vicarious traumatisation: current status and future directions. Br J Guid Counc 2006 Feb;34(1):107-116. [doi: 10.1080/03069880500483166]
50. de la Croix A, Barrett A, Stenfors T. How to…do research interviews in different ways. Clin Teach 2018 Dec;15(6):451-456. [doi: 10.1111/tct.12953] [Medline: 30345639]
51. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: How many interviews are enough. Qual Health Res 2017 Mar;27(4):591-608. [doi: 10.1177/1049732316665344] [Medline: 27670770]
52. Braun V, Clarke V. Thematic Analysis a Practical Guide: Sage Publications Ltd; 2022. [doi: 10.1007/978-3-319-69909-7_3470-2]
53. Franco S, Augusto A. Health professionals' intervention in the context of domestic violence against women: exploring perceptions and experiences of providing healthcare. Health Sociol Rev 2024 Jul;33(2):223-240. [doi: 10.1080/14461242.2024.2354801] [Medline: 38875323]
54. Morgan DL. The Focus Group Guidebook: Sage Publications; 1998. [doi: 10.4135/9781483328164]
55. Krueger RA. Developing Questions for Focus Groups: Sage Publications; 1998. [doi: 10.4135/9781483328126]
56. Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995 Jul 29;311(7000):299-302. [doi: 10.1136/bmj.311.7000.299] [Medline: 7633241]
57. Clarke V, Braun V. Thematic analysis. J Posit Psychol 2017 May 4;12(3):297-298. [doi: 10.1080/17439760.2016.1262613]
58. The Royal Australian College of General Practitioners. RACGP Abuse and Violence: Working with Our Patients in General Practice (the White Book), 5th edition: The Royal Australian College of General Practitioners; 2021. URL: https://www.racgp.org.au/clinical-resources/clinical-guidelines/key-racgp-guidelines/view-all-racgp-guidelines/abuse-and-violence/system-issues/intimate-partner-abuse-education-and-training-for [accessed 2025-10-17]
59. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021 Jul 3;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
60. Ahmed SK. The pillars of trustworthiness in qualitative research. J Med Surg Public Health 2024 Apr;2:100051. [doi: 10.1016/j.glmedi.2024.100051]
61. Watts PI, McDermott DS, Alinier G, et al. Healthcare simulation standards of best practiceTM simulation design. Clin Simul Nurs 2021 Sep;58:14-21. [doi: 10.1016/j.ecns.2021.08.009]
62. Sammut D, Kuruppu J, Hegarty K, Bradbury-Jones C. Which violence against women educational strategies are effective for prequalifying health-care students?: a systematic review. Trauma Violence Abuse 2021 Apr;22(2):339-358. [doi: 10.1177/1524838019843198]
63. World Health Organization. Caring for Women Subjected to Violence: A WHO Curriculum for Training Health-Care Providers: World Health Organization; 2021. URL: https://www.who.int/publications/i/item/9789240039803 [accessed 2026-02-26]
64. IPEC Core Competencies Revision: Version 3 (2021-2023). Interprofessional education collaborative. 2023. URL: https://www.ipecollaborative.org/2021-2023-core-competencies-revision [accessed 2025-02-22]
65. Andrade C, Menon V, Ameen S, Kumar Praharaj S. Designing and conducting knowledge, attitude, and practice surveys in psychiatry: practical guidance. Indian J Psychol Med 2020 Sep;42(5):478-481. [doi: 10.1177/0253717620946111] [Medline: 33414597]
66. Stratton SJ. Quasi-experimental design (pre-test and post-test studies) in prehospital and disaster research. Prehosp Disaster Med 2019 Dec;34(6):573-574. [doi: 10.1017/S1049023X19005053] [Medline: 31767051]
67. Advocacy, communication and social mobilization for TB control: a guide to developing knowledge, attitude and practice surveys. World Health Organization. 2008. URL: https://www.who.int/publications/i/item/9789241596176 [accessed 2025-12-15]

XSL•FO
RenderX

68. Dow A, DiazGranados D, Mazmanian P, Retchin S. IPEC competency self-assessment tool. National Center for Interprofessional Practice and Education. 2016. URL: https://nexusipe.org/advancing/assessment-evaluation/ipec-competency-self-assessment-tool [accessed 2025-07-20]
69. Richardson-Tench M, Taylor B. Research in Nursing: Evidence for Best Practice with Online Study Tools, 5th edition: Cengage; 2014.
70. Steen M, Roberts T. The Handbook of Midwifery Research: Wiley-Blackwell; 2011.
71. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med 2012;1(3):276-282. [doi: 10.11613/BM.2012.031]
72. DePoy E, Gitlin LN. Introduction to Research: Understanding and Applying Multiple Strategies, 5th edition: Elsevier; 2016.
73. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. Niger Postgrad Med J 2015;22(4):195-201. [doi: 10.4103/1117-1936.173959] [Medline: 26776330]
74. Descriptive and inferential statistics. Laerd Statistics. 2018. URL: https://statistics.laerd.com/statistical-guides/descriptive-inferential-statistics.php [accessed 2025-02-20]
75. Dumville JC, Torgerson DJ, Hewitt CE. Reporting attrition in randomised controlled trials. BMJ 2006 Apr 22;332(7547):969-971. [doi: 10.1136/bmj.332.7547.969] [Medline: 16627519]
76. Ribisl KM, Walton MA, Mowbray CT, Luke DA, Davidson WS II, Bootsmiller BJ. Minimizing participant attrition in panel studies through the use of effective retention and tracking strategies: review and recommendations. Eval Program Plann 1996 Feb;19(1):1-25. [doi: 10.1016/0149-7189(95)00037-2]
77. Laerd statistics. McNemar's test using SPSS Statistics.: Laerd; 2018. URL: https://statistics.laerd.com/spss-tutorials/mcnemars-test-using-spss-statistics.php [accessed 2025-02-20]
78. Kishore K. Statistics corner: paired groups. J Postgrad Med Educ Res 2023 May 31;57(2):90-93. [doi: 10.5005/jp-journals-10028-1626]
79. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. BMJ 2010 Sep 17;341:c4587. [doi: 10.1136/bmj.c4587] [Medline: 20851841]
80. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
81. Bamigboye TO, Irinoye OO, Ayandiran EO, Olowokere AE, Adedeji OA, Esan DT. Influence of educational intervention on nurse-midwives' knowledge and management practices of Intimate Partner Violence (IPV) in healthcare Facilities in Ekiti-State, Nigeria. Enferm Clin (Engl Ed) 2023;33(5):316-326. [doi: 10.1016/j.enfcle.2023.07.002] [Medline: 37806710]
82. Baird K, Creedy DK, Saito AS, Eustace J. Longitudinal evaluation of a training program to promote routine antenatal enquiry for domestic violence by midwives. Women Birth 2018 Oct;31(5):398-406. [doi: 10.1016/j.wombi.2018.01.004] [Medline: 29352725]
83. Eckerström J, Mårtensson S, Larsson M, et al. Virtual patient simulation in an interactive educational module on intimate partner violence: nursing students' experiences-a mixed-methods study. Front Digit Health 2025;7:1516379. [doi: 10.3389/fdgth.2025.1516379] [Medline: 40290871]
84. Wagoner K, Dempsey A, Dunn F, Chardukian M. A flipped classroom case to introduce OB/GYN clerkship students to contraception, postpartum care, and intimate partner violence screening. MedEdPORTAL 2025;21:11505. [doi: 10.15766/mep_2374-8265.11505] [Medline: 40206869]
85. Selwyn CN, Moore LP, Thomas R, Mosley BD. Virtual simulation-based educational training enhances nursing students' preparedness to conduct perinatal intimate partner violence screenings. Clin Simul Nurs 2024 Dec;97:101635. [doi: 10.1016/j.ecns.2024.101635]

Abbreviations

ALIGN: Advancing Learning in Gender-Based Violence through Interprofessional Networks
CHERRIES: Checklist for Reporting Results of Internet E-Surveys
IPEC: Interprofessional Education Collaborative
IPV: intimate partner violence
KAP: Knowledge, Attitudes, and Practices
NHMRC: National Health and Medical Research Council
PIS: Participant Information Sheet
REDCap: Research Electronic Data Capture
WHO: World Health Organization

XSL•FO
RenderX

Edited by J Sarvestan; submitted 28.Oct.2025; peer-reviewed by CJ Filmalter, K Baggett; revised version received 25.Jan.2026; accepted 29.Jan.2026; published 24.Mar.2026.

Please cite as:
Jones C, Lovell B, Humphrey T, Brown A
Strengthening Health Care Professionals' Collaborative Responses to Women Experiencing Intimate Partner Violence in Pregnancy: Protocol for an Exploratory Mixed Methods Study
JMIR Res Protoc 2026;15:e86289
URL: https://www.researchprotocols.org/2026/1/e86289
doi:10.2196/86289

© Carly Jones, Belinda Lovell, Tracy Humphrey, Angela Brown. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Etiologic Types and Complications of Diabetes Mellitus in Newly Diagnosed Patients at Health Institutions in Bulawayo, Zimbabwe: Protocol for a Cross-Sectional and Prospective Observational Study

Rudo Gwini[1*], MSc, MMed, MD; Elopy Sibanda[1], MD, PhD; Desmond Mwembe[2*], BSc (Hons), MSc, PhD; Fraser James Pirie[3*], MBChB; Ayesha A Motala[3*], MBChB, MD

[1]Department of Medicine, Faculty of Medicine, National University of Science and Technology Zimbabwe, Vera Road, Mzilikazi, P.O. Box AC939 Ascot, Bulawayo, Zimbabwe

[2]Department of Statistics and Operations Research, Faculty of Applied Sciences, National University of Science and Technology, Bulawayo, Zimbabwe

[3]Department of Diabetes and Endocrinology, School of Clinical Medicine Nelson R Mandela, University of KwaZulu-Natal, Durban, South Africa

[*]these authors contributed equally

Corresponding Author:
Rudo Gwini, MSc, MMed, MD
Department of Medicine, Faculty of Medicine, National University of Science and Technology Zimbabwe, Vera Road, Mzilikazi, P.O. Box AC939 Ascot, Bulawayo, Zimbabwe

Abstract

Background: The prevalence of diabetes mellitus is increasing in sub-Saharan Africa. Data on the prevalence of diabetes in Zimbabwe are scarce, and the etiologic types of diabetes are not well characterized. Classification of diabetes in Zimbabwe relies on clinical criteria at the time of diagnosis, and more detailed phenotype data are lacking. Furthermore, the prevalence of complications at diagnosis of diabetes and the incidence of complications during follow-up are not well documented in Zimbabwe.

Objective: The primary aim of this study is to characterize the etiological types of diabetes in adult and adolescent patients with newly diagnosed diabetes in Bulawayo, Zimbabwe. The secondary objectives are to determine the prevalence of chronic complications of diabetes among adult and adolescent patients with newly diagnosed diabetes and to determine the incidence and risk factors for the development of diabetes complications after a 2-year follow-up in patients with type 2 diabetes who are free of complications at baseline.

Methods: This is a cross-sectional and prospective observational study. The cross-sectional (phase 1) study was conducted in patients presenting for the first time to the diabetes service at 2 referral hospitals and 2 diabetes clinics in Bulawayo, Zimbabwe. Data collected from consenting participants included demographic data, social and medical history, and clinical examination. Laboratory tests included serum urea, creatinine and electrolytes, liver function tests, lipids, plasma glucose, glycated hemoglobin, serum C-peptide, spot urine (dipstick, albumin, and creatinine), and β-cell antibodies (antiglutamic acid decarboxylase, anti-islet antigen, anti-insulin antibodies, anti-islet cell antibodies, and antizinc transporter 8 antibodies). All patients had retinal photography, a 12-lead electrocardiograph, and measurement of carotid intima-media thickness and arterial stiffness. Determination of the incidence of diabetes complications will be conducted through a 2-year follow-up (encompassing 6-mo review) of a subgroup of patients with type 2 diabetes and no diabetes complications at the time of enrollment. At each 6-month visit, in addition to all variables collected at baseline, data on diabetes management and drug therapy compliance will be obtained.

Results: Data collection commenced in October 2021, with 323 participants recruited. Data analysis for phase 1 is ongoing. The study will be completed in October 2026. The results will describe the spectrum of diabetes and complications found at diagnosis (phase 1) and the incidence and risk factors associated with the development of complications of diabetes (phase 2).

Conclusions: The study will provide data on etiologic types of diabetes in patients presenting to health facilities in urban centers in Bulawayo, Zimbabwe. In addition, data on diabetes complications at the time of diagnosis as well as incident complications over 2 years of follow-up will be compared with data from other studies. The data will be used to inform management strategies for patients diagnosed with diabetes in Zimbabwe.

(JMIR Res Protoc 2026;15:e74186) doi:10.2196/74186

KEYWORDS

type 1 diabetes; type 2 diabetes; diabetes complications; etiologic type; autoantibodies; C-peptide

XSL•FO
RenderX

Introduction

Diabetes mellitus (diabetes) is a metabolic disorder that is characterized by hyperglycemia, which develops as a result of defective insulin secretion, insulin action, or both [1]. The burden of diabetes is increasing worldwide. In 2017, 15 million people were estimated to have diabetes in Africa, and this is projected to increase to 40.7 million by 2045 [2]. For sub-Saharan Africa (SSA) in 2017, there were an estimated 14.2 million people with diabetes, and this number is expected to rise to 43.2 million by 2040 [3]. As in other countries in SSA, the prevalence of diabetes in Zimbabwe is reported to have increased from 0.4% in the 1980s to 5.7% in 2014 [4]. Although the burden of diabetes is increasing worldwide, there is a dearth of information on the types of diabetes in SSA in general and in Zimbabwe, in particular, where classification is based mainly on age at presentation. In addition, there are no studies on the incidence of diabetes complications in Zimbabwe.

Classification of diabetes is evolving as more information about the disease is reported. Apart from the age at onset of diabetes, additional tests, such as C-peptide measurement, β-cell autoantibody testing, and genetic testing, may aid in the classification of diabetes and ultimately improve the care of patients through selection of appropriate therapy. The World Health Organization (WHO) and American Diabetes Association classify diabetes mellitus as type 1 (T1D), type 2 (T2D), hyperglycemia in pregnancy, unclassified diabetes, hybrid forms of diabetes, and other types of diabetes [1,5-8]. C-peptide is a useful biochemical measure for assessing pancreatic β-cell function, and it may be used to both classify the type of diabetes and to predict insulin requirement [9,10]. In a Scandinavian study, serum C-peptide >1.51 nmol/L was indicative of insulin resistance [11]. In a few studies reported from countries in SSA, including Ethiopia, Uganda, Sudan, and Côte d'Ivoire, patients with T2D were found to have high levels of circulating C-peptide. Measurement of C-peptide in these studies improved both the classification and treatment of patients [12]. Another useful investigation assisting in the classification of diabetes is β-cell autoantibody testing. The presence of 1 or more β-cell autoantibodies (glutamic acid decarboxylase, zinc transporter 8, insulinoma antigen, and islet cell antibody [ICA]) is suggestive of T1D [5,7]. Measurement of C-peptide and β-cell autoantibodies is expensive and not done routinely in Zimbabwe. Classification of diabetes in Zimbabwe is based on clinical criteria, including mode of presentation, age at diagnosis, and phenotype. Some adult patients may present with features of T1D, and similarly, young patients may present with clinical features of T2D. Misclassification of diabetes may result in the selection of inappropriate therapy. Therefore, this study was undertaken to assess the spectrum of diabetes mellitus and its complications at the time of diagnosis and to investigate the incidence and risk factors associated with the development of complications in a Zimbabwean population in Bulawayo. Establishing the spectrum of diabetes will hopefully assist in formulating policies on diabetes care in Zimbabwe.

The acute and chronic complications of diabetes are well described. Diabetic ketoacidosis (DKA) may be the mode of presentation of undiagnosed diabetes mellitus, and this acute metabolic complication is associated with a high mortality. Worldwide mortality due to DKA is high in T1D (20%), and in older people, it is responsible for deaths in 5%. In SSA, mortality due to DKA is estimated to be between 10% and 30% [8]. Furthermore, ketosis-prone type 2 diabetes (KPD) is reported in several African populations and may lead to confusion regarding the etiologic type of diabetes. KPD is characterized by the onset of diabetes with ketoacidosis, male predominance, low prevalence of β-cell antibodies, and high rates of either remission or insulin independence after initial stabilization [13]. Data on acute diabetes complications, in particular DKA, at the time of diagnosis in patients in Zimbabwe are not available.

Prolonged hyperglycemia leads to the development of chronic micro- (retinopathy, neuropathy, and nephropathy) and macro- (atherosclerosis) vascular complications of diabetes [14,15]. The prevalence of cardiovascular complications was reported to be 17.5%, and the presence of 1 complication was 24% in patients with newly diagnosed diabetes in a Swedish study using data from the Swedish National Diabetes register [16]. In the Verona study, approximately half of the participants had some chronic complications of diabetes at the time of diagnosis, 11.2% had cardiovascular disease complications, 21.2% had neuropathy, 4.9% had retinopathy, 8.8% had chronic kidney disease based on estimated glomerular filtration rate, and 13.2% had albuminuria [17]. In a systematic review of the prevalence of microvascular and macrovascular complications in patients with newly diagnosed T2D in low- to middle-income countries, Aikaeli et al [18] reported that the prevalence of ischemic heart disease was 10%, peripheral artery disease 6%, stroke 2%, diabetic foot 1%, neuropathy 16%, and microalbuminuria 24%. This study also highlighted the paucity of data from the Africa region, and of the 33 reports included for analysis, only 4 were from Africa (1 each from Nigeria and Mauritius and 2 from Egypt). A press release from the International Diabetes Federation 2023 reported that two-thirds of newly diagnosed diabetes patients present with a diabetes-related chronic complication [15,18].

Patients with newly diagnosed diabetes have been noted to present with diabetic foot or arm infection, which may result in amputation of the limb. The global prevalence of diabetic foot is 6.4%. In SSA, it ranges between 4% and 19% [19]. In a cross-sectional Zimbabwean clinic study on 69 patients with known diabetes, the prevalence of diabetic foot disease was 10% [20].

Owing to the lack of reported data on the mode of presentation, etiologic types of diabetes, and prevalence of diabetes complications at the time of diagnosis, this study was developed to fill knowledge gaps in Zimbabwe. In addition, incident complications over a 2-year follow-up period in patients with type 2 diabetes and no complications at diagnosis would provide further insight into the outcome of treatment of patients in Zimbabwe.

Methods

This was a cross-sectional and prospective observational study undertaken on all adult and adolescent patients newly diagnosed

XSL•FO
RenderX

with diabetes who presented to the diabetes clinical service at referral hospitals and 2 city clinics in Bulawayo, Zimbabwe.

Study Design

This is a cross-sectional (baseline study, phase 1) and prospective (2-y follow-up study, phase 2) design.

The primary objective is to characterize the etiological types of diabetes in adult and adolescent patients with newly diagnosed diabetes in Zimbabwe.

The secondary objectives are to determine the prevalence of diabetes complications among adult and adolescent patients with newly diagnosed diabetes and the incidence and risk factors for the development of diabetic complications after a 2-year follow-up in patients with type 2 diabetes, who are complication free at baseline.

Study Sites

The study was undertaken at the Diabetes Clinical Service unit at Mpilo Central Hospital, United Bulawayo Hospitals (UBH), and 2 health clinics in Bulawayo (Luveve and Mpopoma). Mpilo Central Hospital and UBH are tertiary referral hospitals for Bulawayo Metropolitan, Matabeleland North, Matabeleland South, Masvingo, and Midlands provinces. Luveve and Mpopoma clinics are in densely populated residential regions of the Bulawayo Metropolitan Province.

Inclusion Criteria

The inclusion criteria were as follows: (1) patients aged ≥12 years presenting with newly diagnosed diabetes mellitus (WHO criteria) and (2) those who provided written consent or assent if aged <18 years.

Exclusion Criteria

The exclusion criteria were as follows: (1) patients aged <12 years and (2) those who declined to participate in the study.

Recruitment Strategy

The study was conducted at Mpilo Hospital, UBH, and 2 City Health Clinics. Messages regarding the study were disseminated through the hospital's and the city health's public relations offices to the public. The physicians were informed about the study through the Zimbabwe Medical Doctors Association and its affiliated members.

Ethical Considerations

Informed consent and assent were obtained by the research team from the participant or legal guardian. Ethical clearance was granted by the Medical Research Council of Zimbabwe (MRCZ/A/2735) and the Biomedical Research Ethics Committee University of KwaZulu-Natal (BREC 00001442/2020). No compensation was given to the participants for enrolling into the study.

Sample Size

In view of the limited number of studies on the prevalence of newly diagnosed diabetes in the region, a proportion of 0.3 (volume proxy figure) was used to calculate the sample size. Owing to time constraints, a pilot study could not be done. However, using the volume proxy proportion that was obtained in assessing the feasibility of the study, an analysis of patients' registries seeking services at the health institutions where data were collected revealed that, on average, 3 adolescent and adult patients were newly diagnosed with diabetes every 10 days, translating to a prevalence rate of 30% per month. On the basis of the proportion (0.3) of patients with newly diagnosed diabetes mellitus seeking care at the 2 referral hospitals and 2 local clinics per month in Bulawayo, a margin of error of 5% and a Z score of 1.96 using the formula $n=z^2(pxq)/d^2$, the sample size for cross-sectional (phase 1) was estimated to be 323. The acceptability of the study could not be ascertained due to the paucity of research being conducted in our population.

According to the systematic analysis conducted by Aikaeli et al [18], the median prevalence of complications in patients with newly diagnosed diabetes ranged between 10% and 16%. Using a 16% prevalence rate of complications in patients with newly diagnosed diabetes and single group proportion design for binary outcomes, with 95% CI, 5% precision, 80% study power, and an attrition rate of 20% using the formula $N=Z^2xp(1-p)/^{\Delta 2}=1.96^2\times0.16\times(1-0.16)/0.05^2=207$ was the required sample size for the prospective study. After adjusting for a 20% rate for attrition or loss to follow-up, $N^*=N\times1.25=259$ was the calculated sample size. Assessment of the incidence of complications according to the type of treatment received was not assessed in this study. As this is a mixed utility design, the highest sample size of 323 was used to conduct the study.

Statistical Analysis

SPSS 23 (IBM) was used to perform statistical analysis of the data. Descriptive statistical methods were used to analyze categorical variables such as types of diabetes, ankle-brachial index (ABI), carotid intima thickness, BMI, and autoantibodies. Continuous variables (eg, age, blood pressure, hemoglobin, cholesterol levels, random blood sugar, estimated glomerular filtration rate, glycated hemoglobin, and C-peptide levels) were documented as means and SDs. A chi-square analysis will be used to compare variables between males and females. The Cox regression analysis will be conducted to determine the time to the development of complications. Multiple regression analysis was used to identify risk factors for developing chronic complications of diabetes, and sensitivity analysis was used for the identification of the effect of confounders. Multiple imputation for missing values will be conducted to address the issue of missing values. Sensitivity analysis will be done during data analysis.

Study Outline and Data Collection

Overview

The study was undertaken in 2 phases. Adolescent and adult patients presenting to health care facilities with diabetes for the first time had their demographic data on age, sex, level of education, employment status, presenting symptoms, and medical history obtained using a questionnaire (Multimedia Appendix 1). Additionally, a physical examination and special tests such as ABI measurements, 10 g monofilament and 12 Hz tuning fork testing for neuropathy, and 12-lead electrocardiography (ECG) were performed by an experienced research team member. Carotid intima thickness measurements

XSL•FO
RenderX

and retinal photography were done by a trained ultrasound sonographer and a specialist ophthalmologist, respectively. Patients without complications at diagnosis were enrolled into phase 2, and they were monitored for 2 years with clinical evaluations at every 6-month interval as shown in Figure 1. A summary of assessments of patients is presented in Textbox 1.

XSL•FO
RenderX

Figure 1. Study outline.

XSL•FO
RenderX

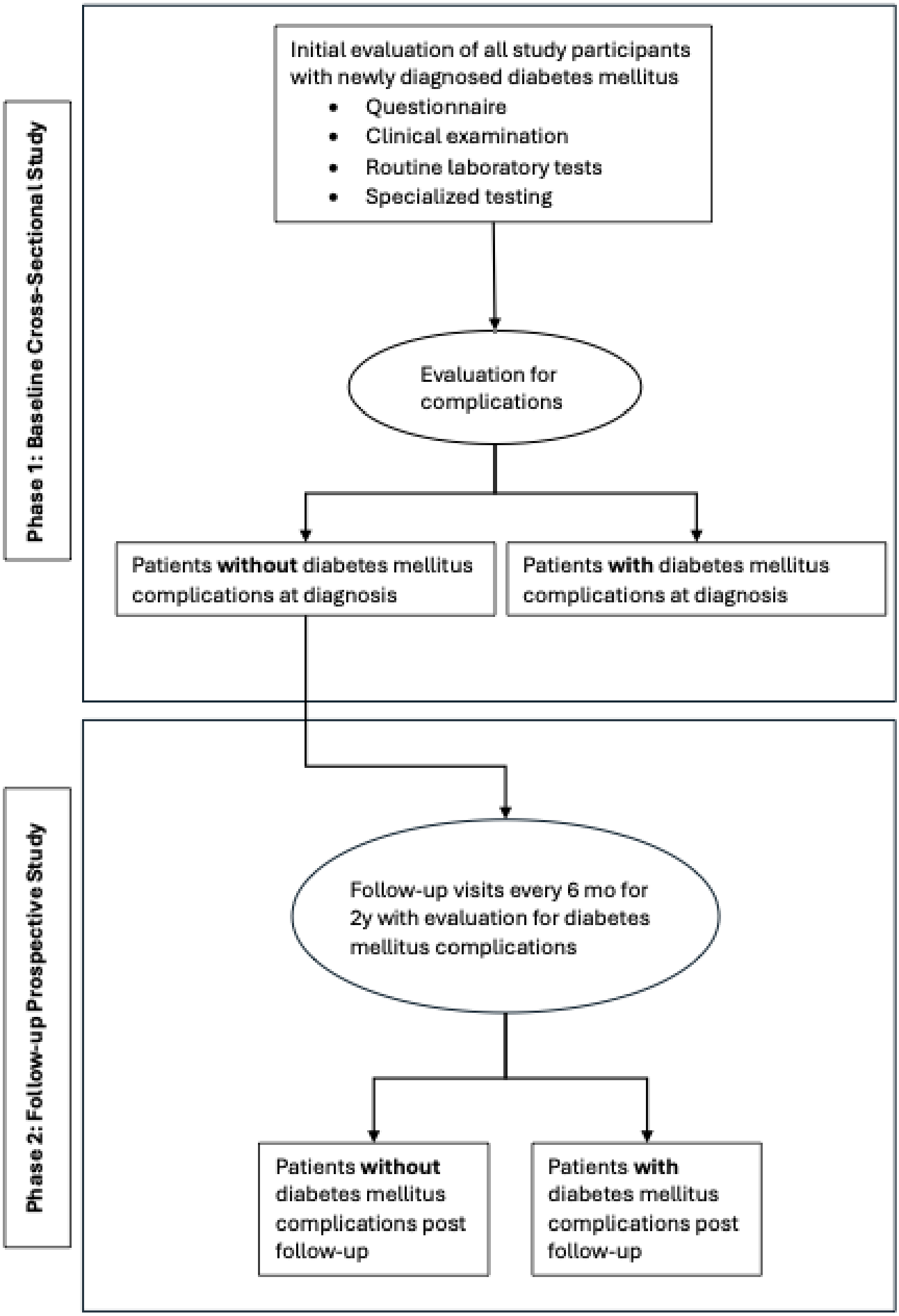

XSL•FO
RenderX

Textbox 1. Summary assessment of participants with newly diagnosed diabetes.

Questionnaire information

- Education
- Employment status
- Alcohol use
- Smoking
- Presenting complaints
- Medical history
- Family history

Clinical examination

- Blood pressure measurement
- Height
- Weight
- Waist circumference
- Hip circumference
- Waist-to-hip ratio
- 10 g monofilament
- 128 Hz tuning fork

Special tests

- Ankle-brachial index
- Body composition assessment
- Carotid intima thickness measurement
- 12-lead electrocardiograph
- Retinal photography

Laboratory tests

- Full blood count
- Urea and electrolytes
- Random blood glucose test
- Capillary blood glucose test
- Point-of-care glycated hemoglobin (HbA_{1c})
- Laboratory HbA_{1c}
- Liver function test
- High-density lipoprotein cholesterol
- Low-density lipoprotein cholesterol
- Triglycerides
- Total cholesterol
- Spot urine analysis

Special laboratory tests

- C-peptide
- Anti-GAD-65 antibody
- Zn 8 transporter antibody
- Islet cell antibody

XSL•FO
RenderX

- Anti-IA2 antibodies

Diagnostic Criteria for Diabetes, DKA, and Hyperosmolar Hyperglycemic Syndrome

Using WHO and American Diabetes Association recommendations, the diagnosis of diabetes mellitus was based on fasting plasma glucose ≥7.0 mmol/L, random plasma glucose ≥11.1 mmol/L in patients with symptoms characteristic of diabetes, and 2-hour plasma glucose ≥11.1 mmol/L during a 75 g oral glucose tolerance test [5,7]. DKA diagnosis was based on pH <7.3, HCO_3 <18 mmol/L, and ≥2 ketones in the urine [21]. The hyperosmolar hyperglycemic syndrome diagnosis was made on the following criteria: blood glucose >33.3 mmol/L, serum osmolality ≥320 mOsm/kg, volume depletion with pH >7.3, and HCO_3 >15.0 mmol/L [22].

Anthropometric measurements were performed using WHO guidelines [23,24].

BMI Measurements

The height and weight were used to calculate the BMI (weight [kg] divided by height [m^2]). According to the WHO classification of BMI, BMI <18.5 kg/m^2 is underweight, BMI between 18.5 and 24.9 kg/m^2 is normal weight, BMI between 25.0 and 29.9 kg/m^2 is overweight, BMI between 30.0 and 34.9 kg/m^2 is class 1 obesity, BMI between 35 and 39.9 kg/m^2 is class 2 obesity, and BMI ≥40 kg/m^2 is class 3 obesity [23].

Waist Circumference

Using WHO and International Diabetes Federation recommendations for waist circumference (WC) cut points, WC is "elevated" if ≥94 cm in men and ≥80 cm in women; WC is "substantially elevated" if there are ≥102 cm in men and ≥88 cm in women, increasing both the metabolic and cardiovascular risks [24].

Hip Circumference

Hip circumference was measured in centimeters, with the individual standing with their feet together. A nonstretching tape measure was placed around the prominent part of the gluteal region parallel to the floor [24].

Waist-to-Hip Ratio

Waist-to-hip ratio (WHR) was obtained by dividing the waist measurement by the hip measurement. WHR is elevated if ≥0.90 in men and ≥0.85 in women. If WHR >1, the risk of cardiovascular disease is high [24].

Blood Pressure

Blood pressure was measured with the patient seated and after resting for at least 15 minutes. The blood pressure was measured 3 times, and the average was recorded as the BP, with BP ≥140/90 mm Hg defined as hypertension [25].

Body Mass Composition

The patient was requested to stand barefoot on the MC TANITA 780MA biomedical impedance machine, with both hands on the grip handle for 5 minutes. The visceral fat, body fat, muscle, bone, and water content values of the patient were measured and compared with gender- and age-specific values using the inbuilt machine formulae [26].

Assessment for Chronic Complications

Diabetic Retinopathy

Assessment for background, proliferative diabetic retinopathy, and maculopathy was undertaken using ophthalmoscopy and retinal photography. In this study, nonmydriatic retinal photography was performed with a portable retinal camera (Microclear) by an ophthalmologist in all patients at diagnosis. Retinopathy was assessed using the Scottish grading system [27].

The Scottish grading system is as follows:

- R0: no retinopathy
- R1: mild background retinopathy (microaneurysms, flame exudates, >4 blot hemorrhages in one or both hemifields, and cotton wool spots)
- R2: moderate background diabetic retinopathy and >4 blot hemorrhages in one hemifield
- R3: severe nonproliferative or preproliferative diabetic retinopathy in both hemifields, intraretinal microvascular anomalies, and venous beading
- R4: Proliferative retinopathy: neovascularization of the disc, neovascularization elsewhere, vitreous hemorrhage, retinal detachment
- M0: No macular findings
- M1: Hard exudates within 1 - to 2-disc diameters of fovea
- M2: Blot hemorrhages or hard exudates

Diabetic Neuropathy

Peripheral neuropathy was evaluated using a 10 g Semmes-Weinstein monofilament examination and vibration test. The monofilament was applied perpendicular to the pulp of the hallux, third toe, and metatarsal phalangeal joints 1, 2, and 3, for 2 seconds per site until it buckles. Absence of pressure sensation in more than 8 sites is regarded as the presence of peripheral neuropathy [28]. A 128 Hz tuning fork was placed at the interphalangeal joint of the hallux. The patient was asked to report any vibration felt. The normal sensation is felt for approximately 60 seconds. The vibration sensation was recorded as absent if the patient failed to perceive any vibration [29].

Specialized Clinical Tests

Carotid Intima-Media Thickness

Carotid intima-media thickness (CIMT) was measured to assess for atherosclerosis using the DC7 Mindray ultrasound machine. A clear water-based gel was applied to the skin in the neck area above the carotid vessels. Each carotid vessel was assessed separately. A 6.5-MHz convex transducer was placed on the common carotid area, and the CIMT was measured. The normal CIMT ranges from 0.59 to 0.95 mm in men and 0.52 to 0.93 mm in women. CIMT >1 mm is associated with significant

XSL•FO
RenderX

atherosclerosis [30]. The measurement of CIMT was undertaken by a qualified sonographer who was trained in CIMT measurements. To improve interrater reliability, CIMT measurements at all sites were carried out by 1 CIMT measurement–trained sonographer.

Ankle-Brachial Index

The patient was allowed to rest and lie in the supine position for 10 minutes. A blood pressure cuff was placed above the malleolus on the lower limb and above the cubital fossa on the arm. BP was measured twice in both limbs. The higher value of the 2 readings was used to calculate the ABI. ABI between 0.90 and 1.30 is normal, whereas ABI <0.90 is indicative of peripheral artery disease and ABI >1.30 is suggestive of calcification of the vessels [31]. To minimize errors in ABI measurements, all personnel underwent training.

Electrocardiography

A resting 12-lead ECG was recorded using a Schiller Cardiovit 102 machine. The patient was put in a supine position on a couch and had limb leads placed symmetrically. Chest leads were placed in order from V1 to V6, in the following sequence: V1 placed in the fourth intercostal space on the right sternal border, V2 on the fourth intercostal space on the left sternal border, V3 placed mid-way between V2 and V4, V4 placed on the fifth intercostal space midclavicular line, V5 placed between V4 and V6, and V6 placed on the fifth intercostal space midclavicular line.

Laboratory Tests

Blood samples were collected from all patients enrolled in the study (phase 1) by a trained laboratory team. Transportation of all samples was undertaken by the laboratory team following standard operating procedures using guidelines provided by the manufacturer of the reagents that were used to analyze specific samples.

Random plasma glucose, glycated hemoglobin, serum urea and electrolytes, serum antibodies, liver function tests, and random serum C-peptide were collected during the first visit [11]. Serum lipids were collected after 12 hours of fasting. Blood for random plasma glucose was collected in a sodium fluoride tube, and for all the other tests, a plain tube was used. A spot urine sample was collected at the initial visit in a plain tube. Laboratory blood tests are summarized in **Table 1**.

XSL•FO
RenderX

Table . Summary: laboratory investigations and methods of newly diagnosed diabetes.

Variable	Sample tube	Assay method
Full blood count	EDTA[a] tube	Beckman Coulter
Random plasma glucose	Sodium fluoride (gray top)	Hexokinase (Vitros 5.1 analyzer)
HbA_{1c}[b] (point-of-care)	—[c]	Nephelometric method (point-of-care SDA_{1C} machine)
HbA_{1c}	EDTA tube (purple top)	Calorimetric (MISPA I2)
Total cholesterol	Plain tube	Calorimetric (Vitros 5.1 analyzer)
Triglyceride	Plain tube	Calorimetric (Vitros 5.1 analyzer)
Cholesterol (HDL[d])	Plain tube	Calorimetric (Vitros 5.1 analyzer)
Cholesterol (LDL[e])	Plain tube	LDL calculated using Friedewald equation
Antiglutamic acid decarboxylase (anti-GAD[f] 65)	Plain tube	ELISA[g] (Euroimmun)
Anti-insulin IA2[h] antibody	Plain tube	ELISA (Euroimmun)
Antizinc transporter 8 antibody	Plain tube	ELISA (Euroimmun)
Anti-islet antibody	Plain tube	Immunofluorescence
Serum C-peptide	Plain tube	MALGUMI C-Peptide (CLIA)
LFT[i]	Plain tube	
Total protein		Calorimetric (Vitros 5.1 analyzer)
Albumin		Calorimetric (Vitros 5.1 analyzer)
AST[j]		Multipoint rate (Vitros 5.1 analyzer)
ALT[k]		Multipoint rate (Vitros 5.1 analyzer)
ALP[l]		Multipoint rate (Vitros 5.1 analyzer)
LDH[m]		Multipoint rate (Vitros 5.1 analyzer)
Total bilirubin		Micro Sensor TM (Vitros 5.1 analyzer)
HIV	EDTA tube (red top)	ELISA
CD4 count	EDTA	Flow cytometry (BD FACSCount)
Spot urine and microalbumin or UACR[n]	Plain tube	Spectrophotometry (Vitros 5.1)

[a]EDTA: ethylenediaminetetraacetic acid.
[b]HbA1c: glycated hemoglobin.
[c]Not available.
[d]HDL: high-density lipoprotein.
[e]LDL: low-density lipoprotein.
[f]GAD: glutamic acid decarboxylase.
[g]ELISA: enzyme-linked immunosorbent assay.
[h]IA2: insulinoma antigen 2.
[i]LFT: liver function test.
[j]AST: aspartate transaminase.
[k]ALT: alanine transaminase.
[l] ALP: alkaline phosphatase.
[m]LDH: lactate dehydrogenase.
[n]UACR: urine albumin-creatinine ratio.

XSL•FO
RenderX

Prospective Study—Phase 2

The prospective (phase 2) component of the study includes all patients with newly diagnosed type 2 diabetes who were without chronic complications at initial presentation (phase 1).

Patients with any diabetes-related complication at presentation were excluded. Follow-up visits at 6, 12, 18, and 24 months after diagnosis aimed to assess for incident diabetes complications.

Similar to the baseline study, at every visit during follow-up, the methodology included questionnaire information, clinical examination, and laboratory investigations; additionally, assessment of medications and diabetes control was undertaken. The questionnaire used in phase 1 was used to assess the incidence of chronic diabetes complications. Retinal examination, body composition, carotid intimal thickness, and ECG were performed at 12 and 24 months. The follow-up of patients and outcomes during the clinic visits are summarized in Table 2. Diabetes complications that occurred between the scheduled visits were reported to the research team by the attending health care personnel. Adverse events related to prescribed medicines were addressed by the attending medical team.

XSL•FO
RenderX

Table . Summary of phase 2 visits and outcomes.

Parameter being assessed	Normal ranges	Baseline	6 mo	12 mo	18 mo	24 mo
Diabetes care	—[a]	—	Insulin or oral antidiabetic agent	Insulin or oral antidiabetic agent	Insulin or oral antidiabetic agent	Insulin or oral antidiabetic agent
BMI (kg/m^2)	18.5 - 24.9	✓[b]	✓	✓	✓	✓
SBP[c] (mm Hg)	<140	✓	✓	✓	✓	✓
DBP[d] (mm Hg)	<90	✓	✓	✓	✓	✓
ABI[e]	0.90 - 1.30	✓	✓	✓	✓	✓
10 g monofilament	Normal sensation in 8 or >8 sites	✓	✓	✓	✓	✓
HbA_{1c}[f] (%/mmol/mol)	<5.7/42	✓	✓	✓	✓	✓
Total cholesterol (mmol/L)	0 - 5.2	✓	✓	✓	✓	✓
LDL[g] (mmol/L)	<2.0	✓	✓	✓	✓	✓
HDL[h] (mmol/L)	1.03 - 1.55	✓	✓	✓	✓	✓
Triglyceride (mmol/L)	0 - 1.69	✓	✓	✓	✓	✓
Urea (mmol/L)	2.5 - 7.9	✓	✓	✓	✓	✓
Creatinine (µmol/L)	55 - 120	✓	✓	✓	✓	✓
eGFR[i] (mL/min/1.73m^2)	≥90	✓	✓	✓	✓	✓
Urinalysis	No abnormality	✓	✓	✓	✓	✓
CIMT[j] (mm)	0.6 - 0.7	✓		✓		✓
Retinal photograph	Normal	✓		✓		✓
ECG[k]	Normal	✓		✓		✓

[a]Not applicable.

[b]Checkmarks indicate which measurements are taken at each time point.

[c]SBP: systolic blood pressure.

[d]DBP: diastolic blood pressure.

[e]ABI: ankle-brachial index.

[f]HbA_{1c}: glycated hemoglobin.

[g]LDL: low-density lipoprotein.

[h]HDL: high-density lipoprotein.

[i]eGFR: glomerular filtration rate.

[j]CIMT: carotid intima media thickness.

[k]ECG: electrocardiograph.

Results

The study describes the characteristics of patients with newly diagnosed diabetes, including the mode of presentation, the prevalence of complications, and the assignment of a specific etiologic type, in patients presenting to 4 urban health facilities in Bulawayo, Zimbabwe. In addition, the prevalence of diabetes complications at the time of diagnosis and the incidence of diabetes complications over 2 years of postdiagnosis follow-up will be determined (Textbox 1). Data collection for the cross-sectional study (phase 1) commenced in October 2021 and ended in October 2024. Data analysis and manuscript writing for the phase 1 study are underway. Data collection for the prospective study (phase 2) is ongoing until October 2026.

XSL•FO
RenderX

Discussion

Principal Findings

Epidemiologic data indicate a rising prevalence of diabetes in Zimbabwe [4], as in other countries in SSA, and therefore, the health services would expect to encounter patients with symptomatic disease more frequently [3]. However, it is not known what the common mode of presentation is in Zimbabwe, nor are there data available on the prevalence of microvascular and macrovascular complications at the time of diagnosis. There are several factors that may influence the increasing prevalence of diabetes, including urbanization, obesity, diet and lifestyle changes, environmental toxins, alcohol, and other factors [2]. The influence of these factors in Zimbabwe is not known, and this study will provide some information in this regard, as the data on lifestyle, mode of presentation, and clinical phenotype will provide details on the association with obesity and lifestyle. In addition, this study incorporates patients from 12 years of age, and it is envisaged that a significant proportion of the younger patients will be diagnosed with type 1 diabetes. The prevalence of type 2 diabetes in young persons in Zimbabwe is not known, and this study will provide details on this diagnosis. Similarly, less common etiologic types, such as ketosis-prone diabetes, may be a form of the disease in Zimbabwe. This study uses diagnostic methods that will assist in assigning an etiology with a level of certainty much higher than has been the case in the past, where reliance was primarily on age at diagnosis and clinical phenotype. While C-peptide may be affected by acute, severe hyperglycemia, due to β-cell glucose toxicity, the combination of both C-peptide and measurement of β-cell autoantibodies will provide a level of diagnostic clarity that has not been possible in the past [5,7,10-12]. It is envisaged that younger patients may have type 1 diabetes, and the lean adults presenting with ketoacidosis may have KPD. As with most other studies, it is considered likely that the majority of the remaining subjects enrolled will have type 2 diabetes. The study will provide data on the body composition of subjects with different etiologic types of diabetes, and while obesity may be increasing in parallel with diabetes, it is not known whether or not this occurs in patients with newly diagnosed diabetes in Zimbabwe. Furthermore, the prevalence of common comorbidities, including hypertension and dyslipidemia, will be provided by the initial phase of this study.

It is possible that there is a high prevalence of undiagnosed diabetes mellitus in Zimbabwe and that only once symptoms develop will the diagnosis be established. This implies prolonged undiagnosed hyperglycemia, during which time complications may develop [14,17,18]. This study will provide data on the prevalence of complications, using investigations, such as retinal photography and carotid intima-media thickness, as well as more widely used clinical tools to improve the detection of these complications. Accurate diagnosis of diabetes complications is important because the implementation of medical management strategies to mitigate progression will alleviate some of the burden on the health care system. Furthermore, the prospective component of the study will provide more data on the incidence of complications in those with type 2 diabetes.

Limitations

There are several limitations to this study. Patients presenting to the selected 4 health care facilities in Bulawayo may not reflect the broader population of Zimbabwe, particularly those residing entirely in rural areas. The study is only able to report on clinical, laboratory, and radiologic parameters in patients presenting to the health services with symptomatic disease and cannot determine factors associated with undiagnosed disease. The follow-up period in the prospective study is relatively short and therefore will not provide a true estimate of the long-term risk of complications. There may be an attrition of follow-up attendance in the prospective phase, but steps will be taken to minimize this, including direct contact with the patients in this phase of the study.

Strengths

The strengths of this study include detailed characterization of the enrolled patients with sophisticated laboratory and radiologic techniques that have not been used in prior studies in Zimbabwe as well as the detailed follow-up evaluation of patients to determine not only incident complications but also adequacy of glycemic control and management of comorbidities.

The importance of this study to public health in Zimbabwe is envisaged as raising awareness of the common modes of presentation and etiologic types of diabetes and informing health authorities and health practitioners to assist in the more appropriate therapeutic measures to use, with the aim of lowering the development of chronic complications.

Acknowledgments

The authors would like to acknowledge Professor Ayesha Motala, Dr Fraser Pirie, and Professor Elopy Sibanda for their valuable support and contributions in the writing up of the protocol. The authors would also like to express their sincere gratitude to Diagnostic Laboratory Services for sponsoring all the laboratory investigations conducted in this research and Greenwood Wholesalers and Pharmaceuticals for supplying glucose testing machines and glucose testing strips to patients who enrolled in phase 2 of the study. It would have been impossible to conduct this research without their valuable support. The authors would like to express their sincere gratitude to the patients and their families for allowing them to use their data. The authors attest that there was no use of generative artificial intelligence technology in the generation of the text, figures, or other informational content of this manuscript.

XSL•FO
RenderX

Funding

The study was funded by Diagnostic Laboratory Services and Greenwood Wholesalers and Pharmaceuticals. The funders had no involvement in the study design, data collection, analysis, interpretation of results, or writing of the manuscript.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Findings of the study will be shared with the relevant department of noncommunicable disease in the Ministry of Health and Child Care by way of presenting a summary report of the main findings. The results of the study will also be disseminated to the greater medical fraternity during scientific conferences.

Authors' Contributions

AAM, FJP, and ES conceptualized the study idea, designed and supervised the study, and reviewed and approved the final manuscript. DM assisted with the formal analysis of data. RG conceptualized the study idea and was responsible for data curation and manuscript writing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire and clinical examination.
[PDF File, 426 KB - resprot_v15i1e74186_app1.pdf]

References

1. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med 1998 Jul;15(7):539-553. [doi: 10.1002/(SICI)1096-9136(199807)15:7<539::AID-DIA668>3.0.CO;2-S] [Medline: 9686693]
2. The Diabetes Atlas, 7th edn. : International Diabetes Federation; 2015 URL: https://www.diabetesatlas.org [accessed 2025-12-10]
3. The Diabetes Atlas, 8th edn. : International Diabetes Federation; 2017 URL: https://www.diabetesatlas.org [accessed 2025-12-10]
4. Mutowo M, Gowda U, Mangwiro JC, Lorgelly P, Owen A, Renzaho A. Prevalence of diabetes in Zimbabwe: a systematic review with meta-analysis. Int J Public Health 2015 Jan;60(1):1-11. [doi: 10.1007/s00038-014-0626-y] [Medline: 25432797]
5. American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes-2019. Diabetes Care 2019 Jan;42(Suppl 1):S13-S28. [doi: 10.2337/dc19-S002] [Medline: 30559228]
6. Definition, diagnosis and classification of diabetes mellitus and its complications: report of a WHO consultation. part 1, diagnosis and classification of diabetes mellitus. : World Health Organization; 1999 URL: https://iris.who.int/handle/10665/66040 [accessed 2025-12-10]
7. Classification of diabetes mellitus. : World Health Organization; 2019 URL: https://iris.who.int/handle/10665/325182 [accessed 2025-12-10]
8. Mbanya JCN, Motala AA, Sobngwi E, Assah FK, Enoru ST. Diabetes in sub-Saharan Africa. Lancet 2010 Jun 26;375(9733):2254-2266. [doi: 10.1016/S0140-6736(10)60550-8] [Medline: 20609971]
9. Jones AG, Hattersley AT. The clinical utility of C-peptide measurement in the care of patients with diabetes. Diabet Med 2013 Jul;30(7):803-817. [doi: 10.1111/dme.12159] [Medline: 23413806]
10. Zhou Patricia D, Kouassi JB, Monteomo GF. Basic C-peptidemia and diabetic patients classification. Sci J Clin Med 2017;6(6):116. [doi: 10.11648/j.sjcm.20170606.14]
11. Berger B, Stenström G, Sundkvist G. Random C-peptide in the classification of diabetes. Scand J Clin Lab Invest 2000 Dec;60(8):687-693. [doi: 10.1080/00365510050216411] [Medline: 11218151]
12. Gill GV, Tekle A, Reja A, et al. Immunological and C-peptide studies of patients with diabetes in northern Ethiopia: existence of an unusual subgroup possibly related to malnutrition. Diabetologia 2011 Jan;54(1):51-57. [doi: 10.1007/s00125-010-1921-7] [Medline: 20890591]
13. Umpierrez GE, Smiley D, Kitabchi AE. Narrative review: ketosis-prone type 2 diabetes mellitus. Ann Intern Med 2006 Mar 7;144(5):350-357. [doi: 10.7326/0003-4819-144-5-200603070-00011] [Medline: 16520476]
14. Sosale A, Prasanna Kumar KM, Sadikot SM, et al. Chronic complications in newly diagnosed patients with type 2 diabetes mellitus in India. Indian J Endocrinol Metab 2014 May;18(3):355-360. [doi: 10.4103/2230-8210.131184] [Medline: 24944931]
15. The Diabetes Atlas, 10th. : International Diabetes Federation; 2021 URL: https://www.diabetesatlas.org [accessed 2025-12-10]

XSL•FO
RenderX

16. Höskuldsdóttir G, Franzén S, Eeg-Olofsson K, Eliasson B. Risk trajectories of complications in over one thousand newly diagnosed individuals with type 2 diabetes. Sci Rep 2022 Jul 11;12(1):11784. [doi: 10.1038/s41598-022-16135-0] [Medline: 35821071]
17. Bonora E, Trombetta M, Dauriz M, et al. Chronic complications in patients with newly diagnosed type 2 diabetes: prevalence and related metabolic and clinical features: the Verona newly diagnosed type 2 diabetes study (VNDS) 9. BMJ Open Diabetes Res Care 2020 Aug;8(1):e001549. [doi: 10.1136/bmjdrc-2020-001549] [Medline: 32819978]
18. Aikaeli F, Njim T, Gissing S, et al. Prevalence of microvascular and macrovascular complications of diabetes in newly diagnosed type 2 diabetes in low-and-middle-income countries: a systematic review and meta-analysis. PLOS Glob Public Health 2022;2(6):e0000599. [doi: 10.1371/journal.pgph.0000599] [Medline: 36962416]
19. Abbas ZG, Archibald LK. Epidemiology of the diabetic foot in Africa. Med Sci Monit 2005 Aug;11(8):RA262-RA270. [Medline: 16049394]
20. Nyika P, Chimusoro A, Tshimanga M, Gombe N, Takundwa L, Bangure D. Risk factors for diabetic complications among diabetic patients, Chirumanzu District, Zimbabwe, 2011. Austin J Public Health Epidemiol 2015 Sep;2:1021 [FREE Full text]
21. Umpierrez GE, Davis GM, ElSayed NA, et al. Hyperglycemic crises in adults with diabetes: a consensus report. Diabetes Care 2024 Aug 1;47(8):1257-1275. [doi: 10.2337/dci24-0032] [Medline: 39052901]
22. Fadini GP, de Kreutzenberg SV, Rigato M, et al. Characteristics and outcomes of the hyperglycemic hyperosmolar non-ketotic syndrome in a cohort of 51 consecutive cases at a single center. Diabetes Res Clin Pract 2011 Nov;94(2):172-179. [doi: 10.1016/j.diabres.2011.06.018] [Medline: 21752485]
23. Body mass index (BMI). World Health Organization. 2025. URL: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/body-mass-index [accessed 2025-12-10]
24. Waist circumference and waist-hip ratio: report of a WHO expert consultation, Geneva, 8-11 December 2008. : World Health Organization; 2011 URL: https://iris.who.int/handle/10665/44583 [accessed 2025-12-10]
25. Whelton PK, Carey RM, Aronow WS, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Hypertension 2018 Jun;71(6):1269-1324. [doi: 10.1161/HYP.0000000000000066] [Medline: 29133354]
26. Tanita. URL: https://www.tanita.com [accessed 2025-12-10]
27. Zachariah S, Wykes W, Yorston D. Grading diabetic retinopathy (DR) using the Scottish grading protocol. Community Eye Health 2015;28(92):72-73. [Medline: 27418727]
28. Boulton AJM, Armstrong DG, Albert SF, et al. Comprehensive foot examination and risk assessment: a report of the task force of the foot care interest group of the American Diabetes Association, with endorsement by the American Association of Clinical Endocrinologists. Diabetes Care 2008 Aug;31(8):1679-1685. [doi: 10.2337/dc08-9021] [Medline: 18663232]
29. Gilman S. Joint position sense and vibration sense: anatomical organisation and assessment. J Neurol Neurosurg Psychiatry 2002 Nov;73(5):473-477. [doi: 10.1136/jnnp.73.5.473] [Medline: 12397137]
30. Chen LY, Leening MJG, Norby FL, et al. Carotid intima-media thickness and arterial stiffness and the risk of atrial fibrillation: the atherosclerosis risk in communities (ARIC) study, multi-ethnic study of atherosclerosis (MESA), and the Rotterdam study. J Am Heart Assoc 2016 May 20;5(5):e002907. [doi: 10.1161/JAHA.115.002907] [Medline: 27207996]
31. Aboyans V, Criqui MH, Abraham P, et al. Measurement and interpretation of the ankle-brachial index: a scientific statement from the American Heart Association. Circulation 2012 Dec 11;126(24):2890-2909. [doi: 10.1161/CIR.0b013e318276fbcb] [Medline: 23159553]

Abbreviations

ABI: ankle-brachial index
CIMT: carotid intima-media thickness
DKA: diabetic ketoacidosis
ECG: electrocardiograph
HbA_{1c}: glycated hemoglobin
KPD: ketosis-prone type 2 diabetes
SSA: sub-Sahara Africa
T1D : type 1 diabetes
T2D: type 2 diabetes
WC: waist circumference
WHO: World Health Organization
WHR: waist-to-hip ratio

XSL•FO
RenderX

Edited by J Sarvestan; submitted 20.Mar.2025; peer-reviewed by AN Ali; revised version received 14.Nov.2025; accepted 17.Nov.2025; published 06.Jan.2026.

Please cite as:
Gwini R, Sibanda E, Mwembe D, Pirie FJ, Motala AA
Etiologic Types and Complications of Diabetes Mellitus in Newly Diagnosed Patients at Health Institutions in Bulawayo, Zimbabwe: Protocol for a Cross-Sectional and Prospective Observational Study
JMIR Res Protoc 2026;15:e74186
URL: https://www.researchprotocols.org/2026/1/e74186
doi:10.2196/74186

© Rudo Gwini, Elopy Sibanda, Desmond Mwembe, Fraser James Pirie, Ayesha A Motala. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Specialists' Perceptions of Workforce Retention Strategies in the Malaysian Ministry of Health and Their Association With Job Satisfaction and Turnover Intention: Protocol for a REDCap-Based National Cross-Sectional Survey

Siti Zubaidah Ahmad Subki[*], MPH, MD; Ili Abdullah Sharin[*], MD; Nor Haniza Zakaria[*], MD; Nursyahda Zakaria[*], BSc, MSc; Pangie Bakit[*], MBBS; Izzuan Khirman Adnan[*], MBBS; Norehan Jinah[*], MD; Kun Yun Lee[*], MBBChBAO, MPH, DrPH

Institute for Health Management, National Institutes of Health, Ministry of Health Malaysia, Block B1, NIH Complex, No. 1 Jalan Setia Murni U13/52, Section U13, Setia Alam, Shah Alam, Malaysia

[*]all authors contributed equally

Corresponding Author:
Ili Abdullah Sharin, MD
Institute for Health Management, National Institutes of Health, Ministry of Health Malaysia, Block B1, NIH Complex, No. 1 Jalan Setia Murni U13/52, Section U13, Setia Alam, Shah Alam, Malaysia

Abstract

Background: Retention of specialists is critical for sustaining health system performance. In Malaysia, the Ministry of Health (MOH) has implemented multiple workforce retention strategies (WRSs) to enhance job satisfaction and reduce turnover among specialists. However, evidence on specialists' awareness, participation, and perceived effectiveness of these strategies remains limited. To address this gap, we plan to administer a standardized survey among specialists in MOH health care facilities to guide evidence-based strategic planning toward specialist retention.

Objective: This protocol describes a national-level online survey to assess MOH specialists' perceptions of WRS and examine their associations with job satisfaction and turnover intention, using REDCap (Research Electronic Data Capture) to ensure real-time data collection within a robust and secure digital platform.

Methods: A cross-sectional mixed methods study will be conducted among 1325 MOH specialists selected through systematic random sampling from the Human Resource Management Information System. Data collection will be implemented using a REDCap-based workflow to support secure and efficient survey administration. REDCap functionalities will be used, including automated survey invitations with nontransferable links to prevent duplicate responses, branching logic to tailor item relevance, auto-reminder scheduling, and real-time data validation to minimize entry errors. The self-administered questionnaire comprises measures of specialists' awareness, participation, and perceived effectiveness of 12 MOH-implemented WRSs, job satisfaction assessed using the Job Satisfaction Survey, and turnover intention measured with the Turnover Intention Scale-6, as well as 2 optional open-ended questions for qualitative input to strengthen the quantitative results. A pilot study will be conducted to assess instrument reliability and REDCap platform usability. Quantitative data will be exported from REDCap for descriptive and inferential analyses, while qualitative responses will undergo thematic analysis using NVivo and be integrated with quantitative findings during interpretation.

Results: This study will generate vital evidence of MOH specialists' engagement with WRSs and how perceptions of these strategies related to their job satisfaction and turnover intention. Participant recruitment and data collection have been completed, and the study is in the data analysis phase. It is expected that the data analysis will be completed in March 2026 and that the results will be published in June 2026.

Conclusions: This protocol establishes a REDCap-based survey for conducting large-scale research focused on MOH specialists. Findings are expected to guide data-driven improvements to WRS implementation within the MOH Malaysia and may serve as a useful methodological model for similar research in public sector workforce studies.

Trial Registration: Malaysian National Medical Research Register NMRR ID-23-03199-3VQ; https://nmrr.gov.my/research-directory/37872157-613e-4ffd-b01b-10fda6b6a736

International Registered Report Identifier (IRRID): DERR1-10.2196/83377

(JMIR Res Protoc 2026;15:e83377) doi:10.2196/83377

XSL•FO
RenderX

KEYWORDS

physician; job satisfaction; turnover intention; retention strategies; sustainability; REDCap

Introduction

Background

Retention strategies are essential for building a sustainable and efficient workforce, particularly in the health care sector, where skilled professionals play a pivotal role in delivering high-quality services [1]. These strategies aim to address ongoing workforce shortages and promote long-term stability by fostering a motivated and committed workforce. As service demands grow and global shortages of health care professionals intensify, effective retention strategies have become increasingly important [1]. Specialists, in particular, are often the focus of retention strategies due to their roles in complex clinical decision-making, procedural care, and mentorship of junior staff [2]. Retaining these highly trained professionals supports continuity of care and reduces the substantial costs associated with recruitment and training [3].

Globally, retention strategies vary depending on a country's economic capacity, health care priorities, and system infrastructure. High-income countries such as the United Kingdom often adopt large-scale approaches like international recruitment, supported by mentorship programs and relocation incentives. However, these strategies frequently face limitations such as high implementation costs, difficulties with professional integration, and limited long-term retention [4,5]. Middle-income countries, including Brazil and Thailand, typically implement targeted strategies such as bonded scholarships, rural service allowances, and performance-based incentives to address workforce shortages in underserved areas [6,7], though these often struggle with administrative complexity and resource constraints. In contrast, low-income countries tend to rely on donor-funded programs and compulsory service schemes, which are frequently undermined by inadequate infrastructure and limited career progression opportunities [8]. These global experiences highlight the importance of context-sensitive, multifaceted strategies that reflect the unique dynamics of each health system [9,10].

In this context, Malaysia has introduced various workforce retention strategies (WRSs) since the early 1990s to improve the retention of specialists in the Ministry of Health (MOH) (see Multimedia Appendix 1; Human Resource Division, Ministry of Health Malaysia, unpublished data, 2023). These strategies encompass financial incentives (eg, on-call allowances, specialist incentive allowance, and Full Paying Patient Scheme), structured career progression pathways (eg, time-based promotion and introduction of Grade 56), flexible working arrangements (eg, flexible working hours), and professional development opportunities (eg, subspecialty training and short-term training abroad). In addition, geographically based allowances are used to support the deployment of specialists to underserved regions, such as Sabah and Sarawak. These initiatives are consistent with international best practices, which emphasize a combination of remuneration, professional support systems, and career advancement opportunities tailored to local workforce dynamics [11].

Although these strategies represent a multifaceted approach to workforce retention, patterns in specialist resignation trends have varied over time. Administrative data from the MOH Human Resource Division showed that overall resignation rates among specialists fluctuated between 2.1% and 5.8% from 2015 to 2023 (Human Resource Division, Ministry of Health Malaysia, unpublished data, 2023). Between 2015 and 2017, observed reductions coincided with the implementation of several major WRSs, including the Full Paying Patient Scheme, the introduction of Grade 56, and flexible working hours (Human Resource Division, Ministry of Health Malaysia, unpublished data, 2023). During this period, overall resignation rates declined from 3.9% to 3.4%, with a more substantial drop observed among surgical specialists, whose rates fell from 6.7% to 4.6% (Human Resource Division, Ministry of Health Malaysia, unpublished data, 2023). While these temporal patterns coincide with the introduction of selected WRSs, the descriptive nature of these data does not permit the attribution of changes in resignation trends to specific strategies. Rising resignation rates in subsequent years underscore the need to examine how contextual factors and implementation conditions of WRSs are associated with specialists' perceptions of the effectiveness and sustainability of these strategies.

To date, there has been a limited systemwide examination of specialists' uptake, engagement, and perceptions of WRSs in MOH, Malaysia, since their inception. Existing studies tend to be narrow in scope, often focusing on individual strategies in isolation, which limit their generalizability and practical relevance for broader policy development [12,13]. A more comprehensive assessment of how specialists engage and how they perceive these strategies in relation to job satisfaction and turnover intention is therefore warranted [14,15]. To guide this assessment, theoretical frameworks offer valuable insights for understanding the relationships between retention strategies and workforce-related outcomes. One of the most widely used is Herzberg's Two-Factor Theory, which distinguishes between hygiene factors (eg, salary, job security, and working conditions), which prevent dissatisfaction, and motivators (eg, recognition, professional growth, and autonomy), which enhance job satisfaction and promote retention [16,17]. This theory underpins models such as Australia's Whole-of-Person Retention Improvement Framework and is supported by systematic reviews on health care workforce stability [14,18]. Applying this framework to Malaysia's public health context may help clarify which components of WRSs are most effective in supporting long-term workforce sustainability.

To address current evidence gaps, a systematic and scalable data collection approach is required. REDCap (Research Electronic Data Capture) is a secure, browser-based platform with customizable survey logic, encryption, audit trails, and real-time validation features that are ideal for online data collection. It has been used in over 140 countries, with a proven

XSL•FO
RenderX

track record in supporting large-scale research [19,20]. Despite its widespread use in clinical and epidemiological studies, REDCap remains underutilized for workforce policy-related research, especially in middle-income countries [21]. Its application in this study will enable standardized, confidential, and geographically distributed data collection across Malaysia's public health system, critical for assessing specialist engagement with WRSs, as well as offering a methodological framework for similar research.

Objectives

This protocol will describe a national-level survey using REDCap to systematically examine Malaysian MOH specialists' awareness, participation, and perceptions of 12 MOH-implemented WRSs. It will outline the methodological approach for examining associations between these engagement dimensions, job satisfaction, and turnover intention. It is intended to support standardized data collection and generate reproducible evidence to inform policy and practice for strengthening specialist retention in Malaysia's public health care system.

Methods

Study Design and Setting

This study will adopt a nationwide, cross-sectional design using a closed online survey administered through the REDCap platform, hosted by the Sector of Biostatistics and Data Repository, National Institutes of Health Malaysia. The study design was developed based on REDCap version 14.0.21. Data collection was conducted using the institutionally supported current long-term support version of REDCap. All study procedures will be centrally administered under a single governance structure, with participants drawn from multiple MOH facilities in the country. To ensure methodological rigor and reporting transparency, the study will be conducted in accordance with the CHERRIES (Checklist for Reporting Results of Internet E-Surveys (Checklist 1) and guided by the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist.

Study Timeline

The study will be conducted over a 24-month period. All stages, from initial recruitment to data analysis and reporting, will be carried out systematically within this timeframe. The timeline will be structured to allow sufficient time for comprehensive data collection, analysis, and interpretation. A detailed breakdown of the study phases is provided in Table 1.

Table . Study timeline.

Timeline	Activity
Months 1 - 3	• Recruitment of study team members • Planning of the full study
Months 4 - 6	• Pilot study commencement
Months 7 - 10	• Refinement of the questionnaire based on pilot test findings • Stakeholders' feedback
Months 11 - 13	• Nationwide data collection
Months 14 - 20	• Data analysis and synthesis
Months 21 - 24	• Result dissemination • Preparation of the final report

Study Population

The study population will comprise all 8953 MOH specialists actively listed in the Human Resource Management Information System 2.0 (HRMIS 2.0) as of December 2023. This time point will be selected to ensure the use of the most current and complete national dataset available prior to sampling and survey distribution. HRMIS 2.0 is the official centralized platform for managing civil service personnel records in Malaysia and is widely used for workforce planning, administrative monitoring, and strategic human resource deployment [22,23]. The system provides standardized information on employment status, job designation, and career trajectory and is accessible via a secure web interface [24]. Leveraging HRMIS 2.0 as the sampling frame will enable comprehensive, nationally representative coverage of the MOH specialist workforce, minimize selection bias, and enhance participant identification accuracy. This approach aligns with best practices in health workforce research and ensures that the sample reflects the diversity of specialist roles and geographic distribution within Malaysia's public health care system. Inclusion and exclusion criteria will subsequently be applied to refine the study sample in accordance with the research objectives.

Inclusion Criteria

The study will include all MOH specialists with full registration who are actively serving in MOH facilities, including hospitals, health clinics, district health offices, state health departments, specialized institutes, and the MOH headquarters. Inclusion

XSL•FO
RenderX

across a wide range of health care settings is intended to ensure comprehensive coverage of specialists.

Exclusion Criteria

Dental specialists, specialists who have not yet obtained full registration (eg, those under gazettement), contract-based specialists, and those on temporary leave will be excluded. Dental specialists are excluded due to differences in practice context, while the remaining criteria ensure the inclusion of individuals eligible for and directly covered by MOH WRSs.

Sample Size Calculation and Sampling Method

The sample size will be calculated based on a 95% CI (z=1.96), a 5% margin of error (d=0.05), and an anticipated turnover intention rate of 46.8% among MOH specialists, as reported in prior studies [12]. This results in a minimum required sample size of 382 participants. After adjusting for an anticipated 40% nonresponse rate, which is commonly observed in web-based surveys, the sample size will be increased to 637 participants to ensure an optimum-achieved sample size [25]. The sample size was inflated by 40% to ensure that the final number of participants who complete the study meets the minimum required sample size for the planned statistical analyses and power considerations.

Considering 2 main groups of specialists (surgical and nonsurgical), the adjusted sample size will be doubled, resulting in a total of 1274 participants (637×2). Based on national workforce distribution data from the Human Resource Division and Medical Development Division, MOH Malaysia, the proportions of nonsurgical and surgical specialists are 74% and 26%, respectively. Applying these proportions to the adjusted sample yields an estimated requirement of 943 nonsurgical specialists (1274×0.74) and 331 surgical specialists (1274×0.26). However, the estimated sample size for the surgical subgroup (n=331) falls below the minimum sample size requirement of 382 derived from the initial prevalence-based calculation; therefore, it is adjusted to 382. Consequently, the final target sample size for this study is 1325 participants.

Participant recruitment will be conducted using the list of specialists in active service obtained from the Human Resource Department, MOH, which is provided in alphabetical order. The HRMIS list will be randomly shuffled using Excel=RAND() before the systematic sampling interval is applied. Following this, a random starting point will be generated for each stratum, followed by a fixed sampling interval. Participants will be selected using systematic random sampling within each specialty stratum (surgical and nonsurgical). To achieve the required samples of 943 nonsurgical and 382 surgical specialists, sampling intervals will be set at 1 in every 7 nonsurgical specialists and 1 in every 6 surgical specialists, reflecting the relative size of each stratum.

Study Tool

Questionnaire

A structured questionnaire will be developed by adapting items from previously validated instruments, including items from Belbin et al [15] to assess retention strategies and the Job Satisfaction Survey (JSS) [26,27] and the Turnover Intention Scale-6 (TIS-6) [28] to measure job satisfaction and turnover intention, respectively. These adaptations aim to ensure conceptual alignment with the study objectives, support content relevance for the target population, and contextualize the instrument for use among MOH specialists in Malaysia. To enhance clarity, contextual appropriateness, and usability, the draft questionnaire will undergo face and content validation by key stakeholders and subject matter experts. Feedback from this process will inform refinements to item wording, logical sequencing, and response formatting. Internal consistency of multi-item domains will be assessed during the pilot phase to guide potential scale refinement. The finalized questionnaire will comprise 5 sections, as summarized in Table 2, and will be administered using the REDCap platform to ensure secure and efficient online distribution. It will include both closed-ended items (eg, Likert-scale and multiple-choice formats) and optional open-text fields to capture contextual insights, thereby supporting a mixed methods analysis approach.

XSL•FO
RenderX

Table . Questionnaire items and measures.

Survey and question number	Type	Responses
Section A: awareness, participation, and perceived effectiveness of 12 WRSs[a] (13 items)		
a. Awareness: Q 1(a) to 12(a)	Dichotomous scale	• Yes or no
b. Participation: Q 1(b) to 12(b)	Dichotomous scale	• Yes or no
c. Effectiveness Q 1(c) to 12(c)	Interval scale	• 1=effective; 2=slightly effective; 3=not effective; 4=don't know
Q 13	Open-ended questions	• Free text ("Any suggestion for other retention strategies that might be beneficial?")
Section B: individual and organizational WRS barriers (3 items)		
Q 1 to 2	Interval scale (5-point Likert scale)	• 1=strongly disagree; 2=disagree; 3=neutral; 4=agree; 5=strongly agree
Q 3	Open-ended questions	• Free text ("Any other barriers that could prevent you from participating in any retention strategies?")
Section C: Job Satisfaction Survey (JSS) (36 items classified into 9 domains, 4 items per domain)		
Q 1 to 36	Interval scale (6-point Likert scale)	• 1=strongly disagree; 2=moderately disagree; 3=slightly disagree; 4=slightly agree; 5=moderately agree; 6=strongly agree • For individual domains, scores of 4 - 12 indicate dissatisfaction, 13 - 16 reflect ambivalence, and 17 - 24 represent satisfaction. • Overall scores of 36 - 108 indicate dissatisfaction, 109 - 144 reflect ambivalence, and 145 - 216 indicate satisfaction
Section D: Turnover Intention (TIS-6) (6 items)		
D1, 3, 4, 6	Interval scale (5-item Likert scale)	• 1=never to 5=always
D2	Interval scale	• 1=to no extent to 5=very large extent
D5	Interval scale	• 1=highly unlikely to 5=highly likely • A total score below 18 reflects a desire to remain with the MOH[b], whereas a score above 18 indicates a desire to leave.
Section E: baseline characteristics (12 items)		
E 1 to 12	Category scale	• Age, gender, marital status, number of children, year of service in MOH, specialty and subspecialty, year of service as a specialist, grade, place of work, retirement scheme: Employees Provident Fund or pension, and the status of National Specialist Register

[a]WRSs: workforce retention strategies.

[b]MOH: Ministry of Health.

REDCap Survey Administration Workflow

REDCap will be implemented through a structured and systematic process that includes automated survey distribution, real-time response monitoring, built-in data validation rules, and secure data export procedures. These features aim to ensure accurate, consistent, and efficient data collection throughout all phases of the study. Table 3 shows each stage of the REDCap-based workflow, highlighting the specific tools used and their roles in the data collection process.

XSL•FO
RenderX

Table . Application of REDCap in a nationwide online cross-sectional study: key steps and implementation process.

Stage and steps	Description	Unique features	REDCap[a] function used
Study design and questionnaire			
Project creation in REDCap	A REDCap project will be initiated to provide a secure and structured system for national-level survey deployment targeting MOH[b] specialists.	Enables secure data storage and standardization across all survey responses.	Project setup and survey module activation
Questionnaire development	The instrument will be developed using the Online Designer, incorporating Likert-type and open-ended questions with embedded validation rules to prevent data entry errors.	Supports uniformity and improves data reliability.	Online Designer and Field Validation Rules
Branching logic and standardization	Branching logic will ensure that only relevant questions appear based on prior responses. For instance, respondents unaware of a particular WRS[c] will skip related questions. Required fields will also be enforced to ensure data completeness.	Minimizes respondent burden and improves data quality.	Branching Logic and Required Fields
Management of study participants			
Importing selected sample	A systematically selected sample of MOH specialists will be imported to restrict access to eligible participants only.	Preserves data integrity and study eligibility criteria	Data import tool
Survey distribution and data collection			
Automated email invitations	Personalized email invitations with unique, nontransferable survey links are used for secure survey access and management of invitations and reminders. Survey responses cannot be linked to email addresses.	Ensures controlled access, supports reminder scheduling, and maintains participant anonymity	REDCap automated survey invitation module and participant list (anonymous mode)
Survey invitation log	This log will be used to monitor the distribution status of survey invitations and reminders to support targeted follow-up.	Mitigate potential nonresponse bias by enabling targeted follow-up	Survey invitation log
Follow-up reminders	Automated reminders will be sent during the second and third weeks to nonrespondents. Reminder emails will cease upon survey completion.	Improves participation while avoiding redundancy.	Automated reminder scheduler
Real-time response monitoring and quality control			
Live response tracking	A real-time dashboard will monitor response rates across subgroups, such as specialty or setting (hospital vs nonhospital).	Enables dynamic monitoring and targeted outreach	Record status dashboard
Data validation and quality control	The system will prompt participants to complete all mandatory questions within each section before submission (1 section per page), with selected fields configured using predefined validation criteria.	Minimize incomplete and invalid data entries	Required and validation fields
Duplicate response detection	Duplicate responses are prevented by using a closed REDCap survey with unique, nontransferable invitation links, each allowing only a single completed submission per recipient.	Prevents duplicate submissions	Survey distribution tool and participant list
Survey accessibility and usability			

XSL•FO
RenderX

Stage and steps	Description	Unique features	REDCap[a] function used
Multidevice compatibility	The survey will be optimized for use on desktops, tablets, and mobile devices to support flexible participation across health care contexts.	Increases accessibility and response convenience	Built-in responsive design (automatically enabled for all surveys)
Data analysis and reporting			
Data export and automated reporting	Survey data will be exported in IBM SPSS Statistics compatible format. No identifiable variables will be included in the dataset used for analysis.	Facilitates secure and structured data analysis	Data export tool and automated report generation

[a]REDCap: Research Electronic Data Capture.

[b]MOH: Ministry of Health.

[c]WRS: workforce retention strategy.

Pilot Study

A pilot study will be conducted involving 133 MOH specialists, representing approximately 10% of the 1325 participants calculated for the main study. This sample size is deemed appropriate to evaluate multiple components of survey implementation, including questionnaire clarity, internal consistency of multi-item domains, and the operational feasibility of data collection procedures [29]. While smaller pilot samples (eg, 30 - 50 participants) may be sufficient for testing usability or basic functionality, a larger sample is warranted in this context to support broader assessments, including psychometric performance and system readiness [29].

The purpose of the pilot study is to assess the anticipated response rate, questionnaire usability, and internal consistency of multi-item domains, as well as to confirm the clarity and cultural appropriateness of survey items. The pilot study will be conducted via preproduction testing using in-built REDCap features to ensure the clarity of all questions and the proper functioning of relevant technical elements, such as branching logic, field validation, and device responsiveness (mobile/desktop). This will establish the usability of the questionnaire and subsequently minimize response bias and data loss. In addition, the operational functionality of the REDCap platform will be tested, including automated email distribution and real-time response capture, replicating the workflow planned for the full-scale national survey. Any necessary modifications will be incorporated prior to national rollout to enhance overall study quality.

Internal consistency metrics (eg, Cronbach α) will be calculated to examine the reliability of multi-item domains and their suitability to the Malaysian public health care and cultural context. An α coefficient equal to or greater than 0.7 will be considered acceptable [30]. Domains with suboptimal internal consistency will be reviewed and revised accordingly. All procedures used in the pilot will mirror those of the main study.

Data Collection

The survey will comprise 7 sequential pages: a participant information sheet (PIS), an informed consent form, and 5 questionnaire sections, as detailed in Table 2. Data will be collected through a structured, self-administered online survey using the REDCap platform, as detailed in Table 3.

Eligible MOH specialists will be invited to participate via email. To minimize selection bias, no prenotifications or prior announcements will be issued. Each invitation will include a cover letter outlining the study objectives and procedures, along with a unique, nontransferable hyperlink to access the survey. This individualized link will restrict access to the intended recipient and prevent duplicate or unauthorized responses.

Upon accessing the survey, participants will first review the PIS and provide informed consent using a nonidentifiable "yes/no" response option. Only participants who provide consent can proceed to the survey items. To ensure data completeness, all items will be mandatory and the survey will be administered in sections (1 section per page). Participants must complete and submit each section before proceeding to the next. Responses may be reviewed and amended prior to section submission; however, once a section is submitted, responses within that section cannot be modified. Participants who are unable to complete a section in 1 sitting may use the "Save & Return Later" function to resume at their convenience.

REDCap paradata (eg, Survey Invitation Log and Record Status Dashboard) will be used to track recruitment and support nonresponse assessment. Survey invitations will be issued through REDCap-generated emails. These will be routed through MOH institutional mail servers and delivered using a secure email infrastructure that adheres to institutional data governance protocols. Due to daily email limits imposed by the MOH system, invitations will be distributed in 3 consecutive monthly batches. Follow-up reminder emails will be sent to nonrespondents during the second and third weeks of each distribution month to optimize response rates.

To maximize participation and minimize nonresponse, several strategies will be implemented, including individualized email invitations, clear and concise study information, appropriate timing of survey distribution, and a concise, mobile-friendly questionnaire design to reduce respondent burden. Response patterns will be monitored throughout the survey period to identify subgroups with lower participation rates and to implement targeted follow-up, including additional reminders via REDCap where appropriate.

XSL•FO
RenderX

Data Management and Analysis

Only completed questionnaires will be included in the analysis. Incomplete surveys will be excluded from the analytic dataset and recorded as noncomplete submissions. The completion rate will be calculated using REDCap survey logs, where the number of completed surveys is divided by the number of surveys started. The partial response rate will be calculated as the number of noncomplete submissions divided by the number of surveys started. Response rate will be calculated as the number of submitted surveys divided by the number of unique invitations sent (calculated sample), based on REDCap invitation and survey status logs. Response and completion rates will be tabulated and reported upon completion of data collection in accordance with CHERRIES (Checklist 1).

Quantitative data will be exported from the REDCap platform to IBM SPSS Statistics Version 29 for analysis. Descriptive statistics will be used to examine the level of awareness, participation, and perceived effectiveness of WRSs, as well as job satisfaction and turnover intention. Binary logistic regression models will be utilized to examine the associations between WRSs and workforce outcomes, namely job satisfaction and turnover intention. Independent variables in this study would include WRS engagement (operationalized into 3 domains: awareness, participation, and perceived effectiveness). Multivariable logistic regression models will adjust for potential confounders, including age, gender, professional grade, disciplines (surgical vs nonsurgical), service duration, facility type, and geographic regions. No subgroup analyses will be conducted; all covariates will be included as adjustment variables in the regression models. Finite-population correction will not be applied, as the primary analyses focus on regression-based association estimates.

To minimize potential nonresponse bias, respondent characteristics will be descriptively compared with the sampling frame on key variables. No weighting or postsurvey statistical correction will be applied, as access to all the necessary variables in our entire study population (8953 specialists) is limited; hence, we would not be able to accurately calculate the necessary weights. Under these conditions, poststratification or raking weights may introduce additional bias and inflate variance without improving representativeness.

Qualitative data from 2 open-ended questions will be imported into NVivo version 12 (Lumivero) for organization and detailed analysis. NVivo will assist in systematically coding responses and grouping them into broader themes, facilitating the generation of meaningful insights. The analytical purpose of these questions is to identify context-specific strategies, implementation gaps, and participation barriers that may not be captured by quantitative questionnaires. Open-ended responses will be analyzed using deductive thematic analysis to complement the quantitative component of the study. The deductively derived qualitative themes will be cross-referenced with quantitative data to provide contextual and explanatory insights into the quantitative findings. The process will adhere to CHERRIES principles to ensure transparency and thorough examination.

Ethical Considerations

Ethics Approval

This study is registered with the National Medical Research Registry and has received ethical approval from the Medical Research and Ethics Committee, MOH Malaysia (23-03199-3VQ). All study procedures will be conducted in accordance with the Malaysian Guidelines for Good Clinical Practice [31] and the principles of the Declaration of Helsinki.

Consent to Participate

Informed consent will be obtained electronically prior to the survey participation. Participants will first be presented with PIS outlining the study objectives, procedures, potential risks and benefits, data handling arrangements, and assurances regarding data confidentiality. Consent will be recorded using a nonidentifiable “yes/no” response option embedded within the survey, and only participants who select “yes” will be able to proceed to the questionnaire. Participation is entirely voluntary, and participants may self-withdraw from the study at any time before the full survey submission without any consequences on their employment status, performance evaluations, or access to workplace benefits. Once the questionnaire has been submitted, withdrawal of data will not be possible, as no personal identifiers or traceable information are collected or linked to survey responses, rendering individual responses unidentifiable. This limitation will be explicitly stated in the PIS to ensure fully informed consent. No compensation will be given to participants.

Anonymization and Data Privacy

Survey invitations will be distributed using email addresses uploaded to the REDCap participant list for recruitment purposes only. Email addresses will be managed within the participant list for invitation and reminder purposes and will not be stored with, nor linked to, survey response data in the dataset exported for analysis.

No personal identifiers will be collected in the questionnaire. The Participant Identifier function will be disabled and Survey Response Status will be automatically set to “Anonymous,” such that responses are not directly identifiable to users within the project interface. As survey responses are anonymous, we will not be able to entertain any request for individual data withdrawal after submission of the completed questionnaire.

Data Security, Access, and Repository

All data will be handled confidentially and stored on secure National Institutes of Health Data Repository System servers for a minimum period of 3 years following study completion. Data governance adheres to the MOH Standard Operating Procedure on Maintenance, Archival and Disposal of Study and Non-Study Files, as required under Medical Research and Ethics Committee approval [32]. Upon completion of the mandated retention period, and where the data are no longer required for regulatory, audit, or research purposes, study records will be securely disposed of in accordance with approved MOH data destruction procedures. Access to the dataset will be restricted to authorized members of the study team and used solely for

XSL•FO
RenderX

the purposes of this study. As such, the data are not publicly available.

Results

The results will present descriptive summaries of participants' sociodemographic characteristics, employment profiles, and levels of awareness, participation, and perceived effectiveness of the 12 MOH-implemented WRSs. Findings from multivariable regression analyses will report associations between WRS engagement and workforce outcomes, including job satisfaction and turnover intention. Findings from the open-ended survey items will be reported narratively, highlighting common themes related to perceived facilitators, barriers, and recommendations for improving WRS implementation, to contextualize the quantitative results. A statistical significance will be set at $P<.05$. Results will be reported in accordance with the STROBE guidelines.

As of the submission of this protocol, participant recruitment and data collection have been completed. The study is currently in the data analysis phase. It is expected that the data analysis will be completed in March 2026 and that the results will be published in June 2026.

Discussion

Anticipated Findings and Policy Implications

This study protocol is important as it systematically outlines a national cross-sectional study to examine the implementation of the 12 WRSs across MOH health care facilities, with the aim of assessing associations, if any, between WRS engagement, job satisfaction, and turnover intention among MOH specialists. By capturing national-level data on awareness, participation, and perceived effectiveness, the study can provide a nuanced view of how WRSs are experienced across diverse clinical settings within the MOH system. This research represents a foundational step toward evidence-informed policymaking aimed at strengthening specialist retention and supporting the human resource sustainability of Malaysia's public health care system.

To our knowledge, this is the first national cross-sectional study focused on specialist engagement with a wide range of MOH-implemented WRSs, encompassing both centrally administered initiatives and those requiring self-initiation or managerial discretion. The study will provide contextual factors, such as institutional readiness, administrative support, and resource availability, which are associated with variations in uptake and perceived effectiveness. By analyzing the variability of implementation across MOH health care settings, the study will offer key insights into how current strategies align with the needs and expectations of specialists and how these strategies can be further refined for equitable and effective applications.

This protocol sets the stage for generating practical evidence that can inform policy revisions and the development of more targeted, context-sensitive retention strategies. Understanding the lived experiences of specialists with respect to WRSs will help identify which elements foster engagement, enhance job satisfaction, and ultimately reduce turnover. The findings from this study may also be used to guide the development of strategic frameworks and best practice models that can be scaled or adapted across different facility types and regions within MOH Malaysia. Moreover, capturing specialists' perspectives will enhance the real-world relevance and applicability of the study's recommendations.

Empowering specialists with well-structured, responsive, and transparent retention strategies is essential for sustaining motivation and workforce stability [14,33]. Insights from this study will inform institutional policies and operational improvements that contribute to a more supportive working environment. By identifying which strategies resonate most with specialists and why, this study has the potential to promote both individual well-being and systemwide efficiency. The potential benefits extend beyond workforce outcomes, including improvements in service continuity, patient care quality, and long-term cost savings through reduced recruitment and training burdens.

Finally, this study will highlight critical gaps in the implementation of WRSs, guiding future research and policy dialogue. It may also generate insights applicable to other health professional groups facing similar retention challenges. By advancing the evidence base for workforce retention in Malaysia's public health system, this study offers a strategic opportunity to enhance human resource planning, specialist well-being, and health system resilience in the years to come.

Integration of REDCap for National-Level Data Collection

This protocol strategically integrates REDCap to facilitate standardized, large-scale data collection from specialists across MOH Malaysia's public health care system. Traditional survey methods, such as paper-based questionnaires or in-person interviews, often pose significant logistical and operational challenges, particularly in geographically dispersed or high-volume clinical settings [21]. These challenges include manual data entry errors, time delays, data loss risks, and the resource-intensive nature of deploying field staff [21]. In contrast, REDCap offers a secure, browser-based solution that enables centralized administration, efficient data capture, and real-time monitoring while maintaining strict data privacy and security standards [34].

The use of REDCap in this study enhances methodological rigor by enabling structured data entry through built-in features, such as branching logic, validation rules, and audit trails [35,36]. These features reduce data entry errors, ensure internal consistency, and support traceability throughout the research process. Compared to general-purpose platforms, such as Google Forms or SurveyMonkey, REDCap offers distinct advantages in data management, including compliance safeguards, user access controls, and robust metadata tracking [37]. These capabilities are especially valuable in health research, where maintaining data integrity and participant confidentiality is essential.

Moreover, REDCap is designed for interoperability with statistical and qualitative analysis software such as SPSS, STATA (StataCorp LLC), and NVivo, allowing for seamless

XSL•FO
RenderX

data export and minimizing the need for manual data handling [20]. This streamlines the data cleaning and analysis workflow and ensures the production of high-quality datasets suitable for both quantitative and qualitative analyses. Its widespread adoption in international health research, including in low-resource and multisite studies, underscores REDCap's scalability and adaptability in diverse research settings [21].

Incorporating REDCap also supports uniform implementation across all participating sites, fostering consistency and transparency in data collection procedures [21]. This is particularly crucial for a national-level study involving multiple facilities with varied institutional capacities. REDCap's deployment is expected to enhance study replicability, improve coordination across research teams, and promote data governance standards aligned with international best practices. Prior evidence has shown that REDCap is well-received by clinicians and researchers alike, with studies reporting high levels of satisfaction related to its reliability, user-friendliness, and capacity to produce complete, accurate datasets [34]. As such, the integration of REDCap is a key enabler of this study's operational feasibility and scientific integrity.

Beyond technical capabilities, ensuring transparency and ethical accountability in digital survey implementation is also a critical element of methodological integrity. This protocol emphasizes the study's adherence to CHERRIES, a standardized reporting framework designed to enhance the transparency, reproducibility, and ethical standards of online research. CHERRIES' compliance ensures the clear reporting of key survey elements, such as consent procedures, response rates, and data handling protocols, while facilitating cross-study comparisons and future meta-analyses [38]. In the context of this nationwide web-based survey, CHERRIES' alignment reinforces the credibility of study findings and contributes to the broader goal of generating ethically sound and scientifically rigorous evidence within the evolving digital health research ecosystem.

Application of Validated Instruments to Support Methodological Rigor in Online Research

To ensure robust and credible measurement of workforce retention dynamics, this study employs a validated, multidimensional set of instruments to examine associations between MOH-implemented WRS, job satisfaction, and turnover intention among specialists in MOH Malaysia's public health care system. Specifically, the study integrates 3 well-established instruments: adapted items from Belbin et al [15] to assess WRS perceptions, the TIS-6, and the JSS. The combined use of these instruments provides a methodologically sound framework to examine relationships between organizational strategies and individual-level outcomes, consistent with best practices in health workforce research [15,26,28].

The TIS-6 is a brief, psychometrically validated scale widely used in health workforce research to assess an individual's intention to leave their current role. It captures both affective and cognitive components of turnover intention, offering insight into short- and longer-term attrition risks [39]. Its brevity and clarity make it particularly suitable for clinician populations, who often face time constraints in research participation. Previous validation studies have confirmed the tool's internal consistency (Cronbach α=0.80) and predictive validity, further supporting its applicability in large-scale health care settings [28].

To measure job satisfaction, this study adopts the JSS, a comprehensive instrument developed to assess satisfaction across 9 key domains of 4 items each, including pay, promotion, supervision, fringe benefits, contingent rewards, operating procedures, coworkers, nature of work, and communication [26,27]. The JSS has demonstrated robust psychometric properties in diverse health systems and cultural settings, including countries such as Malaysia, Jordan, Singapore, Iran, Turkey, and the United States [13,40]. Its multidimensional structure allows for a granular understanding of the factors contributing to professional fulfillment or dissatisfaction among specialists. A systematic review conducted in 2003 reaffirmed its internal consistency and construct validity, underscoring its continued relevance for cross-cultural health care research [41].

Limitations

Several limitations must be acknowledged. First, the reliance on self-reported data may introduce response or recall bias, although anonymity, neutral phrasing, and piloting are expected to mitigate these effects. Second, nonresponse bias remains a concern, particularly among specialists facing high clinical demands. Furthermore, it may be further complicated by the use of mandatory survey items and the exclusion of partially completed questionnaires. To mitigate this risk, practical steps were incorporated to encourage participation, and potential nonresponse bias will be assessed through the descriptive comparison of respondent characteristics with those available from the sampling frame on key variables. Finally, although open-ended items allow for limited qualitative exploration, they may not capture the same depth achievable through interviews or focus group discussions.

Future Directions

This study protocol establishes a foundation for ongoing, data-driven monitoring of WRSs within the MOH health care system. Future research may adopt a mixed methods design incorporating more intensive qualitative methods to deepen insights into WRS experiences. Longitudinal designs are also warranted to assess the long-term impact of evolving WRS policies on specialist retention. Comparative studies between public and private sector settings may also yield valuable insights into contextual differences in WRS effectiveness. In addition, integrating REDCap dashboards into MOH human resource planning systems could support real-time surveillance and predictive modeling of retention risks, enabling more targeted and evidence-informed decision-making.

Conclusion

This study protocol presents a comprehensive framework for the national-level cross-sectional study of WRS targeting specialists within the MOH, Malaysia. By examining the association between WRS implementation, job satisfaction, and turnover intentions, the study seeks to align WRSs more closely with the realities of clinical practice and workforce expectations. The integration of REDCap ensures efficient, secure, and

XSL•FO
RenderX

standardized data collection, while adherence to CHERRIES guidelines strengthens methodological transparency and ethical rigor. By uncovering context-specific barriers and enablers, the study offers practical insights into the implementation and effectiveness of existing WRSs, providing a foundation for more responsive and targeted policy interventions to address specialist attrition. In doing so, it supports evidence-informed decision-making among stakeholders in MOH Malaysia and contributes to broader efforts to mitigate brain drain. Ultimately, this protocol advances the long-term goal of building a resilient, motivated, and sustainable health workforce, essential for the continued strengthening of Malaysia's public health care system.

Acknowledgments

The authors would like to thank the Director General of Health, Malaysia, for the permission to publish this paper.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or nonprofit sectors.

Authors' Contributions

All authors actively contributed to the design of the study protocol, provided valuable feedback on draft versions of the manuscript, and gave their approval for the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of workforce retention strategies for physicians and dentists in the Ministry of Health Malaysia.
[DOCX File, 18 KB - resprot_v15i1e83377_app1.docx]

Checklist 1
CHERRIES checklist.
[DOCX File, 35 KB - resprot_v15i1e83377_app2.docx]

References

1. De Vries N, Lavreysen O, Boone A, et al. Retaining healthcare workers: a systematic review of strategies for sustaining power in the workplace. Healthcare (Basel) 2023 Jun 29;11(13):1887. [doi: 10.3390/healthcare11131887] [Medline: 37444721]
2. de Vries N, Boone A, Godderis L, et al. The race to retain healthcare workers: a systematic review on factors that impact retention of nurses and physicians in hospitals. Inquiry 2023;60:469580231159318. [doi: 10.1177/00469580231159318] [Medline: 36912131]
3. Misra-Hebert AD, Kay R, Stoller JK. A review of physician turnover: rates, causes, and consequences. Am J Med Qual 2004;19(2):56-66. [doi: 10.1177/106286060401900203] [Medline: 15115276]
4. Young R, Noble J, Mahon A, Maxted M, Grant J, Sibbald B. Evaluation of international recruitment of health professionals in England. J Health Serv Res Policy 2010 Oct;15(4):195-203. [doi: 10.1258/jhsrp.2010.009068] [Medline: 20410188]
5. Rolewicz L, Hutchings R, Palmer W, De Los Santos JAA. Stemming the tide: the challenges that must be overcome to reverse the great resignation of healthcare staff. BMJ 2024 Jan 12;384:q81. [doi: 10.1136/bmj.q81] [Medline: 38216209]
6. Phanthunane P, Suwatmakin A, Konglumpun N, Pannarunothai S. Job incentives influencing health professionals working in rural and remote areas in Thailand: finding from discrete choice experiment to policy recommendation. BMC Health Serv Res 2025 Feb 14;25(1):254. [doi: 10.1186/s12913-025-12408-2] [Medline: 39953532]
7. da Silva EN, Ramos MC, Santos W, Rasella D, Oliveira A, Santos LMP. Cost of providing doctors in remote and vulnerable areas: Programa Mais Médicos in Brazil. Rev Panam Salud Publica 2018;42:e11. [doi: 10.26633/RPSP.2018.11] [Medline: 31093040]
8. Makuku R, Mosadeghrad AM. Health workforce retention in low-income settings: an application of the Root Stem Model. J Public Health Policy 2022 Sep;43(3):445-455. [doi: 10.1057/s41271-022-00361-x] [Medline: 35978036]
9. Jinah N, Adnan IK, Bakit P, Sharin IA, Lee KY. Retention strategies for medical doctors in low- and middle-income countries (LMICs): are they effective? A scoping review. BMC Health Serv Res 2024 Dec 30;24(1):1662. [doi: 10.1186/s12913-024-12154-x] [Medline: 39734187]
10. Buykx P, Humphreys J, Wakerman J, Pashen D. Systematic review of effective retention incentives for health workers in rural and remote areas: towards evidence-based policy. Aust J Rural Health 2010 Jun;18(3):102-109. [doi: 10.1111/j.1440-1584.2010.01139.x] [Medline: 20579020]

XSL•FO
RenderX

11. Boone A, Lavreysen O, De Vries N, et al. Retaining healing hands: a transnational study on job retention interventions for the healthcare workforce. Qual Health Res 2024 Nov;34(13):1351-1366. [doi: 10.1177/10497323241254253] [Medline: 38857417]
12. Amir Muhammad Nur AR, Sharifa Ezat WP. Physicians intention to leave from malaysia government hospitals with existing retention strategy. J Public Health Policy Plan 2020;4(3). [doi: 10.35841/public-health-policy.4.3.30-37]
13. Daud F, Abd Ghani NF, Abu Zahid SN. Job satisfaction among specialist in Ministry of Health Malaysia and its associated factors. Malays J Public Health Med 2022 Dec 21;22(3):1-8. [doi: 10.37268/mjphm/vol.22/no.3/art.1177]
14. Willis-Shattuck M, Bidwell P, Thomas S, Wyness L, Blaauw D, Ditlopo P. Motivation and retention of health workers in developing countries: a systematic review. BMC Health Serv Res 2008 Dec;8(1). [doi: 10.1186/1472-6963-8-247]
15. Belbin C, Erwee R, Wiesner R. Employee perceptions of workforce retention strategies in a health system. J Manag Organ 2012 Sep;18(5):742-760. [doi: 10.5172/jmo.2012.18.5.742]
16. Koncar P, Santos T, Strohmaier M, Helic D. On the application of the two-factor theory to online employer reviews. J Data Inf Manag 2022 Mar;4(1):1-23. [doi: 10.1007/s42488-021-00061-3]
17. Ozsoy E. An empirical test of Herzberg's Two-Factor Motivation Theory. Mark Manag Innov 2019;1:11-20. [doi: 10.21272/mmi.2019.1-01]
18. Cosgrave C. The whole-of-person retention improvement framework: a guide for addressing health workforce challenges in the rural context. Int J Environ Res Public Health 2020 Apr 14;17(8):2698. [doi: 10.3390/ijerph17082698] [Medline: 32295246]
19. Van Bulck L, Wampers M, Moons P. Research Electronic Data Capture (REDCap): tackling data collection, management, storage, and privacy challenges. Eur J Cardiovasc Nurs 2022 Jan 11;21(1):85-91. [doi: 10.1093/eurjcn/zvab104] [Medline: 34741600]
20. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
21. Le HT, Vu DVT, Nguyen TNA, et al. Research Electronic Data Capture (REDCap) for population-based data collection in low- and middle-income countries: opportunities, challenges, and solutions. J Med Internet Res 2025 Jun 11;27:e65377. [doi: 10.2196/65377] [Medline: 40499167]
22. Zahari ASM, Harun MA, Hamzah SFM, Salleh SM. The influence of human resource management information system (HRMIS) application towards employees efficiency and satisfaction. Presented at: 1st International Conference on Green and Sustainable Computing (ICoGeS 2017); Nov 25-27, 2017. [doi: 10.1088/1742-6596/1019/1/012077]
23. Sinniah S, Kumar R, Mohamed MH, et al. User satisfaction of human resource management system in rural area. Int J Eng Adv Technol ;8(5C):73-79. [doi: 10.35940/ijeat.E1010.0585C19]
24. Shahibi MS, Saidin A, Izhar TAT. Evaluating user satisfaction on human resource management information system (HRMIS): a case of Kuala Lumpur City Hall, Malaysia. Int J Acad Res Bus Soc Sci 2016;6(10):95-116. [doi: 10.6007/IJARBSS/v6-i10/2338]
25. Wu MJ, Zhao K, Fils-Aime F. Response rates of online surveys in published research: a meta-analysis. Comput Human Behav Rep 2022 Aug;7:100206. [doi: 10.1016/j.chbr.2022.100206]
26. Spector PE. Measurement of human service staff satisfaction: development of the Job Satisfaction Survey. Am J Community Psychol 1985 Dec;13(6):693-713. [doi: 10.1007/BF00929796] [Medline: 4083275]
27. Batura N, Skordis-Worrall J, Thapa R, Basnyat R, Morrison J. Is the Job Satisfaction Survey a good tool to measure job satisfaction amongst health workers in Nepal? Results of a validation analysis. BMC Health Serv Res 2016 Jul 27;16:308. [doi: 10.1186/s12913-016-1558-4] [Medline: 27461030]
28. Bothma CFC, Roodt G. The validation of the turnover intention scale. SA J Hum Resour Manag 2013;11(1). [doi: 10.4102/sajhrm.v11i1.507]
29. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health 2008 Apr;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]
30. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997 Feb 22;314(7080):572. [doi: 10.1136/bmj.314.7080.572] [Medline: 9055718]
31. Malaysian guideline for good clinical practice (fourth edition). : National Pharmaceutical Regulatory Agency (NPRA), Ministry of Health Malaysia; 2018 URL: https://tncpi.upm.edu.my/upload/dokumen/20180521171015MalaysianGuidelineforGoodClinicalPractice-4TH_EDITION.pdf [accessed 2026-03-23]
32. Yee LK. Standard operating procedure: maintenance, archival and disposal of study and non-study files (SOP 5-1), version 2.0. : Medical Research & Ethics Committee (MREC), Ministry of Health Malaysia; 2018 URL: https://nmrr.gov.my/documents?type=others [accessed 2026-03-10]
33. Karaferis D, Aletras V, Raikou M, Niakas D. Factors influencing motivation and work engagement of healthcare professionals. Mater Sociomed 2022 Sep;34(3):216-224. [doi: 10.5455/msm.2022.34.216-224] [Medline: 36310751]
34. Soni H, Ivanova J, Wilczewski H, et al. User preferences and needs for health data collection using Research Electronic Data Capture: survey study. JMIR Med Inform 2024 Jun 25;12:e49785. [doi: 10.2196/49785] [Medline: 38917448]

XSL•FO
RenderX

35. Garcia KKS, Abrahão AA. Research development using REDCap software. Healthc Inform Res 2021 Oct;27(4):341-349. [doi: 10.4258/hir.2021.27.4.341] [Medline: 34788915]
36. Wright A. REDCap: a tool for the electronic capture of research data. J Electron Resour Med Libr 2016 Oct;13(4):197-201. [doi: 10.1080/15424065.2016.1259026]
37. Patridge EF, Bardyn TP. Research Electronic Data Capture (REDCap). jmla 2018;106(1):142-144. [doi: 10.5195/jmla.2018.319]
38. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
39. Ike OO, Ugwu LE, Enwereuzor IK, Eze IC, Omeje O, Okonkwo E. Expanded-multidimensional turnover intentions: scale development and validation. BMC Psychol 2023 Sep 11;11(1):271. [doi: 10.1186/s40359-023-01303-2] [Medline: 37697427]
40. AkbariTabar AA, Nazifi M, Rahi A, MirKamandar E, HoseinPori M. Psychometric properties of Spector's job satisfaction survey in the Iranian population [Article in Persian]. Koomesh 2013 Aug;14(3):335-341 [FREE Full text]
41. van Saane N, Sluiter JK, Verbeek J, Frings-Dresen MHW. Reliability and validity of instruments measuring job satisfaction--a systematic review. Occup Med (Lond) 2003 May;53(3):191-200. [doi: 10.1093/occmed/kqg038] [Medline: 12724553]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
HRMIS: Human Resource Management Information System
JSS: Job Satisfaction Survey
MOH: Ministry of Health
PIS: participant information sheet
REDCap: Research Electronic Data Capture
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TIS-6: Turnover Intention Scale-6
WRS: workforce retention strategy

Edited by J Sarvestan; submitted 02.Sep.2025; peer-reviewed by G Rezniczek; revised version received 21.Jan.2026; accepted 22.Jan.2026; published 01.Apr.2026.

Please cite as:
Ahmad Subki SZ, Abdullah Sharin I, Zakaria NH, Zakaria N, Bakit P, Adnan IK, Jinah N, Lee KY
Specialists' Perceptions of Workforce Retention Strategies in the Malaysian Ministry of Health and Their Association With Job Satisfaction and Turnover Intention: Protocol for a REDCap-Based National Cross-Sectional Survey
JMIR Res Protoc 2026;15:e83377
URL: https://www.researchprotocols.org/2026/1/e83377
doi:10.2196/83377

© Siti Zubaidah Ahmad Subki, Ili Abdullah Sharin, Nor Haniza Zakaria, Nursyahda Zakaria, Pangie Bakit, Izzuan Khirman Adnan, Norehan Jinah, Kun Yun Lee. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 1.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing the Effectiveness of an Awareness Program on Knowledge and Practice Regarding the Mothers' Absolute Affection Program Among Postnatal Mothers in Selected Hospitals: Protocol for a Quasi-Experimental Study

Vaishnavi Dongare, MSc; Kavita Gomase, PhD; Vaishali Taksande, PhD

Smt Radhikabai Meghe Memorial College of Nursing, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Corresponding Author:
Vaishnavi Dongare, MSc
Smt Radhikabai Meghe Memorial College of Nursing, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Abstract

Background: The Mothers' Absolute Affection (MAA) program is a national initiative launched by the Ministry of Health and Family Welfare, Government of India, on August 5, 2016, to promote, support, and protect breastfeeding across the country. The program aims to provide focused breastfeeding promotion and counseling services through the health system. It targets approximately 39 million pregnant and nursing mothers, 880,000 Accredited Social Health Activists, 150,000 subcenters, and 17,000 delivery points across all states and union territories. Despite its wide reach, there remains a lack of valid and reliable tools to assess mothers' knowledge and practices related to breastfeeding and the MAA program.

Objective: The aims of this study are to assess the existing knowledge and practice regarding the MAA program among postnatal mothers, to evaluate the effectiveness of an awareness program on knowledge and practice regarding the MAA program, to correlate knowledge and practice regarding the MAA program among postnatal mothers, and to associate knowledge and practice with selected demographic variables.

Methods: A quasi-experimental one-group pretest and posttest design was used to evaluate the effectiveness of an awareness program for the MAA program among postnatal mothers. The study was conducted at a selected hospital with a sample of 80 postnatal mothers selected through inclusion and exclusion criteria. On Day 1, participants underwent a pretest followed by an awareness session on the MAA program. A posttest was administered on Day 2. Knowledge was assessed using a structured questionnaire consisting of 15 items, while practice was evaluated using a 5-point Likert scale. Data were analyzed using descriptive and inferential statistics, including paired *t* tests and χ^2 tests.

Results: Knowledge scores were classified into three levels—poor (0% - 33.33%), fair (33.34% - 66.66%), and good (66.67% - 100%)—based on the number of correct responses out of 15. Practice was categorized as poor (0% - 50%) or good (51% - 100%) based on the 5-point scale score. The study revealed significant improvements in both knowledge and practice following the awareness intervention, indicating that the program was effective. The scoring procedure provided a comprehensive understanding of the awareness and practical implementation of the MAA program among postnatal mothers.

Conclusions: The awareness program significantly improved the knowledge and practices of postnatal mothers regarding the MAA program, emphasizing the value of educational interventions in promoting breastfeeding.

Trial Registration: Clinical Trials Registry India CTRI/2025/03/081926; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=MTIxMjMy&Enc=&userName=

International Registered Report Identifier (IRRID): RR1-10.2196/73496

(JMIR Res Protoc 2026;15:e73496) doi:10.2196/73496

KEYWORDS

maternal health; health education; infant feeding practices; breastfeeding promotion; behavioral change communication

XSL•FO
RenderX

Introduction

Background

The Ministry of Health and Family Welfare of India developed an enhanced initiative that aims to add to ongoing health system initiatives by placing a clear emphasis on breastfeeding promotion. The program is known as the Mothers' Absolute Affection (MAA) program. After its national launch in August 2016, states and union territories were able to implement it [1]. The MAA program is a national initiative that aims to bring unwavering attention to breastfeeding promotion and offer counseling services to support breastfeeding through health systems [2].

Breastfeeding is a crucial intervention for the survival of children. Thirty percent of newborn deaths may be avoided if breastfeeding begins within an hour of delivery. Children who are not breastfed have a 15-fold increased risk of pneumonia and a 10-fold increased risk of diarrhea compared to those who receive only breast milk. Efforts to instruct mothers on the most effective breastfeeding techniques must be intensified in light of the overwhelming evidence that breastfeeding reduces infant and neonatal mortality. To enhance feeding for infants and young children, Tamil Nadu has thus been implementing the MAA program since August 2016. Breastfeeding is the closest thing to a miracle cure for child survival [3]. To increase breastfeeding rates, the MAA program seeks to reinvigorate efforts to promote, protect, and support breastfeeding methods through health systems [4]. The program aims to encourage the best breastfeeding practices and create an environment that is supportive of breastfeeding by focusing on awareness-raising efforts for family members, society, and expectant and nursing mothers. The program aims to promote breastfeeding as a crucial intervention for the survival and development of young children [5], strengthen the lactation support services offered by qualified community health workers and health care professionals at public health facilities, and reward and acknowledge health care facilities with high breastfeeding rates and lactation control procedures in place [6]. Breastfeeding has many benefits for the baby: early skin-to-skin contact keeps the infant warm, it facilitates the early production of breast milk, and colostrum (also known as the "first milk") helps shield a newborn from illness. In addition, it aids in the development of a close and loving relationship between mother and child, reduces the chances of infants contracting diseases like pneumonia or ear and throat infections, is associated with increased IQ, and guarantees growth and development. There are also benefits for the mother as breastfeeding facilitates placental expulsion and womb contraction; reduces the likelihood of severe bleeding following delivery; lowers the chance of developing ovarian, uterine, and breast cancer; diminishes osteoporosis; increases birth spacing due to lactational amenorrhea; and encourages weight loss following childbirth. In addition, breastfeeding is less expensive than feeding with formula [7].

Study Aim, Objectives, and Hypotheses

The primary aim of this study is to assess the effectiveness of an awareness program on the knowledge and practice regarding the MAA program among postnatal mothers in selected hospitals. Specifically, the study seeks to evaluate the existing knowledge and practice levels of postnatal mothers regarding the MAA program, measure the impact of the awareness intervention on improving these parameters, explore the correlation between knowledge and practice, and determine associations with selected demographic variables. To test these objectives, the study proposes several hypotheses. $H0_1$ states that there is no significant difference between preintervention and postintervention knowledge and practice regarding the MAA program, while $H1_1$ posits a significant difference between preintervention and postintervention knowledge and practice. Similarly, $H1_2$ suggests a significant correlation between knowledge and practice, whereas $H0_2$ proposes no significant correlation between knowledge and practice. This integrated approach is intended to provide a comprehensive understanding of how structured educational interventions can influence maternal behavior and support breastfeeding promotion through the MAA program.

Methods

Study Overview

This quasi-experimental study will be conducted in selected hospitals. The first part of the study (Day 1) involves a researcher providing an explanation of the MAA program to the mother. The posttest will be administered on Day 2. Mothers who participated in this study demonstrated improved knowledge as a result of the awareness program. This study will address the existing gap in MAA-related research by evaluating the effectiveness of structured awareness programs on improving knowledge and practices regarding the MAA program among postnatal mothers, a component that has been underexplored in previous studies [8].

Criteria for Sample Selection

Inclusion criteria for the study were as follows:

- Postnatal mothers who were present at the time of data collection.
- Postnatal mothers who were willing to participate in the study.

Exclusion criteria were as follows:

- Postnatal mothers who are health workers.
- Postnatal mothers who already attended an awareness program for the MAA program.
- Postnatal mothers who had been diagnosed with a postnatal psychological condition.

Sample Calculation

The formula comparing paired before and after proportions is presented here:

Primary variable: knowledge regarding breastfeeding practices = 18%

Knowledge regarding breastfeeding practices before the intervention = 18% = 0.18 (as per reference article [9])

Knowledge regarding breastfeeding practices after the intervention = 38% = 0.38 (20% higher)

XSL•FO
RenderX

Z (table value at α=.05) = 1.96

Z (table value at 1-β=.80) = 0.84

N = 78 participants (rounded to 80 to account for attrition)

Total sample required = 78

Actual sample size = 80

Intervention

On the first day, postnatal mothers received a pretest and awareness program about the MAA program. On the second day, the postnatal mother received a posttest to evaluate the awareness program's efficacy.

The impact of the awareness-raising initiative on MAA program knowledge will be evaluated in a forthcoming study. Knowledge regarding the MAA program was categorized into one of three levels based on the number of correct answers. A total of 15 questions were included. The practice level related to the MAA program was assessed on a 5-point scale. The scoring aimed to capture the frequency and consistency with which mothers followed the key components of the MAA program, thus providing insight into both awareness and real-world application among the participants.

The research results will be released in a journal with indexing. The study began in February 2024.

Statistical Analysis

The data collected for the study were analyzed using Microsoft Excel (version 10; Microsoft Corp). Initially, descriptive statistics such as frequencies, percentages, means ($\bar{X}$), and standard deviations (SD) were used to summarize participants' knowledge and practice scores regarding the MAA program. Inferential statistical tests were then applied to assess the effectiveness of the awareness program. A paired t test was conducted to compare pretest and posttest scores within the same group. Additionally, χ^2 tests were used to examine the association between participants' demographic variables and their knowledge and practice levels. This comprehensive approach allowed for an in-depth understanding of the impact of the intervention.

Reliability of the Tool

Reliability indicates the degree of consistency, and it is major criteria for assessing the quality and accuracy of tools. In this study, reliability was tested using the Guttman split-half method (knowledge: reliability=0.8948, r=0.913; practice: reliability=0.837, r=0.916).

Validity

To determine content validity, the tool was given to 10 experts within the Obstetrics and Gynecology Nursing Department, Smt. Radhikabai Meghe Memorial College of Nursing. Following consultation with the experts, some modifications were made in the framing (ie, wording) of questionnaire items and the same modifications were included in the study.

Ethical Considerations

The study protocol was reviewed and approved by the institutional ethics committee of Smt. Radhikabai Meghe Memorial College of Nursing, Datta Meghe Institute of Higher Education and Research (deemed-to-be-university), Wardha, Maharashtra, India (approval number ECR/440/Inst/MH/2013/RR-2019; February 29, 2024). A framework for participant consent will be provided for the research, which is based on the hospital. Written informed consent was obtained from all postnatal mothers prior to participation in the study. Participation was entirely voluntary, and participants were informed of their right to withdraw from the study at any time without any effect on the care they received. To ensure participant privacy and confidentiality, no personally identifiable information was collected. All data were deidentified and coded prior to analysis. The collected data were stored securely and were accessible only to the research team. The findings are reported in aggregate form to prevent individual identification. No monetary or nonmonetary compensation was provided to participants for their involvement in the study.

Results

The study was conducted between February 2024 and April 2024. Data analysis was completed in May 2024. The final manuscript was submitted for publication in March 2025.

The study showed that, after an awareness intervention, postpartum mothers' knowledge and practice levels of the MAA program significantly improved. A systematic questionnaire consisting of 15 items was used to evaluate knowledge, and the number of correct answers was used to classify the responses as poor (0% - 33.33%), fair (33.34% - 66.66%), or good (66.67% - 100%). Before the intervention, a sizable portion of participants had poor to fair knowledge. However, the number of participants receiving high knowledge ratings significantly increased following the awareness session. Similarly, practice was divided into two categories—poor (0% - 50%) and good (51% - 100%)—based on a 5-point Likert scale. Better adherence to best practices for breastfeeding was indicated by a significant increase in good practice scores in the postintervention data. Chi-square tests demonstrated correlations between demographic factors and results, while paired t tests verified that these changes were statistically significant. These results demonstrate how a well-structured educational program can work to raise awareness and improve the practical application of breastfeeding promotion techniques under the MAA program.

Discussion

Principal Findings

A similar study conducted in a rural setting found that an educational intervention focused on maternal care significantly increased knowledge among postnatal mothers. The results showed that after the intervention, 75% of mothers had increased knowledge, which aligns closely with the findings of another study [8]. Another study evaluated the effectiveness of a breastfeeding program among postnatal mothers and found significant improvements in both knowledge and practice [10]. The mean knowledge score in their posttest increased from 6.5 (SD 2.0) to 11.2 (SD 1.5), and the practice scores improved by 30%, which is consistent with the findings of this study on the

XSL•FO
RenderX

effectiveness of the awareness intervention program regarding the MAA program [11].

Hospitals should encourage breastfeeding because a mother's milk supply may be insufficient if she does not nurse her infant during the initial days following birth. Furthermore, many mothers stop nursing if they believe their milk supply is insufficient. This leads to them adding something other than breast milk to an infant's diet, which may expose the young child to diseases. One limitation of this study is that the promotion of a breastfeeding program at the district hospital level does not reach babies born in other clinics or private hospitals. However, it is realistic to say that, with the limited resources at hand, creating and implementing a policy to support breastfeeding at every private hospital would not be possible. In addition to awareness and knowledge, breastfeeding practices are influenced by a range of external factors such as family support, cultural beliefs, workplace environment, and access to lactation counseling. Future interventions should incorporate strategies to address these contextual elements. For instance, involving family members in educational programs or offering follow-up support through community health workers may enhance the effectiveness of awareness efforts [10].

Conclusion

The study is expected to demonstrate clear distinctions between theoretical knowledge and practical application of the MAA program among postnatal mothers. It will likely emphasize the importance of integrating awareness programs into hospital settings to strengthen both knowledge and practice. Participants are anticipated to demonstrate an improved understanding of the technical aspects of the MAA program, while social elements—such as maternal cooperation and support dynamics—may still present areas for further enhancement. Future research and program planning will aim to bridge these gaps by incorporating more comprehensive components that address both clinical education and sociobehavioral support for breastfeeding practices.

Acknowledgments

We would like to express our sincere gratitude to all those who supported the development of this research protocol. We extend our deepest thanks to our research guide and academic mentors for their expert guidance, insightful feedback, and encouragement throughout the planning and design of this study. We also wish to acknowledge the support of the Smt. Radhikabai Meghe Memorial College of Nursing for providing the necessary resources and administrative assistance required to formulate this protocol. Special thanks to the ethical review committee for reviewing and approving the study plan, and thanks to all the health care staff and officials involved in the implementation of the Mothers' Absolute Affection (MAA) program, whose insights were instrumental in shaping the scope and relevance of this research. Lastly, we are grateful to our peers, friends, and family for their moral support, patience, and motivation, which greatly contributed to the preparation of this research protocol.

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Funding

No funding was received for this study.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

VD contributed to the conceptualization, methodology, investigation, data curation, and original draft preparation of the study. KG provided supervision, assisted with validation, and critically reviewed and edited the manuscript for intellectual content. VT contributed to the formal analysis and interpretation of results and ensured ethical compliance and project administration. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[PDF File, 91 KB - resprot_v15i1e73496_app1.pdf]

References

1. Singh SK, Srivast S. Tracking the progress in breastfeeding after implementation of MAA programme in India. Research Square. Preprint posted online on Dec 30, 2020. [doi: 10.21203/rs.3.rs-134258/v1]

XSL•FO
RenderX

2. Mothers' Absolute Affection (MAA). National Health Mission, Government of Assam. URL: https://dhubri.assam.gov.in/sites/default/files/public_utility/Mothers.pdf [accessed 2026-03-20]
3. Chale LE, Fenton TR, Kayange N. Predictors of knowledge and practice of exclusive breastfeeding among health workers in Mwanza city, northwest Tanzania. BMC Nurs 2016;15:72. [doi: 10.1186/s12912-016-0192-0] [Medline: 28050163]
4. Mothers Absolute Affection. National Health Mission. URL: https://nhm.gov.in/index1.php?lang=1&level=3&sublinkid=1450&lid=799 [accessed 2024-12-13]
5. Garg R, Deepti S, Padda A, Singh T. Breastfeeding knowledge and practices among rural women of Punjab, India: a community-based study. Breastfeed Med 2010 Dec;5(6):303-307. [doi: 10.1089/bfm.2010.0005] [Medline: 20925495]
6. Mother's Absolute Affection. Vikaspedia – Government of India. URL: https://health.vikaspedia.in/viewcontent/health/health-campaigns/mothers-absolute-affection?lgn=en [accessed 2024-12-13]
7. Devi M C, Kaur A, Kaur B, Kaur G. A descriptive study to assess the efficacy of MAA (Mother's Absolute Affection) programme in terms of knowledge and practice among postnatal mothers in selected community areas of Jalandhar, Punjab. International Journal of Current Science :434-443 [FREE Full text]
8. Nayak P. Effectiveness of awareness programme on knowledge regarding home based postnatal practices among primigravid women in selected private hospitals, Udupi. Manipal Journal of Nursing and Health Sciences 2023;9(1) [FREE Full text]
9. Rajak P, Krishna Mandal A, Kumar Jana J, Gayen S. Knowledge of breastfeeding practices among mothers attending a tertiary care setting in East India. Cureus 2023 Apr;15(4):e37146. [doi: 10.7759/cureus.37146] [Medline: 37153262]
10. Gupta R. A study to evaluate the effectiveness of educational package on breastfeeding in terms of knowledge, attitude and practices among primi mothers in selected hospitals of district Ambala, Haryana. IJMPO 2019 Oct 28;5(3):81-88. [doi: 10.18231/j.ijmpo.2019.019]
11. Reddy N S, Dharmaraj A, Jacob J, Sindhu KN. Exclusive breastfeeding practices and its determinants in Indian infants: findings from the National Family Health Surveys-4 and 5. Int Breastfeed J 2023 Dec 20;18(1):69. [doi: 10.1186/s13006-023-00602-z] [Medline: 38124065]

Abbreviations

MAA: Mothers' Absolute Affection
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by J Sarvestan; submitted 05.Mar.2025; peer-reviewed by A Adeoye, A Shadap, J Ogunsakin, T Adebambo; revised version received 20.Jun.2025; accepted 23.Jun.2025; published 30.Mar.2026.

Please cite as:
Dongare V, Gomase K, Taksande V
Assessing the Effectiveness of an Awareness Program on Knowledge and Practice Regarding the Mothers' Absolute Affection Program Among Postnatal Mothers in Selected Hospitals: Protocol for a Quasi-Experimental Study
JMIR Res Protoc 2026;15:e73496
URL: https://www.researchprotocols.org/2026/1/e73496
doi:10.2196/73496

© Vaishnavi Dongare, Kavita Gomase, Vaishali Taksande. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of a Decision Aid to Support Shared Decision-Making on Cannabis Use for Arthritis: Protocol for a Multiphase Study

Heba A T Aref[1], BSc (Pharm), MSc; Yazid N Al Hamarneh[2], BSc (Pharm), PhD; Tony K L Kiang[1], BSc (Pharm), PhD; Elaine Yacyshyn[3], MD; Cheryl A Sadowski[1], BSc (Pharm), PharmD

[1]Faculty of Pharmacy & Pharmaceutical Sciences, University of Alberta, 3-229 Dianne and Irving Kipnes Health Research Academy, 11405 87 Ave. NW, Edmonton, AB, Canada

[2]Department of Pharmacology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

[3]Department of Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

Corresponding Author:
Cheryl A Sadowski, BSc (Pharm), PharmD
Faculty of Pharmacy & Pharmaceutical Sciences, University of Alberta, 3-229 Dianne and Irving Kipnes Health Research Academy, 11405 87 Ave. NW, Edmonton, AB, Canada

Abstract

Background: Conventional treatments have been frequently reported to offer partial relief for some individuals managing arthritis pain and related symptoms, leading many to consider alternative options, such as cannabis. Informed decision-making about cannabis use requires patients to weigh potential benefits and risks in light of their personal values and preferences.

Objective: This study aims to systematically develop a theory-driven, evidence-based, user-centered decision aid (DA) for older adults with arthritis, considering medical cannabis.

Methods: The Ottawa Decision Support Framework, the International Patient Decision Aids Standards, and the User-Centeredness approach guided the DA development process in 4 phases. Phase 1 (completed) involved scoping reviews to explore the development of shared decision-making (SDM) tools and the literature addressing cannabis use in arthritis, guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework. Phase 2 (in progress) consists of semistructured interviews that are conducted with older adults living with arthritis to explore their decisional needs. Phase 3 (in progress) involves semistructured interviews that are conducted with clinicians to explore how they can participate with their patients during cannabis use decision-making. For phases 2 and 3, recruitment has been initiated in a large urban center in Alberta, Canada. Purposive sampling is conducted, and the sample size will be guided by the principle of information power. Phase 4 (not initiated yet) will involve the development of the DA and alpha testing to explore qualification as a DA and acceptability with advisory board members, including patients and clinicians who participate in the development process. Based on the feedback received, the prototype will be refined accordingly. Subsequently, a future study involving a beta testing phase will be conducted to evaluate the usability, accessibility, and comprehensibility of the DA among naive users. The University of Alberta Human Research Ethics Board approved this study. This protocol is based on the latest version of the ethics application, dated December 14, 2024 (Pro00133420).

Results: This project was funded in November 2022. Phase 1 is completed; the advisory board is assembled. Two scoping reviews have been published. The first review showed that the most commonly reported development bases for SDM tools were the Ottawa Decision Support Framework and International Patient Decision Aids Standards criteria, reported in 16% and 53% of tools, respectively. The second review highlighted limited high-quality evidence addressing medical cannabis use for arthritis. Phases 2 and 3 have been initiated. Phase 4 will be initiated when the findings from phases 2 and 3 are available. As of November 2025, we have recruited 10 patients for phase 2 and 7 clinicians for phase 3.

Conclusions: This protocol outlines the systematic development of a patient-centered DA to support SDM about medical cannabis for arthritis pain. The findings will show the feasibility of development and inform the refinement of the DA, guiding its readiness for beta testing and future implementation.

International Registered Report Identifier (IRRID): DERR1-10.2196/76237

(JMIR Res Protoc 2026;15:e76237) doi:10.2196/76237

KEYWORDS

decision support intervention; cannabis; shared decision-making; informed decision-making; decision aid

XSL•FO
RenderX

Introduction

Approximately 1 in 3 adults diagnosed with arthritis diseases (eg, gout, rheumatoid arthritis, and osteoarthritis) have severe chronic joint pain that impacts their quality of life [1]. Many pharmacological agents (eg, nonsteroidal anti-inflammatory drugs, acetaminophen, and opioid analgesics) are available to address such pain [2,3]. However, unsatisfactory responses and adverse reactions have raised interest in using alternative therapies, such as medical cannabis [4-6]. Indeed, arthritic pain has been reported to be one of the most common indications for using medical cannabis in North America [7].

In Canada, patients with arthritis are increasingly inquiring about cannabis as an alternative treatment. This could be explained by the recent legalization of medical cannabis, the aggressive marketing of the cannabis industry, the slow social shift of the stigma of using cannabis for medical purposes, and medical research that reported a limited positive impact of cannabis within specific contexts [7-10]. Despite this increased curiosity, it has been reported that clinicians lack confidence in their knowledge about cannabis [11], as empirically established clinical evidence is inconsistent, and definitive cannabis use guidelines are not yet available [4,12-14]. As such, patients are currently making the decision to use cannabis based on individual preferences and values without having clear guidelines [15,16].

The current lack of cannabis use guidelines, combined with the complexity of cannabis-related decisions, highlights the need for utilizing shared decision-making (SDM) [17,18]. Specifically, patients and clinicians must navigate numerous uncertainties unique to medical cannabis, including (1) the wide variability in cannabinoid formulations, ratios, and delivery methods, each with different onset times and duration of effects [19]; (2) the absence of standardized dosing protocols and the need for individualized titration [15]; (3) potential drug-drug interactions, particularly concerning for older adults with polypharmacy [20]; and (4) weighing potential benefits such as pain reduction and improved sleep quality against risks including cognitive impairment, falls, and cardiovascular effects [21]. These complexities are further compounded by varying legal frameworks, insurance coverage limitations, and persistent stigma surrounding cannabis use [22-24]. Such multidimensional trade-offs and uncertainties make the cannabis use decision for arthritis particularly suited for SDM [25,26]. SDM is the communication process through which patients and clinicians work together collaboratively using specific tools to reach a treatment decision based on integrating clinical evidence with patients' preferences, values, and goals [22]. SDM improves patients' satisfaction and adherence and leads to a better quality of life [23,24]. However, SDM has not been widely incorporated into daily practice due to barriers related to the lack of decision support tools that facilitate SDM within a busy organizational workflow [25-27].

Decision aids (DAs) are structured decision support interventions (DSIs) that systematically present evidence and elicit patient values, especially in situations of preference-sensitive decisions [25,26], such as cannabis use. A DA is designed to present evidence-based information about available options, along with their potential benefits and harms, to support patients in making informed, deliberate, and value-aligned choices [25,27,28]. As such, DAs reduce uncertainty about decision-making and improve the quality of health care decisions, leading to better health outcomes [29,30]. Several DAs have been developed for arthritis diseases [18,31,32]; however, to our knowledge, none of them addressed the topic of cannabis use for patients with arthritis. The need for a DA to facilitate the decision of whether to use cannabis or not is evident, as there is a scarcity of strong recommendations, guidelines, or benefit-to-harm assessments addressing the effectiveness and safety of cannabis for arthritis pain and other symptoms.

Therefore, the objective of this study is to systematically develop a theory-driven, evidence-based, user-centered cannabis DA for older adults with arthritis, guided by the Ottawa Decision Support Framework (ODSF) and the International Patient Decision Aids Standards (IPDAS). We will evaluate the feasibility of developing the first prototype ready for beta testing. Feasibility will be determined through the completion of prespecified milestones, including evidence synthesis, decisional needs assessment, prototype development, iterative refinement through incorporating feedback from the patients and clinicians in the advisory board input, and full documentation following the IPDAS DEVELOPTOOLS reporting checklist [33].

Methods

Study Design and Conceptual Framework

This protocol outlines 4 key phases of development, as illustrated in Figure S1 in Multimedia Appendix 1, with a summary of procedures in Table S1 in Multimedia Appendix 2. The phases comprise phase 1: preliminary work, phase 2: qualitative descriptive study exploring patients' decisional needs, phase 3: qualitative descriptive study exploring clinicians' decisional needs, and phase 4: development and alpha testing. The study is guided conceptually by the ODSF [34,35]—a theory-based framework (grounded in decisional conflict theory [36] and self-efficacy theories [37])—and methodologically by the IPDAS development model [33,38,39].

Outcome Measures

Primary Outcome

The primary outcome of this project is to demonstrate the feasibility of systematically developing a theory-driven, evidence-based DA for a context characterized by limited and evolving evidence—specifically, medical cannabis use for symptom management among older adults with arthritis. Feasibility will be demonstrated through (1), recruitment of at least 80% (n=8-10) of the target sample with ongoing consideration of information power, (2) retention of ≥80% (n=8-10) of participants recruited in phases 2 and 3 to provide feedback for phase 4, (3) adherence to a 12 - to 18-month development timeline (based on comparable studies), (4) the generation of actionable user feedback to guide iterative refinement, and (5) comprehensive reporting of each

XSL•FO
RenderX

development phase in accordance with the IPDAS-recommended DEVELOPTOOLS checklist [33]. Meeting these prespecified benchmarks will indicate the feasibility of developing a prototype ready to progress to beta testing.

Secondary Outcome

The secondary outcome is to obtain acceptability and qualification as data that support proceeding to the beta testing phase. We will conduct the acceptability [40] and the DA qualification assessment [41] as part of alpha testing (iteration with participants). These assessments will provide structured feedback for iterative formative refinement rather than definitive validation for implementation. Measures and scoring procedures are prespecified under phase 4b. This approach will identify aspects requiring refinement before beta testing with naive users and is consistent with iterative development processes for DAs [33,38].

Phase 1: Preliminary Work (Completed)

Phase 1a: Assembling the Advisory Board

The advisory board was assembled under the leadership of the principal investigators, CAS (geriatrics specialized pharmacist) and EY (rheumatologist), who served as the first 2 members of the board. Drawing on their professional networks, they invited members with relevant expertise. Common interests were verified through conversations to ensure a collaborative, multidisciplinary team committed to supporting the development of a DA. The advisory board defined the scope and purpose of the DA.

Phase 1b: Scoping Reviews

Two reviews were conducted to scope the literature on (1) DSIs among patients with arthritis and (2) cannabis use in arthritis [42,43]. Both studies used systematic search and evidence synthesis approaches guided by Arksey and O'Malley's methodological framework and the PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist [44,45].

Phases 2 and 3: Two Qualitative Studies (in Progress)

Overview

Phases 2 and 3 aimed to explore patients' and clinicians' perspectives on cannabis use decision-making. Figure S2 in Multimedia Appendix 3 shows the participant flow throughout the study.

We adopted a qualitative descriptive approach, an inquiry method aiming at providing a comprehensive summary of participants, staying close to the surface of the data without extensive interpretation or theorizing [46-48]. This approach is particularly well suited for exploring the decisional needs of patients and clinicians, as it allows for the direct capture of participants' perspectives in their own words—an essential consideration when the goal is to inform the development of practical, user-centered DA.

Positionality Statement

The research team acknowledges that their professional backgrounds and clinical experiences will inevitably shape how participants' statements are interpreted. The lead researcher (HATA) is a health researcher with training in quantitative and qualitative methodologies and experience in decision-making, geriatrics, and chronic disease management, which may be associated with increased awareness of issues related to medical cannabis safety and evidence uncertainty. Coauthors' stance aligns with the Canadian Rheumatology Association (CRA) Position Statement recognizing insufficient evidence for medical cannabis in arthritis, while calling for empathetic, pragmatic guidance that supports respectful patient-clinician dialogue and prioritizes harm reduction [49]. This stance may predispose the team toward cautious, harm reduction-oriented interpretations of participants' experiences. To address this, we will engage in ongoing reflexive practice, including memoing and analytic discussions, to critically examine how our clinical perspective may influence theme development and framing.

Data Collection

For phase 2, the target population will comprise patient participants who are 55 years of age or older living with any form of arthritis, currently using, interested in, considering, or declining medical cannabis for symptom management. We initiated recruitment in a large urban center in Alberta, Canada. For phase 3, the target population will comprise clinician participants (eg, pharmacists, rheumatologists, family physicians, and nurse practitioners) involved in arthritis care and counseling about symptom management options. We initiated recruitment through advertising in a local medical newsletter.

Purposive sampling will be used, and the sample size will be guided by the principle of information power, wherein the adequacy of the sample depends on the degree to which participants collectively contribute information relevant to the study aim [50]. This approach is consistent with reflexive thematic analysis (RTA), which does not rely on saturation as a marker of adequacy [51,52]. High information power [50] is anticipated, given the narrow research aim (decisional needs regarding medical cannabis for arthritis), the high specificity of the sample, the use of an established theoretical framework (the ODSF) [30], and the expected richness of semistructured interviews guided by a decisional needs assessment workbook [53]. Based on comparable decisional needs studies, we anticipate recruiting approximately 10 to 12 participants per phase. The research team will review data relevant to ODSF constructs and assess information power after every 3 interviews to determine whether additional recruitment is required [50]. Recruitment will cease once consensus among advisory board members indicates that the accumulated data provide sufficient depth, breadth, and relevance to meet the study aim. The inclusion criteria are given in Textbox 1. Eligible participants who agree to participate will be provided with an information letter and consent form outlining the study's purpose, procedures, participants' rights, confidentiality and anonymity protocols, and data storage practices. A trained researcher (HATA) will conduct all one-on-one semistructured interviews using interview guides for patients and clinicians (Multimedia Appendices 4 and 5) informed by the decisional needs assessment approach, the ODSF, and probing techniques [34,53-55]. To support accessibility and accommodate

XSL•FO
RenderX

participants' preferences, interviews will be conducted virtually via a secure video conferencing platform, with telephone or in-person options available as needed. All interviews will be audio-recorded with participant consent and are expected to last approximately 45 to 60 minutes. At the end of each interview, patients and clinicians will be asked to provide demographic information (eg, for patients: age, education level, and cannabis usage status and for clinicians: years of practice, age, and area of practice). Audio recordings will be transcribed verbatim using Quirkos [56] and deidentified using unique participant codes (eg, P01 for patients, C01 for clinicians). Transcripts will be cross-checked against audio recordings for accuracy and completeness. Transcripts, audio files, and the corresponding linking key will then be securely stored on encrypted, password-protected university servers accessible only to the research team and compliant with institutional data protection policies.

Textbox 1. Inclusion criteria

Phase 2 (patients)

- 55 years of age or older (although the World Health Organization generally defines older adults as those aged 60 years and older, chronic conditions, such as arthritis, increase from the mid-50s onward [5,57,58]. For this study, "older adults" are individuals aged 55 years or older, consistent with aging and arthritis research identifying 55 years of age as the onset of heightened chronic disease burden [5,57-61]
- Living with arthritis
- Considering, currently using, have previously used, or have explicitly decided against using cannabis for arthritis symptom management

Phase 3 (clinicians)

- Currently treating patients with arthritis
- Have experience supporting or making decisions about cannabis use (eg, responding to patient questions about cannabis, recommending to use or avoid cannabis)
- Interested in describing their needs to participate with patients in making a decision about cannabis use

Data Analysis

To ensure transparency, rigor, and methodological clarity, both qualitative studies will be reported in accordance with the Standards for Reporting Qualitative Research guidelines [62]. The data from both phases will be analyzed using RTA situated within a qualitative descriptive approach, an analytic combination that enables the researcher to remain close to participants' accounts while acknowledging that meaning-making is inherently interpretive [46-48,51]. RTA provides methodological flexibility suitable for exploring decisional needs in depth, while the qualitative descriptive approach ensures that findings are grounded in the language and experiences of older adults and clinicians [48,51].

A single researcher (HATA) will lead all analytic activities, consistent with RTA's epistemological stance that positions the researcher as the primary analytic instrument and supports sole-coder analysis as a flexible, coherent, and methodologically appropriate practice [51,63]. The analysis procedure will begin with familiarization through repeated reading of verbatim transcripts and verification of transcription accuracy to immerse oneself in the data. A hybrid deductive–inductive approach will be used for coding. An initial coding framework based on the ODSF will be applied deductively to identify predefined decisional needs constructs (eg, knowledge gaps, expectations, values, decisional conflict, and support needs). In parallel, inductive coding will be used to capture unanticipated themes emerging from participants' accounts and the cannabis context. Codes will be iteratively refined and organized into higher-order themes through RTA. Final themes will be explicitly mapped to the DA components, including informational content, values-clarification exercises, and clinician-patient discussion prompts, ensuring that the DA content is directly informed by empirically identified decisional needs of both patients and clinicians. Coding will be conducted in Quirkos [56] to organize excerpts, identify patterned meaning, and support early category development.

Throughout the analysis, HATA will generate initial themes, refining them through iterative cycles of reviewing and returning to the data, critical reflection, and agreement among the research team. This approach will serve to deepen interpretive insight by incorporating the research team's diverse disciplinary standpoints in health services research, geriatric pharmacotherapy, rheumatology, and pharmaceutical sciences, consistent with RTA's emphasis on reflexivity over reliability [51]. Themes will be assessed for coherence, consistency, relevance, and alignment with the research aims and the goal of developing the DA. Finally, themes will be defined and named by refining their boundaries and ensuring clear definitions.

To ensure trustworthiness across both qualitative descriptive studies, multiple strategies will be used to establish credibility, dependability, confirmability, and transferability [64,65]. Credibility will be enhanced through advisory board review of thematic findings, negative-case analysis to identify divergent patterns, and triangulation across patient and clinician datasets to explore convergent, divergent, and complementary insights [66-68]. A comprehensive audit trail with a versioned codebook documenting analytic decisions and coding iterations will be maintained throughout analysis to transparently track coding evolution and methodological reflections supporting dependability, while confirmability will be supported through extensive reflexivity practices, including researcher journaling and reflexive memoing [69-71]. Transferability will be established through thick description of study context,

XSL•FO
RenderX

procedures, and participant characteristics, enabling readers to assess applicability to other settings [72]. The advisory board will provide a final review of the thematic structure to evaluate its relevance and usefulness for informing the DA prototype development.

Phase 4: Developing the DA (Not Initiated Yet)

Overview

Phase 4a focuses on generating the initial prototype, while phase 4b constitutes a formative, embedded alpha-testing phase situated within development rather than a summative validation intended for implementation. The purpose of alpha testing is to obtain structured, actionable feedback to refine the prototype prior to subsequent beta testing with naive users. To enhance methodological transparency and reproducibility, all outcome measures, instruments, scoring procedures, and a priori success thresholds are prespecified below.

Phase 4a: Developing the First Prototype of the DA

Overview

The development of the DA will follow a structured, transparent, and iterative process aligned with the IPDAS Collaboration criteria for content development, balance, and evidence reporting [73,74]. The prototype's first draft will be developed based on phases 1-3, the IPDAS qualifying criteria [41], and recent cannabis evidence literature [43,75]. As it has been reported that DA development processes have been suboptimally reported [33,38], we plan to describe the development process comprehensively and transparently in a stand-alone paper. Witteman et al [33] highlighted 8 studies [32,76-84] as being exemplary in reporting their systematic approach. Therefore, we planned this phase based on these 8 studies, together with leveraging the DEVELOPTOOLS checklist that ensures a rigorous, transparent, systematic approach to DA development [33]. A structured content-governance process is established to ensure rigor and transparency as follows.

First: Cannabis Evidence Synthesis

Due to the current lack of a coherent and homogeneous body of clinical evidence and limited high-quality systematic reviews or clinical practice guidelines on medical cannabis use for arthritis, the current literature does not meet the methodological requirements needed to conduct a Cochrane systematic review or apply evidence-rating frameworks such as GRADE (Grading of Recommendations Assessment, Development, and Evaluation). Therefore, we will instead collect input from patients and clinicians on preferred evidence sources for content. In addition, the synthesis of evidence for content development will be informed by the findings from our team's prior work, together with the updated position statement of the CRA [58]. Consistent with IPDAS guidance for emerging evidence contexts, uncertainty and evidence gaps will be transparently communicated throughout the DA.

Second: Balanced Content Development

To ensure a balanced presentation of potential benefits and harms, we will summarize evidence-based advantages, limitations, and risks of medical cannabis using the highest-quality evidence available at the time of development. In synthesizing evidence, we recognize that the findings from other jurisdictions may not be directly applicable to the Canadian context, where product availability and THC:CBD ratios and potencies differ markedly from settings, such as the United States, which operates under mixed state laws; lacks federal legalization; and exhibits substantial variability in product quality, access, and cannabinoid composition. Accordingly, Canadian data will be prioritized whenever possible, and international evidence will be interpreted cautiously. Evidence will be synthesized using balanced, nondirective, evidence-informed language that presents potential benefits and harms to support shared deliberation between patients and clinicians.

Third: Accessibility, Readability, and Numeracy Standards

Accessibility and readability standards will be considered during the development process and will be assessed later on during beta testing. The findings from phases 2 and 3 will determine patients' and clinicians' preferences for the format of the DA, whether a digital or paper-based. For either format, we plan to ensure that the DA is accessible and comprehensible for older adults. Therefore, we will adopt accessibility and readability standards consistent with health literacy best practices [38,85].

- Accessibility: The DA formats (digital or print) will conform to Web Content Accessibility Guidelines [86]. Features will include adjustable text size, high-contrast design, screen reader compatibility, alternative text for images, clear structure, and captions for all multimedia. For the print version, equivalent accessibility principles will be applied, including the use of large fonts (minimum 14 pt), high-contrast text and background colors, clear headings, sufficient white space, and simplified layout to support legibility for older adults.
- Readability: Content will be written in plain language at a Grade 6 - 8 reading level, verified using the *Flesch–Kincaid Grade Level*. Short sentences, active voice, and simplified terminology will be prioritized. Visuals (icons and infographics) will be used to supplement text and enhance comprehension.
- Numeracy supports: Where feasible, quantitative findings will be reported in multiple formats (eg, consistent denominators, natural frequencies, visual displays) supplemented by plain-language explanations and visual aids to enhance comprehension across varying numeracy levels [87].

Fourth: Intervention Delivery

The delivery characteristics of the DA—including its format, timing, and the individuals involved in its delivery—will be determined based on the findings from phases 2 and 3. Grounding format decisions (digital vs nondigital) in qualitative insights further ensures that the DA accommodates the needs of older adult participants, consistent with standards for DA development. Following this, the DA will undergo further refinement during phase 4b, where advisory board feedback will be systematically integrated.

XSL•FO
RenderX

Fifth: Version Control and Update Cadence

A structured version-control protocol will be implemented to document all content iterations, authorship contributions, revisions, and editorial decisions. In accordance with IPDAS quality criteria, all research team and advisory board members will be required to disclose potential conflicts of interest, which will be reviewed during each update cycle. All funding sources will be transparently reported [88]. Evidence sources will be systematically cataloged and critically reappraised every 2 years, with provisions for earlier review should stronger evidence, regulatory changes, emerging safety considerations, or jurisdiction-specific policy updates (eg, Health Canada or CRA guidance) arise. This governance structure will help ensure that DA's evidence summaries remain current, accurate, and aligned with evolving scientific and regulatory developments.

Phase 4b: Alpha Testing the Prototype With the Advisory Board

Alpha testing will evaluate whether the DA is acceptable to individuals familiar with its development and whether it qualifies as a DA, following the iterative early-testing processes described in IPDAS development literature. Advisory board members will be asked to perform 2 tasks.

Task 1: Structured Assessment of DA Qualification and Acceptability

Each participant will independently review the DA and complete two questionnaires:

1. IPDAS Qualifying Criteria Questionnaire (Multimedia Appendix 6): This questionnaire will be used to evaluate adherence to core DAs standards, including articulation of the decision; identification of target users; balanced presentation of options; clarity regarding benefits and harms; inclusion of values-clarification components; and transparency of evidence sources, authorship, and funding [41]. The DA will be deemed ready for beta testing if ≥75% of participants report that all IPDAS qualifying criteria are met and confirm compliance with bias-reduction standards. Thus, advisory board members (patients and clinicians) will classify whether the prototype qualifies as a DA using the IPDAS qualifying criteria. To qualify, the prototype must meet all 8 essential "qualifying criteria" (eg, defining the decision, presenting options, providing balanced benefits and harms, including a values-clarification exercise) and satisfy 9 criteria that reduce the risk of biased decision-making (eg, balanced presentation, citation of evidence sources, disclosure of conflicts and funding) [41].
2. Acceptability questionnaire (Multimedia Appendix 7): This questionnaire will assess advisory board members' perceptions of the DA's clarity, comprehensibility, usability, neutrality, level of detail, usefulness for decision-making, compatibility with clinical workflow, and overall suitability for adoption in practice [40]. The DA will meet acceptability thresholds if: (1) ≥75% (n=7-9) of participants rate each informational component as "good" or "excellent," (2) no component receives >25% (n=2-3) "poor" ratings, and (3) ≥75% of all questionnaire responses fall within "agree/strongly agree" (ie, mean score ≥4.0) [38,40,74].

Task 2: Semistructured Cognitive Interview

Think-aloud and semistructured cognitive interview will be conducted with all participants to contextualize and corroborate questionnaire results [33,38,89]. Participants will elaborate on reasons for high or low ratings; perceived ambiguities or biases; navigation challenges or areas of cognitive burden; and suggestions for improving clarity, neutrality, and integration into clinical encounters.

Interviews will be audio-recorded, and field notes will capture contextual observations. Interviews will be analyzed to identify usability concerns, refine content, and guide iterative improvements before beta testing.

Quantitative and qualitative findings will be analyzed and used together to identify issues to be addressed through iterative refinement. All changes will be logged in a structured decision log, aligning with DEVELOPTOOLS recommendations for transparent reporting [33].

This alpha-testing stage ensures that the DA is evidence-based, clinically accurate, and responsive to end user perspectives before beta testing. By integrating evidence synthesis, structured content governance, explicit accessibility and readability standards, and predefined acceptability and qualification thresholds, the protocol offers a transparent and user-centered development process [33,38,90]. All modifications will be documented in a design-decision log to ensure transparency and traceability of development decisions.

Beta Testing With Naive Users (Future Study)

Beta testing will be conducted as a future study to evaluate the usability, accessibility, and comprehensibility of the DA among naive users—older adults with arthritis and clinicians who did not participate in its development. Consistent with IPDAS and DEVELOPTOOLS checklist [33], this stage will assess real-world navigation patterns, identify usability and literacy-related barriers, and determine whether the DA functions effectively across diverse levels of technology experience and health literacy. A purposive sample of naive users will be recruited to ensure variation in age, gender, arthritis type, cannabis experience, and literacy or technology proficiency, with no advisory-board or alpha-testing participants included. Progression from beta testing to pilot testing will be guided by predefined acceptance thresholds: Patient Education Materials Assessment Tool understandability ≥70% (indicating adequate clarity and ease of understanding), full compliance with the Web Content Accessibility Guidelines (representing the minimum internationally accepted standard for digital accessibility), successful independent navigation of the DA by at least 75% of the sample that will be recruited for beta-testing in the future. and correct interpretation of at least 80% of icons and visual symbols used within the DA [85]. Participants will complete tasks simulating real-world decision-making (eg, information seeking, values clarification, option evaluation) while engaging in a think-aloud exercise to capture navigation challenges and comprehension issues. Usability and comprehension will be measured through structured observation, brief comprehension questions, and the System Usability Scale [91]. Qualitative and quantitative analyses will identify

XSL•FO
RenderX

readability or usability barriers to inform the iterative refinement of the prototype before pilot testing. A separate protocol manuscript will describe the full beta-testing methodology, pilot-testing procedures, and implementation-evaluation plan to ensure a rigorous, staged, user-centered evaluation aligned with best practices in DA development.

Ethical Considerations

This study received ethical approval from the University of Alberta Health Research Ethics Board – Health Panel (Pro00133420) and is registered on the Open Science Framework . All participants will provide informed verbal consent, including explicit consent for audio recording, prior to beginning the interview. Because the interviews explore cannabis decision-making among older adults—a topic that may involve stigma, sensitive disclosures, or safety-related behaviors—additional safeguards have been incorporated [92]. The interviewer will use nonjudgmental, supportive communication techniques and will avoid probing for details related to illicit procurement or illegal activity, while informing participants that confidentiality is upheld within legal and institutional limits [93]. Any disclosure suggesting imminent risk of harm to self or others will be managed according to mandatory institutional duty-to-report procedures [92].

Participants will receive a CAD $30 (approximately US $22) honorarium per interview as compensation for their time, following ethical guidance on avoiding coercion while maintaining fairness [94]. Remote interviews conducted via encrypted institutional Zoom or telephone will follow privacy safeguards for virtual qualitative research, including confirming that participants are in a private location and ensuring that only audio—not video—is recorded [18,95,96]. Alternative modalities (eg, telephone or, when feasible, in-person interviews) will be offered if privacy cannot be adequately maintained. The interviewer will conduct sessions from secure, private environments to ensure reciprocal confidentiality. To safeguard participant privacy, all identifiable data—including audio recordings, transcripts, and field notes—will be systematically deidentified before analysis. Identifiable information and reidentification keys will be stored separately from deidentified transcripts on encrypted, password-protected institutional servers with restricted access for authorized research personnel only. Deidentified data will be retained for 5 years and then securely deleted. Audio recordings will be transcribed verbatim and deleted within 6 months following verification to minimize identifiable data retention. Any adverse events or unintended consequences that arise during interviews or prototype testing will be documented, reviewed, and, when necessary, reported to the Health Research Ethics Board. Because interview conversations may elicit emotional discomfort related to chronic pain, stigma, or treatment decision-making, the interviewer will offer breaks or pause or discontinue the interview upon the participant's request. Researchers will be ready with information if participants request a resource sheet outlining local arthritis supports, pain management services, and mental health resources; referral pathways will also be available upon request. All participants will receive research team contact information for any follow-up concerns or questions. These integrated safeguards uphold ethical standards and protect participant autonomy and confidentiality.

Results

Phase 1: Preliminary Work

Phase 1a: Assembling the Advisory Board

Phase 1a is completed as the advisory board is assembled, and it includes 1 rheumatologist (EY), 1 pharmacist (CAS), 1 clinical and experimental science researcher (TKLK), 1 patient-oriented research expert (YNA), 1 graduate student researcher trained in qualitative research (HATA), and 1 patient living with arthritis. In alignment with the IPDAS development process, the advisory board collaboratively defined the scope of the DA. Following IPDAS guidance to clarify the target decision, user population, and context, the board identified the decision as preference-sensitive, involving trade-offs between potential benefits, risks, and uncertainties. Accordingly, the DA was scoped to assist patients and clinicians in engaging in SDM regarding the use of cannabis for symptom management.

Phase 1b: Scoping Reviews

Phase 1b is completed. Two scoping reviews were conducted to inform the development of the DA. The first review examined the development and evaluation processes of DSIs in rheumatology [38], identifying 36 DSIs classified as educational tools, DAs, or adaptive conjoint analysis aids. Development was primarily based on frameworks, such as the IPDAS (53%) and the ODSF (16%). Common evaluation measures included the IPDAS qualifying criteria and acceptability assessments. Decision tools varied in format (eg, paper based, web based, interactive) and timing of use (before, during, or after consultations). The second scoping review investigated evidence for cannabis-based medicines in osteoarthritis, revealing mixed evidence with only 60% of studies reporting improvements in pain [58]. Larger studies with longer exposure durations showed no significant benefits. The review highlighted limited high-quality evidence and the need for tailored information addressing patient concerns about safety, efficacy, and potential interactions. Based on the findings from both reviews, we decided to follow established standards (ODSF and IPDAS Development Process Model) emphasizing user-centered methods. These reviews provided actionable insights for developing DSIs tailored to rheumatology patients' needs regarding cannabis use decisions.

Phases 2 and 3: Two Qualitative Studies

Phases 2 and 3 are currently being carried out. Recruitment started in January 2024. As of November 2025, we have recruited 10 patients for phase 2, where the majority self-identified as White, over 60 years of age, and with arthritis. For phase 3, we recruited 7 clinicians, including pharmacists, rheumatologists, and nurses with over 5 years of experience as clinicians. We expect to finalize data collection and analysis by early 2026.

Phase 4: Developing the DA

Phase 4 results will include developing the first draft of the prototype and evaluating its acceptability and qualification as

XSL•FO
RenderX

a DA by the advisory board. To date, this phase has not been initiated as it depends on the completion of phases 2 and 3.

Discussion

Principal Findings

DAs have been widely endorsed by professional rheumatology and patient care associations as an intervention to improve the quality of health care delivery, with SDM now considered a gold standard in clinical practice [72-74].

Prior work has demonstrated that the suboptimal reporting of the development process of DAs can undermine their later implementation. To mitigate this risk, we adopt established IPDAS-aligned development practices and commit to comprehensive, transparent reporting throughout all phases, consistent with IPDAS guidance that the development phase warrants comprehensive reporting similar to that of evaluation [33,38,90]. The findings from phase 1 and preliminary findings from ongoing qualitative studies of phases 2 and 3 highlight a clear need for informational support to assist decision-making in this context. Both patients and clinicians expressed uncertainty and a desire for credible, accessible information to guide medical cannabis decisions. This reinforces the relevance of our project to tailor a DA that addresses decisional needs specific to patients living with arthritis and clinicians treating arthritis.

This protocol outlines a systematic, multistage approach for the development and evaluation of a theory-informed, evidence-based DA to facilitate SDM regarding medical cannabis use for arthritis-related pain. Our proposed intervention integrates theoretical guidance (ODSF and IPDAS) together with empirical evidence [qualitative studies exploring decisional needs] to produce a DA prototype tailored to arthritis care practice. To our knowledge, this will be the first study to apply a transparent, structured methodology for developing a cannabis-specific DA for individuals living with arthritis.

If successful, this work may facilitate better clinical conversations and alignment of the decision with patient values, while laying the groundwork for broader development and implementation of DAs. We anticipate that such an intervention will improve decision quality, promote patient autonomy, and positively impact patient outcomes, including satisfaction with the decision and quality of life. Furthermore, this protocol contributes to the evolving literature on SDM in arthritis care and expands the evidence base for decision-making support in the context of medical cannabis use.

Strengths and Limitations

A major strength of this protocol lies in its use of the ODSF, the IPDAS development model, and user-centeredness principles as the foundation guiding the development process [30,41]. The use of ODSF facilitates the identification of decisional needs, while IPDAS provides a structured, evidence-based framework for study procedures and reporting. Together, these models—alongside user-centered design—maximize intended users' input and enhance the potential usability of the DA [97,98].

The limitations of this study protocol include the fact that the qualitative findings are mostly subjective, and generalizability could be limited. We will try to mitigate this issue by providing a thick description of the findings, setting, and participants. Also, most development decisions are at the discretion of the advisory board, depending on their interpretation of the findings. Therefore, we will transparently report the decisions made by the advisory board with a justification for each decision. Limiting think-aloud alpha testing to advisory board members and other participants involved in DA development may introduce a positive feedback bias. However, such bias is appropriate and expected during alpha-stage refinement, where the goal is to strengthen conceptual clarity and structure before exposing an immature prototype to naive users. Despite these limitations, we are confident that these methods, which are recommended by the IPDAS [33,38] and used by other researchers in the field, together with the expertise of the advisory board members, will lead to quality DA with a high chance of usability and implementation.

Acknowledgments

No generative artificial intelligence tools (eg, ChatGPT or other large language models) were used in the writing, editing, analysis, or preparation of this manuscript. All content was generated exclusively by the authors.

Funding

This study is supported by the Alberta Innovates mCannabis grant #202101031 (https://tinyurl.com/cbbmk5xn). Funding was received by EY and CAS. The funder did not play a role in study design, data collection, analysis, publications, or preparation for the manuscripts.

Data Availability

As this is a study protocol, no datasets have yet been generated; upon completion, qualitative data (audio recordings and transcripts) will contain potentially identifiable and sensitive information related to cannabis decision-making among older adults and therefore cannot be publicly released under the conditions of the University of Alberta Health Research Ethics Board approval (Pro00133420). Deidentified excerpts, coding frameworks, interview guides, and the final decision aid will be made available in a separate publication and deposited in the project's Open Science Framework repository, where appropriate. Researchers may request access to deidentified qualitative data for ethically approved secondary analyses by contacting the principal investigator; access

XSL•FO
RenderX

will require ethics approval, a data-use agreement, and confirmation that the proposed use aligns with participant consent and privacy protections.

Authors' Contributions

Conceptualization: CAS, EY, HATA
Funding acquisition: CAS, EY
Methodology: CAS, EY, HATA
Project administration: HATA
Review & editing: CAS, EY, TKLK, YNA
Supervision: CAS
Writing – original draft: HATA
Writing – review & editing: HATA
All authors approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Phases of developing the decision aid.
[DOCX File, 30 KB - resprot_v15i1e76237_app1.docx]

Multimedia Appendix 2
Table of summary of procedures for each phase.
[DOCX File, 18 KB - resprot_v15i1e76237_app2.docx]

Multimedia Appendix 3
Participant flowchart.
[DOCX File, 26 KB - resprot_v15i1e76237_app3.docx]

Multimedia Appendix 4
Interview guide—patients.
[DOCX File, 25 KB - resprot_v15i1e76237_app4.docx]

Multimedia Appendix 5
Interview guide—clinicians.
[DOCX File, 26 KB - resprot_v15i1e76237_app5.docx]

Multimedia Appendix 6
International Patient Decision Aids Standards (IPDAS) qualifying criteria.
[DOCX File, 33 KB - resprot_v15i1e76237_app6.docx]

Multimedia Appendix 7
Acceptability scale.
[DOCX File, 41 KB - resprot_v15i1e76237_app7.docx]

References

1. Badley EM, Goulart CM, Millstone DB, Perruccio AV. An update on arthritis in Canada—national and provincial data regarding the past, present, and future. J Rheumatol 2019 Jun;46(6):579-586. [doi: 10.3899/jrheum.180147] [Medline: 30770501]
2. van Laar M, Pergolizzi JV Jr, Mellinghoff HU, et al. Pain treatment in arthritis-related pain: beyond NSAIDs. Open Rheumatol J 2012;6:320-330. [doi: 10.2174/1874312901206010320] [Medline: 23264838]
3. Stein C, Baerwald C. Opioids for the treatment of arthritis pain. Expert Opin Pharmacother 2014 Feb;15(2):193-202. [doi: 10.1517/14656566.2014.861818] [Medline: 24283649]
4. Zbib A. Canadian Rheumatology Association (CRA) position statement on medical cannabis use in rheumatic disease. : Canadian Rheumatology Association (CRA); 2021 URL: https://rheum.ca/wp-content/uploads/2019/01/

XSL•FO
RenderX

1-Canadian-Rheumatology-Association-CRA-Position-Statement-on-Medical-Cannabis-Use-in-Rheumatic-Disease.pdf [accessed 2024-04-21]
5. Bobitt J, Kang H, Croker JA III, Quintero Silva L, Kaskie B. Use of cannabis and opioids for chronic pain by older adults: distinguishing clinical and contextual influences. Drug Alcohol Rev 2020 Sep;39(6):753-762. [doi: 10.1111/dar.13080] [Medline: 32469109]
6. Medical cannabis. Arthritis Society Canada. URL: https://arthritis.ca/treatment/medication/medical-cannabis/ [accessed 2023-04-11]
7. Wipfler K, Simon TA, Katz P, Wolfe F, Michaud K. Increase in cannabis use among adults with rheumatic diseases: results from a 2014-2019 United States observational study. Arthritis Care Res (Hoboken) 2022 Dec;74(12):2091-2099. [doi: 10.1002/acr.24752] [Medline: 34269524]
8. Filippini G, Minozzi S, Borrelli F, Cinquini M, Dwan K. Cannabis and cannabinoids for symptomatic treatment for people with multiple sclerosis. Cochrane Database Syst Rev 2022 May 5;5(5):CD013444. [doi: 10.1002/14651858.CD013444.pub2] [Medline: 35510826]
9. Dai H, Richter KP. A national survey of marijuana use among US adults with medical conditions, 2016-2017. JAMA Netw Open 2019 Sep 4;2(9):e1911936. [doi: 10.1001/jamanetworkopen.2019.11936] [Medline: 31539078]
10. Keyhani S, Steigerwald S, Ishida J, et al. Risks and benefits of marijuana use: a national survey of U.S. adults. Ann Intern Med 2018 Sep 4;169(5):282-290. [doi: 10.7326/M18-0810] [Medline: 30039154]
11. Fitzcharles MA, Ste-Marie PA, Clauw DJ, et al. Rheumatologists lack confidence in their knowledge of cannabinoids pertaining to the management of rheumatic complaints. BMC Musculoskelet Disord 2014 Jul 30;15:1-6. [doi: 10.1186/1471-2474-15-258] [Medline: 25080153]
12. Busse JW, Vankrunkelsven P, Zeng L, et al. Medical cannabis or cannabinoids for chronic pain: a clinical practice guideline. BMJ 2021 Sep 8;374:n2040. [doi: 10.1136/bmj.n2040] [Medline: 34497062]
13. Fitzcharles MA, Clauw DJ, Häuser W. Cautious hope for cannabidiol (CBD) in rheumatology care. Arthritis Care Res (Hoboken) 2023 Jun;75(6):1371-1375. [doi: 10.1002/acr.24176] [Medline: 32144889]
14. Connor JP, Stjepanović D, Le Foll B, Hoch E, Budney AJ, Hall WD. Cannabis use and cannabis use disorder. Nat Rev Dis Primers 2021 Feb 25;7(1):16. [doi: 10.1038/s41572-021-00247-4] [Medline: 33627670]
15. Boehnke KF, Scott JR, Litinas E, et al. Cannabis use preferences and decision-making among a cross-sectional cohort of medical cannabis patients with chronic pain. J Pain 2019 Nov;20(11):1362-1372. [doi: 10.1016/j.jpain.2019.05.009] [Medline: 31132510]
16. Zeng L, Lytvyn L, Wang X, et al. Values and preferences towards medical cannabis among people living with chronic pain: a mixed-methods systematic review. BMJ Open 2021 Sep 7;11(9):e050831. [doi: 10.1136/bmjopen-2021-050831] [Medline: 34493521]
17. Wilson M, Klein T, Bindler RJ, Kaplan L. Shared decision-making for patients using cannabis for pain symptom management in the United States. Pain Manag Nurs 2021 Feb;22(1):15-20. [doi: 10.1016/j.pmn.2020.09.009] [Medline: 33139204]
18. Morrison T, Foster E, Dougherty J, Barton J. Shared decision making in rheumatology: a scoping review. Semin Arthritis Rheum 2022 Oct;56:152041. [doi: 10.1016/j.semarthrit.2022.152041] [Medline: 35738040]
19. MacCallum CA, Russo EB. Practical considerations in medical cannabis administration and dosing. Eur J Intern Med 2018 Mar;49:12-19. [doi: 10.1016/j.ejim.2018.01.004] [Medline: 29307505]
20. Alsherbiny MA, Li CG. Medicinal cannabis-potential drug interactions. Medicines (Basel) 2018 Dec 23;6(1):3. [doi: 10.3390/medicines6010003] [Medline: 30583596]
21. van den Elsen GAH, Ahmed AIA, Lammers M, et al. Efficacy and safety of medical cannabinoids in older subjects: a systematic review. Ageing Res Rev 2014 Mar;14:56-64. [doi: 10.1016/j.arr.2014.01.007] [Medline: 24509411]
22. Bottorff JL, Bissell LJL, Balneaves LG, Oliffe JL, Capler NR, Buxton J. Perceptions of cannabis as a stigmatized medicine: a qualitative descriptive study. Harm Reduct J 2013 Feb 16;10:2. [doi: 10.1186/1477-7517-10-2] [Medline: 23414118]
23. Delling FN, Vittinghoff E, Dewland TA, et al. Does cannabis legalisation change healthcare utilisation? A population-based study using the healthcare cost and utilisation project in Colorado, USA. BMJ Open 2019 May 15;9(5):e027432. [doi: 10.1136/bmjopen-2018-027432] [Medline: 31092662]
24. Lapen K, Mishra Meza A, Dee EC, et al. Patient out-of-pocket costs for cannabis use during cancer treatment. J Natl Cancer Inst Monogr 2024 Aug 15;2024(66):305-312. [doi: 10.1093/jncimonographs/lgad030] [Medline: 39108238]
25. O'Connor AM, Stacey D, Rovner D, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2001(3):CD001431. [doi: 10.1002/14651858.CD001431] [Medline: 11686990]
26. Franco JVA, Riganti P, Ruiz Yanzi MV, Kopitowski K. Equipoise is preference sensitive. Can Fam Physician 2020 Aug;66(8):551-552. [Medline: 32817021]
27. O'Connor AM, Llewellyn-Thomas HA, Flood AB. Modifying unwarranted variations in health care: shared decision making using patient decision aids. Health Aff (Millwood) 2004;Suppl Variation:VAR63-VAR72. [doi: 10.1377/hlthaff.var.63] [Medline: 15471770]
28. Alden DL, Friend J, Chun MBJ. Shared decision making and patient decision aids: knowledge, attitudes, and practices among Hawai'i physicians. Hawaii J Med Public Health 2013 Nov;72(11):396-400. [Medline: 24251086]

XSL•FO
RenderX

29. O'Connor A, Stacey D, Saarimaki A, Pardo J, Rader T, Welch V, et al. Ottawa patient decision aid development etraining. Ottawa Hospital Research Institute. URL: https://decisionaid.ohri.ca/eTraining/session.php?section=s10 [accessed 2022-06-29]
30. Stacey D, Légaré F, Boland L, et al. 20th anniversary Ottawa Decision Support Framework: part 3 overview of systematic reviews and updated framework. Med Decis Making 2020 Apr;40(3):379-398. [doi: 10.1177/0272989X20911870] [Medline: 32428429]
31. Search results—A to Z inventory of patient decision aids. Ottawa Hospital Research Institute. URL: https://decisionaid.ohri.ca/AZsearch.php?criteria=arthritis [accessed 2022-11-22]
32. Barton JL, Koenig CJ, Evans-Young G, et al. The design of a low literacy decision aid about rheumatoid arthritis medications developed in three languages for use during the clinical encounter. BMC Med Inform Decis Mak 2014 Nov 25;14:104. [doi: 10.1186/s12911-014-0104-8] [Medline: 25649726]
33. Witteman HO, Maki KG, Vaisson G, et al. Systematic development of patient decision aids: an update from the IPDAS Collaboration. Med Decis Making 2021 Oct;41(7):736-754. [doi: 10.1177/0272989X211014163] [Medline: 34148384]
34. Hoefel L, O'Connor AM, Lewis KB, et al. 20th anniversary update of the Ottawa Decision Support Framework Part 1: a systematic review of the decisional needs of people making health or social decisions. Med Decis Making 2020 Jul;40(5):555-581. [doi: 10.1177/0272989X20936209] [Medline: 32659154]
35. O'Connor AM, Tugwell P, Wells GA, et al. A decision aid for women considering hormone therapy after menopause: decision support framework and evaluation. Patient Educ Couns 1998 Mar;33(3):267-279. [doi: 10.1016/s0738-3991(98)00026-3] [Medline: 9731164]
36. Janis IL, Mann L. Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment: Free Press; 1977. URL: http://betterworldbooks.com/product/detail/decision-making-a-psychological-analysis-of-conflict-choice-and-commitment-9780029161609 [accessed 2026-03-06]
37. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1037//0033-295x.84.2.191] [Medline: 847061]
38. Coulter A, Stilwell D, Kryworuchko J, Mullen PD, Ng CJ, van der Weijden T. A systematic development process for patient decision aids. BMC Med Inform Decis Mak 2013;13 Suppl 2(Suppl 2):S2. [doi: 10.1186/1472-6947-13-S2-S2] [Medline: 24625093]
39. Vaisson G, Provencher T, Dugas M, et al. User involvement in the design and development of patient decision aids and other personal health tools: a systematic review. Med Decis Making 2021 Apr;41(3):261-274. [doi: 10.1177/0272989X20984134] [Medline: 33655791]
40. O'Connor AM, Cranney A. User manual—acceptability. : Ottawa Hospital Research Institute (OHRI) URL: https://decisionaid.ohri.ca/docs/develop/user_manuals/um_acceptability.pdf [accessed 2022-12-30]
41. Joseph-Williams N, Newcombe R, Politi M, et al. Toward minimum standards for certifying patient decision aids: a modified Delphi consensus process. Med Decis Making 2014 Aug;34(6):699-710. [doi: 10.1177/0272989X13501721] [Medline: 23963501]
42. Aref HAT, Turk T, Dhanani R, et al. Development and evaluation of shared decision-making tools in rheumatology: a scoping review. Semin Arthritis Rheum 2024 Jun;66:152432. [doi: 10.1016/j.semarthrit.2024.152432] [Medline: 38554593]
43. Xiao ATY, Turk T, Deol K, et al. Evidence for the use of cannabis-based medicines in osteoarthritis: a scoping review. Clin Rheumatol 2024 Aug;43(8):2375-2390. [doi: 10.1007/s10067-024-07001-7] [Medline: 38853226]
44. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
45. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
46. Neergaard MA, Olesen F, Andersen RS, Sondergaard J. Qualitative description—the poor cousin of health research? BMC Med Res Methodol 2009 Jul 16;9(1):1-5. [doi: 10.1186/1471-2288-9-52] [Medline: 19607668]
47. Kim H, Sefcik JS, Bradway C. Characteristics of qualitative descriptive studies: a systematic review. Res Nurs Health 2017 Feb;40(1):23-42. [doi: 10.1002/nur.21768] [Medline: 27686751]
48. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
49. Fitzcharles MA, Niaki OZ, Hauser W, Hazlewood G, Canadian Rheumatology Association. Position statement: a pragmatic approach for medical cannabis and patients with rheumatic diseases. J Rheumatol 2019 May;46(5):532-538. [doi: 10.3899/jrheum.181120] [Medline: 30647183]
50. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
51. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
52. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exerc Health 2021 Mar 4;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]

XSL•FO
RenderX

53. Jacobsen MJ, O'Connor AM, Stacey D. Decisional needs assessment in populations: a workbook for assessing patients' and practitioners' decision making needs. : The Ottawa Hospital (Patient Decision Aids Program); 2013 URL: https://decisionaid.ohri.ca/docs/implement/Population_Needs.pdf [accessed 2022-02-13]
54. Hoffman A, Crocker L, Mathur A, et al. Patients' and providers' needs and preferences when considering fertility preservation before cancer treatment: decision-making needs assessment. JMIR Form Res 2021 Jun 7;5(6):e25083. [doi: 10.2196/25083] [Medline: 34096871]
55. Brinkmann S, Kvale S. Doing Interviews: SAGE Publications; 2007. [doi: 10.4135/9781529716665]
56. Quirkos. URL: https://www.quirkos.com/ [accessed 2026-02-25]
57. Beard JR, Officer A, de Carvalho IA, et al. The World report on ageing and health: a policy framework for healthy ageing. Lancet 2016 May 21;387(10033):2145-2154. [doi: 10.1016/S0140-6736(15)00516-4] [Medline: 26520231]
58. Livne O, Stohl M, Gilman J, Goldberg TE, Wall MM, Hasin DS. Epidemiology of cannabis use among middle-aged and older adults in the U.S. Am J Prev Med 2026 Mar;70(3):108149. [doi: 10.1016/j.amepre.2025.108149] [Medline: 41101401]
59. Dionigi RA, Eime R, Young BW, et al. Coaching older adults (aged 55+). In: Community Sport Coaching Policies and Practice: Routledge; 2021:147-166 URL: https://psycnet.apa.org/record/2022-04796-009 [accessed 2026-02-25]
60. Fakhry SM, Morse JL, Garland JM, et al. Redefining geriatric trauma: 55 is the new 65. J Trauma Acute Care Surg 2021 Apr 1;90(4):738-743. [doi: 10.1097/TA.0000000000003062] [Medline: 33740785]
61. Han X, Zhang C, Guan B, Zhou H, Kong X, Feng S. Burden of osteoarthritis in older adults (aged ≥55 years) in the United States and China: a comparative analysis of temporal trends, risk factor contributions, and projected burden to 2030 using global burden of disease study 2021 data. Front Med (Lausanne) 2025;12:1636976. [doi: 10.3389/fmed.2025.1636976] [Medline: 41035869]
62. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
63. Clarke V, Braun V. Thematic analysis. J Posit Psychol 2017 May 4;12(3):297-298. [doi: 10.1080/17439760.2016.1262613]
64. Hadi MA, José Closs S. Ensuring rigour and trustworthiness of qualitative research in clinical pharmacy. Int J Clin Pharm 2016 Jun;38(3):641-646. [doi: 10.1007/s11096-015-0237-6] [Medline: 26666909]
65. Amin MEK, Nørgaard LS, Cavaco AM, et al. Establishing trustworthiness and authenticity in qualitative pharmacy research. Res Social Adm Pharm 2020 Oct;16(10):1472-1482. [doi: 10.1016/j.sapharm.2020.02.005] [Medline: 32067887]
66. Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999 Dec;34(5 Pt 2):1189-1208. [Medline: 10591279]
67. Miles MB, Huberman AM, Saldaña J. Qualitative Data Analysis: A Methods Sourcebook, 3rd edition: SAGE Publications; 2014. URL: https://eric.ed.gov/?id=ED565763 [accessed 2026-02-25]
68. Farmer T, Robinson K, Elliott SJ, Eyles J. Developing and implementing a triangulation protocol for qualitative health research. Qual Health Res 2006 Mar;16(3):377-394. [doi: 10.1177/1049732305285708] [Medline: 16449687]
69. Nowell LS, Norris JM, White DE, et al. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1):1-13. [doi: 10.1177/1609406917733847]
70. DeCuir-Gunby JT, Marshall PL, McCulloch AW. Developing and using a codebook for the analysis of interview data: an example from a professional development research project. Field methods 2011;23(2):136-155. [doi: 10.1177/1525822X10388468]
71. Birks M, Chapman Y, Francis K. Memoing in qualitative research: probing data and processes. J Res Nurs 2008;13(1):68-75. [doi: 10.1177/1744987107081254]
72. Drisko JW. Transferability and generalization in qualitative research. Res Soc Work Pract 2025 Jan;35(1):102-110. [doi: 10.1177/10497315241256560]
73. Elwyn G, O'Connor A, Stacey D, et al. Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. BMJ 2006 Aug 26;333(7565):417. [doi: 10.1136/bmj.38926.629329.AE] [Medline: 16908462]
74. Volk RJ, Llewellyn-Thomas H, Stacey D, Elwyn G. Ten years of the International Patient Decision Aid Standards Collaboration: evolution of the core dimensions for assessing the quality of patient decision aids. BMC Med Inform Decis Mak 2013;13 Suppl 2(Suppl 2):S1. [doi: 10.1186/1472-6947-13-S2-S1] [Medline: 24624947]
75. Dubois C, Danielson EC, Beestrum M, Eurich DT. Medical cannabis and its efficacy/effectiveness on the management of osteoarthritis pain and function. Curr Med Res Opin 2024 Jul;40(7):1195-1202. [doi: 10.1080/03007995.2024.2363945] [Medline: 38832841]
76. Ager B, Jansen J, Porter D, et al. Development and pilot testing of a Decision Aid (DA) for women with early-stage breast cancer considering contralateral prophylactic mastectomy. Breast 2018 Aug;40:156-164. [doi: 10.1016/j.breast.2018.05.009] [Medline: 29857282]
77. Ankersmid JW, Siesling S, Strobbe LJA, et al. Supporting shared decision-making about surveillance after breast cancer with personalized recurrence risk calculations: development of a patient decision aid using the International Patient Decision AIDS standards development process in combination with a mixed methods design. JMIR Cancer 2022 Nov 14;8(4):e38088. [doi: 10.2196/38088] [Medline: 36374536]
78. Gabel P, Larsen MB, Kirkegaard P, Edwards A, Andersen B. The LEAD trial—the effectiveness of a decision aid on decision making among citizens with lower educational attainment who have not participated in FIT-based colorectal cancer

XSL•FO
RenderX

screening in Denmark: study protocol for a randomized controlled trial. Trials 2018 Oct 10;19(1):543. [doi: 10.1186/s13063-018-2921-z] [Medline: 30305114]
79. Harmsen MG, Steenbeek MP, Hoogerbrugge N, et al. A patient decision aid for risk-reducing surgery in premenopausal BRCA1/2 mutation carriers: development process and pilot testing. Health Expect 2018 Jun;21(3):659-667. [doi: 10.1111/hex.12661] [Medline: 29281161]
80. Ho ES, Parsons JA, Davidge K, Clarke HM, Lawson ML, Wright FV. Developing a decision aid for youth with brachial plexus birth injuries facing treatment decisions for an elbow flexion contracture. PM R 2022 Aug;14(8):971-986. [doi: 10.1002/pmrj.12677] [Medline: 34259390]
81. Lewis KB, Birnie D, Carroll SL, et al. User-centered development of a decision aid for patients facing implantable cardioverter-defibrillator replacement: a mixed-methods study. J Cardiovasc Nurs 2018;33(5):481-491. [doi: 10.1097/JCN.0000000000000477] [Medline: 29601374]
82. Maguire E, Hong P, Ritchie K, Meier J, Archibald K, Chorney J. Decision aid prototype development for parents considering adenotonsillectomy for their children with sleep disordered breathing. J Otolaryngol Head Neck Surg 2016 Nov 4;45(1):57. [doi: 10.1186/s40463-016-0170-2] [Medline: 27809897]
83. Nota I, Drossaert CHC, Melissant HC, et al. Development of a web-based patient decision aid for initiating disease modifying anti-rheumatic drugs using user-centred design methods. BMC Med Inform Decis Mak 2017 Apr 26;17(1):51. [doi: 10.1186/s12911-017-0433-5] [Medline: 28441950]
84. Santerre-Theil A, Bouchard K, St-Pierre D, et al. Development of a tool to guide parents carrying a BRCA1/2 mutation share genetic results with underage children. J Cancer Educ 2018 Jun;33(3):569-575. [doi: 10.1007/s13187-016-1127-x] [Medline: 27804029]
85. Shoemaker SJ, Wolf MS, Brach C. Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. Patient Educ Couns 2014 Sep;96(3):395-403. [doi: 10.1016/j.pec.2014.05.027] [Medline: 24973195]
86. Web Content Accessibility Guidelines (WCAG) 2.1. World Wide Web Consortium (W3C). URL: https://www.w3.org/TR/WCAG21/ [accessed 2025-11-21]
87. Trevena LJ, Zikmund-Fisher BJ, Edwards A, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med Inform Decis Mak 2013;13 Suppl 2(Suppl 2):S7. [doi: 10.1186/1472-6947-13-S2-S7] [Medline: 24625237]
88. Joseph-Williams N, Abhyankar P, Boland L, et al. What works in implementing patient decision aids in routine clinical settings? A rapid realist review and update from the International Patient Decision Aid Standards Collaboration. Med Decis Making 2021 Oct;41(7):907-937. [doi: 10.1177/0272989X20978208] [Medline: 33319621]
89. Charters E. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. Brock Educ J 2003 Jul;12(2). [doi: 10.26522/brocked.v12i2.38]
90. Stacey D, Lewis KB, Smith M, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2024 Jan 29;1(1):CD001431. [doi: 10.1002/14651858.CD001431.pub6] [Medline: 38284415]
91. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation in Industry: CRC Press; 1996:189-194. [doi: 10.1201/9781498710411]
92. Draucker CB, Martsolf DS, Poole C. Developing distress protocols for research on sensitive topics. Arch Psychiatr Nurs 2009 Oct;23(5):343-350. [doi: 10.1016/j.apnu.2008.10.008] [Medline: 19766925]
93. Kaiser K. Protecting respondent confidentiality in qualitative research. Qual Health Res 2009 Nov;19(11):1632-1641. [doi: 10.1177/1049732309350879] [Medline: 19843971]
94. Largent EA, Eriksen W, Barg FK, Greysen SR, Halpern SD. Participants' perspectives on payment for research participation: a qualitative study. Ethics Hum Res 2022 Nov;44(6):14-22. [doi: 10.1002/eahr.500147] [Medline: 36316972]
95. Barry MJ, Edgman-Levitan S. Shared decision making—pinnacle of patient-centered care. N Engl J Med 2012 Mar 1;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
96. Fraenkel L, Bathon JM, England BR, et al. 2021 American College of Rheumatology guideline for the treatment of rheumatoid arthritis. Arthritis Rheumatol 2021 Jul;73(7):1108-1123. [doi: 10.1002/art.41752] [Medline: 34101376]
97. Walter I, Davies H, Nutley S. Increasing research impact through partnerships: evidence from outside health care. J Health Serv Res Policy 2003 Oct;8 Suppl 2:58-61. [doi: 10.1258/135581903322405180] [Medline: 14596749]
98. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015 Nov;13(6):554-561. [doi: 10.1370/afm.1865] [Medline: 26553895]

Abbreviations

CRA: Canadian Rheumatology Association
DA: decision aid
DSI: decision support intervention
GRADE: Grading of Recommendations Assessment, Development, and Evaluation
IPDAS: International Patient Decision Aids Standards

XSL•FO
RenderX

IPDAS: International Patient Decision Aids Standards
ODSF: Ottawa Decision Support Framework
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
RTA: reflexive thematic analysis
SDM: shared decision-making

Edited by J Sarvestan; submitted 18.Apr.2025; peer-reviewed by C Hoffmann, J Sexton; revised version received 09.Jan.2026; accepted 15.Jan.2026; published 30.Mar.2026.

Please cite as:
Aref HAT, Al Hamarneh YN, Kiang TKL, Yacyshyn E, Sadowski CA
Development of a Decision Aid to Support Shared Decision-Making on Cannabis Use for Arthritis: Protocol for a Multiphase Study
JMIR Res Protoc 2026;15:e76237
URL: https://www.researchprotocols.org/2026/1/e76237
doi:10.2196/76237

© Heba A T Aref, Yazid N Al Hamarneh, Tony K L Kiang, Elaine Yacyshyn, Cheryl A Sadowski. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Identifying Predictors of Gambling Episodes and Craving Using Ecological Momentary Assessment and Smartwatch-Based Physiological Measures: Protocol for a Longitudinal Observational Study

Andreas Maximilian Meyer[1*], MSc; Esther Kim[2*], PhD; Hae Kook Lee[3], MD, PhD; Gwanghyun Jo[4], PhD; Michael Patrick Schaub[1], PhD; Severin Haug[1], PhD

[1]Swiss Research Institute for Public Health and Addiction (ISGF), University of Zürich, Konradstrasse 32, Zürich, Switzerland

[2]Department of Counseling Psychology, Korea Baptist Theological University, Daejeon, Republic of Korea

[3]Department of Psychiatry, Uijeongbu St. Mary's Hospital, The Catholic University of Korea, Seoul, Republic of Korea

[4]Department of Mathematical Data Science, Hanyang University ERICA, Ansan-si, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Andreas Maximilian Meyer, MSc
Swiss Research Institute for Public Health and Addiction (ISGF), University of Zürich, Konradstrasse 32, Zürich, Switzerland

Abstract

Background: Like other addictive behaviors, problem gambling is often chronic and relapsing. While digital interventions offer low-threshold treatment and support, their effectiveness is often limited by small effect sizes, low adherence, and high dropout rates. Progress in digital technology has enabled the development of ecological momentary interventions (EMIs), which provide just-in-time support tailored to users' needs. However, EMIs for addictive behaviors have hardly been developed so far.

Objective: This study aimed to explore and identify relevant predictors of gambling episodes assessed by ecological momentary assessments and physiological smartwatch (Apple Watch) data, which in turn may be used for the further development of EMI-based interventions.

Methods: A total of 109 at-risk gamblers were recruited online in a collaborative study between Switzerland and Korea. Over a period of 28 days, participants were asked to complete brief ecological momentary assessment surveys 3 times a day (morning, afternoon, and evening) asking about their gambling behavior and their levels of craving intensity, sleep quality, physical activity, boredom, vitality, depression, and anxiety. They were instructed to wear an Apple Watch that continuously and passively recorded several physiological indicators (eg, heart rate [variability], sleep metrics, and physical activity). Machine learning techniques and multilevel modeling approaches will be used to develop prediction models for gambling episodes and to identify relevant predictors.

Results: Data collection has been completed since April 2025. In total, 109 participants have been enrolled in both countries (52 in Switzerland, 57 in Korea), and datasets are currently being prepared for analysis. The collected data are expected to enable the development of prediction models for gambling episodes.

Conclusions: Incorporating the relevant predictors found in this study into digital intervention programs and providing just-in-time individually tailored intervention elements could improve program engagement and effectiveness. The approach used in this study is transferable to other digital interventions for addictive behaviors and holds promise to exploit their potential.

International Registered Report Identifier (IRRID): DERR1-10.2196/82782

(JMIR Res Protoc 2026;15:e82782) doi:10.2196/82782

KEYWORDS

gambling; craving; biometric monitoring technologies; ecological momentary assessment; ecological momentary intervention; just-in-time adaptive intervention; wearable; smartwatch; prediction; BioMeTs; EMA; EMI; JITAI

Introduction

Mental disorders and addictive behavior contribute significantly to the global burden of disease [1]. Although depressive disorders and the use of addictive substances continue to be highly prevalent, behavioral addictions are becoming increasingly relevant. Particularly, the availability, participation, and cost of gambling have increased dramatically in recent years, with smartphones being used more and more to access online gambling [2]. Gambling is considered a form of

XSL•FO
RenderX

entertainment characterized by betting mechanisms and monetization aspects, and habitual gambling activity has the potential to cause severe distress or impairment [3,4]. The first gambling research that incorporated burden of disease approaches concluded that the harm associated with gambling was comparable to the harms due to depression and alcohol use disorders and even greater than the harm associated with drug dependency and a number of widespread chronic physical conditions [5]. On the individual level, problem gambling is linked to impaired work or academic performance, social deviance, marital problems, emotional and psychological distress, and a variety of mental health conditions, such as personality disorders, alcoholism, anxiety, and mood disorders [6]. However, the harm negatively impacts not only the gamblers themselves but also their families and the community [7,8]. As problem gambling is associated with debt, loss of productivity, and crime, and disrupts family functioning by eroding the time and attention given to the partner, children, and family responsibilities, a person's gambling harms multiple significant other persons, adding to the total burden of gambling harm in the population [9,10]. Gambling disorder was recently included as a new condition in the *ICD-11* (*International Classification of Diseases, 11th Revision*). Particularly economically and socially disadvantaged populations experience gambling disorder and harms more frequently; therefore, problem gambling is likely to aggravate other health and socioeconomic imbalances and disparities [11].

According to estimates for 2016, the prevalence of problem gambling ranged from 0.2% to 5.3% across various nations [3]. Based on the latest representative data from 2022 and diagnostic criteria of the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]), the estimated lifetime prevalence of at-risk gambling in Switzerland is 5.8%, and the prevalence of pathological gambling is 0.8%. In relation to the past 12 months, the estimated prevalence for at-risk gambling is 3.8% and for pathological gambling, 0.5% [12].

Similar to other addictive behaviors, problem gambling is typically chronic and relapsing, with evidence to suggest that over half of all incident problem-gambling cases are previous problem gamblers who are relapsing [13]. Evidence for gambling interventions that can lower opportunities for potentially hazardous gambling has been found; however, there's a lack of screening and brief interventions providing ongoing support to reduce gambling-related harm and relapse [14].

Although face-to-face psychosocial interventions for the treatment of gambling problems show small but statistically significant reductions in gambling behavior [15], only a minority of affected individuals seek professional treatment services, with fewer than 10% of problem gamblers seeking professional help, as obtained in a German study [16]. Potential barriers to obtaining professional help include the lack of services, the stigma attached to addictive behavior, sociocultural obstacles, privacy concerns, and a preference for self-reliance [17]. In comparison to face-to-face treatment, digital interventions have a variety of potential advantages, including ready accessibility, perceived anonymity, individually tailored contents, and low costs. Recent systematic reviews suggest that internet-based treatment has the potential to reach a large proportion of persons with gambling problems and that this approach shows promising effects on general gambling symptoms, gambling frequency, and amount of money lost gambling [18,19]. The majority of the available internet-based interventions rely on principles derived from cognitive behavioral therapy and motivational interviewing and typically include monitoring of gambling behavior, goal setting as well as information on protective behavioral strategies to cope with stress and cravings [18-20]. Although current internet-based interventions are an effective low-threshold treatment option for problem gamblers, their effect sizes are usually small and program adherence is low, including low program use and a high dropout rate. One potential reason for the limited effectiveness and engagement of current digital interventions for problem gambling is their provision of primarily static content and their lack of dynamic tailoring that aims to provide the right type and amount of support to people in their everyday lives [18,19].

In view of this, there is an urgent need for innovative digital approaches that respond more dynamically to the individual needs of those affected and offer support precisely when it is needed most. One such approach—ecological momentary interventions (EMIs)—delivers behavior change support just in time when users most want or need the support. The addition of EMI components to existing digital interventions for the treatment of addictive behaviors holds the promise of increasing their use and efficacy because of the capacity to track the interaction between one's internal states, such as feelings and cravings and the addictive behavior (eg, a gambling episode). The relapse prevention model, a prominent social-cognitive theory developed to explain relapse in substance use disorders, categorizes elements or circumstances that may trigger or aid relapse or addictive behaviors [21]. This model's major premise is that high-risk situations, such as unfavorable emotional states, interpersonal conflict, social pressure, and cravings, precede lapse and relapse. Recent empirical studies indicate that craving for gambling is a multifaceted construct that is characterized by mental imagery, desire thoughts, and physiological sensations and triggered by a variety of stimuli, including positive affect, negative affect, external cues, mental imagery, and desire thoughts [22]. A recent Australian ecological momentary assessment (EMA) study revealed that gambling cravings predicted gambling episodes and craving self-efficacy predicted gambling expenditure [23]. Based on these findings, Hawker et al [24] recently developed a smartphone-delivered EMI, including EMAs comprising 10 self-report items, assessing major constructs of the relapse prevention model 3 times daily. Although results of this pre-post pilot study showed some beneficial intervention effects, the compliance rates for the EMA (51%) and EMI (15%) were low [24]. This result is in line with other studies indicating high attrition and low retention rates in app-delivered interventions for mental health [25].

Although EMIs hold significant therapeutic promise, low user compliance remains an obstacle. To address this issue and simultaneously ensure continuous monitoring of relevant parameters, passive measurement methods, such as biometric monitoring technologies (BioMeTs), are becoming increasingly important as they require less active involvement from users and thus provide a promising alternative to self-reports for EMA

XSL•FO
RenderX

studies and EMIs. According to Goldsack et al [26], BioMeTs are "connected digital medicine products that process data captured by mobile sensors using algorithms to generate measures of behavioral and/or physiological function." Advantages over self-reports include the potential of BioMeTs to be less burdensome (due to passive data collection), less challenging to complete (due to cognitive or disease-specific impairments), less focused on perceived emotional states (given that they rely on physiological or behavioral measurements), less location-reliant (ie, remote assessment is possible), and less stigmatized (eg, when using a clinical assessment). The challenges include that the monitoring devices (eg, wearables, such as wristbands or smartwatches) have to be worn for a longer period and that the physiological and behavioral indicators collected are still poorly validated in some cases [26]. However, there has been tremendous progress recently both in the validation of valid physiological and behavioral indicators and in the availability of low-cost devices, like wristbands, that can collect these indicators with high reliability. The use of wearable devices like wristbands or smartwatches to track mood and behavior brings enormous potential for EMIs. They provide continuous data on behaviors essential to psychological and health assessments, such as sleep and wake cycles, activity, movement, and stress by passively monitoring motion, heart rate, and other physiological factors [27].

At present, this field of research is in the process of development and there are several studies showing associations between depression and digital features from sleep, physical activity, location, and phone use data [27]. However, there are only a few studies in the field of addictive behavior and so far no studies in the field of gambling, although the passive collection of stress via wearables could be particularly relevant. Gambling and gambling cravings are associated with stress in various ways: On the one hand, gambling behavior can serve as a distraction from and a means to cope with stress; on the other hand, gambling-related losses and harm can become a source of stress, which may lead to a feedback cycle leading to more gambling and more stress [28]. Many studies underline the close association between stress and gambling, which is also expressed in the diagnostic criteria for gambling disorder (ie, "gambling when feeling distressed," which is a *DSM-5* criterion for gambling) or a study among gamblers showing that self-reported stress predicted gambling craving over and above reports of anxiety or depression [29].

This study was designed as an international collaboration between the research teams of the Swiss Research Institute for Public Health and Addiction (Switzerland), Uijeongbu St. Mary's Hospital of the Catholic University, and Korea Baptist Theological University (both in Korea) and aims to identify potential physiological and psychological indicators that can reliably predict gambling episodes and cravings, using data collected via self-reported EMA surveys and an Apple Watch. Unlike previous studies focusing mainly on self-reported data, this study additionally leverages the potential of BioMeTs, such as passively collected physiological data from an Apple Watch (eg, heart rate [variability], sleep metrics, and physical activity), to examine their role as predictors of gambling-related outcomes.

Methods

Objectives

This study aims to explore and identify relevant predictors of gambling episodes in a general population sample of at-risk gamblers by:

1. Collecting physiological and psychological indicators (levels of craving intensity, sleep quality, physical activity, stress, vitality, boredom, depression, and anxiety) 3 times a day using brief online EMA self-reports
2. Collecting physiological indicators (eg, heart rate [variability], sleep metrics, and physical activity) passively and continuously throughout the day using a 2024 2nd generation Apple Watch SE (Special Edition)
3. Developing prediction models for gambling episodes using machine learning (ML) algorithm technique, including time series forecasting (TSF) and recurrent neural networks (RNNs), as well as multilevel modeling approaches, such as generalized linear mixed-effects models (GLMMs) and multilevel vector autoregression (mlVAR) models, based on the collected indicators.

Study Design and Procedure

This longitudinal observational prediction study combined prospective observational and noncontact research methods. Self-reported EMA surveys were conducted over a period of 28 days (4 weeks) using individuals' mobile phones to assess gambling behavior and craving-related internal states as well as physiological and psychological variables (emotions, physical activity, stress, vitality, and sleep status) 3 times a day (morning, afternoon, and evening). Over the same period, physiological variables (eg, activity, heart rate, and sleep status) were continuously and passively measured using a 2nd generation Apple Watch SE, including at night to record sleep data. The study procedure is outlined in Figure 1, which provides an overview of the entire process, including the enrollment of participants, the collection of EMA and physiological smartwatch data, and the debriefing at the end of the study.

XSL•FO
RenderX

Figure 1. Study procedure. EMA: ecological momentary assessment.

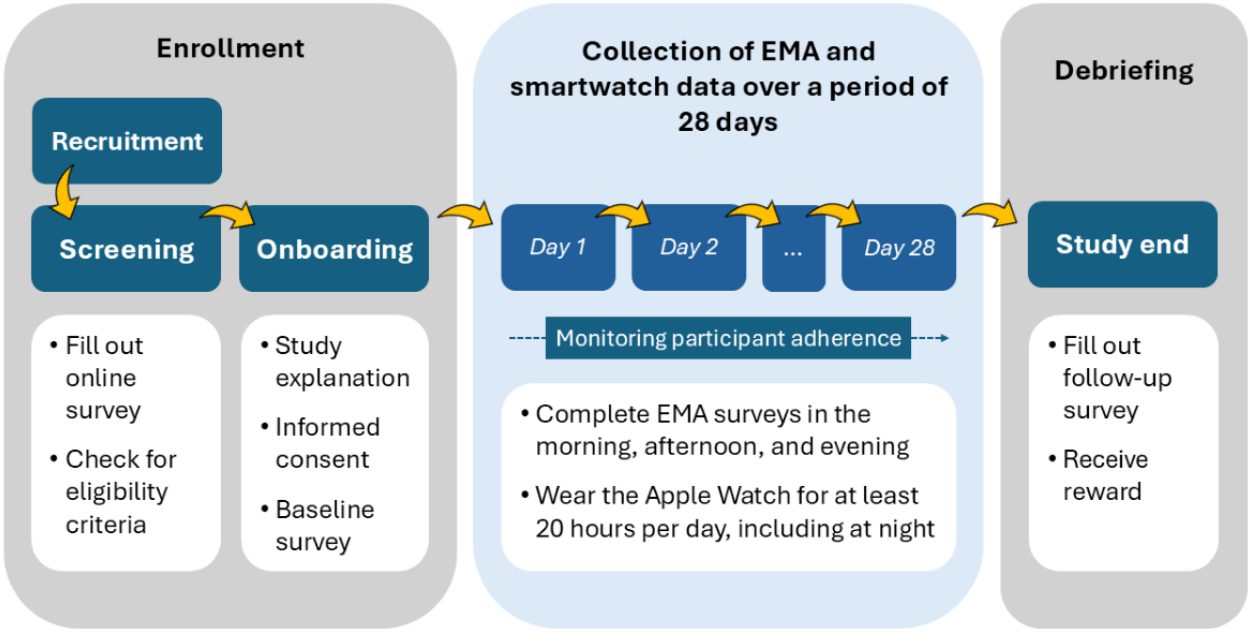

Recruitment and Screening of Study Participants

Between February and April 2025, a total of 52 Swiss participants (mean age 37.4, SD 11.9; range 18 - 79 years; female: n=5, 9.6%) were recruited via a customer satisfaction survey conducted by Swiss Casinos AG that was sent to all online casino guests residing in the German-speaking part of Switzerland. In Korea, a total of 57 participants (mean age 25.6, SD 4.5; range 19 - 33 years; female: n=29, 50.9%) were recruited between October and December 2024 through a professional research agency via online advertisements. Thus, a total number of 109 participants were recruited in both countries, meeting the requirement of reaching the target sample size of 100 participants (refer to "Sample Size" in the "Methods" section for further details regarding the sample size). In both countries, candidates who wished to participate in the study were informed of its aim, and those who consented were screened online for eligibility. The inclusion criteria were (1) age ≥18, (2) a Problem Gambling Severity Index (PGSI) score ≥3, (3) gambling at least once a week, and (4) cognitive capacity to reply to short online EMA surveys and use an Apple Watch. Exclusion criteria were (1) significant suicide risk according to the P4 suicidality screener [30] and (2) current psychiatric illness requiring psychiatric medication or outpatient treatment more than twice a month or current hospitalization in a psychiatric unit.

Onboarding

In both countries, interested participants were screened for eligibility via an online survey to determine if they met the inclusion criteria. Individuals who met the eligibility criteria and expressed interest in participating received detailed information about the purpose of the study, procedures, potential risks and benefits, and participant rights. Those who agreed to participate were invited to a personal onboarding appointment, which was conducted in groups. At the appointment, they were informed of the study, filled out the informed consent form, completed the baseline assessment, and received an Apple Watch. The baseline assessment was a web survey that measured sociodemographic information (eg, age and gender) and problem gambling-related indicators, such as participants' level of problem gambling, gambling symptom severity, frequency of each type of gambling, platform used for gambling (online and offline), and revenue (total profit or loss made gambling). They were also asked about physiological indicators, such as sleep quality, physical activity, and stress, as well as psychological indicators, such as depression and anxiety. Co-occurring symptoms and conditions, such as alcohol use, substance use, and attention deficit and hyperactivity, were also measured (refer to "Baseline and Follow-Up Assessments" in the "Measurements" section for an overview and further details regarding the baseline instruments used in this study). After completing the baseline survey, participants were asked to complete brief EMA self-report surveys, which they received on their smartphones 3 times a day, and wear an Apple Watch for at least 20 hours a day, including at night while sleeping, for 28 days.

Collection of EMA and Smartwatch Data

To capture real-time data about participants' internal states, they received links to brief self-report EMA surveys on their mobile phones three times a day: (1) in the morning from 6 AM to noon, (2) in the afternoon from 3 PM to 5 PM, and (3) in the evening from 8 PM to 10 PM, over a period of 28 days. In case of nonresponding, reminders were sent out at 7:30 AM, 9 AM, and 10:30 AM (for the morning EMA), at 4 PM (for the afternoon EMA), and at 9 PM (for the evening EMA). Participants were asked about their gambling behavior, such as whether they had gambled the previous day or were currently gambling. They were also asked about their craving intensity,

XSL•FO
RenderX

physical activity, sleep quality, current stress levels, vitality and boredom levels, as well as their anxiety and depression levels (refer to "Self-Reported EMA Assessments" in the "Measurements" section for an overview and further details regarding the EMA assessments used in this study). In addition, participants were instructed to wear an Apple Watch for at least 20 hours a day. The device continuously and passively collected their physiological data, such as heart rate, heart rate variability, and sleep metrics, over a period of 28 days (refer to "Measurement of Physiological Indicators" in the "Measurements" section for an overview and further details regarding the physiological indicators collected by the Apple Watch). A 2024 2nd generation Apple Watch SE was used to collect participants' physiological data. It was emphasized that the smartwatch should be worn at night while sleeping in order to record sleep data, such as sleep patterns and cycles. In Switzerland, participants collected relevant physiological data on their Apple Watch using a customized stand-alone app that was specifically designed for the purpose of this study. Swiss participants were required to actively synchronize the data measured and collected throughout the day by the Apple Watch using the study app's synchronization feature once per day (refer to Figure 2 for the data synchronization steps from the participants' perspective). In Korea, participants used a mobile app developed for the collection of both EMA and physiological data. Data from the participants' Apple Watches were accumulated through the mobile app, which was integrated with Apple's Health app. When participants received an EMA notification and accessed the app to complete the brief survey, this action was designed to automatically synchronize the data from the watch.

Figure 2. Customized stand-alone app and its data synchronization steps from the Swiss participants' perspective.

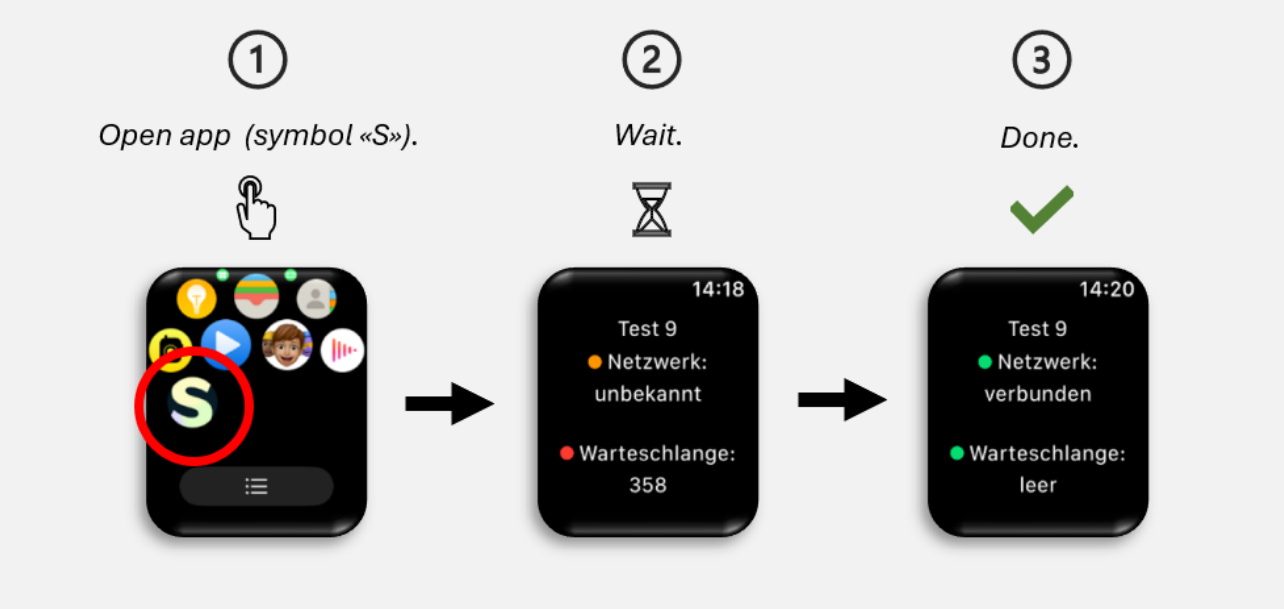

Monitoring Participant Adherence and Compliance

In Switzerland, individuals' adherence to study participation (ie, completing as many of the EMA self-report surveys as possible each day and wearing the Apple Watch for at least 20 hours per day, including at night to record sleep data) was actively monitored. Participants with low adherence were encouraged to complete the EMA self-report surveys and wear the Apple Watch via notifications or proactive phone calls. In Korea, the mobile app was designed to automatically send reminder notifications if a participant failed to complete an EMA survey. Physiological data were also synchronized daily, enabling the research team to monitor if the previous day's data had been collected. For participants who did not meet these collection criteria, a reminder phone call was made to encourage them to wear the Apple Watch.

Debriefing and Follow-Up Assessments

Participants received the follow-up questionnaire after the study period of 28 days had ended (ie, on the 29th day) and used the same measures as in the baseline assessment, along with an additional questionnaire designed to assess whether participants felt influenced by the study itself—for example, whether wearing the Apple Watch affected their sleep or whether the frequent reminders about gambling led to increased gambling behavior (refer to "Baseline and Follow-Up Assessments" in the "Measurements" section for an overview and further details regarding the follow-up instruments used in this study). Participants were then able to choose their participation reward. For active participation, Swiss participants received either a pure monetary reward of up to CHF 400 (US $515) or the Apple Watch, including a pro rata amount of money (the purchase price of the Apple Watch was deducted from the total reward amount), whereas Korean participants received a KRW 100,000 (US $68) cash reward and the study's Apple Watch. The total amount of the reward depended on the level of engagement in the study (eg, the number of daily EMA surveys completed or the number of times participants wore the Apple Watch for at least 20 h/d). In both countries, compensation was determined by compliance, with distinct dropout criteria for EMA and Apple Watch usage. Participants were dropped without compensation if nonadherence—either completing fewer than 3 daily EMAs or wearing the Apple Watch for less than 20 hours—occurred for 2 or more consecutive days or on 3 or more occasions within

XSL•FO
RenderX

a week. These 2 criteria were assessed independently. Participants who avoided dropout and maintained an overall compliance rate of 70% or more.

Hypotheses

Study hypotheses were formulated with respect to the main outcome of interest (ie, whether participants gambled the day before).

1. Psychological indicators (depression, anxiety, and stress) are associated with certain patterns of gambling episodes and cravings, and this information can be used to predict gambling behavior.
2. The physiological indicators (stress, sleep duration and quality, vitality, and physical activity) are associated with certain patterns of gambling episodes and craving, and this information can be used to predict gambling behavior.
3. Physiological indicators continuously measured with the Apple Watch are more predictive of gambling episodes than those measured by a self-reported online EMA survey.

Measurements

Measurements of this study included baseline and follow-up assessments before and after the study, brief self-reported EMA surveys 3 times daily, and continuous recording of the physiological smartwatch data at least 20 hours a day over a period of 28 days. Data were collected independently in both countries using standardized and harmonized procedures and assessment instruments (eg, items of baseline, follow-up, and EMA surveys), allowing for cross-cultural comparison.

Self-Reported EMA Assessments

Overview

Below, the specific EMA items and their respective characteristics are described (refer to Table 1 for an overview of the EMA items and their schedule).

Table . Items for EMA[a] and dissemination schedule.

Item	Time point			Answer options
	Morning (6 AM-noon)	Afternoon (3 PM-5 PM)	Evening (8 PM-10 PM)	
Gambling episode[b]	✓			Dichotomous (yes or no)
Currently gambling	✓	✓	✓	Dichotomous (yes or no)
Craving intensity	✓	✓	✓	Not at all (0) - very severe (5)
Sleep quality[b]	✓			Terrible (0) - excellent (10)
Physical activity[b]	✓			Dichotomous (yes or no)
Stress level	✓	✓	✓	Not at all (0) - very strongly (5)
Vitality level	✓	✓	✓	Not at all (0) - very strongly (5)
Boredom level	✓	✓	✓	Not at all (0) - very strongly (5)
Depression level	✓	✓	✓	Not at all (0) - very severely (5)
Anxiety level	✓	✓	✓	Not at all (0) - very severely (5)

[a]EMA: ecological momentary assessment.

[b]Items "Gambling episode," "Currently gambling," "Sleep quality," and "Physical activity" refer to the previous day and were therefore only asked once per day.

Gambling Episode

The primary outcome of interest, the gambling episode, was assessed during the morning EMA survey. Participants were asked: "Did you engage in gambling yesterday?," a dichotomous variable with response options "Yes" and "No." If participants chose "Yes," they were asked follow-up questions about which platforms they used for gambling (online and offline), the type of gambling they engaged in (eg, casino games, such as blackjack), and the specific times at which they gambled.

Craving

Participants' craving levels were measured 3 times daily (morning, afternoon, and evening) using the item "How strong is your urge to gamble right now?" Responses were recorded on a Likert scale ranging from 0 ("Not at all") to 5 ("Very strong"). Craving was only measured if participants stated that they were not currently gambling at this very moment, using the item "Are you gambling right now?" (answer options: "Yes" or "No").

XSL•FO
RenderX

Physical Activity

Physical activity was measured only during the morning EMA, asking participants "Yesterday, have you done a total of 15 min or more of physical activity, which was enough to raise your breathing rate." Participants could choose from "Yes" or "No" as an answer. When choosing "Yes," a follow-up question that asked participants about how many minutes (range 15-300) they had invested was presented.

Sleep quality

Sleep quality was also assessed only during the morning EMA, asking participants "Today, how would you rate your sleep quality over all (such as how many hours of sleep you got, how easily you fell asleep, how often you woke up during the night, how often you woke up earlier than usual)?" Answers ranged from 0 ("Terrible") to 10 ("Excellent").

Stress

The stress item measured participants' stress level 3 times daily (morning, afternoon, and evening) using the following question: "Stress means a situation in which a person feels tense, restless, nervous, or anxious or is unable to sleep at night because one's mind is troubled all the time. Do you feel this kind of stress right now?" Answers ranged from 0 ("Not at all") to 5 ("Very strongly").

Vitality and Boredom

Vitality and boredom levels will be measured 3 times daily (morning, afternoon, and evening). Questions include "How energetic and rested do you feel right now?" and "How bored are you feeling right now?" Answers ranged from 0 ("Not at all") to 5 ("Very strongly").

Depression and Anxiety

The psychological indicators, depression and anxiety, were measured 3 times daily (morning, afternoon, and evening). Questions include "Right now, how much are you bothered by feeling sad, down, or uninterested in life?" and "Right now, how much are you bothered by feeling anxious and nervous?" Answers ranged from 0 ("Not at all") to 5 ("Very severely").

Measurement of Physiological Indicators

Overview

The physiological indicators measured by the Apple Watch and their specific characteristics are described below (refer to see Table 2 for an overview of the physiological indicators collected by the Apple Watch).

Table . Overview of physiological indicators measured by the smartwatch.

Category and indicator	Measurement frequency	Unit
Vital signs		
Heart rate	Continuously	bpm[b]
Heart rate variability	Sporadically (eg, during resting states)	ms
Sleep		
Time being awake	When sleep state is detected	min
Time being in bed	When sleep state is detected	min
Total time being asleep	When sleep state is detected	min
Core sleep duration	When sleep state is detected	min
Deep sleep duration	When sleep state is detected	min
REM[a] sleep duration	When sleep state is detected	min
Activity		
Stand time	Continuously	min
Step count	Continuously	count
Active energy burned	Continuously	kcal
Distance by walking or running	Continuously	km
Physical activity (eg, workout)	During physical activity (eg, running)	min

[a]REM: rapid eye movement.

[b]bpm: beats per minute.

Vital Signs

Vital signs were recorded using the photoplethysmography sensor of the Apple Watch [31]. Participants' heart rate (HR) was continuously and periodically recorded in beats per minute. When being inactive, HR measures intervals ranging from 5 to 10 minutes. During a workout or movement (eg, when the watch registers being in an active state), the measurement frequency increases to approximately 1-5 seconds. HR data will be used to assess participants' physiological arousal and potential stress responses. The heart rate variability (HRV) was measured using the device's optical sensors (photoplethysmography) to detect interbeat intervals between heartbeats (ie, the time intervals between heartbeats) over a certain period of time (approximately

XSL•FO
RenderX

60 seconds). HRV is calculated by determining the SD of the obtained interbeat intervals in milliseconds. The Apple Watch usually measures HRV during periods of rest or low activity (eg, while falling asleep, upon waking up, or during a breathing exercise). HRV will be used to gain insights into autonomic nervous system functioning and stress regulation.

Sleep Metrics

Sleep metrics were measured in minutes and collected automatically during sleep time using the device's sleep tracking functionality [32]. To estimate sleep stages and sleep-related parameters, the device uses a combination of sensors, including an accelerometer, gyroscope, and HR monitor. The time being awake, the total amount of time a participant is awake during a sleep period, is recorded when the device registers transitions between sleep states and periods of wakefulness. The time being in bed is the total amount of time a participant spends in bed, including both wake and sleep periods. The time being asleep is the total amount of time spent asleep across all 3 sleep stages—core, deep, and REM sleep—all of which are based on HR and movement patterns during the night.

Activity Metrics

Activity metrics were automatically recorded using the Apple Watch SE's built-in sensors, including the accelerometer, gyroscope, and HR monitor [31]. These metrics provide detailed insights into participants' physical activity levels and patterns throughout the day. Stand time measures the amount of time (min) the participant has spent standing. Step count reflects the total number of steps taken by participants throughout the day. Active energy burned indicates the total amount of energy burned through physical activity throughout the day and is measured in kilocalories (kcal). This may include calories burnt from physical workouts or other nonsedentary activities. Both walking and running distances are measured continuously in kilometers (km) and provide information about participants' nonsedentary activities (ie, the distance the user has moved by walking or running). Physical activities, such as running, cycling, or strength training, are actively tracked by the participant or detected automatically by the smartwatch based on movement patterns and HR. Duration (min), type of workout, and intensity were tracked for each recorded session.

Baseline and Follow-Up Assessments

For the purpose of the study, reference periods of the baseline and follow-up instruments were standardized. All instruments were adapted to refer to the past 28 days, regardless of their original wording. Below, the instruments and their specific characteristics of the baseline and follow-up assessment are described (refer to Table 3 for an overview of the baseline and follow-up instruments used in this study and their schedule). In Korea, the respective Korean versions of the instruments were used.

Table . Screening, baseline, and follow-up instruments and their schedule.

Instrument	Screening	Baseline	Follow-up
Sociodemographics		✓	
Problem Gambling Severity Index (PGSI)[a]	✓	(✓)	✓
Gambling Symptom Assessment Scale (G-SAS)		✓	✓
Frequency by Gambling Type		✓	✓
Revenue (Total profit or loss)		✓	✓
Sleep quality		✓	✓
Physical activity (level of exercise)		✓	✓
Perceived Stress Scale (PSS-4)		✓	✓
Patient Health Questionnaire (PHQ-9)		✓	✓
General Anxiety Disorder (GAD-7)		✓	✓
Alcohol Use Disorder Identification Test (AUDIT-C)		✓	✓
Nicotine, Cannabis, and other drugs use		✓	✓
Adult Self-Report ADHD Scale (ASRS)		✓	✓
Suicidality Screener (P4-SCR)	✓		✓
Study Effects Questionnaire			✓

[a]The Problem Gambling Severity Index (PGSI) was used to screen for study eligibility, but also served as a baseline measure.

XSL•FO
RenderX

1. Sociodemographics included gender, age, country of birth (of mother, father, and participant), level of education, partnership status, and marital status.
2. The PGSI contains 9 items that assess a broad array of problems experienced by individuals who engage in problem gambling (eg, feelings of guilt, financial problems, and so on) [33,34]. The PGSI is the most widely used self-report measure of gambling harms in the literature. The total PGSI score was used as a screener in this study to capture participants' level of problem gambling severity, which needed to be ≥3. Answers were given on a 4-point Likert scale ranging from 0 ("Never") to 3 ("Always"). Scores range from 0 to 27 (0=nonproblem gambler, 1 to 2=low risk gambler, 3 to 7=moderate risk gambler, 8+=problem gambler).
3. The Gambling Symptom Assessment Scale (G-SAS) measured the severity of gambling symptoms, focusing on frequency, intensity, and duration of gambling urges, thoughts, behaviors, and both emotional and cognitive impacts [35,36]. Contains 12 items scored 0 (eg, "None") to 4 (eg, "Extreme"). The total score ranges from 0 to 48 (higher scores=greater severity).
4. The Frequency by Gambling Type is a self-developed questionnaire designed to assess participants' engagement in gambling activities, such as lotteries, casino games (eg, blackjack), sports betting, poker, scratch cards, bingo, slot machines outside of casino games (eg, in arcades or bars), and social gambling (eg, private gambling events with friends), tombola, and horse or other racing. The platform used for gambling was also registered (online vs offline). Answers were given on a 5-point Likert scale, ranging from 0 ("Never") to 4 ("4 or more times a week").
5. The Total Profit or Loss item was designed to capture monetary winnings or losses over the last 28 days. In Switzerland, profit or loss was captured in CHF (Swiss Franc), whereas Korea used KRW (Korean Won).
6. Two sleep metrics items were designed to assess participants' sleep quantity and quality, asking participants about how many hours of sleep (0="Less than 5 hours," 5="More than 9 h") they usually get and how they would rate their overall sleep quality during the past 28 days (0="Terrible," 10="Excellent").
7. Two Level of Exercise items assessed participants' physical activity over the last 28 days, asking them on how many days they were physically active for at least 15 minutes and how long they are usually physically active for (in min).
8. The Perceived Stress Scale (PSS-4) is a widely used brief 4-item instrument to assess stress [37,38]. It contains four items scored on a 0 ("Never") to 4 ("Very often") Likert-scale. Scores range from 0 to 16 (higher scores=higher perceived stress).
9. The Patient Health Questionnaire-9 is a reliable and validated 9-question tool developed to assess participants' degree of depression. Answers are given on a 4-point Likert scale (0="never," 3="almost always") [39,40]. Scores range from 0 to 27 (0-4=minimal, 5-9=mild, 10-14=moderate, 15-19=moderately severe, 20-27=severe).
10. The Generalized Anxiety Disorder-7 Scale is a 7-item self-report questionnaire to estimate the severity of generalized anxiety disorder [41,42]. Its items ask about nervousness, inability to stop worrying, excessive worry, restlessness, difficulty relaxing, easy irritation, and the fear of something awful happening. Answers are given on a 4-point Likert scale (0="never," 3="almost every day"). Scores range from 0 to 21 (0-4=minimal, 5-9=mild, 10-14=moderate, 15-21=severe).
11. The Alcohol Use Disorder Identification Test (AUDIT-C) is a 3-item tool screening for risky drinking and alcohol use disorders, focusing on the frequency of drinking, typical amount consumed, and heavy drinking episodes [43,44]. Scores range from 0 to 12 (higher scores=greater risk).
12. Consumption of Nicotine, Cannabis, and Other Drugs assessed single questions measuring the frequency of nicotine, cannabis, and other drug use over the past month. Responses range from 0 ("never") to 4 ("daily or almost daily"). Scores were 0-4 for each substance (higher scores=more frequent use).
13. The Adult Self-Report Scale (ASRS) is a 6-item scale for attention-deficit/hyperactivity disorder (ADHD) screening, including questions about trouble completing tasks and remembering obligations [45,46]. Responses range from 0 ("never") to 4 ("very often") on a 5-point Likert scale. Scores were 0-24 (higher scores=more severe ADHD symptoms).
14. The P4 Suicide Screener is a brief 4-item measure to assess potential suicide risk [30]. If an elevated risk of suicide (scoring ≥ "minimal risk") is recognized at any of the 3 assessments, the participant will not be eligible to participate in the study.
15. A self-developed questionnaire on study effects was designed to assess potential negative impacts of the study on participants' well-being and behavior. Topics include gambling cravings, gambling behavior, and effects from wearing the smartwatch (eg, sleep quality, general well-being, and daily activities). Responses are measured on a 5-point Likert scale (0="Not at all," 4="Very strongly"), with open-ended questions for additional feedback.

Technical Specifications: Data Collection, Storage, and Security

EMA, baseline, and follow-up data were collected and stored using REDCap (Research Electronic Data Capture; Vanderbilt University) version 13.7.13 and MariaDB (MariaDB plc) version 10.11.13. All data were securely stored in Switzerland on servers of our institute's hosting provider (NovaTrend Services GmbH) in accordance with current data protection regulations and using the latest security protocols, including encrypted data transmission (Transport Layer Security) and access controls. In Switzerland, a customized stand-alone app was developed for the 2nd generation Apple Watch SE using Xcode version 16.2 and Swift version 6.0.3 (6.0.3.1.10) to collect the physiological indicators that are relevant for this study (refer to "Measurement of Physiological Indicators" in the "Measurements" section for an overview and further details regarding the physiological indicators collected by the Apple Watch). The data collection process was designed to be entirely independent of a paired iPhone (Apple, Inc), ensuring that all indicators are collected

XSL•FO
RenderX

and processed directly on the Apple Watch. Once collected by the app and synchronized by the participant, the data were transferred to Microsoft Azure Event Hubs, which accepted the incoming data for further processing. The further processed data was eventually stored in a MongoDB (MongoDB, Inc) database.

In Korea, data collection and processing were managed through a dedicated mobile app developed by the Andreia team. Participants installed this app on their phones and completed an in-app consent process to provide access to their health information. Following consent, the app collected physiological data—such as heart rate, heart rate variability, step count, and sleep patterns—from the Apple Watch by integrating with Apple's HealthKit framework. These data were deidentified, linked to a unique participant identifier, and then securely transmitted to an AWS (Amazon Web Services) cloud server running Linux. The entire process was protected by security measures, including Transport Layer Security encryption for data transmission and AWS Security Group with IP-based access controls for the server, ensuring participant privacy in accordance with Apple's data access protocols.

Data Management and Analysis Plan

Overview

The primary outcome of interest will be whether participants report having gambled or not (ie, whether a gambling episode had occurred or not). Thus, the primary goal of our analysis will be to develop models that can reliably predict gambling behavior by using the information from the data that were obtained from the collection of the EMA and physiological smartwatch data. At present, data collection has been completed independently by both the Swiss and Korean research teams. The datasets containing the baseline, follow-up, EMA, and physiological smartwatch data from both countries will be merged to create a combined, cross-national dataset, which will serve as the basis for all analyses outlined in this protocol.

One prediction model we aim to develop will be based on explorative ML techniques, such as TSF or RNNs, to predict gambling episodes. ML models are capable of capturing complex, nonlinear relationships and interactions among multiple psychological and physiological indicators—patterns that may be missed by conventional (linear) approaches. Another advantage is the capability of these models to identify temporally dynamic patterns (eg, changes in craving levels or heart rate variability preceding gambling episodes), allowing us to explore whether short-term fluctuations or longer-term trends in internal states might predict an increased risk for gambling behavior.

Since ML models are sometimes hard to grasp with regard to the specific role of individual predictors (they are often considered "black boxes," hardly seeing which predictor has a stronger effect compared to another), we will also use more classical (linear) models, such as GLMMs, which allow testing for statistical inference and interpretable coefficients and may provide additional insights into whether and how strong a certain predictor (eg, craving) is associated with the outcome (eg, gambling episode). In addition to GLMMs, mlVAR models have been discussed as a promising alternative to other more conventional methods (eg, linear approaches, such as GLMMs) when handling EMA data from addiction research studies [47]. A key advantage of mlVAR in the context of longitudinal data is its ability to explicitly model the temporal and reciprocal relationships between multiple variables across time. Unlike GLMMs, which primarily estimate how predictors at the same time point relate to an outcome in a unidirectional way, mlVAR allows us to examine how changes in psychological and physiological states (eg, craving and HRV) at one time point predict changes in gambling behavior at subsequent time points—and vice versa—providing a dynamic network of bidirectional effects. This enables a more nuanced understanding of dynamic processes and potential feedback loops underlying gambling episodes.

Data Preprocessing

After data collection, the data from both the EMA surveys and the physiological smartwatch sensors undergo several preprocessing steps to ensure data quality and consistency, enabling reliable analysis and modeling. Data preprocessing in this study includes handling missing values, noise filtering, data synchronization, normalization and scaling, feature engineering, and data augmentation.

1. Handling of missing values: Missing values are imputed using simple imputation (mean, median, and mode), time-series interpolation (eg, linear and spline), and ML–based techniques, such as K-Nearest Neighbors and Multiple Imputation by Chained Equations.
2. Noise filtering: Noise from physiological data is filtered using techniques, such as moving averages and Kalman filters.
3. Data synchronization: EMA and physiological data streams are synchronized along a time axis, with resampling applied as needed to create a consistent time series.
4. Normalization and scaling: All variables are adjusted using minimum-maximum scaling, standardization (*z*-scores), and robust scaling.
5. Feature engineering: New features are created from raw data, and the most predictive features are selected.
6. Data augmentation: The dataset is expanded using techniques such as Synthetic Minority Oversampling Technique and various time-series augmentation methods.

Exploratory Data Analysis

Exploratory data analysis (EDA) will be conducted to understand the characteristics of the collected data and identify key variables. EDA will help uncover associations between the data obtained from the EMA assessments (eg, internal states like craving intensity) and the physiological data from the smartwatch (eg, heart rate variability) and assess its significance. EMA and smartwatch data will be visualized to examine distributions, outliers, and missing data to develop a deeper understanding of the data, which in turn may be used for further prediction model development on the primary outcome variable.

Development of the ML Prediction Model

As the biosignal acquired from the smartwatch is basically time-series data, we plan to develop both a deep learning classification model using RNN or a 1D convolutional neural network model and shallow ML models specialized for

XSL•FO
RenderX

time-series dataset, such as TSF to focus on interval information of the data, random interval spectral ensemble to use frequency information, and the shapelet classifier [48] to extract information from phase-independent similarity among biosignals. Those multiple ML classifiers will form a single voting ensemble model to maximize accuracy. The Python (Python Software Foundation) language will be used for EDA and model implementation, with common open-source libraries, including PyTorch, Pandas, and Keras. For model training and validation, a cloud-based server with an Nvidia V100 GPU will be used. The F_1-score, the weighted average of precision and recall, will be used to measure the performance of the classification model, because both accuracy and specificity matter. It depends on the dataset and task, but normally, a classifier with an F_1-score greater than 0.7~0.8 is reported to be good. This classifier can later be used in digital self-help program tasks as well, where participants wearing a smartwatch can receive interventions based on the biosignal containing information about their levels of feelings or cravings measured and uploaded in real time. With the uploaded biosignal, we can send off the participants' necessary nudge interventions according to the level of the internal states of the participants accurately classified by the previously learned ML model.

GLMM Model

Given the hierarchical structure of repeated measurements nested within individuals, we will use GLMMs to analyze our longitudinal data and examine the associations between the collected indicators and gambling outcomes. We will develop models for our primary outcome of interest (ie, the occurrence of gambling episodes) and may develop additional models for secondary outcomes (eg, gambling craving intensity). Predictors will include both time-varying covariates reflecting the EMA data indicators (craving, stress, boredom, vitality, anxiety, and depression) and physiological indicators from the Apple Watch (HR, HRV, sleep metrics, and physical activity levels). The models will include random intercepts to account for individual differences in baseline gambling risk and random slopes to capture individual differences in predictor effects (provided the models achieve satisfactory convergence). Time-lagged versions of predictors will be included to examine temporal precedence. Analyses include both current (same-day effects) and prospective associations (next-day effects) between predictors and outcomes. Country (Switzerland vs Korea) will be included as a between-person factor to explore potential cultural differences in the outcome variables. Model selection will be based on a combination of information criteria (Akaike Information Criterion and Bayesian Information Criterion), theoretical considerations, and model interpretability. A model will be fitted using the *lme4* package (version 1.1.37; [49]) in R (version 4.4.1; R Foundation for Statistical Computing [50]).

mlVAR Model

We will use mlVAR models to examine the dynamic temporal relationships and feedback loops between the indicators collected via EMA (craving, stress, boredom, vitality, anxiety, and depression), physiological indicators collected by the Apple Watch (HR, HRV, sleep metrics, and physical activity levels), and gambling episodes. These models capture both contemporaneous and lagged associations in longitudinal data while accounting for individual differences in these temporal dynamics. The mlVAR approach will model the network of relationships between indicators collected by the EMAs, the physiological indicators collected by the Apple Watch, and gambling episodes. This approach allows us to identify which indicators predict changes in other variables over time (temporal networks), which variables co-occur within the same measurement occasion (contemporaneous networks), and how these relationships vary between participants (between-person networks). We will estimate separate mlVAR models for the Swiss and Korean samples to explore potential cultural differences in the dynamic interplay of predictors. Time-lagged effects will be examined using multiple intervals (t-1, t-2) to capture both immediate and delayed influences on gambling behavior. The analysis will reveal temporal patterns and reciprocal influences between indicators that can be used to develop just-in-time adaptive interventions (JITAIs) or EMIs, respectively. All mlVAR models will be fitted with the *mlVAR* package (version 0.5.2; [51]) in R (version 4.4.1; [50]).

Sample Size

In the case of deep learning, approximately 5000 pieces of learning data per category are required to show acceptable performance [52]. Thus, the data size generally requires 1000-10,000 observations, and using 10,000 observations as a development dataset can detect a 0.1% performance improvement [53]. Therefore, the sample size of this study will be 100 participants (50 in Korea and 50 in Switzerland). If 100 participants are tracked for 28 days, 8400 records are generated (100 participants×28 d×3 EMAs per day), which is sufficient for ML modeling even when missing data is taken into account [54]. For multilevel models, studies suggest that for moderate effect sizes, a minimum of 50 participants (cluster level) with 20-30 observations provides adequate power (>0.80) to detect within-person effects [55,56].

Ethical Considerations

The study has been approved by the Ethics Committee of the Faculty of Arts and Sciences at the University of Zürich (approval number 24.05.11; date of approval June 6, 2024) (Multimedia Appendix 1; Peer Review Report 1; Peer Review Report 2; Peer Review Report 3) and the Institutional Review Board (IRB) of Uijeongbu St. Mary's Hospital, Catholic University of Korea (IRB number UIRB-정20240725 - 022; date of approval July 26, 2024). The study will be executed in compliance with the Helsinki Declaration. All participants will be provided with an informed consent form and given the opportunity to decide whether or not to participate. The informed consent form provides detailed information about the purpose, procedures, potential risks, compensation, and privacy protections associated with the study. All data will be anonymized to ensure that individual subjects cannot be identified. The data collected in this study will be used for research purposes only. Access to the data will be limited to the principal investigators and authorized researchers.

XSL•FO
RenderX

Results

At the time of the protocol submission, recruitment and data collection have been completed by both the Swiss and Korean research teams (IRRID: DERR1). In Switzerland, a total number of 52 Swiss participants (mean age 37.4, SD 11.9; range 18 - 79 years; female: n=5, 9.6%) were recruited between February and April 2025. In Korea, a total number of 57 participants (mean age 25.6, SD 4.5, range 19 - 33 years; female: n=29, 50.9%) were recruited between October and December 2024. The datasets of both countries are currently being merged and prepared for all further statistical analyses (eg, the development of prediction models), which will be performed as outlined in the "Methods" section once the data merging is complete and a combined cross-national dataset is created. Hypothesis testing has not been conducted yet and will be conducted once the dataset is merged and finalized. Results will be published in a scientific peer-reviewed journal. At present, only descriptive statistics (eg, sample characteristics) have been analyzed for each national sample and have been used to present preliminary results at conferences. The Korean team has previously tested their ML model for internal purposes using the dataset from the Korean participants; the Swiss sample was not included. Anonymized study data will be available on request. Participants will be informed via email about study results via a layperson-friendly summary of trial findings if they have requested so at registration.

Discussion

Principal Findings

Building on the findings of the first EMA and EMI studies in the field of gambling [23,24], this study will be the first to examine the influence of potentially relevant physiological indicators that can be collected continuously and nonintrusively using wearable devices (ie, an Apple Watch). In addition to the wealth of EMA and physiological wearable data (BioMeTs) collected, the planned study is innovative in its use of ML methodology to develop prediction models. Unlike traditional regression models, which rely heavily on linear relationships between variables, this study uses advanced ML methodologies, such as RNNs and TSFs, to analyze complex, nonlinear interactions between physiological and psychological indicators. This approach is expected to yield predictive models with higher accuracy, thus contributing to the growing field of digital health interventions for behavioral addictions. Since the task involves building an accurate classification model for use in early intervention (eg, EMIs), we opted for ML methods.

As clients rarely report addiction problems accurately, determining the timing and intensity of appropriate intervention is difficult. Therefore, if a model based on EMA and/or physiological wearable device data could be created that predicts gambling behavior with high accuracy, this would be a valuable technique for enabling early interventions with clients who have (behavioral) addiction problems, paving the way for JITAIs, such as EMIs. Additionally, developing physiological data analysis that can predict behavioral addiction through ML techniques would provide important evidence for the development of an integrated biopsychosocial behavioral addiction model. The use of wearable devices also addresses a significant limitation in addiction research: reliance on self-reported data. Wearables provide a less intrusive, more reliable alternative by passively collecting continuous physiological data. This capability enables more precise and timely detection of changes in participants' stress levels [57,58], sleep patterns [59,60], and physical activity, all of which are closely linked to gambling behavior.

Expected Results

A significant advantage of this research lies in its potential for providing early interventions just in time for individuals who need it the most. It is expected that variables collected via EMAs and wearable devices (in our case, the Apple Watch) will enable the development of valid prediction models for gambling episodes and cravings. The incorporation of the relevant variables in existing digital self-help interventions that address problem gambling and the provision of just-in-time, individually tailored intervention elements could improve program engagement and effectiveness. This allows for objective prediction before a gambling episode becomes "serious," allowing for effective gambling episode prevention. It is expected that stress in particular, measured by HR and its variability, physical activity, and both sleep duration and quality could be good predictors of gambling cravings and gambling episodes [57-60]. We assume that these physiological indicators have a higher predictive value than the corresponding self-reported constructs measured via EMAs because the device continuously, passively, and nonintrusively measures and collects indicators throughout the day, without individuals having to report their current state actively (eg, stress levels). This goes along with the assumption that fluctuations in certain states can be more easily detected compared to self-reported EMA surveys, which usually require individuals to complete short questionnaires several times throughout the day to report their current states.

Challenges and Future Directions

Despite the expectation of promising results, the study may also face several challenges. It is essential to implement robust engagement strategies to ensure that participants comply with the instructions to wear the Apple Watch for extended periods of time and complete the EMA surveys. To mitigate the risk of low participant engagement, we implemented several measures, such as a monitoring strategy, reminders, and proactive follow-ups, as well as an incentive plan.

Moreover, the efficacy of the ML models is contingent upon the quality and consistency of the data collected. The presence of missing or erroneous data could potentially compromise the performance of the model, necessitating the implementation of advanced preprocessing techniques.

Furthermore, as a cross-national study between Switzerland and Korea, this research faces challenges arising from methodological differences. For instance, the data synchronization method and the compensation scheme varied between the 2 countries. In Switzerland, participants had to manually synchronize their data once a day, whereas in Korea,

XSL•FO
RenderX

data were synchronized automatically when participants accessed the app for EMA surveys. Additionally, the Swiss cohort received a graded reward based on their level of engagement, while the Korean cohort was subject to strict dropout criteria, receiving a uniform reward only upon maintaining high compliance. These different incentive structures and synchronization methods may have uniquely influenced participant motivation and data collection patterns in each country. Acknowledging these differences is crucial for interpreting the study's findings.

Future research could investigate the integration of supplementary data sources, such as geolocation or social media activity, with the aim of enhancing the predictive accuracy of the model.

Broader Impact

The insights gained from this study have implications beyond gambling addiction. The methodologies and findings could be applied to other behavioral and substance-related addictions, thereby enhancing the scope of digital health interventions. Wearable technology is already widespread and offers a scalable solution to public health challenges by enabling remote monitoring and intervention. Consequently, this study aligns with global efforts to leverage technology to improve mental health outcomes and reduce health disparities. Identifying good physiological indicators for predicting mental health or addictive behaviors, such as depressive episodes, nicotine use, or alcohol consumption, could take the personalization, adherence, and effectiveness of digital interventions to the next level. In particular, populations that encounter obstacles when accessing conventional mental health services, such as those living in remote regions or individuals who are stigmatized, could benefit greatly from wearable-enabled interventions. The anonymized, low-threshold nature of these interventions enhances their acceptability and potential reach.

Conclusion

This longitudinal, cross-national observational study combines EMAs with continuously collected physiological data from Apple Watches to identify predictors of gambling episodes and gambling craving in at-risk gamblers. Self-reported states (eg, levels of craving, stress, boredom, vitality, anxiety, and depression) will be combined with passively collected biosignals (eg, heart rate, heart rate variability, sleep metrics, and activity indicators). The goal is to describe time-varying risk patterns that precede gambling. Using two complementary analytic approaches—machine learning for high-dimensional prediction and multilevel modeling for interpretable within-person and between-person effects—we will develop and compare prediction models for short-term gambling risk. The findings are expected to enable the development of just-in-time adaptive ecological momentary interventions that provide targeted support when risk is elevated (ie, when participants are in need of support) while keeping user burden low. The cross-national design also allows for examining potential cultural differences in predictors and model performance. Overall, this protocol provides a scalable framework for using wearable and EMA data to improve digital interventions for gambling and other addictive behaviors.

Acknowledgments

The Swiss research team would like to thank Polina Veltmann and Romain Meisterhans from Smilamind for their active support in the development of the stand-alone application for the Apple Watch. The Korean research team would like to thank the Andreia team in Korea for their valuable contributions to the app development.

Funding

Funding for this project was provided by the Swiss National Science Foundation (SNSF; grant number IZKSZ1_218534/1) and the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT South Korea; grant number RS-2023-00233991).

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Decision letter from the funding agency.
[PDF File, 150 KB - resprot_v15i1e82782_app1.pdf]

Peer Review Report 1
Peer-review report nr. 1 from a granting agency.
[PDF File, 93 KB - resprot_v15i1e82782_app2.pdf]

Peer Review Report 2

XSL•FO
RenderX

Peer-review report nr. 2 from granting agency.
[PDF File, 94 KB - resprot_v15i1e82782_app3.pdf]

Peer Review Report 3
Peer-review report nr. 3 from a granting agency.
[PDF File, 96 KB - resprot_v15i1e82782_app4.pdf]

References

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
2. Abbott MW. Gambling and gambling-related harm: recent World Health Organization initiatives. Public Health (Fairfax) 2020 Jul;184:56-59. [doi: 10.1016/j.puhe.2020.04.001] [Medline: 32402594]
3. Calado F, Griffiths MD. Problem gambling worldwide: an update and systematic review of empirical research (2000-2015). J Behav Addict 2016 Dec;5(4):592-613. [doi: 10.1556/2006.5.2016.073] [Medline: 27784180]
4. Santangelo OE, Provenzano S, Firenze A. Risk factors for addictive behaviors: a general overview. Int J Environ Res Public Health 2022 May 28;19(11):6583. [doi: 10.3390/ijerph19116583] [Medline: 35682170]
5. Browne M, Greer N, Rawat V, Rockloff M. A population-level metric for gambling-related harm. Int Gambl Stud 2017 May 4;17(2):163-175. [doi: 10.1080/14459795.2017.1304973]
6. Lorains FK, Cowlishaw S, Thomas SA. Prevalence of comorbid disorders in problem and pathological gambling: systematic review and meta-analysis of population surveys. Addiction 2011 Mar;106(3):490-498. [doi: 10.1111/j.1360-0443.2010.03300.x] [Medline: 21210880]
7. Browne M, Rockloff MJ. Prevalence of gambling-related harm provides evidence for the prevention paradox. J Behav Addict 2018 Jun 1;7(2):410-422. [doi: 10.1556/2006.7.2018.41] [Medline: 29788761]
8. Hing N, Russell AMT, Browne M, et al. Gambling-related harms to concerned significant others: a national Australian prevalence study. J Behav Addict 2022 Jul 13;11(2):361-372. [doi: 10.1556/2006.2022.00045] [Medline: 35895474]
9. Dowling NA, Oldenhof E, Cockman S, Suomi A, Merkouris SS, Jackson AC. Problem gambling and family violence: factors associated with family violence victimization and perpetration in treatment-seeking gamblers. J Interpers Violence 2021 Aug;36(15-16):7645-7669. [doi: 10.1177/0886260519835877] [Medline: 30894040]
10. Latvala T, Lintonen T, Konu A. Public health effects of gambling - debate on a conceptual model. BMC Public Health 2019 Aug 9;19(1):1077. [doi: 10.1186/s12889-019-7391-z] [Medline: 31399026]
11. Sharman S, Butler K, Roberts A. Psychosocial risk factors in disordered gambling: a descriptive systematic overview of vulnerable populations. Addict Behav 2019 Dec;99:106071. [doi: 10.1016/j.addbeh.2019.106071] [Medline: 31473572]
12. Wenger A, Haug S. Geldspiel: verhalten und problematik in der schweiz im jahr 2022 [webpage in german]. Schlussbericht. 2024. URL: https://www.esbk.admin.ch/dam/de/sd-web/iNRy0OflhS69/schlussber-studie-gluecksspielverhalten-2022-d.pdf [accessed 2026-02-05]
13. Smith DP, Battersby MW, Pols RG, Harvey PW, Oakes JE, Baigent MF. Predictors of relapse in problem gambling: a prospective cohort study. J Gambl Stud 2015 Mar;31(1):299-313. [doi: 10.1007/s10899-013-9408-3] [Medline: 24065314]
14. Blank L, Baxter S, Woods HB, Goyder E. Interventions to reduce the public health burden of gambling-related harms: a mapping review. Lancet Public Health 2021 Jan;6(1):e50-e63. [doi: 10.1016/S2468-2667(20)30230-9] [Medline: 33417847]
15. Quilty LC, Wardell JD, Thiruchselvam T, Keough MT, Hendershot CS. Brief interventions for problem gambling: a meta-analysis. PLoS ONE 2019;14(4):e0214502. [doi: 10.1371/journal.pone.0214502] [Medline: 30995229]
16. Loy JK, Grüne B, Braun B, Samuelsson E, Kraus L. Help-seeking behaviour of problem gamblers: a narrative review. Sucht 2018 Dec 1;64(5-6):259-272. [doi: 10.1024/0939-5911/a000560]
17. Clarke D, Abbott M, DeSouza R, Bellringer M. An overview of help seeking by problem gamblers and their families, including barriers to and relevance of services. Int J Ment Health Addiction 2007 Oct 22;5(4):292-306. [doi: 10.1007/s11469-007-9063-y]
18. Rodda SN. A systematic review of internet delivered interventions for gambling: prevention, harm reduction and early intervention. J Gambl Stud 2022 Sep;38(3):967-991. [doi: 10.1007/s10899-021-10070-x] [Medline: 34515903]
19. Sagoe D, Griffiths MD, Erevik EK, et al. Internet-based treatment of gambling problems: a systematic review and meta-analysis of randomized controlled trials. J Behav Addict 2021 Sep 17;10(3):546-565. [doi: 10.1556/2006.2021.00062] [Medline: 34546971]
20. Boumparis N, Haug S, Abend S, Billieux J, Riper H, Schaub MP. Internet-based interventions for behavioral addictions: a systematic review. J Behav Addict 2022 Sep 26;11(3):620-642. [doi: 10.1556/2006.2022.00054] [Medline: 36495471]
21. Marlatt GA, Donovan DM. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors: Guilford press; 2005.

XSL•FO
RenderX

22. Cornil A, Lopez-Fernandez O, Devos G, de Timary P, Goudriaan AE, Billieux J. Exploring gambling craving through the elaborated intrusion theory of desire: a mixed methods approach. Int Gambl Stud 2018 Jan 2;18(1):1-21. [doi: 10.1080/14459795.2017.1368686]
23. Hawker CO, Merkouris SS, Youssef GJ, Dowling NA. Exploring the associations between gambling cravings, self-efficacy, and gambling episodes: an ecological momentary assessment study. Addict Behav 2021 Jan;112:106574. [doi: 10.1016/j.addbeh.2020.106574] [Medline: 32759020]
24. Hawker CO, Merkouris SS, Youssef GJ, Dowling NA. A smartphone-delivered ecological momentary intervention for problem gambling (gamblingless: curb your urge): single-arm acceptability and feasibility trial. J Med Internet Res 2021 Mar 26;23(3):e25786. [doi: 10.2196/25786] [Medline: 33769294]
25. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
26. Goldsack JC, Coravos A, Bakker JP, et al. Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). NPJ Digit Med 2020;3(55). [doi: 10.1038/s41746-020-0260-4] [Medline: 32337371]
27. De Angel V, Lewis S, White K, et al. Digital health tools for the passive monitoring of depression: a systematic review of methods. NPJ Digit Med 2022 Jan 11;5(1):3. [doi: 10.1038/s41746-021-00548-8] [Medline: 35017634]
28. Buchanan TW, McMullin SD, Baxley C, Weinstock J. Stress and gambling. Curr Opin Behav Sci 2020 Feb;31:8-12. [doi: 10.1016/j.cobeha.2019.09.004]
29. Elman I, Tschibelu E, Borsook D. Psychosocial stress and its relationship to gambling urges in individuals with pathological gambling. American J Addict 2010 Jul;19(4):332-339 [FREE Full text] [doi: 10.1111/j.1521-0391.2010.00055.x]
30. Dube P, Kurt K, Bair MJ, Theobald D, Williams LS. The P4 screener: evaluation of a brief measure for assessing potential suicide risk in 2 randomized effectiveness trials of primary care and oncology patients. Prim Care Companion J Clin Psychiatry 2010;12(6):PCC.10m00978. [doi: 10.4088/PCC.10m00978blu] [Medline: 21494337]
31. Using apple watch to measure heart rate, calorimetry, and activity. : Apple Inc; 2024 URL: https://www.apple.com/health/pdf/Heart_Rate_Calorimetry_Activity_on_Apple_Watch_November_2024.pdf [accessed 2026-01-29]
32. Estimating sleep stages from apple watch. : Apple Inc; 2025 URL: https://www.apple.com/health/pdf/Estimating_Sleep_Stages_from_Apple_Watch_Oct_2025.pdf [accessed 2026-01-29]
33. Ferris JA, Wynne HJ. The Canadian Problem Gambling Index. : Canadian Centre on Substance Abuse; 2001 URL: https://www.greo.ca/Modules/EvidenceCentre/files/Ferris%20et%20al(2001)The_Canadian_Problem_Gambling_Index.pdf [accessed 2026-01-29]
34. Kim AY, Cha JE, Kwon SJ, et al. Construction and validation of Korean version of CPGI. Korean Journal of Psychology: General 2011;30(4):1011-1038.
35. Kim SW, Grant JE, Potenza MN, Blanco C, Hollander E. The Gambling Symptom Assessment Scale (G-SAS): a reliability and validity study. Psychiatry Res 2009 Mar 31;166(1):76-84. [doi: 10.1016/j.psychres.2007.11.008] [Medline: 19200607]
36. KIM HJ, KIM JH, SHIN YC, Shin HC, Grant JE, Lee TK. The reliability and validity of the Korean translation of the Gambling Symptom Assessment Scale (KG-SAS). J Korean Neuropsychiatr Assoc 2005;44(6):682-689 [FREE Full text]
37. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
38. Lee J, Shin C, Ko YH, et al. The reliability and validity studies of the Korean version of the Perceived Stress Scale. Korean J Psychosom Med 2012;20(2):127-134 [FREE Full text]
39. Spitzer RL, Kroenke K, Williams JBW. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
40. Choi HS, Choi JH, Park KH, et al. Standardization of the Korean version of the Patient Health Questionnaire-9 as a screening instrument for major depressive disorder. Korean J Fam Med 2007;28(2):114-119 [FREE Full text]
41. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
42. Seo JG, Park SP. Validation of the Generalized Anxiety Disorder-7 (GAD-7) and GAD-2 in patients with migraine. J Headache Pain 2015 Dec;16(1):1-7. [doi: 10.1186/s10194-015-0583-8]
43. Bush K, Kivlahan DR, McDonell MB. The AUDIT Alcohol Consumption Questions (AUDIT-C): an effective brief screening test for problem drinking. Arch Intern Med 1998 Sep 14;158(16):1789. [doi: 10.1001/archinte.158.16.1789]
44. Seong JH, Lee CH, Do HJ, et al. Performance of the AUDIT Alcohol Consumption Questions (AUDIT-C) and AUDIT-K Question 3 alone in screening for problem drinking. Korean J Fam Med 2009;30(9):695. [doi: 10.4082/kjfm.2009.30.9.695]
45. Kessler RC, Adler L, Ames M, et al. The World Health Organization adult ADHD self-report scale (ASRS): a short screening scale for use in the general population. Psychol Med 2005 Feb;35(2):245-256. [doi: 10.1017/S0033291704002892]
46. Kim JH, Lee EH, Joung YS. The WHO Adult ADHD Self-Report Scale: reliability and validity of the Korean version. Psychiatry Investig 2013 Mar;10(1):41-46. [doi: 10.4306/pi.2013.10.1.41] [Medline: 23482673]

XSL•FO
RenderX

47. Serre F, Gauld C, Lambert L, et al. Predictors of substance use during treatment for addiction: a network analysis of ecological momentary assessment data. Addiction 2025 Jan;120(1):48-58. [doi: 10.1111/add.16658] [Medline: 39210697]
48. Lines J, Davis LM, Hills J, Bagnall A. A shapelet transform for time series classification. Presented at: KDD '12: Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining; Aug 12-16, 2012. [doi: 10.1145/2339530.2339579]
49. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Soft 2015;67(1):1-48. [doi: 10.18637/jss.v067.i01]
50. R Core Team. R: a language and environment for statistical computing. R Foundation for Statistical Computing. 2024. URL: https://www.r-project.org/ [accessed 2025-01-29]
51. Epskamp S, van Borkulo CD, van der Veen DC, et al. Personalized network modeling in psychopathology: the importance of contemporaneous and temporal connections. Clin Psychol Sci 2018 May;6(3):416-427. [doi: 10.1177/2167702617744325] [Medline: 29805918]
52. Goodfellow I, Bengio Y, Courville A. Deep Learning: The MIT Press; 2016.
53. Andrew NG. Machine Learning Yearning: Technical Strategy for AI Engineers, in the Era of Deep Learning: Self-publishing; 2018.
54. Yang J, Kershaw KN. Feasibility of using ecological momentary assessment and continuous heart rate monitoring to measure stress reactivity in natural settings. PLoS ONE 2022;17(3):e0264200. [doi: 10.1371/journal.pone.0264200]
55. Bolger N, Laurenceau JP. Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research: Guilford press; 2013.
56. Maas CJM, Hox JJ. Sufficient sample sizes for multilevel modeling. Methodology (Gott) 2005 Jan;1(3):86-92. [doi: 10.1027/1614-2241.1.3.86]
57. Ford M, Håkansson A. Problem gambling, associations with comorbid health conditions, substance use, and behavioural addictions: opportunities for pathways to treatment. PLoS ONE 2020;15(1):e0227644. [doi: 10.1371/journal.pone.0227644] [Medline: 31923269]
58. Pangborn N, Zhang E, Balodis IM. A systematic review of stress physiology in gambling disorder and problem gambling. Curr Behav Neurosci Rep 2024;11(3):182-200. [doi: 10.1007/s40473-024-00279-6]
59. Parhami I, Siani A, Rosenthal RJ, Fong TW. Pathological gambling, problem gambling, and sleep complaints: an analysis of the National Comorbidity Survey Replication (NCS-R). J Gambl Stud 2013 Jun;29(2):241-253. [doi: 10.1007/s10899-012-9299-8] [Medline: 22396174]
60. Loft MH, Loo JMY. Understanding the mechanisms underlying gambling behavior and sleep. J Gambl Stud 2015 Dec;31(4):1273-1286. [doi: 10.1007/s10899-014-9514-x] [Medline: 25381635]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
ASRS: Adult Self-Report Scale
AUDIT-C: Alcohol Use Disorder Identification Test
AWS: Amazon Web Services
BioMeT: biometric monitoring technology
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders* (Fifth Edition)
EDA: exploratory data analysis
EMA: ecological momentary assessment
EMI: ecological momentary intervention
G-SAS: Gambling Symptom Assessment Scale
GLMM: generalized linear mixed-effects model
HR: heart rate
HRV: heart rate variability
***ICD-11*:** *International Classification of Diseases 11th Revision*
IRB: Institutional Review Board
JITAI: just-in-time adaptive intervention
ML: machine learning
mlVAR: multilevel vector autoregression
PGSI: Problem Gambling Severity Index
PSS-4: Perceived Stress Scale
REDCap: Research Electronic Data Capture
RNN: recurrent neural network
TSF: time series forecasting

XSL•FO
RenderX

Edited by J Sarvestan; submitted 01.Sep.2025; this is a non–peer-reviewed article;accepted 05.Dec.2025; published 19.Feb.2026.

Please cite as:
Meyer AM, Kim E, Lee HK, Jo G, Schaub MP, Haug S
Identifying Predictors of Gambling Episodes and Craving Using Ecological Momentary Assessment and Smartwatch-Based Physiological Measures: Protocol for a Longitudinal Observational Study
JMIR Res Protoc 2026;15:e82782
URL: https://www.researchprotocols.org/2026/1/e82782
doi:10.2196/82782

© Andreas Maximilian Meyer, Esther Kim, Hae Kook Lee, Gwanghyun Jo, Michael Patrick Schaub, Severin Haug. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Implementing an eHealth Model of Care for Pediatric Patients and Families at the End of Treatment for Acute Lymphoblastic Leukemia (EMERGE): Type 2 Nonrandomized Hybrid Implementation-Effectiveness Trial Study Protocol

Maria C McCarthy[1,2,3], BA, MA, PhD; Chris Williams[1,2,4], BSN, MA; Michelle Tennant[1,3], BA, BCom, GDip, DPsych; Richard De Abreu Lourenco[5], BEc, MEc, PhD; Hannah Pring[1,4], BA, MA; Katie Moore[1,4,6], MBBS, MA; Jane Templeton[1], BSN; Ken Knight[1,2], BA(Hons), GDip Psych, MA; Peter Downie[4,6,7], MBBS; Stephen Hearps[1,8], BA, PhD; Cinzia De Luca[1,2,3], BA, BSc, MA, PhD

[1]Clinical Sciences, Murdoch Children's Research Institute, 50 Flemington Road, Parkville, Australia

[2]Department of Paediatrics, The University of Melbourne, Parkville, Australia

[3]Children's Cancer Centre, Royal Children's Hospital, Parkville, Australia

[4]Victorian Paediatric Integrated Cancer Service, Parkville, Australia

[5]Centre for Health Economics Research and Evaluation, University of Technology, Sydney, Australia

[6]Children's Cancer Centre, Monash Health, Clayton, Australia

[7]Department of Paediatrics, Monash University, Clayton, Australia

[8]Department of Critical Care, University of Melbourne, Parkville, Australia

Corresponding Author:
Maria C McCarthy, BA, MA, PhD
Clinical Sciences, Murdoch Children's Research Institute, 50 Flemington Road, Parkville, Australia

Abstract

Background: Despite increasing survival rates for childhood cancers, physical and psychological late effects are common. The end-of-treatment period is recognized as a complex transition period, and there are few evidence-based models of care to address patient and family needs during this early survivorship period. The EMERGE model of care has been developed to provide eHealth-delivered, multidisciplinary care to patients and families in the 12 months following treatment for acute lymphoblastic leukemia, the most common type of pediatric cancer.

Objective: The primary aim of this study is to assess the implementation success of the EMERGE model of care into the clinical setting. Secondary aims include evaluating effectiveness and cost consequences.

Methods: The study uses a nonrandomized hybrid implementation-effectiveness design, assessing both implementation and clinical outcomes. Implementation metrics include evaluating the reach, acceptability, feasibility, and maintenance of the EMERGE model. Clinical effectiveness outcomes include parent satisfaction with the EMERGE intervention and pre-post intervention evaluation of parent psychological stress and unmet information needs. The Reach, Effectiveness, Adoption, Implementation, and Maintenance implementation science framework was used to guide study outcomes. Semistructured interviews with clinicians and parents will further evaluate the acceptability and sustainability of the EMERGE model and appraise barriers and facilitators to implementation. Cost analysis will include evaluation of the resources required for program delivery and the impact on subsequent health care service use measured using Medicare data and health service usage collected during the EMERGE intervention.

Results: The trial commenced in December 2022, and recruitment concluded in October 2025, with 81 families recruited. Data collection is ongoing and is anticipated to be completed in Summer 2026.

Conclusion: This study will address a critical gap in multidisciplinary care delivery at the end of treatment for young survivors of acute lymphoblastic leukemia and their families. The EMERGE model has the potential to improve the quality of life of patients and families by providing an early survivorship intervention. Importantly, the usage of an eHealth (telehealth) model will enable distance-delivered care, facilitating family participation regardless of geography. By measuring implementation, clinical, and cost impacts, this study will inform the future development of end-of-treatment models of care that are almost universally lacking in pediatric oncology care.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12623000761662; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=381671&isReview=true

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/85901

(JMIR Res Protoc 2026;15:e85901) doi:10.2196/85901

KEYWORDS

childhood cancer; end-of-treatment; model of care; intervention; telehealth; e-health

Introduction

There is a critical gap in care for young cancer survivors after they complete their primary treatment [1,2]. Typically, young patients and their families transition from a multidisciplinary team of clinicians during active treatment to a model of medical surveillance in the off-treatment phase, with a period of at least 2 years and up to 5 years before they are eligible for "survivorship" programs. This gap in multidisciplinary clinical care represents a crucial missed opportunity to identify those children and adolescents at highest risk of educational, neurocognitive, psychosocial, and physical late effects from their cancer treatment and to implement preventative, targeted referrals and interventions early [2].

The end-of-treatment period is recognized as a potentially difficult transition for patients and families [2-6]. Children and adolescents have reported reduced physical functioning, lowered mood, and difficulties with anxiety, sleep, school re-entry, and social relationships [7-10]. Parents may experience elevated distress, loneliness, and fatigue [3,11,12], feelings of "abandonment" from their health care team [13-15], and an increased sense of personal responsibility, particularly in monitoring their child for relapse or adjustment difficulties [3,16]. Unmet information needs are high during the end-of-treatment period for both parents and young survivors, predominantly related to a lack of understanding of late effects and access to medical and psychosocial information [17-19].

The most common form of childhood cancer is acute lymphoblastic leukemia (ALL). Despite high survival outcomes [20], a substantial proportion of children and adolescents treated for ALL under modern chemotherapy-only regimens experience treatment late effects in the survivorship period (>5 y post treatment). Research indicates up to 50% of survivors will experience cognitive deficits in speed of information processing, working memory, attention, executive functions, and fine motor control [15,21], and behavioral and emotional symptoms including social withdrawal, anxiety, depression, somatization, inattention, and hyperactivity [22,23]. Pediatric ALL treatment also has a significant impact on parents and families more broadly, with disruption to daily routines and nearly every aspect of family life during active treatment [5,24,25]. While many parents and families will adapt over time, some parents experience persistent distress [26,27], and siblings are at risk for psychological distress, increased school absenteeism, and poorer academic outcomes [28,29].

Despite substantial evidence of the difficulties families face at the end of treatment, very little research has examined the provision of care during this period, particularly in relation to addressing the developmental and psychosocial needs of patients and families [9,15]. Therefore, it is currently not clear what is best practice in terms of the timing and mode of delivery of interventions [15,30]. However, it is highly feasible that intervening in this period may prevent the development of long-term deficits and enhance patients' educational, social, and emotional outcomes [2,10,31,32]. It is also feasible that parents may be more able to address their own psychological needs and those of siblings once their child's treatment has finished [6,15,33]. This juncture represents a crucial opportunity to intervene and optimize the recovery trajectory for these young survivors and their families.

In recognition of these difficulties, our research team has developed a novel, telehealth-delivered end-of-treatment model of care called the EMERGE program. EMERGE is designed to be delivered, via an advanced practice nurse and psychosocial clinician, for pediatric and adolescent patients who have completed treatment for ALL (or lymphoblastic lymphoma) within the last 12 months. Participants (parent or caregiver and child) receive 2 telehealth consultations from home and are provided with documentation, including a treatment summary, surveillance roadmap, letters summarizing consultations, and information resources. The EMERGE model was co-designed by an expert research team together with consumers (parents and adolescents) and oncology health care providers. If successful, the EMERGE model is designed to address unmet information needs, provide multidisciplinary clinical assessment of patient and family needs, and provide a preventative and early intervention approach to addressing physical, psychological, and developmental impacts of cancer therapy, thereby potentially reducing the long-term burden of cancer treatment late effects.

This study will address the current gap in systematized, evidence-based end-of-treatment interventions for young pediatric cancer survivors and their families, addressing both the implementation and effectiveness outcomes assessed across 2 pediatric cancer centers.

Within this paper, we present the approach to the design and investigation of the EMERGE program.

Methods

Study Aims

The primary aim of this study is to assess the implementation of the EMERGE model of care in 2 Australian pediatric cancer centers across a range of indicators, including reach, acceptability, satisfaction, maintenance, and barriers and facilitators to participation.

The secondary aim of the study is to evaluate the clinical effectiveness of the EMERGE model of care, including (1) participant satisfaction, (2) reduction in unmet information needs, and (3) reduction in parent or caregiver stress.

XSL•FO
RenderX

An exploratory aim of this study is to evaluate the costs of providing EMERGE and explore a sustainable model to upscale the program to all patients and families diagnosed and treated for childhood cancer in Victoria, Australia.

Setting

This study will be conducted at the Children's Cancer Centers at The Royal Children's Hospital (RCH; parent site) and the Monash Children's Hospital (MCH; satellite site) in Melbourne, Australia. These 2 centers are responsible for coordinating the care of all children and young people aged 0‐18 years diagnosed with cancer in the state of Victoria, Australia.

Outcomes

Implementation Outcomes

We hypothesize that EMERGE will be successfully implemented at both study sites; specifically, the telehealth platform will be highly acceptable to participants, and stakeholders will report high levels of satisfaction with the intervention. We hypothesize that the barriers and facilitators to implementation will differ between the parent site (RCH) and satellite site (MCH) due to available resources and local champions. We also hypothesize that the intervention will be relatively low-cost to deliver, supporting the development of a sustainable business model for ongoing service provision once the research program ceases.

Clinical Effectiveness Outcomes

We hypothesize that the EMERGE program will (1) be rated highly (satisfaction) by participants, (2) decrease unmet information needs for participants, and (3) lead to a reduction in parent-perceived stress.

Trial Design

This implementation study will use a mixed methods approach to develop and evaluate the EMERGE end-of-treatment model of care. A type 2 nonrandomized hybrid trial will be used. A hybrid design advances knowledge of both the intervention and the implementation factors associated with integration into clinical practice [34].

This study has a dual focus of assessing (1) the implementation outcomes of the EMERGE model of care in real-world settings across 2 institutional sites, and (2) the clinical effectiveness of the EMERGE model of care using a within-subject pre-post design.

The study is guided by the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework [35]. RE-AIM emphasizes the translation of research into practice and the importance of considering context in intervention planning and design. It provides a framework to assess what intervention components are effective, for whom and in which settings, and with which implementation strategies [36]. RE-AIM theorizes that the impact of an intervention is a function of 5 systems based on socioecological factors, that is, reach, efficacy, adoption, implementation, and maintenance. The proposed evaluation embedded within this protocol addresses these 5 factors.

Ethical Considerations

Before study activation and participant enrollment, the study protocol was approved by the RCH Human Ethics and Research Committee (78840). Informed consent is obtained from all participants aged 16 years and older. Parental consent will be obtained for children younger than 18 years of age. All data are securely stored on servers and software approved for the use of protected health information. Additionally, data are deidentified upon collection and stored under a study identification number. Participants did not receive compensation for participation in this trial.

Participants

Children and adolescents aged between 2 and 18 years (or older if receiving active surveillance in the pediatric setting), who are completing treatment for pediatric ALL at 1 of the 2 cancer centers (RCH and MCH), and a parent caregiver will be invited to participate in the study. Patients who have completed treatment for lymphoblastic lymphoma are also eligible for this study, as they present with the same disease features and undergo the same treatment protocols as patients with ALL. Given the high similarities shared by these disease entities, and for ease of reference, we will refer to patients with ALL in this protocol. Refer to Table 1 for participant criteria.

XSL•FO
RenderX

Table . Eligibility criteria.

	Inclusion criteria	Exclusion criteria
Patients	• Child aged 2 - 18 y (or older if receiving active surveillance in the pediatric setting). • Diagnosed with ALL[a] or LLy[b] • Treated at the RCH[c] or MCH[d] • Is within the last 3 months of maintenance therapy or has completed treatment within 12 months at time of recruitment.	• Outside this age range • Diagnosed with another form of cancer • With relapsed or refractory disease • Receiving end-of-life treatment
Parent or caregivers	• Parent of a patient aged between 2 and 18 y (or older if receiving active surveillance in the pediatric setting), who has been treated for ALL or LLy at either the RCH or MCH and is in the final period maintenance treatment or has completed treatment within the last 12 months.	• Parent of a child with another form of cancer, or who has relapsed or refractory disease and/or is receiving end-of-life treatment.
Health care providers	• Medical, nursing, social work, allied health, and mental health staff, heads of departments and champions, who work with children treated for ALL or LLy and their families at either the RCH or MCH.	• Medical, nursing, social work, allied health, or mental health staff who do not work directly with pediatric patients with ALL or LLy.
Institutional champions	• Head of Department, Operational Managers, Executive staff	—[e]

[a]ALL: acute lymphoblastic leukemia.

[b]LLy: lymphoblastic lymphoma.

[c]RCH: Royal Children's Hospital.

[d]MCH: Monash Children's Hospital.

[e]Not applicable.

Recruitment

Children and adolescents with ALL, being treated with curative intent (in the nonrelapse or refractory setting) and who have recently completed treatment or are nearing the end of treatment will be identified during routine multidisciplinary team meetings undertaken weekly at the RCH and MCH cancer centers. Families identified as potential participants will be introduced to the program by their treating oncologist and provided with a study brochure. For families who do not opt out of the study, a member of the research team will contact the primary caregiver directly to obtain written consent to participate in the study. Adolescents will be approached to provide consent if they are determined to have appropriate literacy ability and cognitive capacity to understand the requirements of the intervention. Competency will be determined by talking to the treating team and in consultation with parents and caregivers. For culturally and linguistically diverse families, interpreter services can be accessed to assist with the introduction and enrollment of families into the study. We will also encourage families to include a support person in this process if they find this helpful.

Health care providers and organizational stakeholders will be approached directly by the research team to participate in the project. For data that are collected directly from study participants (eg, questionnaires), the participant will need to provide electronic consent via REDCap (Research Electronic Data Capture; Vanderbilt University). REDCap is a secure, web-based software platform designed to support data capture for research studies [37,38].

The EMERGE Intervention Model

The EMERGE model of care will offer families a minimum of 2 telehealth (video conference) consultations during the first 12 months of their child completing treatment for ALL. The EMERGE program is cofacilitated by an advanced practice nurse and psychosocial clinician who have been trained in the implementation of the intervention. A telehealth platform was chosen to address barriers identified for attending appointments (eg, burden of attending the hospital, desire to avoid health settings, and disruption to school attendance) and to increase the opportunity to deliver the intervention regardless of geographical location from the hospital. Telehealth has strong organizational support from the hospitals and offers a flexible, safe, and potentially cost-effective framework to screen and monitor families in the off-treatment period [39,40].

The EMERGE model is based upon the Pediatric Psychosocial Preventative Healthcare Model, a 3-tier model of universal, targeted, and clinical categorization that represents the distribution of psychosocial risk across family populations [41]. The EMERGE model also incorporates the recommendations of the Standards for the Psychosocial Care of Children with Cancer and Their Families, developed in 2015 [42], by specifically including components, such as psychosocial

XSL•FO
RenderX

screening, psychoeducation, monitoring and review, and specialist interventions (Figure 1).

Figure 1. Schematic diagram of the EMERGE model of care. HCP: health care provider; MBS: Medicare Benefits Scheme; PROM: patient-reported outcome measure.

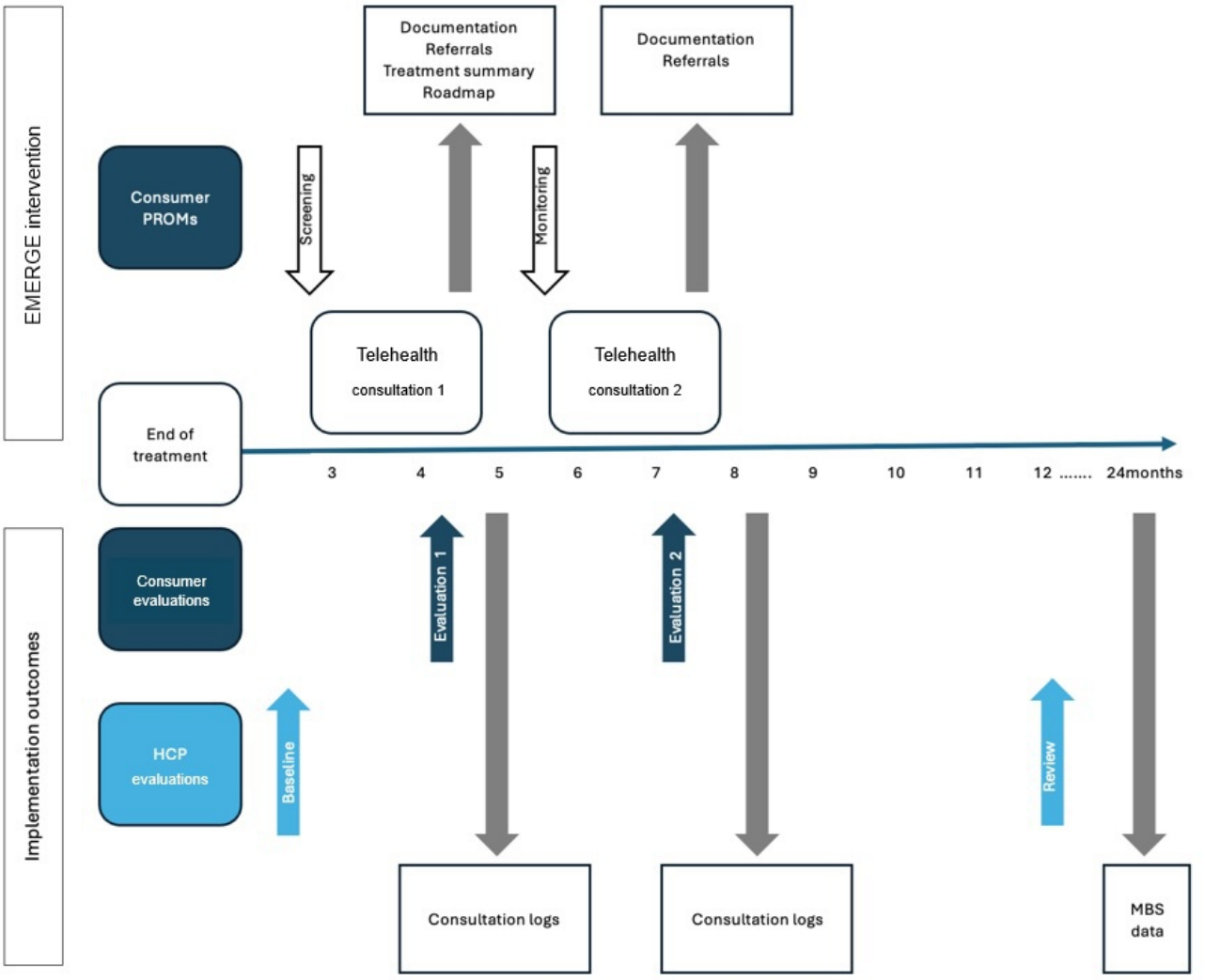

Components of the EMERGE Model

The key components of the EMERGE model are presented below.

First, identification of eligible participants via the clinical team and informed consent obtained by the EMERGE research team.

Second, treatment summary and surveillance roadmap: these documents will be prepared by the EMERGE advanced practice nurse and validated by the patient's primary oncologist. The treatment summary includes the child's initial presentation, significant medical and family cancer history, the treatment protocol delivered, diagnostic criteria (including criteria for risk stratification), cumulative chemotherapy doses, significant treatment toxicities, and any clinical trial participation. The surveillance summary provides a schedule of the necessary tests, investigations, and physical examinations over the corresponding 5-year period. These documents provide a narrative for the nurse to revisit the diagnosis and treatment with the participants, as well as to discuss the prospective treatment and surveillance plan.

Third, screening: a risk profile is established for each patient or family based upon the treatment summary and administration of validated patient-reported outcome measures (PROMs). Screening PROMs include measures of patient symptom burden and quality of life (QoL), parent distress and QoL, and family psychosocial risk and functioning (refer to Table 2 for complete list). These measures are completed by the parent caregiver and patients aged >8 years at preintervention baseline (screening) to inform the first telehealth session, with selected measures repeated before the second telehealth session (monitoring).

XSL•FO
RenderX

Table . Screening measures for EMERGE consultations.

Intervention and health measures	Participant and data source	Timing and data collection	Measure
Psychosocial assessment			
Psychosocial risk	• Parent	• Screening[a]	Psychosocial Assessment Tool, Version 3 (Australian adaptation)
Patient assessment			
Stress	• Parent proxy • Child[b]	• Screening • Monitoring[c]	PROMIS[d] Stress Experiences Scale (4-item)
Summary profile	• Parent proxy • Child[b]	• Screening • Monitoring	PROMIS summary profile (25-item): includes physical function (mobility), anxiety, depressive symptoms, fatigue, peer relationships, pain interference, and pain intensity.
Cognition	• Parent proxy • Child[b]	• Screening • Monitoring	PROMIS Cognitive function (7-item)
Physical function	• Parent proxy • Child[b]	• Screening • Monitoring	PROMIS Physical activity (4-item)
Sleep issues	• Parent proxy • Child[b]	• Screening • Monitoring	PROMIS Sleep-related disturbance (4-item) and PROMIS Sleep-related impairment (4-item)
Meaning and purpose	• Parent proxy • Child[b]	• Screening • Monitoring	PROMIS Meaning and Purpose (4-item)
Friendships	• Parent proxy • Child[b]	• Screening • Monitoring	NIH[e] Toolbox Friendship (5-item)
Social relations	• Parent proxy • Child[b]	• Screening • Monitoring	Neuro-QOL[f] Pediatric Social Relations – interactions with peers (8-item)
Parent assessment			
Stress	• Parent	• Screening • Evaluation 1[g] • Evaluation 2[h]	NIH Toolbox perceived stress module (10 items)
Distress	• Parent	• Screening • Monitoring	Distress Thermometer for Parents (single item plus 34 item problem list)
Family assessment			
Family functioning	• Parent	• Screening	McMaster Family Assessment Device (12 items)

[a]Screening: before consultation 1.

[b]Child aged more than 8 y.

[c]Monitoring: before consultation 2.

[d]PROMIS: Patient-reported Outcomes Measurement Information System.

[e]NIH: National Institutes of Health.

[f]Neuro-QOL: Quality of Life in Neurological Disorders.

[g]Evaluation 1: following consultation 1.

[h]Evaluation 2: following consultation 2.

Fourth, psychoeducation or resource provision: psychoeducation and counseling will be provided directly by the EMERGE advanced practice nurse and psychosocial clinician during the 2 scheduled telehealth sessions. Telehealth sessions will occur at least 3 months apart for families with a minimum of 2 consultations. Information resources will be provided both

XSL•FO
RenderX

electronically and in hard copy if required and will directly address concerns identified in each of the EMERGE telehealth sessions (eg, immunization schedule, sleep, when to call the hospital, fatigue, financial resources, and cancer charities).

Fifth, documentation: a summary letter, personally addressed to the parent and, if applicable, adolescent, will be provided both electronically and in hard copy to families after each EMERGE consultation, along with the treatment summary, roadmap, and information resources. With parent permission, these documents will also be shared with the child's general practitioner or pediatrician. Documents will also be sent to the primary oncologist and uploaded into the patient's electronic medical record.

Sixth, monitoring and review: follow-up surveys after completion of the first EMERGE consultations will evaluate satisfaction with the consultation, parent stress, and information needs. PROMs are readministered before the second EMERGE consultation to enable monitoring of patient and family well-being. Goal or intervention adherence, health service usage, and referral needs will be monitored and discussed in the second EMERGE consultation. Families will be provided access to EMERGE clinicians via a dedicated email and telephone number, which will be monitored regularly by the research team.

Finally, specialist intervention: for patients and families with significant risk factors and identified concerns, referrals will be actioned to hospital or community-based specialists (eg, clinical psychology, consultant pediatrician, neuropsychology, occupational therapy, education liaison, and medical oncologist). Patients and families will also be able to escalate their concerns to activate an additional telehealth or face-to-face session with their EMERGE clinicians or specialist medical consultants as required.

Measures

Screening Measures

Refer to Table 2 for a summary of screening measures and the administration schedule.

Validated Questionnaires Administered

Psychosocial Assessment Tool – Third Edition

The Psychosocial Assessment Tool – Third Edition (PAT 3) [43] is a brief parent report screener of family psychosocial risk. The PAT 3 generates a total score and 7 subscale scores (family structure, social support, child problems, sibling problems, family problems, stress reactions, and family beliefs) by summing the number of endorsed high-risk items. The total score has cut-offs for 3 tiers of risk—universal, targeted, and clinical. The PAT 3 has demonstrated good internal consistency (Kuder-Richardson coefficient=0.81) and construct validity [43].

Patient-Reported Outcomes Measurement Information System

The Patient-Reported Outcomes Measurement Information System [44] measures are a system of highly reliable and precise measures to evaluate patient-reported physical, mental, and social health. They are validated for paper or computer completion by parent proxy (age 5 - 7 y) and by self-report (age >8 y). Measures assess a range of symptom domains, including physical and motor function, sleep, cognition, stress, relationships, physical activity, and life satisfaction. Internal consistency has been shown to be strong (Cronbach α=0.72-0.92) [45], as is construct validity [46,47] and temporal sensitivity [47] in pediatric oncology samples. This study will use the Patient-Reported Outcomes Measurement Information System summary profile (25 items) with additional domains included (eg, cognition, sleep impairment, meaning, and purpose).

Distress Thermometer for Parents

The Distress Thermometer for Parents [48] consists of 3 sections examining parent distress levels and specific problems over the previous week. Parents indicate their distress on a thermometer scale ranging from 0 (*no distress*) to 10 (*extreme distress*); a score of 4 or higher is indicative of clinically elevated distress. They also complete a 34-item problem list spanning 6 domains: practical, social, emotional, physical, cognitive, and parenting. Finally, questions regarding perceived support and understanding from others, parental chronic disease, relationships with medical staff, and desire to consult a health professional about the situation are completed. The Distress Thermometer for Parents has been validated as a short screening tool for identifying parental distress [48].

McMaster Family Assessment Device

The McMaster Family Assessment Device [49] is based on the McMaster model of family functioning and contains statements concerning the family's functioning. The measure requires respondents to indicate their level of agreement with each statement on a 4-point Likert scale *(*strongly agree to strongly disagree). Higher scores indicate poorer family functioning. For this study, only the General Family Functioning Scale will be used (12 items), and it has been shown to have good reliability and validity [50].

Implementation Measures

Table 3 summarizes the implementation measures mapped according to the RE-AIM framework. These measures evaluate implementation processes and any factors (barriers and enablers) influencing implementation of the EMERGE model of care.

XSL•FO
RenderX

Table . Implementation measures using the RE-AIM[a] framework.

Implementation outcome and tool	Source	Phase
Reach		
Percentage uptake among eligible families		
Recruitment log	• Trial database	• End of trial
Representativeness of participants for Victorian ALL[b] population		
Recruitment log	• Trial database	• End of trial
Acceptability of model		
Experience of Care survey	• Consumers	• Evaluation 1[c] • Evaluation 2[d]
Semistructured interviews (subset)	• Consumers • HCPs[e]	• Evaluation 2
Acceptability of the telehealth platform		
Experience of care survey	• Consumers	• Evaluation 1 • Evaluation 2
Semistructured interviews (subset)	• Consumers • HCPs	• Evaluation 2 • 12 months
Effectiveness		
Information needs		
Information Needs survey	• Consumers	• Screening • Evaluation 1 • Evaluation 2
Satisfaction with intervention		
Client Satisfaction Questionnaire	• Consumers	• Evaluation 1 • Evaluation 2
Experience of care		
Experience of Care survey	• Consumers	• Evaluation 1 • Evaluation 2
Semistructured interview (subset)	• Consumers	• Evaluation 2
Health service usage		
Health services usage log	• Consumers	• Evaluation 2
Adoption		
Participation rates		
Recruitment log	• Trial database	• End of trial
Number of staff trained in EMERGE administration		
Training Log	• Trial database	• End of trial
Implementation		
Percent of survivorship care plans provided		
Care plan log	• Trial database	• End of trial

XSL•FO
RenderX

Implementation outcome and tool	Source	Phase
Adherence		
Referral log	• Trial database	• End of trial
Recommendations log	• Trial database	• End of trial
Percent of participants who completed 2 telehealth sessions		
Telehealth session log	• Trial database	• End of trial
Percent of participants who received an information resources pack		
Resource allocation log	• Trial database	• End of trial
Acceptability, appropriateness, and feasibility of the intervention		
Acceptability Intervention Measure, Intervention Appropriateness Measure, and Feasibility of the Intervention Measure	• HCPs	• Baseline[f] • 12 - 24 months
Maintenance		
Staff undertaking procedures		
Telehealth session log	• Trial database	• Evaluation 1 • Evaluation 2
Follow-up work log	• Trial database	• Evaluation 1 • Evaluation 2
Length of consultation		
Telehealth session log	• Trial database	• Evaluation 1 • Evaluation 2
Time spent on follow-up		
Follow-up work log	• Trial database	• Evaluation 1 • Evaluation 2
Sustainability of EMERGE model		
Health Economics Evaluation		
CHU9D[g] and AQoL-4D[h]	• Consumers	• Screening[i] • Evaluation 2 • 12 months
MBS[j] and PBS[k] data	• Services Australia	• 24 months
Cost of model integration into service	• Trial database	• Evaluation 2
HCP and institutional evaluation		

XSL•FO
RenderX

Implementation outcome and tool	Source	Phase
Semistructured interviews	• HCPs	• 12 - 24 months
Semistructured interviews	• Executive staff	• End of trial

[a]RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance.
[b]ALL: acute lymphoblastic leukemia.
[c]Evaluation 1: following consultation 1.
[d]Evaluation 2: following consultation 2.
[e]HCP: health care provider.
[f]Baseline: beginning of trial implementation.
[g]CHU9D: Child Health Utility 9D.
[h]AQoL-4D: Australian Quality of Life Instrument.
[i]Screening: before consultation 1.
[j]MBS: Medicare Benefits Scheme.
[k]PBS: Pharmaceutical Benefits Scheme.

Acceptability of Intervention Measure, Intervention Appropriateness Measure, and Feasibility of Intervention Measure

The Acceptability of Intervention Measure, Intervention Appropriateness Measure, and Feasibility of Intervention Measure [51] are 3 implementation measures designed to measure the degree of fit or match of something to a desired criterion and allow for evaluating the success of the implementation efforts. These measures specifically assess the acceptability, appropriateness, and feasibility of the intervention. Responses are made on a 5-point ordinal scale ranging from "*completely disagree*" to "completely agree." Each of the measures has been shown to have good reliability (Acceptability of Intervention Measure Cronbach α=0.85, Intervention Appropriateness Measure Cronbach α=0.89, and Feasibility of Intervention Measure Cronbach α=0.89) as well as good conceptual validity [51]. These measures will be completed by the health care provider before and after the implementation of the EMERGE intervention.

Semistructured Interviews

A subset of parents (n=20, 25%) and health care providers (n=10 - 15) from the RCH and (n=5 - 10) from MCH will be asked to complete a semistructured interview (with a researcher not involved in delivering the program), 12 - 24 months following the EMERGE intervention being active at their site. Interviews will include questions related to the acceptability of the EMERGE model, experience of care, barriers and facilitators to the delivery, and sustainability of the EMERGE model.

Outcome Measures

Information Needs Survey

An information needs survey has been developed for this study and was adapted from an original survey used in the Adolescent and Young Adult Health Outcomes and Patient Experience (AYA Hope) study [52] and a subsequent Australian study of adolescent and young adult cancer survivors [53]. Parent participants and adolescent participants older than 16 years will be asked to rate their need for more information on items such as their treatment side effects, cancer relapse, physical functioning, survivorship plan, and community-based services. Participants respond to each of the 31 items, indicating (1) *I have enough information*, (2) *I need some more information*, (3) *I need much more information*, and (4) *does not apply*.

National Institutes of Health Toolbox Perceived Stress Scale

The National Institutes of Health Toolbox Perceived Stress Scale [54] is a self-report measure of an individual's perceptions about the nature of events and their relationship to the values and coping resources of an individual. This is a 10-item measure, rated on a 5-point Likert scale (0=*never* to 4=*very often*). It assesses how unpredictable, uncontrollable, and overloaded respondents find their lives. The perceived stress module is normed from 18 to 85 years of age and provides a *t* score with a mean of 50 (SD 10). The Perceived Stress Scale has demonstrated good reliability (Cronbach α=0.91) and content validity [54].

The Client Satisfaction Questionnaire

The Client Satisfaction Questionnaire [55] is a self-report measure of satisfaction with services received by individuals and families in the health and human services settings. The measure has 8 items that target: quality of service, kind of service, met needs, recommend to a friend, amount of help, deal with problems, overall satisfaction, and return intent. Responses are made on a 4-point Likert scale, with high scores indicating higher satisfaction. Reliability and validity testing in health care contexts support good reliability (Cronbach α=0.83-0.93) and validity [55].

Experience of Care Survey

This survey has been developed specifically for the EMERGE study and adapted from other health intervention implementation studies. The survey comprises 7 questions which parents respond to on a 5-point Likert-type scale ranging from strongly agree to strongly disagree. Example items include "*I was able to ask all the questions I had about my child*," "*There was enough time to cover all the areas I wanted to discuss during the session*," and "*The telehealth platform was easy to use*."

XSL•FO
RenderX

Health Service Usage

Medicare Benefits Schedule Data and Psychosocial Service Usage

To capture health service usage during the period of intervention, we will obtain consent (optional) to access the patient's Medicare Benefits Schedule data via Services Australia. This consent will be sought by the research team to align with the family's second telehealth session. The data will provide information relating to Medicare Benefits Schedule–billed medical visits and procedures, and the associated costs (eg, general practitioners, diagnostics, and specialists, including psychosocial services where Medicare is subsidized). Data will be requested for a 24-month period post-treatment.

Child Health Utility 9D

The Child Health Utility 9D (CHU9D) [56] is a preference-based measure of health-related QoL for children and adolescents. Parent-proxy forms will also be used to monitor patient QoL in this study. The questionnaire assesses 9 dimensions—worries, sadness, pain, tiredness, annoyance, schoolwork, sleep, daily routine, and activities. Responses are made on a 5-point ordinal scale of "*no*" to "very." The CHU9D consists of a descriptive system and a set of preference weights, giving utility values for each health state described by the descriptive system, allowing the calculation of quality adjusted life years (QALYs). The measure has been shown to have good reliability [57] and construct validity [58].

The Australian Quality of Life Instrument

The Australian Quality of Life Instrument (AQoL-4D) [59] is a multiattribute utility health-related QoL instrument, which will be used to monitor parent QoL in this study. The AQoL-4D consists of 12 items assessing 5 dimensions: illness, independent living, social relationships, physical senses, and psychological well-being. Each item has 4 response levels. The AQoL-4D uses a descriptive system and set of preference weights that provide utility values for each health state described by the descriptive system. As with the CHU9D, this will allow for the calculation of QALYs [59].

Data Analysis

Sample Size Estimation

The RCH and MCH collectively treat approximately 90 children and adolescents annually who are diagnosed with ALL. Given that approximately 15% (n=9) of these patients may relapse, become palliative, or die, and recruitment will occur over a 24 to 30-month period, we anticipate recruitment of approximately 100 families over the course of the project (allowing for some families opting not to participate and some withdrawals due to disease recurrence).

Data Analysis Plan

Feasibility and acceptability of the EMERGE model will be assessed by calculating recruitment and participation rates, acceptability ratings by clinical staff, and parents' or patients' satisfaction measured post-intervention (eg, length, content, and information). Descriptive statistics will be used to report sample characteristics and for all quantitative measures. Intention to treat will be used to analyze primary outcome measures (satisfaction, information needs, and parent stress ratings). Change between time points in information needs and stress ratings will be analyzed using paired *t* tests. In order to detect a clinically meaningful change from baseline of 0.3 of an SD, with 80% power and error rate α=.05, a sample size of n=100 is required. Univariate and multivariate analyses will be used to identify psychosocial, demographic, and treatment-related variables associated with parent information needs and stress ratings. With a sample of n=100 (80% power, α=.05), we will be powered to detect a medium size of effect (f^2=0.15) in multiple linear regression models with up to 5 predictors. Secondary outcomes such as uptake or adherence to recommendations will be calculated (frequencies and percentages).

Costs and Consequences

The assessment of costs will reflect the resources required for program delivery (as reflected by the type of practitioner involved in the delivery of the program and the time required) and the impact on subsequent health care service use (as reported from Medicare data and patient information for psychology services). Costs will be reported on the basis of total and average costs per patient for program delivery, subsequent health care use, and overall. Program consequences will be expressed in terms of the impact from baseline on unmet needs, satisfaction, and QALYs (as separate outcomes).

Qualitative Analysis

Health care provider and parent interviews will be analyzed deductively to code a priori themes related to implementation domains. Inductive analysis will also be used to identify novel or emerging data, using the thematic approach described by Braun and Clarke [60]. Deidentified interview transcripts will be imported into Lumivero NVivo software [61] for coding and analysis. Parent and health care provider interviews will be analyzed separately. Data analysis will examine differences between stakeholders and participants and across sites and will examine barriers and facilitators to the delivery of the EMERGE intervention.

Results

Funding for this study was received from the Victorian Cancer Agency in April 2021. The trial commenced in December 2022, and recruitment concluded in October 2025, with 81 families recruited. Data collection is ongoing and is anticipated to be completed in Summer 2026. Data analysis and results are expected to be completed in late 2026.

Discussion

Principal Findings

Consistent data from many research studies highlight that the end-of-treatment period is a difficult transition period for many patients and families [1,2]. Furthermore, survivors of childhood cancer are often impacted by disease and treatment-related late effects, but few interventions have been developed that adopt

XSL•FO
RenderX

a preventative approach to identifying issues and proactively providing strategies and resources to optimize child and family well-being [6,62,63]. This protocol details a type 2 nonrandomized hybrid implementation-effectiveness trial that will evaluate an eHealth, multidisciplinary end-of-treatment model of care for pediatric patients who have completed treatment for ALL and their families. The EMERGE intervention addresses a critical gap in models of care at this very early survivorship period and bridges the gap between active treatment and later survivorship models.

The study will assess implementation outcomes using both quantitative and qualitative measures that will provide a detailed assessment of the feasibility and acceptability of the EMERGE model as well as barriers and facilitators. The hybrid methodology, examining implementation and effectiveness outcomes, will provide important data to guide future research and clinical efforts, given the current dearth of end-of-treatment interventions. Health economic analyses will also ascertain the relative costs of the program and contribute to efforts to achieve sustainability of such models of care into the clinical setting.

Strengths and Limitations of This Study

A significant strength of the EMERGE model is that it is delivered by multidisciplinary staff (advanced practice nurse and psychosocial clinician) who are able to respond to medical and psychosocial issues in real time and with the requisite expertise. This is an important feature of the EMERGE model, given the predominance of psychosocial or mental health issues reported by families during this period [3,64]. A further strength of the study is the focus on implementation outcomes, which, given the small number of published evidence-based models of care, will significantly contribute to understanding the barriers and facilitators to establishing end-of-treatment programs. The delivery of the program across 2 hospital sites will provide additional information related to contextual and organizational factors that may influence implementation. Delivery of the program via telehealth will enhance the accessibility of the program to families [39]. This is an important consideration, particularly in the Australian context, where the population is geographically dispersed and traveling to the hospital for additional visits is costly and difficult for many families [65]. We anticipate that the use of telehealth, and also additional resources to support the participation of culturally and linguistically diverse families, will potentially increase the equity and accessibility of the program for all families, and essential considerations in designing any new model of clinical care.

Limitations of the study must be noted. As the study is focused only on patients with ALL and their families, findings will not necessarily be able to translate to other diagnoses. Furthermore, despite the convenience of the telehealth model, this may also be less optimal in terms of engaging young children and adolescents, and some families may indicate a preference for in-person consultations. Finally, the nonrandomized study design means the clinical effectiveness outcomes should be interpreted with caution. The within-subject, pre-post design for clinical measures (parents' stress and information needs) cannot exclude natural recovery or regression to the mean. Future study design considerations should consider ethical and innovative study designs, including randomized controlled designs, to measure whether interventions such as EMERGE prevent or impact short- and longer-term survivorship outcomes for young people and their families.

Conclusion

Addressing the complex needs of a growing population of pediatric cancer survivors is a challenge for health care systems. The EMERGE intervention offers a potentially effective and low-cost approach to addressing the unmet informational, clinical, and psychosocial needs of young survivors of ALL and their families. Importantly, the model also fosters families' re-engagement with primary care providers (eg, general practitioners), which is an important factor in optimizing patient and family well-being as hospital resources are reduced following active treatment. The focus of this study on implementation barriers and facilitators and cost analysis will provide important information to oncology staff and hospital administrators who must continue to promote the evolution of health care models to address the significant treatment-related morbidity associated with increased pediatric cancer survivorship.

Acknowledgments

The authors acknowledge the Victorian Pediatric Integrated Cancer Service for in-kind support to this project.

We thank Siena Barton for assistance in preparing the manuscript.

No AI has been utilized in this study.

Funding

The funding for this study is provided by a fellowship grant from the Victorian Cancer Agency and the Victorian State government (MCRF 20042). The study has been funded for a 4 year period, commencing in 2021. The study was funded for AUD $479,969 (US $318,796). Additional funding is provided by the Monash Children's Hospital Foundation, commencing in February 2024 for the amount of AUD $50,000 (US $33,210).

XSL•FO
RenderX

Data Availability

The datasets will not be publicly available but may be requested from the project principal investigator [MM] upon reasonable request following completion of the trial and publication of the main outcomes.

Authors' Contributions

MCM and CDL conceived the study and design, conducted previous studies and literature reviews, devised the study protocol, assembled the measures, and provided scientific project management. MCM drafted the manuscript with critical review provided by CDL. CW, RDAL, HP, KM, JT, PD, KK, and SH helped to conceive the study, provided expertise in survivorship, health economics, implementation science, statistical analysis, and critically reviewed the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Peer Review Report 1
Peer review report by The Royal Children's Hospital Melbourne and the Australian and New Zealand Children's Haematology and Oncology Group.
[PDF File, 153 KB - resprot_v15i1e85901_app1.pdf]

References

1. von Werthern M, Grigorakis G, Sweet R, Soon K. The continuity of cancer: children and adolescent's experiences of ending cancer treatment - a systematic review. Qual Health Res 2022 Nov;32(13):2006-2018. [doi: 10.1177/10497323221128860] [Medline: 36190313]
2. Conway Keller M, King C, Hart L, et al. The end of cancer treatment experience for children, adolescents, and their parents: a systematic review of the literature. J Psychosoc Oncol 2020;38(5):573-591. [doi: 10.1080/07347332.2020.1769795] [Medline: 32602790]
3. Wakefield CE, McLoone JK, Butow P, Lenthen K, Cohn RJ. Parental adjustment to the completion of their child's cancer treatment. Pediatr Blood Cancer 2011 Apr;56(4):524-531. [doi: 10.1002/pbc.22725] [Medline: 21298736]
4. Lea S, Martins A, Bassett M, et al. Issues experienced and support provided to adolescents and young adults at the end of active treatment for cancer: a rapid review of the literature. Eur J Cancer Care (Engl) 2018 Nov;27(6):30485604. [doi: 10.1111/ecc.12972] [Medline: 30485604]
5. Williams LK, McCarthy MC, Eyles DJ, Drew S. Parenting a child with cancer: perceptions of adolescents and parents of adolescents and younger children following completion of childhood cancer treatment. J Fam Stud 2013 Apr;19(1):80-89. [doi: 10.5172/jfs.2013.19.1.80]
6. Galtieri LR, Perez MN, Barakat LP. Quality of life of caregivers at the end of their child's pediatric cancer treatment: cancer-specific worry and material hardship. J Pediatr Psychol 2024 May 16;49(5):348-355. [doi: 10.1093/jpepsy/jsae014] [Medline: 38457310]
7. Keller MC, Needham A, Holden E, et al. We are done! Now what? Exploring end of treatment needs of childhood cancer survivors and their parents. J Pediatr Hematol Oncol Nurs 2024;41(2):96-106. [doi: 10.1177/27527530231194598] [Medline: 38238967]
8. Parrillo E, Petchler C, Jacobson LA, Ruble K, Paré-Blagoev EJ, Nolan MT. Integrative review of school integration support following pediatric cancer. J Cancer Surviv 2024 Apr;18(2):325-343. [doi: 10.1007/s11764-022-01276-y] [Medline: 36318444]
9. Lea S, Martins A, Cable M, et al. Exploring young people's experience of ending active cancer treatment: when the "Little Cancer Treatment Bubble" pops. Cancer Nurs 2021;44(4):288-294. [doi: 10.1097/NCC.0000000000000798] [Medline: 32195707]
10. Premaratne G, McCarthy M, Tennant M, Downie P, Hearps S, De Luca C. Health-related quality of life in children and adolescents treated for acute lymphoblastic leukaemia (ALL), compared with healthy peers: a longitudinal study of early survivorship. J Cancer Surviv 2024 Dec 26. [doi: 10.1007/s11764-024-01736-7] [Medline: 39722082]
11. Maurice-Stam H, Oort FJ, Last BF, Grootenhuis MA. Emotional functioning of parents of children with cancer: the first five years of continuous remission after the end of treatment. Psychooncology 2008 May;17(5):448-459. [doi: 10.1002/pon.1260] [Medline: 17828715]
12. McCarthy MC, Marks IR, Mulraney M, Downie P, Matson A, De Luca CR. Parental adjustment following their child's completion of acute lymphoblastic leukemia treatment. Pediatr Blood Cancer 2021 Nov;68(11):e29302. [doi: 10.1002/pbc.29302] [Medline: 34411413]

XSL•FO
RenderX

13. Muskat B, Jones H, Lucchetta S, Shama W, Zupanec S, Greenblatt A. The experiences of parents of pediatric patients with acute lymphoblastic leukemia, 2 months after completion of treatment. J Pediatr Oncol Nurs 2017;34(5):358-366. [doi: 10.1177/1043454217703594] [Medline: 28443699]
14. Baenziger J, Hetherington K, Wakefield CE, et al. Understanding parents' communication experiences in childhood cancer: a qualitative exploration and model for future research. Support Care Cancer 2020 Sep;28(9):4467-4476. [doi: 10.1007/s00520-019-05270-6] [Medline: 31927756]
15. Wilford JG, Hopfer S, Wenzel L. Perceptions of changes in clinical, informational, and supportive relationships after end of treatment among parents of young childhood cancer survivors. Psychooncology 2019 Apr;28(4):913-919. [doi: 10.1002/pon.5042] [Medline: 30812064]
16. Inhestern L, Peikert ML, Krauth KA, et al. Parents' perception of their children's process of reintegration after childhood cancer treatment. PLOS ONE 2020;15(10):e0239967. [doi: 10.1371/journal.pone.0239967] [Medline: 33002084]
17. Lea S, Martins A, Fern LA, et al. The support and information needs of adolescents and young adults with cancer when active treatment ends. BMC Cancer 2020 Jul 28;20(1):697. [doi: 10.1186/s12885-020-07197-2] [Medline: 32723357]
18. McCarthy MC, McNeil R, Drew S, Orme L, Sawyer SM. Information needs of adolescent and young adult cancer patients and their parent-carers. Support Care Cancer 2018 May;26(5):1655-1664. [doi: 10.1007/s00520-017-3984-1] [Medline: 29222598]
19. Wakefield CE, Butow P, Fleming CAK, Daniel G, Cohn RJ. Family information needs at childhood cancer treatment completion. Pediatr Blood Cancer 2012 Apr;58(4):621-626. [doi: 10.1002/pbc.23316]
20. Australia's children. : Australian Institute of Health and Welfare; 2020 URL: https://www.aihw.gov.au/getmedia/6af928d6-692e-4449-b915-cf2ca946982f/aihw-cws-69_australias_children_print-report.pdf [accessed 2026-02-12]
21. Kunin-Batson A, Kadan-Lottick N, Neglia JP. The contribution of neurocognitive functioning to quality of life after childhood acute lymphoblastic leukemia. Psychooncology 2014 Jun;23(6):692-699. [doi: 10.1002/pon.3470] [Medline: 24497266]
22. Jacola LM, Edelstein K, Liu W, et al. Cognitive, behaviour, and academic functioning in adolescent and young adult survivors of childhood acute lymphoblastic leukaemia: a report from the Childhood Cancer Survivor Study. Lancet Psychiatry 2016 Oct;3(10):965-972. [doi: 10.1016/S2215-0366(16)30283-8] [Medline: 27639661]
23. De Luca CR, Mulraney M, Anderson V, Downie P, Ma W, McCarthy MC. Behavioral and emotional functioning of children and adolescents at the end of treatment for acute lymphoblastic leukemia compared to healthy peers. J Clin Psychol Med Settings 2022 Jun;29(2):421-431. [doi: 10.1007/s10880-021-09840-x] [Medline: 35113314]
24. Long KA, Marsland AL. Family adjustment to childhood cancer: a systematic review. Clin Child Fam Psychol Rev 2011 Mar;14(1):57-88. [doi: 10.1007/s10567-010-0082-z] [Medline: 21221783]
25. Parker KN, Alexander SW, Jacola LM, et al. Characterizing difficulty and life disruption during B‐acute lymphoblastic leukemia therapy from the perspective of parents: a survey study. Pediatr Blood Cancer 2026 Jan;73(1):e32041. [doi: 10.1002/pbc.32041] [Medline: 40922125]
26. Vrijmoet-Wiersma CMJ, van Klink JMM, Kolk AM, Koopman HM, Ball LM, Maarten Egeler R. Assessment of parental psychological stress in pediatric cancer: a review. J Pediatr Psychol 2008 Aug;33(7):694-706. [doi: 10.1093/jpepsy/jsn007] [Medline: 18287109]
27. Parker C, Schilstra CE, McCleary K, et al. Psychosocial outcomes in parents of children with acute lymphoblastic leukaemia in Australia and New Zealand through and beyond treatment. Cancers (Basel) 2025 Apr 6;17(7):1238. [doi: 10.3390/cancers17071238] [Medline: 40227813]
28. Long KA, Lehmann V, Gerhardt CA, Carpenter AL, Marsland AL, Alderfer MA. Psychosocial functioning and risk factors among siblings of children with cancer: an updated systematic review. Psychooncology 2018 Jun;27(6):1467-1479. [doi: 10.1002/pon.4669] [Medline: 29441699]
29. Wiener L, Devine KA, Thompson AL. Advances in pediatric psychooncology. Curr Opin Pediatr 2020 Feb;32(1):41-47. [doi: 10.1097/MOP.0000000000000851] [Medline: 31790025]
30. Kearney JA, Salley CG, Muriel AC. Standards of psychosocial care for parents of children with cancer. Pediatr Blood Cancer 2015 Dec;62 Suppl 5(Suppl 5):S632-S683. [doi: 10.1002/pbc.25761] [Medline: 26700921]
31. Garas A, McLean LA, De Luca CR, Downie P, McCarthy MC. Health-related quality of life in paediatric patients up to five years post-treatment completion for acute lymphoblastic leukaemia: a systematic review. Support Care Cancer 2019 Nov;27(11):4341-4351. [doi: 10.1007/s00520-019-04747-8] [Medline: 30900055]
32. Walsh S, Mulraney M, McCarthy MC, De Luca CR. Fatigue in children who have recently completed treatment for acute lymphoblastic leukemia: a longitudinal study. Health Qual Life Outcomes 2024 Mar 22;22(1):27. [doi: 10.1186/s12955-024-02241-2] [Medline: 38519964]
33. Keller MC, Ayr-Volta L, Marconi E, et al. Shooting for the STARS: implementing an evidence-based program to meet the end of treatment needs of childhood cancer survivors and their families. J Pediatr Hematol Oncol Nurs 2024 Jul;41(4):240-251. [doi: 10.1177/27527530241245636] [Medline: 39263727]
34. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226. [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]

XSL•FO
RenderX

35. Glasgow RE, Harden SM, Gaglio B, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7(64):64. [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
36. Holtrop JS, Estabrooks PA, Gaglio B, et al. Understanding and applying the RE-AIM framework: clarifications and resources. J Clin Transl Sci 2021;5(1):e126. [doi: 10.1017/cts.2021.789] [Medline: 34367671]
37. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
38. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
39. Bradford NK, Caffery LJ, Smith AC. Telehealth services in rural and remote Australia: a systematic review of models of care and factors influencing success and sustainability. Rural Remote Health 2016;16(4):3808. [Medline: 27744708]
40. Telehealth measures to improve access to psychological services for rural and remote patients under the better access initiative. : Australian Psychological Society URL: https://psychology.org.au/getmedia/4dd9dd91-1617-421b-928c-531d019f05c2/17aps-telehealth-web.pdf [accessed 2026-02-12]
41. Kazak AE. Pediatric Psychosocial Preventative Health Model (PPPHM): research, practice, and collaboration in pediatric family systems medicine. Fam Syst Health 2006;24(4):381-395. [doi: 10.1037/1091-7527.24.4.381]
42. Wiener L, Kazak AE, Noll RB, Patenaude AF, Kupst MJ. Standards for the psychosocial care of children with cancer and their families: an introduction to the special issue. Pediatr Blood Cancer 2015 Dec;62 Suppl 5(Suppl 5):S419-S424. [doi: 10.1002/pbc.25675] [Medline: 26397836]
43. Kazak AE, Hwang WT, Chen FF, et al. Screening for family psychosocial risk in pediatric cancer: validation of the Psychosocial Assessment Tool (PAT) version 3. J Pediatr Psychol 2018 Aug 1;43(7):737-748. [doi: 10.1093/jpepsy/jsy012] [Medline: 29509908]
44. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194. [doi: 10.1016/j.jclinepi.2010.04.011]
45. Hinds PS, Wang J, Stern ED, et al. Voices of children and adolescents on phase 1 or phase 2 cancer trials: a new trial endpoint? Cancer 2017 Oct 1;123(19):3799-3806. [doi: 10.1002/cncr.30782] [Medline: 28581685]
46. Hinds PS, Nuss SL, Ruccione KS, et al. PROMIS pediatric measures in pediatric oncology: valid and clinically feasible indicators of patient-reported outcomes. Pediatr Blood Cancer 2013 Mar;60(3):402-408. [doi: 10.1002/pbc.24233] [Medline: 22829446]
47. Macpherson CF, Wang J, DeWalt DA, Stern ED, Jacobs S, Hinds PS. Comparison of legacy fatigue measures with the PROMIS Pediatric Fatigue Short Form. Oncol Nurs Forum 2018 Jan 1;45(1):106-114. [doi: 10.1188/18.ONF.106-114] [Medline: 29251294]
48. Haverman L, van Oers HA, Limperg PF, et al. Development and validation of the distress thermometer for parents of a chronically ill child. J Pediatr 2013 Oct;163(4):1140-1146. [doi: 10.1016/j.jpeds.2013.06.011] [Medline: 23910979]
49. Epstein NB, Baldwin LM, Bishop DS. The McMaster Family Assessment Device*. J Marital Family Therapy 1983 Apr;9(2):171-180 [FREE Full text] [doi: 10.1111/j.1752-0606.1983.tb01497.x]
50. Byles J, Byrne C, Boyle MH, Offord DR. Ontario Child Health Study: reliability and validity of the general functioning subscale of the McMaster Family Assessment Device. Fam Process 1988 Mar;27(1):97-104. [doi: 10.1111/j.1545-5300.1988.00097.x] [Medline: 3360100]
51. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
52. Keegan THM, Lichtensztajn DY, Kato I, et al. Unmet adolescent and young adult cancer survivors information and service needs: a population-based cancer registry study. J Cancer Surviv 2012 Sep;6(3):239-250. [doi: 10.1007/s11764-012-0219-9] [Medline: 22457219]
53. Darling SJ, De Luca C, Anderson V, McCarthy M, Hearps S, Seal ML. White matter microstructure and information processing at the completion of chemotherapy-only treatment for pediatric acute lymphoblastic leukemia. Dev Neuropsychol 2018;43(5):385-402. [doi: 10.1080/87565641.2018.1473401] [Medline: 29781717]
54. Kupst MJ, Butt Z, Stoney CM, et al. Assessment of stress and self-efficacy for the NIH Toolbox for neurological and behavioral function. Anxiety Stress Coping 2015;28(5):531-544. [doi: 10.1080/10615806.2014.994204] [Medline: 25577948]
55. Attkisson C, Greenfield T. Client satisfaction questionnaire-8 and service satisfaction scale-30. In: Maruish ME, editor. The Use of Psychological Testing for Treatment Planning and Outcome Assessment: Lawrence Erlbaum Associates; 1994:402-420.
56. Stevens K, Ratcliffe J. Measuring and valuing health benefits for economic evaluation in adolescence: an assessment of the practicality and validity of the Child Health Utility 9D in the Australian adolescent population. Value Health 2012 Dec;15(8):1092-1099. [doi: 10.1016/j.jval.2012.07.011] [Medline: 23244812]
57. Lindvall K, Vaezghasemi M, Feldman I, Ivarsson A, Stevens KJ, Petersen S. Feasibility, reliability and validity of the health-related quality of life instrument Child Health Utility 9D (CHU9D) among school-aged children and adolescents in Sweden. Health Qual Life Outcomes 2021 Aug 3;19(1):193. [doi: 10.1186/s12955-021-01830-9] [Medline: 34344386]

XSL•FO
RenderX

58. Ratcliffe J, Flynn T, Terlich F, Stevens K, Brazier J, Sawyer M. Developing adolescent-specific health state values for economic evaluation: an application of profile case best-worst scaling to the Child Health Utility 9D. Pharmacoeconomics 2012 Aug 1;30(8):713-727. [doi: 10.2165/11597900-000000000-00000] [Medline: 22788261]
59. Hawthorne G, Richardson J, Osborne R. The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life. Qual Life Res 1999 May;8(3):209-224. [doi: 10.1023/a:1008815005736] [Medline: 10472152]
60. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
61. NVivo 15 - the most trusted qualitative analysis software (QDA) is even better. Lumiivero. URL: https://lumivero.com/products/nvivo/ [accessed 2026-02-12]
62. Perez MN, Madden RE, Zukin HN, Velazquez-Martin B, Barakat LP. Psychosocial needs of pediatric cancer patients and their caregivers at end of treatment: why psychosocial screening remains important. Pediatr Blood Cancer 2023 Jun 30:e30540. [doi: 10.1002/pbc.30540] [Medline: 37391860]
63. Jones AM, Browne EK, Adams K, et al. The toll of transition: caregiver perceptions of family adjustment during the transition off pediatric cancer therapy. Psychooncology 2022 Jun;31(6):985-994. [doi: 10.1002/pon.5888] [Medline: 35083824]
64. Redkite. The hidden health crisis. 2012 Oct Report no.65.
65. Signorelli C, Wakefield C, McLoone JK, et al. Childhood cancer survivorship: barriers and preferences. BMJ Support Palliat Care 2022 Nov;12(e5):e687-e695. [doi: 10.1136/bmjspcare-2019-002001] [Medline: 31712388]

Abbreviations

ALL: acute lymphoblastic leukemia
AqoL-4D: Australian Quality of Life Instrument
CHU9D: Child Health Utility 9D
MCH: Monash Children's Hospital
PAT 3: Psychosocial Assessment Tool – Third Edition
PROM: patient-reported outcome measure
QALY: quality adjusted life year
QoL: quality of life
RCH: Royal Children's Hospital
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
REDCap: Research Electronic Data Capture

Edited by J Sarvestan; submitted 14.Oct.2025; this is a non–peer-reviewed article;accepted 19.Dec.2025; published 26.Feb.2026.

Please cite as:
McCarthy MC, Williams C, Tennant M, De Abreu Lourenco R, Pring H, Moore K, Templeton J, Knight K, Downie P, Hearps S, De Luca C
Implementing an eHealth Model of Care for Pediatric Patients and Families at the End of Treatment for Acute Lymphoblastic Leukemia (EMERGE): Type 2 Nonrandomized Hybrid Implementation-Effectiveness Trial Study Protocol
JMIR Res Protoc 2026;15:e85901
URL: https://www.researchprotocols.org/2026/1/e85901
doi:10.2196/85901

© Maria C McCarthy, Chris Williams, Michelle Tennant, Richard De Abreu Lourenco, Hannah Pring, Katie Moore, Jane Templeton, Ken Knight, Peter Downie, Stephen Hearps, Cinzia De Luca. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Acceptability of the Cardiff Online Cognitive Assessment for Clinical Screening of Patients With Psychosis: Protocol for a Mixed Methods Study

Amy J Lynham[1], BSc, PhD; Anthony Cope[1,2], BSc; Sarah Milosevic[3], BSc, MSc, PhD; Ian R Jones[1], MRCPsych, PhD; James T R Walters[1,4], MRCPsych, PhD

[1]National Centre for Mental Health, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Hadyn Ellis Building, Maindy Road, Cardiff, United Kingdom

[2]Wales Suicide Prevention, Self Harm, Mental Health & Wellbeing Research and Evidence Network, NHS Wales, Cardiff, United Kingdom

[3]Centre for Trials Research, Cardiff University, Cardiff, United Kingdom

[4]Centre for Neuropsychiatric Genetics and Genomics, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Cardiff, United Kingdom

Corresponding Author:
Amy J Lynham, BSc, PhD
National Centre for Mental Health, Division of Psychological Medicine and Clinical Neurosciences, Cardiff University, Hadyn Ellis Building, Maindy Road, Cardiff, United Kingdom

Abstract

Background: The early detection of cognitive impairments in individuals with psychosis offers a means to support clinical and functional recovery. However, there are significant barriers to assessing cognition in clinical services, including lack of staff time, training, and confidence in administering assessments. We have developed the Cardiff Online Cognitive Assessment (CONCA), aiming to address these barriers, and here present the protocol to assess its acceptability as a clinical tool.

Objective: The aim of this study is to conduct the early-stage testing of the CONCA as a clinical tool to determine whether it is acceptable to young people with a history of psychosis and to health professionals.

Methods: This cross-sectional study will use a mixed methods approach. A total of 100 young people with a history of psychosis will complete the CONCA and an acceptability questionnaire. We will conduct qualitative interviews with a minimum of 20 participants with psychosis and 10 participants with professional experience of working in early intervention in psychosis services to explore opinions on the CONCA as a clinical tool, attitudes toward and barriers or facilitators of implementing the CONCA, and cognitive testing more generally in clinical services. Quantitative data will be analyzed using descriptive statistics and linear or logistic regression. Qualitative interviews will be analyzed using a deductive thematic analysis approach.

Results: The enrollment of study participants started in July 2025 and is expected to end in October 2026. Data analysis is expected to be finalized by March 2027. As of September 15, 2025, we have enrolled 26 participants with psychosis in the quantitative arm of the study, and 4 participants with psychosis and 1 health professional in the qualitative arm.

Conclusions: Our results will provide new data on the acceptability of the CONCA and cognitive testing more generally among patients and clinicians, as well as identify barriers and facilitators to the CONCA's implementation. This will provide the groundwork for a larger hybrid effectiveness-implementation study.

International Registered Report Identifier (IRRID): DERR1-10.2196/84218

(JMIR Res Protoc 2026;15:e84218) doi:10.2196/84218

KEYWORDS

cognition; mental health; online; digital assessment; psychosis; clinical services; acceptability; implementation

Introduction

Background

Patients with psychosis experience significant cognitive impairments, up to 2.5 SDs below average, across a range of cognitive processes, including processing speed, working memory, verbal and visual memory, executive function, and social cognition [1,2]. These impairments impact community functioning, occupational functioning, social skills, and behavior to a greater extent than psychotic symptoms [3,4]. Cognitive problems are present early in the course of illness, have been identified before the onset of psychotic symptoms [5], and persist over time [6]. The early identification of these cognitive impairments enables early intervention to optimize clinical outcomes and support functional recovery.

XSL•FO
RenderX

Early intervention for young people experiencing psychosis is a global priority [7]. For example, in the United Kingdom, the National Institute for Health and Care Excellence recommends that all people presenting with their first episode of psychosis be referred to an early intervention in psychosis service (EIP) [8]. The remit of these services is to provide rapid assessment and tailored treatment for patients with psychosis to prevent relapse and promote recovery. The measurement of cognition as part of a comprehensive clinical assessment is recommended for patients with psychosis in national clinical guidelines across the world [8-11]. Cognitive assessment in EIPs could allow the early detection of cognitive problems and give clinical teams a broader understanding of their patients' presentation, identify those at risk of poor outcomes, and allow for a more personalized approach to care [12]. Despite these potential benefits, formal cognitive assessments are not routinely undertaken due to constraints of staff time and training and the associated costs of available assessments due to licensing [13]. Studies have shown that patients have little insight into their own cognitive abilities, making the self-report of cognitive function unreliable, and thus more formal assessment is required [14,15].

There are numerous barriers to conducting cognitive assessments in clinical services for psychosis. Few cognitive assessments exist that were specifically designed for the clinical assessment of patients with psychosis, although there are several assessments that may be suitable candidates including the Brief Assessment of Cognition in Schizophrenia[16], Brief Cognitive Assessment[17], or the Brief Cognitive Assessment Tool for Schizophrenia[18]. In the United Kingdom, many of the cognitive assessments that are currently available to use in the National Health Service, such as the Montreal Cognitive Assessment, Mini-Mental State Examination, and Repeatable Battery for the Assessment of Neuropsychological Status, were designed to detect cognitive decline in older patients rather than the neurodevelopmental impairments associated with psychosis [11,19]. In addition, staff do not have sufficient time, training, or confidence to administer these tools and identify cognitive impairments in their patients [13,20,21]. Formal testing can be conducted by clinical psychologists, but given their responsibilities for psychological treatment, the time they have for cognitive assessment is minimal. Digital assessments with automated scoring may reduce the burden of cognitive testing for clinical staff. However, a systematic review of digital cognitive assessments for severe mental illness found that these assessments had (1) not been sufficiently validated, (2) had high licensing costs, (3) lacked normative data against which staff could benchmark performance, and (4) lacked data on acceptability among patients and clinical staff [22].

Prior Work

The Cardiff Online Cognitive Assessment (CONCA) was developed as a brief web-based measure of cognition specifically for use in mental health research [23,24]. The CONCA addresses most of the barriers outlined above. It can provide a measure of overall cognition in as little as 10 minutes, so it is brief to administer. It can be completed on any internet-connected device and does not require a trained professional or specialist equipment to administer. The website was codeveloped with people with lived experience of psychosis and other severe mental illness and is simple to use. It has been validated against the gold standard pen and paper research assessment for psychosis, the MATRICS Consensus Cognitive Battery [23]. Performance on the CONCA is associated with the measures of disability [23]. Published normative data are available for the CONCA [24].

Objectives

The CONCA is an established research tool, but less is known about its clinical utility. We have worked with health professionals from EIPs and people with lived experience to develop new features for the CONCA that would support its use in psychiatric services, providing personalized reports for patients and clinical teams to identify cognitive impairments and help them understand their cognitive functioning. However, the CONCA needs to be assessed for clinical acceptability before it can be implemented in services. The aim of this study is to determine whether the CONCA is acceptable to young people with a history of psychosis and health professionals and whether implementation in clinical services is feasible and to support the refinement of the CONCA and design of future testing and implementation. We will also evaluate attitudes toward cognitive assessment more generally. To do this, we will assess the acceptability and usability of the CONCA through questionnaires and qualitative interviews with participants with psychosis and health professionals.

Methods

Relevant Frameworks

Development of Digital Health Technologies

This project has been designed considering guidance from the National Institute for Health and Care Excellence Evidence Standards Framework for Digital Health and Care Technologies [25], the MRC Guidance for Developing and Evaluating Interventions [26], and Mohr et al's [27] accelerated creation-to-sustainment model. The accelerated creation-to-sustainment model aims to accelerate the development of digital mental health technologies through an iterative process of design and evaluation across three phases: (1) Create; (2) Optimization, Effectiveness, and Implementation Hybrid Trial; and (3) Sustainment. Based on this model, the CONCA is currently in the *Create* phase, and the research outlined in this protocol will move the project to the next phase, *OEI Hybrid Trial* (Figure 1), by (1) developing a theory of how the CONCA will support services, (2) modeling outcomes, (3) testing procedures for evaluating the CONCA, and (4) providing data to support estimates of recruitment and retention for a larger study.

XSL•FO
RenderX

Figure 1. Accelerated creation-to-sustainment model adapted from Mohr et al [27], published under a CC BY 4.0 license [28]. OEI: Optimization, Effectiveness, and Implementation.

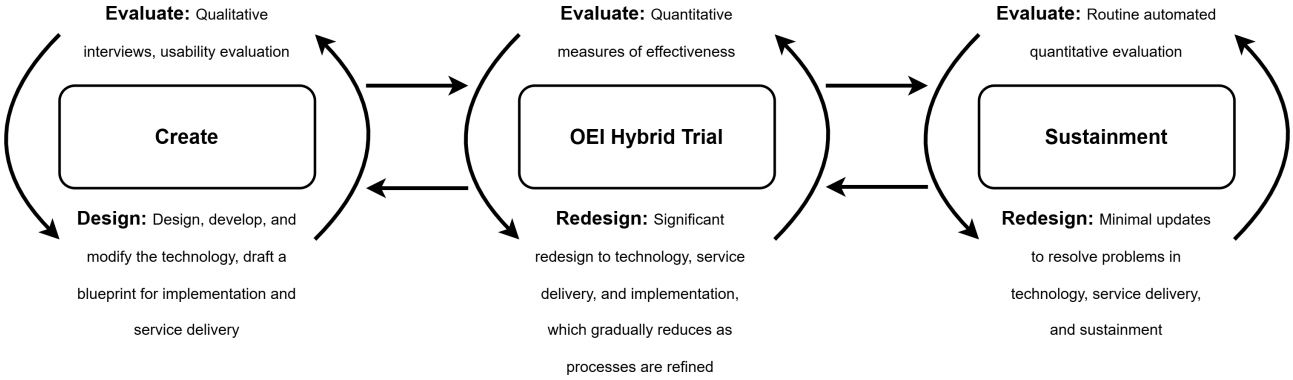

Technology Acceptance

Our measures of acceptability have been selected based on Sekhon et al's [29] theoretical framework for evaluating the acceptability of health care interventions (theoretical framework of acceptability [TFA]). Sekhon et al [29] conceptualized acceptability into the following domains: affective attitude, burden, perceived effectiveness, ethicality, assessment coherence, opportunity costs, and self-efficacy. We will assess the acceptability of the CONCA tasks, the website design and functionality, user profile page and clinical interface, as well as attitudes toward cognitive testing as part of clinical care.

Consolidated Framework for Implementation Research

We will use the Consolidated Framework for Implementation Research (CFIR) [30,31] to support the identification of potential barriers and facilitators to the successful implementation of the CONCA in clinical services. The CFIR includes 5 constructs that might influence implementation, including the innovation itself, inner organizational settings, broader outer settings (such as the political or funding landscape), the individuals involved in implementation, and the process of implementation. We will use these constructs and the existing CFIR tools to develop our qualitative interview topic guide.

Study Design

This cross-sectional study will use a mixed methods approach, where quantitative and qualitative data will be collected in parallel to address the primary research question (equal-status concurrent design [32]). Participants with psychosis will first be recruited into the quantitative arm and will be asked whether they would be interested in completing a longer follow-up interview (qualitative arm). In addition, we will invite health care workers with professional experience working in EIPs to complete the qualitative arm of the study. Consent will be received separately for the 2 arms of the study.

Participants

We will recruit a representative sample of 100 young people with a history of psychosis. The demographic composition of patients in EIPs is documented in the National Clinical Audit of Psychosis, and these data will be used as a guide during recruitment to monitor the representativeness of our sample. We will also compare the demographics of the quantitative and qualitative arms throughout recruitment. We will use a range of approaches to achieve our recruitment target, including (1) working with EIPs in Wales and England to establish and implement ways of offering potential participants the opportunity to take part, (2) liaising with voluntary organizations to promote the research, (3) attending events and conferences to raise awareness of the research, (4) inviting individuals already enrolled in previous or ongoing research studies where participants have indicated their willingness to be approached about future studies, and (5) using communication outlets (social media, website, podcasts, and blogs) to inform the public about the research. To be eligible to take part in the study, individuals must meet the following criteria:

- Aged 16 to 40 years
- Have a diagnosis of psychosis (including schizophrenia-spectrum disorders, transient and acute psychosis, and psychosis secondary to a mood disorder)
- No deficits in sight that would prevent the potential participant from completing the cognitive tasks
- Currently well enough to take part, not currently in hospital or under the care of a crisis resolution and home treatment team

In addition, we will recruit a minimum of 10 participants with professional experience working in EIPs to complete qualitative interviews. Participants will be recruited from a range of disciplines, including psychiatrists, psychologists, community psychiatric nurses, and occupational therapists.

Cardiff Online Cognitive Assessment

Overview

The CONCA is a web-based cognitive battery developed at Cardiff University, consisting of 5 tasks (Table 1). The first 3 tasks form the core battery, and the final 2 tasks are optional extras, which provide data on more complex aspects of cognitive function. The individual tasks were originally developed for web-based platforms by The Many Brains Project [33,34], a not-for-profit organization set up by Harvard University. The brief assessment can be completed in 10 minutes and produces a measure of general cognitive ability that is comparable to more intensive assessments and is associated with functional outcomes [24]. The full assessment (including the optional tasks) takes 30 minutes to complete. The assessment is hosted on a purpose-built website co-designed with people with lived experience of severe mental illness and health professionals working in EIPs.

XSL•FO
RenderX

Table . Cardiff Online Cognitive Assessment (CONCA) tasks.

Domain	Task	Administration time (min)	Description
Brief assessment (10 min)			
Speed of processing	Digit symbol coding	3	During the task, a key of 9 symbols and corresponding numbers is shown at the top of the screen. Participants select the number that corresponds to a target symbol (1, 2, or 3). The outcome measure is the number of correct responses in 90 seconds.
Working memory	Backward digit span	3	Participants are presented with a sequence of numbers and must recall them in the reverse order. The outcome measure is the maximum length of number sequence the participant can recall backward.
Premorbid or crystallized IQ	Vocabulary	3	Participants are shown a target word and asked to select which of 4 words is the closest in meaning to the target word. The outcome measure is total correct out of 20.
Optional extras (up to 20 min)			
Social cognition	Morphed emotion identification	3	Participants are presented with a face and must decide whether the face looks angry, fearful, happy, or disgusted. The outcome measure is the number of correct responses out of 60 faces.
Reasoning and problem solving	Matrix reasoning test	15	Participants must select the image that best completes a logical pattern. The outcome measure was the number of correct responses out of 35 trials.

We have used the CONCA as a research tool since 2017. In that time, we have established the validity of the CONCA against a gold standard, gathered feedback from participants, and collected data on over 5000 participants without any ethical issues (>1200 from the National Centre for Mental Health [23] and >3800 from HealthWise Wales [24]). The CONCA has been well tolerated based on high completion rates and positive feedback from participants.

We have recently developed 2 new features for the CONCA to support clinical administration: (1) an interface for clinicians to allow them to assess patients and support them to interpret their patients' results and identify areas of cognitive difficulties and (2) feedback for both patients and research participants to inform them about their own cognitive ability.

Clinical Interface

The clinical interface is accessible for clinicians through an assigned login page where they can access and search for their patients' cognitive results (patients are assigned a unique ID, so identifiable information is not required). Each cognitive report consists of (1) a guide to interpreting the results; (2) raw scores and age- and gender-corrected percentiles for each task; (3) an overall cognitive score; and (4) information on each task, what it measures, and how it relates to real-world functioning (Figure 2).

XSL•FO
RenderX

Figure 2. Screenshot from Cardiff Online Cognitive Assessment's (CONCA) clinical interface. The results shown are dummy scores for illustrative purposes only.

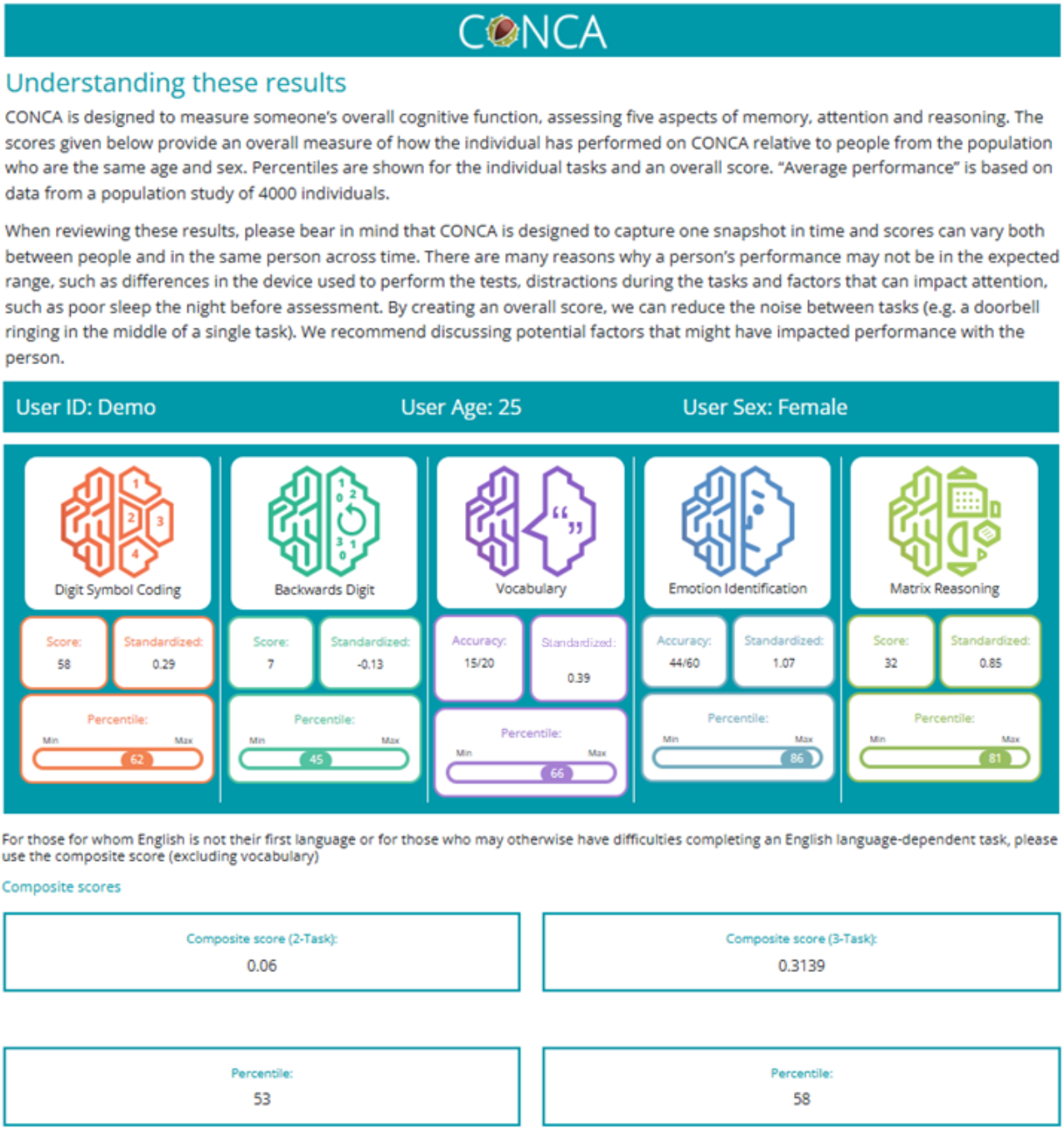

User Profile Page

The user profile page differs from the clinical interface in that the focus is on individual performance rather than comparison with population data. Each profile page includes (1) a guide to interpreting the results; (2) raw scores on the tasks and information on what they mean; (3) graphical representation of the scores, highlighting the user's best score; and (4) a graph tracking repeated attempts at each task so that users can see their performance over time (Figure 3).

XSL•FO
RenderX

Figure 3. Screenshot of user profile page. The screenshot shows a section of the My Profile page as an example. Other features of the page are described above.



Data Collection

Quantitative Data Collection

Participants will complete the assessments on their own devices in an unsupervised setting. For participants who do not have access to the internet or their own device or would like support to complete the study, we will offer the option to complete the assessments on a study device either at home or at our research facility at Cardiff University.

Participants will first complete a brief set of questions on the web-based data collection tool, REDCap (Research Electronic Data Capture), including demographic information, current diagnosis, questions about the device they are using and their internet literacy, and subjective cognitive complaints (PROMIS v2.0 Cognitive Function–Short Form 4a [35]). On the

XSL•FO
RenderX

completion of these questions, participants will be asked to complete all 5 cognitive tasks in the CONCA. The tasks will open in a separate browser tab on the CONCA website [36]. Their participant ID number will be entered automatically to allow us to link the data collected in REDCap with the data collected in the CONCA. Participants will also be instructed to review their results on their CONCA profile page.

On the completion of the CONCA, participants will be directed back to REDCap to complete a feedback questionnaire about the CONCA. We will use an adapted version of Sekhon et al's [37] questionnaire to assess attitudes toward (1) the CONCA website, (2) the CONCA tasks, and (3) the feedback given on the participant profile page. We will also include questions about attitudes toward cognitive testing as part of clinical care. A full list of questions can be found in Multimedia Appendix 1. These questions were developed in collaboration with our public advisory group. Participants will be offered £20 (US $27) by bank transfer or gift voucher as a thank you for their time.

Qualitative Data Collection

For participants with psychosis, the purpose of the qualitative interviews will be to gather information on the acceptability of the CONCA and the profile page that provides feedback on scores, as well as the experiences of cognitive difficulties and attitudes toward cognitive testing in clinical services. For participants with professional experience working in EIPs, we will specifically focus on the clinical reports generated by the CONCA and the feasibility and acceptability of implementing the CONCA in EIPs.

Purposive sampling will be used to select participants for interview to ensure that respondents are as diverse as possible in respect of age, gender, and ethnicity. Participants with psychosis who completed the quantitative study above will be invited to complete a follow-up interview. We will utilize our existing links with EIPs across the United Kingdom to invite health professionals to the study. It is anticipated that our sample sizes will provide sufficient information power [38], as the study aim and sample population are highly specific. Pragmatic decisions on sample size will be made during the data collection process, informed by the assessments of data saturation for each identified theme.

Topic guides were developed with input from our public advisory group and guided by the themes included in the TFA and CFIR. We will adopt an iterative approach to developing our topic guides, refining the questions and prompts as the transcripts are analyzed and themes emerge. For participants with psychosis, interviews will follow a topic guide to address the following areas:

- Digital literacy and confidence
  - Confidence and familiarity using the internet generally
  - Confidence in navigating the CONCA website
- Design of the CONCA
  - Appearance
  - Format
  - Functionality
- Content of the CONCA
  - Was the information on the CONCA useful and clear?
  - What could be changed?
  - What was missing?
- Use of the CONCA
  - How might the information on the CONCA be used by the participant?
  - Concerns about using the CONCA or how the participant's data may be used
- Knowledge of cognitive function
  - Understanding of cognition and different types of thinking skills
  - Experience of cognition in relation to mental health, including cognitive difficulties
- Monitoring cognition in mental health services
  - Experience discussing cognition (memory, concentration, etc) with their care team
  - Attitudes toward cognitive testing in clinics

For participants with professional experience working in EIPs, we will conduct "think-aloud" interviews as participants navigate the CONCA report. Think-aloud interviews capture the thought processes of participants as they use tools and have been used in a range of health intervention studies [39]. Interviews will follow a topic guide to address the following areas:

- Design of CONCA-generated clinical report
  - Appearance and layout
  - Format
  - Functionality
- Content of CONCA-generated clinical report
  - Was the information useful and clear?
  - What could be changed?
  - What was missing?
- Using the CONCA in EIPs
  - Current approaches to assessing cognitive function in EIPs
  - Cognitive testing as a priority in EIPs
  - How would the CONCA be used with patients?
  - Support and infrastructure needed to deliver the CONCA

On the completion of the interview, all participants will be offered £20 (US $27) as a bank transfer or gift voucher as a thank you for their time.

Analysis

Quantitative Data Analysis

To measure protocol adherence and acceptability, we will calculate completion rates for each task and conduct descriptive statistics for the feedback questionnaire (proportions and percentages). Each item in the acceptability questionnaire (Multimedia Appendix 1) is rated on a 5-point Likert scale. Following consultation with our public advisory group, it was agreed that *no opinion* or *neutral* responses would be grouped with positive responses as the evidence of acceptability. We

XSL•FO
RenderX

will report the proportion of positive and negative responses to each item in the acceptability questionnaire. Table 2 shows how each questionnaire item maps onto the TFA domains.

Table . Theoretical framework of acceptability (TFA) domains and data collected.

TFA domain	TFA survey questions	Examples of qualitative themes
Affective attitude	• How comfortable did you feel navigating the CONCA[a] website? • Did you like or dislike completing the CONCA tasks? • Did you like or dislike being able to view your results using the "My Profile" page on the CONCA?	• Positive or negative attitudes toward the CONCA website, tasks, profile page, or clinical interface.
Ethicality	• I have concerns about the CONCA being an online assessment. • I have privacy concerns around completing the CONCA. • The CONCA website is accessible for people with psychosis. • I have concerns about the CONCA tasks. • I have personal, moral, or ethical concerns about the CONCA giving feedback to people on their performance on the tasks.	• Privacy concerns, barriers to accessing the CONCA, and safety concerns (eg, impact of the CONCA on symptom exacerbation).
Self-efficacy	• How confident did you feel navigating the CONCA website? • How confident did you feel completing the CONCA tasks? • The information on the profile page was easy to understand.	• Understanding of the information contained in the CONCA, confidence using the CONCA, and familiarity with cognitive testing.
Burden	• How much effort did it take to complete the CONCA tasks? • How much effort did it take to review your results on the "My Profile" page?	• Amount of information that needs to be processed to use it, time taken to complete, and level of difficulty navigating the CONCA.
Effectiveness	• The CONCA has improved my understanding of my cognition[b].	• How the CONCA has improved the understanding of cognition, how CONCA meets service need, relationship between CONCA and real-world experiences.
Coherence	• It is clear to me how the CONCA tasks measure my cognition[b].	• Clarity of information provided (text and figures), utility of graphs or scores.
Opportunity costs[c]	• I would like my cognition to be assessed as part of my care. • I would be happy for my psychiatrist to see my scores on the CONCA tasks. • My cognition is a priority for me.	• Willingness to give up some appointment time for the CONCA, priority for cognitive testing or CONCA compared to other priorities.

[a]CONCA: Cardiff Online Cognitive Assessment.

[b]Participants are provided with a definition of cognition ("Thinking skills, including memory, concentration and problem-solving").

[c]The suggested question in Sekhon et al's [37] questionnaire was not well received by our public advisory group and did not fit well with the study design. Therefore, these were replaced with questions about whether cognition is a priority for participants as part of their clinical care.

We have developed a set of thresholds to define *acceptable* in consultation with our public advisory group and taking into consideration our existing data on CONCA completion rates and previous feedback [23]. Based on these discussions, we established a *traffic light* system:

- Green: Clear evidence of acceptability (≥80% positive responses)
- Amber: Evidence of acceptability but some aspects require refinement (60% - 79% positive responses)
- Red: Limited evidence of acceptability (<60% positive responses)

We will apply these thresholds to each of the items in the acceptability questionnaire and the overall completion rate of the CONCA battery.

We will conduct secondary analyses to examine factors that predict participants' responses to the acceptability questionnaire, including CONCA scores, current mental state, and demographic

XSL•FO
RenderX

characteristics (binary logistic regression). Finally, we will examine the correlation between participants' ratings of their cognition and performance on each CONCA task.

Qualitative Interview Analysis

Interviews will be professionally transcribed, and transcripts will be analyzed using a deductive thematic analysis approach. A coding framework will be developed based on the interview topic guides through discussion among team members to identify the key themes and subthemes. A sample of the transcripts will be independently double-coded. Analysis will be supported by the computer-assisted qualitative analysis software NVivo 15 (Lumivero). Data will be analyzed thematically, with particular attention to both positive and negative appraisals of the CONCA, as well as the contextual factors influencing participants' perceptions and the perceived barriers and facilitators to implementation.

Acceptability will be defined as the extent to which participants perceive the CONCA as useful, user-friendly, and aligned with their needs or the needs of the EIP. This will be assessed by identifying themes relevant to the TFA, including participants' positive and negative feelings and attitudes toward the CONCA, the burdens and costs associated with using the CONCA, ethical beliefs around the CONCA, perceived benefits of the CONCA, and willingness to use the CONCA.

Triangulation

The data from the quantitative and qualitative arms will be analyzed independently and in parallel (equal status concurrent design [32]). The results will be integrated at the point of interpretation of the data. Thus, we will corroborate the findings from the 2 approaches to address the study questions.

Public Involvement

We have embedded public participation involvement and engagement (PPIe) across the CONCA project lifecycle. This started with the co-design of the CONCA website with people with lived experience of severe mental illness and health professionals from an EIP. This included public contributions to the interface, logo, and graphics and the extensive testing and refinement of the website. The clinical features were also co-designed and tested by our PPIe group, resulting in substantial changes to ensure the feedback for patients and health professionals produced by CONCA was useful for these groups.

At the outset of the current project, we expanded our PPIe group to ensure a wide range of viewpoints from lived experience experts. The functions of this group are wide-ranging and include developing assessment materials, preparing patient consent forms and information sheets (ensuring lay language is used throughout), reviewing and selecting study questionnaires, and co-designing and reviewing updates to the CONCA website. The group will be part of the team developing future protocols and strategies for full implementation in clinical settings. It will also play a key role in disseminating the outputs from the project to a wider audience, including patients, carers, health care professionals, and policy makers.

We have developed a research partner role and appointed one of our most experienced PPIe members. The research partner acts as a PPIe group liaison, providing peer-to-peer support and ensuring group tasks stay on track and all members are kept fully informed and happy with their roles. Our research partner has been involved in a wide range of activities beyond those listed above, including providing training for new staff and contributing to manuscript writing and other forms of dissemination (including talks and podcasts).

We will assess the impact of our public involvement work using the Public Involvement in Research Impact Toolkit [40].

Plans for Dissemination

The results of this study will be disseminated within the scientific community through conference presentations and publications in peer-reviewed journals. We will update our research participants on the progress of the study through blog posts, podcasts, social media posts, hosted events, and annual reports. We will work with our public advisory group and departmental communications team to explore new ways of disseminating our findings to the public, including the National Health Service, service user groups, and charities that serve people with lived experience of psychosis.

Ethical Considerations

The study will be conducted according to the guidelines of the Declaration of Helsinki. The study was approved by Wales Research Ethics Committee 2 on May 20, 2025 (16/WA/0323). Interested individuals will receive an information sheet outlining the study protocol and be given 48 hours to review this prior to screening. Interested individuals will be screened and given the opportunity to ask questions. Informed consent will be received from all participants over the phone or via a video call with a trained researcher. Consent will be received separately for the quantitative and qualitative data collection. Consent will be recorded electronically through the study website hosted on REDCap. All data will be pseudonymized with participant IDs and will not include any identifiable data. All published data, including quotes from qualitative interviews, will be anonymized to ensure participant confidentiality. All participants will be reimbursed for their time at a rate of £20 (US $27) per hour via bank transfer or gift card.

Results

Funding was obtained from Health and Care Research Wales (Welsh Government, grant number: AF 24 01), and the funding commenced on October 1, 2024. Participant recruitment started in July 2025. As of September 15, 2025, we have enrolled 26 participants with psychosis in the quantitative arm of the study, and 4 participants with psychosis and 1 health professional in the qualitative arm. Data collection is expected to be completed by October 2026.

Discussion

Anticipated Findings

The CONCA has the potential to improve service delivery for patients with psychosis by enabling timely, reliable, and clinically informative cognitive testing to take place in clinics to support the early detection of cognitive problems, treatment

XSL•FO
RenderX

planning, and predictions of prognosis. It is our intention to provide the CONCA on a not-for-profit basis to clinical teams. However, it is currently unclear how cognitive testing would fit into the care pathway for individuals with psychosis. There is a need for data on the acceptability of the CONCA and cognitive testing more generally among patients and clinicians.

The findings from our research will be reviewed with our public advisory group and will be used to identify the aspects of the CONCA that could be refined to ensure that it is usable and accessible and provides the data that clinical teams need. We will base any decision on how to move forward with testing the CONCA on our traffic light system of acceptability, along with our identified themes from the qualitative analysis. Acceptability is a multidimensional construct, and we consider each domain to be essential to the future implementation of the CONCA. Therefore, we will assess all TFA domains individually and consider them to hold equal weight. In the case of any domain of acceptability being rated *red*, the further testing of the CONCA will not take place before both quantitative and qualitative feedback is reviewed, the identified issues addressed, and this domain of acceptability has been remeasured. For domains rated *amber*, we will refine the CONCA based on our findings prior to further testing. For domains rated *green*, we will not make any changes in this area and move forward with testing (assuming other aspects of acceptability have been addressed).

This study will provide the groundwork for a larger hybrid effectiveness-implementation study [41]. Working with our public involvement group and health professionals, we will review the data and themes that emerged from our qualitative analysis to identify facilitators and barriers to the CONCA's implementation and reach a consensus on the changes to service delivery we hope to achieve and how we will measure these. We will develop a protocol that measures the outcomes and impacts of interest.

Strengths and Limitations

A key strength of this study is the involvement of people with lived experience of psychosis and other severe mental illnesses. Our public members contributed to (1) the design of the CONCA at every stage; (2) the development of the study assessments, interview topic guides, information sheets, and consent forms; (3) our recruitment strategy; (4) the equality impact assessment for the study; and (5) the funding application process. They will also have opportunities to contribute to the analysis and dissemination of the study findings. We have applied several existing frameworks to ensure the robustness of our study design, including the CFIR, TFA, and regulatory frameworks for developing digital health technology. The use of these frameworks will ensure that we gather the evidence needed to support the future implementation and sustainability of the CONCA.

The use of a digital cognitive assessment risks excluding certain people, including those with low digital literacy, anxiety or paranoia around digital technology, lack of access to an internet-connected device, or existing cognitive impairment. We have put several mitigations in place to ensure that those who wish to take part can do so, regardless of technology access or familiarity. These include providing support to complete the study (in person or by phone) and providing a device if needed. However, there remains the possibility that people will decline to take part in the study due to the issues listed above. We will calculate completion rates as one aspect of acceptability. However, these may not accurately reflect completion in a clinical setting. Participants may be more likely to complete the tasks as part of this study, as they will have to answer questions about the tasks and will be given remuneration for their time. We will compare rates from this study with our existing data collected as part of our nonremunerated, unsupervised studies [23,24].

There is a risk of allegiance bias, as the CONCA is being evaluated by researchers who were involved in the development of the tool. To mitigate this, the qualitative interviews will be independently double coded by one of the authors (SM), a qualitative researcher who was not involved in the development of the CONCA. We have also prespecified the analyses we intend to undertake and outlined clear thresholds that we will use to define acceptability. Finally, we will ensure that negative feedback is clearly reported, including the proportions of negative responses to the acceptability questionnaire and negative appraisals of the CONCA that arise during the qualitative interviews.

Finally, our recruitment of health professionals may be skewed toward those who are receptive to or interested in the use of digital technology or cognitive assessments in their services. We will use purposive sampling to recruit participants across a range of clinical disciplines, seniority levels, and locations to capture diverse perspectives. Researchers will keep a reflective log during recruitment, which will help us monitor and adjust our strategies as needed.

Acknowledgments

This research was conducted with the support of The Many Brains Project, Inc. The web-based tasks remain the intellectual property of The Many Brains Project, which retains the copyright for these tasks and the source code. The authors thank Dr Laura Germine for her support in developing the Cardiff Online Cognitive Assessment.

We thank our public involvement group and the clinicians from Headroom Cardiff and Vale UHB Youth Psychosis Service for contributing to the development of Cardiff Online Cognitive Assessment.

We thank colleagues at the National Centre for Mental Health for supporting this study, particularly Ms Lyvia Jehu and Ms Katalin Danko for participant recruitment, Dr Sahana Baskar and Mr Daniel Oakes for REDCap (Research Electronic Data

XSL•FO
RenderX

Capture) setup and data management, Ms Anna Simpson for research governance and ethics, Dr Sarah Rees for supporting our public involvement work, and Mr Mark Coles for operational support.

Funding

This study is funded by a Health and Care Research Wales Advanced Fellowship competitively awarded to AJL by the Welsh Government (AF-24-01). The development of the Cardiff Online Cognitive Assessment was supported by two Medical Research Council grants (JTRW and IRJ: grant MC_PC_17212; JTRW, AJL, and IRJ: grant MC_PC_17170), the Wellcome Trust (AJL, JTRW, and IRJ: grant 214601/Z/18/Z), and the National Centre for Mental Health, a collaboration between Cardiff, Swansea, and Bangor universities funded by the Welsh Government through Health and Care Research Wales (IRJ and JTRW).

Data Availability

The data from this study will be held in the National Centre for Mental Health's data repository and available through direct application to the National Centre for Mental Health on reasonable request (see [42] for contact details).

Authors' Contributions

AJL is the lead author and principal investigator and was involved in all aspects of the study, including designing the assessment tools and study methodology and drafting and redrafting the manuscript. JTRW is the senior author and academic mentor for AJL's fellowship and was involved in all aspects of the paper. SM contributed to the design of the qualitative arm of the study, including developing topic guides, and aspects relating to implementation science. AC is our research partner and contributed to the development of the Cardiff Online Cognitive Assessment and the study methodology and wrote the section of the manuscript titled Public Involvement. IRJ was involved in the development of the Cardiff Online Cognitive Assessment and advised on the study methodology. All authors critically revised the paper and approved the final version to be submitted.

Conflicts of Interest

The assessment tool evaluated in this study (Cardiff Online Cognitive Assessment) was developed by AJL, AC, IRJ, and JTRW. JTRW has received funding from Takeda Pharmaceuticals for research unrelated to this manuscript. JTRW and IRJ have received funding from Akrivia Health for research unrelated to this manuscript.

Multimedia Appendix 1
Acceptability questionnaire developed for the study.
[PDF File, 136 KB - resprot_v15i1e84218_app1.pdf]

Peer Review Report 1
Peer-review report from the Health and Care Research Wales Advanced Fellowship Review Committee.
[PDF File, 156 KB - resprot_v15i1e84218_app2.pdf]

References

1. Fatouros-Bergman H, Cervenka S, Flyckt L, Edman G, Farde L. Meta-analysis of cognitive performance in drug-naïve patients with schizophrenia. Schizophr Res 2014 Sep;158(1-3):156-162. [doi: 10.1016/j.schres.2014.06.034] [Medline: 25086658]
2. Mesholam-Gately RI, Giuliano AJ, Goff KP, Faraone SV, Seidman LJ. Neurocognition in first-episode schizophrenia: a meta-analytic review. Neuropsychology 2009 May;23(3):315-336. [doi: 10.1037/a0014708] [Medline: 19413446]
3. Fett AKJ, Viechtbauer W, Dominguez MDG, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci Biobehav Rev 2011 Jan;35(3):573-588. [doi: 10.1016/j.neubiorev.2010.07.001]
4. Vesterager L, Christensen T, Olsen BB, et al. Cognitive and clinical predictors of functional capacity in patients with first episode schizophrenia. Schizophr Res 2012 Nov;141(2-3):251-256. [doi: 10.1016/j.schres.2012.08.023] [Medline: 23017825]
5. Dickson H, Laurens KR, Cullen AE, Hodgins S. Meta-analyses of cognitive and motor function in youth aged 16 years and younger who subsequently develop schizophrenia. Psychol Med 2012 Apr;42(4):743-755. [doi: 10.1017/S0033291711001693] [Medline: 21896236]
6. Watson AJ, Harrison L, Preti A, Wykes T, Cella M. Cognitive trajectories following onset of psychosis: a meta-analysis. Br J Psychiatry 2022 Dec;221(6):714-721. [doi: 10.1192/bjp.2022.131] [Medline: 36149012]
7. Bertolote J, Mcgorry P. Early intervention and recovery for young people with early psychosis: consensus statement. Br J Psychiatry 2005 Aug;187(S48):s116-s119. [doi: 10.1192/bjp.187.48.s116]
8. Psychosis and schizophrenia in adults: prevention and management. National Institute for Health and Care Excellence. 2014 Feb 12. URL: https://www.nice.org.uk/guidance/cg178/resources/psychosis-and-schizophrenia-in-adults-prevention-and-management-pdf-35109758952133 [accessed 2021-05-24]

XSL•FO
RenderX

9. Galletly C, Castle D, Dark F, et al. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the management of schizophrenia and related disorders. Aust N Z J Psychiatry 2016 May;50(5):410-472. [doi: 10.1177/0004867416641195] [Medline: 27106681]
10. Keepers GA, Fochtmann LJ, Anzia JM, et al. The American Psychiatric Association practice guideline for the treatment of patients with schizophrenia. Am J Psychiatry 2020 Sep 1;177(9):868-872. [doi: 10.1176/appi.ajp.2020.177901] [Medline: 32867516]
11. Vita A, Gaebel W, Mucci A, et al. European Psychiatric Association guidance on assessment of cognitive impairment in schizophrenia. Eur Psychiatry 2022 Sep 5;65(1):e58. [doi: 10.1192/j.eurpsy.2022.2316] [Medline: 36059109]
12. Cowman M, Lonergan E, Burke T, et al. Evidence supporting the use of a brief cognitive assessment in routine clinical assessment for psychosis. Schizophrenia (Heidelb) 2022 Dec 17;8(1):113. [doi: 10.1038/s41537-022-00322-z] [Medline: 36528607]
13. Nuechterlein KH, Nasrallah H, Velligan D. Measuring cognitive impairments associated with schizophrenia in clinical practice: overview of current challenges and future opportunities. Schizophr Bull 2025 Mar 14;51(2):401-421. [doi: 10.1093/schbul/sbae051] [Medline: 39088730]
14. Burton CZ, Harvey PD, Patterson TL, Twamley EW. Neurocognitive insight and objective cognitive functioning in schizophrenia. Schizophr Res 2016 Mar;171(1-3):131-136. [doi: 10.1016/j.schres.2016.01.021] [Medline: 26811232]
15. Medalia A, Thysen J, Freilich B. Do people with schizophrenia who have objective cognitive impairment identify cognitive deficits on a self report measure? Schizophr Res 2008 Oct;105(1-3):156-164. [doi: 10.1016/j.schres.2008.07.007] [Medline: 18718740]
16. Keefe RSE, Goldberg TE, Harvey PD, Gold JM, Poe MP, Coughenour L. The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. Schizophr Res 2004 Jun 1;68(2-3):283-297. [doi: 10.1016/j.schres.2003.09.011] [Medline: 15099610]
17. Velligan DI, DiCocco M, Bow-Thomas CC, et al. A brief cognitive assessment for use with schizophrenia patients in community clinics. Schizophr Res 2004 Dec 1;71(2-3):273-283. [doi: 10.1016/j.schres.2004.02.027] [Medline: 15474898]
18. Hurford IM, Ventura J, Marder SR, Reise SP, Bilder RM. A 10-minute measure of global cognition: validation of the Brief Cognitive Assessment Tool for Schizophrenia (B-CATS). Schizophr Res 2018 May;195:327-333. [doi: 10.1016/j.schres.2017.08.033] [Medline: 28918221]
19. Bryce S, Allott K. Cognitive screening: a significant unmet need in youth mental health. Aust N Z J Psychiatry 2019 Aug;53(8):813. [doi: 10.1177/0004867419834356] [Medline: 30848664]
20. Allott K, Van‐der‐el K, Bryce S, et al. Need for clinical neuropsychological assessment in headspace youth mental health services: a national survey of providers. Aust J Psychol 2019 Jun 1;71(2):108-116. [doi: 10.1111/ajpy.12225]
21. Saperstein AM, Medalia A, Bello I, Dixon LB. Addressing cognitive health in coordinated specialty care for early psychosis: real-world perspectives. Early Interv Psychiatry 2021 Apr;15(2):374-379. [doi: 10.1111/eip.12966] [Medline: 32307919]
22. Lavigne KM, Sauvé G, Raucher-Chéné D, et al. Remote cognitive assessment in severe mental illness: a scoping review. Schizophrenia (Heidelb) 2022 Mar 5;8(1):14. [doi: 10.1038/s41537-022-00219-x] [Medline: 35249112]
23. Lynham AJ, Jones IR, Walters JTR. Web-based cognitive testing in psychiatric research: validation and usability study. J Med Internet Res 2022 Feb 10;24(2):e28233. [doi: 10.2196/28233] [Medline: 35142640]
24. Lynham AJ, Jones IR, Walters JTR. Cardiff online cognitive assessment in a national sample: cross-sectional web-based study. J Med Internet Res 2023 Sep 13;25:e46675. [doi: 10.2196/46675] [Medline: 37703073]
25. Unsworth H, Dillon B, Collinson L, et al. The NICE evidence standards framework for digital health and care technologies–developing and maintaining an innovative evidence framework with global impact. Digit Health 2021;7:20552076211018617. [doi: 10.1177/20552076211018617] [Medline: 34249371]
26. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
27. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 May 10;19(5):e153. [doi: 10.2196/jmir.7725] [Medline: 28490417]
28. Attribution 4.0 international deed. Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/deed.en [accessed 2026-02-11]
29. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
30. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation Sci 2009 Dec;4(1):50. [doi: 10.1186/1748-5908-4-50]
31. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]

XSL•FO
RenderX

32. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Kolner Z Soz Sozpsychol 2017;69(Suppl 2):107-131. [doi: 10.1007/s11577-017-0454-1] [Medline: 28989188]
33. The Many Brains Project. URL: https://www.manybrains.net [accessed 2025-09-15]
34. TestMyBrain. URL: https://www.testmybrain.org [accessed 2020-10-28]
35. Iverson GL, Marsh JM, Connors EJ, Terry DP. Normative reference values, reliability, and item-level symptom endorsement for the PROMIS® v2. 0 cognitive function-short forms 4a, 6a and 8a. Arch Clin Neuropsychol 2021 Oct 13;36(7):1341-1349. [doi: 10.1093/arclin/acaa128] [Medline: 33454756]
36. Lynham A, Walters J, Jones I. About us. Cardiff Online Cognitive Assessment. URL: https://www.conca.wales/about.html [accessed 2022-05-04]
37. Sekhon M, Cartwright M, Francis JJ. Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. BMC Health Serv Res 2022 Mar 1;22(1):279. [doi: 10.1186/s12913-022-07577-3] [Medline: 35232455]
38. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
39. Wolcott MD, Lobczowski NG. Using cognitive interviews and think-aloud protocols to understand thought processes. Curr Pharm Teach Learn 2021 Feb;13(2):181-188. [doi: 10.1016/j.cptl.2020.09.005] [Medline: 33454077]
40. Edwards M, Seddon K, Baddeley E, et al. Involving patients and the public in cancer associated thrombosis research: a strategy for success. Thromb Update 2025 Mar;18:100196. [doi: 10.1016/j.tru.2024.100196]
41. Landes SJ, McBain SA, Curran GM. An introduction to effectiveness-implementation hybrid designs. Psychiatry Res 2019 Oct;280:112513. [doi: 10.1016/j.psychres.2019.112513] [Medline: 31434011]
42. National Centre for Mental Health. URL: https://www.ncmh.info [accessed 2026-02-18]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
CONCA: Cardiff Online Cognitive Assessment
EIP: early intervention in psychosis service
PPIe: public participation involvement and engagement
REDCap: Research Electronic Data Capture
TFA: theoretical framework of acceptability

Edited by J Sarvestan; submitted 16.Sep.2025; this is a non–peer-reviewed article;accepted 24.Jan.2026; published 03.Mar.2026.

Please cite as:
Lynham AJ, Cope A, Milosevic S, Jones IR, Walters JTR
Acceptability of the Cardiff Online Cognitive Assessment for Clinical Screening of Patients With Psychosis: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e84218
URL: https://www.researchprotocols.org/2026/1/e84218
doi:10.2196/84218

© Amy J Lynham, Anthony Cope, Sarah Milosevic, Ian R Jones, James T R Walters. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 3.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Implementation Science to Improve Health Care Access and Quality for People With Traumatic Brain Injury–Related Morbidity (I-HEAL): Protocol for a Translational Multiproject Program Award

Risa Nakase-Richardson[1,2,3*], PhD; Jolie N Haun[3,4*], EdS, PhD; Megan Moore[5*], MSW, PhD; Jennifer Bogner[6*], PhD; Jeanne M Hoffman[7*], PhD; Kristen Dams-O'Connor[2,8,9,10*], PhD; Jill Coulter[11*], BS; Tracy Kretzmer[12*], PhD; Aaron M Martin[1,12,13*], PhD; Marc A Silva[1,12,13*], PhD; I-HEAL Consortium[3,14,15*]

[1]Department of Neurosurgery, Brain, and Spine, Morsani College of Medicine, University of South Florida, Tampa, FL, United States

[10]Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[11]Stakeholder with living experience of a caregiver for a person with TBI, Boston, MA, United States

[12]Mental Health & Behavior Sciences, James A. Haley Veterans' Hospital, Tampa, FL, United States

[13]Department of Psychiatry and Behavioral Neurosciences, Morsani College of Medicine, University of South Florida, Tampa, FL, United States

[14]See Acknowledgments

[15]

[2]Tampa VA Research and Education Foundation, Tampa, FL, United States

[3]Research and Development Service, James A. Haley Veterans' Hospital, Tampa, FL, United States

[4]Division of Epidemiology, Department of Internal Medicine, Spencer Fox Eccles School of Medicine, University of Utah, Salt Lake City, UT, United States

[5]University of Washington School of Social Work, Seattle, WA, United States

[6]Department of Physical Medicine and Rehabilitation, College of Medicine, The Ohio State University, Columbus, OH, United States

[7]Department of Rehabilitation Medicine, University of Washington School of Medicine, University of Washington, Seattle, WA, United States

[8]Department of Rehabilitation and Human Performance, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[9]Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, United States

[*]all authors contributed equally

Corresponding Author:
Risa Nakase-Richardson, PhD
Department of Neurosurgery, Brain, and Spine, Morsani College of Medicine, University of South Florida, Tampa, FL, United States

Abstract

Background: People with traumatic brain injury (TBI) morbidity (impaired cognition and behavioral regulation) and polytrauma comorbidity (depression, posttraumatic stress disorder [PTSD], chronic pain, and sleep disorders) experience health care inequities. Among Veterans and Service Members (V/SMs), TBI morbidity or polytrauma comorbidity may impact access and meaningful engagement in the high-quality health care needed to reduce poor health care outcomes. The National Academy of Science, Engineering, and Medicine Report on Accelerating Progress in TBI highlights a dearth of implementation science research in TBI that may help overcome health care access challenges. Implementation science uses a mixed methods approach to understand, implement, and examine outcomes associated with using evidence-based care in practice.

Objective: The I-HEAL (Improving Health Care Access and Engagement for Veterans and Service Members with TBI Morbidity) protocol includes 4 synergistic projects with the goal of addressing key knowledge gaps that will improve access and engagement in high-quality, evidence-based health care services for V/SMs with TBI morbidity. Collectively, the 4 projects propose to: (1) adapt existing interventions to promote access and engagement in health care; (2) engage stakeholder communities to maximize uptake and translation; (3) promote research translation that informs policy and practice through knowledge translation products and deliverables targeting key partners (clinicians, V/SMs, caregivers, policymakers, and researchers); (4) facilitate research and implementation to enhance access to high-quality health care for V/SMs with TBI-related morbidity; and (5) foster the development of early/mid-career researchers in advancing implementation science research on access to care for V/SMs with TBI.

Methods: Project 1 will involve the development of a nudge intervention (electronic health care reminder) for providers to engage health care proxies when interacting with cognitive disability at risk for poor health care engagement. Project 2 will involve the development of a provider toolkit of adaptations of guideline-endorsed behavioral health interventions for common

XSL•FO
RenderX

polytrauma comorbidities to meet the needs of cognitively impaired individuals. Project 3 will involve the adaptation and dissemination of evidence-based team interventions for managing maladaptive behaviors after TBI. Project 4 will involve evaluation and recommendations for policy for virtual health modalities among persons with TBI and polytrauma comorbidity.

Results: I-HEAL has been funded as an implementation science Focused Program Award by Congressionally Directed Medical Research Programs, and start-up activities began in October 2023. All 4 projects are currently underway with funding through September 2027. Project 1 has enrolled 48 participants, and project 3 has enrolled 34 participants through September 2025.

Conclusions: TBI is associated with increased health care utilization, comorbid health conditions, and premature mortality. This study has proposed to utilize strategies from the implementation science field to help overcome barriers to physical and psychological health care in order to reduce health care disparities associated with TBI disability.

International Registered Report Identifier (IRRID): DERR1-10.2196/79738

(JMIR Res Protoc 2026;15:e79738) doi:10.2196/79738

KEYWORDS

brain injuries, traumatic; health care quality; health care access; health care evaluation; community-based participatory research; veterans; military personnel

Introduction

Focus Areas

The ability to access and engage in high-quality, evidence-based health care is critical to achieving positive health outcomes. Traumatic brain injury (TBI) morbidities, such as cognitive impairment and behavioral dysregulation, influence a person's ability to access high-quality health care. This proposal uses innovation development to adapt systems of care to improve access and/or engagement in evidence-based care for persons with TBI morbidity in order to achieve improved health outcomes. The proposal focuses on a gap highlighted in the National Academies of Science, Engineering, and Medicine (NASEM) report on accelerating progress in TBI, specifically, promoting research translation using implementation science [1]. The individual studies in the proposal are a series of innovations to improve access to health care by overcoming barriers that have led to unmet care needs, delivery of non–evidence-based care, and subsequent health care disparity for Veterans and Service Members (V/SMs) with TBI-related morbidities (cognition, behavior, and psychological functioning). The individual studies uniquely target supply and demand dimensions that influence health care access to optimize outcomes by utilizing implementation science interventions paired with a community-based participatory action approach. Brief descriptions of the projects are as follows:

- For those with significant cognitive impairments, project 1 will develop a systems intervention to cue providers to include health care proxies in appointments to maximize compliance.
- For those with cognitive impairments, project 2 will develop provider resources to adapt evidence-based behavioral health therapies for common comorbid conditions (posttraumatic stress disorder [PTSD], depression, sleep disorders, and chronic pain) to maximize engagement.
- For those with maladaptive behaviors, project 3 will develop team-based resources to increase workforce capacity in delivering evidence-based behavior management, thus decreasing reliance on pharmacological restraints and denied access to rehabilitation services.
- For those with all types of TBI burdens, project 4 will generate data-driven policy recommendations to adapt universal virtual health care mandates to accommodate the needs of persons with TBI.

Individual projects will interact with 3 robust cores: Implementation Science Core (ISC), Website and Data Management Core (WDMC), and Community Engagement Council (CEC), which will include key partners representing the end users of each innovation proposed. As requested in the funding opportunity announcement, this proposal will address the implementation of evidence-based treatments and engage partners to promote the adoption of evidence-based care for optimal outcomes after military-related TBI. Multimedia Appendix 1 provides information on the impact and relevance of the proposed study for military health.

Overarching Challenge

Our work in the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR) and Veterans Affairs (VA) TBI Model Systems (TBIMS) has highlighted that persons with TBI and greater neurologic morbidity are 6.9 times more likely to die than the general population, resulting in a 12.2-year reduction in life expectancy [2]. When examining service utilization, we found that those with greater neurologic burden experienced a 2-fold increase in rehospitalization rates in the first year after TBI compared to those with a lesser neurologic burden. V/SMs have higher rates of rehospitalization (44%) compared to civilians in the TBIMS [3]. Further, those with greater neurologic burden have higher rates of death and have distinct causes of early mortality (self-inflicted death and preventable injuries) across groups [4,5]. TBI morbidity (impaired cognition and behavioral regulation) coupled with polytrauma comorbidity (depression, PTSD, chronic pain, and sleep disorders) can impact access to and meaningful engagement in high-quality health care to reduce these poor health care outcomes [6]. This disparity is further evidenced by our work highlighting unmet health care needs in chronic stages [7] and comorbidities associated with worse outcomes [8]. The recent NASEM report on accelerating progress in TBI recognizes and calls for action to improve health care access for persons with TBI [1]. Poor access to health care is an

XSL•FO
RenderX

inequity that can be addressed by using implementation science. The NASEM highlights a dearth of implementation science research in TBI that could help improve access and close the quality-of-care chasm experienced by persons with TBI disability [1].

Overarching Solution

There is a growing recognition of significant delays in research translation across scientific fields. The emerging field of implementation science has been recognized as a critical field in the research translation process. Implementation science is the scientific study of methods and strategies that facilitate the uptake of evidence-based practice and research into regular use by practitioners and policymakers. Implementation science is complex and requires the use of mixed methods approaches to understand, implement, and examine outcomes associated with promoting the use of evidence-based care in practice. The field also recognizes that the use of evidence-based practice requires adaptations or innovations in how care is delivered (to retain technical quality and fidelity) across health care systems (eg, VA, Department of Defense [DOD], civilian, rural, and urban settings) and populations (eg, disabled vs healthy cohorts). Our work has shown that persons with TBI-related morbidity and comorbidity experience decreased access to health care in general, including known evidence-based treatments for physical and psychological health care. The purpose of this proposal is to address this health care inequity informed by the VA Quality Enhancement Research Initiative (QUERI) Implementation Roadmap, which defines 3 distinct stages to implementation science research (preimplementation, implementation, and sustainment) [9]. All projects fall in the preimplementation phase (adaptation and innovation to prepare for implementation), which often includes identifying partners, understanding needs, and developing knowledge products to facilitate the next stage (implementation). The measurable primary endpoints for this preimplementation phase are presented as milestones in the statement of work of I-HEAL (Improving Health Care Access and Engagement for Veterans and Service Members with TBI Morbidity) (Multimedia Appendix 2).

Knowledge products that synthesize scientific evidence from the literature or represent innovations in health care are critical to timely translation efforts. The VA QUERI Roadmap highlights phases of translation for bringing evidence-based health care to patients [9]. In the preimplementation phase (focus of this proposal), there is a need for knowledge products that clarify the problem being addressed, best treatment practices, and partners involved. The knowledge products generated across cores and projects (using community-based participatory research [CBPR]) will be informed by conceptual frameworks used in the field of implementation science, as described in the core descriptions. The second implementation phase includes preparing for and conducting implementation to promote the use of evidence-based or innovative care. The sustainment phase evaluates long-term impacts and assignment of implementation ownership. Given the scope of the funding opportunity announcement, the actual implementation (stage 2) and sustainment phases (stage 3) of the research will be addressed in future proposals and are described in the transition plan (Multimedia Appendix 3). Throughout this process, new questions arise, which require the generation of new data and reinitiation of the roadmap to address the evolution of health care delivery. The Focused Program Award (FPA) will generate products relevant for the preimplementation and implementation phases of this model. Knowledge products will be produced targeting key partners in the adoption of the study findings, which include (1) other scientists, (2) clinicians (physicians, behavioral health specialists, physical/occupational/speech therapists, etc), (3) policymakers, (4) hospital administrators, (5) trainees, and (6) persons with TBI and their families. As the products are finalized, they will be hosted on the study website, which will serve as a repository of tools (I-HEAL Toolshed) available to the public.

Overall Program Synergy and Objectives

Developing Knowledge Products to Improve Health Care Access

The I-HEAL proposal includes 4 synergistic projects with the goal of addressing key knowledge gaps that will improve access and engagement in high-quality, evidence-based health care services for V/SMs with TBI morbidity.

Overarching Objectives of the FPA

Collectively, the 4 projects propose to accomplish the following overarching objectives of the FPA: (1) adapt existing interventions to promote access and engagement in health care; (2) engage stakeholder communities to maximize uptake and translation; (3) promote research translation that informs policy and practice through knowledge translation products and deliverables targeting key partners (clinicians, V/SMs, caregivers, policymakers, and other researchers); (4) facilitate research and implementation to enhance access to high-quality health care for V/SMs with TBI-related morbidity; and (5) foster development of early/mid-career researchers in advancing implementation science research on access to care for V/SMs with TBI to expand the implementation science workforce in this understudied, underutilized field.

Framework Informing the Mechanistic Targets and Impacts of Project Interventions

The proposed studies will be organized around a patient-centered conceptual framework of accessing evidence-based health care services to achieve optimal health outcomes [10]. The framework highlights how the health care environment (supply side, top) and the community that it supports (demand side, bottom) are multidimensional and interact to influence health care access [10]. The core elements of this model include having a health care need, the perception of the health care need, the desire for treatment, seeking and reaching health care, and utilizing health care to achieve an optimal outcome. The projects have unique interventional targets in the health care system (provider, team, system, and policy) to improve access and increase impact aligned with the best practices for addressing disparities [11]. The studies proposed in the FPA will target health care access dimensions on either the health care supply or consumer demand side that are known to be barriers arising from TBI-related morbidity, such as cognitive impairment and behavioral dysregulation. Each study will target a health care access dimension with primary and secondary targets. The

XSL•FO
RenderX

interventions address TBI morbidity (impaired cognition and behavioral dysregulation) and high-frequency polytrauma comorbid conditions for military-related TBI, such as chronic pain, sleep disorders, and psychological health [12]. Project deliverables will be used to inform translational steps to promote dissemination and uptake across health care settings (ie, future implementation studies involving federal agencies such as VA, Agency for Healthcare Research and Quality, National Institutes of Health [NIH], and Patient-Centered Outcomes Research Institute [PCORI]). The framework and alignment with proposed projects are presented in Table 1.

Table . Access to the care conceptual framework and alignment with proposed projects.

Focused Program Award proposed projects	Health care supply dimensions of health care access					Consumer demand dimensions of health care access				
	Approachability	Acceptability	Availability and accommodation	Affordability	Appropriateness and technical (high) quality	Ability to perceive the need for services	Ability to seek services	Ability to reach services	Ability to pay for services	Ability to engage in high-quality care
Project 1[a]	Secondary impact	Secondary impact	Secondary impact	Secondary impact	Primary impact	Secondary impact	Secondary impact	Secondary impact	—[b]	Target
Project 2[c]	Secondary impact	Secondary impact	Secondary impact	Secondary impact	Primary impact	—	—	Secondary impact	—	Target
Project 3[d]	Secondary impact	Secondary impact	Target	Secondary impact	Secondary impact	—	—	Primary impact	—	Secondary impact
Project 4[e]	Secondary impact	Secondary impact	Target	Secondary impact	Secondary impact	Secondary impact	Secondary impact	Secondary impact	Secondary impact	Primary impact

[a]Cognitive flag to nudge providers to include family in health care delivery.

[b]Not applicable.

[c]Cognitive adaptation toolkit for behavioral health interventions.

[d]Adapting evidence-based behavioral treatment for managing challenging behavior.

[e]Understanding virtual health care accommodation needs for traumatic brain injury.

Methods

Ethical Considerations

The Institutional Review Board (IRB) approved a combined protocol for projects 1 - 3 on August 12, 2024. Aim 4.1 is a secondary analysis from a larger existing IRB-approved study (December 2, 2019; "Characterization and treatment of chronic pain after moderate to severe traumatic brain injury: a qualitative study" [PR00039496]). All principal investigators (PIs) are currently on the existing IRB-approved study. Completion of aim 4.1 will require a modification to the parent study to acknowledge a new funding source for the secondary analysis. Informed consent will be obtained for all participants in projects 1-3, with the ability to opt out at any point. During consent, participants will be assured of their rights and privacy. Participants in focus groups and interviews will be compensated US $50. Participants in the online survey will be randomly selected to receive US $50. The consent of participants from the original research project covers project 4's secondary analysis of existing data. All reported data will be deidentified and in aggregate form.

Oversight

This research project has been reviewed and found to align with the mission of the VA and be scientifically valid, and has been reviewed by all appropriate subcommittees to ensure the safety of the study participants and VA staff. Approval has been granted by a convened board review of the James A. Haley Veterans' Hospital Research and Development Committee (IRBNet #1804543). I-HEAL will be reviewed at least annually or as otherwise required by the appropriate Research & Development Committee or subcommittees.

Project Core Descriptions

ISC Details

The I-HEAL FPA will extend the successful collaboration among study investigators to advance implementation science within the field of TBI. The goals of the ISC are as follows:

1. Support individual projects from conceptualization and support methodological development, qualitative data collection/analysis, interpretation of the findings, and dissemination using integrated conceptual frameworks to inform methodological approaches and knowledge development.
2. Support implementation mentorship, training, and professional development for early career researchers and TBI-focused scientists to enhance workforce capacity in implementation science.
3. Support dissemination efforts of knowledge translation products and support the overall development of the I-HEAL Toolshed with partners in the CEC.

XSL•FO
RenderX

ISC Individual Project Support

Overview

The ISC will use broad-based efforts within I-HEAL and across projects to inform rigorous implementation science methodologies and outcomes. The ISC will work with the other I-HEAL cores to identify evidence-based strategies that shorten the time between evidence production and implementation into practice and policy, to promote access, and to ensure quality improvement (QI) in health care and outcomes by (1) collaborating with project teams to assist with qualitative data collection, data analysis, and dissemination throughout and after project completion; (2) iteratively assessing and responding to the needs of projects to ensure that the projects optimize implementation and translation opportunities; (3) defining and documenting the best practices in implementation science methods, deliverables, and outcomes; and (4) supporting product and Toolshed development for key stakeholder audiences to accelerate dissemination and implementation of evidence-based practices and maximize impact.

In alignment with the VA QUERI Implementation Roadmap, the proposed I-HEAL objectives and projects are in the preimplementation phase, with a focus on access to care for persons with TBI, as guided by the Levesque Access Framework to support access to care for persons with TBI. As illustrated in Figure 1, the I-HEAL team will achieve the I-HEAL objectives by conducting the proposed projects to support adaptation and development of products and outcomes that increase access to evidence-based care for persons with TBI. To operationalize the I-HEAL projects, protocols, processes, and outcomes, each project will take a transtheoretical approach to integrate (1) CBPR, (2) human-centered design (HCD), and (3) an integrated implementation science framework (Reach, Effectiveness, Adoption, Implementation, Maintenance [RE-AIM]; Consolidated Framework for Implementation Research [CFIR]; and evidence-based implementation strategies).

Figure 1. The multifactorial solution of I-HEAL (Improving Health Care Access and Engagement for Veterans and Service Members with TBI Morbidity) for addressing health care access inequality among Veterans and Service Members (V/SMs) with traumatic brain injury (TBI). CFIR: Consolidated Framework for Implementation Research; EBP: evidence-based practice; RE-AIM: Reach, Effectiveness, Adoption, Implementation, Maintenance.

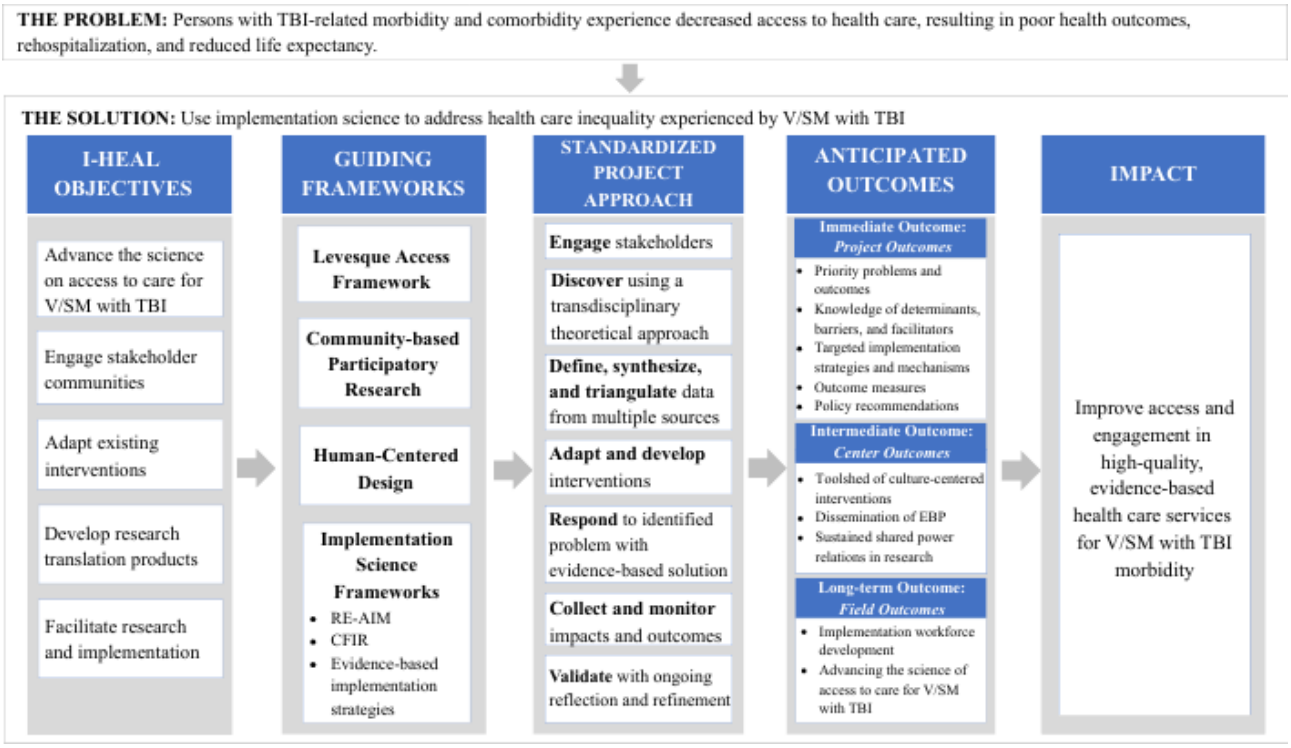

HCD Approach

The proposed I-HEAL projects align with the preimplementation phase of the QUERI Roadmap through the development of knowledge products from published literature and other sources. An HCD process will guide the design and implementation of the proposed I-HEAL projects. HCD is an approach that "centers" end users (ie, the people who will ultimately use or benefit from a product or service) by integrating them into every aspect of the design process, thus strengthening the impact and utilization of the product or service [13-15]. HCD can combine mixed methods with novel efforts to engage end users in an iterative design process. Each of our projects' aims aligns with the 4-phase Double Diamond HCD model [16]. This model directs researchers and designers to engage end users to first *discover* and *define* a problem and then take focused action to *develop* and *deliver* (ie, validate) a solution [17]. HCD supports the process of making high-quality evidence-based tools that are effective, feasible, and aligned with the needs and goals of users and partners.

Implementation Science Framework

I-HEAL and its proposed projects will leverage an integrated implementation science framework that includes RE-AIM, CFIR, and evidence-based implementation strategies. The RE-AIM framework focuses explicitly on issues and steps in

XSL•FO
RenderX

the implementation process that may improve/impede the desired impact. Due to the preimplementation nature of the proposed projects, though not immediately relevant, RE-AIM provides an overarching framework for the lifetime of these projects after the FPA, down the implementation pipeline. Within the phases of RE-AIM, CFIR is a practical evidence-based implementation framework to examine and organize determinant factors that influence implementation outcomes across the following 5 domains: intervention/project characteristics, outer and inner settings, characteristics of individuals, and the process of implementation. At this point, evidence-based implementation strategies can be mapped based on relevance to determinant implementation factors.

Collectively, these theoretical frameworks inform the I-HEAL conceptual lens (see Figure 1) across projects to standardize our approach to implementation over time.

ISC Mentorship, Training, and Professional Development

Overview

The ISC will support mentorship, training, and professional development for the next generation of rehabilitation scientists to support the development of a TBI implementation science workforce. Our objectives are to (1) provide mentorship and training for the next generation of TBI-focused implementation scientists to democratize implementation and knowledge translation expertise in the field of TBI research and (2) identify and facilitate implementation-focused training and professional activities for I-HEAL investigators, which are available within and beyond the VA/DOD network.

These 2 training objectives will be accomplished using the following strategies:

1. Project-based mentorship: Expedite and align the mentorship process by matching prospective mentees to mentors with research portfolios and content expertise, as evidenced by the teams of the 4 projects. These individualized practical experiences will support mentees in developing research programs and co-authoring with mentors.
2. Structured I-HEAL ISC training: Implement a planned structured curriculum through regular I-HEAL meetings that will provide (1) dedicated time to support and train with mentors, (2) dedicated research presentations, (3) training developed by RN-R and JNH [18] (ie, implementation science theoretical models, implementation research logic modeling, implementation planning, and playbook development), and (4) evidence-based resources available from the VA QUERI [19]. The *Using Implementation Facilitation to Improve Healthcare* Manual [20] is an evidence-based training manual designed to "support implementation of evidence-based practices and programs and other clinical innovations." Collectively, the I-HEAL training program will address an array of topics, including but not limited to implementation research designs; conceptual frameworks and theories; mixed methods and evaluation methods; assessment of organizational, provider, and veteran readiness; and customization of strategies to overcome significant barriers.
3. Training opportunity registry: Develop and sustain an I-HEAL training registry to identify training and professional development opportunities (eg, workshops, seminars, webinars, etc) with relevant details (ie, training dates, trainee eligibility, costs associated with training, etc).
4. Professional presentation opportunities: ISC mentor and mentee attendance at professional conferences will be a priority benchmark each year of the funded FPA.
    - ISC training transition plan: Senior investigators will provide regular team training workshops for team members, where early career investigators will attend and learn methods and areas of expertise. Collectively, the I-HEAL group (senior and early career scientists) will work to enhance and tailor training. These trainings will be developed, marketed, and disseminated as workshops to professional venues. These workshops will be co-presented by senior and early career investigators and will eventually transition leadership to early career I-HEAL scientists. These efforts will advance the skills of I-HEAL team members as well as support workforce development in the TBI and rehabilitation professional communities.
5. Mentee research experience translation: I-HEAL will support early career scientists and TBI subject matter experts (SMEs) to successfully translate their research experience with I-HEAL into innovative proposals and manuscripts throughout and after the FPA, by fostering a dynamic, learning environment and providing opportunities to develop core competencies (eg, communication, professional and scholarly development, partnered research, and the application of dissemination and implementation science).
    - Training in partnered research: Mentees are full members of I-HEAL, and the proposed FPA will create a collaborative research environment with a variety of investigators and partners. While participating in partnered research with mentors, mentees will have opportunities to interact with local, FPA CEC partners as well as community, academic, and clinical partners. Investigators will actively assist trainees to develop their own partnerships through a variety of methods, such as modeling behaviors, making introductions, helping them to prepare for meetings with prospective partners, and debriefing after interactions with partners. RN-R, JNH, and JMH have strong backgrounds in professional mentorship and will use their developed mentorship strategies to support I-HEAL mentorship and professional advancement.

ISC Dissemination

Our goal is to spread I-HEAL's impact, focusing on health systems and the delivery of TBI care within VA/DOD and beyond. I-HEAL's dissemination efforts will be led by the ISC, with strategic planning that targets audiences and channels with consideration of the impact. Target audiences are (1) I-HEAL partners, leadership, CEC partners, and investigators; (2) DOD

XSL•FO
RenderX

and VA funding agencies; (3) associated VA leadership; (4) university and educational affiliates; (5) policy groups such as professional organizations and federal agencies; (6) local communities; and (7) V/SM service organizations through the CEC. The channels of communication will minimally include (1) written communication (eg, peer-reviewed publications, issue briefs, formal reports, national magazines, and newsletters), (2) oral communication (eg, scheduled and ad hoc face-to-face and telephone meetings with operations partners and steering committee members, professional meetings, annual VA research day, cyber seminars, and grand rounds), (3) electronic communication (eg, social media and the I-HEAL Toolshed; I-HEAL will have a social media presence and explore implementation and dissemination strategies to amplify dissemination), and (4) product development (eg, toolkits, playbooks, and multimedia products). To optimize dissemination, the ISC has developed a publication pipeline protocol that strategizes and documents benchmarks, timelines, and roles to support accountability and productivity. Impacts of these broad-based efforts include but are not limited to increased awareness, trust, and veteran and employee engagement; distribution of knowledge; and relationship building and networking with stakeholder groups, study participants, mentees, and potential new investigators.

ISC Potential Impact

An overview of the potential impact is provided in Figure 2. Across cores and projects, this proposal has gathered a transdisciplinary team of scientists who have collectively demonstrated a history of rapid development and implementation of projects with the ability to rapidly pivot and assemble strong approaches to evaluate priorities in TBI and implementation. I-HEAL projects will contribute to (1) improving access and outcomes for persons with TBI and relevant partners; (2) tailoring policy recommendations, products, and outcomes for persons with TBI, hosted on the I-HEAL Toolshed; (3) advancing the science of implementation, translation, and dissemination; and (4) expanding the implementation and dissemination of science expertise for I-HEAL TBI scientific team members, which will ultimately support the implementation of professional development for workforces that specialize in TBI research.

Figure 2. Overview of the impact of the Implementation Science Core. I-HEAL: Improving Health Care Access and Engagement for Veterans and Service Members with TBI Morbidity.

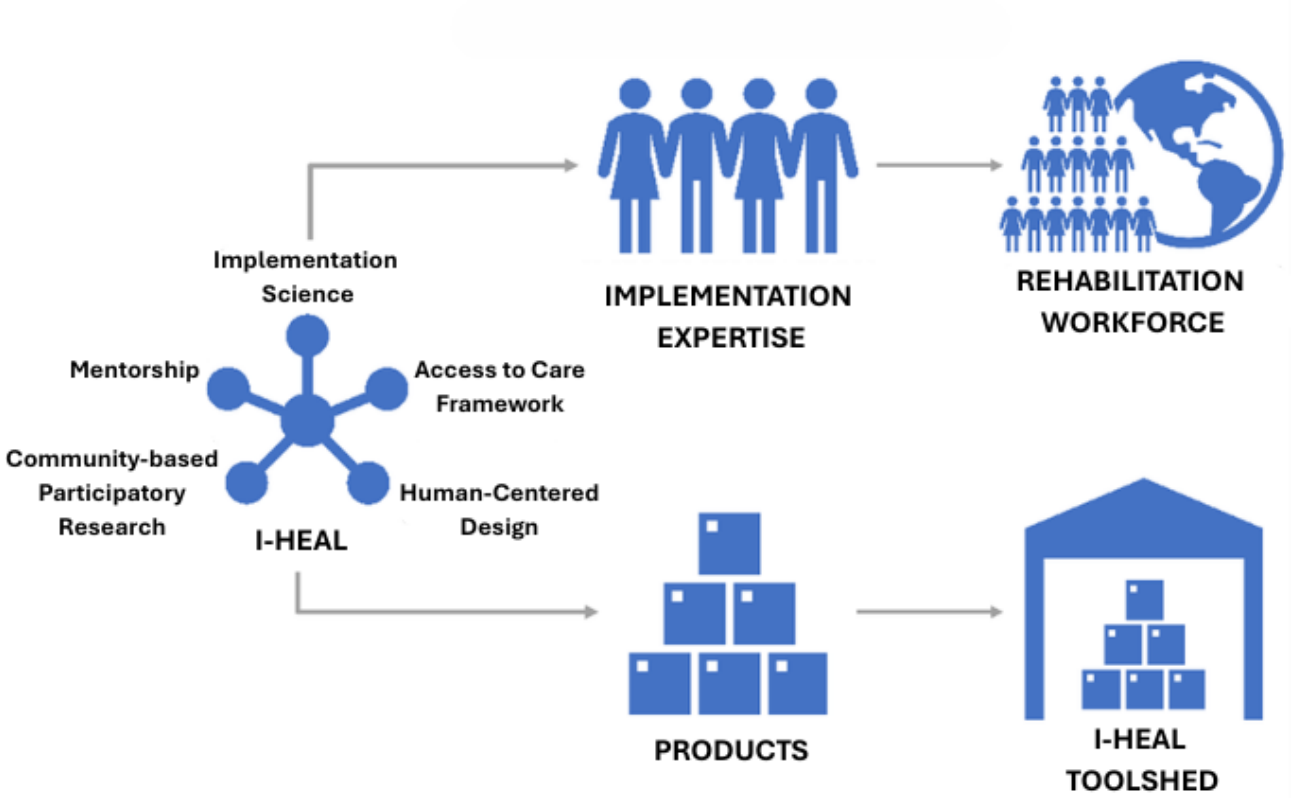

CEC Partnership Plan and Outreach Plan

Overview

This proposal was developed by our team of investigators, policy and professional partners (PPPs), and lived experience partners (LEPs). We have extensive experience engaging in CBPR together. The partnerships outlined here are built on long-term relationships and ongoing collaborations. Aligning with the principles of CBPR, we ground our work in the shared values of co-learning, mutual benefit, and community-driven solutions and practices for addressing the collaboratively identified problem of access to care for persons with TBI [21,22]. The current proposal grew out of years of working closely with professionals and LEPs who identified access to care as the most pressing problem facing V/SMs with TBI and their caregivers. Members of our team have collaborated on several successful projects over many years, including a large-scale randomized controlled trial of a patient-centered care management intervention for persons with TBI (PCORI

XSL•FO
RenderX

R-1511 - 33005), the BeHEALTHY model for chronic brain injury care (NIH-NINDS U24NS095871), Improved Understanding of Medical and Psychological Needs (I-MAP) in V/SMs with chronic TBI (DHA Contracting Office W91YTZ-13-C-0015, HT0014-19-C-0004, HT0014-21-C-0012, and HT0014-22-C-0016), and the BRITE Study (improving transition from inpatient rehabilitation following TBI; NCT03422276).

Building on these existing partnerships and the previously identified priority problem of access to care, we convened in 2022 in Tampa, FL, for a 2-day intensive retreat to develop the I-HEAL objectives, develop the proposed studies, solidify the partnership plan, develop the dissemination and outreach plan, and outline the functions of the CEC. See the CBPR statement for an overview of our engagement process, beginning with our existing partnerships and continuing throughout the stages of the proposal (Multimedia Appendix 4).

CEC Partnership Plan

The CEC partners, purpose, and processes for achieving deliverables, functions, and specific deliverables are outlined below. We have structured our partnership plan to reflect the importance of engagement with a broad range of individuals working to improve the lives of V/SMs with TBI. We have conceptualized the CEC to have 3 core partner groups with synergistic purposes. Important aspects of impactful science to address access to care for V/SMs with TBI are the voices of persons with lived experience of TBI, the system-level and disciplinary knowledge of successful policy and dissemination strategies, and the presence of individuals with expertise on the specific study issues. We have developed a CEC with expertise in each of these core areas. Each partner brings passion and commitment to improving the lives of V/SMs with TBI. They bring unique and complementary knowledge and will engage with all cores and individual projects to achieve the overall I-HEAL objectives.

CEC Partners and Purpose

The composition of the CEC is presented in Figure 3. We have established a CEC comprised of persons from influential organizations representing professionals who will engage with I-HEAL products (PPPs) and individuals who can speak about their lived experience of TBI or caregiving for a loved one with TBI (LEPs). We have representation from members of key professional groups involved in health care policy development for VA and the DOD, and providers representing professional organizations (physiatry, social work, psychology, rehabilitation, and nursing) involved in the care of persons with TBI. We have engaged with a diverse group of LEPs with military service backgrounds and with experience navigating health care in military and civilian environments. Members include persons with TBI (V/SMs and civilians) with greater neurologic morbidity and their spouses. Each member has been selected because of their lived experience of the challenges being addressed in this proposal (individuals with brain injury and cognitive morbidity, legal guardians of V/SMs without decision-making capacity, caregivers of veterans with severe maladaptive behaviors living in the community, and individuals experiencing challenges in navigating the health care system). Many have served as engagement partners on other TBI research projects or as members of the Tampa VA Veteran Engagement Council. The PPPs will focus on dissemination and policy development, and the LEPs will focus on the impact to those with lived experience and the identification of potential facilitators and barriers to successful implementation and outcomes. In addition, individual project engagement partners (PEPs) have been identified by each proposed project team to support the specific needs of the projects.

XSL•FO
RenderX

Figure 3. Community Engagement Council composition.



CEC Processes

The CEC will be involved in all aspects of the proposed research implementing the I-HEAL objectives. Drs John Corrigan, MM, and Chaady Radwan will co-chair the council. Dr John Corrigan will be responsible for chairing the PPP quarterly meetings. The LEPs will elect a chair who will co-lead the monthly LEP meetings with Drs MM and Chaady Radwan. This group will meet more frequently to ensure that each member feels prepared to participate in the co-learning process and has adequate time to contribute their expertise to the research process using a CBPR approach (Multimedia Appendix 4). In years 1 and 4, all members of the council will meet in person with all other participants in center activities to allow for interaction as the studies begin (year 1) and allow for collaboration as results emerge and dissemination and policy development plans are solidified (year 4).

CEC Meeting Schedule

Information on the meeting schedule is provided in Multimedia Appendix 5. First, Dr MM (multiple principal investigator [MPI]) will oversee ongoing, bidirectional communication between the cores, the study teams, and the CEC by attending the all-hands calls, executive committee, steering committee, and core meetings. The co-chair of the LEP group will attend these meetings monthly. Second, Dr John Corrigan, MM, or Chaady Radwan will attend quarterly meetings with the project teams to monitor engagement needs. In addition, the study teams will rotate attending the meetings of the PPPs and LEPs based on their engagement needs as described in their project methodology below. The proposed engagement from study teams will be incorporated into regularly scheduled CEC meetings. It is expected that communication via email will occur between regularly scheduled meetings and that partners will be available on an as-needed basis between regular meetings. These

XSL•FO
RenderX

meetings will be in addition to the in-person meetings in years 1 and 4.

CEC Functions

The CEC will be involved in all aspects of the proposal, including the planning, conduct, and dissemination of results. The planning of this proposal is based on over 25 years of patient-centered research conducted by our team members, an extensive literature review, discussions with our CEC partners, and a 2-day intensive retreat used to solidify the plan and partnership roles. Priority problems and outcomes have been identified through these processes. Moving forward and prior to the initiation of each study, the CEC will advise on consent forms, recruitment processes, study rollout, advertisements, and other materials provided to study participants. The CEC will participate in problem-solving for each study as issues arise. In particular, the LEPs will assist in the identification of barriers and facilitators to successful outcomes via engagement with the study design, progress, and data interpretation. They will provide their views on the results and data interpretation. The PPPs will provide insights into the policies and priorities of the constituents they represent as relating to each study and to the overall access to care for persons with TBI. The overarching goal of their collaboration is to ensure that the study results will be relevant, practical, and acceptable to members belonging to their respective professional organizations. Their involvement will be especially critical to the dissemination and policy development activities described later. The CEC will also assist in community outreach, dissemination, and policy implications. The LEPs will advise on the development of dissemination materials for individuals who have had a TBI and their family members. Together with our outreach and media specialist, they will co-develop social media posts, print materials, and brief multimedia presentations that will be used to disseminate study materials and findings throughout the course of the 4-year project (see the Outreach Plan subsection below). The PPPs will advise on successful dissemination strategies as well as relevant financing or other policy-related issues that influence the implementation and sustainability of products. We will leverage the significant leadership platforms in organizations within the VA and professional trade groups and the leadership of brain injury researchers to understand and advance policy and clinical practice guidelines emanating from the study results. We will work closely with the CEC to ensure engagement remains rooted in the principles of CBPR. We will utilize standard metrics of engagement success, including meeting attendance, self-reported empowerment within the group, and objective measures of engagement success pertinent to the completion of each study (see details in the CBPR statement in Multimedia Appendix 4).

Outreach Plan

Specifically, we will disseminate project updates and results via Twitter, Facebook, and outlets specific to V/SMs recommended by our LEPs. We will utilize multimedia, including video clips and visuals, which are accessible to a wide audience. These outreach efforts will occur at a minimum of once per quarter and as needed when exciting results emerge or other information is determined by the CEC and study teams to need dissemination. The CEC will work with the research teams for traditional academic outreach and dissemination via peer-reviewed papers, conferences, and reports to funders. A focus of our outreach plan is to work closely with our social media specialist to develop materials for social media and community-focused dissemination.

CEC Deliverables

The CEC has both process- and outcome-related deliverables. Process-related deliverables include measurable successful engagement via meeting attendance, participation, and self-reported empowerment. Co-learning, as measured by successful participation in research processes by CEC partners and successful incorporation of CEC advice and suggestions into projects, is a key deliverable. Outcome-related deliverables include inclusive consent and recruitment processes and documents that lead to optimal recruitment for each study; collaborative data analysis that results in meaningful use of data; community-informed dissemination materials to increase reach; and policy and practice guidance for successful dissemination. In addition, each study has identified key needs and deliverables that need to be met by the CEC. See Table 2 for an overview of the deliverables for the CEC and for each project's engagement plan.

XSL•FO
RenderX

Table . Community Engagement Council deliverables.

Deliverable	Overall	Project			
		1	2	3	4
Measurable successful engagement	Yes	Yes	Yes	Yes	Yes
Co-learning related to research and community engagement	Yes	Yes	Yes	Yes	Yes
Optimal recruitment via inclusive consent and recruitment processes and documents	Yes	Yes	Yes	Yes	Yes
Meaningful data interpretation via collaborative data analysis	Yes	Yes	Yes	Yes	Yes
Community Engagement Council–informed dissemination materials	Yes	Yes	Yes	Yes	Yes
Successful dissemination via policy and practice guidance	Yes	Yes	Yes	Yes	Yes
Recruitment of focus group participants	No	Yes	No	Yes	No
Recruitment of survey participants	No	Yes	Yes	No	No
Toolkit or playbook development	No	Yes	Yes	Yes	No
Identification of resources for modifying behavioral health treatments	No	No	Yes	No	No
Formal review of the toolkit prototype or products	No	Yes	Yes	Yes	Yes

I-HEAL WDMC Leveraging the TBI Model System Research Network Infrastructure

TBIMS Network Research Environment

The proposed series of projects will leverage the NIDILRR and VA TBIMS research network infrastructure. The TBIMS program is a longitudinal multicenter study that examines the course of recovery and outcomes following inpatient rehabilitation. This epidemiologic lifetime study, analogous to the Framingham study, has been funded for the past 30 years, with over 18,000 participants enrolled across 16 centers during inpatient rehabilitation for moderate to severe TBI. An elaborate administrative infrastructure developed by the TBIMS investigators supports collaboration via committees (ie, planning, research, knowledge translation, and data) and special interest groups (eg, disorder of consciousness and behavioral health). In addition, multicenter collaborative studies (module studies) are conducted in each funding cycle as part of infrastructure grants. In 2008, the NIDILRR and VA developed an interagency agreement to present the same program as the TBIMS at specialized rehabilitation hospitals within VA. Five polytrauma rehabilitation centers (PRCs) represent the rehabilitation arm of DOD trauma care. Owing to a memorandum of understanding between VA and the DOD, TBI rehabilitation care occurs at these 5 PRCs (Tampa, Richmond, Minneapolis, San Antonio, and Palo Alto). These 5 additional sites have been funded, increasing the TBIMS network to 21 centers, with over 19,000 civilians, veterans, and active-duty service members enrolled in TBIMS research. Most study PIs and co-investigators work within this infrastructure.

TBIMS National Data and Statistical Center

In addition to the network sites, the NIDILRR and VA fund the National Data and Statistical Center (NDSC) in each funding cycle. The research program at Craig Hospital in Denver, CO has successfully competed and won the NDSC award for the past 5 funding cycles (20 years) under senior leadership, including Drs Dave Mellick (director) and Jessica Ketchum (senior biostatistician). Craig Hospital has served as the data coordinating center for several extramural and intramural projects leveraging the TBIMS network. Moreover, it has served as the data management center for 36 module (multicenter) studies within the TBIMS network. Furthermore, it has served

XSL•FO
RenderX

as the data coordinating center for 10 other multicenter studies funded by the DOD, VA, PCORI, NIDILRR, and NIH. The data management infrastructure and experience of investigators at Craig Hospital will be leveraged for this study.

I-HEAL Structure

The staff will plan, develop, manage, and secure the implementation of data systems and conduct all statistical analyses for the successful completion of relevant individual studies. The core will (1) work closely with the individual study PIs to ensure successful execution of the protocol and follow-up of all participants enrolled in individual quantitative data collection efforts across projects through the use of multiple communication strategies; (2) obtain, manage, and analyze data from relevant individual projects; (3) ensure adherence to the research protocol by participants and research staff through in-person and virtual data collector trainings alongside the ISC, regular teleconferences, and constant data validation to monitor the quality of data collection, record keeping, and documentation, and the accuracy of data entry; (4) provide statistical support, expertise, and oversight for the duration of the proposed studies; (5) coordinate communication between individual projects and cores as outlined in the implementation plan; (6) provide administrative and logistical support for the study; and (7) organize, maintain, and submit study data and deidentified study resources to a study repository upon the completion of the study.

Development of the Online Study Dashboard and Toolshed

We intend to create a web presence as a landing space for public and private dissemination of information [23]. The study website will allow for centralized storage of standard operating procedures (SOPs), manuals of operations, IRB documents, form versions, and contact information for staff across all study sites, obviating the need to solely rely on email to deliver information to project members. The study website will also allow all study personnel to securely access up-to-date dynamic study reports for information on the status of enrollment, data errors, and missing data. Automatic email reminders will be sent to all relevant personnel when new participants consent to the study, which will prompt them to schedule focus groups or key informant interviews. Moreover, personnel will receive notifications on the completion of data collection and other data quality updates.

Data Management and Biostatistical Innovations

The NDSC is funded until September 2026 (fifth 5-year funding cycle). The NDSC has had many accomplishments and innovations, including moving the TBIMS National Database to a web-based data management system; providing a web-based database and analytic support for TBIMS Collaborative Module studies; introducing new and innovative statistical techniques for analyzing the TBIMS National Database; developing a structured interview for TBIMS research follow-up; conducting intense quality support visits to individual TBIMS sites; achieving several data collection certifications, including certification for this study's primary outcome; creating a dashboard of dynamic annual data reports; generating several data quality reports; implementing hundreds of computerized data checks; and maintaining an extensive SOP manual. Many of these technical and data management innovations will be incorporated into the procedures of this proposed study.

The WDMC will follow the best reporting standards to effectively display quantitative information visually. Based on the teaching of Tufte [24] and using the statistical package R and Shiny Server, we will produce a dynamic array of data and process reports that can best be understood by the data consumer.

The TBIMS environment is already functioning and is based on the best practices for a public web application, in which the web server sits outside the network in a demilitarized zone on a virtualized server. There will be only one connection between the web server and the database that flows through the firewall, which acts as a logical filter of traffic between the networks. In between these connections is an F5 web application firewall, which provides threat detection. Furthermore, the framework is 100% data-driven, meaning that the data entry forms are derived from information in the data dictionary.

Because the user interface is completely controlled by data, no code needs to be written to roll out a new data collection form, enabling a speedier and more efficient development-to-production time frame. These systems utilize a web front end (Linux web servers using ASP.NET Core, Angular, and web API) and a Microsoft SQL Server, which is encrypted using a FIPS 140.2 algorithm (the necessary level for government databases) and is consistent with the Federal Information Security Modernization Act legislation. Transmission of the data will also be encrypted using Transport Layer Security (TLS). Our technology solution is built on the security model of the CIA triad (confidentiality, integrity, and availability) using web technologies as a front-end system (where data are entered) and a SQL server as a back-end system (where data are stored).

Confidentiality

Confidentiality of the participants will be ensured through multiple measures. First, the database records will contain limited personal identifiers for use in contacting participants (eg, phone number). The web data system is set up in such a way that only authorized individuals using token protocols from ASP.NET 4.0 can access the study website. Additionally, each end user will have specific privileges to read/add/delete the data that are authorized by that center's PI. Staff at the NDSC will have limited access to the data through the use of SQL server roles and active directory group membership. The staff will follow the method of least privilege permissions, in which all users have the minimum amount of system access that still allows them to perform their job functions. Any hardcopy data forms will be stored at the individual study sites and will follow IRB confidentiality protocols.

Representatives of the study sponsor are eligible to review study records, and this will be addressed with potential study participants during the enrollment and informed consent process. No sensitive information that is required to be reported to state or local authorities will be captured as part of this study. We

XSL•FO
RenderX

will leverage the Craig Hospital Information Technology disaster recovery model, in which there are redundancies throughout the system, such as automatic and encrypted backups and co-location of servers in separate areas of the Denver metro area.

Implementation Plan

Communication Strategy

Information on the communication strategy is provided in Multimedia Appendix 6. The management of operations for this proposed study requires close coordination of multiple functions carried out by numerous staff with distinctive qualifications. Forming executive and steering committees is essential to the management process. The executive committee (study MPIs and the overall project manager [PM]; see Multimedia Appendix 7 for project leadership) will oversee overall progress, communication with the funding agency, regulatory issues, and fiscal issues. The steering committee will be comprised of the study MPIs, PIs from all individual studies, core directors, and the centralized PM, who will provide support to both committees. An all-hands call will be established to include all study leaders (MPIs, cores, and study PIs) with all core and co-investigator personnel across the FPA. Individual projects and cores have their own meeting schedule and purpose. In-person meetings will be held in years 1 and 4 to provide an in-depth orientation and review of each project's focus and progress to optimize CEC member input. The WDMC will establish an online secure website and dashboard that will alert the study team to consented participants for scheduling and follow-up.

Milestone Tracking

To ensure that the project will begin without delay, components of the proposal are included in the implementation plan as best practices/innovations by the NIH. Gantt charts for the overall FPA and for individual projects are included in the statement of work (Multimedia Appendix 2).

Regarding the utilization of project management software for milestone tracking for the overall program and individual studies, each core and individual study will utilize software, such as Click-Up or Microsoft Project, as a project management system. The PM will enter goals and tasks based on study milestone tasks (and due dates) and assign responsibility as appropriate. The project will be used for planning, management, tracking meeting agendas and notes, monitoring progress on a regular basis, and supporting other continuous QI activities. This system will be used to ensure that study milestones are met on time and to implement the evaluation plan. Additionally, the project will be linked to Microsoft Outlook and Microsoft Teams, which are available across all study sites for communication of study progress and upcoming events. The Microsoft Teams platform and dashboard will be used to conduct study meetings (to record content for later accessibility) and make available files, documents, and notes during all study-related meetings requiring the use of full-team review features. Decisions made and follow-up actions will be documented and archived for access by all team members. The PM will generate monthly reports to ensure the quality and timeliness of each task and identify tasks (at the PI, site PI, or core level) that require additional attention or staff adjustments. See the statement of work for detailed milestones and work activities across cores and individual projects (Multimedia Appendix 2).

Study Operating Procedures for Data Management and Quality Best Practices Across Projects

Over its 20-year history, the TBIMS NDSC has developed 31 SOPs for all aspects of study conduct (eg, identification of participants with TBI, guidelines for medical record abstraction, strategies for maximizing follow-up, handling unexpected events at follow-up, prevention and investigation of falsification of data, submission processes, performance target monitoring, branding, and authorship policies). Data quality and representativeness are monitored via 15 on-demand reports (eg, enrollment, follow-up, missing data, and underrepresented population recruitment and follow-up). All project PIs and the contact PI RN-R are currently funded TBIMS investigators and are familiar with these processes, which will be adapted for the proposed FPA. The data management infrastructure and SOPs developed at Craig Hospital will be leveraged for this study.

Monitoring of Performance Metrics for the Overall FPA and Individual Projects

During routinely scheduled meetings, as outlined in our communication strategy (Multimedia Appendix 6), updates regarding fiscal and logistical issues will be addressed. Project leadership teams will meet weekly during study initiation and transition to monthly meetings when study operations are well underway. Increased frequency will be added as needed. During these routinely scheduled meetings, the MPIs will resolve fiscal and logistical issues in a timely manner, including plans to proactively evaluate and prioritize study risks and issue corrective responses using TBIMS best practices. During these meetings, progress toward study milestones will be reviewed with the PMs to identify potential risks and issue corrective responses. On-demand reports are available to proactively identify risks. Study-specific reports are generated "on demand" regarding milestone tracking, including screening, enrollment, and the degree of data collection completion. Additionally, the DCC will generate monthly reports regarding follow-up rates, data missingness, and adverse event monitoring, if relevant. These reports will be reviewed during the executive committee meeting, with follow-up during the steering committee meeting. Project teams will also review reports weekly or as generated by the WDMC to proactively address potential risks. This mechanism of feedback has been highly successful in past trials completed in the TBIMS network with regard to the completion of data collection on schedule and the achievement of all study milestones well within the funding period.

Remediation Plans

For underperforming sites, TBIMS SOPs describe that remediation plans should be developed by site PIs to encourage review of local practices by study staff and individually developed plans in order to reach expected benchmarks for enrollment, follow-up data collection, and overall data quality. Remediation plans are expected for sites missing any of the

XSL•FO
RenderX

benchmarks (eg, enrollment, follow-up, and missing data), using TBIMS data quality standards. The steering committee will review remediation plans for approval. Once executed, individual project performance will be monitored monthly for a quarter to determine improvement. If no improvement is observed, the study teams will engage one of the study cores (CEC, WDMC, or ISC) for additional engagement to overcome project challenges. Sites are accustomed to these standards and have had success in overcoming missed benchmarks with a remediation plan policy. RN-R has used this method across several trials, with success in achieving study milestones.

Dissemination of Findings

The ISC, CEC, and WDMC will collaborate throughout the funding period to facilitate communication of findings for various stakeholder communities. Consistent with our history, the overall FPA scientific products will be submitted for publication in leading scientific journals in sleep medicine (eg, *Chest*, *Sleep*, and *Journal of Clinical Sleep Medicine*) and rehabilitation (*Journal of Head Trauma Rehabilitation*, *Journal of Neurotrauma*, and *Archives of Physical Medicine and Rehabilitation*). Utilizing the infrastructure of the TBIMS Study Network, which includes a separately funded Knowledge Translation Center with an international web presence, the study team will propose and develop fact sheets on the Model Systems Knowledge Translation Center (MSKTC) website. These fact sheets are accessed by patients, caregivers, and clinicians from around the world for consumer-friendly information about TBI. Dissemination success is evidenced by the number of downloads annually. The broad reach of the MSKTC is evident by the 21,967 external web pages linked to the MSKTC site from over 4000 unique domains, with 39% in the area of TBI. From January 2022 through October 31, 2022, the MSKTC website was accessed 1,825,439 times (752,047 for TBI-only content). Among those accessing the MSKTC website, 332,065 were from the United States, with the remaining from 48 other countries, highlighting the international success of the MSKTC as a dissemination venue. Many of the products are now available in Spanish and formatted in HTML to be cell-phone accessible. We will propose updates to existing relevant fact sheets and new content in year 4. Finally, standard dissemination activities include presentations at the DOD's annual professional conference, Military Health System Research Symposium, and other venues as indicated by the steering committee.

Strategic Partnership to Rapidly Translate Findings

The strategic selection of partners in the CEC will facilitate the rapid translation of knowledge products developed in the proposal. We have partnered with end users of the products to maximize usability and implementation in real-world VA, DOD, and civilian health care settings. We have partnered with professional organizations that provide guidance to end users in the field to maximize adoption by those professional communities. These include rehabilitation professionals and those in physiatry, social work, rehabilitation psychology, nursing, and neuropsychology. The LEP group is comprised of persons with lived experience of TBI who have navigated VA, DOD, and civilian health care settings. The LEP perspective will be incorporated into all products developed to maximize acceptability by patients with TBI and their families. The PPP group is comprised of representatives from federal agencies (VA, DOD, NIDILRR, and MSKTC) to help align products with policies and vice versa.

Study Closure, Data Sharing, and Biorepository Plans

Data from this project will be made available to the TBI research community by depositing deidentified research data into a nationally available biorepository at the end of the study. Primary data collection will be from providers in the form of qualitative data collection, and the corresponding data will be made available. Qualitative data from our prior work funded by NIDILRR will be archived at the Inter-university Consortium for Political and Social Research Biorepositories. The TBIMS NDSC, which will serve as the WDMC for this study, already has procedures in place to submit its core data to national biorepositories (via previous DOD funding and ongoing NIDILRR funding). They will work with the ISC on qualitative data submission. Funds have been set aside in year 4 for this submission.

Once the data collection phase of individual studies has ended and the datasets have been cleaned and locked for analysis, the data will be deidentified, merged, and securely stored for all study analyses. The final deidentified merged datasets will be securely sent to the study contact MPI (RN-R). The final datasets and documentation will be stripped of identifiers prior to release for sharing to any participating study sites by request once the project has been completed. The datasets will reside with the study contact MPI. Data will be provided in a timely manner to those requesting information for research purposes once a data sharing agreement is in place that meets all the data sharing requirements of all participating sites. To fully protect the human participants, we will evaluate each data request to ensure that special circumstances do not exist that would permit anyone to deduce the identity of individuals from the data. In any such cases, we will share the data based on an agreement stating that (1) the data will be used solely for research, and no individuals will be identified in any manner; (2) the data will be secured by electronic safeguards; and (3) once data analysis is complete, the data will be returned or destroyed.

Individual Project 1: Health Care Disparity Risk Reduction via an Innovative Medical Records Flag for Post-TBI Cognitive Impairment: Cognitive Nudge Decision Support Tool

Problem and Solution

Problem

TBI is a leading cause of long-term disability [25], affecting over 2.8 million individuals annually in the United States [26] and over 458,000 military personnel since the year 2000 [27]. TBI has been described as an *invisible disability* because it causes substantial cognitive impairments [28] without overt physical evidence of injury [29]. During the chronic recovery stage, 48% - 50% of V/SMs were found to need assistance with coordinating and accessing medical services [30,31]. The need for a health care proxy to help access and coordinate health services is likely influenced by the unmet need for cognitive

XSL•FO
RenderX

assistance, as reported by 61% - 70% of V/SMs followed 1 to 5 years after TBI [30]. The cognitive functioning and supervision needs of persons with TBI are not stagnant and indeed fluctuate over time [32], which further complicates health care service delivery. Environmental factors, such as institutional policies and clinical provider attitudes, serve as barriers to addressing health care access and coordination needs [30]. Involvement of a health care proxy (often a family member) has been found to improve access to services [6] and health outcomes among persons with both cognitive impairment and TBI [33,34]. Family involvement during care delivery is espoused by specialized (typically inpatient) rehabilitation programs where neuropsychological assessment informs individually tailored care recommendations [35,36]. However, outside these settings, family or other health care proxies may be excluded from health care encounters when providers are unaware that TBI-related cognitive impairments limit the patient's ability to independently engage, retain information, and adhere to treatments. Unfortunately, 37% of caregivers of persons with TBI reported that they were excluded from health care to the detriment of their loved ones. Cognitive specialists, such as neuropsychologists, often evaluate decision-making capacity and make recommendations about the need for health care proxy involvement [37], but these tailored recommendations are embedded in lengthy reports that may not transfer to other health care settings. Consequently, other providers are unaware of these needs, placing the burden of identifying and accommodating cognitive impairment on clinicians with limited time and resources.

Solution

The proposed project will develop an innovative electronic medical record (EMR) flag (ie, "cognitive nudge" to cue providers to include health care proxies during health encounters when recommended by TBI specialists. This cognitive nudge will be developed using a CBPR approach in partnership with key partners, including health care clinicians, administrators, and persons with lived TBI experience, and it will be pilot tested in a V/SM health care setting [38].

Background

Cognitive morbidity is a high-priority challenge in V/SMs with chronic TBI. In a sample of 120 persons with chronic TBI and 78 family members, 77% reported cognitive difficulties (impaired memory, concentration, processing speed, task initiation and organization, and problem solving), and it was mentioned that there was a high-priority unmet need [39]. These study findings were mirrored in studies of V/SMs with chronic TBI followed 1 to 5 years after injury, with 59% - 61% endorsing ongoing memory difficulties and 47% - 58% endorsing ongoing problem-solving difficulties [30,31].

Cognitive morbidity is associated with worse health outcomes. Survivors of TBI are at increased risk for myriad health morbidities, including neuroendocrine, genitourinary, cardiovascular, cerebrovascular, neurological, and systemic metabolic conditions, with reduced life expectancy [40,41]. Up to 41% of persons with TBI are readmitted to the hospital within the first year after injury [3,42,43], and poorer cognitive functioning has been shown to predict rehospitalization [3,42]. Poorer cognition was associated with both health morbidity [44] and rehospitalization [3,42].

Cognitive morbidity is the leading barrier to engaging in health care for persons with TBI. Persons with cognitive impairment following TBI experience myriad barriers across multiple dimensions of health care access. In a recent qualitative study of TBI survivors and their caregivers, 27% experienced communication issues impacting health care delivery [6]. Studies of TBI survivors at 1, 2, and 5 years after TBI found that 48% - 50% needed assistance with coordinating and accessing medical services within the past year [30,31], with other studies identifying cognitive impairment as a barrier to health care access and optimal utilization [45]. Indeed, when persons with TBI have health care advocates, it is considered a facilitator to accessing needed care [46,47].

A VA physician made the following statement:

> *TBI patients unfortunately forget things, they're not as compliant because they don't remember… You have to be cognizant of that when treating TBI patients. It's the patient population. You give them five or ten things (to do), they're probably not gonna do any of them.*

Providers require mechanisms that alert them to the need for health care proxy involvement. Federal law (eg, the Health Insurance Portability and Accountability Act of 1996 [HIPAA] and the Privacy Rule), while well-intentioned to protect a care recipient's privacy, can unfortunately create an environment that functions to exclude family from health care encounters even when exclusion would be detrimental to the patient. When care providers, such as primary care physicians, are made aware of a patient's cognitive disability, they are better positioned to help ensure that the patient's rights and preferences are prioritized in clinical decision-making and care planning [48]. Indeed, the inclusion of care partners significantly improved health outcomes [49]. Fortunately, designed EMR systems facilitate a process for alerting health care providers to specific needs, patients' preferences, and clinician-recommended actions for optimizing patient care. Feasibility and clinical utility of EMR nudges (or flags) have been demonstrated in prior research (eg, risk of readmission and suicide) [50-52]. However, it is not yet known how best to implement an EMR nudge to alert health care providers of cognitively impaired persons with TBI who may benefit from a health care proxy during health care visits.

With regard to impact, in the short term, this study will create a blueprint to reduce health care inequity and improve care for cognitively impaired V/SMs with TBI, which is an "invisible" disability affecting 5% of Americans and causing life-altering challenges without overt signs of injury [53]. At the conclusion of this study, we will have developed and refined the toolkit required to implement a cognitive nudge intervention in partnership with key partners; these outputs will be disseminated broadly for immediate use. In the long term, this study has the potential to positively impact health care delivery for persons with cognitive impairments outside of TBI. As evidence-based chronic care management protocols are being developed for TBI, similar nudge interventions can facilitate the

XSL•FO
RenderX

implementation of proactive surveillance and tailored disease management interventions.

Objectives

The aims are as follows:

- Aim 1.1 (discover): We aim to engage health care providers, persons with TBI, and their families to identify content, processes, and procedures for developing a cognitive nudge clinical support tool.
  - Aim 1.1a: We aim to determine the best practices across professional communities for the identification of persons with TBI-related cognitive impairments and recommendations for the engagement of health care proxies in medical care. We will conduct online surveys (aim 1.1a1) and focus groups (aim 1.1a2) with professional groups to identify these practices.
  - Aim 1.1b: We aim to identify the tailored needs and operability of a cognitive nudge intervention from the perspective of persons with TBI and their families (LEPs). The ISC will conduct focus groups with persons having lived experience to inform the development of the cognitive nudge intervention.
- Aim 1.2 (define and develop): We aim to co-design a cognitive nudge clinical decision support tool with health care providers, persons with TBI, and their families.
  - Aim 1.2a: We aim to develop a toolkit of resources and an implementation plan for the cognitive nudge clinical decision support tool intervention. In collaboration with the ISC, we will triangulate data from aim 1.1 to inform product development.
  - Aim 1.2b: We aim to engage partners, including CEC partners and individual engagement partners (IEPs), in the formative evaluation of the cognitive nudge toolkit and implementation plan. The ISC will conduct semistructured interviews during regularly scheduled CEC meetings to obtain input on the format, content, and usability to refine the cognitive nudge.
- Aim 1.3 (validate): We aim to implement (pilot) and evaluate the outcome of the cognitive nudge at a major VA medical center and PRC.
  - Aim 1.3a: We will conduct a formative evaluation of the clinical nudge to preliminarily assess acceptability, relative advantage, usability, and complexity. Following a 4-month implementation pilot, we will conduct semistructured interviews with providers, administrators, and patient/caregiver dyads to evaluate implementation outcomes and perceptions of proxy engagement in health care encounters.
  - Aim 1.3b: We aim to disseminate the cognitive nudge toolkit and implementation plan with professional partners. We will refine the toolkit and implementation plan based on the pilot and disseminate it by leveraging professional partners from federal agencies and professional organizations.

Design

This project falls within the preimplementation phase of the QUERI Implementation Roadmap. The goal is to improve access to health care by improving overall engagement of persons with TBI in their health care encounters. Our prior work has shown that caregivers and advocacy organizations are critical to engaging in health care encounters and overcoming communication barriers reported by veterans with lived TBI experience [6]. The proposed mixed method study will identify the best practices to inform cognitive nudge development while considering LEP community preferences (aim 1: discover), develop products for supporting cognitive nudge implementation (aim 2: develop), and conduct a pilot to refine and disseminate the intervention for future implementation and evaluation (aim 3: validate and disseminate).

Aim 1.1: Discover (Study Months 7-24)

Participants

Participants for Aim 1.1a1 (Provider Cohorts)

We will partner with PPPs and IEPs to recruit a national sample of health care providers who treat persons with TBI. Each provider will be required to meet the following inclusion criteria: (1) being a licensed health care provider (eg, medical provider [physician]), a psychologist (rehabilitation or neuropsychology), or other allied health care professional, and (2) spending at least 20% of their professional time in clinical services for patients with moderate-to-severe TBI.

To collect a representative sample across provider types and health care systems, we intend to recruit a minimum of 350 survey responders. This sample size will be large enough to estimate means and proportions from the survey, with sufficiently narrow CIs. A sample size of 350 will estimate a 2-sided 95% CI on standardized means (SD=1) with a margin of error no larger than 0.101 [54] and a 2-sided 95% CI on proportions with a margin of error no larger than 0.107 [55,56].

Partnering with professional organizations specific to this study (National Academy of Neuropsychology), hospital administrators (Chief of Staff, and Chiefs of Medicine, Mental Health, and Rehabilitation), and overall CEC partners increases access to a large pool of professionals involved in the care of persons with TBI for survey purposes, maximizing the positive response rate. We plan to use the Dillman method, which has produced response rates up to 50% [57]. The study team has a successful track record of recruiting providers for surveys involving clinical practice [58].

Participants for Aim 1.1a2 (Provider Focus Groups)

We will conduct up to 5 focus groups (n=8 maximum/group) until thematic saturation is reached. Each provider will be required to meet the following inclusion criteria: (1) being a licensed health care provider (eg, medical provider [physician]) or psychologist (rehabilitation or neuropsychology), and (2) spending at least 20% of their professional time in clinical services for patients with moderate-to-severe TBI. We will partner with volunteers identified by the PPPs to recruit a diverse provider cohort. We will examine CFIR-guided constructs to determine the extent to which the cognitive nudge intervention

XSL•FO
RenderX

is viewed as suitable or likely to be used, and how it can be successfully delivered.

Convenience and snowball sampling methods will be used to identify and recruit participants across the professional organizations engaged. We plan to overenroll underrepresented groups to understand the implementation challenges across racial and ethnic groups.

The number of focus groups has been determined for achieving theme saturation. Theme saturation is defined as the lack of new codes being introduced into codebook development [59]. For determining the number of focus groups needed in qualitative exploratory research, we have referred to the methodological work by Hennink et al [59]. The authors showed that 94% of all codes and 96% of high-prevalence codes are identified by the fourth focus group, with no new codes of this type introduced thereafter. They highlighted that this approach is particularly suitable for designing interventions (the purpose of this aim) but may be limited for other purposes [59]. Therefore, we propose that up to 5 focus groups with ≤8 persons per group is more than sufficient to achieve thematic saturation.

Participants for Aim 1.1b (Lived Experience Focus Groups)

We will conduct up to 5 focus groups (n=8 maximum/group) until thematic saturation is reached. The inclusion criteria for the focus groups are as follows: (1) having moderate-to-severe TBI or being a TBI proxy care partner, and (2) having a reliable internet connection (web-based focus group participation).

We will partner with our LEPs and use TBIMS patient registries to recruit lived experience participants following IRB procedures. This approach has worked well to recruit V/SMs with lived experience and their families in past qualitative TBI research [6].

The number of focus groups has been determined for achieving theme saturation. We have referred to the methodological work by Hennink et al [59], and we propose that up to 5 focus groups with ≤8 persons per group is more than sufficient to achieve thematic saturation.

Data Collection

Data Collection Measures and Tools for Aim 1.1a1

A semistructured survey developed in collaboration with our CEC members (PEPs and LEPs) will be conducted with the goals to: (1) determine current practices that identify persons with cognitive impairments after TBI and policies that require it across health care settings; (2) determine current practices for identifying medical decision-making capacity or inability to participate in health care encounters independently; (3) identify documentation practices and locations in medical record systems for identifying previous goals (1 and 2); (4) review innovative practices for ensuring persons with TBI have health care proxy assistance in medical encounters (open-ended question); and (5) assess provider attitudes regarding satisfaction with existing practices and willingness to trial new approaches to ensure health care proxy assistance in medical encounters. Questions will be developed using both the open-ended (qualitative) approach and the Likert scale rating.

CFIR Guide for Aims 1.1a2 and 1.1b

The CFIR guide will be used to inform the implementation determinants (facilitators and barriers) of the cognitive nudge intervention to determine the extent to which the cognitive nudge intervention is viewed as suitable and likely to be used, and how it can be successfully delivered. The CFIR guide has been used internationally to study all types of health conditions in military and civilian health care settings [60,61]. The overarching access framework domains from the supply/organization side and consumer/demand side will also inform the focus group questions and have been successfully used in our own TBI research [6,46,47,58]. The focus group guide will be pilot tested with people representing those we will engage to ensure comprehension of questions and content.

Study Procedures

Procedures for Aim 1.1a1

Recruitment flyers, emails, social media posts, and telephone scripts will be developed and approved during study initiation. Project 1 team members and PPPs will hold monthly meetings during quarters 3 - 4 of year 1 to develop and field test the survey, which will be deployed (quarter 1 of year 2) to members of professional health care organizations (eg, American Congress of Rehabilitation Medicine, Association of American Physicians, National Academy of Neuropsychology, American Psychological Association's Division of Rehabilitation Psychology, VA, and DOD) via organizational email distribution groups and personal network emails of the research team and core groups. Survey links will be sent up to three times via email. Up to 20 respondents will be randomly selected to receive a US $50 gift card for participation.

Procedures for Aims 1.1a2 and 1.1b

Recruitment flyers, emails, social media posts, and telephone scripts will be developed and approved during study initiation. Up to 10 focus groups (5 focus groups per stakeholder type [professionals and persons with lived TBI experience] will be conducted, with ≤8 people for each focus group). Each focus group will be scheduled for 90 minutes, conducted on online conferencing platforms (eg, Microsoft Teams and Zoom), digitally recorded, and professionally transcribed with quality checks by the study team. Participants will receive a US $50 gift card for participation if permissible.

Analysis Plan

The project team, in consultation with partners (PPPs, LEPs, and IEPs), will analyze the collected data.

Survey Analysis for Aim 1.1a1

Survey data will be summarized descriptively using means, SDs, and percentiles for continuous data, and frequency counts and proportions for categorical data. Survey data will be summarized for the overall sample and by provider type and health care setting. Responses will be compared across provider types and health care systems using chi-square tests, ANOVA, and nonparametric alternatives as necessary (Fisher exact test and Wilcoxon rank sum test). Open-ended survey responses will be exported to an Excel file. Individual responses will be analyzed using a rapid matrix analysis approach by comparing

XSL•FO
RenderX

similar responses and grouping them to create categories [62]. For both the qualitative and quantitative data, sensitivity analyses will examine responses by demographics (age, gender, race, ethnicity, and years in practice).

Qualitative Analysis for Aims 1.1a2 and 1.1b

Focus group and individual interview transcripts will be analyzed using matrix analysis, a rapid assessment approach, to summarize and identify themes. Rapid assessment is a team-based approach to iteratively collect and analyze qualitative data that emphasizes speed of data collection and analysis in relation to focused programmatic questions or problems [62]. Two experienced qualitative researchers will conduct the matrix analysis. Together, they will create codes both deductively from known constructs and inductively from the data, and use matrices to categorize responses under domain names, applying codes that emerge from the data. Codes and preliminary themes will be reviewed by the research team and partners to establish consensus. A matrix analysis approach will facilitate a rapid turnaround of the results to provide immediate feedback to inform the cognitive nudge and supportive implementation materials, such as the implementation blueprint.

Aim 1.2: Design the Toolkit and Develop the Implementation Plan (Study Months 25-37)

Overview

Using a CBPR approach that strengthens knowledge translation efforts [63,64], we will schedule a series of meetings to review the triangulation of aim 1 findings with the CEC, identify needs informing toolkit product development (Multimedia Appendix 8), and evaluate the usability of toolkit content.

Initial Content Identification

Project 1 engagement partners represent key partners in the eventual implementation of the cognitive nudge decision support tool. They will review aim 1 findings to identify needs from their unique perspectives and make suggestions for the content, structure, and language used in products in the toolkit.

Product Development and Prioritization

Project 1 IEPs (VA hospital administration and national VA physical medicine and rehabilitation policy leads) will review product suggestions and prioritize content development. Existing national policies for the use of nudge interventions and products that exist for other clinical reminders in the EMR will be used as a starting point for all products, such that most products will be adaptations to maximize feasibility and acceptability, and to optimize integration into the existing health care system. The study team will then utilize feedback across engagement partner communities to adapt existing content and develop new content for the toolkit.

Formative Evaluation of Toolkit Products and Implementation Plan

The ISC will conduct semistructured interviews with the CEC and IEPs for user testing of toolkit products. No formal analyses will be conducted, and feedback will be used by project investigators to finalize the content.

Aim 1.3: Deliver and Validate (Study Months 38-48)

Overview

We will work with project 1 team members (hospital administration and national physical medicine and rehabilitation EMR leads) to pilot the implementation of the cognitive nudge decision support tool for final refinement prior to dissemination.

EMR Implementation

The study team will follow recommendations for implementing a category 2 patient record flag (ie, nudge) in compliance with Veterans Health Administration Directive 2010‐05 and Patient Record Software Patch v5.3 (March 2019). These directives outline policy and guidance for the proper use of patient record flags to enhance safety for patients, employees, and visitors. Category 2 flags are for nonimmediate risk that can be implemented at a facility level. Category 2 flags denote risk factors (eg, cognitive impairment) that must be known in the initial moments of a patient encounter. Facility-level partners responsible for the implementation of a category 2 flag are involved as IEPs in this proposal. Software patches in the VA EMR system permit entry, editing, and monitoring of patient flags. The team and IEPs will work with the Tampa VA Patient Safety Office, overseeing EMR flags to trial the cognitive nudge intervention.

Staff Training and Pilot Implementation

Veterans Health Administration Directive 2010‐05 requires training on any element of software utility, activation, continuation, and deactivation of the cognitive nudge. Additional areas of training include confidentiality and appropriate responses to a cognitive nudge. Training in these areas will follow the toolkit and implementation plan developed in aim 2. Training for Tampa VA TBI rehabilitation providers and primary consultants will be delivered via existing staff meetings across services. As noted, chiefs of those services serve as project I engagement partners to facilitate local training and unique needs across services. Following training, the cognitive nudge will be piloted for a 4-month period in the inpatient polytrauma rehabilitation unit (18/56 inpatient beds) at the Tampa VA, the largest rehabilitation facility in the VA Healthcare System.

Formative Evaluation of the Implementation of the Cognitive Nudge Clinical Decision Support Tool

At the conclusion of the pilot, ISC staff will conduct semistructured interviews with providers who interacted with the cognitive nudge during the pilot (Medical Director of the brain injury inpatient program and 2 unit neuropsychologists) using key informant interviews. The ISC will also conduct focus groups with allied health providers (case management providers, rehabilitation therapists, nurses, and family therapists). Key informant interviews and focus groups will use the CFIR Interview Guide to identify facilitators and barriers to implementation. Codebook development will include key implementation constructs of acceptability, relative advantage, usability, and complexity. Data will inform final revisions to the toolkit and the implementation plan tailored to rehabilitation settings in the VA/DOD system. Persons with TBI receiving

XSL•FO
RenderX

care during the pilot and their health care proxies will also be approached and consented to complete a semistructured interview with the ISC to understand whether the cognitive flag nudge intervention enhanced health care proxy engagement in the health care encounters during the pilot. Findings will be integrated into the finalization of the toolkit.

Potential Challenges and Solutions

A potential problem is insufficient interest in surveys or focus groups. If we experience difficulty recruiting representative samples, we will inform individual members of the CEC to personally invite their respective networks to participate in surveys and/or focus groups. We will use feedback from participants to optimize the research participation experience and reduce deterrents to participation as applicable. Using methods proposed herein, our team recently recruited over 350 professionals to complete an internet survey in a 4-month time span. By building on our past success, we will minimize recruitment barriers.

Another potential problem is low acceptability of the cognitive nudge among persons with lived experience. We have considered the potential disadvantages of an EMR flag alerting providers to a patient's cognitive challenges, and we will work with our partners to determine the best practices for deferring and/or removing the flag for any reason, including patient/family preferences. A better understanding of what aspects of a nudge intervention may be less attractive to those with lived experience is essential for designing the toolkit proposed herein. Based on extensive data from our team and others, a plurality of TBI survivors living with cognitive impairments report that these impairments pose barriers to care, and this study will determine how best to mitigate those barriers through co-development of a decision support tool in partnership with partners.

Community-Based Participatory Approach

This study embodies the CBPR approach by engaging end users of the cognitive flag intervention as full partners in the conceptual development, development, and dissemination of products developed across study aims.

Individual Project 2: Provider Toolkit for Accommodating Cognitively Impaired Persons in Evidence-Based Treatments

Problem and Solution

Problem

Proposal investigators have robustly shown that cognitive morbidity (ie, cognitive impairment) is the number one rehabilitation need in the chronic stage of TBI [7,30,31], with unmet needs associated with poorer satisfaction with life among V/SMs [7]. Further, V/SMs with greater TBI disability (and greater cognitive morbidity) experience unique barriers to accessing care in the chronic stage of TBI [6]. Our most recent work characterizing health care delivery for the most common comorbid conditions [8,65] in TBI highlights cognitive morbidity [46,47] as a barrier in referral to and receipt of evidence-based behavioral treatments (EBTs), limiting health care access and meaningful engagement in the management of TBI comorbidities that may improve outcomes (ie, suicide risk, PTSD severity, arrest rates, and worsening cognition) [66-70].

Solution

The proposed project will address determinants (ie, facilitators and barriers) that limit a cognitively impaired person's ability to engage in EBTs for the most common comorbid conditions (ie, PTSD, depression, sleep disorders, and chronic pain) in TBI by conducting an environmental scan to identify existing EBT adaptations for persons with cognitive impairments (aim 1), selecting adaptations identified through a CBPR approach (aim 2), and developing and disseminating a provider toolkit for accommodating cognitively impaired persons, which will aid the identification and engagement of cognitively impaired persons in EBTs (aim 3).

Background

Psychological health conditions are common in the chronic stage of TBI. Our group has found high rates of sleep difficulties [71-73], chronic pain [8,74], PTSD [67], and depression [75] in individuals with TBI from months to many years after injury. As high as 46% of those with moderate to severe TBI report ongoing chronic pain from 1 to 30 years after injury. These conditions often co-occur, resulting in constellations of disorders, with signature injuries of military service and deployment making up the polytrauma clinical triad (ie, TBI, chronic pain, and PTSD).

Post-TBI psychological health conditions are associated with worse recovery and an increased risk of suicide. Chronic pain, depression, and PTSD are associated with greater disability and reduced quality of life following TBI [76-79]. Unaddressed, these conditions may exacerbate one another. For instance, the presence of PTSD contributes to avoidance of activity, exacerbating functional limitations [80,81] and resulting in poorer adjustment to ongoing pain [82]. Alone or in combination, the presence of these psychological health conditions is associated with an increased risk of suicide [83,84], making successful engagement in evidence-based treatment in this population an urgent need.

A VA neuropsychologist made the following statement:

> *It is challenging to get comorbidities treated by providers who understand cognitive disability that come along with brain injury…finding a therapist who can take into consideration the cognitive limitations is hard to find.*

Cognitive difficulties have been found to limit referral to and receipt of evidence-based treatments for behavioral health needs in persons with TBI [47]. In our recent work, 67% of providers described cognitive impairment as the number one barrier to receiving evidence-based health care [47]. Provider expectations of limited treatment benefit due to cognitive deficits limit referrals to evidence-based care, which can be adapted to overcome cognitive impairments. In 2022, the American Psychological Association published a new guideline mandating disability accommodations [85]; however, a major gap was the lack of specificity for what accommodations would be enacted and how to integrate them into behavioral health treatments. In the peer-reviewed literature, existing strategies for adapted

XSL•FO
RenderX

behavioral health treatments for disabilities, such as cognitive impairments, are inconsistent and lack specific guidance for providers [86-88] delivering EBTs. To address this gap, this project proposes to develop a toolkit of specific accommodation strategies, reviews of the literature, and specific guidelines for managing cognitive limitations in delivering evidence-based behavioral health care. The toolkit will contain practical strategies consolidated from practice-based surveys, evidence synthesis, and validation by clinicians and persons with lived experience of TBI.

Project investigators [89] have shown that adaptations for TBI morbidity are effective in the delivery of evidence-based, behavioral health treatments for veterans with TBI. The MPI JNH has adapted cognitive behavioral therapy for chronic pain in veterans with TBI and shown improvements in pain interference, a key outcome in this population [89]. Other project investigators have used these adaptations in collaborative care models and the delivery of motivational interviewing to promote the use of continuous positive airway pressure among veterans with TBI and cognitive morbidity. Further, frontline behavioral treatments for PTSD (cognitive processing therapy and prolonged exposure therapy) [90-93] and insomnia [94,95] have also been adapted for persons with TBI and shown to have efficacy. Investigator-described adaptations include the use of concrete examples (avoid abstract constructs), the use of visual aids, and caregiver engagement in sessions. Not all treatments manualized the adaptations but included abbreviated sessions, use of reminder calls, and integration of memory aids [90-93]. To facilitate the utilization of cognitive adaptations of behavioral health treatments for persons with TBI, we propose to use HCD to gather existing resources for adapting treatments for cognitive difficulties, curate those resources to identify the most up-to-date and relevant resources with the support of expert provider partners, develop a toolkit, and conduct a formative review to determine readiness for dissemination. Partnering with the CEC, we will create awareness of the toolkit of adaptation strategies to ultimately increase access to evidence-based treatments for common TBI comorbid conditions known to worsen outcomes.

Objectives

The aims of project 2 are presented in Figure 4. The aims are as follows:

- Aim 2.1 (discover): We aim to conduct an environmental scan of existing recommendations and EBT adaptations for persons with cognitive impairments.
  - Aim 2.1a: We aim to conduct an environmental scan to identify existing resources that fall into three broad categories: (1) tools for assessing cognitive difficulties; (2) general information on adaptations for EBTs, including how to include caregivers/family members when needed; and (3) manuals that include adaptations of existing evidence-based treatments.
  - Aim 2.1b: We aim to summarize the environmental scan results to develop a product grid.
- Aim 2.2 (define and develop): We aim to engage the CEC in prototyping the provider toolkit for accommodating cognitively impaired persons by defining its scope, elements, and function.
  - Aim 2.2a: We aim to convene a SME review panel and include LEPs to (1) select which high-quality materials from the product grid should be included in the toolkit and (2) brainstorm additional toolkit content, layouts, and functions.
  - Aim 2.2b: We aim to synthesize data from aim 2.2a and develop a prototype of the toolkit.
- Aim 2.3 (validate): We aim to activate the CEC to conduct a formal review of the provider toolkit prototype for accommodating cognitively impaired persons to validate its design and identify implementation strategies.
  - Aim 2.3a: We aim to conduct a formal review of the toolkit prototype to determine key strategies for dissemination to practitioners, educational institutions, health care systems, and other identified groups to increase the utilization of the recommendations.
  - Aim 2.3b: We aim to synthesize formal review data and make data-driven revisions to finalize the toolkit.

XSL•FO
RenderX

Figure 4. Project 2 aims. SME: subject matter expert.

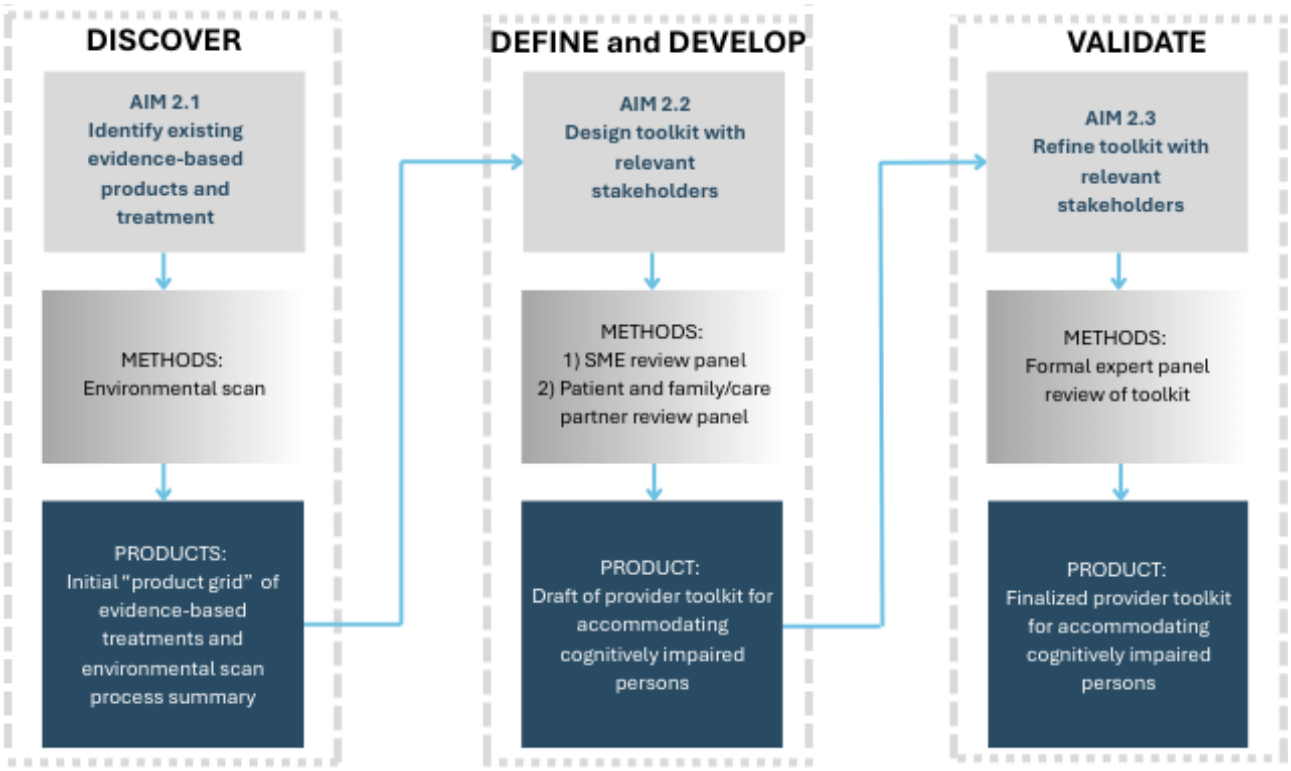

Design

The proposal aligns with the preimplementation phase of the QUERI Roadmap through the development of knowledge products (toolkit) from published literature and other sources. An HCD process will guide the design and implementation of the toolkit. HCD is an approach that "centers" end users (ie, the people who will ultimately use or benefit from a product or service) by integrating them into every aspect of the design process [14,15,17], thus strengthening the impact and utilization of the product or service. HCD can combine familiar qualitative and quantitative data collection methods with novel efforts to engage beneficiaries in an iterative design process. Each of our project aims aligns with the 4-phase Double Diamond HCD model [16]. This model directs researchers and designers to engage end users to first *discover* and *define* a problem and then take focused action to *develop* and *deliver* (ie, *validate*) a solution [16]. HCD supports our process of making high-quality evidence-based tools that are as effortless as possible to find, easy to understand, and esthetically pleasing.

Aim 2.1: Environmental Scan and Product Grid Development (Study Months 1-18)

Initial Toolkit Content Identification (Aim 2.1a)

Following project approval as a QI project, the team will commence a comprehensive environmental scan using a CBPR approach. First, the team will review project goals with CEC members to facilitate identification of provider resources for modifying behavioral health treatments for persons with cognitive impairment. Professional partners will be asked to help identify existing online materials and utilize their professional contacts for finding additional materials and tools. Next, an anonymous Qualtrics survey that queries providers about resources used to provide adaptations for persons with cognitive impairment will be developed. The survey will include a request to submit content to the project (if permitted) to complete the environmental scan. The PPPs within the CEC will be asked to help disseminate the survey to their constituency using professional and email listserv platforms available within their organizations. Social media content will be developed by our social media specialist and shared to help increase the reach of the survey. The WDMC will develop an online platform to store and organize relevant materials about adaptations.

Product Grid Compilation (Aim 2.1b)

Research staff guided by MPIs (AMM, JMH, and Daniel Goldschmidt) will work with the ISC to conduct the comprehensive environmental scan process by (1) compiling collected materials from SMEs on a secure drive, (2) searching extensively via compiled sources (eg, organization websites) and web searches for additional leads on tools, and (3) conducting a literature review with search terms to capture adaptations for cognitive deficits for behavioral health treatments (eg, chronic pain, PTSD, and insomnia). A study research associate will begin this process in quarter 1 of year 1, as they will already be onboarded at the start of the project. Materials identified through the environmental scan will be entered into an Excel document using a product grid template. The product grid, developed by the MPI JNH, uses rows and columns to review, organize, and rate materials (1 product per row) by preestablished qualities or domains (column descriptors). Research staff will review each resource and preliminarily complete a review of products in the product grid. SMEs will serve as IEPs for project 2 to aid in product grid validation. IEPs represent content expertise across commonly comorbid conditions in military personnel after TBI (Multimedia Appendix 9). IEPs will partner with the project team to validate adaptations

XSL•FO
RenderX

proposed and their impacts on specific evidence-based treatment content.

Aim 2.2: Expert Panel Product Grid Review (Study Months 19-36)

Review and Rating of Products by SMEs (Aim 2.2a)

Once the product grid is populated based on a preliminary review by the study team, a secondary SME review will be conducted. The product grid will include general accommodations for cognitive difficulties and be organized by psychological health condition (ie, PTSD, depression, chronic pain, and sleep disorders). SME review panel members will be assigned products corresponding to their primary area of expertise to review. All panel members will review general accommodations. Each line item in the product grid will be evaluated by 2 SMEs, and for any product where there is a discrepancy in ratings, a third SME will evaluate to help reach consensus. Products will be assigned to and independently reviewed by the SME review panel along the following domains: (1) verification of the intended audience (eg, licensed therapist), (2) quality of accommodations made (rating appropriateness from 0 [poor] to 5 [excellent]), and (3) complexity of using the tool (0 [very difficult] to 2 [very easy to implement]).

Convening of the SME Review Panel

Once the SMEs complete their individual reviews of the product grid, the SME review panel will convene to review the products identified as being of high quality, which can be implemented with the greatest ease. Prioritized products will be translated into tools or components of the toolkit. SMEs will collaborate to (1) identify gaps in tools that may require development and (2) provide guidance on how to create tools that need to be developed.

LEP Engagement

In tandem with the SME review, the LEPs will convene to review the identified products to be included in the toolkit to inform appropriateness for the target patient population. Additionally, the LEPs will discuss the appropriateness of language use and orientation for providers when working with persons with TBI. The LEPs will be specifically asked about the appropriateness of language regarding the introduction of tools and the overall approach to guiding providers/end users of the toolkit on integrating modifications. Gaps in the toolkit will be identified through discussions with LEPs based on lived experiences of engaging with behavioral health providers and treatments. For example, feedback about the importance of having a toolkit like this or the frustrations experienced when interacting with behavioral health providers as a person with TBI will be invaluable to the project team for creating content in the “introduction” and “purpose of the toolkit” sections.

Toolkit Compilation, Additional Content Creation, and Design (Aim 2.2b)

The project team will triangulate feedback from aim 2.1a to develop the initial prototype of the toolkit, tentatively titled “The Provider Toolkit for Accommodating Cognitively Impaired Persons.” The team will partner with the ISC product designer to compile tools in the product grid and develop additional tools or toolkit components (see Table 3 for possible components of the toolkit). Details outlined in the product grid categories will inform background narratives, applications, and other framing aspects for each tool.

Table . Sample toolkit content.

Toolkit section	Information
Executive summary	Overview and purpose of the toolkit (gap addressed, applications, and target audience)
Table of contents	Identified sections
Tip sheets	Abbreviated guidance for implementation, including engagement of health care proxies in the treatment process, summarized in an algorithmic approach
Quick references	Source materials, including literature reviews, existing treatment manuals, clinical practice guidelines, and position statements, for relevant topics
Screening resources	Resources for identifying those in need of cognitive adaptations in treatment processes
Tools: introduction and framing of each tool	Target population for each tool, brief background of creation, and utility for specific treatments and disorders, including summary infographics to facilitate understanding
Patient and family materials	Tip sheet explaining behavioral health approaches for persons with cognitive morbidity
Training	PowerPoints, fact sheets, videos, etc
Resources	URLs and references
Toolkit development background	Literature review of the topic (to enhance credibility of the toolkit)
	Project leadership and history
How to share the toolkit	Instructions on how to share toolkit content

XSL•FO
RenderX

Aim 2.3: Engagement of SMEs (Study Months 37-48)

Activate the CEC to Conduct a Formal Review of the Toolkit Prototype (Aim 2.3a)

A formative review will be conducted to review, provide overall impressions, and make recommendations to improve the toolkit [96]. The toolkit prototype completed in aim 2.2b will be circulated several weeks prior to review sessions. Review sessions will take place online, and a facilitator and notetaker will be present. The team will engage the PPPs of the CEC to address organizational insight and high-order messaging, framing, and usability of the tools, as well as identify targets for future dissemination within organizations, educational institutions, and health care systems. The provider SME review panel will assist in validating the toolkit in terms of content validity, treatment fidelity, and usability of the tools across and within EBT approaches. The team will engage the LEPs of the CEC to provide final validation of the appropriateness and sensitivity of the language to guide providers/end users on integrating modifications in a way acceptable to patients and families. Any items still needing resolution will also be reviewed.

Data will be analyzed using the rapid assessment process [62], a team-based approach, which allows for quick turnaround of qualitative data analysis and provides prioritized findings. Review session notes alongside spot checks of audio recordings to clarify any missed notes or important reviewer quotes will be processed and cleaned with quick turnaround (within 3 days of data collection). This data analysis approach uses triangulation and iterative data analysis to quickly generate insights from qualitative data, and intensive team interaction to analyze the data will help to quickly translate reviewer feedback into prioritized modifications for the team to address in aim 2.3b [97].

Synthesizing Formative Review for Toolkit Finalization (Aim 2.3b)

The study team will integrate findings from the formative review to finalize toolkit compilation. The team will work with the social media specialist to create awareness of the project, and milestone completion will be highlighted on social media. The team will work with the WDMC to highlight project milestone completion on the overall FPA online dashboard. The study team will publish data collected across study aims in rehabilitation journals specializing in psychological health and as advised by the PPP members (eg, *Rehabilitation Psychology*, *Archives of Physical Medicine and Rehabilitation*, and *Provider Informational Education Page*). The toolkit and implementation plan (playbook) will be published on the study Toolshed Site hosted by the WDMC. The team will work with professional engagement partners and individual professional engagement partners to create a communication plan for the toolkit with various professional organizations, including those issuing clinical guidance to the field. Future work will evaluate the dissemination and implementation of the toolkit, and the outcome of toolkit use to benefit V/SMs with TBI.

Potential Challenges and Solutions

A potential challenge is the absence of respondents for the survey of the environmental scan. Team members have had success recruiting from professional organizations using informal methods. Formal engagement and buy-in from these organizations from the start of the project will help facilitate commitment to engaging their respective constituencies for input.

Another potential challenge involves the scaling of the project and its timely completion. The project team will prioritize content for high-frequency conditions that are often comorbid with TBI and are known to be associated with poor outcomes.

Community-Based Participatory Approach

The proposed project will use a CBPR approach to engage partners across all aims. The proposed aims will engage SMEs and LEPs in aim 1 to review the product grid, aim 2 to develop toolkit content, and aim 3 to conduct a formative review of the finished products in the toolkit. Collectively, our CBPR approach will ensure that the toolkit and its contents are relevant, useful, and appropriate for the target end users.

Individual Project 3: Empowering Teams to Implement Evidence-Based Behavioral Interventions

Problem and Solution

Problem

Challenging behaviors after moderate-to-severe TBI (eg, agitation and impulsivity) are common (44% - 74%), frequently impacting access and quality of health care [98]. Our preliminary data indicate that these challenging behaviors are associated with denial of access to health care settings for persons with TBI, as well as increased rehabilitation staff injuries and turnover [58,99]. Possibly due to perceived lack of alternative options, one-third of providers worldwide have reported treating challenging behaviors with sedating medications thought to impede adaptive behavior, cognition, and neurorecovery [100,101]. While effective methods for optimizing adaptive behavior while reducing challenging behaviors exist, widespread implementation of evidence-based, nonpharmacological treatment programs has not yet occurred.

For the question “What are the challenges to have effective programs that manage maladaptive behaviors?” an inpatient rehabilitation provider made the following statement:

> *Limited knowledge of the nursing staff regarding TBI, anticipated behavioral issues, and appropriate management strategies.*

Solution

We will develop and pilot a playbook for implementing an interdisciplinary program in TBI inpatient rehabilitation settings to promote adaptive behavior and reduce the challenging behavior of patients with TBI. The Staff Training in Assisted Living Residences-VA (STAR-VA) [102] will be adapted to produce the new program, TeamBI (Team-based Behavioral Interventions). This program will incorporate the principles of applied behavior analysis and positive behavior support [103,104] to prepare interdisciplinary teams in inpatient

XSL•FO
RenderX

rehabilitation settings for setting up the environment to optimize adaptive, positive behaviors and minimize challenging behaviors. The playbook will also provide implementation strategies to assist with establishing the program.

Background

Maladaptive behaviors are prevalent following brain injury. Behavioral disturbances after brain injury occur frequently and often involve the presence of socially unwanted behaviors (eg, agitation, impulsivity, and apathy). The frequency of these behaviors ranges from 44% to 74%, depending on injury severity and the time since injury [98]. A systematic review by Stéfan et al [98] in 2016 examined the prevalence of behavioral disorders following TBI. The most common behavioral disorders following brain injury include agitation (11% - 70%), aggression (25% - 39%), irritability (29% - 71%), and apathy (20% - 71%). Other common behavioral disorders seen following brain injury include impulsivity, confusion, disinhibition, fixation, confabulation, restlessness, sexual inappropriateness, and anosognosia (poor insight/awareness) [105-107].

Maladaptive behaviors negatively impact patient access to health care and long-term outcomes. Brain injury–related challenging behaviors have consistently been associated with disparities in health care access and poor long-term outcomes [58,108-110]. Behavioral disturbances often negatively impact an individual's ability to fully engage in rehabilitation care, thereby limiting the scope and benefit of rehabilitation [111]. Further, behavioral sequelae can compromise discharge placement as many facilities and families are unable to manage challenging behaviors due to safety concerns and the need for extensive direct supervision [100,112,113]. Behavioral disturbances following brain injury significantly limit and may even prevent successful community reintegration, including returning to work and/or social environments [109,114].

Effective interventions exist for reducing challenging behaviors but are not consistently implemented. Clinical guidelines support the use of patient- and family-centered, integrated approaches to manage behaviors that interfere with participation in rehabilitation and with the successful performance of daily functions. Recommended interventions are based on two related schools of thought: "applied behavior analysis" and "positive behavior intervention and support." Systematic reviews indicate that these approaches can be highly effective across a range of populations, including persons with brain injury [103,115-118]. The personalized nature of the interventions limits studies to single-case experimental designs that have inherent methodological challenges, especially when implemented in clinical settings. However, decades of research and clinical applications clearly support their use. The "ABC" model is fundamental to applied behavior analysis, as it examines and identifies the antecedents (what was occurring before the target behavior), behavior (accurately identifying the behavior to be modified), and consequences (what was the result of the behavior). Utilization of this model aims to identify behavior patterns and factors that may be eliciting and reinforcing unwanted behaviors. The positive behavior intervention and support model focuses on antecedents and on the notion that, through collaboration between staff, patients, and family members, the environment can be restructured to promote adaptive behaviors and, by extension, reduce challenging behaviors.

Unfortunately, the consistent application of behavioral interventions, even in units designated to treat persons with brain injury and challenging behaviors, is low [119]. Implementation of behavioral interventions requires staff with behavioral training, a cohesive team approach, and extensive knowledge about brain injury sequelae. In an international survey, only half of clinicians felt that they had adequate training to implement behavioral interventions [100]. In environments where staff rotate, training opportunities and time are limited, and behavioral expertise may not be available, more restrictive measures to manage unwanted behaviors may be used, such as sedating medications [100,120]. Some medications paradoxically increase challenging behaviors [101] and negatively impact cognition and neurorecovery [121,122].

For the question, "What are successful approaches to managing behavioral dysregulation during inpatient rehabilitation?" an inpatient rehabilitation provider made the following statement:

> *Interdisciplinary approaches to behavioral management, with high team involvement, high family involvement, and collaboration across disciplines.*

Preliminary data support the need for systematic implementation of a comprehensive, interdisciplinary behavior intervention program. Our preliminary data gathered from a national survey of over 300 brain injury treatment providers revealed that the most frequently endorsed patient outcomes impacted by maladaptive behaviors included limited options for discharge placement (91%), poor therapy engagement (90%), interference with activities of daily living care delivery (86%), and increased length of stay (85%). The most frequently endorsed system outcomes impacted were an increased number of staff needed (85%) and increased staff burnout (81%) [58,99]. From the health care infrastructure standpoint, facilitators for managing behavior included a culture that prioritized staff collaboration and communication (69%) and specialized staff treating individuals with brain injury (61%). The identified barriers for managing maladaptive behavior included inconsistent treatment strategies (35%) and an ill-prepared, untrained workforce (65%). Some of the most important treatment elements for managing maladaptive behaviors included the use of environmental strategies (68%), the use of individualized care plans (67%), and the implementation of objective measures to monitor progress (35%).

Additional preliminary data collected from 28 PRC staff indicated that they lacked confidence in managing challenging behaviors (77% of team members requested additional training). In another preliminary study with a team at a civilian site, an educational program improved the confidence of rehabilitation staff in addressing behavioral disturbances [123]. These preliminary results highlight key facilitators and barriers that impact access to effective care. The findings provide intervention targets to inform treatment planning, program development, and educational programming.

XSL•FO
RenderX

The STAR-VA program addressed a similar gap in managing the challenging behaviors of persons with dementia receiving care in VA facilities. The STAR-VA program was developed at the request of the VA National Mental Health Director for Psychotherapy and Psychogeriatrics to improve the management of challenging dementia-related behaviors in VA Community Living Centers [102]. The STAR-VA intervention is an interdisciplinary behavioral approach based on the intervention by Teri et al [124], which was developed for training non-VA care workers to improve the care of older adults with dementia. Since the national rollout of STAR-VA over 10 years ago, empirical studies have consistently demonstrated that veterans enrolled in STAR-VA experience significant decreases in the frequency and severity of target behaviors, depression, anxiety, and agitation, with improved staff confidence [125,126].

STAR-VA is a collaborative, interdisciplinary team intervention based on universal behavioral management principles that are applicable to any patient population. The core elements include: (1) setting realistic expectations; (2) effective communication; and (3) identifying antecedents, behaviors, and consequences (ABCs) in order to modify the severity and frequency of challenging behaviors. While these elements are applicable to any patient population, additional elements specific to each target population are needed. As a result, STAR-VA training has not been authorized for other VA programs/providers, despite requests throughout the VA system of care and the need for standardized interventions to manage behaviors in many other patient populations. The goal of this study is to adapt the STAR-VA program for use with rehabilitation teams that care for patients with TBI. The STAR-VA program has consistently demonstrated efficacy in improving staff response to challenging behaviors, and we anticipate similar findings when it is adapted for teams treating veterans with brain injury.

We intend to adapt the STAR-VA manual to target brain injury and related behaviors. We will build upon the concepts presented in the manual to highlight how the social and physical environment can be set up to optimize adaptive behaviors following TBI and how behavioral interventions can reduce the incidence of challenging behaviors. In addition to staff training, the contents of the manual will guide staff in how to acquaint themselves with and train family members so that they can maintain the adaptive behaviors when the individual transitions home. We will also provide guidance and strategies on how to set up and sustain the program. We will be transforming a staff training manual developed for the dementia population into a playbook that will provide not only materials to train staff but also program implementation guidance to secure a more lasting and impactful change in TBI rehabilitation programs.

Objectives

The aims are as follows:

- Aim 3.1 (discover, define, and develop): We aim to co-design a TeamBI Playbook through engagement of rehabilitation team members, persons with brain injury, and family members/care partners to identify areas of the STAR-VA manual that need adaptation and gap areas that need further development.
  - Aim 3.1a: We aim to engage persons with brain injury and family members/care partners to understand participants' lived experiences with participating in rehabilitation and their perspectives regarding acceptable practices for optimizing adaptive behavior and reducing challenging behavior. We will conduct virtual focus groups with veterans having brain injury and their family members/care partners who have experienced rehabilitation.
  - Aim 3.1b: We aim to engage representatives from rehabilitation disciplines to identify the modifications and additions needed to transform the STAR-VA manual into the TeamBI Playbook. We will conduct 12 working meetings with IEPs to adapt the STAR-VA manual and integrate insights from focus groups (aim 3.1a).
  - Aim 3.1c: We aim to produce the TeamBI Playbook. Utilizing the data gathered in aims 3.1a and 3.1b, the investigators and ISC will compile and package the content of the playbook into an accessible online format.
- Aim 3.2 (validate): We aim to validate, refine, and disseminate the playbook. An initial validation study will be conducted through a pilot to determine the impact on rehabilitation team leaders' self-evaluation of capability, opportunity, and motivation to implement the TeamBI program [126]. We will also determine if additional refinement is needed to address barriers to implementation.
  - Aim 3.2a: We hypothesize that exposure to the TeamBI Playbook will change rehabilitation leaders' self-evaluation with regard to their (1) capability to implement a team-based program to reduce challenging behaviors and improve adaptive behaviors, (2) opportunity to implement a team-based program to reduce challenging behaviors and improve adaptive behaviors, and (3) motivation to implement a team-based program to reduce challenging behaviors and improve adaptive behaviors. A sample of rehabilitation leaders will complete a baseline assessment of capability, opportunity, and motivation to implement a team-based behavioral intervention program. They will then be exposed to the playbook and undergo reassessment. Pre- and postexposure responses will be compared to evaluate the impact of the playbook.
  - Aim 3.2b: We aim to determine if the playbook's implementation strategies sufficiently address potential barriers. The rehabilitation leaders will complete an implementation readiness survey, and responses will be used to identify barriers that were not sufficiently addressed in the playbook.
  - Aim 3.2c: We aim to refine, finalize, and disseminate the playbook. We will refine the playbook based on findings from aims 3.2a and 3.2b. Following final review by the CEC, the playbook will be disseminated through the project toolshed, leveraging partners from the VA, other federal agencies, and professional organizations.

XSL•FO
RenderX

Design

This project falls within the preimplementation phase of the QUERI Implementation Roadmap. The goal is to improve access to rehabilitation programs that effectively promote adaptive behaviors and reduce challenging behaviors, thus allowing patients with TBI to fully participate in rehabilitation. The proposed mixed methods study will engage veterans with TBI, family members/care partners, and rehabilitation professionals to develop and design the TeamBI Playbook (aim 1) and conduct a pilot to validate, refine, and disseminate the intervention for future implementation and evaluation (aim 2).

Aim 3.1 (Study Months 3-24)

Participants

The participants in the focus groups will include veterans with brain injury and their family members/care partners. The inclusion criteria are as follows: (1) age 18 or older, (2) participants with TBI who have received rehabilitation and family members/care partners of veterans with TBI who have received rehabilitation within the past 5 years, and (3) online access to participate in a virtual meeting.

We will be conducting focus groups with a homogeneous composition. There will be 3‐6 focus groups with veterans having TBI and 3‐6 focus groups with their family members/care partners. Each focus group will have 6‐8 participants, resulting in a total sample size of approximately 64 participants. The sample size was determined based on the understanding of what it will take to reach thematic saturation in the field [127-129].

The sample will be recruited with the assistance of the CEC, IEPs, and VA PRCs. We will seek to recruit at least 25% of participants representing a racial/ethnic minority and oversample female veterans.

Data Collection

Overview

We will use 3 tools in data collection. The first tool is a brief demographic questionnaire to collect the following information from participants: age, race, gender, and date of rehabilitation. The second tool is the Ohio State University TBI Identification Method (OSU TBI-ID) [130] to ensure consistency in identifying TBI exposure. The OSU TBI-ID has established reliability and validity and has been used with veterans to provide lifetime exposure rather than being limited to injuries experienced during deployment [131-136]. The third tool is a focus group interview guide consisting of 8 questions (1 introductory question, 6 main questions about their inpatient rehabilitation experiences, and 1 concluding question). The focus group interview guide will be reviewed and revised during study initiation to ensure the questions are at an 8th-grade reading level and follow federal plain language standards to promote understanding [137]. The focus group guide will be pilot tested with people representing those we will engage to ensure comprehension of the questions and content, and it will be finalized based on the pilot test data for IRB approval prior to use.

Focus Groups

To begin our process of adapting the existing STAR-VA manual, 2 trained qualitative researchers will conduct virtual focus groups with veterans having brain injury and with their family members/care partners. The goal of the focus groups is to understand participants' lived experiences engaging in rehabilitation as it relates to optimizing adaptive behavior. We will seek to understand the acceptability of various behavioral intervention practices from the perspectives of the veterans and their family members/care partners. Participants will be consented and assured of their rights and privacy. They will complete the demographic questionnaire and OSU TBI-ID in advance of the focus groups. The focus groups will last no longer than 90 minutes and will be audio-recorded with their permission. The virtual platform for the focus groups will include closed captioning and verbal descriptions of visual stimuli to improve accessibility.

The focus groups will be conducted by a facilitator who asks questions, probes for additional information, and manages group dynamics (including the use of cognitive accommodations when needed), and a co-facilitator who documents and summarizes participant responses, probes with clarifying questions, and manages recording technology. Focus group recordings will be transcribed and uploaded to ATLAS.ti version 22 (ATLAS.ti Scientific Software Development GmbH), along with any notes taken during the focus groups.

Working Groups

IEPs and SMEs, including rehabilitation therapists, nurses, behavioral health providers, physicians, and other team leaders, will be invited to a series of 12 one-hour structured working meetings, which will be scheduled over the course of 18 months. The research team will prepare and distribute materials for review 2 weeks prior to each meeting. Meetings will be recorded with permission and facilitated using a structured agenda outlining meeting goals, key questions, and expected deliverables. At the conclusion of each meeting, attendees will be asked to provide simple plus/delta feedback to evaluate and modify meeting processes in a cycle of continued improvement (unpublished data).

Six meetings will evaluate sections of the manual to determine gaps, identify areas for improvement, and determine modifications to be made based on panel members' expertise and data insights from the focus groups (aim 3.1a). Another 6 meetings will focus on measurement to identify appropriate outcome measures (eg, reduction in challenging behaviors, increased discharge to home, and improved staff morale/reduced burnout). In preparation for aim 3.2, input from the IEPs and the entire CEC will be sought to (1) adapt an existing measure of capability, motivation, and opportunity to determine the readiness of rehabilitation professionals to implement the TeamBI program [138] and (2) choose items from the CFIR Interview Guide to solicit open-ended responses regarding barriers and facilitators to implementation of the TeamBI program.

XSL•FO
RenderX

Playbook Production

The content of the e-playbook will be compiled into a web-accessible electronic package. The package will include chapters, tools (eg, checklists, A-B-C forms, and outcome measures), the implementation readiness measure, and a recorded webinar describing the playbook. Our e-playbook will be compiled to ensure relevance, accessibility, acceptability, and utility to all end users, including those with disabilities. Our digital presence will align with Section 508 of the Rehabilitation Act of 1973 [139].

Analysis Plan

Focus group transcripts will be analyzed using matrix analysis, a rapid assessment approach, to summarize and identify themes. Rapid assessment is a team-based approach to iteratively collect and analyze qualitative data that emphasizes the speed of data collection and analysis in relation to focused programmatic questions or problems [62]. Two experienced qualitative researchers will conduct the matrix analysis. Together, they will create codes both deductively from known constructs and inductively from the data, and use matrices to categorize responses under domain names, applying codes that emerge from the data. Codes and preliminary themes will be reviewed by the research team and IEPs to establish consensus. Data insights will be aligned with relevant sections of the playbook (eg, experiences of communication, environment, staff behaviors, and family inclusion in inpatient rehabilitation environments) in a brief, consumable report for IEPs to incorporate into the design of the playbook (aim 3.1b).

Aim 3.2 (Study Months 25-48)

Participants

Participants will include at least 30 rehabilitation professionals in team leadership roles that align with program development and implementation. While we anticipate that most will be employed at a VA PRC, some will also be drawn from civilian facilities that serve veterans. Each included participant must (1) be a rehabilitation professional in one of the key leadership roles for TeamBI program implementation (team leader, administrator, behavioral health provider, and nurse manager) or hold a position that includes program development and implementation in job duties, (2) be employed in a TBI inpatient rehabilitation setting that serves V/SMs, and (3) have online access.

Participants will be recruited with assistance from the CEC. To ensure diversity, we will call upon volunteers from the CEC to assist with outreach. Consent will be obtained online prior to commencement of study procedures.

A sample size of n=30 has at least 80% power at 5% significance to detect a change in means associated with an effect size as small as 0.53 (change in mean/pretest SD) [140,141], which is considered to be a medium effect size.

Data Collection

The Capability, Opportunity, Motivation to Implement a Team-Based Behavioral Intervention (COM-TBI) model, a pre/posttest instrument based on the COM-B model [142,143], will be used to evaluate the impact of exposure to the playbook on rehabilitation professionals' likelihood to implement the TeamBI program. The COM-B model purports that in order for change to occur, the individual needs to be motivated; however, motivation will not lead to behavior change unless the individual perceives themselves as capable and having the opportunity to perform the behavior. Typically, studies of behavior change, including the likelihood that an individual will implement a specific program, require the construction of a study-specific measure. However, a generic 6-item measure was developed specifically for adaptation across different studies [138].

The generic questionnaire was tested with health care professionals and a sample of individuals with low socioeconomic status, and was found to be acceptable, understandable, reliable, and valid [138]. Scores on the scale significantly predicted actual behavior change (delivery of opportunistic behavior change interventions by health care professionals). Confirmatory factor analysis indicated a well-fitting 3-factor model aligning with the 3 constructs (capability, opportunity, and motivation). The Bayesian information criterion and Akaike information criterion indicated that the 3-factor model was superior to a unidimensional model. We will adapt this generic measure, tailoring items to be relevant to the implementation of a team-based behavioral intervention program. Since the participants will not have been exposed to the playbook at the time of the baseline pretest, they will be asked to respond based on the implementation of a generic team-based behavioral intervention program for patients with TBI. The CEC and IEPs will assist with item development and piloting prior to use with the study sample. The psychometrics of the adapted instrument will also be evaluated with the study sample.

The CFIR Interview Guide will be used to identify open-ended questions for use in evaluating implementation readiness and the identification of any determinants (barriers and facilitators) that may not have been sufficiently addressed in the playbook. The CFIR guide has been used internationally to study all types of health conditions in military and civilian health care settings [59,60]. The overarching access framework domains from the supply/organization side and consumer/demand side will also inform the focus group questions, and they have been successfully used in our own TBI research [6,46,47,58]. The interview guide will be pilot tested with people representing those we will engage to ensure comprehension of the questions and content. We will include questions about the TeamBI program's adaptability, complexity, and cost; the design quality and packaging; and the implementation climate at their settings. We will also use the guide's questions about the individual's knowledge and beliefs about the TeamBI program, self-efficacy, and the individual stage of change to implement the program. Responses will be used to further refine the playbook.

Information will be obtained from participants regarding age, gender, race, profession/discipline, job title, type of facility where they are employed, whether the facility routinely serves veterans with TBI, whether the facility is a VA or civilian facility, and whether they have online access.

Participants will be sent a link to the consent form, and after they have successfully completed the consent, a link will open

XSL•FO
RenderX

a screening form to ensure they meet the inclusion criteria. Participants who meet the inclusion criteria will be linked to the baseline version of the COM-TBI to evaluate capability, motivation, and opportunity to implement a generic team-based behavioral intervention. Upon completing the COM-B, they will be sent the playbook; a link to the online webinar that reviews the playbook's contents; and a link to the postexposure version of the COM-TBI to evaluate capability, motivation, and opportunity to implement the TeamBI program. They will also be sent the open-ended questions from the CFIR Interview Guide. Upon completion of the postexposure measures, the participants will receive continuing education credit and a financial incentive (US $50). The total time involved is estimated to be 2 hours per participant.

Analysis Plan

The sample will be used to test the hypothesis that participants' perceptions of their capability, opportunity, and motivation will improve from baseline after exposure to the playbook. Means, SDs, and percentiles will be used to summarize continuous data and frequency counts, and percentages will be used to summarize categorical variables. Sample characteristics (eg, age, gender, race, setting [type and civilian vs VA], job title, and discipline) will also be collected and summarized for participants in this sample. Paired *t* tests will be used to test for a change in capability (hypothesis 1), opportunity (hypothesis 2), and motivation (hypothesis 3) from baseline to postexposure at a significance level of α=.05. We will also explore whether setting (VA vs civilian and facility type) has an impact on improvements in knowledge, confidence, and motivation by examining the means, SDs, and effect sizes of each of the outcome variables using analysis of covariance models.

For aim 3.2b, data from the implementation readiness survey will be compiled to identify barriers that may need to be addressed with additional implementation strategies. Our preliminary study yielded a wealth of information from open-ended questions about determinants to implementing behavior programs, and we anticipate that we will receive similarly detailed responses in this study. Open-ended responses will be analyzed using matrix analysis, a rapid assessment approach, to summarize and identify themes. Rapid assessment is a team-based approach to iteratively collect and analyze qualitative data that emphasize the speed of data collection and analysis in relation to focused programmatic questions or problems [81]. Two experienced qualitative researchers will conduct the matrix analysis. Together, they will create codes both deductively from known constructs and inductively from the data, and use matrices to categorize responses under domain names, applying codes that emerge from the data. Codes and preliminary themes will be reviewed by the research team and CEC to establish consensus.

For aim 3.2c, in consultation with the ISC, investigators will refine the playbook based on the findings from aims 3.2a and 3.2b. Strategies will be added for any barriers that were not sufficiently addressed. Any setting-specific barriers will be called out in the playbook with strategies targeted to the setting. Approval of the final playbook will be obtained from the CEC version of the playbook before moving to wider dissemination via the project toolshed.

Potential Challenges and Solutions

A potential challenge is the possibility of the team facing recruitment difficulties. Recruitment outreach will be expanded using recommendations from the CEC.

Another potential challenge is the possibility that rehabilitation professionals' perspectives of capability, opportunity, and motivation are not improved by exposure to the playbook. The CEC and ISC will be consulted for suggested revisions and guidance.

Community-Based Participatory Approach

Community-based engagement is being conducted across project aims and communities. Engagement throughout the project aims is designed to improve outcomes for V/SMs (better access), improve the health care system (system outcomes), and improve implementation (increase buy-in).

Individual Project 4: Data-Driven Policy Recommendations for Virtual Health Care for Persons With TBI Morbidity

Problem and Solution

Problem

Universal policy mandates for virtual health to enhance health care access exist in VA and the DOD; however, no research exists for understanding its successful implementation among persons with TBI morbidity. Implementation of virtual health resources has been a priority initiative for VA and the DOD for the past 2 decades because virtual technologies can reduce barriers to accessing care. For example, over 5.5 million veterans actively use the VA's electronic health portal (My HealtheVet) [144]. JNH and other virtual health experts in the field have spent more than a decade evaluating users' experiences and implementing HCD efforts to support uptake and sustained use of virtual health resources. However, the COVID-19 pandemic caused a shift toward sustaining access to care, relying heavily on the use of virtual health resources. What was once an option for accessing care has now become a necessity. During the COVID-19 pandemic, implementation efforts and emergency mandates saw a landmark increase in the use of virtual health resources across all health care services [145]. However, the recent universal mandates to increase access do not consider the unique needs of populations who may require accommodations for cognitive impairments, including alternative communication approaches for accessing health care. As the field of TBI systems of care and the health care systems in general embrace protocols that take a "one-model-fits-all" approach to virtual health care resource use, clinicians are identifying barriers that are uniquely present among V/SMs with TBI. To date, there has been little, if any, research focusing on the virtual health care resource needs of persons with TBI.

Solution

This project proposes to leverage existing data to develop a taxonomy of domains and themes relevant to access to virtual health care for persons with TBI morbidity. The study will

XSL•FO
RenderX

leverage an existing qualitative dataset examining facilitators and barriers to chronic pain care in persons with TBI. The use of virtual health resources was commonly noted but not the focus of the primary study. This proposal offers a cost-efficient approach that will launch this untapped area of science into a *discovery* process, which can inform an HCD approach to providing recommendations regarding the appropriate use of virtual health care delivery to persons with TBI. A preliminary review of the relevant dataset indicates a clear need to re-evaluate the data through a lens to assess barriers to the use of virtual health resources among persons with TBI. Although virtual health resource use emerged as a main theme in NIDILRR analyses, facilitators and barriers to using virtual health care were not explored. The focus of the secondary analysis will be on virtual health resource use in the context of delivering chronic pain care, which is the top comorbidity among persons with military TBI [8]. Analysis of this secondary data will inform the identification of (1) necessary virtual health resource accommodations and (2) policy recommendations needed to determine appropriate virtual health resource use for persons with TBI-related morbidity (eg, cognitive, physical, and behavioral impairments).

Background

The COVID-19 pandemic shifted how medical care is provided to patients through a rapid increase in virtual health resources and/or the adoption of virtual health resources to facilitate safe, timely, and accessible care remotely [146,147]. Before the pandemic, V/SMs with chronic TBI reported that distance is a barrier to accessing needed health care services; however, they also indicated that virtual health care is a barrier [6]. Virtual health care is mandated to foster remote access to care, particularly for V/SMs in distant/rural locations [145]. However, universal mandates to increase access do not consider the unique needs of populations who may require accommodations for cognitive impairments, including alternative communication approaches, additional resources, and extra time for appointments, in order to optimally access health care. Qualitative data from provider interviews conducted as part of our NIDILRR-funded study on TBI and chronic pain found that most providers (n=63; rehabilitation therapists, medical doctors and nurses, and psychologists from VA, the DOD, and community health care systems) described the use of virtual resources and telehealth as a main facilitator (93%) to providing care to patients during the COVID-19 pandemic [148]. However, TBI morbidity (ie, cognitive impairments) was reported by 63% of the sample (in thematic analysis) as the primary barrier to delivering health care.

A neuropsychologist made the following statement:

> *[During COVID-19] I did a fair amount of telehealth which included video, actually no, I'll take that back. A lot of telephone contacts. Videoconferencing was not as successful due to the cognitive impairment of the patients.*

Qualitative data from TBI providers highlight the need to better understand how virtual health care interventions require adaptation for TBI survivors with cognitive and physical limitations [46,47]. The use of virtual health care emerged as a main theme in facilitating access to care, but barriers persist for V/SMs with TBI and chronic pain. Understanding and developing a taxonomy of access to virtual health care for persons with TBI and chronic pain can inform the accommodations needed to improve access across relevant access framework dimensions. This taxonomy will inform the development of content needed to inform the creation of policy recommendations for the appropriate use of virtual health care resources among persons with TBI. JNH and Bridget A Cotner will leverage previous experience using mixed-methods and CBPR stakeholder engagement strategies with partners to co-create policy recommendations and the dissemination products needed to support the spread and adoption of stakeholder-driven policy recommendations. This project proposes a QI approach that will only engage CEC PPPs for the purpose of developing recommendations and products relevant to the appropriate use of virtual health care resources among persons with TBI morbidity.

Objectives

This project proposes an HCD and CBPR approach with sequential mixed methods to inform the appropriate use of virtual health care resources among persons with TBI (2.5-year implementation study). This proposed project will leverage existing data archived from a multicenter NIDILRR-funded study of TBI health care needs. The aims of project 4 are presented in Figure 5. The aims are as follows:

- Aim 4.1 (discover and define - secondary analyses; 6 months): We aim to characterize data (facilitators and barriers) and develop a taxonomy to inform virtual health care delivery for persons with TBI morbidity.
  - Aim 4.1a: We aim to access and create datasets. Existing provider interview transcripts will be accessed for secondary analyses. This qualitative dataset has been generated from an NIDILRR-funded multicenter trial and includes 63 semistructured interviews with rehabilitation providers across disciplines and settings. Interview questions were informed by the access to health care conceptual framework and focused on treatment and referral practices for persons with TBI and chronic pain, and facilitators and barriers to accessing treatment.
  - Aim 4.1b: We aim to develop and analyze the codebook. Secondary analyses of this existing dataset will allow the identification and organization of data to inform: (1) necessary accommodations and (2) policy recommendations needed to determine the appropriate use of virtual health resources to improve access and user experiences. Upon creation of the preliminary codebook and analysis of 10% of the transcripts, the CEC (PPPs and LEPs) will be engaged for the first stakeholder panel meeting (n=10 - 15) to participate in codebook construct definition review and interpretation of initial data findings to inform a stakeholder-driven analytic process in order to ensure relevance, accuracy, and usefulness of the data findings. CEC review and feedback will inform the analysis of the remaining 90% of the data transcripts.

XSL•FO
RenderX

- Aim 4.1c: We aim to create data summaries and a taxonomy. Coded archived transcripts will be analyzed to create a taxonomy of data findings, which will inform a multilayered approach to understanding the use of virtual health resources among persons with TBI, including facilitators, barriers, opportunities, and recommendations. A second stakeholder panel meeting will engage them for a review of data summaries and for providing input on product development. Finalized data summaries and the taxonomy with stakeholder input will allow data-driven development of products (eg, executive summary, briefing, presentation slide deck, white paper, testimonial, etc) to inform accommodations and recommendations for virtual health care delivery for persons with TBI.

- Aim 4.2 (develop - stakeholder-driven product and dissemination plan development; 1 year): We aim to engage partners to identify the necessary accommodations and recommendations to inform the development of clinical and policy decision-making support for virtual health resource use among persons with TBI morbidity.
  - Aim 4.2a: We aim to conduct gap analysis and draft content. A gap analysis will be conducted to inform the development of a product grid and dissemination plan to inform clinical and policy decision-making products in order to inform accommodations and recommendations for virtual health resource use among persons with TBI. The initial product development product grid will take a preliminary approach to prioritizing: (1) products and deliverable items, (2) target audiences, (3) purpose, (4) content needed, (5) format, (6) complexity, (7) readiness, (8) weighted score (complexity × readiness), (9) pain points, (10) designated roles of developers, and (11) potential impact.
  - Aim 4.2b: We aim to engage the panel to inform product grid, product, and dissemination plan development. An established protocol for product development and validation will be used to re-engage CEC members (PPPs only, ie, key policymakers in the DOD, VA, and civilian health care systems) to participate in a third stakeholder panel meeting (n=8 - 10) to review and inform the development of a stakeholder-driven prioritized product grid, product, and dissemination plan to facilitate high-impact data-driven clinical and policy decision support tools. Panel activities will inform the review and weighted scores (complexity × readiness) of products to inform prioritized product development.
  - Aim 4.2c: We aim to develop products and a dissemination plan. Upon finalization of the stakeholder-driven product grid, the products and dissemination plan will be developed (eg, executive summary, briefing, presentation slide deck, white paper, testimonial, etc) to inform clinical and policy decision support relevant to the appropriate use of virtual health resources among persons with TBI.

- Aim 4.3 (validate - products and dissemination plan; 1 year): We aim to re-engage CEC PPPs (ie, key policymakers in the DOD, VA, and civilian health care systems) to inform the HCD approach for creating stakeholder-driven products and a dissemination plan to support data-driven virtual health resource use among persons with TBI morbidity.
  - Aim 4.3a: We aim to compile and deliver content to panelists to prepare for a panel meeting. The team will compile a package of stakeholder-driven products and the dissemination plan and deliver the package to PPP panel members from aim 4.2, with instructions for preliminary review to prepare for the final panel meeting to be conducted in 4.3b.
  - Aim 4.3b: We aim to engage the panel to obtain feedback on the final product grid, products, and dissemination plan. A fourth stakeholder panel meeting (n=8 - 10) will be conducted to inform the requirements needed to finalize the product grid, products, and dissemination plan. Panel meeting activities will review the evidence-based process of product development, the final review of products, and the dissemination plan. The PPP panel activities in aim 4 will be the “soft launch” of the dissemination plan, as this aim will rely on key DOD, VA, and civilian health care system representatives who are positioned to champion planned efforts subsequent to this project timeline, to disseminate products for supporting clinical and policy decisions.
  - Aim 4.3c: We aim to finalize the product grid, products, and dissemination plan. Feedback from the fourth, and final, stakeholder panel meeting will inform the finalization of products and the dissemination plan.

XSL•FO
RenderX

Figure 5. Project 4 aims. HCD: human-centered design.

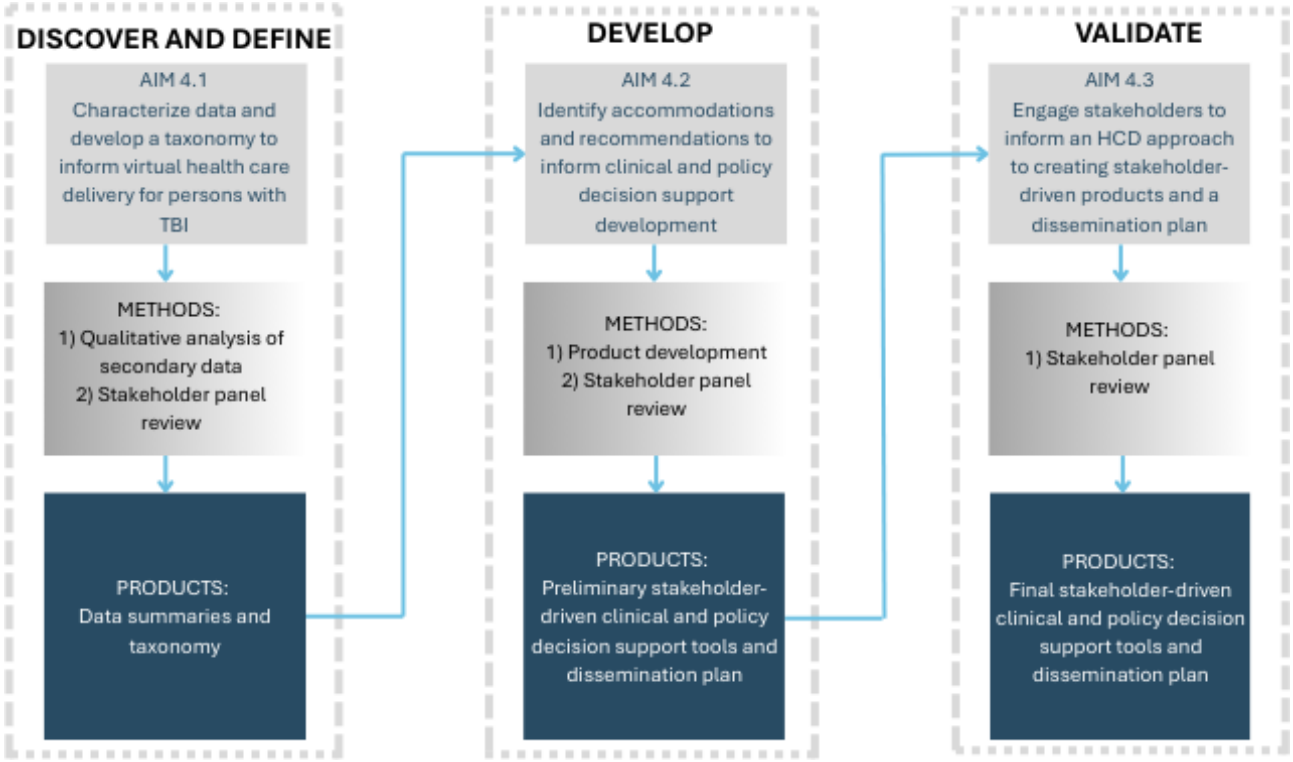

Design

Overview

This project will use HCD and CBPR approaches, adopting sequential mixed methods to complete the QI project. Leveraging existing qualitative interview data from the NIDILRR study provides a unique opportunity to conduct secondary analyses to understand specific barriers to treatment and delivery for persons with TBI and chronic pain [46,47]. Investigators and analysts from the parent study are the investigators for the proposed project, who have familiarity with the dataset leveraged for secondary analyses and methods, thus maximizing the success of the study objectives.

In alignment with the I-HEAL theoretical framework, HCD and CBPR approaches, using sequential mixed methods, will be used to complete the QI project. Descriptive qualitative methods will be used to conduct the secondary qualitative data analysis of semistructured provider interview data that have been recently collected (October 2020-November 2021) as part of the NIDILRR-funded "Characterization and Treatment of Chronic Pain after Moderate to Severe TBI" study. CBPR will be used to facilitate stakeholder engagement activities for partnering with key CEC partners throughout the proposed project activities to complete the project aims.

Aim 4.1 (Discover and Define - Secondary Analyses; 6 Months)

Two data sources will be utilized to conduct the secondary analysis. First, transcripts of semistructured interviews will query providers about their treatment and referral practices for patients with TBI and co-morbid chronic pain, including what improves or impedes access to treatment. These semistructured interviews consist of 14 questions that are informed by the Access to Healthcare Framework [10]. The qualitative data collection team will utilize probing to ensure important concepts related to access are fully discussed. Second, a provider questionnaire consisting of 14 questions will characterize the provider sample, including information on work settings and the types of patients they treat. No new data will be collected for aim 4.1.

Aim 4.1 will be conducted in 3 phases. In aim 4.1a, data will be accessed and managed for analysis. Dr Bridget A Cotner, an anthropologist and qualitative research expert at Tampa VA who led the primary data collection activities and analysis, will lead aim 1 of this project. The qualitative team, consisting of 2 trained qualitative researchers and led by Dr Bridget A Cotner, has access to the interview data and qualitative analysis software, ATLAS.ti (v.22) [149], to be used for the proposed analyses. Data are housed on a secure password-protected network that is accessible only by qualified research team members. In aim 4.1b, codebook development will be completed. To adequately capture virtual health care access barriers among those with TBI, a qualitative codebook will be developed using deductive codes generated from relevant constructs reflected in the Access to Healthcare Framework and CFIR framework (see Figure 2), and it will be developed inductively from interview data and input from CEC partners, including PPPs, LEPs, and IEPs. As such, once the team develops the initial codebook constructs with definitions and 10% of the transcripts are analyzed, Drs JNH and Bridget A Cotner will facilitate an initial 90-minute virtual or in-person stakeholder panel meeting (n=10 - 15) for codebook construct definition review and interpretation of the initial data findings. The stakeholder panel will drive the analytic process to ensure relevance, accuracy, and usefulness of the data findings.

XSL•FO
RenderX

The maximum number of potential participants (PPP, LEP, and IEP group members) is 15, which is beyond the size of a typical group for facilitating a focused group discussion with activities. If the project team determines that the codebook constructs, definitions, and initial analyses are broad and/or complex, the team may opt to divide the PPPs and IEPs into one panel group and the LEPs into another stakeholder panel group. Collectively, the CEC panel meeting review and feedback will inform the analysis of the remaining 90% of the data transcripts.

In aim 4.1c, coded transcripts will be used to create a taxonomy of data findings. Interview transcripts will be recoded by the qualitative team using the codebook augmented by our CEC partner feedback to determine barriers and facilitators to virtual care delivery. Additional codes will emerge inductively from the interview data. The coding team will read one interview transcript separately and discuss the addition of new codes with examples. This process will continue with subsequent transcripts until no new codes are generated. Interrater reliability will be established when the coding team reaches at least 80% coding agreement. Intercoder reliability will be routinely monitored to ensure consistency and limit potential drift in coding. Any discrepancies will be discussed and resolved among PIs, co-investigators, and study staff during weekly meetings to ensure coding and analysis are completed on schedule. Analysts will adopt a constant comparative approach to conduct a matrix analysis [150], a rapid assessment approach [62], which will be used to develop themes for the overall sample by system of care (VA/DOD: n=26; civilian: n=37) and by provider discipline (rehabilitation therapists: n=28; mental health provider: n=15; medical provider: n=15). This analysis will be organized into taxonomies of domains and subdomains that reach saturation. Development of a taxonomy relevant to the informed use of virtual health resources among persons with TBI will provide a formal system for classifying a set of common conceptual domains and dimensions to this multifaceted complex topic. This taxonomy will inform a multilayered approach to understanding: (1) mandated versus actual use of virtual health resources, (2) facilitators and barriers, (3) opportunities and recommendations for virtual health resources for persons with TBI, (4) identified dissemination and implementation strategies, and (5) other factors that surface as relevant to the topic throughout data analysis. A second 90-minute virtual or in-person stakeholder panel meeting will engage them in review of the final data summaries and taxonomy to provide input on product development and dissemination planning. Stakeholder input will inform stakeholder-driven product development (eg, executive summary, briefing, presentation slide deck, white paper, testimonial, etc).

Qualitative data analysis relies on the saturation of data to validate findings, that is, when no new information emerges. The larger parent study, Characterization and Treatment of Chronic Pain after Moderate to Severe Traumatic Brain Injury: A Qualitative Study (PR00039496), provided a sample size justification based on the principle of “saturation.” Due to the specificity of the proposed analysis and interview participants who are knowledgeable of this topic [128], a smaller sample size is needed. As such, for this secondary analysis, themes are expected to reach saturation for the overall sample by system of care (VA/DOD: n=26; civilian: n=37) and by provider discipline (rehabilitation therapist: n=28; mental health provider: n=15; medical provider: n=15) due to a sample size that is more than adequate.

Aim 4.2 (Develop - Stakeholder-Driven Product and Dissemination Plan Development; 1 Year)

Aim 4.2 will be conducted in a multiphase stakeholder engagement process to drive product and dissemination plan development. This aim will involve 3 phases. In aim 4.2a, a gap analysis will be conducted based on aim 4.1 findings and previous experience to inform the development of a product grid to inform accommodations and recommendations for virtual health resource use among persons with TBI. Based on previous experience, the team anticipates that a collective group of products (eg, executive summary, briefing, presentation slide deck, white paper, testimonial, etc) will be required for development to inform clinical and policy decision support relevant to the appropriate use of virtual health resources among persons with TBI. The product development product grid will take a preliminary approach to prioritizing (1) products and deliverable items, (2) target audiences, (3) purpose, (4) content needed, (5) format, (6) complexity, (7) readiness, (8) weighted score (complexity × readiness), (9) pain points, (10) designated roles of developers, and (11) potential impact.

In aim 4.2b, CEC PPP and IEP partners from aim 4.1 will be re-engaged for a third 90-minute in-person or virtual panel meeting (n=8 - 10) to present data summaries, the taxonomy, and the preliminary list of products and dissemination activities. Re-engagement efforts will focus on PPPs, those with leadership experience in the areas of virtual care delivery and TBI, and other panel members, as deemed appropriate by the aim 4.1 data findings. The aim 4.2b panel meeting will be recorded, and the facilitators (led by JNH and Bridget A Cotner) will use methods with a semistructured interview script to solicit respondents’ perceptions about the information and marketing needs and preferences to inform and solicit respondents’ recommendations. During the panel meeting, data summaries and the taxonomy will be reviewed and evaluated for validity among panel members, and the proposed products and dissemination activities will be reviewed. The team will work with partners to ensure that the dissemination plan includes (1) products for engaging partners with targeted messages, (2) an overview of scheduled communication activities, and (3) anticipated outcomes for stakeholder groups. During this panel meeting, the team will address the identification of all partners and specific contact information to ensure all potential influencing partners are identified. Troubleshooting barriers and discussing workarounds for current policies on the use of virtual health care delivery for persons with TBI will be specifically addressed. JNH and other team members will prepare and anticipate potential barriers and troubleshooting, as these topics tend to arise during policy development and implementation. During the panel meeting, the team will conduct an “activity” with panel members to prioritize products by level of complexity, readiness, and potential impact. Data received during the panel and field notes will be documented and integrated using an established qualitative rapid content analysis process. Audio recordings will also be transcribed and analyzed

XSL•FO
RenderX

using rapid content analysis. Initial analyses from the panel meeting will be converged with the analyzed transcribed data, and this approach will support the validity of the findings. This approach will also maximize the relevance and feasibility of recommendation adoption and policy change throughout the DOD, VA, and civilian health care systems.

Brief follow-up calls may be conducted as needed with panel partners (or small groups of 3 - 4 members) to finalize or clarify any unresolved topics that might arise during the virtual panel meeting or during subsequent analysis. This iterative development process allows for the emergence of data and/or recommendations, which, although unforeseen, reflect stakeholder values throughout the study process. Once all data are reviewed and compiled, data and product priority rankings will be used to finalize the project product grid and inform and finalize product and dissemination plan development. In aim 4.2c, product development and dissemination planning will be completed.

Aim 4.3 (Validate Products and the Dissemination Plan; 1 Year)

In aim 4.3, the team will re-engage CEC partners to validate the clinical and policy decision-making support products and the dissemination plan. This process will ensure that the products and plan are relevant and useful, and have optimal potential for high impact. Given the professional roles of PPPs, as key DOD, VA, and civilian health care system representatives, this group will be the “front line” of leaders who are uniquely positioned to support dissemination and policy impact. As such, this aim will be the “soft launch” of the dissemination plan, as this aim will rely on those who are positioned to champion planned dissemination efforts.

Aim 4.3 will be conducted in 3 phases. In aim 4.3a, before the third and final panel meeting is conducted, the product grid, content, and dissemination plan will be sent to panel members as a compiled “package,” with instructions for preliminary review in preparation for the third and final panel meeting. Subsequently, in aim 4.3b, the fourth 90-minute panel meeting (in-person or virtual) with CEC PPP and IEP members (n=8 - 10) will be scheduled to formally present the final product grid (with ratings, product content, and the dissemination plan) to the panel to prompt discussion, collective review, and synthesis of feedback and recommended revisions. As in previous aims, the facilitators will use a semistructured script to guide major discussion points and activities to ensure that all topics are addressed and planned activities are completed. The team will finalize the product grid, products, and dissemination plan with panelists. Other issues, such as stakeholder contacts, will be confirmed as current, and opportunities for troubleshooting and workarounds will be reconfirmed. As in aim 4.2, this iterative process will allow for the emergence of data and/or recommendations, which may be unforeseen but reflect stakeholder values throughout the project phases. As with aim 4.2b, data received during the panel, data from field notes, and information from transcribed data will be converged and analyzed using rapid content analysis to optimize validity and maximize the relevance and feasibility of recommendation adoption and policy change throughout the DOD, VA, and civilian health care systems. In aim 4.3c, once all data are reviewed and compiled, Drs JNH and Bridget A Cotner will work with ISC team members, particularly Mrs Margeaux Chavez, to take a rapid, systematic approach for finalizing the product grid, products, and dissemination plan to prepare for the launch of the dissemination plan with our panel of CEC partners.

Potential Challenges and Solutions

The first potential challenge involves the rapid timeline. We will leverage existing data that was previously analyzed by Bridget A Cotner and leverage established protocols used by JNH and Bridget A Cotner in previous projects.

The second potential challenge is the limitations of secondary data findings. Though secondary data may limit insights, we have established grounds for the secondary analysis based on the primary analysis of data. Moreover, these data can inform follow-up data collection as appropriate for future project efforts.

The third potential challenge is a lack of participation by panel members. Our team has expertise in developing and disseminating resources, with stakeholder input, to support the appropriate use of virtual resource implementation with targeted audiences. Invested partners have expressed a willingness to participate on a panel to provide feedback on the development of products and dissemination planning. We will collaborate with our engagement stakeholder group to invite them to participate on a panel to provide rapid feedback on products and communication strategies to disseminate policy recommendations. Panel member participation will be minimized to 3 - 5 hours overall.

The fourth potential challenge is the limited impact of dissemination efforts. Previous work conducted by JNH has effectively resulted in policy change and redesign efforts of VA’s electronic health record and virtual health resources (ie, secure messaging), and as such, we have the experience and confidence that appropriate data validation and dissemination should have similar results in the proposed effort. Dr JNH will champion the finalized products and the dissemination with PPP members and the ISC to implement and optimize impact.

Community-Based Participatory Approach

In alignment with the I-HEAL framework, this project will leverage a CBPR approach to engage the CEC at multiple timepoints in order to inform the process, activities, and development of products and the dissemination plan. Activities to support ongoing involvement will include a virtual/in-person meeting at the beginning of the project to discuss the project’s aims and activities, and discuss the access framework and qualitative findings from initial themes mapped onto the Access to Healthcare Framework. Partners will meet at designated time points to inform analysis, discuss, create, review decision-making support products and the dissemination plan, and support and inform the dissemination plan.

To ensure panel engagement throughout the lifetime of the proposed project, panelists will receive project updates every 6 months to ensure they are aware of project aims, processes, and outcomes, and are motivated to continue to stay engaged

XSL•FO
RenderX

throughout the project. The first contact with panelists will be their invitation to participate in the project. During this initial contact, team members will inquire about 6-month updates and solicit panelists' preferences for the update format (eg, newsletter, "all call," etc). Panelists' preferred methods will be used to develop the method and protocol for providing project updates. The final project update will be used to share project findings and final products for dissemination. Drs JNH and Bridget A Cotner will work with PPP stakeholder panel members and the ISC to implement the dissemination plan subsequent to the proposed project activities.

Results

I-HEAL was funded as an implementation science FPA by Congressionally Directed Medical Research Programs, and start-up activities began in October 2023. After IRB approval (projects 1 - 3: August 12, 2024; project 4: December 2, 2019), all 4 projects are currently underway, with funding through September 2027. Project 1 has enrolled 48 participants, and project 3 has enrolled 34 participants through September 2025. Project 2 has completed the environmental scan and has identified both general and specific EBTs for inclusion in the toolkit. Project 4 has completed the secondary data analysis and is working on dissemination.

Discussion

Rationale and Expected Impact

The I-HEAL protocol contributes to the field in 4 major ways. First, it supports the overall responsiveness to the health of V/SMs, as previous research has highlighted disparities in populations with TBI and specifically military populations with TBI. Second, it supports scientific advances in the delivery of evidence-based care requiring integration and the use of implementation science in the field of TBI. Third, it creates knowledge products from both synthesized scientific evidence and our project innovations in the health care field that are critical to timely knowledge translation. Fourth, it supports the engagement of V/SMs, their families, and health care system SMEs from VA and the DOD.

Challenges

Potential challenges of the overall I-HEAL effort include (1) interdisciplinary collaboration in a new area of translation, requiring training and novel communication strategies across respective fields of science; (2) a dearth of capacity in the field of TBI and implementation science, thus requiring additional efforts to effectively communicate broadly and transparently to the field; (3) budgetary constraints to respond to new/novel strategies of engagement from established partners during the course of the grant; and (4) funding opportunities to address translational next steps, including sustainment planning for each of the individual projects and core activities.

Dissemination

The cores are actively working to support projects through partner engagement, implementation science consultation, administration, research methodology, and creating dissemination venues such as I-HEAL's website and social media accounts (Multimedia Appendix 3). Additionally, the projects have completed 4 podcasts and 39 presentations at well-known conferences, such as the Military Health System Research Symposium and the American Congress of Rehabilitation Medicine, and will continue to present and publish findings.

Conclusion

The individual preimplementation phase projects in the proposal are a series of innovations to improve access to health care by overcoming barriers that have led to unmet care needs, delivery of non–evidence-based care, and subsequent health care disparity for V/SMs with TBI-related morbidity (cognition, behavior, and psychological functioning). Through HCD, engagement, and iterative evaluation, clear success measures will inform project continuation (Multimedia Appendix 2). Subsequent protocols should focus on the next stages of the implementation and sustainability of the innovations.

Acknowledgments

This material is the result of work supported with resources and the use of facilities at the James A. Haley Veterans' Hospital. The administering institution for this work is the Tampa VA Research and Education Foundation. The views expressed in this manuscript are those of the authors and do not necessarily represent the official policy or position of the Department of Veterans Affairs or any other United States government agency.

Artificial intelligence was not used in any part of the preparation of this manuscript.

Chaady Radwan and Margeaux Chavez were part of the initial proposal.

The I-HEAL Consortium includes Sonia Arellano, Deveney Ching, John Corrigan, Bridget A Cotner, Cassandra Decker, Natalie Gilmore, Daniel Goldschmidt, Leslie Kempthorne, Jessica Ketchum, Dave Mellick, Leah Phillips, Jessica L Ryan, and Amanda Tweed.

Funding

This work is funded by the Department of Defense Congressionally Directed Medical Research Programs Award Number: HT9425-23-1-0621 for Improving Health Care Access and Engagement for Veterans and Service Members with TBI Morbidity (I-HEAL) (IRBNet #: 1804543-1).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Impact and relevance to military health.
[PDF File, 194 KB - resprot_v15i1e79738_app1.pdf]

Multimedia Appendix 2
Statement of work.
[PDF File, 3049 KB - resprot_v15i1e79738_app2.pdf]

Multimedia Appendix 3
Transition plan.
[DOCX File, 20 KB - resprot_v15i1e79738_app3.docx]

Multimedia Appendix 4
Community-based participatory research approach.
[DOCX File, 20 KB - resprot_v15i1e79738_app4.docx]

Multimedia Appendix 5
Community Engagement Council schedule.
[PDF File, 123 KB - resprot_v15i1e79738_app5.pdf]

Multimedia Appendix 6
Communication strategy.
[PDF File, 112 KB - resprot_v15i1e79738_app6.pdf]

Multimedia Appendix 7
Leadership.
[DOCX File, 492 KB - resprot_v15i1e79738_app7.docx]

Multimedia Appendix 8
Cognitive nudge toolkit development.
[PDF File, 96 KB - resprot_v15i1e79738_app8.pdf]

Multimedia Appendix 9
Individual engagement partner areas of subject matter expertise.
[PDF File, 101 KB - resprot_v15i1e79738_app9.pdf]

Peer Review Report 1
Peer review report from the U.S. Army Medical Research and Development Command Congressionally Directed Medical Research Programs, Fiscal Year 2022 Traumatic Brain Injury and Psychological Health Research Program, Focused Program - Healthcare Access Review Panel (Department of Defense, USA).
[PDF File, 322 KB - resprot_v15i1e79738_app10.pdf]

References

1. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Health Care Services, Board on Health Sciences Policy, Committee on Accelerating Progress in Traumatic Brain Injury Research and Care. Traumatic Brain Injury: A Roadmap for Accelerating Progress: National Academies Press; 2022.
2. Greenwald BD, Hammond FM, Harrison-Felix C, Nakase-Richardson R, Howe LLS, Kreider S. Mortality following traumatic brain injury among individuals unable to follow commands at the time of rehabilitation admission: a National Institute on Disability and Rehabilitation Research traumatic brain injury model systems study. J Neurotrauma 2015 Dec 1;32(23):1883-1892. [doi: 10.1089/neu.2014.3454] [Medline: 25518731]
3. Tran J, Hammond F, Dams-O Connor K, et al. Rehospitalization in the first year following veteran and service member TBI: a VA TBI Model Systems study. J Head Trauma Rehabil 2017;32(4):264-270. [doi: 10.1097/HTR.0000000000000296] [Medline: 28195958]

XSL•FO
RenderX

4. Steward KA, Silva MA, Maduri P, et al. Obstructive sleep apnea is associated with worse cognitive outcomes in acute moderate-to-severe traumatic brain injury: a TBI model systems study. Sleep Med 2022 Dec;100:454-461. [doi: 10.1016/j.sleep.2022.09.012] [Medline: 36252414]
5. Byers AL, Li Y, Barnes DE, Boscardin WJ, Peltz CB, Yaffe K. TBI and risk of death in military veterans over 14 years: injury severity, timing, and cause of death. J Psychiatr Res 2022 Dec;156:200-205. [doi: 10.1016/j.jpsychires.2022.09.035] [Medline: 36257114]
6. Cotner BA, Nakase-Richardson R, O'Connor DR, et al. Barriers and facilitators to accessing rehabilitation health care: a Veterans Affairs traumatic brain injury model systems qualitative study. Arch Phys Med Rehabil 2023 Mar;104(3):380-389. [doi: 10.1016/j.apmr.2022.09.020] [Medline: 36265532]
7. Silva MA, VandenBussche Jantz AB, Klocksieben F, et al. Unmet rehabilitation needs indirectly influence life satisfaction 5 years after traumatic brain injury: a Veterans Affairs TBI Model Systems study. Arch Phys Med Rehabil 2021 Jan;102(1):58-67. [doi: 10.1016/j.apmr.2020.08.012] [Medline: 32949552]
8. Noyes ET, Tang X, Sander AM, et al. Relationship of medical comorbidities to psychological health at 2 and 5 years following traumatic brain injury (TBI). Rehabil Psychol 2021 May;66(2):107-117. [doi: 10.1037/rep0000366] [Medline: 33119380]
9. Goodrich DE, Miake-Lye I, Braganza MZ, Wawrin N, Kilbourne AM. The QUERI Roadmap for Implementation and Quality Improvement: Department of Veterans Affairs; 2020. URL: https://www.ncbi.nlm.nih.gov/books/NBK566223/ [accessed 2026-02-10]
10. Levesque JF, Harris MF, Russell G. Patient-centred access to health care: conceptualising access at the interface of health systems and populations. Int J Equity Health 2013 Mar 11;12:18. [doi: 10.1186/1475-9276-12-18] [Medline: 23496984]
11. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Committee on Community-Based Solutions to Promote Health Equity in the United States. Communities in Action: Pathways to Health Equity: National Academies Press; 2017. URL: https://www.ncbi.nlm.nih.gov/books/NBK425848/ [accessed 2026-02-10]
12. Pugh MJV, Finley EP, Copeland LA, et al. Complex comorbidity clusters in OEF/OIF veterans: the polytrauma clinical triad and beyond. Med Care 2014 Feb;52(2):172-181. [doi: 10.1097/MLR.0000000000000059] [Medline: 24374417]
13. Giacino JT, Katz DI, Schiff ND, et al. Practice guideline update recommendations summary: disorders of consciousness: report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and the National Institute on Disability, Independent Living, and Rehabilitation Research. Neurology 2018 Sep 4;91(10):450-460. [doi: 10.1212/WNL.0000000000005926] [Medline: 30089618]
14. The Field Guide to Human-Centered Design: IDEO.org; 2015.
15. Rosinsky K, Murray DW, Nagle K, et al. A review of human-centered design in human services, OPRE report 2022-78. : Office of Planning, Research, and Evaluation (OPRE); 2022 URL: https://www.acf.hhs.gov/sites/default/files/documents/opre/OPRE-HCD-State-of-Field.pdf [accessed 2026-02-10]
16. Banbury A, Pedell S, Parkinson L, Byrne L. Using the Double Diamond model to co-design a dementia caregivers telehealth peer support program. J Telemed Telecare 2021 Dec;27(10):667-673. [doi: 10.1177/1357633X211048980] [Medline: 34726994]
17. Innovating for People: Handbook of Human-Centered Design Methods: Luma Institute; 2012. URL: https://www.luma-institute.com/wp-content/uploads/2021/02/22Innovating-for-People22-Looking.pdf [accessed 2026-02-10]
18. Nakase-Richardson R, Haun JN, Hoffman JM, et al. Building capacity for selecting interventions in implementation research grants: TBIMS chronic pain collaborative project. Presented at: 99th Annual Conference of the American Congress of Rehabilitation Medicine; Nov 8-11, 2022 URL: https://www.eventscribe.net/2022/ACRM/fsPopup.asp?PresentationID=1108485&mode=presinfo [accessed 2026-02-10]
19. Ritchie J, Dollar KM, Miller CJ, et al. Using Implementation Facilitation to Improve Health Care (Version 3): Veterans Health Administration, Behavioral Health Quality Enhancement Research Initiative (QUERI); 2020. URL: https://www.queri.research.va.gov/tools/Facilitation-Manual.pdf [accessed 2022-11-27]
20. Ritchie MJ, Dollar KM, Miller CJ, et al. Using Implementation Facilitation to Improve Healthcare (Version 3.1). Veterans Health Administration, Behavioral HealthQuality Enhancement Research Initiative (QUERI). 2020. URL: https://www.queri.research.va.gov/tools/Facilitation-Manual.pdf [accessed 2026-02-27]
21. Wallerstein N, Oetzel JG, Sanchez-Youngman S, et al. Engage for equity: a long-term study of community-based participatory research and community-engaged research practices and outcomes. Health Educ Behav 2020 Jun;47(3):380-390. [doi: 10.1177/1090198119897075] [Medline: 32437293]
22. Wallerstein NB, Duran B. Using community-based participatory research to address health disparities. Health Promot Pract 2006 Jul;7(3):312-323. [doi: 10.1177/1524839906289376] [Medline: 16760238]
23. I-HEAL. URL: https://ihealbrain.org/ [accessed 2026-02-27]
24. Tufte ER. The Visual Display of Qualitative Information, 2nd edition: Graphics Press; 2001.

XSL•FO
RenderX

25. Jourdan C, Azouvi P, Genêt F, Selly N, Josseran L, Schnitzler A. Disability and health consequences of traumatic brain injury: national prevalence. Am J Phys Med Rehabil 2018 May;97(5):323-331. [doi: 10.1097/PHM.0000000000000848] [Medline: 29016402]
26. Taylor CA, Bell JM, Breiding MJ, Xu L. Traumatic brain injury-related emergency department visits, hospitalizations, and deaths - United States, 2007 and 2013. MMWR Surveill Summ 2017 Mar 17;66(9):1-16. [doi: 10.15585/mmwr.ss6609a1] [Medline: 28301451]
27. Traumatic Brain Injury Center of Excellence: DoD TBI worldwide numbers. Military Health System. URL: https://health.mil/Military-Health-Topics/Centers-of-Excellence/Traumatic-Brain-Injury-Center-of-Excellence/DOD-TBI-Worldwide-Numbers [accessed 2022-11-29]
28. Cristofori I, Levin HS. Traumatic brain injury and cognition. Handb Clin Neurol 2015;128:579-611. [doi: 10.1016/B978-0-444-63521-1.00037-6] [Medline: 25701909]
29. Bracho-Ponce MJ, Grasso-Cladera A, Salas C, Lassaletta A. 'I am not who you think I am': public stigma and invisible brain injury ('No soy quien creen que soy': estigma público y daño cerebral invisible). Studies in Psychology 2022 Jan 2;43(1):35-54. [doi: 10.1080/02109395.2021.2006947]
30. Mahoney EJ, Silva MA, Reljic T, et al. Rehabilitation needs at 5 years post-traumatic brain injury: a VA TBI Model Systems study. J Head Trauma Rehabil 2021;36(3):175-185. [doi: 10.1097/HTR.0000000000000629] [Medline: 33201036]
31. Silva MA, Finn JA, Dillahunt-Aspillaga C, Cotner BA, Stevens LF, Nakase-Richardson R. Development of the traumatic brain injury rehabilitation needs survey: a Veterans Affairs TBI Model Systems study. Disabil Rehabil 2022 Aug;44(16):4474-4484. [doi: 10.1080/09638288.2021.1900930] [Medline: 33756089]
32. Whiteneck GG, Eagle CB, Cuthbert JP, et al. One and five year outcomes after moderate-to-severe traumatic brain injury requiring inpatient rehabilitation: traumatic brain injury report.: Centers for Disease Control and Prevention; 2018. URL: https://stacks.cdc.gov/view/cdc/59524 [accessed 2026-02-10]
33. Finn JA, Klocksieben FA, Smith AN, et al. Family needs after traumatic brain injury: a VA TBI Model Systems study. J Head Trauma Rehabil 2022;37(6):327-337. [doi: 10.1097/HTR.0000000000000799] [Medline: 35687889]
34. Fisher A, Bellon M, Lawn S, Lennon S. Brain injury, behaviour support, and family involvement: putting the pieces together and looking forward. Disabil Rehabil 2020 May;42(9):1305-1315. [doi: 10.1080/09638288.2018.1522551] [Medline: 30653928]
35. Bayley M, Swaine B, Lamontagne ME, et al. INESSS-ONF clinical practice guideline for the rehabilitation of adults with moderate to severe traumatic brain injury. Ontario Neurotrauma Foundation. 2016. URL: https://kite-uhn.com/brain-injury/en [accessed 2022-11-29]
36. Turner-Stokes L, editor. Rehabilitation Following Acquired Brain Injury: National Clinical Guidelines: Royal College of Physicians, British Society of Rehabilitation Medicine; 2003. URL: https://www.headway.org.uk/media/3320/bsrm-rehabilitation-following-acquired-brain-injury.pdf [accessed 2026-02-10]
37. Demakis GJ. Civil Capacities in Clinical Neuropsychology: Research Findings and Practical Applications: Oxford University Press; 2012.
38. Suarez-Balcazar Y, Francisco VT, Rubén Chávez N. Applying community-based participatory approaches to addressing health disparities and promoting health equity. Am J Community Psychol 2020 Dec;66(3-4):217-221. [doi: 10.1002/ajcp.12487] [Medline: 33373469]
39. Borgen IMH, Kleffelgård I, Hauger SL, et al. Patient-reported problem areas in chronic traumatic brain injury. J Head Trauma Rehabil 2022;37(5):E336-E345. [doi: 10.1097/HTR.0000000000000744] [Medline: 34743086]
40. Masel BE. The chronic consequences of neurotrauma. J Neurotrauma 2015 Dec 1;32(23):1833. [doi: 10.1089/neu.2015.29004.bm] [Medline: 26629808]
41. Harrison-Felix C, Pretz C, Hammond FM, et al. Life expectancy after inpatient rehabilitation for taumatic brain injury in the United States. J Neurotrauma 2015 Dec 1;32(23):1893-1901. [doi: 10.1089/neu.2014.3353] [Medline: 25057965]
42. Dams-O'Connor K, Mellick D, Dreer LE, et al. Rehospitalization over 10 years among survivors of TBI: a National Institute on Disability, Independent Living, and Rehabilitation Research traumatic brain injury model systems study. J Head Trauma Rehabil 2017;32(3):147-157. [doi: 10.1097/HTR.0000000000000263]
43. Lercher K, Kumar RG, Hammond FM, et al. Distal and proximal predictors of rehospitalization over 10 years among survivors of TBI: a National Institute on Disability, Independent Living, and Rehabilitation Research traumatic brain injury model systems study. J Head Trauma Rehabil 2023;38(3):203-213. [doi: 10.1097/HTR.0000000000000812] [Medline: 36102607]
44. Kumar RG, Ketchum JM, Hammond FM, et al. Health and cognition among adults with and without traumatic brain injury: a matched case-control study. Brain Inj 2022 Feb 23;36(3):415-423. [doi: 10.1080/02699052.2022.2034190] [Medline: 35143349]
45. Dams-O'Connor K, Landau A, Hoffman J, St De Lore J. Patient perspectives on quality and access to healthcare after brain injury. Brain Inj 2018;32(4):431-441. [doi: 10.1080/02699052.2018.1429024] [Medline: 29388840]
46. Ching D, O'Connor D, Martin A, et al. Organizational facilitators and barriers to providing treatment to persons with TBI and chronic pain. Arch Phys Med Rehabil 2023 Mar;104(3):e41. [doi: 10.1016/j.apmr.2022.12.118]

XSL•FO
RenderX

47. Nakase-Richardson R, Cotner BA, Martin AM, et al. Provider perspectives of facilitators and barriers to reaching and utilizing chronic pain healthcare for persons with traumatic brain injury: a qualitative NIDILRR and VA TBI model systems collaborative project. J Head Trauma Rehabil 2024;39(1):E15-E28. [doi: 10.1097/HTR.0000000000000923] [Medline: 38167719]
48. Douglas J, Bigby C. Development of an evidence-based practice framework to guide decision making support for people with cognitive impairment due to acquired brain injury or intellectual disability. Disabil Rehabil 2020 Feb;42(3):434-441. [doi: 10.1080/09638288.2018.1498546] [Medline: 30451022]
49. Wolff JL, Roter DL. Family presence in routine medical visits: a meta-analytical review. Soc Sci Med 2011 Mar;72(6):823-831. [doi: 10.1016/j.socscimed.2011.01.015]
50. Berg JM, Malte CA, Reger MA, Hawkins EJ. Medical records flag for suicide risk: predictors and subsequent use of care among veterans with substance use disorders. Psychiatr Serv 2018 Sep 1;69(9):993-1000. [doi: 10.1176/appi.ps.201700545] [Medline: 29879873]
51. Baillie CA, VanZandbergen C, Tait G, et al. The readmission risk flag: using the electronic health record to automatically identify patients at risk for 30-day readmission. J Hosp Med 2013 Dec;8(12):689-695. [doi: 10.1002/jhm.2106] [Medline: 24227707]
52. Paterson J, Fernandes J, Hunter K, Hubert S, Strudwick G. Embedding psychiatric risk flags within an electronic health record: initial findings and lessons learned. Healthc Q 2019 Jan;21(4):54-60. [doi: 10.12927/hcq.2019.25739] [Medline: 30946656]
53. 2010 - 2014 ACS 5-year data profile. United States Census Bureau - American Community Survey (ACS). 2014. URL: https://www.census.gov/acs/www/data/data-tables-and-tools/data-profiles/2014/ [accessed 2026-02-10]
54. Hahn GJ, Meeker WQ. Statistical Intervals: A Guide for Practitioners: John Wiley & Sons; 1991. [doi: 10.1002/9780470316771]
55. Fleiss JL, Levin B, Paik MC. Statistical Methods for Rates and Proportions, 3rd edition: John Wiley & Sons; 2003. [doi: 10.1002/0471445428]
56. Newcombe RG. Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med 1998 Apr 30;17(8):857-872. [doi: 10.1002/(sici)1097-0258(19980430)17:8<857::aid-sim777>3.0.co;2-e] [Medline: 9595616]
57. Dillman DA, Smyth JD, Christian LM. Internet, Phone, Mail, and Mixed‐Mode Surveys: The Tailored Design Method: John Wiley & Sons; 2009. [doi: 10.1002/9781394260645]
58. Kretzmer T, Pinto S, Agtarap S, et al. Impact of problematic behaviors on patient and healthcare system outcomes: a survey of brain-injury providers. Arch Phys Med Rehabil 2023 Mar;104(3):e34. [doi: 10.1016/j.apmr.2022.12.096]
59. Hennink MM, Kaiser BN, Weber MB. What influences saturation? Estimating sample sizes in focus group research. Qual Health Res 2019 Aug;29(10):1483-1496. [doi: 10.1177/1049732318821692] [Medline: 30628545]
60. King DK, Shoup JA, Raebel MA, et al. Planning for implementation success using RE-AIM and CFIR frameworks: a qualitative study. Front Public Health 2020;8:59. [doi: 10.3389/fpubh.2020.00059] [Medline: 32195217]
61. Ober AJ, Takada S, Zajdman D, et al. Factors affecting statin uptake among people living with HIV: primary care provider perspectives. BMC Fam Pract 2021 Oct 30;22(1):215. [doi: 10.1186/s12875-021-01563-0] [Medline: 34717560]
62. Beebe J. Rapid Assessment Process: An Introduction: AltaMira Press; 2001.
63. Morgan J, Schwartz C, Ferlatte O, et al. Community-based participatory approaches to knowledge translation: HIV prevention case study of the investigaytors program. Arch Sex Behav 2021 Jan;50(1):105-117. [doi: 10.1007/s10508-020-01789-6] [Medline: 32737658]
64. Greenhalgh T, Jackson C, Shaw S, Janamian T. Achieving research impact through co-creation in community-based health services: literature review and case study. Milbank Q 2016 Jun;94(2):392-429. [doi: 10.1111/1468-0009.12197] [Medline: 27265562]
65. Garcia A, Kretzmer TS, Dams-O'Connor K, et al. Health conditions among special operations forces versus conventional military service members: a VA TBI Model Systems study. J Head Trauma Rehabil 2022;37(4):E292-E298. [doi: 10.1097/HTR.0000000000000737] [Medline: 34698680]
66. Silva MA, Gonzalez AV, Tang X, et al. Examining the relationship between sleep apnea diagnosis and suicide risk in veterans with traumatic brain injury: a VA TBI Model Systems study. J Head Trauma Rehabil 2023;38(5):359-367. [doi: 10.1097/HTR.0000000000000856] [Medline: 36727802]
67. Miles SR, Silva MA, Lang B, et al. Sleep apnea and posttraumatic stress after traumatic brain injury (TBI): a Veterans Affairs TBI Model Systems study. Rehabil Psychol 2021 Nov;66(4):450-460. [doi: 10.1037/rep0000389] [Medline: 34871026]
68. Miles SR, Silva MA, Neumann D, et al. Demographic and mental health predictors of arrests up to 10 years post-traumatic brain injury: a Veterans Affairs TBI Model Systems study. J Head Trauma Rehabil 2021;36(4):E240-E248. [doi: 10.1097/HTR.0000000000000640] [Medline: 33528175]
69. Silva MA, Lee JM, Garcia A, Dams-O'Connor K, Nakase-Richardson R. Research letter: impact of obstructive sleep apnea disease duration on neuropsychological functioning after traumatic brain injury: a Veterans Affairs TBI Model Systems study. J Head Trauma Rehabil 2022;37(6):E496-E501. [doi: 10.1097/HTR.0000000000000797] [Medline: 35687890]

XSL•FO
RenderX

70. Garcia A, Reljic T, Pogoda TK, et al. Obstructive sleep apnea risk is associated with cognitive impairment after controlling for mild traumatic brain injury history: a Chronic Effects of Neurotrauma Consortium study. J Neurotrauma 2020 Dec 1;37(23):2517-2527. [doi: 10.1089/neu.2019.6916] [Medline: 32709212]
71. Bell KR, Bushnik T, Dams-O'Connor K, et al. Sleep after TBI: how the TBI Model Systems have advanced the field. NeuroRehabilitation 2018;43(3):287-296. [doi: 10.3233/NRE-182538] [Medline: 30347631]
72. Holcomb EM, Schwartz DJ, McCarthy M, Thomas B, Barnett SD, Nakase-Richardson R. Incidence, characterization, and predictors of sleep apnea in consecutive brain injury rehabilitation admissions. J Head Trauma Rehabil 2016;31(2):82-100. [doi: 10.1097/HTR.0000000000000230] [Medline: 26959663]
73. Nakase-Richardson R, Schwartz DJ, Drasher-Phillips L, et al. Comparative effectiveness of sleep apnea screening instruments during inpatient rehabilitation following moderate to severe TBI. Arch Phys Med Rehabil 2020 Feb;101(2):283-296. [doi: 10.1016/j.apmr.2019.09.019] [Medline: 31705855]
74. Traumatic Brain Injury Model Systems National Data and Statistical Center. URL: https://pain.tbindsc.org/ [accessed 2026-02-10]
75. Fisher LB, Pedrelli P, Iverson GL, et al. Prevalence of suicidal behaviour following traumatic brain injury: longitudinal follow-up data from the NIDRR Traumatic Brain Injury Model Systems. Brain Inj 2016;30(11):1311-1318. [doi: 10.1080/02699052.2016.1195517] [Medline: 27541868]
76. Lew HL, Pogoda TK, Hsu PT, et al. Impact of the "polytrauma clinical triad" on sleep disturbance in a Department of Veterans Affairs outpatient rehabilitation setting. Am J Phys Med Rehabil 2010 Jun;89(6):437-445. [doi: 10.1097/PHM.0b013e3181ddd301] [Medline: 20489391]
77. Cifu DX, Taylor BC, Carne WF, et al. Traumatic brain injury, posttraumatic stress disorder, and pain diagnoses in OIF/OEF/OND veterans. J Rehabil Res Dev 2013;50(9):1169-1176. [doi: 10.1682/JRRD.2013.01.0006] [Medline: 24458958]
78. Ord AS, Lad SS, Shura RD, Rowland JA, Taber KH, Martindale SL. Pain interference and quality of life in combat veterans: examining the roles of posttraumatic stress disorder, traumatic brain injury, and sleep quality. Rehabil Psychol 2021 Feb;66(1):31-38. [doi: 10.1037/rep0000333] [Medline: 32378923]
79. Balba NM, Elliott JE, Weymann KB, et al. Increased sleep disturbances and pain in veterans with comorbid traumatic brain injury and posttraumatic stress disorder. J Clin Sleep Med 2018 Nov 15;14(11):1865-1878. [doi: 10.5664/jcsm.7482] [Medline: 30373686]
80. Lumley MA, Yamin JB, Pester BD, Krohner S, Urbanik CP. Trauma matters: psychological interventions for comorbid psychosocial trauma and chronic pain. Pain 2022 Apr 1;163(4):599-603. [doi: 10.1097/j.pain.0000000000002425] [Medline: 34338244]
81. Powell MA, Corbo V, Fonda JR, Otis JD, Milberg WP, McGlinchey RE. Sleep quality and reexperiencing symptoms of PTSD are associated with current pain in U.S. OEF/OIF/OND veterans with and without mTBIs. J Trauma Stress 2015 Aug;28(4):322-329. [doi: 10.1002/jts.22027] [Medline: 26194844]
82. Aase DM, Babione JM, Proescher E, et al. Impact of PTSD on post-concussive symptoms, neuropsychological functioning, and pain in post-9/11 veterans with mild traumatic brain injury. Psychiatry Res 2018 Oct;268:460-466. [doi: 10.1016/j.psychres.2018.08.019] [Medline: 30138858]
83. Blakey SM, Wagner HR, Naylor J, et al. Chronic pain, TBI, and PTSD in military veterans: a link to suicidal ideation and violent impulses? J Pain 2018 Jul;19(7):797-806. [doi: 10.1016/j.jpain.2018.02.012] [Medline: 29526669]
84. McIntire KL, Crawford KM, Perrin PB, et al. Factors increasing risk of suicide after traumatic brain injury: a state-of-the-science review of military and civilian studies. Brain Inj 2021 Jan 18;35(2):151-163. [doi: 10.1080/02699052.2020.1861656] [Medline: 33460350]
85. American Psychological Association. Guidelines for assessment of and intervention with persons with disabilities. American Psychologist 2022;67(1):43-62. [doi: 10.1037/a0025892]
86. Gallagher M, McLeod HJ, McMillan TM. A systematic review of recommended modifications of CBT for people with cognitive impairments following brain injury. Neuropsychol Rehabil 2019 Jan;29(1):1-21. [doi: 10.1080/09602011.2016.1258367] [Medline: 27873549]
87. Whiting D, Deane F, McLeod H, Ciarrochi J, Simpson G. Can acceptance and commitment therapy facilitate psychological adjustment after a severe traumatic brain injury? A pilot randomized controlled trial. Neuropsychol Rehabil 2020 Aug;30(7):1348-1371. [doi: 10.1080/09602011.2019.1583582] [Medline: 30789059]
88. Cameron E, Oliver MA, Curvis W. Acceptance and commitment therapy for people with moderate or severe brain injuries. In: Curvis W, Methley A, editors. Acceptance and Commitment Therapy and Brain Injury: Routledge; 2021:53-69. [doi: 10.4324/9781003024408-8]
89. Hoffman JM, Ehde DM, Dikmen S, et al. Telephone-delivered cognitive behavioral therapy for veterans with chronic pain following traumatic brain injury: rationale and study protocol for a randomized controlled trial study. Contemp Clin Trials 2019 Jan;76:112-119. [doi: 10.1016/j.cct.2018.12.004] [Medline: 30553077]
90. Chard KM, Schumm JA, McIlvain SM, Bailey GW, Parkinson RB. Exploring the efficacy of a residential treatment program incorporating cognitive processing therapy-cognitive for veterans with PTSD and traumatic brain injury. J Trauma Stress 2011 Jun;24(3):347-351. [doi: 10.1002/jts.20644] [Medline: 21626573]

XSL•FO
RenderX

91. Ragsdale KA, Gramlich MA, Beidel DC, Neer SM, Kitsmiller EG, Morrison KI. Does traumatic brain injury attenuate the exposure therapy process? Behav Ther 2018 Jul;49(4):617-630. [doi: 10.1016/j.beth.2017.09.008] [Medline: 29937262]
92. Wolf GK, Strom TQ, Kehle SM, Eftekhari A. A preliminary examination of prolonged exposure therapy with Iraq and Afghanistan veterans with a diagnosis of posttraumatic stress disorder and mild to moderate traumatic brain injury. J Head Trauma Rehabil 2012;27(1):26-32. [doi: 10.1097/HTR.0b013e31823cd01f] [Medline: 22218201]
93. Johnston-Brooks CH, Miles SR, Brostow DP. Mental health pocket card for management of patients with posttraumatic stress disorder and mild traumatic brain injury. Arch Phys Med Rehabil 2022 Mar;103(3):611-615. [doi: 10.1016/j.apmr.2021.10.013] [Medline: 34922737]
94. Almklov EA, Rivera GL, Orff H. CBT-i in patients with a history of traumatic brain injury. In: Nowakowski S, Garland SN, Grandner MA, Cuddihy LJ, editors. Adapting Cognitive Behavioral Therapy for Insomnia: Academic Press; 2022:265-290. [doi: 10.1016/B978-0-12-822872-2.00009-8]
95. Mysliwiec V, Martin JL, Ulmer CS, et al. The management of chronic insomnia disorder and obstructive sleep apnea: synopsis of the 2019 U.S. Department of Veterans Affairs and U.S. Department of Defense clinical practice guidelines. Ann Intern Med 2020 Mar 3;172(5):325-336. [doi: 10.7326/M19-3575] [Medline: 32066145]
96. Stetler CB, Legro MW, Wallace CM, et al. The role of formative evaluation in implementation research and the QUERI experience. J Gen Intern Med 2006 Feb;21 Suppl 2(Suppl 2):S1-S8. [doi: 10.1111/j.1525-1497.2006.00355.x] [Medline: 16637954]
97. McNall M, Foster-Fishman PG. Methods of rapid evaluation, assessment, and appraisal. American Journal of Evaluation 2007 Jun;28(2):151-168. [doi: 10.1177/1098214007300895]
98. Stéfan A, Mathé JF, SOFMER group. What are the disruptive symptoms of behavioral disorders after traumatic brain injury? A systematic review leading to recommendations for good practices. Ann Phys Rehabil Med 2016 Feb;59(1):5-17. [doi: 10.1016/j.rehab.2015.11.002] [Medline: 26768944]
99. Nakase-Richardson R, Kretzmer T, Agtarap S, Pinto SM, Reindeau A, Shapiro-Rosenbaum A. Program elements that promote successful management of brain-injury related behaviors: qualitative survey of providers. Arch Phys Med Rehabil 2023 Mar;104(3):e45-e46. [doi: 10.1016/j.apmr.2022.12.131]
100. Carrier SL, Hicks AJ, Ponsford J, McKay A. Managing agitation during early recovery in adults with traumatic brain injury: an international survey. Ann Phys Rehabil Med 2021 Sep;64(5):101532. [doi: 10.1016/j.rehab.2021.101532] [Medline: 33933690]
101. Bogner J, Barrett RS, Hammond FM, et al. Predictors of agitated behavior during inpatient rehabilitation for traumatic brain injury. Arch Phys Med Rehabil 2015 Aug;96(8 Suppl):S274-S281. [doi: 10.1016/j.apmr.2015.04.020] [Medline: 26212403]
102. Karlin BE, Teri L, McGee JS, et al. STAR-VA Intervention for Managing Challenging Behaviorsin VA Community Living Center Residents with Dementia: Manual for STAR-VA BehavioralCoordinators and Nurse Champions: Department of Veterans Affairs; 2017. URL: https://www.mentalhealth.va.gov/healthcare-providers/docs/STAR-VA_Manual.pdf [accessed 2026-02-10]
103. Ylvisaker M, Jacobs HE, Feeney T. Positive supports for people who experience behavioral and cognitive disability after brain injury: a review. J Head Trauma Rehabil 2003;18(1):7-32. [doi: 10.1097/00001199-200301000-00005] [Medline: 12802235]
104. Ylvisaker M, Turkstra L, Coehlo C, et al. Behavioural interventions for children and adults with behaviour disorders after TBI: a systematic review of the evidence. Brain Inj 2007 Jul;21(8):769-805. [doi: 10.1080/02699050701482470] [Medline: 17676437]
105. Lane-Brown AT, Tate RL. Measuring apathy after traumatic brain injury: psychometric properties of the Apathy Evaluation Scale and the Frontal Systems Behavior Scale. Brain Inj 2009 Dec;23(13-14):999-1007. [doi: 10.3109/02699050903379347] [Medline: 19891539]
106. Votruba KL, Rapport LJ, Vangel SJ, et al. Impulsivity and traumatic brain injury: the relations among behavioral observation, performance measures, and rating scales. J Head Trauma Rehabil 2008;23(2):65-73. [doi: 10.1097/01.HTR.0000314525.93381.69] [Medline: 18362760]
107. Steward KA, Kretzmer T. Anosognosia in moderate-to-severe traumatic brain injury: a review of prevalence, clinical correlates, and diversity considerations. Clin Neuropsychol 2022 Nov;36(8):2021-2040. [doi: 10.1080/13854046.2021.1967452] [Medline: 34429014]
108. Bourseau T, Saoût V, Ali P, Pouliquen-Loriferne U, Allain P, Dinomais M. A cross-sectional study of treatments for behavioral disorders following traumatic brain injury: comparison with French recommendations. J Head Trauma Rehabil 2019;34(4):E51-E60. [doi: 10.1097/HTR.0000000000000452] [Medline: 30499933]
109. Donker-Cools B, Daams JG, Wind H, Frings-Dresen MHW. Effective return-to-work interventions after acquired brain injury: a systematic review. Brain Inj 2016;30(2):113-131. [doi: 10.3109/02699052.2015.1090014] [Medline: 26645137]
110. Guevara AB, Demonet JF, Polejaeva E, et al. Association between traumatic brain injury-related brain lesions and long-term caregiver burden. J Head Trauma Rehabil 2016;31(2):E48-E58. [doi: 10.1097/HTR.0000000000000151] [Medline: 26098258]

XSL•FO
RenderX

111. Lequerica AH, Rapport LJ, Loeher K, Axelrod BN, Vangel SJ, Hanks RA. Agitation in acquired brain injury: impact on acute rehabilitation therapies. J Head Trauma Rehabil 2007;22(3):177-183. [doi: 10.1097/01.HTR.0000271118.96780.bc] [Medline: 17510593]
112. Bogner JA, Corrigan JD, Fugate L, Mysiw WJ, Clinchot D. Role of agitation in prediction of outcomes after traumatic brain injury. Am J Phys Med Rehabil 2001 Sep;80(9):636-644. [doi: 10.1097/00002060-200109000-00002] [Medline: 11523965]
113. Tam S, McKay A, Sloan S, Ponsford J. The experience of challenging behaviours following severe TBI: a family perspective. Brain Inj 2015;29(7-8):813-821. [doi: 10.3109/02699052.2015.1005134] [Medline: 25914927]
114. Osborne-Crowley K, McDonald S. A review of social disinhibition after traumatic brain injury. J Neuropsychol 2018 Jun;12(2):176-199. [doi: 10.1111/jnp.12113] [Medline: 27696753]
115. Beaulieu CL, Persel C, Shannon T, et al. Examining the evidence from single-case experimental designs to treat challenging behaviors following moderate to severe traumatic brain injury. J Head Trauma Rehabil 2023;38(2):E126-E135. [doi: 10.1097/HTR.0000000000000795] [Medline: 35687891]
116. Lee A, Gage NA. Updating and expanding systematic reviews and meta‐analyses on the effects of school‐wide positive behavior interventions and supports. Psychol Sch 2020 May;57(5):783-804. [doi: 10.1002/pits.22336]
117. Heinicke MR, Carr JE. Applied behavior analysis in acquired brain injury rehabilitation: a meta-analysis of single-case design intervention research. Behav Interv 2014 Apr;29(2):77-105 [FREE Full text] [doi: 10.1002/bin.1380]
118. Cattelani R, Zettin M, Zoccolotti P. Erratum to: rehabilitation treatments for adults with behavioral and psychosocial disorders following acquired brain injury: a systematic review. Neuropsychol Rev 2011 Jun;21(2):224-224. [doi: 10.1007/s11065-011-9161-2]
119. Cox AD, Pontone M, Gunnarsson KF. Program evaluation of in-patient treatment units for adults with acquired brain injury and challenging behavior. Brain Inj 2022 Feb 23;36(3):321-331. [doi: 10.1080/02699052.2022.2033840] [Medline: 35125040]
120. Carrier SL, Ponsford J, McKay A. Managing agitation during early recovery following traumatic brain injury: qualitative interviews with clinicians. Disabil Rehabil 2023 Nov;45(22):3686-3695. [doi: 10.1080/09638288.2022.2135771] [Medline: 36282105]
121. Williamson D, Frenette AJ, Burry LD, et al. Pharmacological interventions for agitated behaviours in patients with traumatic brain injury: a systematic review. BMJ Open 2019 Jul;9(7):e029604. [doi: 10.1136/bmjopen-2019-029604]
122. Rahmani E, Lemelle TM, Samarbafzadeh E, Kablinger AS. Pharmacological treatment of agitation and/or aggression in patients with traumatic brain injury: a systematic review of reviews. J Head Trauma Rehabil 2021;36(4):E262-E283. [doi: 10.1097/HTR.0000000000000656] [Medline: 33656478]
123. Nakase-Richardson R, Evans CC. Behavioral assessment of acute neurobehavioral syndromes to inform treatment. In: Sherer M, Sander AM, editors. Handbook on the Neuropsychology of Traumatic Brain Injury: Springer; 2014:157-172. [doi: 10.1007/978-1-4939-0784-7_8]
124. Teri L, Huda P, Gibbons L, Young H, van Leynseele J. STAR: a dementia-specific training program for staff in assisted living residences. Gerontologist 2005 Oct;45(5):686-693. [doi: 10.1093/geront/45.5.686] [Medline: 16199404]
125. Curyto KJ, McCurry SM, Luci K, Karlin BE, Teri L, Karel MJ. Managing challenging behaviors of dementia in veterans: identifying and changing activators and consequences using STAR-VA. J Gerontol Nurs 2017 Feb 1;43(2):33-43. [doi: 10.3928/00989134-20160930-01] [Medline: 27711931]
126. Karel MJ, Teri L, McConnell E, Visnic S, Karlin BE. Effectiveness of expanded implementation of STAR-VA for managing dementia-related behaviors among veterans. Gerontologist 2016 Feb;56(1):126-134. [doi: 10.1093/geront/gnv068] [Medline: 26185155]
127. Kreuger RA, Casey MA. Focus Groups: A Practical Guide for Applied Research, 5th edition: Sage Publications, Inc; 2014.
128. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
129. Morse JM. "Data were saturated . . . ". Qual Health Res 2015 May;25(5):587-588. [doi: 10.1177/1049732315576699] [Medline: 25829508]
130. Corrigan JD, Bogner J. Initial reliability and validity of the Ohio State University TBI Identification Method. J Head Trauma Rehabil 2007;22(6):318-329. [doi: 10.1097/01.HTR.0000300227.67748.77] [Medline: 18025964]
131. Bogner JA, Whiteneck GG, MacDonald J, et al. Test-retest reliability of traumatic brain injury outcome measures: a Traumatic Brain Injury Model Systems study. J Head Trauma Rehabil 2017;32(5):E1-E16. [doi: 10.1097/HTR.0000000000000291] [Medline: 28195954]
132. Whiteneck GG, Cuthbert JP, Corrigan JD, Bogner JA. Prevalence of self-reported lifetime history of traumatic brain injury and associated disability: a statewide population-based survey. J Head Trauma Rehabil 2016;31(1):E55-E62. [doi: 10.1097/HTR.0000000000000140] [Medline: 25931187]
133. Bogner JB, Corrigan JD. Reliability and predictive validity of the Ohio State University TBI Identification Method with prisoners. J Head Trauma Rehabil 2009;24(4):279-291. [doi: 10.1097/HTR.0b013e3181a66356] [Medline: 19625867]
134. McKinlay A, Corrigan JD, Bogner JA, Horwood LJ. Obtaining a history of childhood traumatic brain injury using the Ohio State University TBI Identification Method to elicit adult recall. J Head Trauma Rehabil 2017;32(6):E24-E28. [doi: 10.1097/HTR.0000000000000284] [Medline: 28060210]

XSL•FO
RenderX

135. Olson-Madden JH, Brenner L, Harwood JEF, Emrick CD, Corrigan JD, Thompson C. Traumatic brain injury and psychiatric diagnoses in veterans seeking outpatient substance abuse treatment. J Head Trauma Rehabil 2010;25(6):470-479. [doi: 10.1097/HTR.0b013e3181d717a7] [Medline: 20414130]
136. Budd MA, Dixon TM, Barnett SD, Njoh E, Goetz LL, Ottomanelli L. Examination of traumatic brain injury exposure among veterans with spinal cord injury. Rehabil Psychol 2017 Aug;62(3):345-352. [doi: 10.1037/rep0000129] [Medline: 28594193]
137. Plain language guide series. United States General Services Administration. URL: https://digital.gov/guides/plain-language [accessed 2026-02-27]
138. Keyworth C, Epton T, Goldthorpe J, Calam R, Armitage CJ. Acceptability, reliability, and validity of a brief measure of capabilities, opportunities, and motivations ("COM‐B"). British J Health Psychol 2020 Sep;25(3):474-501. [doi: 10.1111/bjhp.12417]
139. Section 508, Rehabilitation Act of 1973. US Department of Labor. URL: https://www.dol.gov/agencies/oasam/regulatory/statutes/section-508-rehabilitation-act-of-1973 [accessed 2026-02-24]
140. Machin D, Campbell MJ, Tan SB, Tan SH. Sample Size Tables for Clinical Studies, 2nd edition: Blackwell Science Ltd; 2008. [doi: 10.1002/9781444300710]
141. Zar JH. Biostatistical Analysis, 2nd edition: Prentice-Hall; 1984.
142. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation Sci 2011 Dec;6(1):42. [doi: 10.1186/1748-5908-6-42]
143. West R, Michie S. A brief introduction to the COM-b model of behaviour and the PRIME theory of motivation. Qeios. Preprint posted online on 2020. [doi: 10.32388/WW04E6]
144. My HealtheVet surpasses 5 million registered users. US Department of Veterans Affair. 2019. URL: https://connectedcare.va.gov/whats-new/awards/my-healthevet-surpasses-5-million-registered-users [accessed 2026-02-10]
145. Title 38, chapter I, part 17: authority of health care providers to practice in VA. Code of Federal Regulations. URL: https://www.ecfr.gov/current/title-38/chapter-I/part-17/subject-group-ECFRdbd8d11c1202212 [accessed 2026-02-10]
146. Bokolo A. Use of telemedicine and virtual care for remote treatment in response to COVID-19 pandemic. J Med Syst 2020 Jul;44(7):132. [doi: 10.1007/s10916-020-01596-5]
147. Wong A, Bhyat R, Srivastava S, Boissé Lomax L, Appireddy R. Patient care during the COVID-19 pandemic: use of virtual care. J Med Internet Res 2021 Jan 21;23(1):e20621. [doi: 10.2196/20621] [Medline: 33326410]
148. Cotner B, Nakase-Richardson R, Agtarap S, et al. COVID-19 impact on delivery of rehabilitation for persons with traumatic brain injury. Arch Phys Med Rehabil 2022 Mar;103(3):e18-e19. [doi: 10.1016/j.apmr.2022.01.050]
149. ATLAS.ti. URL: https://atlasti.com/ [accessed 2026-02-10]
150. Huberman AM, Miles MB. The Qualitative Researcher's Companion: SAGE Publications, Inc; 2002. [doi: 10.4135/9781412986274]

Abbreviations

CBPR: community-based participatory research
CEC: Community Engagement Council
CFIR: Consolidated Framework for Implementation Research
COM-TBI: Capability, Opportunity, Motivation to Implement a Team-Based Behavioral Intervention
DOD: Department of Defense
EBT: evidence-based behavioral treatment
EMR: electronic medical record
FPA: Focused Program Award
HCD: human-centered design
I-HEAL: Improving Health Care Access and Engagement for Veterans and Service Members with TBI Morbidity
IEP: individual engagement partner
IRB: Institutional Review Board
ISC: Implementation Science Core
LEP: lived experience partner
MPI: multiple principal investigator
MSKTC: Model Systems Knowledge Translation Center
NASEM: National Academies of Science, Engineering, and Medicine
NDSC: National Data and Statistical Center
NIDILRR: National Institute on Disability, Independent Living, and Rehabilitation Research
NIH: National Institutes of Health
OSU TBI-ID: Ohio State University TBI Identification Method
PCORI: Patient-Centered Outcomes Research Institute
PEP: project engagement partner

XSL•FO
RenderX

PI: principal investigator
PM: project manager
PPP: policy and professional partner
PRC: polytrauma rehabilitation center
PTSD: posttraumatic stress disorder
QI: quality improvement
QUERI: Quality Enhancement Research Initiative
RE-AIM: Reach, Effectiveness, Adoption, Implementation, Maintenance
SME: subject matter expert
SOP: standard operating procedure
STAR-VA: Staff Training in Assisted Living Residences-Veterans Affairs
TBI: traumatic brain injury
TBIMS: TBI Model Systems
TeamBI: Team-based Behavioral Interventions
V/SM: Veteran and Service Member
VA: Veterans Affairs
WDMC: Website and Data Management Core

Edited by J Sarvestan; submitted 15.Oct.2025; this is a non–peer-reviewed article;accepted 01.Jan.2026; published 06.Mar.2026.

Please cite as:
Nakase-Richardson R, Haun JN, Moore M, Bogner J, Hoffman JM, Dams-O'Connor K, Coulter J, Kretzmer T, Martin AM, Silva MA, I-HEAL Consortium
Using Implementation Science to Improve Health Care Access and Quality for People With Traumatic Brain Injury–Related Morbidity (I-HEAL): Protocol for a Translational Multiproject Program Award
JMIR Res Protoc 2026;15:e79738
URL: https://www.researchprotocols.org/2026/1/e79738
doi:10.2196/79738

© Risa Nakase-Richardson, Jolie N Haun, Megan Moore, Jennifer Bogner, Jeanne M Hoffman, Kristen Dams-O'Connor, Jill Coulter, Tracy Kretzmer, Aaron M Martin, Marc A Silva, I-HEAL Consortium. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Enhancing Caregivers' Quality of Life Through a Web-Based Person-Centered Solution (TechQoL4Carers): Protocol for a Mixed Methods Pilot Trial

Betania Groba, PhD; Patricia Concheiro-Moscoso, PhD; María del Carmen Miranda-Duro, PhD; Manuel Lagos-Rodríguez, MSc; Javier Pereira, PhD; Laura Nieto-Riveiro, PhD

CITIC. Research group TALIONIS, Faculty of Health Sciences, Universidade da Coruña, As Xubias, s/n, A Coruña, Spain

Corresponding Author:
Patricia Concheiro-Moscoso, PhD
CITIC. Research group TALIONIS, Faculty of Health Sciences, Universidade da Coruña, As Xubias, s/n, A Coruña, Spain

Abstract

Background: Informal care is a social challenge that impacts the daily life and quality of life (QoL) of caregivers. While care has evidence of positive aspects, it can also have negative impacts on mental, physical, economic, and social well-being. Nowadays, health, social, and care systems for informal caregivers are needed, from a person-centered perspective, to promote their QoL, health, and empowerment. Technology is a promising tool to provide personalized services.

Objective: This study aims to develop an innovative care-centered technology solution to enhance caregivers' QoL, care impact, occupational balance, health self-management, and empowerment.

Methods: A mixed methods pilot trial was designed as a single-arm open-label study. Participants will be caregivers of people with disabilities or older people, recruited through direct care centers. Caregivers will engage in a participatory process of developing, testing, and validating the TechQoL4Carers project platform, *CuidaconTIC*; a web-based platform to improve their QoL, impact of care, occupational balance, health self-management, and empowerment. The development of this trial will refine the *CuidaconTIC* platform, which is based on person-centered development, participatory design techniques, and an iterative process. A total of 54 caregivers will participate in a 3-month intervention involving the use of the *CuidaconTIC* platform and the Xiaomi Smart Band 7 wristband. Standardized assessment tools (EQ-5D-5L, Care-Related Quality of Life, Zarit Burden Interview, Caregiver Strain Index, Occupational Balance Questionnaire-Spanish Version, Psychological Empowerment Instrument, and the System Usability Scale) and a self-designed tool (Satisfaction-Q) will be used at 3 time points to collect information about usability, satisfaction, and project variables. Continuous information will be obtained from the platform (My week-Q) and the wearable wristband (physical activity and sleep). An interview will be conducted to gain in-depth knowledge about participants' perspectives. This study was approved by the Ethics Committee for Research and Teaching of the Universidade da Coruña (2023_019) and registered on ClinicalTrials.gov. Participation will be entirely voluntary, with informed consent obtained from each participant. Detailed information sheets and informed consent forms will be provided. The data of the participants will be collected in a pseudonymized form. Once the study has been completed, any possibility of participant identification will be eliminated.

Results: Financial support for this project was received on December 1, 2022. This protocol was submitted after data collection but before analysis. Data collection began in May 2024 and ended in March 2025. By October 2025, 62 participants had been recruited. We expect to publish our results in June 2026.

Conclusions: This protocol focuses on the study of the social challenge of quality of life of caregivers, mostly women, and on reducing the digital gap and promoting inclusion, through a service that places people and care at the center.

Trial Registration: ClinicalTrials.gov NCT06226285; https://clinicaltrials.gov/study/NCT06226285

International Registered Report Identifier (IRRID): DERR1-10.2196/86602

(JMIR Res Protoc 2026;15:e86602) doi:10.2196/86602

KEYWORDS

quality of life; informal caregivers; web-based platform; wearable devices; activities of daily living; person-centered care

XSL•FO
RenderX

Introduction

Background and Rationale

Care is considered an occupation in daily life, encompassing a wide range of responsibilities linked to support in activities of daily living and/or instrumental activities of daily living [1,2]. In this regard, informal care, often provided by family members, plays a crucial role in providing support to individuals with long-lasting needs (chronic illnesses, disability, age-related dependency, among others), and in many cases, they must face decision-making together with health and social services [3,4].

In recent years, the combination of social, demographic, and economic factors has increased informal care across Europe [5]. Statistics indicate that approximately 14.4% of the European population aged 18 to 74 are informal caregivers, providing up to 80% of all long-term care for dependent individuals [6]. In particular, Spain is among the European countries with the highest prevalence of informal caregivers, with 15.3% of the population engaged in this role and 52.9% of informal caregivers dedicating more than 20 hours per week to care [7,8]. Furthermore, literature states that the most common profile of an informal caregiver is a woman (n=32,590, 63.28%) aged between 45 and 64 years (n=18,814, 55%) [6,9].

Although care can be a source of great personal satisfaction, it also poses numerous challenges [10]. Caregiving tasks can significantly impact the quality of life (QoL), well-being, and health of informal caregivers. The burden of long-term care may lead to physical and cognitive difficulties, as well as mental health issues [10,11].

Evidence suggests that women, older caregivers, and spouses are likely to experience high levels of stress related to their care roles [9,12]. Factors such as task overload, insufficient support, financial difficulties, restrictions in daily activities, and social exclusion can lead to chronic stress [13,14]. This prolonged situation, when combined with high levels of emotional exhaustion, sleep disturbances, and physical issues, can lead to "caregiver's burnout syndrome or caregivers' burden" phenomenon [13,15]. Additionally, mental health issues are approximately 20% more prevalent among informal caregivers than noncaregivers, especially among those who provide intensive care (more than 20 h/wk) [5,16].

Informal care is also related to occupational balance, habits, and routines of the caregiver. Providing informal care directly affects the caregiver's daily life [17]. Some authors mention that care responsibilities can be a barrier to accessing education and paid employment [17,18]. In fact, many caregivers, due to difficulties in reconciling work and care duties, may be forced to reduce their working hours and/or leave paid employment, leading to financial, social, and personal difficulties [10,19]. Specifically, statistical data suggest that women aged between 35 and 44 are most affected by this situation, finding themselves obliged to reconcile with part-time employment (60.2%) or to stop working (31.4%) [8].

Furthermore, caregivers also face challenges in dedicating time to other meaningful occupations and maintaining their healthy lifestyle habits [2,17]. Specifically, statistical data shows that Spanish caregivers have reduced their leisure activities by 77%, their time for social participation by 47.88%, and their time for personal self-care by 60.34% [8].

In this way, the repercussions of care can make informal caregivers susceptible to health issues that may lead to chronic diseases [11]. Furthermore, the increase in the number of older informal caregivers also highlights the need for social and health care support for them [5]. These circumstances, along with the implementation of person-centered care (PCC) principles, underscore the importance of decision-makers, policymakers, and professionals recognizing, respecting, and supporting informal caregivers by seeking and developing cost-effective, sustainable, and preventive solutions [20,21].

Various studies have focused on the development of intervention programs aimed at reducing care burden, providing social support, or improving caring skills [1,14,19]. However, despite their potential, these programs have some limitations related to difficult access, lack of information, or insufficient support for caregivers to participate effectively [16]. Consequently, there is a need for innovative methods that eliminate these barriers [22]. Thus, Information and Communication Technologies offer the possibility to provide support and services to caregivers in real time [22].

In recent years, the development and evaluation of technological solutions aimed at the informal caregiving population has become an area of growing interest [23]. The literature highlights that tools such as mobile applications, online educational programs on caregiving, health monitoring devices, task management systems, and emotional support networks are accepted by informal caregivers as they facilitate the management of care, improve the quality of care, reduce emotional and physical burden, and enhance their well-being [20,24].

In this regard, Tan et al [25], in their scoping review, report that the use of technological tools by informal caregivers contributes to a reduction in caregiving burden and a slight improvement in the quality of life of caregivers. While various studies emphasize the benefits of digital solutions to improve adherence to treatments for dependent individuals, optimize care task management, and highlight the importance of home assistive technology [26].

Despite advancements in the implementation of technological solutions, significant barriers to their adoption remain, particularly among older caregivers [25]. Studies such as those conducted by Egan et al [27] and Xiong et al [28] point out certain concerns of informal caregivers. These include fears that these technologies may replace human contact, a lack of familiarity with these technologies, the costs associated with some applications or technological services, the lack of technological adaptation, sustainability, the rapid evolution of technology, as well as the design and usability of these systems.

Furthermore, many mobile applications do not offer personalized content, leading caregivers to perceive the information as irrelevant or overly complex for their needs [29]. As noted by Premanandan et al [30], current technological solutions primarily focus on managing the care of dependent individuals but fail to

XSL•FO
RenderX

address the specific needs of the caregivers themselves. This trend is particularly evident in studies centered on caregivers of individuals with dementia, where Information and Communication Technologies solutions have been designed without a comprehensive approach that also considers the emotional and well-being needs of informal caregivers [31]. Additionally, some applications aimed primarily at caregivers, such as the CareFit app, which encourages caregivers to engage in regular physical activity at home, have limitations, including the lack of time caregivers have to carry out activities, difficulties in understanding how to perform the activities, and the feeling of insecurity when carrying out those activities at home [32].

Thus, the literature emphasizes the importance of involving informal caregivers in the design and development process of technological solutions to ensure that these meet their needs effectively, are accessible, easy to use, and sustainable in the long term [22,28]. Person-centered design is crucial for improving the health, maintaining the quality of life, and strengthening the well-being of caregivers, which in turn positively impacts the care provided, at health, social, and economic levels [33].

In this regard, this project aims to address these challenges by creating a technological solution that improves the quality of life and health status of informal caregivers, as well as being aware of the implications and repercussions that continuous care has on informal caregivers.

Aims, Objectives, and Hypotheses

According to the existing literature and the main conclusions of the background, there is a clear need to develop a technological solution specifically designed for informal caregivers. The hypothesis is that this innovative technological solution will have a positive impact on their QoL, impact of care, occupational balance, self-management of health, and empowerment.

The primary objective of this pilot trial is to develop and evaluate an innovative technological solution aimed at enhancing the QoL, impact of care, occupational balance, self-management of health, and empowerment of informal caregivers of older people and/or people with disabilities. The main question is: Will the routine use of the TechQoL4Carers platform have a positive impact on the daily life of informal caregivers?

The specific objectives of this project are (1) to develop an innovative technological solution for informal caregivers aimed to improve their QoL, impact of care, occupational balance, self-management of health, and empowerment; (2) to assess the impact of the platform on the QoL, care, self-management of health, and empowerment of informal caregivers; (3) to test and evaluate the contents, usability, quality, and satisfaction with the platform by informal caregivers; and (4) to identify potential barriers and necessary changes as well as facilitators to use the platform.

Methods

Study Design

A mixed methods pilot trial was designed as a single-arm open-label study. Informal caregivers will participate in a participatory process of developing, testing, and validating a technological platform to improve their QoL, impact of care, occupational balance, self-management of health, and empowerment.

The 3-month intervention will be based on the use of the TechQoL4Carers technology platform, called *CuidaconTIC* [34], and the Xiaomi Smart Band 7 wearable wristband. Formal assessments using standardized tools will be conducted as follows: baseline (preintervention: t_0), mid-intervention (at 6 wk: t_2), and postintervention (t_4).

Throughout the intervention (t_1 to t_3), continuous information will be obtained through data from the platform and the Xiaomi wearable wristband. A qualitative evaluation will be conducted to gain in-depth knowledge about participants' perspectives on the intervention and study procedures (t_4).

The trial has been registered on ClinicalTrials.gov with the identification number NCT06226285 [35]. As this is research with people, in this case, caregivers, a favorable report has been previously obtained from the Ethics Committee for Research and Teaching of the Universidade da Coruña (code: 2023_019).

In addition, the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) statement [36] and the CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) guidelines [37] were adopted for the protocol design. A summary of the trial, based on the World Health Organization (WHO) Trial Registration Data Set [38], can be found in Multimedia Appendix 1.

Participants, Recruitment, and Eligibility Criteria

The participants in this study will be caregivers of people with disabilities or older people. The sample will be recruited through direct care centers for people with disabilities and older people.

Depending on the functioning of the entities, a meeting will be held in the centers, individually or in groups, with the persons preselected by the centers and the research team to provide detailed information on the project and possible participation. In this meeting, the inclusion and exclusion criteria will be explained and a paper information sheet will be distributed.

Individuals may express their interest in participating through the e-mails included in the information sheet or through the reference persons in each entity. Subsequently, an individual meeting will be held with the candidates to clarify any doubts and, if applicable, to confirm participation in the project by signing the informed consent form.

Informal caregivers will be included in the study if they are 18 years of age or older; perform the role of informal caregiver, are the main caregiver; are living with the person receiving the care; regularly use a smartphone; and commit themselves to

XSL•FO
RenderX

wearing an activity wristband regularly during the project. The exclusion criteria were those who had been an informal caregiver for less than 1 year, were the main caregiver of more than 1 person in a situation of dependency, or had modified legal capacity.

Figure 1 shows visually the study flowchart described above.

Figure 1. Study flowchart.

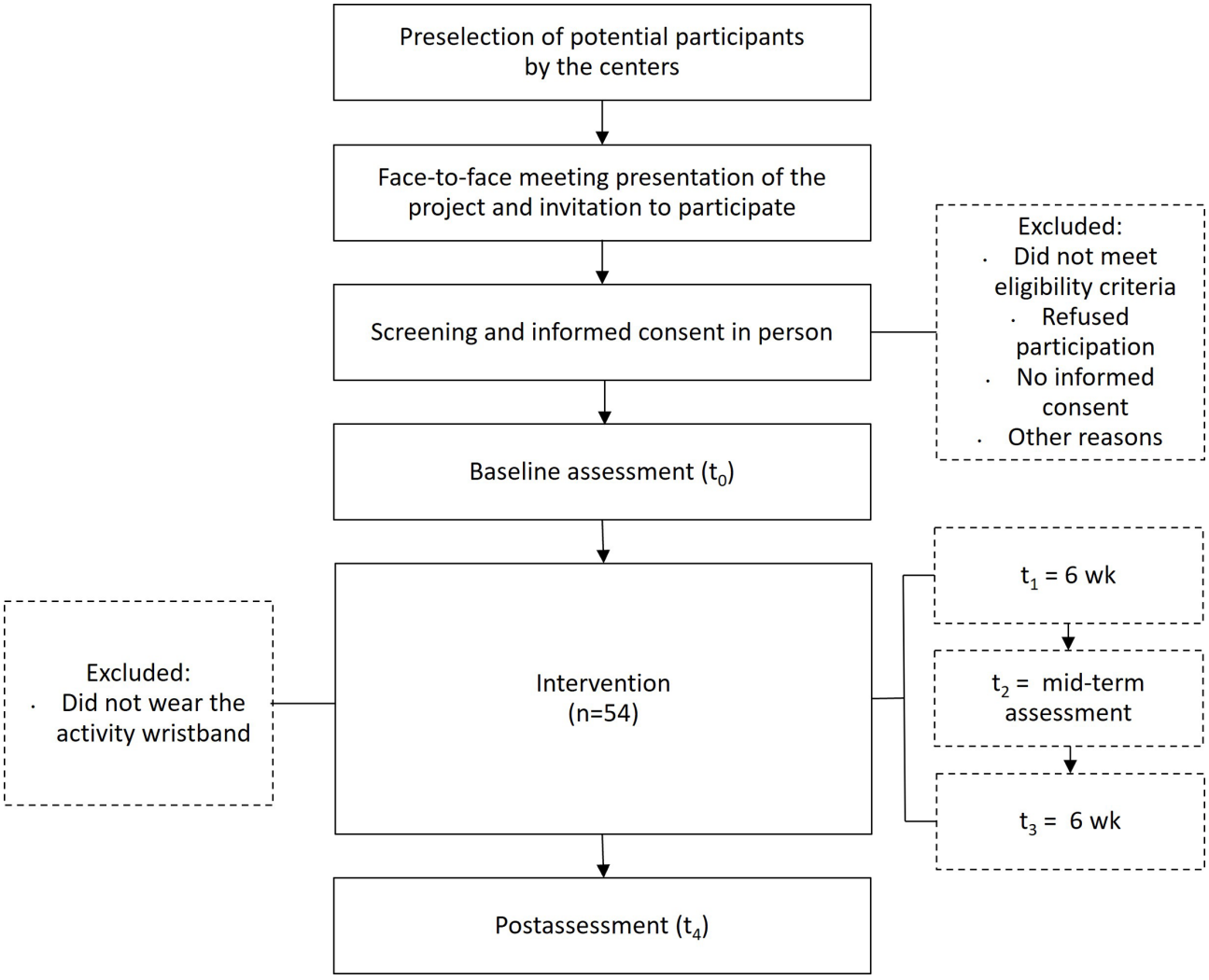

Sample Size

To determine the necessary sample size for this study, the G*Power software was used. Based on similar studies and assuming that caregivers' QoL is the primary outcome of this study (with all other variables treated as secondary and/or exploratory outcomes), power analysis was performed [39]. An a priori power analysis was conducted using a paired t test on the primary outcome, as an improvement in informal caregivers' QoL was expected after the intervention. A moderate effect size (Cohen *d*=0.5), a significance level of .05, and a statistical power of 80% (1−β=.80) were assumed. Under these assumptions, a sample of 54 caregivers who complete the study (*df*=53) provides power greater than 80% to detect the expected effect.

Given the multiplicity of the planned analyses, confirmatory inference will be restricted to the primary outcome. Secondary and exploratory outcomes will be interpreted with caution, without applying a formal adjustment for multiple comparisons, and emphasis will be placed on effect sizes and their 95% CIs.

In addition, an attrition rate of up to 20% is anticipated (due to dropouts, absence from the postintervention assessment, or technical problems with the wearable device or platform). Therefore, we plan to recruit approximately 65 informal caregivers, so that at least 54 are expected to complete the study and be included in the primary analysis.

Ethical Considerations

The project complies with the fundamental principles for research with human beings established in the Declaration of Helsinki, the Oviedo Bioethics Convention, and the Universal Declaration on the Human Genome and Human Rights.

The performance of this study will respect the applicable regulations of ethics regarding research with humans. Also, for data protection, we will follow the Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016, on the protection of natural persons regarding the processing of personal data and on the free movement of such data and the national regulations on data protection.

XSL•FO
RenderX

To guarantee these aspects, the research protocol was submitted to the Ethics Committee for Research and Teaching of the Universidade da Coruña. A favorable report from this committee was obtained (2023_019).

The group of participants in the study will be adults in an informal caregiver's role. Participation in this project will not implicate any risk or damage to the participants, who will (1) cover baseline, mid-term, and end-line surveys, (2) wear a noninvasive wearable device for the duration of the validation, (3) interact and use the platform, and (4) participate in individual semistructured interviews.

Particular attention will be paid to ethical issues: aspects that guarantee data protection and informed and voluntary participation.

The research group is committed to guaranteeing the confidentiality and protection of the data obtained from the participants in the project. The data of the participants will be collected in a pseudonymized form to allow validation tests at different stages of the development process. The interviews will be audio-recorded, with the consent of the participants. A transcript of the interviews will then be made, and the identities of the participants will be coded. The information that might allow identification shall be anonymized. Subsequently, the audios will be deleted.

The access to the identification of the participants will be exclusively carried out by the research team. Once the study has been completed, any possibility of participant identification will be eliminated.

The platform's cybersecurity mechanisms will be identified and implemented following current legislation.

Participation will be completely voluntary, and informed consent will be obtained from each person. For this purpose, detailed information sheets and an informed consent form will be provided to explain the participation in the study. There is no obligation to complete the experiment, so participants can withdraw from it at any time.

Individual participant data will be available for sharing with the scientific community, upon request and after the acceptance for publication of the main findings from the final data set.

The data will be completely anonymized for sharing. The dataset will include only the results of quantitative variables.

Finally, for validation testing of the platform, it will be necessary to obtain a robust prototype. This prototype will be previously tested by professionals of the team to ensure the absence of unexpected behaviors of the application that may generate discomfort to participants.

The cost of the Xiaomi Mi Band 7 wristbands will be supported by the project team, and people will have to authorize the capture of data from this device (physical activity and sleep).

Participants will not receive monetary compensation. As compensation for their time, participants will receive the Xiaomi Mi Band 7 wristband used in the study at no cost (ie, they may keep the device after participation). Compensation is not contingent on completing all study procedures; participants may withdraw at any time without penalty.

Intervention

The intervention lasts 3 months and is based on the use of *CuidaconTIC* and the Xiaomi Smart Band 7 activity wristband. Participants will regularly use *CuidaconTIC* services and wear the Xiaomi Smart Band 7 wristband.

CuidaconTIC is a technological platform for caregivers developed through the TechQoL4Carers project [34]. The research team designed and developed *CuidaconTIC* through an iterative and participatory process with informal caregivers, the platform's end users, and other actors involved in care, such as health science professionals.

The platform is integrated into both a website and a mobile application. The services offered include a summary of the individual's health status, training content, advice, and personalized recommendations based on their QoL, care situation, occupational balance, self-management of health, and empowerment. It also allows for synchronization with the Xiaomi Smart Band 7 activity wristband. *CuidaconTIC*, 1.0 version, will be used for the pilot project. After the trial results are integrated, the final version of the platform will be obtained. The main services of the *CuidaconTIC* platform are: My daily summary, My health, My week, I take care of myself, and I take care of you.

Caregivers are expected to access and use the TechQoL4Carers platform at least once a week to complete the weekly questionnaire and consult the training content related to self-care and care provision at least once a month.

The research team will provide caregivers with the Xiaomi Smart Band 7 wristband to capture health-related data and synchronize it with the platform. Participants are expected to use the wristband daily during the 3-month intervention. These data will be combined with other information collected through subjective measurements to provide personalized content and recommendations.

Outcomes and Measurements

Participants will complete the quantitative evaluations at 3 points in time: baseline (t_0), intermediate (t_2), and final (t_4) through the *CuidaconTIC* platform in its evaluation section. For the final assessment, people may participate in a semistructured interview. The evaluations at these time points will be face-to-face at the reference centers or, if preferred, at the facilities of the research center.

The weekly evaluation will be performed through the *CuidaconTIC* platform and will be reminded through notifications in the mobile application.

The schedule of enrollment, interventions, and assessments, based on the proposal of the SPIRIT statement [36], can be consulted in Table 1.

XSL•FO
RenderX

Table . Schedule of enrollment, interventions, and assessments.

	Enrollment	Preassessment	Intervention (3 mo)			Postassessment
	$-t_1$	t_0 (baseline)	t_1[a]	t_2[b]	t_3[c]	t_4 (post intervention)[d]
Enrollment						
Informed consent	✓					
Configuration						
Wristband delivery	✓					
CuidaconTIC settings	✓					
Intervention						
CuidaconTIC platform						
Xiaomi Band 7						
Assessments						
EQ-5D-5L		✓		✓		✓
CarerQoL[e]		✓		✓		✓
ZBI[f]		✓		✓		✓
CBI[g]		✓		✓		✓
OBQ-E[h]		✓		✓		✓
PEI[i]		✓		✓		✓
Physical activity						
Sleep						
My week-Q[j]						
Satisfaction-Q[k]						✓
SUS[l]						✓
Interview						✓

[a]t_1: 6 weeks of intervention (1-6 wk).

[b]t_2: mid-term assessment.

[c]t_3: 6 weeks of intervention (7-12 wk).

[d]t_4: 1-month post intervention.

[e]CarerQoL: Care-Related Quality of Life.

[f]ZBI: Zarit Burden Interview.

[g]CBI: Caregiver Strain Index.

[h]OBQ-E: Occupational Balance Questionnaire-Spanish version.

[i]PEI: Psychological Empowerment Instrument.

[j]My week-Q: *CuidaconTIC* weekly Questionnaire.

[k]Satisfaction-Q: *CuidaconTIC* Satisfaction Questionnaire.

[l]SUS: System Usability Scale.

XSL•FO
RenderX

Characteristics of Participants

Sociodemographic Questionnaire

A sociodemographic questionnaire was developed with the following variables: age, gender, marital status, educational level, residential environment, household unit, employment situation, years in the role of caregiver, supports, type of support, relationship with the person receiving care, main difficulties, diagnosis, and age of the cared-for person.

Primary Outcome Measures

EQ-5D-5L

EQ-5D-5L is a self-administered scale to assess perceived health-related QoL [40-42].

Health state is obtained from the EQ-5D-5L, through its descriptive system with 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. The dimensions are scored on a 5-point Likert scale reflecting the level of severity (no problem, slight problems, moderate problems, severe problems, and extreme problems or unable to perform). The health profile or health state describes the level of severity that the person perceives in each dimension and it is represented through an individual code for each dimension scored between 1 and 5 or through a single 5-digit code (ranging from 11111 to 55555).

Overall health is derived from a visual analog scale (EQ-VAS) on perceived health, from 0 to 100. In this way, the person scores his or her health on a given day, assigning it a value between 0 (the worst health you can imagine) and 100 (the best health you can imagine).

Care-Related Quality of Life Instrument

Care-related quality of life (CarerQoL) is a self-administered instrument to measure care-related QoL and the impact of the care on caregivers [43].

CarerQoL will be measured through 5 negative and 2 positive dimensions (CarerQoL-7D). The 7 dimensions are fulfillment from caregiving, relational problems, mental health problems, problems with combining daily activities with care, financial issues, perceived support, and physical health problems [44]. They are scored on a Likert scale with 3 values indicating the severity level (no, some, and a lot). For negative dimensions, the values are 0 (a lot), 1 (some), and 2 (no); while for positive dimensions, the values are reversed 0 (no), 1 (some), and 2 (a lot). Score range: 0 to 14. A tariff for the CarerQoL based on Dutch preferences is also available which allows a weighted sum score, ranging from 0 to 100 [44,45]. The higher the score on the dimensions, the better the situation of care.

The CarerQoL-VAS allows rating how happy the person feels at that moment as a measure of general well-being. This construct is scored on a horizontal VAS range from 0 (completely unhappy) to 10 (completely happy). The higher the score, the better the situation in terms of well-being.

Secondary Outcome Measures

Zarit Burden Interview

Zarit Burden Interview is a questionnaire designed for the caregiver population to assess the level of burden and detect possible caregiving stressors [46-49]. The 22-item version is used, scored from 1 to 5.

The score range is 22 to 110. Score interpretation is as follows, with 2 cut-off points: no burden, 22 to 46 points; mild burden, 47 to 55 points; and severe burden, 56 to 110 points.

Caregiver Strain Index

The Caregiver Strain Index or Robinson Index is a 13-item tool for caregivers to measure perceived strain, and detect burden, and possible stressors related to care [50-52]. Each item is answered in a dichotomous way: yes or no.

The score range is 0 to 13. Scores of 7 or more identify caregivers at risk for high caregiving strain or burden.

Occupational Balance Questionnaire, Spanish Version

Occupational balance will be measured through Occupational Balance Questionnaire-Spanish version [53,54]. This self-administered tool of 13 items is scored on a 6-step ordinal scale, ranging from 0 (completely disagree) to 5 (completely agree). It evaluates the person's current experience with occupational balance, considering the quantity, variability, and meaning attributed to activities of daily living.

The score range is 0 to 65. The results show that the higher the score, the higher the occupational balance.

Psychological Empowerment Instrument

The Psychological Empowerment Instrument adapted will be used to assess changes in informal caregivers' perceptions of empowerment.

It is a self-administered instrument that measures psychological empowerment in the work environment, considering 4 cognitive dimensions: competence, meaning, self-determination, and impact [55,56]. The Spanish version of the Psychological Empowerment Instrument comprises 13 items divided among the 4 dimensions and scored on a 7-point Likert scale [57]. This version will be used in the study replacing the references to work and the role of the workers with terms associated with care and caregivers.

The score range is 7 to 91 points. The higher the score, the higher the level of psychological empowerment.

Activity Wristband: Xiaomi Smart Band

In the study, participants will use the Xiaomi Smart Band [58] activity wristband regularly during the 3 months of the intervention.

Version 7 of the Xiaomi Smart Band activity wristband allows the recording and monitoring of health-related variables: physical activity, sleep, heart rate, blood oxygen saturation (SpO_2), women's health, and stress [58].

The wristband will be automatically synchronized with the TechQoL4Carers platform. The variables that will be monitored

XSL•FO
RenderX

during the project are: physical activity (number of steps, distance in meters, and activity in min) and sleep (light sleep, deep sleep, rapid eye movement sleep, and time awake after sleep onset in min).

My Week Questionnaire

My Week Questionnaire is a questionnaire designed by the research team included on *CuidaconTIC*. Participants could self-administer it weekly to identify changes perceived in daily functioning based on the variables studied: QoL, impact of the care, occupational balance, self-management of health, and empowerment.

The development of this questionnaire has been based on previous surveys developed by the research team [59], but it has been adapted to caregivers for this study.

My Week Questionnaire is integrated with 9 items that are answered using a 6-point Likert scale, ranging from 0 (completely disagree) to 5 (completely agree). The score range is 0 to 45 points. The higher the score, the greater the perception of positive changes in QoL-related factors.

CuidaconTIC Satisfaction Questionnaire

The *CuidaconTIC* Satisfaction Questionnaire will be administered at the end of the intervention to assess end-user satisfaction with the project and the technology. The questionnaire has been adapted for this study based on the Technology Satisfaction Factor Loadings questionnaire [60].

It is a self-administered questionnaire consisting of 16 items grouped into 4 factors (competence, assessment, performance, and TechQoL4Carers platform), rated on a 6-point scale, ranging from 0 (completely disagree) to 5 (completely agree). The score range is 0 to 80. The higher the agreement with the statements, the higher the satisfaction with the project.

System Usability Scale

The System Usability Scale is a tool used to measure subjective perceptions of technology usability [61]. It contains 10 items scored on a 5-point Likert scale, from 0 (strongly disagree) to 4 (strongly agree). The items are statements about the effectiveness, efficiency, and satisfaction with a system or product.

The score range is 0 - 100. The higher the score, the higher the perceived level of usability of a product or system. Scores of 70 or more identify acceptable products [62] with a good (70-85) and excellent (85 or more) level of usability.

Semistructured Interview

At the end of the project, a semistructured individual interview will be conducted with each caregiver. The interview guide has been designed to elicit the meaning that the caregivers give to the caregiving experience, to the participation in the project, to the use of technology in the project, and the impact of this project on their daily life (QoL, impact of care, occupational balance, self-management of health, and empowerment).

This interview will take place in the collaborating entities, in a quiet and comfortable place, and at a time that suits the needs of the participants. It will be recorded with a digital device to facilitate the literal transcription. This information will be coded and at the end of the study, the transcriptions will be anonymized and the audio will be deleted.

The interview guide can be found in Appendix B in Multimedia Appendix 2.

Analysis

The statistical analysis of the data will be performed using R software (version 4.4.3 [63]). All data will be cleaned and processed prior to analysis, and a range of statistical procedures will be performed based on the study hypotheses and objectives. The significance level for hypothesis testing is set at 5%. Numerical variables will be presented as mean and SD, along with range, minimum, and maximum values, while categorical variables will be presented as absolute frequency and valid percentage.

After performing a descriptive analysis of the variables, we will determine the distribution of the data using formal tests of normality: the Kolmogorov-Smirnov or Shapiro-Wilk tests, as appropriate. This will help decide whether parametric or nonparametric statistical methods are used. Parametric tests, such as the paired *t* test, ANOVA, and Pearson correlation, will be used to analyze data if normality is confirmed. Otherwise, nonparametric tests, including the Wilcoxon signed-rank test, Friedman and Spearman rank correlation, will be employed. Pre- and postintervention changes in the primary outcome and in the other continuous variables will be analyzed using a paired Student *t* test or, where appropriate, the Wilcoxon signed-rank test. The effect size of within-subject changes will be estimated using Cohen *d*, in order to quantify the magnitude of the observed changes. In addition, 95% CIs for these effect sizes will be reported. It is important to note that QoL, considered the primary outcome of the study, will be used for formal confirmatory inference to control the type I error rate; analyses of other outcomes will be considered secondary or exploratory.

In addition, exploratory correlation analyses will be conducted to examine relationships both among the different outcomes from the tools and between these outcomes and summary measures of the biometrical parameters recorded by the wearable devices. Depending on the distribution of the variables, Pearson or Spearman correlation coefficients will be used. For these analyses, the correlation coefficient r and its 95% CI will be reported as the primary measure of effect size. To explore relationships between categorical variables, the chi-square test will be applied. When expected cell frequencies are low, the Fisher exact test will be used instead.

Furthermore, given the longitudinal nature of the study, data from wearable devices and scores from the weekly questionnaire will be summarized at the weekly level. For wearable data, daily recordings will be checked for validity and aggregated as weekly averages. Based on previous studies [64,65], a week will be considered valid if it contains at least 4 days with usable recordings. Trends in weekly measurements of sleep, heart rate, and step count, as well as weekly questionnaire scores, will be analyzed using repeated-measures ANOVA or the nonparametric Friedman test, depending on data distribution, including only

XSL•FO
RenderX

participants who meet prespecified minimum completeness criteria for the weekly data.

To handle incomplete weekly or wearable data systematically and reduce the risk of biased estimates, the amount and pattern of missing wearable and weekly questionnaire data will be described, and all available valid weeks will be included in the analyses under a missing-at-random assumption. In addition, sensitivity analyses will be performed to assess the robustness of the longitudinal and correlational findings to different assumptions about missing data.

Regarding the qualitative analysis of the final interview, the instructions of the thematic analysis by Braun and Clarke [66] will be followed, who defined 6 phases: (1) familiarizing with the data; (2) generating initial codes; (3) searching for themes; (4) reviewing themes; (5) defining and naming themes; and (6) producing the report. The analysis will be carried out independently by 3 researchers using the previously described inductive approach, and MAXQDA 24 software [67] will be used for the triangulation of the information.

Results

The project was funded in December 2022, and enrollment was completed in March 2025. As of October 2025, this study has completed the recruitment of participants (N=62). The improvements resulting from the iterative co-design process with caregivers have now been implemented on the platform. Figure 2 shows the main menu of the *CuidaconTIC* platform in Spanish with the main sections: My daily summary, My health, My week, I take care of myself, and I take care of you.

Figure 2. Screenshot of the *CuidaconTIC* platform.



The dataset is now being prepared for publication. Data analysis (quantitative and qualitative) will begin next, and the first results are expected to be submitted for publication in June 2026.

Discussion

Anticipated Findings

This study aims to develop and evaluate an innovative technological solution aimed at enhancing the QoL, impact of care, occupational balance, self-management of health, and empowerment of informal caregivers of older people and/or people with disabilities.

Technology, empowerment, and the PCC framework can help to improve the caregiver's health status, QoL, and self-management of their own health. These benefits could extend to the care recipients. Hence, caring for the caregiver has a double effect and greater impact on the social and health systems [68]. The PCC approach is a core element of this research, widely studied in health sciences [21,69,70]. The design and development of the *CuidaconTIC* platform will study the eight PCC dimensions: (1) respect for individuals' values, preferences, and expressed needs; (2) provision of information and education; (3) access to care; (4) emotional support; (5) involvement of family and friends; (6) continuity and safe transitions; (7) physical comfort, and (8) care coordination.

Person-centered technology has significant potential to support informal caregivers. However, most existing technological solutions primarily focus on care recipients, aiming to promote their autonomy, improve their QoL, and consequently, reduce the burden on caregivers [33,71]. Although there are initiatives targeted at caregivers, they tend to be designed for specific groups, such as caregivers of older adults with dementia, or they primarily provide training and informational services without addressing the broader needs of informal care.

Following a comprehensive analysis of health care and social challenges faced by informal caregivers, this study will propose

XSL•FO
RenderX

alternative solutions using person-centered technology. This approach will help advance the field of care and technology as a tool to support caregivers in their daily life.

If the initial hypothesis is confirmed, the study will move toward exploring a low-cost intervention with minimal time requirements, oriented in particular for people with limited availability.

To our knowledge, this is one of the first studies that places informal caregivers at the center of technological development, involving them in an iterative process to create a tool that not only improves care but also directly supports their own QoL, the impact of care, occupational balance, self-management of health, and empowerment.

In summary, this study has the potential to improve caregivers' perception of their own QoL and health status, promote healthy lifestyle habits, and expand knowledge about care and needs from a preventive perspective.

Conclusions

This trial focuses on studying the social challenge of the QoL of the caregiver population, mostly women, and on reducing the digital gap and promoting inclusion, through a service that places people and care at the center.

The results of this study will provide an in-depth understanding of the phenomenon of informal care. In addition, it will explore from the perspective of the caregivers themselves what aspects can be interesting to support their daily life through a technological solution focused on the care for caregivers.

Acknowledgments

The authors acknowledge the recommendations made on the protocol by caregivers and professionals from organizations providing direct care to people with disabilities and older people.

Funding

This work was supported by Universidade da Coruña, Xunta de Galicia and CITIC, Talionis Research Group of the University of A Coruña (grants for consolidation and structuring of competitive research units ED431B 2025/23). CITIC is a center accredited for excellence within the Galician University System and a member of the CIGUS Network (Ref. ED431G 2023/01). Additionally, it is co-financed by the European Union through the FEDER Galicia 2021-2027 operational programme. PC-M received funding for postdoctoral training from the Xunta de Galicia (ED481B-2023-125). This publication is also part of the project TED2021-130127A-I00 - Quality of life for caregivers through a person-centered technological solution, funded by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU"/ PRTR (grant TED2021-130127A-I00). The funders had no role in the study design, in the writing of the manuscript, or in the decision to publish the results.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

BG, PC-M, MdCM-D, ML-R, JP, and LN-R have contributed to the conception and design of the study protocol. BG, PC-M, and LN-R have been mainly responsible for drafting and writing the protocol, with contributions and critical reviewing of important intellectual content from MdCM-D, ML-R, and JP. All authors have been involved in the final approval of the version to be published, and all agree to be accountable for all aspects of the protocol.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of the trial, based on the WHO Trial Registration Data Set. WHO: World Health Organization.
[DOCX File, 17 KB - resprot_v15i1e86602_app1.docx]

Multimedia Appendix 2
Semistructured interview guide.
[DOCX File, 15 KB - resprot_v15i1e86602_app2.docx]

References

1. Czaja SJ, Perdomo D, Lee CC. The role of technology in supporting family caregivers. Presented at: Human Aspects of IT for the Aged Population Design for Aging; Jul 17-22, 2016. [doi: 10.1007/978-3-319-39943-0_17]
2. Malfitano S, Paula A, Sakellariou D. Care and occupational therapy: what kind of care definition do we have? Cad Bras Ter Ocup 2019;27(3):681-685. [doi: 10.4322/2526-8910.ctoARF1886]

XSL•FO
RenderX

3. Hoffmann F, Rodrigues R. Informal carers: who takes care of them? Policy brief 4/2010. : European Centre for Social Welfare Policy and Research; 2010 URL: https://ses.sp.bvs.br/local/File/Informal%20Carers_Who%20Takes%20Care%20of%20Them.pdf [accessed 2026-01-23]
4. Petrini M, Cirulli F, D'Amore A, Masella R, Venerosi A, Carè A. Health issues and informal caregiving in Europe and Italy. Ann Ist Super Sanita 2019;55(1):41-50. [doi: 10.4415/ANN_19_01_08] [Medline: 30968835]
5. Tur-Sinai A, Teti A, Rommel A, Hlebec V, Yghemonos S, Lamura G. Cross-national data on informal caregivers of older people with long-term care needs in the European population: time for a more coordinated and comparable approach. J Biosoc Sci 2023 Mar;55(2):378-382. [doi: 10.1017/S0021932021000742] [Medline: 34986919]
6. Study on exploring the incidence and costs of informal long-term care in the EU. : European Commision; 2021 URL: https://www.infocop.es/pdf/long-term.pdf [accessed 2026-01-23]
7. Living conditions and welfare. persons providing informal care or assistance by sex, age, degree of urbanisation, most frequent activity status and frequency (2016). Eurostat. 2023. URL: https://ec.europa.eu/eurostat/databrowser/view/ILC_ATS18/default/table?lang=en [accessed 2026-01-23]
8. Employed persons, care of dependent persons (children or sick, disabled or older adults) [Web page in Spanish]. Instituto Nacional de Estadística. 2023. URL: https://www.ine.es/ss/Satellite?param1=PYSDetalle&c=INESeccion_C&utm_source=chatgpt.com¶m3=1259924822888&p=%2F&pagename=ProductosYServicios%2FPYSLayout&cid=1259925472720&L=1 [accessed 2026-01-23]
9. Rexhaj S, Nguyen A, Favrod J, et al. Women involvement in the informal caregiving field: a perspective review. Front Psychiatry 2023;14:1113587. [doi: 10.3389/fpsyt.2023.1113587] [Medline: 36741114]
10. Morelli N, Barello S, Mayan M, Graffigna G. Supporting family caregiver engagement in the care of old persons living in hard to reach communities: a scoping review. Health Soc Care Community 2019 Nov;27(6):1363-1374. [doi: 10.1111/hsc.12826] [Medline: 31441179]
11. Stöckel J, Bom J. Revisiting longer-term health effects of informal caregiving: evidence from the UK. The Journal of the Economics of Ageing 2022 Feb;21:100343. [doi: 10.1016/j.jeoa.2021.100343]
12. Cohen SA, Sabik NJ, Cook SK, Azzoli AB, Mendez-Luck CA. Differences within differences: gender inequalities in caregiving intensity vary by race and ethnicity in informal caregivers. J Cross Cult Gerontol 2019 Sep;34(3):245-263. [doi: 10.1007/s10823-019-09381-9] [Medline: 31407137]
13. Van Droogenbroeck F, Spruyt B, Gérain P, et al. Informal caregiving and mental health: results from the Belgian health interview survey 2013 and 2018. BMC Public Health 2025 Jan 2;25(1):15. [doi: 10.1186/s12889-024-20957-0] [Medline: 39748406]
14. Martínez-Campos A, Compañ-Gabucio LM, Torres-Collado L, Garcia-de la Hera M. Occupational therapy interventions for dementia caregivers: scoping review. Healthcare (Basel) 2022 Sep 14;10(9):1764. [doi: 10.3390/healthcare10091764] [Medline: 36141376]
15. Zhou Y, Bai Z, Wan K, Qin T, He R, Xie C. Technology-based interventions on burden of older adults' informal caregivers: a systematic review and meta-analysis of randomized controlled trials. BMC Geriatr 2024 May 4;24(1):398. [doi: 10.1186/s12877-024-05018-w] [Medline: 38704539]
16. Queluz FNFR, Kervin E, Wozney L, Fancey P, McGrath PJ, Keefe J. Understanding the needs of caregivers of persons with dementia: a scoping review. Int Psychogeriatr 2020 Jan;32(1):35-52. [doi: 10.1017/S1041610219000243] [Medline: 30967164]
17. Demers L. Expanding occupational therapy perspectives with family caregivers. Can J Occup Ther 2022 Sep;89(3):223-237. [doi: 10.1177/00084174221103952] [Medline: 36135329]
18. Leven NV, de Lange J, Groenewoud J, van der Ploeg E, Pot AM. Exploring the usefulness of indicators for referring people with dementia and their informal caregivers to activating interventions: a qualitative analysis of needs assessments. BMC Geriatr 2019 Aug 23;19(1):230. [doi: 10.1186/s12877-019-1221-0] [Medline: 31443632]
19. Chou WH, Ko YL, Huang XY. Design of occupational therapy interventions for middle-aged and elderly family caregivers. Healthcare (Basel) 2021 Mar 3;9(3):275. [doi: 10.3390/healthcare9030275] [Medline: 33802540]
20. Hassan AYI. Challenges and recommendations for the deployment of information and communication technology solutions for informal caregivers: scoping review. JMIR Aging 2020 Jul 29;3(2):e20310. [doi: 10.2196/20310] [Medline: 32723720]
21. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. Med Care Res Rev 2013 Aug;70(4):351-379. [doi: 10.1177/1077558712465774] [Medline: 23169897]
22. Hassan A. HTA1 information and communication technology solutions for informal caregivers. Value Health 2022 Jul;25(7):S504 [FREE Full text] [doi: 10.1016/j.jval.2022.04.1135]
23. Kilfoy A, Zaffino I, McAtee E, et al. Understanding the effectiveness and design of parent-oriented mobile health interventions: a systematic review and narrative synthesis. BMC Pediatr 2025 May 10;25(1):372. [doi: 10.1186/s12887-025-05656-y] [Medline: 40349017]
24. Wójcik D, Szczechowiak K, Konopka P, et al. Informal dementia caregivers: current technology use and acceptance of technology in care. Int J Environ Res Public Health 2021 Mar 19;18(6):3167. [doi: 10.3390/ijerph18063167] [Medline: 33808644]

XSL•FO
RenderX

25. Tan YRL, Neo NWS, Siah CR. The impact of digital tool use on older caregivers' burden, depression and quality of life: a systematic review and meta‐analysis. Worldviews Ev Based Nurs 2025 Apr;22(2). [doi: 10.1111/wvn.70021]
26. Espírito-Santo M, Santos S, Estêvão MD. Digital solutions available to be used by informal caregivers, contributing to medication adherence: a scoping review. Pharmacy (Basel) 2024 Jan 23;12(1):20. [doi: 10.3390/pharmacy12010020] [Medline: 38392927]
27. Egan KJ, Clark P, Deen Z, et al. Understanding current needs and future expectations of informal caregivers for technology to support health and well-being: national survey study. JMIR Aging 2022 Jan 27;5(1):e15413. [doi: 10.2196/15413] [Medline: 35084339]
28. Xiong C, D'Souza A, El-Khechen-Richandi G, et al. Perceptions of digital technology experiences and development among family caregivers and technology researchers: qualitative study. JMIR Form Res 2022 Jan 28;6(1):e19967. [doi: 10.2196/19967] [Medline: 35089150]
29. Kelley MM, Powell T, Camara D, et al. Mobile health apps, family caregivers, and care planning: scoping review. J Med Internet Res 2024 May 23;26:e46108. [doi: 10.2196/46108] [Medline: 38781588]
30. Premanandan S, Ahmad A, Cajander Å, Ågerfalk P, van Gemert-Pijnen L. Designing and evaluating IT applications for informal caregivers: scoping review. J Med Internet Res 2024 Oct 23;26:e57393. [doi: 10.2196/57393] [Medline: 39442167]
31. Ye B, Chu CH, Bayat S, Babineau J, How TV, Mihailidis A. Researched apps used in dementia care for people living with dementia and their informal caregivers: systematic review on app features, security, and usability. J Med Internet Res 2023 Oct 12;25(1):e46188. [doi: 10.2196/46188] [Medline: 37824187]
32. Egan KJ, Hodgson W, Dunlop MD, Imperatore G, Kirk A, Maguire R. A novel mobile app ("CareFit") to support informal caregivers to undertake regular physical activity from home during and beyond COVID-19 restrictions: co-design and prototype development study. JMIR Form Res 2021 Oct 1;5(10):e27358. [doi: 10.2196/27358] [Medline: 34406969]
33. Krick T, Huter K, Domhoff D, Schmidt A, Rothgang H, Wolf-Ostermann K. Digital technology and nursing care: a scoping review on acceptance, effectiveness and efficiency studies of informal and formal care technologies. BMC Health Serv Res 2019 Jun 20;19(1):400. [doi: 10.1186/s12913-019-4238-3] [Medline: 31221133]
34. Universidade da coruña, CITIC, talionis, lia2. CuidaconTIC platform. 2025. URL: https://cuidacontic.citic.udc.es [accessed 2024-12-12]
35. Groba B, Nieto-Riveiro L. Quality of life for carers through a person-centred technological solution (TechQoL4Carers). ClinicalTrials.gov. 2024. URL: https://clinicaltrials.gov/study/NCT06226285 [accessed 2025-03-26]
36. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ 2025 Apr 28;389:e081477. [doi: 10.1136/bmj-2024-081477] [Medline: 40294953]
37. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126. [doi: 10.2196/jmir.1923] [Medline: 22209829]
38. WHO data set–WHO Trial Registration Data Set (version 1.3.1). World Health Organization. URL: https://www.who.int/clinical-trials-registry-platform/network/who-data-set [accessed 2025-02-13]
39. Pozet A, Darnis S, Bonnet M, et al. Quality of life and needs in caregivers: results from the prospective multicentric open-label randomized study of informal caregivers of elderly patients. Int J Public Health 2023;68:1605459. [doi: 10.3389/ijph.2023.1605459] [Medline: 37711159]
40. EQ-5D-5L user guide basic information on how to use the EQ-5D-5L instrument. : EuroQol Research Foundation; 2019 URL: https://euroqol.org/wp-content/uploads/2023/11/EQ-5D-5LUserguide-23-07.pdf [accessed 2026-01-23]
41. Ramos-Goñi JM, Craig BM, Oppe M, et al. Handling data quality issues to estimate the Spanish EQ-5D-5L value set using a hybrid interval regression approach. Value Health 2018 May;21(5):596-604. [doi: 10.1016/j.jval.2017.10.023] [Medline: 29753358]
42. Spanish National Health Survey 2011/2012 [Web page in Spanish]. Spanish Ministry of Health. 2012. URL: https://www.sanidad.gob.es/estadEstudios/estadisticas/encuestaNacional/encuesta2011.htm [accessed 2025-01-26]
43. Brouwer WBF, van Exel NJA, van Gorp B, Redekop WK. The CarerQol instrument: a new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations. Qual Life Res 2006 Aug;15(6):1005-1021. [doi: 10.1007/s11136-005-5994-6] [Medline: 16900281]
44. Hoefman RJ, Van Exel NJA, Brouwer WBF. iVICQ iMTA Valuation of Informal Care Questionnaire. : Institute for Medical Technology Assessment (iMTA), Erasmus University Rotterdam; 2013 URL: https://www.imta.nl/assets/uploads/2022/01/iVICQ_UK_version_1.1.pdf [accessed 2026-01-23]
45. Hoefman RJ, van Exel J, Rose JM, van de Wetering EJ, Brouwer WBF. A discrete choice experiment to obtain a tariff for valuing informal care situations measured with the CarerQol instrument. Med Decis Making 2014 Jan;34(1):84-96. [doi: 10.1177/0272989X13492013] [Medline: 23771881]
46. Martín Carrasco M, Salvadó I, Nadal Álava S, et al. Adaptación para nuestro medio de la escala de sobrecarga del cuidador (Caregiver Burden Interview) de Zarit. Rev Gerontol 1996;6(4):338-346.
47. Martin-Carrasco M, Otermin P, Pérez-Camo V, et al. EDUCA study: psychometric properties of the Spanish version of the Zarit Caregiver Burden Scale. Aging Ment Health 2010 Aug;14(6):705-711. [doi: 10.1080/13607860903586094] [Medline: 20544413]

XSL•FO
RenderX

48. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980 Dec;20(6):649-655. [doi: 10.1093/geront/20.6.649] [Medline: 7203086]
49. Zarit SH, Anthony CR, Boutselis M. Interventions with care givers of dementia patients: comparison of two approaches. Psychol Aging 1987 Sep;2(3):225-232. [doi: 10.1037//0882-7974.2.3.225] [Medline: 3268213]
50. López Alonso SR, Moral Serrano MS. Validación del Indice de Esfuerzo del Cuidador en la población española [Article in Spanish]. Enferm Comun 2005;1(1):12-17 [FREE Full text]
51. Odriozola Gojenola M, Vita Garay A, Maiz Alkorta B, Zialtzeta Aduriz L, Bengoetxea Gallastegi L. Índice de esfuerzo del cuidador: test diagnóstico de sobrecarga en cuidadores de enfermos con demencia. Med Fam-SEMERGEN 2008 Nov;34(9):435-438. [doi: 10.1016/S1138-3593(08)75201-8]
52. Robinson BC. Validation of a caregiver strain index. J Gerontol 1983 May;38(3):344-348. [doi: 10.1093/geronj/38.3.344] [Medline: 6841931]
53. Peral-Gómez P, López-Roig S, Pastor-Mira M, et al. Cultural adaptation and psychometric properties of the Spanish version of the occupational balance questionnaire: an instrument for occupation-based research. Int J Environ Res Public Health 2021 Jul 14;18(14):7506. [doi: 10.3390/ijerph18147506] [Medline: 34299957]
54. Wagman P, Håkansson C. Exploring occupational balance in adults in Sweden. Scand J Occup Ther 2014 Nov;21(6):415-420. [doi: 10.3109/11038128.2014.934917] [Medline: 25100158]
55. Spreitzer GM. Psychological empowerment in the workplace: dimensions, measurement, and validation. Acad Manag J 1995 Oct 1;38(5):1442-1465. [doi: 10.2307/256865]
56. Spreitzer GM. Social structural characteristics of psychological empowerment. Acad Manag J 1996 Apr 1;39(2):483-504. [doi: 10.2307/256789]
57. Bretones FD, Jáimez MJ. Spanish adaptation and validation of the Psychological Empowerment Scale [Article in Spanish]. Interdisciplinaria 2022;39(1). [doi: 10.16888/interd.2022.39.1.12]
58. Xiaomi Smart Band 7 [Web page in Spanish]. Xiaomi España. 2024. URL: https://www.mi.com/es/product/xiaomi-smart-band-7/ [accessed 2024-09-27]
59. Concheiro-Moscoso P, Groba B, Martínez-Martínez FJ, et al. Study for the design of a protocol to assess the impact of stress in the ouality of life of workers. Int J Environ Res Public Health 2021 Feb 3;18(4):1413. [doi: 10.3390/ijerph18041413] [Medline: 33546392]
60. Njoroge J, Norman A, Reed D, Suh I. Identifying facets of technology satisfaction: measure development and application. J Learn High Educ 2012;8(2):7-17 [FREE Full text]
61. Brooke J. SUS: a “quick and dirty” usability scale. In: Usability Evaluation In Industry: CRC Press; 1996:4-7. [doi: 10.1201/9781498710411-35]
62. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
63. The R project for statistical computing. R Foundation for Statistical Computing. 2025. URL: https://www.r-project.org/ [accessed 2025-03-25]
64. Lau TY, Ong JL, Ng BKL, et al. Minimum number of nights for reliable estimation of habitual sleep using a consumer sleep tracker. Sleep Adv 2022;3(1):zpac026. [doi: 10.1093/sleepadvances/zpac026] [Medline: 37193398]
65. Hilden P, Schwartz JE, Pascual C, Diaz KM, Goldsmith J. How many days are needed? Measurement reliability of wearable device data to assess physical activity. PLoS ONE 2023;18(2):e0282162. [doi: 10.1371/journal.pone.0282162] [Medline: 36827427]
66. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
67. MAXQDA. URL: https://www.maxqda.com/ [accessed 2025-03-25]
68. Irfan B, Irfan O, Ansari A, Qidwai W, Nanji K. Impact of caregiving on various aspects of the lives of caregivers. Cureus 2017 May 2;9(5):e1213. [doi: 10.7759/cureus.1213] [Medline: 28589062]
69. Gerteis M, Edgman-Levitan S, Walker JD, Stoke DM, Cleary PD, Delbanco TL. What patients really want. Health Manage Q 1993;15(3):2-6. [Medline: 10129190]
70. Gerteis M, Edgman-Levitan S, Daley J, Delbanco TL, editors. Through the Patient’s Eyes: Understanding and Promoting Patient-Centered Care: Wiley; 2002.
71. Palm E. Who cares? Moral obligations in formal and informal care provision in the light of ICT-based home care. Health Care Anal 2013 Jun;21(2):171-188. [doi: 10.1007/s10728-011-0199-3] [Medline: 22179816]

Abbreviations

CarerQoL: care-related quality of life
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
PCC: person-centered care
QoL: quality of life

XSL•FO
RenderX

SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
VAS: visual analog scale

Edited by J Sarvestan; submitted 03.Nov.2025; this is a non–peer-reviewed article;accepted 12.Dec.2025; published 05.Feb.2026.

Please cite as:
Groba B, Concheiro-Moscoso P, Miranda-Duro MDC, Lagos-Rodríguez M, Pereira J, Nieto-Riveiro L
Enhancing Caregivers' Quality of Life Through a Web-Based Person-Centered Solution (TechQoL4Carers): Protocol for a Mixed Methods Pilot Trial
JMIR Res Protoc 2026;15:e86602
URL: https://www.researchprotocols.org/2026/1/e86602
doi:10.2196/86602

© Betania Groba, Patricia Concheiro-Moscoso, María del Carmen Miranda-Duro, Manuel Lagos-Rodríguez, Javier Pereira, Laura Nieto-Riveiro. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Alcohol and Sleep Sensors to Understand Blackout Risk in Young Adults' Natural Settings (The Lights Out Study): Protocol for an Intensive Longitudinal Pilot Study

Veronica L Richards[1,2], MPH, PhD; Ashlea Braun[1,2], RD, PhD; Michael R Sladek[3], PhD; Junru Zhao[1,4], MA, PhD; Thad Leffingwell[5], PhD; Sydney Newell Chesebro[1], BS; Julie M Croff[6], MPH, PhD

[1]TSET Health Promotion Research Center, Stephenson Cancer Center, University of Oklahoma Health Sciences, 4502 E 41st St, Tulsa, OK, United States

[2]Department of Health Promotion Sciences, Hudson College of Public Health, University of Oklahoma Health Sciences, Tulsa, OK, United States

[3]Department of Psychology, The University of Oklahoma, Norman, OK, United States

[4]Department of Pediatrics, College of Medicine, University of Oklahoma Health Sciences, Oklahoma City, OK, United States

[5]Department of Psychology, Oklahoma State University, Stillwater, OK, United States

[6]Department of Rural Health, Oklahoma State University Center for Health Sciences, Tulsa, OK, United States

Corresponding Author:
Veronica L Richards, MPH, PhD
TSET Health Promotion Research Center, Stephenson Cancer Center, University of Oklahoma Health Sciences, 4502 E 41st St, Tulsa, OK, United States

Abstract

Background: Alcohol-induced blackouts (AIBs) are a serious consequence of alcohol use that are strongly associated with experiencing excess alcohol−related harms. AIBs are common and recurrent among young adults who drink. The risk factors for AIBs include dynamics of alcohol use (quantity, speed, duration), alcohol-related behaviors (eg, playing drinking games, not using protective behavioral strategies), and factors related to the subjective experience of alcohol intoxication (eg, expectancies, motivations).

Objective: This study seeks to examine 2 modifiable behaviors that have been shown to impact both alcohol consumption and subjective experiences of intoxication and may therefore be associated with AIB risk: (1) other substance use and (2) sleep.

Methods: Approximately 50 participants will be recruited to participate in this study. Interested individuals will complete an online screening assessment, and those who are eligible (young adults who report recent heavy episodic drinking and AIBs) will be invited to an in-person baseline visit. At the baseline visit, participants will complete a baseline assessment, be fitted with a wrist-worn alcohol sensor (BACtrack Skyn) and a sleep or activity ring sensor (Oura ring), and receive training on the study protocol. Participants will complete a 14-day intensive data collection period consisting of twice daily scheduled mobile surveys and participant-initiated drinking surveys with hourly follow-ups. Participants will also wear the alcohol and sleep or activity sensors continuously during this 14-day period. After the intensive data collection period ends, participants will complete an in-person return visit to return their sensors, complete a follow-up survey, and receive compensation. The data will be processed and cleaned, and analyses will include multi-level structural equation models.

Results: This study was funded in July 2025. Data collection is projected to span January 2026 through June 2026.

Conclusions: This study seeks to understand 2 key modifiable behaviors that may be associated with increased AIB risk by leveraging multiple forms of innovative measurement. The integration of ecological momentary assessments with 2 sensors to capture alcohol use and sleep also supports potential applications in future digital interventions. This study will further enhance our preliminary data on the feasibility and acceptability of these methods, providing opportunities for conducting future research on a larger scale.

(JMIR Res Protoc 2026;15:e83980) doi:10.2196/83980

KEYWORDS

alcohol-induced blackout; alcohol use; cannabis; college student; ecological momentary assessment; heavy drinking; naturalistic research; nicotine; polysubstance use; sleep; wearables; young adult

XSL•FO
RenderX

Introduction

Alcohol-Induced Blackouts

Alcohol-induced blackouts (AIBs) are a serious consequence of alcohol use characterized by anterograde amnesia while the individual is still awake and interacting with their environment and those around them [1]. AIBs are common and recurrent in young adult drinkers, with approximately 50% reporting past-month AIBs and about 80% experiencing at least one during their 4 years of college (for students) [2,3]. Individuals who engage in heavy drinking and report recent AIBs are a particularly "risky" group of young adults who experience AIBs at a rate of about 1 in every 3 drinking days [4]. The frequency with which AIBs occur is of concern given their strong association with experiencing increased numbers of alcohol-related consequences, acutely (eg, sexual violence) and prospectively (eg, alcohol use disorder) [5-12].

A number of risk factors for experiencing AIBs have been identified. The dynamics of alcohol consumption that serve as AIB risk factors include the quantity of alcohol consumed, speed at which intoxication occurs, and the amount of time spent becoming intoxicated [4,13-16]. Behaviors that impact alcohol dynamics have also been identified as AIB risk factors, including pregaming, participating in drinking games, using fewer protective behavioral strategies, and modifying diet to intensify or offset alcohol ingestion [17-21]. Additional evidence suggests that the subjective experience of intoxication, including alcohol use expectancies, daily motivations to use, and the experience of intoxication (ie, how intoxicated one feels), may also serve as important AIB risk factors [22-27]. This study seeks to examine 2 modifiable behaviors that have been demonstrated to impact both alcohol consumption and subjective experiences of intoxication: (1) other substance use and (2) sleep.

Other Substance Use

Relative to single substance use, combined substance use has been associated with greater acute harms (eg, driving while intoxicated, overdose), chronic harms (eg, cancer), as well as heavier use [28-31]. About 20% of young adults who drink report combining alcohol with cannabis or nicotine [32]. The rates of cannabis and nicotine use among young adults continue to increase as markets for cannabis and nicotine (particularly e-cigarettes) rapidly evolve [33,34]. The effects of, and motives for, combined use vary by substance [35,36]. The effects of nicotine on subjective intoxication are mixed, but nicotine is consistently cited to offset the sedative effects of alcohol, allowing the individual to stay awake or out longer and drink more [31,37]. Cannabis is often combined with alcohol to enhance their effects [38] but may result in greater levels of cognitive impairment and impulsive decisions.

The effects of combined substance use on AIB risk have been mixed. Some studies report increased risk for AIBs when combining alcohol with cannabis [39,40] or nicotine [41], while others show no added risk associated with either [8,17,42]. In each of these studies, the timing of other substance use relative to alcohol use (eg, proximity to alcohol; order in which substances combined) may be a driving factor for such discrepancies. In the only study to consider same-day (ie, when alcohol and cannabis are used on the same day without overlapping effects) versus simultaneous alcohol-cannabis use (ie, when overlapping effects are experienced) on AIB risk, simultaneous use predicted AIB risk, but same-day use did not among college students [39].

Sleep

Approximately 20% - 30% of young adults report recent sleep disturbances [43,44], and ~65% report past-month sleep problems [45,46]. Poor sleep has been associated with heavier drinking the next day (higher peak intoxication and higher drink count; ie, within-person) [47] and on average (ie, between-person) [46,48]. The relationship between alcohol quantity and alcohol-related consequences is stronger among individuals who report worse sleep or insomnia, on average [46,48-50]. Only 1 study to date has been published on the topic of sleep and AIB risk, which showed stronger associations between alcohol and AIBs among heavy drinking college students with more severe insomnia (compared to less severe insomnia) [49]. Poor sleep quality has further been associated and found to interact with greater endorsement of drinking motives, resulting in increased consequences [44]. This may be explained by research indicating that poor sleep may negatively impact cognitive functioning [51,52], decrease inhibition [53], and impair decision-making [54]. Most of these studies have been cross-sectional examinations of college students.

Objectives

This study will expand on previous research investigating the associations between substance use, sleep, and AIBs in a sample of college and nonstudent young adults. The first aim is to examine the associations between combined alcohol and other substance use and AIBs. We hypothesize that simultaneous alcohol-cannabis use and simultaneous alcohol-nicotine use will be associated with increased odds of experiencing an AIB, compared to same-day (ie, alcohol and cannabis or nicotine use without overlap) and alcohol-only days. The second aim is to examine the associations between sleep (self-report quality and sensor-based measures; eg, rapid eye movement [REM] sleep) and AIBs. We hypothesize that days with worse past-night's sleep (relative to an individual's average sleep) will be associated with increased odds for AIBs. We will also explore the interaction between combined substance use and sleep on AIB risk via moderation.

Methods

Study Design

This is an intensive longitudinal study comprising an in-person baseline visit, 14-day intensive data collection period involving ecological momentary assessments (EMA) and sensor-wear regimen, and in-person return visit.

In-Person Baseline Visit

Participants will receive a reminder email the day before their scheduled baseline visit and a text message reminder the morning of the visit. The in-person baseline visit is estimated to take approximately 1 hour to complete. Upon arrival, study staff will obtain informed consent. Participants will then

XSL•FO
RenderX

complete a baseline survey on their demographics, recent alcohol and other substance use, alcohol-related consequences, sleep, and more (Table 1) via the Insight mHealth platform [55]. Following the completion of the baseline assessments, a trained study team member will provide an overview of study procedures and assist participants with setting up their devices. Participants will install 3 iOS apps on their personal iPhone (BACtrack Skyn, Oura, and Insight mHealth Platform). The study team member will assist the participant in selecting the correct sizes for each sensor (wrist and ring). The participant will be invited to join the "team" for each sensor's research platform (Skyn, Oura) to share their data. They will connect each sensor to their phones via Bluetooth. The study team member will provide a tutorial regarding how to use each device and the survey platform, including the schedule of surveys. Before leaving the baseline visit, participants will also sign up for a time or date to return their sensors and receive payment within 2 business days (whenever possible) of the end of the data collection period (14 d).

XSL•FO
RenderX

Table . Summary of items assessed for the Lights Out study.

Domain (measures)	Baseline	Scheduled EMAs[a]	Drinking EMAs
Sociodemographics			
Age[b], sex[b], race, ethnicity, height, weight, sexual orientation	✓		
Living arrangement, relationship status, children, student status[b], employment, income, social desirability	✓		
Proximity to study site[b], willingness to wear sensors[b]	✓		
Alcohol use			
Family history of alcohol use disorder [56]	✓		
Age of alcohol use onset, past-year alcohol use disorder identification test [57], past 3-month typical alcohol use[b] [57], past 3-month peak and duration, past 3-month alcohol-induced blackouts [49], past 3-month alcohol-related consequences [58]	✓		
Drink count[c], start and stop time of drinking[c], alcohol-induced blackouts[d] [49], alcohol-related consequences[d, e] [58,59]		✓	✓[f]
Attitudes toward drinking [60]	✓		
Subjective intoxication[c] [22]		✓	✓
Risky drinking behaviors (eg, using beer funnels, playing drinking games)[e]		✓	
Protective behavioral strategies[d, e] [61]		✓	
Situational drinking context (where and who drinking occurred with)		✓	✓
Willingness and expectancies to experience an alcohol-induced blackout[e]		✓	
Motives for drinking[e] [58]62		✓	
Other substance use			
Past 3-month other substance use (cannabis, nicotine, unprescribed stimulant medication, other)	✓		
Current use of other substances (cannabis, nicotine, unprescribed stimulant medication, other), modes of nicotine and cannabis use[f], timing of use relative to alcohol[e]		✓	✓
Cannabis-related consequences (adapted from Morin [63])		✓[e]	
Sleep			
Insomnia severity [63]	✓		

XSL•FO
RenderX

Domain (measures)	Baseline	Scheduled EMAs[a]	Drinking EMAs
Sleep quality[f] [64,65]	✓	✓	
Other measures			
Adverse childhood experiences [66]	✓		
Impulsivity [67]	✓		
Consumption of nonalcoholic beverages [68]		✓[g]	
Diet- and food-related behaviors[f] [69-72]	✓	✓	✓
Mood and energy levels		✓	✓

[a]EMA: ecological momentary assessment.

[b]Assessed only on the screening survey.

[c]Only assessed in the evening EMA if morning EMA was skipped.

[d]Adapted for daily use.

[e]Not assessed on evening EMA.

[f]Past hour drink counts only.

[g]Not assessed on morning EMA.

Intensive Data Collection Period

Overview

The intensive data collection period will begin the morning immediately following the baseline visit and will continue for 14 consecutive days. It will involve three primary components: (1) wearing the wrist-worn alcohol sensor, (2) wearing a finger-worn activity or sleep sensor, and (3) responding to brief surveys on a mobile device. Participants will be asked to wear the alcohol sensor and Oura ring continuously (including while asleep) except for while showering or charging the devices. Each day, participants will receive a reminder through the Insight mHealth Platform to sync their sensors with their respective apps. They will also receive a reminder every fourth day to charge their sensors. If technical issues arise with the sensors or apps, participants will be instructed to contact the study team. Participants will be instructed to continue to wear the devices experiencing the technical issue until they are able to meet with the study team. Technical issues will be addressed first internally and escalated to the respective company as needed. If the technical issue cannot be resolved immediately, devices will be exchanged for another device for the remainder of the study period. Broken devices will be exchanged as needed.

Alcohol Sensor

The alcohol sensor used in this study will be the BACtrack Skyn sensor. This sensor measures transdermal alcohol concentration (TAC) from the skin every 20 seconds, providing a (near) real-time, continuous, objective measure of alcohol intoxication. TAC is analogous and highly correlated with blood or breath alcohol concentration [73]. The sensors have a 10-day battery life and can be offline for extended periods without halting data collection or losing data.

Activity/Sleep Sensor

The Oura Ring will be used to assess sleep. Oura's sleep tracking algorithms show good agreement with polysomnography, with no differences in sleep onset latency, total sleep time, or wake after sleep onset [74]. The ring has a 7-day battery life, can be offline for extended periods without halting data collection or losing data, and is sweatproof and waterproof.

Surveys

All surveys will be administered through the Insight mHealth platform. If Insight experiences any technical issues, surveys will also be programmed in REDCap (Research Electronic Data Capture) and delivered to participants via text message. Participants will be sent 2 surveys daily (see Table 1 for measures) corresponding to the morning and evening. An adaptive random prompting schedule will be programmed to each participant's sleep-wake cycle. The morning survey (~5 - 10 min) will open 30 minutes after typical wake-up time, and the evening survey (~3 - 5 min) will open 7 hours after typical wake-up time. Each survey will be available for 3 hours, after which time the survey will be "locked" and considered missing. Participants will receive up to 2 reminders for each morning and evening survey at hourly intervals. Participants will also be asked to complete brief, self-initiated surveys when initiating alcohol use. Following the first drinking survey, participants will receive hourly follow-up drinking surveys. If the follow-up is not completed within 30 minutes, they will receive 2 additional reminders in 30-minute intervals. Follow-up surveys will end after the participant indicates that they are finished drinking, or they miss all prompts (initial follow-up prompt plus 2 reminders).

In-Person Return Visit

At the end of the data collection period, participants will receive an email reminder the day before their scheduled return and a text message reminder the morning of their return. They will

XSL•FO
RenderX

receive up to 3 reminders if they do not come in at their scheduled time or date. At the return visit, participants will return their sensors, complete a brief (~3 min) survey about their experience in the study, and receive compensation.

Setting and Participants

Recruitment will occur in Tulsa, Oklahoma, and surrounding areas. We will post flyers at businesses in the area (eg, retail stores, restaurants, professional office buildings, community centers). We may also post advertisements on social media platforms (eg, Instagram, Facebook). Flyers and advertisements will include a study email address to contact for more information on participating as well as a link to the study website.

The inclusion criteria for this study include (1) age of 18 - 25 years (college and nonstudent young adults); (2) report past 3-month behaviors that include drinking 4+/5+ drinks (women or men) at least once a week, on average; (3) report experiencing at least 1 AIB in the past 3 months; (4) own an iPhone with internet access (due to app requirements of the alcohol sensor); (5) willing to wear the wrist-worn alcohol sensor and activity or sleep sensor ring for 14 days; (6) live or work within a driving distance (~25 miles) of the study site in Tulsa; and (7) fluent in English. The exclusion criteria include (1) currently participating in another study involving sensor wear or EMA and (2) enrolled in high school.

Study staff will respond to emails from interested individuals with contact information for the project coordinator, a brief description of the research, and a secure link to the screening survey. By clicking on the link, participants will be directed to an informed consent statement describing study procedures. Consenting participants will be routed to a screening survey, estimated to take approximately 2 minutes to complete. The screening survey will assess each aspect of the eligibility or exclusion criteria via REDCap [75,76]. Eligible participants will be contacted to schedule the baseline visit via email and will use a Calendly link to schedule their in-person visit. If they are unable to be reached through email, they will be contacted via phone. Participants who do not meet eligibility criteria will be thanked for their time and told they are not eligible.

Sample Size

We plan to recruit up to 50 participants into this study. Sample size was determined based on feasibility (budget and timeline) related to the pilot funding mechanism. Effect sizes from this pilot will be used to determine appropriate sample sizes and statistical power for future proposals [77].

Ethical Considerations

Ethics approval was obtained from the institutional review board at the University of Oklahoma Health Sciences (#18534) in August 2025. Survey (implied) informed consent will be obtained prior to completing the screening survey. Written informed consent and HIPAA (Health Insurance Portability and Accountability Act) authorization will be obtained at the beginning of the in-person baseline visit. A Certificate of Confidentiality was obtained from the National Institutes of Health in September 2025. Participants can earn up to US $196 for completing all aspects of the study. Participants will be compensated US $20 for attending the baseline visit, with the opportunity of earning an additional US $20 if they attend the baseline visit within 7 days of being contacted about their eligibility. During the intensive data collection period, participants will earn US $2 per survey (morning and evening scheduled surveys only), equaling up to US $56 for perfect survey completion. Participants can earn an additional US $40 if they complete at least 80% of the daily surveys and wear their sensors at least 80% of the time (verified via temperature and movement sensors; see details about processing below). Participants will earn US $40 for returning their devices and an additional US $20 if they return their devices within 2 business days of completing the intensive data collection period.

This study is considered minimal risk. Psychological risks posed by the research are primarily related to the sensitivity of some of the measures, including thoughts, feelings, and personal difficulties that may be private. It is important to note that the data being collected via surveys or sensors are not diagnostic in nature and will not assess harm to self. Although the alcohol sensor collects (near) real-time measurement of alcohol use, it is not directly translatable to blood alcohol concentration nor immediately available to the study team. Neither survey responses nor sensor data will be monitored in real time. Participants are also asked to report on potentially illegal behaviors, such as drinking under the legal age or using illicit substances. Participants are encouraged to contact the investigators at any time to discuss concerns they might have. Upon request, participants will receive a handout with community resources in the area.

The following steps will be taken to protect individuals' identities as research participants. First, responses to web-based questionnaires will be identified only by a unique personal identification number, randomly generated for this study. Second, we will keep a master list of names, addresses, emails, phone numbers, and personal identification numbers so that we can compensate participants who complete the study and contact them about their participation. The master list will be stored separately from collected data on a password-protected computer with restricted access or on the university OneDrive. Only approved team members will have access to this list.

Planned Statistical Analyses: Processing of Intensive Longitudinal Data

Social Days

Data reduction strategies will be used so that the primary temporal unit of analysis for this study is the social day. The social day begins with the morning survey time instead of midnight. We will analyze our data using social days because drinking behavior is not encapsulated neatly in midnight-to-midnight "calendar" days. Drinking often either begins in the evening and extends past midnight or begins after midnight, but it is likely experienced as an extension of the previous day.

Processing of TAC Sensor Data

We will use previously published (and publicly available [78]) algorithms for processing intensive sensor data [4]. First, we

XSL•FO
RenderX

will apply algorithms to filter out observations in which the Skyn device was not worn. These algorithms are based on the Skyn's temperature and movement sensors. The device will be considered worn if: (1) the recorded temperature is over 28 °Cs, (2) the temperature is over 5° greater than the participant's minimum recorded temperature during the study period, or (3) the temperature is more than 3° greater than the participant's minimum recorded temperature and the motion sensor registers motion above 0.01 Gs. Observations indicative of nonwear (low temperature and no movement) will be removed. Second, we will smooth remaining TAC data to remove noise using a 30-minute moving average. The moving average will be centered using data from 15 minutes prior to 15 minutes after to smooth the series. This will result in a variable with a nearly identical mean and reduced variance that is highly correlated with the original. Third, we will bracket data into biological alcohol concentration "episodes": periods of time in which the TAC remained consecutively above a minimum threshold of 5 µg TAC/L air. New episodes will be marked by 2 TAC values: <5 µg TAC/L air followed by a TAC value ≥5 µg TAC/L air. Episode endings will be marked by a TAC value of 5 µg TAC/L air followed by 2 TAC values <5 µg TAC/L air. Fourth, we will filter out false-positive episodes, characteristically defined by the rates of rise and fall in the TAC that are biologically implausible or impossible. Episodes that are (1) shorter than 45 minutes, (2) shorter than 60 minutes and have a peak ≥400 µg TAC/L air, or (3) have rise or fall rates that do not fall between 20 and 300 will be removed. TAC features will then be extracted for each social day. TAC features include peak TAC, rise rate, rise duration, fall rate, total duration, and area under the curve. If at least 80% of the expected data per social day are present, but the features are missing, it will be assumed that no drinking occurred, and features will be recoded as 0 to represent a nondrinking day. Otherwise, data will be left as missing. Day-level TAC features data will be merged with other day-level data by participant ID and social date.

Processing of Oura Ring Sensor Data

Raw data are consolidated via application programming interface and processed using proprietary algorithms. Processed data will be accessed in the Oura Teams environment. The Oura Ring requires at least 3 hours of sleep stage detection during a 24-hour period (between 6 PM and 6 PM) to identify a main sleep cycle. Days without at least 3 hours of sleep stage detection will be considered missing. Potential reasons for missingness may be related to protocol compliance (eg, nonwear or not charging the device) or nights with high levels of movement. Critical components of sleep architecture will be extracted by the social day level using Oura Teams processing pipelines, including sleep latency (minutes to fall asleep), total time and percentage of time in REM sleep, total time and percentage of time in non-REM sleep, total sleep time, total time and percentage of time in deep sleep, and restfulness of sleep. Secondary measures of intoxication (eg, heart rate variability, respiration rate) will be explored. Additional measures related to physical activity (eg, calories, activity level) will also be collected and may be explored in secondary analyses. Day-level sleep data will be merged with other day-level data by participant ID and social date.

General Modeling Framework

The main outcome of interest in this study is AIBs experienced during the 14-day intensive data collection period. To test our primary hypotheses, each main predictor (combined substance use and sleep) will be examined in separate logistic multi-level structural equation models. Data that are characterized by repeated observations of momentary events will be summarized at the day level as Level 1 (within-person) variables, which will be nested within each participant. Person-level (between-person) differences will be Level 2 variables. Level 1 variables will include the assessments of alcohol (TAC and self-report) and other substance use, sleep, and so forth. Level 2 variables include person-level factors, such as demographics. Our analyses use this nested structure to separate estimates of within-person relationships (day-level) from between-person relationships. Level 1 variables will be person-mean centered, and Level 2 variables will be grand-mean centered. Random paths will be allowed, and Bayesian methods will be used if maximum likelihood models do not converge. Indirect effects (the effects of combined substance use or sleep through alcohol use) will be estimated using either (1) bias-corrected bootstrapped confidence intervals (maximum likelihood) or (2) the posterior distribution of the indirect effect (Bayesian). Models with confidence or credible intervals that do not contain 1 will be considered significant. Covariates will include sex, age, race or ethnicity, and student status.

Pre hoc sensitivity analyses replacing sensor-collected data with self-reported data for alcohol use (eg, drink count) and sleep (eg, sleep quality) will be conducted. For exploratory analyses, moderation will be tested by including interaction terms between Level 1 combined substance use and sleep.

Results

This study was funded by the Presbyterian Health Foundation in July 2025. Data collection is projected to span January 2026 through June 2026.

Discussion

Expected Findings

This study seeks to understand 2 key modifiable behaviors that may be associated with increased AIB risk by leveraging multiple forms of innovative measurement. We anticipate that nights with simultaneous use of alcohol and cannabis, and simultaneous use of alcohol and nicotine, will be associated with increased odds of experiencing an AIB (aim 1). We also anticipate that drinking days preceded by worse sleep will be associated with increased odds of experiencing an AIB (aim 2).

Previous research has begun to explore the impact of other substance use [8,17,39-42] and to a lesser extent, sleep [49], on AIBs and other alcohol-related consequences. This study expands on previous research in a number of ways. First, we will include both college and nonstudent young adults. Nearly all of the research on AIBs has focused on college student samples. According to the alcohol-harm paradox, people with lower socioeconomic status (SES) tend to experience greater alcohol-related harm than those with higher SES, despite often

XSL•FO
RenderX

reporting similar (or less) levels of alcohol consumption [79]. Nonstudents are more likely to be of lower SES than those who attend college [80] and may be at increased risk for experiencing alcohol-related harm, in part due to a lack of access to university health resources. Research inclusive of nonstudent young adults is critical to informing alcohol interventions to reduce harm outside of college campuses. Second, we will integrate 2 objective sensors (to measure alcohol use and sleep) with EMA to overcome the limitations of traditional self-report (eg, less reliance on memory). Our approach will allow us to collect objective, fine-grained information on measures of alcohol use (eg, rate of and duration of intoxication) and sleep (eg, REM vs non-REM sleep) otherwise difficult (or impossible) to obtain via self-report. Our use of EMA will also allow us to assess key factors immediately preceding and during a drinking episode. These advanced methodological tools also lend themselves to use in future digital interventions. Each of these tools allows for (near) real-time assessment of our main predictors and outcomes of interest. The use of sensors reduces reliance on self-report, which is frequently inaccurate in inconsistent patterns (ie, heavier levels of drinking often observed in young adults and that are associated with AIBs and examination of measured sleep architecture, including REM sleep), both directly related to the outcome of interest. These methods allow the virtual observation and enhanced understanding of behaviors in real-world contexts.

Limitations

Despite its many contributions, this study has several limitations. First, the intensive data collection period consists of a single 14-day period. It is possible that this period may not be representative of typical behaviors. Second, only about 50 young adults in a single metropolitan area will be enrolled in the study. Third, it is possible that the alcohol sensors may miss lower-level drinking episodes or the sleep sensors may not be perfectly accurate at differentiating stages of sleep.

Conclusions

This study will inform efforts aimed at preventing AIBs and reducing alcohol-related harms. The findings from this study will serve as preliminary data on the feasibility and acceptability of the protocol, providing opportunities for conducting future research on a larger scale.

Funding

This work was supported by the Presbyterian Health Foundation (seed grant), the Oklahoma Tobacco Settlement Endowment Trust (contract #00003615), the University of Oklahoma Health Stephenson Cancer Center via a National Cancer Institute Cancer Center Support Grant (P30CA225520), and the Oklahoma State University Center for Health Sciences via a National Institute of General Medical Sciences grant to support the Center for Integrative Health on Childhood Adversity (P20GM109097).

Conflicts of Interest

None declared.

Peer Review Report 1
Peer review report by the Presbyterian Health Foundation (PHF) Seed Grant Review Committee.
[PDF File, 156 KB - resprot_v15i1e83980_app1.pdf]

References

1. White AM. What happened? Alcohol, memory blackouts, and the brain. Alcohol Res Health 2003;27(2):186-196. [Medline: 15303630]
2. Barnett NP, Clerkin EM, Wood M, et al. Description and predictors of positive and negative alcohol-related consequences in the first year of college. J Stud Alcohol Drugs 2014 Jan;75(1):103-114. [doi: 10.15288/jsad.2014.75.103] [Medline: 24411802]
3. Glenn SD, Turrisi R, Waldron KA, Mallett KA, Russell MA, Reavy RR. Examining the impact of early college experiences on the cumulative number of alcohol-related consequences. Addict Behav 2022 Sep;132:107357. [doi: 10.1016/j.addbeh.2022.107357] [Medline: 35633615]
4. Richards VL, Glenn SD, Turrisi RJ, Mallett KA, Ackerman S, Russell MA. Transdermal alcohol concentration features predict alcohol-induced blackouts in college students. Alcohol Clin Exp Res (Hoboken) 2024 May;48(5):880-888. [doi: 10.1111/acer.15290] [Medline: 38639884]
5. Richards VL, Glenn SD, Turrisi RJ, Altstaedter A, Mallett KA, Russell MA. Does it really matter that I do not remember my night? Consequences related to blacking out among college student drinkers. Alcohol Clin Exp Res (Hoboken) 2023 Sep;47(9):1798-1805. [doi: 10.1111/acer.15157] [Medline: 37533347]
6. Glenn SD, Turrisi RJ, Richards VL, Russell MA, Mallett KA. A dual-process decision-making model examining the longitudinal associations between alcohol-induced blackouts and alcohol use disorder risk among college student drinkers. J Stud Alcohol Drugs 2024 Jan;85(1):73-83. [doi: 10.15288/jsad.23-00063] [Medline: 37768675]
7. Hingson R, Zha W, Simons-Morton B, White A. Alcohol-induced blackouts as predictors of other drinking related harms among emerging young adults. Alcohol Clin Exp Res 2016 Apr;40(4):776-784. [doi: 10.1111/acer.13010] [Medline: 27012148]

XSL•FO
RenderX

8. Merrill JE, Boyle HK, Jackson KM, Carey KB. Event-level correlates of drinking events characterized by alcohol-induced blackouts. Alcohol Clin Exp Res 2019 Dec;43(12):2599-2606. [doi: 10.1111/acer.14204] [Medline: 31557348]
9. Mundt MP, Zakletskaia LI, Brown DD, Fleming MF. Alcohol-induced memory blackouts as an indicator of injury risk among college drinkers. Inj Prev 2012 Feb;18(1):44-49. [doi: 10.1136/ip.2011.031724] [Medline: 21708813]
10. Voloshyna DM, Bonar EE, Cunningham RM, Ilgen MA, Blow FC, Walton MA. Blackouts among male and female youth seeking emergency department care. Am J Drug Alcohol Abuse 2018;44(1):129-139. [doi: 10.1080/00952990.2016.1265975] [Medline: 28032801]
11. Studer J, Gmel G, Bertholet N, Marmet S, Daeppen JB. Alcohol-induced blackouts at age 20 predict the incidence, maintenance and severity of alcohol dependence at age 25: a prospective study in a sample of young Swiss men. Addiction 2019 Sep;114(9):1556-1566. [doi: 10.1111/add.14647] [Medline: 31059161]
12. Valenstein-Mah H, Larimer M, Zoellner L, Kaysen D. Blackout drinking predicts sexual revictimization in a college sample of binge-drinking women. J Trauma Stress 2015 Oct;28(5):484-488. [doi: 10.1002/jts.22042] [Medline: 26401899]
13. Carpenter RW, Merrill JE. How much and how fast: alcohol consumption patterns, drinking-episode affect, and next-day consequences in the daily life of underage heavy drinkers. Drug Alcohol Depend 2021 Jan 1;218:108407. [doi: 10.1016/j.drugalcdep.2020.108407] [Medline: 33257198]
14. Evans-Polce RJ, Stevenson BL, Patrick ME. Daily-level analysis of drinking intensity and acute physical consequences. Addict Behav 2022 May;128:107246. [doi: 10.1016/j.addbeh.2022.107246] [Medline: 35065367]
15. Mallett KA, Varvil-Weld L, Turrisi R, Read A. An examination of college students' willingness to experience consequences as a unique predictor of alcohol problems. Psychol Addict Behav 2011 Mar;25(1):41-47. [doi: 10.1037/a0021494] [Medline: 21219039]
16. Rose ME, Grant JE. Alcohol-induced blackout. Phenomenology, biological basis, and gender differences. J Addict Med 2010 Jun;4(2):61-73. [doi: 10.1097/ADM.0b013e3181e1299d] [Medline: 21769024]
17. Richards VL, Turrisi RJ, Glenn SD, et al. Alcohol-induced blackouts among college student drinkers: a multilevel analysis. Addict Behav 2023 Aug;143:107706. [doi: 10.1016/j.addbeh.2023.107706] [Medline: 37001258]
18. LaBrie JW, Hummer J, Kenney S, Lac A, Pedersen E. Identifying factors that increase the likelihood for alcohol-induced blackouts in the prepartying context. Subst Use Misuse 2011;46(8):992-1002. [doi: 10.3109/10826084.2010.542229] [Medline: 21222521]
19. Wahl S, Sonntag T, Roehrig J, Kriston L, Berner MM. Characteristics of predrinking and associated risks: a survey in a sample of German high school students. Int J Public Health 2013 Apr;58(2):197-205. [doi: 10.1007/s00038-012-0409-2] [Medline: 23007874]
20. Ray AE, Stapleton JL, Turrisi R, Mun EY. Drinking game play among first-year college student drinkers: an event-specific analysis of the risk for alcohol use and problems. Am J Drug Alcohol Abuse 2014 Sep;40(5):353-358. [doi: 10.3109/00952990.2014.930151] [Medline: 25192202]
21. Simons RM, Hansen JM, Simons JS, Hovrud L, Hahn AM. Drunkorexia: normative behavior or gateway to alcohol and eating pathology? Addict Behav 2021 Jan;112:106577. [doi: 10.1016/j.addbeh.2020.106577] [Medline: 32861988]
22. Richards VL, Turrisi RJ, Russell MA. Subjective intoxication predicts alcohol-related consequences at equivalent alcohol concentrations in young adults using ecological momentary assessment and alcohol sensors. Psychol Addict Behav 2024 May;38(3):334-346. [doi: 10.1037/adb0000993] [Medline: 38271080]
23. Marino EN, Fromme K. Alcohol-induced blackouts, subjective intoxication, and motivation to decrease drinking: prospective examination of the transition out of college. Addict Behav 2018 May;80:89-94. [doi: 10.1016/j.addbeh.2018.01.013] [Medline: 29367115]
24. Miller MB, Davis CN, Merrill JE, DiBello AM, Carey KB. Intentions and motives to experience alcohol-induced blackout among young adults in college. Psychol Addict Behav 2020 Sep;34(6):690-698. [doi: 10.1037/adb0000572] [Medline: 32162962]
25. Ward RM, Geyer R, Cleveland M, Perlman E, Messman T. Examining alcohol-related blackouts and drinking motives over time among college women. J Stud Alcohol Drugs 2024 Jan;85(1):84-91. [doi: 10.15288/jsad.22-00285] [Medline: 37650843]
26. Merrill JE, Boyle HK, Peterson R, et al. Factors that increase risk for alcohol-induced blackouts in high-intensity drinking young adults. Alcohol Clin Exp Res (Hoboken) 2025 Jul;49(7):1530-1541. [doi: 10.1111/acer.70074] [Medline: 40395070]
27. Litt DM, Zhou Z, Fairlie AM, et al. A daily-level examination of willingness to pregame, blackout, and hook up across drinking days: associations with respective behavioral outcomes among adolescents and young adults. Addict Behav 2024 Jun;153:108004. [doi: 10.1016/j.addbeh.2024.108004] [Medline: 38457988]
28. Bailey AJ, McHugh RK. Why do we focus on the exception and not the rule? Examining the prevalence of mono- versus polysubstance use in the general population. Addiction 2023 Oct;118(10):2026-2029. [doi: 10.1111/add.16290] [Medline: 37438998]
29. Moss HB, Chen CM, Yi HY. Early adolescent patterns of alcohol, cigarettes, and marijuana polysubstance use and young adult substance use outcomes in a nationally representative sample. Drug Alcohol Depend 2014 Mar 1;136:51-62. [doi: 10.1016/j.drugalcdep.2013.12.011] [Medline: 24434016]

XSL•FO
RenderX

30. Karamouzian M, Cui Z, Hayashi K, et al. Longitudinal polysubstance use patterns and non-fatal overdose: a repeated measures latent class analysis. Int J Drug Policy 2025 Jul;141:104301. [doi: 10.1016/j.drugpo.2023.104301] [Medline: 38182524]
31. McKee SA, Weinberger AH. How can we use our knowledge of alcohol-tobacco interactions to reduce alcohol use? Annu Rev Clin Psychol 2013;9(1):649-674. [doi: 10.1146/annurev-clinpsy-050212-185549] [Medline: 23157448]
32. Cohn AM, Johnson AL, Rath JM, Villanti AC. Patterns of the co-use of alcohol, marijuana, and emerging tobacco products in a national sample of young adults. Am J Addict 2016 Dec;25(8):634-640. [doi: 10.1111/ajad.12456] [Medline: 27706885]
33. Key substance use and mental health indicators in the United States: results from the 2024 National Survey on Drug Use and Health. : Substance Abuse and Mental Health Services Administration; 2024 URL: https://www.samhsa.gov/data/sites/default/files/reports/rpt47095/National%20Report/National%20Report/2023-nsduh-annual-national.pdf [accessed 2025-12-12]
34. Boakye E, Osuji N, Erhabor J, et al. Assessment of patterns in e-cigarette use among adults in the US, 2017-2020. JAMA Netw Open 2022 Jul 1;5(7):e2223266. [doi: 10.1001/jamanetworkopen.2022.23266] [Medline: 35867060]
35. Teter CJ, McCabe SE, Cranford JA, Boyd CJ, Guthrie SK. Prevalence and motives for illicit use of prescription stimulants in an undergraduate student sample. J Am Coll Health 2005;53(6):253-262. [doi: 10.3200/JACH.53.6.253-262] [Medline: 15900989]
36. Stevens AK, Drohan MM, Boyle HK, White HR, Jackson KM. More reasons, more use and problems? Examining the influence of number of motives on consumption and consequences across alcohol-only, cannabis-only, and simultaneous-use days. J Stud Alcohol Drugs 2021 Nov;82(6):782-791. [doi: 10.15288/jsad.2021.82.782] [Medline: 34762038]
37. Verplaetse TL, McKee SA. An overview of alcohol and tobacco/nicotine interactions in the human laboratory. Am J Drug Alcohol Abuse 2017 Mar;43(2):186-196. [doi: 10.1080/00952990.2016.1189927] [Medline: 27439453]
38. Patrick ME, Fairlie AM, Lee CM. Motives for simultaneous alcohol and marijuana use among young adults. Addict Behav 2018 Jan;76:363-369. [doi: 10.1016/j.addbeh.2017.08.027] [Medline: 28915500]
39. Jackson KM, Sokolovsky AW, Gunn RL, White HR. Consequences of alcohol and marijuana use among college students: prevalence rates and attributions to substance-specific versus simultaneous use. Psychol Addict Behav 2020 Mar;34(2):370-381. [doi: 10.1037/adb0000545] [Medline: 31944787]
40. Davis CN, Dash GF, Miller MB, Slutske WS. Past year high-intensity drinking moderates the association between simultaneous alcohol and marijuana use and blackout frequency among college students. J Am Coll Health 2023 Jan;71(1):140-146. [doi: 10.1080/07448481.2021.1880415] [Medline: 33577429]
41. Mallett KA, Turrisi R, Hultgren BA, Sell N, Reavy R, Cleveland M. When alcohol is only part of the problem: an event-level analysis of negative consequences related to alcohol and other substance use. Psychol Addict Behav 2017 May;31(3):307-314. [doi: 10.1037/adb0000260] [Medline: 28182448]
42. Richards VL, Mallett KA, Turrisi RJ, Oliver JA, Croff JM, Russell MA. Effects of nicotine-alcohol co-use on alcohol-induced blackouts and other alcohol-related consequences in college students. Psychopharmacology (Berl) 2025 Dec;242(12):2757-2767. [doi: 10.1007/s00213-025-06830-x] [Medline: 40493073]
43. Grandner MA, Martin JL, Patel NP, et al. Age and sleep disturbances among American men and women: data from the U.S. Behavioral Risk Factor Surveillance System. Sleep 2012 Mar 1;35(3):395-406. [doi: 10.5665/sleep.1704] [Medline: 22379246]
44. Kenney SR, Paves AP, Grimaldi EM, LaBrie JW. Sleep quality and alcohol risk in college students: examining the moderating effects of drinking motives. J Am Coll Health 2014;62(5):301-308. [doi: 10.1080/07448481.2014.897953] [Medline: 24588270]
45. Lund HG, Reider BD, Whiting AB, Prichard JR. Sleep patterns and predictors of disturbed sleep in a large population of college students. J Adolesc Health 2010 Feb;46(2):124-132. [doi: 10.1016/j.jadohealth.2009.06.016] [Medline: 20113918]
46. Kenney SR, LaBrie JW, Hummer JF, Pham AT. Global sleep quality as a moderator of alcohol consumption and consequences in college students. Addict Behav 2012 Apr;37(4):507-512. [doi: 10.1016/j.addbeh.2012.01.006] [Medline: 22285119]
47. Reichenberger D, Chang AM, Russell M. 0005 Bidirectional associations of sleep and alcohol use within and between regularly drinking young adults. Sleep 2022 May 25;45(Supplement_1):A2-A2. [doi: 10.1093/sleep/zsac079.004]
48. Miller MB, DiBello AM, Lust SA, Carey MP, Carey KB. Adequate sleep moderates the prospective association between alcohol use and consequences. Addict Behav 2016 Dec;63:23-28. [doi: 10.1016/j.addbeh.2016.05.005] [Medline: 27395437]
49. Miller MB, Boness CL, DiBello AM, Froeliger B. Insomnia as a moderator of alcohol use and blackout: potential role in acute physiological consequences. Addict Behav 2022 Nov;134:107395. [doi: 10.1016/j.addbeh.2022.107395] [Medline: 35696822]
50. Miller MB, DiBello AM, Carey KB, Pedersen ER. Insomnia moderates the association between alcohol use and consequences among young adult veterans. Addict Behav 2017 Dec;75:59-63. [doi: 10.1016/j.addbeh.2017.06.020] [Medline: 28704711]
51. Benitez A, Gunstad J. Poor sleep quality diminishes cognitive functioning independent of depression and anxiety in healthy young adults. Clin Neuropsychol 2012;26(2):214-223. [doi: 10.1080/13854046.2012.658439] [Medline: 22335237]
52. Brownlow JA, Miller KE, Gehrman PR. Insomnia and cognitive performance. Sleep Med Clin 2020 Mar;15(1):71-76. [doi: 10.1016/j.jsmc.2019.10.002] [Medline: 32005351]
53. Anderson C, Platten CR. Sleep deprivation lowers inhibition and enhances impulsivity to negative stimuli. Behav Brain Res 2011 Mar 1;217(2):463-466. [doi: 10.1016/j.bbr.2010.09.020] [Medline: 20888369]

XSL•FO
RenderX

54. Killgore WDS, Balkin TJ, Wesensten NJ. Impaired decision making following 49 h of sleep deprivation. J Sleep Res 2006 Mar;15(1):7-13. [doi: 10.1111/j.1365-2869.2006.00487.x] [Medline: 16489997]
55. InsightTM mHealth platform. TSET Health Promotion Research. URL: https://healthpromotionresearch.org/Mobile-Health-Technology [accessed 2025-12-12]
56. Rice JP, Reich T, Bucholz KK, et al. Comparison of direct interview and family history diagnoses of alcohol dependence. Alcohol Clin Exp Res 1995 Aug;19(4):1018-1023. [doi: 10.1111/j.1530-0277.1995.tb00983.x] [Medline: 7485811]
57. Collins RL, Parks GA, Marlatt GA. Social determinants of alcohol consumption: the effects of social interaction and model status on the self-administration of alcohol. J Consult Clin Psychol 1985 Apr;53(2):189-200. [doi: 10.1037/0022-006X.53.2.189] [Medline: 3998247]
58. Kahler CW, Strong DR, Read JP. Toward efficient and comprehensive measurement of the alcohol problems continuum in college students: the brief young adult alcohol consequences questionnaire. Alcohol Clin Exp Res 2005 Jul;29(7):1180-1189. [doi: 10.1097/01.alc.0000171940.95813.a5] [Medline: 16046873]
59. Patrick ME, Maggs JL. Short-term changes in plans to drink and importance of positive and negative alcohol consequences. J Adolesc 2008 Jun;31(3):307-321. [doi: 10.1016/j.adolescence.2007.06.002] [Medline: 17651796]
60. DiBello AM, Miller MB, Neighbors C, Reid A, Carey KB. The relative strength of attitudes versus perceived drinking norms as predictors of alcohol use. Addict Behav 2018 May;80:39-46. [doi: 10.1016/j.addbeh.2017.12.022] [Medline: 29329007]
61. Treloar H, Martens MP, McCarthy DM. The Protective Behavioral Strategies Scale-20: improved content validity of the Serious Harm Reduction subscale. Psychol Assess 2015 Mar;27(1):340-346. [doi: 10.1037/pas0000071] [Medline: 25558969]
62. Grant VV, Stewart SH, O'Connor RM, Blackwell E, Conrod PJ. Psychometric evaluation of the five-factor Modified Drinking Motives Questionnaire−Revised in undergraduates. Addict Behav 2007 Nov;32(11):2611-2632. [doi: 10.1016/j.addbeh.2007.07.004] [Medline: 17716823]
63. Morin CM. Insomnia: Psychological Assessment and Management: Guilford Press; 1993. URL: https://psycnet.apa.org/record/1993-98362-000 [accessed 2025-12-12]
64. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
65. Haythornthwaite JA, Hegel MT, Kerns RD. Development of a sleep diary for chronic pain patients. J Pain Symptom Manage 1991 Feb;6(2):65-72. [doi: 10.1016/0885-3924(91)90520-e] [Medline: 2007794]
66. Ford DC, Merrick MT, Parks SE, et al. Examination of the factorial structure of adverse childhood experiences and recommendations for three subscale scores. Psychol Violence 2014 Oct;4(4):432-444. [doi: 10.1037/a0037723] [Medline: 26430532]
67. Zuckerman M, Kuhlman DM, Joireman J, Teta P, Kraft M. A comparison of three structural models for personality: the Big Three, the Big Five, and the Alternative Five. J Pers Soc Psychol 1993 Oct;65(4):757-768. [doi: 10.1037/0022-3514.65.4.757]
68. Bowdring MA, McCarthy DM, Fairbairn CE, Prochaska JJ. Non-alcoholic beverage consumption among US adults who consume alcohol. Addiction 2024 Jun;119(6):1080-1089. [doi: 10.1111/add.16452] [Medline: 38403280]
69. Rahal CJ, Bryant JB, Darkes J, Menzel JE, Thompson JK. Development and validation of the Compensatory Eating and Behaviors in Response to Alcohol Consumption Scale (CEBRACS). Eat Behav 2012 Apr;13(2):83-87. [doi: 10.1016/j.eatbeh.2011.11.001] [Medline: 22365787]
70. Huang Y, Xu YX, Shen YT, et al. Sex-specific association between later circadian timing of food intake and adiposity among Chinese young adults living in real-world settings. Br J Nutr 2024 Dec 28;132(12):1629-1636. [doi: 10.1017/S0007114524001636] [Medline: 39573830]
71. Elsworth RL, Hinton EC, Flynn AN, et al. Development of Momentary Appetite Capture (MAC): a versatile tool for monitoring appetite over long periods. Appetite 2024 Mar 1;194:107154. [doi: 10.1016/j.appet.2023.107154] [Medline: 38081544]
72. Rijnaarts I, de Roos N, Zoetendal EG, de Wit N, Witteman BJM. Development and validation of the FiberScreen: a short questionnaire to screen fibre intake in adults. J Hum Nutr Diet 2021 Dec;34(6):969-980. [doi: 10.1111/jhn.12941] [Medline: 34378249]
73. Fairbairn CE, Kang D, Bosch N. Using machine learning for real-time BAC estimation from a new-generation transdermal biosensor in the laboratory. Drug Alcohol Depend 2020 Nov 1;216:108205. [doi: 10.1016/j.drugalcdep.2020.108205] [Medline: 32853998]
74. de Zambotti M, Rosas L, Colrain IM, Baker FC. The sleep of the ring: comparison of the ŌURA sleep tracker against polysomnography. Behav Sleep Med 2019;17(2):124-136. [doi: 10.1080/15402002.2017.1300587] [Medline: 28323455]
75. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
76. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]

XSL•FO
RenderX

77. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022 Jan 1;60(1):95-103. [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
78. Richards VL, Russell MA. TACLion. Zenodo. 2025. URL: https://zenodo.org/records/17554029 [accessed 2026-01-07]
79. Bloomfield K. Understanding the alcohol-harm paradox: what next? Lancet Public Health 2020 Jun;5(6):e300-e301. [doi: 10.1016/S2468-2667(20)30119-5] [Medline: 32504581]
80. Young adult educational and employment outcomes by family socioeconomic status. : National Center for Education Statistics (NCES), U.S. Department of Education; 2022 URL: https://nces.ed.gov/programs/coe/pdf/coe_tbe.pdf [accessed 2025-12-12]

Abbreviations

AIB: alcohol-induced blackout
EMA: ecological momentary assessment
HIPAA: Health Insurance Portability and Accountability Act
REDCap: Research Electronic Data Capture
REM: rapid eye movement
SES: socioeconomic status
TAC: transdermal alcohol concentration

Edited by J Sarvestan; submitted 17.Sep.2025; this is a non–peer-reviewed article;accepted 02.Dec.2025; published 21.Jan.2026.

Please cite as:
Richards VL, Braun A, Sladek MR, Zhao J, Leffingwell T, Newell Chesebro S, Croff JM
Using Alcohol and Sleep Sensors to Understand Blackout Risk in Young Adults' Natural Settings (The Lights Out Study): Protocol for an Intensive Longitudinal Pilot Study
JMIR Res Protoc 2026;15:e83980
URL: https://www.researchprotocols.org/2026/1/e83980
doi:10.2196/83980

© Veronica L Richards, Ashlea Braun, Michael R Sladek, Junru Zhao, Thad Leffingwell, Sydney Newell Chesebro, Julie M Croff. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Virtual Living Lab Platform Codeveloped for Mental Health in Youth-Onset Type 2 Diabetes (BrightSpark Care Lab): Protocol for a Mixed Methods Study

Mandy M. Archibald[1,2], BSCN, PhD; Brandy Wicklow[2,3], MSc, MD; Elizabeth Sellers[2,3], MSc, MD; Arlene Griffiths[4], BSCN; Linda Diffey[5], MSc, PhD; Jonathan McGavock[2,3], PhD; Leslie E. Roos[6,7], PhD; Alexander M. Clark[8], BSCN, PhD; Jennifer Lopez[9], Dip; Josephine Ho[10,11], MSc, MD; Ernestine Ledoux[4], BA; Gifty Dzorka[1], PhD; Shahina Parvin[1], PhD; Oluwatoyosi Fagbuyi[1], MN; Allison Dart[2,3], MSc, MD

[1]College of Nursing, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

[2]Diabetes Research Envisioned and Accomplished in Manitoba, Children's Hospital Research Institute of Manitoba, Winnipeg, MB, Canada

[3]Department of Pediatrics and Child Health, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada

[4]Community Representative, Winnipeg, MB, Canada

[5]Rady Faculty of Health Sciences, College of Community and Global Health, University of Manitoba, Winnipeg, MB, Canada

[6]Partnering for Research Innovations in Mental Health, Children's Hospital Research Institute of Manitoba, Winnipeg, MB, Canada

[7]Department of Psychology, Faculty of Arts, University of Manitoba, Winnipeg, MB, Canada

[8]Office of the President, University of Athabasca, Athabasca, AB, Canada

[9]Youth Co-researcher, Winnipeg, MB, Canada

[10]Department of Paediatrics, Cummings School of Medicine, University of Calgary, Calgary, AB, Canada

[11]Diabetes Clinic, Alberta Children's Hospital Research Institute, Calgary, AB, Canada

Corresponding Author:
Mandy M. Archibald, BSCN, PhD
College of Nursing
Rady Faculty of Health Sciences
University of Manitoba
Office 317 Helen Glass, 89 Curry Place
Winnipeg, MB, R3T 2N2
Canada
Phone: 1 204 474 6660
Email: mandy.archibald@umanitoba.ca

Abstract

Background: Type 2 diabetes (T2D) is a complex chronic disease that poses significant mental health challenges to affected youth. Despite calls for youth-centered research in this area, qualitative and mixed methods research is lacking, and longitudinal understandings of the mental health experiences of youth have not been generated. Living labs have potential as interactive knowledge exchange and longitudinal research platforms to generate such understandings.

Objective: The proposed research aims to (1) codesign, with youth and parent coresearchers, a virtual living lab platform with an embedded registry of youth with T2D; (2) use this platform to generate longitudinal understandings of youths' mental health experiences; (3) identify youth priorities for research and care based on the thematic data; and (4) codesign an arts-based knowledge translation (KT) resource to communicate these priorities.

Methods: This study proposes a three-stage longitudinal, qualitatively dominant, convergent mixed methods design. Stage 1 involved codesigning an online platform with youth and parent coresearchers over a 4-month period and establishing a user registry of English-speaking youth (age 10-25 years) with T2D (diagnosed at age 18 years or younger) in Canada and their parents or guardians. In stage 2, up to 50 youth were purposively selected from the registry to complete baseline mental health measures, followed by 12 content modules using diary and arts-based response methods. Inductive thematic and mixed methods analysis will inform stage 3. Up to a third of the stage 3 participants will be purposively selected to identify priorities for mental health research and care and codesign arts-based KT resources to impart critical research findings to stakeholder groups identified with

XSL•FO
RenderX

participants and youth coresearchers. This is an experimental modality for data collection, and participant numbers may be fewer; however, methodological insights regarding engagement will be collated and published to support digital strategy in future work.

Results: We recruited 4 youth coresearchers to codesign the BrightSpark online platform—Canada's first virtual living lab for youth and families with T2D—establishing an embedded registry of youth with T2D and creating the educational content of the 12 modules for the research. Recruitment and data collection began in March 2024 and concluded in September 2025. We anticipate study completion by January 2026.

Conclusions: Youth-onset T2D presents a significant challenge to families and health systems, with less than 30% of youth with T2D achieving treatment targets. Self-management in youth with T2D is further complicated by psychosocial morbidity, identity formation, stigma, blame, shame, historical oppression, and structural barriers to health. This study will contribute a sustainable and novel structure to understand this issue, providing opportunities to generate critically needed understandings of youths' mental health experiences to advance family-centered research and care.

International Registered Report Identifier (IRRID): DERR1-10.2196/83865

(JMIR Res Protoc 2026;15:e83865) doi:10.2196/83865

KEYWORDS

type 2 diabetes; mental health; youth; patient-centered research; living labs; arts-based knowledge translation

Introduction

Background

Type 2 diabetes (T2D) is a complex chronic disease that affects an increasing number of youth and impacts the physical, emotional, and mental well-being of youth and families. Worldwide, the incidence of T2D in youth has risen sharply over the past two decades, particularly among marginalized and low-income populations, reflecting broader trends in obesity and health inequities [1-3]. In 2016, the American Diabetes Association highlighted an urgent need for patient-centered research and care in T2D, which presents as a more aggressive disease in youth [4,5]. Compared to adult-onset T2D, youth-onset T2D is characterized by faster beta cell decline; earlier insulin dependence; and higher risk of complications, such as nephropathy and retinopathy, within a few years of diagnosis [3,6].

The need for youth-centered research is amplified by knowledge that 30% of youth with T2D experience comorbid mental health challenges that are poorly understood and rarely prioritized in clinical practice guidelines, despite their negative impact on self-management [6-9]. These mental health challenges commonly include depression, anxiety, diabetes distress, and disordered eating behaviors, which can further complicate disease management and quality of life [10]. Despite increasing recognition of these challenges, few studies have explored how youth themselves conceptualize the relationship between diabetes and mental health [11]. Existing interventions are largely adult oriented or clinically driven rather than youth informed, limiting their effectiveness and relevance [6,12].

The importance of this gap has been reinforced by a participant advisory group (PAG) associated with our collaborating research team [5], which confirmed that mental health is a priority for youth with T2D and called for greater prioritization in research [13]. Yet, no qualitative or arts-based research exists in this area. This suggests an ongoing need for robust youth-centered research in T2D [6,7]. This proposal responds to this need through the cocreation of a living lab (ie, collaborative user-centered research environments that engage patients, clinicians, and researchers in codeveloping, testing, and refining real-world solutions to health care challenges) [14,15] for mental health in youth-onset T2D. We propose a study incorporating codesign, mixed methods [16-19], and arts-based knowledge translation (KT) [20-26] to open new avenues for patient-centered research in youth-onset T2D and mental health not previously considered.

Youth-Onset T2D and Mental Health

The high incidence of youth-onset T2D in Manitoba (ie, >20/100,000/year) may serve as an early indicator of trends in other jurisdictions; the incidence of youth-onset T2D is highest in Manitoba but is growing across Canada [27]. In Manitoba, fewer than 30% of youth with T2D have achieved treatment targets of hemoglobin A1c (HbA1c)<7% despite the majority receiving specialized tertiary care—a comparable finding Canada-wide [27,28]. A critical factor influencing this finding is that we lack evidence on youth perspectives and experiences influencing their self-management and treatment outcomes. What is known is that self-management in youth with T2D is complicated by psychosocial morbidity, identity formation, stigma, blame, shame, historical oppression, and structural barriers to health facing minority groups disproportionately affected by T2D [29,30]. These factors increase the likelihood of short- and long-term complications [5,9,31,32], highlighting a need to better understand youth experiences, particularly in relation to mental health.

Mental health comorbidities, including psychiatric disorders and diabetes-specific psychological issues, affect 30% of youth with diabetes [33] but are often not prioritized by clinical care teams [34]. Compared to youth with type 1 diabetes (T1D), youth with T2D face heightened risk of early-onset depression [31,35], and the prevalence of substantial depressive symptoms in youth with T2D is 15%-22%—twice as high as youth with T1D [32]. These comorbidities are associated with higher risk of adverse microvascular and macrovascular complications [36]. Anxiety, fear [30,37], distress, hopelessness, and poor quality of life are prevalent [7,27] and impede self-management targets designed to reduce these comorbidities. Disconnects between youth and care provider priorities may exacerbate low rates of

XSL•FO
RenderX

self-management and poor attendance in clinical care visits [35]. However, awareness of the complex relationship between mental health and T2D is lacking [38] and contributes to "diagnostic overshadowing" (ie, stigma in action) [39].

Clinical research into the mental health of youth with T2D is grounded in quantitative cross-sectional studies [40]; existing qualitative research into youth mental health in T2D is limited. Notably, Wicklow et al [41] investigated the determinants of renal complications in youth with T2D by developing a PAG comprising youth and caregivers. They identified themes such as blame and shame, stress and mental health, self-care and systemic issue (food security, support) that enriched their study framework and data collection tools and provided significant insights into the design and processes of the study [41]. Given that patient-centered resource and service design hinges on understanding patients' perspectives [40] and that patient-driven strategies can empower patients and improve adherence and treatment outcomes [40,42], building upon these initial qualitative findings is critical to moving toward patient-centered models of research and care that connect with youth [39,43].

Living Labs in Child Health

Living labs are "virtual, collaborative spaces where users create, prototype, and test concepts, innovations, products and systems" [44]. Living labs emerged in the early 2000s as promising collaborative platforms [45,46] and are widely adopted in civil engineering and aging research [14,47,48]. They are often used to support aging-in-place initiatives but have not been widely used in pediatric environments or as knowledge exchange platforms [15]. For instance, a systematic review of living labs for child health [15] provided evidence for living lab approaches for inclusive environments and to reduce health inequalities among vulnerable residents and regions [15]. However, initiatives drawing from living lab principles are poorly reflected in the extant literature.

Although various understandings of living labs have emerged in the literature, the core attributes of living labs have been identified in previous reviews. Technological infrastructure, stakeholder infrastructure, and community and end-user involvement have been identified as key characteristics of living labs [15]. Multimethod approaches, multiple users, user engagement, a real-life setting, and cocreation of an innovation environment are cited as paramount characteristics within the *Living Lab Methodology Handbook* [46]. Various user roles have also been identified, such as developer, contributor, and tester of an innovation; however, previous reviews of living labs demonstrate that although various roles for users can exist, there is an overreliance on the user as a tester of an innovation. Opportunities for users to be involved as developers or contributors to an innovation, or contribute to more bidirectional knowledge exchange, have been less acknowledged [15]. Considering the diversity, structural disadvantage, and regional variation in T2D across Canada, living labs are an ideal platform for patient engagement and integrated knowledge translation (iKT) in this area. They can support subgroup analysis and regionally specific codesign resources that respond to such variation.

Previous work piloted living labs in pediatrics and has shown their potential as a user-centered model for identifying priorities, collecting data, and codeveloping and evaluating resources and interventions [15] to ultimately support health care delivery and behavior modification [49-51]. Despite their relatively limited scope of application in relation to purpose and user roles, in our previous work, we began expanding this scope and using living lab principles as an innovative infrastructure to support patient registries for longitudinal qualitative and mixed methods research [14], enabling cultivation of rich insights into youth and family experiences in a manner not possible through traditional methods. Within the context of youth-onset T2D, registries and cohorts have been established to understand renal complications (eg, Improving Renal Complications in Adolescents with Type 2 Diabetes through Research [iCARE]) [5,9,27,52], and treatment options (eg, the Pediatric Diabetes Consortium) [53]. The current lab will serve as a complement through its focus on youth experiences and priorities.

Study Purpose and Objectives

This study will help address the need for youth-centered research in youth-onset T2D and mental health. It will investigate youth perspectives and experiences influencing their self-management and treatment outcomes, as well as the complex relationship between mental health and T2D. Through longitudinal, qualitatively dominant mixed methods research, the study will generate understandings of youth experiences across a range of domains, helpful for driving new youth-centered inquiry and models of care. The study will also investigate the potential of living labs as an innovative platform for patient engagement and iKT. As such, the study will engage the following objectives to address these gaps:

- Objective 1: Create a virtual living lab platform with embedded registry of youth with T2D.
- Objective 2: Use this platform to generate critically needed understandings of youths' mental health experiences.
- Objective 3: Identify youth priorities for research and care regarding mental health and T2D and communicate these priorities to providers and researchers in Canada using novel arts-based KT methods.

Methods

Study Design

This protocol was prepared in accordance with the GRAMMS (Good Reporting of Mixed Methods Studies) framework to ensure complete and accurate reporting [54] (see Multimedia Appendix 1 for the GRAMMS checklist). The research involves a three-stage longitudinal, qualitatively dominant, convergent mixed methods design that combines qualitative and quantitative data to explore the experiences of mental health and resiliency among youth with T2D using a virtual platform called BrightSpark as a living lab. Qualitatively dominant mixed methods research prioritizes the qualitative study strand, centering in-depth explorations that are supplemented by quantitative data. Stage 1 involved adapting a virtual platform for use as a living lab and recruiting youth with T2D and their parents or guardians to the participant registry through multi-phase purposive recruitment. Stage 2 involved recruiting

XSL•FO
RenderX

up to 50 youth with T2D from the registry to participate in mixed methods data collection on their experiences of mental health and resiliency. Participants completed baseline and monthly content modules on topics identified as important to youth, including self-image and weight, substance use, wellness, coping, and resiliency. Stage 3 will involve using a convergent mixed methods design to analyze and integrate the qualitative and quantitative data collected in stage 2 to gain a comprehensive understanding of youth with T2D. Overall, this work aims to gain an in-depth understanding of the experiences of youth with T2D and provide support for their mental health and resiliency through exploratory, qualitatively dominant mixed methods research.

Conceptual Framework

The Canadian Institute of Health Research (CIHR) Knowledge-to-Action Ethics framework [48] underpins this research and acknowledges often overlooked ethical considerations in the iterative processes of knowledge creation and translation. This is critical to this research, given the ethical nuances of research with youth, the situational oppression and stigma experienced by youth with T2D, and the underrepresentation of youth voices in T2D research. This framework provides a contextually grounded, inclusive, and respectful foundation for engaging and working with stakeholders. Its principles align closely with this research's qualitatively dominant mixed methods design: for objective 1, it will guide the ethical development of a virtual living lab platform; for objective 2, it will ensure respectful and inclusive data collection on youth mental health experiences in the context of T2D; and for objective 3, it will inform the translation of youth priorities into research, practice, and policy through arts-based KT methods. Given the participatory, iterative, and translational nature of this research spanning digital innovation, patient engagement, and arts-based KT, the CIHR Knowledge-to-Action Ethics framework will ensure that ethical reflection, relational accountability, and action-oriented outcomes are integrated throughout each stage of the research.

Study Setting and Inclusion Criteria

This study will be conducted among youth (10-35 years old) in Canada who received their T2D diagnosis before age 18 years. Data collection will be conducted mainly virtually on BrightSpark, an online living lab, which will allow the research team to reach potential participants across the country.

The inclusion criteria are as follows:

- Youth and young adults (10-35 years old) who received a T2D diagnosis as youths (on or before age 18 years)
- Those who speak English
- Those who live in Canada

Recruitment Strategy

The study is designed based on three research objectives, as outlined in Figure 1.

Figure 1. Research objectives outlined in three stages. T2D: type 2 diabetes.

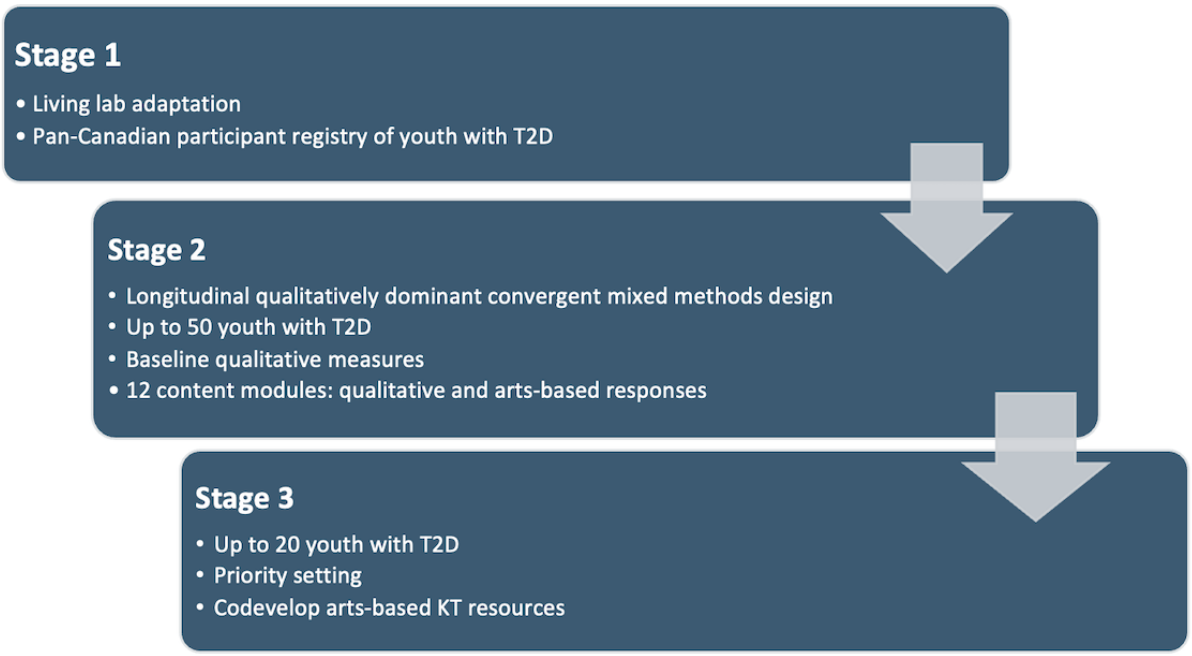

XSL•FO
RenderX

Objective 1: Create Canada's First Living Lab Platform for Youth With T2D

Rationale and Foundational Work

Although patient registries are being developed across Canada by SPOR (ie, Strategy for Patient Oriented Research) networks to facilitate research engagement, patient-oriented research in youth-onset T2D is lacking.*:* Over 12 months, we engaged a PAG of youth, parents, and knowledge users (KUs) to establish platform utility. The PAG indicated that this platform is needed to connect, share experiences of mental health, and build T2D awareness. Youth desire an engaging social media style platform with gaming features, such as avatars and points, to incentivize engagement [55,56]. A systematic review [56] identified seven gamification strategies (goal setting, challenges, feedback, reinforcement, playfulness, progress comparison, social connectivity) that create persuasive system architecture to impact user behavior and experience, providing a strong rationale for their use in the lab [56,57]. A foundational systematic review [15] of living labs in child health was conducted, and the first codesigned clinical living lab for collaborative research and iKT in pediatric rehabilitation [14] enabled us to identify multimethod, user-centered strategies to further optimize platform design.

Codesigning an Engaging Platform

The technology company Recollective helped create the living lab by adapting an existing platform for cultivating online communities. Four youth coresearchers helped codesign the platform. The platform houses a registry where youth with T2D and their guardians (for consent) register in a panel before receiving research study invitations. After consent, registrants become activated participants who at study completion will return to the registry until contacted for a future study. Archive registration preserves all registry and study information between studies, protected by tiers of platform security and data encryption. The BrightSpark URL functions on any smart device (eg, tablet, phone) and presents as a social media app (eg, Facebook) to maximize usability. An evidence-based, youth-directed gamification strategy integrating the seven persuasive strategies [56,57] will incentivize use through points and other customized feedback approaches. Incentives will be redeemed within (eg, activate platform features) and beyond (eg, reimbursement) the platform to optimize engagement. System parameters will prevent incentive overuse for each time period.

Participants and Recruitment

English-speaking youth (10-35 years old) with T2D (diagnosed at age 18 years or younger) in Canada and their guardians (for consent) were invited to join the lab. A broad age range was selected to promote feasibility and enable meaningful subgroup comparisons. Our PAG and the literature reporting sociodemographic characteristics suggest little exclusion bias with English language restriction [5,27].

A research coordinator and graduate research assistant facilitated multiphase purposive recruitment through the Diabetes Education Resource for Children and Adolescents (DER-CA) in Winnipeg, Manitoba, and the national iCARE cohort. DER-CA staff shared study information via onsite posters and clinic visits. The research personnel provided study information to previously consenting iCARE participants using emails and phone numbers provided during iCARE recruitment. Study information was also shared through Facebook and Twitter advertising; Northern communities Buy/Sell; and relevant clinic sites in Canada. Snowball sampling, where participants, KUs, and coresearchers contact potential participants, is being used. Although potentially a source of bias, snowball sampling is appropriate here since T2D often occurs in high rates in communities, clustered in cohorts; the method is sensitive to power relations and social networks [58]. Prospective participants received the BrightSpark URL and created registry profiles with avatars or nonidentifying images. Registry data includes age, sex, gender, ethnicity, geographic location, and years since diabetes diagnosis.

Data Collection in the Living Lab

The platform has embedded data collection and codesign strategies for each study stage, including image-based response; real-time video and time stamping; audio recording; questionnaires; short, long, and diary-answer response formats; and forums.

Objective 2: Investigate Youth Experiences of Mental Health and Resiliency Through Longitudinal, Qualitatively Dominant Mixed Methods Research Within the Living Lab

Rationale

Longitudinal research using interpretive description and arts-based methods is important for generating robust understandings of dynamic youth experiences not captured through short-term studies; for informing iKT, given its emphasis on clinical application of data; and because arts-based methods are developmentally appropriate and efficacious and have been used successfully in a local sample of youth with T2D [22,59].

Participants, Sampling, and Recruitment

Up to 50 youth with T2D were purposively recruited from the BrightSpark registry on the basis of sex, gender, age, years since diagnosis, ethnicity, and geographic location since (1) there is a higher T2D incidence in young females than in males [39]; (2) age and gender impact mental health comorbidities and experiences [60]; (3) years since diagnosis impact the sense of mastery, which is associated with positive mental health [39]; and (4) there are disproportionate rates of T2D and compounding factors impacting mental wellness in minority and remote groups [27,29,39]. Purposive recruitment enables the study of intersectionality since socioeconomic and demographic categories are interacting, not distinct [61]. With the qualitative emphasis, no sample size calculation was performed; efforts are guided by data saturation [62-64] to enable credible subgroup comparisons.

Data Collection

In the first 2 months, baseline data, including sociodemographic and health data (see Table 1), were collected. These measures were selected based on their prevalence and PAG prioritization,

XSL•FO
RenderX

validity, and reliability. The 17-question Diabetes Distress Scale assesses T2D's emotional burden, self-management distress, and social and patient-provider relationship stress and has high reliability (α=.92) and discriminate validity [65]. The Centre for Epidemiological Studies Depression Scale is a 20-item, 4-point scale covering major dimensions of depression, with demonstrated validity from age 6 years through adulthood [66]. The Beck Anxiety Inventory is a 21-item, 4-point scale with strong reliability (α=.92) [67]. Over 12 months, participants also completed monthly content modules (ie, data collection modules; Table 2) including one narrative (video/audio/written) and one arts-based (eg, photographs, drawings) response to exploratory research questions on topics identified from research and PAG/coresearcher engagement. These qualitative data were collected as asynchronous qualitative freeform responses within the BrightSpark platform in response to the domain topic and introductory question. Integration will occur at the moderate level [16,68]: Quantitative data will contextualize qualitative and arts-based findings, inform subgroups for thematic analysis; the extent of missing data will inform the feasibility of future studies; and qualitative analysis will generate hypotheses for future studies within the lead investigator's mixed methods research program.

Table 1. Descriptive characteristics and experiential domains.

Outcome	Data type	Measure/tool	Description
Participant sociodemographic characteristics	Quantitative	BrightSpark registry	Sex, gender, age, years since diagnosis, ethnicity, geographic location, family composition (eg, number of siblings, parents/guardian relationship, parents/guardian participant lives with)
Anthropometric measures	Quantitative	Height and weight (self-reported or measured)	Assesses body composition or general physical status
Glycemic management	Quantitative	Most recent A1c test (self-reported or medical record)	Last-known A1c value (%) to assess blood glucose control over the past 2-3 months
Treatment	Quantitative	Self-reported	Lifestyle interventions (eg, diet, exercise/activities) and medications
Mental health comorbidity	Quantitative	Diabetes Distress Scale [65], Centre for Epidemiological Studies Depression Scale [66], Beck Anxiety Inventory [67]	Assesses diabetes-related distress levels, risk of depression, and anxiety levels
Experiences and challenges	Qualitative	Questionnaire, diary-answer responses, real-time videos, audio recordings, image-based responses	Inductive thematic analysis identifying participant experiences
Coping strategies/facilitators	Qualitative	Questionnaire, diary-answer responses, real-time videos, audio recordings, image-based responses	Themes identified through interpretive description [69] describing strategies/supports used by participants
Integrated insights	Mixed	Joint displays and narrative integration	Synthesizes and juxtaposes quantitative and qualitative findings for a holistic interpretation

XSL•FO
RenderX

Table 2. Monthly content modules for qualitative data collection.

Module	Content
1	Receiving a T2D[a] diagnosis
2	Managing T2D, anxiety, and stress
3	T2D medication by needle or mouth and low and high blood sugar
4	Relationships with family
5	Relationships with friends (social relationships)
6	Relationships with health care professionals
7	School, sports, and work
8	Self-image and weight
9	Substance use
10	Wellness, coping, and resilience
11	Food security and choices
12	Balancing everyday nourishment with special occasions

[a]T2D: type 2 diabetes.

Procedures

A visual timeline prompted activated participants to complete baseline measures, which unlocked the qualitative content modules and a list of free mental health support resources. Participants could then select a minimum of one and a maximum of three modules monthly for 12 months in any order, with the option to suggest new topics to identify additional priorities. Participants provided qualitative responses (written or audio) in response to the module-specific research questions; provided an accompanying artistic response, if desired (eg, photograph, drawing); and received additional evidence-based information following their response provision within the platform. Participants received CA $20 (approximately US $14) honorariums for each module completed.

Data Analysis

A trained graduate research assistant and research coordinator will use a convergent mixed methods data analysis approach. Quantitative data will be descriptively analyzed using IBM SPSS to summarize participant sociodemographic characteristics and outcome data, as described in Table 1. Qualitative data will be inductively analyzed using interpretive description [69] in MAXQDA mixed methods software to identify themes reflecting participant experiences and insights. System-generated transcripts will be read repeatedly to gain a sense of the whole, prior to coding line by line. Open coding will ensue, commencing identification of meaning units at the level of word clusters or sentences, and associated definitions for codes will be created in a codebook. Codes from the first three transcripts will inform a coding framework to apply to the remaining transcripts. New data from the remaining transcripts will refine the coding framework [69]. Codes will be examined for frequency, with focus on within- and across-case coverage, and tentatively grouped into themes. Themes will be substantiated with codes and quotes and reviewed before labeling to enable flexibility with conceptualization and to ensure subgroup comparisons (eg, gender comparisons) are reflected. Cutoff categories for baseline measures will be examined to inform subgroups (eg, low, moderate, high distress); themes will be clustered and thematic labels modified, as needed [70]. Arts-based data will be analyzed by key content, constituent elements, and tone using a framework successfully applied in our previous work [71] and handled holistically with accompanying narrative data. Participants will validate their interpretation. The module selection order will be tallied and descriptively analyzed to suggest content prioritization. Quantitative and qualitative findings will be integrated in joint displays, enabling meta-inferences to be drawn [72-74]. Analysis will generate hypotheses, while producing rich and expressive representations of mental health and T2D to inform stage 3. Recognized trustworthiness [75] and mixed methods legitimation [76] criteria will assure quality. Strategies will include member checking (credibility), analytic memos (dependability), substantive linking of interpretations with data (confirmability, meta-inferential), and considering alternative explanations (authenticity).

Objective 3: Identify and Communicate Youth Priorities Using Arts-Based KT

Rationale

A 2016 consensus document [7] for youth-onset T2D emphasizes the need to prioritize outcomes other than glycemic control. Our PAG identified mental health as a key research domain, yet youth priorities for mental health research and care in T2D remain unknown. As such, we aim to identify research and care priorities of youth with T2D regarding mental health and resiliency. Our research has shown the power of iKT in identifying priorities and communicating findings through arts-based KT (eg, communicating research using storylines based on family experiences) [20,24,26,77]. Such KT resources in youth-onset T2D are needed but lacking.

Participants and Sampling

Up to a third of the participants from stage 2 will be purposively sampled from the registry by age, gender, geography, mental health status, and engagement (eg, >90% module completion

XSL•FO
RenderX

in stage 2) to provide sufficient diversity to generate credible findings regarding priorities and to inform KT resource development. This aligns with the Delphi technique's emphasis on participant qualities over representative samples [78].

Procedures

We will use card sorting [79] with a modified Delphi technique [78]—common and robust priority-setting method—to identify youth priorities from stage 2. Youth will be recruited to a forum within BrightSpark and provided with a series of cards, each representing a theme from stage 2 (first Delphi round). Participants will (1) rank-order the cards by priority (second Delphi round), (2) provide written/verbal feedback and identify priorities not represented (member checking), and (3) amend their ranking following feedback (third Delphi round) [78]. Rank orders will be tabulated to determine up to 10 priority themes. Participants will be led through the following seven steps to codesign up to four arts-based KT resources and two infographics:

- Step 1: Gain KT familiarity, including KT goals relevant to the research (eg, raise awareness, knowledge) [80,81].
- Step 2: Identify target audiences based on the KT principle of tailoring to stakeholder groups (eg, health provider) and evidence of effectiveness of strategies for each group [82].
- Step 3: Identify key messages based on the KT principle of aligning with target audiences and strategies [81-83].
- Step 4: Determine delivery modality (eg, video, comic) in reference to the Archibald arts-based KT classification framework [77] and evidence of effectiveness of arts-based KT strategies [23,84].
- Step 5: Provide storyboard content for iterative revision [24].
- Step 6: Provide feedback at key development stages.
- Step 7: Identify dissemination channels based on preferences and evidence of online analytics [85,86].

At this stage, participants will be provided CA $100 (approximately US $72) honorariums for participation in this stage.

Data Collection and Analysis

We will use three-stage card sorting in the modified Delphi technique for prioritization, video/audio time stamping with narrative feedback for KT development, forum question-and-feedback sessions anchored to each step of the development process, and an inverted rank-order analysis used in our previous work to consolidate priorities [87]. A research assistant will conduct directed qualitative content analysis [88] of forum data using MAXQDA, with purposive subgroups as analytic units focusing on manifest content. A formative coding matrix will be applied to each forum thread, corresponding to each step of KT development. Interim analysis will direct forum questioning to inform the KT resources.

Knowledge Translation Plan

We engaged stakeholders from the onset of study conceptualization [40,42], leveraging their strengths and perspectives, while responding to families' prioritization of mental-health and a critical knowledge gap. We will innovate on iKT by using evidence-based KT principles to codevelop arts-based KT resources (stage 3) to mobilize findings. Dissemination through the lead author's website and YouTube channel will be informed by YouTube and Google Analytics [81]. Resources will be submitted to the CIHR Institute of Human Development, Child and Youth Health (IHDCYH) Talks competition (video form) and resource centers. BrightSpark aligns with the patient-oriented research movement in Canada, providing a rare opportunity to consolidate and communicate youth priorities in T2D to springboard future iKT and youth-centered research of collaborators. It will provide the infrastructure to scale up other investigations into youth-onset T2D and methodological guidance for other living lab initiatives nationally and internationally, with applicability to other chronic disease contexts.

Ethical Considerations

This study was reviewed and approved by the University of Manitoba Health Research Ethics Board (approval number HE2022-0396) and the Shared Health Approval Committee for Privacy, Impact and Access in Research (approval number SH2023:042). Written consent of participation, to record in stage 3 and to use and disseminate data, was collected from all participants on the virtual living lab platform prior to participation in stages 2 and 3. All identifying information will be removed from the data collected through narratives and art-based modules in stage 2 of the project. Similarly, recordings of the group discussion in stage 3 will be destroyed after transcription, and transcriptions will be kept for 5 years in a secure locked file on a secure and password-protected device only accessible to the principal investigator (author MA) and identified personnel of the research team. The module content and presentation were thoughtfully designed to reflect a supportive orientation and were informed with participant advisory input; context-relevant mental health resource lists were made available to participants at multiple time points within the platform, associated with each module.

Results

A name and logo for the living lab was crowdsourced on Hatchwise, and the research team voted to select the name BrightSpark, Care Lab and the logo. Four youth coresearchers and the research team codesigned BrightSpark for youth and families with T2D and created the educational contents of the 12 modules for the research. Data collection and recruitment began in March 2024 and concluded on September 23, 2025, with analysis ongoing through January 2026. Findings from this study will be disseminated through peer-reviewed publications, presentations at national and international conferences, arts-based KT, and digital and social media platforms targeted at youth and health care providers.

Discussion

Summary

This study anticipates that youth with T2D will report a high prevalence of mental health challenges, including stress and anxiety, which are likely to influence their self-management behaviors and treatment outcomes [29,30]. The ordered

XSL•FO
RenderX

identification of pertinent response modules, as indicated by participant-led selection and completion, will provide insight into possible domains of perceived relevance for youth, guiding future exploratory and intervention research. It is also expected that the study will provide insights into the feasibility and acceptability of a virtual youth-centered platform for research participation, highlighting potential strategies for engaging this population in patient-centered care and research.

The existing literature underscores that mental health challenges in youth with T2D are underrecognized and poorly addressed, yet they have substantial implications for adherence, glycemic control, long-term complications, and overall wellness [7-9,29,30,89]. Although quantitative research is generating evidence around the prevalence of such challenges, qualitative investigations are less common. A recent interpretive descriptive study [13] with 22 participants diagnosed with T2D before 18 years of age emphasized the pervasive impacts of mental health on T2D experiences and management, highlighting the intersections between blood sugar stability, mood, mental health, and considerations related to growth and development. Similarly, a focus group study [89] involving eight First Nation adolescents identified common experiences around stigma and shame, and the associated weight of managing a T2D diagnosis. However, the specific domains and facets of such mental health experiences have not been explored. This study will significantly extend the limited qualitative investigations emerging in this area.

The anticipated findings from this study may extend prior work by illuminating the nuanced ways in which mental health intersects with self-management in youth, as well as by demonstrating the potential of technology-mediated approaches for capturing these experiences. Living labs have shown promise in adolescent populations for both research participation and health interventions [90], but evidence on their application for qualitatively dominant research with youth with T2D is limited. This work anticipates generating understandings and perspectives that will inform patient-centered strategies that connect with youth. Concurrently, the work will provide insight into the utility of such a virtual platform for youth-centered research, which has high value in an increasingly user-centered and technologically proficient society.

Strengths and Limitations

A major strength of this study is the use of a qualitatively dominant, convergent mixed methods approach, which allows for the integration of quantitative and qualitative data to provide a more comprehensive understanding of mental health experiences and self-management behaviors in youth with T2D. This approach can reveal patterns not apparent through a single method and support comprehensive understandings, enhancing the validity of the findings. Additional strengths include the use of a virtual platform, which may increase accessibility for diverse participants, and the direct engagement of youth and the PAG ensures that the findings accurately reflect and are relevant to their lived experiences.

Limitations include a potential for selection bias, as youth with higher digital literacy or motivation may be more likely to participate. However, given that the virtual living lab was codesigned with youth and the PAG representing diverse experiences with technology, socioeconomic backgrounds, and health literacy levels, selection bias was likely reduced. Through iterative codesign sessions, the PAG informed the platform's accessibility features, plain-language guidance, navigation, and tone ensuring inclusivity and ease of use. The embedded codesign principles and continuous feedback loops lowered potential digital participation barriers and promoted equitable engagement among youth and parents who might otherwise be underrepresented due to varying comfort levels with virtual platforms or differing levels of intrinsic motivation.

Codesigning the virtual living lab environment with youth and the PAG will also serve to mitigate potential biases related to self-reported data and interpretation of mixed methods data integration. The involvement of advisors in shaping the data collection tools, question framing, and engagement formats will foster psychological safety and participant ownership, which are conditions that reduce social desirability bias and encourage authentic sharing of experiences [91]. The inclusion of multimodal reflection options (eg, written, audio, or video responses) will allow participants to provide input in real time or asynchronously, lessening reliance on retrospective recall and associated recall bias. Furthermore, our engagement with youth and the PAG in defining the key constructs and indicators will ensure conceptual alignment between qualitative and quantitative components, thereby strengthening the methodological coherence.

Future Directions

The insights gained from this study are expected to inform the development of targeted, patient-centered interventions for youth with T2D that address both mental health and self-management needs. Future research could explore additional longitudinal outcomes, evaluate intervention effectiveness, and examine strategies to sustain engagement in digital platforms. Additionally, findings may inform clinical guidelines and support the integration of mental health assessment and support into routine T2D care for youth and adolescents. Engaging participants in the dissemination process, such as through cocreated summaries, arts-based KT resources, or infographics, may further enhance the relevance and impact of the findings.

Conclusion

By exploring the mental health experiences of youth with T2D and the feasibility of a virtual living lab for research engagement, this study aims to generate actionable insights to inform youth-centered care and research strategies. Through this innovative living lab, this study will contribute a sustainable structure to understanding and supporting youths' mental health needs, advancing both research and care practices.

XSL•FO
RenderX

Acknowledgments

MA acknowledges the generous funding received from the SickKids Foundation–Canadian Institute of Health Research–Institute of Human Development, Child and Youth Health New Investigator Research Grants in Child and Youth Health application “A Living Lab for Mental Health in Youth-Onset T2D (UM project number 55215). The funder’s financial support for this study does not indicate their involvement in how the research is conducted.

Data Availability

The datasets generated during this study are not available.

Authors' Contributions

All authors contributed to the conceptualization of the study and writing and editing of the manuscript. All authors also participated in our quarterly meeting and are committed to conducting the proposed research. In addition, all authors have approved the manuscript prior to submission and agree to be held responsible for all aspects of the work described here. The authors attest to not using generative artificial intelligence.

Conflicts of Interest

None declared.

Multimedia Appendix 1
GRAMMS (Good Reporting of Mixed Methods Studies) checklist.
[PDF File (Adobe PDF File), 143 KB - resprot_v15i1e83865_app1.pdf]

Multimedia Appendix 2
Peer review report from SickKids Canada New Investigators Research Grants Committee.
[PDF File (Adobe PDF File), 199 KB - resprot_v15i1e83865_app2.pdf]

References

1. Bombaci B, Passanisi S, Torre A, Sbilordo S, Inì E, Papa M, et al. Youth-onset type 2 diabetes: update on epidemiology, pathophysiology, diagnosis, and management strategies. Diabetology 2025 Sep 01;6(9):90 [FREE Full text] [doi: 10.3390/diabetology6090090]
2. Titmuss A, Korula S, Wicklow B, Nadeau KJ. Youth-onset type 2 diabetes: an overview of pathophysiology, prognosis, prevention and management. Curr Diab Rep 2024 Aug 03;24(8):183-195 [FREE Full text] [doi: 10.1007/s11892-024-01546-2] [Medline: 38958831]
3. Rodriquez IM, O'Sullivan KL. Youth-onset type 2 diabetes: burden of complications and socioeconomic cost. Curr Diab Rep 2023 May 24;23(5):59-67 [FREE Full text] [doi: 10.1007/s11892-023-01501-7] [Medline: 36961664]
4. Young-Hyman D, de Groot M, Hill-Briggs F, Gonzalez JS, Hood K, Peyrot M. Psychosocial care for people with diabetes: a position statement of the American Diabetes Association. Diabetes Care 2016 Dec 22;39(12):2126-2140 [FREE Full text] [doi: 10.2337/dc16-2053] [Medline: 27879358]
5. Dart AB, Wicklow BA, Sellers EA, Dean HJ, Malik S, Walker J, iCARE investigators. The improving renal complications in adolescents with type 2 diabetes through the research (iCARE) cohort study: rationale and protocol. Can J Diabetes 2014 Oct;38(5):349-355 [FREE Full text] [doi: 10.1016/j.jcjd.2014.07.224] [Medline: 25284698]
6. Nadeau K, Mayer-Davis E, Gubitosi-Klug R, Zeitler P, Kahn S, Dabelea D, SEARCH, TODAY, RISE,DISCOVERY study groups. Youth-onset type 2 diabetes: what we’ve learned from key youth-onset type 2 diabetes studies, what we still don’t know, and why it is important. Diabetes Care 2025 Jul 01;48(7):1136-1149 [FREE Full text] [doi: 10.2337/dc25-0001] [Medline: 40272281]
7. Nadeau K, Anderson B, Berg E, Chiang J, Chou H, Copeland K, et al. Youth-onset type 2 diabetes consensus report: current status, challenges, and priorities. Diabetes Care 2016 Sep;39(9):1635-1642 [FREE Full text] [doi: 10.2337/dc16-1066] [Medline: 27486237]
8. Agarwal S, Raymond JK, Isom S, Lawrence JM, Klingensmith G, Pihoker C, et al. Transfer from paediatric to adult care for young adults with Type 2 diabetes: the SEARCH for Diabetes in Youth Study. Diabet Med 2018 Apr 14;35(4):504-512 [FREE Full text] [doi: 10.1111/dme.13589] [Medline: 29377258]
9. McGavock J, Durksen A, Wicklow B, Malik S, Sellers EA, Blydt-Hansen T, et al. Determinants of readiness for adopting healthy lifestyle behaviors among indigenous adolescents with type 2 diabetes in Manitoba, Canada: a cross-sectional study. Obesity (Silver Spring) 2018 May 13;26(5):910-915 [FREE Full text] [doi: 10.1002/oby.22148] [Medline: 29533530]

XSL•FO
RenderX

10. Jaser SS, Roddy MK, Beam AB. Psychosocial and behavioral health among youth and adults with diabetes. In: Lawrence JM, Casagrande SS, Herman WH, Wexler DJ, Cefalu WT, editors. Diabetes in America. Bethesda, MD: National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK); 2023:561-562.
11. Oxlad M, Smith L, McNamara T, Young A, Borrowdale E, Pena A. Psychosocial experiences of adolescents with type 2 diabetes: a systematic review and meta-synthesis. Can J Diabetes 2025 Oct 03 [FREE Full text] [doi: 10.1016/j.jcjd.2025.09.004] [Medline: 41046938]
12. Lee Jia Jia I, Zampetti S, Pozzilli P, Buzzetti R. Type 2 diabetes in children and adolescents: challenges for treatment and potential solutions. Diabetes Res Clin Pract 2024 Nov;217:111879 [FREE Full text] [doi: 10.1016/j.diabres.2024.111879] [Medline: 39369858]
13. Wicklow B, Dart A, McKee J, Griffiths A, Malik S, Quoquat S, et al. Experiences of First Nations adolescents living with type 2 diabetes: a focus group study. CMAJ 2021 Mar 22;193(12):E403-E409 [FREE Full text] [doi: 10.1503/cmaj.201685] [Medline: 33753364]
14. Archibald MM, Wittmeier K, Gale M, Ricci F, Russell K, Woodgate RL. Living labs for patient engagement and knowledge exchange: an exploratory sequential mixed methods study to develop a living lab in paediatric rehabilitation. BMJ Open 2021 May 04;11(5):e041530 [FREE Full text] [doi: 10.1136/bmjopen-2020-041530] [Medline: 33947723]
15. Archibald M, Wiebe S, Rieger K, Linton J, Woodgate R. Protocol for a systematic review of living labs in healthcare. BMJ Open 2021 Feb 05;11(2):e039246 [FREE Full text] [doi: 10.1136/bmjopen-2020-039246] [Medline: 33550226]
16. Archibald MM. Integrating the arts and mixed methods research: a review and a way forward. Int J Mult Res Approaches 2018 Jul 16;10(1):342-355 [FREE Full text] [doi: 10.29034/ijmra.v10n1a23]
17. Archibald MM. Investigator triangulation. J Mixed Methods Res 2015 Feb 17;10(3):228-250 [FREE Full text] [doi: 10.1177/1558689815570092]
18. Archibald M. Arts-based research as an inherently mixed analysis approach. In: Reviewers Guide for Mixed Methods Research Analysis. UK: Routledge; 2021.
19. Archibald MM, Gerber N. Arts and mixed methods research: an innovative methodological merger. Am Behav Sci 2018 Apr 27;62(7):956-977 [FREE Full text] [doi: 10.1177/0002764218772672]
20. Archibald M, Blines J. Metaphors in the making: illuminating the process of arts-based health research through a case exemplar linking arts-based, qualitative and quantitative research data. Int J Qual Methods 2021 Feb 27;20:160940692098795 [FREE Full text] [doi: 10.1177/1609406920987954]
21. Archibald MM, Caine V, Scott SD. Intersections of the arts and nursing knowledge. Nurs Inq 2017 Apr 30;24(2):e12153 [FREE Full text] [doi: 10.1111/nin.12153] [Medline: 27572849]
22. Archibald M, Scott S, Hartling L. Mapping the waters: a scoping review of the use of visual arts in pediatric populations with health conditions. Arts Health 2013 Jan 29;6(1):5-23 [FREE Full text] [doi: 10.1080/17533015.2012.759980]
23. Scott SD, Brett-MacLean P, Archibald M, Hartling L. Protocol for a systematic review of the use of narrative storytelling and visual-arts-based approaches as knowledge translation tools in healthcare. Syst Rev 2013 Mar 20;2:19 [FREE Full text] [doi: 10.1186/2046-4053-2-19] [Medline: 23514237]
24. Archibald MM, Hartling L, Ali S, Caine V, Scott SD. Developing "My Asthma Diary": a process exemplar of a patient-driven arts-based knowledge translation tool. BMC Pediatr 2018 Jun 05;18(1):186 [FREE Full text] [doi: 10.1186/s12887-018-1155-2] [Medline: 29871611]
25. Archibald M, Ambagtsheer R, Lawless MT, Thompson MO, Shultz T, Chehade MJ, et al. Co-designing evidence-based videos in health care: a case exemplar of developing creative knowledge translation “evidence-experience” resources. Int J Qual Methods 2021 Jul 19;20 [FREE Full text] [doi: 10.1177/16094069211019623]
26. Archibald MM, Scott SD. Learning from usability testing of an arts-based knowledge translation tool for parents of a child with asthma. Nurs Open 2019 Oct 10;6(4):1615-1625 [FREE Full text] [doi: 10.1002/nop2.369] [Medline: 31660190]
27. Dart AB, Wicklow B, Blydt-Hansen TD, Sellers EAC, Malik S, Chateau D, et al. A holistic approach to risk for early kidney injury in indigenous youth with type 2 diabetes: a proof of concept paper from the iCARE cohort. Can J Kidney Health Dis 2019 Apr 21;6:2054358119838836 [FREE Full text] [doi: 10.1177/2054358119838836] [Medline: 31041107]
28. Diabetes Canada Clinical Practice Guidelines Expert Committee, Panagiotopoulos C, Hadjiyannakis S, Henderson M. Type 2 diabetes in children and adolescents. Can J Diabetes 2018 Apr;42 Suppl 1:S247-S254 [FREE Full text] [doi: 10.1016/j.jcjd.2017.10.037] [Medline: 29650104]
29. McGavock J, Wicklow B, Dart AB. Type 2 diabetes in youth is a disease of poverty. Lancet 2017 Oct;390(10105):1829 [FREE Full text] [doi: 10.1016/s0140-6736(17)32461-3]
30. Browne JL, Ventura A, Mosely K, Speight J. 'I call it the blame and shame disease': a qualitative study about perceptions of social stigma surrounding type 2 diabetes. BMJ Open 2013 Nov 18;3(11):e003384 [FREE Full text] [doi: 10.1136/bmjopen-2013-003384] [Medline: 24247325]
31. Tryggestad JB, Willi SM. Complications and comorbidities of T2DM in adolescents: findings from the TODAY clinical trial. J Diabetes Complications 2015 Mar;29(2):307-312 [FREE Full text] [doi: 10.1016/j.jdiacomp.2014.10.009] [Medline: 25468310]
32. Viner R, White B, Christie D. Type 2 diabetes in adolescents: a severe phenotype posing major clinical challenges and public health burden. Lancet 2017 Jun;389(10085):2252-2260 [FREE Full text] [doi: 10.1016/s0140-6736(17)31371-5]

XSL•FO
RenderX

33. Diabetes Canada Clinical Practice Guidelines Expert Committee, Robinson DJ, Coons M, Haensel H, Vallis M, Yale JF. Diabetes and mental health. Can J Diabetes 2018 Apr;42 Suppl 1:S130-S141 [FREE Full text] [doi: 10.1016/j.jcjd.2017.10.031] [Medline: 29650085]
34. Huynh E. The mountain of health: the perceptions and perceived expectations of youth living with type 2 diabetes in winnipeg, Manitoba, Canada, through a grounded theory study. Can J Diabetes 2018 Aug;42(4):344-349 [FREE Full text] [doi: 10.1016/j.jcjd.2017.09.006] [Medline: 29128303]
35. Gladstone TR, Beardslee WR, O'Connor EE. The prevention of adolescent depression. Psychiatr Clin North Am 2011 Mar;34(1):35-52 [FREE Full text] [doi: 10.1016/j.psc.2010.11.015] [Medline: 21333838]
36. Lin EHB, Rutter CM, Katon W, Heckbert SR, Ciechanowski P, Oliver MM, et al. Depression and advanced complications of diabetes: a prospective cohort study. Diabetes Care 2010 Feb;33(2):264-269 [FREE Full text] [doi: 10.2337/dc09-1068] [Medline: 19933989]
37. Lawrence J, Standiford D, Loots B, Klingensmith G, Williams D, Ruggiero A, SEARCH for Diabetes in Youth Study. Prevalence and correlates of depressed mood among youth with diabetes: the SEARCH for Diabetes in Youth Study. Pediatrics 2006 Apr;117(4):1348-1358 [FREE Full text] [doi: 10.1542/peds.2005-1398] [Medline: 16585333]
38. Nash M. Mental health service users' experiences of diabetes care by mental health nurses: an exploratory study. J Psychiatr Ment Health Nurs 2014 Oct 18;21(8):715-723 [FREE Full text] [doi: 10.1111/jpm.12140] [Medline: 24548452]
39. Nash M. Diagnostic overshadowing: a potential barrier to physical health care for mental health service users. Ment Health Pract 2013 Dec;17(4):22-26 [FREE Full text] [doi: 10.7748/mhp2013.12.17.4.22.e862]
40. Mcgavock J, Wicklow B, Sellers E, Kroeker K, Blydt-Hansen T, Dart A. 1246-P: mental health among youth living with type 2 diabetes in Manitoba: a cross-sectional study. Diabetes 2020 Jun 1;69(Supplement_1):1246-P [FREE Full text] [doi: 10.2337/db20-1246-p]
41. Wicklow BA, Malik S, Mckee J, Griffiths A, Dart AB. 28 - Engaging patients in pediatric diabetes research: creation of the iCARE participant advisory group. Can J Diabetes 2019 Oct 1;43(7):S11 [FREE Full text] [doi: 10.1016/j.jcjd.2019.07.035]
42. Greenhalgh T, Jackson C, Shaw S, Janamian T. Achieving research impact through co-creation in community-based health services: literature review and case study. Milbank Q 2016 Jun 06;94(2):392-429 [FREE Full text] [doi: 10.1111/1468-0009.12197] [Medline: 27265562]
43. Owen J. Equal treatment: closing the gap. Ment Health Pract 2005;8(5):44. [doi: 10.7748/mhp.8.5.44.s32]
44. Leminen S, Westerlund M. Categorization of innovation tools in living labs. Technol Innov Manag Rev 2017;7(1):15-25 [FREE Full text] [doi: 10.22215/timreview/1046]
45. Guldemond N, Geenhuizen MV. Critical factors in “livings labs” for new health concepts and medical technology. Academia. URL: https://www.academia.edu/17493286/Critical_Factors_in_Livings_Labs_for_New_Health_Concepts_and_Medical_Technology [accessed 2025-10-29]
46. Ståhlbröst A, Holst M. The Living Lab Methodology Handbook. Luleå: Social Informatics at Luleå University of Technology and CDT; 2012.
47. Kim J, Kim YL, Jang H, Cho M, Lee M, Kim J, et al. Living labs for health: an integrative literature review. Eur J Public Health 2020 Feb 01;30(1):55-63 [FREE Full text] [doi: 10.1093/eurpub/ckz105] [Medline: 31169898]
48. Picard R. Co-design in Living Labs for Healthcare and Independent Living: Concepts, Methods and Tools. 1st ed. Hoboken, London: John Wiley & Sons; 2017.
49. Phanareth K, Vingtoft S, Christensen AS, Nielsen JS, Svenstrup J, Berntsen GKR, et al. The epital care model: a new person-centered model of technology-enabled integrated care for people with long term conditions. JMIR Res Protoc 2017 Jan 16;6(1):e6 [FREE Full text] [doi: 10.2196/resprot.6506] [Medline: 28093379]
50. Swinkels ICS, Huygens MWJ, Schoenmakers TM, Oude Nijeweme-D'Hollosy W, van Velsen L, Vermeulen J, et al. Lessons learned from a living lab on the broad adoption of eHealth in primary health care. J Med Internet Res 2018 Mar 29;20(3):e83 [FREE Full text] [doi: 10.2196/jmir.9110] [Medline: 29599108]
51. Konstantinidis EI, Bamparopoulos G, Bamidis PD. Moving real exergaming engines on the web: the webFitForAll Case Study in an active and healthy ageing living lab environment. IEEE J Biomed Health Inform 2017 May;21(3):859-866 [FREE Full text] [doi: 10.1109/jbhi.2016.2559787]
52. Dart AB, McGavock J, Sharma A, Chateau D, Schwartz GJ, Blydt-Hansen T. Estimating glomerular filtration rate in youth with obesity and type 2 diabetes: the iCARE study equation. Pediatr Nephrol 2019 Sep 2;34(9):1565-1574 [FREE Full text] [doi: 10.1007/s00467-019-04250-6] [Medline: 31049718]
53. Tamborlane WV, Chang P, Kollman C, Klingensmith GJ, Ruedy K, Gal RL, Pediatric Diabetes Consortium. Eligibility for clinical trials is limited for youth with type 2 diabetes: insights from the Pediatric Diabetes Consortium T2D Clinic Registry. Pediatr Diabetes 2018 Dec 10;19(8):1379-1384 [FREE Full text] [doi: 10.1111/pedi.12763] [Medline: 30175440]
54. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98 [FREE Full text] [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
55. Kato PM. Video games in health care: closing the gap. Rev Gen Psychol 2010 Jun 01;14(2):113-121 [FREE Full text] [doi: 10.1037/a0019441]

XSL•FO
RenderX

56. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
57. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
58. Noy C. Sampling knowledge: the hermeneutics of snowball sampling in qualitative research. Int J Soc Res Methodol 2008 Oct;11(4):327-344 [FREE Full text] [doi: 10.1080/13645570701401305]
59. MacDonald C, Friedman S, Marks SD. Creative arts diabetes initiative: group art therapy and peer support for youth and young adults transitioning from pediatric to adult diabetes care in Manitoba, Canada. Can J Diabetes 2019 Aug;43(6):377-383.e3 [FREE Full text] [doi: 10.1016/j.jcjd.2019.04.016] [Medline: 31235397]
60. Canadian health survey on children and youth, 2019. Statistics Canada. 2020. URL: https://www150.statcan.gc.ca/n1/daily-quotidien/200723/dq200723a-eng.htm [accessed 2025-10-29]
61. Wemrell M, Bennet L, Merlo J. Understanding the complexity of socioeconomic disparities in type 2 diabetes risk: a study of 4.3 million people in Sweden. BMJ Open Diabetes Res Care 2019;7(1):e000749 [FREE Full text] [doi: 10.1136/bmjdrc-2019-000749] [Medline: 31798898]
62. Collins KMT, Onwuegbuzie AJ, Jiao QG. A mixed methods investigation of mixed methods sampling designs in social and health science research. J Mix Methods Res 2007 Jul 01;1(3):267-294 [FREE Full text] [doi: 10.1177/1558689807299526]
63. Sandelowski M. Sample size in qualitative research. Res Nurs Health 1995 Apr;18(2):179-183 [FREE Full text] [doi: 10.1002/nur.4770180211] [Medline: 7899572]
64. Thorne S. Interpretive Description. New York: Routledge; 2016.
65. Chin YW, Lai PSM, Chia YC. The validity and reliability of the English version of the diabetes distress scale for type 2 diabetes patients in Malaysia. BMC Fam Pract 2017 Feb 20;18(1):25 [FREE Full text] [doi: 10.1186/s12875-017-0601-9] [Medline: 28219325]
66. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977 Jun 01;1(3):385-401 [FREE Full text] [doi: 10.1177/014662167700100306]
67. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56(6):893-897 [FREE Full text] [doi: 10.1037/0022-006X.56.6.893]
68. Greene JC, Caracelli VJ, Graham WF. Toward a conceptual framework for mixed-method evaluation designs. Educ Eval Policy Anal 1989 Sep 01;11(3):255-274 [FREE Full text] [doi: 10.3102/01623737011003255]
69. Charmaz K. Constructing Grounded Theory. London, UK; Thousand Oaks, CA: Sage Publications; 2014.
70. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
71. Archibald MM, Ambagtsheer R, Beilby J, Chehade MJ, Gill TK, Visvanathan R, et al. Perspectives of frailty and frailty screening: protocol for a collaborative knowledge translation approach and qualitative study of stakeholder understandings and experiences. BMC Geriatr 2017 Apr 17;17(1):87 [FREE Full text] [doi: 10.1186/s12877-017-0483-7] [Medline: 28415977]
72. Hitchcock JH, Onwuegbuzie AJ. Developing mixed methods crossover analysis approaches. J Mix Methods Res 2019 Apr 17;14(1):63-83 [FREE Full text] [doi: 10.1177/1558689819841782]
73. Tashakkori A, Teddlie C. Sage Handbook of Mixed Methods in Social & Behavioral Research. 2nd Edition. Thousand Oaks, CA: Sage Publications; 2010.
74. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015 Nov 09;13(6):554-561 [FREE Full text] [doi: 10.1370/afm.1865] [Medline: 26553895]
75. Keller GC. Naturalistic inquiry: by Yvonna S. Lincoln and Egon G. Guba. Beverly Hills, CA: Sage Publications, 1985, 391 pp. Eval Program Plann 1987;10(2):184-185. [doi: 10.1016/0149-7189(87)90056-5]
76. Onwuegbuzie AJ, Johnson RB, Collins KMT. Assessing legitimation in mixed research: a new framework. Qual Quant 2009 Dec 19;45(6):1253-1271 [FREE Full text] [doi: 10.1007/s11135-009-9289-9]
77. Archibald MM, Caine V, Scott SD. The development of a classification schema for arts-based approaches to knowledge translation. Worldviews Evid Based Nurs 2014 Oct 18;11(5):316-324 [FREE Full text] [doi: 10.1111/wvn.12053] [Medline: 25132050]
78. Powell C. The Delphi technique: myths and realities. J Adv Nurs 2003 Feb;41(4):376-382 [FREE Full text] [doi: 10.1046/j.1365-2648.2003.02537.x] [Medline: 12581103]
79. Li C, Fairhurst S, Chege C, Jenks EH, Tsong Y, Golden D, et al. Card-sorting as a tool for communicating the relative importance of supervisor interventions. Clin Superv 2016 Apr 22;35(1):80-97 [FREE Full text] [doi: 10.1080/07325223.2016.1165641]
80. Albrecht L, Archibald M, Snelgrove-Clarke E, Scott SD. Systematic review of knowledge translation strategies to promote research uptake in child health settings. J Pediatr Nurs 2016 May;31(3):235-254 [FREE Full text] [doi: 10.1016/j.pedn.2015.12.002] [Medline: 26786910]

XSL•FO
RenderX

81. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci 2012 May 31;7:50 [FREE Full text] [doi: 10.1186/1748-5908-7-50] [Medline: 22651257]
82. Knowledge translation strategies for different target audiences. Alberta Health Services. 2022. URL: https://www.albertahealthservices.ca/assets/info/amh/if-amh-kt-strategies-for-different-audiences.pdf [accessed 2025-11-25]
83. Knowledge translation planner. Health Canada. 2017. URL: https://www.canada.ca/en/health-canada/corporate/about-health-canada/reports-publications/grants-contributions/knowledge-transfer-planner.html [accessed 2025-11-25]
84. Hartling L, Scott SD, Johnson DW, Bishop T, Klassen TP. A randomized controlled trial of storytelling as a communication tool. PLoS One 2013 Oct 25;8(10):e77800 [FREE Full text] [doi: 10.1371/journal.pone.0077800] [Medline: 24204974]
85. Chambers CT, Dol J, Parker JA, Caes L, Birnie KA, Taddio A, et al. Implementation effectiveness of a parent-directed youtube video (“it doesn’t have to hurt”) on evidence-based strategies to manage needle pain: descriptive survey study. JMIR Pediatr Parent 2020 Mar 04;3(1):e13552 [FREE Full text] [doi: 10.2196/13552] [Medline: 32130190]
86. How to use YouTube Analytics and track the right metrics. Think Media & YouTube. URL: https://www.youtube.com/watch?v=NL-qx_24zCk [accessed 2025-11-25]
87. Archibald MM, Lawless M, Gill TK, Chehade MJ. Orthopaedic surgeons' perceptions of frailty and frailty screening. BMC Geriatr 2020 Jan 16;20(1):17 [FREE Full text] [doi: 10.1186/s12877-019-1404-8] [Medline: 31948387]
88. Assarroudi A, Heshmati Nabavi F, Armat MR, Ebadi A, Vaismoradi M. Directed qualitative content analysis: the description and elaboration of its underpinning methods and data analysis process. J Res Nurs 2018 Feb 10;23(1):42-55 [FREE Full text] [doi: 10.1177/1744987117741667] [Medline: 34394406]
89. Archibald MM, Dart A, Wicklow B, Pundyk KJ, Marks SD, Sellers EAC. Beyond AIC: an interpretive descriptive qualitative study of youth experiences and perceptions of living with type 2 diabetes. J Adv Nurs 2025 Sep 25:70230 [FREE Full text] [doi: 10.1111/jan.70230] [Medline: 40995871]
90. Zipfel N, Horreh B, Hulshof CTJ, de Boer AGEM, van der Burg-Vermeulen SJ. The relationship between the living lab approach and successful implementation of healthcare innovations: an integrative review. BMJ Open 2022 Jun 28;12(6):e058630 [FREE Full text] [doi: 10.1136/bmjopen-2021-058630] [Medline: 35768105]
91. Mogård EV, Rørstad OB, Bang H. The relationship between psychological safety and management team effectiveness: the mediating role of behavioral integration. Int J Environ Res Public Health 2022 Dec 27;20(1):406 [FREE Full text] [doi: 10.3390/ijerph20010406] [Medline: 36612729]

Abbreviations

CIHR: Canadian Institute of Health Research
DER-CA: Diabetes Education Resource for Children and Adolescents
GRAMMS: Good Reporting of Mixed Methods Studies
iCARE: Improving Renal Complications in Adolescents with Type 2 Diabetes through Research
IHDCYH: Institute of Human Development, Child and Youth Health
iKT: integrated knowledge translation
KT: knowledge translation.
KU: knowledge user
PAG: participant advisory group
T1D: type 1 diabetes
T2D: type 2 diabetes

Edited by A Schwartz; submitted 09.Sep.2025; this is a non–peer-reviewed article;accepted 12.Nov.2025; published 19.Jan.2026.

Please cite as:
Archibald MM, Wicklow B, Sellers E, Griffiths A, Diffey L, McGavock J, Roos LE, Clark AM, Lopez J, Ho J, Ledoux E, Dzorka G, Parvin S, Fagbuyi O, Dart A
A Virtual Living Lab Platform Codeveloped for Mental Health in Youth-Onset Type 2 Diabetes (BrightSpark Care Lab): Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e83865
URL: https://www.researchprotocols.org/2026/1/e83865
doi:10.2196/83865
PMID:

©Mandy M. Archibald, Brandy Wicklow, Elizabeth Sellers, Arlene Griffiths, Linda Diffey, Jonathan McGavock, Leslie E. Roos, Alexander M. Clark, Jennifer Lopez, Josephine Ho, Ernestine Ledoux, Gifty Dzorka, Shahina Parvin, Oluwatoyosi Fagbuyi, Allison Dart. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Jan.2026. This is an

XSL•FO
RenderX

open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Traditional Chinese Medicine Syndrome Differentiation of Adult Patients With Type 2 Diabetes and Metabolic Syndrome: Protocol for a Cross-Sectional Study

Jialing Zhang[1,2], PhD; Hoi Ki Wong[1], MChinMed; Zhilin Lin[1], MMed; Shuyan Zhong[1], MChinMed; Minxia Ma[1], MChinMed; Xuejiao Wang[3], MMed

[1]Vincent V.C. Woo Chinese Medicine Clinical Research Institute, School of Chinese Medicine, Hong Kong Baptist University, Hong Kong SAR, China (Hong Kong)

[2]School of Chinese Medicine, Hong Kong Baptist University, Hong Kong SAR, China (Hong Kong)

[3]Children's Diagnosis and Treatment Center of the Affiliated Hospital, Changchun University of Chinese Medicine, Changchun, Jilin, China

Corresponding Author:
Jialing Zhang, PhD
Vincent V.C. Woo Chinese Medicine Clinical Research Institute
School of Chinese Medicine
Hong Kong Baptist University
5/F, David C.Lam Building, 34 Renfrew Road, Kowloon Tong
Hong Kong SAR, 999077
China (Hong Kong)
Phone: 852 34115024
Email: zhangjialing@hkbu.edu.hk

Abstract

Background: The global burden of type 2 diabetes mellitus (T2DM) and metabolic syndrome (MetS) continues to rise, with these conditions significantly increasing risks of cardiovascular disease, disability, and mortality. Traditional Chinese Medicine (TCM) syndrome differentiation, a cornerstone of TCM practice, guides diagnosis and treatment by identifying patterns of disharmony. However, large-scale studies investigating TCM syndrome patterns in T2DM comorbid with MetS remain scarce.

Objective: This cross-sectional study aims to characterize TCM syndrome profiles in a population diagnosed with T2DM and MetS and evaluate their diagnostic relevance.

Methods: This cross-sectional study will enroll a cohort of 470 participants diagnosed with T2DM and MetS. All participants will undergo comprehensive assessments, including the Syndrome Differentiation Questionnaire for T2DM and MetS, demographic and anthropometric measurements, biochemical profiling (eg, fasting glucose, glycosylated hemoglobin, and lipid panel), dietary measurement (Food Frequency Questionnaire), physical activity measurement (International Physical Activity Questionnaire Short Form), sleep quality evaluation (Pittsburgh Sleep Quality Index), quality-of-life assessment (Audit of Diabetes-Dependent Quality of Life), stroke risk estimation (Framingham Stroke Risk Score), and retinal imaging. Latent class analysis will be used to identify the TCM syndrome patterns. Factor analysis will be employed to identify core TCM syndrome factors. Hierarchical cluster analysis will be performed to classify TCM syndrome elements, and logistic regression will examine associations between syndrome differentiation, metabolic parameters, lifestyle factors, and disease progression.

Results: This trial was registered on November 17, 2024. Participant recruitment for this study was initiated in November 2024. As of October 2025, more than 450 eligible participants have been enrolled and have completed data collection. Recruitment is scheduled to conclude on December 31, 2025.

Conclusions: As the first large-scale clinical study to systematically characterize TCM syndrome differentiation in T2DM-MetS comorbidity, this research will establish syndrome profiles associated with metabolic parameters, lifestyle factors, and disease progression. The findings are expected to provide a framework for integrating TCM syndrome differentiation into chronic disease management, ultimately contributing to personalized treatment strategies and improved patient outcomes in integrative medicine.

Trial Registration: ClinicalTrials.gov NCT06703684; https://clinicaltrials.gov/study/NCT06703684

International Registered Report Identifier (IRRID): DERR1-10.2196/86217

(JMIR Res Protoc 2026;15:e86217) doi:10.2196/86217

XSL•FO
RenderX

KEYWORDS

Traditional Chinese Medicine; syndrome differentiation; type 2 diabetes; metabolic syndrome; personalized treatment

Introduction

Diabetes mellitus (DM) presents a critical global health challenge, with its prevalence escalating rapidly. As of 2021, approximately 537 million adults were affected, with projections estimating a rise to 783 million by 2045 without effective preventive measures [1]. Type 2 diabetes mellitus (T2DM) constitutes over 90% of these cases, underscoring its significance in the diabetes epidemic [2]. In Hong Kong, T2DM impacts nearly 10% of the adult population, mirroring global trends [3]. Adults with DM exhibit a 2-4 times higher risk of cardiovascular complications compared to adults with no DM [4]. Alarmingly, 65%-85% of patients with T2DM meet the diagnostic criteria for metabolic syndrome (MetS) [5,6], characterized by abdominal obesity, hypertension, dysglycemia, and dyslipidemia [7,8]. MetS impacts 20%-25% of the global population [9] and serves as an independent predictor of mortality among individuals with diabetes. This coexistence exacerbates metabolic abnormalities and increases cardiovascular risk, underscoring the imperative for early diagnosis and intervention to enhance management and prevent chronic complications [10,11].

Although conventional pharmacotherapy targets individual MetS components (eg, hypoglycemics, statins, and antihypertensives), emerging evidence supports Traditional Chinese Medicine (TCM) as a complementary approach. Clinical trials demonstrate that integrating TCM with Western regimens yields superior improvements in anthropometrics (eg, waist/hip circumference), glycemic control (glucose, glycosylated hemoglobin), insulin resistance index, and triglycerides, with minimal adverse reactions [12]. This aligns with a systematic review of 16 randomized controlled trials, where TCM significantly enhanced weight management and metabolic regulation [13]. Mechanistically, TCM compounds modulate lipid metabolism, vascular endothelial function, and insulin sensitivity via the phosphatidylinositol 3-kinase/protein kinase B signaling pathway, peroxisome proliferator-activated receptor pathway, and AMP-activated protein kinase pathway, offering multitargeted correction of metabolic dysregulation [14]. Together, clinical and pharmacological evidence supports TCM's role in addressing the complex interplay of T2DM and MetS.

TCM practice relies on accurate diagnosis and treatment procedures known as Bian Zheng Lun Zhi (syndrome differentiation followed by treatment procedures). This vital concept involves a comprehensive analysis of clinical information obtained through 4 main diagnostic methods: observation, listening, questioning, and pulse analysis. The integrated data then guides the selection of TCM therapeutic interventions such as acupuncture and herbal formula prescriptions [15]. For DM, different syndromes exhibit distinct metabolic characteristics and corresponding clinical biomarkers, thereby guiding TCM therapies in alignment with this principle [16]. However, consensus on syndrome differentiation among practitioners is often hindered by a lack of standardized terminology [17], and research on syndrome differentiation for T2DM comorbid with MetS remains limited. This study aims to identify prevalent TCM syndromes in individuals with T2DM and MetS. Additionally, it will examine how these syndromes are associated with health-related behaviors (eg, physical activity, diet, sleep) and quality of life. By identifying specific TCM syndrome patterns that will correlate with particular metabolic abnormalities and varying risks of disease progression, this study may enhance diagnosis and guide personalized treatment strategies.

Methods

Study Design

We present a protocol for a cross-sectional study. This study will be conducted at the Chinese Medicine Clinics of Hong Kong Baptist University in Hong Kong Special Administrative Region, China. Data collection will be conducted over a 12-14 month period, commencing in November 2024.

Ethical Considerations

This study has been approved by the Research Ethics Committee of Hong Kong Baptist University (REC/23-24/0564) and has been registered at ClinicalTrials.gov (NCT06703684). Volunteers will be required to provide written informed consent. The reporting of this study follows the recommendations of the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines [18].

Study Procedure and Population

We aim to recruit a cohort of 470 adult participants (age ≥18 years) diagnosed with both T2DM and MetS [19]. The inclusion and exclusion criteria are detailed in Textbox 1. Following screening, enrolled participants will proceed to complete the study assessments. Participants who complete all study procedures will receive a transportation allowance of HK$100.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for participation.

Inclusion criteria

- Age between 18 and 75 years
- Diagnosed with type 2 diabetes mellitus
- Diagnosed with metabolic syndrome according to guidelines for the prevention and treatment of type 2 diabetes mellitus in China (2020 edition) from Chinese Diabetes Society. People who meet 3 of the following diagnostic criteria or more can be diagnosed with metabolic syndrome [19]: (1) abdominal obesity, that is, having waist circumference ≥90 cm for males and ≥85 cm for females; (2) hyperglycemia, that is, having fasting blood glucose ≥6.1 mmol/L or 2-hour postprandial glucose ≥7.8 mmol/L, and/or those who have been diagnosed with hyperglycemia and in treatment; (3) hypertension, that is, those with blood pressure ≥130/85 mm Hg, and/or those who have been diagnosed with hypertension and in treatment; (4) fasting triglyceride ≥1.70 mmol/L; and (5) fasting high-density lipoprotein cholesterol <1.04 mmol/L
- Clearly understand the trial and voluntarily sign the informed consent form

Exclusion criteria

- Diagnosed with type 1 diabetes, steroid-induced diabetes, gestational diabetes, or specific types of diabetes
- Diabetes accompanied by severe complications such as diabetic nephropathy and diabetic ketoacidosis
- Secondary obesity (eg, secondary to pituitary inflammation, tumor)
- Secondary hypertension (eg, pheochromocytoma, renal hypertension)
- Secondary hyperlipidemia (eg, hypothyroidism, nephrotic syndrome)
- With severe heart, liver, or kidney disease or bleeding disorders, or with other serious diseases (eg, cancer, dementia)
- With severe mental disorders, speech, or hearing impairments
- Pregnant or lactating females
- Currently in another clinical trial

Study Outcomes

The primary outcome will be the TCM syndrome differentiation in patients with T2DM and MetS, as assessed by the Syndrome Differentiation Questionnaire for T2DM and MetS (SDQTM). SDQTM is developed according to the Guideline for the Integrated Diagnosis and Treatment of Metabolic Syndrome [20], Guideline for the Integrated Diagnosis and Treatment of T2DM [21], and Guiding Principles for Clinical Research of New Chinese Medicine [22]. It includes 69 items (Textbox 2) that capture disease nature (eg, wind, dryness, yang deficiency) and location (eg, heart, kidney) [23]. Each item is a closed-ended question featuring a ranked scale with descriptors for symptom severity and frequency. Registered Chinese medicine practitioners will administer SDQTM in a face-to-face consultation integrating the 4 TCM diagnostic methods.

XSL•FO
RenderX

Textbox 2. Items of The Syndrome Differentiation Questionnaire for Type 2 Diabetes Mellitus and Metabolic Syndrome (SDQTM).

1. Fatigue and weakness
2. Short of breath and want of speech
3. Spontaneous sweating
4. Pale or sallow yellow complexion
5. Pale lips and fingernails
6. Dizziness
7. Dry throat and mouth
8. Sensation of heat in the palms, soles, and chest
9. Hot flashes or night sweats
10. Aversion to cold and cold limbs
11. Puffy face and edema of the feet
12. Weak cough
13. Susceptibility to common colds
14. Reduced food intake and poor appetite
15. Abdominal distension after eating
16. Lower back and knee weakness
17. Tinnitus and deafness
18. Loose teeth and hair loss
19. Decreased libido
20. Palpitations
21. Insomnia and profuse dreaming
22. Forgetfulness
23. Blurred vision
24. Lusterless nails
25. Numbness in the hands and feet
26. Dry eyes
27. Irritability or depression with frequent sighing
28. Cold pain in the abdomen
29. Thirst with desire for cold drinks
30. Excessive eating and easy hunger
31. Fullness in chest and hypochondrium
32. Distending pain in the abdomen
33. Generalized edema
34. Heavy sensation in the limbs
35. Dry mouth and bitter taste
36. Dusky or dull dark complexion
37. Purplish dark lips
38. Squamous and dry skin
39. Fixed pain
40. Petechiae on the skin
41. Nausea
42. Vomiting
43. Itching skin

XSL•FO
RenderX

44. Sticky, greasy sensation in the mouth
45. Dark yellow or reddish urine
46. Frequent urination in large volume
47. Frequent nocturia
48. Scanty urine
49. Dry hard stools
50. Loose stools
51. Atonic constipation
52. Dark urine and constipation
53. Enlarged tongue with teeth marks
54. Pale tongue
55. Red tongue with scant coating
56. Pale, enlarged tongue with a moist coating
57. Red tongue with yellow coating
58. Yellow and greasy tongue coating
59. Purplish dark tongue, sublingual vein engorgement
60. Slippery and greasy tongue coating
61. Weak and forceless pulse
62. Thready and weak pulse
63. Thready and rapid pulse
64. Deep and slow pulse
65. Rapid pulse
66. Wiry pulse
67. Wiry and slippery pulse
68. Choppy or intermittent pulse
69. Slippery pulse

The secondary outcomes include demographic and anthropometric measurements, blood pressure, fasting glucose, glycosylated hemoglobin, lipid panel (total cholesterol, triglycerides, high-density lipoprotein, low-density lipoprotein), International Physical Activity Questionnaire-Short Form, Food Frequency Questionnaire, Pittsburgh Sleep Quality Index, Audit of Diabetes-Dependent Quality of Life, Framingham Stroke Risk Score (FSRS), retinal imaging, and concurrent medications.

Physical activity levels will be assessed using International Physical Activity Questionnaire-Short Form by recalling the time and days spent on 4 intensity levels (vigorous, moderate, walking, and sitting) over the previous 7 days [24]. Dietary intake will be assessed using a Food Frequency Questionnaire that is developed for older adults in Hong Kong and requires less than 30 minutes to complete [25]. The Pittsburgh Sleep Quality Index, a 19-item instrument, will be used to evaluate sleep quality over a 1-month interval, with its items generating 7 component scores that are summed into a global score ranging from 0 to 21, where higher scores indicate poorer sleep quality [26]. Audit of Diabetes-Dependent Quality of Life assesses diabetes-specific quality of life through 2 overview items and 19 items covering domains that include leisure, work, physical health, social relationships, and living conditions. A mean weighted impact score will be calculated, with lower scores indicating poorer diabetes-specific quality of life [27]. The FSRS will be used to estimate the individual risk of cardiovascular events, calculated based on age, smoking status, use of antihypertensive medications, blood pressure, diabetes status, high-density lipoprotein, and cholesterol levels [28]. Retinal microvasculature will be noninvasively visualized and quantified by fundus photography and computerized image analysis via a cloud-based artificial intelligence platform (Airdoc, Beijing Airdoc Technology Co, Ltd) to analyze its relation to cardiovascular events [29].

Data Management and Monitoring

Participant confidentiality will be maintained using unique identifier codes, with de-identified data entered directly into the secure REDCap (Research Electronic Data Capture) database [30]. Electronic and hard-copy data will be stored on password-protected university servers and in locked cabinets, respectively, and retained for 7 years after the study. As this noninterventional study presents no anticipated risks, a data

XSL•FO
RenderX

monitoring committee is not required, and monitoring will be performed by the principal investigator and the research team.

Sample Size Calculation

The minimum required sample size is calculated using the single-population proportion formula [31], based on T2DM with MetS prevalence of approximately 11.2% in Hong Kong [32]. With a 95% confidence level, a precision of SD 3%, and an assumed nonresponse rate of 10%, the study aims to recruit 470 participants.

Statistical Analysis

Statistical analysis will be performed using SPSS (version 29.0; IBM Corp), Python (version 3.13.9), and R (version 4.5.2; R Foundation for Statistical Computing) software. Missing values will be imputed by multiple imputation using the chained equation method. Quantitative data will be summarized as mean (SD) if normally distributed or as median and interquartile range otherwise. Latent class analysis will be used to identify the TCM syndrome patterns. Factor analysis will be used to reduce the dimensions of TCM syndrome elements and explore their underlying structure. Data suitability will be evaluated using the Kaiser-Meyer-Olkin measure and Bartlett's test of sphericity. Principal component analysis with varimax rotation will be employed to extract the common factors. Based on the factor analysis results, hierarchical cluster analysis will be performed to classify TCM syndrome elements. Cluster boundaries will be determined by identifying large distances between successive fusion levels in the dendrogram. The adjusted Rand index will be used for measuring the agreement between classifications.

For the comparison of quantitative data, one-way analysis of variance will be used if the assumptions of normality and homogeneity of variance are met. Otherwise, the Kruskal-Wallis test will be used. Categorical data will be expressed as numbers and percentages, and comparisons will be conducted using the chi-square test or Fisher exact test, as appropriate. Spearman correlation analysis will be employed to evaluate correlations between continuous variables, and Kendall’s tau-b analysis will be used for ordinal variables. Logistic regression will be performed to examine the relationship between syndrome differentiation and MetS components. A 2-sided value of $P<.05$ will be considered statistically significant.

Results

This trial was registered on November 17, 2024. Participant recruitment for this study was initiated in November 2024. As of October 2025, more than 450 eligible participants have been enrolled and have completed data collection. Recruitment is scheduled to conclude on December 31, 2025.

Discussion

Overview

DM poses a critical global health challenge, with 65%-85% of individuals with T2DM also meeting the diagnostic criteria for MetS [5,6], underscoring a widespread profile of metabolic dysregulation. Given the high comorbidity, this large-scale cross-sectional study will be the first in Hong Kong to systematically characterize the distribution of TCM syndromes in this population. It will explore how TCM syndromes relate to important health-related behaviors such as physical activity, diet, and sleep quality, as well as quality of life. By clarifying the relationship between TCM syndromes and metabolic profiles, this study aims to combine TCM assessments with clinical indicators, support more tailored treatments, and underscore the value of TCM syndrome differentiation in managing and predicting outcomes in patients with T2DM and MetS.

Anticipated Principal Findings

Although TCM syndrome distributions in patients with T2DM and MetS have been previously examined separately [16,33,34], this study specifically focuses on their comorbidity. We anticipate that the TCM syndrome profile of T2DM-MetS comorbidity will demonstrate a distinct distribution pattern compared to either condition alone, exhibiting unique relationships with disease duration and severity. Given that both conditions represent independent cardiovascular risk factors, we further hypothesize that specific TCM syndrome patterns will correlate with particular metabolic abnormalities and varying risks of disease progression. Through the combined application of FSRS and retinal imaging, we will be able to quantitatively characterize cardiovascular risks across different syndrome patterns. Furthermore, we expect these TCM syndromes to be influenced by modifiable health behaviors and associated with differences in the quality of life. Collectively, these findings will provide evidence supporting the clinical value of TCM syndrome differentiation in prognostic stratification and personalized management of T2DM and MetS.

Strengths and Limitations

This study has several strengths, including its design as the first large-scale cross-sectional investigation systematically characterizing TCM syndromes in individuals with T2DM and MetS. This research innovatively integrates clinical indicators such as biochemical profiling and retinal imaging with TCM syndrome differentiation while also incorporating evaluations of health behaviors and quality of life. However, several limitations should be acknowledged. The cross-sectional design inherently prevents establishing causal relationships, and reliance on self-reported data may introduce recall bias. Additionally, potential confounding factors could affect result interpretation, and the generalizability of the findings beyond the studied population requires further validation.

Implications for Practice or Research

The findings may provide practitioners with a practical framework for incorporating TCM syndrome differentiation into T2DM and MetS management. By identifying syndrome patterns associated with metabolic profiles and lifestyle factors, this approach supports personalized treatment strategies that combine biomedical and traditional diagnostic perspectives. The integration of retinal imaging and cardiovascular risk assessments with TCM diagnostics further provides a model for comprehensive evaluation in integrative medicine practice.

Beyond these clinical implications, this study underscores the need for validated TCM syndrome assessment tools to

XSL•FO
RenderX

standardize future research. It also calls for longitudinal research to establish causal relationships between syndrome patterns and disease progression. Further investigation is warranted to explore the biological mechanisms underlying identified syndrome configurations, potentially through multi-omics approaches [35]. Intervention trials targeting specific syndrome patterns would advance evidence for personalized TCM therapies.

Conclusion

This study will first systematically characterize TCM syndromes in patients with T2DM and MetS by integrating clinical indicators with traditional diagnostic approaches. The findings are expected to reveal significant relationships between syndrome differentiation, metabolic parameters, lifestyle factors, and disease progression. This work may establish a framework for incorporating TCM syndrome differentiation into chronic disease management, thereby contributing to the development of personalized treatment strategies and improved patient outcomes in integrative medicine practice.

Dissemination

Results will be disseminated through peer-reviewed journals and presentations at international conferences focused on metabolic disorders, ensuring accessibility to clinicians, researchers, and policymakers.

Acknowledgments

We would like to express our gratitude to authors for their contributions to this study as well as to the editors and reviewers for their insightful feedback and constructive suggestions.

Data Availability

The datasets generated or analyzed for this study are available from the corresponding author on reasonable request.

Funding

This study is supported by the Chinese Medicine Development Fund (23B2_067A_R1). The funder has no role in the study design, data collection, analysis, interpretation, or manuscript preparation.

Authors' Contributions

JZ conceived and designed the study. JZ, HKW, ZL, SZ, MM, and XW drafted the protocol. All authors contributed to the research, read, and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
STROBE checklist.
[PDF File (Adobe PDF File), 60 KB - resprot_v15i1e86217_app1.pdf]

References

1. International Diabetes Federation. URL: https://idf.org/about-diabetes/diabetes-facts-figures/ [accessed 2025-10-15]
2. Lin X, Xu Y, Pan X, Xu J, Ding Y, Sun X, et al. Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. Sci Rep 2020 Sep 08;10(1):14790 [FREE Full text] [doi: 10.1038/s41598-020-71908-9] [Medline: 32901098]
3. Quan J, Li TK, Pang H, Choi CH, Siu SC, Tang SY, et al. Diabetes incidence and prevalence in Hong Kong, China during 2006-2014. Diabet Med 2017 Jul;34(7):902-908. [doi: 10.1111/dme.13284] [Medline: 27859570]
4. Dal Canto E, Ceriello A, Rydén L, Ferrini M, Hansen TB, Schnell O, et al. Diabetes as a cardiovascular risk factor: An overview of global trends of macro and micro vascular complications. Eur J Prev Cardiol 2019 Dec;26(2_suppl):25-32. [doi: 10.1177/2047487319878371] [Medline: 31722562]
5. Alexander C, Landsman P, Teutsch S, Haffner S, Third National HealthNutrition Examination Survey (NHANES III), National Cholesterol Education Program (NCEP). NCEP-defined metabolic syndrome, diabetes, and prevalence of coronary heart disease among NHANES III participants age 50 years and older. Diabetes 2003 May;52(5):1210-1214. [doi: 10.2337/diabetes.52.5.1210] [Medline: 12716754]
6. Tong P, Kong A, So WY, Yang X, Ho CS, Ma RC, et al. The usefulness of the International Diabetes Federation and the National Cholesterol Education Program's Adult Treatment Panel III definitions of the metabolic syndrome in predicting coronary heart disease in subjects with type 2 diabetes. Diabetes Care 2007 May;30(5):1206-1211. [doi: 10.2337/dc06-1484] [Medline: 17259472]

XSL•FO
RenderX

7. Sattar N, McConnachie A, Shaper AG, Blauw GJ, Buckley BM, de Craen AJ, et al. Can metabolic syndrome usefully predict cardiovascular disease and diabetes? Outcome data from two prospective studies. Lancet 2008 Jun;371(9628):1927-1935. [doi: 10.1016/s0140-6736(08)60602-9]
8. Grundy SM. Metabolic syndrome pandemic. Arterioscler Thromb Vasc Biol 2008 Apr;28(4):629-636. [doi: 10.1161/ATVBAHA.107.151092] [Medline: 18174459]
9. Jemal A, Girum T, Kedir S, Bedru A, Mosa H, Assfa K, et al. Metabolic syndrome and its predictors among adults seeking medical care: a trending public health concern. Clin Nutr ESPEN 2023 Apr;54:264-270. [doi: 10.1016/j.clnesp.2023.01.034] [Medline: 36963872]
10. Thomas GN, Schooling CM, McGhee SM, Ho S, Cheung BMY, Wat NMS, Hong Kong Cardiovascular Risk Factor Prevalence Study Steering Committee. Metabolic syndrome increases all-cause and vascular mortality: the Hong Kong Cardiovascular Risk Factor Study. Clin Endocrinol (Oxf) 2007 May;66(5):666-671 [FREE Full text] [doi: 10.1111/j.1365-2265.2007.02798.x] [Medline: 17381490]
11. Cull CA, Jensen CC, Retnakaran R, Holman RR. Impact of the metabolic syndrome on macrovascular and microvascular outcomes in type 2 diabetes mellitus: United Kingdom Prospective Diabetes Study 78. Circulation 2007 Nov 06;116(19):2119-2126. [doi: 10.1161/circulationaha.107.733428]
12. Chen Y, Fu D, Chen Y, He Y, Fu X, Xu Y, et al. Effects of Chinese herbal medicine Yiqi Huaju Formula on hypertensive patients with metabolic syndrome: a randomized, placebo-controlled trial. J Integr Med 2013 May;11(3):184-194. [doi: 10.3736/jintegrmed2013031] [Medline: 23743162]
13. Wu H, Tian J, Dai D, Liao J, Wang X, Wei X, et al. Efficacy and safety assessment of traditional Chinese medicine for metabolic syndrome. BMJ Open Diabetes Res Care 2020 Mar;8(1):e001181 [FREE Full text] [doi: 10.1136/bmjdrc-2020-001181] [Medline: 32220922]
14. Pan Y, Kong L. High fructose diet-induced metabolic syndrome: pathophysiological mechanism and treatment by traditional Chinese medicine. Pharmacol Res 2018 Apr;130:438-450. [doi: 10.1016/j.phrs.2018.02.020] [Medline: 29471102]
15. Jiang M, Lu C, Zhang C, Yang J, Tan Y, Lu A, et al. Syndrome differentiation in modern research of traditional Chinese medicine. J Ethnopharmacol 2012 Apr 10;140(3):634-642 [FREE Full text] [doi: 10.1016/j.jep.2012.01.033] [Medline: 22322251]
16. Wang J, Ma Q, Li Y, Li P, Wang M, Wang T, et al. Research progress on Traditional Chinese Medicine syndromes of diabetes mellitus. Biomed Pharmacother 2020 Jan;121:109565 [FREE Full text] [doi: 10.1016/j.biopha.2019.109565] [Medline: 31704615]
17. Cheng F, Wang X, Song W, Lu Y, Li X, Zhang H, et al. Biologic basis of TCM syndromes and the standardization of syndrome classification. Journal of Traditional Chinese Medical Sciences 2014 Oct;1(2):92-97. [doi: 10.1016/j.jtcms.2014.09.005]
18. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457 [FREE Full text] [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
19. Chinese Chinese Diabetes. Guideline for the prevention and treatment of type 2 diabetes mellitus in China (2020 edition). Chinese Journal of Diabetes Mellitus 2021 Apr;13(4):315-409. [doi: 10.3760/cma.j.cn115791-20210221-00095]
20. Endocrinology and Metabolic Diseases Committee of Integrated Traditional Chinese and Western Medicine Branch of Chinese Medical Doctor Association. Guideline for the integrated diagnosis and treatment of metabolic syndrome. World Chinese Medicine 2023 Nov;18(22):3157-3166. [doi: 10.3969/j.issn.1673-7202.2023.22.001]
21. Pang G, Ni Q, Zhang F. Guideline for the integrated diagnosis and treatment of type 2 diabetes mellitus. Journal of Traditional Chinese Medicine 2021 Feb;62(04):361-368. [doi: 10.13288/j.11-2166/r. 2021.04.019]
22. Zheng X. Guiding Principles for Clinical Research of New Chinese Medicine. Beijing: China Medical Science Press; 2002.
23. Zhao N, Huang X, Qi Z, Yang S. Research on TCM pathologic features of metabolic syndrome based on factor analysis. Journal of Liaoning University of TCM 2014;16(9):91-93. [doi: 10.13194/j.issn.1673-842x.2014.09.031]
24. Liou YM, Jwo CJC, Yao KG, Chiang L, Huang L. Selection of appropriate Chinese terms to represent intensity and types of physical activity terms for use in the Taiwan version of IPAQ. Journal of Nursing Research 2008;16(4):252-263. [doi: 10.1097/01.JNR.0000387313.20386.0a]
25. Chan VW, Zhou JH, Li L, Tse MT, You JJ, Wong M, et al. Reproducibility and relative validity of a short food frequency questionnaire for Chinese older adults in Hong Kong. Nutrients 2024 Apr 11;16(8):1132 [FREE Full text] [doi: 10.3390/nu16081132] [Medline: 38674823]
26. Chong A, Cheung CK. Factor structure of a Cantonese-version Pittsburgh sleep quality index. Sleep and Biological Rhythms 2012 Jan 19;10(2):118-125 [FREE Full text] [doi: 10.1111/j.1479-8425.2011.00532.x]
27. Bradley C, Todd C, Gorton T, Symonds E, Martin A, Plowright R. The development of an individualized questionnaire measure of perceived impact of diabetes on quality of life: the ADDQoL. Qual Life Res 1999;8(1-2):79-91. [doi: 10.1023/a:1026485130100] [Medline: 10457741]

XSL•FO
RenderX

28. D'Agostino RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008 Jan 22;117(6):743-753. [doi: 10.1161/circulationaha.107.699579]
29. Gu C, Wang Y, Jiang Y, Xu F, Wang S, Liu R, et al. Application of artificial intelligence system for screening multiple fundus diseases in Chinese primary healthcare settings: a real-world, multicentre and cross-sectional study of 4795 cases. Br J Ophthalmol 2024 Feb 21;108(3):424-431 [FREE Full text] [doi: 10.1136/bjo-2022-322940] [Medline: 36878715]
30. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
31. Naing L, Nordin RB, Abdul Rahman H, Naing YT. Sample size calculation for prevalence studies using Scalex and ScalaR calculators. BMC Med Res Methodol 2022 Jul 30;22(1):209 [FREE Full text] [doi: 10.1186/s12874-022-01694-7] [Medline: 35907796]
32. Chan NN, Kong APS, Chan JCN. Metabolic syndrome and type 2 diabetes: the Hong Kong perspective. Clin Biochem Rev 2005 Aug;26(3):51-57 [FREE Full text] [Medline: 16450012]
33. Dou Z, Xia Y, Zhang J, Li Y, Zhang Y, Zhao L, et al. Syndrome differentiation and treatment regularity in Traditional Chinese Medicine for type 2 diabetes: a text mining analysis. Front Endocrinol (Lausanne) 2021;12:728032 [FREE Full text] [doi: 10.3389/fendo.2021.728032] [Medline: 35002950]
34. Xia S, Cai J, Chen J, Lin X, Chen S, Gao B, et al. Factor and cluster analysis for TCM syndromes of real‐world metabolic syndrome at different age stage. Evidence-Based Complementary and Alternative Medicine 2020 Jul 08;2020(1):7854325. [doi: 10.1155/2020/7854325]
35. Wang W, Yang L, Li W, Sun Y, Sun H, Chen Y, et al. Multi-omics driven paradigm for construction of traditional Chinese Medicine Zheng (syndrome) diagnosis and treatment model, taking Shi Zheng (syndrome of dampness) as an example. Chin Med 2025 Mar 08;20(1):33 [FREE Full text] [doi: 10.1186/s13020-025-01085-2] [Medline: 40057776]

Abbreviations

DM: diabetes mellitus
FSRS: Framingham Stroke Risk Score
MetS: metabolic syndrome
REDCap: Research Electronic Data Capture
SDQTM: Syndrome Differentiation Questionnaire for Type 2 Diabetes Mellitus and Metabolic Syndrome
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
T2DM: type 2 diabetes mellitus
TCM: Traditional Chinese Medicine

Edited by J Sarvestan; submitted 21.Oct.2025; this is a non–peer-reviewed article;accepted 15.Dec.2025; published 23.Jan.2026.

Please cite as:
Zhang J, Wong HK, Lin Z, Zhong S, Ma M, Wang X
Traditional Chinese Medicine Syndrome Differentiation of Adult Patients With Type 2 Diabetes and Metabolic Syndrome: Protocol for a Cross-Sectional Study
JMIR Res Protoc 2026;15:e86217
URL: https://www.researchprotocols.org/2026/1/e86217
doi:10.2196/86217
PMID:

©Jialing Zhang, Hoi Ki Wong, Zhilin Lin, Shuyan Zhong, Minxia Ma, Xuejiao Wang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Evaluating an AI-Enabled Mobile Mental Health Monitoring Tool Among Family Caregivers of Adults Living With Cancer: Single-Arm Feasibility and Acceptability Trial Protocol

Chiara Acquati[1,2], MSW, PhD; Michael Aratow[3], MD; Tahmida Nazreen[3], PharmD; Arunima Bhattacharjee[1], MSW; Isabella K Marra[4]; Ashley S Alexander[5], MHSA

[1]Graduate College of Social Work, University of Houston, Houston, TX, United States

[2]Department of Family Medicine and Community Health, University of Minnesota, Minneapolis, MN, United States

[3]Ellipsis Health, San Francisco, CA, United States

[4]Fertitta Family College of Medicine, University of Houston, Houston, TX, United States

[5]Kelsey Research Foundation, Houston, TX, United States

Corresponding Author:
Chiara Acquati, MSW, PhD
Graduate College of Social Work
University of Houston
3511 Cullen Blvd
Houston, TX, 77004
United States
Phone: 1 713 743 4343
Email: cacquati@central.uh.edu

Abstract

Background: Psychological distress, particularly symptoms of depression and anxiety (D&A), is highly prevalent among family caregivers of individuals living with cancer, who often assume central roles in care coordination, treatment adherence, symptom monitoring, and emotional support. Rates of distress among caregivers frequently equal or exceed those observed in patients themselves. Despite increased attention to caregivers' mental health needs, routine distress screening remains limited in oncology care settings. Advances in mobile health technology and artificial intelligence (AI) offer opportunities to address these needs by providing accessible and user-driven tools. The Ellipsis Caregiver Assessment Enhancement (eCARE; Ellipsis Health, Inc) is a speech-based, AI-enabled mobile app designed to screen and monitor symptoms of depression and anxiety. By collecting brief voice recordings and in-app survey data, eCARE offers a scalable approach for integrating caregiver distress monitoring into cancer care.

Objective: This single-arm trial will evaluate the feasibility and acceptability of the eCARE app among family members who are the primary caregivers of patients diagnosed with cancer within the past 5 years. Specifically, the study aims to (1) determine feasibility based on platform completion rates, (2) assess acceptability using validated measures, and (3) identify barriers and facilitators influencing the uptake and sustained use of eCARE.

Methods: In Phase 1, a total of 60 United States–based family caregivers will be recruited from community health clinics, cancer and caregiving advocacy groups, and online postings. Screened and enrolled caregivers will complete 6 eCARE sessions over an 8-week period. Pre- and posttest surveys assess depression, anxiety, caregiving burden, and relational processes. Feasibility will be evaluated based on the proportion of participants who complete at least 66% of weekly assessments, and acceptability will be assessed using the acceptability of intervention measure (AIM). In Phase 2, a total of 20 caregivers will be invited to participate in semi-structured online interviews to explore user experience, including perceived benefits, barriers to use, and preferences for future implementation. Qualitative data will be analyzed thematically to inform tool refinement.

Results: The study has received Institutional Review Board approval from the University of Houston. Participant recruitment and enrollment began in June 2024, with data collection expected to conclude by August 2025. Data analysis will begin in December 2025, with preliminary results anticipated by May 2026.

Conclusions: This study will generate preliminary evidence on the feasibility, acceptability, and utility of a speech-based, AI-enabled smartphone tool for monitoring D&A symptoms among family cancer caregivers. Findings will inform the design of a larger, fully powered trial and guide future implementation of remote psychological distress monitoring strategies in oncology

XSL•FO
RenderX

care. By offering a low-burden, caregiver-centered approach, eCARE has the potential to expand access to psychosocial support and facilitate timely identification of needs and coordination of services across cancer care settings.

International Registered Report Identifier (IRRID): DERR1-10.2196/83276

(JMIR Res Protoc 2026;15:e83276) doi:10.2196/83276

KEYWORDS

cancer; caregiving; psychological distress; speech-based monitoring; artificial intelligence (AI); feasibility & acceptability study; mobile health (mHealth); acoustic analysis; semantic analysis; depression; anxiety

Introduction

Background

A cancer diagnosis is often associated with psychological challenges affecting patients and the broader support network, including spouses, family members, and friends who assume the role of caregivers [1-4]. Caregiving entails practical and supportive tasks across the continuum of care: coordinating medical visits, building relationships with clinicians, communicating with the medical team, and contributing to decision-making, adherence to treatment plans, and symptom management [2,3]. These tasks carry physical, psychosocial, emotional, and financial consequences, rendering cancer caregivers highly vulnerable to stress and mental health problems [3,5,6]. Across studies, family caregivers of patients with cancer often experience greater psychological distress than patients [2,4,7-11], and cancer caregiving is rated as more burdensome than other high-intensity caregiving scenarios [12]. Yet, caregivers' mental health needs remain insufficiently addressed in routine cancer care, underscoring the need for targeted psychosocial interventions [13,14].

Caregiver psychological distress is common throughout the illness trajectory, with patterns that differ across diagnosis, active treatment, recurrence, and advanced disease. Between 55% and 95% of caregivers [15-17] report clinically significant distress or mental health conditions. Compared with the general US population, family caregivers of patients with advanced cancer frequently score below 30% on mental health measures [18]. Inadequate preparation for the caregiving role compounds distress, contributing to feelings of helplessness and regret as caregivers witness symptom burden [19,20]. Specific conditions are highly prevalent. Depression ranges from 16% to 67% and tends to increase as prognosis worsens [9,21-24]. Anxiety is even more common; a global meta-analysis estimates 46.5% of family caregivers experience clinically significant anxiety, with rates in advanced disease (~40%-42%) exceeding those observed in patients [8,22,25,26]. In some cancer types (eg, head and neck) up to 57% of caregivers report clinically significant anxiety and/or depression [27].

Multiple factors contribute to this heightened vulnerability. Caregiver risk factors include being female, younger age, financial strain, limited social support, high burden, low caregiving self-efficacy, and prolonged caregiving hours [22,24]. Family caregivers are more likely to experience psychological distress when the patient is younger, when the illness progresses and worsens, when the patient has limited awareness of their prognosis, or when the patient's overall health and daily functioning decline [8,22,24]. Relational and contextual risk factors include being a spousal caregiver, insecure attachment (eg, fear of abandonment or avoidance), and family conflicts, which amplify distress and cause psychological strain among family cancer caregivers [21,22,24].

Despite this evidence, caregiver support remains underresourced and inadequate. Family caregivers frequently report unmet informational and psychological needs [4,11,26,28-31]. Traditional in-person mental health services are often inaccessible due to cost, time, and logistical barriers, limiting opportunities to address caregivers' needs [1]. As a result, family caregivers experience sleep disruption, diminished quality of life, and reduced caregiving capacity, ultimately affecting patient care [1,12,32].

Digital mental health offers a scalable path to support cancer caregivers. Smartphone- and web-based interventions can extend reach, reduce access barriers, and deliver timely, evidence-based care. However, rigorous evaluation in oncology caregiving remains limited. To date, eHealth interventions for family caregivers of individuals living with cancer have been found to reduce depression and modestly improve quality of life versus usual care, supporting the feasibility and potential value of remote delivery for caregiver mental health. Still, effects were small and trials were few, with short follow-up and heterogeneous content, therefore underscoring the need for larger, higher-quality randomized controlled trials [1,6,33-37]. Building on this foundation, AI-enabled depression and anxiety (D&A) screening and monitoring apps have strong potential for flexible, real-time support [34,36,37]. Evidence shows that technology-based approaches can (1) improve health and psychological outcomes, (2) expand access to care, (3) be cost-effective, and (4) be delivered in self-paced, tailored formats [35,38-41]. Yet, family members remain excluded from national distress screening mandates, leaving their needs inconsistently identified and addressed in oncology care settings [42-45]. This evidence emphasizes an urgent need to design, test, and implement family caregiver-specific digital psychosocial interventions that can be integrated into routine cancer care.

Advancements and Gaps in Speech-Based Mental Health Assessment

Technological advancements in automated and AI-enabled speech analysis have significantly improved the ability to detect psychological symptoms [46-49]. Tools using voice data, particularly acoustic features, now offer a promising complement to traditional validated screening instruments. These tools have shown diagnostic performance comparable to

XSL•FO
RenderX

widely used psychometric scales [50,51] with added advantages including reduced user burden, minimal risk of human error (eg, missing data), and the potential for greater user comfort during verbal self-expression [52-54].

Nevertheless, several limitations persist. Most tools continue to rely on either acoustic (ie, how something is said) or semantic content (ie, what is said), without combining both for improved precision [49,55-57]. Additionally, many systems have focused solely on depression, with a smaller fraction targeting anxiety and even fewer addressing both concurrently [55-57]. This gap is particularly concerning for family caregivers of patients with cancer, a population in which anxiety symptoms are both prevalent and frequently co-occur with depressive symptomatology.

This study addresses these gaps by evaluating Ellipsis Caregiver Assessment Enhancement (eCARE; Ellipsis Health, Inc), a next-generation speech-based assessment platform designed to detect both D&A symptom severity using an integrated AI model. Unlike most existing systems, eCARE combines semantic and acoustic data inputs to provide dual assessments, enhancing precision. Additionally, it offers a clinician-facing dashboard to support referrals and care coordination, promoting a person-centered model of supportive care delivery. The Ellipsis Health platform has been technically validated in multiple peer-reviewed publications demonstrating its strong machine learning performance [58-62].

eCARE App

eCARE is a speech-based smartphone app linked to Ellipsis Health, which is a secure, cloud-based AI infrastructure. The app processes a weekly 90-second audio recording from users who respond to prompted topics, including those related to their caregiving stress and emotional well-being. The system analyzes both the acoustic and semantic characteristics of speech using advanced deep learning techniques, including transfer learning. This approach allows for more sophisticated and accurate identification of users' speech signals than traditional models that rely on fixed acoustic features such as pitch or volume [57,60,62].

The platform has demonstrated strong performance metrics, with reported area under the curve values of 0.85 for Patient Health Questionnaire-9 item (PHQ-9) and 0.84 for Generalized Anxiety Disorder-7 item (GAD-7) predictions (binary classification with a cutoff score of 10), root-mean-square error values of 4.25 and 4.47, and mean absolute error values of 3.13 and 3.23, respectively [63]. To date, eCARE remains the only available speech-based screening platform capable of producing concurrent D&A scores for family caregivers of patients with cancer. This dual capability is a critical feature for reducing participant fatigue during long-term monitoring and managing the acute mental health needs of this group.

The eCARE app is available for download on the Apple or the Google Play Store and can be used on iOS and Android mobile devices. After downloading the app on their smartphones, users are able to register in the app with the phone number they provided to the research team. Once a week, users will be prompted to log in to the app and record a 90-second response to a selected topic related to their mental health or caregiving responsibilities. Afterward, users will receive nonclinical feedback on their distress levels. The app also includes longitudinal tracking of responses and mental health crisis hotline information. On the provider side, eCARE features a HIPAA (Health Insurance Portability and Accountability Act)-compliant web portal that visualizes severity of distress trajectories over time, supporting timely mental health referrals and opportunities for direct integration into the cancer care continuum.

Specific Aims

The present study will (1) expand understanding of the psychosocial issues faced by family caregivers of individuals diagnosed with cancer and (2) generate actionable evidence to refine eCARE for responsive delivery in a way that is receptive to preferences for care, format, and optimal timing of screening. Specific Aim 1: establish the feasibility and acceptability of eCARE among family caregivers of patients with cancer in a single-arm prospective cohort study, using platform completion rate (feasibility) and the acceptability of the intervention measure (AIM; acceptability) as indicators. Specific Aim 2: qualitatively evaluate facilitators and barriers to eCARE uptake among a subset of participants (n=20) to inform iterative refinement. Expected outcomes include evidence that a caregiver-driven digital approach can increase the proportion of family caregivers who are screened and monitored, as well as generating concrete recommendations to optimize eCARE for future efficacy testing and implementation in survivorship care.

Methods

Study Design

This single-arm, prospective cohort study will use a mixed methods design to evaluate the feasibility and acceptability of eCARE, a digital AI-enabled tool designed to monitor psychological distress among family caregivers of individuals living with cancer. The study will:

1. Track psychological distress, specifically D&A over an 8-week period among a cohort of 60 family cancer caregivers;
2. Assess adherence to and perceived acceptability of the eCARE tool;
3. Conduct semi-structured qualitative interviews with a subsample of participants (n=20) to explore experiences with eCARE use.

Participants will use eCARE independently, following structured onboarding and usage instructions provided by the research team at baseline. Psychological distress will be monitored using standardized questionnaires and voice-based data collection, with app-based metrics and quantitative survey data supplemented by qualitative feedback (Table 1).

XSL•FO
RenderX

Table 1. Study measures and assessment time points.

Measures	Measure administration time points and personnel			
	T0/Baseline	T1-T6	T7	Administration
Depression (PHQ-9[a]) [64]	✓	✓	✓	Qualtrics/ platform data[b]
Anxiety (GAD-7[c]) [65]	✓	✓	✓	Qualtrics/ platform data
Quality of Life (CDC 4 items[d]) [66]	✓		✓	Qualtrics
Caregiving Burden (Short Form Zarit Burden Interview) [67]	✓		✓	Qualtrics
Closeness (Inclusion of the Other in the Self Scale) [68,69]	✓		✓	Qualtrics
Communication (Social Constraints Scale) [70,71]	✓		✓	Qualtrics
Responsiveness (Perceived Partner Responsiveness Scale) [72,73]	✓		✓	Qualtrics
Communal Motivation to Care (Partner-Specific Communal Motivation Scale) [74,75]	✓		✓	Qualtrics
eCARE[e] App Use		✓		Platform data
Acceptability of intervention measure (AIM) [76]			✓	Qualtrics
User Engagement Scale–Short Form [77]			✓	Qualtrics
Qualitative interviews			✓	RS[f] via Zoom

[a]PHQ-9: Patient Health Questionnaire-9 item.

[b]Platform data: eCARE data collection.

[c]GAD-7: Generalized Anxiety Disorder-7 item.

[d]CDC 4 items: Centers for Disease Control and Prevention 4-item Healthy Days Measure.

[e]eCARE: Ellipsis Caregiver Assessment Enhancement.

[f]RS: research staff.

Participant Eligibility

Participants will be eligible to enroll if they (1) self-identify as the primary caregiver or support person of an individual diagnosed with cancer within the past 5 years, (2) are 18 years of age or older, (3) have access to a smartphone capable of downloading and using the eCARE app, (4) are fluent in English, and (5) are able and willing to provide informed consent. To ensure adequate representation of caregiving experiences, a stratified sampling approach will be used to recruit equal numbers of caregivers who identify as non-Hispanic White (n=15), Black/African American (n=15), Hispanic and Latino (n=15), and Asian American, Native Hawaiian, and Pacific Islander (AANHPI; n=15). Participants will be excluded from the study if they (1) have a cognitive impairment; (2) have a speech impairment; and (3) have a severe mental illness that would impede the ability to provide informed consent or complete study activities. Eligibility will be self-reported through an initial study information survey. Participants will review a list of exclusionary criteria and be asked to self-select out of the study if any listed conditions apply. Eligibility will be verified as necessary via follow-up communication (eg, telephone screening). To minimize fraudulent responses, we will apply minimum time-on-task thresholds, checks for duplicate IPs or devices, and manual pattern review.

Participant Recruitment and Study Procedures

Recruitment Procedures

Recruitment efforts will be conducted in collaboration with the Kelsey Research Foundation and other community partners. A recruitment email will be distributed directly to potential participants via the Kelsey Research Foundation listserv, support group lists, and through oncology social workers. Potential participants may self-refer to the study in one of the following ways: (1) by accessing the study flyer's URL or scanning the QR code, which will link to an online eligibility screening survey containing study information; (2) by contacting study staff, who will provide additional information and instructions to complete the screening survey; and (3) by visiting the study landing page on the Kelsey Research Foundation website and completing a brief informational survey. With appropriate approval, flyers will also be distributed electronically through community partners and legitimate online platforms. Investigators will engage in additional recruitment at educational and support-focused events for family caregivers, including caregiving walks, conferences, and community events. This may include staffing a vendor table or collaborating with event organizers to distribute study recruitment materials through

XSL•FO
RenderX

event newsletters, programs, and digital channels. AANHPI cancer caregivers will also be recruited by leveraging the Collaborative Approach for AANHPI Research and Education (CARE) Registry. The CARE Registry maintains a HIPAA-compliant, institutional review board (IRB)–approved database of AANHPI individuals residing in the United States and US-Affiliated Pacific Islands who have expressed willingness to participate in health-related research. As an established research platform dedicated to reducing disparities in health care studies, CARE provides access to a diverse pool of potential participants. Partnering with this platform allows for a more efficient and targeted recruitment process, ensuring that the study reaches an underrepresented population while maintaining ethical and regulatory compliance. The contact information of the principal investigator, as well as the IRB approval statement and the contact information for the IRB of record, will be included in the informed consent materials of all the electronic surveys and qualitative interviews. Recruitment will span approximately 12-16 months to maximize sample size and ensure sufficient diversity in participant representation.

Enrollment and Participation

Interested individuals who meet initial eligibility criteria will be contacted by study staff to schedule an initial phone screening, after which informed consent will be obtained and baseline measures collected through a pretest electronic survey. Participants will be asked to use the eCARE app weekly for 6 sessions (T1-T6) over an 8-week period. Weekly phone-based reminders will prompt engagement with the platform. If an assessment is missed, the app will automatically issue 2 follow-up notifications within 24-48 hours. Platform usage data will be collected to assess adherence (eg, frequency and duration of app use). A brief posttest assessment (T7) will be conducted approximately 8 weeks after baseline; the survey will include the same instruments presented during the pretest assessment plus the acceptability (AIM questionnaire [76]) and user engagement questionnaires (User-Engagement Scale-Short Form [77]).

Qualitative Component

A purposive subsample of participants (n=20; 5 per racial and ethnic group) will be invited to complete a semi-structured qualitative interview exploring facilitators and barriers to eCARE uptake, in addition to recommended improvements and refinement. Interest in participating in this component of the project will be assessed during the phone screening and pretest survey outlined in Phase 1. Eligible family caregivers will be contacted at the end of the intervention by email and those who agree will be invited to complete a 60- to 90-minute Zoom (Zoom Video Communications, Inc) interview with a member of the research team. Serial recruitment will continue until thematic saturation is achieved, defined as no new codes or themes emerging in at least 2 consecutive interviews. This approach follows empirical and methodological recommendations, as many core themes in relatively homogenous qualitative samples are typically identified within the first 12 interviews [78-80] and sample sizes of 12-20 are generally sufficient for achieving saturation in focused qualitative studies [81-83]. By exceeding the standard threshold for saturation, the study's target sample size allows for a more comprehensive exploration of a diverse group of users' perspectives. Details regarding assessment points and measures are summarized in Table 1.

eCARE Intervention

eCARE is an evidence-based artificial intelligence (AI)–enabled D&A screening and monitoring tool using speech data, which analyzes 90 seconds of participant "conversation" on selected topics. After logging in and hitting "start," participants can "talk to" the app about their mental health by selecting among 3 different prompts. After each recording, users will see a numeric score and a visual interpretation of their results. Then, they will be taken to the next page, where PHQ-9 and GAD-7 will be collected. Caregivers are also presented with mental health and suicidality resources on the last screen of the session.

Participant Safety

An individual safety plan will be implemented for family caregivers of individuals diagnosed with cancer endorsing the suicide item of the PHQ-9 and for those who score above the cutoff scores for anxiety and depression on the selected instruments. In case of elevated distress registered during the use of the eCARE app, as evidenced by scores ≥10 on the GAD-7 and the PHQ-9 instruments, participants will have access to mental health resources both in the app and via separate email within 24 hours. For those endorsing the suicide risk item, the eCARE app provides immediate information to the 988 Suicide and Crisis Lifeline, with the contact number for the helpline and chat option visible at the end of the session. The Suicide & Crisis Helpline offers 24×7 call, text, and chat access to trained crisis counselors who can help people experiencing suicidal ideation, substance use, and/or mental health crisis, or any other kind of emotional distress.

In case of high distress, participants will receive an email about mental health resources. In the case of participants who endorse the suicidality item, the email will contain ad hoc resources such as 988 Suicide & Crisis Lifeline, 211, National Alliance on Mental Illness (NAMI) phone and chat option, as well as 911. These communications will be accompanied by a text message in the app as well to alert the participant of the elevated score and that resources have been shared. Participants can also contact the investigators or let the study team members know about their interest in being referred to these community-based organizations. The online eCARE platform will be monitored weekly; however, the system alerts the investigative team in real time for each missed assessment or for assessments where the suicide risk item is endorsed. Importantly, the eCARE application does not interface with other smartphone applications or services (eg, Apple Health or Android Health) and does not collect or transmit protected health information beyond what is consented to within the study protocol.

Study End Points

The primary feasibility end point will be determined by the number of times each participant completes the eCARE assessment out of a total of 6 possible time points over an 8-week period. Participants who complete the assessment at least 4 times (≥ 66.6% completion rate) will be considered to

XSL•FO
RenderX

have met the feasibility threshold. The feasibility study will be deemed successful if at least 60% of enrolled participants (n≥36) reach this threshold. The secondary acceptability end point will be assessed using the AIM, a psychometrically validated implementation science instrument designed to evaluate participants' perceptions of intervention acceptability. An average AIM score of 4 out of 5 will be interpreted as indicative of acceptable intervention uptake.

Measures

As part of this study, we are collecting a comprehensive set of variables to capture the multifaceted experiences of caregiving.

Sociodemographic Characteristics

Sociodemographic characteristics include participants' age, sex, gender identity, race and ethnicity, educational attainment, relationship status, employment status, insurance coverage, and household income.

Clinical Characteristics (Care Recipient)

Clinical characteristics of the care recipient—such as their age, sex, cancer diagnosis, cancer stage, and type of treatment—are recorded to contextualize the caregiving experience.

Caregiving Characteristics

To further understand the caregiving context, we will gather data on the length of the caregiving period, the caregiver's relationship to the care recipient, primary caregiver status, and cohabitation. We will also assess caregiving intensity through the number of hours of care provided per week, the types of tasks performed, and the care recipient's level of independence in activities of daily living, measured by the Katz Index [84].

Outcome Measures

Psychological Distress

Depression is measured by the PHQ-9, a validated, self-administered instrument widely used to screen for, diagnose, and monitor the severity of depressive symptoms [64]. Respondents rate the frequency with which they have experienced each symptom over the past 2 weeks on a 4-point Likert scale, ranging from 0 ("Not at all") to 3 ("Nearly every day"), yielding a total score between 0 and 27. Interpretation of PHQ-9 scores follows standard clinical guidelines: 0-4 (minimal depression), 5-9 (mild), 10-14 (moderate), 15-19 (moderately severe), and 20-27 (severe). A total score of 10 or higher is commonly used as a cutoff for identifying clinically significant depressive symptoms. Item 9 specifically assesses suicidal ideation and requires immediate clinical attention if endorsed. The PHQ-9 has demonstrated strong reliability and construct validity across diverse populations and care settings, including oncology. The excellent psychometric property of the PHQ-9 has been reported across different age and racial or ethnic groups, endorsing its utility among patients with cancer [85]. The GAD-7 is a widely used, self-administered instrument designed to assess the severity of generalized anxiety symptoms over the past 2 weeks. It includes 7 items corresponding to core symptoms of generalized anxiety disorder. Each item is rated on a 4-point Likert scale ranging from 0 ("Not at all") to 3 ("Nearly every day"), yielding a total score between 0 and 21 [65]. Interpretation of total scores follows established clinical cut points: 0-4 (minimal anxiety), 5-9 (mild), 10-14 (moderate), and 15-21 (severe). A score of 10 or above is commonly used as a threshold for identifying clinically significant anxiety symptoms, warranting further assessment. The GAD-7 demonstrated excellent internal consistency (α=.92), strong test-retest reliability, good criterion, construct, and convergent validity [65,86]. These strong psychometric properties have been confirmed across age groups and racial and ethnic backgrounds, including those diagnosed with cancer [87-89].

Closeness

The Inclusion of Other in the Self (IOS) Scale [68,90] is a widely used, single-item pictorial measure of interpersonal closeness. Grounded in self-expansion theory [91,92], the IOS assesses the degree to which individuals perceive their relationship partner as integrated into their sense of self. In the current study, the IOS was used to evaluate perceived emotional closeness between family cancer caregivers and their care recipients. Participants are shown seven pairs of circles labeled "Self" and "Other," ranging from no overlap (1) to almost complete overlap (7). They are asked to select the pair that best represents their relationship with the care recipient. This visual analog approach provides a rapid and intuitive assessment of relational closeness and has demonstrated validity across diverse populations [69,91,93,94]. The IOS has been linked to a variety of relationship and health outcomes and is commonly used in studies of romantic, familial, and community ties [92]. In this study, caregiver-reported IOS scores provide a snapshot of the emotional bond between caregiver and care recipient, an interpersonal factor that may influence psychological distress, motivation to engage in support tools, and the caregiving experience more broadly.

Caregiving Burden

Caregiving burden is assessed with the Zarit Burden Interview-Short Form (ZBI-12) a 12-item self-report measure used to assess the subjective burden experienced by informal caregivers [67]. It evaluates key dimensions of caregiver strain, including emotional, physical, and social impacts associated with caregiving responsibilities. Each item is rated on a 5-point Likert scale ranging from 0 ("Never") to 4 ("Nearly always"), yielding a total score ranging from 0 to 48, with higher scores indicating greater perceived burden. Standard scoring guidelines categorize burden levels as follows: 0-10 (no to mild burden); 10-20 (mild to moderate burden); >20 (high burden). The ZBI-12 has demonstrated strong reliability and validity across diverse caregiving populations, including those caring for individuals with dementia, cancer, and other serious illnesses.

Communication

The Social Constraints Scale (SCS) is a 15-item self-report instrument developed to assess the extent to which individuals perceive their social environment, particularly close others such as spouses or family members, as inhibiting or discouraging the expression of illness-related thoughts and emotions [70,71]. In the context of cancer, the SCS captures perceived social responses that constrain open communication, such as minimizing concerns, changing the subject, or expressing

XSL•FO
RenderX

discomfort when the patient discusses their cancer-related experiences. Each item is rated on a 4-point Likert scale ranging from 1 (“Never”) to 4 (“Often”), with higher total scores (ranging from 15 to 60), indicating greater perceived social constraint. The SCS has demonstrated strong psychometric properties across diverse populations coping with chronic illness.

Perceived Responsiveness

Caregivers’ perceptions of care recipient responsiveness were measured by the 12-item Perceived Partner Responsiveness Scale (PPRS) [72]. The PPRS measures the degree to which people feel their loved ones are responsive to them. It measures two dimensions of perceived responsiveness in close relationships: (1) understanding (7 items, eg, “My loved one gets the facts right about me;” “My loved one knows me well”) and (2) validation (5 items, eg, “My loved one values and respects the whole package that is the real me;” “My loved one seems interested in what I am thinking and feeling”). Responses were scored on a 9-point Likert scale ranging from 1 (not at all true) to 9 (completely true). A total score is calculated by summing all items [72], with scores indicating greater perceptions of others’ responsiveness.

Communal Motivation to Care

Communal motivation was assessed with the 10-item Partner-Specific Communal Motivation Scale [74]. Caregivers were asked to assess to what extent they were communally motivated to care for the well-being of their care recipients. Responses were scored on a 9-point Likert scale ranging from 1 (not at all) to 9 (extremely). Example items are “Helping my loved one is a high priority for me,” “I would sacrifice very much to help my loved one,” “I would be reluctant to sacrifice for my loved one.” The total score is given by the sum of the items after reversing items 2, 5, and 10, and indicates higher levels of communal motivation.

Quality of Life

The Centers for Disease Control and Prevention (CDC) healthy days measure, commonly referred to as health-related quality of life-4 (HRQOL-4), is a brief, validated instrument designed to assess health-related quality of life in population-based studies and public health surveillance [66]. The HRQOL-4 includes 4 core items, including self-rated general health, physically unhealthy days, mentally unhealthy days, and physical or mental health interference with daily activities. Each item captures a distinct domain of the HRQOL-4, allowing for both individual and composite assessments of overall well-being. The HRQOL-4 has demonstrated strong reliability and validity across diverse populations [66,95].

Acceptability of Intervention

The AIM is a pragmatic, 4-item self-report instrument developed to assess stakeholders’ perceptions of the acceptability of a specific intervention or implementation strategy [76]. The AIM captures the extent to which an intervention is perceived as agreeable, satisfactory, or palatable, with items rated on a 5-point Likert scale ranging from “completely disagree” to “completely agree.” The items include (1) “(Intervention) meets my approval,” (2) “(Intervention) is appealing to me,” (3) “I like (Intervention),” and (4) “I welcome (Intervention).” It has been designed for use across diverse stakeholder groups, including service providers, administrators, and caregivers, and it requires no specialized training for administration or interpretation. While no established cutoff scores exist, higher total scores indicate greater acceptability. The AIM has demonstrated strong psychometric properties, including content and structural validity, reliability, and responsiveness to change [76].

User Engagement

The User Engagement Scale (UES-SF) is a 12-item self-report instrument developed to assess user engagement in human-computer interaction contexts [77]. The short form includes four key subscales, each with 3 items: focused attention, perceived usability, aesthetic appeal, and reward. Participants respond using a 5-point Likert scale ranging from 1 (“Strongly Disagree”) to 5 (“Strongly Agree”). Per scoring instructions, items within the perceived usability subscale are reverse coded. Subscale scores are computed by averaging the responses within each domain, and a total engagement score is calculated by averaging across all 12 items. The UES-SF was designed for high usability, with items written at a fifth-grade reading level and no special training required for administration or interpretation. The instrument has demonstrated sound psychometric properties, including content and structural validity, internal consistency, and responsiveness to user experiences [77].

Data Analysis

Quantitative and qualitative data will be analyzed using a mixed methods approach to evaluate the feasibility and acceptability of the eCARE tool among family caregivers of individuals diagnosed with cancer. Descriptive statistics will be used to summarize participants’ sociodemographic, clinical, and caregiving characteristics, including means, SDs, ranges, and 95% CIs. Graphical methods (eg, histograms, boxplots, and scatterplots) will be used to examine variable distributions and identify the need for transformations. For comparisons, we will use ANOVA for continuous variables and chi-square tests for categorical variables. Differences in psychological distress, caregiving burden, relational closeness, perceived responsiveness, and motivation to provide care will also be explored. Feasibility, the primary outcome of Aim 1, will be assessed based on the number of participants completing at least 4 of 6 total weekly eCARE assessments during the 8-week study period. Participants exceeding 6 sessions due to more frequent use will be classified as having met the feasibility threshold, consistent with those completing at least 4 of 6 sessions. The study will be considered feasible if 60% or more of enrolled participants (≥36 of 60) achieve this end point. Acceptability, the secondary outcome, will be assessed using the AIM [76]. The mean and SD of AIM scores will be calculated, with an average score of ≥4.0 (on a 5-point scale) interpreted as evidence of acceptability. Exploratory latent growth curve modeling will be used to examine changes in depression and anxiety symptoms over time (measured by PHQ-9 and GAD-7). These analyses are considered exploratory due to the limited statistical power associated with the sample size (n=60).

XSL•FO
RenderX

For Aim 2, qualitative data from semi-structured interviews with a purposive subsample of 20 family caregivers (including both completers and noncompleters) will be analyzed using thematic content analysis in ATLAS.ti [96]. Interview transcripts will be transcribed verbatim and independently coded by trained research staff. An initial coding scheme will be developed inductively and applied iteratively. Two independent coders will review all transcripts, with discrepancies resolved through discussion until consensus is achieved [97]. Axial coding and thematic mapping will be used to identify higher-order categories related to eCARE's perceived strengths and limitations, usability, relevance, and reported facilitators or barriers to engagement. Analysis will continue until thematic saturation is reached, defined as the point at which no new codes or themes emerge [79,97,98].

Ethical Considerations

This study was reviewed and approved by the IRB at the University of Houston (Protocol# STUDY00003186) and adheres to the ethical standards outlined in the Declaration of Helsinki. All study procedures will be conducted in accordance with the approved protocol and applicable institutional requirements for human subjects' research.

Participants will provide informed consent prior to initiating any study procedures. Consent will be documented in writing at enrollment either via Qualtrics e-signature functionality or by signing a printed consent form when enrollment occurs in person. For the optional qualitative interview component, participants will complete an additional oral consent procedure prior to the start of the interview, and the study team will request a waiver of written documentation for that interview-specific consent, as applicable.

Participant privacy and confidentiality will be protected through multiple safeguards. Direct identifiers (eg, name, email, phone) will be collected solely for recruitment, scheduling, and compensation purposes and will be stored separately from research data, linked only via a study ID. App-based and survey data will be maintained on encrypted, password-protected platforms with access restricted to authorized study personnel; downloaded data will be stored on password-protected, encrypted university devices. While participants' interest in the qualitative component will be captured as part of the informed consent process, oral informed consent will be obtained again prior to the commencement of the qualitative interviews. Qualitative interviews will be audio and/or video-recorded only with participant permission; recordings will be transcribed and then deleted following transcription. Interview transcripts will be labeled with study IDs. Only approved members of the research team will have access to identifiable information, and the master linkage file will be maintained by the principal investigator and destroyed at study completion per protocol.

Participants will receive incentives in the form of electronic Amazon gift cards. To minimize the risk of undue inducement, gift cards will be distributed at three predetermined time points, including US $30 upon completion of a baseline consent and pretest assessment survey, US $30 at the midpoint assessment (T4) contingent on at least one completed check-in with eCARE, and US $30 upon completion of the posttest assessment (T7). Participants who complete the qualitative interview will receive an additional US $50 e-gift card.

Results

Participant recruitment and enrollment began in June 2024, with data collection expected to conclude in August 2025. Data analysis will commence in December 2025, coinciding with the PI's transition to a new institution, with preliminary findings anticipated by April-May 2026. Table 2 presents the baseline sociodemographic, clinical, and caregiving-related characteristics of enrolled participants, while Table 3 outlines the primary feasibility outcomes and secondary outcomes of the pilot study (Multimedia Appendix 1).

XSL•FO
RenderX

Table 2. Baseline characteristics of the participants in the feasibility study.

Category and variable	Measurement
Sociodemographic characteristics	
Age	Years
Sex	Male, female, or other
Gender	Self-reported
Race and ethnicity	US Census categories; self-reported
Education	Highest level completed
Relationship status	Married, partnered, single, etc
Employment	Full-time, part-time, unemployed, or not applicable
Insurance coverage	Private, public, other or uninsured
Income	Income brackets
Clinical characteristics of care recipient	
Age	Years
Patient sex	Male, female, or other
Cancer diagnosis	As reported by family caregivers
Cancer treatment	Chemotherapy, radiation, surgery, etc
Cancer stage	I-IV or not applicable
Caregiving characteristics	
Length of caregiving period	Months or years
Relationship with care-recipient	Categories based on self-report
Primary caregiver status	Yes or No
Cohabitation with care recipient	Yes or No
Caregiving intensity (hours/week)	Total hours per week
Type of tasks provided	Household, personal, practical, and emotional
Care recipient ADL[a] Independence (Katz Index)	Yes or No; score range 0-6

[a]ADL: activities of daily living.

Table 3. Primary and secondary outcomes of the feasibility study.

Outcome type and measure	Definition/Criteria	Success threshold
Primary outcome		
Feasibility end point (eCARE[a] completion rate)	• Number of eCARE completions during 8-week period • Proportion of participants achieving feasibility end point	• ≥4 completions (≥66.6%) = end point achieved • Study feasible if ≥60% of participants (≥36) reach this threshold
Secondary outcome		
Acceptability of Intervention Measure (AIM)	• Participant-reported acceptability of eCARE using AIM scale (range: 1 to 5)	• Average score ≥4 over 5

[a]eCARE: Ellipsis Caregiver Assessment Enhancement.

Discussion

Principal Findings

Family caregivers play a central role in cancer care, assuming extensive practical, emotional, and coordination responsibilities that contribute to substantial psychological, physical, and financial strain [1,2,11]. Across cancer types and illness stages, caregivers frequently experience clinically significant anxiety and depression, often at rates exceeding those observed in patients themselves, yet their mental health needs remain

XSL•FO
RenderX

insufficiently addressed [4,5,7-9,22]. This single-arm, prospective pilot study is among the first to evaluate the feasibility and acceptability of using artificial intelligence and speech-based analytics through a mobile app (eCARE) to monitor the severity of depression and anxiety among a sample of community-based family caregivers of individuals living with cancer. Traditional methods, such as clinic-based assessments or paper forms, often present challenges with high participant burden and low engagement, which can compromise data accuracy and limit opportunities for timely intervention. It is essential to explore innovative strategies offering more flexible, scalable, and caregiver-centered approaches to distress monitoring. To this end, we hypothesize that (1) eCARE will be feasible and acceptable, and (2) that participants will provide in-depth information related to barriers and facilitators influencing its uptake, in addition to feedback for future refinement and implementation.

Prior research has demonstrated that eHealth interventions for cancer caregivers can reduce depressive symptoms and modestly improve quality of life; however, existing studies are few, heterogeneous, and often limited by short follow-up periods and narrow outcome focus [6,12,34,35,39,99]. Moreover, most digital interventions rely on self-report questionnaires, which can introduce response burden and missing data. By requiring a brief speech sample to generate predictions of depression and anxiety, eCARE represents a minimally burdensome pathway for mental health monitoring. The mobile application integrates natural language processing to analyze spoken words and acoustic modeling to assess vocal features, leveraging an accessible, multimodal approach well-suited for smartphones [37,39,59,99,100]. As Schoenberg [101] highlighted, digital tools that are frictionless and embedded in daily life are more likely to be adopted and sustained in real-world environments.

Findings from this pilot study will provide preliminary data on key indicators of feasibility (eg, completion rates) and acceptability (eg, satisfaction ratings). This information will be instrumental in helping to determine whether a speech-based tool can be successfully integrated into the complex caregiving context, where time constraints, emotional burden, and competing demands often limit participation in mental health initiatives. To further this goal, the study also incorporates qualitative feedback, allowing for a user-informed refinement of the application's content, usability, and perceived value. While eCARE does not currently involve provider feedback loops, future iterations may benefit from exploring how clinicians can access and respond to data trends emerging from the app, potentially enhancing care coordination and timely referral to supportive services.

Should eCARE demonstrate feasibility and acceptability, these findings will directly inform the design of a larger, multisite randomized controlled trial. Such a trial would be designed to robustly compare eCARE with established screening methods, like the National Comprehensive Cancer Network (NCCN) Distress Thermometer or standard electronic symptom checklists, and to evaluate its efficacy on family caregivers' mental health.

Finally, since symptoms of depression and anxiety are among the most prominent predictors of quality of life and long-term health outcomes, effective and early identification is critical. If validated, speech-based assessments like eCARE could serve as a foundation for future interventions aimed at improving caregiver resilience, reducing emotional burden, and ultimately enhancing the caregiving experience.

Limitations

This study has several limitations that should be considered. First, as a single-arm feasibility and acceptability study, it is not designed to evaluate the efficacy of eCARE in reducing depression or anxiety symptoms, nor to support causal inferences. Second, generalizability may be limited, as participation requires access to a smartphone and a minimum level of comfort with mobile technology, potentially excluding family caregivers with lower digital literacy or limited access to digital resources. Third, self-selection bias may be present, as caregivers who choose to enroll may be experiencing higher distress, may be more motivated, or may be more receptive to digital mental health tools, which could inflate estimates of feasibility and acceptability. Finally, the relatively short monitoring period affects the ability to elaborate on sustained engagement and longer-term acceptability, in addition to barriers and facilitators influencing eCARE uptake over time.

Conclusions

This study will generate foundational evidence on the feasibility, acceptability, and perceived utility of a mobile, speech-based, AI-enabled tool for monitoring depression and anxiety among family caregivers of individuals living with cancer. Findings will inform the design of a fully powered clinical trial and guide the optimization of remote mental health screening approaches tailored to the caregiving experience. As a scalable, low-cost, and low-burden intervention, eCARE has the potential to support timely identification of psychological distress, enhance access to supportive care services, and improve coordination of psychosocial care across oncology settings. Beyond cancer caregiving, the underlying model holds promise for adaptation to other chronic or life-limiting conditions and to non–English-speaking populations, pending appropriate validation. Future studies should compare eCARE with established screening tools and examine long-term clinical outcomes, user engagement, and effects on both caregiver and patient well-being, while also evaluating strategies for integrating AI-derived screening results into clinical workflows without increasing provider burden. Dissemination of findings will include sharing results with academic, clinical, and community stakeholders to support continued refinement and implementation of digital mental health solutions for caregivers.

XSL•FO
RenderX

Acknowledgments

We would like to thank the caregivers who will participate in this study for sharing their time and experiences. We are also grateful to our clinical and community partners for their support in recruitment and outreach: our deepest thanks to the Kelsey Research Foundation, Family Caregiver Alliance, the Leukemia and Lymphoma Society, the Collaborative Approach for AANHPI Research and Education Registry, and Friend for Life Cancer Support Network.

The authors declare the use of generative artificial intelligence (GAI) in the writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: Proofreading and editing of the revised manuscript. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Data Availability

The datasets generated or analyzed during this study will be available from the corresponding author on reasonable request.

Funding

This study was supported by a SEED Grant from the University of Houston, awarded to CA. The content is solely the responsibility of the authors and does not necessarily represent the official views of the University of Houston.

Authors' Contributions

Conceptualization: CA, MA, ASA
Methodology: CA, TN, MA, ASA
Investigation: CA, MA, TN, AB, SA
Resources: MA, TN
Software: MA, TN
Project administration: CA
Writing – original draft: CA, AB, IKM
Writing – review & editing: All authors

Conflicts of Interest

MA and TN disclose a conflict of interest as employees and shareholders of Ellipsis Health. All other authors declare no conflicts of interest.

Multimedia Appendix 1
CONSORT checklist.
[PDF File (Adobe PDF File), 281 KB - resprot_v15i1e83276_app1.pdf]

Multimedia Appendix 2
Peer-review report from the University of Houston SEED Grant Review Committee.
[PDF File (Adobe PDF File), 176 KB - resprot_v15i1e83276_app2.pdf]

References

1. Chow PI, Showalter SL, Gerber MS, Kennedy E, Brenin DR, Schroen AT, et al. Use of mental health apps by breast cancer patients and their caregivers in the United States: protocol for a pilot pre-post study. JMIR Res Protoc 2019;8(1):e11452 [FREE Full text] [doi: 10.2196/11452] [Medline: 31344674]
2. Ochoa CY, Buchanan Lunsford N, Lee Smith J. Impact of informal cancer caregiving across the cancer experience: a systematic literature review of quality of life. Palliat Support Care 2020;18(2):220-240 [FREE Full text] [doi: 10.1017/S1478951519000622] [Medline: 31588882]
3. Girgis A, Lambert S. Cost of informal caregiving in cancer care. Cancer Forum 2017;41(2):16-22 [FREE Full text]
4. Taylor J, Fradgley E, Clinton-McHarg T, Byrnes E, Paul C. What are the sources of distress in a range of cancer caregivers? A qualitative study. Support Care Cancer 2021;29(5):2443-2453. [doi: 10.1007/s00520-020-05742-0] [Medline: 32929537]
5. Sklenarova H, Krümpelmann A, Haun MW, Friederich H, Huber J, Thomas M, et al. When do we need to care about the caregiver? Supportive care needs, anxiety, and depression among informal caregivers of patients with cancer and cancer survivors. Cancer 2015;121(9):1513-1519 [FREE Full text] [doi: 10.1002/cncr.29223] [Medline: 25677095]
6. Marzorati C, Renzi C, Russell-Edu SW, Pravettoni G. Telemedicine use among caregivers of cancer patients: systematic review. J Med Internet Res 2018;20(6):e223 [FREE Full text] [doi: 10.2196/jmir.9812] [Medline: 29914858]

XSL•FO
RenderX

7. Grunfeld E, Coyle D, Whelan T, Clinch J, Reyno L, Earle CC, et al. Family caregiver burden: results of a longitudinal study of breast cancer patients and their principal caregivers. CMAJ 2004;170(12):1795-1801 [FREE Full text] [doi: 10.1503/cmaj.1031205] [Medline: 15184333]
8. Nipp RD, El-Jawahri A, Fishbein JN, Gallagher ER, Stagl JM, Park ER, et al. Factors associated with depression and anxiety symptoms in family caregivers of patients with incurable cancer. Ann Oncol 2016;27(8):1607-1612 [FREE Full text] [doi: 10.1093/annonc/mdw205] [Medline: 27177859]
9. Pan Y, Lin Y. Systematic review and meta-analysis of prevalence of depression among caregivers of cancer patients. Front Psychiatry 2022;13:817936 [FREE Full text] [doi: 10.3389/fpsyt.2022.817936] [Medline: 35633789]
10. Cochrane A, Reid O, Woods S, Gallagher P, Dunne S. Variables associated with distress amongst informal caregivers of people with lung cancer: a systematic review of the literature. Psychooncology 2021;30(8):1246-1261. [doi: 10.1002/pon.5694] [Medline: 33945184]
11. Lambert SD, Harrison JD, Smith E, Bonevski B, Carey M, Lawsin C, et al. The unmet needs of partners and caregivers of adults diagnosed with cancer: a systematic review. BMJ Support Palliat Care 2012;2(3):224-230. [doi: 10.1136/bmjspcare-2012-000226] [Medline: 24654195]
12. Cheng Q, Xu B, Ng MS, Duan Y, So WK. Effectiveness of psychoeducational interventions among caregivers of patients with cancer: a systematic review and meta-analysis. Int J Nurs Stud 2022;127:104162. [doi: 10.1016/j.ijnurstu.2021.104162] [Medline: 35121521]
13. Dave R, Friedman S, Miller-Sonet E, Moore T, Peterson E, Fawzy Doran J, et al. Identifying and addressing the needs of caregivers of patients with cancer: evidence on interventions and the role of patient advocacy groups. Future Oncol 2024;20(33):2589-2602 [FREE Full text] [doi: 10.1080/14796694.2024.2387526] [Medline: 39329173]
14. Bergerot CD, Bergerot PG, Philip EJ, Ferrari R, Peixoto RM, Crane TE, et al. Enhancing cancer supportive care: integrating psychosocial support, nutrition, and physical activity using telehealth solutions. JCO Glob Oncol 2024;10:e2400333. [doi: 10.1200/GO-24-00333] [Medline: 39509661]
15. Oechsle K, Ullrich A, Marx G, Benze G, Heine J, Dickel L, et al. Psychological burden in family caregivers of patients with advanced cancer at initiation of specialist inpatient palliative care. BMC Palliat Care 2019;18(1):102 [FREE Full text] [doi: 10.1186/s12904-019-0469-7] [Medline: 31739802]
16. Fujinami R, Sun V, Zachariah F, Uman G, Grant M, Ferrell B. Family caregivers' distress levels related to quality of life, burden, and preparedness. Psychooncology 2015;24(1):54-62 [FREE Full text] [doi: 10.1002/pon.3562] [Medline: 24789500]
17. Halkett GKB, Lobb EA, Shaw T, Sinclair MM, Miller L, Hovey E, et al. Distress and psychological morbidity do not reduce over time in carers of patients with high-grade glioma. Support Care Cancer 2017;25(3):887-893. [doi: 10.1007/s00520-016-3478-6] [Medline: 27841006]
18. Shaffer KM, Jacobs JM, Nipp RD, Carr A, Jackson VA, Park ER, et al. Mental and physical health correlates among family caregivers of patients with newly-diagnosed incurable cancer: a hierarchical linear regression analysis. Support Care Cancer 2017;25(3):965-971 [FREE Full text] [doi: 10.1007/s00520-016-3488-4] [Medline: 27866337]
19. El-Jawahri A, Greer JA, Park ER, Jackson VA, Kamdar M, Rinaldi SP, et al. Psychological distress in bereaved caregivers of patients with advanced cancer. J Pain Symptom Manage 2021;61(3):488-494 [FREE Full text] [doi: 10.1016/j.jpainsymman.2020.08.028] [Medline: 32882355]
20. Kim Y, Carver CS, Spiegel D, Mitchell H, Cannady RS. Role of family caregivers' self-perceived preparedness for the death of the cancer patient in long-term adjustment to bereavement. Psychooncology 2017;26(4):484-492. [doi: 10.1002/pon.4042] [Medline: 26661137]
21. Fasse L, Flahault C, Brédart A, Dolbeault S, Sultan S. Describing and understanding depression in spouses of cancer patients in palliative phase. Psychooncology 2015;24(9):1131-1137. [doi: 10.1002/pon.3777] [Medline: 25707938]
22. Geng H, Chuang D, Yang F, Yang Y, Liu W, Liu L, et al. Prevalence and determinants of depression in caregivers of cancer patients: a systematic review and meta-analysis. Medicine (Baltimore) 2018;97(39):e11863 [FREE Full text] [doi: 10.1097/MD.0000000000011863] [Medline: 30278483]
23. Trevino KM, Prigerson HG, Maciejewski PK. Advanced cancer caregiving as a risk for major depressive episodes and generalized anxiety disorder. Psychooncology 2018;27(1):243-249 [FREE Full text] [doi: 10.1002/pon.4441] [Medline: 28426918]
24. Alam S, Hannon B, Zimmermann C. Palliative care for family caregivers. J Clin Oncol 2020;38(9):926-936. [doi: 10.1200/JCO.19.00018] [Medline: 32023152]
25. Bedaso A, Dejenu G, Duko B. Depression among caregivers of cancer patients: updated systematic review and meta-analysis. Psychooncology 2022;31(11):1809-1820 [FREE Full text] [doi: 10.1002/pon.6045] [Medline: 36209385]
26. Molassiotis A, Wang M. Understanding and supporting informal cancer caregivers. Curr Treat Options Oncol 2022;23(4):494-513 [FREE Full text] [doi: 10.1007/s11864-022-00955-3] [Medline: 35286571]
27. Benyo S, Phan C, Goyal N. Health and well-being needs among head and neck cancer caregivers - a systematic review. Ann Otol Rhinol Laryngol 2023;132(4):449-459 [FREE Full text] [doi: 10.1177/00034894221088180] [Medline: 35549916]
28. Wang T, Molassiotis A, Chung BPM, Tan J. Unmet care needs of advanced cancer patients and their informal caregivers: a systematic review. BMC Palliat Care 2018;17(1):96 [FREE Full text] [doi: 10.1186/s12904-018-0346-9] [Medline: 30037346]

XSL•FO
RenderX

29. Christophe V, Anota A, Vanlemmens L, Cortot A, Ceban T, Piessen G, et al. Unmet supportive care needs of caregivers according to medical settings of cancer patients: a cross-sectional study. Support Care Cancer 2022;30(11):9411-9419. [doi: 10.1007/s00520-022-07379-7] [Medline: 36205779]
30. Wang AW, Kim Y, Ting A, Lam WWT, Lambert SD. Healthcare professionals' perspectives on the unmet needs of cancer patients and family caregivers: global psycho-oncology investigation. Support Care Cancer 2022;31(1):36 [FREE Full text] [doi: 10.1007/s00520-022-07478-5] [Medline: 36520223]
31. Hart NH, Crawford-Williams F, Crichton M, Yee J, Smith TJ, Koczwara B, et al. Unmet supportive care needs of people with advanced cancer and their caregivers: a systematic scoping review. Crit Rev Oncol Hematol 2022;176:103728. [doi: 10.1016/j.critrevonc.2022.103728] [Medline: 35662585]
32. Ullrich A, Marx G, Bergelt C, Benze G, Zhang Y, Wowretzko F, et al. Supportive care needs and service use during palliative care in family caregivers of patients with advanced cancer: a prospective longitudinal study. Support Care Cancer 2021;29(3):1303-1315 [FREE Full text] [doi: 10.1007/s00520-020-05565-z] [Medline: 32632761]
33. Zhong C, Luo X, Tan M, Chi J, Guo B, Tang J, et al. Digital health interventions to improve mental health in patients with cancer: umbrella review. J Med Internet Res 2025;27:e69621 [FREE Full text] [doi: 10.2196/69621] [Medline: 39984165]
34. Su Z, Li X, McDonnell D, Fernandez AA, Flores BE, Wang J. Technology-based interventions for cancer caregivers: concept analysis. JMIR Cancer 2021;7(4):e22140 [FREE Full text] [doi: 10.2196/22140] [Medline: 34783664]
35. Li Y, Li J, Zhang Y, Ding Y, Hu X. The effectiveness of e-Health interventions on caregiver burden, depression, and quality of life in informal caregivers of patients with cancer: a systematic review and meta-analysis of randomized controlled trials. Int J Nurs Stud 2022;127:104179. [doi: 10.1016/j.ijnurstu.2022.104179] [Medline: 35124473]
36. Zhang A, Kamat A, Acquati C, Aratow M, Kim JS, DuVall AS, et al. Evaluating the feasibility and acceptability of an artificial-intelligence-enabled and speech-based distress screening mobile app for adolescents and young adults diagnosed with cancer: a study protocol. Cancers (Basel) 2022;14(4):104179 [FREE Full text] [doi: 10.3390/cancers14040914] [Medline: 35205663]
37. Siegler AJ, Knox J, Bauermeister JA, Golinkoff J, Hightow-Weidman L, Scott H. Mobile app development in health research: pitfalls and solutions. Mhealth 2021;7:32 [FREE Full text] [doi: 10.21037/mhealth-19-263] [Medline: 33898601]
38. Klagholz SD, Ross A, Wehrlen L, Bedoya SZ, Wiener L, Bevans MF. Assessing the feasibility of an electronic patient-reported outcome (ePRO) collection system in caregivers of cancer patients. Psychooncology 2018;27(4):1350-1352 [FREE Full text] [doi: 10.1002/pon.4658] [Medline: 29405538]
39. Shin JY, Kang TI, Noll RB, Choi SW. Supporting caregivers of patients with cancer: a summary of technology-mediated interventions and future directions. Am Soc Clin Oncol Educ Book 2018;38:838-849 [FREE Full text] [doi: 10.1200/EDBK_201397] [Medline: 30231412]
40. Darley A, Coughlan B, Furlong E. People with cancer and their family caregivers' personal experience of using supportive eHealth technology: a narrative review. Eur J Oncol Nurs 2021;54:102030 [FREE Full text] [doi: 10.1016/j.ejon.2021.102030] [Medline: 34531122]
41. Heynsbergh N, Heckel L, Botti M, Livingston PM. Feasibility, useability and acceptability of technology-based interventions for informal cancer carers: a systematic review. BMC Cancer 2018;18(1):244 [FREE Full text] [doi: 10.1186/s12885-018-4160-9] [Medline: 29499663]
42. Lazenby M, Ercolano E, Grant M, Holland JC, Jacobsen PB, McCorkle R. Supporting commission on cancer-mandated psychosocial distress screening with implementation strategies. J Oncol Pract 2015;11(3):e413-e420 [FREE Full text] [doi: 10.1200/JOP.2014.002816] [Medline: 25758447]
43. Corveleyn A, Fann JR, Gray TF, Samolovitch S, Vanderlan J. Addressing mental health in cancer care: optimizing interdisciplinary psychosocial support. J Natl Compr Canc Netw 2025;23(Supplement):e255002. [doi: 10.6004/jnccn.2025.5002]
44. Smith SK, Loscalzo M, Mayer C, Rosenstein DL. Best practices in oncology distress management: beyond the screen. Am Soc Clin Oncol Educ Book 2018;38:813-821 [FREE Full text] [doi: 10.1200/EDBK_201307] [Medline: 30231391]
45. 2020 standards and resources. American College of Surgeons. 2023. URL: https://www.facs.org/quality-programs/cancer-programs/commission-on-cancer/standards-and-resources/2020/ [accessed 2026-01-16]
46. Mundt JC, Snyder PJ, Cannizzaro MS, Chappie K, Geralts DS. Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. J Neurolinguistics 2007;20(1):50-64 [FREE Full text] [doi: 10.1016/j.jneuroling.2006.04.001] [Medline: 21253440]
47. Hashim NW, Wilkes M, Salomon R, Meggs J, France DJ. Evaluation of voice acoustics as predictors of clinical depression scores. J Voice 2017;31(2):256.e1-256.e6. [doi: 10.1016/j.jvoice.2016.06.006] [Medline: 27473933]
48. Galili L, Amir O, Gilboa-Schechtman E. Acoustic properties of dominance and request utterances in social anxiety. J Soc Clin Psychol 2013;32(6):651-673. [doi: 10.1521/jscp.2013.32.6.651]
49. Teferra BG, Rueda A, Pang H, Valenzano R, Samavi R, Krishnan S, et al. Screening for depression using natural language processing: literature review. Interact J Med Res 2024;13:e55067 [FREE Full text] [doi: 10.2196/55067] [Medline: 39496145]
50. Faurholt-Jepsen M, Busk J, Frost M, Vinberg M, Christensen EM, Winther O, et al. Voice analysis as an objective state marker in bipolar disorder. Transl Psychiatry 2016;6(7):e856 [FREE Full text] [doi: 10.1038/tp.2016.123] [Medline: 27434490]

XSL•FO
RenderX

51. Argolo F, Magnavita G, Mota NB, Ziebold C, Mabunda D, Pan PM, et al. Lowering costs for large-scale screening in psychosis: a systematic review and meta-analysis of performance and value of information for speech-based psychiatric evaluation. Braz J Psychiatry 2020;42(6):673-686 [FREE Full text] [doi: 10.1590/1516-4446-2019-0722] [Medline: 32321060]
52. Trevino AC, Quatieri TF, Malyska N. Phonologically-based biomarkers for major depressive disorder. EURASIP J Adv Signal Process 2011;2011(1):42. [doi: 10.1186/1687-6180-2011-42]
53. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri T. A review of depression and suicide risk assessment using speech analysis. Speech Commun 2015;71:10-49 [FREE Full text] [doi: 10.1016/j.specom.2015.03.004]
54. Muaremi A, Gravenhorst F, Grünerbl A, Arnrich B, Tröster G. Assessing bipolar episodes using speech cues derived from phone calls. 2014 Presented at: 4th International Symposium, MindCare 2014; May 8-9; Tokyo, Japan p. 103-114. [doi: 10.1007/978-3-319-11564-1_11]
55. Low DM, Bentley KH, Ghosh SS. Automated assessment of psychiatric disorders using speech: a systematic review. Laryngoscope Investig Otolaryngol 2020;5(1):96-116 [FREE Full text] [doi: 10.1002/lio2.354] [Medline: 32128436]
56. Espinola CW, Gomes JC, Pereira JMS, dos Santos WP. Detection of major depressive disorder using vocal acoustic analysis and machine learning. Res Biomed Eng 2020;37:53-64. [doi: 10.1101/2020.06.23.20138651]
57. Rutowski T, Shriberg E, Harati A, Lu Y, Chlebek P, Oliveira R. Depression and anxiety prediction using deep language models and transfer learning. : IEEE; 2020 Presented at: 7th International Conference on Behavioural and Social Computing (BESC); November 05-07; Bournemouth, United Kingdom p. 1-6. [doi: 10.1109/besc51023.2020.9348290]
58. Harati A, Shriberg E, Rutowski T, Chlebek P, Lu Y, Oliveira R. Speech-based depression prediction using encoder-weight-only transfer learning and a large corpus. : IEEE; 2021 Presented at: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); June 06-11; Toronto, ON p. 7273-7277. [doi: 10.1109/icassp39728.2021.9414208]
59. Ravi V, Wang J, Flint J, Alwan A. Enhancing accuracy and privacy in speech-based depression detection through speaker disentanglement. Comput Speech Lang 2024;86:101605 [FREE Full text] [doi: 10.1016/j.csl.2023.101605] [Medline: 38313320]
60. Rutowski T, Harati A, Lu Y, Shriberg E. Optimizing speech-input length for speaker-independent depression classification. Proc Interspeech 2019:3023-3027. [doi: 10.21437/interspeech.2019-3095]
61. Chlebek P, Shriberg E, Lu Y, Rutowski T, Harati A, Oliveira R. Comparing speech recognition services for HCI applications in behavioral health. 2020 Presented at: UbiComp/ISWC '20: 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and 2020 ACM International Symposium on Wearable Computers; 2020 September 12 - 17; Virtual Event Mexico p. 483-487. [doi: 10.1145/3410530.3414372]
62. Karlin B, Henry D, Anderson R, Cieri S, Aratow M, Shriberg E, et al. Digital phenotyping for detecting depression severity in a large payor-provider system: retrospective study of speech and language model performance. JMIR AI 2025;4:e69149 [FREE Full text] [doi: 10.2196/69149] [Medline: 40605836]
63. Zhao Z, Bao Z, Zhang Z, Deng J, Cummins N, Wang H, et al. Automatic assessment of depression from speech via a hierarchical attention transfer network and attention autoencoders. IEEE J Sel Top Signal Process 2020;14(2):423-434. [doi: 10.1109/jstsp.2019.2955012]
64. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
65. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
66. Moriarty DG, Zack MM, Kobau R. The centers for disease control and prevention's healthy days measures - population tracking of perceived physical and mental health over time. Health Qual Life Outcomes 2003;1:37 [FREE Full text] [doi: 10.1186/1477-7525-1-37] [Medline: 14498988]
67. Bédard M, Molloy D, Squire L, Dubois S, Lever J, O'Donnell M. The zarit burden interview: a new short version and screening version. Gerontologist 2001;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
68. Aron A, Aron EN, Smollan D. Inclusion of other in the self scale and the structure of interpersonal closeness. J Pers Soc Psychol 1992;63(4):596-612. [doi: 10.1037/0022-3514.63.4.596]
69. Helgeson VS, Van Vleet M. Short Report: Inclusion of other in the self scale: an adaptation and exploration in a diverse community sample. J Soc Pers Relat 2019;36(11-12):4048-4056 [FREE Full text] [doi: 10.1177/0265407519848491] [Medline: 34079152]
70. Lepore SJ, Silver RC, Wortman CB, Wayment HA. Social constraints, intrusive thoughts, and depressive symptoms among bereaved mothers. J Pers Soc Psychol 1996;70(2):271-282. [doi: 10.1037//0022-3514.70.2.271] [Medline: 8636882]
71. Lepore S, Ituarte P. Optimism about cancer enhances mood by reducing negative social relations. Cancer Research Therapy and Control 1999;8:165-174.
72. Reis HT, Crasta D, Rogge RD, Maniaci MR, Carmichael CL. Perceived partner responsiveness scale (PPRS). In: Worthington DL, Bodie GD, editors. The Sourcebook of Listening Research. New Jersey: Wiley; 2017.
73. Reis HT, Lee KY, O'Keefe SD, Clark MS. Perceived partner responsiveness promotes intellectual humility. Journal of Experimental Social Psychology 2018;79:21-33. [doi: 10.1016/j.jesp.2018.05.006]

XSL•FO
RenderX

74. Lemay EP, Neal AM. The wishful memory of interpersonal responsiveness. J Pers Soc Psychol 2013;104(4):653-672. [doi: 10.1037/a0030422] [Medline: 23088228]
75. Le BM, Impett EA, Lemay EP, Muise A, Tskhay KO. Communal motivation and well-being in interpersonal relationships: an integrative review and meta-analysis. Psychol Bull 2018;144(1):1-25. [doi: 10.1037/bul0000133] [Medline: 29154556]
76. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
77. O'Brien H, Cairns P, Hall M. A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. Int J Hum Comput Stud 2018;112:28-39 [FREE Full text] [doi: 10.1016/j.ijhcs.2018.01.004]
78. Kerr C, Nixon A, Wild D. Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. Expert Rev Pharmacoecon Outcomes Res 2010;10(3):269-281. [doi: 10.1586/erp.10.30] [Medline: 20545592]
79. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qual Res Sport Exerc Health 2019;13(2):201-216. [doi: 10.1080/2159676x.2019.1704846]
80. Guest G, Bunce A, Johnson L. How many interviews are enough?: an experiment with data saturation and variability. Field Methods 2006;18(1):59-82 [FREE Full text] [doi: 10.1177/1525822X05279903]
81. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017;27(4):591-608 [FREE Full text] [doi: 10.1177/1049732316665344] [Medline: 27670770]
82. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020;15(5):e0232076 [FREE Full text] [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]
83. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
84. Katz S. Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living. J Am Geriatr Soc 1983;31(12):721-727. [doi: 10.1111/j.1532-5415.1983.tb03391.x] [Medline: 6418786]
85. Levis B, Benedetti A, Thombs BD, DEPRESsion Screening Data (DEPRESSD) Collaboration. Accuracy of patient health questionnaire-9 (PHQ-9) for screening to detect major depression: individual participant data meta-analysis. BMJ 2019;365:l1476 [FREE Full text] [doi: 10.1136/bmj.l1476] [Medline: 30967483]
86. Löwe B, Decker O, Müller S, Brähler E, Schellberg D, Herzog W, et al. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. Med Care 2008;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
87. Rutter LA, Brown TA. Psychometric properties of the generalized anxiety disorder scale-7 (GAD-7) in outpatients with anxiety and mood disorders. J Psychopathol Behav Assess 2017;39(1):140-146 [FREE Full text] [doi: 10.1007/s10862-016-9571-9] [Medline: 28260835]
88. Lee B, Kim YE. The psychometric properties of the generalized anxiety disorder scale (GAD-7) among Korean university students. Psychiatry Clin Psychopharmacol 2019;29(4):864-871. [doi: 10.1080/24750573.2019.1691320]
89. Johnson SU, Ulvenes PG, Øktedalen T, Hoffart A. Psychometric properties of the general anxiety disorder 7-item (GAD-7) scale in a heterogeneous psychiatric sample. Front Psychol 2019;10:1713 [FREE Full text] [doi: 10.3389/fpsyg.2019.01713] [Medline: 31447721]
90. Aron A, McLaughlin-Volpe T, Mashek D, Lewandowski G, Wright S, Aron E. Including others in the self. Eur Rev Soc Psychol 2004;15(1):101-132. [doi: 10.1080/10463280440000008]
91. Aron A, Aron EN, Tudor M, Nelson G. Close relationships as including other in the self. J Pers Soc Psychol 1991;60(2):241-253. [doi: 10.1037//0022-3514.60.2.241]
92. Aron A, Lewandowski JG, Mashek D, Aron E. The self-expansion model of motivation and cognition in close relationships. In: Simpson JA, Campbell L, editors. The Oxford Handbook of Close Relationships, Oxford Library of Psychology. England: Oxford University Press; 2013:90-115.
93. Saita E, Acquati C, Kayser K. Coping with early stage breast cancer: examining the influence of personality traits and interpersonal closeness. Front Psychol 2015;6:88 [FREE Full text] [doi: 10.3389/fpsyg.2015.00088] [Medline: 25699003]
94. Mashek D, Cannaday LW, Tangney JP. Inclusion of community in self scale: a single‐item pictorial measure of community connectedness. J Community Psychol 2007;35(2):257-275. [doi: 10.1002/jcop.20146]
95. Dumas SE, Dongchung TY, Sanderson ML, Bartley K, Levanon Seligson A. A comparison of the four healthy days measures (HRQOL-4) with a single measure of self-rated general health in a population-based health survey in New York City. Health Qual Life Outcomes 2020;18(1):315 [FREE Full text] [doi: 10.1186/s12955-020-01560-4] [Medline: 32972402]
96. ATLAS.ti (Version 23). ATLAS.ti. Berlin, Germany; 2023. URL: https://atlasti.com/ [accessed 2026-01-16]
97. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
98. Boddy CR. Sample size for qualitative research. Qual Mark Res 2016;19(4):426-432. [doi: 10.1108/qmr-06-2016-0053]
99. Malgaroli M, Hull TD, Zech JM, Althoff T. Natural language processing for mental health interventions: a systematic review and research framework. Transl Psychiatry 2023;13(1):309 [FREE Full text] [doi: 10.1038/s41398-023-02592-2] [Medline: 37798296]

XSL•FO
RenderX

100. Flanagan O, Chan A, Roop P, Sundram F. Using acoustic speech patterns from smartphones to investigate mood disorders: scoping review. JMIR Mhealth Uhealth 2021;9(9):e24352 [FREE Full text] [doi: 10.2196/24352] [Medline: 34533465]
101. Schoenberg R. Digital health companions: where caregivers can't go. NEJM Catal Innov Care Deliv 2025;6(6). [doi: 10.1056/cat.24.0419]

Abbreviations

AI: artificial intelligence
AIM: acceptability of intervention measure
AANHPI: Asian American, Native Hawaiian, and Pacific Islander
CARE: Collaborative Approach for AANHPI Research and Education
CDC: Centers for Disease Control and Prevention
D&A: Depression and Anxiety
eCARE: Ellipsis Caregiver Assessment Enhancement
GAD-7: Generalized Anxiety Disorder–7
HIPAA: Health Insurance Portability and Accountability Act
HRQOL-4: Health-Related Quality of Life – 4 item Healthy Days Measure
IOS: Inclusion of the Other in the Self Scale
IRB: Institutional Review Board
NAMI: National Alliance on Mental Illness
NCCN: National Comprehensive Cancer Network
PHQ-9: Patient Health Questionnaire–9
PPRS: Perceived Partner Responsiveness Scale
SCS: Social Constraints Scale
UES-SF: User Engagement Scale–Short Form
ZBI-12: Zarit Burden Interview–Short Form (12 items)

Edited by A Schwartz; submitted 30.Aug.2025; this is a non–peer-reviewed article;accepted 31.Dec.2025; published 12.Feb.2026.

Please cite as:
Acquati C, Aratow M, Nazreen T, Bhattacharjee A, Marra IK, Alexander AS
Evaluating an AI-Enabled Mobile Mental Health Monitoring Tool Among Family Caregivers of Adults Living With Cancer: Single-Arm Feasibility and Acceptability Trial Protocol
JMIR Res Protoc 2026;15:e83276
URL: https://www.researchprotocols.org/2026/1/e83276
doi:10.2196/83276
PMID:

©Chiara Acquati, Michael Aratow, Tahmida Nazreen, Arunima Bhattacharjee, Isabella K Marra, Ashley S Alexander. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effectiveness and Cost-Effectiveness of Emergency Department–Based Violence Intervention Programs in the United Kingdom: Protocol for a Quasi-Experimental Study

Simon Moore[1,2], PhD; Sinead Brophy[3], PhD; Amrita Bandyopadhyay[3], PhD; Annemarie Newbury[1], MSc; Megan Hamilton[1], MSc; Adele Battaglia[1], MA Ed; Trudy Lowe[2], MSc; David O'Reilly[4], MD; David Rawlinson[5], MRes; Lara Snowdon[6], MSc; Jonathan Shepherd[1,2], DDSc; Vaseekaran Sivarajasingam[1], PhD; Alan Watkins[7], PhD; Simon Walker[8], MSc; Shainur Premji[8], PhD; Sophie Borgia[1], MSc; Henry Yeomans[1], MSc

[1]Violence Research Group, School of Dentistry, Cardiff University, Cardiff, United Kingdom

[2]Security, Crime and Intelligence Innovation Institute, Cardiff University, Cardiff, Wales, United Kingdom

[3]Health Data Science, Swansea University, Swansea, United Kingdom

[4]Cardiff Liver Unit, School of Medicine, University Hospital of Wales, Cardiff, United Kingdom

[5]Emergency Medical Retrieval & Transfer Service, Wales Air Ambulance, Tŷ Elusen, Ffordd Angel, Llanelli Gate, Dafen, United Kingdom

[6]Public Health Wales, Cardiff, United Kingdom

[7]Medical School, Swansea University, Swansea, United Kingdom

[8]Centre for Health Economics, University of York, York, United Kingdom

Corresponding Author:
Simon Moore, PhD
Violence Research Group
School of Dentistry
Cardiff University
Heath Park
Cardiff, CF14 4XY
United Kingdom
Phone: 1 44 2920874000
Email: mooresc2@cardiff.ac.uk

Abstract

Background: Hospital-Based Violence Intervention Programs (HVIPs), based in Emergency Departments (EDs), have been proposed as a public health response to violence. These programs address the underlying reasons why patients are exposed to violence. In addressing any underlying modifiable risks and vulnerabilities HVIPs can reduce patients' exposure to violence and therefore subsequent unplanned attendance into ED.

Objective: The objectives of this study are to (1) assess whether patient involvement with a HVIP reduces the likelihood of unscheduled ED reattendance, (2) determine whether the presence of the HVIP improves ascertainment of violence in ED attendances, and (3) derive the costs of the HVIP and compare those to the benefits of the intervention and understand whether the HVIP represents value for money from a health service perspective. If an effect is observed, then models will estimate the health impacts, costs and potential savings over a longer time (eg, 10 years) period and for a national roll-out.

Methods: ED patients are eligible for inclusion in the evaluation if they are normally resident in Wales, United Kingdom, aged 11 years and older. A controlled longitudinal natural experiment will be undertaken. The primary outcome is derived from the Emergency Department Dataset, routinely collected for all EDs in Wales, and is subsequent unplanned ED attendance. Case patients will be matched to control patients attending EDs without an HVIP. Analysis will derive the hazard rate for subsequent unplanned ED attendances using recurrent event analysis. The total monthly count of patients identified as attending because of violence in intervention EDs will be compared to the total count of Welsh control EDs in an interrupted time-series analysis to determine whether HVIPS increase violence ascertainment. To determine whether referral, versus no referral, to the HVIP represents value for money, we will undertake a cost-effectiveness analysis from the perspective of the National Health Service. The approval to access and analyze data housed in the Secure Anonymized Information Linkage (SAIL) databank, an ISO (International Organization for Standardization) 27001 certified and UK Statistics Authority accredited secure data environment,

XSL•FO
RenderX

was granted by the SAIL independent Information Governance Review Panel (Ref: 1421). Findings will be presented at local, national, and international conferences and disseminated by peer-reviewed publication.

Results: Design inputs arising from public patient involvement and engagement (PPIE) are reported. As a protocol, no further results are available.

Conclusions: Novel methods are developed to provide the first robust evaluation of Emergency Department Violence Intervention Programs (EDVIPs).

Trial Registration: ISRCTN Registry ISRCTN68945844; https://www.isrctn.com/ISRCTN68945844?q=68945844&filters=&sort=&offset=1&totalResults=1&page=1&pageSize=10

International Registered Report Identifier (IRRID): PRR1-10.2196/86247

(JMIR Res Protoc 2026;15:e86247) doi:10.2196/86247

KEYWORDS

emergency department; assault; violence; injury; prevention; treatment

Introduction

Background

Those who experience serious injury due to violence are likely to attend an Emergency Department (ED). EDs are therefore ideal locations for Hospital-Based Violence Intervention Programs (HVIPs) [1-3]. HVIPs have recently emerged as a public health response to violent victimization [3,4], but despite interest in HVIPs, there has been no rigorous evaluation of this public health approach to violence in the United Kingdom. Moreover, little is known about the effectiveness of patient discharge planning and referral from ED into organizations able to support children involved in violence [5] and patients exposed to domestic violence [6,7], and there is a paucity of studies considering referrals for young men involved in violence, the most dominant population in respect of Assault-Related Attendance (ARA) [8]. There are even fewer studies of the effective support available to victims of sexual violence attending ED [9]. Despite uncertainty over effectiveness in a UK context, HVIPs have begun to be implemented. For example, the Scottish Violence Reduction Unit (VRU) has placed navigators in EDs, typically youth workers who connect with patients aged 25 years and younger [10].

EDs work with patients attending for many reasons, including those presenting with violence-related injuries. EDs also engage in broader violence reduction initiatives, clinical staff typically receive training in adult and pediatric safeguarding, and many EDs have provisions to identify and refer children and victims of domestic and sexual violence. However, methods of ascertainment and referral vary considerably, and formal relationships with the police and other partners can often lack continuity, with multiagency approaches not capturing the entirety of patients' journeys. A public health approach involving the ED would be beneficial in improving overall population health outcomes. It is particularly timely; in 2019 there were 175,764 ARAs at EDs in England and Wales [11] and, despite the pandemic, 119,111 ARAs in 2020 [12]. Knife crime, and therefore serious trauma, has risen by 71% from 2014 to 2018 [13]. Up to 75% of ARAs are unknown to the police [14-16]; therefore, EDs hold exclusive data on assault characteristics, patient vulnerabilities, and modifiable risks and are therefore well situated to play a significant role in the identification of violence, to investigate the circumstances of violence, and to challenge any underlying vulnerabilities or modifiable risks exposing patients to violence, whether that is through direct support, referral, or discharge planning.

The need for HVIPs is aligned to broader UK Government initiatives, which aim to promote a whole system multiagency (WSMA) [17] approach to violence. The 1998 Crime and Disorder Act requires the police, local government, and the National Health Service (NHS) to collaborate on joint crime reduction strategies, and this includes data sharing to inform targeted responses. Violence reduction is further prioritized by the UK Government in its serious violence strategy [18], and the UK Government has allocated funds for the formation of VRUs in England and a violence prevention unit (VPU) in Wales, across 18 Police and Crime Commissioner jurisdictions, with the explicit purpose of promoting the WSMA approach [19]. These initiatives are further aligned with a move toward active population health management, digitally enabled whole-person care, and evidence-based treatment pathways outlined in the NHS future plan [20]. Integrated care systems in NHS England will be expected to specify violence prevention and reduction standards, which are incorporated into the 2021/22 NHS Standard Contract, and there are expectations that hubs will form Violence Prevention Teams (VPTs) similar to the police VRUs and VPU. Furthermore, a public sector duty on partnerships encouraging the prioritization of reducing serious violence has received royal assent as a part of the police, crime, sentencing, and courts bill. This legislation includes a serious violence duty, placing a statutory obligation on organizations to collaborate, communicate, and act.

The overarching aim of the work proposed here is a robust effectiveness and cost-effectiveness evaluation of ED-based VPTs. VPTs represent a formal collaboration between police and health care and embody the WSMA approach. To our knowledge, this is the first formal evaluation of a nurse-led, ED-based HVIP in the United Kingdom and will address significant gaps in current understanding of their effectiveness and thereby facilitate future aspirations for evidence-based referral pathways and discharge planning [20,21]. VPTs main function is to identify and support patients attending ED with assault-related injury. To facilitate this, they engage in broader pedagogical roles, increasing awareness of these patients' needs,

XSL•FO
RenderX

modifiable risks, and opportunities to identify and refer across the ED clinical environment.

Theoretical Framework

The theoretical motivation for a WSMA approach to HVIPs is that there are many modifiable risks and vulnerabilities that, in combination, determine an individual's exposure to violence and subsequently an ARA in an ED. Epidemiologically, these can be usefully described by shared circumstances that in turn signpost opportunities to modify risk or support patients' vulnerability, but responsibility can fall across organizations, including local government, health care, and criminal justice. Risks include the consumption of alcohol and other psychoactive substances [22-25]; criminal and/or sexual exploitation and homelessness [3,26]. Violence tends to be more prevalent in younger, socially disadvantaged groups [27-33], with male, socioeconomically deprived individuals being more likely to endure violence and experience assault-related injury. These characteristics further extend to personality features [34], including mental health status and learning disability, and neurodevelopmental disorders [35]. This complex interplay of factors that promote exposure to violence, and hence lead to an ARA, highlight the need for a WSMA approach. For example, an environment might become synonymous with violence through a process of homophile [36], whereby individuals with shared pursuits who are at risk of violence gather, for example, street drinkers and late-night drinking environments. Mitigation might include challenging reasons for frequenting such an environment, including alcohol and substance misuse counseling. Some environments might involve those who use violence to advance their interests, such as acquisitive crime, sexual assault, or sexual exploitation, in which case criminal justice or safeguarding processes to deter violence might be involved, along with support to victims. Chaotic or otherwise disadvantaged households in which domestic violence or harm to children arises might best be approached from a multiagency process such as the Multi-Agency Risk Assessment Committee (MARAC) and formal investigation (Section 47, Children Act 1989). EDs are primary agencies receiving those who have sustained a serious injury, including those who are motivated to bypass other agencies or whose assailant is motivated to ensure their victim avoids scrutiny.

Treatment for an ARA in ED aims to address symptoms (eg, injury) that may not necessarily characterize the underlying reasons for violence (eg, alcohol dependency), and staff do not always have the resources available to address such modifiable risks and vulnerabilities. However, without addressing them, the risk of repeat unscheduled ED attendance remains, including violence recidivism. For these reasons, services like VPTs that work within a WSMA approach to better understand reasons for ARA are required. Moreover, for those who are most vulnerable, ED may be the only realistic opportunity for patients to enter a system of care. As such, an ARA is often a sentinel event.

Aims and Objectives

The overarching aim of the Emergency Department Violence Intervention Program Evaluation (EDVIPE) is to determine whether VPTs are effective and cost-effective from the perspective of the NHS.

- Objective 1: to assess whether patient involvement with a VPT reduces the likelihood of unscheduled ED reattendance. We consider case and control patients' ED unscheduled reattendance for a minimum of 12 months following the initial ARA.
- Objective 2: to determine whether the presence of the VPT improves ascertainment of ARAs in ED attendances. We will consider the change in identified ED ARAs across intervention implementation in case and control EDs in Wales.
- Objective 3: to derive the costs of the VPT and compare those to the benefits of the intervention and understand whether the VPT represents value for money from an NHS perspective. If an effect is observed, then models will estimate the health impacts, costs, and potential savings over a longer time period (eg, 10 years) and for a national rollout.

Methods

Intervention

A process and implementation evaluation that describes both the planned and implemented VPT intervention is available elsewhere [37] and is further described in a Template for Intervention Description and Replication (TIDieR; Multimedia Appendix 1 [38]).

Intervention as Hypothesized

VPTs, which emerged from the VPU violence prevention strategy, were funded by the UK Home Office and Youth Endowment Fund (YEF), with the funding administered by the VPU and the Office of the South Wales Police and Crime Commissioner (Multimedia Appendices 2 and 3). Other HVIPs in the United Kingdom are volunteer-based, whereas the VPTs are nurse-led. The original implementation for VPTs was focused on identifying and supporting ED patients aged 11-25 years and on formalizing the identification of modifiable risks and vulnerabilities, supporting and advising patients, and to signposting to other services as appropriate. The VPTs also aimed to raise awareness of the service across ED clinical teams, with the aim of it becoming embedded within usual practice, and to train and upskill the clinical team to enable ascertainment and referral.

Intervention as Implemented

Since November 2019, a collaborative VPT between the police and NHS has been operational in a South Wales Type I (consultant-led with resus) ED in Cardiff (the capital and largest city in Wales), and a second VPT began in an adjacent South Wales Type I ED in Swansea (the second-largest city in Wales) in January 2022. The VPTs initially sought to identify patients attending the ED due to violence. This remit was broadened, with VPTs subsequently receiving referrals from across the hospital and other community health care teams (eg, Minor Injury Units [MIUs] and General Practitioners).

XSL•FO
RenderX

The VPTs work with patients to gain an understanding of any circumstances contributing to their exposure to violence. They refer patients into care pathways (primary, secondary, and tertiary care, or third-sector organizations) to address any vulnerabilities or modifiable risks and can work alongside the third sector (nonprofit and charitable enterprise) to provide continual case management. The VPTs also train other staff within the hospital to improve the identification of violence-related injury, to support clinical staff interactions with patients, and to maintain safeguarding procedures. In addition, the EDs at Cardiff and Swansea take part in information sharing to tackle violence, in which anonymous data and intelligence regarding violent incidents are shared with community safety partnerships. These anonymized data enable partner resources to be best used for violence prevention, part of the Cardiff Model for violence prevention [39]. Following implementation, both VPTs expanded the age range of patients to encompass all age groups.

Usual Care

Under usual practice, clinical ED staff are obliged to undertake safeguarding activities and provide for those attending due to violence. However, provision varies across EDs. Under usual practice, when people attend ED with an injury suspected to be caused by violence, their injuries are treated, and the patient is encouraged to contact the police or have the ED contact the police on their behalf. In cases of serious injury involving weapon use, the ED is obliged to contact the police irrespective of patient consent. In terms of general safeguarding, all patient-facing clinical staff are expected to have up-to-date safeguarding training and thus to carry out safeguarding tasks. Patients who are experiencing domestic violence can be referred to an Independent Domestic Violence Advocate (IDVA). Patients attending due to sexual assault can be referred to an Independent Sexual Violence Advocate (ISVA). For the 10 control EDs in Wales, 2 EDs have an IDVA, and 2 others have access to an IDVA not based in their ED. Furthermore, ED staff can also refer patients in a MARAC, typically in cases where the criteria are not met for formal safeguarding but the clinician suspects that something is not right. To facilitate, multiagency referral forms are filled out for children who require safeguarding, and Vulnerable Adult VA1 Forms (to support the referral of vulnerable adults) are completed for vulnerable adults who require safeguarding.

There are IDVAs based in the intervention EDs in Cardiff and Swansea. Only Cardiff has an ISVA. Broadly, usual practice focuses on children and victims of domestic violence. The patients eligible for VPT support are therefore those who are not eligible for support from the IDVA or ISVA and are typically older than 10 years. Apart from the IDVAs in control EDs, none have additional resources specifically dedicated to the role of supporting patients attending due to violence, relying mainly on existing clinical staff to support safeguarding within their departments. This may involve naming an existing member of staff as a safeguarding ambassador or having a nurse act as safeguarding lead for the department. Across Welsh EDs, some control EDs have provisions for victims of domestic violence that includes cards with the “Live Fear Free” helpline that they can give to patients experiencing domestic violence and who do not meet the criteria for a MARAC referral. Similarly, none of the control EDs have processes in place to support patients’ referral to outside agencies. If patients disclose that they are struggling with issues, or if staff suspect patients are experiencing an issue, then referrals will be made by clinical staff. However, this is not the same as VPT members working with patients to identify modifiable risks and vulnerabilities that contribute to their experience of violence. The VPTs in Cardiff and Swansea are therefore unique.

Design and Conceptual Framework

A controlled longitudinal whole population (Wales, United Kingdom) natural experiment. The intervention in Cardiff began in November 2019, and in January 2022 in Swansea. Due to earlier changes in Emergency Department Data Set (EDDS) coding, these data are consistent and available from January 2012. Intervention data collection therefore begins in November 2019 and ends in August 2023, allowing a 12-month follow-up of patients until August 2024.

Objective 1: Effectiveness

We hypothesize that engagement with the VPT will help patients overcome modifiable risks and receive support for vulnerabilities, and that therefore the intervention will reduce the recurrence of unscheduled ED attendance.

Other than those who are most seriously injured, patients will register at ED reception and be triaged, at which point the most appropriate pathway through ED will be determined. At reception patients will be asked about the reason for their attendance, including whether it was due to an assault, data that becomes a part of the Patient Management System (PMS) [39]. Patients may not disclose that the reason for their attendance was assault-related. They might be reluctant, the perpetrator may have accompanied them, or they may wish to avoid scrutiny. One function of the VPTs is to work across clinical teams to improve ascertainment of ARAs. The result is that patients can be stratified according to the extent that they engage with the intervention.

1. Patients identified in the ED PMS data as having attended due to an assault, but with no further contact with the VPT.
2. Patients identified in the ED PMS or VPT data as having attended due to an assault but did not further engage with the VPT.
3. Patients identified in the ED PMS and VPT data and who engaged with the VPT.

The primary analysis concerns group 3. The reasons for patients not engaging are potentially related to underlying characteristics and in secondary analyses we will explore this. However, it is reasonable to assume that the 3 groups represent varying levels of intervention dose, and therefore secondary analyses and analyses using groups 1-3 will be informative.

Objective 2: Ascertainment

Our second hypothesis is that intervention implementation improves ARA ascertainment.

Patients attending ED can do so repeatedly within periods of time. This frequency is likely associated with the modifiable

XSL•FO
RenderX

risks associated with ARA and is of interest here. For Objective 2 it is therefore appropriate to determine the proportion of attendances identified as violence related. This generates time-series data. The outcome of interest is therefore the count of ARAs across all EDs. ARA attendance is defined as ARA in the ED PMS or, in the case of intervention sites, in the ED, PMS, or VPT data. Codes indicating the Provider Site (the ED in its hospital) are available in the EDDS, and this allows comparison between intervention EDs and control EDs. While intervention EDs have been in continual service since before the time-series start dates, this is not so for all Type I EDs in Wales; there have been several changes, with some EDs closing and others opening or being modified to receive additional patients. This, coupled with the intervention sites located at two of the largest hospitals in Wales, reduces the scope for selecting matching control sites [40], leading to potentially important baseline differences in the interrupted time-series data [41]. The counterfactual will therefore be ARAs across all control Type I EDs.

Objective 3: Cost-Effectiveness

We aim to determine whether the VPT represents value for money. The primary outcome will be quality-adjusted life years, which will be estimated based on effectiveness estimates comparing ED attendance, reattendance, and any injuries received for those engaging in the VPT service relative to those who do not. Costs will be captured from an NHS perspective, reflecting the cost of the intervention but also other costs to the NHS due to referral. A secondary cost-effectiveness analysis will undertake a societal perspective, which will explore additional costs across social care, the police, and the third sector.

Secondary Analyses

We aim to co-produce study protocols with collaborators and Patient and Public Involvement and Engagement (PPIE) groups to provide opportunities for them to shape secondary and additional epidemiological analyses. This facilitates opportunities to realize and contribute to what is a rapidly changing policy area. One example is our inclusion of school attainment, exclusions, and attendance in analyses, which have been highlighted in these early discussions. VPUs and VRUs have made little headway working with the education system to challenge the causes of violence, and this has been identified as a priority [42]. Furthermore, additional exploration is planned to characterize those excluded from the intervention but who have available Anonymous Linkage Fields (ALFs).

Population and Data

This is a whole-population evaluation, including all residents of Wales, United Kingdom. Data are housed in the Secure Anonymized Information Linkage (SAIL) databank [43].

Data

Violence Prevention Team Data (Cardiff and Swansea)

Intervention sites record patient details (name, date of birth, and gender), NHS number, extent of engagement with the VPT, and whether any subsequent referral was made.

Administrative Data

Several administrative datasets are available to characterize patients and summarized in Textbox 1.

Table 1 summarizes key patient socioeconomic and demographic characteristics required to enable secondary analyses in relevant subgroups.

Textbox 1. Emergency Department Violence Intervention Program Evaluation (EDVIPE) administrate datasets.

The following routinely collected administrative and population datasets will be linked and analyzed to support the evaluation of the Emergency Department Violence Intervention Program Evaluation (EDVIPE).

Annual District Death Extract (ADDE):

Provides week of birth and date of death, used to describe left- and right-side censoring of study participants.

Outpatient Referral Dataset (OPRD):

Includes data on outpatient referrals from primary care, helping to characterize referral pathways to secondary care. This dataset captures clinical referrals from General Practitioners, general and community dental practitioners, Accident and Emergency Departments, walk-in services, and consultant-to-consultant referrals.

Patient Episode Data for Wales (PEDW):

Covers individuals domiciled in Wales and treated in Welsh and English National Health Service (NHS) Trusts. Includes inpatient and day-case activity data, with spells and episode-level information on hospital admissions, enabling tracking of service use related to an ARA.

Welsh Demographic Service Dataset (WDSD):

Includes all individuals registered with a Welsh General Practitioner and, through anonymization, enables identification of household groups.

Welsh Index of Multiple Deprivation (WIMD):

The Welsh Government's official measure of relative deprivation for small areas in Wales, based on 8 domains, including income, employment, health, and access to services. WIMD is typically grouped into quintiles and is included in several NHS datasets.

2011 and 2021 UK Census:

National censuses are conducted every 10 years that include key demographic information. Due to patient entry into and exit from the EDVIPE, some individuals may be missing from the 2011 Census (eg, born after 2011) or the 2021 Census (eg, died before 2021); therefore, data from both censuses are required.

XSL•FO
RenderX

Table 1. Emergency Department Violence Intervention Program Evaluation (EDVIPE) patient characteristics.

Characteristic	Source
Age (from Week of Birth [WoB])	Welsh Demographic Service (WDS)
Sex	WDS, 2011 and 2021 Census
Ethnicity	2011 and 2021 Census
Asian (Bangladeshi, Chinese, Indian, Pakistani, Other Asian)	
Black (Caribbean, African, Other Black)	
Mixed (White and Asian, White and Black African, White and Black Caribbean, Other Mixed, or Multiple ethnic groups)	
White (English, Welsh, Scottish, Northern Irish, British, Irish, Gypsy or Irish Traveller, Roma, or Other White)	
Other (Arab, Any other ethnic group)	
Quintile of residential deprivation	Welsh Index of Multiple Deprivation (WIMD)
Urban or rural residential classification	2011 and 2021 Census

Secondary Outcomes

To characterize the WSMA involvement of patients involved in the intervention, broader pathways will be explored across several related datasets, summarized in Textbox 2.

Textbox 2. Emergency Department Violence Intervention Program Evaluation (EDVIPE) secondary outcomes.

Children and Family Court Advisory and Support Service (CAFCASS) Wales family justice dataset includes information for residents of England and Wales on cases of divorce, private law, Family Law Act cases, public law, adoption, and other family law applications. It also includes information on marriage and divorce characteristics, the number of children involved in cases, and types of hearings.

Ministry of Justice (MoJ): Data First includes linked administrative datasets covering:

- Magistrates' court defendant data
- Crown Court defendant data
- Criminal courts and prisons data
- Prisoner custodial journey data
- Family Court data

National Pupil Database (NPD) contains 4 broad categories of data: demographics, educational attainment, absence and exclusion, and information on children in need and looked-after children.

Police Data (Police Crime Dataset): Pending applications for police crime data from all 4 Welsh police forces are progressing and will be explored. Data sharing agreements are currently being developed to bring all-Wales police crime data into Secure Anonymized Information Linkage.

Substance Misuse Dataset (SMSD), also known as the Welsh National Database for Substance Misuse (WNDSM), includes data on individuals in Wales presenting for substance misuse treatment, including assessments, referrals, and treatment history.

Welsh Longitudinal General Practitioner Dataset (WLGPD) contains clinical information from general practices in Wales, including diagnoses and referrals to secondary and tertiary care.

Cohort Inclusion and Exclusion Criteria

All residents of Wales who are 11 years of age or older are eligible for inclusion. Residents of Wales will be defined through their identification in the Welsh Demographic Service Dataset (WDSD).

The NHS assigns each patient domiciled in the United Kingdom a unique number. This NHS number links across various NHS data systems. The encrypted and anonymized ALFs are derived from these NHS numbers. Therefore, patients attending ED whose identity cannot be connected to an NHS number (eg, overseas visitors and tourists) will not have a corresponding ALF and will, by necessity, be excluded from analyses.

Allocation

Intervention patients will be identified in the VPT data and will have attended intervention EDs (in Cardiff and Swansea), subject to the above inclusion criteria. Control patients will be identified in the EDDS, and ED attendance was not in an intervention ED.

Progression Criteria

This is a definitive study. As the primary focus of the study uses routinely collected data, which is available for analysis subject

XSL•FO
RenderX

to information governance permissions and extraction, progression criteria are not applicable.

Sampling

From the intervention sites' data (Section 8.2.1), between October 2019 and December 2022, the Cardiff VPT contacted 2312 patients, of whom 77% accepted VPT support (Figure 1). We conservatively estimated that there will be 2500 patients that engaged with the Cardiff VPT across the 4 years (2019-2024) of VPT operation and a further 900 from the 2 years (2021-2024) operation in Swansea. Across the entirety of Wales, there are approximately 1 million ED attendances each year.

Figure 1. Count of patients, by month, identified by the Cardiff violence prevention team (VPT) as attending due to a violence-related injury, with the number that subsequently engaged with the VPT.

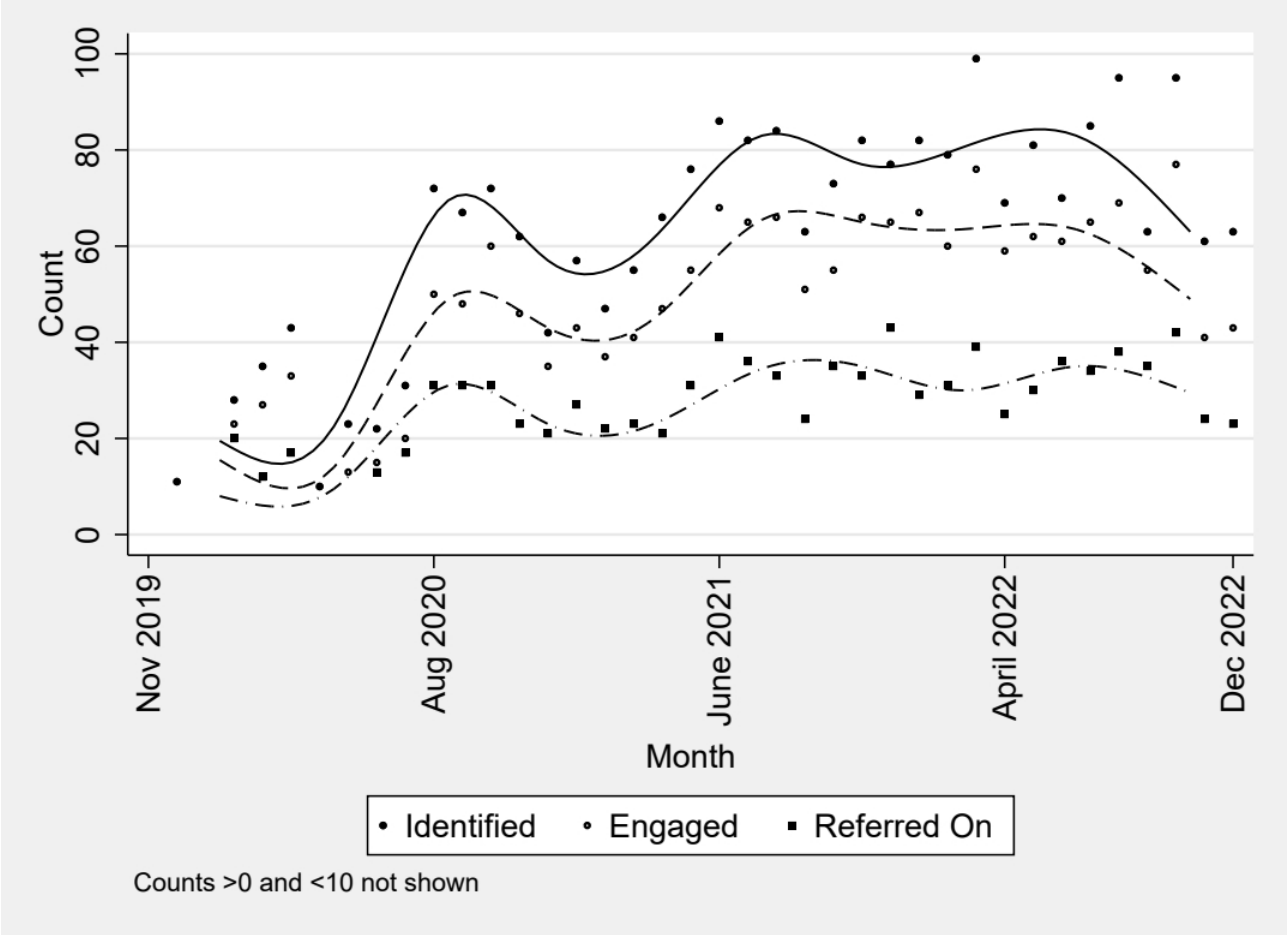

Initial estimates from Cardiff VPT suggest that 3% of those engaging with the VPT reattended ED at least once within one year, compared to 23% of patients who did not engage with the VPT. Data from 2015 and 2016 suggest that the frequency of unscheduled attendances for patients with at least one ARA (mean 2.35 attendances) is greater than for patients making an unscheduled attendance without evidence of an assault (mean attendances 1.73). For a simple Cox survival model, (α=.05; β=.90), and a hazard ratio of 0.8, a total N of 845 is required.

To realize the recurrent nature of analyses, simulation [44] (1000 estimates per point estimate) was used across varying follow-up periods, which suggests a 12-month follow-up period and total N of 300 is adequate to identify a significant effect. By increasing the number of controls, statistical power will be further enhanced [45].

Analytic Strategy

Objective 1: Effectiveness

Effectiveness Overview

Our primary outcome is unplanned ED attendance. It represents the cost to the NHS of serious health care events and acts as a proxy for events eliciting acute health care needs. EDs provide acute care for patients without prior appointment and the aftercare of patients who have received ED treatment, but where there is no alternative provision (eg, for out-of-area tourists). There can, therefore, be follow-up and planned appointments in ED. These appointments in ED will be made where, for example, there is an element of diagnostic uncertainty and a review is required in the ED context, or for patients where other follow-up arrangements are likely to fail (eg, visitors to the area without access to primary care) [46]. Follow-up and planned appointments in ED are typically a continuation of the initial unplanned attendance or referral from another health care provider and are not valid outcomes for EDVIPE, as they are

XSL•FO
RenderX

not elicited in response to acute health care needs. Thus, for Objective 1 we will censor the timeline. Left-side censoring at birth or when someone takes up residence in Wales. Right-side censoring when someone dies or moves out of Wales. We further interval censor the timeline to account for repeat ED attendances associated with a health event, such as a referral from a local emergency hospital to a Major Trauma Centre (MTC).

Discontinuous Risk Interval

Accounting for periods when individuals are not at risk is an essential consideration in repeated time-to-event models [47-50]. The clearest example in the current context is ensuring time at risk does not extend beyond the date of death or originate before birth. Similarly, time at risk will also be left-side censored if patients move into Wales and right-side censored if they move away from Wales. With no adjustment for these discontinuous risk intervals, the time at risk will be incorrect, increasing a greater likelihood of Type II errors.

How patients are routed through emergency care pathways in Wales also influences time at risk. ED attendances are mainly determined by the acuity of the patient's condition and the urgency with which they need to be seen. These decisions can be made by the Welsh Ambulance Service Trust (WAST), a MIU, or in the local ED. It is feasible that a patient initially attends a local ED to be stabilized, is assessed, and requires referral to a MTC or trauma unit (TU). Each MTC and TU is attached to an ED, and therefore, in response to severe injury, patients are registered in more than one ED if they are referred from an ED without trauma facilities to EDs that are attached to a TU or MTC.

In Wales, emergency care is provided in MIUs, EDs (Local Emergency Hospitals, and Rural Trauma Facilities), TUs, and MTCs. There is one MTC in Cardiff University Hospital Wales, which services South and West Wales and South Powys and acts as a TU for the local population. Morriston Hospital in Swansea is a TU, but with additional specialist services (eg, orthoplastics), meaning some patients otherwise destined for the MTC would be referred there instead. North Wales is serviced by the MTC in the Royal Stoke University Hospital in England.

Patient admission is described using spells and superspells. A spell represents patient care under a single specialty, and a superspell comprises multiple spells, where 2 spells are considered part of the same superspell if admission to the second specialty occurs within 48 hours of discharge from the first. ED attendances included within the same superspell (eg, transfer from an ED to an MTC) will therefore be attributable to the same initiating event, whether planned or unplanned. Therefore, only unscheduled attendances at the initiation of a superspell will be included, and additional ED attendances within the same superspell will be analytically censored. The duration of a superspell constitutes a discontinuity in a patient's exposure to risk, which is then accommodated in our analytic approach; this is described further in Figure 2.

Figure 2. A simplified example of discontinuous risk intervals as they relate to unscheduled Emergency Department (ED) attendance and superspells.

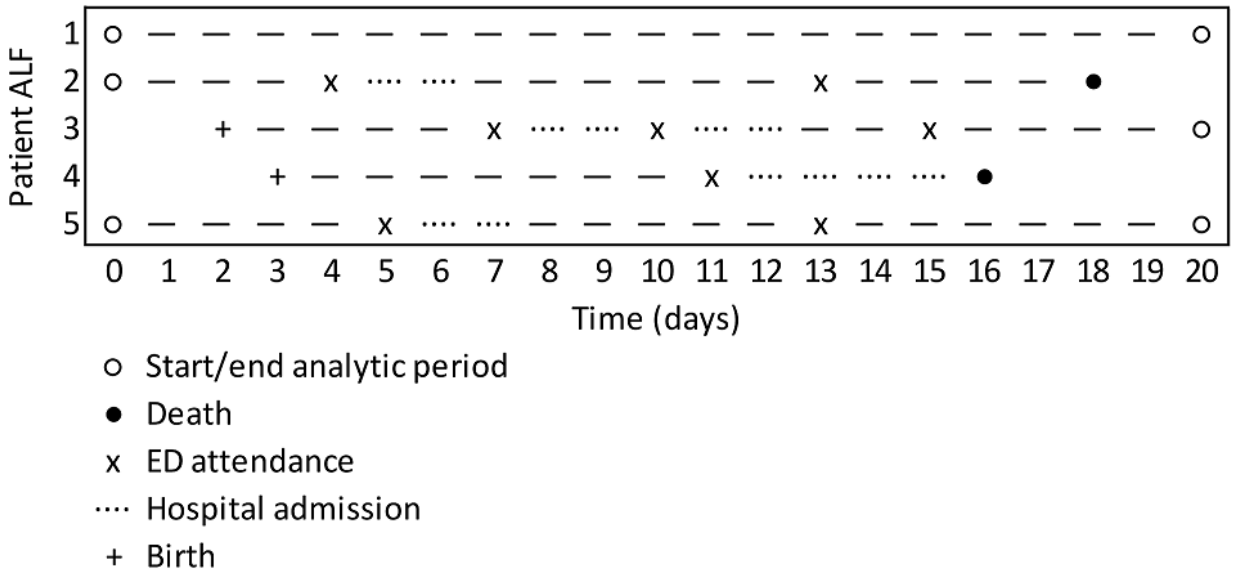

In Figure 2, a hypothetical analytic period runs from day 0 to day 20. Patient 1 lived in Wales for the duration of the study but had no ED attendances. Patient 2 lived in Wales at the start of the analytic period but died on day 18. They had 2 ED attendances, with the first resulting in a 2-day stay in the hospital. Their time at risk for the first ED attendance is 3 days, and for the second 6 days, with a total time at risk of 16 days. Patient 3 was born in Wales on day 2, had 3 ED attendances, and was alive and living in Wales at the end of the analytic period. However, their second ED attendance (eg, transfer to an MTC) occurred in a superspell associated with the first ED attendance, and this is dropped from consideration. They

XSL•FO
RenderX

therefore had 2 ED attendances, with the time at risk for the first as 4 days and the time at risk for the second as 2 days. Their total time at risk is 10 days.

We assume that unscheduled ED attendances are independent and unordered, and therefore assume a common baseline hazard for all events. Thus, the hazard function, , for the kth ED event for each subject (i) is:



However, it is feasible that operational differences across Local Health Boards might influence the likelihood that ED attendance varies across groups of patients. For example, initiatives that influence the likelihood patients are conveyed into emergency care or alternative provision in the community for lower acuity patients. We introduce a random effect, , to describe variation in risk, or frailty, that can be introduced [51]:



Dates and Times

Records on the date and time of events in routine data require consideration. With most events recorded by day, we are obliged to use day as the primary measure of time, and therefore collisions will occur.

In EDDS the date and time of attendance and discharge are recorded, and several events can occur on the same day. For example, it is feasible that a patient attends ED on multiple occasions and on the same day. ED attendances on the same day are collapsed onto one another under the superspell assumption. In the WDSD, while date of birth is approximated to the Monday of the week in which they were born (their Week of Birth), date of death is accurate to that degree of granularity.

For example, a patient might attend ED at 9 AM and die at 2 PM on the same day. To avoid inconsistencies such as this patient entering the study cohort after death, we will adopt the usual approach and add a small value (ε) to dates to preserve the chronological order of events [52].

Matching

It is not possible to manipulate allocation to treatment, and we therefore rely on quasi-experimental methods to infer treatment effects. The purpose of matching control and intervention patients is to allow derivation of the average treatment effect under the assumption that allocation is not conditional on observed confounders. Matching on values of covariates between the treated and untreated groups avoids the bias introduced by covariates that influence the outcome variable. The goal being to find a subset of data that is closest to an exact match on observed covariates [53]. The primary analysis will be conducted using coarsened exact matching, with a minimum 1:1 ratio, with secondary analyses undertaken using propensity score matching.

For categorical covariates, exact matching is a sensible alternative. Here a match is considered acceptable only if the values of all covariates are equal between treated and untreated potential matches. Continuous covariates can, for the purpose of using exact matching, be turned into categorical covariates by defining category intervals; the resulting method is called coarsened exact matching. For matching, covariate categories can be grouped together to form larger (and fewer) categories, improving efficiency [53-55]. This approximates to a fully blocked experiment and requires temporarily coarsening variables.

Propensity score matching reduces the multidimensional space of covariates to a single summary covariate, the propensity score, typically derived using a logistic equation and performing regression on the treatment group for the covariates chosen to match with. The proximity of a potential match to the given subject is estimated in terms of the closeness of their propensity scores. This method simplifies subsequent analyses, as matching refers only to one variate, the propensity score. However, propensity score matching can increase, rather than decrease, the imbalance of the covariates in the samples [53]. A challenge is optimizing the trade-off between the quality of matches and the sample size. A larger sample reduces sampling error and can increase the study power, but including matches of lower quality may lead to greater residual imbalance [56]. The R package Matching Frontier attempts to address these issues by calculating the entire balance-sample size frontier, from which the user can easily choose one, several, or all subsamples to use for their final analysis, given their own choice of imbalance metric and quantity of interest [57,58].

Missing Data

Missing data may arise specific to the use of census data, where it is feasible that records are missing entirely or patients did not disclose personal characteristics. However, given that these characteristics are recorded across multiple datasets, the expectation is that missingness will minimally impact the analyses. Nevertheless, missingness will be considered should it arise in the process of analysis.

Objective 2: Ascertainment

The hypothesis is that a violence attendance is more likely to be ascertained as such with the additional VPT resource present in ED. All unplanned attendances can be coded as ARA or not (1, 0). The time series begins in January 2012, and the 2 sites (Cardiff and Swansea) can be individually compared with all other Type I EDs in Wales. This indicates that a difference-in-difference model [59] is appropriate. We will analyze Cardiff first, using the implementation in Swansea as a future replication of the intervention to facilitate a more robust causal interpretation of any effect. Additional descriptive analysis can explore patient groups (age, gender, and time of attendance) more likely to be ascertained by the VPTs. Unlike Objective 1, we are interested in all ED attendances, as the opportunity to provision safeguarding is applicable across all ED attendances, irrespective of the frequency of attendance.

Objective 3: Cost-Effectiveness Analysis

The aim of the economic evaluation is to understand the costs and consequences of providing a VPT versus no VPT within EDs. For this study, we will focus on the objectives of understanding whether the VPT represents value for money from an NHS perspective.

XSL•FO
RenderX

Methods

A within-study analysis will determine the short-term cost-effectiveness of offering VPT relative to not offering VPT within EDs. Estimates derived from this study will be used to inform a long-term decision analytic model examining the cost-effectiveness of providing, versus not providing, VPT within EDs over a longer time horizon. Alternative perspectives for the analysis will be considered, with the NHS perspective being the base case. Cost-effectiveness will be presented using standard statistics.

Exposure Variable

The exposure variable consists of patients engaging with the VPT.

Outcome Variables

Health-related outcome variables include improved mental health outcomes, measured as a reduction in primary and secondary care visits for a mental health reason, reduced substance use, improved physical health, measured as a reduction in ARAs in ED, and improved health-related quality of life (HRQoL), measured using injury-related estimates from the literature. Data on all outcomes will be sourced from the SAIL databank [43]. Estimates will be measured on a quarterly basis for the follow-up period (12-48 months). HRQoL will be estimated using the literature, in line with other studies on hospital-based violence prevention programs.

Within Study Statistical Analysis

Differences in overall mean outcomes will be analyzed using a fixed effects panel regression model, where we will control for individual and time-varying effects [60].

Covariates of Interest

The following covariates will be included within the fixed effects panel regression model, patient age, gender, severity of injury, number of years the VPT has been operating, year, and quarter. Referral uptake will be treated as a mediating variable and not adjusted for in the models.

Long-Term Cost-Effectiveness Analysis

A Markov decision model with quarterly cycles will be used to estimate the cost-effectiveness of offering VPTs in EDs. To determine cost-effectiveness using a health and social care perspective (primary analysis), the Markov model will be health-state based, with key health states likely to be healthy, injured, or dead. To support a multisector perspective (secondary analysis), the Markov model will be event-based.

Resource Use

Health care resource use will include number and type of visits with

- primary care (eg, General Practitioner),
- secondary care (eg, inpatient stays and outpatient clinic appointments), and
- other services (eg, other health care professional visits, physiotherapy, occupational therapy, mental health therapy, and drug and alcohol treatment).

Cost Data

Cost data will be estimated for health service use based on resource use from the SAIL databank with appropriate unit costs applied. Third sector resource use will be valued using cost data provided in the resource use questionnaires. Other unit cost estimates (eg, NHS reference costs, unit costs of health and social care) will be used to supplement where needed.

Differences in overall mean total health care costs (including primary and secondary care) between groups will be analyzed using generalized linear mixed models. Cost data are known to be left-skewed with a substantial number of zeroes, and there is no single dominant method for analysis [61]. We will test several potential models from the generalized linear mixed model family and identify the correct link function and distribution to use. Potential effect modifiers include age, gender, and deprivation. Mediating variables include intervention engagement. Potential confounding variables include age, gender, deprivation, and the number of years VPT has been operating. The regression model will also include time variables recording the year and quarter.

Costs and benefits will be discounted at 3.5% per annum as recommended by the HM Treasury [62].

Time Horizon

The time horizon for the decision analytic model will extend beyond the time frame of this study to encompass a period of time where, if possible, all relevant costs and outcomes for this appraisal have been accrued [63].

Cost-Effectiveness Analysis Outcomes

For the primary analysis, we will report the incremental cost-effectiveness ratio as the cost per quality-adjusted life-year, the net monetary benefit, and the net health benefit for those being offered versus not offered VPT assessment during their visit to the Accident and Emergency Department. The net monetary benefit and net health benefit will be assessed using a range of values (£15,000-£30,000 [approximately US $20,709-US $41,418, using an exchange rate of 1£≈1.38 US$] per quality-adjusted life year) representing a health decision maker's willingness-to-pay to obtain the distribution of net benefits at different levels of willingness-to-pay.

For the secondary (multisector) analysis, we report the cost per unit of effectiveness in natural units (eg, cost per reconviction avoided). We will also present the results using the extended impact inventory framework and consider alternative methods of aggregation [64].

Sensitivity Analysis

Parameter uncertainty will be assessed using a probabilistic sensitivity analysis, varying key parameters over a range of expected values, and running 1000 Monte Carlo simulations.

Social Costs

Additional exploratory analyses will estimate the broader social costs associated with the VPTs. Social costs will focus specifically on organization costs of supporting patients who have been referred to them by the VPTs. Cost data of organizations will be assessed using top-down gross costing.

XSL•FO
RenderX

Questionnaires or interviews, depending on interviewee preference, will be conducted with members of organizations to assess (1) the total annual expenditure cost for third sector and statutory organizations in the given year [65]. This cost is the total yearly cost of running the service, including costs of supporting patients, staff costs, and consumables. (2) The total number of users supported each year by the third sector and statutory organizations. (3) The percentage of those supported that were referred by the VPTs each year.

The given year will be 2022-2023; this is due to the Swansea team being implemented in January 2022. Top-down gross costing allows data to be aggregated and will allow us to estimate the mean cost of a "typical patient" to a "typical organization." We will include the distribution for each organization but will not attribute a specific cost to a specific organization. Furthermore, we will ask organizations for standard cost data (eg, cost per counseling session); this will be used as a comparison of our mean average. We will ask about waiting lists for support, which will be used to inform our knowledge.

PPIE

Extensive PPIE has been and will continue to be undertaken. The rationale is that many patients managed by the VPTs will be vulnerable, with some at the beginning of their journey in the support they receive. The expectations were that these patients would be unlikely to reflect meaningfully on the VPT within the study timeline and therefore alternative opportunities to explore patients' perceptions were required. Furthermore, follow-up qualitative work with young adults in emergency care, the dominant group in ED, requires considerable resourcing and suffers from high levels of attrition [66]. We therefore sought PPIE engagement in order that those with experience of the emergency health care system were able to feed into the project, co-produce methods, provide their interpretation of the results and assist with interpretation and dissemination. Groups include survivors of domestic violence, carers, and those who have experienced alcohol and drug dependence, homelessness, sexual exploitation, and mental health issues. One PPIE co-investigator was appointed to lead on monitoring equality and diversity, with a second, and experienced, PPIE co-investigator supporting inexperienced PPIE members. PPIE activity was captured using the short-form Guidance for Reporting Involvement of Patients and the Public (GRIPP) [67].

The PPIE groups include lay members with experience of PPIE work and service users from primary and emergency care research, who provided lay perspectives to the research team during the development and conduct of the study; subsequent engagement was undertaken to strengthen the relevance, quality, and dissemination of the research. Additional PPIE members were recruited. These members had lived experience relevant to the patients who are the subject of the intervention (homelessness and domestic violence). Service Users for Primary and Emergency Care Research (SUPER) and the PPIE co-investigators advised on how the research team should engage with those who have lived experience but had not had prior experience of PPIE involvement.

PPIE Activity

SUPER provided feedback on the original proposal and then provided advice on how the materials should be developed for the 2 less experienced PPIE groups. These PPIE groups with lived experience were recruited to give input on the protocol. One group consisted of people with lived experience of homelessness and related conditions, recruited from a charity supporting those who are experiencing homelessness. The second group was comprised of survivors of domestic violence and was recruited through Welsh Women's Aid. The PPIE group members gave feedback on the VPTs and provided their first impressions of the research protocol. The groups also helped to develop a list of potential organizations to disseminate research findings and discussed methods on how this could be achieved. They explained what may be going on in the lives of people who experience violence and where people might turn to for support.

PPIE Design Input

PPIE contributed in different ways to the initial development of the research proposal and the development of the research protocol (Textbox 3).

XSL•FO
RenderX

Textbox 3. Contributions of Patient and Public Involvement and Engagement (PPIE) to study design, protocol development, and data selection.

PPIE contributed to the development of the research proposal and protocol in the following ways:

- Initial consultation in planning the funding application.
- The research proposal was reviewed in July 2021, and the input and advice contributed to a successful funding application. The research team also responded to the formative comments when developing the protocol and continues to build on them.
- PPIE work influenced the types of data being included in the study. For example, an issue was raised about whether some patients would admit to experiencing violence in the Emergency Department (ED), and therefore opportunities for the Violence Prevention Teams to improve ascertainment.
- Data on ethnicity from the 2011 and 2021 Census, which will also have some indication of those residing in refugee centers (and nursing homes, hostels, etc), will be included in the study following recommendations that racial violence and the experiences of asylum seekers should be considered.
- It was further indicated that considering school attendance and exclusions data, free school meals and special educational needs would be valuable, in addition to educational attainment.
- The Emergency Department Violence Intervention Program Evaluation (EDVIPE) Stakeholder Reference Group now includes representation from primary care, secondary care, and the third sector, after working with these sectors was suggested.
- Following feedback, EDVIPE now involves PPIE groups with lived experience of different types of violence.
- Following consultation with PPIE lived experience groups, development and refinement of the protocol and the addition of exploratory work were undertaken to:
- Better understand the pathways patients follow up to their attendance in ED, notably whether General Practitioner consultations were not acted upon.
- To consider patient ascertainment and therefore eligibility for the intervention in ED, as some victims may not realize that they are victims of violence.
- Whether one or two nurses are sufficient to manage an expected high caseload of patients attending ED due to assault, and whether the lack of 24-hour VPT provision would mean some patients are missed.
- Some patients might choose to avoid the police and therefore be less reluctant to receive support from the police or police-aligned services. We might explore this in any VPT referral data available.
- There were concerns that intervention-related activity might become known to the perpetrator, therefore elevating the risk of subsequent harm. We can consider the immediacy of postintervention reattendance by patient group.
- That there are unique challenges for those who are disabled, both in terms of ability to engage and the nature of the support required.
- Some patients, notably those with children, may be less willing to engage, as they would not want to risk losing their home or children. We can consider engagement by gender and presence of dependent children in the patient's residence.
- The judicial system emphasizes the right for both parents to be involved with their children, if any. Involvement of parents in the court system might be associated with a blunted intervention effect.
- That PPIE involvement in future diffusion and dissemination activity would lend credence to the project's communication strategy.

Ethical Considerations

The approval to access and analyze data housed in the SAIL Databank [43], an ISO (International Organization for Standardization) 27001 certified and UK Statistics Authority accredited secure data environment, was granted by the SAIL Independent Information Governance Review Panel (IGRP; Ref: 1421). The IGRP comprises representatives from various organizations and sectors, including the British Medical Association, Welsh Government, Public Health Wales, National Research Ethics Service, Digital Health and Care Wales (DHCW), Swansea Bay University Health Board, and members of the public. All routinely collected anonymized data held in SAIL are exempt from consent due to the anonymized nature of the databank (Section 251, Control of Patient Information; 2006 National Health Service Act). At no time will identifiable data be made available to the research team. ED staff will curate data pertaining to patients' exposure to the intervention, which will be passed to DHCW, where it will be anonymized, and a project-specific ALF added, as will a residential ALF. These data will be passed to SAIL for linkage to anonymized VPT clinical data.

Results

As this is a protocol, there are no results. The study was funded in November 2021, preregistered in June 2022 (ISRCTN 41868) and commenced in November 2023. The current version of the protocol (v4.1) was published by the funders in July 2024 [68]. The collection of routine data used in the current study has no end date, as it is primarily generated by patient interactions with health services. The analyses for Objective 1 were completed in August 2025 with the expectation they will be publicly available in June 2026. The analyses for Objective 2 were completed in July 2025 and are published [69]. The analyses for Objective 3 were completed in January 2026 and are expected to be available in August 2026.

XSL•FO
RenderX

Discussion

Principal Findings

We hypothesize that, for Objective 1, we will find a reduction in ED attendance. Objective 2 is exploratory and arose through PPIE; while we expect that some patients may not be willing to disclose their exposure to violence, the extent of this effect in ED is not known. We hypothesize that the intervention will be cost-effective over a 10-year time horizon and from the perspective of the NHS.

As far as we are aware, this project will not only provide the first robust evaluation of violence prevention activities in ED, but the data available will also provide an unprecedented opportunity to consider some of the pathways that patients follow before, during, and after their attendance in ED. Ongoing PPIE and stakeholder engagement have already shaped the study objectives and methods. However, we are particularly interested in exploring some of the characteristics of patients that expose them to violence. There is growing interest in neurodevelopment conditions [70], and it is feasible that secondary analyses could inform the future design of ED-based violence prevention initiatives. Furthermore, there is a paucity of information on the costs and health consequences of violence. The primary outcomes from this effectiveness and cost-effectiveness evaluation will offer novel insights to the benefit of future research and therefore violence prevention. We anticipate future research might therefore focus on how intersectional vulnerabilities, such as neurodevelopmental conditions, age, gender and ethnicity, and how they interact with residential deprivation to influence the likelihood of serious violence.

Strengths and Limitations

Linkage requires knowledge of the patients' identity, anonymized and encoded as ALFs, to be included in the data return. There are instances where individuals may prefer to remain anonymous. Analyses are therefore limited to those who can be identified in routine data and linked to the WDS and therefore EDDS and related datasets. It is feasible that some may attempt to conceal their identity, and this could correspond to greater vulnerability. Furthermore, our reliance on administrative data means there is no information on the context in which their exposure to violence arose. There may be a marked difference in the modifiable risks and vulnerabilities for patients sustaining injury in and around premises licensed for the sale and onsite consumption of alcohol and those experienced criminal or sexual exploitation. Conversely, our primary hypothesis is that the additionality of the VPT intervention teams means ED is better able to work with patients and therefore address the variation of patients that attend.

Dissemination, Outputs, and Anticipated Impact

A diffusion and dissemination plan will be co-produced with stakeholders and PPIE groups. In so doing, we will define audiences (policy, practitioner, academic, lay, etc) that might find the project outcomes of interest. In so doing, we will determine appropriate modalities to communicate with each, including presentations, briefing documents, and media events, with content informed by audience need.

PPIE, Policy, and Practitioner Focused

An ongoing evaluation of VRUs and VPUs is to recommend that greater attention is paid to evidence-based interventions and that activities should consider the possible involvement of schools and therefore the Department for Education. Education data has been included, and we can therefore contribute to the evidence-based linking of school activity to violence. In addition, our PPIE engagement highlighted the need to consider ARA predictors, and engagement with the Home Office will further identify intermediate outcomes. We therefore aim to undertake interim analyses that will inform the final analysis and respond to significant emerging policy questions as they relate to evidence-based violence prevention and reduction initiatives. While these will likely translate to academic papers, we also seek to produce more accessible outputs and exploit existing networks in that respect.

Conclusion

This protocol describes, in detail, novel methods to undertake a novel and robust effectiveness and cost-effectiveness of Hospital Violence Intervention Programs. Engagement with the public and patients, including those with lived and living experience of violence, together with broad stakeholder engagement, has greatly influenced the study design and methods. For example, it was through patient engagement we became aware that not all patients might disclose their exposure to violence, an exploration of which is now a primary outcome. Furthermore, stakeholders being able to comment on the design of emergency care systems resulted in an analytic need to account for interval censoring in patient timelines. It is feasible that the novelty of this work will identify further matters that might influence the design of subsequent studies.

Acknowledgments

The authors would like to acknowledge Nurses Vicky Lee and Sarah Wilcox, leads for the intervention sites, for their support of this evaluation. Furthermore, we would like to thank the Public, Patient Involvement participants, many of whom had not prior experience working with researchers. Their enthusiasm and insights greatly impacted on this evaluation and the broader research team. Generative Artificial Intelligence was not used in any aspect of this work.

A Study Steering Committee (SSC) was convened to oversee and advise on progress. The committee included the Data Monitoring and Ethics functions and oversees any protocol amendments

XSL•FO
RenderX

Data Availability

The datasets to be analyzed during this study are available in the SAIL Databank repository [43].

Funding

This study is funded by the NIHR (Public Health Research Program [NIHR134055]). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.
The funder had no role in the design of the study, and will not have any role during its execution, analyses, interpretation of the data, or decision to submit results.

Authors' Contributions

SCM, S Brophy, A Watkins, S Walker, A Battaglia, HY, TL, and D O'Reilly contributed to conceptualization, funding acquisition, and writing of the original draft and review and editing. SCM, S Brophy, A Watkins, and S Walker also contributed to investigation, methodology, project administration, and supervision, with SCM and S Walker additionally contributing to resources and visualization. A Bandyopadhyay and S Premji contributed to data curation, formal analysis, investigation, validation, and writing-review and editing. S Borgia and MH contributed to data curation and investigation. A Newbury contributed to investigation and writing-review and editing. D Rawlinson, LS, JS, and VS contributed to funding acquisition, with LS, JS, and VS also contributing to writing-review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
TIDieR checklist.
[DOCX File , 20 KB - resprot_v15i1e86247_app1.docx]

Multimedia Appendix 2
Confirmation of NIHR review and funding. NIHR: National Institute for Health and Care Research.
[PDF File (Adobe PDF File), 115 KB - resprot_v15i1e86247_app2.pdf]

Multimedia Appendix 3
Confirmation of NIHR review and funding. NIHR: National Institute for Health and Care Research.
[PDF File (Adobe PDF File), 142 KB - resprot_v15i1e86247_app3.pdf]

Multimedia Appendix 4
Peer review report by Public Health Research Program (PHR) Funding Committee, NIHR Evaluation, Trials and Studies Coordinating Centre (NETSCC), National Institute for Health Research (NIHR).
[PDF File (Adobe PDF File), 246 KB - resprot_v15i1e86247_app4.pdf]

References

1. Purtle J, Dicker R, Cooper C, Corbin T, Greene M, Marks A, et al. Hospital-based violence intervention programs save lives and money. J Trauma Acute Care Surg 2013;75(2):331-333. [doi: 10.1097/TA.0b013e318294f518] [Medline: 23887566]
2. Cunningham R, Knox L, Fein J, Harrison S, Frisch K, Walton M, et al. Before and after the trauma bay: the prevention of violent injury among youth. Ann Emerg Med 2009 Apr;53(4):490-500. [doi: 10.1016/j.annemergmed.2008.11.014] [Medline: 19162376]
3. Kaufman E, Rising K, Wiebe DJ, Ebler DJ, Crandall ML, Delgado MK. Recurrent violent injury: magnitude, risk factors, and opportunities for intervention from a statewide analysis. Am J Emerg Med 2016;34(9):1823-1830 [FREE Full text] [doi: 10.1016/j.ajem.2016.06.051] [Medline: 27460511]
4. Kao AM, Schlosser KA, Arnold MR, Kasten KR, Colavita PD, Davis BR, et al. Trauma recidivism and mortality following violent injuries in young adults. J Surg Res 2019;237:140-147. [doi: 10.1016/j.jss.2018.09.006] [Medline: 30914191]
5. Newton AS, Hartling L, Soleimani A, Kirkland S, Dyson MP, Cappelli M. A systematic review of management strategies for children's mental health care in the emergency department: update on evidence and recommendations for clinical practice and research. Emerg Med J 2017;34(6):376-384. [doi: 10.1136/emermed-2016-205939] [Medline: 28119350]
6. Hinsliff-Smith K, McGarry J. Understanding management and support for domestic violence and abuse within emergency departments: a systematic literature review from 2000-2015. J Clin Nurs 2017;26(23-24):4013-4027. [doi: 10.1111/jocn.13849] [Medline: 28403521]
7. Ansari S, Boyle A. Emergency department-based interventions for women suffering domestic abuse: a critical literature review. Eur J Emerg Med 2017;24(1):13-18. [doi: 10.1097/MEJ.0000000000000416] [Medline: 27467558]

XSL•FO
RenderX

8. Brice JM, Boyle AA. Are ED-based violence intervention programmes effective in reducing revictimisation and perpetration in victims of violence? A systematic review. Emerg Med J 2020;37(8):489-495. [doi: 10.1136/emermed-2019-208970] [Medline: 32554747]
9. Koenig KL, Benjamin SB, Beÿ CK, Dickinson S, Shores M. Emergency department management of the sexual assault victim in the COVID Era: a model SAFET-I guideline from San Diego county. J Emerg Med 2020;59(6):964-974 [FREE Full text] [doi: 10.1016/j.jemermed.2020.07.047] [Medline: 32951933]
10. Goodall C, Jameson J, Lowe DJ. Navigator: A Tale of Two Cities. Glasgow, Scotland: Violence Reduction Unit; 2017.
11. Sivarajasingam V, Guan B, Page N, Peel E, Moore S, Shepherd J. Violence in England and Wales in 2019. Cardiff, UK: Cardiff University; 2020.
12. Sivarajasingam V, Guan B, Page N, Moore S, Shepherd J. Violence in England and Wales in 2020. Cardiff, UK: Cardiff University; 2021.
13. Introducing public health measures impact assessment 2019 (accessible version). GOV.UK. 2019. URL: https://tinyurl.com/3muws8bp [accessed 2026-01-23]
14. Shepherd JP, Ali MA, Hughes AO, Levers BG. Trends in urban violence: a comparison of accident department and police records. J R Soc Med 1993;86(2):87-88 [FREE Full text] [Medline: 8433313]
15. Sutherland I, Sivarajasingam V, Shepherd JP. Recording of community violence by medical and police services. Inj Prev 2002;8(3):246-247 [FREE Full text] [doi: 10.1136/ip.8.3.246] [Medline: 12226126]
16. Gray BJ, Barton ER, Davies AR, Long SJ, Roderick J, Bellis MA. A shared data approach more accurately represents the rates and patterns of violence with injury assaults. J Epidemiol Community Health 2017;71(12):1218-1224 [FREE Full text] [doi: 10.1136/jech-2017-209872] [Medline: 29056592]
17. Bath R. A Whole-System Multi-Agency Approach to Serious Violence Prevention: A Resource for Local System Leaders in England. London, UK: Public Health England; 2019.
18. Serious violence strategy. GOV.UK. 2018. URL: https://www.gov.uk/government/publications/serious-violence-strategy [accessed 2018-01-23]
19. Violence Reduction Unit interim guidance. GOV.UK. 2020. URL: https://assets.publishing.service.gov.uk/media/5e7de2cfe90e0706ef2db559/12VRU_Interim_Guidance_FINAL__003_2732020.pdf [accessed 2026-01-23]
20. Fit for the Future: 10 Year Health Plan for England. NHS. 2019. URL: https://www.longtermplan.nhs.uk/ [accessed 2026-01-23]
21. Better care for people with co-occurring mental health and alcohol/drug use conditions. Public Health England. 2017. URL: https://assets.publishing.service.gov.uk/media/5a75b781ed915d6faf2b5276/Co-occurring_mental_health_and_alcohol_drug_use_conditions.pdf [accessed 2026-01-23]
22. Duke AA, Smith KMZ, Oberleitner LMS, Westphal A, McKee SA. Alcohol, drugs, and violence: a meta-meta-analysis. Psychol Violence 2018;8(2):238-249. [doi: 10.1037/vio0000106]
23. Babor TF, Berglas S, Mendelson JH, Ellingboe J, Miller K. Alcohol, affect, and the disinhibition of verbal behavior. Psychopharmacology (Berl) 1983;80(1):53-60. [doi: 10.1007/BF00427496] [Medline: 6408672]
24. Malik NS, Munoz B, de Courcey C, Imran R, Lee KC, Chernbumroong S, et al. Violence-related knife injuries in a UK city; epidemiology and impact on secondary care resources. EClinicalMedicine 2020;20:100296 [FREE Full text] [doi: 10.1016/j.eclinm.2020.100296] [Medline: 32300742]
25. Pallett J, Sutherland E, Glucksman E, Tunnicliff M, Keep J. A cross-sectional study of knife injuries at a London major trauma centre. Ann R Coll Surg Engl 2014;96(1):23-26 [FREE Full text] [doi: 10.1308/003588414X13824511649616] [Medline: 24417825]
26. Boyle A, Frith C, Edgcumbe D, McDougall C. What factors are associated with repeated domestic assault in patients attending an emergency department? A cohort study. Emerg Med J 2010;27(3):203-206. [doi: 10.1136/emj.2009.072033] [Medline: 20304886]
27. DeWall CN, Anderson CA, Bushman BJ. The general aggression model. Psychology of Violence 2011;2(3):245-258. [doi: doi: 10.1037/a0023842]
28. Allen JJ, Anderson CA. General aggression model. In: The International Encyclopedia of Media Effects. Hoboken, NJ: Wiley; 2017:1-15.
29. Farrington DP. Early predictors of adolescent aggression and adult violence. Violence Vict 1989;4(2):79-100. [doi: 10.1891/0886-6708.4.2.79]
30. Hawkins JD, Herrenkohl T, Farrington DP, Brewer D, Catalano RF, Harachi TW. A review of predictors of youth violence. In: Howell JC, editor. Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions. Thousand Oaks, CA: Sage Publications, Inc; 1998:106-146.
31. Piquero AR, Jennings WG, Diamond B, Reingle JM. A systematic review of age, sex, ethnicity, and race as predictors of violent recidivism. Int J Offender Ther Comp Criminol 2015;59(1):5-26. [doi: 10.1177/0306624X13514733] [Medline: 24335783]
32. WHO. World Report on Violence and Health. Geneva: World Health Organisation; 2002.

XSL•FO
RenderX

33. Long SJ, Fone D, Gartner A, Bellis MA. Demographic and socioeconomic inequalities in the risk of emergency hospital admission for violence: cross-sectional analysis of a national database in Wales. BMJ Open 2016;6(8):e011169 [FREE Full text] [doi: 10.1136/bmjopen-2016-011169] [Medline: 27558900]
34. Ostrowsky MK. The social psychology of alcohol use and violent behavior among sports spectators. Aggress Violent Behav 2014;19(4):303-310. [doi: 10.1016/j.avb.2014.05.001]
35. Fazel S, Gulati G, Linsell L, Geddes JR, Grann M. Schizophrenia and violence: systematic review and meta-analysis. PLoS Med 2009;6(8):e1000120 [FREE Full text] [doi: 10.1371/journal.pmed.1000120] [Medline: 19668362]
36. Fu F, Nowak MA, Christakis NA, Fowler JH. The evolution of homophily. Sci Rep 2012;2:845 [FREE Full text] [doi: 10.1038/srep00845] [Medline: 23150792]
37. Moore S, Van Godwin J, Moore G, Hamilton M, O'Reilly D. Practitioner experiences of developing and implementing two UK ED-based hospital violence intervention programmes: a process evaluation. Emergency Medicine Journal 2025;42:536-541. [doi: 10.1136/emermed-2024-214333]
38. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 07;348:g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]
39. Florence C, Shepherd J, Brennan I, Simon T. Effectiveness of anonymised information sharing and use in health service, police, and local government partnership for preventing violence related injury: experimental study and time series analysis. BMJ 2011;342:d3313 [FREE Full text] [doi: 10.1136/bmj.d3313] [Medline: 21680632]
40. Linden A. A matching framework to improve causal inference in interrupted time-series analysis. J Eval Clin Pract 2018;24(2):408-415. [doi: 10.1111/jep.12874] [Medline: 29266646]
41. Linden A, Yarnold PR. Using machine learning to evaluate treatment effects in multiple-group interrupted time series analysis. J Eval Clin Pract 2018;24(4):740-744. [doi: 10.1111/jep.12966] [Medline: 29888469]
42. Evaluation of Violence Reduction Units 2020/21. URL: https://tinyurl.com/38pvnwbt [accessed 2026-01-23]
43. Jones KH, Ford DV, Thompson S, Lyons R. A profile of the SAIL databank on the UK secure research platform. Int J Popul Data Sci 2019;4(2):1134 [FREE Full text] [doi: 10.23889/ijpds.v4i2.1134] [Medline: 34095541]
44. Feiveson AH. Power by simulation. Stata J 2002;2(2):107-124. [doi: 10.1177/1536867x0200200201]
45. Hennessy S, Bilker WB, Berlin JA, Strom BL. Factors influencing the optimal control-to-case ratio in matched case-control studies. Am J Epidemiol 1999;149(2):195-197. [doi: 10.1093/oxfordjournals.aje.a009786] [Medline: 9921965]
46. Royal College of Emergency Medicine. Emergency Department Follow-up Clinics. London; 2015. URL: https://tinyurl.com/3skzwd6b
47. Law CG, Brookmeyer R. Effects of mid-point imputation on the analysis of doubly censored data. Stat Med 1992;11(12):1569-1578. [doi: 10.1002/sim.4780111204] [Medline: 1439361]
48. Rücker G, Messerer D. Remission duration: an example of interval-censored observations. Stat Med 1988;7(11):1139-1145. [doi: 10.1002/sim.4780071106] [Medline: 3201039]
49. Turnbull BW. The empirical distribution function with arbitrarily grouped, censored and truncated data. J R Stat Soc Ser B (Methodol) 1976;38(3):290-295. [doi: 10.1111/j.2517-6161.1976.tb01597.x]
50. Guo Z, Gill TM, Allore HG. Modeling repeated time-to-event health conditions with discontinuous risk intervals. Methods Inf Med 2018;47(02):107-116. [doi: 10.3414/me0478]
51. Balan TA, Putter H. A tutorial on frailty models. Stat Methods Med Res 2020;29(11):3424-3454. [doi: 10.1177/0962280220921889] [Medline: 32466712]
52. Gould W. Why can't a subject enter and die at the same time in the Cox model? Stata FAQs. Texas: StataCorp URL: https://www.stata.com/support/faqs/statistics/time-and-cox-model/
53. King G, Nielsen R. Why propensity scores should not be used for matching. Polit Anal 2019;27(4):435-454. [doi: 10.1017/pan.2019.11]
54. King G, Nielsen R, Coberley C, Pope JE, Wells A. Comparative effectiveness of matching methods for causal inference. In: Institute for Quantitative Social Science. Cambridge: Harvard University; 2011.
55. Stuart EA. Matching methods for causal inference: a review and a look forward. Stat Sci 2010;25(1):1-21 [FREE Full text] [doi: 10.1214/09-STS313] [Medline: 20871802]
56. Imbens GW. Matching methods in practice: Three examples. J Human Resources 2015;50(2):373-419. [doi: 10.3368/jhr.50.2.373]
57. King G, Lucas C, Nielsen RA. Matching frontier. Harvard University. 2017. URL: https://projects.iq.harvard.edu/frontier [accessed 2026-01-23]
58. King G, Lucas C, Nielsen RA. The balance - sample size frontier in matching methods for causal inference. Am J Polit Sci 2016;61(2):473-489. [doi: 10.1111/ajps.12272]
59. Wing C, Simon K, Bello-Gomez R. Designing difference in difference studies: best practices for public health policy research. Annu Rev Public Health 2018;39:453-469 [FREE Full text] [doi: 10.1146/annurev-publhealth-040617-013507] [Medline: 29328877]
60. Jones AM. Panel data methods and applications to health economics. In: Palgrave Handbook of Econometrics. Cham: Springer; 2007:557-631.

XSL•FO
RenderX

61. Jones AM, Lomas J, Rice N. Healthcare cost regressions: going beyond the mean to estimate the full distribution. Health Econ 2015;24(9):1192-1212 [FREE Full text] [doi: 10.1002/hec.3178] [Medline: 25929525]
62. The green book: central government guidance on appraisal and evaluation. GOV.UK. 2018. URL: https://www.gov.uk/government/publications/the-green-book-appraisal-and-evaluation-in-central-government [accessed 2026-11-23]
63. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press; 2015.
64. Walker S, Griffin S, Asaria M, Tsuchiya A, Sculpher M. Striving for a societal perspective: a framework for economic evaluations when costs and effects fall on multiple sectors and decision makers. Appl Health Econ Health Policy 2019;17(5):577-590 [FREE Full text] [doi: 10.1007/s40258-019-00481-8] [Medline: 31098947]
65. Špacírová Z, Epstein D, García-Mochón L, Rovira J, Olry de Labry Lima A, Espín J. A general framework for classifying costing methods for economic evaluation of health care. Eur J Health Econ 2020;21(4):529-542 [FREE Full text] [doi: 10.1007/s10198-019-01157-9] [Medline: 31960181]
66. Irving A, Buykx P, Amos Y, Goodacre S, Moore SC, O'Cathain A. The acceptability of alcohol intoxication management services to users: a mixed methods study. Drug Alcohol Rev 2020;39(1):36-43 [FREE Full text] [doi: 10.1111/dar.13002] [Medline: 31769571]
67. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. BMJ 2017;358:j3453 [FREE Full text] [doi: 10.1136/bmj.j3453] [Medline: 28768629]
68. Moore SC, Brophy S, Bandyopadhyay A, Newbury A, Hamilton M, Battaglia A. Protocol: A Quasi-Experimental Effectiveness and Cost-Effectiveness Evaluation of Emergency Department Violence Intervention Programmes in the United Kingdom Southampton.: NIHR; 2024. URL: https://njl-admin.nihr.ac.uk/document/download/2047009 [accessed 2026-01-23]
69. Hamilton M, Snowdon L, Shepherd J, Premji S, Battaglia A, Yeomans H, et al. Nurse-led hospital violence intervention programmes improve emergency department identification of violence-related visits. Am J Emerg Med 2026;99:376-380 [FREE Full text] [doi: 10.1016/j.ajem.2025.10.040] [Medline: 41175441]
70. Whiting D, Lichtenstein P, Fazel S. Violence and mental disorders: a structured review of associations by individual diagnoses, risk factors, and risk assessment. Lancet Psychiatry 2021;8(2):150-161. [doi: 10.1016/S2215-0366(20)30262-5] [Medline: 33096045]

Abbreviations

ALF: Anonymous Linkage Field
ARA: Assault-Related Attendance
DHCW: Digital Health and Care Wales
ED: Emergency Department
EDDS: Emergency Department Data Set
EDVIPE: Emergency Department Violence Intervention Program Evaluation
GRIPP: Guidance for Reporting Involvement of Patients and the Public
HVIP: Hospital-Based Violence Intervention Program
IDVA: Independent Domestic Violence Advocate
IGRP: Independent Information Governance Review Panel
ISO: International Organization for Standardization
ISVA: Independent Sexual Violence Advisor
MARAC: multiagency risk assessment conference
MIU: minor injury unit
MTC: Major Trauma Centre
NHS: National Health Service
PMS: Patient Management System
PPIE: public patient involvement and engagement
SAIL: Secure Anonymised Information Linkage
SUPER: Service Users for Primary and Emergency Care Research
TIDieR: Template for Intervention Description and Replication
TU: Trauma Unit
VPT: violence prevention team
VPU: violence prevention unit
VRU: Violence Reduction Unit
WAST: Welsh Ambulance Service Trust
WDSD: Welsh Demographic Service Database
WSMA: whole system multiagency

XSL•FO
RenderX

YEF: Youth Endowment Fund

Edited by J Sarvestan; submitted 21.Oct.2025; this is a non–peer-reviewed article;accepted 19.Jan.2026; published 18.Feb.2026.

Please cite as:
Moore S, Brophy S, Bandyopadhyay A, Newbury A, Hamilton M, Battaglia A, Lowe T, O'Reilly D, Rawlinson D, Snowdon L, Shepherd J, Sivarajasingam V, Watkins A, Walker S, Premji S, Borgia S, Yeomans H
Effectiveness and Cost-Effectiveness of Emergency Department–Based Violence Intervention Programs in the United Kingdom: Protocol for a Quasi-Experimental Study
JMIR Res Protoc 2026;15:e86247
URL: https://www.researchprotocols.org/2026/1/e86247
doi:10.2196/86247
PMID:

©Simon Moore, Sinead Brophy, Amrita Bandyopadhyay, Annemarie Newbury, Megan Hamilton, Adele Battaglia, Trudy Lowe, David O'Reilly, David Rawlinson, Lara Snowdon, Jonathan Shepherd, Vaseekaran Sivarajasingam, Alan Watkins, Simon Walker, Shainur Premji, Sophie Borgia, Henry Yeomans. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 18.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Targeted Real-Time Assessment of Chronic Pain (TRAC-Pain) in Youth: Protocol for a Digital Biosignature Development Through a Prospective Observational Cohort Study

Anna H Bailes[1], DPT, PhD; Chi-Hung Shu[1], MEng; Alan Chang[1], PhD; Sahrish Masood[1]; Nicole Jehl[1]; Aliyah Davis[1]; Jeremy Giberson[1], MAS; Casey Cashman[2,3]; Allison Hill[3,4], MD; Javed Gill[3]; Ryan S McGinnis[5], PhD; Ellen W McGinnis[6], PhD; Nima Aghaeepour[1], PhD; Laura E Simons[1], PhD

[1]Department of Anesthesiology, Perioperative, and Pain Medicine, Stanford University School of Medicine, Palo Alto, CA, United States

[2]US Pain Foundation, West Hartford, CT, United States

[3]Patient and Pain Advocacy Partner, Palo Alto, CA, United States

[4]Division of Cardiology, Heart Institute, Children's Hospital of Los Angeles, Los Angeles, CA, United States

[5]Biomedical Engineering and Center for Remote Health Monitoring, Wake Forest University School of Medicine, Winston-Salem, NC, United States

[6]Social Sciences and Health Policy and Center for Remote Monitoring, Wake Forest University School of Medicine, Winston-Salem, NC, United States

Corresponding Author:
Anna H Bailes, DPT, PhD
Department of Anesthesiology, Perioperative, and Pain Medicine
Stanford University School of Medicine
3145 Porter Drive
Palo Alto, CA, 94304
United States
Phone: 1 650 665 3253
Email: ahbailes@stanford.edu

Abstract

Background: Approximately 1 in 5 children and adolescents live with chronic pain, with musculoskeletal (MSK) pain being one of the most prevalent subtypes. Unfortunately, some studies show that less than half of youth (<18 years) experience improvements with existing evidence-based treatments. Self-report measures—the current gold standard for monitoring the pain experience—are limited in their use as a single point-of-care assessment and their vulnerability to recall bias. The ubiquitous adoption of wearable technology presents a promising solution for improved monitoring of the pain experience via real-time tracking through a multisystemic lens.

Objective: The purpose of this study is (1) to develop a digital biosignature of the pain experience in youth with chronic MSK pain and (2) to assess the feasibility, acceptability, and appropriateness of the captured data for future applications.

Methods: All aspects of this study were designed in partnership with people with lived experience and patient advocacy partners. This is a longitudinal observational cohort study, with all study activities taking place remotely over a 12-week period. Up to 500 youth (between 14 and 24 years old) with chronic MSK pain will be enrolled through a multipronged recruitment strategy to ensure a representative sample. The participants will wear an Apple Watch throughout the study to continuously monitor physiological, sleep, and physical activity metrics. In addition, participants will complete brief “daily check-in” surveys that include gold-standard measures of the pain experience (eg, pain interference, mood, fatigue) and the option to report a “pain flare” (ie, a temporary but noticeable worsening of the usual pain symptoms). Participants will also complete a modified online Trier Social Stress Test and a 30-second sit-to-stand task to capture individual responses to standardized challenges. Traditional machine learning and deep representation learning methods will be used to develop a digital biosignature of the pain experience. The accuracy of the biosignature will be assessed through measures of model performance as compared to gold-standard self-reports.

Results: This study was funded in September 2024, with data collection beginning in March 2025. As of December 15, 2025, 190 participants are enrolled, with data collection and analysis ongoing.

Conclusions: This is the first study to leverage wearable health technology for real-time monitoring of the pain experience in youth with chronic MSK pain. The resulting digital end points are expected to heighten the rigor of clinical trials and provide opportunities for individually tailored interventions. The second phase of this study will investigate the implementation of

XSL•FO
RenderX

"wellness alerts" triggered by abnormal smartwatch readings. Alerts would empower users toward preemptive self-management strategies, thereby enhancing self-efficacy in those living with chronic MSK pain.

International Registered Report Identifier (IRRID): DERR1-10.2196/84781

(JMIR Res Protoc 2026;15:e84781) doi:10.2196/84781

KEYWORDS

pediatrics; musculoskeletal pain; wearable electronic devices; biosignature

Introduction

Approximately 1 in 5 children and adolescents live with chronic pain, with musculoskeletal (MSK) pain being highly prevalent (25.7%). [1] Chronic MSK pain significantly impacts quality of life, [2] school attendance [3,4], mood [2,4], and physical function [5]. Unfortunately, despite existing evidence-based treatment methods, many children and adolescents (ie, youth) with chronic pain experience little to no reduction in pain or disability [6-8]. In fact, up to 80% of youth with chronic widespread MSK pain become adults with chronic pain [9], emphasizing the urgency of both developing improved interventions and enhancing our ability to provide just-in-time intervention drawn from existing evidence-based tools. The chronic pain experience is not consistent over time; intra- and interdaily fluctuations in symptoms and the presence of pain flares contribute to unpredictability, uncertainty, and greater impairment for youth living with chronic pain [10,11]. Unpredictable fluctuations also present unique challenges for the monitoring and treatment of chronic pain.

Self-report measures are the gold standard for monitoring the pain experience in this population [12]. These measures capture multiple components of the pain experience, including pain interference, mood, fatigue, and physical function. However, self-report metrics are often burdensome and limited by their use as a single point-of-care assessment and vulnerability to recall bias. As such, those living with chronic pain may feel frustrated by the inability to accurately portray the nature of their pain experience [13], while clinicians may be left with a lack of data to make the most effective clinical recommendations for an individual patient. This often leads to inconsistent and insufficient treatment recommendations, to increased health care utilization, and the potential for further frustration and distrust between patients and providers [14].

The nearly ubiquitous consumer adoption of health care technology—particularly wearables [15,16]—presents a promising solution for improved comprehensive monitoring of the pain experience, with opportunities for personalized medicine approaches [17,18]. Unlike self-reports, wearables provide longitudinal and objective measures of health that reflect real-time fluctuations inherent to the pain experience. Wearables also present a unique opportunity for collecting multiple biomarkers (eg, heart rate and sleep data) that can be combined with other domains to inform the development of biosignatures. Through machine learning and related analytic techniques, several digital biosignatures of human health have already been developed using data from wearable technology [17,19,20].

In summary, mechanisms underlying the persistence of chronic pain in youth remain poorly understood. Multisystem involvement (eg, autonomic, sensory, immune) and frequent symptom fluctuations complicate the reliability of traditional self-report monitoring techniques and create barriers to individualized clinical decision-making. Therefore, the purpose of this study, named Targeted Real-Time Assessment of Chronic Pain (TRAC-Pain), is to overcome the limitations of the current standards for pain monitoring through the development of a digital biosignature using consumer-available wearable technology. The expected result will be a digital biosignature of the pain experience in youth with chronic MSK pain. The secondary aim of this study is to assess the feasibility, acceptability, and appropriateness of the captured data for incorporation into future applications of the biosignature.

Methods

Overview

This is a longitudinal observational cohort study, with all study activities taking place remotely over a 12-week period. Apple Health Kit (Apple Inc) and survey data will be collected and stored using a third-party digital health and research platform (Figure 1) approved by the Stanford University Institutional Review Board (IRB-77005). Study design includes multiple layers of data collection (eg, self-report measures, smartwatch metrics, and task-based measures), all of which are essential for biosignature development. Self-report measures encompass current gold-standard metrics of the pain experience and are crucial for biosignature validation. Smartwatch technology leverages the widespread adoption of wearables to capture real-time fluctuations across multiple body systems, which is critical for ensuring a holistic and longitudinal understanding of the pain experience. Task-based measures, such as the online Trier Social Stress Test and 30-second sit-to-stand task (30-s STS), are included as secondary metrics to facilitate participant comparison across standardized challenges. As understood through the biopsychosocial model of pain, multiple systems contribute to the pain experience. Therefore, it is vital to collect multiple layers of data for biosignature development. Machine learning methods are critical to the study design, facilitating an improved interpretation of multidimensional data and enhanced detection of nuanced fluctuations in pain status.

XSL•FO
RenderX

Figure 1. Visualization of third-party app user interface. Daily tasks are easy to identify and complete. Users can monitor their progress in the TRAC-Pain (Targeted Real-Time Assessment of Chronic Pain) study and visualize rewards when completing weekly adherence goals. The interface includes customizable options to view or hide smartwatch summary metrics.

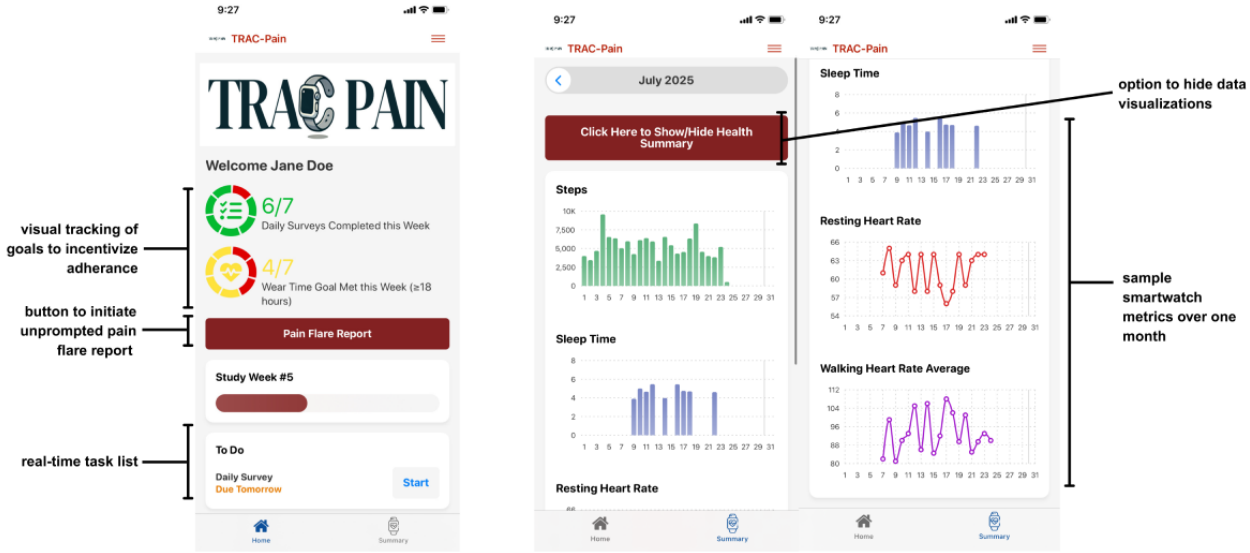

Participants and Setting

Up to 500 youth and adolescents with chronic primary or secondary MSK pain [21,22] will be enrolled in this study. Individuals are eligible to participate if they are between 14 and 24 years old and experience MSK pain (pain in muscles, tendons, bones, joints) in one or more anatomic regions that has persisted or recurred for ≥ 3 months and is associated with significant emotional distress or functional disability. Individuals are excluded if they report only headache, orofacial, or visceral pain, have been hospitalized in the past 30 days for something other than pain, have a significant cognitive impairment that would prevent study participation (eg, severe brain injury), are currently undergoing treatment for cancer, or are pregnant or expect to become pregnant over the study period. This study is listed at ClinicalTrials.gov (NCT06867757) and preregistered on Open Science Framework (10.17605/OSF.IO/EN69C).

Participants are recruited nationwide through a multipronged recruitment strategy to ensure a representative sample. Recruitment within the Stanford health care system occurs through the Research Participation and Engagement Program (RPEP) and Stanford Children's Health Pediatric Pain Management Clinic. Nationwide recruitment occurs through a custom-designed study website, community advertisements to pediatric multidisciplinary pain clinics in the United States, advertisements sent via postal mail, and patient advocacy networks (eg, US Pain Foundation, American Chronic Pain Association, and The Ehlers-Danlos Society). Interested participants are directed to complete a web-based study interest form as the first step of the eligibility screening. Potentially eligible participants are then contacted by the study team for additional screening based on the previously described eligibility criteria.

Ethical Considerations

The participants—and their legally authorized representatives, if applicable—complete an IRB-approved assent/consent process before completing any study activities. All study data, including digital health information, shared medical records, and video submissions, are collected, stored, and shared according to strict confidentiality requirements of the Stanford University IRB (IRB-77005).

Self-Report Measures (Primary)

The study design is outlined in Figure 2. Once enrolled, participants complete a battery of baseline surveys through the third-party app. Baseline surveys include self-report measures assessing pain and pain interference [23-27], fatigue [28,29], mood [30-32], stress [33], sleep [34], and physical activity [35] (Table 1).

XSL•FO
RenderX

Figure 2. Timeline of TRAC-Pain (Targeted Real-Time Assessment of Chronic Pain) study. The study begins with surveys completed at baseline. Participants wear a smartwatch for 12 weeks and complete "daily check-in" surveys about their pain experience. At discharge, participants are invited to complete a modified online Trier Social Stress Test and 30-second sit-to-stand task (30s STS). Finally, implementation outcomes are assessed through a semistructured exit interview to guide future applications. HEAL CDE: Helping End Addiction Long Term Common Data Elements; SDOH: Social Determinants of Health.

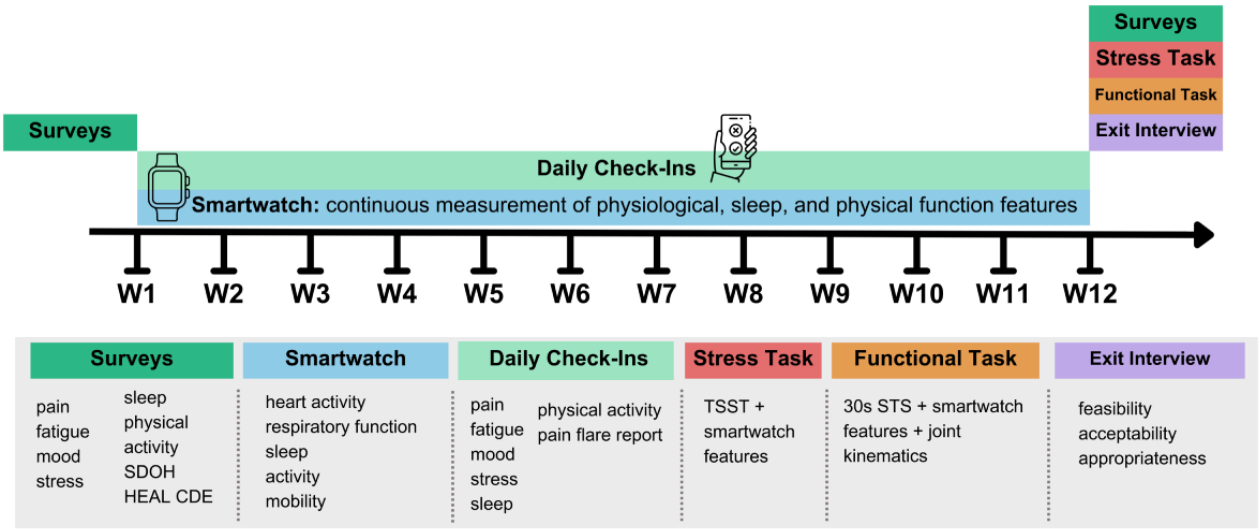

XSL•FO
RenderX

Table 1. Timeline of tests and measures for the Targeted Real-Time Assessment of Chronic Pain (TRAC-Pain) study.

Domain and tests/measures	Time points administered			
	Baseline	Daily	Continuous	Discharge
Candidate Monitoring Features (Apple Health Kit)				
Physiological				
Heart activity				
Heart rate			✔	
HRV[a]			✔	
Respiratory function				
Respiratory rate (sleep only)			✔	
Physical function				
Activity level				
Steps, bouts of activity, standing time			✔	
Mobility				
Gait speed, gait symmetry, step length, 6MWT[b]			✔	
Sleep				
Time in bed, time awake, sleep phases		✔		
Gold Standard self-report metrics of the pain experience				
Pain Interference				
PEG 3-item[c]	✔	✔		✔
CALI-9[d]	✔			✔
BPI-SF[e]	✔			✔
Pain				
PPST[f]	✔			✔
SF-MPQ-2[g]	✔			✔
NRS[h] pain	✔	✔		✔
Fatigue				
PROMIS[i]-fatigue	✔			✔
NRS fatigue	✔	✔		✔
Mood				
PHQ-2[j]	✔	✔		✔
GAD-2[k]	✔	✔		✔
PANAS-10[l]	✔	✔		✔
Stress				
PROMIS-stress	✔			✔
NRS stress	✔	✔		✔
Sleep				
ASWS-SF[m]	✔			✔
NRS sleep quality	✔	✔		✔
Physical activity				
PROMIS-Physical Function SF 8c 7-Day	✔			✔

XSL•FO
RenderX

Domain and tests/measures	Time points administered			
	Baseline	Daily	Continuous	Discharge
NRS activity level[n]	✔	✔		✔
Natural and task-based events				
Pain flare report				
Pain flare qualities, wellness behaviors (eg, deep breathing, pacing), daily events (eg, academic stressor, medical appointment)		✔	✔[o]	
Physiological response to stress task				
Self-report distress metrics, Apple Health Kit metrics collected during TSST-OL[p]				✔
Functional Task Performance				
Number of sit-to-stand-stand repetitions completed, joint kinematics, Apple Health Kit metrics				✔
Additional Variables				
Psychosocial, environmental, and social determinations of health				
PANES[q]	✔			
mMOS-SS-8[r]	✔			
ULS-8[s]	✔			
Everyday Discrimination Scale – short version	✔			
Health Care Discrimination Scale	✔			
Children's HealthWatch Hunger Vital Sign	✔			
HealthBegins upstream risk screening tool	✔			
PSS-10[t]	✔			
CHIS[u]-language spoken at home single-item	✔			
Demographics survey (eg, age, sex, race, gender)	✔			
Medical history				
Prior and current medical and pain history and treatments	✔			✔
HEAL CDE measures				
WHOQOL-2[v]	✔			✔
PCS-6[w]	✔			✔
NIDA-2[x]	✔			✔
PGIC[y]	✔			✔
Implementation outcomes				
Participant feedback				
TFA[z]				✔
NPS[aa]				✔
Exit interview				✔
Adherence metrics				

XSL•FO
RenderX

Domain and tests/measures	Time points administered			
	Baseline	Daily	Continuous	Discharge
Hours of smartwatch wear, % complete of surveys and check-ins				✔

[a]HRV: heart rate variability.
[b]6MWT: 6-minute walk test.
[c]PEG 3-item: Pain, Enjoyment of Life, and General Activity scale.
[d]CALI-9: Child Activity Limitations Interview.
[e]BPI-SF: Brief Pain Inventory – Short Form.
[f]PPST: Pediatric Pain Screening Tools.
[g]SF-MPQ-2: Short-Form McGill Pain Questionnaire-2.
[h]NRS: Numeric Rating Scale.
[i]PROMIS: Patient-Reported Outcomes Measurement Information System.
[j]PHQ-2: Patient Health Questionnaire-2 item.
[k]GAD-2: Generalized Anxiety Disorder-2 item.
[l]PANAS-10: Positive and Negative Affect Schedule.
[m]ASWS-SF: Adolescent Sleep Wake Scale–Short Form.
[n]Not a validated measure, adapted from NRS format (ie, "On a scale from 0 to 10, how would you rate your overall physical activity today?")
[o]Available to complete unprompted, retrospectively or in real time, outside of time-based daily check-in
[p]TSST-OL: Trier Social Stress Test, online/remote.
[q]PANES: Physical Activity Neighborhood Environment Scale.
[r]mMOS-SS-8: 8-item modified Medical Outcomes Study Social Support Survey.
[s]ULS-8: 8-item UCLA Loneliness Scale.
[t]PSS-10: 10-item Perceived Stress Scale.
[u]CHIS: California Health Interview Survey.
[v]WHOQOL-2: World Health Organization Quality of Life-2.
[w]PCS-6: Pain Catastrophizing Scale-6 item.
[x]NIDA-2: Level 2—Substance Use—Child Age 11–17 instrument from the National Institute on Drug Abuse.
[y]PGIC: Patient Global Impression of Change.
[z]TFA: Theoretical Framework of Acceptability.
[aa]NPS: Net Promoter Score.

Once the baseline surveys have been completed, the participants begin to receive "daily check-in" surveys, deployed at 5 pm (participant's local time) and available to complete until 11:59 pm, for the duration of the 12-week study. Participants receive daily prompts through the app interface to complete these "daily check-in" surveys, which include short versions of self-report measures to quantify daily pain interference, pain, fatigue, mood, stress, sleep, and physical activity (Table 1). The daily check-in also asks participants to report whether they had a pain flare that day ("yes/no") and to describe its characteristics. A pain-flare [10,11] is described to participants using the following language: "A temporary but noticeable worsening of your usual pain and/or symptoms (eg, fatigue, stiffness, and emotional distress) that feels more intense than what you typically deal with. Pain flares typically meet the following criteria: (1) are disruptive, (2) last longer than usual pain spikes, (3) affect more than just pain, (4) may be linked to a trigger, and (5) are bigger than normal pain changes." Participants also have the option to initiate an unprompted retrospective or real-time pain flare report outside of the time-based daily check-in. The daily check-in also includes a daily diary for participants to note any exciting or stressful daily events (eg, academic stress, medical appointment) and any pain wellness strategies used (eg, deep breathing, pacing, journaling). The daily self-report measures are used as the gold standard (ie, ground truth) to which novel physiological markers are compared and thus are essential to ensure the clinical relevance of the biosignature.

Smartwatch Metrics (Primary)

If participants do not have an iPhone compatible with iOS 14.1 or later and an Apple Watch compatible with iOS 15 or later, they are mailed study-provided devices (iPhone SE Gen 2, Apple Watch Series 9s) as needed. Once the devices are received, the participants are invited to enroll in the study through an app-based third-party digital health and research platform. Upon study discharge, the participants are required to return any borrowed devices. The following steps are implemented to encourage device return: (1) participants receive a prepaid return envelope and instructions for returning the device(s) upon study discharge; (2) participants explicitly agree to return device(s) as part of the consent process; and (3) frequent touch points with the study team are intended to promote good rapport and increase the likelihood of follow through on agreed returns.

The participants are instructed to wear the Apple Watch for at least 18 hours per day during the 12-week observation period. Apple Health Kit data are collected, uploaded, and stored through the third-party digital health and research platform. Metrics of interest include those related to respiratory function (eg, respiratory rate), heart activity (eg, heart rate, heart rate

XSL•FO
RenderX

variability), physical activity (eg, steps, bouts of activity, standing time), sleep (eg, in-bed time, sleep phases), and mobility (eg, gait speed, walking asymmetry, step length) (Table 1). Survey completion rate and adherence to required watch wear times are monitored by the study team. Participants receive notifications from the study team to encourage high levels of adherence based on completion standards selected a priori (eg, number of daily check-ins completed each week and daily and nightly hours of wear). Participants receive weekly monetary incentives for completing each component of the study, with up to US $210 total. Those who do not meet adherence standards may not be eligible to receive a given week's compensation. Furthermore, those consistently not meeting standards within the first few weeks of the study may be withdrawn in an effort to allocate research team time to adherent participants and prevent excessive amounts of missing data. At the end of the 12-week observation period, the participants complete an additional battery of discharge surveys, once again deployed through the third-party app (Table 1). They also complete a series of task-based measures prior to study discharge.

Task-Based Measures (Secondary)

Overview

The participants complete an additional stress task and a functional task to capture physiological responses to standardized challenges. These tasks are collected on a subset of the study cohort to examine responses to acute stressors. Although not required for the development of the initial biosignature, physiological responses and performance metrics related to these standardized challenge paradigms may improve the sensitivity of the biosignature in future iterations. Additionally, the results will provide unique insight into the pain experience by allowing for a standardized comparison across participants, supporting future hypothesis generation.

Online Trier Social Stress Test

A modified online Trier Social Stress Test (TSST) [36] is used to measure physiological responses to an evaluative stress test known to activate the hypothalamic-pituitary-adrenocortical (HPA) system. The HPA system plays an important role in the pain experience [37]. Although originally developed for in-person testing, an online version of TSST, known as the TSST-OL, has been validated to elicit HPA activation in adolescents [36].

In the TSST-OL, participants are first shown a calming video to induce a standard neutral to positive state. Then, they are given 5 minutes to prepare a speech, as if presenting to a new class. They are instructed to describe themselves to the class and why they would be liked by other students. Next, they have 5 minutes to deliver the prepared speech, via an online conferencing platform, to a panel of neutral or critical "judges." The judge panel will be comprised of study staff who have been trained to maintain a neutral to stern expression during the speech. After speech delivery, the participants are given another 5 minutes to complete a mental arithmetic task (ie, serial subtraction) in front of the same neutral to stern panel of judges. Participants then have a brief, unsupervised rest period prior to task debrief. Traditionally, saliva samples are collected at various points in this protocol to measure cortisol levels as a proxy for HPA activation. However, recording other physiological metrics (eg, heart rate) is more feasible in the online version of the TSST, as these metrics have been shown to fluctuate in response to the TSST [38]. Recording of physiological metrics during the TSST-OL is done via Apple Watch set to "workout" mode to ensure a high sampling rate of heart rate readings. Collecting Apple Health Kit metrics during the TSST-OL provides insights into the systemic response to emotional stressors for each participant. To mitigate psychological risk that may persist following the TSST-OL, the study team completes a participant debrief consistent with the validated TSST-OL protocol [36]. The purpose of the debrief is to further describe the reason for this task and to alleviate any participant concerns. The study's principal investigator (author LES) is a licensed clinical psychologist and is prepared to offer additional support as needed.

30s-STS Task

The 30s-STS is a measure of lower extremity strength, transitional movement mechanics, and functional exercise capacity [39,40], systems that may be impacted by chronic pain. The 30s-STS has not been validated in youth with primary chronic pain disorders. However, this task demonstrates strong psychometric properties for healthy young adults, with strong test-retest reliability (intraclass correlation coefficient=0.93) and convergent validity (Pearson r= –0.78 for the 5x STS and Pearson r=0.51 for the lateral step-down test) [40]. The 30s-STS is also useful in discriminating active versus inactive healthy young adults [40]. Therefore, the 30s-STS is expected to be an appropriate test to quantify functional performance and physiological response to physical challenge in this study population.

Participants receive explicit instructions for remote performance of the 30s-STS (video instructions in Multimedia Appendix 1). For this task, the participants are asked to perform repetitions of an STS movement for 30 seconds. They are instructed to use a sturdy chair and perform repetitions as quickly and safely as possible, while keeping their arms crossed over their chest. The participants are instructed to identify another person to act as a recorder and timekeeper for the task. This person will record the participants' performance and ensure the test is stopped at 30 seconds. Then, the participants upload the video via the secure third-party platform for the study team to access. Established norms are available for a subset of our expected cohort (ages 18 to 24 years) and are used to evaluate participant performance (ie, number of STS repetitions) [40]. Additionally, this study assesses innovative and nuanced metrics of performance through the development and implementation of a modified human pose-estimation algorithm to estimate kinematics (eg, trunk angle, joint angular velocity) [41]. Finally, as with the TSST-OL, participants are instructed to wear their Apple Watch set to the "workout" mode throughout the duration of the 30s-STS to monitor systemic response to a standardized physical performance test.

Exit Interview

Following completion of the 12-week protocol, baseline and discharge surveys, TSST-OL, and 30s-STS, the participants are

XSL•FO
RenderX

invited to complete a semistructured exit interview to garner feedback regarding the feasibility, acceptability, and appropriateness of the captured data. The participants also complete the Theoretical Framework of Acceptability (TFA) [42] and Net Promoter Score (NPS) [43] surveys. Feedback from the exit interview, TFA, and NPS is incorporated into the ongoing study procedures as appropriate and is particularly important for future applications of this work, specifically as a guide for implementation and validation of the digital biosignature (Phase 2).

Collaborations With People With Lived Experience

All aspects of this study were developed in partnership with people with lived experience, including patients, caregivers, and patient advocacy groups. These patient and advocacy partners were and continue to be involved in every aspect of the study, including grant preparation, development of recruitment materials, and selection of outcome measures. For example, based on feedback from patients and advocacy partners, the participant-facing app was modified to enable retrospective reporting of a pain flare, such as when a current pain flare is too severe to report in real time. Additionally, partner feedback prompted the research team to modify a short recruitment video that had previously focused on negative aspects of living with chronic MSK pain. The new video instead focuses on the positive impact and hope inspired by the study. Additionally, patient and advocacy partners noted that including daily visualizations of deterioration in smartwatch metrics via the app interface (eg, poor sleep, reduced activity) may exacerbate participant anxiety and hopelessness surrounding the likelihood of a pain flare. Based on this feedback, the study team modified the app interface to allow participants to hide unwanted data visualizations not required for the study. The participants can also opt out of any health app notifications or alerts for abnormal readings. Alternatively, the study team implemented positively framed visualizations and notifications through the visual tracking of smartwatch wear time goals and notifications for opportunities to implement participant-preferred wellness strategies. Importantly, patient and advocacy partners attended an in-person study launch meeting in December 2024 and continue to provide feedback on recruitment materials, outreach efforts, and research activities via virtual monthly team meetings.

Analysis Plan

Outcome Measures

A digital biosignature will be developed to enable the detection of a change in pain status. The primary outcome used to validate the biosignature (ground truth) will be the participant's self-reported onset of a pain flare. Additional models will be explored to predict secondary outcomes such as a clinically relevant change in pain interference, namely, 2 units on the Pain, Enjoyment of Life, and General Activity (PEG 3-item) scale or the onset of pain flares lasting for more than 2 days only. To improve pain flare prediction, participants may be stratified by demographic features known to impact the pain experience, and models may be adjusted using partial correlation. Accounting for individual factors may be particularly relevant when considering differences in digital access and literacy among participants residing across diverse geographic locations.

Data Reduction

Raw physiological signals (eg, heart activity, step counts, sleep stages) will be synchronized across modalities and aligned to a common timeline while preserving native sampling frequencies. To capture information across temporal scales, features will be engineered using a *multiresolution aggregation strategy*. For each signal, summary statistics will be computed over short windows (seconds to minutes) and further aggregated into intermediate (hourly) and daily summaries. All features derived at finer resolutions will be consolidated into a single daily feature vector, enabling the integration of short-term variability and longer-term physiological trends. The model prediction target and evaluation will be defined at a chosen time scale (eg, daily). Any windowing will only use past data relative to the prediction time to avoid leakage.

In addition to static summaries, *dynamic features* will be generated to characterize within-day temporal structure and behavior. These may include acceleration-based metrics (eg, mean acceleration), measures of temporal change (eg, slopes and rates of change), and circadian rhythm features. Complexity metrics (eg, entropy-based measures) may also be included to quantify signal regularity.

Handling Missing Data

Missing data due to nonwear or sensor dropout will be handled using a hierarchical approach. Short gaps will be interpolated when physiologically appropriate, whereas longer gaps will be left missing at the raw level. Daily features may be imputed using time-series–aware methods (eg, forward/backward filling within individuals). Missingness indicators will be included as additional covariates in model development to account for information loss.

Biosignature Development

Higher-level multivariate models will be developed, beginning with simpler traditional machine learning models before progressing to deep representation learning methods. Traditional machine learning algorithms will be trained using a nested k-fold cross-validation design. Model performance will be assessed using the area under the receiver-operator characteristic curve (AUROC), the area under the precision-recall curve (AUPRC), the F_1-score, Spearman Rho, and/or Mann-Whitney tests of prediction outputs. Additionally, deep learning algorithms will be trained using train/validate/test sets where model loss, AUROC, AUPRC, F_1-score, and/or Spearman Rho will be used to monitor differences between predicted outputs and true values across each epoch of training. Hyperparameter sweeps will be used to optimize the model's performance and reduce overfitting risks. Sweeps will include adjustments to hyperparameters such as model width, model depth, learning rate, optimizer selection, learning rate schedulers, and early stopping. Model interpretability will be aided by methods such as Shapley Additive Explanations (SHAP) to identify key features relevant for clinical interpretation and application. Additionally, Phase 2 of this study is designed for prospective validation in a novel

XSL•FO
RenderX

cohort, allowing for an assessment of the model's validity beyond standard performance metrics.

Sample Size Justification and Power Analysis

Given that the goal of this study is not to test a specific hypothesis, but rather to iterate over an entire data set to develop higher-level multivariate models, sample size and power analysis are less relevant for our study design. However, multiple strategies will be included to ensure data sufficiency. Collecting 12 weeks of daily self-reports and continuous smartwatch-derived features for up to 500 participants will enable multiresolution feature engineering while preserving participant-level separation during model evaluation (nested cross-validation, train-validate-test splits by participant). Outcome-event counts will be monitored to reduce overfitting risk. We expect to include approximately 35 Apple Watch–derived features (eg, physiological and physical function and sleep) in our primary model, alongside demographics (eg, age, sex, social determinants of health), pain-related features (eg, pain diagnoses, medical history), and task-based performance metrics (TSST-OL, 30s-STS) in secondary models. Model complexity controls, such as regularization, feature reduction, and early stopping, will be incorporated as needed. Furthermore, our sample of up to 500 participants will far exceed those found in comparable studies [20,44]. Achieving 80% of the target recruitment (ie, 400 participants) is also sufficient based on outcome-event counts and comparable studies.

Results

This study was funded in September 2024, with data collection beginning in March 2025. As of December 15, 2025, 190 participants have been enrolled, with data collection and analysis ongoing.

Discussion

Expected Findings

The purpose of this study is to develop a digital biosignature of the pain experience in youth and adolescents with chronic MSK pain. Self-report measures are the current gold standard for monitoring the pain experience and informing clinical decision-making. The purpose of developing a digital biosignature is to overcome the limitations of self-report, specifically the tendency to represent a single point in time and its inherent vulnerability to recall bias. A digital biosignature will provide real-time, objective measures of the pain experience through a holistic, multisystem lens.

This study is expected to result in several clinical benefits. First, the development and validation of digital clinical end points facilitates more robust assessment of novel therapeutics and treatment paradigms for youth with chronic MSK pain. The precise monitoring of treatment response would enhance the rigor and effectiveness of clinical trial design, leading to improved treatments, especially for those refractory to existing therapies. Second, a digital biosignature will give clinicians a deeper understanding of their patients' individual pain experience through real-time, longitudinal monitoring of symptom fluctuations. These insights will enhance clinical decision-making and support a shift toward individually tailored interventions. Real-time tracking will also facilitate expedited shifts in a prescribed treatment plan, saving time and resources in helping patients to find the best treatment option for their specific pain presentation. Additionally, an objective biosignature may improve patients' pain experience and interactions with clinicians by reducing the need to repeatedly recount severe pain events, an act that itself may be distressing. Finally, not only will a digital biosignature inform clinical practice, but it will also empower patients to have more control over their symptoms. Specifically, Phase 2 addresses this last point through enhanced "wellness alerts" for improved self-efficacy in living with chronic pain.

Future Directions (Phase 2)

To further the impact of the digital biosignature that will be developed in this study, we expect to conduct a second phase of the study. The goals of Phase 2 are twofold: (1) assess clinical validation of the digital biosignature in a novel sample and (2) evaluate the accuracy of a predictive "wellness alert" system.

Phase 2 includes recruitment of a novel cohort of up to 400 youth with MSK pain. Recruitment criteria will mirror those described in this study. To ensure relevance of the digital biosignature for a representative set of adolescents, this phase will have a particular focus on recruitment of youth experiencing health disparities (eg, rural youth with limited access to tertiary pain clinics and racial and ethnic minorities). Once again, a multipronged recruitment strategy will be implemented as described in this study. Stanford Medicine's Trial Innovation Network will provide additional support for the recruitment of youth and adolescents experiencing health disparities. Participants will use an iPhone and Apple Watch for 12 weeks to monitor physiological, sleep, and physical activity data. Devices will be provided as needed. Metrics such as AUROC, AUPRC, F_1-score, positive predictive value, and negative predictive value will be used in this study phase for prospective validation of the digital biosignature. As in Phase 1, we expect that participants' self-reported pain flare onset will be used as the primary outcome to which the digital biosignature is compared.

Participants enrolled in Phase 2 also have the opportunity to receive enhanced "wellness alerts," which would be triggered by a deterioration in the pain experience as detected through abnormal smartwatch readings. The goal of these "wellness alerts" is to prompt users toward proactive self-management strategies (eg, pacing, sleep hygiene) in an effort to prevent or reduce the severity of the predicted pain flare. Patient and advocacy partners recommended positive framing of these "wellness alerts" (ie, opportunities for enhanced wellness strategies) as opposed to "pain flare predictions," which may be negatively interpreted and lead to increased anxiety and hypervigilance. The accuracy of the enhanced "wellness alerts" will be assessed via comparison to self-reported pain flare onset through measures of accuracy, precision, and true positive and false negative rates. This "wellness alert" system is expected to empower patients by reducing feelings of helplessness surrounding unpredictable fluctuations in pain status. In addition, the alerts may serve as an additional tool to promote pain-related

XSL•FO
RenderX

self-efficacy, which may serve as a protective factor in youth with chronic pain [45].

Access and Dissemination

Although this study uses the Apple Smartwatch as a wearable technology of choice, the algorithms developed can be expanded across multiple monitoring options to promote broader dissemination and encourage improved access to study innovations. This study will introduce opportunities for open-access analysis tools that can be deployed across different wearable devices for improved dissemination. Furthermore, the identification of key physiological features may suggest that simpler, more economical fitness technologies (eg, pedometer or heart rate monitor chest strap) are sufficient for characterizing the pain experience in this population.

Conclusions

The TRAC-Pain study leverages the widespread adoption of consumer wearables to overcome limitations of self-report as the current gold standard for monitoring the pain experience in youth with chronic MSK pain. This study is unique in its multipronged recruitment strategy and intentional collaboration with patient and advocacy partners from the beginning of the study design process. Clinical impacts are expected to span multiple aspects of care, impacting researchers, clinicians, and patients. In terms of research, the development of digital clinical end points will enhance the rigor of clinical trials for the development of improved therapeutics. For clinicians, a longitudinal, real-time assessment of the pain experience will introduce opportunities for individually tailored treatments. Finally, this study is an important step toward promoting patient empowerment through opportunities for proactive self-management and enhanced self-efficacy in living with chronic pain.

Acknowledgments

Thank you to CareEvolution and MyDataHelps for providing the app-based digital platform used in this study. Some wording refinement and literature suggestions were included with assistance from the generative AI tool ChatGPT-4 and ChatGPT-5 by OpenAI. All intellectual content originates from the authors, and the authors have independently evaluated and interpreted all references in this manuscript. Thank you to Dipro Chakraborty and Natasha Harrison for their role as statistical and modeling consultants.

Funding

This work is supported by the National Institute of Neurological Disorders and Stroke (award number 1UG3NS139943), the National Institute of Arthritis and Musculoskeletal and Skin Diseases (award number K24AR078945 awarded to LES), and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (award number T90HD118400 awarded to AHB). The funder had no involvement in the study design, data collection, or writing of this manuscript.

Data Availability

Data generated under this study will be made available in the Open Science Framework Registry.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Link to 30-second sit-to-stand instructional video.
[DOCX File , 13 KB - resprot_v15i1e84781_app1.docx]

Multimedia Appendix 2
Peer review report report from the ZNS1 SRB-R (22) National Institute of Neurological Disorders and Stroke Special Emphasis Panel HEAL Initiative: Digital Biomarker (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 200 KB - resprot_v15i1e84781_app2.pdf]

References

1. Chambers CT, Dol J, Tutelman PR, Langley CL, Parker JA, Cormier BT, et al. The prevalence of chronic pain in children and adolescents: a systematic review update and meta-analysis. Pain 2024 Oct 01;165(10):2215-2234. [doi: 10.1097/j.pain.0000000000003267] [Medline: 38743558]
2. Wrona SK, Melnyk BM, Hoying J. Chronic pain and mental health co-morbidity in adolescents: an urgent call for assessment and evidence-based intervention. Pain Manag Nurs 2021 Jun;22(3):252-259. [doi: 10.1016/j.pmn.2020.12.004] [Medline: 33454204]

XSL•FO
RenderX

3. Jastrowski Mano KE. School anxiety in children and adolescents with chronic pain. Pain Res Manag 2017;2017:8328174-8328179 [FREE Full text] [doi: 10.1155/2017/8328174] [Medline: 29081682]
4. Miró J, Roman-Juan J, Sánchez-Rodríguez E, Solé E, Castarlenas E, Jensen MP. Chronic pain and high impact chronic pain in children and adolescents: a cross-sectional study. J Pain 2023 May;24(5):812-823 [FREE Full text] [doi: 10.1016/j.jpain.2022.12.007] [Medline: 36577459]
5. Wilson AC, Palermo TM. Physical activity and function in adolescents with chronic pain: a controlled study using actigraphy. J Pain 2012 Feb;13(2):121-130 [FREE Full text] [doi: 10.1016/j.jpain.2011.08.008] [Medline: 22099608]
6. Palermo TM, Law E, Zhou C, Holley A, Logan D, Tai G. Trajectories of change during a randomized controlled trial of internet-delivered psychological treatment for adolescent chronic pain: how does change in pain and function relate? Pain 2015 Apr;156(4):626-634 [FREE Full text] [doi: 10.1097/01.j.pain.0000460355.17246.6c] [Medline: 25775441]
7. Simons LE, Sieberg C, Pielech M, Conroy C, Logan D. What does it take? Comparing intensive rehabilitation to outpatient treatment for children with significant pain-related disability. J Pediatr Psychol 2013 Mar;38(2):213-223 [FREE Full text] [doi: 10.1093/jpepsy/jss109] [Medline: 23104827]
8. Simons LE, Sieberg CB, Conroy C, Randall ET, Shulman J, Borsook D, et al. Children with chronic pain: response trajectories after intensive pain rehabilitation treatment with chronic pain: response trajectories after intensive pain rehabilitation treatment. J Pain 2018 Feb;19(2):207-218 [FREE Full text] [doi: 10.1016/j.jpain.2017.10.005] [Medline: 29102693]
9. Kashikar-Zuck S, Cunningham N, Sil S, Bromberg MH, Lynch-Jordan AM, Strotman D, et al. Long-term outcomes of adolescents with juvenile-onset fibromyalgia in early adulthood. Pediatrics 2014 Mar;133(3):e592-e600 [FREE Full text] [doi: 10.1542/peds.2013-2220] [Medline: 24567017]
10. Khanom S, McDonagh JE, Briggs M, Bakir E, McBeth J. Adolescents' experiences of fluctuating pain in musculoskeletal disorders: a qualitative systematic review and thematic synthesis. BMC Musculoskelet Disord 2020 Oct 02;21(1):645 [FREE Full text] [doi: 10.1186/s12891-020-03627-1] [Medline: 33008357]
11. Khanom S, McDonagh JE, Briggs M, McBeth J. Characterizing pain flares in adolescent inflammatory and non‐inflammatory musculoskeletal disorders: A qualitative study using an interpretative phenomenological approach. Eur J Pain 2020 Jul 19;24(9):1785-1796. [doi: 10.1002/ejp.1626]
12. Palermo TM, Walco GA, Paladhi UR, Birnie KA, Crombez G, de la Vega R, et al. Core outcome set for pediatric chronic pain clinical trials: results from a Delphi poll and consensus meeting. Pain 2021 Oct 01;162(10):2539-2547 [FREE Full text] [doi: 10.1097/j.pain.0000000000002241] [Medline: 33625074]
13. Coventry J, Pacey V, Smith M, Williams CM, Ta B, Sturgiss E. How children and adolescents with chronic pain describe their pain experiences: A qualitative systematic review. Patient Educ Couns 2025 May;134:108634. [doi: 10.1016/j.pec.2025.108634]
14. Hess CW, Rosen MA, Simons LE. Looking inward to improve pediatric chronic pain outcomes: a call for team science research. Pain 2022 Nov 30;164(4):690-697. [doi: 10.1097/j.pain.0000000000002836]
15. Marra C, Chen JL, Coravos A, Stern AD. Quantifying the use of connected digital products in clinical research. npj Digit Med 2020 Apr 03;3(1). [doi: 10.1038/s41746-020-0259-x]
16. Witt DR, Kellogg RA, Snyder MP, Dunn J. Windows into human health through wearables data analytics. Curr Opin Biomed Eng 2019 Mar;9:28-46. [doi: 10.1016/j.cobme.2019.01.001]
17. Ho D, Quake SR, McCabe ER, Chng WJ, Chow EK, Ding X, et al. Enabling technologies for personalized and precision medicine. Trends Biotechnol 2020 May;38(5):497-518 [FREE Full text] [doi: 10.1016/j.tibtech.2019.12.021] [Medline: 31980301]
18. Babu M, Lautman Z, Lin X, Sobota MHB, Snyder MP. Wearable devices: implications for precision medicine and the future of health care. Annu Rev Med 2024 Jan 29;75:401-415 [FREE Full text] [doi: 10.1146/annurev-med-052422-020437] [Medline: 37983384]
19. Ravindra NG, Espinosa C, Berson E, Phongpreecha T, Zhao P, Becker M, et al. Deep representation learning identifies associations between physical activity and sleep patterns during pregnancy and prematurity. npj Digit Med 2023 Sep 28;6(1). [doi: 10.1038/s41746-023-00911-x]
20. McGinnis EW, Lunna S, Berman I, Loftness BC, Bagdon S, Danforth C, et al. Discovering digital biomarkers of panic attack risk in consumer wearables data. medRxiv Preprint published online on March 6, 2022 [FREE Full text] [doi: 10.1101/2023.03.01.23286647] [Medline: 36909613]
21. Nicholas M, Vlaeyen JWS, Rief W, Barke A, Aziz Q, Benoliel R, IASP Taskforce for the Classification of Chronic Pain. The IASP classification of chronic pain for ICD-11: chronic primary pain. Pain 2019 Jan;160(1):28-37. [doi: 10.1097/j.pain.0000000000001390] [Medline: 30586068]
22. Perrot S, Cohen M, Barke A, Korwisi B, Rief W, Treede R. The IASP classification of chronic pain for ICD-11: chronic secondary musculoskeletal pain. Pain 2019;160(1):77-82. [doi: 10.1097/j.pain.0000000000001389]
23. Krebs EE, Lorenz KA, Bair MJ, Damush TM, Wu J, Sutherland JM, et al. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. J Gen Intern Med 2009 May 6;24(6):733-738. [doi: 10.1007/s11606-009-0981-1]
24. Holley AL, Zhou C, Wilson AC, Hainsworth K, Palermo TM. The CALI-9: A brief measure for assessing activity limitations in children and adolescents with chronic pain. Pain 2017 Sep 18;159(1):48-56. [doi: 10.1097/j.pain.0000000000001063]

XSL•FO
RenderX

25. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Rehabil Oncol 1995;13(1):29-30. [doi: 10.1097/01893697-199513010-00022]
26. Simons LE, Smith A, Ibagon C, Coakley R, Logan DE, Schechter N, et al. Pediatric Pain Screening Tool: rapid identification of risk in youth with pain complaints. Pain 2015 Aug;156(8):1511-1518 [FREE Full text] [doi: 10.1097/j.pain.0000000000000199] [Medline: 25906349]
27. Melzack R. The short-form McGill Pain questionnaire. Pain 1987 Aug;30(2):191-197. [doi: 10.1016/0304-3959(87)91074-8] [Medline: 3670870]
28. Lai J, Stucky BD, Thissen D, Varni JW, DeWitt EM, Irwin DE, et al. Development and psychometric properties of the PROMIS® pediatric fatigue item banks. Qual Life Res 2013 Feb 2;22(9):2417-2427. [doi: 10.1007/s11136-013-0357-1]
29. Quinn H, Thissen D, Liu Y, Magnus B, Lai J, Amtmann D, et al. Using item response theory to enrich and expand the PROMIS® pediatric self report banks. Health Qual Life Outcomes 2014 Oct 25;12(1). [doi: 10.1186/s12955-014-0160-x]
30. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
31. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
32. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037/0022-3514.54.6.1063]
33. Bevans KB, Gardner W, Pajer K, Riley AW, Forrest CB. Qualitative development of the PROMIS(R) pediatric stress response item banks. J Pediatr Psychol 2012 Nov 03;38(2):173-191. [doi: 10.1093/jpepsy/jss107]
34. LeBourgeois MK, Giannotti F, Cortesi F, Wolfson AR, Harsh J. The relationship between reported sleep quality and sleep hygiene in Italian and American adolescents. Pediatrics 2005 Jan;115(1 Suppl):257-265 [FREE Full text] [doi: 10.1542/peds.2004-0815H] [Medline: 15866860]
35. Tucker C, Bevans K, Teneralli R, Smith A, Bowles H, Forrest C. Self-reported pediatric measures of physical activity, sedentary behavior, and strength impact for PROMIS: item development. Pediatr Phys Ther 2014;26(4):385-392. [doi: 10.1097/pep.0000000000000074]
36. Gunnar MR, Reid BM, Donzella B, Miller ZR, Gardow S, Tsakonas NC, et al. Validation of an online version of the Trier Social Stress Test in a study of adolescents. Psychoneuroendocrinology 2021 Mar;125:105111. [doi: 10.1016/j.psyneuen.2020.105111]
37. Wyns A, Hendrix J, Lahousse A, De Bruyne E, Nijs J, Godderis L, et al. The biology of stress intolerance in patients with chronic pain—state of the art and future directions. J Clin Med 2023 Mar 14;12(6):2245. [doi: 10.3390/jcm12062245]
38. DuPont CM, Pressman SD, Reed RG, Manuck SB, Marsland AL, Gianaros PJ. An online Trier social stress paradigm to evoke affective and cardiovascular responses. Psychophysiology 2022 Apr 16;59(10). [doi: 10.1111/psyp.14067]
39. Jones CJ, Rikli RE, Beam WC. A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. Res Q Exerc Sport 1999 Jun;70(2):113-119. [doi: 10.1080/02701367.1999.10608028]
40. Lein DH, Alotaibi M, Almutairi M, Singh H. Normative reference values and validity for the 30-second chair-stand test in healthy young adults. Int J Sports Phys Ther 2022;17(5):907-914 [FREE Full text] [doi: 10.26603/001c.36432] [Medline: 35949374]
41. Boswell MA, Kidziński M, Hicks JL, Uhlrich SD, Falisse A, Delp SL. Smartphone videos of the sit-to-stand test predict osteoarthritis and health outcomes in a nationwide study. npj Digit Med 2023 Mar 04;6(1). [doi: 10.1038/s41746-023-00775-1]
42. Sekhon M, Cartwright M, Francis JJ. Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. BMC Health Serv Res 2022 Mar 01;22(1). [doi: 10.1186/s12913-022-07577-3]
43. Reichheld F. The one number you need to grow. Harvard Business Review. 2003. URL: https://analyticsconsultores.com.mx/wp-content/uploads/2019/03/The-One-Number-You-Need-to-Grow-Frederick-F.-Reichheld-HBR-2003.pdf [accessed 2026-01-29]
44. Dudarev V, Barral O, Radaeva M, Davis G, Enns JT. Night time heart rate predicts next-day pain in fibromyalgia and primary back pain. Pain Rep 2024 Apr 1;9(2):e1119. [doi: 10.1097/pr9.0000000000001119]
45. Cousins LA, Kalapurakkel S, Cohen LL, Simons LE. Topical review: resilience resources and mechanisms in pediatric chronic pain. J Pediatr Psychol 2015 Oct;40(9):840-845 [FREE Full text] [doi: 10.1093/jpepsy/jsv037] [Medline: 25979085]

Abbreviations

30s-STS: 30-second sit-to-stand task
AUPRC: area under the precision-recall curve
AUROC: area under the receiver-operator characteristic curve
HPA: hypothalamic-pituitary-adrenocortical
IRB: institutional review board
MSK: musculoskeletal
NPS: Net Promoter Score
PEG 3-item: Pain, Enjoyment of Life, and General Activity Scale

XSL•FO
RenderX

SHAP: Shapley Additive Explanations
TFA: Theoretical Framework of Acceptability
TRAC-Pain: Targeted Real-time Assessment of Chronic Pain
TSST: Trier Social Stress Test
TSST-OL: Trier Social Stress Test, online/remote

Edited by J Sarvestan; submitted 29.Sep.2025; this is a non–peer-reviewed article;accepted 23.Dec.2025; published 06.Apr.2026.

Please cite as:
Bailes AH, Shu CH, Chang A, Masood S, Jehl N, Davis A, Giberson J, Cashman C, Hill A, Gill J, McGinnis RS, McGinnis EW, Aghaeepour N, Simons LE
Targeted Real-Time Assessment of Chronic Pain (TRAC-Pain) in Youth: Protocol for a Digital Biosignature Development Through a Prospective Observational Cohort Study
JMIR Res Protoc 2026;15:e84781
URL: https://www.researchprotocols.org/2026/1/e84781
doi:10.2196/84781
PMID:

©Anna H Bailes, Chi-Hung Shu, Alan Chang, Sahrish Masood, Nicole Jehl, Aliyah Davis, Jeremy Giberson, Casey Cashman, Allison Hill, Javed Gill, Ryan S McGinnis, Ellen W McGinnis, Nima Aghaeepour, Laura E Simons. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 06.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Combating Dermatologic Misinformation on Social Media via a Training Program for Health Care Professionals: Protocol for a 2-Phase Pilot Mixed Methods Study

Aliyyat Afolabi[1], BS; Elijah Brown[1], BA; Joanne Chen Lyu[2,3], PhD; Pamela M Ling[4], MD, MPH; Andrea M Rustad[5], MD; Anna E Davis[6], MD; Beth L Hoffman[7], MPH, PhD; Melissa Pugliano-Mauro[6], MD

[1]School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[2]TSET Health Promotion Research Center, Stephenson Cancer Center, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[3]Department of Family and Preventive Medicine, College of Medicine, University of Oklahoma Health Campus, Oklahoma City, OK, United States

[4]Center for Tobacco Control Research and Education and Division of General Internal Medicine, University of California, San Francisco, San Francisco, CA, United States

[5]Department of Dermatology, University of Wisconsin, Madison, WI, United States

[6]Department of Dermatology, University of Pittsburgh Medical Center, Pittsburgh, PA, United States

[7]Department of Behavioral and Community Health Science, University of Pittsburgh School of Public Health, 130 De Soto St, Pittsburgh, PA, United States

Corresponding Author:
Beth L Hoffman, MPH, PhD
Department of Behavioral and Community Health Science, University of Pittsburgh School of Public Health, 130 De Soto St, Pittsburgh, PA, United States

Abstract

Background: The rise of social media has drastically altered the landscape of health information dissemination. While these platforms can promote health literacy, they are also rife with skin care misinformation, often spread by nonexperts. This has led to harmful consequences, including improper product use and delayed treatment. Despite their potential to mitigate misinformation, dermatologists and health care professionals remain underrepresented on platforms such as TikTok and Instagram.

Objective: This protocol describes a pilot study designed to evaluate the feasibility, acceptability, and preliminary efficacy of a social media training program aimed at equipping dermatology-focused health care professionals and students with skills to create accurate, engaging educational content.

Methods: This sequential mixed methods study will be conducted at the University of Pittsburgh. Phase 1 consists of a 5-module online training program completed by 50 health care professionals and students in dermatology or dermatology-adjacent fields. Pretraining and posttraining surveys will be conducted to assess changes in participants' knowledge, confidence, intent to engage in public-facing dermatology education, and satisfaction with the training (posttraining survey only). Phase 2 will involve the selection of 12 participants from phase 1 to apply the training by creating and posting short-form dermatologic educational videos on a study-managed social media channel. Exploratory outcomes include social media engagement metrics, qualitative content analysis of public comments, and individual qualitative interviews with participants to assess audience response and potential knowledge impact.

Results: Recruitment for phase 1 of the study is currently underway. As of January 2026, a total of 46 out of the target 50 participants have been successfully enrolled. Data collection for phase 1 is expected to be completed by the end of April 2026. Data analysis will begin shortly thereafter, with final analyses anticipated to be completed by September 2026. Recruitment for phase 2 is expected to begin in April 2026, with the goal of concluding data collection and analysis by the end of 2028.

Conclusions: This pilot study will inform the feasibility and refinement of social media training interventions designed to empower health care professionals to engage in evidence-based dermatologic health communication online.

International Registered Report Identifier (IRRID): DERR1-10.2196/81003

(JMIR Res Protoc 2026;15:e81003) doi:10.2196/81003

KEYWORDS

dermatology; social media; misinformation; health communication; medical education; TikTok; Instagram; digital literacy

XSL•FO
RenderX

Introduction

In today's digital age, individuals have unprecedented access to skin care information through various online platforms. While this democratization of information has the potential to enhance public knowledge, it has also led to the widespread dissemination of misinformation [1-4]. Social media platforms are inundated with skin care recommendations and quick fixes promoted by individuals lacking formal dermatological training [2,5,6]. One study found that 38% of popular dermatology-related Instagram accounts are managed by health care professionals and just 4% of these are operated by board-certified dermatologists [1,7,8]. This imbalance has facilitated the circulation of unsupported and potentially harmful skin care practices, contributing to widespread misunderstanding of basic dermatological principles and associated negative health outcomes [9]. For example, a case review by Szeto et al [8] that looked at social media in dermatology described patients who adopted harmful skin care practices after encountering inaccurate dermatologic content on social media, including improper chemical peels, excessive exfoliation, and unverified acne treatment [8,10]. These practices resulted in adverse effects such as skin irritation, hyperpigmentation, and worsened acne. Additional research has shown that viral skin care trends on platforms such as TikTok and Instagram have encouraged unsafe behaviors, including the misuse of retinoids, inappropriate chemical exfoliation, and self-diagnosis of conditions [11].

However, when used appropriately, social media offers substantial potential benefits, including interphysician collaboration, patient education, and public health promotion [12]. Prior studies demonstrate that health care professionals' active engagement on social media can improve public understanding of medical topics and help counter the spread of misinformation [12,13].

Despite this potential, dermatologists remain significantly underrepresented in digital spaces compared with nonexpert influencers, creating a persistent gap in accessible, evidence-based skin care guidance [14].

To address this gap, we developed a series of online social media training modules designed for dermatologists, health care professionals in dermatology-adjacent fields, and health professional students [15]. These modules aim to equip participants with the skills and confidence to create engaging, short-form videos that promote evidence-based dermatologic information. Through instruction in storytelling, content structure, production, and platform-specific optimization strategies, the training empowers participants to actively take part in shaping the digital narrative around dermatologic health. The modules were developed in collaboration with a social media expert and consist of concise instructional videos followed by interactive questions to reinforce learning. The overarching goal is to prepare trusted health care professionals to serve as credible digital voices, counter misinformation, and expand access to accurate dermatologic education through independently created content.

This manuscript describes the development of these training modules and presents the protocol for a 2-phase pilot research study. Phase 1 evaluates the acceptability and preliminary efficacy of the training in improving participants' knowledge, confidence, and intent to use social media for public dermatologic education. Phase 2 extends this work by examining how selected participants apply the acquired skills to create and disseminate dermatology-focused educational videos on social media platforms. This phase explores video reach and engagement, as well as how the content creation process influences participants' own confidence and willingness to engage in public health messaging.

While the training program is designed to equip health care professionals and students with practical skills for effective digital communication, the study aims to evaluate the feasibility, acceptability, and preliminary efficacy of the intervention (phase 1) and explore the real-world application of these skills through participant-generated content and audience engagement (phase 2).

Methods

Overview

This study uses a sequential mixed methods design consisting of 2 interrelated phases (Figure 1). To enhance transparency and reproducibility, reporting of the study design, participant flow, and outcome measures will follow the TIDieR (Template for Intervention Description and Replication) checklist, ensuring detailed description of the intervention content, delivery, and evaluation methods (see Multimedia Appendix 1 for the full checklist).

Figure 1. Sequential mixed methods study design. Phase 1 assesses the feasibility, acceptability, and preliminary efficacy of a social media training intervention using pretraining and posttraining surveys. Phase 2 examines the real-world application of training skills using participant-generated content, social media engagement metrics, and semistructured interviews.



XSL•FO
RenderX

Phase 1 Study Design

Phase 1 uses a quantitative, collaborative design for piloting an online training program. The training focuses on teaching physicians and health professional students in dermatology, or fields in which dermatological conditions are commonly seen (eg, pediatrics), how to create engaging social media videos on different popular platforms such as Instagram, TikTok, and YouTube. The study will be conducted at the University of Pittsburgh Medical Center (UPMC) and the University of Pittsburgh. A total of 50 participants will be enrolled in the training program on the Qualtrics platform (Qualtrics International Inc). The modules are hosted on Qualtrics. This platform was selected because it allows for seamless integration of video-based training content with embedded knowledge checks and surveys; supports conditional logic for progressive learning; and enables real-time data collection to evaluate participant engagement, knowledge, and confidence. Qualtrics is also widely used in academic and health care settings, making it well suited for research-based educational interventions. The training consists of 5 short instructional video modules (approximately 8 - 12 minutes each), each followed by a brief quiz containing 2 to 3 questions to reinforce learning and ensure engagement. Phase 1 will be focused on evaluating the training's impact on participants' confidence and knowledge of and engagement with dermatologic content creation, which will be assessed via pre- and posttraining surveys; the postsurvey will also include items to measure satisfaction with the program (Table 1). A sample size of 50 participants was selected to exceed the recommended minimum of 30 participants for pilot studies assessing questionnaire reliability and preliminary efficacy. This approach accounts for anticipated nonresponse and incomplete data and aligns with published methodological guidance for pilot research as outlined by Bujang et al [16]. Participants who withdraw prior to completing phase 1 will not be replaced, consistent with the exploratory nature of a pilot study. A pretest-posttest single-group design was selected due to the pilot nature of the study and its focus on assessing within-participant changes following the intervention. Outcomes including knowledge, confidence, and intent to engage in digital dermatologic education were selected as proximal indicators of participants' readiness to apply training skills.

Table . Phase 1 measures and outcomes assessed in the pre- and posttraining surveys.

Domain	Measure type	Item or description	Scale or response format	Time point
Demographics	Self-report	Age, ethnicity, educational level, student or professional status, field role, and experience	Multiple choice	Baseline
Prior experience	Self-report	"Have you ever created or been involved in creating educational content on social media?"	"Yes," "No," or "Prefer not to say"	Baseline
Knowledge	Quiz	10 multiple-choice questions on video structure, pacing, ethics, and engagement, among other topics	Multiple choice (correct or incorrect)	Baseline and after the training
Satisfaction	Self-report	Satisfaction with the information presented, ability to complete the modules independently, time required, and overall experience	7-point Likert scale (1="completely dissatisfied"; 7="completely satisfied")	After the training

Phase 2 Study Design

Following completion of phase 1, a total of 12 participants will be selected to take part in phase 2 to demonstrate the application of training skills and explore the real-world dissemination of dermatologic educational content. Purposive sampling will be used to ensure diversity in training level, professional background, and engagement with the training modules. The final sample size of 12 participants was determined based on our goal of achieving thematic saturation with individual qualitative interviews as qualitative research guidance indicates that 9 to 17 interviews are typically sufficient to achieve thematic saturation [17].

In this phase, participants will be asked to create at least 3 short-form educational videos over a 1-month period. These videos will be posted to our study-managed social media accounts (Instagram, TikTok, and YouTube) under the Pitt channel, with the goal of correcting common skin care misconceptions using the principles taught in the training modules.

Participants will generate their own scripts using reliable dermatologic information sourced from the American Academy of Dermatology's "Stats and facts" web page as it is considered a trusted source by board-certified dermatologists and is affiliated with the world's largest dermatology organization [18]. All video scripts will be reviewed by the principal investigator, who is also a board-certified dermatologist and Mohs surgeon, in addition to the coauthor, who is a dermatology resident, to ensure medical accuracy prior to posting. Rather than assigning specific pieces of misinformation to rebut, participants will identify commonly encountered dermatologic myths or knowledge gaps based on their clinical experience and training. They will then create original, engaging short-form

XSL•FO
RenderX

videos to educate the public and improve dermatologic health literacy.

To assess digital impact, we will collect quantitative engagement metrics for each video, including likes, comments, shares, saves, views, and engagement rate ([likes + comments + shares + saves]/views) [19]. Because the Pitt study channel is newly created, we expect the initial benchmarks to be zero, and growth over time will serve as an exploratory indicator of how participants applied the training intervention in a real-world digital environment.

We will conduct qualitative content analysis of the videos created and the public comments left on these videos to evaluate how participants select their topics, present themselves, and use techniques to create engaging content. We will also assess how closely the videos align with the training guidance, as well as audience sentiment, perceived credibility, and engagement with dermatologic concepts. This qualitative analysis will complement the quantitative metrics and provide deeper insights into audience knowledge and attitudes in response to the videos.

Finally, we will conduct individual qualitative interviews with all 12 participants to explore their content creation experience, perceived barriers to digital health communication, confidence in public engagement, and perceived benefits of participating in this intervention. Interviews will be conducted by trained research assistants using a semistructured interview guide.

Outcome Measures

For phase 1, the primary outcomes are feasibility, acceptability, and preliminary efficacy of the social media training modules. Feasibility will be assessed using module completion rates and participant retention from the pretraining to posttraining surveys. Acceptability will be evaluated using posttraining survey items on satisfaction with the information presented, ability to complete the modules independently, and perceived relevance of the material using a 7-point Likert scale (1="completely dissatisfied"; 7="completely satisfied"). Preliminary efficacy will be assessed through knowledge gain on 10 multiple-choice questions aligned with the training content, as well as participant-reported confidence and intent to engage in digital dermatologic education. Exploratory measures include participant engagement with each module, such as quiz completion and time spent on each module.

For phase 2, outcomes focus on the real-world application of training skills and the potential digital impact of participant-generated content. Quantitative engagement metrics for each video posted to the study-managed social media accounts (Instagram, TikTok, and YouTube) will include views, likes, shares, comments, saves, and engagement rate ([likes + comments + shares]/views). Content fidelity will be assessed by comparing participant videos to the training guidance, with all scripts reviewed for medical accuracy by study investigators. Qualitative outcomes will be collected through semistructured interviews exploring participants' experiences with content creation, perceived barriers and facilitators, and confidence in public health communication. Additionally, a qualitative content analysis of public comments on the videos will assess audience engagement, sentiment, perceived credibility, and evidence of knowledge uptake or correction of misinformation. Viewer-level changes in knowledge, attitudes, or behavior are beyond the scope of this pilot study but will be considered in future research.

Participants and Inclusion and Exclusion Criteria

For phase 1, a total of 50 participants from the University of Pittsburgh and UPMC will be enrolled in the study. Participants will be eligible if they are physicians (MD or DO), physician associates (PA), clinical registered nurses (RN and BSN), or nurse practitioners (NP) in dermatology or dermatology-adjacent fields such as pediatrics, rheumatology, family medicine, and allergy or students pursuing degrees in these fields (eg, medical, nursing, and physician associates) interested in dermatology or dermatology-adjacent fields.

Exclusion criteria include health students without an expressed interest in dermatology and health professionals whose patient population with dermatological conditions is rare (eg, general surgery, orthopedics, or psychiatry).

Interest in dermatology will be assessed via participant self-report. We will use the following single categorical question to assess professional background and interest in dermatology: "Are you currently a student interested in pursuing a career in the field of dermatology or a health professional working in the field of dermatology or a related health field?" Responses will include four options: (1) student, (2) working in the field or field-related (ie, conducting dermatological research), (3) both (student and working in the field), and (4) neither. Participants selecting options 1, 2, or 3 will be required to specify their role. Students in medical, nursing, or physician associate programs and health care professionals with MD, DO, PA, or NP degrees in rheumatology, allergy, pediatrics, and family medicine will be included in the study.

For phase 2, participants will be eligible if they have completed phase 1 of the study and have access to a smartphone to create videos. If requested, study staff will provide a tripod (valued at US $20) to support video creation.

Recruitment

A multimodal approach will be used to recruit participants for phase 1. Medical student recruitment will be facilitated through collaboration with the University of Pittsburgh Office of Medical Education using direct communication channels (emails and SMS text messages) to ensure comprehensive outreach across academic years, specialty interests, and backgrounds. Recruitment will be coordinated with the health sciences library, which will distribute mass email invitations detailing study objectives, eligibility criteria, and instructions for participation. To recruit practicing health professionals, informational flyers will be distributed at UPMC Children's Hospital of Pittsburgh and adult dermatology clinics affiliated with UPMC. These sites have been selected for their high patient and clinician traffic, allowing for targeted outreach to relevant health care providers.

For phase 2, participants from phase 1 will be sent an email containing a flyer with details about phase 2 of the study. From those who express interest, a purposive sampling approach will be used to select 12 individuals to ensure diversity in training

XSL•FO
RenderX

level, professional background, and engagement with the training modules. Selection will be based on expressed interest and diversity considerations and will not be based on quiz performance, response time, or other performance-based metrics. Selected participants will then be contacted via email with a formal invitation to take part in phase 2, along with a consent form outlining expectations, time commitment, and next steps.

Module Development

Each module was developed by the primary author, who has experience in creating health-related content for social media platforms, and was reviewed by the principal investigators. Prior to development, the first author conducted a content analysis of high-performing videos across TikTok, Instagram, and YouTube to identify strategies for effective audience engagement, including storytelling techniques, visual presentation, and commonly used production equipment. These insights, combined with personal experience in content creation, informed the scripting and design of the modules to ensure that they were concise, engaging, and appropriate for social media dissemination.

The intervention consists of 5 short, interactive modules designed to empower students and health professionals to create engaging and culturally sensitive short-form videos for social media. These modules are led by a medical student with an interest in dermatology and experience in communication and content creation. Each module builds on the last to provide a comprehensive foundation in video creation, health storytelling, ethical considerations, and platform-specific optimization.

The first module, "Fundamentals of Short-Form Video Creation," introduces participants to the core principles of effective short-form videos. It covers content structure (opening, message delivery, and call to action); pacing techniques to retain viewer attention; and strategies for audience engagement, such as using prompts and captions. Visual examples are included to demonstrate each concept (Figure 2). The second module, "Storytelling Techniques for Dermatology," emphasizes the role of narrative in health communication. Participants learn how to use relatable patient stories, show treatment journeys, highlight universal themes (eg, empowerment), and end videos with actionable health messages. The importance of authenticity and obtaining patient consent when using real testimonials is also covered.

Figure 2. Screenshot of one of the modules, titled "Creating Viral Videos to Combat Skincare Misinformation." A QR code is included for direct access to the video module.

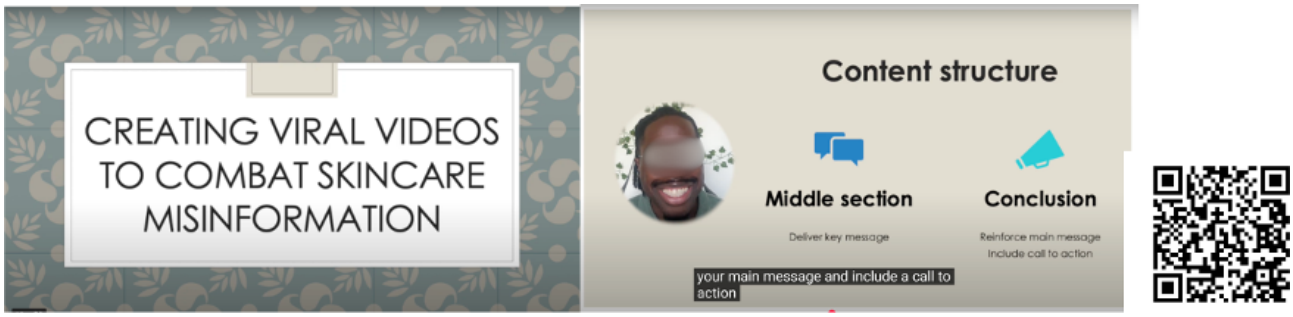

The third module, "Video Production Tools and Techniques," provides an overview of essential production elements, including lighting, sound, editing software, and camera composition. Participants are introduced to accessible tools (eg, video editing, ring lights, and lavalier mics) and guided through the video editing process from raw footage to finalized content.

The fourth module, "Ethical and Culturally Sensitive Practices," addresses ethical video creation with a focus on diversity, inclusion, and respect for cultural practices. Topics include using inclusive language, avoiding stereotypes, honoring traditional practices, obtaining informed consent, and using content to highlight disparities in dermatologic care.

The last module, "Optimizing Videos for Social Media Platforms," guides participants in tailoring content for specific platforms: Instagram, TikTok, and YouTube. It explores best practices for video formatting, captioning, trends, hashtags, and audience targeting. Participants learn how to adapt content tone, length, and structure for each platform to maximize reach and impact.

After each video, a quiz will be administered to ensure that participants actively engage with and retain the material.

Statistical Analysis

All quantitative analyses will be conducted using SPSS Statistics (version 26; IBM Corp). Survey datasets will be screened for missing data and outliers prior to analysis. Descriptive statistics (counts, percentages, means, and SDs) will be calculated for participant demographics, module completion, retention, and satisfaction metrics.

For phase 1, pre- and posttraining survey knowledge items will be analyzed using the McNemar test to evaluate within-participant changes. Likert-scale items assessing satisfaction and perceived relevance will be summarized using frequencies and measures of central tendency. A significance threshold of a P value below .05 will be applied for all quantitative analyses.

For phase 2, social media engagement metrics (views, likes, shares, saves, and comments) will be summarized using means, medians, and engagement ratios as exploratory indicators of public reach and receptivity.

Qualitative Analysis

Qualitative data will be managed and analyzed using the NVivo software (Lumivero). Public social media comments will

XSL•FO
RenderX

undergo qualitative content analysis, with coding for evidence of knowledge uptake, misinformation correction, curiosity, and trust in dermatologic guidance.

Phase 2 interviews will be audio recorded, transcribed verbatim, and analyzed using thematic analysis following Braun and Clarke's 6-phase framework [20]. Two independent coders will familiarize themselves with the transcripts, generate initial codes aligned with the study aims, review and refine themes, and resolve discrepancies through consensus. Recruitment and analysis will continue until thematic saturation is achieved.

Ethical Considerations

This study was reviewed and approved by the University of Pittsburgh Office of Research Protections (approval STUDY24070132). The study was determined to involve minimal risk to participants.

All participants will receive detailed information about the study's aims, procedures, potential risks, and benefits and will provide informed consent electronically prior to participation, ensuring voluntary and fully informed participation in the study. Participants may withdraw at any time without penalty. No personally identifiable information will be collected beyond what is necessary for study purposes. All data will be stored on secure, password-protected institutional servers and will be accessible only to authorized members of the research team. As a thank-you for their time, participants will receive a US $10 gift card upon completion of phase 1. Those selected for and who complete phase 2 will receive a US $50 gift card.

The study will be conducted in accordance with institutional guidelines and applicable ethical standards for research involving human participants.

Results

Recruitment for phase 1 of the study is currently underway. As of January 2026, a total of 46 out of the target 50 participants have been successfully enrolled. Data collection for phase 1 is expected to be completed by the end of April 2026. Data analysis will begin shortly thereafter, with final analyses anticipated to be completed by September 2026. Recruitment for phase 2 is expected to begin in April 2026, with the goal of concluding data collection and analysis by the end of 2028 and publishing results in spring 2029.

Discussion

Expected Findings

This study is among the first to explore an online social media training program designed to equip health professional students and health care providers in dermatology or dermatology-adjacent fields with the practical skills to create engaging, evidence-based digital content. As social media platforms increasingly serve as the primary sources of health information, particularly for younger populations, ensuring the availability of accurate and accessible dermatologic content is increasingly urgent. Due to its pilot nature, this study is not designed to establish causal effects on public knowledge, attitudes, or behavior but, rather, to inform the feasibility and refinement of future interventions.

Although this pilot does not directly measure the causal impact of participant-created videos on viewer outcome, it is grounded on the premise that strengthening dermatology professionals' digital communication skills may increase their confidence and willingness to engage in public-facing education. Over time, this engagement has the potential to contribute to a broader and more sustained presence of accurate dermatologic messaging in the digital space [21]. By training both current and future health care providers to create engaging, evidence-based content, this intervention addresses a critical educational gap while empowering professionals to actively shape the online narrative around skin health [22]. The inclusion of core training components, such as storytelling, ethical content development, and platform-specific optimization, ensures that participants are equipped with both the technical and strategic tools necessary for meaningful digital engagement.

These goals align with recent findings indicating that health care professionals increasingly view social media as a valuable educational tool despite persistent concerns about privacy and misinformation. In a national survey, most dermatologists in the United States reported a willingness to increase their online engagement if better training and institutional support were available [23]. Similarly, structured digital media training has been shown to translate into real-world practice, with nearly 60% of trained providers applying their new skills to create patient-facing content [24].

Phase 2 of the study addresses the potential impact of participant-generated videos through a combination of engagement metrics (eg, likes, shares, and comments), qualitative analysis of viewer comments, and in-depth interviews with participants. Comment analysis will explore indicators such as expressions of learning, curiosity, misinformation correction, and attitudinal shifts. Participants will also reflect on perceived audience reactions and whether their content generated meaningful interactions.

The second phase—focused on video creation and performance metrics—adds a critical dimension by evaluating how educational content performs in real-world digital environments. This allows for preliminary insight into the feasibility of scaling such interventions and highlights the potential for trained health professionals to serve as trusted voices within an oversaturated media landscape. Metrics such as likes, views, shares, and comments will be used as proxies for audience receptivity and public interest in evidence-based dermatologic content. Participant reflections will also help identify key facilitators of and barriers to digital engagement, which will be essential for designing future interventions. While combating misinformation is a central goal, this study also aims to build long-term digital communication capacity by training health care professionals to confidently and ethically engage diverse online audiences.

The study is not without limitations. As a pilot study with a relatively small sample size for the content creation phase (n=12), findings on engagement and reach will be exploratory in nature. Additionally, participant motivation and prior familiarity with social media may influence both the quality of

XSL•FO
RenderX

the content and its performance. Future studies may consider stratifying participants based on baseline digital literacy and including a control group for more robust comparisons.

Despite these limitations, this protocol lays the foundation for a scalable, educationally grounded approach to addressing the growing challenge of dermatologic misinformation online [25]. If successful, this model could be replicated across other medical disciplines and programs and used to cultivate a new generation of digitally fluent, public-facing health professionals.

Conclusions

This pilot study aims to address the growing challenge of dermatologic misinformation on social media by equipping health professional students and providers with the tools and confidence to create accurate, engaging, and accessible educational content. Through a structured training program and real-world application phase, the study will evaluate both knowledge gains and the feasibility of digital content dissemination. By empowering participants to contribute to evidence-based discourse online, this intervention has the potential to enhance public dermatologic literacy; promote digital health equity; and serve as a model for scalable, specialty-specific training across health professions.

Funding

This study did not receive any external funding or financial support. JCL was supported by the Oklahoma Tobacco Settlement Endowment Trust (contract 00003615) and the OU Health Stephenson Cancer Center via a National Cancer Institute Cancer Center Support Grant (P30CA225520).

Authors' Contributions

Authors' Contributions: Conceptualization: AA (lead), MPM (supporting), BLH (supporting)
Data curation: EB (lead), AA (supporting)
Formal analysis: EB (lead), BLH (supporting)
Methodology: AA (lead), PML (supporting), JCL (supporting)
Project administration: AA (lead), AMR (supporting), EB (supporting)
Resources: PML, JCL, AED, BLH, MPM
Supervision: PML, JCL, AED, BLH, MPM
Visualization: AA, EB
Writing – original draft: AA (lead), EB (supporting)
Writing – review & editing: JCL, PML, AMR, AED, BLH, MPM
Authors BLH and MPM are co-senior authors and contributed equally in this role.

Conflicts of Interest

None declared.

Multimedia Appendix 1
TIDieR (Template for Intervention Description and Replication) checklist.
[PDF File, 186 KB - resprot_v15i1e81003_app1.pdf]

References

1. Ranpariya V, Chu B, Fathy R, Lipoff JB. Dermatology without dermatologists? Analyzing Instagram influencers with dermatology-related hashtags. J Am Acad Dermatol 2020 Dec;83(6):1840-1842. [doi: 10.1016/j.jaad.2020.05.039] [Medline: 32416205]
2. Nigro AR, Osman A, Saal RC, Smith RJ. Information regarding dermatology as seen on the social media platform TikTok. JMIR Dermatol 2025 Jan 28;8:e59597. [doi: 10.2196/59597] [Medline: 39874465]
3. Roland-McGowan J, Diehl K, Tobey T, et al. Perception, quality, and accuracy of sunscreen content on TikTok: SkinMedia cross-sectional content analysis. JMIR Dermatol 2025 Dec 1;8:e70010. [doi: 10.2196/70010] [Medline: 41328483]
4. Milanovic S, Rosen C, Gray T. Evaluating bias in social media research using #sunscreen content on Instagram reels. JMIR Dermatol 2025 Nov 5;8:e76829. [doi: 10.2196/76829] [Medline: 41191850]
5. Quartey A, Lipoff JB. Is TikTok the new Google? Content and source analysis of Google and TikTok as search engines for dermatologic content in 2024. J Am Acad Dermatol 2025 Jul;93(1):e25-e26. [doi: 10.1016/j.jaad.2024.10.131] [Medline: 40057279]
6. Subramanyam C, Becker A, Rizzo J, Afzal N, Nong Y, Sivamani R. Visibility of board-certified dermatologists on TikTok. JMIR Dermatol 2024 Jan 5;7:e46085. [doi: 10.2196/46085] [Medline: 38180786]
7. Trepanowski N, Grant-Kels JM. Social media dermatologic advice: dermatology without dermatologists. JAAD Int 2023 May;12:101-102. [doi: 10.1016/j.jdin.2023.05.004] [Medline: 37404249]

XSL•FO
RenderX

8. Szeto MD, Mamo A, Afrin A, Militello M, Barber C. Social media in dermatology and an overview of popular social media platforms. Curr Dermatol Rep 2021;10(4):97-104. [doi: 10.1007/s13671-021-00343-4] [Medline: 34692234]
9. Rey AB, Tan J. Social media in dermatology and skin health: challenges and opportunities. JEADV Clin Pract 2025 Sep;4(S1):S59-S63. [doi: 10.1002/jvc2.70096]
10. Vidal SI, Baek P, Menta N, Murphy E, Friedman A. Evaluating quality and reliability of most-viewed TikTok videos about spironolactone. J Drugs Dermatol 2025 Feb 1;24(2):e7-e9. [Medline: 39913218]
11. Akakpo MG. The role of care-seeking behavior and patient communication pattern in online health information-seeking behavior - a cross-sectional survey. Pan Afr Med J 2022;42:124. [doi: 10.11604/pamj.2022.42.124.33623] [Medline: 36060849]
12. Schick TS, Höllerl L, Biedermann T, Zink A, Ziehfreund S. Impact of digital media on the patient journey and patient-physician relationship among dermatologists and adult patients with skin diseases: qualitative interview study. J Med Internet Res 2023 Sep 22;25:e44129. [doi: 10.2196/44129] [Medline: 37738078]
13. Desta B, Burgess C. Educating through engagement: Instagram as a tool for dermatology advocacy. J Drugs Dermatol 2025 Aug 1;24(8):e56. [Medline: 40773604]
14. Heague M, Ray C, Bowers J, Guckian J, Arents BW, Layton A. Patient and public involvement in dermatology research: a review. Am J Clin Dermatol 2022 May;23(3):319-329. [doi: 10.1007/s40257-022-00680-5] [Medline: 35349092]
15. Auanassova A, Jatuworapruk K, Sarkar M, et al. Social media landscape: a cross-sectional survey of health professionals. Rheumatol Int 2025 Oct 23;45(11):255. [doi: 10.1007/s00296-025-06000-4] [Medline: 41128905]
16. Bujang MA, Omar ED, Foo DH, Hon YK. Sample size determination for conducting a pilot study to assess reliability of a questionnaire. Restor Dent Endod 2024 Jan 10;49(1):e3. [doi: 10.5395/rde.2024.49.e3] [Medline: 38449496]
17. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523. [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
18. Skin conditions by the numbers. American Academy of Dermatology. URL: https://www.aad.org/media/stats-numbers [accessed 2026-01-24]
19. Salah LA, AlTalhab S, Omair A, AlJasser M. Accuracy and quality of YouTube videos as a source of information on vitiligo. Clin Cosmet Investig Dermatol 2022;15:21-25. [doi: 10.2147/CCID.S330015] [Medline: 35027834]
20. Ahmed SK, Mohammed RA, Nashwan AJ, et al. Using thematic analysis in qualitative research. J Med Surg Public Health 2025 Aug;6:100198. [doi: 10.1016/j.glmedi.2025.100198]
21. Wojtara MS. Use of social media for patient education in dermatology: narrative review. JMIR Dermatol 2023 Apr 14;6:e42609. [doi: 10.2196/42609] [Medline: 37632938]
22. Van Ravenswaay L, Parnes A, Nisly SA. Clicks for credit: an analysis of healthcare professionals' social media use and potential for continuing professional development activities. Med Educ Online 2024 Dec 31;29(1):2316489. [doi: 10.1080/10872981.2024.2316489] [Medline: 38359156]
23. Bressler MY, Grudnikoff E, Bressler Y, Tamez R, Zampella JG. Risks and benefits of using social media in dermatology: cross-sectional questionnaire study. JMIR Dermatol 2021 Feb 24;4(1):e24737. [doi: 10.2196/24737] [Medline: 37632799]
24. Van Daele T, Iyer A, Zhang Y, Derry JC, Huh M, Pavel A. Making short-form videos accessible with hierarchical video summaries. Presented at: CHI '24: Conference on Human Factors in Computing Systems; May 11-16, 2024. [doi: 10.1145/3613904.3642839]
25. Zadu AK, Young J, Seltzer JA, Byrd AS, Frey C. The role of TikTok in education on hidradenitis suppurativa in skin of color: cross-sectional analysis. JMIR Dermatol 2026 Feb 27;9:e71566. [doi: 10.2196/71566] [Medline: 41758609]

Abbreviations

TIDieR: Template for Intervention Description and Replication
UPMC: University of Pittsburgh Medical Center

Edited by J Sarvestan; submitted 31.Jul.2025; peer-reviewed by B Eapen, L Vieira; revised version received 04.Mar.2026; accepted 04.Mar.2026; published 08.Apr.2026.

Please cite as:
Afolabi A, Brown E, Lyu JC, Ling PM, Rustad AM, Davis AE, Hoffman BL, Pugliano-Mauro M
Combating Dermatologic Misinformation on Social Media via a Training Program for Health Care Professionals: Protocol for a 2-Phase Pilot Mixed Methods Study
JMIR Res Protoc 2026;15:e81003
URL: https://www.researchprotocols.org/2026/1/e81003
doi:10.2196/81003

XSL•FO
RenderX

© Aliyyat Afolabi, Elijah Brown, Joanne Chen Lyu, Pamela M Ling, Andrea M Rustad, Anna E Davis, Beth L Hoffman, Melissa Pugliano-Mauro. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact of a Prehospital Chest Pain Alert App–Mediated Prehospital–in-Hospital Coordination Model on Treatment Delays and Clinical Outcomes in Patients With ST-Elevation Myocardial Infarction: Protocol for a 4-Year Retrospective Real-World Cohort Study

Shuyuan Chen[1], MBBS; Rensong Wang[1], MBBS; Jiangwei Ma[2], MD; Weichun Xiao[3], MBBS; Changxin Song[4], PhD

[1]Emergency Department, Shanghai Fengxian District Medical Emergency Center, No. 160 Deshun Road, Fengxian District, Shanghai, China

[2]Department of Cardiology, Fengxian District Central Hospital, Shanghai, China

[3]Department of Medical Education, Shanghai Fengxian District Medical Emergency Center, Shanghai, China

[4]School of Artificial Intelligence Application, Shanghai Urban Construction Vocational College, Shanghai, China

Corresponding Author:
Shuyuan Chen, MBBS
Emergency Department, Shanghai Fengxian District Medical Emergency Center, No. 160 Deshun Road, Fengxian District, Shanghai, China

Abstract

Background: The effectiveness of ST-elevation myocardial infarction (STEMI) treatment is highly time-dependent, and the information barrier between prehospital and in-hospital settings remains a key driver of treatment delays. Existing digital coordination tools either have a single function or lack long-term real-world evidence, making it difficult to meet clinical needs. This study adopts a prehospital chest pain alert app developed by the Fengxian District Medical Emergency Center. Mediated through a WeChat-based chest pain center group, the app enables prehospital information synchronization, real-time alerts, multidisciplinary coordination, and feedback on treatment outcomes to form a closed-loop model, overcoming the information barrier.

Objective: This protocol aims to evaluate the impact of the app-mediated prehospital–in-hospital coordination model on treatment delays (eg, time from first electrocardiogram to catheterization laboratory preactivation and door-to-wire time) and clinical outcomes (eg, 30-day major adverse cardiovascular events, and 1-year and 4-year all-cause mortality) in patients with STEMI, and to assess its generalizability in high-risk subgroups.

Methods: This is a single-center retrospective cohort study. Patients with STEMI admitted to Fengxian District Central Hospital from January 1, 2019, to December 31, 2024, will be enrolled and categorized into 3 groups: baseline group (January 1, 2019, to December 31, 2020, without app use), intervention group (January 1, 2021, to December 31, 2024, with app-mediated coordination), and concurrent control group (patients with STEMI who came to the hospital independently without calling an ambulance or were transported by ambulance but not reported via the app during the same period). The primary outcome is door-to-wire time. Secondary outcomes include other treatment delay indicators, clinical prognosis, and app operational efficiency. We will use propensity score matching to control for baseline confounding, segmented linear regression to analyze intervention trend effects, and subgroup analysis to assess generalizability in high-risk populations.

Results: This study is based on 4 years of real-world data from the Department of Cardiology and the STEMI database of Fengxian District Central Hospital. As of April 2026, all 2019-2021 data have been collected; a sample size of 944 or more is expected. Data cleaning and statistical analysis are scheduled from May 2026 to June 2026.

Conclusions: Based on 4 years of real-world data, combined with propensity score matching and interrupted time series analysis, this study aims to provide high-quality observational evidence for the app-mediated prehospital–in-hospital coordination model. The findings are anticipated to offer preliminary references for optimizing regional STEMI care systems and to inform the potential application of digital health technologies in acute coronary syndrome management.

International Registered Report Identifier (IRRID): PRR1-10.2196/90144

(JMIR Res Protoc 2026;15:e90144) doi:10.2196/90144

XSL•FO
RenderX

KEYWORDS

ST-elevation myocardial infarction; STEMI; prehospital-hospital coordination; mobile health; mHealth; treatment delay; clinical outcome; real-world study; retrospective cohort study; alert app

Introduction

ST-elevation myocardial infarction (STEMI) is a life-threatening acute cardiovascular event, and timely reperfusion therapy is crucial for improving prognosis [1]. Guidelines recommend a door-to-wire time (D2W) of 90 minutes or less for patients with STEMI in percutaneous coronary intervention (PCI)-capable centers and a transfer time of 120 minutes or less for those in non-PCI-capable centers transferred to PCI centers [2]. However, asynchronous prehospital and in-hospital information, delayed catheterization laboratory activation [3], and low efficiency of multidisciplinary coordination remain prevalent, leading to widespread treatment delays that significantly increase patient mortality and the risk of adverse outcomes [4]. Current guidelines and prior research highlight the urgency of optimizing prehospital–in-hospital coordination [5,6].

Existing digital intervention tools mostly focus on a single function (eg, electrocardiogram [ECG] transmission) and lack a full-chain design covering “information collection-real-time alert-coordinated disposal-feedback closed loop” [6]. Additionally, evidence regarding their generalizability in high-risk populations (eg, older adult patients and those with diabetes or hypertension) is insufficient. Previous small-scale pilots have shown that the prehospital chest pain alert app (hereafter referred to as the app) can shorten information transmission time, but its long-term real-world effectiveness remains unclear [7].

Leveraging 4 years of real-world data, this retrospective cohort study seeks to evaluate the impact of the app-mediated coordination model on treatment delays and clinical outcomes, as well as its generalizability in high-risk subgroups. The study aims to generate preliminary evidence to inform the optimization of regional STEMI care systems.

Methods

Study Design

This is a single-center retrospective cohort study. The study design, variable definitions, and statistical analysis plan are detailed below, and data collection and analysis will be strictly performed in accordance with this protocol without arbitrary changes to core designs. The study will be conducted at Fengxian District Central Hospital. Patients with STEMI diagnosed from January 1, 2019, to December 31, 2024, will be included and divided into 3 groups: baseline group (January 1, 2019-December 31, 2020, before the app was developed), intervention group (January 1, 2021-December 31, 2024, with app-mediated coordination), and concurrent control group (patients with STEMI who did not use the app during the same period, either self-admitted or transported by ambulance without app reporting). To mitigate potential confounding from temporal trends in STEMI care (eg, accumulated clinical experience of the care team and incremental workflow optimizations of the chest pain center) that might overestimate the app’s intervention effect, we included a concurrent control group consisting of patients with STEMI treated during the same period (2021 - 2024) who did not use the app. This design allows us to compare the app-exposed group against contemporaneous, nonexposed patients, thereby helping to isolate the specific impact of the app intervention from broader system-wide improvements of the chest pain center.

Prehospital Care Pathways

During the study period, all patients with STEMI were managed through 1 of 3 clinically distinct prehospital pathways, which also served as the basis for group classification in this study. The 3 pathways are as follows:

- Pathway A (self-presenting): patients arrive at the emergency department (ED) without using an ambulance. They must first undergo triage by the front-desk nurse; if chest pain is reported, they are immediately transferred to the chest pain center for in-hospital ECG and cardiac biomarker testing. Cardiology consultation is initiated only after abnormalities are confirmed.
- Pathway B (ambulance-transported, standard alert): patients are transported by ambulance but do not use the app. Prehospital personnel provide a verbal chest pain alert via traditional telephone, allowing the patient to bypass ED triage and proceed directly to the chest pain center for immediate in-hospital ECG upon arrival.
- Pathway C (ambulance-transported, app-mediated): patients are transported by ambulance, and the app is used. A prehospital 12-lead ECG is acquired at the scene and transmitted in real time to the hospital chest pain center via the app. If the on-duty cardiologist remotely confirms STEMI, the catheterization laboratory is preactivated before the patient leaves the scene. Upon hospital arrival, the patient bypasses the triage desk and is directly admitted to the chest pain center treatment area within the ED for rapid confirmation and precatheterization procedures (eg, informed consent). The ultimate goal of this pathway is to enable eligible patients to bypass the ED entirely and proceed directly to the catheterization laboratory (Multimedia Appendices 1 and 2).

By explicitly delineating these pathways, we identify the unique and irreplaceable core contributions of the app: real-time prehospital ECG acquisition and remote catheterization laboratory preactivation before patient arrival. These 2 mechanisms are the primary drivers of expected reductions in STEMI treatment delays and directly correspond to the novel door-to-preactivation time (D2P) indicator introduced in this study, which is designed to specifically capture the intervention effect of the app.

We acknowledge that the app-based intervention is a multicomponent integrated care model, encompassing prehospital 12-lead ECG acquisition, real-time data transmission, remote cardiologist diagnosis, and prehospital

XSL•FO
RenderX

catheterization laboratory preactivation. This study is designed to evaluate the composite effect of this coordinated model on reducing STEMI treatment delays, while clarifying that the 2 unique and direct consequences of app use—early acquisition of a diagnostic ECG and subsequent remote catheterization laboratory preactivation—are the core drivers of delay reduction. Due to the inherent limitations of a retrospective cohort study design and the highly integrated nature of these components in real-world clinical practice, we are unable to disentangle the independent contribution of each individual component to the observed reduction in treatment delays.

We acknowledge that the baseline group (2019 - 2020) and the concurrent control group (2021 - 2024; non–app users) may inherently differ not only in app usage but also due to broader temporal trends, including the maturation of chest pain center protocols, accumulated clinical experience, and post-COVID adjustments in acute cardiac care. To mitigate this potential confounding, we used interrupted time series (ITS) analysis to quantify and separate the secular trend from the intervention effect. Additionally, the inclusion of a concurrent control group treated during the same period (2021 - 2024) allows us to compare app-exposed patients with contemporaneous nonexposed patients, thereby partially isolating the app's specific impact from system-wide improvements. Nevertheless, residual confounding cannot be fully excluded due to the retrospective design, and this limitation is acknowledged in the "Discussion" section.

Study Population and Grouping

Inclusion Criteria

Patients were eligible for inclusion if they met all of the following criteria: (1) met STEMI diagnostic criteria (typical symptoms plus ST-segment elevation on ECG or elevated myocardial enzymes), (2) onset time of 12 hours or less, (3) aged 18 years or older, (4) had complete clinical data, and (5) underwent only forward conventional PCI as defined in the "Explanation for Limiting Surgical Methods: Rationale for Limitation" section [6,8].

Exclusion Criteria

Patients were excluded if they met any of the following criteria: (1) had end-stage diseases such as severe liver or kidney failure or malignant tumors; (2) refused PCI or abandoned treatment; (3) were interhospital-transferred patients (due to the unique characteristics of treatment delays in this population, requiring separate research [9]); (4) had severely missing clinical data; or (5) underwent nonforward conventional PCI, including but not limited to retrograde PCI, chronic total occlusion interventional therapy, coronary atherectomy, laser angioplasty, coronary artery bypass grafting combined with PCI, and simultaneous complex interventional therapy for multivessel disease (see "Explanation for Limiting Surgical Methods: Rationale for Limitation" section) [4,8].

Group Definitions

Patients with STEMI are categorized into 3 groups. First, the baseline group (January 2019 to December 2020) uses a conventional care model (only telephone notification and unable to transmit the first ECG). Second, the intervention group (January 2021 to December 2024) uses an app-mediated coordination model (prehospital emergency physicians upload patient information, the first ECG, and disposal measures via the app; the in-hospital team receives real-time alerts, preactivates the catheterization laboratory, and provides feedback on disposal suggestions). Finally, the concurrent control group (January 2021 to December 2024) consists of patients with STEMI not using the app (self-admitted or transported by ambulance without app reporting, ie, the first ECG was obtained within the hospital ED).

Sample Size Calculation

Referring to a systematic review by Gibson et al [6], the app intervention can shorten the mean D2W by 18.6 (SD 32.4) minutes. This study hypothesizes that the app can shorten D2W by 20 minutes, with an α level of .05 (2-sided) and a β of .10 (90% power). Using the sample size formula for a 3-arm design, each group requires 286 cases. Allowing for 10% attrition due to exclusion criteria, the total sample size needs to be 944 or more. Based on the center's annual average of over 300 total chest pain or acute coronary syndrome admissions (with a mean of 260, SD 25 confirmed STEMI cases per year), the cumulative number of cases meets the requirement.

Explanation for Limiting Surgical Methods: Rationale for Limitation

Clinical Dominance of Forward Conventional PCI

Forward conventional PCI is the most mainstream and standardized procedure for primary PCI in patients with STEMI. According to a systematic review by Gibson et al [6] and domestic STEMI diagnosis and treatment data, this procedure accounts for 85% to 90% of total STEMI-related PCI procedures. Its operational process is unified (forward approach puncture-angiography-balloon dilation-stent implantation) without additional complex technical interventions, making it an ideal research object for evaluating the impact of prehospital–in-hospital coordination models on treatment delays.

Avoiding Confounding Effects of Surgical Methods on Treatment Delays

Operational complexity and time consumption vary significantly among different PCI procedures: the average operation time for forward conventional PCI is 30 minutes to 1 hour [4], while that of complex procedures, such as chronic total occlusion intervention and retrograde PCI, can be 1.5 to 4 hours or longer [8]. The inclusion of multiple surgical methods will directly interfere with intergroup comparison of core outcome indicators, such as D2W and total ischemic time, leading to the true effect of the app-mediated coordination model being masked or exaggerated.

Focus on Core Research Objectives

The core goal of this study is to verify the improvement effect of the "app-mediated prehospital–in-hospital coordination model" on treatment delays and clinical outcomes, rather than comparing the advantages or effectiveness of different PCI procedures. By limiting the inclusion to patients undergoing forward conventional PCI, the time variation related to surgical

XSL•FO
RenderX

operations can be homogenized, focusing the research on the process optimization effect of "information transmission–real-time alert–catheterization laboratory preactivation" and significantly improving the internal validity of the study results [6,10].

Intervention Measures

Core Functions of the App

The app's core functions include the following four aspects: (1) information collection—prehospital emergency personnel enter patient basic information, symptoms, vital signs, and 12-lead ECG images; (2) real-time transmission—AES-256 encrypted transmission supporting JPG or PNG format ECG upload, connected to the in-hospital chest pain center WeChat work group; (3) coordinated disposal—in-hospital cardiologists receive real-time alerts, remotely interpret ECGs, activate the catheterization laboratory, and provide feedback on disposal suggestions; and (4) closed-loop management—recording prehospital disposal (eg, dual antiplatelet drug use) and summarizing core in-hospital disposal to achieve a full-chain information closed loop [11].

Technical Parameters of the App

The technical parameters of the app are as follows: it was independently developed by the Fengxian District Medical Emergency Center; its operating platform supports iOS 12.0+ and Android 8.0+; and no commercial software copyright registration has been obtained for this software.

Data Collection and Variable Definitions

Baseline Data

Including demographic characteristics (age, sex, and BMI), underlying diseases (hypertension, diabetes, and history of coronary heart disease), year of admission, and onset period (daytime or nighttime and weekday or holiday).

Treatment Delay Indicators

Treatment delay indicators include the time from the first ECG to catheterization laboratory activation; D2W, which, in reference to the 2025 American College of Cardiology/American Heart Association guidelines, is defined as the time from patient arrival at the emergency triage desk to guidewire passage through the infarct-related artery [1]; D2P, defined as the interval between patient arrival at the hospital door and the time of catheterization laboratory preactivation, where a negative value indicates that preactivation occurred before patient arrival—an outcome only achievable when a definitive STEMI diagnosis is made prehospital via real-time ECG transmission and early alert of the app. This metric directly captures the unique contribution of the app and is not confounded by secular trends (eg, staff experience and routine workflow changes), as such trends cannot transform a positive D2P value into a negative one; and total ischemic time, defined as the time from onset to guidewire passage.

To ensure consistency across the 2019 - 2024 study period, all D2W data were extracted using the same standardized definitions: door time was defined as the patient's arrival at the emergency triage desk (from electronic triage records), and wire time as the first guidewire passage (from PCI procedure logs). Extreme values were verified and handled according to prespecified rules. Thus, although the 2025 guidelines are cited for definitional clarity, the actual measurements are directly comparable throughout the study.

App Performance Metrics

To objectively reflect the real-world operational performance of the app (WeChat mini-program embedded), 3 core performance metrics were defined based on the clinical workflow of STEMI prehospital care, with all time and case data extracted from the app operation logs and chest pain center official records. The operational definitions and calculation methods were as follows. First, app STEMI alert response time was defined as the time interval (in minutes) from the first submission of prehospital ECG and STEMI alert information by prehospital medical staff via the app to the confirmation of receipt and issuance of catheterization laboratory preactivation instructions by the on-duty physician of the chest pain center via the app or coordination group. This was calculated for each individual case as the time difference between the app and clinical records, with the overall distribution summarized for reporting. Second, the timely and complete prehospital information submission rate within 5 minutes was defined as the proportion of STEMI cases in which prehospital medical staff completed the first timely and complete submission of core prehospital information (including the first ECG image, patient vital signs, onset time, and on-site preliminary diagnosis) via the app within 5 minutes of arriving at the patient's side. This was calculated as (number of timely and complete submission cases/total number of app-based prehospital submission cases) × 100%. Finally, prehospital–in-hospital clinical information closed-loop feedback time was defined as the time interval (in hours) from the first prehospital information submission via the app to the post-PCI feedback of core in-hospital information (including angiographic results, surgical method, and postoperative blood flow status) by cardiologists via the app. This was calculated for each individual case as the time difference from the app logs, with the overall distribution summarized for reporting. All app performance metrics will be reported as descriptive statistics (mean, median, IQR, and percentage where applicable). These metrics are purely descriptive and will not be used to support any causal inferences regarding the app's effect on STEMI treatment outcomes.

Clinical Outcome Indicators

Clinical outcome indicators include short-term outcomes, which comprise 30-day major adverse cardiovascular events (including all-cause death, reinfarction, and heart failure) [12,13] and PCI success rate; long-term outcomes, which include 1-year and 4-year all-cause mortality [14,15] and left ventricular ejection fraction [16]; and app operational efficiency metrics, such as information transmission success rate [17,18].

Data Sources

Overview

The data were mainly derived from the Data Reporting Platform of the Fengxian District Chest Pain Center and the background statistical data of the app. The former provides full-process

XSL•FO
RenderX

timeline data including the sending time and response time of the first prehospital ECG, as well as key clinical information such as surgical approaches, subgroup classification, records of adverse events within 30 days, and 1-year and 4-year follow-up data. The latter compiles data on all patients treated through this collaborative system and calculates diagnostic accuracy. For this study, the data scope was strictly defined: we focused only on patients with STEMI managed within the system and excluded misdiagnosed patients and patients with non-STEMI, angina pectoris, and nonacute coronary syndrome.

Brief Description of Data Extraction and Coding Rules

A structured data extraction form (Multimedia Appendix 3) will be used to systematically collect all study variables. The form covers 5 modules: basic identification information, baseline clinical characteristics, treatment delay indicators, clinical outcomes, and app operational efficiency to ensure standardized and complete data collection. Data coding follows unified rules (Multimedia Appendix 3), clarifying the assignment of categorical variables, processing standards for continuous variables, and determination processes for missing and outlier values. All data will be extracted by 2 trained researchers using the double-blind method. After extraction, cross-validation will be performed and discrepancies will be adjudicated by a third cardiologist to ensure data accuracy.

Statistical Analysis

Data Processing

Data will be extracted and verified by professional professors from local universities. EpiData 3.1 (EpiData Association) will be used to establish the database, and missing values will be handled by multiple imputation.

Statistical Methods

All statistical analyses were performed using Stata SE 17.0 (StataCorp LLC). Statistical analyses were planned as follows:

- Descriptive statistics: continuous variables will be expressed as mean (SD) or median (IQR), and categorical variables as n (%).
- Intergroup comparison: independent samples 2-tailed *t* test or rank sum test for continuous variables and chi-square test for categorical variables before propensity score matching (PSM), with matching variables including demographic characteristics (age, sex, and BMI), comorbidities (hypertension, diabetes, and history of coronary heart disease), and key care pathway characteristics (self-presenting vs ambulance-transported) to balance these factors between the baseline and concurrent control groups, and generalized linear model (normal distribution) or quantile regression (nonnormal distribution) after PSM.
- Time series analysis: segmented linear regression (ITS) will be used with the formula $Yt = \beta0 + \beta1 \times time + \beta2 \times intervention + \beta3 \times time \times intervention + \varepsilon t$, with autocorrelation tested (Durbin-Watson test) and corrected using the autoregressive integrated moving average model (1,0,0) if needed [10].
- Survival analysis: Cox proportional hazards regression model will be used to calculate the adjusted hazard ratio and 95% CI.
- Subgroup analysis: stratification by age (≥65 y vs <65 y) [8], diabetes (yes or no) [19], and hypertension (yes or no) [20] to test interaction effects.

Sensitivity Analysis

Given the inherent clinical heterogeneity of the concurrent control group—which comprises 2 distinct patient populations: (1) self-presenting patients with STEMI who must first undergo triage at the emergency triage desk before entering the chest pain center treatment area, and (2) ambulance-transported patients with STEMI not using the app who, via traditional telephone notification, bypass the triage desk and proceed directly to the chest pain center treatment area—these 2 subgroups have clear differences in baseline clinical characteristics and prehospital care pathways. Notably, the triage process itself directly adds time from hospital arrival to entry into the chest pain center treatment area for self-presenting patients. We will conduct a prespecified sensitivity analysis stratifying this group into these 2 subgroups to assess whether the app's effect on core treatment delay indicators differs between these clinically distinct pathways and to evaluate the robustness of our primary findings against potential selection bias arising from these well-defined prehospital pathways and time-delay differences. It is important to note that the number of ambulance-transported patients not using the app in the later study period is extremely limited, as prehospital notification for STEMI cases was changed from "encouraged" to "mandatory" and incorporated into performance indicators after app implementation (with slightly more such cases only in the first year due to initial low compliance). Therefore, the statistical power of analyses focusing on this subgroup may be limited.

Ethical Considerations

Ethical Approval

The study was approved by the medical ethics committee of Fengxian District Central Hospital (2025-KY-116 - 02; January 14, 2026; Multimedia Appendix 4). Due to the use of retrospective deidentified data, informed consent was waived in accordance with the "Measures for the Ethical Review of Biomedical Research Involving Humans" issued by the National Health Commission of the People's Republic of China [21]. No financial compensation was provided to participants in this retrospective study, as all data used were deidentified and retrospectively collected without direct participant interaction. All procedures comply with the Declaration of Helsinki and relevant national laws.

Data Security

For data security, strict protective measures are implemented in 3 key aspects involving deidentification standards, storage and access management, and data backup protocols. Only the last 4 digits of patients' ID numbers are retained, their names are replaced with unique codes, and personal addresses are not recorded. Deidentified data are stored on encrypted servers with hierarchical authorized access, where the principal investigator can access the full dataset while statistical analysts only have

XSL•FO
RenderX

access to coded data. In addition, a dual backup mechanism combined with regular off-site backups is adopted, and all data management practices are in full compliance with the Personal Information Protection Law of the People's Republic of China and the Data Security Law of the People's Republic of China.

Quality Control

A comprehensive set of quality control measures is implemented for data processing, including outlier handling, consistency testing, and regular verification. For outlier handling, extreme values, such as a D2W exceeding 300 minutes, will be traced and verified, and invalid data caused by delayed recording will be excluded. For consistency testing, the double-blind method will be adopted for ECG interpretation, with any discrepancies in results adjudicated by a third independent expert. For regular verification, the completeness of collected data will be checked every 2 weeks to ensure the accuracy of data entry throughout the research process.

Dissemination Plan

The findings of this study will be submitted to JMIR mHealth and uHealth for publication, focusing on the methodological framework of the novel D2P indicator and its clinical application in STEMI care. The practical experience of the prehospital–hospital coordinated treatment model will be shared at national academic conferences on emergency medicine and cardiology. Targeted practical training will also be provided for prehospital emergency personnel and chest pain center teams to promote the implementation of standardized treatment procedures. In addition, deidentified aggregated data will be shared in a public repository within 3 years after publication to facilitate methodological replication and secondary analysis by other researchers.

Results

Study data will be obtained from eligible patients with STEMI admitted between January 1, 2019, and December 31, 2024. As of April 2026, all 2019-2021 data have been collected. Data cleaning and statistical analysis are scheduled from May 2026 to June 2026. This study will evaluate the impact of the app-mediated prehospital–in-hospital coordination model on treatment delays and clinical outcomes in patients with STEMI, with baseline balance verification and subgroup analyses included.

Discussion

Overview

This study is a 4-year real-world retrospective cohort study designed to evaluate the effect of an app-mediated prehospital–in-hospital coordination model on STEMI treatment delays and clinical outcomes, as well as its generalizability in high-risk subgroups. We hypothesize that the primary findings of this study will be as follows: the app-mediated model will significantly reduce key treatment delay indicators, including D2W and the time from first ECG to catheterization laboratory activation, and the intervention group will exclusively present negative D2P values—a novel indicator that directly captures the app's core effect of real-time prehospital ECG acquisition and remote catheterization laboratory preactivation before patient arrival. We also anticipate that the model will improve short-term clinical outcomes (eg, 30-day major adverse cardiovascular events) and reduce 1-year and 4-year all-cause mortality in patients with STEMI, with consistent beneficial effects in high-risk subgroups, including older adult patients (≥65 y) and those with hypertension or diabetes. Stratified sensitivity analysis of the concurrent control group is expected to verify the robustness of these primary findings, while the D2P indicator will effectively isolate the app's intervention effect from confounding by secular trends in STEMI care (eg, accumulated staff experience and routine workflow optimizations). All interpretations of the study findings will be made with appropriate caution, and no causal claims will be overstated beyond the scope of the study's methodology and real-world retrospective design.

Principal Findings

The anticipated principal findings of this study center on the unique and irreplaceable core mechanisms of the app-mediated model in reducing STEMI treatment delays, which are clearly delineated by the 2 clinically distinct prehospital care pathways defined in this study: real-time prehospital 12-lead ECG acquisition and remote catheterization laboratory preactivation before patient hospital arrival. These 2 mechanisms are the direct and exclusive consequences of app use, and their intervention effect is specifically captured by the D2P indicator—negative D2P values in the intervention group will confirm that the app fundamentally alters the traditional STEMI care pathway by moving key diagnostic and catheterization laboratory activation steps from the in-hospital to the prehospital phase, a change that cannot be achieved by secular system-wide improvements of the chest pain center alone. The 3-arm cohort design (baseline, intervention, and concurrent control) is expected to further confirm that the observed reduction in treatment delays is attributable to the app intervention rather than temporal trends, as the concurrent control group (2021 - 2024; nonapp users) allows for comparison with contemporaneous patients receiving standard care in the same mature chest pain center system (established since 2019 in our hospital). We also anticipate that the app-mediated model will show consistent effectiveness in high-risk STEMI subgroups, filling the evidence gap for digital health interventions in these clinically important populations. Furthermore, this study will confirm the composite effect of the app's multicomponent design (information collection, real-time transmission, remote cardiologist diagnosis, and closed-loop management); due to the highly integrated nature of these components in real-world clinical practice and the inherent limitations of retrospective cohort design, the independent contribution of each individual component to treatment delay reduction cannot be disentangled in this study.

Comparison With Prior Work

Existing research on digital health interventions for STEMI care has primarily focused on single-function tools (eg, standalone ECG transmission systems) and small-scale pilot studies with short follow-up periods, lacking long-term real-world evidence

XSL•FO
RenderX

and full-chain design covering the entire prehospital–in-hospital coordination process [6,7]. A recent systematic review and meta-analysis by Gibson et al [6] found that smartphone apps can shorten the mean D2W by 18.6 minutes, but most included studies failed to account for temporal trend confounding or stratify heterogeneous control groups, leading to the potential risk of overattributing treatment delay reduction to digital interventions alone. In contrast, this study addresses these critical limitations by incorporating a concurrent control group and introducing the novel D2P indicator, which directly captures the app's exclusive intervention effect and is not confounded by secular trends such as staff experience accumulation [1,3]. Additionally, previous studies have rarely evaluated the generalizability of digital interventions in high-risk STEMI subgroups [19,20], while this study specifically stratifies analysis by age, hypertension, and diabetes, providing targeted evidence for the clinical application of digital health tools in these populations. The app in this study also differs from existing digital intervention tools in its closed-loop management design, which integrates prehospital disposal records and in-hospital treatment feedback [11], aligning more closely with real clinical workflows and thus improving its practical applicability. Few studies have explicitly delineated prehospital care pathways to identify the core mechanisms of digital interventions for STEMI [22], and this study's clear classification of 3 distinct pathways fills this gap by distinguishing the app's unique contributions (pathway C) from standard care pathways (pathways A and B).

Study Strengths

This study has 3 core strengths. First, the 3-arm cohort design combined with ITS analysis not only controls baseline confounding but also captures the long-term trend effect of the intervention with rigorous methodology. Second, the study focuses on high-risk subgroups, such as older adults and those with underlying diseases, thereby filling the evidence gap regarding digital interventions in this population. Third, the app is aligned with clinical workflows, includes clear technical details, ensures high reproducibility, and facilitates implementation.

A key strength of our study design is the introduction of the D2P as a novel process indicator. Unlike D2W, which can be influenced by numerous downstream factors (eg, PCI procedure complexity and in-hospital team efficiency), this indicator directly reflects the app's core mechanism—enabling prehospital STEMI diagnosis and remote catheterization laboratory preactivation. This effect is exclusive to the app and cannot be confounded by secular trends such as staff experience accumulation or protocol evolution. The magnitude of the negative D2P value is directly dependent on the time elapsed from ambulance arrival at the ED door to patient presentation for care. We further enhanced the clarity of the app's intervention mechanism by explicitly delineating 3 distinct prehospital care pathways and locking in the app's unique core contributions, which are directly captured by the novel D2P indicator we introduced. The anticipated finding that a negative D2P value exists exclusively in the intervention group provides strong, direct evidence that the app fundamentally alters the STEMI care pathway in a way that cannot be explained by secular trends or system-wide improvements of the chest pain center. This metric thus serves as an internal validation of the app's unique intervention effect.

Study Limitations

Single-center data may limit external generalizability, and multicenter verification is needed in the future. Interhospital-transferred patients are not included. This population faces more prominent treatment delay issues, and the use of the app in this scenario needs further exploration [9]. The retrospective design may have residual confounding (eg, patient health care–seeking behavior), which has been minimized through PSM and stratified analysis. Additionally, the concurrent control group has inherent clinical heterogeneity, as it comprises 2 subgroups with distinct prehospital care pathways: self-presenting patients who require nurse-led front-desk triage before entering the chest pain center, and ambulance-transported patients without app use who directly enter the chest pain center via traditional telephone prenotification. These subgroups have distinct baseline clinical characteristics and prehospital care processes, and the nurse-led front-desk triage process directly increases the time from the patient's hospital arrival to entry into the chest pain center treatment area, which introduces potential selection bias into the study. Although we have planned a prespecified sensitivity analysis to stratify the concurrent control group by prehospital pathways and verify the robustness of our primary findings, the statistical power of this stratified analysis is inherently limited: the sample size of ambulance-transported patients without app use is extremely small in the later study period, owing to the mandatory reporting requirement with assessment indicators for STEMI cases after app launch (a slightly larger sample only existed in the first year of app launch due to initial low usage compliance). This study evaluates the composite effect of the app-based multicomponent intervention model, and the inherent limitations of the retrospective cohort study design prevent us from disentangling the independent contribution of each individual component to treatment delay reduction. Residual confounding cannot be fully excluded due to the retrospective study design. Future prospective studies with more homogeneous control groups stratified by prehospital care pathways and larger sample sizes are therefore warranted to validate the conclusions of this study. Notably, the single-center design of this study is a result of the specific medical layout of Shanghai Fengxian District, where our hospital is the only institution with national chest pain center certification and emergency PCI qualification for STEMI treatment—this objective medical context limits the feasibility of a multicenter design in this study. It is critical to clearly distinguish between the generalizable core concept of the app-mediated model (prehospital real-time ECG transmission and remote catheterization laboratory preactivation, an evidence-based strategy recommended by international STEMI guidelines [1,2]) and the locally dependent implementation details (eg, the WeChat-based chest pain center group and the self-developed app); the former is universally applicable to STEMI care systems worldwide, while the latter can be adapted to local medical resources and communication tools. Furthermore, while the concurrent control group and D2P indicator minimize the risk of overattributing effects to the app, the intervention period

XSL•FO
RenderX

overlaps with ongoing refinements of chest pain center protocols, accumulated clinical experience of the care team, and post-COVID operational adjustments in acute cardiac care, and part of the observed improvement in treatment delays may reflect broader system maturation rather than the app alone. Furthermore, the baseline group (2019‐2020) and the concurrent control group (2021‐2024; nonapp users) may inherently differ due to broader temporal trends (eg, COVID-19 impact and system-wide protocol evolution), which cannot be fully eliminated despite the use of ITS analysis and a concurrent control group design. This residual confounding is an inherent limitation of retrospective observational studies. All study conclusions will be restricted to the scope of the real-world retrospective design, and no overgeneralization to other medical settings or populations will be made.

Future Directions

Future research could pursue 3 interconnected directions centered on the D2P indicator and app-mediated model. Starting with inward refinement, the determinants of D2P magnitude—including prehospital, in-hospital, and patient-specific factors—warrant further investigation to optimize ED bypass workflows, with standardized subnode decomposition of D2P's calculation facilitating cross-institutional replication. Moving to outward generalization, multicenter studies are needed to validate context-specific D2P reference ranges and the model's core mechanisms (prehospital ECG transmission and remote catheterization laboratory preactivation) across diverse emergency medical service systems, while adapting implementation details to local resources and replicating the study's statistical framework for methodological consistency. Complementing these efforts, technological advancement could streamline D2P-related workflows by integrating artificial intelligence–assisted ECG diagnosis [23] into the app—accelerating atypical STEMI identification and reducing false activations—while health economic evaluations should anchor on D2P as a key process indicator, linking its reduction to tangible outcomes (eg, shorter length of stay and lower major adverse cardiovascular events–related costs) to maintain alignment with the study's core design.

Acknowledgments

The authors thank the medical ethics committee of the Fengxian District Central Hospital for their guidance on the ethical review. The authors appreciate the data support provided by the Department of Cardiology, the Chest Pain Center Registration Team of Fengxian District Central Hospital, and the data managers of the Fengxian District Medical Emergency Center. The authors give special thanks to the Fengxian District Medical Emergency Center for providing technical support for the development of the prehospital chest pain alert app involved in this study, and thank the study team members for their collaboration, as well as all included patients and their families. The app was independently developed by the Fengxian District Medical Emergency Center. No commercial entities participated in the study design or data analysis, and personnel involved in the app development did not participate in this study. Generative artificial intelligence (AI) tools (Doubao AI, Doubao Large Model 2.0, ByteDance Inc) were used exclusively for auxiliary language polishing, English translation, and formatting support during manuscript preparation. No AI was involved in core academic content, including research design, data analysis, results interpretation, or core academic arguments. All AI-assisted outputs were fully reviewed, manually revised, and verified by all authors, who take full responsibility for the accuracy, originality, and integrity of the entire manuscript.

Funding

This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The study dataset contains sensitive patient health information and cannot be publicly shared immediately. The fully deidentified dataset will be made available to qualified researchers upon reasonable request within 3 years after publication, in compliance with national data management regulations. For data access requests, please contact the corresponding author.

Authors' Contributions

Conceptualization: SC
Data curation: RW
Ethics approval: RW
Formal analysis: CS
Investigation: RW
Methodology: SC
Supervision: JM, WX
Validation: CS
Writing – original draft: SC
Writing – review & editing: JM, WX, CS

XSL•FO
RenderX

All authors confirm that they meet the International Committee of Medical Journal Editors criteria for authorship, have read and approved the final manuscript, and take full responsibility for the integrity and accuracy of the research content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Prehospital data collection interface of the chest pain alert app.
[PNG File, 1506 KB - resprot_v15i1e90144_app1.png]

Multimedia Appendix 2
In-hospital collaborative communication interface (WeChat work group).
[PNG File, 1195 KB - resprot_v15i1e90144_app2.png]

Multimedia Appendix 3
Data extraction form and coding rules.
[DOCX File, 38 KB - resprot_v15i1e90144_app3.docx]

Multimedia Appendix 4
Ethical approval letter of this study (2025-KY-116-02).
[PDF File, 428 KB - resprot_v15i1e90144_app4.PDF]

References

1. Li Y, Han YL. Interpretation of the 2025 ACC/AHA/ACEP/NAEMSP/SCAI guideline for the management of patients with acute coronary syndromes. Zhonghua Xin Xue Guan Bing Za Zhi 2025 Dec 24;53(12):1428-1432. [doi: 10.3760/cma.j.cn112148-20250429-00322] [Medline: 41402108]
2. do Nascimento IJB, Lopes RD, De Lima MVM, et al. Coevolutionary analysis of evidence and recommendations in STEMI clinical practice guidelines: a 33-year meta-research study of ACC, AHA, and ESC. Am Heart J 2026 Apr;294:107326. [doi: 10.1016/j.ahj.2025.107326] [Medline: 41421702]
3. Yiadom M, Gong W, Patterson BW, et al. Influence of time-to-diagnosis on time-to-percutaneous coronary intervention for emergency department ST-elevation myocardial infarction patients: time-to-electrocardiogram matters. J Am Coll Emerg Physicians Open 2024 Jun;5(3):e13174. [doi: 10.1002/emp2.13174] [Medline: 38726468]
4. De Luca G, Suryapranata H, Ottervanger JP, Antman EM. Time delay to treatment and mortality in primary angioplasty for acute myocardial infarction: every minute of delay counts. Circulation 2004 Mar 16;109(10):1223-1225. [doi: 10.1161/01.CIR.0000121424.76486.20] [Medline: 15007008]
5. Cullen KJ, Mir H, Natarajan MK, et al. Exploring the barriers and facilitators to implementing a smartphone app for physicians to improve the management of acute myocardial infarctions: multicenter, mixed methods, observational study. JMIR Mhealth Uhealth 2025 Jul 8;13:e60173. [doi: 10.2196/60173] [Medline: 40627818]
6. Gibson W, Al Kindi D, Akl E, et al. Impact of smartphone apps on reperfusion times and clinical outcomes in acute ST-segment elevation myocardial infarction: systematic review and meta-analysis. JMIR Mhealth Uhealth 2025 Aug 25;13:e66605. [doi: 10.2196/66605] [Medline: 40854299]
7. Mir H, Cullen KJ, Mosleh K, et al. Smartphone app for prehospital ECG transmission in ST-elevation myocardial infarction activation: protocol for a mixed methods study. JMIR Res Protoc 2024 Sep 6;13:e55506. [doi: 10.2196/55506] [Medline: 39240681]
8. Polak G. Percutaneous coronary interventions in nonagenarians: single-centre insights. J Clin Med 2025 Oct 18;14(20):7371. [doi: 10.3390/jcm14207371] [Medline: 41156244]
9. Zahran A, Milhem F, Bdair M, et al. Interhospital transfer versus direct admission for percutaneous coronary intervention in patients with acute ST-segment elevation myocardial infarction: a systematic review and meta-analysis. Clin Res Cardiol 2025 Dec 15. [doi: 10.1007/s00392-025-02814-1] [Medline: 41396303]
10. Bernal JL, Cummins S, Gasparrini A. Interrupted time series regression for the evaluation of public health interventions: a tutorial. Int J Epidemiol 2017;46(1):348-355. [doi: 10.1093/ije/dyw098]
11. Faour A, Pahn R, Cherrett C, et al. Late outcomes of patients with prehospital ST-segment elevation and appropriate cardiac catheterization laboratory nonactivation. J Am Heart Assoc 2022 Jul 5;11(13):e025602. [doi: 10.1161/JAHA.121.025602] [Medline: 35766276]
12. Zhou Y, Du G, Zhang Y, et al. Construction and validation of a prediction model for Killip classes II-IV during hospitalisation in patients with acute ST-segment elevation myocardial infarction. Rev Cardiovasc Med 2025 Sep;26(9):38402. [doi: 10.31083/RCM38402] [Medline: 41089804]

XSL•FO
RenderX

13. Xu D, Liu Y, Ye J, et al. Evaluating the predictive value of the CALLY index for MACE events in STEMI patients: a comparative analysis across admission and discharge time points. J Cardiothorac Surg 2025 Nov 28;21(1):4. [doi: 10.1186/s13019-025-03763-6] [Medline: 41316252]
14. Song Z, Danzeng C, Jiang Y, et al. Association between admission blood pressure and in-hospital mortality and long-term mortality of patients with ST-elevation myocardial infarction undergoing percutaneous coronary intervention: a China Acute Myocardial Infarction registry study. Rev Cardiovasc Med 2025 Aug;26(8):33512. [doi: 10.31083/RCM33512] [Medline: 40927077]
15. Shojaei Y, Sadeghi M, Mehrabani K, et al. Long-term all-cause mortality rate after ST-elevation myocardial infarction and its predictors: ST Elevation Myocardial Infarction Cohort in Isfahan Study. J Res Med Sci 2025;30:53. [doi: 10.4103/jrms.jrms_314_24] [Medline: 41306623]
16. Sun Y, Wu X, Li L, Li T, Wang Z, Yu W. Force reversal during systolic-diastolic transition provides incremental prognostic value over LVEF for heart failure after STEMI. J Clin Med 2025 Nov 11;14(22):7978. [doi: 10.3390/jcm14227978] [Medline: 41303013]
17. Scott G, Olola C, Althoff J, et al. Prehospital aspirin delivery: emergency medical dispatcher-directed versus emergency medical services field provider-directed aspirin administration. Prehosp Disaster Med 2025 Oct;40(5):251-258. [doi: 10.1017/S1049023X25101490] [Medline: 41185427]
18. Huang JF, Zhang SQ, Hsiao YT, Yeh YN, Lin JC, Tsai MJ. Impact of atypical presenting symptoms on door-to-balloon time and mortality outcomes in ST-segment elevation myocardial infarction. BMC Emerg Med 2025 Oct 21;25(1):207. [doi: 10.1186/s12873-025-01368-9] [Medline: 41120915]
19. Sethupathi P, Matetić A, Bang V, et al. Association of diabetes mellitus and its types with in-hospital management and outcomes of patients with acute myocardial infarction. Cardiovasc Revasc Med 2023 Jul;52:16-22. [doi: 10.1016/j.carrev.2023.02.008] [Medline: 36854639]
20. Xie J, Song L, Yang Z, Bai H, Feng C. Development of a risk factor nomogram prediction model for patients with acute coronary syndrome complicated by hypertension using LASSO regression analysis. BMC Cardiovasc Disord 2025 Dec 12;25(1):866. [doi: 10.1186/s12872-025-05317-z] [Medline: 41388248]
21. Measures for the ethical review of biomedical research involving humans. National Health Commission of the People's Republic of China. 2022. URL: https://www.nhc.gov.cn/wjw/c100221/202201/558653bd3d1e4a078904152a30605a03.shtml [accessed 2026-04-02]
22. Nam J, Caners K, Bowen JM, Welsford M, O'Reilly D. Systematic review and meta-analysis of the benefits of out-of-hospital 12-lead ECG and advance notification in ST-segment elevation myocardial infarction patients. Ann Emerg Med 2014 Aug;64(2):176-186. [doi: 10.1016/j.annemergmed.2013.11.016] [Medline: 24368054]
23. Accorsi TAD, Pitta FG, Rompkoski J, et al. Evaluation of AI-enhanced tele-ECG response time and diagnosis in acute chest pain patients. Front Cardiovasc Med 2025;12:1532770. [doi: 10.3389/fcvm.2025.1532770] [Medline: 41357081]

Abbreviations

D2P: door-to-preactivation time
D2W: door-to-wire time
ECG: electrocardiogram
ED: emergency department
ITS: interrupted time series
PCI: percutaneous coronary intervention
PSM: propensity score matching
STEMI: ST-elevation myocardial infarction

Edited by J Sarvestan; submitted 22.Dec.2025; peer-reviewed by S Mirji; revised version received 07.Mar.2026; accepted 09.Mar.2026; published 13.Apr.2026.

Please cite as:
Chen S, Wang R, Ma J, Xiao W, Song C
Impact of a Prehospital Chest Pain Alert App–Mediated Prehospital–in-Hospital Coordination Model on Treatment Delays and Clinical Outcomes in Patients With ST-Elevation Myocardial Infarction: Protocol for a 4-Year Retrospective Real-World Cohort Study
JMIR Res Protoc 2026;15:e90144
URL: https://www.researchprotocols.org/2026/1/e90144
doi:10.2196/90144

XSL•FO
RenderX

© Shuyuan Chen, Rensong Wang, Jiangwei Ma, Weichun Xiao, Changxin Song. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Enhancing Continuous Medication Safety Through e-Prescription and Clinical Decision Support Systems in Outpatient Practices and Pharmacies: Protocol for a Multiperspective Study (eRIKA Study)

Juliana Schmidt[1], MSc; David Lampe[1], MSc; Adriana Poppe[2], MSc; Ingo Meyer[2], MA; Sara Söling[3], PhD; Juliane Köberlein-Neu[3], Prof Dr; Daniel Grandt[4], Prof Dr Med; Lara Düvel[5], MA; Wolfgang Greiner[1], Prof Dr; eRIKA Study Group[6,7,7]

[1]Department of Health Economics and Health Care Management, School of Public Health, Bielefeld University, Universitaetsstrasse 25, Bielefeld, North Rhine-Westphalia, Germany

[2]PMV Research Group, Medical Faculty and University Hospital Cologne, University of Cologne, Cologne, North Rhine-Westphalia, Germany

[3]Centre for Health Economics and Health Services Research, Schumpeter School of Business and Economics, University of Wuppertal, Wuppertal, North Rhine-Westphalia, Germany

[4]Department of Internal Medicine, Klinikum Saarbrücken, Saarbrücken, Saarland, Germany

[5]Department Digital Care/Prevention, BARMER GEK, Wuppertal, North Rhine-Westphalia, Germany

[6]BARMER GEK, Wuppertal, North Rhine-Westphalia, Germany

[7]

Corresponding Author:
Juliana Schmidt, MSc
Department of Health Economics and Health Care Management, School of Public Health, Bielefeld University, Universitaetsstrasse 25, Bielefeld, North Rhine-Westphalia, Germany

Abstract

Background: Increased life expectancy is associated with increasing multimorbidity and polypharmacy, leading to a heightened risk of drug-drug interactions and adverse events, especially when multiple health care providers are involved. To address the urgent need for safer medication management in this population, tools such as medication plans (MP), electronic prescriptions (e-prescriptions), and clinical decision support systems (CDSS) offer valuable support. These instruments have the potential to enhance medication safety by providing physicians and pharmacists with a comprehensive overview of a patient's overall medication regimen and by assisting health care professionals in making informed prescribing decisions.

Objective: This study aims to improve medication therapy safety by combining e-prescriptions, the use of claims data, MPs, CDSS, and interprofessional communication. To comprehensively evaluate this complex intervention, a holistic multiphase study will be conducted, examining (1) the effectiveness of the intervention and (2) health-economic and (3) implementation-related aspects.

Methods: A multiphase study design is used. In the first phase, the intervention is implemented in selected outpatient practices (n=10) and pharmacies (n=10) in 2 regions in Germany as part of a cluster-randomized controlled trial to assess process-related outcomes. The primary outcome is the congruence between the MP and claims data. In phase 2, the intervention is scaled up in 3 regions and evaluated in a quasi-experimental study. The required sample size for the intervention group is 3528 patients, with a synthetic control group matched from existing claims data. The primary outcome is a combined end point of all-cause mortality and hospitalization within 3 months of an index prescription. Quantitative methods (descriptive, regression-based methods using claims data, calculation of the incremental cost-effectiveness ratio, and survey-based analyses of implementation-related aspects) and qualitative methods (interviews and focus groups to capture experiences of health care professionals and patients) are used.

Results: In phase 1, a total of 187 patients were recruited (74 in the intervention group and 113 in the control group) by June 2025. Phase 2 is currently ongoing, with data collection continuing through December 31, 2025. Final analyses are planned by March 2027.

Conclusions: Medication safety in polypharmacy remains a critical challenge in Germany. This study provides multiperspective evidence supporting the nationwide implementation of the eRIKA (e-prescription as an element of interprofessional care pathways for continuous medication therapy management [*eRezept als Element interprofessioneller Versorgungspfade für kontinuierliche AMTS*]) intervention.

XSL•FO
RenderX

Trial Registration: Phase 1: German Clinical Trials Register DRKS00032777; https://www.drks.de/search/de/trial/DRKS00032777/details and Phase 2: German Clinical Trials Register DRKS00035030; https://www.drks.de/search/de/trial/DRKS00035030/details

International Registered Report Identifier (IRRID): DERR1-10.2196/87277

(JMIR Res Protoc 2026;15:e87277) doi:10.2196/87277

KEYWORDS

patient safety; medication therapy management; outpatient; clinical decision support system; polypharmacy; outcome and process evaluation; multiperspective; quasi-experimental design

Introduction

Background

Recent demographic trends project a continued increase in the number of older adults [1]. This is closely linked to a growing global prevalence of chronic diseases [2], often accompanied by complex challenges, such as multimorbidity [3] and polypharmacy [4]. Polypharmacy, commonly defined as the concurrent use of 5 or more medications, is associated with various adverse outcomes [5,6], such as unwanted drug-drug interactions (DDIs), potentially leading to adverse drug events [7]. DDIs are even more likely if multiple providers are involved in prescribing medications to patients with polypharmacy [8]. For instance, a German study found that only 34.7% of patients treated with prescription medications receive their medications from a single physician [7], underscoring the challenges of managing chronic diseases.

A key challenge in managing polypharmacy is the health care providers' lack of comprehensive knowledge of the patients' current medications. This knowledge gap compromises both the safety and effectiveness of pharmacological treatment. To address this, the implementation of comprehensive medication plans (MPs) encompassing prescribed and over-the-counter (OTC) medications in a structured and accessible format for both patients and providers has been widely recommended [9,10]. Empirical evidence supports the positive impact of MPs on medication safety. For example, the use of MPs has been associated with reduced hospitalization rates and fewer incidents of medication misuse [11], as well as a slight decrease in negative DDIs [12].

Despite these benefits, the adoption and consistent use of MPs is still insufficient [7,13-15]. A major barrier is the frequent incompleteness of MPs. For example, a study conducted in Germany found that, in a regional sample, only 6.3% of the nationwide MPs met all formal requirements, including active substance, dosage, or reasoning [9]. Moreover, evaluations of the MP in Germany reveal discrepancies between documented and actual medication use in approximately 78% of cases, particularly regarding OTC medication and dietary supplements [16].

Another intervention promising improvement in medication therapy management is electronic prescriptions (e-prescriptions). They offer a way to significantly contribute to improvements, such as reducing the risk of medication errors and adverse events as well as improving efficiency in medication management if well designed [17,18]. The full potential of MPs and e-prescriptions can be realized when integrated into clinical decision support systems (CDSS). The benefits of CDSS in clinical practice have already been demonstrated in a number of disease conditions [19-22], and their application in medication safety is particularly promising. For example, CDSS can identify DDIs overlooked by physicians, thereby improving the quality of care [22]. Furthermore, CDSS may prompt physicians to engage in more critical and reflective prescribing practices [23,24].

Objectives

The objective of the eRIKA study (e-prescription as an element of interprofessional care pathways for continuous medication therapy management [*eRezept als Element interprofessioneller Versorgungspfade für kontinuierliche AMTS*]) is to improve the safety of medication therapy through the combined use of e-prescriptions, claims data, MPs, CDSS, and interprofessional communication. To comprehensively evaluate the intervention, the study aims to investigate (1) its effectiveness as well as (2) health-economic and (3) implementation-related aspects.

Methods

Overview

This study protocol was written in accordance with the TIDieR (Template for Intervention Description and Replication) reporting checklist (Checklist 1) [25]. To comprehensively evaluate this complex intervention, a mixed perspective, multiphase study design will be employed. Phase 1 is designed as a cluster randomized controlled trial (c-RCT) with a relatively small sample size to allow for initial evaluation under controlled conditions. This is followed by a second phase involving a larger-scale quasi-experimental study with a matched control group (CG), scaling up the intervention and enabling the examination of the intervention's effectiveness, cost-effectiveness, and implementation-related aspects in real-world settings.

Intervention

The eRIKA intervention is designed to guarantee that for every patient receiving pharmacotherapy, an up-to-date and comprehensive electronic MP is generated, updated, and centrally stored, which can be accessed by patients, treating outpatient physicians, and pharmacists. Furthermore, medical information on the patient is electronically extracted from claims data provided by statutory health insurance (SHI) funds and is used to inform outpatient physicians about diagnoses, diagnostic procedures, and treatment history to date.

XSL•FO
RenderX

The eRIKA intervention is tested and implemented in participating outpatient practices and pharmacies. It is a digital software solution that integrates (1) the use of claims data in clinical practice, (2) the generation of an MP, (3) CDSS, (4) e-prescriptions, and (5) interprofessional communication.

Claims data from the SHI funds concerning prior prescriptions and selected clinically relevant information (such as diagnoses, medication prescription history, performed diagnostic and therapeutic procedures, treating outpatient physicians, hospital admissions and treatment, remedies, and aids) are fed into the software to ensure that MPs are complete and up to date at all times. This provides outpatient physicians with a comprehensive overview of the patient's medical and medication history, which enables more efficient and safer medication prescriptions. The provision of medical and pharmaceutical information ensures that pharmacists have an overview of the patient's current and past medications during their pharmaceutical consultations.

By using e-prescriptions via the eRIKA software and claims data, an MP is prefilled to be checked and manually supplemented with OTC medication, thus providing an up-to-date overview of the current medication regimens for use in outpatient practices and pharmacies. Furthermore, this MP is made accessible to patients, providing them with comprehensive information concerning their medication regimen.

Regarding CDSS, the eRIKA software also provides information and alerts for avoidable risks in medication therapy, for example, on contraindications, dosing errors, potentially inadequate medication for age, medication interactions, Dear-Doctor letters (*Rote-Hand-Briefe*), as well as guidance based on the guidelines for drug therapy in multimorbidity (*Leitlinie Arzneimitteltherapie bei Multimorbidität*), all of which are relevant to prescription decisions. Alerts are displayed in real time during prescription entry by the physician and are triggered automatically when rule-based checks detect a mismatch between the prescribed medication and patient-specific characteristics such as age, diagnoses, or concomitant medication. Alerts are categorized by severity (levels 1 - 3) and can be addressed by modifying or canceling the prescription. Alternatively, physicians may override alerts if they consider the prescription clinically appropriate. For high-severity alerts (levels 2 and 3), physicians are required to actively acknowledge the alert and provide a brief justification, whereas lower-severity alerts may be acknowledged without justification. All alerts, physician responses, and overrides are automatically documented within the system, enabling assessment of alert exposure and handling during the intervention. A tailored subset of alerts relevant for pharmaceutical counseling is presented to pharmacists.

e-Prescriptions via the eRIKA software are issued in accordance with the legal requirements under German law.

Regarding interprofessional communication, the software further enables the outpatient physician to document remarks regarding the selected medication and assign corresponding levels of urgency.

An overview of the structured eRIKA process is provided in Figure 1. Following patient consent (1), outpatient physicians use the eRIKA software during routine consultations to access claims-based medical information and the centrally stored MP. After conducting the anamnesis and reviewing the current medication (2.1), supported by CDSS alerts (2.2), new medications are prescribed electronically (2.3). Issued e-prescriptions and imported claims data automatically update the server-side MP (3.1). Furthermore, an up-to-date MP is provided to the patient as a printout (3.2). Next, patients redeem their e-prescriptions in participating pharmacies (4). In this step, pharmacists access the MP and relevant patient information (5.1) and add OTC medications and dietary supplements to the MP if necessary (5.2). The pharmacist-led structured pharmacotherapy safety check is supported by a tailored set of alerts relevant to pharmaceutical counseling (5.3). Pharmacists provide medication counseling (5.4) and, if necessary, contact the treating physician (electronically or via telephone) regarding medication-related remarks or substitution proposals for review (5.5). Following optional physician review, the MP is updated accordingly and made available to all involved health care professionals via the eRIKA software as the server is updated automatically (6), as well as to patients in printed form (7).

XSL•FO
RenderX

Figure 1. Functionalities and processes of the eRIKA (e-prescription as an element of interprofessional care pathways for continuous medication therapy management [*eRezept als Element interprofessioneller Versorgungspfade für kontinuierliche AMTS*]) intervention including the German telematics infrastructure, physician-led medication review, server-sided updated medication plans, and pharmacist-led medication review. ICD: International Statistical Classification of Diseases and Related Health Problems; OTC: over-the-counter.

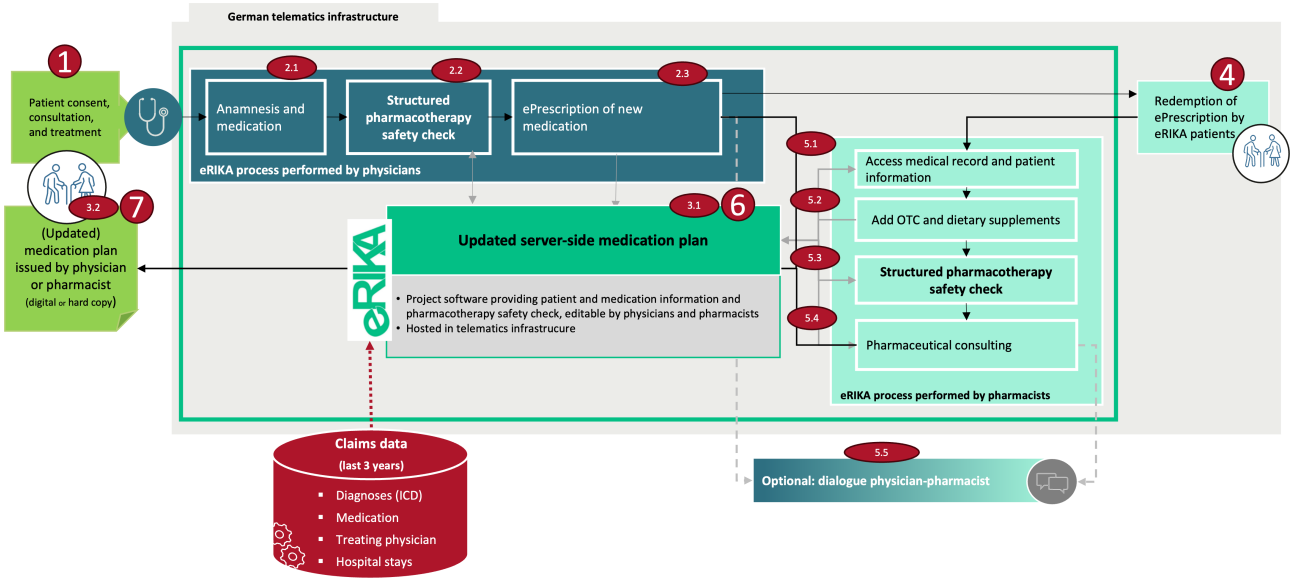

Reimbursement is tied to the completion of defined process steps to support adherence to the intended process. Physicians had to fulfill the following five steps: patient enrollment (1); conducting the anamnesis (2.1) and a structured safety check (2.2) and issuing e-prescriptions (2.3); providing an updated MP (3.1 and 3.2); pharmacists dispensing at least 1 e-prescription (4); and reviewing the corresponding risk analysis (5.1), documenting any required modifications to the MP (5.3), and issuing an updated MP if necessary (7).

Physicians and pharmacists received initial training sessions during which the software developer explained the system's functionality, providing opportunities to ask questions and report technical issues. In addition, a brief process manual was integrated directly into the software for ongoing reference.

Phase 1: Cluster Randomized Controlled Trial

Study Design

Process-related outputs and outcomes are evaluated by a c-RCT from October 2023 to June 2024 following the corresponding extended CONSORT (Consolidated Standards of Reporting Trials) guidelines [26]. Randomization is performed at the cluster level, with each cluster comprising an outpatient practice and at least one geographically proximate pharmacy, assigned to either the intervention group (IG) or the CG. Patients in clusters assigned to the IG are treated with the eRIKA intervention, while patients in the CG receive treatment as usual. Due to the nature of the intervention, blinding is not feasible; thus, health care providers, patients, and outcome analysts are all aware of group allocation.

Eligibility Criteria

Patients (1) aged 18 years or older, (2) insured by a participating SHI fund (BARMER, AOK Nordost), and (3) taking 3 or more long-term prescription medications are included in the study. This includes patients receiving their third medication prescription during the outpatient consultation leading to study participation, thus fulfilling the third criterion from that point onward. Patients participating in the eLiSa (electronic Life Saver) project are excluded from the study. No further specific exclusion criteria are applied.

Outpatient practices are included. No further inclusion or exclusion criteria on the outpatient practice and pharmacy level are applied.

Study Setting and Recruitment

The eRIKA intervention is implemented in selected outpatient practices and pharmacies in 2 regions (Berlin and Westphalia-Lippe). The outpatient practices are selected and approached via the Associations of Statutory Health Insurance Physicians Westphalia-Lippe and the Association of Berlin Physician Networks (*Arbeitsgemeinschaft Berliner Arztnetze*) for the Berlin area. Pharmacies are recruited via the SHI funds in cooperation with regional pharmacist associations. Patients are recruited by their treating outpatient physician.

Outcomes and Measures

Primary Outcome

The primary outcome of phase 1 is the congruence of the MP and claims data (OTC not considered), which is operationalized as the ratio of the medications listed in the MP and in claims data (numerator) to all documented medications according to claims data or MP (denominator). Thus, values range between 0 and 1, with a higher value being associated with a greater alignment between the MP and the medications documented in claims data.

Additional Evaluations

The primary end point is complemented by a process evaluation and a socioeconomic impact assessment (SEIA), each playing a key role in the sustainable adaptation and implementation of the intervention in phase 2 and beyond the study (Table 1).

XSL•FO
RenderX

Table . Overview of additional formative and socioeconomic evaluation of the eRIKA (e-prescription as an element of interprofessional care pathways for continuous medication therapy management [*eRezept als Element interprofessioneller Versorgungspfade für kontinuierliche AMTS*]) intervention in phase 1, including outcomes, data sources, and assessment time points.

Outcome	Data source	Time point
Formative evaluation		
Acceptability, appropriateness, and feasibility [27]	Quantitative data from surveys	Baseline and after 6 months
Adoptability, implementability, and sustainability	Quantitative data from surveys	Baseline and after 6 months
Context characteristics	Quantitative data from surveys	Baseline and after 6 months
Barriers and facilitating factors	Qualitative data from interviews	End of the c-RCT[a]
Intervention fidelity	User data of the software	Continuously over the study period
Socioeconomic impact assessment		
The time needed for implementation correlates positively with expected benefits	Literature research, MP[b], claims data, primary data, and software data	Patient-specific

[a]c-RCT: cluster randomized controlled trial.

[b]MP: medication plan.

Data Collection

Primary Data

Once the patient has provided informed consent, the MP is transmitted to a trust center. In the IG, the software provider extracts the MPs from the software solution and transmits them electronically to the trust center, whereas in the CG, the treating outpatient physician sends the MP via postal mail. The trust center digitizes the MP (for the CG) and subsequently links it with the software and claims data.

Outcomes related to the formative evaluation are obtained from participating outpatient physicians and pharmacists from both study groups through 2 surveys (baseline and after 6 months) and qualitative interviews.

Claims Data

Claims data from September 2023 to December 2024 are used in the analysis. This includes data for the duration of the c-RCT, as well as a preobservation period of 1 year to establish a baseline for the primary end point. The required claims data from participating SHI funds are specified in a coordinated minimal dataset.

Software Data

For the data derived from the software solution used to evaluate the primary end point, the SEIA, and user behavior and interaction with the software (formative evaluation), a minimal data set is coordinated with the software provider.

Data Linkage

Data are linked by the trust center. Claims data and software data are linked with the patients' MP data based on a pseudonymized study ID.

Sample Size

For the evaluation of the primary outcome, 10 clusters of outpatient practices and pharmacies in geographical proximity are randomized into 2 groups (5 IG and 5 CG). Because, to our knowledge, no comparable studies are available to support plausible assumptions regarding the primary outcome, a moderate Cohen *d* effect size of 0.5 is assumed, in line with established conventions for studies with novel outcomes [28,29]. The assumed intracluster correlation coefficient (ICC) of 0.05 is based on the evidence from c-RCTs conducted in similar primary care and outpatient settings, where ICCs typically range from 0.01 to 0.10 [30-32]. Given these assumptions, an α-error of .05 (2-sided), and a power of 0.8, the overall sample size is 138 (2×69). Based on experience from prior studies in comparable outpatient and primary care settings [33,34], as well as to account for possible patient withdrawal and loss of analyzable primary outcome data due to incomplete or manually transmitted MPs, a dropout rate of 40% is expected, increasing the sample size to 230 patients (2×115). The resulting number of patients per practice is 23.

Analyses

Primary Outcome

The evaluation of the primary outcome is initially descriptive. In the next step, appropriate inferential statistical methods (eg, 2-tailed *t* tests, linear mixed models [LMMs], generalized LMMs [GLMMs], or generalized estimating equations to control for potential confounding factors, cluster effects via random intercepts, and, if applicable, repeated measures within individuals) are applied [35]. Analysis is adjusted for multiple testing using the Bonferroni-Holm correction.

Process Evaluation

All process evaluation data are initially analyzed descriptively and exploratively. Qualitative data material is analyzed by combining a broadly deductive analytical approach with an inductive analysis to explore new findings related to the specific intervention and context.

Socioeconomic Impact Assessment

Sustainability planning for the integration of the intervention into standard care is informed by a SEIA. It is a formative

XSL•FO
RenderX

evaluation that aims to answer questions on 3 levels that are visualized in Figure 2.

Figure 2. Levels of the socioeconomic impact assessment (SEIA), illustrating policy, service, and individual or organisational perspectives, along with corresponding evaluation questions. ROI: return on investment; SER: socioeconomic return.

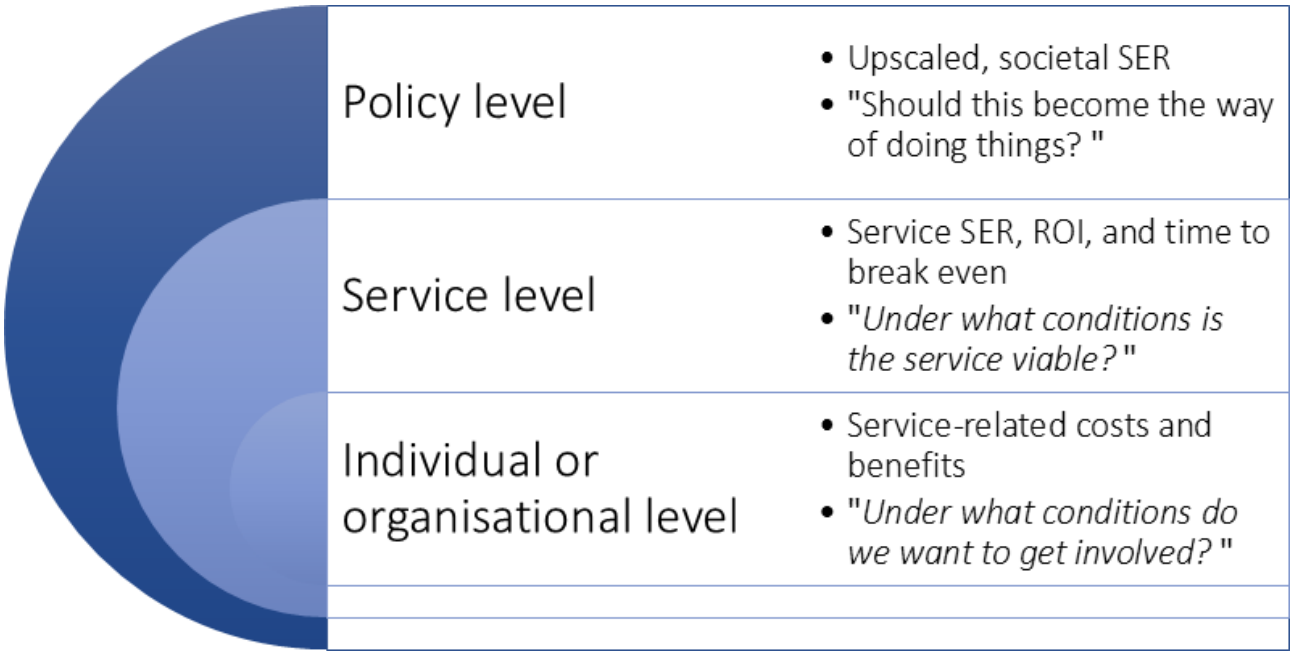

The SEIA methodology is based on cost-benefit analysis as defined by Drummond et al [36] and the recommendations of the UK HM Treasury [37], the Federal Government Commissioner for Information Technology [38], and the White House Office for Management and Budget [39]. An already established framework and associated evaluation software, developed for business model development for IT-based utility services, is applied [40].

Phase 2: Quasi-Experimental Study

Study Design

A quasi-experimental study with an intervention or observational period of 15 months is conducted from July 2024 to December 2025. Eligible IG patients are recruited in the field and treated with the eRIKA intervention, while the CG receives treatment as usual and is derived from claims data using appropriate matching methods. Neither health care providers, patients, nor analysts are blinded to treatment allocation.

Eligibility Criteria

The study population in phase 2 is selected according to the eligibility criteria applied in phase 1. In addition, phase 2 includes a separate study arm comprising patients who take 5 or more long-term prescription medications.

Study Setting and Recruitment

The eRIKA intervention is implemented in participating outpatient practices and pharmacies in 3 regions (Berlin, Saarland, and Westphalia-Lippe). Recruitment of IG outpatient practices and pharmacies is carried out by the collaborating regional Associations of Statutory Health Insurance Physicians, the Association of Berlin Physician Networks (*Arbeitsgemeinschaft Berliner Arztnetze*), and regional pharmacist associations. The University of Wuppertal contacts participating outpatient practices and pharmacies for surveys and qualitative interviews. Outpatient practices and pharmacies involved in phase 1 (IG and CG) are invited to participate in the IG in phase 2. IG patients are recruited by their treating outpatient physician. Recruitment monitoring is implemented in both study phases to enable early identification of deviations and the timely initiation of supportive recruitment measures. The IG is assigned a data-derived CG through the application of appropriate matching methods.

Outcomes and Measures

Primary Outcome

The primary outcome of phase 2 is a combined end point of the incidence of all-cause mortality or all-cause hospitalization in patients with polypharmacy 3 months after an index prescription at baseline.

Additional Evaluations

With the secondary outcomes in phase 2, the eRIKA intervention is evaluated in terms of a summative, socioeconomic impact and health economic evaluation, partly building on the outcomes from phase 1 (SEIA). A comprehensive overview of the secondary outcomes and aspects of the process evaluation in phase 2 is shown in Table 2.

XSL•FO
RenderX

Table . Overview of additional processes, secondary summative outcomes, and socioeconomic evaluations of eRIKA (e-prescription as an element of interprofessional care pathways for continuous medication therapy management [*eRezept als Element interprofessioneller Versorgungspfade für kontinuierliche AMTS*]) in phase 2, including outcomes, data sources, and assessment time points.

Outcome	Data source	Time point or period
Process evaluation		
Acceptability, appropriateness, and feasibility [27]	Quantitative data from surveys	Provider-specific: at baseline, after 3 months and up to 12 months after initial implementation
Adoption, penetration, and sustainability	Quantitative data from surveys	Provider-specific: at baseline, after 3 months and up to 12 months after initial implementation
Context characteristics	Quantitative data from surveys	Provider-specific: at baseline, after 3 months and up to 12 months after initial implementation
Interprofessional collaboration	Qualitative data from interviews and focus groups	Provider-specific: After six months and at the end of phase 2
Intervention fidelity	User data of the software	Continuously over the study period
Secondary summative outcomes		
Congruence of the MP[a] and claims data	Claims data	Patient-specific: at baseline, follow-up quarter
Adverse drug event-related diagnoses and health care resource use (inpatient or outpatient)	Claims data and software data	Patient-specific: continuously over the study period
Potentially inadequate medications	Claims data and software data	Patient-specific: continuously over the study period
Socioeconomic impact assessment		
Socioeconomic return	Literature research, MP, claims data, primary data, and software data	Provider-specific: at baseline, after 3 months and patient-specific: observation period of 3 months
Health economic evaluation		
Health care resource use	Claims data	Patient-specific: observation period of 3 months
Health care costs	Claims data	Patient-specific: observation period of 3 months

[a]MP: medication plan.

Data Collection

Primary Data

Prior to the intervention, the treating outpatient physician transmits the MP to the trust center via postal mail. As part of the patient-specific follow-up assessment in the subsequent quarter, the software provider will extract the MP and transmit it electronically. The linkage and digitization of the MP are carried out in the trust center. MPs are not available for the matched CG.

The outcomes related to the process evaluation are collected from the participating outpatient physicians and pharmacists using a longitudinal survey (3 time points; Table 2), qualitative interviews (at least 10 professionals per region), and focus groups (2 focus groups with 8 - 10 professionals and interviews with at least 15 patients).

Claims Data

Although the intervention runs from July 2024 to December 2025, claims data from January 2023 to December 2025 is used in the analysis, covering the per patient–specific preintervention period as well as the intervention period, including the last patient follow-up of 3 months. In addition to the patient-specific observation period, this includes a preinterventional phase to adjust for preinterventional disease burden and costs. The required claims data from participating SHI funds are specified in a jointly defined minimal dataset.

Software Data

For selected secondary end points, such as the congruence between claims data and MP, as well as user behavior and interaction with the software (process evaluation), a minimal dataset is defined in collaboration with the software provider.

Data Linkage

Data linkage is performed in the same manner as in phase 1.

Sample Size

For the outcome evaluation, the incidence of all-cause hospitalization or all-cause mortality (primary outcome) for patients with polypharmacy is estimated to be 32.9%, based on a preliminary analysis of BARMER claims data. Given the inconsistent evidence regarding the impact of medication management interventions on outcomes such as hospitalization and mortality [11,33,41-45] and the fact that these outcomes are influenced by multiple clinical and contextual factors beyond the intervention itself, a conservative absolute risk reduction of 4 percentage points was assumed, corresponding to a reduction to 28.9%. The evaluation is carried out for 2 groups, each in comparison to a matched CG: for patients with (1) 3 to 4 medications and (2) 5 or more medications. With an α-error of .05 and a power of 1-β=0.80, based on a 2-sided Fisher exact

XSL•FO
RenderX

test (for binary outcomes), this yields a required sample size of 1411 patients per group, resulting in a total of 2822 patients. Accounting for a dropout rate of 20%, the required sample size is adjusted to 3528 participants.

Drawing on experience from previous projects, an average of 10 to 20 patients per practice is expected, corresponding to approximately 175 to 350 outpatient physicians and 116 to 233 outpatient practices and corresponding pharmacies.

The IG is assigned a data-derived CG by selecting and matching comparable insured individuals from the claims data provided by the participating SHI funds. These individuals meet the study's inclusion criteria and have not received treatment in any of the participating outpatient practices or pharmacies during the intervention phase. An iterative matching approach is employed, consisting of a primary procedure and alternative procedures. If the primary method of direct matching based on relevant matching criteria (eg, age, gender, and prescribed medications) does not yield a sufficiently large and balanced matched CG, an alternative procedure based on propensity score matching is applied, using relevant matching criteria as covariates. This approach aligns with standard practice in the application of matching techniques [46], as there is no single optimal choice of matching approach for any given analysis [47]. To assess residual confounding after matching, covariate balance between the IG and the matched CG is evaluated using standardized mean differences and visual diagnostics for all matching variables and key preintervention characteristics (eg, health care use and comorbidity indicators) [48,49]. To address any residual imbalance, outcome analyses are conducted using regression models adjusted for baseline covariates that remain imbalanced after matching. As a sensitivity analysis, alternative weighting approaches are applied to assess the robustness of the results to the choice of matching method [50].

Analyses

Summative Evaluation

The evaluation uses claims data encompassing patients in the IG, as well as a synthetic CG. Primary and secondary end points are initially analyzed descriptively. Subsequently, inferential statistical analyses are conducted using LMMs, GLMMs, or generalized estimating equations, accounting for independent covariates such as age, gender, and comorbidity, as well as for repeated measurements within individuals and cluster effects via random intercepts [35]. Results are adjusted for multiple testing using the Bonferroni-Holm correction.

Health Economic Evaluation

Using claims data, the health economic evaluation is conducted on a per-patient basis from a third-party payer perspective (SHI in Germany), excluding societal costs such as productivity losses. The analysis encompasses health care use and costs related to outpatient, inpatient, rehabilitative and nursing or home care, pharmaceuticals, therapeutic devices, nonphysician specialist services, as well as patient transportation services and SHI-financed sick leave or pay. Intervention-related costs are also included. To quantify the impact of the intervention on total costs, appropriate descriptive methods (eg, mean, median, and SD) are used. As cost data are typically right-skewed and may include extreme values, appropriate regression-based methods such as LMM or GLMM with suitable distributional assumptions (eg, gamma with log-link or log-normal distribution) are applied. Extreme cost values are initially retained and checked for implausibility, and the influence of extreme observations is examined. Relevant variables (eg, prior-year costs) are included to control for potential confounding. To account for clustering of individuals within outpatient practices or pharmacies, cluster-level random intercepts will be incorporated into the model. Cost-effectiveness is determined by means of the incremental cost-effectiveness ratio, calculated by dividing the difference in costs by the difference in health benefit in the IG compared to CG. The effectiveness of the intervention is measured by the primary outcome (hospitalization and/or death). Appropriate sensitivity analyses (eg, deterministic or probabilistic) [51-54] are conducted to account for uncertainties and evaluate their impact on the results.

Process Evaluation

Initially, data are analyzed descriptively and exploratorily. For the qualitative data material, either a content analysis or a qualitative-descriptive analysis is used. A 2-stage procedure is applied. First, deductively, interview paraphrases are assigned to the predefined themes and subthemes of the underlying frameworks or (implementation) theories (eg, innovation readiness, consolidated framework for implementation research, and normalization process theory). Building on this, an inductive processing of the content will occur within these themes. Questionnaire and software data are analyzed descriptively and exploratorily.

Socioeconomic Impact Assessment

The sustainability planning for the transition from phase 1 to standard care continues in phase 2. This process follows the methodology outlined earlier (see *Socioeconomic Impact Assessment* under *Phase 1: Cluster Randomized Controlled Trial* section).

Ethical Considerations

For phase 1, ethical approval was granted by the ethics committee of the Medical Association of Westphalia-Lippe (2023 - 526-f-S). For phase 2, ethical approval was granted by the ethics committee of the Medical Association of Westphalia-Lippe (2024 - 599-f-S) and the Medical Association of Saarland (208/24).

This study is conducted in accordance with the Declaration of Helsinki. The CG in phase 2 is derived from secondary pseudonymized claims data. Informed consent from CG individuals is not obtained. However, the study poses minimal risk, and a waiver of informed consent for the use of pseudonymized secondary data was reviewed and approved by the responsible ethics committee as well as regulatory authorities. All data are collected, pseudonymized, and stored in accordance with German General Data Protection Regulation legislation based on the European General Data Protection Regulation.

XSL•FO
RenderX

Patients in phase 1 and IG patients in phase 2 voluntarily provide written informed consent prior to participation. Participants are informed about the project, the evaluation procedures (including data linkage of MP, software data, and claims data), participation requirements, and their rights. In phase 2, patients additionally consent to participate in interviews. By signing a separate consent form, they agree to be contacted by research staff. No financial compensation is provided to patients in the study.

Participating outpatient practices and pharmacies consent to participate through their project participation agreement, replacing the need for individual staff consent forms. Study staff are informed about study procedures prior to each data collection activity. By completing questionnaires or voluntarily participating in interviews or focus groups, staff consent to the project-related processing, use, and storage of their data.

All participants may withdraw their informed consent at any time without providing reasons and without facing any disadvantages. In the event of withdrawal, no further data will be collected, and previously collected data will be anonymized in accordance with data protection regulations.

Study data are stored in encrypted form with access restricted to authorized project staff. Personal contact data are stored separately from research data and deleted in accordance with data protection regulations.

The protocol was registered with the German Clinical Trials Register database (phase 1: DRKS00032777 and phase 2: DRKS00035030).

Results

The study is currently in phase 2. The data collection for phase 1 was completed by June 30, 2025. In total, 187 patients were successfully recruited for phase 1, comprising 74 individuals in the IG and 113 in the CG (CONSORT flowchart of the progress of clusters and individuals through the phases of a c-RCT in Figure 3). As of October 2025, 344 participants have been enrolled for phase 2. Data collection for this phase will continue through December 31, 2025. Due to a time lag in claims data provision, analyses for both phases will be completed by March 2027.

XSL•FO
RenderX

Figure 3. Flow diagram of the progress of clusters and individuals through the phases of the cluster randomized eRIKA (e-prescription as an element of interprofessional care pathways for continuous medication therapy management [*eRezept als Element interprofessioneller Versorgungspfade für kontinuierliche AMTS*] study (September 2025). tbd.: to be determined.

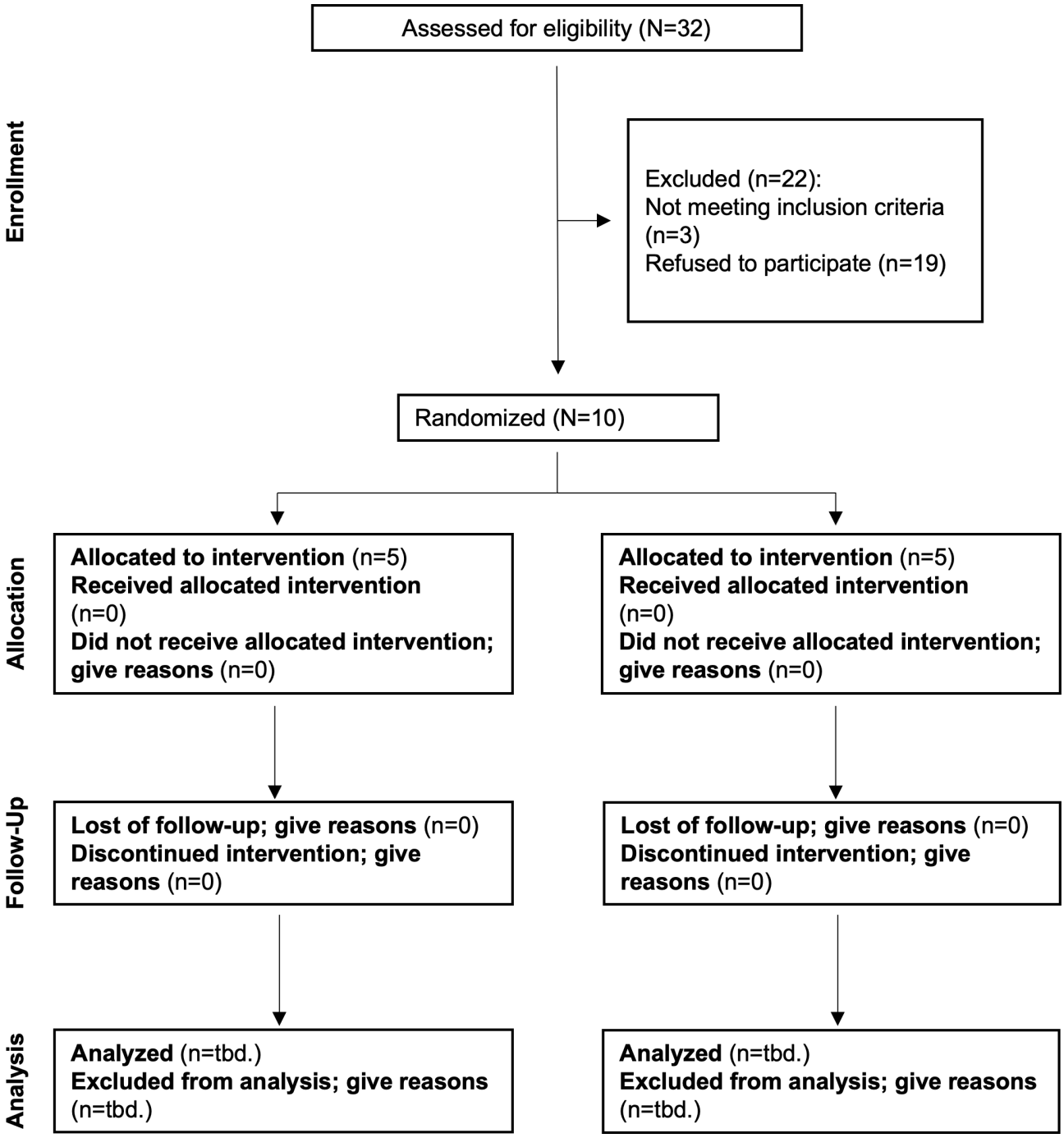

Discussion

Principal Findings

There is an urgent need to improve the safety of medication therapy for patients with polypharmacy. The growing number of older adults is accompanied by rising multimorbidity and polypharmacy [1-4], increasing the risk of DDIs [5-7], particularly when multiple providers are involved [8]. Meanwhile, MPs often remain incomplete or underutilized [7,13-15]. Structured MPs and CDSS have shown positive effects on medication therapy safety, such as reducing DDIs, hospitalizations, and inappropriate medication use [11,12]. However, the widespread implementation of MPs remains limited, highlighting the urgent need for better integration into routine care. The eRIKA software is a promising intervention to improve medication therapy safety for patients with polypharmacy. The combined use of e-prescriptions, claims data, automatically generated MPs, and CDSS has the potential to provide health care providers with an up-to-date overview of the current medication, contribute to the completeness and actuality of the available MP, and support outpatient physicians and pharmacists in their treatment or consultation process. The eRIKA intervention extends forms of care that have already been tested and evaluated in other projects such as the AdAM (Anwendung für digital unterstütztes

XSL•FO
RenderX

Arzneimitteltherapie-Management) [33] or TOP (Transsektorale Optimierung der Patientensicherheit) [34] project by adding the use of e-prescriptions and emphasizing the essential role of pharmacists.

Strengths

A key strength of the study is the alignment of the eRIKA project with the Medical Research Council (MRC) framework for the development and evaluation of complex interventions [55]. This includes a comprehensive approach encompassing intervention development, implementation, and multilevel evaluation, including formative components. Furthermore, the study is structured in 2 phases to ensure both rigor and scalability. Phase 1 is conducted as a c-RCT, allowing for a controlled assessment of the intervention's feasibility within a clinical setting. In phase 2, the study involves a larger sample size to evaluate the effectiveness and large-scale implementation. Moreover, with this comprehensive analysis approach, a multitude of relevant outcomes for the effectiveness of the eRIKA software and process are addressed, as recommended by the literature [56]. The intervention itself is intersectoral in nature, involving coordinated efforts across different sectors of the health care system, which enhances its practical relevance.

Limitations

The study faces limitations that may affect the interpretation and implementation of its findings. In phase 1, the small number of clusters may reduce the generalizability of the results. In addition, the primary outcome in phase 1 (congruence between MP and claims data) is novel, and therefore the underlying effect size and ICC assumptions are based on methodological conventions rather than outcome-specific empirical estimates. Furthermore, claims data are subject to delays and omissions and may not fully reflect actual medication use. Observed discrepancies between the MP and claims data may therefore result from timing differences rather than true inconsistencies, which should be considered when interpreting the congruence outcome in phase 1. Phase 2 is designed as a quasi-experimental study, which, while useful for evaluating interventions in more naturalistic settings, is not exempt from potential confounding and bias. Rigorous (matching) methods are necessary to mitigate such risks. Additionally, the implementation of the intersectoral interventions may present practical challenges, including reliance on service providers, variations in stakeholder engagement, coordination between sectors, technical challenges (at the point of installation or software usage), and contextual factors that could influence intervention fidelity and outcomes. Particularly the implementation of other digital interventions on a national level, such as e-prescriptions in general or the national rollout of electronic patient records, might hinder physicians' recruitment efforts and intervention fidelity. These factors will be considered in the formative evaluation as well as in the statistical analysis by using a time variable.

Conclusions

Improving medication safety for patients with polypharmacy remains a critical health care priority in Germany. This study provides a comprehensive, multiperspective evaluation of the eRIKA intervention, delivering robust evidence to support its nationwide implementation.

Acknowledgments

The authors sincerely thank Katrin Scholz, Peter Ihle, and Jutta Küpper-Nybelen for their invaluable contribution in serving as the trust center which was crucial to the success of this study. The authors would also like to thank all outpatient practices, pharmacies, and patients for their participation in the study. Generative artificial intelligence tools were used solely for language polishing and structural improvements under author supervision.

eRIKA Study Group: Franziska Folter (BARMER), Janet Kornek (BARMER), Till Beckmann (BARMER), Moritz Melcher (University of Wuppertal), Peter Ihle (PMV Research Group), Bettina Meenen (PMV Research Group), Elisabeth Grün (PMV Research Group), Sabrina Jojade (AOK Nordost), Natalie Sinnen (AOK Nordost), Thomas Wulle (AOK Nordost), Sebastian Gohlke (AOK Nordost), Martina Marx (AOK Nordost), Stephanie Sehlen (AOK Nordost), Olga Resch (AOK Nordost), Thomas Berghoff (Association of Statutory Health Insurance Physicians Westphalia-Lippe), Marco Luzius (Association of Statutory Health Insurance Physicians Westphalia-Lippe), Simone Grandt (RpDoc Solutions), and Christof Schelian (RpDoc Solutions).

Funding

This study was funded by the Innovation Fund of the German Federal Joint Committee (G-BA; grant: 01NVF21110).

Data Availability

The data collection forms and datasets generated and analyzed during this study are not publicly accessible. Participant consent agreements, in line with German data protection regulations, limit data use to the specific purposes of this project's evaluation and to members of the authorized research team. Therefore, data sharing with third parties or public repositories is not permitted.

Authors' Contributions

All members of the eRIKA study group have made substantial contributions to the conception of the eRIKA study. JS drafted the first version of the manuscript with input from DL. DL, AP, IM, SS, J Köberlein-Neu, LD, DG, and WG revised the manuscript critically for important intellectual content. Analysis and interpretation of data will be performed by JS, SS, DL, AP, IM, WG,

XSL•FO
RenderX

and J Köberlein-Neu. KS, PI, and J Küpper-Nybelen are responsible for data management and the trust center. WG is the chief investigator of the study. All authors reviewed, read, and approved the final manuscript.

Conflicts of Interest

DG reports grants from BARMER and AOK Nordost during the conduct of the study, and family members of DG work for and hold shares in the IT company involved in the project.

Checklist 1
TIDieR checklist.
[PDF File, 215 KB - resprot_v15i1e87277_app1.pdf]

References

1. Ausblick auf die Bevölkerungsentwicklung in Deutschland und den Bundesländern nach dem Corona-Jahr 2020 [Report in German]. : Statistisches Bundesamt; 2021 URL: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsvorausberechnung/Publikationen/Downloads-Vorausberechnung/bevoelkerung-deutschland-2035-5124202219004.pdf?__blob=publicationFile&utm_source=chatgpt.com [accessed 2026-03-17]
2. Charalampous P, Gorasso V, Plass D, et al. Burden of non-communicable disease studies in Europe: a systematic review of data sources and methodological choices. Eur J Public Health 2022 Apr 1;32(2):289-296. [doi: 10.1093/eurpub/ckab218] [Medline: 35015851]
3. Chowdhury SR, Das DC, Sunna TC, Beyene J, Hossain A. Global and regional prevalence of multimorbidity in the adult population in community settings: a systematic review and meta-analysis. EClinicalMedicine 2023 Mar;57:101860. [doi: 10.1016/j.eclinm.2023.101860] [Medline: 36864977]
4. Pazan F, Wehling M. Polypharmacy in older adults: a narrative review of definitions, epidemiology and consequences. Eur Geriatr Med 2021 Jun;12(3):443-452. [doi: 10.1007/s41999-021-00479-3] [Medline: 33694123]
5. Davies LE, Spiers G, Kingston A, Todd A, Adamson J, Hanratty B. Adverse outcomes of polypharmacy in older people: systematic review of reviews. J Am Med Dir Assoc 2020 Feb;21(2):181-187. [doi: 10.1016/j.jamda.2019.10.022] [Medline: 31926797]
6. Khezrian M, McNeil CJ, Murray AD, Myint PK. An overview of prevalence, determinants and health outcomes of polypharmacy. Ther Adv Drug Saf 2020;11(11):2042098620933741. [doi: 10.1177/2042098620933741] [Medline: 32587680]
7. Grandt D, Lappe V, Schubert I. BARMER Arzneimittelreport 2022 (Schriftenreihe zur Gesundheitsanalyse – Band 36) [Report in German]. : BARMER; 2022 URL: https://www.barmer.de/resource/blob/1138136/c95292f9376da884d3a5de65092c092f/download-report-data.pdf [accessed 2026-03-17]
8. Zazzara MB, Palmer K, Vetrano DL, Carfì A, Onder G. Adverse drug reactions in older adults: a narrative review of the literature. Eur Geriatr Med 2021 Jun;12(3):463-473. [doi: 10.1007/s41999-021-00481-9] [Medline: 33738772]
9. Dormann H, Maas R, Eickhoff C, et al. Der Bundeseinheitliche Medikationsplan in der Praxis [Article in German]. Bundesgesundheitsbl 2018 Sep;61(9):1093-1102. [doi: 10.1007/s00103-018-2789-9]
10. Mueller MA, Opitz R, Grandt D, Lehr T. The federal standard medication plan in practice: an observational cross-sectional study on prevalence and quality. Res Social Adm Pharm 2020 Oct;16(10):1370-1378. [doi: 10.1016/j.sapharm.2020.01.013] [Medline: 32061549]
11. Meid AD, Wirbka L, Moecker R, et al. Mortality and hospitalizations among patients enrolled in an interprofessional medication management program. Dtsch Arztebl Int 2023 Apr 14;120(15):253-260. [doi: 10.3238/arztebl.m2023.0014] [Medline: 37070272]
12. Rudolf H, Thiem U, Aust K, et al. Reduction of potentially inappropriate medication in the elderly. Dtsch Arztebl Int 2021 Dec 27;118(51-52):875-882. [doi: 10.3238/arztebl.m2021.0372] [Medline: 34939917]
13. Killin L, Hezam A, Anderson KK, Welk B. Advanced medication reconciliation: a systematic review of the impact on medication errors and adverse drug events associated with transitions of care. Jt Comm J Qual Patient Saf 2021 Jul;47(7):438-451. [doi: 10.1016/j.jcjq.2021.03.011] [Medline: 34103267]
14. Hammar T, Mzil L, Eiermann B. Discrepancies in patients' medication lists from pharmacies in Sweden: an interview study before the implementation of the Swedish National Medication List. Int J Clin Pharm 2023 Feb;45(1):88-96. [doi: 10.1007/s11096-022-01480-x] [Medline: 36307661]
15. Waltering I, Schwalbe O, Hempel G. Discrepancies on medication plans detected in German community pharmacies. J Eval Clin Pract 2015 Oct;21(5):886-892. [doi: 10.1111/jep.12395] [Medline: 26139566]
16. Amelung S, Bender B, Meid A, et al. Wie vollständig ist der Bundeseinheitliche Medikationsplan? Eine Analyse bei Krankenhausaufnahme [Article in German]. Dtsch Med Wochenschr 2020 Oct;145(21):e116-e122. [doi: 10.1055/a-1212-2836] [Medline: 33022741]

XSL•FO
RenderX

17. Osmani F, Arab-Zozani M, Shahali Z, Lotfi F. Evaluation of the effectiveness of electronic prescription in reducing medical and medical errors (systematic review study). Ann Pharm Fr 2023 May;81(3):433-445. [doi: 10.1016/j.pharma.2022.12.002] [Medline: 36513154]
18. Hareem A, Lee J, Stupans I, Park JS, Wang K. Benefits and barriers associated with e-prescribing in community pharmacy—a systematic review. Explor Res Clin Soc Pharm 2023 Dec;12:100375. [doi: 10.1016/j.rcsop.2023.100375] [Medline: 38145236]
19. Reis WC, Bonetti AF, Bottacin WE, et al. Impact on process results of clinical decision support systems (CDSSs) applied to medication use: overview of systematic reviews. Pharm Pract (Granada) 2017;15(4):1036. [doi: 10.18549/PharmPract.2017.04.1036] [Medline: 29317919]
20. Rittmann B, Stevens MP. Clinical decision support systems and their role in antibiotic stewardship: a systematic review. Curr Infect Dis Rep 2019 Jul 24;21(8):29. [doi: 10.1007/s11908-019-0683-8] [Medline: 31342180]
21. Kruse CS, Ehrbar N. Effects of computerized decision support systems on practitioner performance and patient outcomes: systematic review. JMIR Med Inform 2020 Aug 11;8(8):e17283. [doi: 10.2196/17283] [Medline: 32780714]
22. Schmidtchen L, Villis M, Christoph J, Rödle W. Usability analysis of a medication visualization tool for decision support. Stud Health Technol Inform 2023 Sep 12;307:110-116. [doi: 10.3233/SHTI230701] [Medline: 37697844]
23. Rieckert A, Teichmann AL, Drewelow E, et al. Reduction of inappropriate medication in older populations by electronic decision support (the PRIMA-eDS project): a survey of general practitioners' experiences. J Am Med Inform Assoc 2019 Nov 1;26(11):1323-1332. [doi: 10.1093/jamia/ocz104] [Medline: 31504572]
24. Söling S, Köberlein-Neu J, Müller BS, et al. From sensitization to adoption? A qualitative study of the implementation of a digitally supported intervention for clinical decision making in polypharmacy. Implement Sci 2020 Sep 21;15(1):82. [doi: 10.1186/s13012-020-01043-6] [Medline: 32958010]
25. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 7;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
26. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. CONSORT 2010 statement: extension to cluster randomised trials. BMJ 2012 Sep 4;345:e5661. [doi: 10.1136/bmj.e5661] [Medline: 22951546]
27. Kien C, Griebler U, Schultes MT, Thaler KJ, Stamm T. Psychometric testing of the German versions of three implementation outcome measures. Glob Implement Res Appl 2021 Sep;1(3):183-194. [doi: 10.1007/s43477-021-00019-y]
28. Sedlmeier P, Gigerenzer G. Do studies of statistical power have an effect on the power of studies? Psychol Bull 1989;105(2):309-316. [doi: 10.1037//0033-2909.105.2.309]
29. Schäfer T, Schwarz MA. The meaningfulness of effect sizes in psychological research: differences between sub-disciplines and the impact of potential biases. Front Psychol 2019;10:813. [doi: 10.3389/fpsyg.2019.00813] [Medline: 31031679]
30. Teerenstra S, Eldridge S, Graff M, de Hoop E, Borm GF. A simple sample size formula for analysis of covariance in cluster randomized trials. Stat Med 2012 Sep 10;31(20):2169-2178. [doi: 10.1002/sim.5352] [Medline: 22495809]
31. Chenot JF. Cluster-randomisierte Studien: eine wichtige Methode in der allgemeinmedizinischen Forschung [Article in German]. Z Evid Fortbild Qual Gesundhwes 2009 Jan;103(7):475-480. [doi: 10.1016/j.zefq.2009.07.004]
32. Rose O, Schaffert C, Czarnecki K, et al. Effect evaluation of an interprofessional medication therapy management approach for multimorbid patients in primary care: a cluster-randomized controlled trial in community care (WestGem study protocol). BMC Fam Pract 2015 Jul 22;16:84. [doi: 10.1186/s12875-015-0305-y] [Medline: 26198433]
33. Brünn R, Basten J, Lemke D, et al. Digital medication management in polypharmacy—findings of a cluster-randomized, controlled trial with a stepped-wedge design in primary care practices (AdAM) [Article in German]. Dtsch Arztebl Int 2024 Apr 19;121(8):243-250. [doi: 10.3238/arztebl.m2024.0007] [Medline: 38377330]
34. Meyer S, Söling S, Lampe D, et al. Implementation of an electronic medication management support system in hospitalised polypharmacy patients: study protocol of a stepped-wedge cluster-randomised controlled trial (TOP study). BMJ Open 2025 Apr 15;15(4):e084696. [doi: 10.1136/bmjopen-2024-084696] [Medline: 40233949]
35. Offorha BC, Walters SJ, Jacques RM. Analysing cluster randomised controlled trials using GLMM, GEE1, GEE2, and QIF: results from four case studies. BMC Med Res Methodol 2023 Dec 13;23(1):293. [doi: 10.1186/s12874-023-02107-z] [Medline: 38093221]
36. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes, 3rd edition: Oxford University Press; 2005.
37. The green book: UK central government guidance on appraisal. : HM Treasury; 2022 URL: https://assets.publishing.service.gov.uk/media/698dbcd17da91680ad7f4308/The_Green_Book_2026.pdf [accessed 2026-03-17]
38. WiBe-Fachkonzept 5.0 – Konzept zur Durchführung von Wirtschaftlichkeitsbetrachtungen in der Bundesverwaltung [Report in German]. : Die Beauftragte der Bundesregierung für Informationstechnik; 2014 URL: https://wibe.de/wp-content/uploads/WiBe_50.pdf [accessed 2026-03-28]
39. Circular N0 A-94 Guidelines and discount rates for benefit-cost analysis of federal programs. : Office of Management and Budget; 1992 URL: https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/a94/a094.pdf [accessed 2026-03-17]
40. Hammerschmidt R, Meyer I. Socio-economic impact assessment and business models for integrated ecare. In: Moumtzoglou A, Meyer I, Kubitschke L, Müller S, editors. Achieving Effective Integrated E-Care beyond the Silos Hershey: IGI Global; 2014:136-163 URL: https://www.irma-international.org/viewtitle/111380/?isxn=9781466661387 [accessed 2026-03-28]

XSL•FO
RenderX

41. Bloomfield HE, Greer N, Linsky AM, et al. Deprescribing for community-dwelling older adults: a systematic review and meta-analysis. J Gen Intern Med 2020 Nov;35(11):3323-3332. [doi: 10.1007/s11606-020-06089-2] [Medline: 32820421]
42. Johansson T, Abuzahra ME, Keller S, et al. Impact of strategies to reduce polypharmacy on clinically relevant endpoints: a systematic review and meta‐analysis. Brit J Clinical Pharma 2016 Aug;82(2):532-548. [doi: 10.1111/bcp.12959]
43. Smith SM, Wallace E, O'Dowd T, Fortin M. Interventions for improving outcomes in patients with multimorbidity in primary care and community settings. Cochrane Database Syst Rev 2016 Mar 14;3(3):CD006560. [doi: 10.1002/14651858.CD006560.pub3] [Medline: 26976529]
44. Tecklenborg S, Byrne C, Cahir C, Brown L, Bennett K. Interventions to reduce adverse drug event-related outcomes in older adults: a systematic review and meta-analysis. Drugs Aging 2020 Feb;37(2):91-98. [doi: 10.1007/s40266-019-00738-w] [Medline: 31919801]
45. Dreischulte T, Donnan P, Grant A, Hapca A, McCowan C, Guthrie B. Safer prescribing—a trial of education, informatics, and financial incentives. N Engl J Med 2016 Mar 17;374(11):1053-1064. [doi: 10.1056/NEJMsa1508955] [Medline: 26981935]
46. Stuart EA, Green KM. Using full matching to estimate causal effects in nonexperimental studies: examining the relationship between adolescent marijuana use and adult outcomes. Dev Psychol 2008 Mar;44(2):395-406. [doi: 10.1037/0012-1649.44.2.395] [Medline: 18331131]
47. Greifer N. Matching methods. CRAN. 2025. URL: https://cran.r-project.org/web/packages/MatchIt/vignettes/matching-methods.html#caliper-matching-caliperGreifer [accessed 2026-02-06]
48. Greifer N. Assessing balance. CRAN. 2025. URL: https://cran.r-project.org/web/packages/MatchIt/vignettes/assessing-balance.html#ref-ho2007 [accessed 2026-02-06]
49. Harder VS, Stuart EA, Anthony JC. Propensity score techniques and the assessment of measured covariate balance to test causal associations in psychological research. Psychol Methods 2010 Sep;15(3):234-249. [doi: 10.1037/a0019623] [Medline: 20822250]
50. Matschinger H, Heider D, König HH. A comparison of matching and weighting methods for causal inference based on routine health insurance data, or: what to do if an RCT is impossible. Gesundheitswesen 2020 Mar;82(S 02):S139-S150. [doi: 10.1055/a-1009-6634] [Medline: 32066197]
51. Drummond M, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes, 4th edition: Oxford University Press; 2015.
52. Claxton K, Sculpher M, McCabe C, et al. Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra. Health Econ 2005 Apr;14(4):339-347. [doi: 10.1002/hec.985] [Medline: 15736142]
53. Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmacoeconomics 2008;26(9):781-798. [doi: 10.2165/00019053-200826090-00008] [Medline: 18767898]
54. Siebert U, Jahn B, Mühlberger N, Fricke FU, Schöffski O. 3 Entscheidungsanalyse und Modellierungen. In: Schöffski O, Graf Von Der Schulenburg JM, editors. Gesundheitsökonomische Evaluationen: Springer; 2012:275-324. [doi: 10.1007/978-3-642-21700-5_11]
55. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
56. Lampe D, Grosser J, Grothe D, et al. How intervention studies measure the effectiveness of medication safety-related clinical decision support systems in primary and long-term care: a systematic review. BMC Med Inform Decis Mak 2024 Jul 4;24(1):188. [doi: 10.1186/s12911-024-02596-y] [Medline: 38965569]

Abbreviations

c-RCT: cluster randomized controlled trial
CDSS: clinical decision support system
CG: control group
CONSORT: Consolidated Standards of Reporting Trials
DDI: drug-drug interaction
e-prescription: electronic prescription
eLiSa: electronic Life Saver
eRIKA: e-prescription as an element of interprofessional care pathways for continuous medication therapy management (*eRezept als Element interprofessioneller Versorgungspfade für kontinuierliche AMTS*)
GLMM: generalized linear mixed model
ICC: intracluster correlation coefficient
IG: intervention group
LMM: linear mixed model
MP: medication plan
MRC: Medical Research Council
OTC: over-the-counter

XSL•FO
RenderX

SEIA: socioeconomic impact assessment
SHI: statutory health insurance
TIDieR: Template for Intervention Description and Replication

Edited by J Sarvestan; submitted 07.Nov.2025; peer-reviewed by A Patel; revised version received 20.Feb.2026; accepted 24.Feb.2026; published 07.Apr.2026.

Please cite as:
Schmidt J, Lampe D, Poppe A, Meyer I, Söling S, Köberlein-Neu J, Grandt D, Düvel L, Greiner W, eRIKA Study Group
Enhancing Continuous Medication Safety Through e-Prescription and Clinical Decision Support Systems in Outpatient Practices and Pharmacies: Protocol for a Multiperspective Study (eRIKA Study)
JMIR Res Protoc 2026;15:e87277
URL: https://www.researchprotocols.org/2026/1/e87277
doi:10.2196/87277

© Juliana Schmidt, David Lampe, Adriana Poppe, Ingo Meyer, Sara Söling, Juliane Köberlein-Neu, Daniel Grandt, Lara Düvel, Wolfgang Greiner, eRIKA Study Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 7.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Remote Management of Patients With Diabetic Macular Edema Receiving Long-Term Intravitreal Anti–Vascular Endothelial Growth Factor Therapy (RE-SHINE Study): Protocol for a Real-World Study

Shu Li[1,2], PhD; Yupeng Xu[2,3], MD; Yan Xu[2,3], BMed; Xun Xu[2,3], MD

[1]Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

[2]National Clinical Research Center for Eye Diseases, Shanghai, China

[3]Department of Ophthalmology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, No. 85 Wujin Road, Hongkou District, Shanghai, China

Corresponding Author:
Xun Xu, MD
National Clinical Research Center for Eye Diseases, Shanghai, China

Abstract

Background: Long-term intravitreal anti–vascular endothelial growth factor therapy is the gold standard for treating diabetic macular edema (DME). However, poor treatment adherence in real-world settings often leads to suboptimal visual outcomes. Evidence-based remote management strategies are required to bridge the gap between clinical protocols and patient self-management.

Objective: This study aims to describe a remote management program based on the Behavior Change Wheel model designed to increase treatment adherence and improve disease outcomes for patients with DME receiving long-term anti–vascular endothelial growth factor therapy, and to report the study protocol.

Methods: The RE-SHINE study was structured into three phases: theoretical modeling, digital platform development, and a validation cohort study. The intervention included a smartphone-based app providing personalized health education, self-monitoring of vision (Amsler grid), injection reminders, and direct communication with health care providers. The efficacy of the remote management model was tested in a real-world cohort study. Participants were assessed at baseline and at each follow-up visit over 12 months. The primary outcome measure was best-corrected visual acuity, and secondary outcome measures included central retinal thickness and treatment adherence (injection rate and follow-up frequency).

Results: A total of 1006 patients with DME underwent initial screening, and 958 (95.2%) patients met the inclusion criteria, with only 7 patients declining follow-up. We anticipate that the remote management system will be both feasible and acceptable for patients with DME. It is hypothesized that the intervention group would demonstrate significantly higher levels of treatment adherence and better visual preservation compared to the retrospective control group. In addition, we intend to assess the impact of the digital intervention on patient-reported outcomes, encompassing quality of life and self-efficacy. Participant recruitment began in September 2023 and was completed by the time of submission. The data analysis is yet to begin. The results are expected to be published in 2026.

Conclusions: A Behavior Change Wheel–based remote management approach could be a scalable and sustainable method to enhance adherence in patients requiring long-term intravitreal therapy. If proven effective, this model could be integrated into routine ophthalmic care to mitigate the risk of vision loss due to undertreatment in DME.

Trial Registration: Chinese Clinical Trial RegistryChiCTR2400090348; https://www.chictr.org.cn/showprojEN.html?proj=185739

International Registered Report Identifier (IRRID): DERR1-10.2196/85764

(JMIR Res Protoc 2026;15:e85764) doi:10.2196/85764

KEYWORDS

diabetic macular edema; remote management; anti-VEGF therapy; Behavior Change Wheel; treatment adherence; digital health; anti–vascular endothelial growth factor therapy

XSL•FO
RenderX

Introduction

Diabetic macular edema (DME) is a manifestation of diabetic retinopathy that impairs central vision and has a profoundly detrimental impact on vision-specific function, mobility, independence, and quality of life [1]. Over the past 15 years, intravitreal anti–vascular endothelial growth factor (VEGF) therapy injections have become the first-line therapy for patients with DME [2]. In clinical trials, patients received an average of 6 injections per year, resulting in a visual acuity improvement of 12 letters [3]. In contrast, in clinical practice, patients in China received an average of 2.9 injections during the initial year of treatment, resulting in a visual acuity improvement of 1.1 letters, a figure that was notably lower than the average of 5 letters in other countries [4]. These poor outcomes were attributed to the failure to administer timely and prescribed intravitreal anti-VEGF injections and to adhere to scheduled follow-up appointments [5].

A paucity of research has been conducted on intravitreal anti-VEGF injections, and there has been a lack of theoretical guidance and in-depth investigation into influencing factors [2,5-7]. Consequently, we conducted a study to investigate the factors influencing the timely intravitreal anti-VEGF therapy in Chinese patients [8]. The findings of this study identified that untimely intravitreal anti-VEGF injection was driven by psychological capability (eg, lack of relevant knowledge), social opportunity (eg, effective doctor-patient communication), physical opportunity (eg, high treatment costs), reflective motivation (eg, lack of confidence in the efficacy, unmet expectations), and automatic motivation (eg, fear of injections and fear of blindness). These specific determinants provided the direct rationale for the self-monitoring and education modules in the new platform. The “psychological capability (lack of relevant knowledge)” drives the design of the “structured education” module. The design of the “online Q&A” module is informed by the “social opportunity (effective doctor-patient communication)” initiative. The “reflective motivation (lack of confidence in efficacy, unmet expectations)” informs the design of the “treatment management and feedback” module. The “automatic motivation (fear of injections, fear of blindness)” is a key factor in the design of the “case sharing” module.

Concurrently, clinicians’ emphasis on standardized diagnosis and regimented treatment, with minimal consideration given to patient adherence, has not improved clinical outcomes or reduced the use of health care resources. Consequently, a comprehensive set of measures must be implemented to optimize patient management. Nevertheless, there is a paucity of domestic and international studies on the management of patients with DME. While digital management has been explored in chronic eye diseases such as glaucoma with promising outcomes [9-11] and demonstrating improved cost-effectiveness [12,13], DME presents unique challenges due to its complex injection schedules [14]. Digital health care is currently focused primarily on screening rather than patient management [15-18], and there is no evidence-based remote management framework specifically designed for this patient population.

The objective of this study is to develop a remote management system integrating the Behavior Change Wheel (BCW) framework with the technical features of digital platforms for patients with DME receiving long-term intravitreal anti-VEGF therapy.

Methods

Ethical Considerations

Ethical approval for the study protocol was obtained from the ethics committees of Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine (2022-KY-021). In addition, this study was previously registered (ChiCTR2400090348). All participants were required to provide written informed consent, and their confidentiality was safeguarded through the implementation of anonymized data coding. The study entailed minimal risk and did not involve experimental treatments or invasive procedures. Participants could withdraw from the study at any time without consequence for their medical care.

Design

In this study, researchers conducted a three-stage observational study over 48 weeks, outlined as follows:

1. Theoretical modeling: A management system based on the BCW theory was constructed. The BCW was first proposed by Michie et al [19] as part of the development of the capability, opportunity, motivation–behavior (COM-B) theoretical model into the BCW theoretical framework. This theory has been formulated following a comprehensive synthesis of 19 related behavior change frameworks to support intervention designers in analyzing problems from a behavioral perspective, integrating interventions across three dimensions—capacity, opportunity, and motivation—to systematically select optimal intervention functions.
2. Platform development: Artificial intelligence questions and answers (Q&A), self-monitoring, and other modules were integrated.
3. A 48-week cohort study (remote group vs routine group) assessed the digital remote management model’s effects on visual acuity, retinal thickness, and injection rate. Figure 1 shows the study flow.

XSL•FO
RenderX

Figure 1. Study flowchart.

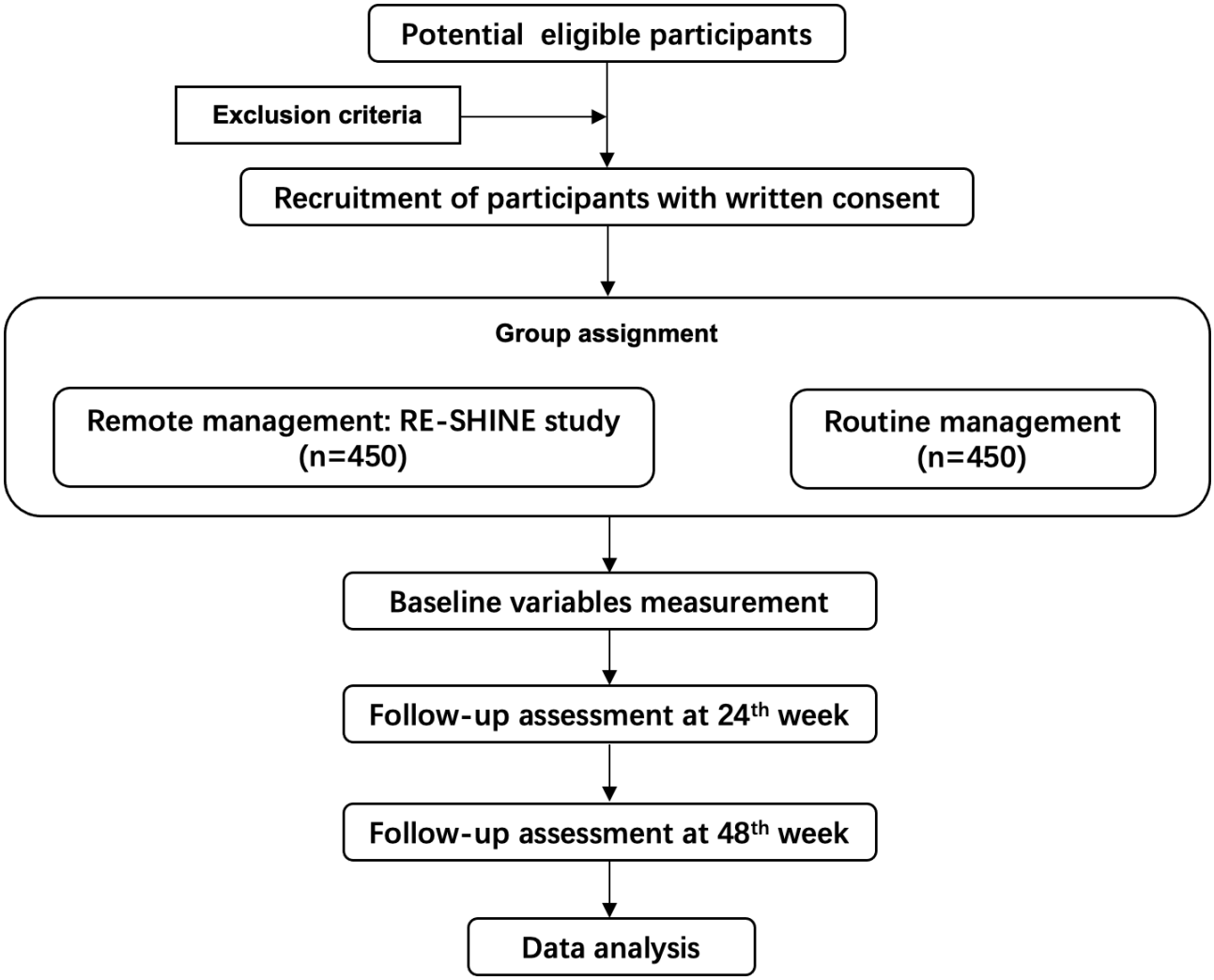

Participation Recruitment and Eligibility Assessment

A convenience sampling methodology was used to recruit participants from 26 cities in 17 provinces between September 2023 and September 2025. The inclusion criteria were as follows: individuals with DME in one or both eyes who are older than 18 years, initiation of intravitreal anti-VEGF injections in the treated eye, central retinal thickness (CRT) of 300 μm or greater in the treated eye, best-corrected visual acuity (BCVA) in the treated eye at 39 to 73 letters using the Early Treatment Diabetic Retinopathy Study (ETDRS) chart 4-meter test (Snellen visual acuity ~20/40 to ~20/160), possession of a smartphone, and ability to use the study app (independently or with assistance from a regular caregiver). Exclusion criteria included previous surgical, laser, or intravitreal hormone therapy for retinal disease in the treated eye; confirmed Alzheimer disease, intellectual disability, severe psychiatric, or cognitive impairments that prevent the independent use of smartphone apps or the provision of informed consent; pregnancy; or absence of local follow-up.

Once suitable patients were identified, the researcher contacted them by telephone to provide an overview of the study, including its purpose, the methodology, data protection, and confidentiality. If suitable participants consented, they were requested to sign an informed consent form, thereby confirming their participation in the study.

Sample Size

For conservative estimation, we consulted a large-scale study of patients with wet age-related macular degeneration (wAMD) [20] and used the ETDRS visual acuity scale as the primary end point (with δ=4.0 and s=16.2) and applied a 2-sided test (with $1 - \beta = .9$ and α=.05); the missed visit rate was set at 20%. Using PASS 2021 software (NCSS, LLC) for estimation, the required minimum total sample size was calculated to be 864 patients (432 patients per group). After considering the uncertainty in key parameter estimates and the feasibility and operational convenience of multicenter recruitment, we made a moderate adjustment to the sample size; the final plan is to enroll 900 patients (450 patients per group). Considering that wAMD and DME share comparable pathophysiological mechanisms (both involving the VEGF pathway) and primary efficacy end points (change in BCVA letter count), differences may exist between patient populations. We therefore conducted a sensitivity analysis based on published DME studies [21-27], adhering to the aforementioned criteria (a 2-sided test, with $1 - \beta = .9$ and α=.05; the missed visit rate was set at 20%). The total sample size we calculated ranges from 58 to 886 patients, demonstrating that the selected parameters in our study,

XSL•FO
RenderX

including effect size and variance, were appropriate. To ensure data comparability, hospitals were matched based on the number of annual ophthalmology outpatient visits. Patients were selected according to previously outlined inclusion and exclusion criteria. The primary outcome for sample size calculation was the mean change in BCVA from baseline to 12 months between the intervention and control groups.

Management Description

Remote Management Description

The patient remote management program comprised five primary domains: patient follow-up, patient education, patient Q&A, patient self-management, and key population management (Figure 2).

Figure 2. Structure of the patient remote management program and definition of key monitoring populations. AI: artificial intelligence; CRT: central retinal thickness; Q&A: questions and answers.

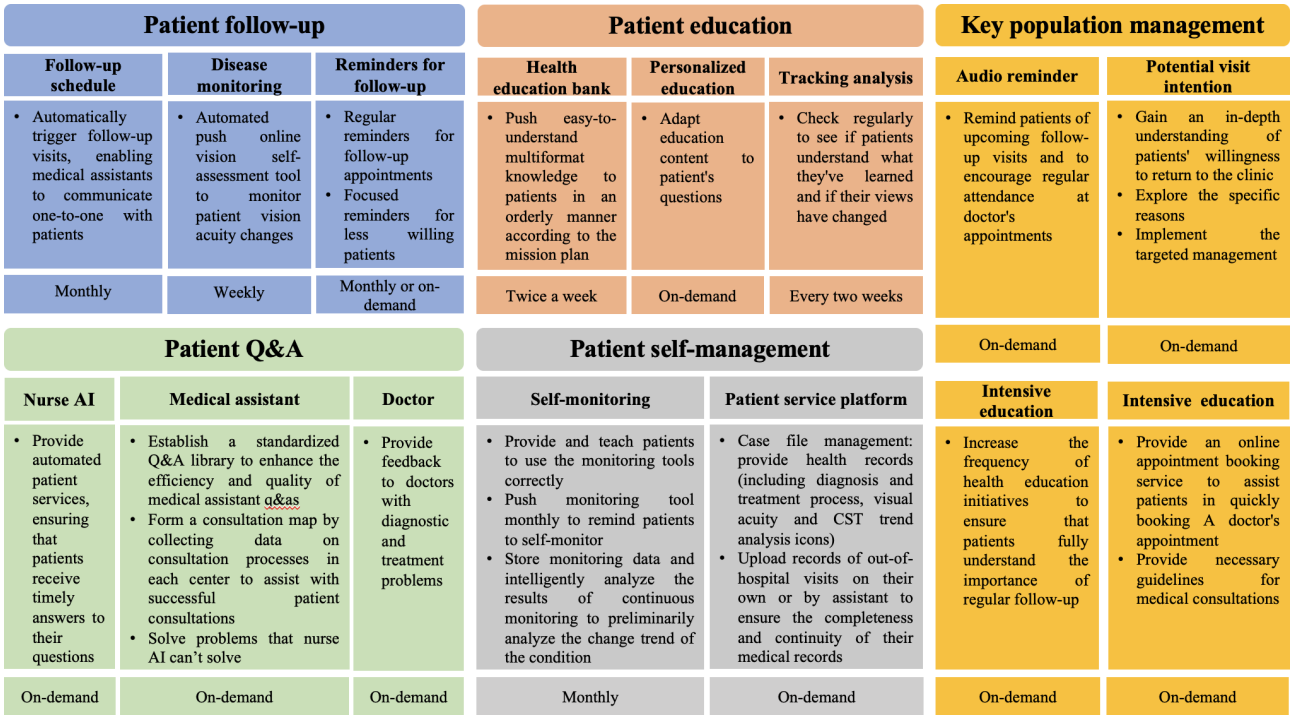

The remote management program categorized patients according to follow-up adherence during different treatment phases of intravitreal anti-VEGF therapy. The loading period was defined as the first 3 months following the initial intravitreal injection. Patients were classified as at risk of poor adherence when the interval between scheduled return visits exceeded 6 weeks. The maintenance period was defined as the period beginning 3 months after the first injection. Patients were considered at risk when the interval between follow-up visits exceeded 12 weeks. Patients were also classified into the risk group if they failed to provide feedback through the digital platform or did not respond after reminder notifications from medical assistants, even when the visit interval had not exceeded the predefined threshold. The remote management system incorporated monitoring, reminder, feedback, and physician-patient communication modules to support adherence and long-term treatment persistence.

Routine Management Description

Routine management measures primarily consisted of hospital medical professionals providing patients with disease education, either verbally or in writing, and informing them of their subsequent follow-up appointment time during each outpatient visit. This management was conducted mainly offline, meaning in-person and within the hospital setting. Patients did not routinely receive online disease education or telephone follow-up.

Data Collection and Management

Data Collection

Table 1 presents the parameters and assessment times for the study in the remote management group.

Table 2 presents the parameters and assessment times for the study in the routine management group.

XSL•FO
RenderX

Table . The parameters and times of assessment for the study in the remote management group.

Parameters	Baseline	Fourth week	Eighth week	12th week	24th week	48th week
Sociodemographic information	✓					
Primary outcomes						
Best-corrected visual acuity	✓	✓	✓	✓	✓	✓
Secondary outcomes						
Central retinal thickness	✓	✓	✓	✓	✓	✓
Number of intravitreal injections		✓	✓	✓	✓	✓
Number of follow-up visits		✓	✓	✓	✓	✓
Compliance rate for follow-up visit		✓	✓	✓	✓	✓
Loss to follow-up rate		✓	✓	✓	✓	✓
Incidence of adverse events		✓	✓	✓	✓	✓

Table . The parameters and times of assessment for the study in the routine management group.

Parameters	Baseline	As prescribed	24th week	48th week
Sociodemographic information	✓			
Primary outcomes				
Best-corrected visual acuity	✓	✓	✓	✓
Secondary outcomes				
Central retinal thickness	✓	✓	✓	✓
Number of intravitreal injections		✓	✓	✓
Number of follow-up visits		✓	✓	✓
Compliance rate for follow-up visit		✓	✓	✓
Loss to follow-up rate		✓	✓	✓
Incidence of adverse events		✓	✓	✓

Sociodemographic Information

The researchers devised a questionnaire to collect the sociodemographic data from the participants. This study collected sociodemographic data from the participants, including their age, marital status, health status, and other relevant demographic information. The questionnaire included a column where participants were asked to list any medicines or supplements they were currently taking, as well as any symptoms or illnesses they were experiencing.

Primary Outcome

The primary outcome was the BCVA as measured using the ETDRS test, which was used to compare the change in visual acuity at each time point before and after medication.

Secondary Outcomes

The secondary outcomes included the following:

- CRT was measured using optical coherence tomography.
- The number of intravitreal injections was determined by calculating the number of intravitreal anti-VEGF injections administered to patients from enrollment to the end of the

XSL•FO
RenderX

follow-up cycle (12 mo), as advised by medical professionals.
- The number of follow-up visits was determined by calculating the number of visits from enrollment to the end of the follow-up cycle (12 mo), as advised by medical professionals.
- The compliance rate for follow-up visits was the proportion of patients whose each follow-up visit was no more than 2 weeks from the scheduled follow-up date, calculated during the observation period after the patient initiates treatment, as advised by medical professionals.
- The loss to follow-up rate was calculated based on the number of patients who did not visit the hospital during the observation period following treatment initiation. This was defined as a period of more than 2 months without a follow-up visit to the hospital during the observation period, until the end of the observation period.
- The incidence of adverse events was defined as the proportion of adverse medical events in a study patient that were not necessarily causally related to the intravitreal injection, such as death or life-threatening events, or those resulting in persistent or significant loss of function.

Data Management

The data was carefully stored in a password-protected, secure electronic research database, which was supported by a backup system, antivirus software, and security protocols to help ensure data safety. Access to the electronic database was thoughtfully limited to a select group of researchers, who were listed in the authorization log of the survey master file. The data manager kindly oversaw regular data quality control checks and generated reports and statistics using tools integrated within the survey distribution platform.

Statistical Analysis

Analyses of Outcomes

The collected data were entered into Excel 16.0 (Microsoft Corporation), and SPSS 25.0 (IBM Corp) was subsequently used for data analysis. The measurement data were presented as means and SDs, and the count data were expressed as percentages. Before primary analysis, continuous variables (eg, visual acuity change values) underwent normality testing using a combination of the Shapiro-Wilk test and Q-Q plots. Should the data be found to be normally distributed, parametric tests will be applied. For data showing significant deviation from normality, appropriate transformations or nonparametric tests were considered. To account for the multicenter design and potential clustering effects across different cities and hospitals, linear mixed-effects models were used, with the treatment group as a fixed effect and the center/hospital as a random effect. Statistical significance was considered as a P value less than .05. When reporting intergroup comparisons, P values were accompanied by corresponding effect sizes (eg, Cohen d) and their 95% CIs to assess the clinical relevance of observed differences.

Handling Missing Data

The missing data were addressed with multiple imputation methods, assuming random data loss. Sensitivity analyses were conducted to assess the robustness of the results, and a full case analysis was performed for comparison. Furthermore, where applicable, data missing due to patient dropout were reported, and any deviations from the original analysis plan were thoroughly documented.

Results

This study integrates the BCW framework with the technical characteristics of digital platforms to systematically design and evaluate interventions. The modeling process comprises three core steps: (1) conducting a behavioral diagnosis and analysis based on the BCW, (2) mapping theory to digital functional design, and (3) developing a digital remote management platform. The platform integrates multiple functional modules, including patient education, treatment reminders, follow-up monitoring, symptom reporting, and doctor-patient communication. These have been designed to support long-term adherence and promote continuous disease management. This system architecture facilitates real-time interaction between patients and health care providers, delivering personalized feedback and remote monitoring throughout the treatment period. A total of 1006 patients with DME underwent initial screening, and 958 (95.2%) patients met the inclusion criteria, with only 7 patients declining follow-up. Participant recruitment began in September 2023 and was completed by the time of submission. The data analysis is yet to begin. The results are expected to be published in 2026.

Because this is a study protocol describing an ongoing intervention, the primary focus of the initial analysis will be on the feasibility of the remote management system, patient engagement with the smartphone app, and the acceptability of the BCW-based intervention content. We have established several benchmarks to define the success of this management model: we anticipate an enrollment rate of over 60% among eligible patients and a retention rate of at least 80% for the 12-month follow-up period. Regarding intervention adherence, we expect that at least 75% of participants will actively use the self-monitoring features (such as the digital Amsler grid) at least once a week and that the majority (>70%) will interact with the personalized health education modules.

We hypothesize that patients enrolled in the remote management program will demonstrate significantly higher treatment adherence, characterized by a lower rate of missed intravitreal injection appointments and more frequent follow-up visits, compared to a historical control group. We also expect that the intervention group will achieve superior visual outcomes, specifically a greater mean improvement in BCVA and a more significant reduction in CRT at the 12-month time point. Furthermore, we will explore the impact of the intervention on patient-reported outcomes, including self-efficacy in managing diabetes and vision-related quality of life. Although this real-world study primarily evaluates clinical effectiveness, we will examine the effect sizes of the remote management model to determine its potential for large-scale clinical implementation.

XSL•FO
RenderX

Discussion

We hypothesize that the RE-SHINE remote management model will significantly enhance treatment adherence and improve clinical outcomes in patients with DME. Specifically, by addressing the barriers identified through the BCW framework, we anticipate a measurable increase in follow-up completion rates and a reduction in missed anti-VEGF injection appointments. These improvements are expected to translate into better long-term preservation of BCVA and a more consistent reduction in CRT compared to patients receiving routine clinical care. Of note, the efficacy of remote management has been demonstrated in several areas, including facilitating improved patient adherence to treatment [28], enhancing disease prognosis [29,30], managing health issues and data, providing personalized self-care advice, supporting patient-provider communication, and decision-making [31]. Despite these promising results, there is a paucity of clinical studies on this subject in patients with DME who received intravitreal anti-VEGF therapy. Building on the team's previous in-depth study, this study was designed to evaluate the utility and effectiveness of management measures using a cohort study.

Interestingly, our study is specifically tailored for the intensive treatment phase of DME. Building on our previous qualitative research that identified key barriers such as injection anxiety and travel burden [4], the RE-SHINE program integrates targeted psychological support and real-time vision monitoring. This transition from "theory-based barrier identification" to "evidence-based digital intervention" represents a significant advancement over generalized teleophthalmology models.

By addressing the "psychological capability" through education, "social opportunity" through patient Q&A, and "physical opportunity" via remote monitoring, the RE-SHINE program directly targets the barriers identified in our framework, aiming to translate theoretical determinants into improved clinical adherence.

The management methods used in this study included a remote virtual approach; a patient relationship management tool to facilitate task shifting; and a scaled, team-based management approach that incorporated physician assistants, nurse practitioners, and physicians who oversaw and intervened in management as required, in addition to self-monitoring of the patient's condition. Given that patients undergoing anti-VEGF therapy spend the majority of their time outside of the hospital settings, remote management can meet their needs when needed. Self-monitoring tools enable patients to acquire specialized health knowledge, track the empowering effects of their health indicators, and improve their ability to access resources [32,33]. This study hypothesizes that self-monitoring will help patients track their disease progression before and after intravitreal injections, improve the efficacy of remote management, and support patients in overcoming challenges during treatment.

The patient's adherence to treatment has been demonstrated to have a significant influence on disease outcomes [29,30]. This study involves remote management at various points throughout the treatment period, with a focus on the changes that occur during the entire therapy cycle.

This study is one of the first systematic attempts to integrate a BCW-based remote management model into real-world anti-VEGF therapy for patients with DME in China. It has the potential to establish a foundational framework and provide a series of concepts that will inform subsequent studies on the management of treatment adherence to anti-VEGF in patients with DME. The findings of this study have the potential to directly influence patients' disease outcomes by enhancing their adherence to treatment regimens. The potential economic benefits of increased adherence may be reflected in a reduced demand for health care services, resulting in lower expenditures.

A major strength of this study is its solid theoretical foundation using the BCW, which ensures that the digital intervention directly targets the modifiable behaviors of patients with DME. Additionally, the multicenter design involving 26 cities enhances the real-world generalizability of our findings. However, several limitations must be acknowledged. First, the sample size estimation for this study was based on hypothetical parameters derived from the wAMD population, which may be subject to inherent limitations. These include potential differences in patient demographics (such as age distribution), real-world treatment patterns and intervals, and treatment adherence or discontinuation behaviors. Such variations may have a significant impact on the estimated effect size and outcome variability. Future studies should focus on obtaining more precise parameter estimates within the DME population. Second, the effect of remote management is only tracked in the first year of treatment, so the validity of its long-term maintenance still needs to be further validated. Third, the study may be limited due to the "digital divide," and patients with low digital literacy or those without access to smartphones may face barriers to participation. Additionally, while the platform's features are evidence-based, a formal standardized usability and acceptability assessment of the app interface was not the primary objective of this protocol phase.

Acknowledgments

The authors acknowledge with gratitude the support of the National Clinical Research for Eye Diseases. The authors confirm that no generative artificial intelligence technologies were used in the creation of this manuscript. The funder had no involvement in the review or approval of the manuscript for publication. The authors are responsible for the content of this work, and it does not necessarily represent the official views of the funder or the authors' employers.

XSL•FO
RenderX

Funding

This research protocol was funded by the National Clinical Research for Eye Diseases.

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[DOCX File, 35 KB - resprot_v15i1e85764_app1.docx]

References

1. Yau JWY, Rogers SL, Kawasaki R, et al. Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care 2012 Mar;35(3):556-564. [doi: 10.2337/dc11-1909] [Medline: 22301125]
2. Angermann R, Hofer M, Huber AL, et al. The impact of compliance among patients with diabetic macular oedema treated with intravitreal aflibercept: a 48-month follow-up study. Acta Ophthalmol 2022 Mar;100(2):e546-e552. [doi: 10.1111/aos.14946] [Medline: 34145756]
3. Liu K, Song Y, Xu G, et al. Conbercept for treatment of neovascular age-related macular degeneration: results of the randomized phase 3 PHOENIX study. Am J Ophthalmol 2019 Jan;197:156-167. [doi: 10.1016/j.ajo.2018.08.026] [Medline: 30148987]
4. Holz FG, Figueroa MS, Bandello F, et al. Ranibizumab treatment in treatment-naive neovascular age-related macular degeneration: results from LUMINOUS, a global real-world study. Retina (Philadelphia, Pa) 2020 Sep;40(9):1673-1685. [doi: 10.1097/IAE.0000000000002670] [Medline: 31764612]
5. Okada M, Mitchell P, Finger RP, et al. Nonadherence or nonpersistence to intravitreal injection therapy for neovascular age-related macular degeneration: a mixed-methods systematic review. Ophthalmology 2021 Feb;128(2):234-247. [doi: 10.1016/j.ophtha.2020.07.060] [Medline: 32763265]
6. Weiss M, Sim DA, Herold T, et al. Compliance and adherence of patients with diabetic macular edema to intravitreal anti-vascular endothelial growth factor therapy in daily practice. Retina (Philadelphia, Pa) 2018 Dec;38(12):2293-2300. [doi: 10.1097/IAE.0000000000001892] [Medline: 29068914]
7. Varano M, Eter N, Winyard S, Wittrup-Jensen KU, Navarro R, Heraghty J. Current barriers to treatment for wet age-related macular degeneration (wAMD): findings from the wAMD patient and caregiver survey. Clin Ophthalmol 2015;9:2243-2250. [doi: 10.2147/OPTH.S92548] [Medline: 26664038]
8. Li S, Pan J, Xu Y, Dai Z, Fang Q. Exploring the factors influencing the timely intravitreal anti-VEGF treatment in patients with diabetic macular edema: a qualitative interview study using the COM-B model. BMC Health Serv Res 2025 Feb 24;25(1):302. [doi: 10.1186/s12913-025-12379-4] [Medline: 39994645]
9. Bhanvadia SB, Brar MS, Delavar A, et al. Assessing usability of smartwatch digital health devices for home blood pressure monitoring among glaucoma patients. Informatics (MDPI) 2022 Dec;9(4):79. [doi: 10.3390/informatics9040079] [Medline: 36873830]
10. Jesus J, Aguiar C, Meira D, Rodriguez-Una I, Beirão JM. Telemonitoring tools for glaucoma patients: a systematic review of current trends and applications. J Clin Med 2025 May 9;14(10):3317. [doi: 10.3390/jcm14103317] [Medline: 40429313]
11. Hamid S, Minakaran N, Igwe C, Baneke A, Pedersen M, Mathew RG. An online glaucoma educational course for patients to facilitate remote learning and patient empowerment. BMJ Health Care Inform 2023 Aug;30(1):e100748. [doi: 10.1136/bmjhci-2023-100748] [Medline: 37620107]
12. Schargus M, Grehn F. The European Glaucoma Society Glaucocard project: improved digital documentation of medical data for glaucoma patients based on standardized structured international datasets. Graefes Arch Clin Exp Ophthalmol 2008 Dec;246(12):1757-1768. [doi: 10.1007/s00417-008-0929-x] [Medline: 18766366]
13. Wright HR, Diamond JP. Service innovation in glaucoma management: using a web-based electronic patient record to facilitate virtual specialist supervision of a shared care glaucoma programme. Br J Ophthalmol 2015 Mar;99(3):313-317. [doi: 10.1136/bjophthalmol-2014-305588] [Medline: 25336582]
14. Horton MB, Silva PS, Cavallerano JD, Aiello LP. Clinical components of telemedicine programs for diabetic retinopathy. Curr Diab Rep 2016 Dec;16(12):129. [doi: 10.1007/s11892-016-0813-8] [Medline: 27796779]

XSL•FO
RenderX

15. Shi L, Wu H, Dong J, Jiang K, Lu X, Shi J. Telemedicine for detecting diabetic retinopathy: a systematic review and meta-analysis. Br J Ophthalmol 2015 Jun;99(6):823-831. [doi: 10.1136/bjophthalmol-2014-305631] [Medline: 25563767]
16. Xiao H, Tang J, Zhang F, et al. Global trends and performances in diabetic retinopathy studies: a bibliometric analysis. Front Public Health 2023;11:1128008. [doi: 10.3389/fpubh.2023.1128008] [Medline: 37124794]
17. Ting DSW, Tan GSW. Telemedicine for diabetic retinopathy screening. JAMA Ophthalmol 2017 Jul 1;135(7):722-723. [doi: 10.1001/jamaophthalmol.2017.1257] [Medline: 28520885]
18. Nguyen HV, Tan GSW, Tapp RJ, et al. Cost-effectiveness of a national telemedicine diabetic retinopathy screening program in Singapore. Ophthalmology 2016 Dec;123(12):2571-2580. [doi: 10.1016/j.ophtha.2016.08.021] [Medline: 27726962]
19. Michie S, van Stralen MM, West R. The Behaviour Change Wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
20. Souied EH, Oubraham H, Mimoun G, et al. Changes in visual acuity in patients with wet age-related macular degeneration treated with intravitreal ranibizumab in daily clinical practice: the TWIN study. Retina (Philadelphia, Pa) 2015 Sep;35(9):1743-1749. [doi: 10.1097/IAE.0000000000000548] [Medline: 25901835]
21. Sugimoto M, Handa C, Hirano K, Sunaya T, Kondo M. Intravitreal aflibercept for diabetic macular edema in real-world clinical practice in Japan: 24-month outcomes. Graefes Arch Clin Exp Ophthalmol 2022 Nov;260(11):3489-3498. [doi: 10.1007/s00417-022-05703-9] [Medline: 35652946]
22. Diabetic Retinopathy Clinical Research Network, Wells JA, Glassman AR, et al. Aflibercept, bevacizumab, or ranibizumab for diabetic macular edema. N Engl J Med 2015 Mar 26;372(13):1193-1203. [doi: 10.1056/NEJMoa1414264] [Medline: 25692915]
23. Wirkkala J, Kubin AM, Ohtonen P, Yliselä J, Siik T, Hautala N. Visual outcomes of observation, macular laser and anti-VEGF in diabetic macular edema in type 1 diabetes: a real-world study. BMC Ophthalmol 2022 Jun 9;22(1):258. [doi: 10.1186/s12886-022-02482-z] [Medline: 35681133]
24. Van Aken E, Favreau M, Ramboer E, et al. Real-world outcomes in patients with diabetic macular edema treated long term with ranibizumab (VISION Study). Clin Ophthalmol 2020;14:4173-4185. [doi: 10.2147/OPTH.S281501] [Medline: 33299294]
25. James DGP, Mitkute D, Porter G, Vayalambrone D. Visual outcomes following intravitreal ranibizumab for diabetic macular edema in a pro re nata protocol from baseline: a real-world experience. Asia Pac J Ophthalmol (Phila) 2019;8(3):200-205. [doi: 10.22608/APO.2018413] [Medline: 31165603]
26. Sahni J, Patel SS, Dugel PU, et al. Simultaneous inhibition of angiopoietin-2 and vascular endothelial growth factor-A with faricimab in diabetic macular edema: BOULEVARD phase 2 randomized trial. Ophthalmology 2019 Aug;126(8):1155-1170. [doi: 10.1016/j.ophtha.2019.03.023] [Medline: 30905643]
27. Singh RP, Barakat MR, Ip MS, et al. Efficacy and safety of brolucizumab for diabetic macular edema: the KINGFISHER randomized clinical trial. JAMA Ophthalmol 2023 Dec 1;141(12):1152-1160. [doi: 10.1001/jamaophthalmol.2023.5248] [Medline: 37971723]
28. Golledge J, Fernando M, Lazzarini P, Najafi B, G Armstrong D. The potential role of sensors, wearables and telehealth in the remote management of diabetes-related foot disease. Sensors (Basel) 2020 Aug 13;20(16):4527. [doi: 10.3390/s20164527] [Medline: 32823514]
29. Blood AJ, Chang LS, Hassan S, et al. Randomized evaluation of a remote management program to improve guideline-directed medical therapy: the DRIVE trial. Circulation 2024 Jun 4;149(23):1802-1811. [doi: 10.1161/CIRCULATIONAHA.124.069494] [Medline: 38583146]
30. Blood AJ, Chang LS, Colling C, et al. Methods, rationale, and design for a remote pharmacist and navigator-driven disease management program to improve guideline-directed medical therapy in patients with type 2 diabetes at elevated cardiovascular and/or kidney risk. Prim Care Diabetes 2024 Apr;18(2):202-209. [doi: 10.1016/j.pcd.2024.01.005] [Medline: 38302335]
31. Liu K, Xie Z, Or CK. Effectiveness of mobile app-assisted self-care interventions for improving patient outcomes in type 2 diabetes and/or hypertension: systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2020 Aug 4;8(8):e15779. [doi: 10.2196/15779] [Medline: 32459654]
32. Wani RT, Nabi S, Nazir U, Haq I, Khan SMS. Effect of a structured health promotional program using the self-support groups on lifestyle behavior: the Ath Waas Interventional Trial. J Educ Health Promot 2024;13:256. [doi: 10.4103/jehp.jehp_1204_23] [Medline: 39309998]
33. Fundoiano-Hershcovitz Y, Ritholz MD, Horwitz DL, Behar E, Manejwala O, Goldstein P. The impact of digital self-monitoring of weight on improving diabetes clinical outcomes: quasi-randomized study. J Med Internet Res 2024 Apr 2;26:e54940. [doi: 10.2196/54940] [Medline: 38564266]

Abbreviations

BCVA: best-corrected visual acuity
BCW: Behavior Change Wheel
COM-B: capability, opportunity, motivation–behavior
CRT: central retinal thickness

XSL•FO
RenderX

DME: diabetic macular edema
ETDRS: Early Treatment Diabetic Retinopathy Study
Q&A: questions and answers
VEGF: vascular endothelial growth factor
wAMD: wet age-related macular degeneration

Edited by J Sarvestan; submitted 12.Oct.2025; peer-reviewed by A Ambresin, N Kato; revised version received 12.Mar.2026; accepted 13.Mar.2026; published 06.Apr.2026.

Please cite as:
Li S, Xu Y, Xu Y, Xu X
Remote Management of Patients With Diabetic Macular Edema Receiving Long-Term Intravitreal Anti–Vascular Endothelial Growth Factor Therapy (RE-SHINE Study): Protocol for a Real-World Study
JMIR Res Protoc 2026;15:e85764
URL: https://www.researchprotocols.org/2026/1/e85764
doi:10.2196/85764

© Shu Li, Yupeng Xu, Yan Xu, Xun Xu. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

WhatsApp-Led mHealth Intervention for Hypertension Management Among Adults Aged 40 to 59 Years in Kerala, India: Protocol for a Mixed Methods Study

Karthika Maniyara[1], BAMS, MPH; Prakash Babu Kodali[1,2], MPH, MPhil, PhD

[1]Department of Public Health and Community Medicine, Central University of Kerala, Tejaswini Hills, Periya, Kasaragod, India

[2]Center for Tobacco Control Research and Education, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Prakash Babu Kodali, MPH, MPhil, PhD
Department of Public Health and Community Medicine, Central University of Kerala, Tejaswini Hills, Periya, Kasaragod, India

Abstract

Background: Medication adherence and lifestyle modification remain key in the management of hypertension. Leveraging mobile technologies to improve patient adherence has been found to be effective in chronic disease management. Given the high mobile penetration and the opportunities it provides for encouraging lifestyle modification, an app such as WhatsApp, which is a free-to-use and real-time messaging platform, could facilitate patient-clinician interactions. We present a protocol for a WhatsApp-led mobile health (mHealth) intervention to improve adherence to medication and healthy lifestyles among adults aged 40 to 59 years in Kerala, India.

Objective: The study aims to assess the impact of the WhatsApp-led mHealth strategy on blood pressure measures and adherence to medication and lifestyle modification. The study also focuses on exploring participants' views on feasibility, facilitators, and challenges, as well as their intention to use this WhatsApp-led mHealth strategy.

Methods: The study will follow a multiphase mixed methods (sequential explanatory) design with an initial quantitative phase (quasi-experimental study design) followed by a qualitative phase (phenomenological approach). Multistage cluster sampling will be adopted. Using a survey checklist and blood pressure measurements, community-dwelling adults aged 40 to 59 years will be screened for hypertension. Approximately 120 participants with hypertension will be assigned each to the intervention and control arms. All the participants will be advised to follow routine hypertension care from health care providers. Additionally, participants in the intervention arm will receive regular health messages and reminders over WhatsApp. A baseline survey will be conducted at enrollment, and a follow-up survey will be conducted after 3 months at the end of the intervention. The qualitative phase will involve follow-up interviews. Participants will be selected through quota sampling from those with the lowest and highest adherence to medication and lifestyle modification after receiving the intervention. Statistical analysis for survey data and thematic analysis for qualitative data will be conducted.

Results: The study received ethics approval (IHEC/CUK/2024/04) from the Institutional Human Ethics Committee of the Central University of Kerala. All 240 participants were recruited by July 2025. The study completed data collection in August 2025 and is expected to finish data analysis by April 2026. The study results are expected to be submitted for publication in July 2026.

Conclusions: The study will be a potential strategy to address nonadherence to treatment and lifestyle regimens, which is a critical bottleneck to controlling hypertension and preventing complications. The study could be proof for a WhatsApp-led mHealth strategy that can be integrated into hypertension management at the community level, especially for improving hypertension awareness and reminders for behavior change.

Trial Registration: Clinical Trials Registry – India CTRI/2024/09/073166; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=MTE0MTg3&Enc=&userName=073166

International Registered Report Identifier (IRRID): DERR1-10.2196/81307

(JMIR Res Protoc 2026;15:e81307) doi:10.2196/81307

KEYWORDS

adherence; hypertension; lifestyle; medication; mobile health; mHealth; WhatsApp

XSL•FO
RenderX

Introduction

Increased blood pressure levels are considered a prominent risk factor for cardiovascular diseases, which account for one-third of the deaths in India [1]. Uncontrolled hypertension is associated with nonadherence to medication and healthy lifestyle practices [2,3]. For the management of chronic conditions such as hypertension, India's National Health Policy 2017 encourages the deployment of digital technologies to supplement routine health care. Digital technologies for hypertension management have been shown to promote treatment adherence and encourage lifestyle modification (ie, increased physical activity and better diet) [4-9].

Adults in the age group of 40 to 59 years are a crucial demographic for hypertension management. They have a greater risk of developing hypertension than their younger counterparts and better odds of controlling blood pressure compared to individuals older than 60 years. Major bottlenecks for hypertension management and control in this group lie in lifestyle modification and adherence to medication. Given the high mobile penetration, better digital literacy, and the opportunities for encouraging lifestyle modification that they provide, mobile-based interventions could be particularly effective among this population group for facilitating chronic disease management. WhatsApp, as a free-to-use and cost-effective app with better ease of use and user acceptance, has been increasingly used in organizational communication, monitoring, and delivery of services. As a real-time messaging platform, it has been shown to facilitate patient-clinician interactions [10,11]. Instant messaging through WhatsApp can facilitate mobile health (mHealth) services by acting as a medium for communicating reminders, health advice, and hypertension management guidelines [5,6,9,12,13].

A significant challenge of disease-specific mHealth apps is limited acceptance and ease of use among end users (patients). Given that hypertension management includes adherence to medicine and lifestyle changes, behavior change communication through WhatsApp-mediated reminders could function as cues to action and subsequently promote hypertension management. The effectiveness and feasibility of such an intervention are yet to be explored.

In this study, we aim to assess the impact of a WhatsApp-led mHealth strategy on blood pressure measures and adherence to medication and lifestyle modification among adults with hypertension aged between 40 and 59 years. The study also explores the participants' intention to use the mHealth strategy and their views on the feasibility, facilitators, and challenges of using a WhatsApp-led mHealth strategy for adhering to medication and a healthy lifestyle for hypertension management. The conceptual framework adopted for the study (Figure 1 [14]) explains the facilitators and challenges of adopting the mHealth strategy for medication and lifestyle adherence using the Andersen-Newman behavioral model of health service use, unified theory of acceptance and use of technology, and health belief model [15-17].

Figure 1. Conceptual framework for the study. Clinically, hypertension control is achieved when systolic blood pressure is below 140 mm Hg and diastolic blood pressure is below 90 mm Hg as a result of lifestyle modification or medication [14]. mHealth: mobile health.

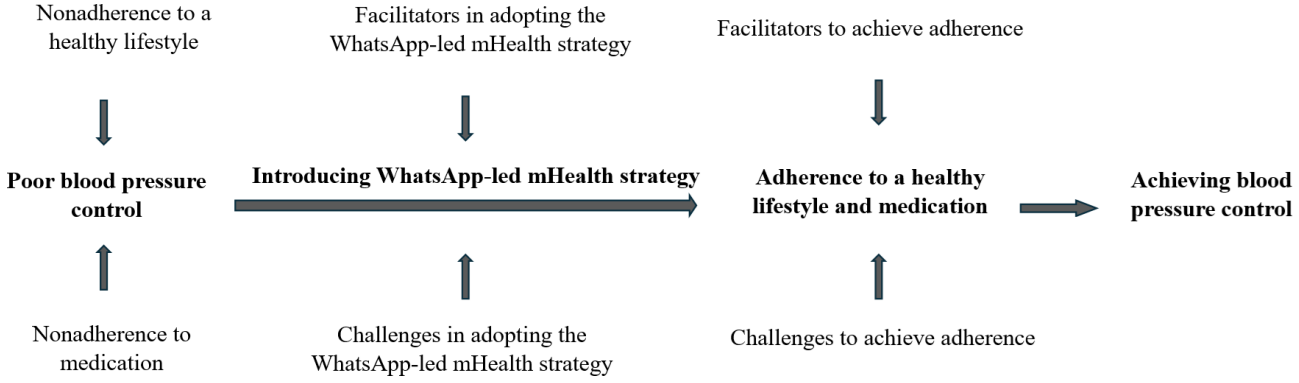

Methods

Study Design and Protocol

The study will adopt a multiphase mixed methods approach with a sequential explanatory design: an initial quantitative phase followed by a qualitative phase. The study protocol follows the Mixed Methods Reporting in Rehabilitation and Health Sciences guidelines, mHealth Evidence Reporting and Assessment checklist, and Transparent Reporting of Evaluations with Nonrandomized Designs statement [18-20].

Study Setting and Population

Given the high prevalence and low control rates of hypertension in the Malabar region of Kerala, India, especially in Kannur [21,22], the study will be conducted among adults aged 40 to 59 years in the Kannur district, Kerala.

Quantitative Phase of the Study

Study Design

The quantitative phase will use a quasi-experimental study design with an intervention group and a control group. This phase is designed to assess the effectiveness of the WhatsApp-led mHealth intervention on blood pressure estimates and medication and lifestyle adherence among individuals with hypertension.

Sample Size

A total study sample size of 240, with 120 each in the intervention and control arms, was estimated based on the

XSL•FO
RenderX

estimates of diastolic blood pressure at baseline (mean 80.9, SD 10.9 mm Hg) and the postintervention time point (mean 75.4, SD 11.5 mm Hg) from a previous randomized controlled trial in Kerala [23]. *OpenEpi* was used for sample size calculation. The sample size was estimated at 80% power, a 95% CI, a 20% nonresponse rate, and a design effect of 1.5.

Sampling

The study followed a multistage cluster sampling method. In the first stage, all taluks in the Kannur district were listed, and then 1 taluk (ie, Payyanur taluk) was selected. One rural (ie, Karivellur-Peralam gram panchayat) and one urban (ie, Payyanur municipality) administrative unit (cluster) from this taluk were then selected for the study (Figure 2). In stage 2, a door-to-door household survey will be conducted in these 2 rural and urban clusters. Household members aged 40 to 59 years will be screened for eligibility after consent is obtained. In the third stage, eligible participants (those with hypertension and access to WhatsApp) will be surveyed. After the baseline survey, the study's recruitment will follow a nonrandomized approach (Figure 3) without matching or blinding as these are not applicable in this study setting. Upon recruitment and consent, the first participant will be allocated to the intervention group, and the next participant will be allocated to the control group. The participants will be recruited into the intervention and control groups sequentially in a 1:1 ratio. This process will be conducted iteratively until the sample size of 240 is reached (ie, 120 in the intervention group and 120 in the control group). At the cluster level, this amounts to 120 participants per cluster (urban and rural), with 60 of them in the intervention group and 60 in the control group.

Figure 2. Study location in the Kannur district of Kerala, India. Source: Wikimedia Commons (2023).

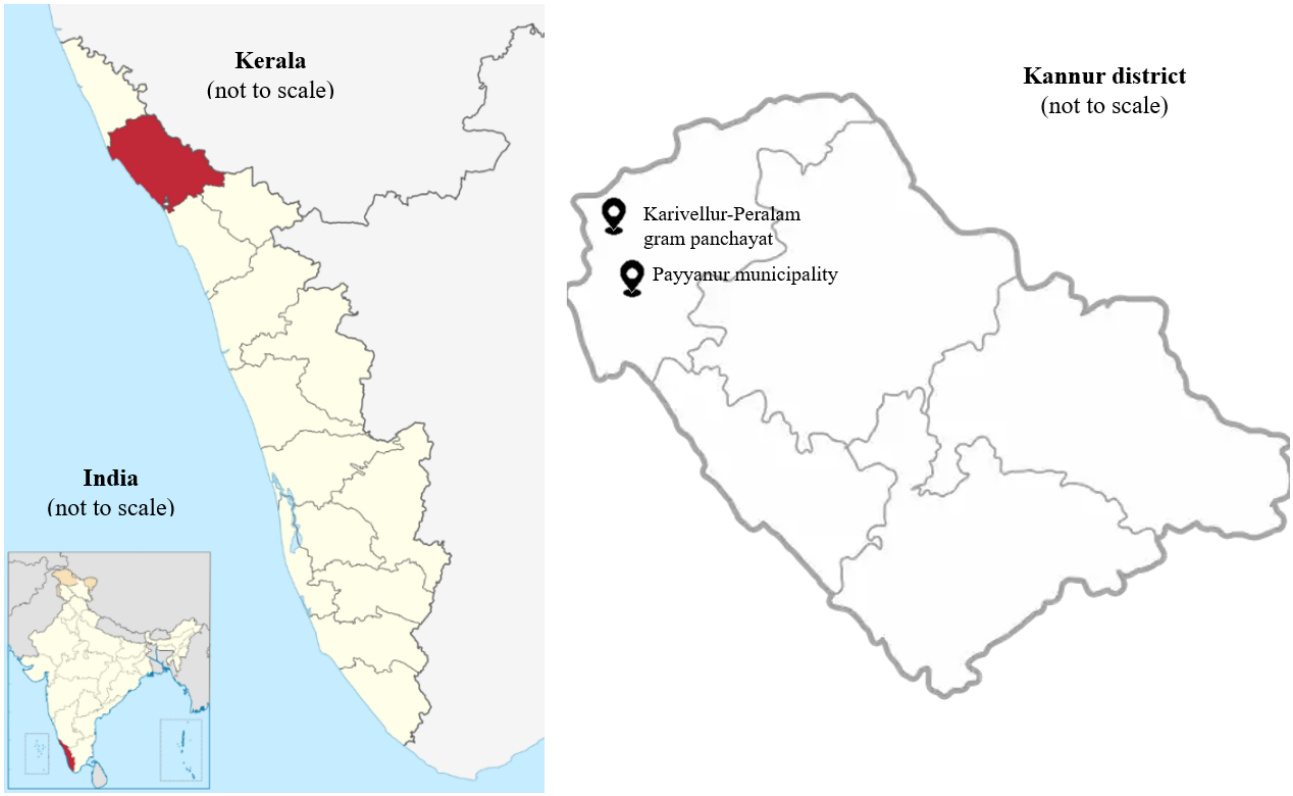

XSL•FO
RenderX

Figure 3. Transparent Reporting of Evaluations With Nonrandomized Designs (TREND) statement flow diagram of the study. DBP: diastolic blood pressure; mHealth: mobile health; SBP: systolic blood pressure.

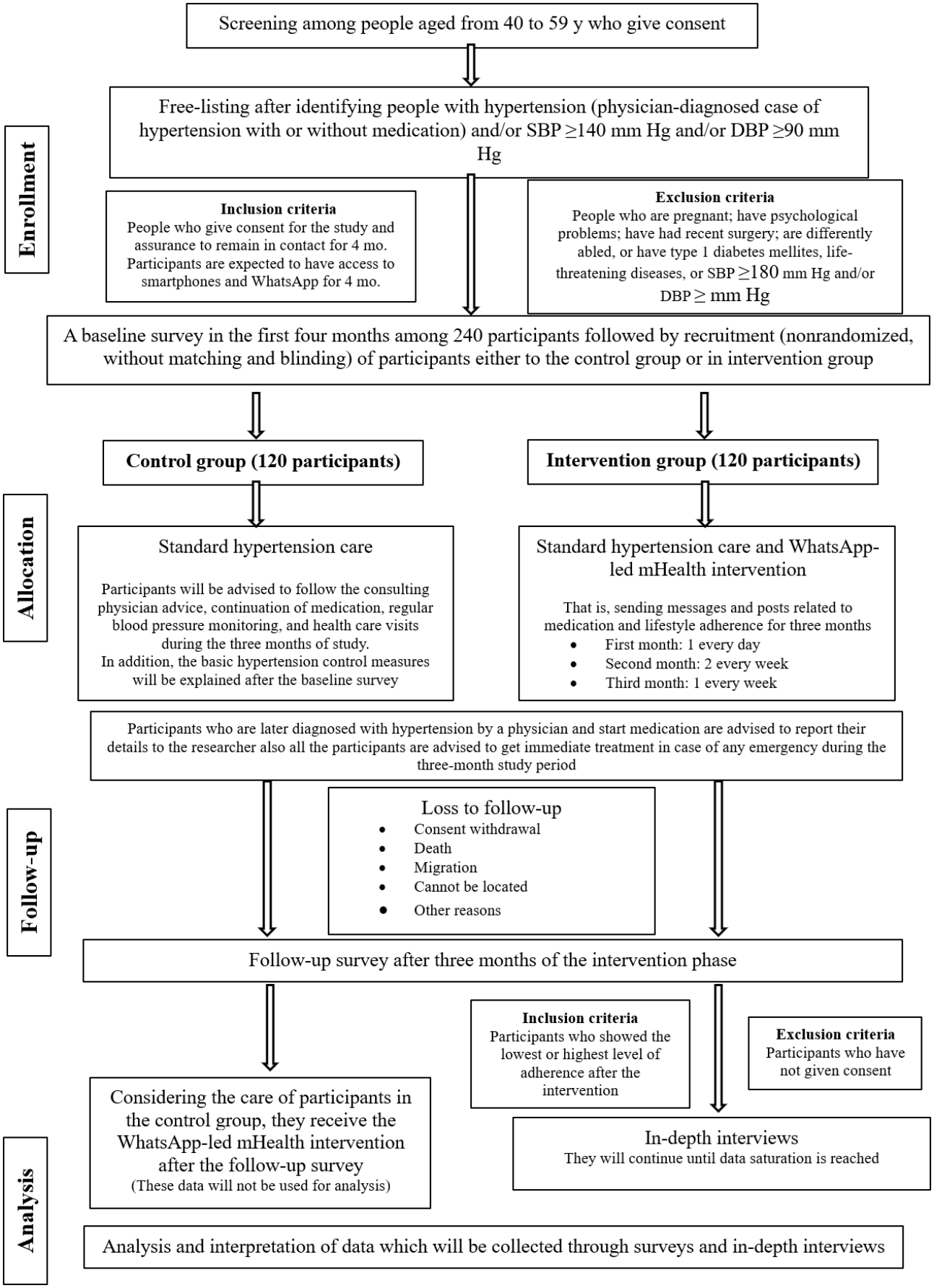

XSL•FO
RenderX

Standard Hypertension Care

All participants will be advised to follow the consulting physician's advice, continue medication, maintain regular blood pressure monitoring, and attend health care visits. Participants will receive health education, including the importance of healthy eating with controlled salt intake and adequate fruit and vegetable intake; regular physical activity; proper medication; and avoiding unhealthy behaviors such as alcohol consumption, tobacco use, and processed food intake. Participants will be reminded about the follow-up survey, which will be conducted 3 months after the baseline survey.

Intervention: WhatsApp-Led mHealth Strategy

A WhatsApp-led mHealth strategy was developed (Figure 4) with participant-specific messages and posters, which were prepared based on standard hypertension management protocols [24-28]. WhatsApp, an instant messaging platform owned by technology conglomerate Meta Platforms, was chosen due to its universal adoption and ease of use and the possibility of sending health messages through text and multimedia (image, voice, and video) formats. The intervention duration was determined to be 3 months based on the review period for hypertension as per the hypertension management protocol of the Department of Health Services, Government of Kerala [12,24].

Figure 4. WhatsApp-led mobile health intervention for medication and lifestyle adherence in participants with hypertension.

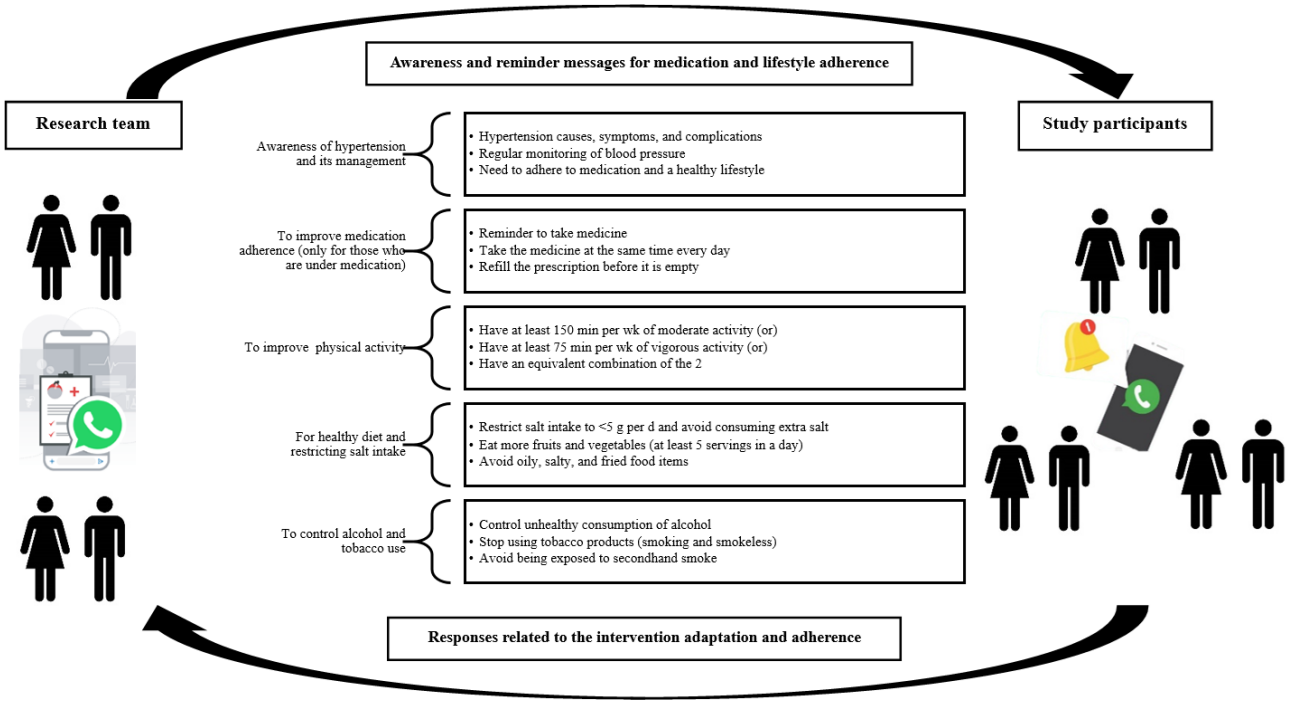

The diagnosis and treatment details of the study participants will be obtained during the baseline survey. Furthermore, the messages will be tailored considering their hypertension treatment status (ie, those who are under medication and those who are not under medication). Among those who are under medication, awareness of healthy lifestyles will be promoted, and they will receive reminders for proper medication adherence, regular consultation with their treating physician, and blood pressure monitoring. Newly diagnosed individuals with hypertension and those who are not in treatment will be sent health messages highlighting the need to adhere to positive health behaviors, consult a physician, and monitor their blood pressure periodically.

The messages for a healthy lifestyle include awareness and reminder messages for keeping healthy eating habits, focusing on the Dietary Approaches to Stop Hypertension diet. The messages also include the need for adequate physical activity and a list of physical activities classified by effort level (mild, moderate, and vigorous). Messages on the importance of reducing mental stress, quitting smoking, avoiding exposure to secondhand smoke, and controlling alcohol intake will also be shared with the participants.

Participants in the intervention group will be advised to follow standard hypertension care. Additionally, participants will be invited to join the intervention through an invitation message and will be listed to receive personalized hypertension management messages via WhatsApp (Figures 5 and 6). The health messages and posters will be shared in the morning at around 6 to 7 AM. Messaging will start within a month following the baseline survey. The intervention will continue from that day for the next 90 days. In the first month (first 30 days), messages will be shared daily; in the second month (middle 30 days), messages will be shared twice a week; and in the third month (last 30 days), messages will be shared once a week (7 days). There will not be a specific day of the week to send messages; days will be calculated from the first day that participants start receiving messages, not from the beginning of a month or a week. The messages will be shared in the participants' preferred language (English or Malayalam), and their responses will be documented to check for participant engagement. The follow-up survey will be conducted after 3

XSL•FO
RenderX

months at the end of the intervention period. Participant engagement with the intervention will be actively monitored throughout the intervention period. A total of four steps will be followed as part of the audit trail to ensure that the messages are seen by the participants: (1) highlighted tick marks for each message shared with the participants, (2) message viewing status from the message information option on WhatsApp, (3) any response (in the form of a like or emoji) or comment from the participants after receiving the message or poster that was shared with them, and (4) monthly confirmation through phone call or home visit to ensure that the participants are receiving the messages. Any instances of disengagement from the participants will be noted and followed up on (through WhatsApp messaging, home visit, or phone call) by the principal investigator to reinforce the importance of being actively involved in the study and address any concerns.

Figure 5. Saved contact and invitation to the participant for the intervention phase. (A) The contact is saved as "Intervention PDM 1" as this is the contact for participant 1 from the intervention group who is hypertensive (physician diagnosed and taking medication). (B) Chat page of participant 1. Translation of the message: "Welcome to this phase of the study. In this phase, you will be given awareness about hypertension and reminders to manage hypertension through WhatsApp. Contact us if you have any confusion or doubts related to hypertension and its management without hesitation."

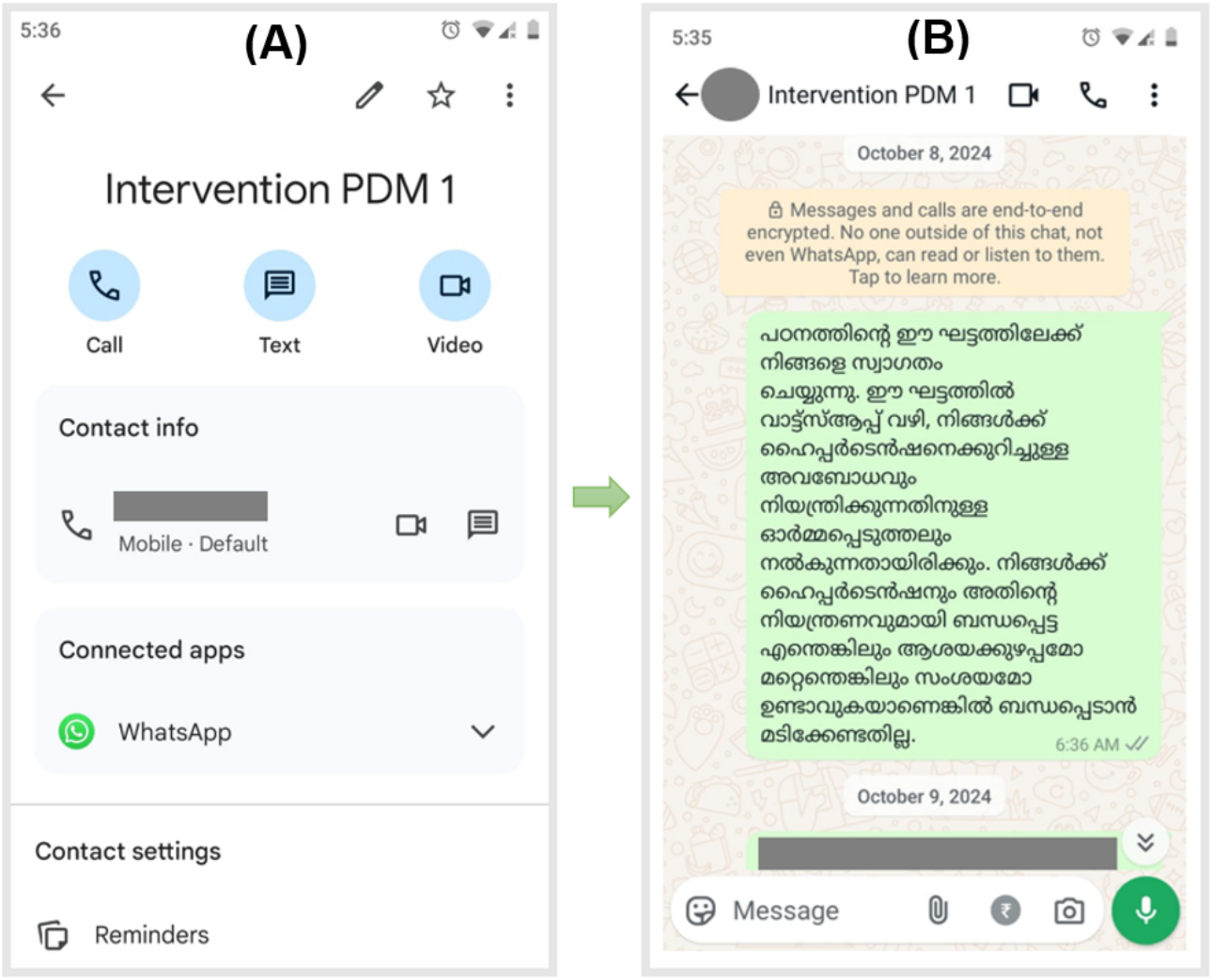

XSL•FO
RenderX

Figure 6. Participant-specific health messages related to healthy lifestyle and medication. (A) Translation of the messages from the chat page of Intervention PDM 7 (a physician-diagnosed participant with hypertension taking antihypertensive medication): "1) Include vegetables, fruits, and whole grains in the diet; 2) The practice of regular physical activities helps to control hypertension, so, fix a time for exercise in the morning or evening; 3) This is a reminder to take your anti-hypertensive medication on time. Try to have an appointment with your consulting doctor at least once every 4 weeks." (B) Translation of the messages from the chat page of Intervention CS 3 (hypertension case found during screening): "Measures to control hypertension—practice exercise daily (at least 150 minutes a week), include fruits and vegetables in the diet (take at least 4-5 servings a day), restrict excess salt intake (consume only less than 5 gm), stop using tobacco, and control consuming alcohol" (from the poster) and "It is important to reduce mental stress to control hypertension" (from the message).

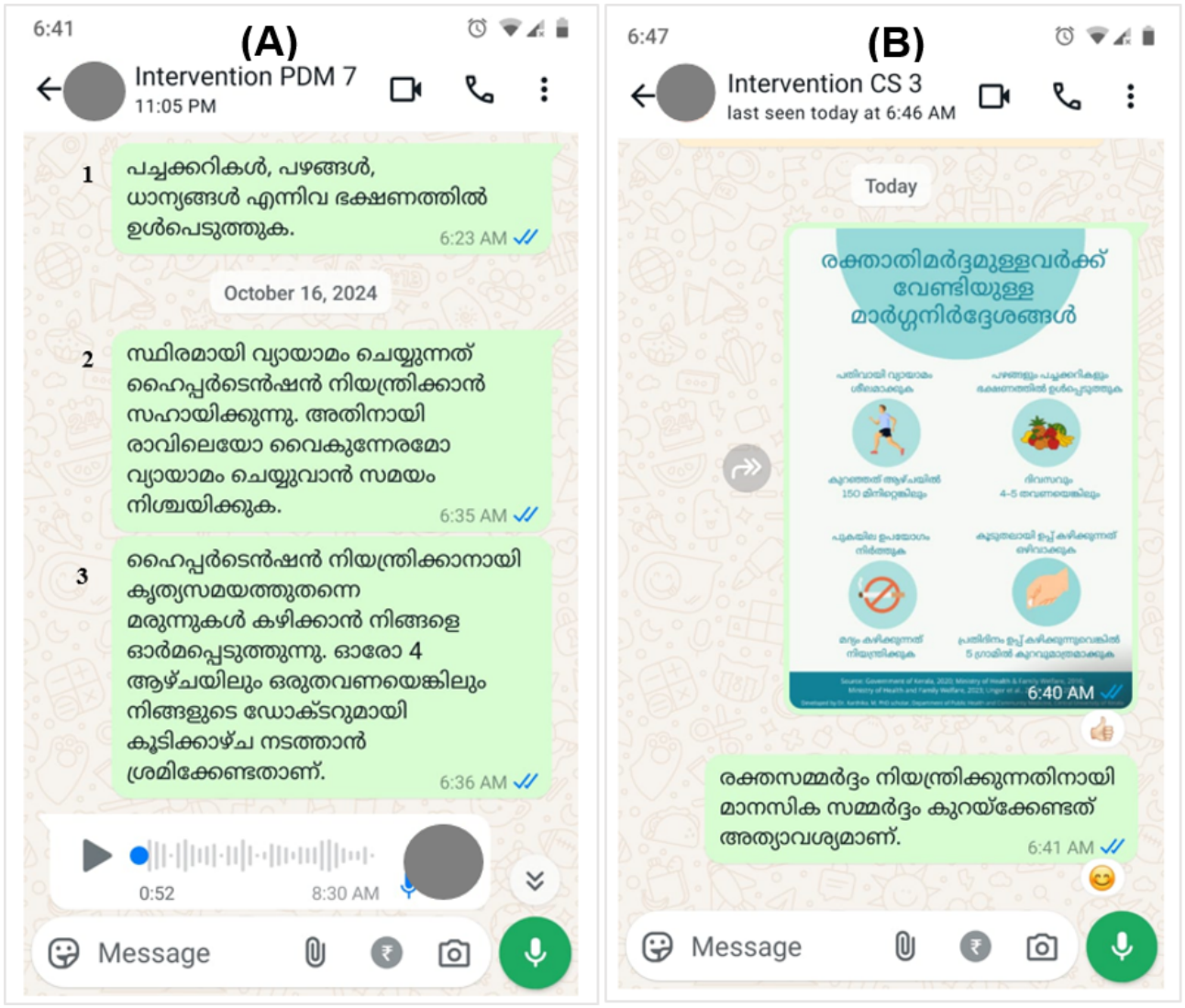

Participants in the control group will be advised to continue standard hypertension care and reminded of the follow-up survey 3 months after the baseline survey. To ensure that the participants in the control group also benefit from being part of the study, they will receive the messages and posters after the follow-up survey.

Data Collection Tools and Techniques

Prospective participants will be screened using a checklist. Blood pressure will be measured as per the guidelines of the World Health Organization and the American Heart Association [29,30] using an Omron HEM-7113 blood pressure monitor [31,32]. The responses to the questionnaire will be self-reported by the participants. Lifestyle factors (ie, diet, physical activity, and tobacco use) and medication adherence will be assessed before and after the intervention period. Lifestyle factors will be assessed through a series of questions on fruit and vegetable intake, salt consumption, processed food intake, physical activity, tobacco use, and alcohol consumption. Medication adherence will be assessed using the Hill-Bone Medication Adherence Scale and a self-developed questionnaire to assess the reasons for nonadherence to medication [33,34]. The *KoboToolbox* application will be used along with the pen-and-paper approach to collect survey data. Additional physical measurements, such as waist and hip circumference, will be taken using the Seca ergonomic retractable tape at baseline and follow-up [31,32,35]. Both the baseline and follow-up surveys will be conducted at the participants' homes. Responses to WhatsApp messages will be recorded in a Microsoft Excel spreadsheet throughout the intervention period. All the tools (Table 1) are prepared in English, translated into Malayalam, and back translated into English. Loss to follow-up will be managed through revisits and prescheduled interviews

XSL•FO
RenderX

based on the participants' convenience except in cases of refusal, migration, or death.

XSL•FO
RenderX

Table . Data collection tools and techniques, along with the operationalization of variables.

Tool or technique and variable	Operationalization	References
Checklist for screening		
Eligibility factors	Age in years; self-reported pregnancy status, disability status, recent surgery, type 1 diabetes mellitus, and life-threatening diseases; access to a smartphone and WhatsApp via a "yes" or "no" question, and if yes, specify whose smartphone they have access to	[4,27,36]
Adapted questionnaires for the baseline survey		
Socioeconomic status	On the basis of the modified Kuppuswamy scale, a score of 26-29 indicates upper socioeconomic class, 16-25 indicates upper-middle socioeconomic class, 11-15 indicates lower-middle socioeconomic class, 5-10 indicates upper-lower socioeconomic class, and <5 indicates lower socioeconomic class	[37]
Sociodemographic characteristics	Self-reported sex, residence, marital status, educational level, occupation, living arrangements, monthly family income, number of children, and household members	[38]
Digital health literacy	Captured using a 6-item Likert-type questionnaire; based on the responses, a higher score denotes a higher digital health literacy	[39]
Intention to use the mHealth[a] strategy	Captured using a 5-domain questionnaire with 20 items; based on the responses, a higher score denotes a higher intention to use the mHealth strategy for hypertension management	[30]
Physical measurements for screening and the follow-up survey		
Hypertension status	A physician-diagnosed case of hypertension and/or when blood pressure is measured and shows SBP[b] ≥140 mm Hg and/or DBP[c] ≥90 mm Hg	[40]
Hypertension emergency	When blood pressure is measured and shows SBP ≥180 mm Hg and/or DBP ≥110 mm Hg	[24]
Physical measurements for the baseline and follow-up surveys		
Waist circumference	Controlled waist circumference ≤90 cm for male individuals and ≤80 cm for female individuals	[35]
Waist-to-hip ratio	Controlled waist-to-hip ratio ≤0.90 for male individuals and ≤0.85 for female individuals	[35]
Adapted questionnaires for the baseline and follow-up surveys		
Awareness of hypertension	A total score will be calculated based on responses to 12 questions; the higher the score, the higher the awareness	[33]
Alcohol consumption and tobacco use	Responses of "yes" or "no"; if yes, then subsequent questions will be asked to assess frequency and pattern of alcohol consumption and the type of tobacco use	[40]
Processed food intake, fruit and vegetable intake, and salt intake	On the basis of the Nova classification system for processed food (based on level of processing) and adherence to the diet	[41,42]
Physical activity	Considered adequate when an individual has at least 150 min per wk of moderate activity or at least 75 min per wk of vigorous activity or an equivalent combination; the leisure-time physical activity will be considered adequate when the individual has at least 300 min per wk	[43]

XSL•FO
RenderX

Tool or technique and variable	Operationalization	References
Medication adherence status and reasons for nonadherence	On the basis of the Hill-Bone Medication Adherence Scale, the higher the total score, the poorer the adherence; in addition, the higher the score, the more reasons for nonadherence to medication	[33,34]
Treatment questionnaire for the baseline survey		
Status of health care service use for diagnosis, treatment, and others	Diagnosis of hypertension and comorbidities, duration of hypertension, frequency of blood pressure monitoring, type of health care facility for service, and status of telemedicine or mHealth service use	[38,40]
Treatment questionnaire for the baseline and follow-up surveys		
Details on antihypertension medication	Medication status, along with frequency and duration of medication; side effects of medication	[38,40]

[a]mHealth: mobile health.

[b]SBP: systolic blood pressure.

[c]DBP: diastolic blood pressure.

Validity and Reliability of the Tools and Measurements

The content and face validity of the tools were improved through literature and peer reviews. All the adapted questionnaires used for the study will be tested for internal consistency using Cronbach α values. Physical measurements and blood pressure measurements will be taken by the principal investigator of the study, who is a trained medical graduate with a Bachelor of Ayurvedic Medicine and Surgery and a Master of Public Health and currently pursuing a PhD in Public Health. Standard equipment and specific cutoffs will be used as per standard protocols for blood pressure [30], waist circumference, and waist-to-hip ratio measurement [35]. The electronic blood pressure monitor will be checked for accuracy using a calibrated sphygmomanometer from the health center.

Study Outcomes

The primary outcomes are the evaluation of the impact of the WhatsApp-led mHealth strategy on blood pressure measurements and adherence to medication and a healthy lifestyle in adults with hypertension aged between 40 and 59 years.

The secondary outcomes are the estimation of the prevalence and control of hypertension; percentage of participants adhering to antihypertensive medication; and prevalence of tobacco use, alcohol consumption, physical activity, diet patterns (such as frequency of fruit and vegetable consumption and processed food intake), and salt intake, as well as the estimation of digital literacy and the intention to use the mHealth strategy among adults with hypertension aged 40 to 59 years.

Statistical Analysis

The survey data collected will be entered into a Microsoft Excel spreadsheet and then exported to SPSS (version 26; IBM Corp) for coding and analysis. Codebooks and frequency tables will be generated to characterize the responses and identify any data entry errors. The missing data will be characterized as missing at random and missing not at random. The data missing at random will be addressed by creating a new variable or recoding an existing variable to account for the missing response. The data missing not at random will be addressed using the last observation carried forward approach (if the missing data were due to loss to follow-up) or regression imputation. Case deletion will be used if there are any missing data observed in multiple outcome variables. The analysis will be conducted adhering to the intention-to-treat approach. Study variables will be developed based on the operationalization plan shown in Table 1.

Descriptive analyses, including frequencies and percentages for categorical variables and means and SDs for continuous variables, will be conducted to understand the sample characteristics. The sociodemographic and health-related characteristics of the intervention and control groups will be compared at baseline. Furthermore, the outcome variables, including medication and lifestyle factors, will be compared between and within the intervention and control groups using baseline and follow-up data.

The comparison of categorical outcomes (medication and lifestyle factors) between and within the intervention and control groups will be conducted using the chi-square test and logistic regression. The regression analysis will be reported as unadjusted odds ratios with 95% CIs. The comparison of continuous outcomes (waist-to-hip ratio, awareness of hypertension, medication adherence score, and duration of physical activity) between the intervention and control groups will be conducted using a 2-tailed independent-sample *t* test (for normally distributed data) and the Mann-Whitney *U* test (for nonnormally distributed data). For comparisons of continuous outcomes within the intervention and control groups, a paired *t* test (for normally distributed data) and a Wilcoxon signed-rank test (for nonnormally distributed data) will be used.

The generalized estimating equations model will be used to assess the effectiveness of the mHealth intervention based on the longitudinal study data (data from the baseline and follow-up surveys). Variables such as time, intervention, time-intervention interaction, and other sociodemographic factors will be considered in the model. The analysis using the generalized estimating equations model will be presented as adjusted odds

XSL•FO
RenderX

ratios with 95% CIs. The significance will be considered at a *P* value of less than .05.

Qualitative Phase of the Study

Study Design

The qualitative phase will follow the phenomenological approach to deeply understand the lived experiences of participants during the period of the intervention.

Sampling

For the qualitative phase of the study, a quota sampling approach will be used among the participants in the intervention group. On the basis of adherence status (ie, high or low adherence after the intervention), participants with hypertension from the intervention group will be categorized into subgroups using a checklist. Medication adherence, as well as adherence to lifestyle modifications (restriction in salt, control of processed food intake, adequate fruit and vegetable intake, and adequate physical activity), will be considered for categorizing participants as having high or low adherence. Sample selection will continue until data saturation is reached.

Data Collection

Face-to-face in-depth interviews will be conducted using an in-depth interview guide developed based on existing models for health care use, medication adherence, behavior adaptation, technology acceptance, and use of mHealth for adherence. The qualitative data will be captured as field notes and voice recordings. The in-depth interviews will be conducted at a location that is convenient for the participants and without the presence of any nonparticipants. The in-depth interview guide will be prepared in English, then translated into Malayalam, and then back-translated into English. All the interviews will be conducted in Malayalam (the vernacular language in the region).

Study Outcomes

The primary outcome is the in-depth understanding of the participants' views on feasibility, facilitators, and challenges of the adoption of the mHealth strategy for achieving adherence.

The secondary outcomes are insights into the facilitators and challenges of adherence to medication and a healthy lifestyle.

Data Analysis

The qualitative data from the in-depth interviews will be analyzed using a thematic approach. The in-depth interview data will be translated and transcribed, and then the transcripts will be read and reread for familiarization. Deductive coding will be conducted based on the aforementioned existing models, and an inductive coding strategy will be adopted when new factors emerge from the in-depth interview transcripts. At least 2 researchers will code and generate themes to explain the phenomena and facts. Member checking with the interview participants and peer debriefing will be conducted to improve the trustworthiness of the qualitative findings. The qualitative analysis software NVivo (version 12; Lumivero) will be used to facilitate coding and theme development.

Ethical Considerations

The study received ethics approval (IHEC/CUK/2024/04) from the Institutional Human Ethics Committee of the Central University of Kerala on August 1, 2024. Informed consent from participants will be obtained upon their enrollment in the study. A personal identification number will be used to ensure the privacy of the participants. The recruited participants will be allowed to withdraw from the study at their request irrespective of the stage at which they are in the study. In case of any unforeseen medical emergencies, the participants will be guided to receive essential treatment from the nearest public health care facility. All the participants from the intervention and control groups will benefit from the intervention at the end of the study. The participants will not receive any financial or material compensation for taking part in the study.

Clinical Trial Registration

The study was registered as a nonrandomized active-controlled trial (phase 1 and phase 2 combined) in the Clinical Trials Registry – India. This study protocol version was approved and received the registration number CTRI/2024/09/073166 on September 2, 2024.

Results

All 240 participants were recruited by July 2025. The study completed data collection in August 2025 and is expected to finish data analysis by April 2026. The study results are expected to be submitted for publication in July 2026. The baseline survey will help understand the participant characteristics, medication status, lifestyle factors, and blood pressure measurements before the intervention. The intention to use the mHealth strategy for hypertension management is also measured at this stage. The follow-up survey data will provide the participant characteristics after the intervention phase. The baseline and follow-up survey data will allow us to study the effectiveness of the intervention on the medication and lifestyle variables. Furthermore, the in-depth interviews will help provide insights into the lived experiences of participants during the period of the intervention, focusing on the adaptation of the mHealth intervention for adherence to medication and a healthy lifestyle. The interviews will also enable the understanding of the perceived effectiveness of the intervention. The study has not received any funding.

Discussion

Expected Findings

The study will develop an mHealth strategy to improve medication and lifestyle adherence for hypertension management. WhatsApp was considered a suitable platform as it is cost-effective and user-friendly. Moreover, it does not need to be separately downloaded and subscribed to like other apps as most people already use it for regular messaging [10,11].

The study seeks to establish a proof of concept for a WhatsApp-led mHealth strategy that can be integrated into hypertension management at the community level. This strategy could function as a means for timely reminders and encourage individuals to change their behavior, improving medication

XSL•FO
RenderX

adherence and lifestyle modifications [40]. Similar approaches could be adopted for individuals with other chronic noncommunicable diseases. Moreover, health care professionals, especially frontline health workers, can contact, provide information to, and receive updates from patients through WhatsApp as instant messaging platforms such as this one can facilitate mHealth services by acting as a medium for communicating health advice and hypertension management guidelines [5,6,9,12,13].

Selecting only 1 taluk from the Kannur district is a major limitation. However, considering that there are similarities in the sociodemographic and economic conditions, linguistic homogeneity (the predominant language is Malayalam), and commonalities in the political and administrative factors, the findings could be generalized to the context of Kerala as a whole. The study adopted a quasi-experimental study design with an intervention group and a control group because the main aspects of a randomized controlled trial, such as randomization, matching, and blinding, cannot be applied in this context. Unlike other trials, this study uses an mHealth intervention, which acts as a platform for sending hypertension awareness messages and reminders for a healthy lifestyle and medication adherence in addition to the routine hypertension care available in health centers. There could be limitations in establishing causal conclusions about the observed effects. However, the qualitative interviews with the participants who received the intervention could help obtain a deeper understanding of their experiences with and adoption of the intervention to adhere to a healthy lifestyle and medication.

The follow-up survey time point of 3 months after the intervention was determined based on the review period (ie, 3 months) for lifestyle and blood pressure measurement changes as per the hypertension management protocol of the Department of Health Services, Government of Kerala [24]. This duration of 3 months is also based on the minimum period required for behavior changes or to develop new habits, including physical activity and dietary changes [12,44,45]. The long-term sustainability of behavior change and the durability of the mHealth intervention will be explored through the qualitative interviews. Even though the intervention ends after 3 months, the ability to keep the messages for an indefinite time, as well as the option to create a backup in WhatsApp, allows the participants to go through the messages again if needed after the intervention period.

WhatsApp is a ubiquitous platform that is not primarily meant for medical communication. However, given the real-time nature of WhatsApp-based intervention and health messaging specific to the individual participant, this study is expected to be effective in improving adherence to medication and healthy lifestyle and have better participant engagement with the mHealth intervention.

The individualized delivery method, with a person-to-person (researcher-to-participant) messaging system rather than a WhatsApp group or community, will help minimize contamination between the intervention and control groups. Additionally, all participants will be recruited from their homes and monitored regularly during the follow-up period until the end-line survey. In addition, during sample size estimation, the researcher accounted for the cluster effect to obtain an adequate sample size to prevent the reduction in the samples' power due to contamination.

The findings from the study will contribute to improving the planning and implementation of mHealth services for managing chronic diseases such as hypertension in India and other similar settings. Tailoring health messages specifically to the individuals could be challenging. However, prestructuring the messages will be helpful for easy dissemination. The researcher will recruit 4 batches of approximately 30 participants each into the control and intervention groups and then start sharing the messages individually with the participants in the intervention group. The entire intervention will be completed in these 4 batches. This is only for the convenience of the participants and the researcher to reduce the gap between the baseline survey and the first day of the intervention. The end-to-end encryption option of WhatsApp helps protect each individual's chat data, and saving the participants' contact information using code names will help protect their privacy.

Dissemination Plan

The findings from the study will be presented at national and international conferences and published in peer-reviewed journals. A study report will be submitted to the Central University of Kerala library, the Institutional Human Ethics Committee of the Central University of Kerala, and the Clinical Trials Registry – India. In addition, study summaries will be provided to the officials of the administrative units selected for the study.

Conclusions

The core focus of our study is to analyze both quantitatively and qualitatively the impact of a WhatsApp-led mHealth strategy on blood pressure measurements and adherence to medication and a healthy lifestyle in adults with hypertension aged between 40 and 59 years. The intervention is expected to improve hypertension control among study participants. Additionally, the participants are expected to have increased awareness, adherence to medication, healthy behaviors, and regular follow-ups or monitoring of blood pressure. The study will also help understand the participants' acceptance of the WhatsApp-led mHealth strategy for the self-management of hypertension.

Acknowledgments

The authors acknowledge the research scholars and faculty members of the Department of Public Health and Community Medicine, Central University of Kerala, Kasaragod, Kerala, India.

XSL•FO
RenderX

Funding

The study has not received any funding.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

KM and PBK conceptualized the study; developed the intervention materials, tables, and figures; and developed and critically reviewed the initial version of the manuscript. Both authors have read, finalized, and approved the final version of the submitted manuscript.

Conflicts of Interest

None declared.

References

1. Hypertension. WHO India. URL: https://www.who.int/india/health-topics/hypertension [accessed 2024-03-26]
2. Balasubramanian A, Nair SS, Rakesh PS, Leelamoni K. Adherence to treatment among hypertensives of rural Kerala, India. J Family Med Prim Care 2018;7(1):64-69. [doi: 10.4103/jfmpc.jfmpc_423_16] [Medline: 29915735]
3. Hamrahian SM, Maarouf OH, Fülöp T. A critical review of medication adherence in hypertension: barriers and facilitators clinicians should consider. Patient Prefer Adherence 2022;16:2749-2757. [doi: 10.2147/PPA.S368784] [Medline: 36237983]
4. Feinberg L, Menon J, Smith R, Rajeev JG, Kumar RK, Banerjee A. Potential for mobile health (mHealth) prevention of cardiovascular diseases in Kerala: a population-based survey. Indian Heart J 2017;69(2):182-199. [doi: 10.1016/j.ihj.2016.11.004] [Medline: 28460766]
5. Cao W, Milks MW, Liu X, et al. mHealth interventions for self-management of hypertension: framework and systematic review on engagement, interactivity, and tailoring. JMIR Mhealth Uhealth 2022 Mar 2;10(3):e29415. [doi: 10.2196/29415] [Medline: 35234655]
6. Khoong EC, Olazo K, Rivadeneira NA, et al. Mobile health strategies for blood pressure self-management in urban populations with digital barriers: systematic review and meta-analyses. NPJ Digit Med 2021 Jul 22;4(1):114. [doi: 10.1038/s41746-021-00486-5] [Medline: 34294852]
7. Li R, Liang N, Bu F, Hesketh T. The effectiveness of self-management of hypertension in adults using mobile health: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Mar 27;8(3):e17776. [doi: 10.2196/17776] [Medline: 32217503]
8. Rehman H, Kamal AK, Morris PB, Sayani S, Merchant AT, Virani SS. Mobile health (mHealth) technology for the management of hypertension and hyperlipidemia: slow start but loads of potential. Curr Atheroscler Rep 2017 Mar;19(3):12. [doi: 10.1007/s11883-017-0649-y] [Medline: 28210974]
9. Xu H, Long H. The effect of smartphone app-based interventions for patients with hypertension: systematic review and meta-analysis. JMIR Mhealth Uhealth 2020 Oct 19;8(10):e21759. [doi: 10.2196/21759] [Medline: 33074161]
10. Busch PA, Hausvik GI, Ropstad OK, Pettersen D. Smartphone usage among older adults. Comput Human Behav 2021 Aug;121:106783. [doi: 10.1016/j.chb.2021.106783]
11. Kamel Boulos MN, Giustini DM, Wheeler S. Instagram and WhatsApp in health and healthcare: an overview. Future Internet 2016 Sep;8(3):37. [doi: 10.3390/fi8030037]
12. Alzahrani SA, Bin Muammar MF, Bin Muammar AF, Alolah A, Almutawa M. The adoption and acceptance of mHealth interventions for self-management of hypertension among adult patients: a systematic review. Cureus 2022 Nov 16;14(11):e31584. [doi: 10.7759/cureus.31584] [Medline: 36540533]
13. Gong K, Yan YL, Li Y, et al. Mobile health applications for the management of primary hypertension: a multicenter, randomized, controlled trial. Medicine (Baltimore) 2020 Apr;99(16):e19715. [doi: 10.1097/MD.0000000000019715] [Medline: 32311957]
14. Crim MT, Yoon SS, Ortiz E, et al. National surveillance definitions for hypertension prevalence and control among adults. Circ Cardiovasc Qual Outcomes 2012 May;5(3):343-351. [doi: 10.1161/CIRCOUTCOMES.111.963439] [Medline: 22550130]
15. Aday LA, Andersen R. A framework for the study of access to medical care. Health Serv Res 1974;9(3):208-220. [Medline: 4436074]
16. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep;27(3):425-478. [doi: 10.2307/30036540]
17. Rosenstock IM. Historical origins of the Health Belief Model. Health Educ Monogr 1974;2(4):328-335. [doi: 10.1177/109019817400200403]

XSL•FO
RenderX

18. Agarwal S, LeFevre AE, Lee J, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
19. Lyles C, Des Jarlais DC, Crepaz N. TREND statement checklist. Centers for Disease Control and Prevention. 2004. URL: https://stacks.cdc.gov/view/cdc/149677 [accessed 2026-03-24]
20. Tovin MM, Wormley ME. Systematic development of standards for mixed methods reporting in rehabilitation health sciences research. Phys Ther 2023 Nov 4;103(11):pzad084. [doi: 10.1093/ptj/pzad084] [Medline: 37672215]
21. Saju MD, Allagh KP, Scaria L, Joseph S, Thiyagarajan JA. Prevalence, awareness, treatment, and control of hypertension and its associated risk factors: results from baseline survey of SWADES family cohort study. Int J Hypertens 2020 Apr 13;2020:4964835. [doi: 10.1155/2020/4964835] [Medline: 32351729]
22. R. K. S, T. R. M, Divakaran B, Karunakaran U, A. K. J. Hypertension: prevalence, awareness, treatment and control in a rural area of North Kerala, India. Int J Community Med Public Health 2017;4(10):3561-3567. [doi: 10.18203/2394-6040.ijcmph20174184]
23. Thankappan KR, Sivasankaran S, Mini GK, Daivadanam M, Sarma PS, Abdul Khader S. Impact of a community based intervention program on awareness, treatment and control of hypertension in a rural Panchayat, Kerala, India. Indian Heart J 2013;65(5):504-509. [doi: 10.1016/j.ihj.2013.08.023] [Medline: 24206872]
24. Hypertension protocol. Government of Kerala. 2020. URL: https://dhs.kerala.gov.in/wp-content/uploads/2020/12/Kerala-A1-Hypertension-Protocol.pdf [accessed 2026-03-24]
25. Hypertension: screening, diagnosis, assessment, and management of primary hypertension in adults in India. : Ministry of Health & Family Welfare, India; 2016 URL: https://nhm.gov.in/images/pdf/guidelines/nrhm-guidelines/stg/Hypertension_full.pdf [accessed 2024-03-25]
26. Operational guidelines: national programme for prevention and control of non-communicable diseases (2023-2030). Ministry of Health & Family Welfare, India. 2023. URL: https://www.nhmmizoram.org/upload/Revised%20Operational%20Guidelines%20of%20NP-NCD%20%282023-2030%29.pdf [accessed 2026-03-24]
27. Unger T, Borghi C, Charchar F, et al. 2020 International Society of Hypertension global hypertension practice guidelines. Hypertension 2020 Jun;75(6):1334-1357. [doi: 10.1161/HYPERTENSIONAHA.120.15026] [Medline: 32370572]
28. Hypertension. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/hypertension [accessed 2023-04-06]
29. Muntner P, Shimbo D, Carey RM, et al. Measurement of blood pressure in humans: a scientific statement from the American Heart Association. Hypertension 2019 May;73(5):e35-e66. [doi: 10.1161/HYP.0000000000000087] [Medline: 30827125]
30. STEPwise approach to NCD risk factor surveillance (STEPS). World Health Organization. URL: https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/steps [accessed 2024-03-04]
31. Mathews E, Sathish T, Joseph A, et al. Effectiveness and implementation of a lifestyle modification intervention for women with isolated impaired fasting glucose: study protocol for a hybrid type 2 study in Kerala, India. Wellcome Open Res 2022;7:62. [doi: 10.12688/wellcomeopenres.17631.1] [Medline: 36865369]
32. STEPS manual. World Health Organization. URL: https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/steps/manuals [accessed 2024-06-10]
33. Pirasath S, Kumanan T, Guruparan M. A study on knowledge, awareness, and medication adherence in patients with hypertension from a tertiary care centre from northern Sri Lanka. Int J Hypertens 2017;2017:9656450. [doi: 10.1155/2017/9656450] [Medline: 29230325]
34. Kim MT, Hill MN, Bone LR, Levine DM. Development and testing of the Hill-Bone Compliance to High Blood Pressure Therapy Scale. Prog Cardiovasc Nurs 2000;15(3):90-96. [doi: 10.1111/j.1751-7117.2000.tb00211.x]
35. Waist circumference and waist-hip ratio: report of a WHO expert consultation. : World Health Organization; 2011 URL: https://www.who.int/publications/i/item/9789241501491 [accessed 2026-03-24]
36. Radhakrishnan M, Nagaraja SB. Modified Kuppuswamy socioeconomic scale 2023: stratification and updates. Int J Community Med Public Health 2023;10(11):4415-4418. [doi: 10.18203/2394-6040.ijcmph20233487]
37. Longitudinal Ageing Study in India (LASI). International Institute for Population Sciences. URL: https://www.iipsindia.ac.in/lasi [accessed 2021-12-18]
38. Nelson LA, Pennings JS, Sommer EC, Popescu F, Barkin SL. A 3-item measure of digital health care literacy: development and validation study. JMIR Form Res 2022 Apr 29;6(4):e36043. [doi: 10.2196/36043] [Medline: 35486413]
39. Seethamraju R, Diatha KS, Garg S. Intention to use a mobile-based information technology solution for tuberculosis treatment monitoring – applying a UTAUT model. Inf Syst Front 2018 Feb;20:163-181. [doi: 10.1007/s10796-017-9801-z]
40. IHCI training manual. Indian Hypertension Control Initiative. URL: https://www.ihci.in/resources/training-modules [accessed 2026-03-24]
41. Dubasi SK, Ranjan P, Arora C, et al. Questionnaire to assess adherence to diet and exercise advices for weight management in lifestyle-related diseases. J Family Med Prim Care 2019 Feb;8(2):689-694. [doi: 10.4103/jfmpc.jfmpc_338_18] [Medline: 30984696]

XSL•FO
RenderX

42. Monteiro CA, Cannon G, Lawrence M, Costa Louzada ML, Pereira Machado P. Ultra-processed foods, diet quality, and health using the NOVA classification system. Food and Agriculture Organization of the United Nations. 2019. URL: https://www.fao.org/3/ca5644en/ca5644en.pdf [accessed 2023-12-13]
43. Physical activity. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2026-03-24]
44. Arlinghaus KR, Johnston CA. The importance of creating habits and routine. Am J Lifestyle Med 2018 Dec 29;13(2):142-144. [doi: 10.1177/1559827618818044] [Medline: 30800018]
45. Gardner B, Lally P, Wardle J. Making health habitual: the psychology of "habit-formation" and general practice. Br J Gen Pract 2012 Dec;62(605):664-666. [doi: 10.3399/bjgp12X659466] [Medline: 23211256]

Abbreviations

mHealth: mobile health

Edited by J Sarvestan; submitted 25.Jul.2025; peer-reviewed by J Menon; revised version received 06.Mar.2026; accepted 06.Mar.2026; published 20.Apr.2026.

Please cite as:
Maniyara K, Kodali PB
WhatsApp-Led mHealth Intervention for Hypertension Management Among Adults Aged 40 to 59 Years in Kerala, India: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e81307
URL: https://www.researchprotocols.org/2026/1/e81307
doi:10.2196/81307

© Karthika Maniyara, Prakash Babu Kodali. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Temporal Associations Between Body Checking and Eating Pathology in Adolescent Girls with Binge-Spectrum Eating Disorders: Protocol for an Ecological Momentary Assessment Study

Jannah Moussaoui[1,2,3], BS; Stephanie Manasse[1,3], PhD; Laura D'Adamo[1,3,4], MS

[1]Center for Weight, Eating, and Lifestyle Science, Drexel University, Philadelphia, PA, United States

[2]Department of Psychological and Brain Sciences, Washington University in St. Louis, 1 Brookings Dr, St. Louis, MO, United States

[3]Center for Healthcare Delivery Science, Nemours Children's Health, Wilmington, DE, United States

[4]School of Medicine, Washington University in St. Louis, St. Louis, MO, United States

Corresponding Author:
Jannah Moussaoui, BS
Center for Weight, Eating, and Lifestyle Science, Drexel University, Philadelphia, PA, United States

Abstract

Background: Up to 92% of individuals with eating disorders (EDs) report engaging in body checking behaviors (eg, repeated self-weighing and pinching of various body parts) to assess their weight and shape. These behaviors contribute to increased body dissatisfaction, negative affect, and dietary restriction, thereby maintaining ED symptomology.

Objective: This study aims to first characterize the types and frequency of body checking behaviors (eg, self-weighing, pinching) reported among adolescent girls with binge-spectrum EDs during a 21-day ecological momentary assessment (EMA) protocol. The second aim is to explore the prospective associations between body checking and cognitive and behavioral ED symptoms, namely body dissatisfaction, fear of weight gain, dietary restraint, dietary restriction, compensatory behaviors, and binge eating. The third aim is to assess whether body checking behaviors show reactive effects (ie, produces change in the behavior subject to monitoring) to EMA, such that they decrease over time.

Methods: The study will recruit 70 adolescent girls aged 14 to 19 years with clinically significant binge eating. Participants will complete a semistructured interview and a series of self-report measures at baseline to assess ED pathology. Then, participants will complete 5 daily EMA surveys to track body checking behaviors and related ED symptoms over 21 days.

Results: Recruitment began in January 2025, with data collection expected to conclude in March 2026. Linear and generalized mixed-effects models will be used to evaluate concurrent and prospective associations between various ED symptoms.

Conclusions: This study will provide insights into the patterns and impacts of body checking behaviors among adolescent girls with binge-spectrum EDs. If body checking behaviors reduce in response to EMA, digital self-monitoring could be a scalable and cost-effective strategy for ED treatment and prevention. The findings may also inform the development of momentary interventions targeting body checking behaviors to mitigate ED symptoms. Future research should extend these observations over longer periods and include male participants to generalize findings across genders.

International Registered Report Identifier (IRRID): PRR1-10.2196/73447

(JMIR Res Protoc 2026;15:e73447) doi:10.2196/73447

KEYWORDS

binge eating; adolescents; body checking; eating disorders; ecological momentary assessment

Introduction

Background

Binge-spectrum eating disorders (EDs), including bulimia nervosa (BN) and binge eating disorder (BED), are serious psychological disorders characterized by recurrent episodes of binge eating (ie, eating an excessive amount of food while experiencing a sense of loss-of-control over one's eating) and may be accompanied by behaviors intended to compensate for these episodes (eg, vomiting or restricting food intake). Binge-spectrum EDs typically start during adolescence [1] and have a 7% to 8% prevalence among adolescents [2,3]. Critically, gold-standard treatments for binge-spectrum EDs fail to produce remission in 60% to 75% of adolescents [4-6], and only about 50% of individuals stay remitted 5 years following treatment [7]. These low remission rates indicate that available treatments do not sufficiently address all mechanisms maintaining the ED.

XSL•FO
RenderX

Indeed, following treatment, many continue to experience ED symptoms, despite no longer meeting diagnostic criteria for an ED [6,8]. In particular, body image disturbances remain elevated following treatment [9,10], and studies have identified elevated body image disturbances as a risk factor for relapse [11]. These findings are consistent with the transdiagnostic model of cognitive behavioral therapy (CBT) for EDs, which points to the overvaluation of weight, shape, eating, and their control as the core psychopathology underlying these disorders [12], thereby suggesting that until overvaluation is addressed, individuals may continue to experience ED symptoms. Given the low remission rates among adolescents with binge-spectrum EDs, it is critical to understand how the overvaluation of shape, weight, and eating manifests among adolescents and how it relates to ED symptomology.

One behavioral manifestation of overvaluation with weight, shape, eating, and their control is body checking [13]. Body checking refers to a broad range of behaviors used to assess one's weight or shape (eg, self-weighing, checking how clothes fit, pinching or feeling various body parts) [14]. Based on studies among adults with EDs, these behaviors have a 90% prevalence [13,15]; are often engaged in compulsively; are brief, typically lasting less than two minutes [13]; and can occur multiple times in a given day. For example, in a transdiagnostic, qualitative study of adults with anorexia nervosa and BN, 92% of adults endorsed body checking and 33% of adults reported body checking 16 or more times in a given day [13]. Likewise, in 2 ecological momentary assessment (EMA; ie, an ecologically valid data sampling approach that prompts participants multiple times over the study period) studies among female college students with high weight and shape concern, participants body checked an average of 16 to 28 times a day [16,17]. However, no study has yet attempted to estimate how frequently body checking occurs among adolescents, let alone those with binge-spectrum EDs.

Critically, body checking behaviors are typically aimed toward disliked body parts [14], and selectively attending to disliked body parts and their perceived changes may contribute to increased ED cognitions, particularly body dissatisfaction and body image disturbances, and behaviors [18]. In transdiagnostic adult ED samples, body checking is associated with increased dietary restriction, meal skipping, use of fluids to curb appetite [19,20], increased negative affect [21,22], and shape and weight overconcern [15]; among those with binge-spectrum EDs, body checking is additionally indirectly tied to binge eating through increased body shame and appearance anxiety [23]. Further, 1 longitudinal study found that in a community sample of adolescents, higher body checking at 1 time point was associated with the development of ED symptoms 4 months later [24], underscoring the need for early and targeted interventions for body checking behaviors in this age group. Indeed, adolescents report high rates of body image dissatisfaction [25], a prominent driver of ED pathology among adolescents [26-28]. In recent years (ie, 2017 - 2022), the prevalence of EDs has dramatically increased among adolescent female participants [29]. Despite this, few studies have examined the associations between body checking behaviors and ED symptomology among adolescents with binge-spectrum EDs.

Given the strong theoretical links between body checking behaviors and ED symptomology, body checking may, itself, be a worthwhile target for intervention. In fact, 2 EMA studies among female college students found small decreases in body checking behaviors as well as body dissatisfaction and internalization of the thin ideal. These reductions can be understood as effects of self-monitoring. In cognitive-behavioral models of psychopathology, the self-monitoring of behaviors and psychological states is key for reducing engagement in undesirable behaviors and cognitions [30], including body dissatisfaction [31], ruminative thinking [32], suicidal ideation [33], depressive symptoms [34], sedentary behaviors [35], and substance use [36]. Self-monitoring theoretically increases the awareness of the frequency, antecedents, and consequences of problematic thoughts and behaviors [30], allowing individuals to make changes either independently or with the support of a clinician. It is possible that during EMA, participants experience reductions in body checking frequency because reporting these behaviors on EMA surveys similarly increases the awareness for the frequency, antecedents, and consequences of these behaviors. Thus, in addition to examining associations between body-checking behaviors and ED symptomology in the real world, it is also worthwhile to investigate how these behaviors may decrease in response to self-monitoring. Notably, if body checking behaviors show reductions during EMA in adolescents with binge-spectrum EDs, and if these reductions are associated with improvements in other ED symptoms, then digital self-monitoring may be a powerful tool to formalize body checking as an intervention target in evidence-based treatments such as enhanced cognitive behavior therapy (CBT-E) or as a stand-alone treatment. Given the large ED treatment gap (eg, less than 20% of individuals with EDs receive care [37]), the scalability of digital self-monitoring could prove valuable among those engaging in body checking [16,17] prior to the receipt of more intensive treatment or as a prevention tool, particularly among adolescents.

Present Study

Despite the relevance of body checking as a maintenance factor for ED pathology, relatively little is known about the nature of body checking and its prospective associations with ED cognitions and behaviors among adolescents with binge-spectrum EDs. Thus, this study aims to fill 3 specific gaps. The first aim of this study is to characterize the types and frequency of body checking behaviors among adolescent girls (ages 14 - 19 years) with binge-spectrum EDs during a 3-week EMA protocol. Although body checking behaviors are present across the ED spectrum, they have most frequently been studied within restrictive EDs, despite the evidence of their presence within binge-spectrum EDs [13] and associations with binge eating [23]. Similarly, while body checking is prevalent among boys [24], consistent with other studies [16,17], we elected to restrict this study to girls because measures in this study would fail to capture the different behaviors used to body check and their associations with ED behaviors in boys [38,39]; further, resource constraints (ie, limited budget and personnel) precluded the recruitment of boys. We expect that participants will report a broad range of body checking behaviors and that they will engage in body checking behaviors several times a day on

XSL•FO
RenderX

average. The second aim of the study is to test prospective associations between body checking and other ED symptoms. Consistent with the CBT model for EDs, we expect that body checking will be concurrently and prospectively associated with cognitive (ie, body dissatisfaction, fear of weight gain) and behavioral (ie, dietary restriction or restraint, compensatory behaviors) ED symptoms. The third aim is to assess whether the self-monitoring of body checking produces decreases in body checking over time. We hypothesize that body checking will decrease over the study period and that participants will report lower frequencies of body checking on the last 3 days of the protocol compared to the first 3 days [16,17].

Methods

Ethical Considerations

Prior to study enrollment, informed consent will be obtained from the parent or from the adolescents if they are 18 years or older. Adolescents under 18 years will also need to provide assent to participate. All study procedures have received approval from the Drexel University Institutional Review Board, under protocol number 2401010341. All participants will be compensated up to US $275 for their participation (see the "Ecological Momentary Assessment" section for more details). Data will be stored securely on Drexel's encrypted servers, and analytic datasets will be stripped of identifiers.

Participants

We will enroll 70 adolescent girls with clinically significant binge eating (ie, engaging in at least 1 binge eating episode per week over the past 12 weeks). Both subjective and objective binge episodes will be counted given literature demonstrating that the feeling of loss of control, rather than the size of the binge episode, is more closely linked to ED severity [40,41]. Thus, to be included in the study, participants will need to be between 14 and 19 years old; be able to speak, read, and write English; have a smartphone they would be willing to complete daily surveys on for the course of the study; and were assigned female at birth. Participants will be excluded if they are currently receiving treatment for binge eating and/or weight loss or if they will be starting treatment during the course of the study, if they have begun a course of or changed the dosage of any medication known to affect appetite or body composition in the past 3 months, if they are currently pregnant or have plans to become pregnant, or if they have a developmental disability (eg, autism spectrum disorder) or are currently experiencing severe psychopathology that would limit their ability to engage in the study or would otherwise pose a risk to their safety (eg, severe depression, substance dependence, active psychotic episode, imminent suicide risk). The final 2 exclusion criteria will be assessed in consultation with a licensed clinical psychologist. Although such criteria may exclude certain individuals who represent the broader clinical population, these criteria are in place to ensure participant safety during this observational study.

Study Procedures

Recruitment and Screening

Participants will be recruited nationwide through online advertising (eg, on social media platforms) and through a registry of adolescents who are either waitlisted or in the screening process for participation in a clinical trial for binge-spectrum EDs at Drexel University. Participants who appear to be eligible following the screening process and are interested in participating will be invited to a 2-hour baseline assessment with study staff held either virtually via Zoom or in person.

Baseline Measures

Participant characteristics, namely race, ethnicity, age, socioeconomic status, year in school, and history of ED diagnoses and treatment receipt, will be measured at baseline through a demographics survey.

Diagnostic status will be established at baseline using the binge and compensatory modules from the eating disorder examination [42], a well-validated interview used to evaluate ED pathology, which will be administered by trained study staff. Participants who report 12 or more binge episodes but fewer than 12 compensatory behaviors will be classified as having BED, whereas those with 12 or more compensatory behaviors will be classified as having BN. If participants report at least 12 objectively large episodes, they will be classified as "full-threshold," whereas those with fewer objectively large episodes will be classified as "sub-threshold" [43].

Body concern will be measured at baseline using the Body Shape Questionnaire [44], a 34-item self-report questionnaire. Items inquire about the frequency of concerns related to one's body and are rated on a 6-point scale ranging from 1 ("Never") to 6 ("Always"). A total score will be calculated by summing all of the items.

Body checking will be measured at baseline using the Body Checking Questionnaire [45], a 23-item self-report measure. Items inquire about the frequency of various body checking behaviors and are rated on a 5-point scale ranging from 1 ("Never") to 5 ("Very often"). The Body Checking Questionnaire contains 3 subscales that assess body checking related to overall appearance, specific body parts, and idiosyncratic rituals. A total score will be calculated by summing all of the subscale items.

ED pathology will be measured at baseline using the Eating Disorder Examination Questionnaire [46], a 28-item self-report questionnaire. Items pertaining to frequency are rated on a 7-point scale ranging from 0 ("No days") to 6 ("Every day"); items pertaining to distress are rated on a 7-point scale ranging from 0 ("Not at all") to 6 ("Markedly"). The Eating Disorder Examination Questionnaire contains 4 subscales measuring eating concern, weight concern, shape concern, and dietary restraint. A global score will be calculated by taking the mean of all the subscale scores.

XSL•FO
RenderX

Ecological Momentary Assessment

EMA surveys will be facilitated by the EMA app Avicenna. Following the completion of the interview and self-report measurements at the baseline visit, research staff will assist with installing the EMA app and how to use it. During the 3-week EMA protocol, signal-contingent (ie, participants are prompted to complete surveys at semirandom times) and event-contingent (ie, participants will be instructed to initiate surveys when they engage in binge eating or a compensatory behavior if it occurs outside signal-contingent survey windows) will be used. The timing of the first and last EMA surveys will be tailored to participants' self-reported weekday or weekend bedtimes and wake times such that they arrive at random times within 1 hour of waking up and within 1 hour of going to bed. The remaining 3 surveys will come between those windows, excluding hours during which adolescents are in school. To promote high compliance, participants will complete a test day ("day zero") and will be provided with feedback from a research assistant. Then, participants will receive 5 signal-contingent surveys a day for 21 days, for a total of 105 surveys. Surveys will pertain to engagement in body checking and both cognitive and behavioral symptoms of EDs (see questions below). Participants will earn US $2.14 for each completed survey and earn a US $50 bonus if they complete 85% or more surveys for a maximum possible compensation of US $275. To promote high response rates, participants will receive personalized reminder texts coupled with memes and GIFs throughout the EMA period [47]. Although such engagement strategies may have unintended effects on responses, at least 1 EMA study with adolescents revealed that such strategies appear to increase compliance without increasing careless responding [48].

Body checking will be measured using the same set of questions included in previous EMA studies [16,17] and will require participants to report the number of times they had engaged in the following behaviors since the last survey: (1) weighing oneself, (2) feeling thighs for fatness or checking to see if thighs spread while sitting down, (3) sucking in the stomach, (4) feeling or pinching stomach to measure fatness, (5) comparing one's body to other individuals, (6) checking body size in the mirror, and (7) checking for fat jiggling. In aim 1 (characterizing body checking), engagement in specific forms of body checking will be examined descriptively. In aim 2 (prospective associations between body checking and ED symptoms), body checking will be collapsed and treated as a binary variable representing engagement in *any* form of body checking at each survey. In aim 3 (body checking reactivity during EMA), body checking will be defined as a count variable representing the number of times each participant engaged in *any* form of body checking in a given day.

Body satisfaction will be measured through a single question: "Right now, how satisfied do you feel with your body shape or weight?" Responses will be rated on a Likert scale ranging from 1 ("Not at all satisfied") to 5 ("Very satisfied"). Body satisfaction will be treated as a continuous variable representing current degree of body satisfaction at each survey.

Fear of weight gain will be measured through a single question: "To what extent are you currently afraid of gaining weight?" Responses will be rated on a Likert scale ranging from 1 ("Not at all afraid") to 5 ("Very afraid"). Fear of weight gain will be treated as a continuous variable representing current degree of fear of weight gain at each survey.

Binge eating will be measured by first asking whether participants have eaten since the last survey. If so, participants will answer "While you were eating, did you feel a sense of loss of control (eg, like a car without brakes that could not stop)?" on a Likert scale ranging from 1 ("Not at all") to 5 ("Extremely"). Consistent with previous EMA studies, participants were counted as having endorsed loss-of-control eating if they reported a "3" or higher on either question at least once during the recording period [49]. Binge eating will be treated as a binary variable representing whether a binge eating episode occurred at each survey.

Dietary restraint will be measured by asking, "Since the last survey, to what extent have you attempted each of the following, even if you were not successful?" Answer choices included (1) "Tried to limit the amount you ate," (2) "Tried to avoid eating certain foods that you like," and (3) "Tried to delay eating." Responses will be rated on a Likert scale ranging from 1 ("Did not attempt") to 5 ("Attempted as hard as I could"). Dietary restraint will be treated as a binary variable representing a response of 3 ("Moderately attempted") or higher on any of the dietary restraint choices.

Dietary restriction will be measured by asking, "Since the last survey, were you successful in actually _____ in order to influence your shape and/or weight?" where the blank field will be replaced with (1) "limiting the amount that you ate," (2) "avoiding eating certain foods that you like," and (3) "trying to delay eating." Responses will be in a "Yes/No" format. Dietary restriction will be treated as a binary variable representing engagement in *any* form of dietary restriction at each survey.

Compensatory behaviors will be measured by asking, "Since the last survey, have you engaged in any of the following? Select all that apply." Options will include (1) "self-induced vomiting," (2) "taking laxatives to influence your shape and/or weight," (3) "taking diet pills to influence your shape and/or weight," (4) "taking diuretics to influence your shape and/or weight," (5) "exercise to influence your shape and/or weight," and (6) "other behavior to influence your shape and/or weight." Compensatory behaviors will be collapsed and treated as a binary variable representing engagement in *any* compensatory behaviors at each survey.

Compliance will be calculated as the number of completed signal-contingent EMA surveys out of 105 possible EMA surveys. To gage engagement, we will additionally report the total number of completed signal-contingent and event-contingent surveys.

Statistical Analysis

Sample Size Determination and Data Exclusions

Sample size was determined through power curve analyses in R version 4.2.3 [50] using the package *EMATools* [51], which indicated that to detect a medium-sized effect with 80% power using mixed-effects models, we would need to recruit a

XSL•FO
RenderX

minimum of 70 participants. To the extent that study resources (eg, budget, staffing) permit, more participants will be recruited to account for poor compliance and/or dropouts.

The data from the test day will be excluded from analyses. To promote high survey compliance, participants who achieve less than 50% compliance on the test day will receive a warning message. Participants who achieve less than 50% compliance on the test day and the first day of the study will be removed from the study.

Study Aims

Aim 1 of this study is to characterize body checking among adolescents with binge-spectrum EDs. To evaluate aim 1, descriptive statistics (ie, means and SDs) will be calculated for the frequency of each type of body checking behavior and the percentage of participants who endorsed each type of behavior over the EMA period, both in the full sample and by diagnostic category (BED-spectrum vs BN-spectrum), diagnostic threshold (full-threshold vs subthreshold), racial group, and gender identity. Additionally, the distribution of body checking frequency will be examined to capture when body checking behaviors most frequently occur throughout the day.

Aim 2 of the study is to assess whether frequency in body checking at 1 observation is associated with other ED symptoms at the same observation and at the next observation. To preserve analytic power, we plan to define body checking frequency as a composite binary variable representing the occurrence of body checking that aggregates counts across all types of body checking behaviors. However, it is possible that body checking (as a composite variable) may be reported so frequently that it fails to have any predictive ability or that certain body checking behaviors may have stronger prospective effects than others. Thus, we will conduct sensitivity analyses that examine each form of body checking separately as opposed to relying on a composite variable. We will use linear mixed-effects models for continuous outcomes (ie, ratings of body dissatisfaction and fear of weight gain) and generalized mixed-effects models using a binomial distribution with a logit link function for binary outcomes (ie, binge eating, compensatory behaviors, dietary restraint, and dietary restriction). All models will include body checking as a fixed predictor and a random intercept per participant. Consistent with other studies [52], models assessing binge eating and dietary restraint or restriction will control for binge eating and dietary restraint or restriction, respectively, at the previous observation. False discovery rate corrections will be used to correct for multiple comparisons.

Aim 3 is to assess whether body checking behaviors are reactive to EMA, such that they decrease over the course of the EMA period. To evaluate aim 3, we will first create a composite count score that aggregates frequencies of all body checking behaviors per person per day. Then, we will fit a generalized linear mixed model with a Poisson distribution to examine whether the count of body checking behaviors decreases as a function of time. We will include a fixed effect of time (ie, day of study) and a random intercept per subject. If the model yields poor fit (ie, there is over- or underdispersion), then a Quasi-Poisson will be used, and data will be examined for both zero-truncation and zero-inflation. To examine differences in body checking counts between the start and end of study participation, we will conduct a Poisson regression to examine associations between counts on day 1 to 3 and day 19 to 21. We elected to examine the first and last 3 days to obtain a more stable metric of frequency. Additionally, we will regress body checking behaviors across 2 time windows, day 1 to day 10 and day 11 to day 21. Given that the frequency of body checking may be directly associated with survey compliance, we will examine whether body checking frequency is correlated with the number of surveys completed. Further, in our Poisson regression model, we will control for the frequency of surveys completed to examine the rate of reduction in body checking behaviors in the first and second halves of the study period.

Results

Recruitment began in January 2025, and data collection is expected to conclude in March 2026. Data analyses and dissemination of results will be completed by August 2026.

Discussion

Study Implications and Future Directions

To date, there is limited literature on the role of body checking in maintaining binge-spectrum EDs among adolescents [12,14,23,53]. Thus, this study will provide valuable insights into the types, frequencies, and effects of body checking behaviors on ED symptoms among adolescents with binge-spectrum EDs. The findings from this study will clarify whether body checking behaviors are prospectively associated with engagement in other ED symptoms and may provide direction for research on mediators and moderators of this relationship, which is critical for better understanding maintaining mechanisms of EDs in adolescents. Furthermore, this study will evaluate whether body checking shows reductions when subject to self-monitoring during EMA. If our hypotheses are supported and body checking behaviors decrease along with other ED symptoms, this study could inform the importance of formally targeting body checking during treatments, such as CBT-E, and could inform clinicians' decisions around which symptoms to prioritize. Further, the findings from this study may provide a stepping stone for the development of scalable, stand-alone interventions to reduce body checking behaviors. Although such stand-alone interventions will not replace gold-standard approaches such as CBT-E, they represent a low-cost, accessible option for individuals with EDs that could be used as an adjunct to available treatments or a brief intervention while patients await the receipt of more comprehensive treatment. Digital self-monitoring could also have high utility as a tool for ED prevention, given literature which supports that higher engagement in body checking in adolescents is longitudinally linked to the development of ED symptoms [24]. Future research should aim to pilot such digital interventions to establish preliminary acceptability and feasibility.

Strengths and Limitations

The strengths of this study include the use of a transdiagnostic sample of adolescents with binge-spectrum EDs, which will

XSL•FO
RenderX

improve the generalizability of the results. Furthermore, the inclusion of a broad range of body checking behaviors, as opposed to assessing body checking with 1 or 2 items as past literature has done [20], will provide critical insights into the frequency and timing of these behaviors among adolescents.

The limitations of this study include the lack of follow-up assessments, which make it difficult to assess reactive effects to self-monitoring, whether body checking continues to decrease, plateaus, or eventually increases back to baseline. Likewise, given that adherence to self-monitoring may have a dose-response relationship with reactive reductions in body checking, it is possible that the incentives used to increase compliance (ie, financial compensation and personalized reminder texts with memes and GIFs) may confound aim 3. Given that incentives are needed to promote compliance, otherwise data missingness would preclude the detection of hypothesized relationships, future studies might consider varying sampling procedures (eg, fewer surveys per day) to more carefully test for a dose-response relationship between self-monitoring via EMA and body checking. Future work should examine these behaviors over longer periods and include follow-up assessments to better understand how effectively self-monitoring intervenes on body checking and how decreases in body checking may correlate with reductions in other indices of ED pathology. Of note, given the EMA design of this study, it is possible that reductions in other ED behaviors may be driven by self-monitoring or may arise from reductions in the overvaluation of shape and weight, as manifest by body checking. This study is further limited by its exclusion of boys. Given that boys use a different set of methods to body check [39], future work should use questionnaires, such as the Male Body Checking Questionnaire, to assess body checking in boys [38]. Likewise, given heterogeneity in body image ideals across cultures [54], it is possible measures used in this study will not appropriately capture body checking behaviors, and caution is warranted when examining racial or ethnic differences in body checking. Finally, although we used EMA items reported in other studies [16,17,52,55], including those with adolescents, some of these items have yet to undergo proper validation efforts, particularly for administration among adolescents.

Conclusion

In conclusion, this study will elucidate the nature and frequency of body checking, body checking's prospective relations to ED symptoms, and the potential reactivity of body checking to digital self-monitoring (via EMA) among adolescent girls with binge-spectrum EDs. The data afforded by this study will be critical for understanding the presentations of body checking in adolescents' daily lives and how it may maintain binge-spectrum EDs. This study will also inform optimal strategies for intervening on body checking and improving long-term ED treatment outcomes in adolescents with EDs.

Acknowledgments

Publication made possible in part by support from the Nemours Grants for Open Access from Library Services.

Funding

Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under award number F31MH138068 (principal investigator: LD). SM was supported by the National Institute of Diabetes, Digestive, and Kidney Diseases (K23DK124514). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The analytic data generated for the present analyses will be stripped of all identifiers and made openly available.

Conflicts of Interest

None declared.

References

1. Stice E, Marti CN, Rohde P. Prevalence, incidence, impairment, and course of the proposed DSM-5 eating disorder diagnoses in an 8-year prospective community study of young women. J Abnorm Psychol 2013 May;122(2):445-457. [doi: 10.1037/a0030679] [Medline: 23148784]
2. Hail L, Le Grange D. Bulimia nervosa in adolescents: prevalence and treatment challenges. Adolesc Health Med Ther 2018;9:11-16. [doi: 10.2147/AHMT.S135326] [Medline: 29379324]
3. Marzilli E, Cerniglia L, Cimino S. A narrative review of binge eating disorder in adolescence: prevalence, impact, and psychological treatment strategies. Adolesc Health Med Ther 2018;9:17-30. [doi: 10.2147/AHMT.S148050] [Medline: 29379325]
4. Hilbert A, Petroff D, Neuhaus P, Schmidt R. Cognitive-behavioral therapy for adolescents with an age-adapted diagnosis of binge-eating disorder: a randomized clinical trial. Psychother Psychosom 2020;89(1):51-53. [doi: 10.1159/000503116] [Medline: 31533113]

XSL•FO
RenderX

5. Peat CM, Berkman ND, Lohr KN, et al. Comparative effectiveness of treatments for binge-eating disorder: systematic review and network meta-analysis. Eur Eat Disord Rev 2017 Sep;25(5):317-328. [doi: 10.1002/erv.2517] [Medline: 28467032]
6. Linardon J, Wade TD. How many individuals achieve symptom abstinence following psychological treatments for bulimia nervosa? A meta-analytic review. Int J Eat Disord 2018 Apr;51(4):287-294. [doi: 10.1002/eat.22838] [Medline: 29417609]
7. Frederick DA, Forbes GB, Grigorian KE, Jarcho JM. The UCLA Body Project I: gender and ethnic differences in self-objectification and body satisfaction among 2,206 undergraduates. Sex Roles 2007 Aug 21;57(5-6):317-327. [doi: 10.1007/s11199-007-9251-z]
8. Tomba E, Tecuta L, Crocetti E, Squarcio F, Tomei G. Residual eating disorder symptoms and clinical features in remitted and recovered eating disorder patients: a systematic review with meta-analysis. Int J Eat Disord 2019 Jul;52(7):759-776. [doi: 10.1002/eat.23095] [Medline: 31169332]
9. Berkman ND, Lohr KN, Bulik CM. Outcomes of eating disorders: a systematic review of the literature. Intl J Eat Disord 2007 May;40(4):293-309. [doi: 10.1002/eat.20369]
10. Eshkevari E, Rieger E, Longo MR, Haggard P, Treasure J. Persistent body image disturbance following recovery from eating disorders. Int J Eat Disord 2014 May;47(4):400-409. [doi: 10.1002/eat.22219] [Medline: 24243423]
11. Sala M, Keshishian A, Song S, et al. Predictors of relapse in eating disorders: a meta-analysis. J Psychiatr Res 2023 Feb;158:281-299. [doi: 10.1016/j.jpsychires.2023.01.002] [Medline: 36623362]
12. Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: a “transdiagnostic” theory and treatment. Behav Res Ther 2003 May;41(5):509-528. [doi: 10.1016/s0005-7967(02)00088-8] [Medline: 12711261]
13. Shafran R, Fairburn CG, Robinson P, Lask B. Body checking and its avoidance in eating disorders. Int J Eat Disord 2004 Jan;35(1):93-101. [doi: 10.1002/eat.10228] [Medline: 14705162]
14. Walker DC, Murray AD. Body image behaviors: checking, fixing, and avoiding. In: Encyclopedia of Body Image and Human Appearance: Elsevier Academic Press; 2012, Vol. 1:166-172. [doi: 10.1016/B978-0-12-384925-0.00025-0]
15. Reas DL, Grilo CM, Masheb RM, Wilson GT. Body checking and avoidance in overweight patients with binge eating disorder. Int J Eat Disord 2005 May;37(4):342-346. [doi: 10.1002/eat.20092] [Medline: 15856496]
16. Stefano EC, Hudson DL, Whisenhunt BL, Buchanan EM, Latner JD. Examination of body checking, body image dissatisfaction, and negative affect using Ecological momentary assessment. Eat Behav 2016 Aug;22:51-54. [doi: 10.1016/j.eatbeh.2016.03.026] [Medline: 27086048]
17. Smith JM, Whisenhunt BL, Buchanan EM, Hudson DL. Evaluating the effectiveness of ecological momentary assessment and intervention targeting body checking behaviors. Eat Disord 2019;27(6):521-537. [doi: 10.1080/10640266.2018.1560850] [Medline: 30664400]
18. Mountford V, Haase A, Waller G. Body checking in the eating disorders: associations between cognitions and behaviors. Int J Eat Disord 2006 Dec;39(8):708-715. [doi: 10.1002/eat.20279] [Medline: 16868998]
19. Goeden AV, Schaefer LM, Crosby RD, et al. Examining the momentary relationships between body checking and eating disorder symptoms in women with anorexia nervosa. Eat Behav 2023 Aug;50:101751. [doi: 10.1016/j.eatbeh.2023.101751] [Medline: 37244020]
20. Lavender JM, Wonderlich SA, Crosby RD, et al. A naturalistic examination of body checking and dietary restriction in women with anorexia nervosa. Behav Res Ther 2013 Aug;51(8):507-511. [doi: 10.1016/j.brat.2013.05.004] [Medline: 23792180]
21. Kraus N, Lindenberg J, Zeeck A, Kosfelder J, Vocks S. Immediate effects of body checking behaviour on negative and positive emotions in women with eating disorders: an ecological momentary assessment approach. Eur Eat Disord Rev 2015 Sep;23(5):399-407. [doi: 10.1002/erv.2380] [Medline: 26095227]
22. Wilhelm L, Hartmann AS, Cordes M, Waldorf M, Vocks S. How do you feel when you check your body? Emotional states during a body-checking episode in normal-weight females. Eat Weight Disord 2020 Apr;25(2):309-319. [doi: 10.1007/s40519-018-0589-8] [Medline: 30288724]
23. Dakanalis A, Carrà G, Timko A, et al. Mechanisms of influence of body checking on binge eating. Int J Clin Health Psychol 2015;15(2):93-104. [doi: 10.1016/j.ijchp.2015.03.003] [Medline: 30487826]
24. Zaitsoff SL, Pullmer R, Coelho JS. A longitudinal examination of body-checking behaviors and eating disorder pathology in a community sample of adolescent males and females. Int J Eat Disord 2020 Nov;53(11):1836-1843. [doi: 10.1002/eat.23364] [Medline: 32815197]
25. Skemp-Arlt KM. Body image dissatisfaction and eating disturbances among children and adolescents: prevalence, risk factors, and prevention strategies. J Phys Educ Recreat Dance 2006 Jan;77(1):45-51. [doi: 10.1080/07303084.2006.10597813]
26. Benowitz-Fredericks CA, Garcia K, Massey M, Vasagar B, Borzekowski DLG. Body image, eating disorders, and the relationship to adolescent media use. Pediatr Clin North Am 2012 Jun;59(3):693-704. [doi: 10.1016/j.pcl.2012.03.017] [Medline: 22643174]
27. Levine MP, Piran N. The role of body image in the prevention of eating disorders. Body Image 2004 Jan;1(1):57-70. [doi: 10.1016/S1740-1445(03)00006-8] [Medline: 18089141]
28. Hogan MJ, Strasburger VC. Body image, eating disorders, and the media. Adolesc Med State Art Rev 2008 Dec;19(3):521-546. [Medline: 19227390]

XSL•FO
RenderX

29. Lin BY, Moog D, Xie H, et al. Increasing prevalence of eating disorders in female adolescents compared with children and young adults: an analysis of real-time administrative data. Gen Psych 2024 Aug;37(4):e101584. [doi: 10.1136/gpsych-2024-101584]
30. Cohen JS, Edmunds JM, Brodman DM, Benjamin CL, Kendall PC. Using self-monitoring: implementation of collaborative empiricism in cognitive-behavioral therapy. Cogn Behav Pract 2013 Nov;20(4):419-428. [doi: 10.1016/j.cbpra.2012.06.002]
31. Cash TF, Hrabosky JI. The effects of psychoeducation and self-monitoring in a cognitive-behavioral program for body-image improvement. Eat Disord 2003;11(4):255-270. [doi: 10.1080/10640260390218657] [Medline: 16864292]
32. Frederiksen LW. Treatment of ruminative thinking by self-monitoring. J Behav Ther Exp Psychiatry 1975 Oct;6(3):258-259. [doi: 10.1016/0005-7916(75)90115-9]
33. Clum GA, Curtin L. Validity and reactivity of a system of self-monitoring suicide ideation. J Psychopathol Behav Assess 1993 Dec;15(4):375-385. [doi: 10.1007/BF00965039]
34. Twomey C, O'Reilly G. The association of "self-monitoring" with depression in help-seeking Facebook users. Int J Soc Psychiatry 2018 Aug;64(5):482-487. [doi: 10.1177/0020764018778700] [Medline: 29852812]
35. Compernolle S, DeSmet A, Poppe L, et al. Effectiveness of interventions using self-monitoring to reduce sedentary behavior in adults: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2019 Aug 13;16(1):63. [doi: 10.1186/s12966-019-0824-3] [Medline: 31409357]
36. Gass JC, Funderburk JS, Shepardson R, Kosiba JD, Rodriguez L, Maisto SA. The use and impact of self-monitoring on substance use outcomes: a descriptive systematic review. Subst Abuse 2021 Oct;42(4):512-526. [doi: 10.1080/08897077.2021.1874595]
37. Hart LM, Granillo MT, Jorm AF, Paxton SJ. Unmet need for treatment in the eating disorders: a systematic review of eating disorder specific treatment seeking among community cases. Clin Psychol Rev 2011 Jul;31(5):727-735. [doi: 10.1016/j.cpr.2011.03.004] [Medline: 21501580]
38. Hildebrandt T, Walker DC, Alfano L, Delinsky S, Bannon K. Development and validation of a male specific body checking questionnaire. Int J Eat Disord 2010 Jan;43(1):77-87. [doi: 10.1002/eat.20669] [Medline: 19247988]
39. Catherine Walker D, Anderson DA, Hildebrandt T. Body checking behaviors in men. Body Image 2009 Jun;6(3):164-170. [doi: 10.1016/j.bodyim.2009.05.001]
40. Keel PK, Mayer SA, Harnden-Fischer JH. Importance of size in defining binge eating episodes in bulimia nervosa. Int J Eat Disord 2001 Apr;29(3):294-301. [doi: 10.1002/eat.1021] [Medline: 11262508]
41. Brownstone LM, Bardone-Cone AM, Fitzsimmons-Craft EE, et al. Subjective and objective binge eating in relation to eating disorder symptomatology, negative affect, and personality dimensions. Int J Eat Disord 2013 Jan;46(1):66-76. [doi: 10.1002/eat.22066] [Medline: 23109272]
42. Cooper Z, Fairburn C. The eating disorder examination: a semi-structured interview for the assessment of the specific psychopathology of eating disorders. Int J Eat Disord 1987 Jan;6(1):1-8. [doi: 10.1002/1098-108X(198701)6:1<1::AID-EAT2260060102>3.0.CO;2-9]
43. Butler RM, Lampe E, Trainor C, Manasse SM. Fear of weight gain during cognitive behavioral therapy for binge-spectrum eating disorders. Eat Weight Disord 2023 Mar 6;28(1):29. [doi: 10.1007/s40519-023-01541-8] [Medline: 36879078]
44. Cooper PJ, Taylor MJ, Cooper Z, Fairbum CG. The development and validation of the body shape questionnaire. Int J Eat Disord 1987 Jul;6(4):485-494. [doi: 10.1002/1098-108X(198707)6:4<485::AID-EAT2260060405>3.0.CO;2-O]
45. Reas DL, Whisenhunt BL, Netemeyer R, Williamson DA. Development of the body checking questionnaire: a self-report measure of body checking behaviors. Int J Eat Disord 2002 Apr;31(3):324-333. [doi: 10.1002/eat.10012] [Medline: 11920995]
46. Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire? Intl J Eating Disorders 1994 Dec;16(4):363-370. [doi: 10.1002/1098-108X(199412)16:4<363::AID-EAT2260160405>3.0.CO;2-%23]
47. EMA data collection: improving compliance rates and giving back to participants. Medium. 2020. URL: https://medium.com/@psychology_schleiderlab/ema-data-collection-improving-compliance-rates-and-giving-back-to-participants-996c00a9b70b [accessed 2024-07-13]
48. Kleiman EM, Glenn CR, Auad EC, et al. A simple way to gamify ecological momentary assessment studies and improve survey adherence with adolescents: the Emoji Game. Psychol Assess 2025;37(6-7):288-296. [doi: 10.1037/pas0001371] [Medline: 40193442]
49. Berg KC, Crosby RD, Cao L, et al. Negative affect prior to and following overeating-only, loss of control eating-only, and binge eating episodes in obese adults. Intl J Eating Disorders 2015 Sep;48(6):641-653. [doi: 10.1002/eat.22401]
50. The R Project for Statistical Computing. 2022. URL: https://www.R-project.org/ [accessed 2026-02-23]
51. Kleiman E. EMAtools: data management tools for real-time monitoring/ecological momentary assessment data. rdrr.io. 2021. URL: https://rdrr.io/cran/EMAtools/ [accessed 2026-03-12]
52. Manasse SM, Lampe EW, Abber SR, Fitzpatrick B, Srivastava P, Juarascio AS. Differentiating types of dietary restraint and their momentary relations with loss-of-control eating. Intl J Eating Disorders 2023 May;56(5):969-977. [doi: 10.1002/eat.23896]
53. Shafran R, Lee M, Payne E, Fairburn CG. An experimental analysis of body checking. Behav Res Ther 2007 Jan;45(1):113-121. [doi: 10.1016/j.brat.2006.01.015] [Medline: 16542638]

XSL•FO
RenderX

54. Abdoli M, Scotto Rosato M, Desousa A, Cotrufo P. Cultural differences in body image: a systematic review. Soc Sci 2024 Jun 5;13(6):305. [doi: 10.3390/socsci13060305]
55. Goldschmidt AB, Smith KE, Crosby RD, et al. Ecological momentary assessment of maladaptive eating in children and adolescents with overweight or obesity. Int J Eat Disord 2018 Jun;51(6):549-557. [doi: 10.1002/eat.22864] [Medline: 29626353]

Abbreviations

BED: binge eating disorder
BN: bulimia nervosa
CBT: cognitive behavioral therapy
CBT-E: enhanced cognitive behavior therapy
ED: eating disorder
EMA: ecological momentary assessment

Edited by J Sarvestan; submitted 04.Mar.2025; peer-reviewed by A Egbert, C Udensi, T Yegateela; revised version received 30.Sep.2025; accepted 01.Oct.2025; published 15.Apr.2026.

Please cite as:
Moussaoui J, Manasse S, D'Adamo L
Temporal Associations Between Body Checking and Eating Pathology in Adolescent Girls with Binge-Spectrum Eating Disorders: Protocol for an Ecological Momentary Assessment Study
JMIR Res Protoc 2026;15:e73447
URL: https://www.researchprotocols.org/2026/1/e73447
doi:10.2196/73447

© Jannah Moussaoui, Stephanie Manasse, Laura D'Adamo. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Wellness4Students Program to Mitigate Stress, Anxiety, and Depression Among Makerere University Students in Uganda: Protocol for a Longitudinal Study

Stephen Ojiambo Wandera[1], PhD; Olivia Nankinga[1], PhD; Patricia Ndugga[1], PhD; Christabellah Namugenyi[2], MS; Claire Ashaba[2], MS; Noeline Nakasujja[3], Prof Dr; Rosco Kasujja[4], Prof Dr; Belinda Agyapong[5], PhD; Vincent Israel Opoku Agyapong[6], Prof Dr

[1]Department of Population Studies, School of Statistics and Planning, College of Business and Management Sciences, Makerere University, Makerere Hill Road, Kampala, Uganda

[2]Department of Planning and Applied Statistics, School of Statistics and Planning, College of Business and Management Sciences, Makerere University, Kampala, Uganda

[3]Department of Psychiatry, School of Medicine, College of Health Sciences, Makerere University, Kampala, Uganda

[4]Department of Mental Health & Community Psychology, School of Psychology, College of Humanities and Social Sciences, Makerere University, Kampala, Uganda

[5]Department of Psychiatry, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, AB, Canada

[6]Department of Psychiatry, Faculty of Medicine, Dalhousie University, Halifax, NS, Canada

Corresponding Author:
Stephen Ojiambo Wandera, PhD
Department of Population Studies, School of Statistics and Planning, College of Business and Management Sciences, Makerere University, Makerere Hill Road, Kampala, Uganda

Abstract

Background: Research indicates that university students experience higher levels of stress, anxiety, and depression (SAD) than the general population. In Uganda, existing psychological interventions for addressing SAD among students are primarily delivered face to face, which limits effective diagnosis and treatment due to stigma, a shortage of counselors, and long waiting times, which contribute to significant unmet mental health needs. Consequently, there is an urgent need for innovative approaches to improve access to mental health services.

Objective: This study aims to assess the feasibility and effectiveness of a supportive SMS text messaging program (Wellness4Students [W4S]) in reducing the prevalence of SAD among university students in Uganda.

Methods: This study has a quantitative longitudinal research design. We will target 4494 undergraduate students aged 18 to 24 years who will be recruited from Makerere University, Uganda. Through advertisements on social media platforms (WhatsApp and X [formerly known as Twitter]) and websites, students will receive a link to the W4S website with information about the intervention. Eligible students will voluntarily self-subscribe to the W4S program by providing their email addresses. The subscription will be an indicator of consent required for participation in receiving the intervention. Subscribers will receive daily supportive messages for 6 months. The data will be collected from subscribers at baseline (program initiation), 1 month, 3 months (midpoint), and 6 months (program completion) through web-based surveys. Students will complete questionnaires regarding SAD and well-being. We will conduct both descriptive and inferential statistical analyses using STATA version 17.

Results: The results of this study are expected within 1 year after program initiation. The data will be collected from February 2026 through October 2026. The development of the W4S app and the survey instrument has been completed. The cocreation of mental health messages with 15 undergraduate students and 2 university psychologists has been completed as of March 23, 2026. The roll-out of the program is scheduled for March 24, 2026.

Conclusions: The W4S intervention is feasible. The W4S program will provide essential insight into the prevalence and correlates of SAD among Makerere University students. We will identify students who are most at risk of poor mental health outcomes.

International Registered Report Identifier (IRRID): PRR1-10.2196/84637

(JMIR Res Protoc 2026;15:e84637) doi:10.2196/84637

KEYWORDS

stress; anxiety; depression; students; Wellness4Students; mental health; Uganda; Africa; email; WhatsApp

XSL•FO
RenderX

Introduction

Background

Mental health is a critical component of overall well-being. Mental health disorders, such as stress, anxiety, and depression (SAD), have been linked to serious health conditions [1], including cardiovascular diseases, cancer, diabetes, respiratory diseases, substance abuse, chronic pain, sleep disorders, and heart diseases [2]. Globally, mental health disorders are among the most burdensome health problems, accounting for more than 970 million cases [3,4]. According to the United Nations Children's Fund, more than 13% of young people are affected worldwide [5]. Anxiety and depression alone comprise about 40% of these cases [5]. While these statistics reflect a worldwide crisis, the burden is disproportionately higher in low- and middle-income countries, where mental health infrastructure is underresourced and access to care is critically limited [2,6,7].

In sub-Saharan Africa, this disparity is glaring. The continent has fewer than 1.5 mental health workers per 100,000 population, compared to a global median of 9, and governments allocate less than 1% of health budgets to mental health [2,4,8]. This scarcity is linked to the treatment gap, estimated at over 75% for common mental disorders in sub-Saharan Africa, leaving the vast majority without any form of care [9]. This systemic neglect exacerbates vulnerabilities among populations already facing socioeconomic challenges, including university students.

Evidence consistently indicates a higher prevalence of SAD among university students compared to the general population [10-12]. Multiple meta-analyses indicate that university or postsecondary students have far higher pooled rates of depression and anxiety than the general or same-age population. For instance, a large pre-COVID meta-analysis estimated depression prevalence among university students at 30.6%, far above the global adult prevalence of 12.9% [13]. A more recent global review (64 studies; 100,187 students) found a pooled prevalence of depression at 33.6% and anxiety at 39% among college students [14]. An umbrella review of anxiety in college or university students across all world regions reported a median anxiety prevalence of 32% (range 7.4% - 55%) [15]. A systematic review covering 40 countries found a prevalence of 26.1% for depression and 24.5% for anxiety among university students and explicitly contrasted this with much lower World Health Organization estimates for the general population (depression: 4.4%, anxiety: 3.6%), concluding that disorders "affect a much higher percentage of university population than the general one." The transition into higher education presents a key developmental period marked by academic pressure, social reorientation, and often financial strain, all of which can significantly impact mental health [16]. These problems not only cause personal distress but also negatively affect students' quality of life and academic performance [17].

Research across East Africa highlights the scale of the issue, though prevalence estimates vary due to methodological heterogeneity in assessment tools and timing. In Tanzania, a study reported a prevalence of depression of 22% among undergraduates [17], while in Kenya, the prevalence reached 41% [18]. In Uganda, a systematic review found that depression among university students ranged from 0.4% to 80% [16], a range underscoring the need for standardized measurements. A study among medical students at Makerere University found a prevalence of 22% [19]. The COVID-19 pandemic remarkably exacerbated this situation; during the lockdowns, one study among Ugandan students reported alarmingly high rates of depression (81%), anxiety (98%), and stress (78%) [20]. This dramatic increase aligns with global findings of a "mental health pandemic" secondary to COVID-19, particularly affecting youth in educational settings [16].

The determinants of SAD among students are multifaceted, arising from an interplay of sociodemographic, academic, and behavioral factors. Key sociodemographic factors include gender, age, and low socioeconomic status [20-22]. Academic and institutional stressors include demanding programs of study, academic pressure, and the transition to university life [23,24]. Behavioral factors include substance use, excessive social media engagement, and relationship problems [20-22]. The "campus culture" and institutional climate, including perceptions of support (or lack thereof) from faculty and administration, have also been identified as significant moderators of student mental health [20-22].

Despite this significant and multifaceted burden, the treatment gap in Uganda remains vast due to systemic, context-specific barriers. The deeply ingrained stigma, often rooted in cultural taboos and spiritual beliefs that attribute mental illness to moral failure or supernatural causes, deters help-seeking [25,26].

Qualitative studies in Uganda specifically cite fear of gossip and being labeled "kizibu" (mad) as primary reasons students avoid formal services. This is mainly linked to low mental health literacy and a crippling shortage of services: Uganda has fewer than 2 psychiatrists, with counselor-to-student ratios often exceeding 1:10,000 [2,4,8,19,27]. Structural and policy gaps limit care, as mental health remains chronically underfunded within national and institutional budgets [25,28]. Yet, these countries experience a significant burden of mental health disorders, with over 24% of individuals qualifying for psychiatric diagnosis [10,16,21,23,27,29-31].

Consequently, current psychological interventions in Uganda are primarily limited to unsustainable person-to-person approaches, hindered by stigma, cultural taboos, long wait times, policy gaps, and insufficient resource allocation [16,19,20,23,27,32]. Available alternative interventions are often Western-based and not tailored to the unique sociocultural and economic context of Ugandan students [33-36]. This lack of cultural adaptation can reduce intervention engagement, relevance, and effectiveness, a known challenge in global mental health transitions [37,38]. There is, therefore, a critical need to develop, adapt, and test scalable, low-cost, and contextually relevant interventions that can bypass traditional barriers to care.

Digital mental health interventions, delivered through widely accessible platforms like email and WhatsApp, offer a promising pathway. Systematic reviews suggest that digital interventions, particularly those based on cognitive behavioral therapy (CBT) principles delivered via text or app, are effective in reducing

XSL•FO
RenderX

symptoms of depression and anxiety in young adults [39,40]. They can provide stigma-free and consistent support, directly addressing gaps in accessibility and human resources. Their scalability makes them particularly suitable for low-resource settings. Preliminary studies in other African contexts, such as SMS-based support for perinatal depression in Kenya, demonstrate feasibility and high acceptability [41,42]. This study responds to this urgent need by assessing the feasibility and effectiveness of a supportive SMS text messaging program (Wellness4Students [W4S]) in reducing the prevalence of SAD among university students in Uganda.

W4S Email and WhatsApp Messaging Intervention

The W4S program is a daily, supportive email and WhatsApp messaging intervention designed to mitigate symptoms of SAD and promote well-being. It is part of the broader ResilienceNHope suite, supported by a web-based platform and implemented through the Global Psychological eHealth Foundation. The initial content is developed by mental health professionals using CBT principles. Crucially, for this study, the program will be contextually adapted. The initial message database contains generic content that will serve only as a starting point for adaptation. Through participatory cocreation workshops with students at Makerere University, this content will be linguistically, culturally, and contextually adapted to ensure that metaphors, proverbs, and messages address specific contextual stressors for the students. This participatory process is central to the intervention's design, ensuring it is not merely imported but is collaboratively shaped by its intended users [25,28,33,43].

Through a partnership with Makerere University, students will be able to self-enroll online. Upon registration, they will receive daily supportive, noninteractive (one-way) messages for 6 months, with the nature of the service clearly explained at the outset. Each day's email and WhatsApp message will be unique, drawn from a large message database.

Examples of messages (not yet adapted) to be delivered to subscribers of the W4S email and WhatsApp messaging program include the following:

- What you do today will determine how you are tomorrow. Rise up and take advantage of whatever opportunities today presents, and success will be yours tomorrow.
- There are 2 days in the week we should not worry about, yesterday and tomorrow. That leaves today, live for today.
- There are no surprises in store. You are in charge of the successes that await you. There may be some challenges, but you have the power to overcome them.

Aims of the Study

This study aims to determine the prevalence and correlates of SAD among university students in Uganda. It also aims to determine if daily supportive text messages can help reduce the prevalence of SAD and improve well-being among university students.

Specific Objectives

This study has the following objectives: (1) to determine the prevalence and correlates of SAD among university students in Uganda, (2) to investigate the effectiveness of the daily W4S supportive messages in reducing levels of SAD and in improving well-being among university students in Uganda, and (3) to investigate the participants' experience and satisfaction associated with the W4S daily SMS text messaging support.

Methods

Study Design

The study used a quantitative longitudinal study design. It will be implemented in three phases: (1) the cocreation and adaptation of the intervention; (2) the baseline survey; and (3) follow-up surveys conducted at 1 month, mid-term follow-up at 3 months, and an endline at 6 months. The cocreation phase was completed during the recess term (July 2025). The next 2 phases, baseline and follow-ups, will be implemented between February 2026 and October 2026. Recruitment will begin in February 2026 and will last for 3 months up to April 2026. There will be an observation period of 6 months for the enrolled participants [28,33,43].

The email and WhatsApp messaging intervention was cocreated and adapted by the study team together with purposively selected students and identified stakeholders. Thereafter, a baseline survey will be conducted before recruitment to estimate the prevalence of SAD. Students at Makerere University will be invited via email and WhatsApp to self-subscribe to the W4S program. All participants who subscribe to the intervention will receive a link to the online baseline survey, daily supportive messages for 6 months, and the follow-up survey links at 1-, 3-, and 6-month postenrollment time points [33,43].

The rationale for the nonrandomized design is based on resource constraints. We have funding to implement a pilot survey and use the preliminary results to apply for a bigger grant to implement a fully randomized controlled trial.

Study Setting

The study will be conducted among Makerere University undergraduate students. The university has a population of over 35,000 students [44]. Since the baseline survey and the follow-up will be conducted virtually, general assumptions are made that all students own smartphones, have an email and WhatsApp, and can access free Wi-Fi at Makerere University.

The Intervention: Cocreation of Email and WhatsApp Messages

For this protocol, the intervention to be delivered to the students at Makerere University is the ResilienceNHope program. The ResilienceNHope suite of interventions delivers daily supportive messages, often accompanied by images and links to mental health literacy resources, to its subscribers. A total of 15 students (8 female and 7 male) were selected across years of study (year 1, year 2, and year 3) at the College of Business and Management Sciences, Makerere University, to ensure representativeness of the cocreation team and to avoid potential biases in the adaptation process.

Supportive text messaging programs based on CBT principles, such as the W4S program within the ResilienceNHope suite, are among the most highly recommended digital interventions

XSL•FO
RenderX

for SAD and have been successfully implemented in both Canada and Zambia [33-35,43]. The study team, together with selected undergraduate students and identified stakeholders, will adapt the W4S program, specifically tailoring messages and images to the Ugandan undergraduate university students' population [43].

There are several adaptation frameworks. These include the ADAPT-ITT (Assessment, Decision, Adaptation, Production, Topical experts, Integration, Training, Testing), intervention mapping (IM) framework, and framework for reporting adaptations and modifications to evidence (FRAME), among others [38,45,46]. For this study, we will adapt the W4S program [33,43] using the IM adaptation framework [47]. The IM adaptation framework has 6 steps. First, establish a detailed understanding of the health problem, the population at risk, the behavioral and environmental causes, and the determinants of these behavioral and environmental conditions and assess available resources. Second, describe the behavioral and environmental outcomes, create objectives for changes in the determinants of behavior and environmental causes, and specify the targets of the intervention program. Third, identify theory- and evidence-based behavior change methods that influence the determinants and translate these to practical applications that fit the intervention context. Fourth, combine the intervention components into a coherent program that uses delivery channels that fit the context. Fifth, develop implementation strategies to facilitate the adoption, implementation, and maintenance of the program. Finally, plan both process and outcome evaluation to assess program implementation, efficacy, or effectiveness [47].

The adaptation is expected to improve the feasibility and acceptability of the W4S intervention [33-35,43]. The cocreation and adaptation will be done when students are in the recess term period (between June and July).

Data Collection

An online baseline survey will be conducted to assess the mental health of the students.

A link to the baseline survey will be sent to the students through their email and WhatsApp. Clicking on the link to the baseline survey will indicate consent required for participation in the survey. The data collection portal is secure and password-protected with dual authentication, which minimizes data breaches. In addition to the demographic questions, participants will fill out the online screening tools for assessing SAD and well-being. It is on this basis that they will be diagnosed (to have or not to have) with at least one of the 3 outcomes [8,48-50]. The app is programmed to produce summaries of the total number of subscribers at any given point in time and how many have completed the different phases.

All participants will be able to view their scores for each mental health outcome (SAD). Depending on the severity of the mental health scores, recommendations will be provided to participants. All students, whether diagnosed with SAD or not, will have the option to enroll in the W4S program and will be made available through a postbaseline subscription link. The baseline survey and intervention enrollment will be conducted from March 2026 to cover the second semester of the academic year 2025/2026 at Makerere University.

Baseline Survey

A structured baseline survey questionnaire will be developed to measure students' mental health outcomes. The questionnaire will include the 10-item Perceived Stress Scale (PSS-10), 7-item Generalized Anxiety Disorder (GAD-7) scale, 9-item Patient Health Questionnaire-9 (PHQ-9) scale, and WHO-5 (World Health Organization-Five Well-Being Index) scale to measure SAD and well-being [20,49,51-53]. These tools have been validated for making comparisons between population subgroups [35]. The questionnaire will also include sociodemographic and economic variables, including age, relationship status, religion, orphanhood status [54], college, sponsorship status, working status, living arrangement, year of study, and current academic performance among others. Participant data will be collected online through self-administered questionnaires on the ResilienceNHope platform [43]. The data will be hosted online on the ResilienceNHope server and downloaded for analysis.

Follow-Up Surveys

A follow-up survey questionnaire will be used to assess improvements in the students' mental well-being using validated scales for SAD, using the PSS-10, GAD-7, and PHQ-9. These surveys will be conducted at 1 month, 3 months (mid-term evaluation), and 6 months of the study (endline evaluation). The app has automated email reminders at 6 weeks, 3 months, and 6 months to boost completion rates. In addition, receiving daily messages with images is an incentive to students, and this was tested during the cocreation process with students.

The questionnaire will also include variables on participants' satisfaction with the W4S program intervention and suggested improvements and recommendations. Participant data will be collected and hosted online through self-administered questionnaires on the ResilienceNHope web platform [43] and downloaded for analysis. We recognize that retention is and has been a challenge for online surveys; part of the mitigation strategy is to make the survey shorter. The number of questions was reduced to the bare minimum.

Participants' Recruitment and Subscription

This study will be conducted among Makerere University students in Kampala, Uganda. We will consider currently enrolled students from year 1 to year 3 of study. We have made an assumption that all university students own smartphones based on a previous study [55]. Internet access is always guaranteed to all university students during the semester when the study will be implemented. We will broadcast information about the study through social media platforms used by university students, including WhatsApp, emails, and Twitter (subsequently rebranded X). We will also share information about the study through posters and by word-of-mouth within accessible groups. We will also engage the staff of the Counselling and Guidance Center at the University Hospital to inform students about this study.

During the broadcast, students will be informed about the possibility of being enrolled in a digital intervention program

XSL•FO
RenderX

(W4S program). The study will be launched through an announcement by the office of the Dean of Students of Makerere University, the Students' Guild, and the Director of Makerere University Hospital. All interested students will be provided with a link to the W4S website to self-subscribe to the study. Potential respondents will be asked to register for the study with their active emails and also consent to being part of the study. Once they have electronically consented to the study, they will receive another link to the baseline study. The cocreated and adapted messages will be sent to subscribers' email and WhatsApp daily at 7 AM East African Time. Each subscriber will receive the W4S program daily for 6 months.

Midterm and Endline Evaluation

The program will be evaluated at 1 month, 3 months (mid-term), and 6 months (endline evaluation) after the students start receiving daily supportive email and WhatsApp messages. The midterm and endline surveys will assess improvements in the students' mental well-being using validated scales for the primary outcomes: SAD [20,48-51,53].

Outcomes and Measures

Primary outcome measures at baseline, 1 month, 3 months, and 6 months will be the mean difference in scores on the PSS-10, GAD-7 scale, and PHQ-9 [49,56-58]. Prevalence estimates for SAD will be calculated as the total number of undergraduate students assessed as likely to have a particular condition divided by the total number of students who completed the baseline survey. In addition, proportions for survey uptake, response rates, and completion rates will be calculated. Finally, the outcomes related to the feasibility or acceptance of the tool, along with measures of engagement and adherence, will be collected at the end of the follow-up survey.

Sample Size Estimation

The total student population at Makerere University, Kampala, Uganda, is about 35,000 [41]. It is estimated that the sample size needed for our prevalence estimates with an expected SAD prevalence of 50% [10,12], a 95% interval, and a 2% margin of error for moderate-to-high stress, moderate-to-high anxiety, and moderate-to-high depression among Makerere University students will be 2247. A sample size calculator was used to determine the sample size [28,59,60].

Based on the previous maximum survey completion rates of 50% observed in similar programs, a minimum of 4494 students will be enrolled in the W4S email and WhatsApp messaging program over 6 months to obtain an estimated 2247 students' completion at baseline. To achieve a power of 80% (β=.2) and a 2-sided significance level of 5% (α<.05) and to detect an effect size of 30% before and after the intervention, a sample size of 90 would be needed [49].

Statistical Analysis

Once data are downloaded from the W4S server, the data from the baseline and follow-up surveys will be exported to STATA version 17 for data analysis. Observations with more than 50% missing responses will be omitted from the dataset. Otherwise, missing data on individual variables will be handled using multiple imputation with baseline and postbaseline variables [61].

At baseline, frequency distributions will be used to summarize categorical variables, while means and SDs will be used for continuous variables. Chi-square tests and/or *F* statistics will be used to measure the association between demographic characteristics and SAD variables. Multivariable logistic regressions will be used to determine the correlates of SAD among MAK undergraduate students. We plan to conduct rural-urban disaggregated analysis to account for diversity among the students.

Paired sample 2-tailed *t* tests and proportion tests will be used to assess significant differences in mean scores and prevalence rates of SAD at baseline, 1 month, 3 months, and 6 months. Profile plots of mean scores will be generated while controlling for background characteristics. All findings will be reported at 95% CI with a significance level of 5% (P<.05). The data will be analyzed using Stata version 17.

Data Quality Assurance

At each stage of data collection, we will implement checks to ensure that we collect high-quality data, enabling us to generate accurate insights. A 3-member data quality committee will be established, which will be responsible for checking quantitative data that are uploaded to the server. All data will be checked for consistency and accuracy. Any errors and inconsistencies will be immediately scrutinized for corrective action. Also, we will use algorithmic checks, such as Mahalanobis distance for multivariate outliers, integrated into STATA scripts to ensure data integrity and reproducibility [62].

Ethical Considerations

Ethical approval for this study was granted by the AIDS Support Organisation Research Ethics Committee (TASO-2025 - 2141) and the Uganda National Council of Science and Technology (UNCST SS4200ES).

Participation in the study will be voluntary, where students will self-subscribe to receive daily supportive emails and WhatsApp messages for 6 months. All respondents will be assured of confidentiality, and participants will be made aware that informed consent will be implied when they complete and submit the survey. Privacy and confidentiality will be guaranteed. The data will be anonymized and deidentified, and protective measures will be taken to safeguard participant information. All students will be compensated with a one-off airtime compensation of CAD $5 (equivalent to US $3) after completing the surveys.

Potential participants will be informed of their right to withdraw from the study. The online questionnaire will automatically end upon their declining to participate in the study. The dataset from the website is programmed to be downloaded without identification particulars (deidentified), and therefore, no personal information will be publicized. The particulars of the respondents are only available to the principal investigator and the core research team.

XSL•FO
RenderX

Students with adverse mental health outcomes will be referred to the university's Counselling and Guidance Center or advised to seek professional help.

Hypotheses

We hypothesize that the W4S email and WhatsApp messaging program will reduce the prevalence and severity of SAD among Makerere University students. We also hypothesize that the prevalence of moderate-to-high stress, moderate-to-high anxiety, and moderate-to-high depression would be comparable to the prevalence of these conditions reported among university students in East Africa (Kenya and Tanzania) and in other jurisdictions. Finally, we hypothesize that demographic factors, socioeconomic factors, organizational factors, and class size will be associated with SAD among students.

Results

The results will be reported following the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist [63]. The W4S email and WhatsApp messaging program was anticipated to be launched in September 2025. There were some delays, and the enrollment will start by March 24, 2026. Enrollment is expected to continue for 6 months up to September 2026. Data collection will continue for another 6 months. The results of the study will be disseminated to stakeholders in the education sector in Uganda, East African countries, and globally through workshops, conference presentations, and peer-reviewed publications.

The progress to date is that ethical approval has been secured. We conducted a cocreation of intervention messages. A pilot survey was conducted with about 20 students. We are going to implement the enrollment and the intervention during the second semester of the academic calendar (February-May 2026). We will implement a survey from March 2026 at different time intervals.

Discussion

Principal Findings

SAD may indirectly hinder university students' academic performance, overall lifestyle, psychological safety, and well-being, as well as lead to other physical health problems. Given their demanding schedules, students need innovative, anonymized, accessible, convenient, and cost-effective programs to support their mental health. This protocol proposes the use of supportive email and WhatsApp messages as a psychological intervention for students experiencing SAD. Those with any of the target outcomes, prior traumatic experiences, or adverse childhood events are expected to benefit particularly from this program. We anticipate a reduction in mental health outcomes, such as reduced SAD among university students at Makerere University.

The study will generate important information on the prevalence and correlates of SAD among students at Makerere University, Kampala, Uganda. It will also provide valuable insights into the application of eHealth approaches for young populations within the education sector. The effectiveness of daily supportive email and WhatsApp messaging in addressing these mental health concerns will be evaluated, offering critical evidence to guide school policies and decisions on psychological interventions for students in low- and middle-income countries.

Strengths and Limitations

The strengths of this protocol include the following: first, the process of cocreation is a core strength of the intervention. This is a novel approach in low- and middle-income country eHealth interventions. This differentiates our protocol from top-down interventions and highlights participatory science.

Second, we expect that the survey will provide insightful data on the prevalence of SAD among university students. This will support the Counselling and Guidance Center of the university in tailoring interventions to the student community.

However, some limitations are anticipated. First, the outcome measures will be assessed at 1, 3, and 6 months. However, the effects of the intervention beyond 6 months remain unknown. It is also uncertain whether the benefits will decline once the daily email and WhatsApp messages end. Second, although the self-reported scales used to measure mental health variables are standardized, they are not diagnostic tools. Third, the demographics of study participants may not represent those of the wider student body at other universities (public and private or rural and urban), limiting the generalizability of the findings. Additionally, web-based surveys distributed via email and WhatsApp typically have response rates below 20%, which may hinder achieving the target sample size. Nonetheless, this will be the first study in Uganda to examine the prevalence and associated factors of SAD among students using an email and WhatsApp messaging program. We plan to use these preliminary data as a pilot for larger trials and invite collaborative follow-ups.

Conclusions

This study will provide vital information on the prevalence and correlates of SAD among students at Makerere University. The study will also provide preliminary evidence of the "feasibility of effectiveness" of the W4S (email and WhatsApp messaging) program, W4S, in reducing mental health disorder symptoms. The knowledge gained from this study will significantly impact the promotion of the management of SAD.

The study outcome will help inform policy decision-making for health care resource allocation in support of advocacy for tech equity in the educational sector, directed toward Ugandan students. In addition, it will address the eHealth-policy nexus. The feasibility data could reshape Uganda's national mental health policy.

Acknowledgments

No generative AI tools were used in the preparation of this manuscript.

XSL•FO
RenderX

Funding

This study was funded by the Global Psychological eHealth Foundation.

Data Availability

The datasets that will be generated and analyzed during this study will be available from the corresponding author upon reasonable request. There is a possibility of committing to open-access principles by naming repositories, such as Figshare, for raw data and protocols for metadata, fostering transparency and enabling secondary analyses by the global research community. Finally, we will have quarterly progress reports to ethics boards, including enrollment dashboards, to demonstrate proactive oversight and adaptability in a dynamic study environment.

Conflicts of Interest

BA and VIOA are board members of the Global Psychological eHealth Foundation, a not-for-profit organization. BA is the President and Chief Executive Officer of the Global Psychological eHealth Foundation. VIOA is the founder and principal investigator of the ResilienceNHope messaging programs. VIOA, BA, SOW, PN, ON, CN, CA, RK, and NN have no financial conflicts of interest related to this paper.

References

1. Stallman HM. Psychological distress in university students: a comparison with general population data. Aust Psychol 2010 Dec 1;45(4):249-257. [doi: 10.1080/00050067.2010.482109]
2. Depressive disorder (depression). World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/depression [accessed 2026-03-23]
3. Moitra M, Owens S, Hailemariam M, et al. Global mental health: where we are and where we are going. Curr Psychiatry Rep 2023 Jul;25(7):301-311. [doi: 10.1007/s11920-023-01426-8]
4. World mental health report: transforming mental health for all. : World Health Organization; 2022 URL: https://iris.who.int/bitstream/handle/10665/356119/9789240049338-eng.pdf?sequence=1 [accessed 2026-03-23]
5. The state of the world's children 2021. On my mind: promoting, protecting and caring for children's mental health. : United Nations Children's Fund (UNICEF); 2021 URL: https://www.unicef.org/media/114636/file/SOWC-2021-full-report-English.pdf [accessed 2026-03-16]
6. Keeley B. The state of the world's children 2021. On my mind: promoting, protecting and caring for children's mental health. : United Nations Children's Fund (UNICEF); 2021 URL: https://www.unicef.org/reports/state-worlds-children-2021 [accessed 2026-03-23]
7. Patel V, Saxena S, Lund C, et al. The Lancet Commission on global mental health and sustainable development. Lancet 2018 Oct 27;392(10157):1553-1598. [doi: 10.1016/S0140-6736(18)31612-X] [Medline: 30314863]
8. Depression and other common mental disorders: global health estimates. : World Health Organization; 2017 URL: https://iris.who.int/server/api/core/bitstreams/6bab42bc-df0f-4f68-a86d-28ebedb85e42/content [accessed 2026-03-16]
9. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022 Feb;9(2):137-150. [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
10. Bantjes J, Lochner C, Saal W, et al. Prevalence and sociodemographic correlates of common mental disorders among first-year university students in post-apartheid South Africa: implications for a public mental health approach to student wellness. BMC Public Health 2019 Jul 10;19(1):922. [doi: 10.1186/s12889-019-7218-y] [Medline: 31291925]
11. Brenneisen Mayer F, Souza Santos I, Silveira PSP, et al. Factors associated to depression and anxiety in medical students: a multicenter study. BMC Med Educ 2016 Oct 26;16(1):282. [doi: 10.1186/s12909-016-0791-1] [Medline: 27784316]
12. Ibrahim AK, Kelly SJ, Adams CE, Glazebrook C. A systematic review of studies of depression prevalence in university students. J Psychiatr Res 2013 Mar;47(3):391-400. [doi: 10.1016/j.jpsychires.2012.11.015] [Medline: 23260171]
13. Deng J, Zhou F, Hou W, et al. The prevalence of depressive symptoms, anxiety symptoms and sleep disturbance in higher education students during the COVID-19 pandemic: a systematic review and meta-analysis. Psychiatry Res 2021 Jul;301:113863. [doi: 10.1016/j.psychres.2021.113863] [Medline: 33984824]
14. Li W, Zhao Z, Chen D, Peng Y, Lu Z. Prevalence and associated factors of depression and anxiety symptoms among college students: a systematic review and meta-analysis. J Child Psychol Psychiatry 2022 Nov;63(11):1222-1230. [doi: 10.1111/jcpp.13606] [Medline: 35297041]
15. Tan GXD, Soh XC, Hartanto A, Goh AYH, Majeed NM. Prevalence of anxiety in college and university students: an umbrella review. J Affect Disord Rep 2023 Dec;14:100658. [doi: 10.1016/j.jadr.2023.100658]
16. Kaggwa MM, Arinaitwe I, Muwanguzi M, et al. Suicidal behaviours among Ugandan university students: a cross-sectional study. BMC Psychiatry 2022 Apr 1;22(1):234. [doi: 10.1186/s12888-022-03858-7] [Medline: 35365105]
17. Whitton SW, Whisman MA. Relationship satisfaction instability and depression. J Fam Psychol 2010 Dec;24(6):791-794. [doi: 10.1037/a0021734] [Medline: 21171780]

XSL•FO
RenderX

18. Othieno CJ, Okoth RO, Peltzer K, Pengpid S, Malla LO. Depression among university students in Kenya: prevalence and sociodemographic correlates. J Affect Disord 2014 Aug;165:120-125. [doi: 10.1016/j.jad.2014.04.070] [Medline: 24882188]
19. Olum R, Nakwagala FN, Odokonyero R. Prevalence and factors associated with depression among medical students at Makerere University, Uganda. Adv Med Educ Pract 2020;11:853-860. [doi: 10.2147/AMEP.S278841] [Medline: 33209071]
20. Najjuka SM, Checkwech G, Olum R, Ashaba S, Kaggwa MM. Depression, anxiety, and stress among Ugandan university students during the COVID-19 lockdown: an online survey. Afr Health Sci 2021 Dec;21(4):1533-1543. [doi: 10.4314/ahs.v21i4.6] [Medline: 35283951]
21. Koly KN, Sultana S, Iqbal A, Dunn JA, Ryan G, Chowdhury AB. Prevalence of depression and its correlates among public university students in Bangladesh. J Affect Disord 2021 Mar 1;282:689-694. [doi: 10.1016/j.jad.2020.12.137] [Medline: 33445093]
22. Ngasa SN, Sama CB, Dzekem BS, et al. Prevalence and factors associated with depression among medical students in Cameroon: a cross-sectional study. BMC Psychiatry 2017 Jun 9;17(1):216. [doi: 10.1186/s12888-017-1382-3] [Medline: 28599624]
23. Kaggwa MM, Najjuka SM, Bongomin F, Mamun MA, Griffiths MD. Prevalence of depression in Uganda: a systematic review and meta-analysis. PLOS One 2022;17(10):e0276552. [doi: 10.1371/journal.pone.0276552] [Medline: 36264962]
24. Rousseau KL, Thompson S, Pileggi LA, Henry M, Thomas KGF. Trends in the prevalence and severity of depressive symptoms among undergraduate students at a South African University, 2016–2019. S Afr J Psychol 2021 Mar;51(1):67-80. [doi: 10.1177/0081246320977759]
25. Castro-Ramirez F, Al-Suwaidi M, Garcia P, Rankin O, Ricard JR, Nock MK. Racism and poverty are barriers to the treatment of youth mental health concerns. J Clin Child Adolesc Psychol 2021;50(4):534-546. [doi: 10.1080/15374416.2021.1941058] [Medline: 34339320]
26. Muhorakeye O, Biracyaza E. Exploring barriers to mental health services utilization at Kabutare District Hospital of Rwanda: perspectives from patients. Front Psychol 2021;12:638377. [doi: 10.3389/fpsyg.2021.638377] [Medline: 33828506]
27. Kaggwa MM, Arinaitwe I, Nduhuura E, et al. Prevalence and factors associated with depression and suicidal ideation during the COVID-19 pandemic among university students in Uganda: a cross-sectional study. Front Psychiatry 2022;13:842466. [doi: 10.3389/fpsyt.2022.842466] [Medline: 35492697]
28. Agyapong B, Shalaby R, Eboreime E, Wei Y, Agyapong VIO. Self-reported alcohol abuse and the desire to receive mental health counselling predict suicidal thoughts/thoughts of self-harm among female residents of Fort McMurray. Int J Environ Res Public Health 2022 Oct 20;19(20):13620. [doi: 10.3390/ijerph192013620] [Medline: 36294198]
29. Andoh-Arthur J. Prevalence and determinants of depressive symptoms among university students in Ghana. J Affect Disord 2015 Jan;171:161-166. [doi: 10.1016/j.jad.2014.09.025]
30. Horwitz AG, McGuire T, Busby DR, et al. Sociodemographic differences in barriers to mental health care among college students at elevated suicide risk. J Affect Disord 2020 Jun 15;271:123-130. [doi: 10.1016/j.jad.2020.03.115] [Medline: 32479307]
31. Kihumuro RB, Kaggwa MM, Kintu TM, et al. Knowledge, attitude and perceptions of medical students towards mental health in a university in Uganda. BMC Med Educ 2022 Oct 20;22(1):730. [doi: 10.1186/s12909-022-03774-0] [Medline: 36266646]
32. Agardh A, Cantor-Graae E, Ostergren PO. Youth, sexual risk-taking behavior, and mental health: a study of university students in Uganda. Int J Behav Med 2012 Jun;19(2):208-216. [doi: 10.1007/s12529-011-9159-4] [Medline: 21590465]
33. Agyapong B, Shalaby R, Wei Y, Agyapong VIO. Can ResilienceNHope, an evidence-based text and email messaging innovative suite of programs help to close the psychological treatment and mental health literacy gaps in college students? Front Public Health 2022;10:890131. [doi: 10.3389/fpubh.2022.890131] [Medline: 36388394]
34. Agyapong VIO, Hrabok M, Vuong W, et al. Implementation and evaluation of a text message-based addiction counseling program (Text4Hope-Addiction Support): protocol for a questionnaire study. JMIR Res Protoc 2020 Nov 17;9(11):e22047. [doi: 10.2196/22047] [Medline: 33200993]
35. Chatterjee K, Edmonds VS, Girardo ME, Vickers KS, Hathaway JC, Stonnington CM. Medical students describe their wellness and how to preserve it. BMC Med Educ 2022 Jun 28;22(1):510. [doi: 10.1186/s12909-022-03552-y] [Medline: 35764972]
36. Nantaayi B, Ndawula RK, Musoke P, et al. Psychological distress and access to mental health services among undergraduate students during the COVID-19 lockdown in Uganda. Front Psychiatry 2022;13:792217. [doi: 10.3389/fpsyt.2022.792217] [Medline: 35722591]
37. Khan N, Butt MG, Awan F, et al. Global mental health commentary: using innovation to create a workforce to deliver and implement culturally adapted CBT in Pakistan. BMJ Ment Health 2025 Mar 4;28(1):e301389. [doi: 10.1136/bmjment-2024-301389] [Medline: 40037899]
38. Day S, Laver K, Jeon YH, Radford K, Low LF. Frameworks for cultural adaptation of psychosocial interventions: a systematic review with narrative synthesis. Dementia (London) 2023 Nov;22(8):1921-1949. [doi: 10.1177/14713012231192360] [Medline: 37515347]

XSL•FO
RenderX

39. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 1;3(1):e7. [doi: 10.2196/mental.4984] [Medline: 26932350]
40. Firth J, Torous J, Nicholas J, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298. [doi: 10.1002/wps.20472] [Medline: 28941113]
41. Green EP, Lai Y, Pearson N, et al. Expanding access to perinatal depression treatment in Kenya through automated psychological support: development and usability study. JMIR Form Res 2020 Oct 5;4(10):e17895. [doi: 10.2196/17895] [Medline: 33016883]
42. Mbuagbaw L, Thabane L, Ongolo-Zogo P, et al. The Cameroon mobile phone SMS (CAMPS) trial: a randomized trial of text messaging versus usual care for adherence to antiretroviral therapy. PLoS One 2012;7(12):e46909. [doi: 10.1371/journal.pone.0046909] [Medline: 23236345]
43. Agyapong B, Shalaby R, Vuong W, et al. Text4Hope effectiveness in reducing psychological symptoms among young adults in Canada: longitudinal and naturalistic controlled program evaluation. J Clin Med 2023;12(5):1942. [doi: 10.3390/jcm12051942]
44. CoBAMS annual report 2023. : Makerere University College of Business and Management Sciences (CoBAMS); 2023 URL: https://news.mak.ac.ug/wp-content/uploads/2024/05/Makerere-CoBAMS-Annual-Report-2023.pdf [accessed 2026-03-16]
45. Crooks N, Debra A, Coleman D, et al. Application of ADAPT-ITT: adapting an evidence-based HIV/STI mother-daughter prevention intervention for Black male caregivers and girls. BMC Public Health 2023 Jul 26;23(1):1426. [doi: 10.1186/s12889-023-16364-6] [Medline: 37491213]
46. Davis T, DiClemente RJ, Prietula M. Using ADAPT-ITT to modify a telephone-based HIV prevention intervention for SMS delivery: formative study. JMIR Form Res 2020 Oct 6;4(10):e22485. [doi: 10.2196/22485] [Medline: 32831178]
47. Fernandez ME, Ruiter RAC, Markham CM, Kok G. Intervention mapping: theory- and evidence-based health promotion program planning: perspective and examples. Front Public Health 2019;7:209. [doi: 10.3389/fpubh.2019.00209]
48. Diagnostic and Statistical Manual of Mental Disorders: DSM-5, 5th edition: American Psychiatric Association Publishing; 2013. [doi: 10.1176/appi.books.9780890425596]
49. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
50. COVID-19 Mental Disorders Collaborators. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. The Lancet 2021 Nov;398(10312):1700-1712. [doi: 10.1016/S0140-6736(21)02143-7]
51. Mohammad S, Almakran IW, Al-Montashri AS, et al. Depression, anxiety and stress and their associated social determinants in the Saudi college students. Health Psychol Res 2020 Dec 30;8(3):9263. [doi: 10.4081/hpr.2020.9263] [Medline: 33553792]
52. Sischka PE, Costa AP, Steffgen G, Schmidt AF. The WHO-5 well-being index—validation based on item response theory and the analysis of measurement invariance across 35 countries. J Affect Disord Rep 2020 Dec;1:100020. [doi: 10.1016/j.jadr.2020.100020]
53. Varma P, Junge M, Meaklim H, Jackson ML. Younger people are more vulnerable to stress, anxiety and depression during COVID-19 pandemic: a global cross-sectional survey. Prog Neuropsychopharmacol Biol Psychiatry 2021 Jul 13;109:110236. [doi: 10.1016/j.pnpbp.2020.110236] [Medline: 33373680]
54. Noor A, Gul N, Dill S, Gulab I. Exploring self-esteem and resilience among orphanage and non-orphanage adolescence. CRSSS 2024;2(2):397-408. [doi: 10.59075/s2eej367]
55. Suhail NA. Exploring mobile phone usage at higher education: a case study of Kampala University, Uganda. Int J Comput Appl 2017;170(2):30-34 [FREE Full text]
56. Agyapong VIO, Hrabok M, Shalaby R, et al. Text4Hope: receiving daily supportive text messages for 3 months during the COVID-19 pandemic reduces stress, anxiety, and depression. Disaster Med Public Health Prep 2022 Aug;16(4):1326-1330. [doi: 10.1017/dmp.2021.27] [Medline: 33551009]
57. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [Medline: 6668417]
58. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
59. Garba B, Mohamed SA, Mohamed MM, et al. Prevalence of depression and its correlates among undergraduate health science students in Mogadishu, Somalia: a cross-sectional study. BMC Psychiatry 2025 Feb 1;25(1):89. [doi: 10.1186/s12888-025-06553-5] [Medline: 39893421]
60. Sample size calculator. Survey Monkey. 2025. URL: https://www.surveymonkey.com/learn/research-and-analysis/sample-size-calculator/ [accessed 2026-03-23]
61. Young-Saver DF, Gornbein J, Starkman S, Saver JL. Handling of missing outcome data in acute stroke trials: advantages of multiple imputation using baseline and postbaseline variables. J Stroke Cerebrovasc Dis 2018 Dec;27(12):3662-3669. [doi: 10.1016/j.jstrokecerebrovasdis.2018.08.040] [Medline: 30297167]

XSL•FO
RenderX

62. Roszkowska E, Filipowicz-Chomko M, Łyczkowska-Hańćkowiak A, Majewska E. Extended Hellwig's method utilizing entropy-based weights and Mahalanobis distance: applications in evaluating sustainable development in the education area. Entropy (Basel) 2024 Feb 25;26(3):197. [doi: 10.3390/e26030197] [Medline: 38539709]
63. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ 2025 Apr 28;389:e081477. [doi: 10.1136/bmj-2024-081477] [Medline: 40294953]

Abbreviations

ADAPT-ITT: Assessment, Decision, Adaptation, Production, Topical experts, Integration, Training, Testing
CBT: cognitive behavioral therapy
FRAME: framework for reporting adaptations and modifications to evidence
GAD-7: 7-item Generalized Anxiety Disorder
IM: intervention mapping
PHQ-9: 9-item Patient Health Questionnaire
PSS: Perceived Stress Scale
SAD: stress, anxiety, and depression
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
W4S: Wellness4Students
WHO-5: World Health Organization-Five Well-Being Index

Edited by A Schwartz; submitted 22.Sep.2025; peer-reviewed by R Nooripour; revised version received 20.Feb.2026; accepted 25.Feb.2026; published 30.Mar.2026.

Please cite as:
Wandera SO, Nankinga O, Ndugga P, Namugenyi C, Ashaba C, Nakasujja N, Kasujja R, Agyapong B, Agyapong VIO
Wellness4Students Program to Mitigate Stress, Anxiety, and Depression Among Makerere University Students in Uganda: Protocol for a Longitudinal Study
JMIR Res Protoc 2026;15:e84637
URL: https://www.researchprotocols.org/2026/1/e84637
doi:10.2196/84637

© Stephen Ojiambo Wandera, Olivia Nankinga, Patricia Ndugga, Christabellah Namugenyi, Claire Ashaba, Noeline Nakasujja, Rosco Kasujja, Belinda Agyapong, Vincent Israel Opoku Agyapong. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Digital Information Environment of Lung Cancer and Lung Cancer Screening: Protocol for a Cross-Platform Social Media Content Analysis

Lisa Carter-Bawa, MPH, MSN, PhD; Ana Guadalupe Vielma, PhD; Gia Nealy, MPH; Diya Vemuganti, BS; Nidhi Patel

Cancer Prevention Precision Control Institute, Center for Discovery & Innovation, Hackensack Meridian Health, 123 Metro Blvd, 6th Floor, 6400 Pod, Nutley, NJ, United States

Corresponding Author:
Lisa Carter-Bawa, MPH, MSN, PhD
Cancer Prevention Precision Control Institute, Center for Discovery & Innovation, Hackensack Meridian Health, 123 Metro Blvd, 6th Floor, 6400 Pod, Nutley, NJ, United States

Abstract

Background: Lung cancer screening (LCS) with low-dose computed tomography reduces mortality by up to 20%, yet uptake in the United States remains below 6% of eligible individuals. Factors contributing to low uptake include lack of awareness, eligibility confusion, stigma associated with smoking history, and nihilistic beliefs about outcomes. Stigma triggers shame-avoidance behaviors, nihilism undermines perceived screening benefit, and misinformation amplifies both by spreading inaccurate eligibility criteria and exaggerated harms. Social media increasingly shapes how individuals encounter health information, form risk perceptions, and make screening decisions. Because platform architectures differ in content modality, algorithmic curation, and user demographics, single-platform studies cannot reliably characterize the digital information environment or identify platform-specific intervention targets.

Objective: This study aims to (1) systematically characterize the clinical accuracy, stigma prevalence, and decision-support quality of lung cancer and screening content across 7 major social media platforms; (2) quantify platform-specific patterns in stigma manifestation and nihilistic messaging; (3) test whether inaccurate or stigmatizing content is associated with disproportionate engagement relative to accurate, nonstigmatizing content; and (4) as an exploratory aim, identify digital opinion leaders who could serve as partners for evidence-based dissemination.

Methods: This cross-sectional content analysis will examine publicly accessible posts from Facebook, Instagram, TikTok, YouTube, X/Twitter, Reddit, and Bluesky. Posts will be identified through predefined search terms across 2 content domains: LCS and lung cancer narratives (diagnosis, treatment, survivorship). The sampling strategy combines relevance-based sampling, targeting approximately 700 - 1000 unique posts after deduplication—a sample size providing 80% power for cross-platform comparisons assuming medium effect sizes. A structured codebook operationalizing constructs from diffusion of innovations theory, attribution theory of stigma, and health misinformation frameworks will assess accuracy, stigma, decision support, and equity. All posts will be dual-coded by trained coders. Interrater reliability will be assessed using Gwet's AC1. Analyses will include descriptive statistics, cross-platform comparisons using chi-square and Kruskal-Wallis tests, and negative binomial regression models testing whether accuracy and stigma characteristics predict engagement.

Results: Data collection began in October 2025 and is projected to be complete by July 2026. As of March 2026, data have been collected from 181 posts across 7 platforms. Results are expected to be published by December 2026. Findings will characterize accuracy patterns, stigma prevalence, benefit-harm framing, and engagement dynamics across platforms, informing clinical communication tools, navigator training, and digital intervention development.

Conclusions: This protocol describes the first multiplatform, theory-informed analysis of lung cancer and LCS content on social media. The study will generate foundational evidence to inform stigma-informed communication strategies, decision support tools, and equitable dissemination approaches. The methodology provides a replicable framework for monitoring health information ecosystems across disease contexts.

International Registered Report Identifier (IRRID): PRR1-10.2196/89479

(JMIR Res Protoc 2026;15:e89479) doi:10.2196/89479

KEYWORDS

lung cancer screening; low-dose computed tomography; social media; digital health; misinformation; stigma; nihilism; health communication; content analysis; implementation science

XSL•FO
RenderX

Introduction

Background

Lung cancer remains the leading cause of cancer-related death in the United States, accounting for approximately 125,000 deaths annually [1]. Screening with low-dose computed tomography (LDCT) of the chest has been shown to reduce lung cancer mortality by 20% among high-risk individuals in the National Lung Screening Trial [2] and by 24% among men in the NELSON trial [3]. Based on this evidence, the United States Preventive Services Task Force (USPSTF) recommends annual LDCT screening for adults aged 50‐80 years with a 20 pack-year smoking history who currently smoke or quit within the past 15 years [4].

Despite this evidence, lung cancer screening (LCS) uptake remains alarmingly low. Recent estimates suggest that fewer than 6% of eligible individuals undergo screening annually [5], with substantial disparities by race, socioeconomic status, geography, and insurance coverage [6]. Multiple factors contribute to low uptake, including limited public awareness of screening availability, confusion about eligibility criteria, uncertainty about benefits and risks, logistical barriers, and clinician-level factors such as competing clinical priorities and insufficient time for shared decision-making [7,8].

Uniquely among cancer screenings, lung cancer carries substantial stigma rooted in its association with tobacco use. This stigma manifests as blame, shame, and social devaluation directed toward individuals with lung cancer or those at risk, regardless of their actual smoking history [9]. Stigma operates at multiple levels—internalized shame among patients, interpersonal blame from family and health care clinicians, and structural discrimination in resource allocation—and has been linked to delayed care-seeking, psychological distress, and poorer quality of life [10,11]. Closely related to stigma is therapeutic nihilism: the belief that lung cancer is uniformly fatal and that screening or treatment is futile [12]. Nihilistic beliefs may discourage both patients and clinicians from pursuing early detection, undermining the potential benefits of screening.

These constructs—stigma, nihilism, and misinformation—operate through distinct but interconnected pathways to undermine screening uptake. Stigma functions at the individual level by triggering shame-avoidance behaviors (avoiding screening to avoid confronting smoking history) and at the interpersonal level by discouraging disclosure and help-seeking [10]. Nihilism, the belief that lung cancer is uniformly fatal, undermines the perceived benefit of early detection, a core prerequisite for screening adoption under rational decision-making models [12]. Misinformation can amplify both constructs: inaccurate eligibility information excludes individuals who might benefit, while exaggerated harms or minimized benefits reinforce nihilistic beliefs. When these constructs converge in high-visibility social media content, they may create a self-reinforcing information ecosystem that systematically disadvantages informed screening decisions.

Social media has become a dominant channel through which individuals encounter health information, including cancer-related content [13]. Platforms such as YouTube, TikTok, Instagram, and Facebook host millions of health-related posts that shape public understanding of disease risk, screening options, and treatment outcomes [14]. For lung cancer specifically, social media serves as a space where survivors share narratives, clinicians disseminate guidance, advocacy organizations promote awareness, and misinformation circulates alongside evidence-based content. The algorithmic curation systems of these platforms may preferentially surface content that generates high engagement, potentially resulting in greater visibility for emotionally charged, sensational, or stigmatizing narratives relative to balanced, accurate information [15].

Despite the growing influence of social media on health perceptions and behaviors, the digital information environment for lung cancer and LCS has received limited systematic attention. Existing research has been confined to single-platform analyses [16,17], narrow constructs such as sentiment or misinformation [18], or cancer types other than lung cancer [19,20]. These limitations are consequential for 3 reasons. First, platform architectures differ fundamentally in content modality (video vs text), algorithmic curation, and user demographics, making cross-platform generalization from single-platform studies unreliable. Second, health information seekers increasingly navigate multiple platforms, encountering potentially contradictory messages that shape cumulative perceptions—a phenomenon invisible to single-platform research. Third, stigma and misinformation may manifest differently across platforms due to content format constraints and community norms. Short-form video platforms such as TikTok and Instagram Reels privilege emotional and narrative content—personal testimonials, fear appeals, and survivor stories—that may convey stigma through visual and affective cues rather than explicit claims. Text-based platforms such as Reddit and X/Twitter enable more detailed informational content, including eligibility discussions that may be accurate or misleading, and community norms that shape whether stigmatizing language is challenged or reinforced. These modality-driven differences are expected to shape both the nature of misinformation and the audience exposed to it, and they draw fundamentally different audiences with distinct information-processing patterns. No prior work has systematically compared clinical accuracy, stigma and nihilism, benefit-harm framing, representation and equity, and engagement dynamics across the dominant social media platforms to identify platform-specific patterns requiring tailored intervention. This gap is particularly consequential given emerging evidence that web-based information exposure influences cancer screening intentions and behaviors [21].

Theoretical Framework

This study is guided by 3 complementary theoretical frameworks that inform the coding scheme and analytic approach. Table 1 provides explicit mapping from theoretical constructs to codebook items.

XSL•FO
RenderX

Table . Theory-to-codebook mapping.

Theory and construct	Codebook item	Operational definition
Diffusion of innovations		
Relative advantage	Benefits mentioned (D1.3)	Mortality reduction, early detection, cure potential mentioned
Compatibility	Values reference (D2.2)	References patient values, preferences, or life context
Complexity	Procedural clarity (D1.6)	Accurate description of LDCT[a] procedure and what to expect
Observability	Survivor narrative (D1.7)	Personal success story or early detection outcome present
Attribution theory of stigma		
Controllability attribution	Explicit blame (D3.1)	Direct statements attributing illness to personal choice (“you did this”)
Responsibility judgment	Redemptive framing (D3.3)	Implies care/sympathy contingent on quitting (“if you quit, you deserve...")
Counter-attribution	Counter-stigma (D3.5)	Acknowledges addiction, structural factors, or challenges blame narratives
Health misinformation frameworks		
Accuracy	Clinical accuracy (D1.1 - 1.5)	Eligibility criteria, procedure, outcomes verified against USPSTF[b] guidelines
Completeness	Benefit-harm balance (D1.4)	Both benefits AND harms (false positives, radiation, anxiety) mentioned

[a]LDCT: low-dose computed tomography.

[b]USPSTF: United States Preventive Services Task Force.

1. Diffusion of innovations theory posits that adoption of new practices, including health behaviors such as cancer screening, is influenced by perceptions of the innovation’s characteristics [22]. Four constructs are particularly relevant: (1) relative advantage (perceived benefits compared with alternatives), (2) compatibility (alignment with values and existing practices), (3) complexity (perceived difficulty of understanding or undertaking the behavior), and (4) observability (visibility of outcomes). We operationalize these constructs to assess whether social media content communicates the advantages of LCS, addresses fit with patients’ lives and values, clarifies or complicates procedural understanding, and highlights screening success stories.
2. Attribution theory of stigma provides a framework for understanding how blame and responsibility attributions shape attitudes toward stigmatized conditions [23,24]. When disease is attributed to perceived controllable behaviors (such as smoking), individuals may be viewed as responsible for their illness and deserving of negative outcomes. This framework informs our coding of explicit and implicit blame, redemptive framing (suggesting that only those who quit “deserve” care), shame imagery, and counter-stigma messaging that challenges blame attributions.
3. Health misinformation frameworks describe how false or misleading health information spreads online and influences beliefs and behaviors [25,26]. Key constructs include accuracy (factual correctness), completeness (presence of important contextual information), source credibility, and emotional framing. We adapt these constructs to evaluate clinical accuracy of screening-related claims, balance in benefit-harm presentation, and the presence of misinformation about eligibility, procedures, or outcomes.

Together, these frameworks enable a comprehensive assessment of how social media content may facilitate or undermine informed screening decisions, perpetuate or challenge stigma, and align with or deviate from evidence-based guidance.

Objectives

This study has 3 primary aims and one exploratory aim.

- Primary aim 1: Characterize the clinical accuracy, completeness, and decision-support quality of LCS-related content across platforms, including eligibility criteria, procedural information, and benefit-harm framing.
- Primary aim 2: Quantify the prevalence, forms, and severity of stigma, blame, and nihilism in lung cancer and LCS content, and identify platform-specific patterns.
- Primary aim 3: Test whether inaccurate or stigmatizing content is associated with disproportionate engagement compared with accurate, nonstigmatizing content, adjusting for platform, content type, and creator characteristics. Engagement metrics are treated as indicators of relative visibility and reach rather than direct measures of algorithmic amplification; observed associations may reflect algorithmic promotion, user preferences, content characteristics, or their interaction.

XSL•FO
RenderX

- Exploratory aim: Document representation patterns in visible content and identify digital opinion leaders (DOLs) who could serve as partners for evidence-based dissemination. DOLs will be categorized by creator type for inclusion in engagement regression models, enabling direct comparison of content characteristics between DOLs and other creators.

Methods

Study Design and Reporting

This is a cross-sectional content analysis of publicly accessible social media posts following established methodological standards for digital health research [27,28]. The protocol established methodological standards for social media research [29] and best practices for content analysis methodology [30]. The study protocol is registered on the Open Science Framework (https://osf.io/krq4h) to enhance transparency and reproducibility.

Platforms

Seven platforms were selected based on user base size, demographic diversity, content modality, and relevance to health information seeking: (1) Facebook, (2) Instagram, (3) TikTok, (4) YouTube, (5) X/Twitter, (6) Reddit, and (7) Bluesky. LinkedIn was excluded due to its professional focus and minimal patient-generated content based on preliminary scoping; Threads was excluded due to insufficient lung cancer content volume at the time of protocol development. This selection captures a range of content formats (long-form video, short-form video, static images, and text-based discussion), algorithmic structures, and user communities. Bluesky is included as an emerging decentralized platform increasingly adopted by health communicators and researchers. We anticipate substantial variation in content volume, format, and community composition across platforms. Bluesky, as an emerging decentralized platform, is expected to yield fewer posts that are more likely news- or academic-oriented, whereas TikTok content is anticipated to be dominated by personal narratives and influencer-generated short-form video. These cross-platform differences are a central feature of the study design—the multiplatform approach is intended precisely to characterize how content and framing vary by platform architecture and community norms—but they may limit statistical power for platform-specific subgroup analyses on lower-volume platforms.

Search Strategy

Search terms were developed through an iterative process involving four sources: (1) clinical guideline terminology from the USPSTF 2021 recommendation statement and associated clinical guidance; (2) prior social media research on lung cancer and cancer screening, from which we extracted search terms used in published studies [16-20]; (3) input from an expert panel comprising LCS researchers, a behavioral scientist specializing in stigma (LC-B), and a patient advocacy representative, who reviewed candidate terms for comprehensiveness and relevance; and (4) known stigma and nihilism signals identified from the lung cancer stigma literature [9-12], which informed the inclusion of Bundle B terms specifically designed to surface blame and fatalism narratives (eg, “I caused my lung cancer,” “smoking and lung cancer blame”). The final term list was refined during the preliminary scoping phase described above, with terms that yielded no relevant content or exclusively irrelevant content removed, and terms that surfaced unanticipated relevant content added. Terms are organized into 2 bundles corresponding to distinct content domains:

- Bundle A (LCS): “lung cancer screening,” “low-dose CT,” “LDCT,” “CT scan for lung cancer,” “lung cancer screening guidelines,” “Medicare lung cancer screening,” “should I get screened for lung cancer,” “lung cancer screening eligibility.”
- Bundle B (Lung Cancer Narratives): “lung cancer,” “lung cancer survivor,” “stage 4 lung cancer,” “non-small cell lung cancer,” “small cell lung cancer,” “quit smoking lung cancer,” “I caused my lung cancer,” “smoking and lung cancer blame.”

Bundle B includes terms specifically designed to surface stigma-laden content (eg, “I caused my lung cancer”) to ensure adequate representation of this construct for analysis.

Search Execution

Search terms and the overall strategy were piloted across all 7 platforms during a preliminary scoping phase conducted in August-September 2025. This scoping assessed content volume, term relevance, and the degree to which search terms surfaced the target constructs (screening information, stigma/blame narratives, misinformation). Based on this scoping, we estimate a screening-to-inclusion ratio of approximately 2:1 to 3:1, meaning approximately 1500 - 3000 posts will be screened to achieve the target sample of 700 - 1000 eligible posts after applying exclusion criteria and deduplication. Should content volume fall below projections for specific platforms, we will use predefined contingency procedures: (1) extending the time window (eg, from 6 to 12 months for lower-volume platforms), (2) supplementing with additional related search terms identified during scoping, and (3) documenting any protocol modifications and their rationale in the final report. All contingency procedures and protocol deviations will be recorded and reported transparently. Search execution will use platform-native search interfaces accessed through standard web browsers in private/incognito mode without authentication to capture the public-view information environment.

Search execution will use platform-native search interfaces accessed through standard web browsers in private/incognito mode without authentication to capture the public-view information environment. This approach reflects real-world information-seeking behavior for users who are not logged in. Browsers will be configured with cleared cookies and cache to partially mitigate algorithmic personalization, though we acknowledge this approach captures the “default” algorithmic presentation rather than eliminating algorithmic influence entirely. For each platform, search queries will be entered verbatim; platform-specific syntax variations (eg, hashtag handling, quotation marks) will be documented in Multimedia Appendix 1 [4,27,28,30-33]. All searches will be conducted within a 72-hour window to minimize temporal variation. Search timestamps, query strings, result counts, and platform version

XSL•FO
RenderX

information will be logged to support reproducibility and contextualization.

Sampling Strategy

The sampling approach combines three complementary methods to capture diverse content types and maximize ecological validity [31]:

1. Relevance-based sampling: For each search term, the first 20 posts sorted by platform-default relevance ranking will be captured. This approach reflects the content most likely to be encountered by users conducting typical searches.
2. Engagement-based sampling: For each search term, the top 20 posts ranked by engagement metrics (likes, comments, shares, views) will be captured. This approach identifies high-visibility content that reaches the largest audiences.
3. Algorithmic recommendation sampling: For anchor terms ("lung cancer screening" and "lung cancer"), 5 posts surfaced through algorithmic recommendations (eg, "For You," "Suggested," "Up Next") will be captured. This approach samples content actively promoted by platform algorithms.

Posts will be deduplicated within and across sampling methods. Based on preliminary scoping and precedent from similar multiplatform content analyses [32,33], we anticipate a final sample of approximately 700 - 1000 unique posts.

This sample size was determined a priori through 3 considerations. First, for cross-platform comparisons (7 platforms), detecting medium effect sizes (Cohen w=0.30) in chi-square analyses with α values of 0.05% and 80% power requires approximately 108 posts per platform. Our target of 700 - 1000 posts meets this threshold. Second, for regression modeling with approximately 12 predictors, a minimum of 10 - 20 observations per predictor requires 120 - 240 observations; our sample substantially exceeds this threshold. Third, content saturation in qualitative synthesis typically occurs by 100 - 200 coded items in health communication research. We acknowledge that rare content types (eg, posts containing explicit eligibility misinformation) may have insufficient prevalence for subgroup analysis; such limitations will be reported transparently. The final sample size will be evaluated relative to both statistical requirements and practical content saturation; should saturation be reached earlier than anticipated or specific platforms yield lower-than-projected content volume, the final sample may fall at the lower end of the target range.

Time Frame

Platform-specific time windows reflect content lifecycle differences. YouTube and Facebook content remains relevant longer; thus, posts from the preceding 12 months will be included. For faster-paced platforms (X/Twitter, Instagram, TikTok, Reddit, Bluesky), posts from the preceding 6 months will be included.

Eligibility Criteria

The inclusion and exclusion criteria were as follows:

- Inclusion criteria: Publicly accessible posts in English containing substantive content about lung cancer or LCS (more than passing mention).
- Exclusion criteria: Private or members-only content; paid advertisements without educational content; non-English posts; duplicate reposts of identical content (shares and retweets are counted as engagement on the original post).

Data Extraction

For each post, the following data will be extracted: platform; posting date; content type (video, image, text, infographic, link); creator type (patient/survivor, caregiver, clinician/health system, advocacy organization, commercial entity, influencer/general creator); sponsorship disclosures; engagement metrics (likes, comments, shares, views as available); and a paraphrased summary of content.

Coding Framework

A structured codebook operationalizes the theoretical framework (Table 1) across 6 domains.

Domain 1: Clinical Accuracy and Completeness

For screening-focused posts, coders will assess whether LDCT is mentioned; whether eligibility criteria (age, smoking history) are stated and accurate; whether benefits (mortality reduction, early detection) are mentioned and appropriately framed; whether harms (false positives, incidental findings, radiation exposure, anxiety) are acknowledged; and whether overgeneralization occurs (eg, implying everyone should be screened). For general lung cancer posts, coders will assess mention of cancer type, stage, and prognosis accuracy, as well as the presence of medical misinformation (eg, unproven cures). A composite accuracy rating (mostly accurate, partially accurate, mostly inaccurate, cannot determine) will be assigned. Accuracy coding distinguishes three categories to prevent conflation of misinformation with opinion: (1) Factual claims about eligibility, procedures, or outcomes that can be verified against clinical guidelines (coded accurate, inaccurate, or incomplete); (2) Incomplete information that omits important context but contains no demonstrably false claims (coded "incomplete" rather than "inaccurate"); (3) Opinion or value statements that do not make verifiable factual claims (coded "not applicable for accuracy"). This taxonomy distinguishes legitimate patient preferences (eg, "I chose not to screen because the uncertainty stressed me") from misinformation (eg, "Screening doesn't reduce mortality"). Decision rules and boundary examples for each category are provided in Multimedia Appendix 2 [4,9,10,22-26,30,34].

Domain 2: Decision-Support Features

Informed by diffusion of innovations theory, coders will assess whether the post acknowledges choices or options; references values and preferences ("what matters to you"); provides appropriate next steps; and presents benefits and harms in a balanced manner. A decision-support strength score (0=none, 1=minimal/implicit, 2=explicit decision support) will be assigned.

Domain 3: Stigma, Blame, and Nihilism

Informed by attribution theory, coders will assess presence and intensity (none, mild, moderate, severe) of: explicit blame ("you

XSL•FO
RenderX

did this to yourself"); implicit blame (emphasis on personal responsibility without acknowledging addiction or structural factors); redemptive framing ("if you quit, you deserve care"); nihilism ("lung cancer is always fatal," "why bother screening"); counter-stigma messaging ("no one deserves lung cancer"); identity-first versus person-first language; and shame or disgust imagery. A composite stigma severity rating will be assigned.

Domain 4: Representation and Equity

Where discernible, coders will note apparent race/ethnicity, gender, and age of individuals depicted. Structural barrier mentions (insurance, transportation, rural access, language barriers, racism, mistrust) will be coded. The focus population (general, specific communities) will be documented. A conservative "not discernible" classification will be applied when demographic attributes cannot be reliably identified.

Domain 5: Tone and Emotional Framing

Coders will assess dominant emotional valence (hopeful, fearful, neutral) and presence of alarmism or sensationalism.

Domain 6: Engagement Metrics

Platform-specific engagement data will be standardized to enable cross-platform comparison. Posts will be categorized into engagement tertiles (high, medium, low) within each platform to account for differences in engagement scale across platforms.

Composite Rating Construction and Reporting

Three composite ratings synthesize item-level assessments within their respective domains. The composite accuracy rating integrates item-level assessments across Domain 1 items (D1.1-D1.5). A post is rated "mostly accurate" when all verifiable claims align with USPSTF guidelines and no demonstrably false claims are present; "partially accurate" when the post contains a mix of accurate and inaccurate or incomplete claims; "mostly inaccurate" when the predominant claims are inconsistent with clinical evidence; and "cannot determine" when insufficient factual content is present for assessment. The composite stigma severity rating integrates Domain 3 items (D3.1-D3.5), ranging from "none" (no stigma indicators present) through "mild" (single implicit indicator), "moderate" (multiple indicators or one explicit indicator), to "severe" (explicit blame combined with nihilistic framing or shame imagery). The decision-support strength score (0‐2) is assigned based on the number and quality of decision-support features present across Domain 2 items: 0 indicates no decision-support features, 1 indicates minimal or implicit features, and 2 indicates explicit decision support. When item-level ratings within a composite do not converge clearly, coders will apply the decision rules detailed in Multimedia Appendix 2 [4,9,10,22-26,30,34], which include worked examples of boundary cases. Unresolved borderline cases will be flagged for adjudication by the principal investigator (LC-B), with reasoning documented in the Coding Decision Log. Additionally, sensitivity analyses described in the Statistical Analysis section include reclassification of borderline composite ratings to assess whether findings are robust to these judgment calls. Item-level distributions will be reported for all accuracy, stigma, and decision-support items alongside composite scores, enabling readers to identify which specific components drive composite patterns.

The complete codebook with operational definitions and decision rules is provided in Multimedia Appendix 2 [4,9,10,22-26,30,34].

Coder Training and Reliability

A minimum of 2 coders will independently code each post. Coders will be research assistants with prior content analysis experience selected based on demonstrated attention to detail and familiarity with social media platforms.

Training will follow a structured 5-phase protocol:

1. Orientation (2 h): Overview of LCS clinical guidelines (USPSTF 2021), eligibility criteria, and stigma/nihilism concepts using published literature [9-12].
2. Codebook familiarization (2 h): Review of operational definitions with worked examples, boundary cases, and common pitfalls.
3. Calibration coding (4 h): Group coding of 20 posts stratified by platform with real-time discussion and codebook refinement based on encountered ambiguities.
4. Pilot reliability testing: Independent coding of 50 posts stratified by platform and content type. Gwet's AC1 [34] will be calculated for each domain; coders achieving AC1≥0.70 on all domains proceed; others receive targeted retraining on low-reliability domains.
5. Drift monitoring: Every 100 coded posts, a random 10% subsample will be recoded by the alternate coder to detect coding drift. Drift exceeding 0.10 decline in AC1 from baseline will trigger a recalibration meeting with codebook clarification.

Discrepancies will be flagged in real-time using a shared coding database. Coders will attempt resolution through discussion within 24 hours. Unresolved discrepancies will be escalated to the principal investigator (LC-B) for final determination; all adjudication decisions and reasoning will be documented in a Coding Decision Log to ensure transparency and support future replication.

Interrater reliability will be calculated using Gwet's AC1 [34] coefficient, which is preferred over Cohen kappa for content analysis due to its stability in the presence of prevalence imbalances and high agreement [34]. The threshold for acceptable reliability is AC1≥0.70 for accuracy and stigma domains. Percent agreement will be calculated for equity markers given the categorical and often conservative ("not discernible") nature of these codes.

Digital Opinion Leader Identification

As an exploratory aim, we will identify DOLs who influence public discourse around lung cancer and LCS. DOL candidates will be identified during content analysis based on creator characteristics and engagement patterns. DOLs will be categorized by creator type for inclusion in engagement regression models, enabling direct comparison of content characteristics (accuracy, stigma) between DOLs and other creators.

A structured scoring rubric (0‐9 points) will assess 5 criteria: (1) Credibility (0‐2 points): Verified status, professional credentials (clinician, researcher), or recognized organizational

XSL•FO
RenderX

affiliation; (2) Reach (0‐2 points): Platform-specific follower/subscriber thresholds (eg, YouTube ≥5000; X/Twitter ≥10,000; TikTok ≥20,000); (3) Engagement rate (0‐2 points): Average engagement≥10%=2 points; 5%‐9%=1 point; (4) Relevance (0‐2 points): 5 or more lung cancer–focused posts in audit period=2 points; 2‐4 posts=1 point; and (5) Multiplatform presence (0‐1 point): Active on 2 or more platforms with relevant content.

DOLs will be categorized into tiers: Tier 1 (7‐9 points)=high priority for partnership; Tier 2 (4‐6 points)=moderate priority; Tier 3 (<4 points)=monitor only. The DOL roster will be validated with community partners and patient advisory groups. The complete scoring rubric is provided in Multimedia Appendix 3 [22,35-40]. We acknowledge that this scoring rubric is pragmatically derived for stakeholder mapping purposes and has not been psychometrically validated; the tiered roster represents a practical output for dissemination planning rather than a validated measurement instrument.

Statistical Analysis

Descriptive analyses will summarize distributions of accuracy, stigma/nihilism, decision-support strength, equity markers, and engagement by platform, content domain (screening vs general lung cancer), and creator type.

Cross-platform comparisons will use chi-square or Fisher exact tests for categorical variables and Kruskal-Wallis tests for ordinal variables to identify significant differences in accuracy, stigma prevalence, and decision-support quality across platforms.

Engagement modeling will use negative binomial regression (for count outcomes) or ordinal logistic regression (for engagement tertiles) to test the hypothesis that content characteristics predict engagement. Predictors will include accuracy rating, stigma severity, creator type (including DOL status), content type, platform, and equity markers. This analysis will test whether inaccurate or stigmatizing content is associated with disproportionate engagement, as specified in Primary Aim 3. We note that significant associations between content characteristics and engagement should be interpreted as evidence of differential visibility rather than causal evidence of algorithmic bias, as multiple mechanisms may contribute to observed engagement patterns.

Subgroup analyses will compare screening-focused versus general lung cancer content; institutional versus individual creators; video versus static posts; and high- versus low-engagement content.

Sensitivity analyses will assess the robustness of findings to reclassification of borderline accuracy ratings, exclusion of posts with missing engagement data, normalization of engagement by time since posting, and comparison of content characteristics between relevance-based and algorithmic samples to assess potential algorithmic bias direction.

Qualitative synthesis will contextualize quantitative findings through thematic analysis of paraphrased content excerpts, identifying illustrative examples of key patterns.

Ethical Considerations

This study was reviewed by the Hackensack Meridian Health Institutional Review Board and determined not to constitute regulated research involving human participants under 45 CFR 46.102(e)(1) as it involves secondary analysis of publicly available data with no direct interaction with individuals and no collection of identifiable private information.

While legally permissible, analysis of public health narratives raises ethical considerations beyond regulatory compliance. We adopt a minimization approach: (1) no direct quotations will appear in publications; content will be paraphrased and aggregated such that individual posters cannot be identified; (2) posts containing acute distress signals (eg, expressions of active suicidal ideation) will be excluded from analysis and not retained; and (3) the research team will not interact with or contact any content creators. These procedures exceed minimum regulatory requirements and align with emerging best practices for ethical social media research [27].

All procedures comply with the platform Terms of Service. No informed consent is required as the study uses only publicly posted content with no direct contact with social media users.

Data Management

Data will be stored on institution-secured servers with role-based access restricted to approved research team members. The data management plan distinguishes three tiers: (1) raw data files containing post URLs and metadata will be retained for verification purposes but will not be shared publicly to comply with platform Terms of Service and protect user privacy; (2) analytic datasets will contain paraphrased content summaries, engagement metrics, and coded variables without direct identifiers linking to specific posts or users; and (3) aggregated summary datasets and the complete codebook will be prepared for public sharing. All data will be retained for 7 years following publication per institutional policy. Upon publication of results, the codebook, Coding Decision Log, aggregated summary statistics, and deidentified analytic datasets will be deposited in the Open Science Framework repository to support replication and secondary analysis.

Results

Data collection began in October 2025, following finalization of the sampling framework and completion of coder training and is projected to be completed by July 2026. As of March 2026, data have been collected from 181 posts across 7 platforms. During the initial phase, posts were systematically identified and captured across all 7 platforms using the predefined search strategy. Dual coding commenced with weekly calibration meetings to ensure consistency in codebook application and maintain interrater reliability standards. Preliminary quality checks indicate strong coder alignment, with reliability metrics meeting prespecified thresholds.

Coding and adjudication will continue until the full dataset is complete. Formal quantitative analyses, including descriptive summaries, cross-platform comparisons, and regression modeling, will begin upon coding completion. The analytic

XSL•FO
RenderX

phase is expected to conclude by mid-2026, with manuscript preparation and dissemination to follow.

Discussion

Principal Findings

This study addresses a critical gap in lung cancer prevention and early detection research. While substantial investment has been directed toward LCS trials, implementation science, and health system interventions, the digital information environment that shapes public understanding, fear, stigma, and nihilism has been largely unexamined. This protocol describes the first systematic, multiplatform, theory-informed analysis of lung cancer and LCS content on social media.

The study is expected to yield several novel contributions. First, by assessing accuracy across platforms and content types, findings will clarify the extent to which publicly visible LCS information aligns with guidelines, where inaccuracies cluster, and which creator types are most likely to disseminate misleading content. Second, the systematic application of a stigma and nihilism lens—operationalized through attribution theory—will provide the first empirical mapping of how blame, shame, and fatalism manifest and circulate in digital spaces. Third, the pairing of content characteristics with engagement metrics will illuminate whether stigmatizing or inaccurate content is associated with disproportionate engagement, which may reflect differential algorithmic visibility, user engagement preferences, or both, offering insights into the attention economy surrounding LCS. Fourth, documentation of representation patterns will quantify gaps in visibility for communities disproportionately affected by lung cancer but potentially underrepresented in online narratives. Importantly, interpretation of cross-platform differences will account for the inherent confounding between platform and content modality. Differences in stigma prevalence or accuracy between, for example, TikTok and Reddit may reflect the constraints and affordances of short-form video versus text-based discussion as much as community-level differences. The planned subgroup analysis comparing video versus static content will help disentangle modality effects from platform effects, though complete separation is not possible in an observational design.

Comparison With Prior Work

Existing research on lung cancer content in social media has been limited in scope. Alban et al [16] analyzed YouTube videos related to lung cancer but did not assess stigma or decision-support features. Zhao et al [17] examined LCS information on social media but focused on a single platform. Our multiplatform approach, theory-informed codebook, and integration of stigma assessment with engagement dynamics extend this work substantially.

Translational Applications

Findings will directly inform several ongoing initiatives. The characterization of dominant misconceptions, nihilistic narratives, and high-engagement misinformation will guide refinements to LungTalk, an evidence-based computer-tailored health communication and decision support tool for LCS. Real-world examples of stigmatizing content will be integrated into training modules for patient navigators and community health workers, enhancing their capacity to recognize and address digital influences on patient perceptions. The dual-coded, reliability-tested dataset will provide training examples that may inform future development of automated stigma detection tools. The labeled examples of explicit and implicit stigma in naturalistic social media contexts could serve as ground truth for supervised learning approaches. However, the feasibility of such applications depends on adequate representation of stigma-positive cases, class balance, and domain-specific model architecture considerations; this study is designed to characterize content patterns and is not intended to develop or validate automated detection systems. Finally, insights about platform-specific patterns will inform dissemination strategies for health systems and advocacy organizations seeking to reach eligible populations with accurate, stigma-informed messaging.

Limitations

Several anticipated limitations warrant acknowledgment. First, the study characterizes the publicly visible information environment—the content most likely encountered by individuals conducting typical health information searches—rather than the totality of lung cancer discussions. This scope is intentional: publicly visible content shapes population-level exposure and is the appropriate target for public health communication interventions. Findings should not be generalized to private groups, direct messages, or content not surfaced by platform algorithms, which may differ substantially in tone and accuracy.

Second, the cross-sectional design captures a specific temporal window in a rapidly evolving digital environment. We address this limitation through three strategies: (1) documenting search timestamps, platform versions, and procedural details to enable contextualization of findings within the specific data collection period; (2) focusing on structural patterns (eg, stigma prevalence differences across platforms, relationship between accuracy and engagement) that are likely to be more stable than specific content; and (3) providing methodological infrastructure (codebook, sampling framework, analytic code) that enables future replication to track temporal trends. We position this study as establishing a baseline against which future shifts can be measured rather than claiming to capture a permanent state.

Third, representation coding is constrained by visual ambiguity; the conservative “not discernible” classification prevents inaccurate assumptions but may undercount certain demographic patterns. Fourth, the DOL scoring rubric is pragmatically derived for stakeholder mapping purposes and has not been psychometrically validated. Fifth, the inclusion of 7 platforms with substantially different content volumes may result in unequal sample sizes across platforms, particularly for Bluesky, which may yield fewer eligible posts. Should any platform produce fewer than the minimum posts needed for meaningful subgroup analysis, platform-specific findings will be interpreted descriptively rather than inferentially, and this will be noted transparently. Sixth, this study analyzes postlevel content and does not examine user comments or replies. Comment sections represent an important and often distinct dimension of public

XSL•FO
RenderX

discourse, sentiment, and narrative framing—for example, stigmatizing or supportive responses to a patient's personal narrative may substantially shape how that content is experienced by other viewers. Future research should examine comment-level content and sentiment to capture the full scope of public engagement with lung cancer and screening-related posts on social media. Despite these limitations, the systematic multiplatform design, dual-coding with reliability assessment, and theory-informed codebook collectively strengthen rigor.

Conclusions

This protocol establishes a rigorous, reproducible framework for characterizing the digital information environment surrounding lung cancer and LCS. By integrating clinical accuracy assessment, stigma and nihilism coding, decision-support evaluation, and engagement analysis within a theory-informed design, the study will generate foundational evidence to guide communication interventions, navigator training, and equitable dissemination strategies. The methodology provides a replicable framework for monitoring health information ecosystems and is adaptable to other cancer types, screening modalities, and emerging platforms.

Funding

LCB is supported in part by the Georgetown Lombardi Comprehensive Cancer Center Support Grant (P30CA051008; Principal Investigator: Louis Weiner).

Data Availability

The complete codebook, Coding Decision Log, and aggregated datasets will be deposited in the Open Science Framework upon publication of the results manuscript. Raw data containing post URLs will not be shared to protect user privacy and comply with platform policies.

Authors' Contributions

LC-B conceptualized the study, developed the methodology, search strategy development, supervised the research team, and wrote the original draft. AGV contributed to methodology development, manuscript review, and editing. GN, DV, and NP contributed to data collection and manuscript review. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Platform-specific search parameters, syntax documentation, and data extraction templates.
[DOCX File, 24 KB - resprot_v15i1e89479_app1.docx]

Multimedia Appendix 2
Complete codebook with operational definitions, coding decision rules, and boundary case examples.
[DOCX File, 34 KB - resprot_v15i1e89479_app2.docx]

Multimedia Appendix 3
Digital opinion leader scoring rubric and identification procedures.
[DOCX File, 22 KB - resprot_v15i1e89479_app3.docx]

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74(1):12-49. [doi: 10.3322/caac.21820] [Medline: 38230766]
2. National Lung Screening Trial Research Team, Aberle DR, Adams AM, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med 2011 Aug 4;365(5):395-409. [doi: 10.1056/NEJMoa1102873] [Medline: 21714641]
3. de Koning HJ, van der Aalst CM, de Jong PA, et al. Reduced lung-cancer mortality with volume CT screening in a randomized trial. N Engl J Med 2020 Feb 6;382(6):503-513. [doi: 10.1056/NEJMoa1911793] [Medline: 31995683]
4. US Preventive Services Task Force, Krist AH, Davidson KW, et al. Screening for lung cancer: US preventive services task force recommendation statement. JAMA 2021 Mar 9;325(10):962-970. [doi: 10.1001/jama.2021.1117] [Medline: 33687470]
5. Fedewa SA, Kazerooni EA, Studts JL, et al. State variation in low-dose computed tomography scanning for lung cancer screening in the United States. J Natl Cancer Inst 2021 Aug 2;113(8):1044-1052. [doi: 10.1093/jnci/djaa170] [Medline: 33176362]

XSL•FO
RenderX

6. Sosa E, D'Souza G, Akhtar A, et al. Racial and socioeconomic disparities in lung cancer screening in the United States: a systematic review. CA Cancer J Clin 2021 Jul;71(4):299-314. [doi: 10.3322/caac.21671] [Medline: 34015860]
7. Raz DJ, Wu GX, Consunji M, et al. Perceptions and utilization of lung cancer screening among primary care physicians. J Thorac Oncol 2016 Nov;11(11):1856-1862. [doi: 10.1016/j.jtho.2016.06.010] [Medline: 27346412]
8. Carter-Harris L, Brandzel S, Wernli KJ, Roth JA, Buist DSM. A qualitative study exploring why individuals opt out of lung cancer screening. Fam Pract 2017 Apr 1;34(2):239-244. [doi: 10.1093/fampra/cmw146] [Medline: 28122849]
9. Hamann HA, Ostroff JS, Marks EG, Gerber DE, Schiller JH, Lee SJC. Stigma among patients with lung cancer: a patient-reported measurement model. Psychooncology 2014 Jan;23(1):81-92. [doi: 10.1002/pon.3371] [Medline: 24123664]
10. Chambers SK, Dunn J, Occhipinti S, et al. A systematic review of the impact of stigma and nihilism on lung cancer outcomes. BMC Cancer 2012 May 20;12:184. [doi: 10.1186/1471-2407-12-184] [Medline: 22607085]
11. Cataldo JK, Jahan TM, Pongquan VL. Lung cancer stigma, depression, and quality of life among ever and never smokers. Eur J Oncol Nurs 2012 Jul;16(3):264-269. [doi: 10.1016/j.ejon.2011.06.008] [Medline: 21803653]
12. Jonnalagadda S, Bergamo C, Lin JJ, et al. Beliefs and attitudes about lung cancer screening among smokers. Lung Cancer 2012 Sep;77(3):526-531. [doi: 10.1016/j.lungcan.2012.05.095] [Medline: 22681870]
13. Social media fact sheet. Pew Research Center.: Pew Research Center; 2024. URL: https://www.pewresearch.org/internet/fact-sheet/social-media/ [accessed 2025-11-01]
14. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85. [doi: 10.2196/jmir.1933] [Medline: 23615206]
15. Vosoughi S, Roy D, Aral S. The spread of true and false news online. Science 2018 Mar 9;359(6380):1146-1151. [doi: 10.1126/science.aap9559] [Medline: 29590045]
16. Alban J, Alban LF, Clayburn A, Khanal A, Feldman L. Video-Based education in lung cancer screening. J Cancer Educ 2021 Apr;36(2):305-309. [doi: 10.1007/s13187-019-01629-9] [Medline: 31729695]
17. Zhao Y, Huo J, Prosperi M, Guo Y, Li Y, Bian J. Exploring lung cancer screening discussions on Twitter. Stud Health Technol Inform 2019 Aug 21;264:2011-2012. [doi: 10.3233/SHTI190760] [Medline: 31438454]
18. Wilner T, Holton A. Breast cancer prevention and treatment: misinformation on Pinterest, 2018. Am J Public Health 2020 Oct;110(S3):S300-S304. [doi: 10.2105/AJPH.2020.305812] [Medline: 33001732]
19. Vraga EK, Stefanidis A, Lamprianidis G, et al. Cancer and social media: a comparison of traffic about breast cancer, prostate cancer, and other reproductive cancers on Twitter and Instagram. J Health Commun 2018;23(2):181-189. [doi: 10.1080/10810730.2017.1421730] [Medline: 29313761]
20. Sutton J, Vos SC, Olson MK, et al. Lung cancer messages on Twitter: content analysis. J Am Coll Radiol 2018 Jan;15(1):210-217. [doi: 10.1016/j.jacr.2017.09.043] [Medline: 30068502]
21. Glanz K, Bishop DB. The role of behavioral science theory in development and implementation of public health interventions. Annu Rev Public Health 2010;31:399-418. [doi: 10.1146/annurev.publhealth.012809.103604] [Medline: 20070207]
22. Rogers EM. Diffusion of Innovations, 5th edition: Free Press; 2003.
23. Weiner B, Perry RP, Magnusson J. An attributional analysis of reactions to stigmas. J Pers Soc Psychol 1988;55(5):738-748. [doi: 10.1037/0022-3514.55.5.738] [Medline: 3210143]
24. Link BG, Phelan JC. Conceptualizing stigma. Annu Rev Sociol 2001 Aug;27(1):363-385. [doi: 10.1146/annurev.soc.27.1.363]
25. Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552. [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]
26. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187. [doi: 10.2196/17187] [Medline: 33470931]
27. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001 Nov 10;323(7321):1103-1105. [doi: 10.1136/bmj.323.7321.1103] [Medline: 11701577]
28. Sinnenberg L, Buttenheim AM, Padrez K, Mancheno C, Ungar L, Merchant RM. Twitter as a tool for health research: a systematic review. Am J Public Health 2017 Jan;107(1):e1-e8. [doi: 10.2105/AJPH.2016.303512] [Medline: 27854532]
29. Kaushal A, Bravo C, Duffy S, et al. Developing Reporting Guidelines for Social Media Research (RESOME) by Using a Modified Delphi method: protocol for guideline development. JMIR Res Protoc 2022 May 9;11(5):e31739. [doi: 10.2196/31739] [Medline: 35532999]
30. Krippendorff K. Content Analysis: An Introduction to Its Methodology, 4th edition: Sage Publications; 2018.
31. Fu J, Li C, Zhou C, et al. Methods for analyzing the contents of social media for health care: scoping review. J Med Internet Res 2023 Jun 26;25:e43349. [doi: 10.2196/43349] [Medline: 37358900]
32. Chen L, Wang X, Peng TQ. Nature and diffusion of gynecologic cancer-related misinformation on social media: analysis of tweets. J Med Internet Res 2018 Oct 16;20(10):e11515. [doi: 10.2196/11515] [Medline: 30327289]
33. Basch CH, Kecojevic A, Wagner VH. Coverage of the COVID-19 pandemic in the online versions of highly circulated US daily newspapers. J Community Health 2020 Dec;45(6):1089-1097. [doi: 10.1007/s10900-020-00913-w] [Medline: 32902813]
34. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 2008 May;61(Pt 1):29-48. [doi: 10.1348/000711006X126600] [Medline: 18482474]

XSL•FO
RenderX

35. Abidin C. Internet Celebrity: Understanding Fame Online: Emerald Publishing; 2018.
36. Freberg K, Graham K, McGaughey K, Freberg LA. Who are the social media influencers? a study of public perceptions of personality. Public Relat Rev 2011 Mar;37(1):90-92. [doi: 10.1016/j.pubrev.2010.11.001]
37. Pilgrim K, Bohnet-Joschko S. Selling health and happiness how influencers communicate on Instagram about dieting and exercise: mixed methods research. BMC Public Health 2019 Aug 6;19(1):1054. [doi: 10.1186/s12889-019-7387-8] [Medline: 31387563]
38. De Veirman M, Cauberghe V, Hudders L. Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude. International Journal of Advertising 2017 Sep 3;36(5):798-828. [doi: 10.1080/02650487.2017.1348035]
39. Araujo T, Neijens P, Vliegenthart R. Getting the word out on Twitter: the role of influentials, information brokers and strong ties in building word-of-mouth for brands. International Journal of Advertising 2017 May 4;36(3):496-513. [doi: 10.1080/02650487.2016.1173765]
40. Israel BA, Eng E, Schulz AJ, Parker EA. Methods for Community-Based Participatory Research for Health, 2nd edition: Jossey-Bass; 2012.

Abbreviations

DOL: digital opinion leader
LCS: lung cancer screening
LDCT: low-dose computed tomography
USPSTF: United States Preventive Services Task Force

Edited by J Sarvestan; submitted 12.Dec.2025; peer-reviewed by B Nickel, S Amin; revised version received 10.Feb.2026; accepted 24.Feb.2026; published 30.Mar.2026.

Please cite as:
Carter-Bawa L, Vielma AG, Nealy G, Vemuganti D, Patel N
The Digital Information Environment of Lung Cancer and Lung Cancer Screening: Protocol for a Cross-Platform Social Media Content Analysis
JMIR Res Protoc 2026;15:e89479
URL: https://www.researchprotocols.org/2026/1/e89479
doi:10.2196/89479

© Lisa Carter-Bawa, Ana Guadalupe Vielma, Gia Nealy, Diya Vemuganti, Nidhi Patel. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Validation of Machine Learning Models for Predicting Early Cognitive Decline Using Home Sensor–Derived Behavioral Data: Sensors in-Home for Elder Wellbeing (SINEW) Cohort Study

James Patrick Moon[1], MBChB, BAO, MHSc; Khalid Abdul Jabbar[2], PhD; Tony Chin Ian Tay[2], BA; Laura Tay[3], MBBS, MMed, MCI; Rathi Mahendran[4], MBBS, MMed; Tze Pin Ng[5], MBBS, PhD; Shian Ming Tan[1], MBBS; Wilbur Zhi Hao Koh[6], MCP; Weng Yan Ying[7], BEng; Ah-Hwee Tan[7], PhD; Tih-Shih Lee[8], MD, PhD; Iris Rawtaer[1,2,9], MBBS, MMed, MCI

[1]Department of Psychiatry, Sengkang General Hospital, Singapore, Singapore

[2]Research Office, Sengkang General Hospital, 110 Sengkang E Way, Singapore, Singapore

[3]Department of Geriatric Medicine, Sengkang General Hospital, Singapore, Singapore

[4]Mind Care Clinic @ SBF, Mind Science Centre, National University of Singapore, Singapore, Singapore

[5]Population Health and Integrated Care Office, Sengkang General Hospital, Singapore, Singapore

[6]Department of Psychology, Sengkang General Hospital, Singapore, Singapore

[7]School of Computing and Information Systems, Singapore Management University, Singapore, Singapore

[8]Laboratory of Neurobehavioral Genomics, Neuroscience and Behavioral Disorders Programme, Duke-NUS Graduate Medical School, Singapore, Singapore

[9]SingHealth Duke-NUS Academic Medical Centre, Singapore, Singapore

Corresponding Author:
Iris Rawtaer, MBBS, MMed, MCI
Department of Psychiatry, Sengkang General Hospital, Singapore, Singapore

Abstract

Background: As the global population continues to age, the prevalence of geriatric conditions, including dementia and frailty, is also increasing. Early identification of individuals at an elevated risk of these conditions, such as those presenting with mild cognitive impairment (MCI) or prefrailty, can provide a critical window for prompt intervention aimed at preventing or reversing disease progression. To promote such early identification, there is a burgeoning interest in the use of digital sensor technology and predictive modeling.

Objective: This study aimed to use a continuous, home-based monitoring sensor system for older adults to distinguish those exhibiting normal aging from those with MCI, early dementia, prefrailty, or frailty, and to predict their transition from normal aging to one of these conditions.

Methods: This longitudinal cohort study will recruit 200 community-dwelling adults aged ≥65 years with normal cognition or MCI at baseline. A multi-sensor system will be installed in participants' homes, including passive infrared motion sensors, door contact sensors, bed sensors, medication box sensors, wearable activity bands, and Bluetooth proximity beacons. These devices will continuously capture spatiotemporal activity patterns, mobility indicators, sleep behaviors, and medication-taking routines. Annual assessments will include standardized cognitive tests (eg, Montreal Cognitive Assessment, Mini-Mental State Examination, Rey Auditory-Verbal Learning Test, digit span, Color Trails Test, semantic fluency, Stroop), frailty measures (modified Fried phenotype, gait speed, grip strength), mental health scales, sleep quality, and psychosocial indicators. Sensor-derived features—such as gait variability, activity regularity, sleep fragmentation, and medication adherence patterns—will be integrated with clinical data to develop supervised machine learning models. Planned approaches include logistic regression, random forests, gradient boosting, and deep learning. Model performance will be evaluated using cross-validation and independent test sets. Primary metrics will include area under the receiver operating characteristic curve, sensitivity, specificity, precision, recall, and F_1-score. Models will be benchmarked against gold-standard clinical diagnoses and validated using temporal subsets of the dataset.

Results: Enrollment for this study started in November 2019 and will continue until March 2030. As of June 2025, we have enrolled 138 participants. Full data analysis has yet to begin.

Conclusions: We aim to develop a reliable and effective sensor system for in-home use that will facilitate the early detection of cognitive and physical decline. In so doing, it will add to our current understanding of digital biomarkers. It is common for

XSL•FO
RenderX

older adults to seek clinical intervention only when their cognitive impairment has already reached an advanced stage. The implementation of readily deployable sensor systems within community settings presents us with opportunities for prompt intervention, which holds the potential for delaying or reversing disease progression and allowing for a greater number of functional and meaningful years.

International Registered Report Identifier (IRRID): DERR1-10.2196/79490

(JMIR Res Protoc 2026;15:e79490) doi:10.2196/79490

KEYWORDS

dementia; neurocognitive disorder; mild cognitive disorder; prefrailty; frailty; internet of things; aged; technology

Introduction

The global population is undergoing significant shifts in its age demographics. By 2050, about 16.3% of the global population will be 65 or older, a sharp increase from 10.0% in 2023 [1]. This shift will pose significant challenges for health care systems due to the economic and medical burdens of providing care for older adults with physical and cognitive limitations. Indeed, frailty and cognitive impairment are two of the most prevalent conditions in the older adult population [2,3].

Frailty, a major syndrome in geriatrics, is characterized by deficits across multiple systems and reduced physiological reserves, leading to increased vulnerability to acute stress and adverse outcomes in older adults. It is associated with a decline in quality of life, heightened disability, mortality, and increased health care use [4]. In high-income countries, the weighted prevalence of frailty was estimated to be 10.7%, whereas the prevalence of prefrailty (PF), a prodromal risk state that can be reversed, was estimated to be 41.6% [5]. Meanwhile, cognitive impairment refers to the deterioration of intellectual abilities such as memory, reasoning, language, and problem-solving. Although it is often observed in older adults, its severity can vary from mild memory lapses to severe dementia. More specifically, mild cognitive impairment (MCI) describes the presence of cognitive decline concomitant with a lack of severity to significantly affect daily functioning [6,7]. MCI affects 6% - 20% of people aged 60 - 89 years [8-10], and it is recognized as a prodromal stage of dementia, with a high probability of progression to dementia within 1 to 3 years [11,12].

The combination of early identification and interventions targeting physical, social, and behavioral aspects of MCI and PF can effectively prevent, delay, or even reverse the progression of dementia and frailty, thereby enabling older adults to age well and maintain a good quality of life within their communities [13-24]. Indeed, screening older adults living in the community and in primary care settings more regularly could help in the identification of symptoms of geriatric conditions at an early stage [25-27]. However, the epidemic of dementia in the aging population cannot be tackled using conventional methods, as these either require reliable informant caregivers or sporadic face-to-face screening by primary health care providers [28].

The shortcomings of these conventional procedures have given way to an increasing demand for the use of digital biomarker technologies, some of which can be deployed effectively and continuously in community settings where older adults reside. There has also been a growing emphasis on the ability of these technologies to monitor cognitive function, physical frailty, and mental health [29-32]. A promising approach related to such technologies for monitoring older adults is the use of in-home sensor systems, including passive infrared (PIR) sensors and door contact magnetic sensors. A home-based study investigated the use of PIR motion sensors and acceleration sensors for the continuous monitoring of activities of daily living [33]. This study found that activities of daily living patterns visualized through activity maps could discriminate between older adults with healthy cognition and those with cognitive impairments. Another study demonstrated that a multi-sensor system including PIR motion sensors, magnetic contact sensors, and bed sensors could detect decreased engagement in activities of daily living and increased sleep interruptions in older adults with MCI [26]. Additionally, a study that measured heart rate and gait variability using wearable sensors found that these measures could detect MCI in a cost-effective manner under real-world conditions [34].

Two notable studies conducted by the Collaborative Aging (in Place) Research Using Technology (also known as CART) group used sensors in home-based settings to assess gait speed and variability and identify cognitive decline [35,36]. The first study, which involved 93 participants, revealed that the group with MCI had a coefficient of variation in median walking speed twice that of the healthy group. Over a follow-up period of approximately 2 years, the researchers identified 3 distinct trajectories of mean weekly walking speeds (fast, moderate, and slow) and found that participants with MCI were more likely to be in the slow-speed group. Additionally, the researchers identified 4 distinct trajectories of walking variability, with nonamnestic MCI participants being less likely to be in the stable coefficient of variation group, suggesting that walking speed variability may serve as an early indicator of cognitive decline. In another study conducted over around 2 years, 64 cognitively healthy older adults were observed; those who developed MCI later were more likely to open their pillboxes later in the day and demonstrated more variability in their medication-taking habits than those who remained cognitively stable [37]. Furthermore, a recent meta-analysis revealed that sensors and digital biomarker technologies could detect MCI and frailty with sensitivities of 84% and 82%, respectively [38].

Despite the existing body of research, a systematic review conducted by Piau et al [29] and the work by the Collaborative Aging Research Using Technology group highlighted the

XSL•FO
RenderX

scarcity of literature on home-based, real-life evaluations [35,36]. The review also underlined that the samples of extant studies are predominantly homogenous, consisting mainly of Caucasian, urban, educated individuals receptive to technology. In addition, most of the studies used a limited approach (eg, using a single device or technology) and were primarily restricted to participants who lived alone. This proposed study will aim to address these gaps by using various digital biomarker technologies to observe and predict behavioral patterns related to MCI, dementia, PF, and frailty, and differentiate them from those associated with normal aging in older adults living in their homes in Singapore.

This study will focus on using a continuous monitoring sensor system within the homes of older adults (aged 65 years and older) to observe changes in their behavioral patterns over a period of 3 to 6 years. This proposed study extends from a recently completed feasibility study that used a similar home-based sensor setup to monitor behavior patterns in community-dwelling senior citizens in Singapore [26]. The study seeks to distinguish those exhibiting normal aging and those with MCI, early dementia, PF, or frailty and to predict the transition from normal aging to one of these conditions.

To achieve this goal, this study has the following aims:

- To monitor changes in behavior associated with the execution of cognitive-related daily tasks and activities among older adults, and to distinguish between changes attributable to normal aging and those indicative of MCI or early dementia.
- To monitor changes in behavior in performing physical function-related daily tasks and activities among older adults, and to distinguish between changes attributable to normal aging and those indicative of PF or frailty.
- To develop and cross-validate machine learning prediction models derived from clinical data and sensor-based, spatiotemporal, daily activity data that will be used to detect individuals at a high risk of developing MCI, early dementia, PF, or frailty. The performance of these prediction models will also be evaluated against gold standard diagnostic outcomes and cross-validated using subsets of the data.

Methods

Ethical Considerations

This study was approved by the SingHealth Centralised Institutional Review Board and registered with the reference number 2019/2026. Written informed consent will be obtained from all participants before enrollment in the study. The study will be conducted in accordance with the Good Clinical Practice guidelines.

Study Design

This study will use a longitudinal cohort design with 200 Singaporean older adults aged 65 years and older who either reside alone or with others. The follow-up period for the study will span 1 to 6 years, contingent on the continuity of research funding. The daily activities of the participants will be monitored using home-based sensors, in conjunction with annual clinical and cognitive evaluations.

Sample Size Calculation

As described in the previous sections, an earlier study with 14 participants comparing walking time variability between people with MCI and healthy controls using data obtained continuously from sensors revealed a significant difference between the groups [35]. A subsequent study was conducted by the same research group with 54 and 39 people with healthy cognition and with MCI, respectively, over a mean follow-up time of around 2 years [36]; this subsequent study identified 3 distinct trajectories of mean weekly walking speeds.

Based on these findings on walking speeds and data from the Singapore Longitudinal Aging Study Waves 1 and 2 (unpublished), we conservatively estimate an incidence of MCI or dementia of 30% over a 3-year period. Given the feasibility and budgetary constraints of the study, a sample of 200 participants without dementia (expected 60+ incident cases of MCI or dementia over the study period) will be sufficient to produce a 2-sided 85% CI with a precision of 5%. The sample size calculation is based on CIs for one proportion from a finite population approach via the exact Clopper-Pearson formula and performed using PASS software (version 14; NCSS, LLC) [39].

Study Eligibility

Inclusion Criteria

To be included in the study, participants must (1) be at least 65 years old, (2) have normal cognition or MCI, (3) live independently in the community (ie, no difficulties with any of the instrumental activities of daily living (IADL), see below), and (4) be able to give written informed consent.

Exclusion Criteria

Individuals will be excluded from the study if they have a diagnosis of (1) dementia, (2) any other psychiatric disorder, or (3) any condition (eg, frailty, stroke, or Parkinson disease) that results in limitations to perform basic activities of daily living; individuals who are (4) unwilling to adhere to the use of wearables or to have sensors installed in designated areas of their home, or (5) currently participating in a cognitive or motor training intervention trial will be excluded.

Participant Recruitment and Study Procedure

Participant recruitment will be conducted in various community and hospital settings, including existing cohorts such as the Singapore Longitudinal Aging Study (also known as SLAS), SingHealth Memory clinics, partner senior care agencies, silver generation offices, and word-of-mouth referrals (Figure 1).

Eligible participants who meet the inclusion criteria will be contacted via phone or in person at senior activity centers. Detailed information about the study will be provided to potential participants, allowing them sufficient time to consider their participation and discuss it with their family members. Interested individuals will undergo a baseline visit to provide informed consent and complete the screening procedures. The study will be thoroughly explained to potential participants using a participant information sheet before informed consent

XSL•FO
RenderX

is obtained. In cases where a Chinese translation of the participant information sheet is required, a witness will be present to ensure comprehension by the potential participant.

To determine eligibility for participation in the study, trained research personnel will administer various questionnaires and assessments to potential participants. These will include basic demographic and health questionnaires, measures of executive functioning, and measures of cognitive health. These measures will be repeated annually for 3 years.

Upon successful completion of the screening process, eligible participants will be enrolled in the study and undergo further baseline assessments, including assessments of mental health, sleep disturbances, and psychosocial health, which have been associated with cognitive impairment. A brief assessment of medication adherence will also be conducted. All these assessments will be repeated annually. The home-based sensors will be installed after the baseline assessment is completed. Participants may drop out at any point in the study and will be withdrawn if they are diagnosed with dementia.

Figure 1. Study design—Timeline and assessments. CFS: Clinical Frailty Scale; IADL: instrumental activities of daily living; MMSE: Mini-Mental State Examination; MNA-SF: Mini Nutritional Assessment Short Form; MoCA: Montreal Cognitive Assessment. *These assessments are only carried out if the participant's Geriatric Depression Scale score is >5. **The following are included in the neuropsychological assessment battery: (**A**) MoCA, (**B**) MMSE, (**C**) Clinical Dementia Scale (CDR), (**D**) Rey Auditory-Verbal Learning Test (RAVLT), (**E**) Digit Span Test, (**F**) Color Trails Test (CTT), (**G**) Block Design Test, (**H**) Semantic Verbal Fluency (Animals) test, (**I**) Victoria Stroop Test (VST), (**J**) Common Objects Memory Test (COMT).

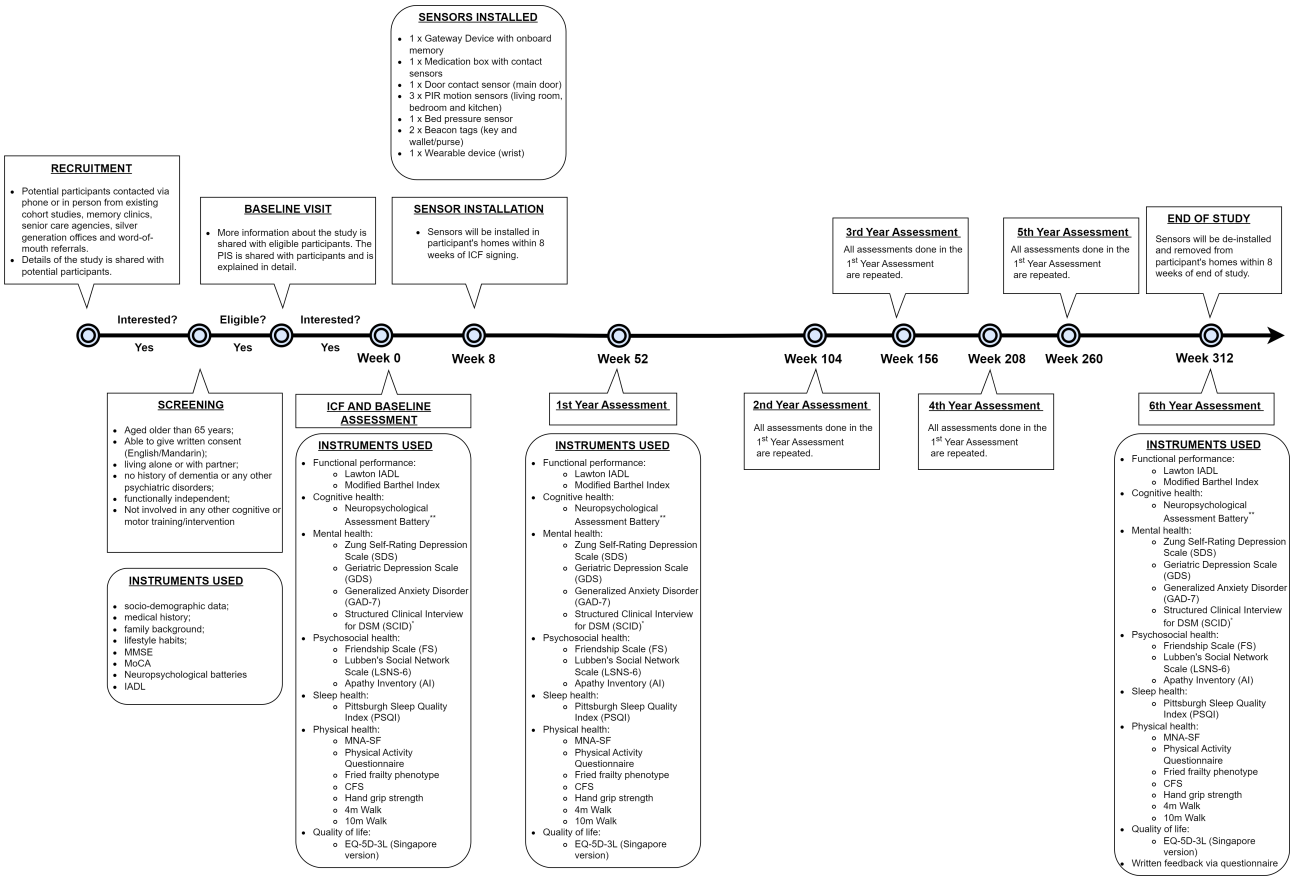

Instruments

Measure of Functional Performance

The Lawton IADL scale will be used to assess independent functional skills. These skills are considered more complex than basic daily activities and are used to identify current functional levels. The scale comprises 8 domains, each of which ranges in score from 0 (low functioning and dependent) to 4 (high functioning and independent), with the total score ranging from 0 to 23. The scale has been used in older adults as a measure of functioning level and is an accurate tool for screening dementia [40,41]. The convergent and divergent validity of the IADL was tested against known strengths and group differences in Singapore. The IADL as a whole explained 87.6% of the variance. The internal consistency for the physical part of the IADL was 0.91, while 0.78 for the cognitive IADL [41].

The modified Barthel Index (MBI) of Activities of Daily Living will be used to assess the basic activities one needs to be independent at home. It has been widely used in various large-scale studies, including studies conducted in Singapore [42,43]. The scale comprises 10 areas: 6 items on feeding, dressing, bladder, bowel, toilet use, and stair climbing are scored using a 3-point scale (0, 1, or 2 points), 2 items on transferring and mobility on a 4-point scale (0, 1, 2, or 3 points), and the last 2 on bathing and grooming on a 2-point scale. The maximum score is thus 20, and the higher the score, the greater the independence [44]. The MBI has been validated against assessments from occupational therapists for older residents of long-term care facilities. The Kappa coefficient of intra-rater reliability was 0.70 - 1.00 for individual tasks; the Kappa coefficients for inter-rater reliability were 0.72 - 1.00. Cronbach coefficient alpha of internal consistency was 0.93 [43].

XSL•FO
RenderX

Measures of Cognitive Health

The Montreal Cognitive Assessment (MoCA) is a brief screening tool for detecting global cognitive changes in patients with MCI and dementia. The effectiveness of the MoCA in detecting MCI and dementia has been validated in Singapore, and it has also been shown to exhibit diagnostic utility in discriminating between mild and major neurocognitive disorders with an area under the curve (AUC) of 0.77 and 0.99, respectively [45].

The Mini-Mental State Examination (MMSE) is the most common screening tool for the early detection of MCI and dementia worldwide. A localized Singapore version can discriminate between older adults with and without dementia, with a high sensitivity of 97.5% [46]. Normative data have also been determined for Chinese older adults and are widely used in Singapore [47]. With a cut-off point of 26, the MMSE has been shown to have a sensitivity and specificity of 61% and 84%, respectively, and a diagnostic accuracy, based on AUC, of 0.78.

The Clinical Dementia Rating Scale is a semi-structured dementia staging instrument that is completed on a 5-point scale used to characterize 6 domains of cognitive and functional performance. The rating for each domain will be used to calculate a global score between 0 (normal) and 3 (severe dementia) and to determine dementia severity. The scale has been shown to provide reliable and valid assessments for community-dwelling older adults even in the absence of an informant, and there was good agreement between the Clinical Dementia Rating Scale and the clinical assessment status of MCI and dementia (internal consistency =0.82 - 0.84; kappa coefficient =0.79) [48].

The Rey Auditory-Verbal Learning Test will be used to evaluate immediate verbal memory, retention of delayed information, learning strategies, proactive and retrospective interference, and delayed recognition. This test has been used to measure immediate and delayed memory performance in older Chinese adults in Singapore [49]. The Rey Auditory-Verbal Learning Test immediate and delayed recall subtests were able to distinguish between converters and nonconverters to dementia with sensitivities of 83% and 85%, respectively, AUCs of 0.75 and 0.76, respectively, and optimal cut-off points of 1 and 2, respectively [50].

The digit span test is part of the Wechsler Adult Intelligence Scale-Fourth Edition and requires participants to read a string of numbers forward (ie, to assess efficiency and capacity for paying attention) and backward (ie, to assess executive function and working memory); its results can show indications of short-term verbal memory. This test has been widely used as a valid tool to discriminate between MCI and dementia among older adults [51], and provided specificities of 99% and 95%, respectively, for MCI and mild dementia, with a cut-off score <4.

The Color Trails Test (CTT) 1 and 2 were designed as culturally appropriate tests in comparison to the Trail Making Test. The CTT 1 requires participants to join 25 numbers in sequence as quickly as possible, whereas CTT 2 requires participants to alternate between colors while joining 25 numbers in sequence. They are used to measure selective attention, processing speed, and cognitive flexibility. Test-retest analysis was performed over a 4-week period to determine the reliability of the CTT, and CTT 1 and CTT 2 showed a Pearson correlation r of 0.90 and 0.91, respectively [52].

The block design test is also part of the Wechsler Adult Intelligence Scale-Fourth Edition, requiring participants to recreate a design shown to them using blocks with different colored sides. The test assesses perceptual reasoning, spatial visualization, and visuo-construction abilities. The block design test has been used for epidemiologic studies among older adults [53], and test performance declines with age across the normative data [54].

The semantic verbal fluency (animals) test will also be administered as a measure of executive function, access to semantic memory, and lexical retrieval. A decline in semantic fluency has been reported as an early indicator of dementia, and it can be used to differentiate the development of dementia from normal age-related cognitive decline [55]. A normative study conducted for the tool showed its reliability, as the number of animals named decreased with an advance in age older than 60 years, and it correlated (r=0.52) with orthographic verbal fluency scores [56].

The Victoria Stroop Test will be used to measure response inhibition and cognitive flexibility [57]. This test is divided into 3 conditions and is composed of 24 items. In the first condition, participants are required to name the colors of the dots. In the second condition, they are required to identify the color of the ink each word is printed in (eg, the word Blue is printed in the blue color), while in the last condition, colored words irrelevant to their printed color are presented. The number of errors and the time taken to complete each condition are recorded and scored. The Victoria Stroop Test is reported to be sensitive to both MCI and dementia [58,59].

Age-related memory impairments were assessed using the Common Objects Memory Test (COMT) [60]. The COMT uses pictures of 10 common objects that participants should be familiar with. It involves 3 learning trials with immediate recall and 2 delayed recall trials, one after 5 minutes and another after 30 minutes of distracted activities. The number of objects correctly recalled is scored. The COMT test is not affected by educational levels, language, or ethnicity, and it could be culturally modified to suit the participants. It has been validated in both Caucasian and Asian older adults with dementia [60], and has been included within the Asian Cohort for Alzheimer Disease Data Collection Packet [61].

Measures of Mental Health

The Zung Self-Rating Depression Scale and the 15-item Geriatric Depression Scale will be used to determine the level of depressive symptoms among participants. The Zung Self-Rating Depression Scale comprises 20 items that measure depressive symptoms, asking respondents how they felt or behaved in the past week using a 4-point scale ranging from 1 (a little of the time) to 4 (most of the time). A high total score indicates worse depressive symptoms. This scale has been

XSL•FO
RenderX

validated among community-dwelling older adults for detecting subsyndromal depression [62]. With cut-off scores of 39 and 40, the scale also showed a sensitivity of 79.2% and an AUC of 0.85 among older adults [63].

The 15-item Geriatric Depression Scale is a shortened form of the original Geriatric Depression Scale and contains 15 questions assessing depressive symptoms over the last 7 days. All 15 items are responded to using a yes-or-no scale, and a score of 5 or above indicates possible depression. This scale showed accuracy in screening for depression among older adults in a meta-analysis, with a pooled sensitivity and specificity of 86% and 76%, respectively, and diagnostic accuracy, based on AUC, of 0.95 [64].

Self-reported anxiety will be captured using the Generalized Anxiety Disorder-7. Participants will respond to 7 questions on how often, over the past 2 weeks, they worried (eg, "Not being able to stop or control worrying," "Worrying too much about different things") and experienced somatic tension (eg, "Trouble relaxing"). A 4-point Likert scale response ("0" denoting not at all, while "3" nearly every day) is used. The maximum score is thus 21, while the minimum score is 0, and scores greater than 10 usually suggest a diagnosis of Generalized Anxiety Disorder. This scale has a sensitivity of 89% and specificity of 82% [65]. It has been validated in both cognitively impaired [66] and general populations [67,68], and has been used in Singapore settings [69,70].

Measure of Sleep Disturbances

The Pittsburgh Sleep Quality Index is used to measure subjective sleep patterns and quality. It includes 4 open-ended questions and 14 frequency- and semantic-scaled questions that assess sleep quality, latency, duration, disturbances, daytime dysfunction, and sleep efficiency. It can be administered verbally to participants, and a high global score of 5 or more indicates poor sleep. Its suitability for use with older adults in Singapore was demonstrated [71]. It showcased an internal consistency of 0.83 and showed a correlation with greater sleepiness on the Epworth Sleepiness Scale (r=0.13) when applied among men [72].

Measures of Psychosocial Health

Social vulnerability, such as a lack of social networks, loneliness, and limited social contact, has been shown to predict cognitive decline among older adults [73]. Apathy symptoms are also associated with the conversion from MCI to dementia and incident dementia among community-dwelling older adults [73]. This highlights the importance of considering social networks and apathy when attempting to understand cognitive decline. As such, the 6-item Friendship Scale, Lubben Social Network Scale, and the Apathy Inventory Clinician Version will be used to measure social isolation and apathy.

The 6-item Friendship Scale measures important dimensions contributing to social isolation and connectedness, requiring participants to describe the frequency with which they experience the situations in each question on a 5-point Likert scale ranging from 1 (almost always) to 5 (not at all). The scale has been shown to exhibit concurrent discriminant validity, suggesting its sensitivity to correlates of social isolation among older adults [74]. It has also been validated for use among older adults with a reliability of 0.83 and an excellent internal structure, as the comparative fit index was 0.99 and the root mean square error of approximation was 0.02 [75].

The Lubben Social Network Scale consists of 2 sets of 3 questions aimed at evaluating kinship and nonkinship ties. Total scores range from 0 to 30, with lower scores indicating lower levels of social support. The scale has been validated in past research, showing a Cronbach α of .74 and a convergent validity of −0.3 [76]. It has been used in Singapore to understand social isolation among community-dwelling older adults [77].

The Apathy Inventory clinician version is a scale operationalized into 3 dimensions, namely emotional blunting, lack of initiative, and lack of interest. Items are responded to on a yes-or-no scale to determine the presence or absence of the behavior in the item; then, a 5-point Likert scale ranging from 0 (no problem) to 4 (major problem) is used to determine the score for each dimension. Total scores range from 0 to 12, with scores of 4 and above indicating the presence of apathy. The scale has been validated in older adults and showed a Cronbach α of 0.90 [78]. It has also been applied to determine chronic behavioral changes in Alzheimer disease [79].

Measures of Physical Health

Physical frailty will be assessed using the modified Fried frailty phenotype, assessing 5 components: low hand grip strength, slow walking speed, low physical activity, self-reported exhaustion, and low BMI (18.5kg/m^2) [80]. Those who screen positive for at least 3 of these criteria are considered frail, while those who meet 1‐2 criteria are classified as prefrail. Participants with a Fried score of 0 are deemed to be robust. Exhaustion will be assessed using the Center for Epidemiological Studies-Depression Scale, which asks how often participants felt the statements "everything I did was an effort" and "I could not get going." An affirmative answer to either or both questions is considered positive for exhaustion. Grip strength will be measured using a hand dynamometer (Jamar Plus+ Digital Hand Dynamometer, USA), with 2 trials for each hand, and the maximum value of 4 trials will be recorded. Gait speed will be assessed using a 10-meter walk at usual pace, with the average of 2 trials recorded. We will adopt the Asian Working Group for Sarcopenia 2019 cut-offs for defining weak hand grip strength and slow gait speed [81]. Physical activity is self-reported using the Physical Activity Vital Sign that quantifies time spent on physical exercise and activities, and values below the lowest quartile from a local community cohort will be deemed physically inactive [82]. The Fried phenotype is the most commonly used of measures for frailty [83]. The frailty phenotype was operationalized using data from the Cardiovascular Health Study, a large-scale study of over 5000 older adults in the United States [80]. It has been used in a number of large-scale longitudinal studies, including one in Singapore, to quantify physical frailty in community-dwelling older adults [84-86].

Participants are also scored on the Clinical Frailty Scale (CFS) for a global clinical assessment of frailty status. CFS is a 9-point scale that ranges from 1 (very fit) to 9 (terminally ill) [87,88].

XSL•FO
RenderX

Specifically, CFS has been highly recommended as the instrument of choice for frailty screening in the community and primary care [89]. The CFS has been validated against the Frailty Index in hospitalized older adults in Singapore with an AUC score of 0.91 (95% CI 0.87 - 0.95; *P*<.001) [90]. A classification tree had been introduced to facilitate the reliable scoring of CFS [91], and this was modified for our local context (Figure 2) [92]. This method uses information gathered from other validated toolkits—specifically the IADL, MBI, and SF-12 (Exhaustion)—to classify physical frailty. This classification tree approach to CFS will provide us with a more nuanced approach to quantifying frailty in older residents living in the community.

Figure 2. Scoring CFS using the modified classification tree [91,92]. BADL: basic activities of daily living; CFS: Clinical Frailty Scale; IADL: instrumental activities of daily living.

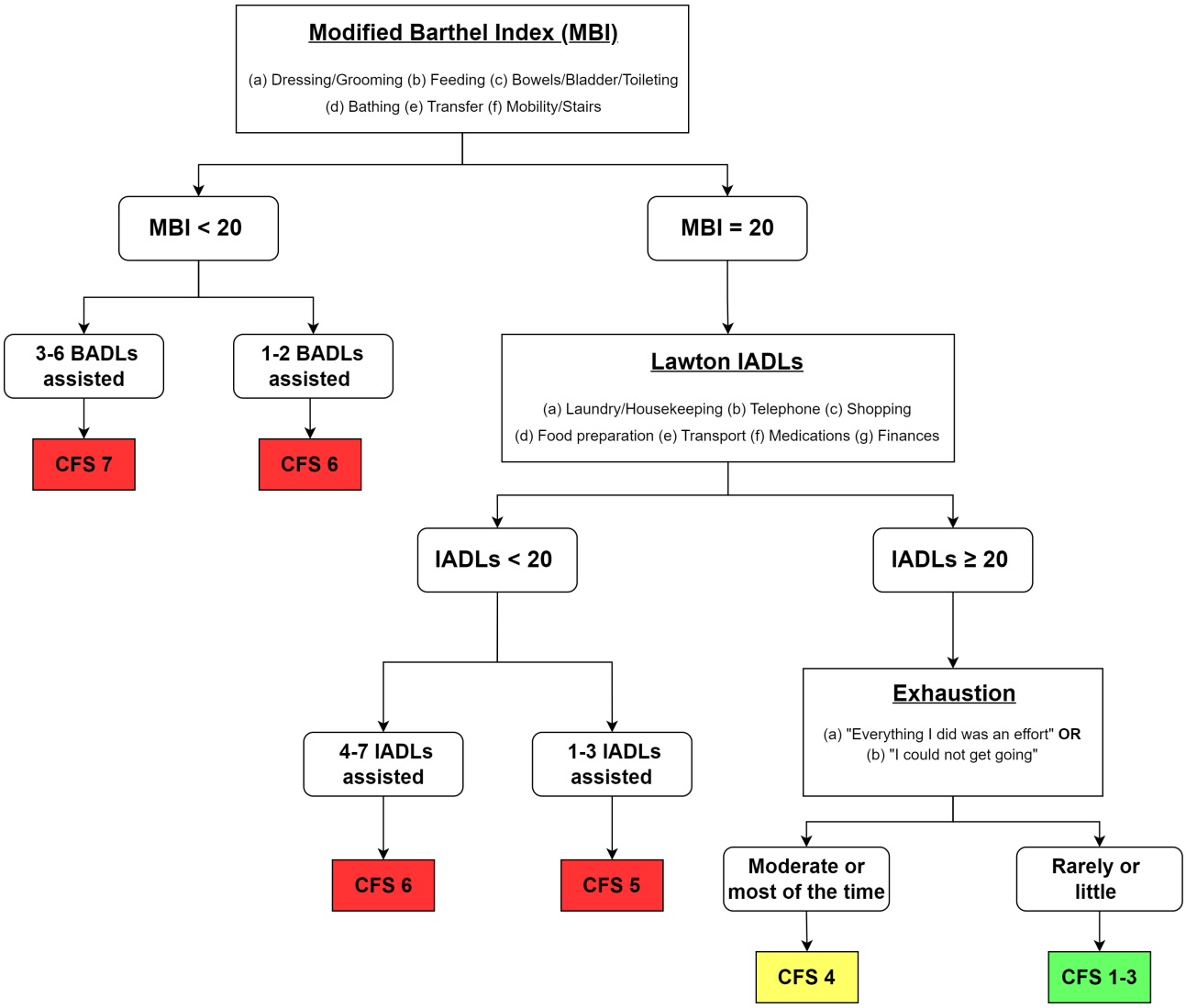

Measures of Quality of Life

Health-related quality of life will be captured using the EQ-5D-3L (Singapore version) questionnaire. This questionnaire has 2 parts: the descriptive system and the visual analog scale (VAS). The descriptive system of the EQ-5D-3L has 5 dimensions: mobility, self-care, usual care pain/discomfort, and anxiety/depression. Each dimension has 3 levels of impairment: no problems (level 1), some problems (level 2), and extreme problems (level 3). An EQ-5D summary index (single utility index), which ranges between −1.00 (worst possible health state) and 1.00 (perfect health state), is derived from the 5-digit health state profile using a weighted method. As for the VAS part, it has a visual scale from 0 to 100, with 0 being the “worst health the participant can imagine” while 100 being the “best health the participant can imagine” [93]. The EQ-5D-3L is widely used around the world and has been translated into more than 170 languages. Additionally, it has been validated among patients with chronic noncommunicable diseases [94-96] as well as persons with cognitive impairments. It has been validated against the Health Utilities Index Mark 3 in dementia patients in Singapore [97]. It has been validated against community-dwelling older adults with cognitive impairment [98]. The EQ-5D-3L was validated against corresponding assessment scales: mobility against the Tinetti Index (correlation: 0.371), self-care against the Barthel Index (0.340), usual activities against the Lawton Index (0.628), pain against the VAS pain scale (0.504), and anxiety against GDS (0.703). Reliability (internal consistency) has been reported as 0.69 (Cronbach α).

XSL•FO
RenderX

The Mini Nutritional Assessment Short Form (MNA-SF) is a 6-item questionnaire that has been used to evaluate the nutritional status of older frail adults [99]. It has been used in large cohort studies involving community-dwelling older adults to investigate malnutrition and frailty [100,101]. The MNA-SF comprises 6 questions that cover anthropometric measurements (BMI and unintentional weight loss), diet changes, neuropsychological problems, acute disease, and mobility. Anthropometric measurements are scored between 0 and 3, while all the others range between 0 and 2. The maximum total score for the MNA-SF is thus 14, and a score of less than 7 indicates malnutrition, 8 to 11 indicates a risk of malnutrition, while a score above 12 indicates normal nutritional status [102]. The MNA-SF has been well-validated in a number of large-scale studies involving community-dwelling older adults and has been included in studies that investigated frailty in the Singapore context [99,103-105]. It has been validated against anthropometric measurements and functional status (IADL) in Malaysian older adults—sensitivity: 93.2%, specificity: 79.4% [106].

The 10-meter Walk Test assesses functional mobility and gait speed. Participants will be instructed to walk at their usual (self-selected) walking pace over a distance of 14 meters—2 meters are added at the beginning and at the end of the test to account for the acceleration and deceleration during the test. Participants will walk 3 times—for the first 2, participants will walk without talking, while for the third walk, which is the dual-task gait speed trial, they will be required to count upwards in multiples of 2, starting from the number 16. A handheld stopwatch and a measuring tape will be used to facilitate this test.

Sensor Setup and Behaviors/Activities Captured

After the baseline assessments are completed, participants will have their homes fitted with a network of sensors (Table 1) for a period of up to 3 to 6 years. Each sensor will regularly monitor the physical environment throughout the day and then wirelessly transmit the collected data to a gateway. The gateway will then send the combined data to the backend server using secure cellular communications for monitoring and processing. A data analytics engine will implement the algorithms for feature extraction of the daily activities of older adults based on the raw sensor data collected by home-based devices. A system monitoring tool will be used to ensure reliable sensor data collection, providing alerts for system downtime. Each data point will only be identifiable through the sensor node identifier, and the mapping between the identifier and location will be securely stored and accessible only to the research project investigators.

The study aims to capture several behaviors of interest using a sensor network. The primary behavior under investigation is forgetfulness, along with in-home activity levels, sleep quality, outings, and physical activity.

Table . Home-based sensors to be deployed.

Type of sensor	Location of sensor	Data collected	Interpretation
Gateway Device (n=1) (eg, Raspberry Pi 4B + Wi-Fi dongle and Z-Wave stick)	Living room	Consolidation/Storage of all data collected and retransmission to secure	N/A[a]
Passive infrared sensor (n=3) (eg, Aeotec motion sensor 5) for single-occupancy homes only	Living room, bedroom, and kitchen	Presence/absence of activities	Activity levels (in-home/out-of-home/at sensor locations)
BLE[b] Receivers (n=3) (eg, ESP32) for multi-occupancy homes only	Living room, bedroom, and kitchen	Presence/absence of activities	Activity levels (in-home/out-of-home/at sensor locations)
Door contact sensor (n=1) (eg, Aeotec door sensor 7)	Main door	Number of opening and closing of main house door - per day, per week, and per month	Duration participant spends time at home and out of home
Medication box with contact sensor (n=1) (eg, Aeotec door sensor 7)	Depends on user's preference	Number of opening and closing of medication box - per day, per week, and per month	Frequency of forgetfulness to take medication
Bed pressure mat + contact sensor (n=1) (eg, Aeotec door sensor 7)	Bedroom	Contact duration	Sleep duration and sleep quality (sleep time, wake time, number of sleep interruptions, and sleep interruption duration)
Beacon tags (n=2) (eg, D15N)	Key and wallet	Locations of key/wallet within the home or out of home	Duration participants spend time at home and out of home. Frequency of forgetfulness to take keys or/and wallet
Wearable sensor (n=1) (eg, Mi Band)	On wrist of participant	Step count and heart rate	Physical activity levels of participant

[a]Not applicable.

[b]BLE: bluetooth low energy.

XSL•FO
RenderX

To measure forgetfulness, a combination of sensor data will be used. Participants will be provided with a sensor-equipped medication box (Figure 3, top-left) to store their prescription medications. Data will be generated each time the box is opened, and this, in combination with expected medication frequency information obtained at baseline, will enable the determination of the frequency of forgetting to take medication at the prescribed timing. Additionally, medication intake timing variation will be trackable.

In-home activity levels and the frequency of outings for individuals living alone will be inferred using door contact sensors (Figure 3, top-middle) and PIR motion sensors (Figure 3, top-right), which will detect the opening and closing of a participant's main door. A bed sensor (Figure 3, bottom-left) based on fiber optics technology will be placed under the participant's mattress to collect data on sleep duration and quality. Participants will be requested to wear a wearable activity band (Figure 3, bottom-middle) to measure heart rate and daily steps. For individuals living with family, the PIR motion sensors will be replaced by 3 signal receivers (such as ESP32) for tracking the location of the individual carrying a wearable or personal beacon. This specific configuration imposes the requirement for additional power points for the 3 USB-powered Bluetooth low energy (BLE) receivers and the participant wearing a smart band or beacon at all times.

Proximity beacons (Figure 3, bottom-right) will be attached to the personal belongings of participants, such as keychains and wallets, allowing for the estimation of the distance between the item and the home gateway. The use of these beacons in conjunction with wearable, PIR motion, and door contact sensors will allow us to determine whether a participant has forgotten to bring these items with them when leaving the home.

All in-home sensors are passive and unobtrusive in nature—they do not require any form of input from the participants. However, future phases of this study could include sensors that might require personalized nudges for the older participants/caregivers. To facilitate this, we will incorporate quick guides and standardize training packages that would be completed by the older participants/caregivers.

To assess frailty, the study will collect data on step count, percentage of sedentary time, percentage of active time, number of outings, movement complexity, and heart rate. Additionally, a new sensor—3DGait (CareCam, Singapore)—will be piloted in this study to capture gait speed during the yearly SPPB assessments. The 3DGait is a Food and Drug Administration–approved innovative software solution that captures human motions using an integrated video acquisition system. It is able to accurately quantify angular, spatial, and temporal parameters related to gait, using a single iPad Pro.

XSL•FO
RenderX

Figure 3. Home-based sensors to be deployed: (A) Contact sensors attached to the medication box. (B) Door contact sensors. (C) Passive infrared sensors. (D) Pressure sensors for beds. (E) Wearable activity band (Xiaomi Mi Band). (F) Proximity sensors on keys and wallets.



Privacy and Qualitative Aspects

None of the home-based sensors will have video-recording capabilities. The motion sensors will detect only the presence or absence of motion, and all other captured activities will be inferred through the use of the various sensors mentioned above. Being mindful of the ongoing conversations, which are indeed much needed, in this field of health monitoring and of concerns with intrusiveness, the researchers will elicit feedback from the participants on their reasons for participating in the study, as well as their experiences with remote monitoring systems.

Outcome Measures and Definitions

The primary outcome measures for this study are the transition of cognitive state (ie, conversion from healthy cognition to MCI or early dementia) and the transition from a robust state to PF or frailty. The secondary outcome measure is global cognitive function deterioration, as indicated by a decline in scores for the MMSE and MoCA, and a composite neuropsychological test battery [107].

MCI and dementia status will be assessed annually through a clinical diagnostic review. An MCI diagnosis will be determined based on the following criteria [108]: (1) self-reported memory and cognitive difficulties or difficulties observed by an informant/clinician; (2) impairment in one or more domains of the MMSE, with a total score ranging from 24 to 27, or a decrease of ≥ 2 points in MMSE scores from baseline; a score 1 to 2 SDs below the age and education-adjusted mean values for at least one neurocognitive domain (ie, attention, memory, executive function, and language or visuospatial abilities), or a decrease of 0.5 SD (vs baseline) during follow-up assessments; (3) a total score of ≥0.5 for the Clinical Dementia Rating Scale; (4) being essentially independent in performing basic activities of daily living; (5) not being diagnosed with dementia.

A dementia diagnosis will require evidence of (1) objective cognitive deficit (ie, a total score for MMSE≤23, or scores for a neurocognitive domain 2 SDs less than the mean values adjusted by age and education); (2) the presence of functional disability (ie, requiring help with at least one basic activity of daily living or a global score for the Clinical Dementia Rating Scale≥1).

Frailty status will be assessed annually using the modified Fried frailty phenotype. Incident PF/frailty will be defined as progression from a state of robustness at baseline to PF/ frailty at follow-up. Progressive frailty will be defined as an incremental Fried score at follow-up relative to baseline.

XSL•FO
RenderX

Additional questionnaires will be performed annually, including the Mini Nutritional Assessment, a Global Physical Activity questionnaire, and a count of the number of near falls/falls over the past year. All assessments, including questionnaires, can be done over separate sessions. Participants have the option to do the assessments at separate sittings if it is tiring for them. However, we will conduct all sessions within a timeframe of 7 days. Additionally, cognitive assessments will be conducted first to avoid physical assessments tiring the participant out.

To mitigate the challenge of high missing and NA rates, daily biomarker values collected from a single participant will be averaged over a weekly or monthly period. The averaging process will produce fewer but higher-quality biomarker records with more reliable values and lower missing/NA rates.

Data Analysis

Sensor Data Processes and Biomarker Extraction

Figure 4 shows the process of extracting digital biomarkers from the raw sensor data. The key steps are elaborated below.

Figure 4. The process of extracting digital biomarkers from the raw sensor data.

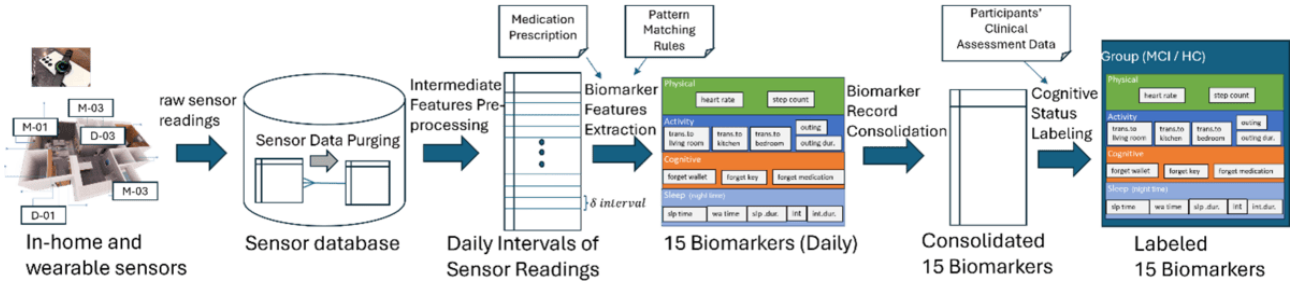

- Biomarker feature extraction: Based on the intervals of sensor readings, each daily biomarker feature is extracted following some predefined pattern-matching rules to detect the participant's state of behavior. For example, to obtain the number of lapses in taking medication (forgetfulness), a rule is used to detect the state that the participant does not take the medication as prescribed based on the absence of readings from the medication box sensor within a range of time intervals in the day.
- Intermediate features preprocessing: Based on the valid sensor data of every participant, each day is divided into intervals of δ minutes (currently, we use δ = 5) in sequential order. The number of sensor readings and their values for every sensor type that occurred within each interval are then recorded.
- Data purging: In this operation, valid raw sensor data are selected based on the periods wherein the in-home sensor gateway system was functional. This step is necessary to filter away unusable sensor data captured when a gateway is down for over a certain period of time.

Daily biomarkers are considered missing when the corresponding behavior states related to the biomarkers cannot be determined. For example, the absence of readings from all the motion sensors during the entire day causes the positional state of the participant to be undetectable, leaving the activity biomarkers of that day missing. Similarly, when the sleeping and awake states cannot be detected because of a prolonged absence of readings from the bed pressure sensor in a day, all sleep biomarkers are considered missing.

On the other hand, biomarker values can be considered as not applicable when the values are not relevant or meaningful under specific circumstances. For example, the sleep time biomarker is considered not applicable when the sleep duration biomarker is 0; forgetting medication is not applicable as well, when the participant does not have any medication prescription.

Statistical Analysis

At baseline, descriptive statistics for demographic information and psychometric scores will be analyzed and presented for all participants. Baseline characteristics will be compared between healthy participants and participants with MCI or PF/frailty using 2 independent sample *t*-tests or Wilcoxon rank sum tests (depending on the normality assumption) for continuous variables, and Pearson chi-square tests or Fisher exact tests (where appropriate) for categorical variables.

Group-based trajectory modeling will be used to classify the study population into distinct subgroups based on the patterns of the individual sensor-derived variables and the coefficient of variations recorded over time. The analysis will be conducted via SAS PROC TRAJ (SAS Inc, Cary, NC, USA), which will be used to perform the analyses, a method described by Jones et al [109] in their seminal work. This procedure integrates 2 statistical models estimated simultaneously via maximum likelihood. The first model develops trajectories for the different latent cohorts as a function of time from the baseline. Quadratic or cubic functions of time may be included if they improve model fitness, as determined by the Bayesian information criterion. The second model is a multinomial regression model, which examines the associations of participants' baseline characteristics with the probability of membership in the homogeneous latent groups. To deal with missing values, complete case analysis is the default method, but SAS PROC TRAJ also allows for the use of multiple imputations to address missingness, where necessary.

Artificial Intelligence Predictive Modeling Methods

To detect individuals at high risk of developing MCI/dementia or PF/frailty, machine learning prediction models will be developed and cross-validated according to the extracted clinical data and spatiotemporal daily activity data (from sensors). Model performance (ie, sensitivity, specificity, and AUC) of these prediction models against gold-standard diagnostic outcomes. The models will undergo rigorous cross-validation (CV) to

XSL•FO
RenderX

ensure robustness, using subsets of the data to prevent overfitting and ensure generalizability. CV will involve partitioning the data into training and testing sets, enabling the models to be trained on one portion while performance is assessed on another. This approach ensures that the models can reliably predict high-risk individuals across different subsets of the population.

We will apply artificial intelligence (AI) methods, namely spatial-temporal neural network models for predictive modeling, using the sensor-derived data and clinical assessment data to classify older adults' behavioral change patterns based on the outcomes assessed at the yearly visits across the 3 participant groups (ie, healthy control, MCI/early dementia, and PF/frailty groups). The goal here will be to build predictive models using the longitudinal data regarding specific daily activities (eg, particularly sleep state, heart rate, medication intake, step count, outings, and room transitions) as learning data for the spatiotemporal neural network models, which will then capture the longitudinal dynamics of the participants' cognitive performance and status.

Two classes of neural network models will be explored and compared, that is, the hierarchical form of self-organizing networks known as fusion adaptive resonance theory (ART) models [110], and deep recurrent neural networks, in particular, long short-term memory [111]. These two neural network models have been shown to provide promising performance in sequence and temporal modeling. For this project, we will formulate the prediction task as a spatial-temporal classification problem, which can typically be addressed via a supervised learning paradigm, wherein sensor-derived data and clinical assessment data collected over a period of time (window) are used as input pattern features, while older adults' cognitive states (ie, considering the healthy control, MCI, and early dementia groups) are used as the output classes (labels). A key feature of fusion ART models is the ability to perform online one-shot learning, which is advantageous for learning from a small number of training samples. This will be critical in our study because of the technical challenge when using machine learning methods, named as "small data problem," which is typically faced when working with user data collected from real-world settings. The performance of the fusion ART and long short-term memory–based models will be benchmarked against the commonly used machine learning models and methods, including k-nearest neighbor, support vector machine, decision tree, random forest, and LightGBM.

AI Model Training and Validation

The AI-based predictive model will be built using data collected from 200 older adults recruited in this project. Then, considering a period of 3 years and using a sliding prediction window of over 3×365 days, this study can yield up to 1095 daily data points for each participant, together with the cognitive and physical frailty status assessments. In total, we expect to have 2,19,000 daily data points across 200 participants in a 3-year period.

Based on the collected daily data, a CV methodology will be adopted to train and evaluate the performance of the predictive model. For example, using 10-fold CV, 90% of the full dataset will be used to train the model, while the remaining 10% will be used to validate the classification accuracy of the model. The evaluation is repeated 10 times so that all data samples are used for validation once across the 10 experiments. Also, CV evaluations will be further performed in both the cross-sectional (ie, across different participants over the entire time period) and temporal holdout manners (ie, for the entire group of participants over different time periods). Specifically, under cross-sectional CV, data samples from 90% of the participants will be used to train the model, while the data of the remaining 10% participants will be used for validation. This allows us to validate the performance of the trained model on new and unseen participants. For temporal holdout CV, the first 90% data collected over time from all participants are used for training the model, and the last 10% data are used for validation. This enables us to validate the performance of the trained model in predicting the cognitive classes of the existing cohort in the future over time. Standard performance measures for classification will be applied to assess the sensitivity, specificity, and AUC of the predictive models against the gold standard diagnostic outcomes.

CV experiments will be conducted using the data collected in the first 2 to 3 years of the study. Data collected in the subsequent stage of the project can be used as an external validation set for evaluating the predictive models developed.

Results

Enrollment for this study started in November 2019 and will continue until March 2030. As of June 2025, we have enrolled 138 participants. Full data analysis has yet to begin.

Discussion

Principal Findings

We anticipate that continuous in-home monitoring using a multi-sensor system will generate behavioral and functional indicators capable of distinguishing normal aging from early cognitive and physical decline. Based on prior evidence, we expect that sensor-derived features—such as gait variability, activity regularity, sleep patterns, and medication-taking behaviors—will differ meaningfully between individuals who remain cognitively and physically stable and those who progress to MCI, early dementia, PF, or frailty. We further hypothesize that supervised machine learning models integrating these sensor-based features with annual clinical assessments will demonstrate strong predictive performance in identifying individuals at elevated risk of transition. These anticipated findings would support the feasibility of unobtrusive, home-based sensing as a scalable approach for early detection and risk stratification in community-dwelling older adults.

This 6-year longitudinal study will be conducted to monitor changes in behavioral patterns among older adults using a home-based sensor system. Results and learning points from this study will be used to develop a dependable and efficient sensor system for in-home use that can detect early signs of cognitive and physical decline. It is common for older adults in Singapore to seek clinical intervention only when their cognitive impairment has reached an advanced stage.

XSL•FO
RenderX

Considering this situation, the implementation of a readily deployable sensor system in community settings presents an opportunity for prompt intervention and holds potential to help delay or even reverse disease progression. The data collected by the sensor system can also be valuably used in supplementing the data from neurocognitive and physical assessments, affording clinicians more reliable information for making a diagnosis. The proposed AI modeling method has several advantages over the methods currently found in the literature. First, the selected behavioral data will be measured using nonintrusive techniques and will be used for building predictive models. While previous studies typically made use of structural neuroimaging or electroencephalogram measurements, our behavioral markers will be collected using simple sensors/wearables and, hence, will be innovative, unobtrusive, and have good potential for upscaling. To our knowledge, this will be a first-of-its-kind study, using longitudinal, sensor-derived behavioral data from people in the community for building predictive models. Second, most prior studies focused on recognizing the cognitive states of participants at a single time point, whereas this study will aim to attempt the more challenging task of predicting the cognitive trajectories of participants over a longer period. If successful, this could contribute to our current understanding of prognostic markers.

This study is planned to evolve into a larger cohort study incorporating serum and neuroimaging biomarkers. We envision that an increasing number of homes will be equipped with clinically studied smart sensors capable of continuously monitoring various health indices related to aging, thereby enabling older adults to enjoy more functional and meaningful lives.

Anticipated Challenges

Privacy

The home-based sensor system that will be used in this study does not have voice or video-recording capabilities. Specifically, PIR sensors can only detect the presence or absence of motion, while the beacon tags lose their connection to the gateway device once they are out of their homes; thus, they do not have tracking capability.

In an earlier feasibility study, there were concerns that the intrusiveness of the continuous monitoring procedures could be an obstacle to recruitment [26]. However, the older adults who lived alone in this cited study expressed enthusiasm to have the system developed to assist them with cognitive monitoring and did not express concerns about intrusiveness, as long as there were no video cameras. Moreover, this study proposal has measures to safeguard data confidentiality. Regardless of these efforts, we will gather qualitative feedback from participants upon study conclusion to better understand the usability and areas for improvement regarding the monitoring system.

Maintenance

Since it is anticipated that the sensors need to run smoothly and continuously (24-7) over an extended period of time, we expect that there would be a need to mitigate hardware failures (eg, issues with battery lifespan, faulty hardware), an online dashboard system will be used to monitor the status of all deployed sensors and to promptly alert our data engineers for attention and necessary actions.

Maintenance visits will be scheduled periodically (every 3 - 4 months) for battery replacement, when the participant calls to inform about faulty devices, or when the study team determines that the sensors require servicing. Maintenance visits will be arranged as soon as the participant's schedule permits.

The research team will also provide participants with regular reimbursements to mitigate the electricity costs of the gateway device, which needs to be plugged into an electrical outlet.

System Complexity Versus Adherence

Our current sensor configuration offers an elaborate setup, consisting of a home gateway, motion sensors, door contact sensors, medication box, a bed pressure mat, BLE beacons, and wearables. As it may be too complex for low-tech literacy homes to handle and create many failure points over the years, a tiered deployment scheme (such as Base: Gateway+ PIR+door; Plus: Base+ beacons+medbox; and Full: Plus+ bed+wearable) could be considered, whereby the seniors may elect to participate with a preferred level of deployment. Augmented with criteria to upgrade/downgrade per household based on early adherence and living context, this tiered deployment will help to reduce dropouts, wastage in hardware, and effort of deployment/maintenance, while preserving data quality.

Wearable Adherence and Contingencies

Long-term wristband use is typically the weakest link of in-home sensing due to various issues such as skin irritation, forgetting, and charging. Guidelines will be provided to the participants regarding predefined minimum wear time KPIs (eg,≥10 h/d on ≥70% of day). We will also provide accessory options, such as spare straps and hypoallergenic bands, and institute regular reminder routines to improve the data quality. Critically, our data processing and AI predictive models will also be designed to tolerate missing wearable-related biomarkers and are still capable of learning a model from readings of other sensors.

Medication Sensing Validity

Adherence to consistently taking medications from the designated box is another challenge, as many elders may decant pills into weekly organizers or multiple storage sites. Possible measures to mitigate such risk include briefings to participants and obtaining user acknowledgment during deployment regarding proper use of medication boxes, running calibration against prescriptions, and recording of "not applicable" days operationally.

Beacon Practicality and False Positives

The use of key/wallet beacons assumes those items are always attached to the items and carried by the older adults. To mitigate the risk, such as adhesive failures and missing beacons, spare beacons and robust attachments will be provided. In addition, a monthly self-check can be instituted to remind elders of the importance of bringing beacons with them outdoors. To reduce false alarms, the "forgetting" features will be triangulated with the door/PIR patterns and medication timing variability.

XSL•FO
RenderX

Multioccupancy Homes and Visitors

Our target expanded cohort will include people both living alone and with others. As the PIR and door sensors are unable to attribute movement to individuals, additional BLE receivers will be installed in place of the PIR sensors to monitor the whereabouts of the specific participants.

Data Completeness and Feasibility Endpoints

For measuring data completeness and quality, various operational feasibility outcomes can be tracked and reported quarterly, such as sensor uptime, proportion of analyzable days per participant, mean devices online/day, time-to-alert-triggered maintenance visit, and device-related complaint rates. These metrics will help us assess real-world scalability beyond a research setting.

Validation Splits and Label Sparsity

With ≈200 participants but up to 219,000 daily rows over 3 years, cross-validation must avoid leakage. We will replace the generic 10-fold CV with participant-level grouped CV and temporal holdouts (train on early windows, test on later windows). In addition, we will also consider quarterly intermediate labels derived from cognitive trajectories to address annual label sparsity for temporal models.

Safety Monitoring and Privacy

To ensure the safety of the older adult participants who live alone, we have incorporated algorithms within our dashboard system to alert the research team if the sensors system does not detect any expected movement (when it was expected, ie, the algorithm "identifies that participant is at home") for a total of 24 h. Such cases are first checked against the participants' travel schedule and indoor/outdoor status based on the prior sensor readings. The verified alert cases will be elevated to the research team who will then contact the participant, followed by caregivers/next of kin and community partners to ensure that the participant is well and safe.

Data Management

Each participant will be assigned a unique participant identifier number upon successful enrollment. Once assigned, the participant will be followed up anonymously using this unique number. Only the principal investigator and the assigned study team member will have access to the participant data link to the identifiers. All consent forms and hardcopy data collection forms will be stored under lock and key, and in areas of the hospital with restricted access. Data will only be accessed by the principal investigator and the assigned study team member.

All research data will be managed in accordance with Good Clinical Practice standards. Study data will be entered into a secure institutional electronic system that incorporates controlled access and routine backups. Only authorized study personnel will have access according to predefined roles. Administrative files (eg, enrolment logs or scheduling sheets) will be maintained in password-protected spreadsheets stored on an encrypted, access-restricted institutional server. These documents are operational in nature and do not constitute the study's primary research database.

Three levels of data monitoring will be used to secure data accuracy. At the first level, research assistants will ensure that the data entry and scoring on the hardcopy forms are accurate, and they will also ensure accuracy in the transfer of the hardcopy data to the electronic datasets. At the second level, another research staff member will go through the hardcopy forms and electronic datasets to ensure that the data comply with each other.

Protocol Deviations

Serious adverse events and unanticipated problems will be reported within 24 hours to the SingHealth Centralized Institutional Review Board from the day the principal investigator becomes aware of the issue.

Dissemination

All findings from this study will be communicated to the scientific community via peer-reviewed publications and conferences. Key findings will be shared with the general public via community workshops and media interviews.

Acknowledgments

The study team would like to acknowledge Allkin Active Aging Center, Anglican Active Aging Center, Asian Women's Welfare Association (Singapore), Chinatown Active Aging Center, Fei Yue Active Aging Center, Fernvale Community Center, Lions Befrienders, Methodist Welfare Services, NTUC Health, People's Association, Precious Active Aging Center, Presbyterian Community Services, Silver Generation Office, Singapore Longitudinal Aging Studies (SLAS) study, Thye Hua Kwan Moral Charities, and the Wellness Kampung for their kind support. We would also like to acknowledge the valuable support from Mr Andy Chan Yong Ann, Ms Josephine Lee Sze Ching, Ms Audrey Ling, Ms Savannah Siew, Dr Budhitama Subagdja, and Mr Tran Quang Thanh.

Funding

This research was supported by the Sengkang General Hospital, Singhealth Foundation, the National Medical Research Council Transition Award (TA22jul-0012), the NMRC Centre Grant Programme (CG21APR3007), the SMU-A*STAR Joint Lab in Social and Human-Centered Computing (grant number SAJL-2022-HAS001), and the Lee Kong Chian Professorship awarded by Singapore Management University. The sponsors had no role in the design of the study, nor will they have any in its conduct;

XSL•FO
RenderX

they will also play no role in the collection, analysis, and interpretation of data; preparation of the manuscript; or review or approval of the manuscript.

Data Availability

Data management and any sharing of study data will adhere to the NMRC Research Data Governance and Sharing Framework, institutional data-management policies, and all applicable legislation governing personal and health information. In accordance with NMRC requirements, anonymized Final Research Data and associated metadata will be deposited in the NMRC Research Data Repository. Access to this data by external researchers will be governed by NMRC's controlled access processes, including submission of a data-request application and review by the Data Contributing Institution and relevant oversight committees. No additional public data release is planned at this stage.

Authors' Contributions

IR and AHT developed the study concept and designed the protocol framework. JPM, TCIT, IR, RM, LT, TPN, KAJ, YWY, TST, LTS, TSM, WK, and AHT contributed to the development of study methods. TCIT and IR secured ethical approvals. IR obtained institutional support and funding to support the study. IR and AHT provided overall guidance and oversight of the protocol development. JPM, IR, TCIT, KAJ, and AHT prepared the initial manuscript draft. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

None declared.

Checklist 1
STROBE Checklist.
[PDF File, 149 KB - resprot_v15i1e79490_app1.pdf]

References

1. World population prospects 2024. United Nations Department of Economic and Social Affairs (DESA) Population Division. 2019. URL: https://population.un.org/wpp/Download/Standard/Population [accessed 2024-09-23]
2. Qiu Y, Li G, Wang X, et al. Prevalence of cognitive frailty among community-dwelling older adults: a systematic review and meta-analysis. Int J Nurs Stud 2022 Jan;125:104112. [doi: 10.1016/j.ijnurstu.2021.104112]
3. Ofori-Asenso R, Chin KL, Mazidi M, et al. Global incidence of frailty and prefrailty among community-dwelling older adults. JAMA Netw Open 2019 Aug 2;2(8):e198398. [doi: 10.1001/jamanetworkopen.2019.8398]
4. Feng L, Zin Nyunt MS, Gao Q, Feng L, Yap KB, Ng TP. Cognitive frailty and adverse health outcomes: findings from the Singapore longitudinal ageing studies (SLAS). J Am Med Dir Assoc 2017 Mar 1;18(3):252-258. [doi: 10.1016/j.jamda.2016.09.015] [Medline: 27838339]
5. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc 2012 Aug;60(8):1487-1492. [doi: 10.1111/j.1532-5415.2012.04054.x] [Medline: 22881367]
6. Allegri RF, Glaser FB, Taragano FE, Buschke H. Mild cognitive impairment: believe it or not? Int Rev Psychiatry 2008 Aug;20(4):357-363. [doi: 10.1080/09540260802095099] [Medline: 18925484]
7. Petersen RC, Negash S. Mild cognitive impairment: an overview. CNS Spectr 2008 Jan;13(1):45-53. [doi: 10.1017/s1092852900016151] [Medline: 18204414]
8. Petersen RC, Lopez O, Armstrong MJ, et al. Author response: practice guideline update summary: mild cognitive impairment: report of the guideline development, dissemination, and implementation subcommittee of the American academy of neurology. Neurology (ECronicon) 2018 Aug 21;91(8):373-374. [doi: 10.1212/WNL.0000000000006042]
9. Jia J, Zhou A, Wei C, et al. The prevalence of mild cognitive impairment and its etiological subtypes in elderly Chinese. Alzheimers Dement 2014 Jul;10(4):439-447. [doi: 10.1016/j.jalz.2013.09.008] [Medline: 24418053]
10. Sachdev PS, Lipnicki DM, Kochan NA, et al. The prevalence of mild cognitive impairment in diverse geographical and ethnocultural regions: the COSMIC collaboration. PLoS ONE 2015;10(11):e0142388. [doi: 10.1371/journal.pone.0142388]
11. Lonie JA, Parra-Rodriguez MA, Tierney KM, et al. Predicting outcome in mild cognitive impairment: 4-year follow-up study. Br J Psychiatry 2010 Aug;197(2):135-140. [doi: 10.1192/bjp.bp.110.077958] [Medline: 20679266]
12. Ward A, Tardiff S, Dye C, Arrighi HM. Rate of conversion from prodromal Alzheimer's disease to Alzheimer's dementia: a systematic review of the literature. Dement Geriatr Cogn Dis Extra 2013;3(1):320-332. [doi: 10.1159/000354370] [Medline: 24174927]
13. Lim WS, Wong CH, Ding YY, Rockwood K, Lien C. Translating the science of frailty in Singapore: results from the national frailty consensus discussion. Ann Acad Med Singap 2019 Jan;48(1):25-31. [Medline: 30788491]

XSL•FO
RenderX

14. Liu X, Ng DM, Seah JT, Munro YL, Wee SL. Update on interventions to prevent or reduce frailty in community-dwelling older adults: a scoping review and community translation. Curr Geri Rep 2019 Jun;8(2):72-86. [doi: 10.1007/s13670-019-0277-1]
15. Travers J, Romero-Ortuno R, Bailey J, Cooney MT. Delaying and reversing frailty: a systematic review of primary care interventions. Br J Gen Pract 2019 Jan;69(678):e61-e69. [doi: 10.3399/bjgp18X700241] [Medline: 30510094]
16. Shimada H, Doi T, Lee S, Makizako H. Reversible predictors of reversion from mild cognitive impairment to normal cognition: a 4-year longitudinal study. Alz Res Therapy 2019 Dec;11(1):1-9. [doi: 10.1186/s13195-019-0480-5]
17. Murukesu RR, Singh DKA, Shahar S, Subramaniam P. A multi-domain intervention protocol for the potential reversal of cognitive frailty: "WE-RISE" randomized controlled trial. Front Public Health 2020;8:471. [doi: 10.3389/fpubh.2020.00471] [Medline: 33014971]
18. Ngandu T, Lehtisalo J, Solomon A, et al. A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): a randomised controlled trial. The Lancet 2015 Jun;385(9984):2255-2263. [doi: 10.1016/S0140-6736(15)60461-5]
19. Williamson JD, Pajewski NM, Auchus AP, et al. Effect of intensive vs standard blood pressure control on probable dementia. JAMA 2019 Feb 12;321(6):553. [doi: 10.1001/jama.2018.21442]
20. Blumenthal JA, Smith PJ, Mabe S, et al. Lifestyle and neurocognition in older adults with cognitive impairments. Neurology (ECronicon) 2019 Jan 15;92(3):e212-e223. [doi: 10.1212/WNL.0000000000006784]
21. Blumenthal JA, Smith PJ, Mabe S, et al. Longer term effects of diet and exercise on neurocognition: 1-year follow-up of the ENLIGHTEN trial. J Am Geriatr Soc 2020 Mar;68(3):559-568. [doi: 10.1111/jgs.16252] [Medline: 31755550]
22. Maharani A, Dawes P, Nazroo J, Tampubolon G, Pendleton N, SENSE-Cog WP1 group. Longitudinal relationship between hearing aid use and cognitive function in older Americans. J Am Geriatr Soc 2018 Jul;66(6):1130-1136. [doi: 10.1111/jgs.15363] [Medline: 29637544]
23. Bucholc M, McClean PL, Bauermeister S, et al. Association of the use of hearing aids with the conversion from mild cognitive impairment to dementia and progression of dementia: a longitudinal retrospective study. A&D Transl Res & Clin Interv 2021 Jan;7(1):e12122 [FREE Full text] [doi: 10.1002/trc2.12122]
24. Piau A, Steinmeyer Z, Cesari M, et al. Intrinsic capacitiy monitoring by digital biomarkers in integrated care for older people (ICOPE). J Frailty Aging 2021;10(2):132-138. [doi: 10.14283/jfa.2020.51] [Medline: 33575701]
25. Alharbi K, van Marwijk H, Reeves D, Blakeman T. Identification and management of frailty in English primary care: a qualitative study of national policy. BJGP Open 2020;4(1):bjgpopen20X101019. [doi: 10.3399/bjgpopen20X101019] [Medline: 32184213]
26. Rawtaer I, Mahendran R, Kua EH, et al. Early detection of mild cognitive impairment with in-home sensors to monitor behavior patterns in community-dwelling senior citizens in Singapore: cross-sectional feasibility study. J Med Internet Res 2020 May 5;22(5):e16854. [doi: 10.2196/16854] [Medline: 32369031]
27. Kua EH, Rathi M. Mental health care in Singapore: current and future challenges. Taiwan J Psychiatry 2019;33(1):6. [doi: 10.4103/TPSY.TPSY_2_19]
28. Bradford A, Kunik ME, Schulz P, Williams SP, Singh H. Missed and delayed diagnosis of dementia in primary care: prevalence and contributing factors. Alzheimer Dis Assoc Disord 2009;23(4):306-314. [doi: 10.1097/WAD.0b013e3181a6bebc] [Medline: 19568149]
29. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. J Med Internet Res 2019 Aug 30;21(8):e12785. [doi: 10.2196/12785] [Medline: 31471958]
30. Reeder B, Meyer E, Lazar A, Chaudhuri S, Thompson HJ, Demiris G. Framing the evidence for health smart homes and home-based consumer health technologies as a public health intervention for independent aging: a systematic review. Int J Med Inform 2013 Jul;82(7):565-579. [doi: 10.1016/j.ijmedinf.2013.03.007] [Medline: 23639263]
31. Taraldsen K, Chastin SFM, Riphagen II, Vereijken B, Helbostad JL. Physical activity monitoring by use of accelerometer-based body-worn sensors in older adults: a systematic literature review of current knowledge and applications. Maturitas 2012 Jan;71(1):13-19. [doi: 10.1016/j.maturitas.2011.11.003]
32. Mc Ardle R, Jabbar KA, Del Din S, et al. Using digital technology to quantify habitual physical activity in community dwellers with cognitive impairment: systematic review. J Med Internet Res 2023 May 18;25:e44352. [doi: 10.2196/44352] [Medline: 37200065]
33. Urwyler P, Stucki R, Rampa L, Müri R, Mosimann UP, Nef T. Cognitive impairment categorized in community-dwelling older adults with and without dementia using in-home sensors that recognise activities of daily living. Sci Rep 2017 Feb 8;7(1):42084. [doi: 10.1038/srep42084] [Medline: 28176828]
34. Chen R, Jankovic F, Marinsek N, et al. Developing measures of cognitive impairment in the real world from consumer-grade multimodal sensor streams. 2019 Jul 25 Presented at: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining; Aug 4-8, 2019. [doi: 10.1145/3292500.3330690] [Medline: 40198015]
35. Hayes TL, Abendroth F, Adami A, Pavel M, Zitzelberger TA, Kaye JA. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. Alzheimers Dement 2008 Nov;4(6):395-405. [doi: 10.1016/j.jalz.2008.07.004] [Medline: 19012864]

XSL•FO
RenderX

36. Dodge HH, Mattek NC, Austin D, Hayes TL, Kaye JA. In-home walking speeds and variability trajectories associated with mild cognitive impairment. Neurology (ECronicon) 2012 Jun 12;78(24):1946-1952. [doi: 10.1212/WNL.0b013e318259e1de] [Medline: 22689734]
37. Dorociak KE, Mattek N, Ferguson JE, et al. Subtle changes in medication-taking are associated with incident mild cognitive impairment. Alzheimer Dis Assoc Disord 2021;35(3):237-243. [doi: 10.1097/WAD.0000000000000439] [Medline: 33538492]
38. Teh SK, Rawtaer I, Tan HP. Predictive accuracy of digital biomarker technologies for detection of mild cognitive impairment and pre-frailty amongst older adults: a systematic review and meta-analysis. IEEE J Biomed Health Inform 2022 Aug;26(8):3638-3648. [doi: 10.1109/JBHI.2022.3185798] [Medline: 35737623]
39. Machin D, Campbell MJ, Tan SB, Tan SH. Sample Size Tables for Clinical Studies: John Wiley & Sons; 2011.
40. Yemm H, Robinson DL, Paddick SM, et al. Instrumental activities of daily living scales to detect cognitive impairment and dementia in low- and middle-income countries: a systematic review. J Alzheimers Dis 2021;83(1):451-474. [doi: 10.3233/JAD-210532] [Medline: 34334407]
41. Ng TP, Niti M, Chiam PC, Kua EH. Physical and cognitive domains of the instrumental activities of daily living: validation in a multiethnic population of Asian older adults. J Gerontol A Biol Sci Med Sci 2006 Jul;61(7):726-735. [doi: 10.1093/gerona/61.7.726] [Medline: 16870636]
42. Htun HL, Wong LH, Lian W, et al. Functional improvement after inpatient rehabilitation in community hospitals following acute hospital care. Ann Acad Med Singap 2022 Jun;51(6):357-369. [doi: 10.47102/annals-acadmedsg.2021507] [Medline: 35786756]
43. Ohura T, Higashi T, Ishizaki T, Nakayama T. Assessment of the validity and internal consistency of a performance evaluation tool based on the Japanese version of the modified barthel index for elderly people living at home. J Phys Ther Sci 2014 Dec;26(12):1971-1974. [doi: 10.1589/jpts.26.1971] [Medline: 25540510]
44. Collin C, Wade DT, Davies S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988 Jan;10(2):61-63. [doi: 10.3109/09638288809164103]
45. Liew TM, Feng L, Gao Q, Ng TP, Yap P. Diagnostic utility of Montreal cognitive assessment in the fifth edition of diagnostic and statistical manual of mental disorders: major and mild neurocognitive disorders. J Am Med Dir Assoc 2015 Feb;16(2):144-148. [doi: 10.1016/j.jamda.2014.07.021] [Medline: 25282632]
46. Ng TP, Niti M, Chiam PC, Kua EH. Ethnic and educational differences in cognitive test performance on mini-mental state examination in Asians. Am J Geriatr Psychiatry 2007 Feb;15(2):130-139. [doi: 10.1097/01.JGP.0000235710.17450.9a] [Medline: 17272733]
47. Feng L, Chong MS, Lim WS, Ng TP. The modified mini-mental state examination test: normative data for Singapore Chinese older adults and its performance in detecting early cognitive impairment. Singapore Med J 2012 Jul;53(7):458-462. [Medline: 22815014]
48. Nyunt MSZ, Chong MS, Lim WS, Lee TS, Yap P, Ng TP. Reliability and validity of the clinical dementia rating for community-living elderly subjects without an informant. Dement Geriatr Cogn Disord Extra 2013;3(1):407-416. [doi: 10.1159/000355122]
49. Ng TP, Nyunt MSZ, Gao Q, Gwee X, Chua DQL, Yap KB. Curcumin-rich curry consumption and neurocognitive function from 4.5-year follow-up of community-dwelling older adults (Singapore longitudinal ageing study). Nutrients 2022;14(6):1189. [doi: 10.3390/nu14061189]
50. Xu Y, Chen K, Zhao Q, Li F, Guo Q. Short-term delayed recall of auditory verbal learning test provides equivalent value to long-term delayed recall in predicting MCI clinical outcomes: a longitudinal follow-up study. Appl Neuropsychol Adult 2020;27(1):73-81. [doi: 10.1080/23279095.2018.1481067] [Medline: 30470140]
51. Clark H, Martin P, Schroeder R. Digit span forward as a performance validity test in dementia evaluations: specificity in mild cognitive impairment, mild dementia, and moderate dementia. Arch Clin Neuropsychol 2019 Aug 28;34(6):837-837. [doi: 10.1093/arclin/acz035.05]
52. Tavakoli M, Barekatain M, Emsaki G. An Iranian normative sample of the color trails test. Psychol Neurosci 2015;8(1):75-81. [doi: 10.1037/h0100351]
53. Lee RZY, Yu J, Rawtaer I, et al. CHI study: protocol for an observational cohort study on ageing and mental health in community-dwelling older adults. BMJ Open 2020 May;10(5):e035003. [doi: 10.1136/bmjopen-2019-035003]
54. Joung H, Yi D, Ahn H, et al. Normative study of the block design test for adults aged 55 years and older in Korean aging population. Psychiatry Investig 2021 Jun;18(6):539-544. [doi: 10.30773/pi.2020.0408]
55. Almeida VN, Radanovic M. Semantic processing and neurobiology in Alzheimer's disease and mild cognitive impairment. Neuropsychologia 2022 Sep;174:108337. [doi: 10.1016/j.neuropsychologia.2022.108337]
56. Tombaugh TN, Kozak J, Rees L. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. Arch Clin Neuropsychol 1999 Feb;14(2):167-177. [Medline: 14590600]
57. Strauss C, Sherman EMS, Spreen O. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary: Taylor & Francis; 2006. [doi: 10.1080/09084280701280502]

XSL•FO
RenderX

58. Joubert S, Brambati SM, Ansado J, et al. The cognitive and neural expression of semantic memory impairment in mild cognitive impairment and early Alzheimer's disease. Neuropsychologia 2010 Mar;48(4):978-988. [doi: 10.1016/j.neuropsychologia.2009.11.019]
59. Nathan J, Wilkinson D, Stammers S, Low JL. The role of tests of frontal executive function in the detection of mild dementia. Int J Geriatr Psychiatry 2001 Jan;16(1):18-26. [doi: 10.1002/1099-1166(200101)16:1<18::aid-gps265>3.0.co;2-w] [Medline: 11180481]
60. Kempler D, Teng EL, Taussig M, Dick MB. The common objects memory test (COMT): a simple test with cross-cultural applicability. J Int Neuropsychol Soc 2010 May;16(3):537-545. [doi: 10.1017/S1355617710000160] [Medline: 20188016]
61. Chow TW, Reyes‐Dumeyer D, Gu Y, et al. Asian cohort for Alzheimer's disease (ACAD) data collection: rationale, design and validation. Alzheimers Dement 2021 Dec;17(S10):e052679. [doi: 10.1002/alz.052679]
62. Rawtaer I, Mahendran R, Yu J, Fam J, Feng L, Kua EH. Psychosocial interventions with art, music, Tai Chi and mindfulness for subsyndromal depression and anxiety in older adults: a naturalistic study in Singapore. Asia Pac Psychiatry 2015 Sep;7(3):240-250. [doi: 10.1111/appy.12201] [Medline: 26178378]
63. Jokelainen J, Timonen M, Keinänen-Kiukaanniemi S, Härkönen P, Jurvelin H, Suija K. Validation of the Zung self-rating depression scale (SDS) in older adults. Scand J Prim Health Care 2019 Sep;37(3):353-357. [doi: 10.1080/02813432.2019.1639923] [Medline: 31286810]
64. Krishnamoorthy Y, Rajaa S, Rehman T. Diagnostic accuracy of various forms of geriatric depression scale for screening of depression among older adults: Systematic review and meta-analysis. Arch Gerontol Geriatr 2020;87:104002. [doi: 10.1016/j.archger.2019.104002] [Medline: 31881393]
65. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
66. Kertz S, Bigda-Peyton J, Bjorgvinsson T. Validity of the generalized anxiety disorder‐7 scale in an acute psychiatric sample. Clin Psychol Psychother 2013;20(5):456-464. [doi: 10.1002/cpp.1802] [Medline: 22593009]
67. Wild B, Eckl A, Herzog W, et al. Assessing generalized anxiety disorder in elderly people using the GAD-7 and GAD-2 scales: results of a validation study. Am J Geriatr Psychiatry 2014 Oct;22(10):1029-1038. [doi: 10.1016/j.jagp.2013.01.076] [Medline: 23768681]
68. Löwe B, Decker O, Müller S, et al. Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. Med Care 2008 Mar;46(3):266-274. [doi: 10.1097/MLR.0b013e318160d093] [Medline: 18388841]
69. Ganasarajah S, Sundström Poromaa I, Thu WPP, et al. Objective measures of physical performance associated with depression and/or anxiety in midlife Singaporean women. Menopause 2019;26(9):1045-1051. [doi: 10.1097/GME.0000000000001355]
70. Chua YCE, Lin YC, Lew JK, et al. Prevalence and risk factors of depression and anxiety in primary care. Ann Acad Med Singap 2024 May 28;53(5):293-305. [doi: 10.47102/annals-acadmedsg.2023195]
71. Chiang GSH, Sim BLH, Lee JJM, Quah JHM. Determinants of poor sleep quality in elderly patients with diabetes mellitus, hyperlipidemia and hypertension in Singapore. Prim Health Care Res Dev 2018 Nov;19(6):610-615. [doi: 10.1017/S146342361800018X]
72. Spira AP, Beaudreau SA, Stone KL, et al. Reliability and validity of the Pittsburgh sleep quality index and the Epworth sleepiness scale in older men. J Gerontol A Biol Sci Med Sci 2012 Apr;67(4):433-439. [doi: 10.1093/gerona/glr172] [Medline: 21934125]
73. van Dalen JW, Van Wanrooij LL, Moll van Charante EP, Richard E, van Gool WA. Apathy is associated with incident dementia in community-dwelling older people. Neurology (ECronicon) 2018 Jan 2;90(1):e82-e89. [doi: 10.1212/WNL.0000000000004767]
74. Hawthorne G. Perceived social isolation in a community sample: its prevalence and correlates with aspects of peoples' lives. Soc Psychiatry Psychiatr Epidemiol 2008 Feb;43(2):140-150. [doi: 10.1007/s00127-007-0279-8] [Medline: 17994175]
75. Hawthorne G. Measuring social isolation in older adults: development and initial validation of the friendship scale. Soc Indic Res 2006 Jul;77(3):521-548. [doi: 10.1007/s11205-005-7746-y]
76. Buckley TD, Becker TD, Burnette D. Validation of the abbreviated Lubben social network scale (LSNS‐6) and its association with self‐rated health amongst older adults in Puerto Rico. Health Social Care Comm 2022 Nov;30(6):e5527-e5538 [FREE Full text] [doi: 10.1111/hsc.13977]
77. Merchant RA, Liu SG, Lim JY, Fu X, Chan YH. Factors associated with social isolation in community-dwelling older adults: a cross-sectional study. Qual Life Res 2020 Sep;29(9):2375-2381. [doi: 10.1007/s11136-020-02493-7] [Medline: 32253669]
78. Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the apathy evaluation scale. Psychiatry Res 1991 Aug;38(2):143-162. [doi: 10.1016/0165-1781(91)90040-v] [Medline: 1754629]
79. Nobis L, Husain M. Apathy in Alzheimer's disease. Curr Opin Behav Sci 2018 Aug;22:7-13. [doi: 10.1016/j.cobeha.2017.12.007] [Medline: 30123816]
80. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [doi: 10.1093/gerona/56.3.m146] [Medline: 11253156]

XSL•FO
RenderX

81. Chen LK, Woo J, Assantachai P, et al. Asian working group for sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc 2020 Mar;21(3):300-307. [doi: 10.1016/j.jamda.2019.12.012]
82. Greenwood JLJ, Joy EA, Stanford JB. The physical activity vital sign: a primary care tool to guide counseling for obesity. J Phys Act Health 2010 Sep;7(5):571-576. [doi: 10.1123/jpah.7.5.571] [Medline: 20864751]
83. Bouillon K, Kivimaki M, Hamer M, et al. Measures of frailty in population-based studies: an overview. BMC Geriatr 2013 Dec;13(1):1-11. [doi: 10.1186/1471-2318-13-64]
84. Tay L, Tay EL, Mah SM, Latib A, Koh C, Ng YS. Association of intrinsic capacity with frailty, physical fitness and adverse health outcomes in community-dwelling older adults. J Frailty Aging 2023;12(1):7-15. [doi: 10.14283/jfa.2022.28] [Medline: 36629078]
85. Tay LB, Chua MP, Tay EL, et al. Multidomain geriatric screen and physical fitness assessment identify prefrailty/frailty and potentially modifiable risk factors in community-dwelling older adults. Ann Acad Med Singap 2019 Jun 15;48(6):171-180. [doi: 10.47102/annals-acadmedsg.V48N6p171]
86. Gale CR, Cooper C. Attitudes to ageing and change in frailty status: the English longitudinal study of ageing. Gerontology 2018;64(1):58-66. [doi: 10.1159/000477169] [Medline: 28848095]
87. Church S, Rogers E, Rockwood K, Theou O. A scoping review of the clinical frailty scale. BMC Geriatr 2020 Oct 7;20(1):393. [doi: 10.1186/s12877-020-01801-7] [Medline: 33028215]
88. Rockwood K, Song X, MacKnight C, Bergman H, Hogan DB, McDowell I. A global clinical measure of fitness and frailty in elderly people. Can Med Assoc J 2005 Aug 30;173(5):489-495. [doi: 10.1503/cmaj.050051]
89. National frailty strategy policy report. Ministry of Health Singapore. 2023. URL: https://www.moh.gov.sg/resources-statistics/reports/frailty-strategy-policy-report [accessed 2024-10-21]
90. Chong E, Ho E, Baldevarona-Llego J, Chan M, Wu L, Tay L. Frailty and risk of adverse outcomes in hospitalized older adults: a comparison of different frailty measures. J Am Med Dir Assoc 2017 Jul 1;18(7):638. [doi: 10.1016/j.jamda.2017.04.011] [Medline: 28587850]
91. Theou O, Pérez-Zepeda MU, van der Valk AM, Searle SD, Howlett SE, Rockwood K. A classification tree to assist with routine scoring of the clinical frailty scale. Age Ageing 2021 Jun 28;50(4):1406-1411. [doi: 10.1093/ageing/afab006]
92. Zhou K, Ng YS, Tay EL, Mah SM, Tay L. Intrinsic capacity assessment using World Health Organization g older adults.. Geriatr Gerontol Int 2024 May;24(5):457-463. [doi: 10.1111/ggi.14869] [Medline: 38597589]
93. Brooks R, Group E. EuroQol: the current state of play. Health Policy 1996 Jul;37(1):53-72. [doi: 10.1016/0168-8510(96)00822-6] [Medline: 10158943]
94. Zhou T, Guan H, Yao J, Xiong X, Ma A. The quality of life in Chinese population with chronic non-communicable diseases according to EQ-5D-3L: a systematic review. Qual Life Res 2018 Nov;27(11):2799-2814. [doi: 10.1007/s11136-018-1928-y]
95. Konerding U, Elkhuizen SG, Faubel R, et al. The validity of the EQ-5D-3L items: an investigation with type 2 diabetes patients from six European countries. Health Qual Life Outcomes 2014 Dec 5;12:181. [doi: 10.1186/s12955-014-0181-5] [Medline: 25479769]
96. Kim SH, Jo MW, Lee JW, Lee HJ, Kim JK. Validity and reliability of EQ-5D-3L for breast cancer patients in Korea. Health Qual Life Outcomes 2015 Dec 23;13:203. [doi: 10.1186/s12955-015-0399-x] [Medline: 26694964]
97. Wang V, Kandiah N, Lin X, Wee H. Comparison of caregiver-rated 3-level Euroqol-5D (EQ-5D-3L) and health utilities index mark 3 (HUI3) in patients with dementia in Singapore. Value Health 2016 Nov;19(7):A810. [doi: 10.1016/j.jval.2016.08.710]
98. Pérez-Ros P, Vila-Candel R, Martin-Utrilla S, Martínez-Arnau FM. Health-related quality of life in community-dwelling older people with cognitive impairment: EQ-5D-3L measurement properties. J Alzheimers Dis 2020;77(4):1523-1532. [doi: 10.3233/JAD-200806] [Medline: 32925071]
99. Kaiser MJ, Bauer JM, Ramsch C, et al. Validation of the t short-form (MNA®-SF): a practical tool for identification of nutritional status.. J Nutr Health Aging 2009 Nov;13(9):782-788. [doi: 10.1007/s12603-009-0214-7] [Medline: 19812868]
100. Kim J, Lee Y, Won CW, Lee KE, Chon D. Nutritional status and frailty in community-dwelling older Korean adults: the Korean frailty and aging cohort study. J Nutr Health Aging 2018;22(7):774-778. [doi: 10.1007/s12603-018-1005-9] [Medline: 30080218]
101. Soysal P, Veronese N, Arik F, Kalan U, Smith L, Isik AT. Mini nutritional assessment scale-short form can be useful for frailty screening in older adults. Clin Interv Aging 2019;14:693-699. [doi: 10.2147/CIA.S196770] [Medline: 31118593]
102. Rubenstein LZ, Harker JO, Salvà A, Guigoz Y, Vellas B. Screening for undernutrition in geriatric practice: developing the short-form mini-nutritional assessment (MNA-SF). J Gerontol A Biol Sci Med Sci 2001 Jun;56(6):M366-M372. [doi: 10.1093/gerona/56.6.m366] [Medline: 11382797]
103. Wei K, Nyunt MSZ, Gao Q, Wee SL, Yap KB, Ng TP. Association of frailty and malnutrition with long-term functional and mortality outcomes among community-dwelling older adults: results from the Singapore longitudinal aging study 1. JAMA Netw Open 2018 Jul 6;1(3):e180650. [doi: 10.1001/jamanetworkopen.2018.0650] [Medline: 30646023]
104. Wei K, Thein FS, Nyunt MSZ, Gao Q, Wee SL, Ng TP. Nutritional and frailty state transitions in the Singapore longitudinal aging study. J Nutr Health Aging 2018;22(10):1221-1227. [doi: 10.1007/s12603-018-1096-3] [Medline: 30498830]

XSL•FO
RenderX

105. Montejano Lozoya R, Martínez-Alzamora N, Clemente Marín G, Guirao-Goris SJA, Ferrer-Diego RM. Predictive ability of the mini nutritional assessment short form (MNA-SF) in a free-living elderly population: a cross-sectional study. PeerJ 2017;5:e3345. [doi: 10.7717/peerj.3345]
106. Suzana S Jr, Siti Saifa H. Validation of nutritional screening tools against anthropometric and functional assessments among elderly people in selangor. Malays J Nutr 2007 Mar;13(1):29-44 [FREE Full text] [Medline: 22692187]
107. Cheung YB, Xu Y, Feng L, et al. Unconditional and conditional standards using cognitive function curves for the modified mini-mental state exam: cross-sectional and longitudinal analyses in older chinese adults in Singapore. Am J Geriatr Psychiatry 2015 Sep;23(9):915-924. [doi: 10.1016/j.jagp.2014.08.008] [Medline: 25260558]
108. Winblad B, Palmer K, Kivipelto M, et al. Mild cognitive impairment--beyond controversies, towards a consensus: report of the international working group on mild cognitive impairment. J Intern Med 2004 Sep;256(3):240-246. [doi: 10.1111/j.1365-2796.2004.01380.x] [Medline: 15324367]
109. Jones BL, Nagin DS, Roeder K. A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. Sociol Methods Res 2001 Feb;29(3):374-393. [doi: 10.1177/0049124101029003005]
110. Tan AH, Subagdja B, Wang D, Meng L. Self-organizing neural networks for universal learning and multimodal memory encoding. Neural Netw 2019 Dec;120:58-73. [doi: 10.1016/j.neunet.2019.08.020]
111. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997 Nov 1;9(8):1735-1780. [doi: 10.1162/neco.1997.9.8.1735]

Abbreviations

AI: artificial intelligence
ART: adaptive resonance theory
AUC: area under the curve
BLE: bluetooth low energy
CFS: Clinical Frailty Scale
COMT: Common Objects Memory Test
CTT: Color Trials Test
CV: cross-validation
IADL: instrumental activities of daily living
MBI: modified Barthel Index
MCI: mild cognitive impairment
MMSE: Mini Mental State Examination
MNA-SF: Mini Nutritional Assessment Short Form
MoCA: Montreal Cognitive Assessment
PF: prefrailty
PIR: passive infrared
VAS: Visual Analogue Scale

Edited by A Schwartz; submitted 25.Jun.2025; peer-reviewed by A Billis, CM Choi; revised version received 24.Jan.2026; accepted 28.Jan.2026; published 01.Apr.2026.

Please cite as:
Moon JP, Abdul Jabbar K, Tay TCI, Tay L, Mahendran R, Ng TP, Tan SM, Koh WZH, Ying WY, Tan AH, Lee TS, Rawtaer I Development and Validation of Machine Learning Models for Predicting Early Cognitive Decline Using Home Sensor–Derived Behavioral Data: Sensors in-Home for Elder Wellbeing (SINEW) Cohort Study
JMIR Res Protoc 2026;15:e79490
URL: https://www.researchprotocols.org/2026/1/e79490
doi:10.2196/79490

© James Patrick Moon, Khalid Abdul Jabbar, Tony Chin Ian Tay, Laura Tay, Rathi Mahendran, Tze Pin Ng, Shian Ming Tan, Wilbur Zhi Hao Koh, Weng Yan Ying, Ah-Hwee Tan, Tih-Shih Lee, Iris Rawtaer. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 1.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Longitudinal and Multimodal Recording System to Capture Real-World Patient-Clinician Conversations for AI and Encounter Research: Protocol for an Observational Study

Misk Al Zahidy[1], MS; Kerly Guevara Maldonado[1], MD; Luis Vilatuna Andrango[1], MD; Ana Cristina Proano[1], MD; Ana Gabriela Claros[1], MD; Maria Lizarazo Jimenez[1], MD; Esteban Gomez-Alvarez[1], MD; David Toro-Tobon[1,2], MD; Victor M Montori[1,2], MSc, MD; Oscar J Ponce-Ponte[1,3], MSc, MD; Juan P Brito[1,2], MSc, MD

[1]Care and AI Laboratory, Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, United States

[2]Division of Endocrinology, Diabetes, Metabolism, and Nutrition, Department of Medicine, Mayo Clinic, 200 First Street SW, Rochester, MN, United States

[3]Derriford Hospital, University Hospitals Plymouth NHS Trust, Plymouth, United Kingdom

Corresponding Author:
Juan P Brito, MSc, MD
Care and AI Laboratory, Knowledge and Evaluation Research Unit, Mayo Clinic, Rochester, MN, United States

Abstract

Background: The promise of artificial intelligence (AI) in medicine depends on its ability to learn from data that reflect what matters to patients and clinicians in the care process. Most existing models are trained on electronic health records (EHRs), which primarily capture biological measures but rarely the interactions and relationships between patients and clinicians. These relationships, central to how care is understood, negotiated, and delivered, unfold across multiple modalities, including voice, text, and video, yet remain largely absent from current datasets. As a result, AI systems trained solely on EHRs risk perpetuating a narrow biomedical view of medicine and overlooking the lived exchanges that define clinical encounters.

Objective: This study aims to design, implement, and evaluate the feasibility of a longitudinal, multimodal system for capturing patient-clinician encounters, linking 360° video or audio recordings with postvisit surveys and EHR data, to create a foundational dataset for downstream AI research.

Methods: This single-site study was conducted in an academic outpatient specialty clinic (Division of Endocrinology, Mayo Clinic, Rochester, Minnesota, United States). Adult patients attending in-person visits with participating clinicians were invited to enroll. Encounters were recorded using a 360° 2D monocular video camera and dual-channel audio. After each visit, patients completed a brief survey assessing relational empathy, satisfaction, visit pace, and treatment burden. Demographic and clinical data were extracted from the EHR. Feasibility was assessed using 5 prespecified end points: clinician consent, patient consent, recording success, survey completion, and data linkage across modalities.

Results: Recruitment began in January 2025. By August 2025, 35 of 36 (97%) eligible clinicians and 212 of 281 (75%) approached eligible patients had consented. Of the consented encounters, 162 (76%) resulted in a complete 360° video recording, and the postvisit surveys were completed for 204 of 212 (96%) consented encounters, reflecting 1 survey per encounter. Data collection is ongoing as of December 2025, and further analyses will be reported in subsequent publications.

Conclusions: This protocol describes a longitudinal multimodal encounter capture system that links 360° audio or video with postvisit surveys and EHR data. The study specifies operational definitions, workflows, feasibility end points, and governance procedures to support implementation and replication in other clinical settings.

International Registered Report Identifier (IRRID): DERR1-10.2196/84688

(JMIR Res Protoc 2026;15:e84688) doi:10.2196/84688

KEYWORDS

artificial intelligence; AI; digital health; electronic health records; endocrinology; feasibility studies; multimodal data capture; patient-clinician communication

Introduction

The patient-clinician encounter is the central space where medicine is practiced, where problems are explored, treatment options are negotiated, and care plans are co-created. These encounters require attention to both biology and biography, encompassing not only laboratory results or physical findings but also the lived experience, values, and goals of patients. The

XSL•FO
RenderX

quality of this interaction—how clinicians and patients listen, respond, and work together—determines whether care fits the realities of patients' lives and enables them to flourish despite illness. As the primary site where medicine unfolds, encounters should also be the primary lens through which we study, evaluate, and improve care [1,2].

The current paradigm in medical artificial intelligence (AI) is shifting from unimodal models, which analyze a single data type, to multimodal systems that integrate diverse inputs such as images, text, and structured data. To contribute meaningfully to patient care, these systems must be trained on data that reflect clinical encounters rather than relying solely on electronic health records (EHRs). Clinical notes, while indispensable for documentation, are written largely for billing and regulatory purposes and, as a result, emphasize biological measures while neglecting relational and contextual aspects of care [3,4]. As a result, they are a poor surrogate for the lived dynamics of encounters and limit the capacity of AI systems to develop a holistic understanding of patients that is necessary to foster patient-centered care [5].

To capture both the biology and the biography of patients, including their experiences, values, and goals of care, future AI models must be trained on encounter data where communication unfolds across multiple modalities.

The primary bottleneck for developing person-centered AI is the lack of high-quality, real-world, longitudinal multimodal data from clinical encounters. No scalable process currently exists to record and integrate this information at the level needed to support AI development and research. Recording real-world visits may disrupt clinical workflows, and both patients and clinicians must be comfortable with how data are captured, stored, and used. Consent processes need to address not only privacy and confidentiality but also future applications in AI research. Institutions must ensure ethical oversight, secure storage, and clarity around ownership and secondary use. With little precedent for routine, large-scale capture of encounter data, a feasibility study is essential to determine whether it can be done in a way that is acceptable, minimally disruptive, and ethically robust [6,7].

To address this gap, we designed a feasibility study to test whether multimodal data from real-world patient-clinician encounters can be captured in a way that is practical, acceptable, and ethically sound. This protocol outlines the development of a replicable system that integrates 360° video and audio recordings with patient-reported surveys and EHR data. By focusing on feasibility end points, including clinician and patient consent, recording success, survey completion, and data linkage, this study aims to establish the foundation for the scalable collection of encounter data. Demonstrating feasibility is a critical step toward creating the longitudinal, multimodal datasets needed to train and validate next-generation AI models that reflect the complexity of clinical care.

Methods

Study Setting and Design

This protocol describes a step-by-step methodology to design and implement a longitudinal, multimodal encounter capture system in an academic outpatient specialty clinic (Division of Endocrinology, Mayo Clinic, Rochester, Minnesota, United States). This study is a prospective feasibility implementation protocol designed to evaluate whether a standardized multimodal encounter capture system can be integrated into routine outpatient care and assessed using predefined feasibility and acceptability end points. The protocol specifies the procedures, end points, and analysis plan for assessing the implementation of the encounter capture system.

Endocrinology was chosen because visits often involve chronic conditions that require longitudinal management, complex decision-making, and sustained patient-clinician collaboration. The investigators' affiliation with the division further facilitated early approvals and integration with practice. However, this protocol can be implemented across multiple health care services.

System Development

Multidisciplinary Design Process

A multidisciplinary team, including encounter researchers, endocrinologists, implementation scientists, AI researchers, and health-services researchers, uses a rapid, iterative design approach to develop a system capable of capturing rich, multimodal data from routine clinical encounters and to define standardized workflow and data management procedures.

Core Data Components

The system captures information from three integrated sources: (1) multimodal video recordings of clinical encounters; (2) postencounter patient surveys assessing satisfaction, relational empathy, and burden; and (3) structured and unstructured EHR data, including clinical notes, patient portal messages, and clinical variables.

Participant Recruitment and Eligibility Criteria

Clinicians

All practicing clinicians in the Division of Endocrinology, including physicians, advanced practice providers, fellows, certified diabetes care and education specialists, and registered nurses, are eligible to participate. Clinicians are approached in person or via email, provided with detailed study information, and consented electronically on a tablet. Upon consent, each clinician is assigned a unique participant ID (eg, ENDOC###) and receives an email to complete a short demographic survey which includes their professional role, age, gender, race or ethnicity, and years in practice.

Patients

Patients are eligible if they are adults (aged ≥18 y), have a scheduled in-person appointment with a participating clinician, and can provide informed consent in English. Exclusion criteria include the presence of cognitive impairment, the need for an

XSL•FO
RenderX

interpreter, or not being the primary individual responsible for managing their care. Enrolled patients are assigned a unique participant ID (ie, ENDOP###) for longitudinal linkage across visits.

Any adult guests present in the exam room (eg, spouse and adult child) must provide verbal consent before recording. If a guest declines to be recorded or if a minor is present, the encounter is not recorded for that visit to ensure privacy and compliance. The full process of recruitment is visualized in Figure 1.

Figure 1. Recruitment and consent workflow for clinicians and patients.

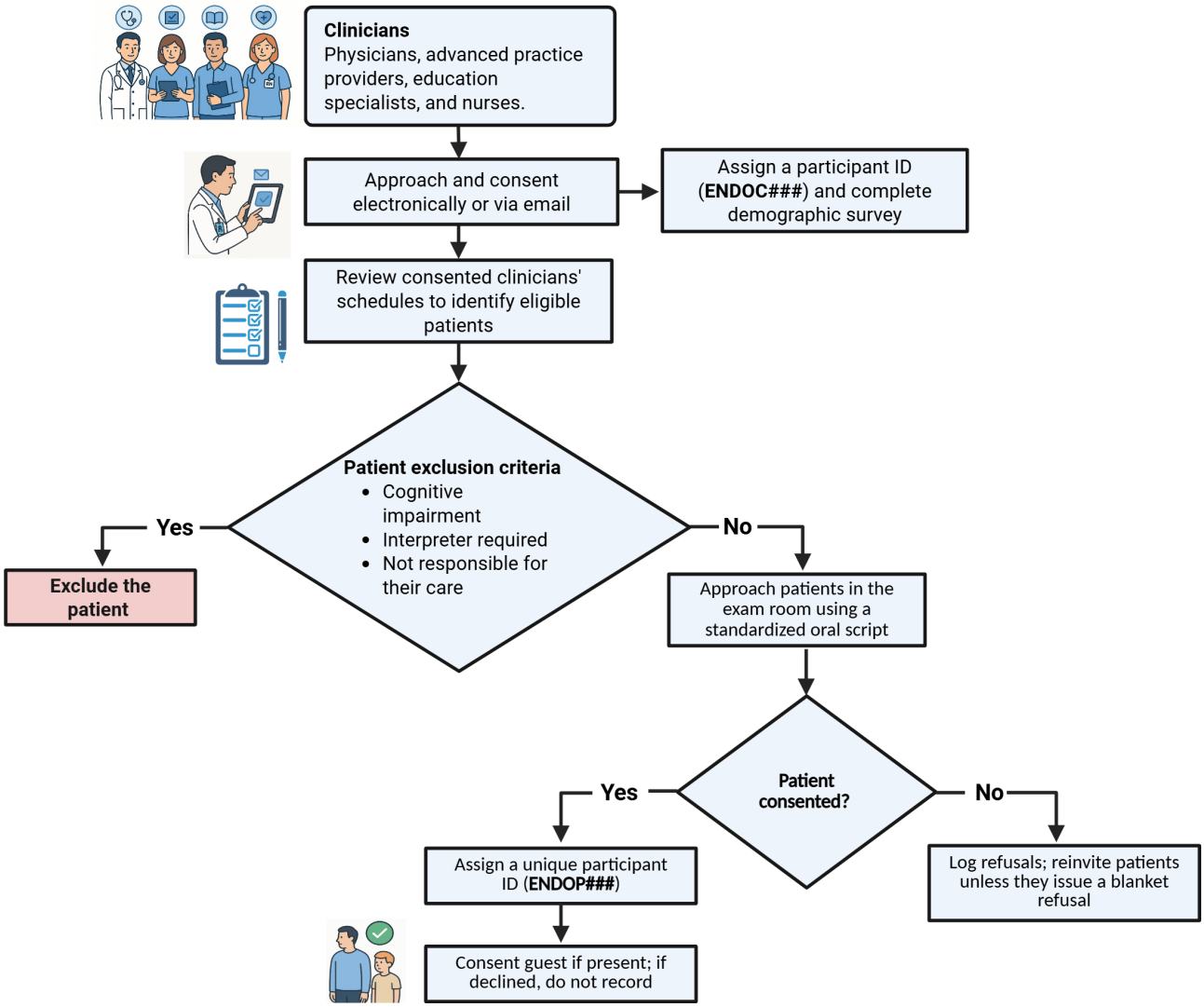

Eligibility, consent, and encounter-level participation are tracked in a recruitment log to support participant flow reporting and feasibility monitoring.

Recruitment Process

A team of 5 trained nonclinical research staff manages all recruitment and data collection activities. The standardized recruitment process is as follows:

- Daily schedule review: research staff review participating clinicians' schedules to identify eligible new patients and previously consented patients returning for follow-up appointments.
- Low-pressure approach: research staff approach eligible patients in the exam room using a standardized, institutional review board (IRB)–approved oral script (Multimedia Appendix 1). The script uses clear, lay language to explain the study's purpose, emphasizing that participation is entirely voluntary, privacy will be protected, and the recording can be paused or stopped at any time without affecting their care.
- Consent and confirmation: new patients provide informed consent on a tablet. For returning patients who have previously consented, the staff obtains verbal confirmation that they are still comfortable being recorded. Full written reconsent is only required if there are protocol changes.
- Declinations: patients may decline recording for any appointment. Reasons for refusal are logged by staff to monitor for potential selection bias. Patients who decline are not re-approached the same day but may be invited to participate again at future visits unless they issue a blanket refusal for all future participation.

Following recruitment and consent, the multimodal encounter capture workflow is implemented as described in the "System Components and Workflow" section.

XSL•FO
RenderX

System Components and Workflow

This section describes the operational workflow used to deliver the multimodal encounter capture system, as well as the standardized procedures used to collect each data stream.

Video Recordings

Encounters are recorded using a 360° video, 2D monocular camera (eg, Insta360×4). The camera is configured at 5.7K resolution and 30 frames per second (fps; ~100 Mbps) to balance image quality with stable performance during longer outpatient visits (eg, reduced risk of overheating). We selected a consumer-grade 360° camera to enable wide-angle capture while keeping the system low-cost and replicable across clinical settings. Recording specifications and quality control criteria used to classify recordings as complete or usable are summarized in Table 1.

Table . Recording specifications and criteria for recordings.

Domain	Specification	Quality control criterion
Video	360° video; 5.7K resolution; 30 fps; approximately 100 Mbps	File opens/plays; no corruption; continuous capture across encounter (no premature stop)
Audio	Stereo dual-channel at 48 kHz; wind reduction disabled; planned downsampling to 16 kHz	Both channels present; speech audible; no sustained dropout or severe clipping preventing analysis
Placement	Concealed tabletop placement; ≥2.5 ft distance	Patient–clinician interaction area visible for most of visit
Seam avoidance	Avoid 360° seam or stitch zone	No seam artifacts affecting the interaction zone
Occlusion	Positioning to avoid obstruction by room objects	No prolonged occlusion that prevents observation or feature extraction
Duration	Recording intended to cover full encounter	Recording duration consistent with visit length; no early termination attributable to device failure
File integrity and upload	Same-day upload to secure server and upload verification	Successful upload verification; retrievable file stored on server

The camera is positioned discreetly on a tabletop stand and aligned so participants remain at an optimal distance (≥2.5 ft) and away from the camera's stitch or seam line, where visual artifacts are most common. Audio is captured in stereo dual-channel at 48 kHz using the camera's "stereo" mode to preserve both patient and clinician voices, with planned downsampling to 16 kHz for speech analytics. "Wind reduction" is disabled to avoid distortion of vocal frequencies relevant for acoustic analysis. To support continuous operation across clinic sessions, batteries are rotated between encounters. Participants are informed that they may pause, redirect, or stop the recording at any time. The 5.7K or 30 fps setting allows for at least 135 minutes of recording time in the latest 360°, 2D monocular camera, sufficient for typical outpatient visits.

A standardized workflow is followed daily to ensure data integrity and security. At the end of each clinic session, the SD card is retrieved, and original camera files (INSV) are transferred to an institutional workstation. Files are then converted to MP4 using the manufacturer's software (when applicable) and uploaded the same day to a restricted-access, Health Insurance Portability and Accountability Act (HIPAA)–compliant institutional server. All files are renamed using a harmonized participant ID and encounter date. After upload verification, raw files are deleted from the SD card and any intermediate workstations. The daily workflow is illustrated in Figure 2. These standardized procedures ensure that recording, transfer, and storage processes are consistent across encounters and reproducible in other clinical settings.

XSL•FO
RenderX

Figure 2. Camera setup: Multimodal encounter capture. The system integrates 3 stages: pre-encounter stage (clinician and patient consent and camera setup), during the encounter (video or audio recording and manual electronic health record [EHR] data extraction), and postencounter (patient survey completion and secure video upload).

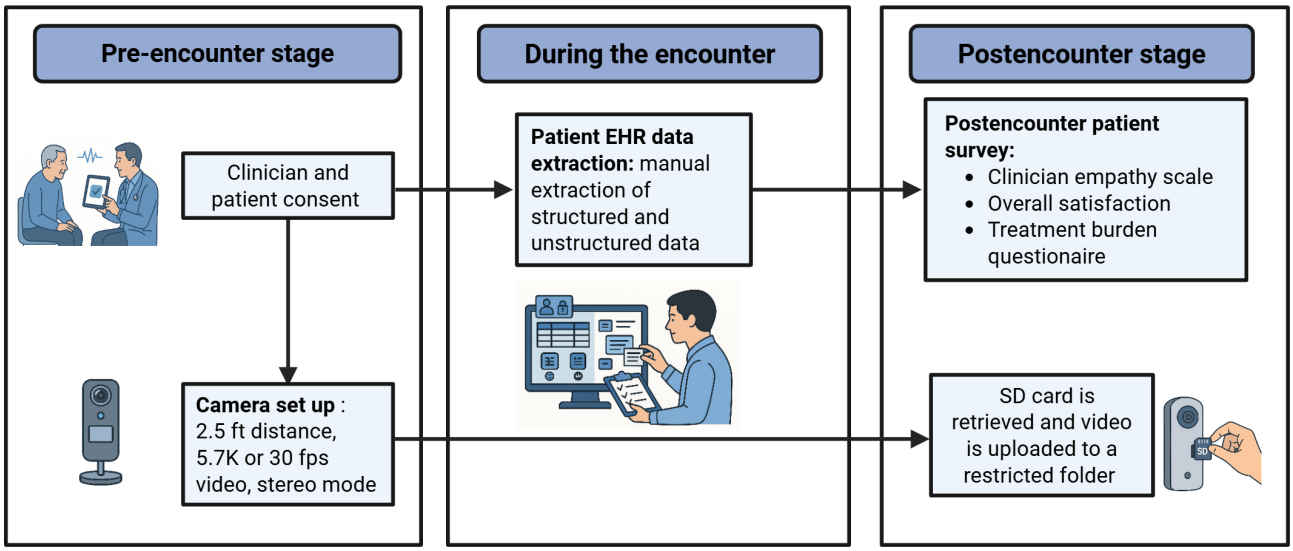

Postencounter Patient Surveys

Immediately following the clinical encounter, patients are invited to complete a brief postvisit survey administered on a tablet. The survey includes the following validated instruments: (1) the Consultation and Relational Empathy (CARE) measure to assess patient-perceived clinician empathy and overall satisfaction, (2) the Unhurried Conversations Assessment Tool (UCAT) to assess perceived visit pace (unhurried vs rushed), and (3) the Treatment Burden Questionnaire plus Digital (TBQ+D) to assess overall treatment burden including burden related to digital care (Multimedia Appendix 2) [8-12]. It also includes the following 2 open-ended questions designed to elicit qualitative feedback:

- "Tell us how, if at all, did your visit with an endocrinology clinician help you?"
- "If you had a magic wand and you could change one thing about your visit today, what would that be?"

CARE and UCAT are scored according to published guidance, with higher scores indicating greater perceived empathy and more unhurried communication, respectively. TBQ+D is scored by summing item responses to generate a total burden score (higher scores indicate greater burden). TBQ+D is administered to all consented patients after each recorded encounter.

Handling of Missing Survey Data

Surveys are classified as complete, partial, or missing. For feasibility reporting, a survey is considered complete when the CARE and UCAT instruments are completed and at least 2 key TBQ+D items (financial burden and arranging medical appointments) are answered. This operational definition ensures that surveys classified as complete include both encounter-experience measures and key treatment workload items. Completion of the full TBQ+D is not required for the survey to be counted as complete. All submitted surveys are retained, including those with incomplete responses.

Instrument scores are calculated using available responses; no imputation or prorating of missing items is performed. When items are missing, scores reflect only the responses provided, and the extent of missingness will be summarized descriptively.

If a survey is not completed in clinic, study staff send up to 3 reminders via email and/or the patient portal within two weeks of the encounter. Surveys completed after reminders are retained and considered completed for feasibility analyses. Completion mode (in-clinic vs remote) and timing are recorded to allow descriptive assessment of response patterns.

EHR Data Extraction

Following a recorded encounter, trained study staff perform a manual extraction of structured (demographic and clinical variables) and unstructured (clinical notes and patient portal messages) data from the patient's EHR. Structured data include age, sex, race or ethnicity, clinical diagnoses, visit type (eg, new or follow-up), and current medications.

Pilot Phase and Iterative Refinement

The system is piloted over 3 weeks, with 30 recorded encounters (~10 per wk) to identify and resolve logistical issues related to camera placement (eg, ensuring the camera is unobtrusive), workflow disruption (eg, additional steps or delays during rooming), and recording quality (eg, interruptions or file-transfer failures). Pilot feedback is reviewed weekly, and the recruitment script, equipment handling, and data processing steps are refined before initiating full-scale recruitment. Early workflow and technical failures are documented and inform refinements to standardized camera placement, battery management, and file verification procedures prior to full-scale recruitment. Recorded encounters from the pilot phase can be included in the final dataset.

The system is refined based on three design constraints to ensure the final system is scalable and replicable: (1) comprehensive video capture: the system must provide 360-degree video coverage without requiring a dedicated camera operator in the room; (2) high-quality audio: the system must capture dual-channel audio at a sample rate sufficient for automated speech recognition and diarization (≥16 kHz); and (3) minimal

XSL•FO
RenderX

clinic disruption: equipment setup and recording initiation must fit within the standard patient rooming process, without extending visit times.

Data Integration and Management

Triangulation and Harmonization

The study is designed to capture three components: (1) 360° video-recorded clinical interactions, (2) post-encounter patient surveys, and (3) structured and unstructured data extracts from the EHR. These sources are harmonized using an identification schema and documented in a central tracking log.

For clinicians, the log tracks their professional role, consent status, and completion of the demographic survey. For patients, it records demographic data (age, sex, and race or ethnicity), eligibility status with documented reasons for ineligibility, guest presence and consent status, recording success, video upload success, postvisit survey completion status (complete, partial, or none), and any challenges noted by the research staff (eg, equipment malfunction and patient discomfort). Recording failure modes are logged in the encounter tracking log.

Each entry is indexed by the unique participant ID, connecting all 3 data streams. All files follow a harmonized naming convention based on the participant ID and encounter date to allow reliable data merging and adherence to the findable, accessible, interoperable, and reusable data principles, maximizing the use of the collected data for future research. Table 2 summarizes the study's data sources and how multimodal data are formatted, linked, and stored securely. This structured approach ensures reliable data merging and supports complex, temporally aligned multimodal analyses.

Table . Multimodal data streams, key variables, and linkage schema.

Data source	Modality	Key variables or format	Linkage key	Storage platform
Camera	Video	360° video, 5.7K or 30 fps, H.265 codec, and .mp4 format	PATIENT_ID_ENCOUNTER_DATE.mp4	Secure institutional server
Camera	Audio	Dual-channel, 48 kHz WAV, and down sampled to 16 kHz	PATIENT_ID_ENCOUNTER_DATE.wav	Secure institutional server
Survey	Patient-reported	Empathy, satisfaction, burden (Likert scales), and open-ended text	PATIENT_ID, ENCOUNTER_DATE	REDCap[a] (Vanderbilt University)
EHR[b]	Structured and unstructured information	Demographics, ICD-10[c] codes, medications, and laboratory values	PATIENT_ID, ENCOUNTER_DATE	REDCap (via manual extraction)

[a]REDCap: Research Electronic Data Capture.

[b]EHR: electronic health record.

[c]ICD-10: *International Classification of Diseases, Tenth Revision.*

Storage and Security

Video and audio recordings are labeled with study-specific IDs and uploaded daily to a secure institutional server with institutional authentication and role-based access controls. Following upload verification, raw files are deleted from the camera's SD card and any intermediate workstations. All study data are stored in accordance with HIPAA regulations and institutional data security protocols and are protected using institutional security safeguards, including encryption protections. Access to identifiable recordings is restricted to authorized study personnel using institutional authentication and role-based permissions, with all access auditable under institutional information security practices.

Survey and EHR data are managed and stored using REDCap (Vanderbilt University), a secure, web-based application designed for research data collection and management. The final analysis dataset is fully deidentified, with the unique participant ID serving as the only link between the different data modalities. Patients who withdrew consent or became ineligible during the study are excluded from the final dataset unless explicit permission is granted to retain their previously collected data.

In accordance with IRB approval and participant authorization, identifiable encounter recordings and linked study data may be retained in a secure research registry for ongoing and future IRB-approved research. Participant authorization to use and share health information is not time-limited and remains in effect unless revoked. Participants may withdraw from the study at any time, after which no additional data will be collected. Participants may also revoke authorization for future use of their information, in which case their information will be removed from repositories where feasible; however, information already distributed for research use cannot be retrieved. Secondary use, including AI or machine learning model development or training, is restricted to IRB-approved purposes and to the scope described in the consent or authorization. External sharing is not planned as part of this feasibility study; any future sharing with collaborators outside the institution would require IRB review and approval and would occur only under a formal data use agreement specifying permitted uses and access restrictions.

After data collection and integration procedures are completed, predefined feasibility and analytic outcomes are assessed as described below.

XSL•FO
RenderX

Data Analysis

The primary outcomes of this feasibility protocol are predefined operational end points reflecting both the feasibility and acceptability of the system, including clinician and patient consent rates, recording success, survey completion, and multimodal data linkage (Table 3).

Table . Primary feasibility end points, definitions, and success thresholds.

End point	Definition (numerator/denominator)[a]	Success threshold (%)
1. Clinician consent rate	Consented clinicians/all approached clinicians	>90
2. Patient consent rate	Consented patients/all approached eligible patients	>50
3. Recording success rate	Complete or usable recordings/all consented encounters	>90
4. Survey completion rate	Completed surveys/all consented encounters	>90
5. Data linkage rate	Encounters with linked video, survey, and EHR[b]/all consented encounters	>90

[a]Feasibility rates are calculated as numerator divided by denominator, where the numerator represents the number of clinicians, patients, or encounters meeting the specified endpoint and the denominator represents the total number of eligible clinicians, patients, or encounters.

[b]EHR: electronic health record.

This protocol reports interim feasibility metrics; additional analytical findings will be reported in a subsequent publication.

Feasibility and Recruitment Metrics

The study's feasibility is measured using 5 predefined end points. These end points, their definitions, and the a priori success thresholds are detailed in Table 3. Descriptive statistics (counts and proportions) are calculated for each end point using data from the recruitment tracking logs and REDCap exports.

All feasibility end points, except clinician and patient consent rates, are calculated at the encounter level, as patients may contribute multiple encounters.

Planned Analysis of Data Components

Overview

The analysis of the 3 components (video recordings, postencounter surveys, and EHR data) will be described, in a future publication, using descriptive statistics. In addition to feasibility end points (Table 3), we will generate encounter-level derived measures from audio or video, survey, and EHR data to support downstream encounter research and AI development. Key derived measures include encounter duration, speech and turn-taking metrics (eg, speaking time per speaker, number of turns, turn length, overlap or interruption frequency, and pause frequency or duration), acoustic or prosodic features (eg, fundamental frequency and intensity summaries), and selected nonverbal behavior features (eg, face presence, gaze direction, head nodding, and smile frequency). Measures sensitive to encounter length will be reported both as absolute values and normalized per minute.

Because patients may contribute multiple encounters and clinicians may contribute encounters with multiple patients, downstream analyses will account for repeated measures and clustering by clinician using mixed-effects models (eg, random intercepts for clinician and patient as appropriate) and/or cluster-robust standard errors. Analyses will also be summarized overall and stratified by visit type (new vs follow-up) and clinician role.

Acoustic Feature Extraction

Video recordings will be automatically transcribed and diarized using the best available models [13,14]. From the audio, key acoustic features will be extracted using the *Parselmouth* Python library (Python Software Foundation), which provides access to the Praat gold-standard algorithms [15]. These include: (1) prosodic features (fundamental frequency mean and SD, and intensity contours), (2) spectral features (formants F1-F3, related to vowel articulation, and (3) temporal features (speech rate, and pause duration and frequency). This allows for quantitative analysis of conversational dynamics such as turn-taking, speech overlap, and backchannels.

Visual Feature Extraction

A planned analysis pipeline using the Open Computer Vision Library (OpenCV) [16] will be used to extract nonverbal cues. This process will involve: (1) face detection using pretrained deep neural network models to locate patient and clinician faces in each frame [16,17], (2) facial landmark detection to identify key points (eg, corners of the eyes and mouth) [16,18], and (3) approximations of facial action units corresponding to specific muscle movements (eg, smiling and brow furrowing) [16,18]. Additional visual features such as smile, nod, gaze, and emotion detection will be extracted using OpenCV and Emotion-Qwen [16,19].

Advanced Emotion and Behavior Analysis

The dataset will be leveraged to train and validate advanced models like Emotion-Qwen [19], a state-of-the-art large multimodal model. This approach will support analysis that moves beyond simple emotion labels (eg, "happy" and "sad") to more complex, context-aware reasoning about the affective dynamics of the conversation, such as identifying moments of empathy, confusion, or shared understanding.

Postencounter Surveys and EHR Data

Data from the postencounter surveys will be used to calculate the mean scores, SDs, and distribution of each scale item. Where relevant, responses will be stratified by key variables of interest,

XSL•FO
RenderX

such as clinician type (eg, physician, fellow, and educator) and visit type (eg, new vs follow-up). For the subset of patients who complete the TBQ+D, total scores and item-level responses will be summarized to describe perceived treatment workload. Analysis will be performed using SAS (version 9.4; SAS Institute).

Structured data from EHRs will be described in the same manner as post-encounter survey results, whereas unstructured data (ie, clinical notes and patient portal messages) will be described in terms of length (number of words).

Ethical Considerations

This study was approved by the Mayo Clinic IRB (number 24 - 012956). All participating clinicians and patients provided informed written consent prior to any recording. Adult guests present in the exam room were asked to provide verbal consent prior to recording; encounters were not recorded if a guest declined or if a minor was present. Participants could request that recording be paused or stopped at any time and could withdraw from the study without affecting their clinical care.

Identifiable audio and 360° video recordings, along with linked survey and electronic health record data, are stored on secure institutional systems. Access to identifiable recordings is restricted to authorized study personnel and is granted only after an individual is added to the IRB-approved study personnel list and receives institutionally approved access. Access provisioning and revocation are managed through institutional information security processes. All access to identifiable data is logged and auditable under institutional security monitoring practices.

Study data, including identifiable encounter recordings, may be retained in a secure research registry without a predefined retention end date, consistent with IRB approval and participant authorization. If a participant withdraws from the study, no additional data are collected; however, data collected prior to withdrawal may be retained and used for research purposes unless the participant explicitly revokes authorization for future use, in accordance with institutional policy.

Secondary use of the data, including use for AI or machine learning research, is permitted only under IRB-approved protocols and within the scope described in the informed consent and authorization. External data sharing is not planned as part of this feasibility study; any future sharing with collaborators outside the institution would require additional IRB review and approval and would occur only under a formal data use agreement specifying permitted uses and access restrictions. No compensation is provided for participation.

Results

The protocol was implemented in the outpatient clinic of the Division of Endocrinology at Mayo Clinic in Rochester, Minnesota, an academic specialty clinic. We selected the Insta360×4 camera because it met all required specifications, including 5.7K resolution at 30 frames per second (~100 Mbps) with dual-channel 48 kHz audio, for a longitudinal 360° 2D encounter capture system. Video files were recorded in INSV format and converted to MP4 using the Insta360 software (Insta360, Shenzhen, China). Pilot-phase recordings meeting the predefined quality control criteria (Table 1) were eligible for inclusion in the final dataset.

Recruitment and encounter recording began in January 2025. By August 2025, 35 of 36 (97%) eligible clinicians and 212 of 281 (75%) approached patients had consented. Of the consented encounters, 162 (76%) resulted in a complete or usable 360° video recording, and the postvisit surveys were completed for 204 of 212 (96%) consented encounters, reflecting 1 survey per encounter. At the time of the interim analysis, 150 of 212 (71%) consented encounters had successfully linked video, survey, and EHR data. Data collection is ongoing as of December 2025, and additional analyses of the multimodal dataset will be reported in subsequent publications.

Discussion

Principal Results

This protocol describes a scalable framework for capturing rich, multimodal, longitudinal, real-world clinical encounters in chronic disease care. The primary objective of this study is to assess feasibility, determined by achieving predefined thresholds for recruitment, data capture, and data linkage. Interim feasibility metrics include 97% clinician consent, 75% patient consent, and 96% postvisit survey completion across consented encounters, reflecting 1 survey per encounter. Although recording success (complete or usable 360° recordings) was 76% of consented encounters to date, common reasons for incomplete or unusable recordings included workflow interruptions (eg, inconsistent camera placement leading to occlusion or seam artifacts, visit-length variability, and battery depletion) and intermittent device malfunction (eg, unexpected shutoffs requiring servicing). These workflow and equipment issues led to protocol refinements, including standardized camera placement guidance to minimize occlusion and seam artifacts, a fixed battery rotation schedule aligned with anticipated visit duration, and same-day verification of file integrity following each encounter. Similarly, linkage across video, survey, and EHR data streams was lower during early implementation primarily because some encounters lacked a complete recording, which is required for multimodal linkage. Improvements to recording workflows and quality control procedures described above are expected to improve linkage rates in subsequent phases.

These implementation metrics describe workflow integration in an outpatient specialty clinic and inform subsequent planned analyses. The collected corpus is designed to support downstream person-centered AI and encounter research by enabling multimodal analysis of real-world clinical interactions.

Comparison With Prior Work

There have been similar efforts to capture multimodal information from natural conversations, but few have focused on real-world health care encounters. To the best of our knowledge, prior studies in health care have been limited to audio-only transcript analysis, such as applying natural language processing to primary care conversations [6], or video capture

XSL•FO
RenderX

in simulated training settings rather than routine clinical care [20]. Outside of health care, multimodal datasets such as the CANDOR corpus have advanced the study of natural conversation, but these do not address the ethical, logistical, and workflow challenges unique to clinical environments, such as HIPAA compliance and the need to avoid disrupting care delivery [7].

This protocol advances the field by embedding multimodal capture directly into routine specialty care, an approach not addressed in previous studies. The iterative piloting with clinical staff illustrates stakeholder engagement, ensuring the system fits existing workflows and minimizes disruption, a principle emphasized in implementation science. The harmonized data linkage strategy enhances feasibility and replicability by reliably integrating video, survey, and EHR data. Finally, the use of low-cost, general-purpose hardware increases scalability and democratizes access to multimodal encounter research, including resource-limited settings.

Next Steps and Future Directions

The immediate next steps involve analyzing the rich dataset that this protocol enables, with initial work focused on correlating patient-reported outcomes with observable communication dynamics. For example, verbal and nonverbal behaviors such as speaking time, turn-taking, and gaze can be examined in relation to patient-reported empathy, satisfaction, and treatment burden. Planned analyses will examine conversational moments characterized by empathy, attentiveness, and shared decision-making in relation to patient-reported empathy, satisfaction, visit pace, and treatment burden. Linking observed communication behaviors to patient-reported outcomes will be explored to characterize patterns associated with higher or lower perceived treatment burden, particularly in relation to navigating services and managing costs.

Beyond these initial projects, the dataset presents significant opportunities for advanced AI research, though not without challenges. Processing each modality introduces unique complexities: audio data requires robust speaker diarization and transcription models tuned for clinical conversations; video data presents challenges in data volume, storage, and the computational cost of extracting nonverbal cues like facial expressions and gestures; and integrating these streams with EHR and survey data requires sophisticated data fusion techniques to handle issues such as missing data and ensure temporal alignment.

Furthermore, while this protocol was implemented in endocrinology, the framework is designed for broader applicability. Future work will involve deploying this system in other clinical settings, such as oncology, where complex treatment decisions and longitudinal patient relationships are also central to care. In oncology, integrating encounter data with genomic, imaging, and clinical data may inform future development of AI models to support shared decision-making and personalize patient communication, augmenting rather than replacing clinical judgment. By systematically addressing these next steps, this work can help build the foundational datasets needed to move beyond EHR-based models and develop AI that truly understands the human context of clinical care.

Finally, it is important to address the question of scalability. While this protocol describes a replicable framework, the long-term vision is not to place a camera in every exam room for every encounter, which may not be practical or desirable. Instead, this foundational research is a crucial step toward the development of more integrated and less obtrusive ambient intelligence technologies. These future systems, embedded seamlessly into the clinical environment, could capture the essential multimodal data of an encounter without the need for standalone recording devices. The insights gained from this study, regarding consent workflows, patient and clinician acceptance, and ethical safeguards, will be invaluable for informing the design of next-generation ambient systems that can capture encounter data at scale while maintaining the trust and privacy essential to the patient-clinician relationship. By systematically addressing these next steps, this work can help build the foundational datasets needed to move beyond EHR-based models and develop AI that truly understands the human context of clinical care.

Limitations

This framework faces potential challenges, and the protocol includes specific strategies to mitigate them. One concern is selection bias, where patients who consent to participate may differ systematically from those who do not—for example, by being more comfortable with technology or having higher trust in the health care system. To address this, the protocol includes logging reasons for refusal, using a low-pressure consent script to build trust, and conducting a comparative analysis of key variables between participants and non-participants to quantify any observed bias. In particular, we will examine whether consent rates vary by visit type (new vs follow-up) and clinician role (eg, physician, fellow, educator, and advanced practice provider), and we will compare available characteristics of consenting versus eligible nonconsenting patients to assess representativeness (eg, age, sex, and race or ethnicity when available). Another potential issue is the Hawthorne effect [21], where the presence of a recording device may alter the natural behavior of participants. The study's longitudinal design serves as the primary mitigation, as reactivity is expected to diminish over time as participants become habituated to the recording process. The use of a small, unobtrusive camera is also intended to minimize the salience of the observation.

The protocol also proactively addresses potential challenges related to the technical reliability of consumer-grade hardware. Because high-resolution 360° video can overheat cameras and exhaust storage, our protocol includes capping resolution at 5.7 K, rotating batteries on a fixed schedule, and formatting SD cards daily. Finally, as a single-site study in a specialized clinic, the findings may not be immediately generalizable. This study is intended as a foundational step, and if feasibility is demonstrated, future multisite trials will be essential to establish broader applicability.

XSL•FO
RenderX

Conclusions

This protocol describes a rigorous, transparent, and replicable method for capturing the multimodal dynamics of clinical encounters in chronic disease care. By detailing operational workflows, technical specifications, feasibility end points (including clinician and patient consent, recording success, survey completion, and data linkage), and mitigation strategies, it offers a practical template for other research teams. These end points use prespecified operational definitions and quality checks, including what constitutes a recorded encounter, a complete or usable 360° recording (full-length capture with acceptable audio or video quality and no critical occlusion or file corruption), and a linked encounter (successful matching of video, survey, and EHR data using the harmonized participant ID and encounter date). Demonstrating the feasibility of this longitudinal approach establishes a resource for developing next-generation, person-centered AI tools that can understand and respond to patients' lived experiences. This work represents a foundational step toward building AI that can appreciate both the "biology" and the "biography" of a patient, helping make health care more effective, equitable, and humane.

Acknowledgments

The authors would like to acknowledge the desk operations specialists and clinicians in the Division of Endocrinology for their support and collaboration during the design and implementation of this study. Their assistance in coordinating clinic logistics and facilitating recruitment was essential to the success of this protocol.

Funding

The authors declare no financial support was received for this work.

Data Availability

The data generated in this study include identifiable patient-clinician encounter recordings and linked electronic health record data. Due to privacy and ethical restrictions, these data cannot be made publicly available. Deidentified data may be made available upon reasonable request to the corresponding author and with appropriate institutional approvals.

Authors' Contributions

MAZ, OJP-P, DT-T, and JPB conceived the study. MAZ led protocol drafting, system design, and preparation of the manuscript. KGM, LVA, ACP, and AGC contributed to patient recruitment, data collection, and drafting sections of the manuscript. MLJ and EG-A assisted with manuscript preparation and figure development. JPB provided overall supervision, clinical oversight, and critical revision of the manuscript. All authors reviewed and approved the final version.

Conflicts of Interest

DT-T was a consultant for Immunovant outside the submitted work in the prior 12 months. The remaining authors have no conflicts of interest to disclose.

Multimedia Appendix 1
Recruitment script.
[DOCX File, 24 KB - resprot_v15i1e84688_app1.docx]

Multimedia Appendix 2
Postencounter patient survey.
[DOCX File, 53 KB - resprot_v15i1e84688_app2.docx]

References

1. Donabedian A. The quality of care. How can it be assessed? JAMA 1988;260(12):1743-1748. [doi: 10.1001/jama.260.12.1743] [Medline: 3045356]
2. May C, Montori VM, Mair FS. We need minimally disruptive medicine. BMJ 2009 Aug 11;339:b2803. [doi: 10.1136/bmj.b2803] [Medline: 19671932]
3. Downing NL, Bates DW, Longhurst CA. Physician burnout in the electronic health record era: are we ignoring the real cause? Ann Intern Med 2018 Jul 3;169(1):50-51. [doi: 10.7326/M18-0139] [Medline: 29801050]
4. Shanafelt TD, Dyrbye LN, Sinsky C, et al. Relationship between clerical burden and characteristics of the electronic environment with physician burnout and professional satisfaction. Mayo Clin Proc 2016 Jul;91(7):836-848. [doi: 10.1016/j.mayocp.2016.05.007] [Medline: 27313121]
5. Tzelepis F, Sanson-Fisher RW, Zucca AC, Fradgley EA. Measuring the quality of patient-centered care: why patient-reported measures are critical to reliable assessment. Patient Prefer Adherence 2015;9(831-5):831-835. [doi: 10.2147/PPA.S81975] [Medline: 26150703]

XSL•FO
RenderX

6. Tran BD, Mangu R, Tai-Seale M, Lafata JE, Zheng K. Automatic speech recognition performance for digital scribes: a performance comparison between general-purpose and specialized models tuned for patient-clinician conversations. AMIA Annu Symp Proc 2022;2022(1072-80):1072-1080. [Medline: 37128439]
7. Reece A, Cooney G, Bull P, et al. The CANDOR corpus: insights from a large multimodal dataset of naturalistic conversation. Sci Adv 2023 Mar 31;9(13):eadf3197. [doi: 10.1126/sciadv.adf3197] [Medline: 37000886]
8. Al Zahidy M, Ridgeway JL, Branda ME, et al. Validation of a novel patient-reported measure of the burden of digital care in diabetes. BMC Med 2025 Dec 20;24(1):45. [doi: 10.1186/s12916-025-04579-w] [Medline: 41419881]
9. Tran VT, Harrington M, Montori VM, Barnes C, Wicks P, Ravaud P. Adaptation and validation of the Treatment Burden Questionnaire (TBQ) in English using an internet platform. BMC Med 2014 Jul 2;12(1):109. [doi: 10.1186/1741-7015-12-109] [Medline: 24989988]
10. Bikker AP, Fitzpatrick B, Murphy D, Mercer SW. Measuring empathic, person-centred communication in primary CARE nurses: validity and reliability of the Consultation and Relational Empathy (CARE) measure. BMC Fam Pract 2015 Oct 23;16(1):149. [doi: 10.1186/s12875-015-0374-y] [Medline: 26493072]
11. Mandhana DM, Jacome CS, Ballard DI, et al. Developing and validating the Unhurried Conversations Assessment Tool (UCAT). Patient Educ Couns 2024 Jun;123:108237. [doi: 10.1016/j.pec.2024.108237] [Medline: 38461793]
12. Krol MW, de Boer D, Delnoij DM, Rademakers J. The net promoter score—an asset to patient experience surveys? Health Expect 2015 Dec;18(6):3099-3109. [doi: 10.1111/hex.12297] [Medline: 25345554]
13. Srivastav V, Majumdar S, Koluguri N, et al. Open automatic speech recognition leaderboard. Hugging Face Spaces. 2023. URL: https://huggingface.co/spaces/hf-audio/open_asr_leaderboard [accessed 2026-03-09]
14. Kuchaiev O, Li J, Nguyen H, et al. NeMo: a toolkit for building AI applications using neural modules. arXiv. Preprint posted online on Sep 14, 2019. [doi: 10.48550/arXiv.1909.09577]
15. Jadoul Y, Thompson B, de Boer B. Introducing Parselmouth: a Python interface to Praat. J Phon 2018 Nov;71:1-15. [doi: 10.1016/j.wocn.2018.07.001]
16. OpenCV – Open Computer Vision Library. 2025. URL: https://opencv.org/ [accessed 2026-03-02]
17. Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees. Presented at: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Jun 23-28, 2014. [doi: 10.1109/CVPR.2014.241]
18. Baltrusaitis T, Zadeh A, Lim YC, Morency LP. OpenFace 2.0: facial behavior analysis toolkit. Presented at: 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018); May 15-19, 2018. [doi: 10.1109/FG.2018.00019]
19. Huang D, Li Q, Yan C, et al. Emotion-Qwen: a unified framework for emotion and vision understanding. arXiv. Preprint posted online on May 10, 2025. [doi: 10.48550/arXiv.2505.06685]
20. Kobayashi M, Katayama M, Hayashi T, et al. Effect of multimodal comprehensive communication skills training with video analysis by artificial intelligence for physicians on acute geriatric care: a mixed-methods study. BMJ Open 2023 Mar 3;13(3):e065477. [doi: 10.1136/bmjopen-2022-065477] [Medline: 36868602]
21. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol 2014 Mar;67(3):267-277. [doi: 10.1016/j.jclinepi.2013.08.015] [Medline: 24275499]

Abbreviations

AI: artificial intelligence
CARE: Consultation and Relational Empathy
EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
IRB: institutional review board
TBQ: Treatment Burden Questionnaire
TBQ+D: Treatment Burden Questionnaire plus Digital
UCAT: Unhurried Conversations Assessment Tool

XSL•FO
RenderX

Edited by J Sarvestan; submitted 23.Sep.2025; peer-reviewed by MH Tsai; revised version received 19.Feb.2026; accepted 20.Feb.2026; published 24.Mar.2026.

Please cite as:
Al Zahidy M, Guevara Maldonado K, Vilatuna Andrango L, Proano AC, Claros AG, Lizarazo Jimenez M, Gomez-Alvarez E, Toro-Tobon D, Montori VM, Ponce-Ponte OJ, Brito JP
Longitudinal and Multimodal Recording System to Capture Real-World Patient-Clinician Conversations for AI and Encounter Research: Protocol for an Observational Study
JMIR Res Protoc 2026;15:e84688
URL: https://www.researchprotocols.org/2026/1/e84688
doi:10.2196/84688

© Misk Al Zahidy, Kerly Guevara Maldonado, Luis Vilatuna Andrango, Ana Cristina Proano, Ana Gabriela Claros, Maria Lizarazo Jimenez, Esteban Gomez-Alvarez, David Toro-Tobon, Victor M Montori, Oscar J Ponce-Ponte, Juan P Brito. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development and Formative Usability Evaluation of a Theory-Driven Progressive Web Application for Young Adult Wellness Engagement (MiCARE): Protocol for a Mixed Methods Study

Ayesha Thanthrige, BSc, MBA; Nilmini Wickramasinghe, PhD

School of Computing, Engineering and Mathematical Sciences, La Trobe University, Kingsbury Drive, Bundoora, Victoria, Australia

Corresponding Author:
Ayesha Thanthrige, BSc, MBA
School of Computing, Engineering and Mathematical Sciences, La Trobe University, Kingsbury Drive, Bundoora, Victoria, Australia

Abstract

Background: Young adults face rising wellness challenges, including prediabetes risk, requiring sustained engagement with preventive health interventions. Digital wellness applications offer promise for promoting healthy lifestyle behaviors, yet high dropout rates and inadequate personalization limit their effectiveness. This paper outlines the technical implementation and formative usability evaluation of MiCARE, a theory-driven progressive web application (PWA) designed to support sustained wellness engagement among young adults through user-centered design.

Objective: This study aims to systematically implement theory-driven design specifications into a functional web application, the MiCARE platform, and to conduct a formative usability evaluation with a convenience sample of 20 university-affiliated young adults aged 18 to 34 years in Victoria, Australia, in both rural and urban areas using the task-technology fit and unified theory of acceptance and use of technology frameworks as organizing lenses to assess usability, usefulness, and satisfaction.

Methods: This is an embedded mixed methods study conducted across 2 phases: phase 3 and phase 4. Phase 3 involves the technical implementation of 6 theory-driven features (ie, empathetic chatbot, learning hub, dynamic goal setting, gamification, personalized reminders, and progress dashboard) using HTML5, CSS3, JavaScript, Google Dialogflow ES, and Firebase services, following the Agile methodology over 6 months with biweekly self-managed sprints and clinical verification. Phase 4 is a 3-month formative usability feasibility evaluation with 20 young adults recruited from La Trobe University (Bundoora and Bendigo campuses). Participants will complete screening and initial, midpoint, and final surveys assessing usability, usefulness, and satisfaction, while real-time use analytics captures engagement patterns. Data analysis will use the task-technology fit and unified theory of acceptance and use of technology frameworks as interpretive guides, with quantitative data analyzed using descriptive statistics (R Studio) and qualitative feedback analyzed through thematic analysis (NVivo). Use analytics will provide descriptive contextual information only. The study has received ethics approval from the La Trobe University Human Research Ethics Committee (HEC24507).

Results: The study will take place between 2025 and 2026. Phase 3 (technical implementation) commenced in October 2025 and is currently ongoing, with core features under active development and verification. Phase 4 (formative usability and feasibility evaluation) is scheduled to commence following completion of phase 3. Evaluation results will be disseminated in academic forums and peer-reviewed publications in early 2027. The findings will enable us to evaluate the feasibility, acceptability, and usability of a theory-driven PWA in this university-affiliated sample, informing refinements and future larger-scale studies.

Conclusions: This study will contribute to the technical implementation and formative usability evaluation of a multitheoretical, user-centered PWA for wellness engagement in preventive health, bridging the gap between conceptual frameworks and deployed interventions.

International Registered Report Identifier (IRRID): PRR1-10.2196/86515

(JMIR Res Protoc 2026;15:e86515) doi:10.2196/86515

KEYWORDS

wellness engagement; young adults; progressive web application; prediabetes prevention; formative usability evaluation; mixed methods study; self-determination theory; task-technology fit; unified theory of acceptance and use of technology; digital health intervention

XSL•FO
RenderX

Introduction

Background

Over recent years, mobile health (mHealth) technologies have become integral to contemporary wellness promotion, with more than 350,000 health applications available globally [1]. Wellness applications represent prevalent mHealth tools designed to promote healthy lifestyle behaviors, such as improved diet and physical activity [2]. In Australia, approximately 88% of the population owns a smartphone [3], offering significant opportunities for large-scale digital health interventions.

Despite this proliferation, sustained engagement remains a critical challenge. Meta-analyses report dropout rates exceeding 43% across mHealth interventions for chronic disease management [4]. Among young adults aged 18 to 34 years, defined by the Australian Communications and Media Authority as digital natives [5], engagement barriers are particularly pronounced due to inadequate personalization [6], cultural insensitivity [7], and a lack of theoretical grounding in application design [8]. Building on the completed phase 1 (systematic literature review of 32 studies, expert consultations [n=6], formative research with young adults [n=20], and rapid review of user interface [UI] or user experience [UX] design practices [35 studies]) and phase 2 (wireframe development and preliminary evaluation using interactive Figma prototypes refined through iterative expert feedback), the MiCARE progressive web application (PWA) aims to address these gaps through a theory-driven, user-centered approach.

Young adults face rising chronic disease risks, particularly prediabetes, which affects approximately 374 million adults globally [9] and can be reversed through early lifestyle intervention [10]. Sedentary behavior compounds these risks, with adults spending an average of 8.8 hours per day sedentary across European countries [11], while Australians sit for approximately 8 hours per day [12]. Poor dietary habits further exacerbate health risks, with nearly half of Australian adults (49%) failing to consume recommended fruit servings and more than 90% not meeting vegetable intake guidelines [13].

Despite these challenges, digital wellness solutions offer unique advantages for young adults, including flexible access to tailored tools and seamless integration into daily routines [14]. However, most commercially available applications lack evidence-based design and theoretical grounding [8,15], compromising personalization and cultural relevance and resulting in poor sustained engagement.

Theoretical Foundation and Prior Development

The MiCARE framework synthesizes multitheoretical models to address engagement gaps: self-determination theory (SDT) [16], which emphasizes autonomy, competence, and relatedness as core constructs underpinning intrinsic motivation; the CARE (compassion, assistance, respect, and empathy) framework [17], which guides culturally sensitive engagement through compassion, assistance, respect, and empathy; user-centered design (UCD) principles [18], which emphasize iterative co-design and responsiveness to user preferences; inclusive design principles [19], which ensure equitable access and representation for diverse populations; design science research methodology (DSRM) [20,21], which provides a systematic approach to developing and evaluating information systems artifacts; and task-technology fit (TTF) [22] and Unified Theory of Acceptance and Use of Technology (UTAUT) [23], frameworks for assessing technology acceptance and sustained use.

Prior systematic work established the foundation for MiCARE development. A systematic literature review of 32 studies identified key engagement barriers, including high dropout rates [24,25], inadequate personalization [26,27], environmental constraints [7,28], and cultural or language barriers [29,30]. Facilitators included cultural tailoring [29,30], personalized feedback [25,31], user-friendly design [32,33], and peer support [34,35].

Building on these findings, a 3-step UCD process included systematic literature review (32 studies), expert consultations (n=6), formative research with young adults (n=20), and rapid review of UI and UX design practices (35 studies). This process generated 5 design objectives: empathy-driven interaction, equity-focused accessibility, culturally sensitive personalization, incremental goal setting, and intuitive onboarding. User preferences identified during earlier formative design research with young adults indicated that 55% preferred grid layouts, 65% favored large-button interactions, and 50% preferred minimalist interface designs, informing the feature specifications.

The Implementation and Evaluation Gap

While conceptual frameworks for digital health interventions are well established [36,37], the technical translation from theory to deployable software constitutes a critical knowledge gap [38]. Published studies typically present either fully developed interventions without implementation details or propose frameworks without functional prototypes. For example, LeSeure et al [29] developed the DiaFriend mobile app for type 2 diabetes management but noted that “the backend is incomplete and requires expert collaboration for full functionality.” Similarly, Curtis et al [39] documented UCD processes for a childhood weight management app but provided limited technical specifications, constraining reproducibility. This implementation gap is particularly problematic for resource-constrained research settings. Comprehensive technical documentation enables future researchers to adapt proven implementations rather than rebuilding from conceptual frameworks, thereby accelerating evidence-based innovation in digital health [40].

Furthermore, rigorous evaluation of theory-driven wellness applications remains limited. While the TTF [22] and UTAUT [23] provide validated frameworks for assessing technology adoption, few studies systematically apply these models to evaluate user experience outcomes in preventive health contexts [38]. The integration of multiple theoretical perspectives (SDT, CARE, UCD, and inclusive design) with evaluation frameworks (TTF and UTAUT) addresses calls for theory-driven development [36,37] and rigorous evaluation methodologies [40] in digital health.

XSL•FO
RenderX

Study Aims

This paper presents a protocol for a 9-month project with the following aims:

- Phase 3 (months 1‐6)—systematically implement theory-driven design specifications into a functional PWA, documenting the technical translation process from design rationale to deployable code
- Phase 4 (months 7‐9)—conduct a formative usability, acceptability, and feasibility evaluation with a convenience sample of university-affiliated young adults from La Trobe University, using the TTF and UTAUT frameworks as organizing lenses to assess usability, usefulness, and satisfaction

We hypothesize that a theory-driven, user-centered digital health platform will demonstrate high usability, perceived usefulness, and user satisfaction among young adults, supporting sustained engagement with wellness content over a 3-month evaluation period. This study assesses formative usability and feasibility to identify design strengths, usability issues, and refinement priorities, informing subsequent development and larger-scale evaluation.

Methods

Overview

The development and evaluation of the platform adopt a co-design, user-centric methodological approach [40], weighing all stakeholders' experiential needs and preferences equally. The implementation process follows the DSRM [20,21], integrating behavioral theory (SDT and CARE framework) with UCD principles [18] and inclusive design [19] to optimize uptake and use of the digital platform.

Integrating digital health technologies into practice has the potential to impact current practices [41]. The design process is guided by multiple theoretical frameworks, including SDT for motivation and autonomy, CARE for empathetic engagement, UCD for iterative co-design, and inclusive design for accessibility, ensuring theoretical fidelity throughout development. The evaluation uses the TTF [22] and the UTAUT [23] frameworks to match stakeholder expectations and digital literacy to the final platform design. The evaluation follows established usability testing guidelines [42], with sample sizes appropriate for identifying usability issues and achieving qualitative saturation [43].

Study Design

The study uses a mixed methods approach across 2 phases conducted over 9 months. An overview of the study design is shown in Figure 1. Phase 1 [44] (systematic literature review, expert consultations, formative research, and UI and UX review) and phase 2 (wireframe development and prototype evaluation, which has been accepted for publication) have been completed, providing the theoretical and design foundation for the subsequent phases. Phase 3 (October 2025 to March 2026) has commenced and is currently underway, involving the technical implementation of the MiCARE platform. Phase 4 (April 2026 to June 2026) will follow the completion of phase 3 and will focus on formative usability and feasibility evaluation with 20 university-affiliated young adults from La Trobe University (Bundoora and Bendigo campuses). This single-site, convenience sample supports heuristic usability discovery and design validation.

XSL•FO
RenderX

Figure 1. Overview of phase 3 implementation and phase 4 evaluation process. AI: artificial intelligence; PWA: progressive web application; TTF: task-technology fit; UTAUT: unified theory of acceptance and use of technology; UX: user experience; WCAG: Web Content Accessibility Guidelines.

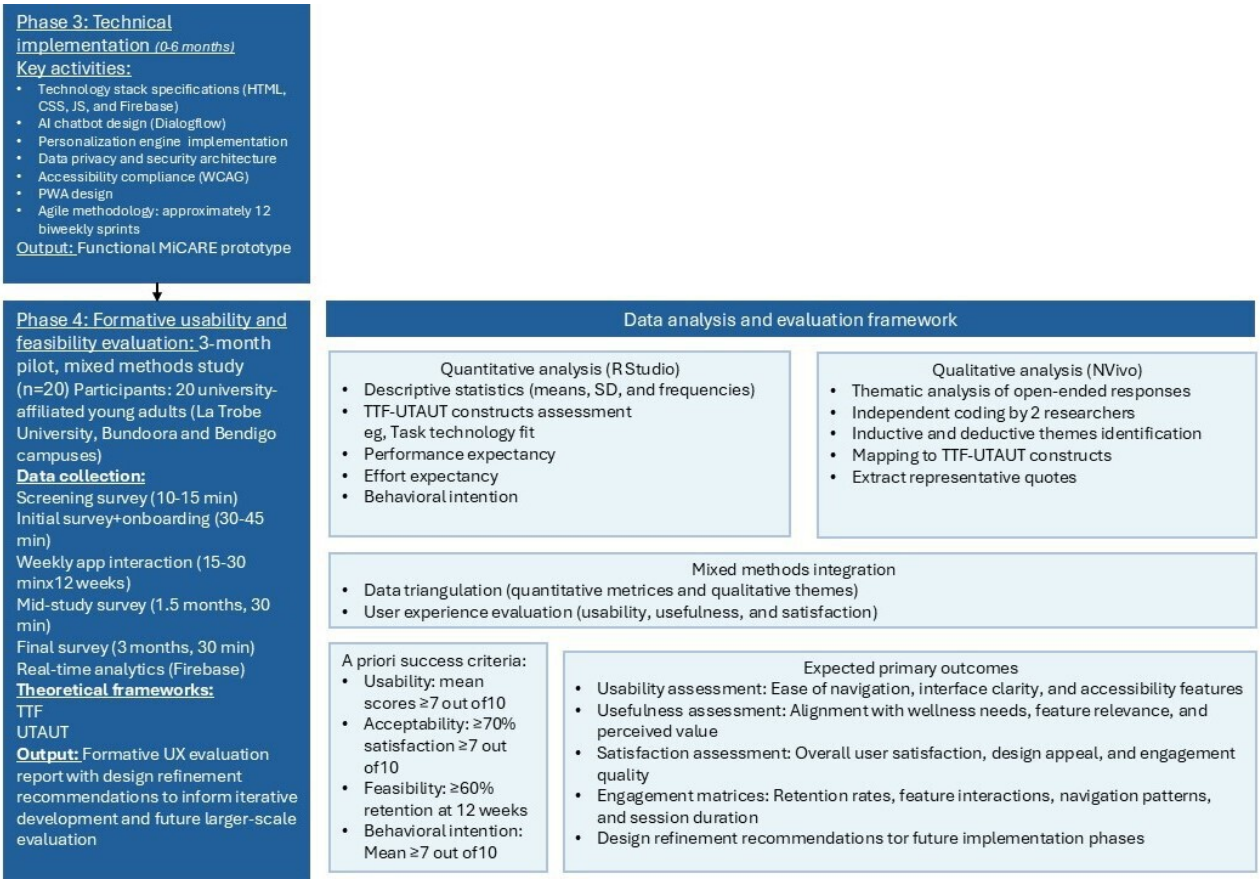

Participants and Setting

The stakeholder group for this study consists of young adults aged 18 to 34 years affiliated with La Trobe University. Participants will be recruited from La Trobe University (Bundoora and Bendigo campuses) through advertisement posters and email invitations.

Inclusion Criteria

Participants will be eligible for inclusion in the study if they meet the following criteria: (1) adults aged 18 to 34 years; (2) ownership of a computer or smartphone with stable internet connectivity; (3) ability to read, write, and understand English; and (4) willingness and ability to provide informed consent.

Exclusion Criteria

Participants will be excluded from the study if they meet any of the following criteria: (1) individuals with compulsive exercise behaviors or a history of disordered eating, as wellness content could be triggering; (2) inability to commit to a 3-month study duration; and (3) lack of basic digital literacy skills necessary to navigate web applications.

Phase 3: Technical Implementation (Months 1-6)

Rationale for Technical Approach

PWAs represent an emerging approach to mHealth delivery, combining the accessibility of web-based platforms with native app–like functionality [45]. Unlike traditional native applications requiring platform-specific development, PWAs use standard web technologies to deliver cross-platform experiences [2]. For digital health research, PWAs offer (1) simplified deployment without the need for app store approval, (2) seamless updates without user intervention, (3) reduced development costs through single codebase maintenance, and (4) broader reach across devices and operating systems [45].

Technology Stack Selection

The front end uses HTML, CSS, and JavaScript modules. This decision prioritizes transparency and reproducibility of digital health interventions. The application uses a modular, component-based architecture, enabling separation of concerns across feature-specific files. Accessibility compliance follows Web Content Accessibility Guidelines (WCAG) 2.1 Level AA standards [46], implementing high-contrast color schemes (dark blue #1E3A8A and white #FFFFFF), dyslexia-friendly fonts (Open Sans, minimum 14 pt), and semantic HTML5 elements, operationalizing inclusive design principles [19]. The backend uses Firebase, and the empathetic chatbot sends requests to Google Dialogflow ES for natural language understanding. Firebase provides the data infrastructure through 4 integrated services: Firestore, authentication (email and password and Google OAuth 2.0), storage, and analytics (interaction tracking).

Development Methodology

Implementation follows the Agile software development methodology [20] with 12 two-week sprints conducted over 6 months, consistent with DSRM principles of iterative design and evaluation [21]. The lead researcher functions as both

XSL•FO
RenderX

product owner and developer, maintaining close alignment between theoretical design specifications and technical implementation. Biweekly sprint reviews involve consultation with the research supervisor to validate theoretical fidelity across all frameworks, while periodic technical reviews with expert panel members (n=6) ensure adherence to accessibility and usability standards. Each sprint produces a deployable increment that is tested locally using Visual Studio (Microsoft Corporation).

The chatbot was initially configured with 50 predefined intents and corresponding responses. These responses were clinically verified against evidence-based sources, including the Diabetes Australia guidelines and the National Health and Medical Research Council recommendations, to ensure accuracy [17]. The key features of the MiCARE platform are summarized in Table 1.

Table . Mapping of design objectives to implemented features.

Design objective	Supporting theory	Implemented features
Empathy-driven interactions	SDT[a] and CARE[b]	Empathetic chatbot with clinically verified responses and personalized reminders with supportive language
Equity-focused accessibility	Inclusive design and CARE	WCAG[c] 2.1 AA compliance and multilingual interface capability at the architectural level, with English as the only language deployed and evaluated in this study
Culturally sensitive personalization	SDT and inclusive design	Symbolic visuals and culturally aligned responses (eg, halal meal suggestions)
Incremental goal setting	SDT	Dynamic goal-setting module, gamification (points and badges), and visual progress tracking
Intuitive onboarding	UCD[d] and UTAUT[e]	Guided video tour, learning hub, and minimalist dashboard with grid-based tiles

[a]SDT: self-determination theory.

[b]CARE: compassion, assistance, respect, and empathy.

[c]WCAG: Web Content Accessibility Guidelines.

[d]UCD: User-Centered Design.

[e]UTAUT: unified theory of acceptance and use of technology.

Core Feature Implementation

A total of 6 core features operationalize the multitheoretical MiCARE design objectives derived from phases 1 and 2.

Feature 1: Empathetic Chatbot

The chatbot integrates Google Dialogflow ES, operationalizing empathy-driven interactions grounded in the CARE framework [17]. User messages are sent from the front end (chatbot.js) to the Node.js backend (server.js), which forwards requests to Dialogflow ES and returns responses. Dialogflow ES is configured with 50 initial intents covering diabetes prevention topics, with fulfillment responses crafted to follow CARE framework principles (ie, compassion, assistance, respect, and empathy).

Feature 2: Learning Hub

The learning hub uses modular content delivery, where educational snippets are stored as JavaScript objects and dynamically rendered based on user authentication state, supporting SDT's competence construct through accessible evidence-based information [16]. Content was developed through a synthesis of Australian Dietary Guidelines (Diabetes Australia, 2024) and National Health and Medical Research Council physical activity recommendations, ensuring cultural relevance consistent with inclusive design principles [19]. Each snippet is concise (50 - 100 words) to align with user preferences for minimal clutter identified during formative research and UCD principles [18].

Feature 3: Dynamic Goal-Setting Module

The user-defined goal-setting module implements through Firestore subcollections, operationalizing SDT's autonomy and competence constructs [16]. Goals include attributes for type, target, progress, time stamps, and reset date.

Feature 4: Gamification System

The gamification system tracks points and badges stored in Firestore within each user's rewards subcollection, supporting SDT's competence and autonomy constructs through achievement recognition [16]. Points and badges are awarded automatically upon goal completion.

Feature 5: Personalized Reminders

Reminders are implemented as dismissible notification cards displayed contextually within the dashboard interface, preserving SDT's autonomy construct through full user control [16].

Feature 6: Progress Dashboard

The dashboard aggregates data from multiple Firestore subcollections (rewards, goals, and interactions) and presents summary statistics through tile-based layout aligned with user preferences identified during the UCD co-design process [18]. Real-time synchronization ensures displayed metrics reflect the

XSL•FO
RenderX

current state without requiring manual refresh through immediate feedback.

Data From Phase 3

Technical implementation data will be documented through (1) sprint logs that record implementation decisions and rationale, (2) a clinical verification report for chatbot responses, (3) expert panel feedback on accessibility compliance, and (4) development timeline.

Phase 4: Formative Usability and Feasibility Evaluation (Months 7-9)

Following completion of phase 3, a systematic formative usability and feasibility evaluation will be conducted with young adults (n=20) over a 3-month period. The evaluation uses the TTF [22] and UTAUT [23] frameworks as descriptive and exploratory organizing lenses.

Recruitment

A total of 20 young adults aged 18 to 34 years will be recruited from La Trobe University (covering both rural and urban areas) through advertisement posters placed in common areas and follow-up emails to those expressing initial interest. The recruitment time frame will span 3 months, allowing sufficient time to reach the target sample size and follow up with potential participants. Sample sizes are informed by established usability testing guidelines [42,47] and qualitative saturation guidelines [43], which suggest that 15 to 20 participants can identify approximately 90% of usability issues. If the recruitment target is not achieved, the recruitment period may be extended.

Data Collection

An online screening survey (10‐15 min) administered via QuestionPro (a platform licensed by La Trobe University) will determine eligibility. Potential participants will receive an email invitation with a link to the survey, accessible on various devices. Survey responses will be reviewed by authorized research personnel to confirm eligibility. Eligible participants will complete an initial survey (30‐45 min) collecting baseline data on familiarity with digital applications, current use patterns of similar applications, and expectations for the web application interface. This will be followed by accessing the web application via a provided link, with guidance from the research team if needed. Participants will be assigned a unique ID code for logging into the web application instead of using email or personal information.

Over the 3-month study period, participants will regularly interact with the web application's interface and features, including navigating information pages, using the reminder system, using the goal-setting feature, and engaging with gamification elements. Weekly interaction is expected to be 15 to 30 minutes per week. These interactions will help evaluate usability and user satisfaction aspects of the design. Midway through the study (after 1.5 months), participants will complete a mid-study survey (30 min) providing feedback on their experience with the interface and specific design elements. At the end of the study, participants will complete a final survey (30 min) evaluating overall user experience and satisfaction with the web application design.

Real-time use analytics will be extracted from Firebase, including retention rates, feature interactions, navigation patterns, and session duration. These data will be analyzed descriptively to contextualize self-reported findings. Over the 3-month period, participants are expected to dedicate approximately 6 to 9 hours in total.

Survey Instruments

Survey instruments have been developed based on the TTF [22] and UTAUT [23] constructs, assessing usability, usefulness, and satisfaction.

Usability

Usability is assessed using the following aspects: ease of navigation, interface clarity and organization, learning curve and intuitiveness, effectiveness of accessibility features, and technical reliability.

Usefulness

Usefulness is assessed using the following aspects: relevance of content and features, alignment with wellness goals, perceived value of personalized feedback, helpfulness of chatbot interactions, and effectiveness of goal-setting tools.

Satisfaction

Satisfaction is assessed using the following aspects: overall satisfaction with interface design, visual appeal and aesthetic preferences, emotional response to interactions, likelihood of continued use, and likelihood of recommending the platform to others.

Quantitative and Qualitative Measures

Surveys use both quantitative scales (1‐10 rating scales and Likert scales) and qualitative open-ended questions that allow participants to share detailed feedback, insights, and suggestions about their experience with the web application interface and design elements.

Evaluation Frameworks

The TTF [22] and UTAUT [23] are used as organizing frameworks to structure data collection and interpretation.

The TTF [22] assesses the degree to which technology assists individuals in performing their tasks. For MiCARE, TTF evaluation examines the alignment between young adults' wellness self-management needs and the platform's features. Key constructs include task characteristics (eg, goal setting and information seeking), technology characteristics (eg, feature functionality and interface design), fit between tasks and technology, and performance impacts (eg, ease of wellness management).

The UTAUT [23] identifies factors influencing technology adoption. For MiCARE, UTAUT evaluation examines performance expectancy (perceived usefulness), effort expectancy (ease of use), social influence (peer recommendations), facilitating conditions (technical support and accessibility), and behavioral intention (likelihood of continued use).

XSL•FO
RenderX

A Priori Success Criteria and Decision Rules

To support transparent interpretation and inform progression decisions, the following a priori success thresholds are defined:

- Usability—mean usability scores of ≥7 out of 10 across navigation, intuitiveness, and reliability items
- Acceptability—≥70% of participants reporting satisfaction scores of ≥7 out of 10
- Feasibility—≥60% participant retention at 12 weeks
- Behavioral intention—mean intention-to-use score ≥7 out of 10

Failure to meet these thresholds will prompt targeted design refinements prior to further evaluation. Meeting these thresholds will justify progression to a larger, more diverse sample evaluation.

Data Analysis

Quantitative Data

Descriptive statistics will be calculated using RStudio, including means, medians, percentages, and frequency distributions for usability, usefulness, satisfaction, and behavioral intention measures.

Qualitative Data

Open-ended survey responses will be analyzed using thematic analysis supported by NVivo software (Lumivero LLC). Initial codes will be deductively derived from the TTF and UTAUT constructs, with inductive codes added for emergent themes. At least 2 researchers will independently code a selection of transcripts to achieve consensus and ensure scientific rigor.

Use Analytics

Descriptive summaries of retention rates, feature engagement frequency, session duration, and navigation patterns will be calculated to contextualize survey findings.

Data Triangulation and Handling of Divergence

Triangulation will be conducted by comparing self-reported survey responses, qualitative feedback, and objective use analytics at the feature level. In cases of convergence, aligned findings will be interpreted as reinforcing evidence of usability or acceptability. In cases of divergence (eg, high self-reported satisfaction but low objective feature engagement), qualitative open-ended responses will be examined to identify explanatory factors such as perceived usefulness without habitual use, usability friction, contextual constraints, or novelty effects. Divergent findings will not be resolved statistically but will be explicitly reported and interpreted as indicators of design refinement needs, consistent with formative mixed methods evaluation principles.

Application of the TTF and UTAUT Frameworks

Given the small convenience sample, the study is not powered to conduct statistical validation, hypothesis testing, or structural modeling of TTF or UTAUT constructs. These frameworks will be applied descriptively and exploratorily only to qualitatively map user feedback, usability perceptions, and engagement experiences to established theoretical domains. No inferential testing of relationships between constructs will be undertaken.

Outcomes

Primary Outcomes

The primary outcomes include the following measures: usability scores (ease of navigation, intuitiveness, and accessibility), usefulness ratings (relevance, perceived value, and feature helpfulness), and satisfaction levels (overall satisfaction, visual appeal, and likelihood of continued use).

Secondary Outcomes

The secondary outcomes include retention rates over 3 months, feature engagement patterns (chatbot use, goal setting, and badge achievements), technical issues encountered, barriers to and facilitators of sustained use, and suggestions for interface improvements.

Ethical Considerations

Overview

Ethics approval was obtained from the La Trobe University Human Research Ethics Committee (HEC24507) prior to commencing the phase 4 evaluation. This study focuses on UX design evaluation, assessing usability, usefulness, and satisfaction of the app. Clinical health outcomes are explicitly excluded from this research scope, and no health-related outcomes are being investigated. Participants will receive detailed information about the study objectives, procedures, data collection methods, and their rights as research participants before enrollment. Written informed consent will be obtained from all participants prior to participation. Participation in the study is voluntary, and participants may withdraw at any time without penalty. All collected data will be deidentified and stored securely to protect participant privacy and confidentiality. Only authorized members of the research team will have access to the data. Participants will not receive financial compensation for participation in this study.

Participants will be fully informed about the nature of the research, the types of data collected, how they will be used, and their rights, including the right to withdraw from the study. Informed consent forms will be provided and signed by each participant.

Data Management

Personal identifiers will be replaced with unique participant codes to anonymize data. All digital data will be stored on secure, encrypted drives (Research DataSpace and the La Trobe University P: Drive), while any incidental nondigital data will be kept in locked cabinets in secure, access-controlled rooms (PS1 212A) at La Trobe University. Access to data will be limited to authorized research team members only. Data will be retained for 5 years after completion of the research and will then be securely destroyed, ensuring ongoing confidentiality and privacy.

Withdrawal

Participants may withdraw at any time until data are fully anonymized (expected 1 week after final survey submission). After this point, individual data cannot be identified or removed from the aggregated dataset.

XSL•FO
RenderX

Dissemination

Findings will be disseminated through (1) peer-reviewed journal publications in digital health, human-computer interaction, and health informatics venues; (2) conference presentations at workshops focusing on digital health, user experience design, and preventive health; (3) summary reports provided to participants highlighting key user experience insights; and (4) recommendations shared with La Trobe University and the broader digital health community.

The research team comprises transdisciplinary specialists in digital health innovation, implementation science, software engineering, human-computer interaction, and health informatics.

Results

The study is scheduled for 9 months between October 2025 and June 2026. Phase 3 (platform development) commenced in October 2025 and is currently ongoing, involving the implementation of 6 core application features. Phase 4 (formative usability testing) is scheduled to occur between April 2026 and June 2026 and will involve approximately 20 participants aged 18 to 34 years recruited from La Trobe University campuses. Data collection will include screening, baseline, midstudy, and final surveys, together with real-time interaction analytics captured through Firebase. Data analysis will involve descriptive statistical analysis of survey responses and thematic analysis of qualitative feedback. The first results are expected to be submitted for publication in early 2027.

Discussion

Anticipated Findings

This protocol documents the systematic technical implementation and formative usability evaluation of MiCARE, a theory-driven PWA designed to support wellness engagement among young adults. Through comprehensive specifications grounded in multiple theoretical frameworks—SDT [16], the CARE framework [48], UCD [18], inclusive design [19], DSRM [20,21], TTF [22] and UTAUT [23]—we demonstrate a replicable methodology for translating multitheoretical frameworks into functional digital health interventions and conducting early-stage usability assessment.

The implementation methodology prioritizes transparency through web technologies, replicability through standardized architectural patterns, and accessibility through WCAG compliance [46]. By documenting code-level design rationale alongside theoretical foundations, this work provides digital health researchers and developers with a practical template for translating engagement frameworks into deployable interventions.

Significance and Innovation

Few published studies document the technical translation from design specifications to functional software [38]. This protocol addresses the gap between conceptual frameworks and deployed interventions. This protocol provides comprehensive documentation of the full implementation process, enabling future researchers to adapt or extend the methodology.

The integration of behavioral theory (SDT and CARE) with implementation frameworks (DSRM, UCD, and inclusive design) and evaluation models (TTF and UTAUT) demonstrates a systematic approach to digital health development. This aligns with calls for theory-driven development in digital health [36,37] and addresses the need for rigorous evaluation methodologies [49]. This work is positioned as a design and implementation protocol accompanied by formative usability evaluation.

The JavaScript approach offers methodological advantages for research contexts. Unlike framework-dependent implementations, our approach remains interpretable without specialized knowledge, facilitating future maintenance by diverse research teams. This transparency aligns with open science principles and reproducibility standards in digital health research [50]. The Firebase-based data architecture demonstrates the practical application of serverless backend patterns in digital health research, minimizing DevOps complexity while maintaining scalability for future deployment.

Limitations

Several limitations constrain interpretation and generalizability. First, the phase 4 evaluation recruits only 20 participants from a single university (La Trobe), all of whom are English-speaking, limiting external validity. This convenience, single-site, student sample supports heuristic usability discovery only and does not permit generalizable claims about acceptability, adoption, effectiveness, equity, or cultural tailoring beyond this specific setting. Accordingly, the TTF and UTAUT frameworks are applied descriptively and exploratorily only, and the study is not powered to test or validate relationships between theoretical constructs. This limitation is appropriate and expected for formative usability testing. Second, the implementation has been tested only in local development environments without deployment to production infrastructure. The planned phase 4 usability evaluation will assess real-world performance under authentic use conditions. Third, the exclusion of non-English speakers from the study is a major limitation, reducing the applicability and usability of the digital platform by culturally and linguistically diverse groups. Although the system architecture is designed to support multilingual features, only the English-language interface is deployed and evaluated in this study.

Future Directions

Future work will build on the MiCARE platform's design and data from phases 3 and 4 by conducting larger-scale evaluation with more diverse samples. Emphasis will be placed on increasing user diversity by including culturally and linguistically diverse populations and individuals with varying levels of digital literacy, addressing key barriers to equitable access identified in digital divide research. Factors such as socioeconomic status, internet access, and device availability, which exacerbate the digital divide, will be assessed to optimize platform uptake and use across diverse populations. A mixed methods approach will be used, combining quantitative data with qualitative insights to evaluate implementation outcomes

XSL•FO
RenderX

and user engagement. To enhance inclusivity, multilingual interface support (expanding beyond English) and simplified navigation for users with lower digital literacy will be integrated, leveraging the scalable Firebase architecture and WCAG 2.1 AA compliance established in phase 3. Recruitment will expand beyond La Trobe University to community-based settings across rural and urban Victoria, Australia, ensuring a more representative sample of young adults. Subsequent phases will incorporate powered hypothesis testing with prespecified behavioral end points and corresponding analysis plans. Findings will inform scalable deployment strategies and contribute to the evidence base for theory-driven digital health interventions, addressing the gap between conceptual frameworks and real-world implementation.

Conclusions

This study contributes a transparent design and implementation protocol for developing and evaluating MiCARE, a theory-driven PWA designed to foster sustained wellness engagement among young adults. By integrating SDT, CARE, UCD, inclusive design, DSRM, TTF, and UTAUT frameworks, this protocol bridges the critical gap between conceptual design and deployable digital health interventions. The systematic implementation of 6 core features, including an empathetic chatbot, learning hub, user-defined goal setting, gamification, personalized reminders, and progress dashboard, alongside a formative mixed methods usability evaluation aims to assess usability, usefulness, and satisfaction among a convenience sample of university-affiliated young adults. This work demonstrates a multitheoretical, user-centered PWA specifically tailored for preventive health engagement, addressing chronic disease risks such as prediabetes through culturally sensitive and accessible digital tools through a formative evaluation.

Acknowledgments

The authors extend their gratitude to the expert panel members for their valuable technical and clinical guidance throughout the development process and to the young adult participants who provided formative feedback that helped shape the user interface design.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AT and NW collaboratively designed the study and were instrumental in obtaining ethics approval from La Trobe University (HEC24507). AT drafted the initial manuscript, and NW reviewed and provided revisions. Both authors contributed to editing and finalizing the manuscript and approved the final version for submission.

Conflicts of Interest

None declared.

References

1. Mescher T, Hacker RL, Martinez LA, Morris CD, Mishkind MC, Garver-Apgar CE. Mobile Health Apps: Guidance for Evaluation and Implementation by Healthcare Workers. J Technol Behav Sci 2025;10(2):224-235. [doi: 10.1007/s41347-024-00441-7]
2. Sousa Basto P, Ferreira P. Mobile applications, physical activity, and health promotion. BMC Health Serv Res 2025 Mar 10;25(1):359. [doi: 10.1186/s12913-025-12489-z] [Medline: 40065339]
3. Oviedo-Trespalacios O, Nandavar S, Newton JDA, Demant D, Phillips JG. Problematic use of mobile phones in Australia…is it getting worse? Front Psychiatry 2019;10:105. [doi: 10.3389/fpsyt.2019.00105] [Medline: 30914975]
4. Meyerowitz-Katz G, Ravi S, Arnolda L, Feng X, Maberly G, Astell-Burt T. Rates of attrition and dropout in app-based interventions for chronic disease: systematic review and meta-analysis. J Med Internet Res 2020 Sep 29;22(9):e20283. [doi: 10.2196/20283] [Medline: 32990635]
5. ACMA. The digital lives of younger and older Australians.: Australian Communications and Media Authority URL: https://www.acma.gov.au/publications/2021-05/report/digital-lives-younger-and-older-australians [accessed 2025-10-05]
6. Lim SL, Tay MHJ, Ong KW, et al. Association between mobile health app engagement and weight loss and glycemic control in adults with type 2 diabetes and prediabetes (D'LITE study): prospective cohort study. JMIR Diabetes 2022 Sep 30;7(3):e35039. [doi: 10.2196/35039] [Medline: 36178718]

XSL•FO
RenderX

7. Nomura A, Noguchi M, Kometani M, Furukawa K, Yoneda T. Artificial intelligence in current diabetes management and prediction. Curr Diab Rep 2021 Dec 13;21(12):61. [doi: 10.1007/s11892-021-01423-2] [Medline: 34902070]
8. Szinay D, Jones A, Chadborn T, Brown J, Naughton F. Influences on the uptake of and engagement with health and well-being smartphone apps: systematic review. J Med Internet Res 2020 May 29;22(5):e17572. [doi: 10.2196/17572] [Medline: 32348255]
9. International Diabetes Federation. Diabetes facts and figures.: International Diabetes Federation URL: https://www.idf.org [accessed 2024-12-03]
10. Reed J, Bain S, Kanamarlapudi V. A review of current trends with type 2 diabetes epidemiology, aetiology, pathogenesis, treatments and future perspectives. Diabetes Metab Syndr Obes 2021;14:3567-3602. [doi: 10.2147/DMSO.S319895] [Medline: 34413662]
11. Loyen A, Clarke-Cornwell AM, Anderssen SA, et al. Sedentary time and physical activity surveillance through accelerometer pooling in four European countries. Sports Med 2017 Jul;47(7):1421-1435. [doi: 10.1007/s40279-016-0658-y] [Medline: 27943147]
12. Sit less, move more. Heart Foundation. URL: https://www.heartfoundation.org.au/healthy-living/physical-activity/sit-less-move-more [accessed 2025-10-05]
13. Poor diet in adults. Australian Institute of Health and Welfare. URL: https://www.aihw.gov.au/reports/food-nutrition/poor-diet/contents/poor-diet-in-adults [accessed 2025-10-05]
14. Aljuboori D, Clary LK, Alomairah SA, et al. Contextual adaptation of digital wellbeing interventions for young people: insights from a project in Saudi Arabia. Front Psychiatry 2024;15:1455962. [doi: 10.3389/fpsyt.2024.1455962] [Medline: 39957755]
15. Bear HA, Ayala Nunes L, Ramos G, et al. The acceptability, engagement, and feasibility of mental health apps for marginalized and underserved young people: systematic review and qualitative study. J Med Internet Res 2024 Jul 30;26:e48964. [doi: 10.2196/48964] [Medline: 39078699]
16. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]
17. Sinclair S, McClement S, Raffin-Bouchal S, et al. Compassion in health care: an empirical model. J Pain Symptom Manage 2016 Feb;51(2):193-203. [doi: 10.1016/j.jpainsymman.2015.10.009] [Medline: 26514716]
18. Norman DA, Draper SW. In: Hillsdale NJ, editor. User Centered System Design: New Perspectives on Human-Computer Interaction: Lawrence Erlbaum Associates; 1986.
19. Clarkson PJ, Coleman R, Keates S, Lebbon C. Inclusive Design: Design for the Whole Population, 1st edition: Springer London; 2003.
20. Hevner AR, March ST, Park J, Ram S. Design science in information systems research. MIS Q 2004 Mar 1;28(1):75-106. [doi: 10.2307/25148625]
21. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manag Inf Syst 2007 Dec;24(3):45-77. [doi: 10.2753/MIS0742-1222240302]
22. Goodhue DL, Thompson RL. Task-technology fit and individual performance. MIS Q 1995 Jun 1;19(2):213-236. [doi: 10.2307/249689]
23. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003 Sep 1;27(3):425-478. [doi: 10.2307/30036540]
24. Bailey-Davis L, Wood GC, Cook A, et al. Communicating personalized risk of diabetes and offering weight reduction program choice: recruitment, participation, and outcomes. Patient Educ Couns 2021 May;104(5):1193-1199. [doi: 10.1016/j.pec.2020.10.017] [Medline: 33097360]
25. Zhang J, Oh YJ, Lange P, Yu Z, Fukuoka Y. Artificial intelligence chatbot behavior change model for designing artificial intelligence chatbots to promote physical activity and a healthy diet: viewpoint. J Med Internet Res 2020 Sep 30;22(9):e22845. [doi: 10.2196/22845] [Medline: 32996892]
26. Sun H, Zhang K, Lan W, et al. An AI dietitian for type 2 diabetes mellitus management based on large language and image recognition models: preclinical concept validation study. J Med Internet Res 2023 Nov 9;25(2):e51300. [doi: 10.2196/51300] [Medline: 37943581]
27. Röhling M, Kempf K, Banzer W, et al. Prediabetes conversion to normoglycemia is superior adding a low-carbohydrate and energy deficit formula diet to lifestyle intervention-a 12-month subanalysis of the ACOORH trial. Nutrients 2020 Jul 7;12(7):2022. [doi: 10.3390/nu12072022] [Medline: 32646010]
28. Wang SCY, Nickel G, Venkatesh KP, Raza MM, Kvedar JC. AI-based diabetes care: risk prediction models and implementation concerns. NPJ Digit Med 2024 Feb 15;7(1):36. [doi: 10.1038/s41746-024-01034-7] [Medline: 38361152]
29. LeSeure P, Chin E, Zhang S. A culturally sensitive mobile app (DiaFriend) to improve self-care in patients with type 2 diabetes: development study. JMIR Diabetes 2024 Oct 21;9:e63393. [doi: 10.2196/63393] [Medline: 39432893]
30. Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. JAMA Intern Med 2023 Jun 1;183(6):589-596. [doi: 10.1001/jamainternmed.2023.1838] [Medline: 37115527]

XSL•FO
RenderX

31. Wu Y, Zhang J, Ge P, et al. Application of chatbots to help patients self-manage diabetes: systematic review and meta-analysis. J Med Internet Res 2024 Dec 3;26(10397):e60380. [doi: 10.2196/60380] [Medline: 39626235]
32. Nguy TT, et al. Designing AI-based conversational agent for diabetes care in a multilingual context. arXiv. Preprint posted online on May 20, 2026 URL: https://www.researchgate.net/publication/351745026_Designing_AI-based_Conversational_Agent_for_Diabetes_Care_in_a_Multilingual_Context [accessed 2026-03-17] [doi: 10.48550/arXiv.2105.09490]
33. Salari R, R Niakan Kalhori S, GhaziSaeedi M, Jeddi M, Nazari M, Fatehi F. Mobile-based and cloud-based system for self-management of people with type 2 diabetes: development and usability evaluation. J Med Internet Res 2021 Jun 2;23(6):e18167. [doi: 10.2196/18167] [Medline: 34076579]
34. Yeh MC, Lau W, Keady CA, et al. Evaluation of feasibility and acceptability of a web-based diabetes prevention program (DPP) for diabetes risk reduction in Chinese Americans in New York City. Front Public Health 2023;11:1199746. [doi: 10.3389/fpubh.2023.1199746] [Medline: 37333528]
35. Al‐Hamdan R, Avery A, Al‐Disi D, Sabico S, Al‐Daghri NM, McCullough F. Efficacy of lifestyle intervention program for Arab women with prediabetes using social media as an alternative platform of delivery. J of Diabetes Invest 2021 Oct;12(10):1872-1880. [doi: 10.1111/jdi.13531]
36. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Internet Res 2014 Jun 5;16(6):e146. [doi: 10.2196/jmir.3077] [Medline: 24905070]
37. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
38. Voorheis P, Bhuiya AR, Kuluski K, Pham Q, Petch J. Making sense of theories, models, and frameworks in digital health behavior change design: qualitative descriptive study. J Med Internet Res 2023 Mar 15;25:e45095. [doi: 10.2196/45095] [Medline: 36920442]
39. Curtis KE, Lahiri S, Brown KE. Targeting parents for childhood weight management: development of a theory-driven and user-centered healthy eating app. JMIR Mhealth Uhealth 2015 Jun 18;3(2):e69. [doi: 10.2196/mhealth.3857] [Medline: 26088692]
40. van Velsen L, Ludden G, Grünloh C. The limitations of user-and human-centered design in an eHealth context and how to move beyond them. J Med Internet Res 2022 Oct 5;24(10):e37341. [doi: 10.2196/37341] [Medline: 36197718]
41. Barr PJ, Haslett W, Dannenberg MD, et al. An audio personal health library of clinic visit recordings for patients and their caregivers (HealthPAL): user-centered design approach. J Med Internet Res 2021 Oct 22;23(10):e25512. [doi: 10.2196/25512] [Medline: 34677131]
42. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. Presented at: CHI '93: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems; May 1, 1993. [doi: 10.1145/169059.169166]
43. Guest G, Bunce A, Johnson L. How many interviews are enough?: An experiment with data saturation and variability. Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
44. Thanthrige A, Wickramasinghe N. Digital health solutions for type 2 diabetes and prediabetes: systematic review of engagement barriers, facilitators, and outcomes. JMIR Diabetes 2026 Mar 12;11:e80582. [doi: 10.2196/80582] [Medline: 41818488]
45. Przewięźlikowska A, Ślusarczyk W, Wójcik K, Ślusarski M. Efficient and scalable architecture for location-based mobile applications using metrica. Autom Constr 2025 Apr;172:106056. [doi: 10.1016/j.autcon.2025.106056]
46. Singh U, Divya Venkatesh J, Muraleedharan A, Saluja KS, J H A, Biswas P. Accessibility analysis of educational websites using WCAG 2.0. Digit Gov: Res Pract 2024 Sep 30;5(3):1-28. [doi: 10.1145/3696318]
47. Turner BO, Paul EJ, Miller MB, Barbey AK. Small sample sizes reduce the replicability of task-based fMRI studies. Commun Biol 2018;1:62. [doi: 10.1038/s42003-018-0073-z] [Medline: 30271944]
48. Malenfant S, Jaggi P, Hayden KA, Sinclair S. Compassion in healthcare: an updated scoping review of the literature. BMC Palliat Care 2022 May 18;21(1):80. [doi: 10.1186/s12904-022-00942-3] [Medline: 35585622]
49. Jakob R, Harperink S, Rudolf AM, et al. Factors influencing adherence to mHealth apps for prevention or management of noncommunicable diseases: systematic review. J Med Internet Res 2022 May 25;24(5):e35371. [doi: 10.2196/35371] [Medline: 35612886]
50. Hekler EB, Klasnja P, Riley WT, et al. Agile science: creating useful products for behavior change in the real world. Behav Med Pract Policy Res 2016 Jun;6(2):317-328. [doi: 10.1007/s13142-016-0395-7]

Abbreviations

CARE: compassion, assistance, respect, and empathy
DSRM: design science research methodology
mHealth: mobile health
PWA: progressive web application

XSL•FO
RenderX

SDT: self-determination theory
TTF: task-technology fit
UCD: user-centered design
UI: user interface
UTAUT: unified theory of acceptance and use of technology
UX: user experience
WCAG: Web Content Accessibility Guidelines

Edited by J Sarvestan; submitted 26.Oct.2025; peer-reviewed by G Jiang; revised version received 27.Jan.2026; accepted 29.Jan.2026; published 24.Mar.2026.

Please cite as:
Thanthrige A, Wickramasinghe N
Development and Formative Usability Evaluation of a Theory-Driven Progressive Web Application for Young Adult Wellness Engagement (MiCARE): Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e86515
URL: https://www.researchprotocols.org/2026/1/e86515
doi:10.2196/86515

© Ayesha Thanthrige, Nilmini Wickramasinghe. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Real-Time and Long-Term Effects of Medical Marijuana on Older Adults: Protocol for a Prospective Cohort Study

Kendall R Robinson[1], BHS; Stella D Seeger[2], MBA; Lauren Nave[3], MPH; Marlin Mejia[1], BS; Maria Vander Meulen[4], MS; Ahmed Rashid[5], MD; Angela M Mickle[6], MS, CCRP; Kimberly T Sibille[5,6], MA, PhD; Zhigang Li[7], PhD; Rene Przkora[5], MD, PhD, FASA; Siegfried Schmidt[8], MD, PhD; Margaret C Lo[9], MD, FACP; Robert L Cook[1], MPH, MD; Yan Wang[1], MS, PhD

[1]Department of Epidemiology, University of Florida, CTRB Bldg, 4th Fl., 2004 Mowry Rd, Gainesville, FL, United States

[2]McCombs School of Business, University of Texas, Austin, TX, United States

[3]Miller School of Medicine, University of Miami, Miami, FL, United States

[4]College of Medicine, University of Florida, Jacksonville, FL, United States

[5]Department of Anesthesiology, University of Florida, Gainesville, FL, United States

[6]Department of Physical Medicine and Rehabilitation, University of Florida, Gainesville, FL, United States

[7]Department of Biostatistics, University of Florida, Gainesville, FL, United States

[8]Department of Community Health and Family Medicine, College of Medicine, University of Florida, Gainesville, FL, United States

[9]Department of Medicine, Division of General Internal Medicine, College of Medicine, University of Florida, Gainesville, FL, United States

Corresponding Author:
Kendall R Robinson, BHS
Department of Epidemiology, University of Florida, CTRB Bldg, 4th Fl., 2004 Mowry Rd, Gainesville, FL, United States

Abstract

Background: Older adults represent the fastest-growing group of medical marijuana (MM) users in the United States, with chronic pain being the most common reason for use. Despite this trend, scientific evidence remains limited regarding the short- and long-term effects of MM on critical health outcomes, including cognitive function, physical and mental health, and overall quality of life, in this population. To better inform clinical practice and public policy, there is a clear need for more rigorous longitudinal studies that examine the impact of real-world MM products over time.

Objective: The Study on Medical Marijuana and Its Long-Term Effects on Older Adults (SMILE) is a prospective cohort study that aims to (1) determine MM's short- and long-term effects on pain, physical, emotional, and cognitive functioning, as well as quality of life, in older adults and (2) identify MM product characteristics and patient subgroups associated with improved outcomes and side effects.

Methods: This study will recruit and follow 440 older adults (aged ≥50 years; ~50% aged >65 years, ~50% male) with chronic pain for 12 months, as some initiate MM use (MM group, n=330) and others do not (comparison group, n=110). Data collection includes quarterly survey questionnaires (longitudinal changes in cannabis use, pain, physical and emotional functioning, side effects, and quality of life); baseline and 12-month cognitive assessments, pain sensory tests, and blood and urine samples for cannabis use; and periodic smartphone- and Fitbit sensor–based measurements to capture MM use patterns, real-time pain, mental health, and objective data on physical activity and sleep. Data will be analyzed using descriptive analyses, generalized linear mixed-effects models, and generalized estimating equation models to assess differences in short- and long-term effects between the MM and comparison groups, as well as subgroups among those initiating MM treatment.

Results: Recruitment for the SMILE study began in July 2022, and all data collection is expected to be completed by 2026. As of October 2025, we have enrolled 399 participants, with 277 in the MM group and 122 in the comparison group. Results are expected to be published starting in 2027.

Conclusions: With multisource data collected in real time and over 12 months, our study will provide much-needed scientific evidence addressing (1) whether MM can reduce pain and improve physical and emotional functioning in the short term among older adults; (2) whether the effects of MM last for 12 months and demonstrate changes in quality of life or cognition; and (3) whether health benefits and consequences differ by MM product type and whether individual differences (eg, sex, baseline pain phenotyping) moderate the relationship. Our findings will offer valuable insights for physicians and patients when considering MM as a treatment option and will help guide more informed, individualized care decisions.

International Registered Report Identifier (IRRID): DERR1-10.2196/78900

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e78900) doi:10.2196/78900

KEYWORDS

medical marijuana; MM; cannabis; older adults; chronic pain; ecological momentary assessment; EMA

Introduction

Medical marijuana (MM) is increasingly accessible in the United States. Its use is rising among the 36 million adults in the United States aged 45 years or older, with chronic pain accounting for 60.6% of qualifying conditions for MM use [1-4]. This increase in MM use could be attributed to state-level legalization of marijuana and older adults seeking MM as an alternative pain treatment, which reduces stigma and increases accessibility [5-7]. However, evidence on marijuana's long-term efficacy and safety in older adults is scarce [1,8-21], as most studies only include participants under the age of 60 [22]. This is concerning, given the association of aging with polypharmacy and substantial changes in pharmacokinetics and pharmacodynamics (eg, drug clearance rate) [5,23]. Therefore, data on short- and long-term health outcomes and side effects among older adults are urgently needed [24].

While marijuana's pharmacological profile supports a potential role in pain management, empirical evidence regarding its efficacy in treating chronic pain remains inconsistent. While some meta-analyses have reported significant beneficial effects of marijuana on chronic pain [25-27], others have found no or limited effect [28-30]. Similarly, the literature has shown that MM may help with chronic pain-related experiences—such as anxiety, depression, and opioid use—while also improving sleep and quality of life among individuals with chronic pain [31-34]. Despite the significant interpersonal differences in responses to pain treatment [35], important biological factors (eg, sex and baseline pain phenotyping), psychological factors (eg, marijuana expectancy and social desirability bias), and the use of objective measures such as sensors or biomarkers to complement self-reported outcomes [36-38] have received limited attention in marijuana and chronic pain research [22,39,40].

Research challenges—such as a reliance on the use of marijuana products not reflective of real-world usage of MM—may explain the inconsistent findings across studies [41]. While randomized controlled trials are considered the "gold standard" of evidence, the products used are often isolated or synthetic cannabinoids (eg, dronabinol) [42,43], which have limited variety and may not be comparable to MM products available in state dispensaries [32]. The ubiquitous access to MM at the state level dictates an urgent need to investigate the effects of these products, which usually have higher concentrations of tetrahydrocannabinol (THC) and/or cannabidiol (CBD) and are consumed via a greater diversity of methods (eg, vape, capsule, tincture, cream). The lack of standard guidance for MM use makes it difficult to accurately characterize the ratio of major components (THC:CBD), administration route, and dosage, all of which may influence the efficacy of MM on chronic pain [1,44,45]. There remains a critical need for systematic investigation of MM's effects on pain, physical and emotional functioning, and overall health outcomes (eg, quality of life), as well as the need to track potential side effects, such as its impact on cognitive functioning in this population [22,46,47].

To address these critical gaps in the literature, this study will build a prospective cohort of older adults aged 50 years or older with chronic pain, including those who initiate MM use and those who do not. This study will integrate both technology-based remote assessments (eg, smartphone-based ecological momentary assessment [EMA], Fitbit sensor) and quarterly in-person visits (eg, surveys, blood tests) to systematically examine the short- and long-term impacts of MM on 6 core chronic pain outcome domains: (1) pain intensity and interference; (2) physical functioning; (3) emotional functioning; (4) cognitive function; (5) participant-rated overall improvement; and (6) adverse events or side effects [48-50], 5 of which are recommended by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT).

The specific aims of this project are as follows: (1) determine whether MM use leads to short-term changes in pain intensity and physical and emotional functioning, as measured in real time by EMA (subjective data) and the Fitbit device (objective data); (2) evaluate whether MM use leads to longer-term changes over a 12-month period, based on self-reported pain intensity or interference, emotional and physical functioning, health-related quality of life, and cellular aging markers, such as telomere length; and (3) among individuals initiating MM, examine which MM product characteristics (ie, THC:CBD ratio, administration route, dose) are associated with greater improvements in the outcomes defined in Aims 1 and 2, or with a greater incidence of side effects. Additionally, the study will explore whether individual differences (eg, sex, baseline pain phenotype) moderate these relationships.

Methods

Ethical Considerations

The University of Florida (UF) Institutional Review Board (IRB) approved this study (IRB202101597) on July 29, 2021, and it was conducted in accordance with the Declaration of Helsinki.

All participants will provide written informed consent prior to any study procedures being conducted, with the option to withdraw consent at any time. The study team will send participants the informed consent form remotely through a secure Research Electronic Data Capture (REDCap) link for review before their baseline visit. Participants have the option to sign the consent form remotely or after discussing study details with study staff during the baseline visit. Following informed consent, participants will complete the Rapid Estimate of Adult Literacy in Medicine, Revised to assess health literacy [51]. If participants score lower than 6 out of 8 points on the Rapid Estimate of Adult Literacy in Medicine, Revised, study staff will re-review the informed consent forms and ask

XSL•FO
RenderX

follow-up questions to ensure that participants fully understand the study protocol.

Only IRB-approved study staff will have access to study information and participant data. Participants are assigned a unique study identifier to maintain confidentiality. Biospecimens are labeled with the unique study identifier, sample type, study time point, and date of collection. Whether a participant is in the MM group or the comparison group will not be disclosed to prevent bias during sample processing. All survey data will be kept in the secure REDCap system during the data collection phase. National Institutes of Health (NIH) Toolbox data will only contain the study ID as an identifier and will be emailed with encryption to the data manager for upload into REDCap. EMA data will only contain the study ID as an identifier and will be downloaded from cloud storage to a secured server at UF on a biweekly basis when active data collection is in process. Any photos of MM products (participants are instructed to avoid including any personal identifiers in the photo) uploaded by study participants during EMA surveys will be stored on a secured server. Fitbit data will be downloaded from a Fitbit database service (ie, Fitabase) via encrypted transmission. Research staff will use the unique study account to match it back to the participant's unique study ID so the Fitbit data can be merged with other data collected from the same participant. A certificate of confidentiality was automatically granted by the NIH to protect participant privacy because the study is funded by the NIH.

Participants will be compensated for their participation, including US $60 for completing the baseline visit, US $40 for each in-person quarterly follow-up visit, and US $60 for the 12-month follow-up visit. Participants are compensated up to US $280 for EMA and Fitbit and are paid in full if they complete 80% of the EMA assessments and wear their Fitbit 80% of the time during each period. If they do not reach this threshold, prorated payments are provided. Travel compensation of an additional US $20 is provided for participants who travel 20 or more miles for the baseline and 12-month visits. Remote options for 3-, 6-, and 9-month visits are offered for individuals experiencing transportation barriers. An additional US $200 payment is provided for participants in the MM group to offset initial costs associated with beginning MM treatment if they delay initiating MM treatment for at least 3 days to provide clean baseline data (without cannabis use) for EMA and Fitbit and then obtain an MM card through the state of Florida. Participants may keep the Fitbit device as part of their compensation after the 12-month follow-up visit or return it if they do not want it. Participants are compensated or given study merchandise of equal monetary value (~US $10, eg, umbrella, mug, hand sanitizer) if they fail screening procedures after enrollment.

Participants

Population

This study plans to recruit 440 older adult participants aged 50 years or older, with at least 50% over the age of 65, and with approximately 50% of the sample being male. We will try to achieve the balanced ratio by monitoring sample demographics throughout the study and adjusting recruitment techniques as needed to target necessary demographics. Among these, 330 participants will be seeking to start MM treatment during the study period, as indicated at baseline when they are enrolled in the study, and 110 participants will have no intention of starting MM treatment for at least 6 months. However, participants are free to initiate or stop MM use at any time during the study and continue participation after they join the study. If a participant decides to withdraw from the study or is lost to follow-up at any point, data will be analyzed up to the point of withdrawal.

Inclusion Criteria

Potential participants will be eligible for the Study on Medical Marijuana and Its Long-Term Effects on Older Adults (SMILE) if they (1) are 50 years or older; (2) have had noncancer chronic musculoskeletal or neuropathic pain [52] for 3 or more months [53], with a pain intensity level in the past week greater than zero ("0=no pain" to "10=most pain imaginable") [54]; (3) are seeking MM treatment (for the MM group) or have no intention to start MM for at least 6 months (for the comparison group); and (4) have access to a smartphone.

Individuals aged 50 years and older were chosen due to the rising use of MM for chronic pain and the need for research on this population [5,6]. Participants must have current noncancer chronic musculoskeletal or neuropathic chronic pain, confirmed by verification of pain for 3 or more months at an intensity level greater than zero. Intention to initiate MM treatment is assessed at recruitment to determine which group enrolled participants will be in. To obtain baseline data before the initiation of MM treatment, participants must not be receiving MM treatment at the time of enrollment. Individuals who are not interested in MM should ideally not initiate MM treatment within the 6 months following enrollment. Access to a smartphone is necessary for completing the EMA and Fitbit aspects of the study. Because all assessments and questionnaires that will be used are in English, participants must be able to read English to complete study measures.

Exclusion Criteria

Potential participants will be excluded if they (1) are currently pregnant or nursing (by self-report); (2) self-report marijuana or CBD use in the past 30 days, or if the urine test is positive for THC; (3) report any history of using MM regularly (daily use); (4) used marijuana regularly (daily use) for over a month in the past 10 years; (5) have current or past substantial or severe substance use disorders [55], as assessed by the Drug Abuse Screening Test [56,57]; (6) are undergoing active cancer treatment, excluding certain cancers such as skin cancer; (7) have a terminal illness; (8) have neurological conditions that impact their ability to consent or complete study procedures (eg, Alzheimer disease, Parkinson disease, multiple sclerosis, head injury, stroke, and/or seizures); or (9) have active, uncontrolled psychiatric disorders (eg, schizophrenia, bipolar disorder, excluding anxiety and depression). Many of these exclusion criteria are imposed to ensure participant safety, the ability to provide consent, and the completion of study procedures.

XSL•FO
RenderX

Recruitment and Consent

The primary recruitment method for both MM and comparison groups will be through UF's health care system contact registry (Consent2Share), where patients have consented to be contacted for research. The research team will first contact patients with chronic pain in the contact registry via email or messages within the MyChart patient portal; those who are interested will then be prescreened for preliminary eligibility over the phone. The UF Pain Research and Intervention Center of Excellence (PRICE) registry (IRB201400126) will also be used to recruit participants who have previously consented to be contacted about future research studies.

Additional participant recruitment will take place at various chronic pain, primary care, and MM clinics across the state of Florida, as well as the Veterans Affairs (VA) Medical Center. The research coordinator or staff has a designated space at clinics to determine preliminary eligibility, explain details of the study, and schedule an in-person baseline visit. Clinic staff or clinicians will distribute study flyers and contact forms to their patients and waiting areas. The VA coordinators will also prescreen potential research participants via VA medical records or physician referrals and contact potential participants via mail, email, or phone calls. Clinics may also send electronic versions of IRB-approved flyers and recruitment materials to their scheduled patients or direct patients to fill out an online form to express interest in the study. We will recruit in community settings via flyers in research facilities and clinics and through radio advertising. All assessments will be conducted at research facilities in Gainesville or Jacksonville, Florida.

Study Design

Overview

All participants will complete an in-person screening, including a urine sample to confirm their marijuana use status, verification of musculoskeletal or neuropathic chronic pain using medication bottles or medical records, and a screening questionnaire. Participants will complete in-person assessments at baseline and 12 months, including survey questionnaires, cognitive testing, blood pressure measurement, and biomarker collection (blood samples for cannabis metabolite tests, telomere length, and inflammation/inflammasome markers). After the initial baseline visit, participants will complete technology-based real-time assessments using a smartphone-based EMA app to measure subjective elements of pain intensity, emotional and physical functioning, and sleep. Participants will also be provided with a wearable Fitbit sensor to record objective data on heart rate, physical activity, and sleep. The EMA surveys and Fitbit data will capture real-time short-term effects during the first month and for 1 week every 3 months over 12 months. Biospecimen collection (urine tests and blood samples) and a follow-up survey will be repeated every 3 months. Participants are encouraged to provide feedback regarding study procedures throughout their participation to inform future revisions to the protocol. Figure 1 provides an overview of the study procedures. A detailed timeline for each measure and assessment conducted throughout the course of the study can be found in Table 1.

Figure 1. Study procedure overview. EMA: ecological momentary assessment; MM: medical marijuana.

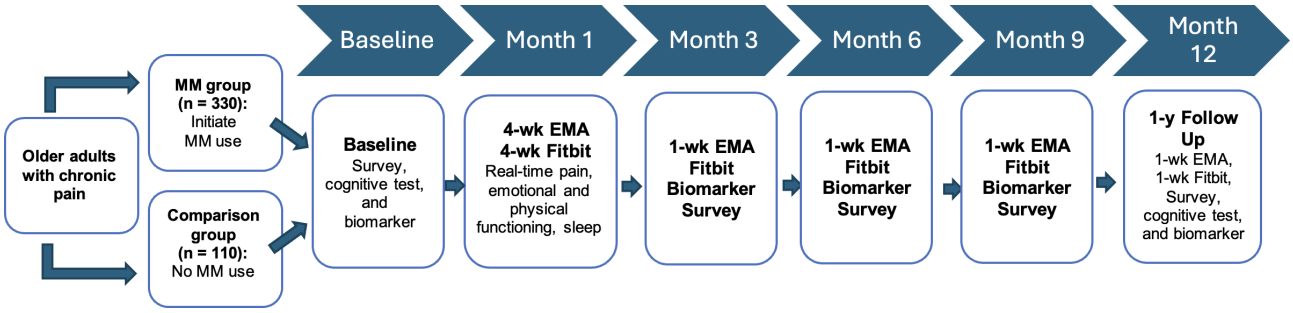

XSL•FO
RenderX

Table . Summary of study measures and assessments.

Measure	Baseline	3, 6, and 9 months	12 month	Source
Screening				
Presence of tetrahydrocannabinol (THC) in urine	✓	✓[a]	✓[a]	Urine test
Current/history of substance use disorders	✓			DAST-10[b] [56,57]
Health literacy	✓			REALM-R[c] [51]
Biometric data/biospecimen				
Height and weight	✓		✓	N/A[d]
Blood pressure	✓	✓	✓	N/A
Telomere length	✓		✓	Blood sample: buffy coat [58-60]
Cannabinoids and metabolites	✓	✓	✓	Blood sample: cannabinoids/metabolites assay [61-63]
Inflammation and neurodegeneration	✓	✓	✓	Blood sample: inflammasome and neurodegeneration biomarker test (exploratory)
Pain sensory tests				
Pain sensitivity	✓		✓	Punctate mechanical stimuli test and pressure pain threshold [64]
NIH[e] Toolbox Cognition Battery				
Executive function	✓		✓	NIH Toolbox dimensional change card sort test and NIH Toolbox Flanker inhibitory control and attention test [65]
Processing speed	✓		✓	NIH Toolbox pattern comparison processing speed test [65]
Episodic memory	✓		✓	NIH Toolbox picture sequence memory test [65]
Working memory	✓		✓	NIH Toolbox list sorting working memory test [65]
Language	✓			NIH Toolbox picture vocabulary test and NIH Toolbox oral reading recognition test [65]
Other assessments				
Real-time outcomes (pain, mood, sleep) and medication/MM[f] use	4 wk	1 wk	1 wk	EMA[g] smartphone app and Fitbit; BSI[h] [66]

[a]Urine test at follow-up does not affect eligibility.

[b]DAST-10: Drug Abuse Screening Test.

[c]REALM-R: Rapid Estimate of Adult Literacy in Medicine, Revised.

[d]N/A: not applicable.

[e]NIH: National Institutes of Health.

[f]MM: medical marijuana.

[g]EMA: ecological momentary assessment.

[h]BSI: Brief Symptom Inventory.

XSL•FO
RenderX

Group Assignment

Participants will be assigned to the MM group or the comparison group based on their responses to baseline screening questions regarding current interest or disinterest in starting MM treatment. Group assignment will remain consistent for each participant, regardless of whether they change their mind throughout the course of the study. Participants interested in initiating MM treatment will be given information about local MM physicians who can certify them for a Florida MM card. Additional compensation for delaying MM treatment to enroll in the study will also be provided to participants who begin MM treatment within 4 weeks after their baseline visit to cover most of the cost for MM physician visits and the initial charge to obtain their MM card in Florida.

Surveys at Baseline and 3-, 6-, 9-, and 12-Month Follow-Ups

Surveys at all time points take approximately 30 to 45 minutes for participants to complete. The length of the visit can be a barrier for older adults, especially those with chronic pain. In the interest of saving participants' time and preventing fatigue during the in-person assessment, participants will have the option to complete the survey remotely before their visit. Participants are offered snacks and beverages during in-person visits and can take as many breaks as needed. The survey includes questions such as demographics, health status measured using the Charlson Comorbidity Index and health history [67], medication and supplement use measured using the Medication Quantification Scale, patient-reported medication symptoms adapted for MM, nonpharmacological treatments [68-70], self-rated health, mental well-being, history of and current nonmedical marijuana use, and substance use. Established scales will be used to assess variables such as pain measured by the Brief Pain Inventory, Graded Chronic Pain Scale (GCPS), Patient Global Impression of Change, and Patient-Reported Outcomes Measurement Information System (PROMIS) Pain Interference Scale [71-74]; marijuana use measured by the comprehensive marijuana motives questionnaire adapted by Lankenau et al [75-78], Medical Cannabis Expectancy Questionnaire, and Cannabis Use Disorders Identification Test-Revised; substance use measured by the World Health Organization-Alcohol, Smoking and Substance Involvement Screening Test [79]; quality of life measured by the PROMIS Global-10 [80]; depression and anxiety measured by the PROMIS Anxiety and Depression Scales [81]; social desirability measured by the Marlowe-Crowne Social Desirability Scale [82]; post-traumatic stress disorder measured by the Primary Care-Post-Traumatic Stress Disorder Scale [83]; and sleep quality measured by the Pittsburgh Sleep Quality Index [84]. A full list of core survey assessments can be found in Table 2, and information on how these measures are related to each study aim is outlined in the *Analytical Strategy* section. Participants will complete the survey independently or with assistance from research staff.

XSL•FO
RenderX

Table . Summary of measures included in the core and quarterly follow-up surveys.

Measure	Baseline	3, 6, and 9 months	12 month	Source
Demographics, income, and insurance	✓			N/A[a]
Medical history	✓	✓	✓	CCI[b] and health history [67]
Current medication use	✓	✓	✓	MQS[c] [68]
Marijuana use				
Marijuana use motive	✓			CMMQ[d]—adapted by Lankenau et al [75,78]
MM[e] attitudes and perceptions	✓		✓	N/A
Positive and negative MM expectations	✓			MCEQ[f] [76]
Cannabis use disorders	✓		✓	CUDIT-R[g] [77]
List of MM products[h]		✓	✓	N/A
Recreational marijuana use	✓	✓	✓	N/A
Marijuana-related adverse events		✓	✓	PRMS[i]—adapted for MM[h] [70]
Pain				
Pain assessment	✓	✓	✓	BPI[j] [71]
Pain severity	✓		✓	GCPS[k] [72]
Change in pain			✓	PGIC[l] [73]
Chronic pain stage	✓		✓	Pain FITT[m] [37,58,85]
Pain interference	✓	✓	✓	PROMIS[n] Pain Interference Scale-Version 4a [74]
Nonpharmacological treatments	✓	✓	✓	Nonpharmacological therapies for chronic pain [69]
Emotion				
Depression	✓	✓	✓	PROMIS Depression Scale-8a [81]
Anxiety	✓	✓	✓	PROMIS Anxiety Scale-8a [81]
Health-related quality of life	✓	✓	✓	PROMIS Global-10 [80]
Social desirability	✓			MCSDS[o] [82]
PTSD[p]	✓			4-item PC-PTSD[q] Scale [83]
Other				
Sleep	✓	✓	✓	Pittsburgh Sleep Quality Index [84]
Substance use	✓		✓	WHO-ASSIST[r] [79]

[a]N/A: not applicable.

[b]CCI: Charlson Comorbidity Index.

[c]MQS: Medication Quantification Scale.

[d]CMMQ: Comprehensive Marijuana Motives Questionnaire.

[e]MM: medical marijuana.

XSL•FO
RenderX

[f]MCEQ: Medical Cannabis Expectancy Questionnaire.

[g]CUDIT-R: Cannabis Use Disorders Identification Test-Revised.

[h]Administered to the Medical Marijuana group only.

[i]PRMS: patient-reported medication symptoms.

[j]BPI: Brief Pain Inventory.

[k]GCPS: Graded Chronic Pain Scale.

[l]PGIC: Patient Global Impression of Change.

[m]Pain FITT: pain frequency, intensity, time, and total number of pain sites.

[n]PROMIS: Patient-Reported Outcomes Measurement Information System.

[o]MCSDS: Marlowe-Crowne Social Desirability Scale.

[p]PTSD: post-traumatic stress disorder.

[q]PC-PTSD: Primary Care PTSD.

[r]WHO-ASSIST: World Health Organization-Alcohol, Smoking, and Substance Involvement Screening Test.

Follow-up surveys will be conducted at 3, 6, 9, and 12 months after the baseline assessments. Quarterly follow-up surveys at 3, 6, and 9 months will collect information on MM products (if applicable), current medications, and any health updates. At the 12-month visit, participants will complete the same core survey as the baseline, excluding demographic questions. Follow-up surveys may be completed remotely.

Biospecimens

At every visit, participants will be asked to provide a urine sample to test for THC. For remote follow-up visits, a drug test cup will be mailed with instructions on how to upload a photo of the results.

We will also collect approximately 4 mL of blood every 3 months from each participant completing an in-person visit. Plasma will be shipped to the University of Colorado laboratory for analysis of various cannabinoids and associated metabolites (6a-OH-CBD, 7-OH-CBD, 11OH-THC, THC, CBD, CBC, CBD-gluc, THC-gluc, 7-CBD-COOH, THC-COOH, THC-COOH-Gluc, CBN, CBG, THCV, CBDV) [61-63]. Backup plasma samples may also be shipped to the University of Miami for inflammasome and/or neurodegeneration biomarker tests for exploratory analysis.

For the blood drawn at baseline and 12-month visits, the buffy coat will be frozen for analysis of telomere length, a biological measure of cellular aging [58-60]. Collected samples will be sent to the Blackburn Lab at the University of California, San Francisco, for analysis at the end of the study.

Anthropometric and Cardiovascular Assessments

Participant's height and weight will be measured at the baseline and 12-month study visits. Every 3 months, study staff will measure the participant's blood pressure using a vital sign monitoring device during in-person visits.

Clinical and Experimental Pain Measures at Baseline and 12-Month Visits

Clinical pain measures will include the GCPS [72], which assesses chronic pain severity, and the Chronic Pain Stage, which is a stage measure based on pain frequency, intensity, time, and total number of pain sites [37,58,85]. Experimental pain measures will include punctate mechanical stimuli and pressure pain threshold to assess pain sensitivity, which has shown predictive utility for pain severity in chronic pain patients [64].

Standardized Neurocognitive Assessments at Baseline and 12-Month Visits

After completing the baseline and 12-month follow-up surveys, participants will complete cognitive assessments using the NIH Toolbox Cognition Battery, a standardized neurocognitive assessment tool available as an iPad app [65]. It contains 7 cognition battery tests assessing 5 cognitive subdomains: Language, Executive Function, Episodic Memory, Processing Speed, and Working Memory [65]. Table 1 shows the full list of assessments used. Three composite scores are created from these tests (Fluid, Crystallized, Total Composites), with normative standards available for comparison [86,87]. At the 12-month follow-up visit, only the fluid composite score will be used.

EMA Survey

Participants starting MM treatment will complete a 1-week EMA using a smartphone app (mEMA, illumivu Inc. or MetricWire) before they start MM and a 3-week EMA after they begin using it. They will also complete a 1-week EMA at 3, 6, 9, and 12 months to track their MM use and longitudinal real-time outcomes. The comparison group will complete an approximately 4-week EMA at baseline and an approximately 1-week EMA at 3, 6, 9, and 12 months to answer questions regarding medication and nonpharmaceutical use.

This study will use 4 random prompts (roughly 1 every 4 hours) to assess real-time pain levels, mood (measured by the Brief Symptom Inventory[66]), and other symptoms (eg, nausea, sleepiness, dizziness) and 1 daily prompt to assess overall pain levels in the past 24 hours, the previous night's sleep quality, and medications taken in the past 24 hours besides MM. After the baseline survey and cognitive assessment, study staff will train participants on how to use the EMA with practice questions. Participants starting MM will also be asked to upload a picture of the label on the MM products they have been using to validate the product information.

Fitbit Activity Tracker

Participants will also wear a Fitbit Charge 5 or 6 (the latest model available) in parallel with EMA for objective measures of physical activity, heart rate, and sleep duration and quality.

XSL•FO
RenderX

Data will be accessed via Fitabase, a service that allows centralized management of Fitbit data.

Quality Assurance

The research team will not be permitted to recruit participants or conduct visits until they receive training and pass evaluation by the study coordinators or principal investigator (PI). EMA and Fitbit data will be assessed multiple times per week to monitor for any missing data or issues with participant compliance. Weekly meetings with the research team and biweekly meetings with the investigators will be held to discuss any protocol deviations or discrepancies. The research coordinators and data manager will conduct biweekly checks of all survey data and visit log notes for discrepancies.

Analytical Strategy

Power and Sample Size

The total sample size was calculated based on Aims 1 and 2, for which statistical analysis will use a "difference-in-differences" approach. Power was calculated to ensure the detection of differences between the baseline and 12-month follow-up survey in Aim 2, which requires a larger sample size than the intensive longitudinal data from EMA assessments. Assuming a 3:1 ratio of MM versus comparison group and within-subject correlation of variables of 0.5, the sample size required to detect an effect size of 0.3 SD between-group difference (small to medium effect size) [88] in the outcome measured by standardized scales (eg, reduction in pain intensity measured by GCPS), with 80% power (type I error=0.05), is 352 (264 vs 88). Assuming 20% attrition, based on our pilot study (79% retention without a preplanned 12-month follow-up), the final sample size required to achieve the power is 440 (330 vs 110) [89]. For Aim 1, we will have more than 80% power to detect an effect size of 0.3 SD in between-group difference because there will be more repeated measurements in Aim 1. For Aim 3, we will have 264 participants for the analysis since only the MM group will be analyzed. After adjusting for multiple testing with the conservative Bonferroni method [90], the adjusted type I error rate becomes 0.05/7=0.007, under which we will have 99% power to detect the Pearson correlation of 0.3.

Aim 1

Aim 1 is to determine whether MM use leads to short-term changes in pain intensity and physical and emotional functioning, as measured in real time by EMA (subjective data) and the Fitbit device (objective data).

Key outcomes of interest include pain intensity, physical functioning, emotional functioning, and sleep. Pain intensity will be assessed from the EMA surveys using self-reported real-time pain intensity on a 0 to 100 visual analog scale. Physical functioning will be assessed from the EMA surveys using self-reported real-time pain interference and from the Fitbit using step counts. Emotional functioning will be assessed from the EMA surveys using self-reported real-time anxiety and depression symptoms (Brief Symptom Inventory).

Separate generalized linear mixed-effects models (GLMM) or generalized estimating equations (GEE) will be used to compare the temporal changes in each of the main outcomes (pain intensity level, physical and emotional functioning) between the MM and comparison groups. A difference-in-difference analysis will be conducted. A propensity score method will account for potential baseline differences between the 2 groups for analysis with inverse probability of treatment weighting. Daily MM use measured via EMA will be a time-varying exposure variable in the GLMM or GEE models to account for crossover between groups. The research team will also include time-varying covariates in these models to account for potential impacts of other non-pharmaceutical pain treatments (eg, category of treatments and number of sessions for each category) and non-MM use (self-reported frequency/quantity).

Aim 2

Aim 2 is to evaluate whether MM use leads to longer-term changes over a 12-month period based on self-reported pain intensity, emotional and physical functioning, health-related quality of life, and cellular aging markers such as telomere length.

Key outcomes of interest include pain intensity and interference, physical functioning, emotional functioning, cognitive function, opioid, benzodiazepine, and other medication use, sleep quality, quality of life, and cellular aging. Pain intensity will be assessed from EMA surveys, similar to Aim 1, as well as Brief Pain Inventory and GCPS responses from the core survey. Physical functioning will be assessed from EMA surveys, similar to Aim 1, and the PROMIS Pain Interference Scale from the core survey. Emotional functioning will be assessed from EMA surveys, similar to Aim 1, and the PROMIS Anxiety and Depression Scales from the core survey. Cognitive function will be assessed using the NIH Toolbox Cognition Battery and self-reported changes in cognition from the core survey. Opioid, benzodiazepine, and other medication use will be assessed using the Medication Quantification Scale from the core survey. Sleep quality will be assessed using self-reported sleep duration and quality from the EMA surveys, the Pittsburgh Sleep Quality Index from the core survey, and Fitbit-recorded sleep duration. Quality of life will be assessed using the PROMIS Global-10 and Patient Global Impression of Change from the core survey. Cellular aging will be assessed by measuring telomere length from buffy coat samples.

The research team will analyze 2 types of data: (1) intensive longitudinal data (EMA, Fitbit, opioid, and benzodiazepine dose) and (2) data collected only at baseline and 12 months (core survey, cognitive tests, biomarkers). For intensive longitudinal data, separate GLMM or GEE models will be built to compare longitudinal trajectories in each outcome (pain intensity, emotional and physical functioning, opioid dose) measured at multiple time points over 12 months between the MM group and the comparison group. As in Aim 1, we will treat MM use as a time-varying exposure variable to account for crossover and include other nonpharmaceutical pain treatments and nonmedical marijuana as time-varying covariates. The analysis will use the exact time (eg, hours) of each measurement (instead of days) for constructing the trajectory and thus does not need to average measurements within each day into a single value. For measures collected only at baseline

XSL•FO
RenderX

and 12 months, generalized linear regression models will be used to compare group differences in the survey data (eg, anxiety/depression, disability, quality of life), cognitive test results, and telomere length. Additionally, milligram morphine equivalent opioid dosage from the prescription drug monitoring program will be obtained through medical records to assess differences in the risk of unintentional overdose between the MM and comparison groups.

Aim 3

Aim 3 is to examine among individuals initiating MM which MM product characteristics (ie, THC:CBD ratio, administration route, dose) are associated with greater improvements in the outcomes defined in Aims 1 and 2, or with a greater incidence of side effects, and explore whether individual differences (eg, sex, baseline pain phenotype) moderate these relationships. Key outcomes of interest are defined in Aims 1 and 2.

Separate GLMM or GEE models using each product characteristic as a predictor will be developed. As in Aims 1 and 2, we will treat MM use as a time-varying exposure variable to account for crossover and include other nonpharmaceutical pain treatments and nonmedical marijuana as time-varying covariates. The research team will construct a multivariate model to include all 3 product characteristics (ie, THC:CBD ratio, administration route, dose) as predictors for each main outcome using stepwise model selection or least absolute shrinkage and selection operator to account for potential interactions within these product characteristics. To determine significant moderating effects of interpersonal differences, separate GLMM or GEE models will be conducted by gender (male vs female) or at each level of the moderator (low/moderate/high stage of pain, low/high marijuana expectancy, low/high social desirability bias) to assess whether the main effect of the MM product characteristic on each main outcome differs by group membership or the level of the moderator.

Missing Data

Missing data will be examined to determine whether information is missing at random by comparing participants with and without missing data. If no significant differences are observed in key demographics, missing at random will be assumed. If missingness exceeds 5%, we will use multiple imputation methods to impute missing responses. Sensitivity analyses will be conducted to assess the impact of the imputed data.

Possible Risks and Discomforts

Adverse Events

Adverse events and discomforts may occur due to participation in the study. In the event of adverse events, the PI will determine whether the event is serious and related to the study procedures. If this is the case, the PI will report the event to the NIH and IRB within 72 hours of the event. If the event is not serious, it will be reported to the NIH and IRB during regular status reports.

Breach of Confidentiality

Our study will collect data on participants' medications, including MM. Breaches of confidentiality could occur through communications with participants, loss of research data, or transfer of data.

Questionnaires and Interviews

The assessments used in this study address some sensitive issues, such as mental well-being and drug use. The major disadvantages of these assessments are the time required to complete them and the potential breach of confidentiality. Our past experience with these measures indicates that they are acceptable to participants.

Cognitive Tests

There are minimal risks associated with these tests. Participants may experience fatigue and feelings of frustration while completing the cognitive function tests. Some participants may also experience stress associated with being tested, though this tends to be quite limited.

EMA Surveys

The EMA involves minimal risk as it only collects self-reported data related to participants' symptoms and their relief, experienced side effects, and other related feelings (eg, anxiety, depression). The major disadvantage of this assessment is the time required to complete daily reports. The possibility of a confidentiality breach is very low because the data collected from the mobile app will be encrypted and pushed to the cloud-based storage database. No data are ever stored on the server's file system but always in the database. Access to the database is gated, so entry is only permitted to users entering through the approved route (ie, you cannot hack your way into the database by guessing the URL).

Pain Sensory Tests

Procedures may be uncomfortable or unpleasant. Participants may experience some temporary discomfort from the pressure and mechanical pain testing. However, if they feel the pain is greater than they wish to tolerate, they can stop any of the procedures at any time.

Urine Sample Collection

The risk is minimal for urine sample collection, although some people may experience embarrassment when providing the sample.

Blood Collection

The risks of drawing blood from a vein include discomfort at the site of puncture; possible bruising and swelling around the puncture site; rarely, an infection; and uncommonly, faintness from the procedure.

Fitbit Data Collection

The risk is minimal when wearing a Fitbit during the EMA period. All data from Fitbit do not include personal identifiers. Data will be stored temporarily on cloud-based storage and will be downloaded regularly onto a secure UF server. Data will be encrypted during the transmission process to ensure security.

Anthropometric and Cardiovascular Assessments

The risk is minimal for collecting these measures, including height, weight, blood pressure, and heart rate. Some participants

XSL•FO
RenderX

may feel slightly uncomfortable or embarrassed about having some of these measures taken.

Results

The SMILE study was funded by the NIH/National Institute on Aging in February 2022. Recruitment for the SMILE study began in July 2022 and is expected to continue until 2025, with the last participants recruited completing study involvement in 2026. As of October 2025, we have enrolled 399 participants, with 277 in the MM group and 122 in the comparison group. Data analysis is currently underway, and results are expected to be published starting in 2027. The anticipated timeline for the study is shown in Table 3.

Table . Study timeline.

Research tasks	2022		2023		2024		2025		2026	
	H1[a]	H2	H1	H2	H1	H2	H1	H2	H1	H2
Recruitment and baseline assessments		✓	✓	✓	✓	✓	✓	✓		
Follow-up assessments (EMA[b]/Fitbit, 1-y visit)			✓	✓	✓	✓	✓	✓	✓	✓
Data/merge analysis and results dissemination				✓	✓	✓	✓	✓	✓	✓

[a]H1 and H2 indicate the first and second halves of the year.

[b]EMA: ecological momentary assessment.

Discussion

Strengths and Innovation

This project is the first prospective cohort study to comprehensively investigate the IMMPACT-recommended core outcomes [48-50] and cognitive function, using multisource (survey, EMA, sensor, and biomarker) subjective and objective data collected in real time and over 12 months to evaluate the effects of MM use on older adults with chronic pain. This project is innovative in a number of ways.

Use of Smartphone and Sensor-Based EMA

Mobile technology has rapidly revolutionized health care and research, but it has not yet been widely applied in efforts to investigate the emerging phenomenon of widespread MM use among older adults [45]. This project is the first to leverage both smartphone-based EMA and Fitbit devices to obtain detailed MM use pattern data, as well as subjective and objective real-time measures of pain and functioning in patients' daily lives. Such data can also provide valuable insights into which MM products are associated with improvements in short-term outcomes [91, 92].

Enhanced Longitudinal Cohort Design

Our study is the first to leverage a measurement burst design within a cohort study to examine the short- and longer-term effects of MM on older adults. The baseline and 12-month follow-up data will provide comprehensive assessments of pain, functioning, and overall quality of life as snapshots at 2 time points, whereas the quarterly EMA measurement bursts will provide detailed data on short-term variability in pain intensity and physical and emotional functioning, as well as the trajectory of change over 12 months. We will include baseline assessments before participants initiate MM treatment and a comparison group who share similar demographic and health characteristics to reach causal inference.

Incorporation of Telomere Length as a Measure to Assess MM's Effect on Cellular Aging Over a 12-Month Period

We and others have shown that telomere length, a measure of cellular aging, is a biomarker of stress-related burdens and buffers [37,38,59,93]. However, no research has investigated how MM may influence cellular aging, as measured by telomere length, in individuals with chronic pain. The proposed study will address this gap by collecting and analyzing telomere length at baseline and after 12 months.

Consideration of Interpersonal Differences, Including Pain Phenotyping at Baseline

Our study will be the first to examine whether MM's effects differ based on biological factors (sex and pain phenotyping) and psychological factors (expectancy and social desirability) using established scales [58,76,82]. Such information has valuable implications for physicians and patients regarding which patient subgroups may experience greater benefits from MM use.

XSL•FO
RenderX

Inclusion of Blood Tests for Multiple Cannabinoids and Associated Metabolites

Our study is the first to test plasma-level cannabinoids and associated metabolites over the course of 1 year for a cohort of older adults with chronic pain. Previous studies have shown that blood tests of cannabinoids and their metabolites can provide valuable information on marijuana use [61-63]. The rich information obtained from repeated blood tests offers objective measures of cannabis exposure to inform how various compounds (eg, THC, CBD) in MM impact health outcomes in this population.

This study will provide valuable scientific evidence to inform physician and patient decisions on using MM as an alternative or supplemental pain treatment option for older adults. With multisource data collected in real time and over 12 months, our results will contribute to the greatly needed evidence on: (1) whether MM can reduce pain and improve physical and emotional functioning in the short term among older adults; (2) whether the effects of MM last for 12 months and demonstrate quality of life improvements; (3) whether health benefits or consequences differ by MM product type; and (4) which subgroups of patients may benefit more from MM use. Should this study confirm that MM helps to improve pain, emotional and physical functioning, and quality of life without significantly reducing cognitive function over a 12-month period, future steps will involve conducting a more focused investigation of specific MM products (eg, sublingual drops with a 1:1 THC:CBD ratio) that are associated with more beneficial outcomes. Future studies are needed to examine treatment side effects as well as optimal dosage to inform more specific recommendations for MM.

Limitations

There are several limitations to our study. Patients can obtain an MM card as early as the same day they meet with a qualifying physician, which may limit the ability to enroll participants prior to the initiation of MM treatment. Conversely, some participants in the MM group, despite expressing an intention to start treatment promptly, may experience delays in initiation due to various barriers, such as financial constraints or access issues. To address this issue, we have implemented additional compensation for participants who delay the initiation of their MM treatment to enroll in the study and begin MM treatment within 4 weeks of their baseline visit. Another potential limitation of the SMILE study is the wide variety of MM products and modes of administration available. As this is an observational study, no limitations are placed on what types of products may be used by participants. Attrition bias is a concern, as study participants who experience side effects or receive no benefits from MM treatment may withdraw from the study, which could result in a more favorable outcome toward MM improving chronic pain. To prevent this, our team emphasizes to the participants that we would like to continue monitoring their health outcomes regardless of whether they continue using MM. Our team may also waive certain data measures (eg, EMA, Fitbit) to make participation easier for individuals who are considering withdrawing from the study. Additionally, individuals with chronic pain may be experiencing pain due to undiagnosed cancer, which could result in misclassification.

Dissemination Plan and Future Research

Results from this study will be published in peer-reviewed research journals and presented at conferences. Authorship for all manuscripts and presentations will follow the International Committee of Medical Journal Editors recommendations. Future research will be needed to expand the literature on the use of MM in older adults with chronic pain. In addition, the use of randomized controlled trials to control for the types of MM products used by participants will be useful in the future.

Acknowledgments

The research team would like to acknowledge the clinics that assisted in recruiting and referring participants to participate in our study. These clinics include CannaMD, DocMJ, Florida Marijuana Doctors, Releafe Now, and UF Health Clinics. Additionally, we would like to acknowledge our data managers and research staff.

Funding

The Study on Medical Marijuana and Its Long-Term Effects on Older Adults (SMILE) study is funded by the National Institutes of Health/National Institute on Aging (R01AG071729 and R01AG071729-01A1S1; principal investigator: YW).

Data Availability

Deidentified data will be available following the initial publication of findings upon reasonable request. Once the main analyses are complete, you may contact the principal investigator, YW (ywang48@ufl.edu), to inquire about data availability.

Conflicts of Interest

None declared.

Checklist 1
SPIROS checklist.
[PDF File, 829 KB - resprot_v15i1e78900_app1.pdf]

XSL•FO
RenderX

References

1. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Committee on the Health Effects of Marijuana: An Evidence Review and Research Agenda. The Health Effects of Cannabis and Cannabinoids The Current State of Evidence and Recommendations for Research: National Academies Press; 2017. URL: https://www.ncbi.nlm.nih.gov/books/NBK423845/ [accessed 2026-03-05] [Medline: 28182367]
2. Hill KP, Palastro MD, Johnson B, Ditre JW. Cannabis and pain: a clinical review. Cannabis Cannabinoid Res 2017;2(1):96-104. [doi: 10.1089/can.2017.0017] [Medline: 28861509]
3. Park JY, Wu LT. Prevalence, reasons, perceived effects, and correlates of medical marijuana use: a review. Drug Alcohol Depend 2017 Aug 1;177:1-13. [doi: 10.1016/j.drugalcdep.2017.03.009] [Medline: 28549263]
4. Dahlhamer J, Lucas J, Zelaya C, et al. Prevalence of chronic pain and high-impact chronic pain among adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Sep 14;67(36):1001-1006. [doi: 10.15585/mmwr.mm6736a2] [Medline: 30212442]
5. Salas-Wright CP, Vaughn MG, Cummings-Vaughn LA, et al. Trends and correlates of marijuana use among late middle-aged and older adults in the United States, 2002-2014. Drug Alcohol Depend 2017 Feb 1;171:97-106. [doi: 10.1016/j.drugalcdep.2016.11.031] [Medline: 28063338]
6. Lloyd SL, Striley CW. Marijuana use among adults 50 years or older in the 21st century. Gerontol Geriatr Med 2018;4:2333721418781668. [doi: 10.1177/2333721418781668] [Medline: 29977980]
7. Boehnke KF, Scott JR, Litinas E, Sisley S, Williams DA, Clauw DJ. Pills to pot: observational analyses of cannabis substitution among medical cannabis users with chronic pain. J Pain 2019 Jul;20(7):830-841. [doi: 10.1016/j.jpain.2019.01.010] [Medline: 30690169]
8. Wang T, Collet JP, Shapiro S, Ware MA. Adverse effects of medical cannabinoids: a systematic review. CMAJ 2008 Jun 17;178(13):1669-1678. [doi: 10.1503/cmaj.071178] [Medline: 18559804]
9. Pletcher MJ, Vittinghoff E, Kalhan R, et al. Association between marijuana exposure and pulmonary function over 20 years. JAMA 2012 Jan 11;307(2):173-181. [doi: 10.1001/jama.2011.1961] [Medline: 22235088]
10. Lee G, Grovey B, Furnish T, Wallace M. Medical cannabis for neuropathic pain. Curr Pain Headache Rep 2018 Feb 1;22(1):8. [doi: 10.1007/s11916-018-0658-8] [Medline: 29388063]
11. Mücke M, Phillips T, Radbruch L, Petzke F, Häuser W. Cannabis-based medicines for chronic neuropathic pain in adults. Cochrane Database Syst Rev 2018 Mar 7;3(3):CD012182. [doi: 10.1002/14651858.CD012182.pub2] [Medline: 29513392]
12. Gruber SA, Sagar KA, Dahlgren MK, et al. The grass might be greener: medical marijuana patients exhibit altered brain activity and improved executive function after 3 months of treatment. Front Pharmacol 2017;8:983. [doi: 10.3389/fphar.2017.00983] [Medline: 29387010]
13. Powell D, Pacula RL, Jacobson M. Do medical marijuana laws reduce addictions and deaths related to pain killers? J Health Econ 2018 Mar;58:29-42. [doi: 10.1016/j.jhealeco.2017.12.007] [Medline: 29408153]
14. Bradford AC, Bradford WD. Medical marijuana laws reduce prescription medication use in Medicare Part D. Health Aff (Millwood) 2016 Jul 1;35(7):1230-1236. [doi: 10.1377/hlthaff.2015.1661] [Medline: 27385238]
15. Finney JW, Humphreys K, Harris AHS. What ecologic analyses cannot tell us about medical marijuana legalization and opioid pain medication mortality. JAMA Intern Med 2015 Apr;175(4):655-656. [doi: 10.1001/jamainternmed.2014.8006] [Medline: 25844747]
16. Hall W, West R, Marsden J, Humphreys K, Neale J, Petry N. It is premature to expand access to medicinal cannabis in hopes of solving the US opioid crisis. Addiction 2018 Jun;113(6):987-988. [doi: 10.1111/add.14139] [Medline: 29468760]
17. Olfson M, Wall MM, Liu SM, Blanco C. Cannabis use and risk of prescription opioid use disorder in the United States. Am J Psychiatry 2018 Jan 1;175(1):47-53. [doi: 10.1176/appi.ajp.2017.17040413] [Medline: 28946762]
18. Hasin DS. US epidemiology of cannabis use and associated problems. Neuropsychopharmacology 2018 Jan;43(1):195-212. [doi: 10.1038/npp.2017.198] [Medline: 28853439]
19. Campbell G, Hall WD, Peacock A, et al. Effect of cannabis use in people with chronic non-cancer pain prescribed opioids: findings from a 4-year prospective cohort study. Lancet Public Health 2018 Jul;3(7):e341-e350. [doi: 10.1016/S2468-2667(18)30110-5] [Medline: 29976328]
20. O'Connell M, Sandgren M, Frantzen L, Bower E, Erickson B. Medical cannabis: effects on opioid and benzodiazepine requirements for pain control. Ann Pharmacother 2019 Nov;53(11):1081-1086. [doi: 10.1177/1060028019854221] [Medline: 31129977]
21. Ware MA, Wang T, Shapiro S, Collet JP, COMPASS study team. Cannabis for the management of pain: assessment of safety study (COMPASS). J Pain 2015 Dec;16(12):1233-1242. [doi: 10.1016/j.jpain.2015.07.014] [Medline: 26385201]
22. Minerbi A, Häuser W, Fitzcharles MA. Medical cannabis for older patients. Drugs Aging 2019 Jan;36(1):39-51. [doi: 10.1007/s40266-018-0616-5] [Medline: 30488174]
23. Abuhasira R, Schleider LBL, Mechoulam R, Novack V. Epidemiological characteristics, safety and efficacy of medical cannabis in the elderly. Eur J Intern Med 2018 Mar;49:44-50. [doi: 10.1016/j.ejim.2018.01.019] [Medline: 29398248]
24. Vyas MB, LeBaron VT, Gilson AM. The use of cannabis in response to the opioid crisis: a review of the literature. Nurs Outlook 2018;66(1):56-65. [doi: 10.1016/j.outlook.2017.08.012] [Medline: 28993073]

XSL•FO
RenderX

25. McDonagh MS, Morasco BJ, Wagner J, et al. Cannabis-based products for chronic pain: a systematic review. Ann Intern Med 2022 Aug;175(8):1143-1153. [doi: 10.7326/M21-4520] [Medline: 35667066]
26. Wang L, Hong PJ, May C, et al. Medical cannabis or cannabinoids for chronic non-cancer and cancer related pain: a systematic review and meta-analysis of randomised clinical trials. BMJ 2021 Sep 8;374:n1034. [doi: 10.1136/bmj.n1034] [Medline: 34497047]
27. Whiting PF, Wolff RF, Deshpande S, et al. Cannabinoids for medical use: a systematic review and meta-analysis. JAMA 2015;313(24):2456-2473. [doi: 10.1001/jama.2015.6358] [Medline: 26103030]
28. Stockings E, Campbell G, Hall WD, et al. Cannabis and cannabinoids for the treatment of people with chronic noncancer pain conditions: a systematic review and meta-analysis of controlled and observational studies. Pain 2018 Oct;159(10):1932-1954. [doi: 10.1097/j.pain.0000000000001293] [Medline: 29847469]
29. Boland EG, Bennett MI, Allgar V, Boland JW. Cannabinoids for adult cancer-related pain: systematic review and meta-analysis. BMJ Support Palliat Care 2020 Mar;10(1):14-24. [doi: 10.1136/bmjspcare-2019-002032] [Medline: 31959586]
30. Aviram J, Samuelly-Leichtag G. Efficacy of cannabis-based medicines for pain management: a systematic review and meta-analysis of randomized controlled trials. Pain Physician 2017 Sep;20(6):E755-E796. [Medline: 28934780]
31. Shover CL, Davis CS, Gordon SC, Humphreys K. Association between medical cannabis laws and opioid overdose mortality has reversed over time. Proc Natl Acad Sci U S A 2019 Jun 25;116(26):12624-12626. [doi: 10.1073/pnas.1903434116] [Medline: 31182592]
32. Raji MA, Abara NO, Salameh H, Westra JR, Kuo YF. Association between cannabis laws and opioid prescriptions among privately insured adults in the US. Prev Med 2019 Aug;125:62-68. [doi: 10.1016/j.ypmed.2019.05.012] [Medline: 31125629]
33. Segura LE, Mauro CM, Levy NS, et al. Association of US medical marijuana laws with nonmedical prescription opioid use and prescription opioid use disorder. JAMA Netw Open 2019 Jul 3;2(7):e197216. [doi: 10.1001/jamanetworkopen.2019.7216] [Medline: 31314118]
34. Piper BJ, Beals ML, Abess AT, et al. Chronic pain patients' perspectives of medical cannabis. Pain 2017 Jul;158(7):1373-1379. [doi: 10.1097/j.pain.0000000000000899] [Medline: 28328576]
35. Edwards RR, Dworkin RH, Turk DC, et al. Patient phenotyping in clinical trials of chronic pain treatments: IMMPACT recommendations. Pain 2016 Sep;157(9):1851-1871. [doi: 10.1097/j.pain.0000000000000602] [Medline: 27152687]
36. Naranjo-Hernández D, Reina-Tosina J, Roa LM. Sensor technologies to manage the physiological traits of chronic pain: a review. Sensors (Basel) 2020 Jan 8;20(2):365. [doi: 10.3390/s20020365] [Medline: 31936420]
37. Sibille KT, Chen H, Bartley EJ, et al. Accelerated aging in adults with knee osteoarthritis pain: consideration for frequency, intensity, time, and total pain sites. Pain Rep 2017 May;2(3):e591. [doi: 10.1097/PR9.0000000000000591] [Medline: 29392207]
38. Johnson AJ, Terry E, Bartley EJ, et al. Resilience factors may buffer cellular aging in individuals with and without chronic knee pain. Mol Pain 2019;15:1744806919842962. [doi: 10.1177/1744806919842962] [Medline: 30900507]
39. Feingold D, Brill S, Goor-Aryeh I, Delayahu Y, Lev-Ran S. Medical marijuana dependence among chronic pain patients suffering from depression and anxiety [Article in Hebrew]. Harefuah 2019 Jul;158(7):438-444. [Medline: 31339242]
40. Aldington D, Eccleston C. Evidence-based pain management: building on the foundations of Cochrane systematic reviews. Am J Public Health 2019 Jan;109(1):46-49. [doi: 10.2105/AJPH.2018.304745] [Medline: 30495991]
41. Häuser W, Finnerup NB, Moore RA. Systematic reviews with meta-analysis on cannabis-based medicines for chronic pain: a methodological and political minefield. Pain 2018 Oct;159(10):1906-1907. [doi: 10.1097/j.pain.0000000000001295] [Medline: 29847473]
42. NIDA's role in providing cannabis for research. National Institute on Drug Abuse (NIDA). 2020. URL: https://nida.nih.gov/research/resources-grants-contracts/nidas-role-in-providing-cannabis-research [accessed 2025-04-03]
43. Carr D, Schatman M. Cannabis for chronic pain: not ready for prime time. Am J Public Health 2019 Jan;109(1):50-51. [doi: 10.2105/AJPH.2018.304593] [Medline: 32941762]
44. Bachhuber MA, Arnsten JH, Starrels JL, Cunningham CO. Willingness to participate in longitudinal research among people with chronic pain who take medical cannabis: a cross-sectional survey. Cannabis Cannabinoid Res 2018;3(1):45-53. [doi: 10.1089/can.2017.0051] [Medline: 29607410]
45. Volkow ND, Weiss SRB. Importance of a standard unit dose for cannabis research. Addiction 2020 Jul;115(7):1219-1221. [doi: 10.1111/add.14984] [Medline: 32083354]
46. Lev-Ran S, Roerecke M, Le Foll B, George TP, McKenzie K, Rehm J. The association between cannabis use and depression: a systematic review and meta-analysis of longitudinal studies. Psychol Med 2014 Mar;44(4):797-810. [doi: 10.1017/S0033291713001438] [Medline: 23795762]
47. Volkow ND, Baler RD, Compton WM, Weiss SRB. Adverse health effects of marijuana use. N Engl J Med 2014 Jun 5;370(23):2219-2227. [doi: 10.1056/NEJMra1402309] [Medline: 24897085]
48. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008 Feb;9(2):105-121. [doi: 10.1016/j.jpain.2007.09.005] [Medline: 18055266]
49. Turk DC, Dworkin RH, Allen RR, et al. Core outcome domains for chronic pain clinical trials: IMMPACT recommendations. Pain 2003 Dec;106(3):337-345. [doi: 10.1016/j.pain.2003.08.001] [Medline: 14659516]

XSL•FO
RenderX

50. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005 Jan;113(1-2):9-19. [doi: 10.1016/j.pain.2004.09.012] [Medline: 15621359]
51. Bass PF, Wilson JF, Griffith CH. A shortened instrument for literacy screening. J Gen Intern Med 2003 Dec;18(12):1036-1038. [doi: 10.1111/j.1525-1497.2003.10651.x] [Medline: 14687263]
52. Dieppe P. Chronic musculoskeletal pain. BMJ 2013 May 16;346:f3146. [doi: 10.1136/bmj.f3146] [Medline: 23709528]
53. Treede RD, Rief W, Barke A, et al. Chronic pain as a symptom or a disease: the IASP Classification of Chronic Pain for the International Classification of Diseases (ICD-11). Pain 2019 Jan;160(1):19-27. [doi: 10.1097/j.pain.0000000000001384] [Medline: 30586067]
54. van de Donk T, Niesters M, Kowal MA, Olofsen E, Dahan A, van Velzen M. An experimental randomized study on the analgesic effects of pharmaceutical-grade cannabis in chronic pain patients with fibromyalgia. Pain 2019 Apr;160(4):860-869. [doi: 10.1097/j.pain.0000000000001464] [Medline: 30585986]
55. Hasin DS, O'Brien CP, Auriacombe M, et al. DSM-5 criteria for substance use disorders: recommendations and rationale. Am J Psychiatry 2013 Aug;170(8):834-851. [doi: 10.1176/appi.ajp.2013.12060782] [Medline: 23903334]
56. Skinner HA. The drug abuse screening test. Addict Behav 1982;7(4):363-371. [doi: 10.1016/0306-4603(82)90005-3] [Medline: 7183189]
57. Yudko E, Lozhkina O, Fouts A. A comprehensive review of the psychometric properties of the Drug Abuse Screening Test. J Subst Abuse Treat 2007 Mar;32(2):189-198. [doi: 10.1016/j.jsat.2006.08.002] [Medline: 17306727]
58. Sibille KT, Steingrímsdóttir Ó, Fillingim RB, et al. Investigating the burden of chronic pain: an inflammatory and metabolic composite. Pain Res Manag 2016;2016:7657329. [doi: 10.1155/2016/7657329] [Medline: 27445627]
59. Sibille KT, Witek-Janusek L, Mathews HL, Fillingim RB. Telomeres and epigenetics: potential relevance to chronic pain. Pain 2012 Sep;153(9):1789-1793. [doi: 10.1016/j.pain.2012.06.003] [Medline: 22770844]
60. Stout SA, Lin J, Hernandez N, et al. Validation of minimally-invasive sample collection methods for measurement of telomere length. Front Aging Neurosci 2017;9:397. [doi: 10.3389/fnagi.2017.00397] [Medline: 29270121]
61. Klawitter J, Sempio C, Mörlein S, et al. An atmospheric pressure chemical ionization MS/MS assay using online extraction for the analysis of 11 cannabinoids and metabolites in human plasma and urine. Ther Drug Monit 2017 Oct;39(5):556-564. [doi: 10.1097/FTD.0000000000000427] [Medline: 28640062]
62. Sempio C, Almaraz-Quinones N, Jackson M, et al. Simultaneous quantification of 17 cannabinoids by LC-MS-MS in human plasma. J Anal Toxicol 2022 Apr 21;46(4):383-392. [doi: 10.1093/jat/bkab030] [Medline: 33754154]
63. Gilman JM, Schmitt WA, Wheeler G, et al. Variation in cannabinoid metabolites present in the urine of adults using medical cannabis products in Massachusetts. JAMA Netw Open 2021 Apr 1;4(4):e215490. [doi: 10.1001/jamanetworkopen.2021.5490] [Medline: 33844003]
64. King CD, Sibille KT, Goodin BR, et al. Experimental pain sensitivity differs as a function of clinical pain severity in symptomatic knee osteoarthritis. Osteoarthritis Cartilage 2013 Sep;21(9):1243-1252. [doi: 10.1016/j.joca.2013.05.015] [Medline: 23973137]
65. Heaton RK, Akshoomoff N, Tulsky D, et al. Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults. J Int Neuropsychol Soc 2014 Jul;20(6):588-598. [doi: 10.1017/S1355617714000241] [Medline: 24960398]
66. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med 1983 Aug;13(3):595-605. [Medline: 6622612]
67. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis 1987;40(5):373-383. [doi: 10.1016/0021-9681(87)90171-8] [Medline: 3558716]
68. Harden RN, Weinland SR, Remble TA, et al. Medication Quantification Scale Version III: update in medication classes and revised detriment weights by survey of American Pain Society Physicians. J Pain 2005 Jun;6(6):364-371. [doi: 10.1016/j.jpain.2005.01.350] [Medline: 15943958]
69. Chang KL, Fillingim R, Hurley RW, Schmidt S. Chronic pain management: nonpharmacological therapies for chronic pain. FP Essent 2015 May;432:21-26. [Medline: 25970869]
70. Weingart SN, Gandhi TK, Seger AC, et al. Patient-reported medication symptoms in primary care. Arch Intern Med 2005 Jan 24;165(2):234-240. [doi: 10.1001/archinte.165.2.234] [Medline: 15668373]
71. Cleeland CS, Ryan KM. Pain assessment: global use of the Brief Pain Inventory. Ann Acad Med Singap 1994 Mar;23(2):129-138. [Medline: 8080219]
72. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. Pain 1992 Aug;50(2):133-149. [doi: 10.1016/0304-3959(92)90154-4] [Medline: 1408309]
73. Ferguson L, Scheman J. Patient global impression of change scores within the context of a chronic pain rehabilitation program. J Pain 2009 Apr;10(4):S73. [doi: 10.1016/j.jpain.2009.01.258]
74. Askew RL, Cook KF, Revicki DA, Cella D, Amtmann D. Evidence from diverse clinical populations supported clinical validity of PROMIS pain interference and pain behavior. J Clin Epidemiol 2016 May;73:103-111. [doi: 10.1016/j.jclinepi.2015.08.035] [Medline: 26931296]

XSL•FO
RenderX

75. Bohnert KM, Bonar EE, Arnedt JT, Conroy DA, Walton MA, Ilgen MA. Utility of the comprehensive marijuana motives questionnaire among medical cannabis patients. Addict Behav 2018 Jan;76:139-144. [doi: 10.1016/j.addbeh.2017.08.001] [Medline: 28820970]
76. Morean ME, Butler ER. The Medical Cannabis Expectancy Questionnaire: adult medical marijuana users' expectancies associated with combustible, vaporized, and edible cannabis use for medical purposes. Cannabis 2019;2(1):1-18. [doi: 10.26828/cannabis.2019.01.001]
77. Adamson SJ, Kay-Lambkin FJ, Baker AL, et al. An improved brief measure of cannabis misuse: the Cannabis Use Disorders Identification Test-Revised (CUDIT-R). Drug Alcohol Depend 2010 Jul 1;110(1-2):137-143. [doi: 10.1016/j.drugalcdep.2010.02.017] [Medline: 20347232]
78. Lankenau SE, Ataiants J, Mohanty S, Schrager S, Iverson E, Wong CF. Health conditions and motivations for marijuana use among young adult medical marijuana patients and non-patient marijuana users. Drug Alcohol Rev 2018 Feb;37(2):237-246. [doi: 10.1111/dar.12534] [Medline: 28434211]
79. WHO ASSIST Working Group. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. Addiction 2002 Sep;97(9):1183-1194. [doi: 10.1046/j.1360-0443.2002.00185.x] [Medline: 12199834]
80. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009 Sep;18(7):873-880. [doi: 10.1007/s11136-009-9496-9] [Medline: 19543809]
81. Schalet BD, Pilkonis PA, Yu L, et al. Clinical validity of PROMIS depression, anxiety, and anger across diverse clinical samples. J Clin Epidemiol 2016 May;73:119-127. [doi: 10.1016/j.jclinepi.2015.08.036] [Medline: 26931289]
82. Reynolds WM. Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. J Clin Psychol 1982 Jan;38(1):119-125. [doi: 10.1002/1097-4679(198201)38:1<119::AID-JCLP2270380118>3.0.CO;2-I]
83. Kimerling R, Trafton JA, Nguyen B. Validation of a brief screen for post-traumatic stress disorder with substance use disorder patients. Addict Behav 2006 Nov;31(11):2074-2079. [doi: 10.1016/j.addbeh.2006.02.008] [Medline: 16574331]
84. Buysse DJ, Yu L, Moul DE, et al. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. Sleep 2010 Jun;33(6):781-792. [doi: 10.1093/sleep/33.6.781] [Medline: 20550019]
85. Antoine LH, Tanner JJ, Mickle AM, et al. Greater socioenvironmental risk factors and higher chronic pain stage are associated with thinner bilateral temporal lobes. Brain Behav 2023 Dec;13(12):e3330. [doi: 10.1002/brb3.3330] [Medline: 37984835]
86. Casaletto KB, Umlauf A, Marquine M, et al. Demographically corrected normative standards for the Spanish language version of the NIH Toolbox Cognition Battery. J Int Neuropsychol Soc 2016 Mar;22(3):364-374. [doi: 10.1017/S135561771500137X] [Medline: 26817924]
87. Casaletto KB, Umlauf A, Beaumont J, et al. Demographically corrected normative standards for the English version of the NIH Toolbox Cognition Battery. J Int Neuropsychol Soc 2015 May;21(5):378-391. [doi: 10.1017/S1355617715000351] [Medline: 26030001]
88. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Routledge; 2013. [doi: 10.4324/9780203771587]
89. Li Z, McKeague IW. Power and sample size calculations for generalized estimating equations via local asymptotics. STAT Sin 2013 Jan 1;23(1):231-250. [doi: 10.5705/ss.2011.081] [Medline: 24478568]
90. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 1995;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
91. Wang Y, Jean Jacques J, Li Z, Sibille KT, Cook RL. Health outcomes among adults initiating medical cannabis for chronic pain: a 3-month prospective study incorporating ecological momentary assessment (EMA). Cannabis 2021 Oct;4(2):69-83. [doi: 10.26828/cannabis/2021.02.006] [Medline: 34671723]
92. Wang Y. Medical marijuana as a pain treatment: preliminary findings from an ecological momentary assessment (EMA) study among older adults with chronic pain. Presented at: Addiction Health Services Research Conference; Oct 16-18, 2019 URL: https://medicine.utah.edu/sites/g/files/zrelqx356/files/migration/media/ahsr-ss2.pdf [accessed 2026-03-16]
93. Sibille KT, Langaee T, Burkley B, et al. Chronic pain, perceived stress, and cellular aging: an exploratory study. Mol Pain 2012 Feb 12;8:12. [doi: 10.1186/1744-8069-8-12] [Medline: 22325162]

Abbreviations

CBD: cannabidiol
EMA: ecological momentary assessment
GCPS: Graded Chronic Pain Scale
GEE: generalized estimating equations
GLMM: generalized linear mixed-effects models
IMMPACT: Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials
IRB: Institutional Review Board
MM: medical marijuana

XSL•FO
RenderX

NIH: National Institutes of Health
PI: principal investigator
PRICE: Pain Research and Intervention Center of Excellence
PROMIS: Patient-Reported Outcomes Measurement Information System
REDCap: Research Electronic Data Capture
SMILE: Study on Medical Marijuana and Its Long-Term Effects on Older Adults
SPIROS: Standardized Protocol Items Recommendations for Observational Studies
THC: tetrahydrocannabinol
UF: University of Florida
VA: Veterans Affairs

Edited by A Schwartz; submitted 18.Jun.2025; peer-reviewed by A AL-Asadi, CY Chen, L Panton; revised version received 17.Nov.2025; accepted 30.Nov.2025; published 24.Mar.2026.

Please cite as:
Robinson KR, Seeger SD, Nave L, Mejia M, Vander Meulen M, Rashid A, Mickle AM, Sibille KT, Li Z, Przkora R, Schmidt S, Lo MC, Cook RL, Wang Y
Real-Time and Long-Term Effects of Medical Marijuana on Older Adults: Protocol for a Prospective Cohort Study
JMIR Res Protoc 2026;15:e78900
URL: https://www.researchprotocols.org/2026/1/e78900
doi:10.2196/78900

© Kendall R Robinson, Stella D Seeger, Lauren Nave, Marlin Mejia, Maria Vander Meulen, Ahmed Rashid, Angela M Mickle, Kimberly T Sibille, Zhigang Li, Rene Przkora, Siegfried Schmidt, Margaret C Lo, Robert L Cook, Yan Wang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Follow-Up of Low-Acuity Patients After Redirection From a Swiss Emergency Department Using an Electronic Triage Application: Protocol for a Single-Center Prospective Cohort Study

Flora Gobet[1,2], MSc; Anne-Laure Feral-Pierssens[3,4], MD; Thomas Castelain[1,2], MSc; Ludovic Galofaro[1,5], MD; Hugo Najberg[6], PhD; Vincent Ribordy[1,2], MD; Youcef Guechi[1,2,4], MD

[1]Faculty of Medicine, University of Fribourg, Chemin du Musée 8, Fribourg, Switzerland

[2]Department of Emergency Medicine, Fribourg Hospital, Chemin des Pensionnats 2/6, Villars-sur-Glâne, Switzerland

[3]Department of Emergency Medicine, Avicenne Hospital, Public Assistance Hospitals Paris, Bobigny, France

[4]Laboratory for Education and Health Promotion, University Paris-Sorbonne, Villetaneuse, France

[5]Department of Anesthesiology and Intensive Care, Hospital Oberwallis, Visp, Switzerland

[6]Directorate of IT Services, University of Fribourg, Fribourg, Switzerland

Corresponding Author:
Youcef Guechi, MD
Faculty of Medicine, University of Fribourg, Chemin du Musée 8, Fribourg, Switzerland

Abstract

Background: Emergency department (ED) overcrowding threatens health care systems worldwide. This poses risks to patient safety, lowers quality of care, and reduces patient satisfaction. Patient input, defined as the caseload of patients presenting to the ED, is one of the factors contributing to overcrowding. Redirecting patients with nonurgent complaints to external health care services could help alleviate ED workload.

Objective: The primary objective of this study is to investigate the safety and efficiency of redirection of low-acuity patients using an electronic clinical decision support system. As a secondary and exploratory aim, we will assess the impact of this intervention on subsequent health care use, patient satisfaction, and health literacy over a 6-month follow-up period.

Methods: A single-center observational study with 2 consecutive cohorts of low-acuity patients presenting to the ED will be conducted. The first cohort will be triaged and offered redirection according to current practice. In the second cohort, after triage, low-acuity patients will be evaluated by an electronic clinical support system to determine eligibility for redirection. If eligible for redirection, an appointment at a nearby clinic will be arranged through the system. The primary end point is any unexpected returns to health care services within 48 hours of triage. Secondary end points are patient satisfaction as well as the number of unexpected consultations and the evolution of health literacy during a 6-month follow-up period.

Results: This study is funded by an internal grant from Hôpital Fribourgeois. Recruitment began on June 20, 2025, with a total of 35 patients enrolled as of August 2025. Data analysis will begin after the recruitment and 6-month follow-up of 300 to 420 patients are completed, which is expected to occur in November 2026. We hope to publish results in December 2026.

Conclusions: We expect the redirection process of low-acuity patients to other health care facilities using an electronic clinical support system to be safe and efficient. If results are positive, application of this reproducible strategy could reduce the number of patients treated in EDs and provide alternative health care pathways for low-acuity patients.

Trial Registration: ClinicalTrials.gov NCT06971419; https://clinicaltrials.gov/study/NCT06971419

International Registered Report Identifier (IRRID): DERR1-10.2196/82195

(JMIR Res Protoc 2026;15:e82195) doi:10.2196/82195

KEYWORDS

emergency service; low-acuity emergency department visit; low-acuity ED visit; nonurgent emergency department visit; nonurgent ED visit; patient acuity; triage; referral and consultation; redirection; LGI Emergency Redirection; health literacy; crowding

Introduction

Emergency department (ED) overcrowding is a global issue with significant consequences for both quality of care and the overall performance of health care systems. EDs operating overcapacity have been associated with poorer patient outcomes, increased mortality, decreased patient satisfaction, elevated staff stress, and lower adherence to clinical guidelines [1,2]. A 3-level model (input, throughput, and output) has been proposed to

XSL•FO
RenderX

describe the causes and potential solutions to ED overcrowding [3]. While most interventions aim to improve patient flow by focusing on efficiency during a patient's ED stay, interventions focusing on patients before ED admission remain largely understudied [4].

In parallel, high numbers of low-acuity patients increase ED workload and may divert resources from critically ill or high-complexity cases. The use of EDs by low-acuity patients can be partially linked to limited access to timely primary care consultations [2,5,6]. Furthermore, low health literacy also contributes to inappropriate use of EDs and is linked to poorer ED outcomes [7-9]. Consequently, low-acuity patients represent a potentially modifiable aspect of ED input, and redirecting those who do not require hospital-based care to external medical services [10] has emerged as a proposed strategy to reduce ED crowding.

Although redirection alone is not sufficient to resolve the underlying structural causes of ED overcrowding, which include a lack of capacity in outpatient services and access block [1], it may contribute to improving patient experience and local workflow. Perceived overcrowding and long wait times are directly associated with a significant decline in real-time patient satisfaction [11]. Redirection alleviates this problem by offering patients with minor conditions a faster alternative, preventing unnecessary waiting time in highly specialized medical centers, while allowing resources to be focused on patients who need them most [12,13]. This restructuring also benefits staff, as excessive congestion is a major cause of emotional exhaustion and job dissatisfaction. Therefore, lightening the load of nonurgent cases can improve working conditions for clinicians.

Redirection criteria must reliably and reproducibly identify patients whose needs can be met by external health care providers. Traditional ED triage systems are limited as they classify patients by vital risk and urgency to determine priority in admission and do not determine eligibility for redirection [14,15]. Previous research at a university hospital in Montreal demonstrated that an electronic clinical support system (LGI Emergency Redirection) could safely redirect patients, with an unexpected return rate below 5% and improvements in ED performance indicators [16,17].

Nonurgent consultations in the ED have been on the rise in the last decade in Switzerland, with more patients being self-referred, and the most frequently reported reason for self-referral is a lack of awareness of alternative care options[6]. There is growing interest in examining the feasibility of selectively redirecting low-acuity patients to alternative health care services [18]. To this end, we aim to assess the safety of implementing an electronic clinical support tool to identify patients eligible for redirection from a Swiss ED to external health care providers. To our knowledge, this is the first time such a redirection strategy has been studied in Switzerland. We aim to show that this strategy will be safe and will prove to be an effective decision support addition to current triage practice. Furthermore, we aim to evaluate the impact of this structured redirection on patient satisfaction, subsequent unplanned consultations, and health literacy, predicting that the formal appointment scheduling provided by the system will lead to improved longitudinal quality of care.

Methods

Study Design and Setting

This study will be a single-center, prospective observational cohort study focused on selective redirection of low-acuity patients, as determined by the electronic clinical support system (LGI Emergency Redirection, LGI Healthcare Solutions). Recruitment will take place between June 2025 and May 2026. The study will consist of 2 consecutive phases, detailed below. Each phase will last approximately 4 to 6 months, with a planned follow-up period of 6 months after the initial visit. The total expected study duration is approximately 15 months.

Participants will be recruited during triage at the ED of the Cantonal Hospital of Fribourg in Switzerland. This is a regional multidisciplinary adult ED managing around 41,000 annual ED visits.

Triage is based on the revised Swiss Emergency Triage Scale (SETS), which categorizes patients into 4 urgency levels based on presenting complaint, medical history, and vital signs [19]. Critical conditions requiring immediate medical attention are classified as SETS level 1, while SETS level 4 represents nonurgent conditions. Some presenting complaints can be classified at multiple levels depending on vital signs and medical history (eg, cardiac arrest=1, chest pain=1 or 2, abdominal pain=2 or 3, and dysuria=3 or 4). Since 2022, our ED has redirected level 4 patients by advising them to consult alternative health care providers (eg, walk-in clinics and general practitioners [GPs]). The safety of this approach has been established in a previous study [20].

Electronic Clinical Decision Support System: Operational Steps and Case Scenario

Overview

The electronic clinical decision support system used in this study (LGI Emergency Redirection) was previously validated in an observational study conducted in Canada to safely redirect low-acuity ED patients [17]. The system works as a 3-step process, which can only be performed after the usual first step of the local triage protocol has been completed. Triage personnel must verify the absence of general criteria precluding redirection (eg, chest pain, active neoplasia, dialysis patients, and unstable vital signs). They then choose from a list of 53 specific minor complaints, which enables access to complaint-specific contraindications to redirection. This 3-step process is illustrated in Figure 1. If contraindications are cleared and the patient consents to redirection, an appointment within 24 hours at an alternative outpatient health care facility is made by the triage nurse.

XSL•FO
RenderX

Figure 1. Example of the 3-step process of the electronic clinical support system for epistaxis (LGI Emergency Redirection).

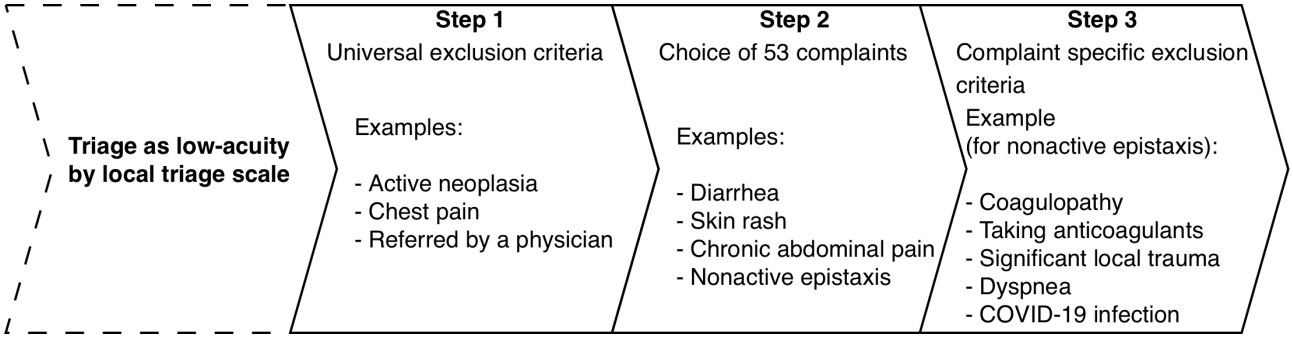

Two main alternative care pathways are available to redirected patients. First, patients can have an appointment at 1 of 5 clinics in the regional outpatient care network directly scheduled through the LGI Emergency Redirection platform. Four of these are walk-in clinics affiliated with the same hospital group. One is a local private clinic affiliated with the regional medical care planning network. The regional on-call general practitioner is a member of the regional care network to which redirected patients can also be referred. All of the abovementioned health care providers, including GPs, have access to on-site laboratories and basic diagnostic imaging facilities. These clinics are distributed across the region of Fribourg Canton, within 2.5 to 29 km from our ED.

Second, patients may be referred to physicians outside the regional outpatient care network, including their own GP. In these cases, direct digital appointment scheduling is not supported, and appointments are coordinated through regular communication channels (eg, phone call). Once the appointment has been arranged, the exact date and time (within 24 h) are communicated to the patient before they leave the ED.

Consider the example of a 28-year-old patient presenting to the ED with a resolved epistaxis. Once admitted by dedicated staff, the patient is assessed by a triage nurse who classifies them as SETS level 4 (nonurgent). Once this triage has been completed, the nurse opens the LGI Emergency Redirection application. After ensuring that there are no universal exclusion criteria (eg, significant difficulty communicating with the patient or their family or acute mental health issues), the triage nurse selects "non-active epistaxis" in the software. The system then displays new contraindications for referral, specific to the complaint (coagulopathy, taking anticoagulants, significant local trauma, etc), to confirm eligibility. Once the patient's eligibility has been verified and their consent for redirection obtained, the nurse uses the application's integrated booking interface to obtain a confirmed appointment at a partner clinic within 24 hours or with a GP by telephone. The process ends with the generation of a referral document for the patient and the secure transmission of triage data to the receiving facility, thereby formalizing the transition of care.

Phase 1: Current Practice

During the first phase, expected to last 4 to 6 months, triage and redirection of patients will follow the current practice at our ED, as described earlier. Additionally, during this phase, patients triaged as SETS levels 3 and 4 will be evaluated for their theoretical eligibility for redirection using the electronic clinical support system. At this stage, redirection will only be offered to SETS level 4 patients and will not be based on the electronic clinical support system. However, patients deemed theoretically suitable for redirection by the clinical support system will be eligible for study participation, regardless of whether redirection is performed (Figure 2). This group represents the patient population that could benefit from redirection using the electronic clinical decision support system.

This phase also allows our triage personnel to familiarize themselves with the electronic clinical support system.

XSL•FO
RenderX

Figure 2. Study groups (phase 1). ED: emergency department; SETS: Swiss Emergency Triage Scale.

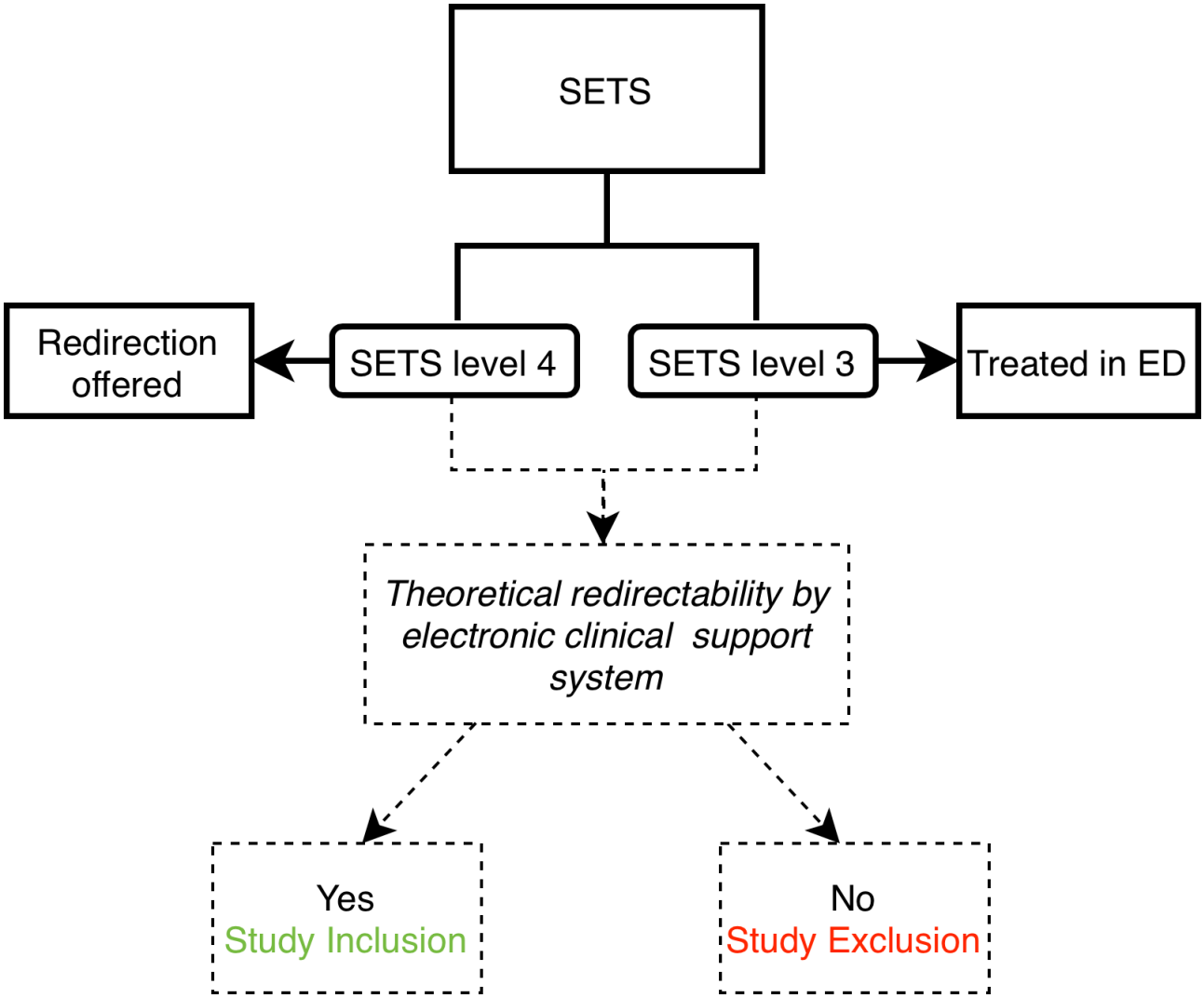

Phase 2: Redirection of Low-Acuity Patients Using the Electronic Clinical Decision Support System

During the second phase, expected to last 4 to 6 months, triage and redirection of patients will be performed by combining the SETS and the electronic clinical decision support system, as illustrated in Figure 3. Patients will first be assessed for clinical acuity using the SETS. Patients classified as low-acuity (SETS levels 3 and 4) will then be assessed by the electronic clinical support system and offered redirection if they are deemed eligible. Patients will be eligible for study inclusion regardless of whether they accept redirection.

XSL•FO
RenderX

Figure 3. Study groups (phase 2). ED: emergency department; SETS: Swiss Emergency Triage Scale.

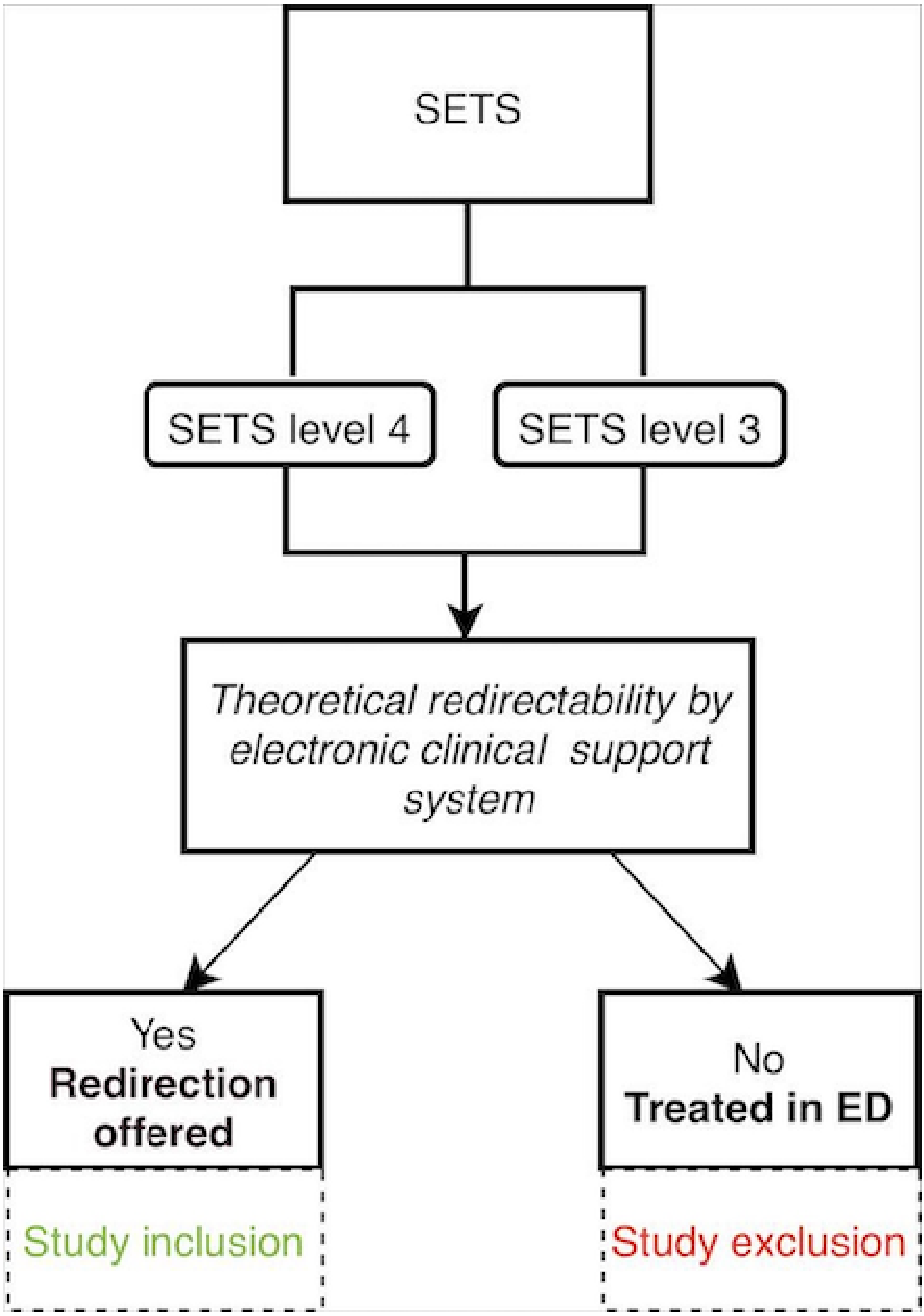

XSL•FO
RenderX

Population and Recruitment

The inclusion and exclusion criteria are presented in Textbox 1.

Eligible patients will be identified by triage nurses after being classified as low acuity using the revised SETS and evaluated as suitable for redirection through the electronic clinical decision support system. If appropriate, the triage nurse will offer the eligible patient the option of being redirected to alternative health care providers. Patients will be informed about the study by a research assistant only after they have made their choice regarding the care pathway (acceptance of redirection or refusal and continued ED consultation).

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Age ≥18 years
- Classified as low-acuity (Swiss Emergency Triage Scale [SETS] levels 3 or 4)
- Deemed eligible for redirection by the electronic clinical decision support system (LGI Emergency Redirection)
- Able to speak and read either French or German
- Oral consent obtained at day 2, followed by written informed consent

Exclusion criteria

- Inability to comply with study procedures (eg, cognitive impairment, significant hearing impairment without hearing aids, acute psychiatric conditions, or unavailability for telephone follow-up over the next 6 months)

Data Collection

Baseline demographic and clinical data will be extracted from electronic medical records. Outcome data will be collected over a 6-month follow-up period, as illustrated in Table 1. Unexpected return visits to any health care facility within 48 hours of triage will be assessed via electronic medical record review and confirmed by the research team through a phone interview. Unexpected returns are defined as any medical visit related to the initial complaint and not scheduled as part of the standard discharge instructions or follow-up care plan during the initial visit. This includes unplanned consultations in any health care setting. Research assistants will conduct phone interviews with participants asking about any unexpected consultations at any health care facility (at day 2, day 7, month 1, month 3, and month 6) and their satisfaction with the redirection process (at day 2). For each study time point, our team will make 3 attempts to reach participants by phone. Health literacy will be evaluated by a written questionnaire at day 7 and month 6.

Table . Study schedule.

	Day 0 (prescreening)	Day 2 (screening)	Day 7	Month 1	Month 3	Month 6
Days	0	2	7 (+1 or −1)	30 (+4 or −4)	91 (+7 or −7)	183 (+7 or −7)
Visit	Emergency department	Phone	Phone	Phone	Phone	Phone
Eligibility	✓[a]	—[b]	—	—	—	—
Study Information	✓ (Oral and written)	—	—	—	—	—
Consent	—	✓ (Oral)	✓ (Written)	—	—	—
Confirmation of eligibility and enrollment	—	✓	—	—	—	—
Study questionnaire by phone	—	✓	✓	✓	✓	✓
HLS-EU-Q16[c] by post or online	—	—	✓	—	—	✓

[a]✓ Refers to data being collected at this timepoint.

[b]Refers to data not being collected at this time point.

[c]HLS-EU-Q16: European health literacy survey questionnaire.

XSL•FO
RenderX

Outcomes

The primary outcome is defined as an unexpected return visit to any health care facility related to the initial complaint and occurring within 48 hours of initial triage assessment. To minimize underreporting, return visits are not restricted to our ED alone, acknowledging that patients may seek subsequent care at other EDs or with other health care providers.

Secondary study end points will include the following:

- Unexpected return visits related to the initial complaint occurring at any health care facility within 7 days of initial triage
- The number of unexpected visits related to any complaint at any health care facility reported at 1, 3, and 6 months after initial triage
- Patient satisfaction with the redirection process, as assessed by a numerical scale ranging from 0 (not at all satisfied) to 10 (perfectly satisfied)
- Patient health literacy, as assessed by the European health literacy survey (HLS-EU-Q16) questionnaire using a 4-point Likert scale, administered 7 days and 6 months after initial triage

Handling of Missing Data

Data will be systematically screened for missing values. Exploratory analyses will assess patterns of missing data in relation to study design variables, such as time of triage, day of the week, and seasonality. Rates of participant dropout and missing data will be reported transparently. Moreover, the study duration may be extended to ensure sufficient data collection through the completion of the second phase.

Statistical Analysis Plan

Statistical analyses will be performed using R (R Foundation for Statistical Computing) base functions if not stated otherwise. Cohen *d* values will be computed using the DescTools R package [21], and noninferiority and nonsuperiority tests will be computed using the TOSTER R package [22,23]. The significance threshold will be $P<.05$.

The season (ie, winter, spring, summer, or autumn), which affects seasonal pathologies, staffing, and workload; age; and pathology category (ie, dermatology, urology, psychiatry, etc) during triage should either be balanced between groups (electronic decision support tool vs control) or should have only a small impact on our outcomes to allow a robust interpretation of our hypotheses.

For the season and pathology categories, both groups are considered balanced if the Cramér V of the contingency table is below 0.3 (small association). For age, the groups are considered balanced if they do not differ by more than a Cohen *d* of 0.4 (small difference).

If the groups are not balanced on a given covariate, its impact on our outcomes (ie, frequency of redirection, amount of unplanned consultations, patient satisfaction, and health literacy) will be assessed using Cohen *d* values, Pearson correlations, and Cramér V values across both groups. These effect sizes should be small, that is, below 0.4 for Cohen *d* values and Pearson correlations, and below 0.3 for Cramér V values.

If the groups are not balanced and the covariate has an impact on an outcome, patients who are further away from their group's average will be excluded until the threshold for balanced groups is met.

The following hypotheses will be tested:

1. We expect patients redirected using the electronic decision support tool to have a lower or equivalent frequency of unexpected return visits in the next 48 hours and 7 days than control low-acuity patients, as indexed by a nonsuperiority test of 2 proportions with an upper bound of 5%.
2. We expect the number of unexpected visits related to any complaints in the 6-month follow-up period to be substantially lower in the electronic decision support tool group than in the control group, as indexed by a 1-sided *t* test.
3. We expect satisfaction to be noninferior in the electronic decision support tool compared with the control group, as indexed by a 1-sided *t* test with a lower margin of 2, which would correspond to the minimum amount for our 0 to 10 scale to have a shift in its interpretation.
4. We expect that the evolution of health literacy will increase significantly more in the electronic decision support tool group than in the control group, as indexed by the interaction term of a 2 × 2 ANOVA with the between-subject variable group (electronic decision tool vs control) and within-subject variable time (day 7 vs month 6). To respect the directionality of the test, the *P* value will be halved if the contrast goes in the expected direction (ie, electronic decision tool>control), and the *P* value will not be interpreted as significant if it goes into the other direction.

While the redirection network includes multiple types of facilities, interprovider variability will not be taken into account, as the study focuses on the overall safety and effectiveness of the redirection process itself rather than on the specific performance of individual partner facilities.

Sample Size

For the first hypothesis, a priori power analysis using the pwr R package [24] indicates that 137 patients per group (274 in total) will be needed to reach 80% power for a 2-proportion test with an alpha of .05 and a smallest effect size of interest of Cohen d=0.3 (small). For the second and third hypotheses, 138 patients per group (276 in total) will be needed for a 1-sided *t* test with the same parameters and a smallest effect size of interest of Cohen d=0.3 (small). For the fourth hypothesis, a priori power analysis computed with G*Power [25] states that 99 patients per group (198 in total) will be needed for a 2 × 2 within-between ANOVA with a smallest effect size of interest of partial η^2=0.01 (small).

On the basis of a retrospective study previously conducted in our department, we estimated that we would be able to include between 150 and 210 patients per phase, bringing the total to

XSL•FO
RenderX

between 300 and 420 participants over the total study duration, which is enough to reach sufficient power for all hypotheses.

Ethical Considerations

The clinical support system used in the study has been proven to be safe in redirecting ED patients of a university hospital, with fewer than 5% of patients returning unexpectedly [17]. The study protocol has been reviewed and approved by the Regional Ethics Committee of Vaud Canton (2023 - 01655). The study will be conducted in accordance with the principles of the Declaration of Helsinki and applicable Swiss regulations. Study participation will not affect patients' orientation or care, as they will be recruited after having accepted or declined redirection. Patients interested in participating in the study will receive a written study information and consent form at triage. On day 2, the study team will contact interested patients by phone to obtain informed oral consent. If the patient agrees, they will be included in the study and asked to return the signed informed consent form. Patients will not be compensated for participation. Each participant will be assigned a unique study identification number, and the corresponding deidentified data will be securely stored and managed using the electronic case report form system REDCap (Research Electronic Data Capture; Vanderbilt University). Study collaborators involved in data collection will have access to identifiable data, which they must maintain in strict confidence. Identifiable data will be securely stored in a locked folder within the Department of Emergency Medicine. Collaborators not involved in data collection, such as the statistical analysis team, will only have access to codified data through the electronic case report form platform. All study-related data will be archived for at least 10 years after study completion or premature termination.

Results

This study was funded by an internal grant from Hôpital Fribourgeois in June 2024. We expect to enroll an average of 1 to 2 patients a day, 5 days a week, over a period of 8 to 12 months, for a total of 300 to 420 patients. Recruitment began in June 2025 and is expected to be completed by May 2026. As of August 2025, a total of 35 patients were enrolled. Data analysis will be conducted after completion of the 6-month follow-up period of included patients, which we expect to end in November 2026. We hope to publish the results in December 2026.

Discussion

Anticipated Findings

Patients redirected using the electronic decision support tool will be compared with a control group of low-acuity patients. The presence of a comparator group strengthens our study by providing data on the outcomes of the 2 cohorts and allowing us to assess the intervention's impact.

This study is, to the best of our knowledge, the first study addressing the impact of redirection using an electronic clinical decision support system on patients' future use of the health care system and health literacy.

However, the study presents some limitations. First, it is a single-center study conducted in an urban regional hospital that benefits from the proximity of multiple walk-in clinics offering suitable care for low-acuity conditions. This may limit the generalizability of findings to settings with fewer outpatient resources, such as rural or underserved areas.

Health literacy will be assessed using a validated questionnaire (HLS-EU-Q16) that our team has translated into the most spoken regional languages. While these translations have undergone expert review and pretesting, they have not yet been formally validated through psychometric evaluation, which may introduce measurement bias. Patient satisfaction with the redirection process will be explored using a quantitative Likert scale questionnaire. However, no qualitative component is planned, which limits our ability to capture patients' in-depth perspectives or subjective experiences with the redirection process.

Selection bias might arise from differences in socioeconomic status or health literacy in participants accepting redirection and those refusing it. While the electronic clinical support system aims to redirect patients based on objective criteria, there might be some variability in criteria interpretation between triage nurses. In addition, outcome assessors will not be blinded to the redirection status of patients, which may introduce interviewer bias during follow-up calls. While we will follow up with patients at regular intervals, some patients might fail to report unplanned health care use, posing a risk of recall bias.

This study does not investigate the strategy's impact on ED crowding or performance indicators, as has been assessed in the literature [16]. Furthermore, none of the study end points directly measure quality of care delivered in the ED. As such, this limits our ability to draw conclusions regarding the redirection strategy's impact on care quality or ED overcrowding.

Management of low-acuity conditions in EDs is generally associated with higher costs compared with treatment in primary care or urgent care settings [26,27]. However, existing redirection interventions have not consistently demonstrated significant reductions in health care costs [28,29]. This study does not include cost-related end points. If the redirection strategy is found to be safe and effective, future studies should explore its economic implications, including cost-effectiveness analyses.

Conclusions

We expect the redirection strategy guided by an electronic clinical decision support tool to be safe and result in less than 5% of unexpected returns within 48 hours of initial presentation. This approach may contribute to optimizing patient pathways by facilitating the redirection of low-acuity patients whose needs can be adequately addressed in alternative care settings, thereby allowing ED resources to remain focused on complex and acute cases. A key strength of this intervention is the standardized use of objective criteria through an electronic tool, enhancing reproducibility and consistency in clinical decision-making. Its design also facilitates transposability to other EDs if their setting permits redirection to alternative health care services.

XSL•FO
RenderX

To our knowledge, this is the first study assessing the impact of this redirection strategy on health care pathways, future unexpected consultations, and health literacy. If the redirection strategy proves to limit future unexpected visits and better health literacy, it would have a lasting impact on patient education and appropriate use of health care resources. Future research should investigate additional outcomes, including the impact of redirection on ED crowding, care quality, patient-reported experiences, and cost-effectiveness. These insights will be essential to fully understand the broader implications and scalability of the intervention.

Acknowledgments

DeepL Pro by DeepL SE was used for the translation of specific terms and grammar correction. The authors would like to acknowledge Laurence Cagnon, Alexandre Sarre, Julien Sauser, and Guliz Vanli-Jaccard of the clinical research center at the Lausanne University Hospital for their methodological support.

Funding

This study is funded by Hôpital Fribourgeois (grant HFR 24-09).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: FG, ALF-P, YG
Formal analysis: HN
Funding acquisition: VR, YG
Methodology: FG, ALF-P, LG, HN, YG
Project administration: YG
Resources: FG, TC
Supervision: VR, YG
Visualization: TC
Writing—original draft: FG, YG
Writing—review and editing: FG, TC, YG

Conflicts of Interest

None declared.

References

1. Morley C, Unwin M, Peterson GM, Stankovich J, Kinsman L. Emergency department crowding: a systematic review of causes, consequences and solutions. PLoS One 2018 Aug 30;13(8):e0203316. [doi: 10.1371/journal.pone.0203316] [Medline: 30161242]
2. Bond K, Ospina MB, Blitz S, et al. Frequency, determinants and impact of overcrowding in emergency departments in Canada: a national survey. Healthc Q 2007;10(4):32-40. [doi: 10.12927/hcq.2007.19312] [Medline: 18019897]
3. Asplin BR, Magid DJ, Rhodes KV, Solberg LI, Lurie N, Camargo CAJ. A conceptual model of emergency department crowding. Ann Emerg Med 2003 Aug;42(2):173-180. [doi: 10.1067/mem.2003.302] [Medline: 12883504]
4. Samadbeik M, Staib A, Boyle J, et al. Patient flow in emergency departments: a comprehensive umbrella review of solutions and challenges across the health system. BMC Health Serv Res 2024 Mar 5;24(1):274. [doi: 10.1186/s12913-024-10725-6] [Medline: 38443894]
5. Moineddin R, Meaney C, Agha M, Zagorski B, Glazier RH. Modeling factors influencing the demand for emergency department services in Ontario: a comparison of methods. BMC Emerg Med 2011 Aug 19;11:13. [doi: 10.1186/1471-227X-11-13] [Medline: 21854606]
6. Diserens L, Egli L, Fustinoni S, Santos-Eggimann B, Staeger P, Hugli O. Emergency department visits for non-life-threatening conditions: evolution over 13 years in a Swiss urban teaching hospital. Swiss Med Wkly 2015 Apr 9;145:w14123. [doi: 10.4414/smw.2015.14123] [Medline: 25856789]
7. Griffey RT, Kennedy SK, D'Agostino McGowan L, Goodman M, Kaphingst KA. Is low health literacy associated with increased emergency department utilization and recidivism? Acad Emerg Med 2014 Oct;21(10):1109-1115. [doi: 10.1111/acem.12476] [Medline: 25308133]
8. Herndon JB, Chaney M, Carden D. Health literacy and emergency department outcomes: a systematic review. Ann Emerg Med 2011 Apr;57(4):334-345. [doi: 10.1016/j.annemergmed.2010.08.035] [Medline: 21035902]

XSL•FO
RenderX

9. Slagman A, Schmiedhofer M, Legg D, et al. Reasons for emergency department use of low-acuity attender: results from the prospective, multicentre, cross-sectional EPICS-9/PiNo-Bund study. BMJ Open 2025 Mar 3;15(3):e090681. [doi: 10.1136/bmjopen-2024-090681] [Medline: 40032406]
10. Sartini M, Carbone A, Demartini A, et al. Overcrowding in emergency department: causes, consequences, and solutions-a narrative review. Healthcare (Basel) 2022 Aug 25;10(9):1625. [doi: 10.3390/healthcare10091625] [Medline: 36141237]
11. Wang H, Kline JA, Jackson BE, et al. The role of patient perception of crowding in the determination of real-time patient satisfaction at emergency department. Int J Qual Health Care 2017 Oct 1;29(5):722-727. [doi: 10.1093/intqhc/mzx097] [Medline: 28992161]
12. O'Keeffe C, Mason S, Jacques R, Nicholl J. Characterising non-urgent users of the emergency department (ED): a retrospective analysis of routine ED data. PLoS One 2018 Feb 23;13(2):e0192855. [doi: 10.1371/journal.pone.0192855] [Medline: 29474392]
13. Phillpotts EL, Coggins J, Goodwin L, et al. Streaming and redirection of lower acuity adult patients attending the ED: patient and staff experience. Emerg Med J 2025 Sep 24;42(10):629-635. [doi: 10.1136/emermed-2024-214767] [Medline: 40571395]
14. Hinson JS, Martinez DA, Cabral S, et al. Triage performance in emergency medicine: a systematic review. Ann Emerg Med 2019 Jul;74(1):140-152. [doi: 10.1016/j.annemergmed.2018.09.022] [Medline: 30470513]
15. Zachariasse JM, van der Hagen V, Seiger N, Mackway-Jones K, van Veen M, Moll HA. Performance of triage systems in emergency care: a systematic review and meta-analysis. BMJ Open 2019 May 28;9(5):e026471. [doi: 10.1136/bmjopen-2018-026471] [Medline: 31142524]
16. Feral-Pierssens AL, Gaboury I, Carbonnier C, Breton M. Redirection of low-acuity emergency department patients to nearby medical clinics using an electronic medical support system: effects on emergency department performance indicators. BMC Emerg Med 2024 Sep 13;24(1):166. [doi: 10.1186/s12873-024-01080-0] [Medline: 39272018]
17. Feral-Pierssens AL, Morris J, Marquis M, et al. Safety assessment of a redirection program using an electronic application for low-acuity patients visiting an emergency department. BMC Emerg Med 2022 Apr 29;22(1):71. [doi: 10.1186/s12873-022-00626-4] [Medline: 35488215]
18. Slagman A, Bremicker A, Möckel M, Eienbröker L, Fischer-Rosinský A, Gries A. Evaluation of an automated decision aid for the further referral of emergency room patients—a prospective cohort study. Dtsch Arztebl Int 2024 Oct 18;121(21):703-709. [doi: 10.3238/arztebl.m2024.0191] [Medline: 39323221]
19. Rutschmann OT, Hugli OW, Marti C, et al. Reliability of the revised Swiss Emergency Triage Scale: a computer simulation study. Eur J Emerg Med 2018 Aug;25(4):264-269. [doi: 10.1097/MEJ.0000000000000449] [Medline: 28099182]
20. Déglise A, Guechi Y, Le Terrier C, Ribordy V, Feral-Pierssens AL, Schmutz T. SAFEty assessment of a REdirection process after TRIage (SAFE RETRI) by a triage nurse in an emergency department: a monocentric cohort study. Swiss Med Wkly 2024 Dec 20;154:4030. [doi: 10.57187/s.4030] [Medline: 39945074]
21. Signorell A, Aho K, Alfons A, et al. DescTools: tools for descriptive statistics. The Comprehensive R Archive Network. 2025. URL: https://cran.r-project.org/web/packages/DescTools/index.html [accessed 2026-02-20]
22. Caldwell AR. Exploring equivalence testing with the updated TOSTER R package. PsyArXiv. Preprint posted online on Nov 18, 2022. [doi: 10.31234/osf.io/ty8de]
23. Lakens D. Equivalence tests: a practical primer for t tests, correlations, and meta-analyses. Soc Psychol Personal Sci 2017 May;8(4):355-362. [doi: 10.1177/1948550617697177] [Medline: 28736600]
24. Champely S, Ekstrom C, Dalgaard P, et al. pwr: basic functions for power analysis. The Comprehensive R Archive Network. 2020. URL: https://cran.r-project.org/web/packages/pwr/index.html [accessed 2026-02-20]
25. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
26. Berthelot S, Mallet M, Simonyan D, et al. PL04: comparison of the cost and the quality of the care provided to low acuity patients in an emergency department and a walk-in clinic. Can J Emerg Med 2019 May;21(S1):S6. [doi: 10.1017/cem.2019.43]
27. Campbell MK, Silver RW, Hoch JS, et al. Re-utilization outcomes and costs of minor acute illness treated at family physician offices, walk-in clinics, and emergency departments. Can Fam Physician 2005 Jan;51(1):82-83. [Medline: 16926958]
28. Smulowitz PB, Honigman L, Landon BE. A novel approach to identifying targets for cost reduction in the emergency department. Ann Emerg Med 2013 Mar;61(3):293-300. [doi: 10.1016/j.annemergmed.2012.05.042] [Medline: 22795188]
29. Morreel S, Homburg I, Philips H, et al. Cost effects of nurse led triage at an emergency department with the advice to consult the adjacent general practice cooperative for low-risk patients, a cluster randomised trial. Health Policy 2022 Oct;126(10):980-987. [doi: 10.1016/j.healthpol.2022.08.002] [Medline: 35963797]

Abbreviations

ED: emergency department
GP: general practitioner
HLS-EU-Q16: European health literacy survey questionnaire
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

SETS: Swiss Emergency Trial Scale

Edited by J Sarvestan; submitted 11.Aug.2025; peer-reviewed by C Patey; revised version received 20.Feb.2026; accepted 23.Feb.2026; published 27.Mar.2026.

Please cite as:
Gobet F, Feral-Pierssens AL, Castelain T, Galofaro L, Najberg H, Ribordy V, Guechi Y
Follow-Up of Low-Acuity Patients After Redirection From a Swiss Emergency Department Using an Electronic Triage Application: Protocol for a Single-Center Prospective Cohort Study
JMIR Res Protoc 2026;15:e82195
URL: https://www.researchprotocols.org/2026/1/e82195
doi:10.2196/82195

© Flora Gobet, Anne-Laure Feral-Pierssens, Thomas Castelain, Ludovic Galofaro, Hugo Najberg, Vincent Ribordy, Youcef Guechi. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of Reduced Single-Photon Emission Computed Tomography Imaging Protocols and Software Variability in 177Lu-DOTATATE Dosimetry: Protocol for an Exploratory Observational Trial

Takayuki Yagihashi[1], PhD; Kenta Miwa[2], PhD; Noriaki Miyaji[2], PhD; Satoru Sugimoto[3], PhD; Hideki Hayakawa[4], BS; Noritoshi Kobayashi[5], MD, PhD; Shoko Takano[6], MD, PhD; Taro Murai[7], MD, PhD

[1]Department of Medical Physics, Shonan Kamakura General Hospital, Kamakura, Japan

[2]Department of Radiological Sciences, Fukushima Medical University, Fukushima, Japan

[3]Medical Science Data-driven Mathematics Team, Division of Applied Mathematical Science, RIKEN Center for Interdisciplinary Theoretical and Mathematical Sciences, Yokohama, Japan

[4]Radiology Division, Shonan Kamakura General Hospital, Kamakura, Japan

[5]Department of Oncology, Yokohama City University Graduate School of Medicine, Yokohama, Japan

[6]Department of Radiation Oncology, Yokohama City University Graduate School of Medicine, Yokohama, Japan

[7]Department of Radiation Oncology, Shonan Kamakura General Hospital, 1370-1 Okamoto, Kamakura, Japan

Corresponding Author:
Taro Murai, MD, PhD
Department of Radiation Oncology, Shonan Kamakura General Hospital, 1370-1 Okamoto, Kamakura, Japan

Abstract

Background: Targeted radiopharmaceutical therapy (TRT) offers a promising approach for cancer treatment by delivering radiation directly to tumor cells while sparing healthy tissues. Accurate dosimetry of organs and tumors is crucial to optimize therapeutic efficacy and minimize toxicity, particularly for dose-limiting organs such as the kidneys. Although routine dosimetry using single-photon emission computed tomography (SPECT) is recommended per guidelines, its widespread clinical application remains limited owing to a lack of consensus on the optimal frequency and timing of SPECT scans for accurate dosimetry, which leads to variability in clinical practice and hinders robust dose–response relationships. Furthermore, absorbed dose calculations rely on software-specific curve-fitting models. Thus, discrepancies in dose estimates among different simulation software programs pose a significant challenge to standardizing dosimetry workflows.

Objective: We aimed to explore the optimal SPECT imaging schedule, evaluate differences among simulation software programs, and establish a standardized protocol for future epidemiological research to define organ tolerance doses and facilitate wider adoption of personalized TRT regimens.

Methods: This exploratory observational trial will determine the optimal SPECT imaging protocol and consistency of dosimetry estimates using software programs for Lu-177-DOTATATE therapy in patients with neuroendocrine tumors. In a single treatment cycle, SPECT imaging will be performed at 4, 24, 96, and 168 hours following Lu-177-DOTATATE administration. The true absorbed dose cannot be directly measured; therefore, doses calculated using all time points (the 4-point method) for each patient will be used as the reference standard and compared with estimates derived from 1- to 3-point methods. Consequently, 15 dose simulations will be conducted per patient. This study implements a 5+5 design (rule-based design). The primary endpoint is the incidence of dose errors for each patient across different dose-calculation software systems. The kidney absorbed dose will be the primary focus of evaluation, with secondary analyses including tumor and other organ dosimetry. In principle, a dose error of <5% will be considered acceptable in each software system, depending on the absolute dose level. Differences among software systems and between patients will be evaluated using descriptive statistical methods. The study was approved by the Tokushukai Group Ethics Committee (Approval No. 2447).

Results: Recruitment began on March 5, 2024, and 4 participants have been enrolled as of September 2025. Data collection is expected to be completed by February 2027, with the study results anticipated in August 2027.

Conclusions: This study will evaluate whether simplified SPECT imaging protocols can maintain dosimetric accuracy while reducing the burden on patients and providers. It will also compare absorbed dose estimates across software programs to assess consistency and reliability as a reference. Findings will be disseminated through open-access peer-reviewed journals and relevant conferences and events.

XSL•FO
RenderX

Trial Registration: UMIN Clinical Trials Registry UMIN000057478; https://center6.umin.ac.jp/cgi-open-bin/ctr/ctr_view.cgi?recptno=R000065636

International Registered Report Identifier (IRRID): DERR1-10.2196/83248

(JMIR Res Protoc 2026;15:e83248) doi:10.2196/83248

KEYWORDS

targeted radiopharmaceutical therapy; cancer; radioactivity; dosimetry; Lutetium-177

Introduction

Targeted radiopharmaceutical therapy (TRT) offers a promising cancer treatment approach by delivering radiation directly to tumor cells while minimizing exposure to surrounding healthy tissue. These radiopharmaceuticals, composed of radioactive isotopes linked to tumor-targeting molecules, are particularly effective for advanced cancers and are being explored for a wider range of malignancies [1-4]. Notable examples include Lutetium-177 (Lu-177)-DOTATATE for neuroendocrine tumors (NETs) and Lu-177-prostate-specific membrane antigen for metastatic castration-resistant prostate cancer [3,4]. Despite TRT's expanding applicability, standardized methods for assessing absorbed doses and defining normal tissue tolerance thresholds remain underdeveloped.

Current guidelines from the European Council Directive 2013/59 (Article 56) and the European Association of Nuclear Medicine/Medical Internal Radiation Dose recommend routine assessment of absorbed radiation doses in each organ and tumor for patients undergoing TRT [5]. This dosimetry typically employs sequential quantitative single-photon emission computed tomography (SPECT) [5]. Accurate image quantification is crucial, as any underestimation or overestimation of absorbed doses can significantly affect clinical decisions and patient safety [5-7]. While SPECT images are routinely corrected for factors such as photon attenuation, scatter, and blurring to ensure quantitative accuracy [5-7], the widespread clinical application of standardized SPECT-based dosimetry remains limited.

Two primary challenges hinder the routine and standardized implementation of TRT dosimetry. First, there is no consensus on the optimal number or timing of SPECT scans. Accurate time–activity curves, which are essential for precise dose estimation, benefit from frequent imaging, particularly shortly after injection [8]. However, increasing SPECT scan frequency also raises patient burden, owing to additional radiation exposure, prolonged imaging time, and psychological stress, and strains health care resources. Consequently, current clinical practices for SPECT timing vary significantly, with protocols ranging from multiple scans (eg, 24, 96, and 168 hours post-injection) to a single scan at a specific time point [5,9-12]. This variability is particularly problematic for dose-limiting organs such as the kidneys, where precise dosimetry is critical to minimize toxicity. Second, absorbed dose calculations in TRT rely on various simulation algorithms, yet their dose estimates often lack universal validation. As a result, a dose reported by one software program may not be directly comparable with that from another, complicating the standardization of dosimetry workflows and hindering the establishment of robust dose–response relationships.

This study aims to identify the optimal SPECT imaging schedule for accurate TRT dosimetry and evaluate the differences in dose estimates among various simulation software programs. Ultimately, we seek to establish a standardized protocol for TRT dosimetry to support future epidemiological research, aid in defining organ tolerance doses, and facilitate the wider adoption of personalized TRT regimens, thereby optimizing therapeutic efficacy and minimizing toxicity.

Methods

Aims and Objectives

This study, conducted at Shonan Kamakura General Hospital, aims to optimize the SPECT imaging schedule, evaluate differences among simulation software programs, and develop a protocol suitable for epidemiological studies to define dose constraints for normal organs in Lu-177-DOTATATE therapy. Specifically, we will investigate whether the standard 4-time-point protocol could be reduced to 1-, 2-, or 3-scan protocols without compromising dosimetric accuracy. The purpose of this study is to explore practical time points for dosimetry within a single treatment cycle. Inter-cycle variability will be examined in subsequent studies after optimal time points are established in this study.

The primary endpoint is to evaluate the effect of scan reduction on the estimated absorbed dose to the kidneys, given that renal toxicity—particularly radiation-induced nephropathy—is the primary dose-limiting factor. Secondary endpoints include comparing tumor and normal organ dosimetry across imaging protocols and analyzing dose variations among software programs to assess consistency and reliability.

Design

This open-label, single-institution exploratory observational study will recruit patients with neuroendocrine tumors undergoing Lu-177-DOTATATE therapy. Gamma emissions from normal organs and tumors will be measured using SPECT imaging at 4, 24, 96, and 168 hours following radiopharmaceutical administration. These images will serve as inputs for simulation software to calculate absorbed doses to normal organs and target tissues.

This study adopts the 4-time-point SPECT protocol (4-point method) as the reference standard for each patient within each software program. Doses calculated after omitting 1 - 3 time points (1-, 2-, and 3-point methods) will be compared with the 4-point dose within each program. The evaluation frameworks

XSL•FO
RenderX

for each method are summarized below. Fifteen dose simulations will be conducted per patient. Data collection will span from March 2024 to February 2027.

For the 4-point method, dosimetry calculation will be performed using SPECT images taken at 4, 24, 96, and 168 hours. For the 3-point method, dosimetry calculation will be performed using SPECT images taken at 24, 96, and 168 hours; 4, 9, and 168 hours; 4, 24, and 168 hours; and 4, 24, and 96 hours. For the 2-point method, dosimetry calculation will be performed using SPECT images taken at 4 and 24 hours; 4 and 96 hours; 4 and 168 hours; 24 and 96 hours; 24 and 168 hours; and 96 and 168 hours. For the 1-point method, dosimetry calculation will be performed using SPECT images taken at 4 hours; 24 hours; 96 hours; and 168 hours post-injection

If dose discrepancies remain within acceptable limits, the simplified protocol will be considered for future trials. Additionally, absorbed doses calculated using each program will be compared across algorithms.

Study Design Considerations

This exploratory observational study will evaluate various SPECT imaging time-point protocols and dosimetry software programs in patients undergoing Lu-177-DOTATATE therapy. Patient randomization is not applicable because the objective is to evaluate dosimetric methods against a reference standard in a real-world clinical setting rather than to compare treatment groups.

This is an open-label study. Blinding of participants or investigators will not be performed, as the primary focus is on technical comparisons of dosimetry methodologies and software programs. Blinding of imaging protocols or software analyses would be impractical and irrelevant for assessing dosimetric accuracy in this context.

Safety and Data Monitoring

This study is designed as an exploratory observational trial and primarily involves the acquisition of additional SPECT/computed tomography (CT) images as part of the standard clinical protocol for Lu-177-DOTATATE therapy. As such, the research interventions are integral to routine clinical care and introduce minimal to no additional risks to participants beyond that associated with standard Lu-177-DOTATATE therapy.

The potential for additional radiation exposure from the extra SPECT/CT scans is considered negligible and will be kept as low as reasonably achievable, in accordance with the recommendations of the International Commission on Radiological Protection [13]. Any adverse events related to Lu-177-DOTATATE therapy will be managed and reported by the treating physicians according to standard clinical practice and existing institutional guidelines [14], and not as part of this study protocol, as the focus of this study is solely on dosimetry methodology.

Given the non-interventional and low-risk nature of this observational study, an independent data monitoring committee will not be established.

Data Handling and Exclusions

All acquired SPECT/CT images that meet technical quality control standards will be included in the dosimetry analysis. Data points, including individual SPECT scans, may be excluded if deemed technically inadequate—owing to motion artifacts, hardware malfunction, or incomplete acquisition—after review by a medical physicist, as these issues could compromise dose calculation accuracy. All exclusions will be thoroughly documented and reported.

Data Management

The data will be entered and stored using Microsoft Excel software, cross-checked in pairs, saved on an encrypted computer, and backed up.

Statistical Analyses and Sample Size

This study implements a 5+5 design, also known as a rule-based design [15]. The primary endpoint is the dose error incidence for each calculation algorithm. Figure 1 presents the decision-making workflow. The assumed dose error incidence in this study is set to be ≤5%; based on this assumption, the required number of eligible patients will be modified using the 5+5 design [16]. If the incidence (dose error per patient ≤ threshold) is ≤5% within the workflow, false-negative results will be observed in ≤23% of trials. If the incidence is ≥30%, false-positive results will occur in ≤17%.

Dose error below the predefined threshold will be considered acceptable for each calculation algorithm. The primary threshold will be set at 5%, following the International Commission on Radiation Units and Measurements Report 83 guidelines [17], which recommend that overall dosimetric uncertainty should ideally be within ±5%. For regions with absorbed doses <5 Gy, where larger errors are generally observed, the threshold will be relaxed to 10% [18]. A simplified method will be deemed “acceptable” if, in all five cases, the absorbed dose estimated by each software program falls within the threshold (5% or 10%, depending on the regional dose) when compared with the reference 4-point method. If multiple algorithms meet the criteria, priority will be given to the algorithm requiring the fewest time points for the subsequent trial. For example, if both the 3-point and 1-point methods meet the criteria, the 1-point method will be selected as the primary candidate. Among algorithms requiring the same number of time points, those incorporating the 4-hour or 24-hour sampling points will be preferred.

XSL•FO
RenderX

Figure 1. Workflow for stepwise evaluation of dose error acceptability. Dose errors are first evaluated in an initial set of five cases. If the dose error is within the predefined threshold in all five cases, the 2-point method is considered acceptable. If four out of five cases meet the criterion, the 4-point method is applied. If three or fewer cases meet the criterion, dosimetry is extended to five additional cases (total n = 10). Final acceptability is determined based on the number of cases within the threshold among all ten cases.

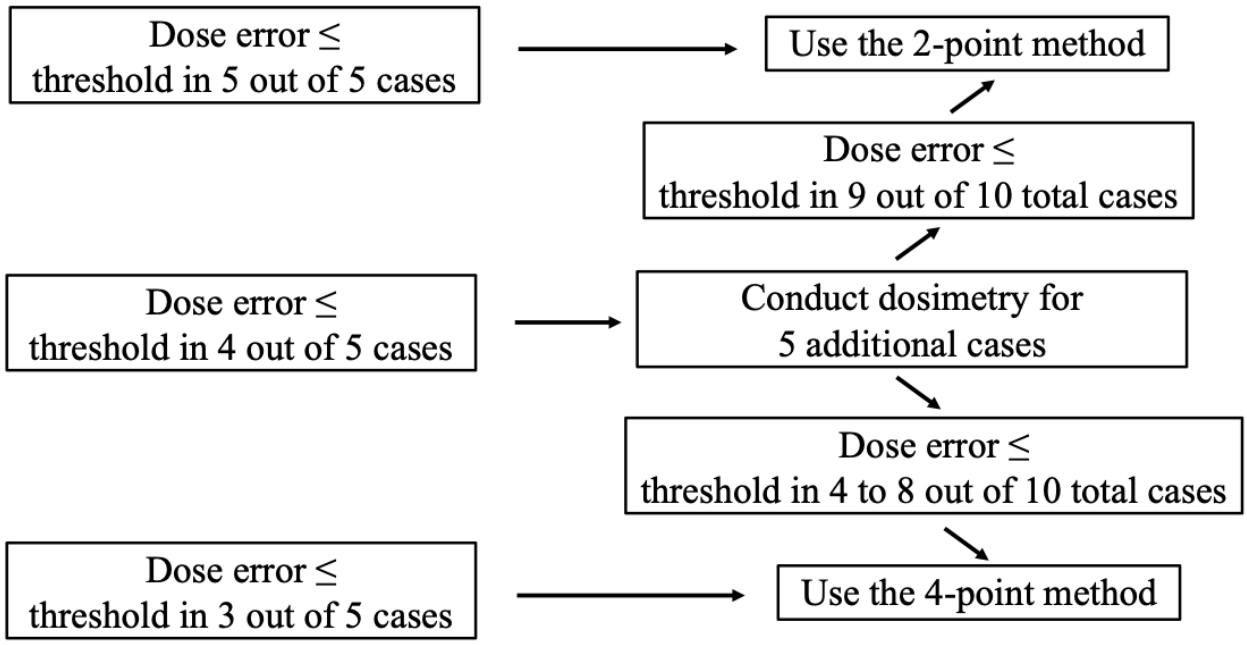

In external beam radiotherapy treatment planning, the total kidney dose threshold is 23 Gy [19], indicating that a dose of 5.5 Gy per treatment cycle is considered clinically acceptable. Therefore, we will relax the absorbed dose threshold from 5% to 10% if kidney doses are <5 Gy.

If the threshold is exceeded in 1 of the 5 patients, an additional 5 cases will be evaluated. The simplified method will be considered “acceptable” if at least 9 of the 10 cases meet the threshold. Conversely, if the dose error exceeds the threshold in 2 or more of the 10 cases, the simplified method will be considered “unacceptable.”

The true absorbed dose cannot be directly measured, and 4-point estimates are derived using software-specific kinetic models and dose-calculation frameworks. Therefore, it cannot be conclusively stated that these estimates are equivalent across different software programs. Consequently, consistency among software systems will be evaluated by comparing the magnitude and distribution of relative dose deviations from the 4-point reference and from the 1- to 3-point estimates using descriptive statistical methods. Visualization of these distributions, such as using box-and-whisker plots, may facilitate assessment of the validity of dose reduction at each time point. As a reference analysis, one-way analysis of variance will be used to compare absorbed doses to each organ of interest across software systems and between patients. The level of statistical significance is set at 0.05. If the *P* value is less than .05, the validity of the candidate calculation algorithm will be re-evaluated.

Inclusion and Exclusion Criteria

The eligibility criteria are presented in Textbox 1 .

XSL•FO
RenderX

Textbox 1. Eligibility criteria

Inclusion criteria

- Age ≥20 years
- Blood tests within 1 month before treatment
- Estimated glomerular filtration rate >50 mL/min
- Hemoglobin level >10 g/dL
- White blood cell count between 3000 and 10,000/μL
- Platelet count >75,000/μL
- Albumin level >3 g/dL
- Bilirubin level <3 mg/dL
- Hasegawa dementia scale ≥21
- Ability to self-feed
- Ability to manage insulin and other medications
- Ability to manage urination (continence aids allowed)
- World Health Organization performance status score of 0 or 1

Exclusion criteria

- Urgent/critical findings on chest/abdominal computed tomography
- Hypersensitivity to Lu-177 DOTATATE, L-lysine HCl, or L-arginine HCl
- Pregnant/breastfeeding or women of childbearing potential without adequate contraception
- Male with partner of childbearing potential without effective contraception
- Psychiatric disorders impairing participation
- Systemic infection requiring ongoing treatment
- Unable to remain in a fixed position for >30 min
- Any condition deemed inappropriate by the principal investigator

TRT Details

Lu-177-DOTATATE will be administered following the protocol previously approved by the Food and Drug Administration [20]. Based on this protocol, TRT will be administered in four treatment cycles per patient, with each cycle scheduled at 8-week intervals. In each cycle, Lu-177-DOTATATE (7.4 GBq) will be administered over 30 minutes. An amino acid solution containing 2.5% lysine and 2.5% arginine will be co-administered via intravenous infusion at approximately 1000 mL over 4 hours, beginning 30 minutes before TRT. This solution facilitates renal protection by promoting urinary excretion of Lu-177-DOTATATE, thereby reducing the dose absorbed by the kidneys. Prior to amino acid infusion, an appropriate antiemetic agent will be administered as an intravenous bolus to prevent nausea and vomiting.

Dosimetry Methods

Dosimetry Protocol

In this study, four quantitative SPECT/CT scans will be performed for dosimetry following the administration of Lu-177-DOTATATE during one of the four treatment cycles. These scans, covering from the neck to the femoral head region, will be acquired over 30 minutes using a Symbia Intevo 6 system (Siemens Healthineers, Erlangen, Germany) equipped with medium-energy general-purpose collimators. To optimize image quality and quantitative accuracy, image reconstruction will be conducted using the ordered subset conjugate gradient minimization algorithm (xSPECT) with a matrix size of

XSL•FO
RenderX

256×256. A Gaussian filter will be applied during image reconstruction. The primary energy peak will be centered at 208 keV (187.6 - 229.2 keV). Scatter correction will use the triple energy window method, and attenuation correction will be based on CT. CT images will be acquired at 130 kV and 20 mA using the adaptive dose modulation technique (CARE Dose 4D; Siemens Healthineers). To enable absolute quantification, the gamma camera will be calibrated using a Lu-177 point source. Dose calibration will be performed using serial SPECT/CT acquisitions of Lu-177 point sources in air and in a tissue-equivalent medium to simulate attenuation. Dead time will be determined and corrected based on wide-spectrum counting rates during image acquisition. To simulate absorbed doses in tumors and normal organs, the injected activity (GBq), date and time of TRT administration, and patient weight (kg) will be entered into the dosimetry software.

Subsequently, organ and target absorbed doses will be simulated using at least five commercially available software programs and one newly developed tool: (1) OLINDA/EXM (Hermes Medical Solutions); (2) Dosimetry Toolkit (GE HealthCare Medical Systems and Solutions); (3) Voxel Dosimetry (Hermes Medical Solutions); (4) MIM SurePlan MRT (MIM Software Inc.); (5) RT-PHITS (Japan Atomic Energy Agency); and (6) a newly developed dose calculation program introduced after 2022 (name and developer to be specified once the software is finalized and becomes available during the study period).

Quality Control and Assurance

To ensure the accuracy and reliability of dosimetry, this study will implement stringent quality control measures. Specifically, to evaluate inter-rater reliability in dose calculations, the following procedures will be carried out.

For all patient data, two independent radiation oncologists will each perform contouring of organs and tumors. If a discrepancy in contouring is observed between the two readings, the oncologists will discuss the findings and determine the final contours.

Dose calculations will be performed by an experienced and certified medical physicist. For quality assurance, a second experienced and certified medical physicist will independently verify the dose calculations.

Recruitment

Participants will be enrolled after being administered at least one TRT cycle to minimize the risk of allergic reactions to the administered agents. Patients will not be enrolled until they have completed at least one treatment cycle and their ability to attend follow-up visits at our hospital has been confirmed. Eligible individuals will be invited to participate and will be provided with detailed study information by their radiation oncologist to support informed decision-making.

Ethical Considerations

Ethics approval and consent to participate were reviewed and granted by the Tokushukai Group Ethics Committee Institutional Review Board (Approval No. 2447; UMIN000057478). Written informed consent will be obtained from all participants prior to study enrollment. All data will be de-identified prior to analysis, and no personally identifiable information will be accessible to the research team outside the treating institution. The study data will be handled in accordance with applicable data protection regulations. Participants will not receive any financial compensation for participation in this study. All study procedures will be conducted in accordance with the principles of the Declaration of Helsinki and applicable regulatory requirements.

Results

Recruitment for this study began on March 5, 2024. As of September 2025, 4 participants have been enrolled. Data collection is expected to be completed by February 2027, and the study results are anticipated to be published by August 2027.

Discussion

TRT is expected to become a standard treatment for various cancer types. While the prescribed dose should ideally be tailored to the tolerance of normal organs, standardized thresholds for normal tissue tolerance have not yet been fully established. Currently, no consensus exists on the optimal method for evaluating absorbed doses, including the appropriate number and timing of SPECT scans. As with external beam radiotherapy, tolerance thresholds for normal tissues should be validated through rigorous epidemiological studies [21]. Given potential regional differences in radiosensitivity, epidemiological data should be collected locally. At present, no prospective studies on Lu-177-DOTATATE dosimetry have been conducted in Japan—an unexpected gap, considering that fluorine-18 fluorodeoxyglucose positron emission tomography has already been optimized for the Japanese population [22].

A previous dosimetry study [23] has reported considerable variability in kidney dose estimates across institutions, with standard deviations as high as 58%. This high interfacility variability in absorbed kidney doses underscores the need for standardized workflows to reduce uncertainty in TRT dosimetry. Such variability impairs the ability to accurately assess and mitigate the risk of nephrotoxicity—a critical concern in TRT—and undermines the statistical power of large-scale epidemiological studies.

Therefore, ensuring the quality of SPECT, image acquisition protocols, and simulation software is essential for future epidemiological studies. Furthermore, workflows should be optimized to minimize the burden on patients and health care providers, as the feasibility of large-scale epidemiological studies depends on the participation of multiple facilities and patients. This protocol may be readily incorporated into routine clinical practice. In this study, we will obtain clinical data to determine whether simplified time points can achieve acceptable dosimetric accuracy. Absorbed doses in tumor tissue and normal organs across multiple calculation programs will be assessed using descriptive statistical methods.

Given the current absence of prospective Lu-177-DOTATATE dosimetry trials in Japan, the data generated from this study will be invaluable. It will serve as foundational epidemiological evidence crucial for establishing organ tolerance doses

XSL•FO
RenderX

specifically for the Japanese population, thereby enhancing the safety and efficacy of TRT in the region. The findings are expected to offer new insights into Lu-177-DOTATATE dosimetry in neuroendocrine tumors and to support future epidemiological studies.

Regarding time-point selection, the maximum blood concentration of Lu-177-DOTATATE is observed within 1 hour. However, peak concentrations in normal organs and tumors occur thereafter. Within 4 hours post-administration, Lu-177-DOTATATE distributes to the kidneys, tumor lesions, liver, and spleen [24]. Although the early fast phase does not contribute substantially to the estimated renal absorbed dose, it can adversely affect the evaluation of the effective half-life. Therefore, it is generally accepted to acquire images after 4 hours and assume instantaneous uptake rather than to perform early imaging and risk misrepresenting the organ time activity–curve [25].

This study is exploratory and is intended to inform decision-making for future multi-institutional validation studies with larger cohorts. Thus, it has several intrinsic limitations. First, the findings may not be fully generalizable to other centers owing to variations in equipment and protocols, including software and imaging time points. The issue must be addressed through multi-institutional discussion prior to subsequent studies.

Second, the small sample size (estimated 5 - 10 patients) may not capture the full range of pharmacokinetic variability or tumor responses among patients. However, given that 4-time-point measurements impose a substantial burden on patients and health care providers, the limited sample size is unavoidable from a practical standpoint. The study design allows for a false-positive probability of 32% if the incidence (dose error per patient≤ threshold) is ≤20% within the workflow. These dose errors are considered to arise from software-specific curve-fitting models and dose-calculation frameworks. The incidence of such errors will be addressed through cross-validation among multiple software systems. Specifically, simplified time-point methods that demonstrate inconsistencies between software systems will be excluded from consideration in subsequent studies. Third, detailed 4-time-point SPECT imaging is performed in only one treatment cycle per patient, which does not account for potential changes in radiopharmaceutical distribution across subsequent cycles. Inter-cycle variability will be examined in subsequent studies after the optimal time points are determined in the present study. Finally, although multiple simulation software programs are evaluated, not all possible variations in image acquisition parameters or software algorithms are explored. These limitations highlight the need for future multi-institutional validation studies with larger cohorts and longitudinal studies to comprehensively standardize TRT dosimetry. Despite these limitations, this study is expected to provide crucial foundational data to guide the development of robust dosimetry protocols essential for advancing personalized TRT and facilitating large-scale epidemiological research.

Acknowledgments

We thank Editage for English language editing [26]. We extend our profound gratitude to Mr. Matsuyoshi Ogawa of the Department of Radiology, Yokohama City University Hospital, for his valuable technical support.

Funding

This work was supported by a Grant-in-Aid for Research on Radiation Oncology from the Japanese Society for Radiation Oncology (2024–2025) and by the Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant Numbers JP25K19125 and JP23K07197).

Data Availability

No public access plans exist for the protocol, dataset, or code. Results will be published in journals, with aggregated data reported. De-identified data may be shared with researchers upon reasonable request, subject to ethics review and a data-sharing agreement.

Data supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: TY, KM, NM, TM.
Funding acquisition: TY, TM.
Methodology: TM, KM, NM, TY.
Project administration: TM.
Resources: TM, KM, NM, SS, HH, NK, ST, TY.
Supervision: TM, SS, NK.
Writing – original draft: TY.
Writing – review & editing: TM, KM, NM, SS, HH, NK, ST, TY.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Sartor O, de Bono J, Chi KN, et al. Lutetium-177-PSMA-617 for metastatic castration-resistant prostate cancer. N Engl J Med 2021 Sep 16;385(12):1091-1103. [doi: 10.1056/NEJMoa2107322] [Medline: 34161051]
2. Strosberg J, El-Haddad G, Wolin E, et al. Phase 3 trial of 177Lu-Dotatate for midgut neuroendocrine tumors. N Engl J Med 2017;376(2):125-135. [doi: 10.1056/NEJMoa1607427] [Medline: 28076709]
3. Lapi SE, Scott PJH, Scott AM, et al. Recent advances and impending challenges for the radiopharmaceutical sciences in oncology. Lancet Oncol 2024 Jun;25(6):e236-e249. [doi: 10.1016/S1470-2045(24)00030-5] [Medline: 38821098]
4. Bela Andela S, Amthauer H, Furth C, et al. Quantitative PSMA-PET parameters in localized prostate cancer: prognostic and potential predictive value. Radiat Oncol 2024 Jul 29;19(1):97. [doi: 10.1186/s13014-024-02483-w] [Medline: 39080696]
5. Ljungberg M, Celler A, Konijnenberg MW, et al. MIRD pamphlet no. 26: joint EANM/MIRD guidelines for quantitative 177Lu SPECT applied for dosimetry of radiopharmaceutical therapy. J Nucl Med 2016 Jan;57(1):151-162. [doi: 10.2967/jnumed.115.159012] [Medline: 26471692]
6. Sandström M, Garske U, Granberg D, Sundin A, Lundqvist H. Individualized dosimetry in patients undergoing therapy with (177)Lu-DOTA-D-Phe (1)-Tyr (3)-octreotate. Eur J Nucl Med Mol Imaging 2010 Feb;37(2):212-225. [doi: 10.1007/s00259-009-1216-8] [Medline: 19727718]
7. Beauregard JM, Hofman MS, Pereira JM, Eu P, Hicks RJ. Quantitative (177)Lu SPECT (QSPECT) imaging using a commercially available SPECT/CT system. Cancer Imaging 2011 Jun 15;11(1):56-66. [doi: 10.1102/1470-7330.2011.0012] [Medline: 21684829]
8. Hu J, Seifert R, Karkampouna S, et al. Influence of dosimetry accuracy on the correlation with treatment outcome in a preliminary PSMA radiopharmaceutical therapy study. Eur J Nucl Med Mol Imaging 2025 Apr;52(5):1649-1657. [doi: 10.1007/s00259-024-07010-3] [Medline: 39704811]
9. Garske U, Sandström M, Johansson S, et al. Minor changes in effective half-life during fractionated 177Lu-octreotate therapy. Acta Oncol 2012 Jan;51(1):86-96. [doi: 10.3109/0284186X.2011.618511] [Medline: 21961497]
10. Willowson KP, Eslick E, Ryu H, Poon A, Bernard EJ, Bailey DL. Feasibility and accuracy of single time point imaging for renal dosimetry following 177Lu-DOTATATE ('Lutate') therapy. EJNMMI Phys 2018 Dec 20;5(1):33. [doi: 10.1186/s40658-018-0232-9] [Medline: 30569328]
11. Heikkonen J, Mäenpää H, Hippeläinen E, Reijonen V, Tenhunen M. Effect of calculation method on kidney dosimetry in 177Lu-octreotate treatment. Acta Oncol 2016;55(9-10):1069-1076. [doi: 10.1080/0284186X.2016.1182642] [Medline: 27219529]
12. Hänscheid H, Lapa C, Buck AK, Lassmann M, Werner RA. Dose mapping after endoradiotherapy with 177Lu-DOTATATE/DOTATOC by a single measurement after 4 days. J Nucl Med 2018 Jan;59(1):75-81. [doi: 10.2967/jnumed.117.193706] [Medline: 28588150]
13. The 2007 Recommendations of the International Commission on Radiological Protection. ICRP publication 103. Ann ICRP 2007;37(2-4):1-332. [doi: 10.1016/j.icrp.2007.10.003] [Medline: 18082557]
14. Hosono M, Ikebuchi H, Nakamura Y, et al. Manual on the proper use of lutetium-177-labeled somatostatin analogue (Lu-177-DOTA-TATE) injectable in radionuclide therapy (2nd ed.). Ann Nucl Med 2018 Apr;32(3):217-235. [doi: 10.1007/s12149-018-1230-7] [Medline: 29333565]
15. Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. J Natl Cancer Inst 2009 May 20;101(10):708-720. [doi: 10.1093/jnci/djp079] [Medline: 19436029]
16. Iwata H, Ishikura S, Murai T, et al. A phase I/II study on stereotactic body radiotherapy with real-time tumor tracking using CyberKnife based on the Monte Carlo algorithm for lung tumors. Int J Clin Oncol 2017 Aug;22(4):706-714. [doi: 10.1007/s10147-017-1123-0] [Medline: 28429140]
17. Hodapp N. Der ICRU-Report 83: Verordnung, Dokumentation und Kommunikation der fluenzmodulierten Photonenstrahlentherapie (IMRT) [Article in German]. Strahlenther Onkol 2012 Jan;188(1):97-100. [doi: 10.1007/s00066-011-0015-x]
18. Sgouros G, Bolch WE, Chiti A, et al. ICRU REPORT 96, dosimetry-guided radiopharmaceutical therapy. J ICRU 2021 Dec;21(1):1-212. [doi: 10.1177/14736691211060117]
19. Emami B, Lyman J, Brown A, et al. Tolerance of normal tissue to therapeutic irradiation. Int J Radiat Oncol Biol Phys 1991 May 15;21(1):109-122. [doi: 10.1016/0360-3016(91)90171-y] [Medline: 2032882]
20. Hope TA, Abbott A, Colucci K, et al. NANETS/SNMMI procedure standard for somatostatin receptor–based peptide receptor radionuclide therapy with 177 Lu-DOTATATE. J Nucl Med 2019 Jul;60(7):937-943. [doi: 10.2967/jnumed.118.230607] [Medline: 31263080]
21. McDonald AM, Schneider CS, Stahl JM, Oster RA, Popple RA, Mayo CS. A focused review of statistical practices for relating radiation dose-volume exposure and toxicity. Radiat Oncol 2023 Mar 24;18(1):57. [doi: 10.1186/s13014-023-02220-9] [Medline: 36964622]
22. Fukukita H, Suzuki K, Matsumoto K, et al. Japanese guideline for the oncology FDG-PET/CT data acquisition protocol: synopsis of version 2.0. Ann Nucl Med 2014 Aug;28(7):693-705. [doi: 10.1007/s12149-014-0849-2] [Medline: 24859759]

XSL•FO
RenderX

23. Uribe C, Peterson A, Van B, et al. An international study of factors affecting variability of dosimetry calculations, part 1: design and early results of the SNMMI dosimetry challenge. J Nucl Med 2021 Dec;62(Suppl 3):36S-47S. [doi: 10.2967/jnumed.121.262748] [Medline: 34857620]
24. Kobayashi N, Takano S, Ito K, et al. Safety and efficacy of peptide receptor radionuclide therapy with 177Lu-DOTA0-Tyr3-octreotate in combination with amino acid solution infusion in Japanese patients with somatostatin receptor-positive, progressive neuroendocrine tumors. Ann Nucl Med 2021 Dec;35(12):1332-1341. [doi: 10.1007/s12149-021-01674-9] [Medline: 34533700]
25. Sjögreen Gleisner K, Chouin N, Gabina PM, et al. EANM dosimetry committee recommendations for dosimetry of 177Lu-labelled somatostatin-receptor- and PSMA-targeting ligands. Eur J Nucl Med Mol Imaging 2022 May;49(6):1778-1809. [doi: 10.1007/s00259-022-05727-7] [Medline: 35284969]
26. Editage. URL: www.editage.jp [accessed 2026-02-02]

Abbreviations

CT: computed tomography
SPECT: single-photon emission computed tomography
TRT: targeted radiopharmaceutical therapy

Edited by J Sarvestan; submitted 29.Aug.2025; peer-reviewed by KY Ko; accepted 23.Jan.2026; published 19.Feb.2026.

Please cite as:
Yagihashi T, Miwa K, Miyaji N, Sugimoto S, Hayakawa H, Kobayashi N, Takano S, Murai T
Evaluation of Reduced Single-Photon Emission Computed Tomography Imaging Protocols and Software Variability in 177Lu-DOTATATE Dosimetry: Protocol for an Exploratory Observational Trial
JMIR Res Protoc 2026;15:e83248
URL: https://www.researchprotocols.org/2026/1/e83248
doi:10.2196/83248

© Takayuki Yagihashi, Kenta Miwa, Noriaki Miyaji, Satoru Sugimoto, Hideki Hayakawa, Noritoshi Kobayashi, Shoko Takano, Taro Murai. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Toward Best Practice Guidelines and Curricula for Virtual Care (Telehealth) in a Cancer Center: Protocol for a Multimethod Study

Sharon Brownie[1,2,3], RN, PhD; Lauren Parkinson-Zarb[4,5], RN, PhD; Andrew Dimech[5], RN, MN; Maria Ftanou[5,6], PhD; Patrick Broman[3,7], MA, PhD

[1]School of Rural Medicine, Charles Sturt University, Leeds Parade, Orange, NSW, Australia
[2]School of Health Sciences, Swinburne University of Technology, Johns Street, Melbourne, VIC, Australia
[3]Centre for Health and Social Practice, Waikato Institute of Technology, Hamilton, New Zealand
[4]Department of Nursing, School of Health Sciences, University of Melbourne, Melbourne, VIC, Australia
[5]Office of the Chief Nurse, Peter MacCallum Cancer Center, Melbourne, VIC, Australia
[6]School of Population and Global Health, University of Melbourne, Melbourne, VIC, Australia
[7]Australian Centre for Student Equity and Success, Curtin University, Perth, WA, Australia

Corresponding Author:
Sharon Brownie, RN, PhD
School of Rural Medicine, Charles Sturt University, Leeds Parade, Orange, NSW, Australia

Abstract

Background: Virtual health care, originally as telephone-based telehealth, has been used for more than 45 years; however, the literature shows limited understanding of the competencies required for safe virtual care practice in nursing and other health-related fields. This has led to a widening education-practice gap.

Objective: The aim of this study is to identify (1) clinical guidelines for nurses and other health professionals undertaking routine virtual health (telehealth) assessment, triage, and follow-up care; and (2) curricula for preparing health professionals for virtual care. Subsequently, data will be collected within a major cancer treatment service to codetermine core competencies and curricula for nurses engaged in telehealth clinics.

Methods: This was a phased multimethod study including reviews of existing literature, followed by qualitative (in-depth interviews, n=20) and quantitative (online survey, n=200) data collection and co-design workshops (n=5) to achieve project aims. Implementation will involve a pilot and an evaluation before full rollout of the developed guidelines and syllabus.

Results: Literature reviews completed in the initial phase of this project confirm a paucity of existing guidelines for virtual health assessment and an urgent need to develop telehealth or virtual care competency frameworks and curricula for health professionals in training or practice. We propose an approach to develop and test these materials in practice. A total of US $52.6 was provided by a philanthropic alumni for support over the full duration of the project. Recruitment and collection from human participants will commence on February 1, 2026, following procurement of ethics clearance. Phase 2 data collection and analysis will occur from February to December 2026. Results will be presented to the ethics committee and clearance gained to implement and evaluate the program in 2027.

Conclusions: Working with health providers, consumers, and academics toward standards of practice and curricula is clearly needed to ensure that the current and future nursing workforce is prepared for the continuing rise in virtual care. Completion of this project will fill an existing gap in the provision of guidelines and education for nurses providing virtual care.

International Registered Report Identifier (IRRID): PRR1-10.2196/81768

(JMIR Res Protoc 2026;15:e81768) doi:10.2196/81768

KEYWORDS

telehealth; telecare; oncology nursing; clinical guidelines; practice guidelines; virtual medicine; nursing education

Introduction

Background

Virtual models of care, including telephone and video-based delivery, have been used for more than 4 decades [1-4]; however, literature describes sluggish uptake in embedding newly emerging telehealth modalities in nursing education and practice [5-7].

Though well-established even before the pandemic [8,9], telehealth service delivery accelerated during COVID-19 in response to the urgent need for innovations in care provision, particularly for those unable to access traditional services under

XSL•FO
RenderX

stay-at-home orders or during periods of community lockdown [9,10]. The trend for greater use of virtual modalities is reflected globally in patient expectations, health service delivery, and clinical practice [11] as well as in the Australian and New Zealand context [12-14].

While various nursing-related telehealth guidelines or pilot initiatives were developed during the COVID-19 pandemic [15,16], and some work has progressed to better define telehealth education and practice standards [8,17,18], there is no international or national agreement as yet on the education of current and future nurses in telehealth. It is unclear which telehealth competencies are needed, and telehealth placements are not universally recognized as part of required clinical hours. The World Health Organization is promoting telehealth initiatives with a focus on improving access to quality care, equity, resources, and universal health coverage. Despite this, there remains no global standard clinical guidelines or competency descriptions for safe telehealth practice [19].

In most contexts, a significant disconnect exists between the reality of practice and professional education, with content related to safe practice in virtual care effectively absent from nursing accreditation and regulatory standards, nursing scopes of practice, and undergraduate nursing curricula [7,20], despite various nurse leaders, researchers, and policymakers having recommended provision of clinically focused telehealth-related education in nursing and health professional programs [17,21,22].

Definitions

Terminology related to virtual care can be confusing. Many terms are linked to the delivery of care outside of traditional face-to-face delivery. It is not uncommon for digital health and telehealth to be used interchangeably, so it is important that the definition is clear. Chen et al [23] propose the term “connected health” as an ecosystem or framework to help with definitions and better understand who and what is connected to whom within the rapidly emerging digital health landscape and multiplicity of terminology.

For the purposes of this project, the term telehealth is used to refer to care provided at a distance (Figure 1). Telehealth fits within a broader framework of digital eHealth, which includes broad-ranging tools and modalities such as mobile health apps, electronic referrals, wearable devices, electronic prescriptions, patient-reported symptom monitoring, and telehealth modalities electronic records [24,25]. Telehealth is a subset of digital health (Figure 1) and a modality by which virtual care or care at a distance is delivered. Telehealth encounters may be telephone-based or video-based. Nurses providing telehealth services can use a range of information sources, assessment techniques, and communication technologies to provide care virtually, that is, in situations where the nurse and the patient are not physically located in the same place [25,26]. This project focuses on telehealth (patient-related care at distance) rather than the broader concept of digital health or eHealth.

XSL•FO
RenderX

Figure 1. Telehealth within digital (e) health. Adapted from OECD Health Working Paper 129 Socha-Dietrich [25]. AI: artificial intelligence; ML: machine learning; OECD: organization for economic cooperation and development.

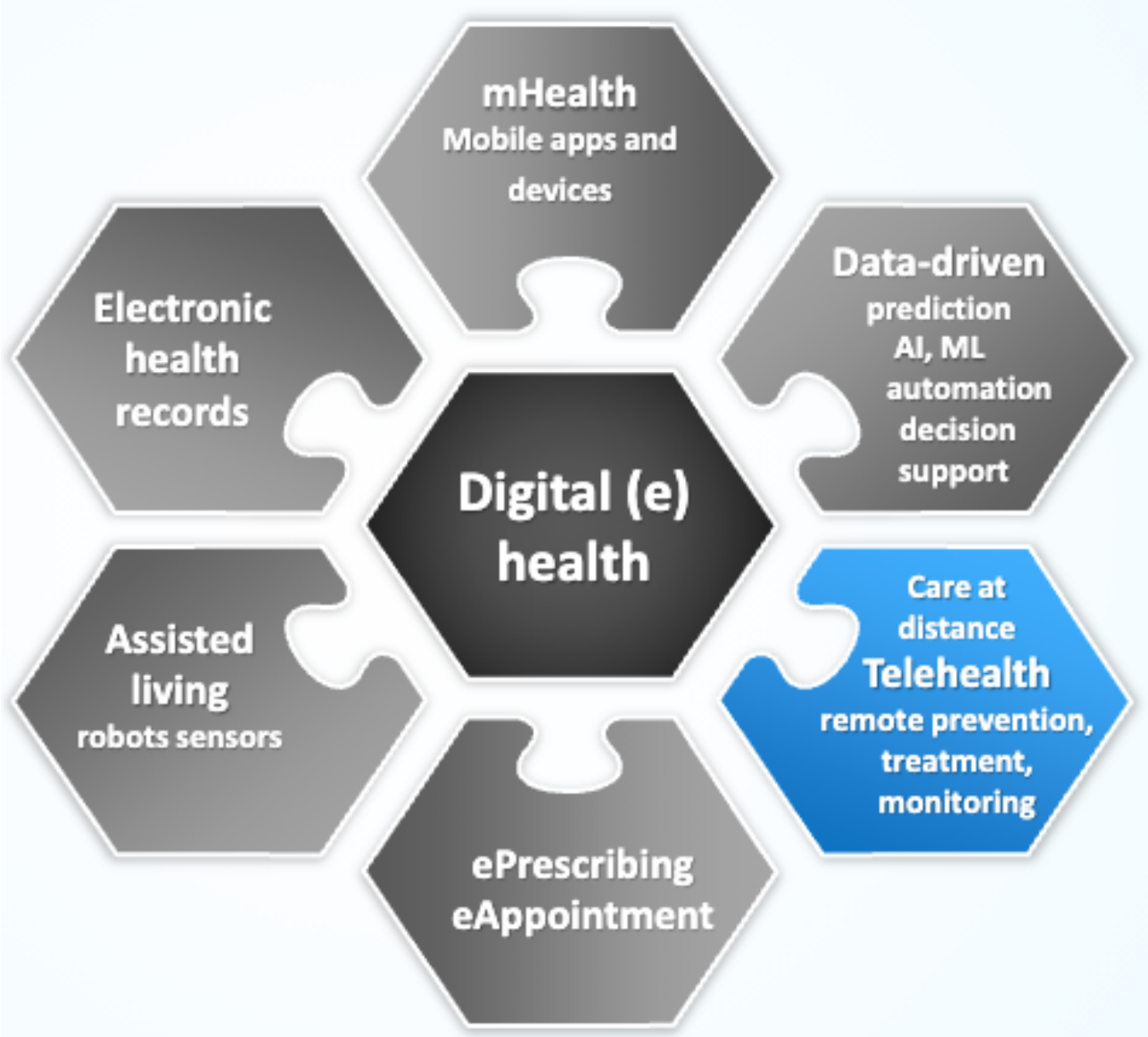

Study Context

The study is performed at the Peter MacCallum Cancer Centre, a large metropolitan cancer treatment center in an Australian city and the leading cancer center in Australia hosting the largest radiation oncology facility in the Southern Hemisphere [27]. It has recently been listed among the top 20 global leaders in cancer care [28]. At Peter MacCallum Cancer Centre, nurse-led telehealth clinics are currently offered across several tumor streams, including gastrointestinal (upper and lower), lung, head and neck, hematology, gynecology, and genitourinary. The total number of nurse-led telehealth consults has grown steadily during 2022 - 2025, with 22,627 consultations provided in the 2024 calendar year (Figure 2).

XSL•FO
RenderX

Figure 2. Remote/telehealth consultations at Peter MacCallum Cancer Centre during July 2022 to June 2025. Consultations by month (time series) and total consultation by tumor stream.

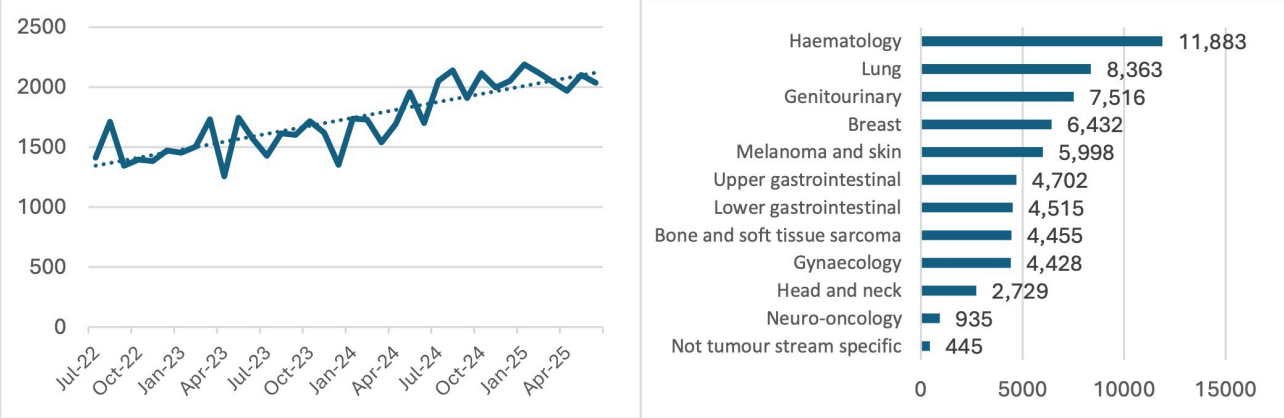

Delivery of telehealth services has increased from 1450 per month in the second half of the year 2022 to 2100 per month in the first half of the year 2025 (Figure 2). While it is the most experienced nurses providing care via this modality, there are currently no formalized guidelines and orientation processes for nurses using telehealth to assess patients and deliver care. Nurses are currently assigned to telehealth services based on their experience and expertise but are not provided with formal orientation or guidelines related to telehealth due to the lack of extant material in the literature or available practice guidelines. This raises several potential quality and safety-related issues, which this project is designed to address.

Project Rationale and Aims

The project rationale is informed by the increasing use of telehealth services and the realization that the use of these technologies is a core competency within the nursing and broader health workforce [9,11,21]. Telehealth involves the use of information, communication, and assessment techniques to provide health care when clinicians and clients are not physically present in the same location [26]. In a virtual assessment context, the clinician is unable to use all senses: “touch” and “smell” are not available. Patient assessment via digital modalities requires different skill sets than in-person assessments; thus, there is a need for guidelines to ensure clinical safety and quality in telehealth delivery. This study protocol was designed to address current gaps in telehealth delivery and formalize orientation and operating guidelines for nurses assigned across the services.

The overarching aim is to develop nursing practice guidelines and syllabi for virtual health assessment and care within a major cancer center. The aims are both clinically practice-oriented and education-focused, as shown in Figure 3.

XSL•FO
RenderX

Figure 3. Telehealth project.

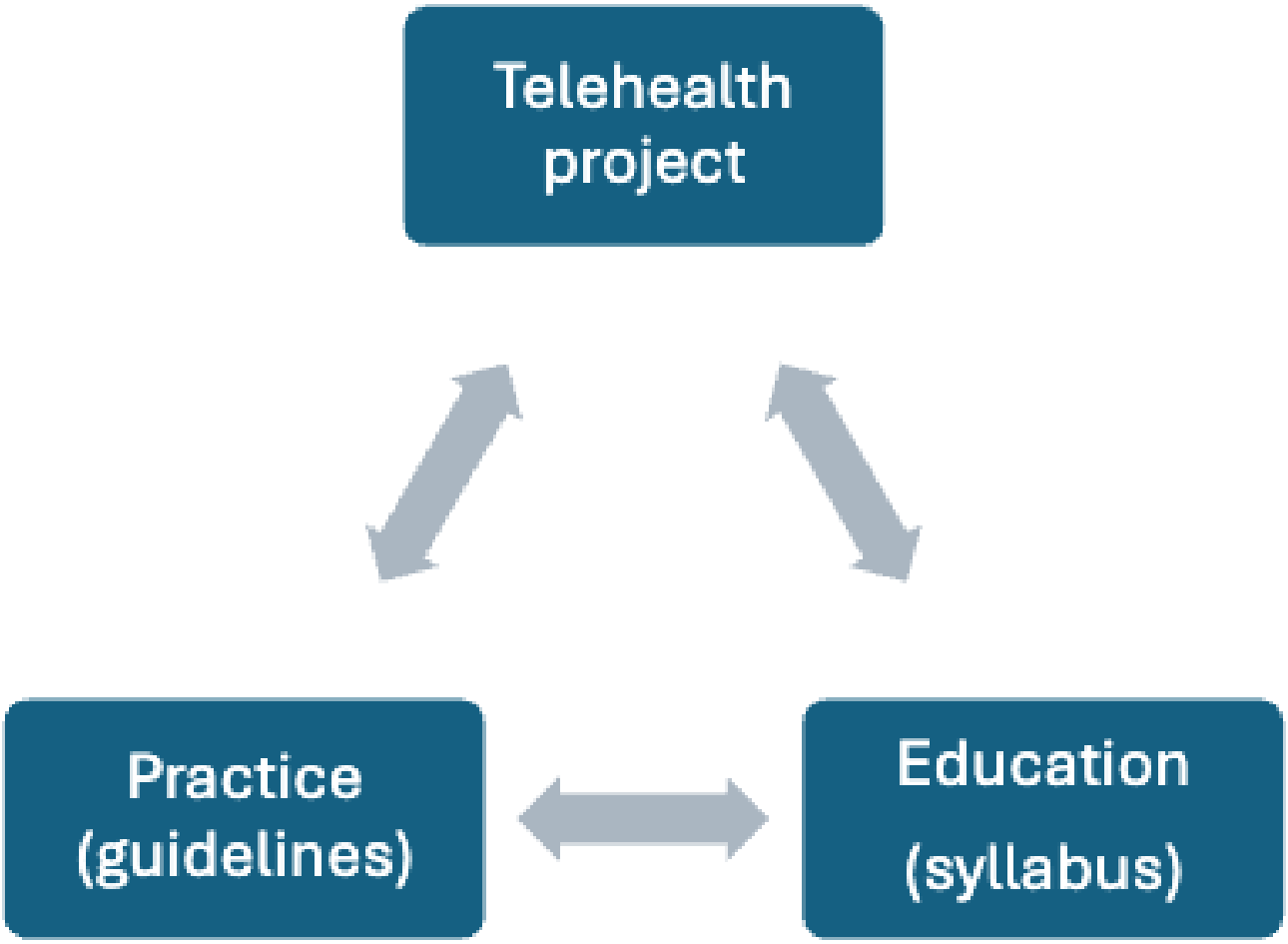

Practice-related aims include the following: (1) identify current practices and practice guidelines in virtual health assessment; (2) identify required nursing competencies for safe practice in telehealth assessment and nursing practice within a cancer center; (3) develop telehealth practice guidelines for senior registered nurses entering telehealth practice in oncology settings, inclusive of triage, health assessment, and routine follow-up; and (4) ensure that the developed guidelines provide the basis for consistent standards and safe practice for telehealth nursing services

Education-related aims include the following: (1) locate existing information about the competencies needed to undertake virtual health assessment and the educational curricula needed to support the development of competencies; (2) detail a syllabus and curriculum-related requirements for the development of the identified oncology telehealth competencies; (3) develop learning objectives, course content, learning outcomes, and assessment tools; (4) design program evaluation guidelines; and (5) pilot and evaluate the developed syllabi and curriculum content.

Methods

Study Method

The study method uses a phased multimethod approach, commencing with a review of existing literature followed by qualitative in-depth interviews, online surveys with quantitative methodologies, and co-design workshops. Implementation will involve a fully evaluated pilot before full rollout of the developed guidelines and syllabus. Generative AI was not used in the generation of the protocol or this manuscript.

A multimethod approach has been chosen to provide a more comprehensive and robust picture in a context of complexity. Multimethod research uses distinct methods within a single study to strengthen validity by collecting various types of data, which can result in deeper and richer insights and is reported to more effectively address complex issues and enhance applicability of findings [29-31]. The study design involves 3 distinct phases as shown in Figure 4, commencing with data-gathering activities (Phase 1) followed by codevelopment processes (Phase 2). A third phase focuses on program implementation, commencing with pilot testing of the developed virtual health practice guidelines, syllabi, and curricula content. Findings from each phase inform the ongoing development of the project.

The sequential/stepped gathering of data is the inputs for the development of the telehealth practice guidelines, teaching syllabi, and curriculum content for piloting in practice.

XSL•FO
RenderX

Figure 4. Phased project design. REDCap: Research Electronic Data Capture.

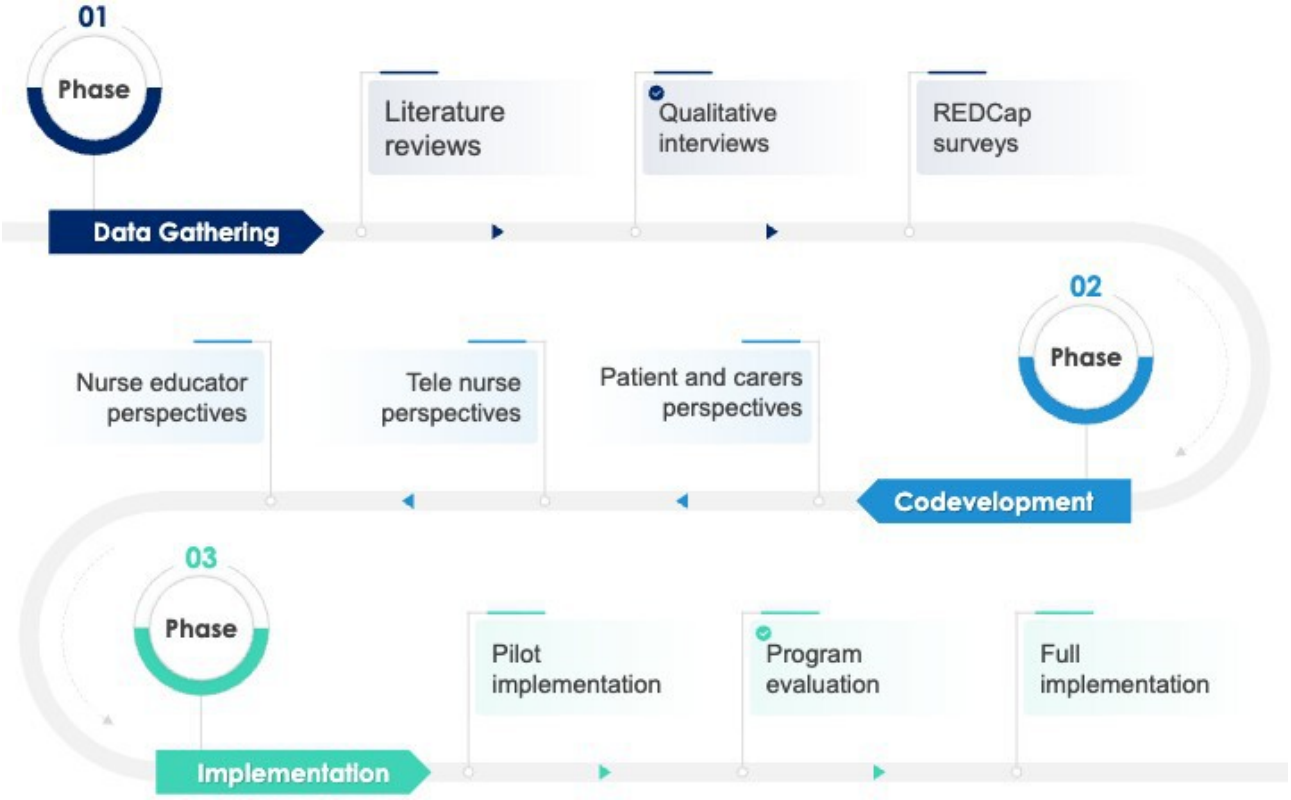

Quality Guidelines

Quality guidelines were integrated throughout the research process to ensure rigor and high standards in the research endeavor, while also enhancing transparency and reproducibility. Subsequently, each phase of the project is designed in accordance with relevant guidelines for quality reporting and conduct of research, including the Joanna Briggs Institute (JBI) guidelines for systematic review and research syntheses [32], PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews): checklist and explanation [33], PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) reporting guidelines [34,35], JBI criteria for qualitative research [36], and the CASP (Critical Appraisal Skills Program) checklist for quality cross-sectional studies as published by the UK-based CASP [37].

Q-Sort methodology will be used to provide structure, transparency, and credibility to the co-design workshops [38-41]. The program logic model (PLM) will be adopted in the framework for the overall implementation and evaluation of the initiative [42,43], with development and evaluation of the educational products supported by the Kirkpatrick conceptual model that provides a useful tool to measure learner reaction, level of learning, behavior, and results [44].

Ethical Considerations

Study Approval

The literature reviews were scheduled to provide insights to inform study design and to protect against duplication of existing knowledge. Reviews were able to proceed without formal ethics approval, following which a formal project design was confirmed, and ethical approval was sought and secured as follows. Project No.: 116964, Ethics No.: 25/60L, HREC Reference: HREC/116964/PMCC.

Informed Consent

To ensure that participants do not feel obligated to participate, the study team will inform invitees that participation is voluntary and that they may stop their involvement in the project at any time. The team will explain that their decision will not impact their care or relationships with hospital staff.

Data Storage and Privacy Issues

Data will be securely stored and only the research team can access it per the National Statement on Ethical Conduct in Human Research [45] and the Australian Code for Responsible Conduct of Research [46]. Participants' identities will stay anonymous, and all data will be destroyed 5 years after publication.

No harm is expected. If a participant is at risk, safety procedures will be followed, and a trusted person or clinician may be informed.

Consumer Compensation

As per the policy guidelines of the health entity, all participating consumer and community representatives will be compensated for their time. For this purpose, $10,000 AUD (US $7007.8) has been ring-fenced in the budget.

Distressed Participants

Distress is not expected during the research. However, if participants become upset, trained project team members will contact them to offer support. Support options may include

XSL•FO
RenderX

Cancer Council Victoria counseling or the Peter Mac Employee Assistance Program.

Phase 1: Data Gathering

Literature Reviews

As previously described, the initial phase of the research involved undertaking 2 literature reviews: (1) a practice-oriented scoping review to identify current practices and practice guidelines in virtual health assessment [47]; (2) an education-focused scoping review to locate existing information about the competencies needed to undertake virtual health assessment and the educational curricula needed to support the development of competencies (Brownie et al, unpublished data, 2025).

The literature searches informing this study adhered to JBI guidelines for systematic review and research syntheses [32] and the PRISMA-ScR: checklist and explanation [33]. Search results were reported in accordance with the PRISMA framework reporting guidelines [34,35]. The rigor and quality of the work were subsequently assessed and confirmed through peer review and publication in Q1-ranked journals (Brownie et al, unpublished data, 2025) [47].

Qualitative Interviews

In this second phase, 20 participants with lived or professional experience—such as consumers, carers, oncology nurse consultants, educators, academics, and instructional designers—will be interviewed about their experience with telehealth. Insights from these interviews will help shape questions for later online surveys.

During the interviews, participants will be asked about their telehealth experiences, the competencies required for telehealth delivery, ways to improve it, and recommendations for better design, teaching, and service delivery. Each interview will last up to 50 minutes and will be audio-recorded and transcribed (Textbox 1).

Textbox 1. Participant inclusion and exclusion criteria: qualitative interviews.

Inclusion criteria:

- Persons with cancer who have completed treatment more than 3 months ago and who have engaged with Peter Mac telehealth services
- Carers of people with cancer
- Telenursing staff and nurse educators working in oncology, with at least 3 years of nursing experience
- Able to engage with researchers in English
- Aged 18 years or older

Exclusion criteria:

- Aged less than 18 years
- Nurses with less than 3 years' experience
- Nurses with less than 3 years of nursing experience
- Non-English speakers

Qualitative interviews will follow the JBI criteria for qualitative research to ensure rigor and quality in this phase of the research process. The use of JBI guidelines helps ensure robust, credible, and practical approaches to qualitative data collection, while supporting a structured, transparent, and rigorous approach to evidence synthesis [36].

Participants for qualitative interviews will be recruited in 2 ways: (1) through health service consumer and carer groups or newsletters; (2) through health service staff who identify potential participants, such as patients who have completed treatment or carers, including those who traveled from regional areas.

Interested participants will receive an email invitation and can respond by phone or email.

All nursing staff involved in outpatient telehealth clinics will be invited to join the interview via email from the chief investigator. Nursing educators and academic faculty will be invited through the Nursing Department Chair.

After giving consent, participants will be contacted to schedule a time for the interview and will receive a REDCap (Research Electronic Data Capture) survey link to provide basic demographic information. This survey, taking about 5 minutes, is only for describing the participant group (eg, in publications). Textbox 2 lists the information that will be collected.

An experienced qualitative researcher will facilitate the interviews. Interviews will be audio-recorded and transcribed for later reference. Only project managers will have access to recordings, in accordance with the National Statement on Ethical Conduct in Human Research 2023 and the Australian Code for Responsible Conduct of Research 2018 [46].

A coding guide will be developed, and transcripts will be coded into patterns, themes, and key insights. Interviews will be coded by 2 or more coders to ensure consistency of coding and interrater reliability.

XSL•FO
RenderX

Textbox 2. Demographic data.

Consumers:

- Gender
- Age
- Country of birth
- Postcode
- Diagnosis
- Date of diagnosis
- Treatments received

Nursing staff, academic faculty, and/or instructional designer:

- Gender
- Age
- Ethnicity
- Years of practice
- Practice area/specialty

Patient REDCap Survey

For the third study phase, the project team aims to survey 200 consumers to explore their experiences with receiving telehealth services with survey questions designed to gain further insight into the themes identified by the preceding qualitative interviews.

The survey instrument will be designed with reference to the CASP checklist for quality cross-sectional studies as published by the UK-based CASP [37]. Use of the checklist is an aid to develop and systematically assess the trustworthiness, quality, and relevance of the survey instrument. The checklist provides a structured framework to identify and avoid methodological limitations and potential bias [48].

The target sample size of 200 respondents was determined to ensure sufficient statistical power to detect meaningful effects in user perceptions of telehealth services. Power calculations were based on conventional thresholds for social and behavioral research [49], assuming a medium effect size (Cohen d=0.5), a 2-tailed α level of .05, and statistical power (1–β) of 0.80. Under these parameters, a minimum of 128 participants is required to detect significant group differences or associations in primary analyses. To account for potential nonresponse, incomplete data, and subgroup analyses, the sample was increased by approximately 50%, yielding a target of 200 completed responses. This sample size affords adequate precision for estimating proportions within a ±7% margin of error at the 95% confidence level (n=196), which is acceptable for consumer research in health services contexts [29].

No comprehensive or validated instruments suitable for use in the survey were located during extensive searches of the literature (Brownie et al, unpublished data, 2025) [47]. Instead, survey items for this domain will be developed by the research team in consultation with telehealth practitioners at the study site. These items relate to existing frameworks and prior literature and aim to capture relevant dimensions of conceptual validity with telehealth practice. To maximize face validity, the survey will undergo pretesting with a small group of users (n=5) to assess clarity, relevance, and comprehensibility of items. Feedback will be used to refine wording, response options, and survey flow.

As per the agreed ethics clearance, the draft survey will be forwarded to the responsible committee for review. The survey will be designed to take no more than 20 minutes to complete, and time will be factored into the survey design so as not to burden consumers.

To maximize face validity, the survey will undergo pretesting with a small group of users (n=5) to assess clarity, relevance, and comprehensibility of items. Feedback will be used to refine wording, response options, and survey flow (Textbox 3).

Over a 15-week period, patients in various telehealth clinics will be asked if they want to take part in a survey about their experience. Interested participants will get a consent form and survey link through REDCap. Participation is voluntary, and participants will be paid according to health service guidelines. Survey results will be analyzed to help plan the next phase of the project.

XSL•FO
RenderX

Textbox 3. Participant inclusion profile Research Electronic Data Capture survey.

Survey participant inclusion profile:

- Persons with cancer who have completed treatment more than 3 months ago
- Persons who have received telehealth sessions from a nurse
- Able to engage with researchers in English
- Aged 18 years or older

Codevelopment Workshops

Following the survey, a series of co-design workshops (Figure 5) will be undertaken, adopting a codevelopment process based on Q-Sort methodology [38,41,50]. These workshops aim to reach agreement on telehealth guidelines, clinical competency requirements, and curricula. Consensus will be built through a series of workshops, using information from Phase 1, including the literature reviews (Brownie et al, unpublished data, 2025) [47], qualitative interviews, and survey findings to inform the development of statements that participants will sort during workshop discussions.

Figure 5. Key project activities.

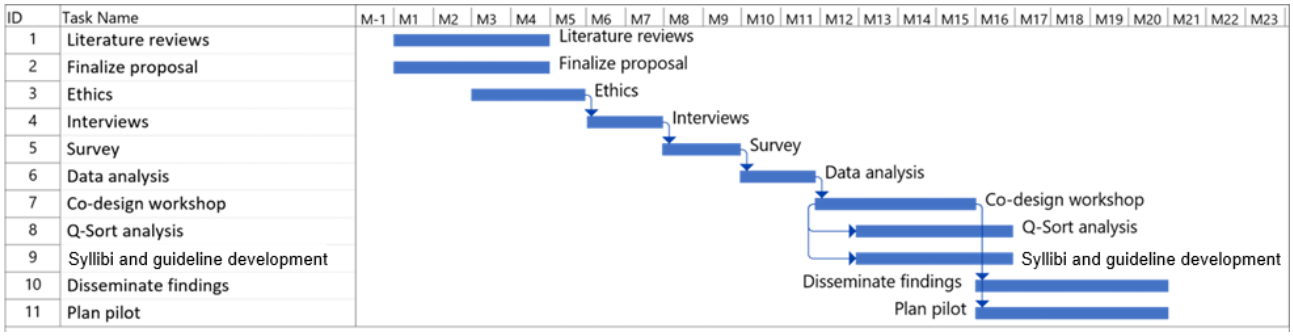

Q-Methodology (Q-Sort) supports systematic information gathering to analyze potentially diverse participant viewpoints in each workshop. Q-Sort is helpful in understanding the perspectives of participants who are likely to have different viewpoints about a topic [38,41]. Q-Sort methodology has been chosen because it will help identify the diversity and commonality of perspectives held by educators, Aboriginal and Torres Strait Islander people, nurses, other health professionals, patients, and families. Q-Sort increases the transparency and replicability of the research. Five basic steps are involved in the Q-Sort methodology [38,41]:

1. Define the domain of focus for discussion
2. Develop a set of statements that participants will sort
3. Select participants representing different groups
4. Q-Sort by participants
5. Analysis and interpretation of information from participants

The workshops will allow participants to view deidentified perspectives which have emerged from Phase 1 findings. Two project team members will facilitate each in-person workshop, which will last about 2 hours (total 10 h). Workshops will be partly guided by participants and will include the following components:

- Presenting the project objectives
- Explanation of the Q-Sort methodology
- Sort and prioritize statements related to the following:
  - Competency requirements for virtual health assessment
  - Requirements for safe practice
  - Curricula content for teaching virtual health assessment
  - Competencies development and upskilling methods
  - Assessment process
  - Delivery and format of curricula
- Opportunity to review materials as they are developed between workshops
- Opportunity for participants to contribute any additional relevant opinions or information

Workshops will be held in person, with options to join by teleconference or through a one-on-one phone discussion if required. Eligibility for participation follows the same criteria as the earlier qualitative interviews (Textbox 1). We plan to recruit 10 consumers (patients and carers) and 10 professionals (nurses, educators, and instructional designers).

Beliefs, views, assumptions, and matters of social or personal importance are inherently subjective [40]. It is anticipated that a range of differing perspectives and viewpoints will exist within Phase 1 findings with potential divergency of views at the workshops. Q-Sort has been chosen as it allows for subjectivity and emphasizes consensus. The approach follows a well-defined and formally structured process [39] without the need for participants to engage in written or spoken activities. Rather, participants rank written statements, objects, pictures, or video clips. Quieter, vulnerable, and potentially marginalized participants can all participate equally in an ethical, nonthreatening, and respectful manner [38,40,51].

Findings from the earlier phases will be consolidated into formal syllabi and practice guidelines, and additional ethics approval will be sought to proceed with the pilot phase. While Q-Sort will provide clear guidance on the priorities for inclusion in both the practice guidelines and curriculum, a conceptual model

XSL•FO
RenderX

is needed to take these through to program development. The Kirkpatrick model will be used to analyze and synthesize sorted priorities, agree with the program objectives and course materials, confirm content relevance, and identify the knowledge required of the educational facilitators who will deliver the program [44].

Phase 1 and Phase 2 Project Step Timelines

Key activities involved in the first two phases of the project will be undertaken across a 20-month period as shown in Figure 5.

Implementation and Evaluation

The final phase of the project involves implementation activities, including the program pilot, evaluation, and subsequent scale-up and full rollout of the developed resources. The pilot will test these materials on a limited scale within the partner university and cancer center.

Evaluation of the overall project phase will be assessed within the overarching PLM—a framework recommended by the Victorian Department of Health that oversees health service delivery at the study setting [43]. The PLM supports the development of a roadmap of inputs, activities, outputs, and outcomes that will guide the entire evaluation process. As per their guidelines for evaluation [42], the study will evaluate project outcomes from the perspective of those impacted by the intervention, specifically consumers, nurses, and educators.

Project outcomes link directly to the research aim. Specifically, project outcomes will result in the development of nursing practice guidelines and syllabi for virtual health assessment and care within a major cancer center. Outcomes will be both clinically oriented and education-focused, as shown in Figure 3.

Evaluation of the developed education and practice products will be supported by the Kirkpatrick conceptual model which provides a useful tool to measure learner reaction, learning, behavior, and results [44]. Feedback from this evaluation will inform subsequent revision of the developed products.

Results

The literature review conducted in the first stage of the study (Brownie et al, unpublished data, 2025) [47] revealed a lack of comprehensive guidelines for virtual health assessment and competency development, highlighting an urgent need to design telehealth curricula that enhance clinical competency.

These early findings support the project rationale and confirm the relevance of its aims. The reviews also provide assurance that the project will extend rather than duplicate existing knowledge. Findings from the reviews have undergone independent peer review and open-access publication, adding rigor to the research process, strengthening the project rationale, and establishing the foundation for the overall project design.

Data from interviews and the consumer survey will guide discussions in co-design workshops. The workshops will produce consensus statements that prioritize content for telehealth guidelines and curricula. These statements will help define telehealth competencies and develop guidelines and curricula, which will be piloted and evaluated before broader use.

A total of US $52.6 was provided by a philanthropic alumni for support over the full duration of the project. Recruitment and collection from human participants commences February 1, 2026, following procurement of ethics clearance. Phase 2 data collection and analysis will occur from February to December 2026. Results will be presented to the ethics committee and clearance gained to implement and evaluate the program in 2027.

Discussion

Anticipated Findings

In 2019, the Australian Federal Government Department of Health and Aged Care commissioned an independent report focused on the educational requirement for the future nursing workforce [52]. This “Educating the Nurse of the Future” report emphasized the way information technology, digital innovation, and telehealth modalities have permeated every aspect of health care development and delivery [52]. It recommended that it would be important to include specific advice within revised registered nurse accreditation standards so that health education providers would be clear about what aspects of digital health technologies and informatics should be included in curricula and the specific skill level and competencies nurses should be expected to achieve [52].

Since the COVID-19 pandemic [8,9], telehealth service delivery has only accelerated in response to the urgent need for innovations in care provision, particularly those unable to access traditional service while under stay-at-home orders and during periods of enforced community lockdown [9,10]. The trend for greater use of digital modalities is reflected globally in patient expectations, health service delivery, and clinical practice [11] across Australia and New Zealand [12-14].

Despite this, many nurses in telehealth practice have had no formal training [53]; nursing academics are reported as having failed in the job of preparing nursing graduates for roles in telehealth [21]; experience and content across the various levels of nursing education programs appears low [5]; and graduates are entering the health workforce without telehealth education and the required virtual practice competencies [20]. An Australian national study funded by the Victorian Higher Education State Investment Fund reported 69.9% of undergraduate nursing (entry to practice) programs as having no digital health content. Of further concern, findings report a significant gap in the confidence and knowledge of nurse academics to teach digital- and telehealth-related content [54].

Regulatory gaps are also evident. While Australia’s recently revised Registered Nurse accreditation standards [55] record feedback detailing industry expectations regarding the inclusion of a “focus on digital health.” Subsequently, the updated standards support inclusion and embedding of 2 subjects within curricula—health informatics and digital health—and that the National Informatics Standards [56] be accessed as the source from which health education providers should elicit

XSL•FO
RenderX

competencies and learning outcomes [55]. Developed by the Australian Nursing and Midwifery Federation (2015), the National Information Standards for Nurses and Midwives focus on 3 key domains: computer literacy, information literacy, and information management [56]. Specific clinical process and competency requirements related to telehealth-based, "virtual clinical health assessment" and service delivery are not included.

Other existing guidelines similarly focus primarily on the technological aspects of health service delivery such as computer literacy, information literacy, and information management [56] and/or digital professionalism, leadership and advocacy, data and information quality, technology, and information-enabled care [13]. The Australian Health Practitioners Regulatory Authority has published telehealth guidance for practitioners, primarily providing a definition of telehealth services and who can use telehealth [57]. In contrast, this project is clinically focused and aims to fill current gaps related to the actual requirements for safe nursing practice in virtual health.

Undoubtedly, various Australia-based schools of nursing will be updating curricula in response to the revised 2019 accreditation standards, but in the absence of standardized Australian Health Practitioners Regulatory Authority or the Australian Nursing and Midwifery Accreditation Council guidelines, developments are likely to be variable. The importance of the issue is emphasized in the Australian Government's 2021 response to the Educating the Nurse of the Future report [12]. Recommendation 20 of 26 states:

> *ANMAC's RN accreditation standards for health informatics and digital health technologies should specify learning outcomes and the level of expertise required. The EN accreditation standards should contain similar specifications. [12]* [p 16]

However, many nurses working in telehealth have received little to no formal training [53]. Nursing educators have been criticized for not adequately preparing graduates for telehealth roles [21], and the amount of telehealth content and practical experience included across nursing education programs remains limited [5]. Subsequently, most graduates are entering the workforce without the essential competencies needed for telehealth practice [20].

The development imperative is compelling. This project is designed to work with clinicians, educators, and consumers in the development of telehealth clinical practice guidelines and curricula to address the gap in virtual care delivery.

Strengths and Limitations

This single-site study limits generalizability, and the authors acknowledge potential bias in co-design workshops wherein some participants may be more vocal than others and/or the workshop participant group may be unintentionally skewed in comparison to the overall population focus of the study. However, the mixed methods approach, with active input from consumers and clinicians, and the use of Q-Sort methods will serve to minimize these limitations.

Conclusions

Working with health providers, consumers, and academics in modernizing standards of practice and curricula and upskilling nurse educators is clearly needed to ensure that the current and future nursing workforce is prepared for the continuing rise in telehealth services. Completion of this industry/education co-design project will fill the existing gap in the provision of clinically oriented guidelines and education for nurses providing telehealth services with care at distance.

Acknowledgments

The authors gratefully acknowledge the expert input of Professor Marie Gerdtz, dean of nursing and midwifery at La Trobe University (previously the University of Melbourne) into the design of this study. Acknowledgments and thanks are also offered to Yvonne Panek-Hudson, nurse practitioner, and Tenille Lewin, and nurse-led research team member, at Peter MacCallum Cancer Center (Peter Mac) for their review of the project ethics application.

Generative AI was not used in the generation of this manuscript.

Funding

This study is supported by an anonymous Swinburne University of Technology alumni donor. The funder was not involved in the production of this work.

Data Availability

Two literature reviews that constituted the first step in this project and informed research design are published in open access journals. Deidentified data associated with the study are available on reasonable request subject to ethics approval conditions.

Authors' Contributions

All authors contributed equally to the idea conception, design, and development of this study. SB, LP-Z, and PB completed the literature review through to publication. MF, SB, and AD prepared documents for ethics and responded to queries through to approval. SB wrote the initial draft of this manuscript. All authors reviewed and edited the manuscript and approved the final version for submission.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Curtis P, Talbot A. The telephone in primary care. J Community Health 1981;6(3):194-203. [doi: 10.1007/BF01323010] [Medline: 7263941]
2. Crouch R, Woodfield H, Dale J, Patel A. Telephone assessment and advice: a training programme. Nurs Stand 1997 Aug 13;11(47):41-44. [doi: 10.7748/ns.11.47.41.s46] [Medline: 9355534]
3. Naurright LP, Moneyham L, Williamson J. Telephone triage and consultation: an emerging role for nurses. Nurs Outlook 1999;47(5):219-226. [doi: 10.1016/s0029-6554(99)90054-4] [Medline: 10546268]
4. Kamei T. Information and communication technology for home care in the future. Jpn J Nurs Sci 2013 Dec;10(2):154-161. [doi: 10.1111/jjns.12039] [Medline: 24373438]
5. Eckhoff DO, Guido-Sanz F, Anderson M. Telehealth across nursing education: findings from a national study. J Prof Nurs 2022;42:308-314. [doi: 10.1016/j.profnurs.2022.07.013] [Medline: 36150875]
6. Goodwin S. Telephone nursing: an emerging practice area. Nurs Leadersh (Tor Ont) 2007;20(4):37-45. [Medline: 18303723]
7. Butler Carroll T, Philips B. Texas Tech University Health Sciences Center frontiers in telemedicine simulation lab and certificate course: facility and curriculum development. : International Institute of Informatics and Systemics, IIIS; 2018 Presented at: Proceedings of The 12th International Multi-Conference on Society, Cybernetics and Informatics (IMSCI 2018); Jul 8-11, 2018; Orlando, Florida, USA URL: https://www.iiis.org/CDs2018/CD2018Summer/papers/HA787IM.pdf [accessed 2026-01-27]
8. van Houwelingen CTM, Moerman AH, Ettema RGA, Kort HSM, Ten Cate O. Competencies required for nursing telehealth activities: a Delphi-study. Nurse Educ Today 2016 Apr;39:50-62. [doi: 10.1016/j.nedt.2015.12.025] [Medline: 27006033]
9. Shaver J. The state of telehealth before and after the COVID-19 pandemic. Prim Care 2022 Dec;49(4):517-530. [doi: 10.1016/j.pop.2022.04.002] [Medline: 36357058]
10. Meyer MA. COVID-19 pandemic accelerates need to improve online patient engagement practices to enhance patient experience. J Patient Exp 2020 Oct;7(5):657-664. [doi: 10.1177/2374373520959486] [Medline: 33294595]
11. Jillian R, Yarbrough W. Telehealth, provider competencies in theory and application. J Health Hum Serv Adm 2023 Jun;46(2):106-132. [doi: 10.37808/jhhsa.46.2.2]
12. Australian government response to the independent review of nursing education: educating the nurse of the future. : Department of Health, Australian Government; 2021 URL: https://www.health.gov.au/sites/default/files/documents/2021/09/government-response-to-the-educating-the-nurse-of-the-future-report.pdf [accessed 2026-01-27]
13. National Nursing and Midwifery Digital Health Capability Framework. : Australian Government, Australian Digital Health Agency; 2020 URL: https://www.digitalhealth.gov.au/sites/default/files/2020-11/National_Nursing_and_Midwifery_Digital_Health_Capability_Framework_publication.pdf [accessed 2026-01-27]
14. Aussies accelerated adoption of virtual care to continue post-pandemic, while government top trusted source of information - PwC Global Health Industry Issues Survey. PwC Australia. 2021. URL: https://www.pwc.com.au/media/2021/pwc-global-health-industry-issues-survey.html [accessed 2026-01-27]
15. Brownie SM, Chalmers LM, Broman P, Andersen P. Evaluating an undergraduate nursing student telehealth placement for community-dwelling frail older people during the COVID-19 pandemic. J Clin Nurs 2023 Jan;32(1-2):147-162. [doi: 10.1111/jocn.16208] [Medline: 35018676]
16. Baxter KE, Kochar S, Williams C, Blackman C, Himmelvo J. Development of a palliative telehealth pilot to meet the needs of the nursing home population. J Hosp Palliat Nurs 2021 Oct 1;23(5):478-483. [doi: 10.1097/NJH.0000000000000784] [Medline: 34225341]
17. Rutledge CM, O'Rourke J, Mason AM, et al. Telehealth competencies for nursing education and practice: the four P's of telehealth. Nurse Educ 2021;46(5):300-305. [doi: 10.1097/NNE.0000000000000988] [Medline: 33481494]
18. Poitras ME, Couturier Y, Beaupré P, et al. Collaborative practice competencies needed for telehealth delivery by health and social care professionals: a scoping review. J Interprof Care 2024;38(2):331-345. [doi: 10.1080/13561820.2023.2213712] [Medline: 37226329]
19. Digit health. World Health Organization. 2025. URL: https://www.who.int/health-topics/digital-health#tab=tab_1 [accessed 2026-01-27]
20. The impact of nurse educators on the evolution of telehealth services. UTA - University of Texas. 2024. URL: https://academicpartnerships.uta.edu/healthcare-nursing-online-programs/msn/education/nurse-educators-on-telehealth-services/ [accessed 2026-01-27]
21. Rutledge CM, Gustin T. Preparing nurses for roles in telehealth: now is the time!. Online J Issues Nurs 2021;26(1). [doi: 10.3912/OJIN.Vol26No01Man03]
22. Búřilová P, Pokorná A, Búřil J, et al. Identification of telehealth nursing approaches in the light of the COVID-19 pandemic: a literature review. J Nurs Manag 2022 Nov;30(8):3996-4004. [doi: 10.1111/jonm.13864] [Medline: 36208135]

XSL•FO
RenderX

23. Chen SCI, Liu C, Hu R, Mo Y, Ye X. “Nomen omen”: exploring connected healthcare through the perspective of name omen. Healthcare (Basel) 2020 Mar 23;8(1):66. [doi: 10.3390/healthcare8010066] [Medline: 32210024]
24. Digital health. Australian Institute of Health and Welfare. 2024. URL: https://www.aihw.gov.au/reports/australias-health/digital-health [accessed 2026-01-27]
25. Socha-Dietrich K. Empowering the health workforce to make the most of the digital revolution, in health working papers no 129. : Organisation for Economic Co-operation and Development (OECD); 2021 URL: https://www.oecd.org/content/dam/oecd/en/publications/reports/2021/07/empowering-the-health-workforce-to-make-the-most-of-the-digital-revolution_4a9f0ad5/37ff0eaa-en.pdf [accessed 2026-01-27]
26. Regulations, standards & guidelines. NZ Telehealth Forum & Resource Centre. 2024. URL: https://www.telehealth.org.nz/regulations-and-policies/regulations-and-standards [accessed 2026-01-27]
27. Peter MacCallum Cancer Centre. 2023. URL: https://www.petermac.org [accessed 2026-01-27]
28. Top 20: Peter Mac named global leader in cancer care. Peter MacCallum Cancer Centre. 2024. URL: https://www.petermac.org/about-us/news-and-events/news/details/top-20-peter-mac-named-global-leader-in-cancer-care [accessed 2026-01-27]
29. Dillman DA, Smyth JD, Christian LM. Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method, 4th edition: Wiley; 2014. [doi: 10.1002/9781394260645]
30. Wasti SP, Simkhada P, van Teijlingen ER, Sathian B, Banerjee I. The growing importance of mixed-methods research in health. Nepal J Epidemiol 2022 Mar;12(1):1175-1178. [doi: 10.3126/nje.v12i1.43633] [Medline: 35528457]
31. Wellman N, Tröster C, Grimes M, Roberson Q, Rink F, Gruber M. Publishing multimethod research in AMJ: a review and best-practice recommendations. Acad Manag J 2023 Aug;66(4):1007-1015. [doi: 10.5465/amj.2023.4004]
32. The Joanna Briggs checklist for systematic reviews and research syntheses: critical appraisal tools for use in JBI systematic reviews. : Joanna Briggs Institute; 2020 URL: https://jbi.global/sites/default/files/2020-08/Checklist_for_Systematic_Reviews_and_Research_Syntheses.pdf [accessed 2026-01-27]
33. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
34. The PRISMA2020 statement: an updated guideline for reporting systematic reviews. Equator Network. 2022. URL: https://www.equator-network.org/reporting-guidelines/prisma [accessed 2026-01-27]
35. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100. [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
36. The Joanna Briggs Institute critical appraisal tools for use in JBI systematic reviews: checklist for qualitative research. : Joanna Briggs Institute; 2017 URL: https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Qualitative_Research2017_0.pdf [accessed 2026-01-27]
37. CASP checklist: CASP cross-sectional studies checklist. Critical Appraisal Skills Program (CASP). 2025. URL: https://casp-uk.net/casp-tools-checklists/cross-sectional-studies-checklist [accessed 2026-01-27]
38. Duncan Millar J, Mason H, Kidd L. What is Q methodology? Evid Based Nurs 2022 Jul;25(3):77-78. [doi: 10.1136/ebnurs-2022-103568] [Medline: 35609957]
39. Damio SM. Q methodology: an overview and steps to implementation. Asian J Univ Educ 2016;12(1):105-122 [FREE Full text]
40. Lundberg A, de Leeuw R, Aliani R. Using Q methodology: sorting out subjectivity in educational research. Educ Res Rev 2020 Nov;31:100361. [doi: 10.1016/j.edurev.2020.100361]
41. Q methodology website. better evaluation knowledge. 2024. URL: https://www.betterevaluation.org/tools-resources/q-methodology-website [accessed 2026-01-27]
42. Holt L, Department of Human Services. Understanding program logic. : Victoria Department of Health; 2009 URL: https://content.health.vic.gov.au/sites/default/files/migrated/files/collections/policies-and-guidelines/u/understanding_program_logic---pdf.pdf [accessed 2026-01-27]
43. Department of Human Services. Understanding program logic. : Victoria Department of Health; 2013 URL: https://www.health.vic.gov.au/publications/understanding-program-logic [accessed 2026-01-27]
44. Kirkpatrick Partners. The Kirkpatrick Model. 2024. URL: https://www.kirkpatrickpartners.com/the-kirkpatrick-model [accessed 2026-02-13]
45. National statement on ethical conduct in human research. : National Health and Medical Research Council (NHMRC); 2023 URL: https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2023 [accessed 2026-01-27]
46. Australian code for responsible conduct of research. : National Health and Medical Research Council (NHMRC); 2018 URL: https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018 [accessed 2026-01-27]
47. Brownie S, Zarb L, Cochrane L, Bonacaro A, Broman P. Clinical guidelines for virtual health assessment and care: a scoping review. J Adv Nurs 2026 Jan;82(1):142-161. [doi: 10.1111/jan.16952] [Medline: 40195584]

XSL•FO
RenderX

48. Ma LL, Wang YY, Yang ZH, Huang D, Weng H, Zeng XT. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? Mil Med Res 2020 Feb 29;7(1):7. [doi: 10.1186/s40779-020-00238-8] [Medline: 32111253]
49. Cohen J. Statistical power analysis. Curr Dir Psychol Sci 1992 Jun;1(3):98-101. [doi: 10.1111/1467-8721.ep10768783]
50. Q-methodology. better evaluation knowledge. 2024. URL: https://www.betterevaluation.org/methods-approaches/methods/q-methodology [accessed 2026-01-27]
51. Q methodology: complete beginner's guide. Jonathan Sandling. 2024. URL: https://jonathansandling.com/q-methodology-complete-beginners-guide/ [accessed 2026-01-27]
52. Schwartz S. Educating the nurse of the future: report of the independent review of nursing education. 2019 URL: https://www.health.gov.au/sites/default/files/documents/2019/12/educating-the-nurse-of-the-future.pdf [accessed 2026-01-27]
53. Marx M. Telehealth education in nursing curricula enhances nursing.: University of North Dakota; 2021 URL: https://commons.und.edu/cgi/viewcontent.cgi?article=1354&context=nurs-capstones [accessed 2026-01-27]
54. Zhao L, Abdolkhani R, Walter R, Petersen S, Butler-Henderson K, Livesay K. National survey on understanding nursing academics' perspectives on digital health education. J Adv Nurs 2024 Dec;80(12):4888-4899. [doi: 10.1111/jan.16163] [Medline: 38558473]
55. Registered Nurse Accreditation Standards 2019. : Australian Nursing and Midwifery Council (ANMAC); 2019 URL: https://anmac.org.au/sites/default/files/2025-10/anmac_rnas2019_evidence_guide_updated_2025.pdf [accessed 2026-01-27]
56. Australian Government Department of Health and Aged Care. National informatics standards for nurses and midwives. : Australian Nursing and Midwifery Federation; 2015 URL: https://www.anmf.org.au/media/kgdb4ozx/national_informatics_standards_for_nurses_and_midwives.pdf [accessed 2026-01-27]
57. Telehealth guidance for practitioners. AHPRA - Australian Health Practitioner Regulation Agency. 2020. URL: https://www.ahpra.gov.au/Resources/Information-for-practitioners-who-provide-virtual-care.aspx [accessed 2026-01-27]

Abbreviations

JBI: Joanna Briggs Institute
PLM: program logic model
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
REDCap: Research Electronic Data Capture

Edited by J Sarvestan; submitted 03.Aug.2025; peer-reviewed by P Koppel, SC Chen; revised version received 27.Dec.2025; accepted 29.Dec.2025; published 20.Feb.2026.

Please cite as:
Brownie S, Parkinson-Zarb L, Dimech A, Ftanou M, Broman P
Toward Best Practice Guidelines and Curricula for Virtual Care (Telehealth) in a Cancer Center: Protocol for a Multimethod Study
JMIR Res Protoc 2026;15:e81768
URL: https://www.researchprotocols.org/2026/1/e81768
doi:10.2196/81768

© Sharon Brownie, Lauren Parkinson-Zarb, Maria Ftanou, Andrew Dimech, Patrick Broman. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Surgical Outcome of Iris Claw and Scleral Tuck Intraocular Lens in Primary Cataract Surgery by Phacoemulsification in a Tertiary Care Center: Protocol for a Randomized Controlled Trial

Mohammed Kaderi*, MBBS; Sachin Daigavane*, MBBS, MS

Department of Ophthalmology, Acharya Vinoba Bhave Rural Hospital, Jawaharlal Nehru Medical College, Datta Meghe Institute of Higher Education and Research, Sawangi, Wardha, Maharashtra, India

*all authors contributed equally

Corresponding Author:
Mohammed Kaderi, MBBS
Department of Ophthalmology, Acharya Vinoba Bhave Rural Hospital, Jawaharlal Nehru Medical College, Datta Meghe Institute of Higher Education and Research, Sawangi, Wardha, Maharashtra, India

Abstract

Background: Cataract remains the leading cause of reversible blindness worldwide, particularly in older adults, in whom capsular weakness during phacoemulsification may preclude posterior chamber intraocular lens (IOL) implantation. In such scenarios, alternative fixation methods such as scleral-sutured and iris-claw lenses are preferred. However, direct comparative evidence regarding their safety, stability, and postoperative visual outcomes during primary cataract surgery is limited.

Objective: This study aims to compare the surgical and visual outcomes between retropupillary iris-claw IOL implantation and scleral-sutured fixation of single-piece foldable IOLs performed during primary cataract surgery by phacoemulsification in patients with inadequate capsular support.

Methods: This is a prospective, 2-arm, open-label randomized controlled trial being conducted at the Department of Ophthalmology, Acharya Vinoba Bhave Rural Hospital, Sawangi, India. A total of 92 patients aged 35 to 75 years with intraoperative posterior capsular compromise will be randomized (block randomization, sealed opaque envelopes) into 2 groups: group A (retropupillary iris-claw IOL) and group B (scleral-sutured IOL). Preoperative, intraoperative, and postoperative data will be collected at day 1, day 15, day 30, and day 90. The primary outcomes are best-corrected visual acuity and lens-induced astigmatism at 3 months. Secondary outcomes include IOL centration or tilt, intraoperative and postoperative complications, and patient satisfaction (assessed using the National Eye Institute Visual Function Questionnaire-25). Data will be analyzed using ANOVA and mixed-effects models.

Results: Recruitment commenced in March 2025 and is expected to conclude by September 2025, with interim data analysis scheduled for November 2025. As of the current submission, patient enrollment is ongoing, and no interim results have been analyzed.

Conclusions: This randomized protocol is designed to generate comparative clinical evidence on the efficacy, safety, and visual stability of iris-claw and scleral-sutured IOL fixation in primary cataract surgery. The findings will contribute to optimizing surgical decision-making in cases with deficient posterior capsular support.

Trial Registration: Clinical Trial Registry — India CTRI/2025/03/083687; https://tinyurl.com/bdexxmjy

International Registered Report Identifier (IRRID): PRR1-10.2196/80943

(JMIR Res Protoc 2026;15:e80943) doi:10.2196/80943

KEYWORDS

retropupillary iris claw; scleral tuck lens; cataract surgery; intraocular lens; phacoemulsification; surgical outcome; blindness

Introduction

Overview

Cataract remains the leading cause of reversible blindness worldwide and continues to impose a major public health burden, particularly among older adults in low- and middle-income countries [1]. In 2020, an estimated 79 million individuals aged 50 years and older were affected globally, representing a 39% increase in prevalence since 1993 [2]. The only definitive treatment is surgical extraction of the opacified lens followed by intraocular lens (IOL) implantation [3]. Phacoemulsification with posterior chamber IOL placement is the preferred modern technique due to its faster recovery, smaller incision size, and superior visual outcomes [4]. However, intraoperative or pre-existing capsular rupture can

XSL•FO
RenderX

complicate surgery and preclude standard posterior chamber fixation, resulting in aphakia and unstable visual rehabilitation [3,5].

When the posterior capsule is compromised, secondary fixation techniques are required to achieve optical stability. The 2 widely practiced approaches are retropupillary iris-claw IOL fixation and scleral-sutured posterior chamber IOL fixation [6]. In iris-claw implantation, the lens is enclavated onto the midperipheral iris, providing a relatively simple and reproducible procedure with shorter operative time [5]. Conversely, scleral-sutured fixation secures the IOL haptics within scleral flaps using 10‐0 polypropylene (Prolene; Ethicon) sutures, offering stable posterior chamber positioning and a reduced risk of endothelial trauma but requiring greater surgical expertise [7]. Both techniques have evolved substantially, yet each carries distinct risks, such as iris chafing, pupil distortion, or suture degradation [8,9].

Comparative evidence between these 2 fixation methods primarily originates from retrospective or heterogeneous secondary IOL implantation studies [7,8]. A limited number of prospective randomized studies have evaluated their performance in primary cataract surgeries where capsular support is inadequate at the index operation. Existing studies report conflicting outcomes: some suggest lower intraoperative complications with iris-claw IOLs [7], while others demonstrate superior centration and long-term stability with scleral fixation [8,9]. Moreover, data on lens-induced astigmatism, patient-reported visual satisfaction, and postoperative quality of life remain scarce. The choice between scleral-sutured and retropupillary iris-claw fixation in primary cataract surgery requires solid evidence on their safety, effectiveness, and patient experience. This randomized trial addresses that gap by directly comparing both techniques under standardized surgical conditions. It will assess visual outcomes, complications, and patient-reported satisfaction to help guide reliable IOL selection when posterior capsular support is inadequate.

Research Question

Among adult patients undergoing primary cataract surgery by phacoemulsification with inadequate posterior capsular support, does retropupillary iris-claw IOL implantation provide comparable or superior visual and surgical outcomes compared with scleral-sutured fixation of single-piece foldable IOLs?

Hypothesis

This randomized controlled trial hypothesizes that retropupillary iris-claw IOL implantation will achieve visual outcomes (measured by best-corrected visual acuity [BCVA] and lens-induced astigmatism at 3 months) that are noninferior or superior to those obtained with scleral-sutured fixation. It is further hypothesized that the iris-claw technique will demonstrate shorter operative time and fewer intraoperative and postoperative complications, while maintaining comparable centration and patient-reported satisfaction scores.

Methods

Study Design

This will be a prospective, hospital-based, 2-arm, open-label randomized controlled trial designed to compare the surgical and visual outcomes between retropupillary iris-claw IOL implantation and scleral-sutured fixation of single-piece foldable IOLs in patients undergoing primary cataract surgery by phacoemulsification. The trial will be conducted at the Department of Ophthalmology, Acharya Vinoba Bhave Rural Hospital (AVBRH), Sawangi, Wardha, India.

Study Setting

All surgical and postoperative procedures will be performed in the dedicated ophthalmic operation theaters of AVBRH under standardized aseptic conditions by trained consultant ophthalmic surgeons. Each participating surgeon will have independently performed at least 30 cases of both iris-claw and scleral-sutured IOL fixation techniques to ensure procedural consistency.

Study Population

Patients aged between 35 and 75 years scheduled for primary cataract surgery by phacoemulsification, in whom intraoperative posterior capsular rupture or deficient zonular support precludes in-the-bag IOL implantation, will be considered for inclusion. Inadequate posterior capsular support was defined as one or more of the following intraoperative conditions: posterior capsular rupture with vitreous disturbance, zonular dialysis exceeding 3 clock hours, or combined zonular-capsular instability preventing safe in-the-bag fixation. All surgeons will follow this standardized definition to ensure uniform eligibility.

Eligibility Criteria

Inclusion and exclusion criteria for study participation are summarized in Textbox 1 (see Multimedia Appendix 1 for the sample size calculation).

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria of the study population.

Inclusion Criteria

- Patients aged 35 to 75 years undergoing primary cataract surgery by phacoemulsification
- Presence of inadequate posterior capsular support identified intraoperatively, including posterior capsular rupture with vitreous disturbance or zonular dialysis exceeding 3 clock hours
- Inability to safely implant an in-the-bag IOL due to capsular or zonular instability
- Both male and female patients
- Willingness to provide written informed consent
- Ability and willingness to comply with postoperative follow-up schedule

Exclusion Criteria

- Pre-existing corneal diseases such as keratitis, corneal dystrophy, or corneal opacities
- History of chronic or recurrent uveitis
- Advanced glaucoma (intraocular pressure greater than 21 mm Hg despite medication or optic nerve cup-to-disc ratio greater than 0.7)
- Retinal pathologies, including diabetic retinopathy, hypertensive retinopathy, age-related macular degeneration, or retinitis pigmentosa
- Traumatic aphakia or lens subluxation secondary to ocular trauma
- Active ocular infection or uncontrolled systemic disease likely to impair wound healing

Randomization, Allocation Concealment, and Blinding

Eligible patients will be randomly allocated in a 1:1 ratio into 2 groups: group A (retropupillary iris-claw IOL) and group B (scleral-sutured fixation). A computer-generated block randomization sequence (block size of 4) will be prepared by an independent statistician. The group assignment will be concealed in sequentially numbered, opaque, sealed envelopes opened only after capsular deficiency confirmation during surgery.

Due to the nature of the interventions, blinding of surgeons and patients will not be feasible. However, postoperative outcome assessors and data analysts will remain blinded to the intervention group to minimize observer bias. Postoperative assessors will follow standardized grading protocols and will remain masked to group assignment to reduce detection bias.

Intervention Procedures

Both procedures will be performed under peribulbar anesthesia using a standardized phacoemulsification technique.

For group A, a retropupillary iris-claw IOL will be implanted using standard enclavation of haptics onto the posterior iris surface after thorough anterior vitrectomy. For group B, a single-piece foldable IOL will be fixated with 10‐0 polypropylene (Prolene) sutures passed through partial-thickness scleral flaps created at the 3- and 9-o'clock positions, ensuring equal centration. All surgeries will adhere to uniform intraoperative and postoperative protocols, including the use of balanced salt solution, intracameral antibiotics, and a topical steroid–antibiotic combination therapy postsurgery. To ensure reproducibility, procedural videos will be periodically reviewed by the data monitoring committee for intersurgeon standardization.

Timeline of Assessments

The schedule of preoperative, intraoperative, and postoperative assessments is outlined in Table 1.

XSL•FO
RenderX

Table . Timeline of clinical assessments and outcome measures.

Study phase	Assessment parameters
Preoperative	Visual acuity (Snellen's chart), slit-lamp biomicroscopy, intraocular pressure (NCT[a]), fundus evaluation, and biometry (A-scan, keratometry)
Day 0 (surgery)	Type of IOL[b] fixation, intraoperative complications, surgical duration, and intraoperative stability
Day 1	Corneal clarity, anterior chamber reaction, intraocular pressure, and early complications
Day 15	BCVA[c], astigmatism (autorefractometry), anterior chamber reaction, and IOL position
Day 30	BCVA, lens centration/tilt (slit-lamp retroillumination as per Calzetti et al [10]), intraocular pressure, and patient comfort
Day 90	Final BCVA, lens-induced astigmatism, postoperative complications, and patient satisfaction (NEI VFQ-25[d])

[a]NCT: noncontact tonometry.

[b]IOL: intraocular lens.

[c]BCVA: best-corrected visual acuity.

[d]NEI VFQ-25: National Eye Institute Visual Function Questionnaire-25.

Outcome Measures

Primary outcomes:

1. BCVA at 3 months using Snellen's chart, converted to logMAR values for statistical analysis.
2. Lens-induced astigmatism, measured through keratometry and autorefractometry.

Secondary outcomes:

1. Intraoperative complications (posterior capsular rupture extension, vitreous loss, surgical duration).
2. Postoperative complications (hyphema, pigment dispersion, raised intraocular pressure, IOL decentration, cystoid macular edema).
3. IOL centration and tilt, assessed by slit-lamp retroillumination following the method of Calzetti et al [10].
4. Patient-reported satisfaction and visual function, measured using the National Eye Institute Visual Function Questionnaire-25 at 3 months.

Adverse Event Reporting and Monitoring

All adverse events (AEs) will be documented at each follow-up. Serious AEs will be reported to the Institutional Ethics Committee within 7 days and nonserious AEs within 30 days. A data monitoring committee comprising 2 ophthalmologists and 1 biostatistician will perform quarterly reviews of safety, recruitment progress, and protocol compliance. Any protocol amendments will require prior institutional ethics committee approval and updating of the Clinical Trial Registry of India.

Data Management and Statistical Analysis

Statistical methods used for data management and analysis are summarized in this section. Continuous variables (eg, BCVA, astigmatism, and intraocular pressure) will be summarized as mean (SD) and analyzed using the independent *t* test for between-group comparisons and repeated-measures ANOVA for longitudinal analysis. Categorical variables (eg, complications and sex) will be compared using the chi-square test or the Fisher exact test. Multivariate analysis will be performed using mixed-effects linear regression to adjust for baseline covariates (age and preoperative vision). Missing data will be handled using multiple imputation under a missing-at-random assumption. Interim analysis will be conducted after 50% enrollment to evaluate safety and adherence; no stopping rules for efficacy are planned. Statistical analyses will be performed using SPSS software (version 27.0; IBM Corp). A P value <.05 will be considered statistically significant.

Ethical Considerations

The study was approved by the Institutional Ethics Committee of Datta Meghe Institute of Higher Education and Research (DMIHER(DU)/IEC/2024/24, dated May 3, 2024) and was prospectively registered with the Clinical Trial Registry of India (CTRI/2025/03/083687) on March 28, 2025. The trial will be conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. Written informed consent will be obtained from all participants prior to enrollment. Participant privacy and confidentiality will be maintained, and all data will be deidentified prior to analysis and publication. Data will be recorded in standardized case report forms and entered into a secure password-protected database. All study data will be anonymized prior to analysis and dissemination. Personally identifiable information, including names and medical record numbers, will be removed or replaced with coded identifiers. The reidentification key will be stored separately in a secure, password-protected database accessible only to the principal investigator.

Results

Patient recruitment began in March 2025 at the Department of Ophthalmology, AVBRH, Sawangi, India, with an estimated

XSL•FO
RenderX

completion date of September 2025. As of this submission, participant enrollment and follow-up are ongoing, and no interim data have been analyzed. Preliminary observations indicate adherence to the standardized surgical protocol and follow-up schedule. Data collection for all preoperative, intraoperative, and postoperative variables, including BCVA, IOL centration, and patient-reported satisfaction, will continue as per the CONSORT flow framework. Final analysis and dissemination of results are expected by early 2026.

Discussion

Management of aphakia during cataract surgery remains a critical surgical challenge, particularly when posterior capsular rupture or zonular instability precludes in-the-bag IOL implantation [11,12]. Although several fixation techniques have been developed, the comparative efficacy and safety of scleral-sutured and retropupillary iris-claw fixation during primary cataract surgery remain insufficiently explored [13]. This protocol aims to prospectively evaluate these two commonly used methods under standardized conditions. The rationale for comparing scleral-sutured and iris-claw fixation stems from their distinct anatomical and procedural advantages. Retropupillary iris-claw lenses offer a posterior chamber position similar to that of an in-the-bag IOL while avoiding contact with the corneal endothelium and anterior chamber angle [2,3]. They are relatively simple to implant, require shorter surgical time, and provide satisfactory centration with minimal postoperative astigmatism [5,7]. Conversely, scleral-sutured fixation restores a near-physiological IOL position behind the iris plane and has been associated with excellent visual outcomes and long-term stability [8]. However, it demands greater surgical skill and carries the potential of suture-related complications, including knot erosion and late dislocation [14,15].

Several retrospective and comparative studies have examined the outcomes of these techniques, but most involve heterogeneous populations or secondary IOL implantations. Madhivanan et al [5] demonstrated that both retropupillary iris-claw and scleral-fixated IOLs achieve significant improvement in BCVA, although scleral fixation required longer surgical duration. Similarly, Durmus et al [7] reported comparable visual outcomes but found that endothelial cell loss was lower with iris-claw lenses than with sutureless scleral fixation. Shahid et al [8] emphasized that scleral fixation provides superior long-term centration yet remains technically demanding and time-intensive. These findings justify the need for a prospective, randomized evaluation focusing specifically on primary cataract cases rather than secondary or complicated aphakia.

This trial's methodological design addresses previous limitations by incorporating randomization, allocation concealment, and masked outcome assessment. Objective parameters such as IOL tilt and decentration will be evaluated using slit-lamp retroillumination and validated anterior-segment optical coherence tomography methods described by Calzetti et al [10], ensuring quantitative reproducibility. In addition, the inclusion of patient-reported outcomes using the National Eye Institute Visual Function Questionnaire-25 will provide valuable insights into subjective satisfaction and functional vision, which have often been underrepresented in earlier studies.

Ethical oversight, AE monitoring, and quarterly data review by an independent committee have been included to maintain methodological rigor and compliance with Good Clinical Practice guidelines. The statistical plan using mixed-effects modeling and repeated-measures ANOVA will allow robust analysis of longitudinal data while accounting for missing values. The trial does not include an economic analysis, as cost-effectiveness was beyond the scope of the present protocol; however, results regarding operative time and complication rates may have indirect implications for resource optimization in tertiary-care settings. Potential limitations of this study include its single-center design and limited sample size, which may influence external generalizability. Nonetheless, the use of standardized surgical techniques, strict eligibility criteria, and a uniform postoperative evaluation schedule are expected to enhance internal validity and reproducibility.

Acknowledgments

No generative artificial intelligence tools were used to generate, analyze, or interpret scientific data for this manuscript. Artificial intelligence–based assistance was used only for language editing, and all scientific content was fully reviewed and verified by the authors. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Funding

This study did not receive any external funding. All research activities were supported by institutional resources.

Data Availability

Data sharing is not applicable to this paper, as no datasets were generated or analyzed during this study.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Sample size calculation.
[DOCX File, 39 KB - resprot_v15i1e80943_app1.docx]

Checklist 1
SPIRIT checklist.
[DOCX File, 17 KB - resprot_v15i1e80943_app2.docx]

References

1. Wan Z, Bai J, Wang W, Peng Q. Global, regional, and national burden of cataract among older adults from 1990 to 2021: a comprehensive analysis based on the Global Burden of Disease Study 2021. Front Med 2025;12:1679828. [doi: 10.3389/fmed.2025.1679828]
2. Drolsum L, Kristianslund O. Implantation of retropupillary iris-claw lenses: a review on surgical management and outcomes. Acta Ophthalmol 2021 Dec;99(8):826-836. [doi: 10.1111/aos.14824] [Medline: 33683012]
3. Seknazi D, Colantuono D, Tahiri R, Amoroso F, Miere A, Souied EH. Secondary sutureless posterior chamber lens implantation with two specifically designed IOLs: iris claw lens versus sutureless trans-scleral plugs fixated lens. J Clin Med 2021 May 20;10(10):2216. [doi: 10.3390/jcm10102216] [Medline: 34065508]
4. de Silva SR, Riaz Y, Evans JR. Phacoemulsification with posterior chamber intraocular lens versus extracapsular cataract extraction (ECCE) with posterior chamber intraocular lens for age-related cataract. Cochrane Database Syst Rev 2014 Jan 29;2014(1):CD008812. [doi: 10.1002/14651858.CD008812.pub2] [Medline: 24474622]
5. Madhivanan N, Sengupta S, Sindal M, Nivean PD, Kumar MA, Ariga M. Comparative analysis of retropupillary iris claw versus scleral-fixated intraocular lens in the management of post-cataract aphakia. Indian J Ophthalmol 2019 Jan;67(1):59-63. [doi: 10.4103/ijo.IJO_326_18] [Medline: 30574894]
6. Thulasidas M. Retropupillary iris-claw intraocular lenses: a literature review. Clin Ophthalmol 2021;15:2727-2739. [doi: 10.2147/OPTH.S321344] [Medline: 34211261]
7. Durmus E, Aykut V, Hepokur M, et al. Visual outcomes, endothelial loss, and complications following iris-claw intraocular lens implantation and trocar-assisted sutureless scleral fixation of intraocular lenses. Beyoglu Eye J 2023;8(4):253-259. [doi: 10.14744/bej.2023.30085] [Medline: 38089072]
8. Shahid SM, Flores-Sánchez BC, Chan EW, et al. Scleral-fixated intraocular lens implants-evolution of surgical techniques and future developments. Eye (Lond) 2021 Nov;35(11):2930-2961. [doi: 10.1038/s41433-021-01571-5] [Medline: 34117388]
9. Park HS, Park SC, Kim JH, Ohn YH, Park TK, Han JW. Surgical outcomes of retropupillary iris-claw intraocular lens implantation for various indications. Ophthalmol Ther 2025 Jul;14(7):1491-1501. [doi: 10.1007/s40123-025-01153-8] [Medline: 40366604]
10. Calzetti G, Bellucci C, Tedesco SA, Rossi M, Gandolfi S, Mora P. Tilt and decentration of posterior and anterior iris-claw intraocular lenses: a pilot study using anterior segment optical coherence tomography. BMC Ophthalmol 2022 May 23;22(1):233. [doi: 10.1186/s12886-022-02430-x] [Medline: 35606746]
11. Kumar M, Chandak N. Managing surgical aphakia and posterior capsular rupture with suture-assisted intraocular lens implantation. Oman J Ophthalmol 2024;17(3):334-336. [doi: 10.4103/ojo.ojo_239_23] [Medline: 39651509]
12. Fiorentzis M, Viestenz A, Heichel J, Seitz B, Hammer T, Viestenz A. Methods of fixation of intraocular lenses according to the anatomical structures in trauma eyes. Clin Anat 2018 Jan;31(1):6-15. [doi: 10.1002/ca.22898] [Medline: 28509425]
13. Moghib K, Salomon I, Abdelglel Y, et al. Anterior chamber and retropupillary iris-claw intar ocular lens fixation comparison of clinical outcomes: a systematic review. Ann Med Surg 2025 Jan;87(1):299-308. [doi: 10.1097/MS9.0000000000002856]
14. Yu T, Yu M, Wu W, et al. Outcomes and complications of sutured scleral-fixated foldable intraocular lens implantation: a retrospective study of 5-year follow-up. J Ophthalmol 2021;2021:5525064. [doi: 10.1155/2021/5525064] [Medline: 34327011]
15. Sun H, Wang C, Wu H. Recent advances and current challenges in suture and sutureless scleral fixation techniques for intraocular lens: a comprehensive review. Eye Vis (Lond) 2024 Dec 30;11(1):49. [doi: 10.1186/s40662-024-00414-0] [Medline: 39736769]

Abbreviations

AE: adverse event
AVBRH: Acharya Vinoba Bhave Rural Hospital
BCVA: best-corrected visual acuity
IOL: intraocular lens

XSL•FO
RenderX

Edited by J Sarvestan; submitted 19.Jul.2025; peer-reviewed by H Tanabe, R Bellucci; revised version received 16.Dec.2025; accepted 17.Dec.2025; published 23.Feb.2026.

Please cite as:
Kaderi M, Daigavane S
Surgical Outcome of Iris Claw and Scleral Tuck Intraocular Lens in Primary Cataract Surgery by Phacoemulsification in a Tertiary Care Center: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e80943
URL: https://www.researchprotocols.org/2026/1/e80943
doi:10.2196/80943

© Mohammed Kaderi, Sachin Daigavane. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Effect of the DemensiaKITA Mobile Health App Intervention on Knowledge, Attitude, Practice, and Burden Level of Dementia Caregivers in Kuala Lumpur and Selangor, Malaysia: Protocol for a Nonrandomized Controlled Trial

Nurulizzah Mahfar[1*], MPH, MD; Nik Nairan Abdullah[1*], MBChB, MPH, PhD; Dalila Roslan[1*], MB BCh BAO, MPH, DrPH; Khalid Ibrahim[1*], MS, MHHSM; Ungku Ahmad Ameen Ungku Mohd Zam[2*], MBBS, M Int Med; Noraliza Noordin Merican[3*], MPH, MD; Xin Wee Chen[1], MPH, MD, DrPH (Epidemiology); Mariam Mohamad[1], M Comm Health, MD

[1]Public Health Medicine Department, Faculty of Medicine, Universiti Teknologi MARA, Sg Buloh Campus, Selangor Branch, Sungai Buloh, Malaysia

[2]Geriatric Unit, Medical Department, Hospital Tengku Ampuan Rahimah, Klang, Malaysia

[3]Elderly Health Sector, Family Health Development Division, Ministry of Health, Putrajaya, Malaysia

[*]these authors contributed equally

Corresponding Author:
Nik Nairan Abdullah, MBChB, MPH, PhD
Public Health Medicine Department, Faculty of Medicine, Universiti Teknologi MARA, Sg Buloh Campus, Selangor Branch, Sungai Buloh, Malaysia

Abstract

Background: Dementia is a growing public health concern, disproportionately affecting low- and middle-income countries. Caregivers of people living with dementia often face significant physical, psychological, social, and financial burdens, with high prevalence rates of caregiver strain in Malaysia. Mobile health (mHealth) apps have demonstrated potential to enhance caregivers' knowledge, attitudes, and practices (KAP) and to reduce burden. However, few culturally tailored solutions exist for Malaysia. The *DemensiaKITA* app was developed to provide locally relevant information, support services, and stress management tools for dementia caregivers.

Objective: The purpose of this single-blinded, nonrandomized controlled trial (NRCT) is to evaluate the effectiveness of the *DemensiaKITA* mHealth app in improving caregivers' KAP and reducing caregiver burden in Selangor and Kuala Lumpur, Malaysia.

Methods: This research will be conducted in 2 phases. Phase 1 involves the adaptation, translation, and validation of 4 instruments: Dementia Knowledge Assessment Tool Version 2, Dementia Attitudes Scale, Caregiver Task Inventory—25 items, and Short-form Zarit Burden Interview—12 items. Content, face, and construct validity, along with reliability (Cronbach α and test-retest intraclass correlation), will be established. Phase 2 is a single-blinded NRCT conducted in 4 public hospitals. A total of 100 caregivers will be recruited. Two intervention hospitals will receive the *DemensiaKITA* app, while 2 control hospitals will receive usual care (Ministry of Health dementia leaflets and video). The primary outcomes (KAP) and secondary outcome (burden level) will be assessed at baseline, 1 month, and 3 months using the validated questionnaires on an intention-to-treat basis. Data will be analyzed with descriptive statistics, chi-square and *t* tests, and generalized estimating equations.

Results: Recruitment and baseline data collection are underway. Recruitment for Phase 1 (instrument validation) occurred from December 2024 until July 2025. It is expected to produce valid and reliable Malay versions of KAP and burden instruments. Recruitment for the Phase 2 NRCT will occur from August 2025 until February 2026. The first follow-up data collection for Phase 2 will occur 1 month after baseline, and the second and final follow-up will occur at 3 months. Data analysis has not yet begun. Phase 2 is expected to show significant improvements in dementia KAP and reduced caregiver burden in the intervention group compared to controls. This study was conducted without any specific external financial support or grants from any public, commercial, or not-for-profit funding agencies. The study findings are expected to be published in December 2026.

Conclusions: This article describes the protocol for a single-blinded NRCT examining a novel mHealth intervention. The *DemensiaKITA* app has the potential to empower caregivers, enhance dementia care practices, and alleviate caregiver burden in Malaysia and may serve as a model for other low- and middle-income countries globally.

Trial Registration: ClinicalTrials.gov NCT07212894; https://clinicaltrials.gov/study/NCT07212894

International Registered Report Identifier (IRRID): DERR1-10.2196/83699

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e83699) doi:10.2196/83699

KEYWORDS

dementia; caregivers; caregiver burden; health knowledge, attitudes, practice; mobile apps; nonrandomized controlled trial

Introduction

Dementia is a progressive syndrome associated with cognitive decline and behavioral and psychological symptoms, requiring substantial support from caregivers [1]. In 2018, an estimated 50 million people were living with dementia globally, with 60% residing in low- and middle-income countries [2]. In Malaysia, the prevalence of dementia among older adults is 8.5% [1,3].

Caregivers often experience significant physical, emotional, social, and financial strain [1,2,4,5], exacerbated in low-income countries where formal support services for people living with dementia remain limited [3,6]. Caregiver burden is highly prevalent, with global estimates at 49.26% and Malaysian estimates reaching 69.4% [7,8]. Additionally, a longitudinal study in Sydney, Australia, demonstrated an increase in caregiver burden from 47.4% at baseline to 56.8% after 3 years [9].

Research consistently shows substantial gaps in dementia knowledge and variable attitudes among caregivers. A survey in China reported low knowledge despite moderate attitudes [10], while a study in India found that many caregivers had only average or below-average knowledge, low to moderate attitudes, and predominantly poor caregiving practices. Evidence indicates these gaps can be improved through targeted interventions. For example, a randomized controlled trial in eastern Taiwan showed that mobile eLearning significantly enhanced dementia knowledge, attitudes, and care competency among home care workers [11]. Positive attitudes have also been associated with better caregiving practices, highlighting the value of interventions that strengthen both knowledge and attitudes through communication training, safe care techniques, behavioral management strategies, psychological support, and information on available services [10].

The rapid expansion of internet access has enabled the use of mobile health (mHealth) apps to support caregivers [3]. Prior studies demonstrate that mHealth tools can provide health information, connect users with local services, and reduce caregiver burden [5,12-14]. However, most dementia-focused apps are developed in high-income countries and predominantly in English, limiting their applicability in diverse cultural contexts [12,15,16]. More rigorous evaluations are also needed to determine how mHealth interventions influence caregivers' knowledge, attitudes, and practices (KAP) and burden level [10,17-19].

DemensiaKITA is a Malaysian mHealth prototype developed to address this gap. The app provides information on dementia, caregiving strategies, management of behavioral symptoms, stress-reduction techniques, and access to local health and community resources. Its content is culturally tailored and delivered in Malay, developed collaboratively by dementia care experts and technology researchers [12]. However, its effectiveness in improving caregivers' KAP and reducing burden has not yet been evaluated.

This study aims to (1) adapt, translate, and validate Malay-language questionnaires assessing dementia caregivers' KAP and burden, and (2) evaluate the effectiveness of the *DemensiaKITA* mHealth app in improving KAP and reducing caregiver burden among dementia caregivers in Selangor and Kuala Lumpur through a nonrandomized controlled trial.

Methods

Study Design

This study consists of 2 sequential phases: (1) adaptation, translation, and validation of KAP and burden-related instruments, and (2) evaluation of the *DemensiaKITA* mHealth intervention via a single-blinded NRCT. The NRCT design was selected due to operational and ethical constraints within participating hospitals. Caregivers are tied to specific clinics based on people living with dementia follow-up, and redistributing them across hospitals for randomization would disrupt continuity of care. Cluster randomization was not feasible due to small eligible populations within each site. Individual randomization within the same clinic would introduce high contamination risks, as caregivers frequently interact and could share intervention materials.

To mitigate potential biases inherent in nonrandom allocation, the study incorporates several methodological safeguards: (1) standardized inclusion and exclusion criteria across all sites; (2) selection of hospitals with comparable geriatric services; (3) baseline comparison of groups; and (4) longitudinal analyses using generalized estimating equations (GEE) to adjust for baseline imbalances and repeated measurements [20,21]. This pragmatic design reflects real-world implementation conditions while preserving internal validity.

Study Population

The target population is dementia caregivers in Malaysia, while the accessible study population comprises primary caregivers of people living with dementia aged 18 years and older in Selangor. The sampling population is caregivers of registered people living with dementia attending the study sites from December 2024 to February 2026. The sampling unit is a consenting primary caregiver providing direct care to registered people living with dementia.

Phase 1: Instrument Adaptation, Translation, and Validation

Overview

Phase 1 involves adapting, translating, and validating 4 instruments into the Malay language: Dementia Knowledge Assessment Tool Version 2 (DKAT-2) [22], Dementia Attitudes Scale (DAS) [23], Caregiver Task Inventory—25 items (CTI-25) [24,25], and the Short-form Zarit Burden Interview—12 items

XSL•FO
RenderX

(ZBI-12) [26-28]. It is scheduled from December 2024 to July 2025.

There will be forward and backward translations of the questionnaires by two independent translators. Two independent translators will translate the original questionnaires into Malay and then back into the original language. This process helps identify and rectify any discrepancies or inconsistencies in meaning. The drafted Malay versions of the KAP questionnaires will be rigorously validated, encompassing content validation, face validation, construct validation, and an assessment of reliability (Figure 1). Eligibility criteria for Phase 1 are given in Textbox 1.

XSL•FO
RenderX

Figure 1. Study flow chart for Phase 1: adaptation, translation, and validation. CTI-25: Caregiver Task Inventory—25 items; CVR: content validity ratio; DAS: Dementia Attitudes Scale; DKAT-2: Dementia Knowledge Assessment Tool Version 2; EFA: exploratory factor analysis; HKL: Hospital Kuala Lumpur; ICC: intraclass correlation coefficient; I-CVI: item-level content validity index; I-FVI: item-level face validity index; KAP: knowledge, attitudes, and practices; KMO: Kaiser-Meyer-Olkin; S-CVI: scale-level content validity index; S-FVI: scale-level face validity index; UMMC: Universiti Malaya Medical Center; ZBI-12: Short-form Zarit Burden Interview—12 items.

XSL•FO
RenderX

Textbox 1. The inclusion and exclusion criteria for Phase 1.

Inclusion criteria

1. A primary caregiver of people living with dementia (a person who has helped in the personal care of a person living with dementia for 1 mo or more during the last 12 mo)
2. Aged 18 years or older
3. Cares for older adult aged 60 years or older with any form and severity of dementia
4. Informal and unpaid caregivers
5. Minimum of at least 4 h of care per day

Exclusion criteria

1. A person who is unable to communicate independently in Bahasa Malaysia
2. Domestic helper

Content Validity

Content validity will be evaluated by a panel of 6 experts comprising public health physicians, geriatricians, and 1 informal caregiver of people living with dementia who will serve as an experiential expert to ensure contextual and practical relevance for the target population. Each panel member will independently assess all the items for relevance, clarity, simplicity, and the absence of ambiguity, following established content-validity assessment procedures [29-32]. The ratings provided will be used to compute item-level and scale-level content validity indices (CVIs).

Face Validity

Face validity will be examined among 30 dementia caregivers who will assess the clarity and comprehensibility of each item. Their feedback will be used to calculate the face validity index in accordance with recommended guidelines [33-35]. This process will ensure that the instrument is clear and understandable to the target population.

Construct Validity

Construct validity will be assessed using exploratory factor analysis (EFA). Based on the recommended participant-to-item ratio of at least 5:1, a minimum of 330 participants is required, with an additional 10% to account for potential attrition, yielding a total target sample of 363 participants. EFA will be performed to evaluate the underlying factor structure and to determine whether the items load appropriately onto the intended domains [36,37]. Participants will be recruited through convenience sampling from the Geriatric Clinic at Universiti Malaya Medical Center (UMMC), Kuala Lumpur, and Hospital Kuala Lumpur. This sample size is expected to provide adequate power and stability for robust factor extraction and model interpretation.

Reliability

Reliability will be assessed through internal consistency and test-retest procedures. Internal consistency will be evaluated using Cronbach α for each domain and for the overall scale. Test-retest reliability will be examined over a 2-week interval using intraclass correlation coefficient, following established guidelines for reliability assessment [38-41]. This combined approach will ensure stability and consistency of the instrument across time.

Phase 2: Effectiveness of the *DemensiaKITA* mHealth App

Overview

In Phase 2, the research assesses the effectiveness of the *DemensiaKITA* mHealth app intervention among dementia caregivers in Selangor, Malaysia. This phase uses an NRCT with 2 parallel arms (intervention and control) and a 1:1 allocation ratio. The study includes 3 assessment points: baseline (T0), 1 month (T1), and 3 months (T2). The intervention period spans from August 2025 to February 2026. Participants will be blinded to group allocation (single-blinded). Hospital-level allocation is applied to minimize contamination (Figure 2). Eligibility criteria for Phase 2 are given in Textbox 2.

XSL•FO
RenderX

Figure 2. Study flowchart for Phase 2: effectiveness trial (nonrandomized controlled trial). CTI-25: Caregiver Task Inventory—25 items; DAS: Dementia Attitude Scale; DKAT-2: Dementia Knowledge Assessment Tool Version 2; GEE: generalized estimating equations; HASA: Hospital Al-Sultan Abdullah; HTAR: Hospital Tengku Ampuan Rahimah; MOH: Ministry of Health; UiTM: Universiti Teknologi MARA; ZBI-12: Short-form Zarit Burden Interview—12 items.

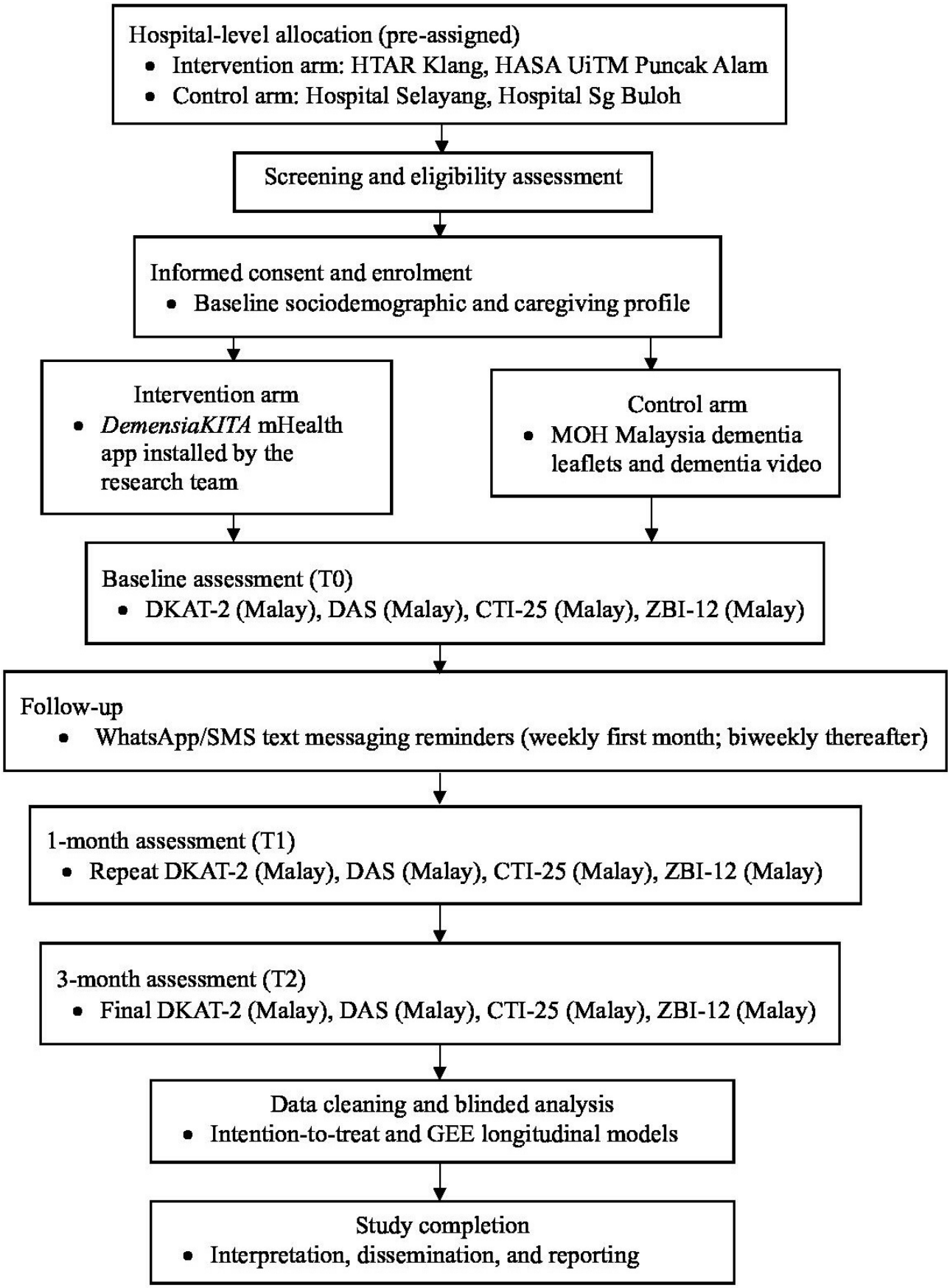

XSL•FO
RenderX

Textbox 2. Inclusion and exclusion criteria for Phase 2.

Inclusion criteria

1. Primary caregiver of people living with dementia who has provided personal care for 1 month or more
2. Aged 18 years or older
3. Caring for an older adult aged 60 years or older with any type of dementia
4. Informal and unpaid caregiver
5. Minimum of at least 4 hours of care per day
6. Possesses an operational Android smartphone with mobile data
7. Able to use basic smartphone functions and apps

Exclusion criteria

1. Unable to communicate independently in Bahasa Malaysia
2. Not an Android smartphone user
3. Employed domestic helper

Study Settings

The trial will be conducted in the geriatric clinics of 4 public hospitals in Selangor, Malaysia: (1) Hospital Tengku Ampuan Rahimah, Klang; (2) Hospital Sungai Buloh; (3) Hospital Selayang; (4) Hospital Al-Sultan Abdullah, Universiti Teknologi MARA (UiTM) Puncak Alam. These settings were chosen due to their established geriatric subspecialty services and continuous caseload of dementia caregivers. To prevent contamination, hospitals were allocated based on geographic distribution and patient-flow patterns—Intervention arm: Hospital Tengku Ampuan Rahimah, Klang, and Hospital Al-Sultan Abdullah, UiTM Puncak Alam; Control arm: Hospital Selayang and Hospital Sungai Buloh (Figure 2).

Because follow-up care rarely overlaps across hospitals, this site-based allocation minimizes cross-exposure. The *DemensiaKITA* app was installed only by the research team on participants' smartphones in the intervention arm. Caregivers were advised not to share the app or its contents during the study period.

Recruitment Strategy

Participants will be identified using purposive sampling among caregivers of registered people living with dementia attending the 4 participating geriatric clinics. Screening and recruitment will be conducted by the research team. Eligible caregivers will receive a full explanation of the study objectives, procedures, confidentiality safeguards, and follow-up requirements.

Allocation and Blinding

Because hospitals were preassigned to their respective study arms, no individual-level randomization will be implemented. The study will be conducted as a single-blinded trial whereby participants will not be informed of their allocation to either the intervention or control group and will instead be informed that they will receive "health education material," as stated in the patient information sheet. In addition, researchers responsible for data analysis will remain blinded to group assignment until all data have been cleaned and finalized to minimize analysis-related bias.

Intervention (Intervention Arm)

Participants assigned to the intervention hospitals will receive the *DemensiaKITA* mHealth app. The app will be installed directly by the research team to prevent unauthorized sharing. Participants may access the app freely during the study period.

DemensiaKITA Mobile Health App Intervention

Overview

The *DemensiaKITA* app is an offline mobile app prototype developed by a team of researchers from Universiti Teknologi MARA, Malaysia, and the Canadian University of Dubai (Figure 3). The app has been granted copyright protection under Universiti Teknologi MARA. It provides essential information on dementia and dementia care, facilitates easy access to nearby health services, optimizes the use of available local resources, and enhances caregiving skills. The app features appropriate and relevant content tailored for the Malaysian context, using the national language (Bahasa Malaysia) as the interface, specifically designed for dementia caregivers in Malaysia (Figure 3 and Textboxes 3 and 4) [12].

XSL•FO
RenderX

Figure 3. Features of *DemensiaKITA* mobile health app.



Textbox 3. Additional features of the DemensiaKITA app (support functions to reduce caregiver burden).

- Stress and emotional management: Resources to help caregivers cope with stress and emotional challenges, potentially reducing their emotional burden.
- Daily motivation: Encouragement and inspiration to support caregivers in their daily routines, potentially reducing their emotional and social burdens.
- Support group: Access to online and in-person support groups, helping caregivers connect with others and reduce feelings of isolation, potentially reducing emotional and social burdens.
- List of available services (including health services, nongovernment organizations, daycare, and nursing home): A comprehensive list of available services, enabling caregivers to find the support they need and potentially reduce their physical burden by facilitating access to care.
- List of available welfare aid: Details on available financial assistance and welfare programs to help alleviate the financial burden of caregiving.

Textbox 4. Basic features of the DemensiaKITA app (information and guidance on dementia care).

The *DemensiaKITA* app aims to improve the caregivers' knowledge, attitudes, and practices in taking care of people living with dementia. It provides information and guidance on various aspects of dementia care:

- Regarding dementia
- Emotional disturbance of people living with dementia
- Basic care of people living with dementia
- Dietary care of people living with dementia
- Oral exercise for people living with dementia
- Basic exercise for people living with dementia
- Physical and memory rehabilitation

Engagement and Adherence Monitoring

To promote engagement with the *DemensiaKITA* app, caregivers in the intervention group will receive a brief, standardized orientation on the app features at enrollment. Thereafter, the research team will send scheduled reminders via WhatsApp or SMS text messaging to encourage continued use of the app and completion of follow-up assessments. Reminders will be sent once weekly during the first month and every 2 weeks thereafter until the 3-month follow-up.

Comparator (Control Arm)

Caregivers in the control arm will receive standard Ministry of Health (MOH) dementia educational materials, which are the dementia leaflets and video. These materials will be explained in detail during the baseline session to ensure consistent educational exposure and engagement comparable to that provided in the intervention arm [42].

Dementia Leaflets and Video

The dementia leaflets and video provided by the MOH Malaysia offer essential information on the nature of dementia, its common symptoms, and the groups at higher risk (Figure 4). They emphasize the importance of early detection and timely diagnosis, describe the available assessments conducted in clinics and hospitals, and outline practical steps to help maintain cognitive and physical well-being. The materials also highlight key measures that may reduce dementia risk, including mentally stimulating activities, regular exercise, healthy lifestyle practices, and effective management of chronic illnesses. In

XSL•FO
RenderX

addition, they provide guidance on daily care, safety, behavioral management, and available support resources to assist people living with dementia and their caregivers [42].

Figure 4. Ministry of Health Malaysia leaflets and video.



Data Collection Method and Study Tool

Data will be collected using a validated self-administered questionnaire at 3 time points: baseline, 1 month postintervention, and 3 months postintervention. The study tool is a Malay-translated and validated questionnaire assessing KAP and burden levels. Participants in the intervention group will have access to the mobile app 2 weeks before the study starts to get familiar with it.

Before the study begins, all participants will be informed of the study objectives, procedures, potential risks, and benefits and will provide informed consent. Participation is voluntary, and participants may withdraw at any time without penalty. To help keep participants engaged, reminders will be sent via WhatsApp or email throughout the study. These reminders will also check on app use for the intervention group and encourage continued participation until 3 months postintervention.

Sample Size Calculation

The sample size for Phase 2 was calculated for comparison of 2 independent means (intervention vs control) assuming equal group sizes, a 2-sided test with a significance level of α=.05, and power (1−β)=0.80. The calculation followed the method described by Rosner [43] and Turnbull et al [44]. The required sample size per group was obtained using the formula:

$$n=(\sigma 12+\sigma 22)(Z1-\alpha/2+Z1-\beta)2\Delta 2$$

where σ1 and σ2 are the SDs in the 2 groups, Z1-α/2 is the standard normal deviate at the 5% (1.96) significance level, Z1-β is the standard normal deviate for 80% (0.84) power, and Δ is the expected mean difference between the groups.

Based on Turnbull et al [44], an expected mean difference in knowledge scores of 1.69 was used, with SDs of 2.53 and 2.76, respectively. Substituting these values:

$$n=(2.532+2.762)(1.96+0.842)21.692\approx 38.5$$

Thus, approximately 39 participants are required per group. To account for an anticipated 30% attrition rate, the sample size was inflated:

$$n_{adjusted}=38.5\times 1.30\approx 50$$

Accordingly, 50 participants will be recruited into each arm (intervention and control), giving a total sample size of 100 caregivers.

Study Outcome

The outcomes of this study will be evaluated using validated instruments measuring caregivers' KAP and burden level in relation to caring for people living with dementia.

Knowledge

Caregiver knowledge will be assessed using the DKAT-2, developed by Toye and colleagues [22] at Curtin University. The DKAT-2 consists of 21 items with Yes/No/Don't Know response options and measures understanding of the dementia trajectory and evidence-based care approaches. Each item is scored as correct or incorrect, with "Don't Know" categorized as incorrect. Total scores reflect overall knowledge, with higher

XSL•FO
RenderX

scores indicating greater understanding of dementia care. The DKAT-2 demonstrates strong psychometric performance, including high internal consistency (Cronbach α=0.79), supporting its use in evaluating caregiver education and guiding patient-centered interventions [22].

Attitude

Attitudes toward dementia will be measured using the DAS, developed by O'Connor and McFadden [23]. The DAS comprises 20 items rated on a 7-point Likert scale (1=strongly disagree to 7=strongly agree), capturing cognitive, emotional, and behavioral components of attitudes toward dementia. Higher scores denote more positive attitudes. The scale has demonstrated excellent reliability (Cronbach α=.83 - .85) and a stable 2-factor structure confirmed through exploratory and confirmatory factor analyses [23]. Its broad conceptual coverage makes it suitable for evaluating the impact of educational and psychosocial interventions.

Practice

Caregiving practices will be assessed using the CTI-25. The CTI-25 is a shortened and refined version of the original 45-item instrument developed by Clark and Rakowski in 1983 [24], with improvements introduced by Lee and Mok [25] to enhance clarity and reduce redundancy. The instrument contains 25 items across 5 domains: learning to cope with a new role; providing care based on the care recipient's needs; managing one's emotional needs; appraising supportive resources; and balancing personal and caregiving demands.

Items are rated on a 3-point Likert scale (0 - 2), with higher scores reflecting greater caregiving difficulty. The CTI-25 demonstrates strong psychometric properties, including high internal consistency (α=.93) and construct validity supported by confirmatory factor analysis [25].

Burden Level

Caregiver burden will be measured using the ZBI-12, a validated abbreviated version of the original 22-item ZBI developed by Zarit in 1980 [26-28]. The ZBI-12, translated into Malay in 2023, evaluates key burden domains including privacy, social life impact, emotional strain, health effects, and time-related pressures. The shorter format enables rapid screening and ongoing monitoring while retaining strong psychometric performance. Total scores are categorized as: 0 - 10 (no to mild burden); 11 - 20 (mild to moderate burden); and 20 or more (high burden) [26-28].

Harms and Safety Monitoring

This study evaluates an educational mHealth app and standard health education materials, which are considered minimal-risk behavioral interventions. No physical risks are anticipated. Harms will therefore be defined as any self-reported serious psychological distress or any adverse event that participants or clinicians consider to be related to use of the app or study procedures. At each follow-up contact, participants will be asked whether they experienced any problems or distress related to the intervention or study participation. Participants may discontinue app use or withdraw from the study at any time on request, without consequences for their routine care. Any serious adverse events judged to be related to the study will be reported to the relevant ethics committees in accordance with institutional policies.

Ethical Considerations

Ethical Approval

This study received ethical approval from multiple institutional review boards. Approval was first obtained from the Faculty Ethics Research Committee, UiTM on November 7, 2024 (REC/11/2024 [PG/MR/549]). The study was subsequently approved by the Medical Research and Ethics Committee, Ministry of Health Malaysia, on December 16, 2024 (24 - 03494-W6I [IIR]), in accordance with the ethical standards of the National Institutes of Health Malaysia. Additional approval was granted by the Medical Research Ethics Committee, UMMC, on May 21, 2025 (2025122 - 14637).

Operational approvals were obtained from the Directors of Hospital Kuala Lumpur and UMMC for Phase 1 and from the Director of Selangor State Health Department and the Directors of the 4 participating hospitals for Phase 2. Permission to use all study instruments was formally granted by the respective copyright holders and instrument developers.

Before the study begins, all participants will be informed of the study objectives, procedures, and potential risks and benefits and will provide informed consent. Written informed consent will be obtained prior to enrollment. No financial or material compensation will be provided to participants for their participation in this study.

Data Management and Security

Participant data will be obtained through both hard-copy questionnaires and electronic forms, then entered into a password-protected database accessible only to authorized research staff. Double data entry and range checks will be performed for key variables to minimize errors. All electronic data will be stored on secure institutional servers with regular backup. Identifiers (names, phone numbers, and identification number) will be stored separately from research data and linked only through a unique study identification code. The *DemensiaKITA* app does not collect or transmit any identifiable health information; it functions as an information and support tool installed locally on participants' smartphones.

Confidentiality

All personal information obtained from participants will be treated as confidential and used solely for research purposes. Identifiable data will not be included in publications or presentations. Hard-copy documents will be stored in locked cabinets within restricted-access offices. Electronic data will be deidentified for analysis, and only the principal investigator and authorized team members will have access to the code key. Research procedures will comply with the Malaysian Personal Data Protection Act and institutional data protection policies.

Data Sharing and Availability

Deidentified individual participant data underlying the main trial results may be made available upon reasonable request to the corresponding author after publication of the primary

XSL•FO
RenderX

findings, subject to institutional data-sharing policies and approval by the ethics committees.

Dissemination Policy

The results of this trial will be disseminated through peer-reviewed publications, conference presentations, and updates to the ClinicalTrials.gov registry record. A lay summary of the main findings will be prepared in Bahasa Malaysia and made available to participating clinics and interested caregivers.

Results

Phase 1: Questionnaire Adaptation, Translation, and Validation

This phase focuses on adapting, translating, and validating the questionnaires on KAP and burden level. The following outcomes are anticipated.

Content Validation

Content validation will be assessed using the CVI and content validity ratio. To ensure rigorous content validation, an item-level CVI of 0.79 or higher is required for item retention, while a scale-level CVI of 0.80 or higher is necessary to establish excellent content validity for the overall scale. A minimum content validity ratio of 0.99 is established based on expert consultation. This rigorous approach ensures that only the most relevant and important items are retained for subsequent phases of instrument development [29-31,34].

Face Validation

Face validation will be assessed by calculating the face validity index (FVI) to quantify the degree of agreement among respondents regarding the clarity and comprehensibility of the questionnaire items. A sample of 30 respondents was used for this process. Items with an item-level FVI and scale-level FVI of 0.80 or higher were retained. This threshold aligns with recommendations for establishing strong face validity [33-35].

Construct Validation

Construct validation will be conducted using EFA to identify the underlying factor structure of the questionnaire. This will confirm that the items are effectively measuring the intended concepts. Additionally, correlational analysis will examine the relationships between different scales within the questionnaire to assess convergent and discriminant validity [34,36,45].

Reliability

To ensure the questionnaire's consistency, 2 key aspects of reliability will be evaluated. First, internal consistency will be assessed using Cronbach α, which measures how well the items within the questionnaire work together to measure the same concept. A Cronbach α of .70 or higher will be considered acceptable. Second, test-retest reliability will be assessed using the intraclass correlation coefficient. This determines the stability of the questionnaire's results over time, with an acceptable value of 0.75 or higher [34,39,41]. This 2-pronged approach ensures the questionnaire produces reliable and consistent results.

Phase 2: Intervention Effectiveness and Data Analysis

This phase focuses on evaluating the effectiveness of the *DemensiaKITA* mHealth app compared to the usual care.

Inferential and Longitudinal Analysis

All analyses will be conducted using a 2-sided significance level of α=.05. Descriptive statistics will summarize sociodemographic and caregiving characteristics by study arm [46]. Categorical variables will be presented as frequencies and percentages, and continuous variables as means and SDs or medians and IQRs, as appropriate. Baseline comparability between the intervention and control groups will be examined using chi-square tests for categorical variables and independent *t* tests (or nonparametric equivalents) for continuous variables [46,47].

The primary effectiveness analysis will use generalized estimating equations models to compare changes in knowledge, attitude, practice, and burden scores between the intervention and control groups over time (baseline, 1 mo, and 3 mo). For each outcome, a GEE model with an appropriate link function and working correlation structure will be fitted, including fixed effects for group (intervention vs control), time, and the group-by-time interaction. The group-by-time interaction term will be interpreted as the main estimate of the intervention effect. To account for potential confounding due to nonrandom allocation, the models will adjust for prespecified baseline covariates (age, sex, and relationship with the people living with dementia, caregiving duration, and baseline score of the corresponding outcome). Robust standard errors will be reported [21,48].

Effect sizes (Cohen *d* for between-group differences in mean change) will be calculated to quantify the magnitude of the intervention effect, in addition to *P* values and 95% CIs [21,48].

Analysis Population, Missing Data, and Sensitivity Analyses

The primary analysis will follow the intention-to-treat principle, in which all participants will be analyzed in the groups to which their hospital was allocated, regardless of their level of app use, exposure to the leaflets, or subsequent adherence. Every effort will be made to obtain outcome data at 1 and 3 months for all enrolled caregivers, including those who discontinue or reduce use of the intervention [49].

Missing outcome data will initially be handled within the GEE framework, which uses all available repeated measurements under a missing-at-random assumption and provides robust standard errors. Patterns and reasons for missing data will be described. If more than a small proportion of outcome data (>20%) is missing, multiple imputation by chained equations will be considered as a sensitivity analysis, with covariates and outcome scores imputed under the missing-at-random assumption. Results from imputed and complete-case analyses will be compared to assess the robustness of the findings [21].

This study was conducted without any specific external financial support or grants from any public, commercial, or not-for-profit funding agencies. Recruitment and baseline data collection are currently underway at the time of manuscript submission (phase

XSL•FO
RenderX

1: 210 participants, phase 2 not yet started at the time of submission). The study findings are expected to be published in December 2026.

Discussion

Principal Findings

Caring for people living with dementia in low-income countries like Malaysia presents significant challenges for caregivers [2,8]. They often experience physical and emotional exhaustion from providing constant care, which can also lead to social isolation as they have less time for other relationships [1,4]. Additionally, dementia care can create a financial burden, especially with limited support services and financial aid available within the health care system [3,6]. This leaves caregivers shouldering the primary responsibility for their loved ones' well-being, often with inadequate resources.

Furthermore, research highlights a lack of knowledge and the prevalence of negative attitudes among dementia caregivers, hindering their ability to provide optimal care [10]. To address these challenges, exploring the potential of mHealth apps to enhance caregiver knowledge, foster positive attitudes, and improve care practices is crucial [3,11]. This exploration holds significant implications for advancing dementia care and support for both people living with dementia and their caregivers [2,5,12,13,19].

Phase 1 of this study anticipates the successful adaptation and validation of questionnaires on KAP and burden levels among dementia caregivers in the Malay language. This adaptation will ensure cultural relevance and accessibility for the Malaysian population. While in Phase 2, the study aims to evaluate the effectiveness of the *DemensiaKITA* mHealth app. This study will rigorously evaluate the *DemensiaKITA* mHealth app to determine its efficacy in enhancing knowledge, fostering positive attitudes, and improving care practices among dementia caregivers in Malaysia. Ultimately, this evaluation aims to demonstrate the app's potential to alleviate caregivers' burden within the Malaysian context.

This study holds significant implications for advancing dementia care in Malaysia by providing culturally relevant and validated questionnaires for assessing KAP and burden levels among dementia caregivers. Furthermore, the rigorous evaluation of the *DemensiaKITA* mHealth app will generate evidence-based decision-making regarding its implementation and potential to support dementia caregivers in Malaysia. The study's findings are expected to contribute to the development of effective strategies for improving KAP and reducing caregiver burden, ultimately enhancing the health and well-being of both people living with dementia and their caregivers in Malaysia and other countries.

In addition, the findings of this study have the potential to guide national policies in Malaysia by providing evidence for the effectiveness of mHealth apps in dementia care. This could lead to the integration of the *DemensiaKITA* app or similar mHealth solutions into national dementia care strategies. Additionally, the findings could inform resource allocation for caregiver support, emphasizing the need for accessible and culturally relevant resources.

In clinical practice, health care providers, particularly those working in geriatric care, could incorporate the *DemensiaKITA* app into routine care by recommending it to caregivers during consultations. This would provide caregivers with easy access to information and support, potentially improving their KAP in dementia care. Effective dementia care can help prevent costly complications, reduce hospitalizations, and delay the need for long-term care. This can lead to significant cost savings for the health care system.

In terms of community impact, the availability and promotion of the *DemensiaKITA* app could have a positive impact on the community by increasing awareness and support for dementia caregivers. This could lead to a greater understanding and acceptance of dementia, fostering a more dementia-friendly society in Malaysia. By supporting caregivers and improving their well-being, the app can enable them to remain in the workforce or continue with their daily activities, contributing to economic productivity.

It is anticipated that there will be several challenges while conducting this study. Recruiting and retaining participants for both phases of the study, particularly in Phase 2, may be challenging, especially considering the sensitive nature of dementia caregiving and the time commitment required for follow-up assessments. Moreover, adapting and validating the questionnaires into the Malay language requires careful consideration of cultural nuances to ensure the accuracy and relevance of the instruments for the Malaysian population. Additionally, the effectiveness of the *DemensiaKITA* app in improving caregiver KAP and reducing burden levels may be influenced by factors such as individual caregiver characteristics, the severity of dementia in the care recipient, and access to technology and support. Besides that, data collection and analysis can be time-consuming and resource-intensive, requiring careful planning and execution to ensure the accuracy and reliability of the results.

Strengths and Limitations

This study acknowledges several limitations. First, the findings may not be fully generalizable to all dementia caregivers in Malaysia. The research focuses on caregivers attending geriatric clinics in Selangor and Kuala Lumpur, potentially excluding those in rural or underserved areas with differing needs and access to resources.

Second, the nonrandomized controlled trial design, while chosen due to anticipated challenges in recruitment and sample size, may introduce selection bias. This limitation could affect the generalizability of the findings to the broader population of dementia caregivers.

In addition, the app's current availability only on the Android platform poses a significant limitation. This restricts access for a portion of caregivers who use iOS or other operating systems, potentially introducing further bias and limiting the study's reach and impact.

XSL•FO
RenderX

Nevertheless, this study possesses several notable strengths. First, the *DemensiaKITA* mHealth app is uniquely positioned as a culturally and linguistically tailored solution for Malaysian dementia caregivers. This addresses a critical gap in existing mHealth apps, which often cater to English-speaking populations in high-income countries, leaving caregivers in Malaysia with limited access to relevant and accessible support tools. The app's comprehensive design includes features aimed at enhancing knowledge, fostering positive attitudes, improving care practices, and reducing caregiver burden across multiple domains (emotional, physical, social, and financial), offering a holistic approach to dementia care support.

Second, the offline functionality is a major advantage in Malaysia, where internet access can be limited in rural areas. This ensures equitable access to support and resources for caregivers across the country. In addition, the app's free availability removes financial barriers, making it accessible to all caregivers, regardless of their socioeconomic background. This promotes inclusivity and reduces health disparities.

Third, the study uses a robust, nonrandomized controlled trial design with follow-up assessments at baseline, 1, and 3 months, enabling a thorough evaluation of the intervention's effectiveness over time. This longitudinal approach provides valuable insights into the app's impact on caregiver outcomes and its potential for long-term benefits. Furthermore, the study's comprehensive evaluation of knowledge, attitude, practice, and burden provides a holistic understanding of the intervention's impact on caregivers' well-being and their capacity to provide quality care to people living with dementia.

Fourth, by focusing on dementia caregivers in Malaysia, the study addresses a population facing a disproportionately high caregiver burden, reported at 69.4% [8]. This highlights the urgent need for effective interventions to support caregivers in this context and emphasizes the study's potential to make a significant contribution to addressing this vital issue.

In addition, while the nonrandomized design and purposive sampling may introduce limitations, they also enhance the study's practical relevance. By mirroring real-world settings in which interventions are often rolled out to specific clinics or populations, the study offers valuable insights into the feasibility and effectiveness of implementing the *DemensiaKITA* app within the existing health care infrastructure in Malaysia.

Conclusion

This study protocol embarks on a crucial mission to improve the health and well-being of dementia caregivers in Malaysia. First, it focuses on adapting, translating, and validating questionnaires to assess their KAP and burden levels. In addition, this study is pioneering research on the effectiveness of the *DemensiaKITA* mHealth app, designed to support dementia caregivers in Malaysia through a nonrandomized controlled trial at baseline, 1-month follow-up, and 3-month follow-up.

As dementia becomes more prevalent, this app offers a crucial lifeline for caregivers facing increasing caregiving challenges. It is culturally tailored for Malaysia and works offline, making it relevant to countries with limited internet access. This research will provide valuable insights into the ways in which this technology can improve dementia care globally. Furthermore, the study will generate culturally relevant and validated questionnaires for researchers and health care professionals to accurately assess KAP and burden levels among dementia caregivers. These data can inform the development of targeted interventions and support programs.

Moreover, the findings of this study can contribute to evidence-based decision-making regarding the implementation of mHealth interventions in dementia care, paving the way for wider adoption and integration into national health care strategies. By enhancing caregiver KAP and reducing their burden, the *DemensiaKITA* app can empower caregivers to provide better care, cope with challenges, and maintain their well-being, leading to improved outcomes for people living with dementia and their caregivers.

In conclusion, this study protocol represents a significant step forward in dementia care research and practice in Malaysia. A rigorous and culturally sensitive approach promises valuable insights into effective strategies for supporting dementia caregivers and improving the lives of those affected by this debilitating disease. This research is a testament to the power of the mHealth app to address pressing health care challenges and promote the well-being of vulnerable populations.

Acknowledgments

The authors extend their deepest appreciation to Dr Rizah Mazzuin binti Razali, Dr Kejal a/p Hasmukharay, Dr Chee Sing Hui, Dr Wan Mazlina binti Wan Mustapha, Dr Aimy binti Abdullah, and Pn Saadah binti Hassan, as well as the dedicated staff of the Geriatric Clinics at Hospital Kuala Lumpur, Pusat Perubatan Universiti Malaya, Hospital Sungai Buloh, Hospital Selayang, Hospital Tengku Ampuan Rahimah Klang, Hospital Al-Sultan Abdullah UiTM Puncak Alam, and Persatuan Usiamas, for their invaluable assistance during the recruitment phase.

They are also grateful for the institutional support provided by the Faculty of Medicine, Universiti Teknologi MARA, Sungai Buloh Campus, and the Faculty of Medicine, Universiti Malaya (UM).

Their sincere thanks go to Dr Nurul Syaireen binti A Rashid, Dr Raudah binti Mohd Yunus, and Dr Muhammad Fadhil bin Mohamad Marzuki for their support and contributions throughout the study.

Generative artificial intelligence tools (ChatGPT and Gemini) were used to assist with language editing of this manuscript.

XSL•FO
RenderX

Funding

This study was conducted without any specific external financial support or grants from any public, commercial, or not-for-profit funding agencies.

Data Availability

The full study protocol and statistical analysis plan will be accessible via the trial registration record and as supplementary material to the published article.

Conflicts of Interest

None declared.

References

1. Clinical practice guidelines. Management of dementia. Third edition. : Malaysian Health Technology Assessment Section (MaHTAS) Medical Development Division, Ministry of Health Malaysia; 2021 URL: https://mymahtas2.moh.gov.my/files/e-CPG_Management_of_Dementia_(Third_Edition)-10_11.pdf [accessed 2026-02-10]
2. Dementia. World Health Organization. 2023. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2026-01-24]
3. Hafdi M, Eggink E, Hoevenaar-Blom MP, et al. Design and development of a mobile health (mHealth) platform for dementia prevention in the prevention of dementia by mobile phone applications (PRODEMOS) project. Front Neurol 2021;12:733878. [doi: 10.3389/fneur.2021.733878] [Medline: 34975710]
4. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues Clin Neurosci 2009;11(2):217-228. [doi: 10.31887/DCNS.2009.11.2/hbrodaty] [Medline: 19585957]
5. Ye B, How TV, Chu CH, Mihailidis A. Dementia care apps for people with dementia and informal caregivers: a systematic review protocol. Gerontology 2021;67(5):633-638. [doi: 10.1159/000514838] [Medline: 33774646]
6. Cheng ST. Dementia caregiver burden: a research update and critical analysis. Curr Psychiatry Rep 2017 Aug 10;19(9):64. [doi: 10.1007/s11920-017-0818-2] [Medline: 28795386]
7. Collins RN, Kishita N. Prevalence of depression and burden among informal care-givers of people with dementia: a meta-analysis. Ageing Soc 2020 Nov;40(11):2355-2392. [doi: 10.1017/S0144686X19000527]
8. Nasreen HE, Tyrrell M, Vikström S, et al. Caregiver burden, mental health, quality of life and self-efficacy of family caregivers of persons with dementia in Malaysia: baseline results of a psychoeducational intervention study. BMC Geriatr 2024 Aug 5;24(1):656. [doi: 10.1186/s12877-024-05221-9] [Medline: 39103767]
9. Connors MH, Seeher K, Teixeira-Pinto A, Woodward M, Ames D, Brodaty H. Dementia and caregiver burden: a three-year longitudinal study. Int J Geriatr Psychiatry 2020 Feb;35(2):250-258. [doi: 10.1002/gps.5244] [Medline: 31821606]
10. Guilin L, Qingran L, Ruiying H, Cuiqiu Z, Yunjia L. An investigation of the knowledge, attitude and practice of family caregivers of patients with Alzheimer's disease. Int J Psychol Brain Sci 2019;4(1):1. [doi: 10.11648/j.ijpbs.20190401.11]
11. Su HF, Koo M, Lee WL, Sung HC, Lee RP, Liu WI. A dementia care training using mobile e-learning with mentoring support for home care workers: a controlled study. BMC Geriatr 2021 Feb 16;21(1):126. [doi: 10.1186/s12877-021-02075-3] [Medline: 33593287]
12. Rashid NSA, Chen XW, Mohamad Marzuki MF, et al. Development and usability assessment of a mobile app (DemensiaKITA) to support dementia caregivers in Malaysia: a study protocol. Int J Environ Res Public Health 2022 Sep 20;19(19):11880. [doi: 10.3390/ijerph191911880] [Medline: 36231181]
13. Ye B, Chu CH, Bayat S, Babineau J, How TV, Mihailidis A. Researched apps used in dementia care for people living with dementia and their informal caregivers: systematic review on app features, security, and usability. J Med Internet Res 2023 Oct 12;25:e46188. [doi: 10.2196/46188] [Medline: 37824187]
14. Yousaf K, Mehmood Z, Saba T, et al. Mobile-health applications for the efficient delivery of health care facility to people with dementia (PwD) and support to their carers: a survey. Biomed Res Int 2019;2019:7151475. [doi: 10.1155/2019/7151475] [Medline: 31032361]
15. Brown EL, Ruggiano N, Page TF, et al. Apps for dementia caregivers: a feasibility study. Res Gerontol Nurs 2016;9(4):193-203. [doi: 10.3928/19404921-20160229-02] [Medline: 29977440]
16. Chelberg GR, Neuhaus M, Mothershaw A, Mahoney R, Caffery LJ. Mobile apps for dementia awareness, support, and prevention - review and evaluation. Disabil Rehabil 2022 Aug;44(17):4909-4920. [doi: 10.1080/09638288.2021.1914755] [Medline: 34034601]
17. Guo Y, Yang F, Hu F, Li W, Ruggiano N, Lee HY. Existing mobile phone apps for self-care management of people with Alzheimer disease and related dementias: systematic analysis. JMIR Aging 2020 Jan 24;3(1):e15290. [doi: 10.2196/15290] [Medline: 32012045]
18. Rathnayake S, Moyle W, Jones C, Calleja P. mHealth applications as an educational and supportive resource for family carers of people with dementia: an integrative review. Dementia (London) 2019;18(7-8):3091-3112. [doi: 10.1177/1471301218768903] [Medline: 29631492]

XSL•FO
RenderX

19. Wang AH, Newman K, Martin LS, Lapum J. Beyond instrumental support: mobile application use by family caregivers of persons living with dementia. Dementia (London) 2022 Jul;21(5):1488-1510. [doi: 10.1177/14713012211073440] [Medline: 35414298]
20. Harris AD, McGregor JC, Perencevich EN, et al. The use and interpretation of quasi-experimental studies in medical informatics. J Am Med Inform Assoc 2006;13(1):16-23. [doi: 10.1197/jamia.M1749] [Medline: 16221933]
21. Hardin JW, Hilbe JM. Generalized Estimating Equations, 2nd edition: CRC Press; 2018.
22. Toye C, Lester L, Popescu A, McInerney F, Andrews S, Robinson AL. Dementia knowledge assessment tool version two: development of a tool to inform preparation for care planning and delivery in families and care staff. Dementia (London) 2014 Mar 1;13(2):248-256. [doi: 10.1177/1471301212471960] [Medline: 24339059]
23. O'Connor ML, McFadden SH. Development and psychometric validation of the Dementia Attitudes Scale. Int J Alzheimers Dis 2010;2010:454218. [doi: 10.4061/2010/454218]
24. Clark NM, Rakowski W. Family caregivers of older adults: improving helping skills. Gerontologist 1983 Dec;23(6):637-642. [doi: 10.1093/geront/23.6.637] [Medline: 6662379]
25. Lee RLT, Mok ESB. Evaluation of the psychometric properties of a modified Chinese version of the Caregiver Task Inventory-refinement and psychometric testing of the Chinese Caregiver Task Inventory: a confirmatory factor analysis. J Clin Nurs 2011 Dec;20(23-24):3452-3462. [doi: 10.1111/j.1365-2702.2011.03729.x] [Medline: 21707805]
26. Bédard M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001 Oct;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
27. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. Gerontologist 1980 Dec;20(6):649-655. [doi: 10.1093/geront/20.6.649] [Medline: 7203086]
28. Zarit SH, Zarit JM. The Zarit Burden Interview--short version (ZBI-12). In: ICON: The Index of Cognitive and Neuropsychological Assessments, 2nd edition: Oxford University Press; 2008, Vol. 41:1-2.
29. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 2006 Oct;29(5):489-497. [doi: 10.1002/nur.20147] [Medline: 16977646]
30. Polit DF, Beck CT. Nursing Research: Generating and Assessing Evidence for Nursing Practice, 10th edition: Lippincott Williams & Wilkins; 2017.
31. Yusoff MSB. ABC of content validation and content validity index calculation. Educ Med J 2019;11(2):49-54. [doi: 10.21315/eimj2019.11.2.6]
32. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007 Jan;60(1):34-42. [doi: 10.1016/j.jclinepi.2006.03.012] [Medline: 17161752]
33. DeVon HA, Block ME, Moyle-Wright P, et al. A psychometric toolbox for testing validity and reliability. J Nurs Scholarsh 2007;39(2):155-164. [doi: 10.1111/j.1547-5069.2007.00161.x] [Medline: 17535316]
34. Rejab N, Muhammad NA, Tohid H, Tohit NM, Kin PW, Drahman I. A Malay version of the attitude to ageing questionnaire: its adaptation, validation, and reliability in the Malaysian older adult population. Ann Geriatr Med Res 2022 Sep;26(3):225-232. [doi: 10.4235/agmr.22.0023] [Medline: 35811114]
35. Yusoff MSB. ABC of response process validation and face validity index calculation. Edu 2019;11(3):55-61. [doi: 10.21315/eimj2019.11.3.6]
36. Clark LA, Watson D. Constructing validity: basic issues in objective scale development. Psychol Assess 1995 Sep;7(3):309-319. [doi: 10.1037/1040-3590.7.3.309]
37. Kyriazos TA. Applied psychometrics: sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. Psych 2018;09(8):2207-2230. [doi: 10.4236/psych.2018.98126]
38. Hajjar ST. Statistical analysis: internal-consistency reliability and construct validity. Int J Quant Qualit Res Methods 2018;6(1):46-57 [FREE Full text]
39. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
40. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use, 5th edition: Oxford University Press; 2015. [doi: 10.1093/med/9780199685219.001.0001]
41. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
42. Dementia. Putrajaya: Bahagian Pendidikan Kesihatan, Kementerian Kesihatan Malaysia. 2021. URL: https://infosihat.moh.gov.my/penerbitan-multimedia/risalah/item/dementia.html [accessed 2026-02-20]
43. Rosner B. Fundamentals of Biostatistics, 8th edition: Cengage Learning; 2015.
44. Turnbull N, Sriruecha C, Kongmant R, Nghiep LK, Tudpor K. A mobile application for enhancing caregiver support and resource management for long-term dependent individuals in rural areas. Healthcare (Basel) 2024 Jul 24;12(15):1473. [doi: 10.3390/healthcare12151473] [Medline: 39120176]
45. Costello AB, Osborne JW. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. Pract Assess Res Eval 2005;10(7):1-9. [doi: 10.7275/JYJ1-4868]
46. Moore DS, McCabe GP, Craig BA. Introduction to the Practice of Statistics, 9th edition: W. H. Freeman; 2017.

XSL•FO
RenderX

47. Pallant J. SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS, 7th edition: New York: Routledge; 2020. [doi: 10.4324/9781003117452]
48. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 2012 Feb;141(1):2-18. [doi: 10.1037/a0024338] [Medline: 21823805]
49. Gupta SK. Intention-to-treat concept: a review. Perspect Clin Res 2011 Jul;2(3):109-112. [doi: 10.4103/2229-3485.83221] [Medline: 21897887]

Abbreviations

CTI-25: Caregiver Task Inventory—25 items
CVI: content validity index
DAS: Dementia Attitudes Scale
DKAT-2: Dementia Knowledge Assessment Tool Version 2
EFA: exploratory factor analysis
FVI: face validity index
GEE: generalized estimating equations
KAP: knowledge, attitudes, and practices
mHealth: mobile health
MOH: Ministry of Health
NRCT: nonrandomized controlled trial
UiTM: Universiti Teknologi MARA
UMMC: Universiti Malaya Medical Center
ZBI-12: Short-form Zarit Burden Interview—12 items

Edited by J Sarvestan; submitted 07.Sep.2025; peer-reviewed by V Dolničar; accepted 12.Dec.2025; published 25.Feb.2026.

Please cite as:
Mahfar N, Abdullah NN, Roslan D, Ibrahim K, Ungku Mohd Zam UAA, Noordin Merican N, Chen XW, Mohamad M
The Effect of the DemensiaKITA Mobile Health App Intervention on Knowledge, Attitude, Practice, and Burden Level of Dementia Caregivers in Kuala Lumpur and Selangor, Malaysia: Protocol for a Nonrandomized Controlled Trial
JMIR Res Protoc 2026;15:e83699
URL: https://www.researchprotocols.org/2026/1/e83699
doi:10.2196/83699

© Nurulizzah Mahfar, Nik Nairan Abdullah, Dalila Roslan, Khalid Ibrahim, Ungku Ahmad Ameen Ungku Mohd Zam, Noraliza Noordin Merican, Xin Wee Chen, Mariam Mohamad. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Acoustic Emission Biomarkers for the Detection and Monitoring of Early Knee Osteoarthritis: Protocol for a Prospective, Single-Center, Exploratory Study

Louis Leuthard[1], MSc; Jérôme Thevenot[2], PhD; Tomás Teijeiro[3], PhD; David Atienza[2], PhD; Vincent A Stadelmann[1], PhD

[1]Department of Research and Development, Schulthess Klinik, Lengghalde 2, Zurich, Switzerland

[2]Embedded Systems Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[3]Basque Center for Applied Mathematics, Bilbao, Spain

Corresponding Author:
Vincent A Stadelmann, PhD
Department of Research and Development, Schulthess Klinik, Lengghalde 2, Zurich, Switzerland

Abstract

Background: Osteoarthritis is a highly prevalent and disabling condition. In early stages, patients are asymptomatic or only experience activity-related pain. When pain intensifies, the disease has often progressed, with few treatment options besides knee arthroplasty. Recently, there has been a growing interest in acoustic emissions generated by poorly lubricated and/or damaged moving joint surfaces, as seen in osteoarthritis. Noninvasive analysis of knee sound could help identify patients with early-stage osteoarthritis at a low cost, without radiation exposure. Thus, preventive measures could be implemented earlier and delay the progression of osteoarthritis.

Objective: This study aims to identify acoustic biomarkers with prognostic value for the early detection of osteoarthritis. To achieve this, a preliminary reference database will be developed using a subpopulation at risk of developing osteoarthritis, incorporating acoustic signals and gold-standard clinical measures for validation.

Methods: A total of 100 patients with previous reconstructive knee surgery and 20 healthy controls will be examined twice at a 9-month interval for both knees. Acoustic emissions (AE) will be recorded with the Inmodi knee brace (École Polytechnique Fédérale de Lausanne; EPFL) during 4 functional tests (unloaded flexion-extension, sit-to-stand test, one-step test, and walk test). Gold-standard osteoarthritis diagnostics will be assessed and evaluated by an experienced radiologist for (1) the magnetic resonance imaging (MRI) Osteoarthritis Knee Score (MOAKS) and (2) the Kellgren-Lawrence (KL) grading scores. Feature analysis and multivariate modeling will be used to study the association between the extracted sound parameters and osteoarthritis (MOAKS and KL). In addition, thermal images will be taken to investigate the state of inflammation. Participants will fill out 5 health-related questionnaires (Oxford Knee Score [OKS], Core Outcome Measures Index [COMI], University of California at Los Angeles [UCLA]activity scale, Knee Injury and Osteoarthritis Outcome Score [KOOS], and Forgotten Joint Knee Score-12 [FJS-12]) at both time points.

Results: Recruitment and data collection started in December 2022. In November 2023, data collection of all 120 participants (female: n=49, 41%) was completed for the first visit. The data collection for the second visit will end in 2024. Preliminary data processing and analysis are ongoing at the time of this writing.

Conclusions: This exploratory study will contribute to a better understanding of the disease progression of osteoarthritis and the use of acoustic biomarkers for predicting osteoarthritis.

Trial Registration: ClinicalTrials.gov NCT06351059; https://clinicaltrials.gov/study/NCT06351059

International Registered Report Identifier (IRRID): DERR1-10.2196/67032

(JMIR Res Protoc 2026;15:e67032) doi:10.2196/67032

KEYWORDS

joint health; multimodal; joint sound analysis; prognosis; explorative study; feature extraction; disease progression

XSL•FO
RenderX

Introduction

Disease Background

Osteoarthritis is a highly prevalent and disabling condition that affects over 7% of people globally (528 million people) [1]. It is significantly limiting their mobility and independent lifestyle. Osteoarthritis is mainly characterized by loss of cartilage, structural changes in bone, and inflammation of the synovium and joint capsule. Common risk factors include aging, obesity, prior joint injury, and repetitive joint stress increasing local joint wear. Osteoarthritis takes several years to develop before the patient sees the doctor when the pain intensifies. In the early stages, patients are asymptomatic or only experience activity-related pain. The pain becomes constant over time with intermittent intense pain episodes. The average annual cost of osteoarthritis for an individual is estimated to be between US $700-$15,600 worldwide (2019) [1]. Consequently, osteoarthritis has a large socioeconomic impact due to its associated high medical cost expenditures, as well as indirect costs, such as early retirement and high absenteeism from work.

Osteoarthritis typically affects the hips, knees, hands, feet, and spine, with a high prevalence of polyarticular involvement. This study focuses on knee osteoarthritis.

Contemporary Standard of Assessment

The contemporary gold-standard assessment of the severity of knee osteoarthritis is based on plain radiography, using the Kellgren-Lawrence (KL) grading system, which was accepted as a standard by the World Health Organization in 1961 [2]. This system does not evaluate the primary affected tissue—cartilage—but only the overall aspect of the joint. It is, therefore, unable to detect early stages of osteoarthritis, and the diagnosis is often established at a late stage, limiting the number of treatment options and typically leading to knee replacement surgery.

Alternative Assessment Methods

Earlier diagnosis and improved patient stratification are required to deploy preventive osteoarthritis treatments. Possible early diagnostic options include magnetic resonance imaging (MRI) and biochemical markers. MRI can be used to assess changes in cartilage volume, thickness, or even quality [3]. However, the use of MRI is limited by its high cost, availability, and the absence of a validated international score [4]. Alternative clinical imaging modalities have also been proposed to address the shortcomings of KL at the time of writing this, such as ultrasound [5] or computed tomography [6]; however, their use remains relatively limited compared to X-rays. Many biochemical markers are under investigation, but so far, they lack specificity to joint tissues and to particular joints, and sensitivity and specificity are still not high enough for widespread clinical use [7]. Recently, there has been a growing interest in acoustic analysis as an alternative to conventional biomarkers, particularly for the knee joint. The assumption is that the affected joint will generate abnormal acoustic events during motion due to poor lubrication or surface damage and that acoustic analysis will detect and quantify these abnormal events as a measure of joint damage [8-17].

Relevance of the Project

A wearable device was developed at the École Polytechnique Fédérale de Lausanne (EPFL) for collecting and processing acoustic and kinematic signals. This device—called the Inmodi knee brace—collects airborne acoustic data from the medial and lateral subpatellar areas and kinematic information from the thigh and the calf, both modalities synchronously [17,18]. The platform is then complemented by thermal information of the knee as a measure of inflammation and uses artificial intelligence–based data analysis to extract and combine relevant biomarkers of joint function from each modality. Preliminary data suggests that this combined approach may enable a noninvasive early diagnosis of osteoarthritis. A preliminary *in vitro* animal study with horses' legs showed a strong association between the joint condition and the power of acoustic emission (AE) signals analyzed [19]. Recently, we tested a prototype of the Inmodi knee brace in 17 patients from our clinic shortly before total knee arthroplasty (TKA). These patients endured severe osteoarthritis on the leg undergoing surgery and mild to moderate on the contralateral one. A machine learning (ML) algorithm was able to discriminate the leg states with a specificity of 0.96 (perfect=1; data not yet published).

The aim of this pilot project is to further explore the potential of the Inmodi system to diagnose osteoarthritis at earlier stages. Therefore, we will develop a preliminary reference database in a subpopulation at risk of developing osteoarthritis, while following recent recommendations [20]. The identification and validation of novel biomarkers with a ML algorithm require a training dataset and a test dataset, both with potential biomarkers and gold-standard clinical outcomes for validation. This database will require a sample size large enough to be statistically sound to prevent statistical overfitting. Unfortunately, power analysis is not possible for most ML methods, due to the impossibility of guaranteeing statistical assumptions on the distributions of the features and their linear and nonlinear dependencies. That is why we rely on leave-one-subject out nested cross-validation for assessing the statistical relevance of the resulting models. This is the most conservative and robust cross-validation strategy for overfitting. The specific sample size of 120 was considered as feasible within the time constraints of the project.

Noninvasive knee health diagnostics should help identify patients at risk at a relatively low cost and without radiation exposure and enable preventive measures to be implemented earlier than with standard radiographic imaging.

Methods

Project Design

This is a prospective, single-center, exploratory study. A cross-sectional and a longitudinal dataset will be collected. A total of 120 participants will be recruited and examined twice at a 9-month interval at the Schulthess Klinik in Zurich (see Table 1 for the measurement procedure). The cross-sectional dataset stems from the first visit of the participants and serves for the identification of AE biomarkers for osteoarthritis. The longitudinal dataset combines the data from the first and second visits and tests whether the osteoarthritis changes can be assessed from the biomarkers. A secondary aim is to evaluate

XSL•FO
RenderX

the ability of the AE biomarkers from the first visit to predict the condition's progression as observed during the second visit.

Table . Summary table with all project visits, including relevant procedures and timelines.

Study period	Information	Screening	1st visit	2nd visit
Time point	Precall (maximum 2 months)	0	0-2 months	9 months (±2 months)
Oral and written information	✓			
Written consent		✓		
Check inclusion/ exclusion criteria		✓		
Anthropometric and patient data			✓	✓
EOS radiography			✓	✓
MRI[a]			✓	✓
Acoustic emission test with Inmodi knee brace			✓	✓
PROMs[b] (at home)			✓	✓

[a]MRI: magnetic resonance imaging.

[b]PROM: patient reported-outcome measurement.

Project Population

This study targets a population at risk of osteoarthritis. The risk of knee osteoarthritis after knee joint injury is up to 5-fold that of noninjured patients. As many as 50% of individuals with an anterior cruciate ligament (ACL) or meniscus tear develop knee osteoarthritis [21]. Therefore, to represent the whole distribution of osteoarthritis stages in the study sample, participants will be recruited from the clinic's Meniscus/ACL/Cartilage (MAC) surgery registry at 2- and 5-year postoperative times. All patients from the MAC registry will be screened regardless of whether they have undergone isolated meniscal tear repair, isolated ligament reconstruction, or a combined meniscal and ligament procedure. Additionally, a healthy control group and a late-stage osteoarthritis group will also be included. The inclusion and exclusion criteria presented in Textbox 1, follow established standards and are consistent with those used in similar studies (eg, NCT02937064).

The total study sample will be 120 patients in 6 groups:

- Group A: 20 healthy participants
- Group B1: 20 patients from the MAC registry at 2 years postoperative with normal outcomes (normal and poor outcomes are defined in section Recruitment)
- Group B2: 20 patients from the MAC registry at 2 years postoperative with poor outcomes
- Group C1: 20 patients from the MAC registry at 5 years postoperative with normal outcomes
- Group C2: 20 patients from the MAC registry at 5 years postoperative with poor outcomes
- Group D: 20 patients booked for in-house unilateral TKA surgery for severe osteoarthritis

Even if it is not possible to perform an exact power analysis for the ML models to be trained, we can use a simple proxy model to estimate the effect size that we can reasonably expect to detect with a sample size of 120. Thus, if we aim to estimate the differences in KL gradings from the AEs, the smallest detectable correlation with 80% power and a significance level of P=.05 would be 0.25. In one-way ANOVA, this corresponds to a minimum explained variance of approximately 7%, which is considered a small-to-medium effect size [22]. Smaller effect sizes are unlikely to be detected within our population.

XSL•FO
RenderX

Textbox 1. Detailed inclusion and exclusion criteria for the groups.

Inclusion criteria

Group A:

- Aged 35-65 years
- People judge themselves to be of good subjective health without any knee problems

Groups B and C:

- Aged 35-65 years
- Unilateral knee surgery (other side as internal control)
- Participants in Meniscus/ACL/Cartilage (MAC) surgery registry
- Completed patient-reported outcome measurements (PROM) questionnaires at 2 years or 5 years within the last 6 months

Group D:

- Aged 35-75 years
- Patient booked for total knee arthroplasty (TKA) due to severe osteoarthritis

Exclusion criteria

- Inability to give consent or follow procedures
- No understanding of the German language
- Open wounds or tissue injuries
- Irritated or infected sections on the limbs
- Class II obesity is defined by BMI ≥35 kg/m2 (comorbidities associated with obesity should be investigated in future studies)
- Uncooperative patients who disregard or cannot follow instructions, including those who abuse drugs and/or alcohol
- Pregnant or with the intention to get pregnant (X-rays)
- Current address outside of Switzerland
- Groups B and C only: revision surgeries at the operated knee, death, known pathologies, or former injuries of the comparator knee

Recruitment

A subject will be considered enrolled in the study only after providing written consent. No study-specific data will be collected until then. The recruitment procedures for the sample groups are described below:

- For Group A, healthy controls will be recruited via the standard outlets, such as the clinic's website and social media accounts.
- For groups B and C, patients' medical files from the MAC registry of our clinic will be screened for inclusion and exclusion criteria. Based on the known distributions of PROM scores (instrument: Core Outcome Measures Index [COMI] [23]) in the registry at 2 and 5 years, the patients' scores will be categorized as normal/poor scores.
- For groups B1 and C1, lists of patients will be generated and ordered by age and sex. Beginning with each patient nearest to age 35, 38, 41, 44, 47, 50, 53, 56, 59, and 62 years, the following recruitment procedure will occur: if no exclusion reason was found, the patient will be contacted by phone by the study nurse and offered study participation. If the patient refuses or cannot be reached after 3 attempts of contact, the next patient within the same age group will be screened and contacted. This ensures a balanced age distribution within sex within the groups. Lists of patients for groups B2 and C2 (by sex) will be generated and ordered by COMI score (decreasing), and patients will be contacted successively.
- The poor outcomes group B2 consists of patients with scores ≥2.
- The normal outcomes group B1 is defined as patients closely around the median COMI scores of their respective group (determined medians are: 2-year, females, 0.9; 2-year, males, 0.7; 5-year, females, 0.5; 5-year, males, 0.5).
- For Group D, all patients scheduled for unilateral TKA surgery at the site and fulfilling the inclusion criteria will be informed about the study and asked for participation. Screening is done by the study manager as soon as the operation is booked. If the patient is eligible, the study nurse will inform the patient by phone about the study and ask for participation. If the patient is interested, a brochure with information about the study will be sent in written form to the patient to consider participation carefully at home for at least 24 hours. This will be done successively until the sample size of 10 males and 10 females is reached.

Study Procedures

The overall study duration is 2 years. Recruitment will last for 6 months. For a patient, the duration of the study is 9 months,

XSL•FO
RenderX

with one visit at baseline and one visit at 9 months. All measurements will occur at the clinic, including anthropometric data, clinical examination, and functional and temperature assessments. Additionally, the participants should fill in PROMs, which are provided as an online questionnaire via the Research Electronic Data Capture (REDCap; Vanderbilt University) application. If preferred, a paper version of the questionnaire can be provided to the patient. Participants can do this at home.

Anthropometric and Patient Data

The following data will be collected from the patient's medical file: sex, age, height, weight, BMI, body fat percentage, comorbidities, and adverse events since the knee operation.

EOS Radiography

Long-leg radiographic images will be acquired in a standing position as stereoradiography using the EOS Edge (EOS X-Ray Imaging Acquisition System). The photon counting detector allows for a dose reduction by a factor between 8 and 10 compared to conventional X-ray radiography [24,25]. EOS imaging of the knee has been validated for osteoarthritis grading using KL grading scores [26].

For all participants, bilateral long-leg, standing EOS images in anteroposterior view will be retrieved and analyzed for:

- Osteoarthritis grading (KL grade 0 - 4)
- Compartmental joint space width
- Leg alignment angles:
  - Knee joint alignment (normal, varus, and valgus) measured with hip-knee-ankle angle
  - Mechanical axis length
  - Proximal tibia width
  - Mechanical axis deviation
  - Joint angles
  - Mechanical lateral distal femoral angle
  - Mechanical medial proximal tibial angle

MRI Technique

Both knees of the participants will be scanned using a 3 Tesla MRI scanner (Magnetom Prisma; Siemens) without any contrast agent (Table 2). MRI allows visualization and assessment of structural changes in joint tissues, such as cartilage, meniscus, and subchondral bone. A radiologist will use the MRI Osteoarthritis Knee Score (MOAKS) system to score these degenerative structures in osteoarthritis semiquantitatively. The MOAKS score has demonstrated very good to excellent reliability [27] and grades tissues in 7 different subregions of the knee [28].

Table . Acquisition parameters of magnetic resonance imaging (MRI) sequences.

MRI[a] sequence	STIR[b]	T1 TSE[c]	PD[d] dixon	PD TSE FS[e]
Orientation	coronal	coronal	3D [4,5]	3D [6,7]
Field of view (mm)	154	160	160	160
Repetition time (ms)	4850	558	5010	4200
Echo time (ms)	31	7.1	44	42
Number of slices	28	28	34	31
Slice thickness (mm)	3	3	2.5	2.7
Distance factor (%)	20	15	15	10
Voxel area (mm^3)	0.6×0.6×3	0.5×0.5×3	0.4×0.4×2.5	0.4×0.4×2.7
Pixel bandwidth (Hz/Px)	260	340	334	260
Flip angles (deg°)	180	180	180	180

[a]MRI: magnetic resonance imaging.

[b]STIR: short tau inversion recovery.

[c]TSE: turbo spin echo.

[d]PD: proton density.

[e]FS: fat saturation.

AE Test With Inmodi Knee Brace

The tests will be done with the Inmodi knee brace (v2; Figure 1) for AE assessment of both knees simultaneously [18]. The brace is designed to not hinder movement and is made of biocompatible materials (polylactic acid/thermoplastic polyurethane) covered by silicon commonly used in orthopedics for direct skin contact. All the embedded electronics are fully isolated, without possible contact with the patient. The device consists of 2 parts to be attached on the limb with the help of straps (Figure 2):

- The main part is to be positioned on the calf. This part has an embedded inertial measurement unit (IMU) module to collect kinematic data on the lower part of the limb and 2 microphones positioned in an enclosure covered by removable silicone. The microphones are placed at each extremity of a flexible arch to be positioned on the medial and lateral subpatellar regions. This mainframe also has a battery, microcontroller, and embedded flash memory to save the raw data.
- A secondary part to be positioned on the thighs and connected by USB to the main part. This secondary part

XSL•FO
RenderX

has an IMU module to collect the kinematics of the upper part of the limb and an interface with buttons to start/stop the recording of all sensors.

XSL•FO
RenderX

Figure 1. Upper part (positioned on the thigh) and lower part (positioned on the calf) of the Inmodi knee brace.

XSL•FO
RenderX

Figure 2. The Inmodi device from the top (left) and the bottom (right). The green silicone covers can be removed from the microphone frames for cleaning.

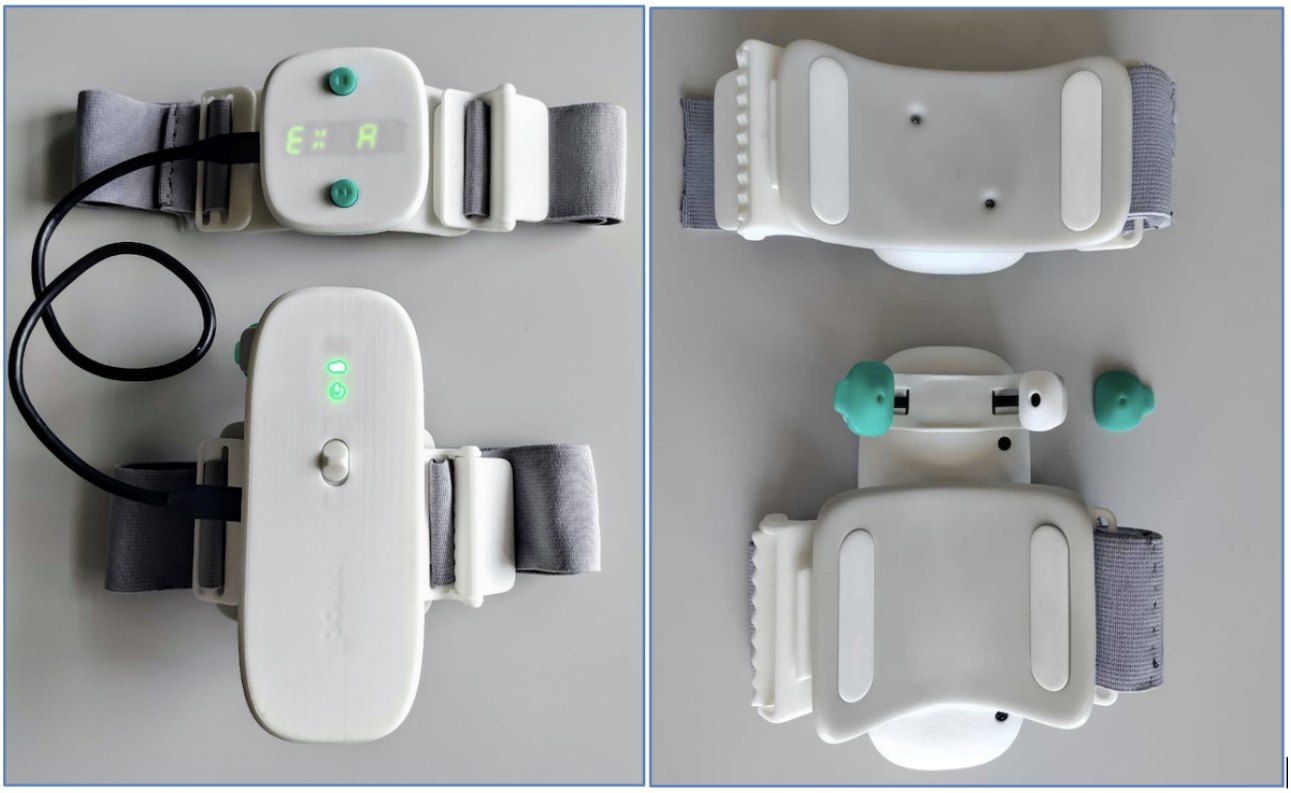

All the sensors collect data in a synchronized manner: the microphones (SPH0645LM4H-B; Knowles) have a sampling rate of 22000 Hz, and the IMU (BNO080; CEVA Hillcrest Labs) has a frequency of 100 Hz. The microphones were chosen to have a sampling rate covering the relevant acoustic information suggested in the literature [9,11,15]. We decided to use airborne microphones to improve the device's usability for easy positioning without needing extra tapes to maintain contact sensors on the skin and potential associated rubbing artifacts from them. The literature has reported that, despite having a lower signal-to-noise ratio, airborne AEs can be collected by Micro-electro-mechanical systems microphones as an alternative to contact sensors to assess knee health [15]. The advantage of using air microphones over contact microphones resides in the clinical applicability of the solution, which requires minimal preparation from the medical workers. However, such an approach requires a design to ensure a consistent distance between the airborne microphone and the skin to keep the microphone's sensitivity to collect knee sounds. The Inmodi device was designed to keep the microphone at a distance of 6.5 mm from the skin while preventing any skin contact with the microphone membrane, thanks to a rigid frame encapsulating the microphone covered by a silicone piece. The silicone piece is the only part in contact with the skin, providing a damping of artifact noise due to possible friction, and it is removable for cleaning between patients.

The tests are based on recommendations by the Osteoarthritis Research Society International (OARSI) for people with knee osteoarthritis [29]:

- Unloaded flexion-extension (F-E) of the knee in a seated position, starting with a knee angle of 90°, followed by a full extension and returning to the starting position (6 times). A video will guide the pace of the motion to normalize the velocity among patients (comfortable speed).
- Sit-to-stand test (STS) at a comfortable pace (6 times). Force plates (FP-BTA; Vernier) will be used during this exercise to assess the ground reaction force for each sole and detect any asymmetrical loading of the joints.
- One-step test (OST) is an exercise step platform, which will be used to perform a complete ascent and descent at a comfortable speed (6 times per leg).
- Walk test at a comfortable pace (10m ×2 times).

The following measurements will be performed together with the AE tests:

- Temperature assessment: thermal distribution of the knee joint will be assessed with a far infrared thermal sensor array (MLX90640; Melexis) to evaluate the state of inflammation. Localized thermal images will be recorded at three time points: (1) at the beginning of the whole test, just after arriving, (2) after warming up (5 minutes on a stationary bicycle), and (3) at the end of all tests.
- Pain assessment: after the test session, the level of knee pain will be assessed using a visual analog scale. Participants will quantify the maximum knee pain experienced during the tests for each knee separately by placing a vertical mark on a 10 cm horizontal line. The line ranges from 0 ("no pain at all") to 10 ("not endurable pain").

PROMs

The following PROMs questionnaires were sent out by email or paper form depending on participants' preference and were completed in their homes.

XSL•FO
RenderX

- Oxford Knee Score (OKS): this instrument contains 12 questions about an individual's activities of daily living and how they have been affected by pain over the last 4 weeks. Its summed score ranges from 0 (severe functional problems) to 48 (excellent knee function) [30,31].
- Core Outcome Measures Index (COMI) knee: assesses the main outcomes of importance to patients with knee problems (pain, function, symptom-specific well-being, quality of life, disability). It is a 6-item multidimensional instrument with 5 domain scores for pain (item 1), function (item 2), symptom-specific well-being (item 3), general quality of life (item 4), and disability (average of items 5 and 6) are averaged to give a COMI score ranging from 0 to 10 (the higher the score, the worse the status) [32].
- University of California at Los Angeles (UCLA): an activity rating scale, which rates the activity level of joint replacement patients globally from 1 to 10. The patient indicates her/his most appropriate activity level (level 1 = "no physical activity, dependent on others"; level 10 = "regular participation in impact sports") [33].
- Knee Injury and Osteoarthritis Outcome Score (KOOS): an instrument to assess the patient's opinion about their knee and associated problems. There are 5 subscales with individual sum scores normalized to 0-100 (0 = "extreme symptoms"; 100 = "no symptoms"): pain, other symptoms, function in daily living, function in sport and recreation, and knee-related quality of life. A total score of all subscales combined has not been validated [34].
- Forgotten Joint Score 12 (FJS-12) is a concise 12-item PROM that evaluates patients' ability to forget their joints daily. Each item is answered within a 5-point Likert scale with the following response options: never (0); almost never (1); seldom (2); sometimes (3); and mostly (4) [35].

Signal Segmentation and Denoising

This involves the isolation of the different movement phases of the acoustic signals, the removal of frequency components of low interest, and the separation of the acoustic signal into different sources (sounds from the knee, ambient sounds, sounds from the device, etc). The segmentation will be based mostly on the kinematic data, using the relative angles between the upper and lower parts of the device to set the beginning and end time of each repetitive movement. For denoising, most previous literature suggests relevant acoustic information is below 20kHz, and typically a bandpass filtering from 100 Hz up to 10 - 16,ooo Hz is applied depending on the study [9,11,15]. We will start from this approach, relying on standard digital filters such as Butterworth filters and complementary transforms such as wavelets. Different cutoff frequencies (3, 6, 8, 10, and 16,000 Hz) will be explored to minimize the amount of data to process and simplify feature extraction. Data separation will be achieved through techniques such as empirical mode decomposition and independent component analysis. To reinforce this separation, a further clustering analysis will be performed with techniques such as Density-based spatial clustering of applications with noise to try to unveil robust causal relations between audio segments and sources.

Feature Extraction

This stage involves the calculation of the outcome parameters that will allow the distinction of different osteoarthritis grades. We will initially rely on state-of-the-art parameters in audio classification, such as spectral analysis (mel-frequency cepstrum coefficients [MFCC]) [36,37] and fractality (the First log-cumulant of the scaling exponent, Global Holder Exponent; paper submitted), average level, peak magnitude (peak amplitude) [38] and click patterns (number of clicks/hits), and number of clicks passing threshold [39].

We will also explore the development of novel features specifically oriented to osteoarthritis classification, such as matched filters of prototype sounds [40].

Feature Analysis and Multivariate Modeling

In this stage we will study the association between the extracted parameters and osteoarthritis (using the KL or MOAKS grading as the dependent variable), both individually for each feature and also by creating linear and nonlinear multivariate models. For this, we will follow a strategy guided by the principles of (1) increasing complexity and (2) clinical explainability. Thus, the initial models we consider are linear discriminant analysis and logistic regression, which would be ideal options in case we have a reduced number of highly informative and discriminative features. In this ideal scenario, the models would be highly interpretable and easily translatable to a clinical description. In case more sophisticated models are required, we prioritize ensembles of decision trees, such as random forest and gradient boosting, due to the possibility of performing advanced feature importance analysis and derivation of statistical decision rules.

All features will be tested for normality by Shapiro-Wilk tests and presented as mean (SD; normally distributed data) or median and range and, if needed, in histograms, box plots, or violin plots (nonnormal data).

Associations between AE parameters and osteoarthritis (KL or MOAKS grading) will be evaluated using Pearson (normally distributed data) or Spearman (nonnormal data) repeated measures correlation coefficients, as well as with logistic regression models. Mixed ANOVA will be used to compare differences in the AE, osteoarthritis grading parameters between groups. Also, we will specifically assess the relevance of potential confounding variables such as BMI, age, activity level, and gender. The significance level will be set to P=.05, and due to the potentially large number of tests, we will apply the Holm post-hoc correction to all P values. For multivariate models, the statistical association will be characterized using leave-one-subject-out nested cross-validation, with a subsequent estimation of a confidence interval for prediction accuracy. This is the most conservative and robust cross-validation strategy for overfitting , and with our sample size (n=120), we will have a good estimation of the model variance. Statistical analyses will be done with R (version 4; R Foundation), Matlab (version 2021b or higher; MathWorks, Inc), and Python (version 3.6 or higher; Python Software Foundation).

XSL•FO
RenderX

Statistics

Since this is an exploratory study, the statistical data analysis will also be mostly exploratory, so multiple techniques and models will be assessed and compared. Overall, the analysis of acoustic signals includes signal segmentation and denoising, feature extraction, feature analysis, and multivariate modeling.

To ensure the reproducibility of the experimental results, we will provide open-source access to the exact feature sets and analysis code for the reported results.

Withdrawal and Discontinuation

The patients can withdraw from the study at any point in time. In the declaration of consent, the patient agrees that the data already collected stays in the study. Patients who have knee surgery or a knee injury within the follow-up time or patients who become otherwise severely ill will not be invited for the second visit. In all cases above, the data will be encrypted and used for evaluation. If the enrollment phase is still ongoing, the withdrawn patient will be replaced with a person of the same sex, comparable age, and study group type (A, B1, B2, C1, C2, and D).

If a participant misses the first visit as well as a substitute appointment, the participant will be withdrawn and replaced. If a participant misses the second visit and substitute appointment, his data collected from the first visit will be used.

If a person becomes pregnant, the study will continue but without EOS radiography.

Handling of Missing Data

Patients who fail to appear for the MRI, EOS, or AE test with the Inmodi knee brace will be invited again. After the second miss, they will be excluded (and replaced if the recruitment phase is still ongoing). Missing source data from MRI or EOS are not expected. In the case of missing readouts in the electronic case report form, the radiologist is asked to do the missing readout or to give reasons for the missing.

In case of unexpected data loss or data damage with the Inmodi brace data collection, the participant will be reinvited.

Concerning PROMs, a maximum of 2 missing items is allowed in OKS. They will be imputed with the mean of all other responses. In case of additional missing items, the patient will be contacted to clarify. For COMI and UCLA, missing data will not be imputed. The patient will be contacted for clarification. For KOOS: If more than 50% of subscale items are missing, the subscale score cannot be calculated. Otherwise, the sum of the items answered is normalized to 0 - 100. For FJS-12, if more than 4 of the answers are missing, the total score cannot be calculated. With fewer items missing, these can be replaced by the mean of the items answered. Generally, if a patient makes more than one entry for a question, the worst applies.

Ethical Considerations

An ethical permit was granted for this study by the Business Administration System for Ethical Committees (BASEC No.: 2022 - 01699) from the ethics committee of Canton Zurich, and written informed consent was obtained from all subjects after the nature and possible consequences of the studies were explained. The participants did not receive any compensation. All study data were stored in a password-protected REDCap database where participants are only identified by a unique identifier. A key list is kept secure by the project leader. An integrated audit trail system maintains a record of initial entries and changes (reason for change, date and time of change, and user identification). All procedures performed in this study were developed internally to suit our specific clinical context and have been validated through prior successful implementation [41] and are in line with the 1964 Helsinki Declaration and its amendments. This study was registered at Clinicaltrials.gov with study identifier NCT06351059.

Results

Recruitment and data collection started in December 2022. In November 2023, data collection has been completed for the first visit. The data collection for the second visit will end in 2024. Data processing and analysis are ongoing at the time of this writing.

Baseline data of study participants are shown in Table 3. The median age for the healthy controls (A) is slightly lower, and for the TKA group (D) slightly higher compared to the 2- and 5-year postoperative patients (B and C). Older patients in the TKA group (D) are reasonable, as TKA surgery is often the last option left for treatment. We do observe higher values for BMI in the poor postoperative outcome groups (B2 and C2) and the TKA group (D). Balanced sex distribution was not achieved in 2- and 5-year postoperative groups due to the low willingness to participate as well as the low numbers of female patients within our registry meeting all the inclusion criteria.

Table . Baseline data of study participants.

Characteristic	A (n=20)	B1 (n=20)	B2 (n=20)	C1 (n=20)	C2 (n=20)	D (n=20)
Age, mean (SD)	46 (7)	50 (8)	54 (8)	52 (7)	52 (7)	67 (7)
BMI, mean (SD)	24.25 (2.82)	24.75 (2.75)	26.05 (2.44)	24.65 (2.31)	26.00 (3.71)	26.35 (3.99)
Sex, n (%)						
Female	10 (50)	9 (45)	8 (40)	6 (30)	6 (30)	10 (50)
Male	10 (50)	11 (55)	12 (60)	14 (70)	14 (70)	10 (50)

XSL•FO
RenderX

Discussion

Principal Findings

This exploratory study aims to identify biomarkers with a prognostic value in the prediction of osteoarthritis. Acoustic, kinematic, and thermal features will be extracted and combined with an artificial intelligence algorithm, and evaluated against gold-standard osteoarthritis diagnosis, such as radiographs and MRI. Some candidate acoustic biomarkers have been identified in a preliminary study [9], using a multifractal analysis method. However, multiple other candidates will be evaluated individually or in combination in our study. Therefore, the primary endpoint can be defined as the area under the curve of the receiver operating characteristic analysis of each biomarker or combination of biomarkers in the regression analysis, finally identifying the parameters with the highest sensitivity and specificity to predict early-stage osteoarthritis in clinical use. AE is patient-dependent, and as such, demographic and baseline health data will be captured and entered as covariates in the regression models. The secondary objective is to evaluate the value of these biomarkers in predicting osteoarthritis evolution. For this objective, the patients will be examined a second time at 9-month intervals, and the new biomarkers will be introduced in regression models to evaluate the predicted performance for osteoarthritis evolution.

Limitations

One major limitation of this study is the sample size and distribution. Although many patients participate in our institutional MAC registry, the inclusion criteria narrow down the eligible patients. Furthermore, the willingness of participants was lower than expected, which influenced the distribution of age and sex within the different groups. Whether these variations in baseline characteristics do influence the identification of acoustic biomarkers with a prognostic value for early-stage osteoarthritis, remains unclear until the data analysis has been completed. Still, the sample size is significantly larger than most comparable studies [20] and in line with the most relevant references in the state-of-the-art [16].

A second limitation of the study is the recruitment of patients postsurgery reconstruction. We decided to choose this patient population, as they have a higher predisposition to develop osteoarthritis; however, it might limit the generalization of the approach to highly active individuals who are more prone to injury. Furthermore, while having a balanced age and BMI between groups reduces the impact of covariates, it also limits the generalization of the approach to populations with different osteoarthritis phenotypes. To further enhance the approach's potential, this pilot study should be applied to other centers for data collection to expand the sample size for validation with other osteoarthritis populations with different backgrounds.

Another concern is that resulting acoustic biomarkers may mostly reflect movement patterns rather than a causal relation to osteoarthritis. This critical aspect has been carefully considered in the design of the experimental protocol. As described in the study procedures section, the exercises (movements) the participants must perform are extremely standardized. Also, we will have the kinematic data associated with any potentially discriminative AEs. This allows us to study any movement patterns that may be associated with acoustic biomarkers and assess how strongly they are correlated with specific movements that may not be specific for osteoarthritis. Also, on top of the IMU in the device, sit-to-stand is performed on force plates, so an imbalanced load could also be accounted for in the models. On the other hand, we believe that finding movement patterns that are strongly correlated with the osteoarthritis gradings would be a very relevant contribution of the study, even if causal links cannot be directly established.

Finally, another limitation is the lack of medical history. Although patients are screened and asked about revision or previous surgeries as well as former injuries of the comparator knee, there is no guarantee a patient has not already been operated on or operated on again at a different hospital.

Conclusion

The reference database that will be gathered by this study in a subpopulation at risk of developing osteoarthritis will help to further explore the potential of acoustic biomarkers to diagnose osteoarthritis at earlier stages. A noninvasive knee sound diagnostic tool could provide a better understanding of the disease and its progression at low cost and without radiation exposure. Therefore, this study will contribute to improved clinical decision-making and enable preventive measures.

Acknowledgments

The authors would like to thank the participants for their time and contributions.

Funding

The trial sponsor is Prof. Atienza (EPFL, Lausanne, Switzerland). The study is funded by Innosuisse, and the investigational device is provided by Sensemodi.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JT, TT, DA, VAS

XSL•FO
RenderX

Data curation: LL, VAS
Formal analysis: JT, TT, VAS
Funding acquisition: JT, DA, VAS
Investigation: LL, VAS
Methodology: LL, JT, TT, VAS
Project administration: LL, JT, DA, VAS
Resources: JT, DA, VAS
Software: JT, TT
Supervision: DA, VAS
Writing – original draft: LL, JT, TT, VAS
Writing – review & editing: JT, TT, DA, VAS
Vincent Stadelmann serves as principal investigator and is responsible for the clinical conduct of the study.

Conflicts of Interest

The authors of this manuscript declare relationships with the following company: DA (CEO), JT (CTO), and TT (Co-Founder) are members of Sensemodi Sàrl. The data collection presented in this manuscript was performed by the Schulthess Klinik personnel without a link to Sensemodi Sàrl. The results presented in this manuscript are descriptive statistics without benefit to Sensemodi Sàrl. The other authors declare that they have no competing interests.

Checklist 1
SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist.
[DOCX File, 34 KB - resprot_v15i1e67032_app1.docx]

References

1. Leifer VP, Katz JN, Losina E. The burden of OA-health services and economics. Osteoarthr Cartil 2022 Jan;30(1):10-16. [doi: 10.1016/j.joca.2021.05.007]
2. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. Ann Rheum Dis 1957 Dec;16(4):494-502. [doi: 10.1136/ard.16.4.494] [Medline: 13498604]
3. Ehmig J, Engel G, Lotz J, et al. MR-imaging in osteoarthritis: current standard of practice and future outlook. Diagnostics (Basel) 2023 Aug 3;13(15):2586. [doi: 10.3390/diagnostics13152586] [Medline: 37568949]
4. Eckstein F, Wirth W. Quantitative cartilage imaging in knee osteoarthritis. Arthritis 2011;2011:475684. [doi: 10.1155/2011/475684] [Medline: 22046518]
5. Nevalainen MT, Uusimaa AP, Saarakkala S. The ultrasound assessment of osteoarthritis: the current status. Skeletal Radiol 2023 Nov;52(11):2271-2282. [doi: 10.1007/s00256-023-04342-3] [Medline: 37060461]
6. Gielis WP, Weinans H, Nap FJ, Roemer FW, Foppen W. Scoring osteoarthritis reliably in large joints and the spine using whole-body CT: osteoarthritis computed tomography-score (OACT-score). J Pers Med 2020 Dec 22;11(1):5. [doi: 10.3390/jpm11010005] [Medline: 33375114]
7. van Spil WE, Szilagyi IA. Osteoarthritis year in review 2019: biomarkers (biochemical markers). Osteoarthr Cartil 2020 Mar;28(3):296-315. [doi: 10.1016/j.joca.2019.11.007]
8. Kalo K, Niederer D, Stief F, et al. Validity of and recommendations for knee joint acoustic assessments during different movement conditions. J Biomech 2020 Aug 26;109:109939. [doi: 10.1016/j.jbiomech.2020.109939] [Medline: 32807320]
9. Choi D, Ahn S, Ryu J, Nagao M, Kim Y. Knee acoustic emission characteristics of the healthy and the patients with osteoarthritis using piezoelectric sensor. Sensors and Materials 2018;30(8):1629. [doi: 10.18494/SAM.2018.1877]
10. Prior J, Mascaro B, Shark LK, et al. Analysis of high frequency acoustic emission signals as a new approach for assessing knee osteoarthritis. Ann Rheum Dis 2010 May;69(5):929-930. [doi: 10.1136/ard.2009.112599] [Medline: 20413570]
11. Goossens Q, Locsin M, Gharehbaghi S, et al. Knee acoustic emissions as a noninvasive biomarker of articular health in patients with juvenile idiopathic arthritis: a clinical validation in an extended study population. Pediatr Rheumatol 2023 Jun 20;21(1):59. [doi: 10.1186/s12969-023-00842-7]
12. Mascaro B, Prior J, Shark LK, Selfe J, Cole P, Goodacre J. Exploratory study of a non-invasive method based on acoustic emission for assessing the dynamic integrity of knee joints. Med Eng Phys 2009 Oct;31(8):1013-1022. [doi: 10.1016/j.medengphy.2009.06.007] [Medline: 19643653]
13. Shark LK, Chen H, Quan W, Goodacre J. Acoustic emission and angular movement variations from early adulthood healthy knees to late adulthood osteoarthritic knees. 2016 Presented at: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Aug 16-20, 2016; Orlando, FL p. 2382-2385. [doi: 10.1109/EMBC.2016.7591209]
14. Befrui N, Elsner J, Flesser A, et al. Vibroarthrography for early detection of knee osteoarthritis using normalized frequency features. Med Biol Eng Comput 2018 Aug;56(8):1499-1514. [doi: 10.1007/s11517-018-1785-4] [Medline: 29392547]

XSL•FO
RenderX

15. Teague CN, Hersek S, Toreyin H, et al. Novel methods for sensing acoustical emissions from the knee for wearable joint health assessment. IEEE Trans Biomed Eng 2016 Aug;63(8):1581-1590. [doi: 10.1109/TBME.2016.2543226] [Medline: 27008656]
16. Machrowska A, Karpiński R, Maciejewski M, Jonak J, Krakowski P, Syta A. Multi-scale analysis of knee joint acoustic signals for cartilage degeneration assessment. Sensors (Basel) 2025 Jan 24;25(3):706. [doi: 10.3390/s25030706] [Medline: 39943344]
17. Karpiński R, Prus A, Jonak K, Krakowski P. Vibroarthrography as a noninvasive screening method for early diagnosis of knee osteoarthritis: a review of current research. Appl Sci (Basel) 2024 Dec 31;15(1):279. [doi: 10.3390/app15010279]
18. Thevenot J, Aminian K, Alonso DA. A wearable to assess knee joint integrity using non-contact acoustic sensors. 2021. URL: https://infoscience.epfl.ch/entities/patent/13b8a9e5-7523-4f10-bcaf-987e30a59ef9 [accessed 2026-02-20]
19. Shakya BR, Tiulpin A, Saarakkala S, Turunen S, Thevenot J. Detection of experimental cartilage damage with acoustic emissions technique: an in vitro equine study. Equine Vet J 2020 Jan;52(1):152-157. [doi: 10.1111/evj.13132] [Medline: 31032989]
20. Kechris C, Thevenot J, Teijeiro T, Stadelmann VA, Maffiuletti NA, Atienza D. Acoustical features as knee health biomarkers: a critical analysis. Artif Intell Med 2024 Dec;158:103013. [doi: 10.1016/j.artmed.2024.103013] [Medline: 39551004]
21. Lohmander LS, Englund PM, Dahl LL, Roos EM. The long-term consequence of anterior cruciate ligament and meniscus injuries: osteoarthritis. Am J Sports Med 2007 Oct;35(10):1756-1769. [doi: 10.1177/0363546507307396] [Medline: 17761605]
22. Casella G. Statistical Design: Springer; 2008. [doi: 10.1007/978-0-387-75965-4]
23. Stephan A, Stadelmann VA, Preiss S, Impellizzeri FM. Measurement properties of PROMIS short forms for pain and function in patients receiving knee arthroplasty. J Patient Rep Outcomes 2023 Feb 28;7(1):18. [doi: 10.1186/s41687-023-00559-x] [Medline: 36854937]
24. Dubousset J, Charpak G, Dorion I, et al. A new 2D and 3D imaging approach to musculoskeletal physiology and pathology with low-dose radiation and the standing position: the EOS system. Bull Acad Natl Med 2005 Feb;189(2):287-297. [doi: 10.1016/S0001-4079(19)33584-8] [Medline: 16114859]
25. Delin C, Silvera S, Bassinet C, et al. Ionizing radiation doses during lower limb torsion and anteversion measurements by EOS stereoradiography and computed tomography. Eur J Radiol 2014 Feb;83(2):371-377. [doi: 10.1016/j.ejrad.2013.10.026] [Medline: 24291000]
26. Koliogiannis VKA, Brandlhuber M, Messerschmidt V, et al. Is the EOS imaging system as accurate as conventional radiography in grading osteoarthritis of the knee? Eur Radiol 2021 May;31(5):3491-3497. [doi: 10.1007/s00330-020-07187-9] [Medline: 33119811]
27. Hunter DJ, Guermazi A, Lo GH, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthr Cartil 2011 Aug;19(8):990-1002. [doi: 10.1016/j.joca.2011.05.004]
28. Nagai K, Nakamura T, Fu FH. The diagnosis of early osteoarthritis of the knee using magnetic resonance imaging. Ann Joint 2018;3(2):110-110. [doi: 10.21037/aoj.2018.12.05]
29. Dobson F, Hinman RS, Roos EM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. Osteoarthr Cartil 2013 Aug;21(8):1042-1052. [doi: 10.1016/j.joca.2013.05.002]
30. Impellizzeri FM, Mannion AF, Leunig M, Bizzini M, Naal FD. Comparison of the reliability, responsiveness, and construct validity of 4 different questionnaires for evaluating outcomes after total knee arthroplasty. J Arthroplasty 2011 Sep;26(6):861-869. [doi: 10.1016/j.arth.2010.07.027] [Medline: 21074964]
31. Naal FD, Impellizzeri FM, Sieverding M, et al. The 12-item Oxford Knee Score: cross-cultural adaptation into German and assessment of its psychometric properties in patients with osteoarthritis of the knee. Osteoarthr Cartil 2009 Jan;17(1):49-52. [doi: 10.1016/j.joca.2008.05.017]
32. Impellizzeri FM, Leunig M, Preiss S, Guggi T, Mannion AF. The use of the Core Outcome Measures Index (COMI) in patients undergoing total knee replacement. Knee 2017 Mar;24(2):372-379. [doi: 10.1016/j.knee.2016.11.016] [Medline: 27979688]
33. Zahiri CA, Schmalzried TP, Szuszczewicz ES, Amstutz HC. Assessing activity in joint replacement patients. J Arthroplasty 1998 Dec;13(8):890-895. [doi: 10.1016/s0883-5403(98)90195-4] [Medline: 9880181]
34. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS) - validation and comparison to the WOMAC in total knee replacement. Health Qual Life Outcomes 2003 May 25;1(1):17. [doi: 10.1186/1477-7525-1-17] [Medline: 12801417]
35. Heijbel S, Naili JE, Hedin A, W-Dahl A, Nilsson KG, Hedström M. The Forgotten Joint Score-12 in Swedish patients undergoing knee arthroplasty: a validation study with the Knee Injury and Osteoarthritis Outcome Score (KOOS) as comparator. Acta Orthop 2020 Feb;91(1):88-93. [doi: 10.1080/17453674.2019.1689327] [Medline: 31711349]
36. Gharehbaghi S, Whittingslow DC, Ponder LA, Prahalad S, Inan OT. Acoustic emissions from loaded and unloaded knees to assess joint health in patients with juvenile idiopathic arthritis. IEEE J Biomed Health Inform 2021 Sep;25(9):3618-3626. [doi: 10.1109/JBHI.2021.3081429] [Medline: 34003759]
37. Sawaryn B, Piaseczna N, Sieciński S, et al. The assessment of the condition of knee joint surfaces with acoustic emission analysis. Sensors (Basel) 2021 Sep 29;21(19):6495. [doi: 10.3390/s21196495] [Medline: 34640815]

XSL•FO
RenderX

38. Whittingslow DC, Zia J, Gharehbaghi S, et al. Knee acoustic emissions as a digital biomarker of disease status in juvenile idiopathic arthritis. Front Digit Health 2020;2:571839. [doi: 10.3389/fdgth.2020.571839] [Medline: 34713044]
39. Nevalainen MT, Veikkola O, Thevenot J, et al. Acoustic emissions and kinematic instability of the osteoarthritic knee joint: comparison with radiographic findings. Sci Rep 2021 Oct 1;11(1):19558. [doi: 10.1038/s41598-021-98945-2] [Medline: 34599226]
40. Nagao M, Konno SI, Kim YH, Yokota O. Frequency response in bone joint acoustic sensor development. THC 2015 Oct 27;23(6):715-727. [doi: 10.3233/THC-151024]
41. Hax J, Leuthard L, Nauer S, Stadelmann VA, Leunig M, Rüdiger HA. Treatment options for persistent lateral femoral cutaneous nerve lesions after total hip arthroplasty via the direct anterior approach: retrospective analysis with clinical assessment. International Orthopaedics (SICOT) 2025 May;49(5):1107-1117. [doi: 10.1007/s00264-025-06512-5]

Abbreviations

ACL: anterior cruciate ligament
AE: acoustic emission
BASEC: Business Administration System for Ethical Committees
COMI: Core Outcome Measures Index
EPFL: École Polytechnique Fédérale de Lausanne
FJS-12: Forgotten Joint Score (with 12 items)
IMU: inertial measurement module
KL: Kellgren-Lawrence
KOOS: Knee Injury and Osteoarthritis Outcome Score
MAC: Meniscus/ACL/Cartilage-surgery registry
ML: machine learning
MOAKS: MRI Osteoarthritis Knee Score
MRI: magnetic resonance imaging
OARSI: Osteoarthritis Research Society International
OKS: Oxford Knee Score
PROM: patient reported outcome measurement
REDCap: Research Electronic Data Capture
TKA: total knee arthroplasty
UCLA: University of California at Los Angeles

Edited by J Sarvestan; submitted 02.Oct.2024; peer-reviewed by C Nichols, MA Hefny, R Karpinski; revised version received 30.Nov.2025; accepted 30.Nov.2025; published 26.Feb.2026.

Please cite as:
Leuthard L, Thevenot J, Teijeiro T, Atienza D, Stadelmann VA
Acoustic Emission Biomarkers for the Detection and Monitoring of Early Knee Osteoarthritis: Protocol for a Prospective, Single-Center, Exploratory Study
JMIR Res Protoc 2026;15:e67032
URL: https://www.researchprotocols.org/2026/1/e67032
doi:10.2196/67032

© Louis Leuthard, Jérôme Thevenot, Tomás Teijeiro, David Atienza, Vincent A Stadelmann. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

AI Models to Reduce Surgical Complications Through Intraoperative Video Analysis: Protocol for a Prospective Cohort Study

António Sampaio Soares[1], MD, PhD; Sophia Bano[2], PhD; Laura T Castro[3], MD; Margarida Pascoal[4], MD; Ricardo Rocha[3], MD; Paulo Alves[3], MD; Paulo Mira[3], MD; Joao Costa[1], MD, PhD; Manish Chand[2], MD, PhD; Danail Stoyanov[2], PhD; Catarina Barata[5], PhD

[1]School of Medicine, University of Lisbon, Av. Prof. Egas Moniz MB, Lisbon, Portugal

[2]University College London, London, United Kingdom

[3]Hospital Professor Doutor Fernando Fonseca, Amadora, Portugal

[4]Instituto Português de Oncologia Francisco Gentil, Lisbon, Portugal

[5]Institute for Systems and Robotics, LARSyS, Instituto Superior Técnico, Lisbon, Portugal

Corresponding Author:
António Sampaio Soares, MD, PhD
School of Medicine, University of Lisbon, Av. Prof. Egas Moniz MB, Lisbon, Portugal

Abstract

Background: Complications following abdominal surgery have a very significant negative impact on the patient and the health care system. Despite the spread of minimally invasive surgery, there is no automated way to use intraoperative video to predict complications. New developments in data storage capacity and artificial intelligence (AI) algorithm creation now allow for this.

Objective: This project aims to develop and validate deep learning models for accurately predicting postoperative complications, classified using the Clavien-Dindo scale. A key objective is to build and share an open-source dataset containing both intraoperative video data and postoperative outcomes.

Methods: This prospective cohort study will collect data reflecting day-to-day surgical practice from 1200 patients, focusing on patient outcomes and intraoperative video. Data will be collected from patients undergoing minimally invasive appendectomy, cholecystectomy, and colorectal resection in the urgent and elective settings. Each video will be annotated at the temporal and semantic level by the study team. Comprehensive data collection will encompass three domains: (1) preoperative variables, including patient demographics, comorbidities, laboratory values, and imaging findings; (2) intraoperative data featuring complete surgical video recordings from laparoscopic or robotic monitors, procedure duration, surgical approach, intraoperative complications, and surgeon-defined technical factors; and (3) 30-day postoperative outcomes classified using the Clavien-Dindo scale (grades I-V). This dataset will be shared under a noncommercial CC BY-NC-SA use license to promote scientific collaboration and innovation, with complete anonymization including metadata removal and out-of-body image blurring. For analysis, the dataset will be split into training, validation, and testing sets. Deep learning algorithms will be developed through supervised learning methodology using 2 parallel approaches: data-derived predictors using fine-tuned surgical video foundational models based on vision transformer architectures and surgeon-defined predictors based on documented intraoperative strategies. Algorithms will be trained on the training set to predict the Clavien-Dindo postoperative complication grade and categorize postoperative outcomes in minimally invasive abdominal surgery. Model performance will be analyzed through sensitivity, specificity, positive and negative predictive values, and area under the receiver operating characteristic curve on the validation and testing sets.

Results: Data collection started in 2024 and is expected to extend throughout 2025. The planned outputs include the publication of a research protocol, main results, and the open-source dataset. Through this initiative, the project seeks to significantly advance the field of AI-assisted surgery, contributing to safer and more effective practice.

Conclusions: Through the creation of an open dataset and the development of state-of-the-art deep learning models, this project seeks to transform the current paradigm in minimally invasive surgery. By providing the surgical AI community with robust, real-world data, the project aspires to catalyze innovations that will enhance surgical safety; refine predictive capabilities; and, ultimately, lead to better clinical outcomes.

International Registered Report Identifier (IRRID): DERR1-10.2196/62734

(JMIR Res Protoc 2026;15:e62734) doi:10.2196/62734

XSL•FO
RenderX

KEYWORDS

artificial intelligence; AI; surgery; dataset; intraoperative video; appendectomy; cholecystectomy

Introduction

The conventional approach to surgical training is deeply rooted in a model that combines theoretical teaching with supervised practice in the operating environment. However, this model has its flaws. One of the most notable is the difficulty in predicting the variability that occurs during surgery itself, whether due to each patient's unique anatomy or the specific pathology that justifies the intervention. Although surgeons accumulate a wealth of information during each procedure, these data often remain confined within the walls of the operating room. This is largely due to the lack of efficient mechanisms for capturing these data and the subsequent inability to analyze them systematically and rigorously. This situation creates a vicious cycle in which learning becomes fragmented and inefficient, negatively affecting clinical outcomes and the evolution of surgical practice.

The emergence of minimally invasive surgery marked a revolution in clinical practice. In addition to minimizing trauma and accelerating patient recovery, this approach provides an unprecedented opportunity: the ability to record the entire surgical procedure on video. Previous works, such as those by Birkmeyer et al [1] and Curtis et al [2], have already provided valuable insights by demonstrating that the quality of surgical gestures is correlated with clinical outcomes. These works were based on a still very rudimentary human assessment that cannot be scaled. With the advent of new technologies in image analysis and artificial intelligence (AI), it is now possible to automate the evaluation of these videos. The objective of this project is to go beyond simple evaluation, developing AI algorithms that can provide real-time support during the surgical act.

Postsurgical complications, both in minimally invasive cholecystectomy and colectomy, have a very high human and economic cost. They increase patient mortality and morbidity and burden health systems with additional costs in postoperative care and prolonged hospital stays [3]. In this context, it is not only desirable but also ethically imperative to seek new ways to reduce the risk of complications and improve clinical outcomes. Currently, postoperative complications are classified according to the Clavien-Dindo complication scale [4]. This scale classifies outcomes according to 5 grades: I (minor deviation from normal recovery), II (requiring pharmacological treatment), III (requiring surgical, endoscopic, or radiological intervention), IV (life-threatening complication requiring intensive care), and V (death of the patient).

The use of AI algorithms for intraoperative video analysis has the potential to revolutionize how safety in surgery is approached, allowing for the identification and correction of problems in real time and, thus, significantly improving the quality of care provided. This prospective dataset is intended to increase the data available for research on AI in surgery reflecting day-to-day practice. It will include outcome data to allow the use of this knowledge to guide the study of the surgical procedure. The aim of this project is the identification of intraoperative factors that predict postoperative complications based on the Clavien-Dindo classification in minimally invasive cholecystectomies, appendectomies, and colorectal resections and the development of AI algorithms for this purpose.

Methods

Dataset Construction

Study Design

This will be a multicentric prospective cohort study for patients undergoing minimally invasive cholecystectomy, appendectomy, or colorectal resection. Recruiting centers will be based in Europe.

Dataset Summary

The dataset will cover both preoperative and intraoperative data, as well as video and 30-day postoperative outcomes, on groups of patients regularly operated in a minimally invasive fashion, as can be seen in Figure 1. These patients will undergo cholecystectomy, appendectomy, and colorectal resection. These are high-volume surgeries, with a significant proportion being performed minimally invasively [5]. Video data will be collected from the minimally invasive device monitor (laparoscopic or robotic). Other variables will be tabular data, as detailed for every variable and procedure in Multimedia Appendices 1 to 3-3.

The datasets will be called Surg_Cloud Cholecystectomy, Surg_Cloud Appendectomy, and Surg_Cloud Colorectal and will include the version number (eg, V1 and V2). The first release is expected by June 2025. It will be shared through a web host, with log-in credentials available for researchers. The dataset will be made available under a CC BY-NC-SA license. This allows for its use for research purposes but restricts its use for commercial purposes. This balance aims to promote scientific collaboration and innovation while protecting the data from unauthorized commercial exploitation.

XSL•FO
RenderX

Figure 1. Surg_Cloud study flowchart.

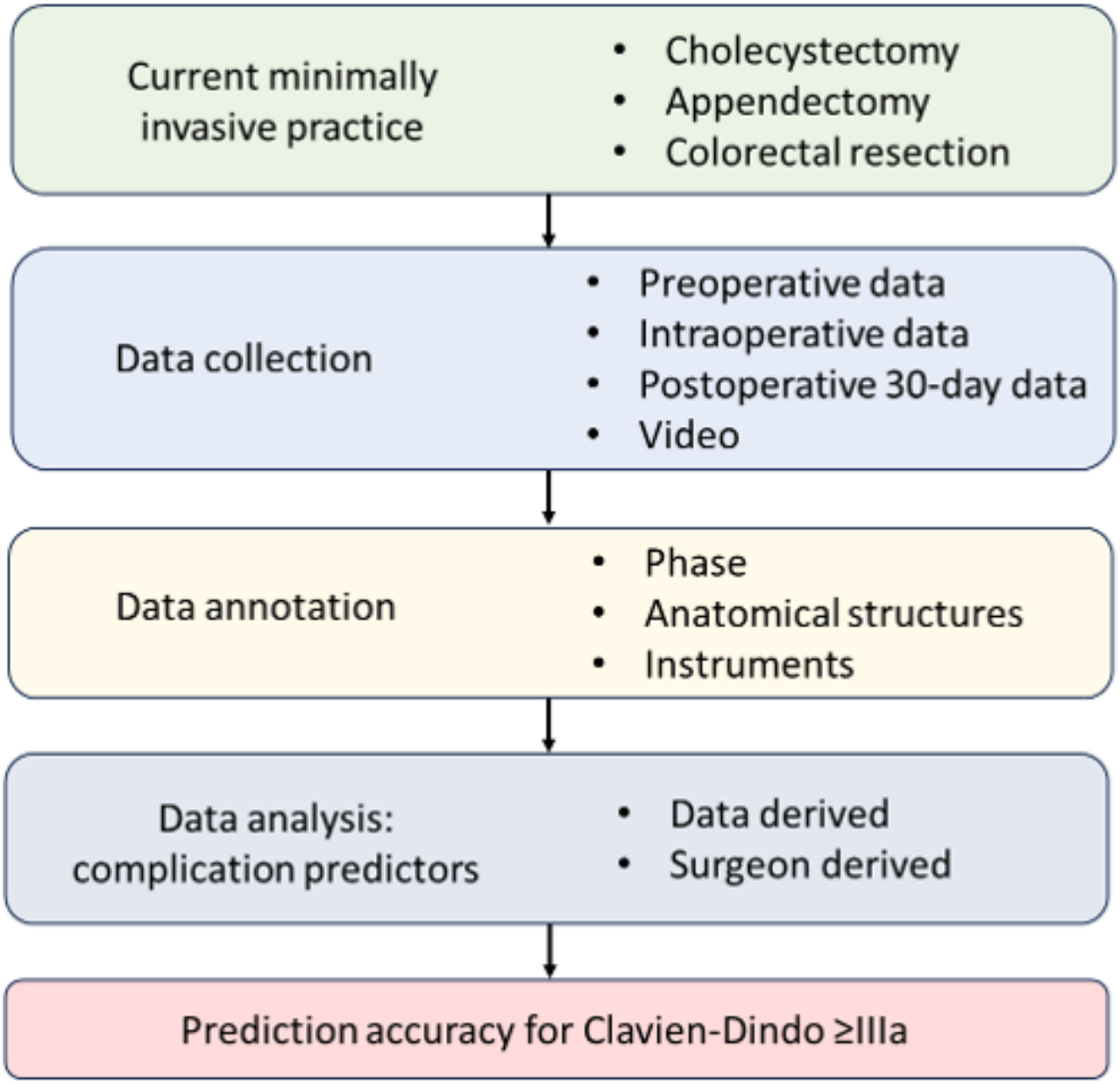

Inclusion Criteria

Data will be collected from patients receiving health care under normal circumstances. All patients undergoing cholecystectomy, appendectomy, and colorectal resection performed minimally invasively will be eligible. This applies to patients undergoing surgery in the urgent and elective settings. No changes in normal clinical practice will occur. These procedures are extremely common in a general surgeon's practice, and therefore, most hospitals can potentially enroll patients. Moreover, disease presentation is extremely heterogeneous across these procedures. This provides an opportunity to assess differential strategies in terms of changes in disease presentation. Patients withdrawing informed consent and for whom videos are not recorded adequately will be excluded.

Data Sampling and Shifts Over Time

The dataset will be composed of anonymized patient data from multiple hospitals. The enrollment of patients will take place under day-to-day clinical conditions, reflecting a convenience sample. It is to be expected that sicker patients will not be as represented among the study participants as more clinically stable ones, given that obtaining preoperative consent for unstable patients tends to be more difficult.

Every dataset released will have a corresponding version number where the details of data acquisition will be made explicit. If any significant changes occur between versions of the release, these will also be detailed in the accompanying documentation. It is expected that most variation will occur in terms of different hospitals' approaches to the specific procedures.

Patients will be grouped according to standard practice as mentioned in the data collection forms, which can be found in Multimedia Appendices 1 to 3-3. Apart from the video and the

XSL•FO
RenderX

data mentioned in this protocol, no other data will be collected from patients. These data will be available for every patient.

Groups at Risk of Disparate Health Outcomes

The dataset comprises various attributes that reflect patient interactions in public hospitals. These attributes were selected to provide insights into common health care practices and patient outcomes in these settings. Each attribute is presented at an individual level, ensuring a granular and detailed analysis. However, it is important to note that no patient-identifiable data have been collected to preserve privacy.

Limitations of the Dataset

The dataset will be created from videos collected in a very diverse range of hospitals. The patient selection will be mostly a convenience sample given the constraints of surgical care provision. Data quality will be variable given the range of devices that collect the data. These specifics will be detailed on a per-patient basis.

Modifications Made to the Data

The only data modifications planned are the blurring of out-of-body images to ensure complete anonymization. No information on the hospital where the data were collected will be shared. No further modifications are planned at this stage. No synthetic data will be used. No missing data will be imputed.

Known or Potential Bias Caused or Exacerbated by Data Acquisition and Processing

Given the complete anonymization of patients and the lack of impact on the patients' clinical care, no bias is expected to be exacerbated by the creation of this dataset.

Formal exclusion criteria have not been defined. Patients less likely to be included in the dataset are those undergoing urgent operations, in which the patients are likely to be sicker. This will necessarily impact the ability of the team to ensure proper informed consent. Due to this, a lower number of urgent patients are expected to be included.

The variables included are well standardized. Postoperative complications will be categorized using the Clavien-Dindo scale [4], which has 5 levels of severity, from I (minor deviations from the desired postoperative course) to V (patient death). This scale provides a structured framework for classifying and comparing complications, allowing for more precise analysis.

Ethical Considerations

Patient inclusion will necessitate ethics approval at every hospital and informed consent from every patient recruited. No patient-identifiable data will be shared in the dataset. All videos will be fully anonymized by removing all metadata and blurring out-of-body images if they appear.

This study has already received approval from the Hospital Professor Doutor Fernando Fonseca Ethics Committee (113/2023) to identify predictors of surgical complications, as well as for dataset curation and open-source publication. Every patient approached will be told the aims of this study, the absence of change in routine clinical care, and that only completely anonymized data will be made public in the dataset. No compensation will be provided to participants.

Patient and Public Participation

A formal patient and public involvement initiative is planned to be developed before dataset publication.

Importance of Reproducibility and Generalizability

Reproducibility of results in research is crucial for the credibility and advancement of science. In accordance with the principles of findability, accessibility, interoperability, and reusability [6], this project aims not only to conduct rigorous studies but also to make data and methods available in a way that allows for their verification and reuse by other researchers. Dataset development was conducted according to the STANDING (Standards for Data Diversity, Inclusivity, and Generalisability) Together recommendations [7]. By doing so, we increase the likelihood that the results are generalizable and, consequently, more useful worldwide. The creation of an open dataset follows the example of successful initiatives such as the Heidelberg Colorectal dataset published in the journal *Nature* [8]. This dataset is a reference in the field, consisting of 30 videos of colorectal surgeries accompanied by annotations and extensive documentation. Other institutions such as the Institut de Recherche contre les Cancers de l'Appareil Digestif (IRCAD) in the case of laparoscopic cholecystectomy [9] and the University of Dresden [10] have also contributed to science by publishing open-source datasets.

Analysis Plan

Objectives

We aim to develop a deep neural network using supervised methodology for predicting the grade of postoperative complications using the Clavien-Dindo scale. The overall aim of this study is the identification of predictors of postoperative complications in the included procedures.

Analysis

In the initial phase, a descriptive statistical analysis of demographic, intraoperative, and postoperative variables will be conducted. This analysis aims to provide a quantitative summary of the data, helping define the profile of the patients and the surgical interventions performed. Measures of central tendency, dispersion, and distribution will be calculated for each variable, such as mean, SD, median, and percentiles.

Subsequently, a regression analysis will be conducted to identify variables that may predict postoperative complications. This regression model will be adjusted to account for potential confounding factors and will allow for the assessment of the independent impact of each predictor variable. The analysis will be conducted for both categorical and quantitative variables. The ability of preoperative variables to predict surgery duration will be assessed.

All statistical tests will be performed with a predefined significance level of P=.05, and the results will be interpreted in the light of the clinical context to ensure that the conclusions are not only statistically significant but also clinically relevant.

XSL•FO
RenderX

Algorithm Development

Compliance With Methodological Guidelines

The development of AI algorithms for this project will be rigorously guided by best practices. The TRIPOD-AI (Transparent Reporting of a Multivariable Model for Individual Prognosis or Diagnosis–Artificial Intelligence) [11] guidance will be followed. This ensures that the development, validation, and implementation of AI algorithms in medical contexts follow strict standards to ensure their efficacy and safety. The publication of the STARD-AI (Standards for Reporting Diagnostic Accuracy–Artificial Intelligence) [12] is expected soon, and according to the timeline of this project, their use in publications originating from it is anticipated.

Supervised Learning Approach

The algorithms will be trained using a supervised learning approach, with the categories of the Clavien-Dindo scale serving as the outcome variable or "label," as can be seen in Figure 1. Supervised learning was chosen to allow for more focused and accurate training considering the specific and complex characteristics of the medical data involved.

The identification of predictors will follow 2 parallel work streams: data-based predictors and surgeon-defined predictors. For the data-based predictors, we will fine-tune a surgical video foundational model [13], adjusting it to predict the categories of the Clavien-Dindo scale. The foundational model has been pretrained on a vast amount of publicly available surgical videos and is expected to provide a discriminative feature space, which can be further fine-tuned to adjust to the desired task. The surgeon-defined predictors will be based on the exhaustive documentation of intraoperative strategies and their correlation with 30-day complication rates. For surgeon-derived predictors, annotated data will be used to automate their detection via the same algorithm design as for data-derived predictors.

All videos will be made available, with annotations performed in a collaborative fashion. These will identify phase, anatomical structures, and instruments. The annotation process will be standardized using validated protocols when available, such as the protocol by Mascagni et al [14].

The collected video data will be divided into 3 sets: training, validation, and testing. The proportions for this division will be 70%, 15%, and 15%, respectively. This structure aims to ensure that the model is trained on a large amount of data while reserving an adequate portion for validation and testing, thereby minimizing the risk of overfitting.

The Python programming language (Python Software Foundation), widely recognized for its versatility and robustness in AI projects, will be used. The specific infrastructure for training the neural networks will be PyTorch (Meta AI), an open-source machine learning library that is highly effective for implementing convolutional neural networks.

The specific type of neural network to be used will be a vision transformer as this is the backbone of the foundational model to be fine-tuned. Vision transformers have shown remarkable performances in different medical image analysis tasks [15]; thus, they are ideal for the task at hand. Additionally, the ability of transformers to perform a region-based analysis can play a relevant role in identifying key elements in surgical videos that correlate with patient outcomes through the inspection of the output of the self-attention blocks, in particular those associated with the "[class]" token. As we will be leveraging a foundational model to develop the algorithm, we have some level of assurance that training will be easier and lead to better performance.

Evaluation and Continual Improvement

After initial training, the algorithm will undergo several rounds of evaluation and adjustment. Metrics such as accuracy, sensitivity, and specificity will be used to assess the model's performance. The results of these evaluations will inform subsequent iterations of the algorithm, allowing for continuous improvement of its performance in line with current best practices [16]. Thus, the project aims to develop a highly effective and reliable AI algorithm that can be used to significantly improve surgical practice and patient outcomes.

Data Storage

Data will be uploaded to a secure cloud storage service. This service will include encrypted storage and access controlled via verified credentials, ensuring that data access is traceable. Additionally, the platform allows for the connection with applications to run the code for training and testing the algorithm. Both clinical and video data are uploaded after a period of 30 days after surgery, ensuring complete anonymization of the uploaded data. Video will be preprocessed before storage to remove patient metadata and blur out-of-body frames using an open-source video processing pipeline.

Authorship

Collaborative authorship will be attributed to members of the study group according to the established guidelines [17]. One hospital lead and up to 3 local collaborators will be eligible per center.

Results

Center recruitment is ongoing. Patient recruitment is expected to take place throughout 2026. Early algorithm development should start in June 2025, with the first outputs expected starting from January 2026, as can be seen in Figure 2. As of January 2025, a total of 3 hospitals have joined the study, with 2 working to secure local ethics committee approval. A preliminary assessment of each site has confirmed that data collection—encompassing both minimally invasive surgical video and corresponding patient information—can be integrated into routine workflows without disrupting patient care. Early interactions with hospital leads have indicated strong support for the study's objectives, suggesting that collaboration and data sharing can continue effectively.

Video quality and anonymization processes were confirmed to meet standards, with out-of-body images blurred to protect patient identities and metadata removed before secure upload. All data were subsequently linked via anonymized identifiers, validating the planned protocol for data management. Feedback from this pilot phase has been documented and will guide the

XSL•FO
RenderX

standardized workflows for other hospitals as they begin data collection.

Every patient who participated in the single hospital actively recruiting provided informed consent after receiving clear, concise explanations of the study's aims, potential benefits, and associated risks. The research team worked collaboratively with each hospital's ethics committee to ensure that the consent forms and protocols met local regulatory requirements. Early discussions with patients and clinicians suggested that the consent process is well understood. No significant ethical or logistical challenges were reported during this preliminary testing, and the approach appears suitable for broader implementation as additional centers come on board.

On the basis of initial enrollment rates and interest from additional sites, it is projected that the dataset will eventually include approximately between 1000 and 1200 patients undergoing minimally invasive cholecystectomy, appendectomy, or colorectal resection. This sample size was determined considering established principles for deep learning model development in medical contexts. For supervised learning tasks using pretrained models, it is recognized that adequate representation of outcome classes is essential for robust model training. Given expected overall complication rates of 15% to 20% across the 3 target procedures (approximately 10% for cholecystectomy, 15% for appendectomy, and 25% for colorectal resection), our planned sample yields approximately 150 to 240 patients with complications classified as Clavien-Dindo grade II or higher. The 70-15-15 training-validation-test split provides sufficient cases across complication grades to support both binary and multiclass classification approaches while maintaining adequate statistical power to detect clinically meaningful differences in model performance. This comprehensive dataset is expected to comprise preoperative and intraoperative variables, postoperative outcomes up to 30 days, and video recordings of each procedure. All video data will undergo a standardized annotation process to identify surgical phases, instruments, and anatomical structures, which should facilitate subsequent supervised learning tasks. The primary analysis will explore how intraoperative factors correlate with the Clavien-Dindo classification of postoperative complications, with secondary objectives focusing on regression modeling and convolutional neural network training to predict complication severity. These methods are designed to generate actionable insights into surgical risks and lay the groundwork for algorithmic tools that can support clinical decision-making. Feasibility assessment based on pilot data from participating centers indicates average monthly recruitment of 15 to 20 eligible patients per site, and with expansion to 5 to 6 centers, target recruitment remains achievable within the planned timeline.

Figure 2. Surg_Cloud project timeline.

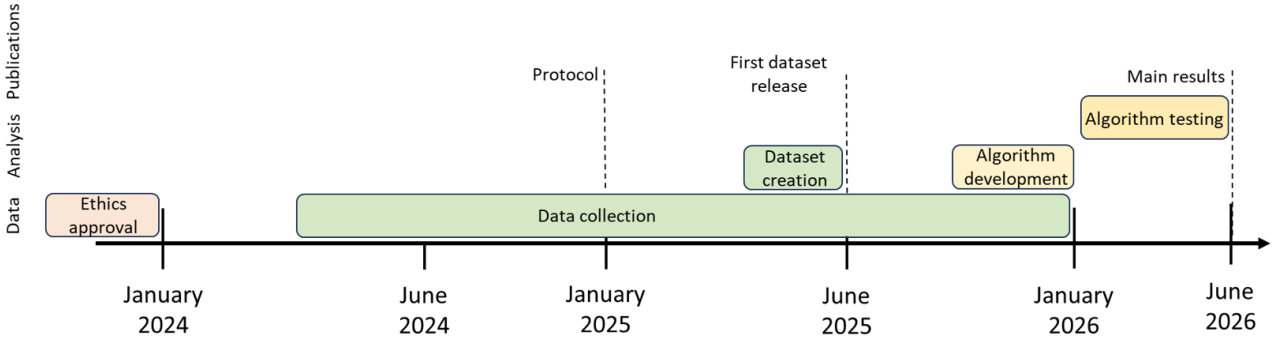

Discussion

Expected Findings

The primary aim of this project is to identify intraoperative factors that predict major postoperative complications following minimally invasive abdominal surgery. By exploring how anatomical structures, surgical instruments, and their interactions contribute to these complications, the project seeks to refine risk stratification and compare these AI-derived insights with existing postoperative complication prediction scores. Moreover, this initiative will result in a shared, open-source dataset that can be interrogated by the wider surgical AI community—maximizing collaborative potential and accelerating advancements in surgical video analysis.

Building on prior efforts—such as those by Birkmeyer et al [1] and Curtis et al [2], who correlated technical performance with clinical outcomes through video review—this initiative adds crucial elements to enhance scalability and applicability. Recent projects, including the Heidelberg Colorectal dataset [6] and open access resources from the Research Institute Against Cancer of the Digestive System [9] and the University of Dresden [10], highlight the value of systematic video collection, annotation, and analysis. However, many existing datasets have limitations, such as small sample sizes or lack of postoperative outcome data, restricting comprehensive analysis. In contrast, this project integrates the Clavien-Dindo classification [3] for robust tracking of postoperative complications and adopts a multicentric, prospective design to capture a broader range of patient populations and clinical scenarios.

Despite these strengths, some limitations warrant attention. First, reliance on convenience sampling across multiple hospitals may underrepresent urgent or high-risk cases due to limited time for obtaining informed consent. Second, variations in video capture technology and differing levels of surgical expertise introduce heterogeneity that lowers internal validity but also reflect real-world conditions—potentially boosting external validity. Third, the project currently emphasizes algorithm

XSL•FO
RenderX

development rather than hypothesis testing, although future iterations will include more rigorous prospective analyses once robust models are established. While other initiatives such as the Operating Room Black Box [18] capture additional operating room data that may influence postoperative complications [19], such technology remains inaccessible to many surgical practices. Consequently, this project focuses on creating a scalable, multicentric data collection and algorithm development pipeline designed for real-world applicability with a broader user base.

Looking ahead, several directions will further strengthen this endeavor. Plans include expanding dataset storage to accommodate increasing volumes of video and metadata from national and international collaborators guided by established frameworks such as STANDING Together [7] and the TRIPOD-AI guidelines [11]. Additionally, federated learning methodologies [20] will be explored to enable multiple institutions to train algorithms collectively while safeguarding patient privacy. Ongoing dissemination efforts will involve publishing open access protocols, inviting broader research contributions, and fostering a collaborative global surgical AI community. Through these initiatives, the project aspires to accelerate advancements in surgical video analysis and refine the predictive capabilities of AI in accurately forecasting postoperative complications.

Conclusions

This project bridges a critical gap in surgical research by integrating high-quality, multicentric intraoperative video data with rigorous postoperative outcome measures. By doing so, it offers a promising avenue for improving the prediction of major complications following minimally invasive abdominal surgery and enhancing surgical safety and quality. The open-source dataset, combined with transparent, reproducible methods, holds the potential to galvanize a broader community of clinicians, data scientists, and AI researchers. Through this collective effort, we anticipate the development of novel AI-driven interventions that will refine surgical practice; reduce complication rates; and, ultimately, improve patient outcomes.

Acknowledgments

The authors would like to acknowledge Pietro Mascagni and Fiona Kolbinger for helpful discussion on the topic of complication prediction. The authors used the generative artificial intelligence tool ChatGPT by OpenAI for minor writing edits in certain sections of the manuscript.

Funding

This work has received funding from a Portuguese Society of Coloproctology research grant and Portuguese Foundation for Science and Technology research grant (2024.07248.IACDC) https://sciproj.ptcris.pt/176816PRJ.

Data Availability

The dataset created through the implementation of this protocol will be made available to the surgical community through a dedicated publication.

Authors' Contributions

Conceptualization: ASS, SB, MC, DS
Data curation: LTC, MP
Methodology: ASS, SB, JC, CB
Supervision: PA, PM
Writing—original draft: ASS, LTC
Writing—review and editing: ASS, LTC, CB

Conflicts of Interest

DS is employed at Medtronic plc. All other authors declare no conflicts of interest.

Multimedia Appendix 1
Case report form for patients undergoing cholecystectomy.
[DOCX File, 17 KB - resprot_v15i1e62734_app1.docx]

Multimedia Appendix 2
Case report form for patients undergoing appendectomy.
[DOCX File, 17 KB - resprot_v15i1e62734_app2.docx]

Multimedia Appendix 3
Case report form for patients undergoing colorectal resection.
[DOCX File, 18 KB - resprot_v15i1e62734_app3.docx]

XSL•FO
RenderX

References

1. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. N Engl J Med 2013 Oct 10;369(15):1434-1442. [doi: 10.1056/NEJMsa1300625] [Medline: 24106936]
2. Curtis NJ, Foster JD, Miskovic D, et al. Association of surgical skill assessment with clinical outcomes in cancer surgery. JAMA Surg 2020 Jul 1;155(7):590-598. [doi: 10.1001/jamasurg.2020.1004] [Medline: 32374371]
3. Healy MA, Mullard AJ, Campbell DA Jr, Dimick JB. Hospital and payer costs associated with surgical complications. JAMA Surg 2016 Sep 1;151(9):823-830. [doi: 10.1001/jamasurg.2016.0773] [Medline: 27168356]
4. Clavien PA, Barkun J, de Oliveira ML, et al. The Clavien-Dindo classification of surgical complications: five-year experience. Ann Surg 2009 Aug;250(2):187-196. [doi: 10.1097/SLA.0b013e3181b13ca2] [Medline: 19638912]
5. Sheetz KH, Claflin J, Dimick JB. Trends in the adoption of robotic surgery for common surgical procedures. JAMA Netw Open 2020 Jan 3;3(1):e1918911. [doi: 10.1001/jamanetworkopen.2019.18911] [Medline: 31922557]
6. Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data 2016 Mar 15;3:160018. [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
7. Alderman JE, Palmer J, Laws E, et al. Tackling algorithmic bias and promoting transparency in health datasets: the STANDING Together consensus recommendations. Lancet Digit Health 2025 Jan;7(1):e64-e88. [doi: 10.1016/S2589-7500(24)00224-3] [Medline: 39701919]
8. Maier-Hein L, Wagner M, Ross T, et al. Heidelberg colorectal data set for surgical data science in the sensor operating room. Sci Data 2021 Apr 12;8(1):101. [doi: 10.1038/s41597-021-00882-2] [Medline: 33846356]
9. Twinanda AP, Shehata S, Mutter D, Marescaux J, de Mathelin M, Padoy N. EndoNet: a deep architecture for recognition tasks on laparoscopic videos. IEEE Trans Med Imaging 2017 Jan;36(1):86-97. [doi: 10.1109/TMI.2016.2593957] [Medline: 27455522]
10. Carstens M, Rinner FM, Bodenstedt S, et al. The Dresden surgical anatomy dataset for abdominal organ segmentation in surgical data science. Sci Data 2023 Jan 12;10(1):3. [doi: 10.1038/s41597-022-01719-2] [Medline: 36635312]
11. Collins GS, Moons KG, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
12. Sounderajah V, Ashrafian H, Golub RM, et al. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. BMJ Open 2021 Jun 28;11(6):e047709. [doi: 10.1136/bmjopen-2020-047709] [Medline: 34183345]
13. Wang Z, et al. Foundation model for endoscopy video analysis via large-scale self-supervised pre-train. In: Greenspan H, Madabhushi A, Mousavi P, editors. Medical Image Computing and Computer Assisted Intervention – MICCAI 2023: Springer Nature; 2023:101-111. [doi: 10.1007/978-3-031-43996-4_10]
14. Mascagni P, Alapatt D, Garcia A, et al. Surgical data science for safe cholecystectomy: a protocol for segmentation of hepatocystic anatomy and assessment of the critical view of safety. arXiv. Preprint posted online on Jun 21, 2021 URL: https://arxiv.org/abs/2106.10916 [accessed 2025-05-25]
15. Shamshad F, Khan S, Zamir SW, et al. Transformers in medical imaging: a survey. Med Image Anal 2023 Aug;88:102802. [doi: 10.1016/j.media.2023.102802] [Medline: 37315483]
16. Jayaraman P, Desman J, Sabounchi M, Nadkarni GN, Sakhuja A. A primer on reinforcement learning in medicine for clinicians. NPJ Digit Med 2024 Nov 26;7(1):337. [doi: 10.1038/s41746-024-01316-0] [Medline: 39592855]
17. Blencowe N, Glasbey J, Heywood N, et al. Recognising contributions to work in research collaboratives: guidelines for standardising reporting of authorship in collaborative research. Int J Surg 2018 Apr;52:355-360. [doi: 10.1016/j.ijsu.2017.12.019]
18. Boet S, Etherington C, Lam S, et al. Implementation of the operating room black box research program at the Ottawa hospital through patient, clinical, and organizational engagement: case study. J Med Internet Res 2021 Mar 16;23(3):e15443. [doi: 10.2196/15443] [Medline: 33724199]
19. Suliburk JW, Buck QM, Pirko CJ, et al. Analysis of human performance deficiencies associated with surgical adverse events. JAMA Netw Open 2019 Jul 3;2(7):e198067. [doi: 10.1001/jamanetworkopen.2019.8067] [Medline: 31365107]
20. Teo ZL, Jin L, Li S, et al. Federated machine learning in healthcare: a systematic review on clinical applications and technical architecture. Cell Rep Med 2024 Feb 20;5(2):101419. [doi: 10.1016/j.xcrm.2024.101419] [Medline: 38340728]

Abbreviations

AI: artificial intelligence
IRCAD: Institut de Recherche contre les Cancers de l'Appareil Digestif
STANDING: Standards for Data Diversity, Inclusivity, and Generalisability
STARD-AI: Standards for Reporting Diagnostic Accuracy–Artificial Intelligence
TRIPOD-AI: Transparent Reporting of a Multivariable Model for Individual Prognosis or Diagnosis–Artificial Intelligence

XSL•FO
RenderX

Edited by A Schwartz; submitted 30.May.2024; peer-reviewed by S Zhao, S Laplante; revised version received 05.Jul.2025; accepted 12.Aug.2025; published 06.Mar.2026.

Please cite as:
Sampaio Soares A, Bano S, Castro LT, Pascoal M, Rocha R, Alves P, Mira P, Costa J, Chand M, Stoyanov D, Barata C
AI Models to Reduce Surgical Complications Through Intraoperative Video Analysis: Protocol for a Prospective Cohort Study
JMIR Res Protoc 2026;15:e62734
URL: https://www.researchprotocols.org/2026/1/e62734
doi:10.2196/62734

© António Sampaio Soares, Sophia Bano, Laura T Castro, Margarida Pascoal, Ricardo Rocha, Paulo Alves, Paulo Mira, Joao Costa, Manish Chand, Danail Stoyanov, Catarina Barata. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digitally Enhanced Routine Outcome Monitoring in Italian Psychotherapy: Protocol for a Longitudinal Implementation Study

Mattia Vincenzo Olive[1], PhD; Antonino La Tona[2], PhD; Vivian Ostwald[3], MSc; Agostino Brugnera[2], PhD; Gianluca Lo Coco[3], PhD; Angelo Compare[2], PhD; Dana Tzur Bitan[4], PhD; Cristina Masella[1], MSc

[1]Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Via Raffaele Lambruschini 4b, Milan, Italy

[2]Department of Human and Social Sciences, Università degli Studi di Bergamo, Bergamo, Italy

[3]Department of Psychology and Educational Sciences, Università di Palermo, Palermo, Italy

[4]Department of Community Mental Health, University of Haifa, Haifa, Israel

Corresponding Author:
Mattia Vincenzo Olive, PhD
Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Via Raffaele Lambruschini 4b, Milan, Italy

Abstract

Background: Routine outcome monitoring (ROM) is an evidence-based methodology in psychotherapy that supports clinical decision-making through the standardized and systematic collection of patient-reported outcome data. Although its benefits are well established, ROM remains poorly integrated into routine psychotherapeutic practice in many countries, including Italy. Structural fragmentation of services, limited digital infrastructure, and cultural resistance within the profession hinder the adoption of data-informed clinical practices.

Objective: This study aims to investigate the implementation of a digital platform designed to support ROM in Italian psychotherapy, with a focus on individual, organizational, and contextual factors influencing adoption and integration into everyday clinical practice.

Methods: This is a longitudinal, mixed methods implementation study embedded within the OutProFeed project, the first randomized controlled trial in Italy focused on ROM in psychotherapy. The study is guided by the Consolidated Framework for Implementation Research, which informed the selection of constructs through a structured expert consensus process. Approximately 30 psychotherapists are included across diverse practice settings, therapeutic orientations, and geographic areas. Data will be collected at three time points (T0 baseline, T1 midimplementation at ~6 mo, and T2 final follow-up at ~12 mo). Qualitative data will be generated through semistructured interviews, while quantitative data will be collected using the Evidence-Based Practice Attitude Scale–36 (EBPAS-36). Qualitative and quantitative data will be linked at the individual level to support integrated interpretation of implementation trajectories.

Results: The study received ethical approval from the Ethics Committee of the University of Bergamo in April 2024. Technical setup of the digital platform and preimplementation engagement with participating psychotherapists were conducted between May and November 2025. Baseline (T0) implementation data collection started on a rolling basis in December 2025 and is planned to continue through March 2026. The midimplementation follow-up (T1; approximately 6 mo after individual baseline) is expected to occur between June and October 2026, with the final follow-up (T2; approximately 12 mo after baseline) planned between December 2026 and April 2027. Data analysis will begin after completion of follow-up assessments, with initial findings expected to be submitted for publication in 2027.

Conclusions: This protocol outlines a theory-informed and context-sensitive approach to studying the implementation of digitally supported ROM in psychotherapy. By combining qualitative and quantitative data within a longitudinal design, the study aims to generate insights into the conditions under which digital outcome monitoring may be integrated into routine psychotherapeutic work in fragmented and professionally autonomous health care systems.

International Registered Report Identifier (IRRID): PRR1-10.2196/82837

(JMIR Res Protoc 2026;15:e82837) doi:10.2196/82837

KEYWORDS

implementation; psychotherapy; routine outcome monitoring; ROM; Consolidated Framework for Implementation Research; CFIR; protocol; Italy

XSL•FO
RenderX

Introduction

Routine outcome monitoring (ROM) is an evidence-based methodology used in psychotherapy to support the continuous assessment and improvement of therapeutic processes through the standardized and systematic collection of clinical data [1-3]. ROM involves the regular administration of validated self-report questionnaires to monitor patient well-being, treatment progress, and the quality of the therapeutic alliance [1-4]. The data collected are typically visualized in graphical form and shared with the psychotherapist and, when appropriate, with the patient, enabling collaborative reflection and timely, data-informed adjustments to the therapeutic plan [3,5].

By integrating a structured feedback loop into clinical practice, ROM fosters greater transparency in the therapeutic relationship and promotes a more individualized and adaptive approach to care [6]. Empirical studies have demonstrated its potential to improve clinical outcomes, increase patient engagement, and detect early signs of stagnation or deterioration in treatment [2,6-8].

Historically, several barriers have limited the adoption of ROM in clinical practice [9]. For instance, paper-based methods often resulted in fragmented and delayed data collection, undermining the timeliness and utility of feedback. The growing availability of digital tools has addressed many of these limitations by enabling real-time data entry, automated analysis, and immediate visualization of results. Digital platforms enhance psychotherapists' ability to track progress over time, prepare for sessions more effectively, and ensure better continuity of care across encounters [5,10-16].

The American Psychological Association has long advocated the routine use of outcome and process monitoring in psychological care [17]. However, ROM implementation remains limited in many contexts, including Italy [18]. In Italy, this challenge is intensified by the organization of the mental health system: psychotherapy is not included among the services guaranteed by the national health system, resulting in a fragmented landscape dominated by private practice and high variability in theoretical orientations and documentation practices [19-21]. This context poses specific challenges for implementing structured practices such as ROM.

From an international perspective, the Italian case is analytically relevant because it makes visible implementation dynamics that are likely to emerge in many psychotherapy and mental health systems characterized by high professional autonomy, limited institutional coordination, and pluralistic practice arrangements, rather than representing an exceptional context.

This study contributes to bridging the gap between the recognized value of ROM and its limited implementation in real-world psychotherapeutic settings. The OutProFeed project, the first randomized controlled trial in Italy focused on ROM in psychotherapy, evaluates the implementation of digitally supported ROM, enabled by a digital platform (Mindy), designed to support psychotherapists in collecting and using clinical data in compliance with privacy regulations. Mindy is a digital platform that supports the administration of routine outcome measures and provides structured feedback through visual dashboards. Its main features, including the data collection interface, feedback visualization, and security-related design choices, are described in detail in a prior usability study [22].

While grounded in the Italian context, the study also offers insights that may inform implementation efforts in other health systems facing similar structural, technological, or cultural barriers [23-25]. By documenting the implementation process, this research provides a practical foundation for the broader adoption of data-informed practices in psychotherapy [26].

Methods

Overview

This study investigates the implementation of digitally supported ROM in Italian psychotherapeutic practice using the Mindy platform. Building on existing international evidence regarding the benefits of ROM in enhancing therapeutic effectiveness, this research focuses on understanding how such a practice can be introduced, adopted, and sustained in a context characterized by a high degree of professional autonomy, heterogeneous clinical approaches, and limited institutional support. The study was designed as a mixed methods implementation study to capture the complexity of this process, addressing not only whether ROM can be used but also how it is actually integrated into psychotherapists' clinical routines and what contextual factors facilitate or hinder this integration. A combined qualitative and quantitative approach was deemed necessary to generate both explanatory insight and empirical grounding for a robust analysis of implementation dynamics.

The Consolidated Framework for Implementation Research and the Rationale Behind Its Employment

To guide the design and evaluation of this mixed methods implementation study, we adopted the Consolidated Framework for Implementation Research (CFIR) [27,28], a widely recognized framework to support the systematic analysis of complex implementation processes. CFIR brings together constructs from multiple theoretical models and organizes them into 5 domains: the characteristics of the innovation, the outer setting, the inner setting, the individuals involved, and the implementation process. CFIR was selected not only for its comprehensiveness but because it is theoretically suited to examine implementation in professional domains characterized by high autonomy, weak formal hierarchy, and discretionary practice, such as psychotherapy. Unlike frameworks that focus primarily on organizational readiness or intervention fidelity, CFIR explicitly accommodates the interaction between individual professionals, loosely structured work settings, and evolving implementation processes. This is particularly relevant for psychotherapy, where adoption of structured practices such as ROM does not occur through formal mandates but through negotiated alignment with professional judgment, therapeutic orientations, and everyday clinical routines. CFIR therefore provides a theoretically appropriate lens to analyze implementation as a multilevel and processual phenomenon, in which professional agency, contextual constraints, and

XSL•FO
RenderX

characteristics of the innovation jointly shape how digitally supported ROM is interpreted and integrated into practice.

In this study, CFIR was used from the outset to inform the development of the research design, the construction of qualitative and quantitative data collection instruments, and the approach for coding and interpretation of findings. The framework's adaptability proved particularly valuable in structuring the analysis of an intervention (ROM) via a digital platform, introduced into a highly fragmented and heterogeneous psychotherapeutic system, such as the one operating in Italy. CFIR's emphasis on capturing the perspectives of those directly involved in the implementation aligns with the study's focus on psychotherapists as both subjects and agents of change.

The framework is not applied at a single point in time but is used to guide data collection and analysis throughout the full implementation trajectory [29]. This continuous engagement with CFIR supports the identification of context-specific dynamics as they evolve, allowing the study to capture temporal variation and the interaction between individual, organizational, and systemic factors. The structured yet flexible nature of CFIR enhances the study's explanatory power and supports the development of insights that are both context-sensitive and transferable beyond the Italian setting.

Implementation Setting

The implementation study is part of the broader *OutProFeed* project, which aims primarily to assess the clinical effectiveness of digitally supported ROM in improving mental health outcomes. Alongside this central objective, the project also includes 2 complementary strands of investigation: one focusing on the usability of the digital platform and the other, addressed in this study, exploring the conditions under which ROM can be successfully adopted and sustained in everyday psychotherapeutic practice. The present work specifically examines the implementation process, taking into account contextual, organizational, and individual-level factors.

The unit of analysis is the individual psychotherapist. This choice reflects the structure of the Italian psychotherapeutic landscape, in which professionals often work autonomously and exercise substantial discretion in adopting new tools and practices [30,31]. By focusing on the psychotherapist as the key actor in the implementation process, the study aims to capture the variability in professional routines, digital readiness, and perceived value of ROM, all of which are expected to influence the integration of digitally supported ROM into routine psychotherapeutic practice.

The study is conducted within the Italian mental health system, where psychotherapy provision is largely fragmented across private practice and under-resourced public services [21,32]. As a result, service provision is predominantly private, with many psychotherapists operating independently and outside institutional structures. Even within the public system, psychotherapy is available only in a limited number of settings, such as family counseling centers and specialized mental health services, which often suffer from limited resources and organizational constraints. This context creates significant variation in how psychotherapy is delivered, regulated, and documented across the country [32,33].

Such heterogeneity offers a unique opportunity to examine implementation in a real-world setting characterized by institutional gaps, inconsistent digital infrastructure, and diverse professional orientations. The insights generated by this study are intended to inform not only national efforts to scale up ROM but also broader international discussions on the conditions needed to support the effective integration of digital tools in mental health care.

Implementation Outcome

The study targets early-stage implementation at the level of professional practice. The primary implementation outcome is the integration of digitally supported ROM into psychotherapists' everyday clinical work. Integration is understood as the extent to which ROM is practically taken up, interpreted, and incorporated into routine therapeutic activities over time, rather than as an indicator of clinical effectiveness or system-level diffusion. The study examines this outcome by exploring how psychotherapists engage with the platform, how ROM aligns with professional routines and norms, and how ROM-related activities are incorporated into ongoing clinical practice. This corresponds to early-stage outcomes such as acceptability, appropriateness, and perceived feasibility of digitally supported ROM at the level of professional practice, rather than to clinical effectiveness or system-level adoption. The CFIR framework is used to identify and interpret the individual, organizational, and contextual conditions that shape this implementation process.

In this study, successful implementation is operationalized in processual terms, rather than through predefined performance thresholds. Meaningful implementation is indicated by psychotherapists' reported incorporation of ROM-related activities into routine clinical work, perceived alignment between ROM and therapeutic practice, and sustained engagement with the digital platform over time. Changes in Evidence-Based Practice Attitude Scale–36 (EBPAS-36) scores are interpreted as complementary indicators of evolving openness toward evidence-based and structured practices, while qualitative data provide the primary basis for assessing how such changes translate into everyday clinical routines. This approach is consistent with early-stage implementation research, where the aim is to understand how practices are integrated and adapted, rather than to quantify adoption or effectiveness.

CFIR Construct Selection and Consensus Process

Given the extensive range of constructs within the CFIR, a structured and transparent process was developed to identify the constructs most relevant to the implementation of digitally supported ROM in psychotherapy. The objective was to ensure a balance between theoretical completeness and practical applicability, adapting the framework to the specific characteristics of the intervention and its context.

The final set of selected constructs is reported in Table 1. The selection was guided by four exclusion criteria, applied in the following order:

XSL•FO
RenderX

- *Relevance to the Italian context:* Constructs whose variation is primarily driven by national-level or cross-country factors were excluded, given the single-country scope of the study.
- *Accessibility of information to participants:* Constructs requiring information not reasonably accessible to practicing psychotherapists were removed in order to ensure the reliability of the data collected.
- *Empirical support from existing literature:* Constructs that have shown limited relevance in comparable implementation studies were deprioritized, while those consistently associated with implementation outcomes were prioritized.
- *Specificity to the case under study:* Residual elements not covered by the above criteria were considered in light of the characteristics of the platform, the type of clinical practice, and the level of professional autonomy.

XSL•FO
RenderX

Table . Selected constructs.

Domain and construct	Average rating	Selected in phase 1	Selected in phase 2
Innovation			
Source	1.67		
Evidence base	2		
Relative advantage	6.33	✓	
Adaptability	6.33	✓	
Trialability	1		
Complexity	5.5	✓	
Design	5.5	✓	
Cost	1.67		
Outer setting			
Critical incidents	1		
Local attitudes	1		
Local conditions	1.67		
Partnerships and connections	1.67		
Policies and laws	5		
Financing	1		
External pressure	1		
Societal pressure	1		
Market pressure	1		
Performance measurement pressure	1		
Inner setting			
Structural characteristics	1		
Physical infrastructure	1		
IT infrastructure	1.83		✓
Work infrastructure	1.5		✓
Relational connections	1.5		
Communications	1.5		✓
Culture	2.5		
Human equality-centeredness	1		
Recipient-centeredness	4.17	✓	
Deliverer-centeredness	1.83		✓
Learning-centeredness	4.33	✓	
Tension for change	6		
Compatibility	6.67	✓	
Relative priority	3.83		✓
Incentive system	5		
Mission alignment	6.67	✓	
Available resources	1		
Funding	1.67		
Space	1		
Individuals			

XSL•FO
RenderX

Domain and construct	Average rating	Selected in phase 1	Selected in phase 2
High-level leaders	1		
Mid-level leaders	1		
Opinion leaders	1		
Implementation facilitators	2		✓
Implementation leads	1		
Implementation team members	1		
Other implementation support	1		
Innovation deliverers	1		
Innovation recipient	1		
Need	1.5		
Capability	6.17	✓	
Opportunity	6.33	✓	
Motivation	3.5		✓
Implementation process			
Teaming	1		
Assessing needs	4.83		
Innovation deliverers	1		
Innovation recipient	4.5	✓	
Assessing context	4.83	✓	
Planning	1		
Tailoring strategies	1		
Engaging	2		
Innovation deliverers (F1)	1.67		
Innovation recipient (F2)	4.33		
Doing	2.17		
Reflecting and evaluating	2		
Implementation (H1)	1		
Innovation (H2)	5.67	✓	
Adapting	3.5		✓

The evaluation process involved 9 contributors, selected for their complementary expertise:

- Three experts in digital health and management of digital transformation, affiliated with academic institutions and research centers
- Six professionals with backgrounds in clinical psychology and psychotherapy, with practical experience in the field and familiarity with the intervention

A two-step consensus process was then conducted:

1. Phase 1—Individual rating: each contributor independently assessed the relevance of each CFIR construct using a 7-point Likert scale (1=low relevance; 7=high relevance) and provided optional qualitative comments to support their evaluations.
2. Phase 2—Plenary discussion: two consensus meetings were held to discuss constructs with intermediate or divergent ratings. This collaborative phase allowed for the reconsideration of constructs based on disciplinary perspectives and case-specific insights.

Following consensus, each selected construct was associated with a specific interview item. These items were iteratively revised to ensure conceptual consistency with the framework, clarity of language, and clinical appropriateness. The resulting model supports a focused and theoretically grounded analysis of the implementation process, adapted to the structure of psychotherapy practice in Italy.

Data Collection and Analysis

Overview

Data will be collected over the study period through a mixed methods design combining semistructured interviews and a validated questionnaire (EBPAS-36). The objective is to

XSL•FO
RenderX

investigate how psychotherapists engage with the digital platform *Mindy* and to identify the factors that facilitate or hinder the implementation of digitally supported ROM across different phases of adoption.

Data collection follows a rolling baseline design (Figure 1). After enrollment, each participant completes a baseline assessment (T0) on a rolling basis between December 2025 and March 2026, with follow-up assessments scheduled approximately 6 months (T1) and 12 months (T2) after their individual baseline. The midimplementation follow-up (T1) is expected to occur between June and October 2026 and the final follow-up (T2) between December 2026 and April 2027. Because enrollment occurs on a rolling basis, calendar dates for follow-ups differ across participants and partially overlap.

Figure 1. Study design and data collection timeline (rolling baseline with longitudinal follow-up). EBPAS-36: Evidence-Based Practice Attitude Scale–36.

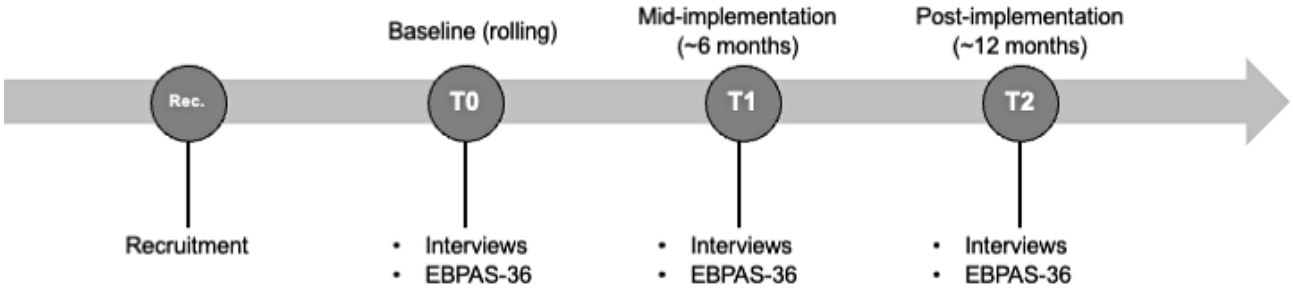

Qualitative Data Collection

Semistructured interviews will be conducted with a sample of around 30 psychotherapists, selected to reflect the diversity of the Italian context in terms of employment setting (private, public, freelance), therapeutic orientation, and geographical location. Recruitment is carried out within the broader OutProFeed project through established collaborations with psychotherapy training schools and their professional networks, complemented by outreach to practitioners in both private practice and organizational settings. The research team monitors the evolving sample composition during enrollment to avoid overrepresentation of specific profiles (eg, a single orientation or geographic area) and to ensure coverage across the predefined variation dimensions. This recruitment approach is intended to reduce selection bias while remaining consistent with the study's purposive, maximum-variation design. While an initial target of approximately 30 participants is set, the final sample size will be determined during the research process. Data collection will continue until all these key dimensions are represented and data saturation is reached.

The target sample of approximately 30 psychotherapists is grounded in the structural and professional heterogeneity of psychotherapy practice in Italy. In particular, the sampling strategy is designed to ensure coverage of the main psychotherapeutic orientations currently practiced in the Italian context, which represent distinct clinical traditions and professional cultures that may shape how ROM is interpreted and integrated into everyday practice. The study aims to include psychotherapists from at least five widely adopted approaches (eg, cognitive-behavioral, psychodynamic, systemic-relational, humanistic, and integrative orientations).

Within this framework, participants are purposively recruited to also reflect variation in employment arrangements (private practice and organizational settings) and geographic macro-areas (North, Center, South/Islands). With a target of approximately 30 participants, the study can include multiple psychotherapists within each major orientation while preserving sufficient diversity across organizational and regional contexts. This sample size supports comparative, process-oriented analysis of implementation trajectories across different professional logics, without aiming for statistical representativeness.

A purposive subsample will be followed longitudinally across T0, T1 (~6 mo), and T2 (~12 mo) to enable within-case process tracing; all participants will be invited at each wave. Meaningful longitudinal analyses will be conducted for participants with repeated observations across waves, with attrition explicitly documented and reported.

The interview guide, based on the CFIR constructs selected during the consensus process, is included in Multimedia Appendix 1. It includes approximately 16 questions for psychotherapists in private practice and 20 questions for those employed in organizations. Interviews will be conducted individually in Italian via Microsoft Teams, in quiet and distraction-free settings. Each session will be led by a pair of researchers, one with clinical expertise in psychotherapy and one with a background in digital transformation, to ensure both contextual sensitivity and methodological coherence.

With informed consent, all interviews will be audio-recorded and transcribed verbatim. Interviewers will adopt an open, nondirective style to encourage participants to reflect freely on their experience. The average duration of interviews is expected to be between 25 and 30 minutes.

Qualitative Data Analysis

Thematic analysis will be applied to the interview transcripts, following the methodology proposed by Braun and Clarke [34]. The CFIR framework will guide the coding process and the organization of emerging themes [35]. Analysis will be conducted using Excel as an analytic workspace. The process will begin with close reading and familiarization, followed by a deductive coding phase that maps participants' responses onto the relevant CFIR domains and constructs. This approach supports a structured interpretation of barriers and facilitators, while also allowing room for emerging contextual elements.

XSL•FO
RenderX

Quantitative Data Collection and Analysis

In parallel with each interview, participants will complete the EBPAS-36, a validated instrument that assesses attitudes toward adopting evidence-based practices. In line with the rolling baseline design, T1 and T2 are scheduled approximately 6 and 12 months after each participant's baseline assessment.

The EBPAS-36 consists of 36 items rated on a 5-point Likert scale and includes multiple subscales. Descriptive statistics will be calculated for each wave of data collection, including mean scores and SDs. Negative subscale items will be reverse-coded so that higher scores consistently indicate more positive attitudes toward innovation and structured clinical practices.

Repeated measures will be analyzed using a descriptive longitudinal approach. For participants providing data at multiple time points, changes in EBPAS-36 total and subscale scores will be explored within individuals, focusing on the direction and magnitude of change across waves. Where appropriate and depending on data completeness and distributional assumptions, nonparametric paired comparisons (eg, Wilcoxon signed rank tests) may be used to explore differences between time points. These analyses are intended to be exploratory and supportive of the qualitative findings, rather than hypothesis testing.

Quantitative data will be linked to the corresponding qualitative data at the level of individual participants. This integration will enable a dual-layered analysis of each psychotherapist's experience, combining narrative accounts with attitudinal profiles. Although not intended as a formal triangulation, this approach will enhance the interpretive depth of the study and contribute to a deeper understanding of implementation trajectories over time.

Data Integration

Qualitative and quantitative data are integrated at the level of individual participants using a convergent, case-oriented mixed methods approach. For each psychotherapist, interview data constitute the primary source for reconstructing the implementation process and its evolution over time. EBPAS-36 scores are used as complementary information to characterize psychotherapists' attitudinal positioning toward evidence-based practices at each time point and to support the interpretation of qualitative changes or continuities. Integration, therefore, occurs through iterative comparison within cases, examining how shifts in narratives of ROM integration relate to stable or changing attitudinal profiles. This strategy prioritizes explanatory coherence over formal triangulation and is aligned with the study's focus on understanding implementation as a situated professional process.

Pilot Test

A pilot phase was conducted with 3 psychotherapists to assess the adequacy and feasibility of the data collection protocol and to support its refinement prior to the main study. This pilot phase was explicitly conceived as a feasibility and protocol-refinement step and not as a source of empirical findings or preliminary implementation results. It provided a pragmatic opportunity to test the interview guide developed around the selected CFIR constructs (Table 1) and to assess the clarity, flow, and acceptability of the interview structure and language.

The interview guide used during the pilot is available in Multimedia Appendix 2. An initial version had been prepared based on the selected constructs. Insights from the pilot led to minor procedural refinements, primarily concerning the sequencing and phrasing of questions to facilitate a more natural conversational flow consistent with clinical interaction. The substantive content of the questions was not altered, and all targeted constructs were retained.

Quantitative data were not collected at this stage. The EBPAS-36 questionnaire, which will be administered in the main phase of the study, is a validated instrument that requires a sufficiently large and diverse sample to yield meaningful results. Given the feasibility-oriented purpose of the pilot and the small number of participants involved, its administration at this stage was not considered appropriate.

Ethical Considerations

The study has been reviewed and approved by the Committee for Integrity and Research Ethics of the University of Bergamo (2024_04_08) in compliance with national regulations and the Declaration of Helsinki. All participants are informed about the aims of the research, the voluntary nature of their participation, and the possibility to withdraw at any time without consequences.

Before each interview, participants receive an information sheet and are asked to provide written informed consent, including permission to record the session. Consent forms are stored securely and separately from the collected data. Interviews are conducted in a confidential setting using secure online platforms. Audio recordings are transcribed verbatim and anonymized to ensure participant privacy.

All personal identifiers are removed or replaced with pseudonyms during transcription. Data are stored in encrypted, password-protected institutional folders accessible only to the research team. Files are regularly backed up and managed in accordance with institutional data protection policies and the European General Data Protection Regulation.

For psychotherapists operating in private practice and using the digital platform, the implementation study does not involve access to identifiable patient-level clinical data. The research team does not access or extract raw clinical content entered into the platform. Data collected for research purposes focus exclusively on psychotherapists' experiences, attitudes, and implementation processes and are analytically and technically separated from any clinical documentation generated within routine practice.

Quantitative data collected through the EBPAS-36 questionnaire are similarly anonymized and securely stored. The integration of qualitative and quantitative data at the individual level is performed using anonymized study codes, ensuring that no identifying information is used in the analysis or dissemination of findings.

XSL•FO
RenderX

Dissemination of results will take place in aggregated form only, and no individual participant will be identifiable in publications or presentations. Participants are also informed that the results may contribute to academic and policy-oriented outputs aimed at improving clinical practice in the field of psychotherapy. No participant compensation is provided.

Results

This implementation study is conducted within the broader OutProFeed project, which received ethical approval from the Ethics Committee of the University of Bergamo in April 2024 (protocol number 2024_04_08); the present work is a predefined substudy covered by this approval. Technical setup of the digital platform and engagement with participating psychotherapists were conducted between May and November 2025. Baseline (T0) implementation data collection started on a rolling basis in December 2025 and is planned to continue through March 2026. The midimplementation follow-up (T1; approximately 6 mo after individual baseline) is expected to take place between June and October 2026, with the final follow-up (T2; approximately 12 mo after baseline) planned between December 2026 and April 2027. As of the time of manuscript submission, no implementation outcome analyses have been completed.

Discussion

Main Contributions and Anticipated Insights

This implementation study is designed to generate practice-relevant insights into how digitally supported ROM is incorporated into routine psychotherapy work in a fragmented and professionally autonomous care system. Specifically, we expect to identify (1) the implementation mechanisms through which psychotherapists interpret, negotiate, and integrate ROM into existing clinical routines, (2) the contextual and organizational conditions that enable or constrain this integration, and (3) how these dynamics vary across therapeutic orientations and practice arrangements. These anticipated contributions are implementation-focused and do not concern clinical effectiveness, which is addressed within the broader OutProFeed trial.

ROM and feedback-informed approaches are supported by a robust body of evidence demonstrating their potential to improve clinical outcomes, patient engagement, and early detection of treatment deterioration [2,6,7]. Despite this evidence, the integration of ROM into routine psychotherapeutic practice remains uneven, particularly in settings where its use requires sustained changes to psychotherapists' workflow, documentation practices, and interaction with patients [3,5,9]. Research on the digital transformation of mental health services further suggests that digital platforms can lower some practical barriers to ROM adoption but do not in themselves resolve deeper professional and organizational challenges related to implementation [10]. In this context, the present protocol explicitly focuses on implementation as a professional and organizational process, rather than on clinical effectiveness, which is addressed within the broader OutProFeed trial.

Despite the evidence base, ROM adoption in psychotherapy is often hindered by barriers that are well documented in implementation and measurement-based care research. These include psychotherapists' concerns that structured monitoring may constrain clinical discretion or undermine therapeutic work, uncertainty about data ownership and potential "surveillance" uses of outcome data, and perceived workload associated with repeated administration and review of measures [3,5,6,9]. In digitally mediated models of care, technology can reduce friction in data capture and feedback, but it can also amplify sensitivities related to confidentiality, governance, and professional autonomy if the tool is perceived as externally imposed or poorly aligned with clinical routines [10]. These barriers are expected to be particularly salient in fragmented systems dominated by independent private practice, such as Italy, where formal organizational mandates and shared infrastructures are limited [18,21].

Relation to Prior Implementation Research

From an implementation science perspective, this study responds to recent calls to strengthen mental health implementation research by attending more closely to contextual variation, professional practices, and the dynamics of early-stage implementation [26]. The Italian psychotherapy ecosystem offers a particularly salient empirical setting for this purpose. Psychotherapy in Italy is characterized by high professional autonomy, strong theoretical pluralism, and limited institutional coordination, with service provision largely fragmented across private practice and under-resourced public services [18,21]. In such a context, the uptake of structured monitoring practices such as ROM is likely to depend less on formal mandates or technical availability and more on how these practices are perceived to fit with psychotherapists' professional identities, therapeutic orientations, and everyday clinical routines.

Consistent with this view, the study conceptualizes implementation as an evolving process of professional integration rather than as a binary outcome of adoption versus nonadoption. Preliminary insights from the pilot phase suggest that variability across therapeutic orientations and practice arrangements is not incidental but central to understanding how ROM is interpreted and enacted in practice. This aligns with implementation research emphasizing nonlinear and processual models of change, in which innovations are adapted, negotiated, and selectively incorporated into routine work over time. The use of the CFIR provides a structured analytical lens to identify multilevel determinants of these processes while enabling comparison with prior implementation studies in digital mental health and psychotherapy [27,28].

Methodological Strengths and Limitations

Methodologically, the longitudinal mixed methods design is intended to capture temporal dynamics in psychotherapists' engagement with ROM, addressing limitations of cross-sectional approaches that treat attitudes and practices as static predictors. Qualitative interviews serve as the primary source for reconstructing implementation trajectories, while repeated measurement of attitudes toward evidence-based practices using the EBPAS-36 offers complementary insight into how openness to evidence-based approaches may relate to evolving experiences

XSL•FO
RenderX

of ROM integration. In line with the study's implementation focus, these data are used for explanatory interpretation rather than for testing causal models or estimating effectiveness. Two limitations should be noted. First, the purposive sample and the focus on one national context support analytic transferability rather than statistical generalizability. Second, the reliance on self-reported accounts and repeated interviews may be subject to social desirability and attrition over time, which we will document and consider in interpretation.

Dissemination

Following completion of data collection, findings will be disseminated through peer-reviewed publications, academic conferences, and stakeholder-facing dissemination activities, contributing to ongoing debates on how data-informed practices can be embedded into routine psychotherapeutic care in contexts characterized by professional autonomy and organizational fragmentation [26].

Funding

This study was funded by PRIN (Progetti di Ricerca di Interesse Nazionale [Italian National Research Program]) 2022. No author has received personal fees or other payments from commercial entities with an interest in the results.

Data Availability

This manuscript reports a study protocol; therefore, no datasets have yet been generated or analyzed at the time of submission. Upon completion of data collection, quantitative data derived from the Evidence-Based Practice Attitude Scale–36 (EBPAS-36) questionnaire will be anonymized and may be made available upon reasonable request from the corresponding author, subject to ethical approval and data protection regulations.

Qualitative interview data will not be publicly shared due to the sensitive nature of the information collected, the professional identifiability of participants, and restrictions imposed by the informed consent procedure and the General Data Protection Regulation. Access to deidentified qualitative excerpts or analytic materials may be considered on a case-by-case basis for bona fide research purposes, subject to approval by the research team and, where required, the relevant ethics committee.

No custom code was developed for this study. Study materials, including the interview guide, are provided as multimedia appendices.

Authors' Contributions

Conceptualization: MVO, DTB, CM
Methodology: MVO, DTB, CM
Investigation: MVO, ALT
Data curation: MVO, ALT
Formal analysis: MVO
Writing – original draft: MVO, VO
Writing – review & editing: MVO, ALT, VO, AB, GLC, AC, DTB, CM
Project administration: MVO
Supervision: GLC, AC, DTB, CM

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview protocol.
[DOCX File, 27 KB - resprot_v15i1e82837_app1.docx]

Multimedia Appendix 2
Summary of insights from the pilot interviews by Consolidated Framework for Implementation Research domain.
[DOCX File, 26 KB - resprot_v15i1e82837_app2.docx]

References

1. Barkham M, De Jong K, Delgadillo J, Lutz W. Routine outcome monitoring (ROM) and feedback: research review and recommendations. Psychother Res 2023 Sep;33(7):841-855. [doi: 10.1080/10503307.2023.2181114] [Medline: 36931228]
2. Lambert MJ, Hansen NB, Finch AE. Patient-focused research: using patient outcome data to enhance treatment effects. J Consult Clin Psychol 2001 Apr;69(2):159-172. [doi: 10.1037/0022-006X.69.2.159] [Medline: 11393594]
3. Mackrill T, Sørensen KM. Implementing routine outcome measurement in psychosocial interventions – a systematic review. Eur J Soc Work 2020 Sep 2;23(5):790-808. [doi: 10.1080/13691457.2019.1602029]

XSL•FO
RenderX

4. Bickman L, Douglas SR, De Andrade ARV, et al. Implementing a measurement feedback system: a tale of two sites. Adm Policy Ment Health 2016 May;43(3):410-425. [doi: 10.1007/s10488-015-0647-8] [Medline: 25876736]
5. Gual-Montolio P, Martínez-Borba V, Bretón-López JM, Osma J, Suso-Ribera C. How are information and communication technologies supporting routine outcome monitoring and measurement-based care in psychotherapy? A systematic review. Int J Environ Res Public Health 2020 May 2;17(9):3170. [doi: 10.3390/ijerph17093170] [Medline: 32370140]
6. Lewis CC, Boyd M, Puspitasari A, et al. Implementing measurement-based care in behavioral health: a review. JAMA Psychiatry 2019 Mar 1;76(3):324-335. [doi: 10.1001/jamapsychiatry.2018.3329] [Medline: 30566197]
7. de Jong K, Conijn JM, Gallagher RAV, Reshetnikova AS, Heij M, Lutz MC. Using progress feedback to improve outcomes and reduce drop-out, treatment duration, and deterioration: a multilevel meta-analysis. Clin Psychol Rev 2021 Apr;85:102002. [doi: 10.1016/j.cpr.2021.102002] [Medline: 33721605]
8. Miller SD, Hubble MA, Chow D, Seidel J. Beyond measures and monitoring: realizing the potential of feedback-informed treatment. Psychotherapy (Chic) 2015 Dec;52(4):449-457. [doi: 10.1037/pst0000031] [Medline: 26641375]
9. Boswell JF, Hepner KA, Lysell K, et al. The need for a measurement-based care professional practice guideline. Psychotherapy (Chic) 2023 Mar;60(1):1-16. [doi: 10.1037/pst0000439] [Medline: 35771518]
10. Bond RR, Mulvenna MD, Potts C, O'Neill S, Ennis E, Torous J. Digital transformation of mental health services. Npj Ment Health Res 2023 Aug 22;2(1):13. [doi: 10.1038/s44184-023-00033-y] [Medline: 38609479]
11. Menachemi N, Collum TH. Benefits and drawbacks of electronic health record systems. Risk Manag Healthc Policy 2011;4(47–55):47-55. [doi: 10.2147/RMHP.S12985] [Medline: 22312227]
12. Schiller B, Kuska M, Becher-Urbaniak S, Wimmer E, Reisinger M, Mörtl K. Online psychotherapy as a first clinical experience during the Covid-19 pandemic: a new generation of psychotherapists in the digital age. Heliyon 2024 Apr 30;10(8):e29464. [doi: 10.1016/j.heliyon.2024.e29464] [Medline: 38638950]
13. Singla DR. Digital mental health interventions: an opportunity to enhance care while preserving human connection. Lancet Reg Health Am 2024 Nov;39:100940. [doi: 10.1016/j.lana.2024.100940] [Medline: 39559429]
14. Smoktunowicz E, Barak A, Andersson G, et al. Consensus statement on the problem of terminology in psychological interventions using the internet or digital components. Internet Interv 2020;21:100331. [doi: 10.1016/j.invent.2020.100331] [Medline: 32577404]
15. The Lancet Digital Health. Mental health in the digital age. Lancet Digit Health 2022 Nov;4(11):e765. [doi: 10.1016/S2589-7500(22)00191-1] [Medline: 36229344]
16. Saxon D, Barkham M, Bee P, Gellatly J, Faija C, Bower P. Telephone treatments in improving access to psychological therapies services: an analysis of use and impact on treatment uptake. BMC Psychiatry 2023 Feb 7;23(1):95. [doi: 10.1186/s12888-022-04404-1] [Medline: 36750788]
17. Wright CV, Goodheart C, Bard D, et al. Promoting measurement-based care and quality measure development: the APA mental and behavioral health registry initiative. Psychol Serv 2020 Aug;17(3):262-270. [doi: 10.1037/ser0000347] [Medline: 31021113]
18. Lampis J, Rocca G. The use of routine outcome monitoring (ROM) among the psychotherapists of the Italian National Health Service. Clin Psychol Psychother 2024;31(6):e70002. [doi: 10.1002/cpp.70002] [Medline: 39510978]
19. de Girolamo G, Bassi M, Neri G, Ruggeri M, Santone G, Picardi A. The current state of mental health care in Italy: problems, perspectives, and lessons to learn. Eur Arch Psychiatry Clin Neurosci 2007 Mar;257(2):83-91. [doi: 10.1007/s00406-006-0695-x] [Medline: 17200877]
20. Adduci A, Belvis AG. Response to the mental health crisis in Italy: the "Psychologist bonus". Health Systems and Policy Monitor (HSPM). 2023 Jul 17. URL: https://eurohealthobservatory.who.int/monitors/health-systems-monitor/updates/hspm/italy-2023/response-to-the-mental-health-crisis-in-italy-the-psychologist-bonus [accessed 2025-04-30]
21. Italy: health system review 2022. : World Health Organization Regional Office for Europe, European Observatory on Health Systems and Policies; 2022 URL: https://iris.who.int/server/api/core/bitstreams/07cc8a18-5dad-4260-b2a8-a3042d29bb6f/content [accessed 2025-04-30]
22. Olive MV, La Tona A, Lo Coco G, Compare A, Cafazzo JA, Masella C. Assessment of a digital platform for routine outcome monitoring in psychotherapy: usability study and thematic analysis. JMIR Med Inform 2025 Sep 30;13:e75885. [doi: 10.2196/75885] [Medline: 41027026]
23. Åhs JW, Ranheim A, Eriksson H, Mazaheri M. Encountering suffering in digital care: a qualitative study of providers' experiences in telemental health care. BMC Health Serv Res 2023 May 1;23(1):418. [doi: 10.1186/s12913-023-09367-x] [Medline: 37127655]
24. Wainberg ML, Gouveia ML, Stockton MA, et al. Technology and implementation science to forge the future of evidence-based psychotherapies: the PRIDE scale-up study. Evid Based Ment Health 2021 Feb;24(1):19-24. [doi: 10.1136/ebmental-2020-300199] [Medline: 33177149]
25. Youn SJ, Jaso B, Eyllon M, et al. Leveraging implementation science to integrate digital mental health interventions as part of routine care in a practice research network. Adm Policy Ment Health 2024 May;51(3):348-357. [doi: 10.1007/s10488-023-01292-9] [Medline: 37615809]
26. McGinty EE, Alegria M, Beidas RS, et al. The Lancet Psychiatry Commission: transforming mental health implementation research. Lancet Psychiatry 2024 May;11(5):368-396. [doi: 10.1016/S2215-0366(24)00040-3] [Medline: 38552663]

XSL•FO
RenderX

27. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
28. Kirk MA, Kelley C, Yankey N, Birken SA, Abadie B, Damschroder L. A systematic review of the use of the Consolidated Framework for Implementation Research. Implement Sci 2016 May 17;11:72. [doi: 10.1186/s13012-016-0437-z] [Medline: 27189233]
29. Ware P, Ross HJ, Cafazzo JA, Laporte A, Gordon K, Seto E. Evaluating the implementation of a mobile phone-based telemonitoring program: longitudinal study guided by the Consolidated Framework for Implementation Research. JMIR mHealth uHealth 2018 Jul 31;6(7):e10768. [doi: 10.2196/10768] [Medline: 30064970]
30. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. Psychiatry Res 2019 Oct;280:112516. [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
31. Christensen A, Jacobson NS. Who (or what) can do psychotherapy: the status and challenge of nonprofessional therapies. Psychol Sci 1994 Jan;5(1):8-14. [doi: 10.1111/j.1467-9280.1994.tb00606.x]
32. Borsci G. La regolamentazione della psicoterapia in Italia: storia della legge 56/1989 e stato attuale delle scuole riconosciute [Article in Italian]. Psicoter Sci Um 2005;39(2):193-222 [FREE Full text]
33. Bosio AC, Lozza E. La costruzione sociale delle professioni psicologiche in Italia: percorsi e agenda building [Article in Italian]. Giorn Ital Psychol 2021;48(2):257-276. [doi: 10.1421/102688]
34. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. Dubois A, Gadde LE. Systematic combining: an abductive approach to case research. J Bus Res 2002 Jul;55(7):553-560. [doi: 10.1016/S0148-2963(00)00195-8]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
EBPAS-36: Evidence-Based Practice Attitude Scale–36
ROM: routine outcome monitoring

Edited by J Sarvestan; submitted 22.Aug.2025; peer-reviewed by A Kurilla, C Remondi, I Benzi, P Bjurner; accepted 21.Jan.2026; published 02.Mar.2026.

Please cite as:
Olive MV, La Tona A, Ostwald V, Brugnera A, Lo Coco G, Compare A, Tzur Bitan D, Masella C
Digitally Enhanced Routine Outcome Monitoring in Italian Psychotherapy: Protocol for a Longitudinal Implementation Study
JMIR Res Protoc 2026;15:e82837
URL: https://www.researchprotocols.org/2026/1/e82837
doi:10.2196/82837

© Mattia Vincenzo Olive, Antonino La Tona, Vivian Ostwald, Agostino Brugnera, Gianluca Lo Coco, Angelo Compare, Dana Tzur Bitan, Cristina Masella. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Real-Time Type 1 Diabetes Self-Management Decision-Making in Adolescents: Protocol for a Longitudinal Mixed Methods Study Using Text Messaging and Continuous Glucose Monitoring

Melissa DeJonckheere[1,2], PhD; Samantha A Chuisano[2], MPH; Juniar Lucien[2], MSc; Fouzaan Amjad[2]; Oorvi Duvvuri[2]; Hasan Khan[2], BS, MBA; Maryam Khan[2]; Rafee Mirza[2], BS; Maya Joy Ollivierre[2]; Timothy Guetterman[2], PhD; Yu Kuei Lin[3], MD; Lorraine R Buis[2,4], MSI, PhD; James E Aikens[2], PhD; Caroline Richardson[2], MD; Joyce M Lee[5,6,7], MPH, MD

[1]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States

[2]Department of Family Medicine, University of Michigan, 1018 Fuller Street, Ann Arbor, MI, United States

[3]Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States

[4]School of Information, University of Michigan, Ann Arbor, MI, United States

[5]Department of Pediatrics, University of Michigan, Ann Arbor, MI, United States

[6]Susan B. Meister Child Health Evaluation and Research Center, University of Michigan, Ann Arbor, MI, United States

[7]Elizabeth Weiser Caswell Diabetes Institute, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Melissa DeJonckheere, PhD
Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States

Abstract

Background: Type 1 diabetes (T1D) requires repeated self-management behaviors and ongoing problem-solving to maintain optimal glucose levels and prevent complications. Despite increasing adoption of continuous glucose monitoring (CGM), which can alleviate some of the constant self-management burden, adolescents struggle to achieve glycemic recommendations and report low engagement with diabetes device data. Previous studies have used retrospective or quantitative approaches to describe adolescent self-management; however, it is unclear how psychosocial influences (eg, mood and distress) and contexts impact adolescent self-management behaviors and engagement with their diabetes devices in everyday life. Exploration of real-time experiences will help to identify potential targets and strategies for future interventions to improve glycemic outcomes in adolescents with T1D using advanced diabetes technologies.

Objective: This study has two aims: (1) to develop a grounded theory of self-management decision-making using diabetes devices among adolescents with T1D and (2) to assess the acceptability and feasibility of longitudinal and real-time qualitative data collection methods in this population.

Methods: We will conduct a mixed methods study informed by the capability, opportunities, and motivation of behavior model. Adolescents (aged 12 - 18 y) with T1D who regularly use CGMs will be recruited from a Midwest pediatric diabetes clinic. Purposive sampling strategy will ensure participants with varied glycemic levels (hemoglobin A_{1c} [HbA_{1c}] ≤9% and HbA_{1c} >9%) and diabetes experiences (eg, diabetes duration, devices used) are included. Using a longitudinal convergent mixed methods design, enrolled participants (n=30 - 40) will complete data collection over 6 weeks including: (1) a baseline survey to capture demographic, clinical, and behavioral characteristics; (2) 30 days of SMS text messaging surveys to describe real-time self-management behaviors, technology use, and decision-making; (3) 30 days of CGM data; and (4) an interview focused on self-management behaviors and technology use. Recruitment will continue until appropriate data completeness and/or theoretical saturation is achieved. Analysis of text responses and interview transcripts will follow a grounded theory approach. Summarized glycemic metrics (eg, time in range) and visuals (ie, ambulatory glucose profile) will be integrated with qualitative findings through participant profiles and joint displays. Integrated findings will be used to refine a grounded theory of daily self-management decision-making using diabetes devices among adolescents with T1D.

Results: As of December 2025, 25 participants have enrolled in this study. We expect SMS text messaging survey completion rates and CGM use near 70% throughout the study period. We anticipate findings to become available in the following several years through conference presentations and peer-reviewed publications.

Conclusions: While routine diabetes self-management behaviors and use of diabetes technologies are important for achieving glycemic goals, adolescents report low adherence to diabetes devices. This real-time mixed methods study will improve our understanding of daily decision-making and influences on diabetes self-management. Findings from this study will identify

XSL•FO
RenderX

facilitators and barriers to optimal T1D self-management. In addition, results will inform future studies using real-time qualitative and mixed methods approaches.

International Registered Report Identifier (IRRID): DERR1-10.2196/83218

(JMIR Res Protoc 2026;15:e83218) doi:10.2196/83218

KEYWORDS

type 1 diabetes; adolescent; youth; continuous glucose monitoring; CGM; diabetes; self-management; SMS text messaging; qualitative research; mixed methods

Introduction

Type 1 diabetes (T1D) is one of the most common chronic conditions among children globally, affecting 3.5 per 1000 youth in the United States [1,2]. T1D requires intensive daily management of blood glucose levels to reach hemoglobin A_{1c} (HbA_{1c}) targets and minimize the risk of severe short- and long-term complications, including diabetic ketoacidosis, neuropathy, retinopathy, and cardiovascular disease [3]. In 2021, the American Diabetes Association recommended glycemic targets of less than 7% for most children and adolescents [4], yet only 19% of youth met this target in 2021 to 2023 [5]. As HbA_{1c} typically peaks in adolescence and young adulthood [6], it is a critical period to address barriers to self-management to reach optimal glycemic levels and prevent common complications that can persist through adulthood [7-9].

Elevated glycemic levels in adolescence result from the complex interplay of physical, emotional, and cognitive development in addition to social and structural determinants of health (eg, access to medical care and health care costs) [3,10-12]. Regular T1D self-management requires frequent blood glucose monitoring, regular insulin use, monitoring carbohydrate intake, modifying insulin use to match diet and activity levels, and troubleshooting unpredictable blood glucose levels. As a result, executive functioning (including planning, decision-making, problem-solving, and reflecting on blood glucose patterns and behaviors) plays a critical role in self-management [6,13,14]. As adolescents begin to take on increased responsibility for their own diabetes care and seek independence from their families, they often struggle to consistently perform these complex self-management behaviors [15,16]. Unsurprisingly, one-third of adolescents with T1D experience high levels of diabetes distress [17,18], which has been linked to poorer adherence to self-management and higher glycemic levels in adolescents [17,19].

Advanced diabetes technologies such as continuous glucose monitoring (CGM) may lessen the burden of diabetes self-management by providing real-time glucose values, alerting adolescents and caregivers to low and high blood glucose levels, and indicating the predicted direction of blood glucose levels in the next hour [20,21]. As of 2023, nearly 80% of adolescents with T1D use CGM [1]; however, the average HbA_{1c} among CGM users in this age group remains above the recommended level [22], suggesting that CGM alone is not enough for youth to meet HbA_{1c} targets.

While adolescents who regularly use CGM report high levels of satisfaction and quality of life [23], they also report concerns about device visibility, feeling "spied on," alarm and device fatigue, feeling as though device data are unreliable or untrustworthy, and being overwhelmed by constant streams of data [24]. The emotional, social, and practical burden of CGM can contribute to distress, reduced motivation for device engagement, and, ultimately, lower device adherence [25]. Fewer than 30% of individuals with T1D report using any real-time CGM features such as event logging or data sharing, and less than 75% have customized their high and low alert thresholds [26]. While reviewing CGM data may support problem-solving and decision-making among adolescents with T1D [27,28], fewer than 20% of adolescents review their diabetes device data independently [23]. Additional research is needed to understand how adolescents are leveraging (or not leveraging) diabetes devices for self-management decisions in daily life, and how the use of diabetes devices impacts their overall health and well-being. Understanding how adolescents use their devices can help health care teams and families to identify key barriers to self-management and support young people to optimize self-management and care.

The overarching purpose of the Qualitative Insights to Type 1 Diabetes in Youth (QUALITY) study described here is to identify potential targets and strategies for future interventions to improve glycemic outcomes in adolescents with T1D using advanced diabetes technologies. The study has two aims: (1) to develop a grounded theory of self-management decision-making using diabetes devices among adolescents with T1D and (2) to assess the acceptability and feasibility of longitudinal and real-time qualitative data collection methods in this population. This observational study leverages a longitudinal, convergent mixed methods design, merging real-time narrative and CGM data to understand daily diabetes self-management behaviors, decision-making, and problem-solving among adolescents with T1D who use CGM. To our knowledge, this is the first application of a real-time and longitudinal mixed methods approach among adolescents with T1D. Thus, a secondary aim of the study is to assess the acceptability and feasibility of the proposed methods.

Methods

Overview

This longitudinal, convergent mixed methods study leverages a real-time approach to develop a theory of self-management among adolescents with T1D who use CGM. We will conduct a grounded theory analysis to describe self-management in this

XSL•FO
RenderX

population, which will inform future intervention development. We aim to enroll approximately 30 to 40 adolescents with T1D who already use CGM to complete 30 days of real-time, narrative data collection via SMS text messaging and 30 days of CGM. Additional data will be collected via a baseline survey and one-time follow-up interview to further develop the theory.

When applied to diabetes, real-time data collection (eg, ecological momentary assessment) can allow us to more accurately assess T1D self-management beliefs, behaviors, and event experiences, reduce recall bias, help researchers to understand behaviors in context, and support the development of timely and tailored interventions [29-32]. With the growth of CGM among adolescents, continuous and real-time glucose data have been incorporated into momentary assessments [1,33] While real-time quantitative assessments help to characterize important variations that influence self-management, the broader context of an individual's experience during those moments remains poorly understood. A mixed methods approach blending real-time and contextual data could address these gaps to provide a more comprehensive understanding of adolescents' experiences and identify potential intervention targets and strategies. In this study, we will use a qualitative real-time approach to further explore the lived experience of T1D self-management from the perspectives of adolescents using diabetes technology.

Conceptual Model

The QUALITY study is informed by the capability, opportunity, and motivation model of behavior (COM-B) [34] by assessing several relevant influences on adolescent self-management behaviors and exploring these behaviors through the routine, reactive, and reflective cycles that are performed in real-time by adolescents with T1D using diabetes devices. The COM-B model [34] serves as a helpful framework for describing T1D self-management behaviors in adolescents who use diabetes technology. According to the COM-B model, to perform a particular behavior, an individual must feel that they are capable of doing so (physically and psychologically), have the opportunity to do so (socially and physically), and be motivated to do so (automatically and reflectively) [34]. COM-B has frequently been used as a tool to develop intervention targets to change individual behaviors and can also be used as a framework to explore and describe influences on diabetes self-management [35-37]. For example, in the context of T1D self-management, the behavior of "checking blood glucose levels" would require the knowledge and skills needed to test through finger pricks or look at a CGM display (ie, capability), access to glucometers or CGM and comfort checking blood glucose levels in front of others (ie, opportunity), and the belief that testing blood glucose levels is desirable and impactful on health (ie, motivation).

T1D requires many interconnected behaviors to optimize self-management. Through an expert panel and literature review, Hamilton et al [38] identified 150 distinct self-management behaviors related to T1D that are performed in three self-regulatory behavioral cycles: (1) routine, (2) reactive, and (3) reflective. Each of these cycles is supported by or inhibited by capability, opportunities, and motivation for a given behavior [38]. Re-examining the example above of checking blood glucose levels, an adolescent with T1D may decide to check their glucose using CGM as a part of their daily self-management routine (ie, routine behavior), in response to hypoglycemia or hyperglycemia symptoms or a CGM alarm (ie, reactive behavior), or by reviewing trends in glucose available on their insulin pump, CGM receiver, or CGM mobile app to make later adjustments (ie, reflective behavior). COM-B highlights the complexity of improving self-management behaviors in adolescence, as each individual behavior is influenced by the interplay of capabilities, opportunities, and motivations in routine, reactive, and reflective contexts which occur every day, in real-time, and in diverse settings. Figure 1 presents the adapted COM-B used to inform the QUALITY study.

To ensure a youth-centered design and improve engagement with this study, the study team will make iterative changes to the study design and data collection tools based on feedback from an online patient advisory council. The Type 1 Diabetes Youth Advisory Council (T1DYAC) was established in 2024 to guide T1D research conducted primarily in our laboratory. T1DYAC members are youth (ages 14 - 24 y) with T1D who have an interest in improving research about T1D in youth through discussion with research teams. T1DYAC members represent a range of ages, diabetes durations and histories, and diabetes devices used. During monthly virtual meetings, the group is presented with a research study and provides feedback on a specific study element (eg, recruitment flyers and approach, research questions, and interpretation of findings). In addition, members complete surveys between monthly meetings to provide written feedback and pointed insights on study materials.

XSL•FO
RenderX

Figure 1. Capability, opportunity, and motivation of behavior (COM-B) conceptual model adapted for Qualitative Insights to Type 1 Diabetes in Youth (QUALITY) study. CGM: continuous glucose monitoring; T1D: type 1 diabetes.

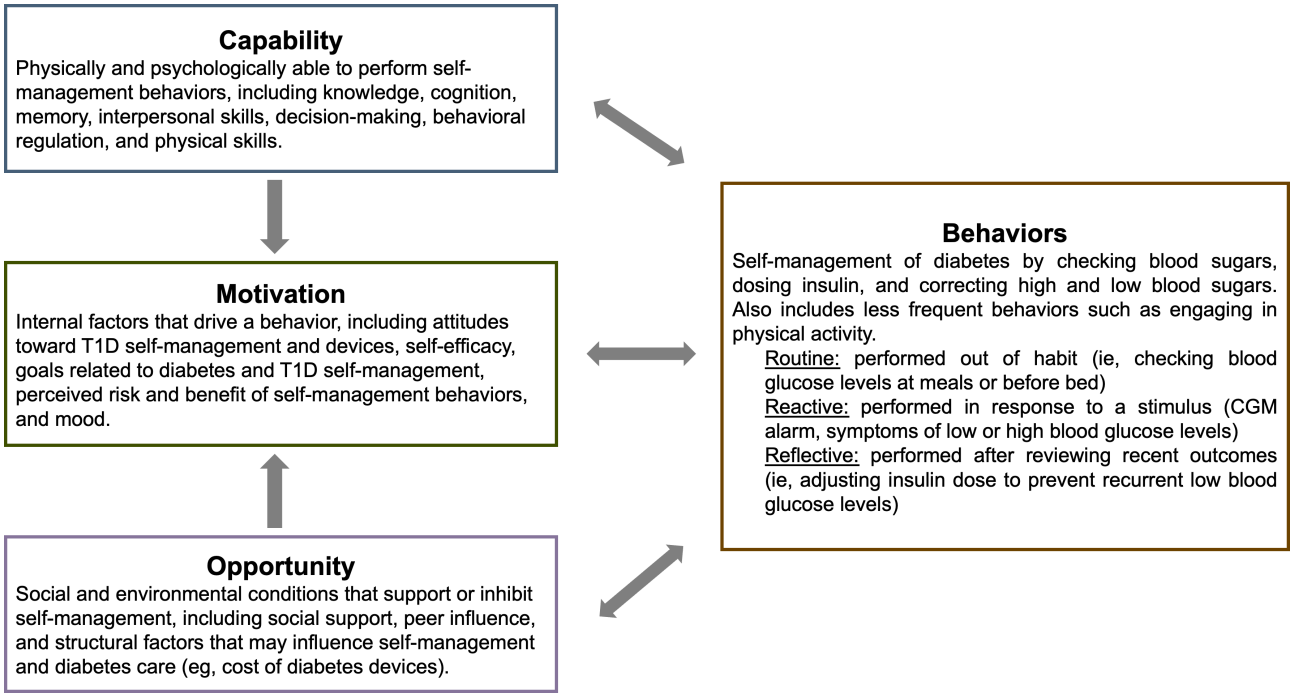

Setting

This study will recruit patients from the University of Michigan (U-M) Pediatric Diabetes Clinic at C.S. Mott Children's Hospital. As of December 2025, there are approximately 700 current patients aged 12 to 18 years with T1D; among these, 357 (50%) patients are male, 554 (79%) are White, 85 (12%) are Black or African American, 12 (2%) are Asian, 36 (5%) are more than one or another race, and 36 (5%) are Hispanic or Latino. In addition, 417 of these patients (62%) have public insurance, and 687 (98%) are English speaking. Likewise, over 92% currently use CGM. The makeup of the clinical population is similar to that of the US population of people with T1D, where 68% are non-Hispanic White, 16% African American, 13% Hispanic, 3% other races, and nearly 80% use CGM [39]. While disparities persist, differences in CGM by age, race and ethnicity, and insurance type have narrowed significantly in the last 15 years [2,5].

Recruitment and Sampling

A list of patients with T1D, aged 12 to 18 years, who (1) use CGM, (2) received care in the U-M Pediatric Diabetes Clinic in the past 12 months, and (3) have at least 1 HbA_{1c} laboratory result from the past 12 months will be generated using DataDirect (University of Michigan Medical School) and Electronic Medical Record Search Engine (EMERSE) [40]—university-developed tools for performing systematic searches of structured and unstructured patient medical records [41]. The list will include basic demographic, clinical, and contact information for identified patients.

Consistent with a grounded theory approach, we will use theoretical sampling to help ensure that adequate data are collected to address our research aims. For our initial sampling strategy (a priori) [42], we seek to enroll participants with diverse diabetes outcomes and experiences. We will intentionally recruit and enroll participants with HbA_{1c} of 9% or less and more than 9% to ensure adolescents with low and high glycemic levels are represented. We aim to recruit and enroll approximately equal groups with high and low glycemic levels. A purposive sampling approach will also be leveraged using demographic and clinical factors including age, gender, years since diagnosis, and other social determinants of health (eg, family income and food insecurity measures) to create a sample with varied experiences and backgrounds. Sociodemographic backgrounds that are under-represented within the clinical population (ie, Hispanic or Latino) in the sample will be intentionally oversampled. As data collection and analysis proceed, recruitment and sampling strategies may shift to explore variation in participant responses, develop relationships between codes and categories, and add details to the descriptions and characteristics of categories in the analysis [42].

Recruitment will be conducted in waves to allow for flexibility and needed changes to meet theoretical sampling goals. During each wave, approximately 100 to 200 identified patients will be contacted via electronic and/or postal mail up to 5 times with study information and a link to an online screening survey to confirm interest and eligibility. Study information will also be shared through (1) flyers and patient newsletters posted by the U-M Pediatric Diabetes Clinic, (2) UMHealthResearch.org, a secure platform at U-M for sharing study information with potential participants, and (3) emails and newsletters shared with patient diabetes councils. If recruitment goals are not reached through these methods, patients may also be approached in person at the U-M Pediatric Diabetes Clinic by nonclinician members of the study team.

Study eligibility criteria include: (1) current patient at Michigan Medicine with at least 1 visit and HbA_{1c} measurement in the last 6 months, (2) aged 12 to 18 years upon enrollment, (3)

XSL•FO
RenderX

diagnosed with T1D for at least 6 months, (4) used CGM for at least 6 months, (5) used CGM for at least 20 days per month, (6) able to communicate in English, and (7) have access to a phone with SMS text messaging capabilities. Eligibility will be assessed via the screening survey and confirmed via medical record review.

A sample size estimate (N=30 - 40) was determined a priori based on estimates of the amount of qualitative data needed to reach data sufficiency across participants with varied glycemic levels and diabetes experiences [43,44]. Due to the longitudinal nature of data collection and multiple data sources in this study, we anticipate having sufficient data to inform our grounded theory. As is common in qualitative research and grounded theory, additional data may be collected as needed to further develop findings and reach theoretical saturation [45]. The same sample will be used for quantitative data collection, including baseline surveys of diabetes characteristics and psychosocial measures, and 30 days of CGM data. Using identical samples in a convergent mixed methods design reduces threats to validity (eg, sample size legitimation) and ensures that appropriate conclusions can be drawn through integration [46,47].

Ethical Considerations

All participants (and a parent or guardian for minors) will complete an enrollment call to review the informed consent document with a member of the study team and provide electronic or written consent or assent to enroll in the study. Throughout the study, survey invitations will include the study team's contact information in the event of any questions or concerns. Participants will also receive contact information for the pediatric diabetes clinic and reminders to reach out to their diabetes care team or the provider on call with diabetes-related questions, concerns, and emergencies. In addition, participants will be instructed not to share any identifiable information through SMS text messaging (ie, name and school) as it is not a secure method of communication. All incoming survey responses will be stored securely, and the exported data will be deidentified for analysis. Participants can earn up to US $150 for completing all study activities, and final payments will be sent as a gift card or a check in the mail, depending on the participant's preference. This study was reviewed and approved by the U-M Institutional Review Board (HUM00230299) prior to beginning recruitment.

Data Collection Procedures

Overall study participation will last approximately 6 weeks. Participants will complete (1) a one-time baseline survey to collect participant demographics, diabetes history, and psychosocial influences on T1D self-management; (2) 30 days of daily SMS text messaging surveys and CGM to evaluate real-time psychosocial influences on T1D self-management and glycemic levels; and (3) an optional semistructured interview to gain additional insights on T1D self-management approaches. Following enrollment and prior to data collection, participants will receive basic training on how to complete the daily surveys using a practice survey that includes sample questions with various formats (open-ended and closed-ended) and provides examples of survey prompts and outro messages. All survey data will be collected and managed using a secure REDCap (Research Electronic Data Capture; Vanderbilt University) database [48,49]. A study timeline and overview of data collection procedures for each participant are presented in Figure 2.

Figure 2. Qualitative Insights to Type 1 Diabetes in Youth (QUALITY) study participant flow. The asterisks indicate the data collection procedures related to study feasibility and acceptability. CGM: continuous glucose monitoring; T1D: type 1 diabetes.

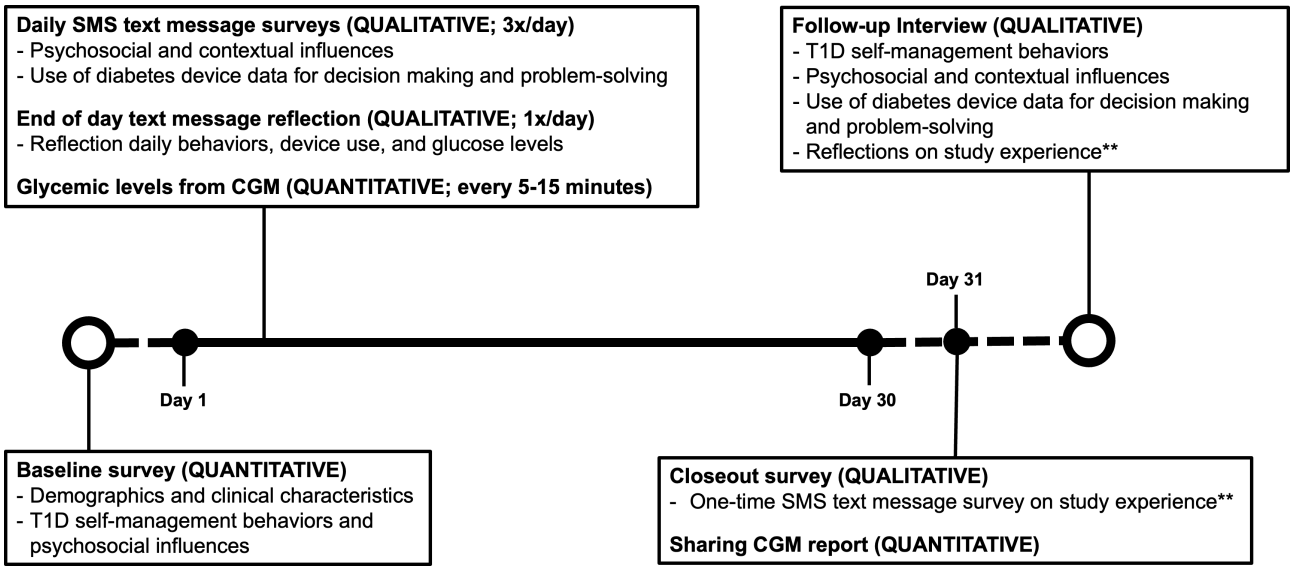

Demographics and Clinical Characteristics

Participant demographic information and clinical characteristics will be collected upon enrollment using an online survey. Demographic measures include age, race and ethnicity, gender, and student and/or employment status. Clinical characteristics include age at diabetes diagnosis, recent (within the last 12 mo) hospitalization for diabetes-related complications, typical medical care for diabetes, and social support for diabetes management. Finally, questions on diabetes technology use include which type and brand of CGM and insulin pump are used, for how long, and typical use (how many d in the mo and use of automated insulin delivery). The participant demographics

XSL•FO
RenderX

and clinical characteristics items are included in Multimedia Appendix 1.

Baseline Survey

Study participants will complete a one-time baseline survey. Measures were selected to align with components of the COM-B model and the literature on psychosocial influences on self-management and diabetes device use. The survey includes measures of diabetes self-management behaviors, depression, diabetes distress, attitudes towards diabetes technology, problem-solving, and self-efficacy with diabetes self-management. A summary of survey instruments, sample items, and their alignment with the COM-B model is included in Table 1. Additional details can be found in Multimedia Appendix 2.

XSL•FO
RenderX

Table . Alignment of capability, opportunities, and motivation of behavior (COM-B) model with select data collection measures.

COM-B construct	Baseline survey items	Daily and end-of-day survey items	Semistructured interview questions
Capability (physically and psychologically able, including knowledge, decision-making, and skills)	• Self-efficacy for T1D[a] management: "[Can you…] adjust your insulin or food accurately based on how much exercise you get." [50]. • CGM[b] self-efficacy: "I am sure I can... adjust insulin dose based on real-time data" [51].	• Did you look at information from your diabetes devices in the last few hours? If no, why not?	• What do you consider to be a high blood sugar? • Why is this number significant to you?
Opportunity (social and environmental conditions, including social support and structural factors)	• Clinical characteristics: For how long have you used CGM? How many days in the last month did you wear a CGM sensor or use CGM?	• What problems, if any, did you encounter today? • Were your diabetes devices useful to you today? How?	• What support do you get from your health care providers that help you to use your technology?
Motivation (internal factors, including mood, attitudes, and beliefs)	• Diabetes Technology Attitudes Scale: "Diabetes technology has made my life better" [52]. • Benefits and burdens of CGM: "CGM alarms are helpful" and "CGM sensor readings cannot be trusted" [53]. • Self-efficacy for T1D management: "[Taking care of my diabetes would] be too much responsibility" [50]. • Problem areas in diabetes—teen: "Not feeling motivated to keep up with my daily diabetes tasks" [54]. • Patient Health Questionnaire (2): "Feeling little interest or pleasure in doing things" [55].	• How did you feel after looking at information from your diabetes device(s)? • What else impacted how you took care of your diabetes today?	• How confident do you feel making decisions in response to your current blood sugar?
Behavior			
Routine	• Diabetes Management Questionnaire: "…adjust your food or insulin before long periods of physical activity?" [56]	• Did you look at information from your diabetes devices in the last few hours? (If no) Why not?	• Tell me about your general day-to-day approach to taking care of your diabetes.
Reactive	• N/A[c]	• What did you learn from your diabetes device(s)? What did you do next? Or how did you use the information?	• When was the last time that you had a high blood sugar? Once you realized, what did you do next?
Reflective	• Diabetes Adolescent Problem-Solving Questionnaire: "When I think a situation will get in the way of taking care of my diabetes, I do something to prevent that from happening" [57].	• Use the daily graph in your CGM graph to review your blood sugars over the last 24 hours. Pick a time period you want to tell us more about (eg, 8 AM to 10 AM). What time did you pick and why? Walk us through what you think impacted your blood glucose levels during this time of day. Do you notice any patterns between your blood glucose levels in the past week? Tell us about it.	• What do you do to try to make CGM more helpful to you?

[a]T1D: type 1 diabetes.

[b]CGM: continuous glucose monitoring.

[c]N/A: not applicable.

XSL•FO
RenderX

Real-Time Daily Data Collection

SMS text messaging survey data will be collected using the "survey as a text conversation" feature in REDCap [48,49], which mimics a typical SMS text messaging conversation where questions are sent one-at-a-time and each question is sent only after a response is received. Surveys will automatically close after completion or after 3 hours if no response is received, and the next survey will automatically be sent at the prescheduled time. SMS text messaging was selected for daily surveys based on discussions with T1DYAC members, the high prevalence of smartphone use among US teens (95%) [58], and high rates of smartphone initiation by the age of 12 years (75%) [59]. In addition, the daily and end-of-day surveys were designed to have no more than 10 questions at a time to reduce the overall participant burden.

Beginning after the baseline survey is completed, participants will receive SMS text messaging survey invitations 4 times per day on a fixed schedule in the morning (10 AM), afternoon (2 PM), early evening (6 PM), and end-of-day (10 PM) [3] The morning, afternoon, and early evening surveys focus on interactions with their diabetes technology, challenges in managing diabetes, and other events that have occurred in the last few hours. For example, 1 question asks, "What have you been doing in the last few hours (ie, exercising, eating, at school, etc)?" In the evenings (10 PM), participants will complete a guided reflection using open-ended questions. During the reflection, they will be asked to open their diabetes device app and review the graph of their glucose levels throughout the day. Qualitative and quantitative data will be integrated during data collection as participants provide written reflections on their CGM data in daily SMS text messaging surveys. In addition, they will share information about what activities they did that day and how they feel events and activities impacted their diabetes. For example, the first question asks, "Overall, how did you feel about your diabetes today?" Sample daily and end-of-day survey question items and their alignment with the COM-B model can be found in Table 1. Complete daily and end-of-day SMS text messaging surveys are included in Multimedia Appendix 3. Participant response rates will be tracked automatically by REDCap throughout the study. In addition, participants will complete an open-ended text messaging survey after the 30 days of daily surveys are completed to qualitatively share their study experiences, burden, and study acceptability.

During the same 30 days, participants will be asked to continue CGM as prescribed to collect glucose levels. Depending on the device manufacturer, glucose levels are measured automatically every 1 to 15 minutes. By not introducing a new device, reactivity is reduced, and we are able to capture typical self-management behaviors and glycemic levels in this population. Likewise, participants do not have to learn to use a different device solely for the purpose of this study. Glycemic data are not blinded from participants to maintain normalcy with their daily routine and to maintain their safety in managing potentially dangerous high and low glucose levels during the study period. After completion of the daily survey portion of the study, participants will receive a secure REDCap survey link to share their CGM data from the previous 30 days downloaded as a .csv file from their CGM platform (eg, Dexcom Clarity). In addition, participants will be asked to reflect on the feasibility and acceptability of the study through 7 survey questions. Questions include study barriers, facilitators, and reflections on the length and timing of data collection (see Multimedia Appendix 4 for full closeout survey).

Semistructured Interview

A subsample of participants (n=20 - 30) will be invited to complete a 1-hour follow-up interview with a member of the study team trained in qualitative data collection and analysis. A maximum variation sampling approach will be leveraged and rely on recent HbA_{1c} levels, baseline survey scores, age, gender, race or ethnicity, diabetes duration, and diabetes device use as well as responses to the SMS text messaging surveys. Interviews are intended to further explore their approach to self-management using their diabetes devices. Interview prompts will focus on the participants' daily approach to managing their diabetes, including decision-making, problem-solving, and interactions with technology. Alignment of interview prompts with the COM-B model is described in Table 1. Likewise, a portion of the interview will be dedicated to exploring study feasibility and acceptability, including perceived study burden and experiences of behavior change during the study period. Interviews will be conducted using Health Insurance Portability and Accountability Act (HIPAA)-compliant web conference tools (eg, Zoom; Zoom Communication, Inc), audio-video recorded, and transcribed for analysis. In line with a traditional qualitative approach, analysis and data collection will be completed simultaneously, and the final sample size will be determined by theoretical saturation.

Analytic Approaches

Overview

Qualitative (open-ended survey responses and interview transcripts) and quantitative (baseline survey and CGM) data will be analyzed simultaneously and integrated to address the study aims. For aim 1, we will use a constructivist grounded theory approach, integrating findings from the daily survey and interview data with CGM and psychosocial measures. For aim 2, we will use a descriptive approach, integrating findings from the closeout survey and follow-up interview to assess feasibility and acceptability. Figure 3 summarizes the analysis and integration plan across data sources.

XSL•FO
RenderX

Figure 3. Overview of data analysis and integration plan. Red, blue, and purple colors indicate qualitative, quantitative, and mixed analyses, respectively. CGM: continuous glucose monitoring; T1D: type 1 diabetes.

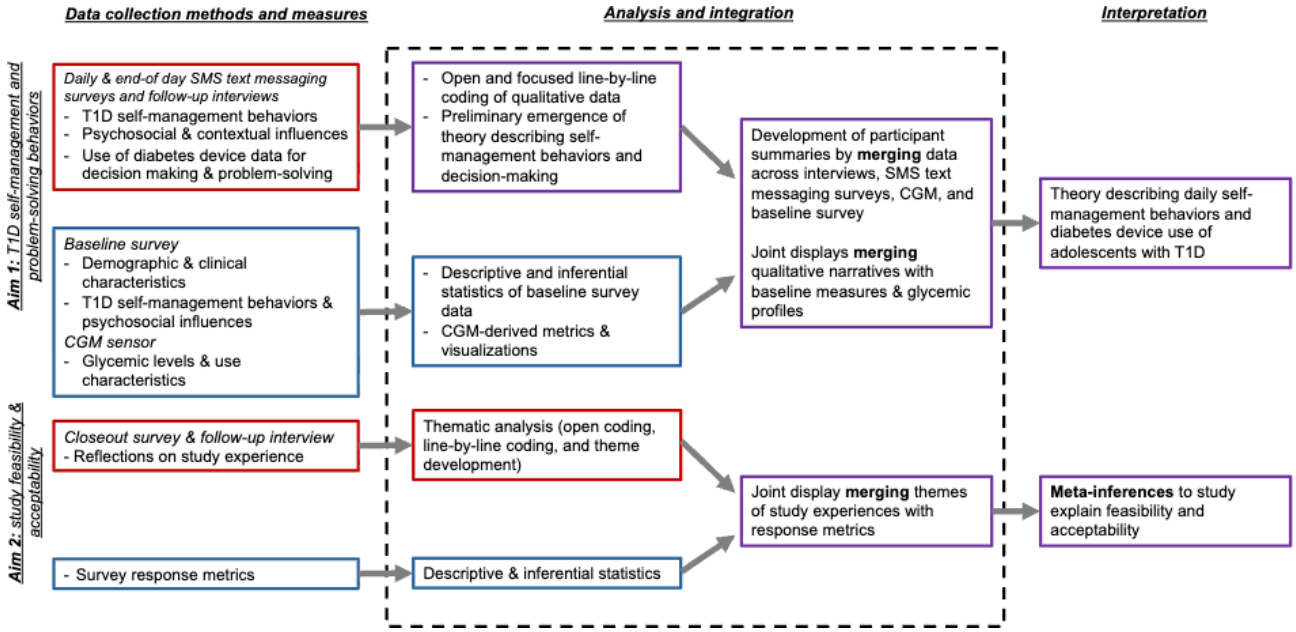

Qualitative Analysis

A .csv file with daily, end-of-day, and closeout survey responses will be downloaded from REDCap for each participant. Columns will include survey timing (eg, morning and afternoon) and study day followed by a column for each survey question. Each row will include responses from a unique survey instance. Interview transcripts and .csv files will be imported and grouped by participant within qualitative and mixed methods data analysis software to support data management (eg, MAXQDA [Max Weber Qualitative Data Analysis]; VERBI GmbH). Participant summaries will be created to merge all data by participant (see also "Mixed Methods Integration" section).

Following a constructivist grounded theory approach informed by Charmaz [42], the principal investigator and research team members will perform at least two iterative rounds of line-by-line coding on the qualitative data: (1) initial coding to identify concepts that are important to participants; and (2) focused coding to synthesize concepts and identify patterns across participants. Theoretical coding can be used to further examine relationships between codes to construct a theory grounded in participant experiences [42]. Throughout the grounded theory analysis, a hybrid inductive-deductive approach will be leveraged, incorporating domains from the COM-B model alongside those identified from participant responses [60] to assist in identifying critical self-management behaviors as they relate to diabetes device use. Findings will be presented as a theory of optimal advanced diabetes technology use for self-management, grounded in the experiences and perspectives of adolescents with T1D. We will refer to COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines Checklist 1to ensure rigor and transparency in reporting [61].

Due to the frequency of data collection in our approach, we anticipate that not all questions will be answered in the real-time daily surveys. We will look for patterns of missing data within and across participants to identify any systematic issues in data collection that may influence analysis and interpretation [62]. Otherwise, any missing data will be treated as a nonresponse. Due to the longitudinal nature of data collection, we anticipate having sufficient responses across daily survey and interview data to reach theoretical saturation. However, additional data collection will occur as needed to further refine the theory [42].

Quantitative Analysis

Baseline survey metrics and clinical characteristics collected from medical records (eg, HbA_{1c}) and participant response rates will be summarized using descriptive statistics. In line with the adapted STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) Checklist for Reporting EMA Studies (CREMAS) from Liao et al [63], we plan to report participant attrition rates according to study days, response latency, overall and daily compliance rates (total responses per participant and average number of responses), and associations of compliance with demographic, clinical, and other time-varying variables.

In this mixed methods study, quantitative measures will be used to characterize the sample and expand understanding of each participant's experience when integrated with qualitative findings. CGM-derived metrics of glycemic control will include those recommended by consensus guidelines from Battelino et al [64] such as sensor use time, time in ranges (ie, above and below), and the glucose management indicator. A minimum of 14 days of CGM data with sensor use 70% of the time has been previously established as sufficient for analysis. The *iglu* R package (R Foundation for Statistical Computing) [65,66] will be used to process and analyze all collected CGM data, and interpolation will be used for small windows of missing data (eg, 45 min). Additional analysis packages (eg, *Best Linear IMPutation*) will be used to impute expected longer periods of missing data (ie, periods of signal loss and sensor malfunction) to reduce bias introduction and evaluate the randomness of missing data and impute missing values, where appropriate [67].

XSL•FO
RenderX

To test for mean differences in psychosocial and behavioral characteristics between glycemic level groups (HbA_{1c}>9% and those with HbA_{1c}≤9%) at baseline, 1-tailed *t* tests will be used. If the data are nonparametric, appropriate alternative tests (ie, the Wilcoxon rank paired test) will be performed. For all participants with at least 28 days of continuous glucose data, we will use linear mixed models to identify risk factors that influence blood glucose levels, model changes in blood glucose over time, and test any intervention effect resulting from longitudinal data collection. Additional statistical tests may be performed as appropriate based on the data collected. Statistical analysis will be performed on the data of all included study participants, and a statistical significance level of 5% (α=.05) will be used to test for differences and associations.

Mixed Methods Integration

Mixed methods findings will help us to better understand how diabetes devices are used for daily T1D self-management among adolescents with T1D and how psychosocial factors and daily self-management behaviors are related. Aligned with the longitudinal, convergent mixed methods design, integration of qualitative and quantitative approaches will occur through sampling, data collection, analysis, and interpretation [68,69]. For example, integration occurs during sampling by collecting data from identical samples in the qualitative and quantitative approaches. During data collection, participants will provide written reflections through SMS text messaging surveys (qualitative) on their glucose data from each day (quantitative). In addition, interview participants will provide narratives of their self-management behaviors that can be compared to CGM metrics.

During mixed methods analysis, we will first create participant summaries that include individual-level responses from interview transcripts, daily SMS text messaging surveys, baseline surveys, and CGM data. Raw and summary level baseline survey responses and summary level CGM metrics will be imported into the qualitative analysis software alongside the coded qualitative data. In addition, ambulatory glucose profiles from CGM data will be imported to provide context of recent and current glucose levels coinciding with real-time survey responses. We will then merge all data sources to create written participant summaries. Example participant profile provided in Multimedia Appendix 5.

We will also merge results to compare qualitative findings to diabetes outcomes (eg, self-management behaviors and diabetes distress) and glycemic levels. Using mixed methods features of MAXQDA software, we will create crosstabs to allow for comparison of qualitative findings by quantitative measures. For example, qualitative findings from interviews and daily SMS text messaging surveys will be compared between those with HbA_{1c} of more than 9% and those with HbA_{1c} of 9% or less. Integration through merging will allow us to identify patterns that can help to explain variation in glycemic levels or other diabetes-related physical and mental health outcomes. A similar approach will be used to integrate feasibility and acceptability data collected through the closeout survey and follow-up interview. Meta-inferences will be created to explain study feasibility and acceptability, including the perceived burden of participating in real-time and longitudinal data collection.

We will also create visual joint displays [69,70] that compare qualitative and quantitative results to generate meta-inferences. For example, one such joint display will examine patterns in qualitative findings by different glycemic outcomes to determine differences in the daily experiences of youth that help to explain glycemic levels and variability. A similar joint display can array qualitative findings by baseline measures of self-reported diabetes distress and self-management behaviors. Integrated findings will be used to further refine our theory of self-management with diabetes devices in this population.

When merging qualitative and quantitative results, we will determine whether the findings are concordant across approaches. Concordant results will be discussed among the research team, returning to the emergent grounded theory and comparing data and analytical categories (in qualitative findings) to quantitative results to examine meta-inferences. Both concordant and discordant results can support the refinement of the theory and generate additional hypotheses for future research [71-73].

Results

This study was reviewed and approved by the institutional review board in January 2025 and enrolled its first participant in April 2025. As of December 2025, 25 participants have enrolled in this study, and we anticipate meeting our enrollment goal of approximately 30 to 40 participants in early 2026. We anticipate findings to become available in the following several years through conference presentations and peer-reviewed publications. Furthermore, findings will inform the development of an intervention to address diabetes self-management and decision-making in adolescents with T1D.

Few previous applications of real-time data collection have leveraged conversational SMS text messaging (where questions are sent one-at-a-time following participant responses), and, as such, there is little data on participant burden and compliance available for this method; however, early studies using closed-ended surveys indicate high satisfaction and compliance rates (≤95%) among adolescents and young adults [74], especially when fewer prompts were used [75,76]. As a result, our study incorporated several recommendations for burden reduction and compliance optimization such as working with a population that feels connected to our study objective, using shorter, less frequent prompts, allowing a window of 2+ hours for responses, allowing the use of participants' own devices, providing a monetary incentive after completing the study, and co-developing surveys with youth with lived experience of T1D from the T1DYAC [77,78]. As a result, we are hopeful for response rates of at least 70% in this study. Likewise, we expect themes emerging from the analysis of feasibility and acceptability responses to highlight overall participant satisfaction and low burden.

We anticipate that sensor use time will be above 75%, as this population will already use CGM prior to study enrollment. In addition, we anticipate that participants with HbA_{1c} of more

XSL•FO
RenderX

than 9% will report lower problem-solving, self-management behaviors, and diabetes device use, and higher diabetes distress and depression compared with those with HbA_{1c} of 9% or less.

Primary qualitative outcomes include thematic description of typical and momentary behaviors and responses to diabetes distress, problem-solving, and decision-making. Historically, reported adherence levels to real-time data collection (eg, quantitative ecological momentary assessment) among children and adolescents have been high (78%), including among youth using mobile devices and wearables (73%) [79]. We anticipate that, even with up to 30% missing data on daily surveys, the combined repeated surveys and in-depth interviews will provide a sufficient sample size to detect differences between adolescents with HbA_{1c} of 9% or less and more than 9%. However, monetary incentives are provided based upon the completion of the study and prorated based on the level of completion to incentivize data completeness. We anticipate that theoretical saturation will likely occur prior to the enrollment of 40 participants. When integrated, the quantitative and qualitative data sources will provide novel insights into daily self-management behaviors for adolescents with T1D.

Discussion

Adolescents with T1D tend to struggle more with their diabetes self-management, have lower engagement with data provided by diabetes devices, and higher glycemic levels compared to all other age groups with this condition. Day-to-day self-management among adolescents may be further complicated by increasing independence from caregivers and ongoing physical, emotional, and cognitive development during this life stage. In particular, the ongoing development of executive functioning may inhibit optimal problem-solving abilities that are required for preventing and treating low and high blood glucose levels that arise throughout the day. Exploring routine, reactive, and reflective behaviors through real-time data collection can help us to better understand how adolescents make decisions and solve problems related to their diabetes during the day.

To our knowledge, this will be the first application of a real-time mixed methods approach to better understand daily decision-making and psychosocial influences surrounding diabetes self-management among adolescents with T1D. The robust data collection strategy has the potential to add significant value to the field as we work to better understand facilitators and barriers to optimal diabetes self-management in adolescents, especially among those who use CGM regularly. Furthermore, the qualitative insights provided will help elucidate intervention targets and preferences to improve self-management behaviors among this age group. Finally, this study will allow us to evaluate the feasibility of leveraging daily SMS text messaging for adolescents with T1D, which can open new potential avenues for both data collection and intervention.

While this study has several strengths, it does have limitations. First, this study protocol was limited by cohort geographics and direct recruitment from an academic medical center. However, based on the list generated from medical records and the flexible recruitment strategies, we intend to enroll a diverse sample of participants based on backgrounds, parental education level, and geographic location. While participants in this study did share structural challenges related to technology use (eg, access), future research should more intentionally explore the impact of these barriers on daily self-management experiences. Because this is one of the few applications of real-time qualitative data collection, there are still methodological questions to address, including potential participant burden, impact of data collection on participant behavior, and limitations to data quality. Findings related to feasibility and acceptability will be used to address these methodological questions and improve future studies using qualitative data collection through SMS text messaging and other real-time techniques. Recruitment and data collection for this study began in Spring 2025 and will likely continue through early 2026.

This study will be among the first to use a mixed methods, real-time approach and leverage SMS text messaging among adolescents with T1D. We anticipate that this approach will bring forth new insights into the daily self-management and problem-solving behaviors of adolescents with T1D. In addition, the combination of several data sources including glucose levels and semistructured interviews will allow for comparison of glycemic outcomes with the self-reported lived experiences of adolescents with T1D. This integrated approach will give a new perspective on self-management by adding contextual information to objective glucose data, supporting identification of key barriers to self-management in this population. Findings will inform future studies and interventions aiming to address diabetes self-management behaviors and problem-solving among adolescents with T1D.

Acknowledgments

The authors thank the members of the Type 1 Diabetes Youth Advisory Council for their feedback on study design including daily survey question and interview guide development. Generative artificial intelligence was not used in the preparation of this manuscript.

Funding

MD and this research are supported by a Career Development Award through the National Institute of Diabetes and Digestive and Kidney Diseases (grant ID K01DK134766). YKL is supported by the National Institute of Diabetes and Digestive Kidney Disease (grant ID K23DK129724). JML is supported by grants from the Michigan Nutrition Obesity Research Center (grant ID P30DK089503), the Michigan Diabetes Research Center (grant ID P30DK020572), and the Michigan Center for Diabetes Translational Research (grant ID P30DK092926) from the National Institute of Diabetes and Digestive and Kidney Diseases. In

XSL•FO
RenderX

addition, JML receives support from the Elizabeth Weiser Caswell Diabetes Institute at the University of Michigan. Research tools used for cohort identification and recruitment (DataDirect, EMERSE, and UMHealthResearch.org) and data collection and management (REDCap [Research Electronic Data Capture]) during this study are maintained and made freely available to researchers by a Clinical and Translational Science Award (grant ID UM1TR004404). The funders had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Authors' Contributions

Conceptualization: MD (lead), JL (supporting), SC (supporting)
Funding acquisition: MD
Methodology: MD (lead), JL (supporting), SC (supporting)
Supervision: JML (lead), TG (supporting), YKL (supporting), LRB (supporting), JEA (supporting), CR (supporting)
Visualization: MD (lead), SC (equal), JL (supporting)
Writing – original draft: MD (lead), SC (supporting), JL (supporting), HK (supporting), OD (supporting), FA (supporting), MJO (supporting), RM (supporting)
Writing – review & editing: MD (lead), SC (supporting), JL (supporting), JML (supporting), TG (supporting), YKL (supporting), LRB (supporting), JEA (supporting), CR (supporting)

Conflicts of Interest

MD is a section editor for *JMIR mHealth and uHealth*. LRB is the editor-in-chief of *JMIR mHealth and uHealth*. CR has received funding from Dexcom and Abbott for focus groups and advisory panels. JML has received grant funding from Lilly Inc. All other authors have no conflicts of interest to declare.

Multimedia Appendix 1
Demographics and clinical characteristics survey from the Qualitative Insights to Type 1 Diabetes in Youth (QUALITY) Study.
[DOCX File, 29 KB - resprot_v15i1e83218_app1.docx]

Multimedia Appendix 2
Validated instruments included in Qualitative Insights to Type 1 Diabetes in Youth (QUALITY) baseline survey.
[DOCX File, 61 KB - resprot_v15i1e83218_app2.docx]

Multimedia Appendix 3
Real-time SMS text messaging survey questions.
[DOCX File, 18 KB - resprot_v15i1e83218_app3.docx]

Multimedia Appendix 4
Qualitative Insights to Type 1 Diabetes in Youth (QUALITY) closeout survey.
[DOCX File, 20 KB - resprot_v15i1e83218_app4.docx]

Multimedia Appendix 5
Sample participant profile.
[DOCX File, 754 KB - resprot_v15i1e83218_app5.docx]

Checklist 1
COREQ checklist.
[PDF File, 208 KB - resprot_v15i1e83218_app6.pdf]

References

1. Fang M, Wang D, Selvin E. Prevalence of type 1 diabetes among US children and adults by age, sex, race, and ethnicity. JAMA 2024 Apr 23;331(16):1411-1413. [doi: 10.1001/jama.2024.2103] [Medline: 38573653]
2. Gong B, Yang W, Xing Y, Lai Y, Shan Z. Global, regional, and national burden of type 1 diabetes in adolescents and young adults. Pediatr Res 2025 Feb;97(2):568-576. [doi: 10.1038/s41390-024-03107-5] [Medline: 38443523]
3. American Diabetes Association Professional Practice Committee. 14. Children and adolescents: standards of care in diabetes-2025. Diabetes Care 2025 Jan 1;48(1 Suppl 1):S283-S305. [doi: 10.2337/dc25-S014] [Medline: 39651980]
4. Redondo MJ, Libman I, Maahs DM, et al. The evolution of hemoglobin A1c targets for youth with type 1 diabetes: rationale and supporting evidence. Diabetes Care 2021 Feb;44(2):301-312. [doi: 10.2337/dc20-1978] [Medline: 33431422]

XSL•FO
RenderX

5. Fang M, Xu Y, Ballew SH, et al. Trends and disparities in technology use and glycemic control in type 1 diabetes. JAMA Netw Open 2025 Aug 1;8(8):e2526353. [doi: 10.1001/jamanetworkopen.2025.26353] [Medline: 40788645]
6. Miller KM, Beck RW, Foster NC, Maahs DM, for the T1D Exchange. HbA1c levels in type 1 diabetes from early childhood to older adults: a deeper dive into the influence of technology and socioeconomic status on HbA1c in the T1D exchange clinic registry findings. Diabetes Technol Ther 2020 Sep 1;22(9):645-650. [doi: 10.1089/dia.2019.0393]
7. Anderzén J, Samuelsson U, Gudbjörnsdottir S, Hanberger L, Åkesson K. Teenagers with poor metabolic control already have a higher risk of microvascular complications as young adults. J Diabetes Complicat 2016 Apr;30(3):533-536. [doi: 10.1016/j.jdiacomp.2015.12.004]
8. Azar S, Maroun Abou Jaoude N, Kędzia A, Niechciał E. Barriers to type 1 diabetes adherence in adolescents. J Clin Med 2024 Sep 24;13(19):5669. [doi: 10.3390/jcm13195669] [Medline: 39407728]
9. Hofer SE, Raile K, Fröhlich-Reiterer E, et al. Tracking of metabolic control from childhood to young adulthood in type 1 diabetes. J Pediatr 2014 Nov;165(5):956-961. [doi: 10.1016/j.jpeds.2014.07.001] [Medline: 25151197]
10. Hill-Briggs F, Adler NE, Berkowitz SA, et al. Social determinants of health and diabetes: a scientific review. Diabetes Care 2020 Nov 2;44(1):258-279. [doi: 10.2337/dci20-0053] [Medline: 33139407]
11. Young-Hyman D, de Groot M, Hill-Briggs F, Gonzalez JS, Hood K, Peyrot M. Psychosocial care for people with diabetes: a position statement of the American Diabetes Association. Diabetes Care 2016 Dec;39(12):2126-2140. [doi: 10.2337/dc16-2053] [Medline: 27879358]
12. Zuijdwijk CS, Cuerden M, Mahmud FH. Social determinants of health on glycemic control in pediatric type 1 diabetes. J Pediatr 2013 Apr;162(4):730-735. [doi: 10.1016/j.jpeds.2012.12.010] [Medline: 23360562]
13. Annan SF, Higgins LA, Jelleryd E, et al. ISPAD Clinical Practice Consensus Guidelines 2022: Nutritional management in children and adolescents with diabetes. Pediatr Diabetes 2022 Dec;23(8):1297-1321. [doi: 10.1111/pedi.13429] [Medline: 36468223]
14. Ding Y, Zhang W, Wu X, et al. Deterioration in glycemic control on schooldays among children and adolescents with type 1 diabetes: a continuous glucose monitoring-based study. Front Pediatr 2022;10:1037261. [doi: 10.3389/fped.2022.1037261] [Medline: 36568429]
15. Baucom KJW, Queen TL, Wiebe DJ, et al. Depressive symptoms, daily stress, and adherence in late adolescents with type 1 diabetes. Health Psychol 2015 May;34(5):522-530. [doi: 10.1037/hea0000219] [Medline: 25798545]
16. Chiang JL, Maahs DM, Garvey KC, et al. Type 1 diabetes in children and adolescents: a position statement by the American Diabetes Association. Diabetes Care 2018 Sep;41(9):2026-2044. [doi: 10.2337/dci18-0023] [Medline: 30093549]
17. Hagger V, Hendrieckx C, Sturt J, Skinner TC, Speight J. Diabetes distress among adolescents with type 1 diabetes: a systematic review. Curr Diab Rep 2016 Jan;16(1):9. [doi: 10.1007/s11892-015-0694-2] [Medline: 26748793]
18. Luo J, Wang H, Li X, et al. Factors associated with diabetes distress among adolescents with type 1 diabetes. J Clin Nurs 2021 Jul;30(13-14):1893-1903. [doi: 10.1111/jocn.15742] [Medline: 33829586]
19. Hagger V, Hendrieckx C, Cameron F, Pouwer F, Skinner TC, Speight J. Diabetes distress is more strongly associated with HbA1c than depressive symptoms in adolescents with type 1 diabetes: results from diabetes MILES youth-Australia. Pediatr Diabetes 2018 Jun;19(4):840-847. [doi: 10.1111/pedi.12641] [Medline: 29383803]
20. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. JAMA 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
21. Ziegler R, von Sengbusch S, Kröger J, et al. Therapy adjustments based on trend arrows using continuous glucose monitoring systems. J Diabetes Sci Technol 2019 Jul;13(4):763-773. [doi: 10.1177/1932296818822539] [Medline: 30666883]
22. DeSalvo DJ, Miller KM, Hermann JM, et al. Continuous glucose monitoring and glycemic control among youth with type 1 diabetes: international comparison from the T1D exchange and DPV initiative. Pediatr Diabetes 2018 Nov;19(7):1271-1275. [doi: 10.1111/pedi.12711] [Medline: 29923262]
23. Huhn F, Lange K, Jördening M, Ernst G. Real-world use of continuous glucose monitoring systems among adolescents and young adults with type 1 diabetes: reduced burden, but little interest in data analysis. J Diabetes Sci Technol 2023 Jul;17(4):943-950. [doi: 10.1177/19322968221081216] [Medline: 35255729]
24. Brew-Sam N, Chhabra M, Parkinson A, et al. Experiences of young people and their caregivers of using technology to manage type 1 diabetes mellitus: systematic literature review and narrative synthesis. JMIR Diabetes 2021 Feb 2;6(1):e20973. [doi: 10.2196/20973] [Medline: 33528374]
25. Tanenbaum ML, Commissariat PV. Experience with burdens of diabetes device use that affect uptake and optimal use in people with type 1 diabetes. Endocr Connect 2023 Oct 1;12(10):e230193. [doi: 10.1530/EC-23-0193] [Medline: 37522857]
26. Dowd R, Jepson LH, Green CR, Norman GJ, Thomas R, Leone K. Glycemic outcomes and feature set engagement among real-time continuous glucose monitoring users with type 1 or non-insulin-treated type 2 diabetes: retrospective analysis of real-world data. JMIR Diabetes 2023 Jan 18;8:e43991. [doi: 10.2196/43991] [Medline: 36602920]
27. Laffel LM, Kanapka LG, Beck RW, et al. Effect of continuous glucose monitoring on glycemic control in adolescents and young adults with type 1 diabetes: a randomized clinical trial. JAMA 2020 Jun 16;323(23):2388-2396. [doi: 10.1001/jama.2020.6940] [Medline: 32543683]

XSL•FO
RenderX

28. Sheikh K, Bartz SK, Lyons SK, DeSalvo DJ. Diabetes device use and glycemic control among youth with type 1 diabetes: a single-center, cross-sectional study. J Diabetes Res 2018;2018:5162162. [doi: 10.1155/2018/5162162] [Medline: 30151393]
29. Divilly P, Zaremba N, Mahmoudi Z, et al. Hypo-METRICS: Hypoglycaemia-MEasurement, ThResholds and ImpaCtS—A multi-country clinical study to define the optimal threshold and duration of sensor-detected hypoglycaemia that impact the experience of hypoglycaemia, quality of life and health economic outcomes: the study protocol. Diabet Med 2022 Sep;39(9):e14892. [doi: 10.1111/dme.14892] [Medline: 35633291]
30. Mulvaney SA, Vaala SE, Carroll RB, et al. A mobile app identifies momentary psychosocial and contextual factors related to mealtime self-management in adolescents with type 1 diabetes. J Am Med Inform Assoc 2019 Dec 1;26(12):1627-1631. [doi: 10.1093/jamia/ocz147] [Medline: 31529065]
31. Nam S, Griggs S, Ash GI, et al. Ecological momentary assessment for health behaviors and contextual factors in persons with diabetes: a systematic review. Diabetes Res Clin Pract 2021 Apr;174:108745. [doi: 10.1016/j.diabres.2021.108745] [Medline: 33713720]
32. Zhang P, Fonnesbeck C, Schmidt DC, White J, Kleinberg S, Mulvaney SA. Using momentary assessment and machine learning to identify barriers to self-management in type 1 diabetes: observational study. JMIR Mhealth Uhealth 2022 Mar 3;10(3):e21959. [doi: 10.2196/21959] [Medline: 35238791]
33. Ehrmann D, Schmitt A, Priesterroth L, Kulzer B, Haak T, Hermanns N. Time with diabetes distress and glycemia-specific distress: new patient-reported outcome measures for the psychosocial burden of diabetes using ecological momentary assessment in an observational study. Diabetes Care 2022 Jul 7;45(7):1522-1531. [doi: 10.2337/dc21-2339] [Medline: 35613338]
34. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6(1):42. [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
35. Butler AM, Hilliard ME, Titus C, et al. Barriers and facilitators to involvement in children's diabetes management among minority parents. J Pediatr Psychol 2020 Sep 1;45(8):946-956. [doi: 10.1093/jpepsy/jsz103] [Medline: 31995219]
36. Cockcroft EJ, Wooding EL, Narendran P, et al. Factors affecting the support for physical activity in children and adolescents with type 1 diabetes mellitus: a national survey of health care professionals' perceptions. BMC Pediatr 2023 Mar 22;23(1):131. [doi: 10.1186/s12887-023-03940-3] [Medline: 36949473]
37. McSharry J, Byrne M, Casey B, et al. Behaviour change in diabetes: behavioural science advancements to support the use of theory. Diabet Med 2020 Mar;37(3):455-463. [doi: 10.1111/dme.14198] [Medline: 31797455]
38. Hamilton K, Stanton-Fay SH, Chadwick PM, et al. Sustained type 1 diabetes self-management: specifying the behaviours involved and their influences. Diabet Med 2021 May;38(5):e14430. [doi: 10.1111/dme.14430] [Medline: 33073393]
39. Smith R, Eisenberg S, Turner-Pfifer A, et al. We are on the verge of breakthrough cures for type 1 diabetes, but who are the 2 million Americans who have it? J Health Econ Outcomes Res 2024;11(2):145-153. [doi: 10.36469/001c.124604] [Medline: 39553538]
40. Hanauer DA, Barnholtz-Sloan JS, Beno MF, et al. Electronic Medical Record Search Engine (EMERSE): an information retrieval tool for supporting cancer research. JCO Clin Cancer Inform 2020 May;4(4):454-463. [doi: 10.1200/CCI.19.00134] [Medline: 32412846]
41. Hanauer DA, Mei Q, Law J, Khanna R, Zheng K. Supporting information retrieval from electronic health records: a report of University of Michigan's nine-year experience in developing and using the electronic medical record search engine (EMERSE). J Biomed Inform 2015 Jun;55:290-300. [doi: 10.1016/j.jbi.2015.05.003] [Medline: 25979153]
42. Charmaz K. Constructing Grounded Theory: Sage Publications; 2006. URL: https://uk.sagepub.com/en-gb/eur/constructing-grounded-theory/book255601 [accessed 2026-02-23]
43. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? Qual Health Res 2017 Mar;27(4):591-608. [doi: 10.1177/1049732316665344] [Medline: 27670770]
44. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
45. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
46. Onwuegbuzie AJ, Johnson RB, Collins KMT. Assessing legitimation in mixed research: a new framework. Qual Quant 2011 Oct;45(6):1253-1271. [doi: 10.1007/s11135-009-9289-9]
47. Perez A, Howell Smith MC, Babchuk WA, Lynch-O'Brien LI. Advancing quality standards in mixed methods research: extending the legitimation typology. J Mix Methods Res 2023 Jan;17(1):29-50. [doi: 10.1177/15586898221093872]
48. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
49. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
50. Iannotti RJ, Schneider S, Nansel TR, et al. Self-efficacy, outcome expectations, and diabetes self-management in adolescents with type 1 diabetes. J Dev Behav Pediatr 2006 Apr;27(2):98-105. [doi: 10.1097/00004703-200604000-00003] [Medline: 16682872]

XSL•FO
RenderX

51. Rasbach LE, Volkening LK, Markowitz JT, Butler DA, Katz ML, Laffel LMB. Youth and parent measures of self-efficacy for continuous glucose monitoring: survey psychometric properties. Diabetes Technol Ther 2015 May;17(5):327-334. [doi: 10.1089/dia.2014.0366] [Medline: 25695341]
52. Tanenbaum ML, Hanes SJ, Miller KM, Naranjo D, Bensen R, Hood KK. Diabetes device use in adults with type 1 diabetes: barriers to uptake and potential intervention targets. Diabetes Care 2017 Feb;40(2):181-187. [doi: 10.2337/dc16-1536] [Medline: 27899489]
53. Messer LH, Cook PF, Tanenbaum ML, Hanes S, Driscoll KA, Hood KK. CGM benefits and burdens: two brief measures of continuous glucose monitoring. J Diabetes Sci Technol 2019 Nov;13(6):1135-1141. [doi: 10.1177/1932296819832909] [Medline: 30854886]
54. Weissberg-Benchell J, Antisdel-Lomaglio J. Diabetes-specific emotional distress among adolescents: feasibility, reliability, and validity of the problem areas in diabetes-teen version. Pediatr Diabetes 2011 Jun;12(4 Pt 1):341-344. [doi: 10.1111/j.1399-5448.2010.00720.x] [Medline: 21443583]
55. Richardson LP, Rockhill C, Russo JE, et al. Evaluation of the PHQ-2 as a brief screen for detecting major depression among adolescents. Pediatrics 2010 May;125(5):e1097-e1103. [doi: 10.1542/peds.2009-2712] [Medline: 20368315]
56. Mehta SN, Nansel TR, Volkening LK, Butler DA, Haynie DL, Laffel LMB. Validation of a contemporary adherence measure for children with type 1 diabetes: the diabetes management questionnaire. Diabet Med 2015 Sep;32(9):1232-1238. [doi: 10.1111/dme.12682] [Medline: 26280463]
57. Mulvaney SA, Jaser SS, Rothman RL, et al. Development and validation of the diabetes adolescent problem solving questionnaire. Patient Educ Couns 2014 Oct;97(1):96-100. [doi: 10.1016/j.pec.2014.07.005] [Medline: 25063715]
58. Sidoti O, Park E, Faverio M, Atske S, Radde K. Teens and internet, device access fact sheet. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/teens-and-internet-device-access-fact-sheet/ [accessed 2025-12-24]
59. Sun X, Haydel KF, Matheson D, Desai M, Robinson TN. Are mobile phone ownership and age of acquisition associated with child adjustment? A 5-year prospective study among low-income Latinx children. Child Dev 2023 Jan;94(1):303-314. [doi: 10.1111/cdev.13851] [Medline: 36056573]
60. Proudfoot K. Inductive/deductive hybrid thematic analysis in mixed methods research. J Mix Methods Res 2023 Jul;17(3):308-326. [doi: 10.1177/15586898221126816]
61. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): A 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
62. Morse JM. The case of the missing data. Qual Health Res 2011 Sep;21(9):1163-1164. [doi: 10.1177/1049732311410359] [Medline: 21821505]
63. Liao Y, Skelton K, Dunton G, Bruening M. A systematic review of methods and procedures used in ecological momentary assessments of diet and physical activity research in youth: an adapted STROBE Checklist for Reporting EMA Studies (CREMAS). J Med Internet Res 2016 Jun 21;18(6):e151. [doi: 10.2196/jmir.4954] [Medline: 27328833]
64. Battelino T, Alexander CM, Amiel SA, et al. Continuous glucose monitoring and metrics for clinical trials: an international consensus statement. Lancet Diabetes Endocrinol 2023 Jan;11(1):42-57. [doi: 10.1016/S2213-8587(22)00319-9] [Medline: 36493795]
65. Broll S, Urbanek J, Buchanan D, et al. Interpreting blood GLUcose data with R package iglu. PLoS ONE 2021;16(4):e0248560. [doi: 10.1371/journal.pone.0248560] [Medline: 33793578]
66. Chun E, Fernandes NJ, Gaynanova I. An update on the iglu software package for interpreting continuous glucose monitoring data. Diabetes Technol Ther 2024 Dec;26(12):939-950. [doi: 10.1089/dia.2024.0154] [Medline: 38885321]
67. Olsen MT, Klarskov CK, Dungu AM, Hansen KB, Pedersen-Bjergaard U, Kristensen PL. Statistical packages and algorithms for the analysis of continuous glucose monitoring data: a systematic review. J Diabetes Sci Technol 2025 May;19(3):787-809. [doi: 10.1177/19322968231221803] [Medline: 38179940]
68. Fetters MD, Molina-Azorin JF. The journal of mixed Methods research starts a new decade: the mixed methods research integration trilogy and its dimensions. J Mix Methods Res 2017 Jul;11(3):291-307. [doi: 10.1177/1558689817714066]
69. Guetterman TC, Manojlovich M. Grand rounds in methodology: designing for integration in mixed methods research. BMJ Qual Saf 2024 Jun 19;33(7):470-478. [doi: 10.1136/bmjqs-2023-016112] [Medline: 38575310]
70. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015 Nov;13(6):554-561. [doi: 10.1370/afm.1865] [Medline: 26553895]
71. Bazeley P. Integrating Analyses in Mixed Methods Research: SAGE Publications; 2018. [doi: 10.4135/9781526417190]
72. Fetters MD. The Mixed Methods Research Workbook: Activities for Designing, Implementing, and Publishing Projects: SAGE Publications; 2020. [doi: 10.4135/9781071909713]
73. Younas A, Pedersen M, Inayat S. Practical strategies to identify and address discordant findings in mixed methods research. Res Methods Med Health Sci 2023 Jan;4(1):24-33. [doi: 10.1177/26320843221122342]
74. Psihogios AM, Li Y, Ahmed A, et al. Daily text message assessments of 6‐mercaptopurine adherence and its proximal contexts in adolescents and young adults with leukemia: a pilot study. Pediatr Blood Cancer 2021 Feb;68(2):e28767. [doi: 10.1002/pbc.28767] [Medline: 33073479]

XSL•FO
RenderX

75. Drexl K, Ralisa V, Rosselet-Amoussou J, et al. Readdressing the ongoing challenge of missing data in youth ecological momentary assessment studies: meta-analysis update. J Med Internet Res 2025 Apr 30;27:e65710. [doi: 10.2196/65710] [Medline: 40305088]
76. Wang S, Yang CH, Brown D, Cheng A, Kwan MYW. Participant compliance with ecological momentary assessment in movement behavior research among adolescents and emerging adults: systematic review. JMIR Mhealth Uhealth 2025 Feb 11;13:e52887. [doi: 10.2196/52887] [Medline: 39933165]
77. Murray AL, Xie T, Power L, Condon L. Recruitment and retention of adolescents for an ecological momentary assessment measurement burst mental health study: the MHIM engagement strategy. Health Expect 2024 Jun;27(3):e14065. [doi: 10.1111/hex.14065] [Medline: 38711174]
78. Smyth JM, Jones DR, Wen CKF, Materia FT, Schneider S, Stone A. Influence of ecological momentary assessment study design features on reported willingness to participate and perceptions of potential research studies: an experimental study. BMJ Open 2021 Jul 30;11(7):e049154. [doi: 10.1136/bmjopen-2021-049154] [Medline: 34330860]
79. Wen CKF, Schneider S, Stone AA, Spruijt-Metz D. Compliance with mobile ecological momentary assessment protocols in children and adolescents: a systematic review and meta-analysis. J Med Internet Res 2017 Apr 26;19(4):e132. [doi: 10.2196/jmir.6641] [Medline: 28446418]

Abbreviations

CGM: continuous glucose monitoring
COREQ: Consolidated Criteria for Reporting Qualitative Research
CREMAS: Checklist for Reporting EMA Studies
EMERSE: Electronic Medical Record Search Engine
HbA_{1c}: hemoglobin A_{1c}
HIPAA: Health Insurance Portability and Accountability Act
MAXQDA: Max Weber Qualitative Data Analysis
QUALITY: Qualitative Insights to Type 1 Diabetes in Youth
REDCap: Research Electronic Data Capture
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
T1D: type 1 diabetes
T1DYAC: Type 1 Diabetes Youth Advisory Council
U-M: University of Michigan

Edited by J Sarvestan; submitted 29.Aug.2025; peer-reviewed by I Milani, M Hilliard; accepted 16.Jan.2026; published 04.Mar.2026.

Please cite as:
DeJonckheere M, Chuisano SA, Lucien J, Amjad F, Duvvuri O, Khan H, Khan M, Mirza R, Ollivierre MJ, Guetterman T, Lin YK, Buis LR, Aikens JE, Richardson C, Lee JM
Real-Time Type 1 Diabetes Self-Management Decision-Making in Adolescents: Protocol for a Longitudinal Mixed Methods Study Using Text Messaging and Continuous Glucose Monitoring
JMIR Res Protoc 2026;15:e83218
URL: https://www.researchprotocols.org/2026/1/e83218
doi:10.2196/83218

© Melissa DeJonckheere, Samantha A Chuisano, Juniar Lucien, Fouzaan Amjad, Oorvi Duvvuri, Hasan Khan, Maryam Khan, Rafee Mirza, Maya Joy Ollivierre, Timothy Guetterman, Yu Kuei Lin, Lorraine R Buis, James E Aikens, Caroline Richardson, Joyce M Lee. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 4.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Digital Dissemination Strategies for Health Promotion Videos in Indigenous Communities in California: Protocol for a Three-Arm Comparative Study

Lucía Abascal Miguel[1], MSc, MD, PhD; Anna E Epperson[2], PhD; Alison B Comfort[3], PhD; Darío León[4], BA; Mary E Garcia[2], BA; Alicia R Riley[4], MPH, PhD; Nadia Diamond-Smith[1,5], MSc, PhD

[1]Institute for Global Health Sciences, University of California, San Francisco, 550 16th St, San Francisco, CA, United States

[2]Psychological Sciences and Health Sciences Research Institute, University of California, Merced, Merced, CA, United States

[3]Department of Obstetrics, Gynecology and Reproductive Sciences, University of California, San Francisco, San Francisco, CA, United States

[4]Department of Sociology, University of California, Santa Cruz, Santa Cruz, CA, United States

[5]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Lucía Abascal Miguel, MSc, MD, PhD
Institute for Global Health Sciences, University of California, San Francisco, 550 16th St, San Francisco, CA, United States

Abstract

Background: Despite the availability of effective vaccines, flu and COVID-19 uptake remains suboptimal, including among Indigenous communities in California who face unique barriers to accessing public health information. While previous research has evaluated health communication message content and design, fewer studies have systematically compared different dissemination strategies for the same intervention, leaving gaps in understanding optimal approaches for reaching marginalized populations.

Objective: This protocol describes a three-arm dissemination study designed to compare the effectiveness of different strategies for distributing COVID-19 and flu vaccine promotion videos among Indigenous Peoples residing in California, including American Indian and Alaska Natives, and migrant Indigenous Peoples from Latin America.

Methods: Following extensive formative work, including cross-sectional surveys, social network analyses, a discrete choice experiment, and qualitative focus groups conducted with guidance from an Indigenous Community Advisory Board, we developed four 30-second vaccine promotion videos available in English and Spanish, two targeting American Indian and Alaska Native communities and 2 targeting Latin American communities. Videos will be translated to four Indigenous languages (Purépecha, Mam, Zapoteco, and Mixteco). The study compared three dissemination strategies implemented over a two-month period: (1) targeted social media advertisements on Facebook and Instagram, (2) distribution through community-based organizations using their established communication channels, and (3) peer-to-peer sharing through Indigenous community members (“seeds”) recruited from previous research. Primary outcomes focus on dissemination effectiveness, defined by measures of reach (eg, link shares, clicks and video views), engagement (eg, watch time and survey completion), and message impact (eg, trust in the message and vaccination intent), assessed using Bitly link tracking, YouTube analytics, and follow-up surveys. Descriptive and comparative analyses will be conducted to assess differences across dissemination strategies.

Results: The study was funded in January 2023. Intervention development was completed in July 2025. Dissemination activities and data collection across all three study arms were conducted between October 6, 2025, to November 28, 2025, aligned with the COVID-19 and influenza vaccination season in the United States. Data consolidation, cleaning, and analyses began January 2026 and primary results are expected to be submitted for publication in Spring 2026.

Conclusions: This study addresses a critical gap in health communication research by providing a systematic methodology for comparing digital dissemination strategies within Indigenous communities. The combination of community-informed recruitment, multilingual accessibility, and comprehensive digital tracking tools offers a replicable model for evaluating how public health messages spread through different channels, particularly for populations historically excluded from traditional outreach efforts.

Trial Registration: ClinicalTrials.gov (NCT07096245); https://clinicaltrials.gov/ct2/show/NCT07096245

International Registered Report Identifier (IRRID): PRR1-10.2196/84674

(JMIR Res Protoc 2026;15:e84674) doi:10.2196/84674

KEYWORDS

health communication; vaccine interventions; social media recruitment; Indigenous populations; native American populations

XSL•FO
RenderX

Introduction

Flu and COVID-19 continue to be leading causes of hospitalization and death, particularly for people 65 years of age and older [1]. Despite the availability of vaccines, uptake has stalled in many communities, including in California [2]. As public health communications become increasingly digital, there is a growing need to understand not only what messages are effective, but how those messages are disseminated and received across diverse population.

A growing body of research has examined the design and content of digital health messages, including message framing, visual style, messenger characteristics, and tone, often using social media analytics or survey-based metrics [3-6]. Within this literature, dissemination has increasingly been identified as a critical determinant of communication effectiveness, influencing who is reached, how messages are engaged with, and whether information is trusted and acted upon [7,8]. However, most existing work evaluates dissemination strategies in isolation or confuses other message attributes with delivery strategy, limiting the ability to study their independent contribution to message effectiveness [9,10]. This gap is particularly salient for Indigenous populations, for whom trust, relational networks, and community-based information channels often play a central role in how health information circulates [11,12]. While community-engaged dissemination and peer-led approaches have been highlighted as promising strategies for Indigenous and other marginalized communities, few studies have directly compared multiple digital dissemination strategies using identical content to assess their relative effectiveness [7,13].

This study addresses an important methodological and knowledge gap by comparing three dissemination strategies using identical, community-informed vaccine promotion content across Indigenous populations in California. We use the term Indigenous to refer to American Indian, Alaska Native, Native American, and migrant Indigenous Peoples from Latin America. US census data estimate that approximately 650,000 - 760,000 individuals in California identify as American Indian or Alaska Native [14]. Population estimates for Indigenous Peoples from Latin America are limited due to census classification practices; community sources suggest roughly 170,000 may reside in California, though this is likely an underestimate [15].

In this paper, we describe the methods for measuring and testing different dissemination approaches for COVID-19 and flu vaccine promotion videos while holding message content constant. The approaches were selected based on prior dissemination and implementation literature and extensive formative work conducted with Indigenous communities (described below), these are: (1) targeted social media advertisements, (2) through community-based organizations (CBOs), and (3) peer-to-peer sharing through community members ("seeds").

This dissemination study is the final part of a collaborative 4-year project (2022 - 2026) that aims to increase flu and COVID-19 vaccine confidence on Indigenous communities in California, including from Tribal communities with historic lands in the United States and Indigenous migrant communities from Latin America. Three University of California campuses have participated including UC San Francisco, UC Merced and UC Santa Cruz. This research has been guided by an Indigenous Community Advisory Board (ICAB) comprised of Indigenous leaders in their communities and draws from principles of implementation science, human centered design and social network theory. Our formative work included:

1. A cross-sectional survey of health information, vaccine hesitancy, and social media platform use among our target population recruited through social media and our community-based partners completed between 2023 and 2024 [16]
2. Social network analyses using an egocentric network approach to collect data from individuals (ie, *egos*) about the individuals they rely on for health advice and information (ie, *alters*) to identify who within their health network matters in influencing future vaccine intentions (eg, family, friends, healthcare worker), what modes of communication between them are most important (eg, social media, in person) and other network characteristics like trust in these individuals. Our analyses also identified the importance of network members' experiences with COVID-19, including being supportive of vaccination, encouraging vaccination, having been sick with COVID-19 and having been vaccinated, completed in 2025. (Comfort A, PhD, unpublished data, December 2025)
3. A discrete choice experiment to evaluate preferences for visual and narrative characteristics of posts on vaccines, including other communication attributes (eg, cartoon vs real-life images, tone, messenger) [17]
4. Qualitative focus groups (n=4) with Indigenous adults in California, recruited through our community partners. They were conducted in English, Spanish, and Purépecha, to explore experiences with COVID-19, vaccination, health information access, and social media use, and to contextualize findings from quantitative components, completed between 2024 and 2025. (Riley A, Epperson A, unpublished data, February 2026)

In this study we aim to compare the effectiveness and reach of different digital delivery strategies for disseminating identical, community-informed COVID-19 and influenza vaccine promotion videos among Indigenous populations in California.

Methods

Overview

This study used a three-arm comparative dissemination design to evaluate different strategies for distributing identical COVID-19 and influenza vaccine promotion videos among Indigenous communities in California (Figure 1). The study was conducted during the 2025 COVID-19 and influenza vaccination season in the United States. A table summarizing the study timeline and schedule of dissemination and assessments is provided in Table 1.

XSL•FO
RenderX

Figure 1. Overview of study design and dissemination strategies.

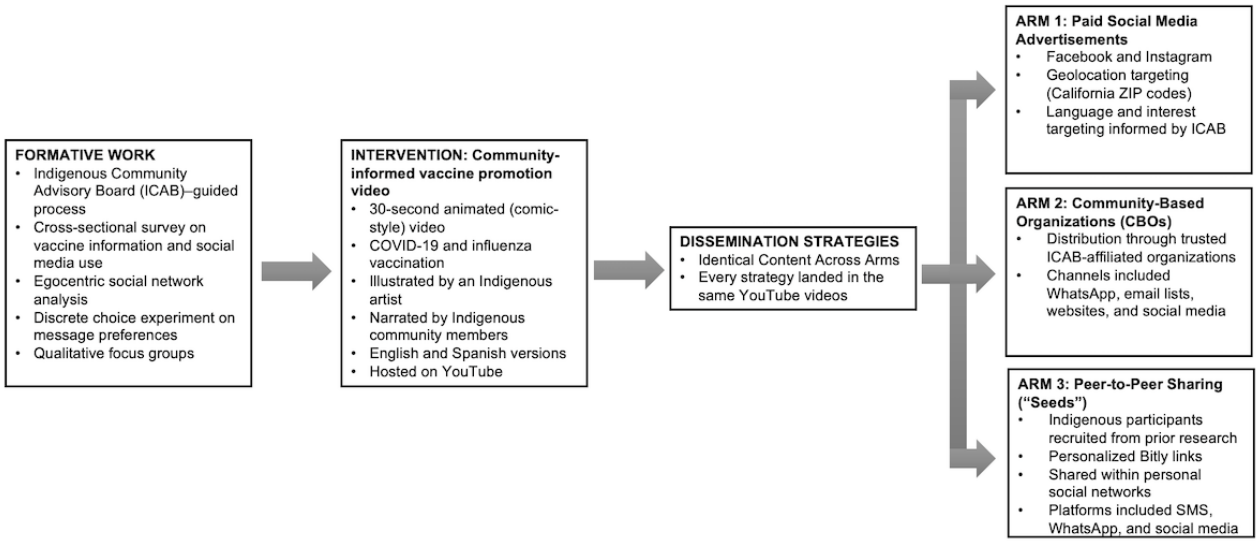

Table . Study timeline and schedule of activities and assessments.

Study phase	Timeframe	Key activities	Data collected
Formative research	2023 - 2025	Cross-sectional survey, egocentric social network analysis, discrete choice experiment, qualitative focus groups, ICAB[a] consultation	Vaccine attitudes, social media use, network characteristics, message preferences
Intervention development	Jan–Jul 2025	Co-creation of vaccine promotion videos with ICAB, illustration and narration by Indigenous artists and community members, translation into Spanish and Indigenous languages	Not applicable
Dissemination period	Oct 6–Nov 28, 2025	Distribution of identical videos through three dissemination arms (paid social media ads, CBO[b] distribution, peer-to-peer "seeds")	Bitly link metrics; YouTube reach and engagement metrics
Post-viewing survey	Oct–Nov 2025	Optional survey linked at end of video	Trust in message, vaccination intent, reactions to video, self-reported sharing
Data consolidation and analysis	Jan 2026–present	Cleaning and integration of Bitly, YouTube, and survey data	Primary and secondary outcome measures
Planned dissemination of results	Spring 2026	Manuscript preparation and submission	Not applicable

[a]ICAB: Indigenous Community Advisory Board.

[b]CBO: community-based organizations.

Intervention Development and Design

Guided by our formative findings and in close collaboration with our ICAB, we developed two 30-second vaccine promotion videos. The videos were co-developed through participatory workshops with ICAB members and were produced in English, Spanish, and four Indigenous languages (Purépecha, Mam, Zapoteco, and Mixteco). For this dissemination study, we tested the English and Spanish versions of one video that features family members narrating their decision to receive COVID-19 and influenza vaccines, motivated by prior illness experiences and concerns about protecting their family. The narrative concludes with family members choosing to get vaccinated together.

The videos use a comic-style visual format illustrated by an Indigenous artist and are narrated by Indigenous community members. The two versions are identical in narrative structure, pacing, and messaging, with minor visual adaptations (eg, clothing and color schemes) to reflect either American Indian and Alaska Native or Indigenous Latin American communities, only the English and Spanish versions were used for testing.

XSL•FO
RenderX

The final intervention was implemented between October 6, 2025, to November 28, 2025, timed for the COVID/flu vaccine season in the United States.

Study Design and Dissemination Arms

In our trial, we will test 3 dissemination strategies for our intervention. The videos will be hosted on our university's institutional YouTube channel, with separate versions in each focal language. Videos will be distributed through the following three dissemination approaches for one month:

1. Paid social media ads: videos will be shared through a paid ad campaign through Facebook and Instagram. They will be shared as posts with a link to the YouTube video. The audience will be targeted using geolocation through zip codes with high Indigenous populations in California (based on ICAB and census data), language including Spanish, and interests identified with our ICAB.
2. Community based organizations (CBOs): ICAB-affiliated CBOs will share the links to the videos using their usual communication channels used with their members (eg, WhatsApp, social media, email lists, CBO websites).
3. Individual seeds: We will recruit Indigenous-identifying participants from our previous survey (listed above) who consented to future contact. Each individual will receive a link and be asked to share the video with five or more contacts in their network using their preferred method (eg, SMS, WhatsApp, social media).

Data Collection

Outcome Measurement

Outcomes include dissemination performance and message impact across dissemination strategies. The primary outcomes were (1) dissemination reach, measured using platform-level analytics and link-level tracking, and (2) message trust and vaccination intent, assessed through a brief post-viewing survey. Secondary outcomes included engagement metrics (eg, views, watch time, and sharing behaviors) and self-reported changes in views or attitudes toward COVID-19 and influenza vaccination following exposure to the video.

Data Collection

Data was collected through three complementary methods to capture both dissemination performance and message impact across strategies:

1. Survey: a clickable Qualtrics survey link was shared at the end of the video to collect post-intervention data. The brief survey assessed likelihood of flu and COVID-19 vaccination, reactions to the video, how the video was shared, and trust and engagement with the video. Participants who completed the survey were entered into a gift card raffle.
2. Link dissemination tracking: Bitly was used to create individual and personalized links for each dissemination strategy (eg bit.ly/socialmedia). We made a single link for all social media ads, one link per organization for the participating CBOs; and one link per participant for the seeds. We used Bitly analytics to track link engagement, clicks, locations, referrals and type of devices used. All links will corresponded to the same YouTube videos.
3. YouTube analytics: we used YouTube metrics to track platform-level spread and engagement including video reach metrics (impressions, click through rate, views, etc), engagement metrics (watch time, retention, likes, comments, shares), audience insights (age and gender, viewers) and follow-ups (end screen and cards).

Analysis

We will conduct both descriptive and comparative analyses to evaluate the effectiveness of each dissemination strategy. Dissemination patterns within each arm will be summarized, including views, link interactions, and survey responses. Where possible, metrics will be standardized (eg, engagement per view or per dissemination unit) to compare across strategies that could have different scales of deliveries.

Differences between arms will be assessed using aggregated platform (YouTube)- and link-level (Bitly) metrics, as well as survey-based outcomes. For survey measures of trust in the video and vaccination intent, distributions will be compared across dissemination strategies using logistical regressions controlled for sociodemographic characteristics (age, gender, and race/ethnicity). Subgroup analyses (eg, by age group, including adults aged ≥65 y) will be conducted where feasible and will depend on the overall sample size and the distribution of participants within each subgroup.

We will examine indicators of dissemination efficiency, such as reach per dissemination unit (eg, per organization, per seed, or per advertising campaign) and engagement. These comparisons are intended to identify dissemination approaches with greater potential for scalability and sustainability in Indigenous communities. The study is not powered to detect predefined effect sizes. Instead, results will provide empirical benchmarks to guide dissemination strategy selection in larger implementation efforts.

Ethical Considerations

This trial was registered at ClinicalTrials.gov (NCT07096245) and approved by the University of California, San Francisco Institutional Review Board (IRB #23 - 38709). No personal or identifiable data were collected through YouTube, Bitly, Facebook, or Instagram analytics; these platforms were used solely to obtain aggregate dissemination and engagement metrics. All survey data were collected through a secure UCSF (University of California, San Francisco) Qualtrics account. Participants who completed the post-viewing survey provided electronic informed consent prior to participation and, if interested, were entered into a raffle for one of forty US $100 gift cards. Individuals recruited as peer-to-peer "seeds" were consented in advance and received $30 in compensation for distributing the intervention content within their networks.

Results

The study was funded in January 2023. Intervention development, including formative work and the co-creation and translation of vaccine promotion videos, was completed in July 2025. Dissemination activities and participant-facing data

XSL•FO
RenderX

collection across all three study arms were conducted between October 6 and November 28, 2025, aligned with the COVID-19 and influenza vaccination season in the United States (Table 1).

Three dissemination strategies were implemented during this period. For the community-based organization arm, five community-based organizations participated in distributing the video links to their members using established communication channels. For the peer-to-peer dissemination arm, 68 Indigenous-identifying participants from prior research were recruited as individual "seeds" and provided with personalized links to share the videos within their social networks. In parallel, targeted paid social media campaigns were conducted on Facebook and Instagram using geolocation, language, and interest-based targeting strategies informed by formative research and guidance from the ICAB.

All dissemination activities concluded on November 28, 2025. Data from Bitly link tracking, YouTube analytics, and follow-up surveys were consolidated and prepared for analysis beginning in January 2026. Outcome analyses are underway, and primary results are expected to be submitted for publication in Spring 2026.

Discussion

This study was designed to compare how identical, community-informed vaccine promotion content performs when disseminated through different digital channels among Indigenous communities in California. We hypothesize that dissemination strategies that leverage trusted relationships, such as community-based organizations and peer-to-peer sharing, will demonstrate higher levels of trust and message impact compared to standard social media advertising, while paid ads may achieve greater overall reach.

As public health communication continues to shift into digital spaces, and as people rely on a growing number of virtual platforms to connect, it becomes more important to understand how health messages are shared and how the messenger impacts how people trust those messages. This project responds to a gap in health research by testing three different strategies to distribute the same vaccine information video to Indigenous communities in California, communities that have often been left out of traditional outreach. We compare broad social media ads, which are commonly used today, to more targeted approaches through trusted community organizations and through individual participants sharing within their own networks.

One of the strengths of this study is the combination of digital tracking tools, like Bitly and YouTube, with a post-intervention survey. This allows us to see how messages move across systems and gives us a way to compare which strategy leads to more views, engagement, and trust. These insights can help shape future campaigns by showing how public health information can be distributed in ways that are more meaningful and relevant for the people they aim to reach. In addition, we ensure accessibility of the message by making it available in four Indigenous languages spoken in California.

We do expect some limitations. For example, the success of the seed-sharing strategy depends on whether participants decide to share the video, and whether the people they share it with choose to watch it and respond to the survey. These steps are voluntary, and it is anticipated that there may be some participant drop-off.

This study will provide a practical and grounded model for testing how digital public health communication interventions are shared, especially when the goal is to reach communities who are not usually centered in public health campaigns, who may have less trust in government resources due to historical experiences, and who may be harder to reach or have other barriers (language, etc).

Funding

This research was primarily supported by funds from the UC Multicampus Research Programs and Initiatives of the University of California, Grant Number M23PR6007. Formative research was supported by the California Collaborative for Pandemic Recovery and Readiness Research (CPR3) Program which was funded by the California Department of Public Health (CDPH). CDPH was not involved in study design, data collection, data analysis and interpretation, manuscript development, or the decision to publish.

Data Availability

Data generated through this study will not be publicly available at the time of publication due to the involvement of Indigenous communities and the sensitive nature of the data. Deidentified, aggregated data may be made available upon reasonable request to the corresponding author, subject to approval by the study team and the Indigenous Community Advisory Board, and in accordance with institutional policies and data governance agreements. Requests will be reviewed to ensure appropriate use and alignment with community priorities.

Authors' Contributions

Conceptualization: LAM, NDS
Data curation: LAM
Formal analysis: AC, AE, AR, LAM, NDS
Funding acquisition: AC, AE, AR, LAM, NDS

XSL•FO
RenderX

Investigation: JV, LAM
Methodology: LAM
Project administration: AE, NDS
Supervision: AE, NDS
Writing – original draft: LAM
Writing – review & editing: AC, AE, AR, JV, LAM, NDS

Conflicts of Interest

None declared.

References

1. Flu and people 65 years and older. Influenza (Flu)—CDC. 2025. URL: https://www.cdc.gov/flu/highrisk/65over.htm [accessed 2025-09-08]
2. Division of Communicable Disease Control. California Department of Public Health. URL: https://www.cdph.ca.gov/Programs/CID/DCDC [accessed 2025-09-08]
3. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520. [Medline: 25083128]
4. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
5. Cascini F, Pantovic A, Al-Ajlouni YA, et al. Social media and attitudes towards a COVID-19 vaccination: a systematic review of the literature. EClinicalMedicine 2022 Jun;48:101454. [doi: 10.1016/j.eclinm.2022.101454] [Medline: 35611343]
6. Patel R, Chang T, Greysen SR, Chopra V. Social media use in chronic disease: a systematic review and novel taxonomy. Am J Med 2015 Dec;128(12):1335-1350. [doi: 10.1016/j.amjmed.2015.06.015] [Medline: 26159633]
7. Kelly D, Ostovar-Kermani TG. Systematic literature review of disseminating health information to BIPOC communities. Sci Posters 2024;12. [doi: 10.14293/P2199-8442.1.SOP-.PUMOO1.v1]
8. Roberts-Lewis S, Baxter H, Mein G, et al. Examining the effectiveness of social media for the dissemination of research evidence for health and social care practitioners: systematic review and meta-analysis. J Med Internet Res 2024 Jun 5;26:e51418. [doi: 10.2196/51418] [Medline: 38838330]
9. LaRocca R, Yost J, Dobbins M, Ciliska D, Butt M. The effectiveness of knowledge translation strategies used in public health: a systematic review. BMC Public Health 2012 Sep 7;12:751. [doi: 10.1186/1471-2458-12-751] [Medline: 22958371]
10. Brownson RC, Eyler AA, Harris JK, Moore JB, Tabak RG. Getting the word out: new approaches for disseminating public health science. J Public Health Manag Pract 2018;24(2):102-111. [doi: 10.1097/PHH.0000000000000673] [Medline: 28885319]
11. Dirks LG, Wanda P. Technology to support collaborative dissemination of research with Alaska Native communities. AMIA Annu Symp Proc 2021;2021:398-407. [Medline: 35308978]
12. Weeks R, White S, Hartner AM, et al. COVID-19 messaging on social media for American Indian and Alaska Native communities: thematic analysis of audience reach and web behavior. JMIR Infodemiology 2022;2(2):e38441. [doi: 10.2196/38441] [Medline: 36471705]
13. Hirchak KA, Oluwoye O, Nadeau M, et al. Coming together for something good: recommendations from a scoping review for dissemination and implementation science to improve indigenous substance use disorder treatment. Front Public Health 2023;11:1265122. [doi: 10.3389/fpubh.2023.1265122] [Medline: 37915816]
14. QuickFacts: California. US Census Bureau. URL: https://www.census.gov/quickfacts/fact/table/CA/PST045224 [accessed 2026-01-30]
15. Who is California's Indigenous Migrant Community? Mixteco. URL: https://mixteco.org/mixtec/ [accessed 2026-01-30]
16. Diamond-Smith N, Comfort A, Epperson A, et al. Social media recruitment in Indigenous and Native American populations: challenges in the AI age. JMIR Public Health Surveill 2025 Sep 17;11(1):e76677. [doi: 10.2196/76677] [Medline: 40962118]
17. Abascal Miguel L, Comfort A, Riley A, et al. Preferences for social media vaccine messaging: a discrete choice experiment in California. JAMA Netw Open (forthcoming) 2026.

Abbreviations

CBO: Community-based organization
ICAB: Indigenous Community Advisory Board
IM: Indigenous migrant
IRB: Institutional Review Board
UCSF: University of California, San Francisco

XSL•FO
RenderX

Edited by A Schwartz; submitted 26.Sep.2025; peer-reviewed by K Gutu; revised version received 30.Jan.2026; accepted 24.Feb.2026; published 17.Mar.2026.

Please cite as:
Abascal Miguel L, Epperson AE, Comfort AB, León D, Garcia ME, Riley AR, Diamond-Smith N
Digital Dissemination Strategies for Health Promotion Videos in Indigenous Communities in California: Protocol for a Three-Arm Comparative Study
JMIR Res Protoc 2026;15:e84674
URL: https://www.researchprotocols.org/2026/1/e84674
doi:10.2196/84674

© Lucía Abascal Miguel, Anna E Epperson, Alison B Comfort, Darío León, Mary E Garcia, Alicia R Riley, Nadia Diamond-Smith. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Development and Protocol for Testing a Co-Created Digital Intervention (Sentinel) to Improve Mental Well-Being and Help Manage and Prevent Trauma in First Responders

Nicola Cogan, MA, DClinPsy, PhD; Alison Kirk, BSc, PhD; Christoph Graf, BA

Department of Psychological Sciences and Health, University of Strathclyde, 40 George Street, Glasgow, United Kingdom

Corresponding Author:
Nicola Cogan, MA, DClinPsy, PhD
Department of Psychological Sciences and Health, University of Strathclyde, 40 George Street, Glasgow, United Kingdom

Abstract

Background: First responders (FRs) and frontline workers are frequently exposed to traumatic events within their professional roles. This exposure places them at risk of experiencing acute stress, posttraumatic stress disorder, burnout, and other adverse mental health outcomes. Despite growing awareness of these risks, there remains a lack of evidence-based digital interventions (DIs) tailored to meet their unique mental health needs.

Objective: This study aims to address this gap by developing and testing Sentinel, an evidence-based, co-created DI to promote mental well-being, build resilience, and help manage and prevent trauma among FRs and frontline workers. The objectives include exploring their experiences of occupational trauma, identifying their preferences for digital mental health tools, and evaluating the feasibility of the Sentinel intervention.

Methods: The development of Sentinel followed a rigorous, 4-phase approach. In phase 1, we conducted market analysis and in-depth qualitative interviews with 54 FRs from fire services, police, and emergency health sectors. The aim was to explore their mental health needs, barriers to accessing support, and views on the potential role of DIs in addressing these needs. In phase 2, we developed the content for Sentinel by integrating findings from phase 1 with existing evidence, policies, and theoretical frameworks, ensuring the intervention reflected the lived experiences of FRs. In phase 3, a high-fidelity clickable prototype of Sentinel will be tested through co-design workshops, iterative development sprints, and usability evaluations. Feedback from FRs and frontline workers during this phase will refine the app's design and functionality. Phase 4 involves a mixed methods, nonrandomized feasibility study to evaluate Sentinel's acceptability, usability, safety, and implementation potential. Quantitative data will be collected from up to 100 FRs, complemented by qualitative interviews with 30 FR participants and 20 health and social care professionals who refer FRs to the intervention.

Results: The co-creation process has proven essential in ensuring Sentinel meets the specific needs of FRs and frontline workers. Preliminary feedback highlights the app's relevance and usability. Our pilot testing began in August 2025 and is planned for completion by August 2026. As of January 16, 2026, a total of 119 participants have completed the initial survey, and 59 have downloaded the Sentinel app, with 24 having used the app for 6 weeks and completed the follow-up survey. Of this group, 13 participants have consented to an interview. We are on track to meet our target recruitment sample, and these data will provide detailed insights to inform refinements and determine readiness for a larger efficacy trial.

Conclusions: Sentinel represents a novel, personalized digital solution designed to address the unmet needs of FRs and frontline workers exposed to occupational trauma. Future work will evaluate its capacity to improve mental well-being, support trauma recovery, and build resilience.

International Registered Report Identifier (IRRID): PRR1-10.2196/72250

(JMIR Res Protoc 2026;15:e72250) doi:10.2196/72250

KEYWORDS

first responders; frontline workers; occupational trauma; digital intervention; participatory methods; mental well-being; trauma management; resilience; co-creation

Introduction

First responders (FRs), including firefighters, police officers, paramedics, and emergency health care workers, as well as other frontline workers, are at increased risk of experiencing traumatic stress-related conditions resulting from exposure to work-based traumas [1-3]. FRs are exposed to unpredictable and/or dangerous environments, demanding workloads, and heavy emotional labor, making such occupations unique in terms of their demands [4-6]. Almost 90% of FRs describe repeated

XSL•FO
RenderX

workplace exposures to incidents that involve direct threats to their lives and/or witnessing the deaths and horrific injuries of others [7,8]. Work-based traumas have been found to have a prolonged negative impact on the mental well-being of FRs, including an increased susceptibility to posttraumatic stress disorder, anxiety, depression, chronic fatigue, insomnia, burnout, and suicide [9-12]. These conditions are associated with higher rates of absenteeism, poorer work productivity, and unhealthy lifestyle behaviors, resulting in poorer physical and mental health [13-15].

The workplace culture of FR organizations strongly esteems strength and self-reliance, which often inhibits FRs from seeking mental health care and support [16]. Seeking help for mental health is often viewed as a form of weakness and/or as a barrier to professional advancement, and more than one-third of FRs experiencing mental ill health do not seek support due to stigmatization, shame, and confidentiality concerns [17-20]. Systematic reviews have synthesized the effectiveness of interventions with FR populations. Smith and Roberts [21] identified 10 intervention studies with FRs and tested their effectiveness, concluding that all interventions lacked quality because of general limitations, such as poor reporting, inadequate sample sizes, low response rates, and sampling bias. A systematic, theory-driven approach to intervention development is recommended to identify modifiable risk factors in at-risk FR groups with prospective studies, followed by the development of interventions to modify core risk factors [22]. The need for a more detailed description of treatment activities or reference to established treatment protocols for interventions with FRs has been recommended [23]. Alden et al [19] found that within 21 intervention studies for FRs, uptake and acceptability of interventions were low, and there was a lack of studies incorporating qualitative data drawn from interviews with FRs to help provide information about which aspects of the intervention were perceived as more or less helpful. Studies included individual or group psychological interventions delivered by clinicians (eg, psychologists) and nonclinicians (eg, experienced supervisors). It was recommended that researchers incorporate FRs' views on their workplace (eg, stressors and support) and draw participants from multiple organizations to collect both qualitative and quantitative information directly from the FRs [24]. The need for studies that incorporate participatory methods that involve FRs in the early stages of intervention development has also been emphasized [19]; this approach provides insight into FRs' views on what they would find most helpful in supporting their mental health needs and the circumstances under which mental health interventions work or do not work.

Recent work has highlighted how FRs experience challenges in accessing and engaging in face-to-face mental health care and support, as well as FRs' anticipation of negative outcomes from treatment, stigma, and how structural barriers inhibit treatment engagement [1]. Intervention research incorporating participatory methods from the onset has been found to improve engagement and uptake [25,26]. Additionally, Winders et al [20] investigated intervention research across FRs and concluded that interventions focusing on the prevention of mental health conditions (as opposed to the treatment) were most effective in reducing the risk of developing such conditions [27]. Although intervention research was limited, many of the systematic reviews agreed that more qualitative research using participatory methods was needed to gauge the specific lived experiences and mental health needs of FRs and that future interventions should aim to address the stigma surrounding mental health help-seeking and prioritize a preventative approach to mental ill health [18-20].

In recent years, digital interventions (DIs) have become popular to overcome existing barriers to accessing mental health support [28,29]. Studies have shown that DIs have been effective in addressing barriers related to stigma and accessibility present within face-to-face interventions [30]. Although DIs in the form of mobile apps or websites have also been shown to present their own problems, as they are dependent on the ability to use and access digital technology [31], they have been shown to be highly effective in the treatment of mental health conditions and have evolved into sophisticated, multifaceted interventions [28,32]. Internet and mobile technologies offer potentially critical ways of delivering mental health support [33]. While there has been very limited research involving DIs with FRs to date, recent studies have reported that DIs have the potential to improve mental health help-seeking behavior among FRs [34,35]. They offer the opportunity to identify early signs of distress and provide timely support, preventing mental health problems from worsening. They can also provide support daily at home or work and at a point when help is most needed and FRs are ready to engage [36]. In addition, they can be implemented on a large scale to meet the increasing demand for mental health support and reduce the burden on mental health services [37]. Nonetheless, research with FRs who are exposed to high rates of occupational trauma is limited, and it remains unclear how DIs can be developed to meet the mental health needs of FRs and ensure that they do not create or exacerbate traumatic experiences [38]. The need for researchers and DI developers to deeply understand users' needs and/or experiences is increasingly recognized [39]. This requires researchers to directly ask participants about their experiences in a trauma-informed way through interviews, focus groups, surveys, and/or workshops [40]. These forms of participant engagement have the potential to cause retraumatization by requiring people to remember and recount traumatic events [41]. In line with the goal of avoiding retraumatization, researchers should carefully consider how their research may be retraumatizing and work to minimize potential harm. A variety of actions and reflections to mitigate the possibility of retraumatizing research participants are recommended, including understanding the differences between truth-telling and healing, designing more expansive processes of informed consent, developing relational-based interviewing skills, and contextualizing trauma for the populations under study [42]. A trauma-informed approach to DI user research requires researchers to be attuned to the power dynamics in researcher-participant relationships; be transparent about the research's goals, expectations, and procedure; have skills in managing participant distress; and give the participants autonomy, flexibility, and opportunity to withdraw their participation if needed [40,43]. Under these conditions, having the opportunity to discuss traumatic experiences can have potential therapeutic benefits and can aid the healing process

XSL•FO
RenderX

[44]. It is also important that researchers conducting qualitative research on trauma need to recognize the potential for vicarious trauma for themselves and be able to implement appropriate protection and self-care strategies [45], as well as to plan for researcher safety and well-being from the onset of the research [46].

We describe a participatory, co-creation approach to developing a DI (Sentinel) that aims to help manage and prevent traumatic stress among FRs and our plans to pilot test this DI. The importance of understanding and accommodating the perspectives of FRs at all stages of its development and testing is central to our research and design process. The first step involves qualitative research with a wide range of people from the target user populations, carried out at every stage of intervention development, from planning to feasibility testing and implementation [47]. This process enables intervention designers to build a deep understanding of users' needs in terms of psychosocial factors as well as behavioral elements of the DI [48]. Insights from this co-creation process can be used to anticipate and interpret intervention usage and outcomes and, importantly, to modify the intervention to make it more acceptable, feasible, credible, and relevant to users [49]. Adopting participatory methods that involve FRs in all stages of intervention development can help highlight the distinctive ways that the intervention will address key context-specific social and behavioral issues facing FRs in meeting their mental health needs [28].

The aim of this paper is to describe the development and evaluation of the Sentinel DI, co-created to manage and prevent trauma among FRs and frontline workers. The research is structured into four distinct phases, each with specific objectives:

- Phase 1: To explore the needs, views, and experiences of FRs and frontline workers regarding the use of the Sentinel DI to support mental health. This phase involves conducting in-depth interviews and assessments to gather qualitative data on their mental health challenges, barriers to seeking support, and perceptions of DIs.
- Phase 2: To develop the DI content in consultation with current evidence, policy, theory, and insights gained from FRs' experiences during phase 1. This collaborative approach ensures that the content is evidence-based, contextually relevant, and tailored to the specific needs of the target population.
- Phase 3: Informed by the outcomes of phases 1 and 2, this phase will involve conducting a series of co-creation workshops and development sprints. The goal is to gather user feedback on a clickable prototype of the DI to test and refine its features, usability, and functionality. Engaging FRs and frontline workers in this iterative process ensures that the DI is user centered and effectively addresses their needs.
- Phase 4: The final phase entails conducting a mixed methods, nonrandomized study to determine the feasibility, acceptability, usability, and safety of the DI. This comprehensive evaluation will inform potential scalability and integration into support systems for FRs and frontline workers.

Given the elevated risk of distress within frontline worker populations, additional safety procedures are in place across both app use and research participation. Sentinel is not a crisis intervention; however, clear signposting to emergency and nonemergency support services is provided within the app and during onboarding. Standardized outcome measures allow for the identification of elevated distress, with participants advised to seek immediate support where risk is indicated. During qualitative interviews, researchers trained in trauma-informed methods will monitor distress, pause or terminate interviews if required, and follow a predefined escalation protocol with senior clinical oversight.

By following this structured, participatory approach, the study aims to develop a DI that is not only effective in managing and preventing trauma but also resonates with and is readily adopted by FRs and frontline workers.

Methods

Intervention: Sentinel

In the following phases, we describe the development and protocol for testing a novel digital app, Sentinel, designed to support both trauma management and trauma prevention among FRs. Sentinel distinguishes between trauma management—supporting regulation and coping in response to existing distress—and trauma prevention—reducing the risk of future harm through early identification, psychoeducation, and adaptive support. These functions are delivered through distinct but complementary artificial intelligence (AI)–driven personalization pathways. The innovation lies in the platform's ability to learn from user interactions and adapt notifications and content over time, enabling the ongoing personalization of interventions that support trauma management and overall health and well-being. The app will be available in a series of sector-specific variants such as Sentinel Green (Health and Social Care), Sentinel Blue (Police), and Sentinel Red (Fire Service), all of which deliver a unique experience specifically designed and tailored to individuals in that sector. Sentinel has been developed through the collaboration of expertise from behavioral science, computer science, physical activity and health, mental health, digital health, clinical psychology, technical product development, and intervention development. This multidisciplinary, combined approach is novel in helping develop an evidence-based and co-created product that is distinct from those currently available in the market.

Personalization Logic, User Data, and Safety Safeguards

Sentinel uses a transparent, rule-based personalization approach to tailor content while prioritizing safety, fairness, and user autonomy. Personalization is achieved through predefined pathways rather than autonomous or self-learning AI systems. User data informing personalization are limited to nondiagnostic, self-reported, and in-app behavioral inputs provided directly by users. These include onboarding self-report information relating to well-being priorities, user-selected goals, patterns of content engagement (eg, viewed, liked, or skipped content), and optional brief mood check-ins that are used to reassess needs over time.

XSL•FO
RenderX

Decision logic is governed by a rule-based persona framework developed through co-creation with FRs and informed by clinical expertise and evidence-based practice. Personas represent broad, nondiagnostic profiles linked to common patterns of occupational stress rather than clinical conditions. On the basis of user inputs and engagement, Sentinel recommends relevant modules and notifications aligned with stated goals and preferences. No diagnostic inference, risk prediction, or unsupervised algorithmic decision-making is performed. Safety and fairness safeguards are embedded throughout the system. Sentinel does not profile users based on protected characteristics, and conservative thresholds are applied when adapting content to avoid abrupt or inappropriate changes. Sentinel is not a crisis intervention; clear signposting to emergency and nonemergency support services is provided within the app and during onboarding. Human oversight is maintained through multidisciplinary clinical and research governance, with personalization rules refined iteratively through feasibility testing and user feedback.

Phase 1: User Engagement and Planning

We conducted an initial market analysis with key stakeholders in March 2024, which highlighted the need for an app such as Sentinel in the current digital apps market. Furthermore, we have seen a huge demand for continual professional development workshops (led by NC) on trauma awareness and management from FRs and frontline workers, and this further supports the market gap. We conducted a qualitative investigation into FRs' exposure to workplace trauma and their views and experiences of using DIs to support their mental health needs. Having gained approval from the University Ethics Committee and informed consent from participants, we conducted in-depth, online one-to-one semistructured interviews with FRs (total, N=54) from a range of organizational contexts (paramedics, n=21; fire and emergency, n=15; emergency health care workers, n=12; and police, n=6) using the Teams platform. All interviewers were provided with in-depth training on conducting qualitative research in trauma, led by a Health & Care Professions Council–registered practitioner consultant clinical psychologist and highly experienced trauma therapist and senior academic. Training included qualitative interview techniques, managing distress, maintaining safety, awareness of vicarious trauma in researchers, and self-care [50]. The training methods used included role-playing, scripted case studies, and online video content on safeguarding, critical incident reporting, and conducting research on sensitive topics [51]. The research team also held regular reflexive meetings to discuss the research process and provide support throughout the conduct of the research. The development of the interview schedule was informed by earlier research [52,53] and focused on FRs' experiences of exposure to trauma in the workplace and its impact, exploring the domains of (1) mental health and well-being, (2) feelings of safety, (3) help-seeking and supports, and (4) interventions. Questions were developed in accordance with guidelines for ensuring the safety and promotion of resilience of research participants and researchers and were co-created in collaboration with FR organizations during the period of project inception [41,50,54]. Interviews were transcribed in full verbatim and analyzed using a reflexive thematic approach [55]. The major themes developed were (1) the pervasive and salient impact of occupational trauma on mental health (self and others), (2) the demands of the job exacerbating the adverse effects of trauma, (3) insufficient support or access to psychological input following exposure to trauma (lack of psychological safety), (4) stigma and fear of judgment as barriers to mental health help-seeking, and (5) need for accessible, credible, and timely trauma-focused interventions and workplace support (acceptability of a DI). We provide a full description of these themes in our earlier work, and the implications emphasize the importance of implementing a strengths-based, nonpathologizing, and destigmatizing approach to trauma in the workplace as experienced by FRs [2]. The importance of overcoming barriers to accessing mental health support and improving access to evidence-based trauma-focused psychological interventions and workplace support was evident across FRs' accounts. The perceived acceptability of using a DI to support their mental health needs was reported; however, participants emphasized the need for safety, privacy, and anonymity when using the DI as well as opportunities to connect with other FRs either face-to-face or on a digital platform as important considerations. Participants also reported that they felt that using a DI should be an option as either a stand-alone intervention or in combination with other treatment options (eg, face-to-face therapy, peer support, and online resources). These findings provided evidence of the impact of exposure to workplace trauma on FRs and the importance of overcoming barriers to accessing mental health support. The acceptability of DIs as a means of providing flexible, accessible, and credible help was supported. These themes informed the next phase of intervention development.

Phase 2: Evidence-Based and User-Informed Content Creation

The second phase of developing Sentinel (creation of content) was completed between April 2024 and March 2025. The first step in designing the Sentinel app consisted of developing a list of evidence-based "personas." These are app user (FR) profiles aligned with information that users input relating to their presenting problems at the onset of using Sentinel (Figure 1A). The "personas" allow Sentinel to decide what intervention modules to give to users, and when, depending on their experiences of traumatic stress and other core presenting problems (comorbid anxiety and/or depressive symptoms; Figure 1B). The second step involved creating the in-app evidence-based module content (Table 1). A knowledge synthesis was adopted to enable the inclusion of a wide range of research to inform content generation and provide a nuanced interpretation and critique of the existing evidence base. The electronic search included the following databases: MEDLINE, PsycINFO, Embase, CINAHL, Scopus, Web of Science, Cochrane Library, and Google Scholar. The following search terms were used: "psychotraumatology," "trauma," "first responders," "healthcare workers," "mental health," "help-seeking," "well-being," "intervention," "digital," and "trauma treatment." The inclusion criteria were all types of articles, including primary qualitative, quantitative, and mixed methods data; case studies; clinical guidelines; reviews; and opinion texts. The exclusion criteria were articles for which the

XSL•FO
RenderX

full text was not available, articles not published in English, or gray literature.

Figure 1. Sentinel Wireframes showing (A) Sentinel asking the user to answer a question to begin determining their "persona"; (B) commencement of a tailored program for the user; and (C) one piece of content within the app.

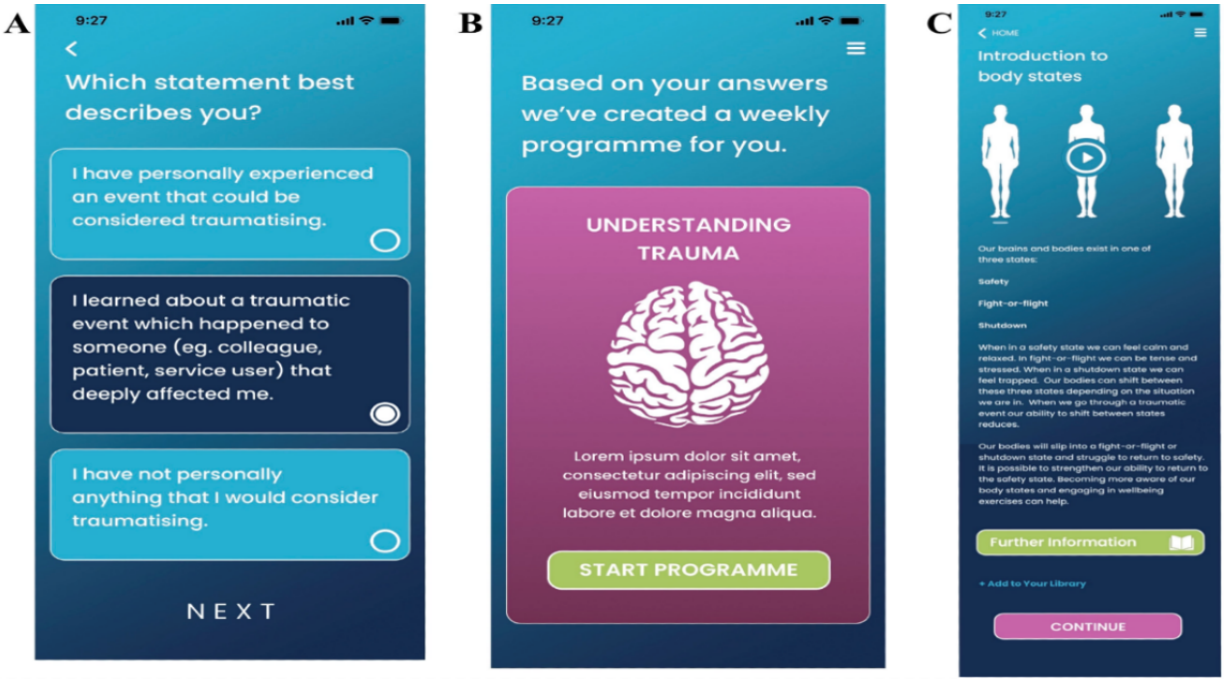

Table . Categories of Sentinel's modules.

Category	Description	Modules, n	Pieces of content, n
Increasing knowledge of trauma and traumatic stress	Provides psychoeducation on primary and secondary trauma and stress	3	7
Increasing knowledge of posttraumatic growth	Provides psychoeducation on posttraumatic growth and an exercise to foster optimism	3	5
Increasing mind-body and emotional state awareness	Provides exercises to bolster awareness of body state and emotions	12	60
Learning strategies to cope with present distress	Provides self-help and problem-solving strategies to cope with stressors	18	88
Improving general mood and affect	Provides short exercises to improve mood and affect	10	39
Improving sleep	Provides psychoeducation on sleep and strategies to improve sleep	5	16
Improving social connectedness	Provides a strategy to increase and strengthen user's social connections	1	3
Increasing levels of physical activity	Provides aerobic, strength, endurance, and resistance exercises to promote physical activity	4	67
Reducing levels of sedentary behavior	Provides psychoeducation on sedentary behavior, tips to reduce it, and short exercises	5	9
Increasing resilience	Provides an exercise to increase resilience to future traumatic events	1	3

The articles retrieved (from March 2004 to February 2024) were identified through a search over a 20-year period. In the first round of the search, additional references were identified by a manual search among the cited references. This included evidence-based therapeutic textbooks relevant to the treatment of trauma. With a review of current evidence [27,56,57], policy [58], theory [59-61], clinical guidelines [62,63], and experts by experience in stage 1 (FRs), a total of 297 pieces of content

XSL•FO
RenderX

were generated, spread across 62 modules (see Figure 1C for an example piece of content). Module content drew upon cognitive behavioral techniques focused on changing maladaptive thoughts, behaviors, and coping strategies; mindfulness and spirituality to enhance stress management and maintain overall well-being; and approaches to develop skills in physiological regulation and social engagement, informed by Polyvagal Theory [64-67].

Integrating such approaches provides a comprehensive framework for addressing the complex and multifaceted nature of trauma, promoting holistic healing [68]. Sentinel adopts a strengths-based and deshaming approach to dealing with trauma, emphasizing the importance of resilience, positive attributes, and the facilitation of posttraumatic growth [36,69]. The aim is to empower FRs, thereby fostering recovery and personal development beyond trauma [70]. Accessibility rules were set in place to ensure content was clear and easily digestible. Specifically, each piece of content aimed to be no longer than 500 characters and to have good readability. Readability was tested using the Hemingway Editor, where "good" readability was considered a score of 9 or below. In addition, all the content was independently reviewed to check readability, as well as fidelity to the evidence base. A taxonomy was created to describe each module better. This included information on category (what does the module focus on; eg, mental health literacy and coping), type (what does the module specifically focus on within the category; eg, providing psychoeducation and teaching a self-help coping strategy), method (how can the content be engaged with; eg, imagery, writing, using senses, meditation, breathing, and an activity), targets (what "persona(s)" will benefit from the module), symptom (what problem(s) does the module aim to tackle), notification (whether the user receives notifications related to the module), and notification type (what does the notification ask of the user; eg, learn more about a topic and engage in an exercise). The modules are classified into 13 different categories, each informed by the themes gathered in stage 1 (Figure 2). The third step was to design how content would be delivered to the user. Sentinel is a personalized DI, as it will ask users which goals they would like to prioritize (eg, "feel less anxious" or "reduce stress") and modify content in line with this. It also adapts content according to a number of person-specific characteristics, including work environment and content preferences (ie, what content the user likes or does not like to engage with). Hence, it will take the user's individual circumstances and preferences into consideration when tailoring its content.

Figure 2. Interview themes informing content generation.

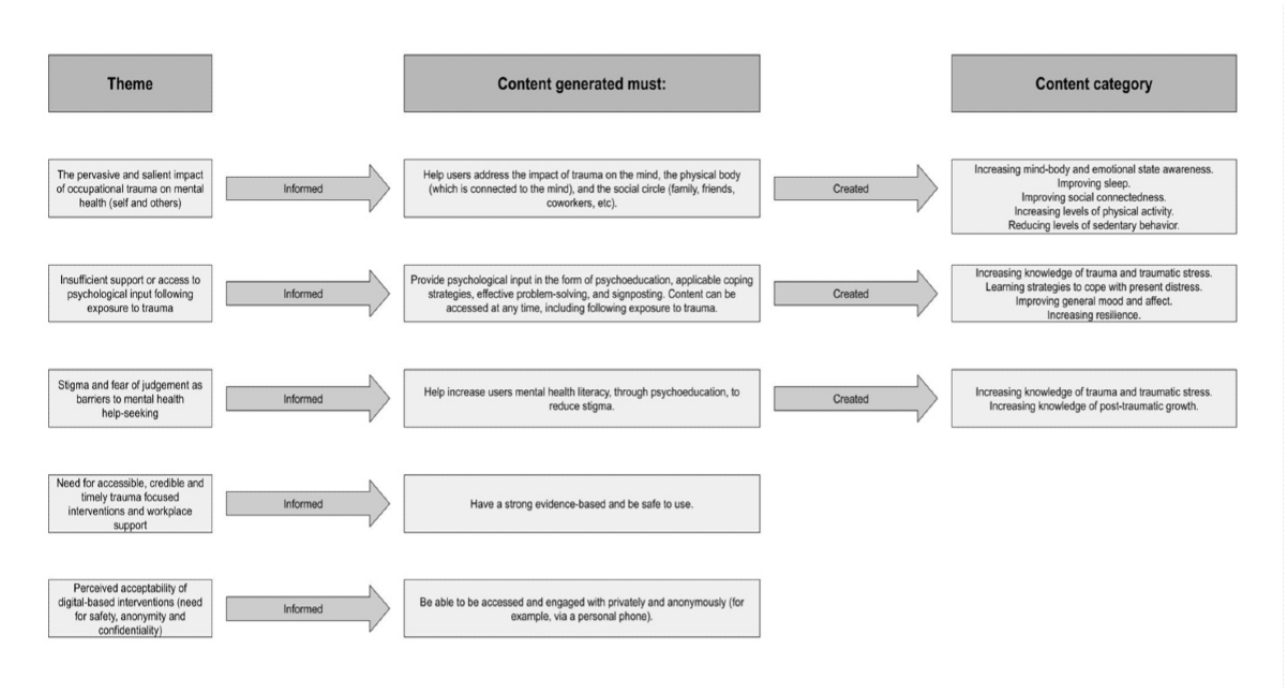

Across the user's journey, Sentinel will also ask a series of questions to allow it to determine which content users should be given to help deal with common problems associated with traumatic stress. For example, the user may be asked if they experience difficulties falling and staying asleep or recurring nightmares and may be given sleep-related content depending on their answer. The app uses push notifications and reminders to increase and maintain user engagement [71]. These notifications are delivered at regular intervals and are designed to boost FRs' overall engagement and well-being. Notifications also serve to reassess traumatic stress symptoms and maintain an adaptive, personalized management approach. It is, therefore, possible for a user's "persona" to change as they use Sentinel.

Phase 3: User-Informed Prototype Refinement

Informed by the outcomes of phases 1 and 2, our next steps will involve initial testing of the prototype for the Sentinel DI. We will conduct 3 co-creation video call sessions with FRs using Teams and simulations of notice boards or Post-its, and this will be complemented by collective discussions. The time immediately after our co-creation meetings will be dedicated to our development "sprints" to refine the Sentinel app. We have used this process previously in the development of DIs

XSL•FO
RenderX

[72,73]. The first co-creation session will present the principles of the DI and then demonstrate and critique a simple initial application prototype. The second co-creation session will review the feedback from meeting 1 and progress during the next design sprint. The third co-creation session will review and finalize the DI design in preparation for the pilot or feasibility study. If a participant cannot attend a group meeting, one-to-one calls will be offered as an alternative. We will recruit participants initially using convenience sampling, followed by snowballing (connections with FR groups and previous survey or interview participants). We will aim to identify 6 to 8 participants for the co-creation groups to maximize the depth of conversation achievable with our discussions [74]. We will aim to include participants representing multistakeholder views across a variety of FR groups, including fire and emergency, paramedics, police, and frontline health and social care workers. Inclusion criteria will include participants aged 18 years and older and interested in contributing to current knowledge of DIs for trauma management in FRs. Ahead of developing the initial minimum viable product of the DI, we created a high-fidelity clickable prototype using a web app mock-up tool (Figma) that is used to gain user insights (phase 3). The prototype allows us to test the basic operation of the DI on various devices with diverse FR user groups (N=100), thereby simulating the user experience of the DI before the intensive app development phase commences [75]. The prototype includes a sample content module, simulated mood reporting and journaling, as well as detailed navigation and functional elements. We have also created an online link to the prototype and a short survey, which will allow us to gather views regarding design, content, and relevance of the DI as well as suggestions for improvements from a larger and more diverse group of potential users.

Co-creation sessions will be guided by structured prompts focusing on perceived relevance and credibility of content, emotional safety, usability and navigation, clarity of language, and appropriateness of personalization and notifications. Participants will be asked what aspects of the prototype feel most and least helpful, whether any content may be distressing or unhelpful, and how the app could better fit the realities of FR work. Feedback will be collated across sessions and reviewed by the multidisciplinary team, with priorities for change agreed upon during postsession debriefs and implemented through iterative development sprints to refine content, functionality, and user flow ahead of pilot testing.

Phase 4: Pilot and Feasibility Testing of Sentinel

The final phase will involve assessing the feasibility and acceptability of Sentinel and exploring participants' experiences of using the DI, with quantitative and qualitative data to be collected. Approval to commence with piloting has been obtained from the University Ethics Committee (REF UEC22/92). All data will be stored on secure servers, in line with appropriate security and compliance standards. The data stored are relevant to the functioning of the app and delivery of relevant personalized experiences. This will comply with General Data Protection Regulation and any relevant institutional data governance policies and meet industry standards for information and data security.

Participants

We aim to pilot test the Sentinel digital app with a target sample of 100 FRs across diverse FR organizations. Participants who report distress associated with exposure to work-related traumas will be recruited from the National Health Service and FR organizations in the United Kingdom. Participants will be recruited through opportunity sampling via online posters distributed within FR organizations, emails sent directly to FR staff members, work intranet bulletin board advertisements, and targeted social media platforms. Interested participants will be eligible if they meet the following criteria: (1) aged 18 years or older, (2) being employed full-time or part-time at the National Health Service and/or an FR organization, (3) understanding the requirements of the study, and (4) having no major mental illness that may impact their capacity to consent to the trial. Given that the focus of the testing phase is to explore the feasibility, acceptability, safety, and usability of the DI, the sample size (N=100) is based on having sufficient variance to examine the feasibility for a larger study across a diverse population [76]. As this study is a feasibility pilot, it is not powered to detect statistically significant pre-post effects. The sample size (N=100) is based on recommendations for feasibility studies to estimate recruitment, retention, acceptability, and variance of outcome measures, which will inform sample size calculations for a future fully powered trial. Quantitative analyses will therefore focus on descriptive statistics, completeness of data, and estimation of effect sizes with CIs rather than hypothesis testing. The nonrandomized design was selected as appropriate for this feasibility study, which prioritizes assessment of acceptability, engagement, usability, and implementation rather than effectiveness. Feasibility will be judged using predefined progression indicators, including successful recruitment within the planned time frame, acceptable retention at follow-up, meaningful engagement with the app over the 6-week period, and satisfactory usability ratings, to inform the design of a future randomized controlled trial. All participants will receive the Sentinel app for 6 weeks. Notifications will encourage FRs to use the app, which is designed to be used in a stand-alone manner alongside routine care.

Procedure

Participants will first be asked to provide informed consent and then complete an online Qualtrics questionnaire including the trial outcome measures. Following completion of the preintervention online questionnaire, participants will be provided with access to Sentinel and instructed to use the app for a 6-week period. Following the 6-week intervention period, participants will complete an online postintervention questionnaire. As well as looking at differences in pre- and post-intervention outcomes among FRs, the possibility that there might be some variations across different types of FRs with respect to responsiveness to the intervention will be explored. Furthermore, a series of open-ended questions will be asked concerning participants' views on whether the Sentinel DI had worked for them and how it has impacted their mental health and well-being. An additional subset of FR participants (n=30) will then be asked to take part in an audio-recorded, one-to-one semistructured interview to gain a more in-depth understanding

XSL•FO
RenderX

of their experiences of using Sentinel. Furthermore, health care professionals and service managers who referred FRs to the trial (n=20) will be interviewed to explore the acceptability and usability of Sentinel. Participants will be selected according to a sampling framework to capture varied demographics, experiences of work-based trauma, and levels of engagement with the Sentinel app. We will explain to all potential participants that participation is voluntary and that they can withdraw their consent at any point during the study without providing a reason. It will be made clear to FRs that, as the intervention does not replace standard care (but is provided in addition to it), withdrawal from the intervention will not impact their ability to continue to access standard care within the referring service or other sources of support they might access contemporaneously.

All participant data, including sensitive mental health information, are stored on secure, encrypted servers in accordance with General Data Protection Regulation and institutional data governance requirements. Data are pseudonymized at the point of collection, with identifying information stored separately and accessible only to authorized members of the research team. Sentinel does not share individual-level data with employers or third parties, and no identifiable app use, journaling, or symptom data are accessible to organizations. Data are retained only for the duration necessary to meet the study objectives and regulatory requirements, after which they will be securely deleted. Participants are informed of these safeguards during consent to ensure transparency and trust.

Outcome Measures

To allow a comprehensive evaluation of the effectiveness of the Sentinel digital app in a future randomized controlled trial, we will work with stakeholders and subject experts to determine relevant outcome measures and then test the feasibility of their inclusion within this pilot study. This approach will also enable us to report an indication of effect and provide data to support sample size calculations for future studies. Potential metrics include employee well-being indices, rates of absenteeism and presenteeism, and overall organizational productivity. We will explore the integration of Sentinel with current organizational systems, particularly human resource platforms, which is essential for seamless implementation. We will assess the feasibility of such integration to ensure that Sentinel complements existing processes and enhances user engagement. Understanding the challenges and motivations that organizations face in evaluating the impact of DIs is vital. Factors such as data privacy concerns, resource allocation, and the need for stakeholder buy-in will be considered. By engaging with stakeholders, we aim to identify and address these challenges, ensuring a comprehensive evaluation strategy. This collaborative effort will guide the development of the outcome measures for our pilot trial, ensuring they are relevant, comprehensive, and aligned with organizational priorities. By involving stakeholders in this process, we aim to create a robust framework for assessing Sentinel's impact, facilitating its successful implementation and adoption within organizations.

As primary outcomes, we will collect information on the total number of participants recruited and retained at different time points, including consent completed, app download, app use over 6 weeks, and follow-up questionnaire completion. We will also collect demographic information, including gender, age, ethnicity, health status, working status, family status, and occupation and recruitment setting. Other proposed psychometric outcome measures to be considered as secondary outcomes include the *Abbreviated Post-Traumatic Stress Disorder Checklist–Civilian* used to measure self-reported trauma symptoms [77]; the *Warwick-Edinburgh Wellbeing Scale* [78]; the *Burnout Measure–Short Version* [79]; the *Depression, Anxiety, and Stress Scale* [80]; the *Neuroception of Psychological Safety* [81]; the *Utrecht Work Engagement Scale* [82]; the *Mobile App Rating Scale* [83]; the *Post-Traumatic Growth Short Form* [84]*;* the *International Physical Activity Questionnaire Short Version* [85]; and the *WHOQOL-BREF* [86]. The validity and reliability of all questionnaires have been well tested [87-89]. More details on the proposed outcome measures have been provided in Table S1 in Multimedia Appendix 1. We also envisage including some physiological outcomes and remote monitoring in a subgroup, and the details of this will be explored during our earlier phases. Following pilot testing, further technology-based adjustments, in addition to delivery components, may require revisions.

Data Analysis

As this is a nonrandomized feasibility study, analyses will focus on describing recruitment, retention, engagement, usability, and safety outcomes rather than testing intervention effectiveness. Feasibility outcomes will be summarized using descriptive statistics, including counts and proportions for recruitment, retention, app download, completion of follow-up measures, and engagement with Sentinel over the 6-week period, with usability and acceptability measures will be reported using means, SDs, medians, and ranges as appropriate. Exploratory pre-post outcomes will be examined descriptively to inform the design of a future fully powered trial, with standardized mean differences (eg, Cohen *d*) and 95% CIs reported for changes in psychometric measures between baseline and follow-up; no formal hypothesis testing will be undertaken. Missing data will be assessed by reporting the extent and patterns of missingness across measures and time points, with no imputation performed, given the feasibility focus of the study. Exploratory subgroup analyses (eg, by FR role or level of engagement) will be conducted using stratified descriptive summaries only and interpreted cautiously to generate hypotheses and inform future trial design rather than to support inferential comparisons. Qualitative interview data will be analyzed using reflexive thematic analysis, with integration of quantitative and qualitative findings used to inform refinement of the intervention and trial procedures.

Ethical Considerations

The study received full ethical approval from the University of Strathclyde (Ref UEC22/92 Cogan). Informed consent will be obtained from all participants and their right to privacy explained. The University of Strathclyde was the study sponsor, including management of compensation.

XSL•FO
RenderX

Results

Our pilot testing began in August 2025 and is planned for completion by August 2026. We anticipate a high attrition rate and will therefore recruit over a full year from several FR organizations using multiple recruitment methods, including organization newsletters and emails, target group presentations, and engagement of key stakeholders within organizations. As of January 16, 2026, a total of 119 participants have completed the initial survey, and 59 have downloaded the Sentinel app, with 24 having used the app for 6 weeks and completed the follow-up survey. Of this group, 13 participants have consented to an interview. We are therefore on track to meet our target recruitment sample.

Once data collection is complete, we expect to report results on the number of eligible FR participants consenting and the total number recruited and retained at different time points, including consent completed, app download, app use over 6 weeks, and follow-up questionnaire completion. We will also report on information about recruitment setting to inform the sampling strategy of a future randomized trial; completeness of outcome measures, including psychometric questionnaires and an indication of effect; attrition rates and reasons for withdrawal; the range of FR organizations offered Sentinel; participant completion rates for Sentinel; Sentinel platform use and engagement analytics; and safety of the Sentinel app and our trial procedures. We also expect to report on perceived barriers and enablers to integration and uptake into existing mental health care provider pathways. Exploratory analyses will be conducted to examine potential variation in intervention responsiveness across different FR groups. Given the feasibility focus and anticipated sample sizes within subgroups, these analyses will be descriptive and exploratory in nature, using stratified summaries and estimation of pre-post change scores by responder group rather than formal hypothesis testing. Findings will be used to inform subgroup specification and analytic strategies for a future fully powered trial.

Discussion

Implication of Work

In this paper, first, we have described the process for the development of the Sentinel DI adopting a participatory co-creation approach, and, second, the protocol for testing the DI. We outlined the phase-based approach we have adhered to in developing the DI and our protocol for ascertaining the feasibility, acceptability, safety, and usability of the DI and how best to integrate it into existing routine care pathways for FRs. As recommended for DI development, we have emphasized the importance of FR user engagement and planning during all phases of the process [90]. To the best of our knowledge, this is the first DI of this type to be co-created with FRs worldwide. We have carried out in-depth qualitative interviews with FRs, held regular advisory group meetings with key stakeholders, incorporated input and feedback from this work into the development of the DI content and user interface, and created evidence-based content for the DI [2]. Ahead of developing the minimum viable product of the DI, we created a high-fidelity clickable prototype that will be used to gain user insights before further app development and pilot testing commences. The user-driven, co-creation approach we have adopted is intended to ensure that the DI is engaging and relevant and reflects the needs and experiences of FRs with a focus on trauma management and prevention, deshaming mental health help-seeking, and fostering resilience and posttraumatic growth [36,91].

The DI has the potential to provide many benefits, including increased awareness of trauma and trauma-related symptoms; accurate identification of primary and secondary trauma; immediate support for FRs and emergency frontline workers; reduced costs and waiting times for therapy; early-stage trauma management and prevention; remote triage leading to self-management; advice on seeking in-person treatment; advanced-stage signposting to mental health professionals; and a unique and tailored experience based on sector, trauma type, and data science [92]. Technology is leading to rapid change in the way mental health interventions are increasingly being delivered. To create inclusive opportunities for people who experience mental health challenges associated with trauma, any DI development needs to not only address the present time but also to anticipate and influence future technological directions [93]. To prevent current inequalities and biases from becoming “hardwired” into DIs and future advancements in algorithms for AI, it is essential that people with lived experiences of mental health challenges are involved in the co-creation of such technologies [94]. People with mental health challenges are one of several groups that are more likely to be digitally excluded (do not have access to or are unable to access the internet). Therefore, they may not have the opportunity to engage in DIs designed to support their mental health needs and experience further exclusion [91]. Developments in technology to help support and treat people experiencing trauma are raising new ethical questions about the current and future rights of people accessing DIs, in particular their rights regarding surveillance, privacy, and what it means to be human [95]. In our work with Sentinel, we aim to consider such broader ethical and moral issues and continue strengthening our partnership working with FRs and other key stakeholders in our co-creation process and future AI innovations [94]. This includes ensuring privacy and confidentiality, addressing potential biases in algorithms, and ensuring that DIs complement, rather than replace, human care [96]. Training and education for mental health professionals and community organizations on the use of DIs are crucial [97]. This ensures that they are equipped to integrate these tools into their mental health practices and can provide support to individuals using such interventions. Training should also emphasize the importance of maintaining a human-centered approach to mental health care [98].

Limitations

As this protocol is for a pilot study, we are unable to conclude whether the Sentinel intervention will bring about any clinically significant change after intervention. However, this study design, with its focus on feasibility, acceptability, safety, and usability, will provide valuable insights that will inform parameters for a future randomized trial, should this be indicated. The inclusion of organizationally led impact outcomes, psychometric

XSL•FO
RenderX

measures, and open-ended questions asking FR users whether the DI worked for them and how it had impacted their mental health and well-being will provide some preliminary evidence of its efficacy. Moreover, feedback from multiple stakeholders, including those who refer FRs to the feasibility trial, will allow for improvements in the Sentinel app, trial procedures, and plans for implementation. In our work, to date, we have not yet engaged with military FRs and/or family members of FRs; we would view this as a priority in our next steps in DI development, given their central role in providing social and/or familial support to FRs [99]. This aligns with emerging work on persistent traumatic stress exposure [100], which highlights the importance of considering ongoing and cumulative trauma within relational and support systems.

Conclusions

In conclusion, the co-creation process has been instrumental in developing Sentinel, a personalized digital app tailored to the unique needs of FRs. This participatory approach has ensured that the intervention is both relevant and user-centered. However, further research is essential to evaluate Sentinel's acceptability, usability, feasibility, and safety across diverse FR service settings. Such evaluations will inform necessary refinements, maximizing the app's effectiveness in trauma management and prevention. Addressing potential concerns and safety issues is crucial to enhance user engagement and ensure the intervention's success. This forthcoming feasibility trial represents a pivotal step in bolstering support for FRs, highlighting a critical area for both research and intervention development. Pilot testing of our DI lays the groundwork for a more definitive solution that can be rigorously tested in future studies and ultimately implemented to support FRs across various service environments.

Funding

This work was funded by University of Strathclyde Stage Gate 1 and 2 funding.

Data Availability

Due to the potentially sensitive and personal nature of the data collected in the conduct of the study, the dataset supporting the results will not be publicly available. However, the lead author will retain the anonymized data on a secure university data storage platform and can share by request for research purposes.

Authors' Contributions

NC, AK, and CG all contributed to conceptualization, methodology, analysis, writing the original draft, and reviewing and editing the paper. Authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

NC holds an advisory leadership role (Chief Scientific Officer) with Sentinel Ltd. NC, AK, and CG are named contributors on research-related intellectual property owned by the University of Strathclyde and may be eligible to receive future royalty payments under the University's totalities policy. The University of Strathclyde retains full ownership of all relevant research intellectual property. This protocol paper reports no commercial activity and has no commercial interests. The authors declare no other competing interests.

Multimedia Appendix 1
Details of proposed outcomes measures.
[DOCX File, 19 KB - resprot_v15i1e72250_app1.docx]

References

1. Cogan N, Kennedy C, Beck Z, et al. ENACT study: what has helped health and social care workers maintain their mental well-being during the COVID-19 pandemic? Health Soc Care Community 2022 Nov;30(6):e6656-e6673. [doi: 10.1111/hsc.13992] [Medline: 36068667]
2. Cogan N, Craig A, Milligan L, et al. "I've got no PPE to protect my mind": understanding the needs and experiences of first responders exposed to trauma in the workplace. Eur J Psychotraumatol 2024;15(1):2395113. [doi: 10.1080/20008066.2024.2395113] [Medline: 39238472]
3. Geuzinge R, Visse M, Duyndam J, Vermetten E. Social embeddedness of firefighters, paramedics, specialized nurses, police officers, and military personnel: systematic review in relation to the risk of traumatization.. Front Psychiatry 2020;11:496663. [doi: 10.3389/fpsyt.2020.496663] [Medline: 33408646]
4. Garner N, Baker J, Hagelgans D. The private traumas of first responders. J Individ Psychol 2016;72(3):168-185. [doi: 10.1353/jip.2016.0015]
5. Johnson CC, Vega L, Kohalmi AL, Roth JC, Howell BR, Van Hasselt VB. Enhancing mental health treatment for the firefighter population: understanding fire culture, treatment barriers, practice implications, and research directions. Prof Psychol Res Pract 2020;51(3):304-311. [doi: 10.1037/pro0000266]

XSL•FO
RenderX

6. Stansfeld SA, Fuhrer R, Head J. Impact of common mental disorders on sickness absence in an occupational cohort study. Occup Environ Med 2011 Jun;68(6):408-413. [doi: 10.1136/oem.2010.056994] [Medline: 21075767]
7. Boffa JW, Stanley IH, Hom MA, Norr AM, Joiner TE, Schmidt NB. PTSD symptoms and suicidal thoughts and behaviors among firefighters. J Psychiatr Res 2017 Jan;84:277-283. [doi: 10.1016/j.jpsychires.2016.10.014] [Medline: 27810667]
8. Johansson UE, Johansson Å, Grimby A. Psychosocial workload of Swedish ambulance and emergency room personnel with high prevalence of dying, death and grieving relatives. A descriptive and comparison study. AJNS 2014;3(5):56. [doi: 10.11648/j.ajns.20140305.11]
9. Back CY, Hyun DS, Chang SJ, Jeung DY. Trauma exposure and suicidal ideation among Korean male firefighters: examining the moderating roles of organizational climate. Saf Health Work 2023 Mar;14(1):71-77. [doi: 10.1016/j.shaw.2022.11.005] [Medline: 36941935]
10. Hoell A, Kourmpeli E, Dressing H. Work-related posttraumatic stress disorder in paramedics in comparison to data from the general population of working age. A systematic review and meta-analysis. Front Public Health 2023;11:1151248. [doi: 10.3389/fpubh.2023.1151248] [Medline: 36969652]
11. Kim W, Bae M, Chang SJ, et al. Effect of burnout on post-traumatic stress disorder symptoms among firefighters in Korea: data from the Firefighter Research on Enhancement of Safety & Health (FRESH). J Prev Med Public Health 2019 Nov;52(6):345-354. [doi: 10.3961/jpmph.19.116] [Medline: 31795611]
12. Vizheh M, Qorbani M, Arzaghi SM, Muhidin S, Javanmard Z, Esmaeili M. The mental health of healthcare workers in the COVID-19 pandemic: a systematic review. J Diabetes Metab Disord 2020 Dec;19(2):1967-1978. [doi: 10.1007/s40200-020-00643-9]
13. Alshehri FF, Alghamdi SA, Alrashoudi AM, et al. Post-traumatic stress disorder and its associated risk factors among emergency healthcare workers: a Saudi cross-sectional analytical study. Cureus 2023 Aug;15(8):e44327. [doi: 10.7759/cureus.44327] [Medline: 37779819]
14. Meckes SJ, McDonald MA, Lancaster CL. Association between physical activity and mental health among first responders with different service roles. Psychol Trauma 2021 Jan;13(1):66-74. [doi: 10.1037/tra0000971] [Medline: 32940520]
15. Williams SD, Williams J. Posttraumatic stress in organizations: types, antecedents, and consequences. Business and Society Review 2020 Mar;125(1):23-40. [doi: 10.1111/basr.12192]
16. Jones S, Agud K, McSweeney J. Barriers and facilitators to seeking mental health care among first responders: “removing the darkness”. J Am Psychiatr Nurses Assoc 2020;26(1):43-54. [doi: 10.1177/1078390319871997] [Medline: 31509058]
17. Drew JM, Martin S. A national study of police mental health in the USA: stigma, mental health and help-seeking behaviors. J Police Crim Psych 2021 Jun;36(2):295-306. [doi: 10.1007/s11896-020-09424-9]
18. Johnston S, Sanderson K, Bowes L, Wild J. Why do ambulance employees (not) seek organisational help for mental health support? A mixed-methods systematic review protocol of organisational support available and barriers/facilitators to uptake. BMJ Open 2022 Oct 10;12(10):e062775. [doi: 10.1136/bmjopen-2022-062775] [Medline: 36216429]
19. Alden LE, Matthews LR, Wagner S, et al. Systematic literature review of psychological interventions for first responders. Work Stress 2021 Apr 3;35(2):193-215. [doi: 10.1080/02678373.2020.1758833]
20. Winders WT, Bustamante ND, Garbern SC, et al. Establishing the effectiveness of interventions provided to first responders to prevent and/or treat mental health effects of response to a disaster: a systematic review. Disaster Med Public Health Prep 2021 Feb;15(1):115-126. [doi: 10.1017/dmp.2019.140] [Medline: 33870882]
21. Smith A, Roberts K. Interventions for post-traumatic stress disorder and psychological distress in emergency ambulance personnel: a review of the literature. Emerg Med J 2003 Jan;20(1):75-78. [doi: 10.1136/emj.20.1.75] [Medline: 12533381]
22. Wild J, El-Salahi S, Esposti MD. The effectiveness of interventions aimed at improving well-being and resilience to stress in first responders. Eur Psychol 2020 Oct;25(4):252-271. [doi: 10.1027/1016-9040/a000402]
23. Haugen PT, Evces M, Weiss DS. Treating posttraumatic stress disorder in first responders: a systematic review. Clin Psychol Rev 2012 Jul;32(5):370-380. [doi: 10.1016/j.cpr.2012.04.001] [Medline: 22561967]
24. Nielsen MBD, Bültmann U, Amby M, Christensen U, Diderichsen F, Rugulies R. Return to work among employees with common mental disorders: study design and baseline findings from a mixed-method follow-up study. Scand J Public Health 2010 Dec;38(8):864-872. [doi: 10.1177/1403494810384424] [Medline: 20855357]
25. MacIntyre G, Cogan N, Stewart A, et al. Understanding citizenship within a health and social care context in Scotland using community-based participatory research methods. In: SAGE Research Methods Cases: SAGE Publications Ltd; 2018. URL: https://methods.sagepub.com/case/citizenship-within-health-social-care-in-scotland-participatory-research [accessed 2026-03-02] [doi: 10.4135/9781526484918]
26. Wallerstein N, Duran B. Community-based participatory research contributions to intervention research: the intersection of science and practice to improve health equity. Am J Public Health 2010 Apr 1;100 Suppl 1(Suppl 1):S40-S46. [doi: 10.2105/AJPH.2009.184036] [Medline: 20147663]
27. Bisson JI, Olff M. Prevention and treatment of PTSD: the current evidence base. Eur J Psychotraumatol 2021;12(1):1824381. [doi: 10.1080/20008198.2020.1824381] [Medline: 34992739]
28. Borghouts J, Eikey E, Mark G, et al. Barriers to and facilitators of user engagement with digital mental health interventions: systematic review. J Med Internet Res 2021 Mar 24;23(3):e24387. [doi: 10.2196/24387] [Medline: 33759801]

XSL•FO
RenderX

29. Schäfer SK, von Boros L, Schaubruch LM, et al. Digital interventions to promote psychological resilience: a systematic review and meta-analysis. NPJ Digit Med 2024 Feb 8;7(1):30. [doi: 10.1038/s41746-024-01017-8] [Medline: 38332030]
30. Gulliver A, Calear AL, Sunderland M, et al. Consumer-guided development of an engagement-facilitation intervention for increasing uptake and adherence for self-guided web-based mental health programs: focus groups and online evaluation survey. JMIR Form Res 2020 Oct 29;4(10):e22528. [doi: 10.2196/22528] [Medline: 33118939]
31. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Barriers and facilitators to patient and public engagement and recruitment to digital health interventions: protocol of a systematic review of qualitative studies. BMJ Open 2016 Sep 2;6(9):e010895. [doi: 10.1136/bmjopen-2015-010895] [Medline: 27591017]
32. Liu S, Ma J, Sun M, Zhang C, Gao Y, Xu J. Mapping the landscape of digital health intervention strategies: 25-year synthesis. J Med Internet Res 2025 Jan 13;27:e59027. [doi: 10.2196/59027] [Medline: 39804697]
33. Ruzek JI, Yeager CM. Internet and mobile technologies: addressing the mental health of trauma survivors in less resourced communities. Glob Ment Health (Camb) 2017;4:e16. [doi: 10.1017/gmh.2017.11] [Medline: 29230312]
34. Newell CJ, Ricciardelli R, Czarnuch SM, Martin K. Police staff and mental health: barriers and recommendations for improving help-seeking. Police Pract Res 2022 Jan 2;23(1):111-124. [doi: 10.1080/15614263.2021.1979398]
35. Yoon S, Goh H, Nadarajan GD, et al. Perceptions of mobile health apps and features to support psychosocial well-being among frontline health care workers involved in the COVID-19 pandemic response: qualitative study. J Med Internet Res 2021 May 31;23(5):e26282. [doi: 10.2196/26282] [Medline: 33979296]
36. Wild J, Greenberg N, Moulds ML, et al. Pre-incident training to build resilience in first responders: recommendations on what to and what not to do. Psychiatry (Abingdon) 2020 Apr 2;83(2):128-142. [doi: 10.1080/00332747.2020.1750215]
37. Taylor CB, Ruzek JI, Fitzsimmons-Craft EE, et al. Using digital technology to reduce the prevalence of mental health disorders in populations: time for a new approach. J Med Internet Res 2020 Jul 24;22(7):e17493. [doi: 10.2196/17493] [Medline: 32706665]
38. Wickersham A, Petrides PM, Williamson V, Leightley D. Efficacy of mobile application interventions for the treatment of post-traumatic stress disorder: a systematic review. Digit Health 2019;5:2055207619842986. [doi: 10.1177/2055207619842986] [Medline: 31019722]
39. Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, Manicavasagar V. Mobile mental health: review of the emerging field and proof of concept study. J Ment Health 2011 Dec;20(6):509-524. [doi: 10.3109/09638237.2011.608746] [Medline: 21988230]
40. Chen JX, McDonald A, Zou Y, et al. Trauma-informed computing: towards safer technology experiences for all. 2022 Apr 29 Presented at: CHI '22; New Orleans LA USA p. 1-20 URL: https://dl.acm.org/doi/proceedings/10.1145/3491102 [accessed 2026-03-08] [doi: 10.1145/3491102.3517475]
41. Isobel S. Trauma-informed qualitative research: some methodological and practical considerations. Int J Ment Health Nurs 2021 Oct;30 Suppl 1:1456-1469. [doi: 10.1111/inm.12914] [Medline: 34310829]
42. Howe K. Trauma to self and other: reflections on field research and conflict. Secur Dialogue 2022 Aug 1;53(4):363-381. [doi: 10.1177/09670106221105710]
43. Wong R. Guidelines to incorporate trauma-informed care strategies in qualitative research. Urban Institute. 2021 Aug 30. URL: https://www.urban.org/urban-wire/guidelines-incorporate-trauma-informed-care-strategies-qualitative-research [accessed 2023-12-05]
44. McMyler B. Why testimony?: trauma, credibility, and the work of truth-telling.. RIVISTA SPERIMENTALE DI FRENIATRIA 2021;2021(1):119-140. [doi: 10.3280/RSF2021-001007]
45. Casas JB, Benuto LT. Breaking the silence: a qualitative analysis of trauma narratives submitted online by first responders. Psychol Trauma 2022 Feb;14(2):190-198. [doi: 10.1037/tra0001072] [Medline: 34410810]
46. Smith E, Pooley JA, Holmes L, Gebbie K, Gershon R. Vicarious trauma: exploring the experiences of qualitative researchers who study traumatized populations. Disaster Med Public Health Prep 2023;17:e69. [doi: 10.1017/dmp.2021.333]
47. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015;17(1):e30. [doi: 10.2196/jmir.4055]
48. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017;19(6):e232. [doi: 10.2196/jmir.7126]
49. Laukka E, Tuunanen T, Jansson M, et al. Value cocreation and codestruction in digital health services: protocol for a systematic review. JMIR Res Protoc 2025 Jan 14;14(1):e63015. [doi: 10.2196/63015]
50. AbiNader MA, Messing JT, Pizarro J, Kappas Mazzio A, Turner BG, Tomlinson L. Attending to our own trauma: promoting vicarious resilience and preventing vicarious traumatization among researchers. Soc Work Res 2023 Nov 17;47(4):237-249. [doi: 10.1093/swr/svad016]
51. Shea SC, Barney C. Teaching clinical interviewing skills using role-playing: conveying empathy to performing a suicide assessment: a primer for individual role-playing and scripted group role-playing. Psychiatr Clin North Am 2015 Mar;38(1):147-183. [doi: 10.1016/j.psc.2014.10.001] [Medline: 25725575]

XSL•FO
RenderX

52. Cogan N, Archbold H, Deakin K, et al. What have we learned about what works in sustaining mental health care and support services during a pandemic? Transferable insights from the COVID-19 response within the NHS Scottish context. Int J Ment Health 2022 Apr 3;51(2):164-188. [doi: 10.1080/00207411.2022.2056386]
53. Greinacher A, Derezza-Greeven C, Herzog W, Nikendei C. Secondary traumatization in first responders: a systematic review. Eur J Psychotraumatol 2019;10(1):1562840. [doi: 10.1080/20008198.2018.1562840] [Medline: 30719236]
54. Alessi EJ, Kahn S. Toward a trauma-informed qualitative research approach: guidelines for ensuring the safety and promoting the resilience of research participants. Qual Res Psychol 2023 Jan 2;20(1):121-154. [doi: 10.1080/14780887.2022.2107967]
55. Braun V, Clarke V. Can I use TA? Should I use TA? Should I not use TA? Comparing reflexive thematic analysis and other pattern‐based qualitative analytic approaches. Couns and Psychother Res 2021 Mar;21(1):37-47. [doi: 10.1002/capr.12360]
56. Bryant RA. Early intervention after trauma. In: Evidence Based Treatments for Trauma-Related Psychological Disorders: A Practical Guide for Clinicians: Springer Nature Switzerland AG; 2022:135-159. [doi: 10.1007/978-3-030-97802-0_7]
57. Headley AM, Witkowski K, Remington C, Ganapati NE, Contreras SL. Trauma‐informed organizational climate and its impact on first responder burnout during COVID‐19. Public Adm Rev 2025;85(1):144-164. [doi: 10.1111/puar.13764]
58. Transforming psychological trauma: a knowledge and skills framework for the Scottish workforce. NES. 2017. URL: https://www.nes.scot.nhs.uk/media/rgxngvpv/nationaltraumatrainingframework-execsummary-web.pdf [accessed 2023-12-05]
59. Dana D. The Polyvagal Theory in Therapy: Engaging the Rhythm of Regulation (Norton Series on Interpersonal Neurobiology): WW Norton & Company; 2018.
60. Griffith B, Archbold H, Sáez Berruga I, et al. Frontline experiences of delivering remote mental health supports during the COVID-19 pandemic in Scotland: innovations, insights and lessons learned from mental health workers. Psychol Health Med 2023 Apr;28(4):964-979. [doi: 10.1080/13548506.2022.2148698] [Medline: 36408950]
61. Porges SW. Polyvagal theory: a science of safety. Front Integr Neurosci 2022;16:871227. [doi: 10.3389/fnint.2022.871227]
62. Martin A, Naunton M, Kosari S, Peterson G, Thomas J, Christenson JK. Treatment guidelines for PTSD: a systematic review. JCM 2021;10(18):4175. [doi: 10.3390/jcm10184175]
63. Mathews RL, Thomas N. Workplace hazards in trauma-informed practice: ensuring helping provider health and well-being. In: Professional's Guide to Trauma-Informed Decision Making: Springer; 2024:309-329. [doi: 10.1007/978-3-031-54626-6_17]
64. Antony J, Brar R, Khan PA, et al. Interventions for the prevention and management of occupational stress injury in first responders: a rapid overview of reviews. Syst Rev 2020 May 31;9(1):121. [doi: 10.1186/s13643-020-01367-w] [Medline: 32475353]
65. Joyce S, Shand F, Lal TJ, Mott B, Bryant RA, Harvey SB. Resilience@Work mindfulness program: results from a cluster randomized controlled trial with first responders. J Med Internet Res 2019;21(2):e12894. [doi: 10.2196/12894]
66. Kick KA, McNitt M. Trauma, spirituality, and mindfulness: finding hope. SWC 2016;43(1):40.
67. Porges SW, Dana D. Clinical Applications of the Polyvagal Theory: The Emergence of Polyvagal-Informed Therapies (Norton Series on Interpersonal Neurobiology): WW Norton & Company; 2018.
68. Ehlers A. Trauma‐focused cognitive behavior therapy for posttraumatic stress disorder and acute stress disorder. In: CBT for Anxiety Disorders: A Practitioner Book: John Wiley & Sons, Ltd; 2013:161-189. [doi: 10.1002/9781118330043]
69. Wang J, Luo Z, Liao X, et al. The levels and related factors of posttraumatic growth among nurses: a systematic review and meta-analysis. J Psychiatr Ment Health Nurs 2024 Apr;31(2):241-254. [doi: 10.1111/jpm.12975] [Medline: 37724379]
70. Rossouw JG, Herlofson J, Geldenhuys DJ, Erieau CL. High adversity resilience training (HART): development for emergency responders and defence. Journal of Applied Neurosciences 2024;3(1):8. [doi: 10.4102/jan.v3i1.8]
71. Saleem M, Kühne L, De Santis KK, Christianson L, Brand T, Busse H. Understanding engagement strategies in digital interventions for mental health promotion: scoping review. JMIR Ment Health 2021 Dec 20;8(12):e30000. [doi: 10.2196/30000] [Medline: 34931995]
72. Egan KJ, Hodgson W, Dunlop MD, Imperatore G, Kirk A, Maguire R. A novel mobile app ("CareFit") to support informal caregivers to undertake regular physical activity from home during and beyond COVID-19 restrictions: co-design and prototype development study. JMIR Form Res 2021 Oct 1;5(10):e27358. [doi: 10.2196/27358] [Medline: 34406969]
73. Haile C, Kirk A, Cogan N, Janssen X, Gibson AM, MacDonald B. Pilot testing of a nudge-based digital intervention (Welbot) to improve sedentary behaviour and wellbeing in the workplace. IJERPH 2020;17(16):5763. [doi: 10.3390/ijerph17165763]
74. Doria N, Condran B, Boulos L, Curtis Maillet DG, Dowling L, Levy A. Sharpening the focus: differentiating between focus groups for patient engagement vs. qualitative research. Res Involv Engagem 2018 Dec;4(1):19. [doi: 10.1186/s40900-018-0102-6]
75. Gottschalk S, Aziz MS, Yigitbas E, Engels G. Design principles for a crowd-based prototype validation platform. In: Wang X, Martini A, Nguyen-Duc A, Stray V, editors. Lecture Notes in Business Information Processing: Springer International Publishing; 2021:205-220. [doi: 10.1007/978-3-030-91983-2_16]
76. Skivington K, Matthews L, Simpson SA, et al. Framework for the development and evaluation of complex interventions: gap analysis, workshop and consultation-informed update. Health Technol Assess 2021 Sep;25(57):1-132. [doi: 10.3310/hta25570] [Medline: 34590577]

XSL•FO
RenderX

77. Lang AJ, Stein MB. An abbreviated PTSD checklist for use as a screening instrument in primary care. Behav Res Ther 2005 May;43(5):585-594. [doi: 10.1016/j.brat.2004.04.005]
78. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes 2007 Dec;5(1):63. [doi: 10.1186/1477-7525-5-63]
79. Malach-Pines A. The Burnout Measure, Short Version. Int J Stress Manag 2005;12(1):78-88. [doi: 10.1037/1072-5245.12.1.78]
80. Lovibond PF, Lovibond SH. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav Res Ther 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-u] [Medline: 7726811]
81. Morton L, Cogan N, Kolacz J, et al. A new measure of feeling safe: developing psychometric properties of the Neuroception of Psychological Safety Scale (NPSS). Psychol Trauma 2024 May;16(4):701-708. [doi: 10.1037/tra0001313] [Medline: 35849369]
82. Schaufeli WB, Salanova M, González-romá V, Bakker AB. The measurement of engagement and burnout: a two sample confirmatory factor analytic approach. J Happiness Stud 2002 Mar;3(1):71-92. [doi: 10.1023/A:1015630930326]
83. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27. [doi: 10.2196/mhealth.3422] [Medline: 25760773]
84. Cann A, Calhoun LG, Tedeschi RG, et al. A short form of the Posttraumatic Growth Inventory. Anxiety, Stress, & Coping 2010 Mar;23(2):127-137. [doi: 10.1080/10615800903094273]
85. Craig CL, Marshall AL, Sj??str??m M, et al. International Physical Activity Questionnaire: 12-country reliability and validity. Medicine & Science in Sports & Exercise 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB]
86. Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A Report from the WHOQOL Group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00]
87. Cogan N, Morton L, Campbell J, et al. Neuroception of psychological safety scale (NPSS): validation with a UK based adult community sample. Eur J Psychotraumatol 2025 Dec 31;16(1):2490329. [doi: 10.1080/20008066.2025.2490329]
88. Haver A, Akerjordet K, Caputi P, Furunes T, Magee C. Measuring mental well-being: a validation of the Short Warwick-Edinburgh Mental Well-Being Scale in Norwegian and Swedish. Scand J Public Health 2015 Nov;43(7):721-727. [doi: 10.1177/1403494815588862] [Medline: 26041133]
89. McCutchan PK, Freed MC, Low EC, Belsher BE, Engel CC. Rescaling the post-traumatic stress disorder checklist for use in primary care. Mil Med 2016 Sep;181(9):1002-1006. [doi: 10.7205/MILMED-D-15-00444]
90. Gan DZQ, McGillivray L, Larsen ME, Christensen H, Torok M. Technology-supported strategies for promoting user engagement with digital mental health interventions: a systematic review. Digit Health 2022;8:20552076221098268. [doi: 10.1177/20552076221098268] [Medline: 35677785]
91. Greer B, Robotham D, Simblett S, Curtis H, Griffiths H, Wykes T. Digital exclusion among mental health service users: qualitative investigation. J Med Internet Res 2019;21(1):e11696. [doi: 10.2196/11696]
92. Mohamed S, Png MT, Isaac W. Decolonial AI: decolonial theory as sociotechnical foresight in artificial intelligence. Philos Technol 2020 Dec;33(4):659-684. [doi: 10.1007/s13347-020-00405-8]
93. Cogan NA, MacIntyre G, Stewart A, et al. "The biggest barrier is to inclusion itself": the experience of citizenship for adults with mental health problems. J Ment Health 2021 May 4;30(3):358-365. [doi: 10.1080/09638237.2020.1803491]
94. Nicola C. Psychotraumatology and artificial intelligence: a public health approach. PHOA 2024;8(1). [doi: 10.23880/phoa-16000284]
95. Morgan P, Cogan NA. Using artificial intelligence to address mental health inequalities: co-creating machine learning algorithms with key stakeholders and citizen engagement. JPMH 2025 Jun 19;24(2):150-165. [doi: 10.1108/JPMH-07-2024-0095]
96. Williamson SM, Prybutok V. Balancing privacy and progress: a review of privacy challenges, systemic oversight, and patient perceptions in AI-driven healthcare. Appl Sci (Basel) 2024;14(2):675. [doi: 10.3390/app14020675]
97. Peters D, Vold K, Robinson D, Calvo RA. Responsible AI—two frameworks for ethical design practice. IEEE Trans Technol Soc 2020;1(1):34-47. [doi: 10.1109/TTS.2020.2974991]
98. Tahvanainen L, Tetri B, Ahonen O. Exploring and extending human-centered design to develop AI-enabled wellbeing technology in healthcare. In: Särestöniemi M, Keikhosrokiani P, Singh D, Harjula E, Tiulpin A, Jansson M, et al, editors. Digital Health and Wireless Solutions: Springer Nature Switzerland; 2024:288-306. [doi: 10.1007/978-3-031-59091-7_19]
99. Johnstone H, Cogan N. 'He messaged me the other night and said you are my saviour': an interpretative phenomenological analysis of intimate partners' roles in supporting veterans with mental health difficulties. Journal of Military, Veteran and Family Health 2021 May 1;7(2):61-70. [doi: 10.3138/jmvfh-2019-0055]
100. Cogan N. Persistent traumatic stress exposure: rethinking PTSD for frontline workers. Healthcare (Basel) 2026 Jan 20;14(2):255. [doi: 10.3390/healthcare14020255] [Medline: 41595391]

XSL•FO
RenderX

Abbreviations

AI: artificial intelligence
DI: digital intervention
FR: first responder

Edited by J Sarvestan; submitted 06.Feb.2025; peer-reviewed by J Maciejewski, S Sagar; revised version received 09.Feb.2026; accepted 11.Feb.2026; published 18.Mar.2026.

Please cite as:
Cogan N, Kirk A, Graf C
The Development and Protocol for Testing a Co-Created Digital Intervention (Sentinel) to Improve Mental Well-Being and Help Manage and Prevent Trauma in First Responders
JMIR Res Protoc 2026;15:e72250
URL: https://www.researchprotocols.org/2026/1/e72250
doi:10.2196/72250

© Nicola Cogan, Alison Kirk, Christoph Graf. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 18.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Utilization and Feasibility of a Wearable Device in Patients With Sedative Effects of Drugs: Protocol for a Prospective Observational Study for the Advanced Respiratory Monitoring Events in Drug Toxicity (ARM-ED) Study

Lisa Christine Dunlop[1], MBChB, BSc(hons); Bruce Henderson[2], MBChB; Osian Meredith[2], BSc, MPhil, PhD; Chris Trueman[2], BSc; Christopher Carlin[3], MBChB, PhD; Robert Docking[4], MBChB; David J Lowe[1,5], MBChB, BMSc, MSc, MD

[1]Emergency Department, Queen Elizabeth University Hospital, 1345 Govan Road, Glasgow, United Kingdom

[2]PneumoWave Ltd, Motherwell, United Kingdom

[3]Department of Respiratory Medicine, Queen Elizabeth University Hospital, Glasgow, United Kingdom

[4]Intensive Care Department, Queen Elizabeth University Hospital, Glasgow, United Kingdom

[5]School of Health and Wellbeing, University of Glasgow, Glasgow, United Kingdom

Corresponding Author:
Lisa Christine Dunlop, MBChB, BSc(hons)
Emergency Department, Queen Elizabeth University Hospital, 1345 Govan Road, Glasgow, United Kingdom

Abstract

Background: Drug-related deaths worldwide are most commonly attributed to opioids. Opioids and other sedative drugs can cause respiratory depression and airway compromise, leading to hypoxia and death. Device technology and artificial intelligence used to detect drug overdose has the potential to improve outcomes. PneumoWave Ltd has developed a small chest-worn respiratory monitoring device to detect concerning breathing patterns and alert an emergency response.

Objective: The aim of this study is to investigate the feasibility of using the PneumoWave device in hospital patients at risk of respiratory depression.

Methods: This 18-month prospective observational study was performed at the Queen Elizabeth University Hospital, Glasgow. The study investigates the use of the device on 3 groups of patients at risk of respiratory depression due to drugs. This includes patients attending the emergency department (ED) due to sedative drug overdose, patients receiving procedural sedation and analgesia in the ED, and patients receiving general anesthesia in theaters. Consenting participants will have the PneumoWave sensor paired with end-tidal CO_2 monitoring and regular recordings of vital signs. Usability will be tested by administering a questionnaire to the patient, the clinician, and the nurse. The primary end point is to determine the feasibility of gathering respiratory data from a wearable respiratory monitoring device in the ED. Statistical analysis includes comparison of biosensor data against reference physiology time course data.

Results: Trial recruitment was completed on December 8, 2023. Twenty-five patients were enrolled in group 1 (acute toxicity); 39 patients in group 2 (procedural sedation), including 1 patient who was subsequently withdrawn; and 14 patients in group 3 (acute toxicity), with 2 patients withdrawn. Publication of the study results is anticipated by December 2025.

Conclusions: This study will explore the feasibility of the PneumoWave device in a variety of clinical contexts in which risk of pharmacologically induced respiratory depression is present. This will provide valuable insight into the use of device technology in individuals at risk of illicit drug–related harm within the relative safety of a hospital setting. A limitation of the study procedure is exclusion of patients with intoxication after sedative drug overdose who lack the capacity to provide informed consent. This study has been designed to acquire feasibility data to demonstrate the potential for continuous respiratory monitoring to improve outcomes for patients who are at risk of drug-induced respiratory depression, inform product development, and inform the design of future pivotal clinical investigations.

Trial Registration: ClinicalTrials.gov NCT05358132; https://clinicaltrials.gov/study/NCT05358132

International Registered Report Identifier (IRRID): DERR1-10.2196/66961

(JMIR Res Protoc 2026;15:e66961) doi:10.2196/66961

XSL•FO
RenderX

KEYWORDS

wearable device; respiratory biosensor; opioid-induced respiratory depression; illicit drug toxicity; drug toxicity; benzodiazepine toxicity; emergency department

Introduction

Background

Over a quarter of a billion people worldwide used illicit drugs at least once during the year 2021, 60 million of which used opioids. Overdose is a significant cause of preventable morbidity and mortality globally, with approximately 128,000 deaths directly due to drug use in 2019 [1]. Within Europe, Scotland reports the highest drug-related death rate at 19.8 per 100,000 population. This study is sited in a tertiary urban hospital in Glasgow, Scotland. National Health Service (NHS) Greater Glasgow and Clyde Health Board has the highest drug-related death rate in Scotland (34 per 100,000 in 2022), and hence, this research is at the epicenter of the drug death crisis [2].

Opioids are the drug group most commonly attributed to drug-related death in the world [3]. In 2021, opioids were implicated in 75% and 74% of drug-related deaths in the United States and Europe, respectively [4,5]. Similarly, in Scotland, postmortem toxicology reports found that the most commonly implicated drugs were opioids (n=1051, 82.5%) and benzodiazepines (n=1051, 57.2%) [2].

Opioids are μ-receptor agonists, and their overdose can lead to reduced consciousness, irregular or slow breathing patterns, and potentially life-threatening or fatal respiratory depression. Death following the use of opioids is due to loss of respiratory reflexes, apnea, and respiratory failure [6,7]. There is an increased risk of death in opiate-naïve patients when opioids are combined with other central nervous system (CNS) depressant drugs such as alcohol, benzodiazepines, and other co-prescribed medications such as gabapentinoids [8,9]. Overdose of benzodiazepines, like opioids, can lead to reduced consciousness and loss of airway reflexes, leading to potential death. This is the case especially when taken in the context of user naivety, high doses used, potent preparations, or co-ingestion [10,11].

Access to emergency services and the use of the opioid antagonist, naloxone, given by any individual can reduce the risk following drug overdose [12]. A user who experiences an overdose must, however, be able to access help and therefore is at higher risk of death when using drugs alone [13].

The use of technology may help provide a solution to reduce the risk of fatal overdose in drug users. There have been several studies investigating the use of device technology in this setting, and these devices may be attached to or positioned somewhere away from the individual. Mobile phone apps are available for prevention and detection of harm. These often rely on the user (who may be unconscious) being able to alert help via the device [14]. A mobile app described by Nandakumar et al [15] found that a mobile phone which was adapted to use sonar to detect respiratory effort and apnea reported higher performance in the detection of postopioid self-administration central apneic events (10 s or longer). However, the device was required to be held by a second person at a 90° angle and at a distance of just 1 m from the individual for its successful use. Other near-patient devices include in-room movement sensors and help buttons; the latter again require the user to be capable of pressing the button [16,17].

Several wearable sensors in development have been described, including those that deploy naloxone [18-20]. The described devices have major barriers to their potential use. Mesa et al [19] describe a sensor that requires implantation of a drug-delivery system, which would likely not be acceptable to many individuals, with a potential increased risk of infection. Chan et al [20] describe a bulky device, which may reduce compliance of use, and they found it was unreliable in deploying the naloxone. Roth et al [21] describe a commercially available biosensor to detect movement. They showed success in collecting accelerometer data; however, they detected no respiratory depression despite 385 drug administrations recorded (>90% involved fentanyl), 27 of which the user described themselves as having “too much”.

There are some potential barriers to the use of a device such as battery life, effective functioning, size, weight, and aesthetics [22,23]. Despite this, potential users of devices are likely to be accepting the device if they understand its ability to reduce harm [24]. As described, the near-patient device described by Nandakumar et al was limited by its requirement to be perpendicular and within 1 m proximity to the patient to achieve stated results [15]. There is clearly a need for innovative technology that can detect individuals at risk of overdose, enabling help to reach them, while also being acceptable to wear.

Wearable Technology

PneumoWave Ltd has developed a small, wearable biosensor that incorporates a microelectromechanical system inertial motion unit capable of continuously monitoring central body motion over time [25]. This physical monitor is linked wirelessly (Bluetooth) to central monitoring systems via either a data gateway that plugs into a nearby electrical socket or a smartphone. The company’s technology is separated into two product lines: DC (data collection) and the Alert (event recognition).

PneumoWave DC is a UKCA (UK Conformity Assessed) Class 1 medical device with the following intended use: PneumoWave DC is a monitoring device intended to capture and store chest motion data continuously for retrospective analysis by a health care professional to document physical movement associated with applications in physiological monitoring. The device is not intended for diagnosis of respiratory conditions. It is indicated for patients deemed to be suitable for chest motion monitoring in health care facilities by a physician (Figure 1).

XSL•FO
RenderX

Figure 1. PneumoWave data collection system. A clinical data collection and analysis platform showing the biosensor, mobile app, or hub (right image), and clinician web portal with visualization of motion data. This system is designed for researchers and clinicians, with a focus on comprehensive data collection and retrospective analysis.



Prototype biosensor units (identical mounting methods and similar dimensions) were subject to two previous studies resulting in 16,476 hours of wear in 138 participants, indicating a high tolerance for the design (NCT04668313 and ISRCTN12060022). The DCM system collects raw X, Y, and Z axis accelerometer movement data with no data processing applied, which transmits via the App or Hub to a cloud environment. By accessing the cloud dashboard, users are able to view and download the accelerometer data.

The ALERT system for real-time collection of data plus processing is in development to enable detection and the potential of immediate intervention through algorithms currently under development and is a key driver for this study (see Figure 2).

Figure 2. PneumoWave ALERT platform technology overview. The biosensor will detect concerning breathing patterns and will send an alert for urgent help for the individual following overdose.

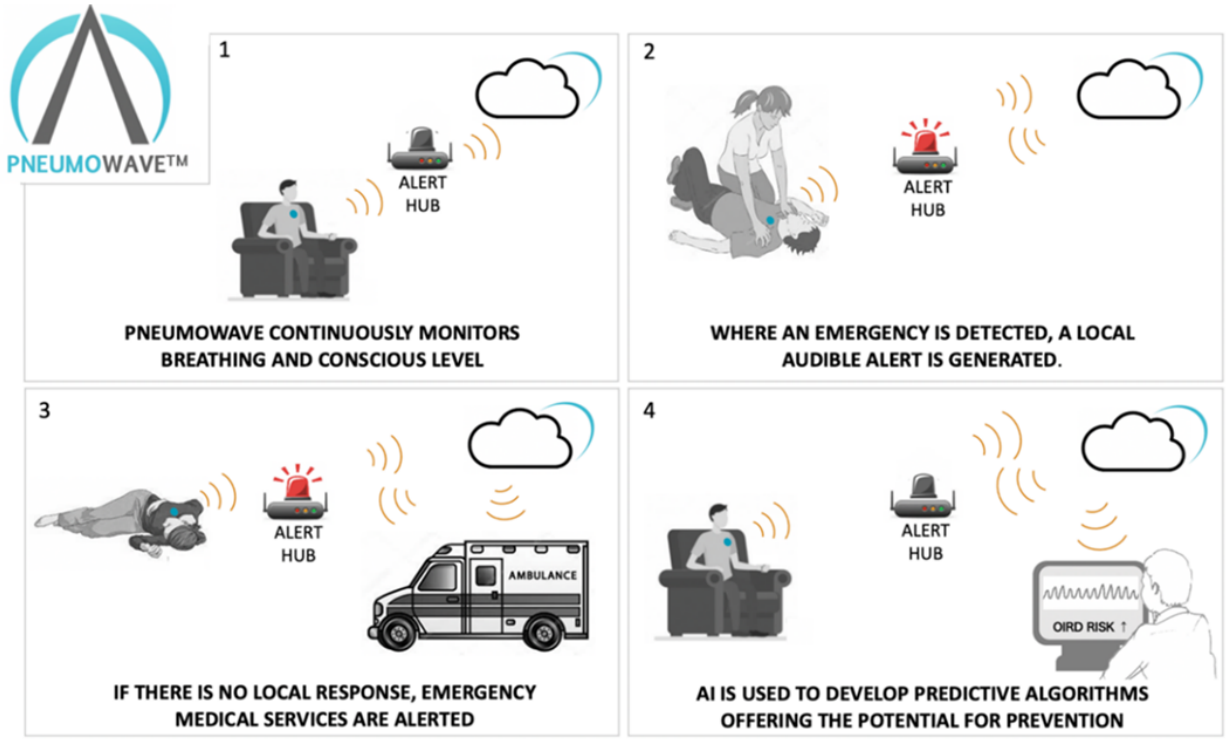

This research investigates the feasibility of the use of the PneumoWave Ltd wearable biosensor in individuals at risk of drug-induced respiratory depression. The safety of a hospital setting has been utilized to understand the feasibility of use.

XSL•FO
RenderX

Risk of respiratory depression occurs in a variety of situations within hospitals. As described, an overdose of opioids or other sedating/respiratory depressant drugs often requires medical care and attending the emergency department (ED) for treatment [26]. This group of patients is included in this trial and provides a pragmatic indication of feasibility of use in patients at risk of respiratory depression. Opioids and sedatives are also administered in hospital for medical purposes in a setting where the patient will need respiratory monitoring. One such example is procedural sedation and analgesia (PSA). This is a technique in which sedation is given with an aim to allow a painful procedure to be performed but not deep enough to cause apnea, though there is a risk, and this occasionally does occur [27]. A further example is that of induction for a general anesthetic (GA) where apnea is either certain (with muscle relaxant) or likely (during GA induction). These two groups of patients will also be studied to understand the use of the device in settings of increased risk or definite apnea due to drug administration. End-tidal CO_2 ($ETCO_2$) is used in hospitals in these settings to monitor respiratory rate by capnography and as such is the gold standard comparator in this study [28].

Study Hypothesis

The PneumoWave device will be well tolerated by patients and will successfully capture respiratory effort and patient movement accelerometer data.

Objectives

The primary objective of this study was to assess the feasibility of collecting respiratory waveform data using the PneumoWave device in ED patients from groups 1 - 3 (see Study Population).

Secondary objectives of the study were as follows: (1) Collate respiratory data from the PneumoWave device and compare to $ETCO_2$ in patients from all 3 groups. Respiratory data will observe normal respiratory effort, reduced respiratory effort due to sedating drugs or anesthetic agents, and response to treatments and interventions required in patients with reduced respiratory effort. (2) Collate waveform and motion artifact data from the PneumoWave device in patients from all 3 groups to observe the relationship with the individual's conscious level. (3) Assess the usability of PneumoWave in the opinion of the clinician, nurse, and patient. (4) Understand the feasibility of gathering respiratory data from a wearable respiratory monitoring device in the ED.

Methods

Study Design

This is a single-center observational cohort study conducted using passive non-invasive data collection.

Setting

This study was undertaken at the Queen Elizabeth University Hospital, Glasgow, United Kingdom, over an 18-month period from June 9, 2022, to December 8, 2023.

Study Population

The overall study population consists of patients at risk of respiratory depression due to exposure to sedating drugs. The study observes three distinct patient groups:

- Group 1 (acute toxicity group): Patients who present to the ED with actual or potential CNS or respiratory depression secondary to toxicological cause
- Group 2 (PSA group): Patients who undergo PSA with respiratory and CNS depressant drugs in the ED
- Group 3 (GA group): Patients who undergo GA in the anesthetic room in theater

Inclusion Criteria

The inclusion criteria are presented in Table 1.

Table . List of inclusion criteria for the Advanced Respiratory Monitoring in the Emergency Department (ARM-ED) study.

Group	Inclusion criteria
All groups	Age >16 years; able and willing (in the investigators' opinion) to comply with all study requirements; able to speak and read English
Group 1 (acute toxicity)	Willing and able to give informed consent (or via Welfare Attorney, Guardian, or Nearest Relative); presenting to ED[a] due to presumed overdose of a drug with potential for respiratory depression (intentional, accidental, recreational, or therapeutic excess); having at least one of the following: Glasgow Coma Score (GCS) <15, respiratory depression, or risk of deterioration in GCS or respiration
Group 2 (PSA[b])	Willing and able to give informed consent; undergoing procedural sedation and anesthesia in ED
Group 3 (GA[c])	Willing and able to give informed consent; undergoing elective general anesthetic of any type in the anesthetic room

[a]ED: emergency department.

[b]PSA: procedural sedation and analgesia.

[c]GA: general anesthetic.

Exclusion Criteria

The exclusion criteria are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. List of exclusion criteria for the Advanced Respiratory Monitoring in the Emergency Department (ARM-ED) study.

- Unable to provide consent and no next of kin to provide consent on participant's behalf
- Impaired consciousness or respiratory suppression most likely due to cause other than acute drug use
- Condition primarily related to alcohol use and no evidence of acute drug use
- Condition due to withdrawal of drugs or alcohol
- Treating clinician deems patient inappropriate to be included in the study
- Pregnant female
- Patient has implantable device (eg, pacemaker)

Withdrawal of Participants

Participants will be able to withdraw from the study at any time, without impact on their clinical management.

Sample Size and Rationale

This is a feasibility study to generate data for future prospective interventional studies; therefore, no formal sample size will be generated. The research group planned a target of 50 patients recruited to both groups 1 and 2 as the number of minutes of data capture would successfully answer the primary outcome question of feasibility. Group 3 was included following trial commencement. Following evidence of the volume and type of data being produced from groups 1 and 2, a target of 12 was used based on the understanding of the number of minutes of data gathered. We aim to recruit a total of 112 patients to the study, which will provide adequate data to establish feasibility of use for larger studies. The sample size breakdown is as follows: group 1: 50 patients; group 2: 50 patients; and group 3: 12 patients.

Primary End Points

The primary end point is understanding the feasibility of gathering respiratory data from a wearable respiratory monitoring device in the ED.

Secondary End Points

The secondary end points are as follows:

- Add to existing respiratory data to inform machine learning and improve sensitivity and specificity of PneumoWave device
- Understand the usability of the device within acute care setting
- Understand the feasibility of gathering respiratory data from a wearable respiratory monitoring device in patients undergoing GA

Outcome Measures

The primary outcome measures will measure the length of time the device is in situ on the patient, the number of times the device is removed by the patient or others, and the ability of the device to collect data while in situ. The secondary outcome measures will observe waveform data from the PneumoWave device and compare those with normal care vital signs and continuous monitoring vital signs extracted from patient monitors. PneumoWave respiratory wave patterns and motion artifact will be compared to clinical events, including normal respiratory patterns, clinical deterioration, and interventions in ED or in anesthetic room in theater, and Glasgow Coma Score.

Stakeholders will be interviewed regarding the usability of the device. These stakeholders are the clinician and nurse caring for the participating patient in question, and the patient themselves, if able to.

Outcome measures and their associated data sources and data capture are shown in Table 2.

XSL•FO
RenderX

Table . Outcome measures and associated data source and data to be collected.

Outcome measure	Data source	Data to be collected
Primary outcome measures		
Length of time the device is in situ on patient	CASTOR EDC [29]	The total time device present on chest as documented by researcher (seconds)
The number of times the device is removed by the patient or others	CASTOR EDC	Absolute number of times the device was removed
The ability of the device to collect data while in situ	PneumoWave sensor	Total amount data recorded (seconds) as the proportion of time the device was on chest (seconds)
Secondary outcome measures		
PneumoWave device waveform data comparison to clinical events	PneumoWave device and CASTOR EDC	Visual representation of PneumoWave data and comparison to the following: • $ETCO_2$[a] waveforms • Documented respiratory rate • Interventions • GCS[b]
Usability of device	CASTOR EDC	• Data from results of interview reported in Likert scale from 1 to 5.

[a]$ETCO_2$: end-tidal carbon dioxide.

[b]GCS: Glasgow Coma Score.

Study Procedures

Potential Participant Identification

Group 1 (Acute Toxicity) and Group 2 (PSA) Patient Identification:

Patients from group 1 and group 2 will be recruited concurrently and opportunistically. Patients will be recruited during the research team by the researcher who will observe ED attendances remotely via electronic health records and identify potential participants who are either intoxicated (group 1) or may undergo procedural sedation (group 2). They will assess the patient for eligibility. ED staff will be given information on study recruitment and alert the research team to assess eligibility for inclusion.

Group 3 (GA) Patient Identification

Patients will be recruited during specified elective theater lists undergoing a GA. Patients will be approached by the anesthetist for willingness to participate, with the researcher assessing for eligibility, providing them with the patient information sheet, and gaining consent from the patient.

Consent

Written consent will be gained from the patient or Welfare Attorney/Welfare Guardian/Nearest Relative for group 1 and from the patient only for group 2 and group 3.

Source Data Collection and Patient Observation

Source data will be collated from multiple sources, including study participant interviews; usual care electronic health records; anesthetic charts; medical, ambulance, and nursing notes; drug charts; and observation charts. Continuous vital sign data will be downloaded from the standard ED or anesthetic monitors or study monitors, PneumoWave sensor data, and manual respiratory rate for study periods using $ETCO_2$ monitor.

During the study episode, data will be collected from the above sources for a period of time specific to each group. Data are all gathered by the researcher. Group 1 patients will be observed by the research team member for 1 hour or from commencement of data collection until they are discharged or moved from the ED, whichever is shorter. Group 2 participants will be observed for the duration of their procedural sedation, length of which may vary, and recovery period for up to 15 minutes. Data are gathered from group 3 participants prior to their GA while in the theater's anesthetic room, during induction of anesthesia and until the patient is moved to the theater.

Usability is investigated by interviewing stakeholders (patient, nurse, and clinician) following each respective research study episode. It is expected that due to the nature of the study, it may not be possible to gather these data for all patients recruited due to reduced conscious level. Responses to domains were on a scale of 1 (least positive) to 5 (most positive). The domains that were investigated were perceptions on device comfort, device intrusiveness, device safety, delay to treatment, acceptability of device in normal practice.

Review of Notes at 28 Days

The participant has a follow-up review of clinical status at a single time point of 28 days postrecruitment via electronic case records. This will identify any adverse outcome relating to their ED attendance or anesthetic, for example, need for critical care, disability, morbidity, or mortality as well as identify any relevant usual care drug screens performed at the time of assessment.

XSL•FO
RenderX

Waveform Analysis and Predictive Modeling Algorithms

Data from PneumoWave are analyzed in parallel with deidentified clinical and vital sign observational data to build artificial intelligence–driven predictive modeling algorithms.

Summary Study Procedure

Table 3 shows a summary of study procedures, including participant enrollment, interventions, and assessments (Checklist 1).

Table . Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) figure of enrollment, interventions, and assessments.

Study procedure	Screening/consent (first visit)	Research episode	28 days' follow-up electronically
Obtain informed consent where relevant	✓		
Review inclusion/exclusion criteria	✓		
Complete case report form	✓	✓	✓
Collect continuous data from standard care vital sign monitor		✓	
Collect waveform respiratory data		✓	
Review retrospective notes for adverse outcome/available toxicology screen			✓

Data Collection and Management

CASTOR (The Netherlands) [29] will provide a platform for data collection via an electronic case report form [29]. CASTOR EDC is compliant with 21 CFR Part 11, ICH E6 GCP, GDPR, and HIPAA. It is ISO27001 and ISO9001 certified.

Safety and Adverse Events

Adverse events and serious adverse events (SAEs) will be managed in line with Good Clinical Practice guidelines. SAEs meeting the above criteria must be reported to the Pharmacovigilance Office immediately (within 24 h) using the SAE form for a non-CTIMP. All SAEs will be reported to the ethics committee within 15 days of the chief investigator or delegate becoming aware of the event.

Planned Analysis

Primary Outcome Measure Analysis

Descriptive statistics will be used to analyze the primary outcome measures and assess the feasibility of studying the use of the PneumoWave device in the three patient groups. The analysis for each primary outcome measure is as shown below.

- Length of time the device is in situ on the patient: The time between the documented time the device is placed on the patient and time off the patient will be calculated (seconds). Any time the device was removed will be deducted from the total time. A final analysis of median, mean, and associated IQR and SD will be reported for each group.
- Number of times the device was removed by the patient or others: This will be reported as median, mean, and associated IQR and SD, respectively, for each group.
- Ability of the device to collect data while in situ: The amount of data gathered for each episode will be recorded and compared to the length of time the device is in situ on the patient's body. This will be reported as a proportion. The median, mean, and associated IQR and SD will be reported for each group.

Secondary Outcome Measure Analysis

The secondary outcome measures study the data that is captured by the PneumoWave device. In order to analyze this, the device data will be visually reviewed alongside the usual care vital signs and continuous monitoring signs. Respiratory rate calculated from the PneumoWave monitor will be compared to respiratory rate counted by a medical professional from the $ETCO_2$ monitor. Accuracy will be analyzed through sensitivity and specificity tests using $ETCO_2$ measurements as the gold standard.

Data points in which there has been a change in respiratory pattern, clinical deterioration, clinical intervention, and Glasgow Coma Score will be visually reviewed through time stamp looking at both usual care vital signs and PneumoWave study waveform data. Results will be presented both through written description and visually to highlight waveform changes if applicable.

Analysis of clinician, nurse, and patient opinions on the use of the device will be recorded as median and ranges.

Feasibility was defined as the device remaining attached to the chest with two or fewer removals during the monitoring period and achieving data capture for more than 75% of the total wear time.

The statistical analysis will be performed using MiniTab.

Ethical Considerations

Ethics Approval and Consent to Participate

Favorable ethical approval was obtained from Scotland A Research Ethics Committee (REC) prior to study commencement on January 19, 2022 (REC reference: 21/SS/0083, Integrated Research Application System project ID: 303922).

A substantial amendment was submitted on March 28, 2023, and gained favorable opinion on April 28, 2023, by Scotland A REC.

XSL•FO
RenderX

The trial was registered on March 30, 2022, on ClinicalTrials.gov (reference: NCT05358132). The protocol information presented here has current ethical approval and is from protocol version 2.0.

Any protocol modifications will be approved with the REC prior to initiation of the change. Consent was obtained according to the protocol from patients or a nominated person by the co-investigators of the trial or a suitably trained member of research staff.

Process for Informed Consent

The process for informed consent will differ between group 1: acute toxicity group and group 2: PSA group. Group 3: GA group will follow the same consenting process as Group 2.

Group 1: Acute Toxicity Group

Because the nature of the condition is being studied, this patient group will most likely be incapacitated and will not be able to provide informed consent. We anticipate that this may lead to difficulty in recruiting patients for this group.

In some cases, the potential study participant will have the capacity either briefly or indefinitely. During this period, the researcher will assess the capacity and will gain informed consent from the participant. The participant will be approached with a verbal explanation from the researcher and provided with a patient information sheet (PIS) to read. They will be given as much time as required to consider the information. When the participant is satisfied and any questions are answered, they will be asked if they will consent to the study. The participant will sign and date the most recent approved version of the informed consent form before any study-specific procedures are performed. The participant will be given a copy of the signed consent form as well as the PIS.

If the patient is unlikely to regain capacity, then as per the Adults with Incapacity (Scotland) Act 2000, consent must be gained from guardian or welfare attorney or nearest relative. The next of kin of the potential participant will be contacted, and verbal and written information regarding the study will be provided. They will be given as much time as required, and a researcher will be available to answer any questions. The next of kin will be required to provide written informed consent and as such must be present in the ED for this.

Group 2 (PSA Group) and Group 3 (GA Group)

These patient groups will only be included if they have the capacity to provide informed consent for this study. The participants will be approached with verbal explanation from the researcher and provided with a PIS to read. They will be given as much time as required to consider the information. When the participant is satisfied and any questions are answered, they will be asked if they will consent to the study. The participant will sign and date the most recent approved version of the informed consent form before any study-specific procedures are performed. The participant will be given a copy of the signed consent form as well as the PIS.

In groups 1, 2, and 3, the participant and their Welfare Attorney, Welfare Guardian, or Nearest Relative are free to withdraw from the study at any time for any reason without prejudice to future care and with no obligation to give the reason for withdrawal.

Consent for publication was achieved during initial participant consent.

Privacy and Confidentiality

Data are anonymized during this study. At study completion, the comprehensive anonymized study dataset will be submitted for inclusion in NHS Greater Glasgow and Clyde SafeHaven. This will allow subsequent deidentified review of all study outcomes, reanalysis of the dataset, and contribution of the dataset to future research work within NHS Scotland.

Compensation Details

No financial or material compensation was provided to participants for their participation in this study.

Updates to Protocol Since Registry

There were no significant protocol deviations during the study period. There were several amendments to the protocol following original ethical approval and trial registry. These amendments were reported and approved through the Integrated Research Application System. Table 4 outlines the amendments to the protocol.

XSL•FO
RenderX

Table . Study amendments following original ethics approval and trial registry.

Sponsor amendment date	Amendment number	Amendment details
March 17, 2022	NSA01	• The electronic case report form platform has been changed from Galen to CASTOR EDC [29]. This was changed due to access to the platform. CASTOR is compliant with 21 Code of Federal Regulations Part 11, International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use E6 GCP[a], General Data Protection Regulation, and Health Insurance Portability and Accountability Act. ISO (International Organization for Standardization) 27001 and ISO 9001 certified. • Following research group consensus, the name of the study was changed from TARS (Toxicology Advanced Respiratory Study) to ARM-ED (Advanced Respiratory Monitoring in the Emergency Department). • In the protocol, the wording was changed from "participants will be studied for the duration of their ED[b] attendance" to "participants will be studied for a period of time during their ED attendance." • The methods were updated to include the requirement of a clinician to count the respiratory rate retrospectively. They would do this by reviewing the end-tidal CO_2 monitor data. The reason for this is to improve the accuracy of respiratory rate counting in the case that the vital signs monitor may not accurately do this.
April 1, 2022	NSA02	• A minor amendment was made to exclude pregnant individuals in groups 1 and 2. The reason for the exclusion is that the physiology of breathing is different during pregnancy, and so the respiratory device may not be reliable in this instance. Separate studies are required to investigate the use of the device on pregnant females. In addition, in the emergency department, pregnant females are very rarely sedated for procedures.
June 7, 2022	NSA03	• The PneumoWave Biosensor was approved as UKCA (UK Conformity Assessed) Class 1. The documentation states that the sensor will be class 2. This was changed in the protocol.
July 21, 2022	NSA04	• Patient information sheets and corresponding consent forms were updated to reflect the correct version date.

XSL•FO
RenderX

Sponsor amendment date	Amendment number	Amendment details
March 28, 2023	SA 05	• A major amendment was approved to add a third arm to the study, named group 3: general anesthetic (GA) substudy. This group aimed to have a recruitment target of 12 patients to investigate the feasibility of the use of the PneumoWave device in patients who are undergoing an elective general anesthetic in the hospital operating theater anesthetic room. The reason that this additional arm was added was to capture respiratory movement waveform data in patients who would likely have a period of apnea during the anesthetic. • The protocol, standard operating procedure, and electronic case report form guidance were updated to exclude patients with an implantable device as per the PneumoWave device instructions for use.

[a]GCP: Good Clinical Practice.

[b]ED: emergency department.

Results

Trial recruitment was completed on December 8, 2023. A total of 25 patients were enrolled in group 1 (acute toxicity), 39 patients in group 2 (procedural sedation), including 1 patient who was subsequently withdrawn, and 14 patients in group 3 (acute toxicity), with 2 patients withdrawn. All withdrawn participants were excluded at an early stage, as planned sedation was not carried out.

Data analysis is currently in progress, and publication of the study results is anticipated by December 2025.

Discussion

Expected Principal Findings

The anticipated main finding of this feasibility study is that the PneumoWave wearable respiratory monitoring device will successfully gather respiratory waveform data and will be well tolerated by patients at risk of drug-related respiratory depression. Specifically, it is expected that the device will provide high-quality respiratory waveform data comparable to the gold standard of $ETCO_2$ monitoring within the relative safety of the hospital environment [27]. These findings would support the feasibility of using wearable respiratory monitoring to enhance clinical surveillance in populations at risk of respiratory compromise.

The study will allow understanding of the ability of the device to measure respiratory rate and detect changes to a patient's condition accurately. Apnea can be difficult to study, and in hospitals, it is often brief, as intervention is swift while in the care of medical professionals due to substantial risk of harm if left untreated. Therefore, insight into the feasibility of studying this will be invaluable.

Comparison to Prior Work

Wearable respiratory monitoring technologies have shown promise in both hospital and community settings, yet are often not acceptable or practical for the patient [15,18-20]. This study, therefore, extends the current literature by focusing on the feasibility of a novel wearable solution that offers continuous and noninvasive respiratory monitoring of the individual at risk of drug-induced respiratory depression.

Strengths and Limitations

A key strength of this feasibility study is its pragmatic design, enabling evaluation of the PneumoWave device in the ED or anesthetic room, where risk to the patient can be minimized. However, several limitations should be acknowledged. Recruitment challenges are anticipated due to legal restrictions under Scottish Adults with Incapacity laws, which preclude consent via a professional legal representative; this constraint may particularly impact enrollment in the acute toxicology cohort (Group 1). This study has, however, been designed to improve patient outcomes who are at risk of drug-induced respiratory depression and inform future development of clinical support and prediction capability. The study will not be powered to assess diagnostic accuracy or clinical outcomes, and conclusions will therefore be limited to feasibility, tolerability, and data acquisition parameters.

Future Directions

Findings from this feasibility study will inform the design of larger scale observational and randomized diagnostic accuracy trials in both hospital and supervised injecting settings. Future studies could expand to assess real-time clinical decision support, predictive analytics for respiratory compromise. Long-term goals include the development of predictive models integrating wearable-derived respiratory data with other physiological or behavioral indicators of overdose risk.

Dissemination

Upon completion of the study, findings will be disseminated through peer-review journal publication and presentation at conferences and engagement with both clinical and public health stakeholders to reduce harm caused by drug-induced respiratory depression.

XSL•FO
RenderX

Conclusion

This hospital-based respiratory monitoring study will provide a safe environment to understand the feasibility of use of a wearable device in detecting drug-induced respiratory depression and inform measures in reducing associated harm.

Acknowledgments

The authors would like to acknowledge that this work uses data provided by patients and collected by the National Health Service (NHS) as part of their care and support. The authors also acknowledge the work of the West of Scotland Safe Haven team at NHS Greater Glasgow and Clyde and the Robertson School for Biostatistics in supporting the extractions and linkage to deidentified NHS patient datasets. Finally, the authors acknowledge NHS Greater Glasgow and Clyde Clinical Research Facility and the West of Scotland Innovation Hub for their research support and the Queen Elizabeth University Hospital Emergency Department for working with the research team to allow this study to take place.

Funding

Funding for this study was provided by PneumoWave Ltd.

Data Availability

Data for this study will be available on reasonable request from the corresponding author. The data are not publicly available due to the sensitive nature of the participants.

Biosensor data and machine learning algorithms are not publicly available and are subject to PneumoWave Ltd Intellectual Property security laws.

Authors' Contributions

LCD conceived, designed, and coordinated the study and wrote the manuscript. DJL is the principal investigator for the study, facilitated conception of the study, and contributed to multiple versions of the manuscript. CC contributed to multiple versions of the protocol and this manuscript. BH, OM, and CT developed the device under the company PneumoWave and contributed to multiple versions of the protocol and this manuscript. RD contributed to multiple versions of the protocol and this manuscript. Data collection was conducted exclusively by researchers affiliated with National Health Service Greater Glasgow and Clyde. Data analyses were performed collaboratively by Lisa Dunlop and employees of PneumoWave. The decision to publish results was made prior to the study commencement by all members of the research team.

Conflicts of Interest

LCD has investigator time funding from the Advanced Respiratory Monitoring Events in Drug Toxicity (ARM-ED) study derived from PneumoWave Ltd funding. DJL has investigator time funding from the ARM-ED (Advanced Respiratory Monitoring in the Emergency Department) study derived from PneumoWave Ltd funding. The author also received investigator time funding from the Chief Scientist Office for prior device validation work. CC received investigator time funding from the Chief Scientist Office for prior device validation work. BH, OM, and CT are all employees of PneumoWave Ltd. RD received no funding from the study but receives honoraria from MSD for invited lectures and webinars.

Checklist 1
SPIRIT checklist.
[DOC File, 126 KB - resprot_v15i1e66961_app1.doc]

References

1. World drug report 2023: executive summary. : United Nations Office on Drugs and Crime (UNODC); 2023 URL: https://www.unodc.org/res/WDR-2023/WDR23_Exsum_fin_SP.pdf [accessed 2026-02-18]
2. Drug-related deaths in Scotland in 2022. : National Records of Scotland; 2023 URL: https://www.drugsandalcohol.ie/39424/1/drug-related-deaths-22-report.pdf [accessed 2026-02-18]
3. Opioid overdose. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/opioid-overdose [accessed 2026-02-18]
4. European drug report 2023: trends and developments. : European Monitoring Centre for Drugs and Drug Addiction (EMCDDA); 2023 URL: https://tls-suchtfragen.de/wp-content/uploads/2024/04/Europaeischer-Drogenbericht-2023_englisch.pdf [accessed 2026-02-18]
5. Drug overdose deaths. Centers for Disease Control and Prevention (CDC). URL: https://www.cdc.gov/nchs/nvss/drug-overdose-deaths.htm [accessed 2026-02-18]

XSL•FO
RenderX

6. Kosten TR, George TP. The neurobiology of opioid dependence: implications for treatment. Sci Pract Perspect 2002 Jul;1(1):13-20. [doi: 10.1151/spp021113] [Medline: 18567959]
7. Pathan H, Williams J. Basic opioid pharmacology: an update. Br J Pain 2012 Feb;6(1):11-16. [doi: 10.1177/2049463712438493] [Medline: 26516461]
8. Hiranita T, Ho NP, France CP. Ventilatory effects of fentanyl, heroin, and d-methamphetamine, alone and in mixtures, in male rats breathing normal air. J Pharmacol Exp Ther 2024 Jan 17;388(2):244-256. [doi: 10.1124/jpet.123.001653] [Medline: 37739803]
9. Pattinson KTS. Opioids and the control of respiration. Br J Anaesth 2008 Jun;100(6):747-758. [doi: 10.1093/bja/aen094] [Medline: 18456641]
10. Kang M, Galuska MA, Ghassemzadeh S. Benzodiazepine toxicity. In: StatPearls [Internet]: StatPearls Publishing; 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK482238/ [accessed 2026-02-18]
11. Corkery JM, Guirguis A, Chiappini S, Martinotti G, Schifano F. Alprazolam-related deaths in Scotland, 2004-2020. J Psychopharmacol 2022 Sep;36(9):1020-1035. [doi: 10.1177/02698811221104065] [Medline: 35912873]
12. van Lemmen M, van der Schrier R, Dahan A, van Velzen M, Sarton E, Niesters M. Pharmacology of viable mechanism agnostic respiratory stimulants for the reversal of drug-induced respiratory depression in humans. Expert Opin Drug Metab Toxicol 2023;19(10):671-679. [doi: 10.1080/17425255.2023.2262386] [Medline: 37795596]
13. Ogeil RP, Dwyer J, Bugeja L, Heilbronn C, Lubman DI, Lloyd B. Pharmaceutical opioid overdose deaths and the presence of witnesses. Int J Drug Policy 2018 May;55:8-13. [doi: 10.1016/j.drugpo.2017.12.020] [Medline: 29433040]
14. Aggarwal M, Borycki EM. Review of mobile apps for prevention and management of opioid-related harm. Stud Health Technol Inform 2019;257:1-8. [Medline: 30741164]
15. Nandakumar R, Gollakota S, Sunshine JE. Opioid overdose detection using smartphones. Sci Transl Med 2019 Jan 9;11(474):eaau8914. [doi: 10.1126/scitranslmed.aau8914] [Medline: 30626717]
16. Welwean RA, Krieg O, Casey G, et al. Evaluating the impact of Brave Technology Co-op's novel drug overdose detection and response devices in North America: a retrospective study. J Urban Health 2023 Oct;100(5):1043-1047. [doi: 10.1007/s11524-023-00779-y] [Medline: 37670172]
17. Buchheit BM, Crable EL, Lipson SK, Drainoni ML, Walley AY. "Opening the door to somebody who has a chance." - The experiences and perceptions of public safety personnel towards a public restroom overdose prevention alarm system. Int J Drug Policy 2021 Feb;88:103038. [doi: 10.1016/j.drugpo.2020.103038] [Medline: 33232885]
18. Dhowan B, Lim J, MacLean MD, et al. Simple minimally-invasive automatic antidote delivery device (A2D2) towards closed-loop reversal of opioid overdose. J Control Release 2019 Jul 28;306:130-137. [doi: 10.1016/j.jconrel.2019.05.041] [Medline: 31158402]
19. Mesa JC, MacLean MD, Ms M, et al. A wearable device towards automatic detection and treatment of opioid overdose. IEEE Trans Biomed Circuits Syst 2024 Apr;18(2):396-407. [doi: 10.1109/TBCAS.2023.3331272] [Medline: 37938943]
20. Chan J, Iyer V, Wang A, et al. Closed-loop wearable naloxone injector system. Sci Rep 2021 Nov 22;11(1):22663. [doi: 10.1038/s41598-021-01990-0] [Medline: 34811425]
21. Roth AM, Tran NK, Cocchiaro B, et al. Wearable biosensors have the potential to monitor physiological changes associated with opioid overdose among people who use drugs: a proof-of-concept study in a real-world setting. Drug Alcohol Depend 2021 Dec 1;229(Pt A):109138. [doi: 10.1016/j.drugalcdep.2021.109138] [Medline: 34781097]
22. Tas B, Walker H, Lawn W, et al. What impacts the acceptability of wearable devices that detect opioid overdose in people who use opioids? A qualitative study. Drug Alcohol Rev 2024 Jan;43(1):213-225. [doi: 10.1111/dar.13737] [Medline: 37596977]
23. Chapman BP, Lucey E, Boyer EW, Babu KM, Smelson D, Carreiro S. Perceptions on wearable sensor-based interventions for monitoring of opioid therapy: a qualitative study. Front Digit Health 2022;4:969642. [doi: 10.3389/fdgth.2022.969642] [Medline: 36339518]
24. Kanter K, Gallagher R, Eweje F, et al. Willingness to use a wearable device capable of detecting and reversing overdose among people who use opioids in Philadelphia. Harm Reduct J 2021 Jul 23;18(1):75. [doi: 10.1186/s12954-021-00522-3] [Medline: 34301246]
25. PneumoWave. URL: https://pneumowave.com/ [accessed 2026-02-18]
26. Lowrie R, McPherson A, Mair FS, et al. Baseline characteristics of people experiencing homelessness with a recent drug overdose in the PHOENIx pilot randomised controlled trial. Harm Reduct J 2023 Apr 4;20(1):46. [doi: 10.1186/s12954-023-00771-4] [Medline: 37016418]
27. Aminiahidashti H, Shafiee S, Zamani Kiasari A, Sazgar M. Applications of end-tidal carbon dioxide ($ETCO_2$) monitoring in emergency department; a narrative review. Emerg (Tehran) 2018;6(1):e5. [Medline: 29503830]
28. ECC Committee, Subcommittees and Task Forces of the American Heart Association. 2005 American Heart Association guidelines for cardiopulmonary resuscitation and emergency cardiovascular care. Circulation 2005 Dec 13;112(24 Suppl):IV1-IV203. [doi: 10.1161/CIRCULATIONAHA.105.166550] [Medline: 16314375]
29. Castor. URL: https://www.castoredc.com/ [accessed 2026-02-18]

XSL•FO
RenderX

Abbreviations

ARM-ED: Advanced Respiratory Monitoring in the Emergency Department
CNS: central nervous system
ED: emergency department
ETCO2: end-tidal CO2
GA: general anesthetic
NHS: National Health Service
PIS: patient information sheet
PSA: procedural sedation and analgesia
REC: Research Ethics Committee
SAE: serious adverse event

Edited by A Schwartz; submitted 30.Sep.2024; peer-reviewed by A Puxty; revised version received 03.Nov.2025; accepted 06.Nov.2025; published 19.Mar.2026.

Please cite as:
Dunlop LC, Henderson B, Meredith O, Trueman C, Carlin C, Docking R, Lowe DJ
Utilization and Feasibility of a Wearable Device in Patients With Sedative Effects of Drugs: Protocol for a Prospective Observational Study for the Advanced Respiratory Monitoring Events in Drug Toxicity (ARM-ED) Study
JMIR Res Protoc 2026;15:e66961
URL: https://www.researchprotocols.org/2026/1/e66961
doi:10.2196/66961

© Lisa Christine Dunlop, Bruce Henderson, Osian Meredith, Chris Trueman, Christopher Carlin, Robert Docking, David J Lowe. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Smartphone-Based Digital Phenotyping to Predict Relapse in Serious Mental Disorders Among Slum Residents in Dhaka, Bangladesh: Protocol for a Machine Learning Study

Nadia Alam[1], MSc; Chayon Kumar Das[2], MPhil; Neelabja Roy[3], MS; Domenico Giacco[2,4], MBBS, PhD; Swaran P Singh[1,4], MBBS, MD; Sagar Jilka[1], PhD

[1]Warwick Medical School, University of Warwick, B 020, Gibbet Hill Road, Coventry, United Kingdom

[2]Department of Clinical Psychology, University of Dhaka, Dhaka, Bangladesh

[3]Department of Psychiatry, National Institute of Mental Health and Neurosciences, Bangalore, India

[4]Coventry and Warwickshire Partnership NHS Trust, Coventry, United Kingdom

Corresponding Author:
Nadia Alam, MSc
Warwick Medical School, University of Warwick, B 020, Gibbet Hill Road, Coventry, United Kingdom

Abstract

Background: Serious mental illnesses (SMIs) are associated with high relapse rates and limited access to continuous care, particularly in low-resource settings such as urban slums. Traditional clinical monitoring is constrained by accessibility and scalability challenges. Digital phenotyping, through passive smartphone data, offers a novel approach to predict relapse by capturing real-world behavioral changes.

Objective: This study aims to evaluate the feasibility and predictive value of smartphone-based digital phenotyping for detecting relapse in individuals with SMIs living in the Korail slum of Dhaka, Bangladesh.

Methods: This prospective 6-month cohort study will recruit 430 participants diagnosed with SMIs who own Android (Google LLC) smartphones. Passive data (eg, screen time, mobility, and call or text frequency) will be continuously collected using a custom-built app (DataDoc). Monthly active data, including symptom and functioning assessments, will be collected via self-report and clinical engagement. Machine learning models will integrate these data to detect early warning signs and predict relapse trajectories.

Results: This study was funded by the NIHR (National Institute for Health and Care Research; award number NIHR200846) in October 2022. Data collection commenced in August 2025 and is ongoing. A total of 14 participants have been recruited, as of January 2026. Preliminary data analysis is ongoing, with expected results to be published in fall 2026.

Conclusions: This study is one of the first to apply smartphone-based digital phenotyping and machine learning for relapse prediction in low- and middle-income countries' slum settings. The findings will inform scalable, low-cost digital interventions to address the mental health treatment gap in underresourced communities.

International Registered Report Identifier (IRRID): PRR1-10.2196/79826

(JMIR Res Protoc 2026;15:e79826) doi:10.2196/79826

KEYWORDS

mental health; digital phenotyping; passive sensing; low-income; urban settings; mobile phone

Introduction

Background

Poor outcomes for patients with serious mental illnesses (SMI) are associated with various factors involving patients, health care providers, and the health care system, in addition to the delivery of inadequate medical services [1-3]. The current clinical ways used to track the progression of SMI may be insufficient or inadequate. Traditionally, health care professionals rely on in-person consultations for assessment and diagnosis [4,5]. However, these clinic-based services come with a variety of logistical difficulties [4], especially for people from impoverished communities [6]. To monitor patients, they must frequently visit a clinical facility, often within restricted operating hours, which can be particularly challenging for individuals dealing with SMI [4]. Additionally, these approaches are demanding in terms of resources because they necessitate one-on-one interactions with a trained health care provider, making widespread adoption difficult. The inherent issues related to accessibility and scalability in these methods create significant obstacles to patient care [7,8].

XSL•FO
RenderX

Remotely collected digital markers may help overcome key challenges in monitoring SMI, such as the reliance on traveling for monitoring and in-person assessments [9,10]. Continuous monitoring of behavioral and physiological data using personal devices has the potential to aid in the identification of personalized indicators for the initiation of illness [4], subsequently enabling the development of tailored treatment strategies [9,11,12]. Smartphones can provide a solution.

Digital Phenotyping (DP)

Smartphones can be used to collect data to understand mental health status. This technique, called digital phenotyping (DP) [13], is already in use in some behavioral problems and mental disorders, such as posttraumatic stress disorder (PTSD) [14], and SMI such as schizophrenia [15-17] and mood disorders [18-20] to predict relapse, symptom exacerbation, or mood fluctuations [21].

Digital phenotypes are collected either actively or passively [5,22]. Active data collection requires the individual's active participation through active engagement [23]. This would include self-reported thoughts and symptoms through surveys or scales, whereas passive data collection does not require an individual to perform any specific action outside their regular activities [24]. Passively collected data, such as screen behavior, including screen time, locking or unlocking events, and screen lighting has been shown to reflect mental health states in adults, offering insights into phases of agitation, depression, and variations in sleep quality and patterns [10,23,24]. Accelerometer data from smartphones has also been linked to mental health concerns such as depression and anxiety due to reduced activity, such as mobility and traveling, while audio data, such as read-aloud recordings, help indicate depressive states [23,25]. Within DP, sleep-related features, such as sleep duration, variability, latency, and nighttime phone use, are increasingly recognized as important phenotypes that can reflect underlying mental states. Disturbed sleep, as captured through these DPs, has been linked to elevated stress and SMI [26]. In bipolar disorder, disruptions in sleep patterns can precipitate manic or depressive episodes and persist even during remission, affecting overall functioning [27-29]. Major depressive disorder (MDD) often features insomnia or hypersomnia, with sleep issues potentially preceding depressive episodes and exacerbating their severity [30]. In schizophrenia, sleep abnormalities frequently emerge before psychotic symptoms and are linked to more severe clinical outcomes [30]. Geolocation data offer insights into individuals' mobility patterns and daily routines [31]. In schizophrenia, reduced GPS-derived mobility, such as spending more time at home and traveling shorter distances, is associated with increased severity of negative symptoms such as social withdrawal and diminished motivation [32]. Similarly, in bipolar disorder, fluctuations in geolocation patterns can reflect mood episodes, with decreased movement often observed during depressive phases [33].

DP has gained significant traction in high-income countries (HICs), where studies have demonstrated its utility in detecting early warning signs (EWS), monitoring symptom fluctuations, and predicting relapse in individuals with SMIs [22,34-37]. For example, in a US-based study, passive sensing data, including GPS location, accelerometer activity, screen use, and call or text logs, was used to identify EWS of psychotic relapse. Machine learning (ML) techniques were applied to detect behavioral anomalies preceding relapse events, demonstrating the potential of DPs for relapse prediction in schizophrenia [22]. A narrative review demonstrated how various digital features, such as geolocation, communication patterns, and activity levels, have been effectively used in schizophrenia research to understand behavioral changes, and studies included were conducted in HICs [16]. In mood disorders, passive data from smartphones and wearable devices, such as GPS location, accelerometer activity, sleep pattern data, and heart rate, were used to estimate depressive and anxiety symptom severity [18,19].

Together, these passively collected DP, such as sleep, mobility, screen use, and activity patterns, offer a continuous, unobtrusive window into individuals' lived experiences. Their integration into mental health monitoring holds promise for early detection of symptom changes, relapse prediction, and timelier, personalized interventions in the care of individuals with SMI.

Relapse

Relapse is a significant concern in the management of SMIs. Relapse can lead to adverse outcomes, including poor psychosocial functioning, increased caregiver burden, and higher health care costs [38-40]. Therefore, continuous monitoring and adherence to treatment are crucial in mitigating the risk of relapse and improving long-term outcomes for individuals with SMIs. Most patients go through a period with changes in behavior, which precedes their psychotic relapse, commonly known as EWS [41]. EWS includes changes in sleep patterns, hallucinations, delusions, hostile behavior, cognitive decline, depression, and paranoia [42]. Developing systems that can identify and monitor these EWS could help clinicians intervene early [41]. In a study, anomalies in passive data, such as geolocation, accelerometer readings, and screen state, were 2.12 times more frequent in the month preceding a relapse compared to nonrelapse periods [40,43]. In another study with participants with schizophrenia, the analysis revealed that anomalies in passive data (geolocation, accelerometer data, and screen state) were significantly more frequent before a relapse [43]. Importantly, models incorporating passive data outperformed those relying solely on active survey data in predicting relapse [43]. These findings show the potential of passive smartphone data to serve as early digital markers of relapse, offering a scalable and nonintrusive method for the timely intervention and continuous monitoring in individuals with SMI.

Low- and Middle-Income Countries (LMICs)

As described above, most of the DP research has been conducted in HICs, raising concerns about generalizability to low- and middle-income countries (LMICs) [44], where digital usage patterns, social determinants of health, and mental health care infrastructure differ substantially [42]. This motivates the need to evaluate whether DP approaches can be adapted and applied effectively in LMIC settings such as Bangladesh, where the burden of untreated SMIs is high and innovative, low-cost monitoring solutions are urgently needed [45,46].

XSL•FO
RenderX

Mental and behavioral disorders account for 12% of the global disease burden, with over 70% of this impact affecting LMICs [47]. In Bangladesh, the treatment gap (ie, the gap between people who need care and the people who obtain care) is over 92% [47], which means that less than 1 in 10 people get the mental health care they need, and the gap is more evident in slums [6,45,47]. People living in slum communities have high rates of SMIs, limited access to mental health services, and conditions of chronic hardship [48]. Therefore, a low-cost, easy-to-use solution to support the identification and monitoring of mental disorders is needed, and DP may provide a solution. DP has paved its way in mental health research [9-11], but most of the research conducted has been in high-income or developed countries; these do not account for findings in LMIC settings. Yet, 80% of people with mental disorders live in LMICs [42]. Social factors such as sudden changes in life, urbanization, and poverty often lead to a high burden of mental illness in several LMICs [42,49-51]. Additionally, there is a significant lack of awareness regarding mental health conditions among LMICs [49]. Slum dwellers here have a high burden of mental health conditions, which remains unnoticed due to a lack of help-seeking or lack of resources [48,49]. The addition of DP will enhance the detection of mental disorders among these communities with little to no cost, as the use of smart devices such as smartphones is common among slum dwellers in Dhaka.

Aims and Objectives

Overview

The proposed study aims to determine whether DPs can effectively predict relapse among residents from the Korail Slum, Dhaka. One of the key outputs of this study will be to provide evidence of the feasibility and reliability of using DP in LMIC settings, particularly slum settings, for more effective mental health monitoring and intervention [47].

Primary Objective

We aimed to develop and validate predictive models using smartphone-based DP data to detect and forecast relapse in individuals with serious mental disorders living in the Korail Slum.

Secondary Objectives

For secondary objectives, our aims were:

- To explore associations between DPs and clinical facets such as mood states, social withdrawal or functioning, sleep disturbances, and activity levels.
- To assess the feasibility, acceptability, and reliability of smartphone-based DP among slum residents in a low-resource setting.
- To identify contextual and user-level factors (eg, socioeconomic, environmental, and technological) that influence data completeness and engagement with DP tools.

Methods

Outcome: Relapse

The primary outcome for prediction modeling will be the occurrence of a relapse event within a 6-month follow-up period defined using prespecified, quantitative, diagnosis-specific criteria based on validated clinical scales, supplemented by structured clinician adjudication. For participants with psychotic disorders, relapse will be defined as either (1) a ≥25% increase from baseline on the Positive and Negative Syndrome Scale (PANSS), sustained for at least 7 days, or (2) deterioration in global functioning as indicated by a decline of ≥10 points on the Global Assessment of Functioning (GAF), confirmed by clinician rating. These thresholds align with relapse definitions used in prior studies in psychosis, where relapse was operationalized using scale-based symptom worsening and functional decline [17,22,43]. For participants with nonpsychotic disorders, including MDD and PTSD, relapse will be defined as clinically significant symptom worsening on the Brief Symptom Inventory (BSI), operationalized as a ≥30% increase from baseline on the relevant symptom subscale and/or meeting caseness on the BSI Global Severity Index (GSI), defined as a total GSI T-score ≥63 together with an increase of at least 7 points from the mean of the 2 preceding assessment scores, indicating a move into or further into the clinically significant distress range [52]. In repeated-measurement and longitudinal studies involving patients with mood, anxiety, and trauma-related disorders, changes in BSI symptom subscales and the GSI have been widely used to capture clinically meaningful worsening over time [52-55]. In these populations, sustained increases from an individual's own baseline, rather than isolated score fluctuations, are commonly interpreted as reflecting true symptom deterioration. To reduce the risk of misclassifying temporary or situational symptom changes as relapse, clinically meaningful worsening is typically operationalized using relative change thresholds or standardized deviations from baseline sustained across multiple assessments.

Across all diagnostic groups, any psychiatric hospital admission or emergency psychiatric presentation during follow-up will be classified as a relapse event. All algorithm-identified relapse events will undergo blind clinician adjudication by an independent clinician using only symptom scale trajectories and clinical information; the clinician will be blinded to all DP features. Clinician judgment may confirm or refute events meeting quantitative thresholds but will not introduce relapse events in the absence of predefined criteria, thereby preventing post hoc outcome redefinition. Each relapse will be timestamped at the earliest point at which the criteria are met, enabling prediction within predefined risk horizons (eg, relapse within 30 d).

Study Design

This study is a 6-month prospective cohort study designed to explore the feasibility and reliability of identifying DP to predict relapse of SMIs among slum residents in Dhaka, Bangladesh. This study is part of the broader NIHR (National Institute for Health and Care Research) initiative, Transforming Access to Care for Serious Mental Disorders in Slums (TRANSFORM project) [47], which provides the overarching framework. The TRANSFORM project is dedicated to improving access to mental health care for individuals with SMIs residing in urban slum environments [47]. The TRANSFORM study aims to bridge the treatment gap by developing, implementing, and assessing innovative community-based interventions [47].

XSL•FO
RenderX

Setting

Participants will be recruited from the Korail slum, one of the largest and most densely populated slums in Dhaka [47]. It is situated between 2 affluent areas and is home to an estimated 200,000 people living in densely populated conditions with limited access to basic services [47,51,56,57]. The community consists of a mix of long-term residents and recent migrants from rural areas, and many inhabitants face precarious employment, housing instability, and health inequities [56,57]. Mental health needs are high, yet access to formal mental health services remains scarce [47].

Participants and Recruitment

Participants will be selected using purposive sampling. Participants will be identified through the TRANSFORM project [47], where community engagement representatives are well known and respected in the area. A community-based approach will be used to identify and enroll participants, leveraging existing partnerships with local stakeholders. We collaborated with community health workers, nongovernmental organizations, and TRANSFORM [47] community representatives, who had well-established relationships with the residents. These trusted individuals facilitated introductions and provided culturally appropriate explanations of this study to potential participants. Importantly, recruitment will not be limited to individuals already enrolled in TRANSFORM. Community representatives will help identify potential participants, introduce this study, provide culturally appropriate explanations, and facilitate trust with potential participants. Participants will progress through clearly defined stages: approached and screened for eligibility, enrolled if inclusion criteria are met, followed up monthly with both active and passive data collection, and retained in the analytic sample provided sufficient data completeness is achieved. Numbers at each stage will be reported to ensure transparency.

This study will recruit participants diagnosed with SMIs, including MDD, psychotic disorders (such as schizophrenia and bipolar disorder), and PTSD, as per the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*) [58]. These conditions were selected due to their high burden in low-resource settings, their recurrent nature, and the critical need for early detection of symptom exacerbation [42,59-61]. Inclusion criteria require participants to be at least 18 years of age, residents of the Korail slum, diagnosed with SMI, and in possession of an Android smartphone. Exclusion criteria include those younger than 18 years of age, those without a smartphone, and individuals diagnosed with substance-induced psychosis or mental illness.

Sample Size

Total Enrollment

Based on findings from the National Mental Health Survey (2019), ≈17% of adults in Bangladesh are estimated to experience a mental health disorder [62]. While specific prevalence rates of SMIs in Korail, Dhaka, are not available, this national estimate provides a useful contextual baseline. Population estimates for Korail vary widely, from 50,000 to 2,00,000 individuals, but for planning purposes, a conservative midpoint estimate of 100,000 residents was used [47,51,57].

To determine an appropriate sample size for estimating a proportion with maximum variability, the standard formula with finite population correction was applied [63-65]:

$$\text{Sample size} = z2 \times p(1-p)e21 + (z2 \times p(1-p)e2N)$$

Where Z=1.96 (for a 95% confidence level), P=0.5 (maximum variability), e=0.05 (margin of error), and n=100,000 (estimated Korail population). Substituting these parameters yields n of ≈384.16. To account for potential nonresponse or incomplete data, a 90% response rate was assumed:

$$\text{nadjusted} = 384.16/0.90 - 426.84\text{nfinal} = 426.84 \approx 430$$

Prediction Model Adequacy (Relapse Events)

To power the prediction model, we planned recruitment of 430 participants and assumed a 15% attrition rate (N is ≈366) over 6 months. The primary justification for the sample size is based on the widely accepted events per predictor (EPP) effective degree of freedom (EDF) principle for prediction model development, which mandates a minimum of 10 events per effective predictor.

Based on previous literature, relapse incidences range between 25% - 40% for psychosis [66,67], 15% - 30% for MDD [68], and 10% - 25% for PTSD [69]. Using these rates, we estimated the expected number of relapse events for our planned cohort. To account for diagnostic heterogeneity within the sample, we used existing diagnostic compositions from the TRANSFORM study [70]: (1) 45% psychosis/42% MDD/20% PTSD) − expected ≈ 66 - 121 relapse events across the incidence ranges, (2) psychosis: 0.40×366 × 0.25 - 0.40=37 - 59 events, (3) MDD: 0.40×366 × 0.15 - 0.30=22 - 44 events, and (4) PTSD: 0.20×366 × 0.10 - 0.25=7 - 18 events.

Concerning anticipated events (E), after accounting for a 15% attrition rate and the calculated relapse incidence ranges (66 - 121 events), the minimum number of events (E) is 66 relapses.

Of the required EDF, given a minimum of E=66, the maximum number of effective predictors (EDF) that the model can reliably support while maintaining the EPP>10 [71] ratio is 6.6 predictors.

About model strategy alignment, our study involves high-dimensional data (numerous digital phenotypes). To ensure compliance with the estimated EDF (≈6 - 7 predictors), the modeling pipeline will reduce the large number of raw features into a smaller set of effective predictors through penalized modeling (eg, LASSO [least absolute shrinkage and selection operator], ridge regression, or regularized tree-based methods) and dimensionality reduction or feature aggregation techniques.

Procedure

Participants will be informed about this study through participant information sheets (PIS) and will be given 48 hours to provide written informed consent. The PIS will be provided in Bengali, the local language, and will clearly explain the purpose of this study, the type of data to be collected, what participation involves, risks and benefits, data security, and the right to

XSL•FO
RenderX

withdraw at any time without affecting access to services. Participants will be contacted through community engagement representatives and ongoing recruitment activities under the TRANSFORM project [47]. These representatives, who have established relationships within the community, will distribute the PIS and introduce this study to potential participants. Trained research assistants will obtain written informed consent at the local field office with a witness present from participants willing to take part in this study. This will be done in a private setting to ensure confidentiality, and the information sheet will be explained in Bengali. After obtaining consent, the DataDoc app will be installed on participants' smartphones to passively collect data. The research team will assist them in downloading the app, granting required permissions (eg, access to device usage), setting up the app to ensure proper functionality, and troubleshooting any technical issues. Participants will be trained on app installation and basic use. Monthly clinical assessments will be conducted either in-person or via telephone, while passive data will be captured continuously in the background. Technical support will be available through community fieldworkers to address issues such as app functioning, battery management, or data sync.

Participants will install DataDoc on their smartphones for data collection as outlined below. This data will be collected via the app automatically and passively once the user gives the app permissions. A link to the app will be sent to participants via email or text (based on their preference). This is an .apk file which they can install on their Android phone. We will support participants to do this (eg, in person in the Korail field office) if needed. Upon downloading the app, they will be asked to give the app permission to access their device usage. Please note, the app also has the potential to connect to music and health data, for example, music data (ie, permission to access their Spotify [Spotify AB] data) and health data (ie, permission to access their Google Health app data), but we will only be using app usage for this study.

Relapse events will be identified using a combined approach. First, thresholds on validated scales (active data) will trigger a potential relapse flag. These events will then be reviewed by 2 independent clinicians and a researcher with expertise in psychiatry and psychology. The researcher will be blinded to the sensor-based DP data and will adjudicate relapse status using only clinical and scale-based information. In case of disagreement, a third senior clinician will review the case to reach a consensus.

DataDoc App

We have developed an app, DataDoc (Figure 1), to collect DP data, which we will use for data collection in this study. Some active data will be collected directly through the DataDoc app, while assessments requiring clinician input will be obtained separately. After giving permissions, the app will sync this data, and then the participant will be presented with a screen asking them to complete the questionnaires outlined below in the active data section. Participants will be asked to complete the questionnaire via the app only. The questionnaire must be completed within 2 weeks of syncing the device, music, and health data. Participants will be called with reminders if they have not completed the questionnaire after 1 week and again after 2 weeks. At this point, we will ask them to resync their data and complete the questionnaire, so all their data is synced to the same time point.

Figure 1. The DataDoc app.



Metrics

Participants will complete a set of baseline assessments and assessments at intervals as shown in Table 1.

XSL•FO
RenderX

Table . Timeline of data collection.

Domain, subdomains, and tools			Assessment time points							
			Baseline	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Continuously
Active data										
	Sociodemographic profile		✓							
	Clinical profile and history									
		Diagnosis	✓							
		Timeline of diagnosis	✓							
	Clinical assessments									
		Global Assessment of Functioning (GAF)	✓	✓	✓	✓	✓	✓	✓	
		WHOQOL (World Health Organization Quality of Life)	✓	✓	✓	✓	✓	✓	✓	
		Brief Symptom Inventory (BSI) or the Positive and Negative Syndrome Scale (PANSS)	✓	✓	✓	✓	✓	✓	✓	
		Perceived Stress Scale	✓	✓	✓	✓	✓	✓	✓	
		Sleep questionnaire	✓	✓	✓	✓	✓	✓	✓	
	Thought patterns									
		Journaling (optional)	✓	✓	✓	✓	✓	✓	✓	
Passive data										
	Device interactions									
		Phone usage								✓
		Application usage								✓
		Screen time								✓
	Digital interaction									

XSL•FO
RenderX

Domain, subdomains, and tools		Assessment time points							
		Baseline	Month 1	Month 2	Month 3	Month 4	Month 5	Month 6	Continuously
	Calls (frequency)								✓
	Duration of calls								✓
	Messages (frequency)								✓
Activities (if applicable)									
	Activities app (sleep and health)								✓
	Sleep								✓
Movement and location									
	Geolocation (longitude and latitude)								✓
	Duration at location								✓
	Movement time								✓

Passive Data

Passive data will be collected continuously, on an event-by-event basis, and will include:

- Device interactions: this study will collect data on how a participant interacts with their digital devices. This includes phone usage, application usage (application history such as application on the foreground, number of application notifications, when an application is installed and removed), and screen time.
- Digital interaction: calls and messaging habits (number of incoming and outgoing messages and calls), frequencies of responses (duration of call; the content of communications will not be accessed or stored).
- Activities (if applicable): sleep and health application installed or removed, and total sleep duration (integrated within DataDoc app).
- Movement and location: geolocation (longitude and latitude), duration at one location, and movement time.

Active Data

Demographic Questionnaire

Each participant will be provided with a demographic questionnaire containing questions regarding age, gender, level of education, occupation, diagnosis profile, and home address.

Active data will consist of the following and will be collected monthly.

Clinical Data

The clinical data consisted of the following:

- GAF is a clinician-rated scale to assess global functioning. GAF represents a latent continuous measure score from 1 (lowest score representing worst symptomology) to 100 (high score representing optimal level of functioning) [72]. This data will be collected from the ongoing TRANSFORM Project [47].
- WHOQOL-BREF (World Health Organization Quality of Life-Brief) is a 26-item self-reported scale to assess quality of life, where scores range from 16 to 112, and a higher score represents a higher quality of life [73].
- BSI is a 53-item self-reported instrument on a 5-point scale used to evaluate psychological symptoms over the past week in patients, nonpatients, and participants involved in experimental research [74]. The BSI consists of 9 main symptom categories: somatization, obsessive-compulsive tendencies, sensitivity in relationships, feelings of depression, anxiety, hostility, phobic anxiety, paranoid thoughts, and tendencies toward psychosis, and higher scores represent more severe symptoms [74].
- The PANSS is a scale used for measuring symptom severity with 3 components of positive, negative, and general psychopathology scales [75]. The total PANSS score is the sum of the scores of the 3 scales, ranging from a minimum of 30 to a maximum of 210, with higher scores indicating more severe symptoms [75].

XSL•FO
RenderX

- Perceived Stress Scale [76] to assess perception of stress with scores ranging from 0 to 40, where higher scores represent higher perceived stress.
- The Nottingham Onset Schedule is a structured interview tool designed to assess the onset of psychosis, and it provides a systematic method of gathering detailed information about the initial presentation and progression of symptoms [77].

Activities

The activities consisted of the following:

- Self-reported sleep patterns through a dedicated general questionnaire for this study derived from the sleep questionnaire for adults (part two) by the National Health Service, which will explore the time of the participant going to bed, hours of sleep, number of times the participant woke up, perceived quality of sleep, etc [78].
- For journaling, participants will be asked to "journal," that is, write down their thoughts and feelings if they wish [79].

Data Privacy and Security

As illustrated in Figure 2, data is first collected on the participant's phone using the DataDoc app. Data are then securely uploaded to a General Data Protection Regulation–compliant cloud server, where they are stored in encrypted form at rest. From there, encrypted anonymized data are transferred to the University of Warwick's secure server for long-term storage and analysis, in line with Warwick's 10-year research data policy. No personally identifiable information (eg, phone numbers and message content) is collected at any stage. Access to both the cloud server and Warwick's secure server is restricted to authorized study personnel using password-protected accounts.

Figure 2. Data flow.

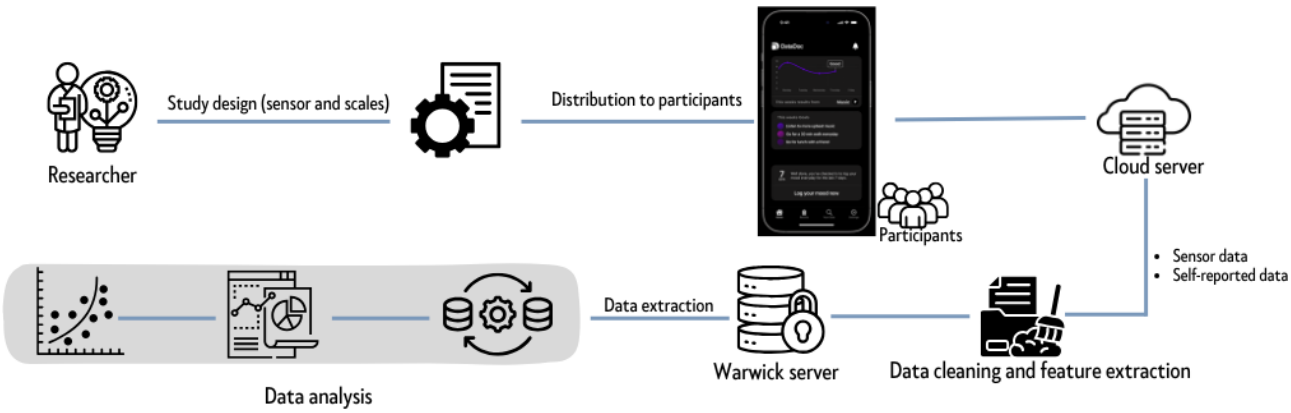

Analysis

Overview

This study adopts a multimodal ML-driven analytical framework to integrate and analyze both active and passive data streams collected via the DataDoc app. The goal is to develop predictive models for relapse detection and to explore associations between specific digital phenotypes and clinical dimensions. The analysis pipeline will involve preprocessing of the collected active and passive data, feature engineering, labeling (when needed), ML model selection, model fitting, and evaluating the predictive validity of our multimodal data, followed by further interpretation and deployment when needed.

Analysis of Active Data

Data Preprocessing

Cleaning and Standardization

We will remove any inconsistencies and standardize the format of the active data to ensure uniformity. For demographic questionnaire variables such as gender, occupation, or diagnosis, any typographical errors will be corrected, and if any absurd numerical value is recorded for age or education, it will be flagged for replacement. For clinical assessment instruments, it must be ensured that item responses are within the appropriate range of the corresponding scales. If not, the out-of-range and/or strange entries will be flagged for further systematic treatments.

Handling Missing Values

Missing responses will be addressed through multiple imputation techniques or median replacement where feasible, and guided by manuals as necessary. For demographic variables, if there are randomly missing variables, then mean or median imputation for numerical variables and mode for categorical variables may be considered. In cases of substantial missingness, a separate category such as "unknown" may be assigned. A threshold for what constitutes "substantial" will be prespecified during analysis planning. For the clinical assessment scales, if there is item-level missingness, then the person's mean imputation will be used. If there is any block-level missingness that affects entire subscales, then the participant would be flagged and dropped if needed.

Encoding Categorical Variables

Categorical variables, such as demographic data like gender, will be converted into numerical representations suitable for computational steps and ML models. Variables such as education level can be mapped to ascending integers, and nominal ones such as occupation can be encoded with 1-hot or label encoding.

XSL•FO
RenderX

Feature Engineering

Extraction of Relevant Features

We will identify and extract features from active data that are indicative of behavioral and psychological factors associated with relapse. Demographic variables such as occupation or education level can be grouped into a factor such as socioeconomic status if deemed necessary for future analysis. From the responses on the clinical assessment instruments such as GAF and others, scoring will be done as per their respective manuals after ensuring all responses are in their appropriate ranges. Additionally, for scales such as WHOQOL-BREF, where specific items such as items 3 and 4 are also reverse-worded, they will be reverse-scored before further processing. Its domains, such as physical health, psychological health, social relationships, and environment, will also have their domain scores from averaging across items of each domain. Similarly, for BSI, subscale scores will be computed for all the symptom dimensions as well as the global indices. For PANSS, subscores of domains (positive, negative, and general psychopathology) will be computed from summing items of each domain, while total PANSS will be the sum of all items. For a scale such as Nottingham Onset Schedule, where key outputs are dates such as that of the onset of prodrome or psychosis, the date-field variables will be converted to the computational datetime format, following which durations such as that of untreated psychosis will be computed comparing as intervals between symptom onset and treatment start dates or likewise.

Symptom Severity Trends

Longitudinal changes in clinical assessments (eg, PANSS and BSI scores) will be modeled to track deterioration or improvement.

Sentiment and Thematic Analysis

Journaling entries will undergo natural language processing (NLP)–based sentiment and thematic analysis to extract emotional patterns. Given the linguistic context of Dhaka slum residents, entries are often written in a mixture of Bengali and romanized Bengali (“Banglish”), and off-the-shelf pretrained models are limited for such code-mixed and colloquial data. We therefore will adopt a staged approach: (1) start with strong multilingual transformers (mBERT and XLM-R); (2) fine-tune Bengali models (BanglaBERT [80] or IndicBERT [81]) on a small, study-specific, annotated code-mixed corpus; and (3) add light normalization (transliteration handling; domain lexicon for common Banglish expressions). Multilingual transformers (especially XLM-R) generally outperform vanilla mBERT on low-resource languages, and Bengali-specific models (BanglaBERT and IndicBERT) provide monolingual priors we can adapt via task-specific fine-tuning. To ensure validity, we will manually label a stratified subsample (≈10% - 20% of entries) for sentiment or topic, compute interannotator agreement, and use this “gold” to calibrate and evaluate the NLP models; NLP features will be treated as exploratory covariates rather than primary end points. Recent work has released Bengali-English code-mixed sentiment datasets (eg, BnSentMix) [82] and shows that fine-tuning or further pretraining on code-mixed data substantially improves over vanilla multilingual models, which supports our feasibility plan.

Analysis of Passive Data

Overview

High-dimensional smartphone sensor features (eg, mobility, sleep, social communication, and screen-use variability) will be condensed into a limited number of aggregated, clinically interpretable composite variables before model training. Analyses will use penalized regression and regularized tree-based algorithms (eg, gradient boosting) to mitigate overfitting.

Data Preprocessing

Noise Reduction

Continuous, high-frequency passive raw data, as captured from the digital sensors, are often noisy, containing irrelevant signals that need to be smoothed out for more reliable analysis and modeling. Therefore, artifacts such as random or closely spaced idle screen event sessions can be merged, and spatial smoothing and clustering can be applied over GPS jitters. Zero-duration, irrelevant calls or text communications will be filtered out. Kalman filters or moving averages will be used for smoothing out sensor spikes for motion data.

Outlier Detection

Tailored to the behavioral context in question, anomaly detection techniques for variables from such passive data sources will be used. To start, the fast and easily interpretable methods, such as IQR-based filtering, will be applied to handle and assess the quality of individual behavioral variables such as screen time, call counts, and their durations, duration of app use or of movement, and of movement distance as well. For more multivariate, time-varying, and longitudinal contexts, techniques of isolation forests and rolling *z* scores will identify and address deviations and outliers from the standard range and/or baseline.

Normalization

We will normalize passive data to ensure consistent scales across different features. To maintain comparability across different devices and participants, minimum-maximum scaling or *z* score normalization will be applied to passive data features.

Feature Engineering

Feature Extraction

We will be extracting relevant information from continuous passive data streams.

ML Steps

In this study, predictive modeling using ML will focus on leveraging the DPs collected through participants’ smartphones as the primary features (predictors), with relapse as the primary target outcome. Clinical symptom scores (eg, PANSS, BSI, and WHOQOL-BREF) will serve as secondary outcomes for exploratory analyses and for validating relapse definitions. By analyzing patterns in the DPs, such as device interactions, sleep patterns, movement, and app usage, the model will learn to identify subtle behavioral changes that may be indicative of

XSL•FO
RenderX

mental health shifts, such as the onset of relapse. By training the model on historical data from participants, where mental health conditions have been assessed via standardized clinical questionnaires, the system will aim to forecast future mental health status or relapse risk with high accuracy. This approach will enable the integration of both passive smartphone data and active clinical assessments to provide a comprehensive and early-warning prediction system for serious mental disorders in low-income settings.

Primary Analysis

Overview

For the ML analysis, we define the prediction horizon as the risk of relapse within the next 30 days. This choice is informed by prior DP research demonstrating elevated behavioral anomalies in the 2 weeks preceding relapse and increased anomaly frequency within a 30-day window surrounding relapse [17,83]. Features from passive and active smartphone data will be derived using a 14-day look-back window before the relapse index date. To reduce bias from the clustering of events, a 30-day washout period after each relapse will be applied, during which no new relapse event will be coded.

The primary supervised modeling approach follows a structured analytical pipeline, beginning with temporal feature aggregation, model training, and validation. Passive DP features, including screen usage, mobility, and sleep-related indicators, will be aggregated across predefined sliding windows (eg, last 1, 7, and 30 d) to condense high-frequency time series data into summary statistics. This aggregation facilitates noise reduction, captures short-term behavioral trends, and converts raw sensor streams into fixed-length feature vectors suitable for supervised learning.

As aggregation may introduce feature redundancy and multicollinearity, dimensionality reduction and feature selection techniques, including principal component analysis, singular value decomposition, regularized LASSO regression, mutual information filtering, and SHAP (Shapley Additive Explanations)-based feature importance assessment, will be applied to retain the most informative behavioral features while preserving interpretability. Interaction features will also be derived to examine nonlinear relationships between digital phenotypes (eg, movement variability and screen usage), enabling representation of higher-level behavioral constructs such as mobility, rest, activity, social behavior, and device use. These derived components will be mapped to labeled relapse and nonrelapse windows and aligned with static demographic or clinical covariates, when applicable.

All primary models will be trained with relapse as the target outcome. To ensure robust evaluation and prevent temporal and subject-level data leakage, the dataset will be divided into time-aware training (70% - 80%) and testing (20% - 30%) partitions, ensuring that future behavioral data are not used to predict earlier outcomes. Model selection and hyperparameter tuning will be conducted using nested cross-validation within training folds. Temporal validation procedures, including rolling window and expanding window validation, will be used to assess model stability over time. Nested cross-validation will also be used to minimize overfitting and optimistic bias.

Given the relatively low frequency of relapse events, class imbalance will be addressed using case-control sampling by time (matching nonrelapse windows from the same observation period) and cost-sensitive learning approaches, with higher penalties assigned to misclassified relapse events. Synthetic oversampling techniques, such as the synthetic minority oversampling technique, will not be applied due to the risk of temporal leakage in longitudinal data.

Primary Model Specific and Evaluation

The prespecified primary supervised model will be gradient boosting (XGBoost or LightGBM), selected a priori for its robust performance with multimodal behavioral data and demonstrated effectiveness in prior DP studies of relapse prediction [7,17]. The gradient boosting model, together with the feature construction and validation strategy described above, constitutes the single prespecified primary modeling pipeline for this study. Model performance will be internally validated using temporal cross-validation and bootstrap optimism correction, with reporting of discrimination (area under the receiver operating characteristic curve) and calibration metrics (calibration slope and Brier score). At a prespecified probability threshold, sensitivity, specificity, positive predictive value, and negative predictive value will also be reported for relapse prediction within the 30-day horizon. Calibration quality will be further assessed using expected calibration error, and decision curve analysis will be conducted to evaluate potential clinical utility across a range of risk thresholds. Explainability methods, including SHAP values and permutation importance, will be applied to identify the most influential digital phenotypes contributing to relapse risk.

Secondary and Exploratory Analyses

Alternative Supervised Models

As sensitivity analyses, the performance of the primary gradient boosting model will be benchmarked against regularized logistic regression and random forest models. These models provide interpretable and clinically acceptable baselines and will be trained and evaluated using the same feature set, outcome definition, and validation framework as the primary analysis.

Unsupervised and Deep Learning Approaches

As exploratory analyses, unsupervised anomaly detection methods (eg, autoencoders and clustering approaches) and time-series deep learning models (eg, long short-term memory and transformer-based architectures) will be examined to explore temporal behavioral patterns and relapse trajectories. These analyses will be contingent on data volume and relapse event frequency and are intended for hypothesis generation and pattern discovery rather than primary inference.

Exploratory Use of Active Data and Data Fusion

Analyses using clinical symptom scores (eg, PANSS, BSI, and WHOQOL-BREF) will be conducted as secondary exploratory models to contextualize relapse prediction and support interpretation of behavioral features. To maximize predictive insight, multimodal fusion approaches will be explored at the feature and decision levels. Feature-level fusion will integrate passive DP data with active clinical assessments, while

XSL•FO
RenderX

decision-level fusion will involve combining outputs from multiple models using ensemble techniques such as stacking and boosting. Multimodal fusion and ensemble methods will be used strictly as secondary analyses to complement, but not replace, the prespecified primary supervised model [7,17,83].

Ethical Considerations

This study was approved by the Biomedical and Scientific Research Ethics Committee at the University of Warwick (reference number Biomedical and Scientific Research Ethics Committee 98/23 - 24). Data collected from participants are securely stored and will not be available for sharing due to privacy considerations. The smartphone app (DataDoc) used for data collection and the feature processing code are available upon request for research purposes.

Participants were informed about the nature, procedures, potential risks, benefits, and voluntary nature of the research before enrollment. Informed consent was obtained from all individual participants before data collection. Participants received detailed information sheets and provided written consent confirming their voluntary participation and understanding of their rights, including the right to withdraw at any time without penalty. The consent process ensured comprehension of this study's aims, data collection procedures, privacy protections, and how data would be used and stored.

To recognize participants' time and engagement in this study, compensation was provided at a rate of 1200 BDT per month for the duration of their involvement. This compensation was disclosed to participants during the consent process and was approved as part of the ethics review.

All data collected from participants are securely stored in accordance with local and international data protection standards and are not publicly available due to privacy and confidentiality considerations. Identifiable personal information has been protected and anonymized where possible to prevent unintended disclosure.

Results

This study received funding from the UK NIHR (award number NIHR200846) under the TRANSFORM Project [47]. Pilot data collection began in August 2025 in the Korail slum, Dhaka, and as of October 2025, recruitment is ongoing. Data analysis, including feasibility and predictive modeling, commenced in December 2025, with preliminary findings anticipated in mid-2026 and expected results to be published in fall 2026. A total of 14 participants have been recruited, as of January 2026.

Discussion

Principal Findings

The results of this study will offer critical insights into the application of DP for predicting the relapse of SMIs in low-resource settings, specifically among slum residents in Dhaka, Bangladesh. By using passive data collected from participants' smartphones, such as device interactions, sleep patterns, movement, and social engagement, this study aims to assess the ability of ML models to detect early signs of relapse. Our findings will contribute to the broader understanding of how DP can be applied in real-world settings to improve mental health care, especially in LMICs where access to traditional mental health services is difficult or limited [42,49].

This study's innovative approach of combining passive data with clinical assessments allows for the development of personalized, dynamic interventions tailored to individuals' behavioral patterns. This could represent a significant advance over static mental health monitoring methods, offering more timely and effective interventions. However, several challenges and limitations must be acknowledged. First, while DP offers a novel way to monitor behavioral changes, the accuracy and validity of these markers in predicting clinical outcomes require further exploration. The reliance on self-reported clinical assessments may introduce bias [4], as participants' perceptions of their mental health may not always align with objective symptoms. Moreover, such assessments often lack ecological validity, as they are typically conducted in clinical settings and may not capture the nuances of daily life. In contrast, DP offers the potential to monitor individuals in real-world contexts, providing richer insights into behavior and symptom fluctuations outside the clinic [12,13]. Furthermore, privacy concerns surrounding the continuous collection of personal data must be addressed [12], even though strict data protection protocols are in place.

The results will also have implications beyond the scope of this study. The potential for DP biomarkers to serve as dynamic tailoring variables in adaptive interventions could reshape how mental health treatments are delivered, particularly in resource-poor settings. As smartphones become increasingly prevalent in LMICs, leveraging these devices for mental health care could provide a scalable, cost-effective solution that reaches underserved populations. The findings from this study will help guide future research and intervention strategies, paving the way for larger, more comprehensive studies in similar settings.

Limitations

A key limitation of this study is the requirement for participants to own an Android smartphone, which may systematically exclude individuals with SMIs who are older, poorer, or have lower educational attainment. These subgroups may be among those most vulnerable to relapse, meaning our sample could underrepresent the most at-risk individuals [84,85]. To mitigate bias, we will document sociodemographic and clinical characteristics of enrolled participants and compare these with available data on the broader Korail slum population. In addition, subgroup analyses stratified by socioeconomic and demographic variables will be conducted to evaluate whether predictive performance varies across groups. These steps will help us interpret findings in light of potential exclusion and inform future efforts to reach harder-to-engage individuals with SMIs.

Finally, our sample size estimation was based on the national prevalence of mental disorders in Bangladesh (17%). We acknowledge that this may not accurately capture the epidemiology of SMI in slum settings, where prevalence may be higher or differently distributed due to socioeconomic disadvantage, overcrowding, and limited access to care. As

XSL•FO
RenderX

such, our study may be under- or overpowered relative to the true prevalence in Korail. This limitation will be transparently acknowledged in reporting and highlights the need for more robust epidemiological studies of SMI within urban slum populations.

Several practical limitations may affect feasibility. First, requiring participants to install the app as an .apk file may present challenges for individuals with low technical literacy. To address this, in-person installation support and illustrated step-by-step instructions in Bengali will be provided. Second, continuous passive data collection, particularly from GPS sensors, has the potential to increase battery consumption. Given that there is potentially unstable access to electricity in slum environments, participants may disable or uninstall the app to conserve power. While the DataDoc app has been optimized for low-frequency sampling to reduce power demand, battery drain remains a possible barrier to adherence. Furthermore, to minimize battery drainage in a setting with intermittent electricity, the app will upload data only when connected to Wi-Fi or charging [86]. Third, concerns about data privacy and security may limit participant engagement. Although strict protocols are in place, including anonymization and secure server storage, participants may still perceive privacy risks, which could affect long-term acceptability. To help mitigate this, a prior qualitative study was conducted with the same Korail community to explore participants' understanding, perspectives, and willingness to share digital data [87]. Insights from this study informed the design of our consent process and data collection strategy. Nevertheless, residual concerns about privacy remain a possible barrier. We will therefore monitor such issues throughout this study and systematically document dropout or data loss attributable to these barriers.

Conclusions

This study represents a step forward in understanding how DP and ML can be used to predict relapse in individuals with serious mental disorders living in low-resource settings. By combining passive smartphone data with clinical assessments, we aim to build a predictive model that can identify the EWS of relapse, allowing for timely interventions.

While there are challenges associated with the use of DP, such as data privacy and the interpretation of behavioral markers, the potential benefits for early detection and intervention in mental health are substantial. The findings from this study will not only contribute to the growing body of research on DP but also offer practical solutions for addressing the mental health treatment gap in underserved populations. Ultimately, this research could help improve the quality of mental health care in LMIC settings and impoverished communities and provide a model for scalable, technology-driven interventions in other low-resource environments. This research has the potential to inform both local and global strategies for improving access to mental health care in low-resource settings.

Acknowledgments

We also acknowledge the work of the Transforming Access to Care for Serious Mental Disorders in Slums (TRANSFORM) project and TRANSFORM consortium. We also acknowledge the support of our local community representatives: Ms Aulia Khatun, Mr Md Abdur Rahman, Ms Ragina Aktar Rumi, Mr Md Abdul Kashem Titu, and Mrs Raseda. The lead author and manuscript guarantor, NA, affirms that this manuscript is an honest, accurate, and transparent account of the study being reported. No important aspects of this study have been omitted, and any discrepancies from the study as originally planned have been clearly explained.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Data Availability

The data generated and analyzed during this study are not publicly available due to privacy and ethical considerations involving sensitive participant information. However, deidentified data or a subset of the data may be made available from the principal investigator upon reasonable request and are subject to appropriate ethical approvals. The smartphone app (DataDoc) and associated data processing code used in this study are available for research purposes on request.

Authors' Contributions

NA and SJ conceived this study, formulated the research questions, and led this study's design, protocol development, and manuscript writing. CKD contributed to the clinical components of the study design and manuscript editing. NR contributed to the statistical and machine learning analysis plan, data pipeline design, and provided critical revisions of this paper. DG provided supervision on study methodology, ethical considerations, and contributed to paper refinement. SPS provided conceptual input during this study's formulation and offered senior academic oversight and critical revisions. All authors reviewed and approved this final paper.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Viron MJ, Stern TA. The impact of serious mental illness on health and healthcare. Psychosomatics 2010;51(6):458-465. [doi: 10.1176/appi.psy.51.6.458] [Medline: 21051676]
2. Agenagnew L, kassaw C. The lifetime prevalence and factors associated with relapse among mentally ill patients at Jimma University Medical Center, Ethiopia: cross sectional study. J Psychosoc Rehabil Ment Health 2020 Dec;7(3):211-220. [doi: 10.1007/s40737-020-00176-7]
3. Ustün TB. The global burden of mental disorders. Am J Public Health 1999 Sep;89(9):1315-1318. [doi: 10.2105/ajph.89.9.1315] [Medline: 10474545]
4. Abdullah S, Choudhury T. Sensing technologies for monitoring serious mental illnesses. IEEE MultiMedia 2018 Jan;25(1):61-75. [doi: 10.1109/MMUL.2018.011921236]
5. Onnela JP, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology 2016 Jun;41(7):1691-1696. [doi: 10.1038/npp.2016.7] [Medline: 26818126]
6. Gruebner O, Khan MMH, Lautenbach S, et al. A spatial epidemiological analysis of self-rated mental health in the slums of Dhaka. Int J Health Geogr 2011 May 20;10:36. [doi: 10.1186/1476-072X-10-36] [Medline: 21599932]
7. Henson P, D'Mello R, Vaidyam A, Keshavan M, Torous J. Anomaly detection to predict relapse risk in schizophrenia. Transl Psychiatry 2021 Jan 11;11(1):28. [doi: 10.1038/s41398-020-01123-7] [Medline: 33431818]
8. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. JMIR Ment Health 2016 Jan;3(2):e16. [doi: 10.2196/mental.5165]
9. Coghlan S, D'Alfonso S. Digital phenotyping: an epistemic and methodological analysis. Philos Technol 2021 Dec;34(4):1905-1928. [doi: 10.1007/s13347-021-00492-1]
10. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. NPJ Digit Med 2019 Sep 6;2(1):88. [doi: 10.1038/s41746-019-0166-1]
11. Oudin A, Maatoug R, Bourla A, et al. Digital phenotyping: data-driven psychiatry to redefine mental health. J Med Internet Res 2023 Oct 4;25:e44502. [doi: 10.2196/44502] [Medline: 37792430]
12. Henson P, Barnett I, Keshavan M, Torous J. Towards clinically actionable digital phenotyping targets in schizophrenia. npj Schizophr 2020 May 5;6(1):13. [doi: 10.1038/s41537-020-0100-1] [Medline: 32372059]
13. Jilka S, Giacco D. Digital phenotyping: how it could change mental health care and why we should all keep up. J Ment Health 2024 Aug;33(4):439-442. [doi: 10.1080/09638237.2024.2395537] [Medline: 39301756]
14. Bourla A, Ferreri F, Ogorzelec L, Guinchard C, Mouchabac S. Évaluation des troubles thymiques par l'étude des données passives : le concept de phénotype digital à l'épreuve de la culture de métier de psychiatre. L'Encéphale 2018 Apr;44(2):168-175. [doi: 10.1016/j.encep.2017.07.007]
15. Difrancesco S, Fraccaro P, van der Veer SN, et al. Out-of-home activity recognition from GPS data in schizophrenic patients. 2016 Presented at: 2016 IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS); Dublin p. 324-328. [doi: 10.1109/CBMS.2016.54]
16. Lane E, D'Arcey J, Kidd S, et al. Digital phenotyping in adults with schizophrenia: a narrative review. Curr Psychiatry Rep 2023 Nov;25(11):699-706. [doi: 10.1007/s11920-023-01467-z] [Medline: 37861979]
17. Barnett I, Torous J, Staples P, Sandoval L, Keshavan M, Onnela JP. Relapse prediction in schizophrenia through digital phenotyping: a pilot study. Neuropsychopharmacol 2018 Jul;43(8):1660-1666. [doi: 10.1038/s41386-018-0030-z]
18. Pedrelli P, Fedor S, Ghandeharioun A, et al. Monitoring changes in depression severity using wearable and mobile sensors. Front Psychiatry 2020;11:584711. [doi: 10.3389/fpsyt.2020.584711] [Medline: 33391050]
19. Jacobson NC, Lekkas D, Huang R, Thomas N. Deep learning paired with wearable passive sensing data predicts deterioration in anxiety disorder symptoms across 17-18 years. J Affect Disord 2021 Mar 1;282:104-111. [doi: 10.1016/j.jad.2020.12.086] [Medline: 33401123]
20. Lu J, Shang C, Yue C, et al. Joint modeling of heterogeneous sensing data for depression assessment via multi-task learning. Proc ACM Interact Mob Wearable Ubiquitous Technol 2018 Mar 26;2(1):1-21. [doi: 10.1145/3191753]
21. Bufano P, Laurino M, Said S, Tognetti A, Menicucci D. Digital phenotyping for monitoring mental disorders: systematic review. J Med Internet Res 2023 Dec 13;25:e46778. [doi: 10.2196/46778] [Medline: 38090800]
22. Adler DA, Ben-Zeev D, Tseng VWS, et al. Predicting early warning signs of psychotic relapse from passive sensing data: an approach using encoder-decoder neural networks. JMIR mHealth uHealth 2020 Aug 31;8(8):e19962. [doi: 10.2196/19962] [Medline: 32865506]
23. Melcher J, Lavoie J, Hays R, et al. Digital phenotyping of student mental health during COVID-19: an observational study of 100 college students. J Am Coll Health 2023 Mar 24;71(3):736-748. [doi: 10.1080/07448481.2021.1905650]
24. Liang Y, Zheng X, Zeng DD. A survey on big data-driven digital phenotyping of mental health. Inf Fusion 2019 Dec;52:290-307. [doi: 10.1016/j.inffus.2019.04.001]
25. Stasak B, Epps J, Schatten HT, Miller IW, Provost EM, Armey MF. Read speech voice quality and disfluency in individuals with recent suicidal ideation or suicide attempt. Speech Commun 2021 Sep;132:10-20. [doi: 10.1016/j.specom.2021.05.004]
26. Ered A, Cooper S, Ellman LM. Sleep quality, psychological symptoms, and psychotic-like experiences. J Psychiatr Res 2018 Mar;98:95-98. [doi: 10.1016/j.jpsychires.2017.12.016] [Medline: 29331930]

XSL•FO
RenderX

27. Harvey AG, Kaplan KA, Soehner AM. Interventions for sleep disturbance in bipolar disorder. Sleep Med Clin 2015 Mar;10(1):101-105. [doi: 10.1016/j.jsmc.2014.11.005] [Medline: 25750600]
28. Robillard R, Hermens DF, Lee RSC, et al. Sleep-wake profiles predict longitudinal changes in manic symptoms and memory in young people with mood disorders. J Sleep Res 2016 Oct;25(5):549-555. [doi: 10.1111/jsr.12413] [Medline: 27242171]
29. Harvey AG, Talbot LS, Gershon A. Sleep disturbance in bipolar disorder across the lifespan. Clin Psychol: Sci Pract 2009 Jun;16(2):256-277. [doi: 10.1111/j.1468-2850.2009.01164.x]
30. Mendlewicz J. Sleep disturbances: core symptoms of major depressive disorder rather than associated or comorbid disorders. World J Biol Psychiatry 2009;10(4):269-275. [doi: 10.3109/15622970802503086] [Medline: 19921968]
31. Palmius N, Tsanas A, Saunders KEA, et al. Detecting bipolar depression from geographic location data. IEEE Trans Biomed Eng 2017 Aug 1;64(8):1761-1771. [doi: 10.1109/TBME.2016.2611862]
32. Raugh IM, James SH, Gonzalez CM, et al. Digital phenotyping adherence, feasibility, and tolerability in outpatients with schizophrenia. J Psychiatr Res 2021 Jun;138:436-443. [doi: 10.1016/j.jpsychires.2021.04.022] [Medline: 33964681]
33. Fraccaro P, Beukenhorst A, Sperrin M, et al. Digital biomarkers from geolocation data in bipolar disorder and schizophrenia: a systematic review. J Am Med Inform Assoc 2019 Nov 1;26(11):1412-1420. [doi: 10.1093/jamia/ocz043] [Medline: 31260049]
34. Currey D, Torous J. Digital phenotyping data to predict symptom improvement and mental health app personalization in college students: prospective validation of a predictive model. J Med Internet Res 2023;25:e39258. [doi: 10.2196/39258]
35. Andrea A, Agulia A, Serafini G, Amore M. Digital biomarkers and digital phenotyping in mental health care and prevention. Eur J Public Health 2020 Sep 1;30(Supplement_5):ckaa165.1080. [doi: 10.1093/eurpub/ckaa165.1080]
36. Hays R, Keshavan M, Wisniewski H, Torous J. Deriving symptom networks from digital phenotyping data in serious mental illness. BJPsych Open 2020 Nov 3;6(6):e135. [doi: 10.1192/bjo.2020.94] [Medline: 33138889]
37. Henson P, Rodriguez-Villa E, Torous J. Investigating associations between screen time and symptomatology in individuals with serious mental illness: longitudinal observational study. J Med Internet Res 2021 Mar 1;23(3):e23144. [doi: 10.2196/23144]
38. Amha H, Getnet A, Munie BM, et al. Relapse rate and predictors among people with severe mental illnesses at Debre Markos Comprehensive Specialized Hospital, Northwest Ethiopia: a prospective follow up study. Eur Arch Psychiatry Clin Neurosci 2024 Sep 18. [doi: 10.1007/s00406-024-01900-1] [Medline: 39292261]
39. Leucht S, Tardy M, Komossa K, et al. Antipsychotic drugs versus placebo for relapse prevention in schizophrenia: a systematic review and meta-analysis. Lancet 2012 Jun 2;379(9831):2063-2071. [doi: 10.1016/S0140-6736(12)60239-6] [Medline: 22560607]
40. Zipursky RB, Menezes NM, Streiner DL. Risk of symptom recurrence with medication discontinuation in first-episode psychosis: a systematic review. Schizophr Res 2014 Feb;152(2-3):408-414. [doi: 10.1016/j.schres.2013.08.001] [Medline: 23972821]
41. Gumley AI, Bradstreet S, Ainsworth J, et al. Digital smartphone intervention to recognise and manage early warning signs in schizophrenia to prevent relapse: the EMPOWER feasibility cluster RCT. Health Technol Assess 2022;26(27):1-174. [doi: 10.3310/HLZE0479]
42. Rathod S, Pinninti N, Irfan M, et al. Mental health service provision in low- and middle-income countries. Health Serv Insights 2017 Jan 1;10. [doi: 10.1177/1178632917694350] [Medline: 28469456]
43. Cohen A, Naslund JA, Chang S, et al. Relapse prediction in schizophrenia with smartphone digital phenotyping during COVID-19: a prospective, three-site, two-country, longitudinal study. Schizophrenia (Heidelb) 2023 Jan 27;9(1):6. [doi: 10.1038/s41537-023-00332-5] [Medline: 36707524]
44. Naslund JA, Deng D. Addressing mental health stigma in low-income and middle-income countries: a new frontier for digital mental health. Ethics Med Public Health 2021 Dec;19:100719. [doi: 10.1016/j.jemep.2021.100719] [Medline: 35083375]
45. Hossain MD, Ahmed HU, Chowdhury WA, Niessen LW, Alam DS. Mental disorders in Bangladesh: a systematic review. BMC Psychiatry 2014 Jul 30;14(1):216. [doi: 10.1186/s12888-014-0216-9] [Medline: 25073970]
46. Ahmed T, Rizvi SJR, Rasheed S, et al. Digital health and inequalities in access to health services in Bangladesh: mixed methods study. JMIR mHealth uHealth 2020 Jul 21;8(7):e16473. [doi: 10.2196/16473] [Medline: 32706736]
47. Singh SP, Jilka S, Abdulmalik J, et al. Transforming access to care for serious mental disorders in slums (the TRANSFORM Project): rationale, design and protocol. BJPsych Open 2022 Oct 13;8(6):e185. [doi: 10.1192/bjo.2022.584] [Medline: 36226591]
48. Gruebner O, Khan MMH, Lautenbach S, et al. Mental health in the slums of Dhaka - a geoepidemiological study. BMC Public Health 2012 Mar 9;12(1):177. [doi: 10.1186/1471-2458-12-177] [Medline: 22404959]
49. Alloh FT, Regmi P, Onche I, Teijlingen EV, Trenoweth S. Mental health in low-and middle income countries (LMICs): going beyond the need for funding. Health Prospect 2018 Jun 19;17(1):12-17. [doi: 10.3126/hprospect.v17i1.20351]
50. Gamieldien F, Galvaan R, Myers B, Syed Z, Sorsdahl K. Exploration of recovery of people living with severe mental illness (SMI) in low/middle-income countries (LMICs): a scoping review. BMJ Open 2021 Mar 24;11(3):e045005. [doi: 10.1136/bmjopen-2020-045005] [Medline: 33762242]

XSL•FO
RenderX

51. Sinthia SA. Analysis of urban slum: case study of Korail Slum, Dhaka. Int J Urban Civil Eng 2020;14(11):416-430 [FREE Full text]
52. Derogatis LR. BSI, Brief Symptom Inventory: Administration, Scoring, & Procedures Manual: National Computer Systems, 1993; 1993.
53. Franke GH, Jagla-Franke M, Küch D, Petrowski K. A new routine for analyzing brief symptom inventory profiles in chronic pain patients to evaluate psychological comorbidity. Front Psychol 2021;12:692545. [doi: 10.3389/fpsyg.2021.692545] [Medline: 34659009]
54. Michel G, Baenziger J, Brodbeck J, Mader L, Kuehni CE, Roser K. The Brief Symptom Inventory in the Swiss general population: presentation of norm scores and predictors of psychological distress. PLoS ONE 2024 Jul 1;19(7):e0305192. [doi: 10.1371/journal.pone.0305192]
55. Mohammadkhani P, Dobson KS, Amiri M, Ghafari FH. Psychometric properties of the Brief Symptom Inventory in a sample of recovered Iranian depressed patients. Int J Clin Health Psychol 2010;10(3):541-551 [FREE Full text]
56. Biplob P, Sarker DC, Sarker RC. Assessment of water supply and sanitation facilities for Korail Slum in Dhaka City. IJCEE-IJENS 2013 [FREE Full text]
57. Bashar R, Tommy SS, Ria AF, Khan NA. Assessing the real-life socio-economic scenario of established slums in Dhaka: the cases of Korail and Sattola. Eur Online J Nat Soc Sci 2020 May 21;9(2):455-466 [FREE Full text]
58. Diagnostic and Statistical Manual of Mental Disorders, 5th edition: American Psychiatric Association; 2013. [doi: 10.1176/appi.books.9780890425596]
59. Patel V. Mental health in low- and middle-income countries. Br Med Bull 2007;81-82:81-96. [doi: 10.1093/bmb/ldm010] [Medline: 17470476]
60. Merchant R, Torous J, Rodriguez-Villa E, Naslund JA. Digital technology for management of severe mental disorders in low-income and middle-income countries. Curr Opin Psychiatry 2020 Sep;33(5):501-507. [doi: 10.1097/YCO.0000000000000626] [Medline: 32520747]
61. Javed A, Lee C, Zakaria H, et al. Reducing the stigma of mental health disorders with a focus on low- and middle-income countries. Asian J Psychiatr 2021 Apr;58:102601. [doi: 10.1016/j.ajp.2021.102601] [Medline: 33611083]
62. Alam MF. National mental health survey 2019. : National Institute of Mental Health; 2021 URL: https://nimh.gov.bd/wp-content/uploads/2021/11/Mental-Health-Survey-Report.pdf [accessed 2026-01-24]
63. Shen Y, Zhao W. Knowledge, attitude and practice of urticaria patients towards urticaria medication treatment: a cross-sectional study at the first affiliated hospital of Kunming Medical University. BMJ Open 2024 Nov 27;14(11):e091425. [doi: 10.1136/bmjopen-2024-091425] [Medline: 39609011]
64. Kim HY. Statistical notes for clinical researchers: sample size calculation 1. comparison of two independent sample means. Restor Dent Endod 2016;41(1):74. [doi: 10.5395/rde.2016.41.1.74]
65. Gogtay NJ. Principles of sample size calculation. Indian J Ophthalmol 2010;58(6):517-518. [doi: 10.4103/0301-4738.71692] [Medline: 20952836]
66. Leucht S, Barnes TRE, Kissling W, Engel RR, Correll C, Kane JM. Relapse prevention in schizophrenia with new-generation antipsychotics: a systematic review and exploratory meta-analysis of randomized, controlled trials. Am J Psychiatry 2003 Jul;160(7):1209-1222. [doi: 10.1176/appi.ajp.160.7.1209] [Medline: 12832232]
67. Bergen J, Hunt G, Armitage P, Bashir M. Six-month outcome following a relapse of schizophrenia. Aust N Z J Psychiatry 1998 Dec;32(6):815-822. [doi: 10.3109/00048679809073871]
68. Kishi T, Sakuma K, Hatano M, et al. Relapse and its modifiers in major depressive disorder after antidepressant discontinuation: meta-analysis and meta-regression. Mol Psychiatry 2023 Mar;28(3):974-976. [doi: 10.1038/s41380-022-01920-0]
69. Purnell L, Graham A, Chiu K, Trickey D, Meiser-Stedman R. A systematic review and meta-analysis of PTSD symptoms at mid-treatment during trauma-focused treatment for PTSD. J Anxiety Disord 2024 Oct;107:102925. [doi: 10.1016/j.janxdis.2024.102925] [Medline: 39270371]
70. Jilka S, Siddiqi B, Winsper C, et al. Influences on help-seeking for serious mental illness in Dhaka, Bangladesh: a mixed-methods study. Soc Psychiatry Psychiatr Epidemiol 2025 Oct 31. [doi: 10.1007/s00127-025-03012-0] [Medline: 41174073]
71. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020 Mar 18;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]
72. Aas IM. Global Assessment of Functioning (GAF): properties and frontier of current knowledge. Ann Gen Psychiatry 2010 Dec;9(1):20. [doi: 10.1186/1744-859X-9-20]
73. The World Health Organization Quality of Life assessment (WHOQOL): position paper from the World Health Organization. Soc Sci Med 1995 Nov;41(10):1403-1409. [doi: 10.1016/0277-9536(95)00112-K]
74. Wideman TH, Sullivan MJL, Inada S. Brief Symptom Inventory. In: Encyclopedia of Behavioral Medicine: Springer; 2013:269-270. [doi: 10.1007/978-1-4419-1005-9_3]
75. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull 1987 Jan 1;13(2):261-276. [doi: 10.1093/schbul/13.2.261]

XSL•FO
RenderX

76. Chan SF, La Greca AM. Perceived Stress Scale (PSS). In: Encyclopedia of Behavioral Medicine: Springer; 2020:1646-1648. [doi: 10.1007/978-3-030-39903-0_773]
77. Singh SP, Cooper JE, Fisher HL, et al. Determining the chronology and components of psychosis onset: the Nottingham Onset Schedule (NOS). Schizophr Res 2005 Dec;80(1):117-130. [doi: 10.1016/j.schres.2005.04.018]
78. Sleep questionnaire for adults. : Oxford University Hospitals NHS Trust; 2025 URL: https://www.ouh.nhs.uk/media/vv1jq1qj/sleep-questionnaire-over-11.pdf [accessed 2026-01-24]
79. MacIsaac A, Mushquash AR, Wekerle C. Writing yourself well: dispositional self-reflection moderates the effect of a smartphone app-based journaling intervention on psychological wellbeing across time. Behav Change 2023 Dec;40(4):297-313. [doi: 10.1017/bec.2022.24]
80. Rahman MH, Uddin MA, Ria ZF, Rahman RM. Optimizing BERT for Bengali emotion classification: evaluating knowledge distillation, pruning, and quantization. CMES 2025;142(2):1637-1666. [doi: 10.32604/cmes.2024.058329]
81. Subhash PM, Gupta D, kanjirangat V, CR K. Indo-Aryan dialect identification using deep learning ensemble model. Procedia Comput Sci 2024;235:2886-2896. [doi: 10.1016/j.procs.2024.04.273]
82. Alam S, Ishmam F, Alvee H, et al. BnSentMix: a diverse bengali-english code-mixed dataset for sentiment analysis. arXiv. Preprint posted online on Dec 10, 2024. [doi: 10.48550/arXiv.2408.08964]
83. Cohen A, Naslund J, Lane E, et al. Digital phenotyping data and anomaly detection methods to assess changes in mood and anxiety symptoms across a transdiagnostic clinical sample. Acta Psychiatr Scand 2025 Mar;151(3):388-400. [doi: 10.1111/acps.13712] [Medline: 38807465]
84. Kirkbride JB, Anglin DM, Colman I, et al. The social determinants of mental health and disorder: evidence, prevention and recommendations. World Psychiatry 2024 Feb;23(1):58-90. [doi: 10.1002/wps.21160] [Medline: 38214615]
85. Clausen W, Watanabe-Galloway S, Bill Baerentzen M, Britigan DH. Health literacy among people with serious mental illness. Community Ment Health J 2016 May;52(4):399-405. [doi: 10.1007/s10597-015-9951-8] [Medline: 26443671]
86. Alam NB, Surani M, Das CK, Giacco D, Singh SP, Jilka S. Challenges and standardisation strategies for sensor-based data collection for digital phenotyping. Commun Med (Lond) 2025 Aug 19;5(1):360. [doi: 10.1038/s43856-025-01013-3] [Medline: 40830260]
87. Alam N, Giacco D, Siddiqi B, Singh SP, Jilka S. Investigating awareness and acceptance of digital phenotyping in Dhaka's Korail slum: qualitative study. JMIR Form Res 2025 Jun 23;9:e65530. [doi: 10.2196/65530] [Medline: 40549359]

Abbreviations

BSI: Brief Symptom Inventory
DP: digital phenotyping
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
EDF: effective degrees of freedom
EPP: events per predictor
EWS: early warning sign
GAF: Global Assessment of Functioning
GSI: Global Severity Index
HIC: high-income country
LASSO : least absolute shrinkage and selection operator
LMIC: low- and middle-income country
MDD: major depressive disorder
ML: machine learning
NIHR: National Institute for Health and Care Research
NLP: natural language processing
PANSS: Positive and Negative Syndrome Scale
PIS: participant information sheet
PTSD: posttraumatic stress disorder
SHAP: Shapley Additive Explanations
SMI: serious mental illness
TRANSFORM: Transforming Access to Care for Serious Mental Disorders in Slums
WHOQOL-BREF: World Health Organization Quality of Life-Brief

XSL•FO
RenderX

Edited by J Sarvestan; submitted 29.Jun.2025; peer-reviewed by H Lv, V Vojdanovski; revised version received 11.Dec.2025; accepted 12.Dec.2025; published 04.Feb.2026.

Please cite as:
Alam N, Das CK, Roy N, Giacco D, Singh SP, Jilka S
Using Smartphone-Based Digital Phenotyping to Predict Relapse in Serious Mental Disorders Among Slum Residents in Dhaka, Bangladesh: Protocol for a Machine Learning Study
JMIR Res Protoc 2026;15:e79826
URL: https://www.researchprotocols.org/2026/1/e79826
doi:10.2196/79826

© Nadia Alam, Chayon Kumar Das, Neelabja Roy, Domenico Giacco, Swaran P Singh, Sagar Jilka. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 4.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating Electroencephalogram-Based Predictive Model for Drowsiness Measurement to Reduce Accident Risk in Active Individuals: Protocol for a Preliminary Monocentric Study

Chloé Boitard[1], PhD; Zoé Mazurie[1], PhD; Khadijeh Sadatnejad[1], PhD; Julien Coelho[1,2], MD, PhD; Patricia Sagaspe[1,2], PhD; Julie Lenoir[1], PhD; Julien Mattei[3], MSc; Pierre Berthomier[3], MSc; Marie Brandewinder[4], PhD; Pierre Philip[2], MD, PhD; Jean-Arthur Micoulaud Franchi[1,2], MD, PhD; Christian Berthomier[3], PhD; Jacques Taillard[1], PhD

[1]Université de Bordeaux, SANPSY, UMR 6033, Groupe Hospitalier Pellegrin 13ème Etage – Aile 3 Place Amélie Raba Léon, Bordeaux, France

[2]Centre Hospitalier Universitaire de Bordeaux, Service Universitaire de Médecine du Sommeil, Bordeaux, France

[3]Physip, Paris, France

[4]ENSAM, Paris, France

Corresponding Author:
Jacques Taillard, PhD
Université de Bordeaux, SANPSY, UMR 6033, Groupe Hospitalier Pellegrin 13ème Etage – Aile 3 Place Amélie Raba Léon, Bordeaux, France

Abstract

Background: Voluntary behaviors and socioeconomic factors, such as social jetlag and shift work, can lead to insufficient or disrupted sleep, resulting in drowsiness among active individuals. In occupational and driving contexts, drowsiness poses a serious safety risk by impairing alertness, slowing reaction times, and increasing the likelihood of accidents. Developing automatic and easy-to-implement tools for drowsiness detection or prediction is essential in the management of sleepy patients or in high-risk environments where sustained vigilance is critical.

Objective: This study aims to validate continuous or predictive methods for assessing drowsiness using automated analysis of a limited number of electroencephalogram (EEG) channels.

Methods: Designed as a single-center, nonrandomized, single-group study, this investigation will evaluate drowsiness and cognitive performance in 40 healthy volunteers exposed to 2 sleep deprivation conditions simulating real-world occupational scenarios. The primary outcome will be the Objective Sleepiness Scale (OSS) and its automated analysis, with a focus on its ability to measure objective wakefulness as assessed by the maintenance of wakefulness test (MWT). Secondary outcomes will include multimodal resting-state EEG markers, subjective and objective sleepiness measures, performance on a simulated driving task, attention, executive function, and vigilance assessments, as well as sleep quality, sleep quantity, and mind-wandering. The influence of sociodemographic and clinical variables on the measurement and prediction of drowsiness will also be systematically examined.

Results: This study received funding from Physip and ANR (Agence Nationale de la Recherche, National Research Agency) in 2019, with ethical committee (Comité de Protection des Personnes, Committee for the Protection of Persons) approval in May 2022. Recruitment began in March 2023 and was completed in May 2025, with a database lock in June 2025. Data analysis started in June 2025 and is still ongoing.

Conclusions: By validating these novel EEG-based measures, this study aims to lay the groundwork for proactive strategies for drowsiness management in occupational, transportation, and clinical settings.

Trial Registration: ClinicalTrials.gov NCT05453643; https://clinicaltrials.gov/study/NCT05453643

International Registered Report Identifier (IRRID): DERR1-10.2196/83969

(JMIR Res Protoc 2026;15:e83969) doi:10.2196/83969

KEYWORDS

cognitive performance; drowsiness; EEG; Objective Sleepiness Scale; prediction; sleep deprivation; electroencephalogram

XSL•FO
RenderX

Introduction

Sleepiness is a physiological and behavioral "need state" or "need for sleep." It plays a key role in the regulation of the sleep/wake cycle, especially in triggering sleep onset at the usual bedtime or during sleep deprivation. Sleepiness facilitates the transition from wakefulness to sleep and hinders the transition from sleep to wakefulness. "Physiological sleepiness" [1], also known as sleep drive, results from an imbalance between processes involved in the regulation of sleep and wake states. It is orchestrated both by the sleep homeostatic process, relying on the prior duration of wakefulness and the prior amount of sleep, and by the circadian rhythm, inducing a forbidden sleep zone at the end of the evening and an imposed sleep zone at the end of the night. Physiological sleepiness is modulated by interindividual differences (chronotype, age, and sex).

Manifest sleepiness [1] is the transformational effect of underlying physiological sleepiness on behavior due to voluntary or involuntary sleep deprivation, circadian misalignment, or pathologies inducing excessive daytime sleepiness. Manifest sleepiness is also referred to as drowsiness, hypoarousal [2], or continuous nonimperative sleepiness [3] by other authors. We will here use the generic term "drowsiness" to qualify manifest sleepiness. Drowsiness is characterized by an inappropriate waking behavior and a lasting inadequate level of arousal, which results in an inability to stay awake, difficulties in maintaining sustained attention and vigilance (ie, brain fog), and affects judgment and decision-making abilities and encourages risk-taking [4]. Drowsiness can thus induce functional consequences in terms of both disability and accident risk [5-7], to the same extent as the effects of alcohol consumption [8]. Drowsiness also redirects attention to thoughts that are not related to the task at hand [9]. This modification in attentional orientation, known as mind wandering, also impacts cognitive performance [9,10]. Finally, introspective sleepiness [1] concerns the individual's self-assessment of their internal state of sleepiness. Manifest and introspective measures are linked because they both reflect underlying sleep pressure and drowsiness and therefore may arise from the same underlying drive state [11]. However, drowsy people often fail to accurately assess their level of drowsiness and therefore overestimate their ability to make important decisions and perform complex tasks [12], despite their decreased cognitive abilities, which can lead to dangerous consequences. Identifying and measuring real-time drowsiness or predicting it in high-risk situations like working or driving is thus a significant public health challenge [13,14].

Drowsiness can result from voluntary behaviors or socioeconomic factors that lead to insufficient or disrupted sleep [15], such as social jetlag [16] and shift work [17]. Drowsiness assessment over time is mainly based on brain activity measurements (electroencephalogram [EEG]) [18]. EEG provides a direct and objective window into neural correlates of alertness, unlike subjective measures or sporadic performance-based assessments. EEG data can reveal subtle transitions between wakefulness and drowsiness that may otherwise go undetected. Thus, the primary index of drowsiness, that is, the ability to stay awake, is assessed during the maintenance of wakefulness test (MWT) [19], an iterative electrophysiological test in which EEG and electrooculogram (EOG) analysis allow the determination of daytime sleep latency, identifying "global" daytime drowsiness. Moreover, the behavioral consequences of drowsiness —that is, impaired cognitive performance and reduced vigilance, characterized by slowed reaction times and an increased rate of commission errors (incorrect responses made when no response is required)—are associated with changes in EEG recordings, including episodic microsleep intrusions, local sleep phenomena, and wake-state instability [4,18]. Thus, EEG analysis, considered the "gold standard" for direct and continuous monitoring of sleepiness [18], represents the most promising way to detect drowsiness in safety-sensitive contexts, such as transportation, health care, and industrial operations, where lapses in vigilance can have serious consequences.

Given the need to provide continuous measurement to combat drowsiness in the workplace, the earliest EEG-based methods developed to continuously assess spontaneous drowsiness in offline conditions include the Karolinska Drowsiness Score [20]. This approach is designed to evaluate sleepiness-related impairments in performance, particularly in driving, by quantifying theta and alpha EEG activity and/or identifying specific slow rolling eye movements via EOG, resulting in a continuous sleepiness score ranging from 0 (not sleepy) to 100 (very sleepy). Currently, numerous EEG-based systems and algorithms—both offline and online—have been proposed for the detection of drowsiness. These methods often rely on various frequency band ratios, including theta/beta, beta/(alpha+beta), theta/(alpha+beta), (theta+alpha)/beta, (theta+alpha)/(alpha+beta), and (gamma+beta)/(sigma+alpha) [14,21-23]. To enhance classification accuracy, some systems integrate additional physiological signals, such as electrocardiogram (ECG), EOG, electromyogram (EMG), respiratory activity, or behavioral data [14,23,24]. However, the practical application of these systems remains limited. Most have been validated only under laboratory conditions, typically during early afternoon sessions, whether after sleep deprivation or not, and their deployment in real-world settings is constrained by the large number of sensors required. To ensure operational feasibility, future development of EEG-based drowsiness detection systems should prioritize the use of a minimal number of sensors, particularly EEG derivations.

The Objective Sleepiness Scale (OSS), developed by Muzet et al [25], was designed to continuously monitor drowsiness and classify sleepiness into 5 levels based on the duration of alpha and/or theta EEG activity and the presence of slow eye movements, assessed over 20-second epochs. Despite its methodological clarity, the OSS is not widely adopted in subsequent research. However, a recent study has reported significant associations between OSS scores and sleepiness-related variables, such as sustained attention and driving performance [26], confirming the potential of the OSS to provide a reliable assessment of sleepiness-related cognitive impairments. The original implementation of the OSS required 4 EEG and 4 EOG derivations, which limited its applicability in operational settings. To address this, the French company Physip, in collaboration with Muzet et al [25], developed

XSL•FO
RenderX

MEEGAWAKE, an algorithm for detecting and assessing drowsiness. Based on the OSS, the algorithm is capable of classifying the 5 stages of sleepiness, relying exclusively on EEG signals recorded from only 2 derivations. The performance of this algorithm was validated on 15 participants in a driving simulator by comparison with visual analysis conducted by a drowsiness reference expert [27]. This streamlined approach facilitates easier integration into workplace environments.

Nevertheless, current real-time spontaneous drowsiness detection systems continue to face several challenges. First, they typically require continuous wear by the operator, which complicates their implementation in real operational settings. Second, the detection of drowsiness often occurs too late to prevent its functional consequences. Therefore, the development of EEG-based tools allowing the prediction of drowsiness and its behavioral consequences—rather than detecting its presence—would mark a significant advancement in ensuring operator safety. In experimental protocols, routine EEG bio-calibrations, performed with eyes open and eyes closed, are critical for assessing signal quality and characterizing an individual's resting-state brain activity. Resting-state EEG recordings, under both eyes-opened and eyes-closed conditions, have been used to evaluate task performance and drowsiness levels [28-31]. An algorithm developed by the University of Leipzig (VIGALL, The Vigilance Algorithm Leipzig) can identify 7 states of somnolence during a 5-minute resting-state EEG but requires the recording of at least 19 EEG leads (25 recommended), making it unusable in real operating conditions [32]. However, these short measures offer a promising avenue for prediction. Indeed, a brief, iterative, and easy-to-administer resting-state assessment could enable early identification of drowsiness and its potential cognitive and behavioral impacts, thereby facilitating timely and effective preventive interventions.

The objective of this study is to validate—under sleep deprivation conditions reflecting real-life occupational scenarios (such as delayed, fragmented, or disturbed sleep, typically experienced during night work, shift work, or on-call duty)—the ability of the OSS and its automatic analysis by the MEEGAWAKE algorithm to accurately assess drowsiness and its associated cognitive and behavioral consequences. In addition, the study aims to evaluate whether automatic analysis of resting-state EEG can reliably predict drowsiness and its functional outcomes, independently of the duration of prior wakefulness. Such predictive capability could enhance risk assessment and contribute to the prevention of sleepiness-related accidents in operational settings.

We hypothesize that both the OSS and resting-state EEG, potentially adjusted for interindividual variability in sleepiness susceptibility, can effectively measure and/or predict drowsiness and its behavioral consequences, including the inability to maintain wakefulness and impairments in sustained attention and vigilance, regardless of the time of day. This approach may provide a robust framework for identifying and mitigating risks associated with drowsiness in safety-critical environments.

Methods

Objectives

This study has 2 main objectives. This study's first aim is to determine whether the OSS criteria can accurately detect manifest sleepiness, compared to the MWT, and to evaluate its temporal accuracy in detecting drowsiness-induced momentary behavioral outcomes, including driving performance, vigilance, and sustained and selective attention.

Second, this study aims to validate robust models based on a multimodal EEG index, derived from resting-state activity during bio-calibration, for predicting functional impairments linked to drowsiness, particularly in driving performance, vigilance, sustained, and selective attention.

Study Design

This research is designed as a preliminary investigation with a single-center, nonrandomized, single-group study, conducted at the Sleep Department of the SANPSY laboratory, Bordeaux, France.

Participants

Sample Size Calculation

This study is exploratory, since, to our knowledge, no other study has used the OSS to predict manifest sleepiness. Thus, because no data would allow us to statistically determine the number of participants required, this number has been arbitrarily set to 40 healthy volunteers. We believe this number will facilitate preliminary exploration at both inter- and intraindividual levels with an appropriate protocol.

Recruitment

Candidates are 40 healthy volunteers recruited from Bordeaux University Hospital's healthy volunteer database and through advertisement via information flyers and internet recruitment. Participants were screened with respect to health status by a physician and filled out specific questionnaires with professional clinicians. The candidates were informed that they would receive financial compensation of €900 (US $1062.02) for participating in the whole protocol, and the physician provided full information about the study (design, potential risks and constraints, answers to potential questions) before collecting the volunteers' informed, written, dated, and signed consent.

The eligibility criteria include the following: individuals (1) aged between 20 and 60 years, with a BMI between 18 and 27, (2) with good French skills and the ability to understand the study, and (3) who are nonprofessional drivers with a valid driver's license (obtained at least one year ago). Individuals (1) experiencing severe psychiatric, neurological, or medical pathology or under psychotropic or cardiotropic drug treatments, (2) afflicted by chronic insomnia disorder, severe diurnal somnolence, or sleeping pathologies that can induce excessive daytime somnolence, (3) declaring substance dependency, alcohol abuse (>2 glasses/day) and/or excessive consumption of coffee, tea, or caffeine-based drinks (such as Coke, >5 cups/day), and (4) who perform night or shift work or being on

XSL•FO
RenderX

care or on-call duty during the last 72 hours before the experimental sessions were excluded from the study.

Once included in the study, participants completed the following self-administered questionnaires, addressing specific items that serve as exclusion criteria for this study, namely the presence of depression and/or anxiety disorders, daytime sleepiness, and sleep-related disorders, such as obstructive sleep apnea syndrome (OSAS) and restless legs syndrome (RLS).

PHQ-4 (Patient Health Questionnaire) is a validated and reliable tool designed to reveal propensity for anxiety and depression [33]. Participants are asked to score 4 items reflecting their tendency to feel anxious (2 items) or depressed (2 items) during the previous 2 weeks (on a Likert scale ranging from 0 (never) to 3 (every day). A score above 2 on any item or a total score above 4 on the anxiety or depression items will lead to the participant's exclusion.

ESS (Epworth Sleepiness Scale) is a common measure used to assess daytime sleepiness [34]. Candidates must score their ability to fall asleep in 8 different daily situations over the past months (on a Likert scale ranging from 0 (never) to 3 (very high probability). A score over 11, reflecting a severe diurnal somnolence, will lead to the participant's exclusion.

STOP-BANG (snoring, tiredness, observed apnea, blood pressure, BMI, age, neck circumference, and gender) questionnaire evaluates the risk of OSAS through several items [35]. A score over 5, reflecting a high probability of being diagnosed with OSAS, will lead to the participant's exclusion.

The RLS questionnaire assesses the risk of being diagnosed with this syndrome through 7 specific binary questions (yes or no answer) [36]. A positive answer to some items, reflecting a restless legs syndrome suspicion, will lead to the participant's exclusion.

Finally, before the selected candidates entered experimental procedures, additional questionnaires were used to assess their chronotype and tendency to daydream (Table 1)(Table 1).

The chronotype was assessed using the 19-question Horne and Östberg Morningness/Eveningness questionnaire, which investigates life preferences (activity, wake/sleep cycle, and meal), somnolence, and tiredness at certain times of the day [37]. Scores vary from 16 to 86 and, with adaptation to the individual's age [38], allow for the identification of evening persons (score ≤42 for those aged 20 - 44 years and ≤53 for participants aged 44 - 60 years), morning persons (score ≥58 and≥64, respectively) and extreme chronotypes (score ≤31 and ≤47 respectively, indicating highly evening persons, and score ≥69, indicating highly morning persons regardless of age). The French version [39] of the Munich Chronotype Questionnaire [40] was also used. It is composed of 7 questions investigating sleep habits on working days and 7 questions investigating sleep habits on free days. The sleep-corrected local time of midsleep on work-free days (MSFsc) and midsleep on working days (MSWsc) [40] will be identified, and the sleep-corrected social jetlag will be calculated (MSFsc – MSWsc) [41].

The candidates' tendency to daydream in everyday situations was assessed using the Daydreaming Frequency Scale (DDFS) [42]. Participants are asked to rate their tendency to daydream during 12 common situations on a 5-point Likert scale (from 1 [very rare] to 5 [very frequent]).

After verification of inclusion and exclusion criteria, participants went through the experimental schedule, which was composed of 2 experimental sessions. Both sessions were designed to increase the sleep pressure through alteration in sleep quantity, while mimicking real-life work experiences. Each candidate's participation lasted 14-35 days, according to the time lapse between the 2 experimental sessions (refer to Table 1). All premature withdrawals (eg, withdrawal of consent, inability to complete the protocol for medical or logistical reasons) will be documented, including date, reason, and any data collected up to the point of discontinuation. Unless a participant requests complete removal of their data, baseline characteristics and available measurements will be retained for analysis. The study aims to include 40 participants completing the full protocol. Replacement of withdrawals may be performed, if necessary, but will be limited to avoid overselection of highly motivated individuals (will not exceed 20% of the initial target sample size). Baseline characteristics of completers and noncompleters will be summarized and compared, along with a brief description of the reasons for withdrawal. Exploratory analyses may assess whether baseline factors predict early withdrawal. Primary analyses will be conducted on participants who complete the full protocol (per-protocol analysis). When feasible, data from partial completers will be included in exploratory models to maximize the use of available information. Mixed-effects models or other appropriate statistical approaches will handle repeated measures and missing data due to withdrawal. Participant flow (screened, enrolled, completed, and withdrawn) will be reported in a CONSORT (Consolidated Standards of Reporting Trials)-style diagram, and all analyses related to dropouts will be documented.

XSL•FO
RenderX

Table . Schedule of participants' enrollment and assessments.

	Study period				
	Interview	Enrollment	Experimental schedule		Close-out
Timepoint	T−1 to −15 days	T0	Session 1: T+1 to T+14 days	Session 2: T+7 to T+30 days	T+7 to T+30 days
Recruitment					
Information about the study	✓				
Eligibility screen	✓	✓			
Informed consent		✓			
Clinical examination		✓			
Inclusion self-questionnaires (PHQ-4[a], ESS[b], STOP-BANG[c], RLS)[d]		✓			
Individual characteristics (chronotype, daydreaming)		✓			
Assessments					
Actimetry			✓	✓	
Polysomnography			✓	✓	
Continuous wake EEG[e] recordings (OSS[f], resting state EEG)			✓	✓	
Karolinska Sleepiness Scale			✓	✓	
Driving simulation			✓	✓	
Maintenance of wakefulness test			✓	✓	
Cognitive tests			✓	✓	
Conscient Experience Characterization			✓	✓	

[a]PHQ-4: Patient Health Questionnaire-4.

[b]ESS: Epworth Sleepiness Scale.

[c]STOP-BANG: snoring, tiredness, observed apnea, blood pressure, BMI, age, neck circumference, and gender.

[d]RLS: Restless Legs Syndrome.

[e]EEG: electroencephalogram.

[f]OSS: Objective Sleepiness Scale.

Experimental Timeline

The experimental schedule was composed of 2 experimental sessions, both designed to increase sleep pressure by altering sleep quantity prior to behavioral testing. Each candidate participated in both sessions in the same order. The first session lasted 3 days and was structured with a fragmented sleep schedule to simulate nocturnal on-call duty. The second session lasted 4 days with 2 full nights of sleep deprivation, separated by a 4-hour recovery period, to simulate a night shift (Figure 1). As shown in Figure 1, gray rectangles indicate periods of recorded sleep, white rectangles correspond to wake periods, hatched rectangles indicate night sleep deprivation.

Several behavioral tests, with established validity for assessing sleepiness and its impact on cognitive performance (manifest sleepiness), described in the following section, were conducted during both sessions to assess the effects of sleep alteration on introspective sleepiness, on the capacity to remain awake (reflecting sleep pressure), on driving performance, on attention and vigilance, and on the mind-wandering experienced during the tests (Figure 2). This figure shows colored arrows, which indicate, before and/or after each test, the registration of the resting state EEG (bio-calibration lasting 2 minutes [BC2], red arrows) and the scoring by the participant of their introspective sleepiness level (KSS, blue arrows) and of their mind wandering during the test (CEC, yellow arrows). The pause from 25-85 minutes allows lunch or dinner to be served to the participants during the relevant blocks (starting at noon and 8 PM). Colored rectangles correspond to the different tests composing the 4-hour test block.

EEG was recorded throughout the entire experimental protocol, during the rest periods to assess sleep duration and quality in both ambulatory and laboratory conditions, and throughout the tests to allow for drowsiness assessment (through OSS visual and automatic scoring). Resting state EEG was recorded during regular bio-calibrations (BC2 or BC5 lasting 2 or 5 minutes, respectively).

XSL•FO
RenderX

Figure 1. Study experimental workflow. (A) Experimental session 1 designed to create sleep deprivation conditions that simulate on-call duty. (B) Experimental session 2 designed to produce sleep deprivation conditions simulating 2 successive working nights.

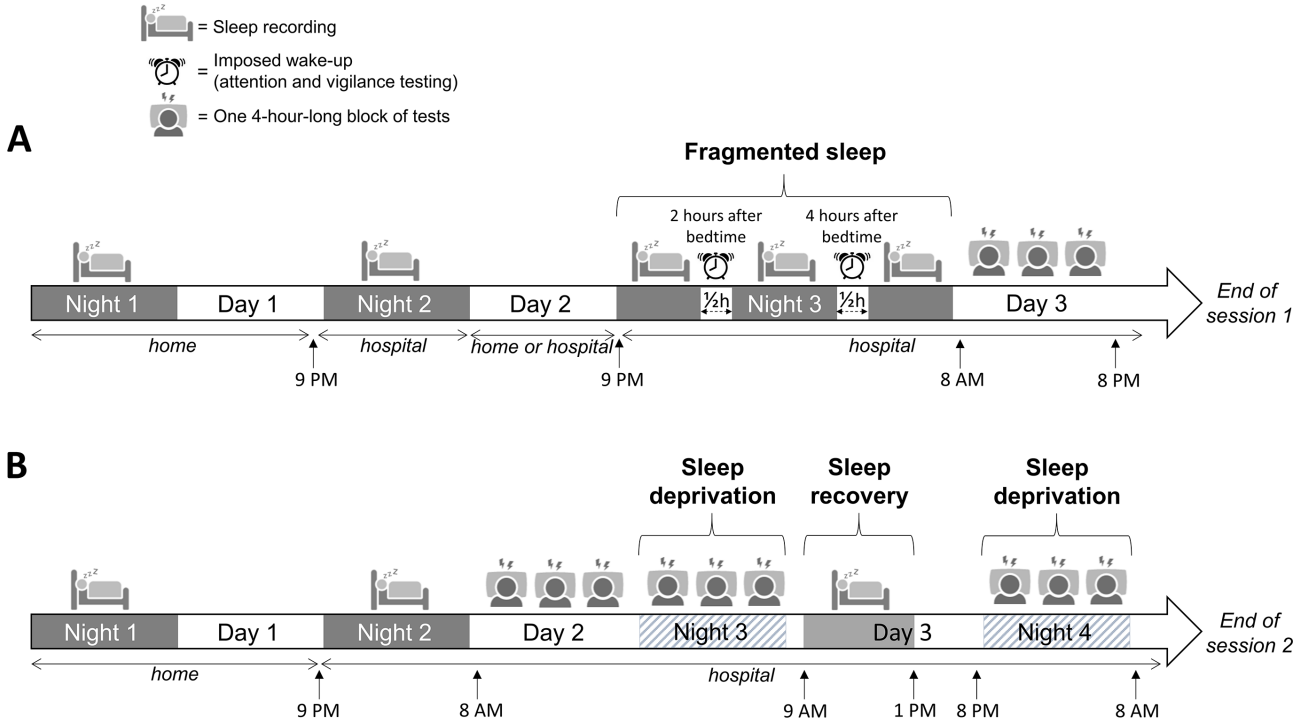

Figure 2. Timeline of one block of tests and questionnaires. CPT: continuous performance test; CEC: Conscient Experience Characterization; EEG: electroencephalogram; KSS: Karolinska Sleepiness Scale; MWT: maintenance of wakefulness test; PVT: psychomotor vigilance task; TAP: Test of Attentional Performance; BC5: 5-minute bio-calibration.

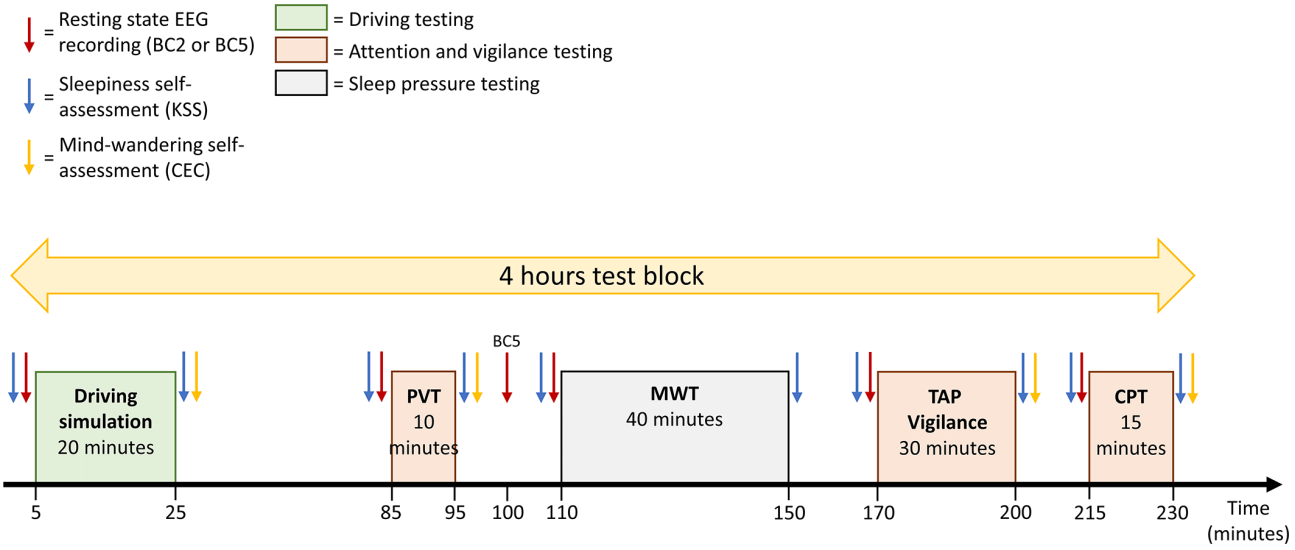

Sleep and Wake Monitoring Setup

Sleep parameters were assessed using ambulatory and laboratory polysomnography coupled with actigraphy. Activity was monitored using a wrist-worn accelerometer (GT9X Link, ActiGraph). Ambulatory polysomnography was performed using an EMBLA Titanium system (Natus; sampling frequency 256 Hz; 16-bit resolution). The montage included 2 EEG channels (Cz and Pz). Laboratory polysomnography was recorded with an Embla NDx amplifier and stored using Natus SleepWorks software (sampling frequency 256 Hz; 16-bit resolution). The montage included at least 2 EOG channels, 1 EMG channel, 6 EEG channels (F4, C4, O2, Cz, Pz, and A1), and 1 ECG channel.

While awake, participants were equipped with a gel-based active electrode system for EEG (Acticap Slim, Brain Products) connected to a compact wireless EEG amplifier LiveAmp (Brain Product: 24-bit resolution). The montage included 8 EEG derivations, namely Fz, Cz, Pz, C4, O2, C3, O1, and P3, and 4 vertical and horizontal EOG. All signals were acquired using a sampling frequency of 500 Hz.

Description of the Tests

Introspective sleepiness was investigated through the Karolinska Sleepiness Scale (KSS), a subjective scale assessing the level of sleepiness during the past 10 minutes [43]. Participants are required to rate their ability to maintain wakefulness on a Likert scale (from 1 = “perfectly awake” to 9 = “severely somnolent,

XSL•FO
RenderX

cannot stay awake"). This score is known to correlate well with EEG measurements and cognitive and behavioral performances linked to the candidate's ability to maintain wakefulness [44].

The ability to stay awake under sleep-promoting conditions was monitored through the MWT, which is considered the gold standard in the measurement of manifest sleepiness [18,45]. The MWT is a 40-minute trial during which the participant is in a comfortable position in bed, in a quiet, dimly lit room. At the beginning of each trial, the candidate is instructed to stay awake and combat against sleep, to keep their eyes open, and to look straight ahead, without looking at the light. During the test, visual analysis of the polysomnographic recording allows for monitoring the participant's state, and the test is stopped when unequivocal sleep is detected (defined by 3 consecutive epochs of stage N1 sleep or 1 epoch of any other sleep stage, according to the American Academy of Sleep Medicine (AASM) recommendations [46]).

Driving performance was assessed during a driving simulation procedure, as it provides an ecological approach that is commonly used to assess the effects of sleepiness on complex cognitive performance and safety, using the fixed-base INRETS-MSIS SIM2 (Institut National de Recherche sur les Transports et leur Sécurité — Modélisations, Simulations et Systèmes d'Information pour la Sécurité — Simulateur version 2; National Institute for Transport and Safety Research — Modeling, Simulation, and Information Systems for Safety — Simulator version 2) driving simulator [47]. In healthy individuals, this simulator has been shown to effectively assess nocturnal driving impairment in a dose-response design, considering extended wakefulness and driving duration, compared to real driving conditions [48]. The driving simulator is composed of a computer and a video game steering wheel with no force feedback applied. The participant's head is positioned at 60 cm in front of the screen. The resolution of the visual scene was 1024×768 pixels, and the update rate was 60 Hz. The simulator generates a highway driving scenario on a 19-inch screen. The car's speed was fixed by the experimenter at 130 km/h, and participants were instructed to drive in the center of the right lane. The test lasts for 20 minutes, during which the participant must maintain the car's trajectory on a monotonous 2-lane highway (reconstruction of the French A62 highway, without any scenery or vehicles).

To evaluate the impact of sleepiness on cognitive performance of increasing complexity, we included 3 tasks ranging from basic attentional measures to more demanding cognitive functions. First, the participants completed the Psychomotor Vigilance Task (PVT), a widely used and validated tool for quantifying manifest sleepiness, particularly through its effects especially on sustained attention and somewhat less on vigilance [18]. The test consists of a one-choice serial reaction time task and was conducted using PC-PVT software (Biotechnology HPC Software Applications Institute), a freely available package for PVT testing, analyzing, and visualizing on a PC, whose performance is comparable to the PVT-192, the gold standard for measuring PVT [49]. The PC-PVT software [50] was installed on a desktop computer running under Windows 10 and equipped with a mouse. The test session lasts 10 minutes, during which a visual stimulus randomly appears on a screen, with interstimulus intervals varying from 2 to 10 seconds. The participant is asked to react as quickly as possible (by a mouse click) to make the stimulus disappear.

The Conners continuous performance test II (CPT-II) is a Go/No-Go task that evaluates the participant's ability to maintain sustained attention and to discriminate relevant stimuli [51]. The test assesses inhibitory control, which is known to be highly sensitive to excessive daytime sleepiness [52-54]. The CPT (continuous performance test) paradigm is based on monitoring and responding to regular targets (here, all alphabet letters except the "X"), while withholding the response to a specific infrequent nontarget (letter "X"). CPT software (Multi-Health Systems Inc, version 5.2) was installed on a desktop computer running under Windows 10 and equipped with a keyboard. The test session lasts for 15 minutes, and the participant is asked to push the space bar when any letter is displayed on the screen, except for the letter "X" (target stimulus).

The Vigilance Test of Attentional Performance (TAP) is a rarely used animated line test, which measures the ability to focus and to maintain mental effort and vigilance on a monotonous task during a 30-minute period [55]. The test displays a horizontal line going up and down on a screen at an irregular speed, and from time to time, this line oscillates with a greater amplitude. TAP Vigilance (Psytest) software was installed on a desktop computer running under Windows 10 and equipped with a specific external response button. The participant is asked to push the response button when they notice a difference in the line oscillation amplitude (target).

Finally, mind-wandering during completion of the test was assessed through a subjective scale, the Conscient Experience Characterization (CEC) [56]. At the end of the cognitive tests, participants were asked to rate their focusing level during the previous task, using a scale of five options: (1) fully focused (attention and thoughts entirely dedicated to the task), (2) focus interference related to the task (attention and thoughts distracted by task characteristics or by their performance), (3) external distraction (attention and thoughts distracted by environmental stimuli with no relation to the task), (4) mind wandering (not focused, with thoughts disconnected from the task or the environment), and (5) empty mind (not focused, without any particular thoughts).

Experimental Session 1: Fragmented Sleep

This session lasted for 3 consecutive nights and days, during which sleep was recorded. In order to induce sleep pressure on day 3, the third night was fragmented (with 2 imposed wake-ups, during which attentional and vigilance tests were carried out). Three blocks of behavioral tests were conducted during the third day (Figure 1A). One block of tests is 4 hours long and includes, in a fixed order, a driving simulation, cognitive tests, evaluation of sleep pressure, and self-assessments of drowsiness and mind-wandering, while continuous EEG is recorded (Figure 2).

The detailed schedule of days and nights is organized as follows:

1. Night 1: In the evening before the first night, the participants arrived at the laboratory to be equipped with Ag-AgCl electrodes (2 EEG derivations: Cz and Pz, connected to the

XSL•FO
RenderX

EMBLA Titanium) to measure their sleep quantity and quality. They returned home for the first night, went to sleep, and woke up according to their usual schedule.
2. Day 1: After waking up and removing the sleep recorder, the participants carried out their usual activities throughout the day before returning to the laboratory in the evening, where training for the cognitive tests and driving simulation was provided.
3. Night 2: The participants were equipped with Ag-AgCl electrodes (6 EEG derivations: F4, C4, O2, Cz, Pz, A1, 1 EMG, 2 EOG, and 1 ECG) connected to the Embla NDx amplifier to measure their sleep quantity and quality. They spent the night in the laboratory, went to sleep, and woke up according to their usual schedule.
4. Day 2: After waking up, the participants either stayed in the laboratory or went to their regular activities before returning in the evening.
5. Night 3: The participants were re-equipped with polysomnographic sensors (as during night 2) and went to bed by midnight. They were awakened twice during the night (2 hours and 4 hours after bedtime) for 30 minutes, during which they underwent the PVT and CPT tests to measure their vigilance along with their ability to maintain sustained attention.
6. Day 3: After waking up at 7 AM, the equipment was removed, and the participants were equipped with a gel-based active electrode system for EEG. The participants underwent 3 test blocks. Block 1 occurred from 8 AM to noon, block 2 from noon to 4 PM, and block 3 from 4 PM to 8 PM After the last block, the EEG helmet was removed, and the participants either stayed in the laboratory for a recovery night or left the hospital, if accompanied, with a recommendation to go home for a recovery night.

Experimental Session 2: Sleep Deprivation

This session lasted for 4 consecutive nights and 3 days, during which sleep was recorded. During the second day, the participants began a series of six 4-hour test blocks, lasting 24 hours and leading to a total sleep deprivation during the third night. On the third day, the participants had a morning recovery sleep period (4 hours) and began a new night of total sleep deprivation with a series of three 4-hour test blocks during the fourth night (Figure 1).

The detailed schedule of days and nights is organized as follows:

1. Night 1: In the evening before the first night, the participants entered the laboratory to be equipped with Ag-AgCl electrodes (2 EEG derivations: Cz and Pz) connected to the EMBLA Titanium to measure their sleep quantity and quality. They returned home for the first night, went to sleep, and woke up according to their usual schedule.
2. Day 1: After waking up and removing the sleep recorder, participants carried out their usual activities throughout the day before returning to the laboratory in the evening.
3. Night 2: The participants were equipped with Ag-AgCl electrodes (6 EEG derivations: F4, C4, O2, Cz, Pz, A1, 1 EMG, 2 EOG, and 1 ECG) connected to the Embla NDx amplifier to measure their sleep quantity and quality. They spent the night in the laboratory and went to bed by midnight.
4. Day 2 and Night 3: After waking up, the equipment was removed, and the participants had breakfast before being equipped with a gel-based active electrode system for EEG. Then, the participants completed six 4-hour test blocks. Block 1 occurred from 8 AM to noon, block 2 from noon to 4 PM, block 3 from 4 PM to 8 PM, block 4 from 8 PM to midnight, block 5 from midnight to 4 AM, and block 6 from 4 AM to 8 AM. After the last block, the EEG helmet was removed.
5. Day 3: The participants underwent a morning recovery sleep period (from 9 AM to 1 PM) during which they were equipped with the Ag-AgCl electrodes and sleep recorder, as described for night 2, to measure their sleep quantity and quality. After waking up, sensors were removed, and the participants remained in the laboratory without sleeping during the afternoon.
6. Night 4: The participants were re-equipped with the gel-based active electrode system for EEG and wireless EEG amplifier, as described for Day 2 and Night 3, and underwent a new series of three 4-hour test blocks. Block 1 occurred from 8 PM to midnight, block 2 from midnight to 4 AM, and block 3 from 4 AM to 8 AM After the last block, the EEG helmet was removed.
7. Day 4: The participant either stayed in the laboratory for a 5-hour recovery sleep or left the hospital, if accompanied, with a recommendation to plan a sleep recovery night at home.

Protocol Assistant

Protocol Assistant is a software tool developed by Physip to manage the timeline of the various behavioral tests and questionnaires during the 4-hour block and to ensure synchronization with the EEG and EOG acquisition performed by a Brain Products amplifier (ie, LiveAmp). It is designed to provide event markers (start and end of each test) with a time precision of approximately 0.1 seconds, without requiring any programming skills. Additionally, it allows for the collection and timestamped recordings of responses to self-assessment questionnaires (KSS and CEC) and any comments the experimenter may make during the experiment.

After a preliminary configuration phase, during which the timeline for tests and questionnaires is set once and for the 4-hour block, the software has to be launched for each acquisition. It opens a graphical interface with windows on 2 separate screens (one for the experimenter and one for the participant), enabling the protocol to be run and the test and questionnaire instructions to be presented to the participant, in written and audio form, while managing the acquisition of EEG signals realized by the EEG amplifier.

The software automatically adds time markers on the fly to the protocol steps and enables the experimenter to follow the progress of the protocol via a graphical timeline and via the list of steps, to visually monitor the tracings in real time, insert customizable predefined event markers, and interrupt or resume the progress of the protocol.

XSL•FO
RenderX

Outcome Measures

Primary Outcome Measure

The primary outcome measure corresponds to the OSS score, which relies on the analysis of electrophysiological variables to assess the instantaneous drowsiness state, corresponding to the participant's level of sleepiness. The OSS score, ranging from 0 (full alertness) to 4 (very drowsy), is assigned every 20 seconds, based on specific EEG activity (amount of beta, alpha, and theta waves) observed simultaneously in 2 regions of the brain during each epoch, and accompanied by distinct eye blinks and movements (normal or slow). An OSS score of 0 corresponds to continuous beta activity, with no alpha or theta activity and no slow eye movements. In contrast, an OSS score of 4 is characterized by the presence of alpha and/or theta rhythms for more than 10 cumulative seconds, associated with slow eye movements.

The OSS was scored both manually, by a trained experimenter—based on EEG (Fz, Cz, Pz, C4, O2, C3, O1, and P3) and EOG (vertical and horizontal) recordings—and automatically, by the MEEGAWAKE algorithm—based on EEG (Cz and Pz)— every 20 seconds throughout all behavioral tests, to reflect the spontaneous drowsiness states.

The MEEGAWAKE algorithm uses data-driven criteria to cope with interindividual variability. After an initial artifact rejection step relying on detecting both temporal and frequential abrupt variations or nonphysiological behavior, the analysis step aims to determine recording-specific thresholds based on EEG power ratios in the usual frequency bands. In addition to this frequency analysis, a temporal localization of EEG events is performed to analyze burst activity. Based on the results of this analysis, the final step is to classify the 20-second epochs into OSS drowsiness states.

During MWT, sleep onset is defined as the start of the first epoch scored 3 or 4. OSS sleep latency (OSS latency) is defined as the time from lights out until the start of the first epoch scored as 3 or 4 (expressed in minutes).

Secondary Outcome Measures

The secondary outcome measures are the following: resting state EEG, longer resting state EEG, objective sleep quantity and quality, objective measurement of manifest sleepiness, driving performance, cognitive performance, and subjective sleepiness and mind wandering.

Resting state EEG, recorded through bio-calibrations, is a simple procedure in which participants are placed in a quiet room and asked to relax while looking at a picture on the wall, to keep their eyes open for 1 minute, refraining from blinking as much as possible, and then close their eyes for 1 minute, without moving (BC2). EEG and EOG channels are recorded continuously during bio-calibrations. In the preprocessing step of the signal analysis, standard signal processing techniques will be used, including notch filtering to remove power-line interference and bandpass filtering to isolate specific frequency subbands. The EEG data will be segmented into 4-second epochs to minimize the effects of nonstationarity in longer signal segments. Spectral power across different EEG frequency subbands and the frequency centroid will be computed for each channel under both eyes-open and eyes-closed conditions. To mitigate the influence of outliers, the median spectral power across epochs within each session will be used. EEG signal processing will be performed using the Python (Python Software Foundation) MNE toolbox (magnetoencephalography and electroencephalography in Python). Following preprocessing, feature vectors representing each session will be constructed, comprising multiple EEG-derived features. These feature vectors will then serve as inputs for a machine learning model to identify the most informative features for distinguishing participants' levels of drowsiness, cognitive, and driving performances as quantified by the different tests conducted after the bio-calibrations.

Longer resting state EEG, recorded through a longer bio-calibration, is a procedure that remains the same as described for BC2, but here the participants are asked to keep their eyes open for 5 minutes (BC5). During the 5 minutes with eyes open, the Fz signal is analyzed in 2-second epochs, each characterized by features including the signal magnitude and spectral power (Fast Fourier Transform, Hanning window) in the frequency bands delta (0 - 4 Hz), theta (4 - 8 Hz), alpha (8 - 12 Hz), sigma (12 - 16 Hz), and beta (16 - 50 Hz). After automatic artifact and blink removal using both temporal and frequential criteria, the spectral power of the artifact-free Fz signal will then be computed in the 6 - 9 Hz band [57] for each 2-second artifact-free epoch. The resulting distributions of spectral power values will then be characterized by their mean, SD, and median. The mean spectral power will be the result of BC5. To compare values between individuals while controlling for interindividual variability, a participant's test values are normalized by their median.

Objective sleep quantity and quality, measured through polysomnography recordings and analyzed by ASEEGA software (Physip), are based on an automated analysis of a single EEG channel (Cz-Pz) using a combination of multiple signal processing and classification techniques [58-60]. The preprocessing phase notably aims to adapt the analysis criteria to cope with EEG interindividual variability. The analysis step uses both frequency and temporal techniques to retrieve numerous 30-second-long and 1-second-long features as well as performing the detection of several sleep microstructural events. The classification phase then reduces and summarizes this information by scoring sleep EEG into conventional sleep stages using artificial intelligence techniques, such as pattern recognition and fuzzy logic. This method was validated in healthy individuals and demonstrated high agreement with visual scoring realized by an experienced scorer, between 82.9% and 96%, depending on the number of vigilance stages considered, allowing its application in several research studies [59,60]. Classical sleep parameters like the total sleep time (TST: the sum of all sleep stages, expressed in minutes), sleep onset latency (SOL: the time, in minutes, from lights out to the first recorded epoch of any sleep stage, N1, N2, N3, or Rapid Eye Movement), sleep efficiency (calculated as the ratio of TST to total time in bed, multiplied by 100), wake after sleep onset (calculated as the total time in bed minus the TST), and time spent in stages N1, N2, N3, and Rapid Eye Movement

XSL•FO
RenderX

(expressed as a percentage of TST and in minutes) will be computed. Objective sleep quantity and quality will also be measured through actigraphy recordings and analyzed by ActiLife software (ActiGraph) using the Cole-Kripke algorithm. The sleep outcomes calculated were TST (minutes), SOL (minutes), and sleep efficiency (%).

Objective measurement of manifest sleepiness will be assessed during the MWT and analyzed visually by an experienced technician according to the AASM criteria [61]. Objective sleepiness is quantified by the SOL (minutes), defined as the time from lights out to the start of the first epoch of any sleep stage (N1, N2, N3, or rapid eye movement). If the patient succeeds in staying awake during the entire duration of the test, the SOL is set to 40 minutes.

Driving performance, assessed using a driving simulator, is quantified by the variability of the car's lateral position (standard deviation of lateral position [SDLP; cm]), indicating weaving of the car, and the number of inappropriate crossings of lateral highway lane markers (Inappropriate Line Crossing).

Cognitive performance is assessed using the PVT, the CPT, and the TAP. For the PVT, outcome measures include mean reaction time (RT; milliseconds), the fastest and slowest 10% of RTs, and the number of lapses (RT>500 ms). For the CPT, outcomes consist of the mean RT (measuring the response execution process), the number of omissions (failure to respond to target stimuli), and commission errors (responses given to nontarget stimuli). For the TAP, outcomes include mean RT and the number of omissions (missed targets).

Subjective sleepiness and mind wandering are assessed using the KSS and the CEC, respectively. The outcome measure for both is the score on the corresponding scale.

Data Analysis

Statistical Analysis

The analyses will be performed using R version 4.4.2 (R Foundation for Statistical Computing) and IBM SPSS Statistics version 27 (IBM Corp) software for Windows. The normality of the data will be assessed using the Shapiro-Wilk normality test, the equality of variance will be evaluated with the Levene test, and the sphericity of the data will be checked using the Mauchly test. Continuous variables will be expressed as means and categorical variables as proportions (%). *P* values less than .05 will be considered statistically significant.

OSS Criteria Accuracy Evaluation

Considering that the collected data consist of either continuous linear data, counts, or proportions, it is very likely that some of the data will not follow a normal distribution. Therefore, the use of parametric tests may be compromised. Moreover, our outcomes are measured repeatedly over time and may be highly dependent on interindividual variability. Thus, we will use a generalized linear mixed model to explore our first main objective, with relevant interindividual data incorporated into the model if needed (including patient ID, age, gender, daydreaming frequency, chronotype, and sleep history).

Using such models, OSS and MWT sleep onset latencies will be compared over time in order to determine whether OSS criteria can accurately detect drowsiness compared to the MWT.

As a second step, behavioral performance based on the OSS score will be assessed to determine whether instantaneous drowsiness states, identified by OSS, can detect momentary functional impairments. The evaluation of OSS temporal accuracy will be conducted by selecting OSS scores greater than 0 at a time T (20-second epoch), surrounded by lower score values (at T-20 seconds and T+20 seconds) [26]. Behavioral outcomes will then be compared using a generalized linear mixed model between T-20 seconds, T, and T+20 seconds of selected OSS scores.

Predictive Models Using Resting State EEG

To predict drowsiness-mediated behavioral performance, the proposed model leverages resting-state EEG recordings obtained prior to task execution. The primary behavioral end point is SDLP, which is dichotomized into 2 classes: "Good Performance" and "Poor Performance" based on the 33rd and 66th percentiles of the SDLP distribution. This percentile-based approach ensures approximately balanced class sizes, and the resulting thresholds and participant categorizations were validated by domain experts.

Considering the nonstationary nature of EEG signals and their susceptibility to interindividual variability, machine learning methods capable of handling such variability will be prioritized. Feature extraction and selection will be guided by SHAP (Shapley Additive Explanations) values computed on the training set. Channels and features with high importance will be selected and then evaluated on a separate validation subset within each fold. Performance will be reported on a completely unseen test set to prevent leakage.

A leave-one-subject-out cross-validation strategy will be applied to provide a robust evaluation under intersubject variability. Within each fold, to optimize hyperparameters and feature selection, a nested validation set will be defined by randomly selecting a subset of participants from the training pool (ensuring no overlap with the test participant). All preprocessing, feature selection, and hyperparameter tuning will be performed exclusively on the training and validation subsets, with the test participant held out to avoid information leakage.

Classification will be performed using CatBoost, an ensemble-based gradient boosting algorithm suitable for tabular and heterogeneous data. To enhance generalizability and benchmark the findings across distinct algorithmic frameworks, additional classifiers, including radial basis function support vector machine, decision trees, logistic regression, and extreme gradient boosting, will also be applied. Demographic and behavioral covariates (eg, age, gender, chronotype, daydreaming frequency, and sleep history) will be optionally included to account for residual interindividual variability.

Model performance will be evaluated using accuracy, precision, recall, and F_1-score, reporting both macroaveraged and weighted-averaged metrics to account for potential class imbalance. Comparing these metrics allows assessment of classifier sensitivity to imbalanced classes. Bootstrapping and

XSL•FO
RenderX

permutation testing will be applied to evaluate the statistical significance and robustness of the results, ensuring that predictive performance is not due to chance.

Ethical Considerations

This study was approved by the French National Ethics Committee (Consultative Committee for the Protection of Persons participating in biomedical research, CPP Sud Est V, on April 14, 2022, under the number N° SI RIPH 2G: 22.00521.000045) and was registered in ClinicalTrials.gov (NCT05453643). The National Agency for the Safety of Medicines and Health Products was notified about this study. All staff members involved in this study ensure that the full research was conducted in accordance with ethical guidelines and regulations on research involving human participants (as stipulated in the Good Clinical Practices [International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use], law no. 2022 - 323 of March 4, 2022, on research involving human participants, and the Declaration of Helsinki).

All participants gave their written informed consent prior to inclusion in the study. This study complies with the General Data Protection Regulation (GDPR) and the MR-001 National Commission on Informatics and Liberty reference methodology, which require data to be deidentified after collection (for privacy protection). Personal information, such as participants' names, addresses, and medical status, will thus be exclusively managed at the examination center and not provided to third parties. All data collected were deidentified, participants being assigned a unique study code, and any information pertaining to personal details is kept in locked filing cabinets in the SANPSY laboratory, only accessible to authorized research staff directly running the study. Relevant data are entered into an electronic database using only the study codes assigned to each participant and are securely stored on an encrypted and secure server made accessible to research staff directly running the study through password-protected computers.

The risk of adverse events is considered low, and the study procedure can be canceled at any time. During the study, participants were monitored by psychologically trained staff, which ensured fast communication of complaints and an immediate response. Volunteers received financial compensation of €900 (US $1062.02) for their participation in the study. The final results of this study will be disseminated through peer-reviewed publications and conferences.

Results

This study was funded by Physip and ANR (Agence Nationale de la Recherche, National Research Agency) in 2019, with ethical committee (Comité de Protection des Personnes, Committee for the Protection of Persons) authorization in May 2022. Candidate recruitment began in March 2023 and was completed in May 2025 (N=40), with a database lock in June 2025. Data analysis started in June 2025 and is still ongoing.

Discussion

Principal Findings

The objective of this study is to validate a new, easily implementable spontaneous drowsiness measurement, derived from OSS criteria and based on the automatic analysis of a limited number of EEG channels, able to reflect manifest drowsiness, particularly the ability to stay awake and the associated cognitive and behavioral consequences, regardless of the time of day. In addition, the study aims to evaluate whether iterative resting-state EEG can reliably predict drowsiness and its functional outcomes, independently of the prior duration of wakefulness. These measurements will enable the early assessment of drowsiness, allowing for the prediction of the ability to maintain wakefulness and/or of both simple and/or complex cognitive impairments associated with drowsiness.

Effortless Installation

The key advantage of our EEG-based drowsiness detection system lies in its design for real-world applications by prioritizing a minimal number of EEG channels (ie, 2) and using short, iterative measurements. By reducing hardware complexity and optimizing data collection, our approach ensures practicality under operational conditions, making it suitable for monitoring in real-time scenarios. This approach not only enhances user comfort and system efficiency but also facilitates seamless integration into sophisticated driver and pilot monitoring systems, contributing to improved safety and performance.

Benefits of 24-Hour EEG

As part of our study protocol, we perform continuous EEG-based monitoring of drowsiness during a 24-hour sleep deprivation period and throughout a subsequent night following daytime recovery sleep. This presents a compelling opportunity to deepen our understanding of the temporal dynamics of sleep pressure and vigilance regulation. Such long-term recordings can capture natural fluctuations in alertness due to circadian and homeostatic processes, as well as individual susceptibility to sleep loss. This comprehensive dataset could inform the development of more robust and generalizable algorithms capable of detecting drowsiness across a wide range of real-world scenarios and time frames.

Moreover, a 24-hour EEG assessment allows the identification of individual variability in drowsiness expression and recovery, potentially enabling personalized predictive models. Such models could dynamically adapt to a user's physiological state, thereby enhancing the precision and timeliness of drowsiness detection. This approach aligns with the growing emphasis on personalized medicine and human-centered design in neurotechnology.

Developing algorithms trained on 24-hour EEG data may also support the integration of drowsiness monitoring into wearable or minimally invasive systems. By learning to identify EEG markers that remain consistent across different levels of activity, posture, and time of day, these algorithms could eventually be embedded in real-time applications, providing continuous monitoring and early warnings to mitigate fatigue-related risks.

XSL•FO
RenderX

Resting-State EEG Predictive Advantage

Unlike conventional automatic drowsiness detection systems that identify spontaneous drowsiness or related performance declines in real-time —often too late to prevent impairments or accidents— our approach might predict drowsiness-induced performance deterioration in advance, enabling timely intervention with effective countermeasures to prevent functional decline. The aim is to assess the cognitive readiness and forecast the likelihood of either good or impaired performance before the task.

Like certain algorithms, we will leverage machine learning analyses while integrating interindividual characteristics, such as age, sex, chronotype, and accumulated sleep deprivation. This personalized approach aims to enhance predictive accuracy.

By leveraging automatic sleep pressure quantification through resting-state EEG data—adjusted for interindividual characteristics— our study will provide valuable insights into individuals' vulnerability to drowsiness before critical lapses in attention or vigilance occur. This predictive capability enables proactive interventions, such as fatigue management strategies, to prevent performance deterioration and reduce the risk of accidents.

Implications for the World of Work and Transport

The ability to forecast drowsiness-related impairments has significant applications in occupational and transport settings [5,62-64]. Workers in industries requiring night shifts, such as in health care, emergency response, and manufacturing, frequently experience sleep deprivation, which can lead to decreased cognitive performance and increased accident risks at work [5,65,66]. By integrating predictive drowsiness assessment tools, employers can enhance safety measures through "fatigue" monitoring and targeted interventions to prevent fatigue-related errors and accidents.

The transport sector, particularly aviation, rail, and road transport, is highly vulnerable to drowsiness-related accidents [67]. Microsleeps and lapses in vigilance can have catastrophic consequences, making early prediction of drowsiness essential for safety. Unlike real-time systems that detect spontaneous drowsiness only at the moment it occurs, our approach will offer a preemptive strategy by identifying individuals at risk of performance decline related to drowsiness beforehand. In other words, our approach systematically evaluates an operator's cognitive readiness and forecasts the probability of optimal or impaired performance due to sleepiness prior to the initiation of a driving task, flight, or other transportation activity. Our EEG-based drowsiness detection systems can facilitate the implementation of drowsiness-monitoring technologies in pilots, train operators, and long-haul drivers, significantly reducing the likelihood of sleep-related incidents occurring. Additionally, our findings could inform regulatory policies mandating proactive drowsiness risk management in safety-sensitive professions.

Continuous real-time drowsiness assessment is essential, especially for validating Driver Drowsiness Warning systems in advanced driver-assistance systems. To detect spontaneous drowsiness during driving or work scenarios, we will develop an automated algorithm that analyzes 2 EEG channels in accordance with OSS scoring.

Clinical Health Applications

Identifying or predicting drowsiness during behavioral or cognitive tasks will primarily be used to assess the effects of medication or other interventions in conditions related to sleepiness or in patients experiencing excessive daytime sleepiness.

Beyond occupational and transport safety, the study's findings have important implications for clinical health. Sleep disorders, such as insomnia, sleep apnea, and narcolepsy, compromise an individual's ability to maintain wakefulness and cognitive function. By leveraging EEG-based biomarkers to measure manifest sleepiness or predict drowsiness susceptibility, clinicians could improve diagnostic accuracy and develop more personalized treatment plans. In fact, the MWT is the reference test for measuring manifest sleepiness (ability to maintain wakefulness under monotonous conditions) in sleepy patients and to determine the efficacy of treatment for sleepiness. In some countries, the MWT has a medico-legal value and is used to determine whether an individual with sleep-related driving risk is fit or unfit to drive after the implementation of therapeutic measures. In this case, the MWT must be repeated annually for heavy vehicle licenses and every 3 years for light vehicle licenses. The MWT is conducted in specialized sleep centers and requires appropriate equipment and trained personnel for accurate execution and interpretation. However, access to these centers may be limited due to a shortage of specialists and restricted availability of necessary equipment. Automatic analysis of sleep latency using OSS criteria during the MWT can facilitate the clinical management of sleepy patients and could add, as microsleep episode, potential additional value in the context of sleepiness assessment [68]. In this sense, the freely available VIGALL (Vigilance Assessment through Graphical Analysis of Linear and Localized) signals algorithm has already demonstrated its ability to provide comparable information on wakefulness regulation as the more resource-intensive and time-consuming Multiple Sleep Latency Test [60]. However, the VIGALL requires 25 EEG derivations, which may still limit its practicality in routine clinical use compared to the MEEGAWAKE single-EEG algorithm. Moreover, automatic analysis of resting-state EEG is simpler, more cost-effective, and easier to administer, making it a potentially valuable alternative to the MWT for the initial assessment of drowsiness in clinical settings. Automatic analysis of resting-state EEG would be a simpler, more cost-effective, and easier to administer assessment, making it a potentially valuable alternative to the MWT for the initial assessment of drowsiness in clinical settings.

Additionally, older individuals, patients recovering from neurological conditions, such as traumatic brain injuries [69] or neurodegenerative diseases [70,71], often experience disrupted sleep-wake cycles and daytime sleepiness. An easy-to-administer and precise assessment of their ability to stay awake and maintain attention could enhance rehabilitation strategies, thereby improving cognitive recovery and overall well-being.

XSL•FO
RenderX

Strengths and Weaknesses

One of the strengths of this study is its rigorous methodology, aimed at validating a drowsiness detection system through a strictly controlled sleep deprivation protocol, simulating realistic conditions of repeated night work or on-call shifts. The assessment of drowsiness throughout the sleep deprivation periods, both during the day and at night, provides a detailed understanding of its fluctuations according to circadian rhythms and the accumulation of sleep pressure. The analysis of behavioral and cognitive consequences is particularly thorough, combining subjective and objective measures through simulated driving tests, simple and complex cognitive tasks, and a scientifically recognized test for evaluating the ability to stay awake. Additionally, the sample is balanced in terms of gender and covers a wide age range (20 - 60 years), enhancing the generalizability of the results to a healthy adult population.

However, this study also has some limitations. Although the laboratory setting minimizes environmental biases and ensures strict standardization of experimental conditions, it does not fully replicate the constraints and dynamics of real-world settings, particularly in professional environments. Furthermore, the relatively small sample size (N=40) limits the statistical power of the conclusions. Another notable limitation is the exclusion of patients diagnosed with sleep disorders or excessive drowsiness, restricting the applicability of the findings to healthy individuals only. As a result, it is not possible to confirm that the tested system would be equally effective in a clinical population. These limitations should be considered before broader implementation of this technology.

Conclusion

This study is expected to contribute to a deeper understanding of the determinants of drowsiness and advance the development of proactive strategies for its monitoring and management across occupational, transport, and clinical contexts. Rather than relying solely on real-time drowsiness detection systems, this work examines how OSS-based criteria may capture spontaneous drowsiness and associated behavioral changes, and how resting-state EEG parameters may predict cognitive readiness and mid-term vulnerability to performance decline. The project aims to identify objective neuromarkers that could support earlier and more effective detection of drowsiness and sleep-related risks. Future work will involve validating these potential neuromarkers across diverse healthy and patient populations and evaluating how they may be integrated into fatigue-risk management tools and clinical assessment frameworks. Together, these efforts aim to support the broader implementation of predictive drowsiness-assessment methods for improving safety, optimizing clinical care, and reducing sleep-related accident risk.

Acknowledgments

The study was conducted within the framework of the French National Research Agency (Agence Nationale de la Recherche; ANR) “LabCom” program, a public-private consortium initiative. We thank the Bordeaux University, the French National Center for Scientific Research (CNRS), and the staff of PRNPP (Pôle de Recherche Neuropsychologie et Psychologie Physiologique; Research Center for Neuropsychology and Physiological Psychology) facility and their leader Pierre Philip for their support in the implementation of this study. The authors declare the use of generative artificial intelligence (GenAI) in the writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: translation proof checking. The GenAI tool used was GPT-4o mini. Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes. This declaration is submitted by collective responsibility. The roles and responsibilities of each partner are defined by a consortium agreement.

Funding

This work was supported by the National Research Agency (Agence Nationale de la Recherche [ANR]; grant ANR-19-LCV1-0004 MEEGASAFE) and by Physip SA. The ANR provides funding to the public research laboratory, while the program promotes long-term collaborations and innovation between the laboratory and an industrial partner. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

All relevant data that will be generated and analyzed during this study will be included in the resulting published articles and the supplementary information files.

Authors' Contributions

Conceptualization: C Berthomier, JT
Data curation: C Boitard, JM, PB
Formal analysis: C Boitard, KS, ZM, JC, C Berthomier
Funding acquisition: MB, PP, C Berthomier, JT
Methodology: KS, PS, C Berthomier, JT
Project administration: PP, JAMF
Software: KS, JM, PB, C Berthomier

XSL•FO
RenderX

Supervision: C Berthomier, JT
Visualization: C Boitard, JT
Writing–original draft: C Boitard, KS, C Berthomier, JT
Writing–review & editing: C Boitard, KS, ZM, JC, PS, JL, PP, JAMF, C Berthomier, JT

Conflicts of Interest

C Boitard, ZM, KS, JC, JL, PS, PP, MB, JAMF, and JT have declared that no competing interests exist. C Berthomier and PB have ownership/directorship of Physip. JM is an employee of Physip.

Checklist 1
SPIRIT checklist.
[PDF File, 331 KB - resprot_v15i1e83969_app1.pdf]

References

1. Carskadon MA, Dement WC. Daytime sleepiness: quantification of a behavioral state. Neurosci Biobehav Rev 1987;11(3):307-317. [doi: 10.1016/s0149-7634(87)80016-7] [Medline: 3684058]
2. Peter-Derex L, Micoulaud-Franchi JA, Lopez R, Barateau L. Evaluation of hypersomnolence: from symptoms to diagnosis, a multidimensional approach. Rev Neurol (Paris) 2023 Oct;179(7):715-726. [doi: 10.1016/j.neurol.2023.07.004] [Medline: 37563022]
3. Lopez R, Micoulaud-Franchi JA, Barateau L, Dauvilliers Y. A multi-dimensional approach to hypersomnolence [Article in French]. Médecine du Sommeil 2022 Dec;19(4):206-220. [doi: 10.1016/j.msom.2022.06.003]
4. Goel N, Rao H, Durmer JS, Dinges DF. Neurocognitive consequences of sleep deprivation. Semin Neurol 2009 Sep;29(4):320-339. [doi: 10.1055/s-0029-1237117] [Medline: 19742409]
5. Dinges DF. An overview of sleepiness and accidents. J Sleep Res 1995 Dec;4(S2):4-14. [doi: 10.1111/j.1365-2869.1995.tb00220.x] [Medline: 10607205]
6. Mitler MM, Carskadon MA, Czeisler CA, Dement WC, Dinges DF, Graeber RC. Catastrophes, sleep, and public policy: consensus report. Sleep 1988 Feb;11(1):100-109. [doi: 10.1093/sleep/11.1.100] [Medline: 3283909]
7. Häkkänen H, Summala H. Sleepiness at work among commercial truck drivers. Sleep 2000 Feb 1;23(1):49-57. [Medline: 10678465]
8. Verster JC, Taillard J, Sagaspe P, Olivier B, Philip P. Prolonged nocturnal driving can be as dangerous as severe alcohol-impaired driving. J Sleep Res 2011 Dec;20(4):585-588. [doi: 10.1111/j.1365-2869.2010.00901.x] [Medline: 21226780]
9. Stawarczyk D, D'Argembeau A. Conjoint influence of mind-wandering and sleepiness on task performance. J Exp Psychol Hum Percept Perform 2016 Oct;42(10):1587-1600. [doi: 10.1037/xhp0000254] [Medline: 27268466]
10. Jubera-Garcia E, Gevers W, Van Opstal F. Local build-up of sleep pressure could trigger mind wandering: evidence from sleep, circadian and mind wandering research. Biochem Pharmacol 2021 Sep;191:114478. [doi: 10.1016/j.bcp.2021.114478] [Medline: 33609561]
11. Kaida K, Takahashi M, Akerstedt T, et al. Validation of the Karolinska sleepiness scale against performance and EEG variables. Clin Neurophysiol 2006 Jul;117(7):1574-1581. [doi: 10.1016/j.clinph.2006.03.011] [Medline: 16679057]
12. Van Dongen HPA, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 2003 Mar 15;26(2):117-126. [doi: 10.1093/sleep/26.2.117] [Medline: 12683469]
13. Philip P, Taillard J, Micoulaud-Franchi JA. Sleep restriction, sleep hygiene, and driving safety: the importance of situational sleepiness. Sleep Med Clin 2019 Dec;14(4):407-412. [doi: 10.1016/j.jsmc.2019.07.002] [Medline: 31640868]
14. Ebrahim Shaik M. A systematic review on detection and prediction of driver drowsiness. Transp Res Interdiscip Persp 2023 Sep;21:100864. [doi: 10.1016/j.trip.2023.100864]
15. Gottlieb DJ, Ellenbogen JM, Bianchi MT, Czeisler CA. Sleep deficiency and motor vehicle crash risk in the general population: a prospective cohort study. BMC Med 2018 Mar 20;16(1):44. [doi: 10.1186/s12916-018-1025-7] [Medline: 29554902]
16. Wittmann M, Dinich J, Merrow M, Roenneberg T. Social jetlag: misalignment of biological and social time. Chronobiol Int 2006;23(1-2):497-509. [doi: 10.1080/07420520500545979] [Medline: 16687322]
17. Booker LA, Magee M, Rajaratnam SMW, Sletten TL, Howard ME. Individual vulnerability to insomnia, excessive sleepiness and shift work disorder amongst healthcare shift workers. a systematic review. Sleep Med Rev 2018 Oct;41:220-233. [doi: 10.1016/j.smrv.2018.03.005] [Medline: 29680177]
18. Taillard J, Micoulaud-Franchi JA, Martin VP, Peter-Derex L, Vecchierini MF. Objective evaluation of excessive daytime sleepiness. Neurophysiol Clin 2024 Apr;54(2):102938. [doi: 10.1016/j.neucli.2023.102938] [Medline: 38401239]

XSL•FO
RenderX

19. Mitler MM, Gujavarty KS, Browman CP. Maintenance of wakefulness test: a polysomnographic technique for evaluation treatment efficacy in patients with excessive somnolence. Electroencephalogr Clin Neurophysiol 1982 Jun;53(6):658-661. [doi: 10.1016/0013-4694(82)90142-0] [Medline: 6177511]
20. Gillberg M, Kecklund G, Akerstedt T. Sleepiness and performance of professional drivers in a truck simulator--comparisons between day and night driving. J Sleep Res 1996 Mar;5(1):12-15. [doi: 10.1046/j.1365-2869.1996.00013.x] [Medline: 8795796]
21. Balandong RP, Ahmad RF, Mohamad Saad MN, Malik AS. A review on EEG-based automatic sleepiness detection systems for drivers. IEEE Access 2018;6:22908-22919. [doi: 10.1109/ACCESS.2018.2811723]
22. Peng Y, Xu Q, Lin S, et al. The application of electroencephalogram in driving safety: current status and future prospects. Front Psychol 2022;13:919695. [doi: 10.3389/fpsyg.2022.919695] [Medline: 35936295]
23. Stancin I, Cifrek M, Jovic A. A review of EEG signal features and their application in driver drowsiness detection systems. Sensors (Basel) 2021 May 30;21(11):3786. [doi: 10.3390/s21113786] [Medline: 34070732]
24. Saleem AA, Siddiqui HUR, Raza MA, Rustam F, Dudley S, Ashraf I. A systematic review of physiological signals based driver drowsiness detection systems. Cogn Neurodyn 2023 Oct;17(5):1229-1259. [doi: 10.1007/s11571-022-09898-9] [Medline: 37786662]
25. Muzet A, Pebayle T, Otmani S, Peters B, Nikolau S, Bekiaris E. Preventing driver drowsiness at the wheel: can steering grip sensor measurement contribute to its prediction. Presented at: European Congress on Intelligent Transportation Systems and Services; May 24-26, 2004 URL: https://trid.trb.org/View/1158614 [accessed 2025-01-18]
26. Giot C, Hay M, Chesneau C, et al. Towards a new approach to detect sleepiness: validation of the objective sleepiness scale under simulated driving conditions. Transportation Research Part F: Traffic Psychology and Behaviour 2022 Oct;90:109-119. [doi: 10.1016/j.trf.2022.08.007]
27. Berthomier C, Muzet A, Berthomier P, Prado J, Mattout J. Real-time automatic measure of drowsiness based on a single EEG channel. Presented at: 19th Congress of the European Sleep Research Society; Sep 9-13, 2008 URL: https://esrs.eu/wp-content/uploads/2019/01/ESRS_2008_Final_Programme.pdf [accessed 2026-01-30]
28. Getzmann S, Gajewski PD, Schneider D, Wascher E. Resting-state EEG data before and after cognitive activity across the adult lifespan and a 5-year follow-up. Sci Data 2024 Sep 10;11(1):988. [doi: 10.1038/s41597-024-03797-w] [Medline: 39256413]
29. Xiang C, Fan X, Bai D, Lv K, Lei X. A resting-state EEG dataset for sleep deprivation. Sci Data 2024 Apr 24;11(1):427. [doi: 10.1038/s41597-024-03268-2] [Medline: 38658675]
30. Jawinski P, Kittel J, Sander C, et al. Recorded and reported sleepiness: the association between brain arousal in resting state and subjective daytime sleepiness. Sleep 2017 Jul 1;40(7):zsx099. [doi: 10.1093/sleep/zsx099] [Medline: 28605521]
31. Khoo SY, Lai WH, On SH, et al. Resting-state electroencephalography (EEG) microstates of healthy individuals following mild sleep deprivation. Sci Rep 2024 Jul 22;14(1):16820. [doi: 10.1038/s41598-024-67902-0] [Medline: 39039219]
32. Hegerl U, Stein M, Mulert C, et al. EEG-vigilance differences between patients with borderline personality disorder, patients with obsessive-compulsive disorder and healthy controls. Eur Arch Psychiatry Clin Neurosc 2008 Apr;258(3):137-143. [doi: 10.1007/s00406-007-0765-8]
33. Löwe B, Wahl I, Rose M, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. J Affect Disord 2010 Apr;122(1-2):86-95. [doi: 10.1016/j.jad.2009.06.019] [Medline: 19616305]
34. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep 1991 Dec;14(6):540-545. [doi: 10.1093/sleep/14.6.540] [Medline: 1798888]
35. Chung F, Yegneswaran B, Liao P, et al. STOP questionnaire: a tool to screen patients for obstructive sleep apnea. Anesthesiology 2008 May;108(5):812-821. [doi: 10.1097/ALN.0b013e31816d83e4] [Medline: 18431116]
36. Walters AS, Aldrich MS, Allen R, et al. Toward a better definition of the restless legs syndrome. Mov Disord 1995 Sep;10(5):634-642. [doi: 10.1002/mds.870100517]
37. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol 1976;4(2):97-110. [Medline: 1027738]
38. Taillard J, Philip P, Chastang JF, Bioulac B. Validation of Horne and Ostberg morningness-eveningness questionnaire in a middle-aged population of French workers. J Biol Rhythms 2004 Feb;19(1):76-86. [doi: 10.1177/0748730403259849] [Medline: 14964706]
39. Coelho J, Micoulaud-Franchi JA, D'incau E, et al. Validation of the French version of the Munich ChronoType questionnaire and associations between chronotype and physiological parameters. Chronobiol Int 2024 Jul;41(7):996-1007. [doi: 10.1080/07420528.2024.2362309] [Medline: 38860541]
40. Roenneberg T, Wirz-Justice A, Merrow M. Life between clocks: daily temporal patterns of human chronotypes. J Biol Rhythms 2003 Feb;18(1):80-90. [doi: 10.1177/0748730402239679] [Medline: 12568247]
41. Jankowski KS. Social jet lag: sleep-corrected formula. Chronobiol Int 2017;34(4):531-535. [doi: 10.1080/07420528.2017.1299162] [Medline: 28318321]

XSL•FO
RenderX

42. Stawarczyk D, Majerus S, Van der Linden M, D'Argembeau A. Using the Daydreaming Frequency Scale to investigate the relationships between mind-wandering, psychological well-being, and present-moment awareness. Front Psychol 2012;3:363. [doi: 10.3389/fpsyg.2012.00363] [Medline: 23055995]
43. Akerstedt T, Gillberg M. Subjective and objective sleepiness in the active individual. Int J Neurosci 1990 May;52(1-2):29-37. [doi: 10.3109/00207459008994241] [Medline: 2265922]
44. Gillberg M, Kecklund G, Akerstedt T. Relations between performance and subjective ratings of sleepiness during a night awake. Sleep 1994 Apr;17(3):236-241. [doi: 10.1093/sleep/17.3.236] [Medline: 7939123]
45. Littner MR, Kushida C, Wise M, et al. Practice parameters for clinical use of the multiple sleep latency test and the maintenance of wakefulness test. Sleep 2005 Jan;28(1):113-121. [doi: 10.1093/sleep/28.1.113] [Medline: 15700727]
46. Krahn LE, Arand DL, Avidan AY, et al. Recommended protocols for the multiple sleep latency test and maintenance of wakefulness test in adults: guidance from the American Academy of Sleep Medicine. J Clin Sleep Med 2021 Dec 1;17(12):2489-2498. [doi: 10.5664/jcsm.9620] [Medline: 34423768]
47. Davenne D, Lericollais R, Sagaspe P, et al. Reliability of simulator driving tool for evaluation of sleepiness, fatigue and driving performance. Accid Anal Prev 2012 Mar;45:677-682. [doi: 10.1016/j.aap.2011.09.046] [Medline: 22269557]
48. Sagaspe P, Taillard J, Akerstedt T, et al. Extended driving impairs nocturnal driving performances. PLoS One 2008;3(10):e3493. [doi: 10.1371/journal.pone.0003493] [Medline: 18941525]
49. Dinges DF, Powell JW. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behav Res Methods, Instrum Comput 1985 Nov;17(6):652-655. [doi: 10.3758/BF03200977]
50. Khitrov MY, Laxminarayan S, Thorsley D, et al. PC-PVT: a platform for psychomotor vigilance task testing, analysis, and prediction. Behav Res Methods 2014 Mar;46(1):140-147. [doi: 10.3758/s13428-013-0339-9] [Medline: 23709163]
51. Conners C, Staff M. Conners' Continuous Performance Test II (CPT II) Computer Program for Windows, Technical Guide and Software Manual: Multi-Health Systems; 2000.
52. Labelle MA, Dang-Vu TT, Petit D, Desautels A, Montplaisir J, Zadra A. Sleep deprivation impairs inhibitory control during wakefulness in adult sleepwalkers. J Sleep Res 2015 Dec;24(6):658-665. [doi: 10.1111/jsr.12315] [Medline: 26087833]
53. Wang M, Shen C, Liu X, et al. Executive function performance in children and adolescent patients with narcolepsy type 1. Sleep Med 2024 Jul;119:342-351. [doi: 10.1016/j.sleep.2024.05.021] [Medline: 38754344]
54. Sagaspe P, Taillard J, Amiéva H, et al. Influence of age, circadian and homeostatic processes on inhibitory motor control: a Go/Nogo task study. PLoS One 2012;7(6):e39410. [doi: 10.1371/journal.pone.0039410] [Medline: 22761784]
55. Zimmermann P, Fimm B. A test battery for attentional performance. In: Applied Neuropsychology of Attention: Psychology Press; 2002.
56. Stawarczyk D, Majerus S, Maj M, Van der Linden M, D'Argembeau A. Mind-wandering: phenomenology and function as assessed with a novel experience sampling method. Acta Psychol (Amst) 2011 Mar;136(3):370-381. [doi: 10.1016/j.actpsy.2011.01.002] [Medline: 21349473]
57. Cajochen C, Brunner DP, Kräuchi K, Graw P, Wirz-Justice A. EEG and subjective sleepiness during extended wakefulness in seasonal affective disorder: circadian and homeostatic influences. Biol Psychiatry 2000 Apr 1;47(7):610-617. [doi: 10.1016/s0006-3223(99)00242-5] [Medline: 10745053]
58. Peter-Derex L, Berthomier C, Taillard J, et al. Automatic analysis of single-channel sleep EEG in a large spectrum of sleep disorders. J Clin Sleep Med 2021 Mar 1;17(3):393-402. [doi: 10.5664/jcsm.8864] [Medline: 33089777]
59. Berthomier C, Drouot X, Herman-Stoïca M, et al. Automatic analysis of single-channel sleep EEG: validation in healthy individuals. Sleep 2007 Nov;30(11):1587-1595. [doi: 10.1093/sleep/30.11.1587] [Medline: 18041491]
60. Berthomier C, Muto V, Schmidt C, et al. Exploring scoring methods for research studies: accuracy and variability of visual and automated sleep scoring. J Sleep Res 2020 Oct;29(5):e12994. [doi: 10.1111/jsr.12994] [Medline: 32067298]
61. Berry RB, Brooks R, Gamaldo C, et al. AASM Scoring Manual Updates for 2017 (Version 2.4). J Clin Sleep Med 2017 May 15;13(5):665-666. [doi: 10.5664/jcsm.6576] [Medline: 28416048]
62. Moradi A, Nazari SSH, Rahmani K. Sleepiness and the risk of road traffic accidents: a systematic review and meta-analysis of previous studies. Transp Res Part F Traffic Psychol Behav 2019 Aug;65:620-629. [doi: 10.1016/j.trf.2018.09.013]
63. Gander P, Hartley L, Powell D, et al. Fatigue risk management: organizational factors at the regulatory and industry/company level. Accid Anal Prev 2011 Mar;43(2):573-590. [doi: 10.1016/j.aap.2009.11.007] [Medline: 21130218]
64. Williamson A, Lombardi DA, Folkard S, Stutts J, Courtney TK, Connor JL. The link between fatigue and safety. Accid Anal Prev 2011 Mar;43(2):498-515. [doi: 10.1016/j.aap.2009.11.011] [Medline: 21130213]
65. Imes CC, Barthel NJ, Chasens ER, et al. Shift work organization on nurse injuries: a scoping review. Int J Nurs Stud 2023 Feb;138:104395. [doi: 10.1016/j.ijnurstu.2022.104395] [Medline: 36481596]
66. Barger LK, Cade BE, Ayas NT, et al. Extended work shifts and the risk of motor vehicle crashes among interns. N Engl J Med 2005 Jan 13;352(2):125-134. [doi: 10.1056/NEJMoa041401] [Medline: 15647575]
67. Marcus JH, Rosekind MR. Fatigue in transportation: NTSB investigations and safety recommendations. Inj Prev 2017 Aug;23(4):232-238. [doi: 10.1136/injuryprev-2015-041791] [Medline: 26929259]
68. Davin A, Tamisier R, Barateau L, et al. Shedding light on microsleep episodes for comprehensive sleepiness assessment: a narrative review. Sleep Med Rev 2025 Aug;82:102113. [doi: 10.1016/j.smrv.2025.102113]

XSL•FO
RenderX

69. Viola-Saltzman M, Watson NF. Traumatic brain injury and sleep disorders. Neurol Clin 2012 Nov;30(4):1299-1312. [doi: 10.1016/j.ncl.2012.08.008] [Medline: 23099139]
70. Maestri M, Romigi A, Schirru A, et al. Excessive daytime sleepiness and fatigue in neurological disorders. Sleep Breath 2020 Jun;24(2):413-424. [doi: 10.1007/s11325-019-01921-4] [Medline: 31444679]
71. Dubessy AL, Arnulf I. Sleepiness in neurological disorders. Rev Neurol (Paris) 2023 Oct;179(7):755-766. [doi: 10.1016/j.neurol.2023.07.005] [Medline: 37598089]

Abbreviations

AASM: American Academy of Sleep Medicine
ANR: Agence Nationale de la Recherche
CEC: Conscient Experience Characterization
CONSORT: Consolidated Standards of Reporting Trials
CPT: continuous performance test
DDFS: Daydreaming Frequency Scale
ECG: electrocardiogram
EEG: electroencephalogram
EMG: electromyogram
EOG: electrooculogram
ESS: Epworth Sleepiness Scale
GDPR: General Data Protection Regulation
INRETS-MSIS SIM2: Institut National de Recherche sur les Transports et leur Sécurité — Modélisations, Simulations et Systèmes d'Information pour la Sécurité — Simulateur version 2
KSS: Karolinska Sleepiness Scale
MSFsc: midsleep on work-free days
MSWsc: midsleep on working days
MWT: maintenance of wakefulness test
OSAS: obstructive sleep apnea syndrome
OSS: Objective Sleepiness Scale
PHQ-4: Patient Health Questionnaire-4
PVT: Psychomotor Vigilance Task
RLS: restless legs syndrome
RT: reaction time
SDLP: standard deviation of lateral position
SHAP: Shapley Additive Explanations
SOL: sleep onset latency
STOP-BANG: snoring, tiredness, observed apnea, blood pressure, BMI, age, neck circumference, and gender
TAP: Test of Attentional Performance
TST: total sleep time
VIGALL: Vigilance Assessment through Graphical Analysis of Linear and Localized

Edited by J Sarvestan; submitted 12.Sep.2025; peer-reviewed by SS Soleimanloo; accepted 17.Dec.2025; published 17.Feb.2026.

Please cite as:
Boitard C, Mazurie Z, Sadatnejad K, Coelho J, Sagaspe P, Lenoir J, Mattei J, Berthomier P, Brandewinder M, Philip P, Micoulaud Franchi JA, Berthomier C, Taillard J
Evaluating Electroencephalogram-Based Predictive Model for Drowsiness Measurement to Reduce Accident Risk in Active Individuals: Protocol for a Preliminary Monocentric Study
JMIR Res Protoc 2026;15:e83969
URL: https://www.researchprotocols.org/2026/1/e83969
doi:10.2196/83969

© Chloé Boitard, Zoé Mazurie, Khadijeh Sadatnejad, Julien Coelho, Patricia Sagaspe, Julie Lenoir, Julien Mattei, Pierre Berthomier, Marie Brandewinder, Pierre Philip, Jean-Arthur Micoulaud Franchi, Christian Berthomier, Jacques Taillard. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is

XSL•FO
RenderX

properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Cognitive and Behavioral Risks and Maintenance Factors of Hikikomori: Protocol for an Ecological Momentary Assessment Study

Alexander MacLellan, PhD; Keisuke Takano, PhD

[1]Human Informatics and Interaction Research Institute, National Institute of Advanced Industrial Science and Technology, 1-1-1 Higashi, Tsukuba, Japan

Corresponding Author:
Alexander MacLellan, PhD
Human Informatics and Interaction Research Institute, National Institute of Advanced Industrial Science and Technology, 1-1-1 Higashi, Tsukuba, Japan

Abstract

Background: Hikikomori is a state of social withdrawal first identified in Japan and is gaining interest globally. Classically, hikikomori is described as a state of isolation within one's home, though recent conceptualizations have proposed a continuum of severity. Hikikomori frequently shares symptoms with depression, social anxiety, autism, and schizophrenia, as well as internet and gaming disorders. Clinical case studies and cross-sectional studies suggest that dysfunctional emotion regulation, familial support, and internet behaviors are proposed to contribute to the onset and maintenance of a withdrawn state, though they have not been explored longitudinally.

Objective: This study aims to investigate affective, behavioral, and cognitive correlates of hikikomori symptoms, and how daily mood, social enjoyment, familial support, and internet usage may maintain a socially withdrawn state.

Methods: A minimum of 84 participants aged between 18 and 60 years will complete self-report measures of hikikomori symptoms, internet addiction, depression, anxiety, autism, and fear of offending others before participating in 14 days of ecological momentary assessment surveys. Surveys will be delivered 5 times per day from 8 AM to 10 PM, measuring mood, internet behavior, familial relationships, social interaction frequency, anticipatory and consummatory enjoyment, sleep quality, and physical activity. Participants will repeat the self-report measure of hikikomori symptoms postmonitoring period.

Results: Recruitment began in November 21, 2025. Data collection and analysis are scheduled to be completed by summer 2026, with the results also scheduled to be available by the end of summer 2026. Correlation and multiple regression analyses will investigate whether internet addiction, social anxiety, expressive suppression, fear of offending others, daily mood, internet use, social enjoyment, and familial support predict hikikomori symptoms. Time-lagged network analyses will explore the temporal dynamics of these relationships, and how these differ in those with high and low levels of hikikomori symptoms. Finally, time-lagged logistic regressions will explore which factors predict future social behavior.

Conclusions: This study will be the first to investigate currently proposed mechanisms underlying hikikomori, while also exploring the time-varying relationships between affect and social behavior. The results will provide initial evidence for factors that predict hikikomori symptoms, explore candidate mechanisms underlying hikikomori, and identify potential maintenance factors as targets for intervention.

Trial Registration: OSF Registries 10.17605/OSF.IO/AFRSB; https://osf.io/afrsb

International Registered Report Identifier (IRRID): PRR1-10.2196/81384

(JMIR Res Protoc 2026;15:e81384) doi:10.2196/81384

KEYWORDS

hikikomori; ecological momentary analysis; social withdrawal; social anxiety; internet addiction

Introduction

Background

Hikikomori is a condition characterized by persistent isolation in one's home alongside functional impairment and distress associated with this isolation [1]. First recognized and examined in Japan, hikikomori is now recognized globally [2]. The isolation is initially voluntary, similar to psychiatric disorders such as autism [3], who may find relief in their withdrawal from social situations; or in anxiety and depression [1], who may experience distress or apathy regarding socializing. However, there are some individuals who socially withdraw and yet do not present with any comorbid psychiatric disorders, raising

XSL•FO
RenderX

questions over whether hikikomori relates to a distinct state or a transdiagnostic phenomenon [4]. Additionally, while classifying an individual as hikikomori has primarily been determined by the length and severity of their social withdrawal (eg, for 6 months), there is growing recognition of the condition existing on a spectrum of severity [1]. To date, much research has proposed onset and maintenance factors from clinical case reports [5] and cross-sectional studies [6]. Affect dynamics [1], social reward processing [7], internet behaviors [8], familial relationships [9], and physical health [10] have all been identified, though little has investigated the time-varying nature of these relationships and whether they predict social withdrawal tendencies in daily life.

Hikikomori shares patterns of disordered affect and emotion regulation with depression, social anxiety, autism, and schizophrenia [1]. Low mood and anhedonia are frequently present in hikikomori patients [1], leading to questions of whether hikikomori may be a subtype or expression of depression [1,2]. Additionally, hikikomori patients often express anger or display higher levels of irritability [5,11]. High irritability and expressions of anger are linked with impaired social outcomes [12], social anxiety [13], and well-being [14]. While sustained negative affect and distress are hallmarks of depression and anxiety and are often examined as traits, emotions are not fixed states and variation is expected. Intensity of emotional change, stability of the emotional state, and how quickly an emotion returns to baseline after a deviation are all predictive of psychological well-being [15,16]. Together, these patterns of emotional dynamics are found to be related to mental health and well-being, with less variable and more stable patterns, both in the frequency of change and the speed of return to baseline, found to be associated with better mental health [16]. This suggests that negative affect and emotional instability play a key role in the onset and maintenance of hikikomori. Differences in emotion regulation styles may underlie these affective variations, with adaptive strategies such as cognitive reappraisal found to be associated with positive affect, while maladaptive strategies such as rumination and suppression are associated with negative affect [17], as well as social anxiety and depression [18].

Equally, anhedonia may play a key role in the onset or maintenance of a socially withdrawn state. In populations with depression, anhedonia is linked with social isolation [19], which may be precipitated by a reduced generation of anticipatory enjoyment [7,20]. Learning from rewarding outcomes (eg, an enjoyable conversation) is also found to be impaired in depression [21] and social anxiety [22], suggesting that social withdrawal may partially result from reduced anticipatory enjoyment, failing to generate motivation to act, or failure to learn from a positive experience preventing an adaptive increase in anticipatory enjoyment at the next occasion. Other forms of communication, such as via the internet, may be more accessible in this case. Taken together, research from populations with both depression and social anxiety suggests that affect, including emotional variability and stability, and reward processing are implicated in social withdrawal. However, it is unclear how these factors relate to hikikomori symptoms and episodes of social withdrawal.

Aside from social withdrawal, behavioral hallmarks of hikikomori include problematic smartphone and internet use [8,23], excessively playing video games [24], a reliance on familial support [1], a more sedentary lifestyle, and poorer sleep quality [10]. Problematic internet use may allow and support a socially withdrawn state by replacing face-to-face connections with virtual interactions. While virtual interactions provide social contact with less risk of stress or challenge [23], this may also prevent individuals from challenging negative beliefs around in-person communication or developing positive resilience and coping strategies. Additionally, given the widespread availability of internet access, it is possible for individuals to work and manage their affairs solely online, again enabling a socially withdrawn state. Previous research has found links between problematic internet use, anhedonia, and well-being [25,26] and therefore is often a target both for research and intervention. Excessive gaming is also commonly researched as either a subtype of internet use or in isolation. Gaming may fill similar roles to internet use, namely as a coping mechanism, and equally may be a factor in the maintenance of social withdrawal [8,24].

Familial relationships are frequently discussed as a risk factor for a hikikomori condition [1]. Dysfunctional family relationships are common in observed cases of hikikomori [27], with both parental abuse [28] and overindulgence [9] linked to the socially withdrawn condition. Familial support may play a critical logistical role in the maintenance [9] and recovery [29] of hikikomori, though how this interacts with other variables is uncertain [9]. It may be that supportive actions (providing food and economic support), though initially made in the best interests of the patient, become maintenance factors that are difficult for the family to stop once started. Conversely, behaviors to challenge the withdrawn condition may also adversely affect hikikomori [28]. Therefore, uncovering how familial behaviors affect social withdrawal and interact with other factors is required.

Finally, sleep patterns, sleep quality, and physical activity may also play a role. Cases of hikikomori often note disordered sleeping patterns [10,30], such as primarily waking at night and sleeping during the day [28]. These patterns may have originated as a way to manage anxiety around other people, though as with defaulting to virtual communication, this may also increase feelings of disconnection and isolation [1]. When combined with a preference for virtual communication and isolation, the sedentary lifestyle noted in hikikomori samples is unsurprising and may contribute to poorer physical health [10].

The Study

Taken together, hikikomori describes a condition that shares symptoms, cognitive profiles, and behavioral profiles with several mental health disorders [1]. While many researchers have proposed maintenance factors, to date, there has been little investigating how observed affective, cognitive, or behavioral factors influence symptoms, social motivation, and social withdrawal over time.

Ecological momentary assessment (EMA) can elucidate the time-varying nature of these relationships by repeatedly sampling daily mood and behaviors. EMA has previously been

XSL•FO
RenderX

used to investigate social withdrawal and depressive symptoms [31], for example, previous studies have found that participants with clinical depression exhibit reduced anticipatory and consummatory enjoyment [32]. EMA therefore offers an opportunity to identify which daily behaviors and mood patterns are indicative of a risk of hikikomori symptoms and social withdrawal. Additionally, the EMA method allows us to investigate proposed factors and behaviors with high ecological validity given the measurements are occurring during the participant's daily life.

This study will investigate affective, behavioral, and cognitive correlates of hikikomori symptoms, and how familial support and internet usage may maintain a socially withdrawn state. Over 2 weeks, participants will be prompted to complete 5 EMA surveys per day, with questions asking about their mood, internet activity, familial interactions, and social behavior. Surveys will be delivered every 3 hours from 8 AM to 8 PM, allowing us to explore temporal dynamics of affect and behavior, and how this relates to hikikomori symptoms.

The study proposes the first research question (RQ1): What factors are related to hikikomori symptoms? The first hypothesis is that "hikikomori symptoms will be positively correlated with internet addiction, social anxiety, expressive suppression, and fear of offending others as measured at baseline" (H1a). The second hypothesis is that "hikikomori symptoms postmonitoring period will be predicted by averaged daily internet usage, sleep quality, sleep duration, physical activity, familial contact, anticipatory and consummatory social enjoyment, mood, and emotional variability, while controlling for depression, autism, and anxiety" (H1b).

Additionally, we will explore the temporal dynamics of momentary mood, internet and familial behaviors, and how these differ in individuals with high and low symptoms of hikikomori. This study also proposes the second research question (RQ2): How do internet, social, and familial behaviors predict future social behavior and enjoyment? The first hypothesis is that "recent social enjoyment will predict the likelihood of future social interaction" (H2a). The second hypothesis is that "recent internet usage will predict the likelihood of future social interaction" (H2b). The third hypothesis is that "recent familial support will predict the likelihood of future social interaction" (H2c). The final hypothesis is that "Recent mood and anger will predict the likelihood of future social interaction" (H2d).

Methods

Study Design

This study will be a longitudinal observational design, using EMA over a 2-week period. An Adapted STROBE Checklist for Reporting EMA Studies (CREMAS) is included in Checklist 1.

Transparency and Openness Statement

Anonymized data and analysis code will be made available from the OSF page for this project [33]. The first draft of the protocol paper was submitted before data collection commenced. The study aims, hypotheses, and analysis plan were pre-registered after data collection commenced, with 15 participants recruited at the point of preregistration [34]. No data had been accessed at the point of pre-registration.

Ethical Considerations

This study received approval from the ethics committee for Life Sciences and Medical Experiments at the National Institute of Advanced Industrial Science and Technology, Japan (S2025-0595). Participants will be fully informed of the study aims prior to giving their consent to participate and are paid a total of 21,000 JPY (US $136.10) for their participation in the wider project, of which this study is one part. Data are anonymous at the point of collection.

Measures

Hikikomori symptoms will be measured at baseline and postmonitoring period, with the 1-month version of the Hikikomori Questionnaire (HQ-25M) [35]. Participants rate their endorsement of statements about social withdrawal symptoms over the last month, and this has been validated in preclinical samples [35]. Items are rated on a 0 ("Strongly Disagree") to 4 ("Strongly Agree") scale, with item scores totaled so that higher scores indicate greater severity of hikikomori symptoms. This scale was developed and validated in nonclinical samples, showing positive correlations with both withdrawal behavior and psychological distress at withdrawal [35], and was considered appropriate for the nonclinical sample in this study.

Depressive symptoms will be measured with the Japanese version of the Patient Health Questionnaire-9 [36]. Participants rate the frequency of depressive symptoms experienced over the previous 2 weeks on a 0 ("Not at all") to 3 ("Almost every day") rating scale. Item responses are totaled with higher scores indicating greater depressive symptom severity. The Japanese version of the Patient Health Questionnaire-9 has been found to be in accordance with neuropsychiatric interviews [36] and suitable for screening and monitoring symptoms in primary care populations [37].

General anxiety symptoms will be measured with the Japanese version of the Generalized Anxiety Disorders Scale [38]. On 7 items, participants rated the frequency of anxiety symptoms from 0 ("Not at all") to 3 ("Every Day"). Item responses are totaled, so that higher scores indicate higher frequency of anxious symptoms. The Japanese Language scale has been validated in clinical and nonclinical samples [38].

Social anxiety symptoms will be measured with the Japanese version of the Liebowitz Social Anxiety Scale [39]. This is a 24-item scale on which participants rate their fear of social situations from 0 ("None") to 3 ("Severe"), and how frequently they avoid those situations from 0 ("Never, 0% of the time") to 3 ("Usually, 67% - 100% of the time). All item responses are summed, with higher scores indicating greater social anxiety. The Japanese version of the scale has been repeatedly investigated and validated in both clinical and nonclinical groups [39,40].

Fear of offending others will be measured with the Taijin Kyofusho Scale [41]. This scale is designed to capture a

XSL•FO
RenderX

culturally specific form of social anxiety centered on the fear of offending others with one's appearance or behavior. Participants rate their agreements with 31 statements from 1 ("Totally Disagree") to 7 ("Totally Agree"). Item responses are totaled so higher scores indicate higher levels of social anxiety and fear of offending others. The scale has been validated in clinical and nonclinical populations [41,42].

Internet addiction will be measured with the Japanese version of the Internet Addiction Test. Participants rate the frequency of internet addiction symptoms on a 0 ("Not at all") to 5 ("Always") rating scale. Item responses are totaled so that higher scores indicate more problematic internet use. The scale has been validated in Japanese populations [43] and has been frequently used in research examining the behavior of hikikomori [8].

Autism traits will be measured with the Japanese version of the Autism Spectrum Quotient short form [44]. Participants rate their agreement with statements consistent with autistic traits from "Strongly Disagree" to "Strongly Agree." Answers consistent with autistic traits, for example,"I find it difficult to understand people's intentions" rated as "Agree," are given a score of 1. Item responses are totaled, with higher scores indicating greater presence of autistic traits. The scale has been validated in clinical [44] and student samples [45].

Emotion regulation style will be measured with the Japanese version of the Emotion Regulation Questionnaire [46]. Participants rate their agreement with 10 statements about emotional regulation strategies (Suppression and Reappraisal) on a 1 ("Strongly Disagree") to 7 ("Strongly Agree") rating scale. Scores for each subscale are totaled, with greater scores indicating greater use of that strategy. The scale has been validated in Japanese samples and against measures of personality, well-being, and emotion experience [46].

EMA Measurements

Mood and emotional variability will be estimated from participant ratings on questions about their general mood ("Very negative" to "Very positive") and how anxious, irritable, sad, and happy ("Not at all" to "Very") they are at that moment on a 0 to 100 visual analog scale (VAS). Additionally, participants will be asked if they have been angry since the last beep and if so, how angry they were on a 0 to 100 VAS scale ("Not at all" to "Very"). From each mood item rating, we will calculate emotional variability as the within-person SD. Emotional stability will be calculated as the mean squared successive difference between scores. Emotional inertia will be calculated as the autocorrelation of moods across time.

Daily emotional variability will be calculated for each emotion as the SD of that emotion's ratings each day, while the overall emotional variability will be the mean of that emotion's daily emotional variability. Overall emotion rating will be calculated as the mean of that emotion's rating across the assessment period.

Anticipatory and consummatory social enjoyment will be assessed using momentary surveys. In the morning survey, participants will be asked to name an event they are looking forward to that day (if one exists), and to rate how much they expect to enjoy it on a 0 to 100 VAS scale ("Not at all" to "Very much"). In the subsequent surveys, participants will be asked whether they have had their social contact since the last survey, whether this was online or in person, how much they enjoyed it ("Not at all" to "Very much," 0 - 100), and whether they desire social contact again that day (or the next day in the evening survey). The enjoyment prediction gap will be calculated as the difference between the anticipated enjoyment rating of an event that day and the consummatory enjoyment rating of the beep after that event occurred.

Sleep quality and physical activity will be estimated from a self-report question about sleep quality in the morning survey ("Very bad" to "Very Good," 0 - 100), what time they went to bed and what time they woke up. Physical activity level will be estimated via step counter and fitness tracker, sleep patterns, duration, and quality will be estimated from the fitness tracker and supported by self-report measures.

Internet and smartphone use will be estimated from participants' self-reports on the number of minutes spent using their phone, how many minutes they have spent browsing apps, and how many minutes they have spent using the internet in the 2 hours prior to the survey. Participants will also be asked for what purposes they have used the internet and smartphone apps in that time (social networking, ordering food or other life management, gaming, shopping, to watch videos, to distract them from boredom).

Familial support questions will ask about the number of interactions with family members since the last notification and the nature of those interactions. Given the first author is not fluent in Japanese, text responses will be translated into English via translation software for coding. To check for accuracy, 10% of the events will be translated back into Japanese, and their codes will be shared with a native Japanese collaborator to check for accuracy. Coding will be both inductive and deductive, as events will either be classified as enabling or challenging a social isolation, but categories will be derived from the data. Textual responses will not be analyzed further if the context is not about social withdrawal. Each event will be assigned to one of the support styles (enabling vs challenging isolation) for statistical analysis.

Recruitment Methods

We will recruit a minimum of 84 participants to this study, as part of a wider project being advertised through the University of Tsukuba's online research recruitment portal, and professional research companies. Participants will be allocated to either "high" or "low" hikikomori symptom groups based on their responses to the HQ-25M, with those scoring ≥42 at baseline being deemed "at risk" of hikikomori, based on cutoffs established in the original scale development paper [47]. This score has been deemed to have "excellent" diagnostic accuracy, sensitivity, and positive predictive value in the original 6-month scale [35]. As items are identical between the original scale and the 1-month version, this potentially represents a valid cutoff in this study and has been used in previous research with the HQ25-M [48]. However, as an HQ25-M cutoff is yet to be validated, we will additionally run sensitivity analyses with a cutoff set as the mean HQ25-M score +1SD, consistent with

XSL•FO
RenderX

previous research [49]. Should our primary recruitment method result in a marked imbalance in group numbers after 30 participants have been recruited, we will also recruit participants online via social media and professional recruitment sites. Participants recruited online who do not meet the necessary group allocation criteria (eg, scoring ≥42 on the HQ-25M if the "high" symptom group is lacking participants) will not be eligible to take part. Participants who are eligible to take part will continue to complete the rest of the baseline measures and will be given instructions on how to download the EMA platform on their smartphone. Participants are eligible for the study if they are between the ages of 18 and 60 years, have a smartphone and a stable internet connection, and can stand up independently. Exclusion criteria for the overarching project are those: (1) affected by cardiac, neurological, cerebrovascular, renal, respiratory, or endocrine diseases; (2) currently receiving counseling or psychiatric medication; (3) who cannot exercise due to illness; (4) who work regular night shifts or irregular working hours; (5) who take drugs which may affect cortisol response; (6) who smoke; (7) who regularly use nail polish or artificial nails; (8) who are pregnant; and (9) who are experiencing irregular menstruation, amenorrhea, or menopause.

Procedure

Participants will be invited into the laboratory to complete all self-report measures and receive instructions on how to download and set up the EMA app, mPath (KU Leuven; [50]), with surveys delivered 5 times per day for 14 days. Surveys will be delivered on a fixed schedule, with the morning survey to be delivered at 8 AM; three more to be delivered at 11 AM, 2 PM, and 5 PM; and the evening survey at 8 PM. There will be a random margin of 15 minutes for the delivery of each prompt (eg, the 11 AM survey may be delivered between 11 AM and 11:15 AM). Participants have a 45-minute response window before that survey is closed and recorded as a missed response. Participants will be nudged to respond if a response has not been recorded within 15 minutes of the first notification being delivered. After the 2-week monitoring period, participants will complete the HQ-25M again.

Statistical Analysis

Sample Size Justification

Given the lack of previous research, our sample size was determined by the resources available for this project, rather than on the basis of previously reported effect sizes. Therefore, we conducted sensitivity power analyses for each of our confirmatory hypothesis tests to determine which effect sizes could be detected with 90% power at our minimum target sample size of 84. With this sample, we will be able to detect effects equivalent to r=0.34 in a correlational analysis for H1a, f^2=0.31 for H1b, and pseudo-R^2 approximately 0.21 to 0.42 for H2a to H2d. Given that these effect sizes are comparable to those found in related populations such as those with social anxiety (eg, correlation between social anxiety and perseverative thinking r=0.55 [51]; difference between hikikomori and controls on depressive symptoms, standardized mean difference=0.78; [52]), we believe this study will be informative to the wider literature on social withdrawal, as well as on hikikomori specifically [53]. Our target of 84 participants completing EMA measures at 5 time points across the 2-week data collection period would result in 70 data points per participant, and a maximum of 5880 data points.

Data Exclusion

We will check and report compliance rates for participants. We will include all participants in our primary analyses, though we will conduct separate sensitivity analyses excluding participants who complete less than 75% of the EMA measures, reporting the results in Multimedia Appendix 1. Additionally, missing data will be assessed for randomness via logistic regression. Each of our baseline and demographic variables will be entered into a multiple logistic regression model to determine whether they influence nonresponse.

Analysis Plan

Overview

Significance will be accepted when P≤.05, after the Benjamini-Hochberg false detection rate (FDR) corrections have been applied within the families of hypothesis tests (eg, H1b, H2a, and so on). For exploratory analyses, no significance threshold is set, but P values will be interpreted alongside effect sizes and CIs [54] to determine evidence of an effect of interest for future research.

RQ1: What Factors are Related to Hikikomori Symptoms?

H1a will be tested with the Pearson correlation analyses, correlating baseline hikikomori symptoms with measures of social anxiety, expressive suppression, cognitive reappraisal, and fear of offending others.

H1b will be assessed with multiple linear regression. Postmonitoring period hikikomori symptoms will be regressed on averaged daily internet usage, averaged sleep quality, averaged sleep duration, averaged physical activity, averaged familial support, averaged anticipatory and consummatory social enjoyment, averaged mood ratings, and overall emotional variability. Symptoms of depression, autism, anxiety, and hikikomori at baseline will be entered as covariates in this model. The hypothesis will be confirmed by the overall model significance and fit, while individual predictor contributions will be assessed by standardized coefficients, FDR corrected P values, and CIs. Exploratory analyses will be conducted with pre-hikikomori symptoms and averaged hikikomori symptoms as the outcome, respectively.

To address our exploratory aim, we will construct time-lagged network analyses. For each hikikomori group, multilevel vector autoregressive models (mlVARs) will be specified for each momentary variable [55,56]. The models will be constructed with 2 levels: within day (level 1) and nested within participants (level 2). Momentary mood, anger, irritability, sadness, happiness, consummatory enjoyment, internet use, and family contact at time t–1 will be entered as our level 1 predictors. The mlVAR model will be estimated using maximum likelihood estimation and taking the form:

$$Mood = \beta_0 + \beta_1 Mood_{t-1} + \sum_{k=2}^{7} \beta_k Angst_{t-1} + \gamma_1 Angst + \sum_{k=2}^{6} \gamma_k Variable_k + (1|participant)$$

where β_0 is the intercept, β_k is the coefficient for a variable k at the lagged time point t–1, γ_k is the coefficient for a variable k

XSL•FO
RenderX

at the current time point *t*, and ε_t is the error term. Coefficients and SEs will be drawn from these models to construct the time-lagged network [56]. Nodes in the model will represent the variables, while edges will represent cross-lagged and contemporaneous relationships.

Edge weights and centrality measures will be compared between groups with 2-tailed *t* tests, though given the lack of previous research, we will not report these results as confirmatory hypothesis tests. Rather, these results will be interpreted to develop a theory and guide future research efforts.

RQ2: How do Internet, Social, and Familial Behaviors Predict Future Social Behavior and Enjoyment?

The hypotheses for this research question will be addressed with four multilevel autoregressive logistic regressions:

1. H2a predictors: consummatory pleasure rating at *t–1*
2. H2b predictors: internet time and usage reason at *t–1*
3. H2c predictors: familial support amount and style (enabling vs challenging) at *t–1*
4. H2d predictors: mood and anger ratings at *t–1*

Regression models will additionally include age, gender, and employment status as covariates, and baseline hikikomori symptoms as moderators. Hikikomori symptom scores will be mean-centered, and significant interaction terms will be examined via a simple slopes analysis. For example, the model for H2b will be:



where β_0 is the intercept, β_k is the coefficient for a variable *k* at the lagged time point *t*–1, *i* represents the participant, and ε_i is the error term. Hypotheses will be confirmed by overall model significance and fit, while the individual predictor contributions will be assessed by odds ratios, FDR corrected *P* values, and CIs.

Exploratory Analyses

We will conduct exploratory multilevel logistic regressions to determine if averaged daily mood, internet use, consummatory enjoyment, and familial support predict future social event intention in the evening survey.

To investigate whether hikikomori symptoms are related to social reward feedback, we will explore whether social enjoyment prediction gap at *d*–1 is related to the change in anticipatory pleasure rating on the morning survey of the current day *d* from the previous day *d*–1, and whether hikikomori symptoms moderate this relationship.

Results

This project was funded in May 2025. Data collection began on November 21, 2025, with 15 participants enrolled at the point of preregistration. Results of the study are expected to be available by the end of summer 2026, and no data have been accessed at the time of this paper's submission.

Discussion

Interpretations and Implications of Results

The results of RQ1 will allow us to identify which factors may be most influential in hikikomori. First, the results of our regression will indicate which factors relate to, and predict, hikikomori symptoms after controlling for the symptoms of depression, anxiety, and autism.

By conducting network analyses in high and low symptom groups and comparing the centrality and edge weights, we will be able to identify influential factors in hikikomori, as well as understand how the strength and structure of networks differ at low and high levels of symptoms, which could have implications when developing new treatments.

The results of RQ2 will allow us to identify a candidate mechanism underlying hikikomori and identify maintenance factors for intervention. Results from the model predicting social enjoyment have implications when positioning hikikomori as a consequence of social anhedonia. Should social enjoyment interact with hikikomori symptoms and predict future social behavior (H2a), this would suggest that social anhedonia is one mechanism underlying withdrawal. This position will be further explored with our exploratory analyses investigating whether the social enjoyment–prediction gap update size predicts change in anticipatory pleasure ratings. A null result in this model could suggest a lack of reinforcement learning from social contexts, indicating that reward processing both at the anticipatory and learning stages is impaired, as in depression [21]. A significant result here would suggest normal learning from a social context.

Subsequent models (H2b-H2d) would identify maintenance factors of hikikomori, as well as suggest targets for cognitive or behavioral intervention. These models are expected to be significant given prior correlational research [1,23]; however, this study will be the first to explore these relationships longitudinally. Should null results be found, this would be unexpected but would have implications for current treatment strategies.

In summary, RQ1 explores the relationship between previously identified psychological and social risk factors and self-reported hikikomori traits, while controlling for mental health symptoms frequently associated with hikikomori and hikikomori risk factors. RQ2 investigates the temporal dynamics of these relationships, seeking to validate previously suggested relationships (eg, between internet use and social withdrawal), and mechanisms proposed in other disorders that have yet to be explored in relation to hikikomori (eg, social consummatory pleasure). Taken together, the results of this study will identify important psychological variables that may underlie hikikomori traits and behaviors that may be targeted for early intervention.

Limitations

The study has several limitations. First, though we are drawing inferences based on symptoms of hikikomori, we are not recruiting a clinical population for this study, and so results may not generalize to clinical samples. For example, given that the isolation in clinical samples will be more marked and prevalent, it may be that the associations found in this study, especially

XSL•FO
RenderX

when examining RQ2, are not replicated. This decision was made for practical reasons, and we will draw conclusions in line with this limitation, specifically how these results may be most applicable to those "at risk" or pre-hikikomori. Additionally, we will propose future work in clinical samples. Second, though items are identical between the original scale and 1-month version of the scale, our chosen cutoff score has not been validated. Further psychometric investigation of the HQ25-M is required to determine if the cutoff of 42 is valid. Third, there are problems inherent in interpreting causation in our time-varying models given that we are establishing a Granger causality rather than causality through experimental manipulation. In addition, our analysis plan specifies an mlVAR model where time is treated as a discrete variable and assumes that measurements are taken at equidistant intervals, which is unlikely given the laxity in both survey delivery and response windows and so may bias estimates [56]. We also have specified a 1-step lag in our models, which may exclude important connections at longer lags [57].

We also must acknowledge the modest target sample size for this study, which is primarily driven by time and resource constraints. Though we have attempted to lend some context on what effect sizes may be detectable, given the lack of previous work using EMA methods or our planned analyses, we were not able to conduct either an *a priori* power analysis to inform a target sample size, nor offer more specific potential effect sizes. We acknowledge the risk of a moderate sample size reducing the positive predictive value of our findings [58], though again reiterate the value in this study both in generating new avenues of future research and in testing confirmatory hypotheses based on previous literature. Finally, though a strength of this study is the EMA methodology, we are not including several passive sensing measures in our protocol (eg, location tracking, app and phone usage data, and so on). These data can be collected without any additional participant burden and have been found to predict episodes of mood disorders (eg, depression [59]), though this evidence is mixed in studies investigating social isolation [60] and both the privacy and ethical concerns must be carefully managed in such studies.

Conclusions

This will be, to our knowledge, the first study investigating the temporal dynamics of proposed maintenance factors of hikikomori and social behavior. Additionally, this study aims to identify influential cognitive, affective, and behavioral factors underlying hikikomori. These results will test and validate previously identified factors from case studies and cross-sectional work and open up avenues for future research in clinical or general samples.

Funding

This study was funded by a Japan Society for the Promotion of Science (JSPS) postdoctoral fellowship (25KF0093) and JSPS grant-in-aid (KAKENHI; 23K25741). The JSPS had no role in the study development, administration, or manuscript.

Authors' Contributions

Conceptualization: AM, KT
Methodology: AM
Project administration: AM
Writing – original draft: AM
Writing – review and editing: AM, KT

Conflicts of Interest

None declared.

Multimedia Appendix 1
Revised sample size estimation and ecological momentary assessment (EMA) study reporting guidelines.
[DOCX File, 20 KB - resprot_v15i1e81384_app1.docx]

Checklist 1
CREMAS reporting checklist.
[DOCX File, 16 KB - resprot_v15i1e81384_app2.docx]

References

1. Kato TA, Kanba S, Teo AR. Hikikomori: multidimensional understanding, assessment, and future international perspectives. Psychiatry Clin Neurosci 2019 Aug;73(8):427-440. [doi: 10.1111/pcn.12895] [Medline: 31148350]
2. Amendola S, Cerutti R, Presaghi F. Symptoms of prolonged social withdrawal, problematic internet use, and psychotic-like experiences in emerging adulthood: a moderation model. Clin Neuropsychiatry 2023 Feb;20(1):29-38. [doi: 10.36131/cnfioritieditore20230104] [Medline: 36936621]
3. Yamada M, Kato TA, Katsuki RI, et al. Pathological social withdrawal in autism spectrum disorder: a case control study of hikikomori in Japan. Front Psychiatry 2023;14:1114224. [doi: 10.3389/fpsyt.2023.1114224] [Medline: 37032930]

XSL•FO
RenderX

4. Orsolini L, Bellagamba S, Volpe U, Kato TA. Hikikomori and modern-type depression in Italy: a new phenotypical trans-cultural characterization? Int J Soc Psychiatry 2022 Aug;68(5):1010-1017. [doi: 10.1177/00207640221099408] [Medline: 35723190]
5. Paulino Ferreira L, Magalhaes M, Figueiredo MM, Conde E. Beyond the locked door: hikikomori through a clinical lens. Cureus 2025 Feb;17(2):e78600. [doi: 10.7759/cureus.78600] [Medline: 39926630]
6. Nonaka S, Sakai M. Psychological factors associated with social withdrawal (hikikomori). Psychiatry Investig 2021 May;18(5):463-470. [doi: 10.30773/pi.2021.0050] [Medline: 34053211]
7. Sherdell L, Waugh CE, Gotlib IH. Anticipatory pleasure predicts motivation for reward in major depression. J Abnorm Psychol 2012 Feb;121(1):51-60. [doi: 10.1037/a0024945] [Medline: 21842963]
8. Tateno M, Teo AR, Ukai W, et al. Internet addiction, smartphone addiction, and hikikomori trait in Japanese young adult: social isolation and social network. Front Psychiatry 2019;10:455. [doi: 10.3389/fpsyt.2019.00455] [Medline: 31354537]
9. Malagón-Amor Á, Martín-López LM, Córcoles D, et al. Family features of social withdrawal syndrome (hikikomori). Front Psychiatry 2020;11:138. [doi: 10.3389/fpsyt.2020.00138] [Medline: 32194459]
10. Yuen JWM, Yan YKY, Wong VCW, Tam WWS, So KW, Chien WT. A physical health profile of youths living with a "hikikomori" lifestyle. Int J Environ Res Public Health 2018 Feb 11;15(2):315. [doi: 10.3390/ijerph15020315] [Medline: 29439488]
11. Frankova I. Similar but different: psychological and psychopathological features of primary and secondary hikikomori. Front Psychiatry 2019;10:558. [doi: 10.3389/fpsyt.2019.00558] [Medline: 31447713]
12. Lemay EP, Overall NC, Clark MS. Experiences and interpersonal consequences of hurt feelings and anger. J Pers Soc Psychol 2012 Dec;103(6):982-1006. [doi: 10.1037/a0030064] [Medline: 22984830]
13. Kashdan TB, Collins RL. Social anxiety and the experience of positive emotion and anger in everyday life: an ecological momentary assessment approach. Anxiety Stress Coping 2010 May;23(3):259-272. [doi: 10.1080/10615800802641950] [Medline: 19326272]
14. Phillips LH, Henry JD, Hosie JA, Milne AB. Age, anger regulation and well-being. Aging Ment Health 2006 May;10(3):250-256. [doi: 10.1080/13607860500310385] [Medline: 16777652]
15. Hauffe V, Vierrath V, Tuschen-Caffier B, Schmitz J. Daily-life reactivity and emotion regulation in children with social anxiety disorder. J Anxiety Disord 2024 Aug;106:102907. [doi: 10.1016/j.janxdis.2024.102907] [Medline: 39059189]
16. Houben M, Van Den Noortgate W, Kuppens P. The relation between short-term emotion dynamics and psychological well-being: a meta-analysis. Psychol Bull 2015 Jul;141(4):901-930. [doi: 10.1037/a0038822] [Medline: 25822133]
17. Boemo T, Nieto I, Vazquez C, Sanchez-Lopez A. Relations between emotion regulation strategies and affect in daily life: a systematic review and meta-analysis of studies using ecological momentary assessments. Neurosci Biobehav Rev 2022 Aug;139:104747. [doi: 10.1016/j.neubiorev.2022.104747] [Medline: 35716875]
18. Dryman MT, Heimberg RG. Emotion regulation in social anxiety and depression: a systematic review of expressive suppression and cognitive reappraisal. Clin Psychol Rev 2018 Nov;65:17-42. [doi: 10.1016/j.cpr.2018.07.004] [Medline: 30064053]
19. Barkus E, Badcock JC. A transdiagnostic perspective on social anhedonia. Front Psychiatry 2019;10:216. [doi: 10.3389/fpsyt.2019.00216] [Medline: 31105596]
20. Zhang RT, Yang ZY, Wang YM, et al. Affective forecasting in individuals with social anhedonia: the role of social components in anticipated emotion, prospection and neural activation. Schizophr Res 2020 Jan;215:322-329. [doi: 10.1016/j.schres.2019.10.006] [Medline: 31611042]
21. Halahakoon DC, Kieslich K, O'Driscoll C, Nair A, Lewis G, Roiser JP. Reward-processing behavior in depressed participants relative to healthy volunteers: a systematic review and meta-analysis. JAMA Psychiatry 2020 Dec 1;77(12):1286-1295. [doi: 10.1001/jamapsychiatry.2020.2139] [Medline: 32725180]
22. Pike AC, Robinson OJ. Reinforcement learning in patients with mood and anxiety disorders vs control individuals: a systematic review and meta-analysis. JAMA Psychiatry 2022 Apr 1;79(4):313-322. [doi: 10.1001/jamapsychiatry.2022.0051] [Medline: 35234834]
23. Muris P, van der Veen A, Luijten B, de Bie C, Meesters C. On your own: an explorative study on the psychopathological and psychosocial correlates of hikikomori symptoms in dutch adolescents and young adults. Child Psychiatry Hum Dev 2025 Mar 18. [doi: 10.1007/s10578-025-01828-0] [Medline: 40100551]
24. Shah H, Helmy M, Vally Z. Hikikomori in the Middle East: the role of problematic gaming, social media use, and loneliness. PLoS One 2024;19(10):e0312818. [doi: 10.1371/journal.pone.0312818] [Medline: 39471202]
25. Cangelosi G, Biondini F, Sguanci M, et al. Anhedonia in youth and the role of internet-related behavior: a systematic review. Psychiatry Int 2025;6(1):1. [doi: 10.3390/psychiatryint6010001]
26. Shen W, Hua M, Wang M, Yuan Y, Shao M, Zhao Y. Anhedonia mediates the link between problematic internet use and psychological well-being. Curr Psychol 2023 Jan;42(3):2023-2030. [doi: 10.1007/s12144-021-01575-0]
27. Teo AR. A new form of social withdrawal in Japan: a review of hikikomori. Int J Soc Psychiatry 2010 Mar;56(2):178-185. [doi: 10.1177/0020764008100629] [Medline: 19567455]
28. Hattori Y. Social withdrawal in Japanese youth: a case study of thirty-five hikikomori clients. J Trauma Pract 2006;4(3–4):181-201. [doi: 10.1300/J189v04n03_01]

XSL•FO
RenderX

29. Nonaka S, Shimada H, Sakai M. Family behavioral repertoires and family interaction influence the adaptive behaviors of individuals with Hikikomori. Front Psychiatry 2020;10:977. [doi: 10.3389/fpsyt.2019.00977] [Medline: 32009999]
30. Amendola S, Hengartner MP. From loneliness to severe social isolation and hikikomori. Asian J Psychiatr 2024 Jul;97:104084. [doi: 10.1016/j.ajp.2024.104084] [Medline: 38815438]
31. Kuczynski AM, Piccirillo ML, Dora J, et al. Characterizing the momentary association between loneliness, depression, and social interactions: insights from an ecological momentary assessment study. J Affect Disord 2024 Sep 1;360:376-386. [doi: 10.1016/j.jad.2024.05.148] [Medline: 38823593]
32. Wu H, Mata J, Furman DJ, Whitmer AJ, Gotlib IH, Thompson RJ. Anticipatory and consummatory pleasure and displeasure in major depressive disorder: an experience sampling study. J Abnorm Psychol 2017 Feb;126(2):149-159. [doi: 10.1037/abn0000244] [Medline: 27936838]
33. Exploring cognitive and behavioural risk and maintenance factors of hikikomori. Open Science Framework. URL: https://osf.io/bmyn6 [accessed 2026-01-30]
34. MacLellan A, Takano K. Exploring cognitive and behavioural risk and maintenance factors of Hikikomori: Protocol for an Ecological Momentary Assessment study. 2025 Dec 9. URL: https://doi.org/10.17605/OSF.IO/AFRSB [accessed 2026-02-04]
35. Kato TA, Suzuki Y, Horie K, Teo AR, Sakamoto S. One month version of hikikomori questionnaire-25 (HQ-25M): development and initial validation. Psychiatry Clin Neurosci 2023 Mar;77(3):188-189. [doi: 10.1111/pcn.13499] [Medline: 36305452]
36. Muramatsu K, Miyaoka H, Kamijima K, et al. The patient health questionnaire, Japanese version: validity according to the mini-international neuropsychiatric interview-plus. Psychol Rep 2007 Dec;101(3 Pt 1):952-960. [doi: 10.2466/pr0.101.3.952-960] [Medline: 18232454]
37. Muramatsu K, Miyaoka H, Kamijima K, et al. Performance of the Japanese version of the Patient Health Questionnaire-9 (J-PHQ-9) for depression in primary care. Gen Hosp Psychiatry 2018;52:64-69. [doi: 10.1016/j.genhosppsych.2018.03.007] [Medline: 29698880]
38. Muramatsu K. Validation and utility of a Japanese version of the GAD‐7. Jpn J Psychosom Med 2009;51:79 [FREE Full text]
39. Asakura S, Inoue S, Sasaki F, et al. Reliability and validity of the Japanese version of the Liebowitz Social Anxiety Scale [Article in Japanese]. Clinical Psychiatry 2002;44(10):1077-1084. [doi: 10.11477/mf.1405902721]
40. Sugawara N, Yasui-Furukori N, Kaneda A, et al. Factor structure of the Liebowitz Social Anxiety Scale in community-dwelling subjects in Japan. Psychiatry Clin Neurosci 2012 Oct;66(6):525-528. [doi: 10.1111/j.1440-1819.2012.02381.x] [Medline: 22988811]
41. Kleinknecht RA, Dinnel DL, Kleinknecht EE, Hiruma N, Harada N. Cultural factors in social anxiety: a comparison of social phobia symptoms and taijin kyofusho. J Anxiety Disord 1997;11(2):157-177. [doi: 10.1016/s0887-6185(97)00004-2] [Medline: 9168340]
42. Fan H, Cai Y, Wang S, Tu D. Development and validation of static short form and adaptive test for the Taijin Kyofusho Scale to measure the severity of culture-bound social anxiety. Curr Psychol 2023 Jan;42(1):542-552. [doi: 10.1007/s12144-021-01497-x]
43. Osada H. Internet addiction in Japanese college students: is Japanese version of Internet Addiction Test (JIAT) useful as a screening tool? Bulletin of Senshu University school of human sciences [Article in Japanese]. Psychology (Irvine) 2013;3:71-80 [FREE Full text] [doi: 10.34360/00005228]
44. Kurita H, Koyama T, Osada H. Autism-Spectrum Quotient-Japanese version and its short forms for screening normally intelligent persons with pervasive developmental disorders. Psychiatry Clin Neurosci 2005 Aug;59(4):490-496. [doi: 10.1111/j.1440-1819.2005.01403.x] [Medline: 16048456]
45. Maeda Y, Kanayama H, Sato H. Investigation on the autistic spectrum tendency of university student: using AQ-J-10 [Article in Japanese]. Kansai University Psychology Research 2017;8:23-29 [FREE Full text]
46. Yoshizu J, Sekiguchi R, Amemiya T. Development of a Japanese version of Emotion Regulation Questionnaire [Article in Japanese]. Japanese Journal of Research on Emotions 2013;20(2):56-62. [doi: 10.4092/jsre.20.56]
47. Teo AR, Chen JI, Kubo H, et al. Development and validation of the 25-item hikikomori questionnaire (HQ-25). Psychiatry Clin Neurosci 2018 Oct;72(10):780-788. [doi: 10.1111/pcn.12691] [Medline: 29926525]
48. Lee S, Kim HS, Hong J, et al. Network structure of social withdrawal symptoms in Asian psychiatric patients at high risk of hikikomori: findings from the REAP-AD3. Asian J Psychiatr 2025 Jun;108:104489. [doi: 10.1016/j.ajp.2025.104489] [Medline: 40250201]
49. Zhang W, Chen MY, A LY, et al. Gender difference in prevalence and network structure of subclinical hikikomori and depression among college students. Int J Soc Psychiatry 2025 Sep;71(6):1114-1126. [doi: 10.1177/00207640251325059] [Medline: 40119503]
50. Mestdagh M, Verdonck S, Piot M, et al. m-Path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. Front Digit Health 2023;5:1182175. [doi: 10.3389/fdgth.2023.1182175] [Medline: 37920867]

XSL•FO
RenderX

51. Adamis AM, Walske S, Olatunji BO. Attention mechanisms of social anxiety in daily life: unique effects of negative self-focused attention on post-event processing. Behav Res Ther 2025 Aug;191:104759. [doi: 10.1016/j.brat.2025.104759] [Medline: 40318602]
52. Nonaka S, Kubo H, Takeda T, Sakai M. Functioning, disability, and health of individuals with hikikomori (prolonged social withdrawal) and their families: a systematic review and meta-analysis of case-control studies. Int J Soc Psychiatry 2025 Jun;71(4):622-641. [doi: 10.1177/00207640241310189] [Medline: 39840577]
53. Lakens D. Sample size justification. Collabra 2022 Mar 22;8(1):33267. [doi: 10.1525/collabra.33267]
54. Muff S, Nilsen EB, O'Hara RB, Nater CR. Response to "why P values are not measures of evidence" by D. Lakens. Trends Ecol Evol 2022 Apr;37(4):291-292. [doi: 10.1016/j.tree.2022.01.001] [Medline: 35078625]
55. Bringmann LF, Vissers N, Wichers M, et al. A network approach to psychopathology: new insights into clinical longitudinal data. PLoS One 2013;8(4):e60188. [doi: 10.1371/journal.pone.0060188] [Medline: 23593171]
56. Vanhasbroeck N, Ariens S, Tuerlinckx F, Loossens T. Computational models for affect dynamics. In: Waugh CE, Kuppens P, editors. Affect Dynamics: Springer International Publishing; 2021:213-260.
57. Shojaie A, Fox EB. Granger causality: a review and recent advances. Annu Rev Stat Appl 2022 Mar;9(1):289-319. [doi: 10.1146/annurev-statistics-040120-010930] [Medline: 37840549]
58. Button KS, Ioannidis JPA, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013 May;14(5):365-376. [doi: 10.1038/nrn3475] [Medline: 23571845]
59. Saeb S, Zhang M, Karr CJ, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015 Jul 15;17(7):e175. [doi: 10.2196/jmir.4273] [Medline: 26180009]
60. Qirtas MM, Zafeiridi E, Pesch D, White EB. Loneliness and social isolation detection using passive sensing techniques: scoping review. JMIR mHealth uHealth 2022 Apr 12;10(4):e34638. [doi: 10.2196/34638] [Medline: 35412465]

Abbreviations

EMA: ecological momentary assessment
FDR: false detection rate
HQ-25M: Hikikomori Questionnaire, 1-month version
RQ: research question
VAS: visual analog scale

Edited by J Sarvestan; submitted 27.Jul.2025; peer-reviewed by S Nonaka; revised version received 11.Dec.2025; accepted 12.Dec.2025; published 17.Feb.2026.

Please cite as:
MacLellan A, Takano K
Exploring the Cognitive and Behavioral Risks and Maintenance Factors of Hikikomori: Protocol for an Ecological Momentary Assessment Study
JMIR Res Protoc 2026;15:e81384
URL: https://www.researchprotocols.org/2026/1/e81384
doi:10.2196/81384

© Alexander MacLellan, Keisuke Takano. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Developing a Culturally Adapted Digital Health Application for Older Hispanic Adults With Type 2 Diabetes: Protocol for a Qualitative and Pilot Study

Joshua Caballero[1], PharmD; Raymond L Ownby[2], MBA, MD, PhD; Henry Nolan Young[1], PhD; Michelle B McElhannon[3], PharmD; N M Mahmudul Alam Bhuiya[1], MPharm; Kenny H Esho[4], MD; Russ H Palmer[5], PhD

[1]Department of Clinical and Administrative Pharmacy, College of Pharmacy, University of Georgia, 250 West Green Street, Athens, GA, United States

[2]Department of Psychiatry and Behavioral Medicine, Dr. Kiran C. Patel College of Osteopathic Medicine, Nova Southeastern University, Fort Lauderdale, FL, United States

[3]Division of Experience Programs, College of Pharmacy, University of Georgia, Athens, GA, United States

[4]Athens Neighborhood Health Center, Athens, GA, United States

[5]Division of Instructional Innovation and Research, College of Pharmacy, University of Georgia, Athens, GA, United States

Corresponding Author:
Joshua Caballero, PharmD
Department of Clinical and Administrative Pharmacy, College of Pharmacy, University of Georgia, 250 West Green Street, Athens, GA, United States

Abstract

Background: Over 40% of Hispanic individuals have below basic health literacy levels, significantly impacting their ability to manage chronic conditions such as type 2 diabetes (T2D). Chronic disease self-management, essential for improving adherence, is often hindered by modifiable factors such as health literacy. Ambulatory care clinical (ACC) pharmacists, through comprehensive medication management (CMM), play a critical role in addressing these barriers; however, current methods for delivering CMM vary significantly, limiting effectiveness. Digital technology may offer significant potential for improving medication adherence, particularly among Hispanic individuals who use smartphones to seek health information.

Objective: This study aims to revise, develop, and pilot test a culturally adapted, individually tailored digital health app designed to enhance health literacy and medication adherence among older Hispanic adults with T2D by integrating interactive educational modules into pharmacist-delivered CMM.

Methods: Our team previously developed a prototype computer-delivered, culturally adapted intervention targeting health literacy and adherence among older Hispanic individuals. The proposed project involves transforming this intervention into a modernized digital health app. We will achieve 2 specific aims. The first aim is to refine and modernize previously developed content into an interactive digital health app suitable for older Hispanic adults with T2D and ACC pharmacists. A total of 20 patients and 10 ACC pharmacists will review and provide feedback on the app modules. Usability and acceptability will be measured using validated tools, including the System Usability Scale, Adjectival Ease of Use Scale, and the technology acceptance model. The second aim is to conduct a pilot test with 40 Hispanic adults (aged ≥50 years) diagnosed with T2D. Participants will be recruited from a local Federally Qualified Health Center where ACC pharmacists manage diabetes care. Patients will complete a baseline CMM session, interact with the app modules, and return for 1 or 2 CMM follow-up visits. We will assess medication adherence, hemoglobin A_{1c}, and fasting glucose levels as primary outcomes.

Results: The digital app has been modernized and completed for acceptability and usability testing. We will begin testing the hypothesis that the app will have above-average usability and acceptability (System Usability Scale: mean score ≥71 and Adjectival Ease of Use Scale: mean score ≥5), with an overall goal that the digital app will be acceptable (via the technology acceptance model) by ≥90% of the patient participants and ACC pharmacists. In addition, anticipated health outcomes for aim 2 include improvements in medication adherence (>80%), hemoglobin A_{1c} (≥0.75% reduction), and fasting glucose (≥25% reduction) from baseline.

Conclusions: When the study is completed, it is expected that the culturally adapted digital health app will be acceptable and usable among both patients and pharmacists. Successful results will enable wider dissemination of the refined app and establish a framework adaptable to other chronic diseases, diverse languages, cultures, and health care settings.

International Registered Report Identifier (IRRID): PRR1-10.2196/76294

(JMIR Res Protoc 2026;15:e76294) doi:10.2196/76294

XSL•FO
RenderX

KEYWORDS

pharmacist; diabetes; digital app; health literacy; Hispanics; Latinx

Introduction

Epidemiology

The Centers for Disease Control and Prevention estimates that over 34 million people in the United States have diabetes [1]. Readmission data from the Centers for Medicare and Medicaid Services show that patients with diabetes as a comorbidity need to be targeted for high-cost and disease management [2]. Diabetes has the third-highest 30-day all-cause readmission rate among Medicaid patients, resulting in nearly 34,000 readmissions and at a cost of approximately US $471 million [3]. In addition, the prevalence of diabetes is higher among minority groups (eg, Hispanic populations). Data show that state-insured Hispanic individuals with type 2 diabetes (T2D) have poorer glycemic control and adherence to newly prescribed diabetes medications (oral or insulin) compared to non-Hispanic White populations [4]. Diabetes medication adherence rates also continue to decrease in the United States. Data suggest that patients with diabetes who had an unscheduled readmission within 90 days were more likely to be lower-income minorities on public insurance [5,6]. As a result, minority patients with T2D are more likely to have poorer health outcomes, higher risk of hospitalizations, and higher economic burden [1,2,7].

Tailored Information and Multimedia Instruction as Interventions

The primary method of provider-patient communication about medications has traditionally been brief oral communications during often hurried medical consultations. Studies suggest that patient retention of information in this context is poor, often less than 50% [8]. The introduction of multimedia presentation of material may increase learning and retention and may be helpful for various conditions and populations [9-12]. Unfortunately, the results of using eHealth interventions for T2D have been mixed, including those focused on Hispanic populations [13-15]. Limitations that commonly plague eHealth interventions in T2D are (1) lack of cultural adaptation for minorities (such as Hispanic individuals); (2) unknown usability or operability by patients and providers; (3) educational health literacy component that is predominantly mono-directional; (4) lack of tailoring to the individual; (5) education principles of multimedia learning are not used; (6) lack of general medication information, or such information is provided by nonspecialists; (7) cost-prohibitive; and (8) lack of sufficient security of patients' personal health information [13-15]. A recent National Institutes of Health (NIH)-funded project (R01DK112322) developed an eHealth intervention using text messaging that achieved a notable 1% reduction in hemoglobin A_{1c} (HbA_{1c}) but did not show improvements in medication adherence [14,16]. While promising, the intervention encountered certain practical challenges, such as limited capacity for individual tailoring and potential security concerns associated with a cloud-based platform [16]. Accordingly, our digital health app aims to address these factors and build on the progress shown by previous interventions.

Hispanic Smartphone and Digital Health Information Use

Similar to non-Hispanic White people, approximately 85% of Hispanic people use a smartphone daily [17]. Hispanic adults are outpacing all other groups in smartphone use growth rate. In addition, similar to other groups, Hispanic adults spend approximately 3 hours per day on a smartphone and 1 hour per day on a tablet. Hispanic individuals are more likely to use smartphones for health-related information compared to non-Hispanic White or Black populations (73% vs 58% - 67%) and more likely to view educational content (45% vs 26% - 32%) [18]. Unfortunately, misinformation across many health topics (eg, noncommunicable diseases and medical treatment) ranges between 30% and 87% [19]. In addition, the ability to interpret information could be a problem, as identified by Powell et al [20], who found that patients often misinterpret online reputable diabetes information (ie, American Diabetes Association myths). The proposed study will enable the investigators to use available technology to provide patients access to correct information, while also enabling health care providers to assess patient interpretation of the information to ensure patients understand. The use of digital apps has become common among all populations in the United States, with Hispanic individuals reporting using an average of approximately 27.9 apps and spent 41.5 hours using mobile applications per month. [21]. Therefore, Hispanic individuals may be more likely to use a digital health app to educate themselves. The digital health app proposed may improve T2D health literacy, promote patient engagement, and allow for precision counseling. Precision counseling is a personalized approach to patient education and adherence that considers social and environmental stressors, medication side effects, and health literacy to tailor interventions for each patient [22]. Similar to precision medicine, it leverages technology to optimize counseling strategies, provide more effective treatment, and potentially improve health outcomes.

Pharmacists' Comprehensive Medication Management (CMM) Services

Data show that health care providers may not adequately promote medication adherence due to a lack of time or by failing to use an effective communication format for the patient (eg, appropriate reading level, language) [23]. There is an increasing role for pharmacists in assisting patients with T2D. Section 10,328 of the Affordable Care Act amended section 1860D-4(c)(2)(ii) of the Act and requires prescription drug plan sponsors (Part D plans) to offer an annual medication review as part of the comprehensive medication management (CMM). CMM is a standard of care that guarantees each medication is individually assessed for appropriateness, effectiveness, and safety for the individual patient. Using CMM can improve chronic disease self-management and related outcomes. Some data suggest pharmacists can positively impact adherence and decrease HbA_{1c} (a measure of blood glucose control over time) through CMM [24]. However, other studies on the benefits of

XSL•FO
RenderX

CMM are inconclusive because of limitations or failure to demonstrate efficacy in chronic disease self-management [25,26]. In these studies, researchers have used digital interventions that may not be readily comparable or optimally defined [24,25]. Limitations of these studies have included time constraints, lack of Hispanic populations, lack of integration into normal workflow, and absence of innovative digital eHealth technology (eg, individual tailoring) [27,28]. Another key limitation in these studies is their focus on adherence as an isolated problem. Given that factors affecting adherence and health outcomes are multidimensional, the content we developed and culturally adapted is multifaceted [22]. It is based not only on providing tailored content, but also on assessing modifiable factors of adherence (eg, depression). As a result, it can enable pharmacists to provide precision counseling to influence areas in which a patient shows a deficiency and be modified based on various factors of adherence.

Synopsis of Current Culturally Adapted Intervention

We have demonstrated that health literacy is impacted by cognitive abilities, academic skills, and health knowledge [29]. Based on these results, we developed and culturally adapted a health intervention for older Hispanic adults with T2D using Adobe Captivate [22]. This intervention suggests favorable results regarding usability among patients and pharmacists providing CMM. Overall, there are 3 modules that provide information, verify understanding of content with queries, and allow the patient to tailor their education (eg, knowledge about medications they take). There are also queries on other factors impacting adherence (eg, depression and cost) that will assist the provider in addressing or referring the patient to other services (eg, mental health support). This differs from other digital technologies that do not incorporate all these measures. For example, some tools take an innovative approach by using issue cards to explain medication classes but do not appear to verify knowledge gained or explain how the tool can be linked to prescribers managing their diabetes in a simple format [30]. Another novel digital app focused on Spanish speakers sends 2 - 3 motivational texts per day, but also does not appear to collect knowledge gained or provide specific information regarding their medications [14,31].

Aims and Hypotheses

There are 2 aims in the proposed study. Aim 1 consists of refining and modernizing developed diabetes content from our previous study [22], developing the Patient and Provider apps, and testing the usability and acceptability by patients and pharmacists. The Aim 1 hypothesis is that the digital apps (ie, Patient and Provider apps) will have above-average usability and acceptability (System Usability Scale [SUS] mean score ≥71 and Adjectival Ease of Use Scale [AEOUS] mean score ≥5) with an overall goal that the digital apps will be acceptable (Davis' technology acceptance model [TAM] ≥90%) to the participant patients and pharmacists.

The goal for Aim 2 is to pilot test the digital health apps to assess adherence and health outcomes in 40 Hispanic patients with T2D. The Aim 2 hypothesis is that the digital health apps will improve adherence (>80% average adherence for the individual), HbA_{1c} (≥0.75% reduction), and fasting glucose (≥25% reduction) from baseline.

Methods

Study Design Overview

The study consists of 3 phases, with Phase I and II addressing Aim 1 and Phase III addressing Aim 2 (Figure 1). During Phase I, the previously developed content will be refined and modernized. Both digital health apps will also be developed. In Phase II, the usability and acceptability testing will be completed for patients and pharmacists. During Phase III, Aim 2 will be assessed by pilot testing the digital health app to assess medication adherence and health outcomes.

Figure 1. Development of patient and provider diabetes digital health apps.

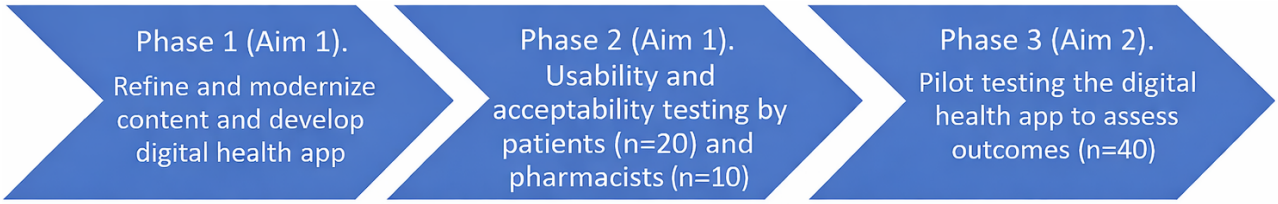

Phase I

During this phase, we want to inform the design of the digital health app with evidence-based instructional principles. Therefore, to maximize the effectiveness of the digital health app, the content should promote and embrace effective learning. As a result, content will be refined to ensure the educational component is bidirectional, educational principles are used to maximize multimedia learning, general medication information is included, and cost and security concerns are addressed. Overall, the content will be improved for digital learning. The content will then be transferred into the Patient digital health app, containing 3 modules. While developing the content, we will adapt Gagne's Nine Events of Instruction (GNEOI) as an overarching evidence-based strategy to support the effectiveness of the new app [32]. GNEOI is aligned with best practices derived from cognitive psychology and information processing theory. It includes the following sequence of events for organizing learning: (1) gain learners' attention, (2) inform learners of objectives, (3) stimulate recall of previous knowledge, (4) present the content, (5) provide learning guidance, (6) elicit performance, (7) provide feedback, (8) assess performance, and (9) enhance retention and transfer. Within the overarching GNEOI strategy, our team will modernize and improve the multimedia learning of the app by tailoring the content following Mayer's 12 principles of cognitive theory of multimedia learning [33]. These principles, also firmly grounded

XSL•FO
RenderX

in information processing theory, include providing text, graphics, and narration that support learning goals, keeping visuals simple, avoiding distractions (eg, unnecessary background music), using human voiceover, and adhering to the redundancy principle (eg, allowing the patient to select graphics and narration vs graphics and text). We will use whiteboard animation software to convert written content into approximately 30 - 60–second videos. Whiteboard animation consists of using simple animation to teach concepts that follow many of Mayer's 12 principles. Data suggest whiteboard animation can assist with increasing patient and caregiver information retention [34]. These animations can be tailored to different populations (eg, by changing skin tones or language and adding closed captioning). By following Mayer's 12 principles, the content will support knowledge construction in which the patient is building a coherent mental representation from the information presented. Please note that for Phase I, no participant recruitment will be needed.

Module Content

There are 3 modules guided by the enhanced information-motivation-behavioral (IMB) model for diabetes, which allows for bidirectional learning and fits squarely within the overarching GNEOI strategy. The IMB skills model provides an actionable plan to design and implement digital health interventions [35,36]. For example, the patient is asked a question about information they just saw or read. If they answer correctly, the module congratulates them and moves on to another topic; if they do not, the module summarizes the content in a follow-up slide or video and asks the question again. (Figure 2). Overall, the IMB model proposes that if individuals receive the information they need, feel motivated to modify their actions, and apply the necessary skills, it may provide sustained behavior change that promotes adherence and health outcomes [35,36].

Figure 2. Bidirectional learning model.

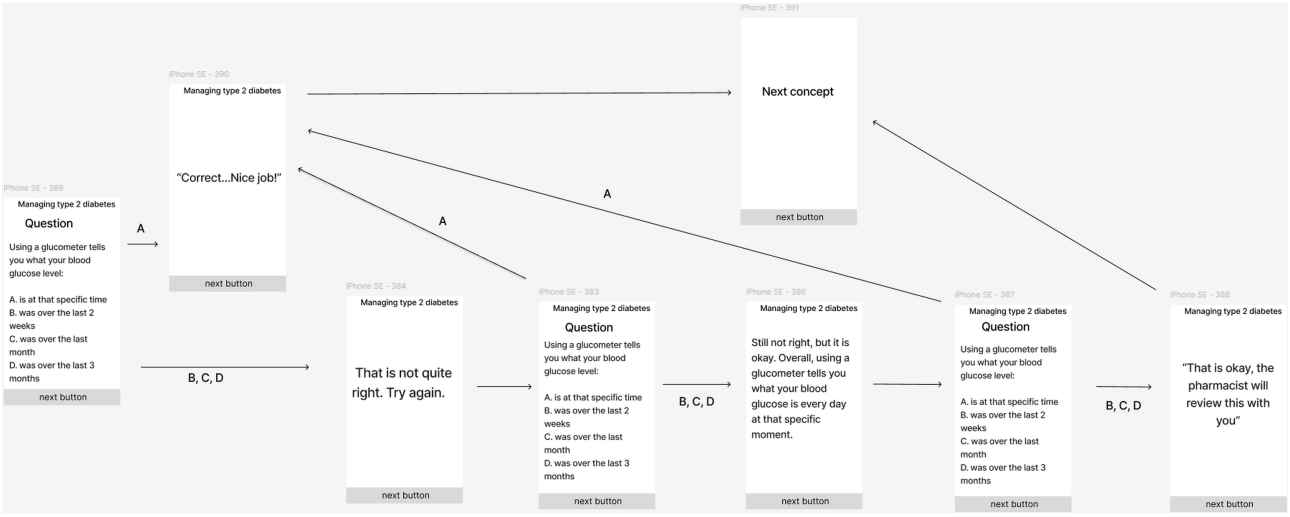

The first module provides general information on T2D (eg, the role of insulin) and asks questions about mental health and family support. The second module provides information on diet, exercise, the benefits of medication adherence, and methods of monitoring T2D. The module also asks the patient if there are specific areas they would like to explore further (eg, diet, exercise, vaccinations, and glucometer use). For example, the module will not only discuss the benefits of using a glucometer but will also ask if they use or need one, thereby alerting the pharmacist of the need to secure one for the patient before the next counseling session. The third module allows patients to select their prescribed medications, provides information about those medications, and asks about their ability to administer them. For example, if a patient states they take insulin, the module will provide information about insulin and then ask if they have fears about injecting it and if they would like additional instructions on proper administration. This information will allow the pharmacist to develop an individualized CMM plan focused on specific needs for each patient.

After the content is refined and modernized, 2 diabetes experts and an instructional design expert will review it and provide feedback. The content will then be updated based on their input, and the app developer will create 2 digital apps: the Patient and Provider apps. The Patient app will house the 3 modules. Patients will be able to easily download it to a mobile device using the app stores on iOS (eg, Apple iPhone) or Android (eg, Google phone). Based on preliminary discussions with the app developer, the app will fit into a 16×9 (1920×1080 pixels) aspect ratio and can be adapted to various formats (eg, smartphone and tablet). A narrated voiceover option will also be included. Therefore, the app will be accessible to patients who may be illiterate or have some visual problems (eg, mild-moderate diabetic retinopathy).

After the patient downloads the digital app and completes the 3 modules, the app will encrypt the patient's responses and generate a QR code. QR codes are used to package and transmit information securely, and their use has become more widespread, even becoming a recent function on Windows, allowing a user to convert a webpage into a QR code. QR codes provide an effective and secure method of transmitting patient responses directly to a provider who uses the Provider app. The QR code will contain all the encrypted information captured from the patient's responses and does not have to be uploaded to the cloud, where the information could be compromised. More importantly, the patient's responses contained within the

XSL•FO
RenderX

QR code cannot be decrypted without the Provider app. The patient will have the option to send their QR code via email ahead of time or provide it at their follow-up appointment. The Provider app will be used by the clinic to decrypt the data stored in the QR code using a concept in computer science known as asymmetric cryptography, which is the industry standard. This technology allows for the creation of a lock made of an unpredictable, large, and randomized series of numbers embedded in the QR Code, which can only be decrypted using a private key (eg, the Provider app). Once the QR code is scanned by the clinic pharmacist using the Provider app, the pharmacist will be able to access the patient's responses, presented in a user-friendly format. Although the modules and patient responses in the Patient app will be in Spanish, the information will be available in the Provider app in either English or Spanish according to the pharmacist's preference.

Phase II

During this phase, usability and acceptability testing by patients and pharmacists will be completed. To screen the patients, we will assess their language dominance using the Language Experience and Proficiency Questionnaire [37]. We will assess the patients' acculturation level using the scale developed by Marin et al [38]. This 12-item acculturation scale has been widely used in research on Hispanic individuals. Acculturation, as measured by this scale, correlates highly with respondents' generation, length of residence in the United States, age at arrival, and ethnic self-identification. We will use criterion sampling to recruit up to 20 older Hispanic patients with T2D and low acculturation.

Patient Usability and Acceptability Testing

A recent systematic review of usability research in educational technology found that many studies are narrowly focused and insufficiently iterative, leading to results that fail to support the user experience as comprehensively as possible [39]. A frequent problem is the application of unimodal approaches that apply a single questionnaire to obtain a summative measure of overall quality. In contrast, our patient usability and acceptability testing will apply a robust set of methods intended to iteratively improve the quality of the design, promoting a highly usable and acceptable experience. This will include repeatedly triangulating results from cognitive "think-aloud" interviews, screen recordings, and the previously validated SUS and AEOUS [40,41]. A previously used questionnaire based on Davis' TAM will also be used to evaluate system acceptability [22,42]. In summary, a set of 5 patients (seen separately) will be observed interacting with the Patient app and will be asked to "think out loud" so that their cognitive processes can be tracked and areas of difficulty can be identified as they review the modules [43,44]. Simultaneously, their screens will be recorded, and transcripts will be produced and timestamped for analysis. After each session, the SUS, AEOUS, and TAM will be administered to provide a snapshot of system quality. The app and content will be revised based on these results, tested with a new group of 5 participants (seen separately), and revised further [45]. This process may be repeated until no further difficulties are identified and benchmark SUS, AEOUS, and TAM scores are obtained, potentially eliciting responses from 15 to 20 participants [46,47]. Previous literature has indicated that small convenience samples are highly effective in software studies [48]. A sample size of 10 should be sufficient based on previous experience [22]; however, since we will also be evaluating the app and content on different devices, we will target up to 20 participants.

Pharmacist Usability and Acceptability Testing

We will test the usability and acceptability of the intervention for incorporation into CMMs with up to 10 pharmacists. Pharmacists will be recruited from our collaborative clinic (Athens Neighborhood Health Center) and other federally qualified health centers that work with Hispanic patients and the CMM process. We will conduct individual interviews with the pharmacists to test the intervention and obtain feedback (eg, SUS, AEOUS, and TAM scores) to improve the ease and practicality of use when conducting CMM. They will also test and evaluate the Provider digital app to identify the optimal method of displaying patient responses. Our previous study tested much of the content with pharmacists and required 1 round (5 individual interviews) [22]. However, since we are also focused on developing the digital app for pharmacists, we budgeted for 10 pharmacists in case 2 rounds of testing are needed to make sure the Provider app is user-friendly.

Phase III

The third phase will test health outcomes: medication adherence, HbA_{1c}, and fasting blood glucose. During this phase, the pharmacist will use the digital health app with 40 adult Hispanic patients with T2D who meet the inclusion requirements (stated under the "Recruitment" section). The study will take place at the Athens Neighborhood Health Center. If by 6 months, we have not recruited at least 25 patients, we will pursue another local federally qualified health center. After consent, patients will be instructed on downloading the digital health app, completing the modules within the first 2 weeks, and generating the QR code. Their preference for either submitting the QR code via email or bringing it to their next appointment will be confirmed. Medication adherence will be measured using scales previously used by the investigators (the Gonzalez-Lu questionnaire and the visual analog scale) [49,50]. The Gonzalez-Lu questionnaire and visual analog scale are being used due to their ease of use and strong reliability and correlation to electronic measurements of medication adherence [51-53]. Baseline HbA_{1c} and blood glucose levels (fasting) will be measured via labs conducted as standard medical practice. A follow-up visit will occur 3 - 4 months later. Depending on patients' responses, an additional visit 3 - 4 months later may be needed. The app would immediately assist a patient who is nonadherent due to not knowing how to take a medication (eg, the app would explain that taking metformin with food optimizes effectiveness and lowers side effects). However, a patient may be nonadherent due to the inability to pay for a particular medication; this would be captured by the app but not addressed immediately because they chose to provide their QR code at the follow-up appointment and not when the issue arose. In such a scenario, the patient would have a follow-up visit an additional 3 - 4 months after appropriate changes are made (eg, selecting another medication). In summary, Phase III will test our Aim

XSL•FO
RenderX

2 hypothesis that the app will improve medication adherence (>80% average adherence for the individual), HbA_{1c} (≥0.75% reduction), and fasting glucose (≥25% reduction) from baseline. These reductions are reasonable target endpoints based on the literature [54-56].

Recruitment

Two types of participants will be recruited for Phase II: patients and pharmacists. Patients (n=20) asked to participate will be Hispanic men and women aged 50 years and older with T2D who have been on diabetes medications for at least 90 days but who are not adherent (ie, <80% adherent) and have increased diabetes markers (ie, HbA_{1c} >7% and fasting glucose >125 mg/dL). In addition, they should be able to meaningfully participate in the informed consent process, demonstrate poor acculturation based on the Marin Acculturation Scale, be comfortable speaking Spanish (during the preconsent or consent process), have and use a smartphone or tablet, and be willing to participate in a single visit lasting approximately 90 minutes. Exclusion criteria will include cognitive or psychiatric difficulties so severe as to make the person unable to meaningfully participate in the informed consent process, diagnosis of Alzheimer disease or any other dementia, or receiving pharmacotherapy for Alzheimer disease, and severe retinopathy or visual impairment that impacts the ability to see module content on a smartphone or tablet. Pharmacist participants (n=10) asked to participate will be licensed pharmacists providing CMM to Hispanic patients with T2D in an ambulatory care setting. In addition, they should be able to meaningfully participate in the informed consent process and be willing to participate in a single visit lasting approximately 90 minutes.

Only patients will be recruited for Phase III. Patients asked to participate will be Hispanic men and women aged 50 years and older with T2D who have been on diabetes medications for at least 90 days but who are not adherent (ie, <80% adherent) and have increased diabetes markers (ie, HbA_{1c} >7% and fasting glucose >125 mg/dL). Inclusion criteria will be similar to those in Phase II. However, they must be willing to participate in a baseline office visit and 1 - 2 subsequent follow-up office visits lasting 60 - 90 minutes, with each visit spaced 3 - 4 months apart. Exclusion criteria are similar to those described in Phase II.

Measures

Data to be collected include participants' demographics (eg, age, gender, and country of origin), self-reported medication adherence (eg, Gonzalez-Lu questionnaire and visual analog scale), and laboratory values (eg, HbA_{1c} and fasting glucose) from the clinic chart. Most self-reported data will be collected on paper and transferred into electronic databases secured on university-issued password-protected computers. Data will only be linked to participants' identities by way of an identification number, and no personally identifiable information other than the identification number will be included with the data. The correspondence between participant identification numbers and their identity will be kept separately. All other data will be kept on password-protected university computers. In addition, during Phase II, individual software testing sessions will be recorded, with computer screen recordings of participant behavior (patients and pharmacists), while notes are taken by the investigator. These recordings will also be kept on password-protected university computers and will only be associated with the identification number and not with any personally identifiable information. Any recordings on software (eg, Lookback Usability Testing software) are also password-protected and will not include any protected health information. Lookback Usability Testing software provides the ability to assess usability by recording the participants' interaction with the digital intervention. During Phase III, the pharmacists who are providing care to patient participants will have the ability to decrypt the QR code and match the patient with the identifier. This will allow the pharmacist to place the electronic file into the patient's protected folder in their electronic health records. These records are password-protected and housed on password-protected computers. Please note this is not any different than the standard of care and normal operating procedures when medical records are received from other sources and securely uploaded to the patient's electronic record.

Procedures

Before the recruitment process, University of Georgia Institutional Review Board (IRB) approval for the proposed human participants research protocols will be secured, as well as any review required by the Athens Neighborhood Health Center. The study will also be registered on the United States government clinical trials website.

Overall, potentially eligible patient participants (for Phase II and III) will be recruited from patients at the Athens Neighborhood Health Center during their regular appointments. In addition, bilingual flyers will be posted in the clinic's waiting room. Patients will be able to contact the investigator or clinician via email or phone call in English or Spanish to obtain more information about the study and/or express their interest in the study. Once patient participants are identified, a telephone screening or face-to-face interview in Spanish will be conducted for each potential participant. During the screening interview, participants will be informed of the nature of the study and may participate if interested and if the eligibility criteria are met. If they choose not to participate, they will be thanked for their interest and reminded that not participating does not impact their standing with the clinic. Potential pharmacist participants for Phase II will be recruited from the clinic site and locally (in the state of Georgia), as well as from other states serving Hispanic populations. As a result, it is important to note that pharmacist participants may be remotely interviewed via a secure server (ie, Zoom Meetings software) during Phase II. While pharmacists will have the option for remote interviews, visits for patient participants will all be conducted in person (face-to-face).

Aim 1

The primary endpoint of Aim 1 is to have a digital app that is user-friendly to both the patient and pharmacist. Regardless of whether the interview is face-to-face (for patients) or virtual (for some pharmacists), the software used (Lookback and Atlas.ti [Lumivero]), along with the recorded interviews, will

XSL•FO
RenderX

allow us to capture information needed to conduct a thorough analysis. For each group of 5 participants, qualitative data (interviews and screen recordings) will be coded to identify features of the intervention that benefited or hindered the system's ease of use and participants' user experiences. Analysis will be based on an existing coding framework designed for medical interventions and expanded based on usability standards [38,55,56]. Analyst triangulation will be conducted to ensure high intercoder agreement. Issues found to have negatively impacted user experiences will be resolved before testing with new groups of 5 participants (for both patients and pharmacists). In addition, the SUS, AEOUS, TAM, and time to complete will be analyzed using descriptive statistics to provide quantitative snapshots of system quality for each group. This process of qualitative and quantitative analysis may be repeated with groups of 5 participants until no further difficulties are identified and benchmark scores are achieved for the SUS, AEOUS, and TAM. In summary, phase II will iteratively refine the digital apps based on qualitative analyses and test our Aim 1 hypothesis that the apps will have above-average usability and acceptability (SUS: mean score: ≥71 and AEOUS: mean score ≥5) with an overall goal that the digital apps will be acceptable (via TAM) by ≥90% of the patient participants and ambulatory care clinical pharmacists. These measurable targets are based on published data [40,57,58].

Aim 2

The primary endpoint of Aim 2 is to improve adherence and health outcomes. We completed power analyses for this aim using a mixed-effects simulation routine in Power Analysis Sample Size (PASS 16; NCSS, LLC) software, and drawing effect size estimates from the literature on the impact of pharmacist counseling on diabetes control [59,60]. This analysis shows our planned sample size of 40 will provide a power of 0.85 to detect a significant change from baseline in outcome measures. Outcome measures include improvements in adherence and a reduction in HbA_{1c} and fasting glucose. We will evaluate differences before and after the intervention (eg, 3 - 8 months later) using repeated-measures mixed-effects models, enabling us to evaluate not only the significance of change but also the impact of key covariates (age, gender, and health literacy) and effect size for use in planning a larger trial.

Evaluation Outcomes

After completion of Phase III, we will have produced 2 refined digital health apps. The Patient app will house the diabetes content for the patient to review and complete, while the Provider app will be able to transform the patients' responses to the modules into an easy-to-read, pragmatic format to provide precision counseling. As a result, both apps will be ready for deployment and testing in a larger-scale study using a control group. The study will also provide a foundational framework for application to different diseases, languages, and cultures, and health care providers, potentially having a wider impact. This study will provide the foundation for our team to conduct a large-scale, robust, and pragmatic trial of the integration of the Patient and Provider apps into ambulatory care clinical pharmacist patient care services in geographically diverse federally qualified health centers to improve T2D medication use and outcomes. We plan to pursue such a larger-scale study, estimated to last 5 years via the R01 funding mechanism, to test the longer-term impact of the apps on health outcomes. In future studies, based on patient responses to these modules, follow-up visits, and as determined by the health care providers, we plan to implement artificial intelligence technology to deliver tailored prompts (via texts) to maintain and promote patient engagement and activation.

Ethical Considerations

All recruitment materials will be provided to the University of Georgia IRB Committee for approval. Study personnel will complete all necessary training in the protection of human participants in research as required by the IRB and available through the Collaborative Institutional Training Initiative program. Patients and pharmacist participants will provide written informed consent obtained by trained study personnel before being enrolled in the study. All protected health information will be maintained in accordance with institutional regulations and policies, and standard ethical practice. The Phase I portion of the study did not require IRB approval since it involved modernizing and refining content. Phase II (patient and pharmacist participants' qualitative analysis) and Phase III (pilot study assessing patient health outcomes) will require University of Georgia IRB approval. Informed written consent will be secured from all participants and appropriately stored.

Results

The tailoring of content and video production has been completed. However, we received notice that our NIH funding was terminated in early May 2025 with the following explanation: "It is the policy of NIH not to prioritize research programs related to DEI (diversity, equity, and inclusion): research programs based primarily on artificial and nonscientific categories, including amorphous equity objectives, are antithetical to the scientific inquiry, do nothing to expand our knowledge of living systems, provide low returns on investment, and ultimately do not enhance health, lengthen life, or reduce illness. Worse, so-called DEI studies are often used to support unlawful discrimination based on race and other protected characteristics, which harms the health of Americans. Therefore, it is the policy of NIH not to prioritize such research programs." As a result, we appealed the decision, stating that we will open the study to all races and ethnicities through a no-cost extension. The appeal was denied in late August 2025. In light of the NIH funding termination, the researchers have used their faculty development funds to continue their work on a smaller scale and have continued to seek other support mechanisms. As of January 2026, submission to the IRB has been completed to begin the qualitative (acceptability and usability) portion of the study (Phase II and Aim 1) in Hispanic adults with T2D and pharmacists. We expect acceptability and usability studies to be completed within 12 months. We anticipate that both digital apps will be acceptable and usable for both patients and pharmacists. In addition, during Phase III (Aim 2), we expect the digital apps to assist patients in improving their health outcomes. Specifically, we expect that the apps will improve adherence (>80% average adherence for the individual), HbA_{1c}

XSL•FO
RenderX

(≥0.75% reduction), and fasting glucose (≥25% reduction) from baseline. These reductions are reasonable target endpoints based on the literature [54-56].

Discussion

Principal Findings

The research project aims to refine and modernize diabetes content into an interactive experience for patients, which will be aimed at developing tailored information. The responses provided by the patient will be provided to the pharmacist in an easy-to-read format that will allow the clinical pharmacist to gather the needed information to provide precision counseling to their patients.

Based on a previous study we completed, 5 themes among older Hispanic participants with T2D were identified, which included: financial considerations, polypharmacy, social and family support, access to medication information, and loneliness or sadness [22]. These themes are being applied to the current digital apps. For example, content and items on social and family support and loneliness or sadness are being implemented. In addition, the Patient app will provide brief information about loneliness and sadness and will have subsequent items exploring patients' mental health (ie, Patient Health Questionnaire–2 [PHQ-2]). If a patient scores high on the PHQ-2 items, after decryption, the Provider app will alert the pharmacist to explore this area with the patient and refer to appropriate services, if needed. Items regarding nonadherence have options related to financial considerations. If the patient selects this as a need, it will trigger the pharmacist to explore a more cost-effective alternative (eg, cheaper medication and a blood glucose monitor).

Digital Content Design and Usability

In addition, we are using software to produce animated videos to support content dissemination. For example, we will develop whiteboard animated videos and determine the types of animation preferred by patients during the acceptability and usability portion of the study (Phase II). Therefore, the qualitative portion of the study (Aim 1) will assist us in having an intervention that is usable and acceptable by patients and pharmacists.

Integration Into Clinical Workflow

We are also hoping the health outcome results of the pilot study are favorable or provide further insight into developing a better understanding of how to improve the content and interface of the digital health apps. If successful, our goal with this project is that the outputs created are provided in an easy-to-read format for the provider, in which they can edit after they meet with the patient and can simply link the document with the patient's electronic medical record. This would allow for a seamless transition and potentially reduce workload.

Implications for Policy and Practice

If successful, this project may assist in supporting the Centers for Medicare and Medicaid Services Diabetes Strategy 2024 [61]. For example, it may assist with telehealth flexibilities for diabetes self-management training, increase support for social determinants of health in identifying adherence barriers (eg, transportation and cost), and potentially decrease readmissions or urgent care visits.

Limitations

While several strengths have been discussed, the study is not without limitations. Limitations include a small sample size, and the current plan is to conduct the study at a single location. Therefore, the generalizability to other populations may remain unclear. In addition, there could be difficulties with technology among older adults (eg, aged ≥65 years); however, we will attempt to mitigate this risk by recruiting participants with T2D across the eligible older adult age range. In addition, there is no control group for the pilot study. Therefore, the significance of the results may be limited when assessing if the health outcomes are attributed specifically to the intervention or other confounding factors.

Conclusions

In summary, this study explores a critical issue in developing technology that engages patients and addresses their specific needs. It also seeks to improve the way pharmacists can review patient information to identify barriers to medication adherence and promote patient well-being. The research aims to expand existing knowledge on how to improve CMM and optimize precision counseling. The findings of this study are expected to provide valuable insights to improve the delivery of care and health outcomes in older patients with T2D.

Acknowledgments

We would like to acknowledge the collaboration with Carlos Lacasa, owner of MKL Designs, for developing the digital apps and assisting with the app development and content integration.

Funding

This study received funding from the National Institutes of Health (NIH), National Institute of Diabetes and Digestive and Kidney Diseases (1R21DK139557-01). However, it was terminated before year 1 completion due to a shift in NIH priorities. Authors have no significant financial relationships to disclose.

XSL•FO
RenderX

Authors' Contributions

JC, RLO, and RP conceived the project, participated in writing the grant applications, prepared the manuscript, and provided revisions. HNY, MAB, KE, and MBM refined intervention content, assisted with manuscript preparation, and provided revisions.

Conflicts of Interest

None declared.

References

1. National diabetes statistics report, 2020. : Centers for Disease Control and Prevention, U.S. Dept of Health and Human Services; 2020 URL: https://diabetesresearch.org/wp-content/uploads/2022/05/national-diabetes-statistics-report-2020.pdf [accessed 2025-03-23]
2. Betancourt JR, Tan-McGrory A, Kenst KS. Guide to Preventing Readmissions among Racially and Ethnically Diverse Medicare Beneficiaries. : Centers for Medicare & Medicaid Services Office of Minority Health (CMS OMH); 2015 Sep URL: https://essentialhospitals.org/wp-content/uploads/2016/01/OMH_Readmissions_Guide.pdf [accessed 2025-02-24]
3. Weiss AJ, Jiang HJ. Overview of clinical conditions with frequent and costly hospital readmissions by payer, 2018. Agency for Healthcare Research and Quality (AHRQ). 2021 URL: www.hcup-us.ahrq.gov/reports/statbriefs/sb278Conditions-Frequent-Readmissions-By-Payer-2018.pdf [accessed 2026-01-11]
4. Fernández A, Quan J, Moffet H, Parker MM, Schillinger D, Karter AJ. Adherence to newly prescribed diabetes medications among insured Latino and white patients with diabetes. JAMA Intern Med 2017 Mar 1;177(3):371-379. [doi: 10.1001/jamainternmed.2016.8653] [Medline: 28114642]
5. Chehal PK, Uppal TS, Turbow S, et al. Continuity of medication use by US adults with diabetes, 2005-2019. JAMA Netw Open 2023 Jan 3;6(1):e2253562. [doi: 10.1001/jamanetworkopen.2022.53562] [Medline: 36716032]
6. Kim H, Ross JS, Melkus GD, Zhao Z, Boockvar K. Scheduled and unscheduled hospital readmissions among patients with diabetes. Am J Manag Care 2010 Oct;16(10):760-767. [Medline: 20964472]
7. Alshamsan R, Majeed A, Vamos EP, et al. Ethnic differences in diabetes management in patients with and without comorbid medical conditions: a cross-sectional study. Diabetes Care 2011 Mar;34(3):655-657. [doi: 10.2337/dc10-1606] [Medline: 21282346]
8. Kessels RPC. Patients' memory for medical information. J R Soc Med 2003 May;96(5):219-222. [doi: 10.1177/014107680309600504] [Medline: 12724430]
9. Mayer RE. Applying the science of learning: evidence-based principles for the design of multimedia instruction. Am Psychol 2008 Nov;63(8):760-769. [doi: 10.1037/0003-066X.63.8.760] [Medline: 19014238]
10. Makoul G, Cameron KA, Baker DW, Francis L, Scholtens D, Wolf MS. A multimedia patient education program on colorectal cancer screening increases knowledge and willingness to consider screening among Hispanic/Latino patients. Patient Educ Couns 2009 Aug;76(2):220-226. [doi: 10.1016/j.pec.2009.01.006] [Medline: 19250791]
11. Reséndiz Lara T, Muñoz Torres AV, Mendoza Salmerón G, et al. Education with a multimedia web platform improves knowledge and HbA1c of Mexican patients with type 2 diabetes. Open clinical trial. Endocrinol Diabetes Nutr (Engl Ed) 2020 Oct;67(8):530-539. [doi: 10.1016/j.endinu.2019.07.011] [Medline: 31690537]
12. Pinillos-Patiño Y, Herazo-Beltrán Y, Rodríguez-Cordero O, et al. User preferences related to multimedia elements of a mobile application to prevent diabetes. Healthc Inform Res 2020 Oct;26(4):295-302. [doi: 10.4258/hir.2020.26.4.295] [Medline: 33190463]
13. Bassi G, Mancinelli E, Dell'Arciprete G, Rizzi S, Gabrielli S, Salcuni S. Efficacy of eHealth interventions for adults with diabetes: a systematic review and meta-analysis. Int J Environ Res Public Health 2021 Aug 26;18(17):8982. [doi: 10.3390/ijerph18178982] [Medline: 34501572]
14. Fortmann AL, Gallo LC, Garcia MI, et al. Dulce Digital: an mHealth SMS-based intervention improves glycemic control in Hispanics with type 2 diabetes. Diabetes Care 2017 Oct;40(10):1349-1355. [doi: 10.2337/dc17-0230] [Medline: 28600309]
15. Heisler M, Choi H, Palmisano G, et al. Comparison of community health worker-led diabetes medication decision-making support for low-income Latino and African American adults with diabetes using e-health tools versus print materials: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S13-S22. [doi: 10.7326/M13-3012] [Medline: 25402398]
16. Fortmann AL, Philis-tsimikas A, Bagsic SRS, et al. 646-P: Dulce Digital-Me: an adaptive mHealth intervention for underserved Hispanics with diabetes. Diabetes 2020 Jun 1;69(Supplement_1):Supp. [doi: 10.2337/db20-646-P]
17. Share of adults in the united states who owned a smartphone from 2011 to 2023, by ethnicity. Statista. 2025. URL: https://www.statista.com/statistics/195001/percentage-of-us-smartphone-owners-by-ethnicity [accessed 2025-03-03]
18. Anderson M. Racial and ethnic differences in how people use mobile technology. Pew Research. 2015 Apr 30. URL: https://www.pewresearch.org/internet/2015/04/30/racial-and-ethnic-differences-in-how-people-use-mobile-technology/ [accessed 2026-01-18]
19. Suarez-Lledo V, Alvarez-Galvez J. Prevalence of health misinformation on social media: systematic review. J Med Internet Res 2021 Jan 20;23(1):e17187. [doi: 10.2196/17187] [Medline: 33470931]

XSL•FO
RenderX

20. Powell D, Keil M, Brenner D, Lim L, Markman EM. Misleading health consumers through violations of communicative norms: a case study of online diabetes education. Psychol Sci 2018 Jul;29(7):1104-1112. [doi: 10.1177/0956797617753393] [Medline: 29757711]
21. Hispanic consumers are leading the way in digital engagement. Nielsen Company. 2015 Jun 18. URL: https://www.nielsen.com/insights/2015/hispanic-consumers-are-leading-the-way-in-digital-engagement/ [accessed 2025-01-18]
22. Caballero J, Jacobs RJ, Ownby RL. Development of a computerized intervention to improve health literacy in older Hispanics with type 2 diabetes using a pharmacist supervised comprehensive medication management. PLoS ONE 2022;17(2):e0263264. [doi: 10.1371/journal.pone.0263264] [Medline: 35139107]
23. Cuddihy RM, Philis-Tsimikas A, Nazeri A. Type 2 diabetes care and insulin intensification: is a more multidisciplinary approach needed? Results from the MODIFY survey. Diabetes Educ 2011;37(1):111-123. [doi: 10.1177/0145721710388426] [Medline: 21131599]
24. Hassali MAA, Nazir SUR, Saleem F, Masood I. Literature review: pharmacists' interventions to improve control and management in type 2 diabetes mellitus. Altern Ther Health Med 2015;21(1):28-35. [Medline: 25599430]
25. Iglay K, Cartier SE, Rosen VM, et al. Meta-analysis of studies examining medication adherence, persistence, and discontinuation of oral antihyperglycemic agents in type 2 diabetes. Curr Med Res Opin 2015;31(7):1283-1296. [doi: 10.1185/03007995.2015.1053048] [Medline: 26023805]
26. Vermeire E, Wens J, Van Royen P, Biot Y, Hearnshaw H, Lindenmeyer A. Interventions for improving adherence to treatment recommendations in people with type 2 diabetes mellitus. Cochrane Database Syst Rev 2005 Apr 18;2005(2):CD003638. [doi: 10.1002/14651858.CD003638.pub2] [Medline: 15846672]
27. Hoover ND, Turner RB, Sampson J, Pye T, Hotan T. Financial sustainability of an Oregon Rural Health, primary care, and pharmacist-run comprehensive medication management program through direct medical billing. J Manag Care Spec Pharm 2020 Jan;26(1):30-34. [doi: 10.18553/jmcp.2020.26.1.30] [Medline: 31880232]
28. Fixen DR, Linnebur SA, Parnes BL, Vejar MV, Vande Griend JP. Development and economic evaluation of a pharmacist-provided chronic care management service in an ambulatory care geriatrics clinic. Am J Health Syst Pharm 2018 Nov 15;75(22):1805-1811. [doi: 10.2146/ajhp170723] [Medline: 30104259]
29. Ownby RL, Acevedo A, Waldrop-Valverde D, Jacobs RJ, Caballero J. Abilities, skills and knowledge in measures of health literacy. Patient Educ Couns 2014 May;95(2):211-217. [doi: 10.1016/j.pec.2014.02.002] [Medline: 24637163]
30. Henderson VA, Barr KL, An LC, et al. Community-based participatory research and user-centered design in a diabetes medication information and decision tool. Prog Community Health Partnersh 2013;7(2):171-184. [doi: 10.1353/cpr.2013.0024] [Medline: 23793248]
31. Philis-Tsimikas A, Fortmann AL, Clark T, et al. Dulce Digital-Me: results of a randomized comparative trial of static versus adaptive digital interventions for Latine adults with diabetes. Ann Behav Med 2025 Jan 4;59(1). [doi: 10.1093/abm/kaae077]
32. Gagné RM, Briggs LJ, Wager WW. Principles of Instructional Design 4th Ed: Harcourt Brace College Publishers; 1992.
33. Mayer RE. Multimedia Learning, 2nd edition: Cambridge University Press; 2009. [doi: 10.1017/CBO9780511811678]
34. Albrecht L, Scott SD, Hartling L. Evaluating a knowledge translation tool for parents about pediatric acute gastroenteritis: a pilot randomized trial. Pilot Feasibility Stud 2018;4:131. [doi: 10.1186/s40814-018-0318-0] [Medline: 30123522]
35. Fisher JD, Fisher WA. Changing AIDS-risk behavior. Psychol Bull 1992 May;111(3):455-474. [doi: 10.1037/0033-2909.111.3.455] [Medline: 1594721]
36. Kim Y, Lee H, Seo JM. Integrated diabetes self-management program using smartphone application: a randomized controlled trial. West J Nurs Res 2022 Apr;44(4):383-394. [doi: 10.1177/0193945921994912] [Medline: 33655794]
37. Marian V, Blumenfeld HK, Kaushanskaya M. The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. J Speech Lang Hear Res 2007 Aug;50(4):940-967. [doi: 10.1044/1092-4388(2007/067)] [Medline: 17675598]
38. Marin G, Sabogal F, Marin BV, Otero-Sabogal R, Perez-Stable EJ. Development of a Short Acculturation Scale for Hispanics. Hisp J Behav Sci 1987 Jun;9(2):183-205. [doi: 10.1177/07399863870092005]
39. Lu J, Schmidt M, Lee M, Huang R. Usability research in educational technology: a state-of-the-art systematic review. Education Tech Research Dev 2022 Dec;70(6):1951-1992. [doi: 10.1007/s11423-022-10152-6]
40. Brooke J. SUS: a quick and dirty usability scale. In: Usability Evaluation In Industry 1996. [doi: 10.1201/9781498710411-35]
41. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact 2008 Jul 29;24(6):574-594. [doi: 10.1080/10447310802205776]
42. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep 1;13(3):319-340. [doi: 10.2307/249008]
43. Barnum CM, Dragga S. Usability Testing and Research: Allyn & Bacon; 2001.
44. Wiklund MD, Kendler J, Strochlic AY. Usability Testing of Medical Devices: CRC Press; 2010.
45. Genov A. Iterative usability testing as continuous feedback: A control systems perspective. J Usability Stud 2005;1:18-27 [FREE Full text]
46. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. Presented at: CHI '93: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing System; Apr 24-29, 1993; Amsterdam, The Netherlands p. 206-213 URL: http://portal.acm.org/citation.cfm?doid=169059 [doi: 10.1145/169059.169166]

XSL•FO
RenderX

47. Nielsen J. Usability engineering at a discount. Presented at: Proceedings of the Third International Conference on Human-Computer Interaction on Designing and Using Human-Computer Interfaces and Knowledge Based Systems; Sep 1, 1989; Boston, Massachusetts p. 394-401 URL: https://dl.acm.org/doi/proceedings/10.5555/92449 [accessed 2026-01-11]
48. Turner CW, Lewis JR, Nielsen J. Determining usability test sample size. In: International Encyclopedia of Ergonomics and Human Factors, 2nd edition: CRC Press; 2006:3084-3088 URL: https://www.researchgate.net/publication/242156700_Determining_Usability_Test_Sample_Size [accessed 2026-01-11]
49. Caballero J, Ownby RL, Jacobs RJ, Pandya N, Hardigan PC, Ricabal LC. Predicting medication adherence in older Hispanic patients with type 2 diabetes. Am J Health Syst Pharm 2018 May 1;75(9):e194-e201. [doi: 10.2146/ajhp170067] [Medline: 29691262]
50. Caballero J, Ownby RL, Jacobs RJ, Thomas JE, Schweizer MS. Association between cognitive tests and antiretroviral medication adherence in older adults with HIV. Ann Pharmacother 2019 Feb;53(2):151-158. [doi: 10.1177/1060028018798327] [Medline: 30175588]
51. Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. Acad Emerg Med 2001 Dec;8(12):1153-1157. [doi: 10.1111/j.1553-2712.2001.tb01132.x] [Medline: 11733293]
52. Gonzalez JS, Schneider HE, Wexler DJ, et al. Validity of medication adherence self-reports in adults with type 2 diabetes. Diabetes Care 2013 Apr;36(4):831-837. [doi: 10.2337/dc12-0410] [Medline: 23204245]
53. Lu M, Safren SA, Skolnik PR, et al. Optimal recall period and response task for self-reported HIV medication adherence. AIDS Behav 2008 Jan;12(1):86-94. [doi: 10.1007/s10461-007-9261-4] [Medline: 17577653]
54. Kirkman MS, Rowan-Martin MT, Levin R, et al. Determinants of adherence to diabetes medications: findings from a large pharmacy claims database. Diabetes Care 2015 Apr;38(4):604-609. [doi: 10.2337/dc14-2098] [Medline: 25573883]
55. Hundal RS, Krssak M, Dufour S, et al. Mechanism by which metformin reduces glucose production in type 2 diabetes. Diabetes 2000 Dec;49(12):2063-2069. [doi: 10.2337/diabetes.49.12.2063] [Medline: 11118008]
56. Sherifali D, Nerenberg K, Pullenayegum E, Cheng JE, Gerstein HC. The effect of oral antidiabetic agents on A1C levels: a systematic review and meta-analysis. Diabetes Care 2010 Aug;33(8):1859-1864. [doi: 10.2337/dc09-1727] [Medline: 20484130]
57. Lewis JR, Sauro J. Item benchmarks for the system usability scale. J Usability Stud 2018;13(3):158-167 [FREE Full text]
58. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
59. Hintze J. PASS 16. NCSS. 2018. URL: https://www.ncss.com/ [accessed 2026-01-11]
60. Presley B, Groot W, Pavlova M. Pharmacy-led interventions to improve medication adherence among adults with diabetes: a systematic review and meta-analysis. Res Social Adm Pharm 2019 Sep;15(9):1057-1067. [doi: 10.1016/j.sapharm.2018.09.021] [Medline: 30685443]
61. CMS diabetes strategy: impact report 2024. : Centers for Medicare & Medicaid Services (CMS), U.S. Department of Health & Human Services; 2024 URL: https://www.cms.gov/files/document/cms-diabetes-strategy-2024-english.pdf [accessed 2025-09-20]

Abbreviations

AEOUS: Adjectival Ease of Use Scale
CMM: comprehensive medication management
GNEOI: Gagne's Nine Events of Instruction
HbA1c: Hemoglobin A1c
IMB: information-motivation-behavioral
IRB: Institutional Review Board
NIH: National Institutes of Health
PHQ-2: Patient Health Questionnaire–2
SUS: System Usability Scale
T2D: type 2 diabetes
TAM: technology acceptance model

XSL•FO
RenderX

Edited by S Basha; submitted 22.Apr.2025; peer-reviewed by A Murzello, D Garcia; revised version received 29.Sep.2025; accepted 23.Oct.2025; published 23.Jan.2026.

Please cite as:
Caballero J, Ownby RL, Young HN, McElhannon MB, Bhuiya NMMA, Esho KH, Palmer RH
Developing a Culturally Adapted Digital Health Application for Older Hispanic Adults With Type 2 Diabetes: Protocol for a Qualitative and Pilot Study
JMIR Res Protoc 2026;15:e76294
URL: https://www.researchprotocols.org/2026/1/e76294
doi:10.2196/76294

© Joshua Caballero, Raymond L Ownby, Henry Nolan Young, Michelle B McElhannon, N M Mahmudul Alam Bhuiya, Kenny H Esho, Russ H Palmer. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Multilevel Network Peer Intervention Among Student Men Who Have Sex With Men Attending University: Protocol for an Implementation-Effectiveness Before-After Cohort Study

Jingpei Xu, MPH; Zhen Dai, MPH; Yushu Qiu, MPH; Liang Wang, MPH

Department of Sexually Transmitted Disease, AIDS Prevention and Control, Chengdu Center for Disease Control and Prevention (Chengdu Institute of Health Supervision), Longxiang Road 4, Wuhou District, Chengdu, Sichuan, China

Corresponding Author:
Liang Wang, MPH
Department of Sexually Transmitted Disease, AIDS Prevention and Control, Chengdu Center for Disease Control and Prevention (Chengdu Institute of Health Supervision), Longxiang Road 4, Wuhou District, Chengdu, Sichuan, China

Abstract

Background: The HIV prevalence among student men who have sex with men (MSM) in China is substantially higher than that in the general student population. However, targeted interventions for this vulnerable population remain limited. While digital technologies and peer-led approaches have shown promise in HIV prevention among MSM, their application in university settings is underexplored.

Objective: This study aims to evaluate the implementation and effectiveness of a multilevel, internet-based peer intervention model in reducing HIV and syphilis incidence, improving prevention behaviors, and increasing uptake of HIV-related services among student MSM attending university in Chengdu.

Methods: This prospective before-and-after self-controlled cohort study will evaluate the effectiveness of a 12-month internet-based intervention targeting university student MSM in Chengdu. A total of 484 HIV-negative student MSM among the MSM population will be recruited through WeChat and Blued. Trained student volunteers, supervised by community-based organization staff, will deliver monthly health education, one-on-one counseling, postexposure prophylaxis navigation, and HIV or sexually transmitted infection (STI) self-testing kits via secure online platforms. Participants will complete behavioral surveys and HIV/syphilis/hepatitis C virus/hepatitis B surface antigen self-tests at baseline, 6 months, and 12 months. Primary outcomes included changes in HIV and syphilis incidence rates; secondary outcomes included changes in sexual behaviors, HIV/STI testing frequency, preexposure prophylaxis/postexposure prophylaxis uptake, and knowledge improvement.

Results: As of May 1, 2025, the intervention materials and training program have been finalized, and recruitment has begun. Enrollment of participants started in April 2025. By the end of April, 127 participants had completed baseline surveys and HIV/STI self-tests.

Conclusions: This study will provide empirical evidence on the feasibility and effectiveness of an internet-based peer intervention for student MSM.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2500100825; https://www.chictr.org.cn/showproj.html?proj=269114

International Registered Report Identifier (IRRID): DERR1-10.2196/77078

(JMIR Res Protoc 2026;15:e77078) doi:10.2196/77078

KEYWORDS

HIV; MSM; university students; internet-based intervention; peer education; cohort study; PEP; men who have sex with men; postexposure prophylaxis

Introduction

National sentinel surveillance data showed a rapid increase in newly reported HIV cases among young students in China from 2010 to 2019. The HIV prevalence among student men who have sex with men (MSM) in China is substantially higher than that in the general student population, with regional comparisons revealing the highest prevalence in the southwest region [1].

As the economic and cultural center of southwest China [2], Chengdu was estimated to have up to 120,000 active MSM in 2018 [3]. In 2021, the estimated number of college MSM registered on social networking platforms in Chengdu reached 8401, accounting for 2.08% of the total male university student population [4]. Although students were not traditionally considered a high-risk population for HIV infection, surveillance data indicate a rapidly increasing proportion of students infected with HIV/AIDS in the 15- to 24-year age group nationwide [5].

XSL•FO
RenderX

Systematic reviews have reported HIV prevalence rates of 3.8% [6] to 4.4% [5] among university student MSM. Risky sexual behaviors are common [7]; moreover, 85% of student MSM in Chengdu reported meeting sexual partners through mobile social networking platforms [4].

To date, research on university student MSM in China has primarily relied on cross-sectional surveys [6], limiting the understanding of factors driving HIV transmission in this population. While intervention trials targeting MSM populations have been conducted, few have specifically focused on student MSM. The HIV prevalence among university student MSM highlights an urgent need for effective, evidence-based public health strategies tailored to this group [5]. Compared to other high-risk populations, prevention efforts targeting university student MSM in China remain insufficient and under-researched [8].

There is growing evidence that online interventions can effectively reduce HIV risk behaviors and promote HIV testing among MSM populations [9,10]. These findings underscore the potential of digital strategies in reducing HIV or sexually transmitted infection (STI) disparities among student MSM. Given this, it is crucial to evaluate the effectiveness of internet-based intervention programs in this population.

In this study, we aim to establish a cohort of university student MSM and implement a comprehensive intervention through social networking platforms. The intervention will include regular behavioral education, HIV testing promotion, service referrals, and postexposure prophylaxis (PEP) subsidies. Previous research has shown that young people are often influenced by their peers and surrounding environments when adopting health behaviors [11,12]. Moreover, based on Social Cognitive Theory, peer educators who share similar identities and experiences can play a vital role in promoting healthy behaviors among student MSM [13].

In recent years, government regulations on community-based organizations (CBOs) serving the MSM population have become increasingly stringent, particularly regarding their direct engagement with student populations. In Chengdu, while the local Centers for Disease Control and Prevention (CDC) and AIDS Prevention Association have recruited university student volunteers for peer education, these efforts have primarily targeted the general student population and have rarely provided comprehensive, MSM-specific interventions.

Therefore, this study proposes an innovative multilevel network peer intervention model, in which the MSM CBO staff will serve as trainers, and student MSM volunteers will act as peer educators. To our knowledge, this will be the first study in China to explore and implement such a targeted intervention model specifically designed for university student MSM.

Methods

Ethical Considerations

Participation in this study will be entirely voluntary, and comprehensive measures will be implemented to ensure participant anonymity and data confidentiality. The study protocol has been reviewed and approved by the Ethics Committee of the Chengdu CDC (Approval No. 2024021).

All participants were required to carefully read the electronic informed consent form prior to enrollment, which will provide detailed information about the study purpose, procedures, potential risks, and benefits. Only those who voluntarily select the “Agree” option will be permitted to proceed with the questionnaire, thereby documenting their informed consent. Participants will have the right to withdraw from the study at any time without any negative consequences.

To prevent coercion or power imbalance between peers, volunteers, and participants were matched across different universities; all volunteers must sign a confidentiality agreement; all participation is anonymous and voluntary. We also conducted dedicated confidentiality training during the pre-project orientation.

To support HIV prevention efforts, the study will provide access to evidence-based online health education, individualized counseling services, and free prevention supplies, including condoms, lubricants, and HIV/STI self-testing kits. Additionally, participants will receive modest financial compensation for completing each survey, in recognition of their time and contribution. Participants received RMB 20 (US $2.83) at baseline, RMB 20 (US $2.83) at the 6-month follow-up, and RMB 20 (US $2.83) at the final 12-month follow-up.

Potential risks primarily will involve the possible inadvertent disclosure of sensitive information related to sexual behavior. However, robust data security and privacy protection protocols will be established, including data encryption, pseudonymization of identifiers, and restricted access to sensitive information, to minimize this risk.

The anticipated benefits of participation—including increased HIV-related knowledge, improved access to prevention services, and facilitation of early detection and timely linkage to care—are expected to outweigh the minimal risks involved.

Findings from this study will be disseminated to local and national public health authorities to inform HIV prevention strategies targeting students belonging to the student MSM population. In addition, study results will be shared with the broader scientific community through peer-reviewed publications and academic conference presentations.

Study Aims

This prospective cohort study aims to evaluate the effectiveness of a comprehensive Internet-based intervention targeting student MSM attending university. Specifically, it seeks to assess the intervention’s impact on reducing HIV-related high-risk behaviors and lowering the prevalence of HIV and other STIs within this population, using a before-and-after controlled trial design.

Study Design

The study evolved from a 2024 local public health intervention pilot. During the ethical review in June 2024, the local ethics committee recommended that, given the sensitivity and vulnerability of university student participants, all eligible

XSL•FO
RenderX

students should receive access to the intervention resources and that establishing a nonintervention control group would be ethically inappropriate. Considering this recommendation and the small, hard-to-reach population size, a randomized controlled trial was deemed infeasible. Therefore, a before-after self-controlled cohort design was adopted, allowing participants to serve as their own controls to assess within-person changes.

Guided by findings from systematic reviews indicating that a 12-month intervention period is optimal for achieving behavioral changes [14], this study will set a 12-month intervention duration, with an interim follow-up assessment at 6 months.

Study Population

Inclusion Criteria

Participants will be eligible for inclusion if they meet all of the following criteria: (1) male, aged 18 years or older; (2) currently enrolled university student (including students pursuing technical diplomas or undergraduate degrees); (3) self-identified MSM, defined as having engaged in sexual behaviors with men, including mutual masturbation, oral sex, or anal intercourse; (4) confirmed HIV-negative status at baseline screening; and (5) willing and able to provide written informed consent to participate in the study.

Sample Size

To estimate the required sample size for a prevalent cohort under an exponential distribution assumption, we adopted the 2-sample Poisson rate comparison framework described by Shiue and Bain [15]. Detailed formulas and derivations are provided in Multimedia Appendix 1.

We assumed a baseline HIV incidence of 4 per 100 person-years (estimated from multiple MSM cohort studies in China [16-18]) and an expected postintervention incidence of 1 per 100 person-years (a 75% relative reduction), with a 2-sided α of .05, 80% power, and an anticipated 10% attrition over 12 months. Based on these assumptions, the required sample size is 484 participants.

The projected reduction to 1 per 100 person-years (75%) was empirically derived. Continuous surveillance of the MSM population in Chengdu observed a 60.9% year-over-year decline in new HIV infections between 2022 - 2023 and 2023 - 2024 (internal monitoring data available upon request). Building upon this trend, our multilevel intervention—integrating digital education, peer counseling, and PEP subsidies—is expected to produce further improvement. Preexposure prophylaxis (PrEP) combination-prevention intervention internationally has achieved 31.5% relative risk reduction among MSM [19], thus a 75% reduction was adopted as an empirically grounded upper-bound planning scenario.

We anticipated a 12-month retention rate of approximately 90% in this online cohort of student MSM. This assumption is informed by previous longitudinal studies among young men who have sex with men and gay/bisexual men, which have reported 12-month follow-up retention of 85% to 91% under structured follow-up procedures and supportive engagement strategies [20,21]. Given that our study uses multiple retention strategies (scheduled online follow-ups, advance notification of all incentive payments at enrollment, and close coordination through student volunteers), a 10% loss to follow-up was considered an ambitious but achievable design target. Sensitivity analyses at 25%, 50%, and 75% relative reductions and attrition rates of 10% to 20% were prespecified at the fixed N to assess feasibility and robustness (see Multimedia Appendix 2 for summary scenarios).

Study Setting

This study is scheduled to commence in June 2025 and conclude in June 2026. The research process will consist of the following sequential phases: participant recruitment, HIV/hepatitis C virus (HCV)/hepatitis B surface antigen (HBsAg)/syphilis testing, baseline survey, intervention implementation, midterm survey, and final follow-up survey.

Enrollment Procedure

Participant recruitment will be conducted exclusively through two internet-based channels to maximize efficiency and ensure privacy. The recruitment period is expected to last for 1 month. To ensure balanced workload distribution, each volunteer will have a preassigned recruitment quota. Once a volunteer has reached their quota, subsequent applicants will be redirected to other available volunteers. Given the estimated total sample size and the involvement of 10 student volunteers, it is calculated that each volunteer will be responsible for recruiting approximately 50 eligible student MSM participants.

Recruitment via WeChat

Recruitment information will be disseminated through a dedicated WeChat official account (WeChat is a Chinese multipurpose messaging, social media, and mobile payment app developed by Tencent Inc), which will publish targeted enrollment announcements clearly specifying the study's inclusion criteria. These announcements will also provide the WeChat contact information of designated student volunteers. Interested individuals will be instructed to contact these volunteers directly via WeChat, where the volunteers will conduct eligibility screening and coordinate preenrollment HIV and STI testing.

Recruitment via Blued

Additional recruitment will be conducted through Blued, a popular location-based social networking application specifically designed for MSM, developed by BlueCity Holdings (Beijing, China). Student volunteers will create recruitment accounts on Blued and prioritize outreach efforts at pilot universities—institutions identified through surveillance data as having historically higher HIV prevalence rates and larger student MSM populations.

Leveraging Blued's geolocation features, volunteers will target MSM individuals within their assigned university campuses. Given Blued's explicitly MSM user base, recruitment through this channel is expected to contribute at least 50% of the total sample size. To ensure consistency in intervention delivery, all participants recruited through Blued will also be required to add the corresponding volunteer's WeChat account, through which the subsequent online intervention activities will be administered.

XSL•FO
RenderX

Incentives

Participants who complete the baseline and follow-up survey will receive a modest monetary incentive. This compensation strategy serves two primary purposes: (1) to acknowledge participants' time and effort and (2) to foster participant retention by establishing a sense of commitment and accountability within the study.

Specifically, participants received RMB 20 (US $2.83) at baseline, RMB 20 (US $2.83) at the 6-month follow-up, and RMB 20 (US $2.83) at the final 12-month follow-up. The incentive amounts were reviewed and approved by the Ethics Committee of the Chengdu CDC (Approval No. 2024021). The total compensation has been carefully calibrated to ensure it remains modest and noncoercive while promoting continued engagement.

This incentive structure also incorporates behavioral-economics principles such as loss aversion, in which participants risk forfeiting future incentives if they withdraw prematurely. This approach encourages consistent participation without exerting undue influence.

HIV-HCV-HBsAg-Syphilis Testing

Overview

Participants will be required to complete self-testing for HIV, HCV, HBsAg, and syphilis at baseline, as well as during both follow-up surveys. Only individuals who test negative for HIV at baseline will be eligible for enrollment in the cohort. All participants are tested at each follow-up using a four-in-one rapid diagnostic kit that simultaneously detects HIV, syphilis, HCV, and HBV. Any reactive screening result triggers referral for confirmatory testing in certified laboratories. If a participant tests positive for any infection marker, they are immediately referred to a certified medical institution for confirmatory diagnosis and treatment, in accordance with national guidelines.

It should be specifically noted that we have also taken into account the issue of the serological window periods. To minimize potential misclassification, we have incorporated several measures including repeat testing for recent exposures and comprehensive risk documentation.

Repeat Testing for Recent Exposures

Participants who report any high-risk sexual exposure within the 3 months preceding a visit are scheduled for repeat rapid testing after 2 - 4 weeks.

Comprehensive Risk Documentation

Behavioral data from the past 3 months are collected at each follow-up survey. Participants reporting recent high-risk exposure but negative results are flagged as "possible window-period" cases and analyzed in sensitivity analyses.

Peer volunteers continuously follow up with referred participants to document their confirmatory results. The same referral and verification process applies to HIV, HCV, and HBV. This workflow ensures that all reported cases represent laboratory-confirmed new infections, minimizing classification bias.

HIV-HCV-HBsAg-Syphilis Self-Testing Procedure

Overview

HIV-HCV-HBsAg-Syphilis self-testing will be conducted using Rong Ai Jian, a WeChat-based point-of-care testing mini-program developed by the Chengdu CDC. This platform integrates CBOs engaged in HIV prevention efforts throughout Chengdu. Participants can request self-testing kits through the Rong Ai Jian mini-program on WeChat. The CBOs will mail the test kits directly to applicants, who will then complete self-testing, upload their results, and receive professional review and feedback via the platform. This mechanism has been operational for over 2 years and has demonstrated stable functionality and effectiveness in practice.

The platform provides a complete traceable workflow, ensuring the accuracy and authenticity of self-test results.

Kit Distribution and Coding

Each four-in-one rapid test kit is prelabeled with a unique ID number by the CBO staff before shipment.

Result Reporting

After completing the self-test according to the instruction sheet, participants are required to upload a clear photo of the used test kit through the Rong Ai Jian WeChat mini-program.

Real-time Verification

CBO staff can view the uploaded image immediately on the platform and verify the result visually. The staff will then record the verified result in the platform system by selecting "Negative," "To Be Retested," or "Invalid."

Participant Confirmation

The verified result automatically appears in the participant's mini-program interface. Participants can screenshot this verification and send it to their peer volunteer as proof of testing completion.

Referral for Confirmation

Cases marked as "To Be Retested" are treated as preliminary reactive results. Both the CBO staff and peer volunteer then refer the participant to a qualified medical institution for confirmatory testing.

Final Confirmation and Data Entry

Only results that are confirmed by certified laboratories (eg, by treponemal and non-treponemal tests for syphilis or confirmatory assays for HIV/HBV/HCV) are entered as positive outcomes in the research database.

HIV-HCV-HBsAg-Syphilis Combo Rapid Testing

The study will utilize the quadruple rapid diagnostic test kit developed by Guangzhou Wondfo Biotech Co., Ltd. This assay enables the simultaneous qualitative detection of HIV-1/2 antibodies, HCV antibody, HBsAg, and Treponema pallidum antibody from whole blood, serum, or plasma samples. The test employs a lateral flow immunochromatographic assay and provides results within 15 minutes.

XSL•FO
RenderX

Each test kit includes a test cassette, desiccant pouch, dropper, and buffer vial. The HIV-HCV-HBsAg-Syphilis Combo Test is approved by the China National Medical Products Administration, ensuring its safety and efficacy for clinical application.

Independent evaluations have reported excellent diagnostic performance, with sensitivities of 99.70% for HIV, 98.71% for HCV, 98.97% for HBsAg, and 98.47% for syphilis, along with corresponding specificities of 100%, 99.33%, 99.78%, and 99.68%, respectively. This highly reliable and user-friendly assay is particularly well-suited for point-of-care testing and large-scale screening initiatives, enabling timely diagnosis and facilitating rapid linkage to appropriate health care services [22].

Questionnaire Survey

Baseline data collection will be conducted via an online self-administered questionnaire. Assigned project volunteers will securely deliver the survey link to participants through encrypted messaging channels.

Electronic Informed Consent

An electronic informed consent process will be embedded at the beginning of the baseline questionnaire. Participants will be required to carefully read the informed consent form, which details the purpose of the study, potential risks and benefits, data confidentiality measures, and their rights as participants. Only after acknowledging their understanding and agreement will participants be permitted to proceed with completing the survey.

Generation of Unique Participant ID and Data Security

To ensure participant anonymity while allowing for longitudinal data linkage, a de-identified participant code will be generated for each individual at enrollment. The participant ID consists of a randomly generated alphanumeric string created through an independent encryption algorithm, ensuring that no elements of phone numbers, timestamps, or other personally identifiable information (PII) are embedded in the code. The mapping file linking IDs with PII will be stored separately in an encrypted database accessible only to authorized Chengdu CDC data managers.

All study data are managed within a comprehensive data protection and security framework, consistent with national and international privacy standards. PII (such as phone numbers, WeChat IDs, and mailing addresses) is collected solely for operational purposes and stored separately from deidentified research data. Both databases are encrypted at rest and during transmission using the Advanced Encryption Standard (AES-256) encryption protocol implemented within the Chengdu CDC secure data environment.

A Data Protection Impact Assessment was performed to identify potential privacy risks and define mitigation strategies, including data-flow mapping, encryption and storage measures, access-control policies, retention and deletion schedules, and oversight procedures.

Access to data is governed by role-based access control, ensuring that each staff member only accesses the data required for their function. Multifactor authentication (eg, password+phone verification) is required for system login. All access and data modification activities are recorded through automated audit logs, which are periodically reviewed by the Chengdu CDC data security officer.

In accordance with the approved data retention and deletion schedule, PII will be retained for up to 6 months after study completion and then permanently deleted from all servers. Deidentified research data will be securely stored for 5 years for potential secondary analysis and subsequently removed after encrypted archiving.

These combined administrative, technical, and procedural safeguards ensure participant anonymity and protect the confidentiality, integrity, and traceability of all data throughout the study lifecycle.

Survey Content and Follow-Up

Participants will complete a baseline questionnaire upon enrollment, which collects detailed information on sociodemographic characteristics, sexual behaviors, HIV-related knowledge, attitudes, practices, and other relevant behavioral factors (see Multimedia Appendix 3 for details).

Follow-up surveys will be administered at two time points post-intervention: 6 months and 12 months. The content of the follow-up questionnaires will be identical to that of the baseline survey to enable assessment of changes over time (see Multimedia Appendix 4 for follow-up questionnaire content). Data collection instruments and operational manuals have been standardized to ensure consistency in measurement and facilitate reproducibility across different investigators.

Intervention Process

Developing Intervention Package

The intervention will be conducted entirely online over a 12-month period. Under the guidance of HIV prevention experts and CBO trainers, we have developed a comprehensive and targeted intervention package for participants. This package includes a standardized online intervention workflow (Figure 1), evidence-based web-based content, and accessible links to self-testing kits and referral services. A detailed schematic of the participant flow and multilevel intervention process is presented in Figure 1. The diagram depicts the sequential stages of recruitment, screening, enrollment, baseline assessment, intervention delivery, and follow-up. Detailed standard operating procedures have been developed to standardize the delivery frequency, message content, and volunteer-participant interactions, ensuring consistency across all intervention components. To ensure consistency and standardization in the delivery of peer-led interventions, the project implements a multilayered training, certification, and supervision system.

Standardized Training Curriculum

All peer volunteers undergo a unified, structured training program jointly conducted by the Chengdu CDC and partner CBOs. The training covers HIV prevention knowledge,

XSL•FO
RenderX

motivational interviewing skills, online communication protocols, and ethical considerations (confidentiality and noncoercive engagement).

XSL•FO
RenderX

Figure 1. Participant flow and intervention process of the study. Rounded rectangles denote the core procedural stages of participant enrollment and intervention and follow-up. Stadium-shaped nodes indicate critical follow-up checkpoints. Circles represent study end points, including completion and loss to follow-up. Square boxes indicate detailed information units (eg, eligibility criteria and intervention components). Solid arrows indicate participant progression through the study procedures, while dashed arrows indicate attrition paths. MSM: men who have sex with men.

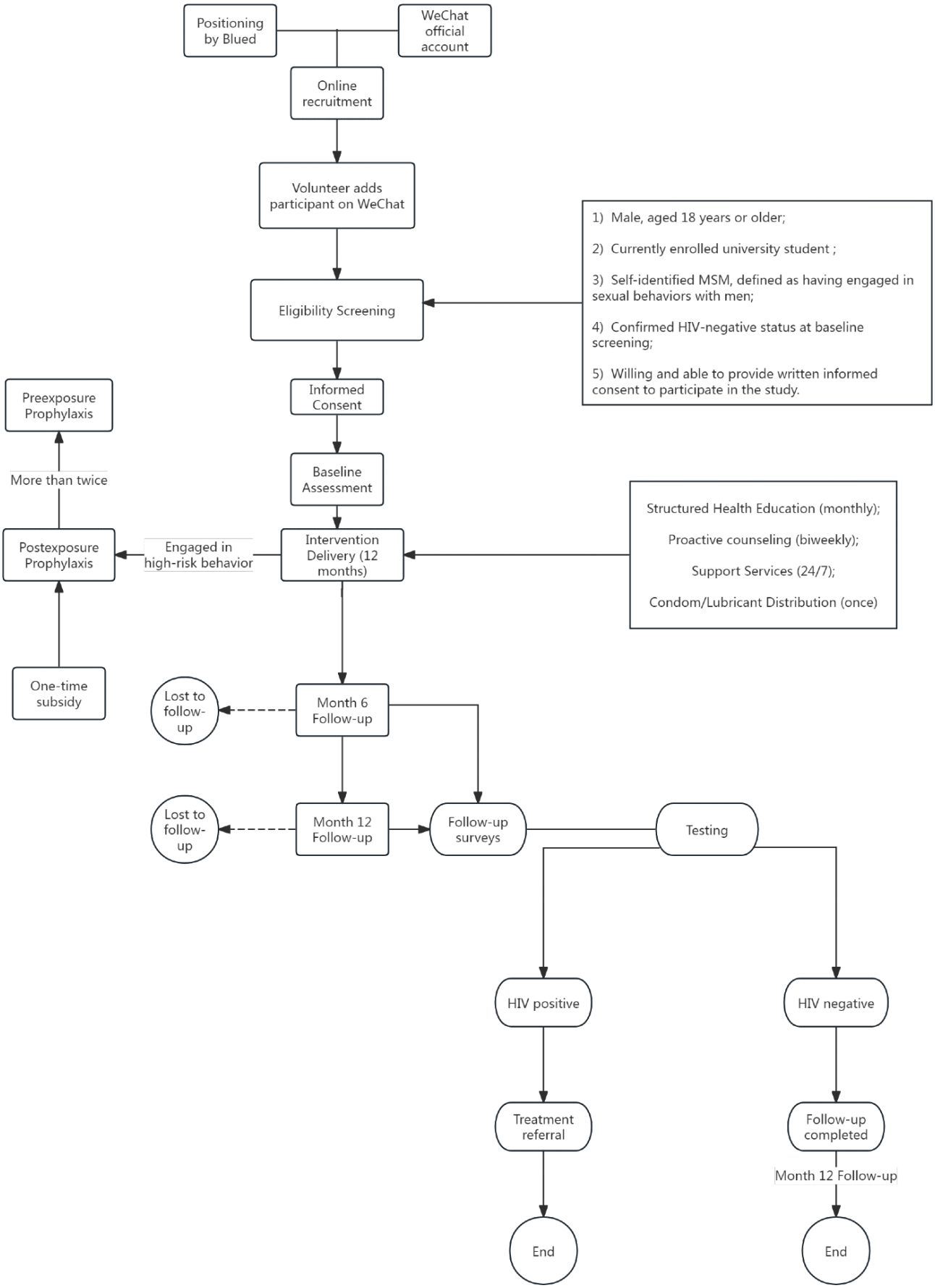

XSL•FO
RenderX

Posttraining Assessment and Certification

After training, each volunteer must pass a competency assessment that evaluates communication, adherence to intervention procedures, and knowledge of referral pathways. Only volunteers who pass this assessment are certified to conduct interventions.

Guideline Support

Each volunteer receives a "Guidelines for Internet-Based HIV Interventions," which includes frequently asked questions and recommended responses. This guideline serves as a reference to ensure uniform content delivery and quality control during intervention sessions.

On-Site and Ongoing Supervision

Experienced CBO staff serve as trainers and provide continuous mentorship to peer volunteers throughout the intervention phase. These trainers, together with CDC supervisors, monitor intervention implementation through regular check-ins, feedback sessions, and review of anonymized communication logs when needed.

Capacity Building for Project Volunteers

To enhance the understanding of the student MSM volunteers on the project rationale and improve their capacity to implement interventions, a structured training program will be provided prior to project initiation. This training will include both theoretical and practical components, such as (1) lectures to be delivered by experts and CBO trainers on HIV prevention strategies (eg, PrEP and PEP), online counseling skills, and psychological communication techniques; (2) practical simulation exercises focusing on real-world scenarios and case discussions; (3) orientation sessions explaining project-specific tasks, performance assessment indicators, and operational guidelines.

Intervention Implementation

Overview

Trained student volunteers will deliver the intervention package to participants via Enterprise WeChat (a secure communication platform). The package will consist of four key components: structured health education modules, proactive counseling services, PEP assistance, and HIV prevention supply distribution.

Structured Health Education Modules

Monthly educational content will be disseminated, featuring curated articles from authoritative HIV prevention platforms operated by CBOs or national organizations (eg, China CDC official accounts). Each module focuses on a specific theme; content development will prioritize cultural relevance and adolescent-friendly design, utilizing infographic-style visual prescriptions tailored to student MSM cognitive preferences.

Educational content will adhere to national guidelines, including Guidelines for HIV Prevention among MSM [23] and Guidelines for Internet-based HIV Interventions [24]. Core content areas will include (1) HIV/STI transmission risks and condom-compatible lubricant usage; (2) PEP/PrEP knowledge and access pathways (eg, initiation within 72 h for PEP); (3) geolocation-enabled testing site information and MSM-specific epidemiological updates; and (4) risks associated with the use of addictive sexual enhancement products.

Proactive Counseling Services

Each participant receives biweekly proactive one-on-one counseling conducted by peer volunteers using motivational-interviewing techniques to reinforce safer sexual behaviors and strengthen adherence to preventive measures. These online sessions typically last 10 to 15 minutes and are delivered through WeChat. Core topics include HIV risk-reduction planning, emotional well-being, stigma coping, and linkage to HIV testing or care services.

In addition, a 24/7 online crisis-support channel is available for participants experiencing psychological distress, substance-use crises, or urgent HIV-related concerns. Volunteers are trained to provide first-line counseling and escalate complex cases to CDC supervisors or professional counselors within 24 h.

PEP Assistance

For participants reporting high-risk exposures (eg, condomless anal intercourse or needle-sharing), volunteers will provide the following:

Rapid linkage to PEP services: Same-day appointment coordination with PEP-prescribing outpatient clinics using a geolocation database.

Financial assistance: A one-time subsidy of RMB 500 (US $70.88) will be provided upon verified PEP completion (submission of clinic-issued proof). All PEP referrals are monitored through CDC-verified reporting channels to ensure timely initiation and adherence support. Participants with repeated high-risk exposures were assessed for PrEP eligibility in accordance with the Expert Consensus on HIV PrEP in China [25].

HIV Prevention Supply Distribution

Participants will receive a discreetly packaged prevention kit via third-party logistics, ensuring privacy protection. The kit will include condoms, water-based lubricants, and STI self-testing kits.

Participants will receive a discreetly packaged prevention kit via third-party logistics, ensuring privacy protection. The kit will include condoms, water-based lubricants, and STI self-testing kits.

Distribution will be limited to a one-time delivery per participant to optimize resource allocation.

Volunteers track delivery confirmation through the digital management platform.

Study End Points

The study has three predefined endpoints:

1. HIV seroconversion: Participants with a reactive HIV test result will receive posttest counseling and referral to the local CDC for confirmatory laboratory testing. If HIV infection is laboratory-confirmed, the participant will be

XSL•FO
RenderX

withdrawn from further study follow-up and intervention activities.

2. Loss to follow-up: Participants will be considered lost to follow-up if they fail to complete follow-up assessments and cannot be contacted despite repeated attempts.
3. Study completion: Participants who complete the 12-month follow-up period and final HIV/STI testing will be considered as having completed the study.

Outcome Measures

Primary Outcome

HIV and syphilis incidence rate: The primary outcome is the change in HIV and syphilis incidence rates before and after the intervention, assessed by comparing the postintervention incidence with the baseline incidence among local student MSM populations.

Secondary Outcomes

Secondary outcome measures include (1) changes in sexual behaviors (eg, condom use, number of sexual partners) before and after the intervention; (2) frequency of biomedical intervention utilization (eg, HIV testing, STI testing, PEP, PrEP use); and (3) changes in HIV and STI prevention knowledge and awareness levels.

Detailed definitions of outcome indicators, measurement methods, and data sources are summarized in Table 1.

XSL•FO
RenderX

Table . Indicators of study outcomes.

Domain	Specific indicator	Definition	Data source
Primary outcomes			
HIV incidence	• Comparison between intervention cohort and local student MSM attending university	• New HIV cases confirmed by medical institutions per 100 person-years vs surveillance data (2025)	Laboratory reports at 12-month follow-up; CDC[b] surveillance data
Syphilis incidence	• Comparison between intervention cohort and local student MSM attending university	• New syphilis cases confirmed by medical institutions per 100 person-years vs surveillance data (2025)	Laboratory reports at 12-month follow-up; CDC surveillance data
Secondary outcomes			
Behavioral Changes			
Risk behaviors	• Condom use during anal intercourse • Commercial sex engagement • Multiple partners	• % reporting “always use condoms” in past 3 months • % reporting paid sex in past 3 months • % with ≥2 sexual partners in past 3 months	Structured questionnaire
Substance use	• Psychoactive substance use during sex	• % reporting use of chemsex drugs in past 3 months	Structured questionnaire
Testing behaviors	• HIV testing frequency • STI[c] screening	• % tested for HIV in past 3 months • % tested for STIs in past 3 months	Structured questionnaire
Biomedical uptake			
Prevention tools	• Condom/lubricant use • PrEP[d]/PEP[e] utilization	• % using condom/lubricant • % PrEP/PEP initiations	Supply logs; pharmacy data
Attitudinal changes			
Prevention intention	• Future condom use intent • Testing willingness	• % planning to use condoms • % willing to undergo future HIV/STI tests	Structured questionnaire
Awareness changes			
Health literacy	• PEP/PrEP knowledge • HIV/STIs awareness	• % correctly describing PEP/PrEP protocols • % choosing the right core knowledge about HIV/STIs	Knowledge assessment

[a]MSM: men who have sex with men.

[b]CDC: Centers for Disease Control and Prevention.

[c]STI: sexually transmitted infection.

[d]PrEP: preexposure prophylaxis.

[e]PEP: postexposure prophylaxis.

Participant Timeline

Time schedule of enrollment, interventions, assessments, and visits for participants are presented in Table 2.

XSL•FO
RenderX

Table . Schematic diagram for the schedule of enrollment, interventions, and assessments.

	Study period								
Timepoint and enrollment	Allocation	Postallocation							Close-out
	M0[a]	M1	M2	M3	M4	M5	M6	M7-11	M12
Enrollment									
Recruitment	✓								
Eligibility screen	✓								
Informed consent	✓								
Intervention									
Health education		✓	✓	✓	✓	✓	✓	✓	
Counseling services		✓	✓	✓	✓	✓	✓	✓	
HIV prevention supply distribution		✓							
PEP[b] assistance		✓	✓	✓	✓	✓	✓	✓	
Assessments									
Baseline survey	✓								
Follow-up survey							✓		✓
Testing							✓		✓

[a]M: Month

[b]PEP: postexposure prophylaxis.

Data Management

Survey data will be collected through Wenjuanxing, an online survey platform developed and operated by Ranxing Information Technology Co., Ltd., based in Changsha, China [26]. Wenjuanxing provides an automated data archiving system designed to ensure the integrity, security, and consistency of submitted responses.

All survey data will be stored on secure servers hosted by Alibaba Cloud in Hangzhou, China. The server infrastructure is protected by enterprise-level firewall systems, with daily automated data backups to prevent potential data loss and ensure disaster recovery capabilities.

Access to the database is strictly controlled through a role-based access control system. Only authorized research personnel, who have signed confidentiality agreements, will be granted access to view or manage the data. User authentication protocols, including secure passwords and encrypted data transmission, are implemented to further enhance data security and prevent unauthorized access.

Statistical Analysis

Upon completion of data collection, all survey data will be exported from the Wenjuanxing platform for data cleaning and preprocessing to ensure accuracy, completeness, and consistency. Data management and analysis will be conducted using established statistical software packages, including SPSS, Stata, or R.

Descriptive statistics will be used to summarize participant characteristics and key study variables. Categorical variables will be presented as frequencies and percentages, while continuous variables will be reported as means with SD or medians with IQR, depending on data distribution. The normality of continuous variables will be assessed using Shapiro-Wilk tests and visual inspection of histograms/Q–Q plots. Univariate analyses, such as chi-square tests for categorical variables and *t* tests or nonparametric tests for continuous variables, will be used to assess preliminary associations between independent variables and outcome measures. Multivariate analysis will be performed using logistic regression models (for binary outcomes) to identify factors independently associated with key study end points, adjusting for potential confounders. If advanced statistical modeling (eg,

XSL•FO
RenderX

Cox proportional hazards models, generalized estimating equations) is deemed necessary to address specific research questions or to handle longitudinal data, additional statistical expertise will be sought to ensure appropriate application and interpretation of these methods. A 2-sided *P* value <.05 will be considered statistically significant.

Feasibility and Acceptability Evaluation

Feasibility will be assessed by examining recruitment and retention rates, completeness of follow-up surveys, and timeliness of testing. Acceptability will be evaluated through participant satisfaction surveys and qualitative feedback collected from peer volunteers regarding the practicality and perceived usefulness of the intervention materials. Descriptive statistics will be used to summarize feasibility and acceptability outcomes, and qualitative comments will be thematically analyzed.

Results

Participant enrollment and data collection are currently ongoing. This trial protocol has been developed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 guidelines [27].

As of May 1, 2025, the intervention materials and training program have been finalized, and recruitment has begun. Enrollment of participants started in April 2025. By the end of April, 127 participants had completed baseline surveys and HIV/STI self-tests.

Discussion

Key Innovation Points

This prospective cohort study will be uniquely designed to evaluate the effectiveness of an internet-based intervention in reducing HIV-related high-risk behaviors and the prevalence of HIV and other STIs among university student MSM. Beyond addressing the primary study objectives, the research will innovatively integrate digital engagement strategies with peer-led support, representing a pioneering approach within this population.

Specifically, the study will establish a dedicated intervention cohort that will utilize social networking platforms to deliver continuous behavioral interventions, promote HIV/STI testing, facilitate service referrals, and provide financial subsidies to support access to PEP.

A key innovation of this study will be the application of the Network Intervention Cohort Peer Model, which will involve recruiting experienced staff from MSM CBOs as trainers and engaging university student MSM volunteers as frontline interventionists. This model is expected to enhance the cultural relevance and acceptability of the intervention while addressing critical limitations of conventional peer education programs.

Furthermore, all interventions will be delivered entirely online, aligning with current trends in MSM-focused health promotion. The intervention package will integrate both behavioral and biomedical components—including the discreet distribution of condoms and lubricants by mail, as well as financial subsidies to support timely initiation of PEP when indicated. Prior research has demonstrated that eliminating financial barriers is a key strategy for empowering MSM to access and sustain engagement with essential HIV prevention services [28].

Recent regulatory tightening surrounding CBO activities—particularly in relation to direct engagement with university students—has significantly constrained the scope of traditional peer education efforts. While Chengdu has previously mobilized college student volunteers through partnerships with local CDCs and AIDS Prevention Associations, such initiatives have generally been broad-based and not specifically tailored to the distinct needs of university student MSM.

Against this backdrop, our project will represent the first known effort in China to design and implement a comprehensive, peer-led network intervention model specifically targeting this demographic.

Despite the study's innovative strengths, several anticipated operational challenges merit careful consideration.

Standardization of Volunteer Training and Mentoring

The project relies on CBO staff serving as trainers and student MSM volunteers acting as interventionists. Ensuring consistency and quality in training and mentoring is critical to maintaining intervention fidelity. Variability in training delivery or gaps in knowledge may lead to inconsistent intervention outcomes.

To address this, trainers will be selected from experienced CBO staff across Chengdu with a demonstrated track record in intervention delivery and volunteer mentoring. All volunteers are required to complete the entire capacity-building training and pass a posttraining assessment before participating in interventions. During the intervention process, volunteers can consult their trainers at any time regarding operational issues, such as making referrals or responding to participant inquiries, and can also seek support through the project's online group chat.

Recruitment and Retention of a Fixed Cohort

Maintaining participant retention throughout the 12-month follow-up period is crucial to ensuring sufficient statistical power and minimizing attrition bias. While multiple recruitment strategies and financial incentives have been planned, challenges in achieving the targeted sample size may still arise.

To mitigate this risk, contingency plans include expanding recruitment channels, such as peer referrals and targeted social media advertisements, if enrollment falls short. Additionally, each participant will be required to provide both their WeChat contact and mobile phone number upon enrollment. Participants will also be informed that this information, along with their mailing address, will be used solely for study-related purposes, such as the confidential delivery of condoms, lubricants, and other prevention materials, which may help improve retention and engagement.

XSL•FO
RenderX

Data Collection and Baseline Assessment: Ensuring Reliability of Self-Reported Data

The baseline survey collects sensitive data, including sexual behaviors, substance use, and HIV-related risk factors. Due to social desirability bias or privacy concerns, participants may underreport or misreport behaviors, potentially affecting data accuracy.

To address this, volunteers will emphasize the project's strict confidentiality protocols when inviting participants to complete the questionnaire. Participants will be clearly informed that no PII is collected in the survey, all data are securely stored in the back-end system, and only the project manager has access to detailed records.

This study has limitations. The before-after self-controlled cohort design does not include a concurrent control group, which limits the ability to make causal inferences. However, a randomized controlled trial was not feasible due to ethical and operational considerations—specifically, the local ethics committee's recommendation that all eligible university students should have access to the intervention resources. To mitigate this limitation, we will compare changes observed in this study with contemporaneous surveillance data from the Chengdu student MSM population, conduct sensitivity analyses to assess secular trends, and interpret outcomes cautiously in light of these contextual factors. Despite these constraints, the design remains appropriate for evaluating real-world implementation and effectiveness under routine public health conditions. While the assumed reduction in incidence may be optimistic, behavioral indicators such as HIV testing and PEP/PrEP uptake serve as sensitive early markers of intervention impact in this short-term cohort. Given the expected rarity of incident infections over 12 months in student MSM cohorts, we interpret incidence changes cautiously while relying on proximal behavioral outcomes with higher short-term sensitivity.

By proactively addressing these anticipated challenges through standardized training, robust recruitment and retention strategies, and comprehensive privacy protections, this study aims to ensure rigorous implementation while generating valuable insights into the feasibility and effectiveness of internet-based, peer-led HIV prevention interventions among university student MSM.

Acknowledgments

We would like to thank the staff of the Chengdu Tongle Social Work Service Center, whose extensive experience helped us develop the protocol for this study. The authors did not use ChatGPT or any other generative artificial intelligence tools in the generation, analysis, or writing of this manuscript.

Funding

This work was supported by the Chengdu Center for Disease Control and Prevention (Chengdu Institute of Health Supervision) as part of its local public health intervention program.

Data Availability

The data generated or analyzed during this study are not yet publicly available because the project is ongoing. Upon study completion, deidentified data will be available from the corresponding author (Liang Wang, Chengdu Center for Disease Control and Prevention) upon reasonable request for academic purposes, subject to ethical approval and data-sharing agreements.

Authors' Contributions

Conceptualization: LW, JX
Methodology: JX, ZD, YQ
Investigation and data collection: ZD, YQ
Formal analysis: JX, LW
Writing – original draft: JX
Writing – review & editing: LW, JX
Supervision: LW
All authors have read and approved the final version of this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Formulas and derivations.
[DOCX File, 21 KB - resprot_v15i1e77078_app1.docx]

Multimedia Appendix 2
Summary scenarios.

XSL•FO
RenderX

[DOCX File, 14 KB - resprot_v15i1e77078_app2.docx]

Multimedia Appendix 3
Baseline questionnaire.
[DOCX File, 15 KB - resprot_v15i1e77078_app3.docx]

Multimedia Appendix 4
Follow-up questionnaire.
[DOCX File, 15 KB - resprot_v15i1e77078_app4.docx]

References

1. Zhihua Z, Jun W, Yue Z, Wei W, et al. Meta analysis of the high risky behaviors and detection rate of HIV infection among MSM students in Mainland China. Chin J School Health 2018;39(5):702-705. [doi: 10.16835/j.cnki.1000-9817.2018.05.018]
2. Chengdu's Seventh National Population Census Bulletin [1] (no 1)— citywide population overview. Chengdu Municipal Statistics Bureau. 2023. URL: https://cdstats.chengdu.gov.cn/cdstjj/uploads/20230222161411zvvby5vtlzo.pdf [accessed 2025-05-07]
3. Dai Z, Wang X, Liu F, Duan Z, Shi Y, Zhang S, et al. Applying multiplier method to estimate the population size of active MSM in Chengdu based on social App data. Modern Prevent Med 2019;46(24):4421-4423 [FREE Full text]
4. Dai Z, He Q, Wang X, Liu F, Zha H. Estimating the MSM cohort size of college students in Chengdu using social software. Chin J Health Stat 2022;39(5):752-754. [doi: 10.11783/j.issn.1002-3674.2022.05.026]
5. Li Y, Xu J, Reilly KH, et al. Prevalence of HIV and syphilis infection among high school and college student MSM in China: a systematic review and meta-analysis. PLOS ONE 2013;8(7):e69137. [doi: 10.1371/journal.pone.0069137]
6. Xu JJ, Reilly KH, Lu CM, et al. A cross-sectional study of HIV and syphilis infections among male students who have sex with men (MSM) in northeast China: implications for implementing HIV screening and intervention programs. BMC Public Health 2011 May 10;11:287. [doi: 10.1186/1471-2458-11-287] [Medline: 21554742]
7. Yang S, Fan C, Li J, Zhang J, Yin W, Zhang J. Present situation and influencing factors of high-risk sexual behaviors among university MSM students in Chengdu. J Preventive Med Inform 2018;34(7):866-870 [FREE Full text]
8. Fan S, Liu Z, Luo Z, et al. Effect of availability of HIV self-testing on HIV testing frequency among men who have sex with men attending university in China (UniTest): protocol of a stepped-wedge randomized controlled trial. BMC Infect Dis 2020 Feb 18;20(1):149. [doi: 10.1186/s12879-020-4807-4] [Medline: 32070297]
9. Cheng W, Xu H, Tang W, et al. Online HIV prevention intervention on condomless sex among men who have sex with men: a web-based randomized controlled trial. BMC Infect Dis 2019 Jul 19;19(1):644. [doi: 10.1186/s12879-019-4251-5] [Medline: 31324233]
10. Wang Z, Chan PSF, Xin M, et al. An online intervention promoting HIV testing service utilization among Chinese men who have sex with men during the COVID-19 pandemic: a quasi-experimental study. AIDS Behav 2024 Oct;28(Suppl 1):77-89. [doi: 10.1007/s10461-023-04100-5] [Medline: 37284925]
11. Li C, Wang X, Cui Z, et al. Association between peer education and HIV testing among student MSM in china: a cross-sectional study. Res Square. Preprint posted online on 2020. [doi: 10.21203/rs.2.22565/v1]
12. Muchiri BW, Dos Santos MML. Family management risk and protective factors for adolescent substance use in South Africa. Subst Abuse Treat Prev Policy 2018 Jun 19;13(1):24. [doi: 10.1186/s13011-018-0163-4] [Medline: 29914541]
13. Dos Santos LM. Promoting safer sexual behaviours by employing social cognitive theory among gay university students: a pilot study of a peer modelling programme. Int J Environ Res Public Health 2020 Mar 10;17(5):1804. [doi: 10.3390/ijerph17051804] [Medline: 32164309]
14. Nguyen LH, Tran BX, Rocha LEC, et al. A systematic review of eHealth interventions addressing HIV/STI prevention among men who have sex with men. AIDS Behav 2019 Sep;23(9):2253-2272. [doi: 10.1007/s10461-019-02626-1] [Medline: 31401741]
15. Shiue WK, Bain LJ. Experiment size and power comparisons for two-sample Poisson tests. Appl Stat 1982;31(2):130. [doi: 10.2307/2347975]
16. Mi GD, Zhang BB, Yu F, et al. HIV incidence among men who have sex with men using geosocial networking smartphone application in Beijing, China: an open cohort study. Infect Dis Poverty 2021 Apr 2;10(1):27. [doi: 10.1186/s40249-021-00814-7] [Medline: 33795011]
17. You X, Gilmour S, Cao W, et al. HIV incidence and sexual behavioral correlates among 4578 men who have sex with men (MSM) in Chengdu, China: a retrospective cohort study. BMC Public Health 2021 Apr 26;21(1):802. [doi: 10.1186/s12889-021-10835-4] [Medline: 33902499]
18. Dai Z, Mi G, Yu F, Chen G, Wang X, He Q. Using a geosocial networking app to investigate new HIV infections and related risk factors among student and nonstudent men who have sex with men in Chengdu, China: open cohort study. J Med Internet Res 2023 Jul 28;25(1):e43493. [doi: 10.2196/43493] [Medline: 37505891]

XSL•FO
RenderX

19. Grulich AE, Guy R, Amin J, et al. Population-level effectiveness of rapid, targeted, high-coverage roll-out of HIV pre-exposure prophylaxis in men who have sex with men: the EPIC-NSW prospective cohort study. Lancet HIV 2018 Nov;5(11):e629-e637. [doi: 10.1016/S2352-3018(18)30215-7] [Medline: 30343026]
20. Saxton PJW, Azariah S, Cavadino A, et al. Adherence, sexual behavior and sexually transmitted infections in a New Zealand prospective PrEP cohort: 12 months follow-up and ethnic disparities. AIDS Behav 2022 Aug;26(8):2723-2737. [doi: 10.1007/s10461-022-03617-5] [Medline: 35167038]
21. Mustanski B, Newcomb ME, Clerkin EM. Relationship characteristics and sexual risk-taking in young men who have sex with men. Health Psychol 2011 Sep;30(5):597-605. [doi: 10.1037/a0023858] [Medline: 21604883]
22. Wondfo HIV-HCV-hbsag-syphilis combo test. : Guangzhou Wondfo URL: https://en.wondfo.com/vancheerfile/files/2023/3/20230313021625582.pdf [accessed 2025-12-06]
23. MSM Intervention Guidelines Information on HIV prevention network interventions for men who have sex with men [Report in Chinese]. : China CDC URL: https://www.chinacdc.cn/jkyj/crb2/yl/azb/jswj_azb/202410/W020241027556171210473.pdf [accessed 2026-01-06]
24. Internet-based HIV intervention guidelines [Report in Chinese]. : China CDC URL: https://ncaids.chinacdc.cn/zxzx/zxzx/202101/W020210128637953503812.pdf [accessed 2026-01-06]
25. Chinese Association for STD and AIDS Prevention and Control Special Committee on HIV/AIDS Drug Prevention and Post-Exposure Prophylaxis. Chinese Expert Consensus on Pre-Exposure Prophylaxis for HIV (2023 Edition) [Article in Chinese]. Chin J AIDS STD 2023;29(9):954-961 [FREE Full text] [doi: 10.13419/j.cnki.aids.2023.09.02]
26. Wenjuanxing. URL: https://www.wjx.cn/ [accessed 2026-01-06]
27. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346:e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
28. Zhou J, Xu Y, Li Q, et al. Understanding awareness, utilization, and the awareness-utilization gap of HIV PrEP and nPEP among young MSM in China. AIDS Behav 2025 Apr;29(4):1327-1339. [doi: 10.1007/s10461-024-04606-6] [Medline: 39779626]

Abbreviations

CBO: community-based organization
CDC: Centers for Disease Control and Prevention
HBsAg: hepatitis B surface antigen
HCV: hepatitis C virus
MSM: men who have sex with men
PEP: postexposure prophylaxis
PII: personally identifiable information
PrEP: preexposure prophylaxis
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
STI: sexually transmitted infection

Edited by J Sarvestan; submitted 07.May.2025; peer-reviewed by S Chen; revised version received 21.Nov.2025; accepted 24.Nov.2025; published 23.Jan.2026.

Please cite as:
Xu J, Dai Z, Qiu Y, Wang L
A Multilevel Network Peer Intervention Among Student Men Who Have Sex With Men Attending University: Protocol for an Implementation-Effectiveness Before-After Cohort Study
JMIR Res Protoc 2026;15:e77078
URL: https://www.researchprotocols.org/2026/1/e77078
doi:10.2196/77078

© Jingpei Xu, Zhen Dai, Yushu Qiu, Liang Wang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of an Ethico-Legal Framework for Quality Improvement and Performance Management in Health Care: Protocol for a Qualitative Study

Kavisha Shah[1], BSc (Hons), LLB; Anna Janssen[1], BA, MA, PhD; Cameron Stewart[2], BEc, LLB (Hons), PhD; Tim Shaw[1], BSc (Hons), PhD; Ian Kerridge[3,4,5], BA, BMed (Hons), MPhil

[1]Faculty of Medicine and Health, University of Sydney, Level 2, Charles Perkins Centre, John Hopkins Drive, Sydney, NSW, Australia

[2]Sydney Law School, The University of Sydney, Sydney, Australia

[3]Sydney Health Ethics, School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[4]Haematology Department, Royal North Shore Hospital, Northern Sydney Local Health District, Sydney, Australia

[5]Philosophy Department, Macquarie University, Sydney, Australia

Corresponding Author:
Kavisha Shah, BSc (Hons), LLB
Faculty of Medicine and Health, University of Sydney, Level 2, Charles Perkins Centre, John Hopkins Drive, Sydney, NSW, Australia

Abstract

Background: The growing digitization of health data has expanded opportunities for professional learning and performance improvement. While they provide new means for improving the quality and safety of health care, these new capabilities for data analysis and performance monitoring come with risks and may exacerbate existing ethico-legal concerns about fairness, accountability, privacy, and more.

Objective: This study aims to develop an ethico-legal framework for the evaluation of professional performance that is cognizant of these concerns and addresses the needs of relevant stakeholders. The study will assess the acceptability, comprehensiveness, and potential utility of the framework from the perspective of end users and subject matter experts.

Methods: This study will use existing evidence on ethico-legal considerations surrounding secondary uses of health data for performance improvement and management to draft the framework. We will conduct 2 focus groups with end users (eg, health professionals and administrators) and subject matter experts (eg, clinical ethicists and legal practitioners). These focus groups will ask participants to reflect on the framework's structure and comprehension, intended audience, comprehensiveness and relevance regarding ethical and legal principles, limitations, and utility and acceptability as a step-by-step guide. Study participants may also opt for one-on-one interviews for any reason. This feedback will be thematically analyzed using open coding and verified by an independent reviewer at the focus groups, followed by constant comparisons of feedback from this study to concepts and interrelationships in data previously collected.

Results: Recruitment for this study is scheduled from August to December 2025. The analysis, compilation, and dissemination of higher-order themes, concepts, and outcomes is planned for after publication of this protocol, after each interview or focus group has been transcribed and coded line by line.

Conclusions: This study seeks to create an actionable tool that is readily translatable to clinical practice in collaboration with end users and subject matter experts. The proposed methodology is a low-resource coapproach that could be iteratively refined to ensure that the proposed framework continues to support robust and efficient use of performance data while respecting the different contexts in which practice analytics may be delivered. This systematic approach to principle-led evaluation of performance and conduct could inform technology-neutral governance capable of addressing perennial concerns about fairness, privacy, and transparency when using health data for professional learning and performance management.

International Registered Report Identifier (IRRID): DERR1-10.2196/82167

(JMIR Res Protoc 2026;15:e82167) doi:10.2196/82167

KEYWORDS

health policy; ethics; privacy; transparency; procedural fairness; law; health data

XSL•FO
RenderX

Introduction

Background

Reflective practice is foundational to professional practice in health care. It directly contributes to the quality and safety of patient care by encouraging health professionals to identify gaps in their knowledge, skill, or expertise and analyze failures, creating opportunities for professional development and growth [1]. Traditionally, health care has relied on self-nomination to identify deficiencies in professional performance [2], but this has limited effectiveness as health professionals frequently struggle to accurately assess their performance [3], identify their clinical or professional deficiencies or limitations, or use appropriate learning styles or tools to address knowledge or skill gaps [4]. In response, professional associations and regulatory bodies internationally have increasingly incorporated "practice analytics," which broadly refers to the use of health data such as clinical outcomes and key performance indicators in performance evaluation as a condition of registration and/or accreditation [5]. While the use of practice analytics as a condition of registration is new, health data have previously been used to assess performance in the form of audit and feedback activities, clinical registries, and ad hoc data analysis [6-8].

While there are good reasons to support the use of practice analytics in these settings, we have collated a number of ethical concerns and medico-legal risks that have limited its application and uptake over the last 3 decades [9]. A scoping review undertaken by the authorship team over 2023 and 2024 in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [10] revealed that the lack of system integration and interoperability and the absence of well-validated, risk-adjusted outcome measures affected perceptions of fairness and justice when managing clinical issues and substandard performance [9]. More specifically, health practitioners were concerned about the risks of using crude measures that fail to capture the dynamic, complex, and heterogeneous nature of clinical practice to discipline individuals and introduce corrective action. The perceived unfairness of using these outcome measures to assess performance was also found to continue to impact trust in practice analytics given the potential consequences of disciplinary action to professional autonomy and reputations [9,11,12]. Beyond these immediate concerns about the use of health data in performance management, there appear to be increasing concerns about public disclosure of outcome measures by health administrators and regulators to the detriment of health practitioners and services [9]. Much of this hesitancy to have health care outcome measures publicly reported appears to relate to the possible misinterpretation and media sensationalization of data disclosed without proper deidentification and contextualization such that the broader clinical narrative remains underrepresented in publicly available information [9]. This reticence likely reflects emerging concerns about access to reflective statements in legal proceedings as punishment of open and honest practitioners [13] despite openness being an established aspect of good medical practice [14]. This treatment of practice reflections was also seen to encourage defensive practices when mistakes are made, thereby undermining professional accountability and compromising patient safety [15].

A review of the relevant literature also revealed that prevention of the escalation of defensive practice was seen to rely on networked forms of governance that meaningfully involved health practitioners in discussions about their performance and accounted for their preferences when deciding upon access to and disclosure of performance information [9]. In this regard, this review affirms the broader literature in that respect for and meaningful engagement of key stakeholders as valued and equal partners drives trust in programs that involve or affect these stakeholders, particularly when co-design and co-production of solutions occur [16]. This effect is particularly pronounced in the context of performance review as health practitioners are not only passive recipients of performance information but also essential to capturing actionable health data at the point of care and analyzing outcome measures using their specialized clinical expertise [17]. Therefore, delivering on the fundamental assurance of practice analytics as an initiative designed to improve patient safety and health care quality [18] depends on overcoming this key dependence and enhancing practitioner willingness to participate.

Our scoping review identified limited knowledge among key stakeholders on how ethical concepts are understood and applied and how legal risks are monitored and minimized in day-to-day practice so as to reduce the impact of perceived ethico-legal barriers on participation. This is reflected in the conflicting viewpoints on "transparency" among health practitioners, administrators, and the public in our scoping review, which we argued have to be reconciled to drive use of performance information [9]. To our knowledge, no study has comprehensively examined how this gap may be closed to guide secondary uses of health data for practice reflection and performance management. We present a research methodology for the development and refinement of an ethico-legal framework to encourage engagement with practice analytics in the workplace.

Aim

This paper briefly outlines the protocol for the development of an ethico-legal framework and a strategy for exploration of the acceptability, comprehensiveness, and potential utility of this developed framework from the perspective of subject matter experts and end users.

Methods

Study Context

This study will involve qualitative semistructured interviews or focus groups conducted in Australia with subject matter experts and end users based in this country. This qualitative study will be reported in accordance with the Standards for Reporting Qualitative Research [19]. Key ethical concepts and medico-legal considerations used in the drafting of the framework have been modeled on feedback from health practitioners working in Australia and health services delivering care in Australia, in addition to the surrounding legal and

XSL•FO
RenderX

regulatory context constructed by Australian governments, professional associations, and regulators.

Development of an Ethico-Legal Framework

An ethico-legal framework for the secondary use of health data for quality improvement and performance management was developed using reviews of relevant literature, laws, governance, and policies governing secondary uses of health data in addition to interviews and focus groups with key stakeholder groups. The full methodologies and results of these reviews, focus groups, and interviews are to be reported elsewhere [9].

Study Design

Reviews were conducted to document the clinical and regulatory rationale for using health data as performance feedback for workforce evaluation and management. These reviews sought to identify key ethico-legal principles used to rationalize the implementation and use of practice analytics for practice reflection and performance management and document the ways in which it is currently operationalized in clinical practice internationally. This included a scoping review of academic manuscripts and conference proceedings available online from the first published materials available on leading medicine and health databases [9], a policy review of relevant guidelines and standards, and a legislative review of applicable laws and regulations. Semistructured qualitative interviews were then conducted with 14 key stakeholders to identify ethico-legal principles that guide implementation of practice analytics, influence professional attitudes toward and engagement with practice analytics in performance evaluation, and/or affect engagement with and impact of the registration requirement positively and negatively. Key stakeholders approached and interviewed included health professionals eligible for registration in Australia; health administrators; and representatives of regulatory bodies, professional associations, and key consumer groups. A key finding from these reviews and interviews was a focus on consistency in the regulation and governance of practice analytics to alleviate worries and a preference for a gradated multistep process for the identification, analysis, and management of clinical variance.

This gradated approach was adapted in the framework as a step-by-step guide consisting of 5 conditional steps with embedded prompts to structure reflections and deliberations on professional performance that use health data as a source of evidence. Descriptions of ethico-legal principles were also extracted and synthesized from these reviews and interviews into a short paragraph no longer than 100 words to enable due consideration of principles during practice reflection and performance management. Many of the prompts explicitly framed ethico-legal principles as questions to encourage discussion on whether these principles are relevant to the reflection and performance review and should be applied. For example, the prompt for trust is “Does the performance/conduct impact public confidence in the quality and safety of healthcare?” and the prompt for honesty is “Were professional responsibilities undertaken honestly, in good faith and with a reasonable degree of care? Was there timely disclosure of unintended consequences to relevant parties?”

The authorship team reviewed this framework, with minor revisions made for comprehension, readability, and clarity, such as removing duplicated prompts, adopting plain language, and including additional features to reduce user error. Multimedia Appendix 1 presents the first draft of the framework for presentation to subject matter experts and end users as outlined in this protocol.

Refining an Ethico-Legal Framework

Two focus groups lasting 1 to 2 hours will be conducted with 5 to 10 purposefully selected participants. Participants will be asked to reflect on six key considerations: (1) framework structure and comprehension, (2) intended audience, (3) comprehensiveness and descriptions of ethical and legal principles, (4) relevance of principles to professional performance and practice more broadly, (5) acceptability and utility of the step-by-step guide, and (6) limitations of the framework. Participants will then be asked to briefly reflect on the potential utility of the framework and when the framework would be most useful in practice through reflections on past experiences or creation of clinical vignettes or case studies.

Participants

Inclusion and Exclusion Criteria

Participants are eligible to take part in the study if they have subject matter expertise relevant to the framework or could be end users of the framework. Potential end users include members of the health workforce involved in performance evaluation, review, and monitoring. This could include registered health professionals, health administrators, and executive staff (eg, directors of clinical governance) working in Australia or supporting the delivery of health services in Australia. Students will be ineligible to take part in the study as students are not obligated to review performance and measure outcomes related to their performance as a condition of their registration or practice [5]. Individuals will also be eligible to take part if they are policymakers or representatives from professional associations and regulatory bodies involved in the enforcement of standards mandating the implementation of practice analytics or outcome measurement in Australia. Subject matter experts include individuals with expertise in health law, clinical ethics, and bioethics due to their specialized knowledge, training, or experience in these fields. Individuals with expertise in health law will be required to have an understanding of Australian statutes, common or judge made, and regulations. Any participant with no connection to health service delivery in Australia or knowledge of the Australian legal and regulatory context will be excluded from eligibility.

Recruitment

Eligible participants and their contact details will be compiled by the authorship team from publicly available information and their personal and professional networks. Each participant will be contacted via email by the lead investigator with a link to the participant information statement and consent form. This email will indicate how this individual was identified as an eligible participant and provide a brief summary of the study and the option to participate in a one-on-one interview for any reason. Participants will also be made aware of their option to

XSL•FO
RenderX

contact a member of the research team to discuss the study before consenting to take part. Participants can self-select for an in-person, online, or hybrid discussion in addition to either an interview or focus group, with in-person interviews and focus groups to be conducted at the University of Sydney. Participants who do agree to take part will be asked to forward information about the study to others who may be interested in participation to assist with snowball recruitment after confirmation with the research team.

Data Collection Procedures

A semistructured guide (Multimedia Appendix 2) was developed consisting of open-ended questions and prompts stratified against 1 or more of the 6 considerations or topics: (1) framework structure and comprehension, (2) intended audience, (3) comprehensiveness and descriptions of ethical and legal principles, (4) relevance of principles to professional performance and practice more broadly, (5) acceptability and utility of the step-by-step guide, and (6) limitations of the framework. Data collection will continue until at least 5 or a maximum of 10 individuals belonging to both stakeholder groups (subject matter experts and end users) have taken part in an interview or focus group. Depending on the rate of recruitment and level of expressed interest in the study, we may keep recruitment open until the minimum number of participants has been reached or cap recruitment at the maximum number of participants.

Focus groups or interviews will be conducted in person and/or online using the Microsoft Teams or Zoom (Zoom Video Communications) videoconferencing platforms depending on participants' preferences. Focus groups or interviews will be recorded using the built-in recording capabilities of the platforms and transcribed verbatim by the facilitator. These recordings and transcripts will be downloaded onto a secure, enterprise-grade network-attached storage hosted by the University of Sydney accessible only to the lead investigator and facilitator. These transcripts will then be deidentified for analysis by the facilitator and members of the authorship team.

Data Analysis

The framework method for qualitative analysis [20] will be adapted to make iterative refinements by linking data obtained in this study to themes and categories obtained from previous data used to draft the framework [21]. This analytical method is favored as it does not prescribe to a theoretical or epistemological approach but, rather, enables the interpretation of study data against predefined topics, as undertaken in this study with the identification of 6 key considerations for feedback on the framework [20]. This methodology also enables us to move along the inductive-deductive continuum to link concepts and themes from this study to those from prior studies while allowing for unexpected responses and temporal and contextually situated feedback from participants in this study [21].

This analysis will commence with immersion in the raw data by the facilitator who will read and reread the transcripts shortly after completion of the interviews or focus groups. This process of familiarization will allow for timely reporting of emerging themes and facilitation challenges to the broader authorship team for review and discussion. These discussions may inform potential revisions to the framework and/or semistructured interview guide to allow the research team to remain agile to evolving study findings and minimize the risk of outdated or repetitive research findings [22]. This preliminary reporting will be followed by open coding, where each interview or focus group transcript will be coded line by line by one author. This open coding will take place in Microsoft Word, with a new comment generated for each line of text. This open coding will then be drawn together based on relationships between line codes and higher-order concepts and outcomes already identified in prior studies or the framework itself, or constant comparison. New concepts and interrelationships will be affirmed by the wider authorship team. A member of the authorship team will be invited to be an independent reviewer in the focus groups to maintain a contemporaneous record of initial impressions, emerging concepts, and theoretical connection. These analytical insights will be combined with feedback from this reviewer to enhance the credibility of the research findings and minimize actual or apprehended bias of this authorship team given our involvement in the development of the draft framework. This reviewer will also be asked to examine the transcripts, codes, and concepts to identify new or miscategorized concepts after completion of data analysis by the research team. These groupings will then be mapped against the predefined considerations [22], namely, the acceptability, utility, and comprehensiveness of the framework as a tool for practice reflection and performance management.

Ethical Considerations

Ethics approval for this study has been received from the University of Sydney Human Research Ethics Committee (2022/643). Each eligible participant will be sent an email invitation to take part in the study. This invitation will contain the participant information statement and participant consent form with a short summary of the study objectives. The focus group or interview will be scheduled with the return of a completed participant consent form. After completion of the focus group or interview, recordings and transcripts will only be accessible to the facilitator, who will then deidentify the information to protect the privacy and confidentiality of participants before analysis.

Results

Enrollment of study participants started in August 2025 and is expected to end in December 2025. An in-depth analysis has commenced in late September 2025 and will likely be completed in early January 2026, as will collation and description of relevant themes for broader dissemination in conference presentations, research seminars, doctoral theses, and journal articles. At manuscript submission, we had conducted 2 focus groups and 6 interviews from October to December 2025. A total of 15 participants were interviewed. Data analysis is underway, with coding and thematic mapping completed for 1 focus group and 2 interviews in November 2025. The remaining data will be analyzed in February 2026 and prepared for publication in March 2026.

XSL•FO
RenderX

Discussion

This research protocol has been designed to enable the development and refinement of an ethico-legal framework to support decision-making related to the secondary use of health data for quality improvement and performance management. The goal of such a framework is to address real and perceived ethico-legal concerns that appear to determine attitudes toward and engagement with the use of practice analytics for practice reflection and performance management [9].

Potential Findings

We expect that this framework will be well received by end users as a useful tool for health practitioners and administrators, but how best to introduce and scale this framework across medical specialties and services, particularly the proposed format and length of the step-by-step guide, is unclear. Possible suggestions may include the development of an electronic tool to reduce the perceived complexity of the framework, which has previously been identified as a critical feature of successful audit and feedback systems and practice change initiatives such as reminders at the point of care [23]. These considerations will likely be critical to widespread uptake of the framework as they will reduce actual or perceived barriers to accessing and applying this tool in day-to-day clinical practice. The findings of the proposed study could have significant implications for quality improvement initiatives and performance management in health care. This study could help focus the conversation away from feasibility and effectiveness testing of each new technology or technique for data collection onto knowledge mobilization and the socialization of performance information using a consistent, harmonized approach.

Comparison With Prior Work

This distinction is increasingly important as barriers to the use of outcome measures and other performance measures appear to depend on data contextualization either through appropriate risk adjustment or full and meaningful engagement of medical practitioners in discussions about their performance [15]. Consequently, a key strength of the proposed study is the fact that the development process involves the systematic identification of ethico-legal concerns regarding practice analytics from a detailed review of the relevant literature and in-depth interviews with relevant subject matter experts and key stakeholders engaged in performance management or subject to it. The collection of in-depth qualitative feedback from key stakeholders has also enabled the research team to tailor research outputs to the local context, including the decision to develop this framework, and consider sociocultural influences and the broader regulatory landscape. The findings of these preceding studies have identified the importance of robust governance and data stewardship for promoting stakeholder engagement with their performance data such that the framework, if well received, will help close an evidence-practice gap.

Furthermore, many of the concerns about privacy and consent appear to be contingent on robust governance over the collection, use, and disclosure of personal and health information, as opposed to focusing on the technical infrastructure used to collect this information alone [9]. The former approach could be taken to inform security assessments of the latter such that technology-specific research may preclude in-depth evaluation of social, cultural, and political influences over medical decision-making and health service planning [24]. Favoring a technology-neutral approach also mimics the prioritization of technology-neutral regulation in the broader legal and regulatory landscape whereby rules are made deliberately flexible to be applicable irrespective of the technology used to reduce compliance costs and burdens to citizens [25], or end users in this case. We acknowledge that this strength may be unique to the Australian context and other jurisdictions where technology-neutral policymaking is prioritized and serves as a limitation on the generalizability of the research outputs. However, this methodological approach does allow for iterative refinements such that outcomes related to earlier studies could be updated periodically as new evidence emerges and repeated with different audiences over time to respond to technological advancements and dynamic contexts.

While many of the studies included in our scoping review identified key ethico-legal considerations surrounding practice analytics [9], principle-focused evaluations are a relatively new program of work in medical education and professional development [26]. This does not mean that guiding principles are not implemented in practice, merely that the nature of their incorporation is highly varied and often dependent on workplace champions to model and disperse a culture of value-led learning and participation [27]. The proposed study seeks to explore a different coapproach to how principles could be actioned in performance evaluations as a systematic exchange of information based on an agreed set of relevant and acceptable ethico-legal values compared to the potential for organic expression in routinized ways of working [28] and whether this systemic approach is capable of driving engagement with performance feedback and practice change. In doing so, the proposed study may be able to help realize the benefits of explicit deliberation on principles when difficult decisions need to be made about professional performance and conduct. This could include procedural improvements, from providing direction and guidance on conducting fair reviews to designing sanctions that are proportionate to the departure from the accepted standard [26], in addition to protecting against arbitrary decision-making through full and meaningful consideration of relevant factors [29].

Limitations

Realizing these hypothesized benefits will depend on effectiveness testing across health care settings (ie, inpatient and outpatient, chronic and acute care, or public and private), research that is beyond the scope of this protocol and an inherent limitation on the impacts of the proposed study. We also note that the use of purposive and snowball sampling may result in the recruitment of highly motivated individuals invested in the successful implementation of the framework. This recruitment and response bias may obscure potential limitations of the framework from the perspective of end users, particularly those who may find principle-led evaluation desirable but not strictly necessary in routine practice.

XSL•FO
RenderX

Conclusions

This study aims to refine the drafted framework with new information about its value and acceptability from implementers of practice analytics. This novel information will inform translation of the resource into day-to-day practice and possibly contribute to improvements in the safety and quality of health services provision.

Acknowledgments

The authors recognize the contributions of Dr Supriya Subramani as a reviewer for this study. No generative artificial intelligence was used in any aspect of manuscript generation. KS declares the use of generative artificial intelligence (under full human supervision; ChatGPT-5.1 model; OpenAI) to develop prompts for the interview guide provided in Multimedia Appendix 2. These prompts were based on research questions developed by the authorship team.

Funding

This research protocol was funded by the Digital Health Cooperative Research Centre Limited. The Digital Health Cooperative Research Centre Limited is funded under the Commonwealth's Cooperative Research Centres program.

Data Availability

The data to be generated or analyzed during the proposed study will not be publicly available but will be available from the corresponding author on reasonable request. Deidentified focus groups may be difficult to analyze without some contextualizing information to help follow the flow of conversation, requiring careful monitoring by the authorship team to mitigate any possible risks to participant privacy and confidentiality.

Authors' Contributions

Conceptualization: KS, AJ, CS, TS, and IK
Data curation: KS
Formal analysis: KS (lead) and AJ (supporting)
Funding acquisition: TS (lead) and AJ (supporting)
Investigation: KS (lead) and AJ (supporting)
Methodology: KS, AJ, CS, TS, and IK
Project administration: KS (lead) and AJ (supporting)
Resources: TS and IK
Supervision: AJ, CS, TS, and IK
Validation: AJ and IK
Visualization: KS, AJ, CS, TS, and IK
Writing—original draft: KS, CS, and TS
Writing—review and editing: KS, AJ, and IK

Conflicts of Interest

None declared.

Multimedia Appendix 1
First draft of the ethico-legal framework.
[DOCX File, 139 KB - resprot_v15i1e82167_app1.docx]

Multimedia Appendix 2
Interview and focus group guide.
[DOCX File, 19 KB - resprot_v15i1e82167_app2.docx]

Checklist 1
SRQR checklist.
[DOCX File, 18 KB - resprot_v15i1e82167_app3.docx]

References

1. Ghaye T, Lillyman S. Learning Journals and Critical Incidents: Reflective Practice for Health Care Professionals, 2nd edition: Quay Books; 1998.

XSL•FO
RenderX

2. Cole M. Learning through reflective practice: a professional approach to effective continuing professional development among healthcare professionals. Res Post Compul Educ 2000 Mar;5(1):23-38. [doi: 10.1080/13596740000200067]
3. Davies H, Powell A, Rushmer R. Healthcare professionals' views on clinician engagement in quality improvement. A literature review. The Health Foundation. 2007. URL: https://www.health.org.uk/sites/default/files/HealthcareProfessionalsViewsClinicianEngagementQualityImprovement.pdf [accessed 2025-05-25]
4. Lewis K. The five 'C's of lifelong learning: CPD, choices, collegiality, challenges and consequences. Br Dent J 2023 Jan;234(1):22-26. [doi: 10.1038/s41415-022-5403-3] [Medline: 36639472]
5. Medical Board of Australia. Professional development plans & types of CPD. AHPRA. 2024. URL: https://www.medicalboard.gov.au/Professional-Performance-Framework/CPD/Professional-Development-Plans.aspx [accessed 2025-08-15]
6. Brown LC, Stanton WC, Paye W. Facing the limits on uses of medical and peer review information: are high technology and confidentiality on a collison course. Whittier L Rev 1997:97-119 [FREE Full text]
7. Boyce MB, Browne JP, Greenhalgh J. The experiences of professionals with using information from patient-reported outcome measures to improve the quality of healthcare: a systematic review of qualitative research. BMJ Qual Saf 2014 Jun;23(6):508-518. [doi: 10.1136/bmjqs-2013-002524] [Medline: 24505110]
8. Kiernan F, Rahman F. Measuring surgical performance: a risky game? Surgeon 2015 Aug;13(4):213-217. [doi: 10.1016/j.surge.2014.12.001] [Medline: 25638721]
9. Shah K, Leow K, Janssen A, Shaw T, Stewart C, Kerridge I. Ethical and legal considerations governing use of health data for quality improvement and performance management: a scoping review of the perspectives of health professionals and administrators. BMJ Open Qual 2025 Apr 17;14(2):e003309. [doi: 10.1136/bmjoq-2025-003309] [Medline: 40246573]
10. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
11. McCabe TJ, Sambrook SA. A discourse analysis of managerialism and trust amongst nursing professionals. Irish J Manage 2019 Dec 30;38(1):38-53. [doi: 10.2478/ijm-2018-0009]
12. Tavares W, Sockalingam S, Valanci S, et al. Performance data advocacy for continuing professional development in health professions. Acad Med 2024 Feb 1;99(2):153-158. [doi: 10.1097/ACM.0000000000005490] [Medline: 37824840]
13. Ahern S, Hopper I, Loh E. Qualified privilege legislation to support clinician quality assurance: balancing professional and public interests. Med J Aust 2019 May;210(8):343-346. [doi: 10.5694/mja2.50124] [Medline: 30924536]
14. Medical Board of Australia. Good medical practice: a code of conduct for doctors in Australia. 2020. URL: https://www.medicalboard.gov.au/Codes-Guidelines-Policies/Code-of-conduct.aspx [accessed 2025-11-24]
15. Whitelock-Wainwright E, Koh JW, Whitelock-Wainwright A, Talic S, Rankin D, Gašević D. An exploration into physician and surgeon data sensemaking: a qualitative systematic review using thematic synthesis. BMC Med Inform Decis Mak 2022 Sep 28;22(1):256. [doi: 10.1186/s12911-022-01997-1] [Medline: 36171583]
16. Grindell C, Coates E, Croot L, O'Cathain A. The use of co-production, co-design and co-creation to mobilise knowledge in the management of health conditions: a systematic review. BMC Health Serv Res 2022 Jul 7;22(1):877. [doi: 10.1186/s12913-022-08079-y] [Medline: 35799251]
17. Alonso V, Santos JV, Pinto M, et al. Health records as the basis of clinical coding: Is the quality adequate? A qualitative study of medical coders' perceptions. Health Inf Manag 2020 Jan;49(1):28-37. [doi: 10.1177/1833358319826351] [Medline: 30744403]
18. Janssen A, Talic S, Gasevic D, Kay J, Shaw T. Exploring the intersection between health professionals' learning and eHealth data: protocol for a comprehensive research program in practice analytics in health care. JMIR Res Protoc 2021 Dec 9;10(12):e27984. [doi: 10.2196/27984] [Medline: 34889768]
19. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Academic Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
20. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13(1):117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
21. Palmer V, Coe A. A best practice guide to qualitative analysis of research to inform healthcare improvement, re-design, implementation and translation: a guidance document. Agency for Clinical Innovation. 2020. URL: https://aci.health.nsw.gov.au/__data/assets/pdf_file/0006/660867/ACI-qualitative-analysis-of-research.pdf [accessed 2025-05-25]
22. Ahmed SK, Mohammed RA, Nashwan AJ, et al. Using thematic analysis in qualitative research. J Med Surg Public Health 2025 Aug;6:100198. [doi: 10.1016/j.glmedi.2025.100198]
23. Pereira VC, Silva SN, Carvalho VK, Zanghelini F, Barreto JO. Strategies for the implementation of clinical practice guidelines in public health: an overview of systematic reviews. Health Res Policy Syst 2022 Jan 24;20(1):13. [doi: 10.1186/s12961-022-00815-4] [Medline: 35073897]
24. Waring J, Allen D, Braithwaite J, Sandall J. Healthcare quality and safety: a review of policy, practice and research. Sociol Health Illn 2016 Feb;38(2):198-215. [doi: 10.1111/1467-9566.12391] [Medline: 26663206]
25. Weber D. Tech neutrality in Australian signature law. J Law Inform Sci 2015(1):101-119 [FREE Full text]

XSL•FO
RenderX

26. Parker K, Sud A. Principles-focused evaluation: a promising practice in the evaluation of continuing professional development. J Contin Educ Health Prof 2023;43(4S):S64-S67. [doi: 10.1097/CEH.0000000000000535] [Medline: 38054494]
27. Bailie J, Laycock AF, Conte KP, et al. Principles guiding ethical research in a collaboration to strengthen Indigenous primary healthcare in Australia: learning from experience. BMJ Glob Health 2021 Jan;6(1):e003852. [doi: 10.1136/bmjgh-2020-003852] [Medline: 33441334]
28. Martin E, Bergeron D, Gaboury I. Champions as catalysts for change: a realist evaluation of their impact in primary care. BMC Health Serv Res 2025 Nov 18;25(1):1472. [doi: 10.1186/s12913-025-13573-0] [Medline: 41254577]
29. Barker M, Nagel A. Legal unreasonableness: life after "Li". AIAL Forum 2015:1 [FREE Full text]

Abbreviations

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
SRQR: Standards for Reporting Qualitative Research

Edited by A Schwartz; submitted 10.Aug.2025; peer-reviewed by E Kuhn, I Aryendu; revised version received 27.Nov.2025; accepted 22.Dec.2025; published 30.Jan.2026.

Please cite as:
Shah K, Janssen A, Stewart C, Shaw T, Kerridge I
Development of an Ethico-Legal Framework for Quality Improvement and Performance Management in Health Care: Protocol for a Qualitative Study
JMIR Res Protoc 2026;15:e82167
URL: https://www.researchprotocols.org/2026/1/e82167
doi:10.2196/82167

© Kavisha Shah, Anna Janssen, Cameron Stewart, Tim Shaw, Ian Kerridge. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Implementation of Digital Device–Assisted and Nurse-Led Case Management to Promote Self-Management in Adults With Noncommunicable Diseases: Protocol for a Single-Arm Intervention Study

Raweewan Pongpoottipatchara[1], MNS; Kanlayawee Anonjarn[2], MNS; Warithorn Prawatwong[1], MNS; Naruemol Singha-Dong[2], MS, PhD

[1]School of Adult and Gerontological Nursing, Institute of Nursing, Suranaree University of Technology, Nakhon Ratchasima, Thailand

[2]School of Community Health Nursing, Institute of Nursing, Suranaree University of Technology, 111 University Avenue, Suranari, Mueang, Nakhon Ratchasima, Thailand

Corresponding Author:
Naruemol Singha-Dong, MS, PhD
School of Community Health Nursing, Institute of Nursing, Suranaree University of Technology, 111 University Avenue, Suranari, Mueang, Nakhon Ratchasima, Thailand

Abstract

Background: Self-management plays a vital role in noncommunicable disease prevention and control. However, it has been challenging for patients and their caregivers to identify how much their lifestyle affects their health and what level of effort they should make to reduce cardiovascular disease (CVD) risks in everyday life. Therefore, knowing their own CVD risk and daily health-related situations will provide relevant information for self-management by those at risk. The need to help individuals understand their relevant information creates an opportunity to investigate whether and how to implement a combined digital and nurse-led self-management intervention in a real-world community setting.

Objective: This study aims to evaluate the effectiveness of a combined approach combining digital device support, including a smartwatch, a mobile app, and a salt meter, with nurse-led case management, on self-management behaviors and clinical outcomes.

Methods: This study uses a combination of a nurse-led self-management with a digital and mobile health innovative approach, including tailored small group face-to-face education sessions, a smartwatch, a smartphone health app, and a salt meter, to increase the self-management behaviors to reduce vascular risk through designing and testing an integrated community-based strategy targeted at adults and older adults at risk of CVD in Thailand. The study uses a single-arm pretest-posttest design to assess the intervention's effects. The intervention will consist of the following components: (1) an interactive face-to-face education session; (2) a real-time knowing your numbers strategy using a smartwatch, a smartphone health app, and a salt meter; (3) a mindfulness-based stress management strategy using Somporn Kantaradusdi-Triamchaisri technique meditation healing exercise; and (4) a self-management diary. Quantitative data will be collected using a smartwatch, a salt meter, a food diary, and questionnaires at baseline and at the end of week 6. Clinical outcomes will be assessed at baseline, primary end point (wk 12), and secondary end point (wk 24).

Results: This study, funded in January 2025, will involve 45 patients. We received ethical approval on May 31, 2024, and began recruitment for participation in May 2025. Researchers will collect, analyze, and synthesize to evaluate the study procedure. We expect to complete data collection by December 2025, with the first results submitted for publication in March 2026.

Conclusions: The implementation of a combined digital device and nurse-led case management may identify the use of digital health to support self-management and improve vascular health. The findings of this study will provide insights for a large-scale randomized controlled trial and for ongoing improvements in the noncommunicable disease care system.

Trial Registration: Thai Clinical Trials Registry TCTR20250701003; https://www.thaiclinicaltrials.org/show/TCTR20250701003

International Registered Report Identifier (IRRID): DERR1-10.2196/80083

(JMIR Res Protoc 2026;15:e80083) doi:10.2196/80083

KEYWORDS

behavioral intervention; CVD prevention; cardiovascular disease; digital health; eHealth interventions; mobile health; NCD; noncommunicable disease; nurse-led intervention; self-management; vascular health

XSL•FO
RenderX

Introduction

Background

Cardiovascular diseases (CVDs) are a leading cause of noncommunicable diseases (NCDs) globally. CVDs consist of ischemic heart disease, rheumatic heart disease, hypertensive heart disease, stroke, and other cardiovascular conditions [1]. CVD caused 20.5 million deaths in 2021, as reported by the World Heart Federation. The death rates from CVDs remain steady in low- and middle-income countries, where over 80% of global CVD deaths occur [2,3]. In 2021, CVD was the leading cause of death and disability-adjusted life years (DALYs) globally, accounting for 26.8% of total deaths and 14.9% of total DALYs 2,3[4]. The American Heart Association reports a significant rise in CVDs worldwide, with heart attacks and strokes being 4 out of 5 causes of CVD deaths [5]. From 1990 to 2021, ischemic heart disease increased by 68%, stroke by 44%, hypertensive heart disease by 87%, and cardiomyopathy and myocarditis by 48% [5]. In Thailand, stroke has significantly increased every year, with an average of 8350 cases per year from 2563 to 2567 and increased to 21,012 cases or 15.28% from 2022 to 2023 and 20.77% death rate in 2024 [6]. Additionally, CVD prevalence in rheumatoid arthritis patients was higher in women with diabetes mellitus, hypertension, and dyslipidemia in Thailand [6].

The risk of CVD is associated with several common risk factors such as age, diabetes, high blood pressure, high cholesterol, a family history of vascular or heart disease, obesity, and smoking, such as an improper perception of health, a diminished sense of control, limited availability of health information, inadequate self-management skills, delayed and inefficient communication between health care providers and patients, and insufficient instrumental support [5,7,8]. Managing these risk factors through self-management intervention can help prevent or minimize the onset of vascular complications [6,9]. Therefore, encouraging self-management is crucial for preventing and controlling CVD. Self-care is a naturalistic decision-making process that addresses both the prevention and management of chronic illness, with core elements including self-care maintenance, self-care monitoring, and self-care management [10]. Self-management refers to a patient's ability to cope with all aspects of a chronic illness, including symptoms, treatment, physical and social consequences, and lifestyle changes [6]. Several literature reviews have found that self-management strategies incorporating lifestyle modifications and behavioral interventions have been shown to improve vascular health [10-13].

Successful management of CVD risk involves interactive group education on the disease and its management, combined with regular monitoring and provision of feedback [9,10]. It also includes providing equipment and lifestyle advice supported by dietary and physical activity guidance, weight reduction, stress management, and medication adherence [11,14]. In NCD self-management, nurses play a crucial role in supporting patients with uncontrolled chronic conditions by providing health education, counseling, coaching, and support [13]. However, it has been challenging for patients and their caregivers to identify how much their lifestyle affects their health and what level of effort and intensity they should make to lower CVD risks in everyday life. Therefore, knowing their own CVD risks and daily health-related situations will provide relevant information and motivation for self-management among those at risk. Digital health interventions (DHIs) have become recognized as an effective complementary strategy [7,9,15-17] to enhance self-management support, as they overcome barriers and time constraints that traditional self-management approaches often encounter.

Interest in using smartwatches and digital technologies to develop DHIs that improve health and well-being has risen in recent years. Several reviews have reported that the effects of DHIs, including telemedicine, mobile health, and decision support tools, have enabled virtual access to clinical care, fostered positive health behaviors, and shown improvements in health outcomes and health care delivery for CVD, thereby improving CVD health [7,15,17-19].

Even though DHIs have proven to be effective in enhancing self-management behaviors and promoting clinical health outcomes, DHIs have been challenging in resource-limited settings [20]. The combination of DHIs and health care providers' interventions can potentially help patients with NCD improve adherence to self-management when compared with stand-alone DHIs [7]. A systematic review found that smartwatch interventions in the everyday lives of patients aged ≥50 years help support lifestyle changes, improve treatment adherence, and improve clinical outcomes for CVD [19,21]. Recent smartwatch technology, which connects to a mobile app, has enabled lifestyle, goal tracking, and interactive communication interventions, thereby fostering virtual access to clinical care and promoting positive health behaviors that lead to achieving clinical outcomes [17,21,22]. The evaluation of the implementation strategy in smartwatch and digital health research studies, especially in Thailand, is limited [21]. This limitation presents an opportunity to investigate whether and how a combined DHI and nurse-led self-management intervention promotes vascular health in a real-world, resource-limited setting with diverse health determinants in Thailand.

Poor eating habits are a significant risk factor for NCD, especially CVDs. The Dietary Approaches to Stop Hypertension (DASH) diet emphasizes vegetables, whole grains, lean protein, and low-fat dairy, while limiting sodium, unhealthy fats, and sugary beverages. It is particularly beneficial for cardiovascular health, especially for those with hypertension [23]. Research shows it effectively lowers blood pressure, with reductions of −3.89 mm Hg systolic and −2.19 mmHg diastolic compared to control diets [24]. Additionally, the DASH diet has a positive impact on other cardiovascular risk factors, including blood pressure, triglycerides, cholesterol levels, body weight, and BMI [23,25,26]. The duration of DASH diet interventions varies. A study found that the DASH diet reduced 10-year atherosclerotic CVD risk scores by about 10% in just 8 weeks [27]. The DASH-sodium trial noted significant reductions in cardiac injury and inflammation markers over 12 weeks [28]. Additionally, a systematic and meta-analysis review suggested that a duration of 12 weeks or more is particularly effective for

XSL•FO
RenderX

achieving reductions in body weight, BMI, total cholesterol, and low-density lipoprotein cholesterol [29].

Overall, these findings suggest that the DASH diet can offer significant cardiovascular benefits within 8 to 12 weeks, with longer adherence potentially enhancing these effects. However, this study uses a 12-week intervention to assess its outcomes. Participants in this study will use digital devices only for the first 6 weeks, then engage in self-management activities for another 6 weeks. These findings highlight the significant benefits of the DASH diet in enhancing cardiovascular health by addressing multiple risk factors and biomarkers associated with hypertension and CVD. Overall, the DASH diet stands as a well-supported dietary intervention for improving cardiovascular health in individuals with hypertension, with evidence indicating its broader applicability across various populations and health conditions.

Excessive salt intake is a significant risk factor for hypertension, a leading cause of NCDs in Thailand [30,31]. Studies suggest that each 1 g increase in sodium intake above the reference level would result in a 2.11 mmHg increment in systolic blood pressure (BP) and a 0.78 mm Hg increment in diastolic BP [31]. Salt meter, a device used to detect sodium content in daily food, is crucial in Thailand for testing and managing NCD, especially hypertension and CVD [31]. Since a salt meter can tell individuals about their sodium intake, it helps them manage their salt consumption habits. A randomized controlled study found that using a salt meter for self-monitoring of salt intake, in conjunction with dietary education, was superior to education alone in patients with hypertension, resulting in better BP control [31].

Excessive salt intake significantly increased the risk of hypertension, stroke, and total CVD [32], a leading cause of NCDs in Thailand [23,24]. A salt meter is crucial for testing and managing NCDs in Thailand, particularly hypertension and CVD [24].

As tailored self-management services and digital health devices are not equally accessible, we need to investigate if, when used with nurse-led self-management, a generic smartwatch and a salt meter can be used to supply individuals with mild to moderate CVD risk with the estimated reference information to guide daily self-management behaviors and whether the use of digital devices for a short duration can offer a carry-on estimation to guide self-management behaviors to lower CVD risk.

This study will address the gap between face-to-face and DHIs for improving self-management behaviors in urban and suburban settings in Thailand by empowering individuals with NCDs to take proactive steps to manage and engage in self-care in their daily lives. This study aims to evaluate the effectiveness of tailored, context-specific education and mobile health engagement, including the use of smartwatches, smartphones, and salt meters, in promoting self-management behaviors. Tailored, case-specific, and context-specific education, digital engagement, real-time information supply, and free primary care access are all critical for improving perceptions of risk, increasing literacy in, and motivating improvements in CVD-related health behaviors. General and reasonable smartwatches may help people get a snapshot of their daily health information, which they will use as a reference to adjust their diet, exercise, stress management, and sleep habits to meet their anticipated everyday milestones. The findings may offer practical guidance for nurses and health care professionals managing CVD in Thailand and contribute valuable insights for shaping future health care policies.

Objectives

This study aims to investigate whether a combined approach involving digital device support, including a smartwatch, a mobile app, and a salt meter, coupled with nurse-led case management, can enhance self-management behaviors and improve clinical outcomes.

Hypotheses

We have three primary hypotheses. First, we hypothesize that patients with chronic diseases who are at mild to moderate cardiovascular risk will report overall acceptable satisfaction with the mobile digital interventions. Second, we hypothesize that participants will use digital-based information as a reference to adjust their self-management behaviors in response to the nurse's tailored interventions over 6 weeks. Third, we hypothesize that participants will exhibit improved self-management behaviors at weeks 6, 12, and 24, better clinical outcomes, and reduced CVD risks at both the end of week 12 (primary endpoint) and 3 months posttreatment (week 24; secondary end point).

Methods

Study Design

This study uses a single-arm pretest-posttest design to assess the effects of the intervention (Figure 1). The study was perofrmed per the SPIRIT checklist (Checklist 1)[33].

XSL•FO
RenderX

Figure 1. Study design flowchart. NCD: noncommunicable disease.

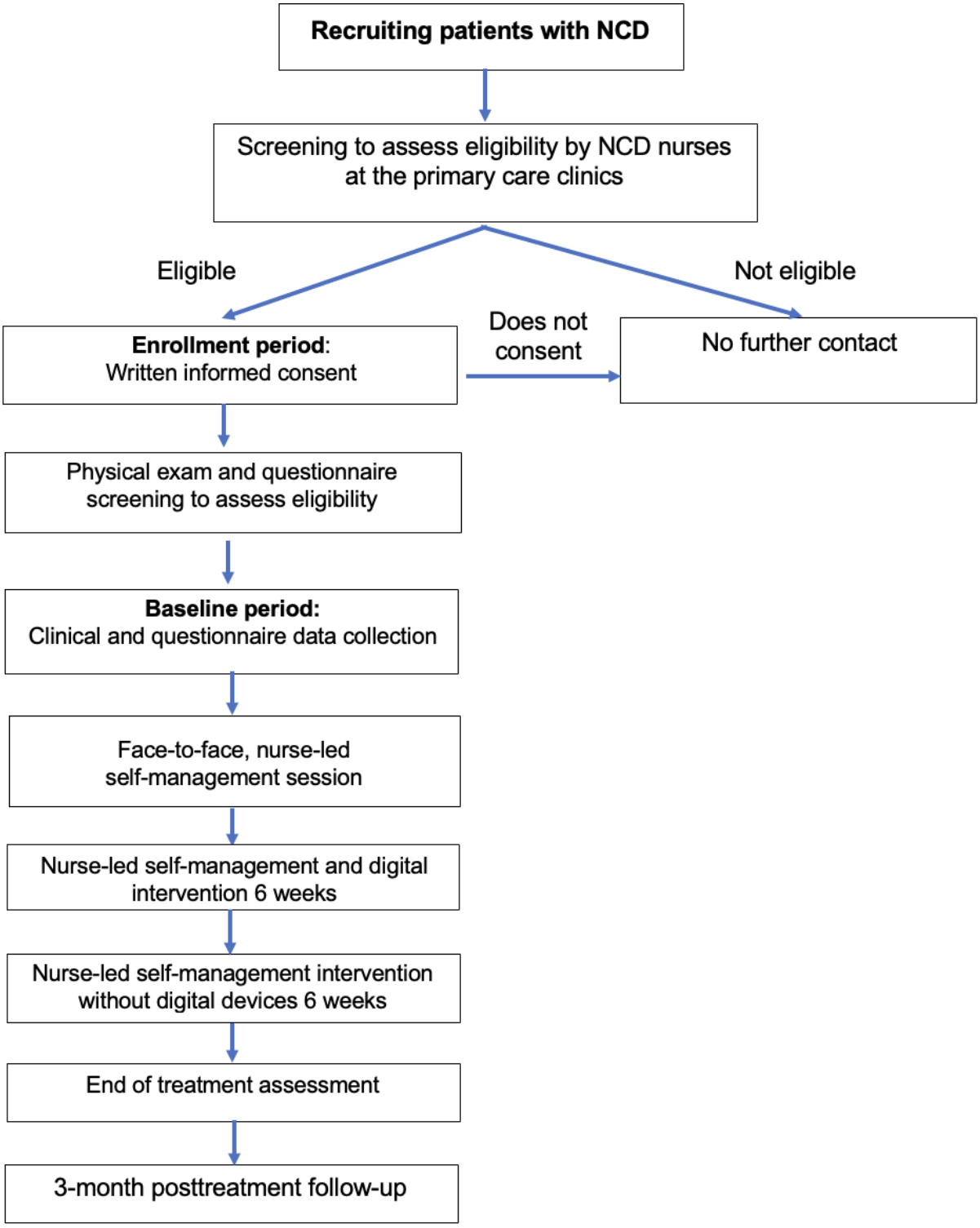

Study Components

Overview

To increase self-management behaviors and reduce vascular risk among adults and older adults at risk of CVD, this study provides interactive group education on vascular health and its management, accompanied by condition monitoring and feedback (Figure 2).

The intervention will consist of the following components: (1) a nurse-led tailored self-management session with an interactive face-to-face education, (2) a real-time knowing your numbers strategy using a smartwatch, a smartphone health app, and a salt meter, (3) a mindfulness-based stress management strategy using the Somporn Kantaradusdi-Triamchaisri (SKT) meditation healing exercise, and (4) a self-management manual. We will ask all participants to engage in self-management behaviors for

XSL•FO
RenderX

6 weeks and to continue their self-management practice without digital devices for an additional 6 weeks.

Figure 2. Timeline of enrollment, assessments, intervention, and follow-up. ABI: ankle-brachial index; CVD: cardiovascular disease; eGFR: estimated glomerular filtration rate; HbA_{1c}: hemoglobin A_{1c}.

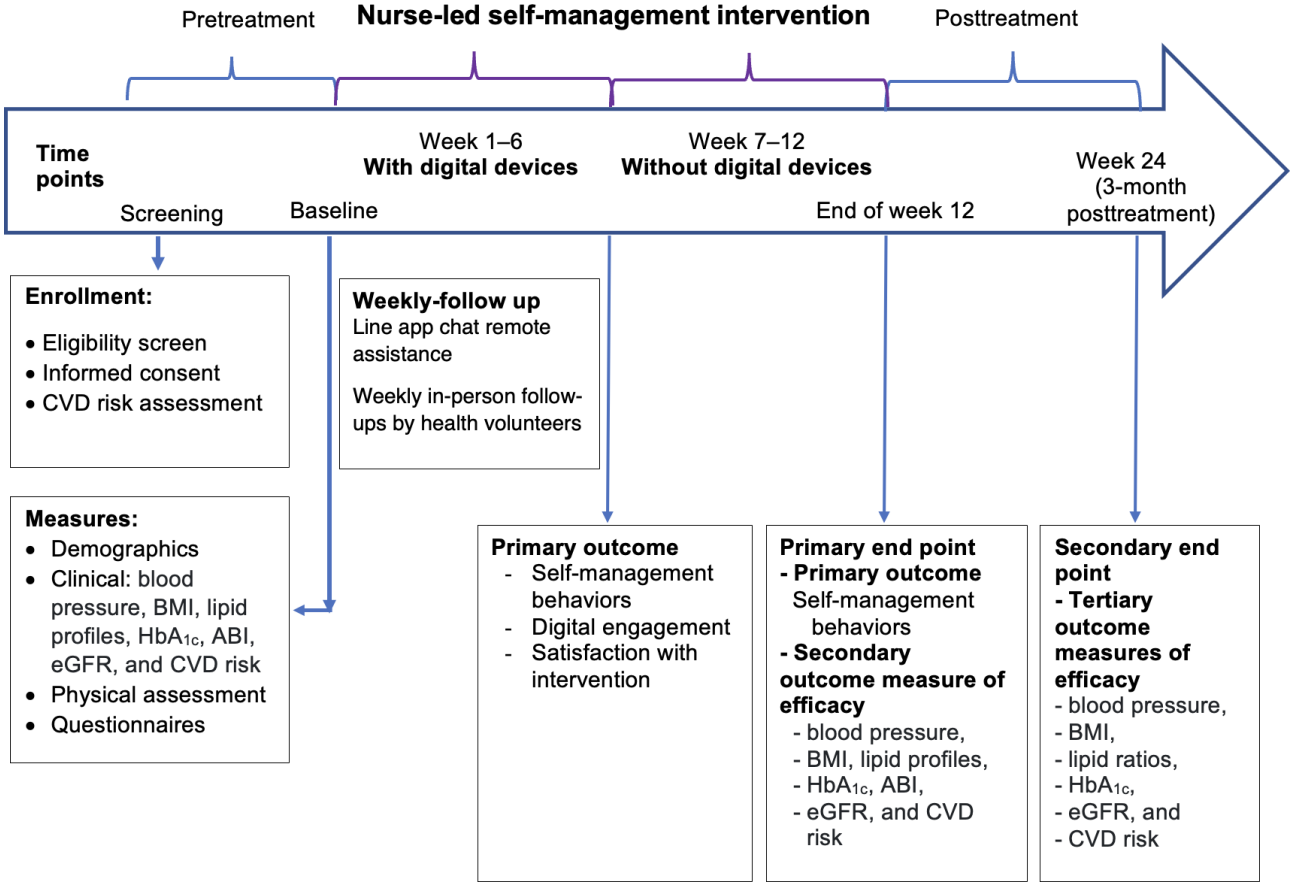

A Nurse-Led Tailored Self-Management Session

This management session will involve (1) a health assessment to identify the needs of participants, (2) an interactive face-to-face education and self-management consultation responsive to individual health needs, (3) a mindfulness-based SKT meditation healing exercise training, and (4) a digital device training. A family doctor will conduct a physical examination, and a nurse will conduct a structured interview. A self-management session focuses on lifestyle management based on the DASH diet, tailored to address the individual needs of individuals with diabetes and chronic renal disease. It includes physical exercise, sleep improvement, and training in self-management skills.

A Real-Time Knowing Your Numbers Strategy

This study component comprises the use of a smartwatch, a smartphone health app, and a salt meter. We will ask participants to use these digital devices for 6 weeks.

A Smartwatch: Huawei Watch Fit 3

This study will use the commercialized Huawei Watch Fit 3 and its associated HUAWEI Health app (Huawei Technologies Co, Ltd). Study participants will use this smartwatch and its smartphone app to collect daily reference data on physical activity, sleep quality, heart rate, and stress levels, which will guide their daily self-management efforts and actions in line with their target setting with the nurses. The device features a built-in interactive function that allows each participant to set goals for exercise, physical activity, and sleep. The health app helps participants track their progress and compare milestones on a weekly and monthly basis. For instance, if a participant misses a walking target, they can take additional steps in their daily routine. Participants can continue using the health app to track their physical activity after using this wearable device. Similarly, participants can use sleep information from a prior night to improve their sleep hygiene.

Portable Salt Measurement Device: CHEM Meter or Salt Meter

CHEM Meter (Quality Assembly Co, Ltd; Figure 3) is a portable salt measurement device using a parts per million/total dissolved solids meter with total dissolved solids expressed in parts per million by measuring the electrical conductivity of sodium chloride (%NaCl) in grams per 100 milliliters (g/100 ml) in food and converting into parts per million unit [34]. The devices use the alternating electrical current, which has a high frequency, to accurately determine the concentration of sodium chloride and display the result in emotional facial graphics; smile, poker face, and frown for sodium chloride content between <0.7%, 0.7% - 0.9%, and >0.9%, respectively [31] (Figure 4). This design is easy to understand for all users. The researchers require all participants to use a salt meter to test sodium levels in their food before meals and adjust their food consumption accordingly

XSL•FO
RenderX

to limit daily sodium intake based on their individual health needs.

XSL•FO
RenderX

Figure 3. Chem meter m device used in this study (Quality Assembly Co, Ltd).

XSL•FO
RenderX

Figure 4. Examples of the CHEM meter facial display intended to show sodium chloride (NaCl) concentration in the interested food; smile for the NaCl <0.7%, poker face for NaCl of 0.7% - 0.9%, and frown for NaCl >0.9%.[35]

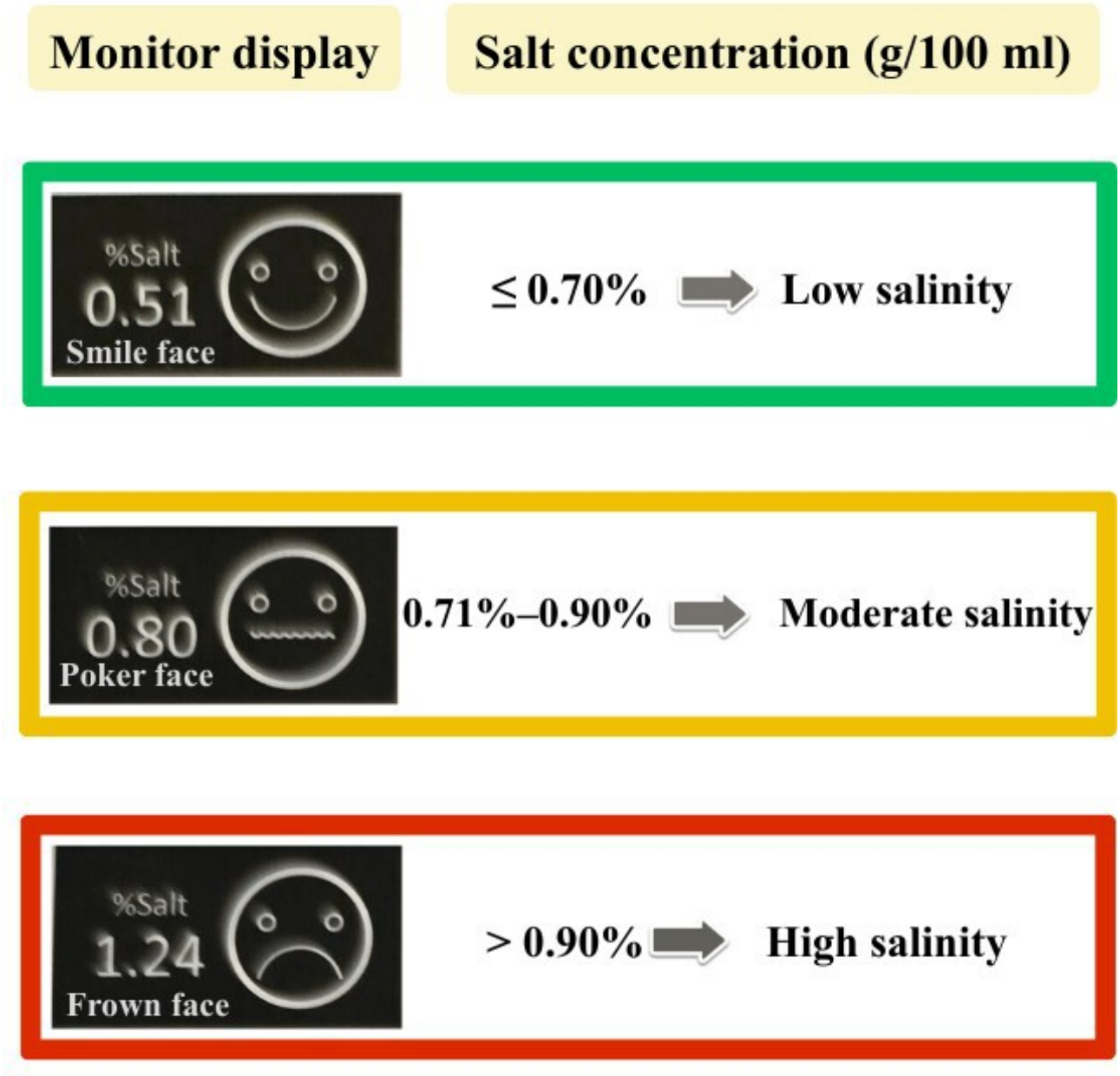

The SKT Meditation Healing Exercise

The SKT is a meditation-based intervention aimed at improving vascular health, particularly in individuals with hypertension [36]. This technique involves a standing deep breathing meditation exercise, which has significantly reduced both systolic and diastolic BP in patients with uncontrolled hypertension and diabetes. The SKT technique is effective when practiced for 20 minutes, twice daily, for 8 weeks. This approach provides a nonpharmacological option for managing BP, a critical factor in maintaining vascular health. A study involving 43 patients with uncontrollable hypertension showed that those who practiced the SKT technique experienced a significant reduction in both systolic and diastolic BP after 8 weeks, compared to a control group that only received medication [37].

A Vascular Care Self-Management Manual

The self-management manual is a color-printed A5-sized manual. The researchers developed this manual based on the study's self-management components to assist the study participants in maintaining the desired management. The manual includes easy-to-use information on the DASH diet, a compatible list of local foods, a step-by-step pictorial SKT meditation healing exercise, information on physical activity and exercise, and an approach to improve sleep quality. The researchers developed this manual based on a literature review. A panel of 5 experts, comprising 2 family medicine doctors, 1 NCD nurse case manager, and 2 nutritionists, reviewed the manual. The researcher revised the manual based on expert suggestions and comments before production. Participants will use digital-based information as a reference to adjust their self-management behaviors in response to the tailored interventions offered by the nurses.

XSL•FO
RenderX

Participants will receive a smartwatch, a salt meter, and printed manuals to support guidance on dietary and physical activity, stress management, and sleep management. All enrolled participants will attend a face-to-face session. The digital components last for 6 weeks. After 6 weeks, we will ask all participants to continue their dietary, physical activity, stress management, and sleep management behaviors based on the information they gained from the digital components during the previous 6 weeks. All participants will record their daily diet and exercise using the printed diary.

Recruitment

Researchers are recruiting participants for this study from primary care clinics in Nakhon Ratchasima, Thailand. The nurses at the primary care clinics will screen potential participants to determine if they meet the inclusion criteria. If the potential participants are deemed eligible and willing to participate in the study, the nurses will refer them to the researcher. Next, the researcher will approach potential participants who meet the inclusion and exclusion criteria to inform them about the study details, obtain written informed consent, and assess their digital health literacy. Those who pass all inclusion and exclusion criteria are eligible for recruitment into the study.

Inclusion and Exclusion Criteria

The inclusion and exclusion criteria are present in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Aged ≥20 years with noncommunicable disease, both man and woman.
- Determined as having a mild to moderate risk for vascular complications, including having at least 1 condition such as uncontrolled diabetes, high blood pressure, hyperlipidemia, a family history of vascular or heart disease, or smoking, or being aged over 50 years. A nurse will assess these risks using the Thai cardiovascular risk score, while a family doctor will confirm eligibility by reviewing the medical history of patients.
- Have access to a smartphone.
- Able to use digital devices and communicate in Thai.
- Having the intention to change self-management behaviors according to verbal screening.
- Able to provide informed consent.

Exclusion criteria

- Acute medical condition requiring hospitalization.
- Cognitive impairment or inability to participate.
- Already enrolled in a similar program.
- Recently diagnosed with strokes, hypertensive heart disease, ischemic heart disease, heart failure, acute myocardial infarction, and psychotic disorder.
- Having inadequate digital health literacy as measured by the eHealth Literacy Scale. The cut-off points of <26 indicate inadequate digital health literacy [38].

Setting and Participant Demographics

Participants include individuals at risk of CVDs, such as uncontrolled hypertension, hyperlipidemia, and diabetes, who live in the urban and suburban communities of Nakhon Ratchasima province of Thailand. Residents of these communities have diverse educational backgrounds and socioeconomic statuses. In this setting, the younger generation typically holds an associate's degree; the middle group is educated primarily at a lower level, often with labor-intensive work, whereas the older generation has slightly better education, with an average ranging from an associate's degree to a bachelor's degree.

Sample Size Calculation

The primary outcome of this study is the change in the self-management behavior score from baseline to 12 weeks postintervention. The sample size was calculated for a paired-samples within-group design using G*Power. Based on previous studies evaluating digital and nurse-led interventions to improve self-management in individuals with chronic conditions [27], a moderate effect size (Cohen d=0.5) was assumed. With a significant level of 0.05, 80% power (1-β=.80), and a Cohen d=0.5 effect size, the required sample size to detect a change is 34. Allowing for a 30% attrition rate, the sample size was adjusted to 45 participants.

Data Collection Procedures

Researchers will collect information on CVD risk, sleep quality, stress levels, and food consumption, as well as self-management behaviors, using physical examinations, questionnaires, and semi-structured interviews at baseline and the conclusion of the study.

1. The research team will collect CVD risk information using health reviews and questionnaires.
2. Digital data will be collected using the Huawei Watch Fit 3, which is associated with the HUAWEI Health app. These data do not include participant identification information. The researchers will access digital health data, including physical activity, sleep patterns, heart rate, and stress levels, through the HUAWEI cloud system. Data encryption will

XSL•FO
RenderX

be used to protect data from unauthorized access, theft, or modification by scrambling it into an unreadable format that can only be decrypted with a special key.

3. Sodium consumption: participants will use a CHEM meter to measure the salt concentration in their food before meals and record their sodium intake in their food diary.
4. Participants will record their mindfulness activity, diet, and exercise data in a self-management diary.

Participants will be monitored weekly for 6 weeks through a health volunteer visit and a web-based survey. For the primary outcome, participants will answer questions about their self-management behaviors during (1) an initial medical screening; (2) at the end of weeks 1, 2, 3, 4, 5, and 6; (3) immediately after digital treatment at the end of week 6; and (4) week 12. For clinical outcomes, the primary end point is assessed at the end of week 12 (short-term outcomes), and the secondary end point is evaluated at week 24 (delayed outcomes) after the intervention is initiated. Clinical outcomes will include BP, BMI, lipid profiles, lipid ratio, HbA_{1c}, ankle-brachial index, estimated glomerular filtration rate, and CVD risk. Acceptability will be assessed using a semistructured in-depth interview by researchers.

Method of Measurement

The research instruments consist of (1) screening measurements, (2) experimental measurements, and (3) data measurements.

Screening Measurements

The eHealth Literacy Scale

Electronic health literacy is "*the ability to seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem*" [39]. Norman and Skinner [39] developed the eHealth Literacy Scale (eHEALS) measurement, which researchers will use to screen the digital health literacy of qualified participants. The eHEALS consists of 8 items rated on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Possible scores range from 8 to 40, with a cut-off point of 26 or higher indicating adequate digital health literacy [38]. The Department of Health, Ministry of Public Health, subsequently translated the eHEALS into Thai. The Thai eHEALS has a test-retest reliability of 0.85 [40].

To ensure participants can safely and effectively engage with the digital intervention, participants will be required to score 26 or higher on the eHEALS to affirm high intent to use digital tools during the screening process. This threshold is applied pragmatically to maximize engagement and minimize attrition during the intervention phase, which is crucial for assessing the intervention's initial efficacy signal.

The Thai Cardiovascular Risk Score

The Thai Cardiovascular Risk score was developed based on cohorts collected from the Electricity Generating Authority of Thailand study and validated using the 2nd and 3rd Electricity Generating Authority of Thailand cohorts, respectively [41]. The Thai health system has been practically using the Thai Cardiovascular Risk score as a reliable tool to forecast the risk of having a cardiovascular event in the future for Thais nationwide. The developer classified the risk as low risk (0%-10%), moderate risk (10%-20%), high risk (20%-30%), very high risk (30%-40%), and extremely high risk (40% or more).

Experimental Measurement

The study will implement a self-management intervention comprising digital health components and a nurse-led face-to-face component among Thai participants. This study will use the Huawei Watch Fit 3 to collect digital data, including physical activity, sleep, heart rate, and stress level, and a CHEM (salt) meter to measure salt concentration in food.

Data Measurements

The Demographic Data Form

Researchers developed this form to collect demographic data. The form includes a structured set of questions asking about gender, age, education, occupation, income, income adequacy, medical history, medication use, and underlying diseases.

The Dietary Behavior Screening (Sweet, Fatty, Salty) Questionnaire

The researchers will use the Dietary Behavior Screening (Sweet, Fatty, Salty) Questionnaire (DBSQ) to examine the dietary behavior of Thais with NCD. The Raipoong Academy, in collaboration with the Royal College of Physicians of Thailand, developed the DBSQ for public use as part of the Prevention & Health Promotion campaign. This measurement consists of 15 items: 5 items for sweet, 5 for fatty, and 5 for salty diet [42]. This measurement uses a 3-category Likert scale, from 1 - 3 (1="not at all" and 3="always" for negative behaviors; and 3="not at all" and 1="always" for positive behaviors). The response choices for all items are standardized on a scale from 5 to 15, with scores of 6 to 9, 10 to 13, and 14 to 15 indicating mild-to-moderate, high, and critical dietary behavior risk, respectively. This study will examine the reliability of the DBSQ using 30 Thais with NCDs and obtain the Cronbach α coefficient before collecting the data.

Sleep Quality: The Thai Pittsburgh Sleep Quality Index

The Pittsburgh Sleep Quality Index (PSQI) is a 19-item self-report questionnaire that evaluates various aspects of sleep over the past month. The questionnaire assesses perceived sleep quality, sleep habits, such as habitual bedtimes and wake times, and sleep disturbances, including difficulty falling asleep, disordered breathing, nightmares, and pain [43]. A habitual bed partner is also required to rate 5 challenging questions about sleep, such as loud snoring, pauses in breathing, leg twitching, and disorientation during sleep. A scoring procedure generates scores for 7 distinct components: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction.

Component scores range from 0 (no difficulty) to 3 (severe difficulty), and their sum produces a global score ranging from 0 to 21, with higher values indicating poorer sleep quality. A PSQI score greater than 5 is considered indicative of poor sleep quality, based on its high sensitivity and specificity in

XSL•FO
RenderX

identifying patients with sleep complaints [43]. Researchers developed the Thai-PSQI using a standardized method to measure major sleep disorders in Thai patients. The Thai-PSQI has a Cronbach α of 0.837 and test-retest reliability (intraclass correlation coefficient=0.89) [44]

The CVD Self-Management Behavior Questionnaire

The primary investigator developed the CVD Self-Management Behavior Questionnaire to examine self-management behaviors among Thais at risk of CVD. This measurement comprises 17 items: 4 related to exercise behaviors, 6 to dietary habits, 5 to treatment compliance (rational medication use), and 2 to smoking behaviors. This measurement uses a 6-category Likert scale, ranging from 0 to 5 (0="not at all" and 5="always"). The response options for all items are standardized, with higher scores indicating better CVD risk-prevention behavior. The researcher will test the CVD Self-Management Behavior Questionnaire for reliability with 30 Thais at risk of CVD and obtain a Cronbach α coefficient before collecting the data.

The Thai Stress Questionnaire (ST-5) or Srithanya Stress Test–5 Items

Srithanya Psychiatric Hospital developed this self-administered brief questionnaire to assess perceived stress levels [45]. The Stress Questionnaire-5 (ST-5) comprises items related to sleep problems, decreased concentration, irritability, boredom, and feelings of isolation. It measures subjective feelings over the past 2 to 4 weeks, using a 4-point Likert scale (0 - 3) for each item. The total score is then categorized to indicate the level of stress experienced by the individual: 0 to 4 indicates mild stress, 5 to 7 indicates moderate stress, 8 to 9 indicates severe stress, and 10 to 15 indicates extremely severe stress. The ST-5 has a Cronbach of 0.85 [45].

An Automatic BP Measurement

A health volunteer will use the Omron HEM-9200T, along with the standard method for measuring BP, as outlined in the 2024 Thai Hypertension Society guidelines [46]. The researchers will require calibration of the automatic BP measurement using the clinical instrument testing center by the standard calibrator lab at the start of the intervention and during the study to ensure accuracy.

Statistical Analysis

The analysis will include both the complete case and intention-to-treat analyses. The researcher will be using SPSS Statistics (IBM Corp) and Jamovi (GNU Affero General Public License and GNU General Public License) open-source statistical software to analyze quantitative data using the following statistical methods:

1. Descriptive statistics, including frequency, percentage, mean, and SD, were used to characterize the demographic data and sample characteristics.
2. Repeated measures ANOVA will analyze changes in self-management behaviors across baseline, week 6, week 12, and week 24. Additionally, clinical outcomes measured at baseline, week 12, and week 24, including those that show significant differences at one or more time points, will be evaluated [47]. This method is based on data that meet assumptions, including normality of the dependent variable, no outliers, sphericity (equal variances across time points), and a balanced number of observations.
3. In case data are missing or observations are unbalanced at each time point, linear mixed effects models will be used to evaluate changes in outcomes over time [48]. Researchers will set time as a fixed effect and individuals as random effects.
4. Post hoc tests will be used to accurately determine the specific differences between time points, using a Bonferroni correction adjustment [47].

The hypothesis will be tested in a hierarchical approach. The primary outcome of self-management behaviors will be tested first at the α=.05 level. If significant, the first key secondary outcome, blood pressure, will be tested, and so on. For the set of secondary outcomes contributing to the main effect claim, the values will be adjusted using the Holm-Bonferroni method to control the family-wise error rate at α=.05.

Ethical Considerations

The study procedures follow the institutional and national ethical standards of the responsible committee on human experimentation and the World Medical Association Declaration of Helsinki. Furthermore, all researchers hold current Good Clinical Practice training certifications. The researchers will use various strategies, including a full-board review and ethical approval from the Human Research Ethics Committee of Suranaree University of Technology (approval 51/2566, dated May 31, 2024; renewed in June 2025). Second, all participants will be informed about the study's objectives, procedures, risks, and benefits. Participation is voluntary, and participants can withdraw at any time. Third, all participants will be required to complete and sign consent forms before participating. The researchers will obtain written informed consent before screening for eligibility. Fourth, participants have the freedom to opt out at any time, before, during, or after participation, without any consequences from the health care system. Fifth, the researchers will maintain the confidentiality of all collected data in accordance with applicable data protection laws [49]. For instance, the researchers will not disclose patients' names, initials, or hospital numbers, and they will guarantee participant anonymity. Each participant will receive travel compensation for every on-site data collection visit.

Data extraction will be performed via a secure, audited call. Participants will be given a unique study ID and code for anonymization to be used throughout the project, including for the smartwatch setup. The pseudonymous dataset will be transferred via Transport Layer Security 1.2 encryption to a dedicated server environment hosted on Huawei. Data encryption will be used both in transit and at rest, with access controls and regular security audits, to protect the data.

Research data will be stored on password-protected computers and in electronic databases. Informed consent forms will be stored separately from research data. Research data will be kept for 5 years after the study ends. Each study participant will receive 100 Thai Baht (US $3.22) for traveling compensation for each clinical visit.

XSL•FO
RenderX

Results

This study received funding in January 2025. A total of 45 patients will participate in this study. The study received ethical approval on May 31, 2024. The ethical committee subsequently granted renewal following the start of participant recruitment in May 2025. The collected data will be analyzed and synthesized to evaluate the implementation of the study procedure. The results will inform the development of a 12-week self-management program for diet, sleep, and physical activity improvement, coordinated with a patient's case management. This program will include both behavior change and clinical outcomes, along with meaningful feedback on participant progress. Study findings will inform the design and ongoing improvement of the NCD care system and will be published in peer-reviewed journals. The study will also examine the implications of using a commercial smartwatch to guide self-management in resource-limited settings. We anticipate completing data collection by December 2025 and submitting the first results for publication in March 2026.

Discussion

Expected Findings

This protocol describes a single-arm, pre- and postintervention design to evaluate the feasibility, acceptability, and preliminary effectiveness of a nurse-led digital support self-management intervention for adults with hypertension, dyslipidemia, and overweight in community settings. NCDs remain burdensome nationally, regionally, and globally, especially in low- and middle-income countries, where the self-management resource gap limits effective control [50,51]. Evidence shows promise that DHIs improve behaviors, behavioral risk factors, patient engagement, and cardiometabolic outcomes [7,52]. However, evidence from real-world community settings in the Southeast Asian region is limited. This study addresses this gap by integrating smartwatch-based monitoring and digital salt intake assessment with nurse-led self-management.

Even though a single-arm design cannot infer causality, it is appropriate to evaluate feasibility preliminarily. Feasibility and pilot studies often use single-arm study designs to assess recruitment, adherence to intervention components, acceptability, and data completeness before conducting controlled trials [53]. The single-arm study is recommended when an intervention is novel, complex, or requires refinement before large-scale testing [54]. In this perspective, single-arm studies provide a valuable estimate of within-person change, intervention responsiveness, and variability needed for future sample size calculation [55].

The intervention designed in this protocol incorporates several strengths. First, it leverages evidence-based components of chronic disease self-management, including self-monitoring, personalized case management, behavioral change support and feedback, delivered by nurses, an approach that has proven effectiveness in improving clinical outcomes [12,56,57]. Second, a smartwatch wearable device provides objective indicators of physical activity, heart rate, and sleep, contributing more accurate behavioral data than self-report [19,58,59]. In addition, nurse-led case management with digital support aligns with World Health Organization (WHO) recommendations for integrating digital tools within primary care to enhance person-centered and continuing care [60-63].

Several limitations must be acknowledged and addressed. In this early phase of a single-arm study, the study is vulnerable to secular trends, regression toward the mean, and the Hawthorne effect (behavioral reactivity due to monitoring) because there is no comparison group [64]. These limitations cannot be eliminated; it is crucial to interpret carefully and avoid making casual claims [53]. Moreover, a short duration may not capture the long-term physiological change required for NCD progression. This study also has limitations: a small patient population, limited to 1 geographic area, which may affect generalizability, and the use of commercial devices, whose accuracy may vary [60]. However, the single-arm study design, which is accepted by the US Food and Drug Administration and the European Medicines Agency, has the advantage of requiring a smaller sample size and a quicker timeline [64].

A significant limitation of this study is selection bias arising from the use of restrictive eHEALS and an intent threshold. This methodology likely filtered out individuals with lower digital proficiency who often experience greater health disparities. Consequently, the reported efficacy and feasibility signs may overestimate the true impact of the intervention in the general population. The results are generalizable only to populations with similar high levels of digital literacy and motivation. A future study will use a lower or zero eHEALS cut-off to assess real-world effectiveness and to analyze digital competence as a moderator of intervention success.

Despite these limitations, this study is positioned to contribute to the evidence on digital health integration in community-based NCD management. Identifying feasibility metrics, adherence patterns, preliminary effect sizes, and patient experiences will guide the refinement of the intervention and inform the design of the following randomized trial. This study also supports broader efforts to strengthen nurse-led digital chronic NCD management within primary care systems in low- to middle-income settings.

Conclusions

This study aims to investigate whether a combined approach involving a nurse-led 12-week self-management intervention coupled with digital device support, including a smartwatch, a mobile app, and a salt meter for the first 6 weeks, can enhance self-management behaviors and improve nurse-led case management clinical outcomes. This intervention was developed using existing evidence-based materials and inputs from important stakeholders. The study will assess changes in self-management behaviors across baseline, week 6, week 12, and week 24. Primary outcomes, including self-management behaviors, urine albumin-to-urea ratio, BP, digital engagement, and satisfaction with the intervention, will be assessed at the end of week 6. Additionally, clinical outcomes, including BP, BMI, lipid profiles, HbA_{1c}, ankle-brachial index, estimated glomerular filtration rate, and CVD risk, will be assessed at baseline, week 12, and week 24.

XSL•FO
RenderX

The results will inform the development of a randomized controlled study for a self-management program coordinated with digital components. This program will include both behavior changes and maintenance strategies, clinical outcome assessment, and participant engagement and progress.

Acknowledgments

The authors would like to sincerely thank all the participants of this research for their valuable time. The authors are grateful to all five professional experts who contributed to validating the intervention program and the self-management support manual. The authors would like to acknowledge Chok Chai Hospital and Ratchasima Hospital, particularly the NCD nurses, family doctors, and nutritionists, for their input and kind collaboration throughout the project. The Practical Solution Public Company Limited provided the digital devices used in this study, a Huawei Watch Fit 3 smartwatch, free of charge as an in-kind donation.

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision:

- Idea generation

- Literature search and systematization

- Writing the literature review - Analysis of market trends and/or patent environment

- Proofreading and editing

The GAI tool used was: Gemini 2.5, Chat GPT5.1, SciSpace, Grammarly AI. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Additional note: The authors used Gemini 2.5, Chat GPT5.1, and SciSpace to search for relevant information. After getting the idea, the authors located original articles to validate the content and manually drafted each topic of this manuscript. Afterward, the authors used Grammarly AI and Microsoft Word writing tools to polish writing. The authors created figures and images for this manuscript without the use of AI. In brief, the authors did not use AI to generate this manuscript; instead, AI tools were only used to aid in searching for information and refining the language. All content was subsequently reviewed, verified, and edited by the authors to ensure accuracy and academic integrity.

Declaration submitted by: Collective responsibility (Naruemol Singha-Dong, Raweewan Pongpoottipatchara, Kanlayawee Anonjarn, Warithorn Prawatwong)

Funding

This project received funding from the Thailand Science Research and Innovation (TSRI) under grant agreement FF8-803-66-24-175(H). The TSRI is not involved in the study design, data collection, data analysis, interpretation of results, decision to publish, or the writing of the manuscript. Only study investigators will have access to the final data set.

Data Availability

The authors are willing to share a complete, cleaned, and de-identified copy of the data upon approval of written requests by the study team, which must include an analytic plan, after publication of the primary and secondary aims in a peer-reviewed journal.

Authors' Contributions

Conceptualization: NS
Funding acquisition: NS, RP
Investigation: NS, RP, KA, WP
Methodology: NS, RP, KA, WP
Project administration: NS
Resources: NS
Software: NS, RP, KA
Supervision: NS
Visualization: NS, RP
Writing–original draft: NS, RP, KA
Writing–review & editing: NS, RP, KA

Checklist 1
SPIRIT checklist [65]
[DOCX File, 35 KB - resprot_v15i1e80083_app1.docx]

XSL•FO
RenderX

References

1. Cardiovascular diseases (cvds). World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2025-06-17]
2. Ruan Y, Bao Q, Wang L, Wang Z, Zhu W, Wang J. Cardiovascular diseases burden attributable to ambient PM2.5 pollution from 1990 to 2019: a systematic analysis for the global burden of disease study 2019. Environ Res 2024 Jan 15;241:117678. [doi: 10.1016/j.envres.2023.117678] [Medline: 37984788]
3. Tan SCW, Zheng BB, Tang ML, Chu H, Zhao YT, Weng C. Global burden of cardiovascular diseases and its risk factors, 1990–2021: a systematic analysis for the global burden of disease study 2021. QJM 2025 Jun 1;118(6):411-422. [doi: 10.1093/qjmed/hcaf022] [Medline: 39847534]
4. Roth GA, Mensah GA, Johnson CO, et al. Global Burden of Cardiovascular Diseases and Risk Factors, 1990-2019: Update From the GBD 2019 Study. J Am Coll Cardiol 2020 Dec 22;76(25):2982-3021. [doi: 10.1016/j.jacc.2020.11.010] [Medline: 33309175]
5. Martin SS, Aday AW, Allen NB, et al. 2025 heart disease and stroke statistics: a report of US and global data from the American Heart Association. Circulation 2025 Feb 25;151(8):e41-e660. [doi: 10.1161/CIR.0000000000001303] [Medline: 39866113]
6. Ketsing N, et al. Cardiovascular disease prevalence and risk in rheumatoid arthritis patients: insights from a nationwide Thai hospital database. Thai J Rheumatol 2025;2(2):1-6 [FREE Full text]
7. Sukpattanasrikul S, Singha-Dong N, Sitthimongkol Y, Anonjarn K. Efficacy and cost-effectiveness of digital health interventions in improving hypertensive outcomes among patients with uncontrolled hypertension: a systematic review. Int J Nurs Sci 2025 May;12(3):225-232. [doi: 10.1016/j.ijnss.2025.04.001] [Medline: 40529464]
8. World health statistics 2024: monitoring health for the SDGs, sustainable development goals. World Health Organization. URL: https://www.who.int/publications/i/item/9789240094703 [accessed 2024-05-21]
9. DiCaro MV, Ogurek I, Tak N, Dawn B, Tak T. Optimizing cardiovascular health: a narrative review of lifestyle, psychobehavioral, and alternative strategies for management and prevention. Heart Mind 2025;9(1):29-39. [doi: 10.4103/hm.HM-D-24-00066]
10. Riegel B, Moser DK, Buck HG, et al. Self‐care for the prevention and management of cardiovascular disease and stroke: a scientific statement for healthcare professionals from the American heart association. J Am Heart Assoc 2017 Aug 31;6(9):e006997. [doi: 10.1161/JAHA.117.006997] [Medline: 28860232]
11. Gosak L, Vrbnjak D, Pajnkihar M. Self-management of chronic diseases: a concept analysis. Nurs 21st Cent [Pielegniarstwo XXI wieku] 2022 Jun 1;21(2):115-121. [doi: 10.2478/pielxxiw-2022-0006]
12. Sukpattanasrikul S, Monkong S, Leelacharas S, Krairit O, Viwatwongkasem C. Comparison of hypertensive outcomes after the implementation of self-management program for older adults with uncontrolled hypertension in Krabi, Thailand: a quasi-experimental study. J Health Res 2022 Jul 4;36(4):641-651. [doi: 10.1108/JHR-12-2020-0626]
13. Tharani A, Van Hecke A, Ali TS, Duprez V. Factors influencing nurses' provision of self-management support for patients with chronic illnesses: a systematic mixed studies review. Int J Nurs Stud 2021 Aug;120:103983. [doi: 10.1016/j.ijnurstu.2021.103983] [Medline: 34147728]
14. Thaiudom S, Posridee K, Liangchawengwong S, et al. Plant-based diet for glycemic control, insulin sensitivity, and lipid profile in type 2 diabetes: a systematic review. Foods 2025 May 28;14(11):1919. [doi: 10.3390/foods14111919] [Medline: 40509445]
15. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52. [doi: 10.2196/jmir.3951] [Medline: 25803266]
16. Redfern J, Coorey G, Mulley J, et al. A digital health intervention for cardiovascular disease management in primary care (CONNECT) randomized controlled trial. NPJ Digit Med 2020;3:117. [doi: 10.1038/s41746-020-00325-z] [Medline: 32964140]
17. Qiu Y, Asselin CY, Kallas D, et al. The emerging role of digital health technology in cardiovascular care. Can J Cardiol 2021 Jul;37(7):939-942. [doi: 10.1016/j.cjca.2021.04.016]
18. Singh V, Kr J, Jacob AG, John O. A scoping review of digital health interventions for cardiovascular diseases in the WHO South-East Asia region. Conn Health Telemed 2023;2(4). [doi: 10.20517/chatmed.2023.07] [Medline: 2000012]
19. Widmer RJ, Collins NM, Collins CS, West CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. Mayo Clin Proc 2015 Apr;90(4):469-480. [doi: 10.1016/j.mayocp.2014.12.026] [Medline: 25841251]
20. Lawrence MR, Fuller BP, Isabirye HK, et al. Health data reporting challenges in the age of DHIS2: a mixed-methods review in the teso sub-region of uganda. medRxiv. Public and Global Health. Preprint posted online on 2024. [doi: 10.1101/2024.09.18.24313921]
21. Triantafyllidis A, Kondylakis H, Katehakis D, et al. Smartwatch interventions in healthcare: a systematic review of the literature. Int J Med Inform 2024 Oct;190:105560. [doi: 10.1016/j.ijmedinf.2024.105560] [Medline: 39033723]

XSL•FO
RenderX

22. Zhou W, Shang S, Cho Y. Associations of wearable activity tracker use with physical activity and health outcomes in patients with cancer: Findings from a population-based survey study. J Med Internet Res 2024 Oct 22;26:e51291. [doi: 10.2196/51291] [Medline: 39436693]
23. Maj D, Wieteska M, Gorecka A, Gorajek A, Wacławska M. The DASH diet—beneficial impacts on cardiovascular system and even more. J Educ Health Sport 2022;12(6):112-117. [doi: 10.12775/JEHS.2022.12.06.011]
24. Sangouni AA, Hosseinzadeh M, Parastouei K. The effect of dietary approaches to stop hypertension (DASH) diet on fatty liver and cardiovascular risk factors in subjects with metabolic syndrome: a randomized controlled trial. BMC Endocr Disord 2024 Jul 25;24(1):126. [doi: 10.1186/s12902-024-01661-x] [Medline: 39054440]
25. Alifah K, Damayanti YAE, Rahmani A. 22. The effect of DASH diet on blood pressure and metabolic profile in hypertensive patients: a systematic review and meta-analysis. J Hypertens (Los Angel) 2023;41(Suppl 2):e6. [doi: 10.1097/01.hjh.0000935468.48116.98]
26. Shoaibinobarian N, Danehchin L, Mozafarinia M, et al. The association between DASH diet adherence and cardiovascular risk factors. Int J Prev Med 2023;14(1):24. [doi: 10.4103/ijpvm.ijpvm_336_21] [Medline: 37033286]
27. Jeong SY, Wee CC, Kovell LC, et al. Effects of diet on 10-year atherosclerotic cardiovascular disease risk (from the DASH Trial). Am J Cardiol 2023 Jan 15;187:10-17. [doi: 10.1016/j.amjcard.2022.10.019] [Medline: 36459731]
28. Belanger MJ, Kovell LC, Turkson-Ocran RA, et al. Effects of the dietary approaches to stop hypertension diet on change in cardiac biomarkers over time: results from the DASH-sodium trial. J Am Heart Assoc 2023 Jan 17;12(2):e026684. [doi: 10.1161/JAHA.122.026684] [Medline: 36628985]
29. Lari A, Sohouli MH, Fatahi S, et al. The effects of the dietary approaches to stop Hypertension (DASH) diet on metabolic risk factors in patients with chronic disease: a systematic review and meta-analysis of randomized controlled trials. Nutr Metab Cardiovasc Dis 2021 Sep 22;31(10):2766-2778. [doi: 10.1016/j.numecd.2021.05.030] [Medline: 34353704]
30. Youngiam W, Kitreerawutiwong N. Reducing salt and sodium intake through health promotion strategies. J Public Health Health Sci Res 2022;4(2):1-21 [FREE Full text]
31. Wiriyatanakorn S, Mukdadilok A, Kantachuvesiri S, Mekhora C, Yingchoncharoen T. Impact of self-monitoring of SALT intake by SALT meter in hypertensive patients: a randomized controlled trial (SMAL-SALT). J Clin Hypertens (Greenwich) 2021 Oct;23(10):1852-1861. [doi: 10.1111/jch.14344] [Medline: 34510703]
32. Strazzullo P, D'Elia L, Kandala NB, Cappuccio FP. Salt intake, stroke, and cardiovascular disease: meta-analysis of prospective studies. BMJ 2009 Nov 24;339:b4567. [doi: 10.1136/bmj.b4567] [Medline: 19934192]
33. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 Statement: Updated Guideline for Protocols of Randomized Trials. JAMA 2025 Aug 5;334(5):435-443. [doi: 10.1001/jama.2025.4486] [Medline: 40294593]
34. Wongsawat Y, Brain computer interface lab researcher teams. CHEM meter: salinity detector in food samples. Mahidol University. 2024. URL: https://int.mahidol.ac.th/en/2024/05/01/chem-meter-en [accessed 2020-12-10]
35. CHEM Meter: . 2024 Nov 6. URL: https://www.facebook.com/photo.php?fbid=1221653225000454 83&set=pb.61551364499706.-2207520000&type=3 [accessed 2025-02-18]
36. Nojai V, et al. Effects of Somporn Kantharadussadee Triumchaisri (SKT) meditation on blood glucose levels of pre-diabetes mellitus. J R Thai Army Nurses 2018;19(3):175-184 [FREE Full text]
37. Chibsamanboon P, Suttineam U. Effects of the Somporn Kantaradusdi-Triamchaisri Technique 2 (SKT 2) on blood pressure levels and biochemical markers. J Boromarajonani Coll Nurs 2013;29(2):122-123 [FREE Full text]
38. Richtering SS, Hyun K, Neubeck L, et al. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. JMIR Hum Factors 2017 Jan 27;4(1):e4. [doi: 10.2196/humanfactors.6217] [Medline: 28130203]
39. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9. [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
40. Plagatong W, et al. Reliability and feasibility of the eHEALS in Thai patients with rheumatic and autoimmune diseases. Thai J Rheumatol 2024;1(3):1-16 [FREE Full text]
41. Tangjitgamol S, Udayachalerm W, Panyarachun S, Wanishsawad C, Kaewwanna W, Manusirivithaya S. The 10-year cardiovascular risk of physicians estimated by the Thai CV risk score and its association with current coronary artery disease: a retrospective study. Health Sci Rep 2024 Apr;7(4):e2009. [doi: 10.1002/hsr2.2009] [Medline: 38629110]
42. Food consumption behavior screening form (sweet, fatty, salty). Thai Health Promotion Foundation. 2024. URL: https://dol.thaihealth.or.th/Media/Index/3a123c3a-39b0-ee11-8100-00155d1aab77 [accessed 2025-03-01]
43. Carpi M. The Pittsburgh Sleep Quality Index: a brief review. Occup Med (Lond) 2025 Apr 4;75(1):14-15. [doi: 10.1093/occmed/kqae121] [Medline: 40190123]
44. Sitasuwan T, Bussaratid S, Ruttanaumpawan P, Chotinaiwattarakul W. Reliability and validity of the Thai version of the Pittsburgh Sleep Quality Index. J Med Assoc Thai 2014 Mar;97 Suppl 3:S57-S67. [Medline: 24772581]
45. Silpakit O. Srithanya Stress Scale. J Ment Health Thailand 2012;16(3):177-185. [doi: 10.14456/dmhjournal.2008.19]
46. 2024 Thai Guidelines on the Treatment of Hypertension. The Thai Hypertension Society. URL: https://thaihypertension.org/wp-content/uploads/2025/07/Guideline2024.pdf [accessed 2024-09-01]
47. Repeated measures ANOVA—understanding a repeated measures ANOVA. Laerd Statistics. URL: https://statistics.laerd.com/statistical-guides/repeated-measures-anova-statistical-guide.php [accessed 2025-12-10]

XSL•FO
RenderX

48. Muhammad LN. Guidelines for repeated measures statistical analysis approaches with basic science research considerations. J Clin Invest 2023 Jun 1;133(11):11. [doi: 10.1172/JCI171058] [Medline: 37259921]
49. Personal data protection act, b.e. 2562 (2019). 2019. URL: https://www.mdes.go.th/law/detail/3577-Personal-Data-Protection-Act-B-E--2562--2019 [accessed 2026-01-12]
50. Potempa K, Rajataramya B, Singha-Dong N, Furspan P, Kahle E, Stephenson R. Thailand's challenges of achieving health equity in the era of non-communicable disease. Pac Rim Int J Nurs Res Thail 2022;26(2):187-197. [Medline: 35432661]
51. Zhou B, Carrillo-Larco RM, Danaei G. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021 Sep 11;398(10304):957-980. [doi: 10.1016/S0140-6736(21)01330-1] [Medline: 34450083]
52. Zhang Y, Ngai FW, Yang Q, Xie YJ. Effectiveness of digital health interventions on sedentary behavior among patients with chronic diseases: systematic review and meta-analysis. JMIR mHealth uHealth 2025 Jun 24;13:e59943. [doi: 10.2196/59943] [Medline: 40554804]
53. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239. [doi: 10.1136/bmj.i5239] [Medline: 27777223]
54. Craig P, et al. Taking account of context in the population health intervention research process in taking account of context in population health intervention research: guidance for producers, users and funders of research.: National Center for Biotechnology Information; 2018. URL: https://www.ncbi.nlm.nih.gov/books/NBK498650/ [accessed 2026-01-12]
55. Lancaster GA. Pilot and feasibility studies come of age!. Pilot Feasibility Stud 2015;1(1):1. [doi: 10.1186/2055-5784-1-1] [Medline: 29611687]
56. Davis KM, Eckert MC, Hutchinson A, et al. Effectiveness of nurse-led services for people with chronic disease in achieving an outcome of continuity of care at the primary-secondary healthcare interface: a quantitative systematic review. Int J Nurs Stud 2021 Sep;121:103986. [doi: 10.1016/j.ijnurstu.2021.103986] [Medline: 34242979]
57. Rahmat R, Haroen H, Juniarti N, Sari SP, Rinawan FR. Effectiveness of community nurse-led intervention in managing older adults with multimorbidity: a systematic review of randomized controlled trials. J Multidiscip Healthc 2025;18:6373-6389. [doi: 10.2147/JMDH.S548950] [Medline: 41064332]
58. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12(1):159. [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
59. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 8;8(9):e18694. [doi: 10.2196/18694] [Medline: 32897239]
60. Labrique A, Agarwal S, Tamrat T, Mehl G. WHO digital health guidelines: a milestone for global health. NPJ Digit Med 2020;3(1):120. [doi: 10.1038/s41746-020-00330-2] [Medline: 33015373]
61. WHO guideline recommendations on digital interventions for health system strengthening: World Health Organization; 2019. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2026-01-12]
62. Mehl G, Tunçalp Ö, Ratanaprayul N, et al. WHO SMART guidelines: optimising country-level use of guideline recommendations in the digital age. Lancet Digit Health 2021 Apr;3(4):e213-e216. [doi: 10.1016/S2589-7500(21)00038-8] [Medline: 33610488]
63. Global strategy on digital health 2020-2025. : World Health Organization; 2021 URL: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf [accessed 2025-12-10]
64. Yao S, Shang Q, Ouyang M, et al. Designing single-arm clinical trials: principles, applications, and methodological considerations. Ann Clin Epidemiol 2025 Mar 14;7(3):90-98. [doi: 10.37737/ace.25011] [Medline: 40697795]
65. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]

Abbreviations

BP: blood pressure
CVD: cardiovascular disease
DALY: disability-adjusted life year
DASH: Dietary Approaches to Stop Hypertension
DBSQ: Dietary Behavior Screening (Sweet, Fatty, Salty) Questionnaire
DHI: digital health intervention
eHEALS: eHealth Literacy Scale
HbA_{1c}: hemoglobin A_{1c}
NCD: noncommunicable disease
PSQI: Pittsburgh Sleep Quality Index
SKT: Somporn Kantaradusdi-Triamchaisri
WHO: World Health Organization
WHO: World Health Organization

XSL•FO
RenderX

Edited by J Sarvestan; submitted 04.Jul.2025; peer-reviewed by A Afzalian; revised version received 01.Dec.2025; accepted 02.Dec.2025; published 22.Jan.2026.

Please cite as:
Pongpoottipatchara R, Anonjarn K, Prawatwong W, Singha-Dong N
Implementation of Digital Device–Assisted and Nurse-Led Case Management to Promote Self-Management in Adults With Noncommunicable Diseases: Protocol for a Single-Arm Intervention Study
JMIR Res Protoc 2026;15:e80083
URL: https://www.researchprotocols.org/2026/1/e80083
doi:10.2196/80083

© Raweewan Pongpoottipatchara, Kanlayawee Anonjarn, Warithorn Prawatwong, Naruemol Singha-Dong. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Home-Based Exercise Cardiac Telerehabilitation on Adherence and Functional Capacity for Patients After Percutaneous Coronary Intervention in Indonesia: Protocol for a Quasi-Experimental Study

Wantiyah Wantiyah[1,2], MN; Anggoro Budi Hartopo[3], PhD; Sri Setiyarini[4], PhD; Budi Yuli Setianto[3], PhD

[1]Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia

[2]Department of Medical Surgical and Critical Nursing, Faculty of Nursing, Universitas Jember, Jember, Indonesia

[3]Department of Cardiology and Vascular Medicine, Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia

[4]Department of Basic and Emergency Nursing, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Gedung Ismangoen, Jl. Farmako Sekip, Yogyakarta, Indonesia

Corresponding Author:
Sri Setiyarini, PhD
Department of Basic and Emergency Nursing, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Gedung Ismangoen, Jl. Farmako Sekip, Yogyakarta, Indonesia

Abstract

Background: Home-based cardiac rehabilitation is a promising approach to improve access and adherence to cardiac rehabilitation (CR), which remains underutilized. However, monitoring patient adherence and safety during home-based exercise is challenging. Therefore, we develop and implement a home-based exercise cardiac telerehabilitation (HBECTR) program, integrating structured exercise guidance with remote monitoring.

Objective: This trial is designed as a pilot feasibility study to evaluate the practicality, safety, and preliminary effects of HBECTR in patients who have undergone percutaneous coronary intervention (PCI). The primary objective is to analyze the effect of HBECTR on patients' adherence to CR, while the secondary objective is to measure the effect of HBECTR on the functional capacity of patients after PCI.

Methods: This study is a quasi-experimental nonequivalent control group pretest-posttest. The eligible participants are adults (≥18 y), patients with coronary artery disease after successful PCI, and patients having active health insurance. Patients will be consecutively recruited from Dr. Sardjito General Hospital, Yogyakarta, Indonesia, with a minimum of 60 patients after PCI as samples, divided into 2 groups: an intervention group and a comparison or control group. The intervention group will receive hybrid cardiac rehabilitation. Patients will be scheduled to do exercise twice a week at the hospital and 3 times a week at their home. When patients exercise at home, they will have to wear the smartwatch "CovWatch," which will be connected to the web-based app (App Rehab Cardio) to monitor their vital signs and the distance of walking. The control group will get usual care (center-based cardiac rehabilitation). The data will be analyzed using the Student *t* test to see if the data comply with the normality assumption. This protocol received approval from the institutional review board of the Faculty of Medicine, Public Health, and Nursing UGM (approval number: KE/FK/0850/EC/2024).

Results: The recruitment of the patients started in August 2024. This study is being funded since February 2025. During the submission, we have recruited 63 patients in both intervention and control groups. The intervention for both groups was finished in June 2025, and the data are being analyzed.

Conclusions: It is expected that the HBECTR program will be feasible and safely implemented in low- and middle-income countries, particularly in Indonesia. By integrating a telerehabilitation model, it is expected that HBECTR will improve patients' adherence to CR and also increase the functional capacity of patients after PCI.

Trial Registration: ANZCTR ACTRN12625001400459; https://anzctr.org.au/ACTRN12625001400459.aspx (retrospectively registered)

International Registered Report Identifier (IRRID): DERR1-10.2196/81067

(JMIR Res Protoc 2026;15:e81067) doi:10.2196/81067

KEYWORDS

adherence; cardiac rehabilitation; functional capacity; home-based exercise cardiac telerehabilitation; percutaneous coronary intervention

XSL•FO
RenderX

Introduction

Background

Cardiovascular disease (CVD) is the leading cause of death and disability worldwide, with the highest death rates (more than 75%) occurring in low- and middle-income countries [1]. Indonesia reported 765,660 deaths from CVD in 2021, placing Indonesia in the top 20% rank of countries with the highest mortality based on age-standardized mortality from CVD. Furthermore, Indonesia has relatively high rates of non-high-density lipoprotein cholesterol in women and tobacco use in men when compared to other countries globally. Indonesia has implemented 4 of 8 key CVD-related policies. It does have a national action plan for CVDs in place and an operational unit within the Ministry of Health with the responsibility for managing non-communicable diseases; however, the action plan to reduce physical inactivity has not been developed [2].

Cardiac rehabilitation (CR), a multicomplex intervention that is a class 1A recommendation from the American Heart Association and the American College of Cardiology, aims to improve the health condition of patients [3,4], reduce mortality, enhance recovery, and support positive behavioral changes to reduce risk factors [4-7]. Although the effectiveness and benefits of the CR program have been proved, it is still underutilized, and its access is limited [5]. Globally, the level of participation and compliance of patients undergoing CR is still low [8], especially in countries with limited resources [9]. The compliance rate for patients with post-heart attack is 14% - 35% [10] and around 24% for patients with coronary heart disease with various conditions [11]. It means that more than 80% of the patients who are referred to CR do not participate in the program [12]. Patients' participation and adherence to CR are essential to determine the success of prevention efforts and improve patients' condition [13].

Exercise-based cardiac rehabilitation (EBCR) is essential for patients undergoing revascularization (postpercutaneous coronary intervention [PCI]) because they often experience decreased exercise capacity and physical activity levels, and both of these indicators are very important for predicting postprocedure morbidity and mortality [14]. Furthermore, these patients also need self-management to improve their health conditions and other clinical outcomes, such as reducing depression and anxiety, mortality, and morbidity and improving their quality of life [15]. EBCR focuses on improving cardiac endurance, exercise capacity, muscle strength, physical activity levels, and quality of life through health education and lifestyle modification [14]. Physical exercise is the main pillar (core stone) of cardiac rehabilitation [13] because it can increase functional capacity and encourage patients to develop and maintain an active lifestyle that can reduce the risk of recurrence in the future [16].

In Indonesia, CR is predominantly implemented in hospital-based or center-based cardiac rehabilitation (CBCR) settings, as the national health insurance scheme—Badan Penyelenggara Jaminan Sosial (BPJS)—currently reimburses only CBCR services. In contrast, home-based cardiac rehabilitation (HBCR), cardiac telerehabilitation, and virtual rehabilitation programs are not yet covered by the insurance system [17]. Although BPJS facilitates access to CBCR, CR services remain limited because not all hospitals in Indonesia have established CR facilities. Similar to other countries in the Southeast Asia region, most available programs are concentrated in tertiary care centers or major urban areas [18]. Dr. Sardjito General Hospital in Yogyakarta is one of the few hospitals in Indonesia that provides CR services and has offered EBCR since 2006. However, the accessibility and continuity of CR remain major challenges, particularly for patients who live far from the hospitals or face transportation and time constraints.

The implementation of EBCR at Dr. Sardjito Hospital also has evolved considerably between the pre- and post-COVID-19 periods, particularly in terms of session frequency and delivery format. Before the pandemic, the EBCR program consisted of 10 - 12 sessions, typically conducted 3 - 5 times per week. During the pandemic, however, the frequency was reduced to twice per week while maintaining the total number of 10 - 12 sessions [17]. At present, most patients participate in CBCR twice weekly for 4 weeks, resulting in approximately 8 sessions overall. The duration and frequency may be extended depending on each patient's clinical condition or physician's recommendation. Nevertheless, the current practice remains below the national guideline, which recommends that patients engage in exercise 4 - 7 times per week to achieve optimal outcomes [19]. Therefore, incorporating home-based exercise as part of a hybrid or fully HBCR program may serve as a feasible and scalable alternative to meet the recommended exercise frequency, enabling patients to continue rehabilitation safely and effectively at home on nonhospital days.

HBCR has a positive effect on patient self-efficacy [7], which will encourage patients to behave positively. Besides, home-based exercises have been proven to be very effective in improving patient conditions because they can empower patients to monitor the exercises they do independently [13]. Moreover, HBCR has the same safety and effectiveness as hospital or CBCR in improving patient functional capacity and various other clinical outcomes [1,5,20-22]. HBCR has emerged as a promising alternative to CBCR [5,23].

Despite the various advantages and benefits of HBCR, there are several obstacles, such as the absence of clear guidelines regarding dosage and components, its limited implementation in certain cases, and its limited implementation in all countries. The research and implementation of HBCR are mostly in developed countries such as the United States, Australia, and England [5,24].

Adding artificial intelligence that can interpret data and monitor the implementation of CR can increase the safety and feasibility of HBCR. Devices installed or worn by patients are expected to identify the patients' toleration of exercise dose and monitor the vital signs to ensure patient safety during exercise at home [20]. The use of information technology in its implementation is predicted to increase the long-term effects of HBCR [24]. Smartphones or mobile devices (20/31, 65%), web-based portals (18/31, 58%), and email or SMS (11/31, 35%) are the most often used information technologies to support CR [12]. In this study, we will use a web-based app (dashboard) to support the

XSL•FO
RenderX

implementation of home-based exercise cardiac telerehabilitation (HBECTR).

Although several studies support the implementation of telerehabilitation, there are still no definitive data on the efficacy of HBECTR compared to CBCR, thus requiring further research [23,25]. Besides, the biggest challenge today is how to improve patient adherence and compliance in CR, the development of HBCR guidelines, and the implementation of information technology in HBCR that can be applied in countries with limited resources, like Indonesia. Therefore, this study is designed as a pilot feasibility trial to evaluate the practicality, safety, and short-term effects of a digitally supported HBECTR program in Indonesian patients after PCI. This is the first pilot study in Indonesia implementing digitally supported HBCR, particularly using a web-based platform, and addressing a critical gap in low- and middle-income country settings where CR utilization remains low. The pilot nature of this study allows us to examine the feasibility of the program, adherence, safety monitoring, and digital health implementation in a resource-limited setting.

Objectives

This study aimed to evaluate the effectiveness of an HBECTR program for patients after PCI. The primary objective was to assess the effect of HBECTR on adherence to CR, while the secondary objective was to determine its effect on patients' functional capacity.

Hypotheses

In this study, we examined two hypotheses:

1. There is no difference in functional capacity between the HBECTR group and the control group (CG; H_0) or HBECTR will improve functional capacity compared to the control (H_1).
2. There is no difference in adherence between the HBECTR group and the CG (H_0) or HBECTR will improve adherence to CR compared to the control (H_1).

Methods

Study Design

This trial is designed as a pilot feasibility study to evaluate the practicality, safety, and preliminary effects of HBECTR using a quasi-experimental nonequivalent CG pretest-posttest. Participants will be assigned to either the intervention group (IG), which will receive the HBECTR program, or the CG, which will receive standard care (CBCR). Pre- and post-intervention data will be collected to assess the impact of the HBECTR program on patients' adherence as primary outcomes and their functional capacity as secondary outcomes.

We use a quasi-experimental design with a nonrandomized control group due to the limited number of eligible participants [26], the logistical and resource constraints, and to bringing the implementation closer to the real world. As a pilot study, we want to identify the feasibility of HBECTR in Indonesia. In fact, there is a strong need to further investigate and incorporate synchronous digital CR into the routine of cardiac patients' secondary prevention routine [27].

Settings

This study was being conducted since August 2024 and June 2025 at Dr. Sardjito Yogyakarta General Hospital, Indonesia, focusing on patients after PCI. The intervention will be done in the CR center at Dr. Sardjito General Hospital, Yogyakarta, and at the patients' homes. The research environment will include outpatient clinics and remote monitoring settings where participants will receive either the HBECTR program or standard care. The setting will ensure access to medical records, availability of telehealth technologies, and close coordination with health care professionals to support the implementation and evaluation of the HBECTR program.

Samples

The population consists of patients after PCI recruited using consecutive sampling. Sample size calculation was performed based on previous studies [15,28], with the minimum difference of 6-minute walk distance as 30 m [29,30]. We determined the patients based on the 95% CI, 80% power, 5% type I error, and 10% dropout rate. A minimum of 60 patients (30 per group) will be included, consistent with recommendations for pilot feasibility studies, which suggest 12 - 30 participants per arm to estimate variability and assess feasibility parameters [31,32]. We will carefully select the samples based on the inclusion criteria: (1) patients with coronary artery disease after successful PCI, (2) aged 18 years or above, (3) covered by active health insurance (BPJS), (4) willing to participate and sign informed consent, (5) having internet connectivity at home (for the CG), (6) living with family or having a family member as a caregiver, and (7) able to read and understand Bahasa Indonesia.

The exclusion criteria were as follows: (1) patients at a high cardiovascular risk as defined by the American Association of Cardiovascular and Pulmonary Rehabilitation (2007), the American Heart Association (2001), Société Française de Cardiologie (2002), and Indonesia based on the Indonesia Heart Association (2019), which included any of the following: (a) resting ejection fraction <30%, (b) history of cardiac arrest or sudden death, (c) complex arrhythmia at rest, (d) multiple imputation (MI) or revascularization procedure with complications, (e) congestive heart failure, (f) symptoms and signs of ischemia post-MI or post-revascularization, or (g) clinical depression; (2) inability to walk independently; and (3) having contraindications to participate in CR, such as (a) unstable angina, (b) resting systolic blood pressure (BP) >200 mm Hg or diastolic BP >110 mm Hg (requiring case-by-case evaluation), (c) orthostatic BP drop >20 mm Hg with symptoms, (d) critical aortic valve stenosis, (e) active systemic illness or fever, (f) uncontrolled atrial or ventricular arrhythmia, (g) uncontrolled sinus tachycardia (>120 bpm), (h) uncompensated heart failure, (i) third-degree atrioventricular block without a pacemaker, (j) active pericarditis or myocarditis, (k) new thromboembolic events, (l) ST-segment depression or elevation at rest (>2 mm), (m) uncontrolled diabetes, (n) severe orthopedic conditions preventing physical exercise, and (o) other metabolic issues such as acute thyroiditis, hypokalemia, hyperkalemia, or hypovolemia. Drop-out criteria included (1) patients who only

XSL•FO
RenderX

participated once or did not complete the full rehabilitation program (loss to follow-up) and (2) patients who have discontinued the program due to clinical reasons (medical advised).

Both groups will be recruited consecutively using identical eligibility criteria, recruitment procedures, and measurement protocols. This study uses a quasi-experimental nonequivalent control group design where the IG was recruited and completed follow-ups before the initiation of the control or comparison group. Sequential, nonconcurrent enrollment was selected due to practical constraints, including a limited number of eligible patients and complete essential preparatory activities and devices (smartwatch and dashboard setup, system testing, and staff training) prior to delivering the hybrid cardiac telerehabilitation intervention. Similar sequential recruitment approaches are common in pragmatic and feasibility trials where randomization or concurrent enrollment is not operationally feasible [26,33].

However, such nonconcurrent enrollment introduces potential time-related biases that statistical adjustment cannot fully eliminate [33]. To mitigate these concerns, we will implement two strategies. First, in the pre-intervention assessment, we will perform demographic and clinical assessments at baseline across both groups. We will carefully select patients based on the eligibility criteria, standardize all outcome measurements (attendance list, 6MWT) using the same instruments and trained team, and apply identical follow-up schedules. These data will help us statistically adjust for any identified differences, although we recognize that such adjustments may not completely eliminate bias. Second, in propensity score matching (PSM), we will plan to use PSM to equate the groups on key variables measured at baseline, which can help balance characteristics across the groups. Baseline demographic and clinical characteristics will be carefully documented, and differences between the groups will be statistically adjusted using multivariable regression and propensity score methods. The use of PSM enabled the construction of a comparison group comprising individuals similar to the treatment group individuals on most observable characteristics [34].

Blinding procedures are often not feasible in quasi-experimental studies due to the absence of randomized assignment and the practical realities of real-world intervention delivery, particularly in exercise, physical activity, and rehabilitation research [35]. In this study, the full blinding of participants and assessors was not possible because of the overt nature of the intervention and logistical constraints. Nevertheless, the risk of bias was mitigated by implementing several strategies: using validated measurement tools; providing objective and standardized outcome assessments; and, wherever feasible, ensuring that data analysts were blinded to group allocation during the data analysis phase. Data were coded such that the analysts were unaware of group assignments until after the completion of initial analysis (see Figure 1).

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.

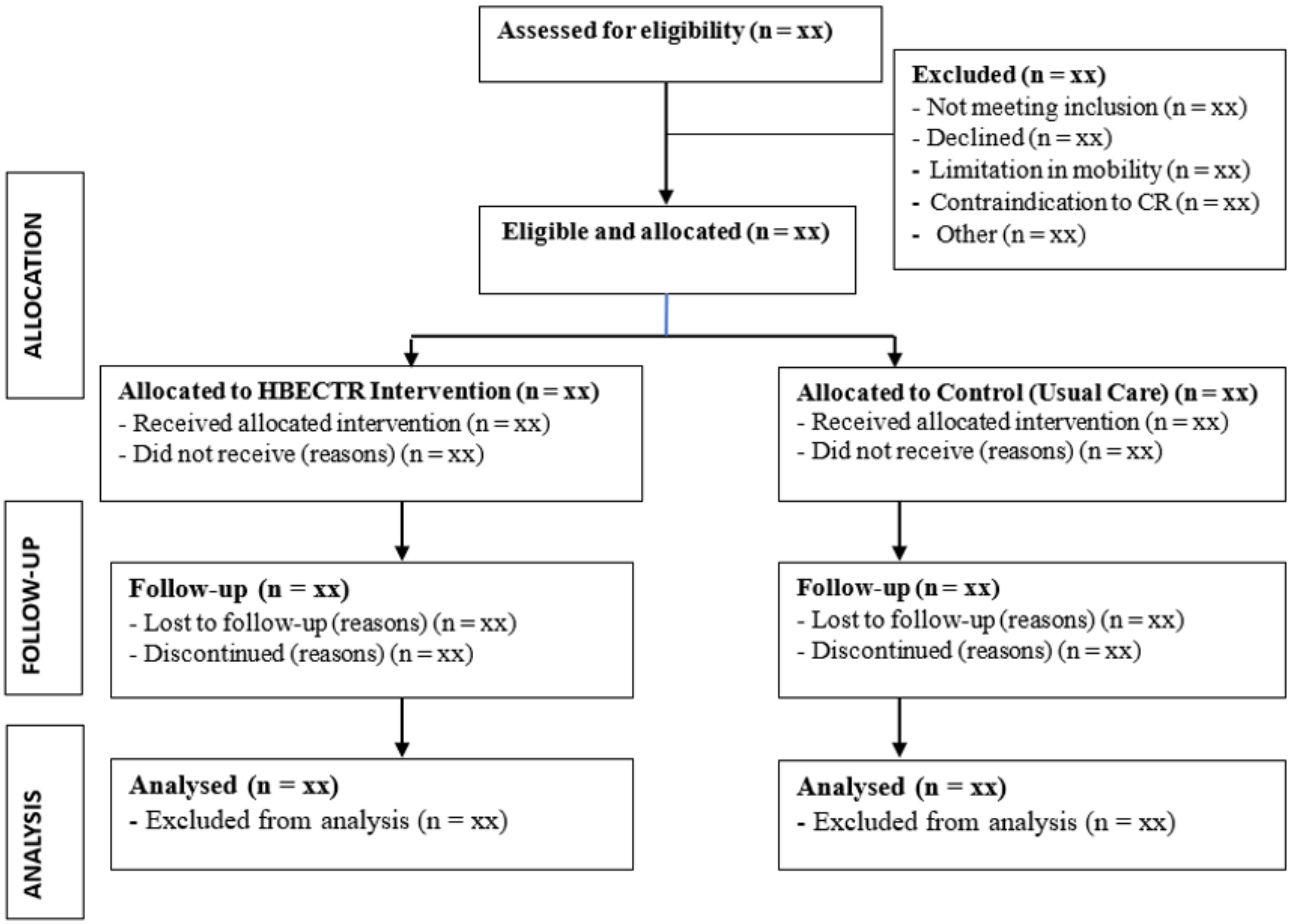

XSL•FO
RenderX

Study Procedures

To assess the efficacy of the HBECTR program, we divided subjects into 2 groups: the intervention group and the control group (Table 1). The CG will receive standard hospital CR or CBCR, while the treatment group will receive HBECTR. We define HBECTR as exercise-based cardiac rehabilitation conducted in a hybrid setting, both in the hospital and at the patient's home, supported by a telemonitoring system.

We will use a telemonitoring system consisting of a smartwatch (CovWatch) produced by PT CovWatch Karya Nusantara, Indonesia, connected via the internet to the web-based app (dashboard), namely "Telerehabilitasi Kardiovaskular (TEKAD)" (App Rehab Cardio) [36]. We developed and conducted usability testing before the dashboard was released (will be reported separately). Prior to the initiation of the study, an accuracy test will be conducted on 10 - 20 patients as part of the CovWatch protocol. A standardized calibration check will be performed at baseline by comparing CovWatch-derived heart rate (HR) with a clinical pulse oximeter at rest and during walking test. Data validity will be monitored through automated system flags for missing data, connectivity loss, and out-of-range physiological values. Mitigation strategies will include repeating measurements, instructing participants on proper device placement (detailed in the guidance book for patient), verifying battery status, and resynchronizing data in case of delayed transmission. All sensor data will be stored with time stamps to identify anomalies and enable manual quality checks.

Before starting the exercise, all patients will undergo a pretest to measure functional capacity using the 6-minute walk test (6MWT) [37]. This pretest is carried out when the patient is still hospitalized or during the first check-up after PCI, according to hospital policy. The cardiologist will prescribe an exercise regime tailored by patients based on this baseline 6MWT, in accordance with the CR guidelines in Indonesia [19]. After the patients have finished all CR programs, they will do the 6MWT for a posttest. The 6MWT for both pretest and posttest will be done at the CR center, Dr. Sardjito General Hospital, Yogyakarta. The intervention of this study will be reported according to the TIDieR (Template for Intervention Description and Replication) [38], in conjunction with the CERT (Consensus on Exercise Reporting Template) [39] and mERA (mHealth Evidence Reporting and Assessment; Checklist 1-4).

Table . Exercise prescription based on frequency, intensity, time, and type (frequency, intensity, time, and type) for the control group and the intervention group.

Dose	CG[a] (CBCR[b])	IG[c] (HBECTR[d])
Frequency	• Standard care included 2 sessions per week at the hospital	• 5 sessions per week, divided into 2 sessions per week at the hospital and 3 sessions per week at home
Intensity	• Moderate	• Moderate
Time	• 45 min • Warm-up: 5 min • Core-exercise: 30 min • Cool-down: 10 min	• 45 min • Warm-up: 5 min • Core-exercise: 30 min • Cool-down: 10 min
Type	• Aerobic (ergo cycle and treadmill)	• Aerobic: ergo cycle and treadmill at the hospital; or walking on flat ground at home
Monitoring	• Hospital: supervised by CR[e] nurse based on vital signs, distance, RPE[f], and %HRR[g]	• Hospital: supervised by CR nurse based on HR[h], BP[i], SaO_2[j], distance, RPE, and %HRR • Home: monitored by smartwatch connected to dashboard (App Rehab Cardio), based on HR, BP, SaO_2, distance, and patients' complaints

[a]CG: control group.

[b]CBCR: center-based cardiac rehabilitation.

[c]IG: intervention group.

[d]HBECTR: home-based exercise cardiac telerehabilitation.

[e]CR: cardiac rehabilitation.

[f]RPE: rate of perceived exertion

[g]HRR: heart rate reserve.

[h]HR: heart rate.

[i]BP: blood pressure.

[j]SaO_2: oxygen saturation.

XSL•FO
RenderX

The Intervention Group

The patients in the treatment group will be prescribed an individually EBCR program at home and in the hospital. The targeted heart rate and walking distance targets will be set individually based on baseline 6MWT. The intensity progression will be evaluated weekly if tolerated. Exercise education and an exercise guidance book will be provided to patients before the program begins. In the hospital, patients do a 45-minute exercise session consisting of a warm-up (5 min), core exercise (30 min) in the form of walking on a treadmill, and cool down (10 min). This exercise is carried out twice a week over 4 weeks in the hospital (a total of 8 sessions) and is directly supervised by a certified cardiac rehab nurse (Table 1).

During exercise at home, subjects will be monitored by a "TEKAD" telemonitoring system. Patients will wear CovWatch that will be connected via the internet to a web-based system (App Rehab Cardio). Prior to the exercise at home, the researcher and the IT team will ensure that the CovWatch is connected to the dashboard App Rehab Cardio to facilitate the real-time monitoring of patients' vital signs during exercise (see Figure 2). Home exercise will be carried out for 45 minutes, consisting of 5 minutes of warm-up (eg, simple stretching or slow walking), 30 minutes of core exercise walking based on the exercise dose and patient's ability, and 10 minutes of cooling down. Vital signs, such as pulse, BP, and oxygen saturation, will be monitored via CovWatch and displayed on the dashboard. Home exercises will be prescribed 3 days a week over 4 weeks (12 sessions in total), with a schedule adjusted by the researchers so as not to coincide with the hospital exercise sessions. Patients will be evaluated and receive weekly exercise reminders directly when they perform exercise at the hospital.

Figure 2. Real-time monitoring in home-based exercise cardiac telerehabilitation (HBECTR). CR: cardiac rehabilitation. App Rehab Cardio can be accessed via the web [36].

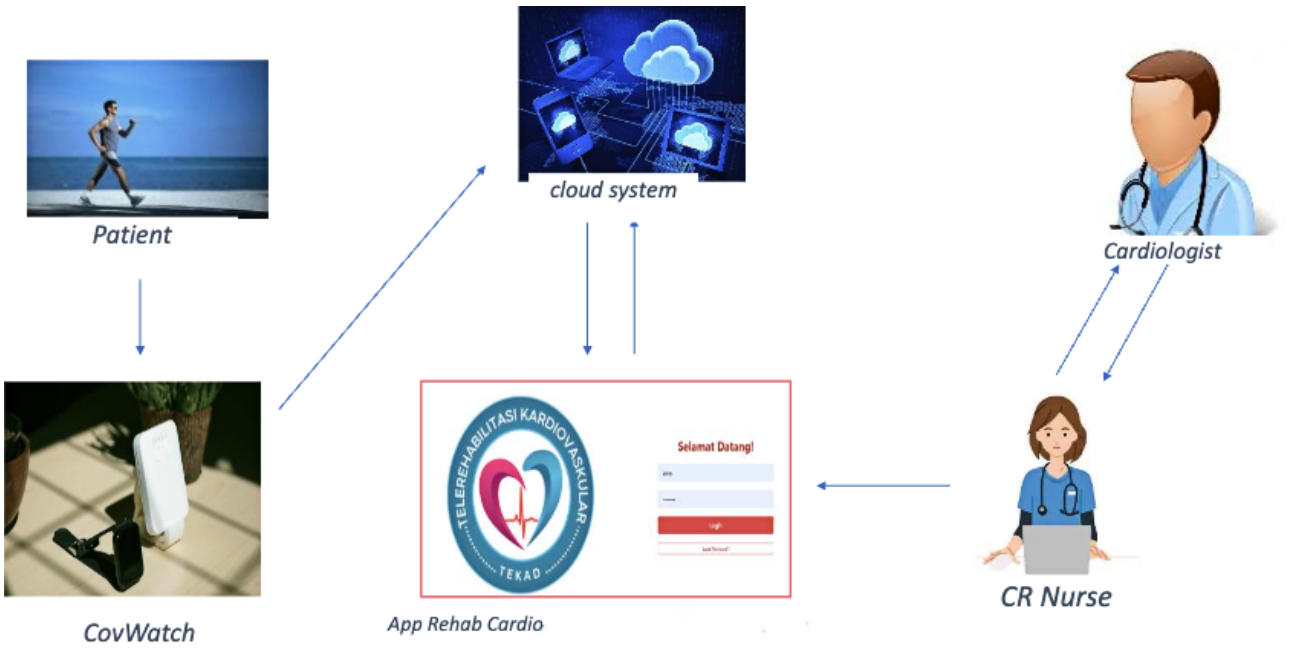

Real-Time Monitoring

During exercise at home, subjects will be monitored by a nurse remotely through telemonitoring (TEKAD). With high-precision sensors, the smartwatch "CovWatch" will continuously track the patient's HR, BP, oxygen saturation level during exercise, and estimated walking distance. The system is connected via the internet to the dashboard (App Rehab Cardio; see Figure 2).

The Control Group

The CG will be enrolled in a cardiac rehabilitation program in the hospital. Patients in this group also will receive exercise education before starting the program. They will be scheduled for 2 sessions per week or 8 sessions in total to do individual exercise at the hospital under a CR nurse's supervision. Based on hospital policy, patients are also suggested to exercise at home independently (without monitoring). Hospital exercises consist of a 5-minute warm-up, 30-minute core exercise (walking on a treadmill), and a 10-minute cool-down (Table 1).

Measurement

In this study, respondents will be asked to fill out the participant characteristics, including age, sex, BMI, and medical history, which were collected at baseline through direct interviews and medical record reviews. Sociodemographic information, such as marital status, educational level, occupation, insurance ownership, and average monthly income, was also gathered. In addition, participants were asked about their smoking history (categorized as current, former, or never smoked) and any past or current illnesses, diseases, or risk factors. Clinical information was obtained, including the patient's medical diagnosis, ejection fraction, prescribed therapies, weight and height, as well as the schedule of heart polyclinic control and cardiac rehabilitation, if applicable.

Adherence to the exercise-based cardiac rehabilitation program as the primary outcome will be measured by recording the number of sessions attended by each participant. Based on a previous study [40], adherence will be calculated as the number of attended sessions divided by the total number of prescribed

XSL•FO
RenderX

sessions, expressed as a percentage, with a minimum of 80% session completion considered adherent. Attendance during in-hospital exercise sessions will be directly monitored by CR nurses and recorded in the CR attendance form, while home-based exercise attendance for IG will be verified through the web-based dashboard system. Adherence will be monitored by smartwatch logs and dashboard analytics. We will apply this comprehensive approach to ensure accurate measurement of each participant's adherence to the prescribed rehabilitation program.

Functional capacity will be assessed using the 6MWT to measure the total distance (in meters) during 6 minutes test (6-minute walk distance) [19,37]. In addition, physiological data such as HR, BP, and oxygen saturation will be continuously monitored using the telemonitoring system (see Figure 2), ensuring the safety and accuracy of HBECTR. To provide a more comprehensive assessment of functional improvement, peak oxygen uptake (VO_2 peak) or metabolic equivalents will be estimated based on the distance covered in the 6MWT, using established predictive equations [41,42]. These measurements collectively will provide robust data on the participants' functional performance following the intervention.

The primary and secondary outcomes will be measured after participants have finished the intervention (4-wk follow-up). We selected a 4-week follow-up because this is a pilot feasibility study designed to assess the early feasibility, safety, and adherence of HBECTR. Short-duration assessments are well established and appropriate for pilot and feasibility trials, particularly in digital and cardiac telerehabilitation research [43,44]. Besides, the guideline of CR in Indonesia based on the Indonesia Heart Association also recommends conducting an evaluation after 1 month (4 wk) of phase II CR [19]. Furthermore, early improvements in functional capacity or cardiopulmonary endurance often occur within the first month [5,45], making a 4-week evaluation appropriate for assessing initial signal of effect.

We consider extending the follow-up (3 - 6 mo follow-up) for our next trial to ensure the sustainability and efficacy of this program. In addition to feasibility and functional outcomes, we will record hospital readmissions during intervention, and patients will be followed up for 3 - 6 months. The ability of patients to do their daily activities will be monitored and recorded using a self-report.

In addition, the implementation outcomes will be assessed using the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) framework [46]. We will report recruitment and representativeness (Reach), adherence and safety (Effectiveness), provider uptake (Adoption), fidelity and dose (Implementation), and short-term intention to continue (Maintenance). Qualitative interviews will be added to explore participants' experiences and barriers during intervention. Furthermore, we will calculate cost-effectiveness by determining the cost per patient, application development, wearable device production, and internet provision.

Statistical Analysis Planning

All data will be analyzed using the IBM SPSS software (version 29). We will use MI for missing data. MI is a widely used technique for handling missing data in various fields, including medical, social, and ecological research.

Data analysis will be performed primarily on participants who complete both baseline and postintervention assessments, regardless of their adherence level during the intervention period (*completer analysis*). Participants who withdraw or are lost to follow-up before the posttest will be excluded from outcome analysis but will be included in the descriptive analysis of recruitment, adherence, and retention rates. This approach is considered appropriate for pilot quasi-experimental studies, where the primary aim is to assess feasibility, acceptability, and signal of effect rather than to test definitive efficacy [31,35].

To minimize potential bias due to nonrandom assignment, baseline characteristics between the groups will be compared using 2-tailed independent *t* tests or chi-square tests. In addition, propensity score adjustment will be applied in exploratory models to control for potential confounding variables (eg, age, sex, baseline functional capacity, comorbidity), following recommendations for quasi-experimental design analysis [33].

All statistical analyses and reporting will follow the principles outlined in the CONSORT-EHEALTH 2011 (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) [47] (Checklist 3). Descriptive statistics will summarize feasibility outcomes (recruitment, adherence, completion rates, safety events), while inferential tests (2-tailed paired *t* tests or analysis of covariance) will be used to explore within- and between-group changes (including means and SDs) in functional capacity and adherence measures. Statistical significance will be interpreted cautiously, given the pilot nature and limited sample size of the study [31,35].

To analyze the effects of HBECTR on patient adherence as the primary outcome and functional capacity as the secondary outcome, there is no formal multiplicity adjustment planned. The data analysis based on normality assumption will be tested using the Shapiro-Wilk test. If the data comply with the normal assumption, we will use the Student *t* test, and vice versa, nonparametric tests will be used if assumptions are violated. In all statistical tests, $P<.05$ (95% CI) will be considered statistically significant. This structured approach will ensure that statistical methods are properly tailored to the data, supporting robust and meaningful interpretation of the results. The statistical analysis plan has been registered within the protocol to the institutional ethics committee board with the approval number KE/FK/0850/EC/2024.

Safety and Monitoring

Patient safety during exercise is crucial and must be taken into consideration. Clinical supervision and efforts to minimize complications during the exercise program must be implemented to ensure the safety of CR [19]. In this study, the efforts to supervise and maintain patient safety during exercise will involve several aspects related to the program and patients. All patients must obtain a referral from a cardiologist and will

XSL•FO
RenderX

undergo an assessment before, during, and after the program. During the recruitment of potential subjects, screening will be conducted, and only patients with a low and moderate risk will be included. We will also educate patients about the signs and symptoms of cardiovascular emergencies, both during exercise at the hospital and home, such as chest pain, irregular heartbeat, weight gain, and shortness of breath. A guidebook will be provided to patients, which details how to perform exercises safely at home and how to use the smartwatch for monitoring. The guidebook also includes the WhatsApp numbers of the researchers and cardiac rehab team and emergency contact details.

Regarding the program, risk stratification screening will be conducted prior to exercise, and patients will be supervised and monitored during exercise. Advances in technology have enabled the remote monitoring of patients, especially when exercise is performed at home or outside the hospital. The safety of remotely delivered CR has been demonstrated by the low long-term rates of reported exercise−related adverse events (AEs), serious adverse events (SAEs), and rehospitalization incidences [48]. Hence, this study will entail the implementation of a telemonitoring system, utilizing the "TEKAD" dashboard for monitoring patients' vital signs during exercise to ensure the safety of the program (Figure 2).

Prior to participation, all patients will receive comprehensive explanations of the program's procedures, safety protocols, and communication flow, which will be reinforced again at the start of the intervention. Participants will receive detailed instructions on the prescribed exercise regimen to be performed both at the hospital and at home. They will be guided on the appropriate type, intensity, and duration of exercise and will be encouraged to report any symptoms or discomfort experienced during exercise. If patients experience any adverse symptoms, such as chest pain, shortness of breath, unusual fatigue, or other discomfort, they will be instructed to immediately stop exercising and rest.

During exercise at home, patients will be able to contact the cardiac rehabilitation team via WhatsApp in case of any complaints or concerns. Conversely, if the team detects abnormal changes in vital signs (eg, HR, BP, or oxygen saturation) during remote monitoring, they will proactively contact the patient to inquire about their condition. Then, if necessary, patients will be advised to stop the exercise and take adequate rest.

Furthermore, the risk of AEs and SAEs in EBCR is low. However, it is increased during vigorous intensity exercises relative to rest [49] and intensive exercise training (≥7 times per week) [50]. Therefore, this study will implement the moderate intensity exercise within 5 times per week. Besides, the exercise prescription will be individualized based on the 6MWT baseline, following the CR guideline in Indonesia [19]. All AEs or SAEs during the study will be reported to the cardiologist who will be supervising this program. These also will be documented in the case report form and reported to the ethical board.

Ethical Considerations

This study protocol adhered to the Declaration of Helsinki and received approval from the institutional review board of the Faculty of Medicine, Public Health, and Nursing, Universitas Gadjah Mada (approval number KE/FK/0850/EC/2024). All participants signed an informed consent form before the study commenced, ensuring their voluntary participation and the protection of their confidentiality throughout the study. The subjects in both groups who will attend on-site rehabilitation sessions at the hospital will receive financial compensation to reimburse transportation costs. This compensation will be provided to support the travel costs incurred during study participation and will not intend to constitute an inducement to participate. The amount and mechanism of reimbursement have been reviewed and approved by the institutional ethics committee.

Emphasizing that the health information of patients is sensitive and the leakage of the information could bring both ethical and economical risks [51], all data obtained will be carefully documented and stored in an encrypted system, and patient confidentiality will be maintained following research ethics standards. Furthermore, the app or dashboard is available only to the researcher team and is password-protected.

This trial was retrospectively registered to the ANZCTR with registry ID: ACTRN12625001400459 (IRRIDDERR1-10.2196/81067). The first enrolment was started at August 14, 2024, while the trial registration was submitted on October 15, 2025 The trial was registered retrospectively due to an administrative and tehnical problem related to the app development.

Results

The recruitment of the subject started from August 2024. This study is being funded since February 2025. During the submission, we have recruited 63 subjects for both IG and CG. The intervention for both groups finished in June 2025, and the data are being analyzed. It is expected that the data analysis will be finished in November 2025 and that the results will be published in December 2025.

Discussion

The low adherence of patients to participate in CR programs, particularly after PCI, remains a significant challenge in Indonesia and other countries. Despite class 1 recommendations for CR as part of secondary prevention, CBCR is often underutilized due to various factors, including transportation difficulties, limited availability, and low patient motivation. This study aims to address these challenges by implementing an HBECTR program to enhance patients' adherence to CR.

This HBECTR intervention protocol offers an innovative approach to CR by leveraging technology to provide continuous monitoring, guidance, and feedback to patients. Previous studies have indicated that home-based cardiac telerehabilitation can be as effective as traditional hospital−based programs in improving functional capacity, adherence, and quality of life [15]. However, there has been limited evidence on home-based

XSL•FO
RenderX

cardiac telerehabilitation implementation in Indonesia, particularly in post-PCI patients.

By integrating a telerehabilitation model in HBCR, this study may address several barriers to traditional CR, including geographical challenges and logistical issues. Digital technologies have revolutionized CR, offering flexible and novel approaches to care. The integration of digital health technologies and artificial intelligence in remote CR is transforming traditional paradigms, providing real-time access to health data, and enhancing patient self-management. Mobile and digital CR models, including synchronous or real-time digital interventions, are addressing accessibility barriers and promoting equity in health care delivery [23,27].

Despite the benefit or positive aspect of HBECTR, there are several challenges such as digital literacy, data privacy, and security that must be addressed to ensure the inclusive implementation of CR. Moreover, the shift toward digital CR raises concerns about cost, safety, and potential depersonalization of therapeutic relationships [27].

Furthermore, several potential limitations should be acknowledged in this quasi-experimental study. First, the intervention may yield null or minimal effects, particularly on adherence to cardiac rehabilitation and functional capacity, due to variability in participant engagement, baseline fitness levels, or external barriers such as comorbidities or limited motivation. Second, feasibility challenges are anticipated, including variable access to technology, differences in digital literacy, difficulties in remotely monitoring safety, and potential missing data from nonadherence or technical issues. Using a wearable device (smartwatch) is preferable to most patients; however, it has challenges in its implementation and needs further exploration to gain a deeper understanding of patients' needs and preferences [52]. Finally, resource limitations may constrain implementation, including limited personnel for supervision, insufficient devices for remote monitoring, costs associated with web-based platforms, and the need for staff training and scheduling coordination.

In addition, this study also has limitations, including its modest sample size, and the short follow-up duration limits its generalizability to the broader population. These are intentional features of a pilot feasibility trial that focuses on testing feasibility, safety, and preliminary efficacy signals rather than providing conclusive evidence of effectiveness [32]. Clinical trials of physical activity and rehabilitation interventions are challenging; therefore, researchers can use pilot and feasibility studies to enhance the rigor of future trials [35].

Despite these challenges, the study provides valuable insights into the practicality and potential impact of HBECTR, informing the design of larger, fully powered trials. Future research is needed to enhance the study in multicenters with multiple outcomes. Longer-term effects, sustainability, and cost-effectiveness should be evaluated in a subsequent definitive trial.

Acknowledgments

The authors would like to express their sincere gratitude to all patients who participated in this study, the health care professionals who supported the implementation of the home-based exercise telerehabilitation program, and all the hospital staff of Dr. Sardjito General Hospital Yogyakarta for their assistance in the research process. The authors also acknowledge the support of the Indonesia Endowment Fund for Education Agency for the scholarship.

Funding

This research was supported by the "Dana Masyarakat (DAMAS)" Grant from Faculty of Medicine, Public Health, and Nursing Universitas Gadjah Mada, Indonesia (number: 3857/UN1/KU.1/PP/PT.01.01/2025). The funder is not involved in the development of the protocol.

Authors' Contributions

Conceptualization: WW.
Funding acquisition: BYS.
Methodology: ABH, BYS, SS, WW.
Project administration: WW.
Resources: ABH, BYS, SS, WW.
Software: WW.
Supervision: SS, ABH, BYS.
Writing – original draft: WW.
Writing – review & editing: ABH, BYS, SS, WW.

Conflicts of Interest

None declared.

Checklist 1
TIDIeR (Template for Intervention Description and Replication) checklist.

XSL•FO
RenderX

[PDF File, 115 KB - resprot_v15i1e81067_app1.pdf]

Checklist 2
CERT (Consensus on Exercise Reporting Template) checklist.
[PDF File, 820 KB - resprot_v15i1e81067_app2.pdf]

Checklist 3
CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online TeleHealth) checklist.
[PDF File, 1096 KB - resprot_v15i1e81067_app3.pdf]

Checklist 4
mERA (mobile health [mHealth] Evidence Reporting and Assessment) checklist.
[PDF File, 112 KB - resprot_v15i1e81067_app4.pdf]

References

1. Cardiovascular diseases (CVDs). World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2025-12-13]
2. Cardiovascular insight for Indonesia. World Heart Observatory. 2025. URL: https://world-heart-federation.org/world-heart-observatory/countries/indonesia/ [accessed 2025-09-23]
3. Woodgate J, Brawley LR. Self-efficacy for exercise in cardiac rehabilitation: review and recommendations. J Health Psychol 2008 Apr;13(3):366-387. [doi: 10.1177/1359105307088141] [Medline: 18420771]
4. Simon M, Korn K, Cho L, Blackburn GG, Raymond C. Cardiac rehabilitation: a class 1 recommendation. Cleve Clin J Med 2018 Jul;85(7):551-558. [doi: 10.3949/ccjm.85a.17037] [Medline: 30004380]
5. Thomas RJ, Beatty AL, Beckie TM, et al. Home-based cardiac rehabilitation: a scientific statement from the American Association of Cardiovascular and Pulmonary Rehabilitation, the American Heart Association, and the American College of Cardiology. J Am Coll Cardiol 2019 Jul 9;74(1):133-153. [doi: 10.1016/j.jacc.2019.03.008] [Medline: 31097258]
6. Chindhy S, Taub PR, Lavie CJ, Shen J. Current challenges in cardiac rehabilitation: strategies to overcome social factors and attendance barriers. Expert Rev Cardiovasc Ther 2020 Nov;18(11):777-789. [doi: 10.1080/14779072.2020.1816464] [Medline: 32885702]
7. Poortaghi S, Baghernia A, Golzari SEJ, Safayian A, Atri SB. The effect of home-based cardiac rehabilitation program on self efficacy of patients referred to cardiac rehabilitation center. BMC Res Notes 2013 Jul 23;6(1):4-7. [doi: 10.1186/1756-0500-6-287] [Medline: 23880148]
8. Ritchey MD, Maresh S, McNeely J, et al. Tracking cardiac rehabilitation participation and completion among Medicare beneficiaries to inform the efforts of a national initiative. Circ Cardiovasc Qual Outcomes 2020 Jan;13(1):1-11. [doi: 10.1161/CIRCOUTCOMES.119.005902] [Medline: 31931615]
9. Lima AP, Nascimento IO, Oliveira ACA, Martins THS, Pereira DAG, Britto RR. Home-based cardiac rehabilitation in Brazil's public health care: protocol for a randomized controlled trial. JMIR Res Protoc 2019;8(11):e13901. [doi: 10.2196/13901] [Medline: 31697246]
10. Bakhshayeh S, Sarbaz M, Kimiafar K, Vakilian F, Eslami S. Barriers to participation in center-based cardiac rehabilitation programs and patients' attitude toward home-based cardiac rehabilitation programs. Physiother Theory Pract 2021 Jan;37(1):158-168. [doi: 10.1080/09593985.2019.1620388] [Medline: 31155986]
11. Beatty AL, Brown TM, Corbett M, et al. Million Hearts cardiac rehabilitation think tank: accelerating new care models. Circ Cardiovasc Qual Outcomes 2021 Oct;14(10):e008215. [doi: 10.1161/CIRCOUTCOMES.121.008215] [Medline: 34587751]
12. Wongvibulsin S, Habeos EE, Huynh PP, et al. Digital health interventions for cardiac rehabilitation: systematic literature review. J Med Internet Res 2021 Feb 8;23(2):e18773. [doi: 10.2196/18773] [Medline: 33555259]
13. Chatzitofis A, Monaghan D, Mitchell E, et al. HeartHealth: a cardiovascular disease home-based rehabilitation system. Procedia Comput Sci 2015;63:340-347. [doi: 10.1016/j.procs.2015.08.352]
14. Prabhu NV, Maiya AG, Prabhu NS. Impact of cardiac rehabilitation on functional capacity and physical activity after coronary revascularization: a scientific review. Cardiol Res Pract 2020;2020:1236968. [doi: 10.1155/2020/1236968] [Medline: 32274208]
15. Zhong W, Fu C, Xu L, et al. Effects of home-based cardiac telerehabilitation programs in patients undergoing percutaneous coronary intervention: a systematic review and meta-analysis. BMC Cardiovasc Disord 2023 Feb 22;23(1):101. [doi: 10.1186/s12872-023-03120-2] [Medline: 36814188]
16. Achttien RJ, Staal JB, van der Voort S, et al. Exercise-based cardiac rehabilitation in patients with coronary heart disease: a practice guideline. Neth Heart J 2013 Oct;21(10):429-438. [doi: 10.1007/s12471-013-0467-y] [Medline: 23975619]

XSL•FO
RenderX

17. Hartopo AB, Arso IA, Ambari AM, Dwiputra B, Radi B. Exercise-based cardiac rehabilitation adaptation protocol during Covid-19 pandemic achieved similar results as compared to non-pandemic usual practice: a single center experience. J Med Sci 2022;54(4):369-375. [doi: 10.19106/JMedSci005404202207]
18. Chowdhury M, Heald FA, Turk-Adawi K, et al. Availability and delivery of cardiac rehabilitation in South-East Asia: How does it compare globally? WHO South East Asia J Public Health 2021;10(2):57-65. [doi: 10.4103/WHO-SEAJPH.WHO-SEAJPH_62_21] [Medline: 35532596]
19. Radi B, Tiksnadi BB, Dwiputra B, Sarvasti D, Ambari AM, editors. Panduan rehabilitasi kardiovaskular (guidelines for cardiovascular rehabilitation). : Perhimpunan Dokter Spesialis Kardiovaskular Indonesia (PERKI) (Indonesian Society of Cardiology); 2019 URL: https://www.inaheart.org/wp-content/uploads/2025/10/Panduan-Rehabilitasi-Kardiovaskular.pdf [accessed 2025-12-13]
20. Stefanakis M, Batalik L, Antoniou V, Pepera G. Safety of home-based cardiac rehabilitation: a systematic review. Heart Lung 2022;55:117-126. [doi: 10.1016/j.hrtlng.2022.04.016] [Medline: 35533492]
21. Anderson L, Sharp GA, Norton RJ, et al. Home-based versus centre-based cardiac rehabilitation. Cochrane Database Syst Rev 2017 Jun 30;6(6):CD007130. [doi: 10.1002/14651858.CD007130.pub4] [Medline: 28665511]
22. Schopfer DW, Whooley MA, Allsup K, et al. Effects of home-based cardiac rehabilitation on time to enrollment and functional status in patients with ischemic heart disease. J Am Heart Assoc 2020 Oct 20;9(19):e016456. [doi: 10.1161/JAHA.120.016456] [Medline: 32954885]
23. Zhang S, Wang Y, Wu J, Ma C, Meng X. Effectiveness of smartwatch device on adherence to home-based cardiac rehabilitation in patients with coronary heart disease: randomized controlled trial. JMIR Mhealth Uhealth 2025 Sep 18;13:e70848. [doi: 10.2196/70848] [Medline: 40966536]
24. Liu J, Wang L, Fang H, Wang X, Wu L, Zhang J. Home-based cardiac rehabilitation: a review of bibliometric studies and visual analysis of CiteSpace (2012–2021). Medicine (Abingdon) 2022;101(49):e31788. [doi: 10.1097/MD.0000000000031788]
25. Antoniou V, Davos CH, Kapreli E, Batalik L, Panagiotakos DB, Pepera G. Effectiveness of home-based cardiac rehabilitation, using wearable sensors, as a multicomponent, cutting-edge intervention: a systematic review and meta-analysis. J Clin Med 2022 Jun 29;11(13):3772. [doi: 10.3390/jcm11133772] [Medline: 35807055]
26. Harris AD, McGregor JC, Perencevich EN, et al. The use and interpretation of quasi-experimental studies in medical informatics. J Am Med Inform Assoc 2006;13(1):16-23. [doi: 10.1197/jamia.M1749] [Medline: 16221933]
27. Pepera G, Antoniou V, Su JJ, Lin R, Batalik L. Comprehensive and personalized approach is a critical area for developing remote cardiac rehabilitation programs. World J Clin Cases 2024 Apr 26;12(12):2009-2015. [doi: 10.12998/wjcc.v12.i12.2009] [Medline: 38680265]
28. Grace SL, Midence L, Oh P, et al. Cardiac rehabilitation program adherence and functional capacity among women: a randomized controlled trial. Mayo Clin Proc 2016 Feb;91(2):140-148. [doi: 10.1016/j.mayocp.2015.10.021] [Medline: 26682921]
29. Gremeaux V, Troisgros O, Benaïm S, et al. Determining the minimal clinically important difference for the six-minute walk test and the 200-meter fast-walk test during cardiac rehabilitation program in coronary artery disease patients after acute coronary syndrome. Arch Phys Med Rehabil 2011 Apr;92(4):611-619. [doi: 10.1016/j.apmr.2010.11.023] [Medline: 21440707]
30. Keteyian SJ, Grimshaw C, Ehrman JK, et al. The iATTEND trial: a trial comparing hybrid versus standard cardiac rehabilitation. Am J Cardiol 2024 Jun 15;221:94-101. [doi: 10.1016/j.amjcard.2024.04.034] [Medline: 38670326]
31. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005 Oct;4(4):287-291. [doi: 10.1002/pst.185]
32. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 6;10:1. [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
33. Tilquin C. Patient classification does work. Dimens Health Serv 1976 Jan;53(1):12-13. [Medline: 1310]
34. White H, Sabarwal S. Quasi-experimental design and methods. : UNICEF Office of Research; 2014 URL: https://www.betterevaluation.org/sites/default/files/Quasi-Experimental_Design_and_Methods_ENG.pdf [accessed 2025-12-13]
35. El-Kotob R, Giangregorio LM. Pilot and feasibility studies in exercise, physical activity, or rehabilitation research. Pilot Feasibility Stud 2018;4(1):137. [doi: 10.1186/s40814-018-0326-0] [Medline: 30123527]
36. App rehab cardio. URL: https://tekad.covwatch.net/ [accessed 2025-12-31]
37. Arena R, Myers J, Williams MA, et al. Assessment of functional capacity in clinical and research settings: a scientific statement from the American Heart Association Committee on Exercise, Rehabilitation, and Prevention of the Council on Clinical Cardiology and the Council on Cardiovascular Nursing. Circulation 2007 Jul 17;116(3):329-343. [doi: 10.1161/CIRCULATIONAHA.106.184461] [Medline: 17576872]
38. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 7;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
39. Slade SC, Dionne CE, Underwood M, Buchbinder R. Consensus on Exercise Reporting Template (CERT): explanation and elaboration statement. Br J Sports Med 2016 Dec;50(23):1428-1437. [doi: 10.1136/bjsports-2016-096651] [Medline: 27707738]

XSL•FO
RenderX

40. McGregor G, Powell R, Begg B, et al. High-intensity interval training in cardiac rehabilitation: a multi-centre randomized controlled trial. Eur J Prev Cardiol 2023 Jul 12;30(9):745-755. [doi: 10.1093/eurjpc/zwad039] [Medline: 36753063]
41. Cahalin LP, Mathier MA, Semigran MJ, Dec GW, DiSalvo TG. The six-minute walk test predicts peak oxygen uptake and survival in patients with advanced heart failure. Chest 1996 Aug;110(2):325-332. [doi: 10.1378/chest.110.2.325] [Medline: 8697828]
42. Tiksnadi BB, Ambari AM, Adriana M. Uji Jalan 6 Menit (UJ6M) pada Pasien Pasca Sindrom Koroner Akut (Six-minute walk test in patients after acute coronary syndrome). Indones J Cardiol 2019;40(1):222-231. [doi: 10.30701/ijc.v40i1.913]
43. Mastorci F, Lazzeri MFL, Ait-Ali L, et al. Home-based intervention tool for cardiac telerehabilitation: protocol for a controlled trial. JMIR Res Protoc 2025 Jan 22;14:e47951. [doi: 10.2196/47951] [Medline: 39841521]
44. Varnfield M, Karunanithi M, Lee CK, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. Heart 2014 Nov;100(22):1770-1779. [doi: 10.1136/heartjnl-2014-305783] [Medline: 24973083]
45. Zhang L, Ge Y, Zhao W, et al. A 4-week mobile app-based telerehabilitation program vs conventional in-person rehabilitation in older adults with sarcopenia: randomized controlled trial. J Med Internet Res 2025 Jan 24;27:e67846. [doi: 10.2196/67846] [Medline: 39854716]
46. Holtrop JS, Estabrooks PA, Gaglio B, et al. Understanding and applying the RE-AIM framework: clarifications and resources. J Clin Transl Sci 2021;5(1):e126. [doi: 10.1017/cts.2021.789] [Medline: 34367671]
47. Eysenbach G, CONSORT-EHEALTH G. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126. [doi: 10.2196/jmir.1923] [Medline: 22209829]
48. Antoniou V, Kapreli E, Davos CH, Batalik L, Pepera G. Safety and long-term outcomes of remote cardiac rehabilitation in coronary heart disease patients: a systematic review. Digit Health 2024;10:20552076241237661. [doi: 10.1177/20552076241237661] [Medline: 38533308]
49. Guy JM, Wilson M, Schnell F, et al. Incidence of major adverse cardiac events in men wishing to continue competitive sport following percutaneous coronary intervention. Arch Cardiovasc Dis 2019 Apr;112(4):226-233. [doi: 10.1016/j.acvd.2018.11.008] [Medline: 30612894]
50. Mons U, Hahmann H, Brenner H. A reverse J-shaped association of leisure time physical activity with prognosis in patients with stable coronary heart disease: evidence from a large cohort with repeated measurements. Heart 2014 Jul;100(13):1043-1049. [doi: 10.1136/heartjnl-2013-305242] [Medline: 24829374]
51. Su JJ, Chan MHS, Ghisi GDM, et al. Real-world mobile health implementation and patient safety: multicenter qualitative study. J Med Internet Res 2025 Apr 29;27:e71086. [doi: 10.2196/71086] [Medline: 40299494]
52. Sarvestan J, Baker KF, Del Din S. Exploring the effect of sampling frequency on real-world mobility, sedentary behaviour, physical activity and sleep outcomes measured with wearable devices in rheumatoid arthritis: feasibility, usability and practical considerations. Bioengineering (Basel) 2024 Dec 28;12(1):18. [doi: 10.3390/bioengineering12010018] [Medline: 39851290]

Abbreviations

6MWT: 6-minute walk test
AE: adverse event
BP: blood pressure
BPJS: Badan Penyelenggara Jaminan Sosial
CBCR: center-based cardiac rehabilitation
CERT: Consensus on Exercise Reporting Template
CG: control group
CONSORT: Consolidated Standards of Reporting Trials
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth
CR: cardiac rehabilitation
CVD: cardiovascular disease
EBCR: exercise-based cardiac rehabilitation
HBCR: home-based cardiac rehabilitation
HBECTR: home-based exercise cardiac telerehabilitation
HR: heart rate
IG: intervention group
mERA: mHealth Evidence Reporting and Assessment
MI: multiple imputation
PCI: percutaneous coronary intervention
PSM: propensity score matching

XSL•FO
RenderX

RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
SAE: severe adverse event
TEKAD: Telerehabilitasi Kardiovaskular
TIDieR: Template for Intervention Description and Replication

Edited by J Sarvestan; submitted 22.Jul.2025; peer-reviewed by G Pepera; accepted 02.Dec.2025; published 16.Jan.2026.

Please cite as:
Wantiyah W, Hartopo AB, Setiyarini S, Setianto BY
Home-Based Exercise Cardiac Telerehabilitation on Adherence and Functional Capacity for Patients After Percutaneous Coronary Intervention in Indonesia: Protocol for a Quasi-Experimental Study
JMIR Res Protoc 2026;15:e81067
URL: https://www.researchprotocols.org/2026/1/e81067
doi:10.2196/81067

© Wantiyah Wantiyah, Anggoro Budi Hartopo, Sri Setiyarini, Budi Yuli Setianto. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Acceptability of Using Virtual Reality to Promote Physical Activity Among Latino, Latina, and Latine Adults With Cardiometabolic Risk Factors and Obesity in Underresourced Settings: Protocol for a Qualitative Focus Group Study

Desiree R Acosta[1,2], MPH; Leslie Aguilar-Hernandez[2], BS; Gael Perez[2], BS; Iris Y Guzman-Ruiz[2], MPH, PhD; Josyel M Castellon[2], MPH; Jailene Cruz[2], BS; Liana Gutierrez[2], BS; Paulina Monteon-Garcia[2], MPH; Vanessa N Torres[2], MPH, PhD; O Kenrik Duru[3], MS, MD; Yelba Castellon-Lopez[2], MS, MD

[1]Program in Medical Education-Leadership and Advocacy (PRIME-LA), David Geffen School of Medicine at UCLA, Los Angeles, CA, United States

[2]Cancer Research Center for Health Equity, Cedars-Sinai Medical Center, 8700 Beverly Blvd, Los Angeles, CA, United States

[3]General Internal Medicine-Health Services Research, David Geffen School of Medicine at UCLA, Los Angeles, CA, United States

Corresponding Author:
Yelba Castellon-Lopez, MS, MD
Cancer Research Center for Health Equity, Cedars-Sinai Medical Center, 8700 Beverly Blvd, Los Angeles, CA, United States

Abstract

Background: Obesity represents a significant public health challenge in the United States, particularly among Latino, Latina, and Latine communities and those in underresourced settings. Virtual reality (VR) is a new and innovative technology that can promote physical activity and has the potential to overcome some structural barriers. However, there are few studies that explore the acceptability of using this new technology among high-risk groups in underresourced settings.

Objective: We outline a community-informed protocol for conducting focus groups with Latino, Latina, and Latine adults who have cardiometabolic risk factors and obesity residing in underresourced communities. The focus groups will assess the acceptability of a culturally aligned VR program to promote physical activity.

Methods: Using a community-engaged approach informed by community health workers and a community advisory board, we delivered an immersive VR dance experience to Latino, Latina, and Latine adult participants with cardiometabolic risk factors and obesity. Following the VR experience, we conducted semistructured focus group interviews to assess acceptability guided by the theoretical framework of acceptability. Data collection included a baseline demographic survey and focus group discussions to evaluate participant experiences and the VR program's acceptability.

Results: As of April 2025, we have completed 7 focus groups with 44 participants across 3 age groups: 18 to 29 years, 30 to 49 years, and 50 to 70 years. Data collection was completed in May 2025, and study findings are expected to be published in February 2026.

Conclusions: The fully analyzed data from this study will offer insights into leveraging VR as an innovative tool for promoting physical activity in underserved populations, contributing to the broader literature on digital health equity and obesity prevention.

International Registered Report Identifier (IRRID): DERR1-10.2196/80534

(JMIR Res Protoc 2026;15:e80534) doi:10.2196/80534

KEYWORDS

virtual reality; physical activity; dance; acceptability; health equity; Latino; Latina; Latine

Introduction

Obesity is a major health concern in the United States as it significantly increases the risk of various chronic diseases, including heart disease, stroke, type 2 diabetes, and cancer [1]. Data from the National Health and Nutrition Examination Survey show that 40.3% of adults in the United States were classified as obese from August 2021 to August 2023 [2]. Obesity prevalence varies significantly based on race and ethnicity, with higher rates among Latino, Latina, and Latine adults than among White adults [3]. Estimates from 2017 to 2018 show that 44.8% of Latino, Latina, and Latine adults and 50.4% of Mexican American adults were classified as obese compared to 42.4% of non-Hispanic White adults [4]. Furthermore, the prevalence of obesity is higher among American adults with lower educational attainment and lower household income—with the lowest prevalence observed among those with a bachelor's degree or higher [2,5,6]. Latino, Latina,

XSL•FO
RenderX

and Latine adults experience a higher prevalence of obesity, which is accompanied by a disproportionate burden of obesity-related chronic diseases such as type 2 diabetes and certain cancers linked to obesity [4]. As Latino, Latina, and Latine adults represent one of the fastest-growing ethnic minority groups, and currently the largest ethnic minority group, in the United States, there is a pressing need for population-specific public health strategies to mitigate the health impacts of obesity-related diseases [7].

Physical inactivity and sedentary lifestyles are key contributors to obesity [8]. Between 2017 and 2020, only 25.3% of adults in the United States reported being physically inactive outside of work [9]. Among Latino, Latina, and Latine adults, this figure was even higher, with 31.7% reporting no physical activity outside of work, resulting in an increased risk of obesity-related chronic diseases [9]. Racial and ethnic disparities in physical activity may be influenced by structural barriers, including lack of access to safe spaces for exercise, walkable neighborhoods, and sufficient time [9,10]. These findings underscore the urgent need for targeted, culturally responsive interventions that can overcome barriers and address obesity, physical inactivity, and sedentary lifestyles within the Latino, Latina, and Latine community [11]. This population encounters a unique combination of cultural, environmental, and structural challenges that standard obesity prevention and physical activity interventions frequently overlook. Language and cultural norms related to diet, family roles, and food traditions can influence preferences and behaviors. Unfortunately, many interventions are not tailored to address these important elements [4,11]. A systematic review of culturally adapted interventions revealed that incorporating bilingual or bicultural staff, involving family, and implementing community-based formats showed better engagement and outcomes among Latino, Latina, and Latine participants compared to nonadapted interventions [12,13]. These disparities and contextual barriers highlight the need for culturally responsive interventions tailored to Latino, Latina, and Latine adults to enhance acceptability, engagement, and effectiveness.

Therapeutic virtual reality (VR) is gaining popularity as a health-promoting strategy to enhance physical activity and exercise [14]. Through a head-mounted display and a close-proximity screen, VR can transport users into a lifelike, 3D experience [15,16]. A VR exergame is an interactive exercise-based game that uses VR to immerse users in a movement-driven experience. By integrating motion tracking and gamification, VR exergames encourage physical activity through engaging scenarios such as sports, dancing, or fitness challenges. These games have been shown to enhance intrinsic motivation, self-efficacy, and physical activity enjoyment and adherence in adolescents and adults [17-21]. Dance- and music-based exergames have been highly rated [22] and show promise in promoting physical activity [23]. Furthermore, qualitative studies have demonstrated that older adults often prefer exergames and dance-based VR experiences over traditional, less interactive physical activity programs, citing greater motivation and perceived benefits (eg, cognitive and psychological), in both general and underserved settings [22,24,25]. VR-based exergaming has been shown to target key drivers of long-term physical activity engagement. For example, a review of 29 studies concluded that exergaming not only enhances enjoyment but can also elicit moderate to vigorous exercise intensity and reduce perceived exertion, compared to traditional exercise formats, indicating that exergames encourage participants to engage in exercise of moderate to vigorous intensity compared to other exercise programs [26,27]. Beyond physical exertion, VR exergaming has also demonstrated potential for cognitive and emotional benefits among older adults, including moderate improvements in cognitive function, memory, and depressive symptoms, suggesting advantages for both mental health and healthy aging [27-29]. However, despite the promise of VR-based exergaming, there are few studies exploring the acceptability of VR among Latino, Latina, and Latine groups with cardiometabolic risk factors and obesity. Addressing this gap is essential to ensure equitable benefits from VR-based interventions that have demonstrated potential to promote physical activity and improve health outcomes.

By enabling individuals to engage in physical activity from the safety and convenience of their homes, VR technology has the potential to reduce barriers and enhance physical activity access in the Latino, Latina, and Latine community. This paper describes a community-informed protocol (version 1) for the Movement Over Virtual Reality study, which aims to assess the acceptability of a VR dance-based exercise program to promote physical activity among Latino, Latina, and Latine adults with cardiometabolic risk factors and obesity.

This protocol was developed in consultation with community health workers (CHWs) from the Latino, Latina, and Latine population and a community advisory board (CAB). If the focus groups demonstrate the acceptability of using VR to promote physical activity among Latino, Latina, and Latine adults with cardiometabolic risk factors and obesity, particularly those in underresourced settings, we plan to conduct a future pilot study to assess the feasibility of at-home VR exercise by providing participants with VR headsets and wearable devices to track use over time.

Methods

Pilot Phase With CHWs

To refine the Movement Over Virtual Reality focus group study protocol, we conducted 2 pilot sessions with 24 CHWs to gather feedback on our proposed study procedures. All CHWs were Spanish-speaking Latino, Latina, and Latine adults between the ages of 30 and 70 years, as well as first-time VR users. The CHWs work independently to engage communities and provide health education, advocacy services, and support. During the 2 pilot sessions, CHWs were paired with trained Spanish-speaking facilitators at a ratio of 1 instructor to 3 CHWs to provide instruction on headset placement, VR hand controller use, and the proposed VR experience. The initial VR experience involved CHWs completing skill-building tasks such as hitting targets to become familiar with VR hand-body coordination and was followed by a dance session with a virtual dance partner. Facilitators remained with participants throughout the VR experience to ensure safety and answer technical questions. After the 30-minute experience, we debriefed CHWs on how

XSL•FO
RenderX

to improve future focus group sessions. CHWs emphasized the need for 1:1 facilitation and suggested adding a group VR orientation session, separate "stations" with physical boundaries, and a nearby chair for those who may experience dizziness or fatigue, as well as shortening the session by removing skill building and proceeding directly to the dance experience. Participants also requested a Spanish version of the VR dance experience. All suggestions were integrated into this study protocol except for a fully Spanish-language VR dance experience due to lack of availability on the Meta platform. CHWs were compensated with US $100 for their support in providing feedback to refine the study protocol.

Incorporating CAB Feedback

A CAB composed of community leaders and the research study team convenes quarterly, and CAB members receive an honorarium of US $125 per hour funded as part of a larger career development grant (K23DK129828). We held 2 meetings to review and adapt study recruitment materials and focus group questions. During these meetings, the study team presented the theoretical framework of acceptability and the proposed focus group interview guide, which included questions in both English and Spanish [24]. The CAB provided recommendations to modify focus group questions to incorporate more lay language and enhance Spanish interpretation. Additionally, the study team reviewed flyers and the step-by-step recruitment approach, resulting in changes to the recruitment materials. These changes included adding a hyperlink and web address to the screening survey alongside the QR code and a phone number for interested participants to call.

Eligibility Criteria for Focus Group Participants

Eligibility criteria for focus group participants were (1) self-identification as Latino, Latina, and Latine; (2) age of ≥18 years; and (3) ability to speak English or Spanish. Given that obesity is a major risk factor for type 2 diabetes [1,5], cardiometabolic risk factors were screened based on a self-reported diagnosis of prediabetes or a moderate to high risk score for developing type 2 diabetes based on the Centers for Disease Control and Prevention's "Do I Have Prediabetes?" screening tool [1,30]. Individuals with physical injuries preventing them from engaging in 30 minutes of physical activity were excluded. Prior VR experience was not required for participation.

CHW-Led Recruitment

Participants were recruited in collaboration with CHWs (Figure 1) in south and east Los Angeles, California, as these are 2 underresourced areas with predominantly Latino, Latina, and Latine populations [31,32]. Recruitment efforts were led by the study team and trained CHWs culturally and linguistically aligned with the anticipated study population. The CHWs spoke English and Spanish and were aged between 30 and 70 years and of a similar cultural background to that of the target study participants. Most CHWs reported prior experience working on research studies and providing health education, resource linkages, and service navigation. Before commencing recruitment, the CHWs received in-person training, in which they were briefed on the study and their expected role in recruitment. After completing the training, 8 CHWs expressed interest in participating, with the goal of recruiting 6 to 8 participants each over a 5-month period. A memorandum of understanding outlining the study goals, the study team commitments, and expectations for CHWs was provided to those amenable to recruiting for the study. CHWs were supplied with recruitment materials, including study flyers, a 1-page frequently asked question sheet, and a Meet the Study Team document based on their recommendations. CHWs were compensated with US $250 for their support with recruitment.

Figure 1. Flowchart of the study participant recruitment and screening process. CDC: Centers for Disease Control and Prevention; CHW: community health worker; REDCap: Research Electronic Data Capture; VR: virtual reality.

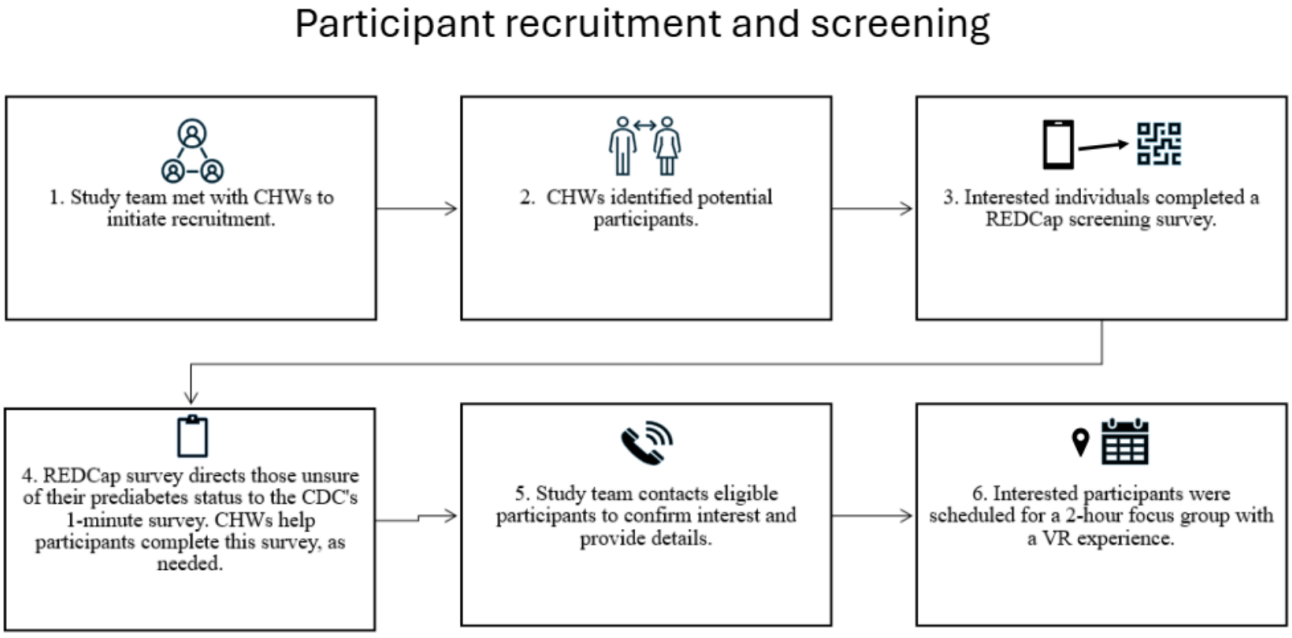

Throughout recruitment, the study team held 30-minute biweekly virtual meetings with the CHWs. During these meetings, the study team provided recruitment and enrollment updates. Additionally, the study team created a WhatsApp group

XSL•FO
RenderX

chat to provide reminders about virtual meetings, aggregate recruitment updates, and address general questions for the group. No personal identifiers or individual-level information were shared on the WhatsApp platform. The study team was also available to answer CHWs' individual questions or concerns via phone and email. Furthermore, the study team also fostered continual engagement by providing monthly workshops on health topics selected by the CHWs, including preventive women's cancer screenings and educational sessions on type 1 and 2 diabetes.

Study Procedures

Upon identifying potential participants, CHWs offered them a QR code or link to a brief eligibility survey, which was received by the study team once it was completed by potential participants. Screening surveys were collected and managed using the secure web-based software platform REDCap (Research Electronic Data Capture; Vanderbilt University) [33,34]. Bot spam was prevented by requiring participants to complete a reCAPTCHA (reCAPTCHA Inc) test before accessing the survey. Participants provided their contact information and age group and confirmation that they had been diagnosed with prediabetes within the previous 12 months. Participants who did not have a prediabetes diagnosis or were unsure of it were asked to complete the Centers for Disease Control and Prevention's "Do I Have Prediabetes?" screening tool through an embedded link to determine their prediabetes risk score. Participants scoring 0 to 4 were considered low risk and ineligible, whereas scores of 5 to 10 were considered moderate to high risk and eligible for study inclusion. The interest survey and risk tool were available in both English and Spanish. Eligible individuals were contacted by the study team via phone or email to schedule their participation in an in-person focus group in their preferred language (ie, English or Spanish).

VR Dance Experience and Focus Group Procedures

VR Equipment

To facilitate the VR dance experience, we used the Meta Quest 3S VR headset (Reality Labs), which includes head and hand position tracking for accurate spatial orientation. Participants interacted with the VR environment using 2 wireless Meta Quest Touch Plus controllers (Reality Labs), which included wrist straps to prevent slippage. These controllers served as motion sensors, enabling precise real-time tracking of participants' full-body movements. We used FitXR as the VR fitness application because it offered an interactive Zumba dance experience with Spanish subtitles. Specifically, participants engaged in a 10-minute beginner Zumba session set to Latin music led by a virtual instructor. To support the VR setup, a Netgear hot spot offering 5G network compatibility and encrypted security to support multiple devices was used to provide wireless internet connectivity.

Stations for the VR Dance Experience

VR sessions took place in a large room in a community center. Each participant was assigned to a "station," which served as the designated area where they could safely engage in the immersive VR dance experience. To promote safety, stations were spaced at approximately 2 arms' length. The study staff used visual boundaries on the floor to create virtual safety boundaries for each VR headset—a built-in safety feature notifying VR users when they approached the boundaries of their area.

Each station was labeled with a number corresponding to the participants' assigned headset. A printed checklist was also displayed at each station for facilitators to reference during setup and throughout the session. Additionally, each station featured a centrally located chair to accommodate participants preferring to be seated during the VR dance experience or who experienced motion sickness or disorientation during their session.

Participant Orientation for the VR Dance Experience

Before beginning the VR dance experience, the study team provided a structured orientation to prepare participants that incorporated CHW feedback from the pilot sessions. Participants received orientation materials in Spanish or English based on their preference, including a slide presentation with a detailed agenda, introduction to study staff, consent forms, pre- and postsurvey instructions, and an overview of the VR equipment. The slide presentation included videos and images to review basic VR skills, including how to adjust the VR headsets; use the hand controller buttons; and set physical space boundaries, a feature integrated into the VR technology to maintain user safety.

1:1 VR Facilitation

Participants were paired with a culturally and linguistically congruent study team member who guided and supported them throughout the VR dance experience. All VR facilitators received training on VR equipment operation and navigation of the VR dance program to ensure effective assistance of participants. At the start of the VR dance experience, each facilitator introduced themselves to their assigned participant and accompanied them to their designated VR station. The facilitator then assisted participants in putting on the VR headset. Each facilitator checked that participants' VR headsets were secure and comfortable and the proper placement of controller wrist straps to prevent accidental drops during the experience. Once participants indicated readiness, the facilitators launched the VR dance program, remaining with the participants throughout to ensure continued safety and provide real-time support. To better assist participants, facilitators also paired VR headsets with a secondary device (eg, phone or tablet) to mirror participants' VR display. This allowed facilitators to respond to questions while seeing participants' mirrored real-time experience in the headset. At the end of the experience, facilitators aided in removing the VR equipment and reviewed VR performance metrics with the participants, including their overall score, best streaks, and points earned.

Data Collection

Participants completed a baseline questionnaire after consenting to study participation and before the VR experience to assess their previous VR use. After the VR experience, participants completed a demographic survey including their age, gender, educational attainment, income, marital status, employment status, country of birth, and preferred language. The survey also included questions about participants' internet and technology

XSL•FO
RenderX

use; neighborhood composition; and self-rated health, sleep quality, and perceived stress levels. Administering the demographic survey after the VR session improved workflow efficiency by allowing participants who finished earlier to begin the survey while others completed their VR dance experience.

Next, focus groups were conducted in person by a primary facilitator and a cofacilitator (both trained study team members) using an interview guide developed in collaboration with the CAB and CHWs. The semistructured interview guide included a standard introduction; ground rules; and a series of open-ended questions designed to elicit participants' perspectives on the VR dance experience and the acceptability of VR as a physical activity modality among Latino, Latina, and Latine adults (see Multimedia Appendix 1). A third study team member took notes to guide future transcript analysis. Focus group discussions were audio recorded using digital recorders. All recordings were transcribed verbatim, and when necessary, a certified Spanish translation and transcription service was used to translate them into English to ensure accuracy. Each focus group lasted approximately 1 hour, allowing for sufficient discussion time. Finally, we administered postintervention surveys immediately after the focus groups. The survey asked participants to rate their satisfaction with their VR dance and overall VR experience. To assess safety, participants were also asked whether they experienced any side effects during VR use. All data were collected in the participants' preferred language and deidentified for privacy.

Data Analysis

Focus group interviews will be coded using Dedoose (SocioCultural Research Consultants), a qualitative data software platform. The study team will develop a coding framework based on the semistructured interview guide informed by the theoretical framework of acceptability, which evaluates the perceptions and attitudes of participants regarding the intervention [35]. Focus group transcripts will be independently coded by 2 trained research team members using a predefined coding framework in alignment with the coding reliability thematic analysis approach. Each week, the study team will meet to review and refine the preliminary codebook. Once finalized, subsequent transcripts will be coded in pairs. Coder discrepancies will be discussed among the research team and resolved through consensus. Field notes will be used to monitor focus group data for saturation. The analysis of the transcripts and coding will continue until no new themes emerge from the data. To arrive at the final themes and subthemes, we will achieve consensus through iterative discussions among research team members, and this process will be documented to ensure thoroughness. Data will be analyzed using thematic analysis, a systematic method for identifying, analyzing, and reporting patterns within the data [36]. Triangulation will be achieved by comparing focus group data with field notes taken during the sessions and other data sources such as the demographic survey and postintervention survey open-ended questions.

Ethical Considerations

This study was approved by the Cedars-Sinai Institutional Review Board (IRB; protocol ID: STUDY00003356). All procedures involving human participants were conducted in accordance with the Cedars-Sinai IRB. Written informed consent was obtained from all participants before initiating the study by trained study staff. Participant data collected during the focus groups were deidentified and stored securely to maintain confidentiality. A US $100 gift card was given to each participant who completed the study as compensation for their time and contributions. No copyrighted surveys were used in this study. This study did not involve research on animals.

Results

This study was funded on July 1, 2024, and received IRB approval on August 9, 2024. Participant recruitment began in January 2025. As of April 2025, a total of 125 individuals have been screened, and 35.2% (44/125) of participants have been enrolled. Seven focus groups have been conducted with enrolled participants across 3 age groups: 18 to 29 years, 30 to 49 years, and 50 to 70 years. Five focus groups were conducted in Spanish, and 2 were conducted in English. Data collection was completed in May 2025, and preliminary analyses began shortly thereafter.

A total of 73% (32/44) of the participants enrolled as of April 30, 2025, identified as female, with an average age of 46 (SD 14.1, range 18-72) years. Most participants (27/44, 61%) preferred Spanish, and 82% (36/44) reported being born outside of the United States (Table 1). Study findings are expected to be published in February 2026 and will be disseminated to the broader scientific community as well as our partner CHWs, the CAB, and the participants.

XSL•FO
RenderX

Table . Characteristics of the focus group participants (n=44).

Participant demographics	Values
Age (y), mean (SD)	46 (14)
Categorical age (y), n (%)	
18 - 29	9 (21)
30 - 49	16 (36)
50 - 70	19 (43)
Sex, n (%)	
Male	12 (27)
Female	32 (73)
Preferred language, n (%)	
Spanish	27 (61)
English	10 (23)
Both	7 (16)
English proficiency, n (%)	
Very well	10 (23)
Well	4 (9)
Not well	23 (52)
Not at all	3 (7)
Missing/Prefer not to answer	4 (9)
Ethnicity[a], n (%)	
Mexican, Mexican American, or Chicano	25 (57)
Salvadoran	3 (7)
Guatemalan	11 (25)
Nicaraguan	2 (5)
Indigenous	1 (2)
Other Hispanic, Latino, or Spanish origin	6 (14)
Born outside of the United States, n (%)	36 (82)
Country of birth, n (%)	
Mexico	18 (50)
Guatemala	9 (25)
El Salvador	3 (8)
Other[b]	4 (11)
Prefer not to respond	2 (6)
Time lived in the United States (y), mean (SD)	27 (16)
Household size, n (%)	
0 - 4	27 (61)
5 - 6	15 (34)
Prefer not to respond	2 (5)
Employment status, n (%)	
Employed	17 (39)
Unemployed	4 (9)
Other[c]	23 (52)

XSL•FO
RenderX

Participant demographics	Values
Household income (US $)[d], n (%)	
<41,600	20 (46)
41,600 - 69,350	8 (18)
>69,350	3 (7)
Prefer not to respond	13 (29)
Marital status, n (%)	
Married	18 (41)
Other	26 (59)
General health status, n (%)	
Excellent	0 (0)
Very good	8 (18)
Good	12 (27)
Fair	22 (50)
Poor	2 (5)
Diagnosis of prediabetes, n (%)	
Yes	33 (75)
No	7 (16)
Not sure	4 (9)

[a]Participants were allowed to select more than 1 option.

[b] Includes Colombia, Jamaica, Nicaragua, and Peru.

[c]Includes "housekeeper," "retired," "disabled," "temporarily laid off," "student," "caregiver," and "other."

[d]Levels based on the US Department of Housing and Urban Development in Los Angeles County based on a family size of 4 persons in 2024.

Discussion

This paper describes a protocol for evaluating the acceptability of a Latin dance–inspired VR exergame aimed at promoting physical activity among Latino, Latina, and Latine adults with cardiometabolic risk factors and obesity in underresourced settings. Our objective is to identify potential barriers and share insights relevant to implementing a technology-driven study in these environments, with a focus on ensuring access for groups that may otherwise lack exposure to new technology. Furthermore, our community-informed protocol may support other research that addresses issues related to digital health equity.

To address these challenges, this study protocol incorporated practical solutions developed with input from CHWs who reflect the study population and were involved in the early stages of protocol development. On the basis of our experience, individualized guidance from trained staff members who were culturally and linguistically aligned with participants was a key component for first-time VR users. In addition, as many participants were first-time VR users with limited digital literacy, we provided clear, hands-on instructions, including how to adjust the headset, use the controllers, and navigate safety features such as the virtual boundary system.

CHWs were crucial in refining the protocol and recruiting participants, leveraging their trust and community ties to engage populations traditionally underrepresented in research [37-39]. Furthermore, CHW involvement offers a sustainable strategy for addressing digital health access and literacy disparities that should be considered [40-42]. To sustain CHW engagement over the 5-month recruitment period, we held biweekly meetings to share recruitment updates and outreach strategy collaboration. Additionally, we fostered a bidirectional colearning environment through health education workshops on topics selected by the CHWs. Working with CHWs to use the VR equipment and participate in the pilot VR sessions before participant recruitment was key to building trust and providing them with the opportunity to share information about the study authentically as part of their recruitment approach.

On the basis of our experience, there are some potential challenges to consider when conducting technology-driven studies with populations with limited or no prior VR experience. These include access to facilitators who can help participants with headset operation and troubleshooting of any technical issues that may arise. Space considerations should include ample room to account for participant movement during headset use, as well as strong Wi-Fi signal to support the simultaneous use of multiple devices. Our decision to implement a short session was influenced by the need to allocate additional time to preparing the VR headsets at study onset. This included setting up generic profiles on the headsets, allowing participants to experience the program without the need to create individual profiles and avatars. While this approach limits personalization,

XSL•FO
RenderX

preventing participants from tailoring their experience to their preferences and fitness levels, it was necessary to streamline the process and enable participants to focus on the VR experience itself. Economizing time and allowing participants to focus on their experience of engaging with the headset outweighed this limitation.

If the proposed qualitative findings demonstrate high acceptability within the target population, the next phase will involve conducting a longitudinal feasibility study. In this future study, participants will be provided with a VR headset to use at home, allowing us to measure physical activity and behavioral outcomes and assess long-term adherence over time. Although this study focuses on evaluating the initial acceptability of the VR dance-based exercise program, future research should incorporate a control group. This addition will facilitate a more rigorous assessment of the efficacy of VR interventions on physical activity levels over time. In addition, we propose the use of wearable technology (ie, Fitbit trackers) in future studies as an objective measure of physical activity. Future studies should aim to equitably address technological literacy and meet the infrastructural needs of resource-limited settings by considering solutions such as providing portable Wi-Fi hot spots, as used in this study. Additionally, identifying or developing a Spanish-language option for a VR dance program should be prioritized to better meet the needs of participants. In this study, we selected the FitXR program because it features Spanish subtitles. It was also the most straightforward to follow with minimal instruction among the available options. However, the limitation of not having a full Spanish-language VR program available may create inequities for those whose preferred language is Spanish. This is an important area to address in future research. Expanding recruitment beyond south and east Los Angeles communities will also be crucial to enhance the generalizability of the findings across diverse Latino, Latina, and Latine populations, ensuring broader applicability and impact.

Findings from this and future studies will provide insight and may serve as a step toward ensuring equitable access to innovative technology-driven health interventions, including VR, for all communities. Future longitudinal studies should explore innovative models that support the sustainable use of VR in both clinical and community settings. By leveraging gamification technology, these models have the potential to promote healthy behaviors and physical activity in high-risk groups through technology-driven approaches. Similar models have been successful, as demonstrated by Medicare's coverage of VR as evidence-based treatment for pain management and mental health, illustrating the potential to integrate cutting-edge technology as a health care benefit [43,44].

Acknowledgments

The authors would like to thank their consultant community organizations, advisory board members, and community health worker partners who helped inform this study protocol.

Funding

YC-L (principal investigator) was funded by the National Institutes of Health and the National Institute on Minority Health and Health Disparities Pilot Award (P50-MD017366 UC END-DISPARITIES), as well as by the National Institutes of Health National Center for Advancing Translational Science (NCATS) through the UCLA Clinical and Translational Science Institute (CTSI; UL1TR001881), which partially supported this pilot award. This study also received funding from the 2024 Cancer Prevention and Control Program Research Developmental Funds Award sponsored by Cedars-Sinai Cancer. The funders had no role in the study design or protocol. The content is solely the responsibility of the authors and does not necessarily represent the official views of the named funders. Research reported in this publication was supported by the National Center for Advancing Translational Science of the National Institutes of Health under the UCLA Clinical and Translational Science Institute grant number UL1TR001881.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: YC-L
Funding acquisition: YC-L and OKD
Methodology: DRA, YC-L, LA-H, and GP
Project administration: YC-L, DRA, LA-H, and GP
Software: YC-L, DRA, LA-H, GP, IG-R, PM-G, LG, and JC
Supervision: YC-L
Visualization: YC-L
Writing—original draft: YC-L, DRA, LA-H, GP, IG-R, PM-G, LG, JC, JMC, and VNT
Writing—review and editing: YC-L, DRA, LA-H, GP, IG-R, PM-G, LG, JC, JMC, VNT, and OKD
All authors have read and agreed to the published version of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

OKD reports receiving consulting fees from Boehringer Ingelheim for work unrelated to this study, and YC-L reports receiving consulting fees from Eli Lilly and Company for work unrelated to this study. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Institutional review board–approved focus group question guide.
[PDF File, 74 KB - resprot_v15i1e80534_app1.pdf]

Checklist 1
SPIRIT checklist.
[PDF File, 685 KB - resprot_v15i1e80534_app2.pdf]

References

1. Lega IC, Lipscombe LL. Review: diabetes, obesity, and cancer-pathophysiology and clinical implications. Endocr Rev 2020 Feb 1;41(1):bnz014. [doi: 10.1210/endrev/bnz014] [Medline: 31722374]
2. Emmerich SD, Fryar CD, Stierman B, Ogden CL. Obesity and severe obesity prevalence in adults: United States, August 2021-August 2023. NCHS Data Brief 2024 Sep(508). [doi: 10.15620/cdc/159281] [Medline: 39808758]
3. Hales CM, Carroll MD, Fryar CD, Ogden CL. Prevalence of obesity and severe obesity among adults: United States, 2017-2018. NCHS Data Brief 2020 Feb(360):1-8. [Medline: 32487284]
4. Alemán JO, Almandoz JP, Frias JP, Galindo RJ. Obesity among Latinx people in the United States: a review. Obesity (Silver Spring) 2023 Feb;31(2):329-337. [doi: 10.1002/oby.23638] [Medline: 36695058]
5. Obesity in Los Angeles County: prevalence and risk factors. Los Angeles County Department of Public Health. 2012. URL: http://publichealth.lacounty.gov/ha/reports/LAHealthBrief2011/Obesity/Obesity_2012_sFinal.pdf [accessed 2025-05-29]
6. Ogden CL, Fakhouri TH, Carroll MD, et al. Prevalence of obesity among adults, by household income and education - United States, 2011-2014. MMWR Morb Mortal Wkly Rep 2017 Dec 22;66(50):1369-1373. [doi: 10.15585/mmwr.mm6650a1] [Medline: 29267260]
7. Funk C, Lopez MH. Hispanic Americans' trust in and engagement with science. Pew Research Center. 2022 Jun 14. URL: https://www.pewresearch.org/science/2022/06/14/hispanic-americans-trust-in-and-engagement-with-science [accessed 2025-12-02]
8. Silveira EA, Mendonça CR, Delpino FM, et al. Sedentary behavior, physical inactivity, abdominal obesity and obesity in adults and older adults: a systematic review and meta-analysis. Clin Nutr ESPEN 2022 Aug;50:63-73. [doi: 10.1016/j.clnesp.2022.06.001] [Medline: 35871953]
9. Adult physical inactivity outside of work. Centers for Disease Control and Prevention. 2025. URL: https://www.cdc.gov/physical-activity/php/data/inactivity-maps.html [accessed 2025-12-02]
10. Adeigbe RT, Ramirez AG. Physical activity in Latino communities. : Institute of Medicine; 2015 URL: http://nam.edu/wp-content/uploads/2015/06/PAandLatinos.pdf [accessed 2025-12-02]
11. Dao MC, Yu Z, Maafs‐Rodríguez A, et al. Perceived intrinsic, social, and environmental barriers for weight management in older Hispanic/Latino adults with obesity. Obes Sci Pract 2022 Aug 31;9(2):145-157. [doi: 10.1002/osp4.631] [Medline: 37034568]
12. Perez LG, Arredondo EM, Elder JP, Barquera S, Nagle B, Holub CK. Evidence-based obesity treatment interventions for Latino adults in the U.S.: a systematic review. Am J Prev Med 2013 May;44(5):550-560. [doi: 10.1016/j.amepre.2013.01.016] [Medline: 23597822]
13. Rosas LG, Lv N, Xiao L, et al. Effect of a culturally adapted behavioral intervention for Latino adults on weight loss over 2 years: a randomized clinical trial. JAMA Netw Open 2020 Dec 1;3(12):e2027744. [doi: 10.1001/jamanetworkopen.2020.27744] [Medline: 33337491]
14. Giakoni-Ramírez F, Godoy-Cumillaf A, Espoz-Lazo S, Duclos-Bastias D, Del Val Martín P. Physical activity in immersive virtual reality: a scoping review. Healthcare (Basel) 2023 May 25;11(11):1553. [doi: 10.3390/healthcare11111553] [Medline: 37297694]
15. Spiegel B. VRx: How Virtual Therapeutics Will Revolutionize Medicine: Basic Books; 2020.
16. Maddox T, Garcia H, Ffrench K, et al. In-home virtual reality program for chronic low back pain: durability of a randomized, placebo-controlled clinical trial to 18 months post-treatment. Reg Anesth Pain Med 2024 May 7;49(5):373-375. [doi: 10.1136/rapm-2022-104093] [Medline: 36427904]
17. Gao Z, Chen S, Pasco D, Pope Z. A meta-analysis of active video games on health outcomes among children and adolescents. Obes Rev 2015 Sep;16(9):783-794. [doi: 10.1111/obr.12287] [Medline: 25943852]
18. Przybylski AK, Rigby CS, Ryan RM. A motivational model of video game engagement. Rev Gen Psychol 2010;14(2):154-166. [doi: 10.1037/a0019440]

XSL•FO
RenderX

19. Andrade A, Correia CK, Coimbra DR. The psychological effects of exergames for children and adolescents with obesity: a systematic review and meta-analysis. Cyberpsychol Behav Soc Netw 2019 Nov;22(11):724-735. [doi: 10.1089/cyber.2019.0341] [Medline: 31697604]
20. Finkelstein S, Suma EA. Astrojumper: motivating exercise with an immersive virtual reality exergame. Presence 2011 Feb 1;20(1):78-92. [doi: 10.1162/pres_a_00036]
21. Farič N, Smith L, Hon A, et al. A virtual reality exergame to engage adolescents in physical activity: mixed methods study describing the formative intervention development process. J Med Internet Res 2021 Feb 4;23(2):e18161. [doi: 10.2196/18161] [Medline: 33538697]
22. Yoong SQ, Wu VX, Chen C, et al. Dance exergames for older adults: a systematic review and meta-analysis with narrative synthesis. J Gerontol A Biol Sci Med Sci 2024 Apr 1;79(4):glae035. [doi: 10.1093/gerona/glae035] [Medline: 38306590]
23. Zhao M, Lei Y, Wei Z, You M. The effect of virtual reality cycling with music on simple obesity in college students: evidence from a randomized controlled trial in China. Front Public Health 2024 Nov 21;12:1466142. [doi: 10.3389/fpubh.2024.1466142] [Medline: 39639896]
24. Schroeder K, Ratcliffe SJ, Perez A, Earley D, Bowman C, Lipman TH. Dance for health: an intergenerational program to increase access to physical activity. J Pediatr Nurs 2017;37:29-34. [doi: 10.1016/j.pedn.2017.07.004] [Medline: 28733128]
25. Wang YH. Understanding senior adults' needs, preferences, and experiences of commercial exergames for health: usability study. JMIR Serious Games 2024 Apr 5;12:e36154. [doi: 10.2196/36154] [Medline: 38578674]
26. Davis JC, Killen LG, Green JM, Waldman HS, Renfroe LG. Exergaming for physical activity: a systematic review. J Am Coll Health 2024 Oct;72(7):2090-2098. [doi: 10.1080/07448481.2022.2103377] [Medline: 35882066]
27. Wei X, Huang C, Ding X, et al. Effect of virtual reality training on dual-task performance in older adults: a systematic review and meta-analysis. J Neuroeng Rehabil 2025 Jun 24;22(1):141. [doi: 10.1186/s12984-025-01675-z] [Medline: 40551196]
28. Li X, Zhang Y, Tang L, Ye L, Tang M. Effects of virtual reality-based interventions on cognitive function, emotional state, and quality of life in patients with mild cognitive impairment: a meta-analysis. Front Neurol 2025 Apr 2;16:1496382. [doi: 10.3389/fneur.2025.1496382] [Medline: 40242620]
29. Yang Q, Zhang L, Chang F, Yang H, Chen B, Liu Z. Virtual reality interventions for older adults with mild cognitive impairment: systematic review and meta-analysis of randomized controlled trials. J Med Internet Res 2025 Jan 10;27:e59195. [doi: 10.2196/59195] [Medline: 39793970]
30. Do I Have Prediabetes?. URL: https://doihaveprediabetes.org [accessed 2025-12-02]
31. Exploring the Latino metropolis. Scalar/University of Southern California. URL: https://scalar.usc.edu/works/latino-metropolis-a-brief-urban-cultural-history-of-us-latinos---1/index [accessed 2025-10-17]
32. Los Angeles County (central)--LA City (Southeast/East Vernon) PUMA, CA. Census Reporter. URL: https://censusreporter.org/profiles/79500US0603745-los-angeles-county-central-la-city-southeasteast-vernon-puma-ca/ [accessed 2025-10-17]
33. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
34. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
35. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017 Jan 26;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
36. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
37. Killough CM, Madaras A, Phillips C, et al. Community health worker insights on promoting research engagement with diverse populations. Front Public Health 2023 Jan 11;10:959504. [doi: 10.3389/fpubh.2022.959504] [Medline: 36711331]
38. Nebeker C, Giacinto RE, Pacheco BA, López-Arenas A, Kalichman M. Prioritizing competencies for "research" promotores and community health workers. Health Promot Pract 2021 Jul;22(4):512-523. [doi: 10.1177/1524839920913548] [Medline: 32228241]
39. Coulter K, Ingram M, McClelland DJ, Lohr A. Positionality of community health workers on health intervention research teams: a scoping review. Front Public Health 2020 Jun 16;8:208. [doi: 10.3389/fpubh.2020.00208] [Medline: 32612967]
40. O'Brien MJ, Perez A, Alos VA, et al. The feasibility, acceptability, and preliminary effectiveness of a Promotora-Led Diabetes Prevention Program (PL-DPP) in Latinas: a pilot study. Diabetes Educ 2015 Aug;41(4):485-494. [doi: 10.1177/0145721715586576] [Medline: 26023095]
41. Ali MK, Echouffo-Tcheugui J, Williamson DF. How effective were lifestyle interventions in real-world settings that were modeled on the Diabetes Prevention Program? Health Aff (Millwood) 2012 Jan;31(1):67-75. [doi: 10.1377/hlthaff.2011.1009] [Medline: 22232096]
42. Hill J, Peer N, Oldenburg B, Kengne AP. Roles, responsibilities and characteristics of lay community health workers involved in diabetes prevention programmes: a systematic review. PLoS ONE 2017 Dec 7;12(12):e0189069. [doi: 10.1371/journal.pone.0189069] [Medline: 29216263]

XSL•FO
RenderX

43. Patients. XR Health. URL: https://www.xr.health/us/dme/patients/ [accessed 2025-10-24]
44. Insurance covers VR therapies for first time. LeadingAge. URL: https://leadingage.org/insurance-covers-vr-therapies-for-first-time/ [accessed 2025-10-24]

Abbreviations

CAB: community advisory board
CHW: community health worker
IRB: institutional review board
REDCap: Research Electronic Data Capture
VR: virtual reality

Edited by J Sarvestan; submitted 12.Jul.2025; peer-reviewed by ER Weinstein, H Milton, S Mehrabi; revised version received 30.Oct.2025; accepted 05.Nov.2025; published 12.Jan.2026.

Please cite as:
Acosta DR, Aguilar-Hernandez L, Perez G, Guzman-Ruiz IY, Castellon JM, Cruz J, Gutierrez L, Monteon-Garcia P, Torres VN, Duru OK, Castellon-Lopez Y
Evaluating the Acceptability of Using Virtual Reality to Promote Physical Activity Among Latino, Latina, and Latine Adults With Cardiometabolic Risk Factors and Obesity in Underresourced Settings: Protocol for a Qualitative Focus Group Study
JMIR Res Protoc 2026;15:e80534
URL: https://www.researchprotocols.org/2026/1/e80534
doi:10.2196/80534

© Desiree R Acosta, Leslie Aguilar-Hernandez, Gael Perez, Iris Y Guzman-Ruiz, Josyel M Castellon, Jailene Cruz, Liana Gutierrez, Paulina Monteon-Garcia, Vanessa N Torres, O Kenrik Duru, Yelba Castellon-Lopez. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cognitive Recovery After Poststroke Delirium (RECOVER): Protocol for a Longitudinal Multimodal Study on Cognitive Assessment in Patients With Stroke

Lena S Geiger[1,2], MD, MSc; Sebastian Klütz[1,2]; Christian Mychajliw[3,4], MSc; Christoph Gäbele[3,4], MSc; Andreas Jooß[1,2], MD; Constanze Single[1,2], MD; Katharina Feil[1], Prof Dr Med; Ulf Ziemann[1,2], Prof Dr Med; Annerose Mengel[1,2], PD Dr Med

[1]Department of Neurology & Stroke, University Hospital Tübingen, Hoppe-Seyler-Street 3, Tübingen, Germany

[2]Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

[3]Geriatric Center, University Hospital for Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany

[4]TuCAN, Tübingen Cognitive Assessment for Neuropsychiatric Disorders, Tübingen, Germany

Corresponding Author:
Lena S Geiger, MD, MSc
Department of Neurology & Stroke, University Hospital Tübingen, Hoppe-Seyler-Street 3, Tübingen, Germany

Abstract

Background: Poststroke delirium (PSD) is a severe complication in patients with acute stroke and is characterized by rapid onset fluctuating symptoms, which affect multiple domains (cognition, motor system, and sleep-wake cycle). Similar to other types of delirium, PSD is associated with longer hospitalization, higher mortality, and a higher disability rate. Behavioral studies on cognitive functioning showed significantly poorer cognitive outcomes in both patients with and without PSD compared to healthy controls. Thus, the distinction between “stroke-related” and “PSD-related” cognitive impairments remains unclear. A frequently affected and highly disabling cognitive domain is memory function. However, imaging studies, particularly task-based functional magnetic resonance imaging (MRI) studies on PSD, are currently scarce and represent a significant gap in the existing literature.

Objective: This longitudinal proof-of-concept study aims to investigate the short- and long-term effects of stroke and PSD on cognitive outcome using a multimodal approach and to outline a protocol for a repeated multimodal cognitive assessment in patients with stroke.

Methods: We developed a longitudinal study protocol to investigate short- and long-term effects of stroke and PSD on cognitive impairment and recovery. Poststroke cognitive impairment (PSCI) will be assessed in a comprehensive digital multimodal approach including a multidomain neuropsychological app to facilitate a standardized, rapid testing, particularly for long-term outcomes, as well as task-based functional MRI (introducing a modified working memory task) during the acute (prior PSD development) and postacute phase post stroke and at a 3-month follow-up. In total, 40 patients with acute stroke, divided into a PSD group and non-PSD group (control), will be examined. In the context of the proof-of-concept study, the eligibility of a modified working-memory task, an app-based neuropsychological assessment, and a multimodal MRI protocol will be evaluated. The primary endpoint is a between-group comparison of the cognitive outcome, defined as global PSCI at a 3-month follow-up. Global PSCI will be classified as normal cognitive functioning, or mild, moderate, or severe cognitive impairment, according to the performance level (norm-referenced *z* scores) on a multidomain neuropsychological assessment.

Results: This study was funded in December 2023. As of May 2025, 15 patients with stroke have been included. Recruitment and data collection were initiated in June 2024 and are ongoing until December 2025. The findings are expected to be published in summer 2026.

Conclusions: This protocol outlines a proof-of-concept study aiming to fill a critical gap by systematically investigating cognitive functions in patients with acute stroke with and without delirium. Generating reliable preliminary data will provide essential groundwork for future large-scale research, ultimately enhancing understanding of the contribution of PSD to cognitive impairment.

Trial Registration: ClinicalTrials.gov NCT06680336; https://clinicaltrials.gov/study/NCT06680336

International Registered Report Identifier (IRRID): DERR1-10.2196/77508

(JMIR Res Protoc 2026;15:e77508) doi:10.2196/77508

XSL•FO
RenderX

KEYWORDS

cognitive assessment; cognitive functioning; cognitive outcome; fMRI; functional magnetic resonance imaging; longitudinal design; memory functions; multimodal approach; poststroke cognitive impairment; poststroke delirium; stroke

Introduction

Delirium is an acute clinical condition characterized by a wide range of symptoms affecting multiple domains. Patients with delirium often present with attention deficits, confusion, hallucinations, agitation, or disturbances in sleep-wake cycle. Typically, delirium occurs during critical illness, such as after surgery or stroke, with rapid onset and fluctuating symptoms [1]. Poststroke delirium (PSD) is one of the most common complications after stroke, with reported prevalence rates reaching up to 40% in older patients [1-3]. Both delirium in general and PSD in particular are associated with longer hospitalization, higher mortality, accelerated cognitive decline, and lower quality of life [1,4-11]. There is a broad consensus of a multifactorial origin [1,12,13], with predisposing factors such as dementia, age, comorbidities, and precipitating risk factors, including infection, stroke, or surgery [1,12,13]. The presence of multiple and fluctuating symptoms is diagnostically challenging, especially in patients with stroke with communication difficulties (eg, aphasia), and requires clinical expertise and standardized screening tools, such as the well-established ICDSC (Intensive Care Delirium Screening Checklist) [14,15]. Despite the high prevalence and the nondeclining mortality rate over the past decades [2,7], objective diagnostic tools and prognostic biomarkers (ie, fluid or proteomic biomarkers), as well as protective markers for long-term cognitive outcome, have not yet been implemented in clinical practice and remain the focus of ongoing studies [5,16-18].

In general, examining delirium patients is a challenging field, which limits the number of available studies. Patients with delirium, such as patients with PSD, represent a vulnerable and critically ill population. Especially during the acute (within 5 d poststroke) and the postacute phase (5 d to 12 wk poststroke), patients may experience secondary deteriorations or be scheduled for invasive diagnostic procedures or surgeries, making the conduction of studies difficult. Additionally, the fluctuating nature of delirium symptoms complicates the timing and consistency of study assessments. Furthermore, during the acute state of delirium, especially cases involving high arousal or hallucinating symptoms, cognitive assessment or magnetic resonance imaging (MRI) assessment becomes impossible, as patients cannot comply with task instructions. This results in a lack of studies and underscores the need for further research in this area. Another challenge in this field concerns the study design, varying with regard to concurrent versus sequential study designs, as well as the choice of an appropriate control group.

Previous studies have reported a higher risk of PSD in total anterior circulation infarcts compared to partial anterior, posterior, and lacunar infarcts [19-21]. Cortical networks in the anterior circulation, such as fronto-parietal networks, are known to be involved in higher-level cognitive functions, including executive functions, attention, and memory. These very cognitive domains are affected in PSD. Memory functions, which represent a critical component of human cognition, are frequently affected and represent a highly disabling cognitive domain in stroke survivors [22]. For example, memory impairments have been significantly associated with poor quality of life of patients with stroke and their caregivers, for example, [22,23], demonstrating the high relevance to daily life and potential burden. Moreover, memory assessment is a well-established component of both neuropsychological testing and in functional magnetic resonance imaging (fMRI) studies, making it an ideal focus for our study.

In the challenging field of cognitive assessment in patients with delirium, the majority of studies focus on behavioral measures including the assessment of multiple cognitive domains such as attention, language, memory, executive functions, and visuospatial construction. Previous studies have demonstrated a significantly poorer cognitive outcome (eg, global cognition and executive functions) in patients with postdelirium compared to healthy controls [24-26].

Additionally, advanced techniques such as MRI have been implemented to provide further insights. Structural MRI studies, which do not require a certain level of patients' compliance, have reported abnormalities in white matter tracts of patients with postdelirium, which are known to be involved in memory and attention [1,13,27]. Since delirium is known to be a disorder of disrupted brain networks [1,12,16,28,29], fMRI studies, which were predominantly conducted as resting-state studies in patients with postdelirium, have demonstrated aberrant connectivity in dorsal prefrontal and subcortical regions compared to healthy controls, which may contribute to the emergence of delirium symptoms [13,28,30]. These findings in network dysfunctions are restricted to the default mode network, and memory functions may not be detected during resting-state fMRI [31]. An essential requirement for studying memory functions and related activity or connectivity dysfunctions in delirium is a memory task-based fMRI study. Unfortunately, these are exceedingly limited.

To date, only 1 fMRI study has been conducted, using a well-known N-Back working memory task [32]. While reporting no association between delirium and brain region activity, several methodological limitations and concerns have been raised, such as the infeasibility of the working memory task (especially at the time of discharge), the infeasible MRI protocol, as well as the absence of a control group. Of note, connectivity analyses have not been conducted. Further research is therefore needed, alongside the development of an adequate and reliable task for working memory assessment during fMRI.

An additional challenge in studying PSD and cognitive outcomes is the unresolved distinction between PSD-related and pure stroke−related cognitive impairments. In detail, cognitive deficits associated with PSD and other types of delirium, for example, attention or memory deficits, have also been reported in patients with stroke without PSD [23]. Previous imaging

XSL•FO
RenderX

studies have shown reduced brain activation as well as abnormal connectivity within fronto-parietal networks in patients without PSD compared to healthy controls [31,33,34]. Thus, the origin of cognitive impairments in patients with PSD (ie, distinction between "stroke-related" vs "PSD-related" impairments) remains unclear.

To summarize, there is a lack of reliable and feasible fMRI tasks for assessing memory functions in patients with PSD. The few existing task-based fMRI studies have either been conducted without control groups or have only compared groups of patients with PSD to healthy controls, limiting the ability to disentangle the distinct cognitive and neural contributions of PSD versus stroke. Therefore, there is a significant gap in the literature regarding appropriate methods to study PSD and its cognitive outcomes, emphasizing the need for comprehensive, longitudinal investigations that combine neuropsychological assessments with functional neuroimaging.

To address this gap, we present a protocol for a comprehensive longitudinal proof-of-concept study that aims to systematically and comprehensively examine poststroke cognitive impairment (PSCI) in a critically ill stroke population, both with and without PSD (control group). Assessments will be conducted during the acute (within 5 d post stroke) and postacute phases (5 - 14 d post stroke), and at a 3-month follow-up using a multicomponent neuropsychological assessment and task-based fMRI. Generating reliable feasibility and preliminary data through this approach will provide essential groundwork for future multicenter studies.

The specific objectives of this study are to (1) compare the global PSCI of patients with and without PSD at a 3-month follow-up, (2) compare domain-specific cognitive outcome measures of patients with and without PSD at a 3-month follow-up, (3) generate preliminary data regarding the differentiation between PSD-related and stroke-related cognitive impairment using behavioral and fMRI connectivity measures, and (4) contribute toward developing protective and prognostic tools for the early detection and treatment of PSD.

Methods

Study Design and General Procedures

We present a protocol of the RECOVER (Cognitive Recovery After Poststroke Delirium) study, which is a monocenter study following a longitudinal design with 3 measurement points (S1-S3) and an individual trial duration of 3 months (Figure 1).

Figure 1. Overview of the RECOVER (Cognitive Recovery After Poststroke Delirium) study procedures. DSM-5: *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*; DTI: diffusion tensor imaging sequence; FLAIR: fluid-attenuated inversion recovery sequence; ICDSC: Intensive Care Delirium Screening Checklist; MPRAGE: T1-weighted magnetization prepared rapid acquisition with gradient echoes sequence; MRI: magnetic resonance imaging; NIHSS: National Institutes of Health Stroke Scale; S1: session 1; S2: session 2; S3: session 3; T2*: T2-weighted echo planar imaging sequence; WM: working memory.

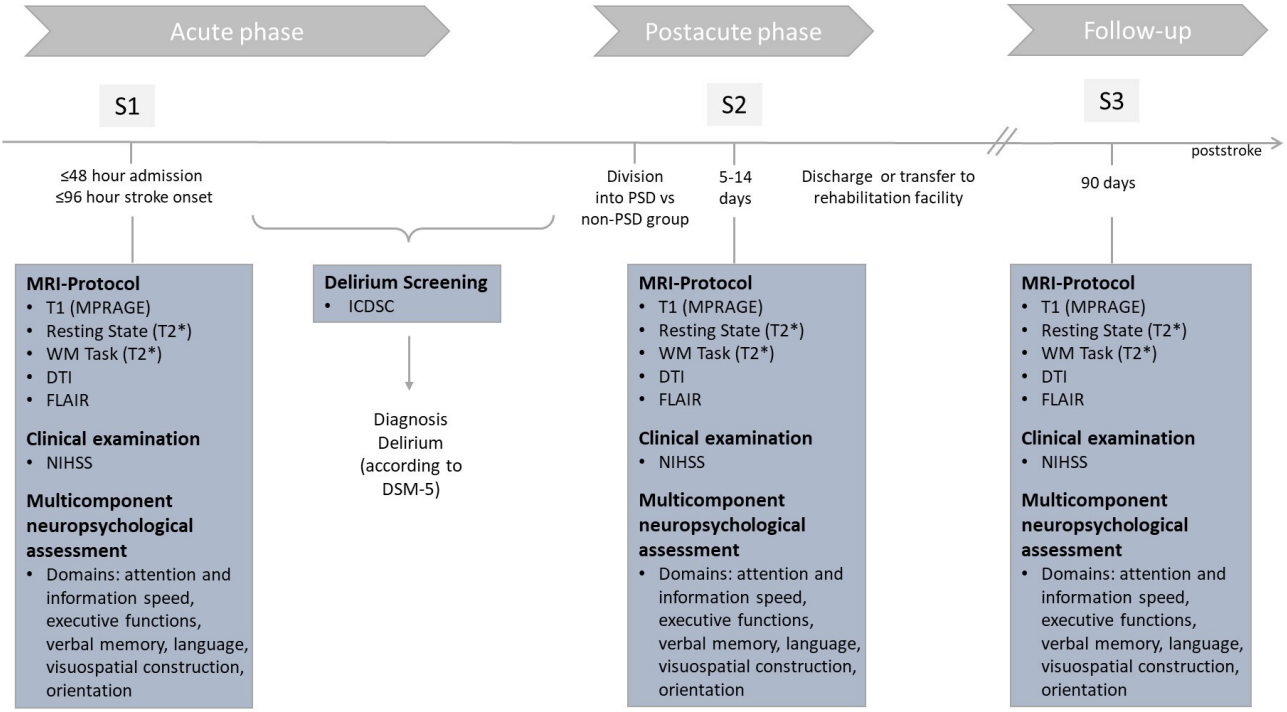

The three measurement points are as follows:

1. S1 during the acute phase poststroke (≤48 h after hospital admission; ≤96 h after stroke onset and before PSD development)
2. S2 during the postacute phase poststroke (5 - 14 d poststroke, depending on delirium recovery)
3. S3 follow-up at 3 months poststroke (approximately 90 d poststroke)

Patients with stroke will be admitted to the Stroke Unit where a screening visit will be conducted, including a review of the inclusion and exclusion criteria. Following the review, written informed consent will be obtained from the patient or, if applicable, the legal guardian. Once consent has been obtained, the patient will be enrolled in the RECOVER study.

Each measurement (S1-S3) consists of a clinical examination, a multicomponent neuropsychological testing, and an MRI

XSL•FO
RenderX

protocol. The MRI protocol is designed to acquire both structural (anatomical) data, as well as functional data (see the *Neuroimaging Protocol* section for more details). The clinical examination includes the National Institutes of Health Stroke Scale (NIHSS) [35] and will take approximately 15 minutes. The multicomponent neuropsychological assessment and the MRI scan (including positioning of patient) will each require 45 - 60 minutes. A more detailed description is provided in the *Neuropsychological Assessment and Clinical Scores* section. S1 and S2 will be conducted as in-patient measurements, and the S3 3-month follow-up measurement will be conducted as an out-patient measurement.

During the course of the longitudinal study, in particular before S2, the cohort of patients with stroke will be subdivided into 2 groups, those with PSD and those without (non-PSD). Delirium screening using the ICDSC [32] will be conducted throughout hospitalization to identify the onset and duration of delirium. To ensure timely detection of delirium onset and recovery, assessments will be performed every 8 hours. A diagnosis will be made according to the DSM-5 criteria by an independent neurologist, who will review the clinical course alongside the ICDSC data. Delirium recovery is defined as 3 consecutive negative ICDSC screenings.

Inclusion and Exclusion Criteria

This study includes patients with stroke with mild-to-moderate stroke symptoms (NIHSS ≤15; eg, [36,37]; Glasgow coma scale [GCS] >8). The prevalence of a PSD increases with age [20], reaching up to 40% among patients aged 60 years and older [1-3]. To ensure an equal distribution into the 2 study groups, only patients aged 60 years or older will be included in the RECOVER study. Textbox 1 gives an overview of all inclusion and exclusion criteria. When patients' conditions or transfers preclude in-hospital assessment, follow-up visits at patients' homes will be arranged.

Textbox 1. Inclusion and exclusion criteria for the study enrollment.

Inclusion criteria

- Aged ≥60 years
- National Institutes of Health Stroke Scale ≤15 [35]
- Acute ischemic stroke
- Stroke onset ≤96 hours
- Hospital admission ≤48 hours
- Informed consent by patient or legal guardian

Exclusion criteria

- Aphasia
- Reduced consciousness, defined as Glasgow Coma Scale ≤8 [38]
- Delirium onset before study enrollment
- Magnetic resonance imaging contraindication such as pacemakers or metal implants
- Preexisting conditions such as prior stroke, seizures, or traumatic head injury within the past 3 months, pre-existing neuropsychiatric disorders (eg, dementia, depression, and schizophrenia) or developmental delay
- High-grade paresis or plegia of the dominant hand

Study Population and Recruitment

This longitudinal proof-of-concept study aims to include a total of 40 patients with acute stroke, equally distributed in a PSD (n=20) and a non-PSD group (n=20). All the patients with acute stroke will be recruited from the wards of the Neurological Department of the University Hospital Tübingen.

Ethical Considerations

The study protocol has been approved by the local Ethics Committee of the Medical Faculty and the University Hospital Tuebingen (reference 444/2023BO2) and has been registered at ClinicalTrials.gov with the number NCT06680336. The study will follow the ethical principles for medical research involving human participants of the Declaration of Helsinki, adopted by the 18th General Assembly of the World Medical Association (World Medical Association, 1964). All patients or their legal guardians will provide informed consent before study enrollment.

This study protocol comes with important ethical considerations, given the vulnerability of patients with acute stroke, particularly those with PSD. Ethical responsibility and patient safety are central. To minimize the burden, neuropsychological testing and MRI sessions are kept as short as feasible and can be paused or rescheduled depending on the patient's condition. Testing is stopped immediately if signs of distress or fatigue are observed. Clinical staff remain available throughout all procedures. We believe that including patients with cognitive vulnerability is ethically justified, as they are often underrepresented in research. However, their inclusion is balanced with stringent safeguards to ensure that participation does not place undue burden on their health or recovery. The data collected during the study will be pseudonymized to ensure the confidentiality of participants. All personal information will be treated with the utmost care and

XSL•FO
RenderX

in accordance with applicable privacy regulations, including data protection laws. Patients are provided with a monetary compensation for their time and effort in accordance with the ethical guidelines of the study.

Neuropsychological Assessment and Clinical Scores

We followed a multicomponent cognitive assessment strategy using standardized tests, questionnaires, and clinical scores in order to ensure a comprehensive assessment of PSCI. The selection of tests and domains aligns with the recommendations of major stroke organizations, such as the European Stroke Organisation and the Canadian Stroke Network [39,40]. The following domains were defined: attention and information speed, executive functions, verbal memory, language, visuospatial construction, and orientation (Table 1). In general, the tests, questionnaires, and clinical scores used in this study protocol are appropriate for our elderly study population, with available normative data. The order of test administration is standardized to ensure consistency across participants and is arranged to alternate between cognitive domains and varies in modality (eg, verbal vs visuospatial) to reduce cognitive load and prevent fatigue. To reduce practice effects across sessions, we use official parallel test forms when available. In consideration of the elderly study population, the selection of tests and questionnaires was also based on font size, or in cases where applicable, the font size was increased. Table 1 illustrates the questionnaires and clinical scores, as well as their targeted components used in this study. The neurological examination comprises the standardized NIHSS [35]. Some parts of the assessment will be digitalized using a tablet app called TuCAN (Tübingen Cognitive Assessment for Neuropsychiatric Disorders) [41] and a digitizer pencil. The TUCAN app was designed for elderly individuals and has previously been evaluated in healthy older populations [42,43]. Furthermore, it has been used to assess cognitive functions in 2 clinical populations: patients with cognitive impairment and patients with stroke [44,45]. To address potential unfamiliarity with digital technology in older patients with stroke, patients complete a short familiarization task before the formal assessment begins, during which they are asked to write a sentence or a few words on the tablet.

XSL•FO
RenderX

Table . Standardized tests, questionnaires, and clinical scores with the corresponding targeted domain.

Target domain and test/clinical score/questionnaire	S1[a]	S2	S3
Attention and processing speed			
TMT[b]—Part A [46]	✓	✓	✓
NAB[c] Module Attention [47]	✓	✓	✓
Executive functions			
WM[d] span within NAB Module Attention [47]	✓	✓	✓
TMT—Part B [46]	✓	✓	✓
Verbal memory			
MIS[e] within MoCA[f] [48]	✓	✓	✓
Language			
LIS[g] within MoCA [48,49]	✓	✓	✓
Visuospatial construction			
VIS[h] within MoCA [48,49]	✓	✓	✓
Orientation			
OIS[i] within MoCA [48,49]	✓	✓	✓
Cognitive impairment			
MoCA [48]	✓	✓	✓
Questionnaires assessing cognitive functioning			
Self-rating CFQ[j] [50]	✓		
Relative questionnaire: IQCODE[k] [51]	✓		
Depression			
GDS[l] [52]	✓	✓	✓
Quality of life			
SF[m]-12 [53]	✓		✓
Global disability			
(p)mRS[n] [54]	✓		✓
Delirium screening			
ICDSC[o] [55]	✓	✓	✓
Stroke symptoms			
NIHSS[p] [35]	✓	✓	✓
Other variables			
Verbal intelligence test (MWT-B[q] [56])	✓		
Demographics	✓		
Stroke location, medical interventions, comorbidities, etc	✓		

[a]S: session.

[b]TMT: trail making test.

[c]NAB: Neuropsychological Assessment Battery.

XSL•FO
RenderX

[d]WM: working memory.
[e]MIS: memory index scale.
[f]MoCA: Montreal Cognitive Assessment.
[g]LIS: language index scale.
[h]VIS: visuospatial index scale.
[i]OIS: orientation index scale.
[j]CFQ: cognitive failure questionnaire.
[k]IQCODE: Informant Questionnaire on Cognitive Decline in the Elderly.
[l]GDS: Geriatric Depression Scale.
[m]SF: Short Form health survey.
[n](p)mRS: (pre)modified Rankin scale.
[o]ICDSC: Intensive Care Delirium Screening Checklist.
[p]NIHSS: National Institutes of Health Stroke Scale.
[q]MWT-B: German intelligence test *Mehrfach-Wortschatz-Intelligenztest*.

Neuroimaging Protocol

Functional and structural MRI data will be acquired using a 3 Tesla Siemens Prisma whole-body imaging system (Siemens), equipped with a 64-channel head coil. Functional and diffusion-weighted images are acquired using the Minnesota CMRR multiband accelerated pulse echo planar imaging (EPI) sequences [57]. For functional imaging, a T2*-weighted EPI sequence with repetition time (TR) of 800 ms, echo time (TE) of 37 ms, flip angle of 52°, 72 axial slices, 2-mm slice thickness, field of view (FoV) of 208 mm (whole-brain coverage was ensured by tilting the FoV to the individual anterior commissure – posterior commissure line) was used. Diffusion tensor imaging data are acquired using an EPI sequence with the following parameters: TR of 2100 ms, TE of 62 ms, flip angle of 72°, 92 slices, 1.50-mm isotropic resolution slice thickness, FoV of 210 mm, and 64 diffusion directions at *b* value of 1500 s/mm^2. For anatomical images, a T1-weighted magnetization-prepared rapid gradient-echo sequence with the following parameters: TR of 2500 ms, TE of 2.22 ms, 208 axial slices, 0.8-mm slice thickness, FoV of 256 mm will be used. Furthermore, to assess the stroke area, a standard Siemens Prisma T2-weighted fluid-attenuated inversion recovery sequence with TR of 9000 ms, TE of 81 ms, 36 axial slices, 4-mm slice thickness, and FoV of 230 mm will be used.

Working Memory Task

The working memory task will be a modified version of a well-known Sternberg task [58-62] with 4 task conditions. Given that our study population comprises older patients with stroke with PSD and without PSD, further modifications of a previously published Sternberg task [58,59,61,62] were necessary. Previous studies have reported difficulties regarding the feasibility of performing a working memory task after delirium [28,32], which we have taken into consideration. To enhance feasibility, the following modifications were implemented: [1] reduction of task difficulty by reducing the number of stimuli in the memory set from 5 to 4 items [2], shortening of the total task duration [3], adjustment of experimental conditions by including objects to enable the examination of the visual-spatial components of working memory.

This modified version of the Sternberg task consists of 4 task conditions with different working memory categories, that is, verbal and visual items. In detail, 2 conditions comprise letters (letter condition [LE] and letter control condition [LC]) and 2 comprise objects (object condition and object control condition), as illustrated in Figure 2. The letters in the letter condition (LE) consist of consonants, and the letters in the control condition (LC) consist of vowels. In order to map the differences in letters (eg, FGMP, as in the letter conditions) onto the differences in objects, several semantic categories were selected for the object conditions, for example, pictures of daily life items, clothes items, kitchen tools, and electrical devices. All objects were derived from the bank of standardized stimuli line drawings [63,64]. To minimize condition-dependent effects such as differences in visual activations, eye movement, or spatial attention, all stimuli are presented on a white screen, at a fixed position and in black font or black line drawings (Figure 2). All conditions are presented 5 times in a pseudo-randomized and counterbalanced order. The experiment will be performed using the Presentation software (version 23.0; Neurobehavioral Systems, Inc) [65].

XSL•FO
RenderX

Figure 2. Modified version of the Sternberg task (working memory task) with 4 task conditions (letters, letter control, objects, and object control) each consisting of an encoding and a retrieval phase.

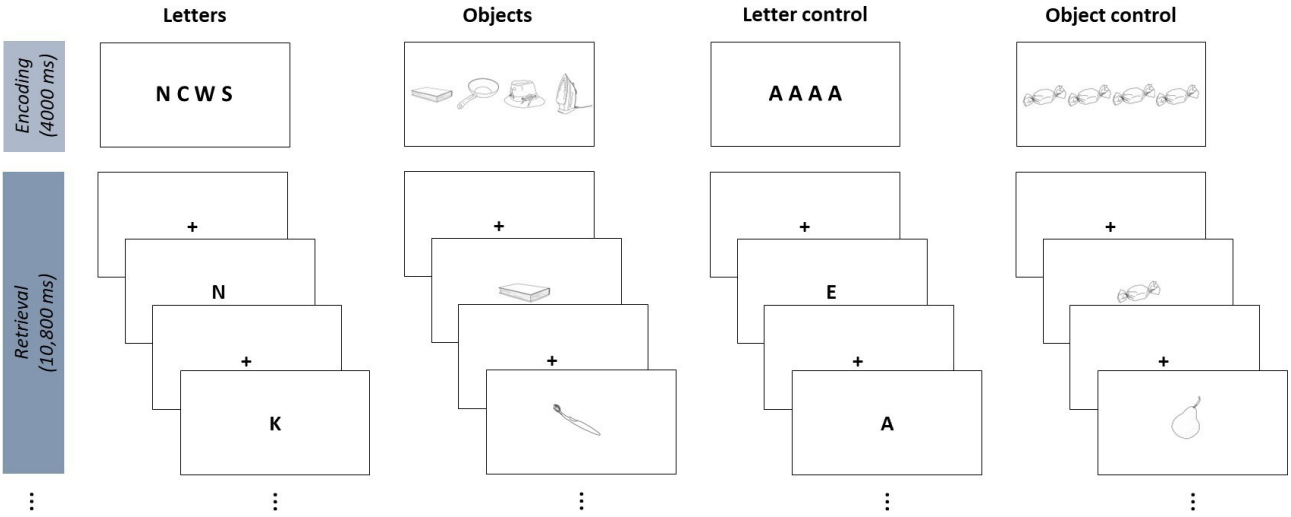

Each trial consists of an encoding phase and a retrieval phase. During the encoding phase, a set of 4 letters or objects (memory set) is presented and has to be memorized. After a variable interstimulus interval (1000 - 1600 ms, randomly jittered in steps of 200 ms), 4 single-letter probes (2 targets, 2 nontargets) are presented for 1800 ms each in the retrieval phase (each followed by a fixation cross for 1200 ms), and the patient has to indicate after each single-letter probe whether the probe was part of the previously shown memory set or not. Patients have to press the left button on an MRI-compatible response pad to indicate that the single-letter probe was part of the memory set or press the right button if it was not. The total duration time of the working memory task is approximately 6.45 minutes. Figure 2 illustrates the conditions and trials of the working memory task.

Statistical Analysis

Multicomponent Neuropsychological Assessment

In general, analysis will be conducted using SPSS Statistics (version 28.0.0.0; IBM Corp) [66] and MATLAB (R2024a; MathWorks Inc) [67].

In a first step, the analysis of the abovementioned tests, questionnaires, and clinical scores will be conducted according to the applicable manual. Raw scores will be transformed into norm-referenced *z* scores if standardizations are available.

In the second step, for the assessment of PSCI, a global and a domain-specific PSCI score will be defined in accordance with the guidelines of the International Society for Vascular Behavioral and Cognitive Disorders [68] and similarly to previous studies [39,69-72]. Further, the distinction between average cognitive functioning and mild, moderate, or severe cognitive impairment will be made according to the performance level on norm-referenced *z* scores, which is in line with the recommendations of multiple neuropsychological societies [73]. In detail, PSCI is a 5-level performance score. Performance level within 1 SD below or above the mean is defined as average cognitive functioning (PSCI=0). A range between 1 and 2 SDs below the mean (between third and 16th percentile) on at least 1 cognitive domain is defined as mild cognitive impairment (PSCI=1). A performance level 2 SDs below the mean (lower than the third percentile) in at least 2 domains is defined as moderate (PSCI=2), while a performance level 3 SDs below the mean (lower than the first percentile) in at least 2 cognitive domains or 2 SDs below the mean in most of the domains is defined as severe cognitive impairment (PSCI=3). Due to increased mortality in the PSD group, and to accommodate clinical realities as well as access feasibility of our study protocol, we use an intention-to-treat analysis for the primary end point. Therefore, we have added a level (PSCI=4), which reflects the death of a patient before the 3-month follow-up session. The following domains are defined: attention and information speed, executive functions, verbal memory, language, visuospatial construction, and orientation (Table 1). If a cognitive domain is represented by more than 1 test, the mean of the *z* scores will be used. We do not apply explicit weighting to individual test components when creating the composite scores for global and domain-specific poststroke cognitive function to ensure that each test within a domain contributes equally to the composite score, maintaining consistency with established neuropsychological standards. The primary endpoint will be a between-group comparison of the cognitive outcome, defined as global PSCI at a 3-month follow-up (S3). The groups will also be balanced for age and sex. Secondary outcome measures will be a between-group comparison of the domain-specific PSCI at a 3-month follow-up (S3).

Further, multivariate linear regression models will be used to analyze multi-item prognostic or predictive tools for cognitive outcome (eg, global PSCI) at the follow-up measurement (S3) by multimodal parameters assessed during the acute phase poststroke (S1), such as NIHSS, acute treatment, and working memory performance during fMRI or fMRI connectivity measures. As some tests will be digitalized using a tablet and a digitizer pencil, new parameters representing search behavior, for example, time on surface and time in air [42,74], will be recorded and might help to differentiate between PSD and non-PSD groups or identify patients with stroke with a high delirium risk.

XSL•FO
RenderX

Missing data will be handled in the following manner. For missing data in a 2-test subdomain, we plan to use the *z* score of the other test within the same subdomain. For missing data in a single test subdomain, we plan to use a multiple imputation approach or maximum likelihood estimation.

Neuroimaging Data

The preprocessing of the structural and functional data will be performed using standard routines of the Statistical Parametric Mapping software (SPM12) [75] within MATLAB [67]. In the first step, lesions are delineated on T2-weighted (fluid-attenuated inversion recovery) or T1*-weighted (T1-weighted magnetization-prepared rapid gradient-echo sequence) images. In the second step, the raw data will be preprocessed including the following steps: realignment, slice timing, coregistration, normalization, and spatial smoothing. This step includes the lesion correction during the normalization process to ensure accurate spatial alignment across subjects, using standard delineation and lesion correction methods [76-78]. In the third step, first-level analysis will be conducted using a general linear model including the following 8 task regressors: 4 regressors for each condition (ie, LE, LC, object condition, and object control condition) in the encoding phase and 4 regressors for each condition (ie, LE, LC, object condition, and object control condition) in the retrieval phase. These regressors will be modeled using a delta (stick) function for the encoding phase (to best account for phasic changes in the neural response to the target set) and a boxcar function for the retrieval phase. Besides brain activation, we also plan to analyze functional connectivity within the frontal-parietal networks or ascending reticular activating system. For the diffusion tensor imaging data, fractional anisotropy, axial or radial diffusivity, mean diffusivity, and so forth will be used to analyze differences in structural connectivity.

For the retrieval phase of the working memory task, the mean reaction time for target and nontarget probes will be calculated for all the 4 task conditions. Accuracy will be calculated for each condition as a percentage of correctly identified probes. Presentation (version 23.0; Neurobehavioral Systems, Inc) [65] will be used for calculating these performance parameters. Further analysis steps will be conducted using SPSS Statistics (version 28.0.0.0; IBM Corp) [66] and MATLAB (R2024a) [67], for example, to compare the performance parameters of the task conditions between the PSD and non-PSD group, for example, multifactor ANOVA.

Results

This study was funded in December 2023. The recruitment of patients was initiated in June 2024 and will continue until September 2025. As of May 2025, 15 patients with stroke have been included. Data acquisition will continue until December 2025. Data analysis will commence at the beginning of 2026, and study findings are expected to be published in summer 2026.

Discussion

Here, we present the protocol of a longitudinal study (RECOVER) designed to investigate both short-term and long-term effects of PSD on cognitive impairment and recovery in patients with stroke. PSCI is assessed using a multimodal approach that includes a multicomponent digital neuropsychological assessment and task-based fMRI at 3 time points: during the acute (prior PSD development) phase, the postacute phase, and at the 3-month follow-up.

Prior behavioral studies showed significantly poorer cognitive outcome (eg, in global cognition or memory decline) in both patients with stroke with PSD and without PSD compared to healthy controls [1,24,33]. Thus, a clear distinction between “stroke-related” and “PSD-related” cognitive impairments has not been established so far.

Brain neuroimaging studies using structural MRI reported dysconnectivity within the cortical-subcortical systems, for example, reticulo-thalamo-cortical network [1,13,79] in patients with delirium. fMRI studies using resting state have further revealed disrupted connectivity in dorsal-prefrontal and subcortical networks in patients with delirium (compared to healthy controls). These alterations may be linked to core delirium symptoms, such as disorientation, altered arousal, attention deficits, or disturbances in the sleep-wake cycle [1,13,16,28,30,80,81].

Memory is a fundamental aspect of human cognition and a frequently affected domain in patients with stroke with PSD and without PSD [22]. However, memory impairments may not be detectable during resting-state fMRI [31]. Unfortunately, task-based fMRI studies that allow for the direct assessment of memory functions are currently very limited. To date, only 1 fMRI study has investigated this challenging area, reporting no significant association between delirium and brain region activity [32]. However, this study raises methodological concerns due to the absence of a control group and the use of an infeasible working memory task. Further research is therefore needed, alongside the development of an adequate and reliable assessment of working memory during fMRI.

Finally, there remains a notable gap in the literature regarding appropriate methods for studying PSD and related cognitive outcomes. This gap has been addressed by this longitudinal study protocol, which includes 3 measurement points and 2 groups of patients with stroke, PSD, and non-PSD (control), to assess PSCI. We followed a multicomponent assessment strategy for the neuropsychological assessment and introduced a suitable digital tablet app tailored for both patients with stroke with PSD and without PSD. This partly tablet−based neuropsychological assessment enables the generation of additional neuropsychological parameters, such as time in air or time on surface [42,74], which may help to detect patients with PSD at an early stage.

In addition, we present an extended (ie, verbal and visual items) and modified version of a well-known working memory task (Sternberg task) [58-61] to examine patients with stroke with PSD and without PSD. This study design enables the direct assessment of memory functions at both the behavior level via neuropsychological testing and advanced parameters derived from digitalized tests, and at the brain level via task-based fMRI, where patients actively perform the working memory task during scanning. Therefore, we can map the neural correlates of

XSL•FO
RenderX

working memory functions and conduct brain-behavior correlation analyses. Additionally, this comprehensive longitudinal design, involving 2 patient groups (PSD and non-PSD), may allow us to differentiate between stroke-related and PSD-related cognitive impairments. Moreover, we aim to contribute preliminary insights toward developing protective and prognostic tools that support the early detection and treatment of PSD, ultimately helping to improve long-term cognitive recovery in patients with stroke.

The findings and experience gained from this proof-of-concept protocol will serve as early, preliminary investigation and provide essential groundwork for the design and implementation of future multicenter studies with a larger sample size. This protocol represents a first step toward expanding previous knowledge and improving the understanding of the contribution of PSD to long-term cognitive impairment after stroke.

These early findings from the proof-of-concept study hold promising clinical implications for the diagnosis and treatment of PSD, with further research required to validate these results on a larger scale. A more profound understanding of the cognitive dysfunctions associated with PSD could significantly influence clinical decision-making and rehabilitation strategies. Recognizing PSD as a distinct clinical condition characterized by specific patterns of cognitive impairment could help clinicians differentiate it from typical poststroke cognitive dysfunction. Specifically, the early identification of PSD-related cognitive deficits may enable the timely initiation of targeted, high-intensity cognitive rehabilitation, including domain-specific interventions such as those focused on attention, memory, and executive function. This approach has the potential to guide the development of individualized care plans tailored to patients' needs, prevent long-term cognitive decline, and promote optimal cognitive recovery.

However, this study protocol has several limitations. First, our study population consisted of patients with stroke with mild-to-moderate symptoms (NIHSS ≤15) and excluded patients with stroke with more severe stroke symptoms, particularly those with a GCS score ≤8, as well as patients with aphasia, which limits the generalizability of our findings to the broader population with stroke. However, patients with global aphasia are often unable to comprehend test instructions or follow simple requests (eg, "close your eyes"), making the assessment of cognitive functions, such as memory functions, extremely challenging, time-consuming, and often frustrating for the patient. Patients with a GCS score ≤ 8 have typically significantly reduced levels of consciousness, making a reliable cognitive assessment not possible. Second, the sample size is 40. Although this proof-of-concept study is underpowered, it is essential for gathering preliminary data and assessing feasibility. Insights gained will provide critical groundwork for designing larger, adequately powered studies in the future. Another important limitation of our study protocol is the variable timing of assessments, especially for S2, which is scheduled within a broad window of 5-14 days post stroke. These flexible time windows were intentionally selected to accommodate clinical realities and to minimize dropout rates and missing data. Regarding S2, this variability is driven by clinical realities, including patient transfers on short notice, fluctuating medical conditions, limited MRI availability, and the unpredictable course of delirium recovery. Unfortunately, the MRI assessment cannot take place during the acute phase of a delirium but only after partial recovery from delirium, a common problem in behavioral and neuroimaging studies on delirium or PSD. While this flexible scheduling approach enhances the feasibility of conducting comprehensive neuropsychological and imaging assessments in a population with acute stroke, it introduces potential confounding effects related to differences in the timing of measurements across patients. To mitigate this issue, we will incorporate the number of days post stroke as a covariate in our statistical analyses. Future large-scale studies should consider using more standardized normalization strategies to handle timing variability at S2, such as time normalization or statistical correction for nonlinear recovery effects in mixed effects models.

Another concern in our multimodal longitudinal study is the increased risk of participant dropout compared to cross-sectional studies. This risk is particularly pronounced among patients with PSD, given the association between PSD, lower quality of life, and higher mortality rates, which may lead to a higher than anticipated dropout rate. To address this, we have implemented several mitigation strategies. First, as described above, we have established predefined time windows to accommodate clinical realities within this vulnerable study population while prioritizing patients' well-being. Assessments are adjusted within predefined time windows based on the patient's condition, balancing protocol adherence with clinical care and comfort. Second, both neuropsychological and MRI assessments include opportunities for breaks or pauses to manage patient fatigue or distress. Third, when the patient's condition or transfer prevents in-hospital assessment, we offer follow-up visits at patients' homes or rehabilitation centers. Collectively, these measures enhance feasibility and data completeness while ensuring patient safety and comfort.

This study protocol involves important ethical considerations due to the vulnerability of patients with acute stroke, particularly those experiencing PSD. Ensuring ethical responsibility and patient safety is of the utmost importance. In an effort to minimize the burden on participants, both neuropsychological testing and MRI sessions are kept as brief as possible and can be paused or rescheduled based on the patient's condition. Testing is immediately halted if any signs of distress or fatigue arise. Throughout the entirety of the procedures, the presence and attentiveness of clinical staff are maintained. The inclusion of cognitively vulnerable patients is ethically justified, as this group is often underrepresented in research. However, their participation is carefully balanced with stringent safeguards to prevent any undue burden on their health or recovery. The presented protocol is critical to provide a comprehensive framework for assessing cognitive outcomes tailored to the challenges of evaluating older patients with stroke with and without PSD.

While this was beyond the scope of this proof-of-concept study, future studies should be initiated as multicenter trials to enhance the generalizability and robustness of the findings. Increasing the sample size will be essential to achieve adequate group sizes and improve statistical power. Additionally, balancing groups

XSL•FO
RenderX

according to key variables, such as lesion location or stroke severity, would strengthen the validity of comparisons.

Another important aspect is the consideration of the amount of cognitive rehabilitation on PSCI. Unfortunately, this was beyond the scope of this study, and this protocol includes only documentation of whether rehabilitation was received and its duration. A detailed evaluation of the effect of cognitive rehabilitation impact on long-term outcomes should be addressed in future studies. Another consideration would be extended follow-ups, for example, after 6 and 12 months, to better capture cognitive trajectories.

This protocol represents a proof-of-concept study designed to systematically and comprehensively examine cognitive functions in critically ill patients with acute stroke with and without delirium. It combines comprehensive neuropsychological testing with a multimodal MRI protocol, including digitalized assessments and task-based functional MRI. By addressing a significant clinical gap and using innovative methodologies, this study lays essential groundwork toward larger, adequately powered studies.

Acknowledgments

The authors thank Jamila Andoh and Michael Erb for their valuable assistance regarding the setup of the experimental task and the magnetic resonance imaging protocol.

Funding

This study is supported by the AKF program of the Medical Faculty of the University of Tuebingen (grant 1004-0-0). The grant application included an independent peer-review process. Publication costs have been partially covered by the Open Access Publishing Fund of University of Tuebingen.

Data Availability

The datasets generated during the current study are not publicly available due to General Data Protection Regulation restrictions but are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: LSG, AM
Investigation: LSG, SK, CS, CM, CG, AJ
Data curation: LSG, SK, CM, CG, AJ
Data analysis: LSG, SK, CM, CG
Research design: LSG, AM
Supervision: AM, KF, UZ
Visualization: LSG
Writing – original draft: LSG, AM
Writing – review & editing: All coauthors

Conflicts of Interest

KF received grants from the University of Tübingen (AKF) and speakers' honoraria/consulting fees from AstraZeneca and BMS/Pfizer, all not related to this work. The rest of the authors declare no conflicts of interest.

References

1. Wilson JE, Mart MF, Cunningham C, et al. Delirium. Nat Rev Dis Primers 2020 Nov 12;6(1):90. [doi: 10.1038/s41572-020-00223-4] [Medline: 33184265]
2. Shaw RC, Walker G, Elliott E, Quinn TJ. Occurrence rate of delirium in acute stroke settings: systematic review and meta-analysis. Stroke 2019 Nov;50(11):3028-3036. [doi: 10.1161/STROKEAHA.119.025015] [Medline: 31554501]
3. Ye F, Ho MH, Lee JJ. Prevalence of post-stroke delirium in acute settings: a systematic review and meta-analysis. Int J Nurs Stud 2024 Jun;154:104750. [doi: 10.1016/j.ijnurstu.2024.104750] [Medline: 38552472]
4. Witlox J, Eurelings LSM, de Jonghe JFM, Kalisvaart KJ, Eikelenboom P, van Gool WA. Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. JAMA 2010 Jul 28;304(4):443-451. [doi: 10.1001/jama.2010.1013] [Medline: 20664045]
5. Shaw R, Drozdowska B, Taylor-Rowan M, et al. Delirium in an acute stroke setting, occurrence, and risk factors. Stroke 2019 Nov;50(11):3265-3268. [doi: 10.1161/STROKEAHA.119.025993] [Medline: 31554500]
6. Zipser CM, Deuel JW, Held JPO, et al. Economic impact of poststroke delirium and associated risk factors. Stroke 2021 Oct;52(10):3325-3334. [doi: 10.1161/STROKEAHA.120.033005]
7. Shi Q, Presutti R, Selchen D, Saposnik G. Delirium in acute stroke: a systematic review and meta-analysis. Stroke 2012 Mar;43(3):645-649. [doi: 10.1161/STROKEAHA.111.643726] [Medline: 22267831]

XSL•FO
RenderX

8. Davis DHJ, Muniz-Terrera G, Keage HAD, et al. Association of delirium with cognitive decline in late life: a neuropathologic study of 3 population-based cohort studies. JAMA Psychiatry 2017 Mar 1;74(3):244-251. [doi: 10.1001/jamapsychiatry.2016.3423] [Medline: 28114436]
9. Aung Thein MZ, Pereira JV, Nitchingham A, Caplan GA. A call to action for delirium research: meta-analysis and regression of delirium associated mortality. BMC Geriatr 2020 Sep 7;20(1):325. [doi: 10.1186/s12877-020-01723-4] [Medline: 32894065]
10. Droś J, Segiet N, Początek G, Klimkowicz-Mrowiec A. Five-year stroke prognosis. Influence of post-stroke delirium and post-stroke dementia on mortality and disability (Research Study-Part of the PROPOLIS Study). Neurol Sci 2024 Mar;45(3):1109-1119. [doi: 10.1007/s10072-023-07129-5] [Medline: 37851292]
11. Gong X, Jin S, Zhou Y, Lai L, Wang W. Impact of delirium on acute stroke outcomes: a systematic review and meta-analysis. Neurol Sci 2024 May;45(5):1897-1911. [doi: 10.1007/s10072-023-07287-6] [Medline: 38182844]
12. van Montfort SJT, van Dellen E, Stam CJ, et al. Brain network disintegration as a final common pathway for delirium: a systematic review and qualitative meta-analysis. Neuroimage Clin 2019;23:101809. [doi: 10.1016/j.nicl.2019.101809] [Medline: 30981940]
13. Nitchingham A, Kumar V, Shenkin S, Ferguson KJ, Caplan GA. A systematic review of neuroimaging in delirium: predictors, correlates and consequences. Int J Geriatr Psychiatry 2018 Nov;33(11):1458-1478. [doi: 10.1002/gps.4724] [Medline: 28574155]
14. Fleischmann R, Warwas S, Andrasch T, et al. Course and recognition of poststroke delirium: a prospective noninferiority trial of delirium screening tools. Stroke 2021 Jan;52(2):471-478. [doi: 10.1161/STROKEAHA.120.031019] [Medline: 33380165]
15. Boßelmann C, Zurloh J, Stefanou MI, et al. Delirium screening in aphasic patients with the Intensive Care Delirium Screening Checklist (ICDSC): a prospective cohort study. Front Neurol 2019;10:1198. [doi: 10.3389/fneur.2019.01198] [Medline: 31781026]
16. Boukrina O, Barrett AM. Disruption of the ascending arousal system and cortical attention networks in post-stroke delirium and spatial neglect. Neurosci Biobehav Rev 2017 Dec;83:1-10. [doi: 10.1016/j.neubiorev.2017.09.024] [Medline: 28963037]
17. Rost NS, Brodtmann A, Pase MP, et al. Post-stroke cognitive impairment and dementia. Circ Res 2022 Apr 15;130(8):1252-1271. [doi: 10.1161/CIRCRESAHA.122.319951] [Medline: 35420911]
18. Wiredu K, Aduse-Poku E, Shaefi S, Gerber SA. Proteomics for the discovery of clinical delirium biomarkers: a systematic review of major studies. Anesth Analg 2023;136(3):422-432. [doi: 10.1213/ANE.0000000000006246]
19. Mukminin MA, Yeh TH, Lin HC, Rohmah I, Chiu HY. Global prevalence and risk factors of delirium among patients following acute stroke: a systematic review and meta-analysis. J Stroke Cerebrovasc Dis 2025 Feb;34(2):108221. [doi: 10.1016/j.jstrokecerebrovasdis.2024.108221] [Medline: 39756631]
20. Kneihsl M, Berger N, Sumerauer S, et al. Management of delirium in acute stroke patients: a position paper by the Austrian Stroke Society on prevention, diagnosis, and treatment. Ther Adv Neurol Disord 2024;17:17562864241258788. [doi: 10.1177/17562864241258788] [Medline: 39161955]
21. Rhee JY, Colman MA, Mendu M, et al. Associations between stroke localization and delirium: a systematic review and meta-analysis. J Stroke Cerebrovasc Dis 2022 Mar;31(3):106270. [doi: 10.1016/j.jstrokecerebrovasdis.2021.106270] [Medline: 34954599]
22. Chiti G, Pantoni L. Use of Montreal cognitive assessment in patients with stroke. Stroke 2014 Oct;45(10):3135-3140. [doi: 10.1161/STROKEAHA.114.004590] [Medline: 25116881]
23. Stolwyk RJ, Mihaljcic T, Wong DK, Hernandez DR, Wolff B, Rogers JM. Post-stroke cognition is associated with stroke survivor quality of life and caregiver outcomes: a systematic review and meta-analysis. Neuropsychol Rev 2024 Dec;34(4):1235-1264. [doi: 10.1007/s11065-024-09635-5] [Medline: 38466357]
24. Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. N Engl J Med 2013 Oct 3;369(14):1306-1316. [doi: 10.1056/NEJMoa1301372] [Medline: 24088092]
25. Müller A, von Hofen-Hohloch J, Mende M, et al. Long-term cognitive impairment after ICU treatment: a prospective longitudinal cohort study (Cog-I-CU). Sci Rep 2020 Sep 23;10(1):15518. [doi: 10.1038/s41598-020-72109-0] [Medline: 32968099]
26. Pendlebury ST, Thomson RJ, Welch SJV, Rothwell PM, Oxford Vascular Study. Cognitive predictors of delirium on long-term follow-up after TIA and stroke: population-based cohort study. Cerebrovasc Dis 2022;51(3):288-295. [doi: 10.1159/000519900] [Medline: 34844242]
27. Morandi A, Rogers BP, Gunther ML, et al. The relationship between delirium duration, white matter integrity, and cognitive impairment in intensive care unit survivors as determined by diffusion tensor imaging: the VISIONS prospective cohort magnetic resonance imaging study*. Crit Care Med 2012 Jul;40(7):2182-2189. [doi: 10.1097/CCM.0b013e318250acdc] [Medline: 22584766]
28. Haggstrom L, Welschinger R, Caplan GA. Functional neuroimaging offers insights into delirium pathophysiology: a systematic review. Australas J Ageing 2017 Sep;36(3):186-192. [doi: 10.1111/ajag.12417] [Medline: 28519903]
29. van den Boogaard M, Slooter AJC. Delirium in critically ill patients: current knowledge and future perspectives. BJA Educ 2019 Dec;19(12):398-404. [doi: 10.1016/j.bjae.2019.09.004] [Medline: 33456864]

XSL•FO
RenderX

30. Choi SH, Lee H, Chung TS, et al. Neural network functional connectivity during and after an episode of delirium. Am J Psychiatry 2012 May;169(5):498-507. [doi: 10.1176/appi.ajp.2012.11060976] [Medline: 22549209]
31. Vinehout K, Schindler-Ivens S, Binder JR, Schmit BD. Task effects on functional connectivity measures after stroke. Exp Brain Res 2022 Feb;240(2):575-590. [doi: 10.1007/s00221-021-06261-y] [Medline: 34860257]
32. Jackson JC, Morandi A, Girard TD, et al. Functional brain imaging in survivors of critical illness: a prospective feasibility study and exploration of the association between delirium and brain activation patterns. J Crit Care 2015 Jun;30(3):653. [doi: 10.1016/j.jcrc.2015.01.017] [Medline: 25769901]
33. Umarova RM, Sperber C, Kaller CP, et al. Cognitive reserve impacts on disability and cognitive deficits in acute stroke. J Neurol 2019 Oct;266(10):2495-2504. [doi: 10.1007/s00415-019-09442-6] [Medline: 31254064]
34. Lam JM, Globas C, Hosp JA, Karnath HO, Wächter T, Luft AR. Impaired implicit learning and feedback processing after stroke. Neuroscience 2016 Feb 9;314:116-124. [doi: 10.1016/j.neuroscience.2015.11.051] [Medline: 26633264]
35. Lyden P, Brott T, Tilley B, et al. Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. Stroke 1994 Nov;25(11):2220-2226. [doi: 10.1161/01.str.25.11.2220] [Medline: 7974549]
36. Briggs DE, Felberg RA, Malkoff MD, Bratina P, Grotta JC. Should mild or moderate stroke patients be admitted to an intensive care unit? Stroke 2001 Apr;32(4):871-876. [doi: 10.1161/01.str.32.4.871] [Medline: 11283385]
37. Ordies S, Peeters G, Lesenne A, et al. Interaction between stroke severity and quality indicators of acute stroke care: a single-center retrospective analysis. Acta Neurol Belg 2022 Feb;122(1):173-180. [doi: 10.1007/s13760-021-01811-1] [Medline: 34604947]
38. Teasdale G, Jennett B. Assessment of coma and impaired consciousness. a practical scale. Lancet 1974 Jul 13;2(7872):81-84. [doi: 10.1016/s0140-6736(74)91639-0] [Medline: 4136544]
39. Hachinski V, Iadecola C, Petersen RC, et al. National Institute of Neurological Disorders and Stroke-Canadian Stroke Network vascular cognitive impairment harmonization standards. Stroke 2006 Sep;37(9):2220-2241. [doi: 10.1161/01.STR.0000237236.88823.47] [Medline: 16917086]
40. Quinn TJ, Richard E, Teuschl Y, et al. European Stroke Organisation and European Academy of Neurology joint guidelines on post-stroke cognitive impairment. Eur J Neurol 2021 Dec;28(12):3883-3920. [doi: 10.1111/ene.15068] [Medline: 34476868]
41. TUCAN – Tübinger Cognitive Assessment for Neuropsychiatric Disorders [Website in German]. URL: https://tucantest.org [accessed 2025-12-08]
42. Polgar S, Mychajliw C, Wortha F, Ninaus M, Holz H. Towards the development of a tablet-based screening for cognitive symptoms of neuropsychiatric disorders. Presented at: MobileHCI '18; Sep 3-6, 2018. [doi: 10.1145/3236112.3236147]
43. Böttinger MJ, Litz E, Gordt-Oesterwind K, et al. Co-creating a digital life-integrated self-assessment for older adults: user experience study. JMIR Aging 2023 Sep 26;6:e46738. [doi: 10.2196/46738] [Medline: 37751274]
44. Mychajliw C, Holz H, Minuth N, et al. Performance differences of a touch-based serial reaction time task in healthy older participants and older participants with cognitive impairment on a tablet: experimental study. JMIR Aging 2024 Mar 21;7:e48265. [doi: 10.2196/48265] [Medline: 38512340]
45. Rosenzopf H, Sperber C, Wortha F, et al. Spatial neglect in the digital age: influence of presentation format on patients' test behavior. J Int Neuropsychol Soc 2023 Aug;29(7):686-695. [doi: 10.1017/S1355617722000790] [Medline: 36303420]
46. Reitan RM. Trail making test: trail making testmanual for administration and scoring. : Reitan Neuropsychological Laboratory; 1992.
47. Petermann F, Jäncke L, Waldmann HC. NAB - Neuropsychological Assessment Battery [Web page in German]. Testzentrale. URL: https://www.testzentrale.de/shop/neuropsychological-assessment-battery.html [accessed 2025-12-14]
48. Bartusch S, Zipper S. Montreal Cognitive Assessment (MoCA). MoCA Cognition. URL: https://mocacognition.com/ [accessed 2025-12-14]
49. Julayanont P, Brousseau M, Chertkow H, Phillips N, Nasreddine ZS. Montreal Cognitive Assessment Memory Index Score (MoCA-MIS) as a predictor of conversion from mild cognitive impairment to Alzheimer's disease. J Am Geriatr Soc 2014 Apr;62(4):679-684. [doi: 10.1111/jgs.12742] [Medline: 24635004]
50. Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The cognitive failures questionnaire (CFQ) and its correlates. Br J Clin Psychol 1982 Feb;21(1):1-16. [doi: 10.1111/j.2044-8260.1982.tb01421.x] [Medline: 7126941]
51. Ehrensperger MM, Berres M, Taylor KI, Monsch AU. Screening properties of the German IQCODE with a two-year time frame in MCI and early Alzheimer's disease. Int Psychogeriatr 2010 Feb;22(1):91-100. [doi: 10.1017/S1041610209990962] [Medline: 19747425]
52. Yesavage JA. Geriatric Depression Scale. Psychopharmacol Bull 1988;24(4):709-711. [Medline: 3249773]
53. Bullinger M, Kirchberger I. Fragebogen Zum Gesundheitszustand: SF-36: Hogrefe; 1998. [doi: 10.1026//0084-5345.28.2.143]
54. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. Scott Med J 1957 May;2(5):200-215. [doi: 10.1177/003693305700200504] [Medline: 13432835]
55. Radtke FM, Franck M, Oppermann S, et al. The Intensive Care Delirium Screening Checklist (ICDSC)—translation and validation of intensive care delirium checklist in accordance with guidelines. Anasthesiol Intensivmed Notfallmed Schmerzther 2009 Feb;44(2):80-86. [doi: 10.1055/s-0029-1202647] [Medline: 19199171]
56. Lehrl S. Mehrfachwahl-Wortschatz-Test (MWT-B): Straube; 1977.

XSL•FO
RenderX

57. Multi-band EPI. Minnesota CMRR. URL: https://www.cmrr.umn.edu/multiband/ [accessed 2025-12-10]
58. van Raalten TR, Ramsey NF, Jansma JM, Jager G, Kahn RS. Automatization and working memory capacity in schizophrenia. Schizophr Res 2008 Mar;100(1-3):161-171. [doi: 10.1016/j.schres.2007.10.035] [Medline: 18155446]
59. Geiger LS, Moessnang C, Schäfer A, et al. Novelty modulates human striatal activation and prefrontal-striatal effective connectivity during working memory encoding. Brain Struct Funct 2018 Sep;223(7):3121-3132. [doi: 10.1007/s00429-018-1679-0] [Medline: 29752589]
60. Sternberg S. High-speed scanning in human memory. Science 1966 Aug 5;153(3736):652-654. [doi: 10.1126/science.153.3736.652] [Medline: 5939936]
61. van Raalten TR, Ramsey NF, Duyn J, Jansma JM. Practice induces function-specific changes in brain activity. PLoS ONE 2008 Oct 1;3(10):e3270. [doi: 10.1371/journal.pone.0003270] [Medline: 18827897]
62. Geiger LS, Wüstenberg T, Zang Z, et al. Longitudinal markers of cognitive procedural learning in fronto-striatal circuits and putative effects of a BDNF plasticity-related variant. NPJ Sci Learn 2024 Nov 28;9(1):72. [doi: 10.1038/s41539-024-00282-2] [Medline: 39604428]
63. Brodeur MB, Dionne-Dostie E, Montreuil T, Lepage M. The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. PLoS ONE 2010 May 24;5(5):e10773. [doi: 10.1371/journal.pone.0010773] [Medline: 20532245]
64. Brodeur MB, Guérard K, Bouras M. Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. PLoS ONE 2014;9(9):e106953. [doi: 10.1371/journal.pone.0106953] [Medline: 25211489]
65. Neurobehavioral Systems. URL: https://www.neurobs.com [accessed 2025-12-08]
66. IBM SPSS Statistics. IBM. URL: https://www.ibm.com/products/spss-statistics [accessed 2025-12-10]
67. MathWorks. URL: https://mathworks.com [accessed 2025-12-14]
68. Sachdev P, Kalaria R, O'Brien J, et al. Diagnostic criteria for vascular cognitive disorders: a VASCOG statement. Alzheimer Dis Assoc Disord 2014;28(3):206-218. [doi: 10.1097/WAD.0000000000000034] [Medline: 24632990]
69. Coenen M, de Kort FA, Weaver NA, et al. Strategic white matter hyperintensity locations associated with post-stroke cognitive impairment: a multicenter study in 1568 stroke patients. Int J Stroke 2024 Oct;19(8):916-924. [doi: 10.1177/17474930241252530] [Medline: 38651756]
70. Biesbroek JM, Weaver NA, Aben HP, et al. Network impact score is an independent predictor of post-stroke cognitive impairment: a multicenter cohort study in 2341 patients with acute ischemic stroke. Neuroimage Clin 2022;34:103018. [doi: 10.1016/j.nicl.2022.103018] [Medline: 35504223]
71. Weaver NA, Kuijf HJ, Aben HP, et al. Strategic infarct locations for post-stroke cognitive impairment: a pooled analysis of individual patient data from 12 acute ischaemic stroke cohorts. Lancet Neurol 2021 Jun;20(6):448-459. [doi: 10.1016/S1474-4422(21)00060-0] [Medline: 33901427]
72. Ferreira J, Pereira G, Alves F, et al. Microemboli detection in acute ischemic stroke could be an early marker of poor cognitive outcome. Stroke 2025 Sep 30. [doi: 10.1161/STROKEAHA.125.052253] [Medline: 41025237]
73. Frei A, Balzer C, Gysi F, Leros J, Plohmann A, Steiger G. Kriterien zur Bestimmung des Schweregrades einer neuropsychologischen Störung sowie Zuordnungen zur Funktions- und Arbeitsfähigkeit. Z fur Neuropsychol 2016 Apr;27(2):107-119. [doi: 10.1024/1016-264X/a000177]
74. Müller S, Preische O, Heymann P, Elbing U, Laske C. Diagnostic value of a tablet-based drawing task for discrimination of patients in the early course of alzheimer's disease from healthy individuals. J Alzheimers Dis 2017;55(4):1463-1469. [doi: 10.3233/JAD-160921] [Medline: 27858717]
75. SPM12 software. Functional Imaging Laboratory. URL: http://www.fil.ion.ucl.ac.uk/spm/software/spm12 [accessed 2025-12-08]
76. Moore MJ, Demeyere N, Rorden C, Mattingley JB. Lesion mapping in neuropsychological research: a practical and conceptual guide. Cortex 2024 Jan;170:38-52. [doi: 10.1016/j.cortex.2023.10.001] [Medline: 37940465]
77. de Haan B, Clas P, Juenger H, Wilke M, Karnath HO. Fast semi-automated lesion demarcation in stroke. Neuroimage Clin 2015;9:69-74. [doi: 10.1016/j.nicl.2015.06.013] [Medline: 26413473]
78. Rorden C, Bonilha L, Fridriksson J, Bender B, Karnath HO. Age-specific CT and MRI templates for spatial normalization. NeuroImage 2012 Jul 16;61(4):957-965. [doi: 10.1016/j.neuroimage.2012.03.020] [Medline: 22440645]
79. Cavallari M, Dai W, Guttmann CRG, et al. Neural substrates of vulnerability to postsurgical delirium as revealed by presurgical diffusion MRI. Brain (Bacau) 2016 Apr;139(4):1282-1294. [doi: 10.1093/brain/aww010]
80. Boukrina O, Kowalczyk M, Koush Y, Kong Y, Barrett AM. Brain network dysfunction in poststroke delirium and spatial neglect: an fMRI study. Stroke 2022 Mar;53(3):930-938. [doi: 10.1161/STROKEAHA.121.035733] [Medline: 34619987]
81. van Montfort SJT, van Dellen E, van den Bosch AMR, et al. Resting-state fMRI reveals network disintegration during delirium. Neuroimage Clin 2018;20:35-41. [doi: 10.1016/j.nicl.2018.06.024] [Medline: 29998059]

Abbreviations

EPI: echo planar imaging sequence
fMRI: functional magnetic resonance imaging

XSL•FO
RenderX

FoV: field of view
ICDSC: Intensive Care Delirium Screening Checklist
LC: letter control condition
LE: letter condition
MRI: magnetic resonance imaging
NIHSS: National Institutes of Health Stroke Scale
PSCI: poststroke cognitive impairment
PSD: poststroke delirium
RECOVER: Cognitive Recovery After Poststroke Delirium
TE: echo time
TR: repetition time
TuCAN: Tübingen Cognitive Assessment for Neuropsychiatric Disorders

Edited by J Sarvestan; submitted 14.May.2025; peer-reviewed by N Sherry, X Ding; accepted 07.Nov.2025; published 12.Jan.2026.

Please cite as:
Geiger LS, Klütz S, Mychajliw C, Gäbele C, Jooß A, Single C, Feil K, Ziemann U, Mengel A
Cognitive Recovery After Poststroke Delirium (RECOVER): Protocol for a Longitudinal Multimodal Study on Cognitive Assessment in Patients With Stroke
JMIR Res Protoc 2026;15:e77508
URL: https://www.researchprotocols.org/2026/1/e77508
doi:10.2196/77508

© Lena S Geiger, Sebastian Klütz, Christian Mychajliw, Christoph Gäbele, Andreas Jooß, Constanze Single, Katharina Feil, Ulf Ziemann, Annerose Mengel. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Self-Management Combined With Digital Health Interventions to Improve Dietary Behavior, Exercise Behavior, Stress Management Behavior, and Blood Pressure Among Thais With Uncontrolled Hypertension: Protocol for an Explanatory Sequential Study

Saowaluck Sukpattanasrikul[1], PhD; Naruemol Singha-Dong[2], PhD; Yajai Sitthimongkol[3], PhD

[1]Department of Fundamental Nursing, Faculty of Nursing, Mahidol University, Nakhon Pathom, Thailand

[2]School of Community Health Nursing, Institute of Nursing, Suranaree University of Technology, 111 University Avenue, Suranari subdistrict, Muang, Nakhon Ratchasima, Thailand

[3]Department of Mental Health and Psychiatric Nursing, Faculty of Nursing, Mahidol University, Nakhon Pathom, Thailand

Corresponding Author:
Naruemol Singha-Dong, PhD
School of Community Health Nursing, Institute of Nursing, Suranaree University of Technology, 111 University Avenue, Suranari subdistrict, Muang, Nakhon Ratchasima, Thailand

Abstract

Background: Uncontrolled hypertension causes substantial morbidity and mortality as well as rising health care costs. Adherence to self-management is critical for minimizing the risk of hypertensive complications. Traditional self-management is hindered by delayed management and insufficient support. Digital health interventions offer a feasible solution for closing these gaps and enhancing hypertension self-management. Little is known about patients' perspectives and experiences concerning how digital health interventions influence their self-management.

Objective: This study aims to determine the effectiveness of a self-management intervention combined with digital health interventions on dietary behavior, exercise behavior, stress management behavior, and blood pressure among Thais with uncontrolled hypertension. The study explored and compared experiences and perceptions among participants with varying levels of blood pressure control.

Methods: This study uses an explanatory sequential design performed in 2 phases comprising (1) a quasi-experimental design with 2 groups using repeated measures to determine the effects of self-management combined with digital health intervention and (2) an in-depth interview approach to explore the perceptions and experiences of 24 participants regarding the combination of self-management and digital health interventions after the intervention. In phase 1, participants were allocated by lottery to either the intervention group, which underwent an 8-week self-management intervention combined with digital health interventions, or the control group. The Dietary Approaches to Stop Hypertension Questionnaire, the Exercise Behavior Questionnaire, the Brief COPE inventory (Thai Version), and an automatic blood pressure measurement were used for data collection at baseline and the 4th and 8th weeks. In phase 2, semistructured interviews were used to conduct in-depth interviews. The analysis will take into account the effects of the interventions on dietary, exercise, and stress management behaviors, as well as blood pressure, using generalized estimating equations and linear mixed-effects modeling. We will perform the method by Colaizzi for the qualitative portion of the analysis.

Results: Funded in December 2024, this study recruited 86 patients with uncontrolled hypertension at the Siriraj Primary Care Unit, Thailand. This study received ethical approval on May 31, 2025, and participant recruitment began in August 2025. In phase 1, this study began recruiting participants in August 2025, with data collection occurring from August through the first half of November 2025. Phase 2 was completed at the end of November 2025. The data analysis is expected to be completed by December 2025. The expected date for the results to be submitted for publication is March 2026.

Conclusions: This study has the potential to address the gap between traditional self-management and digital health interventions for improving self-management behaviors and reducing blood pressure. The findings may offer practical guidance for nurses and other health care providers for managing uncontrolled hypertension in Thailand and contribute valuable insights for shaping future health care policies.

Trial Registration: Thai Clinical Trials Registry (TCTR), via the WHO International Clinical Trials Registry Platform (ICTRP), TCTR20250722001; https://www.thaiclinicaltrials.org/show/TCTR20250722001

International Registered Report Identifier (IRRID): PRR1-10.2196/81148

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e81148) doi:10.2196/81148

KEYWORDS

uncontrolled hypertension; self-management; digital health intervention; web-based application; dietary behavior; exercise behavior; stress management behavior; blood pressure; explanatory sequential design

Introduction

Uncontrolled hypertension (blood pressure [BP]≥140/90 mm Hg) is a substantial health problem that presents a challenge for the global health system, including in Thailand. Approximately 26% of patients with hypertension worldwide have uncontrolled hypertension [1]. Thailand follows this trend, with 24.6% of the population with hypertension having uncontrolled hypertension, often linked to lifestyle-related factors [1,2]. In Thailand, 15% of Thais with uncontrolled hypertension experience complications such as cardiovascular disease, stroke, or chronic kidney disease [3]. Moreover, hypertension is associated with an average cost of 964 (US $30.96) per visit, resulting in a total expenditure of 1.4367 billion (US $46,135,167) [4]. In Thailand, like most other countries, uncontrolled hypertension is associated with poor self-management, such as an unhealthy diet, a lack of exercise, and inappropriate stress management [5-7]. A study by Sodkhomkham [8] found that self-management, including dietary behavior, exercise behavior, stress management, and medication adherence, could predict BP control with an accuracy of 82%. Thus, focusing on self-management is indispensable for addressing uncontrolled hypertension.

Self-management is widely recognized as a pivotal aspect of improving self-management behaviors and clinical outcomes in hypertensive care [9]. Successful hypertension self-management requires interactive group education on hypertension and its management, which can be linked to the condition, monitoring with feedback, the provision of equipment, and lifestyle advice with support related to diet and exercise, weight reduction, stress management, and medication adherence [9]. Nurses play a crucial role in supporting patients with uncontrolled hypertension by providing health education, counseling, coaching, and facilitation of support [10-12]. However, a study by Liang et al [13] illustrated that traditional hypertension self-management is less effective due to the barriers to communication between health care providers and patients, a lack of timely and detailed management, and a lack of instrumental support. Likewise, time constraints and heavy workloads hinder nurses' ability to prioritize hypertension management while providing patient care [14]. To overcome these limitations, digital health interventions (DHIs) serve as a complementary method to support self-management [15].

DHIs have emerged as effective, beneficial, and widely accessible strategies for supporting hypertension self-management [16,17] and promoting health equity for patients with uncontrolled hypertension [17]. The World Health Organization (WHO) [18] emphasized the importance of patient-centered DHIs, including interactive voice response systems, phone calls, web-based telecare platforms, smartphone apps, SMS, and multimedia messaging services. Nurses play a pivotal role in promoting self-management to prevent and control hypertension. They empower individuals to adopt and maintain healthy behaviors through comprehensive education, personalized goal setting, and ongoing support [19]. Nurses meticulously assess risk factors, provide tailored counseling, and leverage digital health tools to help patients effectively monitor their conditions [20]. By consistently following up with patients and coordinating care, nurses enhance patient engagement, adherence, and, ultimately, long-term health outcomes [20]. An updated systematic review by Sukpattanasrikul et al [21] highlighted that the most effective DHIs for patients with uncontrolled hypertension target multiple health behaviors, incorporating key components such as health education, reminders, self-monitoring, feedback, and instrumental support.

Investigation into previous self-management combined with DHIs used for patients with uncontrolled hypertension between 2014 and 2024 revealed that most previous studies conducted research in an effort to control hypertension through self-management education, including health education related to the disease and self-care management [22-24]; self-management skill development through goal setting, decision-making, action plans, self-monitoring, and self-evaluation [23,25,26]; and providing instrumental support such as automatic BP measurement [27], salt meters [28], and graphics displaying BP over time [29]. Concurrently, DHIs being used to support self-management include health education [22,30], medication reminders [22,30-33], and BP monitoring reminders [27,31,33,34], as well as consultation through social media platforms such as WeChat [35] and Line applications [32].

This study identified a gap in knowledge among Thai patients with uncontrolled hypertension concerning self-management combined with DHIs by reviewing the literature over 1 decade (2014 - 2024). Traditional self-management approaches within service units have limitations due to their group-based nature. They may not adequately address the specific challenges and requirements of patients with uncontrolled hypertension. Additionally, providing consultation through the Line application increases the workload for health care personnel, who must manually review information from patients with hypertension and respond [19]. Integrating automated response programs (chatbots) into consultations can help patients with uncontrolled hypertension seek advice or consultations tailored specifically to their needs and receive it more quickly [19].

Literature reviews have also shown that providing health education through applications alone increases patients' knowledge of hypertension [36] but fails to control BP levels within the recommended range (BP<140/90 mm Hg) [37]. Meanwhile, the perceptions and experiences of patients regarding the use of DHIs to support self-management processes remain unclear. Thus, this study addresses the knowledge gap by implementing multiple interventions that integrate

XSL•FO
RenderX

self-management with DHIs for patients with uncontrolled hypertension, emphasizing a hybrid reinforcement model. This approach has the potential to offer nurses and other health care providers practical guidance on how to care for patients with uncontrolled hypertension, thereby improving their hypertensive outcomes. This study could also provide important information leading to the further development of health care policies in Thailand.

The objectives of this study are (1) to determine the dietary, exercise, and stress management behaviors, as well as BP, among Thais with uncontrolled hypertension after receiving a self-management intervention combined with DHIs and (2) to explore their experiences and perceptions after receiving the interventions. More specifically, this study aims to (1) evaluate the effectiveness of self-management combined with DHIs for improving dietary, exercise, and stress management behaviors, as well as BP, among Thais with uncontrolled hypertension and (2) explore the perceptions and experiences of study participants regarding the combined self-management and DHIs after the intervention and to compare these perceptions and experiences with varying levels of BP control.

Methods

Overview

The study uses an explanatory sequential design to determine the effectiveness of self-management combined with DHIs among patients with uncontrolled hypertension, exploring the experiences and perceptions after receiving the intervention among patients with uncontrolled hypertension who received treatment at Siriraj Primary Care Unit. The study achieves its objectives in 2 phases: phase 1, a quasi-experimental design with 2 groups and repeated measures, and phase 2, an in-depth interview approach.

The integration of the 2 phases aligns with an explanatory sequential design. First, the connecting integration was applied to the sampling design. The researcher determined the participants in phase 2 by purposively selecting them based on phase 1 results (eg, BP≥140/90 mm Hg or BP<140/90 mm Hg) and age group, to ensure we capture a range of perceptions and experiences. Second, after both phases are completed, the integration of results will take place. As to the purpose of this protocol, the qualitative themes are used to explain dietary, exercise, and stress management behaviors, as well as BP, among Thais with uncontrolled hypertension. Joint displays and narratives will integrate quantitative and qualitative information, integrating outcome patterns with significant themes to draw conclusions on how self-management and DHIs impact outcomes (Figure 1).

XSL•FO
RenderX

Figure 1. Explanatory sequential design. DHI: digital health intervention.

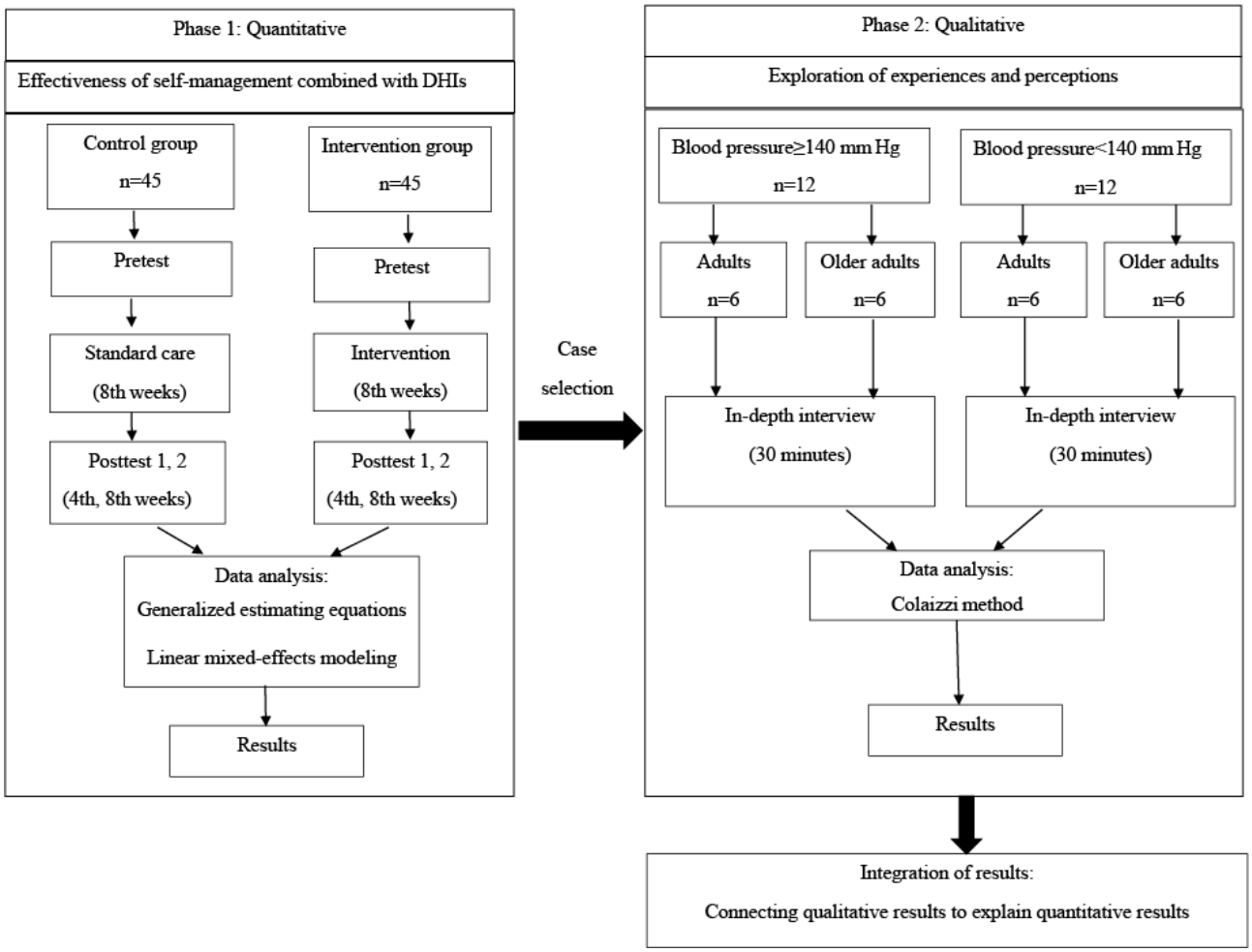

Phase 1: Quantitative Method to Evaluate the Effectiveness of Self-Management Combined With DHIs

This study involves a quasi-experimental design comparing 2 groups with repeated measures to determine the effectiveness of an 8-week self-management program combined with DHIs on dietary behavior, exercise behavior, stress management behavior, and BP.

Recruitment

Participant recruitment included poster advertisements at the Siriraj Primary Care Unit, plus an announcement of the invitation at the Siriraj Primary Care Unit. Patients with uncontrolled hypertension who were interested in participating in the project could express their interest to the nurse at the Siriraj Primary Care Unit or the researchers. The researchers then screened patients with hypertension according to the inclusion and exclusion criteria. The study approached those who passed all inclusion criteria individually to sign an informed consent document. Participants who met the inclusion and exclusion criteria were allocated to either the control or experimental group. The recruitment protocol enphasized that participants could opt out at any time before, during, or after participation, without consequences from the Siriraj Primary Care Unit. In addition, participants who attended fewer than 75% of the activities (less than 6 weeks out of 8 weeks) were withdrawn from this study without consequences from the Siriraj Primary Care Unit.

Participants

The study selected participants by proportional stratified random sampling, followed by purposive sampling adhering to the inclusion criteria. The inclusion criteria were ages 45 years to 79 years old, previous history of a diagnosis of mild to moderate hypertension for at least 1 year and recorded in the database at the Siriraj Primary Care Unit, a determination of uncontrolled hypertension according to the BP criteria (BP 140/90 mm Hg to 179/100 mm Hg) for 2 consecutive measurements among patients with hypertension who received pharmacological and nonpharmacological treatments at the Siriraj Primary Care Unit, had an intention to change self-management behaviors, could be accessed by smartphone, and ability to communicate in Thai. The study excluded those with any severe complications such as cardiovascular, cerebrovascular, and end-stage renal disease; inadequate digital health literacy; cognitive impairment; moderate to total dependency for undertaking basic activities of daily living; an inability to use a telephone or manage their medications; and similar characteristics but were previously exposed to education programs. Additionally, the study excluded participants who experienced abnormal symptoms during the activity, such as dizziness or fainting, and who did not improve after receiving first aid. The study withdrew participants who attended the activities for fewer than 75% of the time (less than

XSL•FO
RenderX

6 weeks out of 8 weeks). Regarding participant allocation, this study used the lottery method to determine the days to assign research participants to either the intervention or control group. The study designated the first and second lottery draws as the control group (Monday and Thursday), with the third and fourth lottery draws designated as the intervention group (Tuesday and Wednesday).

This study requires a total of 90 patients with uncontrolled hypertension to achieve a representative sample for phase 1, and we used the G*power program version 3.1.9.7 [38] to determine the estimated sample size based on repeated-measures ANOVA. An effect size (*f*) of 0.25 was set [39], with a significance level of .05, power of .80, the number of groups set at 2, the number of measurements set at 3, and the correlation among repeated measures set at 0.5. Therefore, the determined total sample size was 86 participants, with 43 participants per group. A previous study showed that patients with uncontrolled hypertension dropped out of the self-management program at a rate of 5% [39]. Finally, the total sample size was determined to be 90. Next, the proportional stratification was calculated using the following formula: sample size = (sample size/population size) × stratum size. Thus, in this study, 18 adults with uncontrolled hypertension and 72 older adults with uncontrolled hypertension were selected as the sample.

Research Measurements

The study used research instruments consisting of (1) screening measurements, (2) experimental measurements, and (3) outcome measurements. The details are presented in the following sections.

Screening Measurements

Cardiovascular Prevention Stage of Change Questionnaire

The study used the cardiovascular prevention stage of change questionnaire to screen the readiness of patients with uncontrolled hypertension to change their behavior, with permission from the developers. Singha-Dong et al [40] developed this measurement. It consists of 13 items, with response options reflecting the stages of behavior change, ranging from 0 points (not considering change) to 5 points (maintaining the behavior for more than 6 months). The total score ranges from 0 to 65. Regarding its psychometric properties, the test-retest reliability is 0.91.

Thai Version of the Mini-Cog

The study used the Thai version of Mini-Cog for screening for cognitive impairment of Thais with uncontrolled hypertension, with permission from the developers and translators. Borson et al [41] developed this measurement, and Trongsakul et al [42] subsequently translated it into the Thai version. It consists of 2 sections, including a memory test (3-item word recall) and a clock-drawing test. The memory test is assigned a score of 1 point per item, giving it a total of 3 possible points. The clock-drawing test is assigned a score of 2 points, giving it a total of 2 possible points. The scores are calculated from these 2 sections. Therefore, the total possible score for the Thai version of the Mini-Cog is 5 points. Possible scores range from 0 to 5, with a cutoff point of ≥3 indicating cognitive intactness. Regarding its psychometric properties, the Mini-Cog has shown high sensitivity (99%) and specificity (96%) [41].

eHealth Literacy Scale (eHEALS)

The study used the eHealth Literacy Scale (eHEALS) to screen the digital health literacy of patients with uncontrolled hypertension, with permission from the developers. Norman and Skinner [43] developed this measurement, and the Office of the Department of Health [44] subsequently translated it into the Thai version. It consists of 8 items rated on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). Possible scores range from 8 to 40, with a cutoff point of ≥26 indicating adequate digital health literacy [45]. Regarding its psychometric properties, the test-retest reliability is 0.85 [46].

Participants must score 26 or higher on the eHEALS to demonstrate their intention to utilize digital tools throughout the screening process. This threshold ensures that participants can safely and effectively engage with the digital intervention. In practice, this threshold is used to maximize participation and minimize dropout rates during the intervention phase. This is crucial for the initial assessment of the intervention's effectiveness.

Thai Version of the Barthel Index

The study used the Barthel index-Thai version to screen activities of daily living of participants who are aged 60 years through 79 years, with permission from the developers and translators. Mahoney and Barthel [47] developed this instrument, and Jitapankul and colleagues [48] subsequently translated it into the Thai version. It consists of 10 items, including feeding, grooming, transfer, toilet use, mobility, dressing, stairs, bathing, bowels, and bladder. The minimum score for the Barthel index is 0, and the maximum score is 20. The Barthel index categorizes dependency as totally dependent (0 to 4 points), severely dependent (5 to 8 points), moderately dependent (9 to 11 points), mildly dependent (12 points or more), and independent (20 points). Regarding psychometric property, the Kappa coefficients for the interrater reliability test were 0.79 and 0.68, respectively [48].

Lawton-Instrumental Activities of Daily Living Scale-Thai Version

The study used the Lawton-Instrumental Activities of Daily Living Scale-Thai version to screen independent living skills for patients who were aged 60 years to 79 years, with permission from the developers and translator. Lawton and Brody [49] developed this instrument, and Phanasathit [50] subsequently translated it into the Thai version. It consists of 8 items: using a telephone, shopping, food preparation, housekeeping, laundry, transportation, responsibility for own medications, and handling finances. The summary score ranges from 0 (low function, dependent) to 8 (high function, independent). Regarding the psychometric property, the Kappa coefficient of interrater reliability test was 0.85, with a Cronbach α reliability of 0.32 (95% CI –0.12 to 0.66), and the test-retest reliability according to the Spearman rank correlation coefficient was 0.46 [50].

XSL•FO
RenderX

Experimental Measurements

The study established a self-management intervention combined with DHIs, including a web-based application, chatbot, salt meter, and smartwatch. The study developed self-management skills combined with DHIs according to the Integrated Theory of Health Behavior Change [51]. The feasibility of this program for objective congruence was examined by 3 experts, including a nurse instructor specializing in self-management, a nurse instructor specialist experienced in hypertensive care, and an instructor specializing in digital health. The index of congruence was 1.00. A team of software developers developed the web-based application and chatbot. Table 1 shows the details of the self-management intervention combined with the activities in the DHIs.

Table . Details of self-management combined with the digital health intervention activities.

Week	Activities
Week 1	Group activity, 5 - 6 participants, 60 min • Assessing the knowledge and health beliefs and then providing accurate information (15 min) • Self-management skills development (20 min) • Accessibility and utilizing support instruments, including a salt meter, a smartwatch, web-based application, and chatbot (25 min)
Weeks 2 - 3	The chatbot follow-ups encourage self-care twice a week (Wednesday and Friday at 8 AM), offering personalized feedback and allowing participants to request one-on-one consultations with the researcher.
Week 4	Group activity, 5 - 6 participants, 60 min • Follow-up and review of self-management skills (20 min) • Training about reading nutrition labels (20 min) • Reviewing the web-based application and its functions, followed by a question-and-answer session (20 min)
Weeks 5, 6, and 7	The chatbot follow-ups encourage self-care twice a week (Wednesday and Friday, 8 AM), offering personalized feedback and allowing participants to request one-on-one consultations with the researcher.
Week 8	Group activity, 5 - 6 participants, 60 min • Knowledge sharing and self-management evaluation with goal comparison, followed by a question-and-answer session (30 min) • Reviewing self-management skills (20 min) • Encouragement to practice self-care at home (10 min)

Outcome Measurements

Overview

At each time point in phase 1, the study collected outcome measurements including (1) the demographic data form (baseline), (2) Dietary Approaches to Stop Hypertension Questionnaire (DASHQ; baseline and 4th and 8th weeks), (3) Exercise Behavior Questionnaire (EBQ; baseline and 4th and 8th weeks), (4) the Brief COPE inventory (Thai version; baseline and 4th and 8th weeks), and (5) an automatic BP measurement (baseline and 4th and 8th weeks).

Demographic Data Form

The researchers developed the demographic data form to collect the participants' demographic data. This semistructured interview form includes questions regarding gender, age, marital status, education, religion, occupation, income, sources of income, adequacy of income, living arrangement, duration of hypertension, received medicines, and underlying diseases.

Dietary Approaches to Stop Hypertension Questionnaire

The study used the DASHQ to examine the dietary behaviors of the participants, with permission from the developers, Siriboonyarit et al [52]. This measurement consists of 13 items: 8 items for the positive approach and 5 for the negative approach. This measurement utilizes a Likert scale with 5 categories, from 1 to 5 (1="not at all" and 5="always"), for items related to the positive approach. The response choices for all items are standardized on a scale from 13 to 65, with a score from 13 to 30.33 reflecting low dietary behavior, a score from 30.34 to 47.67 reflecting moderate dietary behavior, and a score from 47.68 to 65 reflecting high dietary behavior. The content validity was 0.91, with a reliability Cronbach α coefficient of 0.80. This study examined the reliability of the measurement using 30 patients with uncontrolled hypertension and determined it using the Cronbach α coefficient before collecting the data.

Exercise Behavior Questionnaire

The study used the EBQ to examine the exercise behaviors of the participants, with permission from the developers, Pieasakran et al [53]. The EBQ consists of 10 items. The measurement utilizes a Likert scale of 5 categories, from 0 to 4 (0="not at all" and 4="always") for items related to the positive approach. The response choices for all items are standardized on a scale from 0 to 40, with higher scores indicating better exercise behavior. The content validity is 0.80, with a Cronbach α coefficient reliability at 0.87. This study examined the measurement reliability using 30 patients with uncontrolled hypertension and determined it using the Cronbach α coefficient before collecting the data.

XSL•FO
RenderX

Brief COPE inventory (Thai Version)

The study used the Brief COPE inventory (Thai version) to document the stress management behaviors of the participants, with permission from the developers, Numsang and Tantrarungroj [54]. This measurement consists of 28 items, including active coping, planning, positive reframing, acceptance, humor, religion, using emotional support, using instrumental support, self-distraction, denial, venting, substance use, behavioral disengagement, and self-blame. This measurement uses a Likert scale of 4 categories, from 1 to 4 (1="not at all" and 4="always"). The response choices for all items are standardized on a scale from 28 to 112, with higher scores indicating better stress management behavior. Regarding the psychometric properties, the test-retest reliability was 0.70. This study examined the reliability of the measurement using 30 patients with uncontrolled hypertension and determined it using the Cronbach α coefficient before collecting data.

Automatic BP Measurement

This study took an automatic BP measurement using an HEM-9200T (Omron Healthcare Inc), along with the standard method for measuring BP based on the guidelines from the Thai Hypertension Society [55]. The standard calibrator lab at the clinical instrument testing center calibrated the automatic BP measurement for accuracy at the start of the intervention and during the study.

Preparation of the Research Assistant

The research assistant was a nurse who was trained in data collection. Initially, the researcher explained the objectives and methods of the study, including human rights protection of the sample. Second, the research assistant received a questionnaire and instructions for collecting the data. Third, the research assistant was, if not already, trained in BP measurement following the Thai Hypertension Society guidance [55]. After training, the research assistant carried out data collection under real situations to demonstrate an understanding of the instructions.

Data Collection

In phase 1, we used a quasi-experimental design comparing 2 groups with repeated measures to determine the effects of self-management combined with DHIs on dietary, exercise, and stress management behaviors and BP.

First, we asked 45 participants in the control group to complete their demographic data, dietary behavior, exercise behavior, stress-management behavior, and BP measurement as a pretest. They then received standard medical treatments in accordance with the standards of the Siriraj Primary Care Unit. After completion of the intervention (4th week) and a follow-up period (8th week), they completed post-test I and post-test II. The research assistant collected the data.

The study asked the 45 participants in the intervention group to provide their demographic data, as well as dietary behavior, exercise behavior, stress-management behavior, and BP measurement as a pretest. They then received 8 weeks of self-management combined with DHIs. After completion of the intervention (4th week) and a follow-up period (8th week), they completed posttest I and posttest II. The research assistant collected the data.

Data Analysis

In phase 1, the analysis will include both the complete case and intention-to-treat analyses. A statistical software program will be used to analyze the data from this phase. The study uses descriptive statistics (eg, frequency, percentage, mean, and SD) to describe the demographic data and characteristics of the sample. The study will use the χ^2 test or independent *t* test to compare differences in the demographic data and outcome variables between the control and intervention groups at baseline. If differences are identified in the demographic data, the study will set different variables as covariates to control for the confounding factors. This study will use generalized estimating equations to test the effects of self-management combined with DHIs on dietary, exercise, and stress management behaviors, as well as BP, compared between the control and intervention groups. This study will use linear mixed-effects modeling to test the effects of self-management combined with DHIs on dietary, exercise, and stress management behaviors, as well as BP, within the intervention group at baseline, after completion of the intervention, and after completion of a follow-up period.

Phase 2: Qualitative Method to Explore Experiences and Perceptions After Receiving the Interventions

This study conducted in-depth interviews to collect the participants' perceptions and experiences after completing all self-management activities combined with DHIs.

Recruitment

In Phase 2, the research assistant screened potential participants for eligibility based on the inclusion criteria and informed them of the results by telephone. If the potential participants were deemed eligible and willing to participate in the study, the researcher scheduled a meeting. The researcher then screened the qualifications of patients with hypertension according to the inclusion criteria. The study approached those who passed all inclusion criteria individually to sign the informed consent document.

Participants

The study recruited participants through purposive sampling using the inclusion criteria. The inclusion criteria for this phase included individuals aged 45 years through 79 years old who had a BP lower than 140/90 mm Hg after participating in self-management combined with DHIs or individuals aged 45 years through 79 years old who had a BP higher than 140/90 mm Hg after participating in self-management combined with DHIs. The study withdrew participants who experienced abnormal symptoms during the activity, such as dizziness or fainting, that did not resolve after receiving first aid.

Additionally, the study determined the sample size for phase 2 in accordance with the data collection procedure. The minimum sample size for an in-depth interview approach is 12 participants, which aligns with the qualitative research principle that emphasizes data saturation [56]. However, the objective in phase 2 is to explore participants' perceptions and experiences after

XSL•FO
RenderX

receiving self-management combined with DHIs, stratified by age group and BP status. The researcher used a stratified approach, interviewing 6 adults and 6 older adults with uncontrolled hypertension, as well as 6 adults and 6 older adults with controlled BP. Finally, a total number of 24 participants were included in an in-depth interview.

Research Measurements

In phase 2, the researcher used a background information form, semistructured interviews, and an interview guide for data collection. The interview guide was comprised of a set of open-ended questions, including (1) How does hypertension affect your daily life? (2) After participating in the self-management activities combined with the web-based application, how do you feel your self-management behaviors have changed? (3) What are the factors that reinforce or hinder your self-care? (4) How has the web-based application helped you adjust your self-care behaviors? (5) What is the level of your satisfaction with the web-based application? (6) Which function(s) do you use most frequently? And why? and (7) What are the prominent features or any elements of the web-based application that you feel need improvement? The content validity was examined by 3 experts, including a nurse instructor specializing in self-management, a nurse instructor specializing in hypertensive care, and an instructor specializing in digital health. The content validity was 1.00.

Data Collection

In this phase, the study used in-depth and semistructured interviews and an interview guide to collect participants' perceptions and experiences after completing all self-management activities combined with DHIs. The researcher conducted all interviews at a convenient time for the participants at the Siriraj Primary Care Unit for 30 minutes each in Thai, and each interview was audiorecorded.

Data Analysis

We will analyze the in-depth interviews in phase 2 using the method by Colaizzi [57]. The researcher will read the transcribed interviews multiple times to truly understand the feelings of the informants. After that, the researcher will thoroughly review each statement on every page to discern the narrative's significance, understand the content's significance and meaning, and define the meaning of the key phrases. The study will gather and thoroughly explain or describe the results obtained from the phenomenon under study. The study will clarify any ambiguous phenomena as much as possible and return the findings to the informants for verification.

Ethical Considerations

The study was approved by the Human Research Ethics Committee of the Faculty of Nursing and the Human Research Ethics Committee of Siriraj Hospital Faculty of Medicine, which jointly considered and approved the research project in the form of a Memorandum of Understanding (MU-MOU COA number IRB-NS2025/948.3105). The study protocol was registered under a Universal Trial Number (U1111-1324-0264), and approved by the Thai Clinical Trials Registry (TCTR20250722001).

Additionally, we informed all participants about the study's objectives, procedures, risks, and benefits. Participation was voluntary, and participants could withdraw at any time. Further, we required the participants to complete and sign consent forms before participation. Moreover, participants were able to stop attending the intervention at any time, before or after deciding to take part in this study, without any effects on their care at the Siriraj Primary Care Unit. In addition, the participants in the control group had access to the intervention upon request within 1 week after the intervention completion because this approach allowed us to minimize contamination and preserve the internal validity of the study while also ensuring that the participants in the control group could benefit from the intervention afterwards.

Further, the study keeps all collected data confidential following standard data protection laws (eg, participant anonymity was assured as the study would only disclose their participant numbers and initials) [58]. When study results are disseminated, only aggregated findings will be presented, ensuring that no individual participant can be identified. The researcher has stored the collected data in password-protected files on a personal computer, and documents related to the research participants have been kept in a locked cabinet accessible only to the researcher. Data stored on the cloud and audio recording files can only be accessed by the researcher with a password. After the research project is completed, the researcher will immediately destroy the data.

Moreover, the researcher covered each participant's travel expenses, totaling 160 (US $5.14) per trip. The researcher also covered the cost of a flat-rate internet service for each participant in the experimental group, at 300 (US $9.63) per month for 2 months, to support group activities and self-care practices at home. Each participant in the experimental group received compensation in 2 installments of 300 each, paid directly in weeks 1 and 4. In phase 2, the researcher paid each participant's travel expenses for attending 1 in-depth interview, with a fee of 160 per interview. No additional compensation was provided for the interview.

Results

This study received a grant in December 2024 and ethical approval on May 31, 2025. We recruited 86 individuals, including 18 adults and 68 older adults, with uncontrolled hypertension from the Siriraj Primary Care Unit in Thailand. In phase 1, this study began recruiting participants in August 2025. Data were collected between August 2025 and the first half of November 2025. In-depth interviews were conducted in phase 2 and continued through the end of November 2025. The data analysis should be completed by December 2025. The expected date for the results to be submitted for publication is March 2026.

Regarding the results presentation, the quantitative results from phase 1, including eating behavior, exercise behavior, stress management behavior, and BP, will be descriptively summarized by gender and age group. Age group and gender will be included as factors in the analyses, considering the limited sample size of the adult group. In phase 2, this study will use purposive sampling across age groups and BP status, and phase 2 will be

XSL•FO
RenderX

conducted after completion of phase 1 to compare themes between groups and help explain how self-management combined with DHIs influences outcomes. Subsequent academic publications will highlight age-specific findings, including any identified digital challenges.

Discussion

Overview

This protocol describes the rationale and design of an explanatory sequential mixed methods study to evaluate the effectiveness of self-management combined with DHIs while also exploring participants' experiences and perceptions after the intervention among patients with uncontrolled hypertension in Thailand. Uncontrolled hypertension is a significant public health issue in Thailand, leading to elevated morbidity, death, and increasing health care expenses. Previous evidence indicates that DHIs can support self-management and improve BP control [16,17], and their integration with nurses and other health care providers can further enhance adherence to self-management behaviors and BP control [21]. The researchers hypothesize that combining these self-management strategies with DHIs will improve dietary, exercise, and stress management behaviors and lower BP compared with standard care.

Existing research illustrated that DHIs have potential for improving both behavioral and clinical outcomes among patients with uncontrolled hypertension [16,17,22,30]. However, previous interventions in Thailand have emphasized single behaviors rather than adopting a comprehensive, multibehavior change approach [23,32,33]. The systematic review conducted by our team identified that the most effective DHIs for self-management among patients with uncontrolled hypertension address multiple health behaviors [21]. Our protocol addresses this gap by incorporating diet, exercise, and stress management strategies as core components that are often overlooked in interventions. In addition, several DHIs for patients with uncontrolled hypertension have been designed and focus on health education [22,30], medication reminders [22,30-33], and home BP monitoring [27,31,33,34] delivered through mobile apps, text messages, or web platforms. A key innovation of this study is a hybrid self-management model combined with DHIs targeting multiple behaviors.

The use of DHIs, such as web-based applications, chatbots, smartwatches, and salt meters, is not a one-size-fits-all solution. Several factors—including age group, digital health literacy, living area, and infrastructure—influence acceptance and uptake [59]. In this study, adults who are good with technology may find the web-based application and chatbot easy to use, whereas older adults who are digitally challenged may feel anxious or struggle with this technology, making it harder for them to use. To mitigate this, the researchers will incorporate a user-friendly interface with large text and initial in-person training sessions [60]. In addition, participants living in households with good infrastructure (good internet connectivity and affordability) may face fewer barriers than those living in households with poor infrastructure. To address this issue, the researchers will allocate a budget for internet connectivity and ease of internet access [61]. Recognizing these potential differences is important for correctly interpreting the study results.

An explanatory sequential design is used to address variations in users' perceptions and experiences. The quantitative phase will estimate changes in dietary, exercise, and stress management behaviors, as well as BP, associated with the self-management combined with DHIs. The qualitative phase will provide a deeper understanding of why the intervention succeeds or fails in primary care units in Thailand. Researchers planned to purposefully sample both adults and older adults with either controlled or uncontrolled BP for in-depth interviews. This qualitative inquiry will uncover specific barriers to implementation, allowing the researchers to distinguish between intervention success or failure and identify barriers.

This protocol has two main strengths. First, this protocol is a theory-informed, multicomponent design that combines self-management approaches with DHIs to address multiple behaviors for uncontrolled hypertension. These interventions provide evidence-based information, develop self-management skills, and offer instrumental support, as outlined by the Individual and Family Self-Management Theory [51]. Finally, the use of an explanatory, sequential, mixed methods design allows us to examine both behavioral and clinical outcomes in depth, as well as the subtleties of user engagement. Nonetheless, limitations are anticipated. The requirement for participants to use their own smartphone devices, to have adequate digital health literacy, and to have data plans may create a selection bias, thereby omitting the most economically disadvantaged persons. Also, data synchronization may be interrupted in remote regions where the internet is not always available. Another limitation is that it is a single-center study conducted in a primary care unit in Thailand; therefore, generalizability may be limited. Because of ethical limits on collecting data from people who did not provide consent, the researchers could not compare participants with those who were excluded or chose not to take part. This means the findings may not generalize to people who are less comfortable with technology, underscoring how difficult it can be to close the digital divide in eHealth.

Implications and Future Directions

The results of this study will have important effects on how Thailand deals with uncontrolled hypertension. If this combined self-management and DHI approach works, it could serve as a model for primary care units and noncommunicable disease clinics across the country. The explanatory sequence design is essential because it provides more than just data on effectiveness. It also helps us understand how factors like age, ability to use digital technology, and different experiences with health technologies influence outcomes.

Future research will focus on mitigating the gap of DHIs, including web-based applications and chatbots highlighted in this study. Insights from the qualitative phase on the digital divide and user acceptance will be critical for future DHI iterations. We aim to use these findings to develop more personalized, age-friendly interfaces or voice-activated features to further reduce barriers for patients with uncontrolled hypertension who are digitally challenged in Thailand. Furthermore, subsequent studies should evaluate the

XSL•FO
RenderX

cost-effectiveness of this intervention to support policy decisions regarding its integration into the national health security system.

Conclusion

Uncontrolled hypertension remains a critical public health challenge in Thailand, requiring innovative solutions that align with standard care. This protocol describes an innovative approach to addressing the knowledge gap between traditional self-management and DHIs, focusing on enhancing dietary, exercise, and stress management behaviors and reducing BP. This study aims to determine the effectiveness of self-management combined with DHIs and to explore participants' experiences and perceptions after receiving the intervention. The results from this study will be used to determine the effectiveness of self-management combined with DHI solutions in primary care units in Thailand. Moreover, the results will have the potential to offer practical guidance on how to care for patients with uncontrolled hypertension, thereby improving their hypertensive outcomes. These results could also provide important information that will lead to the further development of health care policies in Thailand.

Acknowledgments

The researchers would like to express their gratitude to the entire D43 advisory committee for their constructive comments, criticisms, and valuable suggestions concerning this project.

The authors declare the use of generative AI (GAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: idea generation, proofreading, and editing.

The GAI tools used were ChatGPT-5.2Pro, Gemini-3Pro, and Grammarly AI.

Responsibility for the final manuscript lies entirely with the authors.

GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

This manuscript was delegated to GAI tools under full human supervision. The researchers used the AI ChatGPT 5.2Pro and Gemini 3Pro for idea generation and used Grammarly AI to edit and proofread our manuscript.

Funding

The 2D43 TW009883 D43 Post-Doctoral Program at the School of Nursing, University of Michigan, provided funding support of US $10,000 for this study.

Data Availability

This paper is a study protocol. As such, datasets have not yet been generated nor analyzed. The authors are willing to share a complete, cleaned, and de-identified copy of the data after completing the study. The study team may make this study available upon approval of a written request, which must include an analytic plan, after publication of the primary and secondary aims in a peer-reviewed journal.

Authors' Contributions

Conceptualization: SS, NS, YS
Methodology: SS, NS, YS
Writing – original draft: SS, NS, YS
Writing – review & editing: SS, NS, YS

Conflicts of Interest

None declared.

Checklist 1
SPIRIT Checklist.
[PDF File, 299 KB - resprot_v15i1e81148_app1.pdf]

References

1. Global report on hypertension: the race against a silent killer. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240081062 [accessed 2024-07-01]
2. Sakboonyarat B, Rangsin R, Kantiwong A, Mungthin M. Prevalence and associated factors of uncontrolled hypertension among hypertensive patients: a nation-wide survey in Thailand. BMC Res Notes 2019 Jul 4;12(1):380. [doi: 10.1186/s13104-019-4417-7] [Medline: 31272496]

XSL•FO
RenderX

3. Teza H, Boonmanunt S, Unwanatham N, et al. Evaluation of transitions from early hypertension to hypertensive chronic kidney disease, coronary artery disease, stroke and mortality: a Thai real-world data cohort. Front Cardiovasc Med 2023;10:1170010. [doi: 10.3389/fcvm.2023.1170010] [Medline: 37206104]
4. Khiaocharoen O, Prasertworakul CP, Khattiyod T, et al. Outpatient cost by disease: outputs from the first phase cost per disease project. HISPA Compendium 2023;1(7):100-112. [doi: 10.14456/hispa.2023.7]
5. Aytenew TM, Kassaw A, Simegn A, et al. Uncontrolled hypertension among hypertensive patients in Sub-Saharan Africa: a systematic review and meta-analysis. PLoS ONE 2024;19(6):e0301547. [doi: 10.1371/journal.pone.0301547] [Medline: 38870163]
6. Dewi TP, Ahmad A. The most common predictors of uncontrolled blood pressure among hypertensive elders: a systematic review. International Journal of Advanced Health Science and Technology 2023;3(1):437-443. [doi: 10.35882/ijahst.v3i1.210]
7. Jariyasakulwong P, Wattanakitkrileart D, Pongthavornkamol K, Piaseu N, Roubsanthisuk W. Factors influencing self-care practices and blood pressure among individuals with uncontrolled hypertension: a cross-sectional study. Patient Prefer Adherence 2025;19:1-17. [doi: 10.2147/PPA.S497555] [Medline: 39764477]
8. Sodkhomkham N. Factors affecting blood pressure control among patients with primary hypertension, Samut Songkhram Province. J Res Health Improv Qual Life 2022;2(2):25-36 [FREE Full text]
9. Shahaj O, Denneny D, Schwappach A, et al. Supporting self-management for people with hypertension: a meta-review of quantitative and qualitative systematic reviews. J Hypertens 2019 Feb;37(2):264-279. [doi: 10.1097/HJH.0000000000001867] [Medline: 30020240]
10. Coates V. Role of nurses in supporting patients to self-manage chronic conditions. Nurs Stand 2017 May 17;31(38):42-46. [doi: 10.7748/ns.2017.e10742] [Medline: 28513400]
11. Hussien M, Muhye A, Abebe F, Ambaw F. The role of health care quality in hypertension self-management: a qualitative study of the experience of patients in a public hospital, North-West Ethiopia. Integr Blood Press Control 2021;14:55-68. [doi: 10.2147/IBPC.S303100] [Medline: 33948092]
12. Tharani A, Van Hecke A, Ali TS, Duprez V. Factors influencing nurses' provision of self-management support for patients with chronic illnesses: a systematic mixed studies review. Int J Nurs Stud 2021 Aug;120:103983. [doi: 10.1016/j.ijnurstu.2021.103983] [Medline: 34147728]
13. Liang J, He X, Jia Y, Zhu W, Lei J. Chinese mobile health apps for hypertension management: a systematic evaluation of usefulness. J Healthc Eng 2018;2018:7328274. [doi: 10.1155/2018/7328274] [Medline: 29744027]
14. Thongtaeng P, Seesawang J. Barriers and facilitators to successful hypertension management in older adults from the perspectives of community health nurses: a qualitative study. Nurs Sci J Thail 2022;40(1):35-49 [FREE Full text]
15. Fujiwara T, Tucker KL, McManus RJ. Future of hypertension management through mobile applications. Hypertens Res 2024 Apr;47(4):1087-1089. [doi: 10.1038/s41440-023-01568-5] [Medline: 38242947]
16. Boima V, Doku A, Agyekum F, Tuglo LS, Agyemang C. Effectiveness of digital health interventions on blood pressure control, lifestyle behaviours and adherence to medication in patients with hypertension in low-income and middle-income countries: a systematic review and meta-analysis of randomised controlled trials. EClinicalMedicine 2024 Mar;69:102432. [doi: 10.1016/j.eclinm.2024.102432] [Medline: 38333367]
17. Katz ME, Mszar R, Grimshaw AA, et al. Digital health interventions for hypertension management in US populations experiencing health disparities: a systematic review and meta-analysis. JAMA Netw Open 2024 Feb 5;7(2):e2356070. [doi: 10.1001/jamanetworkopen.2023.56070] [Medline: 38353950]
18. Classification of digital health interventions v1.0. World Health Organization. 2018. URL: https://www.who.int/publications/i/item/WHO-RHR-18.06 [accessed 2024-05-03]
19. Enhancing nursing and midwifery capacity to contribute to the prevention, treatment and management of noncommunicable diseases. World Health Organization. 2012. URL: https://www.who.int/publications/m/item/9789241504157 [accessed 2025-12-10]
20. Sharma A, Singh RK. The role of general nurses in non-communicable disease prevention: community-based interventions in action and their impact on public health outcomes in low resource settings. J Innov Ment Health Psychiatr Nurs 2025;1(2):54-62 [FREE Full text]
21. Sukpattanasrikul S, Singha-Dong N, Sitthimongkol Y, Anonjarn K. Efficacy and cost-effectiveness of digital health interventions in improving hypertensive outcomes among patients with uncontrolled hypertension: a systematic review. Int J Nurs Sci 2025 May;12(3):225-232. [doi: 10.1016/j.ijnss.2025.04.001] [Medline: 40529464]
22. Gong K, Yan YL, Li Y, et al. Mobile health applications for the management of primary hypertension: a multicenter, randomized, controlled trial. Medicine (Baltimore) 2020 Apr;99(16):e19715. [doi: 10.1097/MD.0000000000019715] [Medline: 32311957]
23. Rethanu T, Nuysri M, Jantacumma N. Effects of behavioral modification program on medication adherence among uncontrolled hypertension patients in Nawa District, Nakhon Phanom Province. J Public Health Nurs 2021;35(2):90-107 [FREE Full text]
24. Penpoompaung IJ, Inkoom J, Jaisangad S. Effects of a perceived self-efficacy, disease knowledge, and social support enhancement program on knowledge, self-care behaviors, and blood pressure among the older adults with uncontrolled hypertension. J Med Health Sci 2023;30(3):118-134 [FREE Full text]

XSL•FO
RenderX

25. Yookong K, Kitreerawutiwong N. Effectiveness of a telehealth self-management support program on self-management behavior and blood pressure among essential hypertension patients in Sawankhalok District, Sukhothai Province. Nurs J Minist Public Health 2023;33(3):25-38 [FREE Full text]
26. Sukpattanasrikul S, Monkong S, Leelacharas S, Krairit O, Viwatwongkasem C. Comparison of hypertensive outcomes after the implementation of self-management program for older adults with uncontrolled hypertension in Krabi, Thailand: a quasi-experimental study. JHR 2022 Jul 4;36(4):641-651. [doi: 10.1108/JHR-12-2020-0626]
27. Kario K, Nomura A, Kato A, et al. Digital therapeutics for essential hypertension using a smartphone application: a randomized, open-label, multicenter pilot study. J Clin Hypertens (Greenwich) 2021 May;23(5):923-934. [doi: 10.1111/jch.14191] [Medline: 33484628]
28. Wiriyatanakorn S, Mukdadilok A, Kantachuvesiri S, Mekhora C, Yingchoncharoen T. Impact of self-monitoring of salt intake by salt meter in hypertensive patients: a randomized controlled trial (SMAL-SALT). J Clin Hypertens (Greenwich) 2021 Oct;23(10):1852-1861. [doi: 10.1111/jch.14344] [Medline: 34510703]
29. Leupold F, Karimzadeh A, Breitkreuz T, et al. Digital redesign of hypertension management with practice and patient apps for blood pressure control (PIA study): a cluster-randomised controlled trial in general practices. EClinicalMedicine 2023 Jan;55:101712. [doi: 10.1016/j.eclinm.2022.101712] [Medline: 36386033]
30. Yilmaz E, Uzuner A, Bajgora M, et al. Effect of eTansiyon smartphone application on hypertension control. Prim Health Care Res Dev 2022 Oct 19;23:e64. [doi: 10.1017/S146342362100058X] [Medline: 36259143]
31. Still CH, Margevicius S, Harwell C, et al. A community and technology-based approach for hypertension self-management (COACHMAN) to improve blood pressure control in African Americans: results from a pilot study. Patient Prefer Adherence 2020;14:2301-2313. [doi: 10.2147/PPA.S283086] [Medline: 33262580]
32. Charoenwongsa K, Wirojratana V, Wattanakitkrileart D. The effect of a medication adherence promotion program through the Line application on older persons with essential hypertension. Nurs J CMU 2022;49(2):313-325 [FREE Full text]
33. Promrat S, Kapol P, Kapol N. Effects of a program to promote medication adherence using telepharmacy in patients with hypertension. Thai Bull Pharm Sci 2022;17(1):31-41 [FREE Full text] [doi: 10.69598/tbps.17.1.31-41]
34. McManus RJ, Little P, Stuart B, et al. HOME and Online Management and Evaluation of Blood Pressure (HOME BP) using a digital intervention in poorly controlled hypertension: randomised controlled trial. BMJ 2021 Jan 19;372:m4858. [doi: 10.1136/bmj.m4858] [Medline: 33468518]
35. Li X, Li T, Chen J, et al. A WeChat-based self-management intervention for community middle-aged and elderly adults with hypertension in Guangzhou, China: a cluster-randomized controlled trial. Int J Environ Res Public Health 2019 Oct 23;16(21):4058. [doi: 10.3390/ijerph16214058] [Medline: 31652688]
36. Thajang S, Panyathorn K, Loonprom N, et al. Effects of using media applications on the prevention of complications for knowledge of patients with hypertension. UDRU Sci Tech J 2020;8(2):89-104 [FREE Full text]
37. Samranbua A, Thamcharoentrakul B. The effects of relieving hypertension diet application on health beliefs among patients with hypertension. Thai J Cardio-Thorac Nurs 2021;32(1):228-242 [FREE Full text]
38. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
39. Sukpattanasrikul S, Wirojratana V, Hiransaroj B, et al. The report of health literacy development for patients with uncontrolled hypertension in the communities. 2023 Unpublished report.
40. Singha-Dong N, Pongpoottipatchara R, Anonjarn K, Prawatwong W. Comprehensive integrated program for the prevention and control of non-communicable diseases. 2024 Unpublished report.
41. Borson S, Scanlan J, Brush M, Vitaliano P, Dokmak A. The Mini-COG: a cognitive “vital signs” measure for dementia screening in multi-lingual elderly. Int J Geriatr Psychiatry 2000 Nov;15(11):1021-1027. [doi: 10.1002/1099-1166(200011)15:11<1021::aid-gps234>3.0.co;2-6] [Medline: 11113982]
42. Trongsakul S, Lambert R, Clark A, Wongpakaran N, Cross J. Development of the Thai version of Mini-Cog, a brief cognitive screening test. Geriatr Gerontol Int 2015 May;15(5):594-600. [doi: 10.1111/ggi.12318] [Medline: 25155468]
43. Norman CD, Skinner HA. eHEALS: the eHealth Literacy Scale. J Med Internet Res 2006 Nov 14;8(4):e27. [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
44. Department of Health. Thai version of the eHealth Literacy Questionnaire.: Department of Health, Ministry of Public Health; 2021. URL: https://multimedia.anamai.moph.go.th/infographics/ehealth-literacy-questionnaire/ [accessed 2024-07-20]
45. Richtering SS, Hyun K, Neubeck L, et al. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. JMIR Hum Factors 2017 Jan 27;4(1):e4. [doi: 10.2196/humanfactors.6217] [Medline: 28130203]
46. Plagatong W, Chanapai W, Srisomnuek A, Katchamart W. Reliability and feasibility of the eHEALS in Thai patients with rheumatic and autoimmune diseases. Thai J Rheumatol 2024;1(3):1-16 [FREE Full text]
47. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J 1965 Feb;14:61-65. [Medline: 14258950]
48. Jitapunkul S, Kamolratanakul P, Chandraprasert S, Bunnag S. Disability among Thai elderly living in Klong Toey slum. J Med Assoc Thai 1994 May;77(5):231-238. [Medline: 7869003]
49. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [Medline: 5349366]

XSL•FO
RenderX

50. Phanasathit M. Validity and reliability of lawton instrumental activities of daily living scale (l-IADL) and older people's quality of life questionnaire (OPQOL-brief): Thai version. ResearchGate. 2017. URL: https://www.researchgate.net/publication/342096615_karsuksakhwamtrnglaeakhwamtheiyngkhxngbaebsxbtham_The_Lawton_Instrumental_Activities_of_Daily_Living_IADL_Scale_chbabphasathiy [accessed 2025-12-24]
51. Ryan P. Integrated Theory of Health Behavior Change: background and intervention development. Clin Nurse Spec 2009;23(3):161-170. [doi: 10.1097/NUR.0b013e3181a42373] [Medline: 19395894]
52. Siriboonyarit N, Wattanakitkrileart D, Phligbua W. Effects of nutrition education program through LINE application on dietary behaviors in patients with primary hypertension. Royal Thai Navy Med J 2023;50(2):328-345 [FREE Full text]
53. Pieasakran T, Jiewpattanakul Y, Rerkluenrit J. Effects of a health literacy exercise program on the exercise behaviors and health literacy of older adults with hypertension in the communities, Bangkok Metropolis. Nurs Sci J Thail 2023;41(4):44-55 [FREE Full text]
54. Numsang T, Tantrarungroj T. Validity and reliability of the Brief COPE Inventory: Thai version. J Psychiatr Assoc Thailand 2018;63(2):189-198 [FREE Full text]
55. Thai hypertension guideline on the treatment of hypertension. Thai Hypertension Society. 2024. URL: https://thaihypertension.org/2024-guideline/ [accessed 2025-12-24]
56. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
57. Application of Colaizzi's method of data analysis in phenomenological research. MLU 2021;21(2):914-918 [FREE Full text] [doi: 10.37506/mlu.v21i2.2800]
58. Personal Data Protection Act, BE 2562. Royal Thai Government Gazette. 2019 May 27. URL: https://ratchakitcha.soc.go.th/documents/17088936.pdf [accessed 2025-12-01]
59. Hu S, Song D, Wan S, et al. Digital health: current applications, challenges, and future directions for enhancing healthcare quality and safety. Front Public Health 2025;13:1646802. [doi: 10.3389/fpubh.2025.1646802] [Medline: 41089861]
60. Jiang Q, Deng L, Zhang J, Pengbo Y. User-centered design strategies for age-friendly mobile news apps. SAGE Open 2024 Oct;14(4):21582440241285393. [doi: 10.1177/21582440241285393]
61. Chamwong S, Chomtohsuwan T, Charoenphandhu N. Analysis of inequality in household internet utilization and policy implications. Journal of Demography 2024;40(1):27-48. [doi: 10.56808/2730-3934.1369]

Abbreviations

BP: blood pressure
DASHQ: Dietary Approaches to Stop Hypertension Questionnaire
DHIs: digital health interventions
EBQ : Exercise Behavior Questionnaire
eHEALS: eHealth Literacy Scale
WHO: World Health Organization

Edited by J Sarvestan; submitted 23.Jul.2025; peer-reviewed by A Hebestreit; revised version received 17.Dec.2025; accepted 18.Dec.2025; published 02.Jan.2026.

Please cite as:
Sukpattanasrikul S, Singha-Dong N, Sitthimongkol Y
Self-Management Combined With Digital Health Interventions to Improve Dietary Behavior, Exercise Behavior, Stress Management Behavior, and Blood Pressure Among Thais With Uncontrolled Hypertension: Protocol for an Explanatory Sequential Study
JMIR Res Protoc 2026;15:e81148
URL: https://www.researchprotocols.org/2026/1/e81148
doi:10.2196/81148

© Saowaluck Sukpattanasrikul, Naruemol Singha-Dong, Yajai Sitthimongkol. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Just-in-Time Adaptive Intervention to Promote Walking Behavior and Reduce Stationary Time in Physically Inactive Adults: Protocol for the Walking With JITAIs Study

Cora J Firkin[1*], BS; Ajith Vemuri[2*], PhD; Tanvir Rahman[2*], MS; Barry Bodt[3*], PhD; Elizabeth Orsega-Smith[1*], PhD; Keith Decker[2*], PhD; Gregory M Dominick[1*], PhD

[1]Department of Health Behavior and Nutrition Sciences, College of Health Sciences, University of Delaware, Newark, DE, United States

[2]Department of Computer and Information Sciences, College of Engineering, University of Delaware, Newark, DE, United States

[3]Biostatistics Core Facility, College of Health Sciences, University of Delaware, Newark, DE, United States

[*]all authors contributed equally

Corresponding Author:
Gregory M Dominick, PhD
Department of Health Behavior and Nutrition Sciences
College of Health Sciences
University of Delaware
012 Carpenter Sports Building
26 North College Avenue
Newark, DE, 19716
United States
Phone: 1 3028313672
Fax: 1 3028314261
Email: gdominic@udel.edu

Abstract

Background: A Just-in-Time Adaptive Intervention (JITAI) recognizes the dynamic nature of individuals' states and contexts, predicts support needs, and sends tailored support at more opportune, actionable times.

Objective: This paper outlines the application architecture and protocol for the pilot "Walking With Just-in-Time Adaptive Interventions" (WWJ) study, which uses a JITAI approach to improve walking behavior—duration, speed, and distance—and reduce stationary time, defined as idle sitting or standing.

Methods: This study targets 20 adults who are physically inactive and leverages the Apple Watch to deliver fully automated tailored intervention notifications to "walk faster," "walk longer," or "stand up and move around" based on real-time data and contextual factors, including time-of-day activity patterns, geographic locations (eg, home, work, park, and gymnasium), weather conditions (eg, precipitation, wind speed, and humidity), and receptiveness. The protocol involves a preintervention assessment of demographics, behavior change constructs, anthropometrics, and resting vital signs; a 2-week observation period to establish walking behavior and stationary time baselines; a 2-week just-in-time learning period to evaluate receptiveness to untailored prompts at all applicable times; the 2-week JITAI intervention phase; and a postintervention assessment. Feasibility will be evaluated through protocol fidelity, participant adherence, Apple Watch wear-time compliance, user burden, acceptability ratings, and perceptions of benefits and preferences.

Results: The WWJ architecture development began in spring 2021 and concluded in fall 2022. Participant recruitment and enrollment began in fall 2022. A total of 18 participants were recruited. Upon completion of the analyses, the results of this study are expected to be submitted for publication.

Conclusions: Distinctively, the WWJ just-in-time learning period aims to train the learner based on user receptiveness within contexts by sending interventions whenever a participant meets the predetermined thresholds regardless of the likelihood that the user will be receptive to the notification to prune out nonopportune or "nonactionable" times. This approach may allow for greater customization during the JITAI period.

International Registered Report Identifier (IRRID): DERR1-10.2196/79022

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e79022) doi:10.2196/79022

KEYWORDS

Apple Watch; behavior change; feasibility; iPhone; just-in-time adaptive intervention; physical activity; wearable; mobile phone

Introduction

Background

Regular physical activity (PA) participation attenuates the accumulation and progression of cardiometabolic risk factors such as hypertension, dyslipidemia, overweight and obesity, and insulin resistance that contribute to cardiometabolic disease morbidity and mortality [1,2]. PA, broadly defined as "any bodily movement produced by skeletal muscles that results in energy expenditure," [3] is characterized by 4 dimensions: duration, frequency, intensity or rate of energy expenditure (EE), and PA type [3]. Historically, the 2007 American Heart Association Guidelines for Americans suggested adults 18-65 years of age should participate in moderate-intensity physical activity (MPA) for a minimum of 30 minutes per day for 5 days per week, including only intermittent bouts lasting 10 or more minutes [4]. The 2020 World Health Organization (WHO) recommendations build upon these guidelines, suggesting that the participation in 150-300 minutes of weekly MPA, 75-150 minutes of weekly aerobic vigorous-intensity physical activity, or an equivalent weekly combination of aerobic moderate-to-vigorous-intensity physical activity (MVPA) can be accumulated in bouts which can be of any length [5]. This approach allows for more flexibility in the accumulation of MVPA throughout each day, in which each minute counts toward weekly participation [5]. Nonetheless, a substantial proportion of adults—approximately 30% worldwide and 42% in high-income Western countries—fail to achieve these MVPA recommendations [6].

Moreover, a 2018 review found that adults spend approximately 8 hours participating in sedentary behavior (SB) daily [7]. SB is characterized by EE less than or equal to 1.5 metabolic equivalents of task (MET), with the Sedentary Behavior Research Network further detailing in 2017 that SB occurs while in a seated, reclined, or lying posture [8]. Emerging evidence indicates that excessive SB participation is associated with an increased prevalence of cardiometabolic diseases and risk of all-cause mortality [9] while reducing SB participation and subsequently increasing PA participation—even to a light-intensity—can significantly improve BMI and diastolic and systolic blood pressure (BP) [10-12]. Additionally, while intermittent standing bouts may be a feasible approach to reduce prolonged sitting, reclining, or lying through a postural change, passive or "idle" standing produces limited bodily movement, typically remaining below the 1.5 MET threshold to classify as light-intensity physical activity (LPA) [13].

Thus, reducing stationary behavior (ie, SB and idle standing) is a public health challenge that should be addressed through movement behaviors that reach MET thresholds at or above an LPA level. While specific LPA guidelines have yet to be established, the WHO advises adults to minimize SB time and substitute that time with any type and intensity of PA, including LPA [5]. Given that walking is generally feasible and accessible for adults who ambulate [5,14], it can serve as an ideal mode for not only replacing stationary behavior with LPA participation but also meeting WHO MVPA recommendations [5,14,15]. Walking paces around 0.82-0.89 m/s can exceed the 1.5 MET threshold for LPA [16,17]. To achieve the MPA threshold, walking must be performed at faster paces, typically around 1.3-1.8 m/s [16]. Since the 2020 WHO recommendations recognize that every minute of MVPA—regardless of bout duration—contributes toward an individual's weekly targets [5], promoting walking behavior at faster paces may be an appropriate approach to increasing weekly MVPA volume while also breaking up stationary behavior.

Mobile and Wearable Technology

Mobile and wearable technology advancements have unlocked innovative opportunities for collecting real-time walking and stationary behavior data and delivering individualized, theoretically driven, evidence-based interventions [18]. This technology can impact human behavior within "real-world" contexts by delivering actionable notifications based on real-time data collected from a user's smartphone or wearable device [19,20]. A market-leading consumer-grade wearable—the Apple Watch (AW)—is equipped with a range of features for real-time data collection and notification interaction. Features include triaxial accelerometer, gyroscope, heart rate (HR), and GPS sensors that use Bluetooth to transmit and receive data and interact with other devices or cloud-based services. Notably, systematic reviews have supported the promising validity, reliability, sensitivity, and feasibility of leveraging the AW for assessing PA and SB [21,22]. Prior work has established that the AW can classify movement types (eg, sitting, lying, and walking) with moderate accuracy [20]. With the ability to dynamically detect behavioral and contextual variables in real-time and interact with users through immediate notifications, the AW is a feasible platform for delivering behavioral interventions, such as just-in-time and Just-in-Time Adaptive Interventions (JITAI), to increase walking behavior and decrease stationary time [19,20].

Just-in-Time Intervention

According to the literature, a just-in-time behavioral intervention approach provides real-time support that is typically delivered through mobile or wearable applications or SMS text messaging or "texting" and is tailored to the user's current behavioral, psychological, and physiological state using data via smartphone or wearable sensors [23,24]. This support typically encourages users to further engage in positive, goal-directed health behaviors or discourages them from further engaging in their current behaviors, such as prolonged SB [23,24]. Additionally, some just-in-time interventions have provided notifications only once daily, typically triggered at a preset time or event (eg, 10 AM or postgoal achievement), focusing mainly on summarizing past performance to motivate future behavior, rather than offering continuous or actionable notification [25,26]. The

XSL•FO
RenderX

preprogrammed health and fitness notifications of the AW exemplify this approach by providing stand "reminders" or feedback—cuing users to move every 50 minutes if prolonged stationary behavior without at least a minute movement break was detected—and behavioral reinforcement through celebratory feedback messages when daily goals are met by closing the "move," "exercise," and "stand" rings which indicate aerobic EE, MVPA participation, and frequency of stationary breaks. Nevertheless, these notifications follow a rigid, predefined schedule or pattern rather than dynamically responding to the user and adapting to the various contextual influences that may impact their current behavior (eg, weather or geographic location) [24]. Thus, just-in-time interventions may have a limited impact on behavior change because they do not learn from and adapt to users' states and their various contextual influences, remaining more prescheduled or reactive than predictive and personalized over time.

About JITAI

To determine the most opportune moments for behavior change, JITAI approaches monitor changes in users' behavioral, psychological, and physiological states, and environmental contexts over time to deliver support given the user's state of need and receptivity to real-time actionable notifications [23,24]. This approach recognizes the dynamic nature of individuals' states and contexts, requiring the system to predict (1) whether the person requires support, (2) the type or amount of support needed, and (3) whether or not the support will be acted upon [23]. A 2025 scoping review of JITAIs across several health behaviors (eg, PA, SB, fluid intake, and substance use) and settings revealed that while the included JITAIs were deemed feasible, a majority (N=62, 55%) of approaches solely relied upon self-reporting data, with few studies (21%) collecting passively monitored data via sensors in real time [27]. A 2019 systematic review of JITAIs specifically designed to promote PA and reduce SB participation revealed mixed evidence on their effectiveness [24], emphasizing notable feasibility concerns, such as low adherence and acceptability, and lack of a strong theoretical basis to guide the JITAI approaches or explain behavior changes. Reviews have highlighted the favorable use of the social cognitive theory and transtheoretical model (TTM) in the development and evaluation of mobile and wearable technology-based interventions for promoting PA and reducing SB [18,25,28,29]. To refine and optimize the effectiveness of the JITAI implementation for health behavior change, a thorough evaluation of the feasibility of the JITAI approach and the relevance of social cognitive theory and TTM behavior change constructs is warranted [18,24].

Objective, Aims, and Research Questions

This protocol paper presents the development of the "Walking With Just-in-Time Adaptive Interventions" (WWJ) iPhone and AW app, server, and methodology for a 6-week pilot study among adults who are not meeting the 2020 WHO PA recommendations. The primary aims of the WWJ's study are to examine the feasibility and initial impact of a 2-week JITAI approach for increasing walking behavior (ie, walking speed, distance, and duration) and decreasing stationary behavior (ie, duration of SB and idle standing) compared to a 2-week baseline observation period and a 2-week just-in-time learning period. Thus, our corresponding research questions include: (1) Is the WWJ feasible, regarding recruitment, retention, intervention fidelity, acceptability, adherence, and engagement? (2) To what extent does the just-in-time approach impact walking and stationary behavior compared to baseline? (3) How do changes in walking and stationary behaviors across a 2-week just-in-time and a 2-week JITAI approach compare to baseline?

Methods

Overview of Study Design

This pilot feasibility study will use a 1-way within-subjects repeated-measures design to examine changes in walking and stationary behaviors across the 3 time periods: baseline, just-in-time, and JITAI. This study will consist of a preintervention measurement session, followed by a 2-week observation period that will serve as a baseline for walking and stationary behaviors, a 2-week just-in-time learning period, a 2-week JITAI period, and concluding with a postintervention measurement session (Figure 1). The receptiveness-in-context data from the just-in-time learning period will be used to train and individually adapt the JITAI server by learning the ideal contexts in which the participants are most receptive to each of the 3 intervention notifications, as detailed in the Architecture of the WWJ App and Server section.

XSL•FO
RenderX

Figure 1. CONSORT diagram: the WWJ study approach. CONSORT: Consolidated Standards of Reporting Trials; JITAI: Just-in-Time Adaptive Intervention; PAR-Q+: Physical Activity Readiness Questionnaire for Everyone; WWJ: Walking With Just-in-Time Adaptive Interventions.

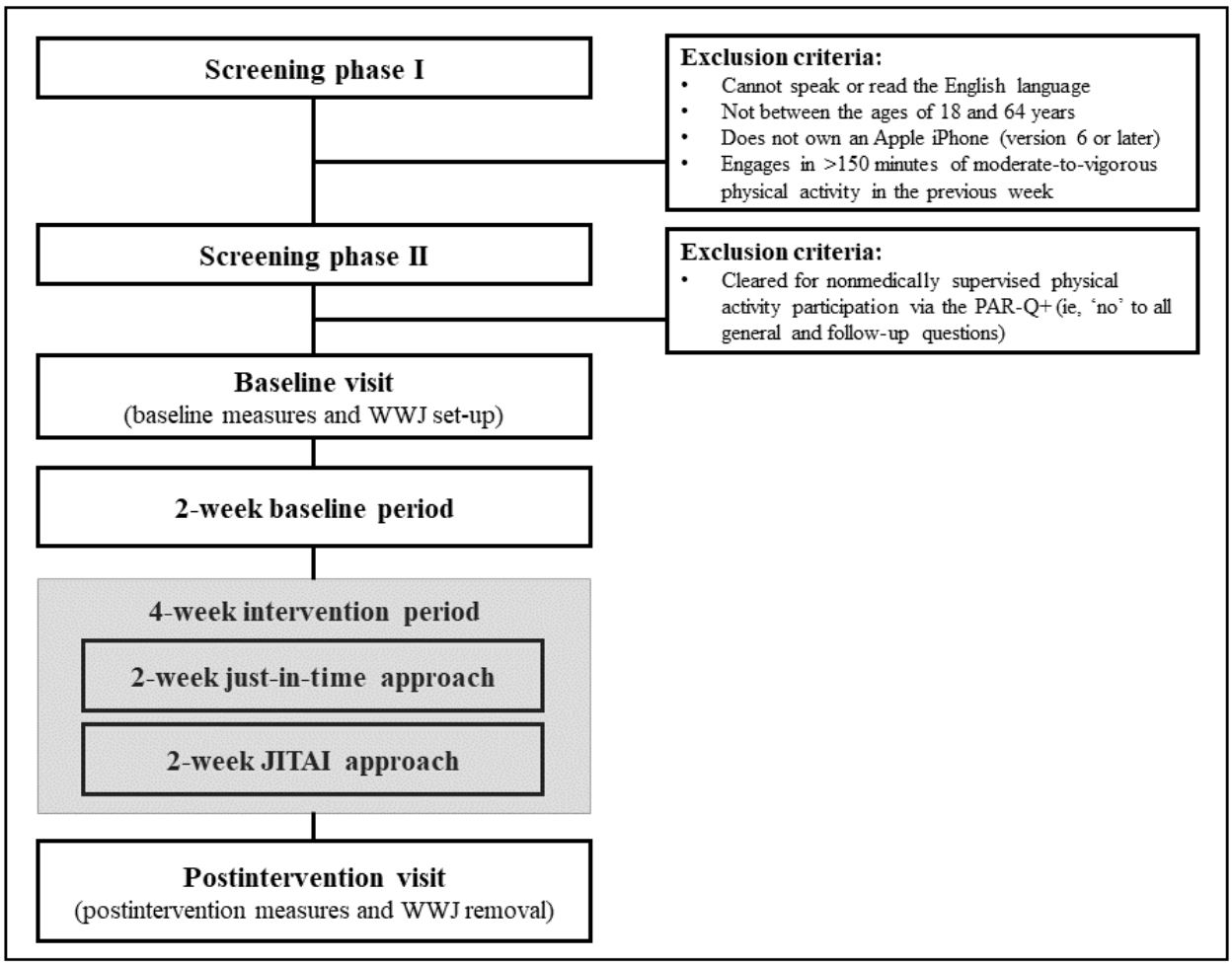

Participant Eligibility and Recruitment

A total of 20 participants will be recruited through community and clinical settings within a small city in New Castle County, Delaware, United States, using flyers, referrals, and social media (eg, Facebook and Instagram). Prospective participants will be directed to an online screening portal for the first screening phase (Figure 1) to determine study eligibility. Inclusion criteria will be adults aged 18-65 years, who own an Apple iPhone version 6s or newer, can speak and read in the English language, and self-report not achieving the minimum MVPA threshold of 150 weekly minutes in the previous week [5]. Exclusion criteria in the second screening phase will include answering "yes" to any general health and follow-up questions from the Physical Activity Readiness Questionnaire for Everyone, indicating that they are not cleared for PA participation without medical consultation and supervision [30]. If the screener indicates that a prospective subject is eligible, the informed consent form will be delivered to prospective participants via REDCap (Research Electronic Data Capture, Vanderbilt University).

Architecture of the WWJ App and Server

The research team has previously detailed a comparable server and application development [19]. The WWJ fully automated multiagent system operated on a per-agent architecture comprising 4 modules—context analyzer, learner, executor, and communicator (Figure 2). In the WWJ agent-level design, learning the support need policy is framed as a fully supervised classification problem rather than as a Markov Decision Problem [31] to account for the time constraints of a 2-week baseline and 2-week just-in-time learning period and the AW's battery-life constraints, which are particularly so when leveraging the HR sensor and GPS networks. Architecture development began in spring 2021 and concluded in fall 2022. Aligning with the CONSORT EHEALTH checklist (V.1.6.1), all updates to the WWJ application and server after this finalized development will be identified by date, version number, and description to allow for the replicability of the intervention [32].

XSL•FO
RenderX

Figure 2. WWJ architecture per agent. AW: Apple Watch; WWJ: Walking With Just-in-Time Adaptive Intervention.

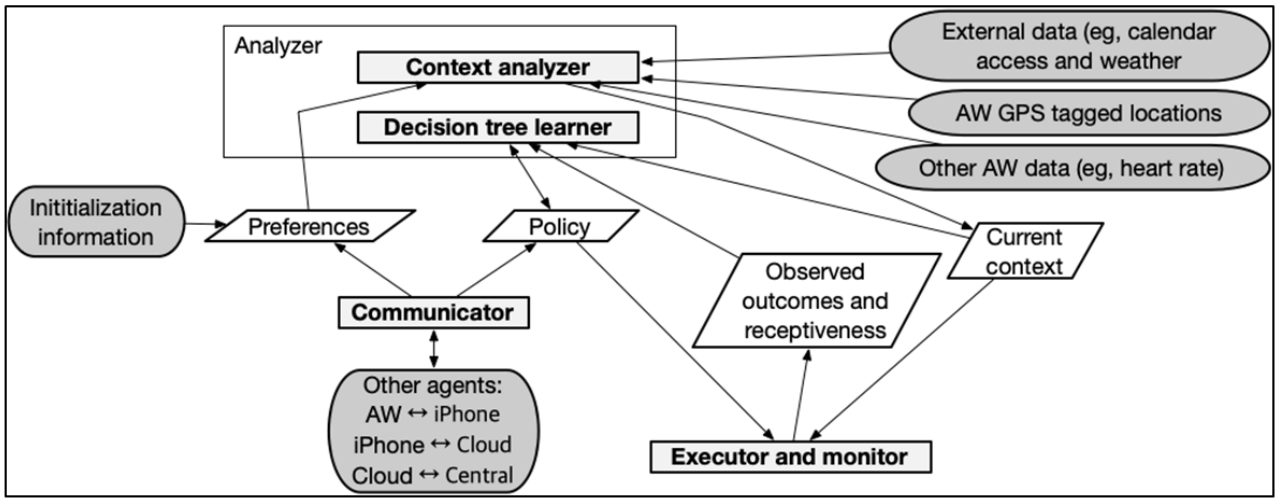

The context analyzer synthesizes external (eg, weather and calendar) and internal (eg, GPS coordinates mapped to tagged locations) contexts into a current context feature vector. At decision points, these contextual features (detailed under "contextual variables") are used in binary classification models to determine whether an intervention notification should be sent based on the predicted likelihood of successful behavior change within these contexts as processed and interpreted by the learner module. Thus, 3 decision trees (policies) will be trained for each user corresponding to the intervention notification types (Figure 3)—walk faster (ie, "You should pick up the pace"), walk longer (ie, "Let's make it a longer walk"), and stand up and move around (ie, "It's time to move around for one minute").

Figure 3. The WWJ intervention notifications: walk faster, walk longer, and stand up. WWJ: Walking With Just-in-Time Adaptive Intervention.



The learner module (Figure 2) also establishes baseline walking and stationary behaviors during the initial 2-week observation period by passively monitoring behaviors and calculating their applicable averages (eg, average walking speed) per bout. During the just-in-time learning period, intervention notifications are triggered at each instance that participants meet a predetermined threshold for walking and stationary behaviors—unless they have marked themselves as "busy" on their calendar to not burden the user during unsuitable times such as work meetings. The threshold for the walking notifications is set at 250 m from the initial start point of a walking detection. This distance was chosen after trial-and-error field testing among 4 authors (CJF, AV, KD, and GMD), which indicated that walking less than 250 m from a given starting location may not align with viable opportunities to walk faster or longer. The WWJ application will send a "walk longer" notification if the system recognizes that the user is in or near a previously tagged location. In all other cases, participants will receive a "walk faster" notification. When stationary behavior is detected for 30 continuous minutes, the application will send the "stand up and move around" notification. The system will determine if these interventions are successful or not (ie, receptiveness) either by comparing the current behaviors in the just-in-time learning period to those in the observation period

XSL•FO
RenderX

within the same contexts for walking behavior or by the AW detecting that the user got up and walked around for a minute for stationary behavior. As shown in Figure 2, data collected by the executor during the just-in-time learning period serves as labeled training data to compute and optimize personalized prompting policies for the JITAI period. At the end of the just-in-time learning period, the learner module uses this labeled data to generate a decision-tree–based policy, predicting when and when not to send the intervention notifications to the user based on their current context, past receptiveness, and predicted support needs. As predicting behavior is uncertain, the learning model will deliberately use intervention notifications even if the user may not be receptive, with a 10% chance of such occurrences, to allow behavior change in previously nonoptimal contexts and for the system to learn continually.

The communicator module will ensure that data will flow seamlessly through connections between the AW and a paired iPhone, the iPhone and Cloud, and the Cloud to our central server for secure data storage, processing, and offline analysis. This transmission process will occur every 2 hours. This timeframe was chosen after the aforementioned trial-and-error field tests to maximize AW battery life. When the WWJ application is running on the AW, the main user interface displays the current behavior (stationary or walking), HR, and tagged location of the participant (Figure 4).

Figure 4. The WWJ application's main user interface. WWJ: Walking With Just-in-Time Adaptive Intervention.



Measures and Materials

Apple iPhone and Watch Device Compatibility and Software

To ensure compatibility with the WWJ application, participants will use an iPhone 6s or newer running iOS 13 or later and an AW Series 3 or newer running watchOS 8.0 or later. If the participant does not own an AW Series 3 or is newer, they will sync a borrowed AW Series 3 running watchOS 8.0 or later with their iPhone. All applicable updates will be conducted for optimal performance. We will monitor iOS and watchOS updates to identify any potential impacts on just-in-time and JITAI notifications and address any issues that arise.

Demographics

Demographics will be collected using an adapted survey based on the demographics section of the 2021 Centers for Disease Control and Prevention–approved Behavioral Risk Factor Surveillance System Survey Questionnaire [33] on the following items: age; sex; race and ethnicity; level of education; employment, marital, and student statuses; annual household income; current housing arrangement (ie, renting, owning, or in a dormitory); and current ZIP code.

Behavior Change Constructs

The stages and processes of PA change from TTM will be assessed using scales developed by Marcus et al [34]. Additionally, the 5-item Self-Efficacy for Physical Activity scale will assess an individual's confidence in engaging in PAs despite barriers, such as tiredness, bad mood, time constraints, vacation time, and unfavorable weather conditions, that is, rain or snow [34]. PA self-regulation will be assessed using the 12-item version of the Physical Activity Self-Regulation scale, which has excellent internal consistency (Cronbach α=0.91) and good construct validity [35].

Anthropometrics and Resting Vital Signs

Body composition will be estimated using BMI (ie, body weight [in kg] divided by height [in m^2]). Participants will remove their shoes and any excessive clothing for all measurements. Weight will be measured using a calibrated, portable Seca scale to the nearest 0.1 kg. Height will be measured using a calibrated, portable stadiometer to the nearest 0.5 cm. Using a standard

XSL•FO
RenderX

stopwatch, radial resting HR will be measured for a full 60 seconds to the nearest beat after the participant is instructed to sit in a relaxed position in a chair with their back supported, feet on the floor, and arms supported at heart-level for 5 minutes in a private room [36]. Resting systolic and diastolic BP will be measured to the nearest mm Hg by a trained researcher using a standard portable BP machine (OMRON Intelli-Sense Series 7 Advanced Accuracy, Model BP760) and cuff size that allows for the bladder of the cuff to encircle at least 80% of the upper arm. The researcher will wrap the cuff firmly around the upper left arm of the participant at heart level and align the cuff correctly with the brachial artery. A second reading for resting HR and BP will be taken after an additional 1-2 minutes. If the BP readings are 5 or more mm Hg apart, a third reading will be obtained after a final 1-2 minutes of rest [36].

Contextual Variables

Busy Times of Day

During the baseline visit, participants will sync all their calendar systems (eg, Google, Microsoft Exchange, and Outlook.com) to their Apple Calendar on their paired iPhone. They will be instructed to use the Apple Calendar to schedule times they cannot be notified during by marking those events as "busy." Participants will be advised not to mark large blocks of time (eg, a full 8-hour workday) as "busy" and to ensure all scheduled physically active times are marked as "free."

Geographic Location

Participants will be instructed to provide a text label of their physical locations via the "tag" button on the WWJ user interface the first time they visit a new location while they are in that precise location (eg, tagging work while seated at one's work desk). After pushing the "tag" button, participants will have the option of selecting home, work, school, gymnasium, park, and other; if the user selects "other," they can type in any new text label (eg, "friend's house" as illustrated in Figure 4). Instructions will be given not to tag: (1) a location in a parking lot, (2) a nonprecise or overly generic location (eg, university campus versus a specific university building), or (3) behavior they are participating in (eg, walking around the neighborhood). Once tagged, the AW system will register each instance when the participant is in this location with a 50-m accuracy whenever data are recorded.

Weather

The ClimaCell application programming interface (API) of watchOS (currently Tomorrow.io API; Tomorrow Companies Inc) will collect participant data on precipitation to the nearest inch, temperature to the nearest Celsius degree, humidity to the nearest percent, and wind speed to the nearest mile per hour at the participant's current precise geographic location, leveraging detected latitudes and longitudes on the AW, with a 50-m accuracy during walking and stationary bouts.

Walking and Stationary Behaviors

Our WWJ app will collect timestamped walking and stationary behaviors once respective thresholds are met (described previously in the Architecture of the WWJ App and Server section). The Apple watchOS behavior detection APIs use its combined location sensors—the GPS, the Globalnaya Navigatsionnaya Sputnikovaya Sistema, Galileo, and Quasi-Zenith Satellite System—and its motion sensors—triaxial accelerometer and gyroscope—to detect walking and stationary behaviors via Apple's proprietary algorithms. These walking behaviors will include duration to the nearest whole second, distance to the nearest meter, and speed calculated from the collected distance divided by duration to the nearest meter per second. In addition, when a walking bout is detected, HR will be collected every 10 seconds. The duration of stationary behavior will be recorded to the nearest whole second.

Feasibility

The feasibility assessment of the WWJ pilot study will be guided by the recommendations outlined by National Institutes of Health's National Center for Complementary and Integrative Health [37]. Key feasibility metrics will include rates of prospective participants screened per month, participants enrolled per month, retention of participants per study measure, fidelity for protocol delivery, and adherence to study protocol. In cases of nonadherence to the WWJ intervention notifications, a follow-up notification will be sent to the participant to inquire why the prompt was not followed [18]. Response options will include "I did not see the notification," "I did not want to respond," "It was not a good time," and an option to choose "other" for inputted responses. Additionally, compliance with wearing AW for at least 10 hours each day for at least 4 hours each week will be determined. An exit survey will include perceived user burden with items adapted from the 2016 User Burden Scale developed by Suh et al [38], acceptability, treatment-specific expectations of benefits, and treatment-specific preferences for consideration for future iterations of WWJ to further detail user feedback [18].

Protocol

Preintervention Visit and WWJ Application Set-Up

The preintervention visit will take place in a private room to explain study procedures and complete baseline assessments (ie, demographics, behavior change constructs, anthropometrics, and resting vital signs). The AW will be initialized according to the participant's age, sex, race, weight, and height by inputting or updating these measures in the Apple Health app. Apple Health and Fitness notifications (eg, stand reminders, goal completions, daily coaching, special challenges, and activity sharing notifications) will be disabled. Participants will download the TestFlight application—an Apple Developer app that allows for testing beta versions of our WWJ iOS and watchOS apps—and be sent an invitation email with a code to access the WWJ app. Participants will be guided through the WWJ app installation on the AW and paired iPhone. After downloading, the WWJ app goes through each data type, including precise location, which will be monitored to obtain user approval. Location services for the WWJ app will be turned to "always on and precise." To familiarize participants with the WWJ app, visually enhanced instructions on how to tag precise geographic locations on the AW, use the Apple Calendar to schedule busy times on the iPhone, and restart the WWJ app will be provided in paper and digital forms, and tasks will be demonstrated on a laboratory iPhone and AW.

XSL•FO
RenderX

WWJ Observation Period, Just-in-Time, and JITAI Periods

For the next six weeks, participants will be instructed to (1) wear the AW on their nondominant wrist during all waking hours, (2) charge the AW during nonwaking hours, (3) tag all locations visited, (4) input and update busy times daily, (5) restart the WWJ app every morning upon waking, and (6) respond to all WWJ app feasibility notifications. During the 2-week observation period, the WWJ app will only monitor walking and stationary behaviors without sending any feasibility or intervention notifications. During the 2-week just-in-time learning period, the WWJ app will (1) monitor walking and SBs, (2) send intervention and feasibility notifications across all appropriate contexts, and (3) assess participant receptiveness to the intervention notifications to begin training for each participant's unique contexts in which intervention notifications led to successful (vs failed) behavioral changes. During the 2-week JITAI period, the WWJ app will (1) monitor walking and SBs and contextual variables; (2) send notifications at the most opportune moments based on the decision tree policy and the extent to which prior notifications were successful [39]; (3) assess participant receptiveness; and (4) elicit reasons for nonreceptiveness.

Postintervention Visit

The postintervention visit will take place in a private room to complete postassessments (ie, behavior change constructs, anthropometrics, and resting vital signs), uninstall the WWJ app, and assist the participant in restoring all original settings on their iPhone and AW. If an AW was borrowed for this study, the device will be returned at this time.

Analytic Plan, Sample Size, and Power Estimates

Data Processing

The WWJ server will store all data for manual processing. Data processing will start by ensuring all timestamps for walking and stationary bouts are sequentially aligned. This alignment will allow for accurate calculations for transitions between these states and other activity modes. Outliers, including device-specific artifacts, such as overlapping timestamps and erroneous bout durations such as those too short to be meaningful (eg, <10 seconds) or exceeding physical limits (eg, >5 hours of continuous walking) and walking speeds surpassing human capacities (eg, >3 m/s) [40], will be identified and reviewed for manual correction or exclusion. Additionally, geographic locations tagged in the data and any nonadherence responses marked as "other" and inputted by participants will be downloaded and coded using an inductive approach to group similar entries into cohesive categories emerging organically from the data. Once fully processed, all data will be stored in a standardized format.

Analytic Plan

Using the latest version of JMP Statistical Software (currently JMP, version 19.01, SAS Institute Inc, 1989-2025), preliminary analyses will include baseline descriptive statistics and exploratory data analysis. Baseline characteristics of the participants (eg, self-reported MVPA minutes from the screener, demographics, behavior change constructs, anthropometrics, and resting vital signs); recruitment, enrollment, and retention rates; implementation fidelity; adherence and compliance rates; user-burden; acceptability; treatment-specific expectations of benefits and preferences will be summarized in frequency tables for categorical variables and means and SDs for quantitative variables. Qualitative responses to applicable feasibility notifications and open-ended exit surveys will be explored.

Walking behavior and stationary time measures through all 3 periods—baseline, just-in-time, and JITAI—support a 1-way within-subjects ANOVA analysis of continuous measures (eg, average bout duration, speed, distance, and stationary time per study period for each participant) to demonstrate the feasibility of effecting a change in the behavior profile throughout the WWJ study. The intended 1-way within-subjects ANOVA assumes limited outliers, normality, and sphericity, which will be assessed graphically using boxplots, the Shapiro-Wilk test, and Mauchly test, respectively. Outliers, if present, will be assessed for their potential to weaken the analysis and removed if necessary. Nonnormal data will be transformed using the Box-Cox procedure to achieve normality. If sphericity is violated, the Greenhouse-Geisser and Huynh-Feldt corrections will be used. Tukey post hoc testing will be used to explain the differences between WWJ periods. If normality cannot be approximately achieved, the nonparametric Friedman test will be used. The model will be extended to allow for covariates if data screening of other collected demographic measures (eg, education) suggests their importance.

During implementation, communication drops or a lack of participant adherence may result in missing observations in some instances that inform the planned ANOVA. It will therefore be necessary to explore missing observations as to missing completely at random, missing at random, and missing not at random by using missing data analytics in JMP to associate instances with other study conditions (eg, software considerations and behavioral context). To the extent that the structure leading to the missing observation is not present or identifiable, missing at random or missing completely at random will be assumed for those records, and a linear mixed model will be used, leveraging maximum likelihood for estimation to achieve the within-subjects ANOVA goals without case-wise deletion. Missingness will be summarized in the analysis.

Secondary analyses will examine the initial effectiveness of the walking and stationary prompts. Specifically, the success or failure of notifications to encourage walking (ie, walking faster or longer) or reduce stationary time (ie, registered walking for at least 1 minute). Interarrival times and bouts of walking and stationary times computed at baseline will be summarized in terms of location and spread for continuous values or as frequency tables for bouts of activity.

Sample Size and Power Estimates

Power calculations for the planned 1-way ANOVA were performed using G*Power Version 3.1.9.6. [34]. With 3 collapsed times of repeated measures (baseline, just-in-time, and JITAI) and assuming a within-subject correlation of 0.5, given the focus on feasibility for this study, the most effective sample size for predicting our primary outcomes (change in

XSL•FO
RenderX

walking behavior and stationary time) was 17 participants with an α level of 0.05, medium-to-large effect size of Cohen f=0.32, and power of 0.8. Allowing for 15% attrition due to dropout or adherence issues, our proposed sample size is 20 participants.

Ethical Considerations

This study protocol aligns with the ethical standards of the American Psychological Association and was approved by the University of Delaware Institutional Review Board (IRBNet ID: 1771712-9). All eligible participants will sign an informed consent form to participate. To protect user privacy and data security, all learning models are run locally on the AW via the WWJ app. Study participants will be compensated US $50.00 via an e-gift card voucher.

Results

Participant recruitment and enrollment began in fall 2022. Data collection was completed in 2024 with a final study sample of 18 participants. Upon completion of the analyses, the results of this study are expected to be submitted for publication. This study was initially funded in March of 2020.

Discussion

Principal Findings

This protocol paper outlines the development of the WWJ server and app architecture and methodology for a pilot 6-week study to promote walking behavior and reduce stationary behavior among adults who are currently not meeting the 2020 WHO MVPA recommendations [5]. The objective of this study will be to examine the feasibility and initial impact of a 2-week JITAI approach on walking and stationary behaviors compared to a 2-week baseline and 2-week just-in-time approach. The cornerstone of a JITAI is the identification of decision points, decision rules, and contextual variables, which are taken into consideration for learning and adapting to provide potential behavioral interventions [24]. Our JITAI approach will identify and integrate both contextual variables and user receptiveness learned during the just-in-time period to provide user-specific intervention notifications to walk fast, walk longer, and stand up and move around at more personalized opportune times.

To provide support for behavior change, a JITAI will distinguish opportune times by predicting whether the person requires support, the type or amount of support needed, and whether or not the support will be acted upon [24,41]. Thus, user-specific receptiveness is a key feature of a JITAI [24]. For instance, a 2023 walking-based JITAI protocol accounted for receptiveness based on users' previous walking behavior occurring within the same 3-hour time window [42]. Our JITAI extends upon this approach by including a 2-week observation and a 2-week just-in-time "learning" period to train the learner based on user receptiveness within contexts. This unique application of the just-in-time approach—sending intervention notifications whenever a participant meets predetermined thresholds, regardless of the likelihood of receptivity—serves to prune out nonopportune or "nonactionable" moments over time. In doing so, it enables greater customization based on user- and context-specific patterns of success and nonsuccess in promoting behavior change. For example, a 2016 walking-based just-in-time study found that delivering intervention notifications when participants "overused" their smartphone, were sedentary "for a long time," were walking, or had "just had their meals" effectively increased walking behavior. Yet, participant feedback suggested that the system should be synced with their calendars to avoid sending intervention notifications at nonopportune—and potentially embarrassing—times [43]. Notably, this study also limited its sample to participants in the TTM's contemplation and preparation stages [43]. In contrast, our dynamic approach will allow for continued learning opportunities during the JITAI by occasionally sending intervention notifications during previously "nonactionable" times. This policy decision assumes that a user's state of need and receptivity may change over time, potentially enabling participants to progress fluidly through applicable stages of the TTM [44].

Similarly, our JITAI approach will not impose a minimum or maximum threshold for the frequency of intervention notifications. For instance, previous JITAIs have capped their maximum daily intervention notifications to delivering only 3 [45,46] or 4 [42] instances, regardless of whether additional opportune moments will arise. Additionally, in the 2023 walking-based JITAI, intervention notifications ceased once participants achieved a predetermined step-count goal [42]. Building on this evidence, our JITAI study will leverage a flexible approach by continuing to deliver intervention notifications even after participants meet the minimum recommendation of 150 weekly minutes of MVPA. This aligns with the WHO recommendation to encourage individuals to exceed the minimum activity thresholds for additional health benefits [5].

Furthermore, our approach extends upon the current literature by passively monitoring minute-level changes in activity patterns and their environmental contexts of busy schedules, location, and weather to personalize support. This contrasts with some previously published JITAIs that rely on self-reported states or ecological momentary assessments for policy decisions on whether to send an intervention notification [27]. For instance, a 2022 JITAI protocol aimed to reduce SB by sending intervention notifications based on preferred weather conditions, which were preidentified by participants during onboarding [47]. The system also considered nearby locations deemed suitable for PA by the researchers to send intervention notifications suggesting that the participant engage in PA rather than SB at one of the reachable locations [47]. In addition, a 2022 JITAI study aimed to promote PA by sending intervention notifications based on individual states aligned with predetermined participant-indicated preferred times of day and weather forecasts [48]. Our JITAI approach extends upon these protocols by considering passively-monitored location, time of day, real-time weather conditions—including precipitation, temperature, humidity, and wind speed—as well as external scheduling logistics (eg, "busy times" due to work restrictions or doctors' appointments). These environmental contexts have been well-documented to impact PA and SB [49,50] and were deemed important to consider by the participants in the 2016 walking-based just-in-time study [43].

XSL•FO
RenderX

Nonetheless, the collection and use of such granular personal and environmental data may raise concerns surrounding data security and privacy [51]. To mitigate these concerns, it is essential to implement robust data protection measures for JITAIs, especially those that leverage users' personal devices. To optimize data security and privacy, our WWJ app will run its learning models locally on the AW before securely sending and storing data on our server. Additionally, all transmitted data will be encrypted end-to-end, and participants' data will only be matched to their participant ID to foster user trust and ultimately support greater engagement and adherence to the JITAI intervention.

Strengths and Limitations

A key strength of our JITAI approach lies in using machine learning techniques to exclusively learn behavioral patterns, environmental contexts, and receptiveness with intervention notifications sent during the just-in-time learning period. This dynamic learning process has great promise to enhance the adaptability of the intervention and potential for behavior change by evolving to meet the unique support needs and preferences of each user. To date, this study is one of only a few true JITAI approaches in which a just-in-time learning period is leveraged to systematically assess, learn from, and dynamically adapt to user receptiveness and contextualized activity patterns in real time [24,27]. A second study strength is the assessment of theoretical psychosocial variables known to be associated with PA and SB behavior change—stages and processes of PA change, PA self-regulation, and PA self-efficacy. Exploring these variables may help identify relevant behavior change constructs to inform the development of a suitable framework for refining behavioral JITAI protocols. This framework could facilitate the proactive integration of psychosocial variables with real-time adaptive learning, further enhancing JITAI personalization.

Despite these strengths, we acknowledge the potential for limitations. One possible limitation is the extent to which participants' engagement with this study influences the potential for the server to learn. For learning to occur, the system must detect the most opportune times to deliver behavioral prompts, which is dependent upon participants meeting the preset minimum activity thresholds. In this study, we set 250 meters as the minimum threshold for detecting walking bouts with viable opportunities to walk faster or longer. As such, participants who do not meet this minimum threshold may not receive any walking intervention notifications during the just-in-time or JITAI periods, and therefore, the potential for the server to learn is likely negligible for participants who do not engage in walking bouts below this threshold. In contrast, participants regularly meeting and exceeding the minimum thresholds will likely have more successful JITAI intervention notifications due to the potential for greater system learning. Although a unique feature of this study is the just-in-time learning period, based on previous literature, the extent to which the perceived user burden of a just-in-time "learning" approach may contribute to potential participant burnout or disengagement over time is unclear. Understanding these responses to intervention notifications is essential for implementing effective behavioral JITAIs that maximize learning, user engagement, and long-term adherence without overwhelming the user. While this study's sample size of 20 participants is relatively small, the primary focus of this study will be establishing feasibility and acceptability, rather than generating generalizable behavioral outcomes. We expect the results and corresponding effect sizes will inform an adequately powered sample size for a future randomized controlled trial.

Anticipated Findings and Future Directions

On Wednesday, March 13, 2024, the methodology for this work was presented as a poster at the Society of Behavioral Medicine's 45th Annual Meeting and Scientific Sessions in Philadelphia, Pennsylvania, earning the Meritorious Award for an outstanding student-authored abstract submission and research presentation. This abstract was published in the *Annals of Behavioral Medicine* [52].

It is expected that our WWJ study will demonstrate adequate feasibility and acceptability. Following this preliminary examination, we expect to expand upon this approach by considering psychosocial constructs to ground our approach in a theoretical framework and additional contextual factors, such as personal affect, personality traits, or social factors, to further impact participant engagement and receptivity for this JITAI approach for adults who are inactive. This refinement could contribute significantly to the broader field of PA interventions, offering more tailored and impactful strategies for promoting long-term behavior change across diverse populations.

Conclusions

In conclusion, the WWJ study represents an innovative step in personalized health interventions, combining real-time contextual data with adaptive notifications to promote walking behavior and reduce stationary time among adults who are currently inactive. By demonstrating feasibility and acceptability, this work lays the foundation for future research to leverage the just-in-time "learning" period to train JITAI approaches, allowing them to dynamically adapt to user receptiveness and contextualized activity patterns in real time. This approach holds promise for developing more personalized and scalable interventions, ultimately contributing to the broader goal of supporting long-term PA and SB behavior change.

Acknowledgments

The Walking With Just-in-Time Adaptive Interventions project was supported by the University of Delaware Center of Innovative Health Research (GMD and KD) and the University of Delaware Graduate College through the Doctoral Fellowship for Excellence (CJF).

XSL•FO
RenderX

Funding

This study was internally funded through the Center for Innovative Health Research (CIHR), University of Delaware.

Conflicts of Interest

None declared.

References

1. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity—a systematic review of longitudinal studies. BMC Public Health 2013;13:813 [FREE Full text] [doi: 10.1186/1471-2458-13-813] [Medline: 24010994]
2. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ 2006;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
3. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131 [FREE Full text] [Medline: 3920711]
4. Haskell WL, Lee I, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007;39(8):1423-1434. [doi: 10.1249/mss.0b013e3180616b27] [Medline: 17762377]
5. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. Br J Sports Med 2020;54(24):1451-1462 [FREE Full text] [doi: 10.1136/bjsports-2020-102955] [Medline: 33239350]
6. Guthold R, Stevens G, Riley L, Bull F. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million participants. Lancet Glob Health 2018;6(10):e1077-e1086 [FREE Full text] [doi: 10.1016/S2214-109X(18)30357-7] [Medline: 30193830]
7. Bauman AE, Petersen CB, Blond K, Rangul V, Hardy LL. The descriptive epidemiology of sedentary behaviour. In: Leitzmann MF, Jochem C, Schmid D, editors. Sedentary Behaviour Epidemiology. Springer Series on Epidemiology and Public Health. Cham: Springer; 2017:73-106.
8. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, SBRN Terminology Consensus Project Participants. Sedentary Behavior Research Network (SBRN) - terminology consensus project process and outcome. Int J Behav Nutr Phys Act 2017;14(1):75 [FREE Full text] [doi: 10.1186/s12966-017-0525-8] [Medline: 28599680]
9. de Rezende LFM, Rodrigues Lopes M, Rey-López JP, Matsudo VKR, Luiz ODC. Sedentary behavior and health outcomes: an overview of systematic reviews. PLoS One 2014;9(8):e105620 [FREE Full text] [doi: 10.1371/journal.pone.0105620] [Medline: 25144686]
10. Frank HR, Mulder H, Sriram K, Santanam TS, Skinner AC, Perrin EM, et al. The dose-response relationship between physical activity and cardiometabolic health in young adults. J Adolesc Health 2020;67(2):201-208. [doi: 10.1016/j.jadohealth.2020.04.021] [Medline: 32571756]
11. Loh R, Stamatakis E, Folkerts D, Allgrove JE, Moir HJ. Effects of interrupting prolonged sitting with physical activity breaks on blood glucose, insulin and triacylglycerol measures: a systematic review and meta-analysis. Sports Med 2020;50(2):295-330 [FREE Full text] [doi: 10.1007/s40279-019-01183-w] [Medline: 31552570]
12. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time: beneficial associations with metabolic risk. Diabetes Care 2008;31(4):661-666. [doi: 10.2337/dc07-2046] [Medline: 18252901]
13. Kowalsky RJ, Stoner L, Faghy MA, Barone Gibbs B. A call to clarify the intensity and classification of standing behavior. Int J Environ Res Public Health 2021;18(16):8460 [FREE Full text] [doi: 10.3390/ijerph18168460] [Medline: 34444209]
14. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. JAMA 2018;320(19):2020-2028 [FREE Full text] [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
15. Morris JN, Hardman AE. Walking to health. Sports Med 1997;23(5):306-332. [doi: 10.2165/00007256-199723050-00004] [Medline: 9181668]
16. Murtagh E, Mair J, Aguiar E, Tudor-Locke C, Murphy M. Outdoor walking speeds of apparently healthy adults: a systematic review and meta-analysis. Sports Med 2021;51(1):125-141 [FREE Full text] [doi: 10.1007/s40279-020-01351-3] [Medline: 33030707]
17. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 compendium of physical activities: a second update of codes and MET values. Med Sci Sports Exerc 2011;43(8):1575-1581. [doi: 10.1249/MSS.0b013e31821ece12] [Medline: 21681120]
18. Aldenaini N, Alqahtani F, Orji R, Sampalli S. Trends in persuasive technologies for physical activity and sedentary behavior: a systematic review. Front Artif Intell 2020;3:7 [FREE Full text] [doi: 10.3389/frai.2020.00007] [Medline: 33733127]
19. Vemuri A, Decker K, Saponaro M, Dominick G. Multi agent architecture for automated health coaching. J Med Syst 2021;45(11):95 [FREE Full text] [doi: 10.1007/s10916-021-01771-2] [Medline: 34562163]
20. Fuller D, Anaraki JR, Simango B, Rayner M, Dorani F, Bozorgi A, et al. Predicting lying, sitting, walking and running using Apple Watch and Fitbit data. BMJ Open Sport Exerc Med 2021;7(1):e001004 [FREE Full text] [doi: 10.1136/bmjsem-2020-001004] [Medline: 33907628]

XSL•FO
RenderX

21. Bai Y, Hibbing P, Mantis C, Welk GJ. Comparative evaluation of heart rate-based monitors: Apple Watch vs Fitbit charge HR. J Sports Sci 2018;36(15):1734-1741. [doi: 10.1080/02640414.2017.1412235] [Medline: 29210326]
22. Fuller D, Colwell E, Low J, Orychock K, Tobin MA, Simango B, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR mHealth uHealth 2020;8(9):e18694 [FREE Full text] [doi: 10.2196/18694] [Medline: 32897239]
23. Naughton F. Delivering "Just-In-Time" smoking cessation support via mobile phones: current knowledge and future directions. Nicotine Tob Res 2017;19(3):379-383. [doi: 10.1093/ntr/ntw143] [Medline: 27235703]
24. Hardeman W, Houghton J, Lane K, Jones A, Naughton F. A systematic review of Just-in-Time Adaptive Interventions (JITAIs) to promote physical activity. Int J Behav Nutr Phys Act 2019;16(1):31 [FREE Full text] [doi: 10.1186/s12966-019-0792-7] [Medline: 30943983]
25. Schembre SM, Liao Y, Robertson MC, Dunton GF, Kerr J, Haffey ME, et al. Just-in-time feedback in diet and physical activity interventions: systematic review and practical design framework. J Med Internet Res 2018;20(3):e106 [FREE Full text] [doi: 10.2196/jmir.8701] [Medline: 29567638]
26. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. Key components in eHealth interventions combining self-tracking and persuasive ecoaching to promote a healthier lifestyle: a scoping review. J Med Internet Res 2017;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
27. Hsu TC, Whelan P, Gandrup J, Armitage CJ, Cordingley L, McBeth J. Personalized interventions for behaviour change: a scoping review of Just-in-Time Adaptive Interventions. Br J Health Psychol 2025;30(1):e12766. [doi: 10.1111/bjhp.12766] [Medline: 39542743]
28. Domin A, Spruijt-Metz D, Theisen D, Ouzzahra Y, Vögele C. Smartphone-based interventions for physical activity promotion: scoping review of the evidence over the last 10 years. JMIR mHealth uHealth 2021;9(7):e24308 [FREE Full text] [doi: 10.2196/24308] [Medline: 34287209]
29. Aldenaini N, Oyebode O, Orji R, Sampalli S. Mobile phone-based persuasive technology for physical activity and sedentary behavior: a systematic review. Front Comput Sci 2020;2:19. [doi: 10.3389/fcomp.2020.00019]
30. Warburton D, Jamnik V, Bredin S, Gledhill N. The Physical Activity Readiness Questionnaire for Everyone (PAR-Q+) and electronic physical activity readiness medical examination (ePARmed-X+). HFJC 2011;4(2):3-17. [doi: 10.14288/hfjc.v4i2.103]
31. Littman ML, Dean TL, Kaelbling LP. On the complexity of solving Markov decision problems. arXiv Preprint posted online on February 20, 2013. [doi: 10.48550/arXiv.1302.4971971]
32. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
33. Behavioral risk factor surveillance system survey questionnaire. Centers for Disease Control and Prevention.: U.S. Department of Health and Human Services; 2021. URL: https://www.cdc.gov/brfss/questionnaires/index.htm [accessed 2025-12-18]
34. Marcus BH, Selby VC, Niaura RS, Rossi JS. Self-efficacy and the stages of exercise behavior change. Res Q Exerc Sport 1992;63(1):60-66. [doi: 10.1080/02701367.1992.10607557] [Medline: 1574662]
35. Levesque CS, Williams GC, Elliot D, Pickering MA, Bodenhamer B, Finley PJ. Validating the Theoretical Structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors. Health Educ Res 2007;22(5):691-702. [doi: 10.1093/her/cyl148] [Medline: 17138613]
36. Liguori G. ACSM's guidelines for exercise testing and prescription. In: American College of Sports Medicine. Waltham, Massachusetts: Lippincott Williams & Wilkins; 2020.
37. National Institutes of Health. Pilot studies: common uses and misuses. National Center for Complementary and Integrative Health. URL: https://www.nccih.nih.gov/grants/pilot-studies-common-uses-and-misuses [accessed 2025-12-18]
38. Suh H, Shahriaree N, Hekler E, Kientz J. Developing and validating the user burden scale: a tool for assessing user burden in computing systems. 2016 Presented at: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA p. 3988-3999. [doi: 10.1145/2858036.2858448]
39. Quinlan JR. Induction of decision trees. Mach Learn 1986;1(1):81-106. [doi: 10.1007/bf00116251]
40. Bohannon R. Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants. Age Ageing 1997;26(1):15-19. [doi: 10.1093/ageing/26.1.15] [Medline: 9143432]
41. Spruijt-Metz D, Wen CKF, O'Reilly G, Li M, Lee S, Emken BA, et al. Innovations in the use of interactive technology to support weight management. Curr Obes Rep 2015;4(4):510-519 [FREE Full text] [doi: 10.1007/s13679-015-0183-6] [Medline: 26364308]
42. Park J, Kim M, El Mistiri M, Kha R, Banerjee S, Gotzian L, et al. Advancing understanding of just-in-time states for supporting physical activity (Project JustWalk JITAI): protocol for a system ID study of Just-in-Time Adaptive Interventions. JMIR Res Protoc 2023;12:e52161 [FREE Full text] [doi: 10.2196/52161] [Medline: 37751237]
43. Ding X, Xu J, Wang H, Chen G, Thind H, Zhang Y. WalkMore: promoting walking with just-in-time context-aware prompts. : IEEE; 2016 Presented at: IEEE Wireless Health (WH); October 25-27, 2016; Bethesda, MD p. 1-8. [doi: 10.1109/wh.2016.7764558]

XSL•FO
RenderX

44. Prochaska JO. Decision making in the transtheoretical model of behavior change. Med Decis Making 2008;28(6):845-849. [doi: 10.1177/0272989X08327068] [Medline: 19015286]
45. van Dantzig S, Geleijnse G, van Halteren AT. Toward a persuasive mobile application to reduce sedentary behavior. Pers Ubiquit Comput 2013;17(6):1237-1246. [doi: 10.1007/s00779-012-0588-0]
46. Müller AM, Blandford A, Yardley L. The conceptualization of a Just-in-Time Adaptive Intervention (JITAI) for the reduction of sedentary behavior in older adults. mHealth 2017;3:37 [FREE Full text] [doi: 10.21037/mhealth.2017.08.05] [Medline: 29184889]
47. Ismail T, Al Thani D. Design and evaluation of a Just-in-Time Adaptive Intervention (JITAI) to reduce sedentary behavior at work: experimental study. JMIR Form Res 2022;6(1):e34309 [FREE Full text] [doi: 10.2196/34309] [Medline: 35080498]
48. Mair JL, Hayes LD, Campbell AK, Buchan DS, Easton C, Sculthorpe N. A personalized smartphone-delivered Just-in-Time Adaptive Intervention (JitaBug) to increase physical activity in older adults: mixed methods feasibility study. JMIR Form Res 2022;6(4):e34662 [FREE Full text] [doi: 10.2196/34662] [Medline: 35389348]
49. Tucker P, Gilliland J. The effect of season and weather on physical activity: a systematic review. Public Health 2007;121(12):909-922. [doi: 10.1016/j.puhe.2007.04.009] [Medline: 17920646]
50. Sallis JF, Floyd MF, Rodríguez DA, Saelens BE. Role of built environments in physical activity, obesity, and cardiovascular disease. Circulation 2012;125(5):729-737 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.110.969022] [Medline: 22311885]
51. Sui A, Sui W, Liu S, Rhodes R. Ethical considerations for the use of consumer wearables in health research. Digit Health 2023;9:20552076231153740 [FREE Full text] [doi: 10.1177/20552076231153740] [Medline: 36756643]
52. Firkin CJ, Vemuri A, Malone S. Development of walking with JITAIs: a Just-in-Time Adaptive Intervention to promote physical activity in adults. 2024 Presented at: 2024 SBM Annual Conference; March 13-16, 2024; Philadelphia, PA p. S321.

Abbreviations

API: application programming interface
AW: Apple Watch
BP: blood pressure
EE: energy expenditure
HR: heart rate
JITAI: Just-in-Time Adaptive Intervention
LPA: light-intensity physical activity
MET: metabolic equivalent of task
MPA: moderate-intensity physical activity
MVPA: moderate-to-vigorous-intensity physical activity
PA: physical activity
REDCap: Research Electronic Data Capture
SB: sedentary behavior
TTM: transtheoretical model
WHO: World Health Organization
WWJ: Walking With Just-in-Time Adaptive Interventions

Edited by E Andrikopoulou; submitted 17.Jun.2025; peer-reviewed by C Chen, AJ Pesola; comments to author 02.Nov.2025; revised version received 24.Nov.2025; accepted 27.Nov.2025; published 07.Jan.2026.

Please cite as:
Firkin CJ, Vemuri A, Rahman T, Bodt B, Orsega-Smith E, Decker K, Dominick GM
Just-in-Time Adaptive Intervention to Promote Walking Behavior and Reduce Stationary Time in Physically Inactive Adults: Protocol for the Walking With JITAIs Study
JMIR Res Protoc 2026;15:e79022
URL: https://www.researchprotocols.org/2026/1/e79022
doi:10.2196/79022
PMID:

©Cora J Firkin, Ajith Vemuri, Tanvir Rahman, Barry Bodt, Elizabeth Orsega-Smith, Keith Decker, Gregory M Dominick. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 07.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which

XSL•FO
RenderX

permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

The Effect of Digital Therapy–Based Aerobic Exercise on Children With Neuroblastoma: Protocol for a Prospective Controlled Trial

Xuan Zhou[1*], MM; Shihao Huang[2*], MM; Lixia Wang[1*], MS; Haibin Guo[1], MM; Sisi Chen[1], MM; Tingrui Li[1], MBBS; Qing Du[1], PhD

[1]Department of Rehabilitation Medicine, Xin Hua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, Shanghai, China

[2]Department of Pediatric Hematologic Oncology, Xin Hua Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, Shanghai, China

[*]these authors contributed equally

Corresponding Author:
Qing Du, PhD
Department of Rehabilitation Medicine
Xin Hua Hospital
Shanghai Jiao Tong University School of Medicine
1665 Kongjiang Road
Yangpu District
Shanghai, Shanghai, 200082
China
Phone: 86 021 2507 8620
Fax: 86 0216503084
Email: duqing@xinhuamed.com.cn

Abstract

Background: Neuroblastoma, one of the most common extracranial solid tumors in children, often results in significant functional impairments, including muscle atrophy, reduced motor function, and sensory disturbances. While aerobic exercise has been shown to support functional recovery in patients with cancer, digital therapy–based aerobic exercise interventions have not yet been explored in pediatric neuroblastoma.

Objective: This study aims to develop and evaluate a digital intervention therapy to provide a safe, feasible, and tolerable home-based physical activity initiation protocol for children with neuroblastoma after chemoradiotherapy, with the main objectives of (1) breaking sedentary behavior habits and increasing daily energy expenditure, (2) gradually improving physical function (flexibility, basic muscle strength and muscle endurance, and balance ability), (3) establishing the habit of regular exercise and improving exercise self-efficacy, and (4) laying the foundation for a possible transition to more standard, albeit low-risk, aerobic exercise.

Methods: A total of 68 participants will be recruited and allocated to 2 groups matched in age and gender, with consideration given to the preferences of participants and their guardians. The control group will receive standard in-hospital medical treatment and health education, whereas the experimental group will receive a 6-month digital therapy–based aerobic exercise intervention in addition to standard care. Outcomes will be assessed at baseline and 3 and 6 months, with data collected for analysis. The primary outcomes are the 6-minute walk distance and the timed scores (in seconds) for the timed up and go and stair climb tests. The secondary outcomes include body composition, grip strength, and quality of life measured using the Pediatric Quality of Life Inventory version 4.0.

Results: Recruitment will be conducted from December 2025 to September 2026. We expect that the intervention group will show significantly greater improvements in physical function, physical activity levels, and fatigue management than the control group, demonstrating the effectiveness of digital therapy for rehabilitation training after radiotherapy and chemotherapy for neuroblastoma in children.

Conclusions: We expect that the results of this study will provide evidence for rehabilitation after chemoradiotherapy in pediatric patients with neuroblastoma to improve physical function, physical activity levels, and fatigue management. In the future, expanding to multiple centers to conduct targeted surveys of different types of childhood cancers will help validate the effectiveness of digital therapy.

International Registered Report Identifier (IRRID): DERR1-10.2196/72698

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e72698) doi:10.2196/72698

KEYWORDS

rehabilitation; aerobic exercise; digital therapy; neuroblastoma; pediatrics

Introduction

Neuroblastoma is one of the most common solid tumors in children, accounting for 8% to 10% of all childhood cancers, with a global incidence of approximately 10.1 cases per million children annually [1,2]. The clinical presentation of neuroblastoma varies depending on the tumor's location and severity, often involving compression or invasion of neural structures. Standard treatments such as radiotherapy, chemotherapy, immunotherapy, and surgery can lead to adverse effects, including fatigue, vomiting, and malnutrition [3,4]. Additionally, children with neuroblastoma frequently experience limb motor dysfunction (eg, hypokinesia), sensory deficits, behavioral and cognitive impairments (eg, learning difficulties and attention-deficit/hyperactivity disorder), mood disorders, and sleep disturbances [5,6]. Given that childhood is a critical period for physical and mental development, the long-term impact of neuroblastoma on cardiopulmonary fitness and quality of life underscores the importance of early exercise interventions [7].

There is a growing body of preclinical evidence suggesting that exercise positively influences the tumor microenvironment, enhances immune function, and supports nervous system health [8,9]. In clinical settings, exercise has been shown to accelerate postoperative recovery, improve physical activity levels and cardiopulmonary fitness, and alleviate cancer-related fatigue in patients undergoing radiation or chemotherapy [10]. Combined exercise interventions alongside chemotherapy and surgery have been reported to reduce tumor growth rates by 31% to 67% and lower the risk of metastasis [11]. Furthermore, preoperative exercise interventions may decrease postoperative complications and shorten hospital stays [12]. Braam et al [13] demonstrated that exercise interventions during and after cancer treatment improve body composition, flexibility, cardiorespiratory fitness, muscle strength, and quality of life in pediatric patients with cancer. A systematic review of exercise interventions in children with acute lymphoblastic leukemia highlighted their benefits in reducing fatigue during chemotherapy and maintaining physical activity, muscle strength, bone density, and functional capacity after treatment [14].

Aerobic exercise, the most commonly studied form of exercise intervention in oncology, has been shown to enhance cardiovascular fitness, boost metabolism, strengthen muscles, improve quality of life, reduce fatigue, and alleviate depression [15-17]. For instance, Leclerc et al [18] reported significant improvements in motor function, body fat percentage, and quality of life in patients with breast cancer following a 3-month multidisciplinary aerobic exercise program. Similarly, Tanriverdi et al [19] found that a virtual reality–based aerobic exercise intervention improved sleep quality in children with acute lymphoblastic leukemia. Despite these benefits, patient adherence remains a significant barrier to the implementation of exercise interventions. Blanchard et al [20] noted that 52.7% to 70.4% of patients with cancer in the United States fail to meet recommended exercise guidelines. Strategies to enhance adherence and intervention efficacy are critical for advancing this field. For example, Le et al [21] demonstrated that a 12-week telephone-based exercise intervention improved physical activity, fatigue, quality of life, and vitality in survivors of cancer. However, such interventions require substantial manpower and lack direct interaction, highlighting the need for more scalable and interactive solutions.

The integration of digital health technologies into health care services has opened up new avenues for innovation. Digital therapeutics (DTx), which use evidence-based software programs to deliver interventions for disease treatment, management, or prevention, represent a promising approach [22]. DTx leverage artificial intelligence to enable effective disease monitoring, provide personalized treatment plans, and enhance health care delivery [23]. In 2019, the World Health Organization endorsed digital rehabilitation as a continuous care model in its guidelines on digital interventions for health system strengthening [24]. DTx for aerobic exercise interventions are now widely used in managing chronic conditions such as heart disease, diabetes, and obesity [25]. Studies have shown that users of digital platforms are more likely to achieve exercise-related benefits compared to nonusers [26], and DTx have demonstrated high engagement and acceptability among children and their parents [27]. Compared with traditional face-to-face aerobic training, DTx-based aerobic exercise offers several distinctive advantages, including real-time feedback, science popularization and patient education, and adaptive intensity adjustment through data analytics. These features contribute to higher user motivation, better adherence, and more precise monitoring of exercise quality and safety [28,29]. By transcending time and space limitations, DTx enable real-time remote guidance, intervention quality monitoring, and reduced hospitalization frequency, thereby improving intervention efficiency and alleviating the burden on health care providers.

While animal studies have demonstrated the positive effects of exercise on physical function, immunity, and tumor prognosis in neuroblastoma models [30], clinical evidence on the impact of DTx-based aerobic exercise in children with neuroblastoma remains scarce. This study aims to address this gap by investigating the effects of a novel aerobic exercise–based digital therapy intervention on motor function and quality of life in children with neuroblastoma.

Methods

Study Design

This study is a single-center, prospective clinical trial designed to evaluate the efficacy of a digital therapy–based aerobic exercise intervention in children with neuroblastoma. The trial adheres to US Food and Drug Administration certification and

XSL•FO
RenderX

management standards for digital therapy. Participants will be divided into 2 groups: the control group will receive routine in-hospital diagnosis and treatment, whereas the experimental group will receive a digital therapy–based aerobic exercise intervention in addition to standard care. Following group allocation, the control group will receive physical activity guidance from clinicians as part of conventional care, whereas the experimental group will follow a personalized digital exercise management program tailored to their disease condition and lifestyle (Figure 1). The trial aims to establish a digital management model for children with neuroblastoma through statistical analysis of clinical outcomes.

Figure 1. Digital management pathway for children with neuroblastoma. 6MWD: 6-minute walk distance; PedsQL 4.0: Pediatric Quality of Life Inventory version 4.0; WHO: World Health Organization.

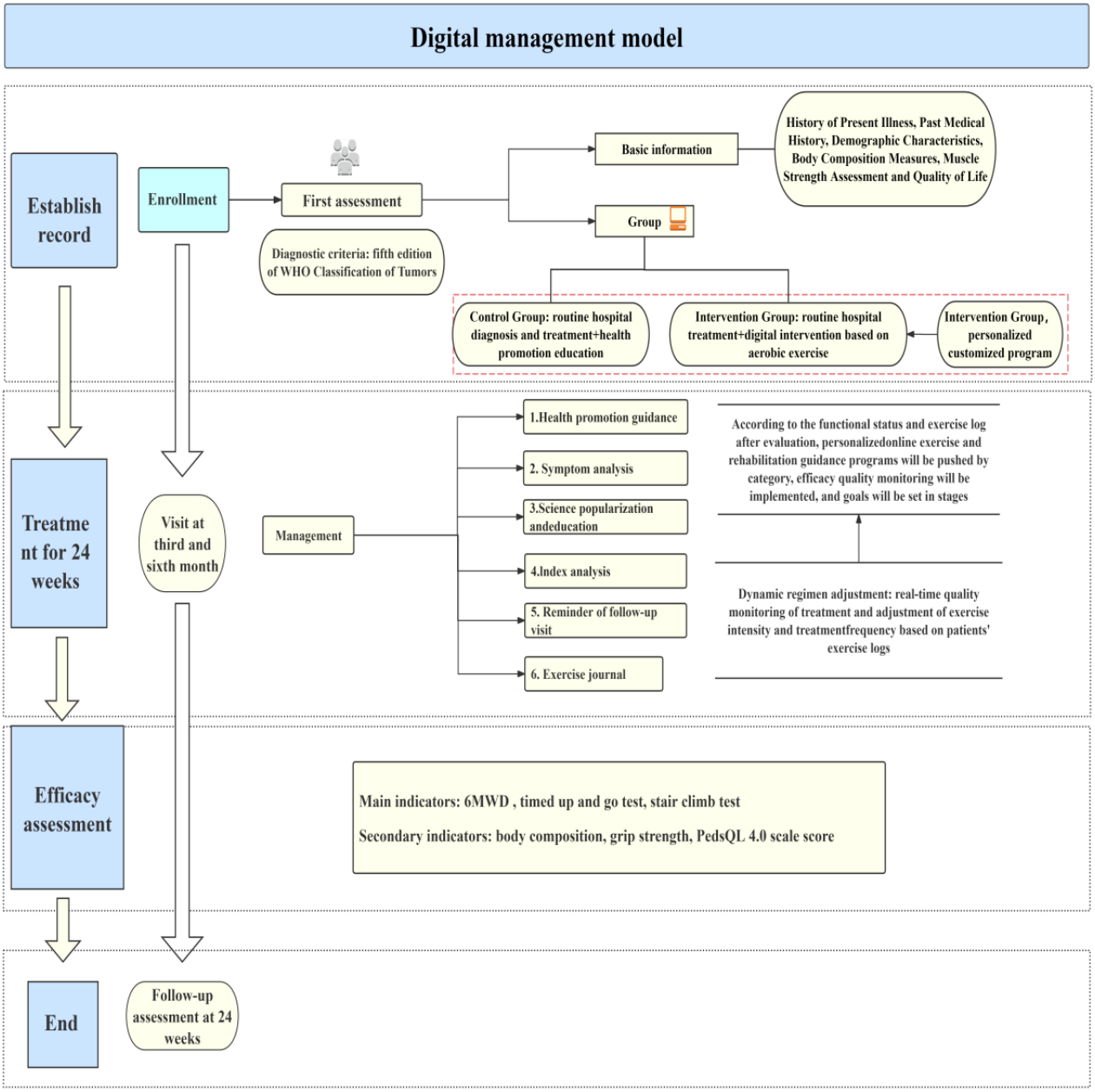

Participants

Children aged 6 to 17 years with a confirmed diagnosis of neuroblastoma will be recruited from outpatient clinics at the research center. Participants will be recruited based on inclusion and exclusion criteria.

Inclusion Criteria

The inclusion criteria are as follows: (1) confirmed diagnosis of neuroblastoma requiring radiotherapy or chemotherapy, (2) age of 6 to 17 years, (3) ability to understand and express themselves to some degree, (4) ability of the patient or their guardian to operate a smartphone app, (5) awareness of the patient's condition and prognosis by the patient or their guardian, and (6) voluntary participation with signed informed consent from the patient and their guardian.

Exclusion Criteria

The exclusion criteria are as follows: (1) preexisting malignant tumors or other types of solid or nonsolid tumors; (2) disease

XSL•FO
RenderX

progression or recurrence at the time of study enrollment; (3) heart rate (HR) exceeding 70% of the age-specific maximum or below 60 beats per minute, systolic blood pressure of >180 mm Hg or orthostatic hypotension, respiratory rate of >50 breaths per minute, body temperature of >37.5 °C, or oxygen saturation of <90%; (4) occurrence within 3 months before enrollment of cardiac insufficiency (New York Heart Association functional class II or higher), severe pulmonary function decline, fracture, hypertension of grade ≥2, or concurrent use of ≥2 antihypertensive medications; (5) history of HIV infection, active hepatitis B virus or hepatitis C virus infection, or uncontrolled systemic infection; and (6) major surgery within 4 weeks before enrollment or planned major surgery during the study period.

Grouping and Masking

Participants with a clear preference for either intervention will be allocated accordingly, whereas those without a preference will be randomly assigned to the experimental or control group. This approach was adopted to accommodate the preferences of the participants and their guardians while preserving the benefits of randomization. Due to the nature of the intervention, participants and intervention providers will not be blinded to group allocation. However, outcome assessors and statisticians will remain blinded to group assignments to minimize bias.

Intervention

Experimental Intervention Program

The experimental group will undergo a baseline assessment using a self-developed questionnaire including demographic and disease information (obtained from hospital records), height, weight, BMI, 6-minute walk distance (6MWD), body composition, grip strength, and Pediatric Quality of Life Inventory version 4.0 (PedsQL 4.0) score. On the basis of the assessment results, a rehabilitation therapist will design a personalized aerobic exercise intervention program, including exercise prescriptions, home exercise routines, physical activity promotion, and daily activity management.

A self-developed digital management system will provide real-time monitoring of exercise logs and deliver tailored exercise strategies. The intensity of aerobic exercise will be adjusted based on individual fatigue levels measured using the Borg fatigue scale. The intervention will be conducted once daily for 30 minutes, 5 days per week over 24 weeks. The digital platform will also offer educational content and personalized information based on the patient's condition and exercise progress.

Outline of the Exercise-Based Digital Therapy

The aerobic digital intervention exercise training program is tailored to participants' Borg fatigue scale scores, with 3 intensity levels designed to accommodate varying levels of fatigue, as shown in Table 1. Each program requires 30 minutes of exercise per day, 5 days per week for 6 months.

XSL•FO
RenderX

Table 1. Exercises included in each program.

Exercises	Description
Low-intensity program (Borg fatigue score 6-10)	
Seated arm stretch	Participants sit upright with their fingers crossed. They straighten their arms and raise their palms above their head. They hold for 30 seconds to feel a stretch in their arms and upper chest.
Backward arm lift	Participants cross their fingers behind their back with their palms facing each other. They slowly turn their elbows inward (palms outward) while straightening their arms. They lift their arms upward and away from their hips until they feel a stretch in their arms, shoulders, or chest. They hold for 30 seconds, keeping their chest lifted and chin tucked.
Seated leg stretch	Participants sit on the edge of a chair with one leg extended and the heel on the floor. They place their hand on the opposite thigh for support. They bend slightly forward at the hips until they feel a stretch in the back of the extended leg. They hold for 10 to 15 seconds and then switch legs.
Supported knee bend	Participants face a wall or stable surface for support. They place their forearm on the wall and rest their head on the back of their hand. They bend one knee and bring it close to the support, keeping the other leg straight. They hold for 30 seconds and then switch legs.
Heel-to-buttock stretch	Participants hold the top of their right foot with their left hand and gently pull their heel toward their buttocks. They hold for 30 seconds and then switch legs.
Waist stretch	Participants stand with their feet shoulder-width apart. They cross their left hand over their waist and raise their right hand above their head. They bend their upper body to the left, hold for 15 seconds, and repeat on the opposite side.
Moderate-intensity program (Borg fatigue score 11-12)	
Standing heel lift	Participants stand upright with their feet shoulder-width apart and arms hanging naturally. They lift both heels off the ground, extending their feet fully. They hold for 5 seconds, then lower their heels and repeat 5 times.
Lunge with arm raise	Participants stand with their feet shoulder-width apart and hands on their hips. They lunge forward with their left leg, bending both knees until their right knee nearly touches the ground. They raise their arms above their head, interlocking their palms. They hold for 15 seconds and then return to the starting position.
One-legged stand	Participants stand near a wall for support. They lift one foot slightly off the ground, bending the other knee. They hold for up to 30 seconds and then switch legs.
Back stretch	Participants position themselves on all fours with their hands under their shoulders and their knees under their hips. They arch their back upward, holding for 10 seconds, and then lower their back slightly. They repeat for 30 to 60 seconds.
Standing shrug	Participants stand with their feet shoulder-width apart, chest up, and core engaged. They let their arms hang naturally at their sides with palms facing in. To perform the movement, they elevate their shoulders toward their ears in a controlled motion, hold the contraction briefly at the top, then slowly lower back to the start.
Kneeling push-ups	Participants place their hands on the ground slightly wider than shoulder-width apart, with their knees hip-width apart. They bend their elbows to lower their chest toward the ground and then push back up.
High-intensity program (Borg fatigue score 13-14)	
In-place high knee raises	Participants stand with their hands on their waist and upper body straight. They lift their right knee to form a right angle with their thigh and calf and then alternate with their left knee in a running motion.

XSL•FO
RenderX

Exercises	Description
Deep squat	Participants stand with their back against a wall and feet slightly apart. They slide down the wall until their thighs are parallel to the floor, keeping their hips higher than their knees. They hold for 5 to 10 seconds and then return to the starting position.
Knee straightening	Participants sit in a chair with their feet together. They straighten one knee, hold for a few seconds, and then slowly lower it and repeat with the other leg.
Wall push	Participants place their hands at shoulder height on a wall with their fingers pointing upward. They bend their elbows to bring their body closer to the wall and then push back until their arms are straight.
Plank support	Participants position themselves on their elbows and toes, keeping their body in a straight line from the shoulders to the ankles, and hold for 30 seconds.
Step-up squat	Participants stand in front of a sturdy bench with one leg bent back and their toes on the bench. Participants bend their forward leg until their thigh is parallel to the floor and then return to the starting position. Participants complete 10 repetitions and then switch legs.

Control Group Program

Children with neuroblastoma who meet the inclusion criteria will be introduced to the study's purpose. After obtaining informed consent, a baseline survey will be conducted using a self-developed questionnaire that includes demographic and disease information extracted from the hospital's information system. Additional baseline measurements include height, weight, BMI, 6MWD, body composition, grip strength, and PedsQL 4.0 score. The control group will receive standard in-hospital medical treatment and health education. At the 3- and 6-month follow-up visits, outcome assessments and questionnaires are administered to the control group (demographic and baseline disease information are not collected).

Compliance Supervision

All personnel involved in the study will be qualified and have undergone prestudy training to ensure role clarity and standardization. Data collectors, data entry personnel, rehabilitation therapists, and digital therapy supervisors will complete training and pass examinations to ensure consistency. Participants in the experimental group will be added to a digital management platform (WeChat app) where they will log daily exercise sessions and receive educational content and virtual or physical rewards based on their adherence. A home training diary will be maintained by patients or their guardians on the digital platform to record exercise time and frequency. Quality control personnel will regularly monitor adherence through the platform. Participants who complete the study will receive additional small toy incentives or follow-up program instruction sessions. Adherence to digital therapy will be measured as the percentage of completed daily log-ins relative to the minimum required (80% adherence threshold). Participants with adherence below 80% will be contacted by the researchers to document reasons for noncompliance.

Withdrawal Criteria

Participants will be withdrawn from the study for the following reasons: (1) serious violation of the test protocol by the participant or their guardian or (2) receipt of other treatments by the participants during the trial that interfere with the study.

Suspension Criteria

This study will be suspended or terminated for participants under the following circumstances: (1) failure to complete the 6-month intervention as required by the study protocol, (2) occurrence of serious safety issues that compromise participant safety, (3) significant protocol deviations that hinder efficacy or safety evaluation, (4) a decision by the ethics committee to suspend or discontinue the study, or (5) a decision by the study sponsor to suspend or halt the study.

Adverse Events

All adverse events (AEs) occurring during the trial will be documented in an AE form. Investigators will assess the relationship between AEs and the intervention and determine whether to discontinue the intervention. AEs will be followed up on until resolution.

Comprehensive Assessment

Efficacy assessments will be conducted by qualified rehabilitation therapists at baseline and 3 and 6 months. Assessments will include (1) demographic and physical developmental indicators (height, weight, and BMI), (2) 6MWD, (3) timed up and go (TUG) test, (4) stair climb test (SCT), (5) body composition, (6) grip strength, and (7) PedsQL 4.0 score (Table 2).

XSL•FO
RenderX

Table 2. Study timeline.

	Screening entry period: visit 0 (day 1)	Treatment period: visit 1, week 12 (–6 days to +6 days)	Treatment period: visit 2, week 24 (–6 days to +6 days)
Informed consent	✓		
Inclusion and exclusion criteria	✓		
Basics	✓		
Measurement of human body parameters	✓	✓	✓
Instrument-based test			
Body composition determination	✓	✓	✓
Muscle power	✓	✓	✓
Quality of life			
Score on the Chinese version of the PedsQL 4.0[a]	✓	✓	✓
Outcome measures			
6-minute walk to test walking distance	✓	✓	✓
Timed up and go test	✓	✓	✓
Stair climb test	✓	✓	✓
Adverse event assessment	✓	✓	✓

[a]PedsQL 4.0: Pediatric Quality of Life Inventory version 4.0.

Primary Outcome Measure

6MWD Measure

The 6-minute walk test is conducted in a flat, 30-m corridor. Participants are instructed to walk as far as possible within 6 minutes without running or jogging. The total distance walked back and forth along the corridor during the 6-minute period is recorded. The distance is measured in meters with an accuracy of 0.1 m. Vital signs, including HR (pretest resting HR and immediate posttest HR), respiratory rate, blood pressure, oxygen saturation, and the Borg subjective fatigue score, are assessed and documented both before and after the test.

For the test environment, a 30 m–long corridor with marked lines at 1-m intervals will be used, and bright-colored strips are attached to the ground to indicate the start and end point of each 60-m segment.

The test requires that participants wear comfortable, walkable shoes; avoid strenuous exercise for at least 2 hours before the test; and rest for a minimum of 10 minutes before the test begins.

If a participant experiences physical discomfort during the test, they may take a short break by leaning against the wall and resume walking once they feel better. The timer will not be paused during this process. If the test is terminated early (before 6 minutes), the following information must be recorded: (1) the time at which the test was stopped, (2) the reason for early termination, and (3) the total distance walked up to that point.

TUG Test

The TUG test will be used to assess functional mobility due to its reliability, cost-effectiveness, and efficiency [31]. During the test, the participants will be instructed to rise from a chair, walk a distance of 3 m, turn around, return to the chair, and sit down again. The total time taken to complete this task will be quantified in seconds. Performance is inversely related to the score, meaning that a shorter time indicates better functional mobility.

SCT Measurement

The SCT will be used to assess functional mobility, including domains such as strength, speed, and axial movement (rotation). The SCT requires participants to ascend a 10-step staircase, promptly turn around at the top, and descend back to the starting point. Children will be instructed to perform the task as quickly and safely as possible without pausing and are permitted to use the handrails. The total completion time, recorded in seconds, serves as the primary outcome measure, with shorter times reflecting better performance [31].

Secondary Outcome Measures

Body Composition Index

A body composition analyzer (InBody 260) will be used to measure the patients' body composition: body fat percentage, basal metabolic rate, skeletal muscle mass, waist-to-hip ratio, body water, protein, and inorganic salt content.

Grip Strength

Grip strength will be measured using a TKK-5401 grip strength meter produced by Takei. The participants' body will be in an upright position, with their feet shoulder-width apart, arms naturally hanging down, and palms inward and one hand holding the grip strength meter at full grip strength. During the measurement, the assessor must ensure that the grip strength meter does not touch the participant's body or clothing. The value displayed on the device is then recorded by the assessor.

XSL•FO
RenderX

The test is performed twice (alternating hands), a 15-second break is allowed between measurements, and the best results are retained for the stronger hand, with grip strength measured in kilograms accurate to 0.1 kg.

Score on the PedsQL 4.0 Scale

The Chinese version of the PedsQL 4.0 scale will be used in this study. This scale has demonstrated good reliability and validity and has been widely adopted in research assessing quality of life among children in China. The PedsQL 4.0 scale comprises 23 items, which are categorized into 4 domains: physical functioning, emotional functioning, social functioning, and school functioning. Each item is rated on a 5-point Likert scale ranging from 0 to 4. During scoring, responses are transformed into a scale from 0 to 100, where higher scores indicate better outcomes. The domain score is calculated by summing the scores of all items within that domain and dividing by the number of items in the domain. Similarly, the total summary score is derived by summing the scores of all 23 items and dividing by the total number of items. Both the domain scores and the total summary score range from 0 to 100, with higher scores reflecting a better quality of life.

Statistical Analysis

Data will be analyzed using SPSS (version 26.0; IBM Corp). Primary models using the intention-to-treat principle will be conducted for all primary and secondary outcomes. Data normality will be assessed initially. Nonnormally distributed data will be presented as medians with IQRs. Between-group comparisons (intervention vs control) for these variables will be conducted using the Mann-Whitney *U* test. Within-group comparisons (baseline vs each follow-up) will be conducted using the Wilcoxon signed rank test. Categorical data will be reported as numbers with percentages and analyzed using the chi-square test. Differences in primary and secondary outcomes between the intervention and control groups across all assessment time points will be evaluated using mixed-model analysis. Statistical significance will be defined as a *P* value of <.05, with multiple comparisons adjusted using the Bonferroni method.

Sample Size Calculation

The sample size was calculated using G*Power (version 3.1.9.7; Heinrich Heine University Düsseldorf). On the basis of a related study [32], an effect size of 0.953 was determined, derived from the difference in the 6MWD between the experimental group (mean 35.89, SD 8.46 laps) and the control group (mean 26.76, SD 10.57 laps) after the intervention. The significance level (α) was set at .05, and the power ($1 - \beta$) was set at 0.95. The calculation indicated that 30 participants were required for each group. To account for a potential 10% loss to follow-up, 34 participants per group were planned for recruitment, resulting in a total sample size of 68 participants.

Quality Control and Quality Assurance

All researchers involved in this study possess appropriate clinical research qualifications. Before the commencement of the clinical trial, the project leader or principal investigator conducted comprehensive training sessions for the research team. Personnel responsible for evaluations were specifically trained in evaluation procedures and operational methods, with their competency regularly assessed. Quality control measures include monthly monitoring and evaluation by inspectors within each group, as well as quarterly monitoring and evaluation by the study team's inspectors. To ensure the integrity of the study, a data and safety monitoring committee was established to oversee the quality of each process. The clinical research unit is responsible for maintaining the quality and safety of the study, ensuring both the safety of participants and the validity of the data. The trial may be terminated if the intervention is found to be ineffective or if significant safety concerns arise based on the assessment of data quality, efficacy outcomes, and AEs. In the event of an AE, the principal investigator will evaluate the situation and determine whether to suspend the intervention.

Ethical Considerations

Ethics and Informed Consent

This study has been approved by the Medical Ethics Committee of Xinhua Hospital, affiliated to Shanghai Jiao Tong University School of Medicine (XHEC-C-2023-062), and all study participants will provide written informed consent.

Dissemination and Publication Policy

Study results will be published in peer-reviewed journals and presented at conferences and relevant stakeholder engagement activities. There will be no information that can identify participants in any publication. Participants who explicitly express a wish to be informed about the research outcome will be contacted and offered to receive an article or poster with a lay summary of the study.

Results

Recruitment will be conducted from December 2025 to September 2026. It is expected that the intervention group will show significantly greater improvements in physical function, physical activity levels, and fatigue management compared to the control group and that the effectiveness of digital therapy for rehabilitation training after radiotherapy and chemotherapy for neuroblastoma in children can be demonstrated. The findings of this study are anticipated to be published by late 2026 or early 2027.

Discussion

Expected Findings

DTx represent an innovative treatment modality that facilitates real-time remote guidance and intervention quality monitoring. This approach reduces the need for frequent hospital visits while enhancing the efficiency of interventions [29]. Aerobic exercise has been widely recognized for its significant role during and after oncology treatment, contributing to improvements in physical fitness, muscle strength, and functional capacity. Additionally, it alleviates symptoms such as pain and diarrhea, shortens hospitalization duration, and promotes mental health and emotional well-being [33]. It should be pointed out that the primary goal of this "aerobic digital intervention" program is not to directly maximize cardiorespiratory endurance (such as maximal oxygen consumption) in the traditional sense but to

XSL•FO
RenderX

provide a safe, feasible, and tolerable home physical activity initiation program for participants with different levels of fatigue to improve physical function, physical activity level, and fatigue management.

The integration of aerobic exercise interventions with DTx has shown promise in enhancing intervention efficacy through personalized, multilevel management of exercise programs. To the best of the authors' knowledge, this study is the first controlled trial to evaluate the impact of an aerobic exercise intervention combined with digital therapy on motor performance in children with neuroblastoma.

The primary objective of this study is to establish a digital management intervention protocol for children with neuroblastoma by comparing the effectiveness of DTx combined with an aerobic exercise intervention against traditional in-hospital visits and health promotion strategies. The study outcomes are expected to demonstrate significant improvements in cardiopulmonary fitness, muscle strength, body composition, and quality of life. These findings will provide an evidence-based rationale for the future implementation of digital therapy–supported aerobic exercise programs in this population.

Limitations

Despite its potential, this study has several limitations that should be considered when interpreting the results. First, the single-center design and small sample size restrict the generalizability of the findings. Future research should involve multiple centers across different regions and include larger samples to enhance external validity. Second, although outcome assessors and statisticians were blinded, participants and care providers were not. Furthermore, for participants with a preexisting preference, allocation was based on that preference rather than randomization. This nonrandomized allocation for a subset of participants introduces potential performance and expectation biases. As a preliminary investigation, this study provides foundational insights that can inform the design of a more rigorous randomized controlled trial in the future. Finally, the digital therapy platform and exercise programs are self-developed, and no formal large-scale usability testing was conducted before the intervention. While the design was informed by pediatric expertise and in-study engagement metrics supported feasibility, future iterations would benefit from structured usability testing to optimize the platform before broader implementation.

Conclusions

This study represents an important first step in aerobic digital therapy for pediatric neuroblastoma. It provides evidence that rehabilitation after chemoradiotherapy in pediatric patients with neuroblastoma can improve physical function, physical activity levels, and fatigue management. In the future, expanding to multiple centers to conduct targeted surveys of different types of childhood cancers will help validate the effectiveness of digital therapy–based aerobic exercise.

Authors' Contributions

XZ and QD designed the study. XZ, SH, and LW drafted the study protocol and developed the recruitment plan and data collection materials. QD, SH, HG, and SC provided guidance on the methodology of this study and codeveloped the intervention protocol. XZ, LW, and TL completed the ethics approval application. All authors critically reviewed and edited the final manuscript.

Conflicts of Interest

None declared.

References

1. Wahba A, Wolters R, Foster JH. Neuroblastoma in the era of precision medicine: a clinical review. Cancers (Basel) 2023 Sep 26;15(19):4722 [FREE Full text] [doi: 10.3390/cancers15194722] [Medline: 37835416]
2. 2Anderson J, Majzner RG, Sondel PM. Immunotherapy of neuroblastoma: facts and hopes. Clin Cancer Res 2022;28(15):3196-3206. [doi: 10.1158/1078-0432.ccr-21-1356]
3. Schirrmacher V. From chemotherapy to biological therapy: a review of novel concepts to reduce the side effects of systemic cancer treatment (Review). Int J Oncol 2019 Feb 10;54(2):407-419 [FREE Full text] [doi: 10.3892/ijo.2018.4661] [Medline: 30570109]
4. Qiu B, Matthay KK. Advancing therapy for neuroblastoma. Nat Rev Clin Oncol 2022 Aug 25;19(8):515-533. [doi: 10.1038/s41571-022-00643-z] [Medline: 35614230]
5. Croteau N, Nuchtern J, LaQuaglia MP. Management of neuroblastoma in pediatric patients. Surg Oncol Clin N Am 2021 Apr;30(2):291-304. [doi: 10.1016/j.soc.2020.11.010] [Medline: 33706901]
6. Zheng DJ, Krull KR, Chen Y, Diller L, Yasui Y, Leisenring W, et al. Long-term psychological and educational outcomes for survivors of neuroblastoma: a report from the Childhood Cancer Survivor Study. Cancer 2018 Aug 01;124(15):3220-3230 [FREE Full text] [doi: 10.1002/cncr.31379] [Medline: 29888538]
7. Butler E, Ludwig K, Pacenta HL, Klesse LJ, Watt TC, Laetsch TW. Recent progress in the treatment of cancer in children. CA Cancer J Clin 2021 Jul;71(4):315-332 [FREE Full text] [doi: 10.3322/caac.21665] [Medline: 33793968]
8. Maddocks M. Physical activity and exercise training in cancer patients. Clin Nutr ESPEN 2020 Dec;40:1-6. [doi: 10.1016/j.clnesp.2020.09.027] [Medline: 33183519]

XSL•FO
RenderX

9. Horowitz AM, Fan X, Bieri G, Smith LK, Sanchez-Diaz CI, Schroer AB, et al. Blood factors transfer beneficial effects of exercise on neurogenesis and cognition to the aged brain. Science 2020 Jul 10;369(6500):167-173 [FREE Full text] [doi: 10.1126/science.aaw2622] [Medline: 32646997]
10. Van Dijk-Lokkart EM, Steur LM, Braam KI, Veening MA, Huisman J, Takken T, et al. Longitudinal development of cancer-related fatigue and physical activity in childhood cancer patients. Pediatr Blood Cancer 2019 Dec 22;66(12):e27949. [doi: 10.1002/pbc.27949] [Medline: 31436372]
11. Buss LA, Dachs GU. Effects of exercise on the tumour microenvironment. Adv Exp Med Biol 2020;1225:31-51. [doi: 10.1007/978-3-030-35727-6_3] [Medline: 32030646]
12. Hojman P, Gehl J, Christensen JF, Pedersen BK. Molecular mechanisms linking exercise to cancer prevention and treatment. Cell Metab 2018 Jan 09;27(1):10-21 [FREE Full text] [doi: 10.1016/j.cmet.2017.09.015] [Medline: 29056514]
13. Braam KI, van der Torre P, Takken T, Veening M, van Dulmen-den Broeder E, Kaspers GJ. Physical exercise training interventions for children and young adults during and after treatment for childhood cancer. Cochrane Database Syst Rev 2013 Apr 30;3(4):CD008796. [doi: 10.1002/14651858.CD008796.pub2] [Medline: 23633361]
14. Coombs A, Schilperoort H, Sargent B. The effect of exercise and motor interventions on physical activity and motor outcomes during and after medical intervention for children and adolescents with acute lymphoblastic leukemia: a systematic review. Crit Rev Oncol Hematol 2020 Aug;152:103004 [FREE Full text] [doi: 10.1016/j.critrevonc.2020.103004] [Medline: 32580035]
15. Santos LS, Rehder MS, Negrao MV, Goes-Santos BR, Toshi Dias E, Paixão CJ, et al. Aerobic exercise training combined with local strength exercise restores muscle blood flow and maximal aerobic capacity in long-term Hodgkin lymphoma survivors. Am J Physiol Heart Circ Physiol 2024 Apr:1462-1468. [doi: 10.1152/ajpheart.00132.2024] [Medline: 38639741]
16. Campbell KL, Winters-Stone KM, Wiskemann J, May AM, Schwartz AL, Courneya KS, et al. Exercise guidelines for cancer survivors: consensus statement from international multidisciplinary roundtable. Med Sci Sports Exerc 2019 Nov;51(11):2375-2390 [FREE Full text] [doi: 10.1249/MSS.0000000000002116] [Medline: 31626055]
17. Rodríguez-Cañamero S, Cobo-Cuenca AI, Carmona-Torres JM, Pozuelo-Carrascosa DP, Santacruz-Salas E, Rabanales-Sotos JA, et al. Impact of physical exercise in advanced-stage cancer patients: systematic review and meta-analysis. Cancer Med 2022 Oct 11;11(19):3714-3727 [FREE Full text] [doi: 10.1002/cam4.4746] [Medline: 35411694]
18. Leclerc AF, Foidart-Dessalle M, Tomasella M, Coucke P, Devos M, Bruyère O, et al. Multidisciplinary rehabilitation program after breast cancer: benefits on physical function, anthropometry and quality of life. Eur J Phys Rehabil Med 2017 Oct;53(5):633-642 [FREE Full text] [doi: 10.23736/S1973-9087.17.04551-8] [Medline: 28322035]
19. Tanriverdi M, Cakir E, Akkoyunlu ME, Cakir FB. Effect of virtual reality-based exercise intervention on sleep quality in children with acute lymphoblastic leukemia and healthy siblings: a randomized controlled trial. Palliat Support Care 2022 Aug 15;20(4):455-461. [doi: 10.1017/S1478951522000268] [Medline: 35289266]
20. Blanchard CM, Courneya KS, Stein K, American Cancer Society's SCS-II. Cancer survivors' adherence to lifestyle behavior recommendations and associations with health-related quality of life: results from the American Cancer Society's SCS-II. J Clin Oncol 2008 May 01;26(13):2198-2204. [doi: 10.1200/JCO.2007.14.6217] [Medline: 18445845]
21. Le A, Mitchell HR, Zheng DJ, Rotatori J, Fahey JT, Ness KK, et al. A home-based physical activity intervention using activity trackers in survivors of childhood cancer: a pilot study. Pediatr Blood Cancer 2017 Feb 12;64(2):387-394. [doi: 10.1002/pbc.26235] [Medline: 27615711]
22. Hong JS, Wasden C, Han DH. Introduction of digital therapeutics. Comput Methods Programs Biomed 2021 Sep;209:106319. [doi: 10.1016/j.cmpb.2021.106319] [Medline: 34364181]
23. Dang A, Arora D, Rane P. Role of digital therapeutics and the changing future of healthcare. J Family Med Prim Care 2020 May;9(5):2207-2213 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_105_20] [Medline: 32754475]
24. Recommendations on digital interventions for health system strengthening. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/9789241550505 [accessed 2025-10-11]
25. Moschonis G, Siopis G, Jung J, Eweka E, Willems R, Kwasnicka D, et al. Effectiveness, reach, uptake, and feasibility of digital health interventions for adults with type 2 diabetes: a systematic review and meta-analysis of randomised controlled trials. Lancet Digit Health 2023 Mar:125-143. [doi: 10.1016/S2589-7500(22)00233-3] [Medline: 36828606]
26. Parker K, Uddin R, Ridgers ND, Brown H, Veitch J, Salmon J, et al. The use of digital platforms for adults' and adolescents' physical activity during the COVID-19 pandemic (our life at home): survey study. J Med Internet Res 2021 Feb 01;23(2):e23389 [FREE Full text] [doi: 10.2196/23389] [Medline: 33481759]
27. Eppes EV, Augustyn M, Gross SM, Vernon P, Caulfield LE, Paige DM. engagement with and acceptability of digital media platforms for use in improving health behaviors among vulnerable families: systematic review. J Med Internet Res 2023 Feb 03;25:e40934 [FREE Full text] [doi: 10.2196/40934] [Medline: 36735286]
28. Li G, Zhou X, Deng J, Wang J, Ai P, Zeng J, et al. Digital therapeutics-based cardio-oncology rehabilitation for lung cancer survivors: randomized controlled trial. JMIR Mhealth Uhealth 2025 Feb 25;13:e60115 [FREE Full text] [doi: 10.2196/60115] [Medline: 39999435]
29. Zhang ZY, Tian L, He K, Xu L, Wang XQ, Huang L, et al. Digital rehabilitation programs improve therapeutic exercise adherence for patients with musculoskeletal conditions: a systematic review with meta-analysis. J Orthop Sports Phys Ther 2022 Nov;52(11):726-739. [doi: 10.2519/jospt.2022.11384] [Medline: 35960507]

XSL•FO
RenderX

30. Rincón-Castanedo C, Martín-Ruiz A, Zazo S, Luis Huertas AL, Valenzuela PL, Morán M, et al. Combined exercise intervention in a mouse model of high-risk neuroblastoma: effects on physical, immune, tumor and clinical outcomes. Exerc Immunol Rev 2023;29:86-110 [FREE Full text] [Medline: 37358366]
31. Sinanoglu B, Ozdemir F. Evaluation of functional mobility, balance, and executive functions in children with epilepsy. Epilepsy Behav 2023 Jun;143:109230. [doi: 10.1016/j.yebeh.2023.109230] [Medline: 37146465]
32. Tanir MK, Kuguoglu S. Impact of exercise on lower activity levels in children with acute lymphoblastic leukemia: a randomized controlled trial from Turkey. Rehabil Nurs 2013;38(1):48-59. [doi: 10.1002/rnj.58] [Medline: 23365005]
33. Knols R, Aaronson NK, Uebelhart D, Fransen J, Aufdemkampe G. Physical exercise in cancer patients during and after medical treatment: a systematic review of randomized and controlled clinical trials. J Clin Oncol 2005 Jun 01;23(16):3830-3842. [doi: 10.1200/JCO.2005.02.148] [Medline: 15923576]

Abbreviations

6MWD: 6-minute walk distance
AE: adverse event
DTx: digital therapeutics
HR: heart rate
PedsQL 4.0: Pediatric Quality of Life Inventory version 4.0
SCT: stair climb test
TUG: timed up and go

Edited by J Sarvestan; submitted 15.Feb.2025; peer-reviewed by LE van den Berg, Y Bai, B Yu; comments to author 18.Jun.2025; accepted 10.Nov.2025; published 08.Jan.2026.

Please cite as:
Zhou X, Huang S, Wang L, Guo H, Chen S, Li T, Du Q
The Effect of Digital Therapy–Based Aerobic Exercise on Children With Neuroblastoma: Protocol for a Prospective Controlled Trial
JMIR Res Protoc 2026;15:e72698
URL: https://www.researchprotocols.org/2026/1/e72698
doi:10.2196/72698
PMID:

©Xuan Zhou, Shihao Huang, Lixia Wang, Haibin Guo, Sisi Chen, Tingrui Li, Qing Du. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 08.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Association Between Mandibular Third Molar Position and Recurrent Pericoronitis: Protocol for a Cross-Sectional Study

Suwarna Dangore- Khasbage[1*], BDS, MDS; Tanaya Teredesai[1*], BDS

Department of Oral Medicine and Radiology, Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research, Wardha, Maharashtra, India

*all authors contributed equally

Corresponding Author:
Suwarna Dangore- Khasbage, BDS, MDS
Department of Oral Medicine and Radiology
Sharad Pawar Dental College and Hospital
Datta Meghe Institute of Higher Education and Research
Sawangi
Wardha, Maharashtra, 442001
India
Phone: 91 9923024036
Email: dangoresuwarna67@gmail.com

Abstract

Background: Pericoronitis is a common pathological condition associated with mandibular third molars that may cause pain and discomfort. This condition may be chronic, exhibiting episodic symptoms that last for a few days to weeks and recurring multiple times in less than a year. The operculum covering the erupting mandibular third molars may become obscured by the eruption of maxillary third molars. Such recurrent traumas might exacerbate the symptoms and lead to ulcerations. With clinical monitoring at regular intervals and with the help of radiographic examinations, clinicians can develop the most effective treatment plan.

Objective: This study aims to determine the association between mandibular third molar position and recurrent pericoronitis.

Methods: This cross-sectional study will include 200 patients having partially impacted mandibular third molar with recurrent pericoronitis. Patients aged 18-40 years with occurrence of pericoronitis will be included in this study. The impacted tooth's side and the symptoms associated with pericoronitis will be recorded during clinical examination. All these patients will be evaluated using panoramic radiographs to assess the position of the unerupted/impacted mandibular third molar.

Results: The duration of this study will be 6 months from October 2025 to April 2026. Approval for this study has been granted by the institutional ethics committee of Datta Meghe Institute of Higher Education and Research (deemed to be university), Sawangi, Wardha (DMIHER (DU)/IEC/2024/53). In panoramic radiographs, the impaction status of the mandibular third molar will be evaluated based on Winter's and Pell and Gregory classification systems. Patients' data will be recorded and analyzed for statistical significance.

Conclusions: The detection of the position and intervention at the early stage for pericoronitis in individuals with impacted third molars is vital.

International Registered Report Identifier (IRRID): PRR1-10.2196/72682

(JMIR Res Protoc 2026;15:e72682) doi:10.2196/72682

KEYWORDS

eruption; molar; orthopantomogram; radiograph; wisdom teeth

Introduction

Pericoronitis is a painful inflammatory condition of the gingival tissue surrounding the crown of a partially erupted tooth. It is often related with the positional relationship of the third molar with the surrounding hard and soft tissues. The condition is more common with mandibular third molars as compared to maxillary third molars [1,2].

The partially erupted tooth is nestled in the posterior recesses of the mandibular arches under the soft tissue or inside the

XSL•FO
RenderX

osseous bounds and adopts a mesioangular, horizontal, or distoangular inclination. This abnormal position of the tooth provides an isolated plaque-retentive situation. Further, this improper placement of the partially erupted third molar creates an environment that is conducive to recurrent episodes of infection, swelling, and discomfort because of the trapping of food particles and harmful bacteria beneath the gingival operculum. This leads to the inflammation of the surrounding soft tissue, that is, pericoronitis. Nevertheless, this condition becomes recurrent due to anatomic inaccessibility for dental hygiene procedures, the pressure from the neighboring teeth, and the trauma from erupting maxillary third molars [1,3].

The common symptoms of pericoronitis include fever, soreness, swelling, pus discharge, and trouble in opening the mouth. This condition may be chronic, exhibiting episodic symptoms that last for a few days to weeks and recurring multiple times in less than a year [4]. If it is not treated in time, it can spread to the surrounding soft as well as hard tissues leading to complications such as maxillofacial space infection, buccal fistula, and osteomyelitis of the jaw [5]. Thus, considering the number of symptoms and accompanying complications, it is recommended to evaluate the position of an impacted third molar that is more at risk to be affected by pericoronitis so as to initiate early intervention.

Few previous studies [1,2,6,7] have reported associations between the position of the mandibular third molar and pericoronitis. However, the results are quite contradictory. Furthermore, despite numerous studies exploring the etiology and management of pericoronitis, a significant research gap remains in understanding the precise correlation between the positional relationship of third molars and the frequency or severity of recurrent pericoronitis. This gap contemplates the need for undertaking this study.

The aim of this study was to determine the mandibular third molar position and its association with recurrent pericoronitis. The research question is as follows: what is the association between the mandibular third molar position and recurrent pericoronitis?

Methods

Ethical Considerations

This cross-sectional study will be performed in the Oral Medicine and Radiology Department, Wardha, Maharashtra, India. The duration of this study will be 6 months from October 2025 to April 2026. Approval for this study has been granted by the Institutional Ethics Committee of Datta Meghe Institute of Higher Education and Research (deemed to be university), Sawangi, Wardha (DMIHER (DU)/IEC/2024/53). The data collected will not include any identifiable information. The objectives of the study will be thoroughly explained to each patient, and informed written consent stating data confidentiality, right to withdraw, and compensation will be obtained prior to any examination.

Study Population

This study recruits patients visiting the outdoor patient department of Oral Medicine and Radiology with a history of pain, hyperemia, edema associated with the mandibular third molar of at least two episodes within a duration of 12 months, and affecting the same partially erupted or impacted tooth.

The inclusion criteria for the study will be patients aged 18-40 years, patients with history of pain, patients with recurrent pericoronitis treated by conservative methods, and patients with teeth in the vicinity of pericoronitis that are healthy and whole. The exclusion criteria will be history of extraction of third molar, missing third molar, fracture due to trauma to impacted or partially erupted third molar, and pregnant or lactating females. The sample size for this study will be determined as shown in Textbox 1 [8,9].

Textbox 1. Sample size estimation.

Sample size

The sample size is calculated based on Daniel formula [9] for sample size estimation:

$n = (Z^2 * P * (1 - P)) / d^2$

Where:

- Z represents confidence level. Confidence level is set at 95%; then the value of Z = 1.96.
- P is the anticipated proportion or prevalence. A common assumption for P is 0.5.
- d is the margin of error or precision. A value of 0.05 is often used to ensure high precision and minimize the error in the estimate.

Z = 1.96

P = prevalence of mandibular third molar with pericoronitis = 54.1% = 0.541

E = acceptable margin of error, which is 7% and equivalent to 0.07

n = 1.96 2 * 0.541 * (1-0.541) 0.072 = 194.7 = 200 patients needed in the study [8]

To ensure adequate power and to account for potential dropouts or loss to follow-up (≈5%) during the 12-month follow-up period, the final sample size was rounded up to 200 participants.

XSL•FO
RenderX

Methodology

A total of 200 patients with recurrent pericoronitis and willing to participate will be selected for this study over a period of 1 year. These patients will be chosen among those reporting to the outpatient department of the Oral Medicine and Radiology department. Recurrent pericoronitis will be typically diagnosed when 2 or more distinct episodes of pericoronitis occur within a duration of 12 months, affecting the same partially erupted or impacted mandibular third molar. After obtaining the informed written consent, clinical as well as radiological evaluation of each patient will be performed.

The first step will be clinical evaluation, which includes a detailed collection of each patient's medical and dental history, information about the duration of the painful episodes, and frequency and severity of painful episodes associated with an impacted mandibular third molar. Clinical examination using diagnostic instruments will focus on identifying the symptoms of pericoronitis, assessing the periodontal and gingival health of the affected third molar with a Williams's periodontal probe, and determining the carious status of adjacent teeth.

The second step will be to analyze the position of the third molar by using panoramic radiography. To evaluate its angulation and position, Winter's and Pell and Gregory classification systems will be applied to categorize partially erupted or impacted third molars [10-12]. These diagnostic measures aim to provide a comprehensive understanding of the clinical and radiographic characteristics of pericoronitis cases, ensuring accurate data for subsequent analysis and interpretation.

The radiographs will be assessed by a postgraduate student who will be blinded and a senior experienced radiologist who will not be blinded. Thus, this study will be single blinded. Interobserver reliability will be calculated by intraclass correlation before the inclusion of the radiographs. Patients in the age group of 18-40 years will be thoroughly examined to ensure comprehensive data collection. All the participants will be followed for future recurrent episodes for a duration of 12 months with an interval of every 3 months or upon patient-reported symptoms.

Results

The duration of this study will be 6 months from October 2025 to April 2026. Approval for this study has been granted by the Institutional Ethics Committee of Datta Meghe Institute of Higher Education and Research (deemed to be university), Sawangi, Wardha (DMIHER (DU)/IEC/2024/53). Only those individuals who meet the inclusion and exclusion criteria will be considered as study participants. The parameters that will be determined are mandibular third molar position and episodes of recurrent pericoronitis. Data analysis will evaluate the primary outcome that includes the association between mandibular third molar position and recurrence of pericoronitis as well as secondary outcomes such as number and severity of episodes, angulation and depth of impaction, and comparison between Winter's and Pell and Gregory classification systems. The distribution of the patients according to episodes, total duration of discomfort, type of impaction, depth of impaction, and Position of Ramus will be evaluated. All the findings will be recorded in a tabular form similar to that shown in Table 1 and Table 2. Statistical analysis will be done using descriptive statistics. The complete analysis dataset will consist of all the participants who have no missing values for any of the parameters.

Table 1. Association between mandibular third molar position according to Winter's classification and recurrent pericoronitis.

Tooth number	Position of the impacted tooth according to Winter's classification	Number of episodes of pericoronitis	Clinical status of mandibular third molar
1			
2			
3			
4			

Table 2. Association between mandibular third molar position according to Pell and Gregory classification and recurrent pericoronitis.

Tooth number	Position of the impacted tooth according to Pell and Gregory classification	Number of episodes of pericoronitis	Clinical status of mandibular third molar
1			
2			
3			
4			

Discussion

Principal Findings

This study aims to determine the association between mandibular third molar position and recurrent pericoronitis by using radiographic evaluation. The primary expectation is that certain positional types, as defined by Winter's and Pell and Gregory classifications, are more frequently associated with recurrent pericoronitis. Identifying these positions may help clinicians predict which partially erupted third molar is at higher risk and requires early intervention.

XSL•FO
RenderX

Pericoronitis is a clinical condition often observed in dental practice. The clinical manifestations range from asymptomatic to quite severe in some cases. Usually, the position or path of eruption of mandibular third molar is responsible for pericoronitis. Therefore, the evaluation of the position of mandibular third molars by radiographic examination is vital for detecting the risk of pericoronitis associated with the erupting or impacted mandibular third molar. Thus, the focus of this study is whether the position of the impacted mandibular third molar by using an orthopantomogram is reliable for diagnosing patients who are at a high risk of developing pericoronitis.

Comparison With Available Literature

A 2023 study by Shirzadeh et al [2] assessed the prevalence of different types of lower third molar impactions in individuals with pericoronitis in 60 participants. Their findings indicated that class B, class II, and mesioangular impaction types can lead to higher incidence of pericoronitis.

Galvão et al [11] performed a systematic review and meta-analysis in 2019 to determine the association between the lower wisdom tooth position and the presence of pericoronitis. Findings showed stronger correlation between pericoronitis and the vertical position of the lower wisdom tooth compared to other Winter's classifications. Additionally, within the Pell and Gregory classification, position A had a higher risk of pericoronitis than position B. Their study suggests that the prophylactic extraction of partially erupted lower wisdom tooth, particularly those in position A, is recommended to reduce the risk of pericoronitis.

Prasad [1] concluded that in case of asymptomatic mandibular impacted third molars, the precautionary approach might not be the most effective option. Ongoing monitoring at regular intervals and with the help of patient-centered decision-making, clinicians can develop the most effective treatment plan.

Indira et al [6] reported that the impacted third molar may be associated with the development of pericoronal infectious process and their potential complications and can be advised for prophylactic removal, particularly in position IA, vertical, and distoangular impactions to prevent surgical risks and patients' morbidity with age.

Singh et al [13] stated that the analysis of the patterns of the impacted mandibular molars may lead to early prediction and diagnosis of the various pathologies associated with the impacted mandibular molar.

Erçal [14] highlighted the significant association between the presence of impacted third molars and various pathological conditions. They also stated the importance of a follow-up to monitor possible changes in the future even though most impacted teeth remained asymptomatic. The third molars were the most prevalent impacted teeth, with approximately 92% of all the impacted teeth, followed by canines and supernumerary teeth. A pathology was detected in 38 out of 335 of the impacted teeth in their study.

Santos et al [15] in their cross-sectional study (2020) evaluated how the position of the mandibular third molars affects periodontal status and clinical severity in patients with acute pericoronitis. They found that teeth in Pell and Gregory position IIA showed greater gingival edema and erythema compared to IA. Increased plaque accumulation, deeper periodontal pockets, and reduced alveolar bone height were observed in adjacent second molars. Their study showed that the third molar position significantly influences both local periodontal health and the clinical severity of pericoronitis.

Nguyen et al [16] analyzed the factors influencing the severity of pericoronitis in patients with impacted mandibular third molars in 2024. Using multivariate analysis, they found that distal radiolucency, right-side involvement in females, upper respiratory infection in males, and menstruation were significantly associated with greater severity. Severity scores correlated with increased pain, swelling, and functional limitation. They concluded that both local radiographic features and systemic conditions contribute to the severity of pericoronitis. Similarly, our study will determine the association between the mandibular third molar position and recurrent pericoronitis. That may help in educating the patient about the condition, thereby aiding in the appropriate treatment approach and halting of the progression of the same.

Study Limitations

This study will use the orthopantomogram, which is one of the popular commonly used imaging modalities in dental practice; however, the use of cone beam computed tomography will give 3D information regarding the position of the impacted mandibular third molar. The examiner variability in clinical assessments could lead to minor observer bias despite standardization measures. The 12-month follow-up period may also result in participant attrition, affecting data completeness, although adjustments were made during sample size estimation. Finally, this study is institution-based, which may limit the generalizability of its findings to broader populations.

Broader Implications and Conclusions

Understanding the association between mandibular third molar position and recurrent pericoronitis has important clinical implications. If specific impaction types are found to predispose to recurrence, clinicians can apply radiographic evaluation as a predictive diagnostic tool to identify high-risk patients. This approach can support personalized management strategies such as early prophylactic extraction for high-risk impactions or conservative monitoring for low-risk cases, thereby minimizing patient morbidity and surgical complications. On a broader scale, these findings could guide evidence-based decision-making in oral and maxillofacial practice, reduce unnecessary extractions, and optimize patient outcomes.

Dissemination of the Study Results

The results of this study will be shared through multiple channels to ensure that all stakeholders, including participants, clinical practitioners, and policymakers, are appropriately informed. Academicians will be informed through presentations at conferences, publications in peer-reviewed journals, and clinical bulletins, facilitating integration of the findings into evidence-based practice. All findings will be presented in a manner that ensures participant confidentiality is fully protected.

XSL•FO
RenderX

Acknowledgments

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article. Generative artificial intelligence was not at all used in any portion of the manuscript generation.

Data Availability

All data generated or analyzed during this study will be made available by the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: SD-K
Formal analysis, supervision, project administration: SD-K
Methodology, software, resources, data curation, funding acquisition: TT
Validation, visualization, investigation: TT and SD-K
Writing—original draft preparation: SD-K
Writing—review and editing: SD-K

Conflicts of Interest

None declared.

References

1. Prasad T. Prophylactic extraction of third molars: justified or not? World J Dent 2019;10(1):72-77 [FREE Full text] [doi: 10.5005/jp-journals-10015-1606]
2. Shirzadeh A, Shirvan SB, Alizadeh O, Grillo R, Vida M, Samieirad S. What is the most prevalent type of third molar impaction in patients with pericoronitis? World J Plast Surg 2023 May 01;12(2):57-63 [FREE Full text] [doi: 10.61186/wjps.12.2.57]
3. Doni BR, Srinivasan A, Issrani R, Agrawal R, Patil S, Alam M. The effect of pain and swelling related to third molars on oral health-related quality of life. Pesqui Bras Odontopediatria Clín Integr 2021;21:e0190 [FREE Full text] [doi: 10.1590/pboci.2021.056]
4. McNutt M, Partrick M, Shugars DA, Phillips C, White RP. Impact of symptomatic pericoronitis on health-related quality of life. J Oral Maxillofac Surg 2008 Dec;66(12):2482-2487. [doi: 10.1016/j.joms.2008.07.005] [Medline: 19022127]
5. Li Z, Fan B, Wu T, Li Z, Yu X, Zhang B. Special diffusion pathway of pericoronitis of the third molar: a case report and literature review. Int J Surg Case Rep 2024 Jun;119:109709 [FREE Full text] [doi: 10.1016/j.ijscr.2024.109709] [Medline: 38718492]
6. Indira A, Kumar M, David M, Rajshekar V, Shashikala. Correlation of pericoronitis and the status of eruption of mandibular third molar: a clinico radiographic study. J Indian Aca Oral Med Radiol 2013;25(2):1-4 [FREE Full text] [doi: 10.5005/jp-journals-10011]
7. Singh R, Devanna R, Tenglikar P, Gautam A, Anubhuti, Kumari P. Evaluation of mandibular third molar position as a risk factor for pericoronitis: a CBCT study. J Family Med Prim Care 2020 Mar;9(3):1599-1602 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_1101_19] [Medline: 32509657]
8. Umaiyal P M, Ramamurthy J. Prevalence of mandibular third molar pericoronitis among smokers and evaluation of its treatment outcomes: a retrospective study. Int J Res Pharm Sci 2020 Sep 12;11(SPL3):452-458. [doi: 10.26452/ijrps.v11iSPL3.2963]
9. Daniel WW, Cross CL. Determination of sample size for estimating proportions. In: Biostatistics: A Foundation for Analysis in the Health Sciences, 11th Edition. Hoboken, NJ: Wiley (John Wiley & Sons); Dec 2018.
10. Kwon G, Serra M. Pericoronitis. In: StatPearls Internet. Treasure Island (FL): StatPearls Publishing; Nov 21, 2022.
11. Galvão EL, da Silveira EM, de Oliveira ES, da Cruz TMM, Flecha OD, Falci SGM, et al. Association between mandibular third molar position and the occurrence of pericoronitis: a systematic review and meta-analysis. Arch Oral Biol 2019 Nov;107:104486. [doi: 10.1016/j.archoralbio.2019.104486] [Medline: 31374491]
12. Sreesha S, Ummar M, Sooraj S, Aslam S, Roshni A, Jabir K. Postoperative pain, edema and trismus following third molar surgery - a comparitive study between submucosal and intravenous dexamethasone. J Family Med Prim Care 2020 May;9(5):2454-2459 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_188_20] [Medline: 32754519]
13. Singh P, Nath P, Bindra S, Rao S, Reddy KR. The predictivity of mandibular third molar position as a risk indicator for pericoronitis: a prospective study. Natl J Maxillofac Surg 2018;9(2):215. [doi: 10.4103/njms.njms_13_17]

XSL•FO
RenderX

14. Erçal P. Prevalence of impacted teeth and related pathologies: a retrospective radiographic study. ED 2023 Dec 19;2(3):108-113. [doi: 10.5152/essentdent.2023.23027]
15. Santos JF, Santos LCR, da Silveira EM, Magesty RA, Flecha OD, Falci SGM, et al. Does the third molar position influence periodontal status and overall condition of patients with acute pericoronitis? A cross-sectional study. Oral Maxillofac Surg 2020 Dec;24(4):447-453. [doi: 10.1007/s10006-020-00871-6] [Medline: 32617697]
16. Nguyen BT, Nguyen-Le C, Nguyen BT, Le SH. Risk factors associated with the severity of pericoronitis of mandibular third molar. Clin Oral Investig 2024 May 11;28(6):307. [doi: 10.1007/s00784-024-05714-2] [Medline: 38733524]

Edited by J Sarvestan; submitted 15.Feb.2025; peer-reviewed by S Kumar, C Yang; comments to author 23.Jul.2025; revised version received 07.Sep.2025; accepted 13.Oct.2025; published 21.Jan.2026.

Please cite as:
Dangore- Khasbage S, Teredesai T
Association Between Mandibular Third Molar Position and Recurrent Pericoronitis: Protocol for a Cross-Sectional Study
JMIR Res Protoc 2026;15:e72682
URL: https://www.researchprotocols.org/2026/1/e72682
doi:10.2196/72682
PMID:41563821

©Suwarna Dangore- Khasbage, Tanaya Teredesai. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Development of a Mobile App (MyLepto App) to Improve Knowledge, Attitude, and Practice Regarding Leptospirosis Among Wet Market Workers in Selangor, Malaysia: Protocol for a Quasi-Experimental Study

Mas Norehan Merican Aljunid Merican[1], MBBS, MPH; Zaleha Md Isa[1], BSc, PhD; Rozita Hod[1], MD, MPH, PhD; Roszita Ibrahim[1], BSc, MSc, PhD; Zamtira Seman[2], MSc; Rusdi Abd Rashid[3], MBBS, MPM

[1]Department of Public Health Medicine, Faculty of Medicine, Universiti Kebangsaan Malaysia, Jalan Yaacob Latif, Bandar Tun Razak, Cheras, Kuala Lumpur, Malaysia

[2]Sector for Biostatistics & Data Repository, National Institutes of Health, Ministry of Health Malaysia, Shah Alam, Malaysia

[3]Department of Psychological Medicine, Faculty of Medicine, University Malaya, Kuala Lumpur, Malaysia

Corresponding Author:
Zaleha Md Isa, BSc, PhD
Department of Public Health Medicine
Faculty of Medicine
Universiti Kebangsaan Malaysia
placeholder
Jalan Yaacob Latif, Bandar Tun Razak, Cheras, Kuala Lumpur, 56000
Malaysia
Phone: 60 0391458796
Email: zms@hctm.ukm.edu.my

Abstract

Background: Leptospirosis is the most common zoonotic cause of mortality, with most of its burden occurring in tropical regions and low-income countries. It is endemic in Southeast and South Asian nations. Leptospirosis outbreaks occur after natural disasters. In Malaysia, the e-notification system of the Communicable Diseases Control Information System recorded 5217 leptospirosis cases in 2019 with 32 fatalities. The incidence rate was 15.61 per 100,000 people. Male individuals comprised 67% of leptospirosis cases, while people aged 25 to 55 years accounted for 45% of the cases. Information and perception are crucial in influencing positive behavior. Nonetheless, information on urban and rural people's knowledge, attitude, and practice (KAP) regarding the incidence of leptospirosis is limited.

Objective: We aimed to develop a mobile app with information on leptospirosis and measure its effectiveness in improving KAP regarding leptospirosis among wet market workers in Selangor, Malaysia.

Methods: A 3-phase study will be conducted and includes development of a mobile app containing information about leptospirosis, analysis of its acceptability, and application of the intervention. Participants will be recruited based on specific inclusion criteria by using purposive sampling. Four wet markets in Hulu Langat district, Selangor, will be selected according to a list provided by local municipal councils. The respondents from each selected wet market will be workers aged 18 years and older. Mobile app development will begin with an idea description, storyboard creation, and content approval through the nominal group technique. The mobile app content will be constructed using the Health Belief Model theory. Subsequently, the usability of the mobile app prototype will be evaluated using the validated Malay version of the System Usability Scale questionnaire for the evaluation of mobile apps. This protocol entails a 12-week intervention stage, in which the baseline assessment is regarded as a pretest evaluation and the follow-up assessment as a posttest evaluation. Participant selection will be based on the inclusion and exclusion criteria. This study will incorporate a set of validated questionnaires created by a group of leptospirosis experts. The validated questionnaire will comprise 9 sections with open-ended questions on sociodemographic data, KAP, and mobile app requirements.

Results: Mobile app development and usability testing were completed between January 2024 and March 2025. Participant recruitment is scheduled in April to May 2025 after submission of this manuscript, with the 12-week intervention and data

XSL•FO
RenderX

collection running from May to July 2025. As of manuscript submission, recruitment, data collection, and data analysis have not yet begun. Data analysis is expected to be completed by September 2025, and results are anticipated for publication in late 2025.

Conclusions: Due to the high number of reported leptospirosis cases in the Hulu Langat district, Selangor, this intervention study will be conducted there. The development of the mobile app may contribute to improving wet market workers' KAP regarding leptospirosis.

International Registered Report Identifier (IRRID): PRR1-10.2196/75809

(JMIR Res Protoc 2026;15:e75809) doi:10.2196/75809

KEYWORDS

leptospirosis; mobile health app; Health Belief Model; wet market workers; public health intervention

Introduction

Background

Leptospirosis is a neglected tropical illness with a large global health burden and is the most common cause of zoonotic disease and mortality. The main leptospirosis disease burden occurs in tropical regions and the poorest countries. Southeast Asia recorded the second-highest leptospirosis incidence after Oceania, with an estimated 1.03 million cases and 58,900 fatalities yearly. Given the dearth of current information, these statistics may have been underestimated. Leptospirosis is endemic in South and Southeast Asia (regions with humid subtropical and tropical climates), with outbreaks occurring after natural disasters [1], especially after a heavy rainy season or flooding.

Malaysia is close to the equator and experiences significant rainfall, especially during the monsoon season. Hence, Malaysia is prone to being leptospirosis-endemic. Local data indicated that Malaysian leptospirosis cases increased from 1.03 cases per 100,000 people in 2004 to 30.2 cases per 100,000 people in 2015. Furthermore, the e-notification system of the Communicable Diseases Control Information System recorded 5217 leptospirosis cases in 2019, with 32 fatalities (case fatality rate: 0.6%), while the incidence rate was 15.61 per 100,000 people (Figure 1) [2].

Figure 1. Incidence rate for brucellosis, Q fever, and leptospirosis per 100,000 population in Malaysia from 2010 to 2019 (reproduced from Ministry of Health Malaysia [2], with permission from Ministry of Health Malaysia).

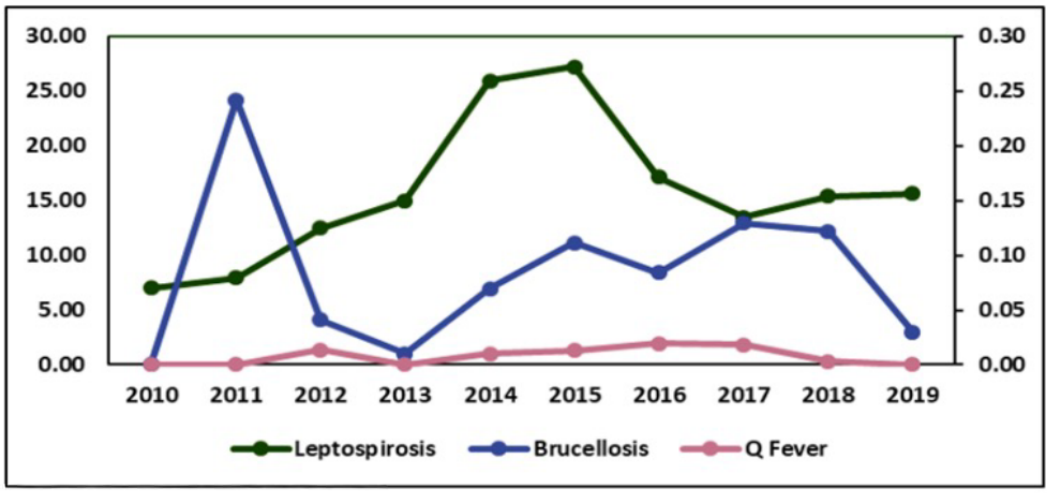

Current Leptospirosis Situation in Malaysia

In Malaysia, leptospirosis must be reported under the Prevention and Control of Infectious Diseases Act of 1988 as of December 9, 2010 [3]. The incidence of leptospirosis in Malaysia has markedly increased for more than a decade, resulting in a significant number of fatalities. In 2020, 2914 leptospirosis cases (incidence rate: 8.63 per 100,000 population) and 38 deaths were recorded (death rate: 0.1 per 100,000 population). Eight leptospirosis outbreaks were recorded in 2020, of which 50% involved households [3]. A total of 1763 leptospirosis cases (incidence rate: 4.81 per 100,000 population) and 25 deaths were recorded in 2021 (death rate: 0.07 per 100,000 population; Figure 2) [4-13]. The occurrence of some zoonoses, particularly leptospirosis, declined during the COVID-19 pandemic, which might have been linked to the Movement Control Order, school closures, and reduced recreational activities. Nevertheless, there is a great need for research and development to identify leptospirosis mitigation methods and address the increasing number of fatalities.

XSL•FO
RenderX

Figure 2. Number of leptospirosis cases and deaths in Malaysia from 2012 to 2021 (reproduced from Ministry of Health Malaysia [2], with permission from Ministry of Health Malaysia).

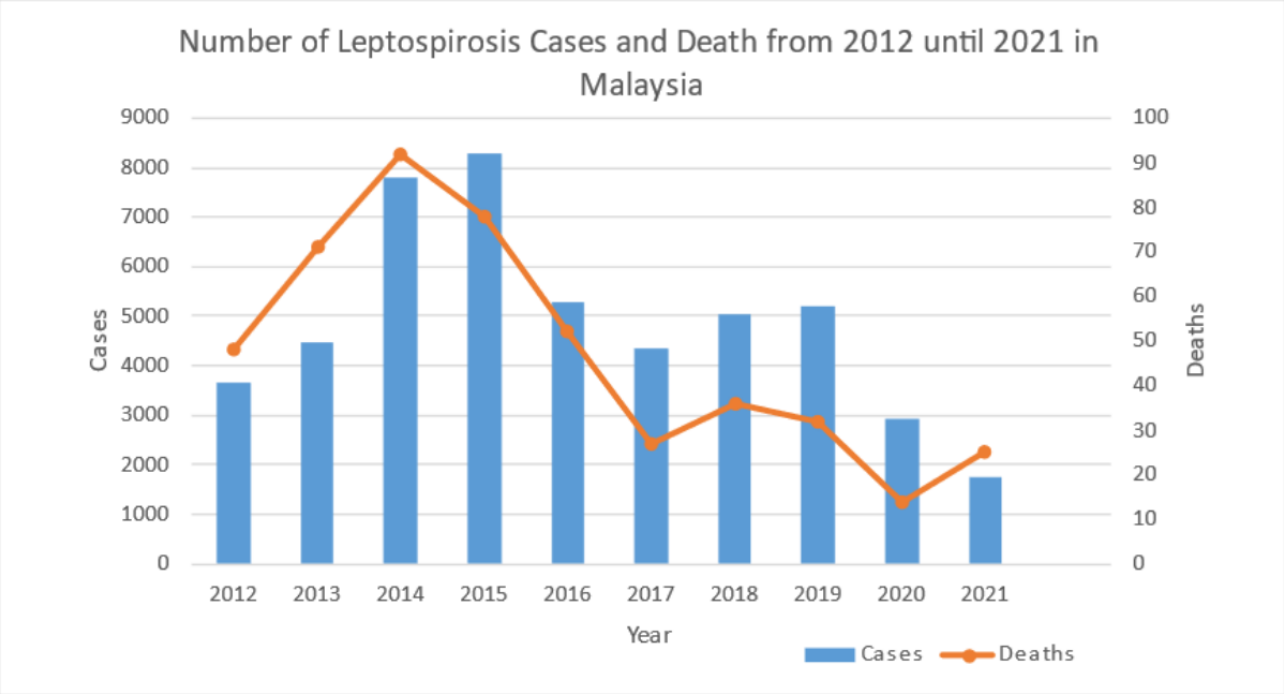

Risk Groups

Exposure to *Leptospira* in contaminated environments, such as wet markets, increases the risk of leptospirosis. An effective antibody against the disease is required to mitigate the leptospirosis infection rate [14]. Wet market workers were suggested to be at risk of contracting leptospirosis, as proven by the high seroprevalence of leptospirosis [15]. In addition, high-risk planters recorded a high seroprevalence of leptospiral antibodies (28.6%) [16], whereas the seroprevalence among wet market workers was 20% to 33.6% [15]. Thus, recognizing interventions to improve leptospirosis knowledge, attitude, and practice (KAP) would aid wet market workers in effectively addressing their risks.

Mobile App Use in Malaysia

The Malaysian Communications and Multimedia Commission Hand Phone Users Survey 2022 reported 94.8% smartphone penetration in Malaysia in 2021 [17] (Figure 3). However, the percentage of people using feature phones is decreasing, with only 7.5% expected to do so by 2021.

Figure 3. Percentage distribution of smartphone and feature phone users in Malaysia from 2017 to 2021 (reproduced from Malaysian Communications and Multimedia Commission [17], with permission from Malaysian Communications and Multimedia Commission).

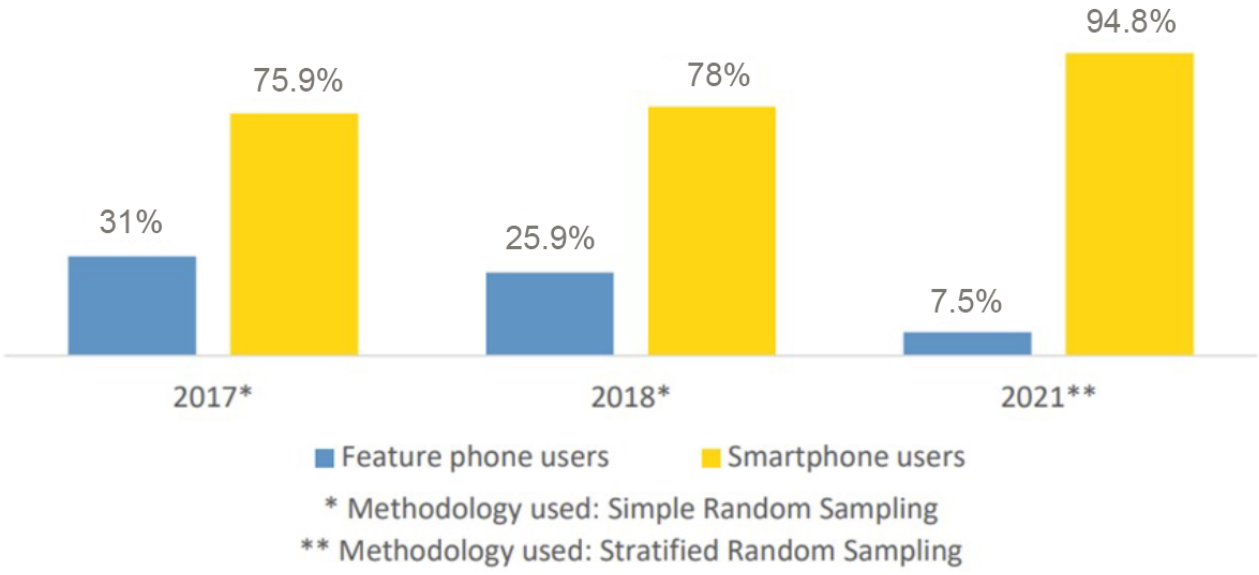

Practicality and commercialization are prevalent in digital health care and awareness, especially after the COVID-19 pandemic outbreak. The 2027 projection indicates that the Asia-Pacific digital health care market value is expected to increase by

XSL•FO
RenderX

21.2%, which would exceed RM 50 billion (US $1=RM 4.08) (Figure 4) [18].

These data reflect the need for research to improve leptospirosis KAP and prevent it among wet market workers in the Hulu Langat district, Selangor. A mobile app is expected to be beneficial for raising leptospirosis awareness among Malaysians, as it will enlighten the community about the disease and its prevention [19].

Figure 4. Projected growth of the Asia Pacific digital health market, 2021–2027. Source: Data adapted from Graphical Research [18].



Why Is This Study Important?

While the availability of scientific data has improved the understanding of leptospirosis in Malaysia, most cases in Selangor, Malaysia, have not been reported. This condition has resulted from the metropolitan community's lack of awareness of the disease [20]. Knowledge and perception are significant motivating factors of constructive behavior [21]. Unfortunately, research on leptospirosis KAP in Malaysia is subject to the lack of understanding of leptospirosis prevalence among urban and rural populations [22]. Thus, this protocol describes a study that aims to develop a mobile app to educate the community on leptospirosis and its prevention, and substitute current information dissemination methods [19].

Objectives

Workers in the agriculture and domestic animal industries are at high risk of contracting leptospirosis. Furthermore, rodent presence and environmental hygiene in wet markets have been related to leptospirosis. Hence, the primary study aim is to identify the requirements for developing a mobile app as a leptospirosis prevention intervention among wet market workers in Hulu Langat. The secondary study aim is to design and develop a mobile app with leptospirosis information to enhance KAP among wet market workers in Hulu Langat and to measure the improvements in the workers' KAP regarding leptospirosis after a 12-week intervention through pretest and posttest evaluations.

Methods

Study Design and Study Area

The study will comprise 3 phases. The first phase will develop a mobile app, the MyLepto app. The second phase will analyze the acceptability of the mobile app, while the third phase will involve the intervention with a follow-up of 12 weeks. The study will be conducted in Hulu Langat and involve preintervention and postintervention evaluations. This study will use a quasi-experimental design. Eligible respondents from 4 wet markets in Hulu Langat will be selected through simple random sampling [23].

This study will be conducted in Hulu Langat, which is well known for having high incidence rates of leptospirosis due to the large concentration of wet markets and other high-risk locations. Hulu Langat was selected given its numerous wet markets where workers regularly encounter possible leptospirosis sources, such as tainted water and rodent populations [24].

Target Population

The target population will be wet market workers in Hulu Langat, who are frequently exposed to risk factors for leptospirosis due to their jobs. High population densities, frequent contact with organic waste, and close contact with live animals are all characteristics of wet markets that increase the risk of zoonotic illnesses, such as leptospirosis [25].

XSL•FO
RenderX

Sample Population

The sample population will comprise wet market workers from 4 selected locations in Hulu Langat who fulfill the inclusion criteria.

Inclusion Criteria

This study will include Malaysian wet market workers aged 18 years or above who have worked at 1 of the 4 selected Hulu Langat wet markets for at least 6 months. The potential participants must also have access to a mobile device compatible with the MyLepto app and be willing to interact with it to participate in the study. All potential participants will be required to provide informed consent before their inclusion in the study.

Exclusion Criteria

The exclusion criteria are people with cognitive or physical disabilities that would make it difficult for them to use a mobile app or complete the study. People who have participated in comparable leptospirosis-related intervention programs within the last 6 months will be excluded to reduce bias. Furthermore, people who do not frequent the 4 wet markets, such as temporary or part-time workers who work less than 2 days per week, will also be excluded. Wet market workers will also be excluded from the study if they refuse to participate or withdraw their consent at any time.

Sample Size Calculation and Sampling Technique

An effect size of 0.26 was derived based on the means (SDs) before and after the intervention: 75.2 (13.29) and 78.65 (13.19), respectively [15]. A calculated sample size of 118 was obtained using 2 dependent means (matched pairs) at a 5% level of significance and 80% power. With the assumption of a 10% dropout rate, the final sample size needed for the study is 130. The figure was obtained using G*Power (version 3.1.9.4).

Participants will be selected using stratified random sampling. Four wet markets in Hulu Langat will be purposively selected based on high-risk profiles and logistical feasibility. Within each market, workers will be stratified according to job roles (vendors, cleaners, and handlers). A proportional random sample will be drawn from each stratum to reduce selection bias and enhance representativeness. Local councils will assist in verifying employment records.

Intervention and Timeline

Textbox 1 [26,27] describes how the 12-week intervention will be conducted.

Textbox 1. Activities and assessments.

Timeline

- Wk 0 (baseline): administration of pretest knowledge, attitude, and practice questionnaire [26] and evaluation of baseline System Usability Scale (SUS) [27]
- Wk 1-12 (intervention): weekly issuance of app-based educational modules, quizzes, and reminder announcements
- Wk 6 (midpoint): interim SUS evaluation of usability
- Wk 12 (postintervention evaluation): administration of the posttest knowledge, attitude, and practice questionnaire, final evaluation of SUS, and the field observation checklist for preventive practices

Phase 1: Mobile App Development

The mobile app development involves 7 main steps. Step 1 involves the idea description, followed by storyboard creation (step 2). In step 3, the content is to be approved by the nominal group technique (NGT), followed by finalizing the content in the Android package file format. The last 2 steps are publication in the Google Play Store and copyright acquisition from the Malaysian Intellectual Property Corporation. Figure 5 depicts the general design process of the mobile app.

The NGT will be used in the development of the mobile app content. Wet market workers will act as the target users to identify vital communication strategies and health information, including the elements for a user-friendly mobile app. In addition, public health specialists, family medicine specialists, and health inspectors will also be included as the target users.

The NGT is a structured meeting that aims to create a systematic process for collecting qualitative data from the target groups most intimately associated with the current issue [28]. In 2006, the Centers for Disease Control and Prevention stated that a small group discussion would aid in decision-making. The participants will be instructed to answer questions presented by a moderator or a researcher to obtain data for the NGT. Next, the participants will be instructed to rank the recommendations or thoughts arising from each group member [29]. This method was created by Delbecq and Van de Ven and comprises 4 important phases: silent generation, round robin, clarification, and voting (rating or ranking) [30].

XSL•FO
RenderX

Figure 5. Development process of the MyLepto mobile application. MOH: Ministry of Health.

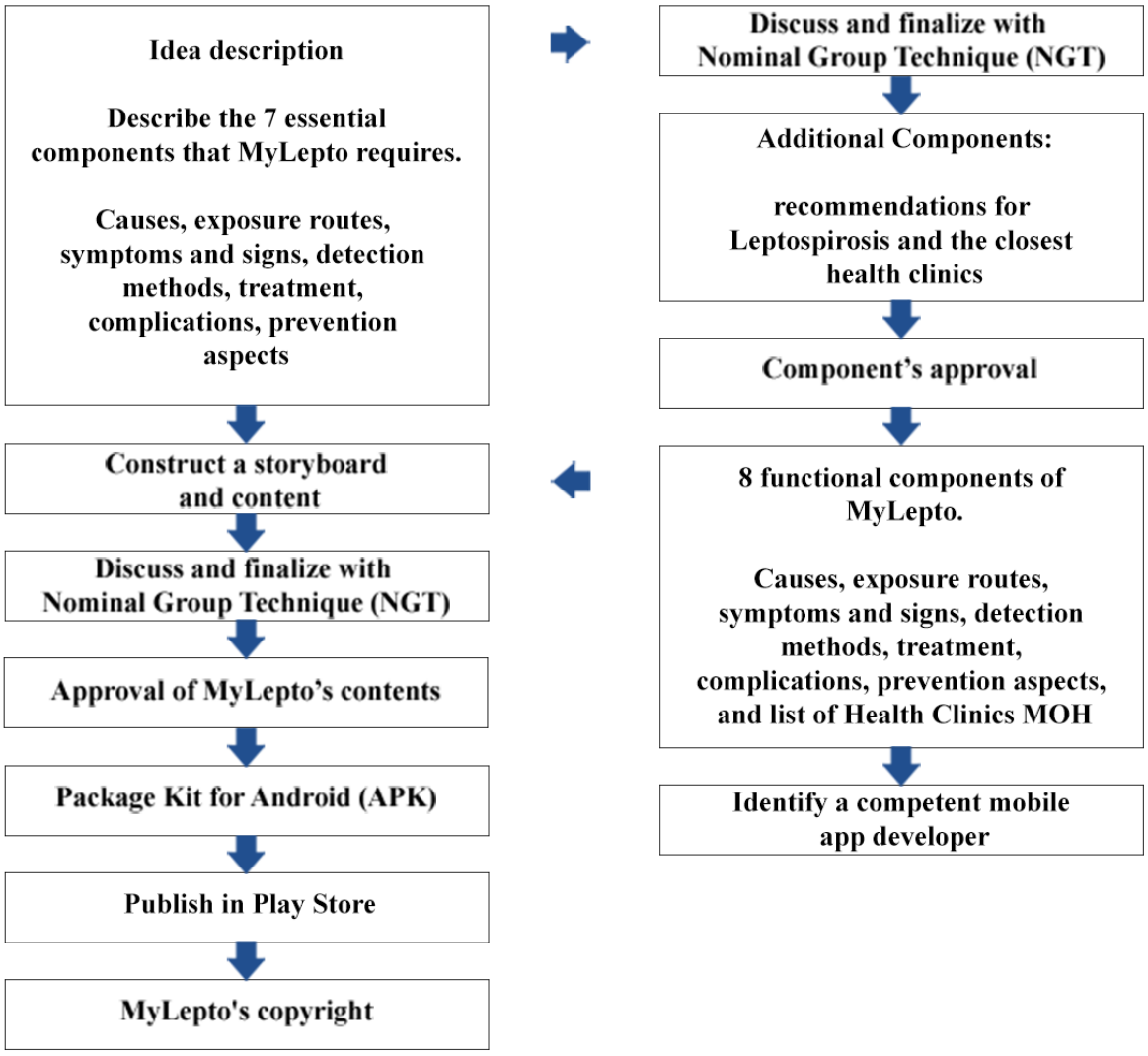

Content Development

The mobile app content will be developed using the Health Belief Model (HBM) theory [31]. The mobile app will undergo several content development stages (Figure 5). The first step reviews all available studies on leptospirosis, including the signs, symptoms, transmission methods, related complications, prevention, and control measures. This step will also gather information on the mobile app features for health education and intervention. The next step in identifying the relevant content for the target users will be based on the appropriate information synthesized. For example, users may find it difficult to express their thoughts if they are unfamiliar with the mobile app features. Hence, researchers will determine the features that users anticipate in a mobile app based on existing data.

The mobile app incorporates HBM constructs to encourage preventive behavior. Interactive modules and animated visuals on leptospirosis risks and complications will address perceived susceptibility and severity. Practical prevention tips will highlight the benefits of simple hygiene practices, while an interactive frequently asked questions section will address common barriers. Cues to action will be provided via weekly notifications and outbreak alerts. Finally, self-efficacy will be supported through step-by-step guidance and real-life examples.

Phase 2: Mobile App Acceptability Analysis

The usability of the mobile app prototype will be evaluated in Malay (Skala Kebolehgunaan Aplikasi Mudah Alih) using the validated Malay version of the System Usability Scale (SUS) questionnaire for the evaluation of mobile apps. The SUS questionnaire was originally designed in English and comprises 10 items, and will be translated into Malay using forward-backward translation. Two translators (1 linguistic expert and 1 field expert) will translate the questionnaire from English to Malay, while another 2 experts (1 linguistic expert and 1 field expert) will back-translate the questionnaire from Malay to English. Finally, the researchers and translators will discuss to finalize the Malay version of the questionnaire. The Malay version will be validated to test the mobile app usability, as described by Mohamad Marzuki et al [26]. The response score will be determined using a 5-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The total score (range 0-100) will be determined by the sum of all item scores and multiplied by 25. The original author suggested a standard usability score of 68 to signify the good usability of an app [32].

XSL•FO
RenderX

The data used in constructing the mobile app were obtained from trusted sources. Initially, information on the introduction, prevention, causes, and treatment of leptospirosis was collected from the Ministry of Health Malaysia MyHealth Portal [33]. Subsequently, the information was reviewed by epidemiology experts, comprising public health and family medicine specialists.

Phase 3: Intervention Study

The study will be conducted in Hulu Langat and involve preintervention and postintervention evaluations. Figure 6 demonstrates that the main procedures in the 12-week intervention stage and study protocol involve a pretest evaluation characterized by baseline assessment, followed by a posttest evaluation. Participants will be recruited based on specific inclusion and exclusion criteria. The local municipal council of the workers' areas will cooperate with the research team to select and educate the participants on the research objectives during the initial recruitment stage. The researcher will organize a meeting at this stage. Participant engagement will be maintained throughout the 12-week intervention through strategies such as weekly push notifications to prompt the completion of educational modules and quizzes; the use of gamified elements (interactive quizzes with instant feedback); and group engagement through WhatsApp groups organized by the market, where weekly summaries and interactive polls will be shared to sustain motivation and peer interaction.

Figure 6. Methods of data collection during the intervention study.

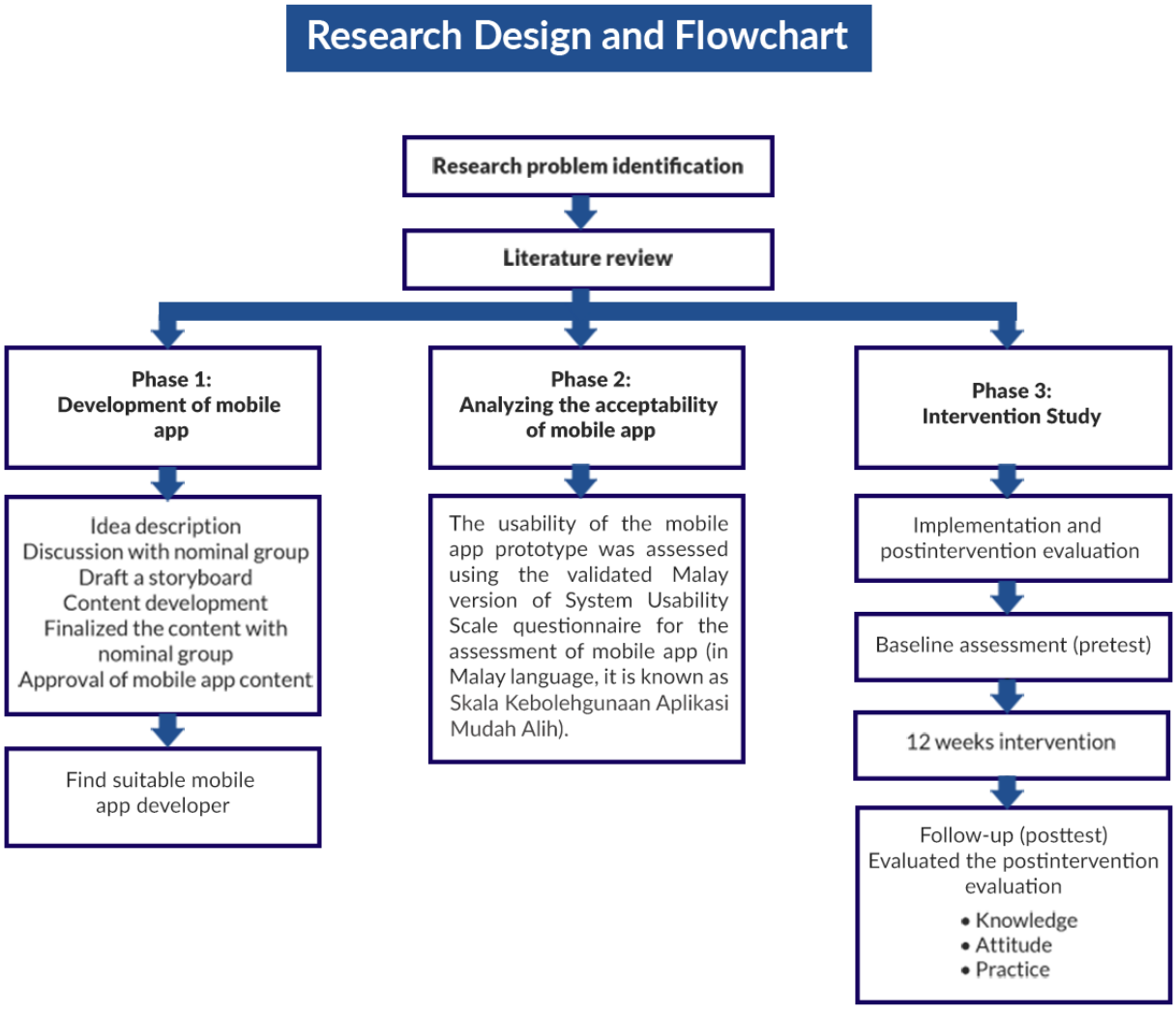

The evaluation framework includes SUS assessments at weeks 0, 6, and 12 to monitor usability, a workplace behavior checklist at week 12 to observe preventive practices, and process measures (app use data, retention rates, and postintervention interviews) to evaluate implementation effectiveness and identify areas for improvement.

Questionnaire

This study will use a validated KAP questionnaire on leptospirosis in Malaysia, which will be adopted and modified as needed [27]. The content validity of the study instrument will be ensured by cross-checking and authentication by experts in the field. Before the study begins, the questionnaire will be pretested with 30 wet market workers from Hulu Langat. The questionnaire contains 9 sections with open-ended questions on sociodemographic data and KAP [27,34].

Four of the 9 sections will focus on KAP, including sociodemographic characteristics and place of residence. The

XSL•FO
RenderX

participants will be required to disclose their age, gender, ethnicity, household income, and highest level of education.

Data Collection

The data will be collected through self-administered questionnaires. Typically, the questionnaire will be completed in 30 minutes. This study will be conducted in Hulu Langat and involve preintervention and postintervention evaluations. Measurements will be recorded from May 2025 to July 2025. App analytics will be used to monitor intervention adherence.

The recruitment and screening will be based on specific inclusion and exclusion criteria. The duration of the intervention phase will be 12 weeks, and the follow-up phase will be a posttest phase (Figure 6). The researcher will conduct a meeting with the participants' leader. During this phase, the leader will cooperate with the research team to select and educate the participants on the research objectives during the initial recruitment stage. The participants will receive a flyer describing the study details and the respondents' eligibility criteria. Subsequently, the participants will be instructed to complete a validated, self-administered preintervention questionnaire.

Data Analysis

All data collected will be entered and analyzed using SPSS (version 24.0; IBM Corp). Means and SDs will be presented for numerical variables, while frequencies and percentages will be presented for categorical data. The adjustment in the outcome measurement between the initial and intervention stages will be compared using a paired 2-tailed *t* test to assess the effectiveness of the mobile app. Usability across time points will be assessed using repeated measures ANOVA. The relationships between sociodemographic variables and KAP improvements will be examined using logistic regression.

Participants and Public Involvement

Wet market workers and an expert group were actively involved in the development process from the beginning to the final stages to ensure that the intervention was applicable, practical, and grounded in behavioral change theories, particularly the HBM. A cross-sectional study of wet market workers was conducted to identify the major gaps in their KAP regarding leptospirosis. Their responses influenced the development of the content of the mobile app. A consensus on the elements and features of the mobile app emerged after an expert group engaged in an NGT session. This process ensured that the mobile app design incorporated end-user requirements and expert-driven insights. During the development stage, wet market workers participated in usability testing sessions to evaluate and enhance the mobile app. During the intervention phase, the wet market workers will assess how effectively the mobile app improves their KAP regarding leptospirosis.

Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki and received ethics approval from the Universiti Kebangsaan Malaysia Research Ethics Committee (UKM PPI/111/8/JEP-2023-069). As the study did not involve data, patients, or facilities under the jurisdiction of the Ministry of Health, approval from the Ministry of Health Medical Research and Ethics Committee was not required. However, the study is registered with the National Medical Research Register (NMRR-ID-23-01054-DVX). Before participation, written informed consent will be obtained from the participants. The data will undergo anonymization and be stored securely. The participants will receive a small token voucher for participation. This study will not use identifiable participant images.

Results

The development of the mobile app, which incorporates HBM-based content and was validated through the NGT, was completed in January 2024. The study preparation timeline included refinement of the app content and navigation based on usability testing from February to March 2025, followed by a 3-day training workshop for research facilitators, and stakeholder meetings in April 2025. Participant recruitment will be conducted from April to May 2025 through flyers, market briefings, and eligibility screening. The 12-week intervention, which includes baseline and posttest KAP assessments, is scheduled from May to July 2025. Data analysis is expected to be completed by September 2025, and the study results are anticipated to be published by the end of 2025.

Discussion

Anticipated Findings

This study aims to develop and evaluate the MyLepto app, which is a mobile health app aimed at improving the leptospirosis-related KAP among Selangor wet market workers. This intervention was guided by the HBM and is expected to increase leptospirosis awareness, improve preventive actions, and reduce occupational risk. The assessment of usability and effectiveness will yield evidence on the role of mobile health interventions in dealing with zoonotic diseases in populations at high risk.

Comparison With Prior Work

Previous studies have demonstrated the potential of mobile apps in improving health outcomes across chronic disease management and infectious disease prevention. To the best of our knowledge, no mobile intervention has specifically targeted the prevention of leptospirosis among wet market workers to date. Thus, this study bridges an important gap by providing context-specific health education to a vulnerable occupational group. Furthermore, integrating the HBM will align this protocol with established approaches in behavioral science and support its potential to lead to significant change.

Strengths and Limitations

The strength of this study is its focus on a high-risk population (wet market workers), which ensures that the intervention is relevant and contextually specific. Furthermore, the 3-phase development process enhances the scientific rigor of the app design. Moreover, the inclusion of the validated KAP and SUS questionnaires will enhance the reliability of outcome measurements.

XSL•FO
RenderX

Nevertheless, this study has several limitations. First, the intervention duration (12 weeks) may not capture long-term behavioral alterations, which would require additional follow-up studies. Second, the participants' differing digital literacy and familiarity with mobile technology may influence their app engagement and, consequently, app effectiveness. Finally, the quasi-experimental design lacks a control group, which may limit causal inferences regarding the effects of the intervention.

Future Directions and Dissemination

Future research should involve a longer-term follow-up to evaluate continual behavioral alteration and examine the integration of gamification or social features to improve engagement. Furthermore, the public health influence of the intervention may be broadened by expanding it to other high-risk occupational groups, such as agricultural workers or sanitation staff. The plan for dissemination includes sharing the results with local health authorities, publishing the results in peer-reviewed journals, and presenting the outcomes at national and international conferences. If the MyLepto app is effective, it may act as a scalable model for mobile health interventions that target neglected tropical diseases in Malaysia and other countries.

Acknowledgments

The authors express their sincere thanks to the Ampang Jaya Municipal Council for the permission and cooperation granted during this study. They also thank the participants and the people directly and indirectly involved in this study. Finally, they thank the Department of Public Health Medicine, Faculty of Medicine, Universiti Kebangsaan Malaysia, and the Ministry of Health Malaysia for their approval and support. This study received the Fundamental Research Grant (grant JEP-2023-069) from the Faculty of Medicine, National University of Malaysia, in January 2024. The authors declare that generative artificial intelligence tools were not used in the preparation of this manuscript.

Authors' Contributions

MNMAM, RH, RI, and ZMI conceived and designed the study. MNMAM, RH, RI, and ZMI conducted the survey. RAR and ZMI participated in data analysis. The results were interpreted with assistance from RAR and ZMI. MNMAM authored the paper. The final manuscript was read and approved by all authors.

Conflicts of Interest

None declared.

References

1. Chadsuthi S, Chalvet-Monfray K, Geawduanglek S, Wongnak P, Cappelle J. Spatial-temporal patterns and risk factors for human leptospirosis in Thailand, 2012-2018. Sci Rep 2022 Mar 24;12(1):5066-5087 [FREE Full text] [doi: 10.1038/s41598-022-09079-y] [Medline: 35332199]
2. Annual report of the Ministry of Health. Ministry of Health Malaysia. 2019. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/LAPORAN%20TAHUNAN%20KKM%202019/mobile/index.html#p=1 [accessed 2025-09-23]
3. Guidelines for the diagnosis, management, prevention and control of leptospirosis in Malaysia. Ministry of Health Malaysia. 2011. URL: https://www.moh.gov.my/moh/resources/auto%20download%20images/589d71cb177d8.pdf [accessed 2025-09-23]
4. Annual report 2012. Ministry of Health Malaysia. 2012. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/MOH_2012_compressed.pdf [accessed 2025-09-23]
5. Annual report 2013. Ministry of Health Malaysia. 2013. URL: https://www.moh.gov.my/moh/images/gallery/publications/Laporan%20Tahunan%202013.pdf [accessed 2025-09-23]
6. Annual report 2014. Ministry of Health Malaysia. 2014. URL: https://myhdw.moh.gov.my/public/documents/20186/150089/Annual+Report+2014.pdf/ea08ccc8-76c8-4602-8c5e-a4b96f88ad11?version=1.0&download=true [accessed 2025-09-23]
7. Annual report 2015. Ministry of Health Malaysia. 2015. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/Annual_Report_MoH_2015_compressed.pdf [accessed 2025-09-23]
8. Annual report 2016. Ministry of Health Malaysia. 2016. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/Annual_Report_MoH_2016_compressed.pdf [accessed 2025-09-23]
9. Annual report 2017. Ministry of Health Malaysia. 2017. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/Annual_Report_MoH_2017_compressed.pdf [accessed 2025-09-23]
10. Annual report 2018. Ministry of Health Malaysia. 2018. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/ANNUAL_REPORT_2018_compressed.pdf [accessed 2025-05-29]
11. Annual report 2019. Ministry of Health Malaysia. 2019. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/LAPORAN_TAHUNAN_KKM_2019.pdf [accessed 2025-09-23]
12. Annual report 2020. Ministry of Health Malaysia. URL: https://www.moh.gov.my/moh/modules_resources/bookshelf/MOH_Annual_Report_2020_min/index.html [accessed 2025-05-29]

XSL•FO
RenderX

13. Annual report 2021. Ministry of Health Malaysia. URL: https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/Annual_Report_MoH_2021-compressed.pdf [accessed 2025-09-23]
14. Zulaikhah ST, Khalimurrosyid A, Jalu M, Jalu M, Maulana F. Risk factors of leptospirosis in Semarang, Central Java Indonesia: a case control study. Int Med J 2020;25(3):1255-1266 [FREE Full text]
15. Ab Rahman MH, Hairon SM, Hamat RA, Jamaluddin TZ, Shafei MN, Idris N, et al. Leptospirosis health intervention module effect on knowledge, attitude, belief, and practice among wet market workers in northeastern Malaysia: an intervention study. Int J Environ Res Public Health 2018 Jul 03;15(7):1396 [FREE Full text] [doi: 10.3390/ijerph15071396] [Medline: 29970799]
16. Ridzuan JM, Aziah BD, Zahiruddin WM. Study on seroprevalence and leptospiral antibody distribution among high-risk planters in Malaysia. Osong Public Health Res Perspect 2016 Jun;7(3):168-171 [FREE Full text] [doi: 10.1016/j.phrp.2016.04.006] [Medline: 27413647]
17. Hand phone users survey 2021 (HPUS). Malaysian Communications and Multimedia Commission. URL: https://web.archive.org/web/20230929072855/https://www.mcmc.gov.my/en/resources/statistics/hand-phone-surveys [accessed 2025-09-23]
18. Hand phone users survey 2017. Malaysian Communications and Multimedia Commission. 2017. URL: https://web.archive.org/web/20231026211034/https://www.mcmc.gov.my/skmmgovmy/media/General/pdf/HPUS2017.pdf [accessed 2025-12-23]
19. Ghazaei C. Pathogenic leptospira: Advances in understanding the molecular pathogenesis and virulence. Open Vet J 2018;8(1):13-24 [FREE Full text] [doi: 10.4314/ovj.v8i1.4] [Medline: 29445617]
20. Abdullah NM, Mohammad WM, Shafei MN, Sukeri S, Idris Z, Arifin WN, et al. Leptospirosis and its prevention: knowledge, attitude and practice of urban community in Selangor, Malaysia. BMC Public Health 2019 May 22;19(1):628 [FREE Full text] [doi: 10.1186/s12889-019-6981-0] [Medline: 31117995]
21. Larson RW, Rusk N. Intrinsic motivation and positive development. In: Lerner RM, Lerner JV, Benson JB, editors. Advances in Child Development and Behavior. New York, NY: Academic Press; 2011:89-130.
22. Nozmi N, Samsudin S, Sukeri S, Shafei MN, Wan Mohd WM, Idris Z, et al. Low levels of knowledge, attitudes and preventive practices on leptospirosis among a rural community in Hulu Langat district, Selangor, Malaysia. Int J Environ Res Public Health 2018 Apr 06;15(4):693 [FREE Full text] [doi: 10.3390/ijerph15040693] [Medline: 29642390]
23. Samsudin S, Saudi SN, Masri NS, Ithnin NR, Hamat RA, Wan Mohd ZW, et al. Awareness, knowledge, attitude and preventive practice of leptospirosis among healthy Malaysian and non-Malaysian wet market workers in selected urban areas in Selangor, Malaysia. Int J Environ Res Public Health 2020 Feb 19;17(4):1346 [FREE Full text] [doi: 10.3390/ijerph17041346] [Medline: 32093098]
24. Malaysia health systems research (vol 1). Ministry of Health Malaysia. URL: https://www.moh.gov.my/moh/resources/Vol_1_MHSR_Contextual_Analysis_2016.pdf [accessed 2025-09-23]
25. World health statistics 2015. World Health Organization. 2015. URL: https://www.who.int/publications/i/item/9789240694439 [accessed 2025-09-23]
26. Mohamad Marzuki MF, Yaacob NA, Yaacob NM. Translation, cross-cultural adaptation, and validation of the Malay version of the system usability scale questionnaire for the assessment of mobile apps. JMIR Hum Factors 2018 May 14;5(2):e10308 [FREE Full text] [doi: 10.2196/10308] [Medline: 29759955]
27. Zahiruddin WM, Arifin WN, Mohd-Nazri S, Sukeri S, Zawaha I, Bakar RA, et al. Development and validation of a new knowledge, attitude, belief and practice questionnaire on leptospirosis in Malaysia. BMC Public Health 2018 Mar 07;18(1):331 [FREE Full text] [doi: 10.1186/s12889-018-5234-y] [Medline: 29514645]
28. Rashid NS, Chen XW, Mohamad Marzuki MF, Takshe AA, Okasha A, Maarof F, et al. Development and usability assessment of a mobile app (Demensia KITA) to support dementia caregivers in Malaysia: a study protocol. Int J Environ Res Public Health 2022 Sep 20;19(19):11880 [FREE Full text] [doi: 10.3390/ijerph191911880] [Medline: 36231181]
29. Gaining consensus among stakeholders through the Nominal Group Technique. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/healthy-youth/php/program-evaluation/pdf/brief7.pdf [accessed 2025-09-23]
30. McMillan SS, King M, Tully MP. How to use the nominal group and Delphi techniques. Int J Clin Pharm 2016 Jun;38(3):655-662 [FREE Full text] [doi: 10.1007/s11096-016-0257-x] [Medline: 26846316]
31. Rosenstock IM. Historical origins of the health belief model. Health Educ Monogr 1974 Dec 01;2(4):328-335. [doi: 10.1177/109019817400200403]
32. Sauro J. Measuring usability with the System Usability Scale (SUS). Measuring U. 2011. URL: https://measuringu.com/sus/ [accessed 2025-09-23]
33. Myhealth. Ministry of Health Malaysia. URL: http://www.myhealth.gov.my/ [accessed 2025-09-23]
34. Anh NQ, Hung LX, Thuy HN, Tuy TQ, Caruana SR, Biggs B, et al. KAP surveys and malaria control in Vietnam: findings and cautions about community research. Southeast Asian J Trop Med Public Health 2005 May;36(3):572-577 [FREE Full text] [Medline: 16124419]

XSL•FO
RenderX

Abbreviations

HBM: Health Belief Model
KAP: knowledge, attitude, and practice
NGT: nominal group technique
SUS: System Usability Scale

Edited by A Schwartz; submitted 14.Apr.2025; peer-reviewed by A Adeoye, S Ismail; comments to author 07.May.2025; revised version received 03.Jul.2025; accepted 09.Oct.2025; published 23.Jan.2026.

Please cite as:
Aljunid Merican MNM, Md Isa Z, Hod R, Ibrahim R, Seman Z, Abd Rashid R
Development of a Mobile App (MyLepto App) to Improve Knowledge, Attitude, and Practice Regarding Leptospirosis Among Wet Market Workers in Selangor, Malaysia: Protocol for a Quasi-Experimental Study
JMIR Res Protoc 2026;15:e75809
URL: https://www.researchprotocols.org/2026/1/e75809
doi:10.2196/75809
PMID:41575778

©Mas Norehan Merican Aljunid Merican, Zaleha Md Isa, Rozita Hod, Roszita Ibrahim, Zamtira Seman, Rusdi Abd Rashid. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Development and Feasibility Assessment of an Intrinsic Capacity Program in Primary Care: Protocol for an Implementation Science Approach

Qing Wang[1], MPP; Mimaika Luluina Ginting[1], MD, MPH; Ezra Ho[1], MA; Siew Fong Goh[1], MGER; Jonathan Gao[1], PhD; Woan Shin Tan[1], PhD; Yew Yoong Ding[1,2,3], MBBS, MPH, PhD; Wei Liang David Ng[4], MBBS, MMed; Jonathan Choon Aik Ng[4], MBBS, MMed; Sing Cheer Kwek[5], MBBS, MMed; Richard Jor Yeong Hui[5], MBBS, MMed; Zhen Sinead Wang[6,7], MBBS, MMed, MCI; Chirk Jenn Ng[6,7], MBBS, MMed, PhD; Grace Sum[1], PhD; The IMPACTFrail Group[8]

[1]Geriatric Education & Research Institute, Singapore, Singapore

[2]Department of Geriatric Medicine, Tan Tock Seng Hospital, Singapore, Singapore

[3]Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[4]National Healthcare Group Polyclinics, Singapore, Singapore

[5]National University Polyclinics, Singapore, Singapore

[6]SingHealth Polyclinics, Singapore, Singapore

[7]Duke-NUS Medical School, Singapore, Singapore

[8]See Acknowledgments

Corresponding Author:
Grace Sum, PhD
Geriatric Education & Research Institute
2 Yishun Central 2
Singapore, 768024
Singapore
Phone: 65 68078020
Email: sum.grace.ce@geri.com.sg

Abstract

Background: The World Health Organization (WHO) public health framework for healthy aging advocates for action on the trajectories of intrinsic capacity (IC) across a person's life course to optimize functional ability. While the WHO integrated care for older people (ICOPE) framework provides guidance on a systematic care pathway on IC screening, clinical assessment to clarify IC deficits and person-centered management, its real-world implementation and evaluation remain nascent. The Intrinsic Capacity Promotion in Primary Care for the Frail (IMPACTFrail) program for mildly frail older adults in Singapore's primary care seeks to operationalize WHO ICOPE and national strategies.

Objective: The objectives of this study are (1) the co-development of IMPACTFrail's core functions and its delivery, as well as selecting, specifying and operationalizing implementation strategies to address anticipated barriers and leverage anticipated facilitators and (2) to conduct a feasibility assessment on the readiness to scale to a main study.

Methods: For the first objective, the co-development process is guided by the United Kingdom Medical Research Council's (MRC's) framework for developing and evaluating complex interventions and the Framework of Actions for Intervention Development (FAID). The identification of contextual barriers and facilitators will draw on the updated Consolidated Framework for Implementation Research (CFIR) and its Outcomes Addendum. To identify individual-level behavior change barriers, we will extend this framework using the Theoretical Domains Framework (TDF). The Expert Recommendations for Implementing Change (ERIC) taxonomy guided our selection and development of implementation strategies. The collaboration involves implementation researchers, clinic leadership, frontline health care providers, and older adults. A 12-month, single-arm feasibility study will recruit 180 older adults aged 60 years and older with mild frailty (Clinical Frailty Scale score 4-5) across 5 public primary care clinics. Feasibility criteria include implementation, acceptability, practicality, and adaptability. We will narratively triangulate findings across study components to enhance the validity and credibility of the feasibility study, including (1) process evaluation using quantitative process indicators, (2) qualitative study to elicit barriers and facilitators to feasibility, sustainability

XSL•FO
RenderX

and scalability, and to assess the attribution of the selected implementation strategies to implementation outcomes, (3) cost analysis, and (4) program description.

Results: The study was funded in September 2024. Data collection for the feasibility assessment commenced in April 2025 and will conclude by March 2026. As of manuscript submission, 98 participants have been recruited across 5 sites. Recruitment, data collection, and analysis are ongoing. Publication of results is expected in early 2027.

Conclusions: This protocol contributes to the literature by providing a detailed protocol on the co-development and feasibility testing of a complex intervention to enhance transparency, fidelity, and replicability. It disseminates knowledge on the integration of frameworks and methodologies to accelerate the translation of evidence to sustainable and scalable programs in practice.

Trial Registration: ClinicalTrials.gov NCT06753643; https://clinicaltrials.gov/study/NCT06753643

International Registered Report Identifier (IRRID): DERR1-10.2196/84257

(JMIR Res Protoc 2026;15:e84257) doi:10.2196/84257

KEYWORDS

aged; world health organization; WHO; integrated care for older people; ICOPE; intrinsic capacity; frailty; feasibility; co-development; implementation science; primary care

Introduction

Background

The World Health Organization (WHO) conceptualizes healthy aging as the "ongoing process of developing and maintaining functional ability that enables well-being in older age" [1]. Intrinsic capacity (IC) refers to an individual's combined physical and mental capacities (locomotion, vision, hearing, vitality, cognition, and psychological), while functional ability is the interaction between IC and the environment [1,2]. Recent literature reveals that IC is a predictor of adverse health outcomes, including functional decline, disability, hospitalizations, and mortality [3-6]. In response, the WHO public health framework for healthy aging advocates for action on trajectories of functional ability and IC across a person's life course, particularly subgroups with stable or decreasing capacity prior to substantial losses [7].

To support the implementation of this approach, the WHO integrated care for older people (ICOPE) framework was introduced to provide guidance on optimizing IC and functional ability through person-centered care that prioritizes personal goals and needs. This reflects a shift from episodic and curative care to a holistic approach to healthy aging [8]. It promotes a systematic care pathway including Step 1 screening for IC deficits, Step 2 on assessments to clarify IC deficits, and Steps 3-5 on the development of personalized care plans, linkage to health and social care services to manage deficits, and engaging caregivers and communities for support [2]. Furthermore, IC is closely linked to frailty, the clinically recognizable state of increased vulnerability due to diminished physiological reserves in physical, cognitive, and social functioning [9,10]. Literature suggests that IC and frailty are interrelated constructs [11]. Declines in IC have been associated with incident frailty, and IC screening may enhance early detection and risk profiling [12-14]. Evidence also suggests that addressing IC declines could be more effective than exercise and nutritional interventions in prefrailty reversal to robustness [15].

In 2023, the Singapore Ministry of Health (MOH) published the National Frailty Strategy Report, which provides guidance on applying the WHO ICOPE framework to support mildly frail older adults in reversing frailty progression and optimizing functional ability [16]. This highlights the need to translate policy into practice. However, there is a dearth in the literature on the implementation and evaluation of the ICOPE in real-world settings [17,18]. International literature on early adopters from France and China identified several barriers. These include low adherence among older adults with more severe IC deficits, those who often need management the most [19]; anticipated challenges in scaling beyond feasibility study due to intense resource use [20]; and limited reach to older adults who are less technologically savvy when telehealth is used as the primary delivery mode [21]. While these findings are insightful, they may not be generalizable to Singapore due to the difference in health care systems and local cultures. IC screening can be implemented across various settings, including community and primary care. While recent local studies have assessed the anticipated determinants and feasibility of implementing IC screening in community settings, no studies have examined its implementation in primary care [18,22,23], which is often the first point of contact for patients accessing health care services [24]. To elaborate, existing health care programs that conduct IC screening are in active aging centers and community clubs to serve community-dwelling older adults in Singapore. These community-based centers are located in close proximity to public residential housing for ease of access [18,25]. They offer a range of services such as befriending services, workshops (eg, cooking and arts and crafts), exercise programs, and chronic condition health screenings. In contrast, there are no health care programs that detect and manage IC decline at public primary care clinics and private general practices in Singapore. Moreover, these studies do not specifically target mildly frail older adults, a group with greater potential for early screening and intervention to prevent further functional decline and a higher likelihood of reverting to robustness compared to their frail counterparts [26].

Additionally, it is common for complex interventions to fall short of achieving sustainment and scalability when implemented in real-world practice settings outside research environments [27,28]. ICOPE programs for community-dwelling older adults are inherently complex, involving multiple components, diverse stakeholders, and the need for integration

XSL•FO
RenderX

with existing workflows and infrastructures. Hence, it is crucial to account for contextual factors, which are the dynamic forces for and against successful implementation across macro, meso, and individual levels [29]. In response, there is growing advocacy for systematic approaches to designing, implementing, and evaluating complex health care interventions, drawing on both complex intervention research and implementation science frameworks [27,30-32]. This study protocol outlines the co-development of an IC management program in Singapore and its feasibility assessment. The program's development and feasibility assessment are guided by complex intervention research and implementation science frameworks, using a phased approach that includes development, feasibility testing, and scaling up for evaluation. This approach is designed to enhance implementation outcomes such as reach, adoption, fidelity, and sustainability. In the long run, it aims to support integration into routine practice, promote adoption by implementers, and improve uptake and adherence among older adults [30,31]. The reporting of this protocol is guided by the adapted SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guideline (see Multimedia Appendix 1) [33,34].

The Intrinsic Capacity Promotion in Primary Care for the Frail Program

Core functions of Intrinsic Capacity Promotion in Primary Care for the Frail (IMPACTFrail) are aligned with the WHO ICOPE framework and the 3-tier approach of the National Frailty Strategy (Figure 1):

1. IC screening will be conducted for older adults aged 60 years and older who are mildly frail, as assigned by a trained health care provider as Clinical Frailty Scale (CFS) score 4 or 5. The CFS was selected by MOH as a national community screening and segmentation tool [16,35]. It is a 9-point frailty screening tool that ranges from CFS 1 (very fit) to 9 (terminally ill) [35]. The policy proposes to target the CFS 4 (very mild frailty) and 5 (mild frailty) groups, as they are anticipated to benefit the most and to support optimal resource allocation. CFS 6-7 individuals (moderate to severe frailty) may benefit from directly undergoing comprehensive geriatric assessment (CGA) [16,36,37], and the CFS 8-9 group requires distinctly different care such as end-of-life support instead of community-based health services.
2. Follow-up clinical assessments to clarify IC deficits. This refers to postscreening assessment for IC deficits using clinical tools and scales by health care providers, such as the short physical performance battery for mobility or the Abbreviated Mental Test for cognition.
3. Establishment of care pathways to health and social care services to manage and monitor IC deficits and other identified needs. These individualized care plans will be overseen by the primary care physicians.

Figure 1. The 3-tier approach in the National Frailty Strategy Report and Intrinsic Capacity Promotion in Primary Care for the Frail (IMPACTFrail) care model.

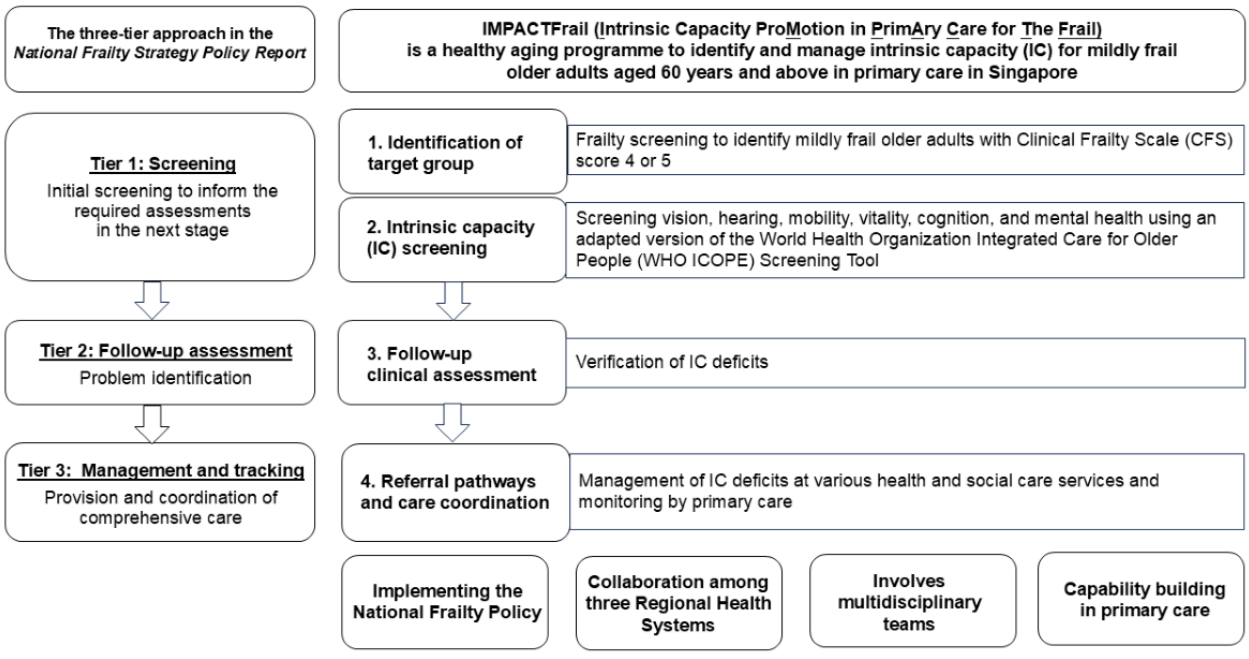

Grounded in the 3-tier approach, IMPACTFrail seeks to translate the National Frailty Strategy specifically to local primary care contexts. First, evidence on the implementation of the ICOPE framework in primary care is still limited and may not be generalizable to Singapore's contexts in terms of health care system infrastructures, policies, and cultures [17,38]. Second, IMPACTFrail's ICOPE Step 1 in primary care may complement early local efforts on IC screening via community-based programs at active aging centers and community care organizations [22,23]. A proportion of community-dwelling older adults do not participate in community-based programs, and public primary care clinics may play a strategic role in

XSL•FO
RenderX

capturing these older adults. Primary care physicians have also been identified as the preferred personnel to coordinate personalized care, and IMPACTFrail may facilitate this relationship by delivering ICOPE Step 1 in primary care at the outset [39,40]. Third, existing primary care national initiatives such as Healthier SG, which promotes a "one patient, one family doctor" model for chronic condition management [41], are complemented by IMPACTFrail's holistic focus on functional ability. IMPACTFrail is also aligned with Age Well SG, a national program on aging well in the community [42].

To facilitate sustainability and scalability, the development of IMPACTFrail involves public primary care clinics across Singapore's 3 integrated health care systems from the outset (Figure 1). Briefly, Singapore's health care system is structured into 3 integrated regional health systems. Primary care is provided by private general practitioners and 26 public clinics. While all 3 health care systems are under MOH, they have autonomy in carrying out their functions [43]. Each of the 3 systems is accountable for the health of their geographically defined populations. There is impetus to accelerate collaborative efforts among all 3 systems to deliver national initiatives cohesively [24]. Second, IMPACTFrail serves to strengthen capability building in primary care (Figure 1). Targeted training is conducted to equip multidisciplinary teams, including doctors, nurses, care coordinators, and allied health care workers (eg, physiotherapists and dieticians) with skills needed for CFS assignment, IC screening, and follow-up clinical assessments. Third, IMPACTFrail facilitates collaborations among multidisciplinary and interprofessional stakeholders. Buy-in from clinic leadership is required for integration of IMPACTFrail within each clinic's ecosystem of programs and workflows. IMPACTFrail involves nonmedical staff, family doctors, family medicine specialists, nurses, and allied health professionals (eg, dieticians and therapists). The clinics build on partnerships with specialist outpatient clinics and community care organizations to manage older adults' IC deficits.

Study Aims and Hypotheses

The overall aim of the protocol is to describe the co-development and feasibility assessment of the implementation of a novel IC detection and management program, along with its implementation strategies, for mildly frail older adults in the primary care setting in Singapore. Specific objectives include:

- Objective 1: to collaboratively develop the form of delivery of IMPACTFrail's core components, and select, specify and operationalize implementation strategies to mitigate anticipated barriers and leverage anticipated facilitators to implementation. The co-development involves implementation researchers as well as health care administrators and frontline health care providers from all three health care systems.
- Objective 2: to conduct a multisite feasibility study to assess the feasibility of implementing IMPACTFrail and its selected implementation strategies in contextualized settings. Specifically, we will conduct a process evaluation using quantitative process indicators, a qualitative study to elicit the barriers and facilitators to feasible, sustainable, and scalable implementation, a cost analysis to understand resources needed to develop and implement the program, and documentation of the program description and adaptations.

We anticipate the following:

- Hypothesis 1: we hypothesize a successful co-development of the program's core components and implementation strategies. This is due to stakeholder engagement and collaboration with clinic leadership teams and frontline implementers from the outset.
- Hypothesis 2: overall, we hypothesize a moderate to high level of feasibility. This is due to the program being supported by the selection, specifying, and development of implementation strategies to mitigate anticipated barriers and leverage anticipated facilitators. We anticipate barriers at the macro, meso, and individual levels, given that IMPACTFrail is a complex intervention involving multiple components and stakeholders. At the patient level, we also anticipate barriers related to behavioral change, such as low motivation and limited opportunity, as older adults may lack health-seeking behaviors or support to follow through on postscreening assessment and interventions. Cost analysis may reveal marginally high resource use from manpower costs for the initial capability building of health care providers. Barriers to sustainability and scalability are anticipated due to competing resource demands and priorities in primary care.

Methods

Integration of Frameworks

Both objectives are guided by integrating complex intervention research and implementation science frameworks (Table 1). First, this study is guided by the overarching United Kingdom Medical Research Council's (MRC's) framework for developing and evaluating complex interventions [30]. IMPACTFrail is a complex intervention with multiple core functions, implementation strategies, and involvement of stakeholders (decision-makers, frontline implementers, and older adults). The MRC framework divides complex intervention research into four phases: (1) development and identification of the intervention, (2) feasibility testing, (3) larger-scale evaluation, and (4) implementation [30]. Phase 1 and 2 correspond to the development of IMPACTFrail and feasibility assessment, respectively. The latter 2 phases of the MRC framework—evaluation and implementation—will be conducted in subsequent studies. Second, Framework of Actions for Intervention Development (FAID) guides Objective 1 [44]. It serves as an extension to the MRC framework's development and identification of an Intervention, and informs the key principles and actions for consideration when developing IMPACTFrail.

XSL•FO
RenderX

Table 1. Integration of frameworks for complex intervention research and implementation research.

Framework and guidelines	What it informs	Elaboration
Integration of the United Kingdom Medical Research Council (UK MRC) framework for developing and evaluating complex interventions [30] and the Framework of Actions for Intervention Development (FAID) [44]	Objectives 1 and 2	• The MRC guidance is a high-level framework that connects the development process to subsequent pilot-testing, implementation, and evaluation. • The FAID informs on the list of actions to undertake in the process of developing an intervention. It provides more detailed guidance on the MRC framework's development stage.
Updated Consolidated Framework for Implementation Research (CFIR) and its Outcomes Addendum [31]	Objectives 1 and 2	• Anchors the implementation research. • Guides the understanding of contextual determinants to attaining successful implementation outcomes.
Integration of Updated CFIR [31], Expert Recommendations for Implementing Change (ERIC) taxonomy [32], and the Proctor framework for specifying and reporting strategies [45]	Objective 1: Study component on selecting, specifying, and developing implementation strategies	• Informs the selection and specifying of implementation strategies to mitigate CFIR-identified barriers and leverage CFIR-identified facilitators.
Guidelines on conducting feasibility studies [30,46-48]	Objective 2	• Guidelines inform the components of a feasibility study and assessment of feasibility criteria.
Proctor framework for specifying and reporting strategies [45]	Objective 2: Study component on assessing fidelity of implementation strategies	• Monitor and track the fidelity of operationalizing the specified domains (actor, action, action target, dose, temporality) of each selected ERIC strategy using a comprehensive list of indicators.
Integration of Updated CFIR [31], the Theoretical domains framework (TDF) [49], and the ERIC taxonomy [32]	Objective 2: Qualitative study component	• Expand CFIR's Individual Characteristics domain by mapping determinants of Capability, Opportunity, and Motivation to TDF domains. • ERIC taxonomy informs the refinement of existing implementation strategies and selection of new strategies.

Third, our study is anchored in the updated Consolidated Framework for Implementation Research (CFIR) and its Outcomes Addendum. The updated CFIR is a determinant framework that guides the understanding of contextual factors that influence successful implementation of complex interventions, while its Outcomes Addendum guides the evaluation of implementation outcomes such as adoption, implementation, sustainability and scalability [31]. CFIR is one of the most widely used implementation science frameworks and this study applies the updated version with 48 constructs across five domains (innovation, outer setting, inner setting, implementation process, and individuals). Additionally, we will apply the Theoretical Domains Framework (TDF) to expand on CFIR's "Characteristics of Individuals" domain [49]. Specifically, the TDF expands on the Capability, Opportunity, Motivation, Behavior (COM-B) model, a behavior change framework that proposes three necessary components for any behavior to occur [49]. The Expert Recommendations for Implementing Change (ERIC) taxonomy and Proctor framework on reporting and specifying implementation strategies will also be used [32,45,50]. The application of these frameworks will be elaborated in the subsequent sections. Finally, published guidelines on conducting feasibility studies will inform the feasibility assessment [30,46-48].

Objective 1: Co-Development of IMPACTFrail and Implementation Strategies (6 Months)

Aim

We aim to collaboratively develop (1) the forms of delivery of IMPACTFrail's core components and (2) implementation strategies to support implementation. This collaboration will involve the research team, health care administrators (clinic leadership teams with decision-making authority), and frontline health care providers, including doctors, nurses, and allied health professionals.

Study Procedure

Prior to this study, the research team conducted 6 months of informal stakeholder engagement (total of 22 interviews and focus group) with the clinics' leadership, frontline health care providers, and older adults. Overall, we aimed to explore their awareness of frailty and IC, elicit anticipated barriers and facilitators of implementing or participating in the program, and understand the contextual operational and logistical constraints at each clinic. Our study builds on the information from the stakeholder engagement. To elaborate, we conducted 3 interviews with 3 older adults who were existing patients of the clinics to elicit expected barriers to participating in IMPACTFrail and to provide suggestions to encourage older people to attend IC screening. We conducted 7 interviews with 8 individuals who held mid-to-senior management positions in the clinics' leadership team and 5 interviews and 5 focus groups with family medicine doctors (n=10), nurses (n=10), allied

XSL•FO
RenderX

health professionals (n=2), care coordinators who assist doctors and nurses (n=4), and nonmedical support staff (n=6). Health care professionals were across different levels of seniority and whom we anticipated to be involved across different steps of IMPACTFrail's patient journey. Sampling was guided by using the principle of sufficiency [51,52]. Additionally, we interviewed 2 private general practitioners to gain some insights from the private sector on implementing an IC detection and management program in a primary care setting.

The individual or paired interviews were for health care administrators and older adults to encourage open sharing, particularly in light of potential power dynamics and different language and cultural backgrounds. Focus groups with multidisciplinary teams were to gather different perspectives on the feasibility of the program. Interviews with older adults were conducted in their preferred language, either English or Chinese, by multilingual research team members. Engagement sessions with health care administrators and frontline health care providers were conducted in English. The engagement aimed to explore participants' awareness of frailty and IC, elicit anticipated barriers and facilitators to implementing or participating in the program, and understand the operational and logistical context at each clinic. The study design and implementation were informed by insights through informal stakeholder engagement and guided by complex intervention research and implementation science frameworks.

Co-Development of the Core Components

The core functions of IMPACTFrail (the "what") are informed by Singapore's National Frailty Strategy, while the forms of delivery (the "who" and "how") will be collaboratively developed with stakeholders in this study. Examples of specific output from this co-development stage may include: (1) IC screening questions adapted from the original WHO ICOPE screening tool (see Multimedia Appendix 2) tailored to each clinic's clinical protocols and preferences; (2) modifications to electronic medical record systems to capture new data required by IMPACTFrail; (3) clear patient journey pathways; (4) identification of relevant clinic staff (doctors, nurses, allied health professionals, clinic coordinators, and support staff) and delegation of roles; and (5) capability building plans to prepare for the upcoming feasibility study.

Co-Development of Implementation Strategies

We will select, specify, and develop implementation strategies to mitigate anticipated barriers and leverage anticipated facilitators to implementation. Implementation strategies are defined as "methods or techniques used to enhance the adoption, implementation, and sustainability of a clinical program or practice" [45]. As we will conduct a smaller scale feasibility study in the next stage, each clinic will only implement a tailored set of 2-5 implementation strategies. These implementation strategies could be from the list of the 73 ERIC strategies or created inductively [32]. Each clinic could select and develop a different set of implementation strategies to align with their specific contexts.

The anticipated barriers and facilitators identified during the informal stakeholder engagement will be mapped to the CFIR determinant constructs. Subsequently, we will use the CFIR-ERIC matching tool to match CFIR-identified determinants to ERIC strategies [53,54]. The research team, polyclinic leadership, and frontline health care providers will collaboratively review the identified barriers and facilitators, the output from the CFIR-ERIC matching tool, and local contextualized considerations (eg, clinic operations, manpower) to prioritize and select implementation strategies. Strategy selection will be guided by 5 literature-informed criteria: relevance, opportunity for improvement, feasibility, level of difficulty, and validity [55]. The selected implementation strategies will be specified using Proctor framework for reporting and specifying implementation strategies across 6 domains, including actor, action target, action, temporality, dose, and implementation outcomes [45], and subsequently developed for operationalization in the feasibility study [46,56].

Objective 2: Multisite Feasibility Study (12 Months)

Overall Aim

The feasibility study is defined as a small-scale assessment aimed at determining whether IMPACTFrail can be implemented at this level, evaluating its potential for scalability, and identifying how it might be adapted for broader implementation [46,56]. In other words, we seek to evaluate the readiness of implementers delivering IMPACTFrail in real-world settings at a larger scale and the preparedness of older persons to participate in the program [30,47,48]. This protocol precedes a subsequent hybrid effectiveness-implementation study that will be conducted at a larger scale with more sites and older adult participants [57,58]. The study was registered on ClinicalTrials.gov on December 22, 2024 (Identifier: NCT06753643).

Study Design

This will be a single-arm feasibility study conducted at five public clinics across all 3 regional health systems in Singapore. Specifically, 2 clinics were selected from each of 2 regional health systems, and 1 from the third System. Each clinic has contextualized barriers and facilitators to implementation, clinic workflows, and infrastructure. Thus, each clinic is conceptualized as a single site for feasibility testing. This study does not include a comparator arm or a pre-post design, as it does not aim to assess effectiveness. Instead, it focuses on a small-scale feasibility assessment to inform the design and implementation of a subsequent, statistically powered hybrid effectiveness-implementation. The phased approach increases the likelihood of successful real-world implementation by enabling early refinement and adaptation in practice settings.

Eligibility Criteria and Sample Size

Eligibility criteria include (1) registered patient of the clinic, (2) aged 60 years and older (3 of 5 clinics) or aged 65 years and older (2 of 5 clinics), (3) mildly frail, assigned by a trained health care provider as CFS score 4 or 5, and (4) able to provide informed consent independently or through a legally authorized representative. Exclusion criteria include (1) administered comprehensive geriatric assessment in the past 12 months or (2) at the time of this writing, being under the care of a geriatrician or patient of a geriatric specialist outpatient clinic.

XSL•FO
RenderX

Older adults who meet these eligibility criteria will be given information on the IMPACTFrail health care program and the research study verbally and via a hardcopy brochure by a paraprofessional care coordinator or nurse. A trained research assistant will obtain written informed consent from each older adult individually and address any queries from potential participants. Consent-taking to be recruited as a participant of the IMPACTFrail study will be conducted prior to intrinsic capacity screening. Recruitment will be conducted over approximately 9 months, continuing until each clinic reaches its target of 36 participants. All older adults will only need to go through this process once. Additionally, a subset of these older adult study participants of IMPACTFrail who participate in in-depth interviews of the qualitative study will undergo a separate informed consent process before their interviews. This informed consent-taking will be facilitated by trained researchers.

This set of consistent and unified inclusion and exclusion criteria will facilitate meaningful comparison of findings across sites and generalizability. Importantly, the broad objective is to translate the National Frailty Strategy to real-world practice. To facilitate this, the intervention will be structured around core functions that are common across all 5 participating clinics, including Clinical Frailty Scale score assignment, IC screening for mildly frail older adults, further clinical assessment, tailored referrals to clinics, specialists or community services, and monitoring and follow-up. While the form of implementation may vary across clinics, such as the precise age cut-offs for inclusion, further clinical assessment tools, and referral workflows, the underlying core functions remain consistent and intentional. This approach is aligned with the core functions and forms paradigm, recognizing that complex interventions must adapt to context while preserving their core purpose [59-61].

For the 2 of 5 clinics that will apply a slightly higher age cut-off of 65 years (instead of 60 years) for inclusion, it serves to facilitate integration of IMPACTFrail's frailty screening with existing health programs and clinical protocols. They anticipated operational challenges and a lack of manpower to intentionally identify older adults 60-64 years old for the feasibility study.

Additionally, 2 clinics under the same regional health system have additional inclusion criteria. These criteria include (1) enrolment in Healthier SG [56], (2) absence of dementia diagnosis, and (3) active chronic disease appointments for any one of the following chronic conditions: Type 2 diabetes mellitus, chronic ischemic heart disease, congestive cardiac failure, osteoporosis, Stage 4 chronic kidney disease, and chronic obstructive pulmonary disease. These criteria are driven by contextual operational considerations within these clinics, including manpower capacity, logistics, and the need to align with ongoing preventive and management care programs for patients with chronic conditions. This more targeted approach was assessed to be more acceptable and pragmatic for small-scale feasibility testing in settings with constrained resources, while keeping the same core frailty screening, IC screening, and management functions.

Overall, this study protocol serves the purpose of disseminating knowledge on co-developing a health care intervention and implementation strategies, with built-in flexibility to accommodate system-level differences across clinics. In line with implementation research, we will subsequently assess how clinics could gradually align in eligibility criteria and protocols as IMPACTFrail scales up nationally.

No formal sample size calculation is performed, as this is a feasibility study not designed to be statistically powered for assessing effectiveness and health outcomes. There is no consensus in the literature on the required sample size for pilot and feasibility studies [62]. Recommendations of sample size vary depending on the primary objective of the study, the feasibility parameter being evaluated, or clinical scale evaluation [63-66]. Published reviews from the United Kingdom have reported median sample sizes of 30 and 36 participants per arm in feasibility studies [67,68]. In addition, clinics leads were also consulted on the practical projections of recruitment and uptake, taking into account the study timeframe and available resources. Hence, our study aims to recruit 36 older adults per clinic. This was computed based on targeting a minimum of 30 participants per site while accounting for an anticipated 20% loss to follow-up (eg, nonattendance at follow-up assessment post screening). Each clinic is conceptualized as an individual site due to contextualized barriers, workflows, and infrastructure.

Feasibility Criteria

Table 2 describes the four feasibility assessment criteria: (1) implementation, (2) acceptability, (3) practicality, and (4) adaptability, along with the corresponding study components mapped to each criterion. The criteria are informed by the literature on feasibility studies [47]. In summary, we prioritize the need to assess the ability to carry out the new screening and integrating the core functions with existing clinical protocols (implementability), how a satisfying or attractive a new program that focuses on domains beyond the conventional chronic conditions is to health care providers and older people (acceptability), whether the resources can support such a new program (practicality), and the degree it can be improved or changed if we progress beyond the pilot (adaptability) [47]. The criteria collectively inform us on the feasibility to progress from the small-scale pilot study to a main study that will be statistically powered to evaluate effectiveness and implementation at scale.

XSL•FO
RenderX

Table 2. Mapping of study components to feasibility assessment criteria.

Criteria	Study components
Implementation: to assess how well the program's core components and implementation strategies are implemented as intended.	• Study component 1: quantitative process indicators • Study component 2: assess fidelity of implementation strategies • Study component 3: qualitative study • Identify deviations or adaptations of the core components • Elicit barriers and facilitators perceived by implementers and older adults on implementing or receiving the program and its implementation strategies • Study component 5: program description • Study component 6: logic model and theory of change
Acceptability: to assess the satisfaction and adoptability of the program's core components and implementation strategies.	• Study component 3: qualitative study to identify barriers and facilitators to acceptability, adoptability, and appropriateness perceived by implementers and older adults on the program and its implementation strategies
Practicality: to assess the ability to carry out the program's core components and implementation strategies in a sustained manner.	• Study component 3: qualitative study to identify barriers and facilitators to (1) delivering the program and implementation strategies with the given resources (eg, manpower, information technology infrastructure, physical infrastructure) and (2) sustainability • Study component 4: cost analysis
Adaptability: to assess the extent that barriers to implementation, acceptability, adoptability, practicality, sustainability and scalability can be mitigated. This section includes assessing if the program and implementation strategies can be adapted and refined, and if we have the ability to develop additional implementation strategies if we scale.	• Study component 3: qualitative study to gain insights on (1) adaptations to the delivery of the program's core components and implementation strategies, to facilitate implementation, acceptability, adoptability, practicality, sustainability, and scalability and (2) selecting and developing additional implementation strategies if we scale

Study Component 1: Quantitative Process Indicators on Care Processes

Aim

We will collect and document a comprehensive list of quantitative process indicators to measure the processes of care (Table 3). This addresses the research question on whether the program activities are delivered as planned and informs us on the follow-through of patients for the entire patient journey.

XSL•FO
RenderX

Table 3. Process indicators.

Process indicators	Measure	Proctor implementation outcomes
Patient recruitment		
Number and proportion of older adults screened for frailty using the Clinical Frailty Scale (CFS)	Volume of older adults screened for frailty using the CFS	• Reach • Adoption
Number and proportion of older adults screened as CFS 4-5	Volume of older adults with mild-to-moderate frailty (CFS 4-5)	• Reach • Adoption
Number and proportion of CFS 4-5 older adults referred to IC screening	Volume of older adults directed to secondary IC screening	• Reach • Adoption
Intrinsic Capacity (IC) screening for CFS 4-5		
Number and proportion of older adults who initiated IC screening	Receptivity to IC screening	• Reach • Adoption
Time between CFS screening to initiation of IC screening (in days)	Efficiency of initiating IC screening after CFS segmentation	• Reach • Adoption
Number and proportion of older adults who completed all 6 domains of IC screening	Completion of IC screening after it was initiated	• Reach • Adoption
Number and proportion of older adults that completed each IC domain	Completion rates of each IC screening domain	• Acceptability
Duration between initiation of IC screening to completion of IC screening of 6 domains (in days or weeks)	Efficiency of completing all 6 domains of IC screening after initiation of IC screening. For example, IC screening of 6 domains may be staggered across two visits. This indicator measures the duration between the screening of at least 1 IC domain to the completion of all 6 IC domains	• Fidelity
Follow-up assessment		
Number and proportion of referrals to follow-up assessment for each IC deficit	This measures care coordination: (1) number of referrals made to services for follow-up assessment for each IC deficit and (2) proportion of referrals to each service with the denominator as the total number of referrals	• Reach • Adoption • Penetration • Fidelity
Number and proportion of actualized referrals to follow-up assessment for each IC deficit	This measures care coordination: (1) number of actualized referrals to services for follow-up assessment for each IC deficit and (2) proportion of actualized referrals to each service with the denominator as the total number of referrals	• Reach • Acceptability • Costs
Downstream care pathways		
Number and proportion of referrals for the management of each IC deficit	This measures care coordination: (1) number of referrals made to services for management of each IC deficit and (2) proportion of referrals to each service with the denominator as the total number of referrals	• Adoption • Penetration • Fidelity
Number and proportion of actualized referrals for the management of each IC deficit	This measures care coordination: (1) number of actualized referrals to services for management of each IC deficit and (2) proportion of actualized referrals to each service with the denominator as the total number of referrals	• Reach • Penetration • Costs
Capability building		
• Number of staff trained to conduct CFS and IC screening, and to provide referrals to follow-up assessments and care pathways • Include breakdown by occupation (doctor, nurse, care coordinator, and patient service associate)	Capability building for screening and delivery of referrals	• Adoption • Feasibility • Acceptability • Costs

XSL•FO
RenderX

Study Sample and Data Source

The sample includes all recruited older adult participants. Each of the 5 clinics will collect and provide data to the research team. Table 3 summarizes the list of process indicators.

Study Component 2: Fidelity Indicators for Implementation Strategies

Aim

We aim to assess the fidelity of the selected implementation strategies.

Study Procedure

We will monitor and track the fidelity of operationalizing each selected ERIC implementation strategy across the specified domains (actor, action, action target, dose, or temporality), using a comprehensive set of indicators [45]. In other words, we will assess whether each strategy is delivered in the feasibility study as originally intended during the co-development phase.

Indicators will be collaboratively designed by the research team and clinic implementation leads. Data will be collected regularly by clinic implementation leads and frontline implementers. The research team will systematically document any deviations from the planned implementation and discuss them with clinic implementation leads. Adaptations to strategies, if any, will also be documented. For example, in the case of the implementation strategy involved educational sessions for capability building on IC screening, we will assess whether the planned elements, such as the trainers (actor), attendees (target), and the frequency and duration of training (dose), are delivered as intended. If discrepancies arise, we will explore whether fidelity can be improved or whether adaptations are needed. In addition, we will focus on sustainability and scalability, such as whether the frequency and dose of the strategies are sustainable within the clinic and adaptable for scaling up to other clinics.

Study Component 3: Qualitative Study

Aim

Broadly, we aim to qualitatively assess:

1. The level of feasibility (implementation, acceptability, adoptability, practicality, and adaptability), sustainability, and scalability of IMPACTFrail's core components and implementation strategies, and barriers and facilitators to their attainment.
2. Effectiveness of implementation strategies. This refers to the attribution of implementation strategies to attaining feasibility, sustainability, and scalability.

Study Sample and Data Source

Study participants will include older adult participants, clinic implementation leads who oversaw the planning and execution of IMPACTFrail, and frontline health care providers involved in the program's core functions (doctors, nurses, care managers, care coordinators, allied health professionals, and support staff). We will conduct purposive sampling and sample sizes will be guided by the concept of sufficiency [51,52]. First, we will include older adults who participated in IMPACTFrail. We aim to interview a minimum of 3-4 older adults per clinic, with a total of at least 15 participants (approximately 10% of the targeted 36 recruited participants per clinic). We will sample the older adults using maximum variation for representation across demographic variables such as sex, ethnic groups, and age groups. Second, the qualitative study component will include clinic implementation leads who oversaw the planning and execution of IMPACTFrail and frontline health care providers involved in the program's core functions (doctors, nurses, paraprofessional support staff). We anticipate only a small number of health care workers being involved in this small-scale pilot study. Hence, we will aim to sample all implementation leads at each clinic, estimated to range from 2 to 4 individuals per site. Additionally, we aim to sample frontline health care workers across every occupation who was involved across each central function of IMPACTFrail. Nurses and paraprofessional support staff in the qualitative study would need to have conducted IC screening for at least 5 older adults and doctors would need to conduct follow-up clinical assessments to assess intrinsic capacity deficits for at least 5 older adults. Interviews and focus groups with clinic implementation leads and frontline health care providers will be conducted in English, while interviews with older adults will be in either English or Chinese.

Study Procedure

As mentioned, both data collection and data analysis will be guided by the 4 feasibility criteria, the updated CFIR and its Outcomes addendum, TDF, and ERIC taxonomy [31,32,49]. We will expand CFIR's Individual Characteristics domain by mapping determinants of capability, opportunity, and motivation to TDF domains [69,70]. The ERIC taxonomy informs the refinement of existing implementation strategies and selection of new strategies [32]. The semistructured topic guides elicit data that inform the 4 feasibility criteria, CFIR-based and TDF-based barriers and facilitators, and suggestions on new ERIC strategies. For example, for the criteria on implementation, we will obtain perspectives from implementers on the barriers and facilitators to delivering IMPACTFrail's core components and implementation strategies (eg, mobilizing program champions, attending training), and perspectives on the extent that the strategies enabled the implementation of IMPACTFrail. Older adults will share about the barriers and facilitators of participating in IMPACTFrail and receiving implementation strategies (eg, patient educational materials) and the extent to which the strategies have facilitated their participation.

We aim to conduct 2 rounds of 14 interviews with clinic implementation leads across all 5 clinics. The first round of interviews will be conducted within the first 2 months of IMPACTFrail's roll out. This will inform early implementation barriers to refine its delivery and inform additional support required. The second round of interviews will be conducted toward the last 4-5 months of the study and after the first 20-25 older adults have been recruited. This is to elicit their perspectives on how implementation has progressed, new barriers and facilitators that arose, and insights on improving the program. Additionally, we will conduct interviews or focus groups with frontline health care providers involved in delivering IMPACTFrail's core functions. In the first instance, we will aim to conduct focus groups, as opposed to interviews,

XSL•FO
RenderX

among the health care providers who worked as a team, if scheduling allows.

Deductive analysis will be guided by the frameworks and concurrent inductive analysis will reveal emergent barriers, facilitators, and themes. Two trained research team members will code verbatim transcripts independently and subsequently reconcile discrepancies. Framework analysis will be conducted [71]. This allows timely charting of data into a framework matrix, generation of themes by linking conceptually related codes, and directly addressing the feasibility criteria (Table 2). We will conduct a thematic analysis using data collected across all 3 regional health systems, as we anticipate cross-cutting barriers, facilitators and learning points across study sites. We will also identify contextual insights specific to each regional health system. This analytic approach allows us to synthesize cross-site patterns with a larger amount of data, while accounting for differences in organizational context, and directly address the feasibility of implementation across settings. A pragmatic approach of concurrent data collection and analysis will be adopted to ensure timely reporting of findings to clinic leadership and stakeholders to support the preparation of the scaled-up phase [72-75]. Microsoft Excel and NVivo 12 (QSR International) will be used for data management.

Study Component 4: Cost Analysis

Aim

We aim to estimate the cost of developing and delivering the program and implementation strategies, with the purpose of informing the amount of resources required to sustain and to expand the program beyond the feasibility study.

Study Procedure and Data Source

First, we will compute developmental costs, which comprise (1) developing IMPACTFrail's core functions, and (2) developing each implementation strategy. The expenditures on manpower (eg, health care providers involved in the co-development phase) and other expenses (materials, printing, etc) will be aggregated. Second, program costs will be derived from the costs of (1) delivering IMPACTFrail's core components, and (2) operationalizing each implementation strategy. The expenditures on manpower, equipment, and operating expenses will be aggregated.

A time-driven, activity-based micro-costing approach will be used to estimate costs related to manpower resourcing [76,77]. Specifically, this will be derived from the unit cost multiplied by the intensity of use and duration of use. The unit cost will be estimated based on financial records or the standardized workforce unit cost for implementers obtained from the clinics. The intensity and duration of use will be derived from self-reported time spent for a select list of activities by implementers to deliver the program [78]. We will only use financial statements or costs self-reported by implementation leads to obtain unit costs on equipment and other operating expenses and their intensity of use.

Within each site, we will present the cost associated with (1) developing IMPACTFrail's core functions, (2) developing each implementation strategy, (3) delivering IMPACTFrail's core functions, (4) delivering each implementation strategy, (5) total developmental cost, (6) total program cost, and (7) total costs.

Study Component 5: Program Description

We aim to produce written program descriptions across clinics guided by frameworks including the Template for Intervention Description and Replication (TIDieR) checklist or the Aims, Ingredients, Mechanism, and Delivery (AMID) framework [79,80]. The program descriptions will facilitate transparency and replicability.

Study Component 6: Logic Model and Theory of Change

The logic model and theory of change depict a visual representation of the resources, key activities, outputs, outcomes, and plausible causal pathways connecting them [81,82]. Figure 2 displays the initial theory of change. During this feasibility study, we will refine the logic model and theory of change. This updated logic model and theory of change will guide our understanding on the mechanisms of change and on the short to longer term health outcomes to evaluate in the future study.

XSL•FO
RenderX

Figure 2. Initial logic model and theory of change. Bold arrow represents the link of an activity to outputs or outcomes. Broken arrow represents the link of an activity to outputs or outcomes. ADL: activity of daily living; CFS: Clinical Frailty Scale; ED: emergency department; GERI: Geriatric Education & Research Institute; HR-QoL: health-related quality of life; iADL: instrumental activity of daily living; IC: intrinsic capacity; LOS: length of stay; SOC: specialist outpatient clinic; SOP: standard operating procedures;.

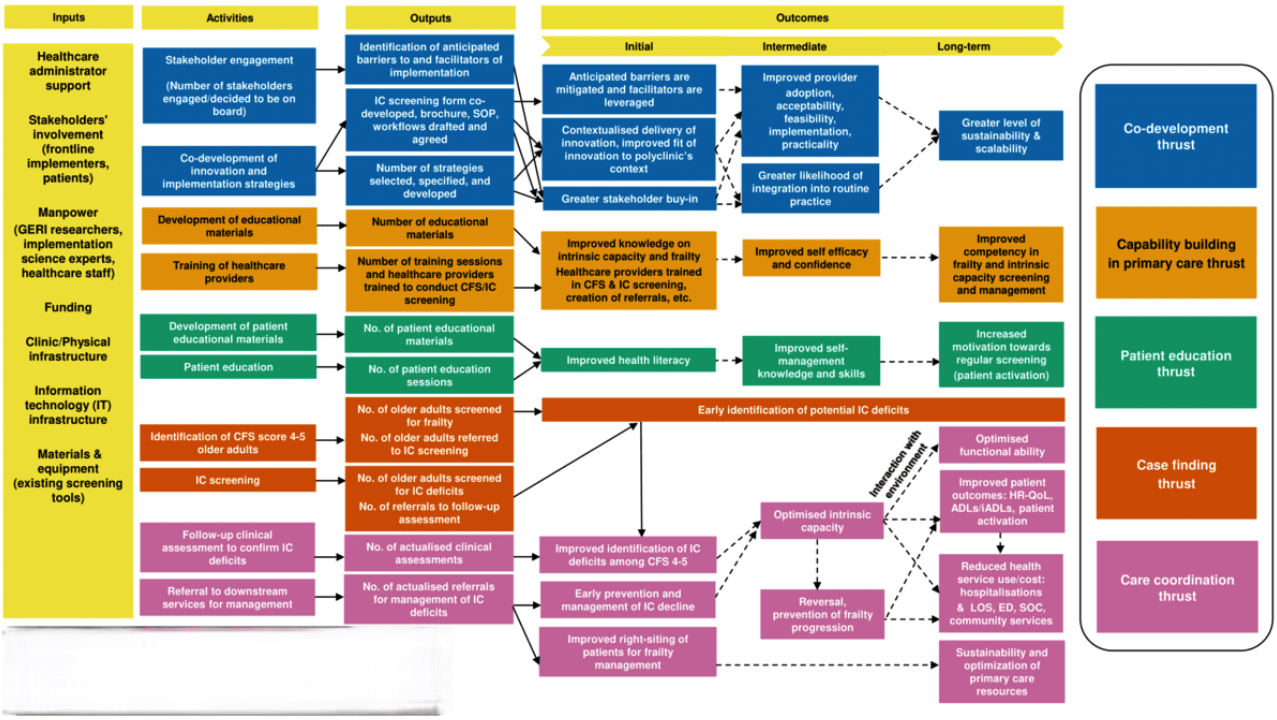

Triangulation of Findings

For each of the 4 feasibility criteria, the findings across study components will be presented and narratively triangulated (Table 2). Next, findings across all criteria will be narratively triangulated and assessed holistically to inform feasibility and readiness to scale. Briefly, for instance, under the criteria on implementation, we will triangulate the quantitative process indicators on completion rates of IC screening, along with CFIR-based barriers to participating perceived by older adults and barriers to implementing IC screening by health care providers. Under the criteria on practicality, we may triangulate the cost analysis of incorporating IC screening into clinic workflows, such as manpower and operating costs, CFIR-based barriers on resource allocation and management, and perspectives on sustainability. Under the criteria on adaptability, we will elicit if the barriers could be mitigated by implementation strategies and refinements to the program. A key feasibility assessment result that will suggest a "go" is at least 80% of the targeted recruitment at each of the participating clinics; that is, at least 29 of the targeted 36 older adults are recruited into the study and complete IC screening. This will demonstrate that the key novel core function of the program, which is to deliver IC screening of 6 health domains, is feasible. Importantly, we will be guided and informed by the findings across the multiple study components and methodologies. Each of the 3 regional health systems in Singapore will be engaged. Overall, we will evaluate the readiness to scale IC screening to more older adults at the 5 clinics and to additional new clinics across the 3 health care systems. Stakeholders will be engaged to review the findings and discuss readiness to scale, including research directors from the 3 health care systems, health care administrators (clinic leadership), and frontline health care providers. This process is aligned with the UK MRC Framework, FAID, and implementation research frameworks.

Ethical Considerations

Ethical approval for this study was obtained from the NHG Domain Specific Review Board (ECOS Ref: 2024-3878). Written informed consent will be taken from all participants prior to their participation in the study. Older adult participants will receive SGD $30 (US $23.65) upon completion of informed consent-taking and IC screening, and SGD $50 (US $39.42) upon completion of the interview. All data collected will be anonymized and securely stored in a restricted-access network accessible only to the research team.

Results

The study was funded in September 2024. Data collection for the feasibility testing of IMPACTFrail commenced in April 2025 and will conclude by March 2026. As of manuscript submission, only 98 participants have been recruited across 5 sites. Recruitment, data collection, and data analysis are still ongoing. The results are expected to be published in early 2027.

Discussion

Principal Findings

IMPACTFrail seeks to translate Singapore's National Frailty Strategy and the WHO ICOPE framework by serving as an IC detection and management program for mildly frail older adults in primary care [16]. It contributes to a holistic approach to healthy aging by enabling early mitigation of frailty progression and optimizing functional ability. At a broader level, it is aligned

XSL•FO
RenderX

with the WHO public health framework on healthy aging [7], strengthens Singapore's contribution to the United Nations Decade of Healthy Ageing (2021-2030) [83], and complements national initiatives on preventative care for chronic conditions in primary care (eg, Healthier SG) and aging well in the community (eg, Age Well SG) [41,42].

The implementation and evaluation of the ICOPE framework in real-world practice is still in its nascence [17]. A systematic review reported that most studies have only used IC primarily as an outcome measure, rather than to inform intervention design [84]. The existing studies have focused on examining the prevalence of IC deficits or tool performance [85-87]. Implementation of the ICOPE programs in the community by early adopters in France and China demonstrated general acceptance of IC screening, but there are still gaps in understanding how to increase uptake and adherence among less activated older adults, and scale up beyond the feasibility testing sites that had more resources and stakeholder buy-in [19-21]. In Korea, Won [38] described the first small-scale ICOPE program implemented in primary care, contextualized to concurrently assess frailty and IC, using a 53-item screening approach. However, scaling up and effectiveness have not been reported.

In response, our adoption of a systematic and staged approach grounded in implementation science aims to accelerate the translation of Singapore's National Frailty Strategy and the WHO ICOPE framework with high reach, fidelity, adoption, sustainability, and scalability. Moreover, embedding implementation considerations from the outset may strengthen the program's design and delivery, hence improving its effectiveness, sustainability, and scalability in real-world settings. If the findings indicate limited effectiveness, the likelihood of confounding from poor implementation and fidelity is lower; conversely, effectiveness elicited under research study conditions tends to be better sustained in real-world settings [88-90]. A second key strength is the integration of complex intervention research and implementation research frameworks to guide the co-development of the program and implementation strategies, the feasibility assessment, and the qualitative study. This integrated approach enhances methodological rigor and increases the likelihood of successful and contextualized implementation of a complex intervention. Third, IMPACTFrail is co-developed from the outset in collaboration with multiple stakeholders across all 3 regional health systems in Singapore. This is imperative in facilitating scalability across more clinics within each cluster and reduces cross-cluster fragmentation of services. Fourth, our feasibility assessment is informed by published guidelines on conducting feasibility studies, precise mapping of study components to feasibility criteria, and adopting a mixed methods approach. This facilitates the depth and quality of insights through triangulation of findings.

There are several potential limitations. First, variations in eligibility criteria across clinics may limit the comparability of findings across sites and reduce generalizability. However, this reflects the real-world implementation context of IMPACTFrail, which is operationalized across three public health care clusters with distinct workflows, resources, and different population needs. As such, contextual adaptations guided by complex intervention research and implementation science frameworks are necessary. While this may limit generalizability, it generates insights to inform the future scale-up of IMPACTFrail and other similar programs in primary care. Second, findings may be less generalizable to settings beyond primary care and to older adults who are not mildly frail. However, insights related to common barriers, such as capability building, resource allocation and patient-level behaviors, will be applicable. We will also report how these insights are used to refine the delivery of the program's core steps and implementation strategies to support future scale-up. Importantly, this article still serves to disseminate generalizable knowledge on the implementation research methodologies and approaches to co-develop an ICOPE-based intervention with multiple stakeholders from the outset, and the structure use of feasibility testing prior to a larger main study. Third, the feasibility study does not include a comparator arm. However, the primary aim is not to assess effectiveness but to apply implementation science methodologies that enhance the likelihood of sustainability and scalability of a complex intervention nationwide. Effectiveness outcomes will be evaluated in a subsequent hybrid effectiveness-implementation study [57,58], which will be statistically powered to evaluate impacts on patient outcomes (eg, function, quality of life), health care use, and cost effectiveness. Lastly, tracking actualization rates of health and social care services for managing IC deficits may be incomplete due to the absence of a common information technology system across public primary care, private specialist clinics, and selected community organizations. Nevertheless, referral actualization within the primary care setting (eg, memory clinic and health and wellness clinic) and most public specialist outpatient clinics will be captured and monitored.

Conclusions

In summary, this study contributes to the literature by providing a detailed study protocol on the co-development and feasibility testing of a complex intervention to enhance transparency, fidelity, and replicability. It demonstrates the application of a systematic implementation science approach, the integration of complex intervention and implementation research frameworks, and a multimethods feasibility assessment to accelerate the translation of evidence to sustainable and scalable programs in real-world practice.

Acknowledgments

We wish to express our gratitude to the following colleagues for supporting the pilot implementation of IMPACTFrail and the design of this study protocol: Shannon Goh Yi Xuan, Ong Hui Chin, Liang Jieying, Lim Xin Hui Priscilla, Fadzlina Binte Sujak,

XSL•FO
RenderX

Torsa Das, Adelene Ho Pei Ee, Noraiani Bte Saidi, Ng Siew Lan, Sim Geok Lan Jacqueline, Jacelyn Tan Li Ping, Chris Goh Wan Teng, Anna Low Ai Yen, and Sandy Tan Mui Kiang.

Members of the IMPACTFrail Program Group:

Jasmine Tan Shi Hui, Geriatric Education & Research Institute, Singapore

Kelly Wong Young Lee, Geriatric Education & Research Institute, Singapore

Robin Choo Wai Munn, Geriatric Education & Research Institute, Singapore

Zheng Dexin, National Healthcare Group Polyclinics, Singapore

Fatimah Binte Endut, National University Polyclinics, Singapore

Chiu Wai Leong, National University Polyclinics, Singapore

Christine Kho Zhen-Bing, SingHealth Polyclinics, Singapore

Isabelle Yoong Rui Wen, SingHealth Polyclinics, Singapore

Laura Tay Bee Gek, Department of General Medicine, Sengkang General Hospital, Singapore

Jinnat Bte Ahmad Marican, SingHealth Polyclinics, Singapore

Joanna Tan Lee Lim, SingHealth Polyclinics, Singapore

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Funding

This study was funded by the Geriatric Education & Research Institute (GERI) Intramural Grant (Project Reference: GERI1637).

Authors' Contributions

Conceptualization: GS, WST, YYD
Funding Acquisition: GS
Methodology: GS, WST, YYD
Project Administration: GS
Resources: GS, DWLN, JCAN, RJYH, SCK, SZW, CJN
Supervision: GS
Visualization: QW
Writing - Original Draft: QW
Writing - Review & Editing: GS, MG, EH, SFG, JG, WST, YYD, DWLN, JCAN, RJYH, SCK, SZW, CJN

Conflicts of Interest

None declared.

Multimedia Appendix 1
SPIRIT 2025 checklist of items to address in a randomized trial protocol.
[PDF File (Adobe PDF File), 180 KB - resprot_v15i1e84257_app1.pdf]

Multimedia Appendix 2
World Health Organization integrated care for older people (WHO ICOPE) screening tool for submission.
[PDF File (Adobe PDF File), 164 KB - resprot_v15i1e84257_app2.pdf]

References

1. World report on ageing and health. In: World Health Organization. Geneva, Switzerland: WHO; 2015.
2. WHO Clinical Consortium on Healthy Ageing. Report of consortium meeting 1-2 December 2016 in Geneva. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/WHO-FWC-ALC-17.2#:~:text=This%20report%20summarizes%20the%20key,for%20older%20people%20(ICOPE) [accessed 2026-01-14]
3. Zhao Y, Jiang Y, Tang P, Wang X, Guo Y, Tang L. Adverse health effects of declined intrinsic capacity in middle-aged and older adults: a systematic review and meta-analysis. Age Ageing 2024;53(7):afae162. [doi: 10.1093/ageing/afae162] [Medline: 39058916]

XSL•FO
RenderX

4. Sánchez-Sánchez JL, Lu WH, Gallardo-Gómez D, del Pozo Cruz B, de Souto Barreto P, Lucia A, et al. Association of intrinsic capacity with functional decline and mortality in older adults: a systematic review and meta-analysis of longitudinal studies. Lancet Healthy Longev 2024;5(7):e480-e492. [doi: 10.1016/s2666-7568(24)00092-8]
5. Pan Y, Li X, Zhang L, Li Y, Tang Z, Ma L. Declined intrinsic capacity predicts long-term mortality in Chinese older adults: Beijing longitudinal study of aging. Maturitas 2024;188:108082. [doi: 10.1016/j.maturitas.2024.108082] [Medline: 39089049]
6. Zhou Y, Ma L. Intrinsic capacity in older adults: recent advances. Aging Dis 2022;13(2):353-359. [doi: 10.14336/ad.2021.0818]
7. Beard JR, Officer A, de Carvalho IA, Sadana R, Pot AM, Michel J, et al. The World report on ageing and health: a policy framework for healthy ageing. Lancet 2016;387(10033):2145-2154 [FREE Full text] [doi: 10.1016/S0140-6736(15)00516-4] [Medline: 26520231]
8. Briggs AM, Valentijn PP, Thiyagarajan JA, Araujo de Carvalho I. Elements of integrated care approaches for older people: a review of reviews. BMJ Open 2018;8(4):e021194 [FREE Full text] [doi: 10.1136/bmjopen-2017-021194] [Medline: 29627819]
9. Kojima G, Liljas AEM, Iliffe S. Frailty syndrome: implications and challenges for health care policy. Risk Manag Healthc Policy 2019;12:23-30 [FREE Full text] [doi: 10.2147/RMHP.S168750] [Medline: 30858741]
10. Lekan DA, Collins SK, Hayajneh AA. Definitions of frailty in qualitative research: a qualitative systematic review. J Aging Res 2021;2021:6285058 [FREE Full text] [doi: 10.1155/2021/6285058] [Medline: 34123425]
11. Belloni G, Cesari M. Frailty and intrinsic capacity: two distinct but related constructs. Front Med (Lausanne) 2019;6:133 [FREE Full text] [doi: 10.3389/fmed.2019.00133] [Medline: 31275941]
12. Yu R, Leung J, Leung G, Woo J. Towards healthy ageing: using the concept of intrinsic capacity in frailty prevention. J Nutr Health Aging 2022;26(1):30-36 [FREE Full text] [doi: 10.1007/s12603-021-1715-2] [Medline: 35067700]
13. Chew J, Lim JP, Yew S, Yeo A, Ismail NH, Ding YY, et al. Disentangling the relationship between frailty and intrinsic capacity in healthy community-dwelling older adults: a cluster analysis. J Nutr Health Aging 2021;25(9):1112-1118 [FREE Full text] [doi: 10.1007/s12603-021-1679-2] [Medline: 34725670]
14. Tay L, Tay EL, Mah SM, Latib A, Koh C, Ng Y. Association of intrinsic capacity with frailty, physical fitness and adverse health outcomes in community-dwelling older adults. J Frailty Aging 2023;12(1):7-15 [FREE Full text] [doi: 10.14283/jfa.2022.28] [Medline: 36629078]
15. Tay L, Tay EL, Mah SM, Latib A, Ng Y. Intrinsic capacity rather than intervention exposure influences reversal to robustness among prefrail community-dwelling older adults: a non-randomized controlled study of a multidomain exercise and nutrition intervention. Front Med (Lausanne) 2022;9:971497 [FREE Full text] [doi: 10.3389/fmed.2022.971497] [Medline: 36341237]
16. Frailty strategy policy report. Ministry of Health Singapore. URL: https://www.moh.gov.sg/resources-statistics/reports/frailty-strategy-policy-report [accessed 2024-07-22]
17. Sum G, Lau LK, Jabbar KA, Lun P, George PP, Munro YL, et al. The World Health Organization (WHO) integrated care for older people (ICOPE) framework: a narrative review on its adoption worldwide and lessons learnt. Int J Environ Res Public Health 2022;20(1):154 [FREE Full text] [doi: 10.3390/ijerph20010154] [Medline: 36612480]
18. Ginting ML, Sum G, Wang SZ, Ding YY, Tay L. Integrated frailty and intrinsic capacity care model for community-dwelling older adults in Singapore: a rapid qualitative study of anticipated implementation barriers and enablers using the consolidated framework for implementation research and its outcomes addendum. Front Health Serv 2025;5:1563686 [FREE Full text] [doi: 10.3389/frhs.2025.1563686] [Medline: 40343232]
19. Berbon C, Rolland Y, Takeda C, Lafont C, Tavassoli N, De Kerimel J, et al. WHO ICOPE programme adherence of 8672 older age people over 2-years of follow-up. J Adv Nurs 2025;81(9):5925-5936. [doi: 10.1111/jan.16740] [Medline: 40195235]
20. Tavassoli N, de Souto Barreto P, Berbon C, Mathieu C, de Kerimel J, Lafont C, et al. Implementation of the WHO integrated care for older people (ICOPE) programme in clinical practice: a prospective study. Lancet Healthy Longev 2022;3(6):e394-e404. [doi: 10.1016/s2666-7568(22)00097-6]
21. Yan Wang N, Liu X, Kong X, Sumi Y, Chhetri JK, Hu L, et al. Implementation and impact of the World Health Organization integrated care for older people (ICOPE) program in China: a randomised controlled trial. Age Ageing 2024;53(1):afad249. [doi: 10.1093/ageing/afad249] [Medline: 38251736]
22. Ma CHK, Chua DQL, Tay L, Teo EWC, Ng WC, Leung AYM. The feasibility of implementing the WHO integrated care for older people (ICOPE) framework in Singapore. J Frailty Aging 2024;13(4):514-521. [doi: 10.14283/jfa.2024.59] [Medline: 39574276]
23. Ma C, Leung AYM, Chua D, Teo WC, Tay LBG, Ng WC. Intrinsic capacity of older adults in Singapore using who integrated care for older people (ICOPE) framework. Innov Aging 2023;7(1):893-894. [doi: 10.1093/geroni/igad104.2875]
24. Primary care. Ministry of Health Singapore. URL: https://www.moh.gov.sg/seeking-healthcare/find-a-facility-or-service/types-of-medical-facilities-and-services/primary-care-services [accessed 2025-05-19]
25. Wang Q, Tan LF, Tan WM, Merchant RA, Sevdalis N. Implementing the ICOPE program amongst community-dwelling older adults in Singapore: a multistage implementation study protocol. Front Public Health 2025;13:1672852 [FREE Full text] [doi: 10.3389/fpubh.2025.1672852] [Medline: 41458401]

XSL•FO
RenderX

26. Low WL, Sultana R, Huda Mukhlis AB, Ho JCY, Latib A, Tay EL, et al. A non-controlled study of a multi-factorial exercise and nutritional intervention to improve functional performance and prevent frailty progression in community-dwelling pre-frail older adults. JAR Life 2021;10:1-7 [FREE Full text] [doi: 10.14283/jarlife.2021.1] [Medline: 36923514]
27. Bauer MS, Kirchner J. Implementation science: what is it and why should I care? Psychiatry Res 2020;283:112376 [FREE Full text] [doi: 10.1016/j.psychres.2019.04.025] [Medline: 31036287]
28. Peters DH, Adam T, Alonge O, Agyepong IA, Tran N. Implementation research: what it is and how to do it. BMJ 2013;347:f6753. [doi: 10.1136/bmj.f6753] [Medline: 24259324]
29. May CR, Johnson M, Finch T. Implementation, context and complexity. Implement Sci 2016;11(1):141 [FREE Full text] [doi: 10.1186/s13012-016-0506-3] [Medline: 27756414]
30. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of medical research council guidance. BMJ 2021;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
31. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
32. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the expert recommendations for implementing change (ERIC) project. Implement Sci 2015;10:21. [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
33. Hróbjartsson A, Boutron I, Hopewell S, Moher D, Schulz K, Collins G, et al. SPIRIT 2025 explanation and elaboration: updated guideline for protocols of randomised trials. BMJ 2025;389:e081660 [FREE Full text] [doi: 10.1136/bmj-2024-081660] [Medline: 40294956]
34. Thabane L, Lancaster G. A guide to the reporting of protocols of pilot and feasibility trials. Pilot Feasibility Stud 2019;5(1):37. [doi: 10.1186/s40814-019-0423-8] [Medline: 30858987]
35. Rockwood K, Theou O. Using the clinical frailty scale in allocating scarce health care resources. Can Geriatr J 2020;23(3):210-215. [doi: 10.5770/cgj.23.463] [Medline: 32904824]
36. Sum G, Choo RWM, Nai ZL, Goh SF, Lim WS, Ding YY, Geriatric Services Hub Program Group. A novel integrated geriatric services hub for frailty identification and comprehensive management of community-dwelling older adults in Singapore: impact on health service utilization. J Am Geriatr Soc 2025;73(5):1576-1587. [doi: 10.1111/jgs.19339] [Medline: 39757993]
37. Lee H, Lee E, Jang IY. Frailty and comprehensive geriatric assessment. J Korean Med Sci 2020;35(3):e16 [FREE Full text] [doi: 10.3346/jkms.2020.35.e16] [Medline: 31950775]
38. Won CW, Ha E, Jeong E, Kim M, Park J, Baek JE, et al. World Health Organization integrated care for older people (ICOPE) and the integrated care of older patients with frailty in primary care (ICOOP_Frail) study in Korea. Ann Geriatr Med Res 2021;25(1):10-16 [FREE Full text] [doi: 10.4235/agmr.21.0025] [Medline: 33794585]
39. Fit for frailty consensus best practice guidance for the care of older people living with frailty in community and outpatient settings. British Geriatrics Society. URL: https://www.bgs.org.uk/sites/default/files/content/resources/files/2018-05-14/fff2_short_0.pdf [accessed 2025-05-19]
40. Rodríguez-Laso Á, Mora M, Sánchez IG, Bouzón CA, Mañas LR, Bernabei R. Updated state of the art report on the prevention and management of frailty. ADVANTAGE Management Frailty 2019:1-48.
41. What is Healthier SG? Healthier SG. URL: https://tinyurl.com/2572yk82 [accessed 2025-05-19]
42. What is Age Well SG? Age Well SG. URL: https://www.agewellsg.gov.sg [accessed 2025-05-19]
43. Our partners. Ministry of Health Singapore. URL: https://www.moh.gov.sg/about-us/our-partners [accessed 2025-05-19]
44. O'Cathain A, Croot L, Sworn K, Duncan E, Rousseau N, Turner K, et al. Taxonomy of approaches to developing interventions to improve health: a systematic methods overview. Pilot Feasibility Stud 2019;5(1):41 [FREE Full text] [doi: 10.1186/s40814-019-0425-6] [Medline: 30923626]
45. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. Implement Sci 2013;8:139 [FREE Full text] [doi: 10.1186/1748-5908-8-139] [Medline: 24289295]
46. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019;5(1):114 [FREE Full text] [doi: 10.1186/s40814-019-0499-1] [Medline: 31608150]
47. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009;36(5):452-457. [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
48. Guidance on applying for feasibility studies. National Institute for Health and Care Research. URL: https://www.nihr.ac.uk/funding-programmes/research-for-patient-benefit/scope-eligibility/guidance-applying-feasibility-studies [accessed 2024-09-17]
49. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the theoretical domains framework of behaviour change to investigate implementation problems. Implement Sci 2017;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
50. Waltz TJ, Powell BJ, Matthieu MM, Damschroder LJ, Chinman MJ, Smith JL, et al. Use of concept mapping to characterize relationships among implementation strategies and assess their feasibility and importance: results from the Expert Recommendations for Implementing Change (ERIC) study. Implement Sci 2015;10(1):109 [FREE Full text] [doi: 10.1186/s13012-015-0295-0] [Medline: 26249843]

XSL•FO
RenderX

51. LaDonna KA, Artino AR, Balmer DF. Beyond the guise of saturation: rigor and qualitative interview data. J Grad Med Educ 2021;13(5):607-611 [FREE Full text] [doi: 10.4300/JGME-D-21-00752.1] [Medline: 34721785]
52. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
53. Shin MH, Montano AL, Adjognon OL, Harvey KLL, Solimeo SL, Sullivan JL. Identification of implementation strategies using the CFIR-ERIC matching tool to mitigate barriers in a primary care model for older veterans. Gerontologist 2023;63(3):439-450. [doi: 10.1093/geront/gnac157] [Medline: 36239054]
54. Weir A, Presseau J, Kitto S, Colman I, Hatcher S. Strategies for facilitating the delivery of cluster randomized trials in hospitals: a study informed by the CFIR-ERIC matching tool. Clin Trials 2021;18(4):398-407 [FREE Full text] [doi: 10.1177/17407745211001504] [Medline: 33863242]
55. Waltz TJ, Powell BJ, Fernández ME, Abadie B, Damschroder LJ. Choosing implementation strategies to address contextual barriers: diversity in recommendations and future directions. Implement Sci 2019;14(1):42 [FREE Full text] [doi: 10.1186/s13012-019-0892-4] [Medline: 31036028]
56. Campbell MK, Piaggio G, Elbourne DR, Altman DG, CONSORT Group. Consort 2010 statement: extension to cluster randomised trials. BMJ 2012;345(sep04 1):e5661-e5661. [doi: 10.1136/bmj.e5661] [Medline: 22951546]
57. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
58. Landes SJ, McBain SA, Curran GM. An introduction to effectiveness-implementation hybrid designs. Psychiatry Res 2019;280:112513 [FREE Full text] [doi: 10.1016/j.psychres.2019.112513] [Medline: 31434011]
59. Perez Jolles M, Lengnick-Hall R, Mittman BS. Core functions and forms of complex health interventions: a patient-centered medical home illustration. J Gen Intern Med 2019;34(6):1032-1038 [FREE Full text] [doi: 10.1007/s11606-018-4818-7] [Medline: 30623387]
60. Harrison MS. Functions and forms framework: a methodology for mechanistic deconstruction and adaptation? J Gen Intern Med 2021;36(9):2809-2811 [FREE Full text] [doi: 10.1007/s11606-020-06533-3] [Medline: 33469751]
61. Terrana A, Viglione C, Rhee K, Rabin B, Godino J, Aarons GA, et al. The core functions and forms paradigm throughout EPIS: designing and implementing an evidence-based practice with function fidelity. Front Health Serv 2023;3:1281690 [FREE Full text] [doi: 10.3389/frhs.2023.1281690] [Medline: 38292916]
62. Lewis M, Bromley K, Sutton CJ, McCray G, Myers HL, Lancaster GA. Determining sample size for progression criteria for pragmatic pilot RCTs: the hypothesis test strikes back!. Pilot Feasibility Stud 2021;7(1):40 [FREE Full text] [doi: 10.1186/s40814-021-00770-x] [Medline: 33536076]
63. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. Pharm Stat 2005;4(4):287-291. [doi: 10.1002/pst.185]
64. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010;10:1 [FREE Full text] [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
65. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health 2008;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]
66. Johanson GA, Brooks GP. Initial scale development: sample size for pilot studies. Educ Psychol Meas 2009;70(3):394-400. [doi: 10.1177/0013164409355692]
67. Billingham SAM, Whitehead AL, Julious SA. An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom clinical research network database. BMC Med Res Methodol 2013;13:104. [doi: 10.1186/1471-2288-13-104] [Medline: 23961782]
68. Totton N, Lin J, Julious S, Chowdhury M, Brand A. A review of sample sizes for UK pilot and feasibility studies on the ISRCTN registry from 2013 to 2020. Pilot Feasibility Stud 2023;9(1):188 [FREE Full text] [doi: 10.1186/s40814-023-01416-w] [Medline: 37990337]
69. O'Donovan B, Kirke C, Pate M, McHugh S, Bennett K, Cahir C. Mapping the theoretical domain framework to the consolidated framework for implementation research: do multiple frameworks add value? Implement Sci Commun 2023;4(1):100 [FREE Full text] [doi: 10.1186/s43058-023-00466-8] [Medline: 37620981]
70. Birken SA, Powell BJ, Presseau J, Kirk MA, Lorencatto F, Gould NJ, et al. Combined use of the consolidated framework for implementation research (CFIR) and the theoretical domains framework (TDF): a systematic review. Implement Sci 2017;12(1):2 [FREE Full text] [doi: 10.1186/s13012-016-0534-z] [Medline: 28057049]
71. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
72. Vindrola-Padros C, Johnson GA. Rapid techniques in qualitative research: a critical review of the literature. Qual Health Res 2020;30(10):1596-1604. [doi: 10.1177/1049732320921835] [Medline: 32667277]
73. Vindrola-Padros C, Brage E, Johnson GA. Rapid, responsive, and relevant?: a systematic review of rapid evaluations in health care. Am J Eval 2021;42(1):13-27. [doi: 10.1177/10982140198869l4]

XSL•FO
RenderX

74. Clark SE, Vera San Juan N, Moniz T, Appleton R, Barnett P, Vindrola-Padros C. A systematic review of the methods used in rapid approaches to research and evaluation. Eval Rev 2025;49(6):951-968 [FREE Full text] [doi: 10.1177/0193841X251338536] [Medline: 40286216]
75. Nevedal AL, Reardon CM, Opra Widerquist MA, Jackson GL, Cutrona SL, White BS, et al. Rapid versus traditional qualitative analysis using the consolidated framework for implementation research (CFIR). Implement Sci 2021;16(1):67 [FREE Full text] [doi: 10.1186/s13012-021-01111-5] [Medline: 34215286]
76. Keel G, Savage C, Rafiq M, Mazzocato P. Time-driven activity-based costing in health care: a systematic review of the literature. Health Policy 2017;121(7):755-763 [FREE Full text] [doi: 10.1016/j.healthpol.2017.04.013] [Medline: 28535996]
77. McBain RK, Jerome G, Warsh J, Browning M, Mistry B, Faure PAI, et al. Rethinking the cost of healthcare in low-resource settings: the value of time-driven activity-based costing. BMJ Glob Health 2016;1(3):e000134 [FREE Full text] [doi: 10.1136/bmjgh-2016-000134] [Medline: 28588971]
78. Huebschmann AG, Trinkley KE, Gritz M, Glasgow RE. Pragmatic considerations and approaches for measuring staff time as an implementation cost in health systems and clinics: key issues and applied examples. Implement Sci Commun 2022;3(1):44 [FREE Full text] [doi: 10.1186/s43058-022-00292-4] [Medline: 35428326]
79. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687 [FREE Full text] [doi: 10.1136/bmj.g1687] [Medline: 24609605]
80. Bragge P, Grimshaw JM, Lokker C, Colquhoun H, AIMD Writing/Working Group. AIMD - a validated, simplified framework of interventions to promote and integrate evidence into health practices, systems, and policies. BMC Med Res Methodol 2017;17(1):38 [FREE Full text] [doi: 10.1186/s12874-017-0314-8] [Medline: 28259155]
81. Chen WW, Cato BM, Rainford N. Using a logic model to plan and evaluate a community intervention program: a case study. Int Q Community Health Educ 1999;18(4):449-458. [doi: 10.2190/jdnm-mnpb-9p25-17cq]
82. Renger R, Titcomb A. A three-step approach to teaching logic models. Am J Eval 2002;23(4):493-503. [doi: 10.1177/109821400202300409]
83. Michel JP, Leonardi M, Martin M, Prina M. WHO's report for the decade of healthy ageing 2021-30 sets the stage for globally comparable data on healthy ageing. Lancet Healthy Longev 2021;2(3):e121-e122 [FREE Full text] [doi: 10.1016/S2666-7568(21)00002-7] [Medline: 36098109]
84. Bevilacqua R, Soraci L, Stara V, Riccardi GR, Corsonello A, Pelliccioni G, et al. A systematic review of multidomain and lifestyle interventions to support the intrinsic capacity of the older population. Front Med (Lausanne) 2022;9:929261 [FREE Full text] [doi: 10.3389/fmed.2022.929261] [Medline: 35911409]
85. Leung AYM, Su JJ, Lee ESH, Fung JTS, Molassiotis A. Intrinsic capacity of older people in the community using WHO integrated care for older people (ICOPE) framework: a cross-sectional study. BMC Geriatr 2022;22(1):304 [FREE Full text] [doi: 10.1186/s12877-022-02980-1] [Medline: 35395736]
86. Meng LC, Hsiao FY, Huang ST, Lu WH, Peng LN, Chen LK. Intrinsic capacity impairment patterns and their associations with unfavorable medication utilization: a nationwide population-based study of 37,993 community-dwelling older adults. J Nutr Health Aging 2022;26(10):918-925 [FREE Full text] [doi: 10.1007/s12603-022-1847-z] [Medline: 36259580]
87. de Oliveira VP, Ferriolli E, Lourenço RA, González-Bautista E, de Souto Barreto P, de Mello RGB. The sensitivity and specificity of the WHO's ICOPE screening tool, and the prevalence of loss of intrinsic capacity in older adults: a scoping review. Maturitas 2023;177:107818. [doi: 10.1016/j.maturitas.2023.107818] [Medline: 37542782]
88. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
89. Roche N, Reddel H, Martin R, Brusselle G, Papi A, Thomas M, Respiratory Effectiveness Group. Quality standards for real-world research. Focus on observational database studies of comparative effectiveness. Ann Am Thorac Soc 2014;11(Suppl 2):S99-S104. [doi: 10.1513/AnnalsATS.201309-300RM] [Medline: 24559028]
90. Tibbits MK, Bumbarger BK, Kyler SJ, Perkins DF. Sustaining evidence-based interventions under real-world conditions: results from a large-scale diffusion project. Prev Sci 2010;11(3):252-262 [FREE Full text] [doi: 10.1007/s11121-010-0170-9] [Medline: 20229358]

Abbreviations

AMID: Aims, Ingredients, Mechanism, and Delivery
CFIR: Consolidated Framework for Implementation Research
CFS: Clinical Frailty Scale
CGA: comprehensive geriatric assessment
COM-B: Capability, Opportunity, Motivation, Behavior
ERIC: Expert Recommendations for Implementing Change
FAID: Framework of Actions for Intervention Development
IC: intrinsic capacity

XSL•FO
RenderX

ICOPE: integrated care for older people
IMPACTFrail: Intrinsic Capacity Promotion in Primary Care for the Frail
MOH: Ministry of Health
MRC: Medical Research Council
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TDF: Theoretical Domains Framework
TIDieR: Template for Intervention Description and Replication
WHO: World Health Organization

Edited by J Sarvestan; submitted 19.Sep.2025; peer-reviewed by P Vatcharavongvan, F Riordan; comments to author 10.Dec.2025; revised version received 05.Jan.2026; accepted 05.Jan.2026; published 02.Feb.2026.

Please cite as:
Wang Q, Ginting ML, Ho E, Goh SF, Gao J, Tan WS, Ding YY, Ng WLD, Ng JCA, Kwek SC, Hui RJY, Wang ZS, Ng CJ, Sum G, The IMPACTFrail Group
Development and Feasibility Assessment of an Intrinsic Capacity Program in Primary Care: Protocol for an Implementation Science Approach
JMIR Res Protoc 2026;15:e84257
URL: https://www.researchprotocols.org/2026/1/e84257
doi:10.2196/84257
PMID:

©Qing Wang, Mimaika Luluina Ginting, Ezra Ho, Siew Fong Goh, Jonathan Gao, Woan Shin Tan, Yew Yoong Ding, Wei Liang David Ng, Jonathan Choon Aik Ng, Sing Cheer Kwek, Richard Jor Yeong Hui, Zhen Sinead Wang, Chirk Jenn Ng, Grace Sum, The IMPACTFrail Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 02.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Contribution of Emotion Dynamics to Adolescent Psychosocial Well-Being: Protocol for a Longitudinal Study

Carli Mastronardi[1], MA; Jade Powers[1], BA; Rosanne Menna[1], PhD; Lance M Rappaport[1], PhD

Department of Psychology, University of Windsor, Windsor, ON, Canada

Corresponding Author:
Lance M Rappaport, PhD
Department of Psychology
University of Windsor
401 Sunset Avenue
Windsor, ON, N9B 3P4
Canada
Phone: 1 519 253 3000 ext 2293
Email: Lance.Rappaport@uwindsor.ca

Abstract

Background: As a critical period in psychosocial development, adolescence is marked by heightened emotion regulation demands as well as increased risk for, and vulnerability to, stress.

Objective: This longitudinal study investigates how dynamic patterns (ie, mean intensity, variability, instability, inertia, and reactivity to stress) in positive and negative affect relate to, and predict change in, broad domains of adolescent psychosocial well-being (ie, mental health, social well-being, and academic motivation). Using a daily diary procedure to capture adolescents' daily naturalistic affective experiences, this study will provide novel insights into how affective processes predict psychosocial well-being over time beyond traditional, static assessment.

Methods: At baseline, adolescents aged 14-17 years from Southwestern Ontario reported on their academic motivation (eg, extrinsic motivation), social well-being (eg, social support and loneliness), and mental health (eg, anxiety syndrome severity) before completing a 35-day smartphone-based daily diary protocol wherein participants reported twice daily in the morning (ie, 7-10 AM) and evening (ie, 8-11 PM) on positive and negative affect, stress, and internalizing symptom severity. Participants then repeat the surveys of academic motivation, social well-being, and mental health 6, 12, and 18 months following baseline assessment to assess change in each domain of psychosocial well-being over time.

Results: Adolescents (N=149) were enrolled into this longitudinal study between April 2023 and November 2024, such that all participants will complete the scheduled 18 months of longitudinal follow-up assessments by May 2026. Primary study analyses will use multilevel modeling, structural equation modeling, multilevel structural equation modeling, and dynamic structural equation modeling to examine how dynamic patterns in positive and negative affect (eg, instability, inertia) concurrently correlate with, and prospectively predict change in, psychopathology and well-being.

Conclusions: This study protocol paper outlines the overarching study objectives and methodology to promote transparency and reproducibility. Through the integration of daily diary methodology within a longitudinal design, this study aims to clarify the potential implications of dynamic affective processing (eg, affective reactivity to daily stress) for both adolescent psychopathology and well-being beyond clinical syndromes.

International Registered Report Identifier (IRRID): DERR1-10.2196/76333

(JMIR Res Protoc 2026;15:e76333) doi:10.2196/76333

KEYWORDS

emotion dynamics; daily diary; psychosocial well-being; adolescence; academic motivation; internalizing syndromes

XSL•FO
RenderX

Introduction

Background

Adolescence presents a critical period for psychosocial and socioemotional development including the development and emergence of many internalizing syndromes [1,2]. For example, epidemiological estimates suggest that 18%-22% of adolescents recently experienced psychopathology or related psychosocial distress [3-5], and psychiatric conditions often emerge during adolescence and then persist into adulthood [6,7]. Adolescent psychosocial distress also impairs multiple areas of functioning including social interactions, academic performance [8], and overall well-being [9]. While longstanding theory documents acute affective distress in adolescence [10], psychosocial distress and related psychopathology may have become particularly prevalent and severe in recent years. For example, the prevalence of internalizing psychopathology rose markedly from 2007 to 2017 [11] and from 1990 to 2023 [12,13]. Similarly, youth visits to emergency departments for mental health–related distress increased substantially from 2009 to 2015 [14].

Historically, research on child and adolescent mental health focused on stable factors that predict psychopathology, such as personality traits [15], temperament [16], and attachment styles [17]. However, despite their theoretical stability and strong test-retest reliability [18,19], each of these seminal theories also encompasses constructs of variability or instability, such as impulsiveness [20], negative affectivity [21,22], or reactivity [23], that confer risk to develop psychopathology. These theories of psychopathology risk suggest that both stable traits and dynamic fluctuations, such as change in response to the demands of one's environment, may provide unique and important insights into adult well-being and adolescent psychosocial and socioemotional development [24-26]. For example, emotional reactivity to stressful events is associated with psychopathology (eg, depression and anxiety) in adolescence [27] and early adulthood [28]. Similarly, depressive symptoms in adults have been implicated in elevated mean quarrelsome and submissive behavior [29], as well as reduced responsiveness of interpersonal behavior to affective [30] and situational cues [31] that typically prompt one to adapt their behavior to the situation [32,33]. However, while considerable extant research examined stable correlates of psychopathology risk in adults and adolescents, there is a paucity of research to examine naturalistic dynamics in adolescence that reflect how adolescents adapt to environmental demands, such as daily stressful experiences. Hence, further research is needed to explore the potential implications of naturalistic dynamics (eg, in emotions) to offer important insights into adolescent psychosocial and socioemotional development above and beyond traditional static, stable indictors of risk.

Extant research on in situ emotional experiences suggests that affective states are inherently dynamic and fluctuate over time (see Kuppens et al [34] for review). Many theories of emotion emphasize the role of acute fluctuations in momentary positive and negative affect to help one (1) locate suitable environments, (2) attend to changes in the environment, (3) appraise the situation, and (4) choose effective compensatory, regulatory behaviors with the goal of returning to emotional equilibrium (see Frijda [35], Gross and Thompson [36], and Scherer [37] for review). Dynamics in one's affective experiences over time may summarize data on a set of momentary affective experiences, such as affective instability or variability over time, to index affective difficulties or dysregulation [38,39]. For example, affective variability over time and instability from one occasion to the next may predict various clinical outcomes (eg, depressive symptoms) and related indicators of psychosocial well-being [24,26].

However, most research on the clinical implications of in situ affective fluctuations and related emotion dynamics focused on adulthood [24]. Less research examined the potential implications of in situ affective dynamics in adolescence [40]. The available evidence suggests that, relative to adults, adolescents may report more intense and variable affective fluctuations [41,42] consistent with heightened challenges during adolescence [10,43]. For example, relative to adults, adolescents reported greater positive and negative affective instability and heightened affective intensity, which may decrease over time as they reach adulthood [41,44]. Given that adolescence may pose a novel challenge for youth to manage more intense, variable, and dynamic emotions [10,41,45], research is critically needed to understand the manifestation of affective dynamics in adolescence and its implications for adolescent well-being including both psychopathology and psychosocial development.

Technological and statistical innovations (eg, the proliferation of smartphones) facilitate the repeated assessment of affect as it manifests naturalistically during adolescents' daily lives. Referred to as experience sampling methodologies, including daily diary procedures, this class of methodology facilitates assessment of affect on the temporal scale required to accurately test many psychological theories about well-being (see Shiffman et al [46] for review). For example, per the self-medication theory, individuals may engage in high-risk alcohol use to reduce momentary affective distress [47]. However, past research largely relied on cross-sectional or longitudinal evidence of high-risk alcohol use among people *who* experience distress (eg, anxiety disorders; see Turner et al [48] for review). In contrast, experience sampling methodology research identified that acute, severe negative affect may predict *when* at-risk young adults engage in high-risk alcohol use [49,50]. Experience sampling methods, such as daily diary procedures, also naturalistically assess participants' current or recent affect, stressful experiences, and social context, which minimize recall biases typically associated with self-report (eg, by minimizing participants' recall time) and provide greater ecological validity than laboratory-based tasks (eg, through the assessment of events in situ) [51,52].

Emotion Dynamics

Emotion dynamics characterize the trajectories, patterns, and regularities of how positive and negative affect vary over time [24,53]. Experience sampling and daily diary methods [54] previously identified 4 dynamics that have been the focus of most extant research: mean affective intensity, variability, instability, and inertia [26]. Mean intensity of positive and negative affect refers to the average level of positive and

XSL•FO
RenderX

negative affect that a participant reported over a set of records completed over a specified period of time (eg, 30 days). Specifically, mean intensity is typically computed separately for positive and negative affect as the mean of all assessments of positive or negative affect provided. Among adults, elevated mean negative affect is consistently associated with psychopathology [24] including depressive and anxiety disorders [55,56]. During adolescence, mean positive and negative affect may be higher [44] and correlated with psychopathology [57,58].

Affective variability and instability capture intraindividual (ie, within-person) fluctuations in positive and negative affect around one's mean level, such as variation over moments or days. Affective variability and instability are typically computed separately for positive and negative affect. Specifically, affective variability is computed as the intraindividual standard deviation while affective instability is computed as the mean squared successive difference (MSSD) between consecutive assessments adjusted for the average time between consecutive assessments [59]. Hence, high variability in negative affect suggests that one's affective experience varies greatly over the occasions studied whereas low affective variability suggests that one's affective experience was consistently close to one's mean level. Similarly, high instability in negative affect suggests frequent large changes in negative affect from one occasion to the next whereas low instability in negative affect would suggest that one's negative affect is relatively static from one occasion to the next. Empirically in adults, instability and variability have been consistently, albeit modestly, associated with lower well-being and psychopathology [24,26,60]. Extensive prior research also consistently implicates affective variability and instability in adolescent psychopathology [61-63]. Given evidence of elevated affective variability and instability in adolescence [10,64], further longitudinal research is needed to examine whether affective variability or instability is associated with change in psychopathology over time. Meanwhile, further research is needed in adolescence to understand the implications of affective variability and instability for other facets of well-being during adolescent psychosocial development (eg, academic motivation).

Like the MSSD, inertia in positive or negative affect reflects the degree to which positive or negative affect at the last assessment (eg, moment or day) persists to influence one's current positive or negative affect [65]. However, whereas the MSSD reflects deviations from the autocorrelation of positive and negative affect over time as affective instability, affective inertia directly assesses the autocorrelation as persistence of positive or negative affect over time. Hence, while conceptually similar, extant research distinguishes between MSSD and inertia, which are moderately correlated [24,66]. Specifically, one can regress each assessment of positive or negative affect on the immediately preceding assessment through multilevel, mixed-effects regression [65] or dynamic structural equation modeling (DSEM) and residual dynamic structural equation modeling [67-69], which has several benefits including latent variable centering [70] and the ability to adjust for possibly unique responding early in a daily diary procedure [71]. Conceptually, elevated inertia in negative affect may indicate difficulty downregulating (ie, reducing) negative affect, which could reflect persistent negative affect that contributes to, or results from, frequent or particularly severe distress [72]. Similarly, although somewhat less emphasized in extant research, reduced inertia in positive affect could reflect difficulty maintaining positive affect. Prior research documented an association of elevated inertia in negative affect with depressive disorders in adulthood [73,74] (see Thompson et al [75] for review), adolescence [76], and childhood [77]. Some extant research examined implications of inertia in positive or negative affect for other psychopathological syndromes among adults (see Dejonckheere et al [24] for review). However, little prior research examined the manifestations of inertia in positive or negative affect during the significant emotional development represented by adolescence or the implications of inertia in positive or negative affect for other internalizing syndromes (eg, anxiety disorders) and facets of psychosocial well-being during adolescence. Moreover, like research on affective variability and instability, longitudinal research is needed to examine whether inertia in positive or negative affect predicts change in psychopathology or other indices of well-being over time.

Beyond the 4 emotion dynamics described thus far, substantial extant research emphasizes the implications of affective reactivity to stress for adult and adolescent psychopathology and well-being. Broadly, reactivity to stress refers to the association of a stressor with an affective, behavioral, or physiological change. For example, using multilevel mixed-effects regression, DSEM, or residual dynamic structural equation modeling, one can regress current positive or negative affect on positive or negative affect at the preceding assessment to assess change over the given time interval as well as regress positive or negative affect on the degree of stress experienced during said interval to assess the association of stress with concurrent change in positive or negative affect. Moreover, the association of stress with positive or negative affect can be examined as both fixed and random effects to evaluate the average impact of stress on positive or negative affect as well as interindividual (ie, between-person) heterogeneity in, and moderation of, the strength of one's reactivity of positive or negative affect to stressful experiences. Extant research on the association of affective reactivity to stress with psychopathology is mixed. Laboratory-based research largely suggests that depressive symptomatology is associated with blunted reactivity to stress (see Bylsma et al [78] for review) whereas naturalistic research (eg, daily diary and ecological momentary assessment methods) suggests that depressive symptomology in adulthood [73,79] and adolescence [80,81] may be associated with elevated affective reactivity to stress. While future research is needed to clarify this apparent discrepancy, additional research is also needed to evaluate whether affective reactivity to stress is specifically associated with depressive disorders relative to other internalizing syndromes [82], particularly in adolescence [83]. Further research is also needed to evaluate the potential implications of affective reactivity to stress prospectively on the development of internalizing psychopathology and psychosocial distress during the high-risk period posed by adolescence [11,84].

XSL•FO
RenderX

This Study

As a period of significant socioemotional development, adolescence is marked by heightened vulnerability to psychosocial distress (eg, internalizing syndromes) and increased affective instability [42,64]. Extant research delineates characteristics of fluctuations in naturalistic affective experiences (ie, mean intensity, variability, instability, inertia, and reactivity to stress) that may correlate with well-being and psychopathology in adults and youth [26,39,45]. However, extant research focused on the implications of putative emotion dynamics (eg, inertia) for psychopathology in adults (see Dejonckheere et al [24] for review). Given the substantial affective development documented during adolescence [10], further research is needed to examine the manifestation and implications of putative emotion dynamics on psychopathology during adolescence [45]. Additionally, although most prior research focused on 5 core emotion dynamics (ie, mean intensity, variability, instability, inertia, and reactivity to stress), future research can extend beyond these indices to examine additional dynamics (eg, emotion-network density [24]).

Moreover, given the substantial psychosocial development documented during adolescence [85], further research is needed to examine the implications of emotion dynamics for adolescent well-being above and beyond psychopathology. Specifically, this study will examine how 5 putative emotion dynamics (ie, mean affective intensity, variability, instability, inertia, and reactivity to stress) correlate with and predict change in adolescent well-being (ie, social well-being, academic motivation, and mental health) over 18 months. Specifically, this study will evaluate hypotheses that elevated mean negative affect, lower mean positive affect, elevated variability and instability in positive and negative affect, elevated inertia in negative affect, lower inertia in positive affect, and elevated reactivity of positive and negative affect to stress will contemporaneously correlate with elevated adolescent psychopathology and prospectively predict increased adolescent psychopathology (eg, depressive and anxiety syndrome severity) over 18 months. For example, this study will evaluate the contemporaneous and prospective association of baseline affective variability, instability [26,31], and reactivity of positive or negative affect to stress [78,83] with elevated psychopathology in adolescence. Similarly, this study will examine whether inertia in positive or negative affect is uniquely associated with elevated depressive symptomatology relative to other psychopathological syndromes (eg, anxiety and eating-related symptomatology [65,72]).

Finally, this study will evaluate hypotheses that elevated mean negative affect, lower mean positive affect, elevated variability and instability in positive and negative affect, elevated inertia in negative affect, lower inertia in positive affect, and elevated reactivity of positive and negative affect to stress will contemporaneously correlate with worsened psychosocial well-being and prospectively predict worsened adolescent psychosocial well-being (eg, academic motivation and parent and peer relationship quality) above and beyond joint associations with psychopathology. For example, this study will evaluate hypotheses that each emotion dynamic (eg, elevated mean negative affect intensity and lower mean positive affect intensity) is correlated with and prospectively predicts worsened social well-being (eg, increased loneliness) and academic motivation (eg, increased amotivation) above and beyond joint associations with psychopathology (eg, depressive symptomatology).

This protocol paper serves to provide a comprehensive record to increase transparency and persistent documentation of the study methodology. For example, the resulting dataset may support additional, secondary projects (eg, on the intraindividual dynamics of change in psychopathology and social relationships). In these cases, this paper serves as a comprehensive account of the study purpose and of all study measures and procedures. Similarly, this protocol paper serves to clearly articulate and register a comprehensive overview of the study aims, background, and methods prior to the completion of data collection. Finally, this protocol paper provides an example framework for longitudinal research to examine the naturalistic manifestation and implications of socioemotional correlates and predictors (eg, emotion dynamics) of adolescent psychopathology and adolescent well-being above and beyond concurrent psychopathology.

Methods

Project Overview

This study uses a longitudinal design to examine the concurrent and prospective association of 5 putative emotion dynamics (ie, affective mean intensity, variability, instability, inertia, and reactivity to stress) with psychosocial well-being in adolescence. Specifically, this study explores how each dynamic correlates with and predicts change in psychopathology (eg, depressive symptomatology) as well as life satisfaction, social well-being, and academic motivation above and beyond psychopathology. Using daily diary assessment, this study assesses the naturalistic manifestation of positive and negative affect, stress, and distress during adolescence to clarify the in situ fluctuation of affect during adolescents' daily lives and the potential implications of dynamic affective fluctuations for adolescents' development.

Recruitment

Adolescent participants aged 14-17 years were recruited from the Windsor-Essex region of Southwestern Ontario from April through May 2023, September through December 2023, February through May 2024, and September through November 2024. Recruitment information was shared with adolescents in the local community through multiple channels to increase the opportunity for adolescents to participate and, thereby, increase the representativeness and generalizability of the resulting sample. Specifically, recruitment information was shared via advertisements on prominent social media (eg, Facebook and Instagram), flyers posted in local community centers, and information provided to students and families by both local school boards. For example, recruitment flyers were (1) posted in schools, (2) sent out to families via email or online message, and (3) shared on school social media accounts. Participants were eligible for the study if they were aged 14-17 years and had access to a device running the iOS or Android OS operating system. There were no explicit exclusion criteria to maximize

XSL•FO
RenderX

sample heterogeneity and generalizability by allowing a broad range of adolescents to participate.

This study initially planned to recruit 200 participants to provide 80% statistical power to detect moderate correlations at baseline using a cautious testwise a value of .005 (r=0.253) [86]. According to a Monte Carlo simulation of a latent change score model, which fixed a value at .05 [87] and accounted for 20% attrition, the intended sample would have also provided 80% statistical power to detect a moderate prospective association (β=.33) of each emotion dynamic (eg, inertia) with change in each well-being domain over each 6-month interval. However, participant enrolment was ended in November 2024 so that all participants could complete the scheduled 18-month follow-up assessment before the end of the project funding in May 2026.

Ethical Considerations

Study procedures were cleared by the University of Windsor Research Ethics Board (number 21-098) on November 8, 2022. During an initial, in-person baseline session, a parent or legal guardian provided informed consent for each adolescent to participate in the study. Each adolescent participant then provided separate informed consent prior to participation in the study. Participants will receive financial compensation up to $140 for time spent participating in each component of the study. Specifically, based on the local minimum wage, each participant received $23.50 for up to 90 minutes spent completing the in-person baseline assessment, received up to $70 based on the number of daily diary surveys completed for time spent completing daily diary surveys, and will receive $15.50 per follow-up survey completed for up to 60 minutes completing each follow-up assessment. Where applicable, each participant also received $10 at the baseline assessment to offset the cost of parking at the university. Data are stored securely through assignment of an arbitrary, person-specific alphanumeric code that facilitates linkage of participant data across study components. Following the end of data collection, data will be deidentified to support secondary use of study data.

Design and Procedure

This study uses a longitudinal design to examine concurrent correlates of, and prospective predictors of change in, adolescents' psychosocial well-being over 18 months (Figure 1). Specifically, each participant was or will be asked to complete measures of social well-being (eg, loneliness and social support), academic motivation, and psychopathology at baseline as well as 6, 12, and 18 months following baseline assessment. For the 35 days immediately following baseline assessment, each participant was also asked to complete a 35-day daily dairy procedure to assess naturalistic dynamics in positive and negative affect (eg, inertia in positive affect and reactivity of negative affect to daily stress).

Figure 1. Study timeline and schedule.

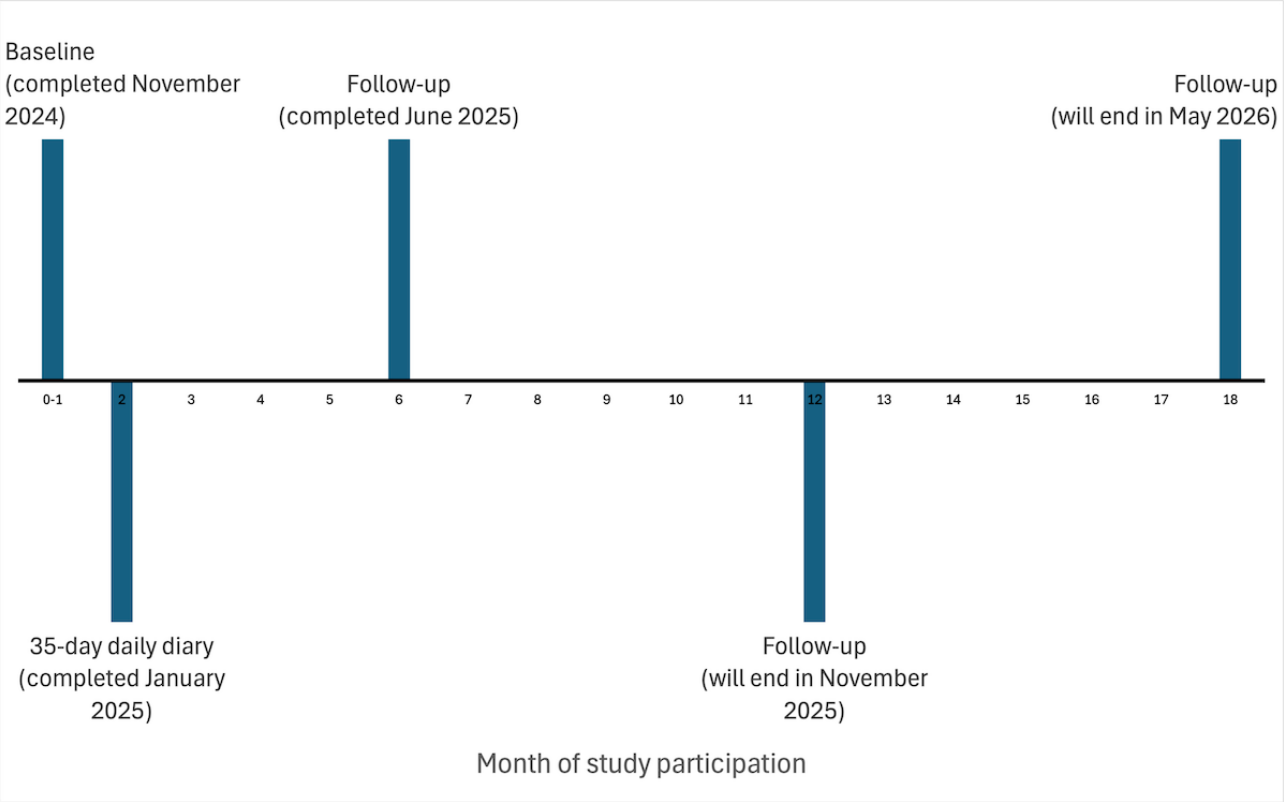

Specifically, following the provision of informed consent, each adolescent participant completed a baseline assessment, during which they completed online questionnaires that assess baseline social well-being, academic motivation, and mental health, and received instruction in the following daily diary procedure. A subset of participants also completed the Mini International Neuropsychiatric Interview for Children and Adolescents [MINI-KID], a structured interview used to assess risk for common psychopathological syndromes (see "Measures" section). Following baseline assessment, all participants completed a daily diary procedure for the next 35 days to assess naturalistic emotion dynamics, such as mean intensity of, and

XSL•FO
RenderX

variability in, positive and negative affect. During this period, participants reported twice daily (ie, 7-10 AM and 8-11 PM) on positive and negative affect, anxiety and depressive symptoms, and stress experienced since waking. Thus, the timing of each assessment in the morning and evening assesses positive and negative affect, anxiety and depressive symptoms, and stress experienced immediately upon waking and over the course of each day, respectively. Three hours were provided for participants to complete each record (eg, 7-10 AM) to support their engagement in the daily diary procedure; variability in response timing will be modeled empirically, such as through adjustment for the interval between consecutive assessments in the computation of MSSD [24] or lagged associations in dynamical structural equation models [67]. To address overlap in the time period covered by morning and evening assessments, evening reports serve as the main measure to capture day-to-day change in affective experiences. Morning assessments were included to examine affective change over the course of each day (see "Statistical Analyses" section). For example, morning assessments facilitate examination of (1) persistence or change in positive and negative affect over the course of each day and (2) persistence or change in positive and negative affect from one day to the next morning. Finally, participants will complete scheduled longitudinal assessments online 6, 12, and 18 months after baseline assessment. Due to the schedule of participant recruitment, longitudinal follow-up assessments will end in May 2026.

Measures

Baseline Assessment

Participants provided brief demographic information including age, biological sex, racial background, current grade level, each parents' level of education, number of people living in the household, and household income. Participants also reported on objective indicators of, and subjective perception of, socioeconomic status. Specifically, participants completed the Family Affluence Scale [88] on which they reported (1) the number of computers at home; (2) whether their family owns a car, van, or truck; (3) whether they have their own bedroom; (4) the number of family holidays taken last year; (5) whether their family has a dishwasher; and (6) the number of bathrooms in their home. Prior studies demonstrate the test-retest reliability, convergent validity, construct validity, internal consistency, and external validity of the Family Affluence Scale in adolescent samples [88]. Similarly, participants reported their perceived socioeconomic status relative to society at large and relative to their school on a ladder from 1 (low) to 10 (high) via the MacArthur Scale of Subjective Social Status—Youth Version [89]. Prior research in adolescent samples demonstrated internal consistency, cross-informant correlation, and test-retest stability of this measure over 2-6 months [89,90].

Academic Motivation

During baseline assessment, participants completed the high school version of the 28-item Academic Motivation Scale (Echelle de Motivation en Education) [91] to assess participants' intrinsic and extrinsic motivation to engage with school including amotivation. Specifically, participants rated their agreement with each item on a 7-point Likert scale ranging from 1 (*does not correspond at all*) to 7 (*corresponds exactly*). Prior research on adolescent samples demonstrates construct validity, internal consistency, test-retest reliability, and discriminant validity of this measure [92].

Social Well-Being

Social well-being is indexed by adolescent report of perceived loneliness, social support, and relationship satisfaction. Specifically, participants indicated the severity of perceived loneliness through endorsement of 20 items on the UCLA Loneliness Scale version 3 [93] using a 4-point Likert rating scale from 1 (*never*) to 4 (*always*). Prior research demonstrates high internal reliability, construct validity, and convergent validity of the full and shortened measure to assess loneliness in adolescent populations [94].

Participants reported on the perceived availability of, and satisfaction with, social support via the 6-item version of the Social Support Questionnaire (SSQ) [95]. Specifically, for each of 6 situations, participants were asked to list all the people who may provide support and then to rate their satisfaction with the support available from 1 (*very dissatisfied*) to 6 (*very satisfied*). Previous research demonstrated the strong internal reliability, test-retest reliability, and construct validity of both the full and 6-item short versions of the SSQ in adolescent and undergraduate samples [95,96].

Finally, participants rated 75 items about their interpersonal relationships to assess overall attachment, trust, communication, and alienation in their relationship with their mother (or a mother figure), father (or a father figure), and peers on the Inventory of Parent and Peer Attachment (IPPA) [97]. Specifically, participants rated the accuracy of each item from 1 (*almost never or never true*) to 5 (*almost always or always true*). Prior research demonstrates strong internal consistency, test-retest reliability, and convergent validity of the IPPA to assess relationship quality in adolescent samples [97,98].

Mental Health

During baseline assessment, participants reported on their mental health as indexed by life satisfaction, perceived stress, difficulties in emotion regulation, broad psychosocial difficulties, conduct problems, hyperactivity- or inattention-related problems, peer relationship–related problems, prosocial behavior, and psychopathology including irritability, anxiety, depressive, and eating disorder syndromes. Additionally, a subset of participants completed the MINI-KID [99] to provide a categorical assessment of potential psychopathology.

Specifically, adolescents reported on life satisfaction by indicating agreement from 1 (*strongly disagree*) to 7 (*strongly agree*) with 5 statements about their overall quality of life or life satisfaction on the Satisfaction with Life Scale [100]. Prior research demonstrates reliability (ie, interitem reliability and internal consistency reliability) and validity (ie, construct validity, convergent validity, and criterion validity) of this measure to assess life satisfaction in adolescent samples [101].

Participants also indicated the severity of, and perceived ability to manage, stress on the 10-item version of the Perceived Stress

XSL•FO
RenderX

Scale [102], which asked participants to rate the frequency with which they have been upset or had difficulty managing stress over the past month on 10 items rated from 0 (*never*) to 4 (*very often*). Prior research documented the convergent validity, internal consistency, criterion validity, and construct validity of the Perceived Stress Scale in adolescent samples [103].

Similarly, participants completed the Difficulties in Emotion Regulation Scale [104], which asked participants to rate the frequency with which they experience difficulties related to distressing emotions on 36 items from 1 (*almost always*) to 5 (*almost never*). Prior research demonstrates the interitem reliability and construct validity of the Difficulties in Emotion Regulation Scale in adolescent samples [105].

Adolescents completed the Strengths and Difficulties Questionnaire (SDQ) [106] to broadly assess psychosocial difficulties, conduct problems, hyperactivity- or inattention-related problems, peer relationship–related problems, prosocial behavior, and emotional distress over the past 6 months. Each participant was specifically asked to rate 25 items as *not true*, *somewhat true*, or *certainly true* and asked 5 questions to assess the impact of any difficulties reported (eg, the degree to which reported difficulties interfere at home or in the classroom). Prior research documented the interitem reliability, test-retest reliability, and construct validity of the SDQ to assess broad and specific domains of psychosocial difficulties in adolescent samples [107,108].

Participants completed the 13-item short version of the Mood and Feelings Questionnaire [109] to assess the severity of depressive symptoms over the past 2 weeks. Specifically, adolescents rated whether they experienced 13 depressive symptoms as *not true* (0), *sometimes* (1), or *true* (2). Prior research on the short version of the Mood and Feelings Questionnaire documented interitem reliability, internal consistency, test-retest reliability, criterion validity, convergent validity, and discriminative validity in adolescent samples [110].

Participants completed the 41-item Screen for Child Anxiety Related Emotional Disorders [111] to assess the severity of school avoidance and 4 anxiety syndromes common in childhood and adolescence: generalized anxiety, social anxiety or phobia, panic or somatization, and separation anxiety. Specifically, adolescents rated their agreement with 41 statements (eg, the ease with which they experience panic or somatization symptoms) from 0 (*not true or hardly ever true*) to 2 (*very true or often true*). The measure was modified slightly to assess the severity of anxiety symptoms over the past 2 weeks to align with participants' concurrent assessment of depressive symptoms and capture change over time as suggested by Albano and colleagues [112] and Caporino and colleagues [113]. Prior research demonstrated strong psychometric properties of this measure to assess the severity of these 5 anxiety syndromes in adolescent samples including interitem reliability, test-retest reliability, concurrent validity, and discriminative validity [114] including over the past 2 weeks [112,113].

Participants also reported on the severity of irritability symptoms they experience via the Affective Reactivity Index (ARI) [115], which asks participants to rate agreement with 6 items that describe the ease with which they become angry or stay angry from 0 (Not True) to 2 (Certainly True). Prior research demonstrated strong psychometric properties of the ARI to assess irritability in adolescent samples including interitem reliability and convergent validity [115,116].

Participants reported on broad eating disorder symptoms (eg, body dissatisfaction and disordered eating behavior) via endorsement of 26 statements on the Eating Attitudes Test (EAT) [117] from 0 (*never*) to 3 (*always*), reported the frequency with which they engaged in 4 disordered eating behaviors (eg, binge eating) from *never* to *once a day or more*, and reported whether or not they lost 20 lb or more in the past 6 months. Prior research documents the interitem reliability, test-retest reliability, and convergent validity of the EAT-26 in adolescent samples [118].

The MINI-KID [99] was administered to assess categorical risk for various psychopathological conditions (eg, major depressive disorder, generalized anxiety disorder, and bulimia nervosa) excluding the following conditions due to relatively low prevalence in a community sample of adolescents [119-121]: tic disorders, psychotic disorders, and autism spectrum disorders. Prior research demonstrated the test-retest reliability, convergent validity, and discriminant validity of the MINI-KID in adolescent samples [99,122]. The MINI-KID was administered to a subset of 54 participants before it was discontinued due to time constraints for the full baseline assessment. Analyses involving the MINI-KID will be limited to participants who completed the interview.

Daily Diary Assessment

Twice daily for 35 days, participants completed brief surveys of positive and negative affect, depressive and anxiety symptoms, and stress on an iOS or Android OS device through the RealLife Exp app (LifeData LLC). Specifically, participants reported on positive and negative affect via completion of the Positive and Negative Affect Schedule—Child Form [123], on which they rated their experience of 30 emotions (eg, excited) since waking that morning from 1 (*not much or not at all*) to 5 (*a lot*). Prior research documented interitem reliability, internal consistency, convergent validity, and divergent validity of the Positive and Negative Affect Schedule—Child Form in adolescent samples [124-127].

Participants also reported twice daily on stressful experiences encountered since the last assessment including the context (ie, whether the stress is related to school, health, an interpersonal relationship, or general stress, such as the weather), severity, and duration of stress on 7 items adapted for adolescents from items previously developed by Dunkley and colleagues [128]. Specifically, adolescents' experience of daily stress will be derived from 3 items that asked participants to reflect on the most bothersome event or issue experienced since the last assessment and then rate how unpleasant the event was from 1 (*not at all*) to 11 (*exceptionally*), how long distress lasted following the event from 1 (*a very brief amount of time*) to 7 (*a very large amount of time*), and how stressful the event was from 1 (*not at all)* to 11 (*exceptionally*). While Mandel and colleagues demonstrated internal consistency and validity of the measure in young adults [129], this study will be the first

XSL•FO
RenderX

to examine the psychometric properties of this assessment of stress in the daily lives of adolescents.

Finally, participants rated depressive and anxiety symptoms twice daily via completion of symptoms drawn from the Dynamic Assessment of Depressive and Anxiety Symptoms [LM Rappaport et al, unpublished data, April 2025]. Specifically, regarding depressive symptoms, participants rated "hopelessness," "worthlessness," "sadness or depressed mood," "feeling down on yourself," "a feeling that things will not improve," and anhedonia ("less interest in things you used to enjoy") on an 11-point scale from 0 (*not at all*) to 10 (*severely*). Similarly, regarding anxiety symptoms, participants rated "nervousness," "restlessness," "difficulty relaxing," "worry," "trouble calming worry," "difficulty breathing or shortness of breath (not due to exercise)," "trembling or shaking," "heart pounding or racing," "dizziness," "tingling or numbness," and "feeling jumpy" on an 11-point scale from 0 (*not at all*) to 10 (*severely*). This study will provide the first psychometric data to assess naturalistic variation in anxiety and depressive symptoms in adolescents' daily lives.

Follow-Up Assessments

Follow-up assessments include demographic information pertaining to age and grade. At each follow-up assessment, participants will report on academic motivation via the high school version of the Academic Motivation Scale [91]. Participants will also report on social well-being, specifically loneliness via the UCLA Loneliness Scale [93]; social support via the 6-item version of the SSQ [95]; and the quality of interpersonal relationships with a father figure, mother figure, and peers via the IPPA [97]. Finally, participants will report on mental health, specifically broad life satisfaction via the Satisfaction With Life Scale [100], irritability via the ARI [115], anxiety syndromes via the Screen for Child Anxiety Related Disorders [111], depressive symptom severity via the Short Mood and Feelings Questionnaire [109], eating disorder–related symptoms via the EAT [117], and broad psychosocial difficulties via the SDQ [106].

Statistical Analyses

Prior to analysis, study data will be examined for outliers and missing data, the mechanism and correlates of which will be examined. From daily diary data, mean intensity of positive and negative affect will be computed for each participant as the mean of all the assessments of positive and negative affect provided. Variability and instability of positive and negative affect will be separately computed as the within-person standard deviation over time and MSSD between consecutive assessments, respectively [26,130]. Inertia in positive and negative affect will be computed by regressing positive and negative affect reported at each nightly assessment on positive and negative affect reported at the previous nightly assessment, respectively [65]. Finally, reactivity of positive and negative affect to stress will be computed via 2 approaches, which will be compared in subsequent data analyses as a sensitivity analysis. Initially, positive or negative affect reported at each nightly assessment will be regressed on stress reported concurrently while adjusting for positive or negative affect reported at the previous nightly assessment, respectively. Second, positive or negative affect reported at each nightly assessment will be regressed on stress reported concurrently while adjusting for positive or negative affect reported at the corresponding morning assessment of the same day, respectively. These 2 approaches effectively examine the degree to which participants' report of stress experienced during the day is associated with change in positive or negative affect relative to affect reported the preceding day or the corresponding morning. Where possible, statistical analyses will use full information maximum likelihood estimation [131], which is robust to common forms of missing data [132]. Additionally, potential loss at follow-up will be addressed by identifying and then adjusting for predictors of attrition to mitigate potential bias in estimates of change in psychosocial well-being (eg, loneliness). To promote transparency and reproducibility and limit potential type II errors, empirical reports will follow statistical recommendations to present unadjusted *P* values [133,134]. Interpretation of empirical results will then adjust for multiple testing of correlated outcomes via the Benjamini-Hochberg correction [135].

Analysis of data from baseline assessment and daily diary assessment will use correlation and regression models, multilevel modeling, structural equation modeling, multilevel structural equation modeling, and DSEM to examine how emotion dynamics (eg, instability and inertia) correlate contemporaneously with each domain of well-being: psychopathology, academic motivation, and social well-being. For example, full and partial correlations will examine the association of baseline psychopathology (eg, irritability severity) with mean intensity of positive and negative affect and variability (ie, intraindividual standard deviation) or instability (ie, MSSD) of positive and negative affect over time while accounting for mean affective intensity, respectively. Regression and structural equation models will examine the association of mean intensity of, variability in, and instability in positive and negative affect with indices of academic motivation (eg, amotivation) and social well-being (eg, loneliness) while adjusting for concurrent psychopathology (eg, irritability severity). Model fit will be evaluated for each structural equation model [136] and multilevel structural equation model [137,138] to ensure that each model adequately represents the underlying data.

Similarly, dynamic structural equation models will examine how between-person variation in affective inertia in negative or positive affect and reactivity of negative or positive affect to stress correlate with adolescent psychopathology (eg, anxiety and depressive symptomatology) at baseline (Figure 2). Dynamic structural equation models can similarly examine the association of between-person variation in affective reactivity to daily stress with academic motivation and social well-being above and beyond concurrent psychopathology, mean overall intensity of affect (eg, negative affect) and stress, and intraindividual variability in affect and stress. Specifically, inclusion of psychopathology (eg, depressive symptom severity) assessed at baseline will facilitate examination of how between-person variation in affective inertia and affective reactivity to stress correlate contemporaneously with psychopathology and whether between-person variation in

XSL•FO
RenderX

affective inertia and affective reactivity to stress are associated with academic motivation or social indices of well-being above and beyond psychopathology [139,140]. As previously suggested [24,141], analyses of emotion dynamics (eg, variability and affective inertia) will adjust for between-person variation in mean positive and negative affect.

Figure 2. Example residual dynamic structural equation model of the contemporaneous association of psychopathology with emotion dynamics. "Time in study" represents the amount of time spent in the daily diary procedure at each record to adjust for potentially unique responding early in a daily diary procedure. ats: stress regressed onto prior affect; ia: inertia in affect; is: inertia in stress; sr: affective reactivity to stress; sta: affect regressed onto prior stress; vna: intraindividual variance in negative (or positive) affect; vstress: intraindividual variance in stress.

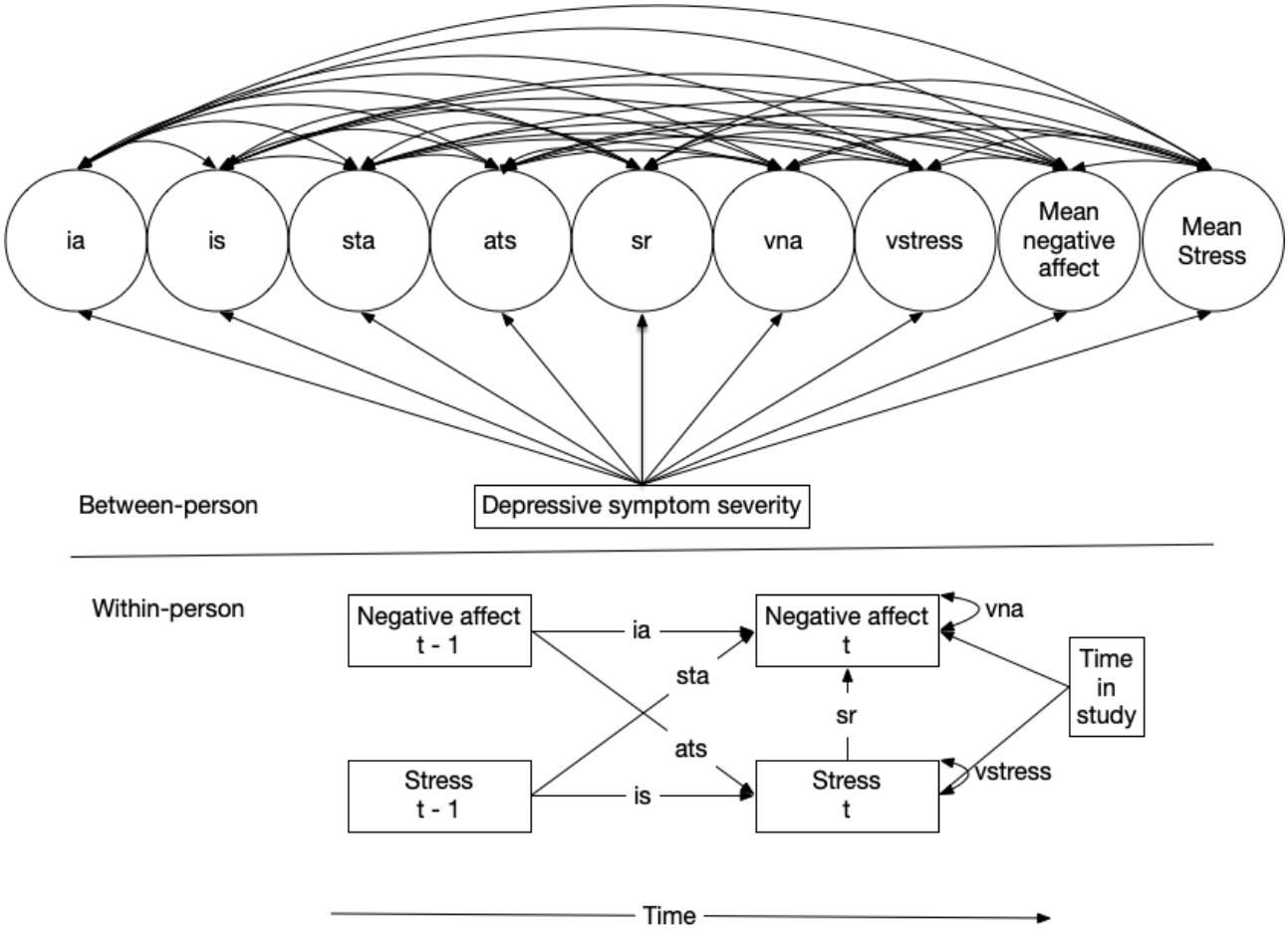

Similarly, analyses of data from longitudinal follow-up assessments will use multilevel modeling, structural equation modeling, multilevel structural equation modeling, and DSEM to examine how each emotion dynamic (eg, instability and inertia) assessed at baseline may predict change in each well-being domain. Specifically, change in each domain will be examined through 2 approaches: estimation of latent growth curves that model trajectories of change (eg, in peer relationship quality) over the 18 months of planned follow-up assessments and estimation of latent change score models that examine change (eg, in academic amotivation) between consecutive follow-up assessments [142]. Similar to analyses of concurrent data from baseline assessment, analyses of emotion dynamics will adjust for mean positive and negative affect [24].

Within the latent growth curve approach, terms that describe the trajectory of change in each well-being domain (eg, intrinsic academic motivation) will be regressed on each emotion dynamic (eg, inertia in positive affect) assessed at baseline to evaluate the degree to which each emotion dynamic may predict gradual change in each well-being domain over the 18-month follow-up period. Subsequent analyses will adjust for psychopathology (eg, symptoms of anxiety syndromes) assessed at baseline and at each follow-up assessment to evaluate the degree to which each emotion dynamic may predict change in each well-being domain above and beyond concurrent, traditional assessment of psychopathology.

Within the latent change score approach, structural equation models and the between-person level of both multilevel and dynamic structural equation models will model change in each well-being domain (eg, social well-being) between consecutive follow-up assessments. Broadly, this model captures within-person change over each 6-month period, which will be, in turn, regressed on each emotion dynamic assessed at baseline to evaluate whether each emotion dynamic (eg, instability in positive affect) may predict change in each well-being domain over the subsequent 6, 12, or 18 months. Finally, secondary analyses will use latent change score models to examine the multivariate patterns of change among psychosocial constructs assessed at baseline and in follow-up assessments (eg, psychopathology and social well-being outcomes). For example, planned analyses will examine whether qualities of adolescents' relationships with parents and peers (eg, trust, communication,

XSL•FO
RenderX

and alienation) prospectively predict or follow from preceding changes in psychopathology over time.

Results

This study was funded in July 2020, but the project was delayed, and funding extended, for 2.5 years due to research restrictions required during the COVID-19 pandemic. Between April 2023 and November 2024, a total of 152 adolescents were enrolled into this longitudinal study such that all scheduled 6-month follow-up assessments were completed by May 2025, 12-month follow-up assessments were completed by November 2025, and 18-month follow-up assessments will be completed by May 2026 (see Figure 1 for a timeline of the study). Three participants (3/152, 1.97%) withdrew, which leaves an effective sample size of 149. Regarding statistical power, the effective sample size reflects 74.5% (149/200) of the target sample, which provides 80% statistical power to detect similarly moderate correlations at baseline using a cautious testwise a value of .005 (r=0.290) and 80% statistical power to detect a moderate prospective association (β=.380) of each emotion dynamic with change in each well-being domain, even after accounting for 20% attrition.

Discussion

Principal Findings

This study aims to examine adolescents' socioemotional development by examining emotion dynamics that describe fluctuations in naturalistic positive and negative affect. Specifically, this study aims to clarify the implications of putative emotion dynamics (ie, mean affective intensity, variability, instability, inertia, and affective reactivity to stress) for psychosocial well-being both concurrently and prospectively over 18 months. For example, by examining the association of emotion dynamics with well-being outcomes (eg, social well-being and academic motivation), this study aims to identify how specific affective patterns (eg, reactivity of negative affect to stress) contribute to adolescent psychosocial adjustment including change in adolescent well-being over time. Finally, this study operationalizes psychosocial well-being broadly to include myriad indicators of psychopathology (eg, symptoms of depressive, eating, and anxiety syndromes) as well as life satisfaction, academic motivation, and social relationships. Traditional approaches to studying adolescent mental health and well-being primarily relied on static, trait-level assessments that may not fully capture the complexity of affective fluctuations in adolescents' day-to-day lives [45]. Therefore, this study seeks to identify the potential contribution of assessing state variation in affect (eg, intraindividual variability in positive and negative affect) to predict psychosocial development during adolescence. For example, this study seeks to identify the potential incremental contribution of assessing state variation in affect over traditional static assessment (eg, baseline self-report assessments) and mean affective intensity.

If study findings align with hypotheses, it will emphasize the contribution of assessing dynamic variation in affect beyond traditional, static assessment, such as traditional self-report assessments of depressive and anxiety symptoms [22]. For example, we hypothesize that naturalistic emotion dynamics derived from in situ affective fluctuations may inform the concurrent model of, and prospective prediction of change in, adolescents' psychosocial well-being. Moreover, this study will clarify whether naturalistic emotion dynamics inform the prospective prediction of change in adolescent psychosocial well-being (eg, depressive symptomatology) above and beyond traditional, static assessments at baseline, such as baseline self-reports of depressive symptomatology. Insights into the importance of assessing dynamic variation could emphasize the importance of understanding and assessing adolescents' daily experiences of positive and negative affect for research and clinical assessment. Such insights may also highlight the potential of emotion dynamics as markers of developmental risk and resilience in adolescence. For example, results from this study may clarify the contribution of dynamic affective variation to adolescent psychosocial development and facilitate more accurate identification of risk for psychopathology including not only *who* is at risk but *when* one enters a period of acute risk (eg, for a depressive episode or panic attack) to support timely and tailored intervention.

This study will also evaluate the degree to which putative affective fluctuations inform change in adolescent well-being beyond clinical symptomatology. Prior research predominantly examined putative emotion dynamics as correlates of psychopathology or broad life satisfaction (see Dejonckheere et al [24] and Houben et al [26] for review). Through examination of academic motivation, social well-being, and life satisfaction, this study will examine the implications of emotion dynamics for functional outcomes previously approximated by clinical symptomatology. If empirical study findings align with hypotheses, this study will highlight the broad relevance of in situ affective fluctuations for adolescent psychosocial development beyond the specific clinical outcomes often studied. Evidence that emotion dynamics contribute to the prediction of change in adolescent psychosocial well-being above and beyond psychopathology would suggest the clinical relevance of emotion dynamics to predict prognosis and functional outcomes above and beyond the existing, traditional assessment of psychopathology. Such evidence would also demonstrate the relevance of emotion dynamics for education, counseling, and social work (eg, through the identification of students at risk for social isolation or decreased academic motivation). Overall, this study uses a broad assessment of adolescent psychopathology and psychosocial well-being in a longitudinal design to assess the breadth and potential relevance of emotion dynamics for adolescent psychosocial development.

Conversely, if study findings do not align with hypotheses, this may suggest that systematic intraindividual variation in emotion adds minimal predictive value for long-term psychosocial outcomes beyond traditional static measures [24]. Although emotion fluctuates over time, weak associations may clarify which dynamics most influence adolescent well-being. Through the use of Bayesian methods to estimate associations, this study will evaluate the relative contributions of emotion dynamics to change in psychopathology and psychosocial well-being over and above concurrent, baseline psychopathology. Evidence of weak associations of emotion dynamics with change in

XSL•FO
RenderX

psychopathology or psychosocial well-being may highlight the relevance of emotion dynamics for short-term outcomes, such as daily functioning, over chronic, long-term psychosocial or clinical outcomes in adolescence. Given the substantial socioemotional developmental challenges posed by adolescence [2], the present examination may clarify which emotion dynamics influence adolescents' mental health, academic well-being, and social functioning contemporaneously or prospectively over time.

Limitations

This study is one of the first and one of the largest longitudinal studies to date to investigate the naturalistic manifestation of emotion dynamics in adolescence as well as their implications for change in adolescent psychosocial well-being over time. However, it is not without limitations. The present sample size represents an initial exploration into the manifestation and potential implications of emotion dynamics. However, this study is designed to inform future research with larger samples to provide greater clarity on the magnitude of each empirical association. As such, empirical findings may be interpreted as a foundation to guide future research.

Similarly, the generalizability of study findings may be limited by sample characteristics, including cultural, socioeconomic, and geographic factors. Participants were drawn from the Windsor-Essex region; factors such as household income, urban versus rural residence, parental education, and access to community resources could influence daily affective patterns and their association with psychosocial outcomes. However, this study is designed to inform future research with larger samples, which could leverage a multisite design to evaluate and increase the generalizability of study empirical findings.

Additionally, this study focuses on daily stressors (eg, daily hassles) rather than traumatic or major life stressors (eg, medical illness, financial hardships, deaths, and victimization). Daily stressors provide a meaningful index of stress in investigations of well-being [143,144] and may predict psychopathology and psychosocial well-being above and beyond major life events [145,146]. However, empirical results regarding associations of daily stressors with well-being should not be generalized to trauma or major stressful events.

Moreover, when designing daily diary and related protocols, one must consider the burden on participants who complete daily diary surveys multiple times daily for many consecutive days (see Shiffman et al [46] for review). Therefore, daily diary surveys are generally restricted in length or frequency, particularly in studies of children and adolescents [147]. This study used a twice daily survey procedure to ask a relatively large set of questions (eg, about affect, stress, and internalizing symptoms) while minimizing the participant burden that could result from many daily assessments. Although this study sampled adolescent participants less often than some extant experience sampling methodology studies of affect in adolescent samples [58,148], daily assessments are consistent with other assessments of affect and stress in adolescence [149,150] and adulthood [151,152]. Therefore, consistent with evidence of affective variation at multiple timescales [153,154], this study is primarily designed for evening assessments to inform interday variation in positive and negative affect while morning assessments provide some information on intraday change in positive and negative affect. We also note that naturalistic assessments of adolescents may often occur less frequently than assessments of adults or young adults (undergraduate students [155]). However, future research is needed to clarify the manifestation and implications of rapid, intraday variation in positive and negative affect to adolescent well-being. Similarly, the 35-day daily diary period may fail to assess longer-term cycles [151], which could be examined in future research on change in emotion dynamics over time [153].

Finally, this study lacks measures of contextual factors that may impact affective fluctuations or individuals' responses to stressors (eg, social support provided in situ). For example, daily and momentary assessments of sleep quality [156], diet [157], and exercise [158] have been associated with well-being. Instead of measuring contextual factors, this study prioritized the assessment of between-person variability in psychosocial factors (eg, social support and self-efficacy) that may be associated with within-person processes (eg, reactivity of positive and negative affect to stress). Therefore, this study was designed to prioritize the examination of between-person correlates of dynamic affective processes. However, future research is needed to develop comprehensive biopsychosocial models of in situ variation in positive and negative affect, such as by integrating active or passive assessments of contextual factors (eg, sleep and exercise).

Conclusions

This protocol paper outlines a longitudinal study designed to examine the developmental implications of naturalistic emotion dynamics (ie, affective intensity, variability, instability, inertia, and affective reactivity to stress). Specifically, this study examined how each dynamic correlates with and prospectively predicts the development of adolescent psychosocial well-being over time including and beyond psychopathology. This paper provides a comprehensive record and overview of the study's aims, background, methodology, and procedures to promote transparency and reproducibility. Additionally, this paper details the integration of daily diary assessment within a longitudinal study design to examine the developmental implications of naturalistic, in situ affective fluctuations. Thereby, beyond serving as a methodological reference, this protocol paper also establishes a framework for future research on adolescent psychosocial and socioemotional development by demonstrating how daily diary assessment can advance the study of dynamic affective processes beyond traditional, static assessment. Particularly following needed replication of any empirical findings, this study has the potential to clarify the manifestation of daily affective fluctuations in adolescence as well as their implications for broad adolescent well-being to, ultimately, inform research and clinical practice in psychology, counseling, and education.

XSL•FO
RenderX

Acknowledgments

The authors wish to thank Rachel Smail-Crevier, Alex Mactavish, Emily Jerome, and Taffy Chan for assistance with study data collection.

Funding

This research was supported by the Social Sciences and Humanities Research Council (grant 430-202-00878).

Conflicts of Interest

RM currently receives grant funding from the Social Sciences and Humanities Research Council and WE-Spark Health Institute. LMR receives or recently received research grant funding from the US National Institutes of Health, Canadian Social Sciences and Humanities Research Council, Government of Ontario Ministries of Health and Long-Term Care, Anxiety Canada, Heart & Stroke Foundation of Canada, and WE-Spark Health Institute. He also recently sat on the scientific advisory committee of Anxiety Canada. Otherwise, the authors have no conflicts of interest to disclose.

References

1. Benton TD, Boyd RC, Njoroge WFM. Addressing the global crisis of child and adolescent mental health. JAMA Pediatr 2021;175(11):1108-1110. [doi: 10.1001/jamapediatrics.2021.2479] [Medline: 34369961]
2. Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. Lancet 2007;369(9569):1302-1313. [doi: 10.1016/S0140-6736(07)60368-7] [Medline: 17434406]
3. Flett GL, Hewitt PL. Disguised distress in children and adolescents "Flying Under the Radar". Why psychological problems are underestimated and how schools must respond. Can J Sch Psychol 2013;28(1):12-27. [doi: 10.1177/0829573512468845]
4. Georgiades K, Duncan L, Wang L, Comeau J, Boyle MH, 2014 Ontario Child Health Study Team. Six-month prevalence of mental disorders and service contacts among children and youth in Ontario: evidence from the 2014 Ontario child health study. Can J Psychiatry 2019;64(4):246-255 [FREE Full text] [doi: 10.1177/0706743719830024] [Medline: 30978138]
5. Verhulst FC, van der Ende J, Ferdinand RF, Kasius MC. The prevalence of DSM-III-R diagnoses in a national sample of Dutch adolescents. Arch Gen Psychiatry 1997;54(4):329-336. [doi: 10.1001/archpsyc.1997.01830160049008] [Medline: 9107149]
6. Costello EJ, Copeland W, Angold A. Trends in psychopathology across the adolescent years: what changes when children become adolescents, and when adolescents become adults? J Child Psychol Psychiatry 2011;52(10):1015-1025 [FREE Full text] [doi: 10.1111/j.1469-7610.2011.02446.x] [Medline: 21815892]
7. de Girolamo G, Dagani J, Purcell R, Cocchi A, McGorry PD. Age of onset of mental disorders and use of mental health services: needs, opportunities and obstacles. Epidemiol Psychiatr Sci 2012;21(1):47-57. [doi: 10.1017/s2045796011000746] [Medline: 22670412]
8. Hughes AA, Lourea-Waddell B, Kendall PC. Somatic complaints in children with anxiety disorders and their unique prediction of poorer academic performance. Child Psychiatry Hum Dev 2008;39(2):211-220. [doi: 10.1007/s10578-007-0082-5] [Medline: 17786552]
9. Costello EJ, Egger H, Angold A. 10-year research update review: the epidemiology of child and adolescent psychiatric disorders: I. Methods and public health burden. J Am Acad Child Adolesc Psychiatry 2005;44(10):972-986. [doi: 10.1097/01.chi.0000172552.41596.6f] [Medline: 16175102]
10. Arnett JJ. Adolescent storm and stress, reconsidered. Am Psychol 1999;54(5):317-326. [doi: 10.1037//0003-066x.54.5.317] [Medline: 10354802]
11. Boak A, Hamilton HA, Adlaf EM, Henderson JL, Mann RE. The Mental Health and Well-being of Ontario Students, 1991-2017: Detailed Findings From the Ontario Student Drug Use and Health Survey (OSDUHS). Toronto, ON: Centre for Addiction and Mental Health; 2018.
12. Farrell AH, Szatmari P, Vaillancourt T. Epidemiology of mental health challenges in children and adolescents. Pediatr Clin North Am 2024;71(6):999-1011. [doi: 10.1016/j.pcl.2024.07.009] [Medline: 39433386]
13. Keyes KM, Platt JM. Annual research review: Sex, gender, and internalizing conditions among adolescents in the 21st century—trends, causes, consequences. J Child Psychol Psychiatry 2024;65(4):384-407. [doi: 10.1111/jcpp.13864] [Medline: 37458091]
14. Santillanes G, Axeen S, Lam CN, Menchine M. National trends in mental health-related emergency department visits by children and adults, 2009-2015. Am J Emerg Med 2020;38(12):2536-2544. [doi: 10.1016/j.ajem.2019.12.035] [Medline: 31902702]
15. Costa PT, McCrae RR. Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people. J Pers Soc Psychol 1980;38(4):668-678. [doi: 10.1037//0022-3514.38.4.668] [Medline: 7381680]
16. Cloninger CR, Svrakic DM, Przybeck TR. A psychobiological model of temperament and character. Arch Gen Psychiatry 1993;50(12):975-990. [doi: 10.1001/archpsyc.1993.01820240059008] [Medline: 8250684]

XSL•FO
RenderX

17. Raja SN, McGee R, Stanton WR. Perceived attachments to parents and peers and psychological well-being in adolescence. J Youth Adolescence 1992;21(4):471-485. [doi: 10.1007/bf01537898]
18. Gourion D, Pélissolo A, Lépine JP. Test-retest reliability of the temperament and character inventory in patients with opiate dependence. Psychiatry Res 2003;118(1):81-88. [doi: 10.1016/s0165-1781(03)00042-8] [Medline: 12759164]
19. Gullone E, Robinson K. The Inventory of Parent and Peer Attachment—Revised (IPPA-R) for children: a psychometric investigation. Clin Psychol Psychother 2005;12(1):67-79. [doi: 10.1002/cpp.433]
20. Costa PT, McCrae RR. Domains and facets: hierarchical personality assessment using the revised NEO personality inventory. J Pers Assess 1995;64(1):21-50. [doi: 10.1207/s15327752jpa6401_2] [Medline: 16367732]
21. Clark LA, Watson D, Mineka S. Temperament, personality, and the mood and anxiety disorders. J Abnorm Psychol 1994;103(1):103-116. [doi: 10.1037/0021-843x.103.1.103]
22. Watson D, Clark LA, Carey G. Positive and negative affectivity and their relation to anxiety and depressive disorders. J Abnorm Psychol 1988;97(3):346-353. [doi: 10.1037//0021-843x.97.3.346] [Medline: 3192830]
23. Rothbart MK, Ahadi SA, Hershey KL, Fisher P. Investigations of temperament at three to seven years: the children's behavior questionnaire. Child Dev 2001;72(5):1394-1408. [doi: 10.1111/1467-8624.00355] [Medline: 11699677]
24. Dejonckheere E, Mestdagh M, Houben M, Rutten I, Sels L, Kuppens P, et al. Complex affect dynamics add limited information to the prediction of psychological well-being. Nat Hum Behav 2019;3(5):478-491. [doi: 10.1038/s41562-019-0555-0] [Medline: 30988484]
25. Fleeson W, Gallagher P. The implications of big five standing for the distribution of trait manifestation in behavior: fifteen experience-sampling studies and a meta-analysis. J Pers Soc Psychol 2009;97(6):1097-1114 [FREE Full text] [doi: 10.1037/a0016786] [Medline: 19968421]
26. Houben M, Van Den Noortgate W, Kuppens P. The relation between short-term emotion dynamics and psychological well-being: a meta-analysis. Psychol Bull 2015;141(4):901-930. [doi: 10.1037/a0038822] [Medline: 25822133]
27. Shapero BG, Abramson LY, Alloy LB. Emotional reactivity and internalizing symptoms: moderating role of emotion regulation. Cognit Ther Res 2016;40(30):328-340 [FREE Full text] [doi: 10.1007/s10608-015-9722-4] [Medline: 27231404]
28. Shapero BG, Farabaugh A, Terechina O, DeCross S, Cheung JC, Fava M, et al. Understanding the effects of emotional reactivity on depression and suicidal thoughts and behaviors: moderating effects of childhood adversity and resilience. J Affect Disord 2019;245:419-427. [doi: 10.1016/j.jad.2018.11.033] [Medline: 30423470]
29. Rappaport LM, Moskowitz DS, D'Antono B. Naturalistic interpersonal behavior patterns differentiate depression and anxiety symptoms in the community. J Couns Psychol 2014;61(2):253-263. [doi: 10.1037/a0035625] [Medline: 24660689]
30. Zuroff DC, Fournier MA, Moskowitz DS. Depression, perceived inferiority, and interpersonal behavior: evidence for the involuntary defeat strategy. J Soc Clin Psychol 2007;26(7):751-778. [doi: 10.1521/jscp.2007.26.7.751]
31. Rappaport LM, Moskowitz DS, D'Antono B. Depression symptoms moderate the association between emotion and communal behavior. J Couns Psychol 2017;64(3):269-279 [FREE Full text] [doi: 10.1037/cou0000194] [Medline: 28221084]
32. Côté S, Moskowitz DS. On the dynamic covariation between interpersonal behavior and affect: prediction from neuroticism, extraversion, and agreeableness. J Pers Soc Psychol 1998;75(4):1032-1046. [doi: 10.1037//0022-3514.75.4.1032] [Medline: 9825533]
33. Tracey TJG. Levels of interpersonal complementarity: a simplex representation. Pers Soc Psychol Bull 2004;30(9):1211-1225. [doi: 10.1177/0146167204264075] [Medline: 15359023]
34. Kuppens P, Oravecz Z, Tuerlinckx F. Feelings change: accounting for individual differences in the temporal dynamics of affect. J Pers Soc Psychol 2010;99(6):1042-1060. [doi: 10.1037/a0020962] [Medline: 20853980]
35. Frijda NH. The Laws of Emotion. Mahwah, NJ: Erlbaum; 2007.
36. Gross JJ, Thompson RA. Emotion regulation: conceptual foundations. In: Gross JJ, editor. Handbook of Emotion Regulation. New York, NY: Guilford Press; 2007.
37. Scherer KR. The dynamic architecture of emotion: evidence for the component process model. Cogn Emot 2009;23(7):1307-1351. [doi: 10.1080/02699930902928969]
38. Dork J, Mangan E, Burns L, Dimenstein E. Affective instability: impact of fluctuating emotions on regulation and psychological well-being. Behav Sci (Basel) 2024;14(9):783 [FREE Full text] [doi: 10.3390/bs14090783] [Medline: 39335997]
39. Kuppens P, Verduyn P. Looking at emotion regulation through the window of emotion dynamics. Psychol Inq 2015;26(1):72-79. [doi: 10.1080/1047840x.2015.960505]
40. Gross JT, Cassidy J. Expressive suppression of negative emotions in children and adolescents: theory, data, and a guide for future research. Dev Psychol 2019;55(9):1938-1950. [doi: 10.1037/dev0000722] [Medline: 31464496]
41. Bailen NH, Green LM, Thompson RJ. Understanding emotion in adolescents: a review of emotional frequency, intensity, instability, and clarity. Emot Rev 2018;11(1):63-73. [doi: 10.1177/1754073918768878]
42. Coe-Odess SJ, Narr RK, Allen JP. Emergent emotions in adolescence. In: LoBue V, Pérez-Edgar K, Buss KA, editors. Handbook of Emotional Development. Cham, Swizerland: Springer International Publishing; 2019:595-625.
43. Silk JS. Context and dynamics: the new frontier for developmental research on emotion regulation. Dev Psychol 2019;55(9):2009-2014. [doi: 10.1037/dev0000768] [Medline: 31464504]

XSL•FO
RenderX

44. Larson R, Csikszentmihalyi M, Graef R. Mood variability and the psychosocial adjustment of adolescents. J Youth Adolesc 1980;9(6):469-490. [doi: 10.1007/bf02089885]
45. Silk JS, Steinberg L, Morris AS. Adolescents' emotion regulation in daily life: links to depressive symptoms and problem behavior. Child Dev 2003;74(6):1869-1880. [doi: 10.1046/j.1467-8624.2003.00643.x] [Medline: 14669901]
46. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
47. Khantzian EJ. The self-medication hypothesis of substance use disorders: a reconsideration and recent applications. Harv Rev Psychiatry 1997;4(5):231-244. [doi: 10.3109/10673229709030550] [Medline: 9385000]
48. Turner S, Mota N, Bolton J, Sareen J. Self-medication with alcohol or drugs for mood and anxiety disorders: a narrative review of the epidemiological literature. Depress Anxiety 2018;35(9):851-860 [FREE Full text] [doi: 10.1002/da.22771] [Medline: 29999576]
49. Kaysen D, Atkins DC, Simpson TL, Stappenbeck CA, Blayney JA, Lee CM, et al. Proximal relationships between PTSD symptoms and drinking among female college students: results from a daily monitoring study. Psychol Addict Behav 2014;28(1):62-73 [FREE Full text] [doi: 10.1037/a0033588] [Medline: 23915369]
50. Rappaport LM, Cusack SE, Sheerin CM, Amstadter AB. Intraindividual association of PTSD symptoms with binge drinking among trauma-exposed students. J Couns Psychol 2021;68(5):571-581 [FREE Full text] [doi: 10.1037/cou0000543] [Medline: 33764116]
51. Mogg K, Mathews A, Weinman J. Memory bias in clinical anxiety. J Abnorm Psychol 1987;96(2):94-98. [doi: 10.1037//0021-843x.96.2.94] [Medline: 3584672]
52. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006;31(1):13-20. [doi: 10.1139/jpn.0602]
53. Kuppens P, Verduyn P. Emotion dynamics. Curr Opin Psychol 2017;17:22-26. [doi: 10.1016/j.copsyc.2017.06.004] [Medline: 28950968]
54. Schneider S, Stone AA. Ambulatory and diary methods can facilitate the measurement of patient-reported outcomes. Qual Life Res 2016;25(3):497-506 [FREE Full text] [doi: 10.1007/s11136-015-1054-z] [Medline: 26101141]
55. Davis PA, Burns GL. Influence of emotional intensity and frequency of positive and negative events on depression. Eur J Psychol Assess 1999;15(2):106-116. [doi: 10.1027//1015-5759.15.2.106]
56. Sultson H, Murd C, Havik M, Konstabel K. Negative affect instability predicts elevated depressive and generalized anxiety disorder symptoms even when negative affect intensity is controlled for: an ecological momentary assessment study. Front Psychol 2024;15:1371115 [FREE Full text] [doi: 10.3389/fpsyg.2024.1371115] [Medline: 38716268]
57. Rojas SM, Leen-Feldner EW, Blumenthal H, Lewis SF, Feldner MT. Risk for suicide among treatment seeking adolescents: the role of positive and negative affect intensity. Cogn Ther Res 2014;39(2):100-109. [doi: 10.1007/s10608-014-9650-8]
58. Silk JS, Dahl RE, Ryan ND, Forbes EE, Axelson DA, Birmaher B, et al. Pupillary reactivity to emotional information in child and adolescent depression: links to clinical and ecological measures. Am J Psychiatry 2007;164(12):1873-1880 [FREE Full text] [doi: 10.1176/appi.ajp.2007.06111816] [Medline: 18056243]
59. Ebner-Priemer UW, Eid M, Kleindienst N, Stabenow S, Trull TJ. Analytic strategies for understanding affective (in)stability and other dynamic processes in psychopathology. J Abnorm Psychol 2009;118(1):195-202. [doi: 10.1037/a0014868] [Medline: 19222325]
60. Panaite V, Rottenberg J, Bylsma LM. Daily affective dynamics predict depression symptom trajectories among adults with major and minor depression. Affect Sci 2020;1(3):186-198 [FREE Full text] [doi: 10.1007/s42761-020-00014-w] [Medline: 36043207]
61. Larson RW, Raffaelli M, Richards MH, Ham M, Jewell L. Ecology of depression in late childhood and early adolescence: a profile of daily states and activities. J Abnorm Psychol 1990;99(1):92-102. [doi: 10.1037/0021-843x.99.1.92]
62. Maciejewski DF, Keijsers L, van Lier PAC, Branje SJT, Meeus WHJ, Koot HM. Most fare well—but some do not: distinct profiles of mood variability development and their association with adjustment during adolescence. Dev Psychol 2019;55(2):434-448. [doi: 10.1037/dev0000650] [Medline: 30507219]
63. Silk JS, Forbes EE, Whalen DJ, Jakubcak JL, Thompson WK, Ryan ND, et al. Daily emotional dynamics in depressed youth: a cell phone ecological momentary assessment study. J Exp Child Psychol 2011;110(2):241-257 [FREE Full text] [doi: 10.1016/j.jecp.2010.10.007] [Medline: 21112595]
64. Gilbert KE. The neglected role of positive emotion in adolescent psychopathology. Clin Psychol Rev 2012;32(6):467-481. [doi: 10.1016/j.cpr.2012.05.005] [Medline: 22710138]
65. Kuppens P, Allen NB, Sheeber LB. Emotional inertia and psychological maladjustment. Psychol Sci 2010;21(7):984-991 [FREE Full text] [doi: 10.1177/0956797610372634] [Medline: 20501521]
66. Ong AD, Wang L, Fang Y, Zhang G, Joiner RJ, Wilcox KT, et al. Affect dynamics and depressive symptomatology: revisiting the inertia-instability paradox. Emotion 2025;25(4):1060-1064. [doi: 10.1037/emo0001465] [Medline: 39636613]
67. Asparouhov T, Hamaker EL, Muthén B. Dynamic structural equation models. Struct Equ Model Multidiscip J 2017;25(3):359-388. [doi: 10.1080/10705511.2017.1406803]

XSL•FO
RenderX

68. Hamaker EL, Asparouhov T, Muthen B. Dynamic structural equation modeling as a combination of time series modeling, multilevel modeling, and structural equation modeling. In: Hoyle RH, editor. Handbook of Structural Equation Modeling. 2nd ed. New York: Guilford Press; 2023.
69. McNeish D, Hamaker EL. A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. Psychol Methods 2020;25(5):610-635. [doi: 10.1037/met0000250] [Medline: 31855015]
70. Asparouhov T, Muthén B. Latent variable centering of predictors and mediators in multilevel and time-series models. Struct Equ Model Multidiscip J 2018;26(1):119-142. [doi: 10.1080/10705511.2018.1511375]
71. Shrout PE, Stadler G, Lane SP, McClure MJ, Jackson GL, Clavél FD, et al. Initial elevation bias in subjective reports. Proc Natl Acad Sci USA 2018;115(1):E15-E23 [FREE Full text] [doi: 10.1073/pnas.1712277115] [Medline: 29255039]
72. Koval P, Brose A, Pe ML, Houben M, Erbas Y, Champagne D, et al. Emotional inertia and external events: the roles of exposure, reactivity, and recovery. Emotion 2015;15(5):625-636. [doi: 10.1037/emo0000059] [Medline: 25844974]
73. Suls J, Green P, Hillis S. Emotional reactivity to everyday problems, affective inertia, and neuroticism. Pers Soc Psychol Bull 1998;24(2):127-136. [doi: 10.1177/0146167298242002]
74. van de Leemput IA, Wichers M, Cramer AOJ, Borsboom D, Tuerlinckx F, Kuppens P, et al. Critical slowing down as early warning for the onset and termination of depression. Proc Natl Acad Sci USA 2014;111(1):87-92 [FREE Full text] [doi: 10.1073/pnas.1312114110] [Medline: 24324144]
75. Thompson RJ, Mata J, Jaeggi SM, Buschkuehl M, Jonides J, Gotlib IH. The everyday emotional experience of adults with major depressive disorder: examining emotional instability, inertia, and reactivity. J Abnorm Psychol 2012;121(4):819-829 [FREE Full text] [doi: 10.1037/a0027978] [Medline: 22708886]
76. Kuppens P, Sheeber LB, Yap MBH, Whittle S, Simmons JG, Allen NB. Emotional inertia prospectively predicts the onset of depressive disorder in adolescence. Emotion 2012;12(2):283-289 [FREE Full text] [doi: 10.1037/a0025046] [Medline: 21988744]
77. Leppert KA, Bufferd SJ, Olino TM, Dougherty LR. A daily diary analysis of preschool depressive behaviors: prospective associations and moderators across 14 days. J Abnorm Child Psychol 2019;47(9):1547-1558 [FREE Full text] [doi: 10.1007/s10802-019-00535-4] [Medline: 30911869]
78. Bylsma LM, Morris BH, Rottenberg J. A meta-analysis of emotional reactivity in major depressive disorder. Clin Psychol Rev 2008;28(4):676-691. [doi: 10.1016/j.cpr.2007.10.001] [Medline: 18006196]
79. Bylsma LM, Taylor-Clift A, Rottenberg J. Emotional reactivity to daily events in major and minor depression. J Abnorm Psychol 2011;120(1):155-167. [doi: 10.1037/a0021662] [Medline: 21319928]
80. Bai S, Repetti RL. Negative and positive emotion responses to daily school problems: links to internalizing and externalizing symptoms. J Abnorm Child Psychol 2018;46(3):423-435 [FREE Full text] [doi: 10.1007/s10802-017-0311-8] [Medline: 28577264]
81. Herres J, Ewing ESK, Kobak R. Emotional reactivity to negative adult and peer events and the maintenance of adolescent depressive symptoms: a daily diary design. J Abnorm Child Psychol 2016;44(3):471-481 [FREE Full text] [doi: 10.1007/s10802-015-0043-6] [Medline: 26084594]
82. Goldschmidt AB, Wonderlich SA, Crosby RD, Engel SG, Lavender JM, Peterson CB, et al. Ecological momentary assessment of stressful events and negative affect in bulimia nervosa. J Consult Clin Psychol 2014;82(1):30-39 [FREE Full text] [doi: 10.1037/a0034974] [Medline: 24219182]
83. Tan PZ, Forbes EE, Dahl RE, Ryan ND, Siegle GJ, Ladouceur CD, et al. Emotional reactivity and regulation in anxious and nonanxious youth: a cell-phone ecological momentary assessment study. J Child Psychol Psychiatry 2012;53(2):197-206 [FREE Full text] [doi: 10.1111/j.1469-7610.2011.02469.x] [Medline: 22176136]
84. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. Arch Gen Psychiatry 2005;62(6):593-602. [doi: 10.1001/archpsyc.62.6.593] [Medline: 15939837]
85. Steinberg L, Morris AS. Adolescent development. Annu Rev Psychol 2001;52(1):83-110. [doi: 10.1146/annurev.psych.52.1.83] [Medline: 11148300]
86. Cohen J. A power primer. Psychol Bull 1992;112(1):155-159. [doi: 10.1037/0033-2909.112.1.155]
87. Muthén BO, Muthén LK. Mplus User's Guide. 8th ed. Los Angeles, CA: Muthén & Muthén; 2017:944.
88. Torsheim T, Cavallo F, Levin KA, Schnohr C, Mazur J, Niclasen B, FAS Development Study Group. Psychometric validation of the Revised Family Affluence Scale: a latent variable approach. Child Indic Res 2016;9:771-784 [FREE Full text] [doi: 10.1007/s12187-015-9339-x] [Medline: 27489572]
89. Goodman E, Adler NE, Kawachi I, Frazier AL, Huang B, Colditz GA. Adolescents' perceptions of social status: development and evaluation of a new indicator. Pediatrics 2001;108(2):e31. [doi: 10.1542/peds.108.2.e31] [Medline: 11483841]
90. Wang Y, Zhu X, Cai L, Wang Q, Wang M, Yi J, et al. Screening cluster A and cluster B personality disorders in Chinese high school students. BMC Psychiatry 2013;13(1):116 [FREE Full text] [doi: 10.1186/1471-244X-13-116] [Medline: 23594882]
91. Vallerand RJ, Pelletier LG, Blais MR, Briere NM, Senecal C, Vallieres EF. The Academic Motivation Scale: a measure of intrinsic, extrinsic, and amotivation in education. Educ Psychol Meas 1992;52(4):1003-1017. [doi: 10.1177/0013164492052004025]

XSL•FO
RenderX

92. Green CD, Thornton K, Fredrick JW, Wiggs KK, Smith ZR, Langberg JM, et al. Longitudinal evaluation of school climate in predicting mood and academic functioning in adolescence: the role of academic motivation. School Ment Health 2023;15(4):1033-1048. [doi: 10.1007/s12310-023-09606-w]
93. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
94. Lasgaard M. Reliability and validity of the Danish version of the UCLA Loneliness Scale. Personal Individ Differ 2007;42(7):1359-1366. [doi: 10.1016/j.paid.2006.10.013]
95. Sarason IG, Sarason BR, Shearin EN, Pierce GR. A brief measure of social support: practical and theoretical implications. J Soc Pers Relatsh 1987;4(4):497-510. [doi: 10.1177/0265407587044007]
96. Dumont M, Provost MA. Resilience in adolescents: protective role of social support, coping strategies, self-esteem, and social activities on experience of stress and depression. J Youth Adolesc 1999;28(3):343-363. [doi: 10.1023/a:1021637011732]
97. Armsden GC, Greenberg MT. The inventory of parent and peer attachment: individual differences and their relationship to psychological well-being in adolescence. J Youth Adolesc 1987;16(5):427-454. [doi: 10.1007/BF02202939] [Medline: 24277469]
98. Jewell T, Gardner T, Susi K, Watchorn K, Coopey E, Simic M, et al. Attachment measures in middle childhood and adolescence: a systematic review of measurement properties. Clin Psychol Rev 2019;68:71-82. [doi: 10.1016/j.cpr.2018.12.004]
99. Sheehan DV, Sheehan KH, Shytle RD, Janavs J, Bannon Y, Rogers JE, et al. Reliability and validity of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). J Clin Psychiatry 2010;71(3):313-326. [doi: 10.4088/JCP.09m05305whi] [Medline: 20331933]
100. Pavot W, Diener E. Review of the Satisfaction With Life Scale. Psychol Assess 1993;5(2):164-172. [doi: 10.1037/1040-3590.5.2.164]
101. Jovanović V. The validity of the Satisfaction With Life Scale in adolescents and a comparison with single-item life satisfaction measures: a preliminary study. Qual Life Res 2016;25(12):3173-3180. [doi: 10.1007/s11136-016-1331-5] [Medline: 27262574]
102. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24(4):385-396. [doi: 10.2307/2136404]
103. Chen W, Liang Y, Cheng J, Xie E, Zhu Q, Liu J. Psychometric properties of the Perceived Stress Scale Short Form (PSS-10) in Chinese adolescents: validity, measurement invariance across gender and left-behind status. Child Youth Serv Rev 2024;156:107252. [doi: 10.1016/j.childyouth.2023.107252]
104. Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. J Psychopathol Behav Assess 2004;26(1):41-54. [doi: 10.1023/B:JOBA.0000007455.08539.94]
105. Weinberg A, Klonsky ED. Measurement of emotion dysregulation in adolescents. Psychol Assess 2009;21(4):616-621. [doi: 10.1037/a0016669] [Medline: 19947794]
106. Goodman R, Meltzer H, Bailey V. The strengths and difficulties questionnaire: a pilot study on the validity of the self-report version. Eur Child Adolesc Psychiatry 1998;7(3):125-130. [doi: 10.1007/s007870050057] [Medline: 9826298]
107. He JP, Burstein M, Schmitz A, Merikangas KR. The Strengths and Difficulties Questionnaire (SDQ): the factor structure and scale validation in U.S. adolescents. J Abnorm Child Psychol 2013;41(4):583-595. [doi: 10.1007/s10802-012-9696-6] [Medline: 23183936]
108. Van Roy B, Veenstra M, Clench-Aas J. Construct validity of the five-factor Strengths and Difficulties Questionnaire (SDQ) in pre-, early, and late adolescence. J Child Psychol Psychiatry 2008;49(12):1304-1312. [doi: 10.1111/j.1469-7610.2008.01942.x] [Medline: 19120709]
109. Angold A, Costello EJ, Messer SC, Pickles A. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. Int J Methods Psychiatr Res 1995;5(4):237-249.
110. Fuseekul N, Orchard F, Reynolds S. Psychometric properties of the Thai Mood and Feelings Questionnaire (MFQ) for adolescent depression. Child Adolesc Psychiatry Ment Health 2021;15(1):17 [FREE Full text] [doi: 10.1186/s13034-021-00372-8] [Medline: 33836780]
111. Birmaher B, Khetarpal S, Brent D, Cully M, Balach L, Kaufman J, et al. The Screen for Child Anxiety Related Emotional Disorders (SCARED): scale construction and psychometric characteristics. J Am Acad Child Adolesc Psychiatry 1997;36(4):545-553. [doi: 10.1097/00004583-199704000-00018] [Medline: 9100430]
112. Albano AM, Comer JS, Compton SN, Piacentini J, Kendall PC, Birmaher B, et al. Secondary outcomes from the child/adolescent anxiety multimodal study: implications for clinical practice. Evid Based Pract Child Adolesc Ment Health 2018;3(1):30-41 [FREE Full text] [doi: 10.1080/23794925.2017.1399485] [Medline: 30906874]
113. Caporino NE, Sakolsky D, Brodman DM, McGuire JF, Piacentini J, Peris TS, et al. Establishing clinical cutoffs for response and remission on the Screen for Child Anxiety Related Emotional Disorders (SCARED). J Am Acad Child Adolesc Psychiatry 2017;56(8):696-702 [FREE Full text] [doi: 10.1016/j.jaac.2017.05.018] [Medline: 28735699]

XSL•FO
RenderX

114. Hale WW, Raaijmakers Q, Muris P, Meeus W. Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED) in the general adolescent population. J Am Acad Child Adolesc Psychiatry 2005;44(3):283-290. [doi: 10.1097/00004583-200503000-00013] [Medline: 15725973]
115. Stringaris A, Goodman R, Ferdinando S, Razdan V, Muhrer E, Leibenluft E, et al. The Affective Reactivity Index: a concise irritability scale for clinical and research settings. J Child Psychol Psychiatry 2012;53(11):1109-1117 [FREE Full text] [doi: 10.1111/j.1469-7610.2012.02561.x] [Medline: 22574736]
116. Mulraney MA, Melvin GA, Tonge BJ. Psychometric properties of the affective reactivity index in Australian adults and adolescents. Psychol Assess 2014;26(1):148-155. [doi: 10.1037/a0034891] [Medline: 24188148]
117. Garner DM, Olmsted MP, Bohr Y, Garfinkel PE. The Eating Attitudes Test: psychometric features and clinical correlates. Psychol Med 2009;12(4):871-878. [doi: 10.1017/S0033291700049163]
118. Maïano C, Morin AJS, Lanfranchi MC, Therme P. The Eating Attitudes Test-26 revisited using exploratory structural equation modeling. J Abnorm Child Psychol 2013;41(5):775-788. [doi: 10.1007/s10802-013-9718-z] [Medline: 23344702]
119. Kelleher I, Connor D, Clarke MC, Devlin N, Harley M, Cannon M. Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies. Psychol Med 2012;42(9):1857-1863. [doi: 10.1017/s0033291711002960]
120. Lyall K, Croen L, Daniels J, Fallin MD, Ladd-Acosta C, Lee BK, et al. The changing epidemiology of autism spectrum disorders. Annu Rev Public Health 2017;38(1):81-102 [FREE Full text] [doi: 10.1146/annurev-publhealth-031816-044318] [Medline: 28068486]
121. Tabori Kraft J, Dalsgaard S, Obel C, Thomsen PH, Henriksen TB, Scahill L. Prevalence and clinical correlates of tic disorders in a community sample of school-age children. Eur Child Adolesc Psychiatry 2012;21(1):5-13. [doi: 10.1007/s00787-011-0223-z] [Medline: 22038343]
122. Duncan L, Georgiades K, Wang L, Van Lieshout RJ, MacMillan HL, Ferro MA, et al. Psychometric evaluation of the Mini International Neuropsychiatric Interview for Children and Adolescents (MINI-KID). Psychol Assess 2018;30(7):916-928. [doi: 10.1037/pas0000541] [Medline: 29199837]
123. Laurent J, Catanzaro SJ, Joiner TE, Rudolph KD, Potter KI, Lambert S, et al. A measure of positive and negative affect for children: scale development and preliminary validation. Psychol Assess 1999;11(3):326-338. [doi: 10.1037/1040-3590.11.3.326]
124. Ahmed W. Psychometric properties of Positive and Negative Schedule Scale for Children—Short Form in diverse American adolescents. J Psychoeduc Assess 2023;42(1):3-13. [doi: 10.1177/07342829231206576]
125. Chorpita BF, Daleiden EL. Tripartite dimensions of emotion in a child clinical sample: measurement strategies and implications for clinical utility. J Consult Clin Psychol 2002;70(5):1150-1160. [doi: 10.1037//0022-006x.70.5.1150] [Medline: 12362965]
126. Ebesutani C, Okamura K, Higa-McMillan C, Chorpita BF. A psychometric analysis of the positive and negative affect schedule for children-parent version in a school sample. Psychol Assess 2011;23(2):406-416. [doi: 10.1037/a0022057] [Medline: 21381834]
127. Wilson K, Gullone E, Moss S. The youth version of the positive and negative affect schedule: a psychometric validation. Behav Change 2014;15(3):187-193. [doi: 10.1017/s0813483900003077]
128. Dunkley DM, Zuroff DC, Blankstein KR. Self-critical perfectionism and daily affect: dispositional and situational influences on stress and coping. J Pers Soc Psychol 2003;84(1):234-252. [doi: 10.1037/0022-3514.84.1.234]
129. Mandel T, Dunkley DM, Moroz M. Self-critical perfectionism and depressive and anxious symptoms over 4 years: the mediating role of daily stress reactivity. J Couns Psychol 2015;62(4):703-717. [doi: 10.1037/cou0000101] [Medline: 26191980]
130. Jahng S, Wood PK, Trull TJ. Analysis of affective instability in ecological momentary assessment: indices using successive difference and group comparison via multilevel modeling. Psychol Methods 2008 Dec;13(4):354-375. [doi: 10.1037/a0014173] [Medline: 19071999]
131. Pritikin JN, Hunter MD, von Oertzen T, Brick TR, Boker SM. Many-level multilevel structural equation modeling: an efficient evaluation strategy. Struct Equ Modeling 2017;24(5):684-698 [FREE Full text] [doi: 10.1080/10705511.2017.1293542] [Medline: 29606847]
132. Little TD, Jorgensen TD, Lang KM, Moore EWG. On the joys of missing data. J Pediatr Psychol 2014;39(2):151-162. [doi: 10.1093/jpepsy/jst048] [Medline: 23836191]
133. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990;1(1):43-46. [doi: 10.1097/00001648-199001000-00010]
134. Streiner DL, Norman GR. Correction for multiple testing: is there a resolution? Chest 2011;140(1):16-18. [doi: 10.1378/chest.11-0523] [Medline: 21729890]
135. Blakesley RE, Mazumdar S, Dew MA, Houck PR, Tang G, Reynolds CF, et al. Comparisons of methods for multiple hypothesis testing in neuropsychological research. Neuropsychology 2009;23(2):255-264 [FREE Full text] [doi: 10.1037/a0012850] [Medline: 19254098]
136. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model Multidiscip J 1999;6(1):1-55. [doi: 10.1080/10705519909540118]

XSL•FO
RenderX

137. Hsu HY, Kwok O, Lin JH, Acosta S. Detecting misspecified multilevel structural equation models with common fit indices: a Monte Carlo Study. Multivariate Behav Res 2015;50(2):197-215. [doi: 10.1080/00273171.2014.977429] [Medline: 26609878]
138. Rappaport LM, Amstadter AB, Neale MC. Model fit estimation for multilevel structural equation models. Struct Equ Modeling 2020;27(2):318-329 [FREE Full text] [doi: 10.1080/10705511.2019.1620109] [Medline: 32774078]
139. Liu Y. The longitudinal relationship between Chinese high school students' academic stress and academic motivation. Learn Individ Differ 2015;38:123-126. [doi: 10.1016/j.lindif.2015.02.002]
140. Sigfusdottir ID, Kristjansson AL, Thorlindsson T, Allegrante JP. Stress and adolescent well-being: the need for an interdisciplinary framework. Health Promot Int 2017;32(6):1081-1090 [FREE Full text] [doi: 10.1093/heapro/daw038] [Medline: 27153917]
141. Baird BM, Le K, Lucas RE. On the nature of intraindividual personality variability: reliability, validity, and associations with well-being. J Pers Soc Psychol 2006;90(3):512-527. [doi: 10.1037/0022-3514.90.3.512] [Medline: 16594835]
142. McArdle JJ. Latent variable modeling of differences and changes with longitudinal data. Annu Rev Psychol 2009;60(1):577-605. [doi: 10.1146/annurev.psych.60.110707.163612] [Medline: 18817479]
143. Asselmann E, Wittchen HU, Lieb R, Beesdo-Baum K. A 10-year prospective-longitudinal study of daily hassles and incident psychopathology among adolescents and young adults: interactions with gender, perceived coping efficacy, and negative life events. Soc Psychiatry Psychiatr Epidemiol 2017;52(11):1353-1362. [doi: 10.1007/s00127-017-1436-3] [Medline: 28889251]
144. Zheng H, Cooke EM, Li K, Zheng Y. Capturing hassles and uplifts in adolescents' daily lives: links with physical and mental well-being. J Youth Adolesc 2023;52(1):177-194. [doi: 10.1007/s10964-022-01682-6] [Medline: 36180660]
145. Chamberlain K, Zika S. The minor events approach to stress: support for the use of daily hassles. Br J Psychol 1990;81(4):469-481. [doi: 10.1111/j.2044-8295.1990.tb02373.x] [Medline: 2279232]
146. Kanner AD, Coyne JC, Schaefer C, Lazarus RS. Comparison of two modes of stress measurement: daily hassles and uplifts versus major life events. J Behav Med 1981;4(1):1-39. [doi: 10.1007/bf00844845]
147. Heron KE, Everhart RS, McHale SM, Smyth JM. Using mobile-technology-based ecological momentary assessment (EMA) methods with youth: a systematic review and recommendations. J Pediatr Psychol 2017;42(10):1087-1107. [doi: 10.1093/jpepsy/jsx078] [Medline: 28475765]
148. Nock MK, Prinstein MJ, Sterba SK. Revealing the form and function of self-injurious thoughts and behaviors: a real-time ecological assessment study among adolescents and young adults. J Abnorm Psychol 2009;118(4):816-827 [FREE Full text] [doi: 10.1037/a0016948] [Medline: 19899851]
149. White ME, Shih JH. A daily diary study of co-rumination, stressful life events, and depressed mood in late adolescents. J Clin Child Adolesc Psychol 2012;41(5):598-610. [doi: 10.1080/15374416.2012.706518] [Medline: 22889264]
150. Maciejewski DF, van Lier PAC, Branje SJT, Meeus WHJ, Koot HM. A daily diary study on adolescent emotional experiences: measurement invariance and developmental trajectories. Psychol Assess 2017;29(1):35-49. [doi: 10.1037/pas0000312] [Medline: 27099977]
151. Chow SM, Ram N, Boker SM, Fujita F, Clore G. Emotion as a thermostat: representing emotion regulation using a damped oscillator model. Emotion 2005;5(2):208-225. [doi: 10.1037/1528-3542.5.2.208] [Medline: 15982086]
152. Bolger N, Zuckerman A, Kessler RC. Invisible support and adjustment to stress. J Pers Soc Psychol 2000;79(6):953-961. [doi: 10.1037//0022-3514.79.6.953] [Medline: 11138764]
153. McKee KL, Rappaport LM, Boker SM, Moskowitz DS, Neale MC. Adaptive equilibrium regulation: modeling individual dynamics on multiple timescales. Struct Equ Modeling 2018;25(6):888-905 [FREE Full text] [doi: 10.1080/10705511.2018.1442224] [Medline: 30416330]
154. Boker SM. Adaptive equilibrium regulation: a balancing act in two timescales. J Pers Oriented Res 2015;1(1-2):99-109 [FREE Full text] [doi: 10.17505/jpor.2015.10] [Medline: 27066197]
155. Kim J, Nakamura T, Kikuchi H, Sasaki T, Yamamoto Y. Co-variation of depressive mood and locomotor dynamics evaluated by ecological momentary assessment in healthy humans. PLoS One 2013;8(9):e74979. [doi: 10.1371/journal.pone.0074979] [Medline: 24058642]
156. Triantafillou S, Saeb S, Lattie EG, Mohr DC, Kording KP. Relationship between sleep quality and mood: ecological momentary assessment study. JMIR Ment Health 2019;6(3):e12613 [FREE Full text] [doi: 10.2196/12613] [Medline: 30916663]
157. Morshed MB, Kulkarni SS, Saha K, Li R, Roper LG, Nachman L, et al. Food, mood, context: examining college students' eating context and mental well-being. ACM Trans Comput Healthc 2022;3(4):1-26. [doi: 10.1145/3533390]
158. Pham LT, Hernandez R, Spruijt-Metz D, Gonzalez JS, Pyatak EA. Movement matters: short-term impacts of physical activity on mood and well-being. J Behav Med 2023;46(5):781-790 [FREE Full text] [doi: 10.1007/s10865-023-00407-9] [Medline: 36939975]

Abbreviations

ARI: Affective Reactivity Index

XSL•FO
RenderX

DSEM: dynamic structural equation modeling
EAT: Eating Attitudes Test
IPPA: Inventory of Parent and Peer Attachment
MINI-KID: Mini International Neuropsychiatric Interview for Children and Adolescents
MSSD: mean squared successive difference
SDQ: Strengths and Difficulties Questionnaire
SSQ: Social Support Questionnaire

Edited by A Schwartz; submitted 28.Apr.2025; peer-reviewed by R Nooripour, AKC Chau; comments to author 07.Jul.2025; accepted 16.Jan.2026; published 18.Feb.2026.

Please cite as:
Mastronardi C, Powers J, Menna R, Rappaport LM
Contribution of Emotion Dynamics to Adolescent Psychosocial Well-Being: Protocol for a Longitudinal Study
JMIR Res Protoc 2026;15:e76333
URL: https://www.researchprotocols.org/2026/1/e76333
doi:10.2196/76333
PMID:

©Carli Mastronardi, Jade Powers, Rosanne Menna, Lance M Rappaport. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 18.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effectiveness of a Digital Awareness App in HIV/AIDS Mitigation Among Transgender Individuals in Rawalpindi District: Protocol for a Quasi-Experimental Study

Muhammad Mudassar Farooq[1], PhD; Shaheer Ellahi Khan[1], PhD; Ali Ahmed[2], PhD

[1]Department of Public Health, Health Services Academy, Islamabad, Pakistan

[2]School of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego, San Diego, CA, United States

Corresponding Author:
Ali Ahmed, PhD
School of Medicine
Division of Infectious Diseases and Global Public Health
University of California, San Diego
UCSD 9500, Gilman Drive, 92039
La Jolla, California
San Diego, CA, MC0507
United States
Phone: 1 619 9539767
Email: aliahmed@health.ucsd.edu

Abstract

Background: HIV/AIDS is a disease associated with stigma and discrimination. This can hinder the adoption of preventive and treatment methods, especially in vulnerable populations, such as the transgender community.

Objective: The primary objectives of this study are to explore awareness barriers related to HIV/AIDS, develop and pilot a mobile-based HIV awareness app, and evaluate its acceptability and usability within the transgender community.

Methods: The research will employ a quasi-experimental design, utilizing a pre- and posttest comparison between an intervention group that will use the mobile app and a comparison group that will not. Phase 1 involves a situational analysis, including key informant interviews, focus group discussions, and a cross-sectional survey. An app will be designed and developed in Phase 2. Phase 3 will comprise a preintervention assessment recruiting 150 transgender people, implementation of the app on the cell phones of 75 transgender people, and a postapp assessment. Statistical techniques will be employed to analyze the captured data and assess the effectiveness of the app.

Results: The recruitment began on August 25, 2025, for the first phase, with the subsequent phases to follow. The data collection and analysis will be completed and finalized by August 31, 2026, following the intervention deployment. No funding was received from any external source for this study.

Conclusions: The results of this study will reveal the effectiveness of a mobile app for the transgender community. These results will determine the continuation and further scale-up of this intervention. The findings will create evidence to inform favorable strategies for vulnerable populations.

International Registered Report Identifier (IRRID): DERR1-10.2196/84610

(JMIR Res Protoc 2026;15:e84610) doi:10.2196/84610

KEYWORDS

HIV/AIDS; digital awareness; mobile app; HIV treatment and prevention; key populations; transgender individuals

Introduction

Overview

The purpose of this study is to promote HIV/AIDS awareness, especially among key populations such as transgender people, helping not only in this high-risk population but also the public adopt prevention methods and reduce the stigma associated with this disease. An estimated 42.3 million people lived with HIV/AIDS globally in mid-2024, among whom 0.33 million were estimated to live in Pakistan [1]. The National AIDS

XSL•FO
RenderX

Control Program (NACP) registered 74,619 people living with HIV at 94 antiretroviral treatment (ART) centers across Pakistan, out of which 51,821 are currently adherent to receive medicine as of December 31, 2024 [2]. The Joint United Nations Programme on HIV/AIDS (UNAIDS) has set ambitious "95-95-95 targets" [1] to reduce the transmission of HIV by the end of 2030. Achieving these targets is challenging with the current strategies, as progress shows [3], particularly due to a lack of proper awareness about the disease and prevention methods—not only in the general population but also in the target populations. The 95-95-95 HIV testing, treatment, and viral suppression targets aim to close gaps in HIV treatment coverage and outcomes [4]. According to these targets, 95% of all people living with HIV globally should have a diagnosis, 95% of people diagnosed should be taking lifesaving ART, and 95% of people on treatment should achieve a suppressed viral load by the end of 2025 [5]. Globally, these percentages were 86%, 77%, and 72% in 2023, respectively. In Pakistan, these statistics are low—by the end of 2022, only 25.8% knew their status, 16.6% were receiving ART, and 6% were virally suppressed [3]. A comprehensive awareness of the causes of HIV may result in reducing new infections in the country.

A rapid rise in HIV prevalence among the transgender community was observed during the latest Integrated Biological and Behavioral Surveillance (IBBS) round (for 2023 and 2024) [6], as compared to the previous survey held in 2016 and 2017 [7]. Recent annual infection trends among key populations show that transgender individuals are more likely to spread infection and are not easily traceable [8]. Viable strategies like condom use and pre-exposure prophylaxis to reduce new infections are necessary [9]. However, awareness and knowledge of the disease are missing components. Even high school students involved in sexual activities without preventive measures put themselves at risk of HIV transmission [10]. No effective solution has been suggested to reduce the risk and adopt strong preventive measures for adolescents. Pakistan continues to experience a rising number of new cases each year from all key and general populations due to a lack of awareness of HIV and a low number of testing services [11]. The government can reduce HIV cases by regulating health care practices, promoting public awareness, supporting counselling, and minimizing transactional sex work [12]. People should be aware of HIV, and access to health facilities should be readily available [13]. However, studies in the country context have not specified how HIV awareness programs should be implemented. Districts including Multan, Chiniot, Sargodha, Rawalpindi, and Dera Ghazi Khan have the highest HIV infection cases [14]. Training health care workers can help reduce stigma and discrimination toward people living with HIV [15]. However, no studies have been found that provide an absolute intervention for reducing the stigma-related factors. Psychological interventions, such as therapeutic alliance and empathy, can improve treatment adherence [16], but techniques for implementing these interventions remain inadequate. Limited household income and a lack of formal education are the main reasons for delayed HIV treatment initiation [17].

Although hospitals provide free testing and treatment for people living with HIV, supported by the Global Fund and the Government of Pakistan [18,19], factors like social stigma, forgetfulness, drug side effects, anxiety, depression, substance abuse, and financial difficulties are the reasons behind poor adherence to ART [20,21]. Only 58% of the registered patients follow up on their treatment in Pakistan, which is a low ratio [22]. Data from the NACP show that 69% of people living with HIV were following their treatment as of September 2024 [23]. Depression is one of the major causes of nonadherence to ART [24,25]. Telehealth improved treatment adherence in people living with HIV, showing a direct relationship between adherence and the quality of their lives [20,26]. Phone call interventions for people living with HIV can help them with treatment adherence, which is a good telehealth technique, but this intervention may not address the stigma around HIV [27]. Patient management, including individual and group counselling, may improve adherence at ART centers [28,29]. Feeling depressed or overwhelmed from daily dosing and a desire to forget about having HIV are the main reported reasons for missing appointments [30]. This study also showed that employed patients are more careful than unemployed patients about their medication. In a review of digital apps for people living with HIV, usability and cost were shown to be hindrances in the African context, despite their effectiveness [31]. Digital interventions for people living with HIV have positively affected adherence to treatment, prevention, and social and behavioral issues [32].

Researchers do not formally publish the studies conducted in the health sector of Pakistan, particularly interventional study designs [33]. Due to a lack of scientific evidence, it is difficult to determine the effectiveness of such interventions. Despite efforts to raise awareness and reduce stigma and discrimination related to HIV/AIDS, barriers to prevention, testing, and treatment—particularly among transgender individuals in Pakistan—are prevalent. Digital interventions can provide a confidential and accessible means to bridge the gaps in preventive measures.

A web-based digital platform, namely Sehat Dost [34], was recently launched but does not specifically address key populations, especially transgender individuals. The proposed awareness mobile app in our study is mobile-based and linked with the web to allow frequent information updates, which will help remove the sociocultural and health care access barriers for key populations at risk of HIV. Audiovisual content in this proposed app will make it easy for people without formal education to understand and gain relevant knowledge and awareness. HIV is a sensitive issue associated with stigma and discrimination, and the existing platforms do not adequately protect user privacy. The proposed app will adopt enhanced data protection protocols. The existing digital information focuses on testing rather than awareness and behavior change, whereas HIV prevention is a main component that needs to be addressed. This study hypothesizes that a mobile-based awareness intervention will enhance knowledge about HIV/AIDS in the transgender community.

Theoretical Framework

The Health Belief Model, originally developed in the 1950s and updated into the 1980s by Rosenstock [35], is widely used to

XSL•FO
RenderX

predict health behaviors. Our digital awareness app will be designed to address this model's constructs. Informational content in the app may increase perceived susceptibility and severity by providing real-life examples. Practical tips and access to services may reduce perceived barriers and enhance self-efficacy, as presented in Table 1.

Table 1. Conceptual table linking the Health Benefit Model constructs to the proposed app's features and outcome indicators.

Perceived Construct	App Feature	Output
Susceptibility	Risk quiz	Awareness about personal risk
Severity	Testimonials	HIV seriousness understanding
Benefits	Information on prevention and testing	Motivation for testing, prevention, and treatment
Barriers	List of health facilities, FAQs[a]	Reduces fear and misconceptions
Self-efficacy	Quiz, guide	Increases confidence in adapting to prevention
Cues to action	Reminders and alert messages	Participation increase

[a]FAQ: frequently asked question.

Intervention theories of change consider factors that predict behaviors and propose general mechanisms of behavior change [36]. The theory of change will follow a pathway from input to impact trail. App development, content creation, and community engagement are the input variables, whereas digital outreach, awareness messaging, and feedback loops are core activities. The number of users reached, content viewed, and user feedback will be outputs of the intervention. Improved knowledge of the community, increased awareness, and a positive attitude will be considered as short-term outcomes of the mobile-based app, whereas assessing stigma reduction, behavior change, and safer sex practices requires a longer period. Reduction in risky behaviors and new HIV infections presents a long-term impact of this digital intervention. With this pathway in mind, this study focuses on measuring knowledge and awareness as the short-term changes. A 3-to-4-month period of implementation, as proposed in this study, is useful for short-term changes, whereas an extended period is required for bigger behavioral changes. The theory of change will provide a useful lens to interpret the findings. These findings will guide future interventions.

Rationale

In the current digital era, there is a need to introduce digital strategies for delivering disease-related messages and guiding people, especially high-risk communities, on how to prevent disease. Transgender and other communities, including men who have sex with men (MSM), female sex workers, and people who inject drugs, are among underserved communities that need special attention to improve awareness of HIV disease, as they are at higher risk due to behavioral factors.

Objectives

This research aims to improve the well-being of transgender individuals through digital intervention in increasing awareness and promoting the adoption of preventive measures against HIV. More specifically, we have the following objectives: (1) to explore existing awareness and barriers regarding HIV/AIDS in the transgender community through situational analysis; (2) to develop and pilot a mobile app for awareness and stigma-related factors associated with HIV among the transgender community in Rawalpindi district; and (3) To evaluate the acceptability and utilization of a digital tool in improving public health awareness.

Methods

Overview

This research will comprise 3 phases: situational analysis, intervention design, and execution and evaluation. Figure 1 shows an overview of the research design. A transgender leader working closely with the community, a government official from the HIV program, and an IT expert who is an app developer have been consulted and will be on board throughout the study. This board will provide guidance through all study phases.

XSL•FO
RenderX

Figure 1. Flow diagram of a quasi-experimental study on HIV mitigation in the transgender community of Rawalpindi, Pakistan, conducted from August 2025 to August 2026.

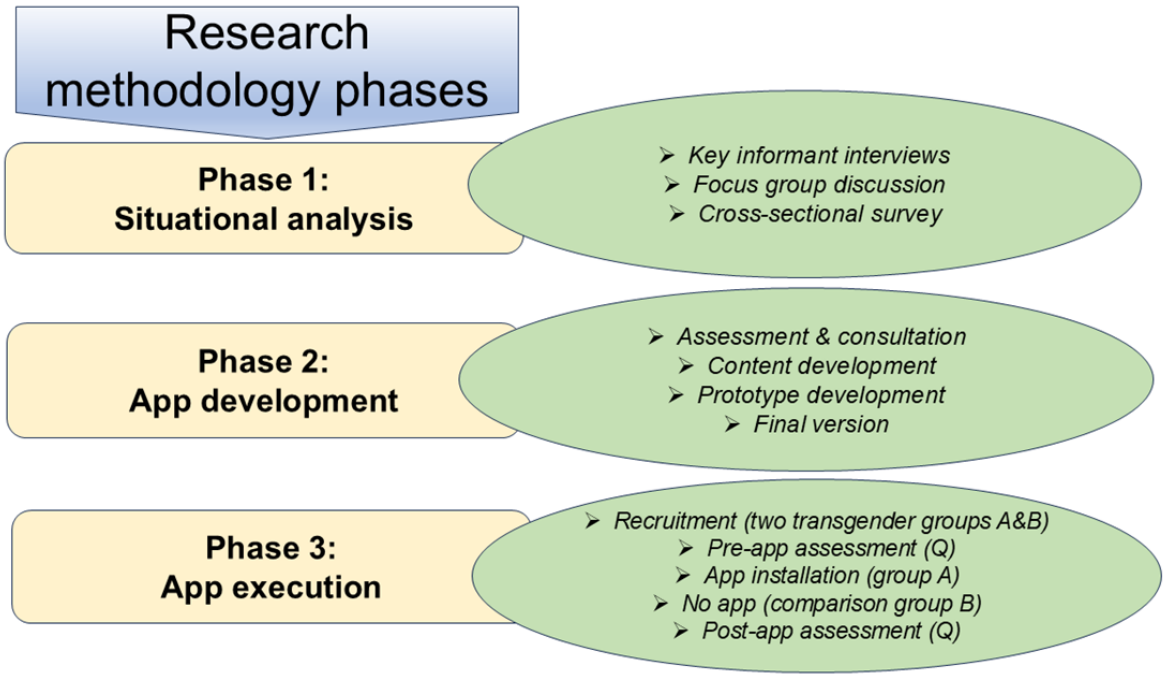

Phase 1: Situational Analysis

Explanation

The main objective of situational analysis is to explore existing awareness and barriers regarding HIV/AIDS among the transgender community. Identifying the key stakeholders is an important step in situational analysis, wherein concerned national and international organizations like the Association for Social Development, the Association of People Living with HIV, UNAIDS, the World Health Organization (WHO), the Dareecha Male Health Society (DMHS), and the Sathi Foundation will be contacted to explore the ideas and ground realities. Trends and locations/hotspots will be acquired from the NACP and DMHS. Purpose sampling techniques are suitable for key informant interviews and focus group discussions [37,38], whereas snowball sampling for a cross-sectional survey [39] will be adapted in this phase of the study.

Key Informant Interviews

In this component, 10 key persons, including transgender community leaders, health care providers, and nongovernmental organization (NGO) representatives, will be engaged to gain an understanding of the challenges and barriers related to HIV/AIDS awareness and health care access. The participants will be categorized into transgender leaders (4 interviews), government officials (2 interviews), United Nations stakeholders, and NGOs (2+2 interviews). A purposive sampling technique will be used to contact the interviewees. Open-ended questions will be asked related to community knowledge, behaviors, HIV services and policies, and adaptation of the digital intervention.

Focus Group Discussion

Two groups of transgender individuals will be engaged to explore information related to behaviors and mobile apps. These groups will be selected from key community members in the Rawalpindi district using purposive sampling. The first group will comprise 5 participants aged between 16 and 30 years, and the second group will have a similar number of participants aged 30 years and up. Based on the findings of this focus group discussion, a questionnaire will be developed to administer a cross-sectional survey, which will help us gain a better understanding of the community's lived experiences.

Cross-Sectional Survey

A small sample size cross-sectional survey will be conducted by recruiting 50 transgender individuals as a fixed sample from hotspots in the Rawalpindi district, with data collected through a snowball sampling technique at a single time point. Several cross-sectional studies have been conducted with sample sizes ranging from 30 to 50 as pilot or exploratory works [40,41]. This practice will assess the basic variables, including mobile phone availability and the knowledge, attitudes, and behaviors of the transgender community. The sample size is enough to estimate the preliminary study outcomes. A small sample size is more suitable in studies involving sensitive diseases [42]. The questionnaire will be tested with 10 individuals with similar characteristics to ensure clarity.

Thematic Analysis

Thematic analysis will be applied to the qualitative data obtained from the key informant interviews and focus group discussions to extract deep insights and ensure that the digital intervention is designed based on real needs and experiences. Descriptive statistics, including percentages, mean, and median, will be applied to the cross-sectional survey data. Furthermore,

XSL•FO
RenderX

chi-square tests and regression analysis will be used to explore associations between knowledge levels and demographic factors. This analysis will establish our baseline understanding for designing an effective digital intervention.

Phase 2: Intervention Development

Explanation

The second phase of research will focus on designing and developing an intervention, a new mobile-based app. The following steps will be taken in the second phase of the study.

Assessment and Consultation

Based on Phase 1 and its findings, a mobile app will be developed to provide culturally appropriate, easy-to-understand information on HIV/AIDS prevention, treatment, and testing services. Needs, preferences, messages, and other related material will be consulted with stakeholders. Furthermore, the awareness gap, app language, message format, access to smartphones and the internet, and most importantly, user privacy and confidentiality, will be considered. Health information specialists and IT personnel will also be consulted for designing and developing the app.

Content of the App

Content development is the core component of this app, as it can either support awareness or discourage the use of the digital platform on this sensitive issue. A cross-sectional survey from Phase 1 and meetings with key stakeholders will support the development of content for the mobile app. HIV transmission and prevention, ART adherence and treatment literacy, testing and linkage care centers, mental health and stigma support, self-assessment tool, quiz, push notifications, and frequently asked questions (FAQs) are the main modules of this app. Recommendations from key stakeholders and transgender individuals will be valuable for incorporation into the app. The audiovisual content will be designed to be accessible to all users, including those with and without formal education, and will be available in both English and Urdu.

Prototype Development

A minimum viable product of the digital intervention will be designed by the IT personnel hired for this purpose and implemented for a small sample size of 10 transgender individuals. Basic content sections, including awareness messages, chatbot, offline access, and adherence support, will be included. This pilot will give feedback about the security, compatibility, and user-friendliness of the app.

Final Version

At this stage, feedback from a small-scale test will be integrated to develop the final version of the mobile app. The app developer will fix bugs, optimize app speed, and ensure compatibility with Android devices. A manual for this app will also be developed along with back-end data protocols.

Phase 3: Intervention and Evaluation

Explanation

The third phase of research will comprise the intervention execution. The following methods will be adopted.

Study Design

A quasi-experimental study will be adopted to evaluate the effectiveness of the digital intervention on HIV awareness among key populations, particularly the transgender community. Propensity score matching (PSM) will be applied to create statistically similar baseline groups for the intervention and nonintervention categories. This technique will be used for group allocations to reduce bias, not as a sampling method. The difference-in-differences technique will be used to evaluate the effectiveness of the digital HIV awareness app. The app will not be installed in the cell phones of the comparison group during the observation period, although other routine services will be provided accordingly. Both groups will be monitored during the intervention period to ensure data authenticity.

Area and Duration of the Study

Two groups of transgender individuals from Rawalpindi district will be enrolled with their consent; participants will have some awareness of HIV and engage in behaviors that place them at risk of contracting HIV. They will be interviewed, and data will be collected using a structured questionnaire. Then, the digital app will be installed on the cell phones of the intervention group. The facilitator will provide 1 hour of hands-on training to the intervention group, scheduled in convenient sessions, regarding the use of the app. After 3 months, the same questions will be asked, and any changes in their behavior will be observed across both the intervention and comparison groups. During the first month of this phase, a baseline survey will be conducted to assess existing awareness levels in both the intervention group and the comparison group. The intervention phase will last for approximately 3 months, during which participants in the intervention group will actively engage with the app's educational content. A follow-up survey will be conducted with the same participants within 1 month. This will measure changes in awareness and compare the results between the 2 groups. The final month will be devoted to data analysis and presentation of findings for review.

Data Sources

Data from the national program will be obtained to review current estimates and the registration status of people living with HIV. The hotspots will be identified based on patient burden in the area. Primary data will be collected from the transgender community of Rawalpindi for quantitative analysis.

Study Population

Preliminary results from IBBS (round 2023-2024) showed a 2.8% increase in transgender individuals compared to the 2016-2017 round, along with increases of 1.5% in MSM and 1.6% in female sex workers and a 10.6% decrease in people who inject drugs [6,7]. This indicates that the transgender community is less likely to adopt preventive measures compared to other key populations. Rawalpindi, Kasur, Sheikhupura, Sialkot, and Hyderabad are cities with a high number of new cases among transgender individuals in 2024 [2]. The registered cases at the DMHS are residents of Rawalpindi, while community-based organizations (CBOs) in other high-burden districts also report transfer-in cases from various cities. Therefore, transgender individuals from high-risk areas of

XSL•FO
RenderX

Rawalpindi will be selected based on predefined criteria. Transgender individuals from Rawalpindi will be recruited from different hotspots, CBOs, and outreach services. A purposive sampling technique applying the PSM method will be used.

Technique and Size Calculation

A purposive sampling technique will be used, and PSM will be applied to balance the 2 groups for a more valid comparison. A standard power of 80% and a 95% confidence level will be used to minimize errors, and the expected effective size will be set at 0.5. High statistical power will be used, which means an 80% chance of detecting the impact of awareness that is high. A moderate effect size (d=0.5) means that the expected impact of the digital app on HIV/AIDS awareness is noticeable and meaningful but not extreme. This level suggests that, on average, the intervention group will score 0.5 standard deviations higher than the comparison group. It is assumed that $\sigma=1$ for the sample size calculation, using it as an assumption, not as a measured value. The exact standard deviation will be found after collecting baseline data.

The sample size formula resulted in 63 individuals in each group for the study. Considering the dropout scenarios, 75 participants will be recruited.

The sample size calculation: where



$Z_{\alpha/2}$ = z score for the chosen confidence level (eg, 1.96 for 95% confidence)

Z_{β} = z score for power (eg, 0.84 for 80% power)

= estimated standard deviation of the outcome variable, assumed 1

Δ = expected difference between group means (effect size × standard deviation)





Sample size = 63

Sample Recruitment

This study will recruit the participants into 2 groups based on certain inclusion and exclusion criteria, particularly applicable to the intervention group. The eligibility criteria are as follows: (1) transgender individuals aged 18 years and above; (2) residents of Rawalpindi; (3) own a smartphone; and (4) are willing to install the mobile app. The participants must have sufficient digital literacy to use the app and provide informed consent. Individuals who have prior training in HIV/AIDS awareness programs, those unwilling to provide consent, and those lacking basic operational literacy for smartphone apps will be excluded.

Data Collection Tool

A structured questionnaire will be developed and adapted based on coordination, meetings, and formal interviews of the key stakeholders working in the HIV community. WHO assessment tools will also assist in preparing the questionnaire. This primary data collection tool will include basic demographics, multiple-choice questions, and Likert scale items to evaluate the participants' baseline and postintervention knowledge, attitudes, and perceptions about HIV/AIDS. Consent forms for each participant will include thumbprints and signatures.

Data Collection Techniques

A structured questionnaire will be developed and administered to participants to gather data for this study. This questionnaire will be given to both intervention and comparison groups to assess baseline knowledge of HIV/AIDS before app installation and to measure changes in awareness after the intervention. The survey will be conducted using well-designed forms to ensure efficient data collection. Community members will be recruited through outreach, and the 2 groups will be evaluated for effectiveness in a quasi-experimental design [43]. Participants will be recruited through a community-based outreach program and collaboration with key stakeholders involved in the HIV and health sectors. A pretest of the questionnaire will be conducted with a subset of the population to ensure its reliability and clarity. Any ambiguities or difficulties in the questionnaire will be addressed before it is fully deployed. One month will be allocated to gather baseline information and conduct meetings before installing the app. Participants will be recruited from community-based organizations, homes, gatherings of key populations, and some designated health centers. Measures will be taken to ensure consistency and reduce bias throughout the process.

Pretest or Pilot Study

The structured questionnaire will be administered to a subgroup of 10 people who meet the same inclusion and exclusion criteria to assess clarity, relevance, and ease of understanding. The results of this pilot study will help to refine the study design, ensuring that questions are unambiguous, the intervention is effective, and the data collection process runs smoothly. Necessary modifications will be made before launching the full study.

Reliability and Validity

Both the questionnaire and digital intervention will be pretested in a pilot study with 20 participants of similar characteristics to ensure consistent understanding. A test-retest method will check consistent responses from the subset of participants at different time intervals. The scores from both times will be compared on the Pearson correlation coefficient (r), with values greater than 0.6 considered acceptable for reliability. The Cronbach alpha will be used to measure the internal consistency of the questionnaire. This technique is commonly used to assess internal consistency, and expert reviews will ensure content validity [44]. Cronbach alpha is considered more reliable than other methods for measuring internal consistency [45]. If a reliability test yields an $\alpha \geq 0.7$ based on full baseline data, this will indicate that our tool is suitable. The WHO HIV knowledge assessment tool will also be considered when preparing the questionnaire. Pre- and posttest scores will be analyzed to

XSL•FO
RenderX

validate whether the digital intervention effectively improves awareness.

Data Analysis Plan

Descriptive and inferential statistical methods will be applied to the collected data to assess the effectiveness of the digital intervention in improving HIV/AIDS awareness. Descriptive statistics, including mean, standard deviation, frequency, and percentage, will be used to summarize the participants' basic demographic characteristics. Pre- and posttest scores within the intervention group will be compared using paired *t* tests. Independent *t* tests will be used to compare outcomes between the study groups. PSM will be applied to control for baseline differences between the 2 groups. Cronbach alpha will be calculated to measure internal consistency and evaluate the reliability of the questionnaire. A test-retest method will be applied, and the Pearson correlation coefficient will assess the stability of the questionnaire over time. Regression analysis will be carried out to identify the factors influencing HIV/AIDS awareness improvement. SPSS and Microsoft Excel (IBM Corp) software will be used for data analysis. The results will be presented in tables, graphs, and charts for interpretation. $P<.05$ will be considered statistically significant.

Sustainability

The proposed mobile-based app will be linked to a website to enable content updates. Login details and usage information stored on the website will be used to measure app performance and success. A feedback survey on the app content will help guide improvements. This app will be monitored on various indicators, including number of app downloads, user registration or logins, session frequency and duration, content completion rates, clickstream data, activity participation, push notification response rates, feedback submitted, and repeat usage or return rate.

Expected Outcomes

The targeted community is expected to gain an improved understanding of HIV and its transmission. The participants will become more aware of prevention methods, treatment access, care, and support services. Relevant, accessible, and inclusive content will enhance community engagement in digital public health initiatives. It is anticipated that knowledge and behavior-based indicators will be improved. Positive feedback from stakeholders will indicate readiness to scale up this intervention.

Ethical Considerations

This research was approved by the Ethical Review Board and Graduate Research Management Council of Health Services Academy (00022/HSA/PhD-2022) on July 22, 2025. The consent form, available in both English and Urdu, will cover 2 sections, including information and consent. Participant confidentiality will be strictly upheld, and their data will be safeguarded. A clear message within the app stating, "your data is encrypted," will help to build trust and encourage engagement.

Results

Ethical Review Board approval was obtained on July 22, 2025. Recruitment of the 10 key informant interviews occurred between August 25 and 30, 2025, followed by the focus group discussion on September 8, 2025. Recruitment of the 50 transgender people for the cross-sectional survey took place from October 1 to 20, 2025. The app development process occurred between December 1, 2025, and January 5, 2026. The recruitment of 150 participants and the app installation process for 75 participants will be completed by February 28, 2026, followed by an assessment period until June 14, 2026. A follow-up survey will be conducted from June 15 to 30, 2026, and the report will be completed on August 31, 2026. Pre- and postintervention surveys will be compared using the appropriate statistical methods. This analysis will determine the intervention's effectiveness. The study results will provide a foundation for policymakers to serve at-risk and underserved communities. Given the sensitivity of stigma-related disease, understanding the community context is important for carrying out any intervention. Descriptive statistics will guide us about the demographics of the participants, and inferential statistics will give us key information about awareness levels and behaviors. The results will be displayed in both tables and graphs to provide a clear understanding of awareness levels and changes over time in both the intervention and control groups.

Discussion

Expected Findings

It is anticipated that a mobile-based app for the transgender community in Rawalpindi will increase community awareness and guide them to access prevention and treatment services. The culturally sensitive content of this app will enhance the participation and acceptability level among transgender individuals on digital platforms. The intervention group is expected to become more aware of the disease and its prevention and treatment methods as compared to the control group. App users are also expected to share HIV-related knowledge with other community members and guide them to adopt preventive measures by avoiding risky behaviors.

Studies on digital platforms for marginalized communities show some improvement in their desired objectives. The availability of studies for transgender individuals in the country context is limited. Therefore, specific evidence is not available in this context about the success of the interventions. However, findings from available studies helped build the concept of our study. The results of this study will generate evidence on the use of mobile apps within the transgender community. Insights and cultural aspects of the transgender community will support the development of policies for their well-being.

Limited HIV awareness among the key population, particularly transgender individuals in Pakistan, highlights the need for a digital app considering stigma sensitivity. HIV is an urgent and sensitive public health issue that needs to be addressed by introducing new but authentic interventions. This is the first community-based and locally contextual mobile-based app that will improve awareness, encourage HIV testing, and reduce

XSL•FO
RenderX

misconceptions about HIV. While a mobile-based app for transgender individuals is a new idea for this community, the rigorous methodology applied in this study will support science-based decision-making.

A mixed methods approach is the main strength of this study, as the situational analysis provides the baseline for the quasi-experimental study of a hard-to-reach community. This research also emphasizes ethical considerations aligned with community norms, culture, and sensitivities, strengthening the expected outcomes. Technological literacy issues may hinder the process, but training participants and audiovisual aids will mitigate such issues. The study findings have the potential to support scale-up of the intervention to other key populations, including MSM and female sex workers.

The results of this study will be disseminated in the first week of September 2026 in a seminar to be held at the Health Services Academy, to which focal people from the NACP, UNAIDS, WHO, CBOs, and other key stakeholders will be invited. Further, the results will be submitted to a peer-reviewed journal for publication.

Limitations and Mitigation

Recruiting transgender individuals using the probability sampling method is impractical due to hidden hotspots, stigma, discrimination, and cultural aspects. Therefore, this study adopted snowball sampling for the cross-sectional survey. A comprehensive baseline data set on relevant covariates, along with cofounders, will be collected by ensuring sufficient overlap and appropriate matching methods. If PSM does not achieve an adequate balance between the intervention and control groups, the study will proceed with stratified analysis to reduce bias. In this approach, the sample will be divided into homogeneous subgroups called strata based on age group, education, or baseline knowledge. Multivariable regression analysis may also be used to identify any imbalance in covariates, which will isolate the effect of the app from other factors.

Conclusion

This study is expected to provide insights into improving HIV/AIDS awareness among transgender individuals in Rawalpindi through a mobile app–based intervention. This study also has the potential to support scale-up of the intervention in subsequent phases to achieve improved health outcomes. The findings are expected to guide future planning in the public health sector related to HIV/AIDS, particularly for vulnerable populations. It will also help reduce stigma and discrimination by strengthening community engagement.

Acknowledgments

The authors would like to thank the Health Services Academy for granting permission to conduct this study among transgender individuals. We extend our gratitude to all individuals and organizations, including the Dareecha Male Health Society and the National HIV Program, for their positive involvement in the proposal. We also acknowledge the University of California for supporting open access in the publication of the research protocols. No external financial support or grants were received for this study from any individual or organization. The authors declare the use of generative artificial intelligence (AI) in the research and writing process. According to the Generative AI Delegation Taxonomy (2025), proofreading and editing tasks were delegated to generative AI tools under full human supervision. The generative AI tool used was ChatGPT-5. The responsibility for the final manuscript lies entirely with the authors. The generative AI tools are not listed as authors and do not bear responsibility for the final outcomes. This declaration was submitted under collective author responsibility. Additionally, no generative AI was used for producing scientific content, generating ideas, interpreting results, or any other research-related tasks. The authors solely created all the scientific content. ChatGPT-5 was used only to improve the clarity of a few sentences.

Data Availability

Data confidentiality and availability will be addressed in the informed consent form. The data will be available upon request after the personal information has been anonymized.

Authors' Contributions

MMF, as the principal investigator, developed the concept, designed the methodology, drafted the protocol, and managed the study. SEK supervised the work, provided input on study design, and guided the protocol development. AA reviewed the manuscript, provided feedback, and managed the journal submission. All authors read and approved the final manuscript. For collaboration-related inquiries, readers may also contact the principal author at mudassar9999@yahoo.com.

Conflicts of Interest

None declared.

References

1. HIV and AIDS. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/hiv-aids [accessed 2025-03-08]

XSL•FO
RenderX

2. Current statistics. Government of Pakistan (Common Management Unit).: Government of Pakistan; 2025. URL: https://www.cmu.gov.pk/nacp-national-aids-control-programme/ [accessed 2025-03-07]
3. An evaluation of the contribution of the UNAIDS joint program to strengthening HIV and primary health care outcomes. UNAIDS. URL: http://www.unaids.org/en/whoweare/evaluation [accessed 2024-11-16]
4. Understanding measures of progress towards the 95-95-95 HIV testing, treatment and viral suppression targets. UNAIDS. URL: https://www.unaids.org/sites/default/files/media_asset/progress-towards-95-95-95_en.pdf [accessed 2024-11-16]
5. HIV and AIDS. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/hiv-aids [accessed 2024-11-16]
6. Bhatti W. Data of Pakistan's $3.2m HIV survey under fire for 'fabrication'. The News. Islamabad; 2025. URL: https://www.thenews.com.pk/print/1322070-data-of-pakistan-s-3-2m-hiv-survey-under-fire-for-fabrication.com [accessed 2025-06-17]
7. Integrated biological and behavioral surveillance. Government of Pakistan (Common Management Unit). 2017. URL: https://www.cmu.gov.pk/resource-center/ [accessed 2025-06-16]
8. Govender RD, Hashim MJ, Khan MA, Mustafa H, Khan G. Global epidemiology of HIV/AIDS: a resurgence in North America and Europe. J Epidemiol Glob Health 2021 Sep;11(3):296-301 [FREE Full text] [doi: 10.2991/jegh.k.210621.001] [Medline: 34270183]
9. Mody A, Sohn AH, Iwuji C, Tan RKJ, Venter F, Geng EH. HIV epidemiology, prevention, treatment, and implementation strategies for public health. Lancet 2024 Feb;403(10425):471-492. [doi: 10.1016/s0140-6736(23)01381-8]
10. Obeagu E, Obeagu G, Ede M, Odo E, Buhari H. Translation of HIV/AIDS knowledge into behavior change among secondary school adolescents in Uganda: A review. Medicine 2023;102(49). [doi: 10.1097/md.0000000000036599]
11. Marfani W, Khan H, Sadiq M, Outani O. The rise in HIV cases in Pakistan: Prospective implications and approaches. Ann Med Surg (Lond) 2022 Sep;81:104492 [FREE Full text] [doi: 10.1016/j.amsu.2022.104492] [Medline: 36106067]
12. Aizaz M, Abbas FA, Abbas A, Tabassum S, Obeagu EI. Alarming rise in HIV cases in Pakistan: Challenges and future recommendations at hand. Health Sci Rep 2023 Aug 28;6(8):e1450 [FREE Full text] [doi: 10.1002/hsr2.1450] [Medline: 37520460]
13. Ullah N, Adil M, Khan M, Junaid M. The health education/health awareness of communal members regarding HIV/AIDS. Int J Health Educ Soc 2024;7(3):66-80 [FREE Full text]
14. Amjad M, Rehman Y, Hussain Babar Y, Hussain Ali Shah S, Ali M, Awais Fiaz M, et al. Revealing the magnitude of HIV prevalence and unraveling its risk factors in the enclave of district DG Khan. J Asian Dev Studies 2024 Dec 01;13(4):886-898. [doi: 10.62345/jads.2024.13.4.71]
15. Obeagu E, Obeagu G. Unmasking the truth: Addressing stigma in the fight against HIV. Elite J Public Health 2024;2(1):8-22.
16. Spaan P, van Luenen S, Garnefski N, Kraaij V. Psychosocial interventions enhance HIV medication adherence: A systematic review and meta-analysis. J Health Psychol 2020 Sep 08;25(10-11):1326-1340 [FREE Full text] [doi: 10.1177/1359105318755545] [Medline: 29417851]
17. Ali H, Zakar R, Junaid K, Khan A, Fischer F. Frequency and reasons for delayed treatment initiation after HIV diagnosis: cross-sectional study in Lahore, Pakistan. BMC Public Health 2021 May 27;21(1):1000 [FREE Full text] [doi: 10.1186/s12889-021-11031-0] [Medline: 34044793]
18. Ahmed A, Saqlain M, Umair MM, Hashmi FK, Saeed H, Amer M, et al. Stigma, social support, illicit drug use, and other predictors of anxiety and depression among HIV/AIDS patients in Pakistan: A cross-sectional study. Front Public Health 2021 Sep 30;9:745545 [FREE Full text] [doi: 10.3389/fpubh.2021.745545] [Medline: 34660521]
19. Shah SA, Altaf A, Vermund SH. Challenges of antiretroviral treatment for human immunodeficiency virus infection in Pakistan. J Pak Med Assoc 2021 Mar;71(3):788-790 [FREE Full text] [Medline: 34057921]
20. Esmaeili ED, Azizi H, Dastgiri S, Kalankesh LR. Does telehealth affect the adherence to ART among patients with HIV? A systematic review and meta-analysis. BMC Infect Dis 2023 Mar 17;23(1):169 [FREE Full text] [doi: 10.1186/s12879-023-08119-w] [Medline: 36932376]
21. Ahmed A, Dujaili JA, Jabeen M, Umair MM, Chuah L, Hashmi FK, et al. Barriers and enablers for adherence to antiretroviral therapy among people living with HIV/AIDS in the era of COVID-19: a qualitative study from Pakistan. Front Pharmacol 2021 Jan 28;12:807446 [FREE Full text] [doi: 10.3389/fphar.2021.807446] [Medline: 35153763]
22. Ahmed A, Dujaili JA, Chuah LH, Hashmi FK, Le LKD, Chatha ZF, et al. Cost-effectiveness analysis of pharmacist adherence interventions in people living with HIV/AIDS in Pakistan. Healthcare (Basel) 2023 Sep 01;11(17):2453 [FREE Full text] [doi: 10.3390/healthcare11172453] [Medline: 37685487]
23. National AIDS control program. Government of Pakistan (Common Management Unit). 2024. URL: https://www.cmu.gov.pk/nacp-national-aids-control-programme/ [accessed 2024-11-08]
24. Abbas Q, Nisa M, Khan MU, Anwar N, Aljhani S, Ramzan Z, et al. Brief cognitive behavior therapy for stigmatization, depression, quality of life, social support and adherence to treatment among patients with HIV/AIDS: a randomized control trial. BMC Psychiatry 2023 Jul 25;23(1):539 [FREE Full text] [doi: 10.1186/s12888-023-05013-2] [Medline: 37491185]
25. Gokhale RH, Weiser J, Sullivan PS, Luo Q, Shu F, Bradley H. Depression prevalence, antidepressant treatment status, and association with sustained HIV viral suppression among adults living with HIV in care in the United States, 2009-2014. AIDS Behav 2019 Dec 31;23(12):3452-3459. [doi: 10.1007/s10461-019-02613-6] [Medline: 31367965]

XSL•FO
RenderX

26. Saragih ID, Tonapa SI, Osingada CP, Porta CM, Lee B. Effects of telehealth-assisted interventions among people living with HIV/AIDS: A systematic review and meta-analysis of randomized controlled studies. J Telemed Telecare 2021 Dec 30;30(3):438-450. [doi: 10.1177/1357633x211070726]
27. Osingada CP, McMorris B, Piehler TF, Tracy MF, Porta CM. Acceptability, feasibility, and efficacy of telehealth interventions for HIV testing and treatment in adult populations. Telemed J E Health 2023 Jul 01;29(7):1001-1013. [doi: 10.1089/tmj.2022.0424] [Medline: 36472470]
28. Reif LK, Abrams EJ, Arpadi S, Elul B, McNairy ML, Fitzgerald DW, et al. Interventions to improve antiretroviral therapy adherence among adolescents and youth in low- and middle-income countries: a systematic review 2015-2019. AIDS Behav 2020 Oct 09;24(10):2797-2810. [doi: 10.1007/s10461-020-02822-4] [Medline: 32152815]
29. Okonji EF, Mukumbang FC, Orth Z, Vickerman-Delport SA, Van Wyk B. Psychosocial support interventions for improved adherence and retention in ART care for young people living with HIV (10-24 years): a scoping review. BMC Public Health 2020 Dec 01;20(1):1841 [FREE Full text] [doi: 10.1186/s12889-020-09717-y] [Medline: 33261566]
30. de Los Rios P, Okoli C, Punekar Y, Allan B, Muchenje M, Castellanos E, et al. Prevalence, determinants, and impact of suboptimal adherence to HIV medication in 25 countries. Prev Med 2020 Oct;139:106182 [FREE Full text] [doi: 10.1016/j.ypmed.2020.106182] [Medline: 32593732]
31. Ukoaka BM, Ugwuanyi EA, Ukueku KO, Ajah KU, Udam NG, Daniel FM, et al. Digital tools for improving antiretroviral adherence among people living with HIV in Africa. J Med Surg Public Health 2024 Apr;2:100077. [doi: 10.1016/j.glmedi.2024.100077]
32. Barzegary A, Karimi A, Mehraeen E, Vahedi F, Qaderi K, M. Tantuoyir M, et al. Mobile applications in HIV self-management: A systematic review of scientific literature. AIDSRev 2022 Mar 16;24(1). [doi: 10.24875/aidsrev.21000025]
33. Kazi AM, Qazi SA, Ahsan N, Khawaja S, Sameen F, Saqib M, et al. Current challenges of digital health interventions in Pakistan: mixed methods analysis. J Med Internet Res 2020 Sep 3;22(9):e21691. [doi: 10.2196/21691]
34. Sehatdost. UNDP. 2025. URL: https://www.sehatdost.pk/home [accessed 2025-03-03]
35. Joorbonyan H, Ghaffari M, Rakhshanderou S. Peer-led theoretically Desinged HIV/AIDS prevention intervention among students: a case of health belief model. BMC Public Health 2022 Jan 05;22(1):8 [FREE Full text] [doi: 10.1186/s12889-021-12445-6] [Medline: 34983478]
36. Meiksin R, Melendez-Torres GJ, Falconer J, Witzel TC, Weatherburn P, Bonell C. Theories of change for e-health interventions targeting HIV/STIs and sexual risk, substance use and mental ill health amongst men who have sex with men: systematic review and synthesis. Syst Rev 2021 Jan 11;10(1):21 [FREE Full text] [doi: 10.1186/s13643-020-01523-2] [Medline: 33423693]
37. Pahwa M, Cavanagh A, Vanstone M. Key informants in applied qualitative health research. Qual Health Res 2023 Dec 30;33(14):1251-1261 [FREE Full text] [doi: 10.1177/10497323231198796] [Medline: 37902082]
38. Nyimbili F, Nyimbili L. Types of purposive sampling techniques with their examples and application in qualitative research studies. Br J Multidiscip Adv Stud 2024 Feb 16;5(1):90-99. [doi: 10.37745/bjmas.2022.0419]
39. Gillani AH, Arshad H, Arshed M, Jairoun A, Shukar S, Akbar J, et al. Cross-sectional assessment of perception and attitude of pharmacy students towards pharmaceutical promotion: a study from developing country, Pakistan. Front Med (Lausanne) 2024 Nov 1;11:1424352 [FREE Full text] [doi: 10.3389/fmed.2024.1424352] [Medline: 39554505]
40. Cendrero-Luengo C, Jiménez-Palomares M, Rodríguez-Mansilla J, Garrido-Ardila EM. Cross-sectional descriptive pilot study on the risk of psychotic disorders among adolescents. Children (Basel) 2021 Oct 14;8(10):916 [FREE Full text] [doi: 10.3390/children8100916] [Medline: 34682181]
41. Hernan V, Duran-Muños I, Jurado M, Faraldo-Cabana A, Oliva-Hernando M, Ortuño-Sorinao I. Quality of life and perceived care of patients in advanced chronic kidney disease consultations: a cross-sectional descriptive study. arXiv Preprint posted online on Jan 2, 2025.
42. Althubaiti A. Sample size determination: A practical guide for health researchers. J Gen Fam Med 2023 Mar 14;24(2):72-78 [FREE Full text] [doi: 10.1002/jgf2.600] [Medline: 36909790]
43. Arije O, Titus R, Olaniran A, Dadi A, Garba D, Okeke E, et al. Effectiveness of community mobilisation models in improving HIV testing services uptake among women and children in Nigeria: A quasi-experimental study. Glob Public Health 2023 Jan 28;18(1):2284880 [FREE Full text] [doi: 10.1080/17441692.2023.2284880] [Medline: 38015746]
44. Trabelsi K, Saif Z, Driller MW, Vitiello MV, Jahrami H. Evaluating the reliability of the athlete sleep behavior questionnaire (ASBQ): a meta-analysis of Cronbach's alpha and intraclass correlation coefficient. BMC Sports Sci Med Rehabil 2024 Jan 02;16(1):1 [FREE Full text] [doi: 10.1186/s13102-023-00787-0] [Medline: 38167160]
45. Amirrudin M, Nasution K, Supahar S. Effect of variability on Cronbach alpha reliability in research practice. J Math Stat Comput 2020 Dec 23;17(2):223-230. [doi: 10.20956/jmsk.v17i2.11655]

Abbreviations

ART: antiretroviral treatment
CBO: community-based organization
DMHS: Dareecha Male Health Society

XSL•FO
RenderX

IBBS: Integrated Biological and Behavioral Surveillance
MSM: men who have sex with men
NACP: National AIDS Control Program
NGO: nongovernmental organization
PSM: propensity score matching
UNAIDS: Joint United Nations Programme on HIV/AIDS
WHO: World Health Organization

Edited by J Sarvestan, A Mavragani; submitted 22.Sep.2025; peer-reviewed by MD Butt, O Igunma; comments to author 17.Nov.2025; revised version received 29.Nov.2025; accepted 05.Dec.2025; published 18.Feb.2026.

Please cite as:
Farooq MM, Khan SE, Ahmed A
Effectiveness of a Digital Awareness App in HIV/AIDS Mitigation Among Transgender Individuals in Rawalpindi District: Protocol for a Quasi-Experimental Study
JMIR Res Protoc 2026;15:e84610
URL: https://www.researchprotocols.org/2026/1/e84610
doi:10.2196/84610
PMID:

©Muhammad Mudassar Farooq, Shaheer Ellahi Khan, Ali Ahmed. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 18.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Digital Intervention for Capturing Real-Time Health Data for Epilepsy Seizure Forecasting: Protocol for the ATMOSPHERE Study

Lauren Thompson[1], BSc, MSc, PhD; Emily Nielsen[1], BSc, MSc, PhD; Emily E V Quilter[1], BSc, MSc; Amelia Slay[1], BSc, MSc; Marceli Wac[2], BSc, PhD; Matthew Wragg[1], BSc, PhD; Samuel Downes[1], BSc, MSc; Phil Tittensor[3], MSc; Leandro Junges[4], BSc, MSc, PhD; Liz Stuart[5], BSc, PhD; Rosie Charles[6], BSc, MSc; Peter Kissack[4,7], MMath; Yasser Qureshi[4,8], BEng, PhD; Amberly Brigden[1], BSc, MSc, PhD

[1]School of Engineering Mathematics and Technology, University of Bristol, Bristol, United Kingdom

[2]Colinear Software Ltd, Bristol, United Kingdom

[3]The Royal Wolverhampton NHS Trust, Wolverhampton, United Kingdom

[4]Centre for Systems Modelling and Quantitative Biomedicine, University of Birmingham, Birmingham, United Kingdom

[5]School of Computing, Ulster University, Belfast, United Kingdom

[6]Neuronostics, Bristol, United Kingdom

[7]School of Mathematics, University of Birmingham, Birmingham, United Kingdom

[8]School of Engineering, University of Warwick, Coventry, United Kingdom

Corresponding Author:
Lauren Thompson, BSc, MSc, PhD
School of Engineering Mathematics and Technology
University of Bristol
1 Cathedral Square
Bristol, BS1 5DD
United Kingdom
Phone: 44 0117 42 82343
Email: lauren.thompson@bristol.ac.uk

Abstract

Background: Epilepsy is a chronic neurological disorder marked by recurrent and apparently unpredictable seizures and associated with premature death, injury, and diminished quality of life. The unpredictability of seizures is a major concern for people with epilepsy. Thus, developing tools for seizure prediction is a research priority. The Artificial Intelligence to Optimise Seizure Prediction to Empower People With Epilepsy (ATMOSPHERE) project focuses on the development and evaluation of seizure forecasting technology involving mobile technology and machine learning to provide personalized seizure forecasting (risk of seizure in the near future). The project is informed by complex intervention frameworks, which recommend phases of development, feasibility study, clinical evaluation, and implementation.

Objective: Objective 1 aims to conduct a feasibility study to test and refine the trial methods for a future clinical trial. Objective 2 aims to test and refine the data collection technology, considering usability and technical performance. Objective 3 aims to collect longitudinal data on seizures and their precipitants to refine seizure forecasting.

Methods: This study is a single-arm, mixed methods feasibility study, testing a prototype of the data collection technology, with phase 2 testing a minimum viable product. In total, 60 participants will be recruited via specialist National Health Service epilepsy clinics. Inclusion criteria are adults with epilepsy, experiencing seizures twice per month, able to consent, and engage with technology. Clinicians will screen and gain consent to contact, with researchers obtaining full consent. Participants will be invited to complete the following study procedures: (1) onboarding, (2) use the data collection technology (phase 1 or 2) in their lived context for up to 6 months, (3) complete patient-reported outcome measures and capture clinical-reported outcome measures at baseline and 3 months, and (4) complete a qualitative interview exploring their views of the data collection technology. A study flow diagram will report recruitment rates (outcome 1), diversity of the recruited sample (outcome 2), barriers and facilitators to recruitment (outcome 3), retention rates (outcome 4), and barriers and facilitators to retention (outcome 5). To assess the data collection technology, quantitative technology use data and qualitative interview data will be analyzed to assess usability (outcome

XSL•FO
RenderX

6) and technical performance (outcome 7) of the data collection technology. These outcomes will inform iterative minimum viable product development and testing cycles with stakeholders.

Results: Recruitment is planned to begin in quarter 1 of 2026, with data collection expected to be completed by quarter 2 of 2027. Data analysis will take place during quarter 3, and the results will be published in quarter 4.

Conclusions: This project aims to improve clinical outcomes for people with epilepsy through seizure forecasting technology. To evaluate clinical outcomes, robust trial methodology is critical. This feasibility study will optimize methods for a future full-scale clinical trial as well as refine the seizure forecasting intervention.

International Registered Report Identifier (IRRID): PRR1-10.2196/85993

(JMIR Res Protoc 2026;15:e85993) doi:10.2196/85993

KEYWORDS

epilepsy; seizure forecasting; data science; artificial intelligence; machine learning; data collection technology

Introduction

Background

Epilepsy is a chronic neurological disorder, which is characterized by recurrent and unpredictable seizures [1]. Although there have been significant advancements in treatment, up to one-third of people with epilepsy still experience recurrent seizures [2]. Epilepsy has a significant impact on the well-being and quality of life of people with epilepsy, in which fear and the unpredictability of seizures are considered greatly debilitating characteristics of the condition [3,4], with the unpredictability of seizures being identified as having the greatest impact on the lives of people with epilepsy [5]. To address the problem of unpredictability, technology to enable the prediction of seizures has been identified as a research priority by the UK Epilepsy Priority Setting Partnership [6]. To tackle this issue, there is a growing body of research investigating the potential for technology to provide seizure forecasts [7]. This technology holds promise for helping people with epilepsy take proactive steps to reduce the risk of injury and lessen the psychological burden caused by the unpredictability of seizures [8].

Current technologies for epilepsy have primarily focused on detection (identifying a seizure that is already happening) rather than predicting or forecasting future seizures. Electroencephalography (EEG) remains the leading method for monitoring brain signals to detect seizure activity. However, it is not well-suited for seizure forecasting, as EEG is impractical for long-term use outside clinical settings [9]. Ultra–long-term EEG monitoring is an alternative, which allows continuous EEG data capture outside clinical settings and in individuals' real-world context. However, ultra–long-term EEG monitoring requires surgical implantation, and the associated risks and costs may limit its accessibility for many users [10]. Less invasive alternatives such as scalp EEG are available; however, their adoption is also limited due to low acceptance among users [4].

With the increasing prevalence of wearable technologies such as smartwatches [11], there is significant potential for their application in epilepsy management and seizure forecasting. These devices, which are already embedded in many individuals' daily routines, offer a convenient, noninvasive means of continuously capturing real-time physiological and behavioral data [12]. The Empatica E4 wristband has shown to be acceptable to users for epilepsy management and exhibits promise for seizure detection [9,13]. Advancements in machine learning and artificial intelligence have enabled the use of physiological data collected from wrist-worn devices to inform artificial intelligence algorithms for seizure forecasting, which have shown promising results [14,15]. However, despite the value provided by prior work, there remain limitations such as a narrow scope focusing only on circadian cycles and phases [16] or challenges with accuracy and generalizability [17].

The ATMOSPHERE (Artificial Intelligence to Optimise Seizure Prediction to Empower People With Epilepsy) project [18] aims to develop a seizure forecasting system using accessible data collection technology and machine learning. The data collection technology collects real-time data on evidence-based seizure precipitants (or triggers), for example, sleep quality, stress, and medication adherence [19,20]. These data are collected using sensors on a smartwatch and self-report surveys via a companion app using ecological momentary assessment (EMA) [21]. The system then uses machine learning for predictive analytics, creating an individualized risk score for the likelihood of seizures in the near future, which is presented to users through the companion app. The data collection technology was co-designed with clinicians and people with epilepsy (21 consultations). Their input informed the iterative design of the data collection technology used to collect the real-time data underpinning the seizure forecasting system. The resultant prototype was then usability tested and refined through a mixed methods, in-the-wild [15], longitudinal deployment study (n=14) [18]. In parallel, the seizure forecasting technology based on machine learning algorithms is in development. The development of the seizure forecasting system has been informed by state-of-the-art work in human-computer interaction and data science. The ultimate aim of the technology is to improve clinical outcomes for people with epilepsy. To determine this, the complex digital health intervention will need to be evaluated as part of a fully powered clinical trial investigating the clinical and cost-effectiveness of the technology.

A complex health intervention is defined by the UK Medical Research Council (MRC) as an intervention that has several interacting components or is dependent on the behaviors of those delivering and receiving the intervention, has a range of possible outcomes, or requires tailoring to different contexts

XSL•FO
RenderX

and settings [22]. The MRC and the National Institute for Health and Care Research have developed a framework to provide guidance on complex health intervention development, recommending an approach of moving through 4 distinct phases: from development or identification of the intervention, to feasibility, then evaluation, and finally to the implementation phase [23]. The feasibility phase is a critical component, used to test trial methods on a small scale to assess whether a future, fully powered clinical trial can be undertaken [24]. Recruitment and retention of participants are typically the major challenges in the delivery of clinical trials [25,26], and mixed methods data during feasibility studies can be used to identify issues and optimize procedures to maximize these elements. In relation to recruitment, clinical trials need to recruit both a sufficient sample size and also sufficient diversity, for example, in relation to socioeconomic status, ethnicity, and age since a lack of diversity can exacerbate health inequalities [27]. Therefore, it is important to identify and reduce any modifiable barriers to participation, with particular attention to inclusive procedures for those groups typically underrepresented in health research [28]. With respect to retention and completion of outcome measures, it is crucial to avoid issues that lead to underpowered studies, as these can introduce bias in the interpretation of treatment effects. This depends on the extent of, and underlying reasons for, missing data, and how it is distributed across trial arms [29]. Therefore, it is essential to monitor attrition and completeness of the outcome measures (referring to the availability of data for specific measures or items) across arms of the trial. This allows for the assessment of retention and completion rates, as well as the identification of reasons for any missing data, which can inform potential strategies to address barriers.

In addition to optimizing trial procedures, the MRC-National Institute for Health and Care Research complex intervention framework recognizes that the feasibility phase can be used to optimize the intervention. Mixed methods data, such as quantitative use data and qualitative data on user experience, can inform intervention refinement to enhance usability and the clinical utility of the system. Further to usability refinement, the Clinical Trials Transformation Initiative recommends that feasibility studies should be used to assess, and if needed, correct technical performance, assessing elements such as issue-related data quality, unanticipated challenges, and weaknesses of the selected system [30].

Aims and Objectives

Building on formative complex intervention development work, this study focuses on the second phase defined as feasibility testing. Initially, this single-arm, single-site feasibility study aims to test and refine the trial methods that will be used within a future fully powered clinical trial of clinical effectiveness. Second, it aims to test and refine the data collection technology. This is the component of the intervention designed to capture the real-time data underpinning the ATMOSPHERE seizure forecasting system.

The first objective is to assess and optimize trial processes. First, recruitment rates will be assessed (outcome 1) to understand the feasibility of recruiting people with epilepsy into the study. Specifically, we will record the number of people with epilepsy attending the clinic, screened by a clinician, eligible, providing consent to contact, and providing full informed consent. Second, we will assess the diversity of the recruited sample (outcome 2) by reporting the clinical and demographic characteristics of those consenting to participate. Third, the barriers and facilitators to recruitment will be explored with a focus on inclusive trial design for those underrepresented in health research (outcome 3). The number of participants leaving the study at each point of the recruitment phase as well as demographic characteristics and reason for withdrawal (where known) will be represented in the study flow diagram. If needed, Patient and Public Involvement (PPI) sessions with people with epilepsy and clinicians will be conducted to identify modifiable barriers and inform future recruitment strategies. Fourth, retention and outcome measure completion rates will be evaluated (outcome 4) by monitoring attrition and withdrawal rates throughout the study and reporting competition rates for each study procedure. Finally, barriers and facilitators to retention and outcome measure completion will be examined, with a focus on inclusive trial design (outcome 5). The number of participants who discontinue participation at each study procedure, together with their demographic characteristics and reasons for withdrawal (where known), will be summarized in the study flow diagram. Additional PPI sessions with people with epilepsy and clinicians will be undertaken to address modifiable barriers and inform future retention strategies.

The second objective is to assess and optimize the data collection technology of the seizure forecast system. Usability (outcome 6) will be assessed using both quantitative and qualitative methods. Quantitative technology use data, including onboarding rates and engagement metrics, will be analyzed to understand how participants interact with the system. In addition, qualitative interviews will be conducted to explore user experiences, perceptions of the technology, and suggestions for improvement. The technical performance (outcome 7) of the minimum viable product (MVP) and safety of the technology system (outcome 8) will be evaluated through a built-in function that allows users to report technical performance and safety issues within the app in real time. In addition, users' perceptions of the app's technical performance and safety issues encountered with the technology will be explored through qualitative interviews.

The third objective is to add the longitudinal data on seizures and seizure precipitants (data collected from the data collection technology) to the existing ATMOSPHERE dataset and share with the team developing the forecasting technology to refine and improve algorithmic performance.

Methods

Study Design

This study is a single-arm mixed methods feasibility study to assess and refine trial design (objective 1) and the data collection technology component (objective 2) to inform the future clinical evaluation of the seizure forecasting system. Recruitment is expected to begin in quarter 1 of 2026. The feasibility study will initially test a prototype of the data collection technology

XSL•FO
RenderX

used to collect the data (phase 1), transitioning to testing an MVP of the data collection technology (phase 2). The data collected will be added to the existing dataset and used to refine the seizure forecasting algorithm (objective 3). This study protocol is reported in accordance with the CONSORT (Consolidated Standards of Reporting Trials) guidelines (see checklists in Multimedia Appendices 1 and 2).

Participants and Recruitment

Overview

Potential participants will be identified during routine clinical appointments by clinicians from specialist National Health Service epilepsy clinics operated by the Royal Wolverhampton NHS Trust (RWT). RWT provides neurology services to almost 1 million people. Its core services are based in a large, multicultural inner city with significant levels of socioeconomic deprivation. However, RWT is also contracted to provide neurology services to people living in Shropshire and parts of Powys in Wales. These are largely rural localities. The diverse populations will enable assessment of the intervention across a range of geographical areas, as well as ethnic and socioeconomic groups. Clinicians will screen each patient at the clinic against the inclusion or exclusion criteria, maintaining a screening log. For those who are eligible, the clinician will provide a brief introduction about the study, including a participant information sheet, and obtain oral consent (1) to share their contact details with the research team and (2) for a member of the research team to contact them to discuss their participation in the study (consent to contact). Oral consent to contact will be documented in clinical records and screening log.

Eligible participants will be contacted by a member of the research team who will share further information about the study and invite participants to ask any questions they may have. For those choosing to take part, they will be asked to complete an online consent form via Microsoft Forms.

Inclusion or Exclusion Criteria

Adults (aged 18 years and older) will be eligible for inclusion if they have a diagnosis of epilepsy; are likely to have a seizure at least twice per month (seizure data required for modeling); are fluent in English (to be able to engage with the technology and qualitative interview); can provide informed consent; have access to an email address; can wear a wrist wearable device (smartwatch), supplied as part of the study; can input data into a smartphone app; own a smartphone that can download apps from either the Apple Store or the Google Play Store (for phase 1); and own an iPhone that can download apps from the Apple Store (for phase 2).

Adults will be excluded if they have a major cardiac condition, have a diagnosis of a major psychiatric condition (defined as needing secondary care mental health services support), and have a diagnosis of functional dissociative seizures or reflex epilepsy.

The aim is to screen all participants attending the clinics, with all participants who are eligible being given initial information and invited to consent to be contacted. All participants who consent to contact will be offered a meeting with a researcher and an opportunity to consent to the study. This is to reduce bias in the study. Where possible, reasons for nonscreening, nonconsent to contact, and nonconsent will be recorded, so that recruitment processes can be monitored.

Sample Size

Several potential positive outcomes are anticipated from the use of the seizure forecasting system. It is hypothesized that these include improved clinical outcomes, better seizure control, reduced seizure-related injuries, enhanced locus of control, improved quality of life, and better mental health. Additionally, a reduction in health care resource use, such as fewer acute injury-related admissions, is also expected.

We will engage with stakeholders (people with epilepsy and clinicians) to identify which of these outcomes should be designated as the primary outcome for evaluation. This decision will be made once the complete seizure forecasting system has been finalized and is ready for clinical evaluation. Following this, the sample size required for the full-scale evaluation will be calculated based on the minimal clinically important difference for the selected primary outcome, ensuring that the trial is adequately powered to detect a clinically meaningful effect.

For this feasibility study, the sample size was determined by the 3 objectives:

- Objective 1 includes exploring the rate of recruitment and the diversity of the sample recruited. This will be used to inform a fully powered trial, for example, guiding the selection of additional sites and the length of the trial. This feasibility study will be used to determine the time frame needed to recruit 60 participants and the diversity of this sample. A sample size of 60 was selected to meet the requirements of objectives 2 and 3 (see below) and aligned with the typical sample size and recommendations for feasibility studies [31,32].
- Objective 2 focuses on refining the data collection technology of the seizure forecasting system. To do this, purposeful sampling will be used to identify 9-12 participants to invite them to take part in qualitative interviews to explore their experiences of using the MVP of the technology and identify any usability issues. Sampling for diversity will take place, with a focus on including individuals with lower digital literacy and from lower socioeconomic status groups. This is to ensure that the data collection technology is designed to include the needs of these groups to reduce the risks of digital exclusion and exacerbations of health inequalities. A sample size of 9-12 is considered sufficient to identify most usability issues in qualitative usability testing [33], ensuring a rich understanding of user experience and potential areas for improvement.
- Objective 3 seeks to add the data collected to the existing dataset to support the optimization of the seizure forecasting system. As the forecasting will be personalized and the system will be tailored to individual participants, the focus will be on obtaining accurate and continuous longitudinal data from each participant.

XSL•FO
RenderX

The Data Collection Technology

Overview

The technology for collecting real-time data underpinning the seizure forecasting system consists of a smartwatch and a companion smartphone app referred to as the data collection technology. Previous iterations of the smartphone app prototype have undergone co-design and usability testing [18]. In this study, participants will be given either version 1 (the prototype) or version 2 (the MVP) of the technology, depending on their point of entry into the study. The primary difference between the prototype and MVP is that the prototype uses a third-party app (Labfront) integrated with the smartwatch, whereas the MVP will be a novel smartphone app developed by the research team. Both versions will collect the same data points, as outlined in Table 1.

XSL•FO
RenderX

Table 1. Data collection protocol.

Data point	Labfront data collection schedule	MVP[a] data collection schedule
Passive data collection via Garmin device		
Sleep score[b]	Daily summary	Daily summary
Sleep stage[b]	Continuous	Continuous
Sleep respiration	Continuous, 1-minute sampling rate	Continuous
Stress[b]	Continuous, 10-second sampling rate	Continuous
Intensity minutes	Continuous, 15-minute sampling	Continuous
Motion intensity	Continuous, 15-minute sampling	Continuous
Body battery[b]	Continuous, 3-minute sampling rate	Continuous
Active calories burned	Continuous, 15-minute sampling	Continuous
Step count	Continuous, 1-minute sampling rate	Continuous
Actigraphy	Continuous, 1-minute sampling rate	Continuous
Heart rate	Continuous, 10-second sampling rate	Continuous
Heart rate interval	Continuous, each beat	Continuous
Respiration rate	Continuous, 3-minute sampling rate	Continuous
Pulse oximetry blood oxygen saturation	Continuous, 3-minute sampling rate	Continuous
Accelerometer	Continuous, 15-minute sampling rate	Continuous
User-initiated[c] data via smartphone app (EMA[d])		
Seizure event (type, timing, duration, severity)	Twice daily notification[e]	A twice daily notification and ad hoc user-initiated
Prodromal symptoms	Twice daily (notification)	Twice daily notification and ad hoc user-initiated
Postictal symptoms	Up to twice daily (notification; optional if seizure event reported)	As required (if seizure event reported)
Epilepsy medication use: routine medication	Once daily (notification)	Customized notification and ad hoc user-initiated
Epilepsy medication use: emergency medication	Once daily (notification)	As required, ad hoc user-initiated
Emotional states and emotional intensity	Twice daily (notification)	Twice daily (notification) and ad hoc user-initiated
Sleep duration and quality	Once daily (notification)	Once daily (notification)
Acute illness or infection	Twice daily (notification)	Twice daily notification and ad hoc user-initiated
Menstrual cycle	One off series of questions (optional)	One off series of questions (optional), which can be updated
Alcohol	N/A[f]	Ad hoc user-initiated
Personalized trigger	N/A	Ad hoc user-initiated

[a]MVP: minimum viable product.

[b]Compound measures through the Garmin application programming interface.

[c]User-initiated: users can complete the EMA item any time within the app.

[d]EMA: ecological momentary assessment.

[e]Notification: smartphone notification to prompt the user to complete the EMA item.

[f]N/A: not available.

Phase 1: Prototype

In phase 1, data collection will take place through a customized Labfront smartphone app installed on participants' smartphones and connected via Bluetooth to a Garmin VivoActive 5 worn by participants (as outlined in Figure 1). Quantitative data captured passively through the Garmin wearable as well as compound measures, such as stress level and sleep score (Table

XSL•FO
RenderX

1), will be securely transmitted to the encrypted Labfront cloud by Garmin's Health application programming interface. The use of Garmin's application programming interface enables access to the raw data from the smartwatch. All participant data are anonymized before entering the Labfront system.

The EMA surveys that participants will receive in the Labfront app have been developed for participants to record data on seizures (timing, duration, type, and severity), seizure precipitants (triggers), prodromal symptoms, and postictal symptoms. EMA survey completion will be prompted twice daily by notification to the participant at a set morning and evening time. An optional menstrual cycle set-up survey may be completed at the start of participation. Completed surveys will be uploaded to the secure Labfront cloud, where they can be accessed by permissioned members of the core research team.

Figure 1. An overview of the phase 1 interactive prototype. API: application programming interface.

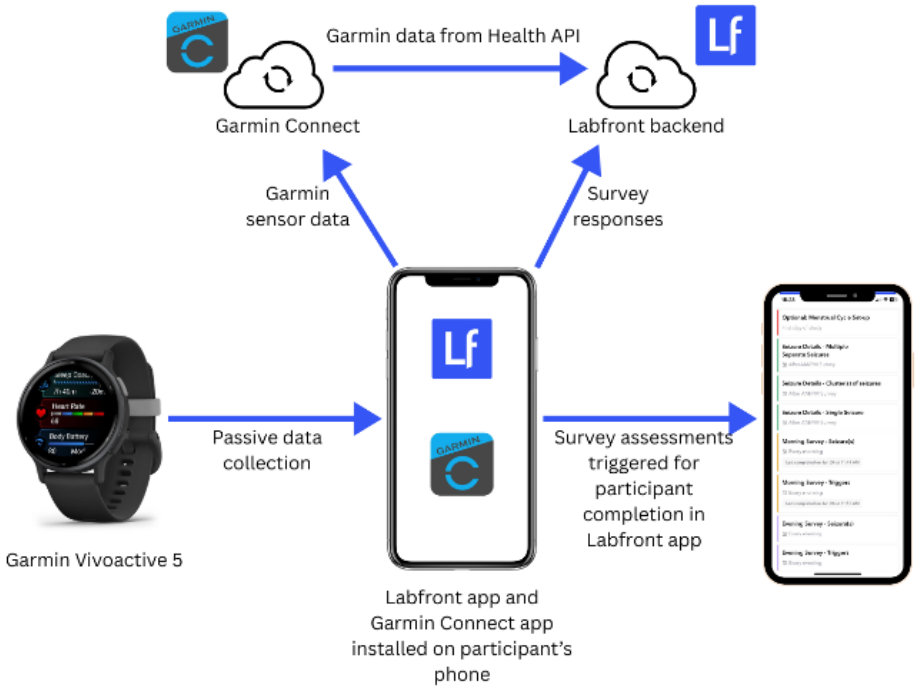

Phase 2: MVP

The design and development of the MVP were informed by earlier co-design and usability studies [18]. The anticipated development plan includes the creation of a custom-built smartphone app for phase 2, which will be developed using Swift, Apple's native programming language for iOS. An agile software development methodology will be used, encompassing several iterative development and feedback cycles with stakeholders. Operating under a licensing agreement, the app will use the Garmin Health Standard Software Development Kit, which enables health and fitness data, such as heart rate and sleep, to be transferred from Garmin wearable devices to the app. Data being collected via the MVP are defined in Table 1. Quantitative data from the connected Garmin Vivoactive 5/6 device will be passively sent to the MVP app installed on an Apple iPhone. This will be relayed to an encrypted cloud database hosted in a secure digital environment in Amazon Web Services (AWS), independent of Garmin's own services or databases. User self-reported data relating to EMA data defined in Table 1 will be similarly relayed to the same encrypted cloud database. Authentication to upload and download data from this database will be managed through AWS Cognito, which adopts the industry standard OAuth 2.0 protocol.

Algorithm Development

As part of this project, we are developing a seizure forecasting algorithm that uses physiological time series data collected from a smartwatch. The algorithm calculates time- and frequency-domain feature statistics, which are combined with seizure timings, obtained from the companion app to train and test a combination of machine learning models to extract seizure propensity scores in different time scales. The data collected via the data collection technology will be added to the existing ATMOSPHERE dataset and used to refine and advance the seizure forecasting algorithm.

Study Procedures

Study procedures are illustrated in the template study flow diagram in Figure 2.

XSL•FO
RenderX

Figure 2. Study flow diagram showing participant flow through the study. Recruitment phases in pink and yellow, retention phases in blue and green, and a qualitative interview for a subgroup of participants in purple. CROM: clinician-reported outcome measure; PROM: patient-reported outcome measure.

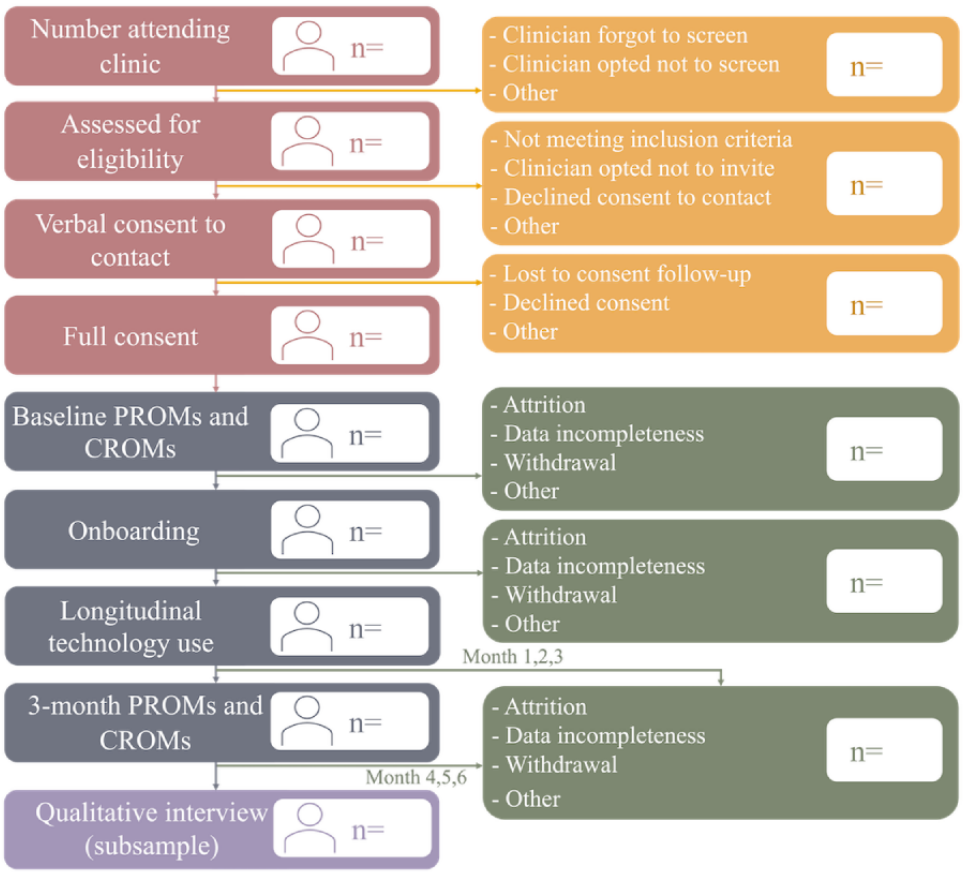

Procedure 1: Deployment of Prototypes

A pack containing a smartwatch and instructions for using the smartwatch and companion app will be sent to participants. An onboarding session will be held online with each participant to support the setup of the prototype. Participants will be invited to use the technology for up to 6 months. Each participant's data will be reviewed monthly, and continuation of technology use for the following month will be contingent upon the participant wearing the smartwatch for some time and responding to the app questions on at least 21 days during the current month.

During the prototype deployment, incoming data will be monitored by the research team through dashboards generated within the Labfront software. Each week, participants will be sent an email summarizing their engagement with the prototype, including device wear time and survey completion. If missing data are noted, such as a survey completion rate falling below 80%, an in-app notification will be sent to participants reminding them to respond. If issues persist, follow-up via email will be conducted by the research team to offer troubleshooting support.

During the MVP deployment phase, incoming data will be monitored, and, if issues with use are identified, the MVP will be refined accordingly.

Procedure 2: Patient-Reported Outcome Measures and Clinician-Reported Outcome Measures

The following demographic and clinical information will be collected about participants at baseline by the clinical team based at the recruiting site. This will be taken from the clinical notes where possible, with additional data collected from the participant if it is not available in the notes: ethnicity, age, highest level of education completed, total annual household income, seizure type, and epilepsy medications and treatments.

Currently, there are no internationally accepted core outcomes for epilepsy, although one is under development [34]. The International Consortium for Health Outcomes Measurement (ICHOM) formed a global working group, including clinicians, people with epilepsy, and advocates, to define a standardized minimum set of outcomes for clinical practice [35]. The clinician-reported outcome measures (CROMs) and patient-reported outcome measures (PROMs) in this study capture these essential outcomes, including seizure- and non–seizure-related outcomes.

Participants will be asked to complete a series of PROMs at baseline and at 3 months. These include Patient-Reported Outcomes Measurement Information System (PROMIS) Cognitive Function Short Form [36], Patient Health Questionnaire (PHQ)-9 [37], Generalized Anxiety Disorder (GAD)-2 [38], PROMIS Sleep Disturbance—Short Form [39],

XSL•FO
RenderX

and Quality Of Life in Epilepsy (QOLIE)-10 [40]. Participants will also be asked to complete a digital literacy scale, the Mobile Device Proficiency Questionnaire-16 [41], at baseline only.

A member of the clinical team will also complete a series of CROMs at baseline and at 3 months. These include: mortality; seizure frequency, seizure freedom, emergence of new seizure type, and seizure severity; emergency and unplanned health care services use; seizure control during pregnancy, pregnancy, and delivery complications; and memory.

Participants will have up to 1 month to complete the measures. These can be done either online or in paper format. An email reminder will be sent within the first 2 weeks after the expected completion date. If the measures remain incomplete after 2 weeks, a member of the research team will follow up by phone to assist with completion. Participants may be contacted by phone up to 3 times.

Procedure 3: Qualitative Interview and Usability Questionnaire (Phase 2 Only)

At the end of the technology deployment, participants who used version 2 of the prototype (MVP) will be invited to participate in a qualitative interview with a member of the research team. These interviews will be scheduled for up to 60 minutes and use a semistructured topic guide (Multimedia Appendix 3) designed to explore participants' views and experiences of using the technology.

Participants will also be asked to complete the mHealth App Usability Questionnaire (MAUQ), an 18-item questionnaire, which consists of 3 subscales that focus on ease of use (5 items), interface and satisfaction (7 items), and usefulness (6 items), rated on a 7-point Likert scale (1=strongly disagree to 7=strongly agree) [42].

Outcome Measures and Planned Analyses

Objective 1: Feasibility Outcomes Measures and Analysis: Trial Procedures

Overview

Quantitative and qualitative data will be used to assess and optimize the trial procedures and intervention component. Using the screening logs, consent outcomes, and the number of participants completing study procedures (onboarding, use of the technology longitudinally, PROMs at baseline and 3 months, and qualitative interview), we will create a study flow diagram to present descriptive statistics on the participant flow through the study, as per Figure 2.

Outcome 1: Recruitment Rates

The number of participants at each recruitment phase will be represented in the study flow diagram, as illustrated in Figure 2. The feasibility of recruitment will be assessed through the following indicators: the number of patients (1) attending the clinic, (2) screened by a clinician, (3) eligible to participate, (4) providing consent to contact, and (5) providing full consent. To determine the feasibility to progress to a larger study and full trial, we expect to reach a minimum of 70% of the recruitment target.

Outcome 2: Diversity of Recruited Sample

The clinical and demographic characteristics of those consenting to the study will be reported to evaluate the diversity of the recruited sample. Characteristics will include seizure type, age, gender, ethnicity, socioeconomic status, and digital literacy. This will help assess whether the recruitment approach was effective in enrolling a diverse sample and be used to inform a purposive recruitment strategy for a full trial if necessary.

Outcome 3: Barriers and Facilitators to Recruitment, With a Focus on Inclusive Trial Design for Those Underrepresented in Health Research

Quantitative Descriptive Statistics

The number of participants leaving the study at each point of the recruitment phase, along with their demographic characteristics and reason for withdrawal (were known), will be represented in the study flow diagram (Figure 2).

Stakeholder Consultations

If issues are identified during the recruitment phase, we will engage with people with epilepsy and the key members of the clinical and research teams to identify and document any barriers or problems related to recruitment. In response, practical, context-specific solutions aimed at improving recruitment processes will be developed and implemented collaboratively with the clinical team. This iterative approach will ensure that emerging challenges are addressed promptly and that recruitment strategies remain acceptable and feasible within the clinical setting. A full trial is unlikely to be feasible if we do not reach 70% of our recruitment target, and the qualitative data suggest that recruitment cannot be improved any further.

Outcome 4: Retention

To assess participant retention, the number of participants remaining in the study at each study procedure will be reported in the study flow diagram (Figure 2). Specifically, the following will be reported: the number of participants who completed (1) onboarding, (2) the technology deployment phase (either phase 1 or 2 version) in their lived context for up to 6 months, (3) PROMs at baseline and 3 months, and (4) a qualitative interview exploring their views of the technology (conducted with a subsample of 9-12 participants). To determine the feasibility to progress to a larger study and full trial, we expect to reach 80% retention at follow-up.

Outcome 5: Barriers and Facilitators to Retention, With a Focus on Inclusive Trial Design for Those Underrepresented in Health Research

The number and demographics of participants who leave the study, along with reasons for leaving (where known), will also be reported. If issues are identified during the retention phase, we will conduct PPI sessions with people with epilepsy to identify possible solutions to any barriers. A full trial is unlikely to be feasible if retention is <80%, and the qualitative data suggest that retention cannot be improved any further.

XSL•FO
RenderX

Objective 2: Feasibility Outcomes Measures and Analysis: Refining the Intervention

Outcome 6: Usability

Quantitative Use Metrics

Usability will be assessed through descriptive statistics of wear time (assessed using the presence of heart rate data), proportion of participants who complete the scheduled questionnaires at each time point. These quantitative indicators will offer insight into the feasibility and acceptability of both the data collection technology and associated data collection procedures.

To determine feasibility to progress to a larger study and full trial, we will consider the study feasible if at least 80% (≥48/60) of participants achieve an average wear time of ≥19 hours per day across the study period and a total average score of 5 on the MAUQ. This threshold was informed by established usability benchmarks for the System Usability Scale, which has been shown to be strongly correlated with the MAUQ [42].

Quantitative User Experience (Phase 2 Only)

Usability will be assessed quantitatively by calculating the total and average scores of the MAUQ items.

Qualitative User Experience (Phase 2 Only)

Usability will also be assessed through qualitative interviews with participants who use the MVP version of the app. Interview data will be analyzed using thematic analysis to identify key themes related to the acceptability and feasibility of the app in its current form. In addition, tables of change will be used to systematically track suggested improvements and user feedback, highlighting specific areas where the app may require refinement. It is unlikely that a full trial will be feasible if the data suggest that the technology is not acceptable to users.

Outcomes 7 and 8: Technical Assessment and Safety Reporting (Phase 2 Only)

The number of issues reported by participants will be presented. The types of issues reported will be described following a content analysis of the issue reports. The study's safety officer (PT; clinical lead) will review the issues reported and will determine whether the study should be stopped.

Objective 3: Refining the Algorithm Development

The longitudinal data on seizures and seizure precipitants (data collected from the data collection technology) will be added to the existing ATMOSPHERE dataset, shared with the modeling team to support the continued development of the seizure forecasting algorithm.

The modeling team will evaluate a combination of machine learning algorithms for seizure forecasting. Continuous physiological time series will be windowed (with segments ranging between 1 and 36 hours), and features will be calculated for each segment. Features will include statistical measures in the time domain (eg, mean, median, variance, skewness, and kurtosis) and the frequency domain (eg, mean frequency and peak frequency).

Features will be used to inform machine learning models to predict seizures in a period of time after each segment (forecasting window—ranging from 1 to 24 hours). Models will include 1-class support vector machine, isolation forest, and extreme gradient boosting. Model performance will be evaluated using well-established metrics such as accuracy, balanced accuracy, sensitivity, and specificity. We will also use a user-centered approach to evaluate performance, training the algorithm to minimize false positive or negative rates according to user preference. Model accuracy will be validated using a 5-fold cross-validation approach.

Ethical Considerations

This study was reviewed and granted ethics approval by the South East Scotland NHS committee (REC reference: 25/SS/0116). Participants will be asked to provide informed consent by completing an online consent form administered through Microsoft Forms. There is a minor risk that participants will find the smartwatch device uncomfortable to wear. Participants will be made aware of this risk and that they can stop wearing the device at any time via the participant information sheet. If participants report finding the watch uncomfortable, an alternative soft watch strap will be offered as a replacement. There is a risk that participants will experience a seizure when in the presence of the research team, for example, during a qualitative interview. The research team has read an Epilepsy First Aid document written by the charity Epilepsy Action and provided by our Consultant Nurse for the Epilepsies. The research team will also have the contact details of a health care professional with sufficient expertise (such as an epilepsy nurse) with epilepsy in the event that additional support is needed. At the start of any meeting with participants, researchers will make a joint plan at the beginning to identify what the participant would like the researcher to do in the event of a seizure. Participants will be asked to share information about emotional states so there is a possibility that this could cause participants to reflect on negative states. Participants will be made fully aware of this and provided information about this procedure, including that they will be able to stop engaging in this procedure at any point. Participants are being asked to contribute time and effort to the study—the time and effort of wearing the smartwatch, using the smartphone app, engaging in qualitative interviews, and completing outcome measures. Participants will be made aware of this burden in the information sheet, and they will be free to withdraw from study procedures at any time. These will be stored on a separate server to research data so that they cannot be linked. Participants will be allocated a unique research ID, and all research data will be labeled with the research ID and no identifiable data. A Microsoft Excel file linking identifiable information (participants' names) with their research ID will be maintained to enable pseudonymization. The Excel file will be stored securely within the University of Bristol environment, with restricted access to the research team and stored separately from research data. For qualitative research data, audio recordings of qualitative interviews will be deleted once transcription is complete. Transcripts will be pseudonymized with all personal or identifiable information redacted. Baseline characteristics, outcome measure forms (PROMs and CROMs), and MAUQ will be captured via Microsoft Forms. PROMs and CROMs will be entered into Microsoft Forms directly by participants and by the clinical

XSL•FO
RenderX

team and will be accessed by members of the research team. The data collection system has been designed to be secure. Labfront is a data collection tool designed for academic researchers. It is a secure solution for collecting high-resolution physiological data from wearables and survey responses. Storage of data will be in the United States, meeting the appropriate security and privacy regulations and certifications. To ensure data privacy and protection, a Data Protection Impact Assessment was conducted, which aims to evaluate and minimize risks to the rights and freedoms of individuals' personal data (aligned with General Data Protection Regulation) [43], which was approved by the information compliance manager and data protection officer at the university. The pseudonymized data will be shared with the modeling collaborators via the secure Research Data Storage Facility to ensure controlled and protected access. The information from the app will be hosted in the cloud using AWS. The development of this will follow industry standard methods of data handling, storage, and transfer, including data encryption on device and in transit using HTTPS, in accordance with the UK General Data Protection Regulation. At this stage, the Data Protection Impact Assessment will be renewed to reflect the respective changes to data flows and storage. Participants will be given approximately £50 (approximately US $68) in shopping vouchers per month that they are involved in the study, including a further £25 (approximately US $34) voucher for participating in a qualitative interview.

Results

Overview

Recruitment is planned to begin in quarter 1 of 2026, with data collection expected to be completed by quarter 2 of 2027. Data analysis will take place during quarter 3, and the results will be published in quarter 4.

Dissemination of Findings

Feasibility outcomes (objectives 1 and 2) will be analyzed and written up for publication in quarter 3 of 2027. Objective 3 is to add the longitudinal data on seizures and seizure precipitants (data collected from the data collection technology, the smartwatch, and smartphone app) to the existing ATMOSPHERE dataset to optimize the machine learning seizure forecasting algorithm. The machine learning analysis will be reported separately. The outcomes of the feasibility study will be reported according to the CONSORT extension to pilot and feasibility trials [44,45].

The outcomes of this study will be disseminated through open-access peer-reviewed publications via research journals and conferences. Research findings will also be disseminated to people with epilepsy and public networks via mechanisms such as workshops, plain English summaries, and blog posts.

Discussion

Principal Findings

This feasibility study aims to lead to optimized methods for a full-scale clinical evaluation (objective 1) and a refined version of the mobile data collection component MVP (objective 2). In parallel to this study, the data science research and development will refine the seizure forecasting algorithm, which will be optimized through data from this feasibility study (objective 3). As such, this feasibility study is a critical step in finalizing the seizure forecasting technology and delivering a fully powered trial to determine clinical effectiveness.

Strengths and Limitations

The study was designed to test trial procedures, including recruitment, retention, and outcome measures (objective 1). A key strength is the focus on optimizing the trial process to include individuals who are typically underrepresented in health research, such as those from socioeconomically disadvantaged backgrounds and ethnic minority groups [46]. At this initial single site, the representation of these groups will be monitored, and findings will directly inform the selection of additional sites for the full-scale clinical trial in order to improve diversity and reach. A limitation associated with objective 1 is that the methodology used in this smaller-scale study is likely to differ in several respects from the anticipated design of the full-scale trial, meaning that certain elements of the full trial will not have been feasibility-tested. Specifically, this study is being conducted at a single site using a single-arm design, whereas the full trial is expected to adopt a randomized controlled trial design. Additionally, outcome measures are being collected at baseline and at 3 months, while the full-scale trial is anticipated to involve a longer follow-up period (eg, up to 12 months). As a result, multisite recruitment, randomization procedures, and long-term outcome measurement will remain untested; however, it is typical for feasibility studies to use a simplified design. Furthermore, outcome measures in this study have been derived from the ICHOM clinical outcomes framework, as no research core outcome sets for epilepsy currently exist [35]. The Epilepsy Outcome Set for Effectiveness Trials (EPSET) project [34] is developing an epilepsy outcome set for effectiveness trials, which may become available in time for the fully powered trial. If so, the outcome measures will be amended to align with the EPSET Core Outcome Sets. This would result in a discrepancy between the outcome measures assessed during the feasibility study and those used in the larger trial. However, it is likely that ICHOM clinical outcomes and EPSET will be largely aligned; therefore, it is expected that the measures used in this study will likely still produce relevant feasibility insights. Finally, as the primary outcome will be determined upon finalization of the intervention, a power calculation for the full trial has not yet been conducted. However, this feasibility study will provide estimates of recruitment rates, which will inform realistic projections of timelines and the number of sites required to achieve the desired sample size once it is defined.

The study was also designed to refine the intervention (objective 2). A strength of this phase is that sampling for the qualitative usability testing has been structured to prioritize inclusivity. This is particularly important, given that underrepresented groups are at increased risk of digital exclusion and health inequalities [47], as well as experiencing a much higher burden of epilepsy [48]. Usability and technical performance are being assessed through both qualitative and quantitative data,

XSL•FO
RenderX

providing insights into objective use across the full sample, alongside in-depth data on user perspectives and experiences.

Finally, this study is evaluating 1 component of the seizure forecasting technology, the data collection component. Valuable insights will be generated regarding various aspects of the technology, including the onboarding process, the design and functionality of the frontend interface, and the technical performance of the system. However, the final version of the technology will incorporate algorithm integration and model outputs within the backend. As such, further rounds of user testing, technical performance evaluation, and safety assessment will be required prior to the full-scale trial. Nonetheless, testing at this early stage is considered a strength, as it enables a more agile development process. Early learning about data collection challenges can occur in parallel with the ongoing development of the data science components and human-computer interaction elements.

Conclusions

Seizure forecasting technology holds significant promise for improving clinical and cost-effectiveness outcomes for individuals living with epilepsy. This mixed methods feasibility study is designed to assess the feasibility of the proposed trial procedures and to refine the forecasting technology in a real-world setting. Insights gained from this study will be critical for optimizing the design and implementation of a future full-scale clinical trial, ensuring that it is both methodologically robust and patient-centered.

Acknowledgments

The authors are grateful to Labfront and Garmin for their continued support. The authors also thank the staff at Royal Wolverhampton NHS Trust for their essential time and effort in supporting the setup of this study.

Funding

This work has been supported by the LEAP Digital Health Hub, which has been funded by the Engineering and Physical Sciences Research Council (grant EP/X031349/1).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request..

Authors' Contributions

LT drafted the initial manuscript and contributed on study design and National Health Service ethics applications. EN, EEVQ, and AS supported the draft of the initial paper and led the formative usability study. M Wac, M Wragg, and SD led on software development and contributed to the software development sections of the manuscript. PT contributed to study design and ethics applications and is the clinical advisor for ATMOSPHERE (Artificial Intelligence to Optimise Seizure Prediction to Empower People With Epilepsy). LJ contributed to the study design and is the data science lead for ATMOSPHERE. LS, RC, PT, and YQ are all part of the ATMOSPHERE project team of the data science and/or human-computer interaction research. AB conceptualized the study and contributed on study design and had input on the iterative version of the manuscript. All authors reviewed and contributed to the paper.

Conflicts of Interest

RC is a data scientist at Neuronostics. Neuronostics develops technologies for the diagnosis, treatment, and monitoring of neurological conditions, including epilepsy. PT is part of the Epilepsy Outcome Set for Effectiveness Trials steering group.

Multimedia Appendix 1
CONSORT checklist.
[PDF File (Adobe PDF File), 144 KB - resprot_v15i1e85993_app1.pdf]

Multimedia Appendix 2
CONSORT abstract checklist.
[PDF File (Adobe PDF File), 164 KB - resprot_v15i1e85993_app2.pdf]

Multimedia Appendix 3
Interview topic guide.
[DOCX File , 22 KB - resprot_v15i1e85993_app3.docx]

References

XSL•FO
RenderX

1. Epilepsies in children, young people and adults. NICE Guideline. 2022. URL: https://www.nice.org.uk/guidance/ng217 [accessed 2025-01-30]
2. Chen Z, Brodie MJ, Liew D, Kwan P. Treatment outcomes in patients with newly diagnosed epilepsy treated with established and new antiepileptic drugs: a 30-year longitudinal cohort study. JAMA Neurol 2018 Mar 01;75(3):279-286 [FREE Full text] [doi: 10.1001/jamaneurol.2017.3949] [Medline: 29279892]
3. Arthurs S, Zaveri HP, Frei MG, Osorio I. Patient and caregiver perspectives on seizure prediction. Epilepsy Behav 2010 Nov;19(3):474-477. [doi: 10.1016/j.yebeh.2010.08.010] [Medline: 20851054]
4. Stirling RE, Cook MJ, Grayden DB, Karoly PJ. Seizure forecasting and cyclic control of seizures. Epilepsia 2021 Feb;62(Suppl 1):S2-S14. [doi: 10.1111/epi.16541] [Medline: 32712968]
5. Community survey. Epilepsy Foundation Ei2. 2016. URL: https://www.epilepsy.com/sites/default/files/atoms/files/community-survey-report-2016%20V2.pdf [accessed 2026-01-31]
6. Epilepsy PSP top 10. James Lind Alliance. URL: https://www.jla.nihr.ac.uk/priority-setting-partnerships/epilepsy#tab-27091 [accessed 2026-01-31]
7. Carmo AS, Abreu M, Baptista MF, de Oliveira Carvalho M, Peralta AR, Fred A, et al. Automated algorithms for seizure forecast: a systematic review and meta-analysis. J Neurol 2024 Oct;271(10):6573-6587. [doi: 10.1007/s00415-024-12655-z] [Medline: 39240346]
8. Ghosh S, Sinha JK, Ghosh S, Sharma H, Bhaskar R, Narayanan KB. A comprehensive review of emerging trends and innovative therapies in epilepsy management. Brain Sci 2023 Sep 11;13(9):1305 [FREE Full text] [doi: 10.3390/brainsci13091305] [Medline: 37759906]
9. Regalia G, Onorati F, Lai M, Caborni C, Picard RW. Multimodal wrist-worn devices for seizure detection and advancing research: focus on the Empatica wristbands. Epilepsy Res 2019 Jul;153:79-82. [doi: 10.1016/j.eplepsyres.2019.02.007] [Medline: 30846346]
10. Viana PF, Pal Attia T, Nasseri M, Duun-Henriksen J, Biondi A, Winston JS, et al. Seizure forecasting using minimally invasive, ultra-long-term subcutaneous electroencephalography: individualized intrapatient models. Epilepsia 2023 Dec;64(Suppl 4):S124-S133. [doi: 10.1111/epi.17252] [Medline: 35395101]
11. Masoumian Hosseini M, Masoumian Hosseini ST, Qayumi K, Hosseinzadeh S, Sajadi Tabar SS. Smartwatches in healthcare medicine: assistance and monitoring; a scoping review. BMC Med Inform Decis Mak 2023 Nov 03;23(1):248 [FREE Full text] [doi: 10.1186/s12911-023-02350-w] [Medline: 37924029]
12. Kamišalić A, Fister I, Turkanović M, Karakatič S. Sensors and functionalities of non-invasive wrist-wearable devices: a review. Sensors (Basel) 2018 May 25;18(6):1714 [FREE Full text] [doi: 10.3390/s18061714] [Medline: 29799504]
13. Nasseri M, Nurse E, Glasstetter M, Böttcher S, Gregg NM, Laks Nandakumar A, et al. Signal quality and patient experience with wearable devices for epilepsy management. Epilepsia 2020 Nov;61(Suppl 1):S25-S35. [doi: 10.1111/epi.16527] [Medline: 32497269]
14. Nasseri M, Pal Attia T, Joseph B, Gregg NM, Nurse ES, Viana PF, et al. Ambulatory seizure forecasting with a wrist-worn device using long-short term memory deep learning. Sci Rep 2021 Nov 09;11(1):21935 [FREE Full text] [doi: 10.1038/s41598-021-01449-2] [Medline: 34754043]
15. Meisel C, El Atrache R, Jackson M, Schubach S, Ufongene C, Loddenkemper T. Machine learning from wristband sensor data for wearable, noninvasive seizure forecasting. Epilepsia 2020 Dec;61(12):2653-2666. [doi: 10.1111/epi.16719] [Medline: 33040327]
16. Karoly PJ, Cook MJ, Maturana M, Nurse ES, Payne D, Brinkmann BH, et al. Forecasting cycles of seizure likelihood. Epilepsia 2020 Apr;61(4):776-786. [doi: 10.1111/epi.16485] [Medline: 32219856]
17. Cheng JC, Goldenholz DM. Seizure prediction and forecasting: a scoping review. Curr Opin Neurol 2025 Apr 01;38(2):135-139. [doi: 10.1097/WCO.0000000000001344] [Medline: 39831588]
18. Quilter EEV, Downes S, Deighan MT, Stuart L, Charles R, Tittensor P, et al. A digital intervention for capturing the real-time health data needed for epilepsy seizure forecasting: protocol for a formative co-design and usability study (the ATMOSPHERE study). JMIR Res Protoc 2024 Sep 19;13:e60129 [FREE Full text] [doi: 10.2196/60129] [Medline: 39298757]
19. Samsonsen C, Mestvedthagen G, Uglem M, Brodtkorb E. Disentangling the cascade of seizure precipitants: a prospective observational study. Epilepsy Behav 2023 Aug;145:109339. [doi: 10.1016/j.yebeh.2023.109339] [Medline: 37413785]
20. Haut SR, Hall CB, Masur J, Lipton RB. Seizure occurrence: precipitants and prediction. Neurology 2007 Nov 13;69(20):1905-1910. [doi: 10.1212/01.wnl.0000278112.48285.84] [Medline: 17998482]
21. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
22. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
23. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]

XSL•FO
RenderX

24. Bell ML, Whitehead AL, Julious SA. Guidance for using pilot studies to inform the design of intervention trials with continuous outcomes. Clin Epidemiol 2018;10:153-157 [FREE Full text] [doi: 10.2147/CLEP.S146397] [Medline: 29403314]
25. Tudur Smith C, Hickey H, Clarke M, Blazeby J, Williamson P. The trials methodological research agenda: results from a priority setting exercise. Trials 2014 Jan 23;15(1):32 [FREE Full text] [doi: 10.1186/1745-6215-15-32] [Medline: 24456928]
26. Bower P, Brueton V, Gamble C, Treweek S, Smith CT, Young B, et al. Interventions to improve recruitment and retention in clinical trials: a survey and workshop to assess current practice and future priorities. Trials 2014 Oct 16;15:399 [FREE Full text] [doi: 10.1186/1745-6215-15-399] [Medline: 25322807]
27. Pomeranz H, Horvath T. Promoting diversity to eliminate health disparities. J Phys Assist Educ 2017 Oct;28(Suppl 1):S43-S48. [doi: 10.1097/JPA.0000000000000155] [Medline: 28961622]
28. Improving inclusion of under-served groups in clinical research: guidance from INCLUDE project. National Institute for Health and Care Research. 2020. URL: https://www.nihr.ac.uk/support-and-services/support-for-delivering-research/improving-inclusion-under-served-groups-clinical-research-guidance-include-project [accessed 2026-01-31]
29. Akl EA, Briel M, You JJ, Sun X, Johnston BC, Busse JW, et al. Potential impact on estimated treatment effects of information lost to follow-up in randomised controlled trials (LOST-IT): systematic review. BMJ 2012 May 18;344:e2809. [doi: 10.1136/bmj.e2809] [Medline: 22611167]
30. Bakker JP, Goldsack JC, Clarke M, Coravos A, Geoghegan C, Godfrey A, et al. A systematic review of feasibility studies promoting the use of mobile technologies in clinical research. NPJ Digit Med 2019;2:47 [FREE Full text] [doi: 10.1038/s41746-019-0125-x] [Medline: 31304393]
31. Teare MD, Dimairo M, Shephard N, Hayman A, Whitehead A, Walters SJ. Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. Trials 2014 Jul 03;15:264 [FREE Full text] [doi: 10.1186/1745-6215-15-264] [Medline: 24993581]
32. Totton N, Lin J, Julious S, Chowdhury M, Brand A. A review of sample sizes for UK pilot and feasibility studies on the ISRCTN registry from 2013 to 2020. Pilot Feasibility Stud 2023 Nov 21;9(1):188 [FREE Full text] [doi: 10.1186/s40814-023-01416-w] [Medline: 37990337]
33. Nielsen J, Landauer TK. A mathematical model of the finding of usability problems. 1993 Presented at: CHI '93: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems; April 24-29, 1993; Amsterdam, The Netherlands p. 206-213. [doi: 10.1145/169059.169166]
34. Mitchell JW, Noble A, Baker G, Batchelor R, Brigo F, Christensen J, et al. Protocol for the development of an international Core Outcome Set for treatment trials in adults with epilepsy: the EPilepsy outcome Set for Effectiveness Trials Project (EPSET). Trials 2022 Nov 17;23(1):943 [FREE Full text] [doi: 10.1186/s13063-022-06729-4] [Medline: 36397081]
35. Mitchell JW, Sossi F, Miller I, Jaber PB, Das-Gupta Z, Fialho LS, et al. Development of an international standard set of outcomes and measurement methods for routine practice for adults with epilepsy: the International Consortium for Health Outcomes Measurement consensus recommendations. Epilepsia 2024 Jul;65(7):1916-1937. [doi: 10.1111/epi.17971] [Medline: 38738754]
36. Iverson G, Marsh JM, Connors EJ, Terry DP. Normative reference values, reliability, and item-level symptom endorsement for the PROMIS® v2.0 Cognitive Function-Short Forms 4a, 6a and 8a. Arch Clin Neuropsychol 2021 Oct 13;36(7):1341-1349. [doi: 10.1093/arclin/acaa128] [Medline: 33454756]
37. Kroenke K, Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. Psychiatr Ann 2002 Sep;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
38. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 06;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
39. Cella D, Choi SW, Condon DM, Schalet B, Hays RD, Rothrock NE, et al. PROMIS adult health profiles: efficient short-form measures of seven health domains. Value Health 2019 May;22(5):537-544 [FREE Full text] [doi: 10.1016/j.jval.2019.02.004] [Medline: 31104731]
40. Cramer JA, Perrine K, Devinsky O, Meador K. A brief questionnaire to screen for quality of life in epilepsy: the QOLIE-10. Epilepsia 1996 Jun;37(6):577-582. [doi: 10.1111/j.1528-1157.1996.tb00612.x] [Medline: 8641236]
41. Roque NA, Boot WR. A new tool for assessing mobile device proficiency in older adults: the Mobile Device Proficiency Questionnaire. J Appl Gerontol 2018 Feb;37(2):131-156 [FREE Full text] [doi: 10.1177/0733464816642582] [Medline: 27255686]
42. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth App Usability Questionnaire (MAUQ): development and validation study. JMIR Mhealth Uhealth 2019 Apr 11;7(4):e11500 [FREE Full text] [doi: 10.2196/11500] [Medline: 30973342]
43. Regulation P. General data protection regulation. Intouch 2018;25:1-5. [doi: 10.1002/9781119436775.ch18]
44. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ 2016 Oct 24;355:i5239 [FREE Full text] [doi: 10.1136/bmj.i5239] [Medline: 27777223]
45. Lancaster GA, Thabane L. Guidelines for reporting non-randomised pilot and feasibility studies. Pilot Feasibility Stud 2019;5:114 [FREE Full text] [doi: 10.1186/s40814-019-0499-1] [Medline: 31608150]

XSL•FO
RenderX

46. Erves JC, Mayo-Gamble TL, Malin-Fair A, Boyer A, Joosten Y, Vaughn YC, et al. Needs, priorities, and recommendations for engaging underrepresented populations in clinical research: a community perspective. J Community Health 2017 Jun;42(3):472-480 [FREE Full text] [doi: 10.1007/s10900-016-0279-2] [Medline: 27812847]
47. Gann B. Transforming lives: combating digital health inequality. IFLA J 2019 May 07;45(3):187-198. [doi: 10.1177/0340035219845013]
48. Bush K, Cullen E, Mills S, Chin RFM, Thomas RH, Kingston A, et al. Assessing the extent and determinants of socioeconomic inequalities in epilepsy in the UK: a systematic review and meta-analysis of evidence. Lancet Public Health 2024 Aug;9(8):e614-e628 [FREE Full text] [doi: 10.1016/S2468-2667(24)00132-4] [Medline: 39095136]

Abbreviations

ATMOSPHERE: Artificial Intelligence to Optimise Seizure Prediction to Empower People With Epilepsy
AWS: Amazon Web Services
CONSORT: Consolidated Standards of Reporting Trials
CROM: clinician-reported outcome measure
EEG: electroencephalography
EMA: ecological momentary assessment
EPSET: Epilepsy Outcome Set for Effectiveness Trials
ICHOM: International Consortium for Health Outcomes Measurement
MAUQ: mHealth App Usability Questionnaire
MRC: Medical Research Council
MVP: minimum viable product
PPI: Patient and Public Involvement
PROM: patient-reported outcome measure
RWT: Royal Wolverhampton NHS Trust

Edited by A Schwartz; submitted 16.Oct.2025; peer-reviewed by P Bhatnagar; comments to author 14.Nov.2025; revised version received 26.Jan.2026; accepted 27.Jan.2026; published 20.Mar.2026.

Please cite as:
Thompson L, Nielsen E, Quilter EEV, Slay A, Wac M, Wragg M, Downes S, Tittensor P, Junges L, Stuart L, Charles R, Kissack P, Qureshi Y, Brigden A
A Digital Intervention for Capturing Real-Time Health Data for Epilepsy Seizure Forecasting: Protocol for the ATMOSPHERE Study
JMIR Res Protoc 2026;15:e85993
URL: https://www.researchprotocols.org/2026/1/e85993
doi:10.2196/85993
PMID:

©Lauren Thompson, Emily Nielsen, Emily E V Quilter, Amelia Slay, Marceli Wac, Matthew Wragg, Samuel Downes, Phil Tittensor, Leandro Junges, Liz Stuart, Rosie Charles, Peter Kissack, Yasser Qureshi, Amberly Brigden. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

HealthData@MAD-R&I: Protocol for Design and Development of a Regional Health Data Infrastructure to Enable Secondary Use of Health Data in Research and Innovation

Montserrat León-García[1], PharmD, PhD; Sergio Álvarez-Pérez[1], PhD; Janire Gesto-Gómez[1], MSc; Clara Urbano-Molina[1], PhD; Sonia Soto-Díaz[2], MD; Juan Cárdenas-Valladolid[2,3,4], RN, PhD; Luis Rodríguez-Rodríguez[5], MD, PhD; Antonio Díaz-Holgado[6], MSN; Isabel del Cura-González[6,7,8,9,10], MD, PhD; Javier De La Cruz-Bertolo[11], MD; Noelia García-Barrio[11], MSc; Juan Luis Cruz-Bermúdez[11], PhD; Cristina García-Fernández[12], BEng, MSc; Carlos Rodríguez-Antolín[13], PhD; Laila García-Aldars[2], PhD; Elsa María Moreda-Sánchez[2], PhD; Álvaro Roldán López[2], MSc; José María Veganzones Alonso-Cortés[12], MD, MSE; Ana Isabel Gonzalez Gonzalez[2,8,9], MD, PhD; Miguel A Salinero-Fort[2,4], MD, PhD; HealthData@MAD-R&I Working Group[14]

[1]Fundación para la Investigación e Innovación Biosanitaria de Atención Primaria (FIIBAP), Madrid, Madrid, Spain

[2]Dirección General de Investigación y Docencia, Consejería de Sanidad de la Comunidad de Madrid, Madrid, Madrid, Spain

[3]Universidad Alfonso X el Sabio, Madrid, Madrid, Spain

[4]Frailty, Multimorbidity Patterns and Mortality in the Elderly Population Residing in the Community, Hospital La Paz Institute for Health Research (IdiPAZ), Madrid, Madrid, Spain

[5]Grupo de Patología Musculoesquelética, Fundación para la Investigación Biomédica del Hospital Clínico San Carlos, Instituto de Investigación Sanitaria San Carlos (IdISSC), Madrid, Madrid, Spain

[6]Gerencia Asistencial de Atención Primaria, Servicio Madrileño de Salud, Consejería de Sanidad de la Comunidad de Madrid, Madrid, Madrid, Spain

[7]Departamento de Especialidades Médicas y Salud Pública, Universidad Rey Juan Carlos, Madrid, Madrid, Spain

[8]Network for Research on Chronic Diseases, Primary Care, and Health Promotion (RICAPPS), Madrid, Madrid, Spain

[9]Instituto de Investigación Sanitaria Gregorio Marañón, Madrid, Madrid, Spain

[10]Ageing research center, Karolinska Institute, Solna, Stockholm, Sweden

[11]Grupo de Investigación e Innovación en Transformación Digital e Ingeniería Biomédica, Hospital Universitario 12 de Octubre, Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Madrid, Spain

[12]Dirección General de Salud Digital, Consejería de Digitalización de la Comunidad de Madrid, Madrid, Madrid, Spain

[13]Instituto de Genética Médica y Molecular, Instituto de Investigación Sanitaria del Hospital Universitario La Paz, Madrid, Madrid, Spain

[14]See Acknowledgments

Corresponding Author:
Ana Isabel Gonzalez Gonzalez, MD, PhD
Dirección General de Investigación y Docencia
Consejería de Sanidad de la Comunidad de Madrid
Calle Aduana, 29, 3rd Fl
Madrid, Madrid, 28013
Spain
Phone: 34 91 5290742
Email: healthdata-ri@salud.madrid.org

Abstract

Background: The exponential growth of electronic health records (EHRs), together with the recent entry into force of the European Health Data Space (EHDS) Regulation, highlights the urgent need for secure, interoperable environments that support the secondary use of health data. In response, HealthData@MAD-R&I emerges as a pioneering initiative in Madrid (Spain), aligned with the EHDS strategy and the European Commission's vision for data sovereignty and trustworthy data reuse.

Objective: This study aims to design and implement HealthData@MAD-R&I, a regional health data space that enables responsible access to high-quality health data to support clinical research, health care innovation, and evidence-informed decision-making.

XSL•FO
RenderX

Methods: HealthData@MAD-R&I aims to establish an ethically governed, scalable, and sustainable health data space. The project adopts a structured, iterative methodology based on the Data Management Association (DAMA) framework. and it is organized into 9 work packages across three thematic areas: (1) project management and sustainability, (2) governance and technological infrastructure, and (3) validation through 4 real-world use cases. The technical architecture adopts a hybrid federated model built with open-source components, and data harmonization is performed using the Observational Medical Outcomes Partnership (OMOP) common data model to ensure semantic and syntactic interoperability. Artificial intelligence, machine learning, natural language processing, and privacy-preserving techniques are applied for data curation and secure access.

Results: As of November 2025, the main achievements include (1) the development of a data governance model that articulates principles of quality, transparency, and regulatory compliance; (2) the design of a secure, interoperable technological architecture with federated capabilities based on international standards (DAMA and OMOP); and (3) the implementation of 4 use cases—optimizing rheumatology referrals, characterizing care pathways for long-term survivors of breast cancer, predicting unplanned hospitalizations, and evaluating the effectiveness of statins in older adults—to validate the data space while addressing diverse clinical and policy challenges. Together, these components demonstrate the potential of regional data spaces to support evidence-based clinical practice and public policy.

Conclusions: HealthData@MAD-R&I seeks to strengthen Madrid's role in digital health innovation and contribute to the broader European health data ecosystem by promoting interoperable, privacy-compliant secondary use of health data. The project's evaluation framework includes indicators for data quality, research outputs, and health care system impact.

International Registered Report Identifier (IRRID): DERR1-10.2196/82815

(JMIR Res Protoc 2026;15:e82815) doi:10.2196/82815

KEYWORDS

secondary use; interoperability; digital health; advanced predictive models; European Health Data Space; federated data; clinical research; innovation; public health

Introduction

The exponential growth of electronic health records (EHRs) and other digital health data sources has created unprecedented opportunities to improve health care delivery, foster scientific discovery, and support evidence-informed policymaking [1-7]. The COVID-19 pandemic further underscored the need for timely, high-quality health data to support public health responses and accelerate research [5]. In this context, the European Commission launched the European Health Data Space (EHDS) strategy in 2022, aiming to establish a unified health data market that enhances Europe's competitiveness, promotes data sovereignty, and supports the secure and ethical reuse of health data across borders and sectors [3-10]. The publication of the EHDS regulation in the *Official Journal of the European Union* on March 5, 2025, marked a key milestone in the transition from strategy to implementation [11].

Health data spaces are central to this vision, serving as the foundational infrastructure for secure, interoperable, and trustworthy access to health data across member states. They have the potential to enhance population health by enabling large-scale research, improving disease prevention and treatment, advancing precision medicine, and supporting more efficient health system management [4,6,7]. A key component of this paradigm is the secondary use of health data—information originally collected during routine clinical care and subsequently repurposed for research, innovation, or policymaking. Despite its promise, secondary use remains limited by persistent challenges related to interoperability, governance, data quality, privacy protection, and ethical oversight [12-16].

HealthData@MAD-R&I emerges in response to these challenges as a regional initiative in Madrid, Spain, designed to develop a federated, secure, and ethically governed health data space for the secondary use of health data. The project is funded by the Ministry for Digital Transformation and Civil Service and by the European Union through the European Recovery Instrument ("Next Generation EU"), under Spain's Recovery, Transformation, and Resilience Plan. It builds upon existing digital health infrastructure in the region, including the Infobanco project [3] and the Hipócrates secondary-use platform, which integrates large-scale clinical data from multiple providers. The initiative brings together health care institutions, research centers, technology companies, and public authorities within a collaborative ecosystem aimed at enhancing the accessibility, usability, and trustworthiness of health data [3,4,6,12-16].

The objective of this project is to design and implement HealthData@MAD-R&I, a regional health data space that enables the secure and responsible secondary use of health data to support research, innovation, and evidence-informed policymaking, in alignment with European and national data governance frameworks. This manuscript presents the implementation protocol for the development of this infrastructure, with a focus on governance, interoperability, and validation through real-world use cases as part of a systems-level digital health innovation initiative.

Methods

Overview

The implementation of HealthData@MAD-R&I follows a structured, phased methodology articulated through 9 interconnected work packages (WP1-WP9), each addressing a

XSL•FO
RenderX

critical dimension of the project to ensure effective management, governance, security, communication, and long-term scalability (Table 1 and Figure 1). This comprehensive approach integrates data governance, technological infrastructure, and scientific validation while ensuring full compliance with ethical principles and both national [17] and European regulatory frameworks [11].

Figure 1 shows how the 9 WPs of HealthData@MAD-R&I interact with each other, highlighting dependencies, information flows, and oversight. It clarifies roles, reduces perceived overlap, and shows how coordination, governance, technology, communication, and sustainability are connected throughout the project life cycle.

The system architecture is based on a hybrid federated data-space model built using open-source components, which enables flexibility, scalability, and interoperability across institutions. The methodological approach is aligned with the Data Management Association (DAMA) framework [18] and incorporates best practices in data governance, standardization, and advanced analytics, including artificial intelligence (AI), machine learning (ML), and natural language processing (NLP). These elements collectively support the development of a secure, standardized, and analytically robust environment for secondary use of health data.

To streamline execution and ensure coherence across diverse activities, the 9 WPs are organized into three thematic areas: (1) core operational pillars, comprising project management, communication, professional training, and long-term scalability and sustainability; (2) governance and technical architecture, addressing regulatory alignment, data management rules, and the design of the secure processing infrastructure; and (3) validation through real-world use cases, which assess the functionality, clinical relevance, and public-health impact of the data space.

Table 1. Overview of the structure of the HealthData@MAD-R&I project.

Thematic areas	Work packages (WPs)
A1. Core aspects of the project	• WP1. Project management • WP8. Communication and professional training. • WP9. Long-term scalability and sustainability and scalability of the data space
A2. Design and architecture of the data space	• WP2. Development of a robust data governance framework • WP3. Technical architecture of the data space and data curation processes
A3. Validation of the data space through real-world use cases	• WP4. Use case number 1: optimization of referral pathways for patients with rheumatic and musculoskeletal diseases • WP5. Use case number 2: longitudinal analysis of care pathways for women who are long-term survivors of breast cancer • WP6. Use case number 3: development of a predictive model for unplanned hospitalizations • WP7. Use case number 4: evaluation of the effectiveness of statins in the primary prevention of cardiovascular events among individuals aged 75 years and older

Figure 1. Work packages interaction flow. WP: work package.

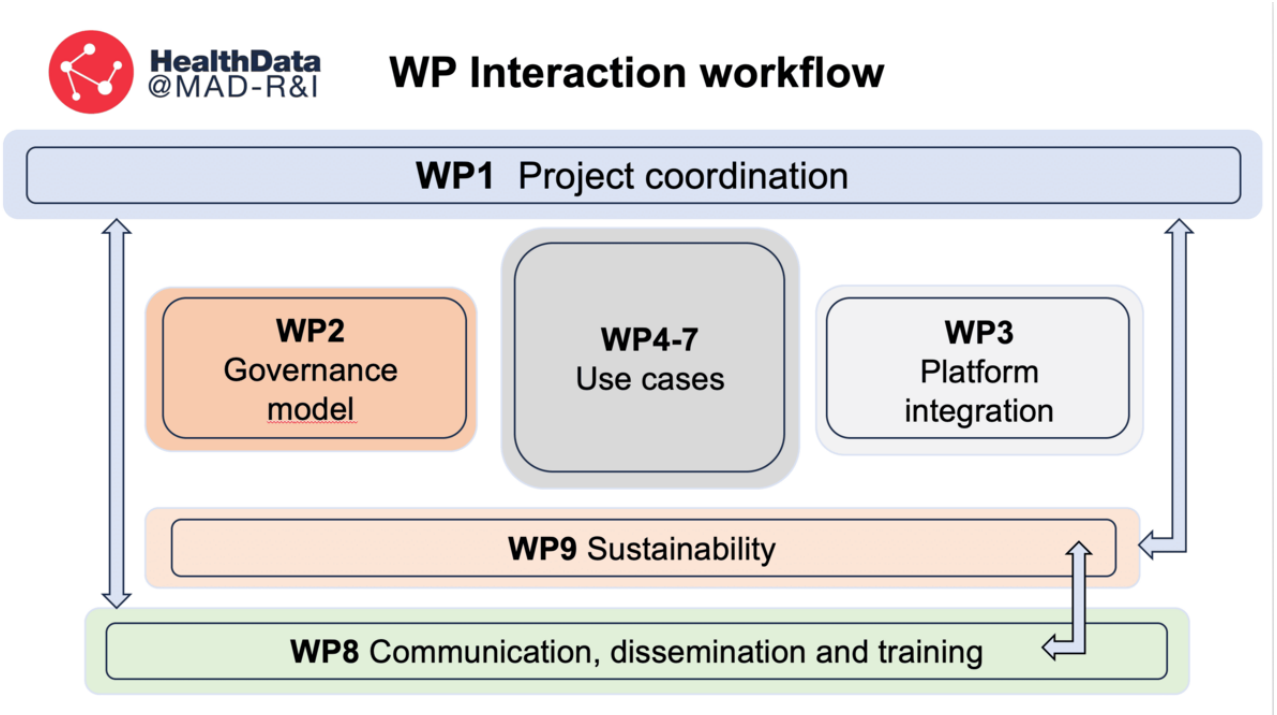

XSL•FO
RenderX

Area 1: Core Aspects of the Project: Management, Communication, Professional Training, and Scalability and Sustainability (WPs 1, 8, and 9)

This first thematic area addresses the foundational elements necessary for the successful deployment and long-term viability of the HealthData@MAD-R&I health data space. It encompasses project management, communication strategies, professional training, and sustainability planning. These components are operationalized through 3 highly intertwined WPs (WP1, WP8, and WP9).

WP1 focuses on project management, establishing the organizational structure, coordination mechanisms, and oversight procedures required to guide the implementation process. The management approach is grounded in the Agile methodology [19], which supports iterative planning, continuous evaluation, and stakeholder participation. This allows for flexible and efficient workflows, fostering collaboration among technical, scientific, and administrative teams. Key activities include the identification and mitigation of risks, the systematic monitoring of milestones and deliverables, and the facilitation of cross-functional coordination to ensure alignment across all components of the project.

A central objective of WP1 is to build and maintain public trust. To ensure meaningful citizen involvement, WP1 includes a structured public and stakeholder engagement plan consisting of (1) the establishment of a patient and citizen advisory board with representation from patient associations and community organizations; (2) periodic public consultations to gather feedback on governance, ethical frameworks, and data-access policies; and (3) open communication channels through the project website and the RIMASalud (institutional repository of the health department) portal for submitting questions, suggestions, or concerns. These mechanisms promote transparency, support co-creation of governance decisions, and ensure that societal expectations and values are reflected in the design and operation of the health data space.

WP8 addresses the project's communication strategy and capacity-building activities. It sets out strategies for communicating project progress, raising stakeholder awareness, and improving digital health literacy, particularly regarding the purpose and value of secondary health data use. Communication activities target multiple audiences (the general public, health care professionals, researchers, and policymakers) and include scientific dissemination through publications and conference presentations, engagement through digital media, and outreach initiatives tailored to patients and citizens.

In parallel, the professional training component delivers workshops, seminars, and user guides designed for researchers, health care professionals, data scientists, and other relevant actors to foster capacity-building in the effective use of the data space. An evaluation framework will assess the impact of these efforts by measuring visibility, stakeholder engagement, user satisfaction, and knowledge acquisition.

WP9 ensures the long-term scalability and sustainability of HealthData@MAD-R&I beyond the initial implementation phase. It defines a roadmap for the progressive incorporation of new real-world use cases and for the continued evolution of the infrastructure. WP9 also guarantees strategic alignment with the EHDS, enabling future interoperability and supporting cross-border data exchange.

To guarantee sustainability, the project adopts a mixed economic model that combines regional public funding with cost-sharing agreements among participating institutions. In addition, a controlled data-access service fee will apply to external research requests—particularly from industry or nonaffiliated organizations—under transparent and equitable governance rules. WP9 also explores public-private partnerships under open-innovation frameworks, allowing private entities to contribute technical capabilities or infrastructure support while preserving public oversight and data safeguards.

A reinvestment mechanism will channel revenues from data-access services into continuous improvement activities, including infrastructure maintenance, cybersecurity reinforcement, quality-assurance processes, and ongoing capacity-building for professionals. Long-term operation of the platform will be supported through framework agreements with the General Directorate for Digital Health (DGSD) [17] and integration with future national and European digital health programs.

All documentation, technical specifications, and progress reports are published through the institutional repository of the health department (RIMASalud) and will continue to be updated as the infrastructure matures.

Area 2: Design and Architecture of the HealthData@MAD-R&I Space: Governance and Technology (WPs 2 and 3)

The second thematic area focuses on the governance and technological foundations required to build a secure, interoperable, and ethically sound data space for the secondary use of health data. It encompasses the design and implementation of a governance model (WP2) and the technical architecture and data curation processes that underpin the infrastructure (WP3). Together, these WPs ensure that HealthData@MAD-R&I meets the highest standards of regulatory compliance, data quality, and operational functionality.

WP2 is focused on developing a robust data governance framework aligned with both national and European regulatory requirements [8,11,17]. The governance framework establishes the lawful basis for data processing as the performance of a task carried out in the public interest for health research and innovation, pursuant to Articles 6(1)(e) and 9(2)(j) of the General Data Protection Regulation (GDPR) [17]. A Data Protection Impact Assessment (DPIA) will be conducted during implementation, under the oversight of the Data Protection Officer of the Madrid Health Service (Regional Health System of Madrid (Servicio Madrileño de Salud [SERMAS]).

Processing activities will follow a controlled workflow with traceable audit logging, and data made available within the secure processing environment (SPE) will undergo privacy-preserving transformations—including

XSL•FO
RenderX

pseudonymization, date shifting, suppression of quasi-identifiers, and aggregation of rare events—in accordance with ISO (International Organization for Standardization)/IEC (International Electrotechnical Commission) 20889:2018 [20,21]. Differential-privacy techniques will be applied to aggregated outputs.

Information security and confidentiality procedures are aligned with the National Security Framework (ENS) and ISO/IEC 27001 [20], ensuring compliance with Spanish and European cybersecurity requirements.

In alignment with GDPR [17] and the citizen rights provisions of the EHDS [8,11], the infrastructure will include a fully automated consent and opt-out management service integrated with institutional EHR systems. Patient choices will be securely stored as standardized Health Level Seven Fast Healthcare Interoperability Resources (HL7 FHIR) [22] and propagated across all participating institutions through FHIR-based interoperability services. Each processing request will query the consent registry in real time, ensuring that opt-out or withdrawal decisions are enforced consistently across the system. This mechanism guarantees traceability, preserves patient autonomy, and provides a practical and interoperable solution for system-wide compliance with data-participant rights.

The governance model also defines the roles and responsibilities of health data holders, health data access bodies, and data users across the lifecycle, following the principles of the DAMA framework [18], to promote responsible stewardship and institutional accountability.

WP3 addresses the technical architecture and data-curation processes that underpin the functionality, scalability, and interoperability of the data space (Figure 2). The architecture is built using open-source components and deployed through the Cloudera platform (Cloudera, Inc), under the supervision of the DGSD.

Figure 2, illustrates the three main layers: (1) the data layer, which integrates clinical and administrative datasets (eg, CMBD [Conjunto Mínimo Básico de Datos], SELENE, and FARMADRID) harmonized to the OMOP (Observational Medical Outcomes Partnership) common data model; (2) the platform layer, including the Federated Coordination Node, ETL and deidentification modules, metadata registry, and SPEs; and (3) the governance layer, encompassing access management, audit logging, and compliance oversight by the Directorate General for Research and Teaching and the Directorate General for Digital Health. Data flow proceeds from local institutional sources to standardized repositories through privacy-preserving pipelines, ensuring traceability, interoperability, and lawful reuse for research and innovation.

To ensure legal and technical interoperability across institutions with varying levels of data maturity, HealthData@MAD-R&I establishes a Federated Coordination Node (FCN) hosted within the regional cloud environment. The FCN synchronizes metadata, validates schemas, orchestrates governance workflows, and supports federated queries. While each participating institution retains control of its source data, the FCN provides shared registries for metadata, terminology mappings, and audit logs. This hybrid federated–centralized model allows heterogeneous systems to interoperate without transferring raw data to a single repository and ensures compliance with the National Interoperability Framework (ENI) and the technical specifications of the EHDS.

The system ingests data from 4 primary domains, including primary care (AP MADRID), hospital care (SELENE and Healthcare Information System [HCIS; DXC Technology]), population-based health registries (including the minimum basic dataset CMBD, CIBELES, and the National Statistics Institute [INE]), and medication-related datasets (eg, FARMADRID). Ingested data are harmonized and transformed into a standard semantic and syntactic format using the OMOP model [21] enabling integrated analyses across heterogeneous sources.

Although storage and computation occur within a secure regional cloud operated by Madrid Digital, governance, stewardship, and access decisions remain decentralized across institutions. This preserves institutional data sovereignty while enabling standardized, privacy-preserving access and analytics through shared coordination services. The OMOP common data model and the DAMA framework support semantic interoperability, metadata governance, and quality assurance, ensuring alignment with both national [17] and EHDS requirements [8,11].

The infrastructure is also designed in accordance with the Spanish ENS. A preliminary security impact assessment has been completed to determine the system's classification (medium–high category), and risk-based controls are being implemented to achieve ENS certification. These include penetration testing, role-based access control, incident-response protocols, and regular third-party security audits coordinated by Madrid Digital's information-security office.

Structured data are automatically processed and encoded using internationally recognized terminologies such as *ICD-10 (International Statistical Classification of Diseases and Related Health Problems 10th Revision)*, SNOMED-CT (Systematized Nomenclature of Medicine Clinical Terms), and LOINC (Logical Observation Identifiers, Names, and Codes). In parallel, unstructured clinical data—such as free-text notes and reports—are processed using advanced NLP and ML techniques to extract meaningful information and enrich data completeness. This dual curation process ensures that both structured and unstructured data can be leveraged for research and decision-making.

The extract-transform-load process is fully documented and version-controlled to ensure traceability of data lineage. Mapping specifications follow the OMOP model conventions, including concept coverage analysis for local codes and terminologies. Deidentification follows a standardized pipeline comprising date shifting, pseudonymization of identifiers, suppression of quasi-identifiers, and aggregation of rare categories. NLP outputs undergo automated review to prevent inadvertent reidentification. These processes comply with ISO/IEC 20889:2018 principles [20] on privacy-enhancing data deidentification.

XSL•FO
RenderX

To ensure high-quality, interoperable data across systems such as CMBD, SELENE, and FARMADRID, HealthData@MAD-R&I implements a multilayered data-quality and harmonization pipeline. This includes automated ingestion validation (format, structure, and referential integrity), rule-based cleaning to correct inconsistencies and handle missingness, and semantic normalization using controlled vocabularies. A dedicated quality registry stores metrics—completeness, consistency, conformity, timeliness, and plausibility—and ML-based anomaly detection identifies outliers or unexpected patterns. This pipeline ensures that only validated, high-quality datasets are made available within the SPE, supporting robust, reproducible analytics across all use cases.

Figure 2. Architecture and governance structure of the HealthData@MAD-R&I infrastructure. OMOP: Observational Medical Outcomes Partnership.

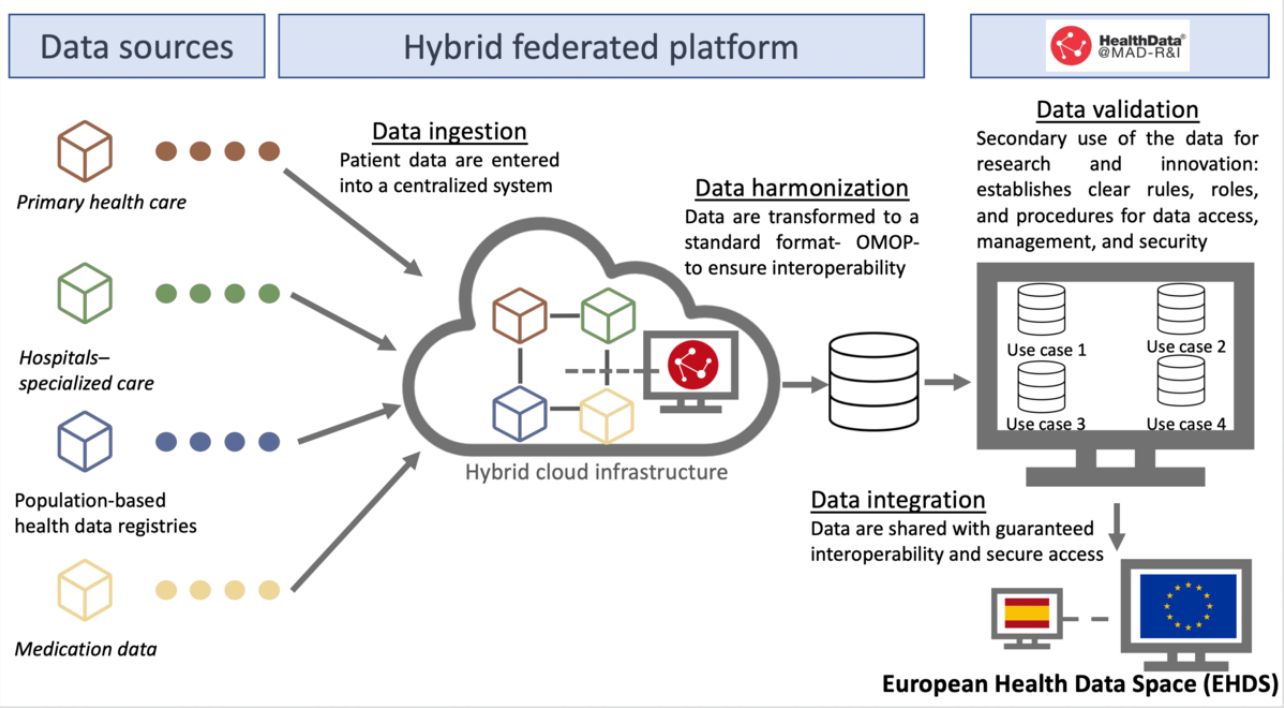

Area 3: Validation of the HealthData@MAD-R&I Space Through Real-World Use Cases (WPs 4-7)

The third thematic area is devoted to validating the functionality, relevance, and impact of the HealthData@MAD-R&I space through the implementation of 4 real-world, data-driven use cases. These use cases serve both as proof-of-concept and as practical demonstrations of how secondary use of health data can support clinical decision-making, improve health care outcomes, and generate evidence to guide public health policy. Each use case is implemented within a dedicated work package (WP4-WP7), allowing for domain-specific focus while ensuring methodological consistency.

Data Analysis

All use cases follow a standardized analytical framework to ensure methodological consistency, reproducibility, and compliance with data protection and governance requirements. Data analyses are conducted within the Secure Processing Environment of the HealthData@MAD-R&I space.

Structured datasets are harmonized using the OMOP common data model and undergo preprocessing procedures, including data cleaning, normalization, and feature engineering. Datasets are partitioned into training (70%), validation (15%), and testing (15%) sets. For NLP-based analyses, unstructured clinical text is semiautomatically annotated using labeling schemes aligned with SNOMED CT entities.

Depending on the specific objectives of each use case, a range of statistical and ML methods are applied, including gradient-boosted decision trees, regularized regression models, and neural network architectures. Model development and tuning are performed using cross-validation techniques. Model performance is evaluated using standard metrics such as area under the receiver operating characteristic curve, F_1-score, precision–recall metrics, and calibration curves.

Model interpretability and transparency are addressed through Shapley Additive Explanations (SHAP)–based feature attribution analyses to support clinical relevance and trustworthiness. All analytical workflows are version-controlled and monitored to ensure auditability, reproducibility, and compliance with GDPR and EHDS requirements.

WP4 aims to optimize referral pathways for patients with rheumatic and musculoskeletal diseases (RMDs). The primary objective is to develop and validate predictive tools to support early diagnosis and appropriate stratification of patients at both the primary and specialist care levels. Given the high prevalence and burden of RMDs, coupled with significant diagnostic delays and misclassifications [23-25], this use case seeks to enhance referral accuracy and reduce the time to effective treatment. By integrating real-world data from EHRs across care levels—including demographic, clinical, and treatment information—this WP will identify referral patterns and clinical markers associated with optimal care pathways. The

XSL•FO
RenderX

implementation of this WP is multicentric, involving collaboration between primary care teams and several university hospitals in Madrid.

WP5 focuses on the longitudinal analysis of care pathways for women who are long-term survivors of breast cancer. Using data from the SURBCAN-Madrid cohort, this use case investigates multimorbidity, health care use, and the evolution of health outcomes among women with long-term survival (>5 years after cancer remission). The objective is to identify predictive factors associated with better outcomes and to inform personalized follow-up protocols. The study integrates data from hospital and primary care settings, as well as registries and pharmacy records, and includes both survivors of breast cancer and matched control groups. The findings will support the design of evidence-based survivorship care strategies.

WP6 is dedicated to developing a predictive model for unplanned hospitalization. Building upon earlier risk stratification efforts developed in the Infobanco project [3], this use case aims to refine and validate predictive tools that identify individuals at high risk of avoidable admissions. By integrating EHRs, medication data, and hospital activity records, this WP seeks to support clinicians with decision-making tools for proactive intervention. The model is expected to improve care coordination, optimize resource allocation, and reduce unnecessary hospitalizations.

WP7 evaluates the effectiveness of statins in the prevention of cardiovascular events among individuals aged 75 years and older with no prior cardiovascular disease. This population is underrepresented in clinical trials, and the effectiveness of statins for primary prevention in this age group remains uncertain [26]. Using linked clinical, pharmaceutical, and mortality data, this use case will assess health outcomes associated with statin use, identify subgroups that benefit most, and evaluate the cost-effectiveness of current prescribing practices. The findings will contribute to more personalized prescribing strategies and inform clinical guidelines for primary prevention of cardiovascular events in older adults.

All 4 use cases follow a common analytical and governance framework, including standardized data harmonization, secure data access, and outcome evaluation using established clinical and system-level indicators. The results will be synthesized into actionable recommendations, contributing to evidence-based clinical practice and informing policy at both regional and national levels.

Evaluation

The evaluation framework for HealthData@MAD-R&I follows a SMART approach—specific, measurable, achievable, relevant, and time-bound—to assess the platform's performance and impact. Indicators are structured around 3 dimensions, including data quality, research output, and health care improvement. Each includes baselines, target values, and timelines for assessment. For example, data completeness and interoperability will be measured quarterly, with a baseline of current system coverage and a target of ≥90% standardized datasets by 2026. Research indicators (eg, number of publications, funded projects) will be benchmarked annually. To assess health care impact, quasi-experimental designs such as "Interrupted Time Series" and "Difference-in-Differences analyses" will be applied to evaluate changes in outcomes attributable to the platform rather than to secular trends.

Evaluation metrics are structured into 3 primary categories, including data quality, research impact, and health care improvement. Ensuring the quality of the data available within the health data space is fundamental to its success. Key indicators in this area include the completeness of datasets—for example, measured by the percentage of missing values in critical variables—alongside the accuracy of data entries validated against reference standards. Timeliness is also evaluated by tracking the speed of data ingestion and processing, while interoperability is assessed through the number of datasets successfully converted to standard models such as OMOP and made available for use. Data usage metrics, including the number of executed queries and frequency of dataset access by researchers and health care providers, further indicate the utility and accessibility of the platform.

To measure research impact, the project will monitor the number of peer-reviewed publications, conference presentations, and technical reports that derive from analyses conducted within the data space. Additional indicators include the volume of funded research projects leveraging the platform, the establishment of new collaborations among health care institutions, academic centers, and industry, and reductions in the time required to generate actionable insights compared to conventional research methods.

Health care improvement is evaluated by analyzing the influence of data-driven insights on clinical practice and health system performance. This includes reductions in unnecessary referrals and increased diagnostic precision for RMDs, enhanced coordination and follow-up in survivorship care for patients with breast cancer, lower rates of unplanned hospitalizations, and better understanding of medication effectiveness—particularly statins—in older populations. Improvements in patient outcomes, such as earlier diagnoses, fewer complications, and increased survival rates, will serve as key indicators of impact.

In addition to performance metrics, a set of broader success indicators has been defined to evaluate the overall sustainability and systemic influence of the initiative. These include the extent to which the data space becomes a valuable resource for biomedical and clinical research, the validation and clinical implementation of predictive models and decision-support tools developed through the platform, and improvements in health care delivery driven by access to high-quality, real-world data. Evidence of enhanced care pathways, greater adoption of personalized medicine approaches, and measurable cost reductions will be crucial outcomes.

Stakeholder engagement and satisfaction will also be tracked. High levels of participation from health care providers, researchers, and patients, as well as the successful inclusion of new stakeholders over time, will indicate a dynamic and sustainable ecosystem. Compliance with data protection regulations, the absence of significant data breaches, and a

XSL•FO
RenderX

demonstrable level of public trust in the platform are equally important for long-term success.

Finally, the project's contribution to the broader data ecosystem will be assessed through indicators such as the stimulation of innovation and entrepreneurship, the emergence of new products and services, the development of revenue streams to support the sustainability of the platform, and Spain's strengthened positioning as a European leader in secondary health data use.

Through continuous monitoring and iterative evaluation based on these indicators, HealthData@MAD-R&I aims not only to ensure internal quality improvement but also to provide a robust foundation for scaling and replicating the model across other regions and sectors.

Ethical Considerations

The protocol was approved by the Research Ethics Committees of Hospital Universitario 12 de Octubre (applications 25/442 and 25/526), Comité de Ética de la Comunidad de Madrid (application 01/25), and Hospital Clínico San Carlos (application 24/259-E). The study complies with national and European ethical and legal standards, including the GDPR (EU [European Union] 2016/679) and the European Health Data Space Regulation (EU 2025/327).

Results

As of November 2025, the HealthData@MAD-R&I project, funded on November 1, 2024, has achieved 3 major milestones: (1) the drafting of a comprehensive data governance model that articulates principles of quality, transparency, and regulatory compliance; (2) the development of a secure, interoperable technological architecture with data federation capabilities, based on international standards such as DAMA and OMOP and incorporating artificial intelligence, ML, and NLP tools; and (3) the design and implementation of 4 real-world use cases—optimization of rheumatology referrals, follow-up of long-term survivors of breast cancer, prediction of unplanned hospitalizations, and evaluation of statin effectiveness in older adults—which will validate the health data space while addressing clinical and policy challenges. Together, these milestones demonstrate the project's progress toward creating a sustainable, evidence-driven regional health data infrastructure.

Discussion

The HealthData@MAD-R&I project constitutes a significant step forward in the development of health data spaces for secondary use in Europe. By establishing a federated, secure, and ethically governed data infrastructure in the Madrid region, this initiative aligns with the strategic objectives of the EHDS [8] and contributes to the creation of a harmonized framework for responsible data sharing. Its primary innovation lies in improving the usability of real-world health data—originally collected for clinical care—for purposes of research, innovation, and health system strengthening, while maintaining compliance with both Spanish and European regulatory frameworks.

A core contribution of HealthData@MAD-R&I is the development of a robust data governance model, which defines the roles and responsibilities of health data holders, access bodies, and users. This governance framework introduces oversight mechanisms to ensure transparency, equity, and security in data access and usage. By applying privacy-preserving techniques and reinforcing ethical standards, the initiative lays the groundwork for building public trust, an essential factor for the success of any secondary use health data initiative [4,10,27].

Equally important is the project's use of advanced data harmonization processes, based on OMOP interoperability standards, which enable the transformation of heterogeneous datasets into a common semantic and syntactic structure [21,28]. This allows for scalable, reproducible analysis across care settings and institutions, addressing one of the most pressing barriers to the secondary use of health data fragmentation. Furthermore, by validating its architecture through 4 real-world use cases—covering referral optimization, survivorship care, hospitalization prediction, and medication effectiveness—the project demonstrates not only technical feasibility but also practical value in improving health care delivery.

Within the EHDS framework, HealthData@MAD-R&I is positioned primarily as a regional health data holder that provides secure and interoperable SPEs for authorized research and innovation activities. It will interoperate with national and EU-level Health Data Access Bodies through standardized metadata exchange and harmonized governance procedures. The infrastructure operates under the authority of the SERMAS), with governance jointly coordinated by the General Directorate of Health Research and teaching (DGID) and the DGSD. This structure ensures alignment with regional and European regulatory requirements while maintaining institutional autonomy and accountability.

The design and governance model of HealthData@MAD-R&I were informed by an analysis of other regional health data infrastructures in Spain, including BARDENA (Aragón), PADRIS and SIDIAP (Catalonia), BIGAN (Basque Country), VALMED (Valencia), and MED-P (Andalusia). While these initiatives have advanced secondary use of health data within their respective regions, HealthData@MAD-R&I introduces a stronger integration of governance, interoperability, and real-world validation through use cases. Its hybrid federated–centralized architecture is specifically designed to maximize scalability, reproducibility, and alignment with the EHDS regulatory framework [8], while preserving regional data sovereignty and institutional control over datasets.

Building on these regional contributions, HealthData@MAD-R&I is also positioned within a broader European context, where health data spaces are increasingly recognized as key enablers of research, innovation, and digital transformation. The EHDS provides the regulatory and technical framework to facilitate data access while safeguarding individual rights [7].

Health data spaces are expected to accelerate scientific discoveries, support evidence-based policymaking, and enhance public health preparedness [7,14]. The COVID-19 pandemic demonstrated the need for structured, interoperable, and cross-border data systems to enable real-time decision-making.

XSL•FO
RenderX

Beyond crisis response, secondary health data can drive personalized medicine, improve population health surveillance, and stimulate economic growth through public–private partnerships and digital innovation [12,27].

Looking ahead, the project anticipates compliance with the wider EU digital and data governance framework, including the Data Act, Data Governance Act, Artificial Intelligence Act, Medical Devices Regulation, and Health Technology Assessment Regulation. Alignment with these instruments will ensure that the infrastructure not only supports lawful and ethical data use but also provides a foundation for trustworthy AI, transparent algorithmic evaluation, and interoperability with future cross-border health data services under the EHDS.

At the same time, these opportunities are accompanied by significant ethical, legal, and technical challenges. Privacy and data protection concerns remain at the forefront. Although regulations such as the European Union's GDPR provide legal safeguards, ambiguities persist around the practical application of these principles within health data spaces [7]. The potential for reidentification of anonymized data, lack of explicit consent mechanisms, and growing citizen concerns about surveillance or commercial exploitation may undermine public trust. In parallel, issues of transparency in data governance—especially regarding access by private entities—raise concerns about fairness, accountability, and societal benefit [29].

Technical interoperability is another persistent barrier, particularly in countries with fragmented health care IT systems. Despite the adoption of international standards like FHIR and OMOP, differences in data quality, structure, and availability continue to hinder seamless integration. Furthermore, the governance and operational roles of health data access bodies require further clarification and consistency across jurisdictions to ensure equitable access and regulatory coherence.

Within this evolving ecosystem, HealthData@MAD-R&I demonstrates several strengths. Its federated architecture allows data to remain under the control of health institutions while enabling secure and efficient secondary access. The project integrates a wide range of clinical and administrative datasets across care levels and leverages AI and ML to unlock clinically relevant insights. Its validation through real-world use cases ensures direct applicability to pressing health challenges and facilitates the development of data-driven interventions that improve patient outcomes. The project's alignment with national digital transformation policies and its compatibility with EHDS standards make it a scalable model for replication.

Nonetheless, the project is not without limitations. While federated systems enhance privacy, they require significant investment in infrastructure, standardization, and coordination. The heterogeneity of electronic health records and clinical terminologies in the Madrid region remains a challenge despite the harmonization efforts. Moreover, although the governance model is designed to ensure transparency, the use of opt-out consent mechanisms may provoke resistance from patients and clinicians who are concerned about autonomy and data misuse [1,29,30]. Finally, the risk of disproportionate benefit by private actors must be actively managed to preserve equity and public trust.

In conclusion, HealthData@MAD-R&I offers a pioneering approach to unlocking the potential of secondary health data for research, clinical decision-making, and public health. Its integrated governance, advanced technical architecture, and validation through impactful use cases position Madrid as a leader in digital health innovation. The long-term success of the project will depend on its ability to address persistent challenges related to interoperability, ethical governance, and public engagement. If these are met, the initiative can serve as a blueprint for other regions aiming to build transparent, secure, and effective health data ecosystems across Europe.

Acknowledgments

The authors would like to acknowledge the health care professionals of the Madrid Regional Health Service (SERMAS), whose daily clinical work and systematic recording of health information in electronic health records make this project possible. We thank the HealthData@MAD-R&I working group for their contributions to the development of the protocol and the data infrastructure. The project is supported by the institutional commitment of the Ministry of Health and the Ministry of Digitalization of the Regional Government of Madrid and coordinated by the Fundación para la Investigación e Innovación Biosanitaria de Atención Primaria (FIIBAP). We are also grateful to the Plataforma de Organizaciones de Pacientes (POP), Farmaindustria, the Fundación para la Investigación Biomédica of Hospital Clínico San Carlos (FIBHCSC), the Gerencia Asistencial de Atención Primaria, the Instituto de Investigación Sanitaria del Hospital 12 de Octubre (Imas12), and both Directorates General of the Regional Government of Madrid for their support and institutional engagement.

The name HealthData@MAD R&I has been officially registered with the Spanish Patent and Trademark Office. This registration was undertaken solely to protect the project identity and prevent third-party use. The registration number is not provided, as it is linked to the legal identification of the registering entity (Fundación para la Investigación e Innovación Biosanitaria de Atención Primaria, Madrid, Spain).No generative artificial intelligence tools (such as ChatGPT or other large language models) were used in the writing, editing, or submission of this manuscript.

The contributors associated with HealthData-MAD R&I working group are as follows:

Ana García Sacristán, Laura Varo González, Idoia Ramos Cuevas, Leticia Cárdenas Pozuelo, Mónica Luque Rodriguez, Maria Sanchez Besga, María Beneito Durá, Diego Trancho Pérez, Ainhoa Caballero Cardenete, Miguel Ángel Jiménez Peralta, María Martínez Alonso, Cristina Muntañola Valero, Clara Muñoz Capilla, Tomas Navarro Alvarez, Olatz Alonso Del Cura, Isabel

XSL•FO
RenderX

Amaya Rodriguez, Patricia Medina Burgos, Juan Luis González Rodríguez, Laura Rivero Lago, Lucia Rocío Camacho Montaño, Francisco Javier Somolinos Simón, Carlos Arias Alcaide, M Pilar Losantos Albacete, Gloria Villar Acebedo: Fundación para la Investigación e Innovación Biosanitaria de Atención Primaria (FIIBAP), Madrid, Madrid, Spain; Paula Rubio Mayo: Grupo de Investigación e Innovación en Transformación Digital e Ingeniería Biomédica, Hospital Universitario 12 de Octubre, Instituto de Investigación Sanitaria Hospital 12 de Octubre (i+12), Research Institute Hospital 12 de Octubre, Madrid, Madrid, Spain; Miguel Ángel Sicilia Urbán: Computer Science Deparment, University of Alcalá, Alcalá de Henares, Madrid, Spain.

Funding

This project ("HealthData@MAD-R&I", code TSI-100121-2024-79) is funded by the Ministry for Digital Transformation and Public Service and by the European Union through financial resources derived from the European Recovery Instrument ("Next Generation EU"), within the framework of the Recovery, Transformation, and Resilience Plan.

However, the views and opinions expressed are solely those of the authors of this manuscript and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

This project has received funding from Horizon Europe programme under grant agreement no. 101080711 (SPIDERR).

Data Availability

Data sharing is restricted while the HealthData@MAD-R&I infrastructure remains under development. However, metadata, project documentation, and public deliverables are available through the RIMASalud repository. Upon completion of the project, additional non-sensitive materials, including technical documentation and example scripts, will also be deposited in RIMASalud for open access. Requests for access to restricted datasets can be submitted through the governance procedures of the Madrid Regional Health System (SERMAS) once the Secure Processing Environments become operational.

Authors' Contributions

Conceptualization: AIGG, ML-G, SA-P, JG-G, CU-M, SS-D, JC-V, LR-R, AD-H, IC-G, JLC-B, NG-B, CG-F, CR-A, LG-A, EMM-S, ARL, JMVA-C, MAS-F
Data curation: ML-G, SA-P, AIGG
Formal analysis: ML-G, SA-P, AIGG
Funding acquisition: AIGG, MAS-F
Investigation: ML-G, SA-P, AIGG
Methodology: ML-G, SA-P, JG-G, CU-M, SS-D, JC-V, LR-R, AD-H, IC-G, JLC-B, NG-B, CG-F, CR-A, LG-A, EMM-S, ARL, JMVA-C, AIGG, MAS-F, HealthData@MAD-R&I Working Group
Project administration: AIGG, JC-V, EMM-S, ARL, JMVA-C
Resources: HealthData@MAD-R&I Working Group
Supervision: AIGG
Writing – original draft: AIGG, ML-G, SA-P
Writing – review & editing: AIGG, ML-G, SA-P, JG-G, CU-M, SS-D, JC-V, LR-R, AD-H, IC-G, JLC-B, NG-B, CG-F, CR-A, LG-A, EMM-S, ARL, JMVA-C, MAS-F

Conflicts of Interest

The authors declare no conflicts of interest related to this work. The project described in this paper is funded by the Ministry for Digital Transformation and the Civil Service of Spain under the Data Spaces Program (Exp. No. EDS2023/XXXX) and co-coordinated by the Directorate General for Research and Teaching and the Directorate General for Digital Health of the Madrid Regional Health System (SERMAS). None of the authors hold any financial interests that could be perceived to influence the outcomes reported in this manuscript.

References

1. Naude K, Snowdon DA, Parker E, McNaney R, Srikanth V, Andrew NE. Sharing data matters: exploring the attitudes of older consumers on an emerging healthy ageing data platform using electronic health records for research. BMJ Health Care Inform 2025;32(1):e101126 [FREE Full text] [doi: 10.1136/bmjhci-2024-101126] [Medline: 39753271]
2. Schneider ALC, Ginestra JC, Kerlin MP, Shashaty MGS, Miano TA, Herman DS, et al. The complete inpatient record using comprehensive electronic data (CIRCE) project: a team-based approach to clinically validated, research-ready electronic health record data. Learn Health Syst 2025;9(1):e10439 [FREE Full text] [doi: 10.1002/lrh2.10439] [Medline: 39822919]
3. Pedrera-Jiménez M, García-Barrio N, Frid S, Moner D, Boscá-Tomás D, Lozano-Rubí R, et al. Can openEHR, ISO 13606, and HL7 FHIR work together? An agnostic approach for the selection and application of electronic health record standards

XSL•FO
RenderX

to the next-generation health data spaces. J Med Internet Res 2023;25:e48702 [FREE Full text] [doi: 10.2196/48702] [Medline: 38153779]
4. Peeters LM. The arisal of data spaces: why I am excited and worried. Front Immunol 2024;15:1461361 [FREE Full text] [doi: 10.3389/fimmu.2024.1461361] [Medline: 39502694]
5. Staunton C, Shabani M, Mascalzoni D, Mežinska S, Slokenberga S. Ethical and social reflections on the proposed European health data space. Eur J Hum Genet 2024;32(5):498-505 [FREE Full text] [doi: 10.1038/s41431-024-01543-9] [Medline: 38355959]
6. Mateus M, Loureiro M, Fernandes AR, Oliveira M, Cruz-Correia R. Implementation status of the proposal for a regulation of the European health data space in Portugal: are we ready for it? Stud Health Technol Inform 2023;302:48-52. [doi: 10.3233/SHTI230062] [Medline: 37203607]
7. Raab R, Küderle A, Zakreuskaya A, Stern AD, Klucken J, Kaissis G, et al. Federated electronic health records for the European health data space. Lancet Digit Health 2023;5(11):e840-e847 [FREE Full text] [doi: 10.1016/S2589-7500(23)00156-5] [Medline: 37741765]
8. European Commission. European health data space (EHDS). European Commission. URL: https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en [accessed 2026-02-18]
9. Marco C, Atienza M. El espacio europeo de datos sanitarios: la propuesta de reglamento de la Unión Europea. Derecho y Salud. 2024. URL: https://www.ajs.es/sites/default/files/2024-11/Revista_DS_Vol34_Extraordinario.pdf [accessed 2026-02-18]
10. Marelli L, Stevens M, Sharon T, Van Hoyweghen I, Boeckhout M, Colussi I, et al. The European health data space: too big to succeed? Health Policy 2023;135:104861 [FREE Full text] [doi: 10.1016/j.healthpol.2023.104861] [Medline: 37399677]
11. Regulation (EU) 2025/327 of the European parliament and of the council of 11 February 2025 on the European health data space and amending directive 2011/24/EU and regulation (EU) 2024/2847. EUR-Lex. 2025. URL: http://data.europa.eu/eli/reg/2025/327/oj [accessed 2026-02-18]
12. Tan AC, Webster AC, Libesman S, Yang Z, Chand RR, Liu W, et al. Data sharing policies across health research globally: cross-sectional meta-research study. Res Synth Methods 2024;15(6):1060-1071. [doi: 10.1002/jrsm.1757] [Medline: 39275943]
13. van Panhuis WG, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst AJ, et al. A systematic review of barriers to data sharing in public health. BMC Public Health 2014;14:1144 [FREE Full text] [doi: 10.1186/1471-2458-14-1144] [Medline: 25377061]
14. Tamuhla T, Lulamba ET, Mutemaringa T, Tiffin N. Multiple modes of data sharing can facilitate secondary use of sensitive health data for research. BMJ Glob Health 2023;8(10):e013092 [FREE Full text] [doi: 10.1136/bmjgh-2023-013092] [Medline: 37802544]
15. Lewerenz S, Moen A, Martins H. Public value and digital health: the example of guiding values in the national digital health strategy of France. Int J Med Inform 2025;196:105794 [FREE Full text] [doi: 10.1016/j.ijmedinf.2025.105794] [Medline: 39862565]
16. Lee WB, Choi SJ. Secondary use provisions in the European health data space proposal and policy recommendations for Korea. Healthc Inform Res 2023;29(3):199-208 [FREE Full text] [doi: 10.4258/hir.2023.29.3.199] [Medline: 37591675]
17. Secretaría GDSD. Estrategia de salud digital. Secretaría General de Salud Digital, Información e Innovación para el SNS. URL: https://www.mscbs.gob.es/areas/saludDigital/doc/Estrategia_de_Salud_Digital_del_SNS.pdf [accessed 2026-02-20]
18. Sigi AL. Designing data governance with DAMA DMBOK framework. Institut Teknologi Sepuluh Nopember 2024;8(2):79-89 [FREE Full text] [doi: 10.12962/j24609463.v8i2.1408]
19. Broca A, Gutierrez T, Jimenez G. AGILE methodology in human management. Latitude: Multidisciplinary Research Journal 2023;1(17):90-100. [doi: 10.55946/latitude.v1i17.214]
20. ISO/IEC 20889:2018(en) Privacy enhancing data de-identification terminology and classification of techniques. ISO TIO for SI the IEC. 2018. URL: https://www.iso.org/standard/69373.html [accessed 2026-02-18]
21. Frid S, Pastor Duran X, Bracons Cucó G, Pedrera-Jiménez M, Serrano-Balazote P, Muñoz Carrero A, et al. An ontology-based approach for consolidating patient data standardized with European norm/international organization for standardization 13606 (EN/ISO 13606) into joint observational medical outcomes partnership (OMOP) repositories: description of a methodology. JMIR Med Inform 2023;11:e44547 [FREE Full text] [doi: 10.2196/44547] [Medline: 36884279]
22. Bossenko I, Randmaa R, Piho G, Ross P. Interoperability of health data using FHIR mapping language: transforming HL7 CDA to FHIR with reusable visual components. Front Digit Health 2024;6:1480600 [FREE Full text] [doi: 10.3389/fdgth.2024.1480600] [Medline: 39749099]
23. Picavet HSJ, Schouten JSAG. Musculoskeletal pain in the Netherlands: prevalences, consequences and risk groups, the DMC3-study. Pain 2003;102:167-178. [doi: 10.1016/s0304-3959(02)00372-x]
24. Hensvold A, Klareskog L. Towards prevention of autoimmune diseases: the example of rheumatoid arthritis. Eur J Immunol 2021;51(8):1921-1933 [FREE Full text] [doi: 10.1002/eji.202048952] [Medline: 34110013]

XSL•FO
RenderX

25. Cieza A, Causey K, Kamenov K, Hanson SW, Chatterji S, Vos T. Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the global burden of disease study 2019. Lancet 2021;396(10267):2006-2017 [FREE Full text] [doi: 10.1016/S0140-6736(20)32340-0] [Medline: 33275908]
26. Lahoz C, Cárdenas-Valladolid J, Salinero-Fort M, Mostaza JM. Use of statins and associated factors in nonagenarians in the community of Madrid, Spain. Aging Clin Exp Res 2022;34(2):439-444 [FREE Full text] [doi: 10.1007/s40520-021-01945-5] [Medline: 34363590]
27. Horgan D, Hajduch M, Vrana M, Soderberg J, Hughes N, Omar MI, et al. European health data space-an opportunity now to grasp the future of data-driven healthcare. Healthcare (Basel) 2022;10(9):1629 [FREE Full text] [doi: 10.3390/healthcare10091629] [Medline: 36141241]
28. Hripcsak G, Duke J, Shah N, Reich CG, Huser V, Schuemie MJ, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. Stud Health Technol Inform 2015;216:574-578 [FREE Full text] [doi: 10.3233/978-1-61499-564-7-574] [Medline: 26262116]
29. Baines R, Stevens S, Austin D, Anil K, Bradwell H, Cooper L, et al. Patient and public willingness to share personal health data for third-party or secondary uses: systematic review. J Med Internet Res 2024;26:e50421 [FREE Full text] [doi: 10.2196/50421] [Medline: 38441944]
30. Skovgaard L, Ekstrøm CT, Svendsen MN, Hoeyer K. Survey of attitudes in a Danish public towards reuse of health data. PLoS One 2024;19(12):e0312558 [FREE Full text] [doi: 10.1371/journal.pone.0312558] [Medline: 39724084]

Abbreviations

AI: artificial intelligence
CMBD: Conjunto Mínimo Básico de Datos
DAMA: Data Management Association
DGID: General Directorate of Health Research and Teaching
DGSD: General Directorate of Digital Health
DPIA: Data Protection Impact Assessment
EHDS: European Health Data Space
EHR: Electronic Health Record
ENI: National Interoperability Framework
ENS: National Security Framework
EU: European Union
FCN: Federated Coordination Node
FHIR: Fast Healthcare Interoperability Resources
GDPR: General Data Protection Regulation
HCIS: Healthcare Information System
HL7: Health Level 7
ICD-10: International Statistical Classification of Diseases and Related Health Problems 10th Revision
IEC: International Electrotechnical Commission
INE: National Institute of Statistics
ISO: International Organization for Standardization
LOINC: Logical Observation Identifiers, Names, and Codes
ML: machine learning
NLP: natural language processing
OMOP: Observational Medical Outcomes Partnership
RMD: rheumatologic musculoskeletal disease
SERMAS: Madrid Health Service (Servicio Madrileño de Salud)
SHAP: Shapley Additive Explanations
SNOMED cT: Systematized Nomenclature of Medicine Clinical Terms
SPE: secure processing environment
WP: work package

XSL•FO
RenderX

Edited by J Sarvestan; submitted 25.Aug.2025; peer-reviewed by F Estupinan-Romero, K Adegoke; comments to author 06.Oct.2025; accepted 27.Dec.2025; published 20.Mar.2026.

Please cite as:
León-García M, Álvarez-Pérez S, Gesto-Gómez J, Urbano-Molina C, Soto-Díaz S, Cárdenas-Valladolid J, Rodríguez-Rodríguez L, Díaz-Holgado A, del Cura-González I, De La Cruz-Bertolo J, García-Barrio N, Cruz-Bermúdez JL, García-Fernández C, Rodríguez-Antolín C, García-Aldars L, Moreda-Sánchez EM, Roldán López Á, Veganzones Alonso-Cortés JM, Gonzalez Gonzalez AI, Salinero-Fort MA, HealthData@MAD-R&I Working Group
HealthData@MAD-R&I: Protocol for Design and Development of a Regional Health Data Infrastructure to Enable Secondary Use of Health Data in Research and Innovation
JMIR Res Protoc 2026;15:e82815
URL: https://www.researchprotocols.org/2026/1/e82815
doi:10.2196/82815
PMID:41861842

©Montserrat León-García, Sergio Álvarez-Pérez, Janire Gesto-Gómez, Clara Urbano-Molina, Sonia Soto-Díaz, Juan Cárdenas-Valladolid, Luis Rodríguez-Rodríguez, Antonio Díaz-Holgado, Isabel del Cura-González, Javier De La Cruz-Bertolo, Noelia García-Barrio, Juan Luis Cruz-Bermúdez, Cristina García-Fernández, Carlos Rodríguez-Antolín, Laila García-Aldars, Elsa María Moreda-Sánchez, Álvaro Roldán López, José María Veganzones Alonso-Cortés, Ana Isabel Gonzalez Gonzalez, Miguel A Salinero-Fort, HealthData@MAD-R&I Working Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Mobile Imaging–Based Machine Learning for Dental Caries, Sealants, and Fluorosis: Protocol for a Cross-Sectional Model Development and Validation Study

Sang Mok Park[1*], PhD; Semin Kwon[1*], PhD; Shaun G Hong[1*], MS; Yuhyun Ji[1], PhD; Sreeram P Nagappa[1], MS; Jung Woo Leem[1], PhD; Mei Lin[2], MPH, MD; Eugenio D Beltrán-Aguilar[3,4], DMD, MPH, DrPH, DABDPH; Susan O Griffin[5], PhD; Young L Kim[1,6,7], PhD

[1]Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, United States

[2]Dental Public Health Consultant, Atlanta, GA, United States

[3]Department of Epidemiology and Health Promotion, New York University, New York, NY, United States

[4]Department of Oral Health Sciences, Kornberg School of Dentistry, Temple University, Philadelphia, PA, United States

[5]TIS Consulting Group, State College, PA, United States

[6]Purdue Institute for Cancer Research, Purdue University, West Lafayette, IN, United States

[7]Regenstrief Center for Healthcare Engineering, Purdue University, West Lafayette, IN, United States

[*]these authors contributed equally

Corresponding Author:
Young L Kim, PhD
Weldon School of Biomedical Engineering
Purdue University
206 S. Martin Jischke Drive
West Lafayette, IN, 47907
United States
Phone: 1 765 496 2445
Email: youngkim@purdue.edu

Abstract

Background: Assessing dental caries, sealants, and fluorosis is essential for public health surveillance, providing critical data to evaluate national prevention programs. Standard methods performed by dental professionals are often limited by affordability, accessibility, and scalability for both population-level and individualized assessments. Mobile health (mHealth) approaches to concurrently detect caries, sealants, and fluorosis have remained largely unexplored, especially at the population level.

Objective: This study leverages mHealth technologies that integrate computer vision using machine learning and deep learning with images captured by smartphone cameras and low-cost intraoral cameras. The primary objective is to develop and validate models for detecting caries lesions, identifying sealants, and quantifying fluorosis severity from standardized dental images, using standardized visual clinical examinations as the reference standard.

Methods: The proposed study population will include approximately 1000 adolescents in Colorado, United States, living in communities with naturally elevated fluoride levels in the public water system. Participants will undergo standardized clinical dental examinations and imaging using intraoral cameras and smartphones. Supervised learning models will incorporate reference chart–based color correction, radiomic spatial and textural features, and neural network classifiers. The reference standard will be standardized visual clinical examinations performed by trained and calibrated dental professionals. Two models will be developed and evaluated: one to detect caries lesions and sealants and another to assess fluorosis severity. Model performance will be evaluated against clinical assessments by dental professionals using stratified cross-validation and multiclass performance metrics while minimizing bias and accounting for confounders common to human examiners.

Results: A standardized dental examination, an intraoral imaging protocol, and a smartphone imaging protocol are used to assess all 8 permanent molars for caries and sealants, as well as the 6 upper anterior teeth for fluorosis severity. Pilot studies were conducted to test study logistics and calibrate 3 examiners in person, supplemented by debriefings, mobile app training, and a web-based calibration module. The study was funded in September 2022 with supplemental funding awarded in June 2024. The study launched in May 2024, and as of January 2026, data have been collected from approximately 300 participants.

XSL•FO
RenderX

Conclusions: The integration of computer vision and mobile device imaging will enable affordable, scalable, population-level assessments for detecting caries and sealants and quantifying fluorosis severity among adolescents. mHealth technologies have been increasingly incorporated into dentistry for both clinical decision support and at-home use. This protocol will further help establish a structured methodological framework for acquiring, processing, and analyzing mobile imaging data for dental health surveillance and epidemiological studies.

International Registered Report Identifier (IRRID): DERR1-10.2196/91239

(JMIR Res Protoc 2026;15:e91239) doi:10.2196/91239

KEYWORDS

mobile health; machine learning; computer vision; dental caries; dental fluorosis; dental sealants; intraoral camera; smartphone imaging; adolescent oral health

Introduction

Dental caries (tooth decay) remains one of the most prevalent diseases of childhood [1,2]. If left untreated or unmanaged, it can impair a child's ability to eat, speak, learn, and play [3-10]. Importantly, tooth decay is largely preventable with the use of sealants and fluorides. Reducing the burden of tooth decay and improving access to sealants and community water fluoridation are national health priorities [11-15]. However, excessive fluoride intake from drinking water, toothpaste, and other sources during the critical period of tooth development (up to approximately 6-8 years of age for most teeth) increases the risk of fluorosis, which manifests as enamel changes ranging from faint white striations in mild cases to pitting in severe forms [11-15].

Monitoring the prevalence of caries, sealants, and fluorosis has primarily been conducted at the national level through the National Health and Nutrition Examination Survey (NHANES) [16]. Some states have also collected data on caries and sealants using the Basic Screening Survey [17]. These surveillance efforts are costly, as they require trained dental professionals and portable examination equipment. Moreover, decisions regarding public health sealant programs and community water fluoridation are typically made at the local level, where relevant data are often limited or unavailable. As a result, affordable, valid, and reliable alternatives to traditional surveillance methods are urgently needed [18].

Dental clinical photography plays a critical role in diagnosis, monitoring, and treatment management [19-29]. Digital and mobile dental health applications using mobile devices (eg, smartphones and tablets) are increasingly used in telemedicine and remote care, particularly in home-based settings and in regions with limited access to dental care infrastructure [30-34]. Mobile health (mHealth) approaches leverage smartphones, tablets, and intraoral cameras to facilitate screening, treatment planning, remote examinations, and patient education, particularly in resource-limited settings and underserved populations [35,36]. Mobile devices are also well-suited for population-based dental public health surveillance in diverse settings at both local and national levels. The use of low-cost intraoral cameras and smartphones offers practical advantages for scalability, as these tools are widely available, user-friendly, and far less resource intensive than traditional mobile dental units and examinations performed by dental professionals. Their cost-effectiveness makes them especially attractive for public health programs operating under tight budget constraints.

Automated computer vision using machine learning and deep learning is an active area of research in oral health, often focused on caries detection [37-43]. However, evidence supporting the validity and feasibility of computer vision–integrated mHealth tools for jointly detecting caries, sealants, and fluorosis remains limited, especially for population-level use. Fluorosis detection, in particular, requires extensive examiner training and standardization, as current assessment measures are subjective [44,45]. Building on the promise of mHealth dental applications for both large-scale screening and individualized care [46-51], the mobility, simplicity, and affordability of mobile devices further support the integration of machine learning into population-level surveillance and public health monitoring. To date, no comprehensive mHealth protocol has proposed computer vision methods that combine caries and sealant detection using affordable intraoral cameras with smartphone fluorosis assessment for population-level surveillance and epidemiological studies.

The primary objective of this study is to develop and validate machine learning models using dental images for assessing dental health among middle- and high-school students (aged approximately 13-15 years), compared with the status quo of standardized visual clinical examinations. Specifically, the study aims to develop automated models for (1) detecting caries lesions and sealants and (2) detecting the presence and severity of fluorosis. With an emphasis on population-level surveillance, the models will focus on (1) caries and sealants in the permanent first and second molars and (2) fluorosis in the upper anterior teeth using Dean's Index–based fluorosis ratings [45,52].

The central hypothesis is that supervised learning models trained on intraoral camera images of molar occlusal surfaces and smartphone images of the upper anterior teeth can achieve detection performance comparable to the status quo of direct clinical evaluation by dental professionals. The reference standard for the supervised learning models will be the results of standardized visual clinical examinations performed by trained and calibrated dental professionals. Affordable intraoral cameras will be used for assessing posterior teeth due to limited physical space, as they provide higher diagnostic accuracy for caries than onboard smartphone cameras [53]. For fluorosis detection in anterior teeth, color correction using an application-specific color reference chart can enable reliable

XSL•FO
RenderX

assessments with conventional smartphone cameras [54,55]. Smartphones are also chosen for their scalability and efficiency, as a single photograph can capture multiple anterior teeth simultaneously, supporting high-throughput population-level surveillance.

Methods

Overall Study Design

This is a noninterventional, cross-sectional observational study involving the acquisition of intraoral camera images of molar occlusal surfaces and smartphone images of the upper anterior teeth alongside standardized visual clinical examinations in adolescents. Assessments of caries and sealants and fluorosis severity will be used solely for research purposes. All data collection will take place during a single study visit; thus, a retention plan is not required.

Setting and Recruitment

Data will be collected in selected middle and high schools from communities in Colorado, United States, with natural fluoride levels higher than 1 ppm to ensure a wide range of clinical presentations. According to data from the US Centers for Disease Control and Prevention (CDC) and the National Center for Education Statistics [56], 25 counties in Colorado meet these criteria, each with at least one qualifying water system and one school. Participant recruitment will be supported through coordination with school administrators and local community health coordinators via informational flyers, emails, and brief in-person presentations.

Caries and sealants are common among adolescents aged 12 to 15 years with a prevalence of 56.8% and 51.7%, respectively [57], whereas moderate to severe fluorosis affects fewer than 4% of US adolescents overall [58]. As fluorosis prevalence and severity increase with higher fluoride exposure, recruitment from areas with drinking water fluoride concentrations of ≥1.0 ppm is needed to obtain an adequate number of moderate or severe fluorosis images for model training. Among the 20 state oral health programs funded by the CDC, Colorado serves the largest population on public water systems with higher naturally occurring fluoride, and some of the earliest studies linking water fluoride content to fluorosis severity were conducted in these counties [52].

Study Participants

The study aims to enroll up to 1000 students between the ages of 13 and 15 years. This age group is of particular interest due to the high prevalence of caries, the likelihood that sealants have already been applied, and the full eruption of the upper anterior teeth with varying degrees of fluorosis. In addition, adolescents in this age range are typically better able to follow instructions and cooperate during data collection procedures, enabling effective implementation of standardized protocols for both intraoral camera imaging and smartphone imaging.

The proposed sample size is anticipated to provide a sufficient number of outcome-positive cases to train the proposed models, mitigate overfitting risk, and support robust validation. We will split the dataset at the participant level (70% for training and 30% for testing); thus, a target sample size of 1000 yields approximately 700 participants for model development and approximately 300 independent participants for separate testing. The testing dataset will not be used or seen during model training and learning. This split helps ensure the test set contains enough positive cases while preventing data leakage during model selection. In similar photo-based studies, we successfully developed and validated machine learning models using data from approximately 400 to 500 participants [59,60].

Inclusion and Exclusion Criteria

Consenting participants must agree to have digital images of their teeth captured using an intraoral camera and a smartphone. Potential participants will be excluded if they have fixed orthodontic appliances or other dental treatments that obstruct clear visualization of the anterior buccal and posterior occlusal tooth surfaces or if they have underlying medical conditions that make it unsafe or impractical to conduct dental assessments in a school setting (eg, conditions requiring antibiotic prophylaxis that cannot be administered onsite). These criteria ensure that all participants can safely and comfortably complete the clinical and imaging components of the study.

Standardized Clinical Evaluation Procedure

The clinical protocol is adapted from the dental examination method used in the NHANES, ensuring the standardized dental public health surveillance procedures and facilitating adequate access to the intraoral environment within a constrained time frame. First, participants will be given simple instructions on how to place a disposable dental retractor in the mouth to facilitate data collection. Second, the examiner will visually inspect all 8 permanent molars for caries lesions and the presence of sealants, categorizing each tooth as one of the following: sound; sealed; untreated cavitated decay; and restored, including crowns, missing, or unerupted or cannot be assessed. The occlusal surfaces of the molars will be dried using gauze prior to the assessments of caries and sealant presence. Third, the examiner will assess the 6 upper anterior teeth to determine fluorosis severity using the Dean's Index, classifying each tooth’s buccal surface as one of the following: normal/questionable; very mild; mild; moderate; severe; or cannot be assessed. Participants will be instructed to moisten the surfaces of the anterior teeth using their tongue and saliva in accordance with the Dean's Index evaluation protocol [45,52,61]. All clinical data will be registered in the data collection form (Figure 1).

XSL•FO
RenderX

Figure 1. Dental screening form.

Examiner Initials:	Date (MM/DD/YYYY):
School Code:	Grade:
Student Pseudo ID:	Sex: ☐ Female ☐ Male / Age (Years):

Caries and Sealant Status of Permanent First and Second Molars

Tooth	Tooth Number	Sound	Sealed	Untreated cavitated decay	Restoration including crown	Missing	Unerupted or cannot be assessed
UR2M	2	☐	☐	☐	☐	☐	☐
UR1M	3	☐	☐	☐	☐	☐	☐
UL1M	14	☐	☐	☐	☐	☐	☐
UL2M	15	☐	☐	☐	☐	☐	☐
LL2M	18	☐	☐	☐	☐	☐	☐
LL1M	19	☐	☐	☐	☐	☐	☐
LR1M	30	☐	☐	☐	☐	☐	☐
LR2M	31	☐	☐	☐	☐	☐	☐

Fluorosis Severity of Upper Anterior Teeth

Tooth	Tooth Number	Normal/Questionable	Very Mild	Mild	Moderate	Severe	Missing, unerupted, restoration/crown or other reason cannot be assessed
URC	6	☐	☐	☐	☐	☐	☐
URLI	7	☐	☐	☐	☐	☐	☐
URCI	8	☐	☐	☐	☐	☐	☐
ULCI	9	☐	☐	☐	☐	☐	☐
ULLI	10	☐	☐	☐	☐	☐	☐
ULC	11	☐	☐	☐	☐	☐	☐

Treatment Urgency:

☐ Routine
☐ Early
☐ Urgent

Notes

UR2M: Upper right second permanent molar. UR1M: Upper right first permanent molar. UL1M: Upper left first permanent molar. UL2M: Upper left second permanent molar. LL2M: Lower left second permanent molar. LL1M: Lower left first permanent molar. LR1M: Lower right first permanent molar. LR2M: Lower right second permanent molar.

URC: Upper right cuspid. URLC: Upper right lateral incisor. URCI: Upper right central incisor. ULCI: Upper left central incisor. ULLI: Upper left lateral incisor. ULC: Upper left cuspid.

Intraoral Camera Imaging Procedure

The examiner will capture images of the occlusal surfaces of the molars using an intraoral camera wirelessly connected to a smartphone. For each participant, the examiner will capture 2 to 3 images of each of the 8 permanent molars after drying the teeth with gauze, generating a total of approximately 24 images per participant. A disposable, transparent protective sleeve will be fitted over the distal tip of the intraoral camera for infection control purposes. To ensure affordability and scalability, we selected a consumer-grade intraoral camera (Protector International) designed for home use. The device includes a built-in white LED light source and is available for approximately US $50 per unit. As a key specification, we evaluated the spatial resolution of an intraoral camera using the edge method [62,63], resulting in a spatial resolution of 37 µm. In a pilot study conducted as part of the data collection project described in Multimedia Appendix 1, images taken by an intraoral camera with the protective sleeve in place provided a clear visualization of the occlusal surfaces, which is essential for detecting caries and assessing sealants (Figure 2).

XSL•FO
RenderX

Figure 2. Representative images of the permanent molars. An affordable intraoral camera, wirelessly connected to a smartphone, captures detailed images of the occlusal surfaces of permanent molars. Computer vision and machine learning models using intraoral camera images are developed to automatically and interpretably detect caries and sealants. LL1M: lower left first molar; LL2M: lower left second molar; LR1M: lower right first molar; LR2M: lower right second molar; UL1M: upper left first molar; UL2M: upper left second molar; UR1M: upper right first molar; UR2M: upper right second molar.

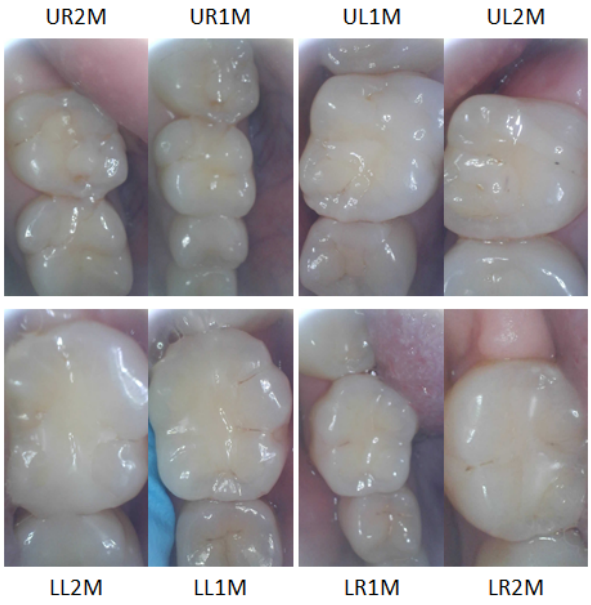

Smartphone Camera Imaging Procedure

The examiner will take images of the 6 upper anterior teeth from 3 angles (left, front, and right) using a smartphone, immediately after the participant moistens the surfaces of their teeth using their tongue and saliva, with the retractor in place. The front-angle image will capture the 2 central incisors, while the left and right angles will each include one lateral incisor and one cuspid (Figure 3). For each participant, 2 to 3 images will be taken from each of the 3 angles, resulting in a total of approximately 9 images per participant. We evaluated the spatial resolution of a typical smartphone camera (Samsung Galaxy S21) that will be used for the study, using the edge method in a laboratory setting [62,63], yielding a spatial resolution of 170 µm at a typical distance of 100 mm between the camera and the participant. Images captured using smartphone cameras can exhibit color variability due to differences in device models, light conditions, and image file formats. Achieving accurate and precise color quantification is a complex task that goes beyond simple white balancing or perceptual color constancy [54,64], requiring high-fidelity color standardization and canonicalization.

Figure 3. Representative images of upper anterior teeth alongside a customized color chart. A conventional smartphone camera captures detailed images of a participant's anterior teeth with a retractor in place. The customized color chart is used to recover the absolute color of the teeth, regardless of device models, light conditions, or image file formats. Computer vision and machine learning models using smartphone camera images are developed to automatically and interpretably assess fluorosis severity.

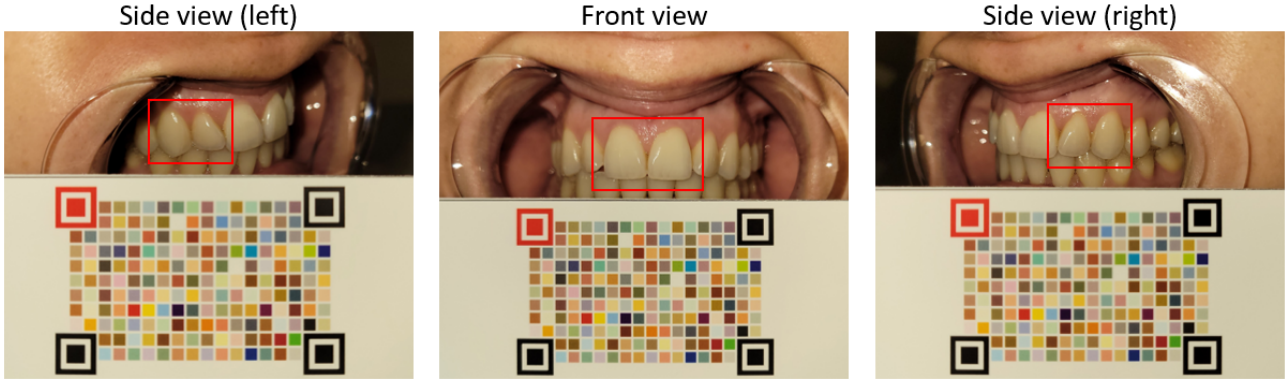

The image acquisition protocol will incorporate a custom-designed color reference chart to enable the recovery of absolute color values and the quantification of color changes for fluorosis assessment (Figure 3). Instead of commercially available color reference charts (eg, Macbeth ColorChecker or ColorChecker Classic Mini), a disposable custom-designed color chart will be used, which can be mass-produced using a commercially available inkjet printer [55]. As implemented in our previous studies [54,55], the customized color reference charts will be printed using the same inkjet printer (Canon imagePROGRAF PRO-1000) and the manufacturer-recommended paper (Canon Photo Paper Premium Fine Art Smooth). Print-to-print reproducibility and photostability testing of the printed charts support high

XSL•FO
RenderX

consistency across printing batches and strong resistance to color fading under prolonged light exposure, demonstrating the chart's practicality and scalability. Approximately the size of a business card, this chart is designed to be included in each image to facilitate color correction. To reduce the risk of infection and ensure accurate participant tracking, each color chart will include a participant ID number and will be appropriately disposed of after use. In a pilot study conducted as part of the data collection project described in Multimedia Appendix 1, images of the anterior teeth, juxtaposed with the customized color chart and taken with a Samsung Galaxy S21, provided a clear visualization of the anterior teeth from left, front, and right angles.

Data Collection Mobile App

We will use a custom-developed data collection mobile app to facilitate the collection and secure transfer of dental images in a similar manner to our previous study [65]. The app enables examiners to upload both intraoral camera and smartphone images, which are transferred to a Health Insurance Portability and Accountability Act (HIPAA)–compliant cloud server and accessed through a high-security portal. The app requires completion of designated form fields before selecting photos from the smartphone gallery for upload. The app also supports offline functionality: in the event of network interruption, images are temporarily saved in a local folder and automatically uploaded in the background, one at a time, once connectivity is restored to minimize data congestion. Once all uploads are complete, the temporary folder is deleted from the device.

Training and Calibration of Dental Examiners

Field dental examiners will be trained using standardized procedures to ensure consistent application of assessment criteria for identifying the presence of caries lesions, sealants, and fluorosis, in accordance with the protocols used by the NHANES [66]. Before data collection, all field examiners will complete a 3-day training program led by a designated reference examiner. To support multiple sites and examiners, we will develop a web-based calibration module in Qualtrics (Qualtrics, LLC; Figure 4) in which examiners will evaluate dental images for caries and fluorosis on a monitor at home under controlled conditions (with recommended monitor display and ambient light settings). The purpose of this web-based calibration is to ensure that color-corrected image evaluations align with examiners' in-person visual assessments and to enable examiners to calibrate against the reference examiner's image-reading evaluations. The evaluation time will be limited to 30 seconds for all 6 teeth to mimic the timing of in-person assessments. A background color similar to gingival tissue will be used in the web-based interface, and the tooth images will be color corrected using the custom color reference chart to ensure color consistency.

XSL•FO
RenderX

Figure 4. Representative web-based calibration module using Qualtrics. Field dental examiners evaluate color-corrected images on a home computer monitor under standardized viewing conditions. Evaluations are time-limited to mimic in-person participant assessments. This module uses a gingiva-like background color, and tooth images are color corrected using the custom color reference chart to ensure visual consistency. URC-6: upper right canine; URLI-7: upper right lateral incisor; URCI-8: upper right central incisor; ULCI-9: upper left central incisor; ULLI-10: upper left lateral incisor; ULC-11: upper left canine.

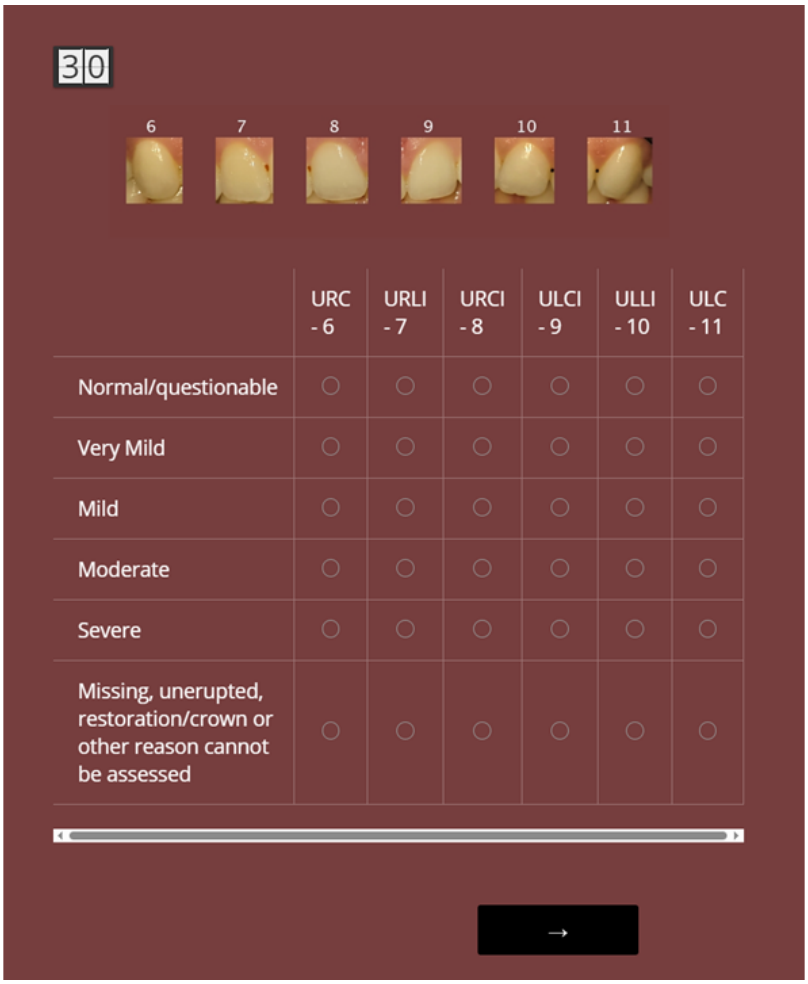

mHealth Detection Model Development

First, we will develop several mHealth detection models for caries and sealants using computer vision and machine learning approaches. The input will consist of intraoral camera images of the occlusal surfaces of the molars. The output of supervised learning models will be the clinical evaluation categorizing each tooth as one of the following: sound; sealed; untreated cavitated decay; or restored, including crowns, missing, or unerupted or cannot be assessed. We will leverage convolutional neural networks (CNNs), including the Visual Geometry Group network, a mask region–based CNN, and a residual network. We will apply data augmentation methods, including brightness, contrast, and saturation adjustments, as well as image flipping and random rotations. Our transfer learning strategy will incorporate pretrained models based on large-scale datasets for object detection, segmentation, and captioning, such as the Microsoft Common Objects in Context dataset and the ImageNet Large Scale Visual Recognition Challenge dataset. In addition, we will use radiomics, a technique widely applied across various medical imaging modalities. Radiomics enables the extraction of detailed spatial features (eg, texture, shape, and intensity) through predefined mathematical descriptors, providing insights beyond direct visual interpretation [59,60].

Second, we will develop a series of detection models for fluorosis using computer vision and machine learning approaches. The input will consist of smartphone images of the anterior teeth, and the output from supervised learning models will be fluorosis severity classifications based on the Dean's Index. Each tooth will be categorized as one of the following: normal/questionable; very mild; mild; moderate; severe; or cannot be assessed. The fluorosis assessment pipeline will include two main steps: (1) color correction will be applied to extract absolute color representations of the anterior teeth, ensuring consistency across different smartphone models, light conditions, and image file formats [54,55], and (2) CNN-based models will be trained using data augmentation and transfer learning to improve their robustness and accuracy. In addition, radiomics-based models will be developed to extract spatial and textural features specific to fluorosis assessment [59,60].

Performance Evaluation

We will evaluate the models exclusively on the testing dataset (30%) to prevent data leakage and reduce the risk of overfitting. The testing dataset refers to an independent dataset reserved for external validation and will not be used or accessed during model training or development. Cross-validation will also be performed to evaluate model performance across different

XSL•FO
RenderX

training data subsets. Specifically, k-fold cross-validation will be performed within the training dataset by partitioning the data into k approximately equal folds. In each iteration, the model will be trained on k−1 folds and evaluated on the remaining fold, rotating so that each fold serves as the validation set once. Performance will be summarized by averaging metrics across the k iterations, providing a robust estimate of model performance and supporting model selection and hyperparameter tuning. The held-out testing dataset will remain fully independent for final evaluation.

We will ensure comprehensive evaluations of each model's performance across all classes in the multiclass classification tasks against the reference standard performed by trained and calibrated dental professionals [67,68]. While accuracy provides a basic measure of performance, representing the proportion of correctly predicted instances across all classes, it can be misleading in the presence of class imbalance. Thus, precision, recall, and F_1-score will be computed for each class, along with their macro, micro, and weighted averages. Specifically, macroaveraging treats all classes equally, microaveraging aggregates contributions across all classes, and weighted averaging adjusts for class imbalance by weighting each class according to its support. In addition, confusion matrices will be used to visualize patterns of misclassification and to understand class-specific model performance better. Furthermore, TRIPOD-AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis-Artificial Intelligence) [69] and CREMLS (Consolidated Reporting Guidelines for Prognostic and Diagnostic Machine Learning Modeling Studies: Development and Validation) [70] will be used for reporting our machine learning−based diagnostic results.

Confidentiality, Data Storage, and Security

All study data will be stored and accessed in compliance with the HIPAA regulations. Specifically, image data will be labeled using participant identification numbers to maintain confidentiality. Demographic and clinical information recorded on paper forms by site personnel will be scanned and uploaded via the custom data collection app developed for this study, which transmits information to the secure HIPAA-compliant server. Access to the server is restricted to study investigators and authorized personnel. Electronic records will be stored on password-protected systems, and physical paper records will be kept in locked cabinets accessible only to authorized study staff.

Ethical Considerations

This study has been approved by the CDC Institutional Review Board (protocol 7494). Purdue University and the Colorado Department of Public Health and Environment are under a deferral to the CDC Institutional Review Board. Our study involves recruiting participants from vulnerable populations aged 13 to 15 years. Written consent will be obtained from parents or legal guardians, and written assent will be required from all adolescent participants to confirm their willingness to participate. This approach ensures clear communication about the study's objectives, procedures, and privacy protections, fostering a supportive and well-informed environment for both students and their families. All collected data will be deidentified and cannot be linked to individual participants. Participants will receive a free dental examination and an oral hygiene kit (including a toothbrush, toothpaste, and dental floss) as a token of appreciation for their participation.

Results

Protocol Development

We established the standardized school-based dental examination to assess all 8 permanent molars for caries and sealants and the 6 upper anterior teeth for fluorosis severity. The intraoral imaging protocol was developed in which examiners capture occlusal-surface images of the molars using an intraoral camera wirelessly connected to a smartphone (Figure 2). The smartphone imaging protocol was also established for capturing images of the 6 upper anterior teeth from 3 angles (ie, left, front, and right; Figure 3).

Pilot Studies

We conducted pilot studies in Colorado communities in May and September 2024. The first pilot study focused on test runs of the study logistics and provided in-person, hands-on training in the procedures for 3 dental hygienists as field dental examiners. Using printed images representing diverse cases of fluorosis, caries, and sealants, as well as 15 volunteer participants, a certified and experienced dental clinical and surveillance expert provided examiner training and calibration. The second pilot study further ensured examiner training and calibration in school-based settings over 3 days. It involved 55 volunteer participants and focused on acquiring high-quality photos and practicing data upload to the designated server. Debriefing sessions helped ensure that the clinical examiners were consistent and calibrated. During the pilot studies, the examiners also received training on how to use the app for data transfer. In addition, using this web-based calibration module (Figure 4), we verified that the expert's color-corrected image evaluations aligned with the expert's in-person visual assessments. We also validated that dental field examiners could score dental caries, sealant presence, and fluorosis severity across a range of cases using color-corrected images with expert feedback provided to refine scoring consistency.

Status

The procedure and protocol development were initially funded in September 2022, and supplemental funding was awarded in June 2024. The data collection began in May 2024. As of January 2026, approximately 300 participants had been enrolled. Results will be disseminated following completion of target enrollment and model development.

Discussion

To the best of our knowledge, this will be the first study to develop computer vision models for the joint surveillance of dental caries, sealants, and fluorosis severity using mobile devices among adolescents. Specifically, computer vision–assisted surveillance of fluorosis will be critical for monitoring population-level impact. The US Public Health

XSL•FO
RenderX

Service (USPHS) advises that US Department of Health and Human Services agencies continue to prioritize the development of valid and reliable measures for fluorosis assessment. In December 2012, the USPHS revised its recommendation for the optimal fluoride concentration in drinking water to 0.7 ppm for the prevention of tooth decay [11]. This level was intended to balance protection against caries with minimizing the risk of fluorosis, emphasizing the importance of monitoring both caries and fluorosis. In population-level studies, fluorosis assessment is typically performed by human examiners; however, examiner training and standardization are complex and often introduce bias, even when image-based scoring methods are used [45,71]. Recent cycles of US fluorosis data have exhibited substantial variation across survey cycles that cannot be explained by changes in national fluoride intake [16,18,58,72], indicating a need for more objective assessments.

Most previous studies using mobile devices, with or without computer vision (machine learning or deep learning), have focused on caries detection in resource-limited settings, including at-home and rural environments. Multiple studies have validated the use of smartphone—and intraoral camera—acquired images for assessing caries in both children and older adults with interpretation by dental professionals [37-41,73-77]. The use of machine learning for reliable caries detection has also been demonstrated in at-home settings [36,39,40,78-83]. For example, a study of parent-child dyads has demonstrated the feasibility of using smartphone images of the anterior and posterior teeth for at-home detection of caries with high parental acceptability [41,84]. Another study using intraoral images has shown that a deep learning model can simultaneously detect caries and fissure sealants with performance comparable to early-career dentists [42]. In addition, fluorescence imaging and other optical modalities have been explored for fluorosis detection [85,86], but these methods have yet to be applied in real-world field settings. Upon successful completion of this study, mHealth technologies can potentially be integrated into existing national surveillance efforts to strengthen the monitoring of progress toward national oral health objectives. Combining computer vision with clinical photos from mobile devices can further support low-cost, scalable, population-wide screening to detect caries, identify sealant presence, and assess fluorosis severity.

Acknowledgments

The authors thank the members of the Health Access Branch at the Colorado Department of Public Health and Environment (Robyn Maestas, Lennon Miell, Melissa Byrnes, and Ashleigh Kirk) for collecting preliminary data and the following experts for their valuable input on the data collection protocols: Dr Margherita Fontana (University of Michigan), Dr John Warren (University of Iowa), Dr Steven Levy (University of Iowa), Dr Jayanth Kumar (California Department of Public Health), and Dr Ekaterina Mauritson (Centers for Disease Control and Prevention). The generative AI tools ChatGPT (OpenAI) and Gemini (Google LLC) were used solely for grammar checking and language refinement during the preparation of this manuscript. The authors reviewed, verified, and finalized all outputs generated by these tools. YLK, SOG, ML, and EDB-A acknowledge support from the Centers for Disease Control and Prevention as former employees.

Funding

The protocol development and preliminary analyses were primarily supported by the US Centers for Disease Control and Prevention. Additional ancillary work was supported by the US National Institutes of Health Technology Accelerator Challenge Prize.

Data Availability

Data sharing is not applicable to this paper, as no datasets were analyzed during this study.

Authors' Contributions

SOG, ML, EDB-A, and YLK conceived the overall study design and protocol. SMP, SK, SGH, YJ, SPN, and JWL conducted device characterization and pilot computer vision and machine learning experiments. EDB-A provided clinical insights and training for the examiners. SOG, ML, and EDB-A contributed to the clinical and public health aspects of the study. SMP, SK, SPN, YLK, and EDB-A supported pilot training studies in the field. SMP, SK, and SGH prepared the initial draft of the paper. SOG, ML, EDB-A, and YLK contributed substantially to the main writing. SOG, ML, and YLK prepared the paperwork and secured Centers for Disease Control and Prevention funding. SOG and YLK supervised the project. All authors read and approved the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Federal Register notice describing the broader data collection project underlying this study: evaluating deep learning algorithm assessment of digital photographs for dental public health surveillance. Department of Health and Human Services, Centers for Disease Control and Prevention.
[PDF File (Adobe PDF File), 198 KB - resprot_v15i1e91239_app1.pdf]

References

1. Uribe SE, Innes N, Maldupa I. The global prevalence of early childhood caries: a systematic review with meta-analysis using the WHO diagnostic criteria. Int J Paediatr Dent 2021 Nov;31(6):817-830. [doi: 10.1111/ipd.12783] [Medline: 33735529]
2. Maklennan A, Borg-Bartolo R, Wierichs RJ, Esteves-Oliveira M, Campus G. A systematic review and meta-analysis on early-childhood-caries global data. BMC Oral Health 2024 Jul 24;24(1):835 [FREE Full text] [doi: 10.1186/s12903-024-04605-y] [Medline: 39049051]
3. Oral Health in America: A Report of the Surgeon General.: U.S. Department of Health and Human Services, U.S. Public Health Service, Office of the Surgeon General, National Institute of Dental and Craniofacial Research; 2000. URL: https://www.nidcr.nih.gov/sites/default/files/2017-10/hck1ocv.%40www.surgeon.fullrpt.pdf [accessed 2026-03-18]
4. Beltrán-Aguilar ED, Barker LK, Canto MT, Dye BA, Gooch BF, Griffin SO, Centers for Disease Control and Prevention (CDC). Surveillance for dental caries, dental sealants, tooth retention, edentulism, and enamel fluorosis--United States, 1988-1994 and 1999-2002. MMWR Surveill Summ 2005;54(3):1-43 [FREE Full text] [Medline: 16121123]
5. Dye BA, Li X, Thornton-Evans G. Oral health disparities as determined by selected healthy people 2020 oral health objectives for the United States, 2009-2010. NCHS Data Brief 2012;104:1-8 [FREE Full text] [Medline: 23101968]
6. Wei L, Griffin SO, Robison VA. Disparities in receipt of preventive dental services in children from low-income families. Am J Prev Med 2018;55(3):e53-e60 [FREE Full text] [doi: 10.1016/j.amepre.2018.04.039] [Medline: 30017612]
7. National Institute of Dental and Craniofacial Research. Oral Health in America: Advances and Challenges. Bethesda, MD: National Institutes of Health; Dec 2021.
8. Williams S, Wei L, Griffin SO, Thornton-Evans G. Untreated caries among U.S. working-aged adults and association with reporting need for oral health care. J Am Dent Assoc 2021;152(1):55-64 [FREE Full text] [doi: 10.1016/j.adaj.2020.09.019] [Medline: 33413851]
9. Griffin SO, Thornton-Evans G, Wei L, Griffin PM. Disparities in dental use and untreated caries prevalence by income. JDR Clin Trans Res 2021;6(2):234-241 [FREE Full text] [doi: 10.1177/2380084420934746] [Medline: 32585114]
10. Fontana M, Eckert GJ, Katz BP, Keels MA, Levy BT, Levy SM, et al. Predicting dental caries in young children in primary health care settings. J Dent Res 2023;102(9):988-998 [FREE Full text] [doi: 10.1177/00220345231173585] [Medline: 37329133]
11. US. Department of Health and Human Services Federal Panel on Community Water Fluoridation. U.S. public health service recommendation for fluoride concentration in drinking water for the prevention of dental caries. Public Health Rep 2015;130(4):318-331 [FREE Full text] [doi: 10.1177/003335491513000408] [Medline: 26346489]
12. Healthy People 2030. Office of Disease Prevention and Health Promotion. URL: https://odphp.health.gov/healthypeople [accessed 2026-03-06]
13. Mariño R, Zaror C. Economic evaluations in water-fluoridation: a scoping review. BMC Oral Health 2020;20(1):115 [FREE Full text] [doi: 10.1186/s12903-020-01100-y] [Medline: 32299417]
14. Cronin J, Moore S, Harding M, Whelton H, Woods N. A cost-effectiveness analysis of community water fluoridation for schoolchildren. BMC Oral Health 2021;21(1):158 [FREE Full text] [doi: 10.1186/s12903-021-01490-7] [Medline: 33765985]
15. Paemanukornruk Y, Luksamijarulkul N, Gaewkhiew P. Resin-based sealant effectiveness in high-caries risk children: a systematic review. BMC Oral Health 2025;25(1):768 [FREE Full text] [doi: 10.1186/s12903-025-06158-0] [Medline: 40410795]
16. Junger ML, Canto MT, Beltrán-Aguilar ED. Data Quality Evaluation of the Dental Fluorosis Clinical Assessment Data From the National Health and Nutrition Examination Survey, 1999-2004 and 2011-2016. Hyattsville, MD: National Center for Health Statistics; 2019.
17. Oral Health Data Systems: Child Indicators. Centers for Disease Control and Prevention. 2024 Jul 17. URL: https://www.cdc.gov/oral-health-data-systems/about_oral_health_data/child-indicators.html [accessed 2026-03-06]
18. Taylor CL, Madans JH, Chapman NN, Woteki CE, Briefel RR, Dwyer JT, et al. Critical data at the crossroads: the National Health and Nutrition Examination Survey faces growing challenges. Am J Clin Nutr 2023;117(5):847-858 [FREE Full text] [doi: 10.1016/j.ajcnut.2023.03.007] [Medline: 36907514]
19. Loiacono P, Pascoletti L. Photography in Dentistry: Theory and Techniques in Modern Documentation. Chicago, IL: Quintessence Publishing; 2012.
20. Sheridan P. Clinical Photography in Dentistry: A New Perspective. Chicago, IL: Quintessence Publishing; 2017.
21. Ortiz MA. LIT: The Simple Protocol for Dental Photography in the Age of Social Media. Chicago, IL: Quintessence Publishing; 2019.

XSL•FO
RenderX

22. Webb LR, Simms ML, Atkin PA. Inclusion of clinical photographs with general dental practitioner electronic referrals of patients with oral mucosal disease to secondary care: a service review. Br Dent J 2024 Jul 19. [doi: 10.1038/s41415-024-7608-0] [Medline: 39030367]
23. Moussa C, Hardan L, Kassis C, Bourgi R, Devoto W, Jorquera G, et al. Accuracy of dental photography: professional vs smartphone's camera. Biomed Res Int 2021;2021:3910291 [FREE Full text] [doi: 10.1155/2021/3910291] [Medline: 34957302]
24. Shahrul AI, Shukor N, Norman NH. Technique for orthodontic clinical photographs using a smartphone. Int J Dent 2022;2022:2811684 [FREE Full text] [doi: 10.1155/2022/2811684] [Medline: 35103062]
25. Teruya K, Park JH, Bay C. Dental photography using digital single-lens reflex cameras vs smartphones. AJO-DO Clinical Companion 2025;5(1):26-34. [doi: 10.1016/j.xaor.2024.11.004]
26. Hasbini O, Bourgi R, Kharouf N, Cuevas-Suárez CE, Kharma K, Moussa C, et al. Digital dentistry and imaging: comparing the performance of smartphone and professional cameras for clinical use. Prosthesis 2025;7(4):77 [FREE Full text] [doi: 10.3390/prosthesis7050109]
27. Nandhra K, Woolley J, Bister D, Sherriff M, Jeremiah H. Smartphones or digital SLRs for clinical dental photography: is there a difference? Eur J Orthod 2025;47(4):cjaf033. [doi: 10.1093/ejo/cjaf033] [Medline: 40501276]
28. Shahrul AI, Mohd Shukor N, Norman NH. Does the camera type affect the quality of orthodontic photographs? Mirrorless and smartphone cameras versus digital single lens reflex (DSLR) camera. J Orthod 2025;52(4):315-325 [FREE Full text] [doi: 10.1177/14653125251358837] [Medline: 40819226]
29. Ku JCK, Mao K, Wang F, Carreiro AFP, Lam WYH, Yu OY. Accuracy of clinical photography for the detection of dental caries: a systematic review and meta-analysis. J Dent 2025;157:105737 [FREE Full text] [doi: 10.1016/j.jdent.2025.105737] [Medline: 40204147]
30. Gurgel-Juarez N, Torres-Pereira C, Haddad AE, Sheehy L, Finestone H, Mallet K, et al. Accuracy and effectiveness of teledentistry: a systematic review of systematic reviews. Evid Based Dent 2022:1-8 [FREE Full text] [doi: 10.1038/s41432-022-0257-8] [Medline: 35804195]
31. Talla PK, Inquimbert C, Dawson A, Zidarov D, Bergeron F, Chandad F. Barriers and enablers to implementing teledentistry from the perspective of dental health care professionals: protocol for a systematic quantitative, qualitative, and mixed studies review. JMIR Res Protoc 2023;12:e44218 [FREE Full text] [doi: 10.2196/44218] [Medline: 37494093]
32. Tsai C, Savran A, Chau Y, Hurrell L, Forsyth C, Kumar H. Effectiveness of telehealth in providing preventive oral health care: a pilot evaluation of patient, carer, and clinician experience. J Public Health Dent 2023;83(4):331-339 [FREE Full text] [doi: 10.1111/jphd.12580] [Medline: 37675863]
33. Nascimento da Silva Mulder J, Ramos Pinto M, Aníbal I, Dornellas AP, Garrido D, Huanca C, et al. Teledentistry applied to health and education outcomes: evidence gap map. J Med Internet Res 2024;26:e60590 [FREE Full text] [doi: 10.2196/60590] [Medline: 39602783]
34. Alavi SE, Gholami M, Malik L, Matti R, Oktaei S, Al-Najafi F, et al. Assessment of teledentistry in improving access to dental care: a systematic review. Aust Dent J 2025;70(1):4-41 [FREE Full text] [doi: 10.1111/adj.13045] [Medline: 39440545]
35. Wang L, Ren J, Fiscella KA, Bullock S, Sanders MR, Loomis EL, et al. Interprofessional collaboration and smartphone use as promising strategies to improve prenatal oral health care utilization among US underserved women: results from a qualitative study. BMC Oral Health 2020;20:333 [FREE Full text] [doi: 10.1186/s12903-020-01327-9] [Medline: 33228617]
36. Xiao J, Fiscella KA, Meyerowitz C. mDentistry: a powerful tool to improve oral health of a broad population in the digital era. J Am Dent Assoc 2021;152(9):713-716 [FREE Full text] [doi: 10.1016/j.adaj.2021.06.004] [Medline: 34454644]
37. Kohara EK, Abdala CG, Novaes TF, Braga MM, Haddad AE, Mendes FM. Is it feasible to use smartphone images to perform telediagnosis of different stages of occlusal caries lesions? PLoS One 2018;13(9):e0202116 [FREE Full text] [doi: 10.1371/journal.pone.0202116] [Medline: 30188900]
38. Zhang Y, Liao H, Xiao J, Jallad NA, Ly-Mapes O, Luo J. A smartphone-based system for real-time early childhood caries diagnosis. In: Medical Ultrasound, and Preterm, Perinatal and Pediatric Image Analysis. Cham, Switzerland: Springer; 2020:233-242.
39. Duong DL, Kabir MH, Kuo RF. Automated caries detection with smartphone color photography using machine learning. Health Informatics J 2021;27(2):1-17 [FREE Full text] [doi: 10.1177/14604582211007530] [Medline: 33863251]
40. Pham DD, Müller J, Aggarwal P, Khatri A, Sharma M, Zesch T, et al. Fully vs. weakly supervised caries localization in smartphone images with CNNs. In: Pattern Recognition. ICPR International Workshops and Challenges. Cham, Switzerland: Springer; 2021:321-336.
41. Al-Jallad N, Ly-Mapes O, Hao P, Ruan J, Ramesh A, Luo J, et al. Artificial intelligence-powered smartphone application, AICaries, improves at-home dental caries screening in children: Moderated and unmoderated usability test. PLOS Digit Health 2022;1(6):e0000046 [FREE Full text] [doi: 10.1371/journal.pdig.0000046] [Medline: 36381137]
42. Xiong Y, Zhang H, Zhou S, Lu M, Huang J, Huang Q, et al. Simultaneous detection of dental caries and fissure sealant in intraoral photos by deep learning: a pilot study. BMC Oral Health 2024;24(1):553 [FREE Full text] [doi: 10.1186/s12903-024-04254-1] [Medline: 38735954]
43. Rathod R, Dean S, Sproat C. The effectiveness of a novel artificial intelligence (AI) model in detecting oral and dental diseases. BDJ Open 2025;11(1):62 [FREE Full text] [doi: 10.1038/s41405-025-00336-6] [Medline: 40588504]

XSL•FO
RenderX

44. Rozier RG. Epidemiologic indices for measuring the clinical manifestations of dental fluorosis: overview and critique. Adv Dent Res 1994;8(1):39-55. [doi: 10.1177/08959374940080010901] [Medline: 7993559]
45. Cruz-Orcutt N, Warren JJ, Broffitt B, Levy SM, Weber-Gasparoni K. Examiner reliability of fluorosis scoring: a comparison of photographic and clinical examination findings. J Public Health Dent 2012;72(2):172-175 [FREE Full text] [doi: 10.1111/j.1752-7325.2012.00315.x] [Medline: 22316120]
46. Boye U, Willasey A, Walsh T, Tickle M, Pretty IA. Comparison of an intra-oral photographic caries assessment with an established visual caries assessment method for use in dental epidemiological studies of children. Community Dent Oral Epidemiol 2013;41(6):526-533. [doi: 10.1111/cdoe.12049] [Medline: 23566100]
47. Tom K. Patient's experience, attitude, and satisfaction toward the use of a digital intraoral camera in a dental screening campaign. J Oral Health Community Dent 2020;14(2):44-47. [doi: 10.5005/jp-journals-10062-0071]
48. Ong JMD, Crasto GJ, Anwar EJ, Brooke R, Kang P. A standardized approach to extra-oral and intra-oral digital photography. J Vis Exp 2022(185):e63627 [FREE Full text] [doi: 10.3791/63627] [Medline: 35938814]
49. Kargozar S, Jadidfard MP. Teledentistry accuracy for caries diagnosis: a systematic review of in-vivo studies using extra-oral photography methods. BMC Oral Health 2024;24(1):828 [FREE Full text] [doi: 10.1186/s12903-024-04564-4] [Medline: 39039499]
50. Olvera G, Garrett K, A'see K. The role of intra-oral photography in dental patients' acceptance of treatment plans. Dent Rev 2024;4(3):100103 [FREE Full text] [doi: 10.1016/j.dentre.2024.100103]
51. Mania L, Nanobashvili K, Manjavidze T, Benashvili M, Astamadze I. Diagnostic accuracy of intraoral mobile photography for oral health screening in children: a pilot study. BMC Oral Health 2025;25(1):1144 [FREE Full text] [doi: 10.1186/s12903-025-06500-6] [Medline: 40646524]
52. Dean HT. Endemic fluorosis and its relation to dental caries. 1938. Public Health Rep 2006;121 Suppl 1:213-9; discussion 212. [Medline: 16550781]
53. Aly NM, El Kashlan MK, Giraudeau N, El Tantawi M. Comparison of intraoral cameras and smartphones in early childhood caries detection: a diagnostic accuracy study. J Evid Based Dent Pract 2024;24(4):102041 [FREE Full text] [doi: 10.1016/j.jebdp.2024.102041] [Medline: 39631974]
54. Park SM, Kwon S, Ji Y, Sakthivel H, Leem JW, Kwak Y, et al. Machine reading and recovery of colors for hemoglobin-related bioassays and bioimaging. Sci Adv 2025;11(23):eadt4831 [FREE Full text] [doi: 10.1126/sciadv.adt4831] [Medline: 40465724]
55. Kwon S, Park SM, Ji Y, Sakthivel H, Leem JW, Kim Y. Hyperspectral information extraction with full resolution from arbitrary photographs. IEEE Trans Image Process 2025;34:5429-5441. [doi: 10.1109/tip.2025.3597038]
56. Hamilton EK, Griffin SO, Espinoza L. Effect of CDC adjustment of state-reported data on community water fluoridation statistics. J Public Health Dent 2023;83(3):320-324. [doi: 10.1111/jphd.12579] [Medline: 37401874]
57. Lin M, Griffin SO, Gooch BF, Espinoza L, Wei L, Li CH, et al. Oral Health Surveillance Report: Trends in Dental Caries and Sealants, Tooth Retention, and Edentulism, United Sates: 1999-2004 to 2011-2016. Atlanta, GA: Centers for Disease Control and Prevention; 2019.
58. Beltrán-Aguilar ED, Barker L, Dye BA, Tinanoff N. Prevalence and Severity of Dental Fluorosis in the United States: 1999-2004. Hyattsville, MD: National Center for Health Statistics; 2010.
59. Hong SG, Park SM, Kwon S, Sakthivel H, Nagappa S, Leem JW, et al. Smartphone conjunctiva photography for malaria risk stratification in asymptomatic school age children. NPJ Digit Med 2025;8(1):151 [FREE Full text] [doi: 10.1038/s41746-025-01548-8] [Medline: 40065117]
60. Hong SG, Park SM, Kwon S, Sakthivel H, Leem JW, Steinhubl SR, et al. Radiomic identification of anemia features in monochromatic conjunctiva photographs in school-age children. Biophotonics Discov 2025 Apr;2(2):022303. [doi: 10.1117/1.bios.2.2.022303] [Medline: 40535314]
61. Dean HT. The Investigation of Physiological Effects by the Epidemiological Method. Washington, DC: American Association for the Advancement of Science; 1942.
62. Visbal Onufrak MA, Konger RL, Kim YL. Telecentric suppression of diffuse light in imaging of highly anisotropic scattering media. Opt Lett 2015;41(1):143-146 [FREE Full text] [doi: 10.1364/ol.41.000143]
63. Viallefont-Robinet F, Helder D, Fraisse R, Newbury A, van den Bergh F, Lee D, et al. Comparison of MTF measurements using edge method: towards reference data set. Opt. Express 2018;26(26):33625-33648 [FREE Full text] [doi: 10.1364/oe.26.033625]
64. Badano A, Revie C, Casertano A, Cheng WC, Green P, Kimpe T, et al. Consistency and standardization of color in medical imaging: a consensus report. J Digit Imaging 2015;28(1):41-52 [FREE Full text] [doi: 10.1007/s10278-014-9721-0] [Medline: 25005868]
65. Sakthivel H, Park SM, Kwon S, Kaguiri E, Nyaranga E, Leem JW, et al. Machine learning of blood haemoglobin and haematocrit levels via smartphone conjunctiva photography in Kenyan pregnant women: a clinical study protocol. BMJ Open 2025;15(5):e097342 [FREE Full text] [doi: 10.1136/bmjopen-2024-097342]
66. Dye BA, Afful J, Thornton-Evans G, Iafolla T. Overview and quality assurance for the oral health component of the National Health and Nutrition Examination Survey (NHANES), 2011-2014. BMC Oral Health 2019;19:95 [FREE Full text] [doi: 10.1186/s12903-019-0777-6] [Medline: 31142316]

XSL•FO
RenderX

67. Grandini M, Bagli E, Visani G. Metrics for multi-class classification: an overview. ArXiv. Preprint posted online on August 13, 2020 2026. [doi: 10.48550/arXiv.2008.05756]
68. Rainio O, Teuho J, Klén R. Evaluation metrics and statistical tests for machine learning. Sci Rep 2024;14(1):6086 [FREE Full text] [doi: 10.1038/s41598-024-56706-x] [Medline: 38480847]
69. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024;385:e078378. [doi: 10.1136/bmj.q902] [Medline: 38636956]
70. Klement W, El Emam K. Consolidated reporting guidelines for prognostic and diagnostic machine learning modeling studies: development and validation. J Med Internet Res 2023;25:e48763 [FREE Full text] [doi: 10.2196/48763] [Medline: 37651179]
71. Warren JJ, Levy SM. Measurement and distribution of dental fluorosis. In: Burt BA, Eklund SA, editors. Dentistry, Dental Practice, and the Community, 7th Ed. St. Louis, MO: Elsevier; 2021:218-226.
72. Dye BA, Li X, Beltrán-Aguilar ED. Selected oral health indicators in the United States: 2005-2008. Hyattsville, MD: National Center for Health Statistics; 2012.
73. Farook TH, Ahmed S, Jamayet NB, Dudley J. Computer vision with smartphone microphotography for detection of carious lesions. Intell Based Med 2023;8:100105 [FREE Full text] [doi: 10.1016/j.ibmed.2023.100105]
74. Hossain MS, Rahman MM, Syeed MM, Hannan UH, Uddin MF, Mumu SB. CaViT: early stage dental caries detection from smartphone-image using vision transformer. Presented at: 2023 4th International Conference on Artificial Intelligence, Robotics and Control (AIRC); May 9-11, 2023; Cairo, Egypt p. 9-14 URL: https://ieeexplore.ieee.org/document/10303012 [doi: 10.1109/airc57904.2023.10303012]
75. Pandey P, Jasrasaria N, Bains R, Singh A, Manar M, Kumar A. The efficacy of dental caries telediagnosis using smartphone: a diagnostic study in geriatric patients. Cureus 2023;15(1):e33256 [FREE Full text] [doi: 10.7759/cureus.33256] [Medline: 36741615]
76. Priyank H, Verma A, Zama Khan DU, Prakash Rai N, Kalburgi V, Singh S. Comparative evaluation of dental caries score between teledentistry examination and clinical examination: a systematic review and meta-analysis. Cureus 2023;15(7):e42414 [FREE Full text] [doi: 10.7759/cureus.42414] [Medline: 37637546]
77. Lamas-Lara VF, Mattos-Vela MA, Evaristo-Chiyong TA, Guerrero ME, Jiménez-Yano JF, Gómez-Meza DN. Validity and reliability of a smartphone-based photographic method for detection of dental caries in adults for use in teledentistry. Front Oral Health 2025;6:1470706 [FREE Full text] [doi: 10.3389/froh.2025.1470706] [Medline: 40443702]
78. Ding B, Zhang Z, Liang Y, Wang W, Hao S, Meng Z, et al. Detection of dental caries in oral photographs taken by mobile phones based on the YOLOv3 algorithm. Ann Transl Med 2021;9(21):1622 [FREE Full text] [doi: 10.21037/atm-21-4805] [Medline: 34926666]
79. AlShaya M, Farsi D, Farsi N, Farsi N. The accuracy of teledentistry in caries detection in children - a diagnostic study. Digit Health 2022;8:1-14 [FREE Full text] [doi: 10.1177/20552076221109075] [Medline: 35774249]
80. Adnan N, Faizan Ahmed SM, Das JK, Aijaz S, Sukhia RH, Hoodbhoy Z, et al. Developing an AI-based application for caries index detection on intraoral photographs. Sci Rep 2024;14(1):26752 [FREE Full text] [doi: 10.1038/s41598-024-78184-x] [Medline: 39500993]
81. Huang B, Estai M, Pungchanchaikul P, Quick K, Ranjitkar S, Fashingbauer E, et al. Mobile health assessment of traumatic dental injuries using smartphone-acquired photographs: a multicenter diagnostic accuracy study. Telemed J E Health 2024;30(10):2592-2600. [doi: 10.1089/tmj.2024.0012] [Medline: 38946689]
82. Moharrami M, Farmer J, Singhal S, Watson E, Glogauer M, Johnson AEW, et al. Detecting dental caries on oral photographs using artificial intelligence: a systematic review. Oral Dis 2024;30(4):1765-1783 [FREE Full text] [doi: 10.1111/odi.14659] [Medline: 37392423]
83. Boy AF, Akhyar A, Arif TY, Syahrial S. Development of an artificial intelligence model based on MobileNetV3 for early detection of dental caries using smartphone images: A preliminary study. Adv Sci Technol Res J 2025;19(4):109-116 [FREE Full text] [doi: 10.12913/22998624/200308]
84. Xiao J, Luo J, Ly-Mapes O, Wu TT, Dye T, Al Jallad NA, et al. Assessing a smartphone app (AICaries) that uses artificial intelligence to detect dental caries in children and provides interactive oral health education: protocol for a design and usability testing study. JMIR Res Protoc 2021;10(10):e32921 [FREE Full text] [doi: 10.2196/32921] [Medline: 34529582]
85. Liu Z, Goodwin M, Ellwood RP, Pretty IA, McGrady M. Automatic detection and classification of dental fluorosis in vivo using white light and fluorescence imaging. J Dent 2018 Jul;74 Suppl 1:S34-S41 [FREE Full text] [doi: 10.1016/j.jdent.2018.04.021] [Medline: 29929587]
86. Kashirtsev F, Tressel J, Simon JC, Fried D. High contrast imaging of dental fluorosis in the short wavelength infrared. J Biophotonics 2021;14(10):e202100145 [FREE Full text] [doi: 10.1002/jbio.202100145] [Medline: 34302430]

Abbreviations

CDC: Centers for Disease Control and Prevention
CNN: convolutional neural network

XSL•FO
RenderX

CREMLS: Consolidated Reporting Guidelines for Prognostic and Diagnostic Machine Learning Modeling Studies: Development and Validation
HIPAA: Health Insurance Portability and Accountability Act
mHealth: mobile health
NHANES: National Health and Nutrition Examination Survey
TRIPOD-AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis-Artificial Intelligence
USPHS: US Public Health Service

Edited by A Schwartz; submitted 11.Jan.2026; peer-reviewed by WYH Lam; comments to author 04.Feb.2026; accepted 06.Mar.2026; published 30.Mar.2026.

Please cite as:
Park SM, Kwon S, Hong SG, Ji Y, Nagappa SP, Leem JW, Lin M, Beltrán-Aguilar ED, Griffin SO, Kim YL
Mobile Imaging–Based Machine Learning for Dental Caries, Sealants, and Fluorosis: Protocol for a Cross-Sectional Model Development and Validation Study
JMIR Res Protoc 2026;15:e91239
URL: https://www.researchprotocols.org/2026/1/e91239
doi:10.2196/91239
PMID:

©Sang Mok Park, Semin Kwon, Shaun G Hong, Yuhyun Ji, Sreeram P Nagappa, Jung Woo Leem, Mei Lin, Eugenio D Beltrán-Aguilar, Susan O Griffin, Young L Kim. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Personalized, Home-Based, Multidisciplinary Outpatient Clinic for Managing Breathlessness in Chronic Obstructive Pulmonary Disease: Protocol for a Single-Arm, Mixed-Methods Cohort Study

Jonathan Shung-Yi Lee[1], BExPhys, MPT; Belinda Cochrane[1,2], BMedSci, MBBS (Hons), DM; Tracy A Smith[3,4], BSc, MBBS, PhD; Sheree M Smith[5], BNurs, MSocPlanDev, PhD; Soo Foo[1,2], MBBS (Hons), DM; Luisa Garcia[6], MNurs (Clinical Nursing and Teaching); Teresa Kemp[1], GradDipNurs (Acute Care)

[1]Department of Respiratory Medicine, Camden and Campbelltown Hospitals, Sydney, Australia

[2]School of Medicine, Western Sydney University, Sydney, Australia

[3]Westmead Clinical School, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[4]Department of Respiratory and Sleep Medicine, Westmead Hospital, Sydney, New South Wales, Australia

[5]School of Nursing and Midwifery, College of Health, Adelaide University, Adelaide, South Australia, Australia

[6]School of Nursing and Midwifery, Western Sydney University, Sydney, Australia

Corresponding Author:
Belinda Cochrane, BMedSci, MBBS (Hons), DM
Department of Respiratory Medicine
Camden and Campbelltown Hospitals
Campbelltown Hospital
Therry Road, Campbelltown
Sydney, 2560
Australia
Phone: 61 02 46344001
Email: Belindacochrane@bigpond.com

Abstract

Background: Breathlessness is a distressing symptom for people with chronic obstructive pulmonary disease (COPD), impacting independence and community participation, resulting in reduced quality of life. With optimized treatment, nonpharmaceutical interventions deployed in a home-based breathlessness intervention service may reduce breathlessness to a manageable level.

Objective: The aim of this study is to investigate whether patients with COPD experience reduced burden of chronic breathlessness and health care use following participation in the Macarthur Breathless Clinic (MBC), a multidisciplinary home-based health service intervention and to explore participant and carer experiences.

Methods: The proposed research is a single-site, mixed methods, single-arm cohort study evaluating MBC. Eligible participants will have chronic breathlessness as defined by a modified Medical Research Council score of at least 2 and COPD confirmed by spirometry. Initially, participants will undergo a holistic medical assessment to understand factors contributing to breathlessness and to optimize therapy. The MBC intervention will be a 9-week, home-based, multidisciplinary program (with contributions from medical, nursing, physiotherapy, and occupational therapy clinicians). Central to this program will be coaching for breathlessness self-management using individualized, multicomponent strategies. Breathlessness impact will be assessed using questionnaire instruments measuring breathlessness mastery, symptom burden, quality of life, and psychological distress, with the Chronic Respiratory Questionnaire's mastery subscale as primary outcome. Assessments will be undertaken at baseline, 9 weeks, and 52 weeks; 12-month exacerbation rates, health care use, and rescue medication use will be monitored and measured against participants' historical data. Participant and carer experiences will be evaluated using qualitative methods to inform development of the service.

Results: This project was funded in stages according to defined milestones, with the final milestone payment received in October 2025. Data collection commenced in March 2022. On submission of this manuscript, recruitment is complete with 92 participants having completed the protocol by March 2024. Data analysis is ongoing, and results are expected to be reported by early 2027.

XSL•FO
RenderX

Conclusions: This research protocol is the first to investigate long-term outcomes following a health service intervention targeting chronic breathlessness for patients with COPD. Positive results may see breathlessness strategies from MBC or other similar health service interventions incorporated into existing models of chronic COPD care.

Trial Registration: Australian and New Zealand Clinical Trial Registry ACTRN12620001330932; https://tinyurl.com/35fec8h4

International Registered Report Identifier (IRRID): DERR1-10.2196/85766

(JMIR Res Protoc 2026;15:e85766) doi:10.2196/85766

KEYWORDS

chronic obstructive pulmonary disease; COPD; breathlessness; multidisciplinary; chronic disease management; breathlessness intervention service

Introduction

Chronic obstructive pulmonary disease (COPD) is a multisystem disease, primarily affecting the lungs and characterized by incompletely reversible airflow obstruction [1]. COPD is common, affecting 10% of adults worldwide and 14% of Australians aged >40 years [2]. Despite treatment according to guidelines, people with COPD experience a range of symptoms. Breathlessness is often prominent among these and may contribute to physical and psychological distress, and reduced functional status and community participation [1]. Breathlessness is responsible for a high proportion of ambulance presentations, medical admissions, and health care costs [3,4]. However, patients with refractory breathlessness and their caregivers feel that there is a lack of awareness of the symptom burden among health care providers and insufficient provision to address this burden within health care services [2,5,6].

Breathlessness intervention services have been clearly shown effective in reducing the burden of breathlessness for populations with malignant and mixed etiology disease [7]. Existing breathlessness intervention service research primarily targets participants with malignant disease, in Europe, over a 4- to 6-week period, led by clinicians whose expertise is in palliative care [7]. An Australian protocol for a single-blind, randomized, parallel-group study investigating the impact of a breathlessness intervention service on people with COPD has been published; however, the study outcomes are not yet fully reported [8].

The primary aim of this study is to assess whether participants living with COPD will experience reduced breathlessness impact following participation in a bespoke 9-week intervention, the Macarthur Breathless Clinic (MBC), focused on nonpharmacological interventions such as breathing retraining, energy conservation, physical activity, and use of a handheld electric fan. A secondary aim is to monitor participants' exacerbations, health care use, and rescue medication use after the intervention for comparison with past patterns of disease activity. Furthermore, we will use qualitative methods to assess patient and caregiver experiences.

The hypothesis is that patients with COPD will experience (1) reduced impact of breathlessness (in terms of coping ability, symptom burden, quality of life, and psychological distress); (2) reduced COPD-related health care use; and (3) report a positive consumer experience following participation in the MBC.

Methods

Study Design

This research is a prospective, single-arm cohort study of MBC, a breathlessness intervention service, assessed using mixed methods. The intervention is designed to adapt to incorporate components found to work well and to discard or revise those found to be less effective, informed by the experience of conducting the program.

Quantitative outcomes will be evaluated with participants acting as their own controls, before and after the clinic intervention. A 12-month follow-up with sequential cross-sectional assessments will provide an opportunity to assess breathlessness impact, exacerbations, health care use, and durability of outcomes. Qualitative outcomes for patients and their carers, if applicable, will be sought via an interviewer-administered survey in the months following completion of the clinic intervention. A flow diagram overview of the study design is shown in Figure 1.

XSL•FO
RenderX

Figure 1. Study design. Allocated percentages are estimates only. ECOPD: exacerbation of chronic obstructive pulmonary disease; HCU: health care use.

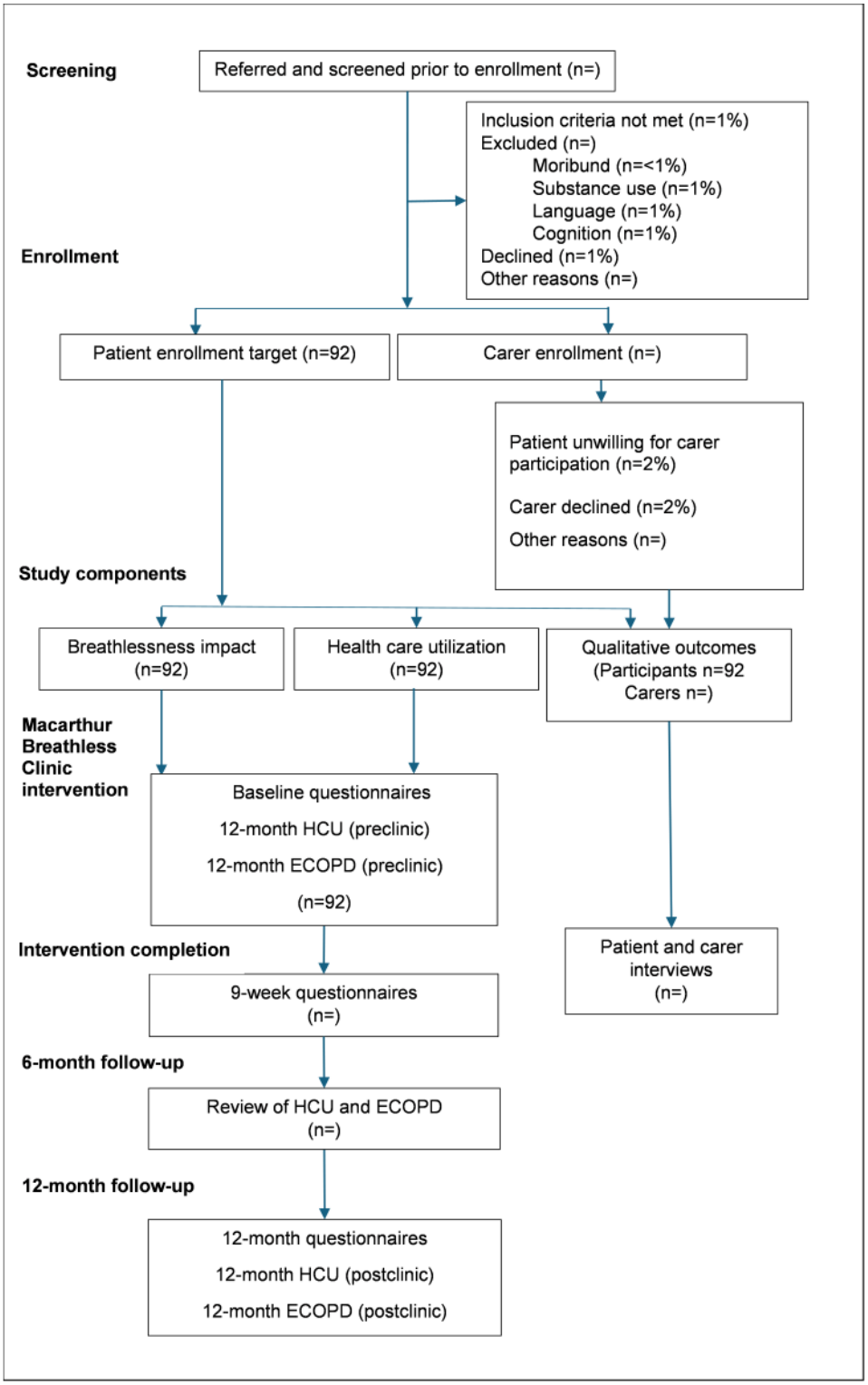

Ethical Considerations

This study protocol has been approved by the South Western Sydney Local Health District's Human Research Ethics Committee on December 1, 2020, and is listed on the Australian and New Zealand Clinical Trial Registry (ACTRN12620001330932). The research will be conducted in accordance with the principles of the World Medical Association Declaration of Helsinki, ensuring participant safety, data confidentiality, and adherence to all ethical guidelines. Written informed consent will be obtained from each participant after they have been fully advised of the study aims, methods, potential risk and benefits, their right to withdraw at any time without reprisal, and the measures to protect their privacy and confidentiality. Participation will be entirely voluntary, and no

XSL•FO
RenderX

monetary compensation, gifts, or incentives will be provided to participants in exchange for their time or information.

Setting

The study will be conducted predominantly within participants' homes across the Campbelltown, Camden, and Wollondilly local government areas in Sydney, Australia. This is an area on the fringe of metropolitan Sydney that accommodates a population with ethnic and linguistic diversity, social disadvantage, and a comparatively high proportion of Aboriginal Australians. In addition to home visits, a maximum of 2 scheduled health care interactions will occur in selected outpatient sites within the Macarthur Health Service.

Recruitment

Clinician referrals will be sought via outpatient services or after stabilization following hospital admission for COPD-related illness. Referrals will be screened by research personnel to ensure eligibility. Previous piloting of this mode of referral and recruitment has demonstrated feasibility for achieving the recruitment target over 12 months.

Where applicable and with patient agreement, carers will be invited to participate in the trial to contribute qualitative data to reflect on their support person role and perceptions of the service.

Sample Size

Recruitment target estimation is constrained by limited precedent data, most of which originate from research undertaken in the United Kingdom, studying participant populations, interventions, and outcomes that differ to ours [7,8]. However, pilot data from Westmead Hospital's Breathlessness Clinic for patients with COPD (similar target population to that proposed in this study) demonstrated statistically significant and clinically relevant outcome improvements in preliminary analysis of only 10 participants [9]. The currently proposed model has been informed by their prior experience. To allow for the substantial losses to follow-up commonly seen in clinical trials in moderate to severe patients with COPD, we have set our recruitment goal at 92 participants, in anticipation of 30% (n=31) discontinuation or dropout.

Data Collection Methods

All participant clinical and demographic data, as well as baseline and subsequent outcome data, will be collected by nonclinician MBC staff to minimize social acceptance bias. To ensure consistency in data evaluation, the same staff members will, where possible, collect outcome data for individual patients throughout the study.

A combination of outcomes has been chosen to comprehensively assess the intervention. Breathlessness impact will be measured using standardized questionnaires. These questionnaires will be administered in a pen-and-paper format in the home setting by interviewers without detailed knowledge of the intervention to further reduce bias. Participants will attempt questionnaire responses uninformed about their previous results but can be prompted with this information, if needed, in the event that the participant is otherwise unable to respond.

Health care use and exacerbation data will be gathered through structured interviews with participants. These interviews will be designed to capture detailed information about health care visits, hospitalizations, and other relevant health care interactions over the study period. The self-reported data from these interviews will be systematically cross-referenced with participants' health records to improve accuracy.

Qualitative data will be collected through in-depth interviews using a predefined semistructured interview script for both participants and carers. After obtaining permission, interviews will be recorded and transcribed verbatim for subsequent analysis. After transcription, recordings will be destroyed, and transcripts will be stored in the participant research files.

The purpose of surveying participant and carer experiences is partly a pragmatic attempt to inform development and adaptation of the clinical intervention to better suit end users and should also provide a platform for future research. The interviews will aim to capture participants' experiences and perceptions of the intervention, providing valuable insights about any impact on their daily lives. Similarly, for the carers, we hope to explore their perception of how the intervention affects their care recipient's coping ability and whether the caregiver notices any change in the burden of care.

The semistructured interview transcript can be found in Multimedia Appendix 1.

Inclusion and Exclusion Criteria

The proposed participant eligibility criteria are shown in Textbox 1.

XSL•FO
RenderX

Textbox 1. Participant eligibility criteria.

Inclusion criteria

- Clinical and spirometric diagnosis of moderate to severe chronic obstructive pulmonary disease (COPD) with forced expiratory volume in 1 second (FEV1) <80% predicted and FEV1/forced vital capacity ratio <70%
- Severe breathlessness with modified Medical Research Council score ≥2
- Aged ≥40 years

Exclusion criteria

- Bed-bound or moribund
- Cognitive impairment that precludes ability to engage and participate in education strategies
- Current active diagnosis of cancer, substance abuse (other than nicotine dependence), or other uncontrolled medical or psychiatric disorder
- Moderate or severe exacerbation of COPD (defined as an escalation of treatment to include systemic corticosteroids and/or antibiotics [10]) within 4 weeks
- Insufficient command of the English language to comprehend and engage with interventions
- Current enrollment in pulmonary rehabilitation program (may participate if still eligible upon program completion or if involved in a long-term maintenance program)

Description of Intervention: MBC

The proposed intervention, the MBC, is a 9-week program to manage chronic refractory breathlessness. The MBC team will include a respiratory physician, occupational therapist, specialist respiratory nurse, and physiotherapist with referrals as needed for dietetic and speech pathology services. The MBC itself will comprise 9 health care interactions of 60 to 90 minutes, with most visits occurring within the patient's home as outlined in Table 1.

Table 1. Macarthur Breathless Clinic 9-week program.

Schedule	Location	Activity
Week 1	Patient's home	• Initial assessment, respiratory nurse
Week 2	Outpatient clinic	• Medical assessment by respiratory physician • Multidisciplinary team assessment of breathlessness, contributing factors and impacts to develop recommendations and management plan • Introduction to nonpharmacologic strategies, including provision and use of a battery-operated handheld fan • Fact sheets on managing breathlessness
Week 3	Patient's home	• Physiotherapist
Week 4	Patient's home	• Occupational therapist
Week 5	Patient's home	• Respiratory nurse
Week 6	Patient's home	• Physiotherapist
Week 7	Patient's home	• Respiratory nurse
Week 8	Patient's home	• Final assessments, research assistant
Week 9	Outpatient clinic	• Closure and follow-up arrangements, multidisciplinary team

An initial holistic medical assessment will be performed by an experienced multidisciplinary team, including a respiratory physician, specialist respiratory nurse, and respiratory physiotherapist. The holistic assessment framework that will be used for breathlessness is systematic and encompasses the skill set of the multidisciplinary team ensuring the whole person is considered. The assessment template in Multimedia Appendix 2 outlines the domains considered. The assessment recognizes the multifactorial nature of breathlessness, with possible contributions from many interconnected parts, including medical history, comorbid diseases, medications, mental and social factors, understanding and beliefs about breathlessness, self-management strategies, inhaler technique, current functional capacity, exercise history, and even incontinence.

Following initial assessment and optimization of medications, nonpharmacological therapies will commence. Through the MBC, patients and carers will receive education about

XSL•FO
RenderX

breathlessness management and daily coping strategies to enhance independence and productivity. The clinical team will coach participants through individualized strategies to improve management of breathlessness in daily activities (Table 2) based on Cambridge Breathlessness Intervention Service's "breathing, thinking, functioning" clinical model [11]. A broad repertoire of standard interventions will be tailored to individual patients' needs, goals, and preferences.

Table 2. Nonpharmacologic breathlessness interventions.

Theme	Examples
Beliefs	• Validation of breathlessness • Clarifying mistaken beliefs—for example, breathlessness is not always due to low oxygen levels and breathlessness itself is not dangerous • Motivational interviewing
Strategies	• Use of a handheld battery-operated fan • Breathing techniques such as "breathing around a rectangle," "pursed lip breathing," "flickering the candle," "paced breathing," or "diaphragmatic breathing" explained with an accompanying fact sheet • Positions to relieve breathlessness • Sputum clearance techniques
Exercise	• Walking program with use of a pedometer • Strengthening exercises
Aids/efficiency	• Shower chair • 4-wheel walker • Energy conservation • Home modifications • Referral to dietitian if underweight or malnourished
Education and self-management	• Inhaler technique • Disease education (including medications and other therapies) • Written breathlessness action plans • Written chronic obstructive pulmonary disease exacerbation action plan • Breathlessness DVD to reinforce key interventions • Relaxation CD • Fact sheets on managing breathlessness
Miscellaneous	• Carer support • Assessment and management of incontinence

The role of each clinician and coaching topics covered is illustrated in Figure 2, noting that any cross-over of topics between clinicians will serve as reinforcement of strategies provided or will allow colleague cross-cover.

XSL•FO
RenderX

Figure 2. Clinician roles and coaching topics. COPD: chronic obstructive pulmonary disease; MDT: multidisciplinary team.



There will be fortnightly multidisciplinary team case conferencing of all participants during their 9-week intervention. Case conference discussion will aim to review and/or revise patient goals and optimize, coordinate, troubleshoot, and plan intervention activities. At the patient's final outpatient clinic review, the multidisciplinary team will revisit patient goals, discuss next steps, and seek informal feedback on the program.

Outcome Measures

Breathlessness Outcomes

The proposed primary outcome is mastery of breathlessness, as measured by the Chronic Respiratory Questionnaire (CRQ) subscale of mastery [12].

The proposed secondary outcomes, described below, will measure the impact of the intervention using several scales to

XSL•FO
RenderX

provide a full understanding of change across a number of dimensions.

The CRQ is a validated measure of quality of life used in chronic lung diseases, with a particular emphasis on breathlessness. It consists of 4 subscales: dyspnoea, fatigue, emotion, and mastery. In each scale, higher scores reflect better health. The minimal clinically important difference (MCID) for each subscale individually is an increase of 0.5 on a 7-point scale, with a change of 1 representing a moderate change and 1.5 a large change [12].

For breathlessness intensity, unpleasantness, and confidence numerical rating scales, international guidelines recommend assessing both intensity and unpleasantness of breathlessness [13]. The MCID for these measures is a change of 1 unit [14].

The COPD Assessment Test (CAT) is a validated 8 item tool that provides a reliable measure of the impact of COPD on a patient's health status. Higher scores reflect worse health with an MCID of 2 [15].

The Hospital Anxiety Depression Scale is a 14 item self-reported measure with anxiety and depression subscales each with 7 items. An MCID of 1.5 points for Hospital Anxiety Depression Scale in COPD populations has been suggested [16].

The EQ-5D-5L is a standardized measure of health status that provides a simple, generic measure of health for clinical and economic appraisal. The scale explores the degree of difficulty across 5 domains: mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. A weighted index value will be derived, ranging between −0.59 and 1, which can be converted to quality-adjusted life years based on the Australian values set. Higher scores reflect better health, and the mean MCID is 0.051. It also includes a visual analog scale 0 to 100, rating overall health from worst to best with a mean minimum important difference of 6.9 (range 6.5-8.0) [17].

Pedometer readings of average number of steps per day over the past 7 days will be recorded. There is no consensus regarding the MCID for this measure.

The Dalhousie Dyspnoea and Perceived Exertion Scale tracks dyspnea and perceived exertion. The scale consists of 7 pictures representing breathing effort, chest tightness, throat closure, and leg effort, with increasing severity depicted in the pictures, traversing from left to right [18].

Health Care Resource Use Outcomes

Self-reported exacerbation, emergency department (ED) visit, and hospitalization data will be collected. South Western Sydney Local Health District's electronic medical records will also be systematically reviewed to verify participant self-reports. Data will be updated at 6 and 12 months from MBC enrollment and will include the annual number of the following:

- ED presentations (all presentations), lasting <24 hours, that do not result in hospitalization, as well as breathlessness-related ED presentations, lasting <24 hours, that do not result in hospitalization
- Unplanned hospitalizations, including unplanned hospitalizations due to breathlessness
- Hospital length of stay in bed days (all unplanned hospitalizations), as well as breathlessness-related hospital bed days
- Exacerbations of COPD counted by patient-reported activation of COPD action plan, along with occasions of prescription of antibiotics and/or corticosteroid for the treatment of symptoms of respiratory exacerbation during hospitalized events

Statistical Analysis Plan

Descriptive statistics will be applied to the study population to characterize them in terms of demographic and baseline clinical parameters. Normally distributed continuous variables will be expressed as mean (SD), nonparametric variables will be defined by median and range, while categorical variables will be expressed as percentages.

Breathlessness outcomes will be compared across 3 time points: baseline, 9 weeks, and 52 weeks using repeated measures ANOVA. Predictors of response will be sought using linear regression models with sex, forced expiratory volume (FEV) in 1 second, ratio of FEV to forced vital capacity, BMI, Montreal Cognitive Assessment, and baseline anxiety and depression scores as prespecified variables. Health care use measures and exacerbation data for the 12 months prior to MBC enrollment will be compared with the corresponding values for the 12 months after MBC enrollment, using Student 2-tailed *t* tests or nonparametric tests as appropriate. Protocol completer and dropout populations will be compared in terms of baseline characteristics (sex, age, FEV in 1 second, pack years, BMI, CRQ-Mastery subscale, CAT score, and Montreal Cognitive Assessment) to investigate potential selection bias.

Finally, comparison will be made with patients' baseline data from the year prior to enrollment in the breathless clinic, benchmarked against nonparticipant comparator hospital deidentified patient medical records with (1) ED presentations for reasons of breathlessness, triaged as "shortness of breath—respiratory" or (2) hospitalized with exacerbations of COPD or heart failure as per diagnostic related group assignment. These benchmarking comparisons will use deidentified electronic medical records and will contribute to exploratory analyses undertaken to better understand the health care use of our sample cohort in the context of the overall health care use patterns of our local population "breathless" groups.

Qualitative Outcomes

Patients and carers (where relevant) will undertake semistructured, audio-recorded interviews up to 9 weeks after completion of MBC. These interviews will explore participant perceptions of the MBC, any particularly helpful components, and seek recommendations for future versions of the service (Multimedia Appendix 1). To identify key concepts and themes, thematic analysis will be conducted [19]. Interviews will be coded by 2 researchers, one of whom has been involved in the conduct of the MBC and a researcher from the Westmead Breathlessness Service, to provide comprehensive assessment of the data. Input from additional researchers, experienced in qualitative methods, will provide triangulation of the results to enhance trustworthiness of the analysis. The qualitative data

XSL•FO
RenderX

will complement the quantitative findings and offer a more complete understanding of the intervention's effects.

Results

Data collection commenced in March 2022, and recruitment is now complete, with 92 participants finishing the protocol by the end of March 2024. Data analysis is ongoing, and results are anticipated in early 2027. This project has been funded incrementally according to achievement of predefined milestones. The final milestone payment was received in October 2025.

Discussion

Anticipated Findings

This study adds to the literature by evaluating the change from baseline after a 9-week multidisciplinary, home-based health intervention to address the symptom of breathlessness in a cohort of Australian patients who have moderate to severe COPD. The study based on this protocol will provide important insights into applying this type of intervention to patients with COPD, a chronic disease. The authors anticipate that this study will help to inform models of care for patients with COPD enduring breathlessness as a persistent and distressing symptom, providing information about the potential durability of these interventions, and the characteristics of patients with COPD who might be most likely to benefit.

Comparison With Prior Work

Breathlessness intervention services have been shown to be effective for breathlessness over the short term [7] The first examples of these services targeted patients with malignant disease, using complex, multicomponent, nonpharmacological interventions for the management of breathlessness [20]. Subsequent trials enrolled heterogeneous populations of patients with a mixture of malignant and nonmalignant diseases, including some patients with COPD. The authors are aware of a recently completed randomized controlled trial in Australia targeting a population of patients with COPD using a breathlessness intervention model similar to our own. Breathlessness in COPD often remains persistent and distressing despite best pharmacological treatments. However, COPD is a chronic disease, and so patients may live for many years with this burden. For these patients, not only are effective treatment strategies needed but benefit must be sustained over the long term. The proposed protocol uses a multidisciplinary home-based complex health service intervention similar in type to those known to confer significant short-term benefits for breathless patients in mixed disease populations, aiming to confirm its applicability to breathless patients with COPD, and also exploring the durability of benefits and best selection of patients.

Strengths and Limitations

The study's focus on a homogeneous disease population of patients with moderate to severe COPD will ensure that the results are highly relevant to this specific group. This targeted approach will allow for more precise conclusions regarding the usefulness in this cohort of the intervention being tested. That said, we will recruit a wide range of patients with COPD in terms of severity and exacerbation susceptibility, ensuring broad applicability to this important, and large, patient group. Thus, our results will complement the existing evidence base, as most previously published data include more heterogeneous participant populations; patients with malignant disease, and/or patients with a variety of restrictive and obstructive nonmalignant respiratory disease.

To our knowledge this is the first long-term study protocol for this type of intervention service. By spanning a longer period, an entire year, the study will mitigate the potential confounding effects of seasonal variations in COPD symptoms and exacerbations. This will ensure that the observed outcomes are more likely attributable to the intervention rather than external seasonal factors. Additionally, the proposed intervention requires behavior change, which can take time to implement. By having a longer duration than previous interventions, we hope to enable patients to embed into their routine the suggested nonpharmacological strategies and to thus realize a durable effect beyond the 9 weeks of the program.

The 12-month follow-up period will also provide insight into the durability of the intervention response over a prolonged period. To date, the scientific literature contains few reports of outcomes for this type of intervention extending beyond 2 or 3 months. Showing sustained symptomatic benefits for this type of population with nonmalignant, chronic disease is perhaps much more important than demonstrating a short-term response.

Finally, the use of the CAT as a key assessment tool will underscore the study's relevance to clinical practice. The CAT is a validated, disease-specific instrument that provides a comprehensive evaluation of COPD symptoms and their impact on patients' quality of life. This may allow for the MBC to be ranked against other treatment interventions used in COPD, so that any positive study findings can be integrated into existing treatment paradigms, thereby facilitating the translation of research into practice.

However, this study has several limitations that must be acknowledged. First and foremost is the nonrandomized study design; all recruited participants will receive the intervention and act as their own controls. This design could introduce both time and attrition bias and is a clear limitation on the level of evidence produced. The authors were aware of a randomized controlled trial [8] underway in another Australian center using a similar model to our protocol, and we therefore chose a nonrandomized study design to address clinical need while still demonstrating translatability to a different population and setting. Given that COPD is a chronic and progressive disease, evaluating the impact of breathlessness over time may still yield valuable insights, provided the results are interpreted carefully. While potential mortality among participants with more severe disease may skew study results, on the other hand, worsening disease trajectory may mask benefits when determining results over the longer term. Considering this, even if patients appear to be stable with respect to outcome measures at 12 months, this may actually represent improvement compared with the expected trajectory in COPD, which is one of deterioration. In

XSL•FO
RenderX

fact, clinically significant breathlessness is a known marker of poor prognosis for patients with COPD [21], and so our target population will be expected to deteriorate more rapidly than the COPD group as a whole.

The study's proposed setting, a single metropolitan center situated within a large semirural area, may limit the general applicability of our results. However, the authors are aware of concurrent research on this topic being conducted at other centers and health districts throughout Australia, which should help to address this.

Our study protocol excludes patients based on English language proficiency, which represents another important limitation. Ideally language should not prevent equitable access to medical care nor opportunities to participate in health research. The decision to include only English-speaking participants was made due to significant ethnic and cultural differences in the expression and description of breathlessness. Additionally, some of the outcome measures are only validated in English. The exclusion of people who do not speak English will limit the generalizability of our results to the English-speaking population. The authors estimate the impact of exclusion due to English proficiency to be small as shown in Figure 1, noting that exclusion percentages are estimates based on census data. The 2021 census data from the Australian Bureau of Statistics inform that the Macarthur and Wollondilly local government areas have, when compared with Sydney and New South Wales, a higher concentration of English-only speakers [22]. To mitigate the impact on generalizability and perhaps to provide an alternative means to characterize and express the severity of breathlessness in future iterations of the program, the Dalhousie pictorial scales, which help bypass literacy and language barriers, have been included as an outcome measure.

Finally, regarding COPD exacerbations, the authors recognize that patient recall of these events is notoriously poor, particularly for milder events, either self-treated or treated via primary care in the community. Moreover, Australia's national guidelines, COPD-X, recommend that such patients have an exacerbation plan, along with prescriptions for relevant rescue medications making it even more complicated to establish the timing of these events. We will seek to confirm more severe events, resulting in an ED visit or hospitalization, systematically using local hospital electronic medical records. However, patient reports of exacerbation and health care use events will be sought nonetheless, given the possibility of capturing additional events occurring outside our health district, interstate, or within primary care, while acknowledging that the reporting and therefore capture of these events will be less reliable.

Future Directions

Breathlessness intervention services are now established as effective models to reduce the burden of breathlessness in cohorts with both malignant and nonmalignant disease. However, questions remain about which components are most important (clinician expertise or strategy type), durability of benefits, patient selection, and the most cost-effective models. While we have preliminary answers, future research should aim to fine-tune these interventions to provide the best breathlessness solutions for both patients and health systems.

Conclusions

The proposed study will be an important addition to the growing body of evidence about the use of nonpharmacological interventions for chronic refractory breathlessness in patients with COPD. Positive results would support incorporation of strategies from breathlessness intervention services, such as MBC into chronic disease models for COPD care, with potential to lighten the burden of chronic breathlessness for patients with COPD.

Funding

Funding for this study was provided by GSK plc (NCT11092). GSK was provided the opportunity to provide a courtesy review of the preliminary version of this publication for accuracy only, but the authors are solely responsible for final content and interpretation.

Authors' Contributions

All listed authors meet the criteria for authorship per the International Committee of Journal Editors Guidelines. BC and TAS contributed to the study design. JS-YL, TK, BC, and TAS drafted the paper. SMS, SF, and LG made substantial contributions to the critical review of the paper. All authors read and approved the submitted version of the paper.

Conflicts of Interest

BC received an investigator-sponsored study grant from GSK plc to conduct the research for this manuscript; personal adviser or consultancy fees from GSK and Sanofi; and speaker fees from Chiesi, Moderna, Astra Zeneca, Inside Practice, and RX Global and is a member of Lung Foundation Australia's COPD Coordinating Committee, none of which represent a direct conflict of interest (COI) with the work presented here. This is relevant to the current manuscript but does not represent a COI. SMS declares membership of the TSANZ Conference Committee but has no specific conflicts related to this work. SF has received speaker fees from Astra Zeneca and travel and accommodation support from GSK, Chiesi, Arizona, and Boehringer Ingelheim, none of which represent a direct COI with the work presented here. TK declares membership of the Respiratory Nurses Special Interest Group but has no specific conflicts related to this work. JS-YL, TAS, and LG have no potential COIs to declare.

XSL•FO
RenderX

Multimedia Appendix 1
Qualitative outcomes—semistructured interview questions.
[DOCX File , 35 KB - resprot_v15i1e85766_app1.docx]

Multimedia Appendix 2
Initial assessment template—Macarthur Breathless Clinic.
[DOCX File , 36 KB - resprot_v15i1e85766_app2.docx]

References

1. Yang I, Hancock K, George J, McNamara R, McDonald C, McDonald V. COPD-X Handbook: Summary Clinical Practice Guidelines for the Management of Chronic Obstructive Pulmonary Disease (COPD). Milton, Queensland: Lung Foundation Australia; 2014. URL: https://lungfoundation.com.au/resources/copd-x-handbook/ [accessed 2026-03-20]
2. Toelle BG, Xuan W, Bird TE, Abramson MJ, Atkinson DN, Burton DL, et al. Respiratory symptoms and illness in older Australians: the Burden of Obstructive Lung Disease (BOLD) study. Med J Aust 2013 Feb 18;198(3):144-148. [doi: 10.5694/mja11.11640] [Medline: 23418694]
3. Australian Burden of Disease Study. Australian Institute of Health and Welfare. 2024. URL: https://www.aihw.gov.au/reports/burden-of-disease/australian-burden-of-disease-study-2024/contents/summary [accessed 2026-03-11]
4. Hutchinson A, Pickering A, Williams P, Bland JM, Johnson MJ. Breathlessness and presentation to the emergency department: a survey and clinical record review. BMC Pulm Med 2017 Mar 20;17(1):53 [FREE Full text] [doi: 10.1186/s12890-017-0396-4] [Medline: 28320369]
5. Schunk M, Schulze F, Bausewein C. What constitutes good health care for patients with breathlessness? Perspectives of patients, caregivers, and health care professionals. J Palliat Med 2019 Jun;22(6):656-662. [doi: 10.1089/jpm.2018.0319] [Medline: 30589622]
6. Johnson MJ, Yorke J, Hansen-Flaschen J, Lansing R, Ekström M, Similowski T, et al. Towards an expert consensus to delineate a clinical syndrome of chronic breathlessness. Eur Respir J 2017 May 25;49(5):1602277 [FREE Full text] [doi: 10.1183/13993003.02277-2016] [Medline: 28546269]
7. Brighton LJ, Miller S, Farquhar M, Booth S, Yi D, Gao W, et al. Holistic services for people with advanced disease and chronic breathlessness: a systematic review and meta-analysis. Thorax 2019 Mar 29;74(3):270-281 [FREE Full text] [doi: 10.1136/thoraxjnl-2018-211589] [Medline: 30498004]
8. Smith TA, Roberts MM, Cho J, Klimkeit E, Luckett T, McCaffrey N, et al. Protocol for a single-blind, randomized, parallel-group study of a nonpharmacological integrated care intervention to reduce the impact of breathlessness in patients with chronic obstructive pulmonary disease. Palliat Med Rep 2020 Dec 01;1(1):296-306 [FREE Full text] [doi: 10.1089/pmr.2020.0081] [Medline: 34223489]
9. Smith T, Roberts M, Luckett T, Cho JG, Kirby A, Wheatley J. A multidisciplinary, non-pharmacologic intervention improves breathlessness in COPD: an RCT. ERS Publications 2024;64(Suppl(68)):OA5542 [FREE Full text] [doi: 10.1183/13993003.congress-2024.OA5542]
10. Disease G. GOLD report. Global Initiative for Chronic Obstructive Lung Disease. URL: https://goldcopd.org/ [accessed 2026-03-11]
11. Spathis A, Burkin J, Moffat C, Tuffnell R, Barclay S, Mant J, et al. Cutting through complexity: the breathing, thinking, functioning clinical model is an educational tool that facilitates chronic breathlessness management. NPJ Prim Care Respir Med 2021 May 10;31(1):25 [FREE Full text] [doi: 10.1038/s41533-021-00237-9] [Medline: 33972569]
12. Guyatt G. Chronic Respiratory Diseases Questionnaire. American Thoracic Society. 2002 Jul 2. URL: http://qol.thoracic.org/sections/instruments/ae/pages/crq.html [accessed 2026-03-11]
13. Mahler DA, Selecky PA, Harrod CG, Benditt JO, Carrieri-Kohlman V, Curtis JR, et al. American College of Chest Physicians consensus statement on the management of dyspnea in patients with advanced lung or heart disease. Chest 2010 Mar;137(3):674-691. [doi: 10.1378/chest.09-1543] [Medline: 20202949]
14. Ekström M, Currow D, Johnson M. Outcome measurement of refractory breathlessness endpoints and important differences. Current opinion in supportive and palliative care 2015;9(3):e [FREE Full text] [doi: 10.1097/spc.0000000000000147]
15. Jones I, Kirby A, Ormiston P, Loomba Y, Chan KK, Rout J, et al. The needs of patients dying of chronic obstructive pulmonary disease in the community. Fam Pract 2004 Jun 01;21(3):310-313. [doi: 10.1093/fampra/cmh317] [Medline: 15128695]
16. Lovibond S, Lovibond P. Manual for the Depression Anxiety Stress Scales, Second Edition. Sydney, Australia: Psychology Foundation; 1995.
17. Nolan CM, Longworth L, Lord J, Canavan JL, Jones SE, Kon SSC, et al. The EQ-5D-5L health status questionnaire in COPD: validity, responsiveness and minimum important difference. Thorax 2016 Jun 30;71(6):493-500 [FREE Full text] [doi: 10.1136/thoraxjnl-2015-207782] [Medline: 27030578]

XSL•FO
RenderX

18. Pianosi PT, Zhang Z, Hernandez P, Huebner M. Measuring dyspnea and perceived exertion in healthy adults and with respiratory disease: New Pictorial Scales. Sports Med Open 2016 Jan 7;2(1):17 [FREE Full text] [doi: 10.1186/s40798-015-0038-4] [Medline: 26770885]
19. Braun V, Clarke V. Is thematic analysis used well in health psychology? A critical review of published research, with recommendations for quality practice and reporting. Health Psychol Rev 2023 Dec 19;17(4):695-718. [doi: 10.1080/17437199.2022.2161594] [Medline: 36656762]
20. Higginson IJ, Bausewein C, Reilly CC, Gao W, Gysels M, Dzingina M, et al. An integrated palliative and respiratory care service for patients with advanced disease and refractory breathlessness: a randomised controlled trial. The Lancet Respiratory Medicine 2014 Dec;2(12):979-987. [doi: 10.1016/s2213-2600(14)70226-7]
21. Nishimura K, Izumi T, Tsukino M, Oga T. Dyspnea is a better predictor of 5-year survival than airway obstruction in patients with COPD. Chest 2002 May;121(5):1434-1440. [doi: 10.1378/chest.121.5.1434] [Medline: 12006425]
22. idprofile. Western Sydney (LGA) Community profile. Who Are We? Language Used at Home. URL: https://profile.id.com.au/cws/language [accessed 2026-03-11]

Abbreviations

CAT: COPD Assessment Test
COPD: chronic obstructive pulmonary disease
CRQ: Chronic Respiratory Questionnaire
ED: emergency department
FEV: forced expiratory volume
MBC: Macarthur Breathless Clinic
MCID: minimal clinically important difference

Edited by J Sarvestan; submitted 19.Oct.2025; peer-reviewed by A Cund, K Heslop-Marshall, L Pirhonen; comments to author 08.Jan.2026; accepted 04.Mar.2026; published 02.Apr.2026.

Please cite as:
Lee JSY, Cochrane B, Smith TA, Smith SM, Foo S, Garcia L, Kemp T
A Personalized, Home-Based, Multidisciplinary Outpatient Clinic for Managing Breathlessness in Chronic Obstructive Pulmonary Disease: Protocol for a Single-Arm, Mixed-Methods Cohort Study
JMIR Res Protoc 2026;15:e85766
URL: https://www.researchprotocols.org/2026/1/e85766
doi:10.2196/85766
PMID:

©Jonathan Shung-Yi Lee, Belinda Cochrane, Tracy A Smith, Sheree M Smith, Soo Foo, Luisa Garcia, Teresa Kemp. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 02.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Pediatric Emergency Agitation Care Enhancement: Protocol for a Prospective Mixed Methods Study

Lillian Klein[1], BS; Wendy Pomerantz[1], MS, MD; Teresa Pestian[1], MS, PhD; Yin Zhang[1], MS; Drew Barzman[1], MD; Lisa Vaughn[1], PhD; Nancy Daraiseh[1], PhD; Ryan Siders[1], BSN, RN; Holly Hanson[1], MS, MD; Bijan Ketabchi[2], MPH, MD; Lynn Babcock[1], MS, MD

[1]Emergency Medicine, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH, United States

[2]Children's Hospital of Philadelphia, Philadelphia, PA, United States

Corresponding Author:
Lillian Klein, BS
Emergency Medicine, Cincinnati Children's Hospital Medical Center, 3333 Burnet Ave, Cincinnati, OH, United States

Abstract

Background: Children and adolescents presenting to emergency departments (EDs) with mental and behavioral health (MBH) concerns frequently exhibit agitation, which poses safety risks for patients and staff and strains ED resources. Validated tools for agitation risk stratification in pediatric MBH populations are lacking, and evidence-based, risk-informed management strategies remain underdeveloped. Addressing these gaps is critical to reducing the need for emergent interventions, enhancing safety, and optimizing care delivery.

Objective: This study aims to (1) evaluate the predictive ability of a brief pediatric agitation risk prediction tool, a 5-item version of the Brief Rating of Aggression by Children and Adolescents (BRACHA-S), and (2) codevelop an associated risk-based management plan bundle.

Methods: This is a single-center prospective mixed methods study. First, we conducted a prospective cohort study enrolling children and adolescents aged 5 to 18 years who presented to the ED with MBH concerns. At ED triage, nurses completed the BRACHA-S risk assessment tool. The primary outcome, agitation requiring intervention, was defined as the use of pharmacologic agents, physical restraints, or mechanical restraints for events of agitation or aggression, as extracted from the electronic health record and confirmed through chart review. Second, we will use qualitative participatory methodologies, including key informant interviews, group-level assessments, and participatory design workshops, to codevelop a multifaceted, tiered, risk-based management plan designed to mitigate agitation and aggression in pediatric ED patients. Insights will be gathered from diverse stakeholders, including patients, parents or guardians, and members of the care team.

Results: Funding was secured in July 2024. Initial data collection commenced in October 2024 and is projected to conclude in 2026. For aim 1, enrollment was completed on November 1, 2025, achieving the targeted sample size of 472 participants. We hypothesize that the BRACHA-S will demonstrate strong predictive validity (area under the curve >0.70) for agitation requiring intervention. Data collection for aim 2 began in October 2025, with analysis and results anticipated by June 2026. As of April 2026, 10 patients and parents/guardians were enrolled for key informant interviews, 30 staff members were enrolled for group-level assessment sessions, and 9 participants were enrolled for participatory design workshops. For this aim, we will codevelop a stakeholder-informed, tiered, risk-based agitation management pathway aligned with BRACHA-S risk strata.

Conclusions: This study outlines a pragmatic approach to early agitation risk identification and the management of agitation in pediatric ED settings. If BRACHA-S demonstrates predictive validity and is paired with a stakeholder-informed, tiered management pathway, this strategy could fill critical gaps in validated tools and structured workflows, ultimately enabling timely, least-restrictive interventions and improving patient and staff safety.

International Registered Report Identifier (IRRID): DERR1-10.2196/92452

(JMIR Res Protoc 2026;15:e92452) doi:10.2196/92452

KEYWORDS

emergency service; hospital; pediatrics; mental disorders; behavioral symptoms; agitation; risk assessment; aggression; restraint; physical; qualitative research; mixed methods

XSL•FO
RenderX

Introduction

The emergency department (ED) plays a pivotal role in screening children with acute mental and behavioral health (MBH) concerns. In the United States, over the past decade, ED encounters for pediatric MBH concerns have increased by 60%, reaching 13% of total visits in 2020 [1-3]. In a subset of these encounters, up to 10% involve agitation or aggression requiring intervention (ARI), inclusive of the use of pharmacologic, physical, or mechanical restraints [4-8]. Agitation poses safety risks to patients and staff and strains ED resources because agitated patients often require staff for 1-on-1 observation and designated psychiatric-safe rooms. Predicting patient agitation is essential to optimize care; however, validated agitation risk-stratification tools for children and adolescents cared for in the ED are lacking [9].

The Brief Rating of Aggression by Children and Adolescents (BRACHA) is a 14-item assessment conducted during ED mental health evaluations and is highly predictive of aggressive behaviors during inpatient psychiatric admission [10,11]. Typically completed by trained MBH specialists 1 to 2 hours after arrival at the ED or before admission, the BRACHA's delayed administration, lengthy nature, and requirement for specialist completion render it ineffective as an early triage tool for predicting aggression in the ED. In a 9-year retrospective cohort study involving 32,091 children and adolescents presenting to the ED with MBH concerns, a shortened 5-item version of BRACHA (BRACHA-S; Table 1) demonstrated strong discriminative power (area under the curve 0.80) for predicting ARI in the ED [8]. This preliminary work suggests that a brief risk prediction tool may support earlier identification of children at risk for agitation and aggression.

Table . Items of the 5-item Brief Rating of Aggression by Children and Adolescents and scoring scheme.

Item description	Score (no=0 or yes=1)
History of psychiatric hospitalizations	No or yes
Frequency of physically aggressive acts toward others	Never or often
Threats or physical aggression toward self or others in the past 24 hours	No or yes
Impulsivity or agitation during evaluation	No or yes
Intrusiveness toward others during evaluation	No or yes

In addition to the BRACHA, other brief observational tools have been proposed to identify imminent aggression in various care settings [12]. The Brøset Violence Checklist (BVC), created for psychiatric inpatient units, predicts violent incidents within 24 hours by assessing 6 behaviors associated with escalating agitation (confusion, irritability, boisterousness, verbal threats, physical threats, and attacking objects) [13-16]. The BVC has been used in more than 62 studies across various settings, primarily in adult and inpatient psychiatric populations [17]. Another tool, the STAMP framework, developed from qualitative work with emergency nurses, focuses on 5 observable behaviors that may signal escalating risk: staring, changes in tone or loudness of voice, visible anxiety, mumbling, and pacing [18]. To date, STAMP has not been systematically evaluated in ED patient populations, and its predictive properties are unknown. Conceptually, both STAMP and the BVC can be applied at triage and repeated throughout an ED visit as brief, momentary assessments of agitation and violence risk. However, neither tool has been prospectively validated in pediatric ED settings for predicting the need for agitation-related interventions, underscoring the need for rigorously tested, evidence-based violence and agitation risk assessment tools for pediatric emergency care.

Current guidelines for pediatric agitation management emphasize early, least-restrictive care, including verbal de-escalation and environmental modification, to reduce the need for emergent pharmacologic or physical interventions while maintaining safety [9,17,19-23]. As recommended by the Joint Commission, patients in the ED with MBH concerns are routinely screened for suicide risk using brief, validated tools at triage [24-27] to stratify the risk of self-harm and initiate tiered responses that enhance safety. However, these approaches fail to identify individuals who may pose a risk of aggression toward others or property [28]. Collectively, these findings highlight the need for a stakeholder-informed, risk-based approach to agitation and aggression, guided by a validated risk assessment at triage and linked to a tiered ED management plan, which could support safer and more consistent care while streamlining downstream decision-making and resource use.

The Pediatric Emergency Agitation Care Enhancement study is a single-center, prospective, mixed methods study designed to address this gap. The objectives are to evaluate a brief pediatric agitation risk prediction tool (BRACHA-S) and to develop an associated risk-based management plan bundle. The primary quantitative outcome is the identification of an ARI event. Qualitative participatory methods will be used to co-design a tiered, risk-based care plan. We will triangulate prospectively collected quantitative and qualitative data to fulfill our objectives through the following specific aims:

1. Aim 1 (quantitative) is to evaluate the predictive ability of the BRACHA-S administered by non-MBH specialists early in the ED course for assessing agitation or aggression in children with MBH concerns by conducting a prospective cohort study. This study hypothesizes that the BRACHA-S will have robust predictive ability (area under the curve >0.70) for ARI.
2. Aim 2 (qualitative) is to develop a multifaceted, tiered, risk-based management plan designed to mitigate agitation or aggression by gathering insights from a diverse group of stakeholders, including patients, parents or guardians, and members of the care team, using qualitative methodologies, including group-level assessment (GLA)

XSL•FO
RenderX

sessions, key informant (KI) interviews, and participatory design workshops (PDWs). The product to be created in this study is an ED risk-based agitation or aggression management plan.

Methods

Overview

This is a single-center, prospective, mixed methods study, funded by the Place Outcomes Research Award grant. The study period spans July 2024 to July 2026. Reporting of the quantitative cohort for aim 1 will follow the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines and incorporate relevant items from the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement [29,30]. Reporting of the qualitative participatory aim 2 will follow the Standards for Reporting Qualitative Research (SRQR) [31].

Aim 1: Quantitative Prospective Cohort

Approach and Population

This was a prospective cohort study that enrolled children and adolescents presenting to the ED with MBH concerns, as identified by the triage nurse and documented with a chief complaint of "psychiatric evaluation," "behavior concern," "aggression," or a similar MBH-related complaint, in accordance with predefined inclusion and exclusion criteria in Textbox 1.

Textbox 1. Aim 1 inclusion and exclusion criteria.

Inclusion criteria

- Patients aged 5-18 years
- Presentation with a mental and behavioral health concern

Exclusion criteria

- Life-threatening emergencies
- Need for immediate intervention to mitigate agitation or aggression upon arrival at the emergency department

Study Procedures

Before recruitment, nursing staff underwent training on the BRACHA-S, BVC, and STAMP tools, including recommended assessment methods for each tool, with regular retraining throughout the recruitment period. Screening windows for the 1-year recruitment period were generated using a random block method based on historical MBH patient arrival data and Clinical Research Coordinator (CRC) coverage. During these windows, CRCs reviewed the ED track board and triage documentation to identify eligible patients and approached a convenience sample. Triage or initial care team nurses completed the BRACHA-S and the initial BVC and/or STAMP shortly after patient arrival, facilitated by the CRC using tablets. Repeat BVC or STAMP scores were completed by care team members, including behavioral health specialists, patient care attendants, nurses, and clinicians, at regular intervals (approximately every 3 to 4 hours) during the patient's ED stay. Declines to complete the tools were documented.

Additional data were extracted from the electronic health record (EHR), including demographics, restraint application, administered medications, MBH specialist-acquired BRACHA scores and diagnoses, length of stay, disposition, ED diagnosis, and relevant timestamps, and merged into a comprehensive database. Staff injury data were collected using a tablet-based questionnaire adapted from a prior injury monitoring tool [32]. CRCs asked staff involved in the patient's care to complete the survey approximately once per shift during the patient's ED stay. In addition, injury events linked to enrolled patients and reported to our institutional injury reporting system were obtained from the institution's occupational safety team and uploaded into the study database.

Study Variables and Outcomes

The primary predictor was the BRACHA-S total score, with additional predictors including the initial BVC and STAMP scores recorded at triage. The primary outcome, ARI, was defined as agitation or aggression posing a risk of harm to self or others that resulted in pharmacologic management or physical or mechanical restraint during the ED visit. Intervention data were extracted from the EHR, including medications administered, use of the agitation order set, and restraint orders, and were confirmed through chart review for each agitation or aggression event. Exploratory outcomes to be assessed will include severe ARI (intramuscular pharmacologic interventions and physical or mechanical restraints), changes in BVC or STAMP scores over time, and staff injuries categorized by type, mechanism, and severity. Key covariates will include age, sex, race, ethnicity, insurance status, presenting complaint, prior MBH visits, ED length of stay, and disposition.

Statistical Analysis Plan

Aim 1 Sample Size

On the basis of prior analysis, we observed rising proportions of ARI, with rates of 1.5%, 3%, 8%, 17%, 30%, and 45% corresponding to BRACHA-S scores of 0 to 5. The sample sizes across these BRACHA-S scores were proportioned at 0.8, 1, 1, 1, 1, 1, 0.3, and 0.15, reflecting the expected relative distribution of patients across score categories. Sample size estimates were derived from logistic regression, aiming for a type 1 error rate of 5% and >80% power to detect pairwise differences across various score groupings. A sample size of 472 observations was estimated to detect differences across 4 score groups, with the estimated number of patients in each group as follows: scores 0 to 1 (199/472, 42.2% patients), score

XSL•FO
RenderX

2 (111/472, 23.5%), score 3 (111/472, 23.5%), and scores 4 to 5 (51/472, 10.8%). With more than 3500 MBH patients annually having documented BRACHA scores, a 12‐ to 18-month recruitment period using screening windows was expected to suffice.

Aim 1 Analysis

Patient characteristics, predictors, and outcomes will be summarized using descriptive statistics. Associations between demographic variables (eg, age category and sex) and ARI (primary outcome) will be examined using chi-square tests. The association between BRACHA-S scores and first ARI in the ED will be evaluated using logistic regression. Discriminative performance will be assessed with receiver operating characteristic curves and performance metrics (sensitivity, specificity, positive predictive value, and negative predictive value) at selected BRACHA-S thresholds. Calibration will be examined by comparing observed and predicted ARI rates across BRACHA-S risk groups. Sensitivity and subgroup analyses will be conducted to evaluate the robustness of the model when there is potentially population heterogeneity across age groups, sex, ED site, and workflow period. Supplementary models may include demographic variables and the initial BVC or STAMP scores as additional predictors.

Comparative analyses will explore how BRACHA-S, BVC, and STAMP scores relate to one another and differ in their ability to predict ARI. To assess the relationship between BRACHA-S and agitation-related behaviors over time, distributions and correlations between BRACHA-S and BVC or STAMP scores will be examined. A generalized linear mixed model with an appropriate count distribution (eg, Poisson distribution) and correlation structure will be considered to analyze associations between BRACHA-S and repeated BVC or STAMP scores during the ED stay, with demographic variables included as fixed effects. Additional exploratory analyses may include associations between BRACHA-S scores and staff injuries and interrater reliability of BRACHA-S component scores between nurses and MBH specialists. Analyses may be repeated using the severe ARI outcome, defined by intramuscular pharmacologic medications typically used for agitation or aggression and physical or mechanical restraints.

The extent and pattern of missing data for BRACHA-S, BVC, STAMP, outcomes, and key covariates will be described. Primary analyses will use available case data; if substantial missingness is identified for BRACHA-S or ARI, sensitivity analyses using appropriate methods such as multiple imputation may be conducted.

Aim 2: Qualitative

Approach and Population

Aim 2 uses a constructivist, stakeholder-engaged qualitative and participatory design that combines individual interviews, GLA sessions, and PDWs to codevelop a risk-stratified agitation management pathway.

Stakeholders will be recruited using purposive sampling, through in-person, email, text message, or other virtual means, to capture diverse perspectives across patients, caregivers, and ED care team members. Up to 70 stakeholders will be recruited according to the criteria in Textbox 2, including approximately 10 to 15 patients and parents or guardians for KI interviews, 15 to 20 stakeholders per GLA session (up to 3 sessions), and up to 10 stakeholders for a PDW.

Textbox 2. Aim 2 inclusion and exclusion criteria.

Inclusion criteria

- Patients aged 5-18 years who have experienced care for mental and behavioral health concerns in an emergency department setting, and/or their parents or guardians
- Stakeholders involved in emergency department mental and behavioral health care processes (ie, clinicians, nurses, social workers, counselors, behavioral health specialists, and protective services personnel)

Exclusion criteria

- Non–English-speaking participants
- Patient stakeholders with a history of severe neurodevelopmental issues or other conditions that prevent optimal participation in discussions
- Parents or guardians who do not reside with the patient participant at least 50% of the time
- Patients who are in the custody of a county child welfare agency

Study Procedures

Overview

An overview of the study procedures involved in aim 2 is presented in Table 2, with further details provided later.

XSL•FO
RenderX

Table . Qualitative methodologies and objectives.

Item number	Content	Type or sample	Objective
1	Symptoms, treatment, and experiences with restraint for agitation	Up to 10 KI[a] interviews	Conduct interviews with patients and their parents to gather insights to inform the development of a tiered agitation management plan
2a	In-depth understanding of responses and ideas from KI interviews	3 GLA[b] sessions (~30 stakeholders per session; total ~90 stakeholders)	Provide additional context and insights on KI interview findings and engage a wide range of stakeholders in generating, analyzing, and prioritizing ideas
2b	Initial draft of the care management plan	Action-planning session (7-step GLA process; 10-12 individuals)	Synthesize findings from aim 1, GLA sessions, and KI interviews
3	Development of a tiered agitation care algorithm and tailored care pathways	5 participatory design workshop sessions (5-7 individuals per session)	Create a consensus-based pediatric agitation risk prediction tool and associated care plan bundle

[a]KI: key informant.

[b]GLA: group-level assessment.

KI Interviews

These interviews will gather open‐ended insights from MBH patients and their parents or guardians about calm‐promoting strategies and experiences with agitation-related care. Patients will be recruited through the ED, and interviews will be conducted in person or through secure web conferencing. Interviews will be conducted by trained members of the research team using a semistructured guide with open-ended questions and prompts to allow participants to elaborate on their experiences [33]. Each KI interview is expected to take approximately 20 to 40 minutes to complete.

GLA Method

GLA is a qualitative, participatory method designed to engage large groups of diverse stakeholders who generate, interpret, synthesize, and evaluate data in participatory sessions [33-35]. Stakeholders, including health care providers and other staff involved in MBH care, will be recruited through the ED and personal invitations to participate in virtual GLA sessions. Up to 3 sessions will be arranged, each with up to 20 unique participants. Sessions are hosted on a secure video-based platform and recorded. All participants will complete a brief survey capturing demographics (eg, age group, sex, role, and years in role).

A trained qualitative methodologist will facilitate the 7-step GLA process [33-35]. Each session will begin with an introduction outlining its purpose, confidentiality, recording guidelines, and group rules (step 1: climate setting). Stakeholders will respond to open-ended prompts on Padlet (version 223.3.0; Wallwisher Inc; an online platform that mimics flip charts) regarding symptoms, treatments, and experiences, with a focus on developing a plan to promote the least restrictive means of restraint for agitation (step 2: generating). Participants will then examine the totality of responses to the prompts (step 3: appreciating), followed by individual reflections on others' contributions (step 4: reflecting). From the large group of stakeholders, smaller groups of 4 to 6 individuals will distill and summarize data from assigned prompts into themes to report out (step 5: understanding). The facilitator will keep a running list of themes in the participants' own words. Themes will then be prioritized by participants (step 6: selecting). The study team will synthesize the findings across the 3 GLA sessions. Select GLA participants will be invited to a separate action-planning session (step 7: GLA) to inform them of the design priorities for the risk prediction tool and care plan bundle. Attendees (n=10-12) of the action-planning session will synthesize the findings from aim 1 and the GLA sessions, aiming to develop an initial draft of a care management plan for future work. Each GLA session is expected to take approximately 90 minutes to complete.

PDW Sessions

A group of 5 to 7 key stakeholders (involved in prior GLA sessions) will participate in up to five 2-hour PDW sessions, either virtually or in person [36-38]. These workshops will focus on collaboratively integrating quantitative and qualitative data to create a consensus-based pediatric agitation risk prediction tool and associated care plan bundle, emphasizing risk mitigation and least-restrictive interventions. Facilitated by a trained methodologist, the PDWs will use structured consensus-building techniques, idea generation, and iterative discussion to guide bundle development [39,40]. Personas representing diverse patient profiles (informed by quantitative and qualitative data) will be collaboratively developed to enhance the understanding of patient experiences and inform the design of tailored care pathways [41-44]. Stakeholders will iteratively refine the tool and care plan bundle through role-play, experiential activities, and discussion.

Research Teams and Qualitative Data Management

The multidisciplinary research team (emergency medicine, psychiatry, nursing, human factors, and qualitative methods) will leverage their collective experience with ED MBH care. Brief team debriefings and analytic notes will be used to reflect on roles and assumptions. All KI interviews, GLA sessions, action-planning sessions, and PDWs will be audio- and/or video-recorded using secure platforms. Recordings will be transcribed by trained staff or a Health Insurance Portability and Accountability Act (HIPAA)–compliant transcription

XSL•FO
RenderX

service and checked for accuracy. Transcripts, Padlet responses, session notes, and other qualitative materials will be deidentified before analysis and managed using qualitative data analysis software on secure institutional servers.

Reimbursement

Participants will be reimbursed for time, effort, and travel. Participation will not be contingent on accepting payment.

Statistical Analysis Plan

Aim 2 Sample Size

Drawing from previous work led by the coinvestigator (LV), it is anticipated that 1 to 3 GLA sessions involving 15 to 20 participants each will provide ample data for thematic analysis and meet the study's objectives. Likewise, a group comprising 5 to 7 stakeholders is typically conducive to achieving consensus and finalizing the risk-stratified care pathway.

Aim 2 Analysis

KI interviews will be analyzed using an established thematic analysis approach, which uses a 6-step inductive process of coding and distilling patterns in the data (themes) [45,46]. Two analysts will code an initial subset of transcripts, develop a shared codebook, and apply it to the remaining interviews. Data generated in GLA sessions will be incorporated iteratively, with participants collaboratively analyzing and prioritizing ideas during the sessions. In the PDW, qualitative data, including field notes, participant discussions, writings, and session materials, will also be analyzed using thematic analysis [45,46]. Themes from KI, GLA, and PDW activities will be compared and integrated to inform the development and refinement of the multifaceted, tiered, risk-based management plan. Trustworthiness will be supported through triangulation across methods and stakeholder groups and the maintenance of an audit trail of analytic decisions. The final bundle will be determined through group consensus-building to reflect the collective insights and perspectives of stakeholders.

Ethical Considerations

This study was approved by Cincinnati Children's Hospital Institutional Review Board (2024-0413) and conducted in compliance with institutional review board and HIPAA regulations [47]. This study involves minimal risk to patients, families, and participating clinicians.

For aim 1, enrollment was conducted under a waiver of informed consent and HIPAA authorization. Electronic medical records were used to screen for eligibility. Identifiable data (eg, medical record number, visit time, demographics, and clinical data), agitation risk assessments (BRACHA-S, BVC, or STAMP), and linked staff injuries were stored in a REDCap (Research Electronic Data Capture) research database. No study-specific procedures altered clinical management, and interactions related to data collection occurred only with the treating team. Staff injury surveys were voluntary, and responses were recorded without personal identifiers other than clinical role. No compensation was provided for participation in aim 1.

For aim 2, written or electronic informed consent and, when applicable, assent will be obtained from patient and parent or guardian participants before the start of KI interviews. Health care team members and parents or guardians involved in GLA and PDW activities will participate under a waiver of documented consent. They receive a study information sheet and provide verbal consent or assent before study procedures. At the start of each group session, facilitators will review ground rules emphasizing confidentiality, respectful discussion, and voluntary participation. Participants may decline to answer any question or withdraw at any time without impact on clinical care or employment. Each participant will receive up to US $40 per session.

All study data, including EHR extracts, survey responses, and staff injury reports, will be stored in accordance with HIPAA and institutional data security policies [48]. Data will be entered into secure, password-protected REDCap databases or other approved institutional systems with role-based access. Identifiers will be removed or replaced with unique study codes at the earliest possible time, and linkage files will be stored separately in restricted-access locations. Only authorized study personnel will have access to identifiable data, and all reports and publications will present aggregate, deidentified results.

Results

Funding was secured in July 2024. Enrollment for aim 1 was completed on November 1, 2025, with 472 participants enrolled. Data cleaning and preliminary descriptive summaries are underway, and complete analysis will be conducted in early 2026. For aim 2, stakeholder recruitment and qualitative data collection began in October 2025 and are ongoing, with completion anticipated by May 2026. As of April 2026, 10 patients and parents/guardians were enrolled for KIs, 30 staff members were enrolled for GLAs, and 9 participants have been enrolled for PDWs. Analyses for aim 2 will be finalized by July 2026, and results will be disseminated in subsequent publications.

Discussion

Anticipated Findings

This study protocol outlines a mixed methods approach to enhance care for children and adolescents presenting to the ED with MBH concerns by prospectively evaluating the predictive performance of the BRACHA-S tool at triage and developing a tiered, risk-based management plan. With completion of this study, and assuming adequate predictive performance, the results are expected to support BRACHA-S as an early risk-stratification tool for agitation in the ED for patients with MBH concerns and to yield a stakeholder-derived, tiered management plan ready for testing in future implementation studies. This risk-stratified approach has the potential to reduce reliance on emergent pharmacologic and physical interventions, improve patient and staff safety, and enhance operational efficiency through early recognition of elevated risk, more strategic allocation of ED resources, and more targeted interventions.

Evidence supporting brief agitation or violence risk tools in ED settings remains limited. Literature syntheses of studies

XSL•FO
RenderX

conducted in adult EDs identify more than 10 brief violence risk tools and conclude that, although these instruments appear feasible for short-term risk stratification, psychometric evidence is sparse, heterogeneous, and insufficient to support recommending any single tool [49-52]. In the pediatric ED setting, far fewer instruments have been studied [53]. Studies on the Pediatric Violence/Aggression Assessment Tool [54] and the Dynamic Appraisal of Situational Aggression, Youth Version [55] suggest that higher scores are associated with increased aggression and greater restraint use in the ED; however, external validation is limited, and outcome definitions (eg, aggression events, restraint use, and medication use) vary across studies. These gaps underscore the need for rigorous testing of agitation risk tools such as BRACHA-S within real-world pediatric ED workflows. This study will also generate exploratory data on the BVC and STAMP framework in pediatric emergency care. BVC is validated for short-term aggression prediction in psychiatric units and has been implemented in adult EDs as part of workplace violence initiatives; however, pediatric ED performance data remain sparse [16,56,57]. STAMP, developed from qualitative observation in adult EDs, is widely cited in violence-prevention resources but lacks formal validity and reliability testing and has not been evaluated as a predictive tool in pediatric ED populations. Collecting these measures alongside BRACHA-S and ARI will provide comparative estimates to inform future validation studies and potential refinement of ED-specific agitation risk tools for children and adolescents.

Staff injury related to patient aggression is a central occupational safety concern in pediatric ED agitation care. National survey data show that 96% of pediatric emergency physicians have experienced workplace violence, and 10% have sustained injuries requiring medical attention or time off, with approximately half of the incidents unreported [58]. In pediatric inpatient settings, a high-risk flag identified a small subgroup that accounted for a disproportionate share of staff injuries [59], supporting exploration of whether ED risk stratification aligns with occupational risk. Furthermore, quality improvement work in inpatient settings suggests that staff injuries may be reduced through system-level responses, such as behavioral response teams for children and adolescents at risk of aggression [60]. Accordingly, this protocol will explore linkages among agitation risk scores and staff injury surveillance in the ED to inform future prevention strategies.

Current best-practice recommendations for management of agitated children or adolescents emphasize early, least-restrictive strategies, including verbal de-escalation, environmental modification, and proactive engagement of behavioral health resources, with stepwise escalation to medication and restrictive interventions only when needed to maintain safety [19,61,62]. Standardized pathways and order sets may reduce variation by clarifying roles, aligning team communication, and supporting timely selection of nonpharmacologic and pharmacologic options [9]. However, the evidence base remains heterogeneous and largely single-site, and few approaches explicitly integrate prospective risk stratification at triage to tailor the intensity of response. Results from this study may address this gap by pairing early risk identification with standardized, risk-appropriate actions in a pediatric ED setting.

Prior qualitative and mixed methods studies of pediatric behavioral emergencies describe persistent barriers to delivering least-restrictive care in ED settings, including variable staff training and confidence with de-escalation, unclear roles and communication during escalation, environmental constraints, and staff safety concerns [17,20,63]. Most published pediatric ED agitation pathways have been developed by multidisciplinary clinical teams, with limited description of patient or caregiver co-design [9]. Parent-inclusive stakeholder processes have been used to define pediatric ED agitation quality measures, demonstrating the feasibility and value of incorporating patient and caregiver perspectives in this domain [64]. In this protocol, participatory design methods extend this literature by incorporating stakeholder perspectives on risk and selection of management strategies to build a tiered approach for caring for agitated children and adolescents in pediatric ED settings.

Strengths and Limitations

The integration of both quantitative and qualitative methodologies will provide a comprehensive understanding of predictive accuracy and contextual implementation needs. A key strength of this study is its pragmatic design, which evaluates the BRACHA-S tool in real-world ED settings using frontline nursing staff. This enhances the generalizability and potential scalability of the findings. The prospective cohort design supports robust evaluation of the predictive validity of the tool, while the participatory design methods ensure that the resulting care plan bundle is informed by the lived experiences and insights of diverse stakeholders, including patients, families, and care providers.

This study has several limitations. It is conducted within a single pediatric health system, which may limit generalizability. Aim 1 uses a convenience sample during predefined screening windows, and nurse refusal to complete BRACHA-S, BVC, or STAMP may introduce selection bias. ARI is ascertained from electronic orders and documentation, so incomplete recording of behaviors, brief holds, or de-escalation efforts may lead to outcome misclassification. Additional de-escalation interventions were not systematically recorded, limiting understanding of how agitation risk scores operate within the broader spectrum of agitation management. Exploratory analyses of BVC and STAMP are not powered for precise estimates of predictive performance. For aim 2, qualitative findings rely on an English-speaking, self-selected sample from a single institution, which may limit transferability to other settings.

Conclusions

Early identification and intervention are critical for managing agitation and aggression in pediatric MBH patients; however, practical tools and workflows in emergency settings remain limited. This study positions BRACHA-S as a brief, nurse-administered triage tool for early quantification of agitation risk among children and adolescents presenting to the ED with MBH concerns. Paired with a stakeholder-informed, tiered management pathway, it establishes a structured

XSL•FO
RenderX

foundation for safer, more consistent, and patient-centered agitation care in pediatric emergency settings.

Acknowledgments

The authors thank the study research coordinators, coinvestigators, and clinical staff members who helped enroll patients. The authors also acknowledge the contributions of study participants, including patients and their guardians, health care providers, and other stakeholders.

Funding

This work was funded by the Place Outcomes Research Award through Cincinnati Children's Hospital Medical Center.

Data Availability

Deidentified data generated during this study are not publicly available but may be shared upon reasonable request to the corresponding author. Any data sharing will require review and approval by the institutional review board and compliance with the Health Insurance Portability and Accountability Act and institutional policies.

Authors' Contributions

Conceptualization: LB (lead), WP, BK (supporting)
Data curation: LK (lead), TP (supporting)
Formal analysis: YZ (lead), LV, TP (equal), LB, LK (supporting)
Funding acquisition: LB
Investigation: LK, LB, LV, TP
Methodology: LB (lead), LV (equal)
Project administration: LK (lead), TP, LB (supporting)
Resources: WP, DB, ND, HH
Supervision: LB (lead), LV (equal), WP (supporting)
Validation: LB, LV, YZ
Visualization: YZ (lead), LV (lead), LK, LB, TP (supporting)
Writing—original draft: LK (lead), LB (supporting)
Writing—review and editing: LB (lead), TP (supporting), LV (supporting), WP, HH, DB, ND, YZ, RS

Conflicts of Interest

None declared.

Peer Review Report 1
Peer review report by Place Outcomes Research Award Review Committee, Cincinnati Children's Hospital Medical Center.
[PDF File, 269 KB - resprot_v15i1e92452_app1.pdf]

References

1. Lo CB, Bridge JA, Shi J, Ludwig L, Stanley RM. Children's mental health emergency department visits: 2007-2016. Pediatrics 2020 Jun;145(6):e20191536. [doi: 10.1542/peds.2019-1536] [Medline: 32393605]
2. Bommersbach TJ, McKean AJ, Olfson M, Rhee TG. National trends in mental health-related emergency department visits among youth, 2011-2020. JAMA 2023 May 2;329(17):1469-1477. [doi: 10.1001/jama.2023.4809] [Medline: 37129655]
3. Hoffmann JA, Stack AM, Samnaliev M, Monuteaux MC, Lee LK. Trends in visits and costs for mental health emergencies in a pediatric emergency department, 2010-2016. Acad Pediatr 2019;19(4):386-393. [doi: 10.1016/j.acap.2019.02.006] [Medline: 30797896]
4. Foster AA, Porter JJ, Monuteaux MC, Hoffmann JA, Hudgins JD. Pharmacologic restraint use during mental health visits in pediatric emergency departments. J Pediatr 2021 Sep;236:276-283.e2. [doi: 10.1016/j.jpeds.2021.03.027] [Medline: 33771581]
5. Nash KA, Tolliver DG, Taylor RA, et al. Racial and ethnic disparities in physical restraint use for pediatric patients in the emergency department. JAMA Pediatr 2021 Dec 1;175(12):1283-1285. [doi: 10.1001/jamapediatrics.2021.3348] [Medline: 34515764]
6. Hoffmann JA, Corboy JB, Liu L, et al. Use of electronic health record-based measures to assess quality of care for pediatric agitation. Hosp Pediatr 2024 May 1;14(5):319-327. [doi: 10.1542/hpeds.2023-007532] [Medline: 38618654]
7. Tang AS, Shieh MS, Pekow PS, Prentiss KA, Lindenauer PK, Westafer LM. Treatment of pediatric behavioral health patients with intravenous and intramuscular chemical restraints: results from a nationwide sample of emergency departments. Acad Emerg Med 2023 Oct;30(10):1029-1038. [doi: 10.1111/acem.14754] [Medline: 37259900]

XSL•FO
RenderX

8. Ketabchi B, Babcock L, Zhang Y, Barzman D, Pomerantz WJ, Cincinnati Children's Hospital Medical Center Prehospital Care Committee. Predicting agitation in the emergency department. Pediatrics 2025 Jul 1;156(1):e2024068727. [doi: 10.1542/peds.2024-068727] [Medline: 40494553]
9. Hoffmann JA, Pergjika A, Liu L, et al. Standardizing and improving care for pediatric agitation management in the emergency department. Pediatrics 2023 Jul 1;152(1):e2022059586. [doi: 10.1542/peds.2022-059586] [Medline: 37317809]
10. Barzman DH, Brackenbury L, Sonnier L, et al. Brief Rating of Aggression by Children and Adolescents (BRACHA): development of a tool for assessing risk of inpatients' aggressive behavior. J Am Acad Psychiatry Law 2011;39(2):170-179. [Medline: 21653259]
11. Barzman D, Mossman D, Sonnier L, Sorter M. Brief Rating of Aggression by Children and Adolescents (BRACHA): a reliability study. J Am Acad Psychiatry Law 2012;40(3):374-382. [Medline: 22960920]
12. Väätäinen L, Björkqvist M, Li Y, Pelto-Piri V, Ferreira A, Lantta T. Instruments for short-term (24 h) violence risk assessment and strategies for managing violence risk among adolescents with risk for violent behaviour: a systematic review. Int J Ment Health Nurs 2025 Jul;34(4):e70110. [doi: 10.1111/inm.70110] [Medline: 40757747]
13. Almvik R, Woods P, Rasmussen K. The Brøset Violence Checklist: sensitivity, specificity, and interrater reliability. J Interpers Violence 2000;15(12):1284-1296. [doi: 10.1177/088626000015012003]
14. Woods P, Almvik R. The Brøset Violence Checklist (BVC). Acta Psychiatr Scand Suppl 2002(412):103-105. [doi: 10.1034/j.1600-0447.106.s412.22.x] [Medline: 12072138]
15. Ilarda E, McIlveen P, Tynan A, Senz A. Emergency department staff experiences of the Bröset Violence Checklist. J Adv Nurs 2024 May;80(5):2027-2037. [doi: 10.1111/jan.15961] [Medline: 37975436]
16. Hvidhjelm J, Berring LL, Whittington R, Woods P, Bak J, Almvik R. Short-term risk assessment in the long term: a scoping review and meta-analysis of the Brøset Violence Checklist. J Psychiatr Ment Health Nurs 2023 Aug;30(4):637-648. [doi: 10.1111/jpm.12905] [Medline: 36718598]
17. Lelonek G, Crook D, Tully M, Trufelli K, Blitz L, Rogers SC. Multidisciplinary approach to enhancing safety and care for pediatric behavioral health patients in acute medical settings. Child Adolesc Psychiatr Clin N Am 2018 Jul;27(3):491-500. [doi: 10.1016/j.chc.2018.03.004] [Medline: 29933797]
18. Luck L, Jackson D, Usher K. STAMP: components of observable behaviour that indicate potential for patient violence in emergency departments. J Adv Nurs 2007 Jul;59(1):11-19. [doi: 10.1111/j.1365-2648.2007.04308.x] [Medline: 17543010]
19. Gerson R, Malas N, Feuer V, Silver GH, Prasad R, Mroczkowski MM. Best practices for evaluation and treatment of agitated children and adolescents (BETA) in the emergency department: consensus statement of the American Association for Emergency Psychiatry. West J Emerg Med 2019 Mar;20(2):409-418. [doi: 10.5811/westjem.2019.1.41344] [Medline: 30881565]
20. Dalton EM, Worsley D, Krass P, et al. Factors influencing agitation, de-escalation, and physical restraint at a children's hospital. J Hosp Med 2023 Aug;18(8):693-702. [doi: 10.1002/jhm.13159] [Medline: 37401165]
21. Critical crossroads: pediatric mental health care in the emergency department. Health Resources & Services Administration. 2019. URL: https://www.hrsa.gov/critical-crossroads [accessed 2026-04-07]
22. Shah H, Somaiya M, Chauhan N, Gautam A. Clinical practice guidelines for assessment and management of children and adolescents presenting with psychiatric emergencies. Indian J Psychiatry 2023 Feb;65(2):159-174. [doi: 10.4103/indianjpsychiatry.indianjpsychiatry_494_22] [Medline: 37063627]
23. Roppolo LP, Morris DW, Khan F, et al. Improving the management of acutely agitated patients in the emergency department through implementation of Project BETA (Best Practices in the Evaluation and Treatment of Agitation). J Am Coll Emerg Physicians Open 2020;1(5):898-907. [doi: 10.1002/emp2.12138] [Medline: 33145538]
24. Latif F, Patel S, Badolato G, et al. Improving youth suicide risk screening and assessment in a pediatric hospital setting by using The Joint Commission guidelines. Hosp Pediatr 2020 Oct;10(10):884-892. [doi: 10.1542/hpeds.2020-0039] [Medline: 32928898]
25. Scudder A, Rosin R, Baltich Nelson B, Boudreaux ED, Larkin C. Suicide screening tools for pediatric emergency department patients: a systematic review. Front Psychiatry 2022;13:916731. [doi: 10.3389/fpsyt.2022.916731] [Medline: 35903632]
26. Lowry NJ, Goger P, Hands Ruz M, Ye F, Cha CB. Suicide risk screening tools for pediatric patients: a systematic review of test accuracy. Pediatrics 2024 Mar 1;153(3):e2023064172. [doi: 10.1542/peds.2023-064172] [Medline: 38356410]
27. The Joint Commission National Patient Safety Goal on Suicide Prevention in Health Care Settings.: The Joint Commission URL: https://www.jointcommission.org/en-us/knowledge-library/suicide-prevention [accessed 2026-04-15]
28. Gvion Y, Apter A. Aggression, impulsivity, and suicide behavior: a review of the literature. Arch Suicide Res 2011;15(2):93-112. [doi: 10.1080/13811118.2011.565265] [Medline: 21541857]
29. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Bull World Health Organ 2007 Nov;85(11):867-872. [doi: 10.2471/blt.07.045120] [Medline: 18038077]
30. Collins GS, Moons KG, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]

XSL•FO
RenderX

31. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
32. Macaluso M, Summerville LA, Tabangin ME, Daraiseh NM. Enhancing the detection of injuries and near-misses among patient care staff in a large pediatric hospital. Scand J Work Environ Health 2018 Jul 1;44(4):377-384. [doi: 10.5271/sjweh.3739] [Medline: 29777614]
33. Vaughn LM, Lohmueller M. Calling all stakeholders: group-level assessment (GLA)-a qualitative and participatory method for large groups. Eval Rev 2014 Aug;38(4):336-355. [doi: 10.1177/0193841X14544903] [Medline: 25147354]
34. Vaughn LM. Group level assessment methodology as a liberating structure within qualitative and participatory research. Qual Health Res 2025 Jul;35(8):827-836. [doi: 10.1177/10497323241240654] [Medline: 38648084]
35. Vaughn LM, DeJonckheere M. Methodological progress note: group level assessment. J Hosp Med 2019 Oct 1;14(10):627-629. [doi: 10.12788/jhm.3289] [Medline: 31433767]
36. Moser A, Korstjens I. Series: practical guidance to qualitative research. Part 5: co-creative qualitative approaches for emerging themes in primary care research: experience-based co-design, user-centred design and community-based participatory research. Eur J Gen Pract 2022 Dec;28(1):1-12. [doi: 10.1080/13814788.2021.2010700] [Medline: 35037811]
37. Smith S, Winkler S, Towne S, Lutz B. Utilizing CBPR Charrette in community-academic research partnerships – what stakeholders should know. J Particip Res Methods 2020;1(1). [doi: 10.35844/001c.13179]
38. Sutton SE, Kemp SP. Integrating social science and design inquiry through interdisciplinary design charrettes: an approach to participatory community problem solving. Am J Community Psychol 2006 Sep;38(1-2):125-139. [doi: 10.1007/s10464-006-9065-0] [Medline: 16794876]
39. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ 1995 Aug 5;311(7001):376-380. [doi: 10.1136/bmj.311.7001.376] [Medline: 7640549]
40. Cantrill JA, Sibbald B, Buetow S. The Delphi and nominal group techniques in health services research. Int J Pharm Pract 1996;4(2):67-74. [doi: 10.1111/j.2042-7174.1996.tb00844.x]
41. Haldane V, Koh JJ, Srivastava A, et al. User preferences and persona design for an mHealth intervention to support adherence to cardiovascular disease medication in Singapore: a multi-method study. JMIR Mhealth Uhealth 2019 May 28;7(5):e10465. [doi: 10.2196/10465] [Medline: 31140445]
42. Husain L, Finlay T, Husain A, Wherton J, Hughes G, Greenhalgh T. Developing user personas to capture intersecting dimensions of disadvantage in older patients who are marginalised: a qualitative study. Br J Gen Pract 2024 Mar;74(741):e250-e257. [doi: 10.3399/BJGP.2023.0412] [Medline: 38242714]
43. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
44. Vaughn LM, DeJonckheere M, Pratap JN. Putting a face and context on pediatric surgery cancelations: the development of parent personas to guide equitable surgical care. J Child Health Care 2017 Mar;21(1):14-24. [doi: 10.1177/1367493516645858] [Medline: 27282548]
45. Braun V, Clarke V. Thematic analysis: a practical guide. QMiP Bull 2022;1(33):46-50. [doi: 10.53841/bpsqmip.2022.1.33.46]
46. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
47. Cincinnati Children's Hospital Medical Center Policy Tech. CCRF research compliance and oversight record retention and storage. URL: https://cchmc.navexone.com/content/docview/?docid=13534&app=pt&source=unspecified [accessed 2026-04-08]
48. Cincinnati Children's Hospital Medical Center Policy Tech. Uses and disclosures of protected health information in research. URL: https://cchmc.navexone.com/content/docview/?docid=13475&app=pt&source=unspecified [accessed 2026-04-08]
49. Mesbah H, Rafique Z, Moukaddam N, Peacock WF. Predicting aggressive behavior in psychiatric patients in emergency department: a systematic literature review. Am J Emerg Med 2024 Jun;80:44-50. [doi: 10.1016/j.ajem.2024.02.040] [Medline: 38507846]
50. Anderson KK, Jenson CE. Violence risk-assessment screening tools for acute care mental health settings: literature review. Arch Psychiatr Nurs 2019 Feb;33(1):112-119. [doi: 10.1016/j.apnu.2018.08.012] [Medline: 30663614]
51. Sammut D, Hallett N, Lees-Deutsch L, Dickens GL. A systematic review of violence risk assessment tools currently used in emergency care settings. J Emerg Nurs 2023 May;49(3):371-386.e5. [doi: 10.1016/j.jen.2022.11.006] [Medline: 36585335]
52. Kamarova S, Davidson SR, Williams CM, Leite M, Kamper SJ. Predictive validity of violence screening tools in emergency and psychiatric services: a systematic review. Trauma Violence Abuse 2025 Aug 17:15248380251358224. [doi: 10.1177/15248380251358224] [Medline: 40819240]
53. Koh LL, Day A, Klettke B, Daffern M, Chu CM. The predictive validity of youth violence risk assessment tools: a systematic review. Psychol Crime Law 2020;26(8):776-796. [doi: 10.1080/1068316X.2020.1734200]
54. Mancl ME. Improving Safety in the Pediatric Emergency Department Through Early Violence/Aggression Assessment [doctoral thesis]. : University of Maryland School of Nursing; 2020 URL: https://archive.hshsl.umaryland.edu/server/api/core/bitstreams/2dd81c1a-886d-46a0-9152-285160346cac/content [accessed 2026-04-07]
55. Sund G, Kirkpatrick T, Choi KR. Applying the DASA-YV for aggression risk reduction in pediatric acute care. J Pediatr Nurs 2025;82:e23-e30. [doi: 10.1016/j.pedn.2025.03.013] [Medline: 40118749]

XSL•FO
RenderX

56. Partridge B, Affleck J. Predicting aggressive patient behaviour in a hospital emergency department: an empirical study of security officers using the Brøset Violence Checklist. Australas Emerg Care 2018 Feb;21(1):31-35. [doi: 10.1016/j.auec.2017.11.001] [Medline: 30998863]
57. Mitra B, Settle K, Koolstra C, Talarico C, Smit DV, Cameron PA. Introduction of the Broset Violence Checklist in the emergency department: a retrospective cohort study. Emerg Med Australas 2025 Feb;37(1):e14546. [doi: 10.1111/1742-6723.14546] [Medline: 39668676]
58. Huang CJ, Boulos AK, Field S, Wang VJ, Yen K. Workplace violence in the pediatric emergency department: a national survey of physicians in the United States. Pediatr Emerg Care 2024 Apr 1;40(4):249-254. [doi: 10.1097/PEC.0000000000002944] [Medline: 38563972]
59. Vaughn A, Daraiseh NM, Aeschbury M, et al. Screening tool for predicting patient aggressive behavior and staff injury at a pediatric hospital. J Psychiatr Res 2025 Jun;186:373-378. [doi: 10.1016/j.jpsychires.2025.04.029] [Medline: 40300265]
60. Martorana JA, Griffith DM, Eiger C, et al. Reducing employee injuries from aggressive patient behavior at children's hospital by implementing a behavioral response team. Pediatr Qual Saf 2025;10(1):e790. [doi: 10.1097/pq9.0000000000000790] [Medline: 39839744]
61. Saidinejad M, Foster AA, Santillanes G, et al. Strategies for optimal management of pediatric acute agitation in emergency settings. J Am Coll Emerg Physicians Open 2024 Aug;5(4):e13255. [doi: 10.1002/emp2.13255] [Medline: 39183940]
62. Georgadarellis AG, Baum CR. De-escalation techniques for the agitated pediatric patient. Pediatr Emerg Care 2023 Jul 1;39(7):535-539. [doi: 10.1097/PEC.0000000000002989] [Medline: 37383009]
63. Hoffmann JA, Kshetrapal A, Pergjika A, Foster AA, Wnorowska JH, Johnson JK. A qualitative assessment of barriers and proposed interventions to improve acute agitation management for children with mental and behavioral health conditions in the emergency department. J Acad Consult Liaison Psychiatry 2024;65(2):167-177. [doi: 10.1016/j.jaclp.2023.12.001] [Medline: 38070778]
64. Hoffmann JA, Johnson JK, Pergjika A, Alpern ER, Corboy JB. Development of quality measures for pediatric agitation management in the emergency department. J Healthc Qual 2022;44(4):218-229. [doi: 10.1097/JHQ.0000000000000339] [Medline: 35067536]

Abbreviations

ARI: aggression requiring intervention
BRACHA: Brief Rating of Aggression by Children and Adolescents
BRACHA-S: shortened 5-item version of Brief Rating of Aggression by Children and Adolescents
BVC: Brøset Violence Checklist
CRC: Clinical Research Coordinator
ED: emergency department
EHR: electronic health record
GLA: group-level assessment
HIPAA: Health Insurance Portability and Accountability Act
KI: key informant
MBH: mental and behavioral health
PDW: participatory design workshop
REDCap: Research Electronic Data Capture
SRQR: Standards for Reporting Qualitative Research
STAMP: Systems Theoretic Accident Model and Processes
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

Edited by J Sarvestan; submitted 29.Jan.2026; this is a non–peer-reviewed article;accepted 13.Mar.2026; published 21.Apr.2026.

Please cite as:
Klein L, Pomerantz W, Pestian T, Zhang Y, Barzman D, Vaughn L, Daraiseh N, Siders R, Hanson H, Ketabchi B, Babcock L
Pediatric Emergency Agitation Care Enhancement: Protocol for a Prospective Mixed Methods Study
JMIR Res Protoc 2026;15:e92452
URL: https://www.researchprotocols.org/2026/1/e92452
doi:10.2196/92452

© Lillian Klein, Wendy Pomerantz, Teresa Pestian, Yin Zhang, Drew Barzman, Lisa Vaughn, Nancy Daraiseh, Ryan Siders, Holly Hanson, Bijan Ketabchi, Lynn Babcock. Originally published in JMIR Research Protocols

XSL•FO
RenderX

(https://www.researchprotocols.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Locus-Coeruleus Norepinephrine Functioning as a Predictor of Childhood Mental Health (LOCUS-MENTAL): Protocol for a Longitudinal Study

Iskra Todorova, MSc; Paula R Neubauer, MSc; Nico Bast, PhD

Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Goethe-University, University Hospital Frankfurt, Deutschordenstraße 50, Frankfurt am Main, Hessen, Germany

Corresponding Author:
Iskra Todorova, MSc
Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, Goethe-University, University Hospital Frankfurt, Deutschordenstraße 50, Frankfurt am Main, Hessen, Germany

Abstract

Background: Mental health disorders (MHDs) remain a leading cause of the global burden of diseases. Early identification of neurobiological mechanisms mediating a risk for MHDs is key to reducing a lifetime burden. Recent findings emphasize the locus coeruleus-norepinephrine (LC-NE) system as a neuromodulator of arousal translating acute stress responses into neuronal excitability. We propose that individual differences in LC-NE functioning can explain a differential susceptibility to psychological adversity, which mediates the development of transdiagnostic psychopathology in early childhood.

Objective: The primary objective of LOCUS-MENTAL is to assess LC-NE functioning in preschoolers as a predictor of later psychopathology. This will be applied to generate an objective tool of individual early risk prediction, supporting targeted prevention of MHD.

Methods: LOCUS-MENTAL includes 4 work packages. The centerpiece is an accelerated longitudinal study, in which a cohort of 300 preschool-aged children (aged 4 - 6 years) will be recruited and followed up across 3 assessment waves, each one year apart. This will characterize developmental trajectories from 4 to 8 years of age. The primary outcome is the prediction of transdiagnostic psychopathology by pupillometry-derived LC-NE functioning. Transdiagnostic psychopathology is assessed by the Child Behavior Checklist (CBCL), while LC-NE functioning is assessed with pupillometry that includes core metrics of baseline pupil size (BPS) and stimulus-evoked pupillary response (SEPR). Cross-lagged panel models will be applied to the longitudinal data for quantifying causal effects between LC-NE functioning, childhood adversity, and psychopathology. Normative modeling and classification approaches will estimate an individual risk prediction based on pupillometric metrics of LC-NE functioning. The pupillometric battery has 4 passive auditory and visual paradigms. This battery will be validated with a multitrait-multimethod design that combines pupillometry with neurophysiological measures in electroencephalography, behavioral and cognitive measures, and neurocognitive paradigms. The study was approved by the Ethics Committee of the Faculty of Medicine at Goethe University Hospital (2024 - 2160). Written informed consent will be obtained from caregivers and verbal assent by the children.

Results: The study was funded in September 2024. Participant enrollment for the validation phase commenced in August 2025. As of February 2026, 82 participants were assessed, with a target of reaching 90 by the end of February 2026. Statistical analysis of the validation phase is planned for March 2026, with results aimed for publication in a peer-reviewed journal by the end of 2026. The longitudinal study is scheduled to start in April 2026 with completion in March 2029.

Conclusions: The LOCUS-MENTAL study will establish whether LC-NE functioning provides a validated biomarker for detecting early psychopathology. The associated pupillometry provides a feasible and scalable test for identifying high-risk developmental trajectories in preschool children, which can be translated to clinical practice. The resulting risk assessment tool could facilitate a shift toward objective screening for mental health risks that enables targeted prevention with evidence-based treatment before diagnosis onset. This could prevent sequential comorbidity and reduce the lifetime burden of MHDs.

Trial Registration: German Clinical Trials Register DRKS00035866; https://drks.de/search/en/trial/DRKS00035866

International Registered Report Identifier (IRRID): DERR1-10.2196/89237

(JMIR Res Protoc 2026;15:e89237) doi:10.2196/89237

KEYWORDS

biomarker; pupillometry; transdiagnostic psychopathology; child development; locus coeruleus

XSL•FO
RenderX

Introduction

Background

Mental health disorders (MHDs) describe enduring altered cognition, emotion, and behavior and remain a leading cause of the global burden of diseases. Recently, the locus coeruleus norepinephrine system (LC-NE) has been established as a modulator of sensory processing that translates acute stress responses into neuronal excitability. We propose this LC-NE functioning to explain a differential susceptibility to psychosocial adversity that underlies dispositional risks for MHD (Figure 1). LOCUS-MENTAL will put this to the test. We will assess LC-NE functioning in preschoolers as a predictor of later psychopathology. We will use pupillometry as a feasible neuroimaging technique to quantify LC-NE functioning in preschool children. LOCUS-MENTAL ultimately aims to generate an objective tool of individual risk prediction that informs a targeted prevention of lifetime MHD.

Figure 1. Theoretical model of LOCUS-MENTAL. Locus coeruleus-norepinephrine (LC-NE) functioning as a neurobiological mechanism in a developmental model of dispositional risks and stress, in which LC-NE functioning predicts psychopathology. LC-NE: locus coeruleus-norepinephrine.

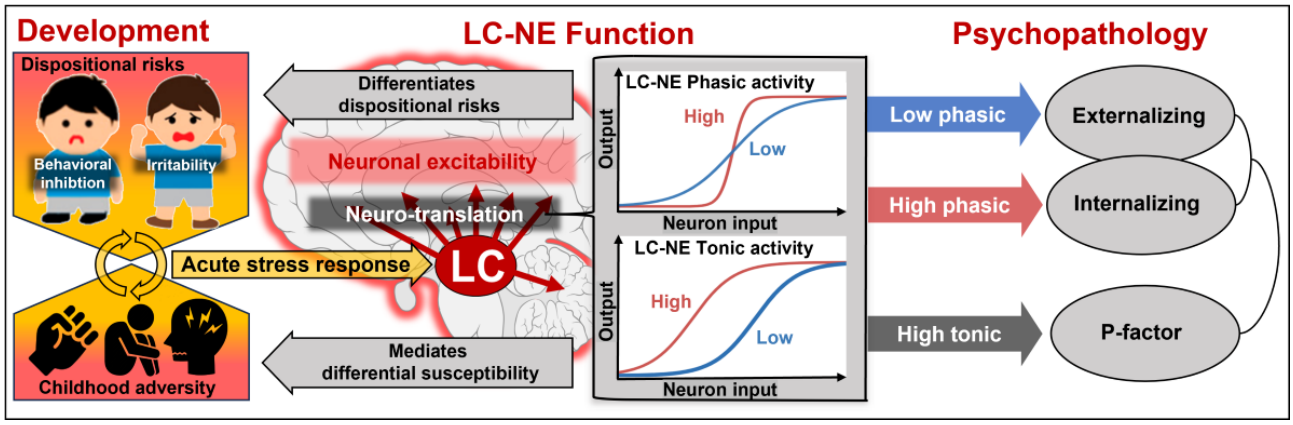

LOCUS-MENTAL uses a transdiagnostic, dimensional, and developmental approach to investigate LC-NE functioning as a predictor of MHD. Variations of LC-NE functioning have been associated with phenotypes of psychopathology [1-3]. Our work validated pupillometry as a feasible method to assess LC-NE functioning, which will now be scaled to the application in an outpatient clinic for early detection in the general population [4,5]. Preschool children will be assessed in an accelerated longitudinal study to reveal prospective effects of LC-NE functioning on psychopathology in the school age. These key concepts are discussed below.

Advancing Mental Disorders to a Transdiagnostic, Developmental Psychopathology

Comorbidity describes the co-occurrence of more than one diagnosis of MHD. Comorbidity is common, associated with more impairment and worse lifetime outcomes [6]. We lack the evidence to effectively treat MHD comorbidity at the time of this writing [7]. This may originate from conceiving MHD as independent diagnoses in clinical classification systems, which emphasized a focus on isolated diagnoses. In contrast, transdiagnostic psychopathology uses MHD comorbidity patterns to derive overarching factors. A bi-factor model with 3 factors showed the best fit in children [8]. An internalizing factor describes problems within individuals (depression and anxiety), whereas an externalizing factor describes problems between individuals (conduct problems and attention deficit hyperactivity disorder). Positive correlations between internalizing and externalizing established a superordinate p-factor that describes a liability for MHD. These 3 factors are continuous dimensions, in which manifestations below clinical thresholds are of prognostic relevance [9]. This transdiagnostic and dimensional perspective is expected to expose etiological mechanisms across MHD [10].

Comorbidity of MHD may further represent an endpoint where a secondary diagnosis arises because of a preceding primary diagnosis. This sequential comorbidity emphasizes a developmental perspective on MHD [11]. Most adults presenting comorbidity received an initial diagnosis in childhood, and half of all initial diagnoses occur before adolescence [6]. The preschool age may provide a sensible phase to identify mechanisms that influence developmental trajectories of lifetime psychopathology [12]. In addition, preschoolers differ in their risk for psychopathology even for shared psychosocial adversity [13]. This has been attributed to a differential susceptibility that explains an interaction of dispositional risks and stress [14]. We aim to identify a neurobiological mechanism of differential susceptibility [15] that mediates the impact of stress and contributes to the development of psychopathology.

LC-NE in the Translation of Sensory Input to Neurophysiological Excitability

The LC-NE system is a neuromodulatory network orchestrating the brain's response to sensory stimuli [16]. The LC-NE comprises the bilateral pontine locus coeruleus (LC) and its projections, which is the primary source of cerebral norepinephrine. Neuroimaging has suggested a bidirectional connectivity of the LC with the anterior cingulate cortex (ACC), thalamus, cerebellum, and the temporoparietal junction [17]. Compared to the substantia nigra, another neuromodulatory nucleus, the LC has shown increased connectivity to the visual, parietal, and motor cortices and the cerebellum [18]. This suggests a relative focus of the LC in earlier stages of sensory processing that perpetuates cognition and behavior.

XSL•FO
RenderX

The Adaptive Gain Theory (AGT) describes LC-NE functioning in cognition and behavior as a performance optimizer with 2 interchanging modes [19]. An exploration mode describes LC-NE tonic activity with a variable low frequency (1 - 5 Hz) that corresponds to arousal levels [20]. In contrast, an exploitation mode describes LC-NE phasic activity with transient frequency and amplitude bursts that enhance the sensory processing of task-relevant stimuli [21]. At the system level, LC-NE activity reflects alerting attention [22]. At the neuron level, this has been elaborated to a mechanism of sensory selectivity. LC-NE tonic and phasic activity release norepinephrine which causes an autoinhibition of spontaneous neuronal activity and an excitation of input-driven activity [23]. This specific mechanism of neuronal gain modulation emphasizes a sensory processing of salience [24]. In the glutamate-emphasizes-noradrenergic-effects model [25], a salience is neurochemically indicated by a release of glutamate and can be independent of top-down control [24]. Under relative arousal, LC-NE activity initiates a positive feedback loop of glutamate and norepinephrine interactions that leads to excitatory hot spots of functional activity and biases large-scale networks [26] toward sensory processing of salient stimuli [27]. It has been observed that, in mice, a moderate stimulation of LC-NE activity enhanced thalamic feature selectivity, increased information transmission, and led to improved perceptual performance [28]. LC-NE functioning has been established as a key modulator of arousal and sensory processing with subsequent effects on perception and behavior [16,29].

Pupillometry Reliably Measures LC-NE Functioning in Clinical Settings

LC-NE activity causes a dilation of the eye's pupil. This pathway is indirect, as the LC-NE inhibits the Edinger-Westphal nucleus that typically emphasizes pupil constriction [30]. Pupil dilation has historically been used as an indicator of arousal and cognitive load [31]. Recently, pupil dilation has been differentiated into a baseline pupil size (BPS) and a stimulus-evoked pupillary response (SEPR), which index changes in LC-NE tonic and phasic activity, respectively [32]. These indices are assessed by pupillometry in video-based eye tracking. A study of combined functional magnetic resonance imaging and pupillometry showed a correlation of the LC's signal with pupil dilation during the resting state (BPS) and with pupil dilation to oddball stimuli (SEPR) [33]. In a rodent model, an LC-NE stimulation predicted pupil dilation after 200 - 300 milliseconds [34]. Other subcortical regions elicited pupil dilation with larger delays, which indicated an indirect activation of the LC [35]. In an optogenetic mouse model, pupil dilation marked relative changes in LC-NE activity, whereas variability challenges pupil size as a readout of absolute LC-NE activity [36]. This can be addressed by the application of baseline-corrected metrics [37]. In humans, the assessment of LC-NE activity is difficult with magnetic resonance imaging and electroencephalography (EEG) due to the LC's location and small size [38]. This is pronounced in children that often induce more movement artifacts. Pupillometry now provides an index of LC-NE tonic and phasic activity (BPS and SEPR) that can be reliably measured with eye tracking. This innovative neuroimaging allows us to assess LC-NE functioning even in preschoolers in a clinical setting [5].

LC-NE Functioning in an Etiological Model of Dispositional Risks and Stress

Psychopathology often develops by an interaction of dispositional risks and stress [39]. Dispositional risks describe internal features that increase a latent burden. In children, stress often refers to adverse childhood experiences like neglect and further extends to psychosocial domains like peer victimization [40]. If sufficient stress occurs upon dispositional risks, a psychopathology may develop. The preschool age is sensitive to this interaction and describes a precursor phase of lifetime psychopathology [15]. In preschoolers, dispositional risk profiles are observed, which map to polygenic scores for MHD that develop much later [11]. We propose that the LC-NE functioning in preschoolers underlies dispositional risks and shapes risk trajectories across diagnoses.

Dispositional Risks for Psychopathology and the LC-NE

Temperament describes a child's innate sensory processing of affective stimuli and can represent an early dispositional risk [41]. Negative emotionality is a temperament with a proneness to negative affect that relates to later psychopathology and neuroticism. Negative emotionality has been associated with an elevated neurophysiological reactivity to reward and punishment, which explained an association between MHD polygenic scores and the p-factor in children [42]. This elevated reactivity has been reflected in an increased activity of the amygdala and the salience network and decreased activity of the ACC [43]. The LC-NE connectivity with the amygdala and ACC [44] suggests an effect of LC-NE functioning on negative emotionality.

Irritability

Irritability is a temperament trait of negative emotionality that is described by anger after frustration and attenuated fear under threat [45]. Irritability in children is the most pressing problem reported by help-seeking families [46]. Increased irritability predicted later conduct problems [47] and has been associated with a genetic risk for attention deficit hyperactivity disorder [48]. In contrast, irritability also predicted later internalizing [49]. This multifinality has been differentiated by error-related negativity (ERN) in EEG. In preschoolers with high irritability, increased ERN predicted later internalizing, while decreased ERN predicted later externalizing [50]. The ERN indicates an automated error detection that is modulated by LC-NE phasic activity [51]. Attenuated LC-NE phasic activity further attenuates a limbic fear response, which might promote externalizing behavior in irritable children [52]. We propose LC-NE phasic activity as an irritability correlate that differentiates between later externalizing and internalizing.

Behavioral Inhibition

Behavioral inhibition is another temperament trait of negative emotionality that is described by wariness and avoidance. Behavioral inhibition in young children predicted later anxiety and internalizing psychopathology as adults [53]. Behavioral inhibition in children has also been associated with an increased ERN [54] and an increased neurophysiological reactivity to

XSL•FO
RenderX

novelty [55]. It indicates elevated LC-NE phasic activity as an amplifier of the limbic fear response that emphasizes a general threat perception and promotes behavioral inhibition [56]. We will investigate patterns of LC-NE activity that correspond to the early dispositional risks of irritability and behavioral inhibition. We expect differentiating profiles of LC-NE functioning, where elevated LC-NE phasic activity predicts later internalizing and attenuated LC-NE phasic activity predicts later externalizing psychopathology.

LC-NE Activity as a Neuro-Translator of Acute Stress Responses

Stress like adverse childhood experiences is the other important determinant of psychopathology. An acute stress response is an internal coping mechanism to stress triggered by a neuronal network including the amygdala [57]. First, an autonomic response increases sympathetic activation for a momentary fight-flight-or-freeze response. A hypothalamic-pituitary-adrenal (HPA) axis then releases glucocorticoids like cortisol for sustained physical readiness. This endocrine stress response initially supports stress coping but cumulatively upregulates the HPA axis, which ultimately wears the body [58]. It explains the negative impact of stress over time. However, the acute stress response is also translated to a neuronal excitability that shapes our immediate perception in the face of stress [59].

This immediate neuro-translation of stress is moderated by LC-NE functioning [60]. The LC-NE has been established to initiate the autonomic stress response [61]. In addition, the HPA axis expresses a corticotropin-releasing hormone that binds to LC receptors and upregulates LC-NE tonic activity early in the endocrine stress response [62]. The upregulation of LC-NE tonic activity attenuates LC-NE phasic activity, which increases stimulus reactivity at the expense of sensory selectivity [63]. It reflects hyper-arousal [64] that intensifies the impact of stress [13]. For example, a liability to social stress in rodents was explained by an LC-NE modulation [65]. In humans, anxiety and stress-related disorders have been associated with increased LC-NE tonic activity [53,66]. Excessive LC-NE tonic activity in response to stress-inducing adversity might be an etiological mechanism of anxiety symptoms [62,66] that warrants an empirical evaluation in longitudinal studies.

In addition, LC-NE phasic activity is involved in the stress response by an LC projection to the amygdala in a subliminal response to fear [67]. LC-NE phasic activity emphasizes a neurophysiological response to emotional stress [68]. This has been explained by a different distribution of adrenergic receptors between the prefrontal cortex and the amygdala, where LC-NE phasic activity causes an inhibition of executive control and an excitation of the limbic fear response [69]. It has been observed in mice, pharmacological blocking of LC-NE phasic activity reduced amygdala activity during fear conditioning [70]. In a seminal study with healthy humans, LC-NE phasic responsivity and functional coupling with the amygdala during a neurocognitive task predicted internalizing symptoms following stress-inducing life events [2]. Accordingly, posttraumatic stress disorder has been associated with elevated LC-NE phasic activity in response to loud noises [71]. This supports increased LC-NE phasic activity as an amplifier of the neurophysiological fear response in the face of stress.

LC-NE activity in stress regulation can, thus, be described as a (neuro-)translation of the acute stress response to neuronal excitability [72]. We propose that LC-NE functioning as a mechanism in young children that shapes a subjective perception of stress. Distinct LC-NE functioning profiles may emphasize psychological states (eg, threat perception) that increase the psychopathological impact of stress. Thus, interindividual differences in LC-NE functioning are expected to mediate the negative impact of childhood adversity. We expect that elevated LC-NE tonic activity corresponds to hyperarousal under stress that represents a liability to transdiagnostic psychopathology (p-factor). We further expect deviant LC-NE phasic activity to differentiate the development of internalizing versus externalizing psychopathology in response to childhood adversity. Elevated LC-NE tonic and phasic activity have been associated with internalizing symptoms in adults (eg, anxiety and negative affect), whereas blunted LC-NE phasic activity might indicate a subtype at risk for conduct problems [73]. LOCUS-MENTAL will investigate LC-NE functioning in preschoolers that mediates a differential susceptibility to adversity and differentiates future profiles of transdiagnostic psychopathology (see Figure 1).

Preliminary Work

Validity of Pupillometry in Clinical Samples

We established a model that introduced LC-NE activity as a modulator of attentional function in autism [74,75]. We manipulated task utility to induce changes in pupillary responses that related LC-NE functioning to performance [1]. We established the LC-NE to influence memory [76] and social attention [77]. At the time of this writing (DFG number: 492582254), we validate LC-NE measures by a combined pupillometry and EEG study. In data of 150 children, we showed a higher SEPR for oddball versus standard stimuli in a passive auditory oddball task (manuscript under review). In a recent publication, we used the European Autism Interventions Longitudinal European Autism Project sample [78] to analyze combined pupillometric and EEG data in a large cohort of deeply phenotyped autistic and nonautistic individuals. We showed that baseline BPS and SEPR as pupillometric measures drive the event-related potential of mismatch negativity in the response to oddball sounds [4]. This well-investigated auditory oddball task is applied as a reference task in this research program (Figure 2A).

XSL•FO
RenderX

Figure 2. Preliminary work on auditory oddball task. (A) Stimulus-evoked pupillary response (SEPR) of children in an auditory oddball task. (B) SEPR to oddball stimuli as an index of LC-NE phasic activity is associated with salivary cortisol levels. (C) SEPR to oddball stimuli differentiated between factors of transdiagnostic psychopathology within clinically relevant ranges.

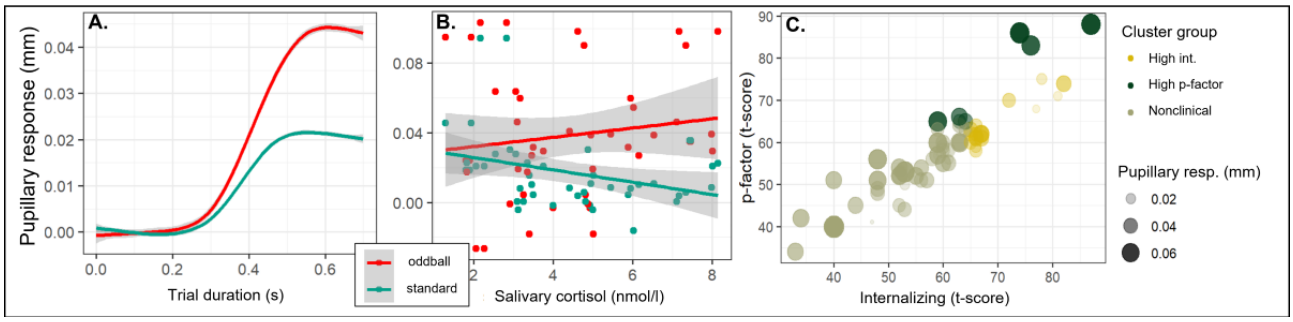

Reliability of Pupillometry in Preschoolers

We showed the feasibility of our methodology in early childhood by acquiring longitudinal pupillometric data in an eye-tracking battery of neurocognitive tasks in 100+ preschool children. Two cross-sectional analyses have been published [79,80], while a longitudinal study applied pupil responses as a mediating mechanism in early intervention outcomes [5]. We assessed the reliability of pupillometry in 26 participants with 3 measurements (baseline, after 6 months, and after 12 months). BPS showed high correlations between measurement time points (*rs*=0.69−0.81) and an intraclass correlation of ICC=0.90, 95% CI 0.83-0.95. SEPR showed lower correlations between measurement time points that might be induced by intervention effects.

Pupillometry and Stress Response

We further assess pupillometry and stress responses during an oddball task in adolescents (32-01/22). In preliminary data (n=32), BPS was associated with an attenuated downregulation of salivary cortisol (β=.56, 95% CI 0.38-0.74; R^2=0.28), while pupillary response to oddballs (SEPR) was associated with higher hair (Δβ=.24, 95% CI 0.07-0.42; R^2=0.15) and salivary cortisol levels (Figure 2B) supporting the LC-NE in stress responses. Finally, SEPR differentiated psychopathology above clinical cutoffs (Figure 2C). Taken together, our preliminary work underlines the feasibility and reliability of pupillometry in preschoolers, outlines our capacity to recruit burdened preschoolers, and emphasizes pupillometric responses as an index of LC-NE activity in stress and psychopathology.

Main Objectives

The primary aim of LOCUS-MENTAL is to establish pupillometric measures of LC-NE functioning as a predictor of transdiagnostic psychopathology in childhood. We hypothesize that LC-NE functioning represents a fundamental neurobiological mechanism in the development of MHD that could serve as a biomarker for early risk detection. Transdiagnostic psychopathology is assessed with the caregiver-reported Child Behavior Checklist (CBCL), where clinical relevance is defined as reaching established cutoffs (T>60). LC-NE functioning is assessed with pupillometry. The study outcome could improve targeted prevention and intervention for MHD comorbidity in children. To achieve this aim, LOCUS-MENTAL is structured into 4 consecutive objectives, each with specific tasks within the following project plan (Figure 3).

Figure 3. Work program of LOCUS-MENTAL with time flow, required personnel, and the associated tasks.

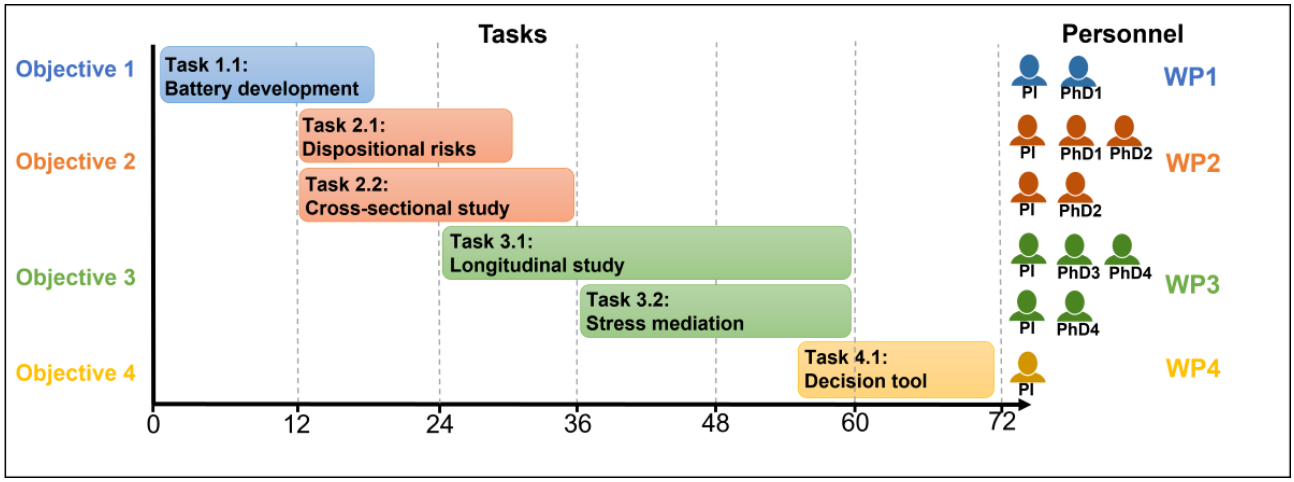

Objective 1

This study aims to validate a concise test battery assessing LC-NE system functioning in preschool children using pupillometric measures during passive tasks. The test battery will be evaluated for convergent and discriminant validity.

XSL•FO
RenderX

Objective 2

This study aims to associate LC-NE functioning with correlates of psychopathology in a cross-sectional study. Task 2.1 will determine whether LC-NE functioning profiles serve as markers of dispositional risks and differentiate between irritability and behavioral inhibition. Task 2.2 will examine whether LC-NE functioning further mediates the association between childhood adversity and transdiagnostic psychopathology.

Objective 3

This study aims to evaluate the prospective effect of LC-NE functioning in a developmental model of vulnerability and stress. Task 3.1 will investigate whether pupillometric measures of LC-NE functioning in preschool age predict risk for psychopathology at school age. Task 3.2 will assess LC-NE functioning as a mediator between early stress responses and later psychopathology.

Objective 4

This study aims to develop an objective tool of risk prediction in the preschool ages based on the pupillometric measures of LC-NE functioning. LC-NE function profiles in preschoolers define a normative development space, where individual deviations indicate quantitative risks for MHD.

Methods

Overview

LOCUS-MENTAL will achieve its objectives by combining pupillometric indices of LC-NE functioning and using a prospective panel study design (Figure 4). Pupillometry is assessed by video-based eye tracking during basic paradigms that are feasible in preschoolers. Transdiagnostic psychopathology is assessed by CBCL externalizing, internalizing, and total T-scores. Childhood adversity is assessed by interviewing caregivers, children, and clinicians. Pupillometric indices, transdiagnostic psychopathology, and childhood adversity are repeatedly measured in an accelerated longitudinal study: We assess preschoolers (aged 4 - 6 years) in an initial assessment with reassessments after one year (aged 5 - 7 years) and after 2 years (aged 6 - 8 years). This accelerated design allows revealing causative effects of LC-NE functioning on psychopathology from 4 to 8 years of age. In addition, this allows quantifying LC-NE functioning as a mediator in the effect of childhood adversity on psychopathology. In distinct subtasks, LC-NE functioning is elaborated as a neurophysiological mechanism of established dispositional risks, while LC-NE functioning in stress regulation is explored by combined pupillometry and analysis of hair cortisol levels and salivary cortisol after a social stress test. This will characterize LC-NE functioning profiles in the interaction of dispositional risk and stress. Finally, LOCUS-MENTAL applies normative modeling on the pupillometric data to develop a tool for risk prediction of developmental trajectories in transdiagnostic psychopathology. An overview of all measures is provided below (Table 1). The reporting of this study protocol follows the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for longitudinal cohort studies [81]. The completed checklist is available in Checklist 1.

Figure 4. LOCUS-MENTAL study design with task integration. The flowchart illustrates the longitudinal study design consisting of a pilot phase and 3 primary assessment waves (**T1, T2, AND T3**) each one year apart. Each assessment wave spans over 2 years. The longitudinal study will follow a cohort of 300 preschool children (initially aged 4 - 6 years). Dark blue boxes indicate longitudinal study elements, and light blue boxes indicate cross-sectional tasks within the longitudinal study. Dotted lines indicate selection of a specific subsample at specific time points, including a randomized subsample (n=160) either from T2 or T3 for Task 3.2. Small “n” represents the subsample number, while “N” represents the primary study sample.

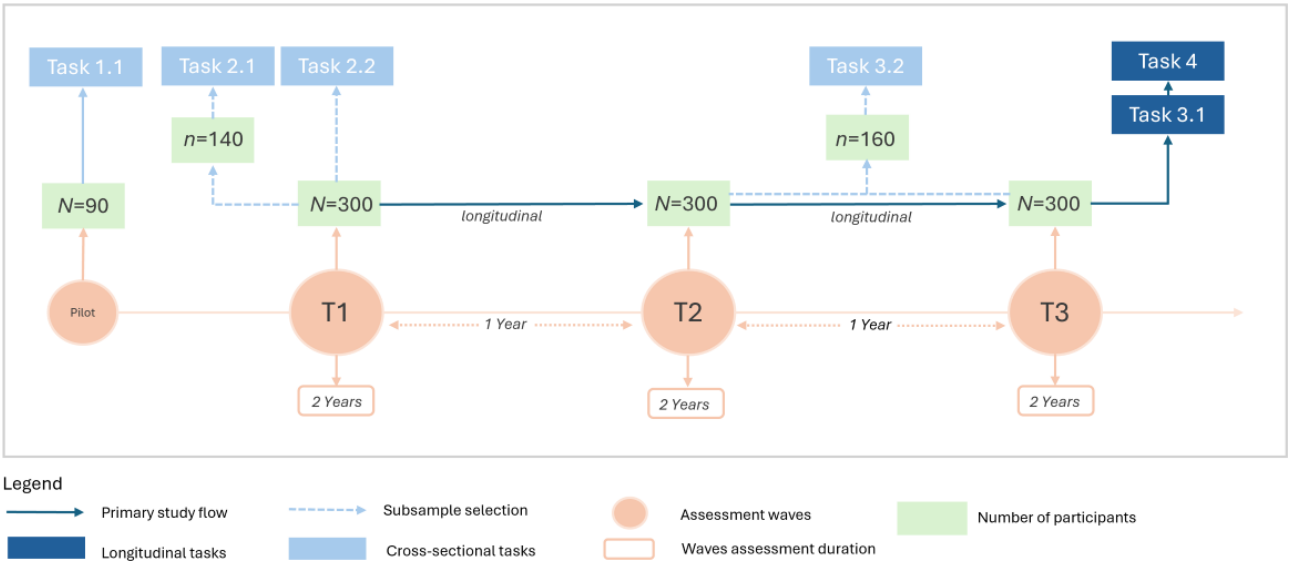

XSL•FO
RenderX

Table . Assessment measures and questionnaires of LOCUS-MENTAL.

Modality, instrument and task	Construct	Objective
Questionnaires (caregiver)		
core set	Demographic data	1.1. - 3.2
CBCL[a] [4-18,82]	Psychopathology	2.1 - 3.2
BIQ[b] [83]	Behavioral inhibition	2.1 - 3.2
ARI[c] [84]	Irritability	2.1 - 3.2
CBQ-VSF[d] [85]	Temperament	2.1 - 3.2
ACE+[e] [86]	Childhood Adversity	2.2 - 3.2
F-SoZu 6[f] [87]	Social support	2.2 - 3.2
Interview (child)		
CLES-P[g] [88]	Childhood adversity	2.2 - 3.2
VEX-Preschool[h] [89]	Violence exposure	2.2 - 3.2
Pupillometry battery		
Auditory oddball [90]	LC-NE activity	1.1 - 3.2
Visual oddball [91]	LC-NE activity	1.1 - 3.2
Cued visual search [92]	LC-NE activity	1.1 - 3.2
Rapid sound sequences [24]	LC-NE activity	1.1 - 3.2
Cognitive tests		
WPPSI[i] subtest–vocabulary [93]	Verbal IQ estimate	1.1 - 3.2
WPPSI subtest–block design [93]	Nonverbal IQ estimate	1.1 - 3.2
Bio samples		
Hair sample	Retrospective cortisol level	2.2 - 3.2
Saliva sample	Acute cortisol level	2.2 - 3.2
Assessment		
Child stress test [94]	Stress reaction	3.2
EEG[j]		
See pupillometry tasks	LC-NE activity	1.1 - 2.1
Go/NoGo Task [50]	Error Related Negativity (ERN)	2.1
Neurocognitive tasks		
Dot-probe tasks [95]	Irritability	2.1
Frustrative reward learning task [96]	Irritability	2.1

[a]CBCL: Child Behavior Checklist
[b]BIQ: Behavioral Inhibition Questionnaire
[c]ARI: Affective Reactivity Index
[d]CBQ-VSF: Children's Behavior Questionnaire–Very Short Form
[e]ACE+: Adverse Childhood Experiences Questionnaire Plus
[f]F-SoZu 6: Social Support Questionnaire
[g]CLES-P: Childhood Life Events Scale–Preschool
[h]VEX-Preschool: Violence Exposure Scale–Preschool
[i]WPPSI: Wechsler Preschool and Primary Scale of Intelligence
[j]EEG: electroencephalogram.

XSL•FO
RenderX

Recruitment

The main recruitment avenues will be the outpatient clinic for early detection at the Department for Child and Adolescent Psychiatry at the Goethe University Hospital. In the help-seeking population, externalizing disorders might be overrepresented, which is why we will overrecruit preschoolers with internalizing symptoms. We aim for a representative sample by recruitment across all levels of psychopathology. This is why we will also recruit in the general population by distributing flyers with study information at local health care and public institutions, pediatricians, and day care facilities and promoting the study in targeted social media groups and channels. Exclusion criteria will be known genetic syndromes and neurological diseases. Caregivers will receive a summary report of their child's assessment (eg, IQ evaluation) upon completion of each assessment wave.

General Setting

The experimental assessments will be carried out at our combined eye-tracking and EEG lab that we established at the clinic. The procedure is framed as a play experience. The child watches a sequence of their favorite movie on the stimulus presentation screen. This is followed by the assessments where the child may sit on their caregiver's lap. Each paradigm is designed to enable breaks. Questionnaire measures may be filled out by the caregivers via our in-house online tool to decrease the participant burden, whereas structured clinical interviews and child questionnaires will be carried out by trained clinical staff.

Sample Size and Power

Sample sizes (n) were determined based on recruitment feasibility and the ability to detect moderate effects with a power of $\beta>.80$. For each task, we estimated the power by Monte Carlo simulations (k=1000) within respective models [97,98]. In the following sections, the sample size and power are reported, respectively, for each specific objective or task.

Work Package 1 (Objective 1): Validation of a Pupillometric Battery of LC-NE Functioning

Participants

We will recruit 90 preschool-aged children to participate in a passive pupillometry test battery designed to measure LC-NE activity.

Sample Size and Power

A power analysis using a linear mixed model (LMM) indicates a power of $\beta=.82$ for 90 participants to detect a moderate stimulus effect (b=0.3) in the auditory oddball task (100 trials, 20% oddballs).

Experimental Design

This task uses a multitrait, multimethod design combining pupillometry, EEG, and behavioral assessments to validate the pupillometry battery.

Battery Design

Overview

The battery consists of 4 individual tasks. The primary task is a passive auditory oddball task, and the 3 additional tasks are a rapid sound sequence, a visual oddball, and an audio-cued visual search task. At the beginning of each task, a fixation cross in the center of the screen is presented for 5 seconds and can be used as a global baseline. After the last task, there is an additional measurement of a 5-second baseline. All tasks are presented in a fixed order, and in between the tasks, a short children's cartoon (duration=30 s) is presented to maintain children's attention and motivation. The overall duration of the battery is approximately 16 minutes. Prior to the task, we conduct a 6-point calibration. The source code of the tasks as a Python (Python Software Foundation) implementation can be found on the GitHub repository "locusmental_wp1_tasks" of Nico Bast.

Auditory Oddball Task

A passive auditory oddball task that is feasible in preschoolers (see 1.3, Figure 2) will be applied as a reference measure [90]. The auditory oddball task consists of a series (k=100) of frequent ("standard," 80%) or infrequent ("oddball," 20%) pure tones. Pure tones will have a short duration (50 ms) and will be presented with an interstimulus interval of a random duration between 1.8 and 2 seconds, while the pitch of the pure tones (500 or 750 Hz) will be counterbalanced between participants. Attention will be directed to the screen center by a continuous cartoon video without sound and controlled for by gaze-contingent trial presentation. Deviant gazes will evoke an attention-grabbing stimulus to redirect attention.

Rapid Sound Sequences Task

The Rapid Sound Sequences task consists of auditory stimuli 6 seconds-long sequences of subsequent tone pips (50 ms duration), drawn from a pregenerated pool of 20 fixed frequencies. Each sequence is followed by a 2-second interstimulus interval. The tone-pip sequences are arranged according to varying patterns and are randomly generated for each participant and on each trial. Conditions vary by levels of auditory regularity and irregularity with 2 control and 3 transition conditions. Previous research showed that transitions from regularity to irregularity evoked pupillary responses [24].

- Control condition regular 10 (REG10): randomly selecting 10 frequencies from the pool, arranging them in a fixed sequence pattern, and iterating the sequence to create a regular pattern.
- Control condition random 20 (RAND20): randomly selecting 20 frequencies with replacement from the pool and playing these fully randomized.
- Transition condition REG10-RAND20: the first 3 seconds are presented as REG10, followed by a transition to RAND20 for 3 seconds.
- Transition condition RAND20-REG10: the first 3 seconds are presented as RAND20, followed by a transition to REG10 for 3 seconds.
- The transition condition RAND20-REG1: the sequence starts as RAND20 for 3 seconds and transitions into REG1,

XSL•FO
RenderX

a single tone pip selected from the pool and iterated for the final 3 seconds.

Each control condition will be presented 5 times, and each transition condition 10 times, in randomized order. Attention will be directed to the screen center by a fixcross or a video and controlled for by gaze-contingent trial presentation. Deviant gazes will evoke an attention-grabbing stimulus to redirect attention. A silent cartoon will be presented during the interstimulus interval.

Visual Oddball Task

The visual oddball task consists of 72 trials displaying frequent (standard size: 2.9° visual degrees; 80%) and infrequent (oddball size: 4.6° visual degrees; 20%) blue circles. Each stimulus is displayed for a 150-millisecond duration and is preceded by a 1.5-second interstimulus interval. Attention is maintained using gaze-contingent trial presentation.

Audio-Cued Visual Search Task

Each trial presents a visual search display consisting of 4 colored circles (red, yellow, or green) arranged in an imaginary circle around the center of the screen. One circle (target circle) has a deviant color from the 3 distractors, and its position changes randomly across the 30 trials. The visual search phase has a duration of 1.5 seconds. An auditory alerting cue, a simple brief beep, is presented in 50% of the trials. The cue duration varies randomly between 200 and 300 milliseconds and is presented with a random delay of 0‐100 milliseconds within a 400-millisecond interval prior to the visual search onset. Each trial starts with a 1.5-second interstimulus interval with a fixation cross in the center of the screen. Attention is maintained using gaze-contingent trial presentation.

Neurophysiological Measures

BPS and SEPR are assessed for each trial in each task. BPS is defined as the pupil size prior to stimulus onset, while SEPR represents a pupil size change in response to stimuli. BPS and SEPR will be evaluated in the auditory oddball task in response to standard trials as neurophysiological habituation [99] and to oddball trials as sensory selectivity [33].

Cognitive Measure

To assess the verbal and cognitive abilities, we will conduct 2 subtests from the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) [93].

Statistical Analysis Plan

We will validate the pupillometric indices by a multitrait-multimethod design. In a homo-method analysis, the BPS and SEPR during the auditory oddball task (Figure 5A) will be compared to BPS and SEPR in tasks of LC-NE functioning with irregularity in rapid sound sequences [24] (Figure 5B), a visual oddball task [91] (Figure 5C), and an audio-cued visual search [92] (Figure 5D).

In a hetero-method analysis, the BPS and SEPR during the auditory oddball task will be compared to event-related potentials (ERPs) in concurrent EEG. We investigate oddball ERPs that are likely associated with LC-NE activity (P300, MMN) [27,100]. Finally, the pupillometric measures will be compared to divergent measures of cognitive and language ability as assessed by WPPSI-IV [93]. Patterns of high intercorrelations will be used to create a concise pupillometric battery of LC-NE functioning (~15 min). BPS and SEPR are reliably assessed in few oddball trials (k=20).

Figure 5. Experimental tasks. (**A**) Auditory oddball task (see description above). (**B**) Irregularity in rapid sound sequences; sequences either have a pattern (regular=REG), have no pattern (irregular=IRR), or change regularity within the sequence (REG to IRR, IRR to REG). (**C**) Visual oddball task. (**D**) Cued and uncued visual search task. IRR: irregular; REG: regular.

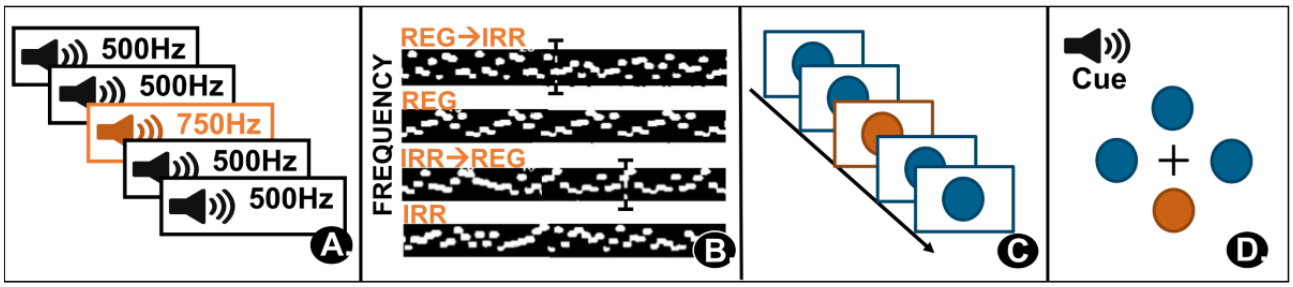

Expected Outcome and Study Hypotheses

A successful task manipulation will be indicated by a higher SEPR in oddball versus standard trials (ie, sensory selectivity) and declining BPS across standard trials (ie, habituation). We expect convergent validity with moderate-to-high correlations between the auditory oddball SEPR and the pupillometric measures in the validation tasks (including SEPR to a change from regularity to irregularity, SEPR to visual oddball stimuli, and SEPR before target detection). We expect convergent validity in the hetero-method analyses, whereas we expect divergent validity with cognitive and language development.

Risk Assessment

We only include paradigms that are feasible in young children. This shall ensure a sufficient completion rate in preschoolers (>80%). The concurrent presentation of visual stimuli to maintain attention during the auditory oddball task may impede the SEPR. However, we recently showed in independent data that visual presentation during an auditory oddball task still induces distinct task-evoked pupillary responses.

XSL•FO
RenderX

Work Package 2 (Objective 2): Association of LC-NE Functioning With Correlates of Psychopathology

Task 2.1: LC-NE Functioning Corresponds to Early Dispositional Risks for Psychopathology

Participants

A sample of 140 preschool-aged children will be recruited in accordance with the general recruitment strategy of the study.

Sample Size and Power

A power analysis based on an LMM indicates that a sample of 140 participants provides power of β=.84 to detect a moderate correlation (b=0.3) between LC-NE functioning and dispositional risks.

Experimental Design

Cross-sectional analysis of combined pupillometry, EEG, and behavioral and cognitive assessments.

Neurophysiological Measures

The validated pupillometric battery, as described in Objective 1, will be used to assess BPS and SEPR across multiple tasks. In a concurrent EEG, the ERP of ERN will be assessed using an age-appropriate Go/No-Go task [50]. Neurocognitive paradigms of irritability will be examined using a dot-probe task [95] and a frustrative reward learning task [96].

Cognitive Measure

To assess the verbal and cognitive abilities, we will conduct 2 subtests from the WPPSI [93].

Behavioral Measures

Temperament traits will be measured via caregiver reports with the Behavioral Inhibition Questionnaire (BIQ) [83] and Affective Reactivity Index (ARI) [84]. Negative emotionality will be assessed via Child Behavior Questionnaire-Very Short Form (CBQ-VSF) [85]. The 3 transdiagnostic factors of psychopathology will be assessed with the Child Behavior Checklist (CBCL 4‐18) [82].

Statistical Analysis Plan

We will apply LMM analysis with participants as random intercepts. These will allow us to explore associations between variables on a per-task level and control for potential covariates (age, gender, and IQ).

Expected Outcome and Study Hypotheses

Behavioral inhibition (BIQ score) is expected to correlate positively with both BPS and SEPR in the pupillometric battery of LC-NE functioning. In contrast, irritability (assessed via ARI score) is expected to show differential profiles of LC-NE functioning. We expect high irritability scores and high SEPR to be associated with higher internalizing symptoms, while high irritability scores and low SEPR will be associated with higher externalizing symptoms—these expected outcomes are supposed to match ERN profiles [101]. Furthermore, in the neurocognitive paradigms, participants with high-irritability-low-SEPR profiles are further expected to show increased attentional bias toward threat cues (threat perception) and decreased attentional shifts during frustration (frustrative reward learning). These patterns would suggest a subgroup with elevated risk for conduct disorders. We will further explore LC-NE functioning in children with both high irritability and high behavioral inhibition, who are expected to exhibit higher negative emotionality and an increased BPS as an index of an elevated LC-NE tonic activity.

Risk Assessment, Participant Burden, and Ethical Considerations

This task is associated with the highest participant burden due to its combined EEG assessments and neurocognitive task demands. To mitigate this, we will provide additional breaks for the participants and offer food and drinks between assessments. We assume a higher attrition rate for this task compared to other tasks, estimated at approximately 30%. Thus, this task is designed as more optional, ensuring that potential dropouts will not compromise subsequent study objectives.

Task 2.2: LC-NE Function Mediates the Association Between Childhood Adversity and Transdiagnostic Psychopathology

Participants

A sample of 300 preschool-aged children (aged 4‐6 years) will be recruited in accordance with the general recruitment strategy of the study.

Sample Size and Power

A power analysis based on a structural equation model indicates that a sample of 300 children provides a power of β=.80 to detect a moderate path coefficient (a=0.2) between latent variables of LC-NE functioning and psychopathology.

Experimental Design

Cross-sectional analysis of combined pupillometry, behavioral, and cognitive assessments will be used.

Neurophysiological Measures

The validated pupillometric battery, as described in Objective 1, will be used to assess BPS and SEPR across multiple tasks.

Cognitive Measure

To assess the verbal and cognitive abilities, we will conduct 2 subtests from the WPPSI [93].

Behavioral Measures

Transdiagnostic psychopathology factors (externalizing, internalizing, and p-factor) will be assessed with CBCL 4‐18. In addition, we will assess childhood adversity based on multiple sources. Specifically, a caregiver interview that captures adverse childhood experiences and other psychosocial domains (ACE+) [40], a respective child interview [86], validated child-report measures of adversity [88,89], and a report by the respective clinician from the outpatient clinic for those children recruited via the outpatient recruitment leg. This assessment approach will provide an index of cumulative adversity or will be differentiated into specific adversity domains.

Statistical Analysis Plan

Structural equation models will be applied to estimate the associations between the latent constructs (LC-NE functioning,

XSL•FO
RenderX

transdiagnostic psychopathology, and childhood adversity). A best-fit path model will be used to reduce the complexity of the structural equation in the longitudinal study of Objective 3. A latent mediation model, in which LC-NE functioning is expected to mediate an association between previous childhood adversity and transdiagnostic psychopathology, will be applied. Independent models for BPS, SEPR, and each transdiagnostic factor will be estimated. Classification approaches will be applied in exploratory analyses to identify LC-NE function profiles in subgroups of psychopathology.

Expected Outcome and Study Hypotheses

We hypothesize associations between retrospective childhood adversity, LC-NE function, and transdiagnostic psychopathology. For example, lower LC-NE tonic activity is expected to represent a resilience to the negative impact of childhood adversity on psychopathology. In addition, we expect profiles of LC-NE functioning that are associated with patterns of transdiagnostic psychopathology; BPS is expected to positively correlate with the p-factor, whereas increased and decreased SEPR are expected to positively correlate with internalizing and externalizing symptoms, respectively.

Risk Assessment, Participant Burden, and Ethical Considerations

We anticipate that some children will require immediate support, and we will use our clinical infrastructure to provide further case management. In addition, reports of adversity will be evaluated according to our in-house reporting guidelines, which align with the German child protection laws. This includes evaluation in a joint discussion and may result in reporting to child protection authorities when imminent danger is identified.

Work Package 3 (Objective 3): Prospective Effect of LC-NE Functioning in a Developmental Model of Vulnerability and Stress

Task 3.1: LC-NE Functioning Predicts Later Transdiagnostic Psychopathology.

Participants

The study will follow the sample of 300 preschool-aged children, initially assessed in Task 2.2 at 4 - 6 years of age across 3 time points.

Sample Size and Power

For the core longitudinal study with a cross-lagged panel model, power is β=.82 for 300 participants to detect a moderate cross-lagged effect of LC-NE functioning on psychopathology (a=0.2; see Figure 6).

Figure 6. Autoregressive cross-lagged panel model to investigate prospective effects (**A**) of LC-NE functioning on psychopathology (PSY). b and c: autoregressive effects; e2i, e3i, e4i, and e5i: time-specific measurement errors; LC: locus coeruleus; PSY: psychopathology.

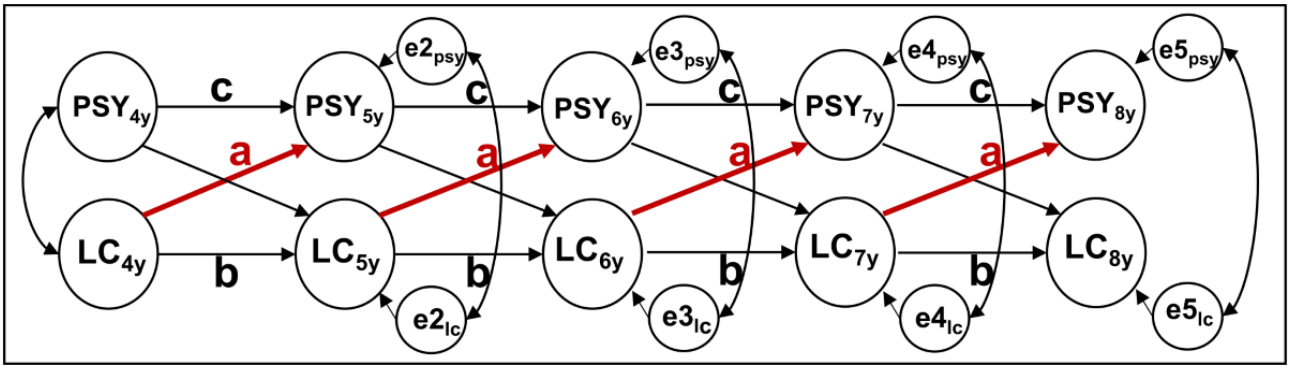

Experimental Design

The study uses an accelerated longitudinal design with 3 assessment waves and time points: the baseline or initial assessment as part of Task 2.2 (T1; aged 4 - 6 years), a 1-year follow-up (T2; aged 5 - 7 years), and a 2-year follow-up (T3; aged 6 - 8 years). At each time point, we assess neurophysiological (pupillometry), behavioral, and cognitive measures.

Neurophysiological Measures

At each time point, the validated pupillometric battery, as described in Objective 1, will be used to assess BPS and SEPR across multiple tasks.

Cognitive Measure

To assess the verbal and cognitive abilities, we will conduct 2 subtests from the WPPSI [93].

Behavioral Measures

At each timepoint assessment of the behavioral measures, as described in Task 2.2, will include caregiver questionnaires (CBCL 4 - 18) and the multiple-source assessment of childhood adversity (see Table 1).

Statistical Analysis Plan

The accelerated longitudinal design allows estimation of an autoregressive cross-lagged panel model that will quantify the effect of previous LC-NE functioning (BPS and SEPR) on later transdiagnostic psychopathology factors (CBCL) across the age of 4-8 years (see Figure 6). We will control for potential cohort effects (initial age: 4 vs 5 vs 6 years) by including cohort as a random intercept. Each model fit will be evaluated by the predefined indices of CFI, TLI, RMSEA, and SRMR.

XSL•FO
RenderX

Expected Outcome and Study Hypotheses

This core of LOCUS-MENTAL will be informed by the outcome of task 2.2. We expect (see Figure 1) that higher BPS as an index of LC-NE tonic activity predicts a higher p-factor. Increased and decreased SEPR as an index of LC-NE phasic activity is expected to predict higher internalizing and externalizing symptoms, respectively. This would predict future psychopathology based on pupillometric markers. The stability of the investigated constructs is controlled for by autoregressive effects. The structural equation models will be explored for the covariate effects of childhood adversity, previous psychopathology, and received clinical support (eg, in the context of the outpatient clinic for early detection).

Risk Assessment

A major risk is a low retention rate. A meta-analysis identified a mean retention rate of 73.5% and identified strategies to increase it [102]. This includes barrier reduction by including remote data collection (eg, offer data collection via phone and online), tracing strategies by locator documents with contact information, and community-building by sending participants thank-you, birthday, and holiday cards. We will implement these strategies in LOCUS-MENTAL to maximize our retention rate while maintaining ethical standards.

Task 3.2: LC-NE Functioning Mediates the Effect of Stress Responses on the Development of Transdiagnostic Psychopathology

Participants

Participants in this study will comprise a subsample of 5 - 6-year-olds (n=160) at one randomly selected timepoint (T2 or T3) from the large longitudinal cohort (Task 3.1).

Sample Size and Power

A power analysis using a bivariate latent growth curve model indicates a power of β=.82 for 160 participants to detect a moderate path coefficient (b=0.2) of cortisol reactivity on changes in psychopathology.

Experimental Design

This task uses a cross-sectional assessment of endocrine stress responses within the longitudinal study.

Measures

Endocrine stress responses provide a more proximate measure of stress compared to childhood adversity. We will assess stress responses through hair cortisol levels [103] and salivary cortisol measurements following a developmentally appropriate social stress test [94]. Hair cortisol levels will quantify the retrospective stress levels, while salivary cortisol will quantify the acute stress response.

For hair cortisol, we will collect 3 centimeters in length of scalp-nearest hair; this will contain cortisol information over the preceding 3 months. Salivary cortisol will be collected via fiber chewing before and after a social stress test (after 30-40 minutes). In the social stress test, the children are instructed to match stickers to animals to win a prize. Time pressure will be applied by a commenting experimenter and a toy traffic light that turns red and emits a buzzer sound [94,104]. Contractors will process salivary (Daacro) and hair (Dresden Lab service) samples.

Statistical Analysis Plan

The cortisol measures will be applied in a mediation model to examine LC-NE functioning in the association of stress response and psychopathology. We will further quantify the effect of LC-NE functioning in the association of childhood adversity and stress response.

Expected Outcome

We expect a positive association between cortisol measures and psychopathology. Based on our model (Figure 1), we assume BPS and SEPR to mediate this direct effect. Higher BPS and SEPR are expected to explain an increased effect of hair cortisol levels and salivary cortisol response on the internalizing symptoms and the p-factor. Lower BPS is expected to be a protective factor on the negative impact of stress on psychopathology, which could represent a mechanism of resilience. Based on our preliminary work (see above), we expect an attenuated downregulation of salivary cortisol in response to the social stress test (ie, inefficient acute stress regulation) to be associated with higher internalizing symptoms and a higher p-factor. We expect this effect to be mediated by BPS as an index of LC-NE tonic activity.

Risk Assessment

Childhood adversity might not be disclosed by caregivers if they are also the perpetrators, while child reports on adversity might be unreliable. To mitigate the underreporting biases, we will implement the multiple-sources approach to adversity. In addition, any reports on adversity will be evaluated according to our in-house reporting guidelines and may lead to further actions.

Work Package 4 (Objective 4): Development of an Objective Tool of Risk Prediction in the Preschool Age Based on the Pupillometric Measures of LC-NE Functioning

Participants

The full sample was recruited during work packages 2 and 3.

Sample Size

The full sample consists of the 300 participants assessed in task 2.2.

Experimental Design

The findings of Objectives 2 and 3 will be used to enrich a probabilistic model in the full sample (n=300) that assigns individual risk based on LC-NE functioning in the preschool age. This will be evaluated as a tool of quantitative risk prediction for future psychopathology. We will operationalize risk probability within a normative model of transdiagnostic psychopathology [105]. A reference cohort will be defined from a subsample of participants that did not develop clinically relevant transdiagnostic psychopathology. This cohort will be applied to estimate a normative development model (ie, output) as a function of LC-NE tonic and phasic activity profiles. It

XSL•FO
RenderX

defines a range of LC-NE functioning variation that is associated with a low risk. The model can then be applied to withheld data, that is, individuals who did develop transdiagnostic psychopathology. This allows one to quantify a numeric deviation of LC-NE functioning from normative ranges. The numeric deviation is then applied as a predictor of psychopathology in linear models that represent an individual risk prediction. The performance of individual risk prediction based on LC-NE functioning will be compared to the predictive validity of available covariates, confirmed by v-fold validation, and implemented in the outpatient clinic for external validation in independent data.

Expected Outcome

We expect LC-NE functioning to provide added predictive validity, with an increase in explained variance compared to previous psychopathology, childhood adversity, and cognitive ability. We assume a moderate accuracy of the normative model and substantial associations of normative range deviations with transdiagnostic psychopathology. We further expect that the spatial direction of the range deviation within the normative model differentiates the factors of transdiagnostic psychopathology. The simplicity of the model emphasizes its feasibility for use in clinical settings. The final task of LOCUS-MENTAL will provide a tool for quantitative risk prediction in at-risk populations of preschoolers.

Risk Assessment

We acknowledge that LC-NE functioning is not the only determinant in the complex development of psychopathology, but given this, the time is ripe to test a promising neurophysiological mechanism as an objective predictor that improves the assessment of clinical risk over the current practice of subjective evaluation.

Ethical Considerations

The study was approved by the Ethical Committee (reference number: 2024 - 2160) and the Data Protection Officer of the Medical Faculty of Goethe University Frankfurt. To date, no adverse events have been reported by any comparable study, and thus we expect no risks nor harm to our participants. The study is planned and will be conducted and analyzed according to Good Scientific Practice (GSP) guidelines, the Declaration of Helsinki, and the European Data Protection Directive (DSGVO). Study participation is on a voluntary basis, and participants can withdraw from study participation at any time without any adverse consequences concerning their medical or psychiatric treatment in our clinic. Participants will also receive monetary compensation for their time. All relevant information regarding data protection and the voluntary basis of participation will be communicated to participants and their legal caregivers verbally, as well as in the form of written informed consent forms. We also obtain verbal assent from the preschool participants. The child's right to withdraw at any time will be prioritized over parental consent. During assessments, the child will be closely monitored by the experimenter that is present at all times, and signs of acute distress (such as crying, physical withdrawal, or verbal expression of a desire to stop) will lead to termination of the assessment. In cases of suspected or disclosed adversity, the research team will inform the principal investigator immediately (licensed psychotherapist for children), who offers a voluntary counseling appointment to discuss potential adversity with the family and the child. In line with German child protection laws and in accordance with our clinical guidelines, the quantity and quality of suspected adversity might result in further support by our department or even a report to authorities (eg, imminent danger of ongoing adversity). We acknowledge that the development of a risk assessment tool could potentially lead to stigmatization and misuse. With our data management plan outlined below, we made sure that individual risk prediction cannot be associated with personally identifiable information. We further refrain from any nonscientific use of the prognostic risk marker. The rights to the information of prognostic risk remain with the participants in accordance with DSGVO. We are confident that a prognostic risk marker also provides the opportunity to shift from a health care system that focuses on intervention after manifestation of psychopathology to a focus on targeted prevention. Children at an elevated risk for mental health problems as identified by LOCUS-MENTAL could receive further case management at our outpatient clinic, where we decide on useful next steps on a case-by-case basis with evidence-based treatments. For example, caregivers of children at risk could be offered the opportunity to receive parent trainings that are established in the treatment of mental health problems in children. This may aid in the early prevention of the initial manifestation of psychopathology. Our newly established outpatient clinic for early detection could be the ideal environment to use risk assessment information for positive clinical outcomes.

LOCUS-MENTAL applies a noninvasive neuroimaging methodology that allows indexing neuronal activity by video-based pupillometry. Previous work showed feasibility and established this methodology in clinical samples of children. In addition, we apply further noninvasive assessments with electroencephalography. These measures introduce a burden for the participants and can induce a state of irritability but do no harm. This also applies to the application of an age-appropriate social stress test. The test induces a temporary impression that the child's performance was insufficient in a task, which has been associated with a cortisol release response that needs to be quantified to achieve the objective of task 3.2. The social stress test thus includes a comprehensive disclosure strategy after the test to alleviate any negative effects. In task 3.2, we further collect tissue as hair samples (cutting off hair strands) and saliva samples (chewing an artificial fiber). These methods of tissue collection are also entirely non-invasive. The tissue samples will be processed by experienced contractors and subsequently destroyed. In the project, we further assess personally identifiable information that is handled according to the descriptions in section 2.4. We expect that this study does not involve an immediate risk of yielding knowledge, products, or technology that could intentionally be misused to cause substantial harm.

Questionnaire and interview data will be collected via our on-premises online tool according to the European General Data Protection Regulation (GDPR). We further do video-based eye tracking, apply EEG, and collect biological samples. Hair

XSL•FO
RenderX

samples will be analyzed by Dresden Labservice with online solid-phase extraction in liquid chromatography-tandem mass spectrometry. Saliva samples will be analyzed by Daacro Contract Research with the ELISA Salimetrics Assay in a double estimation. Metadata will be generated by a printed case report form. Data are transferred to electronic databases that allow for automated quality controls and sanity checks. Eye tracking and EEG data will also receive quality control by peer-reviewed preprocessing pipelines. We established standard operating procedures to document research data, personnel will be trained in GSP, and all data handling will be in accordance with the GDPR. We differentiate between person-identifying information and research data that are separated in different networks, both behind the clinic's firewall. The data can only be combined by a person identification code that is restricted in access to authorized personnel and stored on the clinical network. This is separated from a pseudonymized research network. All databases run on a Microsoft Access front- and back-end with an Active Directory system. Data are stored on on-premises servers with backups in a different building. After the project's end, anonymized research data will be made available according to Open Science standards and archived for 10 years. The author (IT) will administer the project-specific databases. PhD students will have read and write rights with log protocols. A dedicated database manager with a permanent position is the administrator of all databases.

The LOCUS-MENTAL study uses targeted strategies across its work packages to manage missing data and ensure robust longitudinal findings. Work package 1 focuses on feasibility measures by using passive and child-friendly paradigms to achieve a projected completion rate of over 80%. Statistical analysis will exclude children that have per-task data validity rates of less than 50% or more than one missing task. The auditory oddball task is a marker task that is required to be included in the statistical analysis. We established standardized procedures of pupillometry data preprocessing to address within-trial missing data due to inattention. Work package 2 addresses high-burden EEG assessments by designating them as "more optional" in the context of the overall project, which allows for an expected 30% attrition rate without hindering subsequent study objectives. For the longitudinal assessments in work package 3, the study implements retention strategies like online data collection and locator documents. In addition, the statistical framework of work package 3 includes LMM and structural equation modeling to handle intermittent missingness across time points. Questionnaire data on an item level with per-measure missing rates up to 20% will be addressed by Multiple Imputation by Chained Equations (MICE) that preserves statistical power under the missing at random assumptions. However, children with increased irritability or increased burden may be more likely to drop out (ie, missing not at random or MNAR). Thus, pattern mixture models will be applied as a sensitivity analysis to determine if the developmental trajectories of LC-NE functioning are biased by specific attrition patterns. In addition, Inverse Probability Weighting (IPW) in the longitudinal data analysis can be used to adjust for non-random attrition by weighting the remaining participants based on baseline characteristics (eg, socioeconomic status or initial psychopathology scores) that predict dropout. Missing bio samples will be addressed by case-wise exclusion in the specific subanalysis. Finally, work package 4 uses v-fold validation to evaluate the predictive stability of the risk prediction tool.

Results

The project was launched upon funding in November 2024. The validation phase (work package 1) began participant enrollment in August 2025. As of February 2026, 82 participants have been assessed. Recruitment is ongoing, with a target of 90 participants by the end of February 2026. Data collection for the validation phase will be followed by comprehensive analysis of the psychophysiological and behavioral data. Results from this phase are intended for publication in 2026. Based on the validated paradigms, the core accelerated longitudinal study (work package 2) is scheduled to commence in April 2026. This longitudinal component will follow a cohort of 300 preschool-aged children across 3 annual assessment waves till March 2029. The entire LOCUS-MENTAL project is designed as a 6-year study. Final data analysis, synthesis of findings, and dissemination for the full longitudinal cohort are planned to conclude by the end of 2030.

Discussion

Principal Findings

The LOCUS-MENTAL protocol outlines a longitudinal study design with the primary aim to establish LC-NE functioning as a neurobiological predictor of later psychopathology in childhood. We anticipate that the validation phase (Objective 1) will demonstrate convergent and discriminant validity of the pupillometric markers BPS and SEPR with sufficient reliability. By developing and validating a child-friendly, pupillometry-based tool, this aims to bridge a critical gap between neurobiological mechanism and clinical risk prediction, which is based solely on behavioral observation. Accordingly, this is a prerequisite to ensure the method sensitivity and specificity. This would provide a concise test battery that assesses a neurobiological mechanism that could be implemented in future clinical studies capturing a potential biomarker of psychopathology in early childhood. Furthermore, we hypothesize that BPS and SEPR will cross-sectionally differentiate between transdiagnostic risk profiles, such as irritability and behavioral inhibition, and serve as a mediator between childhood adversity and psychopathology (Objective 2). This would provide an underlying neurobiological basis for temperament dimensions that are also based on caregiver reports. In the longitudinal data, as a core outcome, we expect to reveal baseline LC-NE functioning at age 4 - 6 years to predict the development of psychopathology at school age (Objective 3). This central component of the LOCUS-MENTAL project would provide causative evidence for an impact of LC-NE functioning on transdiagnostic mental health conditions. These findings will be used in forming a normative model as a pupillometry-based risk prediction tool, establishing a key neurobiological susceptibility mechanism that mediates the impact of early adversity on psychopathology (Objective 4). This will allow a translation of the experimental findings to clinical practice. For

XSL•FO
RenderX

example, children associated with an elevated risk for psychopathology based on LC-NE functioning could be supported by evidence-based treatments that are effective before the onset of a mental health condition. This would emphasize a targeted prevention approach. We conclude that the identification of a neurobiologically grounded risk marker could transform our mental health infrastructure from interventions that are applied years after the onset of mental health problems to early and targeted support. Ultimately, this aims to reduce the lifetime burden of mental health problems.

Strengths

LOCUS-MENTAL focuses on a critical developmental window, the preschool period (aged 4 - 6 years), a key time for the emergence of transdiagnostic risk factors, offering a unique opportunity for targeted prevention. This developmental period has been largely neglected in experimental research on lifetime mental health, as preschoolers have limited capabilities to follow task instructions. In this project, this is addressed by the application of feasible pupillometry as a primary method that is feasible in young and burdened children. The study provides a robust design, combining an initial multitrait-multimethod validation phase with a subsequent large-scale accelerated longitudinal study. This design strengthens internal validity through paradigm optimization and allows for examination of causal pathways and developmental trajectories over time.

Limitations

The novelty of the paradigms for preschoolers is a limitation since established pupillometric paradigms for assessing specifically LC-NE functions have not been established in young children. The dedicated validation phase (working package 1) is specifically designed to address this issue. Another limiting factor is the longitudinal study design. This design comes with the risk of high participants' attrition, which may affect the final statistical power. To address this, the study will use retention strategies (eg, regular contact, family-friendly appointment scheduling, and compensation) and will use appropriate statistical methods to handle missing data.

Future Directions

If the pupillometric markers are validated as successful predictors, future research could focus on clinical implementation. The natural next step would be the development of a pupillometry tool that can be applied outside the lab, for which head-mounted devices with built-in eye trackers are a promising avenue. Ultimately, this would deliver a child-friendly and low-cost automated screening tool for pediatricians or for educational settings.

Dissemination Plan

The results of the study will be disseminated through peer-reviewed publications in high-impact journals. Findings will be presented at national and international conferences. Following Open Science principles, data analysis will be made available in online repositories, while deidentified data can be shared upon reasonable request in line with data privacy considerations. In addition, we engage in active patient and public involvement by media outreach and promoting our research on social media and our Mind-Child lab web page.

Funding

This work was supported by Emmy Noether Program grant number [German Research Foundation reference number: 548343771]. Preliminary work was funded by the German Research Foundation (reference number: 492582254), the Daimler and Benz Foundation (reference number: 32-01/22) and funds by the medical faculty of the University Hospital Frankfurt for outstanding early career researchers (F38/21, R139/2021). The ongoing project synergized and is supported by the DYNAMIC center, which is funded by the LOEWE program of the Hessian Ministry of Science and Arts (grant LOEWE1/16/519/03/09.001[0009]/98).

Disclaimer

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to generative AI tools under full human supervision:

- Proofreading and editing

The generative AI tool used was: Google AI Studio.

Responsibility for the final manuscript lies entirely with the authors.

Generative AI tools are not listed as authors and do not bear responsibility for the final outcomes.

Declaration submitted by: IT

Additional note: Google AI Studio was used to assist with proofreading for improving reading comprehension only, but none of the prompts, responses, or conversation transcripts were saved or recorded.

Data Availability

The data set generated during this study will be made available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: NB

XSL•FO
RenderX

Funding acquisition: NB
Investigation: IT, PRN
Methodology: NB
Project administration: IT, PRN, NB
Resources: NB
Software: IT, PRN, NB
Supervision: NB
Visualization: IT, PRN, NB
Writing – original draft: IT
Writing – review & editing: PRN, NB

Conflicts of Interest

NB receives royalties for lecturing at institutes for training in psychotherapy and psychiatric hospitals and receives research funding by the German Research Foundation (DFG), the Johanna Quant Young Academy (JQYA), and the Daimler & Benz Foundation.

Checklist 1
Completed STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist for the LOCUS-MENTAL longitudinal study protocol.
[PDF File, 178 KB - resprot_v15i1e89237_app1.pdf]

Peer Review Report 1
Peer-review report by Emmy-Noether Program Review Panel, Excellence Research Funds of the German Research Foundation (DFG; Germany).
[PDF File, 95 KB - resprot_v15i1e89237_app2.pdf]

References

1. Bast N, Boxhoorn S, Supér H, et al. Atypical arousal regulation in children with autism but not with attention-deficit/hyperactivity disorder as indicated by pupillometric measures of locus coeruleus activity. Biol Psychiatry Cogn Neurosci Neuroimaging 2023 Jan;8(1):11-20. [doi: 10.1016/j.bpsc.2021.04.010] [Medline: 33930603]
2. Grueschow M, Stenz N, Thörn H, et al. Real-world stress resilience is associated with the responsivity of the locus coeruleus. Nat Commun 2021 Apr 15;12(1):2275. [doi: 10.1038/s41467-021-22509-1] [Medline: 33859187]
3. Llorca-Torralba M, Camarena-Delgado C, Suárez-Pereira I, et al. Pain and depression comorbidity causes asymmetric plasticity in the locus coeruleus neurons. Brain (Bacau) 2022 Mar 29;145(1):154-167. [doi: 10.1093/brain/awab239] [Medline: 34373893]
4. Bast N, Ahmad J, Mason L, et al. Locus coeruleus tonic upregulation increases selectivity to inconspicuous auditory information in autistic compared to non-autistic individuals: a combined pupillometry and electroencephalography study. Mol Autism 2025 Aug 21;16(1):41. [doi: 10.1186/s13229-025-00678-w] [Medline: 40841687]
5. Bast N, Polzer L, Raji N, et al. Early intervention increases reactive joint attention in autistic preschoolers with arousal regulation as mediator. Eur Child Adolesc Psychiatry 2025 Oct;34(10):3161-3174. [doi: 10.1007/s00787-025-02738-1] [Medline: 40347242]
6. Caspi A, Houts RM, Ambler A, et al. Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin birth cohort study. JAMA Netw Open 2020 Apr 1;3(4):e203221. [doi: 10.1001/jamanetworkopen.2020.3221] [Medline: 32315069]
7. Weisz JR, Kuppens S, Ng MY, et al. What five decades of research tells us about the effects of youth psychological therapy: a multilevel meta-analysis and implications for science and practice. Am Psychol 2017;72(2):79-117. [doi: 10.1037/a0040360] [Medline: 28221063]
8. Lahey BB, Krueger RF, Rathouz PJ, Waldman ID, Zald DH. A hierarchical causal taxonomy of psychopathology across the life span. Psychol Bull 2017 Feb;143(2):142-186. [doi: 10.1037/bul0000069] [Medline: 28004947]
9. Karsten J, Hartman CA, Smit JH, et al. Psychiatric history and subthreshold symptoms as predictors of the occurrence of depressive or anxiety disorder within 2 years. Br J Psychiatry 2011 Mar;198(3):206-212. [doi: 10.1192/bjp.bp.110.080572] [Medline: 21357879]
10. Delgadillo J, Lutz W. A development pathway towards precision mental health care. JAMA Psychiatry 2020 Sep 1;77(9):889-890. [doi: 10.1001/jamapsychiatry.2020.1048] [Medline: 32459326]
11. Thapar A, Riglin L. The importance of a developmental perspective in psychiatry: what do recent genetic-epidemiological findings show? Mol Psychiatry 2020 Aug;25(8):1631-1639. [doi: 10.1038/s41380-020-0648-1] [Medline: 31959848]

XSL•FO
RenderX

12. Vasileva M, Graf RK, Reinelt T, Petermann U, Petermann F. Research review: a meta - analysis of the international prevalence and comorbidity of mental disorders in children between 1 and 7 years. Child Psychology Psychiatry 2021 Apr;62(4):372-381 [FREE Full text] [doi: 10.1111/jcpp.13261]
13. Gee DG. Early adversity and development: parsing heterogeneity and identifying pathways of risk and resilience. AJP 2021 Nov;178(11):998-1013. [doi: 10.1176/appi.ajp.2021.21090944]
14. Slagt M, Dubas JS, Deković M, van Aken MAG. Differences in sensitivity to parenting depending on child temperament: a meta-analysis. Psychol Bull 2016 Oct;142(10):1068-1110. [doi: 10.1037/bul0000061] [Medline: 27513919]
15. Ioannidis K, Askelund AD, Kievit RA, van Harmelen AL. The complex neurobiology of resilient functioning after childhood maltreatment. BMC Med 2020 Feb 13;18(1):32. [doi: 10.1186/s12916-020-1490-7] [Medline: 32050974]
16. Poe GR, Foote S, Eschenko O, et al. Locus coeruleus: a new look at the blue spot. Nat Rev Neurosci 2020 Nov;21(11):644-659. [doi: 10.1038/s41583-020-0360-9] [Medline: 32943779]
17. Liebe T, Kaufmann J, Li M, Skalej M, Wagner G, Walter M. In vivo anatomical mapping of human locus coeruleus functional connectivity at 3 T MRI. Hum Brain Mapp 2020 Jun 1;41(8):2136-2151. [doi: 10.1002/hbm.24935] [Medline: 31994319]
18. Zhang S, Hu S, Chao HH, Li CSR. Resting-state functional connectivity of the Locus Coeruleus in humans: in comparison with the Ventral Tegmental Area/Substantia Nigra Pars Compacta and the effects of age. Cereb Cortex 2016 Aug;26(8):3413-3427. [doi: 10.1093/cercor/bhv172] [Medline: 26223261]
19. Aston-Jones G, Cohen JD. AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: adaptive gain and optimal performance. Annu Rev Neurosci 2005;28(1):403-450. [doi: 10.1146/annurev.neuro.28.061604.135709] [Medline: 16022602]
20. Hayat H, Regev N, Matosevich N, et al. Locus coeruleus norepinephrine activity mediates sensory-evoked awakenings from sleep. Sci Adv 2020 Apr;6(15):eaaz4232. [doi: 10.1126/sciadv.aaz4232] [Medline: 32285002]
21. Minzenberg MJ, Watrous AJ, Yoon JH, Ursu S, Carter CS. Modafinil shifts human locus coeruleus to low-tonic, high-phasic activity during functional MRI. Science 2008 Dec 12;322(5908):1700-1702. [doi: 10.1126/science.1164908] [Medline: 19074351]
22. Petersen SE, Posner MI. The attention system of the human brain: 20 years after. Annu Rev Neurosci 2012;35(1):73-89. [doi: 10.1146/annurev-neuro-062111-150525] [Medline: 22524787]
23. Dahl MJ, Mather M, Werkle-Bergner M. Noradrenergic modulation of rhythmic neural activity shapes selective attention. Trends Cogn Sci (Regul Ed) 2022 Jan;26(1):38-52. [doi: 10.1016/j.tics.2021.10.009]
24. Zhao S, Chait M, Dick F, Dayan P, Furukawa S, Liao HI. Pupil-linked phasic arousal evoked by violation but not emergence of regularity within rapid sound sequences. Nat Commun 2019 Sep 6;10(1):4030. [doi: 10.1038/s41467-019-12048-1] [Medline: 31492881]
25. Mather M, Clewett D, Sakaki M, Harley CW. Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. Behav Brain Sci 2016 Jan;39:e200. [doi: 10.1017/S0140525X15000667] [Medline: 26126507]
26. Pfeffer T, Keitel C, Kluger DS, et al. Coupling of pupil- and neuronal population dynamics reveals diverse influences of arousal on cortical processing. Elife 2022 Feb 8;11:e71890. [doi: 10.7554/eLife.71890] [Medline: 35133276]
27. Vazey EM, Moorman DE, Aston-Jones G. Phasic locus coeruleus activity regulates cortical encoding of salience information. Proc Natl Acad Sci U S A 2018 Oct 2;115(40):E9439-E9448. [doi: 10.1073/pnas.1803716115] [Medline: 30232259]
28. Rodenkirch C, Liu Y, Schriver BJ, Wang Q. Locus coeruleus activation enhances thalamic feature selectivity via norepinephrine regulation of intrathalamic circuit dynamics. Nat Neurosci 2019 Jan;22(1):120-133. [doi: 10.1038/s41593-018-0283-1] [Medline: 30559472]
29. O'Callaghan C. Noradrenaline drives learning across scales of time and neurobiological organisation. Trends Cogn Sci (Regul Ed) 2026 Mar;30(3):264-278. [doi: 10.1016/j.tics.2025.07.003]
30. Samuels E, Szabadi E. Functional Neuroanatomy of the Noradrenergic Locus Coeruleus: Its Roles in the Regulation of Arousal and Autonomic Function Part I: Principles of Functional Organisation. CN 2008:235-253. [doi: 10.2174/157015908785777229]
31. Kahneman D, Beatty J. Pupil diameter and load on memory. Science 1966 Dec 23;154(3756):1583-1585. [doi: 10.1126/science.154.3756.1583] [Medline: 5924930]
32. Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI. Rational regulation of learning dynamics by pupil-linked arousal systems. Nat Neurosci 2012 Jun 3;15(7):1040-1046. [doi: 10.1038/nn.3130] [Medline: 22660479]
33. Murphy PR, O'Connell RG, O'Sullivan M, Robertson IH, Balsters JH. Pupil diameter covaries with BOLD activity in human locus coeruleus. Hum Brain Mapp 2014 Aug;35(8):4140-4154. [doi: 10.1002/hbm.22466] [Medline: 24510607]
34. Joshi S, Li Y, Kalwani RM, Gold JI. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron 2016 Jan 6;89(1):221-234. [doi: 10.1016/j.neuron.2015.11.028] [Medline: 26711118]
35. Joshi S, Gold JI. Pupil Size as a Window on Neural Substrates of Cognition. Trends Cogn Sci (Regul Ed) 2020 Jun;24(6):466-480. [doi: 10.1016/j.tics.2020.03.005]
36. Megemont M, McBurney-Lin J, Yang H. Pupil diameter is not an accurate real-time readout of locus coeruleus activity. Elife 2022 Feb 2;11:e70510. [doi: 10.7554/eLife.70510] [Medline: 35107419]

XSL•FO
RenderX

37. Reilly J, Kelly A, Kim SH, Jett S, Zuckerman B. The human task-evoked pupillary response function is linear: Implications for baseline response scaling in pupillometry. Behav Res 2019 Apr;51(2):865-878. [doi: 10.3758/s13428-018-1134-4]
38. Liu KY, Marijatta F, Hämmerer D, Acosta-Cabronero J, Düzel E, Howard RJ. Magnetic resonance imaging of the human locus coeruleus: a systematic review. Neurosci Biobehav Rev 2017 Dec;83:325-355. [doi: 10.1016/j.neubiorev.2017.10.023]
39. Monroe SM, Simons AD. Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. Psychol Bull 1991 Nov;110(3):406-425. [doi: 10.1037/0033-2909.110.3.406] [Medline: 1758917]
40. Finkelhor D, Shattuck A, Turner H, Hamby S. Improving the adverse childhood experiences study scale. JAMA Pediatr 2013 Jan;167(1):70-75. [doi: 10.1001/jamapediatrics.2013.420] [Medline: 23403625]
41. Rothbart MK, Bates JE. Temperament. In: Eisenberg N, Damon W, Lerner RM, editors. Handbook of Child Psychology: Social, Emotional, and Personality Development Perament, 6th edition: John Wiley and Sons, Inc; 2006:99-166.
42. Pat N, Riglin L, Anney R, et al. Motivation and cognitive abilities as mediators between polygenic scores and psychopathology in children. J Am Acad Child Adolesc Psychiatry 2022 Jun;61(6):782-795. [doi: 10.1016/j.jaac.2021.08.019] [Medline: 34506929]
43. Mincic AM. Neuroanatomical correlates of negative emotionality-related traits: a systematic review and meta-analysis. Neuropsychologia 2015 Oct;77:97-118. [doi: 10.1016/j.neuropsychologia.2015.08.007] [Medline: 26265397]
44. Zerbi V, Floriou-Servou A, Markicevic M, et al. Rapid reconfiguration of the functional connectome after chemogenetic locus coeruleus activation. Neuron 2019 Aug 21;103(4):702-718. [doi: 10.1016/j.neuron.2019.05.034] [Medline: 31227310]
45. Leibenluft E. Pediatric irritability: a systems neuroscience approach. Trends Cogn Sci (Regul Ed) 2017 Apr;21(4):277-289. [doi: 10.1016/j.tics.2017.02.002]
46. Evans SC, Corteselli KA, Edelman A, Scott H, Weisz JR. Is irritability a top problem in youth mental health care? A multi-informant, multi-method investigation. Child Psychiatry Hum Dev 2023 Aug;54(4):1027-1041. [doi: 10.1007/s10578-021-01301-8] [Medline: 35064392]
47. Ezpeleta L, Granero R, de la Osa N, Trepat E, Domènech JM. Trajectories of oppositional defiant disorder irritability symptoms in preschool children. J Abnorm Child Psychol 2016 Jan;44(1):115-128. [doi: 10.1007/s10802-015-9972-3] [Medline: 25630822]
48. Riglin L, Eyre O, Cooper M, et al. Investigating the genetic underpinnings of early-life irritability. Transl Psychiatry 2017 Sep 26;7(9):e1241-e1241. [doi: 10.1038/tp.2017.212] [Medline: 28949337]
49. Vidal-Ribas P, Brotman MA, Valdivieso I, Leibenluft E, Stringaris A. The status of irritability in psychiatry: a conceptual and quantitative review. J Am Acad Child Adolesc Psychiatry 2016 Jul;55(7):556-570. [doi: 10.1016/j.jaac.2016.04.014]
50. Kessel EM, Meyer A, Hajcak G, et al. Transdiagnostic factors and pathways to multifinality: the error-related negativity predicts whether preschool irritability is associated with internalizing versus externalizing symptoms at age 9. Dev Psychopathol 2016 Nov;28(4pt1):913-926. [doi: 10.1017/S0954579416000626] [Medline: 27739383]
51. Grueschow M, Kleim B, Ruff CC. Role of the locus coeruleus arousal system in cognitive control. J Neuroendocrinology 2020 Dec;32(12):e12890. [doi: 10.1111/jne.12890]
52. Sears RM, Fink AE, Wigestrand MB, Farb CR, de Lecea L, Ledoux JE. Orexin/hypocretin system modulates amygdala-dependent threat learning through the locus coeruleus. Proc Natl Acad Sci U S A 2013 Dec 10;110(50):20260-20265. [doi: 10.1073/pnas.1320325110] [Medline: 24277819]
53. Tang A, Crawford H, Morales S, Degnan KA, Pine DS, Fox NA. Infant behavioral inhibition predicts personality and social outcomes three decades later. Proc Natl Acad Sci USA 2020 May 5;117(18):9800-9807. [doi: 10.1073/pnas.1917376117]
54. Lahat A, Lamm C, Chronis-Tuscano A, Pine DS, Henderson HA, Fox NA. Early behavioral inhibition and increased error monitoring predict later social phobia symptoms in childhood. J Am Acad Child Adolesc Psychiatry 2014 Apr;53(4):447-455. [doi: 10.1016/j.jaac.2013.12.019]
55. Reeb‐Sutherland BC, Vanderwert RE, Degnan KA, et al. Attention to novelty in behaviorally inhibited adolescents moderates risk for anxiety. Child Psychology Psychiatry 2009 Nov;50(11):1365-1372. [doi: 10.1111/j.1469-7610.2009.02170.x]
56. Fox NA, Buzzell GA, Morales S, Valadez EA, Wilson M, Henderson HA. Understanding the emergence of social anxiety in children with behavioral inhibition. Biol Psychiatry 2021 Apr 1;89(7):681-689. [doi: 10.1016/j.biopsych.2020.10.004] [Medline: 33353668]
57. Russell G, Lightman S. The human stress response. Nat Rev Endocrinol 2019 Sep;15(9):525-534. [doi: 10.1038/s41574-019-0228-0] [Medline: 31249398]
58. McEwen BS, Akil H. Revisiting the stress concept: implications for affective disorders. J Neurosci 2020 Jan 2;40(1):12-21. [doi: 10.1523/JNEUROSCI.0733-19.2019] [Medline: 31896560]
59. Snyder HR, Young JF, Hankin BL. Chronic stress exposure and generation are related to the P-factor and externalizing specific psychopathology in youth. J Am Acad Child Adolesc Psychiatry 2019 Mar 4;48(2):306-315. [doi: 10.1080/15374416.2017.1321002]
60. Devilbiss DM, Waterhouse BD, Berridge CW, Valentino R. Corticotropin-releasing factor acting at the locus coeruleus disrupts thalamic and cortical sensory-evoked responses. Neuropsychopharmacol 2012 Aug;37(9):2020-2030. [doi: 10.1038/npp.2012.50]

XSL•FO
RenderX

61. Godoy LD, Rossignoli MT, Delfino-Pereira P, Garcia-Cairasco N, de Lima Umeoka EH. A comprehensive overview on stress neurobiology: basic concepts and clinical implications. Front Behav Neurosci 2018;12:127. [doi: 10.3389/fnbeh.2018.00127] [Medline: 30034327]
62. McCall JG, Al-Hasani R, Siuda ER, et al. CRH engagement of the locus coeruleus noradrenergic system mediates stress-induced anxiety. Neuron 2015 Aug 5;87(3):605-620. [doi: 10.1016/j.neuron.2015.07.002] [Medline: 26212712]
63. Devilbiss DM. Consequences of tuning network function by tonic and phasic locus coeruleus output and stress: regulating detection and discrimination of peripheral stimuli. Brain Res 2019 Apr 15;1709:16-27. [doi: 10.1016/j.brainres.2018.06.015] [Medline: 29908165]
64. Hermans EJ, van Marle HJF, Ossewaarde L, et al. Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. Science 2011 Nov 25;334(6059):1151-1153. [doi: 10.1126/science.1209603] [Medline: 22116887]
65. Isingrini E, Perret L, Rainer Q, et al. Resilience to chronic stress is mediated by noradrenergic regulation of dopamine neurons. Nat Neurosci 2016 Apr;19(4):560-563. [doi: 10.1038/nn.4245] [Medline: 26878672]
66. Morris LS, McCall JG, Charney DS, Murrough JW. The role of the locus coeruleus in the generation of pathological anxiety. Brain Neurosci Adv 2020;4:2398212820930321. [doi: 10.1177/2398212820930321] [Medline: 32954002]
67. Méndez-Bértolo C, Moratti S, Toledano R, et al. A fast pathway for fear in human amygdala. Nat Neurosci 2016 Aug;19(8):1041-1049. [doi: 10.1038/nn.4324] [Medline: 27294508]
68. Fast CD, McGann JP. Amygdalar gating of early sensory processing through interactions with locus coeruleus. J Neurosci 2017 Mar 15;37(11):3085-3101. [doi: 10.1523/JNEUROSCI.2797-16.2017] [Medline: 28188216]
69. Hermans EJ, Henckens M, Joëls M, Fernández G. Dynamic adaptation of large-scale brain networks in response to acute stressors. Trends Neurosci 2014 Jun;37(6):304-314. [doi: 10.1016/j.tins.2014.03.006] [Medline: 24766931]
70. Giustino TF, Ramanathan KR, Totty MS, Miles OW, Maren S. Locus coeruleus norepinephrine drives stress-induced increases in basolateral amygdala firing and impairs extinction learning. J Neurosci 2020 Jan 22;40(4):907-916. [doi: 10.1523/JNEUROSCI.1092-19.2019] [Medline: 31801809]
71. Naegeli C, Zeffiro T, Piccirelli M, et al. Locus coeruleus activity mediates hyperresponsiveness in posttraumatic stress disorder. Biol Psychiatry 2018 Feb 1;83(3):254-262. [doi: 10.1016/j.biopsych.2017.08.021] [Medline: 29100627]
72. Ross JA, Van Bockstaele EJ. The locus coeruleus- norepinephrine system in stress and arousal: unraveling historical, current, and future perspectives. Front Psychiatry 2020;11:601519. [doi: 10.3389/fpsyt.2020.601519] [Medline: 33584368]
73. Burley DT, van Goozen SHM. Pupil response to affective stimuli: a biomarker of early conduct problems in young children. J Abnorm Child Psychol 2020 May;48(5):693-701. [doi: 10.1007/s10802-020-00620-z] [Medline: 31982978]
74. Bast N, Poustka L, Freitag CM. The locus coeruleus–norepinephrine system as pacemaker of attention – a developmental mechanism of derailed attentional function in autism spectrum disorder. Eur J of Neuroscience 2018 Jan;47(2):115-125. [doi: 10.1111/ejn.13795]
75. Boxhoorn S, Bast N, Supèr H, Polzer L, Cholemkery H, Freitag CM. Pupil dilation during visuospatial orienting differentiates between autism spectrum disorder and attention‐deficit/hyperactivity disorder. Child Psychology Psychiatry 2020 May;61(5):614-624. [doi: 10.1111/jcpp.13179]
76. Bast N, Gaigg SB, Bowler DM, Roessner V, Freitag CM, Ring M. Arousal-modulated memory encoding and retrieval in adults with autism spectrum disorder. Autism Res 2022 Sep;15(9):1609-1620. [doi: 10.1002/aur.2784] [Medline: 35906845]
77. Bast N, Mason L, Ecker C, et al. Sensory salience processing moderates attenuated gazes on faces in autism spectrum disorder: a case-control study. Mol Autism 2023 Feb 9;14(1):5. [doi: 10.1186/s13229-023-00537-6] [Medline: 36759875]
78. Loth E, Charman T, Mason L, et al. The EU-AIMS Longitudinal European Autism Project (LEAP): design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. Mol Autism 2017;8(1):24. [doi: 10.1186/s13229-017-0146-8] [Medline: 28649312]
79. Polzer L, Freitag CM, Bast N. Pupillometric measures of altered stimulus-evoked locus coeruleus-norepinephrine activity explain attenuated social attention in preschoolers with autism spectrum disorder. Autism Res 2022 Nov;15(11):2167-2180. [doi: 10.1002/aur.2818] [Medline: 36111843]
80. Polzer L, Schenk M, Raji N, et al. Temporal progression of pupil dilation and gaze behavior to emotion expressions in preschoolers with autism spectrum disorder. Sci Rep 2024 Apr 3;14(1):7843. [doi: 10.1038/s41598-024-58480-2] [Medline: 38570565]
81. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
82. Schmeck K, Poustka F, Döpfner M, et al. Discriminant validity of the Child Behaviour Checklist CBCL-4/18 in German samples. Eur Child Adolesc Psychiatry 2001 Dec 1;10(4):240-247. [doi: 10.1007/s007870170013]
83. Bishop G, Spence SH, McDonald C. Can parents and teachers provide a reliable and valid report of behavioral inhibition? Child Dev 2003;74(6):1899-1917. [doi: 10.1046/j.1467-8624.2003.00645.x] [Medline: 14669903]
84. Stringaris A, Goodman R, Ferdinando S, et al. The Affective Reactivity Index: a concise irritability scale for clinical and research settings. Child Psychology Psychiatry 2012 Nov;53(11):1109-1117 [FREE Full text] [doi: 10.1111/j.1469-7610.2012.02561.x]

XSL•FO
RenderX

85. Putnam SP, Rothbart MK. Development of short and very short forms of the Children's Behavior Questionnaire. J Pers Assess 2006 Aug;87(1):102-112. [doi: 10.1207/s15327752jpa8701_09] [Medline: 16856791]
86. Finkelhor D, Shattuck A, Turner H, Hamby S. A revised inventory of adverse childhood experiences. Child Abuse Negl 2015 Oct;48:13-21. [doi: 10.1016/j.chiabu.2015.07.011]
87. Lin M, Hirschfeld G, Margraf J. Brief form of the Perceived Social Support Questionnaire (F-SozU K-6): Validation, norms, and cross-cultural measurement invariance in the USA, Germany, Russia, and China. Psychol Assess 2019 May;31(5):609-621. [doi: 10.1037/pas0000686] [Medline: 30589275]
88. Coddington DR. Coddington life events scales (CLES). Multi-Health Systems. 1999. URL: http://www.mhs.com [accessed 2026-03-29]
89. Shahinfar A, Fox NA, Leavitt LA. Preschool children's exposure to violence: relation of behavior problems to parent and child reports. Am J Orthopsychiatry 2000 Jan;70(1):115-125. [doi: 10.1037/h0087690] [Medline: 10702856]
90. Mather M, Huang R, Clewett D, et al. Isometric exercise facilitates attention to salient events in women via the noradrenergic system. Neuroimage 2020 Apr 15;210:116560. [doi: 10.1016/j.neuroimage.2020.116560] [Medline: 31978545]
91. Kim Y, Kadlaskar G, Keehn RM, Keehn B. Measures of tonic and phasic activity of the locus coeruleus-norepinephrine system in children with autism spectrum disorder: an event-related potential and pupillometry study. Autism Res 2022 Dec;15(12):2250-2264. [doi: 10.1002/aur.2820] [Medline: 36164264]
92. Kleberg JL, Del Bianco T, Falck-Ytter T. How infants' arousal influences their visual search. Child Dev 2019 Jul;90(4):1413-1423. [doi: 10.1111/cdev.13198] [Medline: 30597547]
93. Wechsler D. The Wechsler Preschool and Primary Scale of Intelligence—Fourth Edition, Wechsler Intelligence Scale for Children— Fifth Edition, and Wechsler Individual Achievement Test—Third Edition, 1st edition: Hogrefe; 2018. URL: https://psycnet.apa.org/record/2018-36604-009 [accessed 2025-07-29]
94. Send TS, Bardtke S, Gilles M, et al. Stress reactivity in preschool-aged children: evaluation of a social stress paradigm and investigation of the impact of prenatal maternal stress. Psychoneuroendocrinology 2019 Mar;101:223-231. [doi: 10.1016/j.psyneuen.2018.11.002] [Medline: 30471571]
95. Salum GA, Mogg K, Bradley BP, et al. Association between irritability and bias in attention orienting to threat in children and adolescents. Child Psychology Psychiatry 2017 May;58(5):595-602. [doi: 10.1111/jcpp.12659]
96. Deveney CM, Connolly ME, Haring CT, et al. Neural mechanisms of frustration in chronically irritable children. Am J Psychiatry 2013 Oct;170(10):1186-1194. [doi: 10.1176/appi.ajp.2013.12070917]
97. Wang YA, Rhemtulla M. Power analysis for parameter estimation in structural equation modeling: a discussion and tutorial. Adv Methods Pract Psychol Sci 2021 Jan;4(1):2515245920918253. [doi: 10.1177/2515245920918253]
98. Green P, MacLeod CJ. SIMR: an R package for power analysis of generalized linear mixed models by simulation. In: Nakagawa S, editor. Methods Ecol Evol 2016 Apr;7(4):493-498 [FREE Full text] [doi: 10.1111/2041-210X.12504]
99. McDiarmid TA, Bernardos AC, Rankin CH. Habituation is altered in neuropsychiatric disorders—a comprehensive review with recommendations for experimental design and analysis. Neurosci Biobehav Rev 2017 Sep;80:286-305. [doi: 10.1016/j.neubiorev.2017.05.028]
100. Murphy PR, Robertson IH, Balsters JH, O'connell RG. Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. Psychophysiology 2011 Nov;48(11):1532-1543. [doi: 10.1111/j.1469-8986.2011.01226.x] [Medline: 21762458]
101. Pasion R, Barbosa F. ERN as a transdiagnostic marker of the internalizing-externalizing spectrum: A dissociable meta-analytic effect. Neurosci Biobehav Rev 2019 Aug;103:133-149. [doi: 10.1016/j.neubiorev.2019.06.013]
102. Teague S, Youssef GJ, Macdonald JA, et al. Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. BMC Med Res Methodol 2018 Nov 26;18(1):151. [doi: 10.1186/s12874-018-0586-7] [Medline: 30477443]
103. Khoury JE, Bosquet Enlow M, Plamondon A, Lyons-Ruth K. The association between adversity and hair cortisol levels in humans: a meta-analysis. Psychoneuroendocrinology 2019 May;103:104-117. [doi: 10.1016/j.psyneuen.2019.01.009] [Medline: 30682626]
104. Kryski KR, Smith HJ, Sheikh HI, Singh SM, Hayden EP. Assessing stress reactivity indexed via salivary cortisol in preschool-aged children. Psychoneuroendocrinology 2011 Sep;36(8):1127-1136. [doi: 10.1016/j.psyneuen.2011.02.003] [Medline: 21388745]
105. Marquand AF, Kia SM, Zabihi M, Wolfers T, Buitelaar JK, Beckmann CF. Conceptualizing mental disorders as deviations from normative functioning. Mol Psychiatry 2019 Oct;24(10):1415-1424. [doi: 10.1038/s41380-019-0441-1] [Medline: 31201374]

Abbreviations

ACC: anterior cingulate cortex
AGT: Adaptive Gain Theory
ARI: Affective Reactivity Index
BIQ: Behavioral Inhibition Questionnaire
BPS: baseline pupil size

XSL•FO
RenderX

CBCL: Child Behavior Checklist
CBQ-VSF: Child Behavior Questionnaire-Very Short Form
DSGVO: European Data Protection Directive
EEG: electroencephalography
ERN: error-related negativity
ERP: event-related potential
GDPR: General Data Protection Regulation
GSP: Good Scientific Practice
HPA: hypothalamic-pituitary-adrenal
IPW: Inverse Probability Weighting
LC: locus coeruleus
LC-NE: locus coeruleus-norepinephrine
LMM: linear mixed models
MHD: mental health disorder
MICE: Multiple Imputation by Chained Equation
RAND20: random 20
REG 10: regular 10
SEPR: stimulus-evoked pupillary response
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WPPSI: Wechsler Preschool and Primary Scale of Intelligence

Edited by J Sarvestan; submitted 15.Dec.2025; this is a non–peer-reviewed article;accepted 03.Mar.2026; published 20.Apr.2026.

Please cite as:
Todorova I, Neubauer PR, Bast N
Locus-Coeruleus Norepinephrine Functioning as a Predictor of Childhood Mental Health (LOCUS-MENTAL): Protocol for a Longitudinal Study
JMIR Res Protoc 2026;15:e89237
URL: https://www.researchprotocols.org/2026/1/e89237
doi:10.2196/89237

© Iskra Todorova, Paula R Neubauer, Nico Bast. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Care Pathways and Patient Experiences Among Patients With Post COVID-19 Condition: Study Protocol for a Mixed-Methods Study in Germany

Daniela Gesell[1,2], Dr rer biol hum; Paula Lienesch[2], MSc; Yanyan Shi[2,3], PhD; Ralf Strobl[2,4], Dr rer biol hum; Martin Tauscher[5], Dr rer biol hum; Roman Gerlach[5], Dr rer biol hum; Eva Grill[2,4], Prof Dr; Daniela Koller[1,2], Prof Dr

[1]Department of Orthopaedics and Trauma Surgery, Musculoskeletal University Center Munich (MUM), University Hospital, LMU Munich, Marchioninistr. 15, Munich, Germany

[2]Institute of Medical Information Processing, Biometry and Epidemiology, Faculty of Medicine, LMU Munich, Marchioninistr 15, Munich, Bavaria, Germany

[3]Institute for Medical Biometry, Informatics and Epidemiology, University Hospital Bonn, University of Bonn, Bonn, Germany

[4]German Center for Vertigo and Balance Disorders, LMU University Hospital, Munich, Germany

[5]Bavarian Association of Statutory Health Insurance Physicians, Munich, Germany

Corresponding Author:
Daniela Gesell, Dr rer biol hum
Department of Orthopaedics and Trauma Surgery, Musculoskeletal University Center Munich (MUM), University Hospital, LMU Munich, Marchioninistr. 15, Munich, Germany

Abstract

Background: The COVID-19 pandemic has a lasting impact on health care utilization, as both the acute infection and post COVID condition (PCC) can lead to increased demand for medical services due to ongoing symptoms.

Objective: The aim of this study is to systematically examine health care utilization among individuals after acute SARS-CoV-2 infection in Bavaria, Germany, with a particular focus on PCC. The study combines claims data analysis with qualitative interviews to improve the understanding of objective care pathways and patients' subjective experiences within the health care system.

Methods: The research project 'SOLongCOVID' employs a mixed-methods design consisting of two subprojects: (1) a retrospective cohort study using claims data from the Bavarian Association of Statutory Health Insurance Physicians (KVB) to analyze care pathways through state sequence analysis, (2) a qualitative study based on semistructured interviews and focus groups with patients with PCC concerning their subjective care experiences. A synthesis process involving a focus group discussion will combine the information from the two subprojects, providing a comprehensive understanding of the care processes of patients with PCC.

Results: The study was funded by the German Federal Joint Committee Innovation Fund in October 2024. Statutory health insurance claims data cover the period from 2019 to 2022, and qualitative interview data collection is planned from May 2025 to August 2026. As of manuscript submission, study preparation and ethics approvals have been completed, and 14 participants have been recruited for the qualitative interviews. Study findings are anticipated to be published from July 2026 to August 2027.

Conclusions: The results are expected to enhance the understanding of existing barriers and challenges and to support evidence-based recommendations for improving care pathways for patients with specific care needs.

Trial Registration: DRKS - Deutsches Register Klinischer Studien DRKS00038446; https://drks.de/search/de/trial/DRKS00038446/details

International Registered Report Identifier (IRRID): DERR1-10.2196/91976

(JMIR Res Protoc 2026;15:e91976) doi:10.2196/91976

KEYWORDS

PCC; long COVID; ambulatory care; healthcare pathways; patient experiences; mixed-methods; post COVID condition

Introduction

According to the World Health Organization, more than 770 million people worldwide have been infected with Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), with approximately 6.9 million deaths [1]. While symptoms often subside within 14 days, many patients continue to have symptoms after the acute illness [2,3]. The generally used term 'long COVID,' medically described as *post-COVID-19 condition* (PCC) is defined as the presence of symptoms that persist or newly develop three months after a SARS-CoV-2 infection and last for at least two months without an alternative explanation

XSL•FO
RenderX

[4]. The symptoms experienced by patients with PCC are highly diverse and reflect the multifaceted nature of the condition. These symptoms can affect multiple systems of the body, further highlighting the complexity of the condition [3,5,6]. However, there are certain symptoms that occur more frequently. These include fatigue, myalgia, neurological symptoms such as cognitive impairment, loss of smell and taste, as well as headaches, mental health issues, including depression and anxiety, and cardiopulmonary symptoms like dyspnea, chest pain, and palpitations [5,7-11].

The broad range and duration of these symptoms, combined with the diverse nature and unpredictability of PCC, often lead to increased and prolonged health care utilization in both ambulatory and hospital settings, as demonstrated in several studies [12-15]. The increase in utilization has a major impact on the health care system, which is faced with rising costs and increasing demand for specialized treatment [14,16,17]. A study from the United States showed that SARS-CoV-2 infection was associated with an overall 4% increase in health care utilization within six months of infection, primarily driven by virtual contacts and emergency department visits [18]. Specifically for the ambulatory setting, recent studies on PCC health care utilization have mainly focused on treatment by specialists or consultation with primary care physicians. However, many of these studies are limited by follow-up periods of less than six months [5,8,10,11]. At the same time, international research is increasingly addressing potential barriers to care and patient experiences in accessing PCC-related health care services. These studies indicate that access to care is often limited by long waiting times, a shortage of specialists, and a lack of specialized treatment services. They also underscore the urgent need for patient-centered approaches to improve care for those affected [19-23].

Evidence on long-term health care utilization, particularly with respect to different types of physician services, remains limited. Gaining a more comprehensive understanding of health care utilization and the specific needs of patients following a SARS-CoV-2 infection is essential for policymaking and clinical management of affected patients.

To improve the understanding of both objective care pathways and subjective patient experiences, claims data derived from the routine records of the Bavarian Association of Statutory Health Insurance Physicians (KVB) will be analyzed, and qualitative interviews and a focus group will be conducted to examine the individual care experiences of people affected by PCC. The aim of the study is to systematically examine health care utilization among patients after an acute SARS-CoV-2 infection, in Bavaria, Germany, with a particular focus on PCC. The study combines claims data analysis with qualitative interviews to improve the understanding of objective care pathways and patients' subjective experiences within the health care system.

Methods

To ensure a high scientific standard and transparency, the reporting of this study protocol and the study results will adhere to the following guidelines: the *Strengthening the Reporting of Observational Studies in Epidemiology* (STROBE) guideline and the *Consolidated criteria for reporting qualitative research* (COREQ) Checklist [24-26]. The study is registered in the German Clinical Trials Register (registration number DRKS00038446).

Ethical Considerations

The claims data analysis and the qualitative study were approved by the Local Research Ethics Committee of Ludwig-Maximilians-University Munich (reference number 25 - 0173-KB and 25 - 0276). Written informed consent is obtained from each participant after they had been informed about the study's aims, procedures, voluntary nature, and data protection measures. Participant privacy and confidentiality are strictly maintained throughout the study. Personal identifiers are removed or not collected where not essential, and all data are handled and stored securely in accordance with institutional ethics and data protection guidelines. No identifiable personal information is reported in this manuscript. Copies of the signed consent forms are securely stored in accordance with institutional ethics guidelines. No compensation is provided to participants.

Study Design

The study follows a convergent mixed-methods design aiming to describe and analyze the trajectories of patients following an acute SARS-CoV-2 infection in the ambulatory setting (see Figure 1). Quantitative analyses of statutory health insurance claims data and qualitative interview analyses will be conducted in parallel and independently. Data and results from two subprojects will be used to facilitate the synthesis.

XSL•FO
RenderX

Figure 1. Mixed-methods study approach.

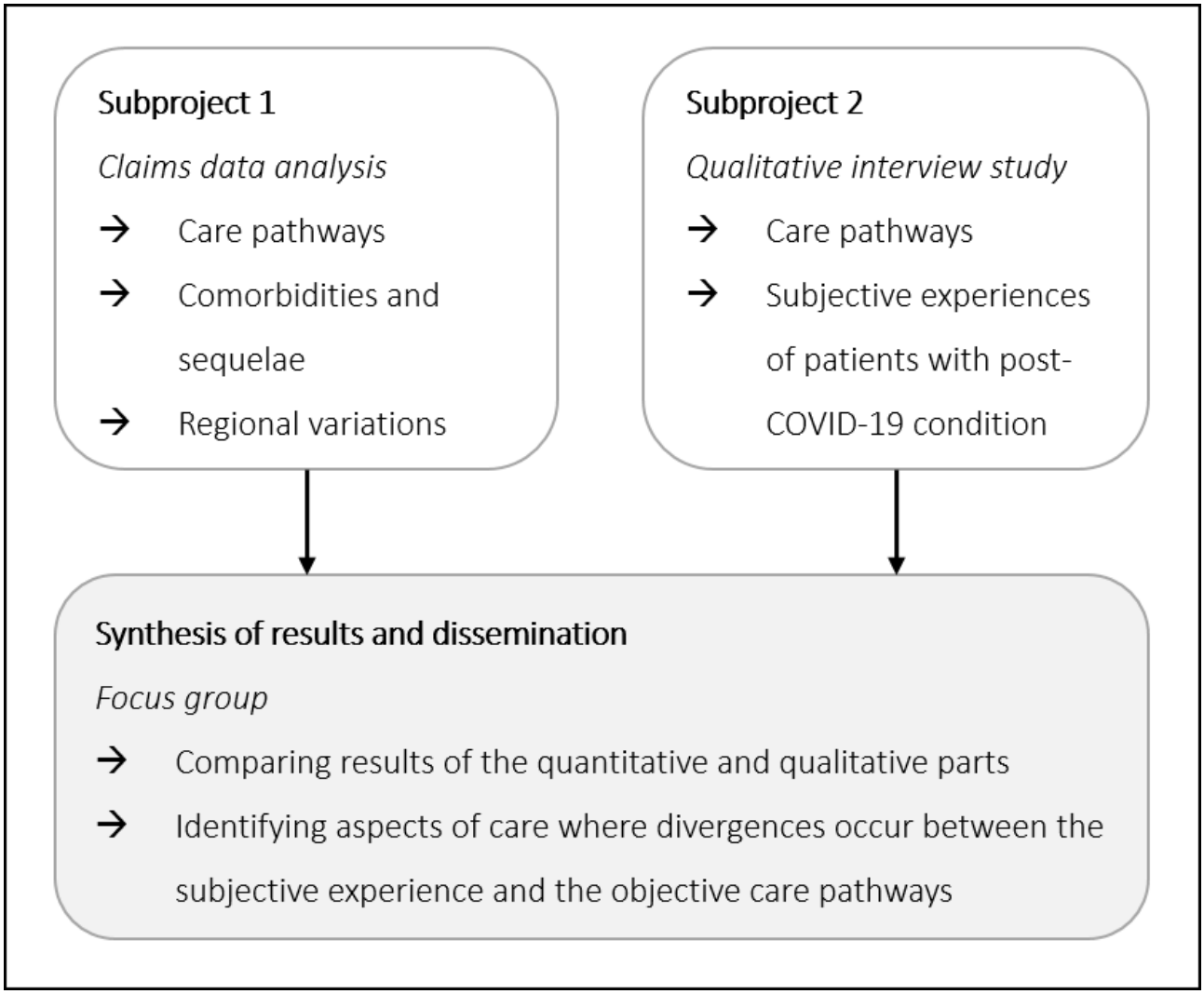

First, in a retrospective cohort study with claims data (Subproject 1) objective care pathways will be reconstructed based on routine data from the KVB. To do so, three cohorts will be defined, a cohort of patients with SARS-CoV-2 infection and a cohort of patients with other respiratory tract infection during the pandemic phase 2020 till 2022. To measure the impact that the pandemic had on the overall health care utilization, we will also form a historic control group with respiratory tract infections in the year 2019.

Second, a qualitative study (Subproject 2) will be conducted, including semistructured interviews with people affected by PCC concerning their individual subjective experiences and personal motivations for (non-) utilization of care.

The integration of the results will take place at the interpretation stage through systematic triangulation. Quantitatively identified care pathways will be compared with themes derived from patient interviews to identify areas of convergence, complementarity, and divergence. Joint displays will be developed that compare quantitative care trajectories with qualitative interview data, enabling structured comparisons across data sources. This will be achieved by using the software MAXQDA (version 24.9.1), as referenced by Kuckartz [27,28].

All results will be synthesized in a focus group in order to record and understand the care processes holistically and to classify them in the light of the existing evidence. The focus group will serve to validate and refine the results of the individual interviews, to contrast individual experiences within a group setting, and to enable in-depth exploration of shared experiences and interpretive patterns. Moreover, it supports the development of evidence-informed recommendations for care improvement. The synthesis will be published in scientific journals and made accessible to affected individuals and care providers to ensure broad dissemination and practical impact.

Participants and Data Collection

The subprojects will be informed by a systematic literature review performed before the project (manuscript forthcoming).

Subproject 1: KVB claims data will be used for the analysis. The KVB data includes all claims for ambulatory physicians in the State of Bavaria made by insured persons residing in Bavaria. The analysis is based on the *Bavarian COVID-19 Cohort* which was established from anonymized billing records of all patients with statutory health insurance in Bavaria. Both general practitioners and medical specialists submit details on patient encounters, including treatments, diagnoses, and billing information, to the KVB in each quarter. The dataset includes

XSL•FO
RenderX

pseudonymized patient identifiers, demographic variables (age, sex, region), diagnoses (International Classification of Diseases, 10th Revision, ICD-10 codes with diagnostic certainty) and billed services (EBM codes, specialty, and practice identifiers). This structure allows longitudinal tracking of health care utilization, comorbidities, and treatment patterns across cohorts.

Out of 9,202,930 individuals with any respiratory infection, 4,247,455 had a confirmed COVID-19 diagnosis and the remained any other respiratory infection. For the analysis, three cohorts were randomly drawn:

1. COVID-19 (2020–2022): 196, 000 individuals with a confirmed diagnosis of COVID-19, ie, *ICD* codes U07.1 or U08.9 or U09.9 (secured and suspected diagnoses)
2. Control (2020–2022): 140,000 individuals matched to the COVID-19 cohort by age and sex who received a confirmed ambulatory diagnosis of a respiratory infection (*ICD* J00*–J22*, J40*–J42*) in the same quarter.
3. Historical control (2019): 140,000 individuals with a confirmed respiratory infection (*ICD* J00*–J22*, J40*–J42*) diagnosed in 2019, matched by age and sex to the COVID-19 cohort

Subproject 2: Patients with PCC will be asked to participate in individual interviews about their experiences with health care delivery. Inclusion criteria comprise adult patients from the Munich region and the wider Bavaria area who either have a confirmed diagnosis of PCC or self-report having experienced PCC in the past or currently. Participants must have received or be receiving ambulatory treatment. Recruitment will be conducted through PCC support groups in Munich. These groups bring together individuals affected by PCC who meet regularly to exchange experiences and provide mutual support. A semistructured interview guideline will be developed and tested, taking into account individual experiences and care processes. Findings from the systematic review will be incorporated into the development of the guidelines. The guideline will establish a thematic framing and consists of all relevant topics that need to be addressed in the interview. It ensures better comparability of the data and structures the entire communication process [29]. Potential interview partners for the qualitative part of the project will be recruited utilizing a snowball sampling approach wherever possible [30]. This process follows the principle of maximum variation, aiming to capture a diverse range of perspectives and experiences. The qualitative interviews and the focus group will be conducted in person or via video conference. During the interview, patient journey maps will be developed [31,32]. The number of interviews is determined according to the methodological criterion of theoretical saturation [33]. The sample is expected to consist of around 20 to 25 participants. Should saturation be achieved earlier, indicated by the emergence of no new themes, a reduced number of interviews will be conducted. Conversely, if saturation is not attained, the sample size will be stepwise increased until the required number is reached.

The study will include a focus group, comprising 5 - 10 participants. The discussion will take place face-to-face in a university meeting room and will be facilitated by two researchers. If an in-person participation is not possible due to severeness of disease, an additional online-based option will be made available. Results of both the data analysis and the synthesis of the individual interviews will be presented. Discussions will follow a semistructured guide with open-ended questions and are planned to last between 60 and 90 minutes. Participants will be selected from the pool of individuals who previously took part in the individual interviews, applying the same inclusion criteria. Recruitment will occur through renewed email contact. Group composition will be arranged to allow discussion of both shared and divergent care experiences. Participation will be voluntary, with withdrawal possible at any time without consequences. Before the discussion begins, sociodemographic information will be collected and written informed consent for participation and audio recording will be obtained on site. The moderator will lead the discussion, while the co-moderator will oversee time management, observe group dynamics, and take field notes. Data will be collected through audio recordings supplemented by field notes. All data will be pseudonymized, and deleted after transcription is completed.

Data Analysis

Subproject 1

For the claims data, control groups will be matched by age, gender, and time of diagnosis (year/quarter). We will analyze the utilization patterns longitudinally for the cohorts and will focus specifically on the health care service billing group. We will estimate the incidence of PCC diagnoses (*ICD* codes: U09.9, U07.4), as well as newly occurring comorbidities, which may also be in the context of PCC. Expected outcomes are: (1) SARS-CoV-2 related symptoms, (2) newly developed diseases (cardiovascular diseases, autoimmune diseases, mental health diseases). Here, a special focus will be on diseases related to PCC, irrespective of a specific PCC diagnosis (eg, ICD-10 codes: R53, G93.3, F06.7, R43, M79.1, I26, R05, F48.0, F41) [4,34]. Health care utilization is analyzed using a *State Sequence Analysis* (SSA) [35]. Analyses are conducted for the cohort of SARS-CoV-2 infections and the control groups. In the analysis, the health care utilization for primary and specialist care as well as their respective diagnoses will be analyzed in the time before and after the SARS-CoV-2 infection. A specific focus will be on the appearance of a PCC diagnosis or diagnoses of PCC-related conditions. SSA identifies individual care trajectories represented as ordered sequences of quarterly states, where each state reflects contacts with general practitioners, COVID-19 specialists, or other specialists. Dissimilarities between trajectories are quantified using a theory-driven cost matrix, and clusters of patients with similar utilization patterns are derived through an unsupervised medoid-based clustering procedure. The optimal number of clusters is selected based on established quality indices (eg, average silhouette width). Factors associated with the resulting care pathway clusters are subsequently analyzed using multivariable regression models.

Additionally, the potential health care utilization depends on regional factors. Therefore, the place of residence of the cohort is recorded at the level of the districts and independent cities of Bavaria as well as the degree of urbanization of the municipality of residence. The regional distributions of the cohort and utilization patterns are visualized and analyzed using

XSL•FO
RenderX

geographical information systems (GIS). Differences are analyzed using hierarchical regression models.

Subproject 2

All audio recordings from the interviews will be transcribed verbatim to ensure an accurate representation of participants' perspectives. The evaluation process of the qualitative interviews employs thematic content analysis, as outlined by Mayring, to identify patterns and themes that emerge from the narratives [33]. This analytical process will involve both deductive category application based on the interview guide and inductive category development derived from the data itself. To support transparency, traceability, and efficient data management, MAXQDA (version 24.9.1) software will be used for coding and category refinement. Throughout the analysis, regular team discussions will be conducted to enhance reflexivity and ensure the reliability and coherence of the coding framework.

Furthermore, the results of the quantitative and qualitative subprojects are compared, and aspects of care are identified where there are divergences between the qualitative experience and the quantitative analyses. The prepared results of the care pathways of the claims data analysis and the experiences of the interview study are presented to the participants of the focus group and discussed. In the focus group, recommendations for action will be developed. The transcript will be analyzed using qualitative content analysis following Mayring's approach [33], which enables a systematic, rule-guided development of categories and the stepwise interpretation of the material.

Results

The study was funded by the German Federal Joint Committee Innovation Fund in October 2024 and is progressing according to the planned timeline, with data collection and analysis set to commence thereafter to generate meaningful results. Subproject 1: Statutory health insurance claims data covering the period from 2019 to 2022 were provided by the KVB. The first access to the routine data was in March 2025, with corrections made in April 2025. Analysis of the routine data began in July 2025 with quality control and descriptive analyses, with the first results expected by May 2026. Further analyses will be conducted until April 2027. Subproject 2: The recruitment of participants for the qualitative interviews and focus groups began in May 2025, and data collection and recruitment are expected to be finished by June 2026. As of manuscript submission, study preparation and ethics approvals have been completed, and 14 participants have been recruited for the qualitative interviews. First results of the qualitative interviews and focus group are expected in August 2026. The synthesis of the results and the development of recommendations are expected to be finished in August 2027.

Discussion

This mixed-methods study will contribute to deepening and expanding the current state of research by combining systematic evidence synthesis, qualitative and quantitative approaches of subjective care experiences and objective care pathways of patients with PCC. The mixed-methods design was chosen to capture individual experiences of care that can be triangulated with the results of the routine data analysis, interpreted and embedded in the current state of research. The combination of different methods is a strength of this study, as it offers the opportunity to gain a deeper and more comprehensive understanding of ambulatory care practices for PCC. While there may be different descriptions or explanations, they complement each other and contribute to a broader picture [27,36].

It is expected that patients with PCC will show increased health care utilization. The wide range of symptoms observed in this population, combined with the heterogeneity of available health care services, presents a major challenge in identifying appropriate care pathways. Therefore, this study determines, in comparison with control cohorts, which services were utilized and what outcomes were achieved. In addition, it seeks to identify approaches that could improve care delivery. While the expected findings are currently limited to COVID-19 and PCC, they may offer valuable insights for managing other post-infectious conditions and for supporting patients with still persistent PCC symptoms. This could lead to a more efficient use of resources in the ambulatory sector and support earlier improvements in patients' health. Through the integration of the systematic review, as well as subjective and objective care pathways and experiences, the study offers a substantial opportunity for an improved care provision.

As for all studies, some limitations need to be considered. First, in analyzing claims data, it must be acknowledged that clinical information is only available to a limited extent, as coding is primarily based on billing-related aspects. This can result in inaccuracies, particularly regarding *ICD* codes, due to the lack of mandatory coding standards, especially in general practice settings [37]. In addition, the dataset is limited to the region of Bavaria, which may introduce regional bias in observed health care utilization patterns. Temporal inaccuracies can also occur, as the recorded dates often reflect billing processes rather than the exact time of care provision [38].

Second, for the qualitative study, interviews capture subjective perspectives rather than objective realities. Recruitment via support groups and snowball sampling may lead to an overrepresentation of individuals who are more engaged or better connected within patient networks. Individuals with PCC who are less socially connected or socioeconomically disadvantaged may be underrepresented. While qualitative research does not aim for statistical representativeness, it seeks to provide in-depth insights into experiences, perceptions, and contextual factors [39]. The interview setting may introduce selectivity or artificiality, especially with personal topics like health. Moreover, the interpretative nature of the method involves a degree of subjectivity, and inherent trade-offs, such as between narrative depth and comparability which must be carefully managed [40]. The methodology for preparing and conducting the interviews will therefore be carefully designed to ensure consistency, taking into account all aspects of the research to enhance both the validity of the interviews and the overall understanding of the study. Furthermore, the focus only on patients in the Munich region may not fully capture health care experiences and developments in other geographical or

XSL•FO
RenderX

more rural areas. A potential bias may also arise among focus group participants, as individuals involved in such networks may be more proactive or reflective about their situation.

Based on the expected results, this study is intended to provide concrete evidence for decision-makers in health care policy, particularly with regard to the future planning and organization of care structures. The more systematic inclusion of the patient-centered perspective can improve the relevance and effectiveness of future care concepts.

Acknowledgments

The authors thank the Bavarian Association of Statutory Health Insurance Physicians (Kassenärztliche Vereinigung Bayerns, KVB) for providing access to the anonymized health insurance data utilized in this study. We further thank all participants who join in in the qualitative interviews for their time and valuable insights. The authors declare their use of generative AI in the research and writing process. The following tasks were completed by ChatGPT 5.2 under full human supervision: proofreading, editing and translation. The authors take full responsibility for the final manuscript.

Funding

The study is funded by the German Federal Joint Committee Innovation Fund (Innovationsfonds des Gemeinsamen Bundesausschusses, grant number 01VSF24046) and has been peer reviewed by the funding body. DG and PL were supported by the funder (URL: https://innovationsfonds.g-ba.de/projekte/solongcovid.683). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

Due to confidentiality and data protection regulations, the qualitative interview data and health insurance claims data analyzed in this study are not publicly available. Interview data may be made available from the corresponding author upon reasonable request, whereas claims data cannot be shared due to legal restrictions. Generated analysis code can be provided upon reasonable request for purposes of scientific transparency.

Authors' Contributions

Conceptualization: DK, EG, RG, MT
Funding acquisition: DK, EG, RG, MT
Investigation: DK, EG
Methodology: YS, RS, PL
Project administration: DK, EG
Visualization: DG, PL
Writing – original draft: DG, PL, RS, DK, EG
Writing – review & editing: PL, RS, YS, MT, RG, EG, DK

Conflicts of Interest

None declared.

References

1. WHO COVID-19 dashboard. World Health Organization. 2024. URL: https://data.who.int/dashboards/covid19/cases?n=c [accessed 2025-04-10]
2. Chen C, Haupert SR, Zimmermann L, Shi X, Fritsche LG, Mukherjee B. Global prevalence of post-coronavirus disease 2019 (COVID-19) condition or long covid: a meta-analysis and systematic review. J Infect Dis 2022 Nov 1;226(9):1593-1607. [doi: 10.1093/infdis/jiac136] [Medline: 35429399]
3. O'Mahoney LL, Routen A, Gillies C, et al. The prevalence and long-term health effects of Long Covid among hospitalised and non-hospitalised populations: a systematic review and meta-analysis. EClinicalMedicine 2023 Jan;55:101762. [doi: 10.1016/j.eclinm.2022.101762] [Medline: 36474804]
4. Post COVID-19 condition (long COVID). World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/post-covid-19-condition-(long-covid) [accessed 2025-03-11]
5. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. Sci Rep 2021 Aug 9;11(1):16144. [doi: 10.1038/s41598-021-95565-8] [Medline: 34373540]
6. Greenhalgh T, Sivan M, Perlowski A, Nikolich J. Long COVID: a clinical update. The Lancet 2024 Aug;404(10453):707-724. [doi: 10.1016/S0140-6736(24)01136-X]
7. Natarajan A, Shetty A, Delanerolle G, et al. A systematic review and meta-analysis of long COVID symptoms. Syst Rev 2023 May 27;12(1):88. [doi: 10.1186/s13643-023-02250-0] [Medline: 37245047]

XSL•FO
RenderX

8. Raveendran AV, Jayadevan R, Sashidharan S. Long COVID: an overview. Diabetes Metab Syndr 2021;15(3):869-875. [doi: 10.1016/j.dsx.2021.04.007] [Medline: 33892403]
9. Sudre CH, Murray B, Varsavsky T, et al. Attributes and predictors of long COVID. Nat Med 2021 Apr;27(4):626-631. [doi: 10.1038/s41591-021-01292-y] [Medline: 33692530]
10. Luo D, Mei B, Wang P, et al. Prevalence and risk factors for persistent symptoms after COVID-19: a systematic review and meta-analysis. Clin Microbiol Infect 2024 Mar;30(3):328-335. [doi: 10.1016/j.cmi.2023.10.016] [Medline: 37866679]
11. Su S, Zhao Y, Zeng N, et al. Epidemiology, clinical presentation, pathophysiology, and management of long COVID: an update. Mol Psychiatry 2023 Oct;28(10):4056-4069. [doi: 10.1038/s41380-023-02171-3] [Medline: 37491461]
12. Castriotta L, Onder G, Rosolen V, et al. Examining potential Long COVID effects through utilization of healthcare resources: a retrospective, population-based, matched cohort study comparing individuals with and without prior SARS-CoV-2 infection. Eur J Public Health 2024 Jun 7;34(3):592-599. [doi: 10.1093/eurpub/ckae001] [Medline: 38243748]
13. DeVoss R, Carlton EJ, Jolley SE, Perraillon MC. Healthcare utilization patterns before and after a long COVID diagnosis: a case-control study. BMC Public Health 2025 Feb 10;25(1):514. [doi: 10.1186/s12889-025-21393-4] [Medline: 39930426]
14. Tene L, Bergroth T, Eisenberg A, David SSB, Chodick G. Risk factors, health outcomes, healthcare services utilization, and direct medical costs of patients with long COVID. Int J Infect Dis 2023 Mar;128:3-10. [doi: 10.1016/j.ijid.2022.12.002] [Medline: 36529373]
15. Huang BZ, Creekmur B, Yoo MS, Broder B, Subject C, Sharp AL. Healthcare utilization among patients diagnosed with COVID-19 in a large integrated health system. J Gen Intern Med 2022 Mar;37(4):830-837. [doi: 10.1007/s11606-021-07139-z] [Medline: 34993879]
16. Koumpias AM, Schwartzman D, Fleming O. Long-haul COVID: healthcare utilization and medical expenditures 6 months post-diagnosis. BMC Health Serv Res 2022 Aug 8;22(1):1010. [doi: 10.1186/s12913-022-08387-3] [Medline: 35941617]
17. Musheyev B, Boparai MS, Kimura R, et al. Longitudinal medical subspecialty follow-up of critically and non-critically ill hospitalized COVID-19 survivors up to 24 months after discharge. Intern Emerg Med 2023 Mar;18(2):477-486. [doi: 10.1007/s11739-023-03195-x] [Medline: 36719540]
18. Tartof SY, Malden DE, Liu ILA, et al. Health care utilization in the 6 months following SARS-CoV-2 infection. JAMA Netw Open 2022 Aug 1;5(8):e2225657. [doi: 10.1001/jamanetworkopen.2022.25657] [Medline: 35960522]
19. Gamillscheg P, Łaszewska A, Kirchner S, Hoffmann K, Simon J, Mayer S. Barriers and facilitators of healthcare access for long COVID-19 patients in a universal healthcare system: qualitative evidence from Austria. Int J Equity Health 2024 Oct 23;23(1):220. [doi: 10.1186/s12939-024-02302-4] [Medline: 39443935]
20. Gardner E, Lockrey A, Stoesser KL, et al. Challenges in receiving care for long COVID: a qualitative interview study among primary care patients about expectations and experiences. Ann Fam Med 2024 Sep 23;22(5):369-374. [doi: 10.1370/afm.3145] [Medline: 39191462]
21. Schmachtenberg T, Königs G, Dragaqina A, et al. "There is no one who helps you with it": experiences of people with long COVID regarding medical care, therapeutic measures, and barriers in the German healthcare system: results of a qualitative study with four focus groups. BMC Health Serv Res 2023 Oct 26;23(1):1160. [doi: 10.1186/s12913-023-10170-x] [Medline: 37884993]
22. MacEwan SR, Rahurkar S, Tarver WL, et al. Patient experiences navigating care coordination for long COVID: a qualitative study. J GEN INTERN MED 2024 Jun;39(8):1294-1300. [doi: 10.1007/s11606-024-08622-z] [Medline: 38308155]
23. Turk F, Sweetman J, Chew-Graham CA, et al. Accessing care for long Covid from the perspectives of patients and healthcare practitioners: a qualitative study. Health Expect 2024 Apr;27(2):e14008. [doi: 10.1111/hex.14008] [Medline: 38481384]
24. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
25. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg 2014 Dec;12(12):1495-1499. [doi: 10.1016/j.ijsu.2014.07.013] [Medline: 25046131]
26. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
27. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
28. Kuckartz U. Integrative datenanalyse und joint displays mit MAXQDA. In: Mixed Methods: Methodologie, Forschungsdesigns Und Analyseverfahren: Springer Fachmedien Wiesbaden; 2014:123-154. [doi: 10.1007/978-3-531-93267-5_5]
29. Misoch S. Qualitative Interviews: Walter de Gruyter GmbH & Co KG 2019.
30. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. Soc Probl 1997 May;44(2):174-199. [doi: 10.1525/sp.1997.44.2.03x0221m]
31. Trebble TM, Hansi N, Hydes T, Smith MA, Baker M. Process mapping the patient journey: an introduction. BMJ 2010 Aug 13;341:c4078. [doi: 10.1136/bmj.c4078] [Medline: 20709715]
32. Davies EL, Bulto LN, Walsh A, et al. Reporting and conducting patient journey mapping research in healthcare: a scoping review. J Adv Nurs 2023 Jan;79(1):83-100. [doi: 10.1111/jan.15479] [Medline: 36330555]

XSL•FO
RenderX

33. Mayring P. Qualitative inhaltsanalyse. In: Flick U, Keupp H, S W, editors. Handbuch qualitative Forschung : Grundlagen, Konzepte, Methoden und Anwendungen: München: Beltz - Psychologie Ver. Unnion; 1991:209-213.
34. Donnachie E, Hapfelmeier A, Linde K, et al. Incidence of post-COVID syndrome and associated symptoms in outpatient care in Bavaria, Germany: a retrospective cohort study using routinely collected claims data. BMJ Open 2022 Sep 22;12(9):e064979. [doi: 10.1136/bmjopen-2022-064979] [Medline: 36137635]
35. Gabadinho A, Ritschard G, Mueller NS, Studer M. Analyzing and visualizing state sequences in R with TraMineR. J Stat Softw 2011;40(4):1-37. [doi: 10.18637/jss.v040.i04]
36. Tariq S, Woodman J. Using mixed methods in health research. JRSM Short Rep 2013 Jun;4(6):2042533313479197. [doi: 10.1177/2042533313479197] [Medline: 23885291]
37. Slagman A, Hoffmann F, Horenkamp-Sonntag D, Swart E, Vogt V, Herrmann WJ. Analyse von Routinedaten in der Gesundheitsforschung: Validität, Generalisierbarkeit und Herausforderungen. Z Allg Med 2023 Apr;99(2):86-92. [doi: 10.1007/s44266-022-00004-0]
38. Swart E, Gothe H, Geyer S, et al. Good Practice of Secondary Data Analysis (GPS): guidelines and recommendations. Gesundheitswesen 2015 Feb;77(2):120-126. [doi: 10.1055/s-0034-1396815] [Medline: 25622207]
39. Pyo J, Lee W, Choi EY, Jang SG, Ock M. Qualitative research in healthcare: necessity and characteristics. J Prev Med Public Health 2023 Jan;56(1):12-20. [doi: 10.3961/jpmph.22.451] [Medline: 36746418]
40. Hohl J. Das qualitative interview. J Public Health (Oxf) 2000 Jun;8(2):142-148. [doi: 10.1007/BF02962637]

Abbreviations

COREQ: Consolidated criteria for reporting qualitative research
GIS: Geographical information system
ICD-10: International Classification of Diseases, 10th Revision
KVB: Association of Statutory Health Insurance Physicians in Bavaria (Kassenärztliche Vereinigung Bayern)
PCC: Post-COVID-19 Condition
SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2
SSA: State Sequence Analysis
STROBE : Strengthening the Reporting of Observational Studies in Epidemiology

Edited by J Sarvestan; submitted 22.Jan.2026; this is a non–peer-reviewed article;accepted 16.Mar.2026; published 22.Apr.2026.

Please cite as:
Gesell D, Lienesch P, Shi Y, Strobl R, Tauscher M, Gerlach R, Grill E, Koller D
Care Pathways and Patient Experiences Among Patients With Post COVID-19 Condition: Study Protocol for a Mixed-Methods Study in Germany
JMIR Res Protoc 2026;15:e91976
URL: https://www.researchprotocols.org/2026/1/e91976
doi:10.2196/91976

© Daniela Gesell, Paula Lienesch, Yanyan Shi, Ralf Strobl, Martin Tauscher, Roman Gerlach, Eva Grill, Daniela Koller. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Establishment, Implementation, and Impacts of the Observatory on Student Mental Health in Higher Education in Quebec, Canada: Protocol for a Mixed Methods Research Program

Julie Lane[1], PhD; Benjamin Gallais[2], PhD; François Lauzier-Jobin[3], PhD; Sèverine Lanoue[4], PhD; Félix Guay-Dufour[4], PsyD; Rémi Paré-Beauchemin[5], MSc; Rachel Guertin[4], PhD; Christiane Bergeron-Leclerc[6], PhD; Saliha Ziam[7], PhD; Marie-Claude Lallier Beaudoin[8], PhD; Martine Shareck[9], PhD; Véronique Boudreault[10], PsyD, PhD; Marie-Ève Langelier[6], MD; Steve Geoffrion[11], PhD; Nancy Beauregard[12], PhD; Bianca B-Lamoureux[13], PhD; Magaly Brodeur[14], MD, PhD; Esther McSween-Cadieux[7], PhD; Audrey Dupuis[15], PhD; Jonathan Smith[16], PhD; Isabelle Thibault[17], PhD; Noémie Beauregard[11], MSW

[1]Department of School and Social Adaptation Studies, University of Sherbrooke, 2500 University Boulevard, Sherbrooke, QC, Canada
[10]Faculty of Physical Activity Sciences, University of Sherbrooke, Sherbrooke, QC, Canada
[11]School of Psychoeducation, University of Montreal, Montreal, QC, Canada
[12]School of Industrial Relations, University of Montreal, Montreal, QC, Canada
[13]Department of Education, TELUQ University, Montreal, QC, Canada
[14]Department of Family Medicine and Emergency Medicine, University of Sherbrooke, Sherbrooke, QC, Canada
[15]Faculty of Education, University of Moncton, Moncton, NB, Canada
[16]Department of Preschool and Primary Education, University of Sherbrooke, Sherbrooke, QC, Canada
[17]Department of Psychoeducation, University of Sherbrooke, Sherbrooke, QC, Canada
[2]Department of Health Sciences, University of Quebec at Chicoutimi, Chicoutimi, QC, Canada
[3]School of Nursing Science, University of Sherbrooke, Sherbrooke, QC, Canada
[4]University of Sherbrooke, Sherbrooke, QC, Canada
[5]Faculty of Education, University of Sherbrooke, Sherbrooke, QC, Canada
[6]Department of Humanities and Social Sciences, University of Quebec at Chicoutimi, Chicoutimi, QC, Canada
[7]School of Business Administration, TELUQ University, Montreal, QC, Canada
[8]Department of Psychology, University of Sherbrooke, Sherbrooke, QC, Canada
[9]Faculty of Medicine and Health Sciences, University of Sherbrooke, Sherbrooke, QC, Canada

Corresponding Author:
Julie Lane, PhD
Department of School and Social Adaptation Studies, University of Sherbrooke, 2500 University Boulevard, Sherbrooke, QC, Canada

Abstract

Background: Research is needed to better understand mental health (MH) problems among higher education (HE) students and how to address them. The Observatory on Student Mental Health in Higher Education (OSMHHE) brings together 350 members across Quebec (Canada) and internationally. Its mission is to develop, promote, and disseminate knowledge to foster a culture that supports student MH in HE.

Objective: This study aims to describe the OSMHHE's research protocol, which consists of three objectives: (1) establishing a portrait of students' MH and its determinants, (2) identifying and evaluating a variety of MH practices, and (3) assessing the implementation and impacts of the OSMHHE as a knowledge mobilization infrastructure.

Methods: Objective 1 will be achieved through a provincial survey using a cross-sectional, repeated-measures design with 2 data collections (November 2024 and 2026) targeting the entire Quebec HE student population. Dimensions, indicators, and scales were selected based on conceptual frameworks, a systematic literature review, and Delphi methods. Analyses will include descriptive statistics by education levels; inferential analyses comparing subpopulations; multiple regressions, logistic models, and linear mixed models to identify MH determinants; and repeated-measures ANOVA to examine temporal changes. Objective 2 will evaluate the implementation, sustainability, scale-up, and impacts of MH practices using mixed methods. Analyses may include descriptive and comparative statistics, correlations, structural equations modeling (path analysis), and qualitative content or thematic analyses. Objective 3 will draw on the framework of Ziam et al to assess knowledge mobilization strategies. A developmental evaluation approach and convergent mixed methods design within a case study will be used to assess the OSMHHE's

XSL•FO
RenderX

implementation and impacts. Qualitative data will include semistructured individual and group interviews with OSMHHE members, addressing topics such as roles, decision-making processes, facilitators and barriers, and outcomes. Additional qualitative sources will include diverse documents (eg, meeting agendas, reports). Quantitative data will come from questionnaires completed by members examining levels of engagement and satisfaction, challenges and barriers, and impacts of the OSMHHE's activities and knowledge mobilization practices. Qualitative data will be analyzed using content analyses. Quantitative data will be examined using descriptive, comparative, and correlational analyses.

Results: This project is funded from February 2023 to February 2028. The first provincial survey took place in November 2024, collecting data from 32,212 students in 77 HE institutions. Analyses are underway, and a first report was released in November 2025. Approximately 20 student MH practices are currently being evaluated.

Conclusions: The OSMHHE provincial survey will provide portraits of students' MH in HE in Quebec and its determinants to better guide MH practices and institutional decision-making. Evaluating MH practices will advance knowledge of their effectiveness. Assessing the implementation of the OSMHHE will help deepen our understanding of knowledge mobilization infrastructures designed to support student MH in HE.

International Registered Report Identifier (IRRID): DERR1-10.2196/83225

(JMIR Res Protoc 2026;15:e83225) doi:10.2196/83225

KEYWORDS

observatory; higher education; mental health; impacts; students; knowledge mobilization

Introduction

Background

Higher education institutions (HEIs) are internationally recognized as environments that support the mental health (MH) of their communities [1,2]. This is in line with the Healthy Universities and Health Promoting Universities movement [3] and the Okanagan Charter, to which 45 countries adhere [4]. Despite this recognition, several studies have highlighted the precarious MH of students in higher education (HE) [5-8]. A study involving 14,348 students from 19 universities in 8 countries showed that 38.4% of respondents had experienced at least 1 MH problem (eg, anxiety disorder, depressive disorder, substance use disorder) over the course of their studies [8]. These MH problems can impair students' cognitive function, psychosocial adjustment, motivation, and academic performance, while being associated with reduced well-being, lower social participation and academic satisfaction, as well as increased risks of dropping out of school and of suicide [9-14].

In Quebec (Canada), a 2020 survey of 1815 college students found that 33.7% had clinical symptoms consistent with a depressive disorder and that 24.9% had clinical symptoms of generalized anxiety disorder [15]. A province-wide survey carried out by the Quebec Student Union found that 58% of university students were experiencing a high level of psychological distress and that 1 in 5 had depressive symptoms severe enough to warrant medical or psychological support [16]. Another survey by the College Students Federation of Quebec reported that 63.8% of students felt their psychological health had deteriorated during the pandemic [17]. Finally, a Quebec Student Union survey also conducted during the pandemic showed that 51% of the student community felt that their level of distress had increased during this period [18].

Although the surveys presented above highlight MH issues among students in HE, further research is needed to deepen knowledge, monitor changes over time, and reduce certain methodological biases. Indeed, many surveys rely on recruitment strategies (eg, via social networks) that may limit sample representativeness [19,20]. Moreover, some surveys use nonstandardized instruments, use methodologies that vary across institutions, or rely on cross-sectional designs. These limitations make it difficult to develop a reliable province-wide portrait of student MH, identify its risk and protective factors, conduct longitudinal follow-up, and formulate actionable recommendations. Finally, few studies adopt a holistic approach that considers both the multiple dimensions of student MH and the systemic factors or determinants that influence it, whether individual (eg, identity, psychosocial skills), social (eg, social support, discrimination), or environmental (eg, sense of belonging to the university, campus climate).

In response to MH issues, HEIs report difficulties coping with increased demand for consultations and the growing severity of problems encountered [21-23]. A meta-analysis has shown that traditional psychological services struggle to meet the diverse needs of students [24]. Moreover, these services are often fragmented, tend to respond to needs in a reactive and ad hoc manner, are underfunded, and, consequently, require substantial additional resources [25].

On the other hand, it remains difficult to ascertain which practices are most effective or promising for supporting students' MH in HE [26-32]. In Quebec, as elsewhere, little information is available on the implementation and evaluation of these interventions [33,34]. Despite these limitations, several meta-analyses have concluded that the interventions assessed (eg, mindfulness meditation) are effective, showing small-to-moderate effect sizes on symptoms of stress, anxiety, and depression [26-29,31,32,35,36]. However, Huang et al [30] note that interventions targeting the general student population do not demonstrate significant efficacy for more severe MH problems. This finding is consistent with another meta-analysis reporting that, among students reporting high levels of stress, more structured and longer-term intervention programs are particularly effective [35]. Finally, several stakeholders in the field have observed a proliferation of practices and innovations

XSL•FO
RenderX

being implemented in HEIs without undergoing any evaluation [37].

Initiatives to improve students' MH require a broad cultural transformation within academic environments, spanning from the promotion to implementation of practices to foster MH [38-41]. Producing research results, and even disseminating, synthesizing, or contextualizing them within HEIs to promote their use, is not sufficient. There is also a need to support the adaptation of practices and the development of both individual and collective leadership related to MH. HEIs would benefit from receiving support to promote, implement, evaluate, sustain, and scale up MH practices across multiple institutions [42]. A systematic review on the sustainability of public health interventions implemented in educational settings identified key enabling conditions (eg, support from leaders, coconstruction approaches) and contextual barriers (eg, insufficient human and financial resources) that can promote or hinder their implementation [43]. However, to date, little is known about the implementation of promising practices to promote MH in educational settings, or about the factors that influence their adoption and sustainability [44,45]. In light of this, supporting knowledge mobilization (KM) to promote and maintain a culture conducive to student MH in HE is clearly relevant.

Objective of the Study

To overcome these limitations, the Observatory on Student Mental Health in Higher Education (OSMHHE) (*Observatoire sur la santé mentale étudiante en enseignement supérieur*), funded by the Fonds de recherche du Québec (FRQ, 2023 - 2027), aims to contribute to the advancement of knowledge and the KM development and practice in order to promote and maintain a culture conducive to enhancing student MH in HE [46]. This study describes the OSMHHE research protocol, as submitted to the FRQ, which is akin to a research program. The study presents the OSMHHE logic model and its components, the objectives of the research protocol, the research methodology, the anticipated results, and a discussion.

OSMHHE Logic Model

Overview

Drawing on the Quebec government's 2021 - 2026 Action Plan on Student Mental Health in Higher Education (*Plan d'action sur la santé mentale étudiante en enseignement supérieur*) [47], the OSMHHE aims to (1) develop scientific knowledge through research projects and a survey on student MH and promising MH practices. Since there are few evidence-based practices in the field of student MH in HE, we use the concept of promising practice. A promising practice must have demonstrated a minimal effect (eg, observed positive changes toward goal attainment), evidence of correct quality (eg, in the early stages of evaluation), or high potential to produce similar positive results in other contexts [48]. In addition, the concept of practice encompasses initiatives, programs, measures, services, and so forth. Furthermore, the OSMHHE aims to (2) contribute to training and mobilizing students with regard to MH in order to create positive leaders; (3) seek out and monitor promising practices in MH promotion, awareness, prevention, and intervention; and (4) promote KM by implementing activities to support changes in culture and practices. The OSMHHE logic model is illustrated in Figure 1, and more information can be found on the OSMHEE website [46].

XSL•FO
RenderX

Figure 1. Observatory on Student Mental Health in Higher Education (OSMHHE) logic model. FRQ: Fonds de recherche du Québec; HEI: higher education institution; KM: knowledge mobilization; MH: mental health.

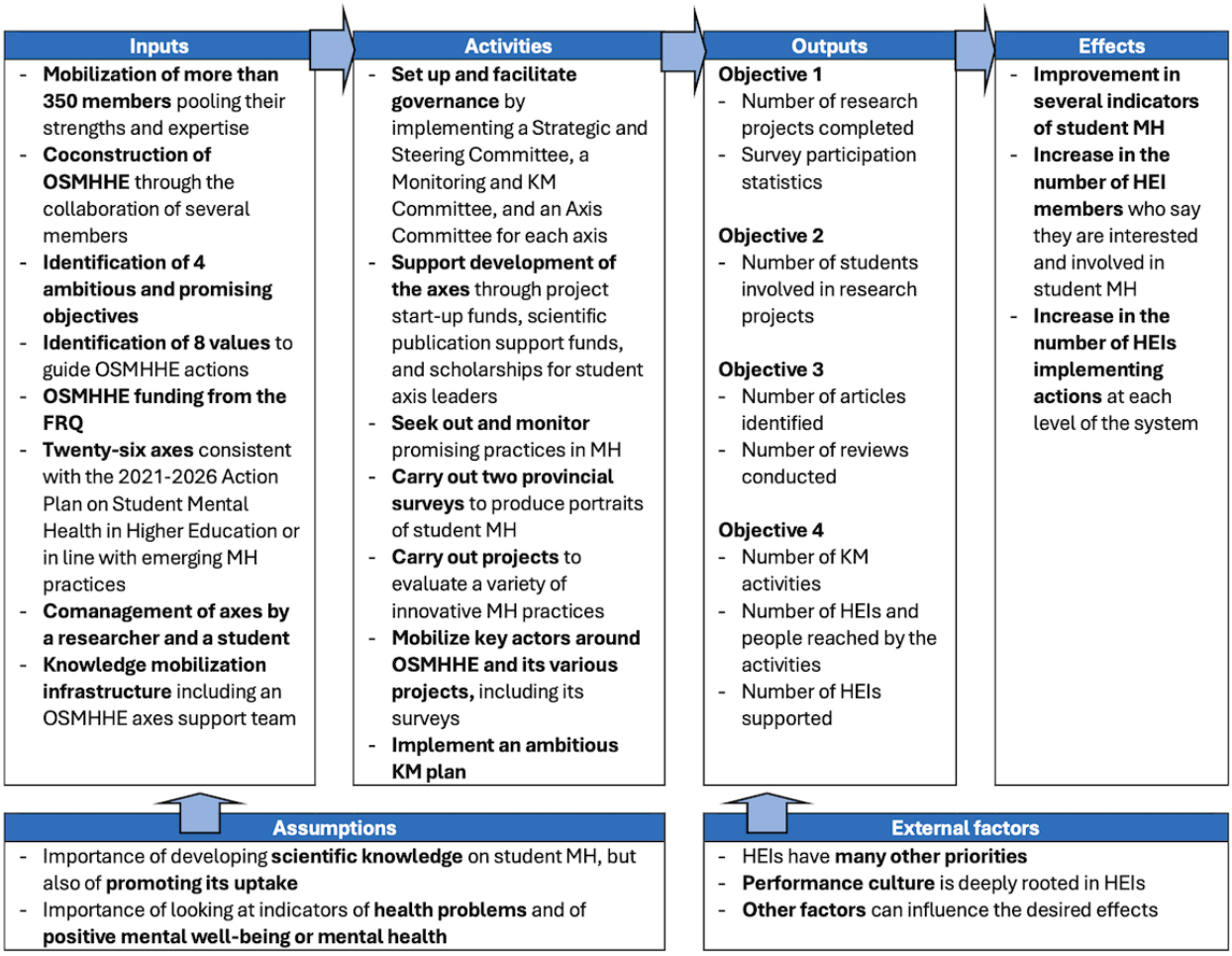

OSMHHE Axes and Governance

The OSMHHE brings together more than 350 professors, researchers, students, and collaborators, as well as a number of cross-sectoral partners (eg, government bodies, HEIs, community organizations). Eight values guide their actions and projects: (1) democratization of access to knowledge; (2) nonhierarchization of knowledge and human relationships; (3) benevolence and authenticity; (4) collaboration; (5) equity, diversity, and inclusion; (6) commitment and determination; (7) creativity; and (8) excellence.

These members collaborate around 26 mission, population, and theme axes (Figure 2). Each axis is supported by a coinvestigator and a student coleader and brings together a variety of members within an axis committee. The OSMHHE has also established a strategic committee and a monitoring and KM committee.

XSL•FO
RenderX

Figure 2. The 26 mission, population, and thematic axes of the Observatory on Student Mental Health in Higher Education (OSMHHE).

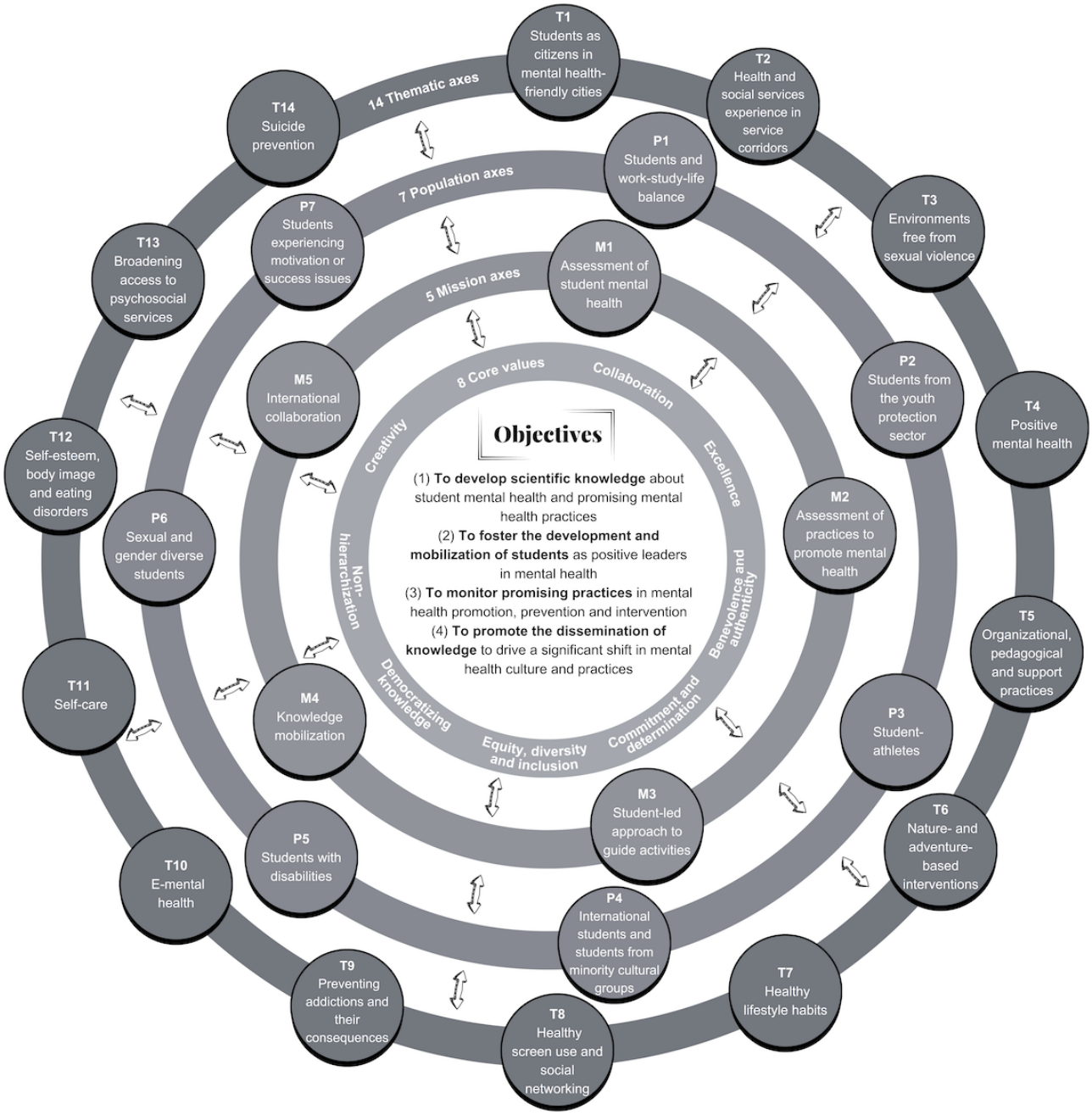

The Four Pillars of OSMHHE

Overview

To achieve its objectives, the OSMHHE is grounded on four pillars: (1) adopting a holistic conception of student MH, (2) fostering a participatory "by-and-for students" approach, (3) implementing governance based on the Teal organizational model [49], and (4) defining itself as a knowledge mobilization infrastructure (KMI) to support change in culture and practices in HEIs. These pillars make the OSMHHE a unique infrastructure that stands out from others, such as Canada's Student Mental Health Network [50,51] and the Community Engagement Course and Action Network [52]. Indeed, these networks are more oriented toward KM, whereas the OSMHHE combines a dual mission of advancing knowledge through both research and KM.

A Holistic Conception of Student MH and Its Determinants

The OSMHHE adopts a holistic conception of MH, defined as "a state of well-being in which the individual realizes his or her own abilities, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to his or her community" (p. 1) [53]. It therefore does not consist solely of the absence of disease [53]. Keyes' [54] complete state model of MH proposes the combination of a first continuum representing positive MH—ranging from languishing MH to flourishing MH—and a second continuum ranging from the absence of distress through to the presence of symptoms, or even an MH disorder. Although complementary, these dimensions are distinct [55-57]. The concomitant analysis of these 2 continua provides a more complete description of individuals' MH status [58,59], particularly in HE [60]. The positive dimension of MH includes psychological well-being,

XSL•FO
RenderX

life satisfaction, meaning in life, and recovery, while the negative dimension presents a continuum that includes psychological distress, early warning signs, and psychiatric symptoms [56,61].

In addition, understanding and intervening on student MH requires adopting a developmental approach [62-66]. This approach integrates concepts from several other theoretical models and has the advantage of simultaneously considering the roles of several groups of interrelated factors and their interactions, with the goal of better understanding individual development in relation to the environment [67,68]. Considering these factors is essential, as students' MH is influenced by social determinants of mental health (SDMH), which include individual, social, environmental, economic, cultural, and political determinants [53]. Some determinants may act as risk factors for developing MH disorders or as aggravating factors in their progression [69,70], while others may act as protective or promotional factors (eg, having a clear and mature identity, having a sense of authenticity and meaning in life, adopting healthy lifestyle habits, benefiting from social support, and evolving in a positive university climate) and prevent the onset of disorders among students [15,71-75]. The OSMHHE recognizes that SDMH must be understood in the light of structural and historical power relations. Thus, the project aims to document the cumulative and differentiated effects of systemic exclusion, stigmatization, racism, sexism, ableism, homophobia, and other forms of discrimination in student trajectories.

In concrete terms, the OSMHHE embodies this conception across all its activities. For example, in its KM efforts, members are working to raise awareness within HEIs about the importance of addressing multiple SDMH. This conception is also reflected in the various themes of the OSMHHE's 26 axes. Likewise, in its research activities, particular attention is paid to documenting both the positive and negative dimensions of MH and the various SDMH.

A Participatory "By-and-for Students" Approach

Anchored in a participatory "by-and-for students" approach, the OSMHHE aims to work in concert with this population by involving them at every stage, from conceptualization to interpretation, and ultimately, to KM [76]. This central role assigned to students helps the OSMHHE move away from the hierarchical relationship that typically exists within HEIs and ensures that student voices remain at the core of the project. Despite inherent challenges (eg, determining the depth of their participation, diminishing the hierarchization of knowledge) [77,78], this approach has the potential to promote better alignment with community needs; coconstruct best practices; and support their adoption, use, and sustainability, as well as share power while fostering capacity and agency of those involved [76,79]. Students can participate across different spheres, including training (eg, with active instruction [80]), educational practices (eg, the *student voice* method [81]), and research, which falls within the broader field of participatory approaches [76,82].

In concrete terms, the OSMHHE was coconstructed with significant contributions from students, including provincial student associations. As a result, this community is actively involved in its governance committees, where it is at the heart of decision-making. To ensure that the perspectives of those most affected by structural inequalities in MH are heard and valued, the OSMHHE ensures that representatives from marginalized groups are included in its governance structures.

Students can take on a number of responsibilities, including participating in discussions and decision-making, helping to plan the activities of each axis, organizing OSMHHE events (eg, conferences, scientific symposia), and evaluating applications for start-up funds that finance OSMHHE flagship projects. In addition, a community of practice is dedicated specifically to student members. Led by 2 student coleaders, it aims to promote the MH and well-being of its members, as well as the development of their knowledge and skills.

Governance Based on the Teal Organizational Model

The OSMHHE implements the Teal governance model (as proposed by Laloux [49]), which is characterized by self-governance, holism, and an evolutionary purpose. This type of organizational model has been associated with several benefits, including: (1) higher job satisfaction, attributed to the increased level of autonomy that people enjoy [83]; (2) greater freedom to innovate, which positively affects their engagement [84]; and (3) greater team effectiveness [85]. Such governance has created space for collaborative and shared leadership, both individual and collective, as well as for the embodiment of values that are shared and coconstructed by the active members of the organization [86]. While these findings are encouraging, further studies are needed given their limited number [85,87].

Accordingly, the OSMHHE encourages its axes and their members to exercise leadership that drives its activities forward. For example, suggestions are made to guide their work (eg, annual frequency of meetings, agenda) and update the role of axis leaders. These proposals are intended to support members while allowing them freedom of action. An OSMHHE support team accompanies them, offering assistance with initial setup, agenda preparation, and facilitation. Finally, various mechanisms have been put in place to foster the development of projects (eg, project start-up funds) and the dissemination of results (eg, financial assistance for publication).

A KM Infrastructure to Support Change in Culture and Practices

The OSMHHE is defined as a KMI, that is, a set of interconnected components (governance, human and financial resources, KM strategies and activities, processes, etc) that promotes KM by stimulating sharing among various actors (researchers, knowledge users, etc), and facilitates the use of research knowledge to enhance its impact [88-90]. Several KMIs have emerged since the early 2000s [91-93]. These are increasingly recognized as a strategy to support practice improvement and evidence-based decision-making [94-97]. Despite their emergence, very few empirical studies have evaluated them [98]. Current research is more focused on evaluating specific KM activities and initiatives that rely on individuals, such as knowledge brokers, than on assessing KMIs as systems [99-101]. To date, little is known about the roles, governance structures, responsibilities, processes, and activities

XSL•FO
RenderX

of KMIs [98]. This is further complicated by the fact that each KMI has unique characteristics that enable it to remain in constant motion as a complex system [102] influenced by a wide range of political, social, organizational, human, and financial factors [92,96,103]. Thus, it is difficult to determine how these KMIs function to produce their impacts [104,105].

A recent evaluation of Canada's Student Mental Health Network highlights the lack of a standardized evaluation process for assessing this type of KMI [51], despite the acknowledged importance of continuously evaluating these networks' collaborative efforts to improve partner synergy and ensure effective project functioning [52]. Given this state of knowledge, publications such as this one can contribute to building a body of knowledge that identifies optimal ways to evaluate KMIs. This contribution draws on the KM expertise of several OSMHHE members in evaluating KM strategies or KMI and in supporting the development of professional practices.

Operationally, the OSMHHE KMI includes a support team composed of an OSMHHE coordinator who provides high-level coordination of all activities and supports certain axes, 2 support staff and scientific advisors who assist specific axes and carry out cross-sectoral mandates, a communications specialist, a knowledge broker specializing in KM, an individual responsible for ethical considerations and liaison with the granting agency, as well as students, supporting specific mandates. This team facilitates the work of members and cross-sectoral activities.

In addition, the OSMHHE has a dedicated KM axis whose role has been to help develop an initial KM plan, inspire the axes, evaluate activities, and assess the implementation and impacts of OSMHHE. The initial plan, drawing on knowledge brokers whose work aligns with broad categories of brokering activities [106,107], is ambitious and evolving. For the time being, it includes the following activities: (1) knowledge dissemination and management activities (eg, website, newsletter, member information, *Être et Liberté* podcast by and for students, collection of practice guides, scientific publications), (2) mobilization and networking activities (eg, conferences, colloquia, assemblies); (3) capacity-building activities (eg, interactive workshops or training courses to foster uptake of our guides or productions), and (4) change-support activities (eg, training and support for intermediaries who play a role in supporting HEIs).

Objectives of the Research Protocol

With the aim of remaining attentive, responsive, and agile with the knowledge it will develop and mobilize in response to the needs of students, institutions, and its members, the OSMHHE has equipped itself with a unique infrastructure and set far-reaching objectives that require both flexible and innovative conceptualization and methodology. The research protocol has three objectives:

1. Develop portraits of student MH and the determinants that support it
2. Evaluate a variety of MH practices
3. Evaluate the implementation and impacts of the OSMHHE as a KMI implementing governance inspired by the Teal organizational model and a participatory approach

Methods

Methods for Objective 1: Develop Portraits of Students' MH

Overview

Portraits will be developed based on surveys and complementary research projects. In particular, a first province-wide survey on the MH of students in HE was conducted in Quebec's HEIs in the fall of 2024 and will be repeated in the fall of 2026.

The specific objectives of these surveys are to:

- Establish a portrait of students' MH, both positive and negative, in Quebec's HEIs, and track its evolution over time using the 2 measurement periods
- Generate subpopulation portraits that account for the diversity of individual characteristics and educational trajectories of student populations in Quebec's HEIs
- Provide each participating institution with portraits of the MH of their student population
- Assess the need for, access to, and use of psychosocial support services within and outside HEIs

Creating an Integrated Conceptual Model and Selecting Indicators

To select the dimensions, indicators, and scales to be included in the survey, our team developed a conceptual framework that integrates 3 existing frameworks [47,61,108,109] and conducted a systematic review of the literature specific to the MH of students in HE.

These sources were used to produce a comprehensive portrait of the objects and dimensions to be studied in an MH survey. Inspired by Delphi methodologies [110], they were first subjected to cross-sectional analysis. They were then grouped, merged, or recategorized according to logical criteria to develop a conceptual model specific to our research object, and ultimately, to select the indicators to be included in the survey. This process took into account, among other things, the priorities identified by OSMHHE members.

The selection of variables (scales, questionnaires, items, and questions) assessing the selected indicators was supported by a second literature review focusing on the psychometric qualities of the measurement instruments (validity and reliability), the availability and validated English and French versions as well as the short versions, the royalty-free options, and the number of items.

Design of the Provincial Survey on the MH of Students in HE

The survey is based on a cross-sectional, repeated-measures design (2024 - 2026). A census-type strategy is used, meaning that the entire Quebec HE student population is surveyed. All 154 post-secondary institutions are targeted, including private, public, and subsidized or nonsubsidized sectors.

To minimize participation biases and obtain satisfactory statistical power, for a population of 415,000 students, a 20% sample size is targeted to give a margin of error of 0.5%, a risk

XSL•FO
RenderX

of α=5, and statistical power greater than 99.9%. However, a 30% sample would be much preferred, in hopes of achieving better representation of specific populations in line with the OSMHHE population axes. To achieve this rate, several strategies are used to encourage active collaboration from institutions, such as premeetings with representatives of HEIs, collaboration with student associations, and participation incentives through a prize draw.

The survey includes questions on students' individual characteristics (eg, age, gender identity, sexual orientation, immigration status), MH (eg, positive MH, anxiety symptoms, depressive symptoms, suicidal ideation, and risks), and MH determinants (eg, perception of MH, identity development, authenticity, meaning in life, substance use, screen time, sleep, physical activity). It also covers several systemic contextual variables (eg, social relationships, campus climate, sense of belonging), sexual violence (eg, reporting sexual violence experienced since arriving at the HE institution), and care pathways (eg, help sought from a psychosocial support service, knowledge of psychosocial support services). Examples of survey questions are provided in Multimedia Appendix 1, along with the observational study protocol checklist by Low et al [111] completed for the survey study (Table S1 in Multimedia Appendix 2).

Data Collection and Analysis

An invitation to complete an online questionnaire hosted on the Limesurvey platform is sent by email to all students enrolled in HEIs.

At each of the two data-collection periods, descriptive analyses will be carried out to present the results of the various indicators by students' postsecondary education levels. Inferential analyses will also be conducted to compare the different education levels as well as subpopulations defined by individual characteristics (eg, students with disabilities, immigrant background) and academic contexts (eg, full-time or part-time). Multiple regression and logistic models will be used to target the factors (individual and academic context characteristics, student experience, lifestyle habits) that may predict MH variables. To track changes in MH indicators, repeated-measures analyses of variance will be used to determine whether positive or negative changes have occurred over time. Finally, more complex linear mixed-effects models will be used to determine whether certain determinants explain the changes observed in MH indicators across time.

Methods for Objective 2: Evaluate MH Practices

Overview

The second general objective of the OSMHHE is to evaluate MH practices, programs, and initiatives. More specifically, this involves assessing both their impacts and their implementation and scale-up, as considering these 2 types of impacts often provides clues that could help explain implementation failures or successes [43,45,112,113]. Projects related to this objective will be implemented gradually as promising practices, programs, or initiatives are identified.

Evaluation of the Implementation, Sustainability, and Scale-Up of Practices

In line with this objective, the OSMHHE encourages its members to evaluate the implementation, sustainability, and scale-up of promising practices identified. This type of evaluation seeks to highlight the conditions that influence these practices' adoption, scale-up, and sustainability within HEIs. Targeted initiatives will be prioritized using an analysis tool developed by Ben Charif et al [114], which assesses a practice's potential for scale-up.

To achieve this objective, the OSMHHE will implement a mixed approach based on the work of several authors, as suggested by Schultes [45] for evaluating the implementation of school-based interventions [115,116]. On the one hand, Proctor et al [115] propose implementation impact analysis as a promising strategy for identifying indicators of success (eg, program accessibility, acceptability, adoption by the community, relevance, feasibility). On the other hand, Damschroder et al [116,117] propose the Consolidated Framework for Implementation Research, which helps identify optimal conditions for implementing promising practices or programs. This mixed approach has already been used by the first author of this paper to evaluate a program aimed at preventing anxiety in young people. Table S1 in Multimedia Appendix 3 provides an overview of the methodological approach adapted for program evaluation in HE.

Although data analysis will vary from project to project, depending on the evaluation context, quantitative descriptive and correlational analyses could be used [118,119]. In addition, drawing on implementation-evaluation tools, path analysis using structural equations may be performed [120]. Finally, thematic analyses could be carried out using qualitative data [121].

Evaluation of the Impacts of Different Student MH Practices

The OSMHHE aspires to contribute to the development of a culture of evaluation in HEIs by ensuring that a substantial number of practices are evaluated each year, either independently or with support. The analysis of their impacts will pay specific attention to the differentiated effects associated with membership in historically marginalized groups. The intention is to make sure that interventions do not replicate, or even exacerbate, existing inequalities and that they support true equity in the observed effects. To achieve this ambitious goal, members of the OSMHHE will help support and equip institutions wishing to carry out these evaluations themselves. Initially, a first toolkit will be developed and distributed to guide institutions in their evaluation of training workshops aimed at improving student MH. Adaptations of this toolkit will subsequently be proposed for other types of initiatives (eg, peer support). Support will also be offered to institutions that wish to use the toolkit.

In addition, certain axes—in particular, axis M2 (evaluation of practices to promote MH)—will conduct in-depth evaluations of selected practices based on the following criteria: (1) project scope and cross-sectoral nature, (2) potential project impacts, (3) potential contribution to the advancement of scientific knowledge, (4) nature of the targeted practice (eg, consistency

XSL•FO
RenderX

with policy directions, innovation), and (5) presence of conditions likely to facilitate the conduct of the research (eg, level of engagement of the communities proposing the projects).

Data analysis for this component will vary depending on the practices, methodological approach, and tools selected for the evaluation by the institutions. In general, quantitative descriptive, correlational, and comparative analyses [119] could be conducted. In addition, content analyses [76,122] or thematic analyses [121] could be carried out with qualitative data. Table S1 in Multimedia Appendix 4 provides a concrete example of this methodological approach, which is currently being used to evaluate a program aimed at developing the psychosocial skills of students in HE.

Methods for Objective 3: Evaluate the OSMHHE Implementation and Impacts

Overview

The third objective is to evaluate the implementation and impacts of OSMHHE as a KMI that operated with governance inspired by the Teal organizational model and a participatory approach that fosters student mobilization. The intention is to document the innovative nature of the OSMHHE structure by evaluating the development, functioning, and progress toward achieving its objectives. It will be based on the framework of Ziam et al [123] for assessing KM strategies. This framework identifies four components to be considered: (1) the internal context, (2) the external context, (3) the implementation process, and (4) the short-term and medium-term impacts. This flexible framework targets the components and indicators that relate to the conditions for implementation and the impacts of the KM strategy.

Design for the Evaluation of the OSMHHE Implementation and Impacts

The design proposed for this evaluation is the case study [124], that is, an in-depth investigation of a specific phenomenon (the case) within its natural context (real-life setting). It will take the form of an embedded case study, with each axis, committee, and work team constituting a unit of analysis [124]. It will rely on data triangulation [76,124] using different sources of information. To avoid overburdening members, data collection will be aligned with the OSMHHE's regular activities, including accountability reporting.

Evaluative Approach

The evaluative approach used is grounded in the principles [125] of developmental evaluation [126], which aims to assess social innovations embedded in complex, dynamic environments. In this approach, the evaluator is involved in the program and seeks "to support the development of complex and dynamic interventions in real time" (p. 2) [127]. This creates strong coherence between the evaluative approach and the establishment of the OSMHHE, whose structure is innovative and whose actions are guided by a shared set of values. This approach is also consistent with the proposed design, as it encourages feedback loops in the development of the infrastructure.

Data Collection and Analysis

A convergent mixed methods design will be used, combining qualitative (eg, individual and group interviews, documents, observations) and quantitative data (eg, questionnaires) [128]. Qualitative and quantitative data will be collected and analyzed separately, and the results will then be compared and interpreted jointly.

Approximately 15 semistructured individual interviews [129,130] will be conducted with key members of the OSMHHE, including codirectors, coordinator, support team members, and axis leaders and members until theoretical saturation is reached. Three semistructured group interviews [2] will also be conducted with other OSMHHE members. Depending on the timing of the interviews, the protocols will address topics such as participants' understanding and perceptions of the OSMHHE's mission and objectives; strategic planning of actions to achieve these objectives; roles, decision-making processes, and governance; level of engagement and the nature of collaborations; planned or implemented KM strategies; anticipated or observed facilitators and barriers; anticipated or observed outcomes and indicators of success (eg, knowledge development, student training, monitoring of promising practices, etc). All interviews will be recorded and transcribed verbatim.

Qualitative data collection will also include documents already collected by the OSMHHE for operational monitoring (eg, meeting agendas, email newsletters, annual reports, publications). This will involve analyzing the observation notes of the support team members, in line with the participant observation research methodology, where the observer is a recognized participant, in a real-life setting [129,131,132]. Notes will be recorded in a logbook to facilitate systematic documentation. In addition, content analysis will be carried out from discussions on the Teams platform used by the student community of practice to evaluate both its implementation and impacts. Public documents generated by the OSMHHE (eg, meeting minutes) will also be integrated as information sources, consistent with a factist approach [133].

Regarding quantitative data collection, several questionnaires will be developed by the research team and tailored to the context and realities of the OSMHHE and to the different groups of participants surveyed (researchers, clinicians, students, partner institutions). The questionnaires will include both closed-ended and open-ended questions. The data will be collected from 150 members and partners of the OSMHHE. Participants will be surveyed about their level of engagement and satisfaction, the extent to which they feel the OSMHHE's values are embodied, the challenges and barriers encountered, and the impact of the OSMHHE's activities and KM practices (eg, MH awareness, improved knowledge on promising practices, changes in practices within HE institutions, implementation of actions, and perceived cultural change in HE).

The qualitative data [122] will be analyzed using content analysis. The data will be transferred to qualitative analysis software (NVivo 11; QSR International) to facilitate data management, coding, and extraction [134]. Quantitative data will be analyzed using IBM SPSS software [119] to generate

XSL•FO
RenderX

descriptive statistics. Comparative and correlational analyses will also be conducted to assess the influence of certain variables and to identify trends, including analyses based on sociodemographic factors and correlations between variables related to students' MH.

Ethical Considerations

Regarding the provincial survey on the MH of students in HE, ethical approval was obtained from 2 designated research ethics committees (Université de Sherbrooke [2023-3736/Lane] and Cégep de Jonquière [CER-2324-30]). Approval was also secured from all participating institutions, in accordance with institutional requirements and prevailing ethical standards. Although the survey is deemed to be of minimal risk, several ethical considerations are carefully addressed. For example, participation is contingent upon informed consent. As participants create a code for anonymous identification before answering the questionnaire, their participation is completely anonymous, as are their data. Participants were entirely free to choose whether or not to answer the different questions in the survey. A button allowing them to withdraw from the study and not submit their responses was visible and available on every page of the survey. However, because the data were anonymous, it was not possible to withdraw participants' responses after they had completed and submitted the survey. MH support resources are embedded throughout the questionnaire, and a comprehensive list of available services is provided at both the beginning and the end of the survey. To promote inclusivity and representativeness of the Quebec student population, participants may complete the questionnaire in either French or English. In addition, the principles of equity, diversity, and inclusion are considered and integrated into all documents. This strategy provides nonprobability sampling with voluntary participation. Finally, two CAD $500 (US $353.08) gift cards and eight CAD $100 (US $70.62) gift cards were drawn among all the students who participated in the survey.

Results

This research project is funded from February 2023 to February 2028. The first of the 2 OSMHHE periodic provincial surveys took place in November 2024. The data were collected from 77 HEIs and 32,212 students. Analyses are currently underway, and the final report was released in November 2025. This survey will be used to produce detailed portraits of the MH status of students in HE, in terms of well-being (positive MH) and of MH difficulties. Identifying MH determinants through this questionnaire will guide HEIs in implementing practices and measures. The survey will contribute to determining whether certain populations are more vulnerable to MH issues and to developing tailored winning strategies for intervention and even prevention. It will also enable a more effective response to the diverse needs of the student population and ensure equitable access to available services.

The 2 survey waves will make it possible to assess changes in the student MH portrait. While maintaining confidentiality, specific reports could be provided to participating HEIs, or to student and professional associations concerned with the issues affecting specific populations (eg, students with disabilities, student athletes). The use of standardized questionnaires, the projected sample size, and the large number of variables identified in the integrated conceptual framework will support the development of a body of scientific knowledge unprecedented in Quebec.

Moreover, around 20 student MH practices are currently being evaluated in terms of their impacts and the conditions under which they are implemented, sustained, and scaled up. Evaluation reports will be published beginning in 2026. These impact evaluations will contribute to understanding their effectiveness and help identify promising practices. Exploring implementation conditions will deepen our understanding of the factors involved in implementing and adopting these practices, and foster their sustainability and scale-up. These findings can also be incorporated into the KM plan. Overall, these evaluations are expected to promote student MH and advance scientific knowledge in this field. A further aim is that the projects will foster change in the culture of practice evaluation within HEIs.

Finally, evaluating the implementation of the OSMHHE infrastructure will help advance knowledge about this type of infrastructure and its functioning. It will also enable feedback loops consistent with the principles-based evaluation approach. In other words, the results generated from the various data collections will allow for continuous monitoring of OSMHHE and its activities, which is essential to the infrastructure's agility, optimize its operation, and promote its impact. This evaluation will also help advance knowledge on the benefits such infrastructures can bring, including concrete cultural and practice changes in both research (culture of collaboration, interdisciplinarity, and KM) and practice, notably the MH of HE students. It is hoped that this initiative will inspire the development of similar infrastructures elsewhere in the world to foster this cultural change for the benefit of students' MH.

Discussion

Principal Findings

This protocol describes the development of the OSMHHE, a KMI designed to generate, mobilize, and apply scientific and practical knowledge to improve student MH in HE. It offers a unique infrastructure that combines a research mission, aimed at advancing knowledge, with a KM mission focused on supporting cultural and practice change within HEIs. In a context where institutions worldwide are facing an intensification of MH problems among students and increasing calls for systemic approaches, the OSMHHE integrates large-scale data collection, province-wide surveillance, implementation evaluation, assessment of practices, participatory student engagement, and KM within a single coordinated infrastructure. In this regard, it contributes to emerging provincial, national, and international efforts to operationalize learning health principles in HE settings.

First, the objective of developing scientific knowledge through research projects and 2 provincial surveys on student MH and promising practices is consistent with ongoing studies emphasizing the importance of continuing research in this area using rigorous methodologies [135]. These surveys will make

XSL•FO
RenderX

a significant contribution to the understanding of student MH by simultaneously considering the positive and negative dimensions of MH and by exploring the explanatory factors that may influence it, including individual, social, and environmental determinants. Repeated data collection may also allow for the monitoring of trends and changes over time and the identification of evolving needs. Because these data will be collected across all HEIs in Quebec, it will be possible to assess the effects of policy orientations, such as the Quebec government's 2021 - 2026 Action Plan on Student Mental Health in Higher Education [47]. Our findings may inform policymakers concerned with student MH in HE. Providing HEIs with personalized reports will enable them to adjust their actions to better support the MH of their student populations. In doing so, those institutions embody the learning health system approach, as proposed by Broglia and Barkham [135] for HE.

Our studies will also help identify the most effective practices (eg, programs, initiatives), in line with our objective of seeking out and monitoring promising practices in MH promotion, awareness, prevention, and intervention in Quebec and internationally. These contributions directly address concerns raised by several authors regarding the difficulty of deciding which are the most effective practices for MH among HE students [26-32].

Second, the objective of supporting the training and mobilizing students involved in MH, with the aim of developing positive leaders, relies fundamentally on our participatory "by-and-for students" approach. The OSMHHE adopts a participatory approach that involves them at every stage: design, interpretation, and dissemination of the results [76]. By prioritizing coconstruction, this approach helps to align actions with real needs, promote knowledge uptake, build capacity, and support genuine empowerment. It is agreed that it is important to involve those most directly concerned—in this case, students—in initiatives, such as the OSMHHE. Indeed, some studies emphasize the importance of involving key members of the target community [2,136,137]. In the KM field, there is shared recognition of the value of interactive approaches that engage stakeholders in various phases, such as knowledge production, adaptation, and dissemination [138]. Partnership and coconstruction are also common features of many KMIs [92,93]. In addition, in terms of the sustainability of actions in academic environments, Herlitz et al [43] emphasize the importance of tailoring interventions to students' needs and that their involvement is a key factor in the implementation and sustainability of initiatives. The participatory approach adopted by the OSMHHE, therefore, has the potential to improve alignment with local needs, support the coconstruction and uptake of knowledge, and foster capacity-building and empowerment among those involved [76,79]. However, as previously noted, participatory approaches also present a number of challenges [77,78]. To support full student participation in governance, research, and KM actions, the OSMHHE has set up a student community of practice. This initiative fosters their commitment while helping to train the next generation of researchers and leaders in the field of MH.

Third, our goal of fostering KM through the implementation of activities to support cultural and practice changes implies a collective awareness of the importance of such changes. It is widely acknowledged that implementing practices aimed at improving student MH entails substantial cultural change, especially in academic environments [38-41]. Contributing to this transformation requires more than producing research findings: their application must also be promoted. Institutions need support to innovate, implement, scale up, evaluate, and sustain these practices. Considering that HEIs have limited data on promising MH practices and the optimal conditions for their implementation, adoption, sustainability, and scale-up [44,45], OSMHHE studies and KM plans can meaningfully contribute to this needed cultural and practice change.

Finally, it is important to identify the strengths and limitations of this research program. Its major strengths are the scope of the proposed research program and the mobilization around it, the systemic approach, and the participatory "by-and-for students" approach, which enable action across multiple levels, thereby increasing the likelihood of achieving meaningful results. The main limitation is that the current funding is insufficient to fully implement this large-scale research program. Several OSMHHE researchers are seeking (and some have obtained) additional funding to support the projects.

Conclusions

The OSMHHE's mission—advancing knowledge and supporting KM to promote and maintain a culture conducive to student MH in HE—is ambitious and driven by its 350 members and collaborators across Quebec and internationally. Thanks to FRQ funding, the OSMHHE has 5 years to accomplish this mission. The coming years will allow us to assess whether the extensive efforts invested in establishing this infrastructure and its activities will bear fruit and generate the intended impacts. The hope is that these impacts will extend well beyond scientific outputs and will contribute to improving student MH and fostering a culture supportive of student MH in HE in Quebec and elsewhere for the benefit of the communities in which these students learn, live, and work.

Acknowledgments

The authors want to thank all the members of the Observatory on Student Mental Health in Higher Education for contributing to its mission. No generative artificial intelligence tools were used at any stage in the preparation of this manuscript.

XSL•FO
RenderX

Funding

This research was supported by grant 324014 (CAD $2,809,980 [US $2,059,996]) from the Fonds de recherche du Québec (2023-2027). The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Data Availability

Consistent with open access practices, the data related to the protocol presented in this study are accessible according to a procedure that protects confidential information and regulates the use of such data. Reasonable requests for access to datasets should be addressed to osmees@usherbrooke.ca.

Authors' Contributions

JL, BG, FL-J, FG-D, RP-B, and SL drafted the first version of the research protocol that was submitted to the Fonds de recherche du Québec. JL and BG developed the study concept and methodology of the research protocol. FL-J, FG-D, RP-B, and SL did a literature review. CB-L, SZ, MCLB, VB, MEL, MB, EM-C, AD, JS, and IT reviewed the research protocol before it was submitted to the Fonds de recherche du Québec. FL-J, SL, JL, RP-B, and BG wrote this manuscript based on the protocol. RP-B has updated the literature review. All authors reviewed the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of questions in the provincial survey on the mental health of students in higher education.
[PDF File, 176 KB - resprot_v15i1e83225_app1.pdf]

Multimedia Appendix 2
Observational study protocol checklist.
[PDF File, 260 KB - resprot_v15i1e83225_app2.pdf]

Multimedia Appendix 3
Example of a methodological approach to evaluate program implementation, sustainability, and scale-up.
[PDF File, 142 KB - resprot_v15i1e83225_app3.pdf]

Multimedia Appendix 4
Example of a methodological approach to evaluate the zenith program aimed at developing the psychosocial skills of students in higher education.
[PDF File, 135 KB - resprot_v15i1e83225_app4.pdf]

Peer Review Report 1
Peer-review report from the Fonds de recherche du Québec.
[PDF File, 292 KB - resprot_v15i1e83225_app5.pdf]

References

1. Tafireyi CGS, Grace JM. Status of the Health Promoting University (HPU) globally and its relevance for emerging African HPUs: an integrative review and bibliometric analysis. Glob Health Promot 2024 Sep;31(3):42-59. [doi: 10.1177/17579759241235109] [Medline: 38520053]
2. Suárez-Reyes M, Muñoz Serrano M, Van den Broucke S. How do universities implement the Health Promoting University concept? Health Promot Int 2019 Oct 1;34(5):1014-1024. [doi: 10.1093/heapro/day055] [Medline: 30052965]
3. Sánchez-Ordóñez JM. Healthy and health promoting university: concept, characteristics and challenges. In: Liamputtong P, editor. Handbook of Concepts in Health, Health Behavior and Environmental Health: Springer; 2025:1-15. [doi: 10.1007/978-981-97-0821-5_54-1]
4. Okanagan Charter: an international charter for Health Promoting Universities and Colleges. : International Conference on Health Promoting Universities & Colleges; 2015 URL: https://open.library.ubc.ca/media/stream/pdf/53926/1.0132754/7 [accessed 2026-03-12]
5. Biasi V, Patrizi N, Mosca M, De Vincenzo C. The effectiveness of university counselling for improving academic outcomes and well-being. Br J Guid Counc 2017 May 27;45(3):248-257. [doi: 10.1080/03069885.2016.1263826]

XSL•FO
RenderX

6. Wiens K, Bhattarai A, Pedram P, et al. A growing need for youth mental health services in Canada: examining trends in youth mental health from 2011 to 2018. Epidemiol Psychiatr Sci 2020 Apr 17;29:e115. [doi: 10.1017/S2045796020000281] [Medline: 32299531]
7. Sheldon E, Simmonds-Buckley M, Bone C, et al. Prevalence and risk factors for mental health problems in university undergraduate students: a systematic review with meta-analysis. J Affect Disord 2021 May 15;287:282-292. [doi: 10.1016/j.jad.2021.03.054] [Medline: 33812241]
8. Auerbach RP, Mortier P, Bruffaerts R, et al. Mental disorder comorbidity and suicidal thoughts and behaviors in the World Health Organization World Mental Health Surveys International College Student Initiative. Int J Methods Psychiatr Res 2019 Jun;28(2):e1752. [doi: 10.1002/mpr.1752] [Medline: 30450753]
9. Martineau M, Beauchamp G, Marcotte D. Efficacité des interventions en prévention et en promotion de la santé mentale dans les établissements d'enseignement postsecondaire [Article in French]. Sante Ment Que 2017;42(1):165-182. [doi: 10.7202/1040249ar]
10. Tong Y, Wang S, Cao L, et al. School dropouts related to mental disorders: a systematic review and meta-analysis. Asian J Psychiatr 2023 Jul;85:103622. [doi: 10.1016/j.ajp.2023.103622] [Medline: 37172421]
11. Duffy A, Keown-Stoneman C, Goodday S, et al. Predictors of mental health and academic outcomes in first-year university students: identifying prevention and early-intervention targets. BJPsych Open 2020 May 8;6(3):e46. [doi: 10.1192/bjo.2020.24] [Medline: 32381150]
12. Gili M, Castellví P, Vives M, et al. Mental disorders as risk factors for suicidal behavior in young people: a meta-analysis and systematic review of longitudinal studies. J Affect Disord 2019 Feb 15;245:152-162. [doi: 10.1016/j.jad.2018.10.115] [Medline: 30390504]
13. Lipson SK, Eisenberg D. Mental health and academic attitudes and expectations in university populations: results from the healthy minds study. J Ment Health 2018 Jun;27(3):205-213. [doi: 10.1080/09638237.2017.1417567] [Medline: 29265935]
14. Markoulakis R, Kirsh B. Difficulties for university students with mental health problems: a critical interpretive synthesis. Rev Higher Education 2013 Sep;37(1):77-100. [doi: 10.1353/rhe.2013.0073]
15. Gallais B, Blackburn M, Paré J, Maltais A, Brassard H. Adaptation psychologique et adaptation aux études à distance des étudiants collégiaux face à la crise de la COVID-19 [Report in French]. : Écobes-Recherche et transfert, Cégep de Jonquière; 2022 URL: https://ecobes.cegepjonquiere.ca/media/tinymce/RapportAdapCovidPhase1_1.pdf [accessed 2025-01-25]
16. Enquête « sous ta façade »: enquête panquébécoise sur la santé psychologique étudiante [Report in French]. : Union étudiante du Québec; 2019 URL: https://unionetudiante.ca/Media/publicDocuments/Rapport-enquete-sous-ta-facade.pdf [accessed 2025-01-25]
17. Rapport de l'enquête nationale: derrière ton écran. Une enquête nationale sur les impacts de la COVID-19 sur la condition étudiante au collégial [Report in French]. : Fédération étudiante collégial du Québec; 2021 URL: https://stationsme.ca/wp-content/uploads/2023/11/Rapport-final-DTE_109eCo_Zoom.pdf [accessed 2025-01-25]
18. Un plan pour nous: enquête courte sur la santé psychologique étudiante au Québec en temps de pandémie de la COVID-19 [Report in French]. : Union étudiante du Québec; 2021 URL: https://unionetudiante.ca/Media/publicDocuments/c7b08d33-a8ad-4f9b-9c2f-ded46f2e6d4a.pdf [accessed 2025-01-25]
19. Frippiat D, Marquis N. Les enquêtes par Internet en sciences sociales: un état des lieux [Article in French]. Population (Paris) 2010 Nov 29;65(2):309-338. [doi: 10.3917/popu.1002.0309]
20. Lane J, Therriault D, Dupuis A, et al. The impact of the COVID-19 pandemic on the anxiety of adolescents in Québec. Child Youth Care Forum 2022;51(4):811-833. [doi: 10.1007/s10566-021-09655-9] [Medline: 34658612]
21. Brown P. The invisible problem? Improving students' mental health. : Higher Education Policy Institute; 2016 URL: https://www.hepi.ac.uk/wp-content/uploads/2016/09/STRICTLY-EMBARGOED-UNTIL-22-SEPT-Hepi-Report-88-FINAL.pdf [accessed 2025-01-25]
22. Cooke M, Huntley J. White paper on postsecondary student mental health. : The Coordinating Committee of Vice Presidents Students; 2015 URL: https://occccco.files.wordpress.com/2015/05/ccvps-white-paper-on-postsecondary-student-mental-health-april-2015.pdf [accessed 2025-02-25]
23. Gallais B, Bikie Bi Nguema N, Turcotte A. Portrait de l'intervention psychosociale dans le réseau collégial québécois: point de vue des intervenants sur leur pratique, leurs collaborations et sur les programmes d'accompagnement psychologique existants [Report in French]. : ÉCOBES – Recherche et transfert, Cégep de Jonquière; 2021 URL: https://ecobes.cegepjonquiere.ca/media/tinymce/Rapport_IntervenantsPsy_Cegeps_Gallais_2021_FINAL_2.pdf [accessed 2025-01-25]
24. Conley CS, Shapiro JB, Kirsch AC, Durlak JA. A meta-analysis of indicated mental health prevention programs for at-risk higher education students. J Couns Psychol 2017 Mar;64(2):121-140. [doi: 10.1037/cou0000190] [Medline: 28277730]
25. Towards a comprehensive mental health strategy: the crucial role of colleges and universities as partners. : Ontario University & College Health Association; 2009 URL: http://oucha.ca/pdf/mental_health/2009_12_OUCHA_Mental_Health_Report.pdf [accessed 2025-01-25]
26. Bamber MD, Morpeth E. Effects of mindfulness meditation on college student anxiety: a meta-analysis. Mindfulness 2019 Feb;10(2):203-214. [doi: 10.1007/s12671-018-0965-5]

XSL•FO
RenderX

27. Breedvelt JJF, Amanvermez Y, Harrer M, et al. The effects of meditation, yoga, and mindfulness on depression, anxiety, and stress in tertiary education students: a meta-analysis. Front Psychiatry 2019;10:193. [doi: 10.3389/fpsyt.2019.00193] [Medline: 31068842]
28. Dawson AF, Brown WW, Anderson J, et al. Mindfulness-based interventions for university students: a systematic review and meta-analysis of randomised controlled trials. Appl Psychol Health Well Being 2020 Jul;12(2):384-410. [doi: 10.1111/aphw.12188] [Medline: 31743957]
29. Halladay JE, Dawdy JL, McNamara IF, et al. Mindfulness for the mental health and well-being of post-secondary students: a systematic review and meta-analysis. Mindfulness 2019 Mar;10(3):397-414. [doi: 10.1007/s12671-018-0979-z]
30. Huang J, Nigatu YT, Smail-Crevier R, Zhang X, Wang J. Interventions for common mental health problems among university and college students: a systematic review and meta-analysis of randomized controlled trials. J Psychiatr Res 2018 Dec;107:1-10. [doi: 10.1016/j.jpsychires.2018.09.018] [Medline: 30300732]
31. Lattie EG, Adkins EC, Winquist N, Stiles-Shields C, Wafford QE, Graham AK. Digital mental health interventions for depression, anxiety, and enhancement of psychological well-being among college students: systematic review. J Med Internet Res 2019 Jul 22;21(7):e12869. [doi: 10.2196/12869] [Medline: 31333198]
32. Yusufov M, Nicoloro-SantaBarbara J, Grey NE, Moyer A, Lobel M. Meta-analytic evaluation of stress reduction interventions for undergraduate and graduate students. Int J Stress Manag 2019;26(2):132-145. [doi: 10.1037/str0000099]
33. Ebersold S, Cabral LSA. Enseignement supérieur, orchestration de l'accessibilité et stratégies d'accompagnement [Article in French]. Éducation francophonie 2016;44(1):134-153. [doi: 10.7202/1036176ar]
34. Nsamba A, Makoe M. Evaluating quality of students' support services in open distance learning. Turk Online J Distance Educ 2017;18(4):91-103. [doi: 10.17718/tojde.340391]
35. Amanvermez Y, Rahmadiana M, Karyotaki E, et al. Stress management interventions for college students: a systematic review and meta-analysis. Clin Psychol Sci Pract 2023;30(4):423-444. [doi: 10.1111/cpsp.12342]
36. Harrer M, Adam SH, Baumeister H, et al. Internet interventions for mental health in university students: a systematic review and meta-analysis. Int J Methods Psychiatr Res 2019 Jun;28(2):e1759. [doi: 10.1002/mpr.1759] [Medline: 30585363]
37. Massuard M, Lauzier-Jobin F. Stimuler les liens sociaux favorables a` la sante´ mentale e´tudiante [Report in French]. : Observatory on Student Mental Health in Higher Education in Quebec; 2024 URL: https://bp-net.ca/wp-content/uploads/2025/04/OSMEES_Guide_liens_sociaux_2025_01_14.pdf [accessed 2025-05-03]
38. Beasy K, Emery S, Dyer L, et al. Writing together to foster wellbeing: doctoral writing groups as spaces of wellbeing. High Educ Res Dev 2020 Sep 18;39(6):1091-1105. [doi: 10.1080/07294360.2020.1713732]
39. Boulter J, Orozco Morales ML, Principe N, Tilsed CM. What is kindness in science and why does it matter? Immunol Cell Biol 2023 Feb;101(2):97-103. [doi: 10.1111/imcb.12580] [Medline: 36006827]
40. Driscoll DL, Leigh SR, Zamin NF. Self-care as professionalization: a case for ethical doctoral education in composition studies. Coll Compos Commun 2020 Feb 1;71(3):453-480. [doi: 10.58680/ccc202030503]
41. Jones CH, Whittle R. Researcher self-care and caring in the research community. Area 2021 Jun;53(2):381-388. [doi: 10.1111/area.12703]
42. Brown AD, Ross N, Sangraula M, Laing A, Kohrt BA. Transforming mental healthcare in higher education through scalable mental health interventions. Glob Ment Health (Camb) 2023;10:e33. [doi: 10.1017/gmh.2023.29] [Medline: 37854434]
43. Herlitz L, MacIntyre H, Osborn T, Bonell C. The sustainability of public health interventions in schools: a systematic review. Implement Sci 2020 Jan 6;15(1):4. [doi: 10.1186/s13012-019-0961-8] [Medline: 31906983]
44. Owens JS, Lyon AR, Brandt NE, et al. Implementation science in school mental health: key constructs in a developing research agenda. School Ment Health 2014 May 1;6(2):99-111. [doi: 10.1007/s12310-013-9115-3] [Medline: 26413173]
45. Schultes MT. An introduction to implementation evaluation of school-based interventions. Eur J Dev Psychol 2023 Jan 2;20(1):189-201. [doi: 10.1080/17405629.2021.1976633]
46. Observatoire sur la santé mentale étudiante en enseignement supérieur. 2025. URL: https://osmees.ca [accessed 2025-02-02]
47. Plan d'action sur la santé mentale étudiante en enseignement supérieur 2021-2026 [Report in French]. : Ministère de l'Enseignement supérieur, Gouvernement du Québec; 2021 URL: https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/education/publications-adm/enseignement-superieur/PASME.pdf [accessed 2025-01-25]
48. Fazal N, Jackson SF, Wong K, Yessis J, Jetha N. Élaboration de critères définissant les pratiques prometteuses en promotion de la santé et en prévention des maladies pour le Portail canadien des pratiques exemplaires [Article in French]. Promot Santé Prévent Mal Chron Can 2017;37(11):433-440. [doi: 10.24095/hpcdp.37.11.03f]
49. Laloux F. Reinventing Organizations: A Guide to Creating Organizations Inspired by the Next Stage of Human Consciousness, 1st edition: Nelson Parker; 2014. URL: https://www.goodreads.com/book/show/20787425-reinventing-organizations [accessed 2026-03-12]
50. Ecclestone A, Linden B, Monaghan C, Zheng S. Canada's Student Mental Health Network: protocol for a comprehensive program evaluation. JMIR Res Protoc 2023 Jun 22;12:e41521. [doi: 10.2196/41521] [Medline: 37347511]
51. Ecclestone A, Linden B, Rose J, Kullar K. Mobilizing health promotion through Canada's Student Mental Health Network: concurrent, mixed methods process evaluation. JMIR Form Res 2025 Feb 27;9(1):e58992. [doi: 10.2196/58992] [Medline: 40014047]

XSL•FO
RenderX

52. Akintobi TH, Barrett R, Hoffman L, et al. The community engagement course and action network: strengthening community and academic research partnerships to advance health equity. Front Public Health 2023;11:1114868. [doi: 10.3389/fpubh.2023.1114868] [Medline: 37404270]
53. Comprehensive mental health action plan 2013-2030. : World Health Organization; 2013 URL: https://iris.who.int/server/api/core/bitstreams/69921758-6229-49ba-bd3d-c24736e35829/content [accessed 2025-01-25]
54. Keyes CLM. Mental illness and/or mental health? Investigating axioms of the complete state model of health. J Consult Clin Psychol 2005 Jun;73(3):539-548. [doi: 10.1037/0022-006X.73.3.539] [Medline: 15982151]
55. Knoll AD, MacLennan RN. Prevalence and correlates of depression in Canada: findings from the Canadian Community Health Survey. Can J Psychol 2017;58(2):116-123. [doi: 10.1037/cap0000103]
56. Provencher HL, Keyes CLM. Complete mental health recovery: bridging mental illness with positive mental health. J Public Ment Health 2011 Mar 15;10(1):57-69. [doi: 10.1108/17465721111134556]
57. Iasiello M, van Agteren J. Mental health and/or mental illness: a scoping review of the evidence and implications of the dual-continua model of mental health. Evid Base 2020;2020:1-45. [doi: 10.21307/eb-2020-001]
58. Vaingankar JA, Chong SA, Abdin E, et al. Understanding the relationships between mental disorders, self-reported health outcomes and positive mental health: findings from a national survey. Health Qual Life Outcomes 2020 Mar 4;18(1):55. [doi: 10.1186/s12955-020-01308-0] [Medline: 32131837]
59. Westerhof GJ, Keyes CLM. Mental illness and mental health: the two continua model across the lifespan. J Adult Dev 2010 Jun;17(2):110-119. [doi: 10.1007/s10804-009-9082-y] [Medline: 20502508]
60. Peter T, Roberts LW, Dengate J. Flourishing in life: an empirical test of the Dual Continua Model of Mental Health and Mental Illness among Canadian university students. Int J Ment Health Promot 2011 Jan;13(1):13-22. [doi: 10.1080/14623730.2011.9715646]
61. Cadre de référence sur la santé mentale étudiante [Report in French]. : Ministère de l'Enseignement supérieur, Gouvernement du Québec; 2022 URL: https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/education/publications-adm/enseignement-superieur/PASME_cadre-reference.pdf [accessed 2025-01-25]
62. Côté JE. The enduring usefulness of Erikson's concept of the identity crisis in the 21st century: an analysis of student mental health concerns. Identity 2018 Oct 2;18(4):251-263. [doi: 10.1080/15283488.2018.1524328]
63. Co^te´ JE. Youth Development in Identity Societies: Paradoxes of Purpose: Routledge; 2019. URL: https://www.taylorfrancis.com/books/mono/10.4324/9780429433856/youth-development-identity-societies-james-cote [accessed 2026-03-12] [doi: 10.4324/9780429433856]
64. Andersen R, Holm A, Côté JE. The student mental health crisis: assessing psychiatric and developmental explanatory models. J Adolesc 2021 Jan;86:101-114. [doi: 10.1016/j.adolescence.2020.12.004] [Medline: 33429156]
65. Bantjes J, Hunt X, Stein DJ. Public health approaches to promoting university students' mental health: a global perspective. Curr Psychiatry Rep 2022 Dec;24(12):809-818. [doi: 10.1007/s11920-022-01387-4] [Medline: 36399235]
66. Bantjes J, Hunt X, Stein DJ. Anxious, depressed, and suicidal: crisis narratives in university student mental health and the need for a balanced approach to student wellness. Int J Environ Res Public Health 2023 Mar 9;20(6):4859. [doi: 10.3390/ijerph20064859] [Medline: 36981766]
67. Kerig PK, Alexander AR. Models of psychopathology. In: Troop-Gordon W, Neblett E, editors. Encyclopedia of Adolescence, 2nd edition, Vol. 2024:321-333. [doi: 10.1016/B978-0-323-96023-6.00040-3]
68. Kerig PK. The developmental psychopathology approach. In: Kerig PK, Ludlow A, Wenar C, editors. Developmental Psychopathology: From Infancy Through Adolescence, 6th edition: McGraw-Hill Education; 2012:1-33.
69. Compton MT, Shim RS. The social determinants of mental health. FOCUS 2015 Oct;13(4):419-425. [doi: 10.1176/appi.focus.20150017]
70. Compton MT, Shim RS. Why employers must focus on the social determinants of mental health. Am J Health Promot 2020 Feb;34(2):215-219. [doi: 10.1177/0890117119896122c] [Medline: 31931601]
71. Wickham SR, Amarasekara NA, Bartonicek A, Conner TS. The big three health behaviors and mental health and well-being among young adults: a cross-sectional investigation of sleep, exercise, and diet. Front Psychol 2020;11:579205. [doi: 10.3389/fpsyg.2020.579205] [Medline: 33362643]
72. Sutton A. Living the good life: a meta-analysis of authenticity, well-being and engagement. Pers Individ Dif 2020 Jan;153:109645. [doi: 10.1016/j.paid.2019.109645]
73. Schwartz SJ, Zamboanga BL, Luyckx K, Meca A, Ritchie RA. Identity in emerging adulthood: reviewing the field and looking forward. Emerg Adulthood 2013;1(2):96-113. [doi: 10.1177/2167696813479781]
74. Padilla-Walker LM, Memmott-Elison MK, Nelson LJ. Positive relationships as an indicator of flourishing during emerging adulthood. In: Padilla-Walker LM, Nelson LJ, editors. Flourishing in Emerging Adulthood: Positive Development during the Third Decade of Life: Oxford University Press; 2017:212-236. [doi: 10.1093/acprof:oso/9780190260637.003.0018]
75. Thaivalappil A, Stringer J, Burnett A, Papadopoulos A. Institutional factors affecting postsecondary student mental wellbeing: a scoping review of the Canadian literature. Psych 2023;5(3):630-649. [doi: 10.3390/psych5030040]
76. Patton MQ. Qualitative Research & Evaluation Methods, 4th edition: SAGE Publications; 2015. URL: https://us.sagepub.com/en-us/nam/qualitative-research-evaluation-methods/book232962 [accessed 2026-03-12]

XSL•FO
RenderX

77. Fleming J. Young people's involvement in research: still a long way to go? Qual Soc Work 2011;10(2):207-223. [doi: 10.1177/1473325010364276]
78. Lauzier-Jobin F. Évaluation participative guidée par la théorie de programme: retour sur une recherche évaluative des pratiques de trois organismes communautaires venant en aide aux jeunes en situation d'itinérance [Article in French]. Presented at: Proceedings of the Colloque étudiant 2016 sur la recherche partenariale; Nov 25, 2016 URL: https://chairerp.uqam.ca/fichier/document/Actes_du_colloque_26_juil._2017.pdf [accessed 2025-01-25]
79. Barr ML, McNamara J. Community-based participatory research: partnering with college students to develop a tailored, wellness-focused intervention for University Campuses. Int J Environ Res Public Health 2022 Dec 6;19(23):16331. [doi: 10.3390/ijerph192316331] [Medline: 36498404]
80. Prince M. Does active learning work? A review of the research. J Engineering Edu 2004 Jul;93(3):223-231. [doi: 10.1002/j.2168-9830.2004.tb00809.x]
81. Cook-Sather A. Student voice across contexts: fostering student agency in today's schools. Theory Pract 2020 Apr 2;59(2):182-191. [doi: 10.1080/00405841.2019.1705091]
82. Paturel D. La recherche participative comme enjeu pour la recherche en travail social [Article in French]. Le sociographe 2014;7(5):103-124. [doi: 10.3917/graph.hs07.0103]
83. Kryczka A. Job satisfaction in the context of TEAL Organization. Acta Univ Nicolai Copernici Zarz 2019;46(3):29-40. [doi: 10.12775/AUNC_ZARZ.2020.012]
84. Raetzer S, Self RL, Klein N. Closing the gap: reinventing organizations for innovation and the disruptors who drive it. CERN IdeaSquare J Exp Innov 2019;3(2):48-57. [doi: 10.23726/CIJ.2019.920]
85. Rzepka A. TEAL organizations in times of Industry 4.0. Eur Res Stud J 2021 Jun 1;XXIV(Special Issue 2):60-71. [doi: 10.35808/ersj/2191]
86. Gagnon S, Mailhot C, Ziam S. The role and contribution of an intermediary organisation in the implementation of an interactive knowledge transfer model. Evid Policy 2019;15(1):07-29. [doi: 10.1332/174426418X15166967955544]
87. Wyrzykowska B. Teal organizations: literature review and future research directions. Cent Eur Manag 2019 Dec 15;27(4):124-141. [doi: 10.7206/cemj.2658-0845.12]
88. Bannister J, O'Sullivan A. Knowledge mobilisation and the civic academy: the nature of evidence, the roles of narrative and the potential of contribution analysis. Contemp Soc Sci 2013 Nov;8(3):249-262. [doi: 10.1080/21582041.2012.751497]
89. Cooper A, Ben L. Some Canadian contributions to understanding knowledge mobilisation. Evid Policy 2010;6(3):351-369. [doi: 10.1332/174426410X524839]
90. Massuard M, Manceau LM, Lane J. Guide de mise en œuvre d'une infrastructure de soutien à l'implantation et au transfert de connaissance – soutien à l'implantation de pratiques prometteuses et au transfert de connaissances dans le réseau de la santé et des services sociaux [Report in French]. : Direction des communications, ministère de la Santé et des Services sociaux; 2023 URL: https://publications.msss.gouv.qc.ca/msss/fichiers/2023/23-516-01W.pdf [accessed 2025-01-25]
91. Curran JA, Grimshaw JM, Hayden JA, Campbell B. Knowledge translation research: the science of moving research into policy and practice. J Contin Educ Health Prof 2011;31(3):174-180. [doi: 10.1002/chp.20124] [Medline: 21953658]
92. Haynes A, Rowbotham S, Grunseit A, et al. Knowledge mobilisation in practice: an evaluation of the Australian Prevention Partnership Centre. Health Res Policy Sys 2020 Dec;18(1):13. [doi: 10.1186/s12961-019-0496-0]
93. Sketris IS, Carter N, Traynor RL, Watts D, Kelly K, On behalf of the following contributing members of the CNODES Knowledge Translation Team: Pierre Ernst, JM Gamble, Brenda Hemmelgarn, Colleen Metge, Michael Paterson, Robert Platt W. and Gary Teare. Building a framework for the evaluation of knowledge translation for the Canadian Network for Observational Drug Effect Studies. Pharmacoepidemiol Drug Saf 2020 Jan;29(S1):8-25. [doi: 10.1002/pds.4738]
94. Armitage D, Berkes F, Dale A, Kocho-Schellenberg E, Patton E. Co-management and the co-production of knowledge: learning to adapt in Canada's Arctic. Glob Environ Change 2011 Aug;21(3):995-1004. [doi: 10.1016/j.gloenvcha.2011.04.006]
95. Dawes SS, Cresswell AM, Pardo TA. From "need to know" to "need to share": tangled problems, information boundaries, and the building of public sector knowledge networks. Public Adm Rev 2009 May;69(3):392-402. [doi: 10.1111/j.1540-6210.2009.01987_2.x]
96. Gagliardi AR, Webster F, Brouwers MC, Baxter NN, Finelli A, Gallinger S. How does context influence collaborative decision-making for health services planning, delivery and evaluation? BMC Health Serv Res 2014 Nov 19;14(1):545. [doi: 10.1186/s12913-014-0545-x] [Medline: 25407487]
97. Green LW. Making research relevant: if it is an evidence-based practice, where's the practice-based evidence? Fam Pract 2008 Dec;25 Suppl 1(Suppl 1):i20-i24. [doi: 10.1093/fampra/cmn055] [Medline: 18794201]
98. Fitzgerald L, Harvey G. Translational networks in healthcare? Evidence on the design and initiation of organizational networks for knowledge mobilization. Soc Sci Med 2015 Aug;138:192-200. [doi: 10.1016/j.socscimed.2015.06.015] [Medline: 26117554]
99. Kasonde JM, Campbell S. Creating a knowledge translation platform: nine lessons from the Zambia Forum for Health Research. Health Res Policy Sys 2012 Dec;10(1):31. [doi: 10.1186/1478-4505-10-31]
100. MacKillop E, Quarmby S, Downe J. Does knowledge brokering facilitate evidence-based policy? A review of existing knowledge and an agenda for future research. Policy Polit 2020;48(2):335-353. [doi: 10.1332/030557319X15740848311069]

XSL•FO
RenderX

101. Rycroft-Malone J, Burton C, Wilkinson J, et al. Collective action for implementation: a realist evaluation of organisational collaboration in healthcare. Implementation Sci 2015 Dec;11(1):17. [doi: 10.1186/s13012-016-0380-z]
102. Best A, Holmes B. Systems thinking, knowledge and action: towards better models and methods. Evid Policy 2010;6(2):145-159. [doi: 10.1332/174426410X502284]
103. Elizabeth K, Wallace A, Crosland A. A model for collaborative working to facilitate knowledge mobilisation in public health. Evid Policy 2015;11(4):559-576. [doi: 10.1332/174426415X14298767491657]
104. Gagliardi AR, Berta W, Kothari A, Boyko J, Urquhart R. Integrated knowledge translation (IKT) in health care: a scoping review. Implementation Sci 2015 Dec;11(1):38. [doi: 10.1186/s13012-016-0399-1]
105. Tantchou Dipankui M. Collaboration between researchers and knowledge users in health technology assessment: a qualitative exploratory study. Int J Health Policy Manag 2016;6(8):437-446. [doi: 10.15171/ijhpm.2016.153] [Medline: 28812843]
106. McSween-Cadieux E, Lane J, Kadio K, Ridde V, Dagenais C. Un processus de transfert de connaissances pour favoriser l'utilisation des résultats d'évaluation [Book in French]. In: Rey L, Quesnel JS, Sauvin V, editors. L'Évaluation En Contexte de Développement: Enjeux, Approches, et Pratiques: Les Éditions JFD inc; 2022:441-464 URL: https://www.editionsjfd.com/static/uploaded/Files/9782897993016_v18_int.pdf [accessed 2026-03-12]
107. Albers B, Metz A, Burke K. Implementation support practitioners—a proposal for consolidating a diverse evidence base. BMC Health Serv Res 2020 May 1;20(1):368. [doi: 10.1186/s12913-020-05145-1] [Medline: 32357877]
108. Outil synthèse: les facteurs d'influence de la santé mentale étudiante, des leviers pour agir! [Report in French]. : Direction régionale de santé publique de Montréal, Gouvernement du Québec; 2024 URL: http://www.santecom.qc.ca/bibliothequevirtuelle/Montreal/9782550970927.pdf [accessed 2025-01-25]
109. Proposition d'un modèle conceptuel concernant la surveillance de la santé mentale, des troubles mentaux courants et de leurs déterminants au québec. État des connaissances: synthèse des connaissances [Report in French]. : Institut national de santé publique du Québec; 2023 URL: https://www.inspq.qc.ca/sites/default/files/2023-07/3365-cadre-conceptuel-surveillance-sante-mentale-troubles-mentaux.pdf [accessed 2025-01-25]
110. Hutchings A, Raine R, Sanderson C, Black N. A comparison of formal consensus methods used for developing clinical guidelines. J Health Serv Res Policy 2006 Oct;11(4):218-224. [doi: 10.1258/135581906778476553] [Medline: 17018195]
111. Low GK, Subedi S, Omosumwen OF, et al. Development and validation of observational and qualitative study protocol reporting checklists for novice researchers (ObsQual checklist). Eval Program Plann 2024 Oct;106:102468. [doi: 10.1016/j.evalprogplan.2024.102468] [Medline: 39029287]
112. Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C, Mittman B. Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. Adm Policy Ment Health 2009 Jan;36(1):24-34. [doi: 10.1007/s10488-008-0197-4] [Medline: 19104929]
113. Damschroder LJ. Clarity out of chaos: use of theory in implementation research. Psychiatry Res 2020 Jan;283:112461. [doi: 10.1016/j.psychres.2019.06.036] [Medline: 31257020]
114. Ben Charif A, Zomahoun HTV, LeBlanc A, et al. Effective strategies for scaling up evidence-based practices in primary care: a systematic review. Implement Sci 2017 Nov 22;12(1):139. [doi: 10.1186/s13012-017-0672-y] [Medline: 29166911]
115. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
116. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation Sci 2009 Dec;4(1):50. [doi: 10.1186/1748-5908-4-50]
117. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
118. Tabachnick BG, Fidell LS. Using Multivariate Statistics, 7th edition: Pearson; 2019. URL: https://www.pearsonhighered.com/assets/preface/0/1/3/4/0134790545.pdf [accessed 2026-03-12]
119. Field AP. Discovering Statistics Using IBM SPSS Statistics, 4th edition: SAGE Publications Ltd; 2013. URL: https://sadbhavnapublications.org/research-enrichment-material/2-Statistical-Books/Discovering-Statistics-Using-IBM-SPSS-Statistics-4th-c2013-Andy-Field.pdf [accessed 2026-03-12]
120. Muthén LK, Muthén BO. Mplus: Statistical Analysis With Latent Variables: User's Guide, 8th edition: Muthe´n & Muthe´n; 2017. URL: https://www.amazon.com/Mplus-Statistical-Analysis-Latent-Variables/dp/0982998325 [accessed 2026-03-12]
121. Paillé P, Mucchielli A. L'analyse Qualitative en Sciences Humaines et Sociales [Book in French], 4th edition: Armand Colin; 2016. URL: https://shs.cairn.info/analyse-qualitative-en-sciences-humaines-et-social--9782200614706?lang=fr [accessed 2026-03-12]
122. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
123. Ziam S, Lanoue S, McSween-Cadieux E, et al. A scoping review of theories, models and frameworks used or proposed to evaluate knowledge mobilization strategies. Health Res Policy Syst 2024 Jan 10;22(1):8. [doi: 10.1186/s12961-023-01090-7] [Medline: 38200612]

XSL•FO
RenderX

124. Yin RK. Case Study Research and Applications: Design and Methods, 6th edition: SAGE Publications Inc; 2017. URL: https://www.amazon.in/Case-Study-Research-Applications-Methods/dp/1506336167 [accessed 2026-03-12]
125. Patton MQ. Principles-Focused Evaluation: The Guide: Guilford Press; 2017. URL: https://www.amazon.in/Principles-Focused-Evaluation-Michael-Quinn-Patton/dp/1462531822 [accessed 2026-03-12]
126. Patton MQ. Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use: Guilford Press; 2011. URL: https://www.routledge.com/Developmental-Evaluation-Applying-Complexity-Concepts-to-Enhance-Innovation-and-Use/Patton/p/book/9781606238721 [accessed 2026-03-12]
127. Rey L, Brousselle A, Dedobbeleer N, Tremblay MC. Les défis de l'évaluation développementale en recherche : une analyse d'implantation d'un projet «hôpital promoteur de santé» [Article in French]. Can J Program Eval 2013 Mar;28(1):1-26. [doi: 10.3138/cjpe.028.001]
128. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research, 3rd edition: SAGE Publications Inc; 2018. URL: https://www.vitalsource.com/products/designing-and-conducting-mixed-methods-research-john-w-creswell-vicki-l-v9781506394671?utm_source=chatgpt.com [accessed 2026-03-12]
129. Robson C, McCartan K. Real World Research, 4th edition: John Wiley & Sons Inc; 2016. URL: https://www.amazon.in/Real-World-Research-Colin-Robson/dp/111874523X [accessed 2026-03-12]
130. Ayres L. Semi-structured interview. In: Encyclopedia of Qualitative Research Methods: SAGE Publications; 2008:811-812 URL: https://sk.sagepub.com/ency/edvol/research/chpt/semistructured-interview#_ [accessed 2026-04-01]
131. Martineau S. L'observation en situation: enjeux, possibilités et limites [Article in French]. Rech qual 2005:5-17 [FREE Full text]
132. Gauthier B. Recherche Sociale: De La Problématique à La Collecte Des Données [Book in French], 5th edition: Presses de l'Université du Québec; 2009.
133. Gaudet S, Robert D. L'aventure de La Recherche Qualitative: Du Questionnement à La Rédaction Scientifique [Book in French]: Les Presses de l'Université d'Ottawa; 2018. URL: https://www.amazon.com/Laventure-recherche-qualitative-Praxis-French/dp/2760327310 [accessed 2026-03-12] [doi: 10.1515/9782760327320]
134. Bassett R. Computer-based analysis of qualitative data: NVIVO. In: Mills AJ, Durepos G, Wiebe E, editors. Encyclopedia of Case Study Research: SAGE Publications Inc, Vol. 2010:193-194. [doi: 10.4135/9781412957397.n71]
135. Broglia E, Barkham M. Adopting the principles and practices of learning health systems in universities and colleges: recommendations for delivering actionable data to improve student mental health. Cogent Ment Health 2024;3(1):2301339. [doi: 10.1080/28324765.2023.2301339] [Medline: 41262654]
136. Budgen C, Morrison H, Sullivan K, et al. Creating a healthier campus community using action research and health promotion strategies: students and organizational leaders as partners. Int J Health Wellness Soc 2011;1(3):155-176. [doi: 10.18848/2156-8960/CGP/v01i03/41183]
137. Hansen-Ketchum P, MacDonald J, Austen EL, Gilham CM, McKinnon M, Gougeon L. Healthy settings: barriers and facilitators to a healthy campus community. Int J Health Wellness Soc 2021;11(1):1-22. [doi: 10.18848/2156-8960/CGP/v11i01/1-22]
138. Lemire N, Souffez K, Laurendeau MC. Animer un processus de transfert des connaissances: bilan des connaissances et outil d'animation [Report in French]. : Institut National De Santé Publique Du Québec; 2009 URL: https://www.inspq.qc.ca/sites/default/files/publications/1012_animertransfertconn_bilan.pdf [accessed 2025-01-25]

Abbreviations

FRQ: Fonds de recherche du Québec
HE: higher education
HEI: higher education institution
KM: knowledge mobilization
KMI: knowledge mobilization infrastructure
MH: mental health
OSMHHE: Observatory on Student Mental Health in Higher Education
SDMH: social determinants of mental health

XSL•FO
RenderX

Edited by A Schwartz; submitted 30.Aug.2025; this is a non–peer-reviewed article;accepted 11.Feb.2026; published 22.Apr.2026.

Please cite as:
Lane J, Gallais B, Lauzier-Jobin F, Lanoue S, Guay-Dufour F, Paré-Beauchemin R, Guertin R, Bergeron-Leclerc C, Ziam S, Lallier Beaudoin MC, Shareck M, Boudreault V, Langelier MÈ, Geoffrion S, Beauregard N, B-Lamoureux B, Brodeur M, McSween-Cadieux E, Dupuis A, Smith J, Thibault I, Beauregard N
Establishment, Implementation, and Impacts of the Observatory on Student Mental Health in Higher Education in Quebec, Canada: Protocol for a Mixed Methods Research Program
JMIR Res Protoc 2026;15:e83225
URL: https://www.researchprotocols.org/2026/1/e83225
doi:10.2196/83225

© Julie Lane, Benjamin Gallais, François Lauzier-Jobin, Sèverine Lanoue, Félix Guay-Dufour, Rémi Paré-Beauchemin, Rachel Guertin, Christiane Bergeron-Leclerc, Saliha Ziam, Marie-Claude Lallier Beaudoin, Martine Shareck, Véronique Boudreault, Marie-Ève Langelier, Steve Geoffrion, Nancy Beauregard, Bianca B-Lamoureux, Magaly Brodeur, Esther McSween-Cadieux, Audrey Dupuis, Jonathan Smith, Isabelle Thibault, Noémie Beauregard. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Gut Bacterial Resistome in the First Two Years of Life: Protocol for a Longitudinal Observational Birth Cohort Study

Justine Fri[1*], PhD; Idris Njanje[1*], PhD; Tjale Cloupas Mahopo[1,2], PhD; Lufuno Grace Mavhandu-Ramarumo[1,3], PhD; Pascal Obong Bessong[1,4], PhD; Microbiome and Antimicrobial Resistance Collaborators[5,6*]

[1]SAMRC-UNIVEN Antimicrobial Resistance and Global Health Research Unit, Institute for Pathogen and Global Health Research, University of Venda, University Road, Private Bag X5050, Thohoyandou, South Africa

[2]Department of Nutrition, Faculty of Health Sciences, University of Venda, Thohoyandou, South Africa

[3]Department of Virology, Sefako Makgatho Health Sciences University, Pretoria, South Africa

[4]Center for Global Health Equity, University of Virginia, Charlottesville, VA, United States

[5]See Acknowledgments

[6]

[*]these authors contributed equally

Corresponding Author:
Pascal Obong Bessong, PhD
SAMRC-UNIVEN Antimicrobial Resistance and Global Health Research Unit, Institute for Pathogen and Global Health Research, University of Venda, University Road, Private Bag X5050, Thohoyandou, South Africa

Abstract

Background: Antimicrobial resistance (AMR) is a global health threat that increases the burden of infectious diseases and disproportionately affects communities of low socioeconomic status. Despite the call for community-level AMR data, prospective studies from rural sub-Saharan African communities to inform appropriate targeted interventions remain scarce. Given the role of enteric bacteria in AMR transmission dynamics, there is a need to understand the timing, risk factors, and ecological drivers of gut resistome acquisition and development during infancy.

Objective: This study aimed to characterize the temporal dynamics of enteric bacterial resistomes during the first 2 years of life and to identify drivers of AMR acquisition and development in a community-based, prospective, observational birth cohort study in a rural South African community.

Methods: The study aims to enroll 200 newborns and their mothers within 17 days post partum. Data on key exposures and variables include sociodemographics; perinatal and anthropometrics; feeding practices and dietary exposures; illness, medication, and vaccination history; breast milk metabolomic profiles; household socioeconomic status; maternal psychosocial and behavioral factors; hygiene and sanitation practices; and environmental exposures including hydro-meteorological variables, in-house livestock and pets, and drinking water quality. Biological samples include stools from monthly collections and diarrhea episodes for metagenomic analysis and breast milk for metabolomics. Planned analyses include assessing the infant microbiome and resistome structure (diversity, abundance, and composition) across time points and modeling associations between risk factors and AMR outcomes. Additionally, a cross-sectional community survey on knowledge, attitudes, and practices regarding antimicrobial use is conducted to inform knowledge translation through responsive dialogues, thereby developing ethnographically relevant packages for community-level AMR stewardship.

Results: Participant identification and enrollment began in August 2023. By October 2025, 167 newborns had been enrolled, with 20 having completed the 24-month follow-up. The characteristics of the enrolled participants are presented in this protocol.

Conclusions: This study will offer a unique opportunity to generate longitudinal resistome data from a rural sub-Saharan African setting. The study is expected to contribute knowledge on the microbiome and resistome structure dynamics and trajectories associated with key risk factors of acquisition and development. In addition, co-produced ethnographically tailored educational packages, informed by knowledge, attitudes, and practices and bacterial resistome data, will drive sustainable community-centered AMR awareness interventions.

International Registered Report Identifier (IRRID): DERR1-10.2196/86058

(JMIR Res Protoc 2026;15:e86058) doi:10.2196/86058

KEYWORDS

longitudinal birth cohort; antimicrobial resistance; bacterial gut resistome; ESKAPE pathogens; community AMR stewardship; Escherichia coli

XSL•FO
RenderX

Introduction

Bacterial antimicrobial resistance (AMR) is a major threat to the treatment of infectious diseases and a significant contributor to morbidity and mortality worldwide. AMR disproportionately affects low- and middle-income countries. A systematic analysis revealed that in 2019 alone, bacterial AMR was responsible for close to 1.3 million deaths globally [1] and 250,000 deaths in the World Health Organization (WHO) African region alone [2]. The effects of AMR are reflected in poor health outcomes, absenteeism from work and school, an increase in health care expenditure for families and the state, and a decline in food security [3,4]. A concerted effort to mitigate the spread of antibiotic resistance is of high priority, and it is required to achieve certain sustainable development goals and global health imperatives. However, more often, efforts to tackle this challenge are concentrated in health care settings. The WHO working group on pathogen priority list of antibiotic resistance has emphasized the dearth of prospective data on the burden of antibiotic resistance from communities in low- and middle-income countries, to guide discovery and interventions [5]. Available datasets have limited power to contribute to national and regional understanding of AMR burden at the community level, and for policy development [6]. In addition, sustaining efforts to mitigate AMR through AMR stewardship requires investments in communities for the co-production and implementation of evidence [7,8].

The nonjudicious use of antibiotics in human and animal health, and the application of antibiotics in agriculture, drive the development of antibiotic resistance [9]. Poor infection control and prevention, poor hygiene and sanitation, and inadequate wastewater treatment and disposal contribute to the spread of AMR in the environment and its transmission to humans. Of interest is the relationship between changing meteorological conditions, such as higher temperatures, and the spread of pathogens, favoring human infection by hitherto mainly zoonotic microorganisms [10]. A multipronged approach is desirable to effectively mitigate AMR, given that it is driven by multiple factors impacting several dimensions of human life [11-13].

In a community-based cross-sectional study, we observed multiclass antibiotic-resistant *Escherichia coli* in children less than 4 months old with no prior exposure to antibiotics [14]. In the same community, it was later shown that enteroaggregative *E coli* was the most frequently exposed pathogen to antibiotics [15]. We have also identified bacterial taxa that could serve as indicators to monitor the repair of gut microbiota in malnourished children [16]. The human gut is a complex microbial ecosystem in which antibiotic resistance genes (ARGs) are embedded within microbial communities. A bacterial resistome may be shaped by the surrounding (“background”) microbial communities, as complex interactions between different bacterial species within these communities can create opportunities for horizontal gene transfer and contribute to resistome enrichment [17]. It is therefore plausible that the composition of the gut microbiota could influence the diversity and abundance of the gut bacterial resistome. To understand the risk factors and trends in enteric bacterial AMR, with the ability to measure the impact of interventions, *E coli* can be used innovatively as a model organism in apparently healthy populations. This is because *E coli* causes different diseases that require management with antibiotics, it is less fastidious, patterns of its resistome vary geographically, and it is a designated AMR priority organism.

For a better understanding of the landscape of enteric AMR to inform targeted interventions in communities of low socioeconomic status, the following questions need to be addressed: (1) At what age in a community of low socioeconomic status does the acquisition of enteric resistant *E coli* and other clinically relevant bacteria occur? (2) What is the composition of *E coli* resistomes and associated microbiota? Considering that several variables could contribute to the acquisition of enteric AMR, a plausible hypothesis is that the acquisition and carriage of resistant *E coli* and ESKAPE (*E coli* and *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter spp*) bacteria begin early in life, and that colonization with resistant organisms is influenced by background microbiota and external exposures.

Reports have shown a strong association between hydrometeorological variables and enteric viral infections, and that earth observation data products, available at a global scale and at subdaily resolution, can be combined with longitudinal surveillance data to test hypotheses about the routes and drivers of transmission [18]. Also, it has been reported that the risk of infection with enteropathogens increases with an increase in soil moisture, and that humidity is associated with increases in bacterial infections [19]. If hydrometeorological variables impact the transmission of enteropathogens, they may also affect transmission dynamics and diversity of the enteric bacterial resistome. Therefore, do hydrometeorological indicators affect trends in the acquisition of community-level enteric resistomes? Using *E coli* as a model for the gut resistome, this question will be investigated through the following hypothesis: Seasonal changes in hydrometeorological indicators affect trends in the acquisition and composition of *E coli* resistome.

We have shown that rural low-socioeconomic communities can be prepared and organized to participate in intensive, prospective, observational studies [20]. Also, community-based projects can be leveraged to advance the sustainable development aspirations of communities [21]. Therefore, we hypothesize that engaged, involved, and mobilized communities are critical in the co-production of ethnographically relevant educational materials for antimicrobial resistance stewardship at the community level.

Considering the above hypotheses, the overarching aim of the study is to conduct a community-based, prospective birth cohort surveillance to understand the acquisition, composition, and evolution of enteric bacteria resistome. The specific objectives include the following:

1. To establish a prospective birth cohort of 200 mother and child pair participants.
2. To determine bacteria gut resistomes (total gut resistome and ESKAPE resistomes) in the first 2 years of life.
3. To determine risk factors of bacterial gut resistome acquisition, considering exposures and covariates.

XSL•FO
RenderX

4. Using *E coli* as a model to determine the association of background gut microbiota on gut resistome.
5. Using *E coli* as a proxy to study gut resistome to determine the impact of seasonal hydrometeorological indicators on *E coli* resistome.
6. To co-produce, with study communities, through knowledge translation and application, ethnographically relevant community-oriented AMR educational and awareness packages for community-level AMR stewardship.

The study, therefore, adopts a longitudinal observational cohort design with repeated follow-up from birth to 24 months of age, which allows direct observation of the timing, acquisition, and drivers of the infant gut resistome. Longitudinal mixed-effects and compositional analytical approaches are used to model temporal trajectories and assess associations between exposures and resistome outcomes. Shotgun metagenomic sequencing of repeated stool samples provides comprehensive outcome measures, including ARGs (abundance, diversity, and composition), as well as resistomes of sentinel organisms (*E coli* and ESKAPE pathogens) of public health importance. Detailed longitudinal data on maternal, infant, household, environmental, and hydrometeorological exposures are collected to align with the hypothesized drivers of resistome acquisition. Leveraging *E coli* as a model organism enables linkage of its resistome profiles with background microbiota signatures and the potential drivers. This approach provides a framework for generating locally relevant evidence on early-life acquisition of AMR with implications for AMR stewardship strategies at the community level.

Methods

Study Design

The primary design of the study is a prospective, observational birth cohort study to describe the acquisition, patterns, and burden of gut bacterial resistome in the first 24 months of life. This study protocol adheres to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for reporting observational studies [22], adapted to capture metagenomic-specific methodological details (STROBE-metagenomics) [23]. The STROBE checklist is provided in Checklist 1.

Study Setting

The study is conducted in the Dzimauli community in the Limpopo Province, northeastern South Africa. The geography and climatic conditions of Dzimauli have been previously described [20,24]. Briefly, Dzimauli is rural and approximately 50 km from Thohoyandou, the main city of the Vhembe district, Limpopo Province. The population predominantly practices subsistence agriculture and petty trading. The area experiences heavy rainfall and flooding due to variations in climatic conditions [24].

Sample Size

A sample size of 200 mother-child pairs was selected based on adequacy for the longitudinal study objectives. The cohort size is sufficient to estimate resistome carriage and burden, and to support mixed-effects and compositional analyses of resistome trajectories and associated exposures across repeated measurements at the community level. It also accommodates anticipated attrition (15% - 20%) over the 24-month period, while maintaining analytical robustness and feasibility. Although conducted in a small rural community, recruiting up to 200 participant pairs with extensive follow-up would enable population-level interpretation of findings or application of the study approach within a similar setting.

Study Participants and Recruitment

Pregnant women in their third trimester are identified through household and outpatient antenatal clinic visits. Upon consent, they are screened and enrolled in the study at the birth of the child. Inclusion criteria for newborns include (1) participant should be available in the study community for at least 6 months, (2) birth weight ≥1500 g, (3) absence of immediate postnatal hospitalization or treatment for severe illness, and (4) mothers should be aged at least 16 years. Infants are excluded if (1) the child is from a multiple pregnancy, (2) the mother has another child enrolled in the AMR study, (3) the mother plans to relocate from the community within 6 months of enrollment, or (4) the mother is unable to consent or declines participation.

Data Collection

Field workers were recruited from the study community and trained to screen and enroll participants and conduct surveillance activities. Data collection spans a 24-month period and includes structured surveillance questionnaires, biological samples, and clinical assessments to capture longitudinal information on maternal and child health, feeding practices, growth, socioeconomic status, hygiene, and environmental exposures [25]. A comprehensive summary of data collection domains, time points, and laboratory analyses is provided in Table 1. Briefly, at enrollment (≤17 d), maternal and child demographic, perinatal, and anthropometric data, as well as treatment, hospitalization, and breastfeeding practices, are collected. Child anthropometrics are subsequently collected monthly for 12 months and quarterly for another 12 months. The study adopted instruments developed and used in the same community by the MAL-ED (The Etiology, Risk Factors and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development) study [25]. However, for the purpose of this study, household visits were conducted once a week for 24 months, during which time data on illness, antibiotic use, and other treatments received, feeding patterns, and vaccines received were captured by a 7-day recall questionnaire.

XSL•FO
RenderX

Table . Schedule of data collection measures, time points, participant groups, and outcomes in the 24-month longitudinal birth cohort study.

Measures	Time point	Group	Type of assay/analysis/outcome
Surveillance			
Recruitment, screening, and enrollment	≤17 days	Mother, child	Inclusion in a prospective study
Baseline demographic	≤17 days	Mother, child	Maternal and child demographics, perinatal and anthropometrics
Socioeconomic characteristics	6, 12, 18, and 24 months	Mother	Questionnaire on socioeconomic variables
Maternal assessment	≤17 days, 3, 6, 9, 12, 15, 18, 21, and 24 Months	Mother	Anthropometry, medication, and behavioral characteristics
Hygiene practices	1 - 12, 15, 18, 21, and 24 months	Mother	Measures of food, water, and hygienic practices
Sanitation	6, 12, 18, and 24 months	—	Measures of sanitation
Hydrometeorological variables	Monthly from 1 - 24months	—	Mean monthly precipitation volume, daily total surface runoff, soil moisture, surface pressure, wind speed, relative humidity, solar radiation, specific humidity, and average daily temperatures
Nutrition and health			
Food security	≤17 days, 6, 12, 18, and 24 months	Mother	Household food security variables
Infant Feeding	1 - 24 months	Child	Report on infant feeding practices Exclusive breastfeeding Complementary Report on dietary diversity
Infant Food Diary	9 - 12, 15, 18, and 24 months	Child	Record of infant food consumption in 24 h period, food sources, nutrient intake and adequacy
Child Growth Measures	1 - 12-, 15, 18, 21, and 24 months	Child	Infant growth trajectory
Psychosocial	7, 15, and 24 months	Mother	Perceived stress, depression symptoms, and anxiety
Vaccine and treatment	1 - 12-, 15, 18, 21, and 24 months	Child	Vaccination compliance and medication history
Laboratory assessment			
Stool	1-week, 3-, 6-, 12-, and 18-month	Child, Mother	Gut resistome
Breastmilk	1 - 12-, 15, 18, 21, and 24 months	Mother	Metabolomics
Drinking water sample	Once during the study	—	Household drinking water quality and zoonotic *E coli* resistance
Saliva	Once between 1 - 2 years	Child	Drug transport and metabolism
In-house livestock and pet fecal samples	Once during the study	—	Zoonotic *E coli* resistance

Surveillance data are collected with the mother or primary caregiver at various timepoints as detailed in Table 1. For infants aged 1 - 8 months, a qualitative 24-hour dietary recall is adapted from the Demographic and Health Survey, which was developed for the same community as part of the MAL-ED study [25]. The qualitative dietary recall complements the once-weekly dietary reports by capturing a broader variety of foods consumed, including specific fruits and vegetables, grains, roots, and animal source foods. A quantitative 24-hour recall, used in the same community from the MAL-ED study [26] was also used from 9 to 24 months, to estimate energy and nutrient intake from nonbreast milk foods during the weaning process. In addition, to assess the household food access insecurity status. To assess food access insecurity, our survey included the nine-question household food insecurity access scale adapted in 2006 by the Food and Nutrition Technical Assistance project for use in low-resource settings [27] and used in the same population in the eight-countryside MAL-ED study by Psaki et

XSL•FO
RenderX

al [28]. On a monthly basis, we assess water, sanitation, and hygiene practices of mothers or caregivers, incorporating the 5 key principles and the practical steps to achieve universal access to quality care WHO standards [29], as well as indicators related to antibiotic use and potential drivers of antibiotic resistance. At baseline, 7, 15, and 24 months, we also administer a Self-Reporting Questionnaire screening instrument that effectively measures depression in postpartum women, an instrument previously developed and used in the same population [30]. The Self-Reporting Questionnaire comprises 20 items that assess psychological disturbances related to depressive symptoms occurring within the previous 4 weeks. Questions are answered with a simple "yes" or "no," and trained field assistants administer the instrument.

Mothers are trained to collect stool and breast milk samples as required. A stool sample is collected from the child and mother at enrollment, and then monthly, and whenever there is diarrhea, for 12 months. Thereafter, stool is collected quarterly, and whenever there is diarrhea, for a further 12 months. Breast milk is collected at enrollment and monthly, and for as long as the mother breastfeeds, for a total duration of 12 months.

Environmental samples, including household drinking water and fecal matter from in-household livestock and pets, are collected once during the study to assess zoonotic *E coli* resistome. Hydrometeorological exposure data such as precipitation, temperature, humidity, and soil moisture will be extracted from version 2.1 of the Global Land Data Assimilation System.

DNA Extraction and Sequencing

The main goal is to generate metagenomic sequence data using next-generation sequencing, bioinformatic, and correlational analyses to understand infant microbial composition and resistance dynamics. Total genomic DNA (gDNA) is extracted from infant stool samples using a QIAamp Fast DNA stool min Kit (Qiagen) following the manufacturer's protocol, with the inclusion of a bead-beating step, to the protocol. Extracted gDNA is quantified with a Qubit dsDNA High Sensitivity Assay, and library preparation is conducted using the Nextera XT DNA Library Prep Kit (Illumina). Sequencing will be performed on the Illumina NextSeq 2000 platform with paired-end reads (2×150 bp) using the P4 flow cell.

Analysis of Enteric Bacterial Resistomes and Associated Risk Factors

The study will first characterize the total gut resistome and then focus on *E coli* resistome, where the organism will be used as a proxy organism for monitoring AMR in the gut microbiota due to its ubiquity, early colonization in the infant gut, and well-characterized genomic features. In addition, the ESKAPE pathogens' resistome will be included to broaden our understanding of clinically relevant bacteria. By leveraging *E coli* and ESKAPE pathogens as sentinel organisms, this study aims to provide a focused yet robust characterization of the early-life enteric resistome, enabling deeper insights into AMR acquisition in low-resource community settings.

Metagenomic Sequence Analysis

Overview

Bioinformatic analysis of metagenomic data from stool samples includes quality control and filtering, removal of host contaminants, taxonomic classification of reads, metagenome assembly and binning of assembled contigs, evaluation of bin completeness and contamination, taxonomic classification of bins, detection of AMR genes, and quantification of gene abundance through read mapping. Pipelines for sequence processing include BBduk (quality filtering) [29], BBmap (host decontamination and read mapping) [30], Kaiju (read-level taxonomic classification) [31], MetaSPAdes (assembly) [32], MetaBat (binning) [33], CheckM (bin quality assessment) [34], GTDB-Tk (bin classification) [35], and AMRfinder (ARG detection) [36]. ARGs will be classified by antibiotic class and resistance mechanism. The resulting metagenomic data presents an opportunity to study the total gut resistome, resistomes of *E coli* and ESKAPE pathogens, and other South African priority pathogens such as *Salmonella* and *Shigella*.

Association of Enteric Gut Resistome and Potential Risk Factors

Based on the major objectives, key analysis will be to (1) characterize the infant gut microbiome in terms of taxonomic composition, within and between sample diversity, and longitudinal trajectories across the first 24 months of life; (2) resistome profiles by assessing the abundance, diversity, and distribution of ARGs, including the identification of key resistance determinants; (3) investigate associations between resistome outcomes and host of metadata variables (sociodemographics; perinatal and anthropometrics; nutritional exposures, and feeding practices; illness, medication, and vaccination; breastmilk metabolome; household socioeconomic status; maternal psychosocial and behavioral factors; hygiene and sanitation practices; environmental exposures including hydro-meteorological variables, in-house livestock and pets, and drinking water quality); and (4) evaluate the temporal dynamics of the microbiome and resistome by leveraging longitudinal sampling to assess population-level trends over time. This will uncover temporal patterns and trajectories in gut microbial and resistome dynamics that may inform future interventions for AMR mitigation in early life.

The R software (R Foundation for Statistical Computing) is used for statistical analyses and graphical representations [37]. Descriptive statistics is used to summarize the mother-child pair characteristics of the cohort. The microbiome and resistome profiles are characterized using alpha (eg, Shannon index and Simpson index) and beta diversity (eg, Bray-Curtis dissimilarity) metrics, ordination (eg, principal component analysis using centered log-ratio-transformation data) to visualize compositional structure, and permutational multivariate analysis of variance testing for group-based differences, such as mode of delivery or nutrition. Core microbiome and resistome features are defined as taxa or ARGs present in ≥90 % of samples within a group. These are used to identify persistent and dominant features across time and metadata categories. Longitudinal changes in microbiome and resistome composition are modeled using linear mixed-effects models to account for repeated

XSL•FO
RenderX

measures. Dirichlet-multinomial modeling is used to capture compositional shifts over time. For ARG abundance, exponential decay models are fitted to assess temporal trends, particularly in early life, where ARG load is expected to decline with age. Associations between microbiome or resistome features and covariates (eg, birth type, gender, dietary profiles, and meteorological data) are assessed using multivariable regression models. Spearman correlation is used for nonparametric associations. Where appropriate, machine learning models (eg, random forests) are used to identify predictive relationships between metadata and resistome features.

Differential abundance of taxa and ARGs is assessed using compositionally aware methods such as Analysis of Compositions of Microbiomes with Bias Correction. These models account for sampling bias and compositional constraints inherent in metagenomic data. Finally, core ARGs and taxa will be compared across groups to identify key features associated with metadata and time. Where appropriate, missing nonbiological covariate data will be addressed using complete-case analysis or multiple imputation, depending on missingness, and sensitivity analyses will be performed to assess the robustness of findings based on assumptions.

Knowledge, Attitude, and Practice on the Use of Antimicrobials

To generate data that will contribute to the development of relevant educational packages, a community survey on Knowledge, Attitude, and Practice (KAP) on the use of antimicrobials is underway. A previous study in the Dzimauli area (in the Thulamela municipality in the Vhembe District of South Africa) from 2009 to 2012 reported an estimated population of 9000 [19]. Assuming an annual growth rate of 2.23% based on the population growth rate calculated from the Thulamela municipality population from 2011 to 2022 [38], the population of Dzimauli was projected to be 12,256 by 2023. The sample size for the community KAP survey was estimated as 466 targeted participants using the single-proportion formula with finite population correction (n=12,256), assuming P=0.5, d=0.05, and 95% CI and assuming 80% data completeness. A structured questionnaire is used to collect information on participant demographics, knowledge of antibiotics and antibiotic resistance, as well as attitudes and practices related to antibiotic use. The survey is conducted door-to-door and is administered in English or Tshivenda (the local language), depending on the participant's preference. Inclusion criteria are those aged 16 years or older. The KAP component consists of 29 closed-ended questions (14 knowledge, 7 attitude, and 8 practice-related items). There is also a question to provide sources of information on antibiotics and 2 open-ended questions on general views. A 4-point Likert scale ranging from "Strongly disagree" to "strongly agree" was adopted to capture participant responses to closed-ended questions. Responses will be analyzed only for participants with complete data sets. The responses will be collapsed into a dichotomous set based on correct or incorrect responses. Total scores will be used to classify "high" (≥80%), "moderate" (60% - 79%), and "low" (<60%) KAPs on antibiotic use [39,40]. Frequency distribution for each KAP will be determined to understand the level of KAP in the community regarding antibiotics. Logistic regression analysis will be conducted to identify factors associated with KAP regarding antibiotic use, and odds ratios with 95% CIs will be reported. Statistical significance will be set at a P value <.05.

Open-ended responses from KAP surveys will be analyzed using thematic analysis. Responses will be transcribed where necessary, coded inductively, and grouped into thematic themes related to antibiotic knowledge, perceptions, practices, and contextual drivers of inappropriate antibiotic use. Qualitative findings will be used to contextualize and interpret quantitative KAP. Integration of qualitative and quantitative data will be conducted informally during interpretation, to support knowledge translation and the co-production of ethnographically relevant AMR stewardship interventions.

In addition to the community KAP survey, an end-assessment KAP survey is administered at 24 months, or when a participant exits the study. This latter survey aims at evaluating whether participation in a longitudinal AMR study has influenced maternal KAP related to antimicrobial use. Table 2 summarizes the different KAP assessments, the target population, and timing.

Table . Summary of the Knowledge, Attitude, and Practice (KAP) surveys, target groups, time points, and assessment focus in the Dzimauli birth cohort study.

Measures	Time point	Group	Type of analysis or outcome
Community KAP	Once off	Community	General Community KAP on antibiotic use
End assessment KAP	24 months	Mothers	Impact of participation in a prospective study on KAP

Development of Ethnographically Appropriate Educational Packages for AMR Awareness Through Responsive Dialogues and Knowledge Translation

This activity draws from the KAP and enteric bacteria resistome datasets. During conversation events, participants will share their first-hand experiences on AMR and collectively brainstorm ideas and co-create solutions to tackle AMR challenges. Engagement with stakeholders, such as policy and decision makers, will take place during symposiums, which began during the conception of the project. Policy and decision makers are drawn from district, provincial, and national government health agencies charged with infection prevention and control.

Research Data Integrity

We describe here how the study guarantees data quality, ensures reproducibility, and the ethics of doing research by tracking research integrity, misconduct, and misuse of research findings. The principal investigator reviews informed consent forms on

XSL•FO
RenderX

the day they are signed, for correctness and completeness, and stores them securely. The research assistants are trained on the study protocol, data collection tools, and ethical conduct of research prior to commencement of fieldwork and through ongoing refreshers as needed. Case report forms are reviewed within 24 hours and signed by a supervisor. Corrected data on case report forms remains legible, signed, and dated. Repeated data capture errors trigger re-training for the worker or the group. Field and laboratory supervisors hold weekly meetings with staff to review data quality plans, challenges, and progress. Monthly, a supervisor does 5% of the monthly assessments for quality assurance and performs a dry-run training. Approved data are captured in REDCap (Research Electronic Data Capture; Vanderbilt University) system. Devices for anthropometric measurements are maintained and calibrated as recommended. Biorepositories are fitted with temperature monitoring devices and are monitored daily.

A protocol deviation form is completed for every out-of-schedule field assessment, laboratory test, or missing data. Case report forms are scanned and backed up electronically. In total, 85% of completeness and accuracy is the minimum targeted threshold for all data points. Data management is governed by policies and ethics on data creation, storage, sharing, and access. Training on ethics, integrity, and responsible conduct of research emphasizes data falsification, fabrication, plagiarism, and the reputation of the study community, investigators, institutions, and funders.

Potential tangible risks related to research misuse include the risk of misuse of (1) personal data and identifiers of study participants, and (2) analysis of DNA sequence data for reasons other than those of the current research. Soon after consent to participate is obtained, all data collection, tests, and assessments will include a participant identifier on the data collection tool. Samples also have sample identifiers. Through this, participants are delinked from samples, assessments, and datasets. In addition, access to data is password-protected. Study investigators, field, and laboratory workers are trained on the responsible conduct of research involving human participants.

The study generates next-generation sequencing data potentially containing human genes. It is worthwhile to note that it is difficult to deidentify a genome, because a genome by itself is in many ways an identifier. However, in this study, investigators avoid examining parts of the generated sequences that are not of interest to the study in order to avoid incidental findings, such as stumbling on variants that are not of interest to the study but could be of interest to a study participant. Incidental findings pose a dilemma. If this sort of event arises, the matter will be referred to the institutional research ethics committee for guidance on how to proceed. Generally, misuse of data is governed and managed according to legislation and guidelines regulating access to personal information and protection of human dignity.

Ethical Considerations

The study protocol has approval from the Human and Clinical Trials Research Ethics Committee of the University of Venda, South Africa (FSEA/22/MBY/04/0302). Permission was granted by the Limpopo Provincial Department of Health, South Africa, to access local health care facilities in the study community for identification of potential participants (LP_2023-03-014). A Memorandum of Understanding regulates the relationship with the Dzimauli Tribal Authority. Several meetings were held with the community leadership to obtain community approval prior to individual consent. A community advisory committee was established to guide the values of respect, honesty, fairness, and justice vis-à-vis community participation, protection, and expectations from the study. Individually signed informed consent is obtained for all forms of participation in the study. To protect participants' privacy and confidentiality, specimens and surveillance data are stripped of personal identifiers and coded. Ethical principles according to the Helsinki Declaration are adhered to [41]. The study participants did not receive compensation.

Results

Study Progress and Timeline

The study was funded in June 2022. Ethical approval was obtained in February 2023 from the Human and Clinical Trials Research Ethics Committee of the University of Venda, South Africa (FSEA/22/MBY/04/0302). Recruitment of mother-child dyads commenced in August 2023, with an expected end date in June 2026. By October 2025, 167 mother-child dyads had been enrolled, and baseline demographic data collection had been completed for the enrolled participants. Stool sample collection continues in accordance with the study follow-up schedule. Laboratory DNA extraction commenced in January 2024, while sequencing of the baseline samples for all participants began in August 2025, with initial data quality control and descriptive analyses expected by May 2026. Comprehensive integrative resistome-microbiome and qualitative analyses are planned following completion of data collection in December 2027. Dissemination of the principal study findings is anticipated through peer-reviewed publications and structured community feedback activities between June 2026 and December 2028.

Characteristics of Currently Enrolled Participants

Among the 167 newborns enrolled by October 2025, 20 participants have completed 24 months of follow-up. Baseline characteristics of these participants are summarized by total numbers with percentages and mean (SDs) to give an overview of participant characteristics. Proportions for each variable are based on the total number of participants with complete data on the assessment or question on the instrument. An attrition rate of 9.6% has been recorded. Figure 1 shows the enrolled strategy, while Tables 3 and 4 show selected baseline descriptive characteristics of the 167 participant-dyads. The majority of the infants (129/167, 77.2%) were enrolled within 14 days after birth. There were no significant differences in sex (χ^2_1=0.15; *P*=.70) among enrolled babies, indicating an approximately equal representation of males and females. The mean weight of the babies at enrollment was 3.4 (SD 0.6) kg, compared to 3.1 (SD 0.5) kg recorded at birth. The mean length of babies at enrollment was 49.5 (SD 4.4) cm. None of the babies were hospitalized prior to enrollment, and 95.8% (160/167) received colostrum at birth. The length-for-age *z* scores indicated that

XSL•FO
RenderX

3.6% (6/165) were severely stunted, and weight-for-length *z* scores revealed a 5% (8/161) in the category of severe wasting. At enrollment, the mean age of the mothers was 27.7 (SD 7.1) years. Overall, 29.9 % (50/167) of the mothers had completed at least 12 years of schooling. Of those below 25 years, 30.6% (26/85) were attending at least high school. At enrollment, a baseline stool sample was collected from 96.4% (158/167) of the infant participants. For a robust statistical power calculation, a minimum threshold of 85% of complete baseline data collection for each of: sociodemographics, anthropometrics, treatment, hospitalization, and breastfeeding is being met.

Figure 1. Flowchart of study cohort establishment: screening criteria and current enrollment.

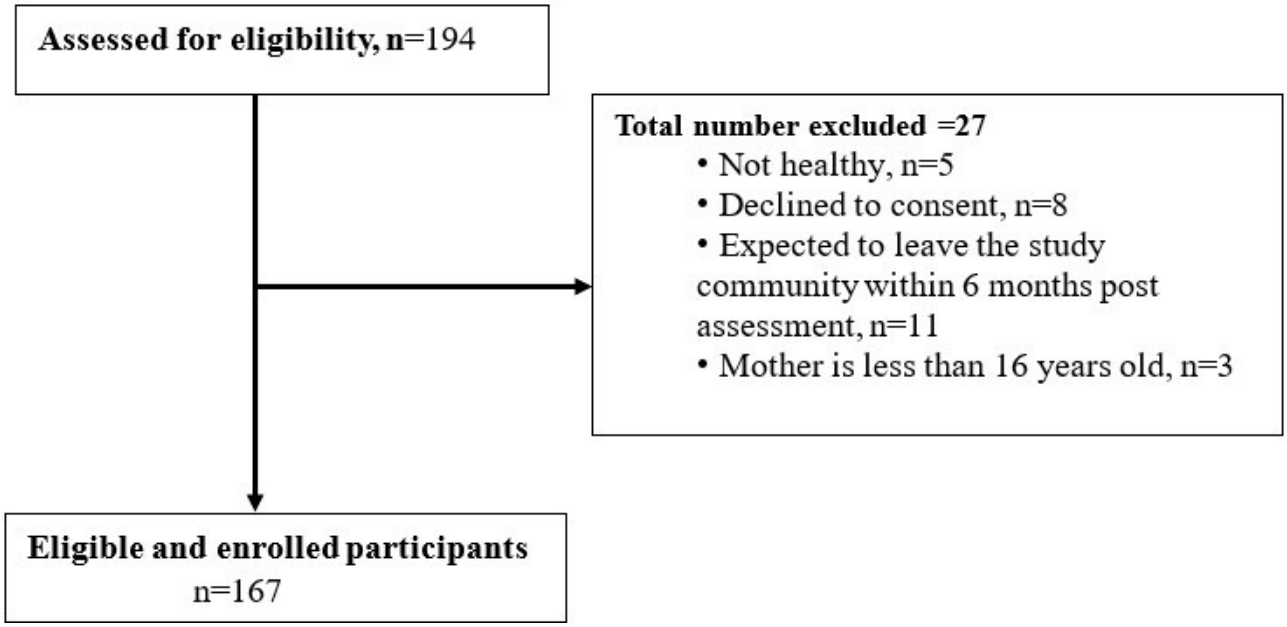

Table . Baseline characteristics of enrolled participants.

Variables	Mean (SD)	Range (min-max)
Birthweight (kg)	3.1 (0.5)	1.9 - 4.4
Weight at enrollment (kg)	3.4 (0.6)	2.1 - 5
Length at birth (cm; n=160[a])	49.0 (2)[b]	33 - 57
Length at enrollment (cm)	49.5 (4.4)	33 - 57
Head circumference at birth (cm; n=161[a])	34.6 (1.60)	30 - 38.1
Head circumference at enrollment (cm; n=166[a])	35.3 (1.5)	30.6 - 39
Child age (days)	9.6 (4.6)	1 - 19[c]
Age of mother in years	27.7 (7.1)	16 - 44
Number of live births of the mother	2.6 (1.6)	1 - 8
Number of years of schooling of mother (n=164[a])	11.1 (2.2)	2 - 16

[a]N, when not equal to 167

[b]Median (IQR)

[c]There were 3 deviations from the target timeline of enrollment within 17 days of birth. (One infant was enrolled at day 19 after birth, and two infants at 18 days after birth).

XSL•FO
RenderX

Table . Distribution of enrolled infants and their mothers based on selected baseline variables.

Categories	Participants, n (%)
Gestational age at birth	
Term	160 (95.8)
Preterm	7 (4.2)
Infant sex	
Male	86 (51.5)
Female	81 (48.5)
Place of birth	
Health facility	163 (97.6)
Home	4 (2.4)
Birth weight group (kg)	
>2.5	144 (86.2)
1.9 - 2.49	23 (13.8)
WLZ[a] at enrollment (n=161[b])	
−1≤ WLZ ≤1 (normal)	71 (44.1)
−2≤ WLZ <−1 (mild wasting)	22 (13.7)
−3≤ WLZ<−2 (moderate wasting)	9 (5.6)
WLZ < −3 (severe wasting)	8 (5)
1 <WLZ ≤2 (mild overweight)	41 (25.5)
2≤ WLZ ≤3 (overweight)	9 (5.6)
WLZ >3 (obese)	1 (0.6)
Length-for-age (LAZ)[c] (n=165[b])	
LAZ ≥ −1 (normal)	113 (68.5)
−2≤ LAZ<−1 (mild stunting)	32 (19.4)
−3≤ LAZ<−2 (moderate stunting)	14 (8.5)
LAZ < −3 (severe stunting)	6 (3.6)
Child fed first milk (colostrum)	
Yes	160 (95.8)
No	7 (4.2)
Prelacteal feeding	
Yes	67 (40.1)
No	100 (59.9)
Parity	
1 (low)	50 (29.9)
2 - 3 (moderate)	73 (43.7)
>3 (high)	44 (26.3)
Mother’s marital status	
Never married	99 (59.3)
Married	66 (39.5)
Divorced	1 (0.6)
Widowed	1 (0.6)
Mother attended school	

XSL•FO
RenderX

Categories	Participants, n (%)
Yes	164 (98.2)
No	3 (1.8)

[a]WLZ: weight-for-length *z* scores.

[b]N, when not equal to 167.

[c]LAZ: length-for-age *z* scores.

Discussion

Knowledge Gaps Being Addressed

This study protocol describes a prospective observational birth cohort designed to investigate the acquisition, composition, and transmission of ARGs during the first 24 months of life in a South African setting of low socioeconomic status. Early life exposures, particularly within the first 24 months of life, are critical for microbiome establishment and resistome development and shape long-term immunity, metabolism, and resistance profiles [31,32]. By integrating longitudinal metagenomic surveillance with external exposures, including hydrometeorological and contextual data, the study aims to address key gaps in understanding the temporal dynamics of microbiome-resistome and AMR emergence and dissemination outside health care settings. In parallel, the study incorporates KAP assessments on antibiotic use [33] and the co-production of intervention materials, with the goal of translating biological findings into contextually relevant AMR mitigation strategies [34].

Potential Study Findings

Accordingly, a multidimensional longitudinal birth cohort in Dzimauli, a rural South African setting, as a model for community-level AMR stewardship is being established. The cohort is designed to integrate biological, social, and contextual data to provide comprehensive interventional data for community-level AMR. Integrative data generated from this cohort is expected to provide several important contributions to AMR evidence-based interventions. First, the study is expected to generate critical prospective datasets describing the enteric resistome in a low-socioeconomic community, addressing a major gap in community-level AMR surveillance. Longitudinal metagenomic enteric profiles will reveal critical data, including acquisition, composition, persistence, and evolution of ARGs; community-level intervention hotspots; emerging community-level critical pathogens of clinical relevance and potential transmission pathways, including vertical and horizontal transmissions within households and the broader environment. We capture a wide range of exposures, contextual factors, and potential confounders to clearly identify drivers of AMR. Integrative metagenomic and exposure data will provide intervention insights. This approach ensures that intervention strategies are grounded in local realities and informed by community perspectives. Finally, the study will ensure enhancement of the development of leadership in integrative AMR epidemiology in low-income communities for local and global relevance.

Current evidence on AMR in low-resource settings is dominated by cross-sectional studies, which provide only snapshots of resistance patterns at single time points [35,36]. Although such studies are valuable in terms of estimating prevalence, they are limited in their ability to capture proper AMR baseline and temporal dynamics data, including when ARGs are first acquired, which bacteria greatly contribute the most ARGs, how they persist or change over time, and how their transmission pathways evolve. The emergence of AMR itself is complex, shaped by the interaction of biological, environmental, behavioral, and social factors [37,38]. However, a plethora of studies focus on a limited subset of potential drivers, such as antibiotic use or selected environmental variables, without accounting for broader contextual influences [39,40,42]. This narrow focus may bias interpretations of AMR drivers and underestimates the contribution of diverse structural and social determinants. Furthermore, while many AMR studies generate detailed microbiological data, limited studies explicitly link these findings to community knowledge, behaviors, lived experiences, and the production of setting or regional-specific interventions [34,43,44]. These limitations are particularly pronounced in sub-Saharan Africa, where longitudinal AMR data from community settings, especially among infants and young children, remain scarce [45]. As a result, the establishment of an interdisciplinary birth study cohort is key to addressing these limitations and generating the longitudinal, contextualized data necessary for effective AMR surveillance and stewardship.

Community-Based Challenges

It is worthwhile to indicate challenges faced so far in the cohort establishment, biospecimen and surveillance data collection, and the strategies being applied to mitigate the challenges. First, due to socioeconomic factors, there is a fair number of relocations of community residents out of the study area. Therefore, effort is required to identify potential eligible participants who aim to stay long enough. This approach contributes to lower attrition and meaningful contributions to the datasets envisaged. Second, the topology of the study community makes the mobility of the field workers to access households difficult. To address this, efforts were made to increase the number of field workers such that the number of households that each field worker is responsible for is reduced. The added resources to address the challenge of mobility pay off by enhancing timely data collection, completeness, and accuracy. Third, the display of fatigue in providing samples prospectively. This behavior was diagnosed as an inadequate understanding of why a specimen, for example, stool, is collected repeatedly. This is addressed through conversations with relevant participants for the purpose of repeated collection and through continuous consenting.

XSL•FO
RenderX

Strengths and Limitations

The cohort study has major strengths, including its (1) prospective design and community-based nature, which address a critical evidence gap in community engagement and co-production of scientific evidence in the mitigation of AMR, (2) its broad longitudinal exposure data, including the integration of hydrometeorological and nutritional variables, would enable a holistic assessment of potential drivers of resistome acquisition beyond antibiotic exposure alone, and (3) the use of shotgun metagenomic sequencing to determine the resistome makes it a strong methodological approach to obtain large datasets necessary for the discovery of novel niches and relevant priority organisms driving AMR. Overall, it brings to the fore the role of enteric resistome in child health and livelihoods for global health imperatives.

Despite these strengths, the cohort study should be considered with certain potential limitations. These include the introduction of recall bias, increased risk of missing or delayed data due to the prospective nature and the intensity of data collection; factors such as undocumented exposures, such as over-the-counter antibiotic use, may be missed; the management of attrition rate in some communities, for example, pastoral communities, may be difficult. Furthermore, although metagenomic data provides a high level of granularity, it may not be feasible in resource-constrained research environments.

Conclusions

Large prospective datasets on AMR are necessary for decision-making and intervention in sub-Saharan African settings, particularly in low-socioeconomic settings. Therefore, a comprehensive assessment of the multiple and interacting determinants that drive AMR dynamics remains an essential aspect in AMR stewardship. This protocol is a proposal for AMR research at the community level to identify drivers important for AMR mitigation involving local communities in data and evidence generation.

Acknowledgments

The collaboration of the leadership of the Dzimauli community and of the study participants is highly appreciated. We acknowledge the contribution of the field assistants.

Microbiome and Antimicrobial Resistance Collaborators is made up of the following members:

Pascal Bessong, University of Venda, South Africa

Justine Fri, University of Venda, South Africa

Cloupas Mahopo, University of Venda, South Africa

Idris Njanje, University of Venda, South Africa

Amidou Samieie, University of Venda, South Africa

Angelina Maphula, University of Venda, South Africa

Cebisa Nesamvuni, University of Venda, South Africa

Tirhani Masia, University of Venda, South Africa

Makondelele Radzilani-Makatu, University of Venda, South Africa

Hilda Shilubane, University of Venda, South Africa

Edwin Madala, University of Venda, South Africa

Emmanuel Sikitime, University of Venda, South Africa

Martin Nyaga, University of the Free State, South Africa

Stefano Cacciatore, International Centre for Genetic Engineering and Biotechnology (ICGEB), Cape Town, South Africa

Maphoshane Nchabeleng, Sefako Makgatho Health Sciences University, South Africa

Lufuno Mavhandu-Ramarumo, Sefako Makgatho Health Sciences University, South Africa

Andrew Musyoki, Sefako Makgatho Health Sciences University, South Africa

Gloria Selabe, Sefako Makgatho Health Sciences University, South Africa

Teke Apalata, Walter Sisulu University, South Africa

Rian Pierneef, University of Pretoria, South Africa

Adeline Ngie, North-West University, South Africa

Joel Bazira, Mbarara University of Science and Technology, Uganda

XSL•FO
RenderX

Barbra Tuhamize, Mbarara University of Science and Technology, Uganda

Stella Smith, Nigerian Institute of Medical Research, Nigeria

Funding

The study is supported by the South African Medical Research Council (SAMRC). Additional support was received from the National Research Foundation of South Africa (Grant No. RCHDI241107279719). Support for JF and IN is through the SAMRC, Division of Research Capacity Development under the Research Capacity Development Initiative, on funding received from the South African National Treasury. It is also based on research supported in part by the National Research Foundation of South Africa (Grant number: RCHDI241107279719). LGM-R is supported by the Future Professors Programme of South Africa. The content hereof is the views of the authors and not necessarily those of the funders or affiliated institutions.

Data Availability

Surveillance and data collection forms used in this study are available upon request. Data generated from this study will be made publicly available upon completion of the primary analyses as open access peer-reviewed publications. Deidentified quantitative and sequencing data will be deposited in publicly accessible repositories, with accession details provided in subsequent publications. Data that cannot be publicly shared due to ethical or intellectual property considerations will be made available upon reasonable request to the principal investigator (corresponding author).

Authors' Contributions

POB conceived the study, sourced funding, and wrote the first draft of the manuscript. JF, IN, and LGM-R are laboratory supervisors, and TCM is the field supervisor. JF, IN, and TCM conducted the preliminary analyses and developed the manuscript. JF, IN, TCM, LGM-R, and POB reviewed and edited the manuscript. All authors approved the submission of the manuscript. All authors have reviewed the final version of the manuscript and consent to its publication.

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[DOCX File, 35 KB - resprot_v15i1e86058_app1.docx]

Peer Review Report 1
Peer-review report from the South African Medical Research Council (SAMRC).
[PDF File, 373 KB - resprot_v15i1e86058_app2.pdf]

References

1. Murray CJL, Ikuta KS, Sharara F, et al. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. Lancet 2022 Feb;399(10325):629-655. [doi: 10.1016/S0140-6736(21)02724-0] [Medline: 35065702]
2. Sartorius B, Gray AP, Davis Weaver N. The burden of bacterial antimicrobial resistance in the WHO African region in 2019: a cross-country systematic analysis. Lancet Glob Health 2024 Feb;12(2):e201-e216. [doi: 10.1016/S2214-109X(23)00539-9] [Medline: 38134946]
3. Larsson S, Svensson M, Ternhag A. Production loss and sick leave caused by antibiotic resistance: a register-based cohort study. BMC Public Health 2022 Mar 17;22(1):527. [doi: 10.1186/s12889-022-12947-x] [Medline: 35300650]
4. Ferri M, Ranucci E, Romagnoli P, Giaccone V. Antimicrobial resistance: a global emerging threat to public health systems. Crit Rev Food Sci Nutr 2017 Sep 2;57(13):2857-2876. [doi: 10.1080/10408398.2015.1077192] [Medline: 26464037]
5. WHO bacterial priority pathogens list, 2024: bacterial pathogens of public health importance, to guide research, development, and strategies to prevent and control antimicrobial resistance. : World Health Organization; 2024 URL: https://www.who.int/publications/i/item/9789240093461 [accessed 2026-03-06]
6. Antibacterial agents in clinical and preclinical development: an overview and analysis. : World Health Organization; 2021 URL: https://www.who.int/publications/i/item/9789240047655 [accessed 2026-03-06]
7. Mitchell J, Cooke P, Ahorlu C, et al. Community engagement: the key to tackling antimicrobial resistance (AMR) across a one health context? Glob Public Health 2022 Nov;17(11):2647-2664. [doi: 10.1080/17441692.2021.2003839] [Medline: 34882505]
8. Mathew P, Chandy SJ, Ranjalkar J. Community engagement to mitigate antimicrobial resistance in low-and middle-income countries - an essential strategy for implementation of national action plans on AMR. Lancet Reg Health Southeast Asia 2024 May;24:100379. [doi: 10.1016/j.lansea.2024.100379] [Medline: 38510334]

XSL•FO
RenderX

9. Manyi-Loh C, Mamphweli S, Meyer E, Okoh A. Antibiotic use in agriculture and its consequential resistance in environmental sources: potential public health implications. Molecules 2018 Mar 30;23(4):795. [doi: 10.3390/molecules23040795] [Medline: 29601469]
10. Burnham JP. Climate change and antibiotic resistance: a deadly combination. Ther Adv Infect Dis 2021;8:2049936121991374. [doi: 10.1177/2049936121991374] [Medline: 33643652]
11. Laxminarayan R, Duse A, Wattal C, et al. Antibiotic resistance-the need for global solutions. Lancet Infect Dis 2013 Dec;13(12):1057-1098. [doi: 10.1016/S1473-3099(13)70318-9] [Medline: 24252483]
12. Ahmad I, Malak HA, Abulreesh HH. Environmental antimicrobial resistance and its drivers: a potential threat to public health. J Glob Antimicrob Resist 2021 Dec;27:101-111. [doi: 10.1016/j.jgar.2021.08.001] [Medline: 34454098]
13. Magnano San Lio R, Favara G, Maugeri A, Barchitta M, Agodi A. How antimicrobial resistance is linked to climate change: an overview of two intertwined global challenges. Int J Environ Res Public Health 2023 Jan 17;20(3):1681. [doi: 10.3390/ijerph20031681] [Medline: 36767043]
14. DeFrancesco AS, Tanih NF, Samie A, Guerrant RL, Bessong PO. Antibiotic resistance patterns and beta-lactamase identification in Escherichia coli isolated from young children in rural Limpopo Province, South Africa: the MAL-ED cohort. S Afr Med J 2017 Feb 27;107(3):205-214. [doi: 10.7196/SAMJ.2017.v107i3.12111] [Medline: 28281425]
15. Rogawski McQuade ET, Brennhofer SA, Elwood SE, et al. Frequency of bystander exposure to antibiotics for enteropathogenic bacteria among young children in low-resource settings. Proc Natl Acad Sci U S A 2022 Sep 6;119(36):e2208972119. [doi: 10.1073/pnas.2208972119] [Medline: 36037372]
16. Raman AS, Gehrig JL, Venkatesh S, et al. A sparse covarying unit that describes healthy and impaired human gut microbiota development. Science 2019 Jul 12;365(6449):eaau4735. [doi: 10.1126/science.aau4735] [Medline: 31296739]
17. Moura de Sousa J, Lourenço M, Gordo I. Horizontal gene transfer among host-associated microbes. Cell Host Microbe 2023 Apr 12;31(4):513-527. [doi: 10.1016/j.chom.2023.03.017] [Medline: 37054673]
18. Colston JM, Zaitchik B, Kang G, et al. Use of earth observation-derived hydrometeorological variables to model and predict rotavirus infection (MAL-ED): a multisite cohort study. Lancet Planet Health 2019 Jun;3(6):e248-e258. [doi: 10.1016/S2542-5196(19)30084-1] [Medline: 31229000]
19. Colston JM, Zaitchik BF, Badr HS, et al. Associations between eight earth observation-derived climate variables and enteropathogen infection: an independent participant data meta-analysis of surveillance studies with broad spectrum nucleic acid diagnostics. Geohealth 2022 Jan;6(1):e2021GH000452. [doi: 10.1029/2021GH000452] [Medline: 35024531]
20. Bessong PO, Nyathi E, Mahopo TC, Netshandama V, MAL-ED South Africa. Development of the Dzimauli community in Vhembe District, Limpopo province of South Africa, for the MAL-ED cohort study. Clin Infect Dis 2014 Nov 1;59 Suppl 4:S317-S324. [doi: 10.1093/cid/ciu418] [Medline: 25305304]
21. Moyo CS, Francis J, Bessong PO. Perceptions of community-based field workers on the effect of a longitudinal biomedical research project on their sustainable livelihoods. BMC Public Health 2017 Mar 17;17(1):267. [doi: 10.1186/s12889-017-4138-6] [Medline: 28302095]
22. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. BMJ 2007 Oct 20;335(7624):806-808. [doi: 10.1136/bmj.39335.541782.AD] [Medline: 17947786]
23. Bharucha T, Oeser C, Balloux F, et al. STROBE-metagenomics: a STROBE extension statement to guide the reporting of metagenomics studies. Lancet Infect Dis 2020 Oct;20(10):e251-e260. [doi: 10.1016/S1473-3099(20)30199-7] [Medline: 32768390]
24. Netshisaulu KH, Chikoore H, Chakwizira J, Kom Z. Sustainable livelihood options adopted by rural communities in response to climate change dynamics: a case study approach in Vhembe District, South Africa. Sustainability 2025;17(3):1284. [doi: 10.3390/su17031284]
25. Richard SA, Barrett LJ, Guerrant RL, Checkley W, Miller MA, MAL-ED Network Investigators. Disease surveillance methods used in the 8-site MAL-ED cohort study. Clin Infect Dis 2014 Nov 1;59 Suppl 4(Suppl 4):S220-S224. [doi: 10.1093/cid/ciu435] [Medline: 25305290]
26. Caulfield LE, Bose A, Chandyo RK, et al. Infant feeding practices, dietary adequacy, and micronutrient status measures in the MAL-ED study. Clin Infect Dis 2014 Nov 1;59 Suppl 4(Suppl 4):S248-S254. [doi: 10.1093/cid/ciu421] [Medline: 25305294]
27. Swindale A, Bilinsky P. Household dietary diversity score (HDDS) for measurement of household food access: indicator guide. : Food and Nutrition Technical Assistance Project, Academy for Educational Development; 2006 URL: https://www.fantaproject.org/sites/default/files/resources/HDDS_v2_Sep06_0.pdf [accessed 2026-03-06]
28. Psaki S, Bhutta ZA, Ahmed T, et al. Household food access and child malnutrition: results from the eight-country MAL-ED study. Popul Health Metr 2012 Dec 13;10(1):24. [doi: 10.1186/1478-7954-10-24] [Medline: 23237098]
29. WASH in health care facilities: practical steps to achieve universal access to quality care. : World Health Organization; 2018 URL: https://www.who.int/publications/i/item/9789241515511 [accessed 2026-03-06]
30. Pendergast LL, Scharf RJ, Rasmussen ZA, et al. Postpartum depressive symptoms across time and place: structural invariance of the Self-Reporting Questionnaire among women from the international, multi-site MAL-ED study. J Affect Disord 2014;167:178-186. [doi: 10.1016/j.jad.2014.05.039] [Medline: 24981251]

XSL•FO
RenderX

31. Lebeaux RM, Coker MO, Dade EF, et al. The infant gut resistome is associated with E. coli and early-life exposures. BMC Microbiol 2021 Jul 2;21(1):201. [doi: 10.1186/s12866-021-02129-x] [Medline: 34215179]
32. DuPont HL, Salge MMH. The importance of a healthy microbiome in pregnancy and infancy and microbiota treatment to reverse dysbiosis for improved health. Antibiotics (Basel) 2023 Nov 11;12(11):1617. [doi: 10.3390/antibiotics12111617] [Medline: 37998819]
33. Farley E, van den Bergh D, Coetzee R, Stewart A, Boyles T. Knowledge, attitudes and perceptions of antibiotic use and resistance among patients in South Africa: a cross-sectional study. S Afr J Infect Dis 2019;34(1):118. [doi: 10.4102/sajid.v34i1.118] [Medline: 34485456]
34. Mitchell J, Arjyal A, Baral S, et al. Co-designing community-based interventions to tackle antimicrobial resistance (AMR): what to include and why. BMC Res Notes 2023 Oct 24;16(1):290. [doi: 10.1186/s13104-023-06449-1] [Medline: 37875996]
35. Gandra S, Alvarez-Uria G, Turner P, Joshi J, Limmathurotsakul D, van Doorn HR. Antimicrobial resistance surveillance in low- and middle-income countries: progress and challenges in eight South Asian and Southeast Asian countries. Clin Microbiol Rev 2020 Sep 16;33(3):10-1128. [doi: 10.1128/CMR.00048-19] [Medline: 32522747]
36. Iskandar K, Molinier L, Hallit S, et al. Surveillance of antimicrobial resistance in low- and middle-income countries: a scattered picture. Antimicrob Resist Infect Control 2021 Mar 31;10(1):63. [doi: 10.1186/s13756-021-00931-w] [Medline: 33789754]
37. Sakalauskienė GV, Radzevičienė A. Antimicrobial resistance: what lies beneath this complex phenomenon? Diagnostics (Basel) 2024 Oct 18;14(20):2319. [doi: 10.3390/diagnostics14202319] [Medline: 39451642]
38. Bennett C, Russel W, Upton R, Frey F, Taye B. Social and ecological determinants of antimicrobial resistance in Africa: a systematic review of epidemiological evidence. Antimicrob Steward Healthc Epidemiol 2024;4(1):e119. [doi: 10.1017/ash.2024.375] [Medline: 39257424]
39. Al-Khalaifah H, Rahman MH, Al-Surrayai T, Al-Dhumair A, Al-Hasan M. A one-health perspective of antimicrobial resistance (AMR): human, animals and environmental health. Life (Basel) 2025 Oct 13;15(10):1598. [doi: 10.3390/life15101598] [Medline: 41157271]
40. Li X, Brejnrod A, Thorsen J, et al. Differential responses of the gut microbiome and resistome to antibiotic exposures in infants and adults. Nat Commun 2023 Dec 22;14(1):8526. [doi: 10.1038/s41467-023-44289-6] [Medline: 38135681]
41. International ethical guidelines for health-related research involving humans. : Council for International Organizations of Medical Sciences (CIOMS); 2016 URL: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf [accessed 2026-03-06]
42. Salehi M, Laitinen V, Bhanushali S, et al. Gender differences in global antimicrobial resistance. NPJ Biofilms Microbiomes 2025 May 19;11(1):79. [doi: 10.1038/s41522-025-00715-9] [Medline: 40389466]
43. Asiimwe BB, Kiiru J, Mshana SE, et al. Protocol for an interdisciplinary cross-sectional study investigating the social, biological and community-level drivers of antimicrobial resistance (AMR): holistic approach to unravel antibacterial resistance in East Africa (HATUA). BMJ Open 2021 Mar 8;11(3):e041418. [doi: 10.1136/bmjopen-2020-041418] [Medline: 34006022]
44. Ho J, Yeoh YK, Barua N, et al. Systematic review of human gut resistome studies revealed variable definitions and approaches. Gut Microbes 2020 Nov 9;12(1):1700755. [doi: 10.1080/19490976.2019.1700755] [Medline: 31942825]
45. Bargheet A, Noordzij HT, Ponsero AJ, et al. Dynamics of gut resistome and mobilome in early life: a meta-analysis. EBioMedicine 2025 Apr;114:105630. [doi: 10.1016/j.ebiom.2025.105630] [Medline: 40048849]

Abbreviations

AMR: antimicrobial resistance
ARG: antibiotic resistance gene
ESKAPE: *Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumannii, Pseudomonas aeruginosa, and Enterobacter spp.*
KAP: Knowledge, Attitude, and Practice
MAL-ED: The Etiology, Risk Factors, and Interactions of Enteric Infections and Malnutrition and the Consequences for Child Health and Development
REDCap: Research Electronic Data Capture
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WHO: World Health Organization

XSL•FO
RenderX

Edited by A Schwartz; submitted 18.Oct.2025; this is a non–peer-reviewed article;accepted 28.Jan.2026; published 15.Apr.2026.

Please cite as:
Fri J, Njanje I, Mahopo TC, Mavhandu-Ramarumo LG, Bessong PO, Microbiome and Antimicrobial Resistance Collaborators
The Gut Bacterial Resistome in the First Two Years of Life: Protocol for a Longitudinal Observational Birth Cohort Study
JMIR Res Protoc 2026;15:e86058
URL: https://www.researchprotocols.org/2026/1/e86058
doi:10.2196/86058

© Justine Fri, Idris Njanje, Tjale Cloupas Mahopo, Lufuno Grace Mavhandu-Ramarumo, Pascal Obong Bessong, Microbiome and Antimicrobial Resistance Collaborators. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Role of Dairy Consumption on the Growth and Development of Canadian Children: Protocol for a Longitudinal Mixed Method Research

Karim Karbin[1], MSc; Mojtaba Shafiee[1,2,3], PhD; Zoe L Longworth[1], MPH; Adam Baxter-Jones[4], PhD; Marta C Erlandson[4], PhD; Ginny Lane[3], PhD; Ling-Ling Tsao[3], PhD; Walter L Siqueira[5], PhD; Hassan Vatanparast[1,2], MD, PhD

[1]College of Pharmacy and Nutrition, University of Saskatchewan, 107 Wiggins Road, Saskatoon, SK, Canada

[2]School of Public Health, University of Saskatchewan, Saskatoon, SK, Canada

[3]Margaret Ritchie School of Family and Consumer Sciences, College of Agricultural and Life Sciences, University of Idaho, Moscow, ID, United States

[4]College of Kinesiology, University of Saskatchewan, Saskatoon, SK, Canada

[5]College of Dentistry, University of Saskatchewan, Saskatoon, SK, Canada

Corresponding Author:
Hassan Vatanparast, MD, PhD
College of Pharmacy and Nutrition, University of Saskatchewan, 107 Wiggins Road, Saskatoon, SK, Canada

Abstract

Background: Childhood is a critical period for physical and cognitive development, in which nutrition plays a fundamental role. Dairy products, rich in calcium, vitamin D, and high-quality protein, are essential for bone health, body composition development, and cognitive performance. However, evidence on the long-term impact of dairy consumption on children's growth and development, particularly in the Canadian context, is limited. Additionally, recent revisions to the Canadian Food Guide, which classify dairy as part of the protein group rather than as a separate food group, may influence caregiver perceptions and children's dairy intake.

Objective: This protocol describes a study designed to examine the longitudinal relationship between dairy consumption and key indicators of growth, bone health, and cognitive development in Canadian children aged 5 to 11 years. The secondary objective is to assess caregivers' and children's knowledge, attitudes, practices, and perceived facilitators and barriers to dairy consumption.

Methods: A longitudinal convergent parallel mixed methods study is being conducted with 4 data collection points over 2.5 years (baseline, and approximately 10, 20, and 30 months). Eligible participants include children aged 5 to 11 years at baseline, residing in the Saskatoon area, with no diagnosed conditions or medication use affecting bone metabolism or cognitive development. Quantitative data include dietary intake, anthropometrics, body composition, bone density and microarchitecture, subjective and objective measures of physical activity, cognitive assessments, and a salivary biomarker. The qualitative component aims to investigate facilitators and barriers to dairy consumption through interviews with caregivers. Mixed-effects models will be applied to assess longitudinal associations, and thematic analysis will be used to analyze qualitative data. Quantitative and qualitative findings will be integrated at the interpretation stage to inform overall conclusions.

Results: Ethics approval was obtained from the University of Saskatchewan Research Ethics Board (Bio-3339) in 2022. Recruitment began in September 2022 and concluded in February 2025. A convenience sample of 166 volunteer children from Saskatoon was recruited, with efforts made to achieve demographic diversity. Data collection is ongoing and expected to finish by August 2027. Data analysis is underway, with preliminary findings expected to be published in 2026.

Conclusions: This protocol outlines a comprehensive study that aims to generate context-specific evidence on the role of dairy in child development, with implications for developing responsive dietary guidance and public health policy in Canada.

International Registered Report Identifier (IRRID): DERR1-10.2196/87415

(JMIR Res Protoc 2026;15:e87415) doi:10.2196/87415

KEYWORDS

dairy consumption; child development; bone health; cognitive development; diet quality

XSL•FO
RenderX

Introduction

Childhood is a period of rapid growth and development, during which a balanced diet is essential in supporting physical growth, cognitive development, and immune function [1]. Sufficient nutrient intake during this period helps maintain stable energy levels, supports bone and tooth development, and promotes healthy body composition [2]. In contrast, inadequate intake can restrict growth, impair cognitive performance, increase susceptibility to infections, and elevate the risk of obesity and chronic diseases later in life [2,3].

Dairy products are a rich source of nutrients such as high-quality protein, calcium, vitamin D, magnesium, and other minerals that are essential for skeletal development and overall growth [4,5]. Studies show that regular dairy consumption supports linear growth, influences body composition, and contributes to bone mineral accrual, potentially through growth factors such as insulin-like growth factor-1 (IGF-1) [6-8]. In many developed countries, dairy products are a major source of nutrients for children, supplying over 45% of calcium, 34% of iodine, 20% of potassium, 20% - 40% of vitamin B12, and 20% of vitamin A [4]. These nutrients are essential not only for growth and bone health but also for nerve function, muscle activity, and immune defense [9,10].

Emerging research considers dairy as a complex nutrient matrix (dairy matrix) where interactions between nutrients can improve bioavailability and produce synergistic effects, particularly regarding bone development [11,12]. Meta-analyses and randomized controlled trials have suggested that dairy consumption enhances bone mineral content (BMC) and bone mineral density (BMD), with the strongest effects observed among children with lower baseline calcium intake [6,7,13]. However, findings for body composition are inconsistent; some studies suggest associations with favorable weight outcomes, while systematic reviews report largely nonsignificant effects on lean or fat mass [14-16]. Preliminary evidence has also linked dairy intake with cognitive outcomes, such as executive function [17], but most available evidence is from cross-sectional data, which limits causal interpretation and complicates distinguishing dietary effects from normal maturation processes [4].

Growth during childhood is influenced by more than age and nutrition. Biological maturation, which varies widely among children, affects growth rate, bone development, and body composition [18]. Without accounting for maturation status, it becomes challenging to determine whether observed differences in growth are attributable to nutrition or developmental timing.

In the Canadian context, national dietary trends show declining dairy intake among children, coinciding with increased use of plant-based alternatives that differ in nutrient density and bioavailability [19,20]. At the same time, Canada's 2019 Food Guide repositioned dairy into a broader "protein foods" category, reducing its distinct positioning in national dietary recommendations [21]. While intended to promote dietary variety, sustainability, and ultimately to support optimal public health, this change may unintentionally downplay the distinct nutritional contributions of dairy, particularly for children [22]. Because caregivers' perceptions and eating behaviors strongly shape children's dietary practices, this shift could lead some caregivers to view dairy as interchangeable with other protein sources, potentially reducing children's intake and influencing child health outcomes [12]. However, these behavioral and perceptual dimensions remain largely unexplored.

This study is informed by life course theory and a systems-oriented perspective to investigate child nutrition. Life course theory identifies childhood as a critical period during which dietary exposures, such as milk and dairy intake, may affect growth, bone health, and long-term health outcomes [23]. In addition, from the perspective of the socioecological framework of health behavior, this study positions children's dairy consumption as part of a network of influences shaped by both individual nutritional requirements and caregiver practices influenced by environmental and cultural factors [24].

Guided by these frameworks, the study examines multiple domains of child development, including growth, body composition, bone density and microarchitecture, and cognitive function, while accounting for dietary intake, physical activity, biological maturation, and sociodemographic factors. In parallel, qualitative assessments from caregivers capture modifiable behavioral and contextual determinants of dairy intake, such as perceived risks and benefits and barriers and facilitators toward dairy consumption and perceptions of dietary guidance.

So far, limited research has explored how dairy intake connects to various aspects of growth and development in school-aged children [17,25], as well as how caregiver perceptions might affect these results. To address these gaps, this study will use a longitudinal mixed methods design to track dairy consumption and developmental indicators among Canadian children aged 5 to 11 years. This approach allows for the modeling of individual growth trajectories and helps distinguish nutrient-specific effects from age-related changes, such as biological maturation. By following children over time, the study aims to strengthen causal inference and provide more detailed insight into how dairy intake may influence physical and cognitive development. The findings will contribute to evidence-based dietary guidance and public health strategies that support optimal growth and development in school-aged children.

Methods

Study Design

This study uses a longitudinal convergent parallel mixed methods design with both quantitative and qualitative components. The longitudinal framework captures developmental changes and the cumulative effects of diet over time. Quantitative measures include dietary intake, growth trajectories, bone health and body composition parameters, cognitive assessments, and a structured knowledge, attitudes, and practices (KAP) questionnaire, allowing statistical evaluation of how dairy intake and dairy perceptions relate to developmental outcomes. The qualitative component explores facilitators and barriers to dairy consumption among caregivers. Participants will be followed for 2.5 years, with assessments at 4 points approximately 10 months apart. Figure 1 presents the

XSL•FO
RenderX

study timeline. This study protocol is reported in accordance with the ObsQual Protocol Checklists, a structured guidance tool for observational study protocols [26]. The checklist is available in Multimedia Appendix 1.

Figure 1. Timelines of the study.

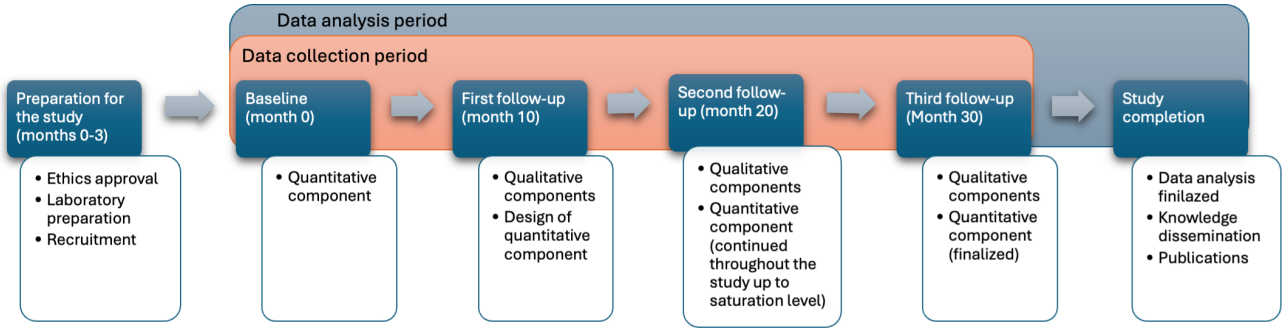

Study Population and Sample Size

Children aged 5 - 11 years were selected as the target population to capture a developmental stage characterized by steady linear growth, rapid bone mineral accrual, and meaningful cognitive advancement before the onset of puberty [2,27]. During this middle childhood period, biological maturation is variable but less pronounced than in adolescence, allowing clearer modeling of nutritional influences without the confounding effects of pubertal growth surges [28]. This age range also represents a stage in which caregivers continue to exert substantial control over children's dietary behaviors, making it an optimal window for assessing both dietary exposures and caregiver-related influences [29].

The target population for this study consists of children aged 5 - 11 years residing in Saskatoon, Canada. According to the 2016 Canadian census, Saskatoon had a total of 44,305 children aged 0 to 14 years, with 14,975 in the 5- to 9-year age group and 13,705 in the 10- to 14-year age group. The total child population consisted of 14,765 (51.5%) boys and 13,915 (48.5%) girls and included 6160 immigrant children [30].

The study sample size was determined to ensure sufficient statistical power to detect a clinically meaningful relationship between dairy intake and linear growth in children aged 5 to 11 years. Linear growth, defined as the change in height over time, served as the primary outcome. The study used a repeated-measures design with 4 data collection points at baseline, 10, 20, and 30 months. Based on prior literature and expert input, the analysis assumed a 2-sided α of .05% and 80% power. The residual SD of height was estimated at 0.5 cm, and the intrasubject correlation was set at 0.8, reflecting expected consistency in individual height measurements over time [31]. The SD for daily dairy intake was approximated at 1 serving per day [32,33]. A clinically meaningful effect was defined as an additional 0.2 cm of height gain over 30 months per extra daily serving of dairy, corresponding to approximately 0.0067 cm per month per serving. This estimate was informed by meta-analytic findings, suggesting a 0.4 cm per year increase in height associated with an additional 245 mL of milk consumed daily [34]. Using a power analysis approximation for a linear mixed-effects model that incorporates the effect of a continuous covariate (dairy intake) on growth rate, the required sample size was estimated at 87 participants. To account for an anticipated 25% attrition over the 30-month study period, the recruitment target was increased to 116 participants.

A combination of convenience and snowball sampling methods was used to recruit participants, reflecting the demographic characteristics of children in Saskatoon. Recruitment was monitored to maintain a balance between boys and girls, and demographic variables such as age and newcomer status were tracked to ensure proportional representation approximating the actual population. Inclusion criteria are (1) children aged 5 to 11 years, (2) healthy status (as determined by parental report and study screening), and (3) no current treatment for malnutrition or other serious medical conditions. Exclusion criteria include (1) chronic diseases or medical conditions affecting growth and development, (2) chronic medication use, and (3) medications known to affect bone metabolism or calcium balance. Retention efforts will rely on flexible scheduling, reminders, and small incentives.

Study Setting

The study is being conducted in Saskatoon, Canada, leveraging various recruiting channels, including libraries, recreational and community centers, social media, and online platforms, to ensure a diverse and representative sample. Assessments are being conducted in research facilities at the University of Saskatchewan.

Ethical Considerations

Ethics approval has been obtained from the University of Saskatchewan Ethics Board (Bio-3339). Written informed consent was obtained from the parents or legal guardians of all participating children. During the consent process, parents or guardians received detailed information about the study's purpose, procedures, potential risks, and benefits and had the opportunity to ask questions before signing the consent form. Age-appropriate assent was also obtained from each child to ensure understanding and voluntary participation. Children were informed that they could withdraw from the study at any time without consequence. To protect participant privacy and confidentiality, all data are stored on secure, access-restricted servers housed at the University of Saskatchewan. Identifiable information is maintained separately from study data and replaced with unique coded identifiers. Only authorized members of the research team have access to the linkage file. Any published results will present aggregated findings only,

XSL•FO
RenderX

ensuring that no individual participant can be identified. Parents and participating children will receive 1 gift card per follow-up visit as compensation, provided in accordance with institutional guidelines. Compensation is distributed uniformly to all participants and is not dependent on full study completion, which minimizes the risk of coercion or undue influence.

Data Collection

Overview

Data are collected at months 0, 10, 20, and 30 using both quantitative and qualitative methods outlined in Table 1. These repeated measurements allow the study to track changes over time and capture comprehensive information on the participants.

XSL•FO
RenderX

Table . Data collection instruments and time points.

Type of data and variable	Instrument or method	Month 0	Month 10	Month 20	Month 30
Contact information and consent	Caregiver consent form and contact information form	✓			
Quantitative					
Sociodemographic and socioeconomic status	Questionnaire (caregiver-reported)	✓			
Food security questionnaire	HFSSM[a]	✓			
Anthropometric measures	Standardized measurement protocol	✓	✓	✓	✓
Dietary intake	Multiple 24-hour dietary recalls (interviewer-administered)	✓	✓	✓	✓
Habitual dairy intake	Dairy Food Frequency Questionnaire	✓	✓	✓	✓
Physical activity (self-report)	PAQ-C[b]	✓	✓	✓	✓
Physical activity (Objective)	Accelerometers		✓		✓
Sun exposure	Sun exposure questionnaire	✓	✓	✓	✓
Skin pigmentation	Konica Minolta spectrophotometer	✓			✓
Body composition and bone health	DXA[c]	✓	✓	✓	✓
Bone microarchitecture and density	HR-pQCT[d]	✓	✓		
Biomarker of growth and development	Saliva sample for IGF-1[e] analysis	✓		✓	✓
Cognitive function	KABC-II[f]		✓		✓
Child dairy knowledge, attitudes, and practices toward dairy product consumption	KAP[g] survey		✓		✓
Caregiver knowledge, attitudes, and practices toward dairy product consumption	KAP survey		✓		
Qualitative					
Facilitators and barriers influencing child dairy consumption	Semistructured interview with caregivers		✓		

[a]HFSSM: Household Food Security Survey Module.

[b]PAQ-C: Physical Activity Questionnaire for Children.

[c]DXA: dual-energy x-ray absorptiometry.

[d]HR-pQCT: high-resolution peripheral quantitative computed tomography.

[e]IGF-1: insulin-like growth factor 1.

[f]KABC-II: Kaufman Assessment Battery for Children, Second Edition.

[g]KAP: knowledge, attitudes, and practices.

XSL•FO
RenderX

Quantitative Component

Overview

We assess dietary intake, physical activity, anthropometrics, body composition (including bone, fat, and muscle mass), cognitive performance, and a salivary biomarker of growth and development. At baseline, we also collect sociodemographic and socioeconomic data using standardized questionnaires from the Canadian Community Health Survey (CCHS) [35]. These instruments record age, sex, ethnicity, household income, education, source of income, immigration status, and duration of residence in Canada. Using the CCHS format allows us to compare our sample with national population data, strengthening the generalizability of the results.

Dietary Assessment

To obtain a reliable estimate of individual dietary intake, a combination of serial 24-hour dietary recalls and a dairy-specific Food Frequency Questionnaire (FFQ) is administered at each of the 4 data collection time points [36,37]. Three 24-hour dietary recalls are scheduled approximately 10 to 14 days apart and on different days of the week to capture day-to-day variation. These recalls follow the multiple-pass method, with the primary caregiver reporting as a proxy to ensure accuracy [38,39]. The mean of the three 24-hour recalls at each time point provides an estimate of usual dietary intake, and the ESHA Food Processor (version 11.6.522; ESHA Research) is used to determine the nutrient composition of the foods. The dairy FFQ, which assesses the frequency of dairy consumption over the previous 6 months, is adopted from the Canadian Health Measures Survey Food Frequency Questions that are specific to milk and dairy products and plant-based dairy alternatives. This questionnaire comprehensively assesses the dairy food group compared to 24-hour recall data at the national level [40].

Food Security

At baseline, food security status is being assessed using the US Department of Agriculture Household Food Security Survey Module [41] and interpreted using the PROOF guide for measurement in Canadian contexts [42]. The Household Food Security Survey Module contains 18 questions about household food security over the past 12 months. Ten of the items are specific to the experiences of adults in the household, and the remaining 8 items are specific to the experiences of children (<18 years). Responses are scored in accordance with US Department of Agriculture item-affirmative criteria, and food security categories (food secure, marginal, moderate, and severe) are assigned using thresholds adapted per the PROOF guide [42]. Both household-level and child-level food insecurity are derived from the corresponding adult and child items. This assessment will allow an analysis of the potential role of food security status in food choices, particularly dairy foods.

Anthropometric Measurements

Anthropometric data are being collected at all time points following standardized protocols. Standing height is measured to the nearest 0.1 cm using a wall-mounted stadiometer. Participants stand with their shoes off, heels against the stadiometer, arms by their sides, and eyes aligned with the Frankfurt horizontal plane, taking a deep breath [43]. Sitting height is measured from the head's vertex to the seat surface and will be used to estimate age at peak height velocity (PHV), a proxy for somatic maturity status. PHV reflects the timing of the adolescent growth spurt and is a widely used noninvasive indicator of biological maturity in children [44]. Body weight is recorded to the nearest 0.01 kg using a calibrated SECA electronic scale, with participants wearing light indoor clothing and no footwear or jewelry [45]. An average of 3 measurements taken for all anthropometric variables is used to minimize random measurement error [45].

Body Composition

Body composition, including lean mass, fat mass, and percent body fat, is measured using dual-energy x-ray absorptiometry (DXA; Hologic Discovery-Wi) at each assessment time point. DXA is the gold standard method for evaluating body composition in children and adolescents, given its minimal radiation exposure and noninvasive nature [46]. All scans are performed by certified radiology technicians following the manufacturer's recommended calibration and positioning procedures.

Bone Health Assessments

Dual-Energy X-Ray Absorptiometry

BMC and areal BMD are measured using DXA at the total body, hip, and lumbar spine during each of the 4 data collection time points. DXA provides highly reproducible estimates of areal BMD and BMC, enabling the evaluation of skeletal mass and mineralization [46].

High-Resolution Peripheral Quantitative Computed Tomography

Bone microarchitecture and compartment-specific volumetric BMD are being assessed with high-resolution peripheral quantitative computed tomography (HR-pQCT) at the distal radius and tibia. This low-radiation, noninvasive technique generates high-resolution 3D images of trabecular and cortical compartments, allowing detailed analysis of volumetric BMD, bone geometry, and microstructural integrity, which are important indicators of bone strength and fracture risk [47].

Physical Activity Assessment

Physical activity is measured using both subjective and objective methods. For subjective assessment, the Physical Activity of Children and Youth module from the CCHS is administered [48]. This module forms part of the health components within the CCHS and collects information on physical activity behaviors among children and youths. It is used nationally to monitor movement behaviors and complements other health and nutrition data collected through the survey. To complement self-reported data, accelerometers are used to provide objective measures of physical activity. Worn on the hip for 7 consecutive days, these devices record continuous data on movement patterns, including time spent in sedentary, light, moderate, and vigorous activity levels.

Skin Tone and Sun Exposure

Using the related Canadian Health Measures Survey, outdoor activity and duration of sun exposure are measured based on the number of hours and the period of the day. Further, because

XSL•FO
RenderX

the study includes children from different ethnic backgrounds with varying skin pigmentation (important for UV absorption and vitamin D synthesis), a compact portable spectrophotometer (CM-600d, Konica Minolta) is used with skin analysis software to measure skin pigmentation [49].

Salivary Biomarker of Growth and Development

In this study, salivary IGF-1 is examined as a noninvasive biomarker of growth and development in children [50]. Saliva samples are collected using the Oasis Pure•SAL-C Kit between 1 PM and 4 PM to maintain uniform conditions and minimize diurnal variation in IGF-1 levels. The collected saliva samples are analyzed using enzyme-linked immunosorbent assay techniques. Enzyme-linked immunosorbent assay is a highly sensitive and specific method that allows for the quantification of IGF-1 concentrations in biological samples [51].

Cognitive Assessments

Cognitive functioning in children is assessed using the Kaufman Assessment Battery for Children, Second Edition (KABC-II), a standardized, individually administered instrument designed to evaluate a broad range of cognitive abilities in children aged 3 to 18 years [52]. The KABC-II is grounded in contemporary theories of intelligence, including the Cattell-Horn-Carroll model and Luria's neuropsychological framework, making it particularly appropriate for use in diverse populations [53]. It provides composite scores across key cognitive domains, including sequential processing, simultaneous processing, learning, planning, and knowledge, as well as an overall Fluid-Crystallized Index [53]. In this study, the KABC-II is administered to children at 2 time points (months 10 and 30). The assessment typically requires 90 to 150 minutes to complete and is conducted in a quiet, child-friendly environment by trained research personnel. Administration and scoring follow standardized procedures outlined in the KABC-II manual [53].

KAP Survey

Two separate KAP questionnaires on dairy foods were developed for caregivers and children following the Food and Agriculture Organization guidelines for assessing nutrition-related KAP [54] and grounded in the health belief model [55]. The caregiver questionnaire is administered once, at month 10, to capture baseline KAP regarding dairy consumption, while the child questionnaire is administered at 2 time points (months 10 and 30) to assess changes over time. Both questionnaires were developed by the research team and reviewed for content relevance, clarity, and age appropriateness. In combination with children's dietary intake data, these questionnaires capture KAP regarding dairy foods within participating families.

Qualitative Component

The qualitative component of the study consists of semistructured interviews with a subset of primary caregivers involved in the study. The interviews explore facilitators and barriers to milk and dairy product consumption. Participants are selected from the overall study sample using purposive sampling to ensure representation across ethnic and socioeconomic groups, with interviews continuing until data saturation is achieved [56,57]. Interviews are conducted in private, face-to-face settings convenient for caregivers, and are audio-recorded. Each interview lasts about 30 minutes with informed consent.

Data Management

All data are managed in accordance with the project's Data Management Plan and institutional ethics protocols, ensuring confidentiality, data integrity, and controlled access for authorized personnel only. Researchers may request access to nonrestricted anonymized data by submitting a proposal to the principal investigator (HV). As this is a protocol, no data currently exist.

Quantitative data are captured using REDCap (Research Electronic Data Capture; Vanderbilt University), a secure web-based system. Quantitative data from field and laboratory measurements (eg, dietary recalls, anthropometry, sociodemographics, food security, salivary IGF-1, DXA, and HR-pQCT) are entered directly into REDCap. Additional quantitative datasets, such as KABC-II results and accelerometry outputs, are exported from specialized software as digital or CSV files and stored securely in the institutional repository. Dietary recalls data are coded with the ESHA Food Processor software and exported as CSV files. Qualitative data, including audio-recorded interviews and REDCap interview notes, are transcribed, anonymized, and structured, then imported into NVivo (QSR International) for analysis, with the original audio files archived in secure formats.

Quality Assurance and Monitoring

The study incorporates a comprehensive quality assurance and monitoring strategy to ensure the collection of high-quality and reliable data. A foundational element of this approach is the rigorous training of research assistants (RAs). All RAs participate in training sessions that include an overview of the study protocol, ethical procedures, data collection methods, as well as instruction in using REDCap for data entry and NVivo for qualitative data analysis. The training combines theoretical instruction with hands-on demonstrations to build competency, with refresher sessions held periodically to reinforce protocol adherence and address emerging challenges.

Ongoing supervision and regular data audits are conducted to uphold data integrity throughout the study. Supervisors perform routine checks on field activities to ensure compliance with research protocols and provide immediate, constructive feedback to RAs. Randomly selected data entries are audited for accuracy, completeness, and consistency. These audits include a thorough examination of missing values, outlier handling, and the calculation of derived variables. Any discrepancies identified are promptly corrected and documented, ensuring transparency and continued quality control.

Outcomes

The primary outcome of this study is the association between dairy consumption and bone and body composition outcomes in children aged 5 to 11 years. Bone outcomes are assessed using DXA, including BMC and BMD at the lumbar spine, femur, and total body, as well as lumbar spine bone mineral apparent density. Body composition outcomes are also derived

XSL•FO
RenderX

from DXA and include total body muscle mass, fat mass, and percent body fat. These measures are used to describe how dairy intake may be linked to skeletal development and overall body composition during middle childhood.

Secondary outcomes include broader indicators of growth and development. The study evaluates how milk and dairy foods contribute to total dietary intake and overall diet quality. It also examines associations between dairy consumption and cognitive development, as well as hormones and biomarkers connected to growth, including IGF-1. These outcomes help clarify how dairy intake may relate to developmental processes beyond bone health.

The study also investigates KAP related to dairy consumption among children and their caregivers. This includes an assessment of factors that influence dairy intake over time, such as household routines, access to foods, and perceived supports or challenges. Quantitative measures are used to identify facilitators and barriers to dairy consumption among Canadian children. Together, these outcomes provide a detailed picture of how dairy intake fits within patterns of growth, development, dietary behavior, and family context. By addressing these primary and secondary outcomes, the study aims to provide comprehensive insights into the role of dairy consumption in the growth, development, and overall health of Canadian children. The findings will contribute to evidence-based dietary recommendations and interventions aimed at improving child health outcomes.

Data Analysis Plan

The research uses a mixed methods approach, using both quantitative and qualitative data. Quantitative data analyses follow the internal data processing protocol, beginning with data cleaning and preparation. This process includes checking for data entry errors, examining distributions, and identifying missing data and outliers. Data entry issues are evaluated by inspecting frequencies, identifying misspelled entries for categorical variables, and examining minimum and maximum values, means, histograms, and ranges for continuous variables. Loss to follow-up will be addressed by comparing baseline characteristics of participants who complete the study with those who withdraw. These comparisons will draw on standardized mean differences and logistic regression to identify differences in demographics, dairy intake, and growth measures [58]. Missing data are addressed using multiple imputation or maximum likelihood methods, depending on the pattern of missingness [59].

Once cleaned, the dataset is used to generate derived variables relevant to the study objectives. These include anthropometric indicators such as BMI, BMI *z* scores, height-for-age, and weight-for-age *z* scores; dietary measures such as total dairy intake, calcium intake, and diet quality indices; and consumption groupings based on children's dairy intake relative to national dietary guidelines. Descriptive statistics summarize participant characteristics and key outcomes. Continuous variables are reported as means with SDs or as medians with IQRs, and categorical variables are summarized as counts and percentages.

Longitudinal associations between dairy intake and outcomes such as somatic growth, body composition, bone health, and cognitive development are evaluated using mixed-effects models and generalized estimating equations. Mixed-effects models account for repeated measures within participants and include random intercepts and slopes where appropriate [60]. Generalized estimating equation estimates population-averaged effects while accounting for within-subject correlation over time [60]. Covariates, including age, sex, socioeconomic status, physical activity, and total energy intake, are incorporated to control for potential confounding. Subgroup analyses will be conducted for male and female participants separately to explore sex-specific associations and developmental trajectories. Interaction effects by sex and baseline diet quality are also examined. Sensitivity analyses assess the robustness of findings, including alternative categorizations of dairy intake and different approaches to handling missing data. Exploratory analyses may be used to examine associations between diet quality indices and secondary outcomes such as bone microarchitecture using regression or correlation models. All quantitative analyses are performed using R (R Foundation for Statistical Computing) or SPSS (IBM Corp) software, in accordance with internal protocols to ensure reproducibility and transparency.

Qualitative data from caregiver interviews and child surveys are analyzed using inductive thematic content analysis. The analysis begins with familiarization through transcript review, followed by iterative coding using NVivo software. Codes are grouped into preliminary themes, which are then reviewed and refined. To ensure consistency and reduce bias, a second coder independently codes approximately 15% of the data using a predefined codebook. Intercoder agreement is quantified using Cohen κ, and any discrepancies are resolved through discussion. Thematic saturation is monitored throughout, and analysis continues until saturation is achieved.

Knowledge Translation

The study uses a comprehensive knowledge translation strategy involving strong collaboration with key partners such as Osteoporosis Canada, public health agencies, community-based child health organizations, and immigrant and cultural health advocacy groups. A dedicated working group will coordinate stakeholder engagement and tailor dissemination strategies for specific audiences, including the public, service providers, policymakers, and academics. Efforts will prioritize cultural sensitivity, particularly for immigrant and Indigenous communities. Dissemination will include policy briefs, workshops, peer-reviewed publications, and conference presentations. Continuous evaluation will guide adaptive improvements to ensure broad impact and sustained use.

Results

Recruitment began in September 2022 and concluded in February 2025, following ethics approval by the University of Saskatchewan Research Ethics Board (Bio-3339) in October 2022. A total of 116 children and their primary caregivers were enrolled. Participants are being followed across 4 visits spaced approximately 10 months apart. Retention rates have remained strong, with 102 participants completing the second visit, 47

XSL•FO
RenderX

completing the third visit, and 25 completing the fourth visit to date. Data collection is ongoing and is expected to continue through August 2027. Preliminary study findings are expected to be published beginning in fall 2026, with initial descriptive results on dietary patterns and KAP, providing early insights into dairy consumption behaviors among the study population.

Discussion

To the best of our knowledge, this is the first study in Canada investigating the longitudinal relationship between dairy consumption and child growth, bone health, and cognitive development using a mixed methods design. This study is expected to clarify how dairy intake relates to child growth, bone health, and cognitive development. It also considers the social and behavioral factors that shape dietary patterns. The design brings together quantitative measures of dietary intake, anthropometry, body composition, bone health, and cognitive assessments, with qualitative insights from caregivers. Taken together, the mixed methods design allows the study to describe both the biological outcomes and the circumstances that influence how children consume dairy foods.

While the majority of prior research on dairy intake in children originates from cross-sectional studies [4], this longitudinal study enables the assessment of patterns across multiple time points, providing a more comprehensive understanding of the evolution of dietary intake and growth and development outcomes. This approach addresses the knowledge gap in the longitudinal studies. For example, the Iowa Fluoride Study in the United States found that consistent milk consumption from age 2 to 17 years was associated with a small increase in adult height, specifically 0.39 cm for each extra 8-ounce daily serving [61]. The Generation R Study in the Netherlands showed that adherence to a “dairy and whole grains” pattern was positively associated with BMD and area-adjusted BMC in children aged 8 to 10 years [62]. In China, a randomized trial in preschool children reported an increase in arms BMD and BMC with milk supplementation [6]. However, few studies have followed children over time while also considering how caregivers shape food choices. By focusing on Canadian children aged between 5 and 11 years, this study captures a period of consistent growth and bone development before the onset of maturation complicates nutritional models. It also extends findings from US preschool cohorts, in which children who consumed more milk were taller but also showed higher BMI. Validated tools ensure data quality, including DXA/HR-pQCT for bone, multiple 24-hour recalls or FFQs for diet, and standardized cognitive assessments [36,37,39,52].

Despite its strengths, the study has potential limitations that could affect the validity of its findings. Selection bias could occur if participants are healthier or of higher socioeconomic status than the general population; this risk is mitigated through targeted recruitment weighted against census data [58,63]. Measurement bias is also a concern, particularly in dietary recalls or anthropometric assessments [64]. For example, 24-hour dietary recalls in children underestimate intake by 15% - 24% due to omissions, especially of snacks, as validated against observational methods [65]. To minimize these risks, rigorous training for research staff, standardized data collection protocols, and validated dietary assessment tools have been implemented. Differences in age and maturity, such as early puberty in girls influencing growth and bone development, are addressed by controlling for age at PHV [66]. Residual confounding, attrition, and dairy heterogeneity (eg, milk vs yogurt) remain challenges, which can be addressed using mixed-effects models and sensitivity analyses.

The extent to which findings from this study can inform broader Canadian dietary guidance also needs to be considered. While CCHS data show that suboptimal adherence to dietary recommendations is common across regions [67], Saskatoon represents a mid-sized Prairie urban context with distinct demographic, socioeconomic, and cultural characteristics that may shape dairy intake behaviors [68]. Therefore, the study population should be viewed as illustrative rather than fully representative of all Canadian children. Nevertheless, insights into longitudinal dietary patterns, developmental outcomes, and caregiver-reported barriers provide meaningful evidence on mechanisms relevant to national policy, particularly when interpreted alongside nationally representative data.

Overall, this study provides evidence on the effects of dairy consumption on child development and identifies facilitators and barriers to dairy intake from the perspectives of both caregivers and children. The findings are expected to inform responsive dietary guidance, influence public health policy, and guide future interventions to support optimal child nutrition and development.

Acknowledgments

This study is funded by the Dairy Farmers of Canada (DFC) and Mitacs. The financial support provided by DFC and Mitacs facilitates the comprehensive data collection and analysis required for this research. The research team maintains full independence in the design, implementation, analysis, and reporting of the study. The funding organizations have no influence over the study's outcomes or the interpretation of the data.

Funding

This study is funded by the Dairy Farmers of Canada and Mitacs, which facilitate the comprehensive data collection and analysis required for this research. The research team maintains full independence in the design, implementation, analysis, and reporting of the study. The funding organization has no influence over the study’s outcomes or the interpretation of the data.

XSL•FO
RenderX

Authors' Contributions

HV, ZLL, and MS led the conceptualization and design of the study and developed the initial research proposal in collaboration with coinvestigators GL, AB-J, MCE, and WLS. KK was responsible for drafting and revising the manuscript, coordinating internal reviews, and overseeing project coordination.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Facilitators and barriers to dairy product consumption—interview guide.
[DOCX File, 21 KB - resprot_v15i1e87415_app1.docx]

Peer Review Report 1
Peer review report by the Dairy Farmers of Canada.
[PDF File, 585 KB - resprot_v15i1e87415_app2.pdf]

References

1. Mummert A, Schoen M, Lampl M. Growth and life course health development. In: Halfon N, Forrest CB, Lerner RM, Faustman EM, editors. Handbook of Life Course Health Development: Springer International Publishing; 2018:405-429. [doi: 10.1007/978-3-319-47143-3_17]
2. Saavedra JM, Prentice AM. Nutrition in school-age children: a rationale for revisiting priorities. Nutr Rev 2023 Jun 9;81(7):823-843. [doi: 10.1093/nutrit/nuac089] [Medline: 36346900]
3. DiGirolamo AM, Ochaeta L, Flores RMM. Early childhood nutrition and cognitive functioning in childhood and adolescence. Food Nutr Bull 2020 Jun;41(1_suppl):S31-S40. [doi: 10.1177/0379572120907763] [Medline: 32522121]
4. Dougkas A, Barr S, Reddy S, Summerbell CD. A critical review of the role of milk and other dairy products in the development of obesity in children and adolescents. Nutr Res Rev 2019 Jun;32(1):106-127. [doi: 10.1017/S0954422418000227] [Medline: 30477600]
5. Dror DK, Allen LH. Dairy product intake in children and adolescents in developed countries: trends, nutritional contribution, and a review of association with health outcomes. Nutr Rev 2014 Feb;72(2):68-81. [doi: 10.1111/nure.12078] [Medline: 24330063]
6. Zhao ZF, Li BY, He Q, et al. Impact of dairy supplementation on bone acquisition in children's limbs: a 12-month cluster-randomized controlled trial and meta-analysis. Arch Osteoporos 2024 Jul 24;19(1):65. [doi: 10.1007/s11657-024-01422-2] [Medline: 39043915]
7. Hidayat K, Zhang LL, Rizzoli R, et al. The effects of dairy product supplementation on bone health indices in children aged 3 to 18 years: a meta-analysis of randomized controlled trials. Adv Nutr 2023 Sep;14(5):1187-1196. [doi: 10.1016/j.advnut.2023.06.010] [Medline: 37414219]
8. Hohoff E, Perrar I, Jankovic N, Alexy U. Dairy intake and long-term body weight status in German children and adolescents: results from the DONALD study. Eur J Nutr 2022 Mar;61(2):1087-1096. [doi: 10.1007/s00394-021-02715-9] [Medline: 34718860]
9. Terrell K, Choi S, Choi S. Calcium's role and signaling in aging muscle, cellular senescence, and mineral interactions. Int J Mol Sci 2023 Dec 1;24(23):17034. [doi: 10.3390/ijms242317034] [Medline: 38069357]
10. Sharma D, Singh A, Prashar Y. Vitamin B12: a cornerstone of cognitive and metabolic well-being in young adults. Obes Med 2025 Jan;53:100563. [doi: 10.1016/j.obmed.2024.100563]
11. Thorning TK, Bertram HC, Bonjour JP, et al. Whole dairy matrix or single nutrients in assessment of health effects: current evidence and knowledge gaps. Am J Clin Nutr 2017 May;105(5):1033-1045. [doi: 10.3945/ajcn.116.151548] [Medline: 28404576]
12. Wang K, Zhao X, Yang S, Qi X, Li A, Yu W. New insights into dairy management and the prevention and treatment of osteoporosis: the shift from single nutrient to dairy matrix effects—a review. Compr Rev Food Sci Food Saf 2024 Jul;23(4):e13374. [doi: 10.1111/1541-4337.13374] [Medline: 38847750]
13. van den Heuvel E, Steijns J. Dairy products and bone health: how strong is the scientific evidence? Nutr Res Rev 2018 Dec;31(2):164-178. [doi: 10.1017/S095442241800001X] [Medline: 29560832]
14. Kouvelioti R, Josse AR, Klentrou P. Effects of dairy consumption on body composition and bone properties in youth: a systematic review. Curr Dev Nutr 2017 Aug;1(8):e001214. [doi: 10.3945/cdn.117.001214] [Medline: 29955717]
15. Kang K, Sotunde OF, Weiler HA. Effects of milk and milk-product consumption on growth among children and adolescents aged 6-18 years: a meta-analysis of randomized controlled trials. Adv Nutr 2019 Mar 1;10(2):250-261. [doi: 10.1093/advances/nmy081] [Medline: 30839054]

XSL•FO
RenderX

16. Babio N, Becerra-Tomás N, Nishi SK, et al. Total dairy consumption in relation to overweight and obesity in children and adolescents: a systematic review and meta-analysis. Obes Rev 2022 Jan;23 Suppl 1:e13400. [doi: 10.1111/obr.13400] [Medline: 34881504]
17. Zeng X, Cai L, Gui Z, et al. Association between dairy intake and executive function in Chinese children aged 6-12 years. Front Nutr 2022;9:879363. [doi: 10.3389/fnut.2022.879363] [Medline: 35898711]
18. Sanders JO, Qiu X, Lu X, et al. The uniform pattern of growth and skeletal maturation during the human adolescent growth spurt. Sci Rep 2017 Dec 1;7(1):16705. [doi: 10.1038/s41598-017-16996-w] [Medline: 29196711]
19. Desmond MA, Fewtrell MS, Wells JCK. Plant-based diets in children: secular trends, health outcomes, and a roadmap for urgent practice recommendations and research—a systematic review. Nutrients 2024 Mar 1;16(5):723. [doi: 10.3390/nu16050723] [Medline: 38474851]
20. Vatanparast H, Islam N, Shafiee M. Consumption of milk and alternatives decreased among Canadians from 2004 to 2015: evidence from the Canadian Community Health Surveys. BMC Nutr 2021 Nov 1;7(1):63. [doi: 10.1186/s40795-021-00465-9] [Medline: 34719398]
21. Canada's Dietary Guidelines for Health Professionals and Policy Makers: Health Canada; 2019.
22. Fernandez MA, Bertolo RF, Duncan AM, et al. Translating "protein foods" from the new Canada's Food Guide to consumers: knowledge gaps and recommendations. Appl Physiol Nutr Metab 2020 Dec;45(12):1311-1323. [doi: 10.1139/apnm-2020-0192]
23. Gallegos D. The life-course, developmental origins of health and disease, and nutrition. In: Food and Nutrition: Routledge; 2026:367-378. [doi: 10.4324/9781032696133-31]
24. El-Ammari A, El Kazdouh H, Bouftini S, El Fakir S, El Achhab Y. Social-ecological influences on unhealthy dietary behaviours among Moroccan adolescents: a mixed-methods study. Public Health Nutr 2020 Apr;23(6):996-1008. [doi: 10.1017/S1368980019003641] [Medline: 31896381]
25. Vogel KA, Martin BR, McCabe LD, et al. The effect of dairy intake on bone mass and body composition in early pubertal girls and boys: a randomized controlled trial. Am J Clin Nutr 2017 May;105(5):1214-1229. [doi: 10.3945/ajcn.116.140418] [Medline: 28330908]
26. Low GK, Subedi S, Omosumwen OF, et al. Development and validation of observational and qualitative study protocol reporting checklists for novice researchers (ObsQual checklist). Eval Program Plann 2024 Oct;106:102468. [doi: 10.1016/j.evalprogplan.2024.102468] [Medline: 39029287]
27. Sopher AB, Fennoy I, Oberfield SE. An update on childhood bone health: mineral accrual, assessment and treatment. Curr Opin Endocrinol Diabetes Obes 2015 Feb;22(1):35-40. [doi: 10.1097/MED.0000000000000124] [Medline: 25517023]
28. Salami S, Bandeira PFR, Mashhadi MR, et al. Associations between biological maturation, chronological age, body mass index, sex, and motor competence in prepubertal children: a network analysis. Children (Basel) 2024 Sep 20;11(9):1143. [doi: 10.3390/children11091143] [Medline: 39334675]
29. Yee AZH, Lwin MO, Ho SS. The influence of parental practices on child promotive and preventive food consumption behaviors: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017 Apr 11;14(1):47. [doi: 10.1186/s12966-017-0501-3] [Medline: 28399881]
30. Census profile, 2016 census—Saskatoon, city [census subdivision], Saskatchewan and Canada [country]. Government of Canada. 2017. URL: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CSD&Code1=4711066&Geo2=PR&Code2=01&Data=Count&SearchText=4711066&SearchType=Begins&SearchPR=01&B1=All&Custom=&TABID=3 [accessed 2026-04-02]
31. Kelly A, Winer KK, Kalkwarf H, et al. Age-based reference ranges for annual height velocity in US children. J Clin Endocrinol Metab 2014 Jun;99(6):2104-2112. [doi: 10.1210/jc.2013-4455] [Medline: 24601728]
32. Fiorito LM, Mitchell DC, Smiciklas-Wright H, Birch LL. Dairy and dairy-related nutrient intake during middle childhood. J Am Diet Assoc 2006 Apr;106(4):534-542. [doi: 10.1016/j.jada.2006.01.005] [Medline: 16567149]
33. Huh SY, Rifas-Shiman SL, Rich-Edwards JW, Taveras EM, Gillman MW. Prospective association between milk intake and adiposity in preschool-aged children. J Am Diet Assoc 2010 Apr;110(4):563-570. [doi: 10.1016/j.jada.2009.12.025] [Medline: 20338282]
34. de Beer H. Dairy products and physical stature: a systematic review and meta-analysis of controlled trials. Econ Hum Biol 2012 Jul;10(3):299-309. [doi: 10.1016/j.ehb.2011.08.003] [Medline: 21890437]
35. Census profile, 2016 census. Government of Canada. 2016. URL: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E [accessed 2026-04-02]
36. Thompson FE, Kirkpatrick SI, Subar AF, et al. The National Cancer Institute's dietary assessment primer: a resource for diet research. J Acad Nutr Diet 2015 Dec;115(12):1986-1995. [doi: 10.1016/j.jand.2015.08.016] [Medline: 26422452]
37. Thomson CA, Giuliano A, Rock CL, et al. Measuring dietary change in a diet intervention trial: comparing food frequency questionnaire and dietary recalls. Am J Epidemiol 2003 Apr 15;157(8):754-762. [doi: 10.1093/aje/kwg025] [Medline: 12697580]
38. Moshfegh AJ, Rhodes DG, Baer DJ, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. Am J Clin Nutr 2008 Aug;88(2):324-332. [doi: 10.1093/ajcn/88.2.324] [Medline: 18689367]

XSL•FO
RenderX

39. Foster E, Bradley J. Methodological considerations and future insights for 24-hour dietary recall assessment in children. Nutr Res 2018 Mar;51:1-11. [doi: 10.1016/j.nutres.2017.11.001] [Medline: 29673539]
40. Hosseini Z, Whiting SJ, Vatanparast H. Canadians' dietary intake from 2007 to 2011 and across different sociodemographic/lifestyle factors using the Canadian Health Measures Survey cycles 1 and 2. J Nutr Metab 2019;2019:2831969. [doi: 10.1155/2019/2831969] [Medline: 30867963]
41. Quality of life indicator: food security. Government of Canada. 2023. URL: https://www.statcan.gc.ca/hub-carrefour/quality-life-qualite-vie/prosperity-prosperite/food-security-securite-alimentaire-eng.htm [accessed 2026-04-02]
42. Household food insecurity in Canada: a guide to measurement and interpretation. PROOF: Research to Identify Policy Options to Reduce Food Insecurity. 2021. URL: https://proof.utoronto.ca/wp-content/uploads/2018/11/Household-Food-Insecurity-in-Canada-A-Guide-to-Measurement-and-Interpretation.pdf [accessed 2026-04-02]
43. Warrier V, Krishan K, Shedge R, et al. Height Assessment: StatPearls Publishing; 2023.
44. Mirwald RL, Baxter-Jones ADG, Bailey DA, Beunen GP. An assessment of maturity from anthropometric measurements. Med Sci Sports Exerc 2002 Apr;34(4):689-694. [doi: 10.1097/00005768-200204000-00020] [Medline: 11932580]
45. Casadei K, Kiel J. Anthropometric Measurement: StatPearls Publishing; 2022.
46. Lohman TG, Harris M, Teixeira PJ, Weiss L. Assessing body composition and changes in body composition. Another look at dual-energy X-ray absorptiometry. Ann N Y Acad Sci 2000 May;904(1):45-54. [doi: 10.1111/j.1749-6632.2000.tb06420.x] [Medline: 10865709]
47. Zhou B, Wang J, Yu YE, et al. High-resolution peripheral quantitative computed tomography (HR-pQCT) can assess microstructural and biomechanical properties of both human distal radius and tibia: ex vivo computational and experimental validations. Bone 2016 May;86:58-67. [doi: 10.1016/j.bone.2016.02.016] [Medline: 26924718]
48. Canadian Community Health Survey—Nutrition (CCHS). Government of Canada. 2017. URL: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5049 [accessed 2026-04-02]
49. Jacobs LC, Hamer MA, Verkouteren JAC, Pardo LM, Liu F, Nijsten T. Perceived skin colour seems a swift, valid and reliable measurement. Br J Dermatol 2015 Oct;173(4):1084-1086. [doi: 10.1111/bjd.13888] [Medline: 25940018]
50. Almalki A, Thomas JT, Khan ARA, et al. Correlation between salivary levels of IGF-1, IGFBP-3, IGF-1/IGFBP3 ratio with skeletal maturity using hand-wrist radiographs. Int J Environ Res Public Health 2022 Mar 21;19(6):3723. [doi: 10.3390/ijerph19063723] [Medline: 35329407]
51. Yüksel B, Özbek MN, Mungan N, et al. Serum IGF-1 and IGFBP-3 levels in healthy children between 0 and 6 years of age. J Clin Res Pediatr Endocrinol 2011;3(2):84-88. [doi: 10.4274/jcrpe.v3i2.17] [Medline: 21750637]
52. Loomis D, Holt R, Kaufman JC, Kaufman AS. Kaufman Assessment Battery for Children. In: Weiner IB, Craighead WE, editors. The Corsini Encyclopedia of Psychology, 1st edition: Wiley; 2010:1-2. [doi: 10.1002/9780470479216]
53. Drozdick LW, Singer JK, Lichtenberger EO. The Kaufman Assessment Battery for Children—second edition and KABC-II normative update. In: Contemporary Intellectual Assessment: Theories, Tests, and Issues, 2nd edition: The Guilford Press; 2018.
54. Marías Y, Glasauer P. Guidelines for Assessing Nutrition-Related Knowledge, Attitudes and Practices: Food and Agriculture Organization of the United Nations (FAO); 2014.
55. Skinner CS, Tiro J, Champion VL. The health belief model. In: Health Behavior: Theory, Research, and Practice, 5th edition: John Wiley & Sons; 2015.
56. Thompson CM, Elliott SJ, Meyer S, Leatherdale ST, Majowicz SE. Attitudes, subjective norms and perceived behavioural control factors influencing Canadian secondary school students' milk and milk alternatives consumption. Public Health Nutr 2024;27(1):e92. [doi: 10.1017/S1368980024000661]
57. Kegler MC, Raskind IG, Comeau DL, Griffith DM, Cooper HLF, Shelton RC. Study design and use of inquiry frameworks in qualitative research published in Health Education & Behavior. Health Educ Behav 2019 Feb;46(1):24-31. [doi: 10.1177/1090198118795018] [Medline: 30227081]
58. Nohr EA, Liew Z. How to investigate and adjust for selection bias in cohort studies. Acta Obstet Gynecol Scand 2018 Apr;97(4):407-416. [doi: 10.1111/aogs.13319] [Medline: 29415329]
59. Afkanpour M, Hosseinzadeh E, Tabesh H. Identify the most appropriate imputation method for handling missing values in clinical structured datasets: a systematic review. BMC Med Res Methodol 2024;24(1):188. [doi: 10.1186/s12874-024-02310-6]
60. Vazquez O, Xu X, Yu Z. Introduction to linear, generalized linear, and mixed-effects models. In: Statistical Methods in Epilepsy: Chapman and Hall/CRC; 2024:103-123. [doi: 10.1201/9781003254515-5]
61. Marshall TA, Curtis AM, Cavanaugh JE, Warren JJ, Levy SM. Higher longitudinal milk intakes are associated with increased height in a birth cohort followed for 17 years. J Nutr 2018 Jul 1;148(7):1144-1149. [doi: 10.1093/jn/nxy071] [Medline: 29924327]
62. van den Hooven EH, Heppe DHM, Kiefte-de Jong JC, et al. Infant dietary patterns and bone mass in childhood: the Generation R Study. Osteoporos Int 2015 May;26(5):1595-1604. [doi: 10.1007/s00198-015-3033-1] [Medline: 25792489]
63. Profile table, census profile, 2021 census of population—Saskatoon, city (cy) [census subdivision], Saskatchewan. Government of Canada. 2022. URL: https://www12.statcan.gc.ca/census-recensement/2021/dp-pd/prof/details/page.

XSL•FO
RenderX

cfm?Lang=E&SearchText=saskatoon&GENDERlist=1&STATISTIClist=1&DGUIDlist=2021A00054711066&HEADERlist=0 [accessed 2026-04-02]
64. Burrows TL, Ho YY, Rollo ME, Collins CE. Validity of dietary assessment methods when compared to the method of doubly labeled water: a systematic review in adults. Front Endocrinol (Lausanne) 2019;10(850):850. [doi: 10.3389/fendo.2019.00850] [Medline: 31920966]
65. Arsenault JE, Moursi M, Olney DK, Becquey E, Ganaba R. Validation of 24-h dietary recall for estimating nutrient intakes and adequacy in adolescents in Burkina Faso. Matern Child Nutr 2020 Oct;16(4):e13014. [doi: 10.1111/mcn.13014] [Medline: 32337835]
66. Malina RM, Kozieł SM, Králik M, Chrzanowska M, Suder A. Prediction of maturity offset and age at peak height velocity in a longitudinal series of boys and girls. Am J Hum Biol 2021 Nov;33(6):e23551. [doi: 10.1002/ajhb.23551] [Medline: 33314450]
67. Ngueta G, Blanchette C, Mondor M, Moubarac JC, Lucas M. Canadians adults fail their dietary quality examination twice. Nutrients 2023 Jan 26;15(3):637. [doi: 10.3390/nu15030637] [Medline: 36771346]
68. Shafiee M, Al-Bazz S, Lane G, Szafron M, Vatanparast H. Exploring healthy eating perceptions, barriers, and facilitators among urban Indigenous peoples in Saskatchewan. Nutrients 2024 Jun 25;16(13):2006. [doi: 10.3390/nu16132006] [Medline: 38999754]

Abbreviations

BMC: bone mineral content
BMD: bone mineral density
CCHS: Canadian Community Health Survey
DXA: dual-energy x-ray absorptiometry
FFQ: Food Frequency Questionnaire
HR-pQCT: high-resolution peripheral quantitative computed tomography
IGF-1: insulin-like growth factor-1
KABC-II: Kaufman Assessment Battery for Children, Second Edition
KAP: knowledge, attitudes, and practices
PHV: peak height velocity
RA: research assistant
REDCap: Research Electronic Data Capture

Edited by A Schwartz; submitted 08.Nov.2025; this is a non–peer-reviewed article;accepted 25.Feb.2026; published 08.Apr.2026.

Please cite as:
Karbin K, Shafiee M, Longworth ZL, Baxter-Jones A, Erlandson MC, Lane G, Tsao LL, Siqueira WL, Vatanparast H
Exploring the Role of Dairy Consumption on the Growth and Development of Canadian Children: Protocol for a Longitudinal Mixed Method Research
JMIR Res Protoc 2026;15:e87415
URL: https://www.researchprotocols.org/2026/1/e87415
doi:10.2196/87415

© Karim Karbin, Mojtaba Shafiee, Zoe L Longworth, Adam Baxter-Jones, Marta C Erlandson, Ginny Lane, Ling-Ling Tsao, Walter L Siqueira, Hassan Vatanparast. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Improving Registration and Dataflows Between Pediatric Oncology Clinics and the Childhood Cancer Registry of Switzerland: Protocol for SwissPedCancer Quality Assurance Study

Yara Shoman[1,2], PhD; Lorenz Leuenberger[1,3], Dr med; Grit Sommer[1,2], PD; Julia Anna Bielicki[4,5], Prof Dr Med; Pierluigi Brazzola[6], Dr med; Sophia della Valle[2], MMed; Manuel Diezi[7], Dr med; Daniel Drozdov[8], Dr med; Fabienne Gumy-Pause[9,10], Dr med; Ana Guerreiro Stücklin[11], Prof Dr Med; Ursula M Kuehnel[2], MA; Katrin Scheinemann[12,13], MSc, Prof Dr med; Christina Schindera[1,14], PD Dr Med, PhD; Freimut Schilling[15], Dr med; Nicolas Waespe[1,10,16], PD Dr Med, PhD; Ben D Spycher[1,2], Prof Dr; Luregn J Schlapbach[17,18], Prof Dr Med; Claudia E Kuehni[1,2,16*], Prof Dr Med; Fabiën N Belle[1*], PhD; SwissPedHealth Consortium[19,21]

[1]Institute of Social and Preventive Medicine, University of Bern, Mittelstrasse 43, Bern, Switzerland

[10]CANSEARCH Research Platform for Pediatric Oncology and Hematology, Department of Pediatrics, Gynecology and Obstetrics, Faculty of Medicine, University of Geneva, Geneva, Switzerland

[11]Division of Oncology and Children's Research Center, University Children's Hospital of Zurich, Zurich, Switzerland

[12]Division of Hematology/Oncology, Children’s Hospital of Eastern Switzerland, St. Gallen, Switzerland

[13]Faculty of Health Sciences and Medicine, University of Lucerne, Lucerne, Switzerland

[14]Depaertment of Pediatric Oncology/Hematology, University Children’s Hospital Basel, Basel, Switzerland

[15]Children’s Hospital of Central Switzerland, Lucerne, Switzerland

[16]Division of Pediatric Hematology and Oncology, Department of Pediatrics, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

[17]Division of Intensive Care Oncology and Neonatology, and Children's Research Center, University Children’s Hospital Zurich, University of Zurich, Zurich, Switzerland

[18]Child Health Research Centre, The University of Queensland, Brisbane, Queensland, Australia

[19]See Acknowledgments

[2]Childhood Cancer Registry of Switzerland (ChCR), Bern, Switzerland

[21]

[3]Graduate School for Health Sciences, University of Bern, Bern, Switzerland

[4]Department of Paediatric Infectious Diseases, University Children’s Hospital Basel, Basel, Switzerland

[5]Centre for Neonatal and Paediatric Infection, St George's University, London, United Kingdom

[6]Pediatric Institute of Italian Switzerland, Bellinzona, Bellinzona, Switzerland

[7]Lausanne University Hospital (CHUV), University of Lausanne, Lausanne, Switzerland

[8]Division of Pediatric Hematology Oncology, Children’s Hospital, Kantonsspital Aarau (KSA), Aarau, Switzerland

[9]Division of Pediatric Oncology and Hematology, Department of Women, Child and Adolescent, Geneva University Hospitals (HUG), Geneva, Switzerland

[*]these authors contributed equally

Corresponding Author:
Claudia E Kuehni, Prof Dr Med
Institute of Social and Preventive Medicine, University of Bern, Mittelstrasse 43, Bern, Switzerland

Abstract

Background: Cancer registries are essential to monitor cancer incidence and survival to provide better quality cancer data for research. In Switzerland, the pediatric oncology units within pediatric hospitals actively report cancer cases, and the coding and registration team of the Childhood Cancer Registry (ChCR) enters data manually from medical files into the registry database. There are no automated data transfers or feedback loops between the pediatric oncology clinics and the ChCR. This ongoing process is time-consuming, inefficient, and a source of potential errors.

Objective: SwissPedCancer aims to explore the options for automated data transfers from clinical data warehouses and feedback loops to make cancer registry processes more efficient.

Methods: SwissPedCancer is a nested project within the national data stream initiative, the Swiss Pediatric Personalized Research Network (SwissPedHealth). Since September 2022, SwissPedHealth has developed and piloted structures to make routine clinical data from pediatric oncology clinics available for monitoring, benchmarking, and research in an interoperable, standardized, and

XSL•FO
RenderX

quality-controlled way. SwissPedCancer expects to include approximately 2800 patients diagnosed with cancer before the age of 20 years between 2017 and 2023. The pediatric oncology clinics' data and the manually validated ChCR data will be delivered separately to a secure national computing network for health-related data (Biomedical Information Technology). We will compare these two data sources to assess completeness (case ascertainment), accuracy (validity), and timeliness of cancer registration in the ChCR. We will evaluate data on diagnosis, treatments, underlying genetic disease, remission, relapse, and late effects. SwissPedCancer will provide a framework for optimizing standardized and uniform data transfers between pediatric oncology clinics and the ChCR and for other registries within Switzerland.

Results: The project was funded in September 2022 and received ethics exemption in October 2023. Data extraction from participating hospitals and the ChCR is expected to commence in January 2026. Study results are anticipated to be available in summer 2026.

Conclusions: SwissPedCancer aims to reduce manual workload while improving the completeness, accuracy, timeliness, and comparability of childhood cancer data in Switzerland. The project will contribute to a robust, interoperable, and sustainable national infrastructure supporting high-quality cancer registration, timely analyses, and evidence-based decision-making.

International Registered Report Identifier (IRRID): PRR1-10.2196/87007

(JMIR Res Protoc 2026;15:e87007) doi:10.2196/87007

KEYWORDS

cancer registration; cancer diagnosis; case ascertainment; chemotherapy; data quality; radiotherapy; Europe

Introduction

Cancer is one of the leading causes of death in childhood and adolescence [1]. In Switzerland, approximately 400 new cancer cases are diagnosed every year in children and adolescents aged younger than 20 years [2]. The incidence rate in children or adolescents in Switzerland is comparable to that of neighboring countries [3,4]. According to the Childhood Cancer Registry of Switzerland (ChCR), the incidence rate of childhood cancer is approximately 108.2 cases per million children annually [5]. Cancer registries play a crucial role in understanding and addressing the burden of cancer among children and adolescents [6]. They enable the collection of standardized data on cancer incidence, prevalence, spectrum of diagnoses, mortality, survival, treatments, and the course of the disease [7]. The Swiss Childhood Cancer Registry was founded in 1976 by the Swiss Pediatric Oncology Group [8]. Before 2020, registration of childhood cancer cases was done on a voluntary basis by all pediatric oncology clinics but already with a high completeness and data quality for children aged less than 15 years [8-14]. Adolescents aged between 16 and 20 years are often treated at adult oncology clinics where registration was less systematic in the registry. On January 1, 2020, the Swiss Federal Act on the Registration of Cancerous Diseases (Cancer Registration Act [CRA]; RS 818.33) and its related Cancer Registration Ordinance (CRO) came into force, mandating the systematic, uniform, and comprehensive registration of all cancer cases across the country. The act aims to improve understanding of cancer patterns, support research, and enhance the quality of diagnosis, treatment, and care. Health care providers are legally obliged to report relevant cancer data to cantonal registries for adults and to the national ChCR for children and adolescents [15]. The ChCR supports evidence-based public health planning, ultimately contributing to better care and prognosis for children with cancer [2,16].

As part of the growing national and international efforts to standardize cancer registration data [17-19] and automate dataflows [20,21], SwissPedCancer represents a key initiative to improve the quality and efficiency of childhood cancer registration in Switzerland. SwissPedCancer is a nested project of the pediatric national data stream (NDS) initiative, the Swiss Pediatric Personalized Research Network (SwissPedHealth), which supports personalized health research through a harmonized dataset across Switzerland [22]. In line with global efforts in countries such as Sweden, Denmark, Norway, the United Kingdom, and the United States [17,19,20,23,24]—where integrated, interoperable systems allow for timely cancer data capture—SwissPedCancer seeks to modernize data transfers from Clinical Data Warehouses (CDWs) to the ChCR. Currently, hospital staff manually enter clinical information from medical files into the clinical information system of CDWs of reporting hospitals. Hospital-based clinical research associates send medical files of patients with cancer to the ChCR, using methods such as secure email or online transfer tools. This process varies across pediatric oncology clinics: It occurs at the time of diagnosis, monthly, or in even larger time intervals, and it can involve individual or multiple patient files. At the ChCR, trained staff enter and code personal and cancer-related data manually. These procedures entail a high manual workload, lead to duplication of work (entering data twice at the CDWs and ChCR) with considerable delays from event to registration, and increase the risk of errors. SwissPedCancer aims to harmonize and automate these dataflows, minimizing delays in patient registration, reducing redundancy in data entry, and enhancing data security by limiting the number of individuals handling sensitive health information. Automation could save time and resources and improve the accuracy, timeliness, and completeness of registry data. As the ChCR is Switzerland's oldest national pediatric disease registry with high case completeness, the advancements achieved through SwissPedCancer could serve as a model for similar national and international registries.

SwissPedCancer aims to assess five quality indicators for cancer registration in the ChCR and improve dataflows between pediatric oncology clinics and the ChCR:

XSL•FO
RenderX

1. Assess adherence of cancer data to national and international standards: we will evaluate whether cancer data registered at the pediatric oncology clinics and at the ChCR align with the national and international guidelines for cancer registration and the standardization of practices.
2. Assess completeness of case ascertainment and case completeness at the ChCR from the nine pediatric clinics: we will evaluate if all eligible children with cancer, who should be registered according to the inclusion criteria of the ChCR, are registered within the ChCR dataset. We will evaluate how registered and unregistered cases differ and the reasons for variation:
   - by period, comparing the periods before and after the CRA came into force (January 1, 2020) when notification of cancer cases became compulsory for medical professionals and
   - by other potential predictors of registration, for example, hospital or medical subspecialty overseeing patients' treatments, region in Switzerland, cantonal cancer registry catchment area, cancer type, and age group.
3. Assess data accuracy (validity): we will evaluate the quality of registered data in the ChCR by comparing data from CDWs and ChCR. Data accuracy relies on the quality of data entry in the clinical information systems, the fidelity of corresponding data in CDWs, and the level of proficiency in abstracting, coding, and recording, both at the pediatric oncology clinics and the ChCR level.
4. Assess timeliness: we will evaluate timeliness, defined as the time interval between the date of cancer diagnosis (date of incidence) and the date of inclusion in the registry and official cancer statistics.
5. Optimize the dataflow between pediatric oncology clinics, the CDWs, and the ChCR: we will develop standard operating procedures to improve registration, set up dataflows between CDWs and the ChCR, and develop feedback loops between pediatric oncology clinics and the ChCR.

Methods

Ethical Considerations

SwissPedCancer is a national audit and quality improvement project. SwissPedCancer has been exempted from the requirement to obtain ethical approval by the Cantonal Ethical Committee Bern (Req-2023 - 01081); thus, individual informed consent was waived. All data used for this study were deidentified before analysis. Data access and handling complied with applicable Swiss data protection regulations, and appropriate technical and organizational measures were implemented to safeguard participant privacy and confidentiality (the Federal Act on the Registration of Cancerous Diseases in Switzerland, known as the CRA [15]. No financial or other compensation was provided to participants, as no direct participant contact or intervention occurred.

Inclusion Criteria

SwissPedCancer will include data from inpatients and outpatients younger than 20 years at the time of cancer diagnosis who were treated at a Swiss pediatric oncology clinic: Kantonsspital Aarau, Aarau; Universitäts-Kinderspital beider Basel, Basel, Ente Ospedaliero Cantonale, Bellinzona; Inselspital, Bern, Hôpitaux Universitaires de Genève, Geneva; Centre hospitalier universitaire vaudois, Lausanne; Luzerner Kantonsspital, Luzern; Ostschweizer Kinderspital, St. Gallen; or Kinderspital Zürich, Zurich. SwissPedCancer will initially include patients diagnosed between January 1, 2017, and December 31, 2023. We will include patients diagnosed with malignant neoplasms and low-grade central nervous system tumors, excluding skin basalioma, according to *International Statistical Classification of Diseases, Tenth Revision* (*ICD-10*), to align with the ChCR inclusion criteria and the CRO (Table 1). SwissPedCancer is a nested study of the NDS SwissPedHealth, which has been described in detail elsewhere [22].

Table . Codes of cancer diagnoses registered in the Childhood Cancer Registry.

ICD-10[a] codes	Notifiable cancers
C codes	
C00-C97	Malignant neoplasms
D codes	
D00-D03, D05-D09	Carcinoma in situ
D32, D33, D35(D35.2, D35.3, D35.4)	Benign neoplasms of the meninges, brain, and other parts of the central nervous system and benign neoplasms of endocrine glands of the head or brain (pituitary gland, epiphysis, ductus craniopharyngealis)
D37-D48	Neoplasms of uncertain or unknown behavior
D61	Unspecified aplastic anemia
D76	Other unspecified diseases involving lymphoreticular tissue and the reticulohistiocytic system

[a]*ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*.

XSL•FO
RenderX

Study Design

This is a retrospective comparative study that aims to compare information available in CDWs with the dataset of the national ChCR. We expect to include approximately 2800 patients from the nine pediatric oncology clinics for this period. This represents approximately 80% of the total childhood cancer cases in Switzerland [2,8,10-14,25,26].

Data Sources

Childhood Cancer Registry of Switzerland

Swiss pediatric oncologists started in 1976 to nationally register children diagnosed with leukemia, lymphoma, central nervous system tumors, malignant solid tumors, or Langerhans cell histiocytosis to facilitate inclusion in and comparison with international studies. Initially, only data from patients enrolled in clinical trials were documented in the ChCR. Since 1981, the ChCR also included non–study participants [27,28]. In 2020, the legal framework changed, and the registry is owned by the Federal Government. Under the new law, medical professionals are legally obliged to notify childhood cancer cases to the ChCR [15]. The ChCR registers all children and adolescents with a primary diagnosis before 20 years of age and who belong to the permanent resident population of Switzerland. The ChCR records longitudinal data about the course of disease, treatments, and follow-up information. Patients or legal representatives have the right to object to the registration (veto). The ChCR analyzes the data to monitor the incidence and survival of cancer in children and adolescents, to use them for national health reporting, and to make data available for researchers [2]. The ChCR is aligned with international and national standards and uses the Swiss coding handbook [2]. It collects the following data: identifying data, diagnosis coded according to *ICD-10*, International Classification of Diseases for Oncology (ICD-O-3), and International Classification of Childhood Cancer (ICCC-3) codes, staging and grading, treatments, underlying genetic disease, progression, remission, relapse, late effects, prognostic factors, and study participation in clinical trials (Table 2) [8]. Since 2009, the ChCR published biannual reports [2,8,10-14,25,26,29].

XSL•FO
RenderX

Table . Transfer of data from routine clinical information of pediatric oncology clinics for the project SwissPedCancer and overlapping variables of the Childhood Cancer Registry (ChCR) in Switzerland.

SwissPedCancer CDWs[a] data[b]	SwissPedCancer standard	Nonexhaustive list of overlapping ChCR variables	ChCR standard
Administrative case, care handling	—[c]	First treatment complex, treatment end date, accuracy of treatment end date, diagnostic institutions, treatment institution(s)	—
Administrative sex	SNOMED CT[d]	Sex	—
Age, birth	—	Date of birth, accuracy for date of birth	—
Data provider	UID[e]	Diagnostic institution(s), treatment institution, first treatment complex institution(s)	UID
Death	—	Vital status, date for vital status, accuracy for date of vital status	—
Drug administrative event	GTIN[f], SNOMED CT	ATC[g] code(s), standard drug combinations, start date of treatment, accuracy for start date of treatment, treatment end date, accuracy of treatment end date, study patient, type of study, study protocol, regimen, treatment institution, first treatment complex institution(s)	ATC
Drug prescription	SNOMED CT	Start date of treatment, accuracy for the start date of treatment, treatment end date, accuracy of treatment end date, study patient, type of study, study protocol, regimen, (first) treatment goal(s)	—
Diagnosis, billed diagnosis	*ICD-10*[h], ICD-O-3[i], SNOMED CT	*ICD* version, *ICD* code, inherited predispositions, type of medical condition, medical condition *ICD* version, medical condition *ICD* code, *ICD* version for causes of death, type of recurrence(s)/transformation(s), diagnostic institution(s)	*ICD-10*, ICD-O-3, ICCC-3[j]
Billed procedure	CHOP[k]	CHOP treatment code	CHOP
Health care encounter	—	Treatment end date, accuracy of treatment end date, diagnostic institution(s), treatment institution	—
Oncology diagnosis	ICD-O-3	ICD-O version, ICD Topography, ICD-O Morphology, ICD-O Behavior, ICD-O Histological grade, molecular or cytogenetic marker(s) tested, molecular or cytogenetic marker(s) test result, topography of metastases at diagnosis	ICD-O-3
CDW subject pseudo identifier	—	Study ID, NCID[l]	—
Nationality	ISO[m] code (ISO 3166)	Nationality	—
Consent, study participation	—	Study patient, type of study, study protocol, regimen	—

[a]CDW: clinical data warehouse.

[b]CDW data are based on Swiss Personalized Health Network concepts. Swiss Personalized Health Network concepts are generalizable building blocks, which can be used in different contexts. Each concept contains all information necessary to understand it, and concepts can be combined to composed concepts, which again can be combined to more complex compositions [30].

[c]Not available.

XSL•FO
RenderX

[d]SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms.
[e]UID: standardized business identification number.
[f]GTIN: Global Trade Item Number.
[g]ATC: Anatomical Therapeutic Chemical Classification.
[h]*ICD-10*: *International Statistical Classification of Diseases, Tenth Revision*.
[i]ICD-O-3: International Classification of Diseases for Oncology.
[j]ICCC: International Classification of Childhood Cancer.
[k]CHOP: diagnoses and the Swiss surgical classification.
[l]NCID: National Case Identifier.
[m]ISO: International Organization for Standardization.

Swiss Personalized Health Network: Data From CDWs

In Switzerland, the Swiss Personalized Health Network (SPHN) has built a shared standard infrastructure to make health-related data harmonized and interoperable, ease the work of data providers, and enhance the quality of routinely collected data [31]. SPHN aims to use and exchange health-related data between data providers and researchers in a FAIR (Findable, Accessible, Interoperable, and Reusable) manner [32]. SPHN supports the SwissPedHealth [22], a joint pediatric NDS initiative. The SwissPedHealth NDS will include data from the CDWs of all Swiss pediatric oncology clinics in a standardized interoperable format based on the Resource Description Framework [22,33]. SwissPedCancer is a nested project of the SwissPedHealth NDS initiative that aims to pilot this data infrastructure in the context of quality improvement (Table 2). This data structure incorporates existing ontologies to ensure standardized and interoperable data representation from all available CDW data, including data from pediatric and adult departments, inpatients, outpatients, and pathology reports [22].

Data Protection

Health-related personal data in this project are pseudonymized for the research team (Table 3). Only ChCR registry staff members with the relevant legal mandate have access to identifying information. To enable patient linkage, the CDWs generate a subject-specific pseudo identifier for each patient, which serves as the sole means of identification during analysis. ChCR employees will link the subject pseudo identifier with the ChCR Study ID using the Swiss personal social security insurance number (Old-Age and Survivors' Insurance [OASI] number) or, if not available, using names, sex, and date of birth. The process is multifaceted to satisfy all legal and data protection obligations.

Table . Summary of definitions of terms.

Term	Definition
ChCR[a] team	Staff of the ChCR including employees in cancer registration, medical coders, and statisticians
SwissPedCancer team	Researchers working for the SwissPedCancer project at the Institute of Social and Preventive Medicine, University of Bern
ChCR study ID	ID given to the participants of this specific study. This ID differs from the ID used within the ChCR (PidN[b]), so researchers cannot identify patients
CDW[c] subject pseudo identifier from pediatric oncology clinic	ID given to patients in the clinical data warehouse of a pediatric oncology clinic
PidN	ID given to patients at the ChCR

[a]ChCR: Childhood Cancer Registry.of Switzerland
[b]PidN: patient identification number.
[c]CDW: clinical data warehouse.

Data Governance

We used the project-specific Data Project Consortium Agreements from the governance structure of SwissPedHealth, which consists of a steering level with an overarching Steering Committee, an Infrastructure Executive Board and a Scientific Advisory Board, a management level comprising the Executive Office (Executive Board and Management Office), and an operational level with work packages covering the data integration and research programs [22]. The Data Project Consortium Agreements of SwissPedCancer focus on project-specific governance aspects, including project aims, sponsorship, project partners, data needs, and timelines. Principles of data governance and regulatory compliance respect primary ownership of routine data by the hospitals with their respective CDWs and research-only data by the respective research teams. Requests to access data are governed accordingly at the level of the SwissPedHealth Data Access Request Committee of the SwissPedHealth Infrastructure Executive Board [22].

Dataflows, Patient Linkage, and Data Storage

The dataflows within SwissPedCancer are divided into several phases (Figure 1).

XSL•FO
RenderX

Figure 1. Dataflows within SwissPedCancer (steps 1 - 5) and the ongoing process of cancer registration (steps I-III). *Merge the CDWs and ChCR datasets in the NP2 B-space of BioMedIT with the subject pseudoidentifier of the CDWs and the study ID of the ChCR: In the first stage, only ChCR employees have permission to access the SwissPedCancer B-space to erase those not eligible. Not eligible are non-Swiss residents, those who gave a veto against cancer registration, and those patients who initially had a suspicion of cancer but in whom cancer was excluded later. In the second stage, the SwissPedCancer research team can access the SwissPedCancer B-space. The dataset is now ready to assess quality assurance. **Administrative data: names, sex, date of birth, and Swiss personal social security insurance number (OASI number). In green color, ongoing process of registration at the Childhood Cancer Registry of Switzerland: data transfer from pediatric oncology clinics to ChCR. In blue color, SwissPedCancer processes: automated transfers from CDWs of hospitals to ChCR. BioMedIT: Biomedical Information Technology; CRA: Cancer Registration Act; CRO: Cancer Registration Ordinance.

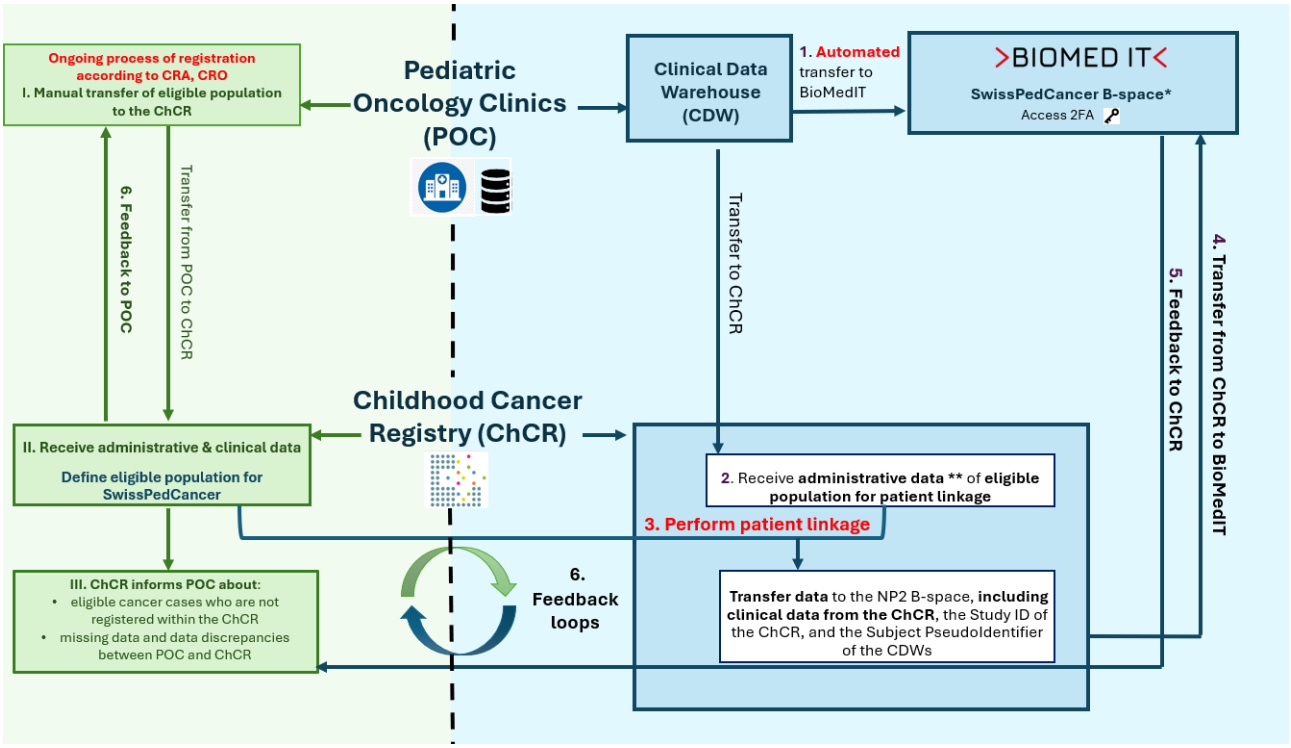

1. Automated transfers from CDWs to Biomedical Information Technology (BioMedIT): The CDWs of each pediatric oncology clinic will identify eligible patients (Table 1). *Selection of eligible population*: Pediatric oncology clinics will identify eligible patients in their CDW in line with the CRA and CRO [15] (Table 1). The CDWs will convert clinical data to SPHN standards and the Resource Description Framework format [34]. The CDWs will encrypt and transfer two datasets (Table 2) for the SwissPedCancer project. The *first dataset* is the full dataset of *clinical data*, which is transferred to the secure SwissPedHealth project space, so-called "NDS B-space" on the BioMedIT platform [35]. Clinical data include information on diagnosis (*ICD-10*, ICD-O-3), medical procedures, and death status. The NDS B-space serves as a data repository where data quality assessments and remediations are performed. After quality assessments by the central data manager of the NDS B-space, data are filtered and transferred to the SwissPedCancer project-specific B-space, so-called "workspace" [22].
2. Administrative data of eligible population: The second dataset is a minimal administrative dataset, which is transferred to the ChCR, including the subject-specific pseudo identifier, OASI number, name, sex, and date of birth for patient linkage.
3. Patient linkage within the ChCR: The ChCR team will perform a one-time patient linkage between the minimal administrative dataset from the CDWs and the dataset from the ChCR. After patient linkage, the ChCR will prepare a CSV file with the subject pseudo identifier of the pediatric oncology clinics, the ChCR Study ID, and all ChCR clinical data with overlapping information with CDW variables (Tables 2 and 3). ChCR staff will pseudonymize this dataset before transferring it to BioMedIT by deleting identifying information such as OASI numbers or names.
4. Data transfer from ChCR to BioMedIT: ChCR staff will transfer the pseudonymized dataset to BioMed IT. The data preparation in the BioMed IT workspace will contain two stages. Logging in to this secure workspace requires 2-factor authentication, and access to data is only possible via trusted networks. In the first stage, only ChCR employees have access to the workspace. ChCR employees will combine the pseudonymized ChCR dataset with the full dataset from the CDWs. According to the CRA and CRO, only ChCR staff are allowed to handle the vetoes against cancer registration. As the ChCR is not allowed to share the identity of patients who submitted a veto against cancer registration—only patients with cancer and their legal representatives hold this right—ChCR employees will exclude all cases at the time of diagnosis who are not eligible. In summary, not eligible are non-Swiss residents, patients who vetoed against cancer registration, and patients with an initial cancer suspicion that was later ruled out. In the second stage, the SwissPedCancer research team will

XSL•FO
RenderX

obtain access to the dataset from the ChCR research team after the exclusion of noneligible cases; the dataset is now ready for analysis.

5. Feedback to the ChCR, data evaluation: The SwissPedCancer team will evaluate data completeness and accuracy. The SwissPedCancer team will share their findings with the ChCR, the pediatric oncology clinics, and the CDWs on, for example, missing cases, missing data, or discrepancies, to improve data quality.
6. Feedback to pediatric oncology clinics to improve cancer registration according to CRA: The ChCR will contact the pediatric oncology clinics to request a defined dataset for eligible children who have not yet been reported to the registry.

Data Analysis Plan

SwissPedCancer will assess five quality indicators for cancer registration in the ChCR, proposed by the International Agency for Research on Cancer (IARC) and by the National Agency for Cancer Registration (NACR) [36,37]: (1) adherence of the data to national and international standards, (2) completeness of case ascertainment, (3) case completeness, (4) accuracy or validity of the data, and (5) timeliness of cancer registration.

Aim 1: Adherence of Registered Data to the National and International Standards

We will evaluate the extent to which data follow the national (NACR) standards and the international (IARC) recommendations [36,37]. We will assess how standard terminologies such as the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT), *ICD-10*, and ICD-O-3 are used across data sources, examining whether key variables are coded consistently. This evaluation will allow us to determine the interoperability of the datasets and their readiness for national and international comparisons.

Aim 2: Completeness of Case Ascertainment and Case Completeness

We will focus on evaluating the completeness of case ascertainment and case completeness in a multidimensional approach. To achieve this, we will compare patient data recorded during clinical care in pediatric oncology clinics to corresponding records in the ChCR. This includes diagnosis, presence of metastases, treatment details, genetic predisposition, remission status, relapse events, and late effects. Various computational methods [38] will identify missing cases and assess the extent and characteristics of incomplete data. For the assessment of completeness of case ascertainment and case completeness, we will consider the CDWs as the standard reference because they are the source of data.

We will conduct the following assessments:

1. Quantitative completeness of case ascertainment and case completeness to see whether all reportable cases are registered in the ChCR and if all variables for each case are captured in the ChCR.
2. Qualitative or semiquantitative completeness to reflect the thoroughness of each case record captured in the CDWs. We will assess the percentage and importance of missing variables or information of variables (eg, for diagnosis, are all codes available, eg, *ICD-10*, ICD-O-3, ICCC-3; Table 2). The completeness checks will include:
   - Variable presence: we will check the completeness of the extracted CDW data and the ChCR dataset, for example, if all variables listed in Table 2 are present and captured in both datasets. For example, apart from treatment received, the start and end date of treatment, the cumulative doses, and the location (body part) of radiotherapy are also captured in both datasets.
   - Data variable agreement: we will compare CDW data to ChCR data to evaluate if information is the same or compatible (eg, diagnosis, according to ICD-10 and ICD-O-3 codes, or treatment).
3. Temporal completeness: we will evaluate how completeness evolves over time, particularly comparing periods before and after the implementation of the CRA and CRO [39].

Aim 3: Data Accuracy (Validity)

The third objective is to assess the accuracy (validity) of selected variables in CDWs and the ChCR. For this aim, the ChCR will be considered the reference. The SwissPedCancer team will perform these analyses, and all findings will be validated by a medical coder from the ChCR.

Accuracy checks will include:

1. Face validity: Assessment of clinical plausibility of data entries. For instance, any diagnosis not typically found in the pediatric population as first primary tumors (eg, breast or prostate cancer) or implausible age at diagnosis (eg, over 20 years) will be flagged for review.
2. Qualitative assessment of disagreements:
   - Cataloging and characterization: Discrepancies between datasets will be documented and described, including the proportion of observations affected and the types of inconsistencies.
   - Prioritization: Disagreements will be categorized as high, medium, or low priority based on two criteria: the percentage of affected observations and the clinical utility of the variable for the cancer registry. Variables with high disagreement and high relevance—such as diagnosis and treatment—will be prioritized for investigation.
3. Quantification of agreement: We will compute statistical measures such as Cohen κ, sensitivity, specificity, positive predictive value, and negative predictive value to evaluate the level of agreement between the two datasets. Cohen κ will be applied to assess concordance in the registration of high-priority variables, such as diagnostic coding.

Aim 4: Timeliness

We will evaluate the timeliness of childhood cancer registration, defined as the time interval between the date of cancer diagnosis (date of incidence) and the date of inclusion in official cancer statistics. We will assess data from 2016 to 2023, covering four years before and after the CRA. To provide a granular understanding of this timeline, we will decompose the overall

XSL•FO
RenderX

interval into the following three distinct phases, as outlined in previous literature:

1. First registration: This period spans from the date when the cancer diagnosis is made and known (date of incidence) to the date when the case is first notified to and entered in the ChCR.
2. Final registration: The second interval begins with the initial recording of the case in the ChCR database and ends with the completion of internal quality assurance procedures. These procedures include the medical coding of the case by one medical coder and an independent cross-check by a second coder to ensure accuracy and consistency.
3. Publication phase: The final phase covers the time from the submission of aggregated registry data to the NACR until the publication of the data in the first statistical report. This interval reflects the broader dissemination of cancer statistics for public health monitoring and research use.

Timeliness will be reported as median time intervals with IQRs for each phase. Additionally, subgroup analyses will be conducted to explore differences in timeliness by age group, ICCC-3 main diagnostic group, and type of notifying institution, for example, diagnostic centers and pediatric oncology clinics.

Aim 5: Dataflow Processes

We will focus on improving the dataflow processes between CDWs and the ChCR, building upon the findings from aim 1 (comparability) and aim 2 (completeness). On the basis of identified discrepancies, missing data, and inconsistencies, we will propose targeted strategies to enhance efficiency, consistency, and automation of data transfers. These strategies include:

1. Harmonization and standardization of upstream data collection: For variables that exhibit low data quality or inconsistent registration across CDWs, we will recommend the implementation of more structured data entry formats. Harmonizing the standardized data extraction across Swiss pediatric oncology clinics and the ChCR will reduce variability and improve overall data quality.
2. Direct data integration: To minimize manual data entry, which is both labor-intensive and error prone, we will explore the feasibility of direct dataflows from CDWs to the ChCR. Automating this process would significantly reduce the burden on clinical and registry staff while improving data timeliness and accuracy.

In parallel, we will develop tools that complement the IARC and ENCR (European Network for Cancer Registries) check tools to support automated detection of data errors within the ChCR. Here, a custom R script or machine learning model will be designed to identify data errors systematically, defined as either missing variables or incomplete or inconsistent values for captured variables. These automated tools will flag problematic entries that can be addressed during subsequent data updates from pediatric oncology clinics. By incorporating automated quality checks into the data pipeline, the ChCR can proactively detect and resolve data integrity issues, leading to higher-quality datasets for analysis and reporting.

Missing Data

Missing data in the ChCR and the pediatric oncology clinics' clinical data will be described by variable, hospital, and incidence year. Subgroup analyses (eg, by cancer type, age group, or hospital) will be considered exploratory. To minimize overinterpretation, the number of subgroups will be limited, and findings will be interpreted using effect estimates and CIs rather than relying solely on *P*-values.

Results

The project was funded in September 2022 and received ethics exemption in October 2023. Data extraction from participating hospitals and the ChCR is expected to commence in January 2026. Study results are anticipated to be available in Summer 2026.

Discussion

Main Findings

Completeness of registration of cancer cases—referred to as *case ascertainment*—is essential for drawing reliable conclusions about the development, progression, and treatment efficacy of the different cancer types. In Switzerland, case ascertainment in the ChCR is greater than 95% complete for children aged 0 to 15 years at diagnosis, as these patients are typically treated in specialized pediatric oncology clinics where cancer cases are routinely reported to the ChCR [40,41]. However, case ascertainment is lower for adolescents aged 16 to 19 years, whose care is fragmented across several pediatric and adult hospitals [40,41]. In 2010, 22% of individuals aged younger than 20 years and recorded in regional cancer registries were not registered in the ChCR. These missing individuals were mainly infants or adolescents [42]. Infants may die before being admitted to pediatric oncology clinics, making registration only possible via linkages with death certificate notifications of the Swiss Federal Statistical Office, whereas adolescents are often treated in adult oncology clinics with different registration procedures. Achieving complete case ascertainment across all age groups is critical to ensure accurate monitoring of pediatric cancer incidence, prevalence, survival, and quality of care. Comprehensive and reliable cancer registration enables timely cancer monitoring and supports public health planning [43]. Improvements in cancer registration—such as those proposed by SwissPedCancer—could enhance the ChCR's data quality through the efficient and cost-effective transfer of CDW data and by using standardized and uniform procedures across pediatric oncology clinics.

Currently, there is no automated extraction of data from CDWs to the ChCR. Validation feedback loops are manual and cumbersome, resulting in delays in registering new patients and collecting follow-up data. Furthermore, it is redundant work, as the same data must be entered both in the primary record systems of the pediatric oncology clinics (from where it is automatically transferred into local CDWs) and again into the registry. The ChCR uses internationally standardized formats to code medical oncology data, which are often not used in primary record systems and associated CDWs, which further

XSL•FO
RenderX

increases manual workload during data extraction and entry. The automated dataflows proposed by SwissPedCancer aim to reduce this manual effort while improving the completeness, accuracy, timeliness, and comparability of childhood cancer data in Switzerland. Ultimately, these enhancements aim to create a more robust, interoperable, and sustainable infrastructure for childhood cancer data in Switzerland—supporting accurate registration, timely analysis, and evidence-based decision-making.

Strengths and Limitations

SwissPedCancer will include all nine pediatric oncology clinics in Switzerland, ensuring full national coverage. The comparison between CDW and ChCR data is a novel approach to assessing the completeness and quality of ChCR data from the nine pediatric oncology clinics. SwissPedCancer is the first project to describe the changes in quality indicators of cancer registration before and after the CRA came into force in 2020, which could have influenced completeness and timeliness of cancer registration.

One limitation of SwissPedCancer is that the COVID-19 pandemic started in 2020 which may have resulted in delays and interruptions to childhood cancer diagnosis and treatment [44]. However, in Switzerland, currently available evidence suggests that the effect of the COVID-19 pandemic on childhood cancer diagnosis and treatment was minimal, as in other high-income countries with healthcare systems that adapted quickly to maintain essential pediatric oncology services and avoid important delays in care [45]. Finally, as childhood cancer is a rare disease, which increases the chance of reidentification of the patients based solely on medical data, the procedures proposed in SwissPedCancer are designed to protect patients' data by avoiding double data entry and minimizing the risk of data breach, which is a strength of this study.

Outlook

SwissPedCancer presents a unique opportunity to assess the completeness and quality of cancer registration data in the ChCR compared to an external reference—CDWs from pediatric oncology clinics. It provides the opportunity to use existing data sources, which are rich in clinical details, for quality assurance. SwissPedCancer enables evaluation not only of case ascertainment and case completeness in the ChCR but also of data validity in pediatric oncology clinics and the ChCR.

Monitoring cancer at the population level has traditionally been labor-intensive, relying on manual record review and data entry by certified registrars [46]. Advances in information technology have expanded the availability of administrative databases, which are increasingly important for epidemiological research, quality assessment, and monitoring of disease outcomes [47]. CDWs, already in use at pediatric oncology clinics, have the potential to reduce the manual workload for registry staff and medical coders. While discrepancies may exist between pediatric oncology clinics and registry data—due to different objectives and methods of data collection—a shared objective remains: the accurate and complete data collection of cancer cases from the nine pediatric oncology clinics [48]. Data linkage and comparison between pediatric oncology clinics and the ChCR can generate new insights, broadening the potential impact of either data source when used alone. Finally, while the findings of this study are specific to childhood cancer, they are expected to inform the scalability of the approach toward other noncommunicable and communicable medical conditions, which may benefit from surveillance and clinical information system–enabled registries. By minimizing data handling, SwissPedCancer strengthens data privacy and supports efficient national data aggregation. The project's standardized approach could serve as a model for other pediatric networks and broader national and international registry initiatives.

Assess adherence of cancer data to national and international standards: we will evaluate whether cancer data registered at the pediatric oncology clinics and at the ChCR align with the national and international guidelines for cancer registration and the standardization of practices.

Acknowledgments

The Swiss Pediatric Personalized Research Network (SwissPedCancer) is supported by the Swiss Childhood Cancer Registry. The authors would like to thank the following team members of the Swiss Childhood Cancer Registry: Alex Junghans and Eleftheria Michalopoulou. Patients or the public were involved in the design, conduct, reporting, or dissemination plans of the national data stream SwissPedHealth of which SwissPedCancer is a part of [22]. The proposal for this study was peer-reviewed by Swiss Personalized Health Network-Personalized Health and Related Technologies Review Committee (Switzerland).

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT (Generative AI Delegation Taxonomy; 2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing. The GAI tool used was ChatGPT 5.2. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. This declaration was submitted by YS.

The SwissPedHealth consortium includes the following:

Andrea Agostini, Department of Computer Science, Institute for Machine Learning, ETH Zurich, Zurich, Switzerland;

Anita Rauch, Institute of Medical Genetics, University of Zurich, Zurich, Switzerland;

Anna Hartung, Inselspital, Bern University Hospital, University of Bern, Switzerland;

XSL•FO
RenderX

Audrey van Drogen, PHRT Swiss Multi-Omics Centre (SMOC), ETH Zurich, Zurich, Switzerland & Institute of Translational Medicine (ITM), Department of Health Sciences and Technology (D-HEST), ETH Zurich, Zurich, Switzerland;

Aurélie Martin Necker, Patient and Family Advisory Committee, SwissPedHealth,

Ben D Spycher, Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland;

Christian Kahlert, Ostschweizer Kinderspital, St Gallen, Switzerland;

Christopher B Forrest, Center for Applied Clinical Research, Children's Hospital of Philadelphia, Philadelphia, USA;

Claudia E Kuehni, Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland & Division of Paediatric Respiratory Medicine and Allergology, Children's University Hospital, Inselspital, University of Bern, Bern, Switzerland;

Cornelia Hagman, Department of Intensive Care and Neonatology and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland;

D Sean Froese, Division of Metabolism and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland,

Daphné Chopard, Department of Computer Science, Institute for Machine Learning, ETH Zurich, Zurich, Switzerland & Department of Intensive Care and Neonatology and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland;

Dylan Lawless, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland & Department of Intensive Care and Neonatology and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland;

Effy Vayena, Department of Health Sciences and Technology, Institute of Translational Medicine, ETH Zurich, Zurich, Switzerland;

Eirini I Petrou, Department of Health Sciences and Technology, Institute of Translational Medicine, ETH Zurich, Zurich, Switzerland; Emanuele Palumbo, Department of Computer Science, Institute for Machine Learning, ETH Zurich, Zurich, Switzerland;

Eric Giannoni, Clinic of Neonatology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland;

Fabiën N Belle, Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland;

Ioannis Xenarios, PHRT Swiss Multi-Omics Centre (SMOC), EPFL, Lausanne, Switzerland & Department of Computational Biology, University of Lausanne, Lausanne 1015, Switzerland & Health 2030 Genome Center, Foundation Campus Biotech, Geneva, Switzerland;

Jacques Fellay, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland & Biomedical Data Science Center, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland;

Jana Pachlopnik Schmid, Division of Immunology and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland;

Julia A Bielicki, Paediatric Research Center, University l Children's Hospital Basel (UKBB), University of Basel, Basel, Switzerland & Centre for Neonatal and Paediatric Infection, St George's, University of London, London, UK;

Julia E Vogt, Department of Computer Science, Institute for Machine Learning, ETH Zurich, Zurich, Switzerland;

Kathrin Hofmann, Patient and Family Advisory Committee, SwissPedHealth,

Katrin Männik, PHRT Swiss Multi-Omics Centre (SMOC), EPFL, Lausanne, Switzerland & Center for Integrative Genomics, University of Lausanne, Lausanne, Switzerland & Health 2030 Genome Center, Foundation Campus Biotech, Geneva, Switzerland;

Keith Harshman, PHRT Swiss Multi-Omics Centre (SMOC), EPFL, Lausanne, Switzerland & Health 2030 Genome Center, Foundation Campus Biotech, Geneva, Switzerland;

Kelly Ormond; Department of Health Sciences and Technology, Institute of Translational Medicine, ETH Zurich, Zurich, Switzerland & Department of Genetics, Stanford University School of Medicine, Stanford, California, USA;

Klara Posfay-Barbe, Hôpitaux Universitaires de Genève, Geneva, Switzerland;

Léa Ho Dac, Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland;

Lorenz M Leuenberger, Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland;

XSL•FO
RenderX

Luregn J Schlapbach, Department of Intensive Care and Neonatology, and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland & Child Health Research Centre, The University of Queensland, Brisbane, Australia;

Manon Jaboyedoff, Pediatric Infectious Diseases and Vaccinology Unit, Service of Pediatrics, Department Mother-Woman-Child, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland;

Mariam Ait Oumelloul, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland;

Martin Stocker, Luzerner Kantonsspital, Luzern, Switzerland;

Matthias R Baumgartner, Division of Metabolism and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland;

Nicola Zamboni, PHRT Swiss Multi-Omics Centre (SMOC), ETH Zurich, Zurich, & Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland;

Nicole Goebel, Research and Analyses Services, Digitalisation & ICT division, University Hospital Basel, Basel, Switzerland;

Patrick G A Pedrioli, PHRT Swiss Multi-Omics Centre (SMOC), ETH Zurich, Zurich, Switzerland & Institute of Translational Medicine (ITM), Department of Health Sciences and Technology (D-HEST), ETH Zurich, Zurich, Switzerland & Swiss Institute of Bioinformatics, Lausanne, Switzerland & Department of Biology, Institute of Molecular Systems Biology, Swiss Federal Institute of Technology/ETH Zürich, Zurich, Switzerland;

Philipp Latzin, Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland;

Rebeca Mozun, Department of Intensive Care and Neonatology, and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland;

Roger Lauener, Ostschweizer Kinderspital, St. Gallen, Switzerland;

Sandra Goetze, PHRT Swiss Multi-Omics Centre (SMOC), ETH Zurich, Zurich, Switzerland & Institute of Translational Medicine (ITM), Department of Health Sciences and Technology (D-HEST), ETH Zurich, Zurich, Switzerland;

Seraina Prader, Division of Immunology and Children's Research Centre, University Children's Hospital Zurich;

Simon Boutry, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland;

Sven Schulzke, Department of Neonatology, University Children's Hospital Basel, University of Basel, Basel, Switzerland;

Tatjana Welzel, Paediatric Research Center, University Children's Hospital Basel, University of Basel, Basel, Switzerland;

Thomas M Sutter, Department of Computer Science, Institute for Machine Learning, ETH Zurich, Zurich, Switzerland;

Varvara Dimopoulou, Clinic of Neonatology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland;

Vito RT Zanotelli, Division of Metabolism and Children's Research Centre, University Children's Hospital Zurich, University of Zurich, Zurich, Switzerland;

Xeni Deligianni, Research and Analyses Services, Digitalization & ICT division, University Hospital Basel, Basel, Switzerland;

Xenia Bovermann, Division of Pediatric Respiratory Medicine and Allergology, Department of Pediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland; and

Yara Shoman, Institute of Social and Preventive Medicine (ISPM), University of Bern, Bern, Switzerland

Funding

This work was supported by the grant NDS-2021-911 (SwissPedHealth) from the Swiss Personalized Health Network and the Strategic Focal Area "Personalized Health and Related Technologies" of the ETH Domain (Swiss Federal Institutes of Technology). The grant was peer reviewed during a meeting, and the meeting minutes are provided as feedback from reviewers (Peer Review Report 1). LJS received support from the NOMIS Foundation and the Thomas & Doris Ammann Foundation. The funder had no involvement in the data collection, analysis, interpretation, or the writing of the manuscript. However, the Swiss Personalized Health Network uses concepts that incorporate existing ontologies and terminologies [30], which were used in SwissPedHealth [22].

Authors' Contributions

Conceptualization: FNB, CEK, LJS, JAB

XSL•FO
RenderX

Funding acquisition: FNB, CEK, LJS, JAB
Methodology: YS, FNB, CEK, LJS, JAB, BDS
Project administration: FNB, CEK
Resources: SDV, GS, FNB, CEK, UMK
Supervision: FNB, CEK
Validation: All authors and the SwissPedHealth consortium contributed to the validation of the study
Writing – original draft: All authors contributed to writing the original draft
Writing – review & editing: All authors contributed to preparation, reviewing and editing the final draft

Conflicts of Interest

NW reports advisory board and travel reimbursement from Swedish Orphan Biovitrum AB, advisory board and travel reimbursement from NovoNordisk, and advisory board reimbursement from Novartis. None of these activities were in relation to the content of this manuscript. All other authors declare no competing interests.

Checklist 1
SPIRIT checklist.
[DOCX File, 34 KB - resprot_v15i1e87007_app1.docx]

Peer Review Report 1
Peer review report by the Swiss Personalized Health Network (SPHN)-Personalized Health and Related Technologies (PHRT) Review Committee (Switzerland).
[PDF File, 404 KB - resprot_v15i1e87007_app2.pdf]

References

1. Childhood cancer. World Health Organization. 2025 Feb 4. URL: https://www.who.int/news-room/fact-sheets/detail/cancer-in-children [accessed 2026-02-04]
2. Annual report of the Swiss Childhood Cancer Registry 2022 [Report in German]. : Swiss Childhood Cancer Registry; 2022 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2023/11/Jahresbericht-2022_KiKR_oeffentlich-1.pdf [accessed 2026-02-04]
3. Lorez M, Nanieva R, Arndt V, Rohrmann S. Benign and malignant primary brain tumours in the Swiss population (2010-2014). Schweizer Krebsbulletin 2018;38(2):188-196 [FREE Full text]
4. Swiss cancer report 2021: rise in number of cancer patients coincides with decline in mortality. : Federal Statistical Office; 2021 URL: https://www.bfs.admin.ch/asset/en/19204988 [accessed 2026-02-04]
5. Statistics and reports. Childhood Cancer Registry of Switzerland. 2025. URL: https://www.childhoodcancerregistry.ch/statistics-and-reports [accessed 2026-02-04]
6. The Lancet. Cancer registries: the bedrock of global cancer care. Lancet 2025 Feb 1;405(10476):353. [doi: 10.1016/S0140-6736(25)00189-8] [Medline: 39892893]
7. Concept for analysing cancer registry data. : National Agency for Cancer Registration; 2024 URL: https://onec.ch/download/upload/downloads/Berichte%20und%20Statistiken/Auswertungskonzept/2024_C_1.PDF [accessed 2026-02-04]
8. Annual report of the Swiss Childhood Cancer Registry 2017/2018. : Swiss Childhood Cancer Registry; 2019 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2019/09/FINAL-Annual-Report_2017_2018_with-CC-license.pdf [accessed 2026-02-04]
9. International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation I: history and theoretical development. Am J Epidemiol 1995 Nov 15;142(10):1047-1058. [doi: 10.1093/oxfordjournals.aje.a117559] [Medline: 7485050]
10. Annual report of the Swiss Childhood Cancer Registry 2011-2012. : Swiss Childhood Cancer Registry; 2013 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2019/09/FINAL-Annual-Report_2011_2012_with-CC-license.pdf [accessed 2026-02-04]
11. Annual report of the Swiss Childhood Cancer Registry 2013-2014. : Swiss Childhood Cancer Registry; 2015 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2019/09/FINAL-Annual-Report_2013_2014_with-CC-license.pdf [accessed 2026-02-04]
12. Annual report of the Swiss Childhood Cancer Registry 2014-2015. : Swiss Childhood Cancer Registry; 2016 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2019/09/FINAL-Annual-Report_2014_2015_with-CC-license.pdf [accessed 2026-02-04]
13. Annual report of the Swiss Childhood Cancer Registry 2015-2016. : Swiss Childhood Cancer Registry Team; 2017 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2019/09/FINAL-Annual-Report_2015_2016_with-CC-license.pdf [accessed 2026-02-04]

XSL•FO
RenderX

14. Annual report of the Swiss Childhood Cancer Registry 2020 [Report in German]. : Swiss Childhood Cancer Registry; 2020 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2023/11/210806_Jahresbericht_KiKR_oeffentlich_FINAL.pdf [accessed 2026-02-04]
15. Cancer Registration Act (CRA). Fedlex Die Publikationsplattform des Bundesrechts. 2020. URL: https://www.fedlex.admin.ch/eli/cc/2018/289/de [accessed 2025-02-25]
16. Steliarova-Foucher E, Colombet M, Ries LAG, et al. International incidence of childhood cancer, 2001-10: a population-based registry study. Lancet Oncol 2017 Jun;18(6):719-731. [doi: 10.1016/S1470-2045(17)30186-9] [Medline: 28410997]
17. Gjerstorff ML. The Danish Cancer Registry. Scand J Public Health 2011 Jul;39(7 Suppl):42-45. [doi: 10.1177/1403494810393562] [Medline: 21775350]
18. Larsen IK, Småstuen M, Johannesen TB, et al. Data quality at the Cancer Registry of Norway: an overview of comparability, completeness, validity and timeliness. Eur J Cancer 2009 May;45(7):1218-1231. [doi: 10.1016/j.ejca.2008.10.037] [Medline: 19091545]
19. Sigurdardottir LG, Jonasson JG, Stefansdottir S, et al. Data quality at the Icelandic Cancer Registry: comparability, validity, timeliness and completeness. Acta Oncol 2012 Sep;51(7):880-889. [doi: 10.3109/0284186X.2012.698751] [Medline: 22974093]
20. Artificial intelligence and natural language processing in the cancer registry field. North American Association of Central Cancer Registries (NAACCR). URL: https://education.naaccr.org/products/artificial-intelligence-and-natural-language-processing-in-the-cancer-registry-field [accessed 2026-02-04]
21. Henson KE, Elliss-Brookes L, Coupland VH, et al. Data resource profile: National Cancer Registration Dataset in England. Int J Epidemiol 2020 Feb 1;49(1):16-16h. [doi: 10.1093/ije/dyz076] [Medline: 31120104]
22. Mozun R, Belle FN, Agostini A, et al. Paediatric Personalized Research Network Switzerland (SwissPedHealth): a joint paediatric national data stream. BMJ Open 2024 Dec 26;14(12):e091884. [doi: 10.1136/bmjopen-2024-091884] [Medline: 39725440]
23. Ekström Smedby K, Eloranta S, Wästerlid T, et al. The National Swedish Lymphoma Register - a systematic validation of data quality. Acta Oncol 2024 Jul 10;63:563-572. [doi: 10.2340/1651-226X.2024.40431] [Medline: 38988133]
24. Kerr SM, Campbell A, Marten J, et al. Electronic health record and genome-wide genetic data in Generation Scotland participants. Wellcome Open Res 2017;2:85. [doi: 10.12688/wellcomeopenres.12600.1] [Medline: 29062915]
25. Annual report of the Swiss Childhood Cancer Registry 2021 [Report in German]. : Childhood Cancer Registry in Switzerland; 2021 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2023/11/20220711_Jahresbericht2021_KiKR_oeffentlich.pdf [accessed 2026-02-04]
26. Annual report of the Swiss Childhood Cancer Registry 2023 [Report in German]. : Childhood Cancer Registry in Switzerland; 2023 URL: https://www.kinderkrebsregister.ch/wp-content/uploads/sites/2/2025/01/KiKR_Jahresbericht-2023_oeffentlich.pdf [accessed 2026-02-04]
27. Michel G, von der Weid NX, Zwahlen M, et al. The Swiss Childhood Cancer Registry: rationale, organisation and results for the years 2001-2005. Swiss Med Wkly 2007 Sep 8;137(35-36):502-509. [doi: 10.4414/smw.2007.11875] [Medline: 17990137]
28. Michel G, von der Weid NX, Zwahlen M, et al. Incidence of childhood cancer in Switzerland: the Swiss Childhood Cancer Registry. Pediatr Blood Cancer 2008 Jan;50(1):46-51. [doi: 10.1002/pbc.21129] [Medline: 17226849]
29. Childhood Cancer Registry. URL: https://www.childhoodcancerregistry.ch [accessed 2025-05-12]
30. The SPHN semantic interoperability framework. SPHIN - Swiss Personlized Health Network. 2024. URL: https://sphn.ch/network/data-coordination-center/the-sphn-semantic-interoperability-framework [accessed 2026-02-04]
31. Lawrence AK, Selter L, Frey U. SPHN - the Swiss Personalized Health Network Initiative. Stud Health Technol Inform 2020 Jun 16;270:1156-1160. [doi: 10.3233/SHTI200344] [Medline: 32570562]
32. Touré V, Krauss P, Gnodtke K, et al. FAIRification of health-related data using semantic web technologies in the Swiss Personalized Health Network. Sci Data 2023 Mar 10;10(1):127. [doi: 10.1038/s41597-023-02028-y] [Medline: 36899064]
33. Resource description framework. Semantic Web Standards. URL: https://www.w3.org/RDF/ [accessed 2025-03-05]
34. van der Horst E, Unni D, Kopmels F, et al. Bridging clinical and genomic knowledge: an extension of the SPHN RDF schema for seamless integration and fairification of omics data. : SWAT4HCLS 2024; 2024 URL: https://dblp.uni-trier.de/rec/conf/swat4ls/HorstUKATFCCO24.html [accessed 2026-02-17] [doi: 10.20944/preprints202312.0373.v1]
35. BioMedIT. URL: https://www.biomedit.ch/ [accessed 2026-02-10]
36. Lorez M, Wagner U. New data quality concept within the framework of the federal law on the registration of cancer (cancer registration act, CRA; SR 813.33). : National Agency for Cancer Registration (NACR); 2022 URL: https://www.nacr.ch/upload/downloads/Berichte%20und%20Statistiken/Datenqualit%C3%A4t/nacr-data-quality-concept-v1.pdf [accessed 2026-02-04]
37. Parkin DM, Chen VW, Ferlay J, et al. Comparability and quality control in cancer registration: IARC technical report No 19. : International Agency for Research on Cancer, World Health Organization; 1994 URL: https://publications.iarc.who.int/Book-And-Report-Series/Iarc-Technical-Publications/Comparability-And-Quality-Control-In-Cancer-Registration-1994 [accessed 2026-02-04]

XSL•FO
RenderX

38. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc 2013 Jan 1;20(1):144-151. [doi: 10.1136/amiajnl-2011-000681] [Medline: 22733976]
39. Bray F, Z A, Cueva P. Quality control at the population-based cancer registry: IARC technical report, No. 43. In: Planning and Developing Population-Based Cancer Registration in Low- or Middle-Income Settings: International Agency for Research on Cancer; 2014. URL: https://www.ncbi.nlm.nih.gov/books/NBK566960/ [accessed 2026-02-04]
40. Adam M, von der Weid N, Michel G, et al. Access to specialized pediatric cancer care in Switzerland. Pediatr Blood Cancer 2010 May;54(5):721-727. [doi: 10.1002/pbc.22426] [Medline: 20108340]
41. Le cancer en suisse, rapport 2021 état des lieux et évolutions [Report in French]. : Office fédéral de la statistique; 2021 URL: https://oncosuisse.ch/wp-content/uploads/Krebsbericht-2021-fr-1.pdf [accessed 2026-02-04]
42. Zanetti R, Schmidtmann I, Sacchetto L, et al. Completeness and timeliness: cancer registries could/should improve their performance. Eur J Cancer 2015 Jun;51(9):1091-1098. [doi: 10.1016/j.ejca.2013.11.040] [Medline: 24393522]
43. White MC, Babcock F, Hayes NS, et al. The history and use of cancer registry data by public health cancer control programs in the United States. Cancer 2017 Dec 15;123(Suppl 24):4969-4976. [doi: 10.1002/cncr.30905] [Medline: 29205307]
44. Majeed A, Wright T, Guo B, Arora RS, Lam CG, Martiniuk AL. The global impact of COVID-19 on childhood cancer outcomes and care delivery - a systematic review. Front Oncol 2022;12:869752. [doi: 10.3389/fonc.2022.869752] [Medline: 35463381]
45. Graetz D, Agulnik A, Ranadive R, et al. Global effect of the COVID-19 pandemic on paediatric cancer care: a cross-sectional study. Lancet Child Adolesc Health 2021 May;5(5):332-340. [doi: 10.1016/S2352-4642(21)00031-6] [Medline: 33675698]
46. Rollison DE, M Levin G, Warner JL, et al. Current and emerging informatics initiatives impactful to cancer registries. J Registry Manag 2022;49(4):153-160. [Medline: 37260815]
47. Nørgaard M, Skriver MV, Gregersen H, Pedersen G, Schønheyder HC, Sørensen HT. The data quality of haematological malignancy ICD-10 diagnoses in a population-based hospital discharge registry. Eur J Cancer Prev 2005 Jun;14(3):201-206. [doi: 10.1097/00008469-200506000-00002] [Medline: 15901987]
48. Field K, Kosmider S, Johns J, et al. Linking data from hospital and cancer registry databases: should this be standard practice? Intern Med J 2010 Aug;40(8):566-573. [doi: 10.1111/j.1445-5994.2009.01984.x] [Medline: 19460066]

Abbreviations

BioMedIT: Biomedical Information Technology
CDW: clinical data warehouse
ChCR: Childhood Cancer Registry of Switzerland
CRA: Cancer Registration Act
CRO : Cancer Registration Ordinance
FAIR: Findable, Accessible, Interoperable, and Reusable
IARC: International Agency for Research on Cancer
ICCC-3 : International Classification of Childhood Cancer–Third Edition
ICD-10 : International Statistical Classification of Diseases, Tenth Revision
ICD-O-3: International Classification of Diseases for Oncology
NACR : National Agency for Cancer Registration
NDS : national data stream
PHRT : Personalized Health and Related Technologies
SNOMED CT: Systematized Nomenclature of Medicine Clinical Terms
SPHN : Swiss Personalized Health Network
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SwissPedHealth: Swiss Pediatric Personalized Research Network

Edited by J Sarvestan; submitted 04.Nov.2025; this is a non–peer-reviewed article;accepted 07.Jan.2026; published 10.Mar.2026.

Please cite as:
Shoman Y, Leuenberger L, Sommer G, Bielicki JA, Brazzola P, della Valle S, Diezi M, Drozdov D, Gumy-Pause F, Guerreiro Stücklin A, Kuehnel UM, Scheinemann K, Schindera C, Schilling F, Waespe N, Spycher BD, Schlapbach LJ, Kuehni CE, Belle FN, SwissPedHealth Consortium
Improving Registration and Dataflows Between Pediatric Oncology Clinics and the Childhood Cancer Registry of Switzerland: Protocol for SwissPedCancer Quality Assurance Study
JMIR Res Protoc 2026;15:e87007
URL: https://www.researchprotocols.org/2026/1/e87007
doi:10.2196/87007

XSL•FO
RenderX

© Yara Shoman, Lorenz Leuenberger, Grit Sommer, Julia Anna Bielicki, Pierluigi Brazzola, Sophia della Valle, Manuel Diezi, Daniel Drozdov, Fabienne Gumy-Pause, Ana Guerreiro Stücklin, Ursula M Kuehnel, Katrin Scheinemann, Christina Schindera, Freimut Schilling, Nicolas Waespe, Ben D Spycher, Luregn J Schlapbach, Claudia E Kuehni, Fabiën N Belle, SwissPedHealth Consortium. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cardiovascular Health and Financial Hardship: Protocol for a Qualitative Citizen Science Study

Dagmar Niewold[1]; Evy C E van Gestel[2,3], MSc; Latifa Abidi[3], PhD; Klaske Tiemstra[3], MSc; Myrte Thoolen[2], PhD; Gera E Nagelhout[2,3], PhD

[1]Independent scholar, Amersfoort, Utrecht, The Netherlands

[2]Centre of Expertise Perspective in Health, Avans University of Applied Sciences, Lovensdijkstraat 63, Breda, North Brabant, The Netherlands

[3]Department of Health Promotion (CAPHRI), Maastricht University, Maastricht, Limburg, The Netherlands

Corresponding Author:
Gera E Nagelhout, PhD
Centre of Expertise Perspective in Health, Avans University of Applied Sciences, Lovensdijkstraat 63, Breda, North Brabant, The Netherlands

Abstract

Background: Cardiovascular disease (CVD) is the leading cause of death worldwide. Individuals with lower income or experiencing financial hardship face a significantly higher risk of developing CVD. However, there is a lack of in-depth insight into their experiences with CVD, and specific attention to women is essential.

Objective: The In a Heartbeat study aims to understand the relationship between CVD and financial hardship and enable earlier recognition and prevention of CVD among both women and men. In this study protocol, we describe our citizen science study, in which we unravel the mechanisms and contexts through which financial problems lead to the development and late recognition of CVD.

Methods: The main data for this study are collected by citizen scientists through qualitative lifeline interviews. All citizen scientists have experience with financial hardship, and some also have experience with CVD. We hold weekly project team meetings with citizen scientists and professional scientists in which we use methods such as photo elicitation, critical-creative hermeneutic analysis, design thinking, a dynamic learning agenda, and regular individual and group evaluations of the citizen science process.

Results: The study was funded in October 2024 and started in January 2025. Data collection started in November 2025 and is expected to end halfway through 2026. Four qualitative lifeline interviews had been conducted as of December 6, 2025. Data analyses are planned for 2026. Manuscripts reporting findings on the central research question and the process evaluations will be submitted for publication in 2027.

Conclusions: Toward the end of the study in 2027, we will develop and disseminate concrete recommendations for various stakeholders to prevent CVD and recognize CVD earlier among people with financial hardship. In all our analyses and recommendations, we will consider sex and gender differences. Our study could contribute to a reduction in health inequalities.

International Registered Report Identifier (IRRID): DERR1-10.2196/89101

(JMIR Res Protoc 2026;15:e89101) doi:10.2196/89101

KEYWORDS

cardiovascular disease; citizen science; socioeconomic status; qualitative research; women’s health

Introduction

Cardiovascular disease (CVD) is the leading cause of death worldwide. In 2019, an estimated 17.9 million people died from CVD, accounting for 32% of all global deaths. Of these, 85% were due to myocardial infarction and stroke [1]. In addition, looking at global causes of death from noncommunicable diseases among individuals younger than 70 years, 38% were attributed to CVD in 2019 [2]. Individuals with a lower socioeconomic position (SEP) face a significantly higher risk of developing CVD than those with a higher SEP worldwide [3,4]. This pattern is consistent across world regions, with approximately 80% of CVD deaths occurring in low- and middle-income countries [1]. Additionally, studies in both Australia and the Netherlands show that people in areas or groups with the lowest SEP have myocardial infarction and stroke rates 1.5 to nearly 2 times higher than those in areas or groups with the highest SEP [5-10]. Looking at socioeconomic differences in CVD death rates, the number of deaths attributed to CVD is higher for households with lower financial prosperity [11].

XSL•FO
RenderX

Several mechanisms explain the link between a lower SEP, lower income, or financial hardship and increased CVD risk. First, chronic psychosocial stress due to financial hardship represents an important mechanism leading to increased risk of CVD [12,13]. Financial insecurity, debt, unemployment, and social isolation cause chronic stress, which elevates stress hormones such as cortisol, triggers inflammatory responses, and results in vascular damage [14,15]. This chronic stress activates both the immune system and the hypothalamic-pituitary-adrenal axis, leading to chronic inflammation, hypertension, insulin resistance, abdominal fat accumulation, and other metabolic dysfunctions—all of which are known contributors to CVD development and progression [13,15,16].

Second, behavioral factors can play a role. Financial scarcity affects decision-making capacity by reducing cognitive bandwidth, forcing individuals to prioritize immediate needs while limiting resources for health-promoting behaviors. This scarcity mindset creates trade-offs between essential expenses and health care, contributing to treatment nonadherence and delayed care seeking [17]. Financial hardship restricts healthy choices and increases the likelihood of unhealthy behaviors such as tobacco use, poor diet, physical inactivity, and medication nonadherence, often used as coping mechanisms [8].

Third, limited access to care is a possible mediator between low income or financial hardship and CVD risk. Financial barriers hinder access to preventive care, necessary medications, treatments, and rehabilitation after cardiovascular events [8,18]. Finally, environmental factors can play a role. People with low incomes are more likely to live in environments with more stressors, such as food deserts, social isolation, and lack of safe spaces for physical activity, perpetuating unhealthy behaviors [10]. All the aforementioned factors can interact and reinforce each other, collectively increasing the risk and severity of CVD [8,18,19].

Although statistical links between SEP and CVD are well described and some of the mechanisms are known, there is a lack of in-depth understanding of the lived experiences and stories of people with CVD who have difficulty making ends meet. This is supported by a review by Powell-Wiley et al [20], which recommended future research to “incorporate qualitative approaches to better understand how individual lived experiences of marginalization affect cardiovascular health outcomes.” Specifically, little is known about how these individuals perceive and manage their health daily or how various contextual factors and mechanisms interact. Research on CVD recognition among women with financial difficulties is also scarce. Women have a poorer prognosis after cardiovascular events and experience higher postevent mortality rates. Additionally, women more frequently present with atypical symptoms (eg, extreme fatigue, shortness of breath, or dizziness) that are recognized later, contributing to delayed diagnosis and treatment [21,22]. These symptoms can resemble emotional and mental stress caused by financial hardship. Stress, a risk factor for CVD, can hinder symptom recognition, making women in this group particularly vulnerable to late diagnosis [8,23]. Finally, factors such as gender discrimination, a greater socioeconomic burden, and constraints on physical mobility among women may reduce their access to optimal cardiovascular care [22].

In this study, we focus on an underresearched group of people in a vulnerable position with an increased risk of CVD: people who have difficulty making ends meet. Our study is developed and conducted in cocreation with citizen scientists who have personal experience with financial hardship and/or CVD. Their experiential knowledge enriches all phases of the research, ensuring better alignment with the lived reality and needs of this group [24]. This participatory approach aligns with established frameworks for community health promotion that emphasize the importance of involving people with lived experience as equal partners in research [25]. This cocreation process increases the relevance of the research questions and the practical applicability of the results.

This study is conducted in the Netherlands. In 2024, almost half of Dutch households were classified as financially unhealthy or vulnerable, which is reflected in a lower healthy life expectancy for this group [26]. In the Netherlands, 23% of deaths in 2022 were attributed to CVD [27]. The aim of the In a Heartbeat study is to recognize CVD earlier and reduce CVD among people who have difficulty making ends meet. Our central research question is, How and in what groups does CVD develop among people who have difficulty making ends meet? Our research subquestions are (1) How can CVD be recognized early in this group, and what are the differences between women and men? (2) What are the best approaches to prevent CVD in this group, and do these differ for women and men? Answering these questions will result in developing concrete recommendations for various stakeholders, namely, people who have difficulty making ends meet themselves, health care and social professionals, health promoters, and policymakers. For all recommendations, we consider sex and gender differences. Consequently, these recommendations could contribute to a reduction in health inequalities.

Methods

Citizen Science

The study project team consists of 9 citizen scientists and 5 professional scientists working together as equals from the start. Eight citizen scientists are involved in an existing citizen advisory group, and 1 (the first author of this paper) is a certified expert by experience in poverty and health who also serves as co–project leader. All 9 citizen scientists have or have had experience with financial hardship, and some also have experience with CVD. The project team is supported by a supervisory committee with expertise in cardiology, sex and gender differences, citizen science, and socioeconomic health disparities. All citizen scientists are involved in every research phase as the entire process is carried out in cocreation. The collaboration is based on equality, open communication, and flexibility. This collaborative model follows principles of participatory systems approaches that have proven effective in community health promotion with people in a vulnerable position, emphasizing flexibility and shared decision-making [25]. We also follow the 10 principles of citizen science from

XSL•FO
RenderX

the European Citizen Science Association [28]. Every project team member is paid for the time that they contribute to the project through an employment contract, a self-employed arrangement, or a volunteer allowance. The study duration is 2.5 years, and the study started in January 2025.

Research Design

The research focuses on unraveling the mechanisms and contexts through which financial problems lead to the development of CVD, with special attention to sex and gender, health behaviors, and mental health. The main data for this study are collected through qualitative lifeline interviews. Additionally, we have weekly project team meetings in which methods such as photo elicitation, critical-creative hermeneutic analysis, design thinking, a dynamic learning agenda, and regular evaluations are used.

Qualitative Interviews

We aim to conduct 40 to 60 interviews with adults (≥18 years) who have difficulty making ends meet (self-report) and have CVD (diagnosis). At recruitment, we ask potential participants whether they have difficulty making ends meet ("moeite met rondkomen" in Dutch) and whether they have a diagnosis of CVD (including high blood pressure or high cholesterol). Within the study population, the aim is to have an equal number of women and men and diversity in age, marital status, educational level, and cultural background. Interviewees are recruited through networks of citizen scientists and professional scientists, through relevant organizations, with the help of media attention for our study, and also through snowball sampling.

Interviews are conducted in duos consisting of a professional scientist and a citizen scientist or 2 citizen scientists. During the preparation phase of the study, citizen scientists receive interview training until they are confident enough to perform the interviews. This training includes general conversation techniques and concludes with practice interviews with mock interviewees.

Interviews take place at a location of choice by the interviewee, which can be at their home or, for example, at a community center or the university. Interviewees complete a short questionnaire including demographic characteristics (gender, age, educational level, household composition, and employment situation), their health situation (blood pressure, presence of diabetes, height, weight, and loneliness), lifestyle behaviors (tobacco use, alcohol use, and sedentary behavior), and financial situation (using the Psychological Inventory of Financial Scarcity [29]). During the interview, a timeline of life events related to CVD and financial challenges is created by one of the interviewers based on the input of the interviewee. On this "lifeline," interviewees describe both their difficulty making ends meet and their CVD severity on a scale from 0 to 10, spanning from birth to the present. This timeline approach was chosen to specifically examine the relationship and trajectories between financial difficulties and CVD over time. The timeline allows for a retrospective lens to elicit autobiographical data covering personal disease trajectories [30]. The topic list includes questions about the financial situation of the interviewees, their experiences with CVD, health behaviors, sex- and gender-specific risks, mental health, living environment, and their interactions. The interviews are estimated to take 1 to 1.5 hours to complete. After the interview, emotional support for interviewers and interviewees is available when needed. This is provided by the expert by experience in the team and, if necessary, a psychologist.

Photo Elicitation

Photo elicitation is a technique in which photographs are incorporated into research interviews [31]. Including photographs during interviews can lead to deeper reflections about participants' experiences than what would be achieved through verbal questions [32]. It also helps build trust and improve communication during interviews, especially when experiences are emotionally charged [33]. The process to perform photo elicitation varies between studies as it is possible for interviewees as well as researchers to generate the photographs used [34].

Within this project, photo elicitation is used as both an interview and analysis technique. For the first 20 interviews, citizen scientists conducting the interviews take one or more photographs after each interview that capture their interpretation of the experiences of the interviewees. These photographs are discussed within the project team and used in data analysis. After the first 20 interviews, we decide whether it is feasible to add photo elicitation to the subsequent interviews, in which the photographs of the citizen scientists can serve as prompts for discussion with interviewees. This decision depends on how the previous interviews went, the length of these interviews, the quality of data collection thus far, and whether citizen scientists feel comfortable using photographs during interviews.

The goal of this approach is to enrich the data collection and analyses by using photographs. From a practical standpoint, on the one hand, our process minimizes the burden on interviewees (because they do not have to take the photos) while still incorporating photos that seek to capture the essence of the conducted interviews and, on the other hand, ensures that the photographs can be validated by subsequent participants.

Data and Analysis

Qualitative interviews are transcribed and rewritten into anonymized readable stories, and the lifelines are digitized and anonymized. The stories and lifelines are checked by the interviewees before they are used in the analysis. The analysis is performed by the entire research team and follows the critical-creative hermeneutic framework by Cardiff and van Lieshout [35]. After the first interviews are performed, everyone reads or listens to the first anonymized stories and examines the lifelines and photographs. During creative analysis sessions, citizen scientists and professional scientists each make a creative display of the most important elements that they see in the stories, lifelines, and photos. Subsequently, they show their creative displays to the team and search for recurring elements in them, which they code. Phases of analysis are followed by, and inform, subsequent phases of interviewing in an iterative process.

After the creative analysis sessions, the original transcripts are used for thematic coding [36] based on the codes made during

XSL•FO
RenderX

the joint sessions. We examine whether results differ for women and men by making matrices with themes and codes separately for both groups and inspecting these for subgroup differences. If feasible, we additionally search for patterns between other demographic characteristics, health situations, lifestyle behaviors, and financial situations (all measured using the questionnaire) and the themes and codes of the qualitative interviews. Additionally, descriptive statistical analyses are performed on the questionnaires. Finally, the lifelines are examined visually for recurring patterns over time and for relationships between difficulty making ends meet and CVD severity.

Creative Work Sessions

Creative work sessions are organized using design thinking, drawing on principles of participatory research and co-design that facilitate the collaborative integration of participants' experiences and insights in addressing complex challenges [37]. Its participatory and iterative nature is particularly valuable in citizen science as cocreation ensures that research tools and outcomes remain closely aligned with the experiences and perspectives of those involved [38]. In addition, these methods foster inclusive community engagement, actively involving underrepresented groups in the research process [39].

Within this project, design thinking is first applied in the co-design of a lifeline tool, developed together with the entire team to guide the interview process. The tool is collaboratively shaped to ensure that it provides a supportive structure for conducting the interviews and aligns with the interview topic list while also providing a supportive structure for the interviewers, helping them feel confident and facilitating both data collection and their comfort during data collection.

Following the creative and thematic analyses described above, creative work sessions are conducted to collaboratively explore ideas; discuss potential approaches; and iteratively generate recommendations for people, practice, and policy focused on the prevention and early detection of CVD among people experiencing financial hardship. In these work sessions, findings are collectively interpreted and synthesized using affinity mapping, clustering, and thematic synthesis. Citizen scientists and professionals play a key role in assessing the feasibility and relevance of these recommendations, ensuring that they reflect the perspectives and needs of the people themselves and the professionals who work with them. These sets of recommendations constitute the design output and will be presented in formats jointly determined by citizen and professional scientists, such as videos, infographics, exhibitions, or podcasts, with the aim of ensuring that the outputs are meaningful, relevant, and applicable in real-world contexts.

Dynamic Learning Agenda

Throughout the project, the team collectively explores what is already known and what we still want or need to learn using a dynamic learning agenda that is adjusted based on current learning questions [40]. Learning questions can be answered in several ways, for example, by inviting experts to explain more about the topic, by assigning students to perform a literature review on the topic, by organizing training sessions, or through self-study.

Evaluation of Collaboration

The collaboration and progress of our project team are evaluated approximately 6 or 7 times over the course of this project. During each evaluation round, individual interviews take place first, in which every citizen scientist is interviewed by a professional scientist and vice versa. Next, the most important points from these individual evaluation interviews are used as input for a group evaluation. Topics that are covered during the evaluation interviews include collaboration, personal development, challenges, and expectations. In addition, there is room for the interviewees to bring up their own topics that they would like to discuss and use as input for a group evaluation.

The goal of this continuous evaluation process is 2-fold. First, it allows for improvement in collaborating as a group, as various challenges may arise during the different phases of the project. It ensures that every member of the team can share their doubts, needs, and ideas for improvement. Second, it generates valuable data about citizen science as a method and how it is experienced by the different members of the group at different stages of the project.

Ethical Considerations

Both the citizen scientists in the research team and the interviewees in the study belong to a potentially vulnerable group, which makes it particularly important to consider the ethical aspects of the study. Participation as citizen scientists and as interviewees is strictly voluntary. Reciprocity is an important principle in this study. Participation should not only benefit the research but also provide value to those involved. This means that citizen scientists are given opportunities to develop research skills, reflect on their own experiences, and contribute meaningfully to the research process. Interviewees are offered a space to share their stories in a respectful and safe setting, and their insights are used to inform concrete outcomes that may benefit the wider community. Participation in the study is confidential, with the identities (names and email addresses) of the interviewees known to only 2 of the professional scientists and the interviewer (ie, citizen scientist) for purposes of contacting individuals for participation in the interview. Interviewees will not be identifiable in any publications or other dissemination of research results. This will be explained to interviewees in the study information letter before they provide informed consent.

Interviewees will participate after providing written informed consent, and consent will be checked verbally before each interview. Consent is an ongoing process and is also checked during the interview if a participant becomes upset or talks about a particularly sensitive issue. After the qualitative interviews, aftercare is offered to both the citizen scientist interviewer and the interviewee if needed. This is provided by the professional expert by experience in the project team and, if necessary, a psychologist. For interviewers, there is also space during or after project meetings to talk about any triggering things said

XSL•FO
RenderX

during interviews. Interviewees receive a voucher of €25 (US $29.51) for their time.

For the evaluation of the citizen science process, citizen scientists and professional scientists should feel safe to talk about other members of the project team when necessary. Therefore, everyone can choose with whom they want to do the evaluation interviews, and transcripts are anonymized. The Dutch code of conduct for academic integrity is explained during project meetings; the university integrity officer has been introduced, and team members are informed that any integrity concerns can be discussed with this officer.

The study received ethics approval from the ethics committee of Avans University of Applied Sciences (25.004) and the Ethics Committee of the Faculty of Health, Medicine, and Life Sciences at Maastricht University (FHML-REC/2025/049).

Results

The study was funded in October 2024 and started in January 2025. Data collection started in November 2025 and is expected to end halfway through 2026. Four qualitative lifeline interviews had been conducted as of December 6, 2025. Data analyses are planned for 2026. Manuscripts reporting findings on the central research question and the process evaluations will be submitted for publication in 2027.

Discussion

Citizen Science With an Underrepresented Group

This research focuses on a group that often remains underrepresented in research: people who have difficulty making ends meet and, therefore, face an increased risk of CVD. By collaborating with citizen scientists who have personal experience with financial hardship and/or CVD, the study gains a unique perspective [41]. Their experiential knowledge ensures that the research resonates with the everyday realities of those living in these circumstances and that the findings and recommendations from the study will be more societally relevant. This cocreative approach not only enhances the relevance of the research questions and the practical applicability of the results but also fosters stronger connections with the people being interviewed, which could lead to more honest answers [41,42]. In addition, involving representatives from a group that is often underrepresented in research in a central role throughout the project may help empower the priority group and ensures that their perspectives are being heard and amplified [42-44].

A notable methodological feature of this study is the use of design thinking principles. Although increasingly applied in health research, their explicit use within citizen science remains relatively uncommon. Design thinking offers advantages in this context as it emphasizes the role of citizens, allows for meaningful engagement even with small groups, and generates tangible outputs that can be reviewed and adopted by the community [38]. However, applying this approach in potentially vulnerable settings, such as with people experiencing financial hardship, may present specific challenges. While design thinking typically involves structured planning of activities, it is anticipated that flexibility will be essential in this study to accommodate participants' varying capacities, availability, and circumstances without compromising the objectives of the research.

Challenges

Although we expect to encounter several challenges during the study, each of these also presents valuable opportunities for growth and improvement. For example, there is significant differentiation within the group of citizen scientists in terms of educational background, digital literacy, and research experience. This diversity requires accessible communication and clear agreements to ensure that everyone can participate equally [41,45]. By being transparent about the approach and working together on an equal footing, this project contributes to a more positive image of citizen science and reduced stigma regarding financial hardship.

Another challenge concerns the quality of the qualitative interviews, which are conducted by trained citizen scientists who are not professionally trained researchers [46,47]. While this may influence the scientific rigor of the interviews, it also allows participants to better relate to the interviewer and feel more comfortable sharing their stories [46,47]. There is sometimes skepticism about the quality of the data collected by citizen scientists [46]. However, research demonstrates that citizen scientists can gather data of quality to professional settings when proper training, protocols, and quality control measures are implemented [47].

A significant challenge that requires particular attention is the reality of health-related attrition among citizen scientists. Due to their (lived) experience with financial hardship and their own health problems or health crises among family members and close contacts, project team members regularly face periods when they cannot fully participate. This pattern of intermittent availability is, unfortunately, more prevalent in our project group as they experience higher rates of illness, hospitalization, social problems, and bereavement than people who do not have these risks. While experiential knowledge provides the project with invaluable insights and authenticity, it also means that we must accommodate the realities of these health and social challenges. To address this, our approach emphasizes maximum flexibility: team members can take a break when needed, can re-engage when their circumstances allow, and receive ongoing support from the project team [25]. We foster a culture of mutual understanding and support, recognizing that the health challenges that our citizen scientists face are inherently connected to the very issues that our research also seeks to address. This flexibility is not viewed as a limitation but as an essential ethical consideration when working with people in a vulnerable position in participatory research. By maintaining inclusive practices that accommodate varying levels of participation, we ensure that the voices of those affected by the intersection of financial hardship and CVD or other health issues remain central to our research even when their participation must be temporarily reduced.

Compensation and contracts for citizen scientists present additional complexities. Paying citizen scientists who receive social benefits is challenging in practice due to regulations in

XSL•FO
RenderX

the Netherlands. Although volunteer contracts offer a solution, they come with limitations, such as a limit on the amount of compensation, restrictions on travel expense reimbursement, or having to prefinance travel expenses. In addition, these volunteer reimbursements are seen as expense claims and not as a salary by the university and, therefore, are not always paid on the same day of the month or an exact number of days since submitting the expense claim, causing additional financial stress for citizen scientists who already have difficulty making ends meet. This situation demands new procedures and customization from research organizations.

Finally, language use is essential in reducing stigma and disparities. When the language used resonates with the experiences of people, it fosters recognition and trust [41,45]. This connection is crucial for encouraging individuals to share their personal stories and for truly reaching those who might otherwise remain unseen [42,45]. Accessible language also increases the chances for everyone to understand all the information and participate actively, whether as research participants or as citizen scientists [41,45,48]. Therefore, it is important that clear and understandable language is used not only during the research process but also in final knowledge products such as articles and recommendations [41,48]. This approach ensures that findings are relevant to and usable by a broad audience, contributes to equal opportunities in health, and supports the goal of inclusive science [41,45]. For that reason, a visual summary of the study protocol has been added to this publication (Figure 1).

XSL•FO
RenderX

Figure 1. Visual summary of the study.

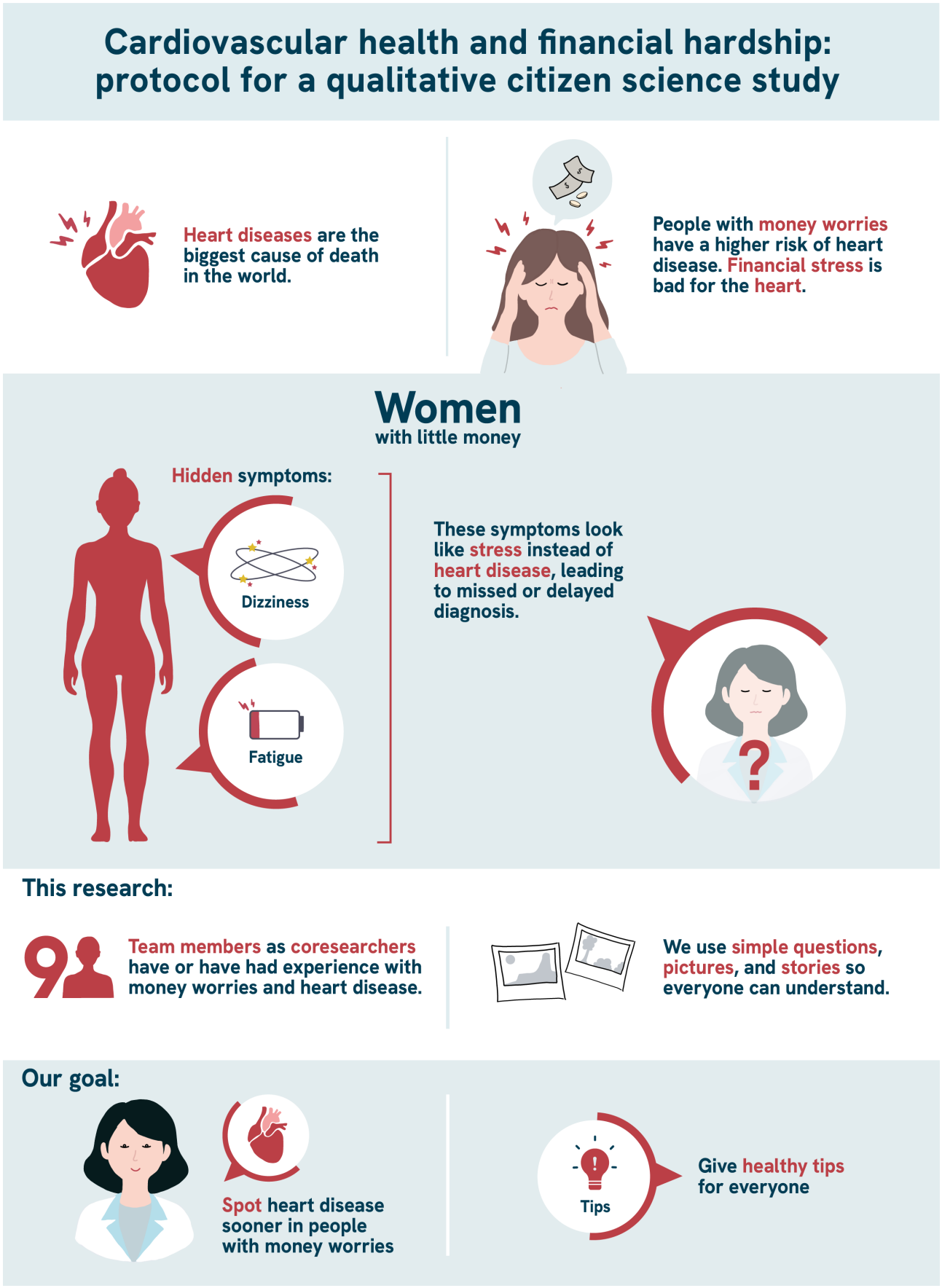

XSL•FO
RenderX

Recommendations and Dissemination

Our study aims to examine multiple determinants of health considering not only medical factors but also lifestyle, mental health, living and working environment, and socioeconomic circumstances [49]. This broad approach leads to a better understanding of the complex challenges that people face. Additionally, this approach will help in formulating recommendations for various stakeholders (people who have difficulty making ends meet themselves, health care and social professionals, health promoters, and policymakers) to prevent CVD and recognize CVD earlier among people having difficulty making ends meet.

The process and results of our study are disseminated in a joint, participatory, and accessible manner [42,50]. For example, a visual summary of the study protocol has been added to this publication (Figure 1), and an accessible website was made in cocreation with citizen and professional scientists and a website designer [51]. Dissemination is conducted through open access academic articles, blogs, practical articles in Dutch, social media posts, media interviews, and recommendations for policy and practice involving both citizen and professional scientists in the process. The advisory committee also plays a role in increasing impact through disseminating our work in their professional networks. This approach not only builds and shares knowledge but also works to strengthen the position of people with experiential knowledge [42,43,49].

Acknowledgments

The authors thank all citizen scientists involved for their valuable contributions to the development, implementation, and ongoing evaluation of this study on cardiovascular health and financial insecurity. Their experiential knowledge and commitment are essential for the depth and relevance of this research. The authors also appreciate the advisory committee for its expert guidance regarding the design and realization of this study. They intend to include other citizen scientists from their team as coauthors in future publications. Finally, the authors acknowledge the Dutch Heart Foundation and Netherlands Organization for Health Research and Development for providing funding and support for this research with citizen scientists. The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: literature search and systematization, translation, and reformatting. The GenAI tool used was Perplexity Pro (Perplexity AI, Inc). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This work was supported by the Netherlands Organization for Health Research and Development and the Dutch Heart Foundation (grant 04270042410003).

Authors' Contributions

GEN, DN, ECEvG, and LA conceived the study together with the other 8 citizen scientists. DN and GEN drafted the manuscript. All authors contributed to the writing of the manuscript, revised it critically for important intellectual content, and read and approved the final version.

Conflicts of Interest

None declared.

Peer Review Report 1
Peer review report by the Netherlands Organization for Health Research and Development.
[PDF File, 165 KB - resprot_v15i1e89101_app1.pdf]

References

1. Cardiovascular diseases (CVDs). World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2026-02-24]
2. Noncommunicable diseases. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2026-02-24]
3. Mackenbach JP, Cavelaars AE, Kunst AE, Groenhof F. Socioeconomic inequalities in cardiovascular disease mortality; an international study. Eur Heart J 2000 Jul;21(14):1141-1151. [doi: 10.1053/euhj.1999.1990] [Medline: 10924297]
4. Rosengren A, Smyth A, Rangarajan S, et al. Socioeconomic status and risk of cardiovascular disease in 20 low-income, middle-income, and high-income countries: the Prospective Urban Rural Epidemiologic (PURE) study. Lancet Glob Health 2019 Jun;7(6):e748-e760. [doi: 10.1016/S2214-109X(19)30045-2] [Medline: 31028013]

XSL•FO
RenderX

5. Indicators of Socioeconomic Inequalities in Cardiovascular Disease, Diabetes and Chronic Kidney Disease: Australian Institute of Health and Welfare; 2019. URL: https://www.aihw.gov.au/reports/social-determinants/indicators-socioeconomic-inequalities/summary [doi: 10.25816/5ebca9d1fa7e0]
6. Kist JM, Vos RC, Mairuhu ATA, et al. SCORE2 cardiovascular risk prediction models in an ethnic and socioeconomic diverse population in the Netherlands: an external validation study. EClinicalMedicine 2023 Feb 16;57:101862. [doi: 10.1016/j.eclinm.2023.101862] [Medline: 36864978]
7. Muthalaly RG, Nelson AJ, Baradi A, Mehta OH, Wilson AM, Nasis A. Socioeconomic determinants of health, traditional risk factors and cardiovascular outcomes in Australia. Int J Cardiol Cardiovasc Risk Prev 2023 Apr 7;17:200184. [doi: 10.1016/j.ijcrp.2023.200184] [Medline: 37131952]
8. Schultz WM, Kelli HM, Lisko JC, et al. Socioeconomic status and cardiovascular outcomes: challenges and interventions. Circulation 2018 May 15;137(20):2166-2178. [doi: 10.1161/CIRCULATIONAHA.117.029652] [Medline: 29760227]
9. Stirbu I, Looman C, Nijhof GJ, Reulings PG, Mackenbach JP. Income inequalities in case death of ischaemic heart disease in the Netherlands: a national record-linked study. J Epidemiol Community Health 2012 Dec;66(12):1159-1166. [doi: 10.1136/jech-2011-200924] [Medline: 22685304]
10. Stringhini S, Carmeli C, Jokela M, et al. Socioeconomic status and the 25 × 25 risk factors as determinants of premature mortality: a multicohort study and meta-analysis of 1·7 million men and women. Lancet 2017 Mar 25;389(10075):1229-1237. [doi: 10.1016/S0140-6736(16)32380-7] [Medline: 28159391]
11. Otten F, Bosma H, Arts K. Sociaaleconomische ongelijkheid in cardiale sterfte. Centraal Bureau voor de Statistiek. 2025. URL: https://www.cbs.nl/nl-nl/longread/statistische-trends/2025/sociaaleconomische-ongelijkheid-in-cardiale-sterfte/3-resultaten [accessed 2026-02-24]
12. Connors K, Flores-Torres MH, Cortés-Valencia A, et al. Serious financial difficulties, psychological stress, and subclinical cardiovascular disease in Mexican women. Ann Epidemiol 2022 Jul;71:38-43. [doi: 10.1016/j.annepidem.2022.03.001] [Medline: 35292402]
13. Tawakol A, Osborne MT, Wang Y, et al. Stress-associated neurobiological pathway linking socioeconomic disparities to cardiovascular disease. J Am Coll Cardiol 2019 Jul 2;73(25):3243-3255. [doi: 10.1016/j.jacc.2019.04.042] [Medline: 31248544]
14. Puterman E, Haritatos J, Adler NE, Sidney S, Schwartz JE, Epel ES. Indirect effect of financial strain on daily cortisol output through daily negative to positive affect index in the Coronary Artery Risk Development in Young Adults Study. Psychoneuroendocrinology 2013 Dec;38(12):2883-2889. [doi: 10.1016/j.psyneuen.2013.07.016] [Medline: 23969421]
15. Steptoe A, Kivimäki M. Stress and cardiovascular disease. Nat Rev Cardiol 2012 Apr 3;9(6):360-370. [doi: 10.1038/nrcardio.2012.45] [Medline: 22473079]
16. Zhou X, Han D, Xu R, et al. A model of metabolic syndrome and related diseases with intestinal endotoxemia in rats fed a high fat and high sucrose diet. PLoS ONE 2014 Dec 11;9(12):e115148. [doi: 10.1371/journal.pone.0115148] [Medline: 25502558]
17. de Bruijn EJ, Antonides G. Poverty and economic decision making: a review of scarcity theory. Theory Decis 2022 Feb;92(1):5-37. [doi: 10.1007/s11238-021-09802-7]
18. Lisipeki S, Masters-Awatere B, Hodgetts D, Liew TV. Reciprocal relations between cardiovascular disease, employment, financial insecurity, and post cardiac event recovery among Māori men: a case series. J Med Case Rep 2023 Nov 12;17(1):468. [doi: 10.1186/s13256-023-04202-7] [Medline: 37951947]
19. Steptoe A, Kivimäki M. Stress and cardiovascular disease: an update on current knowledge. Annu Rev Public Health 2013;34:337-354. [doi: 10.1146/annurev-publhealth-031912-114452] [Medline: 23297662]
20. Powell-Wiley TM, Baumer Y, Baah FO, et al. Social determinants of cardiovascular disease. Circ Res 2022 Mar 4;130(5):782-799. [doi: 10.1161/CIRCRESAHA.121.319811] [Medline: 35239404]
21. Lv Y, Cao X, Yu K, et al. Gender differences in all-cause and cardiovascular mortality among US adults: from NHANES 2005-2018. Front Cardiovasc Med 2024 Feb 14;11:1283132. [doi: 10.3389/fcvm.2024.1283132] [Medline: 38420264]
22. Vogel B, Acevedo M, Appelman Y, et al. The Lancet women and cardiovascular disease commission: reducing the global burden by 2030. Lancet 2021 Jun 19;397(10292):2385-2438. [doi: 10.1016/S0140-6736(21)00684-X] [Medline: 34010613]
23. Mosca L, Barrett-Connor E, Wenger NK. Sex/gender differences in cardiovascular disease prevention: what a difference a decade makes. Circulation 2011 Nov 8;124(19):2145-2154. [doi: 10.1161/CIRCULATIONAHA.110.968792] [Medline: 22064958]
24. Okop KJ, Murphy K, Lambert EV. Community-driven citizen science approach to explore cardiovascular disease risk perception, and develop prevention advocacy strategies in sub-Saharan Africa. Res Involv Engagem 2021;7:11. [doi: 10.1186/s40900-020-00246-x]
25. Wagemakers A, Elkhuizen S, Bindels A, et al. A participatory systems approach in community health promotion: lessons learnt from an overall evaluation of a program to reduce health inequities. Eval Program Plann 2025 Dec;113:102695. [doi: 10.1016/j.evalprogplan.2025.102695] [Medline: 40884876]
26. Financiële gezondheid van Nederlandse huishoudens in 2024: bijna de helft van Nederland financieel kwetsbaar, ondanks de stappen vooruit. Deloitte. URL: https://www.deloitte.com/nl/nl/Industries/financial-services/research/financiele-gezondheid-van-nederlandse-huishoudens.html [accessed 2026-02-24]

XSL•FO
RenderX

27. Koop Y, Wimmers RH, Bots ML, Vaartjes I. Annual figures cardiovascular diseases 2022 [Jaarcijfers hart- en vaatziekten 2022]. 2024. URL: https://www.hartenvaatcijfers.nl/jaarcijfers/jaarcijfers-hart-en-vaatziekten-2022-9a67f [accessed 2026-03-03]
28. ECSA 10 principles of citizen science. OSF. 2019. URL: https://osf.io/xpr2n/overview [accessed 2026-03-03]
29. van Dijk WW, van der Werf MMB, van Dillen LF. The Psychological Inventory of Financial Scarcity (PIFS): a psychometric evaluation. J Behav Exp Econ 2022 Dec;101:101939. [doi: 10.1016/j.socec.2022.101939]
30. Berends L. Embracing the visual: using timelines with in-depth interviews on substance use and treatment. Qual Rep 2011;16(1):1-9. [doi: 10.46743/2160-3715/2011.1036]
31. Harper D. Talking about pictures: a case for photo elicitation. Vis Stud 2002;17(1):13-26. [doi: 10.1080/14725860220137345]
32. Harrison B. Seeing health and illness worlds – using visual methodologies in a sociology of health and illness: a methodological review. Sociol Health Illness 2002 Nov;24(6):856-872. [doi: 10.1111/1467-9566.00322]
33. Shaw D. A new look at an old research method: photo-elicitation. TESOL J 2013;4(4):785-799. [doi: 10.1002/tesj.108]
34. Kyololo OM, Stevens BJ, Songok J. Photo-elicitation technique: utility and challenges in clinical research. Int J Qual Methods 2023;22. [doi: 10.1177/16094069231165714]
35. Cardiff S, van Lieshout F. Het kritisch-creatief hermeneutisch dataanalyseraamwerk: analyseren met in plaats van voor participanten. KWALON 2014 Jun;19(2):30-36. [doi: 10.5117/2014.019.002.030]
36. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology: American Psychological Association; 2012:57-71. [doi: 10.1037/13620-004]
37. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008;4(1):5-18. [doi: 10.1080/15710880701875068]
38. Goi HC, Tan WL. Design thinking as a means of citizen science for social innovation. Front Sociol 2021 May;6:629808. [doi: 10.3389/fsoc.2021.629808] [Medline: 34026900]
39. Bazzano AN, Noel LA, Patel T, et al. Improving the engagement of underrepresented people in health research through equity-centered design thinking: qualitative study and process evaluation for the development of the grounding health research in design toolkit. JMIR Form Res 2023 Feb 28;7:e43101. [doi: 10.2196/43101] [Medline: 36649162]
40. Mierlo BC, Regeer B, Amstel M, et al. Reflexive Monitoring in Action: A Guide for Monitoring System Innovation Projects: Wageningen UR; 2010. URL: https://edepot.wur.nl/149471 [accessed 2026-03-03]
41. Abidi L, van Koeveringe J, Smolka M, et al. Perceptions, barriers and facilitating strategies of inclusive research: a qualitative study with expert interviews. J Underrepresented Minor Prog 2024;8(2):237-263. [doi: 10.32674/cwa36g66]
42. Robinson D, Delany J, Sugden H. Beyond science: exploring the value of co-created citizen science for diverse community groups. CSTP 2024;9(1):13 [FREE Full text] [doi: 10.5334/cstp.682]
43. Barragán CDC. Democratizing STEM: citizen science as a tool for engagement and social change. Am J STEM Educ 2025;11:1-8. [doi: 10.32674/gznjkg56]
44. Pateman RM, West SE. Citizen science: pathways to impact and why participant diversity matters. Citiz Sci 2023;8(1):50. [doi: 10.5334/cstp.569]
45. Fernandez-Sanchez H, Marfo EA, Santa Maria D, Mumba M. Language matters: exploring preferred terms for diverse populations. Glob Qual Nurs Res 2024 Sep 3;11:23333936241275266. [doi: 10.1177/23333936241275266] [Medline: 39233769]
46. Villa G, Rosa D, Marcomini I, et al. Community-driven research: exploring the potential of citizen science in nursing. Int J Nurs Stud 2025 Sep;169:105125. [doi: 10.1016/j.ijnurstu.2025.105125] [Medline: 40466320]
47. Kosmala M, Wiggins A, Swanson A, Simmons B. Assessing data quality in citizen science. Front Ecol Environ 2016;14:551-560. [doi: 10.1002/fee.1436]
48. Ortega P, Osman-Krinsky M, Silva D. Slow and steady: using inclusive language to enhance academic medicine's social accountability. Acad Med 2024 Dec 1;99(12):1323-1327. [doi: 10.1097/ACM.0000000000005792] [Medline: 39601600]
49. Braveman P, Egerter S, Williams DR. The social determinants of health: coming of age. Annu Rev Public Health 2011;32:381-398. [doi: 10.1146/annurev-publhealth-031210-101218] [Medline: 21091195]
50. Campos R, Monteiro J, Carvalho C. Engaged citizen social science or the public participation in social science research. J Sci Commun 2021;20(6):A06. [doi: 10.22323/2.20060206]
51. Home page. In a Heartbeat Project. URL: https://inaheartbeatproject.nl/ [accessed 2026-02-28]

Abbreviations

CVD: cardiovascular disease
SEP: socioeconomic position

XSL•FO
RenderX

Edited by J Sarvestan; submitted 06.Dec.2025; this is a non–peer-reviewed article;accepted 23.Jan.2026; published 13.Mar.2026.

Please cite as:
Niewold D, van Gestel ECE, Abidi L, Tiemstra K, Thoolen M, Nagelhout GE
Cardiovascular Health and Financial Hardship: Protocol for a Qualitative Citizen Science Study
JMIR Res Protoc 2026;15:e89101
URL: https://www.researchprotocols.org/2026/1/e89101
doi:10.2196/89101

© Dagmar Niewold, Evy C E van Gestel, Latifa Abidi, Klaske Tiemstra, Myrte Thoolen, Gera E Nagelhout. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Use of Vending Machines to Deliver Oral Rapid HIV Self-Tests to Veterans: Protocol for a Pilot Study

Tessa Rife-Pennington[1,2,3], PharmD; Michael P Douglas[1,2], MS; Wendy Xie[1,2], BS; Jennifer Cocohoba[2], PharmD, MAS

[1]Pharmacy Service, San Francisco VA Health Care System, 4150 Clement St (119), San Francisco, CA, United States

[2]Department of Clinical Pharmacy, School of Pharmacy, University of California, San Francisco, San Francisco, CA, United States

[3]Department of Pharmacy Practice, Thomas J Long School of Pharmacy, University of the Pacific, San Francisco, CA, United States

Corresponding Author:
Tessa Rife-Pennington, PharmD
Pharmacy Service, San Francisco VA Health Care System, 4150 Clement St (119), San Francisco, CA, United States

Abstract

Background: California has the largest number of people living with HIV in the United States, and in 2022, there were 4882 new diagnoses. Veterans with histories of substance use, viral hepatitis, sexually transmitted infections, and homelessness carry substantial HIV burden. Testing is essential, yet approximately 12% of Californians with HIV were undiagnosed in 2020, and 50% of veterans in care had never been tested as of 2023. HIV self-tests (HIVSTs) can mitigate stigma, confidentiality, and access barriers, and vending machines (VMs) offer private, convenient distribution. However, the use of VM-dispensed HIVST has not been evaluated for veterans or within Veterans Affairs (VA) settings.

Objective: We describe a Reach, Evaluation, Adoption, Implementation, and Maintenance–guided pre-implementation protocol to evaluate VM-dispensed HIVSTs in Northern California VA clinics and supportive housing settings.

Methods: Fifteen VMs will stock oral-fluid HIVSTs (n=900). Program data (de-identified dispense logs), veteran electronic surveys (n=90), and qualitative interviews (n=15) will quantify reach (uptake), early effectiveness proxies (use, results, and next steps), adoption (machine/site dispensing), implementation (stockouts, restocking interval, and costs), and maintenance (dispensing trends).

Results: Ethics approval activities (study material development and Institutional Review Board submission) began 2 months prior to the receipt of award funding (January to February 2025). Following funding, the project is planned for over a total of 18 months (12 months original project period + 6 months no-cost extension; March 2025 to August 2026). Ethics approval was obtained in August 2025. Veteran feedback was incorporated into study materials, and HIVSTs were purchased and packaged in September to November 2025. HIVSTs were added to VMs, and data collection is projected to occur from December 2025 through June 2026. Results are anticipated to be available in August 2026.

Conclusions: This study will generate practice-ready evidence on the feasibility, acceptability, and early behavioral impacts of VM-dispensed HIVSTs for veterans. By pairing a stigma-responsive delivery channel with pragmatic measures, findings can inform equitable scale-up across VA and community settings, guide comparative evaluations of distribution channels (VMs, mail-to-home, or clinic pick-up), and support privacy-preserving linkage strategies to confirmatory testing, HIV pre-exposure prophylaxis, and treatment. Results will address a critical evidence gap for veteran-focused HIV prevention and provide parameters for multi-site evaluations.

International Registered Report Identifier (IRRID): PRR1-10.2196/84317

(JMIR Res Protoc 2026;15:e84317) doi:10.2196/84317

KEYWORDS

veterans; HIV; self-testing; sexually transmitted diseases; vending machine

Introduction

San Francisco and Alameda Counties, core jurisdictions in the United States Ending the HIV Epidemic initiative, continue to experience concentrated HIV burden among populations served by the San Francisco Veterans Affairs Health Care System (SFVAHCS) [1]. Despite national recommendations for routine and risk-based screening, testing within Veterans Affairs (VA) care centers remains suboptimal, with many veterans never having been tested [2,3]. Addressing this persistent screening gap requires approaches that extend beyond traditional clinic workflows and that better reflect how veterans prefer to seek prevention services.

Multiple barriers impede routine testing in VA and community settings alike. Veterans report concerns about confidentiality

XSL•FO
RenderX

and stigma, competing demands during clinical visits, uncertainty about eligibility or need for testing, and logistical hurdles such as travel and wait times [4]. Clinicians cite limited time, competing priorities, and variable comfort initiating testing discussions [5]. Framed through a behavioral lens, these barriers cluster around privacy, autonomy, convenience, and perceived control—constructs linked to testing avoidance and delayed diagnosis.

HIV self-tests (HIVSTs) are a recommended complement to facility-based testing and can mitigate several of these barriers by enabling private, user-directed testing [1,6-11]. Distributing HIVSTs through vending machines (VMs) further increases anonymity and convenience, while leveraging existing low-touch infrastructure. Early experience from non-VA settings suggests VM distribution is feasible and acceptable; however, its utility for veterans, and its implementation within VA facilities and partnered housing sites, have not been evaluated [12-16].

This protocol responds to this gap by describing a mixed-methods evaluation of VM-dispensed HIVSTs within the SFVAHCS and affiliated supportive housing locations. Guided by the Reach, Evaluation, Adoption, Implementation, and Maintenance (RE-AIM) framework, this study will characterize reach, adoption, and implementation at the site level and assess early individual-level behaviors related to test use [17]. By pairing a stigma-responsive delivery channel with pragmatic measures, this work aims to generate actionable evidence for VA and community partners considering VM-enabled HIV screening.

Methods

Study Aims and Objectives

This is a single-arm, mixed-methods pilot implementation evaluation that aims to evaluate the RE-AIM of VM-dispensed HIVSTs among Veterans located in Northern California (Table 1) [17].

Table . Study aims and hypotheses.

Aim	Description	Hypothesis
Aim 1: Reach	Characterize HIV risk factors among veterans who obtained ≥1 VM[a]-dispensed HIVST[b].	Among survey respondents, commonly reported HIV risk indicators will include male-to-male sexual contact, condomless sex, and unregulated drug use in the past 6 months.
Aim 2: Effectiveness (early outcomes; not causal)	Describe early, self-reported outcomes after obtaining a VM-dispensed HIVST (eg, intended use, completed use, results, and planned next steps) and the perceived usefulness of embedded HIV resources.	Most participants will report intending to use or having used the HIVST, will report reviewing the included post-test resources, and will report an intention to seek confirmatory testing and/or prevention services (eg, HIV pre-exposure prophylaxis).
Aim 3: Adoption	Identify which VM locations have the highest HIVST dispensing volume and when dispensing occurs (eg, clinic vs supportive housing; weekday vs after hours).	HIVST dispensing will be the highest at locations with higher baseline VM utilization and/or higher concentration of veterans experiencing housing instability (eg, supportive housing sites or outpatient clinics serving veterans experiencing homelessness).
Aim 4: Implementation and Maintenance	Assess implementation outcomes (eg, direct costs) and acceptability/experience of veterans with VM-dispensed HIVST (eg, convenience, privacy, satisfaction, and preference compared with other access points).	Participants will rate the VM delivery method as highly acceptable (eg, convenient, private, and satisfactory) and will describe VM access as a low-barrier, low-stigma option for obtaining HIV testing; dispensing will be feasible to maintain over the pilot period with measurable restocking workflows.

[a]VM: vending machine.

[b]HIVST: HIV self-test.

To accomplish this, we will collect veteran self-report data from 15 VMs via an electronic web-based Questionnaire 1 (n=90), a semi-structured qualitative interview paired with Questionnaire 2 (n=15), and HIVST distribution data (n=900) via the EMS software (VendNovation LLC). Our study lacks a comparison condition. Accordingly, the “effectiveness” outcomes are limited to early, self-reported behaviors and perceived usefulness/acceptability; causal inference regarding changes in HIV testing, diagnosis, or linkage to care is not possible.

Timeline

The study is expected to be completed over 20 months (reflecting 2 months of pre-funding activities, 12 months for the original project period, and a requested 6-month no-cost extension). During months 1 - 2 (January to February 2025), research team members designed initial study materials and submitted Institutional Review Board (IRB) protocols for review. During months 3 - 8 (March to August 2025), requested edits to IRB protocols were incorporated, and IRB approval was attained. During months 9 - 11 (September to November 2025), veteran feedback was obtained about the study materials, which were further refined and finalized; HIVSTs and VM coils

XSL•FO
RenderX

were purchased; and HIVSTs were packaged with the study materials. HIVSTs were added to the VMs in December 2025. Data collection is projected to occur in months 12 - 18 (December 2025 to June 2026), and data analysis is anticipated to be performed in months 19 - 20 (July to August 2026).

Setting

This study will be conducted at the SFVAHCS, which serves over 310,000 US Veterans at the San Francisco VA Medical Center and 9 community-based outpatient clinics located in Downtown San Francisco, Oakland, San Bruno, Santa Rosa, Clearlake, Ukiah, and Eureka, California. HIVSTs will be stocked in pre-existing SFVAHCS VMs: 2 are located at the San Francisco VA Medical Center, 7 in community-based outpatient clinics (all the cities listed above except San Bruno), and 6 in supportive housing buildings where veterans live (San Francisco and Daly City, California). Research study team members will conduct qualitative interviews with veterans at the San Francisco VA Medical Center and supportive housing buildings. The research study team will also be available via telephone and video to assist with recruitment, enrollment, and collection of study measures.

Participant Eligibility and Recruitment

To be eligible for Questionnaire 1, participants must be a US Veteran (of any service discharge status), must be registered for VM access, and must have taken a HIVST from one of the 15 SFVAHCS VMs. To be eligible for the qualitative interview and Questionnaire 2, veterans must meet the eligibility criteria for Questionnaire 1 and must have personally used a VM-dispensed HIVST. Non-veterans who obtain a VM-dispensed HIVST from a veteran will be ineligible. Recruitment fliers will be taped on the VMs, HIVSTs, and posted in common areas of supportive housing buildings. The research team will attend veteran community meetings at supportive housing buildings to introduce the study and recruit participants.

Ethical Considerations

We obtained approval by the University of California IRB and the San Francisco Veterans Affairs Human Subjects Committee in August 2025 (No. 25-43931). The research study team will obtain electronic or written informed consent before beginning the study questionnaire and qualitative interview. Participants will receive a US $20 electronic gift card after completion of Questionnaire 1 (maximum of US $40 per participant). Veterans who completed Questionnaire 2 would receive a US $100 gift card. Each veteran advocate will receive a US $1000 honorarium for their time and expertise (about 5 hours of work). HIVST dispensing data will be de-identified and tracked regularly throughout the study using HRVM software. All participant data collected will be de-identified to ensure privacy and confidentiality.

Study Measures and Outcomes

Study measures and outcomes will be mapped onto the RE-AIM framework for evaluating individual and organizational factors that determine public health impacts of a program or policy (Table 2) [17].

Table . Application of the RE-AIM[a] framework to the proposed study.

Dimensions	Measures/Outcomes	Data source	Data level
Reach the target population	Number of participants who access ≥1 HIVST[b]	HRVM software (EMS, VendNovation LLC)	Organizational
	Participant demographics and characteristics, sexual partners and practices, drug use practices	Electronic Questionnaires 1 and 2	Individual
Effectiveness or efficacy	Participant HIVST usage plan, results, and steps after testing	Electronic Questionnaire 1	Individual
Adoption by target staff, settings, systems, and communities	Number of HIVSTs dispensed, location, date, and time	HRVM software (EMS, VendNovation LLC)	Organizational
Implementation consistency, costs, and adaptations made during delivery	Direct costs (vending machine coils and HIVSTs)	Purchase order invoices	Organizational
Maintenance/sustainment of intervention effects in individuals and settings over time	Participant reported outcome measures	Electronic Questionnaire 1; semi-structured qualitative interview	Individual

[a]RE-AIM: Reach, Evaluation, Adoption, Implementation, and Maintenance.

[b]HIVST: HIV self-test.

The RE-AIM framework has been utilized in similar studies evaluating health care provider views on HIVSTs, workplace HIV testing, integrating HIV prevention services into family planning, barriers and facilitators to HIV pre-exposure prophylaxis uptake, and provision of conditional financial incentives to people living with HIV [18-22].

Study Instruments

Questionnaires 1 and 2 will be developed in Qualtrics in January 2026, and both will evaluate self-reported sociodemographic characteristics, including age; gender identity; race/ethnicity; sexual orientation; branch of military service; service-connected disability rating; past 6-month housing status, sexual practices, and street drug use; and if ever heard of, willing to take, or has

XSL•FO
RenderX

taken HIV pre-exposure prophylaxis, HIV post-exposure prophylaxis, or doxycycline PEP in the past year. These questions will be adapted from the Collecting Demographic Data at Syringe Services Programs guide, Core Questionnaire, internal VA guidance on taking a brief sexual health history, and self-efficacy scale for HIV risk behaviors [23-25]. Questionnaire 1 will include additional questions, including HIVST usage plan, results, and next steps. Added questions will be derived from the Theoretical Framework of Acceptability Questionnaire, which evaluates affective attitude, burden, ethicality, intervention coherence, opportunity costs, perceived effectiveness, and self-efficacy of health care interventions, as well as one item from the Net Promotor Score "How likely is it that you would recommend to a friend or a colleague?"; they will be asked to rate convenience, privacy, satisfaction, reduced stigma, preference over other testing methods, ability to pay for testing, and other recommended self-tests to add to the VMs (eg, chlamydia/gonorrhea self-swabs) [26,27].

A semi-structured qualitative interview guide will be developed. Participants will be asked about prior HIV testing experiences, knowledge of self-testing, experience accessing via VMs, perceived ease of use, motivations for self-testing, concerns, impact of stigma, support needs, and recommendations for other self-test distribution via VMs [28].

Veteran Advocates

We will use participatory action research methodology and recruit 4 veteran advocates with lived/living expertise of HIV who reside in California to provide feedback and recommendations on study materials and assist with interpretation and dissemination of study findings [29-31]. Veterans will provide feedback and recommendations on study recruitment materials, education handouts, and instruments (Questionnaires 1 and 2 and the qualitative interview guide) via a feedback form and a virtual focus group (Multimedia Appendix 1). Feedback will be incorporated prior to resource utilization. Each veteran advocate will receive a US $1000 honorarium for their time and expertise (about 5 hours of work).

Study Administration and Participant Reimbursement

Veterans who dispensed a HIVST via VMs may anonymously complete Questionnaire 1 via web-based link located on the HIVST. Participants will receive a US $20 electronic gift card after completion (maximum of US $40 per participant). Reimbursement will be available for a total of 90 completed questionnaires.

Veterans (n=15) will meet with the research team members for approximately 1 hour to complete an in-person semi-structured qualitative interview. Veterans will need to complete Questionnaire 2 via an iPad, and those who did so would receive a US $100 gift card. Veterans may complete Questionnaire 1 and/or the semi-structured qualitative interview visit (both are not required).

Data Collection and Storage

Interviews will be audio recorded via Microsoft Teams, validated by research team members, and submitted for transcription [32]. Collected Qualtrics questionnaire and semi-structured qualitative interview data will be secured in password-protected files on the SFVAHCS network in January 2026, and access will be restricted to research study staff. Participants will be assigned alpha-numeric identifiers, and files with identifying information will be destroyed upon completion. HIVST dispensing data will be de-identified and tracked regularly throughout the study using HRVM software (EMS, VendNovation LLC).

Sample Size and Precision

This pilot implementation evaluation is designed to estimate feasibility and acceptability outcomes rather than to test efficacy. Funding was allocated for the purchase of 900 HIVSTs. Assuming 100% distribution during the data collection period, we estimated a ~10% optional survey response rate based on a similar HIVST distribution study and targeted up to 90 completed surveys [16]. With 90 survey responses, descriptive estimates of key feasibility/acceptability proportions (eg, intended/actual test use, satisfaction, and perceived privacy) will have acceptable precision (eg, for proportions near 50%, an approximate 95% CI half-width of ~10 percentage points). We will conduct up to 15 qualitative interviews, which is expected to be sufficient to achieve thematic saturation for a focused inquiry on acceptability, barriers/facilitators, and implementation experiences.

Data Analysis

Descriptive statistics will be used to evaluate participant characteristics and organizational-level data. For the semi-structured qualitative interviews, an inductive coding process will be used employing thematic analysis [33]. Research team members will independently code three veteran transcripts, resolve discrepancies, and refine the initial code book collaboratively. Once the final codebook is generated, transcripts will be divided evenly among team members for coding using qualitative research software (Dedoose, Sociocultural Research Consultants LLC, Version 9.2.22). Coding will be verified via a 10% cross-check to ensure consistency across team members. Patterns in coding, categories, and themes/concepts will be examined to draw conclusions about the data [33].

Results

Ethics approval was obtained in August 2025. Veteran written (n=4) and focus group (n=3) feedback was obtained in September 2025, and study materials were refined by November 2025. HIVSTs and VM coils were purchased, HIVSTs were packaged with study materials, and all VMs were anticipated to be stocked with HIVSTs by December 2025. Data collection from VM dispensing and semi-structured qualitative interviews are projected to occur December 2025 through June 2026, with final analyses in July to August 2026.

Discussion

Principal Findings and Comparison With Previous Works

This protocol addresses a persistent gap in HIV screening among veterans despite US and national VA recommendations for

XSL•FO
RenderX

routine or risk-based testing [2,3,34]. With California's high-burden context, only half of the veterans in care have ever been tested, underscoring the need for alternative, low-barrier testing strategies that can extend reach beyond clinic workflows [35-37]. While HIVSTs are accurate and acceptable across diverse settings, distribution via VMs has not been evaluated in veteran populations, representing a vital evidence gap this study will begin to fill [8-11].

VM-dispensed HIVSTs are a promising, technology-enabled approach precisely because they target known behavioral and structural barriers to testing, including stigma, confidentiality concerns, and competing clinical priorities. By enabling private, on-demand access outside traditional encounters, VM delivery may increase autonomy, convenience, and reduce disclosure concerns, mechanisms frequently identified as drivers of testing avoidance and delayed diagnosis. Empiric work in non-VA settings indicates that VM access to HIVSTs is feasible, acceptable, and capable of engaging first-time testers across public venues and community spaces; adapting this model to VA facilities has potential to extend reach while ensuring privacy and choice [12-16].

Our mixed-methods design, mapped onto the RE-AIM framework, is intended to generate implementation-ready knowledge [17]. Quantitatively, dispensing data from 15 VMs and participant surveys will estimate reach (who accesses tests), early effectiveness proxies (intended/actual use and next steps), adoption (dispensing volume, time, and location), and implementation (direct costs, fidelity, and adaptations). Qualitative interviews will illuminate context, equity, and mechanisms, including how privacy, convenience, and autonomy shape decisions to test and seek confirmatory services, providing critical explanatory depth that can guide scale-up interventions.

Thus, this protocol advances a low-barrier, stigma-responsive strategy to expand HIV screening among veterans by pairing the privacy and convenience of self-testing with the scalability of VM infrastructure. Findings on reach, acceptability, adoption, and early post-test behaviors will provide actionable guidance for VA and community partners considering VM-enabled HIV prevention programs, while delineating the analytic and operational groundwork for larger, multi-site trials.

Limitations

This study protocol has several limitations. First, participation requires proximity and access to SFVAHCS VMs, which may preferentially include veterans who are already engaged with VA services or comfortable with public VM use. As such, reach estimates may under-represent veterans who avoid clinic spaces or public machines. Second, the evaluation is limited to a single VA health system with clinics and supportive housing partners in Northern California; differences in housing policies, facility layouts, and community norms may constrain adoption and transferability to other VA and non-VA settings. Third, private self-testing and reliance on self-reported survey/interview data introduce recall bias, social desirability bias, and limit the verification of reactive results or care linkage. Fourth, inequities in technology access, health literacy, or trust in research staff may shape who participates, potentially biasing acceptability and feasibility estimates. Finally, as a 12-month pilot study, outcomes will emphasize uptake, acceptability, and implementation (RE-AIM) of HIVST distribution via VMs, rather than characterize HIV incidence, engagement in care, or sustained outcomes.

Future Directions

Findings from this pilot should inform a multi-site VA evaluation that tests scale-up strategies and the comparative effectiveness of distribution channels (VMs vs mail-to-home vs clinic pick-up) on reach, equity, and care linkage. A pragmatic stepped-wedge or cluster randomized design across VA and supportive housing partners would allow the estimation of system-level effects while accommodating staged implementation. Future work should also examine reductions in stigma, increased autonomy/convenience, and changes in self-testing efficacy among veteran subgroups (eg, women and gender-diverse veterans, those experiencing homelessness, rural vs urban settings, and diverse racial/ethnic groups).

To strengthen downstream outcomes, next-phase studies could embed privacy-preserving navigation options (eg, quick response [QR] codes or links inside HIVSTs that directly connect veterans to follow-up testing and HIV pre-exposure prophylaxis). Future economic evaluations (implementation costs, and budgetary impact) will be vital for sustainability decisions. Implementation research should compare site-level strategies (staff champions, restocking workflows, and accessibility features) and assess maintenance over time (supply chain reliability, vandalism, theft, policy, and contracting steps). Finally, bundling HIVSTs with sexual health resources (pregnancy tests or syphilis self-tests) warrants further testing to evaluate whether integrated harm-reduction VMs increase first-time HIV testing and accelerate linkage to preventive care and treatment.

Acknowledgments

This manuscript was reviewed using OpenAI's ChatGPT (version GPT-5, August 2025) for proofreading, clarity, grammar, and consistency. No scientific content, data, analyses, or references were generated by the tool; all text was verified and approved by the authors.

This work represents the views of the authors and does not necessarily represent the position or policy of the United States Government, Department of Veterans Affairs, or any other affiliation.

XSL•FO
RenderX

Funding

This work is funded by a 2024-2025 California HIV-AIDS Research Program (CHRP) Low Barrier Technology Interventions for HIV Prevention and Care grant (H25TC8886S) and the American Academy of HIV Medicine Dr. Dawn K. Smith HIV Prevention Clinical Fellowship, both to Dr. Tessa Rife-Pennington. The CHRP award underwent external peer review; the summary statment and score are provided in Peer Review Report. .

Data Availability

No data were generated for this study protocol.

Authors' Contributions

TR-P: conceptualization, methodology, software, resources, writing (original draft, review, and editing), visualization, supervision, project administration, funding acquisition. MPD: conceptualization, methodology, software, writing (review and editing), project administration. WX: investigation, writing (review and editing). JC: writing (review and editing), supervision.

Multimedia Appendix 1
A focus group guide that will be used to interview 4 veteran advocates. The guide includes an overview of the focus group (participants, materials, roles and responsibilities), bulleted outline, facilitator script, references, confirmation of notice of participation, recruitment tracking form, and reviewer feedback form.
[PDF File, 320 KB - resprot_v15i1e84317_app1.pdf]

Peer Review Report 1
Peer Review report from California HIV/AIDS Research Program (CHRP) Review Committee.
[PDF File, 274 KB - resprot_v15i1e84317_app2.pdf]

References

1. Office of Infectious Disease and HIV/AIDS Policy, U.S. Department of Health & Human Services. EHE Overview. 2025 Mar 20. URL: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview [accessed 2026-01-22]
2. U.S. Department of Veterans Affairs. Fact Sheet – Human Immunodeficiency Virus (HIV) – VA is the Largest Single Provider of HIV Care in the United States. 2020 Mar. URL: https://www.hiv.va.gov/pdf/HIV-program-factsheet.pdf [accessed 2026-01-22]
3. Centers for Disease Control and Prevention. Getting Tested for HIV. 2025 Feb 11. URL: https://www.cdc.gov/hiv/testing/index.html [accessed 2026-01-22]
4. Valdiserri RO, Nazi K, McInnes DK, Ross D, Kinsinger L. Need to improve routine HIV testing of U.S. Veterans in care: results of an Internet survey. J Community Health 2010 Jun;35(3):215-219. [doi: 10.1007/s10900-010-9233-x] [Medline: 20146092]
5. Gant-Clark V, Skipper J, McCrackin M, Gaillard S, Coxe DN. Veterans and HIV screening rates: why veterans are not getting tested, a Veterans Affairs Nursing Academic Partnership (VANAP) study. J Assoc Nurses AIDS Care 2015;26(4):485-491. [doi: 10.1016/j.jana.2015.03.002] [Medline: 26066698]
6. Division of HIV Prevention, National Center for HIV, Vital Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention. Self-Testing. 2023 Jun 12. URL: https://www.cdc.gov/stophivtogether/hiv-testing/self-testing.html [accessed 2026-01-22]
7. World Health Organization. WHO recommends optimizing HIV testing services. 2023 Jul 22. URL: https://www.who.int/news/item/22-07-2023-who-recommends-optimizing-hiv-testing-services [accessed 2026-01-22]
8. Gupta-Wright A, Barnabas RV, Ingold H, Duneton P, Abubakar I. HIV self-testing: lessons learnt and priorities for adaptation in a shifting landscape. BMJ Glob Health 2021 Jul;6(Suppl 4):e004418. [doi: 10.1136/bmjgh-2020-004418] [Medline: 34275867]
9. Krause J, Subklew-Sehume F, Kenyon C, Colebunders R. Acceptability of HIV self-testing: a systematic literature review. BMC Public Health 2013 Aug 8;13:735. [doi: 10.1186/1471-2458-13-735] [Medline: 23924387]
10. Vashisht S, Jha S, Thakur N, et al. Comparing the effects of oral HIV self-testing with those of standard HIV testing for men who have sex with men (MSM): a systematic review and meta-analysis. Cureus 2022 Aug;14(8):e28157. [doi: 10.7759/cureus.28157] [Medline: 36158381]
11. Witzel TC, Eshun-Wilson I, Jamil MS, et al. Comparing the effects of HIV self-testing to standard HIV testing for key populations: a systematic review and meta-analysis. BMC Med 2020 Dec 3;18(1):381. [doi: 10.1186/s12916-020-01835-z] [Medline: 33267890]
12. McGuire M, de Waal A, Karellis A, et al. HIV self-testing with digital supports as the new paradigm: a systematic review of global evidence (2010-2021). EClinicalMedicine 2021 Sep;39:101059. [doi: 10.1016/j.eclinm.2021.101059] [Medline: 34430835]

XSL•FO
RenderX

13. Gobin M, Dhillon S, Kesten JM, et al. Acceptability of digital vending machines to access STI and HIV tests in two UK cities. Sex Transm Infect 2024 Feb 19;100(2):91-97. [doi: 10.1136/sextrans-2023-055969] [Medline: 38302411]
14. Young SD, Daniels J, Chiu CJ, et al. Acceptability of using electronic vending machines to deliver oral rapid HIV self-testing kits: a qualitative study. PLoS ONE 2014;9(7):e103790. [doi: 10.1371/journal.pone.0103790] [Medline: 25076208]
15. Woods WJ, Lippman SA, Agnew E, Carroll S, Binson D. Bathhouse distribution of HIV self-testing kits reaches diverse, high-risk population. AIDS Care 2016;28 Suppl 1(sup1):111-113. [doi: 10.1080/09540121.2016.1146399] [Medline: 26883730]
16. Stafylis C, Natoli LJ, Murkey JA, et al. Vending machines in commercial sex venues to increase HIV self-testing among men who have sex with men. Mhealth 2018;4:51. [doi: 10.21037/mhealth.2018.10.03] [Medline: 30505849]
17. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
18. Jones W, Somerset S, Evans C, Whittingham K, Middleton M, Blake H. Test@work: evaluation of workplace HIV testing for construction workers using the RE-AIM framework. BMC Public Health 2021 Sep 24;21(1):1737. [doi: 10.1186/s12889-021-11739-z] [Medline: 34560853]
19. Brant AR, Dhillon P, Hull S, et al. Integrating HIV pre-exposure prophylaxis into family planning care: a RE-AIM framework evaluation. AIDS Patient Care STDS 2020 Jun;34(6):259-266. [doi: 10.1089/apc.2020.0004] [Medline: 32484743]
20. Kadiamada-Ibarra H, Hawley NL, Sosa-Rubí SG, Wilson-Barthes M, Franco RR, Galárraga O. Barriers and facilitators to pre-exposure prophylaxis uptake among male sex workers in Mexico: an application of the RE-AIM framework. BMC Public Health 2021 Nov 27;21(1):2174. [doi: 10.1186/s12889-021-12167-9] [Medline: 34837988]
21. Chrestman S, Patel T, Lass K, et al. Examining the implementation of conditional financial incentives using the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) framework to improve HIV outcomes among persons living with HIV (PLWH) in Louisiana. Int J Environ Res Public Health 2022 Aug 2;19(15):9486. [doi: 10.3390/ijerph19159486] [Medline: 35954839]
22. Nangendo J, Wanyenze RK, Obeng-Amoako GO, et al. Health provider perspectives of village health team-delivered oral HIV self-testing among men in Central Uganda: a qualitative evaluation using RE-AIM framework. Res Sq 2024 Jan 26:rs.3.rs-3816613. [doi: 10.21203/rs.3.rs-3816613/v1] [Medline: 38343851]
23. University of Washington Supporting Harm Reduction Programs (sharp) team collecting demographic data at syringe services programs. Collecting Demographic Data at Syringe Services Programs. 2023 May. URL: https://sites.uw.edu/sharpta/collecting-demographic-data [accessed 2026-01-22]
24. Drug Policy Alliance. Core questionnaire for supervised consumption services (SCS) evaluations. Presented at: Data Harmonization Meeting; Jan 11-12, 2018 URL: https://astrosstudy.org/SCS_Core_questionnaire.pdf [accessed 2026-01-22]
25. Smith KW, McGraw SA, Costa LA, McKinlay JB. A self-efficacy scale for HIV risk behaviors: development and evaluation. AIDS Educ Prev 1996 Apr;8(2):97-105. [Medline: 8727650]
26. Sekhon M, Cartwright M, Francis JJ. Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. BMC Health Serv Res 2022 Mar 1;22(1):279. [doi: 10.1186/s12913-022-07577-3] [Medline: 35232455]
27. Reichheld FF. The one number you need to grow. Harv Bus Rev 2003 Dec;81(12):46-54. [Medline: 14712543]
28. Iliyasu Z, Haladu ZA, Iliyasu BZ, et al. A qualitative study of HIV testing experiences and HIV self-testing perspectives among men in Northern Nigeria. Nurs Res Pract 2024;2024:8810141. [doi: 10.1155/2024/8810141] [Medline: 38682002]
29. Gray C, Crawford G, Lobo R, Maycock B. Co-designing an intervention to increase HIV testing uptake with women from Indonesia at-risk of HIV: protocol for a participatory action research study. Methods Protoc 2019 May 23;2(2):41. [doi: 10.3390/mps2020041] [Medline: 31164620]
30. Salazar ZR, Vincent L, Figgatt MC, Gilbert MK, Dasgupta N. Research led by people who use drugs: centering the expertise of lived experience. Subst Abuse Treat Prev Policy 2021 Sep 20;16(1):70. [doi: 10.1186/s13011-021-00406-6] [Medline: 34544478]
31. Closson K, McNeil R, McDougall P, et al. Meaningful engagement of people living with HIV who use drugs: methodology for the design of a Peer Research Associate (PRA) hiring model. Harm Reduct J 2016 Oct 7;13(1):26. [doi: 10.1186/s12954-016-0116-z] [Medline: 27717364]
32. Rev. URL: https://www.rev.com/ [accessed 2026-01-21]
33. Dill L. National Institutes of Health, Office of Disease Prevention. Engaging in Qualitative Research Methods: Opportunities for Prevention and Health Promotion. URL: https://prevention.nih.gov/education-training/methods-mind-gap/engaging-qualitative-research-methods-opportunities-prevention-and-health-promotion [accessed 2026-01-22]
34. HHRC indicators: HIV testing. VHA HIV, Hepatitis, and Related Conditions (HHRC) Team. 2024 Jan. URL: https://dvagov.sharepoint.com/sites/VHAOPHLDD/Shared%20Documents/Indicators-%20HIV%20Testing.aspx [accessed 2026-01-22]
35. Estimated HIV incidence and prevalence in the United States, 2018–2022. : Centers for Disease Control and Prevention; 2024 May URL: https://www.cdc.gov/hiv-data/nhss/estimated-hiv-incidence-and-prevalence.html [accessed 2026-01-22]
36. California HIV surveillance report – 2022. : California Department of Public Health, Center for Infectious Diseases, Office of AIDS; 2022 Feb 9 URL: https://www.cdph.ca.gov/Programs/CID/DOA/CDPH%20Document%20Library/California-HIV-Surveillance-Report-2022.pdf [accessed 2026-01-22]

XSL•FO
RenderX

37. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. HIV Prevention to End the HIV Epidemic in the United States – California. URL: https://www.cdc.gov/hivpartners/media/pdfs/2025/06/cdc_hiv_o3pc_factsheet_california_2025-508.pdf [accessed 2026-01-22]

Abbreviations

HIVST: HIV self-test
IRB: Institutional Review Board
QR: quick response
RE-AIM: Reach, Evaluation, Adoption, Implementation, and Maintenance
SFVAHCS: San Francisco VA Health Care System
VA: Veterans Affairs
VM: vending machine

Edited by J Sarvestan; submitted 17.Sep.2025; this is a non–peer-reviewed article;accepted 05.Jan.2026; published 29.Jan.2026.

Please cite as:
Rife-Pennington T, Douglas MP, Xie W, Cocohoba J
Use of Vending Machines to Deliver Oral Rapid HIV Self-Tests to Veterans: Protocol for a Pilot Study
JMIR Res Protoc 2026;15:e84317
URL: https://www.researchprotocols.org/2026/1/e84317
doi:10.2196/84317

© Tessa Rife-Pennington, Michael P Douglas, Wendy Xie, Jennifer Cocohoba. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 29.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Bedside Clinical Ultrasound Performed by Family Physicians in Adult Patients With Abdominal Pain in a Hospital Emergency Department: Protocol for a Pilot Quasi-Experimental Study

Laura Carbajo Martín[1,2,3], BChD; Ignacio Párraga-Martínez[4,5,6], BChD; Luis M Beltrán-Romero[7,8], BChD; Máximo Bernabeu Wittel[7,8], BChD; Northern Huelva Health Management Area Research Group[9]

[1]Emergency Department, Northern Huelva Health Management Area, Avda de la Esquila, Minas de Riotinto, Spain

[2]Emergency and Continuous Care Group of the Spanish Society of Family and Community Medicine, Barcelona, Spain

[3]Spanish Red Cross Nursing School, University of Seville, Sevilla, Spain

[4]Health Center Zone VIII Albacete, Integrated Care Management Albacete, Albacete, Spain

[5]Preventive Medicine and Public and Community Health, Faculty of Medicine of Albacete, University of Albacete, Albacete, Spain

[6]IDISCAM, Castilla-La Mancha Institute of Health Research, Albacete, Spain

[7]Internal Medicine Department, Virgen del Rocío University Hospital, Sevilla, Spain

[8]Department of Medicine, University of Seville, Sevilla, Spain

[9]See Acknowledgments

Corresponding Author:
Laura Carbajo Martín, BChD
Emergency Department, Northern Huelva Health Management Area, Avda de la Esquila, Minas de Riotinto, Spain

Abstract

Background: Point-of-care ultrasound is a valuable bedside tool that, with adequate training, can reduce diagnostic uncertainty and improve clinical accuracy. Abdominal pain is a frequent complaint in emergency departments and often requires imaging for appropriate management.

Objective: This study aims to assess the impact of bedside clinical ultrasound performed by family physicians on length of stay, number of basic radiological tests, and need for further diagnostic evaluations in adult patients with abdominal pain.

Methods: This is a pilot quasi-experimental study assessing feasibility and viability, with a nonrandomized control group, to be conducted in the Emergency Department of Hospital Comarcal de Riotinto. Adult patients (≥18 y) presenting with abdominal pain will be included. Both groups will receive standard care. In the intervention group, bedside ultrasound will be performed by trained family physicians; in the control group, ultrasound will be performed by radiologists only if deemed necessary. The primary outcome is the improvement in quality of care, assessed through a reduction in emergency department length of stay, fewer basic radiology tests requested, and diagnostic concordance. Secondary outcomes include the need for additional diagnostic studies and the appropriateness of referrals, evaluated through 1-month follow-up and reconsultation.

Results: The first phase of the project began in 2023 with the validation of the data collection form. Subsequently, the patient satisfaction questionnaire was validated, and the results were published in the journal *Care Primary*. The study has received external funding, and patient recruitment is currently ongoing and expected to be completed in December 2025.

Conclusions: This study aims to demonstrate the clinical and organizational benefits of implementing bedside ultrasound by family physicians in emergency care.

International Registered Report Identifier (IRRID): DERR1-10.2196/82393

(JMIR Res Protoc 2026;15:e82393) doi:10.2196/82393

KEYWORDS

abdominal pain; emergency; personal satisfaction; quality of health care; ultrasonography

Introduction

Ultrasound is a diagnostic technique based on the emission and reception of high-frequency waves, so-called ultrasound. It is a patient-friendly technique but requires an interpretation of possible artifacts, as well as progression through a learning curve. Ultrasound in the clinical setting complements physical examination in the same medical procedure, as it is accessible and fast and can be reproduced [1].

The use of ultrasound in emergencies began between 1991 and 1992 by surgeons in the United States, and later in 1996, the acronym FAST (Focused Abdominal Sonography for Trauma)

XSL•FO
RenderX

was adopted [2]. From then on, the term POCUS (point-of-care ultrasound) also became commonly used [3].

A review in 2019 showed that POCUS was used for a variety of conditions, although most focused on abdominal and obstetric indications with a good level of diagnostic accuracy [4]. Therefore, it seems appropriate to focus this study on abdominal pain, a common pathology encountered in the emergency department. In fact, a 2010 review indicated that the percentage of nonspecific pain was 24% - 44%, followed by acute appendicitis (15.9% - 28.1%), acute biliary disease (2.9% - 9.7%), and bowel obstruction or diverticulitis in older adults [5].

The pathology related to abdominal pain is highly varied, and training in clinical ultrasound requires a minimal learning curve. Consequently, a number of studies have assessed the training and education of professionals in the detection of biliary tract−related diseases. In this study [6], an acceptable level of agreement was found in the detection of hepatobiliary pathology among junior emergency medicine trainee doctors and radiologists after a short training program. In fact, ultrasound performed by emergency department clinicians affects the diagnostic certainty of abdominal pain in the right hypochondrium, particularly when requesting additional diagnostic studies or in patients with uncertain pathology [7]. The ability of emergency department physicians to diagnose cholelithiasis with bedside ultrasound has also been evaluated [8] and more recently also to rule out cholecystitis [9]. Furthermore, a recent review supports the use of POCUS for gallbladder examination by emergency physicians [10], and another review also suggests that this scan may help in arranging outpatient follow-ups if symptoms have disappeared [11].

In the diagnosis of nephrourological pathology such as renal colic, clinical ultrasound is a good diagnostic tool with a high level of accuracy in detecting urinary tract dilation [12]. These patients may also benefit from an improved level of risk stratification [13]. An observational study has already validated an algorithm for patients with flank pain in the examination and detection of urolithiasis [14]. The use of computed tomography for this pathology has increased, so a review was carried out to compare ultrasound as an alternative and safe test for the patient, which concluded that there is sufficient evidence to determine a more rational algorithm for the management of renal colic [15]. The use of POCUS was also found to be associated with a shorter emergency department stay compared to computed tomography [16].

In the emergency department, the length of stay is determined by many factors, such as the patient's age, the need for laboratory tests, or referral for consultation, but another factor to be taken into account is the need for ultrasound in abdominal pain [17]. In view of the limited time in the emergency department, ultrasound examination by family doctors is regarded as helping to improve the overall diagnosis and treatment of patients [18]. A recent study from October 2022 also analyzed the effect of POCUS in the initial assessment of patients with acute abdominal pain by evaluating the diagnosis and the decrease in patient waiting time without significant changes in the cost of the service [19].

Many of these patients presenting with abdominal pain at the emergency department require the performance of various complementary tests such as simple radiography. However, the diagnostic efficiency of abdominal radiography in urgent pathologies is not related to the high number of examinations performed as reported in this publication [20]. It is sometimes performed as a routine test despite recommendations to reduce radiation doses to patients and to select pathologies requiring radiography more carefully. This review has already indicated that the use of plain radiography in abdominal pain should be limited to very specific indications [21]. The overuse of these complementary tests leads to unnecessary exposure and increased waiting times in the emergency department. This has been supported by another systematic review [22] where 38 original studies were collected, and the conclusion was that, in most cases, their indiscriminate routine use is not indicated.

Finally, the assessment of the quality of health care is complex and highly subjective as it depends on both the care itself and the patient's expectations. This analysis of the patient's experience was carried out using a patient questionnaire, in which the patient's experience of receiving ultrasound scanning at the point of care was rated positively [23]. However, among the multidimensional models for the assessment of the quality of services provided are the SERVQUAL and SERVPERV questionnaires. The SERVPERF model has been used to assess the perception of quality of care in an emergency department [24], outpatient departments [25], hospital units [26], primary care [27], and even reassessment during the COVID-19 pandemic [28]. The score is calculated by the sum of the scores, and the quality of service is deemed to be as high as the total of the sum of the scores [29]. The SERVPERF model allows the level of quality to be measured by patient assessment.

Accordingly, considering that POCUS can assist professionals in the emergency department to exclude serious causes of abdominal pain−related pathologies, as well as to detect other pathologies that could be followed up at home, this study is proposed to assess the impact of its implementation in a hospital emergency department. Furthermore, the implementation of clinical ultrasound is intended to improve the quality of care and analyze the degree of patient satisfaction.

Methods

Study Design and Setting

A quasi-experimental pilot study of feasibility and viability with a control group will be carried out at Hospital Riotinto's Emergency Department. This center is located in the north of Huelva (Spain) and serves a population of almost 70,000 inhabitants with a high geographical distribution, as some localities are more than 80 km away from the hospital.

This study does not require registration in ClinicalTrials.gov as it does not meet the criteria for a clinical trial according to the International Committee of Medical Journal Editors and US Food and Drug Administration definitions. The project is a quasi-experimental feasibility and implementation study conducted in a real-world clinical setting, with no randomization, no experimental interventions, and no

XSL•FO
RenderX

administration of investigational drugs or devices. The study aims to evaluate the impact and utility of point-of-care ultrasound performed by family physicians in emergency departments as part of standard clinical practice. Therefore, it is considered a health services research study rather than a clinical trial, and formal registration is not applicable.

The implementation of bedside clinical ultrasound performed by family physicians in the emergency department is hypothesized to reduce patient length of stay as well as the need for additional diagnostic tests, such as plain radiography or specialist-performed ultrasound, thereby improving the overall quality of patient care. The primary objective is to analyze the impact of this intervention on length of stay, the number of simple radiology tests requested, and the necessity for further examinations by radiodiagnostic specialists.

Specific objectives include determining the length of stay for patients receiving clinical ultrasound by family physicians compared to usual care, quantifying and comparing the number of radiographs and ultrasounds performed, and analyzing patients' diagnostic and clinical progression over at least 1 month of follow-up. In cases where the process is not resolved, the patient can be monitored for the next 6 months. Additionally, the study aims to evaluate the number and outcomes of specialist-requested ultrasounds in relation to those performed by family physicians, identify factors influencing care quality related to patient health and sociodemographic characteristics, perform diagnostic concordance analyses, and assess patient satisfaction with the service provided.

The sample will consist of patients over the age of 18 years presenting with abdominal pain in the emergency department, who will be assigned to the experimental group (POCUS is performed) or the control group (no clinical ultrasound is performed) consecutively. Patients readmitted for the same pathology will be ruled out. Patients will be invited by the investigators to participate freely, and the recruitment will end when the minimum sample size is reached in each cohort.

The inclusion and exclusion criteria are presented in Textbox 1.

Textbox 1. The inclusion and exclusion criteria.

Inclusion criteria

- Patients treated at the Emergency Department of Riotinto Hospital during the study and follow-up period
- Patients presenting with acute de novo abdominal pain, defined as a clinical condition causing pain localized by the patient between the thorax and the pelvis. Additionally, patients must not have been admitted for the same condition in the previous 3 months.
- Age ≥18 years
- Patients with sufficient capacity to provide written informed consent
- Experimental group: Patients will be attended by physicians from the Emergency Department of Riotinto Hospital and by professionals with demonstrated training in clinical ultrasound
- Control group: Patients will be attended by physicians from the Emergency Department of Riotinto Hospital and by professionals without demonstrated training in clinical ultrasound

Exclusion criteria

- Patients with readmission due to abdominal pain
- Patients who leave the emergency department before completing care
- Patients followed in other regional health systems or private health care
- Pregnant women
- Patients with morbid obesity
- Patients with severe mental illness
- Patients whose clinical severity prevents obtaining informed consent
- Intercurrent illness that makes participation impossible

The reasons for study discontinuation are as follows: (1) patient withdrawal, (2) withdrawal of informed consent, (3) death, (4) protocol violation, and (5) intercurrent illness preventing continued participation.

Informed consent will be obtained by the attending physician at the time of patient evaluation, prior to any study-related procedures.

For the sample size calculation, emergency department length of stay was used as the primary variable, as it is the only parameter that has been studied to date (19). A 25% reduction in the mean length of stay is expected in the intervention group compared with the control group. According to the multicenter trial [30], the mean length of stay for patients with nontraumatic abdominal pain in the emergency department was 5.2 (SD 3) hours. Assuming a 25% reduction (1.3 h) in the intervention group, with a 2-sided significance level of 0.05 and a power of 80%, and applying a Student *t*-test for independent samples, a minimum of 84 patients per group is required. Considering an

XSL•FO
RenderX

anticipated 15% attrition rate, the total number of participants to be recruited will be 194 patients (97 per group).

There will be no randomization of the sample, as the 2 groups will receive the same care as normal for their pathology; however, in the experimental group, the intervention (clinical ultrasound) will also be carried out to test its effectiveness and assess whether it improves the quality of care. It is important to consider that clinical ultrasound is not a regulated ultrasound, but rather a response to a clinical question; therefore, it is not a complex intervention, and with the usual pressure of the service, it can be performed at any time of the year.

To achieve the target sample size, recruitment will be conducted consecutively among eligible patients presenting with abdominal pain at the Emergency Department. All family physicians participating in the study have been informed and trained to identify and invite potential participants. The collaboration of the entire emergency care team will be essential to ensure continuous enrollment during the study period. Regular follow-up meetings will be held to monitor recruitment progress and address potential barriers.

Prior to this, training will be provided to professionals, and a common understanding will be reached on the possible main findings. This is intended to reduce clinical variability and improve the external validity of the study.

Given the limited sample size, short duration, and minimal risk involved in this pilot quasi-experimental study, a formal independent data monitoring committee has not been established. The principal investigator and the co-investigators will be responsible for monitoring trial conduct, data accuracy, and participant safety. As no data monitoring committee is planned, further details regarding its charter are not applicable.

No interim analyses are planned for this study due to its exploratory nature and short timeframe. Consequently, no predefined stopping rules have been established. The final decision to complete or terminate the study will rest with the principal investigator, based on recruitment feasibility, protocol adherence, or unforeseen safety concerns.

To ensure data integrity, a quality control procedure has been designed, including predefined field validation rules within the case report form. Participating clinicians were trained during project initiation to minimize data entry errors. Data entry will follow the case report form sequence, and filters will prompt the confirmation of extreme or illogical values. The principal investigator will perform monthly audits, and a second investigator will periodically review the database or a sample of its records. Questionnaire variables will be coded numerically, and preliminary analyses will be performed to detect inconsistencies or missing data. All efforts will focus on minimizing missing data and correcting errors promptly to reduce potential biases and ensure the reliability of the findings.

All individual participant data will be deidentified and securely stored in the REDCap platform hosted by semFYC. Access to the data dictionary, statistical analysis code, and related materials will be available upon reasonable request to the corresponding author, in accordance with ethical and data protection regulations. These materials will be used exclusively for academic or research purposes.

Table 1 shows the summary with the most frequent coded findings in clinical ultrasound.

XSL•FO
RenderX

Table . Summary with the most frequent coded findings in clinical ultrasounds.

	Encoding	Image
Renal ultrasound		
Kidney cyst	E1	
Renal pelvis dilatation	E2	
Bladder polyp	E3	
Kidney stones	E4	
Other (specify)	E5	
Hepatobiliary ultrasound		
Cholelithiasis	H1	
Biliary sludge	H2	
Cholecystitis	H3	
Liver angioma	H4	
Metastasis	H5	
Liver cyst	H6	
Other (specify)	H7	
Abdominal trauma		
Abdominal free liquid	T1	
Other (specify)	T2	

The study protocol is represented in Figure 1.

XSL•FO
RenderX

Figure 1. Flowchart depicting study protocol. POCUS: point-of-care ultrasound.

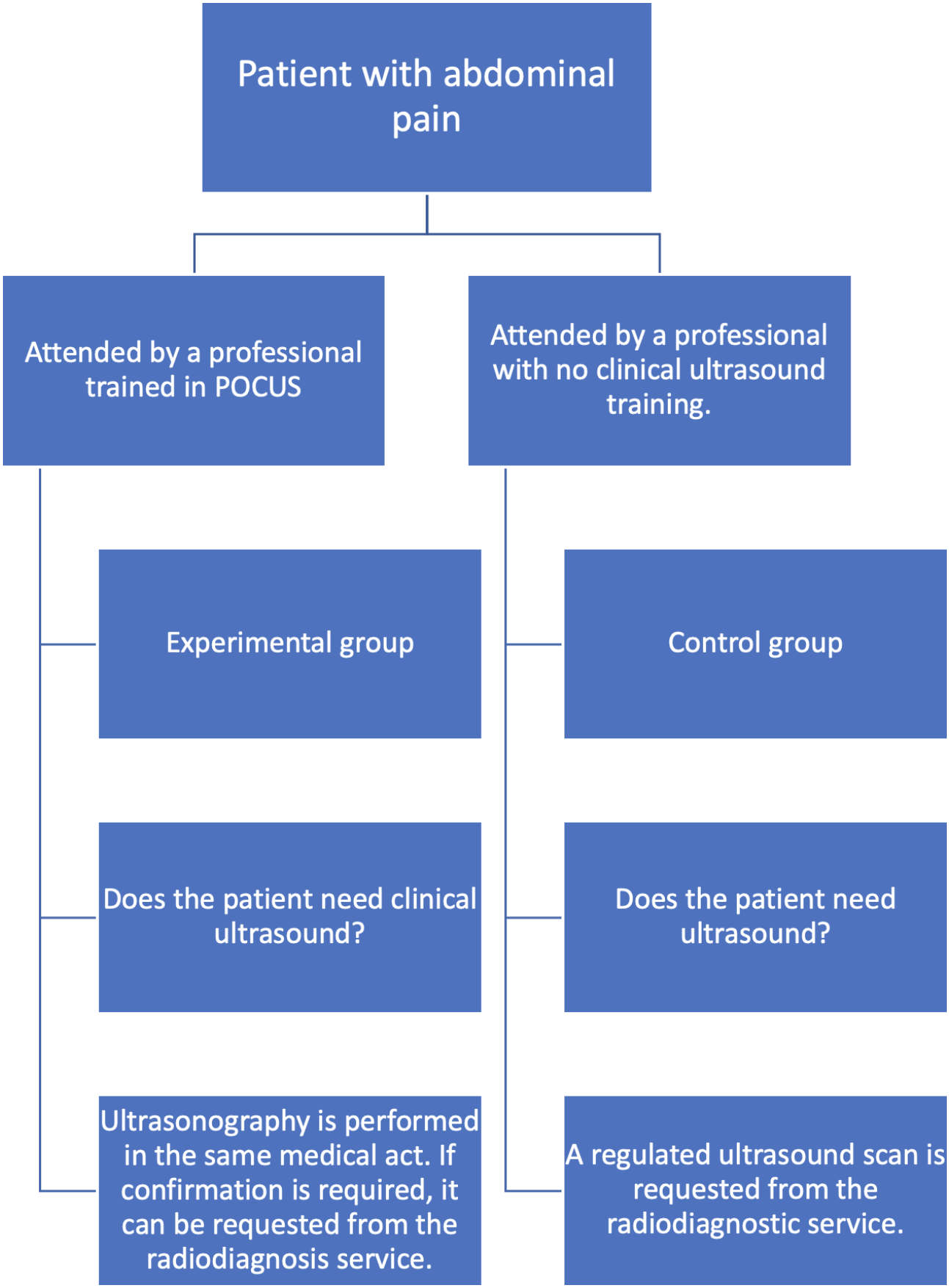

Subjects will be observed for a period of 1 month to determine whether they have returned for a repeat consultation for the same reason, and if so, to check the degree of alignment of the final diagnosis with the initial care. Only in cases where the process is not resolved can the patient be monitored for the next 6 months.

XSL•FO
RenderX

Statistical Analysis

A descriptive analysis will be carried out including proportions, measures of central tendency, and measures of dispersion, according to the nature of the variables (quantitative variables expressed as mean and SD and qualitative variables as a percentage). The analysis will include description and statistical analysis to identify any significant differences between men and women.

Comparison of the variables of interest and the stratification and potentially confounding variables at the start of the study will be carried out in both groups. The baseline values of the study variables will be tested for their homogeneity between the two groups (Student *t*-tests and chi-square tests, depending on the nature of the variables). There will also be a normality check for variables such as length of stay and those that require it.

The emergency department length of stay will be compared using a Student *t* test for independent samples, and a multiple linear regression model will be applied to adjust for potential confounding factors. The number of plain radiographs and complementary imaging studies in the Radiology Department will be analyzed using Poisson or logistic regression models, as appropriate. To account for multiple hypothesis testing, a Holm-Bonferroni correction will be applied.

The possible existence of potential confounding factors or the interaction of other variables in the relationship between the proposed intervention and the outcome variables will be assessed using logistic regression models. For continuous variables, it will be assessed by multiple linear regression.

In all cases, a statistical significance of 5% ($P<.05$) is required. The data will be processed and analyzed with SPSS (version 25.0; IBM Corp).

Variables

The main outcome variables aligned with the study objectives include the length of stay in the emergency department (measured in minutes), the number and type of radiological tests requested, the performance of bedside ultrasound (POCUS), diagnostic concordance, and the patient's satisfaction with the care received. Sociodemographic variables such as age, sex, and distance from home, as well as clinical and procedural variables, will also be collected.

A complete list of all study variables, including their measurement units or response categories, is provided in Table 2, grouped into thematic blocks: sociodemographic, clinical, diagnostic, and outcome variables.

XSL•FO
RenderX

Table . Study protocol variables.

Block and variable		Unit or format	Notes or response options
Sociodemographic			
	Identification code	Alphanumeric	Unique ID per patient
	Gender	Categorical	Male or female or other
	Date of birth	Date (dd/mm/yyyy)	Age will be calculated
	Distance to home	Kilometers	Estimated by postal code
Clinical			
	Health problems	Free text or ICD-10[a]	Comorbidities
	Current medications	Free text	List of active treatments
	Vital signs	Numeric (per parameter)	BP[b] (mm Hg), HR[c] (bpm), EVA[d], temperature (°C), $SatO_2$ (%)
	Reason for consultation	Free text	Chief complaint
Care process			
	Attending doctor	Code or initials	Family physician ID
	Date or time of arrival	Time stamp	Format: dd/mm/yyyy hh:mm
	Date or time of discharge	Time stamp	Format: dd/mm/yyyy hh:mm
	Origin	Categorical	Self-referred or primary care or ambulance or other
	POCUS[e] performed	Binary	Yes or no
	POCUS findings	Free text	Based on clinical protocol
	Other tests requested	Categorical or multiple	Blood tests or ECG[f] or other
	Radiology test performed	Categorical	X-ray or formal ultrasound or CT[g] or none
	Interconsultation	Binary	Yes or no; specialty if applicable
Outcome			
	Diagnosis	Free text or ICD-10	Final ED[h] diagnosis
	Discharge referral	Categorical	Home or hospitalization or specialist or follow-up
	Length of stay	Minutes	Calculated from entry and discharge times
	Satisfaction questionnaire	Likert scale or categorical	1 - 5 scale or qualitative responses
	Diagnostic concordance	Binary	Yes or no (based on follow-up or second opinion)

[a] *ICD-10: International Classification of Disease, Tenth Revision.*

[b]BP: blood pressure.

[c]HR: heart rate.

[d]EVA: Escala Visual Analógica (VAS: visual analog scale).

[e]POCUS: point-of-care ultrasound.

[f]ECG: electrocardiogram.

[g]CT: computed tomography.

[h]ED: emergency department.

Table 3 shows the results of the satisfaction questionnaire based on the SERVPERF model.

XSL•FO
RenderX

Table . Satisfaction questionnaire based on the SERVPERF model[a].

Dimensions and questions		Score range (strongly disagree to strongly agree)
Tangible elements		
1.	The facilities are suitable for patient care	1-7
1.	Technological equipment is adequate	1-7
1.	The emergency department has sufficient capacity to attend to the population	1-7
Responsiveness dimension		
1.	If you have any questions, you will receive a response within a reasonable time	1-7
1.	The attention of the staff provides a fast and high-quality service	1-7
1.	You have been dealt with quickly and promptly	1-7
Reliability dimension		
1.	When staff commit to doing something within a certain time frame, it is fulfilled	1-7
1.	If you have a problem, the staff show an interest in solving it	1-7
1.	Staff perform the service well the first time	1-7
1.	Professionals are ready to help when they are needed	1-7
Security dimension		
1.	Staff behavior imparts confidence	1-7
1.	There is a capacity to resolve doubts accurately	1-7
1.	Staff are knowledgeable when responding to queries	1-7
Empathy dimension		
1.	Care is offered that responds to the needs of the users	1-7
1.	Staff care about your interests and needs	1-7
1.	The emergency service provides a solution to the health needs of users	1-7
Access dimension		
1.	The waiting time for service was adequate	1-7
1.	Waiting time for tests and results was adequate	1-7
1.	The overall waiting time was adequate	1-7

[a]Likert scale (1-7) from least to most satisfaction.

Table 4 shows the work schedule for the study protocol.

XSL•FO
RenderX

Potential harms will be defined as any adverse event or unintended consequence resulting from the use of bedside ultrasound by nonradiologist physicians. Harms will be assessed systematically, using a structured reporting form completed by the attending physician and monitored through follow-up at 1 month. Any serious adverse events will be documented and reviewed by the principal investigator and the ethics committee if needed.

XSL•FO
RenderX

Table . Timeline of the study.

		Year 1												Year 2											
	Phase 0	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Ultrasound training	✓																								
Literature review update		✓																							
Presentation of the research project in the emergency department			✓																						
Protocol development				✓	✓																				
Design of the data collection form					✓	✓																			
Training and consensus on key pathologies							✓	✓																	

XSL•FO
RenderX

	Year 1						Year 2											
Data collection			✓	✓	✓	✓	✓	✓										
Data cleaning, analysis, and results generation									✓	✓	✓	✓	✓					
Results dissemination and manuscript preparation														✓	✓	✓	✓	✓

Ethical Considerations

This study has been approved by the Huelva Provincial Independent Ethics Committee (reference code PEIBA 1923-N-21). Any significant modifications to the protocol will be submitted for approval to the same ethics committee and promptly communicated to all relevant stakeholders, including participating investigators and, where necessary, trial registries.

Informed consent will be obtained by the attending family physicians or authorized research team members prior to the inclusion of any participant. Participants will receive both oral and written information about the study objectives, procedures, risks, and benefits and will be given the opportunity to ask questions before providing written consent. No biological specimens will be collected, and no ancillary studies requiring additional consent are planned. Participants will not receive any financial compensation or incentives for their participation in the study.

Personal data will be collected and stored using anonymized codes to ensure confidentiality. Access to identifiable information will be restricted to the research team and stored securely in password-protected systems in accordance with national data protection regulations. Data will be treated with strict confidentiality before, during, and after the trial, and only aggregated, deidentified data will be reported in publications or presentations.

Given the minimal-risk nature of this study, no specific post-trial care provisions are foreseen. However, if any harm were to occur as a result of participation, appropriate medical care and any necessary support will be provided by the institution in accordance with applicable regulations and institutional policies. The application procedures and the study were approved by the Huelva Provincial Ethics Committee (Comité de Ética de Investigación de Huelva) (Code PEIBA 1923-M1-21).

Results

The first phase of the project began in 2023 with the validation of the data collection form. Subsequently, the patient satisfaction questionnaire was validated, and the results were published in the journal Care Primary. The study has received external funding, and patient recruitment is currently ongoing and expected to be completed in December 2025.

The project was initiated in 2023. Between March and July 2023, an initial validation phase of the data collection form was conducted. Preliminary findings were presented as a clinical communication at the 44th semFYC National Congress, held in Barcelona, 14 - 16 November 2024. A total of 65 patients were included in this first sample, with a mean age of 56.6 (SD 20.0) years, of whom 53.8% (n=35) were women and 46.2% (n=30) were men. The validation of the data collection form showed that 76.9% (n=50) of the ultrasound examinations were performed as point-of-care clinical ultrasounds (POCUS), whereas only 6.2% (n=4) were formal imaging studies referred to the Radiology Department.

Based on this pilot experience, the study protocol was further developed. Funding was subsequently obtained through the 2024 Call of the Fundación Progreso y Salud (Andalusian Public Health System) under the Research and Innovation Projects

XSL•FO
RenderX

category in Primary Care, Regional Hospitals, and High-Resolution Healthcare Centers (grant ID: AP-0561 - 2024-C5-F2).

Further work focused on refining the study protocol and validating the patient satisfaction questionnaire was published in [31].

Patient recruitment is currently ongoing and is expected to be completed by December 2025. We anticipate that the implementation of POCUS in patients presenting with abdominal pain in emergency settings will lead to a reduction in length of stay and a decrease in the use of additional diagnostic tests, such as plain radiography or formal radiology−performed ultrasound.

Discussion

Principal Findings

POCUS has emerged as a valuable tool in emergency medicine, particularly for its potential to reduce diagnostic uncertainty and enhance clinical accuracy in decision-making. Despite these recognized advantages, there is limited evidence regarding its overall impact on the quality of care, and even less from the patient's perspective. This study protocol is based on the hypothesis that the implementation of clinical ultrasound performed by family physicians in emergency departments can reduce length of stay and decrease the need for complementary diagnostic tests, such as plain radiography or radiology-performed ultrasound. Collectively, these improvements may lead to a measurable enhancement in both perceived and actual quality of care.

The geographic context of this study adds a relevant dimension to the analysis. The target population resides in an area characterized by wide territorial dispersion and significant population aging, factors that place additional demands on emergency services. In such settings, clinical practice often prioritizes the resolution of health problems in a single visit to avoid repeated travel—especially for patients with reduced mobility or limited access to transportation. POCUS performed by emergency physicians may therefore not only optimize diagnostic workflows but also align more closely with the health care needs of the community.

One of the main methodological challenges in this study is interoperator variability in the interpretation of ultrasound findings. To minimize this, a reference image repository has been developed, alongside simplified visual guides to standardize the use of clinical ultrasound in patients presenting with abdominal pain. In addition, prestudy training sessions were conducted in collaboration with the radiology department to promote consistency in image acquisition and interpretation among participating physicians.

Beyond evaluating diagnostic efficiency, the study also aims to identify the key factors contributing to improved quality of care from the health care professional's perspective. This approach will provide both qualitative and quantitative insights into the clinical, organizational, and human elements that influence better health outcomes.

Finally, this research initiative may pave the way for broader applications of clinical ultrasound beyond hospital-based emergency settings, including prehospital care and primary health care services. If the results are favorable, they could support the development and implementation of standardized protocols for the use of POCUS by family physicians, with potential implications for health care system efficiency and patient-centered outcomes.

Conclusion

This study aims to assess whether the implementation of POCUS performed by family doctors in the emergency department reduces the length of stay in the department, as well as other diagnostic tests such as simple radiology or regulated ultrasound, improving the quality of patient care.

Acknowledgments

The authors would like to acknowledge the technical and methodological support provided by the Fundación Andaluza Beturia para la Investigación en Salud, especially Alina Rigabert, for her guidance and valuable contributions throughout the development of the study. We also extend our sincere thanks to the entire Emergency Department team at Hospital Comarcal de Riotinto for their collaboration, involvement, and commitment to the design and implementation of this research. Their support has been essential in making this study possible. Generative artificial intelligence tools, including ChatGPT (OpenAI), were used to improve the language and clarity of the English text during manuscript preparation. The authors reviewed, verified, and take full responsibility for all content and conclusions.

Members of the Northern Huelva Health Management Area Research Group are as follows: Aguas Santas Benítez La Fuente, Ana María Benjumeda González, Julio César Caballero Morgado, José Antonio Cabañas Arija, Laura Carbajo Martín, María del Carmen Castaño Durán, Julia Cristina Chávez Sánchez, Esperanza de Cos Maestre, Marta Delgado Moya, Celia Domínguez Rite, Selene Fernández Burgos, Ester Martín Brioso, Inmaculada Martín Santos, Begoña Medina García, José Luis Merchán Geniz, Luis Núñez Vaquero, Claudia Ollado Darriba, Esperanza Macarena Ortega Hidalgo, Laura Quirante Pérez, Juan Jesús Rosado Cabral, Rocío Romero Cortés, Juan Luis Rosario Ventura, María José Rosende Domínguez, Gonzalo Tena Santana, Jacqueline Trueba Carreón, and Noelia María Trujillo Díaz.

XSL•FO
RenderX

Funding

Project was funded in the call for Research and Innovation Projects in the field of Primary Care and Regional Hospitals and High-Resolution Hospitals of the Public Health System of Andalusia. Code: AP-0561-2024-C5-F Call 2024. This work has been carried out with the support of the Andalusian Society of Family and Community Medicine Foundation – SAMFyC (File Ref. TD 235/24). This study was funded by the Fundación Progreso y Salud (Andalusia, Spain) under the 2024 Research and Innovation Call in Primary Care and Regional Hospitals (project code: AP-0561-2024-C5-F2). The funding body had no role in the study design, data collection, analysis, interpretation, or manuscript preparation.

Data Availability

Data collection for this study is currently ongoing. Preliminary datasets have already been generated and stored securely in compliance with institutional and ethical regulations. Upon the completion of data collection and analysis, all deidentified datasets supporting the findings of this study will be made available without access restrictions.

The datasets will be available from the corresponding author upon reasonable request and will also be deposited in a publicly accessible data repository once the final analyses are completed.

The deidentified dataset and statistical code will be made available under a CC BY-NC 4.0 license.

All data used in the research will be available without access restrictions.

We are committed to disseminating the trial results to all relevant stakeholders, including study participants, health care professionals, the general public, and other interested parties. Dissemination will be conducted through multiple channels such as updating the trial registry with results, providing a plain language summary accessible to nonspecialist audiences, and publishing findings in peer-reviewed scientific journals.

The datasets generated or analyzed during the current study are available from the corresponding author on reasonable request. The datasets will be available from the corresponding author upon reasonable request and will also be deposited in a publicly accessible data repository once the final analyses are completed.

Authors' Contributions

Conceptualization: LCM.
Data curation: LCM.
Formal analysis: LCM, MBW.
Funding acquisition: LCM.
Investigation: IPM, LCM, LMBR, MBW, NHHMARG.
Methodology: IPM, LCM, LMBR, MBW.
Project administration: LCM.
Resources: IPM, LCM, LMBR, MBW.
Supervision: IPM, MBW.
Validation: IPM, LMBR, MBW.
Visualization: MBW.
Writing – original draft: LCM.
Writing – review & editing: IPM, LCM, LMBR, MBW.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[DOCX File, 23 KB - resprot_v15i1e82393_app1.docx]

Peer Review Report 1
Peer review report by the 2024 Call for R&I Projects in Primary Care, Regional Hospitals, and CHARES Review Committee, Fundación Andaluza Beturia para la Investigación en Salud (FABIS; Beturia Andalusian Foundation for Health Research, Spain.
[PDF File, 13 KB - resprot_v15i1e82393_app2.pdf]

References

1. Kameda T, Taniguchi N. Overview of point-of-care abdominal ultrasound in emergency and critical care. J Intensive Care 2016;4:53. [doi: 10.1186/s40560-016-0175-y] [Medline: 27529029]

XSL•FO
RenderX

2. De Cabo Francés FM. Utilidad de la ecografía en la emergencia. Prehosp Emerg Care 2008 Jul;1(3):295-300 [FREE Full text]
3. Choi W, Cho YS, Ha YR, et al. Role of point-of-care ultrasound in critical care and emergency medicine: update and future perspective. Clin Exp Emerg Med 2023 Dec;10(4):363-381. [doi: 10.15441/ceem.23.101] [Medline: 38225778]
4. Andersen CA, Holden S, Vela J, Rathleff MS, Jensen MB. Point-of-care ultrasound in general practice: a systematic review. Ann Fam Med 2019 Jan;17(1):61-69. [doi: 10.1370/afm.2330] [Medline: 30670398]
5. Grundmann RT, Petersen M, Lippert H, Meyer F. Das akute (chirurgische) Abdomen – Epidemiologie, Diagnostik und allgemeine Prinzipien des Managements. Z Gastroenterol 2010 Jun;48(6):696-706. [doi: 10.1055/s-0029-1245303]
6. Seyedhosseini J, Nasrelari A, Mohammadrezaei N, Karimialavijeh E. Inter-rater agreement between trained emergency medicine residents and radiologists in the examination of gallbladder and common bile duct by ultrasonography. Emerg Radiol 2017 Apr;24(2):171-176. [doi: 10.1007/s10140-016-1468-0] [Medline: 27878404]
7. Bektas F, Eken C, Soyuncu S, Kusoglu L, Cete Y. Contribution of goal-directed ultrasonography to clinical decision-making for emergency physicians. Emerg Med J 2009 Mar;26(3):169-172. [doi: 10.1136/emj.2008.059220] [Medline: 19234003]
8. Cox C, MacDonald S, Henneberry R, Atkinson PR. My patient has abdominal and flank pain: identifying renal causes. Ultrasound 2015 Nov;23(4):242-250. [doi: 10.1177/1742271X15601617] [Medline: 27433264]
9. Sharif S, Vlahaki D, Skitch S, et al. Evaluating the diagnostic accuracy of point-of-care ultrasound for cholelithiasis and cholecystitis in a Canadian emergency department. CJEM 2021 Sep;23(5):626-630. [doi: 10.1007/s43678-020-00068-6] [Medline: 34491557]
10. Dupriez F, Geukens P, Penaloza A, Vanpee D, Bekkering G, Bobbia X. Agreement of emergency physician-performed ultrasound versus RADiology-performed UltraSound for cholelithiasis or cholecystitis: a systematic review. Eur J Emerg Med 2021 Oct 1;28(5):344-351. [doi: 10.1097/MEJ.0000000000000815] [Medline: 33758146]
11. Ross M, Brown M, McLaughlin K, et al. Emergency physician-performed ultrasound to diagnose cholelithiasis: a systematic review. Acad Emerg Med 2011 Mar;18(3):227-235. [doi: 10.1111/j.1553-2712.2011.01012.x] [Medline: 21401784]
12. Bourcier JE, Gallard E, Redonnet JP, et al. Ultrasound at the patient's bedside for the diagnosis and prognostication of a renal colic. Ultrasound J 2021 Nov 22;13(1):45. [doi: 10.1186/s13089-021-00246-2] [Medline: 34807330]
13. Lee WF, Goh SJ, Lee B, Juan SJ, Asinas-Tan M, Lim BL. Renal point-of-care ultrasound performed by ED staff with limited training and 30-day outcomes in patients with renal colic. Can J Emerg Med 2024 Mar;26(3):198-203. [doi: 10.1007/s43678-023-00645-5] [Medline: 38219263]
14. Kartal M, Eray O, Erdogru T, Yilmaz S. Prospective validation of a current algorithm including bedside US performed by emergency physicians for patients with acute flank pain suspected for renal colic. Emerg Med J 2006 May;23(5):341-344. [doi: 10.1136/emj.2005.028589] [Medline: 16627832]
15. Dalziel PJ, Noble VE. Bedside ultrasound and the assessment of renal colic: a review. Emerg Med J 2013 Jan;30(1):3-8. [doi: 10.1136/emermed-2012-201375] [Medline: 22685250]
16. Orosco E, Terai H, Lotterman S, et al. Point-of-care ultrasound associated with shorter length of stay than computed tomography for renal colic. Am J Emerg Med 2024 May;79:167-171. [doi: 10.1016/j.ajem.2024.02.024] [Medline: 38452429]
17. Dadeh AA, Phunyanantakorn P. Factors affecting length of stay in the emergency department in patients who presented with abdominal pain. Emerg Med Int 2020;2020:5406516. [doi: 10.1155/2020/5406516] [Medline: 32566303]
18. Abdolrazaghnejad A, Rajabpour-Sanati A, Rastegari-Najafabadi H, Ziaei M, Pakniyat A. The role of ultrasonography in patients referring to the emergency department with acute abdominal pain. Adv J Emerg Med 2019;3(4):e43. [doi: 10.22114/ajem.v0i0.152] [Medline: 31633098]
19. Durgun Y, Yurumez Y, Guner NG, Aslan N, Durmus E, Kahraman Y. Abdominal pain management and point-of-care ultrasound in the emergency department: a randomised, prospective, controlled study. J Coll Physicians Surg Pak 2022 Oct;32(10):1260-1265. [doi: 10.29271/jcpsp.2022.10.1260] [Medline: 36205268]
20. Martín JMA, de Gracia MM, Torres CR, Martínez DM, Herranz PP. Radiografía del abdomen en Urgencias. ¿Una exploración para el recuerdo? Radiología 2015 Sep;57(5):380-390. [doi: 10.1016/j.rx.2015.06.002]
21. Wolfe C, Halsey-Nichols M, Ritter K, McCoin N. Abdominal pain in the emergency department: how to select the correct imaging for diagnosis. Open Access Emerg Med 2022;14:335-345. [doi: 10.2147/OAEM.S342724] [Medline: 35899220]
22. Smith JE, Hall EJ. The use of plain abdominal x rays in the emergency department. Emerg Med J 2009 Mar;26(3):160-163. [doi: 10.1136/emj.2008.059113] [Medline: 19234001]
23. Andersen CA, Brodersen J, Rudbæk TR, Jensen MB. Patients' experiences of the use of point-of-care ultrasound in general practice - a cross-sectional study. BMC Fam Pract 2021 Jun 18;22(1):116. [doi: 10.1186/s12875-021-01459-z] [Medline: 34144701]
24. Torres Navarro C, Jélvez Caamaño A, Vega Pinto F. Evaluación de la calidad de un servicio de urgencia utilizando el modelo servperf. Universidad, Ciencia y Tecnología 2014;18:38-49 [FREE Full text]
25. Campoverde RE, Baldeón M, González V, Montero M. Calidad de servicios médicos ambulatorios: un análisis confirmatorio del modelo SERVPERF. Agosto 2020;41(31):33-45 [FREE Full text]
26. Duc Thanh N, Quynh Anh P, Thi Huyen Chang P, Minh Nguyet HT. Cross-cultural adaption and validation of SERVPERF tool for measuring healthcare quality in an oncology public hospital, Vietnam. Inquiry 2023;60:469580221146826. [doi: 10.1177/00469580221146826] [Medline: 36680357]

XSL•FO
RenderX

27. Gómez-Besteiro MI, Somoza-Digón J, Einöder-Moreno M, Jiménez-Fraga MN, Castiñeira-Pereira S, Feijoo-Fuentes ML. Calidad de atención en salud percibida por los pacientes en un centro de salud urbano de A Coruña. Enferm Clín 2012 Jul;22(4):182-190. [doi: 10.1016/j.enfcli.2012.04.004]
28. Babroudi NEP, Sabri-Laghaie K, Ghoushchi NG. Re-evaluation of the healthcare service quality criteria for the Covid-19 pandemic: Z-number fuzzy cognitive map. Appl Soft Comput 2021 Nov;112:107775. [doi: 10.1016/j.asoc.2021.107775] [Medline: 34377110]
29. Ibarra Morales LE, Casas Medina EV. Aplicación del modelo Servperf en los centros de atención Telcel, Hermosillo: una medición de la calidad en el servicio. Contad Adm 2015;60:229-260 [FREE Full text]
30. Slagman A, Schenk L, Huscher D, et al. Management of non-traumatic abdominal pain in the emergency department: a multicentre, stepped-wedge, cluster-randomised trial. Lancet Reg Health Eur 2025 Aug;55:101362. [doi: 10.1016/j.lanepe.2025.101362] [Medline: 40678036]
31. Carbajo Martín L, Párraga-Martínez I, Rosado Cabral JJ, Beltrán Romero LM, Bernabeu Wittel M, Grupo de Investigación AGS Norte de Huelva. Hospital de Riotinto. Validation of the SERVPERF tool for quality analysis in emergency care with clinical ultrasound. Aten Primaria 2025 Jun;57(6):103204. [doi: 10.1016/j.aprim.2024.103204] [Medline: 39731851]

Abbreviations

POCUS: point-of-care ultrasound

Edited by J Sarvestan; submitted 14.Aug.2025; this is a non–peer-reviewed article;accepted 17.Nov.2025; published 02.Jan.2026.

Please cite as:
Carbajo Martín L, Párraga-Martínez I, Beltrán-Romero LM, Bernabeu Wittel M, Northern Huelva Health Management Area Research Group
Bedside Clinical Ultrasound Performed by Family Physicians in Adult Patients With Abdominal Pain in a Hospital Emergency Department: Protocol for a Pilot Quasi-Experimental Study
JMIR Res Protoc 2026;15:e82393
URL: https://www.researchprotocols.org/2026/1/e82393
doi:10.2196/82393

© Laura Carbajo Martín, Ignacio Párraga-Martínez Sr, Luis Matías Beltrán Romero, Máximo Bernabeu Wittel, Northern Huelva Health Management Area Research Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Biopsychosocial Determinants, Diet Quality, Gastrointestinal Health, and Disease Activity in Adults With Rheumatoid Arthritis: Cross-Sectional Descriptive Study

Maureen McGarrity-Yoder[1], RN, PhD; Emily K Cope[2], PhD; Heidi A Wayment[3], PhD; Alicia Rodriguez-Pla[4], MPH, MD, PhD; Tracy E Crane[5], RDN, PhD

[1]College of Nursing, Northern Arizona University, Flagstaff, United States

[2]Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ, United States

[3]Department of Psychological Sciences, Northern Arizona University, Flagstaff, AZ, United States

[4]Sierra Pacific Arthritis and Rheumatology Centers, Fresno, CA, United States

[5]Department of Medical Oncology, Miller School of Medicine, University of Miami, Miami, FL, United States

Corresponding Author:
Maureen McGarrity-Yoder, RN, PhD
College of Nursing
Northern Arizona University
202 E. Pine Knoll Dr.
Flagstaff, 86011
United States
Phone: 1 523 8936
Email: maureen.yoder@nau.edu

Abstract

Background: Rheumatoid arthritis (RA) causes pain, fatigue, joint deformity, disability, and an increased risk for serious sequelae, often despite treatment, in 1.3 million Americans. RA is affected by numerous biopsychosocial determinants, which greatly complicate treatment, including altered efficacy.

Objective: The purpose of this study is to examine associations between individual biopsychosocial determinants, diet quality, gastrointestinal (GI) health, and disease activity in adults with RA.

Methods: This cross-sectional, descriptive study has been approved by the Northern Arizona University Internal Review Board (# 2111208-12). We will include 96 adults with RA recruited from across Arizona using social media and community events (through the Arthritis Foundation) and various primary care and rheumatology practices in Flagstaff and the greater Phoenix metro area. Individual biopsychosocial factors will be measured with a demographic survey and direct measures. The Arizona Food Frequency Questionnaire will measure dietary intake for the past 6 months, and Healthy Eating Index-2020 scores will be calculated from these data. The Automated Self-Administered 24-hour diet recall will measure recent dietary intake. Fecal analyses for gut microbiome diversity and composition and fecal calprotectin will measure current GI health. Disease activity will be measured by the Health Assessment Questionnaire–Disability Index and pain scale, Disease Activity Score of 28 Joints, and hematology results (C-reactive protein and erythrocyte sedimentation rate). In addition to descriptive statistics, hierarchical linear regression will examine hypothesized associations between diet quality, GI health, and disease activity. We hypothesize that individual biopsychosocial determinants will be associated with diet quality, which will be indirectly associated with disease activity through gut microbiome diversity and level of GI inflammation in adults with RA.

Results: This study was funded in February 2024. As of December 19, 2025, a total of 80 individuals have been recruited. Data analysis has not yet commenced at the time of manuscript submission. Study results are expected to be published in fall 2026.

Conclusions: RA is a complicated disease that impacts millions. Few individuals reach sustained remission, even while following provider recommendations. A better understanding of the various factors that impact this complicated disease has the potential to support changes in research and care that will improve the lives of people with RA. The knowledge gained in this study will provide a foundation to inform future interventional research targeting diet quality to support GI health and decrease RA disease activity. Further, the details of this research plan provide methodological resources for other RA researchers, and research results have the potential to improve communication between rheumatology providers and patients.

International Registered Report Identifier (IRRID): PRR1-10.2196/79889

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e79889) doi:10.2196/79889

KEYWORDS

diet quality; gastrointestinal health; gut microbiome; inflammation; rheumatoid arthritis

Introduction

Background

Rheumatoid arthritis (RA) is a debilitating autoimmune disease that affects approximately 1.3 million Americans [1]. RA is characterized by intermittent, chronic inflammation that causes decreased physical function [2]. Treatment typically focuses on pharmacologic management to reach the goal of remission, but, unfortunately, many patients do not reach despite following provider recommendations [3]. Increased RA disease activity can result in decreased quality of life due to chronic pain, fatigue, and disability [4], impacting family, social, and work life [5]. Prolonged inflammation often results in bone and joint damage, deformity, altered psychological and social function, and an increased risk of serious sequalae such as hypertension, depression, and cardiovascular disease, again, often despite treatment [1,4,6,7]. RA also causes a high financial burden due to excessive medical costs, disability, and a decreased ability to work [4].

The exact etiology of RA remains uncertain, but it is understood that an underlying genetic risk is activated in the presence of an environmental stimulus [8]. The activation and progression of RA, characterized by intermittent local and systemic inflammation, ultimately results in joint deformity and bone breakdown if the disease is not well managed [8]. Localized RA inflammation includes redness, warmth, decreased function, pain, and edema of affected joints [9]. Systemic inflammation, measured with hematologic proinflammatory biomarkers [9], is often associated with neurological manifestations through the bidirectional neuroimmune loop. These manifestations are categorized as “sickness behavior”: fatigue, malaise, nausea, and withdrawn behavior [7,10,11]. Systemic inflammation in RA is also associated with poor sleep quality [12], mood changes and depression [13], and an increased risk of several comorbidities including mental health disorders (eg, anxiety and depression), interstitial lung disease, hypertension, and cardiovascular disease [7,8]. The objective of this funded project is to examine associations between individual biopsychosocial determinants, diet quality, gastrointestinal (GI) health, and disease activity in adults with RA. The long-term goal of this research is to improve rates of sustained remission to allow better health outcomes and quality of life for all diagnosed with RA.

The Biopsychosocial Model of Disease Activity in RA

RA inflammation is impacted by various biological, psychological, and sociological factors. The evidence-based Biopsychosocial Model of Disease Experience in Rheumatoid Arthritis (BDRA) (Figure 1) considers these determinants of disease activity in RA. The BDRA, developed by the primary author [14], combines concepts from Engel’s Biopsychosocial Model of Health [15] and the Revised Symptoms Management Conceptual Model [16] to demonstrate these influential relationships in RA to better inform research and patient care. This model serves as the theoretical underpinnings of the proposed project and considers determinants individual impact; their effects on one another; and their impacts on neurological, immunological, endocrine, and GI function; and, finally, disease activity and experience in RA.

Figure 1. Biopsychosocial Model of Disease Experience in Rheumatoid Arthritis (BDRA) [1].

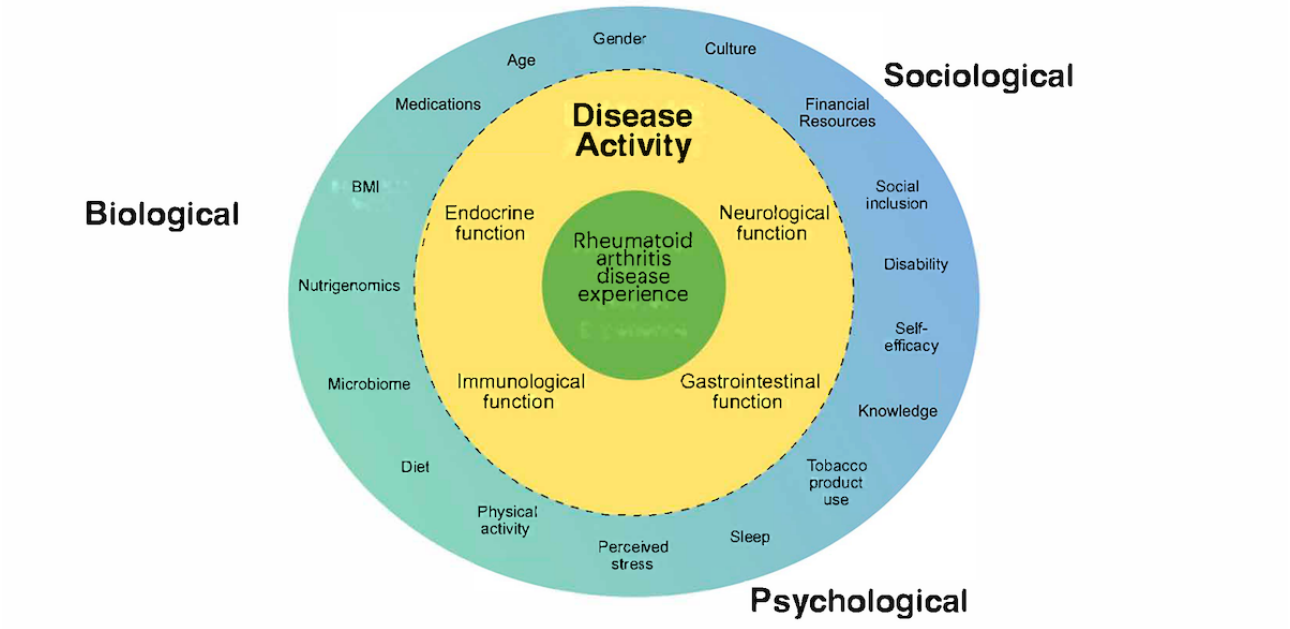

XSL•FO
RenderX

Diet, Gut Microbiome and Diversity, GI Inflammation, and RA Disease Activity

There is significant patient interest in how diet may be related to disease activity in RA [17]. Although a few studies have reported associations between dietary intake and generalized inflammation, these associations are unclear, possibly due to the unreliable nature of self-reported dietary intake [18,19]. The ambiguity in the literature may also contribute to inadequate communication about diet recommendations between rheumatology providers and patients [20]. To date, researchers have yet to consider the effect of biological data to explain associations between dietary intake and RA disease activity. GI health (gut microbiome composition and diversity and GI inflammation) may be mechanistic underpinnings to explain these potentially important associations. Diet is known to impact the composition of the gut microbiome, and individuals with RA are more likely to exhibit gut dysbiosis, which is an imbalance of gut microbiota [21-23]. Gut dysbiosis has also been associated with RA development [24]. Fecal calprotectin, a protein released by neutrophils in the presence of inflammation, has been associated with various autoimmune diseases, including spondyloarthritis [25]. Further, increased serum calprotectin levels are associated with increased disease activity in RA [26].

Biopsychosocial Factors

Overview

Background biopsychosocial factors known to be associated with RA disease activity include gender, age, socioeconomic status (education and household income), ethnicity, perceived stress, BMI, smoking, and medication use. Women are 3 times more likely to develop RA and have higher disability rates [27,28]. Functional disability and poor outcomes increase with age [28-30]. Worsened health outcomes in RA have been associated with poverty, decreased education, and a lack of basic health care services [31]. RA rates are significantly higher in minority populations [32]. Perceived stress has significant biological effects on disease activity through psychoneuroimmunological mechanisms [7,33]. Increased BMI and smoking are associated with higher rates, worse disease activity, and poor outcomes in RA [34,35]. RA medications disrupt various inflammatory pathways of the immune system [36].

Socioeconomic Status

Lower socioeconomic status, underserved communities, racial and ethnic minority groups, and sexual and gender minority groups are at risk for less health care access according to the National Institutes of Health (NIH) [37]. As with other disease processes, these individuals are at risk for poor outcomes in RA. Worsened RA outcomes [31] and decreased rates of remaining under rheumatology care have been associated with poverty [38,39]. Individuals with RA are at risk for social exclusion, which is often worsened with low socioeconomic status [31,40]. Increased disease activity in RA can further increase financial burden due to high medical costs and a decreased ability to work [4,31].

Gender

Women are 3 times more likely than men to be diagnosed with RA [27] and report worse pain and fatigue and exhibit greater evidence of disease activity [41]. Women are more likely to experience decreased productivity and missed work [42] and have higher rates of long-term functional disability [28].

Age

Functional disability [28] and rates of serious sequelae increase with age [30]. Further, increased rates of RA diagnosis occur in older adults [30].

Ethnicity

Individuals of color are noted to have higher rates of RA diagnosis [32]. Additionally, individuals of color also experience higher rates of poor outcomes [32,43].

BMI

Obesity is significantly associated with increased rates of RA [34]. Obesity also results in inflammation, worse disease activity, and poorer clinical outcomes [44].

Smoking

Cigarette smoking increases the risk for RA [45] and is associated with worsened disease severity, increased joint damage, and risk for sequalae (eg, malignancies, atherosclerosis, and vasculitis in RA) [35].

Medications

Medications used to treat RA include nonsteroidal anti-inflammatories, corticosteroids, and biologics (conventional, biologics, and Janus Kinase-inhibitors). They decrease inflammation by disrupting various inflammatory pathways associated with RA [36].

Potential Biopsychosocial Factors: Diet and GI Health

GI health may be the bridge needed to clarify potential associations. Diet is known to impact the composition of the gut microbiome, and individuals with RA are more likely to exhibit gut dysbiosis [21-23]. Further, gut dysbiosis has been associated with RA development [24]. Fecal calprotectin, an indicator of GI inflammation, has been associated with disease activity in various autoimmune diseases, including spondyloarthritis [25], and increased serum calprotectin levels were associated with increased disease activity in RA [26].

Diet

Dietary metabolites can directly affect cellular production of proinflammatory and anti-inflammatory cytokines through signaling pathways [46]. Proinflammatory signaling pathways are directly stimulated by metabolites of refined grains and omega-6 polyunsaturated fatty acids and directly inhibited by metabolites of antioxidants and omega-3 polyunsaturated fatty acids [46]. Cancer, hypertension, heart disease, and type 2 diabetes are associated with a typical Western diet, characterized by high sugar, fat, protein, and salt content [8]. Cyclic citrullinated peptide antibody, a key biological mediator in RA, has been found to react with dietary triggers (eg, food and microbial proteins) and induce RA symptoms [47].

XSL•FO
RenderX

The GI Tract and Immune Function

The GI tract plays a significant role in immune function. Approximately 70% of lymphocytes in the human body are located within the GI tract [9]. Symbiotic and pathogenic organisms, located throughout the GI tract cause an almost constant potential for disruption of gut homeostasis [48]. The heavy presence of immune cells assists in protection against toxins and pathogens [48]. The GI tract further influences immune function through the brain–gut–microbiome axis, impacting the inflammatory response [48].

The Gut Microbiome

The gut microbiome is a key component of health and disease and significantly impacts immune function [49]. Patients with RA were found to have a significantly different composition and associated dysbiosis of gut microbiota compared with healthy controls [50]. This included an increased quantity of *Prevotella*, which is generally associated with GI mucosal inflammation and an increased production of proinflammatory cytokines [49]. Individuals in preclinical stages of RA were found to have an abundance of *Prevotella* spp, supporting the involvement of intestinal dysbiosis in the development of RA [51]. Patients with RA were found to have increased *Alloprevotella*, *Parabacteroides*, and *Prevotella-2*, which were previously associated with increased systemic inflammation in other diagnoses [24]. Interestingly, disease-modifying antirheumatic medications partially restore eubiosis, or a normal balance of gut microbiome composition, in patients with RA [52]. Lastly, gut microbiome dysbiosis and associated GI tract damage may result in passage of inappropriate molecules and microorganisms, referred to as a "leaky gut" [53]. This may be a mechanism through which dietary intake alters disease activity of various diseases, including RA. Interestingly, DNA from gut bacteria has been found in synovial fluid of patients with RA and healthy control patients, although in differing amounts based on RA diagnosis [54]. This offers preliminary evidence that the gut microbiome may impact the synovial microbiome and disease activity in RA.

GI Tract

GI inflammation is closely associated with gut microbiome balance. Calprotectin is a damage-associated molecular pattern protein released by neutrophils in the presence of inflammation. It is measurable in both serum and fecal samples. Fecal calprotectin indicates acute inflammation of the intestines and is significantly associated with other measures of inflammation in inflammatory bowel disease and spondyloarthritis [25]. Although fecal calprotectin has not been compared to disease activity in RA, serum calprotectin levels have been associated with increased disease activity in RA [26].

The proposed study fills an important gap in the literature by examining how background factors and biological measures of GI health (gut microbiome diversity and GI inflammation) may improve our understanding of the hypothesized relationship between dietary intake and RA disease activity (Figure 2).

Figure 2. Study model. GI: gastrointestinal; RA: rheumatoid arthritis.

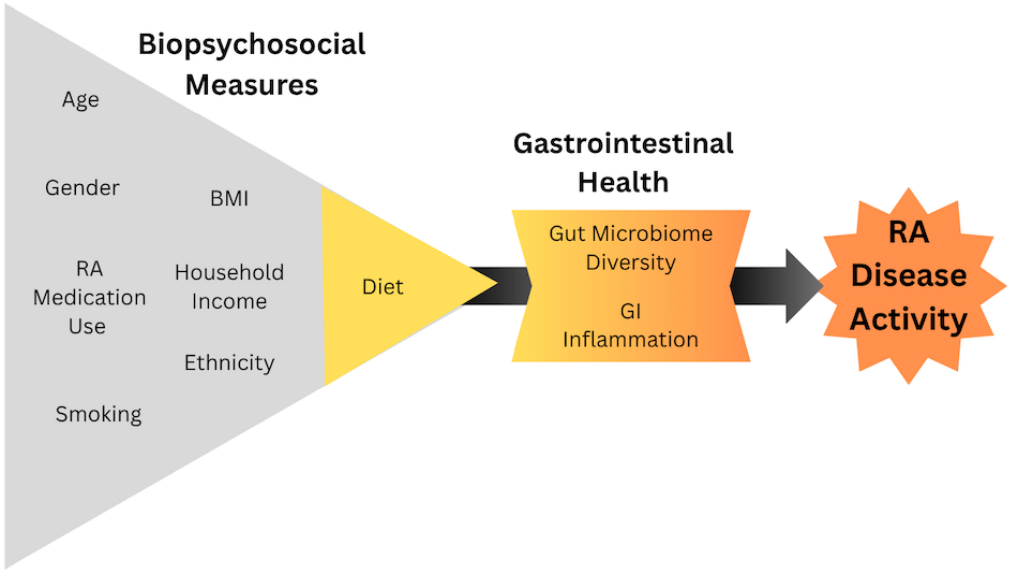

This Study

In the southwest United States, approximately 28,000 Arizonans are diagnosed with RA, including many Black, Hispanic, and Native American individuals. The rate of RA diagnosis is higher in women, and this rate increases for women of color. In Hispanic populations, for example, this rate is doubled compared with White populations [32]. Additionally, the prevalence of RA has been found to be significantly higher for Native Americans [43]. Further, poverty rates are high for many Native Americans. The Navajo Nation, for example, has a reported poverty rate of 38% [55], which is significantly higher than the Arizona average of 16.4% [56]. Unfortunately, RA studies continue to focus on individuals of European ancestry. This is a serious concern for patient outcomes as well as for the development and implementation of personalized medicine focused on genome-specific care [32]. The lack of diversity in RA research leads to a significant breach in patient care.

There are two aims for this study:

- Aim 1: Examine associations between diet quality, other important individual biopsychosocial factors, and GI health

XSL•FO
RenderX

in adults with RA. Hypothesis 1: We expect individual biopsychosocial factors will be associated with diet quality, and most participants will report dietary intake that is either poor or fair quality per preliminary Healthy Eating Index (HEI)–2020 scores. We also expect that individuals who report lower diet quality will have associated decreased GI health.

- Aim 2: Examine predicted associations between diet quality, gut microbiome diversity and composition, and GI inflammation on RA disease activity. Hypothesis 2: We expect that diet quality will not be directly associated with RA disease activity; instead, we will find a significant indirect relationship between diet quality and RA disease activity through relationships between diet quality and biological measures of GI inflammation and gut microbiome diversity, and between gut health measures and RA disease activity. We will test our hypotheses with 4 indicators of RA disease activity. Past research has shown that the Disease Activity Score using 28 joint counts (DAS-28), HAQ-DI: Health Assessment Questionnaire–Disability Index (HAQ-DI) and Pain Scale, Erythrocyte Sedimentation Rate (ESR), and high-sensitivity C-reactive Protein (hsCRP) are moderately correlated. Our hypotheses regarding the role of gut health will be reinforced if either hypothesis is supported.

Methods

Research Design

This study uses a cross-sectional design to describe associations between background factors and diet quality and to test 2 a priori hypotheses regarding the relationship between diet quality, gut microbiome diversity, and GI inflammation on disease activity in RA. The inclusion of GI measures addresses an important gap in the literature and inform future research.

Sample and Recruitment

The target population for this proposed study are adults aged ≥18 years diagnosed with RA. We will use posters and business cards (Figures 3 and 4) to recruit adults with RA from a variety of locations including the Arizona Arthritis and Rheumatology Associates offices in Arizona (Flagstaff, Prescott, Avondale, Sun City, Glendale, Phoenix, Mesa, Gilbert, San Tan Valley), the Mayo Clinic Collaborative Research in Rheumatology (Scottsdale), Northern Arizona University, North Country Healthcare with locations in northern Arizona (Flagstaff, Bullhead City, Grand Canyon, Kingman, Show Low, Payson, Holbrook, Williams, Kingman, Round Valley, Winslow, Lake Havasu City), and the Arizona office of the Arthritis Foundation (including social media posts on Instagram and Facebook and at the Arthritis Run). The posters and business cards include a link to a website and a QR code that directs interested individuals to a secure eligibility survey through REDCap (Research Electronic Data Capture; Vanderbilt University). Eligible individuals will receive access to an online informed consent form. They will be directed to carefully read the informed consent form and will have the option to click “yes” to agree to participate in the study or “no” to decline. If they choose “no,” they will receive a message thanking them for their time and consideration, as well as information directing any questions or concerns to the principal investigator. Consented participants will receive access to all online surveys and timeline (Figure 5). Upon completion of all online questionnaires, participants will receive a link for the final survey to sign up for an in-person appointment on a convenient date, time, and location (Northern Arizona University Flagstaff or Phoenix North Valley Campuses). We will also offer specific days focusing on Spanish-speaking individuals. Inclusion criteria are that participants must be aged ≥18 years, able to read and write in English or Spanish, and have a diagnosis of RA that is managed by a rheumatology health care provider. Exclusion criteria for this study include age <18 years, pregnancy, inability to read or write in English or Spanish, or a diagnosis of any other autoimmune disease.

XSL•FO
RenderX

Figure 3. Recruitment poster. NAU: Northern Arizona University.

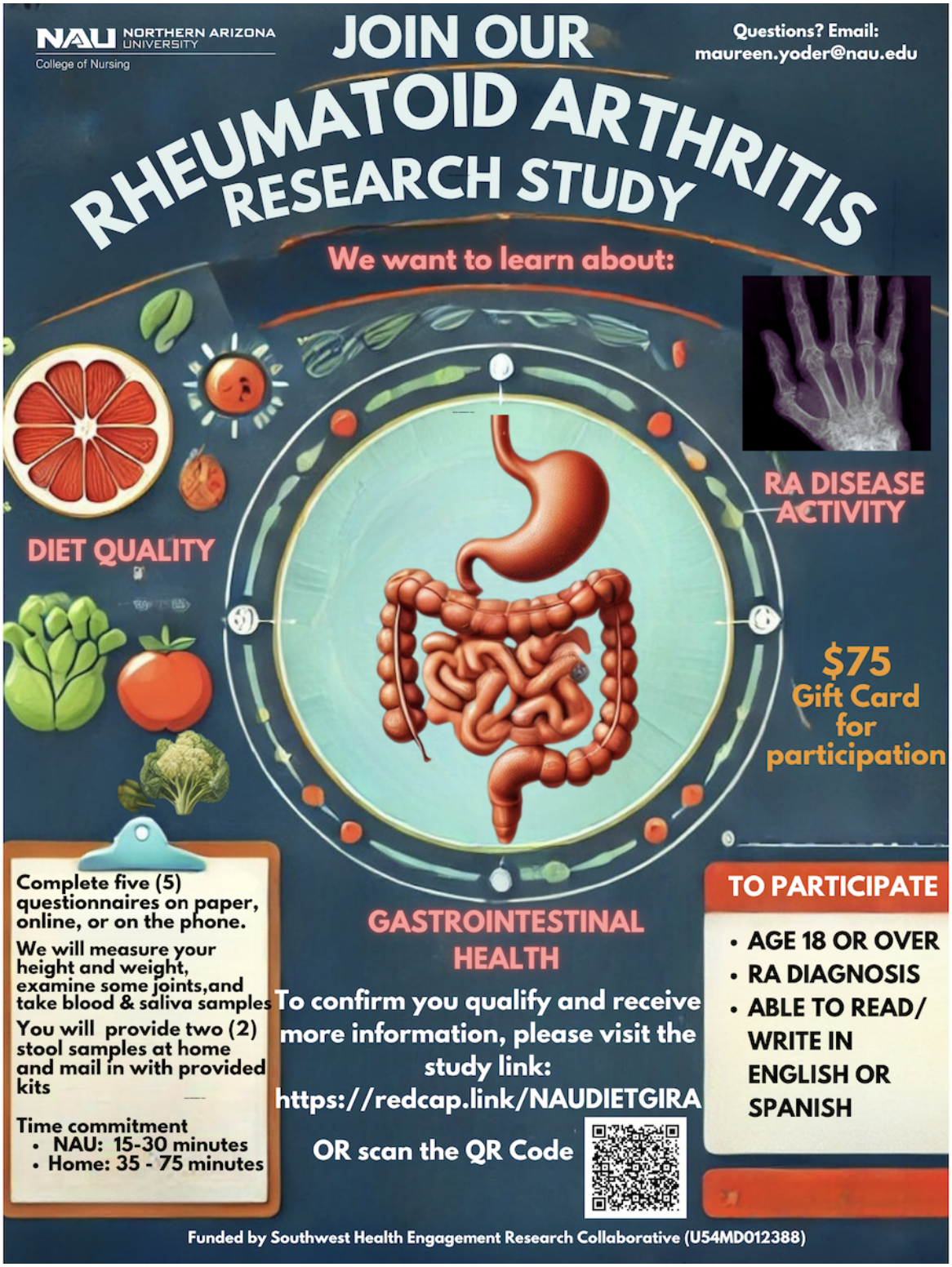

XSL•FO
RenderX

Figure 4. Recruitment business card.



XSL•FO
RenderX

Figure 5. Study timeline for participants. RA: rheumatoid arthritis.



Each data collection venue will be set up with 4 stations, beginning with a welcome table. On the appointment day, participants will first meet with the principal investigator or trained research assistant and receive a tote bag with instructions, stool collection kits, and an addressed, stamped shipping envelope for self-collection of stool samples. At this station, participants will also receive all directions to complete the study at home. Next, a researcher will work with the participant to measure height and weight, followed by the DAS-28, a physical exam to determine the number of swollen and tender joints, and pain level. Next, a blood sample will be drawn to measure ESR and C-reactive protein (CRP) levels, followed by meeting with a research assistant to complete the Automated Self-Administered 24-Hour Dietary Assessment Tool (ASA24), a dietary assessment. Figure 6 shows the informational checklist to be provided to participants.

Figure 6. Checklist for data collection day. CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; RA: rheumatoid arthritis.

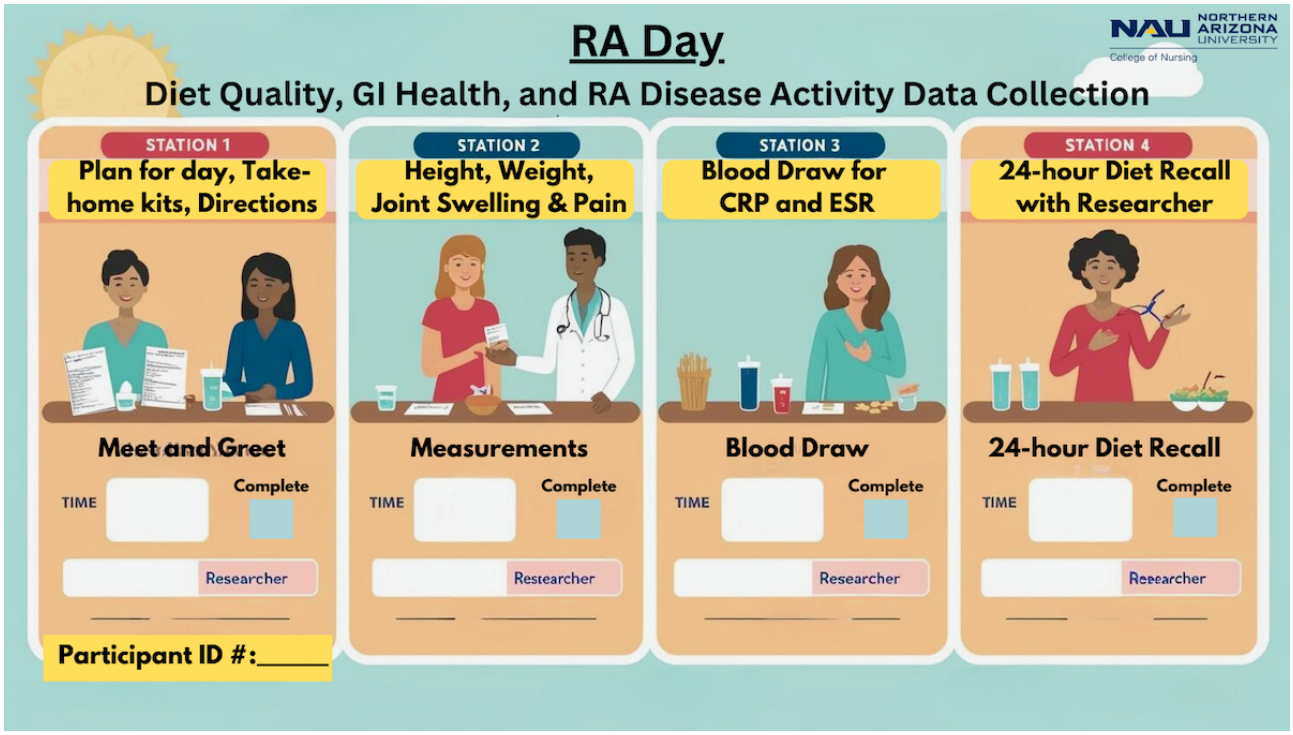

XSL•FO
RenderX

Data Management

Quantitative data will be examined for completeness as they are received. All data will then be cleaned and checked for errors before data analysis. Quantitative data will be analyzed descriptively using frequencies and proportions for the total sample. For nonnormal distributions of continuous data, medians and IQRs will be used. Categorical data will be presented using proportions and frequencies.

Data Analysis Plan

Data from the above-described measures will be combined into a single dataset with any identifying information removed before analysis. All data analyses (and visualizations) will be conducted with SPSS (version 29; IBM Corp) and R (version 4.2.3; R Foundation for Statistical Computing). Before analyses data will be cleaned and examined for outliers, and on the basis of distributions, appropriate data transformations will be computed to meet assumptions for parametric statistical tests. If not, nonparametric tests will be performed where needed. For Aim 1, we will conduct appropriate parametric or nonparametric statistical tests with self-report measures to examine how our primary model variables are differentially associated with background factors. Bonferroni-adjusted *P* values will be used to control for Type I errors. In our analyses, we will use the HEI-2020, HAQ-DI/Pain Scale, and DAS-28 scales as continuous measures. For ease of comparison with existing literature, we will recode these measures into the published categorical designations to describe our sample. In addition to the examination of *P* values, CIs will also be reported. This study will use a hierarchical linear regression to test 2 separate hypotheses that diet quality will be indirectly related to RA disease activity through its expected relationships with gut diversity and gut inflammation. Each model will regress one of the RA disease activity variables onto the following measures: step 1, previously identified significant demographic characteristics; step 2, continuous measure of diet quality; and step 3, two measures of GI health.

Power

To determine sample size, several multiple regression power analyses (fixed model, R^2 deviation from zero), were completed using G*Power software (version 3.1.9.6) [56]. The analyses were calculated using power specified at 0.80 across 3 different effect sizes (f^2) and 3 different α levels (0.05, and 2 types of Bonferroni-adjusted values) [57]. Two effect size approximations were based on published correlations on measures similar to those proposed in this study (diet quality, GI inflammation, gut diversity) with DAS-28 (f^2=0.120) and HAQ-DI/pain (f^2=0.083) measures, and one effect size (0.15) was a general recommendation for a medium effect [58]. We determined that a sample size of 96 would be a reasonable sample size to test our two study hypotheses (f^2=0.12; power=0.80; α=0.05) and slightly higher than the sample size of 91 determined using a conservative α with Cohen medium effect size (f^2=0.15; power=0.80; α=0.025). A post hoc power analysis (fixed model, R^2 increase) also indicated that a sample size of 96 is adequate to insure at least 80% power to detect the ability of our two primary predictor variables (GI inflammation and gut diversity) to contribute at least 8% additional variance in our outcome measures (RA disease activity) above and beyond the variance (20%) expected to be predicted by our 6 background measures.

Measures

Background Measures

Background measures will be assessed with a demographic questionnaire that includes questions regarding yearly household income (less than US $25,000, US $25,000-US $49,999, US $50,000-US $74,999, US $75,000-US $99,999, US $100,000-US $124,999, more than US $125,000, do not wish to answer, or do not know for certain), sex (male or female), age, zip code years since diagnosis of RA, preferred language for communication, highest educational level (elementary, high school or General Educational Development, vocational training, some college or associate's degree, bachelor's degree, master's degree, doctoral degree, or do not wish to answer), marital status (single, married, separated, divorced, widowed, other, or do not wish to answer), race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or other Pacific Islander, White, more than one race, other, or do not wish to answer), current tobacco use, current cannabis use, 2 questions regarding food security, and current medication use (disease-modifying antirheumatic agents, corticosteroids in the last month, nonsteroidal anti-inflammatories in the last week). The demographic questionnaire also includes questions regarding dietary beliefs associated with RA symptoms.

Diet Quality

The Arizona Food Frequency Questionnaire (AFFQ) will be used to measure dietary intake. The 175 item self-reported AFFQ is a modified version of the National Cancer Institute Health Habits and History Questionnaire and will be used to determine self-reported intake (foods, portion sizes, beverages, and vitamins and minerals) for a 4-week period [57]. Reliability of the AFFQ demonstrated a correlation of 0.63, and test-restest reliability correlation of 0.48 (adjusted for energy-related nutrients) and 0.54 (unadjusted nutrients) [57]. The AFFQ will be used to determine participant intake (foods, beverages, and portion sizes) to estimate intake for a total of 34 variables [57]. The AFFQ is available in English and Spanish. Diet quality will be scored with the HEI-2020, which is the most recent update of the HEI based on the comparison of dietary intake to the 2015-2020 Dietary Guidelines for Americans [58,59]. The HEI-2020 was developed in a collaborative effort of the National Cancer Institute and the US Department of Agriculture to score diet quality based on comparison of dietary intake to the 2015-2020 Dietary Guidelines for Americans [58,59]. The HEI-2020 is scored out of 100 points: less than 51 is poor, 51-80 is fair, and above 80 is good diet quality [9]. Americans have an average score of 59 out of 100 [60]. AFFQ analysis and HEI-2020 score calculations will be completed at the Behavioral Measurement Intervention Shared Resources at the University of Arizona.

The ASA24 will be used to measure dietary intake for the previous 24-hours. The self-reported ASA24 was initially based on the previously validated Automated Multiple-Pass Method

XSL•FO
RenderX

[61], and this validated, reliable tool has been used to complete over 1 million 24-hour diet recalls, with collected data used in over 1000 peer-reviewed publications [62]. Various usability and cognitive tests have been completed on sequential versions of the ASA24, and indicate similar results to interview-administered, standardized 24-hour diet recalls [63]. Research has indicated that approximately 80% of consumed items are reported with the ASA24 [64]. The ASA24 is available in English and Spanish.

Gut Microbiome

Fecal microbiome composition and diversity will be assessed using 16S rRNA gene sequencing. Sequencing will be performed by the Translational Genomics Research Institute North Pathogen and Microbiome Division in Flagstaff, Arizona. 16S rRNA gene sequencing is one of the most widely used applications to examine bacterial microbiota [65] and is commonly used in RA research [66-72]. Stool samples will be self-collected at home by participants, who will receive a toilet accessory, disposable gloves, a sterile specimen cup, an OMNIgene GUT DNA RNA Fecal Collection Tube, 2 biohazard bags, 1 shipping box, and 1 plastic FedEx Clinical Pak. Fecal samples will be collected in OMNIgene GUT Fecal collection kits, which stabilize DNA and RNA for transport at room temperature for several days. Participants will place the samples in the appropriate shipment box, enclose each with the plastic FedEx Clinical Pak, and attach the appropriate shipping label. The samples must be shipped via FedEx Monday through Wednesday. Samples will be stored at the Pathogen and Microbiome Institute at Northern Arizona University in Co-I Cope's lab at −80°C until transported to and analyzed at the Translational Genomics Research Institute in Flagstaff. Samples will be extracted using a Kingfisher Flex system that extracts 96 samples in a single batch and sequenced on the same Illumina MiSeq run. If interim analyses are necessary, we will include 5 replicate samples on each run and mock communities to assess batch effects.

Microbiome Bioinformatics

Microbiome bioinformatics will be performed with QIIME 2 (Quantitative Insights into Microbial Ecology version 2) [52] using Denoising Algorithm Directed by Amplicon Variants for sequence quality control and definition of amplicon sequence variants to provide the highest possible taxonomic specificity. Alpha diversity (community richness) metrics will be computed with QIIME 2, including Faith's Phylogenetic Diversity, observed operational taxonomic unit, and Shannon diversity. Beta diversity (community dissimilarity) metrics will also be computed with QIIME 2, including Bray-Curtis dissimilarity, Jaccard, weighted UniFrac, and unweighted UniFrac distances. These diversity metrics, along with taxonomic profiles of samples, will be used to compare microbiome compositions associated with categorical variables. Group comparisons of alpha diversity will be performed with nonparametric Kruskal-Wallis tests, and group comparisons of beta diversity will be performed with nonparametric PERMANOVA. Since identifying amplicon sequence variants and taxa that are differentially abundant across disease groups is currently a very active area of research, we will assess available methods at the time of analysis for performing this step but expect to work with multiple methods, including established methods such as analysis of compositions of microbiomes with bias correction. All *P* values will be corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate correction.

GI Inflammation

GI inflammation will be determined with fecal calprotectin levels using enzyme-linked immunosorbent assay. Fecal calprotectin values in adults of less than 50 µg/g are normal, 50-100 µg/g indicate moderate inflammation, and greater than 150 µg/g indicate significant inflammation [73]. Samples will be collected and shipped as described for the gut microbiome analysis. Samples will be assessed using the Human IL-6 enzyme-linked immunosorbent assay kit by Sigma Aldrich (RAB0306) as per manufacturer's instructions.

Health Assessment Questionnaire–Disability Index and Pain Scale

The HAQ-DI and Pain Scale will be used as a self-reported measure of RA disease activity, as it is frequently used in research and clinical practice [74]. The HAQ-DI and Pain Scale is recommended by the American College of Rheumatology [74]. It was chosen due to participant experience, low burden, and ease of scoring [75]. The HAQ-DI and Pain Scale is a 9-item self-report (8 items associated with HAQ-DI and 1 item for pain) that determines the overall health or disease activity of patients with RA [74,76]. Test-retest correlation has been reported as 0.87-0.99, and HAQ-DI score had a correlation of 0.71-0.95 with task performance [75]. Validity of the HAQ-DI has been found "in literally hundreds of studies" [75]. Total scores range from 0 to 3: where 0-1 indicates mild to moderate disability, 1-2 indicates moderate to severe disability, and 2-3 indicates extremely severe disability [74].

Disease Activity Score Using 28 Joint Counts

The DAS-28 is frequently used in practice and research, as it offers ease of scoring and low participant burden [77]. The DAS-28 is a valid, reliable tool that uses hematologic biomarkers, examination, and participant-reported information to calculate disease activity. DAS-28 CRP and DAS-28 ESR scores have been highly correlated (r=0.92; P<.001) [78], and this project will use the DAS-28 ESR, as was used in a recent study [79]. The DAS-28 has had high internal consistency reliability (Cronbach α=0.719) and is significantly correlated with disability as measured by a health assessment questionnaire (P<.001) in other work [77]. Preliminary work offered evidence of significant correlation between erythrocyte sedimentation levels, CRP levels, and DAS-28 scores; therefore, DAS-28 scores will be used as evidence of inflammation in the proposed project. Scores of <2.6 indicate remission, scores of 2.6-3.2 indicate low disease activity, and scores >3.2 indicate moderate to severe disease activity [80].

Hematologic Biomarkers

Hematologic proinflammatory biomarkers will be used as evidence of disease activity in this study, as they are a direct measure of inflammation. These biomarkers included the ESR and hsCRP analyses. The ESR indicates generalized inflammation by the rate of erythrocyte sediment that

XSL•FO
RenderX

accumulates in a test tube over 1 hour; the normal range for adults is 0-20 mm per hour (male) and 0-30 mm per hour (female) [81]. The hsCRP indicates generalized inflammation, as CRP is an acute-phase protein produced by the liver in response to increased levels of interleukin-6 in the blood; values of less than 1.0 are normal, 1.0-3.0 indicate moderate inflammation, and values greater than 3.0 indicate high levels of inflammation [7,81]. While the ESR and hsCRP indicate generalized inflammation in the body, not specific to RA, both are used to determine severity of disease activity in patients with RA [81].

Ethical Considerations

This study has been approved by the Northern Arizona University Institutional Review Board (2111208-12). Eligible participants will receive online access to the informed consent form in REDCap in their choice of English or Spanish. They will be directed to carefully read the form and click "yes" to agree to participate in the study, or "no" to decline. If they choose to participate, they will receive access to all online surveys (their choice of English or Spanish); at the end, they will be directed to self-schedule an in-person appointment. Those who choose not to participate will receive a message of thanks for their time and consideration. Participants can withdraw from the study at any time. There are no risks associated with participant withdrawal. As this is a cross-sectional study, participants who choose to withdraw will not be asked to continue in any other activities related to this research study.

Researchers will use the secure REDCap, Qualtrics, and NIH platforms for participant questionnaires. Each participant will receive an online link or code to access online platforms. Questionnaires will be uploaded into REDCap, with the exception of the AFFQ, which is only available in Qualtrics, and the ASA-24, which is located on a secure NIH website. Questionnaires collected on paper will have the participant ID code placed in the top right corner of each document. All questionnaires, including the demographic questionnaire, will remain nonidentifiable. Stool and hematology samples will be labeled with participant ID number only. Nonidentifiable data will be securely stored on an encrypted, password-protected drive. The principal investigator will have access to the data. Non–personally identifiable data will be kept indefinitely for future analyses.

Once the completed surveys and both stool samples are received, participants will be sent a US $75 gift card (either an electronic or physical card, based on participant preference).

Results

This study was funded in February 2024. As of December 19, 2025, a total of 80 individuals have been recruited. Data analysis has not yet commenced at the time of manuscript submission. Study results are expected to be published in fall 2026.

Discussion

Overview

Biopsychosocial factors such as socioeconomic status, age, ethnicity, gender, BMI, smoking, and RA medication use are known factors that can affect eating habits and therefore may influence the gut microbiome, but this research question has not yet been examined in patients with RA. Our exploratory analyses will examine the extent to which background variables are associated with dietary quality and biological markers of gut health, including GI inflammation and gut microbiome composition and diversity, known to be altered in individuals with RA [24]. We have reasonable expectations that low socioeconomic status will be associated with poor diet quality, and most participants will report dietary intake that is either poor or fair quality per preliminary HEI-2020 scores [17]. We also expect individuals reporting low diet quality will have associated gut microbiome dysbiosis and higher fecal calprotectin levels, indicating increased inflammation [21-23,26]. Significant results from analyses will be used in subsequent analyses to account for patterns related to diversity of experience and background in patients with RA. There is little evidence that self-reported diet quality, is a reliable predictor of RA disease activity. Preliminary work completed by the primary author examined associations between diet quality and disease activity in 50 adults with RA. Diet quality scores were lower in participants (average score 56) compared to the national average (average score 59) [79]. Women (P=.003) and those with increased age (P=.02) were more likely to report a higher diet quality [79]. A total of 44% of participants reported that they believe their diet affects RA disease activity, and these participants were significantly more likely to change their diet (P<.001). Higher education level (at least some college) was associated with this belief (β=–1.535; P=.02). Participants with lower diet quality had higher pain (β=–0.396; P=.02) and inflammation (P=.02) [79]. Significantly higher HAQ-DI scores, an indicator of disability, were noted in women (β=.570, P=.001) [79]. Although decreased levels of proinflammatory biomarkers were previously associated with a higher quality diet [80], no other associations were found between diet quality and overall disease activity [79]. This lack of evidence may be accounted for by the diet quality scores of participants in this preliminary work, as only 4% (n=2) individuals exhibited "good" diet quality [79]. Other considerations are that this was a relatively homogeneous (74% White) sample. Diet quality may directly impact functional ability (eg, shopping, preparation, and even eating food) in RA [9].

By including biological markers of gut health (presumed to be affected by diet quality), we will be able to test 2 hypotheses regarding how gut health may be a key link to understanding the relationship between diet quality and disease activity in patients with RA. We expect that diet quality will not be directly associated with RA disease activity; instead, we will find a significant indirect relationship between diet quality and RA disease activity through relationships between diet quality and biological measures of GI inflammation and gut microbiome diversity, and between GI health measures and RA disease activity. Our hypotheses regarding the role of gut health will

XSL•FO
RenderX

be reinforced if either hypothesis is supported. To the best of our knowledge, this project is the first to consider associations between background biopsychosocial variables, diet quality, GI health, and disease activity in adults with RA. An innovative advantage specific to this study is the comparison of an existing, vetted self-reported dietary measure with measures of GI health.

Possible Limitations and Alternate Solutions

Obstacles may occur with data collection. Participants may have difficulty completing paper or online questionnaires if they have significant pain or deformity of their wrists, hands, or fingers. To address these concerns, participants will have the option of a telephone interview for all surveys. A further obstacle may be participant recruitment, as a total of 96 participants are needed. A total of 55 participants were previously recruited at one rheumatology office in 50 days. The sample previously recruited was homogeneous, as 68% reported some form of higher education (at least some college), and 74% identified as White. To address this, we will recruit from across Arizona using social media through the Arthritis Foundation, as well as from rheumatology practices across the greater Phoenix metro area, which, according to the research manager, serves a diverse and underserved community.

Conclusion

RA is a complicated disease that impacts millions [14]. Few reach sustained remission, even while following provider recommendations [3]. A better understanding of the various factors that impact this complicated disease has the potential to support changes in research and care that will improve the lives of people with RA. We believe biopsychosocial factors will be associated with self-reported diet quality, which will be indirectly associated with disease activity through gut microbiome diversity and level of GI inflammation in adults with RA. The knowledge gained in this study will be a foundation to inform future interventional research targeting diet quality to support GI health and decrease RA disease activity. Further, the details of this research plan provide methodological resources for other RA researchers.

Acknowledgments

The authors declare the use of generative AI in the research and writing process. In accordance with the Generative Artificial Intelligence Disclosure Taxonomy (GAIDeT; 2025), a structured framework for transparently reporting the use of generative AI in scholarly work, specific tasks were delegated to generative AI tools under full human supervision. These tasks included idea generation and reformatting of text. The generative AI tools used were ChatGPT 5.1 and Canva Pro. Responsibility for the accuracy, integrity, and final content of the manuscript lies entirely with the authors. Generative AI tools are not listed as authors and do not bear responsibility for the final outcomes. This disclosure was submitted by MMY. Additionally, figures related to recruitment and participant information were generated with the assistance of ChatGPT 5.1, and the Canva platform was used solely for image editing. Generative AI was not used in any other aspect of this manuscript.

Funding

This research received funding from the Southwest Health Engagement and Research Collaborative at Northern Arizona University (U54MD012388), which is sponsored by the National Institute on Minority Health and Health Disparities. Analyses of data will also be supported by the National Cancer Institute of the National Institutes of Health under award number P30 CA023074.

Data Availability

The datasets generated and analyzed during this study will be available in the OpenKnowledge@NAU repository at Northern Arizona University.

Authors' Contributions

All authors took part in the development, editing, and approval of the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from the Pilot Project Program (PPP) Award Committee, Investigator Development Core, Southwest Health Engagement and Research Collaborative (SHERC) at Northern Arizona University.
[PDF File (Adobe PDF File), 295 KB - resprot_v15i1e79889_app1.pdf]

References

1. Winkvist A, Bärebring L, Gjertsson I, Ellegård L, Lindqvist HM. A randomized controlled cross-over trial investigating the effect of anti-inflammatory diet on disease activity and quality of life in rheumatoid arthritis: the anti-inflammatory diet In rheumatoid arthritis (ADIRA) study protocol. Nutr J 2018;17(1):44 [FREE Full text] [doi: 10.1186/s12937-018-0354-x] [Medline: 29678183]

XSL•FO
RenderX

2. Tanaka S. Regulation of bone destruction in rheumatoid arthritis through RANKL-RANK pathways. World J Orthop 2013;4(1):1-6 [FREE Full text] [doi: 10.5312/wjo.v4.i1.1] [Medline: 23362468]
3. Ajeganova S, Huizinga T. Sustained remission in rheumatoid arthritis: latest evidence and clinical considerations. Ther Adv Musculoskelet Dis 2017;9(10):249-262 [FREE Full text] [doi: 10.1177/1759720X17720366] [Medline: 28974987]
4. Michelsen B, Uhlig T, Sexton J, Hammer H, Kristianslund E, Haugeberg G, et al. THU0281 do improvements in health-related quality of life during DMARD treatment differ between psoriatic and rheumatoid arthritis patients? Data from the prospective observational NOR-DMARD study, including baseline comparisons with Norwegian general population controls. Ann Rheum Dis 2018;77:359. [doi: 10.1136/annrheumdis-2018-eular.1758]
5. Kłak A, Raciborski F, Samel-Kowalik P. Social implications of rheumatic diseases. Reumatologia 2016;54(2):73-78 [FREE Full text] [doi: 10.5114/reum.2016.60216] [Medline: 27407283]
6. Ikdahl E, Wibetoe G, Rollefstad S, Salberg A, Bergsmark K, Kvien TK, et al. Guideline recommended treatment to targets of cardiovascular risk is inadequate in patients with inflammatory joint diseases. Int J Cardiol 2019;274:311-318. [doi: 10.1016/j.ijcard.2018.06.111] [Medline: 30007486]
7. Daruna JH. Introduction to Psychoneuroimmunology. London: Elsevier; 2012.
8. Gibofsky A. Epidemiology, pathophysiology, and diagnosis of rheumatoid arthritis: a synopsis. Am J Manag Care 2014;20(7 Suppl):S128-S135 [FREE Full text] [Medline: 25180621]
9. Berube LT, Kiely M, Yazici Y, Woolf K. Diet quality of individuals with rheumatoid arthritis using the Healthy Eating Index (HEI)-2010. Nutr Health 2017;23(1):17-24. [doi: 10.1177/0260106016688223] [Medline: 28112038]
10. D'Mello C, Swain M. Immune-to-brain communication pathways in inflammation-associated sickness and depression. Curr Top Behav Neurosci 2017;31:73-94. [doi: 10.1007/7854_2016_37] [Medline: 27677781]
11. Louveau A, Smirnov I, Keyes TJ, Eccles JD, Rouhani SJ, Peske JD, et al. Structural and functional features of central nervous system lymphatic vessels. Nature 2015;523(7560):337-341 [FREE Full text] [doi: 10.1038/nature14432] [Medline: 26030524]
12. Milrad SF, Hall DL, Jutagir DR, Lattie EG, Ironson GH, Wohlgemuth W, et al. Poor sleep quality is associated with greater circulating pro-inflammatory cytokines and severity and frequency of chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME) symptoms in women. J Neuroimmunol 2017;303:43-50 [FREE Full text] [doi: 10.1016/j.jneuroim.2016.12.008] [Medline: 28038892]
13. Sun Y, Wang D, Salvadore G, Hsu B, Curran M, Casper C, et al. The effects of interleukin-6 neutralizing antibodies on symptoms of depressed mood and anhedonia in patients with rheumatoid arthritis and multicentric Castleman's disease. Brain Behav Immun 2017;66:156-164. [doi: 10.1016/j.bbi.2017.06.014] [Medline: 28676350]
14. McGarrity-Yoder M, Insel K, Crane T, Pace T. Interdisciplinary rheumatoid arthritis research and patient care: an introduction and critique of the biopsychosocial model of disease experience in rheumatoid arthritis. Am J Lifestyle Med 2024:15598276241298937. [doi: 10.1177/15598276241298937] [Medline: 39540189]
15. Engel GL. The clinical application of the biopsychosocial model. Am J Psychiatry 1980;137(5):535-544. [doi: 10.1176/ajp.137.5.535] [Medline: 7369396]
16. Dodd M, Janson S, Facione N, Faucett J, Froelicher ES, Humphreys J, et al. Advancing the science of symptom management. J Adv Nurs 2001;33(5):668-676. [doi: 10.1046/j.1365-2648.2001.01697.x] [Medline: 11298204]
17. McGarrity-Yoder ME, Insel KC, Crane TE, Pace TW. Diet quality and disease activity in rheumatoid arthritis. Nutr Health 2022;28(4):581-590. [doi: 10.1177/02601060211044311] [Medline: 34668452]
18. Mitka M. Do flawed data on caloric intake from NHANES present problems for researchers and policy makers? JAMA 2013;310(20):2137-2138. [doi: 10.1001/jama.2013.281865] [Medline: 24281443]
19. Dhurandhar NV, Schoeller D, Brown AW, Heymsfield SB, Thomas D, Sørensen TIA, et al. nergy balance measurement: when something is not better than nothing. Int J Obes (Lond) 2015;39(7):1109-1113 [FREE Full text] [doi: 10.1038/ijo.2014.199] [Medline: 25394308]
20. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr 2015;145(12):2639-2645 [FREE Full text] [doi: 10.3945/jn.115.219634] [Medline: 26468491]
21. García-Vega Á, Corrales-Agudelo V, Reyes A, Escobar JS. Diet quality, food groups and nutrients associated with the gut microbiota in a nonwestern population. Nutrients 2020;12(10):2938 [FREE Full text] [doi: 10.3390/nu12102938] [Medline: 32992776]
22. Laitinen K, Mokkala K. Overall dietary quality relates to gut microbiota diversity and abundance. Int J Mol Sci 2019;20(8):1835 [FREE Full text] [doi: 10.3390/ijms20081835] [Medline: 31013927]
23. Guerreiro CS, Calado Â, Sousa J, Fonseca JE. Diet, microbiota, and gut permeability-the unknown triad in rheumatoid arthritis. Front Med (Lausanne) 2018;5:349 [FREE Full text] [doi: 10.3389/fmed.2018.00349] [Medline: 30619860]
24. Sun Y, Chen Q, Lin P, Xu R, He D, Ji W, et al. Characteristics of gut microbiota in patients with rheumatoid arthritis in Shanghai, China. Front Cell Infect Microbiol 2019;9:369 [FREE Full text] [doi: 10.3389/fcimb.2019.00369] [Medline: 31709198]
25. Bjarnason I. The use of fecal calprotectin in inflammatory bowel disease. Gastroenterol Hepatol (N Y) 2017;13(1):53-56 [FREE Full text] [Medline: 28420947]

XSL•FO
RenderX

26. Jarlborg M, Courvoisier DS, Lamacchia C, Martinez Prat L, Mahler M, Bentow C, et al. Serum calprotectin: a promising biomarker in rheumatoid arthritis and axial spondyloarthritis. Arthritis Res Ther 2020;22(1):105 [FREE Full text] [doi: 10.1186/s13075-020-02190-3] [Medline: 32375861]
27. Krasselt M, Baerwald C. Sex, symptom severity, and quality of life in rheumatology. Clin Rev Allergy Immunol 2019;56(3):346-361. [doi: 10.1007/s12016-017-8631-6] [Medline: 28795307]
28. Young A, Dixey J, Cox N, Davies P, Devlin J, Emery P, et al. How does functional disability in early rheumatoid arthritis (RA) affect patients and their lives? Results of 5 years of follow-up in 732 patients from the Early RA Study (ERAS). Rheumatology (Oxford) 2000;39(6):603-611. [doi: 10.1093/rheumatology/39.6.603] [Medline: 10888704]
29. Oláh C, Kardos Z, Sepsi M, Sas A, Kostyál L, Bhattoa HP, et al. Assessment of intracranial vessels in association with carotid atherosclerosis and brain vascular lesions in rheumatoid arthritis. Arthritis Res Ther 2017;19(1):213 [FREE Full text] [doi: 10.1186/s13075-017-1422-x] [Medline: 28950911]
30. Ward H, Toledano M, Shaddick G, Davies B, Elliott P. Oxford Handbook of Epidemiology for Clinicians. Oxford: Oxford University Press; 2012.
31. Schneider M, Manabile E, Tikly M. Social aspects of living with rheumatoid arthritis: a qualitative descriptive study in Soweto, South Africa - a low resource context. Health Qual Life Outcomes 2008;6:54 [FREE Full text] [doi: 10.1186/1477-7525-6-54] [Medline: 18651986]
32. Díaz-Peña R, Quiñones LA, Castro-Santos P, Durán J, Lucia A. Latin American genes: The great forgotten in rheumatoid arthritis. J Pers Med 2020;10(4):196 [FREE Full text] [doi: 10.3390/jpm10040196] [Medline: 33114702]
33. Lutgendorf SK, Costanzo ES. Psychoneuroimmunology and health psychology: an integrative model. Brain Behav Immun 2003;17(4):225-232. [doi: 10.1016/s0889-1591(03)00033-3] [Medline: 12831823]
34. Zhou Y, Sun M. A meta-analysis of the relationship between body mass index and risk of rheumatoid arthritis. EXCLI J 2018;17:1079-1089 [FREE Full text] [doi: 10.17179/excli2018-1763] [Medline: 30585275]
35. Ruiz-Esquide V, Sanmartí R. Tobacco and other environmental risk factors in rheumatoid arthritis. Reumatología Clínica (English Edition) 2012;8(6):342-350. [doi: 10.1016/j.reumae.2012.06.007]
36. Louthrenoo W, Kasitanon N, Katchamart W, Aiewruengsurat D, Chevaisrakul P, Chiowchanwisawakit P, et al. 2016 updated Thai rheumatism association recommendations for the use of biologic and targeted synthetic disease-modifying anti-rheumatic drugs in patients with rheumatoid arthritis. Int J Rheum Dis 2017;20(9):1166-1184. [doi: 10.1111/1756-185X.13130] [Medline: 28730640]
37. Disparities NIoMHaH. Overview on Minority Health and Health Disparities. 2022. URL: https://www.nimhd.nih.gov/about/what-are-health-disparities [accessed 2021-12-10]
38. Baldassari AR, Cleveland RJ, Luong MLN, Jonas BL, Consortium for the Longitudinal Evaluation of African Americans with Early Rheumatoid Arthritis, Conn DL, et al. Socioeconomic factors and self-reported health outcomes in African Americans with rheumatoid arthritis from the Southeastern United States: the contribution of childhood socioeconomic status. BMC Musculoskelet Disord 2016;17:10 [FREE Full text] [doi: 10.1186/s12891-016-0882-5] [Medline: 26754747]
39. Barber CEH, Lacaille D, Croxford R, Barnabe C, Marshall DA, Abrahamowicz M, et al. A population-based study evaluating retention in rheumatology care among patients with rheumatoid arthritis. ACR Open Rheumatol 2022;4(7):613-622 [FREE Full text] [doi: 10.1002/acr2.11442] [Medline: 35514156]
40. Waltz M, Kriegel W, van't Pad Bosch P. The social environment and health in rheumatoid arthritis: marital quality predicts individual variability in pain severity. Arthritis Care Res 1998;11(5):356-374. [doi: 10.1002/art.1790110507] [Medline: 9830880]
41. Intriago M, Maldonado G, Cárdenas J, Ríos C. Clinical characteristics in patients with rheumatoid arthritis: differences between genders. ScientificWorldJournal 2019;2019:8103812 [FREE Full text] [doi: 10.1155/2019/8103812] [Medline: 31354388]
42. Puolakka K, Kautiainen H, Pekurinen M, Möttönen T, Hannonen P, Korpela M, FIN-RACo Trial Group. Monetary value of lost productivity over a five year follow up in early rheumatoid arthritis estimated on the basis of official register data on patients' sickness absence and gross income: experience from the FIN-RACo trial. Ann Rheum Dis 2006;65(7):899-904 [FREE Full text] [doi: 10.1136/ard.2005.045807] [Medline: 16291811]
43. McDougall C, Hurd K, Barnabe C. Systematic review of rheumatic disease epidemiology in the indigenous populations of Canada, the United States, Australia, and New Zealand. Semin Arthritis Rheum 2017;46(5):675-686 [FREE Full text] [doi: 10.1016/j.semarthrit.2016.10.010] [Medline: 27914688]
44. Levitsky A, Brismar K, Hafström I, Hambardzumyan K, Lourdudoss C, van Vollenhoven RF, et al. Obesity is a strong predictor of worse clinical outcomes and treatment responses in early rheumatoid arthritis: results from the SWEFOT trial. RMD Open 2017;3(2):e000458 [FREE Full text] [doi: 10.1136/rmdopen-2017-000458] [Medline: 28879052]
45. Albano SA, Santana-Sahagun E, Weisman MH. Cigarette smoking and rheumatoid arthritis. Semin Arthritis Rheum 2001;31(3):146-159. [doi: 10.1053/sarh.2001.27719] [Medline: 11740796]
46. Carney RM, Steinmeyer BC, Freedland KE, Rubin EH, Rich MW, Harris WS. Baseline blood levels of omega-3 and depression remission: a secondary analysis of data from a placebo-controlled trial of omega-3 supplements. J Clin Psychiatry 2016;77(2):e138-e143 [FREE Full text] [doi: 10.4088/JCP.14m09660] [Medline: 26930527]

XSL•FO
RenderX

47. Tsuda R, Ozawa T, Kobayashi E, Hamana H, Taki H, Tobe K, et al. Monoclonal antibody against citrullinated peptides obtained from rheumatoid arthritis patients reacts with numerous citrullinated microbial and food proteins. Arthritis Rheumatol 2015;67(8):2020-2031. [doi: 10.1002/art.39161] [Medline: 25892475]
48. Dinan TG, Cryan JF. The microbiome-gut-brain axis in health and disease. Gastroenterol Clin North Am 2017;46(1):77-89. [doi: 10.1016/j.gtc.2016.09.007] [Medline: 28164854]
49. Larsen JM. The immune response to Prevotella bacteria in chronic inflammatory disease. Immunology 2017;151(4):363-374 [FREE Full text] [doi: 10.1111/imm.12760] [Medline: 28542929]
50. Scher JU, Nayak RR, Ubeda C, Turnbaugh PJ, Abramson SB. Pharmacomicrobiomics in inflammatory arthritis: gut microbiome as modulator of therapeutic response. Nat Rev Rheumatol 2020;16(5):282-292 [FREE Full text] [doi: 10.1038/s41584-020-0395-3] [Medline: 32157196]
51. Alpizar-Rodriguez D, Lesker TR, Gronow A, Gilbert B, Raemy E, Lamacchia C, et al. Prevotella copri in individuals at risk for rheumatoid arthritis. Ann Rheum Dis 2019;78(5):590-593. [doi: 10.1136/annrheumdis-2018-214514] [Medline: 30760471]
52. Hammad DBM, Hider SL, Liyanapathirana VC, Tonge DP. Molecular characterization of circulating microbiome signatures in rheumatoid arthritis. Front Cell Infect Microbiol 2019;9:440 [FREE Full text] [doi: 10.3389/fcimb.2019.00440] [Medline: 32039040]
53. Johnson K. Leaky gut syndrome. 2013. URL: https://link.springer.com/referenceworkentry/10.1007%2F978-1-4419-1698-3_29 [accessed 2025-12-10]
54. Hammad DBM, Liyanapathirana V, Tonge DP. Molecular characterisation of the synovial fluid microbiome in rheumatoid arthritis patients and healthy control subjects. PLoS One 2019;14(11):e0225110 [FREE Full text] [doi: 10.1371/journal.pone.0225110] [Medline: 31751379]
55. Arizona T. Navajo Nation community profile. 2016. URL: https://navajoprofile.wind.enavajo.org/ [accessed 2025-12-10]
56. Bureau USC. Community Facts: Arizona. 2018. URL: https://www.census.gov/quickfacts/fact/table/AZ/PST045224 [accessed 2025-12-10]
57. McGarrity-Yoder ME, Insel KC, Crane TE, Pace TWW. Diet quality and disease activity in rheumatoid arthritis. Nutr Health 2022 Dec;28(4):581-590. [doi: 10.1177/02601060211044311] [Medline: 34668452]
58. The Epidemiology and Genomics Research Program NCI. Basic steps in calculating HEI scores. Secondary Basic Steps in Calculating HEI Scores. 2018. URL: https://epi.grants.cancer.gov/hei/calculating-hei-scores.html [accessed 2025-10-12]
59. Krebs-Smith SM, Pannucci TE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, et al. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet 2018;118(9):1591-1602 [FREE Full text] [doi: 10.1016/j.jand.2018.05.021] [Medline: 30146071]
60. USDA. Healthy Eating Index (HEI). Secondary Healthy Eating Index (HEI). 2019. URL: https://www.fns.usda.gov/resource/healthy-eating-index-hei [accessed 2025-12-10]
61. Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, et al. Structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol 2003;158(1):14-21; discussion 22. [doi: 10.1093/aje/kwg091] [Medline: 12835281]
62. National Cancer Institute. Automated Self-Administered 24-Hour (ASA24) Dietary Assessment Tool. 2025. URL: https://epi.grants.cancer.gov/asa24/ [accessed 2025-12-10]
63. National Cancer Institute. ASA24 Evaluation & Validation. 2022. URL: https://epi.grants.cancer.gov/asa24/respondent/validation.html [accessed 2025-12-10]
64. Kirkpatrick S, Subar A, Douglass D, Zimmerman TP, Thompson FE, Kahle LL, et al. Performance of the automated self-administered 24-hour recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. Am J Clin Nutr 2014;100(1):233-240 [FREE Full text] [doi: 10.3945/ajcn.114.083238] [Medline: 24787491]
65. Gupta S, Mortensen MS, Schjørring S, Trivedi U, Vestergaard G, Stokholm J, et al. Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. Commun Biol 2019;2:291 [FREE Full text] [doi: 10.1038/s42003-019-0540-1] [Medline: 31396571]
66. Jubair WK, Hendrickson JD, Severs EL, Schulz HM, Adhikari S, Ir D, et al. Modulation of inflammatory arthritis in mice by gut microbiota through mucosal inflammation and autoantibody generation. Arthritis Rheumatol 2018;70(8):1220-1233 [FREE Full text] [doi: 10.1002/art.40490] [Medline: 29534332]
67. Beyer K, Zaura E, Brandt BW, Buijs MJ, Brun JG, Crielaard W, et al. Subgingival microbiome of rheumatoid arthritis patients in relation to their disease status and periodontal health. PLoS One 2018;13(9):e0202278 [FREE Full text] [doi: 10.1371/journal.pone.0202278] [Medline: 30231060]
68. Picchianti-Diamanti A, Panebianco C, Salemi S, Sorgi ML, Di Rosa R, Tropea A, et al. Analysis of gut microbiota in rheumatoid arthritis patients: disease-related dysbiosis and modifications induced by etanercept. Int J Mol Sci 2018;19(10):2938 [FREE Full text] [doi: 10.3390/ijms19102938] [Medline: 30261687]
69. Pianta A, Arvikar S, Strle K, Drouin EE, Wang Q, Costello CE, et al. Evidence of the immune relevance of prevotella copri, a gut microbe, in patients with rheumatoid arthritis. Arthritis Rheumatol 2017;69(5):964-975 [FREE Full text] [doi: 10.1002/art.40003] [Medline: 27863183]

XSL•FO
RenderX

70. Chen J, Wright K, Davis JM, Jeraldo P, Marietta EV, Murray J, et al. An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. Genome Med 2016;8(1):43 [FREE Full text] [doi: 10.1186/s13073-016-0299-7] [Medline: 27102666]
71. Maeda Y, Kurakawa T, Umemoto E, Motooka D, Ito Y, Gotoh K, et al. Dysbiosis contributes to arthritis development via activation of autoreactive T cells in the intestine. Arthritis Rheumatol 2016;68(11):2646-2661. [doi: 10.1002/art.39783] [Medline: 27333153]
72. Christensen H, Andersson A, Jorgensen S, Vogt J. 16S rRNA amplicon sequencing for metagenomics. In: Introduction to Bioinformatics in Microbiology. Cham: Springer; 2018:135-161.
73. Waugh N, Cummins E, Royle P, Kandala NB, Shyangdan D, Arasaradnam R, et al. Faecal calprotectin testing for differentiating amongst inflammatory and non-inflammatory bowel diseases: systematic review and economic evaluation. Health Technol Assess 2013;17(55):xv-xix, 1 [FREE Full text] [doi: 10.3310/hta17550] [Medline: 24286461]
74. Bruce B, Fries J. The stanford health assessment questionnaire: dimensions and practical applications. Health Qual Life Outcomes 2003;1:20 [FREE Full text] [doi: 10.1186/1477-7525-1-20] [Medline: 12831398]
75. NIAID-Sponsored Expert Panel, Boyce JA, Assa'ad A, Burks AW, Jones SM, Sampson HA, et al. Guidelines for the diagnosis and management of food allergy in the United States: report of the NIAID-sponsored expert panel. J Allergy Clin Immunol 2010;126(6 Suppl):S1-58 [FREE Full text] [doi: 10.1016/j.jaci.2010.10.007] [Medline: 21134576]
76. Health N. The Health Assessment Questionnaire (HAQ) Disability Index (DI) of the Clinical Health Assessment Questionnaire (Version 96.4). 2025. URL: https://www.niehs.nih.gov/research/resources/assets/docs/haq_instructions_508.pdf [accessed 2025-12-10]
77. Salaffi F, Cimmino MA, Leardini G, Gasparini S, Grassi W. Disease activity assessment of rheumatoid arthritis in daily practice: validity, internal consistency, reliability and congruency of the Disease Activity Score including 28 joints (DAS28) compared with the clinical disease activity index (CDAI). Clin Exp Rheumatol 2009;27(4):552-559. [Medline: 19772784]
78. Kuriya B, Schieir O, Lin D, Pope J, Boire G, Haraoui B, et al. FRI0066 The 28-joint disease activity score (DAS28) Using C-reactive protein yields lower thresholds compared to conventional DAS28 with erythrocyte sedimentation rate in early rheumatoid arthritis. Annals of the Rheumatic Diseases 2016;75:449-450. [doi: 10.1136/annrheumdis-2016-eular.2937]
79. McGarrity-Yoder M, Insel K, Crane T, Pace T. Diet quality and disease activity in rheumatoid arthritis. Nutr Health 2022;28(4):581-590. [doi: 10.1177/02601060211044311] [Medline: 34668452]
80. Bärebring L, Winkvist A, Gjertsson I, Lindqvist HM. Poor dietary quality is associated with increased inflammation in swedish patients with rheumatoid arthritis. Nutrients 2018;10(10):1535 [FREE Full text] [doi: 10.3390/nu10101535] [Medline: 30340316]
81. Richey M. Understanding Rheumatoid Arthritis Lab Tests and Results. 2018. URL: https://www.hss.edu/conditions_understanding-rheumatoid-arthritis-lab-tests-results.asp [accessed 2012-12-10]

Abbreviations

AFFQ: Arizona Food Frequency Questionnaire
ASA-24: Automated Self-Administered 24-Hour Dietary Assessment Tool
BDRA: Biopsychosocial Model of Disease Experience in Rheumatoid Arthritis
CRP: C-reactive protein
DAS-28: Disease Activity Score using 28 joint counts
ESR: Erythrocyte Sedimentation Rate
GI: gastrointestinal
HAQ-DI: Health Assessment Questionnaire–Disability Index
HEI: Healthy Eating Index
hsCRP: high-sensitivity C-reactive Protein
NIH: National Institutes of Health
QIIME 2: Quantitative Insights into Microbial Ecology version 2
RA: rheumatoid arthritis
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

Edited by J Sarvestan; submitted 30.Jun.2025; this is a non–peer-reviewed article;accepted 04.Dec.2025; published 08.Jan.2026.

Please cite as:
McGarrity-Yoder M, Cope EK, Wayment HA, Rodriguez-Pla A, Crane TE
Biopsychosocial Determinants, Diet Quality, Gastrointestinal Health, and Disease Activity in Adults With Rheumatoid Arthritis: Cross-Sectional Descriptive Study
JMIR Res Protoc 2026;15:e79889
URL: https://www.researchprotocols.org/2026/1/e79889
doi:10.2196/79889
PMID:

©Maureen McGarrity-Yoder, Emily K Cope, Heidi A Wayment, Alicia Rodriguez-Pla, Tracy E Crane. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 08.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Death Literacy and Death Competence in Undergraduate Clinical and Allied Health Education: Protocol for a Mixed Methods Study

Sora Lee[1], PhD; Kerrie Noonan[1], PhD; Andrea Grindrod[1], PhD; Sumina Shrestha[1], PhD; Niki Read[1], BA

La Trobe University, Bundoora, Victoria, Australia

Corresponding Author:
Sora Lee, PhD
La Trobe University
Plenty Rd
Bundoora, Victoria, 3086
Australia
Phone: 61 3 9479 1773
Email: sora.lee@latrobe.edu.au

Abstract

Background: End-of-life care is a critical competency for the health care workforce, yet evidence suggests that many health care professionals feel unprepared to engage with death, dying, and bereavement. Death literacy and death competence are emerging frameworks for assessing readiness to provide high-quality, compassionate care. Although validated tools exist, little is known about the preparedness of final-year undergraduate health care students in Australia. Understanding their current levels of death literacy and death competence is essential for informing curriculum design and strengthening workforce capacity.

Objective: This study aims to (1) measure death literacy and death competence among final-year students in medicine, nursing, and allied health programs in Australian universities; (2) explore students' reflections on how undergraduate training has shaped their preparedness for end-of-life care; and (3) identify educational needs and opportunities for curriculum enhancement.

Methods: A mixed methods design will be used. An online survey (15-20 minutes) will be distributed to final-year students across multiple Australian universities. The survey includes the Death Literacy Index, the Death Competency Scale, and open-ended reflection questions. Quantitative data will be analyzed using descriptive and inferential statistics (in SPSS and Stata), with subgroup comparisons across disciplines and benchmarking against national professional datasets. Qualitative responses will be analyzed thematically. In phase 2, up to 20 students will participate in 2 focus groups (60-90 minutes each). The focus groups will explore survey findings and students' perceptions of training, preparedness, and gaps. Data will be transcribed, anonymized, and analyzed thematically using NVivo.

Results: Data collection for the national survey is scheduled from September 2025 to December 2025, with an anticipated sample of 60 to 120 final-year students across medicine, nursing, and allied health disciplines. Data analysis will begin in March 2026, and findings are expected to be published in late 2026. The findings will establish baseline measures of death literacy and death competence among final-year health care students and identify strengths and gaps in current curricula. Results will be synthesized to provide actionable insights for educators and to inform future intervention studies.

Conclusions: By providing the first Australian pilot data on death literacy and death competence among final-year health care students, this study will inform curriculum development and workforce planning. The findings have the potential to enhance educational strategies, improve the preparedness of graduates for delivering end-of-life care, and contribute to the development of a death-literate health system.

International Registered Report Identifier (IRRID): DERR1-10.2196/86867

(JMIR Res Protoc 2026;15:e86867) doi:10.2196/86867

KEYWORDS

death literacy; death competency; end-of-life care; health care education; palliative care; curriculum development

XSL•FO
RenderX

Introduction

Background

Population aging and the rising prevalence of chronic, life-limiting illnesses mean that most health care professionals will inevitably care for patients at the end of life. Delivering high-quality, compassionate care in these circumstances requires not only clinical competence but also effective communication, psychosocial support, and emotional resilience. However, many medical and nursing graduates report feeling underprepared to manage death and dying in practice [1]. Therefore, the preparedness of the future workforce is a pressing concern for both health care education and health system planning.

Two complementary frameworks have become increasingly important in assessing readiness for death-related work: death literacy and death competence. Death literacy refers to the knowledge, skills, and experiential learning that enable individuals and communities to understand, navigate, and engage with death, dying, and bereavement [2]. It incorporates both practical and social dimensions, including the ability to draw on networks of support and to foster open discussion about death [3]. Death literacy is emerging as a valuable concept for professionals and the wider community, contributing to more compassionate and participatory approaches to end-of-life care.

Death competence, while related, emphasizes the personal and professional capacities required to care effectively for dying patients. These capacities include self-efficacy, communication skills, the ability to cope with emotional demands, and the ability to sustain compassionate care under pressure [4]. Higher levels of death competence have been linked to greater practitioner confidence, reduced death anxiety, and enhanced quality of care [5-8]. Validated tools such as the Death Competency Scale (DCS) [4,9] provide structured means of assessing these capacities among both practicing professionals and trainees.

Although these constructs are increasingly studied internationally, a clear gap remains in the Australian context regarding undergraduate health care education. Existing research has typically examined practicing professionals or early graduates, often highlighting deficiencies in end-of-life preparedness and recommending greater emphasis on structured education and experiential learning [1,10]. Furthermore, while there has been one published study on death literacy among psychology students [11], research on the death literacy and death competence of students at the point of graduation is lacking in an Australian context. Understanding the baseline levels of preparedness among final-year students is crucial for informing curriculum development, aligning educational strategies with workforce needs, and ultimately improving the delivery of end-of-life care.

International Perspectives

Research from the United Kingdom shows that many medical students and junior doctors feel poorly prepared to care for dying patients [12]. Surveys consistently demonstrate low levels of confidence and high levels of distress among graduates, despite the inclusion of palliative care in undergraduate curricula [13]. A cross-institutional review of UK medical schools found substantial variability in palliative care teaching, with weekly teaching hours ranging from less than 1 to more than 16 [14]. Similarly, 88% of US medical residents reported that they had little or no training during residency [15].

In European contexts, particularly Germany and the Netherlands, studies have highlighted inconsistent integration of palliative care into undergraduate education. Although awareness of the importance of end-of-life training is growing, student exposure to real clinical encounters with dying patients remains uneven [14]. Research with nursing students in Turkey suggests that high levels of spirituality may influence death literacy, hence recommending that nursing students have access to a curriculum that encourages exploration of death literacy and spirituality [16]. Nursing curricula that specifically provide both didactic and exploratory instructions along with practical experience have also been shown to improve death competence and death literacy [17].

In Asia, cultural taboos surrounding death further complicate training. In Hong Kong, for example, students often struggle with death-related conversations, influenced both by societal norms and by insufficient curriculum space for end-of-life care [18]. Similarly, research from Taiwan and mainland China suggests that health and social care students experience heightened anxiety and reduced confidence when discussing death, reinforcing the importance of context-specific training approaches [10,19].

North America has a strong tradition of conducting death education programs in undergraduate health and medical training [20,21]. Simulation-based training has been adopted to enhance student preparedness. A national study demonstrated that simulated experiences with dying patients improved both confidence and communication skills among medical trainees [22]. Despite these innovations, gaps persist in graduates' ability to manage the emotional and ethical complexities of caring for people who are dying and their families.

Together, these international studies highlight a common pattern: although death literacy and death competence are increasingly recognized as essential, systematic assessments of student preparedness at the point of graduation remain limited. Comparative data also suggest that curriculum alone is insufficient unless combined with experiential learning that is meaningful and structured and provides an opportunity for personal reflection.

Rationale and Study Aims

Despite international advances in death literacy, there are no published studies examining the skills, knowledge, and death competence among medical, nursing, and allied health students in Australia. Addressing this gap is important for three reasons: (1) educational insight to establish baseline measures of knowledge, skills, and preparedness for end-of-life care among graduating students, benchmarked against practicing professionals and the general population; (2) curriculum development to identify strengths and gaps in existing training programs, as perceived by students, and to generate evidence for guiding curriculum reform; and (3) workforce readiness to provide actionable information on how well prepared the future

XSL•FO
RenderX

workforce is to deliver compassionate and competent end-of-life care, thereby strengthening Australia's death-literate health system.

Accordingly, this study aims to accomplish the following objectives:

1. Examine the levels of death literacy and death competence among final-year undergraduate health care students
2. Explore qualitatively how students reflect on their training, preparedness, and perceived gaps in education related to death and dying
3. Generate recommendations for enhancing health care professional curricula to better prepare the future workforce for end-of-life care

By combining validated survey instruments with focus group discussions, this mixed methods protocol will generate both benchmarking data and rich insights into students' lived educational experiences. Together, these findings will provide a foundation for targeted curriculum innovations and policy strategies to build a death-literate and death-competent health care workforce in Australia.

Benchmarking of Death Literacy Index Scores

The proposed benchmarking against national professional datasets is intended as a *contextual comparison* and not a statistical equivalence test. Benchmark data will be drawn from published aggregate results in the Australian National Death Literacy Index survey (2024), which includes a subset of practicing clinicians and related professional practice studies [23].

As the Australian benchmark data were collected using the same validated Death Literacy Index (DLI) under similar online conditions, mean score comparisons will be descriptive in nature. These comparisons will be accompanied by effect-size estimates and narrative interpretation rather than formal significance testing.

Methods

Study Design

This study uses a mixed methods design combining a cross-sectional online survey (15-20 minutes) with qualitative focus group discussions (60-90 minutes each). The survey component will establish baseline measures of death literacy and death competence among final-year health care students, whereas the qualitative component will provide depth and contextual understanding of students' experiences and perceptions of preparedness for end-of-life care. Insights from both components will be analyzed to inform and strengthen curriculum development in this area.

Participants

Eligible participants are final-year undergraduate students enrolled in medicine, nursing, and allied health programs (including psychology, paramedicine, midwifery, and social work) across Australian universities. The following 9 universities have agreed to support recruitment: the University of Sydney, Western Sydney University, Charles Sturt University, the University of Newcastle, La Trobe University, Swinburne University, Victoria University, Monash University, and the University of Melbourne. The study population, sample size, and eligibility criteria are outlined as follows:

1. Survey sample size of 60-120 respondents (pilot estimate, with a 10%-20% expected response rate)
2. Focus group sample size of up to 20 participants, divided into 2 groups of 8 to 10 each, providing sufficient qualitative data for thematic analysis
3. Inclusion criteria are final-year enrollment in one of the target programs, age ≥18 years, and consent to participate
4. Exclusion criteria are enrollment in earlier years of study or outside the specified institutions

Sample Size and Analytic Scope

The initial recruitment target of 60 to 120 final-year students is appropriate for a pilot study designed to estimate variability and assess feasibility rather than to conduct fully powered inferential analyses. We acknowledge that subgroup analyses (eg, medicine vs nursing and allied health; metropolitan vs rural) may be underpowered and will therefore be treated as exploratory and descriptive rather than confirmatory. Results will focus on precision and CI estimates, with formal hypothesis testing deferred to subsequent, larger studies once effect-size estimates are available. This approach is consistent with standard pilot study methodology and the emphasis on proportionality of design to purpose in the 2023 National Statement on Ethical Conduct in Human Research [24].

Recruitment

Recruitment will occur through professional networks and university course coordinators who will circulate email invitations and postlearning management system announcements.

For survey recruitment, students will receive a link to the online questionnaire via QuestionPro, version 2024 (QuestionPro Inc). Participation will be anonymous and will not identify individual students or universities. Consent will be obtained electronically before participation, with students indicating their agreement by selecting the option to begin the survey.

For focus group recruitment, at the end of the survey, students will be invited to complete a separate expression-of-interest form. Those selected (based on program representation and availability) will be contacted by the research team, provided with participant information and consent forms, and scheduled for either an in-person or online session.

Course coordinators will not be informed of which students choose to participate, and researchers will have no direct teaching or assessment responsibilities for the eligible student cohorts.

Instrument

The online survey will include (1) demographics, including age, gender, program of study, university, and previous experience with death and dying; () the DLI a validated 29-item measure assessing knowledge, skills, and experiential learning related to death and dying [23]; (3) the Bugen Death Competence Scale

XSL•FO
RenderX

(Coping with Death Scale), short version, a validated 9-item questionnaire measuring self-efficacy, emotional coping, and preparedness for death work based on the original 30-item scale [4,10,25]; and (4) open-ended questions, comprising items to reflect on students' training, the perceived strengths and gaps, and suggestions for curricular improvements.

Focus groups will follow a semistructured guide, exploring reactions to descriptive survey findings, perceptions of preparedness for end-of-life care, experiences of training that were most or least helpful, and recommendations for educational improvements. Sessions will be facilitated by experienced palliative care researchers, audio recorded (with consent), and transcribed verbatim.

Data Analysis

Quantitative Data

Descriptive and inferential analyses will be conducted using SPSS (version 29.0; IBM Corp) and Stata (version 18.0; StataCorp LLC). Scaled mean scores on DLI, and mean scores on the DCS will be calculated, and subgroup analyses will compare results across disciplines (medicine, nursing, and allied health) and geographic location (metropolitan and rural). Results will be benchmarked against national DLI data from practicing professionals.

Qualitative Data

Open-ended survey responses and focus group transcripts will be analyzed thematically following the 6-step framework by Braun and Clarke [25]. NVivo (version 14; Lumivero) will be used to support coding and theme development. Analysis meetings will document coding decisions to ensure reflexivity and consistency.

Integration

A mixed methods synthesis will compare quantitative benchmarks with qualitative themes, identifying areas of convergence (eg, perceived strengths) and divergence (eg, unmeasured concerns about training). Illustrative quotations will be included in the reporting to enrich interpretation.

Ethical Considerations

This study has received ethics approval from the La Trobe University Human Research Ethics Committee (HEC 25370). All participants will provide informed consent before participation. For the online survey, consent will be indicated electronically via an information and consent page. For focus groups, written consent will be obtained before participation, with verbal consent reconfirmed at the beginning of each session. While there is no remuneration for survey participants, each focus group participant will be compensated Aus $50 (US $32.50). Participation is voluntary, and students may withdraw at any time before survey submission or at any point before or during focus group discussions. All data will be deidentified during transcription, and confidentiality will be maintained through secure storage on password-protected university servers. Participants will also be provided with information about available support services, given the sensitive nature of discussions around death and dying.

Timeline

As shown in Table 1, the timeline for the project spans October 2025 to February 2026, encompassing survey data collection (October 2025 to December 2025), qualitative data collection (November 2025 to February 2026), analysis (March 2026 to September 2026), and dissemination (October 2026 to December 2026).

Table 1. Study timeline and milestones.

Timeline, year (quarter)	Activity and milestones	Output and deliverables
2025 (Q3)	Ethics approval obtained	Approved study protocol
2025 (Q3)	Survey preparation (September)	Online survey finalized
2025 (Q4)	Survey data collection (October-December)	Online survey responses collected
2025 (Q4) to 2026 (Q1)	Focus group data collection (November 2025-February 2026)	Audio recordings and transcripts
2026 (Q1)	Quantitative data analysis (January-March)	Descriptive or comparative analysis
2026 (Q1)	Qualitative data analysis (March-June)	Thematic coding and analytic memos
2026 (Q2)	Integration of findings and manuscript preparation	Draft manuscripts for submission
2026 (Q3)	Dissemination	Journal submission and conference abstracts

Results

Study Timeline

Funding for this project was awarded in August 2025. Recruitment and data collection for the online survey are planned for October 2025 to December 2025, with a target sample of 60 to 120 final-year students enrolled in Australian medicine, nursing, or allied health programs. Focus group recruitment will follow immediately after the survey closes, with sessions anticipated for November 2025 and February 2026. Data cleaning and analysis will commence in March 2026, and dissemination of findings is planned for late 2026.

Quantitative Outcomes

The survey will generate descriptive statistics on death literacy and death competence levels across medicine, nursing, and allied health disciplines. Subgroup analyses are anticipated to reveal variation by discipline (eg, higher preparedness among nursing students relative to other groups, consistent with

XSL•FO
RenderX

international findings) and by geographic location (ie, metropolitan and rural). An exploratory study on benchmarking against existing professional population data from the DLI will provide insight into how well undergraduate training prepares students compared to practicing clinicians.

Qualitative Outcomes

Open-ended survey responses and focus group discussions are expected to provide rich insights into student perceptions of preparedness. Anticipated themes include strengths of current training (eg, clinical placements and communication skills workshops), gaps or barriers in curricula (eg, limited experiential exposure to dying patients and lack of reflective spaces), emotional and cultural challenges in engaging with death and dying, and recommendations generated by students for curriculum enhancement.

Integrated Insights

By combining quantitative benchmarks with qualitative reflections, the study will identify both measurable outcomes and contextualized experiences. This mixed methods integration is expected to highlight areas where curricula can be further development to prepare graduates for working with people with life-limiting illnesses.

Dissemination Plan

Findings will be disseminated through peer-reviewed journal publications and conference presentations in palliative care, health education, and public health forums. A summary of results will also be shared with participating universities to inform curriculum development and training strategies.

Long-Term Impact

Although exploratory in scope, this study will provide a baseline dataset that can inform larger-scale studies and intervention trials. Ultimately, the anticipated outcomes will contribute to building a death-literate health care workforce, better equipped to support patients, families, and communities in end-of-life contexts.

Discussion

Anticipated Findings

This study is expected to show meaningful variation in death literacy and death competency levels among final-year medicine, nursing, and allied health students. We anticipate that students with greater exposure to palliative or end-of-life education will demonstrate higher DLI and DCS scores and that gaps will emerge between discipline groups. These anticipated findings will help identify priority areas for curriculum strengthening and workforce readiness.

This pilot study addresses a clear evidence gap by assessing death literacy and death competence among final-year health care students in Australia. International research has consistently shown that many students feel ill-equipped to care for dying patients [13,14]; however, no comparable Australian data examining death literacy currently exists. By situating student preparedness within the frameworks of death literacy and death competence, this study positions end-of-life education as both a professional requirement and a public health priority.

Based on previous national surveys and training evaluation research, we expect to observe moderate overall levels of death literacy, with notable gaps in system navigation, communication readiness, and community engagement components. Differences between disciplines and training environments are also anticipated. Previous Australian DLI population studies [26] indicate wide variability in death literacy across the general population. Existing research with early-career clinicians highlights insufficient preparation for end-of-life communication and care planning. This study aims to extend that evidence by providing pregraduation baselines among future clinicians, which has not yet been mapped at scale.

The findings will provide actionable evidence for supplementary education on end-of-life care in medical, nursing, and allied health programs. They will help educators identify which aspects of training best support preparedness and where additional strategies—such as simulation, reflective practice, and structured communication skill building—may be needed [22]. Beyond the classroom, stronger graduate readiness is likely to improve the quality of communication and care for patients and families and to contribute to building death-literate health systems that normalize end-of-life discussions.

Strengths

A key strength of this project is its mixed methods approach, combining validated survey measures with qualitative focus groups to provide both breadth and depth. The multi-institutional recruitment strategy enhances the diversity of student perspectives, while the involvement of an interdisciplinary research team ensures methodological rigor.

Limitations

As a pilot, the study is not designed for large-scale generalization. The modest sample size may limit the precision of subgroup comparisons, and voluntary participation may introduce self-selection bias. Focus groups are restricted to 2 sessions, which may not capture the full spectrum of student experiences. Furthermore, given the sensitivity of the topic, some participants may choose not to disclose personal reflections. Other limitations include possible underrepresentation of certain disciplines and variability in exposure to death and dying curricula.

Future Directions

Despite these limitations, the study will establish a baseline for future, larger-scale research and intervention trials. The mixed methods framework can be expanded nationally to track changes in student preparedness over time and across institutions. Longer term, the findings may inform accreditation standards and professional development frameworks, embedding death literacy and death competence more systematically into health care education.

Conclusions

This study will provide the first national snapshot of death literacy and death competence levels among final-year Australian health care students. By establishing baseline

XSL•FO
RenderX

measures and identifying educational gaps, the project will lay the groundwork for evidence-informed curriculum design and targeted end-of-life training for the future health workforce.

Acknowledgments

The authors acknowledge the support of the Care Economy Research Institute, which provided funding through the Industry Partner Development Fund. They also thank the participating universities and course coordinators for their assistance with recruitment.

Funding

This study is supported by the Care Economy Research Institute (CERI) Industry Partner Development Fund (Aus $15,248 [US $9911.20]). The funding body has no role in the study design, data collection, analysis, interpretation, or manuscript preparation.

Data Availability

Data generated in this study will not be publicly available due to confidentiality requirements.

Authors' Contributions

SL (principal investigator) conceived the study, led the design, and drafted the protocol. KN contributed to conceptualization, provided expertise in death literacy, and supported study design and tool selection. AG offered input on palliative care expertise and will contribute to data analysis. SS contributed to study design, recruitment planning, and the qualitative analysis approach. NR contributed to conceptualization, recruitment strategy, and qualitative data collection. All authors reviewed, edited, and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Deravin L, Karin Anderson J, Croxon L. Are newly graduated nurses ready to deal with death and dying? - a literature review. Nurs Palliat Care 2016;1(4):89-93 [FREE Full text] [doi: 10.15761/npc.1000123]
2. Noonan K, Horsfall D, Leonard R, Rosenberg J. Developing death literacy. Prog Palliat Care 2016 Feb 18;24(1):31-35 [FREE Full text] [doi: 10.1080/09699260.2015.1103498]
3. Graham-Wisener L, Toner P, Leonard R, Groarke JM. Psychometric validation of the death literacy index and benchmarking of death literacy level in a representative UK population sample. BMC Palliat Care 2022 Aug 13;21(1):145 [FREE Full text] [doi: 10.1186/s12904-022-01032-0] [Medline: 35962383]
4. Robbins RA. Death competency: Bugen's coping with death scale and death self-efficacy. In: Neimeyer RA, editor. Death Anxiety Handbook: Research, Instrumentation, And Application. Milton Park, UK: Taylor & Francis; 1994.
5. Cheung JT, Au DW, Chan WC, Chan JH, Ng K, Woo J. Self-competence in death work among health and social care workers: a region-wide survey in Hong Kong. BMC Palliat Care 2018 Apr 20;17(1):65 [FREE Full text] [doi: 10.1186/s12904-018-0317-1] [Medline: 29678168]
6. An R, Zhang SF, Huang XX, Zhao XY, Cao T, Bai L, et al. Self-reported practices, competence and difficulties towards palliative care among nurses: a cross-sectional study. Eur J Cancer Care (Engl) 2022 Nov 15;31(6):e13688. [doi: 10.1111/ecc.13688] [Medline: 35971281]
7. Wang Y, Huang Y, Zheng R, Xu J, Zhang L, Zhu P, et al. The contribution of perceived death competence in determining the professional quality of life of novice oncology nurses: a multicentre study. Eur J Oncol Nurs 2023 Feb;62:102273. [doi: 10.1016/j.ejon.2023.102273] [Medline: 36709716]
8. Guo Q, Wang Y, Zheng R, Wang J, Zhu P, Wang L, et al. Death competence profiles and influencing factors among novice oncology nurses: a latent profile analysis. BMC Nurs 2024 Dec 20;23(1):939 [FREE Full text] [doi: 10.1186/s12912-024-02641-1] [Medline: 39707293]
9. Bugen LA. Coping: effects of death education. Omega 1981 Oct 01;11(2):175-183. [doi: 10.2190/jqpa-qfhw-vq7a-mcak]
10. Lin Z, Lou VW, Chan WC. Validating the self-competence in death work scale for end-of-life care volunteers. BMC Palliat Care 2025 Feb 04;24(1):35 [FREE Full text] [doi: 10.1186/s12904-025-01666-w] [Medline: 39905427]
11. Vivekananda K, Parratt C, Tucker M, Leonard R. The impact of online death literacy education on psychology students to have better end‐of‐life conversations. Aus Psychol 2021 Mar 11;55(4):410-422 [FREE Full text] [doi: 10.1111/ap.12444]
12. Moloney C, Nuzum D, Lyons M, Steele C, Forrest C, Lynch E, et al. Dying, death and doctors-in-training: survey. BMJ Support Palliat Care 2025 Oct 13:spcare-2025-005631. [doi: 10.1136/spcare-2025-005631] [Medline: 41087034]

XSL•FO
RenderX

13. Gibbins J, McCoubrie R, Forbes K. Why are newly qualified doctors unprepared to care for patients at the end of life? Med Educ 2011 Apr;45(4):389-399 [FREE Full text] [doi: 10.1111/j.1365-2923.2010.03873.x] [Medline: 21401687]
14. Weber M, Schmiedel S, Nauck F, Alt-Epping B. Knowledge and attitude of final - year medical students in Germany towards palliative care - an interinstitutional questionnaire-based study. BMC Palliat Care 2011 Nov 23;10:19 [FREE Full text] [doi: 10.1186/1472-684X-10-19] [Medline: 22112146]
15. Patterson L, Decker A, King A, Roman A, Bolkan C, Weaver RH. Preparing for the inevitable: a scoping review of death and dying education in U.S. medical schools. Acad Med 2025 Dec 03:1-42 [FREE Full text] [doi: 10.1093/acamed/wvaf003]
16. Göktuna G, Dağcan Şahin N, Gürol Arslan G. Death literacy and related factors among nursing students in Turkey: the role of spiritual well-being. J Relig Health 2025 Oct 24;64(5):4278-4292 [FREE Full text] [doi: 10.1007/s10943-025-02421-4] [Medline: 40849847]
17. Vitorino J, Querido A, Semeão I, Laranjeira C. Promoting death literacy in palliative care nursing education using narrative pedagogy. In: Proceedings of the 10th International Conference on Higher Education Advances. 2024 Presented at: HEAd '24; June 18-21, 2024; Valencia, Spain URL: https://www.researchgate.net/publication/381651030_Promoting_Death_Literacy_in_Palliative_Care_Nursing_Education_Using_Narrative_Pedagogy [doi: 10.4995/head24.2024.17338]
18. Chan WC, Tin AF. Beyond knowledge and skills: self-competence in working with death, dying, and bereavement. Death Stud 2012;36(10):899-913 [FREE Full text] [doi: 10.1080/07481187.2011.604465] [Medline: 24563946]
19. Zhu JJ, Cui XS. A study of death education for nursing students: scoping review. Nurs Commun 2024;8:e2024019 [FREE Full text] [doi: 10.53388/IN2024019]
20. Corr CA, Corr DM. Death education. In: Bryant CD, editor. Handbook of Death and Dying. Thousand Oaks, CA: SAGE Publications; 2003:292-301.
21. Dickinson GE, Mermann AC. Death education in U.S. medical schools, 1975-1995. Acad Med 1996 Dec;71(12):1348-1349. [doi: 10.1097/00001888-199612000-00018] [Medline: 9114895]
22. Schmit JM, Meyer LE, Duff JM, Dai Y, Zou F, Close JL. Perspectives on death and dying: a study of resident comfort with end-of-life care. BMC Med Educ 2016 Nov 21;16(1):297 [FREE Full text] [doi: 10.1186/s12909-016-0819-6] [Medline: 27871287]
23. Noonan K, Grindrod A, Shrestha S, Lee S, Leonard R, Johansson T. Progressing the Death Literacy Index: the development of a revised version (DLI-R) and a short format (DLI-9). Palliat Care Soc Pract 2024 Sep 17;18:26323524241274806 [FREE Full text] [doi: 10.1177/26323524241274806] [Medline: 39314871]
24. National Statement on Ethical Conduct in Human Research 2023. National Health and Medical Research Council. 2023. URL: https://www.nhmrc.gov.au/about-us/publications/national-statement-ethical-conduct-human-research-2023 [accessed 2025-12-23]
25. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
26. Galiana L, Oliver A, De Simone G, Linzitto JP, Benito E, Sansó N. A brief measure for the assessment of competence in coping with death: the Coping with Death Scale short version. J Pain Symptom Manage 2019 Feb;57(2):209-215 [FREE Full text] [doi: 10.1016/j.jpainsymman.2018.11.003] [Medline: 30447381]

Abbreviations

DCS: Death Competency Scale
DLI: Death Literacy Index

Edited by J Sarvestan; submitted 31.Oct.2025; this is a non–peer-reviewed article;accepted 12.Dec.2025; published 19.Jan.2026.

Please cite as:
Lee S, Noonan K, Grindrod A, Shrestha S, Read N
Death Literacy and Death Competence in Undergraduate Clinical and Allied Health Education: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e86867
URL: https://www.researchprotocols.org/2026/1/e86867
doi:10.2196/86867
PMID:

©Sora Lee, Kerrie Noonan, Andrea Grindrod, Sumina Shrestha, Niki Read. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Early-Life Exposures, Neurodevelopment, and Health Outcomes: Protocol for a Birth Cohort Study

Tor A Strand[1,2], Prof Dr Med; Maria Averina[3,4], MD, PhD; Kjersti Sletten Bakken[1,2], PhD; Sudha Basnet[5], MD, PhD; Yvonne Böttcher[6,7], PhD; Sandra Huber[3], PhD; Mari Hysing[8], PhD; Ingrid Kvestad[1], PhD; Torben Lüders[6], PhD; Adrian McCann[1], PhD; Dhiraj Pokhrel[9]; Suman Ranjitkar[8,10]; Arun K Sharma[2,5], MD; Merina Shrestha[5], MD; Manjeswori Ulak[5,10], MD, PhD; Ram K Chandyo[10,11], MD, PhD

[1]Innlandet Hospital Trust, Lillehammer, Norway

[2]Centre for International Health, Faculty of Medicine, University of Bergen, Bergen, Norway

[3]Department of Laboratory Medicine, University Hospital of North Norway, Tromsø, Norway

[4]Department of Clinical Medicine, The Arctic University of Norway, Tromsø, Norway

[5]Department of Pediatrics, Maharajgunj Medical Campus, Institute of Medicine, Kathmandu, Nepal

[6]Clinical Molecular Biology, EpiGen, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[7]EpiGen, Medical Division, Akershus University Hospital, Lørenskog, Norway

[8]Department of Psychosocial Science, Faculty of Psychology, University of Bergen, Bergen, Norway

[9]LEADERS Nepal, Kathmandu, Nepal

[10]Siddhi Memorial Hospital, Bhaktapur, Nepal

[11]Kathmandu Medical College, Kathmandu, Bagmati Province, Nepal

Corresponding Author:
Tor A Strand, Prof Dr Med
Innlandet Hospital Trust
Anders Sandvigsveg
Lillehammer, 2615
Norway
Phone: 47 90971086
Email: tors@me.com

Abstract

Background: Negative early-life exposures, particularly during the first 1000 days of life, may disrupt organ development and lead to lifelong negative health consequences.

Objective: Using an exposome and deep phenotyping framework, this study aims to characterize established early-life risk factors, including environmental pollutants and nutritional status during pregnancy and infancy, and identify associated short- and long-term health and developmental outcomes.

Methods: We leverage a pregnancy cohort of 800 mother-infant pairs in Bhaktapur, Nepal, nested within a randomized controlled trial (ClinicalTrials.gov: NCT03071666) that evaluated daily vitamin B12 supplementation from before 15 weeks of gestation until 6 months post partum. The primary outcomes of the original trial were linear growth and neurodevelopment at 12 months. In this follow-up, children will be evaluated up to school age to obtain more robust estimates of long-term health outcomes. Exposures include clinical, dietary, cognitive, demographic, and anthropometric variables during pregnancy and infancy, as well as analyses of environmental pollutants, inflammation, micronutrient status, and hormonal status. Outcomes comprise neurodevelopment, morbidity, mental health, vaccine responses, thyroid function, growth, body composition, lung function, and biomarkers of health and development. Our main research questions for this phase of the project are: (1) what are the most common environmental pollutants among Nepalese women and children? (2) Is there a social gradient in exposure to these pollutants? (3) To what extent are these exposures associated with nutritional status, growth, neurodevelopment, and clinical outcomes? Associations will be examined using cross-sectional, case-control, and cohort designs applying advanced statistical methods to address confounding and complex exposure patterns.

Results: Enrollment began in March 2017, and the first child was born in August of the same year. More than 90% of the original cohort (734/800, 91.8%) have provided data up to the children's fourth birthday. By December 2025, the project will have funding until July 2027, and the papers addressing the main research questions will be submitted for publication before the end of 2026.

XSL•FO
RenderX

Conclusions: This study draws on a well-characterized mother-child cohort in a South Asian setting with repeated biological samples from blood, breast milk, and urine and extensive high-quality longitudinal data on health, growth, and neurodevelopment. By integrating data on environmental exposures, nutrition, inflammation, and biological responses, the project aims to improve understanding of early-life determinants of health and inform policies and potential interventions to protect vulnerable women and children in marginalized settings. While the exploratory nature of exposome analyses entails a risk of spurious associations, careful interpretation and transparent communication of uncertainty will be prioritized.

International Registered Report Identifier (IRRID): DERR1-10.2196/78593

(JMIR Res Protoc 2026;15:e78593) doi:10.2196/78593

KEYWORDS

Nepal; Developmental Origins of Health and Disease; DOHaD; early child development; nutrition; environmental pollutants

Introduction

The Golden 1000 Days

The period from conception until the second birthday is considered crucial for normal organ development and growth [1]. For example, these first 1000 days of life constitute a period of massive and finely tuned brain development in which the brain is particularly vulnerable to adverse influences [2]. Negative exposures in early life not only may have immediate apparent effects but can also result in developmental cascades or "snowball" effects, where small deficits in brain development can become larger over time and spread to other cognitive domains [3]. Ensuring good health and growth in early life is important to secure optimal functioning throughout the life span [4].

Environmental Pollutants and Health

We are exposed to environmental contaminants through the air, water, food, and many consumer products such as cosmetics and plastics. Environmental pollutants include persistent and nonpersistent substances, for example, polychlorinated biphenyls, organochlorine, and other pesticides; polybrominated diphenyl ethers; per- and polyfluoroalkyl substances (PFAS); heavy metals; phthalates; and bisphenols [5]. Even at low levels, these exposures have raised substantial concerns because they may adversely affect our health, including growth, development, and immunity [5,6]. For example, a study from Denmark found that prenatal PFAS exposure was associated with an increased risk of children being hospitalized for infectious diseases [7]. Another study from the Faroe Islands found that each doubling of plasma PFAS concentration was associated with up to 40% reduction in the concentrations of antibodies to common childhood vaccines [8]. An association between PFAS exposure and poor vaccine response has also been found in low- and lower-middle–income country children [9], which could explain why many vaccines underperform in such settings.

Altered immunity, poor vaccine responses, and increased susceptibility to infections are particularly critical for marginalized children. Unfortunately, environmental pollutants are likely to have disproportionate effects across socioeconomic strata [10], adding to the already high burden of disease among the poorest children of the world. Exposure to many environmental pollutants has also been linked to poor development [5]. There is, for instance, ample evidence of a relationship between lead or mercury exposure and impaired neurodevelopment [11], and a recent study found a link between poor infant motor development and elevated PFAS concentration [12]. However, the evidence is conflicting, possibly due to poor study quality, such as not capturing confounding and interacting factors or poor specificity of the exposures and outcomes [5]. It is also important to remember that the effects of environmental pollutants on growth or development need not be direct. These substances can act indirectly through impaired immunity, increased inflammation, reduced nutrient availability, or disrupted hormonal pathways such as thyroid hormone production. Repeated infections, inflammation, oxidative stress, and deficiencies of specific nutrients are linked to an increased risk of poor growth and development and represent potential mediators for the effects of toxic exposures [13-15]. The simultaneous manifestation of both undernutrition and overweight is increasing, particularly in poor populations [16]. This increase in the double burden of malnutrition has been attributed to access to cheap, ultraprocessed foods; micronutrient deficiencies; and other nutrition-related factors. However, environmental pollutants are also associated with alterations in childhood growth and could contribute to this double burden of malnutrition [17].

Environmental Pollutants in Nepal

There are many studies highlighting concerns related to the presence of environmental contaminants in indoor dust [18], groundwater [19], rice [20], and soils [21] in Nepal. However, there are not many studies that have described environmental pollutant levels in humans. A recent review on environmental health in Nepal concluded that it is essential to conduct more studies on the consequences of environmental pollutants as they are of great concern and the knowledge is sparse [22].

Diet, Nutrition, and Environmental Pollutants

Inadequate nutrition has consequences for child health and development [23]. Micronutrient deficiencies may interact with and aggravate the impacts of environmental pollutants. This can be exemplified by 2 recent studies suggesting that folic acid mitigates the adverse health effects of air pollutants and pesticides in relation to neurodevelopmental disorders [24,25]. Moreover, the risk of thyroid disease from iodine deficiency increases in the presence of endocrine disruptors such as perchlorate, thiocyanate, and nitrate [26]. Crucially, impaired thyroid function in early life has catastrophic consequences for the developing brain. Excess intake of nutrients also poses a health threat. For example, iodine excess results in the same

XSL•FO
RenderX

health consequences as iodine deficiency, and high intakes of macronutrients increase the risk of overweight and obesity.

Other Factors Affecting Child Health and Development

In addition to the aforementioned factors, many poverty-related variables can influence a child's development. These include but are not limited to a child's home environment, maternal psychological distress, and responsive stimulation [27]. Many of these factors can act as important confounders of the association between exposures and outcomes and, accordingly, are essential to capture when estimating the effects of environmental pollutant exposure.

As described above, exposure to environmental pollutants is of great concern, especially during early life. As new pollutants continuously enter our environment, describing current exposure levels among pregnant and lactating women and infants and the consequences of such exposures is needed for targeted prevention strategies. Obtaining data from studies or biobanks within health facilities would enable the monitoring of changes and rapid identification of new and emerging concerns. Recent reviews have also called for more research on the relationships between cocktails of environmental pollution, nutritional status, and health in vulnerable populations such as children and pregnant women in an "exposome approach" [28]. Novel epidemiological and data-driven statistical methods could enable simultaneous assessment of the impact of multiple environmental pollutants and nutritional factors on child health and development.

Aims and Project Objectives

The main objectives are to identify current exposure to environmental pollutants during pregnancy and infancy and investigate the consequences of these for infant health and development. Our three hypotheses are as follows: (1) infants and pregnant women in Nepal are exposed to several environmental pollutants (hypothesis 1); (2) exposures to environmental pollutants are associated with other risk factors for poor health (hypothesis 2); and (3) pregnant women and infants face several adverse exposures that, alone or in combination, may have consequences for their health and development.

Hypothesis 1 encompasses the following research questions (RQs):

- RQ 1a: what are the most common environmental pollutants found in Nepalese women?
- RQ 1b: what are the most common environmental pollutants found in Nepalese infants?
- RQ 1c: are the environmental pollutants found in the mothers also identified in their children?

Hypothesis 2 encompasses the following RQs:

- RQ 2a: is there a social gradient in the exposure to environmental pollutants?
- RQ 2b: to what extent do exposures to environmental contaminants tend to co-occur?
- RQ 2c: to what extent do exposures to environmental contaminants occur with other risk factors?

Hypothesis 3 encompasses the following RQs:

- RQ 3a: is exposure to environmental pollutants related to executive functioning and general abilities?
- RQ 3b: is exposure to environmental pollutants associated with poor vaccine response in children?
- RQ 3c: are environmental pollutants related to thyroid function in mothers and children?
- RQ 3d: is exposure to environmental pollutants associated with stunting and underweight?
- RQ 3e: to what extent is exposure to environmental pollutants associated with maternal and child overweight?
- RQ 3f: are environmental pollutants related to impaired lung function at school age in children?
- RQ 3g: is there a recognizable gradient of lung function in children exposed to increasing concentration of environmental pollutants?

Methods

Setting, Participants, and Enrollment Procedures

The Exposome Concept

The exposome concept includes the integration of (1) the measurement of several exposures in the human external environment, (2) the measurement of a broad spectrum of biological responses in the human body's internal environment, and (3) addressing the exposome's dynamic life cycle nature [29]. The exposome approach may lead to a greater understanding of the role of environmental risk factors in child health, growth, and development. Indeed, a recently completed population-based exposome study using data from 6 European countries highlighted the need for more well-characterized cohorts to measure the association between toxic exposures and health outcomes [30].

Because of the high plasticity and vulnerability of the brain during the first 1000 days of life, measurements of development during this period are less reliable than those later in life. Furthermore, for specific domains and functions, such as scores on general abilities (IQ), the predictive ability for future academic achievements increases with age, and scores should preferably be obtained after 4 to 5 years. Taken together, a longitudinal design that investigates exposures during the critical first 1000 days and neurodevelopmental outcomes during preschool is warranted to capture seemingly small but important consequences for the brain. It is also important to measure several potentially mediating factors such as inflammation and infections to not overlook possible mediators or important consequences of environmental exposures.

The RQs will be addressed in a pregnancy cohort with 800 mother-infant pairs in Bhaktapur, Nepal. This cohort is nested within a randomized controlled trial on daily vitamin B12 supplementation during pregnancy. The latter study was designed to address other objectives in a cohort design until the age of 2 years in the children. Details on data collection, funding, and primary objectives have been published elsewhere [31,32].

XSL•FO
RenderX

The RQs for hypotheses 1 and 2 will be addressed using cross-sectional designs, whereas for the RQs for hypothesis 3, we will use cohort and case-control designs (Figure 1). In the cohort design, we will use the entire sample and fewer but targeted biomarkers (see below). In the case-control design, we will additionally analyze the metabolome and include a broader panel of environmental pollutants in 100 high-risk infants and 100 randomly selected controls. The definition of high-risk infants will be determined using neurodevelopmental test scores and growth indexes. A random process will be used to select the controls. In longitudinal analyses, we will undertake the same biomarker assessments repeatedly in a subset of 50 randomly selected mother-infant pairs from the case-control sample (Figure 1). This analysis aims to assess the extent to which biomarkers vary between mother-infant pairs and whether their concentrations differ across pregnancy, infancy, and the postpartum period. At 7 years of age, we will collect new blood and urine samples from the children and their mothers.

Figure 1. Outline of the study design in a cohort study in Bhaktapur, Nepal. Bayley-III: Bayley Scales of Infant and Toddler Development, third edition; CC: case-control; SDQ: Strengths and Difficulties Questionnaire; TIMP: Test of Infant Motor Performance; WPPSI: Wechsler Preschool and Primary Scale of Intelligence, fourth edition.

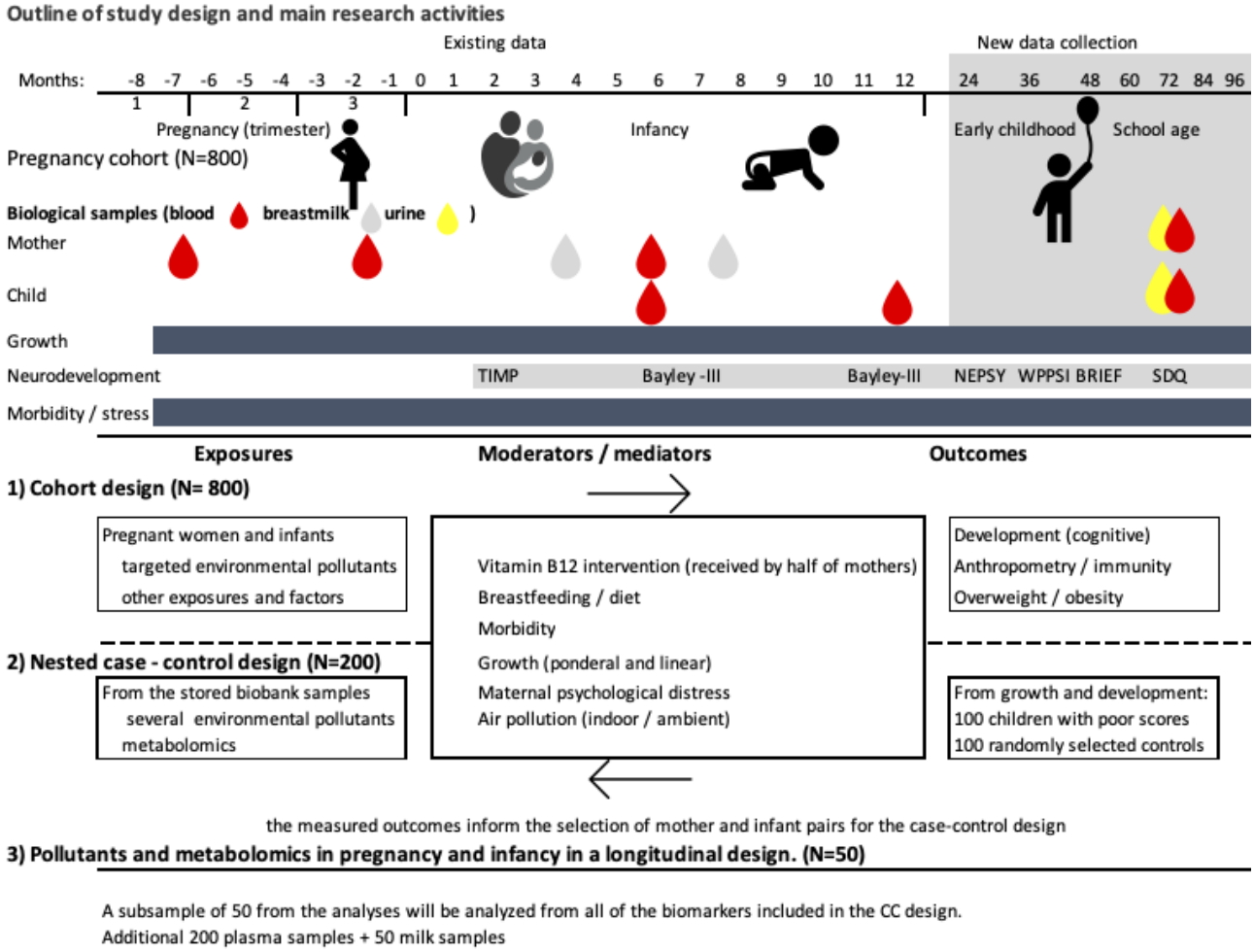

Biospecimen Collection

The study is being conducted in the Bhaktapur district in the Kathmandu Valley, Nepal. Bhaktapur is a periurban agriculturally based community inhabited by people from all over Nepal and the most densely populated district in Nepal due to recent migration. At the field site, we have access to a fully functional field office staffed with relevant health personnel and field workers. The blood and urine samples, which are collected for analysis of environmental pollutants, are processed immediately and stored at −70 °C locally, during transport, and in biobanks in Norway.

The study population has already been well described, including variables related to socioeconomic status; mental health; maternal cognitive abilities; vaccination status; air pollution; intrauterine growth; birth outcomes; child, maternal, and paternal anthropometry; and others.

Mothers and Infants

Venous blood samples were collected from the mothers at enrollment, 8 months of gestation, and 6 months after delivery and from the infants at 6 and 12 months. Approximately 3 to 4 mL of blood were collected into vials containing ethylenediaminetetraacetic acid. The hemoglobin concentration was estimated immediately using the HemoCue Hb 201 system calibrated per the manufacturer's guidelines. Blood was centrifuged at approximately 2000 g at room temperature for 10 minutes, and plasma was stored at −70 °C at the field site and always transported on dry ice. We also collected blood samples from the mother-child pairs when the children were

XSL•FO
RenderX

aged 7 years. These samples were analyzed for glycated hemoglobin and hemoglobin immediately following collection and stored at −60 or −80 °C in our biobank in Nepal.

Analyses of Environmental Exposures

Fully validated analytical methods for analyzing environmental pollutants, heavy metals, and trace elements will be applied in the Environmental Pollutant Laboratory at the University Hospital of North Norway. This laboratory uses automated sample preparation methods that can handle small sample volumes for a broad range of analyses. The instrumental analysis will be performed using gas, ultrahigh-performance liquid, and supercritical fluid chromatography combined with different ionization techniques (ie, atmospheric pressure, electrospray, and UniSpray ionization) and coupled to tandem mass spectrometers. The analytical methods for persistent organic pollutants and PFAS in plasma are well established [33]. A multielement method (>30 elements) based on dilution and direct introduction to the inductively coupled plasma mass spectrometer for plasma, whole blood, and urine will be used at the University Hospital of North Norway. This analysis will yield both toxic elements (such as heavy metals) and micronutrients.

Targeted Metabolomics

We will choose targeted over untargeted metabolomics because untargeted or discovery-based metabolomics focuses on global detection and relative quantitation of metabolites in a sample and, as a result, typically captures only abundant metabolites. The targeted metabolic profiling will include both direct and functional markers of nutritional status (including B vitamins such as B6 and folate; total homocysteine; and the fat-soluble vitamins A, D, E, and K), as well as other nutrients, including essential amino acids. Central to investigating the effect of exposure to environmental pollutants on growth, development, and health outcomes, we will also target a wide array of complementary biomarkers and proxy markers linked to inflammation (eg, C-reactive protein, serum amyloid A, calprotectin, neopterin, the kynurenine-to-tryptophan ratio, and the PAr index of vitamin B6 catabolism) [34-36]. In addition, we will target metabolites involved in defined metabolic pathways related to nucleotide synthesis, methylation reactions, mitochondrial function, energy homeostasis, and amino acid metabolism, including one carbon metabolism, the tricarboxylic acid cycle, and the kynurenine pathway. The targeted metabolomic analyses will be conducted using a methodology based on gas chromatography–tandem mass spectrometry, liquid chromatography–tandem mass spectrometry, and matrix-assisted laser desorption–ionization time-of-flight mass spectrometry (proteins and their proteoforms) and microbiological assays at Bevital AS.

Gut Microbiome–Derived Metabolites

Gut-derived metabolites in maternal and infant blood will also be measured. These include tryptophan metabolites, the indoles, and short-chain fatty acids, as well as choline derivatives that are produced by gut bacteria. These microbiome-derived metabolites play a key role in immune function, brain development, and metabolic health [37-39] and are invaluable for understanding how maternal and infant gut microbiomes shape long-term health and disease trajectories. Again, targeted metabolomic analyses will be used to reliably detect and quantitate systemic concentrations of these low-abundance metabolites using gas chromatography–tandem mass spectrometry and liquid chromatography–tandem mass spectrometry methods at Bevital AS.

Telomere Length

Genomic DNA were isolated from whole-blood samples from both mothers and infants (all time points) using an in-house–developed method for automated isolation on the QIAsymphony system (Qiagen). Telomere length was measured using real-time polymerase chain reaction [40]. The 36B4 was used as a reference gene as described previously [41]. Analyses were carried out on a QuantStudio 6 Pro real-time polymerase chain reaction system (Applied Biosystems). Telomere length analyses were performed at the Department of Clinical Molecular Biology (EpiGen), a joint research facility between Akershus University Hospital and the University of Oslo.

Neurodevelopment: Early Child Development, Cognitive Development, and Mental Health

Figure 1 provides a brief account of neurodevelopmental assessments already conducted in the cohort and assessments up to the age of 7 years. Neurodevelopmental outcomes were assessed using state-of-the art and age-appropriate instruments. At approximately 3 months, the Test of Infant Motor Performance was used to measure infant motor development. The Bayley Scales of Infant and Toddler Development, third edition (BSID-III), were used to assess early child development at 6, 12, 24, and 36 months. The BSID-III have previously been shown to have good psychometric properties at our field site and are widely used worldwide to assess development in young children [42].

For the follow-up, we have also included assessment of specific cognitive functions such as executive functioning through the NEPSY neuropsychological test battery [43] and the parent-reported Behavior Rating Inventory of Executive function–Preschool Version at 4 and 5 years [44]. We also included the Wechsler Preschool and Primary Scale of Intelligence, fourth edition, to measure intellectual functioning in the children at 4 and 5 years. As a measure of cognitive function, we also included a digital tool when the children are aged 6 years. The Developmental Assessment on an E-Platform is a cognitive ability assessment tool for children covering 6 cognitive domains presented to children on a tablet with a gamified approach.

Mental health outcomes are assessed using the BSID-III Social-Emotional scale via interview with caregivers when the children are aged 6 to 36 months. In the follow-ups at 4, 5, and 7 years, the parent-reported Strengths and Difficulties Questionnaire, a mental health screening questionnaire for children, has been included, and at 7 years, the Autism Spectrum Quotient [45], a screening tool for autism spectrum disorders, is included. The latter screening tool is included as autism spectrum disorders have previously been found to be associated with neurotoxic exposures [34]. In this study, we have also

XSL•FO
RenderX

included a child health sleep questionnaire to measure the quality and quantity of child sleep at 4 and 7 years.

Maternal Mental Health, Sleep, and Functioning

From the cohort, there are repeated measures of maternal psychological distress using the Self-Reporting Questionnaire–20 and the Reduced Tension Checklist and of maternal sleep through the Bergen Insomnia Scale. We have included the Raven Progressive Matrices as a measure of maternal nonverbal intellectual functioning. As measures of the quality of parental caregiving, we have included the Home Observation for Measurement of the Environment inventory and the Caregiver Knowledge of Child Development Inventory. These measures will be included in the statistical analyses as potential moderators or mediators. A comprehensive list of the variables included in the project is provided in our study web page [46].

Vaccine Response

Vaccine coverage in the selected population is excellent, and our previous studies have shown that nearly all children have received their scheduled vaccine doses. In the parent project, all children also received an oral rotavirus vaccine at 6 and 10 weeks. We will measure plasma anti–rotavirus immunoglobulin A antibodies using an enzyme-linked immunosorbent assay in samples drawn at 6 and 12 months. We will also include other vaccine responses (such as measles and polio) as outcomes using enzyme-linked immunosorbent assay methods. For these analyses, we will use the clinical laboratories of the partner institutions (see below) in Nepal.

Thyroid Function

The markers of maternal thyroid function during early pregnancy have already been measured at the Research Laboratory of the Department of Clinical Biochemistry, Institute of Medicine, Tribhuvan University, Kathmandu, Nepal. Plasma concentrations of free triiodothyronine, free tetraiodothyronine, total triiodothyronine, total tetraiodothyronine, thyroid-stimulating hormone, thyroglobulin, thyroglobulin antibody, thyroid peroxidase antibody, thyroid-stimulating hormone receptor antibodies, and total human chorionic gonadotrophin were analyzed using a Maglumi 800 chemiluminescence auto-analyzer (Snibe Diagnostic).

Anthropometry (Growth and Body Composition)

Dedicated staff have measured length, height, weight, middle upper arm circumference, and head circumference in the children monthly for the first year of life and biannually thereafter. We will obtain anthropometric measurements biannually at least until 84 months after birth.

Lung Function

Lung function will be assessed in the original child cohort using spirometry and oscillometry. All children are of school age and can perform reliable spirometry, which will be conducted using open-circuit spirometers in a child-friendly setting with infection control measures. Incentive spirometers will encourage best efforts, and calibration will be done daily with a 3-L syringe. Testing will be delayed if children recently took bronchodilators (6 hours), had a large meal (2 hours), or exercised (30 minutes). Children will wear light, nonrestrictive clothing; sit comfortably on a stool; and use a nose clip.

Each child will undergo pre- and postbronchodilator oscillometry followed by spirometry. Oscillometry will involve 5 sequences of tidal breathing (≥30 seconds). Following oscillometry, the child will complete at least 3 acceptable, repeatable forced expiratory maneuvers for spirometry, with rest breaks as needed. Up to 10 attempts may be made unless the child tires. After the prebronchodilator test, 400 µg of salbutamol will be administered via metered-dose inhaler using a spacer. The sequence will be repeated 15 to 20 minutes later for postbronchodilator measures.

Spirometry will measure forced vital capacity (FVC), forced expiratory volume in 1 second (FEV1), ratio of FEV1 to FVC (percentage), and forced expiratory flow between 25% and 75% of FVC. A minimum of 3 acceptable, reproducible efforts are required. Spirograms must be artifact free, show a good start and end, and meet repeatability criteria (eg, FVC and FEV1 within 150 mL). Measurements will be made under body temperature, pressure, and saturated conditions and interpreted using Global Lung Function Initiative references following American Thoracic Society and European Respiratory Society guidelines.

Impulse oscillometry will use sound waves over tidal breathing with the child seated, the head neutral, the nose clipped, and the cheeks supported. Children will perform 5 sequences of 30-second tidal breathing. Total respiratory impedance, resonant frequency, respiratory system resistance, respiratory system reactance, and reactance area will be measured and interpreted based on published references for oscillometry in children [47,48].

Data Management

Data management encompasses capture, storage, sharing, and later use of data. Three individuals in Nepal work solely with data integrity during data collection for this project.

Data Capture

Data were captured electronically using tablets through the iFormBuilder system (Zerion Software [49]) and case report forms (paper forms). All paper forms are entered twice into a Microsoft Access database by 2 computer operators, and their entries are compared concurrently for errors.

Data Storage

Data are stored on secure servers at the field site in Nepal and in Norway. Information that could identify participants is stored on separate servers to ensure their privacy. Data used for statistical analyses are always deidentified. Person-identifiable data are stored on a separate server in Nepal with limited and restricted access. Only the data managers and principal investigators have access to this server.

Statistical Approach and Sample Size

Within the project group, we maintain a registry of identified RQs that can be addressed within the project. Before an RQ is

XSL•FO
RenderX

addressed, we have defined the following steps that should be followed:

1. Define and spell out the RQ.
2. Define a core writing group (CWR) and a list of potential authors and suggest the key positions of the list of authors (first, second, last, and corresponding).
3. The CWR will prepare the first version of the plan of analysis (PoA), which will be circulated and agreed upon among all authors. This plan will define the main exposure and outcome variables and their expected distributions (eg, continuous and categorical). It will also include a description of other variables to be included in the statistical models or of the process for deciding which variables to include. The PoA will also include information on how missing data will be handled and whether and how sensitivity analyses will be undertaken.
4. When the PoA is finalized, it will be uploaded to a suitable repository, such as Zenodo [50] or a clinical trial registry.
5. The CWR will undertake the analyses and present the results and a summary (abstract) of the work for the entire group.
6. The CWR will lead the writing of the final report.
7. Once all suggested authors have had the opportunity to contribute, we will define the final authorship list.

During this process, we will follow established guidelines, such as the Strengthening the Reporting of Observational studies in Epidemiology guidelines and the Vancouver guidelines, when defining who to include in different scientific manuscripts.

To estimate the association between the selected exposures and outcomes, we will use different regression methods. We will explore dose-response relationships using generalized additive models, cubic splines, and similar techniques and measure potential mediation using structural equation models. In addition to the more traditional regression-based approaches, we will evaluate and develop machine learning approaches to address our RQs. The major advantage of machine learning is the ability to deal with the plethora of exposures and the extensive number of combinations; a feature that traditional regression-based methods struggle with. However, the resulting machine learning models can be more challenging to interpret. To develop models that are both interpretable and high performing, we will consider intrinsically interpretable models such as least absolute shrinkage and selection operator regression and provide post hoc explanations for more complex methods such as neural networks.

For most of the RQs, we will have a sample size of 800 participants. This way, we will be able to describe relevant prevalences and exposure-outcome associations precisely. For each RQ, we will calculate the expected precision of the effect measure estimates when preparing the analysis plans. Limited statistical power and precision will be adequately discussed when relevant.

Ethical Considerations

All phases, including the follow-up described in this manuscript and the measurement of leucocyte telomere length, environmental pollutants, and metabolomics, have been approved by the Nepal Health Research Council (253/2016) and the Regional Committee for Medical and Health Research Ethics in Norway (project 2016/1620/REK vest). After thorough information was provided to the parents, we obtained written informed consent or a thumbprint from illiterate individuals (in the presence of an impartial witness). We keep a registry of all women and witnesses if this is required. Personal identifiable information (such as names, addresses, and dates of birth) will be stored separately from research data on a secure server not connected to the internet and in a locked cabinet with restricted access. Participants may withdraw from the study at any time and request that their data be deleted. The study team will ensure that such requests are complied with in accordance with applicable regulations. Personal identifiable data will not be shared with third parties beyond authorized data handlers involved in the study.

In the consent form, we included a question regarding the use of data for purposes other than the original outcomes of the randomized controlled trial and the use of the biobank for other analyses. Only those who opted in for this purpose will be included in subsequent analyses. When the objectives of the secondary analyses are not covered in the original consent form, we request an updated consent form that includes the new objectives. As of December 2025, most of the participants have completed 4 consent forms. We keep a record of what all participants have consented to since the project started in 2017.

The study participants in this phase of the project were compensated for their travel time to the study site. As part of the main study, participants received a small gift once a year. The maximum value of these gifts is US $50.

Results

As of December 2025, we were still following up with 91.8% (734/800) of the original mother-infant pairs included in the study, and the study is funded until July 2027. We have completed 662, 430, 190, and 20 for the 5-, 6-, 7-, and 8-year follow-ups, respectively. We have also analyzed metabolomic responses, trace element and heavy metal concentrations, and PFAS concentrations from baseline plasma samples in all enrolled women. By May 2026, all children in the cohort will have reached their fifth birthday, and we will have complete follow-up data on key outcomes and exposures for >90% of these participants. The timing of the different phases of the study is shown in Figure 2.

XSL•FO
RenderX

Figure 2. Outline of the different phases of a pregnancy cohort study in Nepal. The lighter shade of gray indicates planned activities (preadolescence and beyond). PFAS: per- and polyfluoroalkyl substances.

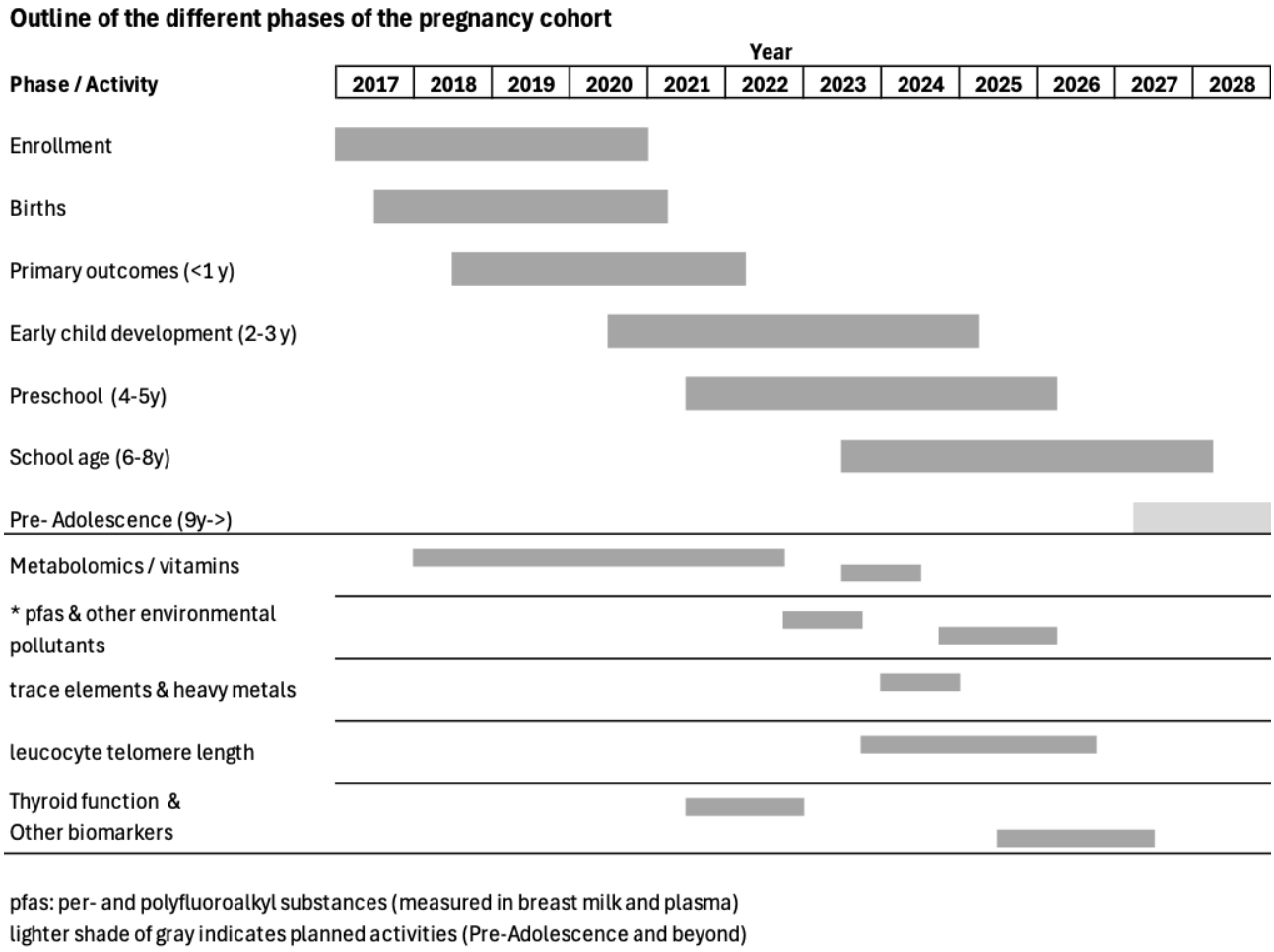

Discussion

This well-characterized mother-child cohort with extensive longitudinal data and repeated biological samples will serve as a valuable resource for studying exposure to environmental pollutants and nutritional status in a densely populated district in Nepal that includes both rural and urban residents. We will also assess the relationships between these exposures and various health-related outcomes. For instance, data from this project have been used to describe B vitamin and selenium status and examine the relationship between selenium levels and neurodevelopment [51,52].

While the broad scope and detailed phenotyping approach offer valuable opportunities to identify potential determinants of long-term outcomes, many analyses are inherently exploratory and involve multiple correlated exposures and outcomes. As a result, the observed associations should be viewed as hypothesis generating rather than causal. We will focus on effect sizes and uncertainty intervals instead of significance testing with dichotomous *P* values. As *P* values depend on effect size, variability, measurement error, and sample size, they have limited usefulness for communicating uncertainty. Additionally, our findings from this cohort may not directly apply to other populations or settings with different exposure profiles, nutritional patterns, or health systems.

Even when following well-defined and prespecified analysis plans, the large number of associations that can be tested may lead to false positives and spurious results. At the same time, even sample sizes of up to 800 might be too small to detect clinically meaningful associations with the necessary precision. Additionally, variables differ greatly in how effectively they measure what they are intended to capture. For example, we are confident in our anthropometric measurements, including the methods for obtaining them and what they represent; however, many factors can influence our neurodevelopmental assessments, such as the children's fatigue level, ongoing illnesses, mood, temperament, and current home environment. These factors increase the uncertainty about the interpretation of the findings, which we will carefully discuss when we communicate our results.

The dissemination of results will be conducted carefully and supported by a dedicated communication work package, with designated personnel responsible for engaging media and lay audiences. This communication strategy ensures that both the identified associations and their inherent uncertainties, as well as the contextual nature of the findings, are conveyed accurately, transparently, and without overstating their significance.

XSL•FO
RenderX

Funding

This study was funded by the Research Council of Norway (336566) and Innlandet Hospital Trust (0303-150504 and 0303-150403).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author or the emails below on reasonable request. To meet ethical requirements for the use of confidential patient data, requests must be approved by the Nepal Health Research Council and the Norwegian Regional Ethics Committee. Requests for data can also be sent to the Child Health Research Project, Department of Child Health, Institute of Medicine, Tribhuvan University (chrp2015@gmail.com), or the Department of Global Public Health and Primary Care at the University of Bergen (post@igs.uib.no). Upon completion of the study, we will anonymize the data and share the code and metadata on appropriate servers for further use in accordance with national rules in Norway and Nepal.

Authors' Contributions

TS, MH, IK, RC, SR, SB, MS, MU, and DP contributed to overall project planning and execution. YB, TL, and TS contributed to description and planning of leucocyte telomere length measurements and DNA extraction. TS, MH, IK, RC, MU, SR, and MS contributed to neurodevelopmental outcomes and measurements related to mental health. TS, KSB, SH, and MA contributed to measurements of environmental exposures. AM and TS contributed to the metabolomic panel. KSB, AM, AS, SH, MA, and TS contributed to measurements of other biomarkers. AS, TS, and SB contributed to measurement of lung function.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review comments from the Research Council of Norway (Forskningsrådet).
[PDF File (Adobe PDF File), 118 KB - resprot_v15i1e78593_app1.pdf]

References

1. Keunen K, Counsell SJ, Benders MJ. The emergence of functional architecture during early brain development. Neuroimage 2017 Oct 15;160:2-14. [doi: 10.1016/j.neuroimage.2017.01.047] [Medline: 28111188]
2. Fox SE, Levitt P, Nelson 3rd CA. How the timing and quality of early experiences influence the development of brain architecture. Child Dev 2010;81(1):28-40 [FREE Full text] [doi: 10.1111/j.1467-8624.2009.01380.x] [Medline: 20331653]
3. Masten AS, Cicchetti DJ. Developmental cascades. Dev Psychopathol 2010 Aug 24;22(3):491-495. [doi: 10.1017/S0954579410000222] [Medline: 20576173]
4. Barker DJ. The developmental origins of chronic adult disease. Acta Paediatr Suppl 2004 Dec;93(446):26-33. [doi: 10.1111/j.1651-2227.2004.tb00236.x] [Medline: 15702667]
5. Vrijheid M, Casas M, Gascon M, Valvi D, Nieuwenhuijsen M. Environmental pollutants and child health-a review of recent concerns. Int J Hyg Environ Health 2016 Jul;219(4-5):331-342. [doi: 10.1016/j.ijheh.2016.05.001] [Medline: 27216159]
6. von Holst H, Nayak P, Dembek Z, Buehler S, Echeverria D, Fallacara D, et al. Perfluoroalkyl substances exposure and immunity, allergic response, infection, and asthma in children: review of epidemiologic studies. Heliyon 2021 Oct;7(10):e08160 [FREE Full text] [doi: 10.1016/j.heliyon.2021.e08160] [Medline: 34712855]
7. Dalsager L, Christensen N, Halekoh U, Timmermann CA, Nielsen F, Kyhl HB, et al. Exposure to perfluoroalkyl substances during fetal life and hospitalization for infectious disease in childhood: a study among 1,503 children from the Odense Child Cohort. Environ Int 2021 Apr;149:106395 [FREE Full text] [doi: 10.1016/j.envint.2021.106395] [Medline: 33508532]
8. Grandjean P, Andersen EW, Budtz-Jørgensen E, Nielsen F, Mølbak K, Weihe P, et al. Serum vaccine antibody concentrations in children exposed to perfluorinated compounds. JAMA 2012 Jan 25;307(4):391-397 [FREE Full text] [doi: 10.1001/jama.2011.2034] [Medline: 22274686]
9. Timmermann CA, Jensen KJ, Nielsen F, Budtz-Jørgensen E, van der Klis F, Benn CS, et al. Serum perfluoroalkyl substances, vaccine responses, and morbidity in a cohort of Guinea-Bissau children. Environ Health Perspect 2020 Aug;128(8):87002 [FREE Full text] [doi: 10.1289/EHP6517] [Medline: 32772733]
10. do Nascimento FP, de Almeida MF, Gouveia N. Individual and contextual socioeconomic status as effect modifier in the air pollution-birth outcome association. Sci Total Environ 2022 Jan 10;803:149790. [doi: 10.1016/j.scitotenv.2021.149790] [Medline: 34481165]
11. Saavedra S, Fernández-Recamales Á, Sayago A, Cervera-Barajas A, González-Domínguez R, Gonzalez-Sanz JD. Impact of dietary mercury intake during pregnancy on the health of neonates and children: a systematic review. Nutr Rev 2022 Jan 10;80(2):317-328. [doi: 10.1093/nutrit/nuab029] [Medline: 33954792]

XSL•FO
RenderX

12. Varsi K, Torsvik IK, Huber S, Averina M, Brox J, Bjørke-Monsen AL. Impaired gross motor development in infants with higher PFAS concentrations. Environ Res 2022 Mar;204(Pt D):112392 [FREE Full text] [doi: 10.1016/j.envres.2021.112392] [Medline: 34800539]
13. Kosek MN, MAL-ED Network Investigators. Causal pathways from enteropathogens to environmental enteropathy: findings from the MAL-ED birth cohort study. EBioMedicine 2017 Apr;18:109-117 [FREE Full text] [doi: 10.1016/j.ebiom.2017.02.024] [Medline: 28396264]
14. Kvestad I, Taneja S, Hysing M, Kumar T, Bhandari N, Strand TA. Diarrhea, stimulation and growth predict neurodevelopment in young North Indian children. PLoS One 2015;10(3):e0121743 [FREE Full text] [doi: 10.1371/journal.pone.0121743] [Medline: 25826376]
15. Jiang NM, Tofail F, Ma JZ, Haque R, Kirkpatrick B, Nelson CA, et al. Early life inflammation and neurodevelopmental outcome in Bangladeshi infants growing up in adversity. Am J Trop Med Hyg 2017 Sep;97(3):974-979 [FREE Full text] [doi: 10.4269/ajtmh.17-0083] [Medline: 28722635]
16. Popkin BM, Corvalan C, Grummer-Strawn LM. Dynamics of the double burden of malnutrition and the changing nutrition reality. Lancet 2020 Jan 04;395(10217):65-74 [FREE Full text] [doi: 10.1016/S0140-6736(19)32497-3] [Medline: 31852602]
17. Vrijheid M, Fossati S, Maitre L, Márquez S, Roumeliotaki T, Agier L, et al. Early-life environmental exposures and childhood obesity: an exposome-wide approach. Environ Health Perspect 2020 Jun;128(6):67009 [FREE Full text] [doi: 10.1289/EHP5975] [Medline: 32579081]
18. Chandra Yadav I, Devi NL, Li J, Zhang G. Examining the role of total organic carbon and black carbon in the fate of legacy persistent organic pollutants (POPs) in indoor dust from Nepal: implication on human health. Ecotoxicol Environ Saf 2019 Jul 15;175:225-235. [doi: 10.1016/j.ecoenv.2019.03.048] [Medline: 30903878]
19. Pokhrel D, Bhandari BS, Viraraghavan T. Arsenic contamination of groundwater in the Terai region of Nepal: an overview of health concerns and treatment options. Environ Int 2009 Jan;35(1):157-161. [doi: 10.1016/j.envint.2008.06.003] [Medline: 18723222]
20. Wang L, Han J, Katuwal HB, Xia P, Xu X, Feng X, et al. Occurrence of total mercury and methylmercury in rice: exposure and health implications in Nepal. Ecotoxicol Environ Saf 2021 Nov 22;228:113019 [FREE Full text] [doi: 10.1016/j.ecoenv.2021.113019] [Medline: 34823213]
21. Pokhrel B, Gong P, Wang X, Chen M, Wang C, Gao S. Distribution, sources, and air-soil exchange of OCPs, PCBs and PAHs in urban soils of Nepal. Chemosphere 2018 Jun;200:532-541. [doi: 10.1016/j.chemosphere.2018.01.119] [Medline: 29501890]
22. Pathak KP, Gaire T. Nepal: country report on children's environmental health. Rev Environ Health 2020 Mar 26;35(1):53-56. [doi: 10.1515/reveh-2019-0086] [Medline: 32112705]
23. Georgieff MK, Ramel SE, Cusick SE. Nutritional influences on brain development. Acta Paediatr 2018 Aug 22;107(8):1310-1321 [FREE Full text] [doi: 10.1111/apa.14287] [Medline: 29468731]
24. Schmidt RJ, Kogan V, Shelton JF, Delwiche L, Hansen RL, Ozonoff S, et al. Combined prenatal pesticide exposure and folic acid intake in relation to autism spectrum disorder. Environ Health Perspect 2017 Sep 08;125(9):097007 [FREE Full text] [doi: 10.1289/EHP604] [Medline: 28934093]
25. Goodrich AJ, Volk HE, Tancredi DJ, McConnell R, Lurmann FW, Hansen RL, et al. Joint effects of prenatal air pollutant exposure and maternal folic acid supplementation on risk of autism spectrum disorder. Autism Res 2018 Jan;11(1):69-80 [FREE Full text] [doi: 10.1002/aur.1885] [Medline: 29120534]
26. Bajaj JK, Salwan P, Salwan S. Various possible toxicants involved in thyroid dysfunction: a review. J Clin Diagn Res 2016 Jan;10(1):FE01-FE03 [FREE Full text] [doi: 10.7860/JCDR/2016/15195.7092] [Medline: 26894086]
27. Hamadani JD, Baker-Henningham H, Tofail F, Mehrin SF, Grantham-McGregor S. Learning opportunities and responsive caregiving. Lancet Glob Health 2019 Jun;7(6):e709 [FREE Full text] [doi: 10.1016/S2214-109X(19)30191-3] [Medline: 31097274]
28. LaKind JS, Lehmann GM, Davis MH, Hines EP, Marchitti SA, Alcala C, et al. Infant dietary exposures to environmental chemicals and infant/child health: a critical assessment of the literature. Environ Health Perspect 2018 Sep;126(9):96002 [FREE Full text] [doi: 10.1289/EHP1954] [Medline: 30256157]
29. Vrijheid M. The exposome: a new paradigm to study the impact of environment on health. Thorax 2014 Sep;69(9):876-878. [doi: 10.1136/thoraxjnl-2013-204949] [Medline: 24906490]
30. Julvez J, López-Vicente M, Warembourg C, Maitre L, Philippat C, Gützkow KB, et al. Early life multiple exposures and child cognitive function: a multi-centric birth cohort study in six European countries. Environ Pollut 2021 Sep 01;284:117404 [FREE Full text] [doi: 10.1016/j.envpol.2021.117404] [Medline: 34077897]
31. Chandyo RK, Kvestad I, Ulak M, Ranjitkar S, Hysing M, Shrestha M, et al. The effect of vitamin B12 supplementation during pregnancy on infant growth and development in Nepal: a community-based, double-blind, randomised, placebo-controlled trial. Lancet 2023 May 06;401(10387):1508-1517. [doi: 10.1016/S0140-6736(23)00346-X] [Medline: 37031691]
32. Chandyo RK, Ulak M, Kvestad I, Shrestha M, Ranjitkar S, Basnet S, et al. The effects of vitamin B12 supplementation in pregnancy and postpartum on growth and neurodevelopment in early childhood: study protocol for a randomized placebo

XSL•FO
RenderX

controlled trial. BMJ Open 2017 Aug 29;7(8):e016434 [FREE Full text] [doi: 10.1136/bmjopen-2017-016434] [Medline: 28851784]
33. Huber S, Brox J. An automated high-throughput SPE micro-elution method for perfluoroalkyl substances in human serum. Anal Bioanal Chem 2015 May;407(13):3751-3761. [doi: 10.1007/s00216-015-8601-x] [Medline: 25795026]
34. Bach AM, Xie W, Piazzoli L, Jensen SK, Afreen S, Haque R, et al. Systemic inflammation during the first year of life is associated with brain functional connectivity and future cognitive outcomes. Dev Cogn Neurosci 2022 Feb;53:101041 [FREE Full text] [doi: 10.1016/j.dcn.2021.101041] [Medline: 34973509]
35. Prendergast AJ, Rukobo S, Chasekwa B, Mutasa K, Ntozini R, Mbuya MN, et al. Stunting is characterized by chronic inflammation in Zimbabwean infants. PLoS One 2014;9(2):e86928 [FREE Full text] [doi: 10.1371/journal.pone.0086928] [Medline: 24558364]
36. Sturgeon JP, Njunge JM, Bourke CD, Gonzales GB, Robertson RC, Bwakura-Dangarembizi M, et al. Inflammation: the driver of poor outcomes among children with severe acute malnutrition? Nutr Rev 2023 Nov 10;81(12):1636-1652 [FREE Full text] [doi: 10.1093/nutrit/nuad030] [Medline: 36977352]
37. Mallick R, Basak S, Das RK, Banerjee A, Paul S, Pathak S, et al. Roles of the gut microbiota in human neurodevelopment and adult brain disorders. Front Neurosci 2024;18:1446700 [FREE Full text] [doi: 10.3389/fnins.2024.1446700] [Medline: 39659882]
38. Yang W, Cong Y. Gut microbiota-derived metabolites in the regulation of host immune responses and immune-related inflammatory diseases. Cell Mol Immunol 2021 Apr;18(4):866-877 [FREE Full text] [doi: 10.1038/s41423-021-00661-4] [Medline: 33707689]
39. Wang J, Zhu N, Su X, Gao Y, Yang R. Gut-microbiota-derived metabolites maintain gut and systemic immune homeostasis. Cells 2023 Mar 02;12(5):793 [FREE Full text] [doi: 10.3390/cells12050793] [Medline: 36899929]
40. Cawthon RM. Telomere measurement by quantitative PCR. Nucleic Acids Res 2002 May 15;30(10):e47 [FREE Full text] [doi: 10.1093/nar/30.10.e47] [Medline: 12000852]
41. Kroustallaki P, Lirussi L, Carracedo S, You P, Esbensen QY, Götz A, et al. SMUG1 promotes telomere maintenance through telomerase RNA processing. Cell Rep 2019 Aug 13;28(7):1690-702.e10 [FREE Full text] [doi: 10.1016/j.celrep.2019.07.040] [Medline: 31412240]
42. Ranjitkar S, Kvestad I, Strand TA, Ulak M, Shrestha M, Chandyo RK, et al. Acceptability and reliability of the Bayley scales of infant and toddler development-III among children in Bhaktapur, Nepal. Front Psychol 2018 Jul 24;9:1265. [doi: 10.3389/fpsyg.2018.01265] [Medline: 30087639]
43. Brooks BL, Sherman EM, Strauss E. Test review: NEPSY-II: a developmental neuropsychological assessment, second edition. Child Neuropsychol 2010;16(1):80-101. [doi: 10.1080/09297040903146966]
44. Velligan D, Prihoda T, Dennehy E, Biggs M, Shores-Wilson K, Crismon ML, et al. Brief psychiatric rating scale expanded version: how do new items affect factor structure? Psychiatry Res 2005 Jun 30;135(3):217-228. [doi: 10.1016/j.psychres.2005.05.001] [Medline: 15993949]
45. Heys M, Gibbons F, Haworth E, Medeiros E, Tumbahangphe KM, Wickenden M, et al. The estimated prevalence of autism in school-aged children living in rural Nepal using a population-based screening tool. J Autism Dev Disord 2018 Oct 31;48(10):3483-3498 [FREE Full text] [doi: 10.1007/s10803-018-3610-1] [Medline: 29855757]
46. Vitabeginning. B 12. URL: https://www.vitabeginning.org/ [accessed 2025-05-25]
47. Thamrin C, Robinson PD, Farah CS, King GG, authors of the ERS Technical Standards for Respiratory Oscillometry. Technical standards for respiratory oscillometry and bronchodilator response cut-offs. Eur Respir J 2022 Mar;59(3):2102663 [FREE Full text] [doi: 10.1183/13993003.02663-2021] [Medline: 35058246]
48. Graham BL, Steenbruggen I, Miller MR, Barjaktarevic IZ, Cooper BG, Hall GL, et al. Standardization of spirometry 2019 update. An official American Thoracic Society and European Respiratory Society technical statement. Am J Respir Crit Care Med 2019 Oct 15;200(8):e70-e88 [FREE Full text] [doi: 10.1164/rccm.201908-1590ST] [Medline: 31613151]
49. Mobile, offline data collection for field inspections, research, and compliance. Zerion. URL: https://zerionsoftware.com/ [accessed 2026-01-21]
50. Home page. Zenodo. URL: https://zenodo.org/ [accessed 2025-05-29]
51. Ranjitkar S, Kvestad I, Chandyo RK, Strand TA, Bakken KS, Ulak M, et al. The association between prenatal maternal selenium concentration and neurodevelopment in early childhood, results from a mother-child cohort study. J Nutr 2025 Jun;155(6):1962-1969 [FREE Full text] [doi: 10.1016/j.tjnut.2025.04.005] [Medline: 40222584]
52. Schwinger C, Sharma S, Chandyo RK, Hysing M, Kvestad I, Ulak M, et al. Cobalamin and folate status in women during early pregnancy in Bhaktapur, Nepal. J Nutr Sci 2021;10:e57 [FREE Full text] [doi: 10.1017/jns.2021.53] [Medline: 34422259]

Abbreviations

BSID-III: Bayley Scales of Infant and Toddler Development, third edition
CWR: core writing group
FEV1: forced expiratory volume in 1 second

XSL•FO
RenderX

FVC: forced vital capacity
PFAS: per- and polyfluoroalkyl substances
PoA: plan of analysis
RQ: research question

Edited by A Schwartz; submitted 05.Jun.2025; this is a non–peer-reviewed article;accepted 05.Jan.2026; published 11.Feb.2026.

Please cite as:
Strand TA, Averina M, Bakken KS, Basnet S, Böttcher Y, Huber S, Hysing M, Kvestad I, Lüders T, McCann A, Pokhrel D, Ranjitkar S, Sharma AK, Shrestha M, Ulak M, Chandyo RK
Early-Life Exposures, Neurodevelopment, and Health Outcomes: Protocol for a Birth Cohort Study
JMIR Res Protoc 2026;15:e78593
URL: https://www.researchprotocols.org/2026/1/e78593
doi:10.2196/78593
PMID:41493372

©Tor A Strand, Maria Averina, Kjersti Sletten Bakken, Sudha Basnet, Yvonne Böttcher, Sandra Huber, Mari Hysing, Ingrid Kvestad, Torben Lüders, Adrian McCann, Dhiraj Pokhrel, Suman Ranjitkar, Arun K Sharma, Merina Shrestha, Manjeswori Ulak, Ram K Chandyo. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Quantitative Understanding of Advanced Novel Imaging Techniques for Fasciitis and Biosignature Yield (Quantify): Protocol for a Cross-Sectional Diagnostic Study

Zahra Amerian[1], MPT; Timothy Fleagle[1], DPT; Utsav Tuladhar[2], PhD; Rachel Watson[3], BA, BS; Micah Wong[1], DPT; Barbara Van Gorp[1,4], MPT, DPT; Brian J Smith[3], PhD; Michael Richards[2], PhD; Mederic Hall[5], MD; Joe Donnelly[6], DHS; Jessica Danielson[7], DScD; Renata Vidal Leao[8], MD; David Stanley[8], BS; Vincent Magnotta[8], PhD; James H Holmes[8], PhD; Kathleen Sluka[1,9], PT, PhD; Ruth L Chimenti[1], DPT, PhD

[1]Department of Physical Therapy and Rehabilitation Science, University of Iowa, Iowa City, IA, United States
[2]Department of Biomedical Engineering, Rochester Institute of Technology, Rochester, NY, United States
[3]Department of Biostatistics, College of Public Health, University of Iowa, Iowa City, IA, United States
[4]Iowa Orthopedics, Sports Medicine and Rehabilitations, University of Iowa, Iowa City, IA, United States
[5]Department of Orthopedics and Rehabilitation, University of Iowa, Iowa City, IA, United States
[6]Physical Therapy Program, College of Rehabilitative Sciences, University of St Augustine for Health Sciences, Coral Gables, FL, United States
[7]Institute for Clinical and Translational Science, University of Iowa, Iowa City, IA, United States
[8]Department of Radiology, Carver College of Medicine, University of Iowa, Iowa City, IA, United States
[9]Department of Neuroscience and Pharmacology, University of Iowa, Iowa City, IA, United States

Corresponding Author:
Ruth L Chimenti, DPT, PhD
Department of Physical Therapy and Rehabilitation Science
University of Iowa
500 Newton Rd, 1-252 Medical Education Building
Iowa City, IA, 52242-1190
United States
Phone: 1 319 335 9802
Email: Ruth-chimenti@uiowa.edu

Abstract

Background: Myofascial pain remains an underdiagnosed contributor to musculoskeletal pain conditions, including plantar heel pain, which is the most common source of foot pain. The current standard for diagnosing myofascial pain is a clinical examination using manual palpation. However, this approach lacks quantitative thresholds for precise assessment of myofascial pain, highlighting the need for validated biomarkers.

Objective: This protocol describes the development of a diagnostic imaging biosignature of myofascial pain using both ultrasound and magnetic resonance imaging to differentiate individuals with plantar heel pain from those with other kinds of foot pain and matched pain-free controls. The study will also explore whether diagnostic accuracy is enhanced by creating a composite biosignature that includes psychological factors.

Methods: In this cross-sectional study, 100 participants will be recruited: 50 with plantar heel pain, 25 with insertional Achilles tendinopathy, and 25 pain-free controls. Participants will undergo a clinical examination of 5 calf and foot muscles to identify sites of abnormal myofascial tissue. The primary imaging outcomes will capture biochemical properties (T1ρ of muscle and fascia), biomechanical properties (shear wave speed of the muscle, shear strain of the plantar fascia during passive movement), and structural profile (fat fraction of the muscle, thickness of the plantar fascia). Patient-reported outcomes will include the National Institutes of Health's Helping to End Addiction Long-term (HEAL) Initiative Common Data Elements and additional psychological measures.

Results: This study is supported by grant R61AT012275 from the National Center for Complementary and Integrative Health and the National Institute of Neurological Disorders and Stroke, awarded in September 2024. Participant enrollment began in May 2025. As of November 2025, a total of 55 participants have been enrolled. Enrollment is expected to conclude no later than July 2026. The anticipated study completion date is August 2026. Data will be shared within 1 year of completing the study or upon publication, whichever occurs first.

XSL•FO
RenderX

Conclusions: This protocol provides novel mechanistic insight into myofascial pain through advanced imaging techniques, offering a biopsychosocial framework for improving the diagnosis and treatment of plantar heel pain and related conditions. We anticipate that combining imaging and psychosocial measures will improve the diagnostic accuracy of the biosignature and provide a more comprehensive understanding of myofascial pain.

Trial Registration: ClinicalTrials.gov NCT06803056; https://clinicaltrials.gov/study/NCT06803056; OSF Registries osf.io/nxqfj; https://osf.io/nxqfj

International Registered Report Identifier (IRRID): DERR1-10.2196/87613

(JMIR Res Protoc 2026;15:e87613) doi:10.2196/87613

KEYWORDS

fasciitis; plantar; Achilles tendon; biomarker; myofascial; ultrasound imaging; magnetic resonance imaging; pain

Introduction

Plantar heel pain, the most common source of foot pain, affects approximately 1 out of every 10 adults and accounts for approximately 2 million medical visits annually in the United States [1,2]. Individuals with plantar heel pain frequently experience reduced quality of life and diminished satisfaction with activities of daily living [3,4]. Plantar heel pain is characterized by pain along the sole, most noticeable during initial steps following a period of inactivity and during prolonged weight-bearing activities [5]. The term “plantar fasciitis” is often used to describe this condition and suggests inflammation as the main cause of heel pain. However, the broader term “plantar heel pain” is preferred, reflecting the multifactorial nature of the underlying causes—including myofascial changes, adjacent tissue pathology, and psychosocial factors [5-9].

One potentially underrecognized component of plantar heel pain is myofascial pain, which is characterized by the presence of trigger points in the calf and foot muscles that can reproduce the individual’s pain [10,11]. Trigger points are defined as hypersensitive areas within a tight band of skeletal muscle that elicit local or referred pain when compressed, contracted, or stretched. The current standard for diagnosing myofascial pain relies on a clinical examination using manual palpation [12,13], which lacks objectivity and quantitative precision. Identifying a diagnostic imaging biomarker [14] for trigger points could significantly improve the accuracy of diagnosing myofascial pain and inform targeted treatment strategies.

Imaging is a powerful, yet underused, diagnostic tool that can be used to identify muscles with trigger points [15]. Ultrasound and magnetic resonance imaging (MRI) are both commonly used clinically to evaluate soft tissues [16-18]. Ultrasound is a noninvasive imaging technique with minimal contraindications that provides real-time assessments of soft tissue, including muscles and fascia [19]. In particular, ultrasound shear wave elastography (SWE) demonstrates high diagnostic accuracy in capturing abnormal biomechanical properties of myofascial tissue [15]. MRI is often regarded as the gold standard for evaluating soft tissues due to its ability to provide superior contrast, high resolution, and multiplanar images [18]. Previous studies have shown promise using T1 signal intensity and T2 mapping to detect structural alterations in trapezius muscles in women with migraines [20-22]. To build on previous foundational work, further research is needed to examine the diagnostic capabilities of both ultrasound and MRI within the same sample and to expand imaging approaches for identifying trigger points in the foot and ankle.

The primary aim of this proposed study is to develop a diagnostic imaging biosignature of myofascial pain to differentiate individuals with plantar heel pain from those with other kinds of foot pain (specifically insertional Achilles tendinopathy) and from matched pain-free controls. We hypothesize that a diagnostic biosignature, which uses both advanced ultrasound and MRI to quantify biochemical (muscle and fascia T1ρ), biomechanical (muscle SWE and fascia strain), and structural (muscle fat fraction and fascia thickness) tissue properties, will more accurately identify myofascial pain in the foot and ankle than a single imaging biomarker alone. We will also explore the potential of composite biosignatures that combine multiple imaging biomarkers with psychological factors, such as kinesiophobia, to provide a more comprehensive assessment. The development of a diagnostic biosignature may provide novel mechanistic insights into myofascial pain and offer a framework to facilitate the diagnosis of myofascial pain contributing to plantar heel pain.

Methods

Participants

This cross-sectional study will enroll a total of 100 participants, including 50 individuals diagnosed with plantar heel pain, along with 2 control groups. The first control group will consist of 25 individuals with insertional Achilles tendinopathy, for whom targeted myofascial treatments are not considered first-line interventions [23]. The second control group will include 25 pain-free individuals with a similar age, BMI, and sex distribution as the plantar heel pain and insertional Achilles tendinopathy groups. To ensure balanced representation, enrollment will be structured so that neither sex comprises more than 60% of each group. Eligibility and group assignment will be based on the inclusion and exclusion criteria outlined in Textbox 1. Screening for study eligibility will occur in three phases: (1) an online screening form, (2) a video visit with a physical therapist, and (3) an in-person clinical examination by a physical therapist. Individuals who provide informed consent and meet all criteria will be enrolled in this study.

XSL•FO
RenderX

Textbox 1. Eligibility criteria.

Inclusion criteria

- For the plantar heel pain and insertional Achilles tendinopathy group:
- Clinical diagnosis of plantar heel pain [5] or insertional Achilles tendinopathy [23].
- Duration of pain ≥3 months.
- Severity of pain ≥3/10.

Exclusion criteria

- For all groups:
- Age <18 years.
- History of invasive procedures to the foot or ankle on the side of interest.
- History of lower extremity injections, dry needling, extracorporeal pulse activation treatment, or extracorporeal shockwave therapy within the past 3 months on the side of interest.
- Contraindications for MRI (eg, non–magnetic resonance–compatible implanted devices, claustrophobia, inability to remain still comfortably for 1 hour in a supine position [24])
- Clinically unstable medical or psychiatric issues.
- Comorbidities associated with changes in musculoskeletal imaging, including rheumatologic and inflammatory conditions that directly impact the musculoskeletal system (eg, rheumatoid arthritis, spondyloarthropathy, and gout), osteoarthritis of the foot or ankle, diabetes, neuromuscular diseases (eg, Charcot-Marie-Tooth [CMT] disease), familial hypercholesterolemia, history of foot or ankle fracture, neoplasms involving the foot, plantar fibromatosis, or recent (within the past 1 year) infection of the foot or ankle (eg, infectious fasciitis and calcaneal osteomyelitis).
- For the plantar heel pain and insertional Achilles tendinopathy group:
- Other sources of heel pain (eg, tarsal tunnel syndrome, peripheral neuropathy, lumbar radiculopathy, calcaneal stress fracture, fibromyalgia, Morton’s neuroma, or Baxter neuropathy).
- Additional exclusion criteria for insertional Achilles tendinopathy group: co-occurring plantar heel pain and Achilles tendinopathy.
- For the pain-free control group:
- Persistent or recurrent leg pain in the past 6 months.

Data Collection Procedures

Overview

This cross-sectional study will include 3 main components of data collection: a clinical evaluation, ultrasound imaging, and MRI (Figure 1).

XSL•FO
RenderX

Figure 1. Study flow diagram. This diagram outlines the sequential study procedures, starting with participant recruitment, consent, and eligibility assessment. It proceeds through a clinical examination of the calf and foot muscles to identify trigger points. Following the clinical assessment, participants undergo ultrasound and magnetic resonance imaging (MRI) with minimal delay to ensure spatial correspondence of identified trigger points across imaging modalities. The entire study visit lasts approximately 3 hours in a single session. MR: magnetic resonance; US: ultrasound.

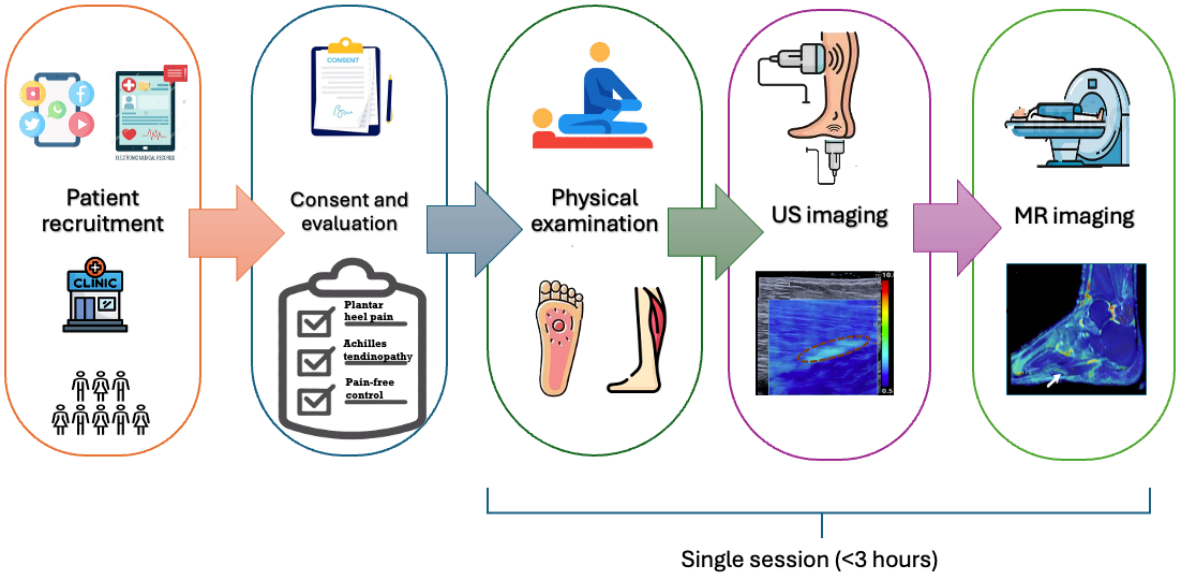

Clinical Examination

The clinical examination will be performed by a physical therapist with advanced training in dry needling for myofascial pain and will include (1) movement tasks, (2) trigger point identification, and (3) pressure pain threshold testing. In addition, participants will complete a series of patient-reported outcome measures. These outcomes align with the HEAL Common Data Elements and are supplemented by additional psychological surveys, as well as foot- and ankle-specific pain and function outcomes [25] (Table 1).

XSL•FO
RenderX

Table 1. Patient-reported outcome measures.

Outcome	Description	Psychometrics
Pain		
PEG[a]	A 3-item instrument that evaluates pain intensity and its impact on enjoyment and general activity. Each item is scored from 0 to 10, and the 3 scores are averaged to yield a total PEG score, with higher values indicating greater pain and interference [26].	• Internal consistency: Cronbach α=0.73 • Validity: r=0.60-0.89 (in chronic musculoskeletal pain populations)
PROMIS[b]-Pain Interference Short Form	Assesses pain interference, which quantifies the self-reported impact of pain on social, cognitive, emotional, physical, and recreational domains. This instrument yields raw scores from 6 to 30, based on 6 items scored from 1 (“not at all”) to 5 (“very much”). Scores are converted to T-scores (mean 50, SD 10; range ~30 to over 80) standardized to the US general population. T-scores below 55 are within normal limits, 55-60 suggest mild, 60-70 moderate interference, and scores >70 reflect severe interference [27-29].	• Internal consistency: Cronbach α=0.96-0.99 • Test-retest reliability: ICC[c]=0.85 • Convergent validity with the Brief Pain Inventory: r=0.80-0.85
Physical function		
PROMIS[b]-Physical Function Short Form 6b	Assesses the broader influence of pain on life domains, with raw scores ranging from 6 to 30; each item is scored from 1 (“unable to do”) to 5 (“without any difficulty”). Converted T-scores (mean 50, SD 10) range approximately from 20 to >80. T-scores >45 are within normal limits, scores from 40 to 45 suggest mild limitation, scores from 30 to 40 indicate moderate limitation, and scores <30 reflect severe limitation [29-31].	• Internal consistency: Cronbach α=0.96 for person and 0.99 for item • Construct validity: r=0.657 (in patients undergoing foot and ankle surgery)
FAAM[d]	Provides targeted assessment of foot and ankle function across two subscales: (1) activities of daily living (ADL), which includes 21 items scored from 0 (“unable to do”) to 4 (“no difficulty”), with a total score range of 0-84; and (2) sports, which includes 8 items with a total score range of 0-32. Scores are converted to percentages, with higher percentages reflecting better function [32].	• Test-retest reliability: ICC=0.89 for the ADL; ICC=0.87 for the Sports. • Correlations with the SF-36 physical function subscale: r=0.84-0.78 and physical component summary score: r=0.78-0.80
VISA-A Sedentary[e]	Evaluates the severity of Achilles tendinopathy in sedentary individuals. It consists of 8 variably scored items addressing symptoms and activity impact, with a total score ranging from 0 to 100. A score of 100 indicates full function without symptoms, whereas scores <70 suggest notable impairment [33].	• Test–retest reliability: ICC=0.991 for symptom and ICC=0.999 for activity • Internal consistency: Cronbach α=0.72 • Construct validity: r=0.42 for symptoms; r=0.40 for activity
Psychological factors		
PCS-6[f]	A short form of the Pain Catastrophizing Scale that measures catastrophic thinking about pain. It consists of 6 items across 3 subscales—2 items each for “rumination,” “helplessness,” and “magnification”—each scored from 0 to 4, yielding a total score of 0-24, with higher scores indicating greater catastrophic thinking [34,35].	• Internal consistency: Cronbach α=0.88 • Correlation with the full 13-item PCS: r=0.94 in individuals with chronic pain conditions
TSK-11[g]	Measures fear of movement using 11 items from 1 (“strongly disagree”) to 4 (“strongly agree”), yielding a total score of 11-44. Scores <22 indicate minimal fear, 23-28 low fear, 29-35 moderate fear, and ≥36 high fear [36,37].	• Convergent validity: association with pain catastrophizing, severity of expected pain with movement, and willingness to complete tendon loading exercises in individuals with Achilles tendinopathy • Test-retest reliability: ICC=0.81 (in individuals with back pain)
PHQ-2[h]	Screens for depression over the past 2 weeks using 2 items scored from 0 (“not at all”) to 3 (“nearly every day”), yielding a total range of 0-6. Scores of 0-2 suggest minimal or no depression, while scores ≥3 indicate possible depression requiring further evaluation [36].	• Sensitivity: 83% • Specificity: 92%

XSL•FO
RenderX

Outcome	Description	Psychometrics
PROMIS[b] Sleep Disturbance	Evaluates sleep quality and disturbances across 8 items, each scored from 1 (“not at all”) to 5 (“very much”), yielding raw scores of 8-40. Scores are converted to T-scores (mean 50, SD 10), ranging approximately from 20 to >80. T-scores of 55-60 indicate mild sleep disturbance, 60-70 suggest moderate disturbance, and >70 reflect severe sleep disturbance [37,38].	• Internal consistency: Cronbach α=0.96 • Convergent validity with the Pittsburgh Sleep Quality Index: r=0.85 • Test-retest reliability: ICC=0.62-0.71 (in fibromyalgia populations)
PROMIS[b] Self-Efficacy for Managing Symptoms 8a	Assess the subject’s level of confidence in managing their symptoms in different settings (home, public place, and an unfamiliar place) and in keeping their symptoms from interfering with work, sleep, relationships, or recreational activities. Each item scored from 1 (“not at all confident”) to 5 (“very confident”), yielding raw scores from 8 to 40. Converted T-scores below 30 suggest very low, 30-40 low, 40-60 average, 60-70 high, and 70-80 reflect very high self-efficacy [39].	• Internal consistency: Cronbach α=0.92 • Correlations with Chronic Pain Self-Efficacy Scale: r=0.74 (in patients with chronic conditions)
GAD-2[i]	Screens for generalized anxiety disorder by asking about the frequency of feeling nervous and unable to control worrying over the past 2 weeks. It uses 2 items, each scored from 0 (“not at all”) to 3 (“nearly every day”), totaling 0 to 6. Scores of 0 to 2 indicate minimal or no anxiety, while scores of 3 or higher suggest possible anxiety requiring further assessment [40].	• Sensitivity: 76% • Specificity: 81% • Internal consistency: Cronbach α=0.89
Comorbidities		
TAPS-1[j]	Screens for substance use in the past 12 months. Scores range from 0 to 4 (0=“never” and 4=“daily or almost daily”) for each substance category, with higher scores indicating more frequent use [41].	• Sensitivity: 81%-96% • Specificity: 78%-91%
CCI[k]	Predicts 10-year mortality for patients with multiple comorbidities based on 19 weighted conditions, scored from 1 to 6 points, with a total range of 0 to >37. Higher scores correspond to a greater comorbidity burden [42-44].	• Test-retest reliability: ICC=0.94 (for self-report data compared to chart reviews) • Inter-rater reliability: ICC=0.80
WPI[l]	Assesses pain distribution across 19 body areas, assigning a score of 0 if there is no pain and 1 if pain is present in each area. A total score of 7 or more typically indicates widespread pain. The SSS[m] scale evaluates symptom severity in 2 components: the first assesses fatigue, unrefreshing sleep, and cognitive difficulties, while the second derives a general symptom severity score based on the number of symptoms reported from a list of 41 symptoms. Patients receive a score of 0 for no symptoms, 1 for 1-10 symptoms, 2 for 11-24 symptoms, and 3 for 25 or more symptoms. The total SSS score ranges from 0 to 12, calculated by summing the scores from the first component (range 0-9) and the second (range 0-3) [45].	• Internal consistency: Cronbach α=0.34 for the WPI, 0.83 for the SS, and 0.82 for the combined use of both tools. • Sensitivity: 100% and Specificity: 81% for WPI • Sensitivity: 89.2% and Specificity: 83.8% for SSS

[a]PEG: Pain, Enjoyment, General Activity.

[b]PROMIS: Patient-Reported Outcomes Measurement Information System.

[c]ICC: intraclass correlation coefficient.

[d]FAAM: Foot and Ankle Ability Measure.

[e]VISA-A Sedentary: Victorian Institute of Sport Assessment–Achilles (sedentary version).

[f]PCS-6: Pain Catastrophizing Scale–6.

[g]TSK-11: Tampa Scale for Kinesiophobia–11.

[h]PHQ-2: Patient Health Questionnaire–2.

[i]GAD-2: Generalized Anxiety Disorder–2.

[j]TAPS-1: Tobacco, Alcohol, Prescription medication, and other Substance use.

[k]CCI: Charlson Comorbidity Index.

[l]WPI: Widespread Pain Index.

[m]SSS: Symptoms Severity Scale.

XSL•FO
RenderX

Movement Tasks

Pain intensity will be reported using the numeric rating scale (NRS), where 0 indicates no pain, and 10 indicates the worst pain imaginable [46,47]. The NRS has been shown to have excellent test-retest reliability (intraclass correlation coefficient=0.95) and concurrent validity with good-to-excellent correlation with the visual analog scale (r=0.941) [48]. Participants will rate their level of pain at rest. In addition, they will report their heel pain (plantar fascia or Achilles tendon, depending on group assignment) during the following three movement tasks: (1) the "Windlass test" will be performed in a standing position with passive extension of the participant's great toe to the end range while bearing weight on the affected foot. A positive test will be indicated by reproduction or exacerbation of plantar heel pain [5,49]. The weight-bearing Windlass test demonstrates 100% specificity and 31.8% sensitivity [49]; (2) for the "single-leg heel raise test," participants will stand on the affected leg with their fingertips on a wall in front of them to provide balance. Keeping the knee straight, they will then lift and lower their heel as high as possible. They will perform as many repetitions as they can while keeping pace with a metronome beat played from a phone application (60 Hz). The task will continue until muscle fatigue or symptom exacerbation prevents further movement. The test will be discontinued if the participant is unable to maintain proper technique, such as bending the knee, flexing the trunk, or failing to keep pace with the metronome. If a participant is unable to perform single-limb heel raises, the test will be modified to a double-limb heel raise, instructing them to shift as much weight as possible onto the involved side [5,50,51]; and (3) the "single-leg drop landing test" will be conducted by asking participants to step onto a platform while leading with their affected leg and then step back down, landing only on the affected foot [52]. If this movement does not provoke pain or symptoms, participants will be instructed to add a vertical jump immediately after landing on the involved foot [53].

Trigger Point Identification

Participants will lie in the prone position with knees flexed at 15° and ankles in their naturally relaxed plantarflexed posture. A physical therapist will examine 5 muscles: 3 in the calf (medial gastrocnemius, lateral gastrocnemius, and medial soleus) and 2 in the foot (abductor hallucis and flexor digitorum brevis). To facilitate consistency in the examination and imaging, the assessment will be limited to the following muscle zones, which Simons [54] has described as common trigger point locations.

Calf zones will be centered at 50% of the leg length measured from the fibular head, with the gastrocnemius zone extended proximally by 40% of the leg length, while the soleus zone will be extended distally by 30% of the leg length. For foot muscles, the abductor hallucis zone will be centered along a vertical line from the anterior medial malleolus, with medial and lateral boundaries extending 5% of the leg length [55]. The flexor digitorum brevis zone will begin 2 cm distal to the medial calcaneal tubercle and span 10% of the leg length distally. Leg length will be measured from the fibular head to the lateral malleolus. Using standardized zones based on leg length ensures that similar regions of the muscle belly are captured and helps ensure that the distance between the most proximal and distal trigger points is no greater than 20 cm, which corresponds to the MRI field of view.

The physical therapist will palpate within the 5 zones to determine if a trigger point is present and mark the tautest band using a waterproof marker in a cross shape in each zone. This approach allows for the same locations to be assessed with ultrasound imaging and MRI. Trigger point identification will follow the gold standard diagnostic criteria established by Simons [54], who defines a trigger point as a taut band of skeletal muscle that is tender to palpation. The physical therapist will apply constant pressure with palpation for 5 seconds on the most tender taut band in each of the 5 zones, with a 20-second rest between each spot. Participants will be asked whether the pressure reproduces the same heel pain [54] and whether the pain is felt locally at the site of palpation or referred to another area [10] at each site. For comparison of imaging outcomes between trigger points and healthy tissue, at least 1 of the 5 marks will indicate healthy tissue using the same cross shape. Imaging researchers will remain blinded to the trigger point versus healthy tissue identification to ensure unbiased image acquisition.

Pressure Pain Threshold Test

Following the completion of ultrasound and MRI, a pressure pain threshold test will be conducted using a pressure algometer (Somedic Algometer Type II) on each of the 5 previously marked locations on the leg and foot [56-60]. The algometer will be positioned perpendicularly to these sites, and pressure will be consistently applied downward at a constant rate of 50 kPa/s [61,62] using a 1-cm^2 round rubber tip probe. Participants will be instructed to press a handheld button at the moment they first perceive the pressure sensation as painful, defined as anything greater than 0/10 on the NRS [63]. Three measurements will be taken per site, with at least 30 seconds between repetitions to reduce temporal summation effects. Participants with even and odd ID numbers will be tested in different sequences of muscle groups to minimize order effects.

Ultrasound Imaging

Ultrasound imaging will be performed by a physical therapist experienced in ultrasound imaging of the foot and ankle. To provide consistent positioning across the ultrasound and MRI scans, the participant will be in the prone position with the knee flexed to 15° and a strap around the ball of the foot applying slight pressure to maintain the ankle in a slightly plantarflexed position (approximately 20°). This ankle position replicates the 1.5-inch lift under the heel provided during the MRI. Ultrasound imaging will be acquired using the ACUSON Sequoia Ultrasound System (Siemens Healthineers). For SWE, a 10L4 linear array transducer (2.9-9.9 MHz) will be positioned parallel with the muscle fibers, centered over the 5 marks identified during the clinical examination. Because the plantar fascia lies immediately superficial to the flexor digitorum brevis, its shear wave speed will also be captured. In addition, as an exploratory measure, the transducer will also be positioned parallel with the Achilles tendon to capture the region immediately proximal to the calcaneus. A standoff pad will be used for Achilles tendon and abductor hallucis muscle imaging to allow for a similar

XSL•FO
RenderX

imaging depth as the calf muscles, which have a greater amount of superficial adipose tissue. Three images per location will be collected. To quantify shear wave velocity, a custom MATLAB code will be used to allow for the selection of the tissue region of interest, with analysts blinded to the overlay of the shear wave velocity map. The average of the 3 replicates will be used to estimate a representative shear wave velocity for that tissue [64].

The thickness of the plantar fascia will be measured using the maximum of 2 B-mode images, a long-axis and a short-axis view, at the insertion to the calcaneus. For the plantar fascia strain, an 18L6 linear array transducer (5.5-18 MHz) will be positioned parallel with the plantar fascia for a long-axis view. A cine loop will be collected while the participant's great toe is passively extended from neutral to the end range of motion over 5 seconds. Simultaneously, a video of the foot will be captured on a tablet camera placed approximately 1 foot medial to the participant's ankle, with the foot centered in the video frame and the screen positioned at a 90° angle relative to the floor. This passive range of motion task and imaging protocol will be repeated 3 times. A nonrigid image registration–based strain estimator will be used to measure the accumulated displacement field in the fascia throughout the first toe passive extension task [65]. The displacement estimates will then be used to quantify the average shear and principal strains within the plantar fascia during the passive movement. The magnitude of toe extension will be quantified using Kinovea software (open-source, version 0.9.5; Joan Charmant).

MRI

To ensure spatial alignment between clinical examination, ultrasound imaging, and MRI, magnetic resonance–visible fiducial markers will be placed over the marked trigger points identified during clinical examination. MRI will be conducted using a 3T whole-body MRI (Signal Premier, GE Healthcare). Scanning will begin using an 8-channel foot-ankle receive RF coil (GE Healthcare, Waukesha) for imaging the foot, followed by a 32-channel flexible blanket array receive RF coil (GE Healthcare) for imaging the calf. Imaging will be performed on the affected side in participants with plantar heel pain and insertional Achilles tendinopathy, and on a single side in pain-free controls, with all participants scanned in the supine position with slight knee flexion and ankle plantar flexion to accommodate a 1.5-inch heel lift between the heel and the RF coil (approximately 20° of plantar flexion).

The imaging will include the following sequences: T2-weighted fast spin-echo sequences for T2 mapping and evaluation of inflammation and cysts [66]; zero echo time proton density sequences to detect bony abnormalities [67,68]; quantitative T1ρ imaging, which provides sensitivity to multiple features such as chemical exchange between extracellular water, macromolecules and pH [69,70]; chemical shift–encoded Iterative Decomposition of water and fat with Echo Asymmetry and Least-squares estimation for T2* and fat fraction quantification to assess tissue composition [71,72]; intravoxel incoherent motion (IVIM) diffusion imaging to evaluate tissue perfusion; and T1-weighted imaging to measure plantar fascia and fat pad thickness.

MRI data will be analyzed using both quantitative and qualitative approaches. Regions of interest (ROI) will be measured based on the fiducial markers corresponding to the locations identified during the clinical examination. Water and fat images will be generated using the Iterative Decomposition of water and fat with Echo Asymmetry and Least-squares estimation (IDEAL) for the separation of fat and water species [71, 72], Specifically, these will be produced using the vendor supplied chemical shift–encoded 3D gradient-echo sequence (IDEAL-IQ, GE Healthcare, WI, USA).

The IDEAL-IQ gradient-echo images will be used for marker identification and tissue segmentation, and will serve as the reference images for spatially registering the other MRI images to a common image-space. The ROIs will be delineated using the IDEAL-IQ images, guided by the locations of the fiducial markers that are superficial to the trigger points. Detailed MRI scan parameters for the foot and lower leg, respectively, are provided in Tables S1 and S2. The corresponding ROIs will then be used to measure T1ρ, T2*, perfusion fraction, and pseudo-perfusion from IVIM, as well as fat–water fractions. Measurements from each ROI will include the mean, median, and SD. Quantitative maps of fat–water fraction and T2* will be automatically generated by the scanner using the IDEAL-IQ sequence, while voxel-wise parametric maps for T1ρ, T2, and IVIM will be produced using in-house software.

Qualitative image assessments will be performed by a radiologist to exclude differential diagnoses for plantar heel pain and to confirm accurate identification of leg and foot muscles.

Statistical Analysis

Data Analysis

In the statistical analysis, quantitative measures will be summarized with means, SDs, and quantiles, and categorical measures, including the presence of trigger points (trigger points vs healthy tissue), will be summarized with counts and percentages. Data will be described for 3 calf muscles (medial gastrocnemius, lateral gastrocnemius, and medial soleus) and 2 foot muscles (abductor hallucis and flexor digitorum brevis).

The primary study aim is to examine the diagnostic performance of 20 candidate biomarkers (biomechanical, biochemical, and structural) in distinguishing between tissue with and without myofascial pain. Performance will be quantified with the area under the receiver operating characteristic curve (ROC AUC), pooled over the 5 muscle groups assessed on each participant for imaged metrics and trigger points. Five-fold cross-validation will be used to estimate predictive performance, and bootstrap resampling applied to obtain 95% CIs. Participants will be treated as the sampling units in cluster cross-validation and bootstrap sampling algorithms to account for repeated muscle measurements in each participant. To identify the most promising biomarkers and account for multiple comparisons, false discovery rates will be computed with the method of Benjamini and Hochberg to rank the biomarkers in order of diagnostic performance [73]. Up to 5 top-performing biomarkers will be selected to constitute a single biosignature to achieve a target AUC >0.7, sensitivity >60%, and specificity >60%.

XSL•FO
RenderX

In an exploratory aim, composite biosignatures will be developed from imaging biomarkers, patient demographics, and psychological factors. Traditional statistical and machine learning techniques will be used. Logistic regression models, fit with the method of general estimating equations, will be used specifically to test for interactions between imaging biomarkers and sex. Composite models will be built from the full set of predictors using stepwise variable selection as well as the machine learning methods of Lasso, gradient boosting, and random forests. The latter 2 methods have the distinction of performing automated encoding of nonlinear and interaction effects. ROC AUC will be estimated as the measure of diagnostic performance for each model, as described for the primary aim. In addition, variable importance will be reported to quantify the relative contributions of predictor variables, and partial dependence plots will be constructed to show the functional forms of their effects.

Sample Size

The target sample size of 50 individuals with plantar heel pain, 25 individuals with insertional Achilles tendinopathy, and 25 pain-free controls was chosen to provide a 95% CI half-width of 0.10 for an AUC of 0.70. This calculation assumes 5 muscle measurements for each participant, an intraclass correlation coefficient of 0.5 among them, and an overall active trigger point prevalence of 36% in individuals with plantar heel pain [10,74]. Further assuming that the 20 candidate biomarkers have AUCs uniformly distributed from 0.5 to 0.8, the study will have at least 85% power to obtain a biosignature meeting the desired AUC threshold.

Alterations to Study After Initiation

To minimize missing data and maximize data quality, we have made some changes to our study protocol since study initiation. For trigger point identification, we added the step of checking to ensure that the maximum distance between the most proximal and the most distal trigger point is no more than 20 cm, to ensure that they are captured within the magnetic resonance field of view. Also, we adjusted the relative proportions of the regions of ROIs for the gastrocnemius and soleus muscles to minimize inclusion of the gastrocnemius within the soleus ROI during ultrasound imaging. For shear wave ultrasound imaging, we added the use of a gel pad over the abductor hallucis muscle to have a more equivalent imaging depth as the muscles that have a thicker superficial adipose tissue layer. For MRI, we similarly updated the slice number to routinely cover 20 cm in the calf. If global fat-water swaps are encountered, we reduce the readout matrix to 160 to reduce the first echo time, as well as reduce the phase encode number to 160 to maintain consistent voxel dimensions. We also added a B1 mapping sequence to monitor and reduce sensitivity to B1 variations [75]. ROI segmentation was updated to better accommodate oblique scan orientations and data resampling based on registration to the 3D Iterative Decomposition of water and fat with Echo Asymmetry and Least-squares estimation acquisitions. The length of the ROIs was consistent with the diameter of the fiducial markers (1.5 cm).

Ethical Considerations

This study was preregistered on ClinicalTrials.gov (NCT06803056) and the Open Science Framework (https://osf.io/nxqfj). Ethical approval (IRB 202409110) was obtained from the Biomedical Research Committee (IRB-01) at the University of Iowa. Informed consent will be obtained from all participants before enrollment, and the potential risks and benefits of the study will be clearly explained. Participants will receive compensation for their involvement. All study data will be securely stored in REDCap, a password-protected database hosted by the University of Iowa College of Medicine. The REDCap platform is managed and supported by the Biomedical Informatics Group in the Institute for Clinical and Translational Science at the University of Iowa. REDCap is HIPAA (Health Insurance Portability and Accountability Act)- and 21 CFR 11–compliant, which provides the necessary data security for the proposed clinical trial.

Results

This study is supported by a grant R61AT012275 from the National Center for Complementary and Integrative Health and the National Institute of Neurological Disorders and Stroke, awarded in September 2024. Participant enrollment began in May 2025. As of November 2025, a total of 55 participants have been enrolled. Enrollment is expected to conclude by no later than July 2026. The anticipated study completion date is August 2026. Data will be shared within 1 year of study completion or upon publication, whichever occurs first.

Discussion

Principal Findings

This study is designed to develop a diagnostic imaging biosignature for myofascial pain that differentiates individuals with plantar heel pain from those with other foot pain, such as insertional Achilles tendinopathy, and from pain-free controls. We anticipate that combining imaging and psychosocial measures will improve diagnostic accuracy and provide a more comprehensive understanding of myofascial pain.

Comparison With Previous Work

The novelty of this study is supported by a multimodal approach, integrating ultrasound and MRI techniques to provide complementary insights into myofascial structures. Despite MRI being considered the gold standard for evaluating soft tissues due to its ability to provide high-resolution detail and multiplanar assessments, its application in myofascial pain research has been limited [15,18]. Previous studies analyzed either T1 signal intensity or T2 mapping, suggesting the promise of these techniques to detect the altered structure of the trapezius muscle in the migraine population [20-22].

Multimodal Imaging and Biosignature Development

Building on this foundational work, our study will use multiple advanced quantitative MRI techniques, including T1ρ, to capture a broader range of biochemical properties associated with myofascial pain. In addition, the inclusion of matched pain-free controls and individuals with insertional Achilles tendinopathy,

XSL•FO
RenderX

a common differential diagnosis for plantar heel pain, enhances the clinical relevance and enables a nuanced understanding of biomarker variations across conditions.

Psychosocial Factors in Myofascial Pain

In addition to examining biological factors contributing to myofascial pain, the study will also explore how psychosocial factors may enhance the accuracy of an imaging biosignature. Psychosocial factors have been shown to influence pain perception and treatment response, further complicating diagnosis. By expanding the definition of a "biosignature" to include psychological variables such as depression, anxiety, and fear of movement, this study aligns with emerging trends in biomarker research, which recognize the interplay of biological and psychological factors in pain chronicity [76].

Limitations

A limitation of this study is the absence of a universally accepted definition of a trigger point across body regions and between professions. In this study, the definition of a trigger point is based on criteria commonly cited in the field of physical therapy [12,13]. Yet other clinicians and researchers examining myofascial pain may select other definitions. As evidence on biosignatures of myofascial pain continues to emerge, these findings will help refine and standardize the definition of trigger points across diagnoses and professions.

Implications for Future Research and Clinical Practice

The anticipated findings will lay the groundwork for future large-scale investigations, including clinical trials focused on nonpharmacological and nonsurgical interventions for plantar heel pain, such as dry needling. By establishing a robust biosignature for myofascial pain, this research has the potential to refine diagnostic criteria, improve treatment selection, and contribute to a more personalized approach to pain management.

Conclusions

This protocol provides novel mechanistic insights into myofascial pain through advanced imaging, offering a biopsychosocial framework for improved diagnosis and treatment of plantar heel pain and related conditions. We anticipate that combining imaging and psychosocial measures will improve the diagnostic accuracy of the biosignature and provide a more comprehensive understanding of myofascial pain.

Acknowledgments

The contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Center for Complementary and Integrative Health, National Institute of Neurological Disorders and Stroke (NINDS), or the National Institutes of Health (NIH). We would like to acknowledge Paul DiCamillo and Dan Thedens for their valuable contributions to the development of the magnetic resonance imaging methods. The authors declare the use of generative artificial intelligence (AI; Microsoft Copilot 2.2) during the research and writing process. In accordance with the GAIDeT taxonomy (2025), generative AI tools were used exclusively for reformatting tasks, including grammar correction and sentence structure refinement, under full human supervision. Generative AI tools were not involved in the conceptualization, data analysis, interpretation, or drafting of original content. These tools are not listed as authors and do not bear responsibility for the outcomes.

Funding

This publication was made possible by a grant (R61AT012275) from the National Center for Complementary and Integrative Health and NINDS. The funders had no involvement in the study design, data collection, analysis, interpretation, or manuscript writing.

Data Availability

In compliance with the findability, accessibility, interoperability, and reusability principles, data will be deposited in the University of Iowa's open access institutional repository, Iowa Research Online. This repository is maintained by the University of Iowa Libraries to preserve and share scholarly work. Individual participant data will be shared upon publication. Primary outcome data and all information necessary to validate the research findings of this study will be preserved and made available to support reproducibility and facilitate further research that builds upon these results.

Conflicts of Interest

The University of Iowa receives research support from GE Healthcare. RLC and MR receive software from Siemens Medical Solutions USA, Inc. MH reports serving on the Medical Advisory Board for Trice Medical and holds stock in both Trice Medical and Sonex Health. He also receives royalties from UpToDate, Inc. The other authors declare no competing interests related to this study.

Multimedia Appendix 1

XSL•FO
RenderX

Peer review report by ZAT1 SH (14) - National Center for Complementary and Integrative Health (NCCIH) Special Emphasis Panel HEAL Initiative: Toward Developing Quantitative Imaging and Other Relevant Biomarkers of Myofascial Tissues for Clinical Pain Management (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 161 KB - resprot_v15i1e87613_app1.pdf]

References

1. Thomas MJ, Whittle R, Menz HB, Rathod-Mistry T, Marshall M, Roddy E. Plantar heel pain in middle-aged and older adults: population prevalence, associations with health status and lifestyle factors, and frequency of healthcare use. BMC Musculoskelet Disord 2019;20(1):337 [FREE Full text] [doi: 10.1186/s12891-019-2718-6] [Medline: 31325954]
2. Riddle DL, Schappert SM. Volume of ambulatory care visits and patterns of care for patients diagnosed with plantar fasciitis: a national study of medical doctors. Foot Ankle Int 2004;25(5):303-310. [doi: 10.1177/107110070402500505] [Medline: 15134610]
3. Alshammari SA, Alshwieer MA, Dammas SS, Alrasheed A, Alasmari MA, Alahmari MM. Impact of plantar fasciitis on foot-specific and generic health-related quality of life in King Khalid University Hospital, Saudi Arabia. Cureus 2023;15(7):e41912 [FREE Full text] [doi: 10.7759/cureus.41912] [Medline: 37583733]
4. Torpil B, Önal G, Kaya O. is there a relationship between perceived activity performance-satisfaction level and quality of life in activities of daily living in adults with plantar fasciitis. Sağlık Bilimlerinde Değer 2023;13(1):15-21. [doi: 10.33631/sabd.1092158]
5. Koc TA, Bise CG, Neville C, Carreira D, Martin RL, McDonough CM. Heel pain - plantar fasciitis: revision 2023. J Orthop Sports Phys Ther 2023;53(12):CPG1-CPG39. [doi: 10.2519/jospt.2023.0303] [Medline: 38037331]
6. Landorf K. Plantar heel pain and plantar fasciitis. BMJ Clin Evid 2015;2015:1111 [FREE Full text] [Medline: 26609884]
7. San-Antolín M, Rodríguez-Sanz D, López-López D, Romero-Morales C, Carbajales-Lopez J, Becerro-de-Bengoa-Vallejo R, et al. Depression levels and symptoms in athletes with chronic gastrocnemius myofascial pain: a case-control study. Phys Ther Sport 2020;43:166-172. [doi: 10.1016/j.ptsp.2020.03.002] [Medline: 32179495]
8. Kashif M, Tahir S, Ashfaq F, Farooq S, Saeed W. Association of myofascial trigger points in neck and shoulder region with depression, anxiety and stress among university students. J Pak Med Assoc 2021;71(9):2139-2142. [doi: 10.47391/JPMA.375] [Medline: 34580502]
9. Healy GM, Finn DP, O'Gorman DA, Maharaj C, Raftery M, Ruane N, et al. Pretreatment anxiety and pain acceptance are associated with response to trigger point injection therapy for chronic myofascial pain. Pain Med 2015 Oct;16(10):1955-1966. [doi: 10.1111/pme.12781] [Medline: 26309134]
10. Ortega-Santiago R, Ríos-León M, Martín-Casas P, Fernández-de-Las-Peñas C, Plaza-Manzano G. Active muscle trigger points are associated with pain and related disability in patients with plantar heel pain: a case-control study. Pain Med 2020;21(5):1032-1038. [doi: 10.1093/pm/pnz086] [Medline: 30986304]
11. Renan-Ordine R, Alburquerque-Sendín F, de Souza DPR, Cleland JA, Fernández-de-Las-Peñas C. Effectiveness of myofascial trigger point manual therapy combined with a self-stretching protocol for the management of plantar heel pain: a randomized controlled trial. J Orthop Sports Phys Ther 2011;41(2):43-50. [doi: 10.2519/jospt.2011.3504] [Medline: 21285525]
12. Fernández-de-Las-Peñas C, Dommerholt J. International consensus on diagnostic criteria and clinical considerations of myofascial trigger points: a Delphi study. Pain Med 2018;19(1):142-150. [doi: 10.1093/pm/pnx207] [Medline: 29025044]
13. Donnelly J. Travell, Simons & Simons' Myofascial Pain and Dysfunction: the Trigger Point Manual. Waltham, Massachusetts: Lippincott Williams & Wilkins; 2018.
14. FDA-NIH Biomarker Working Group. Best (Biomarkers, Endpoints, and Other Tools) Resource. Bethesda: National Institutes of Health (US); 2016.
15. Mazza DF, Boutin RD, Chaudhari AJ. Assessment of myofascial trigger points via imaging: a systematic review. Am J Phys Med Rehabil 2021;100(10):1003-1014 [FREE Full text] [doi: 10.1097/PHM.0000000000001789] [Medline: 33990485]
16. Koh RG, Dilek B, Ye G, Selver A, Kumbhare D. Myofascial trigger point identification in b-mode ultrasound: texture analysis versus a convolutional neural network approach. Ultrasound Med Biol 2023;49(10):2273-2282. [doi: 10.1016/j.ultrasmedbio.2023.06.019] [Medline: 37495496]
17. Jacobson JA. Musculoskeletal ultrasound and MRI: which do I choose? Semin Musculoskelet Radiol 2005;9(2):135-149. [doi: 10.1055/s-2005-872339] [Medline: 16044382]
18. Evans V, Behr M, Gangwar A, Noseworthy MD, Kumbhare D. Potential role of MRI imaging for myofascial pain: a scoping review for the clinicians and theoretical considerations. J Pain Res 2021;14:1505-1514 [FREE Full text] [doi: 10.2147/JPR.S302683] [Medline: 34079365]
19. Sikdar S, Shah JP, Gebreab T, Yen R, Gilliams E, Danoff J, et al. Novel applications of ultrasound technology to visualize and characterize myofascial trigger points and surrounding soft tissue. Arch Phys Med Rehabil 2009;90(11):1829-1838 [FREE Full text] [doi: 10.1016/j.apmr.2009.04.015] [Medline: 19887205]
20. Sollmann N, Mathonia N, Weidlich D, Bonfert M, Schroeder SA, Badura KA, et al. Quantitative magnetic resonance imaging of the upper trapezius muscles - assessment of myofascial trigger points in patients with migraine. J Headache Pain 2019;20(1):8 [FREE Full text] [doi: 10.1186/s10194-019-0960-9] [Medline: 30658563]

XSL•FO
RenderX

21. Luiza da Silva Queiroz M, Bezerra Diniz PR, Nepomuceno Montenegro EJ, de Souza Costa Neto JJ, Santana de Albuquerque Filho E, Valença MM, et al. MRI in migraineurs: are there abnormalities in the area where the myofascial trigger points are palpable and in volume measurements? J Bodyw Mov Ther 2020;24(3):260-266. [doi: 10.1016/j.jbmt.2020.02.025] [Medline: 32825998]
22. Sollmann N, Schandelmaier P, Weidlich D, Börner C, Urban G, Lang M, et al. Patients with episodic migraine show increased T2 values of the trapezius muscles - an investigation by quantitative high-resolution magnetic resonance imaging. Cephalalgia 2021;41(8):934-942 [FREE Full text] [doi: 10.1177/0333102421996374] [Medline: 33615841]
23. de Vos R, van der Vlist AC, Zwerver J, Meuffels DE, Smithuis F, van Ingen R, et al. Dutch multidisciplinary guideline on Achilles tendinopathy. Br J Sports Med 2021;55(20):1125-1134 [FREE Full text] [doi: 10.1136/bjsports-2020-103867] [Medline: 34187784]
24. Carr MW, Grey ML. Magnetic resonance imaging. Am J Nurs 2002;102(12):26-33. [doi: 10.1097/00000446-200212000-00012] [Medline: 12473927]
25. Wandner LD, Domenichiello AF, Beierlein J, Pogorzala L, Aquino G, Siddons A, NIH Pain Consortium InstituteCenter Representatives. NIH's helping to end addiction long-term initiative (NIH HEAL Initiative) clinical pain management common data element program. J Pain 2022;23(3):370-378 [FREE Full text] [doi: 10.1016/j.jpain.2021.08.005] [Medline: 34508905]
26. Krebs EE, Lorenz KA, Bair MJ, Damush TM, Wu J, Sutherland JM, et al. Development and initial validation of the PEG, a three-item scale assessing pain intensity and interference. J Gen Intern Med 2009;24(6):733-738 [FREE Full text] [doi: 10.1007/s11606-009-0981-1] [Medline: 19418100]
27. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, PROMIS Cooperative Group. The patient-reported outcomes measurement information system (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010;63(11):1179-1194 [FREE Full text] [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
28. Amtmann D, Cook KF, Jensen MP, Chen WH, Choi S, Revicki D, et al. Development of a PROMIS item bank to measure pain interference. Pain 2010;150(1):173-182. [doi: 10.1016/j.pain.2010.04.025] [Medline: 20554116]
29. Driban JB, Morgan N, Price LL, Cook KF, Wang C. Patient-reported outcomes measurement information system (PROMIS) instruments among individuals with symptomatic knee osteoarthritis: a cross-sectional study of floor/ceiling effects and construct validity. BMC Musculoskelet Disord 2015;16:253 [FREE Full text] [doi: 10.1186/s12891-015-0715-y] [Medline: 26369412]
30. Rose M, Bjorner J, Becker J, Fries J, Ware J. Evaluation of a preliminary physical function item bank supported the expected advantages of the patient-reported outcomes measurement information system (PROMIS). J Clin Epidemiol 2008;61(1):17-33. [doi: 10.1016/j.jclinepi.2006.06.025] [Medline: 18083459]
31. Hung M, Baumhauer JF, Latt LD, Saltzman CL, SooHoo NF, Hunt KJ. Validation of PROMIS ® physical function computerized adaptive tests for orthopaedic foot and ankle outcome research. Clin Orthop Relat Res 2013;471(11):3466-3474 [FREE Full text] [doi: 10.1007/s11999-013-3097-1] [Medline: 23749433]
32. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the foot and ankle ability measure (FAAM). Foot Ankle Int 2005;26(11):968-983. [doi: 10.1177/107110070502601113] [Medline: 16309613]
33. Norris R, Cook JL, Gaida JE, Maddox T, Raju J, O'Neill S. The VISA-A (sedentary) should be used for sedentary patients with Achilles tendinopathy: a modified version of the VISA-A developed and evaluated in accordance with the COSMIN checklist. Br J Sports Med 2023;57(20):1311-1316 [FREE Full text] [doi: 10.1136/bjsports-2022-105547] [Medline: 36927742]
34. McWilliams L, Kowal J, Wilson K. Development and evaluation of short forms of the pain catastrophizing scale and the pain self-efficacy questionnaire. Eur J Pain 2015;19(9):1342-1349. [doi: 10.1002/ejp.665] [Medline: 25766681]
35. Baxter NA, Hoch C, Reid JJ, Scott DJ, Gross CE. Pain catastrophizing scale associated with other patient-reported outcome measures in plantar fasciitis and chronic ankle instability patients. Foot Ankle Int 2022;43(10):1340-1345. [doi: 10.1177/10711007221106472] [Medline: 35794824]
36. Kroenke K, Spitzer RL, Williams JBW. The patient health questionnaire-2: validity of a two-item depression screener. Med Care 2003;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
37. Buysse DJ, Yu L, Moul DE, Germain A, Stover A, Dodds NE, et al. Development and validation of patient-reported outcome measures for sleep disturbance and sleep-related impairments. Sleep 2010;33(6):781-792 [FREE Full text] [doi: 10.1093/sleep/33.6.781] [Medline: 20550019]
38. Chimenti RL, Rakel BA, Dailey DL, Vance CGT, Zimmerman MB, Geasland KM, et al. Test-retest reliability and responsiveness of PROMIS sleep short forms within an RCT in women with fibromyalgia. Front Pain Res (Lausanne) 2021;2:682072 [FREE Full text] [doi: 10.3389/fpain.2021.682072] [Medline: 35295526]
39. Gruber-Baldini AL, Velozo C, Romero S, Shulman LM. Validation of the PROMIS measures of self-efficacy for managing chronic conditions. Qual Life Res 2017;26(7):1915-1924 [FREE Full text] [doi: 10.1007/s11136-017-1527-3] [Medline: 28239781]

XSL•FO
RenderX

40. Plummer F, Manea L, Trepel D, McMillan D. Screening for anxiety disorders with the GAD-7 and GAD-2: a systematic review and diagnostic metaanalysis. Gen Hosp Psychiatry 2016;39:24-31. [doi: 10.1016/j.genhosppsych.2015.11.005] [Medline: 26719105]
41. Hartz SM. Validation of the TAPS-1: a four-item screening tool to identify unhealthy substance use in primary care. J Gen Intern Med 2017;32(9):990-996. [doi: 10.1007/s11606-017-4119-6]
42. Moltó A, Dougados M. Comorbidity indices. Clin Exp Rheumatol 2014;32(5 Suppl 85):131-134.
43. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson comorbidity index: a critical review of clinimetric properties. Psychother Psychosom 2022;91(1):8-35 [FREE Full text] [doi: 10.1159/000521288] [Medline: 34991091]
44. Hall SF, Groome PA, Streiner DL, Rochon PA. Interrater reliability of measurements of comorbid illness should be reported. J Clin Epidemiol 2006;59(9):926-933. [doi: 10.1016/j.jclinepi.2006.02.006] [Medline: 16895815]
45. Galvez-Sánchez CM, de la Coba P, Duschek S, Reyes Del Paso GA. Reliability, factor structure and predictive validity of the Widespread Pain Index and Symptoms Severity scales of the 2010 American College of Rheumatology criteria of fibromyalgia. J Clin Med 2020;9(8):2460 [FREE Full text] [doi: 10.3390/jcm9082460] [Medline: 32752048]
46. Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. J Clin Nurs 2005;14(7):798-804. [doi: 10.1111/j.1365-2702.2005.01121.x] [Medline: 16000093]
47. Serlin R, Mendoza TR, Nakamura Y, Edwards KR, Cleeland CS. When is cancer pain mild, moderate or severe? Grading pain severity by its interference with function. Pain 1995;61(2):277-284. [doi: 10.1016/0304-3959(94)00178-H] [Medline: 7659438]
48. Alghadir A, Anwer S, Iqbal A, Iqbal Z. Test-retest reliability, validity, and minimum detectable change of visual analog, numerical rating, and verbal rating scales for measurement of osteoarthritic knee pain. J Pain Res 2018;11:851-856 [FREE Full text] [doi: 10.2147/JPR.S158847] [Medline: 29731662]
49. De Garceau D, Dean D, Requejo SM, Thordarson DB. The association between diagnosis of plantar fasciitis and Windlass test results. Foot Ankle Int 2003;24(3):251-255. [doi: 10.1177/107110070302400309] [Medline: 12793489]
50. Saban B, Masharawi Y. Three single leg standing tests for clinical assessment of chronic plantar heel pain syndrome: static stance, half-squat and heel rise. Physiotherapy 2017;103(2):237-244. [doi: 10.1016/j.physio.2016.06.003] [Medline: 27908459]
51. Visser TSS, Neill SO, Hébert-Losier K, Eygendaal D, de Vos R. Normative values for calf muscle strength-endurance in the general population assessed with the Calf Raise Application: a large international cross-sectional study. Braz J Phys Ther 2025;29(3):101188 [FREE Full text] [doi: 10.1016/j.bjpt.2025.101188] [Medline: 40020545]
52. Morikawa M, Maeda N, Komiya M, Hirota A, Mizuta R, Kobayashi T, et al. Contribution of plantar fascia and intrinsic foot muscles in a single-leg drop landing and repetitive rebound jumps: an ultrasound-based study. Int J Environ Res Public Health 2021;18(9):4511 [FREE Full text] [doi: 10.3390/ijerph18094511] [Medline: 33922807]
53. Arima S, Maeda N, Tsutsumi S, Ishihara H, Ikuta Y, Ushio K, et al. Association of the intrinsic foot muscles and plantar fascia with repetitive rebound jumping and jump landing in adolescent athletes: an ultrasound-based study. Gait Posture 2023;99:146-151. [doi: 10.1016/j.gaitpost.2022.11.003] [Medline: 36436441]
54. Simons D. Myofascial Pain and Dysfunction. Amsterdam, Netherlands: Wolters Kluwer; 1999.
55. Baellow A, Jaffri AH, Hertel J, Higgins MJ, Rangecroft CM, Hryvniak DJ, et al. Intrinsic foot muscle size and quality in a single leg weight bearing position across foot posture types in individuals with patellofemoral pain compared to healthy. Phys Ther Sport 2022;54:58-64. [doi: 10.1016/j.ptsp.2022.01.002] [Medline: 35085911]
56. Park G, Kim CW, Park SB, Kim MJ, Jang SH. Reliability and usefulness of the pressure pain threshold measurement in patients with myofascial pain. Ann Rehabil Med 2011;35(3):412-417 [FREE Full text] [doi: 10.5535/arm.2011.35.3.412] [Medline: 22506152]
57. Fischer AA. Pressure threshold meter: its use for quantification of tender spots. Arch Phys Med Rehabil 1986;67(11):836-838. [Medline: 3778185]
58. Hong CZ. Algometry in evaluation of trigger points and referred pain. J Musculoskelet Pain 2010;6(1):47-59. [doi: 10.1300/j094v06n01_04]
59. Pöntinen PJ. Reliability, validity, reproducibility of algometry in diagnosis of active and latent tender spots and trigger points. J Musculoskelet Pain 2010;6(1):61-71. [doi: 10.1300/j094v06n01_05]
60. Kinser AM, Sands WA, Stone MH. Reliability and validity of a pressure algometer. J Strength Cond Res 2009;23(1):312-314. [doi: 10.1519/jsc.0b013e31818f051c] [Medline: 19130648]
61. Ostermann S, Olesen JL, Holden S, Riel H. Stretching and relaxing the plantar fascia may change plantar fascia thickness but not pressure pain thresholds: a cross-sectional study of patients with plantar fasciopathy. BMC Musculoskelet Disord 2020;21(1):804 [FREE Full text] [doi: 10.1186/s12891-020-03833-x] [Medline: 33272236]
62. Plaza-Manzano G, Ríos-León M, Martín-Casas P, Arendt-Nielsen L, Fernández-de-Las-Peñas C, Ortega-Santiago R. Widespread pressure pain hypersensitivity in musculoskeletal and nerve trunk areas as a sign of altered nociceptive processing in unilateral plantar heel pain. J Pain 2019;20(1):60-67 [FREE Full text] [doi: 10.1016/j.jpain.2018.08.001] [Medline: 30121357]

XSL•FO
RenderX

63. Linde LD, Kumbhare DA, Joshi M, Srbely JZ. The relationship between rate of algometer application and pain pressure threshold in the assessment of myofascial trigger point sensitivity. Pain Pract 2018;18(2):224-229. [doi: 10.1111/papr.12597] [Medline: 28440895]
64. DeWall RJ, Slane LC, Lee KS, Thelen DG. Spatial variations in Achilles tendon shear wave speed. J Biomech 2014;47(11):2685-2692 [FREE Full text] [doi: 10.1016/j.jbiomech.2014.05.008] [Medline: 24933528]
65. Chimenti RL, Flemister AS, Ketz J, Bucklin M, Buckley MR, Richards MS. Ultrasound strain mapping of Achilles tendon compressive strain patterns during dorsiflexion. J Biomech 2016;49(1):39-44 [FREE Full text] [doi: 10.1016/j.jbiomech.2015.11.008] [Medline: 26655590]
66. Sikdar S, Ortiz R, Gebreab T, Gerber LH, Shah JP. Understanding the vascular environment of myofascial trigger points using ultrasonic imaging and computational modeling. In: Annu Int Conf IEEE Eng Med Biol Soc. 2010 Presented at: Annual International Conference of the IEEE Engineering in Medicine and Biology; August 31, 2010; Buenos Aires, Argentina p. 5302-5305 URL: https://europepmc.org/abstract/MED/21096064 [doi: 10.1109/IEMBS.2010.5626326]
67. Wiesinger F, Sacolick LI, Menini A, Kaushik SS, Ahn S, Veit-Haibach P, et al. Zero TE MR bone imaging in the head. Magn Reson Med 2016;75(1):107-114. [doi: 10.1002/mrm.25545] [Medline: 25639956]
68. Puel U, Lombard C, Hossu G, Louis M, Blum A, Teixeira PAG, et al. Zero echo time MRI in shoulder MRI protocols for the diagnosis of rotator cuff calcific tendinopathy improves identification of calcific deposits compared to conventional MR sequences but remains sub-optimal compared to radiographs. Eur Radiol 2023;33(9):6381-6391. [doi: 10.1007/s00330-023-09602-3] [Medline: 37014406]
69. Johnson CP, Heo H, Thedens DR, Wemmie JA, Magnotta VA. Rapid acquisition strategy for functional T1ρ mapping of the brain. Magn Reson Imaging 2014;32(9):1067-1077 [FREE Full text] [doi: 10.1016/j.mri.2014.07.010] [Medline: 25093630]
70. Magnotta VA, Xu J, Fiedorowicz JG, Williams A, Shaffer J, Christensen G, et al. Metabolic abnormalities in the basal ganglia and cerebellum in bipolar disorder: a multi-modal MR study. J Affect Disord 2022;301:390-399 [FREE Full text] [doi: 10.1016/j.jad.2022.01.052] [Medline: 35031333]
71. Reeder SB, Wen Z, Yu H, Pineda AR, Gold GE, Markl M, et al. Multicoil dixon chemical species separation with an iterative least-squares estimation method. Magn Reson Med 2004;51(1):35-45. [doi: 10.1002/mrm.10675] [Medline: 14705043]
72. Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB. Multiecho water-fat separation and simultaneous R2* estimation with multifrequency fat spectrum modeling. Magn Reson Med 2008;60(5):1122-1134 [FREE Full text] [doi: 10.1002/mrm.21737] [Medline: 18956464]
73. Benjamini Y. and Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 1995;57(1):a-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
74. Zale KE, Klatt M, Volz KR, Kanner C, Evans KD. A mixed-method approach to evaluating the association between myofascial trigger points and ankle/foot pain using handheld sonography equipment. J Diagn Med Sonogr 2015;31(4):210-220 [FREE Full text] [doi: 10.1177/8756479315586396]
75. Sacolick LI, Wiesinger F, Hancu I, Vogel MW. B1 mapping by Bloch-Siegert shift. Magn Reson Med 2010;63(5):1315-1322 [FREE Full text] [doi: 10.1002/mrm.22357] [Medline: 20432302]
76. Sluka KA, Wager TD, Sutherland SP, Labosky PA, Balach T, Bayman EO, et al. Predicting chronic postsurgical pain: current evidence and a novel program to develop predictive biomarker signatures. Pain 2023;164(9):1912-1926 [FREE Full text] [doi: 10.1097/j.pain.0000000000002938] [Medline: 37326643]

Abbreviations

HIPAA: Health Insurance Portability and Accountability Act
IDEAL: Iterative Decomposition of water and fat with Echo Asymmetry and Least-squares estimation–Quantitative Imaging
IVIM: intravoxel incoherent motion
MRI: magnetic resonance imaging
NRS: numeric rating scale
ROC AUC: area under the receiver operating characteristic curve
ROI: regions of interest
SWE: shear wave elastography

XSL•FO
RenderX

Edited by J Sarvestan; submitted 17.Nov.2025; this is a non–peer-reviewed article;accepted 16.Dec.2025; published 11.Feb.2026.

Please cite as:
Amerian Z, Fleagle T, Tuladhar U, Watson R, Wong M, Van Gorp B, Smith BJ, Richards M, Hall M, Donnelly J, Danielson J, Leao RV, Stanley D, Magnotta V, Holmes JH, Sluka K, Chimenti RL
Quantitative Understanding of Advanced Novel Imaging Techniques for Fasciitis and Biosignature Yield (Quantify): Protocol for a Cross-Sectional Diagnostic Study
JMIR Res Protoc 2026;15:e87613
URL: https://www.researchprotocols.org/2026/1/e87613
doi:10.2196/87613
PMID:

©Zahra Amerian, Timothy Fleagle, Utsav Tuladhar, Rachel Watson, Micah Wong, Barbara Van Gorp, Brian J Smith, Michael Richards, Mederic Hall, Joe Donnelly, Jessica Danielson, Renata Vidal Leao, David Stanley, Vincent Magnotta, James H Holmes, Kathleen Sluka, Ruth L Chimenti. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 11.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Differential Factors Associated With the Presence of Persistent Symptoms in Individuals Diagnosed With Long COVID: Protocol for a Longitudinal Matched Case-Control Study

David Lerma-Irureta[1,2], MPH, MD; Fátima Méndez-López[3], MSc, PhD; Elena Navarro-Matías[4]; Diego Lerma-Puertas[5,6], MD, PhD; Rosa Magallón-Botaya[1], MD, PhD

[1]Department of Medicine, Psychiatry and Dermatology, University of Zaragoza, Zaragoza, Spain, Zaragoza, Spain

[2]Instituto de Investigación Sanitaria Aragón, Zaragoza, Spain

[3]Universidad de Zaragoza, Faculty of Health Sciences, Department of Physiatry and Nursery, Zaragoza, Spain

[4]Instituto de Investigación Biomédica de Salamanca, Salamanca, Spain

[5]Obstetrics Department, Hospital Clínico Universitario Lozano Blesa, Zaragoza, Spain

[6]Department of Surgery, Gynecology, and Obstetrics, Universidad de Zaragoza, Zaragoza, Spain

Corresponding Author:
Fátima Méndez-López, MSc, PhD
Universidad de Zaragoza
Faculty of Health Sciences
Department of Physiatry and Nursery
Facultad de Ciencias de la Salud (Campus San Francisco, cerca de Pedro Cerbuna 12).
Zaragoza
Spain
Phone: 34 976 76 17 19
Email: fmendez@unizar.es

Abstract

Background: Long COVID (postacute sequelae of SARS-CoV-2 infection) is a heterogeneous condition with persistent multisystem symptoms and substantial functional burden. Integrative longitudinal studies combining clinical phenotyping, lifestyle factors, and immunobiological markers are needed to clarify determinants of symptom persistence and inform risk stratification and targeted interventions.

Objective: This study aims to identify clinical, biological/immunological, and sociodemographic factors associated with Long COVID status by comparing individuals with persistent symptoms to matched recovered controls, and to evaluate longitudinal changes in symptoms and secondary outcomes over follow-up.

Methods: ARALongCOV is a longitudinal matched case-control observational study conducted in Aragón, Spain, including adults (≥18 years of age) with confirmed SARS-CoV-2 infection. Participants are recruited through 3 sources: the Long COVID Aragón Patient Association, the Aragón Health Service database, and primary care consultations. Long COVID and recovered participants are individually matched 1:1 (without replacement) on sex/gender (exact), age (±3 years), and date of acute COVID-19 diagnosis (±30 days). Outcomes include persistent symptoms and functioning and patient-reported outcomes (quality of life, physical activity, diet, sleep, mental health, functional status, cognitive performance, pain catastrophizing, and fatigue), alongside clinical variables and biochemical/immunological markers (including inflammatory and cytokine profiles, SARS-CoV-2 antispike immunoglobulin G serology, and viral reactivation serologies). Measurements are obtained at baseline (T0) and repeated at follow-up (T1) using standardized procedures.

Results: The study received ethics approval from the Clinical Research Ethics Committee of Aragón (PI21/278). Funding was provided by Instituto de Salud Carlos III through project PI22/01070 (cofunded by the European Union) for the period 2023-2027. Baseline assessments (T0) were initiated in late 2022/early 2023. As of February 2026, a total of 200 participants have been enrolled (n=100 Long COVID; n=100 recovered controls) and have completed T0; T1 assessments are scheduled for late 2025/early 2026 (~3-year follow-up for the earliest enrolled participants). Primary analyses will be conducted after completion of T1 assessments, with dissemination planned from the second half of 2026 and continuing through 2027.

Conclusions: This protocol describes a comprehensive, multidimensional longitudinal study designed to clarify determinants of Long COVID by integrating clinical, functional, lifestyle, and immunobiological data in matched cohorts. Findings are expected

XSL•FO
RenderX

to support risk stratification, phenotype discovery, and identification of prognostic markers to inform preventive, diagnostic, and rehabilitative strategies.

Trial Registration: ISRCTN Registry ISRCTN27312680; https://tinyurl.com/33cbysrk

International Registered Report Identifier (IRRID): DERR1-10.2196/67133

(JMIR Res Protoc 2026;15:e67133) doi:10.2196/67133

KEYWORDS

biomarkers; cohort study; lifestyle factors; Long COVID; machine learning; persistent symptoms; postacute sequelae of SARS-CoV-2 infection; quality of life

Introduction

Overview

The COVID-19 pandemic has substantially reshaped the global health burden by producing not only acute morbidity and mortality but also a spectrum of persistent postinfectious sequelae that continue to affect large numbers of survivors [1-3]. One of the most clinically and socioeconomically consequential of these sequelae is Long COVID, frequently termed postacute sequelae of SARS-CoV-2 infection (PASC), a heterogeneous syndrome comprising persistent, relapsing, or newly emergent symptoms after the acute phase of infection [4-6]. Population and cohort analyses show appreciable heterogeneity in estimates of prevalence across definitions, follow-up intervals, and dominant variants, with many studies reporting persistent symptoms in the approximate 10%-30% range among infected persons depending on case definition and follow-up timing [7-9]. Typical symptom clusters include debilitating fatigue, dyspnea/exertional intolerance, and cognitive dysfunction (often described as “brain fog”), and these complaints occur in individuals across the spectrum of acute disease severity, including those who had mild acute illness [2,8,9].

Mechanistic concepts for PASC are multifactorial and emphasize an interplay of viral, immune, and host-environment factors, including potential viral persistence or residual antigen, dysregulated or protracted immune activation, chronic systemic and tissue-specific inflammation, perturbations of metabolic pathways, and, in some patients, autoimmune sequelae [5,10,11]. Proteomic and metabolomic studies and longitudinal immune profiling have reported persistent alterations in circulating proteins and immune mediators—including reported associations with proinflammatory cytokines such as interleukin-6 (IL-6) and tumor necrosis factor-α—although results vary across cohorts and time points [10,12,13]. At the same time, psychosocial determinants (eg, preexisting psychiatric morbidity, chronic stress, and social isolation) and baseline comorbidity burden (eg, obesity, cardiopulmonary disease, and chronic kidney disease) have been associated with greater acute severity and may influence recovery trajectories, symptom perception, and care seeking after SARS-CoV-2 infection [14-17].

Despite rapid expansion of descriptive and mechanistic work, the extant literature remains fragmented: many investigations target single domains (immunologic, imaging, symptom inventories, or psychosocial measures) or selected subpopulations (hospitalized cases, clinic-referred patients, or pediatric cohorts), limiting integrative inference and generalizability [5,18]. Large electronic health record–driven and population-based studies have mapped cooccurrence patterns and identified candidate risk factors, but frequently lack prospectively collected multimodal biomarker and psychosocial data required to link biological signals to clinically meaningful longitudinal trajectories [6,19]. Neurocognitive sequelae exemplify this gap: observational and translational work implicates persistent neuroinflammation and systemic immune perturbations as plausible contributors to post–COVID-19 cognitive complaints and to accelerated decline among susceptible individuals, but definitive, integrative longitudinal evidence is limited [20-22]. Taken together, these observations emphasize the need for prospective, multimodal cohorts that combine standardized clinical phenotyping, serial biosampling, and structured psychosocial measurement to permit hypothesis-driven biomarker discovery and etiologic modeling of PASC [3,23].

Demographic factors and baseline comorbidities modulate both acute disease course and postacute outcomes: older age and multimorbidity are consistently associated with a higher risk of persistent symptoms and adverse long-term outcomes in many cohort studies, and prior cognitive impairment has been identified as a vulnerability factor for prolonged neuropsychiatric sequelae [9,24-26]. Obesity and chronic respiratory disease have been linked to greater acute severity and higher likelihood of prolonged symptoms, and these associations have implications for targeted surveillance and rehabilitation strategies [17,27,28]. Collectively, these epidemiologic and mechanistic observations underscore the need for health systems to develop evidence-based approaches to identify at-risk populations early and to tailor prevention, monitoring, and rehabilitation resources accordingly [3,23,29].

In summary, although the body of research on Long COVID has expanded quickly, there remains an urgent requirement for coordinated, integrative studies that synthesize clinical phenotyping, longitudinal immunology, objective physiologic assessment, and psychosocial characterization, with the ultimate goal of deriving reproducible biomarkers, mechanism-informed subtypes, and actionable interventions for PASC [3,5,6,12,23].

Hypothesis and Rationale

We hypothesize that antecedent and periacute factors (eg, sociodemographic, preinfection comorbidities, and acute COVID-19 severity indicators) are associated with Long COVID status, and that baseline postacute measures collected at T0 (eg, patient-reported outcome measures and biochemical/immunological markers) primarily differentiate

XSL•FO
RenderX

Long COVID from recovered controls as concurrent phenotypic correlates rather than implying causality. By comparing individuals with Long COVID to carefully matched recovered controls and repeating standardized assessments longitudinally, we aim to identify differential factors across clinical, immunological, and lifestyle domains. These findings are expected to support biomarker discovery, mechanism-informed subtyping, and clinically actionable risk stratification in future studies.

Study Objectives

Primary Objective

The primary objective is to estimate associations between antecedent/periacute clinical and sociodemographic factors and Long COVID status, and to characterize postacute T0 clinical, biomarker, and psychosocial profiles as phenotypic differentiators between cases and matched recovered controls.

Secondary Objectives

The secondary objectives are as follows:

1. To characterize the Long COVID cohort with respect to demographics, comorbidity burden, and lifestyle habits.
2. To evaluate associations between lifestyle factors and preexisting conditions and the presence and severity of persistent symptoms and functional impairment.
3. To examine longitudinal change in symptoms and patient-reported outcomes between T0 and T1 within each group and compare trajectories between groups.

This protocol therefore describes a longitudinal matched case-control study designed to integrate clinical, functional, lifestyle, and immunobiological data to better understand determinants of persistent post–COVID-19 symptoms.

Methods

Study Design

This study is a longitudinal matched case-control study design combining a retrospective cohort with a prospective cohort. The retrospective cohort comprises individuals diagnosed during the early stages of the pandemic and provides long-term follow-up. In parallel, the prospective cohort will enroll newly identified COVID-19 cases, with follow-up assessments conducted after the baseline evaluation. This combined approach allows examination of symptom persistence for up to 3 years in the retrospective cohort while also enabling real-time tracking of changes over time in the prospectively enrolled participants. The study is registered in the International Standard Randomized Controlled Trial Number (ISRCTN) Registry (ISRCTN27312680), and reporting will adhere to the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines.

Study Setting and Population

The study will be conducted in Aragón, Spain, leveraging the region's established primary care network and health care infrastructure. Eligible participants will be adults (≥18 years of age) with a confirmed diagnosis of COVID-19. Participants will be identified through three complementary sources: (1) the Long COVID Aragón Patient Association (patients already registered with the association), (2) the Aragón Health Service database (individuals identified from electronic health records), and (3) primary care consultations (direct referrals from general practitioners and other health care professionals).

The Long COVID Aragón Patient Association is a regional, patient-led organization based in Aragón, Spain, that provides peer support, advocacy, and information for individuals experiencing persistent symptoms following SARS-CoV-2 infection. The association was established in 2020 in response to the emergence of postacute sequelae of COVID-19 and maintains a secure registry of members who voluntarily enroll and provide contact information and basic clinical descriptors (eg, date of acute infection, persistence and type of symptoms, and care pathways). Members who indicate willingness to be contacted for research constitute an important recruitment frame for this study, enabling identification of community-dwelling individuals with persistent symptoms who may not be fully captured through health care databases alone. For eligible individuals recruited through the association, study participation will proceed via the same standardized eligibility assessment and consent procedures used for all recruitment sources, and (where applicable and permitted) self-reported information will be verified or complemented using clinical records.

Eligibility Criteria

Participants will be included if they have confirmed SARS-CoV-2 infection by polymerase chain reaction, antigen testing, or serology, and are aged ≥18 years. For the Long COVID group, eligibility will require persistence of symptoms for ≥12 weeks after infection onset, consistent with the World Health Organization (WHO) definition. For the control group, complete recovery will be defined as the absence of persistent or new symptoms after COVID-19. Individuals will be excluded if they refuse to participate, report preexisting symptoms similar to Long COVID prior to SARS-CoV-2 infection, or present severe psychiatric disorders or cognitive impairment that preclude informed participation.

Sample Size Calculation

Because the primary aim of this study is hypothesis-generating—that is, to identify differential factors associated with Long COVID status rather than to test a single prespecified end point—sample size was chosen to provide adequate power to detect moderate associations between candidate predictors and case status within a matched case-control framework. Power was evaluated using the standard normal approximation for comparing 2 independent proportions (Fleiss et al [30]), with a 2-sided significance level of α=.05 and 90% power. Under a conservative scenario in which a candidate factor has an expected prevalence of 20% among controls, we considered an association of odds ratio 2.3 to be clinically meaningful; this corresponds to an expected prevalence of approximately 36.5% among cases. Under these assumptions, the required analyzable sample size is approximately 154 participants per group, rounded to 155 per group. To accommodate incomplete matching, potential attrition between T0 and T1, and planned multivariable and subgroup

XSL•FO
RenderX

analyses, the overall recruitment target will be increased to approximately 400 participants [30].

Matching Strategy

Recovered controls will be individually matched 1:1 to Long COVID cases without replacement to improve baseline comparability. Matching will require exact matching on sex/gender, with flexibility of ±3 years for age and ±30 days for date of acute COVID-19 diagnosis. When multiple eligible controls are available for a given case, the selected control will be the one minimizing the absolute age difference; if ties remain, the absolute diagnosis date difference will be minimized, with random selection used only if ties persist. Each matched set will be assigned a unique pair identifier (pair_id) and retained in the analytical dataset; participants who cannot be matched within the prespecified calipers will be retained in the study database and considered in prespecified sensitivity analyses.

Data Collection and Management

After written informed consent is obtained, each participant will be assigned a unique internal study code to ensure confidentiality. The code key linking personal identifiers to the study code will be accessible only to the principal investigator and will be stored separately from study data, in accordance with the General Data Protection Regulation (GDPR) and applicable Spanish data protection legislation. All study information will be recorded and managed using secure electronic data capture procedures; self-administered questionnaires will be completed via REDCap (Research Electronic Data Capture; Vanderbilt University) software [31]. Study datasets will be deidentified/anonymized for analysis and will be stored securely, with participants linked only through the assigned study code.

Appointments and Procedures

Data collection will be organized into 2 study appointments. During the first appointment, fasting blood samples will be collected for routine laboratory analyses, and participants will sign informed consent covering both study participation and biobank sample storage. Participants will also receive standardized instructions for completing the self-administered questionnaires online through REDCap [31]. The second appointment will include completion of exploratory tests and the collection of any additional clinical or sociodemographic data not captured at the first visit. Standardized procedures will be used at each visit to minimize measurement variability and ensure comparability across participants and across time points. The participant flow and assessment schedule are summarized in Figure 1.

XSL•FO
RenderX

Figure 1. Participant flow diagram. 1MSTST: 1-minute sit-to-stand test; 6MWT: 6-minute walk test; ACC: accuracy; AUC: area under the receiver operating characteristic curve; FSS: Fatigue Severity Scale; HADS: Hospital Anxiety and Depression Scale; IPAQ-SF: International Physical Activity Questionnaire-Short Form; ML: machine learning; PCFSS: Post–COVID-19 Functional Status Scale; PCS: Pain Catastrophizing Scale; PHC: primary health care; PHQ-9: Patient Health Questionnaire-9; PREDIMED: Prevención con DIeta Mediterránea; PSQI: Pittsburgh Sleep Quality Index; RF: random forest; SF-12: 12-Item Short Form Health Survey.

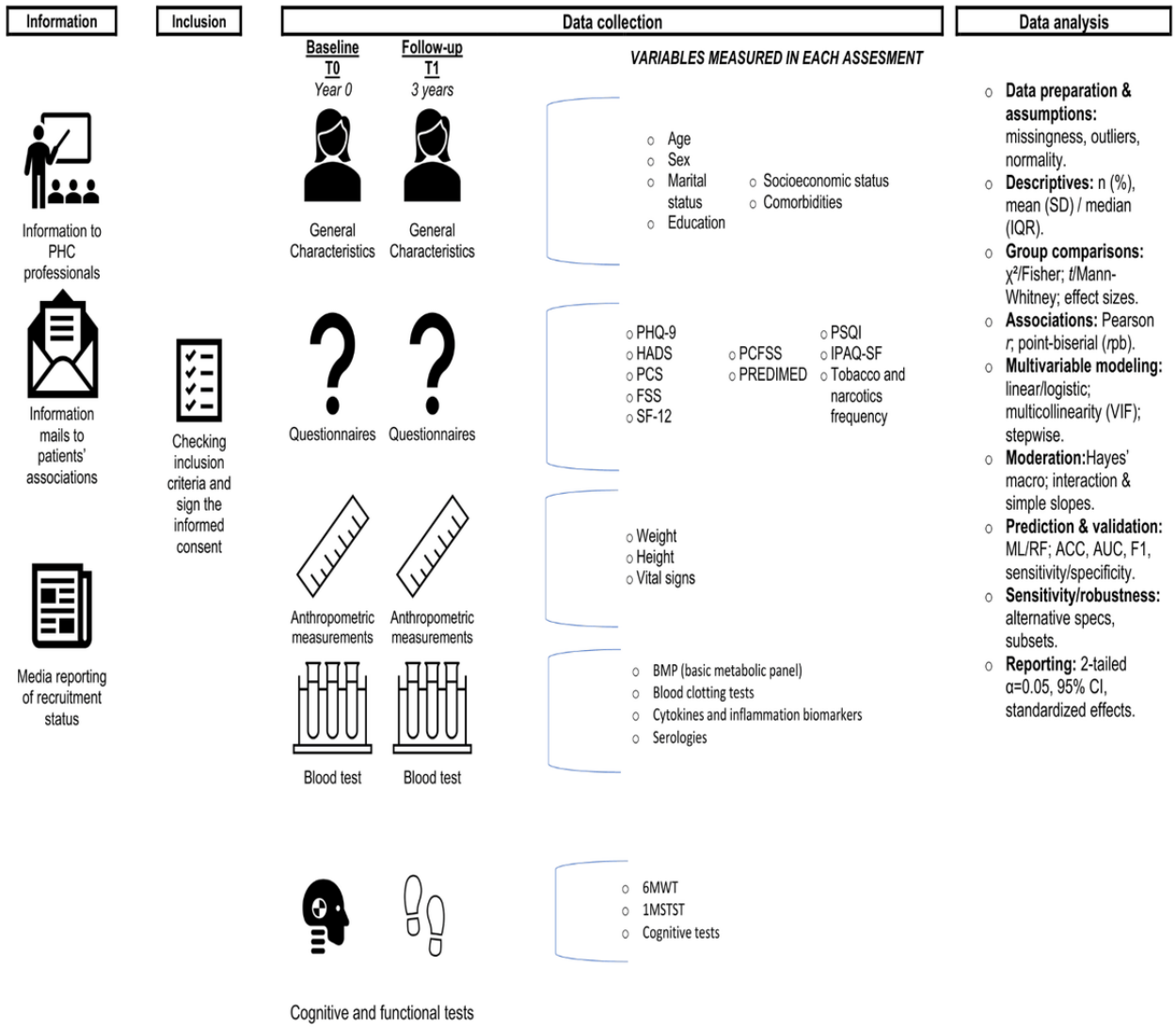

Variables and Measurements

Overview

Study variables will include outcomes related to persistent symptoms and functioning, additional patient-reported outcomes, cognitive performance measures, and laboratory biomarkers, alongside core sociodemographic and clinical covariates. Measurements will be obtained at T0 (baseline) and repeated at T1 (follow-up) using the same instruments and procedures to support longitudinal comparisons. Textbox 1 summarizes the study variables collected for sociodemographic, clinical, lifestyle, psychological/functional, and biological domains.

XSL•FO
RenderX

Textbox 1. Study variables.

Sociodemographic

- Age, sex/gender, place of residence, occupation, education level, and ethnicity

Clinical (acute COVID-19 episode)

- Date of symptom onset, symptom profile, hospitalization, intensive care unit admission, and diagnostic method (polymerase chain reaction, antigen, and serology)

Comorbidities

- Cardiovascular disease, diabetes mellitus, chronic respiratory disease, and other chronic conditions

Lifestyle

- Physical activity level (International Physical Activity Questionnaire), dietary habits (Prevención con Dieta Mediterránea), sleep quality (Pittsburgh Sleep Quality Index), smoking status, and alcohol consumption

Psychological and functional

- Quality of life (12-item Short Form Health Survey), depression (Patient Health Questionnaire-9), anxiety (Generalized Anxiety Disorder-7), functional status (Post–COVID-19 Functional Status Scale), cognitive flexibility and attention (Stroop Test), processing speed (Symbol Digit Modalities Test), visuospatial memory (Rey-Osterrieth), pain catastrophizing (Pain Catastrophizing Scale), and fatigue severity (Fatigue Severity Scale)

Biological

- Complete blood count, basic biochemistry, coagulation profile, inflammatory markers (C-reactive protein and interleukin-6), cytokine profile, and SARS-CoV-2 serology

Primary Outcome

The primary outcome will be Long COVID status (case vs recovered control). Cases will be defined as individuals with confirmed SARS-CoV-2 infection who report persistent symptoms for ≥12 weeks after infection onset, consistent with the WHO definition. Controls will be defined as individuals with a confirmed SARS-CoV-2 infection who report complete recovery, operationalized as the absence of persistent or new symptoms after COVID-19.

Secondary Outcomes

Secondary outcomes will encompass patient-reported and performance-based domains including health-related quality of life, physical activity, dietary pattern/adherence, sleep quality, depressive and anxiety symptoms, post–COVID-19 functional status, cognitive performance, as well as pain catastrophizing, fatigue severity, and laboratory biomarkers (biochemistry/hematology/coagulation, inflammatory markers including C-reactive protein [CRP] and IL-6, cytokine profiling, SARS-CoV-2 serology, and viral reactivation serologies such as Epstein-Barr virus [EBV], cytomegalovirus, and herpes simplex virus type 1 [HSV-1]). The presence and severity of persistent symptoms consistent with Long COVID (eg, fatigue, dyspnea, and cognitive complaints), will be operationalized using standardized symptom assessment procedures within the Long COVID group and contrasted with the absence of persistent symptoms in the control group. Symptom severity will be quantified using the prespecified instruments and scales described in the “Sociodemographic and Clinical Variables” section, enabling both between-group comparisons at each time point and within-participant change from T0 to T1.

These outcomes will be analyzed as complementary indicators of disease burden and functioning and as potential correlates of persistent symptom status and its evolution over time.

Sociodemographic and Clinical Variables

Key sociodemographic variables will include age, sex/gender, place of residence, occupation, education, and ethnicity. Clinical history related to the acute SARS-CoV-2 infection will include onset date, acute symptoms, hospitalization, and intensive care unit (ICU) admission (if applicable), and the method used for diagnostic confirmation. Preexisting and incident comorbidities will be recorded, including (but not limited to) cardiovascular disease, diabetes, chronic respiratory disorders, and other relevant conditions, to support confounding control and stratified analyses.

Lifestyle and Psychosocial Variables

Overview

All participant-reported outcomes will be collected at the baseline visit (T0) and repeated at the follow-up visit (T1; ~3 years after T0) using the same administration procedures. Self-administered questionnaires were completed in a standardized order, with staff available to resolve comprehension issues without coaching responses. Neurocognitive tests were administered by trained personnel following published manuals and standardized instructions. A detailed summary of instruments, scoring, recall periods, and key measurement properties is provided in Multimedia Appendix 1 [32-47].

Health-Related Quality of Life

Health-related quality of life was assessed using the Spanish 12-item Short Form Health Survey (SF-12; a 12-item short-form measure derived from the 36-item Short Form Health Survey

XSL•FO
RenderX

[SF-36]) that captures 8 health domains and yields 2 standardized summary scores: the physical component summary and mental component summary. Scores are computed using Spanish population weights and interpreted against population norms, with higher scores indicating better health status. In the Spanish validation work, the SF-12 showed acceptable metric performance and construct validity; notably, SF-12 items explained approximately 91% of the variance of SF-36 summary components, supporting the adequacy of the short form for estimating physical component summary and mental component summary in Spanish samples. Internal consistency for SF-12 summary measures has been reported in an acceptable-to-good range in Spain (with component reliability typically around $\alpha \approx .78$-$.85$) [32].

Physical Activity

Physical activity over the preceding week was measured using the International Physical Activity Questionnaire (IPAQ). The IPAQ quantifies time spent in activities of different intensities during the last 7 days, allowing the calculation of energy expenditure (eg, metabolic equivalent of task minutes per week) and categorical classifications of activity level. In the Spanish population, the IPAQ demonstrated good reproducibility and expected validity patterns: test-retest reliability was strong (Spearman $\rho \approx 0.81$; intraclass correlation coefficient [ICC]≈ 0.74), while criterion validity versus objective accelerometry was modest (approximately $\rho \approx 0.27$) and improved when compared with activity diaries (approximately $\rho \approx 0.38$), consistent with typical performance of self-reported physical activity measures [33].

Adherence to the Mediterranean Diet

Mediterranean diet adherence was assessed with the 14-item Mediterranean Diet Adherence Screener (MEDAS) developed within the Prevención con Dieta Mediterránea (PREDIMED) framework. The MEDAS is a brief checklist covering key Mediterranean diet targets; responses are summed to a total score from 0 to 14, with higher scores indicating greater adherence. As a dietary pattern screener (rather than a unidimensional psychometric scale), "validity" is primarily supported through its relationship with external criteria: in PREDIMED, the 14-item tool was used alongside a full-length food frequency questionnaire and showed a strong, graded association with obesity-related indices, supporting its capacity to capture meaningful differences in diet quality in high-risk adults [34].

Sleep Quality

Sleep quality was measured using the Pittsburgh Sleep Quality Index (PSQI), a 19-item questionnaire assessing sleep during the previous month and yielding 7 component scores (subjective sleep quality, latency, duration, efficiency, disturbances, sleep medication use, and daytime dysfunction) and a global score ranging from 0 to 21; higher scores indicate worse sleep quality. The original PSQI demonstrated strong clinical use and has been widely used in research settings [35]. In the Spanish version, clinometric evaluation supports its use, with internal consistency around $\alpha \approx 0.81$ and evidence of diagnostic performance using standard thresholds (with sensitivity in the high 80% range and specificity around the mid-70% range, depending on the criterion) [36].

Depressive Symptoms

Depressive symptom severity was assessed with the Patient Health Questionnaire-9 (PHQ-9), a 9-item self-report measure aligned with the *DSM* (*Diagnostic and Statistical Manual of Mental Disorders*) symptom criteria and anchored to the preceding 2 weeks. Each item is scored 0 ("not at all") to 3 ("nearly every day"), producing a total score from 0 to 27, where higher scores indicate more severe depressive symptoms. In Spanish primary care validation, the PHQ-9 showed strong internal consistency (reported as McDonald $\omega \approx 0.89$) and good screening performance for major depressive disorder, with high sensitivity and moderate specificity depending on the selected cut-point and reference standard [37].

Anxiety Symptoms

Anxiety symptom severity was measured with the Generalized Anxiety Disorder-7, a 7-item self-report scale also anchored to the preceding 2 weeks, using the same 0-3 response options and yielding a total score from 0 to 21; higher scores indicate more severe anxiety. The Spanish adaptation demonstrated excellent internal consistency ($\alpha \approx 0.94$) and strong short-term stability (test-retest ICC≈ 0.84), alongside evidence of concurrent validity through strong correlations with established anxiety and disability measures [38].

Post–COVID-19 Functional Status Scale

Functional limitations were assessed using the Post–COVID-19 Functional Status Scale (PCFS), administered as a self-completed web-based questionnaire (REDCap). The PCFS is an ordinal measure grading limitations from 0 (no functional limitations) to 4 (severe functional limitations), with higher grades indicating greater restriction in usual activities and lifestyle. Because PCFS is an ordinal functional status grade rather than a multi-item psychometric scale, internal consistency is not applicable; instead, measurement quality is supported by agreement and construct validity [39]. The Spanish web-based PCFS has demonstrated substantial test-retest reliability ($\kappa = 0.63$) and high agreement between web- and paper-based formats (88% agreement), alongside strong construct validity through expected correlations with health status and activity limitation measures (eg, EQ-5D-5L questionnaire and Global Activity Limitation Indicator) [40].

Cognitive Performance

Cognitive domains relevant to attention/executive control, processing speed, and visuoconstructive memory were assessed with standardized neuropsychological tests. Selective attention and inhibitory control were assessed using the Stroop Color and Word Test, generating condition scores and an interference index reflecting executive control under competing stimulus demands [41]. Processing speed and attention were assessed with the Symbol Digit Modalities Test, which measures rapid visual scanning, sustained attention, and speeded substitution under timed conditions [42]. Visuoconstruction and visual memory were evaluated using the Rey-Osterrieth Complex Figure, including copy accuracy and recall performance using standardized scoring procedures [43]. These instruments have

XSL•FO
RenderX

extensive normative use in clinical and experimental contexts, and standardized administration enhances reliability across assessment waves.

Pain Catastrophizing

Maladaptive pain-related cognition was measured with the Pain Catastrophizing Scale (PCS), a 13-item measure covering rumination, magnification, and helplessness. Items are typically rated on a 0-4 scale and summed to a total score from 0 to 52, where higher scores reflect greater catastrophizing. The original PCS demonstrated robust psychometric development and validation [44], and the Spanish version has shown excellent internal consistency (α≈.94) and strong test-retest reliability (ICC≈0.88), supporting its use in Spanish-speaking clinical populations [45].

Fatigue Severity

Fatigue was assessed using the Fatigue Severity Scale (FSS), a 9-item questionnaire capturing the impact and severity of fatigue on daily functioning. Responses are typically scored from 1 to 7 and summarized as a mean score (commonly interpreted with higher values indicating greater fatigue severity) [46]. In post–COVID-19 recovery clinic populations, the FSS showed excellent internal consistency (α≈0.96) and expected construct validity through correlations with quality of life and depressive symptoms, supporting its use for quantifying clinically relevant fatigue in this context [47].

Biochemical and Immunological Markers

Fasting venous blood samples will be collected at baseline (T0) and follow-up (T1) to characterize participants' biochemical, hematological, coagulation, inflammatory, and immunological profiles. All measurements will be performed in the routine hospital clinical laboratory according to standard operating procedures. Standard panels will include basic biochemistry, a complete blood count, and a coagulation profile. Systemic inflammation will be assessed using circulating inflammatory markers, including CRP and IL-6, and broader immune signatures will be characterized through cytokine profiling using a predefined panel. Humoral response to SARS-CoV-2 will be assessed using quantitative antispike immunoglobulin G serology, recorded as a continuous measure. Potential viral reactivation will be explored by measuring serological markers for latent herpesviruses, including EBV, cytomegalovirus, and HSV-1. Biomarker variables will be analyzed as candidate correlates of persistent symptoms and as longitudinal measures to evaluate change between T0 and T1.

Statistical Analysis

Overview

All analyses will be performed using IBM SPSS Statistics (version 27) and R software (R Foundation for Statistical Computing). Statistical inference will emphasize effect sizes with 95% CIs and clinically interpretable magnitudes; 2-sided tests will be used. Given the observational design, analyses will be interpreted as associations rather than causal effects. For interpretability, candidate "differential factors" will be grouped by temporality: (1) antecedent factors (eg, sociodemographic and preinfection comorbidities), (2) periacute factors (eg, markers of acute COVID-19 severity such as hospitalization/ICU admission and related clinical features), and (3) concurrent correlates measured at baseline in the postacute phase (eg, patient-reported outcome measures and biomarkers at T0). Antecedent and periacute factors will be interpreted as candidate determinants, whereas concurrent measures will be interpreted primarily as phenotypic correlates differentiating Long COVID from recovered controls.

Analysis of Populations and Data Structure

The primary analytical dataset will consist of 1:1 matched pairs of Long COVID cases and recovered controls, with matching performed without replacement on sex/gender (exact), age (±3 years), and date of acute COVID-19 diagnosis (±30 days), and a retained pair_id. Data will be organized (1) in wide format for descriptive summaries by group and time point and (2) in long format (1 row per participant per time point) for longitudinal models incorporating T0 and T1.

Objective 1 (Primary): Identify Factors Associated With Long COVID Status (Case vs Recovered Control)

Consistent with the study hypothesis that biopsychosocial characteristics, biomarkers, lifestyle factors, and comorbidities are associated with symptom persistence, the primary inferential analyses will model Long COVID status as the dependent variable.

As an initial descriptive step, we will summarize within-pair distributions and conduct paired comparisons as appropriate (eg, McNemar tests for binary variables; paired *t* tests, or Wilcoxon signed-rank tests for continuous variables). These comparisons will be used to describe patterns of differences and will not be used as the sole criterion for variable inclusion in multivariable models.

Adjusted associations with Long COVID status will be estimated using conditional logistic regression stratified by pair_id (matched sets), reporting adjusted odds ratios with 95% CIs. Model building will follow a prespecified, interpretable structure. A core adjustment set defined a priori will be included to address confounding not already controlled by matching (eg, acute disease severity indicators such as hospitalization/ICU admission where available, vaccination status/epoch if captured, and comorbidity burden). Candidate predictors will then be entered in clinically coherent blocks (eg, lifestyle/psychosocial block and biomarker block) to support interpretation of how associations change after adjustment for confounding and correlated predictors. Continuous predictors will be assessed for appropriate functional form (eg, linearity in the logit); flexible specifications (eg, splines) or clinically meaningful categories will be used when needed. Collinearity will be assessed (eg, variance inflation), and influential observations will be examined using standard diagnostics.

Biomarkers (routine laboratory tests, inflammatory markers including CRP and IL-6, cytokine panel, quantitative antispike immunoglobulin G, and viral reactivation serologies such as EBV/cytomegalovirus/HSV-1) will be treated as candidate correlates of Long COVID status. Continuous biomarkers will be summarized with appropriate transformations if skewed (eg,

XSL•FO
RenderX

log transformation), and extreme values will be checked for plausibility and laboratory flags prior to analysis.

To assess whether associations differ across clinically relevant strata, interaction terms (eg, predictor × sex/gender; predictor × acute severity; predictor × cohort type) will be evaluated where justified by the objective and sample size; interactions will be reported as stratum-specific estimates with uncertainty.

Objective 2: Characterize the Long COVID Cohort and Compare Profiles With Matched Recovered Controls

Baseline (T0) characteristics will be reported for cases and matched controls using paired summaries (within-pair differences where informative) and standard descriptive statistics (means [SD] or medians [IQR] for continuous variables and frequencies and percentages for categorical variables). Distributional assumptions will be checked (eg, Shapiro-Wilk for guidance, complemented by graphical inspection).

Objective 3: Evaluate Symptom and Outcome Evolution From T0 to T1

Longitudinal changes in continuous secondary outcomes (eg, questionnaire scores, cognitive test metrics, and selected continuous biomarkers) will be modeled using mixed-effects models including fixed effects for group (case vs control), time (T0 vs T1), and the group × time interaction to estimate differential change over time, with random intercepts for participants and appropriate correlation structures. Binary/ordinal outcomes will use generalized mixed-effects models with appropriate links (logit for binary; ordinal models for ordered scales where applicable). Primary longitudinal contrasts will be (1) within-group change and (2) between-group difference in change (interaction term), each reported with 95% CIs.

Exploratory Prediction/Risk Profiling (Secondary, Hypothesis-Generating)

To complement explanatory regression, predictive models will be developed to explore multivariable risk profiles. Random forest models will be used to rank variable importance and detect nonlinearities/interactions, followed by parsimonious regression-based models for interpretability. Model performance will be evaluated with internal validation (eg, k-fold cross-validation), reporting discrimination (area under the curve) and classification metrics; matched pairs will be kept together within the same resampling fold to avoid leakage across training/testing splits.

Missing Data, Multiplicity, and Sensitivity Analyses

Missingness will be summarized by variable and time point and assessed for patterns. If missingness is nontrivial for key predictors/outcomes, multiple imputation will be considered for analyses where assumptions are plausible; complete-case analyses will be presented as sensitivity checks. For families of secondary outcomes (multiple correlated questionnaire and biomarker measures), multiplicity control will be applied using a false discovery rate approach (eg, Benjamini-Hochberg) and reported alongside unadjusted estimates to preserve interpretability. Sensitivity analyses will include assessment of incomplete matching and, where necessary, inclusion of unmatched participants with regression adjustment for matching variables. Additional sensitivity analyses will evaluate robustness to alternative model specifications (eg, excluding highly collinear predictors and alternative transformations of skewed biomarkers) and to cohort type (retrospective vs prospective).

Ethical Considerations

The study was approved by the Clinical Research Ethics Committee of Aragón, Spain (Comité de Ética de la Investigación Clínica de Aragón; PI21/278) and will be conducted in accordance with the Declaration of Helsinki and applicable local regulations. The collection and processing of personal data will comply with the GDPR (European Union) 2016/679 and the Spanish Organic Law 3/2018 of December 5 on the protection of personal data and guarantee of digital rights.

Biological samples and associated data will be managed through the Biobank of the Aragón Health System (National Biobank Register B.0000873; PT20/00112), integrated within the Instituto de Salud Carlos III (ISCIII) Biobanks and Biomodels Platform. Samples and data will be handled according to approved standard operating procedures and under the oversight and authorization of the relevant ethics and scientific committees and will be used exclusively for the purposes described in the approved protocol.

All participants will provide written informed consent prior to any study procedures, including specific consent for biobanking where applicable. Signed consent forms will be securely archived by the principal investigator. No participant compensation is planned due to limited funding.

Privacy and Confidentiality

Participants will be assigned a unique internal study code at enrollment. Direct identifiers will be stored separately from research data, and the reidentification key linking identifiers to study codes will be accessible only to the principal investigator and stored in a restricted, encrypted location. Analytical datasets will be deidentified (coded) and will not include direct identifiers. Data will be collected and managed using secure electronic data capture systems with role-based access controls and audit trails and stored on secure institutional servers with restricted access. Data transfers, if required, will occur only via encrypted channels. Study results will be reported in aggregate form to minimize the risk of reidentification.

Results

Study Status, Approvals, and Registration

ARALongCOV is a registered observational study (ISRCTN27312680) with ethics approval from the Clinical Research Ethics Committee of Aragón (PI21/278). The study is conducted under GDPR and applicable Spanish data protection legislation, with biospecimens and associated data managed through the Biobank of the Aragón Health System (PT20/00112) following approved standard operating procedures. Funding was provided by the ISCIII through project PI22/01070 (cofounded by the European Union) for the period 2023-2027.

XSL•FO
RenderX

Timeline and Data Collection Milestones

The study combines (1) a retrospective cohort providing long-term follow-up (enabling evaluation of symptom persistence for up to 3 years) and (2) a prospective cohort enrolling newly identified Long COVID cases with follow-up assessments after baseline. Baseline assessments (T0) were initiated in late 2022/early 2023, and follow-up assessments (T1) are scheduled for late 2025/early 2026 for the earliest enrolled participants. Study assessments are conducted at 2 time points (baseline [T0] and follow-up [T1]) using the same instruments and standardized procedures to support longitudinal comparisons. Data collection is organized into 2 appointments per assessment wave, including fasting blood sampling for routine laboratory analyses and completion of online questionnaires via REDCap.

Recruitment and Follow-up Progress

The planned recruitment target is approximately 400 participants, exceeding the minimum analyzable sample size of 155 per group derived from power considerations for detecting moderate associations in the matched case-control framework. As of February 2026, a total of 200 participants have been enrolled (n=100 Long COVID; n=100 recovered controls) and have completed T0 assessments; T1 assessments are scheduled for late 2025/early 2026.

Planned Reporting of Findings

Primary analyses will compare matched Long COVID and recovered participants using methods that account for the paired design (eg, conditional logistic regression with pair strata) and will evaluate longitudinal change using mixed-effects models including group, time, and group × time interaction. The first manuscripts reporting baseline comparisons and longitudinal outcomes are planned for submission in the second half of 2026, followed by mechanistic/biomarker and predictive modeling outputs.

Discussion

Overview

In this protocol, we describe a longitudinal matched case-control study design that combines retrospective and prospective data collection to investigate the multifaceted determinants of Long COVID in adult populations; such mixed longitudinal approaches have proven useful for characterizing recovery trajectories and for linking clinical phenotypes with immune and proteomic signatures in prior cohort work [3,48,49]. Our design explicitly aims to overcome limitations of single-dimension studies by integrating clinical, functional, and laboratory domains within a single, longitudinal framework to delineate temporal relationships among symptoms, objective physiologic measures, and circulating biomarkers and thereby to identify candidate mechanistic pathways and potential intervention targets [10,14,18].

Strengths of the Study

This is a longitudinal matched case-control (1:1 matching) design. The prospective 3-year follow-up combined with retrospective characterization permits within-participant trajectory analysis and between-group comparisons that are less susceptible to the inferential constraints of purely cross-sectional or clinic-based samples [50].

The multimodal phenotyping and biospecimen through systematic collection of validated patient-reported instruments, functional measures, and comprehensive biomarker panels enhance the potential to identify biologically coherent PASC subtypes and to relate subjective symptoms to measurable biological perturbations [10,14,49,51].

In particular, including a matched recovered control group reduces confounding by age, sex, and calendar time (including variant and vaccination epoch) and helps distinguish sequelae specifically attributable to prior SARS-CoV-2 infection from those related to general post-illness or secular effects [5,18,24].

The use of machine learning and multivariable modeling (eg, random forests for variable importance and receiver operating characteristic analyses for predictive performance) enables exploration of high-dimensional interactions [52].

Limitations

Despite these strengths, several important limitations merit emphasis. First, the targeted sample size of roughly 400 participants may limit power to detect associations for low-prevalence outcomes and to perform highly stratified subgroup analyses or fully corrected high-dimensional biomarker discovery without external validation, a limitation recognized in therapeutic and mechanistic PASC research [3,23]. Second, geographic confinement to a single region (Aragón, Spain) may reduce external validity because regional differences in variant circulation, vaccination patterns, sociodemographic composition, and health care delivery can influence both exposure and postacute outcomes; this limitation underscores the rationale for multicenter and multinational replication [29]. Third, recruitment through patient associations, health-service registries, and clinical referrals risks selection bias because individuals with persistent symptoms or heightened health care engagement may be overrepresented, an issue documented across online and clinic-recruited PASC cohorts [53]. Fourth, retrospective self-report carries the potential for recall and misclassification bias, and longitudinal attrition can induce informative censoring if participants who drop out differ systematically from those retained, both of which are well-recognized challenges in longitudinal PASC investigations [4,48,50,54].

Future Research Directions

This protocol serves as a foundation for broader etiologic and interventional efforts. Key priorities for subsequent work include (1) expansion to multicenter/multinational networks to improve power and generalizability and to permit pooled analyses across variable systems and variant epochs [29], (2) harmonization of phenotyping and biospecimen protocols across cohorts to enable external validation and meta-analytic biomarker discovery [3,12,19], (3) incorporation of objective physiologic and imaging modalities (eg, cardiac and pulmonary imaging, cardiopulmonary exercise testing, and ambulatory wearable sensors) to corroborate and contextualize self-reported symptoms and to identify organ-level dysfunctions amenable

XSL•FO
RenderX

to targeted therapies [55], and (4) prolonged longitudinal follow-up beyond 1 year to capture late-emerging sequelae and to quantify long-term risks of incident cardiometabolic or autoimmune conditions that have been reported in some cohorts [13,56].

Broader Implications

Systematic identification of clinical, biological, and psychosocial determinants of Long COVID has immediate implications for clinical care and public health. Clinically, evidence of differential risk and mechanism-specific subtypes can inform stratified follow-up algorithms, personalized rehabilitation plans, and selection criteria for interventional trials [3,57]. From the public health and systems perspective, robust longitudinal data are necessary to develop standardized case definitions, to guide resource allocation for post–COVID-19 care pathways and rehabilitation services, and to underpin evidence-based policy and clinical guidance [23,29].

Conclusions

This longitudinal matched case-control study, with mixed retrospective-prospective protocol, is designed to generate multimodal, longitudinal evidence on clinical, biological, and psychosocial correlates of Long COVID, leveraging matched controls, standardized phenotyping, and modern analytic methods to identify reproducible markers of symptom persistence and potential targets for intervention [3-5,18,23,52]. The study acknowledges limitations related to sample size, geographic scope, and selection or information biases and therefore positions its findings as a basis for validation in larger, multicenter consortia and for informing longer-term follow-up studies and therapeutic trials [29]. Ultimately, integrated longitudinal investigations of the type described here are essential to reduce the long-term burden of SARS-CoV-2 infection through improved prevention, diagnosis, and management of PASC [3,5,6,58].

Acknowledgments

We would like to thank the Aragonese Primary Care Research Group (B21_23R), a part of the Department of Innovation, Research and University at the Government of Aragón (Spain), and the Research Network on Chronicity, Primary Care and Health Promotion (RD24/0005/0004) that is part of the Results-Oriented Cooperative Research Networks in Health (Resultados en Salud-Orientados a la Cooperación en Redes; ISCIII). In addition, we would like to thank other members of the Biocomputing Unit, IACS, Aragon Health Research Institute, and the Biobank of the Aragon Health System (PT20/00112) integrated in the ISCIII Biobanks and Biomodels Platform, for their continued support and interest. We particularly want to acknowledge all the participants for their collaboration in this study. We also thank the Long Covid Aragón Patients Association for their contributions and collaboration in carrying out the study.

The authors used generative artificial intelligence (AI) tools to support language editing and summarization during manuscript preparation, under full human supervision, in accordance with the GAIDeT taxonomy (2025). All AI-assisted outputs were reviewed, edited as needed, and verified by the authors. No generative AI tool was used for data collection, data management, statistical analysis, or the generation/interpretation of study results. The tools used were NotebookLM and Scite. Responsibility for the content and integrity of the final manuscript rests entirely with the authors; generative AI tools are not listed as authors and bear no responsibility for the work.

Funding

This study has been funded by the Instituto de Salud Carlos III through the project PI22/01070 and cofunded by the European Union.

Funders had no role in the study design, data collection and analysis, decision to publish, or manuscript preparation. The research will be audited once a year by the funding organization.

Data Availability

Because the study includes personal health data collected under GDPR and Spanish data protection law and involves biobank-linked samples, individual-level participant data cannot be made openly available without compromising confidentiality. A reproducibility package (study protocol, data dictionary/codebook, analysis plan, and analysis code/scripts) will be deposited in Zenodo; the public record will include a DOI and versioning to support transparent reuse. Access to deidentified individual-level datasets may be considered for bona fide research purposes on reasonable request and subject to (1) approval by the relevant ethics and/or scientific committees and biobank governance bodies, (2) a signed data use agreement, and (3) compliance with applicable legal requirements.

Authors' Contributions

Conceptualization: RM-B
Methodology: DL-I, FM-L, RM-B
Writing — original draft: RM-B, DL-I
Writing — review & editing: DL-I, FM-L, EN-M, DL-P, RM-B

XSL•FO
RenderX

Project administration: RM-B
Supervision: RM-B

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of validated instruments, scoring, and psychometric properties.
[DOCX File , 3058 KB - resprot_v15i1e67133_app1.docx]

Multimedia Appendix 2
Original peer reviewed protocol and grant.
[PDF File (Adobe PDF File), 332 KB - resprot_v15i1e67133_app2.pdf]

References

1. Al-Aly Z, Bowe B, Xie Y. Long COVID after breakthrough SARS-CoV-2 infection. Nat Med 2022;28(7):1461-1467 [FREE Full text] [doi: 10.1038/s41591-022-01840-0] [Medline: 35614233]
2. Sharma SK, Mohan A, Upadhyay V. Long COVID syndrome: an unfolding enigma. Indian J Med Res 2024;159(6):585-600 [FREE Full text] [doi: 10.25259/IJMR_1449_23] [Medline: 39382470]
3. Bonilla H, Peluso MJ, Rodgers K, Aberg JA, Patterson TF, Tamburro R, et al. Therapeutic trials for Long COVID-19: a call to action from the interventions taskforce of the RECOVER initiative. Front Immunol 2023;14:1129459 [FREE Full text] [doi: 10.3389/fimmu.2023.1129459] [Medline: 36969241]
4. Sneller MC, Liang CJ, Marques AR, Chung JY, Shanbhag SM, Fontana JR, et al. A longitudinal study of COVID-19 sequelae and immunity: baseline findings. Ann Intern Med 2022;175(7):969-979 [FREE Full text] [doi: 10.7326/M21-4905] [Medline: 35605238]
5. Zhang H, Zang C, Xu Z, Zhang Y, Xu J, Bian J, et al. Data-driven identification of post-acute SARS-CoV-2 infection subphenotypes. Nat Med 2023;29(1):226-235 [FREE Full text] [doi: 10.1038/s41591-022-02116-3] [Medline: 36456834]
6. Thompson RC, Simons NW, Wilkins L, Cheng E, Del Valle DM, Hoffman GE, Mount Sinai COVID-19 Biobank Team, et al. Molecular states during acute COVID-19 reveal distinct etiologies of long-term sequelae. Nat Med 2023;29(1):236-246 [FREE Full text] [doi: 10.1038/s41591-022-02107-4] [Medline: 36482101]
7. Lewnard JA, Malden DE, Hong V, Skela J, Feldstein LR, Saydah S, et al. Comparative risk of post-acute sequelae following SARS-CoV-2 or influenza virus infection: a retrospective cohort study among United States adults. PLoS Med 2025;22(10):e1004777 [FREE Full text] [doi: 10.1371/journal.pmed.1004777] [Medline: 41066528]
8. Ballouz T, Menges D, Kaufmann M, Amati R, Frei A, von Wyl V, et al. Post COVID-19 condition after Wildtype, Delta, and Omicron SARS-CoV-2 infection and prior vaccination: pooled analysis of two population-based cohorts. PLoS One 2023;18(2):e0281429 [FREE Full text] [doi: 10.1371/journal.pone.0281429] [Medline: 36812215]
9. Richard SA, Pollett SD, Fries AC, Berjohn CM, Maves RC, Lalani T, Epidemiology, Immunology,Clinical Characteristics of Emerging Infectious Diseases With Pandemic Potential (EPICC) COVID-19 Cohort Study Group. Persistent COVID-19 symptoms at 6 months after onset and the role of vaccination before or after SARS-CoV-2 infection. JAMA Netw Open 2023;6(1):e2251360 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.51360] [Medline: 36652247]
10. Talla A, Vasaikar SV, Szeto GL, Lemos MP, Czartoski JL, MacMillan H, et al. Persistent serum protein signatures define an inflammatory subcategory of long COVID. Nat Commun 2023;14(1):3417 [FREE Full text] [doi: 10.1038/s41467-023-38682-4] [Medline: 37296110]
11. Block T, Kuo J. Rationale for nicotinamide adenine dinucleotide (NAD+) metabolome disruption as a pathogenic mechanism of post-acute COVID-19 syndrome. Clin Pathol 2022;15:2632010X221106986 [FREE Full text] [doi: 10.1177/2632010X221106986] [Medline: 35769168]
12. Agius JE, Johnson-Mackinnon JC, Fong W, Gall M, Lam C, Basile K, et al. SARS-CoV-2 within-host and genomic variability and sub-genomic RNA levels indicate differences in viral expression between clinical cohorts and culture. Front Microbiol 2022;13:824217 [FREE Full text] [doi: 10.3389/fmicb.2022.824217] [Medline: 35663867]
13. Rahmati M, Yon DK, Lee SW, Udeh R, McEVoy M, Kim MS, et al. New-onset type 1 diabetes in children and adolescents as postacute sequelae of SARS-CoV-2 infection: a systematic review and meta-analysis of cohort studies. J Med Virol 2023;95(6):e28833. [doi: 10.1002/jmv.28833] [Medline: 37264687]
14. Huapaya JA, Gairhe S, Kanth S, Tian X, Demirkale CY, Regenold D, et al. Alterations in the plasma proteome persist ten months after recovery from mild to moderate SARS-CoV-2 infection. Front Immunol 2024;15:1448780 [FREE Full text] [doi: 10.3389/fimmu.2024.1448780] [Medline: 39324144]
15. Wang S, Quan L, Ding M, Kang JH, Koenen KC, Kubzansky LD, et al. Depression, worry, and loneliness are associated with subsequent risk of hospitalization for COVID-19: a prospective study. Psychol Med 2023;53(9):4022-4031 [FREE Full text] [doi: 10.1017/S0033291722000691] [Medline: 35586906]

XSL•FO
RenderX

16. Nishimi K, Neylan TC, Bertenthal D, Seal KH, O'Donovan A. Association of psychiatric disorders with clinical diagnosis of Long COVID in US veterans. Psychol Med 2024;54(9):2024-2032 [FREE Full text] [doi: 10.1017/S0033291724000114] [Medline: 38311905]
17. Carter SJ, Baranauskas MN. Why obesity and psychological stress matter in recovery of post-acute sequelae of SARS-CoV-2. Obesity (Silver Spring) 2022;30(6):1136-1138 [FREE Full text] [doi: 10.1002/oby.23442] [Medline: 35352508]
18. Zang C, Hou Y, Schenck EJ, Xu Z, Zhang Y, Xu J, et al. Identification of risk factors of Long COVID and predictive modeling in the RECOVER EHR cohorts. Commun Med (Lond) 2024;4(1):130 [FREE Full text] [doi: 10.1038/s43856-024-00549-0] [Medline: 38992068]
19. Zang C, Zhang Y, Xu J, Bian J, Morozyuk D, Schenck EJ, et al. Data-driven analysis to understand Long COVID using electronic health records from the RECOVER initiative. Nat Commun 2023;14(1):1948 [FREE Full text] [doi: 10.1038/s41467-023-37653-z] [Medline: 37029117]
20. Andronescu LR, Richard SA, Scher AI, Lindholm DA, Mende K, Ganesan A, et al. SARS-CoV-2 infection is associated with self-reported post-acute neuropsychological symptoms within six months of follow-up. PLoS One 2024;19(4):e0297481 [FREE Full text] [doi: 10.1371/journal.pone.0297481] [Medline: 38626117]
21. Krishna VD, Chang A, Korthas H, Var SR, Seelig DM, Low WC, et al. Impact of age and sex on neuroinflammation following SARS-CoV-2 infection in a murine model. Front Microbiol 2024;15:1404312 [FREE Full text] [doi: 10.3389/fmicb.2024.1404312] [Medline: 39077737]
22. Thakur KT, Chu VT, Hughes C, Kim CY, Fleck-Derderian S, Barrett CE, et al. Risk factors for new neurologic diagnoses in hospitalized patients with COVID-19: a case-control study in New York City. Neurol Clin Pract 2022;12(4):e66-e74 [FREE Full text] [doi: 10.1212/CPJ.0000000000200006] [Medline: 36382116]
23. Admon AJ, Wander PL, Iwashyna TJ, Ioannou GN, Boyko EJ, Hynes DM, et al. Consensus elements for observational research on COVID-19-related long-term outcomes. Medicine (Baltimore) 2022;101(46):e31248 [FREE Full text] [doi: 10.1097/MD.0000000000031248] [Medline: 36401423]
24. Oelsner EC, Sun Y, Balte PP, Allen NB, Andrews H, Carson A, et al. Epidemiologic features of recovery from SARS-CoV-2 infection. JAMA Netw Open 2024;7(6):e2417440 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.17440] [Medline: 38884994]
25. Bota SE, McArthur E, Naylor KL, Blake PG, Yau K, Hladunewich MA, et al. Long-term morbidity and mortality of coronavirus disease 2019 in patients receiving maintenance dialysis: a multicenter population-based cohort study. Kidney360 2024;5(8):1116-1125 [FREE Full text] [doi: 10.34067/KID.0000000000000490] [Medline: 39151048]
26. Ferraz T, Benton SJ, Zareef I, Aribaloye O, Bloise E, Connor KL. Impact of co-occurrence of obesity and SARS-CoV-2 infection during pregnancy on placental pathologies and adverse birth outcomes: a systematic review and narrative synthesis. Pathogens 2023;12(4):524 [FREE Full text] [doi: 10.3390/pathogens12040524] [Medline: 37111410]
27. Choi YJ, Kim HJ, Park J, Lee M, Kim S, Koyanagi A, et al. Acute and post-acute respiratory complications of SARS-CoV-2 infection: population-based cohort study in South Korea and Japan. Nat Commun 2024;15(1):4499 [FREE Full text] [doi: 10.1038/s41467-024-48825-w] [Medline: 38802352]
28. Wrigley Kelly NE, Kenny G, Cassidy FC, Garcia-Leon AA, De Barra C, Mallon PWG, et al. Individuals with obesity who survive SARS-CoV-2 infection have preserved antigen-specific T cell frequencies. Obesity (Silver Spring) 2022;30(10):1927-1931 [FREE Full text] [doi: 10.1002/oby.23526] [Medline: 35766325]
29. Gagliotti C, Banchelli F, De Paoli A, Buttazzi R, Narne E, Ricchizzi E, et al. The incidence and risk factors of selected drug prescriptions and outpatient care after SARS-CoV-2 infection in low-risk subjects: a multicenter population-based cohort study. Front Public Health 2023;11:1241401 [FREE Full text] [doi: 10.3389/fpubh.2023.1241401] [Medline: 37860802]
30. Fleiss JL, Levin B, Paik MC. Statistical Methods for Rates and Proportions. Hoboken, NJ: Wiley; 2003.
31. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap Consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
32. Vilagut G, Valderas JM, Ferrer M, Garin O, López-García E, Alonso J. [Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components]. Med Clin (Barc) 2008;130(19):726-735. [doi: 10.1157/13121076] [Medline: 18570798]
33. Roman‐Viñas B, Serra‐Majem L, Hagströmer M, Ribas‐Barba L, Sjöström M, Segura‐Cardona R. International Physical Activity Questionnaire: reliability and validity in a Spanish population. Eur J Sport Sci 2010;10(5):297-304. [doi: 10.1080/17461390903426667]
34. Martínez-González MA, García-Arellano A, Toledo E, Salas-Salvadó J, Buil-Cosiales P, Corella D, PREDIMED Study Investigators. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. PLoS One 2012;7(8):e43134 [FREE Full text] [doi: 10.1371/journal.pone.0043134] [Medline: 22905215]
35. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]

XSL•FO
RenderX

36. Royuela A, Macías JA. Propiedades clinimétricas de la versión castellana del cuestionario de Pittsburgh. Vigilia-Sueño 1997;9(2):81-94 [FREE Full text]
37. Muñoz-Navarro R, Cano-Vindel A, Medrano LA, Schmitz F, Ruiz-Rodríguez P, Abellán-Maeso C, et al. Utility of the PHQ-9 to identify major depressive disorder in adult patients in Spanish primary care centres. BMC Psychiatry 2017;17(1):291 [FREE Full text] [doi: 10.1186/s12888-017-1450-8] [Medline: 28793892]
38. García-Campayo J, Zamorano E, Ruiz MA, Pardo A, Pérez-Páramo M, López-Gómez V, et al. Cultural adaptation into Spanish of the Generalized Anxiety Disorder-7 (GAD-7) scale as a screening tool. Health Qual Life Outcomes 2010;8:8 [FREE Full text] [doi: 10.1186/1477-7525-8-8] [Medline: 20089179]
39. Klok FA, Boon GJAM, Barco S, Endres M, Geelhoed JJM, Knauss S, et al. The Post-COVID-19 Functional Status Scale: a tool to measure functional status over time after COVID-19. Eur Respir J 2020;56(1):2001494 [FREE Full text] [doi: 10.1183/13993003.01494-2020] [Medline: 32398306]
40. Sacristán-Galisteo C, Del Corral T, Ríos-León M, Martín-Casas P, Plaza-Manzano G, López-de-Uralde-Villanueva I. Construct validity of the Spanish version of the Post-COVID-19 Functional Status Scale and validation of the web-based form in COVID-19 survivors. PLoS One 2022;17(6):e0269274 [FREE Full text] [doi: 10.1371/journal.pone.0269274] [Medline: 35648770]
41. Golden CJ. Stroop Color and Word Test: A Manual for Clinical and Experimental Uses. Chicago, IL: Stoelting Co; 1978.
42. Smith A. Symbol Digit Modalities Test (SDMT) Manual. Los Angeles, CA: Western Psychological Services; 1968.
43. Rey A. L'examen psychologique dans les cas d'encephopathie traumatique. Arch Psychol 1941;28:286-340 [FREE Full text]
44. Sullivan MJL, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychol Assess 1995;7(4):524-532. [doi: 10.1037/1040-3590.7.4.524]
45. García Campayo J, Rodero B, Alda M, Sobradiel N, Montero J, Moreno S. Validation of the Spanish version of the Pain Catastrophizing Scale in fibromyalgia. Med Clin (Barc) 2008;131(13):487-492. [doi: 10.1157/13127277] [Medline: 19007576]
46. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The Fatigue Severity Scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46(10):1121-1123. [doi: 10.1001/archneur.1989.00520460115022] [Medline: 2803071]
47. Naik H, Shao S, Tran KC, Wong AW, Russell JA, Khor E, et al. Evaluating fatigue in patients recovering from COVID-19: validation of the Fatigue Severity Scale and single item screening questions. Health Qual Life Outcomes 2022;20(1):170 [FREE Full text] [doi: 10.1186/s12955-022-02082-x] [Medline: 36575437]
48. Keeshan A, Galipeau Y, Heiskanen A, Collins E, McCluskie PS, Arnold C, et al. Results of the Stop the Spread Ottawa (SSO) cohort study: a Canadian urban-based prospective evaluation of antibody responses and neutralisation efficiency to SARS-CoV-2 infection and vaccination. BMJ Open 2023;13(10):e077714 [FREE Full text] [doi: 10.1136/bmjopen-2023-077714] [Medline: 37907304]
49. Collins E, Galipeau Y, Arnold C, Bosveld C, Heiskanen A, Keeshan A, et al. Cohort profile: Stop the Spread Ottawa (SSO) a community-based prospective cohort study on antibody responses, antibody neutralisation efficiency and cellular immunity to SARS-CoV-2 infection and vaccination. BMJ Open 2022;12(9):e062187 [FREE Full text] [doi: 10.1136/bmjopen-2022-062187] [Medline: 36691221]
50. Klein C, Borsche M, Balck A, Föh B, Rahmöller J, Peters E, et al. One-year surveillance of SARS-CoV-2 transmission of the ELISA cohort: a model for population-based monitoring of infection risk. Sci Adv 2022;8(15):eabm5016 [FREE Full text] [doi: 10.1126/sciadv.abm5016] [Medline: 35427158]
51. Saito S, Shahbaz S, Luo X, Osman M, Redmond D, Cohen Tervaert JW, et al. Metabolomic and immune alterations in Long COVID patients with chronic fatigue syndrome. Front Immunol 2024;15:1341843 [FREE Full text] [doi: 10.3389/fimmu.2024.1341843] [Medline: 38304426]
52. Futterman ID, McLaren R, Friedmann H, Musleh N, Haberman S. Use of machine learning to identify clinical variables in pregnant and non-pregnant women with SARS-CoV-2 infection. Methods Inf Med 2022;61(3-04):61-67. [doi: 10.1055/s-0042-1756282] [Medline: 36096142]
53. Durstenfeld MS, Peluso MJ, Peyser ND, Lin F, Knight SJ, Djibo A, et al. Factors associated with Long COVID symptoms in an online cohort study. Open Forum Infect Dis 2023;10(2):ofad047 [FREE Full text] [doi: 10.1093/ofid/ofad047] [Medline: 36846611]
54. Selvakumar J, Havdal LB, Drevvatne M, Brodwall EM, Lund Berven L, Stiansen-Sonerud T, et al. Prevalence and characteristics associated with post-COVID-19 condition among nonhospitalized adolescents and young adults. JAMA Netw Open 2023;6(3):e235763 [FREE Full text] [doi: 10.1001/jamanetworkopen.2023.5763] [Medline: 36995712]
55. Osada SS, Szeghy RE, Stute NL, Province VM, Augenreich MA, Putnam A, et al. Monthly transthoracic echocardiography in young adults for 6 months following SARS-CoV-2 infection. Physiol Rep 2023;11(1):e15560 [FREE Full text] [doi: 10.14814/phy2.15560] [Medline: 36597212]
56. Peghin M, De Martino M, Palese A, Chiappinotto S, Fonda F, Gerussi V, et al. Post-COVID-19 syndrome 2 years after the first wave: the role of humoral response, vaccination and reinfection. Open Forum Infect Dis 2023;10(7):ofad364 [FREE Full text] [doi: 10.1093/ofid/ofad364] [Medline: 37520419]

XSL•FO
RenderX

57. Mercadé-Besora N, Li X, Kolde R, Trinh NT, Sanchez-Santos MT, Man WY, et al. The role of COVID-19 vaccines in preventing post-COVID-19 thromboembolic and cardiovascular complications. Heart 2024;110(9):635-643 [FREE Full text] [doi: 10.1136/heartjnl-2023-323483] [Medline: 38471729]
58. World Health Organization. A Clinical Case Definition of Post COVID-19 Condition by a Delphi Consensus. Geneva, Switzerland: World Health Organization; 2021.

Abbreviations

CRP: C-reactive protein
DSM: Diagnostic and Statistical Manual of Mental Disorders
EBV: Epstein-Barr virus
FSS: Fatigue Severity Scale
GDPR: General Data Protection Regulation
HSV-1: herpes simplex virus type 1
ICC: intraclass correlation coefficient
ICU: intensive care unit
IL-6: interleukin-6
IPAQ: International Physical Activity Questionnaire
ISCIII: Instituto de Salud Carlos III
ISRCTN: International Standard Randomized Controlled Trial Number
MEDAS: 14-item Mediterranean Diet Adherence Screener
pair_id: pair identifier
PASC: postacute sequelae of SARS-CoV-2 infection
PCFS: Post–COVID-19 Functional Status Scale
PCS: Pain Catastrophizing Scale
PHQ-9: Patient Health Questionnaire-9
PREDIMED: Prevención con Dieta Mediterránea
PSQI: Pittsburgh Sleep Quality Index
REDCap: Research Electronic Data Capture
SF-12: 12-item Short Form Health Survey
SF-36: 36-item Short Form Health Survey
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
T0: baseline assessment time point
T1: follow-up assessment time point
WHO: World Health Organization

Edited by A Schwartz; submitted 18.Dec.2025; this is a non–peer-reviewed article;accepted 20.Feb.2026; published 24.Mar.2026.

Please cite as:
Lerma-Irureta D, Méndez-López F, Navarro-Matías E, Lerma-Puertas D, Magallón-Botaya R
Differential Factors Associated With the Presence of Persistent Symptoms in Individuals Diagnosed With Long COVID: Protocol for a Longitudinal Matched Case-Control Study
JMIR Res Protoc 2026;15:e67133
URL: https://www.researchprotocols.org/2026/1/e67133
doi:10.2196/67133
PMID:

©David Lerma-Irureta, Fátima Méndez-López, Elena Navarro-Matías, Diego Lerma-Puertas, Rosa Magallón-Botaya. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Development of a Line of Care for the Health of People Who Engage in Chemsex: Protocol for a Multimethod Study

Lariane Angel Cepas[1]; Isadora Silva de Carvalho[1]; Alvaro Francisco Lopes de Sousa[2,3]; Caíque Jordan Nunes Ribeiro[4]; Shirley Verônica Melo Almeida Lima[4]; Anderson Reis de Sousa[5]; Carlos Arterio Sorgi[6]; Ricardo Nakamura[7]; Sumaé Embaló[8]; Ivone Aparecida de Paula[9]; Naila Janilde Seabra Santos[9]; Rodrigo Pierre de Freitas[10]; João Geraldo da Silva Netto[10]; Talita Morais Fernandes[11]; Ana Paula Morais Fernandes[1,3]

[1]Department of General and Specialized Nursing, Ribeirão Preto School of Nursing, University of São Paulo, Ribeirão Preto, SP, Brazil

[2]Campus de Três Lagoas, Universidade Federal de Mato Grosso do Sul, Três Lagoas, Mato Grosso do Sul, Brazil

[3]Public Health Research Centre, Comprehensive Health Research Center (CHRC), REAL, NOVA University Lisbon, Lisbon, Portugal

[4]Department of Nursing, Federal University of Sergipe, Aracaju, Sergipe, Brazil

[5]School of Nursing, Federal University of Bahia, Salvador, Bahia, Brazil

[6]Department of Chemistry, Ribeirão Preto Faculty of Philosophy, Science, and Letters, University of São Paulo, Ribeirão Preto, SP, Brazil

[7]Department of Computer Engineering and Digital Systems, Polytechnic School, University of São Paulo, São Paulo, SP, Brazil

[8]Institute of Mathematics, Statistics and Computer Science (IME), University of São Paulo, São Paulo, SP, Brazil

[9]São Paulo State Health Department, Reference and Training Center for STDs/AIDS, São Paulo, SP, Brazil

[10]Multiverso Institute, São Paulo, Brazil

[11]Higher Institute of Lisbon and Tagus Valley, Lisbon, Portugal

Corresponding Author:
Alvaro Francisco Lopes de Sousa
Campus de Três Lagoas
Universidade Federal de Mato Grosso do Sul
Av. Ranulpho Marques Leal, 3484
Três Lagoas, Mato Grosso do Sul, 79613-000
Brazil
Phone: 55 16997966962
Email: sousa.alvaromd@gmail.com

Abstract

Background: Chemsex, defined as the intentional use of psychoactive substances to enhance sexual experiences, is associated with increased risk of sexually transmitted infections (STIs), mental health harms, and disruptions in continuity of care, particularly among men who have sex with men. In Brazil, health services lack an integrated, stigma-sensitive care pathway (CP) for prevention, clinical management, and harm reduction. Embedding a CP within the Health Care Network (HCN) is essential to organize access, continuity, and quality of care.

Objective: This study aims to develop an evidence-informed CP for people who engage in chemsex, grounded in the needs of users, managers, and health professionals within the HCN, and supported by educational tools and implementation strategies.

Methods: This sequential multimethod study is guided by the Knowledge-to-Action framework. Phase 1 (predevelopment) focuses on establishing partnerships and mapping the local context. Phase 2 (Knowledge Creation) includes (1) cross-sectional surveys with adults from the general population (target n≈1600) to estimate chemsex prevalence and associated factors, and with managers (n≈54) and health professionals (n≈135) to assess knowledge, attitudes, practices, barriers, and stigma; (2) a scoping review following established methodological guidelines; and (3) triangulation and concept mapping to integrate quantitative and qualitative findings. Data will be collected via REDCap (Research Electronic Data Capture), online and in person, during testing events. Quantitative analyses will involve descriptive statistics and regression models. Interviews and focus groups will undergo thematic analysis supported by Nvivo software. Educational products will be developed and evaluated for appearance, content, and usability by experts using a content validity index ≥0.78 as a cutoff. Phase 3 (Action Cycle) includes adapting knowledge to the local context, co-designing and validating the CP with stakeholders, piloting implementation in priority municipalities, and establishing monitoring processes. Process indicators (eg, number of trained professionals, educational activities, rapid tests

XSL•FO
RenderX

performed, and app engagement) and outcome indicators (eg, STI testing and linkage to care, service use across the HCN) will be monitored through official information systems.

Results: Ethical approval was obtained in April 2025. Recruitment and data collection began in June 2025 through online and in-person strategies. By December 2025, 3061 individuals had been screened online, and 1723 participants had undergone rapid testing for STIs. Data collection is expected to conclude by June 2026, followed by data cleaning and preliminary analyses between April and June 2026, inferential and qualitative analyses between July and September 2026, and CP development and validation between October and December 2026.

Conclusions: This protocol will generate a CP tailored to chemsex and validated educational resources to support harm reduction, mental health, and STI prevention within the HCN. Findings are expected to inform inclusive policies, reduce stigma, and strengthen care coordination within Brazil's Unified Health System.

International Registered Report Identifier (IRRID): DERR1-10.2196/84068

(JMIR Res Protoc 2026;15:e84068) doi:10.2196/84068

KEYWORDS

Brazil; chemsex; harm reduction; Health Care Network; HIV prevention; Knowledge-to-Action framework; line of care; men who have sex with men; public health; sexualized drug use

Introduction

Overview

The term chemsex refers to the intentional use of psychoactive substances with the aim of intensifying and/or prolonging sexual experiences [1]. Although documented in diverse populations, the phenomenon is particularly frequent among men who have sex with men (MSM) [2], with elevated and heterogeneous prevalence rates reported in Brazil [3]. Review studies [4] indicate that, among MSM samples, chemsex prevalence ranges from 27% to 69.9%, varying according to substance profiles and recruitment contexts [5]. These data position chemsex as a public health challenge and justify the development of contextually informed and socially grounded care responses.

At the international level, reviews [6,7] underscore that the specific mix of drugs involved in chemsex is not static but depends on local scenes and markets. Nonetheless, crystal methamphetamine, gamma-hydroxybutyrate (GHB), gamma-butyrolactone (GBL), and mephedrone constitute the core most frequently described in European studies [7]. These patterns are associated with high-risk sexual behaviors [8] (eg, condomless sex with multiple partners and prolonged encounters), a higher burden of sexually transmitted infections (STIs) [7,8], and cooccurrence with events such as the mpox outbreak [9], as well as mental health harms, anxiety, depression, substance-induced psychotic symptoms, and social vulnerabilities [10,11].

In Brazil, understanding chemsex requires attention to specific market dynamics and local cultures of use. In addition to the "classic" trio described in European settings, cocaine powder, crack, ketamine, 3-, 4-methylenedioxymethamphetamine, cannabinoids, poppers (alkyl nitrites), and the adjuvant use of pharmaceuticals such as sildenafil/tadalafil, often combined with alcohol and other substances in party circuits and app-mediated encounters, are relevant [12,13]. Brazilian studies point to widespread availability and use of cocaine/crack and polysubstance use patterns that shape the sexualization of substance use among urban segments, including MSM, but not limited to them [5,14,15]. This drug ecology, which is more diverse than that depicted in the "classic" chemsex literature, reinforces the need for care pathways (CPs) that integrate harm reduction, mental health, and clinical emergency management (eg, GHB/GBL intoxication and stimulant-associated cardiovascular events).

It is also important to recognize that the sexualization of substance use is not limited to MSM. Studies have estimated the pooled prevalence of sexualized substance use (SSU) and chemsex in the general population at around 19.9%, with variations by sex and context, suggesting that sexualized use practices occur across different groups and settings [16]. Although Brazil lacks robust population-based estimates for SSU, the international evidence broadens the scope of surveillance and planning of interventions, including for populations not classified as "key" yet potentially exposed to biomedical and psychosocial risks.

From a service perspective, the research problem arises from the convergence of three gaps: (1) informational—fragmented data on prevalence, polysubstance use patterns, and therapeutic itineraries; (2) care-related—the absence of standardized screening and care flows for early detection, clinical/psychosocial management, and continuous linkage; and (3) organizational—limited integration across points of care (primary care, specialized services, mental health, emergency care, and harm reduction), with consequences for continuity and coordination of care [17,18]. These gaps are exacerbated by stigma and perceived confidentiality barriers, which inhibit disclosure of use and hinder engagement [5,7,10]. In parallel, the intervention literature suggests that multicomponent strategies, such as structured screening, brief counseling, peer educators, digital technologies, and care navigation, are promising for reducing harms and improving outcomes [19].

Brazil has a normative framework capable of guiding integrated responses through the Unified Health System (SUS)/Health Care Network (HCN), established to overcome care fragmentation and organize thematic CPs (eg, HIV/AIDS, viral hepatitis, STIs, mental health, psychosocial care, and harm reduction), with primary health care serving as the coordinator of care [20]. Within this framework, the articulation between

XSL•FO
RenderX

surveillance and care, including through regional surveillance groups and bodies, is designed to ensure timely access, longitudinality, comprehensiveness, and effective communication among network points [21]. Despite this, specific CPs for chemsex have not yet been consolidated, creating gaps in STI prevention, management of intoxications and related harms, and the provision of mental health support and psychosocial rehabilitation [19,20].

Thus, the development of a CP for chemsex adapted to the SUS/HCN is justified by (1) high prevalence among MSM and evidence of SSU in the general population; (2) local drug compositions that amplify clinical risks (stimulants, central nervous system depressants, and drug interactions) and psychosocial vulnerabilities; (3) gaps in screening, management, and linkage; and (4) the opportunity to integrate knowledge and practices (clinical care, harm reduction, mental health, testing, and combination STI prevention) into a clear care trajectory with process and outcome indicators [5,17]. Such a CP should include identification tools (screening questions and warning signs), person-centered care packages (including self-testing and linkage), protocols for toxicological emergencies (eg, GHB/GBL), psychosocial support and relapse prevention, and care navigation to reduce losses along the care continuum.

Accordingly, this study aims to develop and pilot-validate a CP for people who engage in chemsex within the SUS/HCN, informed by evidence and sensitive to stigma, articulating surveillance, primary care, specialized services, mental health, and harm-reduction strategies, with clearly defined process and outcome indicators. The research questions guiding this study are as follows: (1) What is the prevalence of chemsex, and what factors are associated with it, in priority municipalities (with an emphasis on MSM and comparisons with other groups)? (2) What gaps, barriers/facilitators, and training needs do managers and professionals identify across HCN care points regarding prevention, testing, toxicological emergency management, mental health, and linkage? (3) Which components and implementation strategies (screening, brief counseling, peer educators, digital tools, navigation) are feasible, acceptable, and scalable in the SUS?

Objective

To develop a CP for the health of individuals who engage in chemsex, based on the analysis of health needs and the experiences of users, managers, and professionals within the HCN.

Methods

Study Design

The study uses a multimethod design with a sequential approach, structured into three main phases: (1) predevelopment, (2) Knowledge Creation, and (3) the Action Cycle, as proposed by the Knowledge-to-Action (KTA) framework [22]. This model, grounded in evidence-informed change theories from the fields of health, social sciences, education, and management, has been widely used to implement health interventions.

The study will be conducted in the state of São Paulo, focusing on both urban and regional contexts. Specifically, the research involves the general population within São Paulo, particularly individuals with an active sexual life, in order to estimate the prevalence of chemsex and identify associated factors. Additionally, the study targets health care professionals and managers working in public health services within the HCN. This includes key personnel from the epidemiological surveillance groups (ESG) across the state, who are responsible for monitoring epidemiological data and coordinating health interventions, including those related to STIs, HIV/AIDS, and other public health concerns. The diverse settings in São Paulo provide a comprehensive context for understanding chemsex practices and the corresponding health care responses at both individual and systemic levels.

Study Population

Study Population, Eligibility, and Recruitment Procedures

The study population will comprise two target groups, aligned with the specific objectives:

1. Individuals from the general population with an active sexual life, to estimate the prevalence of chemsex and identify associated factors.
2. Health managers and health care professionals working in public services linked to the HCN.

General Population

Inclusion Criteria

Participants aged ≥18 years, residing in Brazil, being sexually active in the past 12 months, and willing to answer sensitive questions regarding sexual behavior, psychoactive substance use, mental health, and access to health care services are eligible for this study. Participation will occur only after written informed consent. For online respondents (REDCap [Research Electronic Data Capture]; Vanderbilt University), consent will be obtained electronically: participants will read the informed consent form (ICF) and indicate agreement by marking the mandatory checkbox before accessing the questionnaire.

For participants recruited in person, the ICF will be presented and signed physically by both the participant and the responsible researcher. In both modalities, consent will be formally documented, and a copy will be provided, digital for online respondents and printed for in-person participants.

Exclusion Criteria

Individuals with significant cognitive or emotional impairments that prevent comprehension or completion of the instrument, and those who refuse to participate or decline to sign the ICF, are excluded from participation. Participation is not contingent upon engaging in chemsex; the sample must include individuals who do and do not practice chemsex to enable unbiased prevalence estimates and analysis of associated factors.

In-Person Recruitment and Participant Procedures

To recruit participants for the survey, chain-referral sampling (snowball sampling) adapted to the virtual environment will be used, a method commonly used with hard-to-reach populations. In this approach, eligible participants receive the study link and

XSL•FO
RenderX

are invited to share it through their social networks and contacts with other individuals who meet the criteria.

In parallel, targeted dissemination will be carried out on social media platforms (Facebook and Instagram). The research team will create a fixed post on the project's web page containing a brief description of the study and an invitation to participate, including a link to the questionnaire. This post will be boosted (paid promotion) to reach all regions of the state of São Paulo and will remain active until the planned sample size is reached. Considering the higher prevalence of chemsex reported among MSM in the literature, the Hornet platform, popular within this community, will also be used. Hornet's geolocation features will allow more targeted outreach to users across different regions, enhancing reach. Another key factor is the familiarity and comfort many MSM have with this platform, commonly used for social interaction, which may increase response rates. All recruitment materials will use neutral, nonstigmatizing language, and no personally identifiable information will be collected through dissemination channels.

For in-person questionnaire administration, individuals will be invited to participate voluntarily. After acceptance and signing of the ICF, one-on-one interviews will be conducted in settings that ensure privacy. Confidentiality and anonymity will be guaranteed throughout the entire data collection process, and data will be coded to protect participant identity. After completing the questionnaire, participants will be oriented and referred for rapid testing for HIV, syphilis, and viral hepatitis. Each data collection session, including rapid testing and questionnaire administration, will last approximately 30 minutes, allowing sufficient time to ensure privacy and participant support. During this interaction, participants will also receive counseling and health education on harm reduction and STI/HIV/AIDS prevention.

Context and Systematic Recruitment Procedures (General Population)

To recruit participants for the survey, chain-referral sampling (snowball sampling) adapted to the virtual environment will be used, a method commonly used with hard-to-reach populations. In this approach, eligible participants receive the study link and are invited to share it through their social networks and contacts with other individuals who meet the criteria.

In parallel, targeted dissemination will be carried out on social media platforms (Facebook and Instagram). The research team will create a fixed post on the project's web page containing a brief description of the study and an invitation to participate, including a link to the questionnaire. This post will be boosted (paid promotion) to reach all regions of the state of São Paulo and will remain active until the planned sample size is reached. Considering the higher prevalence of chemsex reported among MSM in the literature, the Hornet platform, popular within this community, will also be used. Hornet's geolocation features will allow more targeted outreach to users across different regions, enhancing reach. Another key factor is the familiarity and comfort many MSM have with this platform, commonly used for social interaction, which may increase response rates. All recruitment materials will use neutral, nonstigmatizing language, and no personally identifiable information will be collected through dissemination channels.

For in-person questionnaire administration, individuals will be invited to participate voluntarily. After acceptance and signing of the ICF, one-on-one interviews will be conducted in settings that ensure privacy. Confidentiality and anonymity will be guaranteed throughout the entire data collection process, and data will be coded to protect participant identity. After completing the questionnaire, participants will be oriented and referred for rapid testing for HIV, syphilis, and viral hepatitis. Each data collection session, including rapid testing and questionnaire administration, will last approximately 30 minutes, allowing sufficient time to ensure privacy and participant support. During this interaction, participants will also receive counseling and health education on harm reduction and STI/HIV/AIDS prevention.

HCN Managers and Professionals

The study will include managers/coordinators and frontline health professionals (nursing, medicine, psychology, social work, pharmacy, peer educators, among others) who are formally affiliated with a public service within the HCN. Coordination will occur with support from the STD/AIDS Reference and Training Center of São Paulo state (CRT-DST/AIDS-SP) during periodic meetings with the 27 STI/AIDS focal points from the ESG of the state of São Paulo, enabling identification and agreement with municipalities interested in participating.

The study will be conducted in the city of São Paulo and in 27 municipalities selected from the ESG. It will include administration of questionnaires to the 27 regional STI/AIDS focal points of the ESG and to 1 STI/AIDS program coordinator from each selected municipality. Additionally, questionnaires will be applied to health professionals in these municipalities across mental health, STI/AIDS care, emergency and urgent care, primary care, and social assistance.

Municipal selection will be based on the highest incidence of HIV or AIDS, allowing the study to focus on areas where prevention efforts are insufficient or where incidence is rising. This approach will enable the investigation of how the knowledge of managers and professionals influences municipal services and the proposition of tailored interventions for this population. By focusing on localities where prevention and care challenges are most critical, the study will provide a robust and targeted analysis of gaps in health policies.

Recruitment will occur through institutional channels (emails and mailing lists from health departments and coordinations). Formal invitations will be sent to managers and professionals to participate voluntarily, including a link to the ICF and the questionnaire. The ICF will be presented online, and participants must accept it to proceed. Once accepted, they will be automatically directed to the questionnaire to be completed individually on the REDCap platform.

Data Collection Platform and Instrument

The questionnaires, conducted both in person and online, will be hosted on the REDCap platform (project with HTTPS

XSL•FO
RenderX

connection, audit trails, and access controls). The main survey is structured into five sections, with skip logic and input validations:

1. Eligibility, introduction, and ICF (study information and electronic consent)
2. Social and demographic characteristics
3. Sexual behaviors and substance use (lifetime and past 12 months)
4. Chemsex practices and harm reduction (frequency, contexts, substance combinations, adverse events, and care-seeking)
5. Well-being (World Health Organization-5 Well-Being Index [WHO-5]) and quality of life (EQ-5D-5L)

To ensure data integrity, REDCap will be configured to allow only one response per IP address and will include mandatory fields and logical checks to reduce missing data and inconsistencies. No direct personal identifiers will be collected; geographic variables will be aggregated to avoid reidentification.

Instrument Validation and Pretesting

Prior to fieldwork, the questionnaire underwent content and face validation conducted by 5 expert judges (topics: chemsex/SSU, STIs/HIV, harm reduction, mental health, and health service organization). Judges evaluated the relevance, clarity, and comprehensiveness of the items; their suggestions informed revisions to wording, ordering, and response options. Subsequently, a pretest was performed with 10 participants from the target audience to assess comprehension, flow, and completion time, following predefined criteria for adjustments (eg, >10% “don’t know”/missing responses on any item or recurrent comprehension problems).

Sample Size and Monitoring

For the general population survey, the primary outcome is the prevalence of chemsex in the past 12 months (yes/no). The sample size was calculated to estimate this proportion with a 95% confidence level and an absolute precision (margin of error) of 3 percentage points, adopting the conservative estimate P=.50, which maximizes sample size when true prevalence is unknown [23-25]. The standard formula for prevalence studies was applied:



where $Z_{0.95}$=1.96, P=0.50, d=0.03. Thus,



Considering the virtual recruitment method using snowball sampling and targeted campaigns, a design effect (deff) of 1.4 was adopted to accommodate potential intranetwork/region dependence, resulting in n ≈ 1068 × 1.4 ≈ 1495. To account for losses, ineligible cases, and incomplete questionnaires, an additional 10% was added, yielding n≈1645. Based on this justification and logistical constraints, a target sample of approximately 1600 participants was defined, a value that preserves the desired precision and enables stratified analyses by relevant subgroups for the protocol (eg, MSM vs others) and by macroregions [16].

The calculation was verified in G*Power (version 3.1.9.7; Heinrich Heine Universität Düsseldorf), using conservative scenarios (assumed prevalence of 50%, 95% CI, 3-percentage-point precision) and scenarios comparing proportions (MSM vs non-MSM). It is noteworthy that selecting P=.50 is a recommended practice when local prevalence is uncertain, as it maximizes sample size and safeguards the planned precision [24,25].

For the module involving managers and health care professionals within the HCN, sampling will be conducted by convenience, operationalized in partnership with the ESG of the state of São Paulo, with targets of ≥3 respondents per service (multiprofessional composition). Although this is not a probabilistic sample, the target of approximately 135 professionals from 5 key health care sectors (mental health, STI/HIV/AIDS, emergency care, primary care, and social assistance) will be included, with a focus on those working with the population who engage in chemsex, that allows descriptive estimates with reasonable precision. The sample for managers will be selected by convenience and will consist of 54 individuals, including 27 regional STI/HIV/AIDS coordinators and 27 additional managers from selected municipalities based on HIV incidence.

Study Phases

The study is structured according to the KTA framework, a comprehensive model designed to guide the translation of evidence into practice. Although the KTA cycle comprises two overarching and interlinked components (Knowledge Creation and the Action Cycle), this study operationalizes this framework into 3 sequential phases to better organize the development of the Chemsex Care Pathway within Brazil’s HCN. This framework has been widely applied in health, social sciences, and implementation research, drawing on theories of organizational change, knowledge translation, and evidence-based practice [22]. In this protocol, the model is explicitly adapted to support the development, implementation, and evaluation of a CP tailored to chemsex-related needs.

The predevelopment phase lays the groundwork for the study by identifying the problem, engaging stakeholders, and assessing contextual factors relevant to implementation. This preparatory stage ensures that the study is anchored to real-world demands and enables the effective progression of the subsequent phases. The Knowledge Creation phase encompasses primary knowledge investigation, evidence synthesis, and the generation of best-practice tools and products that will inform the CP. Finally, the Action Cycle phase focuses on implementing, monitoring, and sustaining the CP, ensuring that the knowledge generated is systematically contextualized and translated into practical applications within health care settings (Figure 1 [11]).

XSL•FO
RenderX

Figure 1. Synoptic overview of the study design adapted from the Knowledge-to-Action framework.

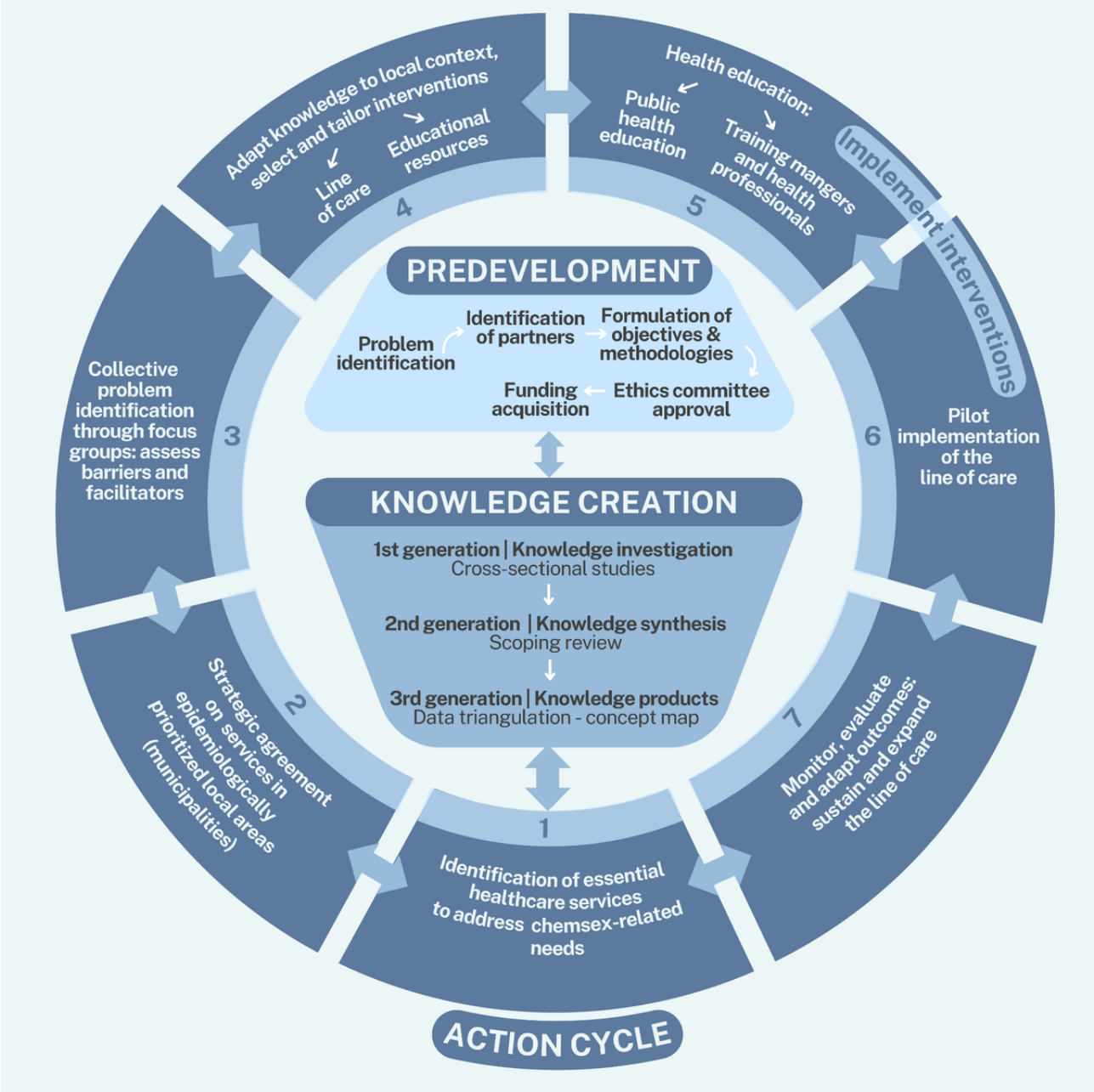

Predevelopment

The predevelopment phase establishes the foundational conditions required for successful implementation. In this study, it begins with a comprehensive process of problem identification and contextual assessment, which includes the review of scientific literature, policy documents, and epidemiological indicators (such as STI rates, mental health outcomes, and emergency care use) alongside preliminary consultations with key stakeholders. This process allows the research team to identify system gaps, stigma-related barriers, service fragmentation, and contextual determinants that may influence the feasibility and effectiveness of the intervention.

In parallel, stakeholders (professionals and chemsex users) are systematically mapped and engaged, including municipal and state health authorities, ESG, primary care managers, psychosocial care teams, emergency services, nongovernmental organizations, and community representatives. Their early involvement supports the definition of shared expectations, roles, and resource requirements, thereby strengthening feasibility and fostering joint ownership of the CP.

During this phase, the study also completes ethical preparation and methodological alignment. All study materials (including surveys, interview guides, review protocols, and data collection instruments) are finalized and submitted for ethical approval. This stage ensures coherence between the study's objectives, methodological approach, and governance structure, thereby establishing a solid basis for the subsequent phases of Knowledge Creation and the Action Cycle.

Knowledge Creation

The Knowledge Creation phase is structured into 3 progressively refined generations of knowledge production, each designed to address specific methodological requirements of the KTA

XSL•FO
RenderX

framework and to provide the level of detail necessary for replication and external evaluation. This phase follows the logic of funneling broad empirical inputs into increasingly targeted and actionable products that will inform the Chemsex Care Pathway.

The first generation (primary knowledge inquiry) consists of a set of methodologically distinct but interrelated cross-sectional investigations targeting 3 key populations: individuals who engage in chemsex, health care managers, and health professionals. For people who engage in chemsex, the survey instrument is designed to capture multiple analytical domains: (1) substance-use characterization, including granular details on substance types, administration routes, polysubstance combinations, temporal sequencing of consumption during sexual encounters, and adverse effects; (2) behavioral and contextual determinants, such as sexual partnering dynamics, motivations, environmental settings, social networks, and indicators of harm (eg, blackouts, nonintentional substance mixing, or emergency care use); (3) biopsychosocial profiling, including mental health indicators, health care use patterns, perceived barriers, and service needs; and (4) biological vulnerability markers, specifically incidence or self-reported history of HIV, STIs, viral hepatitis, syphilis, and mpox.

Surveys for managers and health professionals are structured to address three domains: (1) knowledge, including familiarity with chemsex-related substances, clinical risks, and care protocols; (2) attitudes toward people who engage in chemsex, with measurement of explicit and implicit stigma and discrimination; and (3) practices, including clinical management, referral patterns, institutional barriers, documentation procedures, and training needs. All instruments undergo content validation by experts, followed by cognitive interviews to ensure clarity and construct alignment. Data will be collected via REDCap using both online and in-person recruitment, and statistical analyses will include descriptive modeling, bivariate testing, and regression models to estimate prevalence ratios or odds ratios with 95% CIs.

The second generation (knowledge synthesis) consists of a comprehensive scoping review whose protocol has been published separately as “Healthcare Access Among Individuals Who Practice Chemsex in Brazil: A Scoping Review Protocol” [17]. In accordance with this protocol, the review follows the Joanna Briggs Institute methodology [26-28] and Arksey and O’Malley’s [29] foundational framework. The review question was formulated using the population-concept-context structure, and the search strategy was developed in collaboration with a health sciences librarian to ensure precision and reproducibility. Searches will be conducted in the electronic databases PubMed/MEDLINE, Embase, Scopus, SciELO, LILACS, and PsycINFO, as specified in the published protocol, using free-text terms related to chemsex, sexualized drug use, substance use, and health care access. All retrieved records will undergo dual independent screening at the title/abstract and full-text levels using Covidence, with disagreements resolved by a third reviewer. Data extraction will follow a structured form capturing study characteristics, population details, chemsex-related behaviors, health service use, barriers and facilitators to care, and reported outcomes. Evidence will be synthesized using a mixed methods approach, allowing the integration of quantitative and qualitative findings to comprehensively map existing knowledge. Reporting will follow the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [30], ensuring transparency in search, screening, extraction, and synthesis procedures. As detailed in the published protocol [17], this methodological structure provides a rigorous and reproducible foundation for understanding health care access among populations who engage in chemsex in Brazil.

The third generation (triangulation and concept mapping) provides the analytical integration needed to translate empirically derived and synthesized evidence into a coherent framework for decision-making. Guided by methodological triangulation principles [31,32], data from the cross-sectional studies, qualitative strands, and the scoping review will be integrated to identify convergence, complementarity, and divergence across sources. A reflexive thematic analysis will be conducted following Braun and Clarke’s [33] established 6-phase framework, as well as their more recent methodological guidance emphasizing reflexivity, analytic transparency, and researcher positionality [34]. The analytic process will include familiarization with the dataset, systematic coding, theme construction, theme review, theme definition, and final reporting. Coding will be performed iteratively by multiple researchers, and regular peer-debriefing meetings will be held to compare coding decisions, refine codebooks, and ensure analytic coherence. Analytic memos and reflexive journals will be maintained throughout to document decision-making processes, positionality considerations, and interpretive influences. To enhance credibility, confirmability, dependability, and transferability, the analysis will incorporate triangulation across stakeholder groups, an audit trail documenting analytic decisions, maintenance of a coding consistency log, and the use of thick descriptions to contextualize findings within the operational realities of the HCN. Following thematic analysis, qualitative themes will be systematically integrated with quantitative findings to identify systemic weaknesses, care discontinuities, risk amplifiers, and context-specific barriers. This integration will support the development of a conceptual map using structured concept mapping approaches [35,36], which will visually articulate the relationships among care needs, service gaps, contextual barriers, and organizational levers for change. The resulting conceptual map will function as a transitional knowledge tool bridging the Knowledge Creation and Action Cycle phases, defining both the theoretical and operational foundations of the Chemsex Care Pathway and guiding the selection of implementation strategies [31,32,35,36].

Action Cycle

Overview

The Action Cycle, as defined within KTA, provides a structured, iterative, and evidence-informed process for translating knowledge into practice and ensuring that interventions are effectively implemented, monitored, and sustained in real-world settings. Conceptually, the KTA model is grounded in well-established theories of change widely applied in health, social sciences, education, and management. These include the

XSL•FO
RenderX

Promoting Action on Research Implementation in Health Services framework, which emphasizes the interaction between evidence, context, and facilitation in successful implementation; the Diffusion of Innovations theory, which explains patterns of adoption and spread of new practices; and organizational learning and knowledge use theories, which describe how systems absorb, adapt, and routinize new knowledge [22,37,38].

The Action Cycle begins with a clear identification of the problem, the available evidence, and the contextual factors shaping the local health system. Subsequent stages follow the core steps of the KTA process: adapting knowledge to the local context; assessing barriers and facilitators to knowledge use; selecting and tailoring interventions to meet local needs; implementing evidence-informed strategies; monitoring knowledge use and fidelity; evaluating outcomes and service performance; and sustaining and scaling successful practices. In this protocol, these steps are operationalized through a series of structured activities, including territorial epidemiological assessment, stakeholder agreements, multisectoral focus groups, development and validation of the CP and educational tools, pilot implementation in prioritized municipalities, professional training, and systematic monitoring through national health information systems. Each stage contributes to ensuring that the CP is contextually relevant, institutionally feasible, and capable of improving the quality and continuity of care for individuals who engage in chemsex.

Identification of Essential Health Care Services to Address Chemsex-Related Needs

In alignment with KTA, this stage consists of a structured and operationally feasible process for selecting priority municipalities and defining which services within the HCN will participate in the pilot implementation of the Chemsex Care Pathway. Selection will begin with the identification of priority municipalities within the 27 ESG of the state of São Paulo. This process will use routinely available indicators from the Notifiable Diseases Information System, the Viral Hepatitis Information System, and official epidemiological bulletins from the State Health Secretariat. Priority will be given to municipalities with consistently higher incidence of HIV, viral hepatitis, and syphilis over the past 5 years, as well as areas reporting higher demand for STI services, emergency care related to substance intoxication, and psychosocial care needs. This approach relies solely on datasets already used in statewide planning and, therefore, is directly feasible within the project's scope. Next, a practical territorial assessment will be conducted for each candidate municipality. This includes reviewing the organization of the local HCN (primary care units, specialized STI/HIV clinics, psychosocial care centers, emergency services, and hospital referral units), availability of human resources, and existing care flows documented in municipal or regional planning instruments. These data are routinely compiled by municipal and state managers.

The combination of epidemiological burden and service availability will allow the identification of municipalities with both high need and adequate minimum structure for piloting the CP. This ensures feasibility while maintaining transparency in the selection criteria. Municipalities with clear service gaps but high epidemiological burden may also be included if the local management expresses interest and capacity to participate, reflecting the collaborative nature of implementation within SUS.

Finally, the hierarchical structure of the HCN (comprising primary care, specialized outpatient care, high-complexity care, and emergency/urgent services) will be mapped in each selected territory. This mapping will identify the specific services that will integrate the pathway, their roles within the intervention, and the coordination mechanisms between levels of care. Figure 2 illustrates the governance structure of SUS and its multilevel coordination across federal, state, regional, and municipal spheres, which underpins the organization of the pilot CP.

XSL•FO
RenderX

Figure 2. Illustration of the institutional and governance structure of Brazil's Unified Health System and the Health Care Network, highlighting the coordination across national, state, regional, and municipal levels. This hierarchical organization ensures integrated (national, regional, and local) health care delivery.

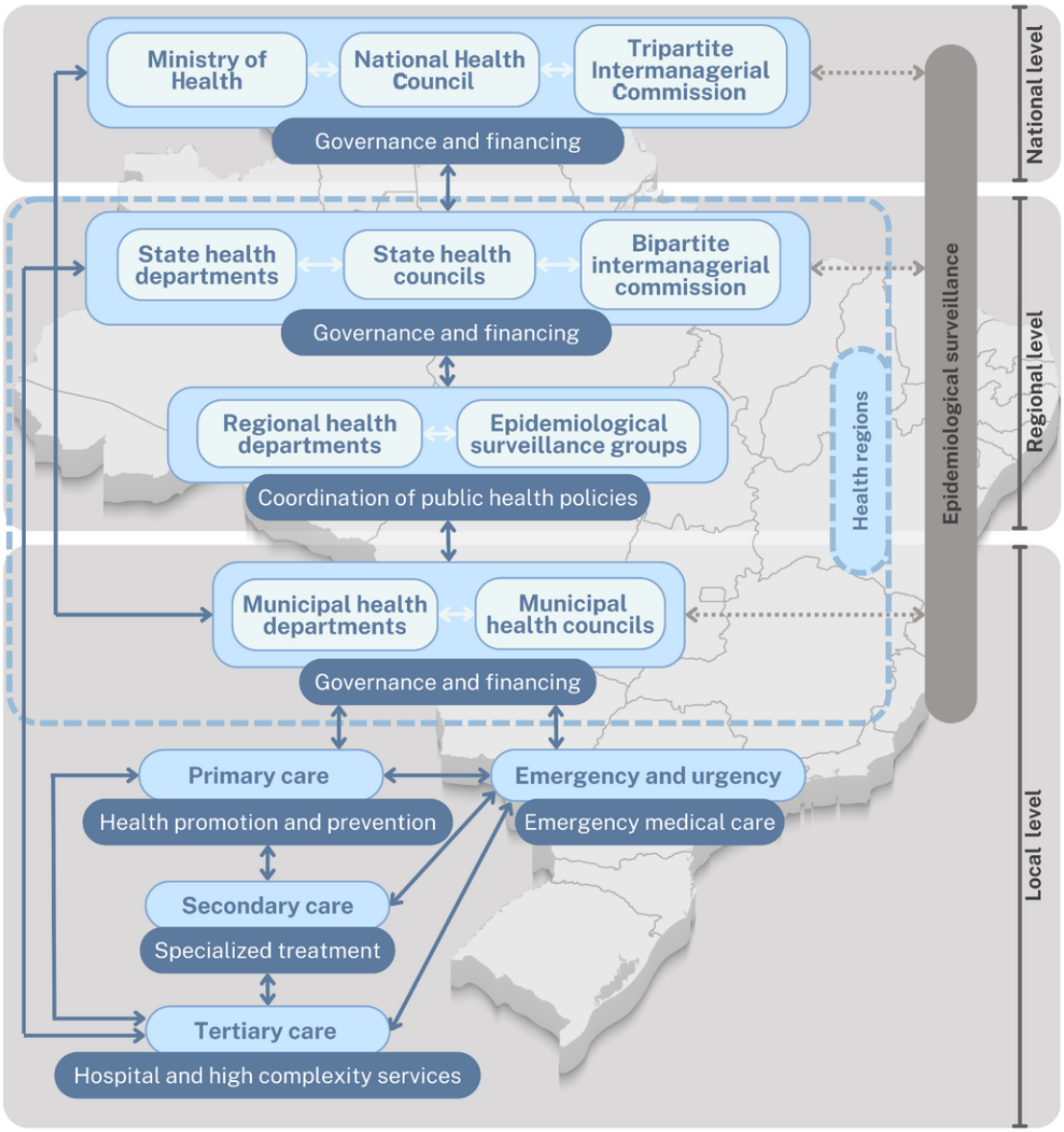

Strategic Agreement on Services in Epidemiologically Prioritized Local Areas (Municipalities)

Following the identification of priority municipalities, formal agreements will be established directly with municipal health managers and the municipal STI program coordinators, in coordination with the regional ESG of the state of São Paulo. This process corresponds to the KTA steps of adapting knowledge to the local context and identifying barriers and facilitators, and is fully aligned with existing SUS planning and governance mechanisms.

These agreements will be carried out through structured virtual or in-person meetings involving the municipal health secretariat, the STI/HIV program coordinator of each municipality, and the respective ESG technical teams. During these meetings, the scope of municipal participation will be defined considering (1) availability and current capacity of STI/HIV rapid testing, (2) organization of primary care teams, (3) presence and functioning of psychosocial care centers, (4) availability and structure of emergency care services, and (5) existing clinical and managerial protocols related to sexual health, substance use, and harm reduction. Because these components are already part of municipal responsibilities, the agreements focus on strengthening and integrating existing service points rather than creating new structures.

Jointly defined operational goals will guide implementation, such as increasing testing availability, expanding harm reduction counseling, improving referral pathways between HCN levels, and formally adopting the Chemsex Care Pathway. A routine monitoring plan will also be established, using feasible process indicators already captured by SUS information systems (eg, number of rapid tests performed, number of professionals trained, or number of referrals completed), ensuring a pragmatic and sustainable approach within municipal workflows.

XSL•FO
RenderX

Collective Problem Identification Through Focus Groups: Assess Barriers and Facilitators

This stage uses a structured qualitative approach to ensure that the CP is aligned with the operational realities of the HCN. Focus groups will be conducted with municipal health managers, STI/HIV program coordinators, primary care professionals, psychosocial care providers, emergency care staff, nongovernmental organizations, and individuals who engage in chemsex. Participants will be invited through formal communications issued via municipal health secretariats and the regional ESG, ensuring representation across the main service points implicated in chemsex-related care. Each stakeholder group will participate in 2 sequential focus groups. The first session will present the findings from the primary studies and the preliminary conceptual map, and will use a semistructured analytical matrix to identify operational barriers (such as fragmented referral pathways, insufficient training, limited harm reduction practices, stigma, and gaps in mental health integration) and local facilitators, including existing protocols, experienced teams, or ongoing prevention initiatives. Analytical techniques, including priority ranking and elements of the nominal group method, will be applied to elicit and organize participant perspectives systematically. The second session will focus on the validation of the preliminary CP logic model following the procedures proposed by McLaughlin and Jordan [39], examining the model's assumptions, internal coherence, feasibility, and alignment with municipal workflows. All sessions will be recorded, transcribed, and subjected to thematic analysis, decision-tracking matrices, and the maintenance of an audit trail. The outputs of this stage will include a refined list of priority problems, a structured assessment of contextual barriers and facilitators, and a stakeholder-validated logic model that will guide subsequent adaptation and implementation steps.

The semistructured interview and focus group topic guides are available as supplementary material (Multimedia Appendices 1 and 2).

Health Education

Overview

This stage corresponds to the KTA steps of implementing selected interventions and supporting knowledge use through targeted educational strategies directed both at the general population and at health care professionals. The health education component is designed to strengthen awareness, improve risk perception, and promote harm reduction practices among individuals who engage in chemsex, while simultaneously enhancing the technical capacity of professionals across different levels of the HCN.

Public Health Education

Public health education will be carried out using a multimodal communication strategy that relies on existing municipal infrastructures and partnerships. Educational content (developed based on the triangulated findings from the primary studies, scoping review, and focus groups) will address risks associated with psychoactive substance use during sexual activity, signs of acute intoxication, safer-use strategies, STI/HIV prevention, mental health warning signs, and recommended pathways for seeking care. Dissemination will leverage municipal health department communication channels, social media platforms, and collaboration with dating applications commonly used by communities (eg, Hornet, Grindr, Scruff, and Tinder), enabling targeted outreach without requiring new technological infrastructures. Partnerships with nongovernmental organizations and community organizations will support community-based educational interventions tailored to diverse and vulnerable populations, including migrants, trans people, sex workers, and individuals with previous experiences of substance use or mental health conditions.

Training Managers and Health Professionals

Professional training will be implemented in parallel and will follow a blended learning model that combines in-person workshops with online modules integrated into the continuing education structures already established within the SUS. Training activities will be developed collaboratively with municipal health secretariats and the ESG, ensuring alignment with regional epidemiological profiles and existing service capacities. The educational guide and mobile application developed in Phase 4 will serve as core pedagogical tools, offering standardized, evidence-based content on harm reduction, chemsex-specific clinical management, acute toxicity recognition, STI/HIV prophylaxis pathways, and mental health support strategies. Training sessions will be scheduled to accommodate municipal workflows and will prioritize frontline teams, including primary care professionals, psychosocial staff, urgent and emergency service providers, and specialized STI clinics (ensuring that services most frequently accessed by chemsex users are systematically capacitated). Evaluation of training effectiveness will incorporate pre- and posttraining assessments, feedback forms, and monitoring of knowledge retention and self-reported confidence to apply guidelines in clinical encounters. Together, these population- and professional-directed educational strategies provide the foundation for sustained knowledge use and support the successful implementation of the CP across diverse municipal contexts.

In addition, a structured 20-hour online continuing education course will be developed and made available through digital learning platforms, targeting health professionals at different levels of care (primary care, specialized services, and mental health and STI/HIV programs). This course will include asynchronous modules, recorded lectures, case studies, and practical guidelines, enabling continuous professional development, broad territorial reach, and long-term sustainability of training actions.

Pilot Implementation of the CP

Overview

The pilot implementation stage corresponds to the KTA steps of applying tailored interventions and supporting the initial use of knowledge within routine service settings. Implementation will take place in municipalities that formalize participation through agreements with local health managers and STI program coordinators, in coordination with the ESG. Instead of requiring new structures, the pilot will embed the Chemsex Care Pathways

XSL•FO
RenderX

into the existing municipal service network, including primary care units, specialized STI clinics, psychosocial staff from municipal mental health services, and urgent and emergency care teams. This approach ensures feasibility by relying on already-established HCN infrastructures and workforce capacities. This strategy will allow the CP to address the needs of both urban and semirural contexts, considering the organizational characteristics, service capacity, and epidemiological profile of each local HCN. Implementation will occur in consecutive and integrated phases.

Territorial Selection and Situational Diagnosis

Municipalities interested in implementing the CP will be selected based on epidemiological indicators, service infrastructure, institutional commitment, and willingness to adopt harm reduction approaches. A situational diagnosis will be conducted in each site to map available services, care flows, gaps in assistance, and existing intersectoral partnerships.

Governance and Institutional Alignment

A formal alliance will be established between CRT-DST/AIDS-SP, local health unit managers, municipal and state health authorities, and key stakeholders. This governance structure will define responsibilities, communication flows, referral protocols, and monitoring mechanisms. Technical committees will be created locally to ensure alignment between policy, management, and care practice.

Adaptation of the CP Model to the Local Context

Although based on a standardized care model, the CP will be adapted to the operational, cultural, and territorial realities of each municipality. This includes adjustments in referral flows, integration with existing services (primary care, STI/HIV services, mental health, harm reduction teams, and social care), and tailoring communication strategies to local populations.

Training and Capacity Building of Health Teams

Health professionals and managers involved in the CP will receive structured training, including in-person workshops and the 20-hour online continuing education course developed within the project. Training will focus on chemsex, harm reduction, trauma-informed care, sexual health, STI/HIV prevention, and integrated care approaches.

Operationalization of Care Flows

The CP will be implemented through structured CPs linking community-based entry points, primary health care, specialized services, mental health services, and social support networks. These pathways will ensure early identification, welcoming, risk assessment, clinical follow-up, harm reduction interventions, access to preexposure prophylaxis/postexposure prophylaxis, STI diagnosis and treatment, psychosocial care, and referral when needed.

Monitoring and Evaluation

Continuous monitoring will be conducted using quantitative and qualitative indicators, including service use, user flow, adherence to care, professional performance, and user satisfaction. Periodic evaluation meetings will be held with local stakeholders to assess challenges, identify best practices, and adjust the implementation process dynamically.

Sustainability and Scaling-Up Strategy

Based on the pilot results, a scaling-up plan will be developed to support expansion to other municipalities. The flexibility of the CP model will allow its replication in cities with distinct profiles, provided there is institutional commitment and minimum service capacity.

Monitor, Evaluate, and Adapt Outcomes: Sustain and Expand the CP

This stage corresponds to the KTA steps of monitoring knowledge use, evaluating outcomes, and sustaining knowledge application over time. Monitoring and evaluation activities will be carried out in municipalities participating in the pilot implementation and coordinated jointly with local health managers, STI program coordinators, and ESG. To ensure feasibility, the monitoring and evaluation plan is fully integrated into existing SUS information systems and municipal governance structures, requiring no new platforms or parallel data systems.

Each municipality will establish a CP monitoring committee composed of representatives from primary care, specialized STI services, psychosocial staff, urgent and emergency care professionals, and local management teams. This committee will review implementation progress, assess operational challenges, and guide ongoing adjustments. Monitoring will rely on routinely collected indicators derived from national health information systems, including Notifiable Diseases Information System-STI notifications (SINAN), Laboratory Test Control System (SISCEL), and Medication Logistics Control System-HIV care continuum (SICLOM), e-SUS APS (primary care records), the National Immunization Program Information System-vaccinations (SI-PNI), and Outpatient Information System of the Unified Health System (SIA/SUS) or Hospital Information System of the Unified Health System-service use (SIH/SUS). These sources provide real-time or periodic data on service performance and population outcomes without introducing additional burdens to municipal teams.

Key process indicators will include the number of rapid STI/HIV tests performed, referrals originated and completed, uptake of preexposure prophylaxis/postexposure prophylaxis, documented counseling sessions on harm reduction, and engagement with psychosocial services when indicated. Outcome indicators will focus on improved linkage to STI/HIV care, timeliness of treatment initiation, recurrent emergency visits for drug-related complications, and the accessibility of mental health care for individuals who engage in chemsex. App engagement metrics (eg, active users and most accessed content categories) will complement clinical indicators by assessing the reach and acceptability of digital educational tools.

Evaluation will follow a mixed methods approach, combining quantitative service data with qualitative feedback from municipal focal points and frontline teams. Regular virtual or in-person meetings will be held to discuss emerging barriers, identify service gaps, and refine workflows. Any required adjustments to the CP will be documented systematically to ensure traceability and to strengthen the adaptive capacity of the intervention.

XSL•FO
RenderX

Sustainability will be supported through continued collaboration with municipal managers and ESG leadership, incorporation of CP-related activities into routine municipal planning cycles, and dissemination of results through reports, scientific publications, and presentations in state-level governance forums such as Council of Municipal Health Secretariats, the State Health Secretariat technical committees, and intermunicipal planning meetings. At the end of the pilot period, consolidated findings will be presented to the State Health Secretariat to inform decisions regarding the feasibility of scaling the CP to additional municipalities. The final products, including the CP, recommendation guide, and mobile app, will be integrated into educational and academic programs to promote long-term capacity building and institutionalization. Through this iterative, data-driven process, the CP is expected to become a sustainable, evidence-informed component of comprehensive care for individuals who engage in chemsex within the HCN.

Data Analysis

The quantitative analysis will include descriptive statistics to characterize participants and estimate the prevalence of chemsex, as well as logistic regression to identify associated factors (prevalence ratio), 95% CI; $P<.05$). Analyses will be conducted using SPSS (IBM Corporation) or Stata (Stata Corp) software.

Qualitative data from interviews and focus groups will be subjected to thematic analysis [31], supported by NVivo software (Lumivero). Triangulation will integrate quantitative and qualitative findings, along with results from the scoping review, to inform the development of the CP and educational materials.

During the implementation phase, process and outcome indicators will be monitored using official systems (SINAN, Mortality Information System, SIH/SUS) to evaluate the impact of the CP in pilot municipalities.

Ethical Considerations

The study complies with resolution numbers 466/2012 and 510/2016 of the Brazilian National Health Council. Aligned with the Declaration of Helsinki, it was approved by the research ethics committee of the Ribeirão Preto College of Nursing from the University of São Paulo (85545824.2.0000.5393) and CRT-DST/AIDS-SP (85545824.2.3002.5375) in April 2025 and is set to begin in June 2025. All participants will be informed of the study's objectives and will sign the ICF.

Data Management Plan

Participants will sign the ICF, and data will be accessible only to the authorized research team, to be used exclusively for the purposes outlined in the study protocol. All information will be stored securely, with restricted access, and participants' anonymity will be guaranteed throughout all phases of the research.

Results

Participant recruitment and data collection began in June 2025 through a web-based platform and in-person strategies. By December 2025, a total of 3061 individuals had been screened online for eligibility and completion of the survey instruments.

In parallel, in-person recruitment and biological testing were conducted in health services and community-based settings. By December 2025, a total of 1723 participants had undergone rapid testing for STIs, including HIV, syphilis, and viral hepatitis.

Data collection is scheduled to be completed by June 2026. Data cleaning and preliminary descriptive analyses are scheduled for April-June 2026. Inferential analyses and qualitative data analysis are scheduled for July-September 2026. The development and validation of the Chemsex Care Pathway, including stakeholder workshops and logical model validation, are scheduled for October-December 2026. Dissemination of the main findings is expected in 2027.

Discussion

We anticipate identifying higher chemsex prevalence among MSM than in other population segments, as well as polysubstance-use patterns that combine, in addition to the "classical" core described in European contexts (GHB/GBL, methamphetamine, and mephedrone/3-, 4-methylenedioxymethamphetamine), substances widely circulating in Brazil, such as cocaine/crack, ketamine, cannabinoids, and poppers, often associated with the adjuvant use of proerectile medications [5,7]. We further hypothesize associations between chemsex and greater vulnerability to sexual risk behaviors, worse well-being scores (WHO-5) and quality of life scores (EQ-5D-5L), as well as access barriers linked to stigma and perceived confidentiality [10,19]. Among managers and professionals, we expect heterogeneous CPs and low standardization of screening and management, training gaps (eg, toxicological emergencies related to GHB/GBL, counseling, and harm reduction), and incipient integration between surveillance and care, elements that should guide the development of a feasible, stigma-sensitive CP.

These expectations are aligned with recent literature describing heterogeneous chemsex prevalence among MSM in Brazil and emphasizing that substance use mixes depend on local markets and scenes, requiring that care responses be contextually adapted [5,7]. In parallel, global estimates indicate that SSU practices are not limited to MSM, with nonnegligible prevalence in the general population, expanding the scope of surveillance and planning [16]. In this sense, the hypothesis of more plural Brazilian patterns of polysubstance use is consistent with local availability and reports of differentiated risk profiles, for example, GHB/GBL intoxications and cardiovascular events associated with stimulants, which demand specific clinical routines and referral flows [10,19].

From an organizational standpoint, the architecture of Brazil's HCN provides guidance for overcoming fragmentation and organizing thematic CPs with primary health care coordinating care [20]. The protocol's planned collaboration with the ESG in São Paulo is expected to add territorial reach and operational feasibility to the analyses, enabling identification of bottlenecks across network points (primary health care–specialized

XSL•FO
RenderX

care-mental health-emergency care), standardizing screening, and establishing protocols for acute events and for combined STI prevention. Findings from managers and professionals are expected to qualify, with service- and territory-level granularity, the "critical nodes" already suggested in reviews on chemsex interventions and harm reduction [19], strengthening the coproduction of care flows and educational materials.

The study presents methodological strengths that enhance the expected usefulness of the results. The sequential multimethod design combines a population survey, organizational diagnostics in services, and coproduction with stakeholders, allowing triangulation across prevalence/associated factors, installed capacity, and acceptability of CP components [5,19]. The target survey sample size was calculated to ensure precision in prevalence estimation and adequate power for comparisons among relevant subgroups, while the analytical plan using robust Poisson regression follows recommendations for estimating prevalence ratios in cross-sectional studies [40,41]. The projected events-per-variable adequacy supports stable, transportable models, adhering to contemporary criteria for events per variable and model planning [42-44].

Nevertheless, limitations must guide interpretation. Virtual sampling through snowball techniques and targeted campaigns may introduce selection biases (digital access and users of specific apps), requiring caution when generalizing findings. Self-reported data may introduce information bias for behaviors and substance use; data-quality procedures (mandatory fields, logic checks, and duplicate-response control) will mitigate, but not eliminate, this risk. The cross-sectional survey design does not permit causal inference between chemsex and well-being or quality-of-life outcomes; analyses will be interpreted as associations. In the managers/professionals module, convenience sampling limits population-level inferences; results should be understood as an organizational diagnostic useful for CP design and implementation.

Considering these strengths and limitations, potential implications for practice and policy include incorporating brief standardized screening at strategic HCN points; designing referral and counterreferral flows that integrate combined STI prevention, mental health, and harm reduction; establishing protocols for toxicological emergencies, particularly those related to GHB/GBL and stimulants, and implementing care-navigation strategies to reduce losses along the CP [19,20]. At a system level, findings may support state and municipal guidelines, multiprofessional training plans, and routine monitoring through process and outcome indicators, with a transversal emphasis on stigma reduction and equity.

Future directions include evaluating the CP from an implementation-science perspective, fidelity, acceptability, adoption, and cost, and measuring user outcomes such as testing coverage, linkage, and retention in care, and adverse events associated with substance use. Longitudinal studies could explore trajectories of use and mental health, as well as impacts on clinical outcomes (STIs, reinfections, and hospitalizations) and CP cost-effectiveness. Expansion to other states and nonurban contexts will test transferability and adaptation needs; digital interventions integrated into the CP, including peer education and just-in-time approaches, could be assessed in pragmatic trials.

The dissemination plan will include technical reports and policy briefs for SUS managers, educational materials and training packages for HCN teams, regional feedback sessions with ESG, academic outputs (preprints, articles, and conferences), and deposition of instruments and REDCap specifications in an open repository, as well as communication with civil society organizations in nonstigmatizing language. In summary, by combining population survey, organizational diagnostics, and coproduction, the protocol aims to inform and structure a feasible and acceptable chemsex CP within the SUS/HCN, enabling conclusions consistent with the proposed methodology, regarding the magnitude and associated factors of the phenomenon, organizational bottlenecks, and operational components of the CP, and creating a foundation for future effectiveness evaluations at scale [5,10,11,16,19,20,40-44].

Acknowledgments

The authors would like to acknowledge the institutional support provided by the National Council for Scientific and Technological Development and the Coordination for the Improvement of Higher Education Personnel, Brazil.

During the preparation of this manuscript, the authors used ChatGPT (GPT-5.1) to support linguistic refinement and the organization of sections. The authors critically reviewed, edited, and approved all content and assume full responsibility for the accuracy, integrity, and originality of the final version of this publication.

Funding

This study was supported by the National Council for Scientific and Technological Development, Brazil (CNPq grant 405741/2024-3), São Paulo Research Foundation (FAPESP grant 2024/15320-5 and 2025/04763-6); and by the Coordination for the Improvement of Higher Education Personnel, Brazil (CAPES finance code 001). The funding agencies had no role in the study design; data collection, analysis, or interpretation; manuscript preparation; or the decision to submit the manuscript for publication.

XSL•FO
RenderX

Authors' Contributions

LAC and ISdC contributed to the study conception and protocol drafting. AFLdS and APMF coordinated the overall study design, provided methodological guidance, and supervised manuscript preparation. CJNR, SVMAL, and ARdS contributed to the development of data collection instruments and qualitative study components. IAdP and NJSS assisted with institutional partnerships and stakeholder engagement strategies. CAS, RN, SE, RPdF, JGdSN, and TMF contributed to literature review, data analysis planning, and manuscript editing. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Focus group topic guide.
[DOCX File , 28 KB - resprot_v15i1e84068_app1.docx]

Multimedia Appendix 2
Semistructured interview guide (English).
[DOCX File , 16 KB - resprot_v15i1e84068_app2.docx]

Multimedia Appendix 3
Peer review report from the National Council for Scientific and Technological Development (CNPq, Brazil).
[PDF File (Adobe PDF File), 189 KB - resprot_v15i1e84068_app3.pdf]

References

1. Moreno-Gámez L, Hernández-Huerta D, Lahera G. Chemsex e psicose: uma revisão sistemática. Behav Sci (Basel) 2022;12(12):516 [FREE Full text] [doi: 10.3390/bs12120516] [Medline: 36546999]
2. Wang H, Jonas KJ, Guadamuz TE. Chemsex e uso de substâncias associado ao chemsex entre homens que fazem sexo com homens na Ásia: uma revisão sistemática e meta-análise. Drug Alcohol Depend 2023;243:109741 [FREE Full text] [doi: 10.1016/j.drugalcdep.2022.109741] [Medline: 36630807]
3. Sehnem GD, Merigo GK, Mario Saraçol DN, Souza Nunes LD, Peres Rezer JF. Chemsex e seu impacto na saúde. Open Sci Res X 2023:593-608. [doi: 10.37885/230111695]
4. Georgiadis N, Katsimpris A, Vatmanidou MA, Vassilakou T, Beloukas A, Sergentanis TN. Prevalence of chemsex and sexualized drug use among men who have sex with men: a systematic review and meta-analysis. Drug Alcohol Depend 2025;275:112800. [doi: 10.1016/j.drugalcdep.2025.112800] [Medline: 40759067]
5. Sousa AFLD, Camargo ELS, Mendes IAC. Chemsex and its repercussions on the health of men who have sex with men (MSM): a global health perspective. Rev Bras Enferm 2023;76(3):e20230004 [FREE Full text] [doi: 10.1590/0034-7167-2023-0004] [Medline: 37556700]
6. Amundsen E, Muller AE, Reierth E, Skogen V, Berg RC. Chemsex among men who have sex with men: a systematic scoping review of research methods. J Homosex 2024;71(6):1392-1418 [FREE Full text] [doi: 10.1080/00918369.2023.2170757] [Medline: 36939142]
7. Coronado-Muñoz M, García-Cabrera E, Quintero-Flórez A, Román E, Angel VA. Sexualized drug use and chemsex among men who have sex with men in Europe: a systematic review and meta-analysis. J Clin Med 2024;13(6):1812 [FREE Full text] [doi: 10.3390/jcm13061812] [Medline: 38542036]
8. Gertzen M, Strasburger M, Geiger J, Rosenberger C, Gernun S, Schwarz J, et al. Chemsex: a new challenge in addiction medicine and infectious diseases. Nervenarzt 2022;93(3):263-278. [doi: 10.1007/s00115-021-01116-x] [Medline: 33852029]
9. Santos GRDS, Ribeiro CJN, Lima SVMA, Neto JC, de Sousa AR, Bulcao CDS, et al. Chemsex among men who have sex with men during the Mpox health crisis in Brazil: a nationwide web survey. Public Health Nurs 2024;41(3):589-601. [doi: 10.1111/phn.13308] [Medline: 38528749]
10. Capodieci S, Zaccarelli M, Stingone C, Giuliani E, Cafaro V, Gianserra L, et al. Chemsex, uso sexualizado de drogas e vulnerabilidade emocional entre homens que fazem sexo com homens (HSH): um estudo de um mês em um centro de DST/HIV. BMC Public Health 2025;25(1):600 [FREE Full text] [doi: 10.1186/s12889-025-21792-7] [Medline: 39948546]
11. Coronado-Muñoz M. Uso sexualizado de drogas e chemsex entre homens que fazem sexo com homens na Europa: uma revisão sistemática e meta-análise. J Clin Med. 2024 2024;13(6):1812. [doi: 10.11606/d.5.2014.tde-20102014-143507]
12. Del Rei M, De Camargo Ferreira Adorno R. Redução de danos e chemsex em aplicativos voltados para HSH. PLURAL Rev Psicol UNESP Bauru 2024;3:e024a07 [FREE Full text] [doi: 10.59099/prpub.2024.48]
13. Souza DCD, Rodrigues IM, Araújo Neto DSD. Chemsex entre HSH: revisão integrativa. Rev Bras Sex Humana 2023;34:1136 [FREE Full text] [doi: 10.35919/rbsh.v34.1136]

XSL•FO
RenderX

14. Sousa Á, Queiroz AAFLN, Lima SVMA, Almeida PD, Oliveira LBD, Chone JS, et al. Prática de chemsex entre homens que fazem sexo com homens (HSH) durante período de isolamento social por COVID-19: pesquisa online multicêntrica. Cad Saúde Pública 2020;36(12):e00202420. [doi: 10.1590/0102-311x00202420]
15. Nikobin R, Mella-Cobas J, Veras MA, Bastos FI, Barbarà JC, Ávila VS, et al. Patterns of drug combinations used by men who have sex with men in Brazil. Braz J Psychiatry 2025;47:e20243874 [FREE Full text] [doi: 10.47626/1516-4446-2024-3874] [Medline: 39827456]
16. Ramazanova M, Turdaliyeva B, Igissenova AI, Zhakupova M, Izbassarova AS, Seifuldinova M, et al. [Prevalence of sexualized substance use and chemsex in the general population and among women: a systematic review and meta-analysis of cross-sectional studies]. Healthcare (Basel) 2025;13(8) [FREE Full text] [doi: 10.3390/healthcare13080899] [Medline: 40281847]
17. Carvalho IS, Cepas LA, Sousa Á, Fernandes TM, Luz TG, Silva JCS. Acesso à saúde entre indivíduos que praticam chemsex no Brasil: um protocolo de revisão de escopo. Nurs Rep 2025;15:353 [FREE Full text] [doi: 10.3390/nursrep15100353.Accessed]
18. Medeiros FAF, Campana PDS, Cuba GT. Avaliação do conhecimento sobre chemsex em profissionais de saúde do departamento de moléstias infecciosas e parasitárias de um complexo hospitalar. Braz J Infect Dis 2022;26(Suppl 2):102483. [doi: 10.1016/j.bjid.2022.102483]
19. Del Pozo-Herce P, Tovar-Reinoso A, Martínez-Sabater A, Chover-Sierra E, Sufrate-Sorzano T, Saus-Ortega C, et al. Exploration of mental health elements in chemsex behavior: interventions, influencing factors, and attitudes towards risk and harm reduction in Spain. Nurs Rep 2024;14(3):2313-2326 [FREE Full text] [doi: 10.3390/nursrep14030172] [Medline: 39311180]
20. Tofani LFN, Furtado LAC, Guimarães CF, Feliciano DGCF, Silva GRD, Bragagnolo LM, et al. Caos, organização e criatividade: revisão integrativa sobre as redes de atenção à Saúde. Ciênc Saúde Coletiva 2021;26(10):4769-4782. [doi: 10.1590/1413-812320212610.26102020]
21. Portaria nº 2.436, de 21 de setembro de 2017. Aprova a Política Nacional de Atenção Básica - PNAB. Brasília: Ministério da Saúde. Brasília: Ministério da Saúde; 2017. URL: https://bvsms.saude.gov.br/bvs/saudelegis/gm/2017/prt2436_22_09_2017.html [accessed 2025-11-27]
22. Graham ID, Tetroe JM. The knowledge to action framework. In: Models and Frameworks for Implementing Evidence-Based Practice: Linking Evidence to Action. Oxford: Wiley-Blackwell; 2010:222.
23. Cochran WG. Sampling Techniques. New York: John Wiley & Sons; 1977.
24. Lwanga SK, Lemeshow S. Sample Size Determination in Health Studies: A Practical Manual. Geneva: World Health Organization; 1991.
25. Nagington M. The Moral Lessons of Chemsex: A Critical Approach. Abingdon: Taylor & Francis; 2025.
26. Joanna Briggs Institute. The Joanna Briggs Institute Reviewers' Manual 2014. Adelaide: Joanna Briggs Institute; 2014.
27. Peters M, Godfrey C, McInerney P, Baldini SC, Khalil H, Parker D. Scoping reviews. In: Joanna Briggs Institute Reviewer's Manual. Adelaide: Joanna Briggs Institute; 2017.
28. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Capítulo 11: revisões do escopo (versão 2020). In: JBI Manual for Evidence Synthesis. Adelaide: JBI; 2020.
29. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
30. Tricco A, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
31. Teixeira JC, Nascimento MCR, Antonialli LM. Profile of administration studies that used methodological triangulation: an analysis of the annals of EnANPAD in the period of 2007-2011. Rev Adm (Sao Paulo) 2013;48(4):800-812 [FREE Full text]
32. Basso JA, Persch L, Kiekow A, Seben P, Gubert F, Tondolo V. Triangulação: uma ferramenta de validade e confiabilidade. Sinergia 2016;20(1):19-28.
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
34. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a researcher. Int J Transgend Health 2023;24(1):1-6 [FREE Full text] [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
35. Bardin L. Análise de Conteúdo. São Paulo: Edições; 2016.
36. Novak JD. The Theory Underlying Concept Maps and How to Construct Them. Ithaca (NY): Cornell University; 2008.
37. Kitson A, Harvey G, McCormack B. Enabling the implementation of evidence based practice: a conceptual framework. Qual Health Care 1998;7(3):149-158 [FREE Full text] [doi: 10.1136/qshc.7.3.149] [Medline: 10185141]
38. Kitson AL, Rycroft-Malone J, Harvey G, McCormack B, Seers K, Titchen A. Evaluating the successful implementation of evidence into practice using the PARiHS framework: theoretical and practical challenges. Implement Sci 2008;3:1 [FREE Full text] [doi: 10.1186/1748-5908-3-1] [Medline: 18179688]

XSL•FO
RenderX

39. McLaughlin JA, Jordan GB. Logic models: a tool for telling your programs performance story. Eval Program Plann 1999;22(1):65-72. [doi: 10.1016/S0149-7189(98)00042-1]
40. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003;3:21 [FREE Full text] [doi: 10.1186/1471-2288-3-21] [Medline: 14567763]
41. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004;159(7):702-706. [doi: 10.1093/aje/kwh090] [Medline: 15033648]
42. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49(12):1373-1379 [FREE Full text] [doi: 10.1016/s0895-4356(96)00236-3] [Medline: 8970487]
43. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol 2007;165(6):710-718. [doi: 10.1093/aje/kwk052] [Medline: 17182981]
44. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ 2020;368:m441. [doi: 10.1136/bmj.m441] [Medline: 32188600]

Abbreviations

CP: care pathway
CRT-DST/AIDS-SP: STD/AIDS Reference and Training Center of São Paulo state
ESG: epidemiological surveillance groups
GBL: gamma-butyrolactone
GHB: gamma-hydroxybutyrate
HCN: Health Care Network
ICF: informed consent form
KTA: Knowledge-to-Action
MSM: men who have sex with men
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
REDCap: Research Electronic Data Capture
SIA/SUS: Outpatient Information System of the Unified Health System
SICLOM: Medication Logistics Control System-HIV care continuum
SIH/SUS: Hospital Information System of the Unified Health System-service use
SINAN: Notifiable Diseases Information System-STI notifications
SI-PNI: National Immunization Program Information System-vaccinations
SISCEL: Laboratory Test Control System
SSU: sexualized substance use
STI: sexually transmitted infection
SUS: Unified Health System
WHO-5: World Health Organization-5 Well-Being Index

Edited by A Schwartz; submitted 15.Sep.2025; this is a non–peer-reviewed article;accepted 27.Jan.2026; published 26.Mar.2026.

Please cite as:
Cepas LA, Carvalho ISD, de Sousa AFL, Ribeiro CJN, Lima SVMA, Sousa ARD, Sorgi CA, Nakamura R, Embaló S, Paula IAD, Santos NJS, Freitas RPD, Silva Netto JGD, Fernandes TM, Fernandes APM
Development of a Line of Care for the Health of People Who Engage in Chemsex: Protocol for a Multimethod Study
JMIR Res Protoc 2026;15:e84068
URL: https://www.researchprotocols.org/2026/1/e84068
doi:10.2196/84068
PMID:

©Lariane Angel Cepas, Isadora Silva de Carvalho, Alvaro Francisco Lopes de Sousa, Caíque Jordan Nunes Ribeiro, Shirley Verônica Melo Almeida Lima, Anderson Reis de Sousa, Carlos Arterio Sorgi, Ricardo Nakamura, Sumaé Embaló, Ivone Aparecida de Paula, Naila Janilde Seabra Santos, Rodrigo Pierre de Freitas, João Geraldo da Silva Netto, Talita Morais Fernandes, Ana Paula Morais Fernandes. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Maternal Perception vs Actual Breast Milk Supply: Protocol for an Observational Cross-Sectional Study

Bailey Bruckner[1], MSc; Rachael Taylor[2], PhD; Carmen Parata[3], MA; Jillian Haszard[4], PhD; Barry Taylor[5], MBChB; Samantha Bevin[1], MAppSc; Kassidy Gooding[1], BSc; Anne-Louise Heath[1], PhD; Ioanna Katiforis[6], PhD; Lisa Daniels[1], PhD

[1]Department of Human Nutrition, University of Otago, Dunedin, New Zealand

[2]Department of Medicine, University of Otago, Dunedin, New Zealand

[3]Department of Public Health, University of Otago, Wellington, New Zealand

[4]Haszard Biostatistics, Otago, New Zealand

[5]Department of Paediatrics and Child Health, University of Otago, Dunedin, New Zealand

[6]School of Exercise and Nutrition Sciences, Deakin University, Burwood, Australia

Corresponding Author:
Lisa Daniels, PhD
Department of Human Nutrition
University of Otago
70 Union Place
Dunedin, 9016
New Zealand
Phone: 64 3 556 5071
Email: lisa.daniels@otago.ac.nz

Abstract

Background: It is well known that breastfeeding provides favorable health outcomes for both mother and baby. However, many mothers struggle to meet global recommendations to exclusively breastfeed for 6 months. Of those who cease breastfeeding early, one third attribute this to perceived insufficient milk supply. Currently, it is uncertain how the perception of insufficient milk supply relates to physiological milk volume or nutrient composition.

Objective: The Māmā and Baby Breastfeeding Study aims to estimate human milk volumes produced by a diverse sample of breastfeeding mothers at 3 months post partum with differing perceptions of milk supply and to investigate human milk composition in relation to milk volume.

Methods: We plan to recruit a sample of 150 mother–infant dyads in this observational study in Dunedin, Aotearoa New Zealand. Participants will represent a diverse range of ethnicities and socioeconomic backgrounds. Human milk volume will be assessed using the dose-to-mother stable isotope (deuterium oxide) technique. Mother participants will consume an accurate dose (30 g) of deuterium oxide after baseline saliva samples are collected from both mother and infant. Subsequent postdose samples will be collected over 3 time points to determine deuterium enrichment over a 14-day period using Fourier transform infrared spectrometry. Human milk macronutrient (energy, fat, carbohydrate, crude protein, and true protein) and mineral and trace element (sodium, magnesium, phosphorus, potassium, calcium, iron, copper, zinc, selenium, and iodine) composition of 1 full milk expression from 1 breast will be analyzed using the MIRIS Human Milk Analyzer (Miris AB) and inductively coupled plasma mass spectrometry, respectively. Potential predictors and maternal perception of milk supply will be assessed via questionnaire. Infant BMI will be calculated from measures of weight and length at 3 different time points over 4 weeks, using standard techniques. These, alongside anthropometric measurements collected at the infant's Well Child Tamariki Ora visits, will be used to assess infant growth trajectory in the first 6 months of life. Regression models will be used to assess the associations between maternal perception of milk supply, human milk volumes, and composition.

Results: This study was funded in April 2023 by the Health Research Council of New Zealand (grant 23/461). Recruitment for this study began in February 2025 and is anticipated to conclude in June 2026, with analysis expected to be completed by February 2027. As of January 17, 2026, a total of 91 participants have been enrolled. Final results are anticipated to be disseminated in late 2027 following completion of data analysis.

XSL•FO
RenderX

Conclusions: This research will provide new knowledge on whether maternal perception of milk supply aligns with actual human milk volume or nutrient composition. Such information will be extremely useful for health professionals working with breastfeeding mothers with milk supply concerns and for informing the design of breastfeeding support programs and resources.

Trial Registration: Australian and New Zealand Clinical Trial Registry ACTRN12625000180415; https://www.anzctr.org.au/ACTRN12625000180415.aspx

International Registered Report Identifier (IRRID): DERR1-10.2196/84776

(JMIR Res Protoc 2026;15:e84776) doi:10.2196/84776

KEYWORDS

breastfeeding; human milk composition; human milk; infants; lactation; milk supply; perceived insufficient milk supply; PIMS; stable isotope

Introduction

Overview

While national and international advisory groups recommend exclusive breastfeeding for the first 6 months of life and continued breastfeeding for children aged 2 years or older [1,2], many mothers in New Zealand and globally are not meeting these recommendations. In New Zealand, 77% of infants are exclusively breastfed at discharge from maternity services [3]; however, this drops to under 60% by 4 months of age and less than 10% by 6 months of age [4]. These percentages are lower than exclusive breastfeeding rates at 6 months in the United States (24.9% in 2019) [5] but higher than recent estimates in Australia (less than 1% in 2021) [6]. It is evident that work is needed to improve exclusive breastfeeding rates for the first 6 months of life in New Zealand and other high-income countries, especially considering the global goal of 60% of infants being exclusively breastfed to 6 months of age by 2030 set by the World Health Assembly [7].

There are many reasons why mothers may not meet recommended breastfeeding targets, such as pain, tiredness, inconvenience, or return to work or education [8,9]. However, perceived insufficient milk supply (PIMS)—a mother's perception that her production of milk is not adequate to satisfy her infant's needs—is reported to be the leading cause of early cessation of breastfeeding [10-12]. Gatti [10] concluded that, of women who discontinue breastfeeding early, 35% describe PIMS as the primary reason. It is also a known barrier to continuing breastfeeding in New Zealand [13,14]. However, while PIMS is the most commonly cited reason for early cessation of breastfeeding, it is not currently known how many experience PIMS but continue to breastfeed despite these challenges.

Factors that commonly contribute to this perception often include unsettled infant behaviors, such as frequent crying or short sleep duration, as well as low maternal satisfaction and self-confidence in breastfeeding [15]. However, the perception of insufficient milk supply can be a dangerous cycle that can lead to actual low milk supply through disruption of the supply-demand feedback loop [16]. If supplementation with other milks, such as infant formula, occurs due to the perception of low milk supply, this can cause a decrease in feeding at the breast and therefore a reduction in milk production [17,18]. Maternal feelings of personal failure, guilt, anxiety, and reduced maternal self-efficacy can occur as a consequence of perceived breastfeeding failure [9,19], potentially exacerbating the symptoms of PIMS.

A major gap in the literature exists in that, while PIMS is very common and leads to a large proportion of early breastfeeding cessation, how perception of insufficient milk supply relates to actual physiological milk volume is uncertain. Similarly, there appears to be no research investigating the perception of high milk volume (oversupply) in this context. Only 1 study appears to have investigated the relationship between PIMS and human milk volume [20], using the 24-hour test-weighing method in first-time mothers only. The study found that, at 2 weeks post partum, there was no significant association between perceived and actual insufficient milk supply. No other research has examined older infants, particularly infants aged 3 months, immediately before breastfeeding rates decrease in New Zealand [4]. This is perhaps due in part to the fact that accurately measuring human milk intake is complex. Standard approaches often assume a fixed volume of intake across infants of a specific age. For example, a commonly used "one size fits all" approach estimates individual intake at 780 mL/day for infants aged 0 to 6 months [21]. However, such estimates mask considerable variation observed between mother-infant pairs, with intakes ranging from 400 to 1400 mL/day in infants aged 2 to 5 months [22]. Such standard approaches clearly undermine the ability to assess true intake in breastfed infants.

Other approaches include 24-hour test-weighing, in which the mother weighs the infant directly before and after all breastfeeds over a 24-hour period. Test-weighing, while more individualized than the "one size fits all" approach, may impact usual feeding patterns and interfere with the physiological and behavioral aspects of lactation [23,24]. Measuring pumped (expressed) milk volume has also been used to estimate human milk production [25-27]. However, this is not considered an accurate indicator of infant intake largely due to differences between maternal production capacity and infant need [27]. Pumping output is also highly variable, influenced by factors such as device efficiency, maternal pumping technique, and individual responsiveness to pumping [28,29]. As a result, expressed volume often reflects the effectiveness of the tools and methods used rather than the mother's underlying physiological milk production.

Another common method of determining human milk intake is the "minutes per feed" approach, which assumes that a set

XSL•FO
RenderX

amount of milk is transferred to the baby per minute of breastfeeding. However, ultrasound imagery has previously found no relationship between the length of the breastfeed and the amount of milk transferred [30], suggesting considerable inaccuracy with this method. A promising approach to estimating human milk intake is the use of predictive modeling [31]; although, to date, only limited data are available and the method has yet to be validated.

Finally, the dose-to-mother (DTM) stable isotope technique [24,32,33] is considered the gold-standard method for estimating human milk volume. This method can estimate human milk intake over a 14-day period by giving the mother an accurate oral dose of deuterium oxide (D_20; a stable isotope, also known as "heavy water") and tracing the transfer of D_2O from the mother to the infant through mother and infant saliva or urine samples. This method has been used to estimate human milk intake of infants worldwide, most commonly in early infancy [23,34-38], but also in older infants aged 7 to 10 months in New Zealand [31]. While it is more expensive than other methods, it is relatively noninvasive and provides an accurate estimate of human milk intake [33]. To date, no research has used the gold-standard DTM technique to investigate the relationship between PIMS and actual milk intake in infants.

To add to the complexity of measuring human milk intake, infant intake can be influenced by a range of factors. Well-known maternal predictors of milk intake include higher education, which is associated with higher intakes [31,39], as well as current employment and high BMI, both associated with lower intakes [31,40]. Well-known infant predictors include age [23,31,39,41,42], with intake varying over the course of lactation, as well as high body weight [22,31,34,39,43], demand vs scheduled feeding [31,44], and a higher average number of breastfeeds [31,45], which are associated with higher intake in infants. However, many factors that may impact infant milk intake have not been well investigated. For example, maternal cannabis or nicotine use has been suggested to be associated with lower infant milk intake [46,47], as well as certain maternal medications [48], poor maternal mental health [49], high maternal stress [50], and infant male sex [23], among others. These potential predictors have not been thoroughly investigated, and further exploration is needed to understand their impact.

It is also not known how differences in nutritional composition influence the perception of milk supply or volume. Determining human milk adequacy in terms of both volume (quantity) and composition (quality) will provide valuable information and has been highlighted as a known gap in the literature [51]. For mothers with lower milk volume, it may be important to know whether the nutritional composition of the milk, particularly total energy, differs from that of a mother with higher milk volume. Furthermore, determining the nutrient composition of human milk in New Zealand is particularly important because of unique local nutritional issues. In particular, iodine and selenium are low in the New Zealand food supply due to poor soil content [52] and are critical for infant development, as both support thyroid hormone production and regulation, and therefore brain maturation [53,54].

The aim of the Māmā and Baby Breastfeeding Study (MABBS) is to describe human milk volumes produced by a diverse sample of lactating mothers of infants aged 3 months with differing perceptions of milk supply. We aim to compare the lactating mothers' perceptions of milk supply with the volume of milk they produce, as well as the nutrient composition of that milk.

Primary Objective

The primary objective is to estimate the human milk intakes of fully breastfed infants aged 3 months using the gold standard DTM D_2O technique.

Secondary Objectives

The secondary objectives, at 3 months post partum, are to (1) estimate how human milk volumes differ between lactating mothers with differing perceptions of milk supply, (2) explore the characteristics that may predict human milk volumes produced by a diverse range of lactating mothers, (3) explore the relationship between human milk volumes produced and infant growth as a potential means of assessing human milk adequacy, (4) investigate the association between human milk composition and human milk volumes, and (5) assess how demographic characteristics might modify the relationship between perception vs actual milk supply of lactating mothers.

While the term mother is used throughout the manuscript, we acknowledge that not all lactating individuals identify as mothers. However, due to the name of the methodology used (dose-to-mother stable isotope technique) the word mother is used in this manuscript with the inclusion of all breastfeeding, chestfeeding, and human milk–feeding individuals in mind, regardless of how they identify.

Methods

Ethical Considerations

The study has been approved by the University of Otago Human Ethics Committee (Health) Te Pae Matatika Tangata (Haurora) Ōtākou Whakaihu Waka (H24/0225), and written informed consent will be obtained at the first appointment. The study is registered with the Australian New Zealand Clinical Trials Registry (registration number ACTRN12625000180415). Upon enrollment, participants will be assigned a unique identifying code (participant ID). To preserve confidentiality, all data collected will be recorded against the participant ID. Participants will be given a NZ $100 (US $60) grocery voucher as a thank-you for their involvement.

Study Design

The MABBS is an observational, cross-sectional study of human milk intakes in infants aged 2 to <4 months in Dunedin, New Zealand. Participants will be included if they are: a lactating mother aged 16 years or older; living in Dunedin, New Zealand; have an infant aged 2 to <4 months at the time of their first visit; are fully breastfeeding (ie, their infant is receiving human milk only) for 2 weeks prior to recruitment; and intend to continue fully breastfeeding for the duration of the study (4 weeks).

We will exclude participants who cannot provide informed consent or complete study procedures in English, are tandem

XSL•FO
RenderX

breastfeeding (simultaneously breastfeeding another child), or are breastfeeding multiples (twins or higher-order multiples). To our knowledge, the DTM technique has not been specifically validated for use in twins or tandem-fed children; therefore, the study design was based on estimating human milk intake in singleton dyads.

Participants and Recruitment

A total of 150 mother–infant dyads will be recruited. We initially planned to target recruitment to mothers with infants aged 2 to 3 months, so that they would participate within the study age range of 3 to <4 months. This time point was selected because the prevalence of breastfeeding drops dramatically around 4 months of age in New Zealand [4], allowing estimation of human milk intake just before this decline to understand factors that may contribute to the decrease around 4 months. While some literature suggests that PIMS is most common early in the breastfeeding journey [8,55], research in New Zealand shows that PIMS is a primary reason for early breastfeeding cessation around 4 to 5 months of age [14].

Furthermore, around 3 to 4 months of age, many PIMS-related factors may contribute to this decline, including changes in infant feeding behaviors, such as an increased rate of cluster feeding (close intervals between breastfeeds), and maternal changes, such as softer breasts and less milk leakage as the body becomes more efficient at producing milk. This time point also occurs before the majority of mothers return to work in New Zealand [56] and before the recommended age for introducing complementary foods, both of which events which influence infant feeding.

However, when recruitment commenced, we found that many mothers had already begun mixed feeding before 3 months of age, potentially due to PIMS. As many of these women were from ethnic groups other than New Zealand European, the 1-month age band (3 to <4 months of age) appeared to limit the ethnic diversity of our sample. Thus, from August 29, 2025, we revised the minimum age of inclusion to 2 months rather than 3 months of age, resulting in a final inclusion period of infants aged between 2.0 and <4 months, and therefore the recruitment age range expanded to infants aged 1.5 to 2 months.

Recruitment will occur through advertisement on social media, word of mouth, and referral through health professionals (eg, midwives, Well Child Tamariki Ora providers, and lactation consultants) where possible, and by study posters in parent groups and community spaces as needed. We aim to recruit a diverse range of participants in terms of ethnicity and socioeconomic status, including approximately 20% Māori (indigenous people of New Zealand, to reflect local statistics of the region in which data collection is occurring) [57]. Some targeted recruitment (identifying and approaching groups to promote the study) may be required to achieve an ethnically diverse sample. This will occur through using networks within the maternity and health care sector as well as with Māori research colleagues.

Interested individuals will complete an online eligibility questionnaire administered through REDCap (Research Electronic Data Capture; Vanderbilt University) [58], which includes access to the information sheet about the study and the consent form to review. Once eligibility is confirmed, the potential participant will be sent an email offering them the opportunity to ask any questions and attaching the information sheet and consent form. A follow-up phone call or text will be set around 3 days later to confirm they have read the information, answer any questions about the study, obtain verbal consent, and schedule their appointments.

Data Collection

Overview

Participation in the study will involve 5 contacts over 4 weeks following recruitment, as depicted in Table 1. Data collection will be conducted in the participant's home or in the research clinic (Department of Human Nutrition, University of Otago, Dunedin, New Zealand), depending on the participant's preference. The first visit (visit 1) will involve obtaining written consent, completion of the main questionnaire (demographic information, breastfeeding history and self-efficacy, perception of milk supply, maternal and infant sleep, maternal diet and health history, and maternal stress), obtaining anthropometric measurements (infant weight and length and maternal weight and height), obtaining baseline saliva samples from both the mother and the infant, and administration of a 30-gram dose of stable isotope (D_2O) to the mother. The mother will be asked to take her own (mother only) at-home saliva sample 3 to 4 hours after the dose is administered (to be retrieved at visit 2). Postdose saliva samples will be taken by researchers at the second visit (visit 2), occurring 2 to 3 days after administration of the isotope, as well as at the third visit (visit 3), 6 to 7 days after the isotope, and at the fourth visit (visit 4), 13 to 14 days after isotope administration. Infant anthropometric measurements (weight and length) will be repeated at the fourth visit (visit 4). After the fourth visit (visit 4), the mother will be asked to express a full milk expression from a single breast. The final visit (visit 5) will be conducted by an International Board-Certified Lactation Consultant (IBCLC), who will take the final anthropometric measurements from the infant; conduct an observed breastfeed with a one-off, single test weighing (pre- and postfeed weights); conduct an oral assessment of the infant; and ask the mother to complete a questionnaire about dysphoric milk ejection reflex (D-MER) [59].

XSL•FO
RenderX

Table 1. Overview of participant contacts.

Study stage	Procedures
Recruitment: infants aged 0-2 months	• Recruitment of potential participants
Screening: infants aged 6 weeks to 2 months	• Eligibility questionnaire
Visit 1 (baseline): infants aged 2 to <4 months	• Written consent • Baseline saliva samples (mother and infant) • Isotope administration (mother only) • Anthropometry (mother and infant) • Main questionnaire • Note: the mother is asked to collect a saliva sample (mother only) 3-4 hours after the dose is administered
Visit 2: 2-3 days after baseline	• Follow-up (postdose) saliva samples (mother and infant) • Collect postdose saliva sample, that was self-collected by the mother after visit 1 (mother only)
Visit 3: 6-7 days after baseline	• Follow-up (postdose) saliva samples (mother and infant)
Visit 4: 13-14 days after baseline	• Final (postdose) saliva samples (mother and infant) • Anthropometry (infant only)
Human milk collection: approximately 3 days after visit 4	• Collection of 12 mL subsample of full milk expression from single breast
Visit 5: 2 weeks after visit 4	• Anthropometry (infant only) • Single test-weigh = weight before and after a breastfeed (infant) • Breastfeeding assessment • Oral assessment • D-MER[a] questionnaire

[a]D-MER: dysphoric milk ejection reflex.

Demographic Data

Self-reported maternal and mother-reported child ethnicity will be collected as part of the eligibility questionnaire, using standard New Zealand census questions [60]. Area-based household socioeconomic deprivation (New Zealand Index of Deprivation 2023) will be determined using the participant's postal address [61]. New Zealand Index of Deprivation 2023 combines 9 variables from the 2023 New Zealand national census to determine a socioeconomic deprivation score for each statistical area (a geographic area defined by Statistics New Zealand containing approximately 100 to 200 people). The deprivation score estimates the relative material and social deprivation for the area where the participant lives, where decile 1 represents areas with the least deprivation and decile 10 represents areas with the most deprivation. Maternal education, maternal work status, childcare use, and the number of children and adults living in the household will be collected in the main questionnaire administered at the first visit.

Maternal Perception of Milk Supply

Perception of milk supply will be assessed in the main questionnaire using 2 approaches. First, using the Perceived Infant Breastfeeding Satiety Subscale of the Hill and Humenick Lactation Scale [15] in the main questionnaire, which is a validated scale designed to determine PIMS. The Perceived Infant Breastfeeding Satiety Subscale is a 5-question subscale with responses recorded using a 7-point Likert scale (1=strongly disagree to 7=strongly agree). Higher scores represent higher perceived milk supply.

Second, a single question developed by the research team will be administered to assess maternal perception of milk supply: "How do you currently feel about the amount of breast milk you make for your baby/pēpi? (Think about over the past week.)" with answers from a 5-point Likert scale (1=not enough milk, 2=sometimes not enough milk, 3=enough milk, 4=sometimes too much milk, and 5=too much milk). This continuous variable will be dichotomized to create a variable for comparing low (insufficient) perceived milk supply and high (oversupply) perceived milk supply, combining 1=not enough milk and 2=sometimes not enough milk to define low (insufficient), and 4=sometimes too much milk and 5=too much milk to define high (oversupply), relative to 3=sufficient supply.

Participants will also be asked whether they have had any prior concerns about their milk supply and, if so, when these concerns occurred, what caused these concerns, and whether the participant did anything to address these concerns. Additionally, at visits 1 to 4, participants will be asked whether there have been any new concerns about low milk supply to determine whether their perception changes over the course of the DTM period.

XSL•FO
RenderX

DTM Stable Isotope Technique

Human Milk Volume

Human milk volume will be determined using the gold-standard stable isotope DTM technique [24,32,33]. Baseline saliva samples (approximately 2 mL) will be collected from the mother and infant participants at the first visit. Participants will be asked to ensure that neither they nor their baby has consumed any food or drink for 30 minutes prior to saliva sample collection (to prevent contamination of the sample with any food or drink residue).

After ensuring appropriate sanitization, the researcher, using gloved hands, will place a dental cotton roll (Premium Dental Cotton Rolls, Henry Schein) into the mother's mouth to soak up saliva, without the cotton roll being touched by the mother's hands. Once sodden, the cotton roll will be placed into a 10 mL syringe and squeezed into a labeled tube (1.5 mL; screw top) until at least 1 mL of saliva is extracted. This will be repeated to fill a second tube as a spare, in case of any measurement errors or spills during analysis.

For the infant's saliva sample, the same process will be followed; however, the researcher will hold 1 cotton roll and move it around the lower left and right sides of the infant's mouth and under the tongue until sodden. Filled sample and spare tubes will be placed into separate "mother" and "baby" labeled plastic ziplock bags and into a cooler box with an ice pack before being transferred into a –20 °C freezer until analysis.

After the baseline samples have been taken, an accurately measured dose of D_2O (Aldrich, 99.8 atom% deuterium) will be administered to the mother participant as a drink of water through a straw. The dose amount will be 30 g measured to the nearest 0.01 g (ie, doses will range from a minimum of 30 g to a maximum of 30.09 g). To ensure the full dose is consumed by the participant, the dose bottle will be filled twice with 50 mL of regular drinking water, inverted, and consumed by the mother through the same straw.

The elimination of D_2O from the mother and transfer to the infant will be traced over the 14-day collection period through analysis of deuterium enrichment of the saliva samples. Postdose (D_2O) saliva samples will be collected using the same methods described above from the mother and infant at 3 time points: 2 to 3 days postdose, 6 to 7 days postdose, and 13 to 14 days postdose. The reduced sampling protocol from the original International Atomic Energy Agency (IAEA) protocol will be used [24], which has shown high sensitivity and specificity [33].

Infant morbidity will be assessed at visits 3 and 4 with a one-off question: "Has baby been unwell in the last week?" to determine whether any illness occurred over the DTM period.

Any participants expressing breast milk will be asked to weigh and record the amount of milk expressed and indicate whether this was fed back to the infant, either fully or partially, during the 14-day DTM period. An expression log and electronic kitchen scale (Salter, model 1023) will be provided to the participant for recording purposes. Participants will be asked not to feed the infant any expressed milk that had been stored prior to the consumption of D_2O, as this cannot be tracked; therefore, infants will only consume expressed milk from the 14-day DTM period. Specific expression storage bags will be provided to participants to ensure accurate tracking of milk expressed during the DTM period and to minimize any confusion with milk expressed and stored outside of this time.

Maternal Fat Mass and Fat-Free Mass

Maternal body composition will be measured through the DTM technique, providing calculated fat mass and fat-free mass [24]. Participants will be asked to collect their own saliva sample 4 hours (or at least between 3 and 4 hours) after administration of the stable isotope, using the same methods used by the researcher at visit 1. They will be provided with equipment and instructions and asked to store the sample in their refrigerator until it is collected by a researcher at the second visit.

Human Milk Sample Collection

At the end of the DTM assessment period, a subsample of a complete human milk sample (full expression) will be collected from the participant in their home. At visit 4, participants will either be provided with a sterile breast pump (manual or electric) and collection bottle or given the option of using their personal breast pump, depending on preference. For those using their own pump, a microwave steam sterilizer bag (Medela or New Beginnings) will be provided so that the participant can sterilize the pump parts prior to expression, to minimize contamination.

Participants will be asked to, prior to expression, to clean their hands and the breast area with demineralized water and a cleaning gauze as provided, in order to remove any trace elements from creams that may have been applied to the hands or nipple area (eg, barrier creams, which commonly contain zinc). Participants will be asked to express a full milk sample from 1 breast in the morning (before 12 PM), at least 2 hours after the last feeding or pumping session on that breast, to standardize the collection period.

This method of human milk collection has been shown to be representative of the milk that an infant consumes, and by standardizing the collection period for all participants, the samples will be appropriate for comparisons between participants at the same time point post partum [62]. Participants will be asked to gently invert their full milk sample 3 times immediately after collection and then pour a 12-mL subsample into a collection container provided by the study that has been labeled with the participant ID and to indicate on the container the volume collected. The 12-mL amount was chosen to balance analytical needs (this is the volume required for the planned analyses) with sensitivity to mothers who may have concerns about their milk supply (their infant can be fed the remainder of the expressed milk).

Breastfeeding Observation

A one-off breastfeeding will be observed in the final visit (visit 5) by an IBCLC. The IBCLC will use the Bristol Breastfeeding Assessment Tool [63] to assess adequate latch and transfer of milk by the infant. This tool has 4 main elements, scoring a breastfeeding on positioning, sucking, swallowing, and attachment. A score of 0 (poor feeding elements) to 8 (good feeding elements) will be recorded by the IBCLC on the standardized Bristol Breastfeeding Assessment Tool form [63].

XSL•FO
RenderX

Test Weighing

During the breastfeeding observation by the IBCLC, a single test weighing of human milk transfer will be taken using an electronic baby scale (Seca model 334 or Marsden model M-300). Single test weighing, the process of weighing the infant before and after a single breastfeed, is often used clinically by health care professionals working with breastfeeding dyads to assess adequate milk transfer. The volume of milk transferred in the single test weighing will be compared with the human milk volume estimated by the DTM technique to investigate whether a one-off test weighing is indicative of actual human milk volume over a 2-week period.

Anthropometry

Infant weight and length will be measured using an electronic baby scale (Seca model 334) and a 99-cm measuring mat (Seca model SE210). Infants will be weighed naked with only a dry nappy of known weight. Maternal weight (minimal clothing, no shoes) will be measured using an electronic scale (Tanita model HD-351), and height will be measured using a portable stadiometer (Wedderburn model WMHM200P or Seca model 213), depending on whether the visit is at the participants' home or in the clinic. Measurements will be conducted by trained research staff in duplicate, following the World Health Organization (WHO) anthropometric protocol [64]. A third measurement will be taken if the difference between the first and second measurements exceeds the maximum allowable difference (0.1 kg and 0.5 kg for infant and maternal weight, respectively, and 0.7 cm for length and height).

At visit 5 (the final visit), photographs will be taken of anthropometric measurements collected during Well Child Tamariki Ora visits. These will be entered into REDCap. Well Child Tamariki Ora visits are free health assessments that are offered to all New Zealand families and conducted between ages 6 weeks and 5 years [65]. These anthropometric data will be used alongside those directly measured in the study to assess infant growth trajectory over time, which will be compared with the WHO Child Growth Standards [66]. Infant weight and length data will be used to calculate BMI and determine BMI-for-age *z* score using sex-specific reference data from the WHO Child Growth Standards [66].

Perception of Infant and Maternal Sleep

Infant sleep will be assessed using questions from a New Zealand sleep study [67], modified to include questions about infant sleep location and changes to infant sleep patterns. Participants will be asked a question about their own sleep quality and sleep quantity, developed by the research team, in the main questionnaire to assess maternal sleep perception.

Potential Predictors of Human Milk Intake

Overview

Human milk intake is complex and can be influenced by a range of different factors. A selection of potential maternal and infant factors that may influence human milk intakes has been identified through literature searches and advice from breastfeeding experts. These are summarized in Table 2.

XSL•FO
RenderX

Table 2. Potential predictors (maternal and infant) of human milk intakes.

Factors that potentially predict human milk intakes	Explanation of potential impact or hypothesis	References
Maternal factors		
Breastfeeding self-efficacy	Mothers who have greater self-efficacy may have infants with higher intake.	• Khresheh and Ahmad [68]
Breast or chest trauma	Mothers who have experienced breast or chest trauma may have infants with lower intake.	• Kraut et al [69] • Haifeng et al [70] • Stopenski et al [71] • Leal et al [72] • Holmgren et al [73]
Cannabis	Mothers who use cannabis may have infants with lower intake.	• Castro-Navarro et al [46]
D-MER[a]	Mothers who experience D-MER (sudden, intense negative emotions that occur immediately before and during milk let-down) may have infants with lower intake.	• Ureño et al [74]
Early return of menses	Mothers who have had return of menses may have infants with lower intake.	• Spencer et al [49]
Education	Mothers who are more educated may have infants with higher intake.	• Bandara et al [39] • Haszard et al [31]
Employment (current)	Mothers who are in current employment may have infants with lower intake.	• Haszard et al [31]
Following specific diet	Mothers who are consuming specific dietary patterns (eg, low fat, dairy-free, and vegetarian) may have infants with differing intake.	• Spencer et al [49]
Galactagogues or antigalactagogues	Mothers who are consuming galactagogues (a substance that is thought to increase milk volume) may have infants with higher intake, and mothers who are consuming antigalactagogues may have infants with lower intake.	• Brodribb et al [75] • Johnson et al [76]
Hormonal birth control	Mothers who are taking hormonal birth control may have infants with lower intake.	• Berens et al [77]
Maternal BMI	Mothers with higher BMI may have infants with lower intake.	• Hilson et al [40] • Haszard et al [31]
Maternal fat mass	Mothers who have higher fat mass may have infants with lower intake.	• Diana et al [78]
Parity	Mothers who have previously given birth may have infants with lower intake.	• Whitehead et al [41]
Previous breastfeeding experience	Mothers with previous breastfeeding success may have infants with higher intake, while those who did not have success previously may have infants with lower intake.	• Huang et al [79]
Maternal morbidities	Mothers who have medical conditions may have infants with lower intake.	• Spencer et al [49]
Maternal sleep	Hypothesis that maternal sleep may impact infant milk intake.	• N/A[b]
Maternal use of medications	Mothers on certain medications may have infants with differing intake.	• Farah et al [48]
Mental health	Mothers who experience challenges with mental health may have infants with lower intake.	• Spencer et al [49]
Pain while breastfeeding	Mothers who experience pain while breastfeeding may have infants with lower intake.	• McGuire [80]
Smoking or nicotine	Mothers who use nicotine may have infants with lower intake.	• Napierala et al [47]
Stress	Mothers with higher levels of stress may have infants with lower intake.	• Nagel et al [50]
Infant factors		
Age	It is known that infant intake varies over the course of lactation.	• Bandara et al [39] • da Costa et al [23] • Whitehead et al [41] • Brown et al [42] • Haszard et al [31]

XSL•FO
RenderX

Factors that potentially predict human milk intakes	Explanation of potential impact or hypothesis	References
Body weight	Infants with smaller body weight may have lower intake.	• Bandara et al [39] • Daniels et al [22] • Agne-Djigo et al [43] • Haisma 2003 [34] • Haszard et al [31]
Demand vs schedule feeding	Infants fed on demand rather than by schedule may have higher intake.	• Dewey et al [44] • Haszard et al [31]
Growth trajectory	Infants with greater growth velocity may have higher intake.	• Olga et al [36]
Household temperature and humidity	Higher household temperature and humidity may be associated with higher intake.	• Baby et al [81]
Number of breast feeds (average per day)	Infants feeding more often may have higher intake.	• Brown et al [42] • Haszard et al [31]
Sex	Male infants may have higher intake than female infants.	• da Costa et al. 2010 [23]
Spilling (average quantity)	Hypothesis that intake may differ in infants who spill.	• N/A
Term, preterm, or early term	Infants born early may have lower intake.	• Lapillonne et al [82] • Altuntas et al [83] • Boies and Vaucher [84]

[a]D-MER: dysphoric milk ejection reflex.

[b]N/A: not applicable.

Breastfeeding Self-Efficacy

Maternal breastfeeding self-efficacy will be assessed using the Breastfeeding Self-Efficacy Scale–Short Form [85] in the main questionnaire. This is a validated 14-item questionnaire with a 5-point Likert scale (1=not confident to 5=very confident) to determine the participant's confidence in their ability to breastfeed. Total scores range from 14 to 70, with higher scores indicating higher level of breastfeeding self-confidence.

Health History

Maternal and infant health history will be collected in the main questionnaire. Questions about birth and breastfeeding experience, breastfeeding support and practices, infant gestational age, parity, current medications or medical conditions, use of supplements, maternal dietary pattern and appetite, breast surgery, and return of menses will be included to determine their potential impact on human milk intake.

The main questionnaire will also include questions about breastfeeding practices, such as the average number of breastfeeds in 24 hours, whether participants are breastfeeding on demand or on a schedule, how much the infant spills, and questions on previous breastfeeding experience. Maternal stress will be assessed using the Global Measure of Perceived Stress Questionnaire [86].

Questions about nipple or breast trauma will be asked at visit 1 and repeated at subsequent visits (visits 2 to 4) to determine whether any changes occurred during the study period. Breastfeeding pain and maternal experience with D-MER will be assessed at visit 5 (final) with the IBCLC.

Breastfeeding pain will be the mother's self-reported usual pain during a breastfeed on a scale of 0 (no pain) to 10 (worst pain). If pain is reported as greater than 0, follow-up questions will be asked about the type of pain, timing, frequency, and its impact on breastfeeding. This will be assessed separately from the questions around nipple or breast trauma, as those focus on damage and injury rather than general pain. Presence of D-MER will be assessed using the validated D-MER questionnaire [59]. Those who are experiencing symptoms of pain or D-MER will be offered further support by the IBCLC based on clinical judgement.

Nicotine and Cannabis Use

Questions about nicotine use will be assessed in the main questionnaire. Due to the rising rates of electronic cigarette ("vape") use in New Zealand [87] and internationally [88], as well as evidence that use could be negatively associated with breastfeeding duration [89], questions on cigarette and electronic cigarette use will both be included. The number of cigarettes per day, as well as the use, nicotine level, and flavor of electronic cigarettes, will be assessed via questionnaire.

Cannabis use will be assessed in the main questionnaire. As cannabis is illegal in New Zealand, care will be taken when asking these questions to ensure that confidentiality is maintained and participants feel safe responding. Participants will be reminded that all responses are confidential and no identifying information will be kept alongside their responses. The questions about cannabis will be optional, allowing participants to leave this section blank if they feel more comfortable. Cannabis use will be assessed via questions used

XSL•FO
RenderX

previously [90], modified to include questions about use after birth, along with use during pregnancy.

Household Temperature and Humidity

To explore the impact of potential heat stress on infant milk intake, household temperature and humidity will be measured using a digital hygrometer (Niome) with a temperature accuracy of ±1 °C and humidity accuracy of ±5%. For participants who choose to have their visits in the home that the infant is most commonly in, the digital hygrometer will be placed in the main room of the home where the visit takes place. The household temperature and humidity will be recorded after 15 to 20 minutes, which gives the hygrometer enough time to obtain an accurate reading of the room.

Laboratory Analyses

Human Milk Volume

The deuterium enrichment in saliva samples will be determined by Fourier transform infrared spectrometry (Thermo Fisher Nicolet Summit, Thermo Fisher Scientific) with a liquid accessory attachment (SPECAC Pearl and CaF_2 100-µm pathlength top and bottom window plates). Enrichment will be determined using a modified method described by the IAEA [24]. Previously frozen saliva samples will be thawed at room temperature for 30 minutes, inverted 4 times, and centrifuged at 1000 × g for 10 minutes.

A trained researcher will pipette 45 µL of sample (baseline or postdose) onto the cell window, close the top cell window, and measure the absorption spectrum based on settings described in the IAEA protocol [24]. Each sample will be measured in duplicate unless the difference between the 2 sample measurements is greater than 3 mg/kg, in which case a third or fourth measure will be taken until 2 measurements are less than 3 mg/kg apart. The mean of the 2 closest measurements (within 3 mg/kg) will be the calculated D_2O at each time point.

Following the determination of D_2O enrichment, the calculation of human milk volume, infant nonmilk water intake, and classification of exclusive breastfeeding will be determined using a validated method [32,33]. The mother's measured height and weight and the infant's measured length and weight (collection described above) will be used in the calculation of human milk volume.

Human Milk Composition

Macronutrients

Human milk samples will be collected from the participant on the same day of expression. Samples will be carefully mixed and aliquoted into a 10-mL collection tube and frozen within 48 hours of collection, at –20 °C. Frozen samples will be defrosted at room temperature and analyzed within 2 hours. Once samples are defrosted, they will be placed into a heating bath where they will sit for approximately 20 minutes to warm to 40 °C. When the samples have reached 40 °C, they will be mixed by inversion and homogenized using an ultrasonic processor (Miris Ultrasonic Processor).

Macronutrient composition will be analyzed using a human milk analyzer (HMA, Miris). A total of 3 mL of sample will be slowly syringed into the HMA and repeated in triplicate (3 analyzes per sample) to ensure accuracy of the reading.

The HMA uses a combination of midinfrared transmission spectroscopy principles to measure concentrations of fat, carbohydrate (lactose and oligosaccharides), and both crude and true protein in human milk, as well as providing calculated values for total solids and energy. The HMA software uses internal calibration based on chemical standard methods validated for analysis of human milk. The Röse-Gottlieb method is used for total fat [91], and Kjeldahl [92] for crude protein. True protein is crude protein multiplied by 0.8 to correct for nonprotein nitrogen. The total carbohydrate content is calculated as total solids minus fat, protein, and ash [93]. Total solids are measured by drying oven [94].

Minerals and Trace Elements

Minerals and trace elements will be determined following methodology used previously [22]. Specifically, minerals (sodium, magnesium, phosphorus, potassium, and calcium) and trace elements (iron, copper, zinc, selenium, and iodine) will be analyzed in the Centre for Trace Element Analysis, Department of Chemistry, University of Otago, New Zealand, by inductively coupled plasma mass spectrophotometry (Agilent 7900; Agilent Technologies). Briefly, 10 mL of quartz sub-boiling 14 N nitric acid and 1 mL of 30% hydrogen peroxide will be added to approximately 1 g of the human milk sample. Samples will be microwave-digested at 200 °C for 15 minutes and dried down to approximately 1 mL. The final sample will be diluted to 25 mL with 2% aqueous nitric acid. Precision and accuracy will be checked using in-house pooled samples and multielement reference standards (SRM 1549a, whole milk powder) from the National Institute of Standards and Technology.

Data Management

Upon enrollment, participants will be assigned a unique identifying code (participant ID). To preserve confidentiality, during data collection all data will be recorded against the participant ID. The key linking the participant ID with the participant's name and address will be kept secure. Although it will need to be accessible to research staff so they can attend home visits (name and address) and collect data (participant ID), research staff will only have access to health information when it is essential for specific study-related tasks.

Responses to online questionnaires will be linked to participant ID numbers. Participant IDs will also be written on biological sample tubes and recording sheets. All study data will be stored in the secure REDCap database [58]. In the unlikely event that data cannot be entered directly into the REDCap database, they will be entered on a hard-copy datasheet and later transferred into REDCap. All paper copies will be stored in a locked office.

Data will be collected following rigorous protocols and cleaned according to standard data entry procedures. Quality control visits (observations of research staff conducting visits) will be conducted by senior research staff to ensure that staff follow data collection protocols consistently. Data quality checks will be run on all entered data to check for accuracy, consistency, and completeness. The results database will be stored on

XSL•FO
RenderX

research team members' computers, all of which are password protected.

Statistical Analysis

Sample Size

We anticipate a dropout rate, including incomplete data, of 20% based on recruitment and completion data from a similar recent study [95]. Therefore, we aim to recruit 150 mother–infant dyads, which is expected to result in approximately 120 with complete data. With 120 participants, mean human milk intake can be estimated to a 95% CI of SD 0.18. If the distribution of intakes is similar to previous data, this corresponds to around ±27 mL/day (ie, 3% of an estimated intake of 787 mL/day at 3 months [22]). A final sample of 120 will provide a robust sample size to estimate the correlation between human milk intake and PIMS, as well as to estimate differences between groups [96].

Data Analysis Plan

Human milk intakes of infants aged 3 months will be described. Perception of milk supply will be collected as a continuous scale and dichotomized to create a variable for comparing low (insufficient) perceived milk supply and high (oversupply) perceived milk supply. Regression models will be used to assess how maternal perception of human milk supply (categorical independent variable) is associated with actual human milk volumes produced by lactating participants at 3 months post partum (dependent variable). Mean differences and 95% CIs will be estimated. Similarly, the relationships between predetermined potential predictors (Table 1) of human milk intakes and actual milk intakes will be described using regression modeling. Residuals of regression models will be plotted to assess homoscedasticity and normality. Results will be presented for descriptive purposes; therefore, no adjustment for confounders will be made. However, moderation by demographic characteristics will be investigated descriptively by presenting stratified results. Data will be illustrated graphically where appropriate.

To understand whether certain factors predict breastfeeding problems, the breastfeeding assessment tool will be used. These data will be analyzed using prediction modeling techniques, including a combination of linear regression models (reporting R^2), and least absolute shrinkage and selection operator regression, to understand which factors might combine to predict breastfeeding problems. In addition, if there are any modifiable factors indicated to cause breastfeeding problems, a causal inference approach to analysis will be used to determine which confounders to include in the model, including the development of a directed acyclic graph.

BMI and weight velocity between birth and age up to 5 months will be calculated from the Well Child Tamariki Ora data collected. An initial investigation will be conducted into whether infant milk intakes and infant growth over time (birth to age up to 5 months) can be used to assess human milk inadequacy (actual low milk volume) using regression models.

Statistical analyses will be conducted under the guidance of a biostatistician using Stata (version 18; StataCorp LLC) software.

Results

This study was funded in April 2023 by the Health Research Council of New Zealand (grant 23/461). Recruitment for this study began in February 2025 and is currently underway. As of January 17, 2026, a total of 91 participants have been enrolled. Recruitment is anticipated to conclude in June 2026, with analysis expected to be completed by February 2027. Final results are anticipated to be disseminated in late 2027 following completion of data analysis.

Discussion

This study will explore the relationship between maternal perception and actual milk quantity and quality, providing significant implications for improving breastfeeding support. If PIMS is linked to specific physiological characteristics of human milk, such as lower volume or suboptimal nutrient concentrations, this could provide a biological basis for potentially individualized interventions to support the continuation of breastfeeding. Alternatively, if perception is not linked to measurable differences, this would reinforce the need for better psychosocial support and education to sustain breastfeeding, or the development of tools that can be used to assess milk supply and reassure mothers that their supply is sufficient. Because perception is a modifiable factor, identifying its relationship to human milk volume and composition offers a critical opportunity to intervene early, support breastfeeding continuation, and improve infant and maternal health outcomes.

To the best of our knowledge, there is currently no published research using the DTM technique to investigate whether PIMS is related to actual milk insufficiency, nor has the association with nutrient composition been investigated, a gap in the literature as highlighted by the 2023 Lancet Series on Breastfeeding [51]. While 1 study has investigated the relationship between perceived and actual milk insufficiency [20], it was restricted to first-time mothers and used test-weighing methods to estimate infant milk intake, both of which inherently limit the interpretation of the findings.

The main strength of MABBS is the use of the DTM stable isotope technique for estimating human milk volume in relation to mothers' perception of milk supply. This methodology is considered to be the gold standard, allowing the most accurate reporting of infant milk intake [33]. This is particularly important given our large sample size (at least 120 participants), which allows us to determine daily infant milk intake to an accuracy of ±27 mL/day. Finally, we aim to recruit a diverse population with good ethnic representation to ensure that the outcomes are relevant to all people living in New Zealand.

This study also has some limitations. A narrow age range of participants (those with babies aged 2 to <4 months) was required because of the variation in milk intake at different stages of lactation. While this is necessary for the methodology of this study, it may impact our ability to capture participants with a wide range of perception of milk supply. Furthermore, while the research team aim to make the results relevant to the New Zealand population, due to differences in populations, the

XSL•FO
RenderX

findings of this study may not be generalizable to people in other countries outside New Zealand.

We will disseminate the results of this study through multiple channels. For academic dissemination, the results will be published in open-access and high impact journals, as well as presented at relevant conferences. However, it is of high importance to members of the research team to ensure that results are also useful to inform practice at a clinical level. With help from the advisory group and experts in this space, results will be shared with those working with breastfeeding families to help shape clinical practice. This may include webinars or resources specifically designed for health professionals to inform personalized intervention strategies for breastfeeding parents.

Acknowledgments

We are grateful to the Māmā and Baby Breastfeeding Study (MABBS) advisory group for providing suggestions and feedback about the study design.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical restrictions related to the consent provided by participants but are available from the corresponding author on reasonable request.

Funding

This research is funded by the Health Research Council of New Zealand (grant 23/461). BB is supported by a University of Otago doctoral scholarship.

Authors' Contributions

BB contributed to study design, data collection, and wrote the first and subsequent drafts of the manuscript. RT, CP, JH, BT, ALH, and IK contributed to the study design and critical review of the manuscript. SB and KG contributed to data collection and critical review of the manuscript. LD (primary investigator of MABBS) was responsible for study conception, study design, data collection, drafting of the manuscript, and primary responsibility for the final content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from HRC.
[PDF File (Adobe PDF File), 67 KB - resprot_v15i1e84776_app1.pdf]

References

1. Health eating guidelines for New Zealand babies and toddlers (0-2 years old). Ministry of Health. 2021. URL: https://www.health.govt.nz/system/files/2021-09/healthy-eating-guidelines-for-new-zealand-babies-and-toddlers-nov21-v3.pdf [accessed 2026-02-05]
2. Breastfeeding. World Health Organization. URL: https://www.who.int/health-topics/breastfeeding#tab=tab_1 [accessed 2026-03-03]
3. National infant feeding data at discharge report 2024. New Zealand Breastfeeding Alliance. 2024. URL: https://www.babyfriendly.org.nz/fileadmin/Documents/National_Infant_Feeding_Data_at_Discharge_2024_FINAL.pdf [accessed 2026-02-05]
4. New Zealand health survey annual data explorer 2024/2025. Ministry of Health. 2026. URL: https://minhealthnz.shinyapps.io/nz-health-survey-2024-25-annual-data-explorer/_w_9c480a17f1fe4910999ae79ee111d5e6/#!/explore-indicators [accessed 2026-02-05]
5. Breastfeeding report card. Centres for Disease Control and Prevention. 2022. URL: https://www.cdc.gov/breastfeeding-data/media/pdfs/2024/06/2022-Breastfeeding-Report-Card-508.pdf [accessed 2026-05-02]
6. Netting MJ, Moumin NA, Knight EJ, Golley RK, Makrides M, Green TJ. The Australian feeding infants and toddler study (OzFITS 2021): Breastfeeding and early feeding practices. Nutrients 2022;14(1):206 [FREE Full text] [doi: 10.3390/nu14010206] [Medline: 35011081]
7. World health assembly re-commits to global nutrition targets and marketing regulations. World Health Organization. URL: https://www.who.int/news/item/27-05-2025-world-health-assembly-re-commits-to-global-nutrition-targets-and-marketing-regulations [accessed 2025-06-26]
8. Holmes W, Thorpe L, Phillips J. Influences on infant-feeding beliefs and practices in an urban aboriginal community. Aust N Z J Public Health 1997;21(5):504-510 [FREE Full text] [doi: 10.1111/j.1467-842x.1997.tb01743.x] [Medline: 9343896]

XSL•FO
RenderX

9. Brown CRL, Dodds L, Legge A, Bryanton J, Semenic S. Factors influencing the reasons why mothers stop breastfeeding. Can J Public Health 2014;105(3):e179-e185 [FREE Full text] [doi: 10.17269/cjph.105.4244] [Medline: 25165836]
10. Gatti L. Maternal perceptions of insufficient milk supply in breastfeeding. J Nurs Scholarsh 2008;40(4):355-363 [FREE Full text] [doi: 10.1111/j.1547-5069.2008.00234.x] [Medline: 19094151]
11. Sakalidis VS, Rea A, Perrella SL, McEachran J, Collis G, Miraudo J, et al. Longitudinal changes in wellbeing amongst breastfeeding women in Australia and New Zealand during the COVID-19 pandemic. Eur J Pediatr 2022;181(10):3753-3766 [FREE Full text] [doi: 10.1007/s00431-022-04580-y] [Medline: 35976413]
12. Hauck YL, Fenwick J, Dhaliwal SS, Butt J. A western Australian survey of breastfeeding initiation, prevalence and early cessation patterns. Matern Child Health J 2011;15(2):260-268. [doi: 10.1007/s10995-009-0554-2] [Medline: 20077131]
13. Essex C, Smale P, Geddis D. Breastfeeding rates in New Zealand in the first 6 months and the reasons for stopping. N Z Med J 1995;108(1007):355-357. [Medline: 7566772]
14. Glover M, Manaena-Biddle H, Waldon J. Influences that affect maori women breastfeeding. Breastfeed Rev 2007;15(2):5-14. [Medline: 17695072]
15. Hill PD, Humenick SS. Nurs Res 1996;45(3):136-140. [doi: 10.1097/00006199-199605000-00003] [Medline: 8637793]
16. Huang Y, Gau M, Huang C, Lee J. Supplementation with cup-feeding as a substitute for bottle-feeding to promote breastfeeding. Chang Gung Med J 2009;32(4):423-431 [FREE Full text] [Medline: 19664349]
17. Kent JC, Mitoulas LR, Cregan MD, Ramsay DT, Doherty DA, Hartmann PE. Volume and frequency of breastfeedings and fat content of breast milk throughout the day. Pediatrics 2006;117(3):e387-e395. [doi: 10.1542/peds.2005-1417] [Medline: 16510619]
18. Giugliani ERJ. Growth in exclusively breastfed infants. J Pediatr (Rio J) 2019;95 Suppl 1:79-84 [FREE Full text] [doi: 10.1016/j.jped.2018.11.007] [Medline: 30594467]
19. Borra C, Iacovou M, Sevilla A. New evidence on breastfeeding and postpartum depression: The importance of understanding women's intentions. Matern Child Health J 2015;19(4):897-907 [FREE Full text] [doi: 10.1007/s10995-014-1591-z] [Medline: 25138629]
20. Galipeau R, Dumas L, Lepage M. Perception of not having enough milk and actual milk production of first-time breastfeeding mothers: Is there a difference? Breastfeed Med 2017;12:210-217. [doi: 10.1089/bfm.2016.0183] [Medline: 28326807]
21. Nutrient reference values for Australia and New Zealand including recommended dietary intakes. Australian Government National Health and Medical Research Council and New Zealand Ministry of Health. 2006. URL: https://www.nhmrc.gov.au/about-us/publications/nutrient-reference-values-australia-and-new-zealand-including-recommended-dietary-intakes [accessed 2026-02-05]
22. Daniels L, Gibson RS, Diana A, Haszard JJ, Rahmannia S, Luftimas DE, et al. Micronutrient intakes of lactating mothers and their association with breast milk concentrations and micronutrient adequacy of exclusively breastfed Indonesian infants. Am J Clin Nutr 2019;110(2):391-400 [FREE Full text] [doi: 10.1093/ajcn/nqz047] [Medline: 31152543]
23. da Costa TH, Haisma H, Wells JC, Mander AP, Whitehead RG, Bluck LJ. How much human milk do infants consume? Data from 12 countries using a standardized stable isotope methodology. J Nutr 2010;140(12):2227-2232 [FREE Full text] [doi: 10.3945/jn.110.123489] [Medline: 20980653]
24. Stable isotope technique to assess intake of human milk in breastfed infants. International Atomic Energy Agency. 2010. URL: https://www-pub.iaea.org/MTCD/Publications/PDF/Pub1429_web.pdf [accessed 2026-02-05]
25. Kent JC, Gardner H, Lai C, Hartmann PE, Murray K, Rea A, et al. Hourly breast expression to estimate the rate of synthesis of milk and fat. Nutrients 2018;10(9):1144 [FREE Full text] [doi: 10.3390/nu10091144] [Medline: 30135368]
26. Lai CT, Hale TW, Simmer K, Hartmann PE. Measuring milk synthesis in breastfeeding mothers. Breastfeed Med 2010;5(3):103-107. [doi: 10.1089/bfm.2009.0074] [Medline: 20433368]
27. Roznowski DM, Wagner EA, Riddle SW, Nommsen-Rivers LA. Validity of a 3-hour breast milk expression protocol in estimating current maternal milk production capacity and infant breast milk intake in exclusively breastfeeding dyads. Breastfeed Med 2020;15(10):630-638 [FREE Full text] [doi: 10.1089/bfm.2019.0182] [Medline: 32700964]
28. Morton J, Hall JY, Wong RJ, Thairu L, Benitz WE, Rhine WD. Combining hand techniques with electric pumping increases milk production in mothers of preterm infants. J Perinatol 2009;29(11):757-764. [doi: 10.1038/jp.2009.87] [Medline: 19571815]
29. Prime DK, Garbin CP, Hartmann PE, Kent JC. Simultaneous breast expression in breastfeeding women is more efficacious than sequential breast expression. Breastfeed Med 2012;7(6):442-447 [FREE Full text] [doi: 10.1089/bfm.2011.0139] [Medline: 23039397]
30. Ramsay DT, Kent JC, Owens RA, Hartmann PE. Ultrasound imaging of milk ejection in the breast of lactating women. Pediatrics 2004;113(2):361-367. [doi: 10.1542/peds.113.2.361] [Medline: 14754950]
31. Haszard JJ, Heath AM, Taylor RW, Bruckner B, Katiforis I, McLean NH, et al. Equations to estimate human milk intake in infants aged 7 to 10 months: Prediction models from a cross-sectional study. Am J Clin Nutr 2024;120(1):102-110 [FREE Full text] [doi: 10.1016/j.ajcnut.2024.04.009] [Medline: 38890036]
32. Liu Z, Diana A, Slater C, Preston T, Gibson RS, Houghton L, et al. Development of a nonlinear hierarchical model to describe the disposition of deuterium in mother-infant pairs to assess exclusive breastfeeding practice. J Pharmacokinet Pharmacodyn 2019;46(1):1-13 [FREE Full text] [doi: 10.1007/s10928-018-9613-x] [Medline: 30430351]

XSL•FO
RenderX

33. Liu Z, Diana A, Slater C, Preston T, Gibson RS, Houghton L, et al. Development of a parsimonious design for optimal classification of exclusive breastfeeding. CPT Pharmacometrics Syst Pharmacol 2019;8(8):596-605 [FREE Full text] [doi: 10.1002/psp4.12428] [Medline: 31215140]
34. Haisma H, Coward WA, Albernaz E, Visser GH, Wells JCK, Wright A, et al. Breast milk and energy intake in exclusively, predominantly, and partially breast-fed infants. Eur J Clin Nutr 2003;57(12):1633-1642. [doi: 10.1038/sj.ejcn.1601735] [Medline: 14647230]
35. Lopez-Teros V, Limon-Miro A, Astiazaran-Garcia H, Tanumihardjo S, Tortoledo-Ortiz O, Valencia M. 'Dose-to-Mother' deuterium oxide dilution technique: An accurate strategy to measure vitamin a intake in breastfed infants. Nutrients 2017;9(2):169 [FREE Full text] [doi: 10.3390/nu9020169] [Medline: 28230781]
36. Olga L, Vervoort J, van Diepen JA, Gross G, Petry CJ, Prentice PM, et al. Associations between breast milk intake volume, macronutrient intake and infant growth in a longitudinal birth cohort: The Cambridge baby growth and breastfeeding study (CBGS-BF). Br J Nutr 2023;130(1):56-64 [FREE Full text] [doi: 10.1017/S0007114522003178] [Medline: 36259139]
37. Urteaga N, San Miguel JL, Aguilar AM, Muñoz M, Slater C. Nutritional status and human milk intake of exclusively breast-fed infants at high altitude in La Paz, Bolivia. Br J Nutr 2018;120(2):158-163. [doi: 10.1017/S0007114518001137] [Medline: 29947320]
38. Cloostermans L, Allegaert K, Smits A, Van Neste M. The deuterium oxide dilution method to quantify human milk intake volume of infants: A systematic review-a contribution from the ConcePTION Project. Nutrients 2024;16(23):4205 [FREE Full text] [doi: 10.3390/nu16234205] [Medline: 39683598]
39. Bandara T, Hettiarachchi M, Liyanage C, Amarasena S, Wong WW. The deuterium oxide-to-the-mother method documents adequate breast-milk intake among Sri Lankan infants. J Nutr 2015;145(6):1325-1329 [FREE Full text] [doi: 10.3945/jn.115.211771] [Medline: 25904731]
40. Hilson JA, Rasmussen KM, Kjolhede CL. High prepregnant body mass index is associated with poor lactation outcomes among white, rural women independent of psychosocial and demographic correlates. J Hum Lact 2004;20(1):18-29. [doi: 10.1177/0890334403261345] [Medline: 14974697]
41. Whitehead R, Rowland M, Hutton M, Prentice A, Müller E, Paul A. Factors influencing lactation performance in rural Gambian mothers. Lancet 1978;2(8082):178-181. [doi: 10.1016/s0140-6736(78)91920-7] [Medline: 78383]
42. Brown KH, Black RE, Robertson AD, Akhtar NA, Ahmed G, Becker S. Clinical and field studies of human lactation: Methodological considerations. Am J Clin Nutr 1982;35(4):745-756. [doi: 10.1093/ajcn/35.4.745] [Medline: 7072626]
43. Agne-Djigo A, Kwadjode KM, Idohou-Dossou N, Diouf A, Guiro AT, Wade S. Energy intake from human milk covers the requirement of 6-month-old Senegalese exclusively breast-fed infants. Br J Nutr 2013;110(10):1849-1855. [doi: 10.1017/S0007114513001074] [Medline: 23578434]
44. Dewey KG, Lönnerdal B. Infant self-regulation of breast milk intake. Acta Paediatr Scand 1986;75(6):893-898. [doi: 10.1111/j.1651-2227.1986.tb10313.x] [Medline: 3564971]
45. Brown KJ, Beck KL, von Hurst P, Heath A, Taylor R, Haszard J, et al. Adherence to infant feeding guidelines in the first foods New Zealand study. Nutrients 2023;15(21):4650 [FREE Full text] [doi: 10.3390/nu15214650] [Medline: 37960303]
46. Castro-Navarro I, McGuire MA, Williams JE, Holdsworth EA, Meehan CL, McGuire MK. Maternal cannabis use during lactation and potential effects on human milk composition and production: A narrative review. Adv Nutr 2024;15(4):100196 [FREE Full text] [doi: 10.1016/j.advnut.2024.100196] [Medline: 38432590]
47. Napierala M, Mazela J, Merritt TA, Florek E. Tobacco smoking and breastfeeding: Effect on the lactation process, breast milk composition and infant development. A critical review. Environ Res 2016;151:321-338. [doi: 10.1016/j.envres.2016.08.002] [Medline: 27522570]
48. Farah E, Barger MK, Klima C, Rossman B, Hershberger P. Impaired lactation: Review of delayed lactogenesis and insufficient lactation. J Midwifery Womens Health 2021;66(5):631-640. [doi: 10.1111/jmwh.13274] [Medline: 34596953]
49. Spencer B, Campbell S, Chamberlain K. Core Curriculum for Interdisciplinary Lactation Care. Burlington, MA: Jones & Bartlett Learning; 2024.
50. Nagel EM, Howland MA, Pando C, Stang J, Mason SM, Fields DA, et al. Maternal psychological distress and lactation and breastfeeding outcomes: A narrative review. Clin Ther 2022;44(2):215-227 [FREE Full text] [doi: 10.1016/j.clinthera.2021.11.007] [Medline: 34937662]
51. Pérez-Escamilla R, Tomori C, Hernández-Cordero S, Baker P, Barros AJD, Bégin F, 2023 Lancet Breastfeeding Series Group. Breastfeeding: Crucially important, but increasingly challenged in a market-driven world. Lancet 2023;401(10375):472-485 [FREE Full text] [doi: 10.1016/S0140-6736(22)01932-8] [Medline: 36764313]
52. Thomson CD. Selenium and iodine intakes and status in New Zealand and Australia. Br J Nutr 2004;91(5):661-672. [doi: 10.1079/BJN20041110] [Medline: 15137917]
53. Stråvik M, Gustin K, Barman M, Skröder H, Sandin A, Wold AE, et al. Infant iodine and selenium status in relation to maternal status and diet during pregnancy and lactation. Front Nutr 2021;8:733602 [FREE Full text] [doi: 10.3389/fnut.2021.733602] [Medline: 34988107]
54. Calcaterra V, Cena H, Scavone IAM, Zambon I, Taranto S, Ricciardi Rizzo C, et al. Thyroid health and selenium: The critical role of adequate intake from fetal development to adolescence. Nutrients 2025;17(14):2362 [FREE Full text] [doi: 10.3390/nu17142362] [Medline: 40732987]

XSL•FO
RenderX

55. Gunn TR. The incidence of breast feeding and reasons for weaning. N Z Med J 1984;97(757):360-363. [Medline: 6589521]
56. Shakked N, Isabelle S. The drivers of mothers' parental leave decisions. Ministry of Social Development. 2021. URL: https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/research/parental-leave-decisions/index.html [accessed 2026-03-03]
57. Census national and subnational usually resident population counts (by ethnic group, age, and Māori descent) and dwelling counts. Statistics New Zealand. URL: https://www.stats.govt.nz/information-releases/2023-census-population-counts-by-ethnic-group-age-and-maori-descent-and-dwelling-counts/ [accessed 2026-02-05]
58. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
59. Žutić M, Matijaš M, Nakić Radoš S. Dysphoric milk ejection reflex: Measurement, prevalence, clinical features, maternal mental health, and mother-infant bonding. Breastfeed Med 2025;20(2):133-139. [doi: 10.1089/bfm.2024.0172] [Medline: 39501787]
60. Standard for ethnicity v1.0.0. Statistics New Zealand. URL: https://aria.stats.govt.nz/aria/?gr=2246863185795301626484791686858159833048&gc=1245187816264847[illegible]/stats.govt.nz/cms/StatisticalStandard/vv0ovwUoTSSVDhpt [accessed 2025-04-08]
61. Atkinson J, Salmond C, Crampton P, Viggers H, Lacey K. NZDep2023 Index of Socioeconomic Deprivation: Research Report. 2023. URL: https://www.otago.ac.nz/__data/assets/pdf_file/0026/593135/NZDep2023-Research-Report-31-October-2024.pdf [accessed 2026-02-05]
62. Leghi GE, Middleton PF, Netting MJ, Wlodek ME, Geddes DT, Muhlhausler BS. A systematic review of collection and analysis of human milk for macronutrient composition. J Nutr 2020;150(6):1652-1670 [FREE Full text] [doi: 10.1093/jn/nxaa059] [Medline: 32240307]
63. Ingram J, Johnson D, Copeland M, Churchill C, Taylor H. The development of a new breast feeding assessment tool and the relationship with breast feeding self-efficacy. Midwifery 2015;31(1):132-137 [FREE Full text] [doi: 10.1016/j.midw.2014.07.001] [Medline: 25061006]
64. de Onis M, Onyango AW, Van den Broeck J, Chumlea WC, Martorell R. Measurement and standardization protocols for anthropometry used in the construction of a new international growth reference. Food Nutr Bull 2004;25(1 Suppl):S27-S36. [doi: 10.1177/15648265040251S104] [Medline: 15069917]
65. Well child tamariki ora visits, health information and services. Health New Zealand. URL: https://info.health.nz/services-support/pregnancy-birth-and-children-services/well-child-tamariki-ora/well-child-tamariki-ora-visits [accessed 2025-08-21]
66. World Health Organization. WHO Child Growth Standards: Length/height-For-Age, Weight-For-Age, Weight-For-Length, Weight-For-Height and Body Mass Index-For-Age: Methods and Development. Switzerland: World Health Organization; 2006.
67. Fangupo LJ, Haszard JJ, Russell-Camp T, Taylor RW, Richards R, Galland BC, et al. The measurement of young children's nocturnal sleep health and the development of the perception of infant and toddler sleep scale (PoITSS) in aotearoa New Zealand whānau (families). Sleep Health 2024;10(5):567-575 [FREE Full text] [doi: 10.1016/j.sleh.2024.07.004] [Medline: 39129127]
68. Khresheh RM, Ahmad NM. Breastfeeding self efficacy among pregnant women in Saudi Arabia. Saudi Med J 2018;39(11):1116-1122 [FREE Full text] [doi: 10.15537/smj.2018.11.23437] [Medline: 30397711]
69. Kraut RY, Brown E, Korownyk C, Katz LS, Vandermeer B, Babenko O, et al. The impact of breast reduction surgery on breastfeeding: Systematic review of observational studies. PLoS One 2017;12(10):e0186591 [FREE Full text] [doi: 10.1371/journal.pone.0186591] [Medline: 29049351]
70. Haifeng G, Yi Z, Ying C, Xin L, Yan L, Songtao D. Exploring the efficacy of comprehensive management of breast abscesses in restoring milk volume. J Hum Lact 2022;38(1):171-176. [doi: 10.1177/08903344211005400] [Medline: 33818190]
71. Stopenski S, Aslam A, Zhang X, Cardonick E. After chemotherapy treatment for maternal cancer during pregnancy, is breastfeeding possible? Breastfeed Med 2017;12:91-97. [doi: 10.1089/bfm.2016.0166] [Medline: 28170295]
72. Leal SC, Stuart SR, Carvalho HDA. Breast irradiation and lactation: A review. Expert Rev Anticancer Ther 2013;13(2):159-164. [doi: 10.1586/era.12.178] [Medline: 23406557]
73. Holmgren T, Lee AHX, Hocaloski S, Hamilton LJ, Hellsing I, Elliott S, et al. The influence of spinal cord injury on breastfeeding ability and behavior. J Hum Lact 2018;34(3):556-565. [doi: 10.1177/0890334418774014] [Medline: 29787691]
74. Ureño TL, Buchheit TL, Hopkinson SG, Berry-Cabán CS. Dysphoric milk ejection reflex: A case series. Breastfeed Med 2018;13(1):85-88. [doi: 10.1089/bfm.2017.0086] [Medline: 29115857]
75. Brodribb W. ABM clinical protocol #9: Use of galactogogues in initiating or augmenting maternal milk production, second revision 2018. Breastfeed Med 2018;13(5):307-314. [doi: 10.1089/bfm.2018.29092.wjb] [Medline: 29902083]
76. Johnson HM, Eglash A, Mitchell KB, Leeper K, Smillie CM, Moore-Ostby L, Academy of Breastfeeding Medicine. ABM clinical protocol #32: Management of hyperlactation. Breastfeed Med 2020;15(3):129-134. [doi: 10.1089/bfm.2019.29141.hmj] [Medline: 32031417]

XSL•FO
RenderX

77. Berens P, Labbok M, Academy of Breastfeeding Medicine. ABM clinical protocol #13: Contraception during breastfeeding, revised 2015. Breastfeed Med 2015;10(1):3-12. [doi: 10.1089/bfm.2015.9999] [Medline: 25551519]
78. Diana A, Haszard JJ, Houghton LA, Gibson RS. Breastmilk intake among exclusively breastfed Indonesian infants is negatively associated with maternal fat mass. Eur J Clin Nutr 2019;73(8):1206-1208 [FREE Full text] [doi: 10.1038/s41430-019-0458-1] [Medline: 31235956]
79. Huang Y, Ouyang YQ, Redding SR. Previous breastfeeding experience and its influence on breastfeeding outcomes in subsequent births: A systematic review. Women Birth 2019;32(4):303-309. [doi: 10.1016/j.wombi.2018.09.003] [Medline: 30274877]
80. McGuire TM. Drugs affecting milk supply during lactation. Aust Prescr 2018;41(1):7-9 [FREE Full text] [doi: 10.18773/austprescr.2018.002] [Medline: 29507453]
81. Baby J, Kaestel P, Preston T, Duffull SB, Liu Z, Diana A, et al. Effects of variations in atmospheric temperature and humidity on the estimation of exclusive breastfeeding status using the deuterium oxide dose-to-mother technique. Front Pediatr 2023;11:1188811 [FREE Full text] [doi: 10.3389/fped.2023.1188811] [Medline: 38034831]
82. Lapillonne A, Bronsky J, Campoy C, Embleton N, Fewtrell M, Fidler Mis N, ESPGHAN Committee on Nutrition. Feeding the late and moderately preterm infant: A position paper of the European society for paediatric gastroenterology, hepatology and nutrition committee on nutrition. J Pediatr Gastroenterol Nutr 2019;69(2):259-270 [FREE Full text] [doi: 10.1097/MPG.0000000000002397] [Medline: 31095091]
83. Altuntas N, Kocak M, Akkurt S, Razi HC, Kislal MF. LATCH scores and milk intake in preterm and term infants: A prospective comparative study. Breastfeed Med 2015;10(2):96-101. [doi: 10.1089/bfm.2014.0042] [Medline: 25548967]
84. Boies EG, Vaucher YE. ABM clinical protocol #10: Breastfeeding the late preterm (34-36 6/7 weeks of gestation) and early term infants (37-38 6/7 weeks of gestation), second revision 2016. Breastfeed Med 2016;11:494-500. [doi: 10.1089/bfm.2016.29031.egb] [Medline: 27830934]
85. Dennis C. The breastfeeding self-efficacy scale: Psychometric assessment of the short form. J Obstet Gynecol Neonatal Nurs 2003;32(6):734-744. [doi: 10.1177/0884217503258459] [Medline: 14649593]
86. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24(4):385-396. [Medline: 6668417]
87. Trends in smoking and vaping: New Zealand health survey. Ministry of Health New Zealand. URL: https://www.health.govt.nz/statistics-research/surveys/new-zealand-health-survey/publications/202324-survey-publications/trends-in-smoking-and-vaping [accessed 2025-08-21]
88. Jerzyński T, Stimson GV, Shapiro H, Król G. Estimation of the global number of e-cigarette users in 2020. Harm Reduct J 2021;18(1):109 [FREE Full text] [doi: 10.1186/s12954-021-00556-7] [Medline: 34688284]
89. McBride M, Haile ZT. Association between electronic nicotine delivery systems use and breastfeeding duration. Breastfeed Med 2021;16(11):886-893. [doi: 10.1089/bfm.2021.0132] [Medline: 34297610]
90. Ng JH, Rice KK, Ananth CV, Brandt JS. Attitudes about marijuana use, potential risks, and legalization: A single-center survey of pregnant women. J Matern Fetal Neonatal Med 2022;35(24):4635-4643. [doi: 10.1080/14767058.2020.1858279] [Medline: 33292026]
91. International Organization for Standarization. 1211:2010 Milk - Determination of fat content - Gravimetric method (reference method). Geneva, Switzerland: International Organization for Standarization; 2010. URL: https://cdn.standards.iteh.ai/samples/51348/9269698203534b37aad233d47ed801aa/ISO-1211-2010.pdf [accessed 2026-05-02]
92. 8968-1:2014 Milk and milk products - Determination of nitrogen content - Part 1: Kjeldahl principle and crude protein calculation. International Organization for Standarization.: International Organization for Standarization URL: https://cdn.standards.iteh.ai/samples/61020/3e2cd95323b94be1b5d51dd2f15d0062/ISO-8968-1-2014.pdf [accessed 2026-02-05]
93. Merril AL, Watt BK. Energy value of foods, basis and derivation. Human Nutrition Research Branch, Agricultural Research Service. URL: https://www.scirp.org/reference/referencespapers?referenceid=1769288 [accessed 2026-02-05]
94. 6731:2010 milk, cream and evaporated milk - Determination of total solids content (reference method). International Organization for Standarization. Geneva, Switzerland: International Organization for Standarization; 2010. URL: https://cdn.standards.iteh.ai/samples/56815/f00171b3d7b74f23a396e49ead04f4df/ISO-6731-2010.pdf [accessed 2026-02-05]
95. Haszard JJ, Heath AM, Taylor RW, Bruckner B, Katiforis I, McLean NH, et al. Equations to estimate human milk intake in infants aged 7 to 10 months: Prediction models from a cross-sectional study. Am J Clin Nutr 2024;120(1):102-110 [FREE Full text] [doi: 10.1016/j.ajcnut.2024.04.009] [Medline: 38890036]
96. Wilson Van Voorhis CR, Morgan BL. Understanding power and rules of thumb for determining sample sizes. TQMP 2007;3(2):43-50. [doi: 10.20982/tqmp.03.2.p043]

Abbreviations

D2O: deuterium oxide
D-MER: dysphoric milk ejection reflex
DTM: dose-to-mother
HMA: human milk analyzer

XSL•FO
RenderX

IAEA: International Atomic Energy Agency
IBCLC: international board-certified lactation consultant
MABBS: Māmā and Baby Breastfeeding Study
PIMS: perceived insufficient milk supply
REDCap: Research Electronic Data Capture
WHO: World Health Organization

Edited by A Schwartz; submitted 24.Sep.2025; this is a non–peer-reviewed article;accepted 19.Feb.2026; published 31.Mar.2026.

Please cite as:
Bruckner B, Taylor R, Parata C, Haszard J, Taylor B, Bevin S, Gooding K, Heath AL, Katiforis I, Daniels L
Maternal Perception vs Actual Breast Milk Supply: Protocol for an Observational Cross-Sectional Study
JMIR Res Protoc 2026;15:e84776
URL: https://www.researchprotocols.org/2026/1/e84776
doi:10.2196/84776
PMID:

©Bailey Bruckner, Rachael Taylor, Carmen Parata, Jillian Haszard, Barry Taylor, Samantha Bevin, Kassidy Gooding, Anne-Louise Heath, Ioanna Katiforis, Lisa Daniels. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Taunakitanga Takitini, Reframing Self-Management Support for All in Aotearoa New Zealand: Protocol for a Participatory Case Study Program of Research

Leigh Hale[1*], PhD; Bernadette Huatau Jones[2*], MPH; Ofa Dewes[3*], PhD; Shirley Keown[4*], PhD; Brigit Mirfin-Veitch[5*], PhD; Amanda Wilkinson[1*], PhD; Donna Keen[1*], MTour; Pauline Norris[6*], PhD; Eileen McKinlay[7*], MA (App); Meredith Perry[1*], PhD; Kate Diesfeld[8*], PhD; Henrietta Trip[9*], PhD; Tristram Ingham[2*], MBChB; Nancy Longnecker[10*], PhD; Roberta Francis Watene[5*], PhD

[1]School of Physiotherapy, University of Otago, Dunedin, Otago, New Zealand

[2]University of Otago Wellington, Department of Medicine, University of Otago, Wellington, New Zealand

[3]The Cause Collective, Auckland, New Zealand

[4]Te Hauora a Tūranganui a Kiwa, Tūranga Health, Gisborne, New Zealand

[5]Donald Beasley Institute, Dunedin, New Zealand

[6]Va'a o Tautai – Centre for Pacific Health, University of Otago, Dunedin, New Zealand

[7]University of Otago Wellington, Centre for Interprofessional Education, University of Otago, Wellington, New Zealand

[8]School of Community and Public Health, Auckland University of Technology, Auckland, New Zealand

[9]Department of Nursing, University of Otago, Christchurch, New Zealand

[10]Department of Botany, University of Otago, Dunedin, New Zealand

[*]all authors contributed equally

Corresponding Author:
Leigh Hale, PhD
School of Physiotherapy
University of Otago
School of Physiotherapy
325 Great King Street
Dunedin, Otago, 9054
New Zealand
Phone: 64 34795425
Email: leigh.hale@otago.ac.nz

Abstract

Background: Increasingly, people are living with multimorbidity and long-term conditions or permanent impairments, which contribute substantially to health loss and additional health inequity. Critical to managing this health crisis is the appropriate support provided by health and social services. International research has informed the implementation of supported self-management knowledge into Aotearoa New Zealand. Meagre attention has been given to how New Zealand health care organizations can appropriately support people to self-manage their lifelong conditions within their own life contexts and cultures. Currently in New Zealand, those experiencing the greatest health inequities are Māori and Pacific peoples—Tāngata Whaikaha Māori (Māori living with disability) and people with learning (intellectual) disability who live with long-term health conditions or permanent impairments. This research aims to challenge and reframe supported self-management in New Zealand to enable our underserved populations and their whānau (families) to receive appropriate support to live well.

Objective: We aim to reconceptualize supported self-management in New Zealand as a holistic approach to the provision of support and care within the complexities of the lived human and whānau experience. In total, three objectives will be considered across 3 key populations of people with long-term conditions or permanent impairments (Māori, Pacific Peoples, and those with learning disability) to (1) understand our community partners' and their communities' aspirations for living well and how best to support these desires, (2) develop innovative models of support by working alongside communities to enable supported self-management within their own context, and (3) implement models and evaluate outcomes.

Methods: Embracing social construction of reality, this participatory case study program uses mixed methods and implementation evaluation design and is underpinned by Whānau Tuatahi (Māori) and Fonua (Tongan) research models and the Transformative

XSL•FO
RenderX

Paradigm. Moreover, 3 case studies, 1 for each population group, will apply the same program objectives. Objectives 1 and 2 will be addressed with qualitative methodologies underpinned by relevant participatory designs. Objective 3 will use appropriate implementation frameworks.

Results: Funded in October 2023, we have completed 2 years of this 5-year program grant. These first 2 years were focused on relationship building, ethical applications, research capability, and capacity building. Substantial, progressive consultation with the respective communities of each case study was undertaken.

Conclusions: New knowledge generated across our program has the potential to inform New Zealand policy and practice about service delivery acceptable to the people to whom it matters (particularly Māori, Pacific Peoples, and people with a learning disability) and places emphasis on well-being promotion. This approach focuses on the inherent strengths and abilities of people, rather than the deficits or problems, acknowledging the wealth of expertise and experience people living with long-term conditions or permanent impairments bring, and builds relationships and partnerships between people and health care partners.

International Registered Report Identifier (IRRID): PRR1-10.2196/89658

(JMIR Res Protoc 2026;15:e89658) doi:10.2196/89658

KEYWORDS

self-management; support self-management; Māori People; Pacific Island People; learning disability; disability; wellbeing

Introduction

Background

This paper describes the overarching protocol for a program of research aimed at building Aotearoa New Zealand–specific knowledge regarding supported "self"-management in the context of long-term health conditions or permanent disabilities. Intentionally, the research focuses on underserved populations, namely Māori and Pacific Peoples, Tāngata Whaikaha Māori (Māori with lived experience of disability) and people with a learning (intellectual) disability and their whānau (extended family) and how they can be optimally supported to live the best life possible. In this research, we seek to disrupt current thinking and understandings of supported self-management and reframe the concept specifically for the New Zealand context.

Long-Term Conditions and Self-Management

Living with a long-term health condition or impairment (hereon collectively referred to as a long-term condition) impacts an individual personally (physically, mentally, emotionally, spiritually, and economically), their family (whānau), and the wider society and economy [1-3].

The impact of long-term conditions on the health of populations, and the associated costs of these, is an important determinant of public policy and public spending [4]. Underlying the interest in and recognition of the impact and cost of long-term conditions are neoliberal discourses [5], resulting in attempts to reduce the burden and attached societal costs arising from long-term conditions [6]. Thus, "defining and supporting patient self-management [has become] an important task of health services. Self-management [is now] a "policy relevant" construct, clearly within the remit of the health system and [therefore] one of the daily tasks of patients and health professionals in their encounters" [6]. Greaney and Flaherty [7], however, propose that

> *[A]an under resourced self-management agenda is arguably a means of covert rationing and morally unjust. When self-care is not performed by the individual, the health service does not always replace or supplement self-care activities … incomplete self-care becomes care left undone and renders the already vulnerable individual at risk.*

Such discourses require critical scrutiny, given the demands that increasing numbers of people living with multimorbidity and long-term conditions will place on the New Zealand health system [7]. The prevalence and cost of long-term conditions are high [8,9], plus relatively higher for Māori and Pacific ethnic groups, older people, and those who face socioeconomic disadvantage [10].

Latest information on the Te Whatu Ora (Health New Zealand) website [11] refers to the New Zealand Ministry of Health's recently developed suite of resources and tools for primary care teams, which provide information to build understanding of self-management support and how to embed this in primary care. This website then provides a link to resources located on the Healthify website (a New Zealand website run by a charitable trust providing evidence-based accessible health information and self-help resources) [12]. Here, self-management is defined as "taking charge of your own health and health-related decisions." The reported key points about self-management and self-management support are (1) having enough knowledge and confidence to be actively involved in what happens to you, and how your health conditions are treated or managed; (2) understanding your conditions and the medicines you are taking and why. This involves working with your health care provider to set goals to achieve what you want for your own well-being and quality of life; and (3) it enables you to make informed choices about your health care and to live as well as you can with your health conditions. While this information suggests endorsement of the neoliberal approach to health previously referred to, it also infers that support is required for people to manage their health. Indeed, international research considers support provided to a person or persons to people to manage their health and wellness as a crucial component to enabling well-being [13].

Supported Self-Management

As our genesis in this research, we refer to "supported self-management." By this, we currently mean how health care

XSL•FO
RenderX

teams and systems collaborate with an individual and their whānau to support their self-management endeavors, and the actions and informed decisions they take to care for and maintain their own or their loved one's health and well-being [14]. This focus on the crucial support from others required by people in their efforts to live well contrasts with the more neoliberal rhetoric of "self-management," a term engendering philosophical principles of self-responsibility, independence, autonomy, and self-determination, as generally endorsed by health systems [14-16]. Supported self-management applied with holistic support and care has the potential to enable flourishing mauri (the essential quality and vitality of a being) and reduce health inequity in New Zealand, particularly for those populations experiencing health disparities or living precarious lives [17-19].

The concept of self-management has had extensive global research, and numerous self-management programs have been developed [20-22]. The benefits of these programs are evidenced [13], but participation by those living in lower socioeconomic situations and indigenous cultures is low [23]. Barriers to participation include multistep referrals and uptake processes, environment and financial barriers, and discord between individual health beliefs and program content, which focus on symptom management and patient activation, and not on what matters to the person [24,25].

The term "self" implies individual responsibility and personal agency, a perception criticized for diminishing the need for social responsibility and collective action to support people with long-term conditions [15,26]. Furthermore, self-management is considered a Western construct necessitating adaptation for indigenous and ethnic minority groups [26]. Evidence now supports a capabilities-perspective approach to support self-management by developing a collaborative, supportive, and enabling relationship between the person with the long-term condition, their whānau and communities, their health care professionals or organizations, and the wider public health system [26-28]. Some good examples of such a holistic approach exist within primary health care in New Zealand [29-32]; however, supported self-management per se largely occurs within communities. Few primary care services or general practices have sufficient or prioritized resources to provide adequate holistic support or preventative habilitation to achieve and maintain well-being, resulting in intergenerational cycles of multimorbidity and fatalism within whānau [8].

From a rights-based lens, at the individual, family or whānau, and community level, what does this mean and what would be the optimal support required to enable a well life? Addressing this question lies at the core of our program of research.

Community-Based Research

Answering our research question necessitates community engagement and involvement. Communities are important; they directly experience local challenges and develop local solutions to address health inequalities [29]. Health care in New Zealand in recent years has rapidly transformed and has strengthened the importance of community involvement. The New Zealand health reforms in 2022, driven by the health and disability inequities evidenced in the Waitangi Tribunal Report 2019 (WAI 2575) [33] and outlined in the Pae Ora (Healthy Futures) Bill in 2022 [34], presented new community-level health opportunities. An example was the creation of Te Aka Whai Ora, a government agency established in 2022 to lead and improve Māori health outcomes, responsible for commissioning Māori health services and monitoring the overall system for equity. Of note, it was disestablished in 2024. Other examples include the xrevised Ministry for Pacific Peoples, and Whaikaha – Ministry of Disabled People. Recently, the latter was significantly downsized in terms of responsibilities and resources [35]. While the COVID-19 pandemic in 2020-2022 exposed areas of health and social inequities and failings in accessibility of health care delivery, it also identified local community solutions to these challenges that informed future health service delivery. Māori, Pacific, and health and disability services directly experienced these health care changes and were forced to adjust their care delivery accordingly. These services thereby became experienced in identifying strength-based solutions. Furthermore, as these services are embedded within their communities, they have trusting relationships with and comprehensive knowledge of the people they serve.

Valuing the authority of the experiential knowledge of community-based Māori, Pacific, and health and disability services goes beyond community engagement and involvement to address research questions and calls for an approach underpinned by participatory action research [36]. Providing a research platform based on the principles of participatory action research was recognized as the most appropriate and effective approach to supporting community health services to engage with their communities. This type of methodology would enable these services to construct what they perceive is needed to best inform and support the health and well-being of their community members living with long-term conditions. From such an approach, innovative community-led solutions and advocacy can be developed [37]. These solutions can provide transformative health delivery representative of the priorities and values of health and social services and the communities they support.

Upholding New Zealand's Te Tiriti o Waitangi

Research undertaken in New Zealand must uphold Te Tiriti o Waitangi (New Zealand's founding document) and is guided by Vision Mātauranga (a policy framework designed to unlock the innovation potential of Māori knowledge, resources, and people). Mātauranga Māori refers to Māori knowledge, ways of knowing and knowledge generation practices, and is a broad system that encompasses time, space, place, and discipline; it is thus a knowledge-generating system. This knowledge-generation should be done with care. As described by Mercier [37], it should

> *…address a problem of shared concern so that there can be equal input from contributors; cannot be solved by one knowledge system alone; have equitable outcomes; build capability and capacity; are underpinned by Treaty principles such as protection and partnership; have Māori in leadership roles; and crucially, are injected with human values of honesty, truth-seeking, kindness, generosity and humility.*

XSL•FO
RenderX

Thus, in researching alongside community-based partners, mindful observation of the 5 principles of Te Tiriti o Waitangi (tino rangatiratanga [self-determination], equity, active protection, partnership, and options) and Te Mana Raraunga (Māori data sovereignty guidelines) [38] is critically important.

Our research program is a collaborative partnership between community partners and academic researchers who will provide the research platform for guidance, support, and mentorship as requested by our partners. Upholding rangatiratanga, the community partners have developed their own studies within this program of research and hold sovereignty, autonomy, control, and independence over their study's aims, methodologies, processes, and data.

This program's kaupapa (principles and ideas which act as a base or foundation for action) began with a clean slate agenda in 2019 at an in-person hui (meeting) of researchers and community partners who all held existing relationships and concurring philosophies about supporting self-management for health and well-being. The research evolved over regular and numerous in-person partner visits and Zoom (Zoom Communications, Inc) videoconferencing hui. Consequently, the core of the program is focused on 3 community partners—a Māori Health Provider, a Pacific community organization, and Ngā Tāngata Tuatahi, People First New Zealand (People First), supported by a national, independent, nonprofit organization which undertakes disability research and education.

During the early stages, we desired an appropriate program title to reflect its kaupapa. The program's Māori researchers and advisors met to discuss. Included in this hui was a rangatahi (youth; Kāi Tahu, Ngāpuhi) raised in Kōhanga Reo and Kura Kaupapa Māori who suggested "Taunakitanga Takitini: reframing self-management support for all in Aotearoa" as a formal title, along with the whakatuaki (proverb) "E hara taku toa I te toa takitahi, ēngari he toa takitini" ("My successes are not mine alone, they are ours, the greatest successes we will have are from working together"). Later, the Kaumātua (elder) of the Māori health provider endorsed this title.

Research Program's Purpose

Our program endeavors to build mātauranga Māori and Pacific and disability knowledge and understanding of supported self-management for people living with long-term conditions within the context of New Zealand. As supported self-management is a human right [39], enhancing how the health system provides such support will, by respecting and enabling people or group-led solutions relevant to their context, contribute to reducing health inequities in New Zealand. For this, we must continue to challenge and evolve conventions of "self"-management, adopting holistic, culturally appropriate, and disability-centered approaches focused on people and whānau priorities and solutions.

Research Program Aims

This program of research aims to reconceptualize "supported 'self'-management" in New Zealand as a holistic and evolving approach to the provision of support and health care within the complexities of lived human and whānau experiences. The following specific objectives will be applied across three key populations of people living with long-term conditions, (1) Māori, (2) Pacific peoples, and (3) Tāngata Whaikaha Māori and people with learning (intellectual) disability to, first, understand the program's community partners' and their communities' needs and aspirations for living well and how best to support people living with long-term conditions or disability. Second, develop innovative models of how health and social services can support and work alongside whānau and communities to enable supported "self"-management within their own context. And third, implement these models and evaluate outcomes.

More specific program research questions are presented in Textbox 1.

Textbox 1. Specific program research questions.

- What are the aspirations and challenges to living well of Māori, Pacific peoples, Tāngata Whaikaha Māori and people with a learning (intellectual) disability living with long-term health conditions?
- How do these populations think health care services can best support them?
- What are the aspirations and challenges of the health, social, and advocacy organizations who currently support these populations?
- How can community health, kaiāwhina (support workers), and allied and nursing health care professionals best provide support to enable these aspirations?
- What can we learn from our community-based health care research partners who already have developed innovative holistic models of care for their service users, especially during these current transitions in health delivery?
- Together with our research health care and advocacy partners can we refine these models of health care or develop new ones to meet the aspirations of our targeted populations group?
- How can we measure the impacts of these refined or new models of health care?
- How beneficial are these refined or new models of health care?
- How can community providers use data sovereignty guidelines to safeguard the valuable data collected from their communities?
- And thus, what does effective "supported 'self'-management" look like in New Zealand?

XSL•FO
RenderX

Methods

Overall Philosophical Approach

This research takes a sociocritical approach, applying a critical social construction worldview within the Transformative Paradigm [40]. The Transformative Paradigm combines diverse philosophical strands associated with feminism, critical theory, Indigenous and postcolonial theories, and disability and deafness rights theories [36]. Thus, this program of research is based on a participatory case study approach [41] that applies a mixed method design [42] and implementation evaluation [43].

Our Methodological Approach

Given our philosophical approach, our methodological stance goes beyond cocreation or co-design with our 3 community partners to that akin to a participatory approach. Notwithstanding the lack of explicit definitions of cocreation or co-design in health care [44,45], these 2 methodologies imply an equal partnership between the communities concerned and the researchers. In our work, the 3 community partners have control and sovereignty of their own studies, including how their data are collected, analyzed, and used. The research team members' roles are to support and facilitate the process; prepared to be reflective, flexible, accept uncertainties and pivot as required [36].

Thus, we are taking a participatory case study approach based on the social construction of reality [41]. Our program comprises 3 discrete case studies, representing our 3 targeted population groups, each sharing the same 3 program objectives. We will leverage the strengths of each case study to showcase what best practice and integrated support can look like as examples of New Zealand evidence-informed practice ("homegrown") rather than imported initiatives and models that are not contextually relevant or understood. Our community partners' current practices have been developed from working with their communities within their lived contextual experience. Our program of research is designed to generate a deeper understanding of what our partners do, what their communities aspire to, and what they want the health system to respond to. This information will enable exploration of existing gaps or successful strategies for the achievement of equitable health outcomes. This interconnecting and intersectional research will foster a New Zealand–focused community of practice, strengthening knowledge transfer between academics, partners, and government agencies, and with traditional health care delivery models (eg, general practice), acknowledging deficits of a "one size fits all" approach for good health outcomes.

While overall the research has as its foundation the Transformative Paradigm [40], each case study is underpinned by appropriate methodological approaches. For case study 1, this is Kaupapa Māori Research [46] and the holistic health model of Te Wheke (Māori) [47-49]. The Fonua health model [50] underpins case study 2. Case study 3 is based on the Transformative Paradigm. Using holistic models will enable us to explore the multifaceted and interactive components of mauri ora (healthy individual), whānau ora (healthy families), and wai ora (healthy environments; a collective Tongan concept is mo'ui lelei [51]) that determine or contribute to supporting self-management with the underlying intent of increasing social transformation.

Ensuring our Research Is Culturally Safe

For our overall program of work, we are guided by the Whānau Tuatahi Research principles to ensure culturally safe research processes and effective outcomes for all involved [46]. This framework enables a practical adaptation and application of Westernized research methods within a Kaupapa Māori methodology and is formed by the following concepts—whakawhirinaki (trust), whakawhanaungatanga (building relationships), whakamana (empowerment), ngāwari (flexibility), utu (reciprocity), and hurihuringa (reflexivity).

Our Conceptual Method

These concepts are woven through each of our 3 studies, providing a supportive whariki (mat). Figure 1 depicts an unfinished woven mat, a traditional art form creating functional implements for both Māori and Pacific Peoples. While most of the mat is woven and supportive, acknowledging the foundational strengths and knowledge that currently exist in our community partner organizations, the unfinished sections represent gaps in knowledge for which strands still need to be woven to provide more synergistic and supportive linkages to enhance resilience for our 3 population groups.

XSL•FO
RenderX

Figure 1. Taunakitanga Takitini (commissioned and created by Emily Fruean from Si'usega, Samoa).



Our research program endeavors to progress toward completion of this mat by tightening the intersectional weave. While traditionally used as a ground mat, woven mats are also used as sails in traditional boats. The turbulent seas of the transforming health services and the pandemic challenges require a strongly woven mat that ensures we sail and land together safely, especially for oppressed and underserved populations with long-term conditions.

Case Study 1: Kaupapa Māori Research

Kaupapa Māori Research methodology underpins much of this program of research. It is a transformative methodology that is conducted by, with, and for Māori to benefit the Indigenous Māori people of Aotearoa New Zealand [52]. It is grounded in Māori philosophical principles and seeks to address the negative impact of Western research practices that have historically misrepresented and exploited Indigenous peoples [53]. Kaupapa Māori Research seeks to give voice to Indigenous Māori who have been historically silenced, particularly those most negatively impacted, such as disabled Māori and women's subgroups [54]. This methodology promotes Māori self-determination over the entire research process with the focus on privileging the voice of Māori and supporting mana-enhancing outcomes [52].

A range of Māori models of health will be drawn on throughout this program, including Te Wheke. The Te Wheke model is a holistic Māori health and well-being framework that uses the wheke (octopus) as its central metaphor [47]. The head of the octopus represents the whānau, emphasizing its central role in supporting health and well-being. The eyes of the octopus symbolize waiora (total well-being for both the individual and the whānau). Each of the 8 tentacles represents a specific and interconnected element of health—wairuatanga (spirituality), hinengaro (the mind), taha tinana (physical), whanaungatanga (extended family), mauri (the life force), mana ake (unique identity), hā a koro ma, a kui ma (the breath of life from ancestors), and whatumanawa (the expression of emotion). The interwoven nature of the tentacles signifies that these dimensions are deeply connected, acknowledging that health is an integrated concept in Te Ao Māori [the Māori world]. This holistic approach contrasts with Western biomedical models by prioritizing the link between the physical, the mind, spirit, whānau, and the wider world [47].

Interpretation and analysis of data generated from this program will also be cognizant of holistic Māori models of health, well-being, and disability that privilege the voices of participants in this program of research. For example, Te Pae Mahutonga, a Te Ao Māori health promotion framework, was used to interrogate health and disability outcomes for disabled Māori and thereby inform equity in health policy and services [55].

Māori models of health and well-being are highly relevant for this program of research because they are collective in nature and prioritize the interconnectedness of individuals within their social and environmental contexts. These models encompass critical cultural practices and concepts, including tikanga (protocols), te reo Māori (the Māori language), mātauranga (knowledge), whakapapa (genealogy), and social connectedness. They are flexible in their application and can also be adapted to encompass and respect other cultures and collective communities, such as disabled groups and Pacific Peoples.

XSL•FO
RenderX

While existing models can be a useful guide, this research aims to explore the extent to which these models meet the needs and aspirations of the targeted populations and whether we need to refine or develop new models to deliver maximum benefits to our health care delivery.

Case Study 2: The Fonua Health Model

The Pacific component of our research program is characterized by (1) leadership-driven (because of the mana-enhancing social structures in Pacific families and communities), (2) systematic engagement of relevant stakeholders (because of the need to involve our stakeholders in the change process to achieve wide coverage and sustainability), (3) transferable processes (to other populations living in persistent disadvantage), and (4) culturally mediated and community-delivered (because of the pivotal role that Pacific communities have on cultural norms).

Our cultural approach is to ensure the voices of Pacific communities, especially those living with long-term conditions and the services that currently support them, are listened to, understood, respected, and assisted as to how best to reframe supported self-management from their perspective. Active listening will build confidence, courage, and trust between participants, the services, and researchers, resulting in mutual understanding and meaningful engagement and communication. Moreover, it will encourage and activate creative thinking about new ways of providing supported self-management that will enable people living with long-term conditions to live their lives in the best possible way. Bilingual talanoa (a process of inclusive, participatory, and transparent dialogue) will align with Pacific and New Zealand Health Research Guidelines [56] on communal relationships, cultural sensitivity and reciprocity, and the Fonua health model.

The Fonua model is a Tongan framework comprising 5 dimensions of life that are interwoven like a mat and interdependent and complementary of each other [49]. The 5 dimensions are Laumalie (spiritual), ‘Atamai (mental), Sino (physical), Kāinga (collective or community), and ‘Atakai (environment). Within these 5 dimensions are five structural levels that are (1) Taautaha (individual), (2) Famili or Kāinga (family), (3) Kolo (local community or village), (4) Fonua (nation), and (5) Mamani (global society) [50]. Thus, our laakanga [weaving] phases of research will be guided by the following processes.

1. Kumi Fonua (consultation): We will establish connections with our community leaders and key stakeholders face-to-face. Talanoa will comprise information to guide the cocreation of a quality framework for interaction during this activity. More specifically, we will seek to understand what supported self-management might look like for people living with long-term conditions, with a view to then creating a model of health care that will best develop, deliver, maintain, and support a community-wide generic program of supported self-management.

2. Langa Fonua (data collection): We will ascertain what data should be collected and analyzed. Our design will be augmented by a community-based social movement framework that will include people living with long-term conditions and their primary caregivers who live with them.

3. Tauhi Fonua (data analysis): Data will be analyzed for common themes, discussed, and refined. Reflections on the knowledge generated will enable deeper understanding. Furthermore, the notion of sustainability, along with the roles of support and community workers, and health care professionals in this support, will be explored and made explicit. The outcome will be a refined or new model of supported self-management, with decisions made on the evaluation methods for the new model. Final analysis will include recommendations on how the results may be delivered for implementation and evaluation.

4. Tufunga Fonua (dissemination or reformation): Final analyses will include recommendations on how the results and the new model of supported self-management may be delivered for implementation and evaluation. This will include feedback to stakeholders, presentations at community fonos (meetings) and conferences, publication via local media networks (print and radio), and scientific journals.

Case Study 3: The Transformative Paradigm

Case study 3 focuses on people with a learning (intellectual) disability. The relationally based, inclusive approach of the transformative paradigm was considered appropriate as researchers and the community (participants) collaborate on an equal basis. The transformative paradigm has a social justice orientation and advocacy for marginalized voices, and it aims to achieve social change through research [40]. This framework purposively includes the “hidden” or seldom-heard voices, and contests aspects of power and privilege to promote social justice. Crucial to this approach is the building of relationships of trust and the valuing of lived expertise, not only from the standpoint of lived experience, but of acknowledging leadership, self-determination, and appropriate remuneration, and the reshaping of the power imbalances [40,57,58].

The researcher in the Transformative Paradigm is a disruptor for social justice who challenges inequalities while assuming shared responsibility and acting with humility [59]. The central tenet of the Transformative Paradigm is that power must be attended to at each stage of the research process; Mertens [40] challenges researchers to ethically examine who they are and who they are in relation to the community in which they are working. Relationship-based research practice underpins the Transformative Paradigm, ensuring authentic and caring relationships between the researchers and those with lived-expertise in participatory-based research [60,61]. The Transformative Paradigm provides a strong theoretical basis for the program’s study involving people with learning disabilities.

Individual Backgrounds to the 3 Case Studies

Overview

The background to each case study is described below. The detailed methodologies and methods each study will take will be dependent on their initial and ongoing consultations with their communities and will evolve with time, and thus, these

XSL•FO
RenderX

details will be presented in later disseminations as determined by each community partner involved.

Case Study 1: Hikoi Kite Oranga Tangata Katoa – Pakeke Matihiko

Turanga Health, a long-established and trusted iwi health provider in Tairāwhiti [62], will undertake the Māori-specific case study.

We were cognizant that supported self-management needs to be flexible and responsive to the sociocultural contexts of different people and cultures. Within Te Ao Māori, the concept of whānau encapsulates the whānau whānui (the broader or extended family and friends). The premise of supported self-management's focus on "self" appears incongruent in a Kaupapa Māori context [63]; and maybe better defined as "whānau-support." However, in current times, does this assumption of whānau-support hold true? Do time-poverty, financial stresses, urbanization, and the need for more whānau members to be in employment reduce the level of support that whānau experience and provide? Turanga Health considered that a better understanding of these aspects for different people would inform them (and other providers) as to how they could better support whānau wellness and overall well-being.

Working toward hauora (well-being) is a journey, and often a long journey for many people living with a long-term condition. Some whānau have a large and supportive whānau network able to engage, awhi (support), and tautoko (endorse) their whānau on their journey. Some whānau live without any support, while others have family members who, for various reasons, may not be able to provide the necessary or desired assistance (eg, due to work commitments, lack of space, or geographical distance). The uniqueness of these journeys was amplified during the COVID-19 pandemic. This study provides an opportunity for mana Motuhake (whānau determining their "what and how" for their hauora journey), rather than a "Westernized" model of supported self-management imposed on whānau.

The COVID-19 experience and the recent health reforms have caused reflection on current models of health care and how these could be modified based on recent experiences and learnings, and the future directions of health care. Turanga Health's care is underpinned by the fundamental belief that understanding and being connected to the past are important for both the present and the future, as reflected in the whakatauki (proverb)—"*Turanga mua, Turanga tika, Turanga muri, Turanga he*." (*Those who move forward and accept the challenge will bear the fruits of success. Those who do not advance will not develop*.)

Case Study 2: Lūsia Ki Taulanga

Lusia ki taulanga is a Tongan proverb referring to seafarers' struggle through stormy weather to reach land, striving to achieve despite life's challenges.

A Pacific case study will be undertaken in collaboration with The Cause Collective (TCC). TCC is a Pacific community organization in South Auckland "that aims to create the conditions for communities to thrive from this generation to the next." TCC is interested in the overall outcomes of Pacific Peoples across Aotearoa New Zealand, underserved populations, and South Auckland communities [64]. With a vision that "This Generation will Win the Next Generation," TCC is poised to accelerate the pace of change by co-designing, testing, and prototyping scalable breakthrough solutions to complex problems and achieve its goal of "healthy and thriving families and communities." To get to the heart of pervasive and complex problems, TCC uses social innovation tools, critical thinking, and cultural perspectives.

TCC is also the only Pacific-led primary health care organization in New Zealand. Established in 2010, TCC is passionate about improving health equity and well-being for Pacific Peoples and other persistently disadvantaged populations. TCC provides tailored tools, guidance, and support to their network of 46 General Practices with funding, contract management, clinical and administrative tasks, education, business and clinical modeling, and service delivery and reporting so that they can achieve equitable outcomes for their >140,000 enrolled patients and local communities.

Under the umbrella of TCC, Tātou Collective was appointed by the government in July 2025 as the only Whānau Ora commissioning agency for Pacific Peoples in Aotearoa. Whānau Ora is a culturally based and whānau-centered approach to well-being focused on whānau (family group) as a whole, as the decision-makers who determine their goals and aspirations [65]. As an independent Pacific agency, Tātou Collective is focused on getting support to those with the greatest need, guided by transparency, integrity, and the best available evidence.

Aligning with Pacific and New Zealand Health Research Guidelines and bilingual Talanoa [56], the Tongan Fonua [50] research approach will be implemented. Fonua is a cyclic, dynamic model of interdependent relationships (va) between humanity and its ecology for the ultimate purpose of harmony in health and well-being. Fonua's values are Fe'ofa'ofani (love), Fetokoni'aki (reciprocity), Faka'apa'apa aki (respect), and Fakapotopoto (wisdom and prudence) [50,56]. This study will draw on the strengths of Pacific Peoples, in which capabilities and knowledge abound. The Pacific capabilities lens suggests that, despite lower educational levels overall, Pacific Peoples, as a group, hold a wealth of knowledge, values, and experiences [49]. Gifting this to people living with a long-term condition is a resource and a means of empowering pathways of future-thinking and transformational impact.

Case Study 3: Supported Decision-Making to Increase and Enhance Supported Self-Management With People With Learning (Intellectual) Disability

Our community partners in this case study are an independent disability research organization, the Donald Beasley Institute [66] and Ngā Tāngata Tuatahi – People First New Zealand (People First) [67]. People First is the only learning disability disabled person's organization (DPO) in New Zealand. DPOs have special status in New Zealand as organizations and are defined as a DPO if they are exclusively governed by and for the group of disabled people they represent. People First

XSL•FO
RenderX

endorsed both the wider program, and case study 3, which is specifically concerned with and involves people with a learning disability. They are centrally involved in the research, including nominating People First members for a research advisory group, providing direction and advice throughout, and engaging in co-design and coproduction activities at all phases of the study, including the development, delivery, and implementation of the intervention.

People with learning (intellectual) disabilities comprise about 2% of New Zealand's population [68], and experience poor health outcomes [68] mostly attributable to inequities in access to health information, services, and support [70]. People with learning disabilities are frequently perceived to lack the capacity to make decisions, including health-related decisions [71]. International research has also confirmed that people with learning disabilities experience inequities in accessing supported self-management programs [72]. This inequity is seen as a direct consequence of disability and health professionals perceiving people with learning disabilities as not having the capacity to receive or apply health information, or to make competent health-related decisions.

The presumed and conditioned incapacity experienced by people with learning disabilities is a significant barrier to supported self-management initiatives. Despite this, research has shown that people with learning disabilities can intentionally engage in supported self-management initiatives with successful outcomes [73-75]. Disability and health care professionals could facilitate supported self-management endeavors with individuals and their whānau (or other close supporters) [76]. A potential strategy to facilitate supported self-management is supported decision-making (SDM). SDM is a framework whereby disabled people with cognitive impairment are supported to make decisions that exercise their legal capacity and promote their rights, will, and preferences [77-80]. In prioritizing individual's will and preference, SDM ensures disabled people remain central to decisions about them.

SDM challenges a "best interests" approach whereby decisions, including health decisions, are made by others according to their perceptions of what is "best" for the disabled person. There is an obvious synergy between supported self-management and SDM. Recent research in the learning disability and supported self-management arena points both directly [79] and indirectly [80] to the potential of SDM, but it is yet to be comprehensively studied. This learning disability–specific case study is designed to explore the potential of SDM to increase and enhance access to supported self-management with people with learning disabilities, inclusive of tāngata whaikaha with learning disabilities. Ensuring supported self-management is available to all as a human right meets aspirations of Te Tiriti o Waitangi and the United Nations Convention on the Rights of Persons with Disabilities.

Overall Methods and Designs

Overview

This is a participatory case study research program [41] with an embedded QUAL quant mixed method [42] and implementation evaluation design [43]. Our 3 discrete case studies share the same 3 program objectives and are broadly structured around a similar phased design. This enacts our metaphor of the sail as a weave in process, but recognizing the need to first consider our threads and materials (objectives 1 and 2), then collaboratively weave together those threads, and testing them for strength and resilience (objective 3), with the hope that they will see us through to the next step of our journey.

This weaving metaphor operates within our guiding Whānau Tuatahi Research principles that ensure culturally safe research processes and effective outcomes for all involved [46]. This framework enables a practical adaptation and application of Westernized research methods within a Kaupapa Māori methodology and, as noted earlier, is formed by the following concepts: whakawhirinaki (trust), whakawhanaungatanga (building relationships), whakamana (empowerment), ngāwari (flexibility), utu (reciprocity), and hurihuringa (reflexivity). These concepts underpin our interactions with one another in our participatory approach to this research. These concepts are strong threads that underpin our 3 case studies and who we support through the mechanisms of hui, newsletters, strong communication, open leadership, advisory groups, and developing potential and capacity.

To address objectives 1 and 2, in each case study, the community partner will consult and explore their service users' and communities' perceptions, needs, aspirations, and desired health outcomes using qualitative methodological approaches, appropriate to the community in question. Based on these findings, each case study will then refine or develop health care models that can more appropriately support the aspirations of their people living with long-term conditions to live well.

To address objective 3, each case study will implement its developed or refined model and evaluate outcomes over 1 year. Given the research's kaupapa (guiding principles), we were unable to predetermine which research methodologies and methods would work best in each study to achieve this third objective. While all 3 case studies will broadly follow the implementation process described below, each study will have a nuanced approach. Based on guidance from extensive community consultation to date, the following description for addressing objective 3 provides a suggested method, acknowledging that flexibility and agility will be required to be responsive to each case study's communities' directions.

Design

We followed the 12-step process proposed by Hudon et al [41] for conducting case studies with a participatory case study approach. To enact step 6, choosing a theoretical framework to guide data collection and analysis, we chose implementation science methodology as (1) the health environment, systems, and delivery are changing too rapidly to justify lengthy randomized controlled trials and (2) our previous experiences of randomized controlled trials in collectivism-based populations has identified the aversion and mistrust these communities have in such a positivist, individual-focused scientific enquiry [81].

As a starting point, we chose the RE-AIM (reach, effectiveness, adoption, implementation, and maintenance) model as a framework to evaluate site community impact [43,82]. The

XSL•FO
RenderX

RE-AIM model is a robust, frequently used starting point for evaluation [83]. The RE-AIM dimensions ensure a holistic evaluation of health care interventions based on both qualitative and quantitative outcomes. Ways to evaluate these outcomes will be modified depending on each community's wishes and requirements. Using a quasi-experimental design with each case study as a control group is not feasible given the diverse nature of the partners' services.

Participants and Recruitment

While these details will be dictated by the desires of our community partners (finalized while addressing objectives 1 and 2), each implementation study will most likely include up to 20 participants who experience long-term impairments or conditions and their whānau. This sample size was dictated by (1) previous studies using a participatory case study approach with people with long-term complex health needs, which had sample sizes ranging from 10 [84] to 30 [85] participants, and (2) the size and capacity of our 3 case study community partners. Appropriate ethical procedures will be followed to invite participation and gain informed consent.

Intervention and Outcomes

The refined or new model of health care will be implemented within each community. As the program of research aims to reconceptualize "supported 'self'-management" in New Zealand as a holistic approach to the provision of support and care to enable people with long-term conditions to live well, the primary outcomes of well-being and vitality will be evaluated using qualitative longitudinal research [86,87] with qualitative data collected longitudinally every 3 months over a 1-year period. We will also simultaneously collect quantitative data; the quantitative outcome measures to be used will be identified based on community consultation outcomes. Thus, for each case study, 80 individual semistructured interviews are likely to be undertaken; however, this is participatory research and the actual methods of data collection, as detailed in step 8 of the approach proposed by Hudon et al [41], will be finalized by our community partners during their consultation stages. The quantitative data collected via questionnaires will be contextually applied, as the small sample sizes will likely preclude comparative analysis.

For example, case study 1 (*Hikoi kite oranga tangata katoa – Pakeke matihiko*) may use qualitative data gathered via audio-recorded focus groups and semistructured interviews, photovoice, and pūrākau (storytelling), underpinned by a Kaupapa Māori Research paradigm and quantitative questionnaires, such as Hua Oranga (A Māori Measure of Mental Health Outcome) and Whānau Capacities [88].

Case study 2 (*Lūsia ki taulanga*) may use qualitative data gathered by the Pacific oral tradition of Talanoa, likely to be via audio-recorded focus groups and semistructured interviews, photovoice, and digital storytelling workshops [89]. Quantitative data may be collected via the Pacific Identity and Wellbeing Scale [90-92]. A culturally appropriate self-report measure assessing 5 factors (perceived familial well-being, perceived societal well-being, pacific connectedness and belonging, religious centrality and embeddedness, and group membership evaluation) of Pacific identity and well-being. Face, content, and construct validity are established for this measure, and exploratory and confirmatory factor analyses support the model.

For both case studies 1 and 2, other questionnaires that may be used include the World Health Organization-Five Well-Being Index [93,94]. This is a short self-reported measure of current mental well-being for use in primary care settings that has robust psychometric properties across a range of countries. And the Assessment of Life Habits (version 4.0) assesses life habits from daily activities to social participation across 12 domains [95,96].

Within the frameworks of inclusive disability and participatory research, the case study 3 team will evaluate the uptake and impact of SDM as a tool to support self-management of long-term health conditions. Specifically, new disability-led education and training for people with a learning disability and for supporters, including family and whānau, support workers, and health professionals, will be explored using RE-AIM principles to evaluate the potential of SDM in advancing supported self-management approaches that recognize and respond to people with a learning disability [97].

Data Analysis

Demographic and questionnaire data will be analyzed with descriptive statistics (means, medians, ranges, and frequencies). Differences in living arrangements, migration status, contact with home nations, economic status, and degree of the long-term condition that shape experiences of supported self-management may be explored. Audio-recordings will be transcribed verbatim. The qualitative data analysis will be undertaken by 2 culturally appropriate, trained research team members (Māori, Pacific, and learning disability) for each case study, following the reflexive thematic analysis approach described by Braun and Clarke [98,99]. Although this is a "Western" approach to analysis, notwithstanding reservations expressed by some Māori and Pacific researchers [100], many ingenious studies have used it, considering that its processes are appropriate and complementary [101-103]. Similarly, it is used frequently in research into learning disabilities [104,105]. To gain a sense of the overall themes, the analysis will begin with repeated listening of the audio-recordings and multiple readings of the transcripts. Important statements will be highlighted and provided with a preliminary naming code. With ongoing discussion and reflection by the research team, these initial codes will be collapsed and categorized to create a final coding scheme, which will then be applied across all transcripts within each case study. Subsequent case study research team reflective discussions will collate these codes into themes. Multiple team discussions, ongoing presentations, and discussions with the study advisory groups and community members will enhance the reflexivity, trustworthiness, and robustness of the analysis. Analysis and reporting will be in accordance with the Big Q Qualitative Reporting Guidelines. These guidelines articulate a values-based, rather than consensus-based, framework for reporting and evaluating qualitative research. They were developed to support methodological congruence and reflexively open evaluation and reporting of Big Q Qualitative research. Big Q Qualitative research includes both the techniques and

XSL•FO
RenderX

values of qualitative research and thus rejects objectivist and positivist assumptions, values, and norms [106].

Each study may choose other uniquely diverse methods of evaluating outcomes appropriate to their context and what their communities considered best reflected living well for them, for example stories, dance, artwork, kapa haka (living Māori performance art form encompassing traditional songs, dances, and actions that express language, culture, and heritage), waiata (Māori song), theatre, music, or songs.

Hui and fono (meetings) will be held with participants and research teams to moderate analyzes and interpret the impact of implementation and evaluation, and kōrero or talanoa (discuss) about values, successes, challenges, and share improvement recommendations.

Ethical Considerations

The broad program proposal has been registered and approved by the University of Otago Human Ethics Committee (HE24/002). Ethics for each phase of the research will be specifically applied for and each phase will be undertaken in accordance with the ethical standards of the responsible ethics committee and with the WMA Declaration of Helsinki. To date, this has included approval for the co-design research pertaining to objectives 1 and 2: University of Otago Human Ethics Committee (Health), category A: "Taunakitanga Takitini: reframing 'self'-management support for all in Aotearoa" (H24/002); "Lūsia kiTaulanga: A Pacific study in 'supported' self-management" (H25/0403); and "Exploring Turanga Health values and 'capabilities' for evaluating program outcomes" (H24/0260). The research entitled "Reimagining supported self-management of health and wellbeing with people with learning disability in Aotearoa New Zealand" was approved by the Central Health and Disability Ethics Committee (New Zealand Ministry of Health; 2025 FULL 22164).

Ethical approval from the University of Otago School of Physiotherapy Ethics Committee ("The 'Taunakitanga Takitini' research team members' reflections and engagements about self-management: SoP/EC/2024/09") has also been approved to allow us to capture how our broad team of research members and partners change and grow into their aspirations and understandings of what reframing of self-management support could mean.

In all above ethically approved projects, informed written consent will be obtained from participants. Potential participants are informed of their right to refuse to participate in the research or to withdraw consent to participate at any time without reprisal. They are also informed that every precaution will be taken to protect their privacy and the confidentiality of their personal information. Participants will be reimbursed in recognition of expenses incurred by participating in the study, usually by way of a voucher.

In the research involving adults with a learning disability, as legally required in New Zealand, participants will be consented in a range of ways depending on their level of understanding, including (1) providing independent signed consent, (2) their legal guardian will provide consent, or (3) a consent statement will be provided by a significant other in consultation with the participant. We will provide a plain English study information sheet and consent forms approved by the Ministry of Health's Health and Disability Ethics Committee. Every effort will be made to secure freely given informed consent, which may be verbal or written as appropriate, that participants have actively provided, and we will ensure that they have the time and opportunity to access support in their decision-making, for example, by discussing their choice with a trusted adult or relative.

Results

This 5-year program grant was funded in October 2023. The first 2 years have been devoted to relationship building, ethical applications, research capability and capacity building, and data collection and consultation with the respective communities of each case study. We have completed our first program objective and partially completed our second objective.

Discussion

Principal Findings

Our program of research aims to build New Zealand–specific knowledge about supported "self"-management to enable those who experience the greatest health inequities, that is, underserved populations with long-term conditions, such as Māori and Pacific Peoples, Tāngata Whaikaha Māori and people with a learning (intellectual) disability and their whānau to be supported to live the best possible life. We anticipate that our 3 case studies will develop innovative approaches to support people living with long-term health conditions or permanent disabilities within their communities to live well; approaches that their community members perceive to be acceptable and beneficial.

We aim to disrupt current understandings of supported "self"-management and reframe the concept specifically for New Zealand to enable the vision of Pae ora (good health and well-being). Increasingly, people are living with multimorbidity, long-term conditions, or permanent impairment, which contributes substantially to health loss and additional health inequity [107]. Health and social services providing appropriate support is critical to managing this health crisis [108].

To date, international research has largely informed self-management knowledge and implementation in New Zealand, with little attention given to how New Zealand health care organizations and social services appropriately support people to self-manage their long-term conditions within their own life contexts. Reframing supported "self"-management with the New Zealand world view to include perspectives of Te Ao Māori, Pacific Peoples, and people with learning disabilities is crucial if we truly wish to enable the vision of the Pae ora (healthy futures or healthy life) Act. The current transformational health context, driven by the 2575 Waitangi Tribunal Report and the Pae Ora Act, the COVID-19 pandemic and subsequent economic crises, and the Royal Commission into abuse and neglect of children, young people and adults in the care of the State and faith-based institutions in New Zealand between 1950 and 1999, along with changing governmental

XSL•FO
RenderX

agendas, presents a unique opportunity to reimagine health care access for people most at risk.

Dissemination Plan

A clear pathway of planned activities will generate outputs and outcomes for research end users and knowledge transfer both within our program and wider. Reports and presentations will be iteratively constructed within each case study to ensure the voices and paradigms of the participants remain clear throughout. Rigorous processes will uphold community partners' sovereignty and data ownership. Dissemination processes of hui, fono, meetings, and presentations will verify and shape findings. Appropriate findings will be gifted by the community partners and their communities to participants, key stakeholders, policy advisors, the funder and other health service providers.

Future Research

The dissemination processes are expected to increase the impact of this research and ensure reciprocity through bidirectional learning, bicultural communication, and bidirectional feedback. Such exchanges provide a pathway for impactful transformational health knowledge transfer, whereby mindsets of health and disability providers and policymakers, organizations, professionals, and individuals or whānau are empowered for future health and well-being thinking. Future research should focus on the evaluation of these expected impacts on the perceptions, processes, and policies of health and disability providers and organizations and health professionals. Additionally, research to establish the social return on investment in delivering the innovative approaches to supporting people living with long-term health conditions or permanent disabilities is warranted, and this, in turn, will inform health policies.

Strengths and Limitations

The strength of this research lies in its equity-focused methodology. It explicitly targets health equity by addressing issues identified by and for communities experiencing ongoing health and disability inequities; empowering them to drive the research kaupapa (agenda), building research capacity and capability, and developing health care models contextually relevant for New Zealand that have the potential to positively impact the health and well-being of disadvantaged populations. Impacts will be achieved by working in close partnership with health and disability providers and advocacy services, and by targeted dissemination of findings back to communities, providers, and policymakers.

Furthermore, our work program has a strength-based stance. Evidence from "Western" studies now supports a capabilities approach to supported self-management (developing collaborative, supportive, and enabling relationships between individuals with long-term conditions and health care partners) for improved health outcomes [13,19,25,26]. Reframing this approach from the perspectives of Tāngata Whenua, New Zealand's Indigenous Māori population, and Pacific Peoples, collective approaches and those with learning disabilities have huge potential to foster self-actualization, maintaining and enhancing health and well-being for these populations to live richer, fuller lives; populations frequently underserved in the current delivery of health and disability services. Importantly, drawing on the wisdom of people who emanate from collective and supportive societies with deeply held values and binding genealogical, land, history and spiritual ties [109] will reframe supported self-management in ways that will be beneficial to enhance health and well-being and improve equitable outcomes for all living in New Zealand.

New knowledge generated across our program has the potential to inform policy and practice about service delivery acceptable to the people to whom it matters (particularly Māori, Pacific Peoples, and those with learning disability) and places more emphasis on promotion of well-being. This approach focuses on the inherent strengths and abilities of people, rather than the deficits or problems, acknowledges the wealth of knowledge and experience people with such conditions have, and builds relationships and partnerships between people and health care partners.

Potential limitations of this work need to be acknowledged. The case study approach may limit the generalizability of the findings to other regions and cultures in New Zealand, and internationally. To undertake value-based research of this type takes time and cannot be rushed—time to get to know the communities and develop mutual respect and trust—yet the sociopolitical landscape can rapidly change and render the findings less influential as political standpoints and polices transform.

We firmly believe that if health care is culturally appropriate and made accessible and equitable for our 3 targeted groups, applying the concept of universal design, it will enable all people in New Zealand to have improved health. Lessons learnt from our 3 case studies can thus be extrapolated to provide learning points for New Zealand health services with the aim of improving health at the community (primary care) level, which in turn may reduce the health care required at secondary and tertiary levels. We emphatically hold the view that people are the best experts in their own health conditions—they have the lived experiences within their own life contexts. However, for many people, especially the population groups we seek to work alongside, mainstream policies (pertaining to both health and beyond) and health and disability initiatives frequently leave them behind, and they simply become "the forgotten" [110].

Acknowledgments

We fully acknowledge the wisdom and input from our wider Taunakitanga Takitini research team, our national and international advisory group members, and importantly, the multiple communities we are working with and alongside.

No generative artificial intelligence was used in the writing of this manuscript.

XSL•FO
RenderX

Funding

Our program grant was funded by the Health Research Council of New Zealand (23/448).

Authors' Contributions

LH: conceptualization, funding acquisition, methodology, and writing – original draft
BHJ, OD, SK, BMV, PN, EMK, MP, KD: conceptualization, funding acquisition, methodology, and writing – reviewing and editing original draft
HT, TI, NL, RFW: conceptualization, funding acquisition, and methodology
DK, AW: conceptualization, funding acquisition, methodology, project administration, and writing – reviewing and editing original draft

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from the Health Research Council of New Zealand.
[PDF File (Adobe PDF File), 116 KB - resprot_v15i1e89658_app1.pdf]

Multimedia Appendix 2
23-448-Hale 53 Applicant Peer Review.
[PDF File (Adobe PDF File), 105 KB - resprot_v15i1e89658_app2.pdf]

Multimedia Appendix 3
23-448-Hale 81 Applicant Peer Review.
[PDF File (Adobe PDF File), 105 KB - resprot_v15i1e89658_app3.pdf]

Multimedia Appendix 4
23-448-Hale 91 Applicant Peer Review.
[PDF File (Adobe PDF File), 146 KB - resprot_v15i1e89658_app4.pdf]

References

1. Hogan S, Song Z. The Cost of Long-Term Conditions in New Zealand: Review of the Evidence -2021 Update. Wellington: New Zealand Institute of Economic Research; 2022. URL: https://www.health.govt.nz/publications/the-cost-of-long-term-conditions-in-new-zealand [accessed 2026-02-17]
2. Holland E, Matthews K, Macdonald S, Ashworth M, Laidlaw L, Cheung KSY, et al. The impact of living with multiple long-term conditions (multimorbidity) on everyday life - a qualitative evidence synthesis. BMC Public Health 2024;24(1):3446 [FREE Full text] [doi: 10.1186/s12889-024-20763-8] [Medline: 39696210]
3. Underwood L, Teng A, Bowden N, Dewes O, Marek L, Milne B. Long-term health conditions among household families in Aotearoa New Zealand: Cross-sectional analysis of integrated Census and administrative data. N Z Med J 2024;137(1596):20-34. [doi: 10.26635/6965.6370] [Medline: 38843547]
4. Kendall E, Ehrlich C, Sunderland N, Muenchberger H, Rushton C. Self-managing versus self-management: Reinvigorating the socio-political dimensions of self-management. Chronic Illn 2011;7(1):87-98. [doi: 10.1177/1742395310380281] [Medline: 20921037]
5. De Vos P, Malaise G, De Ceukelaire W, Pérez D, Lefèvre P, Van der Stuyft P. Participation and empowerment in primary health care: From Alma Ata to the era of globalization. Social Med 2009;4(2):121-127. [doi: 10.71164/socialmedicine.v4i2.2009.269]
6. Furler J, Harris M, Rogers A. Equity and long-term condition self-management. Chronic Illn 2011;7(1):3-5. [doi: 10.1177/1742395310386978] [Medline: 21398320]
7. Health and disability system review – final report – pūrongo whakamutung. Health and Disability System Review. 2020. URL: https://www.health.govt.nz/system/files/2022-09/health-disability-system-review-final-report.pdf [accessed 2026-02-17]
8. Ministry of Health. Briefing to the incoming Minister of Health. In: 2017: The New Zealand Health and Disability System. Wellington, New Zealand: Ministry of Health; 2017.
9. Schoen C, Osborn R, Squires D, Doty M, Pierson R, Applebaum S. New 2011 survey of patients with complex care needs in eleven countries finds that care is often poorly coordinated. Health Aff (Millwood) 2011;30(12):2437-2448. [doi: 10.1377/hlthaff.2011.0923] [Medline: 22072063]

XSL•FO
RenderX

10. Stanley J, Semper K, Millar E, Sarfati D. Epidemiology of multimorbidity in New Zealand: A cross-sectional study using national-level hospital and pharmaceutical data. BMJ Open 2018;8(5):e021689 [FREE Full text] [doi: 10.1136/bmjopen-2018-021689] [Medline: 29794103]
11. Self-management support resources. Health New Zealand (Te Whatu Ora). URL: https://www.tewhatuora.govt.nz/for-health-professionals/clinical-guidance/diseases-and-conditions/long-term-conditions/self-management-support-resources [accessed 2026-02-17]
12. Healthify. URL: https://healthify.nz/health-a-z/s/self-management-and-self-management-support [accessed 2026-02-17]
13. Dineen-Griffin S, Garcia-Cardenas V, Williams K, Benrimoj SI. Helping patients help themselves: A systematic review of self-management support strategies in primary health care practice. PLoS One 2019;14(8):e0220116 [FREE Full text] [doi: 10.1371/journal.pone.0220116] [Medline: 31369582]
14. Gallant MP, Spitze G, Grove JG. Chronic illness self-care and the family lives of older adults: A synthetic review across four ethnic groups. J Cross Cult Gerontol 2010;25(1):21-43 [FREE Full text] [doi: 10.1007/s10823-010-9112-z] [Medline: 20177963]
15. Morgan HM, Entwistle VA, Cribb A, Christmas S, Owens J, Skea ZC, et al. We need to talk about purpose: A critical interpretive synthesis of health and social care professionals' approaches to self-management support for people with long-term conditions. Health Expect 2017;20(2):243-259 [FREE Full text] [doi: 10.1111/hex.12453] [Medline: 27075246]
16. Wilkinson M, Whitehead L, Crowe M. Nurses perspectives on long-term condition self-management: A qualitative study. J Clin Nurs 2016;25(1-2):240-246. [doi: 10.1111/jocn.13072] [Medline: 26769211]
17. Hale L, Oosman S, Stewart AV. Editorial: Challenging the concept of self-management support in unique and diverse populations. Front Rehabil Sci 2022;3:999528 [FREE Full text] [doi: 10.3389/fresc.2022.999528] [Medline: 36189014]
18. Hale L, Higgs C, Keen D, Smith C. It is "All About Relationships" in lifestyle programmes for adults living with type two diabetes underpinned by a person/whānau-centred care approach. Front Rehabil Sci 2022;3:829542 [FREE Full text] [doi: 10.3389/fresc.2022.829542] [Medline: 36189031]
19. McKinlay E, McDonald J, Darlow B, Perry M. Social networks of patients with multimorbidity: A qualitative study of patients' and supporters' views. J Prim Health Care 2017;9(2):153-161 [FREE Full text] [doi: 10.1071/HC16062] [Medline: 29530227]
20. Bridges self-management. Bridges Self-Management. URL: https://www.bridgesselfmanagement.org.uk/ [accessed 2026-02-17]
21. The Flinders Program. Flinders University. URL: https://www.flindersprogram.com.au/ [accessed 2026-02-17]
22. Build your chronic conditions toolbox. Stanford University. URL: https://healthyliving.stanford.edu/classes/register/webinar.php?cc=bcbh-30r [accessed 2026-02-17]
23. Hardman R, Begg S, Spelten E. What impact do chronic disease self-management support interventions have on health inequity gaps related to socioeconomic status: A systematic review. BMC Health Serv Res 2020;20(1):150 [FREE Full text] [doi: 10.1186/s12913-020-5010-4] [Medline: 32106889]
24. Newman S, Steed L, Mulligan K. Chronic Physical Illness: Self-Management and Behavioural Interventions. Maidenhead: Open University Press; 2008.
25. Ehrlich C, Kendall E, Parekh S, Walters C. The impact of culturally responsive self-management interventions on health outcomes for minority populations: A systematic review. Chronic Illn 2016;12(1):41-57. [doi: 10.1177/1742395315587764] [Medline: 26026156]
26. de Longh A, Fagan P, Fenner J, Kidd L. A practical guide to self-management support. In: Key Components for Successful Implementation. London, UK: The Health Foundation; 2015.
27. Palumbo R, Annarumma C. Empowering organizations to empower patients: An organizational health literacy approach. International Journal of Healthcare Management 2016;11(2):133-142. [doi: 10.1080/20479700.2016.1253254]
28. Sheridan NF, Kenealy TW, Fitzgerald AC, Kuluski K, Dunham A, McKillop AM, et al. How does it feel to be a problem? Patients' experiences of self-management support in New Zealand and Canada. Health Expect 2019;22(1):34-45 [FREE Full text] [doi: 10.1111/hex.12823] [Medline: 30244514]
29. Rong T, Ristevski E, Carroll M. Exploring community engagement in place-based approaches in areas of poor health and disadvantage: A scoping review. Health Place 2023;81:103026 [FREE Full text] [doi: 10.1016/j.healthplace.2023.103026] [Medline: 37084705]
30. McKinlay E, Young J, Gray B. General practice and patients' views of the social networks of patients with multimorbidity. J Prim Health Care 2018;10(3):258-266 [FREE Full text] [doi: 10.1071/HC17050] [Medline: 31039940]
31. Young J, Egan T, Williamson M, Jaye C, Kenrick K, Ross J, et al. An exercise to map patient-centred care networks. Clin Teach 2016;13(6):448-450. [doi: 10.1111/tct.12459] [Medline: 26748899]
32. Young J, Jaye C, Egan T, Williamson M, Askerud A, Radue P, et al. Communities of clinical practice in action: Doing whatever it takes. Health (London) 2018 Mar;22(2):109-127. [doi: 10.1177/1363459316688515] [Medline: 28111987]
33. Hauora: Report on stage one of the health services and outcomes kaupapa inquiry. Waitangi Tribunal Report (WIA 2572). URL: https://www.tewhatuora.govt.nz/assets/For-the-health-sector/Specific-life-stage/Health-of-older-people/Hauora-2023-W.pdf [accessed 2026-02-17]

XSL•FO
RenderX

34. Pae ora (Healthy Futures) Act. Ministry of Health. 2022. URL: https://www.health.govt.nz/about-us/new-zealands-health-system/overview-and-statutory-framework/pae-ora-healthy-futures-act [accessed 2026-02-17]
35. Donald Beasley Institute. Disabled person-led monitoring of the UNCRPD: My experiences, my rights: Disability supports and services report 1 - Whaikaha - Ministry of Disabled People. Disabled People's Organisation Coalition. 2025. URL: https://www.donaldbeasley.org.nz/assets/projects-file/Report-1-Whaikaha-l-Ministry-of-Disabled-People-Full-Report-Word-Doc.pdf [accessed 2026-02-17]
36. Cornish F, Breton N, Moreno-Tabarez U, Delgado J, Rua M, de-Graft Aikins A, et al. Participatory action research. Nat Rev Methods Primers 2023;3(1):1-14. [doi: 10.1038/s43586-023-00214-1]
37. Mercier OR. Mātauranga and science. NZ Sci Review 2018;74(4):83-90. [doi: 10.26686/nzsr.v74i4.8477]
38. Te Mana Raraunga: Maori data sovereignty network. URL: https://www.temanararaunga.maori.nz/ [accessed 2026-02-17]
39. Convention on the rights of persons with disabilities (CRPD); Article 25 - Health. United Nations Department of Economics and Social Affairs Disability. URL: https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/article-25-health.html [accessed 2026-02-17]
40. Mertens DM. Transformative research: Personal and societal. Int J Ther Rehabil 2017;4(1):18-24. [doi: 10.1515/ijtr-2017-0001]
41. Hudon C, Chouinard M, Bisson M, Danish A, Karam M, Girard A, et al. Case study with a participatory approach: Rethinking pragmatics of stakeholder engagement for implementation research. Ann Fam Med 2021;19(6):540-546 [FREE Full text] [doi: 10.1370/afm.2717] [Medline: 34750129]
42. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research. Thousand Oaks, CA: SAGE Publications; 2011.
43. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. Am J Public Health 1999;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
44. Agnello DM, Anand-Kumar V, An Q, de Boer J, Delfmann LR, Longworth GR, et al. Co-creation methods for public health research - characteristics, benefits, and challenges: A Health CASCADE scoping review. BMC Med Res Methodol 2025;25(1):60 [FREE Full text] [doi: 10.1186/s12874-025-02514-4] [Medline: 40050729]
45. Talevski J, Kulnik ST, Jessup RL, Falls R, Cvetanovska N, Beauchamp A. Use of co-design methodology in the development of cardiovascular disease secondary prevention interventions: A scoping review. Health Expect 2023;26(1):16-29 [FREE Full text] [doi: 10.1111/hex.13633] [Medline: 36366855]
46. Jones B, Ingham T, Davies C, Cram F. Whānau tuatahi: Māori community partnership research using a kaupapa Māori methodology. Mai Review J 2010;3 [FREE Full text]
47. Pere RR, Pere RT. Te Wheke: A Celebration of Infinite Wisdom. New Zealand: Ao Ako Global Learning New Zealand; 1997.
48. Hawaikirangi L. An exploration of wellbeing in Hapū Wānanga through a Te Wheke framework analysis (Thesis, Master of Social Sciences). The University of Waikato, Hamilton, New Zealand; 2021. URL: https://hdl.handle.net/10289/14430 [accessed 2026-02-17]
49. Love C. Extensions on Te Wheke (Working Papers No. 6-04). Lower Hutt, New Zealand: The Open Polytechnic of New Zealand; 2004.
50. Tu'itahi S. Fonua: A model for pacific health promotion. Symposium Conducted at the Meeting of the Health Promotion Forum of New Zealand. URL: https://hpfnz.org.nz/pacific-health-promotion/pacific-health-models/ [accessed 2026-02-17]
51. Whakamaua: Māori health action plan 2020-2025. Ministry of Health. Wellington, NZ: Ministry of Health; 2020. URL: https://www.health.govt.nz/publications/whakamaua-maori-health-action-plan-2020-2025 [accessed 2026-02-17]
52. Smith GH. The development of kaupapa Māori : Theory and praxis. Doctoral dissertation, The University of Auckland. A foundational text that articulated Kaupapa Māori theory. 1997. URL: https://tewaharoa.victoria.ac.nz/discovery/fulldisplay/alma99774124002386/64VUW_INST:VUWNUI [accessed 2026-02-17]
53. Smith LT. Decolonizing Methodologies: Research and Indigenous Peoples. London: University of Otago Press; 1999.
54. Wilson D, Mikahere-Hall A, Sherwood J. Using indigenous kaupapa Māori research methodology with constructivist grounded theory: Generating a theoretical explanation of indigenous womens realities. International Journal of Social Research Methodology 2021;25(3):375-390. [doi: 10.1080/13645579.2021.1897756]
55. Ingham TR, Jones BH, Perry MA, Sporle A, Elliott T, King PT, et al. Māori health, wellbeing, and disability in aotearoa New Zealand: A national survey. Int J Environ Res Public Health 2025;22(6):829 [FREE Full text] [doi: 10.3390/ijerph22060829] [Medline: 40566257]
56. Ola manuia: Pacific health and wellbeing action plan 2020-2025. National Academies of Ministry of Health. URL: https://www.health.govt.nz/publications/ola-manuia-pacific-health-and-wellbeing-action-plan-2020-2025 [accessed 2026-02-17]
57. Clifton S, Cooper E, Bourke J, Connor S, Denton S, Dominish B, et al. A framework for disability lived expertise. Evid Policy 2025:1-21. [doi: 10.1332/17442648Y2025D000000069] [Medline: 41073073]
58. Jackson KM, Pukys S, Castro A, Hermosura L, Mendez J, Vohra-Gupta S, et al. Using the transformative paradigm to conduct a mixed methods needs assessment of a marginalized community: Methodological lessons and implications. Eval Program Plann 2018;66:111-119 [FREE Full text] [doi: 10.1016/j.evalprogplan.2017.09.010] [Medline: 29091786]

XSL•FO
RenderX

59. Mertens DM. Transformative paradigm: Mixed methods and social justice. Journal of Mixed Methods Research 2007;1(3):212-225. [doi: 10.1177/1558689807302811]
60. Ellem K, Venables J. Relationship-based practice with people with a mild intellectual disability who have been socially marginalised and excluded. In: Bigby C, Hough A, editors. Disability Practice. Singapore: Palgrave Macmillan; 2024.
61. Chalachanová A, Østby M, Power A, Tilley L, Walmsley J, Westergård B, et al. Building relationships in inclusive research in diverse contexts. Scand J Disabil Res 2020;22(1):147-157. [doi: 10.16993/sjdr.681]
62. Turanga Health. URL: http://www.turangahealth.co.nz/ [accessed 2026-02-17]
63. Willing E, Paine SJ, Wyeth E, Te Ao B, Vaithianathan R, Reid P. Indigenous voices on measuring and valuing health states. AlterNative: An international journal of indigenous peoples 2019;16(1):3-9. [doi: 10.1177/1177180119885418]
64. The Cause Collective. 2024. URL: https://thecausecollective.org.nz/wp-content/uploads/2024/11/TCC_AnnualReport_2024-Web-Version.pdf [accessed 2026-02-17]
65. Whānau Ora. 2025. URL: https://whanauora.nz/about-us [accessed 2026-02-17]
66. Donald Beasley Institute. URL: https://www.donaldbeasley.org.nz/ [accessed 2026-02-17]
67. People First. URL: https://www.peoplefirst.org.nz/ [accessed 2026-02-17]
68. New Zealand disability survey. Statistics New Zealand. 2023. URL: https://www.stats.govt.nz/information-releases/disability-statistics-2023/ [accessed 2026-02-17]
69. Health indicators for New Zealanders with intellectual disability. Ministry of Health. Wellington, New Zealand: Ministry of Health; 2011. URL: https://www.health.govt.nz/system/files/2011-10/health-indicators-nzders-intellectual-disability.pdf [accessed 2026-02-17]
70. Emerson E, Hatton C. Health Inequalities and People with Intellectual Disabilities. Cambridge: Cambridge University Press; 2013.
71. Kripke C. Supported health care decision-making for people with intellectual and cognitive disabilities. Fam Pract 2016;33(5):445-446 [FREE Full text] [doi: 10.1093/fampra/cmw060] [Medline: 27432358]
72. House A, Latchford G, Russell AM, Bryant L, Wright J, Graham E, et al. Development of a supported self-management intervention for adults with type 2 diabetes and a learning disability. Pilot Feasibility Stud 2018;4:106 [FREE Full text] [doi: 10.1186/s40814-018-0291-7] [Medline: 29862037]
73. Hale LA, Trip HT, Whitehead L, Conder J. Self‐management abilities of diabetes in people with an intellectual disability living in New Zealand. Policy Practice Intel Disabi 2011;8(4):223-230. [doi: 10.1111/j.1741-1130.2011.00314.x]
74. Trip H, Conder J, Hale L, Whitehead L. The role of key workers in supporting people with intellectual disability in the self-management of their diabetes: A qualitative New Zealand study. Health Soc Care Community 2016;24(6):789-798. [doi: 10.1111/hsc.12262] [Medline: 26099802]
75. Whitehead LC, Trip HT, Hale LA, Conder J. Negotiated autonomy in diabetes self-management: The experiences of adults with intellectual disability and their support workers. J Intellect Disabil Res 2016;60(4):389-397. [doi: 10.1111/jir.12257] [Medline: 26840793]
76. Jones M, MacGillivray S, Kroll T, Zohoor AR, Connaghan J. A thematic analysis of the conceptualisation of self-care, self-management and self-management support in the long-term conditions management literature. J Nursing Healthcare Chronic Illn 2011;3(3):174-185. [doi: 10.1111/j.1752-9824.2011.01096.x]
77. Browning M, Bigby C, Douglas J. Supported decision making: Understanding how its conceptual link to legal capacity is influencing the development of practice. Research and Practice in Intellectual and Developmental Disabilities 2014;1(1):34-45. [doi: 10.1080/23297018.2014.902726]
78. Kohn NA, Blumenthal JA. A critical assessment of supported decision-making for persons aging with intellectual disabilities. Disabil Health J 2014;7(1 Suppl):S40-S43. [doi: 10.1016/j.dhjo.2013.03.005] [Medline: 24456684]
79. Sullivan WF, Heng J. Supporting adults with intellectual and developmental disabilities to participate in health care decision making. Can Fam Physician 2018;64(Suppl 2):S32-S36 [FREE Full text] [Medline: 29650742]
80. Maine A, Dickson A, Truesdale M, Brown M. An application of Bandura's 'Four Sources of Self-efficacy' to the self-management of type 2 diabetes in people with intellectual disability: An inductive and deductive thematic analysis. Res Dev Disabil 2017;70:75-84 [FREE Full text] [doi: 10.1016/j.ridd.2017.09.004] [Medline: 28918307]
81. Stokes T, Wilkinson A, Jayakaran P, Higgs C, Keen D, Mani R, et al. Implementation of the diabetes community exercise and education programme (DCEP) for the management of type 2 diabetes: qualitative process evaluation. BMJ Open 2022;12(5):e059853 [FREE Full text] [doi: 10.1136/bmjopen-2021-059853] [Medline: 35623756]
82. Holtrop JS, Rabin BA, Glasgow RE. Qualitative approaches to use of the RE-AIM framework: Rationale and methods. BMC Health Serv Res 2018;18(1):177 [FREE Full text] [doi: 10.1186/s12913-018-2938-8] [Medline: 29534729]
83. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: A systematic review of use over time. Am J Public Health 2013;103(6):e38-e46. [doi: 10.2105/AJPH.2013.301299] [Medline: 23597377]
84. Danish A, Chouinard MC, Aubrey-Bassler K, Burge F, Doucet S, Ramsden VR, et al. Protocol for a mixed-method analysis of implementation of case management in primary care for frequent users of healthcare services with chronic diseases and complex care needs. BMJ Open 2020;10(6):e038241 [FREE Full text] [doi: 10.1136/bmjopen-2020-038241] [Medline: 32487584]

XSL•FO
RenderX

85. Hudon C, Chouinard MC, Aubrey-Bassler K, Burge F, Doucet S, Ramsden VR, et al. Case management in primary care for frequent users of healthcare services with chronic diseases and complex care needs: An implementation and realist evaluation protocol. BMJ Open 2018;8(11):e026433 [FREE Full text] [doi: 10.1136/bmjopen-2018-026433] [Medline: 30478129]
86. Holland J, Thomson R, Henderson S. Qualitative Longitudinal Research: A Discussion Paper. London: London South Bank University; 2006.
87. Owen D, Green A. A scoping study for a longitudinal survey of ethnic minorities for the UK. A report presented to the national longitudinal strategy committee. Warwick: University of Warwick; 2004. URL: https://www.understandingsociety.ac.uk/wp-content/uploads/working-papers/2008-02.pdf [accessed 2026-02-17]
88. Kingi T, Durie M. Hua Oranga. A Māori Measure of Mental Health Outcome. Massey University, School of Maori Studies, Te Pumanawa Hauora Palmerston North. URL: https://www.massey.ac.nz/documents/521/T_Kingi__M_Durie_Hua_Oranga_A_maori_measure_of_mental_health_outcome.pdf [accessed 2026-02-17]
89. Dewes O, Milne B, Sporle A. Intergenerational, integrative and intellectual Pacific properties and pathways for life (IPforLife): A study protocol. J Prim Health Care 2022;14(2):173-178 [FREE Full text] [doi: 10.1071/HC22004] [Medline: 35771695]
90. Sopoaga F, Nada-Raja S, Leckie T, Samaranayaka A. Mental health and wellbeing of Pacific students : Protocol of a Pacific perspective study. Pac Health Dialog 2020;21(5):272-280. [doi: 10.26635/phd.2020.626]
91. Manuela S, Sibley CG. The pacific identity and wellbeing scale (PIWBS): A culturally-appropriate self-report measure for pacific peoples in New Zealand. Social Indicators Research 2013;112(1):83-103 [FREE Full text] [doi: 10.1007/s11205-012-0041-9]
92. Manuela S, Sibley CG. The Pacific Identity and Wellbeing Scale - revised: Comparisons across Pacific groups. Cultur Divers Ethnic Minor Psychol 2015;21(1):146-155. [doi: 10.1037/a0037536] [Medline: 25111555]
93. Fougeyrollas P, Noreau L, Bergeron H, Cloutier R, Dion SA, St-Michel G. Social consequences of long term impairments and disabilities: Conceptual approach and assessment of handicap. Int J Rehabil Res 1998;21(2):127-141. [doi: 10.1097/00004356-199806000-00002] [Medline: 9924676]
94. Wellbeing measures in primary health care/the depcare project. WHO. WHO Regional Office for Europe: Copenhagen; 1998. URL: https://iris.who.int/bitstream/handle/10665/349766/WHO-EURO-1998-4234-43993-62027-eng.pdf [accessed 2026-02-17]
95. Noreau L, Fougeyrollas P, Vincent C. The LIFE-H: Assessment of the quality of social participation. TAD 2002;14(3):113-118. [doi: 10.3233/tad-2002-14306]
96. Noreau L, Desrosiers J, Robichaud L, Fougeyrollas P, Rochette A, Viscogliosi C. Measuring social participation: Reliability of the LIFE-H in older adults with disabilities. Disabil Rehabil 2004;26(6):346-352. [doi: 10.1080/09638280410001658649] [Medline: 15204486]
97. Bamberg M. Narrative discourse and identities. In: Meister JC, Schernus W, editors. Narratology Beyond Literary Criticism. New York: Walter de Gruyter; 2004:213-237.
98. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exercise Health 2019;11(4):589-597 [FREE Full text] [doi: 10.1016/j.jsat.2009.06.002]
99. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qualitative Research in Psychology 2020;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
100. Rangiwai B, Enari D, Masae C, Paea D, Tahilanu-Mapili L, Vailahi V. Lost in translation: Reflexive thematic analysis in research with Pacific peoples. Te Kaharoa 2021;14(1). [doi: 10.24135/tekaharoa.v17i1.354]
101. McCarty G, Maclennan K, Derrett S, Sullivan T, Crengle S, Wyeth E. Kia Whakatōmuri te haere whakamua: Applying an indigenous qualitative method with rangatahi māori and whānau. MAI Journal 2025;14(2):251-264. [doi: 10.20507/maijournal.2025.14.2.11]
102. Manuel AR, Smith J, Lowe T, O'Beirne GA. Kaupapa Māori development of te reo Māori assessments and culturally responsive hearing healthcare in Te Waipounamu: Findings from wānanga. J R Soc N Z 2025;55(3):687-703 [FREE Full text] [doi: 10.1080/03036758.2024.2418999] [Medline: 39989648]
103. Enari D. Methodology marriage: Merging western and Pacific research design. Proc. J Interdiscip. Res 2021. [doi: 10.26021/10641]
104. Elise F, Irvine B, Brinkert J, Hamilton C, Farran EK, Milne E, et al. Perceptual experiences of autistic people with an intellectual disability and people with williams syndrome: A reflexive thematic analysis. J Appl Res Intellect Disabil 2025;38(1):e13326. [doi: 10.1111/jar.13326] [Medline: 39611267]
105. Cashin A, Morphet J, Wilson NJ, Pracilio A. Barriers to communication with people with developmental disabilities: A reflexive thematic analysis. Nurs Health Sci 2024;26(1):e13103. [doi: 10.1111/nhs.13103] [Medline: 38408760]
106. Braun V, Clarke V. Reporting guidelines for qualitative research: A values-based approach. Qualitative Research in Psychology 2024;22(2):399-438. [doi: 10.1080/14780887.2024.2382244]
107. Sarangi S. Towards a communicative mentality in medical and healthcare practice. Commun Med 2004;1(1):1-11. [doi: 10.1515/come.2004.002] [Medline: 16808684]

XSL•FO
RenderX

108. Lynch R, Hanckel B, Green J. The (failed) promise of multimorbidity: Chronicity, biomedical categories, and public health. Crit Public Health 2022;32(4):450-461 [FREE Full text] [doi: 10.1080/09581596.2021.2017854] [Medline: 38013883]
109. Reweti A. Understanding how whānau-centred initiatives can improve Māori health in Aotearoa New Zealand. Health Promot Int 2023;38(4):daad070. [doi: 10.1093/heapro/daad070] [Medline: 37493232]
110. They have forgotten us. https://reliefweb.int/organization/sightsavers. URL: https://reliefweb.int/report/world/they-do-not-care-about-us-they-have-forgotten-us-new-research-shows-disability [accessed 2026-02-17]

Abbreviations

DPO: disabled person's organization
RE-AIM: reach, effectiveness, adoption, implementation, and maintenance
SDM: supported decision-making
TCC: The Cause Collective

Edited by J Sarvestan; submitted 16.Dec.2025; this is a non–peer-reviewed article;accepted 20.Feb.2026; published 08.Apr.2026.

Please cite as:
Hale L, Jones BH, Dewes O, Keown S, Mirfin-Veitch B, Wilkinson A, Keen D, Norris P, McKinlay E, Perry M, Diesfeld K, Trip H, Ingham T, Longnecker N, Francis Watene R
Taunakitanga Takitini, Reframing Self-Management Support for All in Aotearoa New Zealand: Protocol for a Participatory Case Study Program of Research
JMIR Res Protoc 2026;15:e89658
URL: https://www.researchprotocols.org/2026/1/e89658
doi:10.2196/89658
PMID:

©Leigh Hale, Bernadette Huatau Jones, Ofa Dewes, Shirley Keown, Brigit Mirfin-Veitch, Amanda Wilkinson, Donna Keen, Pauline Norris, Eileen McKinlay, Meredith Perry, Kate Diesfeld, Henrietta Trip, Tristram Ingham, Nancy Longnecker, Roberta Francis Watene. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 08.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Kenyan Neonatal Mortality Risk Predictor: Protocol for a User-Centered Design Evaluation

Ronald Danny Nyatuka[1], PhD; Paul Macharia[1], PhD; Kakhata Esther[1], PhD; Faith Siva[1], MSc; Betsy Muriithi[1], PhD; Md Shafiqur Rahman Jabin[2,3], PhD

[1]School of Computing and Engineering Sciences, Strathmore University, Nairobi, Nairobi, Kenya

[2]Department of Medicine and Optometry, Linnaeus University, Universitetsplatsen 1, Växjö, Sweden

[3]Faculty of Health Studies, University of Bradford, Bradford, England, United Kingdom

Corresponding Author:
Md Shafiqur Rahman Jabin, PhD
Department of Medicine and Optometry, Linnaeus University, Universitetsplatsen 1, Växjö, Sweden

Abstract

Background: Neonatal mortality remains a major public health challenge in low- and middle-income countries (LMICs), particularly in sub-Saharan Africa, where health systems often lack effective triage mechanisms to identify and prioritize high-risk neonates. Existing clinical tools frequently fail to support timely decision-making during the critical early postnatal period. A previous machine learning (ML)–based neonatal risk prediction model developed using multicountry LMIC datasets demonstrated high predictive accuracy for neonatal mortality in the Indian context, achieving an area under the curve above 0.80. The model incorporates 11 neonatal parameters assessed from delivery through day 2 of life, with birth weight identified as the strongest predictor.

Objective: This study aims to evaluate the contextual feasibility, applicability, and usability of the neonatal risk predictor variables embedded in the ML model in Kenyan health care facilities to inform potential adoption aligned with national health priorities and global targets, including Sustainable Development Goal 3.2.

Methods: A mixed methods feasibility study will be conducted through the real-world preimplementation of a neonatal risk assessment tool in 3 Kenyan health facilities that provide neonatal services. Qualitative and quantitative approaches will be combined to strengthen methodological rigor and enhance the credibility of the findings. The study will be implemented in 3 sequential phases. Phase 1 will involve key informant interviews to capture contextual insights on existing workflows, stakeholder perceptions, and implementation considerations. Phase 2 will comprise a 4-month intervention period during which a paper-based neonatal risk tool will be integrated into routine clinical workflows to assess feasibility, applicability, and usability. Phase 3 will involve postintervention surveys conducted longitudinally at 2 time points to evaluate outcomes, user experience, and implementation barriers.

Results: The neonatal risk predictors are expected to demonstrate strong contextual feasibility, allowing for integration into existing triage systems with minimal workflow disruption. Their applicability is expected to enable early identification of high-risk neonates within the first 48 hours of life, thereby enabling timely clinical decision-making. Usability assessments are expected to indicate positive user experience, acceptability, and perceived usefulness among health care workers. Six participants (frontline neonatal health workers) had been recruited by the time of initial submission of the manuscript (August 2025). Data analysis is ongoing and is expected to be concluded in March 2026. Publication of study findings is expected by June 2026.

Conclusions: The study is expected to generate actionable evidence to support the translation of ML-based neonatal risk prediction into routine clinical practice in LMIC settings, thereby reducing neonatal mortality and advancing progress toward Sustainable Development Goal target 3.2.

International Registered Report Identifier (IRRID): DERR1-10.2196/81996

(JMIR Res Protoc 2026;15:e81996) doi:10.2196/81996

KEYWORDS

neonatal mortality; machine learning; low- and middle-income countries; feasibility; predictive accuracy

XSL•FO
RenderX

Introduction

Overview

Global Neonatal Health Situation

In 2022, approximately 2.3 million newborns died within the first 28 days of life, averaging 6500 deaths per day and representing 47% of all deaths under the age of 5 years worldwide [1]. Neonatal mortality remains a pressing global health concern, particularly in low- and middle-income countries (LMICs) with limited access to quality maternal and newborn care. Sub-Saharan Africa (SSA) and South and Central Asia bear the highest burden, with SSA recording 27 deaths per 1000 live births and South and Central Asia recording 21 deaths per 1000 live births [1]. Although neonatal deaths declined from 5 million in 1990 to 2.3 million in 2022, progress toward Sustainable Development Goal (SDG) target 3.2, which aims to reduce neonatal mortality to ≤12 per 1000 live births by 2030, remains slow and uneven. Newborns in the highest-mortality countries face a 60-fold higher risk of dying within 28 days than those in high-income countries [1]. These disparities reflect systemic inequities in health systems, socioeconomic conditions, and access to essential care, emphasizing the need to strengthen maternal and neonatal health services in high-burden LMICs.

Neonatal Health in LMIC Settings

Neonatal mortality in LMICs is largely driven by preventable conditions such as diarrhea, infections, and respiratory complications [2]. Over 30% of neonatal deaths occur within 24 hours of hospital admission, often due to inadequate triage systems, late presentation, insufficient medical equipment, and personnel shortages [2,3]. Risk factors such as prematurity, respiratory distress, infections, and congenital anomalies exacerbate these challenges. Timely identification of high-risk neonates is critical for rapid emergency interventions to avert preventable deaths. Global neonatal care guidelines advocate for rapid triage, evidence-based treatment, and continuous monitoring; however, implementation in LMICs remains limited [4]. Overcrowded emergency systems; long travel distances; extended waiting times; and shortages of specialists, including neonatologists and pediatricians, further impede effective care [5,6]. These gaps underscore the urgent need for scalable, data-driven tools to support early identification and prioritization of high-risk neonates.

Neonatal Health in Kenya

In Kenya, the neonatal mortality rate (NMR) stands at 22 per 1000 live births, reflecting limited access to quality intrapartum and postnatal care [7,8]. Leading causes of neonatal deaths include intrapartum complications; preterm birth; infections; and low birth weight, which affects 8% of Kenyan newborns [9-11]. Infants with low birth weight face mortality rates up to 20 times higher than their counterparts in high-income countries. Over 60% of facilities under the Kenyan Ministry of Health (MOH) lack standardized resuscitation protocols, and staff shortages further compromise timely care. Neonatal triage often relies on subjective clinical judgment, delaying recognition of high-risk neonates [12]. Similar challenges are reported in Uganda and India [13,14]. Addressing these gaps requires automated, evidence-based decision support systems to improve neonatal outcomes.

LMIC-specific innovations such as the Every Newborn Action Plan emphasize scalable health ITs, including machine learning (ML), to reduce neonatal deaths and stillbirths [15]. For example, Rwanda's intensive care unit network achieved a 30% reduction in neonatal mortality using data-driven protocols [16]. To meet SDG target 3.2, Kenya must leverage technology to accelerate the annual NMR decline from 3.2% to at least 6.5% [17].

Role of ML Tools in Reducing Neonatal Mortality in LMICs

ML tools offer potential to enhance triage and clinical decision-making through early prediction, risk stratification, and personalized interventions. ML models developed using multicountry LMIC datasets have demonstrated high accuracy in identifying high-risk neonates, supporting early interventions and improved outcomes [13,18]. ML-enabled mobile health apps such as Ubenwa, which detects neonatal asphyxia from cry analysis, demonstrate practical applicability in resource-limited settings [19]. Integrating ML with health information systems such as DHIS2 further supports real-time, data-driven neonatal care [20].

Despite this promise, ML adoption in LMICs is hindered by digital infrastructure gaps, poor data quality, and limited technical capacity [21]. Effective implementation requires context-specific, user-friendly solutions that align with existing health care workflows. Recognizing this, this study proposes a user-centric evaluation of the neonatal risk predictor variables incorporated into the ML model by Shukla et al [13] (Table 1), assessing their feasibility, applicability, and usability within Kenyan MOH facilities. The study will leverage these predictors from a multicountry LMIC database to determine the model's potential adoption to reduce neonatal mortality in Kenya.

XSL•FO
RenderX

Table . Key neonatal variables for predicting perinatal and neonatal risk in low- and middle-income countries (LMICs) according to the machine learning model by Shukla et al [13] (postdelivery scenarios)[a].

Risk variable	Description and clinical relevance	Contribution to predictive accuracy
Birth weight	Neonatal weight at birth; strongest independent predictor	Highest importance; major impact on model AUC[b]
Bag and mask resuscitation	Need for resuscitation at birth using bag and mask ventilation	Moderate contribution; associated with increased risk
Cluster perinatal mortality	Perinatal mortality rate within the cluster or facility	Contributes to contextual risk stratification
Gestational age at enrollment	Weeks of gestation at the first prenatal visit	Adds modest predictive value
Maternal age	Age of the mother at delivery	Minor contribution to increased risk
Parity	Number of previous births	Minimal change in AUC; included in model
Multiple births	Twin or higher-order deliveries	Minor contribution
Antenatal corticosteroids	Use of steroids before delivery	Slight incremental predictive value
Maternal education: none	Lack of formal maternal education	Weak influence
Hospital delivery	Birth occurring in a hospital setting	Slight predictive value
Maternal antibiotics	Use of antibiotics during pregnancy	Negligible effect
Delivery by physician	Birth attended by a physician	Minor contribution

[a]The identified postdelivery variables, particularly birth weight and resuscitation requirements, contributed the most to the model's predictive accuracy, with an area under the curve exceeding 0.80. Prenatal-only variables (ie, maternal age, parity, and antenatal corticosteroids) contributed modestly. The table highlights variables that are considered most actionable for early identification of high-risk neonates in LMIC health facilities.

[b]AUC: area under the curve.

Significance of the Study

This study is significant for its potential to reduce neonatal mortality in Kenya, where neonatal deaths account for nearly half of all cases of child mortality under the age of 5 years, highlighting the urgent need for targeted interventions. Despite global and national efforts, progress toward SDG target 3.2, which aims to end preventable newborn deaths and deaths under the age of 5 years by 2030, remains slow, particularly in resource-limited settings. The study evaluates the feasibility, applicability, and usability of neonatal risk predictor variables incorporated into the ML model proposed by Shukla et al [13], which was initially validated in India. By testing these predictors in Kenyan MOH facilities, the study assesses their adaptability to local clinical and infrastructural contexts. Through a user-centered approach, it also explores health care professionals' experiences, perceptions, and interactions with the tool. If applicable, the model could enable early identification of high-risk neonates, support timely interventions, inform evidence-based policies, promote digital health adoption, and strengthen neonatal care systems, accelerating progress toward SDG 3.2.

Study Aims and Objectives

Overview

This study aims to assess the contextual feasibility, applicability, and usability of neonatal risk predictors incorporated into the ML model proposed by Shukla et al [13] for identifying high-risk neonates in the Kenyan context. The objective is to inform the model's potential adoption in LMIC settings in alignment with both national priorities and global goals, including SDG target 3.2. The study will evaluate the applicability of these predictors in Kenyan MOH facilities that provide neonatal care and will be conducted between August 2025 and November 2025.

Specific Objectives

The specific objectives of the study will be as follows:

- To assess the feasibility of implementing the neonatal risk predictors from the ML model by Shukla et al [13] within triage systems in Kenyan MOH facilities as an LMIC setting
- To determine the applicability of the neonatal risk predictor variables incorporated into the ML model to accurately identify high-risk neonates and their potential contribution to reducing neonatal mortality in line with SDG target 3.2
- To evaluate the usability and user experience (UX) of the ML model's neonatal risk predictors among health care professionals in identifying high-risk neonates

Methods

Overview

A mixed methods approach will be used, combining qualitative and quantitative methods to enhance generalizability and statistical rigor and strengthen the credibility of the findings [22]. The study will be conducted in 3 sequential phases. First, key informant interviews (KIIs) will be conducted to capture contextual insights. The second phase is an intervention period, during which a quantitative neonatal risk survey tool is integrated into clinical workflows to evaluate the contextual feasibility, applicability, and usability of the ML risk predictor variables proposed by Shukla et al [13]. Third, a

XSL•FO
RenderX

postimplementation questionnaire will be administered to assess whether the survey tool's intended outcomes were achieved. Phase 2 (intervention period) and phase 3 (postintervention survey) will follow a longitudinal design, observing the same research participants at 2 time points over a 4-month period to examine changes in outcomes associated with the intervention prior to implementation [23,24]. Therefore, the study will fully leverage the strengths of both qualitative and quantitative methods. These include triangulation, methodological flexibility, in-depth insights, and a comprehensive understanding of the research phenomenon, all of which enhance the validity of the study's findings compared with a single-method approach [25-27].

Phase 1: Conducting KIIs

This will be the preintervention phase, which will involve conducting KIIs with key informants (KIs), including facility heads and neonatal department heads at selected MOH facilities providing neonatal care. The KIIs will capture contextual insights into participants' lived experiences, current workflows, perceptions of neonatal risk interventions, potential benefits, anticipated implementation barriers, and preferred mode of deployment (paper based or electronic). These insights will enable the researchers to better understand the study context and stakeholder expectations, thereby informing the study's planning and execution.

Phase 2: Intervention Period

This will be the intervention phase, which will involve providing frontline neonatal care workers at selected MOH facilities with a paper-based neonatal risk assessment tool integrated into routine clinical workflows. Nominal-scale surveys will be used to operationalize and assess the applicability of the risk predictor variables incorporated into the ML model proposed by Shukla et al [13]. The tool will be deployed during routine neonatal care activities over a 4|-month period to enable real-world evaluation of its impact [28]. Longitudinal outcomes will be assessed through observations made during both the intervention and postintervention surveys, focusing on the tool's feasibility, applicability, usability, and utility. Findings from this phase will inform the potential feasibility of implementing the intervention across Kenyan MOH facilities.

Phase 3: Postintervention Survey

This will be the postintervention phase, during which quantitative surveys will be conducted with neonatal health workers. The primary objective is to assess whether the intervention's intended outcomes were achieved, gain a deeper understanding of implementation barriers, highlight areas for improvement, and inform future deployment strategies [29]. Both Likert and semantic differential scales will be used to evaluate the predictors' usability (ease of use and efficiency), UX (satisfaction and engagement), utility (perceived applicability for identifying high-risk neonates), and feasibility (practicality of implementation) within Kenyan MOH facilities. Both standardized and custom validation tools will be used to comprehensively assess the usability, UX, interaction satisfaction, and utility of the intervention [30-33]. The instruments will include the Usability Scale Questionnaire, Questionnaire for User Interaction Satisfaction, and a Post-Study Neonatal Risk Predictor Utility Questionnaire. The Usability Scale Questionnaire evaluates UX with the neonatal risk predictor variables using a 5-point Likert scale (1="strongly disagree," 2="disagree," 3="neutral," 4="agree," and 5="strongly agree"). The Questionnaire for User Interaction Satisfaction measures UX and interaction satisfaction using a semantic differential scale with bipolar adjectives such as "terrible"-"wonderful" and "difficult"-"easy." The Post-Study Neonatal Risk Predictor Utility Questionnaire, a custom questionnaire developed specifically for this study, assesses how well the tool meets user needs and its overall utility. It includes items on ease of use, clarity of instructions, and relevance of content and perceived effectiveness rated on a 7-point Likert scale (1="strongly agree," 2="agree," 3="slightly agree," 4="neutral," 5="slightly disagree," 6="disagree," and 7="strongly disagree").

The findings from these tools will provide a comprehensive understanding of the intervention's applicability and performance and support evidence-based decisions regarding the adoption and implementation of the proposed solution.

Target Population and Study Setting

The target population will consist of health care workers providing neonatal care services in Kenyan MOH facilities. Specifically, 3 health facilities will be selected from 3 different counties: Kibera Community Health Centre – Amref (Nairobi County), Nazareth Mission Hospital (Kiambu County), and Matuu Level 4 Hospital (Machakos County). To enhance the national representativeness of the study findings, these facilities are located in diverse contexts across Kenya, an LMIC setting, reflecting the unique environmental and socioeconomic conditions in which they operate.

Kibera Community Health Centre is a public, community-based facility located within Kibera, one of Africa's largest informal settlements characterized by widespread poverty and limited access to basic services, including health care. Nazareth Mission Hospital is a nonprofit, level 5 tertiary referral hospital in Kiambu County. It serves patients from across Kenya and neighboring regions, offering a wide range of specialized services. Matuu Level 4 Hospital, located in Machakos County, is a public level 4 health facility operating in a semiarid region characterized by low-income households, extreme temperatures, water scarcity, and food insecurity.

Sampling Procedure

Overview

Textbox 1 summarizes the inclusion and exclusion criteria, detailing the specific conditions that determine participant eligibility for inclusion in or exclusion from the study.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for study participants.

Inclusion criteria

- Aged 18 years or older
- Health care professionals (eg, clinicians, nurses, midwives, and pediatricians) directly involved in neonatal care
- Current assignment to neonatal care duties at one of the selected study facilities during the study period
- At least 6 months of continuous experience providing neonatal care
- Willingness and ability to provide informed consent

Exclusion criteria

- Temporarily rotating through neonatal units for training, internship, or short-term placement
- Exclusive assignment to nonneonatal departments (eg, outpatient, maternity only, or administrative roles) during the study period
- On extended leave or otherwise unavailable during the intervention period
- Prior direct involvement in the design or development of the neonatal risk assessment tool
- Less than 6 months of continuous experience providing neonatal care

Qualitative Sample

Participants will include KIs from the selected health facilities. These will primarily include senior staff (ie, heads of neonatal units or senior clinicians involved in neonatal care). Owing to their administrative roles and clinical experience, these individuals are well positioned to provide informed insights into the study context, including the perceived value, potential impact, and implementation barriers of the neonatal risk screening tool [34]. Owing to the limited number of eligible KIs per facility, a purposive sampling approach will be used. While qualitative sample sizes are commonly determined by data saturation [35,36], a minimum of 3 KIIs will be conducted, one in each study site (ie, Kibera, Nazareth, and Matuu). Restricting participation to 1 KI per facility is considered appropriate as qualitative findings will be supplemented with quantitative data.

Quantitative Sample

Potential participants will primarily comprise health care workers involved in providing neonatal care at the selected facilities, including neonatal pediatricians, clinicians, general practitioners, nurses, and midwives. Across the 3 study sites, 43 staff members provide neonatal care, as shown in Table 2. However, due to staffing constraints, the same personnel often rotate across multiple departments, including neonatal, maternity, and postnatal units, and therefore are not permanently stationed within neonatal units.

Table . Number of neonatal health care workers at the selected facilities[a].

Facility name	Qualitative sample (senior health care workers; n=3), n (%)	Quantitative sample (frontline health care workers), n (%) n=N1+N(e2)	
		Frontline care workers (n=43)	Required sample size, n (%)
Kibera Community Health Centre	1 (33.3)	11 (25.6)	11 (26.8)
Nazareth Mission Hospital	1 (33.3)	17 (39.5)	16 (39.0)
Matuu Level 4 Hospital	1 (33.3)	15 (34.9)	14 (34.1)

[a]The qualitative sample comprises purposively selected senior neonatal health care workers in leadership or supervisory roles. The quantitative sample includes frontline health care workers directly providing neonatal care at each facility during the study period. Sample totals reflect eligible and available participants at the time of data collection. Frontline health workers will collect data on the specified predictive variables from 400 newborns across the 3 selected facilities, distributed as follows: 100 in Kibera Community Health Centre, 100 in Nazareth Mission Hospital, and 200 in Matuu Level 4 Hospital.

Given the study's focus on assessing the applicability of neonatal risk parameters in routine clinical settings, only staff assigned to neonatal care units during the study period will be eligible to participate. On the basis of this criterion, a purposive sample of 10 health care workers will be recruited across the 3 sites: Kibera (n=4, 40%), Matuu (n=4, 40%), and Nazareth (n=2, 20%). This approach ensures the inclusion of staff actively providing neonatal care during the study period by integrating the proposed neonatal risk parameters into routine clinical workflows [37].

Furthermore, according to the literature, total population sampling is appropriate when the population with the specific characteristics of interest is very small [38-41], which is the case here. The researchers believe that neonatal care workers in the selected facilities are uniquely equipped with the knowledge, skills, experience, and exposure necessary to provide meaningful responses that can inform the study. The total number of frontline neonatal care workers across the 3 facilities is 43, as shown in Table 2. Applying the Yamane [42] formula to calculate the required quantitative sample size, the study

XSL•FO
RenderX

anticipates recruiting 11 participants from Kibera, 14 from Nazareth, and 16 from Matuu, yielding a total sample of 41 participants (Table 2).

Data Collection and Analysis

Data will be collected using both qualitative and quantitative methods, including KIIs, neonatal risk assessment tool and questionnaire, and postintervention questionnaires.

Data Capture and Management Plan

Overview

Data will be entered into a secure electronic database by trained personnel and subjected to routine verification and quality checks. Access to the data will be restricted to authorized study staff to ensure confidentiality, privacy, and data integrity. All data handling procedures will comply with Kenyan data protection laws governing the privacy and confidentiality of personal information. Paper-based records will be securely stored in locked cabinets at each facility, whereas electronic data will be encrypted and password protected. Study data will be used solely for research purposes and will not be accessible to individuals or entities outside the project team.

Qualitative Data Collection

KIIs will be conducted with KIs at the selected facilities in 2 stages. The initial KIIs will be carried out to collect baseline data aimed at understanding current workflows, existing challenges, and expectations related to the use of the neonatal risk predictors proposed in the ML model by Shukla et al [13]. Follow-up KIIs will also be conducted at the end of the intervention period to explore participants' lived experiences, perceived benefits, challenges encountered, and suggestions for improving the effectiveness of the predictors.

Quantitative Data Collection

The study will recruit 1 co–principal investigator from among neonatal care clinicians or staff at each selected facility. The co–principal investigators will be responsible for administering a paper-based survey tool (questionnaire) containing the neonatal risk predictors incorporated in the study by Shukla et al [13] on intrapartum stillbirth and neonatal mortality. The tool will be used by fellow staff members as part of routine neonatal care operations during the intervention period. These risk factors are considered significant neonatal predictors and are assessed at 2 time points: delivery (day 1) and after delivery (day 2). They are grouped into 5 key categories, namely, maternal, pregnancy, delivery, neonatal, and after delivery.

It is essential to note that the data collected during the study will not be used to input into the specific predictive ML model to generate actual risk scores. Instead, the study will focus on gathering input data to assess the feasibility, usability, and applicability of the risk predictors within the Kenyan context prior to full implementation of the ML prediction tool. After the intervention period, participants will be administered postsurvey questionnaires comprising closed-ended questions and several measurement scales. These surveys will collect feedback on the feasibility, usability, UX, and perceived utility of the neonatal risk predictors to inform future decisions regarding the tool's adoption.

Data Analysis

Qualitative data will be analyzed using thematic analysis to identify patterns and themes emerging from the KIIs. The analysis will follow an inductive coding approach involving systematic identification and classification of concepts, categories, and subcategories to generate overarching themes grounded in the data [5,43,44]. Data analysis will be conducted using the NVivo qualitative data analysis software (Lumivero). Thematic analysis is particularly valued for its flexibility and accessibility, making it a versatile qualitative research method. It enables in-depth exploration of complex datasets and supports the generation of rich, nuanced insights while remaining relatively straightforward to learn and implement [45].

Quantitative data obtained from the surveys will be analyzed using descriptive statistical techniques, including calculating frequencies and percentages and presenting results in tables and charts. Descriptive statistics are particularly useful for summarizing complex numerical data into manageable and interpretable formats, thereby facilitating the identification of key patterns and trends [46]. Data analysis will be conducted using Stata (StataCorp). This approach enhances clarity, improves communication of findings, and supports timely, evidence-based decision-making.

Ethical Considerations

Ethics approval (SU-ISERC2949/25) was received from the Strathmore University Institutional Scientific and Ethical Review Committee in Nairobi, Kenya, on July 28, 2025, prior to this study. Various ethical considerations related to key research stakeholders will be carefully addressed, including concerns involving the researcher, research participants, and sponsoring organization [47]. The study will strictly adhere to these ethical guidelines to ensure compliance and integrity throughout the research process. Informed consent was obtained before engaging participants, after explaining the aim, benefits, and risks associated with the study and the mechanisms put in place to address them. In addition, privacy and confidentiality were ensured by omitting identifying details.

Results

Status

The study was funded in March 2025. Six participants (frontline neonatal health workers) had been recruited by the time of initial submission of the manuscript (August 2025). Data analysis is ongoing and is expected to be concluded in March 2026. Publication of study findings is expected by June 2026.

Feasibility of Implementation

The research will provide a clear understanding of whether and how the neonatal risk predictors developed by Shukla et al [13] will be practically integrated into existing triage systems within Kenyan health care facilities. This includes identifying both enablers of and barriers to implementation in LMIC settings, assessing health system readiness, the availability of resources such as skilled personnel and infrastructure, and the level of institutional support for adoption.

XSL•FO
RenderX

Applicability and Clinical Impact

The study will explore the applicability and clinical impact of the tool in accurately identifying high-risk neonates. It will investigate whether early risk identification using the tool improves neonatal outcomes or enables more timely interventions. Ultimately, the research will assess the tool's potential to reduce neonatal mortality, aligning with SDG target 3.2, which aims to end preventable deaths of newborns.

Usability and UX

The study will assess the usability and UX of the ML model's neonatal risk predictors by examining how health care professionals interact with and perceive the usefulness of these variables in real-world clinical settings. This assessment will evaluate the predictors' learnability and efficiency and overall user satisfaction, as well as the tool's seamless integration into routine clinical workflows. Additionally, user feedback will provide valuable insights into the relevance, clarity, and acceptability of the predictors, helping inform potential improvements and ensure that they meet the practical needs of their intended users.

Research Impact, Use, and Dissemination of Results

Evaluating the utility of the ML tool's predictors is crucial for understanding their applicability in identifying neonatal risks. This evaluation will inform strategies for potential implementation and support the tool's sustained use, maximizing its impact and application in practice. The tool is expected to have a positive impact on neonatal care processes across health facilities in Kenya. Anticipated benefits include improved triage decisions, enhanced applicability in neonatal care delivery, increased operational efficiency, and greater productivity among health care providers.

Dissemination of the study findings will occur at 2 levels. First, preliminary results will be presented to the stakeholders at a local conference. Second, the results will be shared through 2 peer-reviewed scientific journal publications. Throughout the dissemination process, the project will adhere to key principles aligned with responsible publication practices. These include maintaining ethical research conduct throughout the study, respecting intellectual property rights by properly acknowledging original authors and their work, and upholding academic integrity. The latter involves conducting the publication process with honesty, fairness, transparency, respect, and accountability, ensuring that the research makes a credible contribution to the broader scientific and health care communities.

Discussion

Expected Findings

Neonatal health outcomes remain a significant public health concern in LMICs, where approximately 98% of global neonatal deaths occur due to preventable causes such as preterm birth, intrapartum complications, and infections [48]. SSA bears a disproportionate burden, with NMRs often exceeding 25 per 1000 live births [49]. Socioeconomic disparities are key contributors as infants born to mothers with low literacy and lower income levels are more vulnerable to poor health outcomes. Access to prenatal care and essential neonatal services is often hindered by financial constraints, transport challenges, and low health literacy [50].

Moreover, fragile health care systems in LMICs lack adequate infrastructure, including electricity, medical equipment, and essential medical commodities and supplies, which compromises the quality of neonatal care. Inadequate infection control further exacerbates neonatal risk, with sepsis remaining a leading cause of mortality. Despite these challenges, promising interventions have emerged, such as community health worker–led programs and decentralized neonatal care units, which have shown substantial reductions in mortality. Achieving SDG 3.2 will require integrated efforts focused on health system strengthening; equitable access to care; and targeted community-based strategies, including digital health interventions such as ML tools.

The impact of contextually informed usability studies in LMICs extends beyond tool optimization to strengthening health system resilience. Effective usability testing has enabled the scalable deployment of artificial intelligence–driven diagnostic aids in Uganda, where simplified interfaces have reduced clinician cognitive load and improved diagnostic accuracy by 25% [51,52]. Moreover, integrating usability feedback into national digital health policies, as piloted in Rwanda's health information exchange, fosters sustainable innovation [53]. Future efforts must prioritize capacity building for local usability researchers and leverage mixed methods data by combining quantitative metrics (such as task success rates) with qualitative insights to advocate for resource allocation. Ultimately, embedding usability science in LMIC health initiatives bridges this significant gap in technology adoption, advancing equity in global health delivery.

Several well-established frameworks guide the measurement of the feasibility of health care interventions such as digital health tools. Notably, the reach, effectiveness, adoption, implementation, and maintenance framework comprehensively addresses feasibility by evaluating the real-world effectiveness of interventions [54]. Additionally, the Medical Research Council framework for complex interventions provides guidance on the iterative assessment of interventions' practicality, effectiveness, and sustainability in diverse contexts [55]. Other frameworks include implementation mapping, which systematically connects implementation determinants with specific strategies, providing a structured approach to assessing feasibility and developing tailored solutions [56,57]. The Consolidated Framework for Implementation Research also provides a detailed framework for analyzing feasibility, considering elements such as intervention characteristics, inner and outer settings, and stakeholder engagement, often in conjunction with pragmatic tools such as the Pragmatic Context Assessment Tool and the Clinical Sustainability Assessment Tool [58].

Successful adoption of ML tools within health care systems is influenced by stakeholder engagement and organizational readiness. The literature highlights the critical role of clinician and organizational perceptions in determining the feasibility,

XSL•FO
RenderX

acceptability, and appropriateness of new technologies [59-61]. Complex systems thinking frameworks emphasize that feasibility assessments must account for dynamic interactions among technology, workflows, and human factors within health care environments [47,62,63]. Research indicates that early and sustained stakeholder engagement across clinical, administrative, and IT teams, as well as end user populations such as patients and caregivers, significantly enhances the feasibility of adopting ML tools. Additionally, assessing organizational readiness by evaluating knowledge, attitudes, and contextual factors (such as existing infrastructure and resource constraints) ensures that ML tools are tailored to local needs, thereby enhancing their practical integration and sustained use within health care systems [64-66].

The proposed evaluation of the ML model by Shukla et al [13] for neonatal risk prediction in Kenya is vital in informing its adoption and alignment with SDG target 3.2. It must address local challenges such as delayed triage, limited resources, inadequate staff training, and weak digital infrastructure. Mobile health platforms, already used in Kenyan CHW programs, present a scalable solution [67,68]. Integrating the tool with Kenya's electronic medical record systems could enable real-time alerts for high-risk neonates. A community validation in Kakamega County found that CHWs achieved 85% accuracy with digital triage tools after brief training [69]. Evaluating usability in Kenya may improve triage, referrals, and resource use, thus reducing neonatal mortality. This initiative can guide scale-up, infrastructure investment, and policy alignment, supporting digital health transformation in LMICs.

Strengths and Limitations of the Study

The project will adhere to established quality standards throughout its life cycle, encompassing planning, data collection and analysis, and reporting of findings. Quality assurance in research refers to the techniques and strategies used to ensure proper care and control during the research process [70]. To uphold quality assurance in research, the project will implement measures to enhance research credibility, accuracy, and transparency [71,72]. First, staff competence will be ensured by providing relevant training to equip personnel with the necessary skills to perform their roles effectively. Project planning and documentation will be transparent, with all essential information made accessible and understandable to stakeholders, promoting open communication. Standard operating procedures will be established to guide project execution, including clear objectives, prioritized activities, appropriate methodologies, clear communication, and the protection of participant and data confidentiality.

However, this study has several limitations inherent to its exploratory preimplementation design. Neonatal health care staff will collect data by integrating a paper-based checklist into routine workflows, which may introduce observer and Hawthorne effects. The absence of a control group and blinding due to limited staffing and purposive recruitment of active neonatal care providers restricts causal inference and may introduce bias, although participants' prior experience allows for meaningful before-and-after comparisons. To mitigate these limitations, methodological triangulation using both qualitative and quantitative approaches [73], including KIIs, surveys, and expert reviews, will be applied, alongside defined data quality standards. Data will be securely managed through restricted access, routine verification, encryption, and locked storage in compliance with Kenyan data protection laws. Ethical oversight, participant validation, expert review, and peer-reviewed publication will further strengthen the study's credibility and trustworthiness.

Acknowledgments

The authors would like to sincerely thank the UK Royal Academy of Engineering for funding the study; without this support, this work would have been impossible.

Funding

A Frontiers seed funding grant from the Royal Academy of Engineering has been received to support this study.

Data Availability

All data generated or analyzed during this study will be included in the published article (and its supplementary information files).

Authors' Contributions

RDN contributed to conceptualization, writing and editing—original draft, and overall coordination. PM reviewed the initial research study concept for examining, evaluating, and analyzing the manuscript idea to assess its quality, accuracy, relevance, and effectiveness. KE contributed to conceptualization and literature review. FS contributed to the design of the study tools, investigation, coordination, and data curation. BM contributed to conceptualization, literature review, and investigation. MSRJ contributed to conceptualization, review, editing, and manuscript structure refining.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Newborn mortality. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/newborn-mortality [accessed 2026-03-01]
2. Santhanam I, Moodley P, Jayaraman B, et al. Triage and resuscitation tools for low and middle income countries: how to catch the killer? Arch Dis Child Educ Pract Ed 2022 Feb;107(1):71-76. [doi: 10.1136/archdischild-2021-321981] [Medline: 34112664]
3. Robertson MA, Molyneux EM. Triage in the developing world--can it be done? Arch Dis Child 2001 Sep;85(3):208-213. [doi: 10.1136/adc.85.3.208] [Medline: 11517102]
4. Paediatric emergency triage, assessment and treatment: care of critically-ill children. : World Health Organization; 2016 URL: https://www.who.int/publications/i/item/9789241510219 [accessed 2026-03-01]
5. Nyatuka DR. A framework for cloud-based patient-centred health and wellbeing services in Kenya [PhD thesis]. : Cape Peninsula University of Technology; 2021 URL: https://africacommons.net/artifacts/20310564/a-framework-for-cloud-based-patient-centred-health-and-wellbeing-services-in-kenya/21211088/ [accessed 2026-03-01]
6. Milesky J, Rosen M, Sharma R, Zhang A, Bass EB. Acute care nursing staff shortages that compromise patient-to-nurse ratios: rapid response. In: Making Healthcare Safer IV: A Continuous Updating of Patient Safety Harms and Practices [Internet]: Agency for Healthcare Research and Quality (US); 2025. [Medline: 40498891]
7. Liu J, et al. The role of explainable artificial intelligence in human-in-the-loop healthcare. Artif Intell Med 2019;98:103-109. [doi: 10.1016/j.artmed.2019.05.017]
8. Imbo AE, Mbuthia EK, Ngotho DN. Determinants of neonatal mortality in Kenya: evidence from the Kenya Demographic and Health Survey 2014. Int J MCH AIDS 2021;10(2):287-295. [doi: 10.21106/ijma.508] [Medline: 34938597]
9. Levels & trends in child mortality. : United Nations Children's Fund; 2023 URL: https://data.unicef.org/wp-content/uploads/2024/03/UNICEF-2023-Child-Mortality-Report-1.pdf [accessed 2026-03-01]
10. Nugent R, et al. Economic effects of the double burden of malnutrition. Lancet 2019;395(10218):156-164. [doi: 10.1016/S0140-6736(19)32473-0]
11. Kenya Demographic and Health Survey 2022. Key indicators report. : Kenya National Bureau of Statistics and ICF; 2023 URL: https://dhsprogram.com/pubs/pdf/PR143/PR143.pdf [accessed 2026-03-01]
12. Wachira BW, Mwai M. A baseline review of the ability of hospitals in Kenya to provide emergency and critical care services for COVID-19 patients. Afr J Emerg Med 2021 Jun;11(2):213-217. [doi: 10.1016/j.afjem.2021.01.001] [Medline: 33495726]
13. Shukla VV, Eggleston B, Ambalavanan N, et al. Predictive modeling for perinatal mortality in resource-limited settings. JAMA Netw Open 2020 Nov 2;3(11):e2026750. [doi: 10.1001/jamanetworkopen.2020.26750] [Medline: 33206194]
14. Levenets O, et al. Coping with cancer in post-communist Europe: a systematic literature review. Health Policy Plan 2021;36(10):1690-1704. [doi: 10.1093/heapol/czab121]
15. Every newborn action plan. : World Health Organization; 2024 URL: https://www.who.int/initiatives/every-newborn-action-plan [accessed 2026-03-01]
16. Binagwaho A, et al. Rwanda's role in global health — lessons for COVID-19 and future pandemics. N Engl J Med 2020;382(25):2391-2402 [FREE Full text]
17. Kenya Demographic and Health Survey 2022. : Kenya National Bureau of Statistics; 2022 URL: https://www.knbs.or.ke/reports/kdhs-2022/ [accessed 2026-03-01]
18. Silva G, Wichmann RM, da Silva Junior FC, Chiavegatto Filho AD. Development and evaluation of machine learning training strategies for neonatal mortality prediction using multicountry data. Sci Rep 2025 Jul 7;15(1):24278. [doi: 10.1038/s41598-025-04066-5] [Medline: 40624031]
19. Onu CC, Udeogu I, Ndiomu E, et al. Ubenwa: cry-based diagnosis of birth asphyxia. arXiv. Preprint posted online on Nov 17, 2017. [doi: 10.48550/arXiv.1711.06405]
20. Piper S, et al. Implementing a digital health model of care in Australian youth mental health services: protocol for impact evaluation. BMC Health Serv Res 2021;21 [FREE Full text] [doi: 10.1186/s12913-021-06394-x]
21. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2014 Nov 15;13:8-17. [doi: 10.1016/j.csbj.2014.11.005] [Medline: 25750696]
22. Dever JA. Combining probability and nonprobability samples to form efficient hybrid estimates: an evaluation of the common support assumption. Presented at: 2018 Federal Committee on Statistical Methodology (FCSM) Research Conference; Mar 7-9, 2018 URL: https://nces.ed.gov/FCSM/pdf/A4_Dever_2018FCSM.pdf [accessed 2026-03-01]
23. Vogl S. Mixed methods longitudinal research. Forum Qual Soc Res 2023 Jan;24(1). [doi: 10.17169/fqs-24.1.4012]
24. Plano Clark VL, Anderson N, Wertz JA, Zhou Y, Schumacher K, Miaskowski C. Conceptualizing longitudinal mixed methods designs: a methodological review of health sciences research. J Mix Methods Res 2015;9(4):297-319. [doi: 10.1177/1558689814543563]
25. Tenny S, Brannan JM, Brannan GD. Qualitative study. In: StatPearls [Internet]: StatPearls Publishing; 2022.
26. Molina-Azorin JF. Mixed methods research: an opportunity to improve our studies and our research skills. Eur J Manag Bus Econ 2016 May;25(2):37-38. [doi: 10.1016/j.redeen.2016.05.001]
27. Mertens DM, Hesse-Biber S. Triangulation and mixed methods research: provocative positions. J Mix Methods Res 2012;6(2):75-79. [doi: 10.1177/1558689812437100]

XSL•FO
RenderX

28. Ray-Sannerud BN, Dolan DC, Morrow CE, et al. Longitudinal outcomes after brief behavioral health intervention in an integrated primary care clinic. Fam Syst Health 2012 Mar;30(1):60-71. [doi: 10.1037/a0027029] [Medline: 22288398]
29. Newall N, Lewin G, Boldy D. The importance of collecting pre-and post-intervention practice data when conducting a randomised controlled trial. Wound Pract Res 2016;24(1):7-13 [FREE Full text]
30. Chin JP, Diehl VA, Norman LK. Development of an instrument measuring user satisfaction of the human-computer interface. Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 15-19, 1988. [doi: 10.1145/57167.57203]
31. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, Weerdmeester B, McClelland IL, editors. Usability Evaluation in Industry: CRC Press; 1996:189-194.
32. Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int J Hum Comput Interact 1995;7(1):57-78. [doi: 10.1080/10447319509526110]
33. Lewis JR. The System Usability Scale: past, present, and future. Int J Hum Comput Interact 2018;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
34. Kibuacha F. Key informant interviews: an in-depth guide for researchers. GeoPoll. 2024. URL: https://www.geopoll.com/blog/key-informant-interviews/ [accessed 2026-03-01]
35. Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. Field Methods 2006;18(1):59-82. [doi: 10.1177/1525822X05279903]
36. Kerr C, Nixon A, Wild D. Assessing and demonstrating data saturation in qualitative inquiry supporting patient-reported outcomes research. Expert Rev Pharmacoecon Outcomes Res 2010 Jun;10(3):269-281. [doi: 10.1586/erp.10.30] [Medline: 20545592]
37. Rai N, Thapa B. A study on purposive sampling method in research. Kathmandu Sch Law 2015;5:8-15 [FREE Full text]
38. de Ruyter K, Scholl N. Positioning qualitative market research: reflections from theory and practice. Qual Mark Res 1998 Apr 1;1(1):7-14. [doi: 10.1108/13522759810197550]
39. Etikan I, Musa SA, Alkassim RS. Comparison of convenience sampling and purposive sampling. Am J Theor Appl Stat 2015;5(1):1-4. [doi: 10.11648/j.ajtas.20160501.11]
40. Patton MQ. Qualitative Research & Evaluation Methods, 3rd edition: Sage Publications; 2002.
41. Purposive sampling. Laerd Dissertation. URL: https://dissertation.laerd.com/purposive-sampling.php [accessed 2026-03-01]
42. Yamane T. Elementary Sampling Theory: Prentice Hall; 1967.
43. Elliott V. Thinking about the coding process in qualitative data analysis. Qual Report 2018;23(11):2850-2861. [doi: 10.46743/2160-3715/2018.3560]
44. Braun V, Clarke V. What can “thematic analysis” offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014 Oct 16;9:26152. [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
45. Dawadi S. Thematic analysis approach: a step-by-step guide for ELT research practitioners. J NELTA 2020;25(1-2):62-71. [doi: 10.3126/nelta.v25i1-2.49731]
46. Rani R. Graphical representation of data. 2020. URL: https://byjus.com/maths/graphical-representation
47. Rahman Jabin MS, Steen M, Wepa D, Bergman P. Assessing the healthcare quality issues for digital incident reporting in Sweden: incident reports analysis. Digit Health 2023 May;9:20552076231174307. [doi: 10.1177/20552076231174307] [Medline: 37188073]
48. Oza S, Lawn JE, Hogan DR, et al. Cause-of-death estimates for neonatal periods: 2000–2013. 2014 URL: https://pubmed.ncbi.nlm.nih.gov/33292564/
49. Tolossa T, Fekadu G, Mengist B, Mulisa D, Fetensa G, Bekele D. Impact of antenatal care on neonatal mortality among neonates in Ethiopia: a systematic review and meta-analysis. Arch Public Health 2020 Nov 10;78(1):114. [doi: 10.1186/s13690-020-00499-8] [Medline: 33292564]
50. Islam MA, Biswas B. Socio-economic factors associated with increased neonatal mortality: A mixed-method study of Bangladesh and 20 other developing countries based on demographic and health survey data. Clinical Epidemiology & Global Health 2021;11:100801. [doi: 10.1016/j.cegh.2021.100801]
51. Odera P, et al. AI diagnostics and usability in low-resource hospitals. Nat Med 2024;30(1):89-97.
52. Pan D, Nilsson E, Rahman Jabin MS. A review of incidents related to health information technology in Swedish healthcare to characterise system issues as a basis for improvement in clinical practice. Health Informatics J 2024;30(3):14604582241270742. [doi: 10.1177/14604582241270742] [Medline: 39116887]
53. Ngabo F, Ndiaye S. Development, multi-institutional external validation, and algorithmic audit of an artificial intelligence-based Side-specific Extra-Prostatic Extension Risk Assessment tool (SEPERA) for patients undergoing radical prostatectomy: a retrospective cohort study. Lancet Digit Health 2023;5(7):e435-e445.
54. Glasgow RE, Harden SM, Gaglio B, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019 Mar 29;7:64. [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
55. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655. [doi: 10.1136/bmj.a1655] [Medline: 18824488]
56. Fernandez ME, Ten Hoor GA, van Lieshout S, et al. Implementation mapping: using intervention mapping to develop implementation strategies. Front Public Health 2019 Jun 18;7:158. [doi: 10.3389/fpubh.2019.00158] [Medline: 31275915]

XSL•FO
RenderX

57. Jabin MS. The need for a refined classification system and national incident reporting system for health information technology-related incidents. Front Digit Health 2024 Jul 26;6:1422396. [doi: 10.3389/fdgth.2024.1422396] [Medline: 39131183]
58. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4:50. [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
59. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
60. Jabin MS. Operational disruption in healthcare associated with software functionality issue due to software security patching: a case report. Front Digit Health 2024 Mar 14;6:1367431. [doi: 10.3389/fdgth.2024.1367431] [Medline: 38550716]
61. Braithwaite J, Churruca K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation science: a theoretical and empirical analysis of systems change. BMC Med 2018 Apr 30;16(1):63. [doi: 10.1186/s12916-018-1057-z] [Medline: 29706132]
62. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
63. University Corporate Services. University of Reading. URL: https://www.reading.ac.uk/about/governance/professional-services-directorates/university-corporate-services [accessed 2026-03-02]
64. Jabin MS, Wepa D, Hassoun A. A case report of system configuration issue in medical imaging due to system upgrade-changes in hardware and software. Front Digit Health 2024 Sep 2;6:1371761. [doi: 10.3389/fdgth.2024.1371761] [Medline: 39347445]
65. Reynolds J, Kizito J, Ezumah N, Mangesho P, Allen E, Chandler C. Quality assurance of qualitative research: a review of the discourse. Health Res Policy Syst 2011 Dec 19;9:43. [doi: 10.1186/1478-4505-9-43] [Medline: 22182674]
66. Jabin MSR, Nilsson E, Nilsson AL, Bergman P, Jokela P. Digital health testbeds in Sweden: an exploratory study. Digit Health 2022 Feb 14;8:20552076221075194. [doi: 10.1177/20552076221075194] [Medline: 35186314]
67. Quality assurance and enhancement. University of Stirling. URL: https://www.stir.ac.uk/about/professional-services/student-academic-and-corporate-services/academic-registry/quality-assurance-and-enhancement/# [accessed 2026-03-02]
68. Wepa D, Thomas S, Jabin MS. The experience and impact of digital technologies on indigenous populations in New Zealand during the COVID-19 pandemic and Cyclone Gabrielle: the Kaupapa Māori methodology. JMIR Aging 2025 Oct 28;8:e73974. [doi: 10.2196/73974] [Medline: 41150869]
69. Jabin MS, Pan D, Nilsson E. Characterizing patient details-related challenges from health information technology-related incident reports from Swedish healthcare. Front Digit Health 2024 Feb 6;6:1260521. [doi: 10.3389/fdgth.2024.1260521] [Medline: 38380372]
70. Noble H, Heale R. Triangulation in research, with examples. Evid Based Nurs 2019 Jul;22(3):67-68. [doi: 10.1136/ebnurs-2019-103145] [Medline: 31201209]
71. Muinga N, et al. Designing a mobile health solution to facilitate the transition from hospital to home care for mothers and neonates in Kenya: user-centered design study. JMIR Mhealth Uhealth 2020;8(3):e15586 [FREE Full text]
72. Ouma P, et al. Development and validation of a machine-learning model to predict mortality and hospital admission in a population with suspected COVID-19: a prospective cohort study. Lancet Digital Health 2022;4(4):e214-e224. [doi: 10.1016/S2589-7500(22)00008-8]
73. Kumar R. Research Methodology: A Step-by-Step Guide for Beginners, 3rd edition: SAGE Publications; 2010.

Abbreviations

KI: key informant
KII: key informant interview
LMIC: low- or middle-income country
ML: machine learning
MOH: Ministry of Health
NMR: neonatal mortality rate
SDG: Sustainable Development Goal
SSA: sub-Saharan Africa
UX: user experience

XSL•FO
RenderX

Edited by J Sarvestan; submitted 27.Aug.2025; peer-reviewed by JK Schaffzin, S Prakash; revised version received 03.Feb.2026; accepted 05.Feb.2026; published 27.Mar.2026.

Please cite as:
Nyatuka RD, Macharia P, Esther K, Siva F, Muriithi B, Rahman Jabin MS
Kenyan Neonatal Mortality Risk Predictor: Protocol for a User-Centered Design Evaluation
JMIR Res Protoc 2026;15:e81996
URL: https://www.researchprotocols.org/2026/1/e81996
doi:10.2196/81996

© Ronald Danny Nyatuka, Paul Macharia, Kakhata Esther, Faith Siva, Betsy Muriithi, Md Shafiqur Rahman Jabin. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparative Evaluation of Effectiveness of Oxygen-Releasing Gel With Scaling and Root Planing (SRP) With That of Scaling and Root Planing Alone in the Management of Chronic Periodontitis Patients: Study Protocol for Cross-Sectional Survey

Geeta Bhandari, BDS; Priyanka Jaiswal, MDS; Shweta Bhagat, BDS; Sakshi Kotecha, BDS

Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research (DMIHER(DU)), Sawangi, Wardha, India

Corresponding Author:
Geeta Bhandari, BDS
Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research (DMIHER(DU)), Sawangi, Wardha, India

Abstract

Background: Gingival recession and periodontal pockets are the results of the gradual deterioration of the alveolar bone and periodontal ligament caused by periodontitis, which is caused by inflammation of the tissues supporting the tooth. The primary goal of periodontal therapy is to eradicate these aberrant traits. Despite being a popular treatment, scaling and root planing (SRP) has limitations, such as difficulty accessing deeper pockets and root concavities.

Objective: The aim of this study is to assess the effectiveness of oxygen-releasing gel as a supplement to nonsurgical treatment for chronic periodontitis.

Methods: Twenty-two systemically healthy people with sites of chronic periodontitis were the participants of a split-mouth randomized controlled experiment. For every patient, two locations were chosen and divided into two groups at random. In total, 44 sites were chosen and divided into two groups: Group I (the test group) consisted of 22 test sites, while Group II (the control group) consisted of 22 control sites. While group II underwent SRP alone, group I underwent SRP followed by the use of BlueM oral gel. Plaque index, full mouth bleeding score, clinical attachment level, and probing pocket depth were measured at baseline, one month, and three months, and were compared appropriately.

Results: All the clinical measurements will be recorded. It can be anticipated that SRP followed by the use of BlueM oral gel will provide additional benefits in the management of chronic periodontitis by promoting better healing and clinical outcomes.

Conclusions: Within the constraints of this study, it can be anticipated that by encouraging improved healing and clinical results in chronic periodontitis, oxygen-releasing gel is more advantageous when used in conjunction with SRP than when SRP is done alone.

Trial Registration: Clinical Trials Registry India CTRI/2025/05/086768; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=MTI2NTA1&Enc=&userName=

International Registered Report Identifier (IRRID): DERR1-10.2196/75795

(JMIR Res Protoc 2026;15:e75795) doi:10.2196/75795

KEYWORDS

chronic periodontitis; oxygen-releasing gel; antimicrobial properties; wound healing; tissue regeneration; blue m gel; scaling and root planing

Introduction

Background

The hallmark of Stage II Grade A periodontitis, by Panos N. Papapanou et al, a chronic multifactorial inflammatory disease associated with dysbiotic plaque biofilms, is the progressive deterioration of the tooth-supporting framework. Its primary features include gingival bleeding, periodontal pocketing, loss of alveolar bone as seen on radiography, and the loss of periodontal tissue support, as seen by clinical attachment level (CAL) [1]. Removing bacteria and halting the disease's progression are two primary goals of periodontal therapy. Scaling and root planing (SRP) and other conventional nonsurgical procedures aim to permanently disrupt the biofilm and its maintenance [2]. As it is difficult to lower the bacterial burden because of the complex structure of the root and the location of the deposits, locally injected antimicrobial medications are used either alone or in conjunction with conventional therapy to target the areas of localized periodontal

XSL•FO
RenderX

degeneration [3]. It has been recognized as a means of reducing the emergence of resistance, allergic responses, and systemic issues in contrast to systemic antibiotics [4]. At least 700 bacterial species have been found in the oral cavity, making it one of the most complex and varied human microbiomes. A number of these microorganisms are crucial for preserving the health and regular operation of the mouth. Oral conditions like caries, gingivitis, and periodontal disease can result from changes in the ecology of the oral microbiota [5]. Pathogenic bacteria like Capnocytophaga, Leptotrichia, and Fusobacterium can flourish in the anoxic conditions at the biofilm center, whereas oxidative species like Haemophilus, Aggregatibacter, and Neisseriaceae inhabit the oxygen-rich periphery [6]. Gram-negative bacteria, including *Porphyromonas gingivalis*, *Tannerella forsythia*, and *Treponema denticola*, can control the subgingival biofilm in anoxic periodontal pockets due to interactions between proteins and lectins and carbohydrates. Bacterial cells must communicate and signal one another to form biofilms [6]. The immune system and periodontal pathogens interact intricately to define the activity of periodontal disease. The onset and progression of periodontitis have long been linked to changes in the subgingival microbiome. Changes in the polymicrobial population cause disruptions in host homeostasis in vulnerable individuals, and a number of bacteria are commonly linked to periodontal diseases [7]. In the subgingival region, gram-negative anaerobic bacteria interact with the host's inflammatory responses to create a hypoxic or reduced oxygen environment. Therefore, oxygen is a necessary molecule for life. A variety of physiological or pathological conditions are linked to variations in tissue oxygen requirements [8]. Hypoxia triggers a local defense or protective cellular responses in a chronic inflammatory illness. Such hypoxic reactions could be the pathophysiology of inflammation, which contributes to the pathogenesis and progression of the disease, if the cause of inflammation cannot be eliminated [9]. In the Netherlands, a group of dental surgeons under the direction of Dr. Peter Blijdorp created a product that relied on the production of active oxygen [10]. By increasing the oxygen content at the wound site, this oxygen-releasing gel helps to speed up the healing process. Additionally, this high concentration of active oxygen aids in wound healing, bleeding gums, and pocket depth reduction. Using this special mixture lowers the risk of irritation and infections while also improving a person's dental hygiene. Oxygen is essential for many wound healing processes, including collagen synthesis, bacterial oxidative death, re-epithelialization, and angiogenesis [11]. In order to encourage wound healing, topical and hyperbaric oxygen have both been studied as therapeutic approaches. Therefore, the primary goals of topical oxygen therapy using active oxygen-releasing gel are to promote wound healing and avoid dysbiosis [12]. The purpose of this study is to evaluate how well oxygen-releasing gel works as a local medication to treat chronic periodontitis [13].

Objectives

1. To compare the reduction in probing depth following scaling and root planing (SRP) with and without the application of oxygen-releasing gel in the treatment of chronic periodontitis
2. To evaluate the gain in clinical attachment level after SRP with and without the adjunctive use of oxygen-releasing gel
3. To assess the reduction in Full Mouth Bleeding Score (FMBS) as an indicator of improved periodontal health following SRP with and without the oxygen-releasing gel

Methods

Overview

This study aims to compare the clinical and microbiological outcomes of SRP with and without the adjunctive use of an oxygen-releasing gel in the treatment of chronic periodontitis.

Recruitment

Participants in this clinical trial will be those receiving treatment at the Department of Periodontics and Implantology of Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, or those referred for treatment at this institute, who have periodontal pockets larger than or equal to 4 mm in at least ten permanent teeth. All clinical procedures will be performed here. It is thought that this will enhance trial adherence and recruitment because they will already be receiving treatment.

Inclusion Criteria

Participants eligible for the study were motivated to undergo therapy and complete follow-up visits. Individuals who provided written informed consent were included. Participants were required to have a minimum of 15 natural teeth and be in good physical and systemic health. Only patients with clinical attachment loss ≥1 mm and diagnosed with chronic generalized periodontitis, presenting with at least 2–3 sites exhibiting periodontal pocket depths ≥4 mm, were selected for the study.

Exclusion Criteria

Patients who had undergone periodontal therapy within the previous 6 months or had received antibiotic therapy during the preceding 3 months were excluded. Individuals who were pregnant or lactating, those with known allergies to the study medications, and regular users of mouthwash were also excluded. Additionally, chronic alcohol users were not included in the study.

Sample size estimations

The sample size was calculated using the formula for comparison of two independent means:

$$n_1 = \frac{(\sigma_1^2 + \sigma_2^2/\kappa)(Z_{1-\alpha/2} + Z_{1-\beta})^2}{\Delta^2}, \quad n_2 = \kappa n_1$$

where n_1 and n_2 represent the sample sizes of groups 1 and 2, respectively; σ_1 and σ_2 denote the standard deviations of groups 1 and 2; Δ represents the expected difference in group means; κ is the allocation ratio (n_2/n_1); $Z_{1-\alpha/2}$ corresponds to the standard normal deviate for a two-sided significance level (1.96 for a 95% confidence interval); and $Z_{1-\beta}$ represents the standard normal deviate corresponding to the study power (0.84 for 80% power).

Based on previous findings reported by Anuradha et al [14], the mean Gingival Index was 0.52 in group I and 0.88 in group

XSL•FO
RenderX

II, with standard deviations of 0.51 and 0.33, respectively. The expected mean difference (Δ) was therefore calculated as 0.36 (0.88 − 0.52). Assuming an allocation ratio (κ) of 1, a power of 80%, and a significance level of 5% (two-sided), the calculated sample size was 22 participants per group.

Randomization

44 participants each will receive a different treatment as allocated to Groups 1 and 2 after scaling and root planing are completed. For the participant referring to Group I, oxygen-releasing gel will be administered locally, and for Group II, the supra and subgingival areas will be treated with SRP therapy alone. The flip coin approach will be used for randomization in order to distribute participants into Groups I and II. A separate researcher who is not involved in the methods will create the allocation sequence.

Procedure

Participants will be enrolled and randomly assigned to interventions by the researchers in charge of the therapy application. Researchers will undergo training before data collection and evaluations, following the procedures and outcome measures outlined. The UNC-15 probe, which measures plaque index (PI), FMBS, Probing Pocket Depth (PPD), and CAL, will be used to evaluate participants' periodontal clinical characteristics. Every patient will get instructions on the proper brushing technique as well as a demonstration of good oral hygiene practices.

At the first appointment, each patient will receive full-mouth supragingival scaling. After 1 week, on the second appointment, under local anesthesia of 2% lidocaine with 1:100,000 epinephrine, subgingival root planing will be done using Gracey curettes (Hu-Friedy). Following subgingival SRP, the experimental site will be segregated for patients in Group I. Oxygen-releasing gel will be administered in all sites with a probing depth of ≥4 mm over the course of five to ten minutes [7].

Regarding patients in Group II, they will be treated with subgingival SRP alone.

Statistical Analysis

The calculation of mean and SD will be done, and the mean data will be analyzed using a standard statistical method to determine statistical significance. Examiners performing outcome assessments are blinded to group allocation. The study includes assessor blinding. Outcome assessors will be unaware of participant group allocation to reduce potential bias. All assessments will be conducted using standardized tools to maintain consistency and objectivity. Data from the baseline, one month, and three months for each group will be evaluated using descriptive and inferential statistics, including the χ^2 test, Student paired and unpaired t test, one-way ANOVA, and Tukey test. The analysis will be conducted using SPSS (version 27.0; IBM Corp) and GraphPad Prism (version 7.0) software, with a significance threshold of $P<.05$. A P value below 0.05 will be regarded as significant, but a P value above 0.05 will be regarded as not statistically significant.

Evaluation Outcomes

The primary outcome of this study is to measure probing pocket depth, and secondary outcomes will be PI, FMBS, and CAL.

Secondary outcomes would complement the primary clinical measurements, providing a more comprehensive understanding of how the oxygen-releasing gel contributes to the overall treatment of chronic periodontitis when added to SRP.

Ethical Considerations

Ethics committee approval was obtained for this study (Ref No. DMIHER(DU)/IEC/2025/507) from the Datta Meghe Institute of Higher Education & Research (Deemed To Be University). A written informed consent form approved by the university's institutional ethics committee must be read, comprehended, and signed by research participants. The consent form will be obtained by the same researcher, a dentist, who is in charge of the treatments. The researcher will go over all the steps that need to be taken, the trial's goal, the advantages and disadvantages of taking part in the study, and the fact that volunteers are free to decide whether or not to continue with the study at any point without facing any consequences. Participants will be provided with an informed consent form, clearly describing the study's aim, procedures, risks, benefits, and their right to opt out at any time. Consent for publication of data will be obtained and will be used for all analyses, including secondary analyses. We will confirm that the original consent and institutional review board approval explicitly cover the use of primary and secondary data. Collection of data and entering the data into the database for screening and randomization will be done by the primary investigator. The privacy of prospective and enrolled participants' personal information will be upheld to ensure confidentiality before, during, and after the trial. We ensure that the data collected will be fully anonymized, with no personally identifiable information retained or linked to any participant. As there are no interventions involved and as only a cross-sectional survey will be conducted, in which the data will be collected via an interview and anthropometric measurements, participants will not be provided any compensation in this study.

Results

The study was not funded. Recruitment began in April 2025 and is ongoing. As of January 2026, 22 participants have been enrolled. Data collection is ongoing, and statistical analysis has not yet been completed. The results are expected to be published in Winter 2026. All the clinical measurements will be recorded at the baseline, 1 month, and 3 months. PIFMBS, PPD, and CAL will be used to assess the patient's dental hygiene state. Oxygen-releasing gel may exhibit a higher reduction in clinical parameters.

Discussion

Principal Findings

In recent years, adjunctive therapies have been explored to enhance the effects of SRP. One promising approach involves the use of oxygen-releasing gels, which have been shown to

XSL•FO
RenderX

possess antimicrobial and tissue-healing properties. These gels release oxygen upon contact with moisture in the periodontal pocket, creating an unfavorable environment for anaerobic bacteria while promoting oxygenation of the periodontal tissues, which may aid in healing and tissue regeneration. The increased oxygen concentration may also help reduce the inflammatory response and facilitate faster wound healing.

A new oxygen-rich mixture called BlueM gel was developed by a team of dentists in the Netherlands. Its main components, sodium perborate and honey enzymes, have combined to create amazing results. This composition is unique in its ability to provide wounded tissues with therapeutically high oxygen concentrations. Along with its oxygen-releasing properties, BlueM gel contains compounds that have antibacterial properties. Its strong bactericidal and bacteriostatic qualities, along with its well-known gradual release characteristic—that is, substantivity within the oral cavity—are what give it its effectiveness [14]. Han et al claim that this gel promotes neovascularization, speeds up the production of new blood cells, and increases stem cell synthesis to produce new fibroblasts [15]. Aqua, alcohol, glycerine, silica, sodium saccharin, sodium perborate, citric acid, PEG-32, sodium gluconate, lactoferrin, xanthan gum, and cellulose gum are the components of the BlueM oral gel recipe, which was developed under Peter Blijdrop's direction. Each of these chemicals has a specific purpose. BlueM products use oxygen as an active ingredient. Studies show that toothpaste with lactoferrin and active oxygen has similar antiplaque and antigingivitic properties as toothpaste with triclosan. Furthermore, the application of oxygen is essential for a number of healing processes, such as collagen synthesis, angiogenesis, and re-epithelialization [16-18]. Molecular oxygen, one of the wound's most vital nutrients, is necessary for the healing process [19]. For collagen to be released from cells during the reparative process, oxygen acts as a substrate for the hydroxylation of proline and lysine during collagen synthesis [20]. Oxygen also affects the rates of epithelialization and angiogenesis. When oxygen is given at 1 - 2 atmospheres absolute, the oxygen supply regulates the rate of epithelialization in both ischemic and normal wounds [10]. An oxygen gradient stimulates angiogenesis; high arterial PO2~ drives angiogenesis toward areas with low oxygen levels [21]. Molecular oxygen is necessary for leukocytes' oxidative bacterial killing processes. Bacterial membranes are destroyed by the production of free radicals originating from oxygen. Local oxygen tension is intimately correlated with the rate at which harmful radicals are produced and the capacity to kill through oxidative processes [22]. Investigating topical oxygen treatment may help to lessen wound inflammation and hasten the healing process. According to research, topically administering oxygen promotes the generation of reactive oxygen species (ROS), which are directly related to oxygen levels. Because it may readily pass through cellular membranes, hydrogen peroxide (H2O2) is one of the most prevalent and stable ROS. In tiny quantities, it acts as an intracellular messenger. This property of H2O2 promotes collagen production, stimulates the expression of growth factors such asVEGF (vascular endothelial growth factor) and EGF (epidermal growth factor), controls TGF-β signaling (transforming growth factor–beta), and facilitates neutrophil movement [23]. This study's findings were comparable to those of a survey by Anuradha et al [14], which compared the effects of curcumin gel as an adjuvant to SRP (Test group) and SRP (Control group). Additionally, Doungudomdacha et al [24] showed statistically insignificant changes in PI and GI (gingival index) in both groups when they compared patients with periodontitis with those with periodontally healthy locations. On the other hand, Koul et al's survey [10] had different findings that mean difference in GI was greater in the test group than in the control group, even though there was no statistically significant difference in PI between the control and test groups when compared intergroup.

Oxygen-releasing gels provide a slow and controlled release of molecular oxygen directly to the affected area, promoting healing and inhibiting anaerobic bacteria with minimal tissue irritation. In contrast, ozone therapy uses ozone gas (O_3), a highly reactive form of oxygen, which rapidly kills microbes but carries risks of toxicity if not carefully handled [25]. Hydrogen peroxide irrigation works by rapidly releasing oxygen and reactive species upon contact with tissues, which can help clean wounds but may damage healthy cells if used excessively. Unlike ozone and H2O2, oxygen-releasing gels offer a sustained, gentler oxygen supply, making them safer for long-term or repeated use [26].

Strengths of This Study

A randomized controlled trial design, which is regarded as the gold standard for clinical research, is used in this investigation. Dividing individuals into the experimental and control groups at random lowers biases and improves the reproducibility of the findings. Each participant can function as their control thanks to the split-mouth design. As each patient receives both the normal treatment (SRP) and the experimental treatment (oxygen-releasing gel) on separate sides of their mouth, this reduces interpatient variability. Measurable clinical outcomes that are objective indicators of periodontal health and treatment efficacy, like the PI, FMBS, PPD, and CAL, are probably used in this study. This makes it possible to evaluate the oxygen-releasing gel's effectiveness with clarity.

Limitations

Although a split-mouth design minimizes patient variability, clinical assessment of outcomes, such as probing depth or attachment level, can be prone to bias. The study may be limited in terms of its generalizability to all patients with chronic periodontitis. Factors such as age, severity of the disease, and comorbidities might influence the applicability of the results to other populations or settings. If there are other factors influencing the outcome (eg, smoking, systemic health conditions), the results could be confounded.

Conclusion

It can be anticipated that oxygen-releasing gel has the potential to be employed as a safe and effective adjunct to nonsurgical periodontal therapy for the treatment of chronic periodontitis.

XSL•FO
RenderX

Acknowledgments

We gratefully acknowledge the Department of Periodontics for providing logistical support throughout the conduct of this study.

Funding

All authors declare that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this paper.

Authors' Contributions

Conceptualization: GB
Data curation: GB
Formal analysis: PS
Investigation: GB
Methodology: GB
Project administration: SB
Resources: SB
Supervision: SB
Validation: PS
Visualization: SK
Writing – original draft: GB
Writing – review & editing: PS, SK

Conflicts of Interest

None declared.

References

1. Papapanou PN, Sanz M, Buduneli N, et al. Periodontitis: Consensus report of workgroup 2 of the 2017 World Workshop on the Classification of Periodontal and Peri‐Implant Diseases and Conditions. J Periodontol 2018 Jun;89(S1). [doi: 10.1002/JPER.17-0721]
2. Boutaga K, van Winkelhoff AJ, Vandenbroucke-Grauls C, Savelkoul PHM. Comparison of real-time PCR and culture for detection of Porphyromonas gingivalis in subgingival plaque samples. J Clin Microbiol 2003 Nov;41(11):4950-4954. [doi: 10.1128/JCM.41.11.4950-4954.2003] [Medline: 14605122]
3. Sadaf N, Anoop B, Dakshina B, Shweta B. Evaluation of efficacy of tetracycline fibers in conjunction with scaling and root planing in patients with chronic periodontitis. J Indian Soc Periodontol 2012 Jul;16(3):392-397. [doi: 10.4103/0972-124X.100918] [Medline: 23162335]
4. Unsal E, Akkaya M, Walsh TF. Influence of a single application of subgingival chlorhexidine gel or tetracycline paste on the clinical parameters of adult periodontitis patients. J Clin Periodontol 1994 May;21(5):351-355. [doi: 10.1111/j.1600-051x.1994.tb00725.x] [Medline: 8034781]
5. Genco RJ, LaMonte MJ, McSkimming DI, et al. The subgingival microbiome relationship to periodontal disease in older women. J Dent Res 2019 Aug;98(9):975-984. [doi: 10.1177/0022034519860449] [Medline: 31329044]
6. Leventis M, Deliberador T, Alshehri F, Alghamdi H. Topical oxygen therapy as a novel strategy to promote wound healing and control the bacteria in implantology, oral surgery and periodontology: a review. Saudi Dent J 2024 Jun;36(6):841-854. [doi: 10.1016/j.sdentj.2024.04.004] [Medline: 38883907]
7. Plemmenos G, Piperi C. Pathogenic molecular mechanisms in periodontitis and peri-implantitis: role of advanced glycation end products. Life (Basel) 2022 Jan 30;12(2):218. [doi: 10.3390/life12020218] [Medline: 35207505]
8. Wang XX, Chen Y, Leung WK. Role of the hypoxia-inducible factor in periodontal inflammation. In: Hypoxia and Human Diseases: InTechOpen; 2017. [doi: 10.5772/66037]
9. Semenza GL. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. Annu Rev Pathol Mech Dis 2014 Jan 24;9(1):47-71. [doi: 10.1146/annurev-pathol-012513-104720]
10. Koul A, Kabra R, Chopra R, Sharma N, Sekhar V. Comparative evaluation of oxygen releasing formula (Blue-M Gel®) and chlorhexidine gel as an adjunct with scaling and root planing in the management of patients with chronic periodontitis –A clinico-microbiological study. JDS 2020 Mar 28;7(2):111-117. [doi: 10.18231/j.jds.2019.026]
11. Knighton DR, Hunt TK, Scheuenstuhl H, Halliday BJ, Werb Z, Banda MJ. Oxygen tension regulates the expression of angiogenesis factor by macrophages. Science 1983 Sep 23;221(4617):1283-1285. [doi: 10.1126/science.6612342]
12. Niveda R, Kaarthikeyan G. Effect of oxygen releasing oral gel compared to chlorhexidine gel in the treatment of periodontitis. JPRI 2020 Aug 26:75-82. [doi: 10.9734/jpri/2020/v32i1930711]

XSL•FO
RenderX

13. Varghese SS, Thomas H, Jayakumar ND, Sankari M, Lakshmanan R. Estimation of salivary tumor necrosis factor-alpha in chronic and aggressive periodontitis patients. Contemp Clin Dent 2015 Sep;6(Suppl 1):S152-S156. [doi: 10.4103/0976-237X.166816] [Medline: 26604566]
14. Anuradha BR, Bai YD, Sailaja S, et al. Evaluation of anti-inflammatory effects of curcumin gel as an adjunct to scaling and root planing: a clinical study. J Int Oral Health 2015 Jul;7(7):90-93. [doi: 10.3389/fphys.2016.00219] [Medline: 26229378]
15. Katayama K, Ishida K, Saito M, Koike T, Ogoh S. Hypoxia attenuates cardiopulmonary reflex control of sympathetic nerve activity during mild dynamic leg exercise. Exp Physiol 2016 Mar;101(3):377-386. [doi: 10.1113/EP085632] [Medline: 27094223]
16. Hyun SW, Jung YS. Hypoxia induces FoxO3a-mediated dysfunction of blood-brain barrier. Biochem Biophys Res Commun 2014 Aug 8;450(4):1638-1642. [doi: 10.1016/j.bbrc.2014.07.055] [Medline: 25044111]
17. Mittal S, Soni H, Sharma DK, Mittal K, Pathania V, Sharma S. Comparative evaluation of the antibacterial and physical properties of conventional glass ionomer cement containing chlorhexidine and antibiotics. J Int Soc Prev Community Dent 2015;5(4):268-275. [doi: 10.4103/2231-0762.161754] [Medline: 26310885]
18. Han SJ, Jeong SY, Nam YJ, Yang KH, Lim HS, Chung J. Xylitol inhibits inflammatory cytokine expression induced by lipopolysaccharide from Porphyromonas gingivalis. Clin Diagn Lab Immunol 2005 Nov;12(11):1285-1291. [doi: 10.1128/CDLI.12.11.1285-1291.2005] [Medline: 16275942]
19. Eisenbud DE. Oxygen in wound healing: nutrient, antibiotic, signaling molecule, and therapeutic agent. Clin Plast Surg 2012 Jul;39(3):293-310. [doi: 10.1016/j.cps.2012.05.001] [Medline: 22732377]
20. Cunha EJ, Auersvald CM, Deliberador TM, et al. Effects of active oxygen toothpaste in supragingival biofilm reduction: a randomized controlled clinical trial. Int J Dent 2019;2019:3938214. [doi: 10.1155/2019/3938214] [Medline: 31354823]
21. LaVan FB, Hunt TK. Oxygen and wound healing. Clin Plast Surg 1990 Jul;17(3):463-472. [doi: 10.1016/S0094-1298(20)30621-0] [Medline: 2199137]
22. Hutton JJ, Tappel AL, Udenfriend S. Cofactor and substrate requirements of collagen proline hydroxylase. Arch Biochem Biophys 1967 Jan;118(1):231-240. [doi: 10.1016/0003-9861(67)90302-5]
23. Hunt TK, Pai MP. The effect of varying ambient oxygen tensions on wound metabolism and collagen synthesis. Surg Gynecol Obstet 1972 Oct;135(4):561-567. [Medline: 5077722]
24. Doungudomdacha S, Rawlinson A, Walsh TF, Douglas CW. Effect of non-surgical periodontal treatment on clinical parameters and the numbers of Porphyromonas gingivalis, Prevotella intermedia and Actinobacillus actinomycetemcomitans at adult periodontitis sites. J Clin Periodontol 2001 May;28(5):437-445. [doi: 10.1034/j.1600-051x.2001.028005437.x] [Medline: 11350507]
25. Elvis AM, Ekta JS. Ozone therapy: a clinical review. J Nat Sci Biol Med 2011 Jan;2(1):66-70. [doi: 10.4103/0976-9668.82319] [Medline: 22470237]
26. Sibbald RG, Williamson D, Orsted HL, et al. Preparing the wound bed--debridement, bacterial balance, and moisture balance. Ostomy Wound Manage 2000 Nov;46(11):14-22. [Medline: 11889735]

Abbreviations

CAL: clinical attachment level
FMBS: full mouth bleeding score
GI: gingival index
H202: hydrogen peroxide
PI: plaque index
PPD: probing pocket depth
ROS: reactive oxygen species
SRP: scaling and root planing

Edited by GRML Rosa; submitted 10.Apr.2025; peer-reviewed by A Elfana, A Pejcic; revised version received 22.Jul.2025; accepted 24.Jul.2025; published 02.Mar.2026.

Please cite as:
Bhandari G, Jaiswal P, Bhagat S, Kotecha S
Comparative Evaluation of Effectiveness of Oxygen-Releasing Gel With Scaling and Root Planing (SRP) With That of Scaling and Root Planing Alone in the Management of Chronic Periodontitis Patients: Study Protocol for Cross-Sectional Survey
JMIR Res Protoc 2026;15:e75795
URL: https://www.researchprotocols.org/2026/1/e75795
doi:10.2196/75795

XSL•FO
RenderX

© Geeta Bhandari, Priyanka Jaiswal, Shweta Bhagat, Sakshi Kotecha. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perioperative Antibiotic Prophylaxis in Cesarean Section and the Maternal Gut Microbiome: Protocol for a Remote Observational Cohort Study

Elisabeth AL Feles[1,2,3], MSc; Frauke Mattner[1,2], MD

[1]Department of Human Medicine, Faculty of Health, Witten/Herdecke University, Alfred-Herrhausen-Strasse 50, Witten, Germany

[2]Institute of Hygiene, University Hospital of Witten/Herdecke, Cologne Merheim Medical Centre, Cologne, Germany

[3]Central Pharmacy, University Hospital of Witten/Herdecke, Cologne Merheim Medical Centre, Cologne, Germany

Corresponding Author:
Frauke Mattner, MD
Department of Human Medicine, Faculty of Health, Witten/Herdecke University, Alfred-Herrhausen-Strasse 50, Witten, Germany

Abstract

Background: Cesarean section (CS) requires perioperative antibiotic prophylaxis (PAP) for the prevention of surgical site infections. However, systemic antibiotics during the peripartum period may induce compositional perturbations of the maternal gut microbiome, a system already characterized by reduced resilience. Data on maternal gut microbiome dynamics after CS with PAP are scarce, largely due to logistical and feasibility barriers that limit the participation of pregnant women and new mothers in conventional clinical studies.

Objective: This protocol primarily aims to evaluate the feasibility of a fully decentralized, remote study design for longitudinal gut microbiome research in the peripartum period. Secondary exploratory objectives include the comparative analyses of microbiome composition between CS with PAP and vaginal delivery (VD) without antibiotic exposure to inform future adequately powered studies.

Methods: The MAMA (Microbiome Changes Due to Antibiotic Prophylaxis in Mothers at Birth) study is a prospective, 2-arm observational cohort study conducted entirely off-site. Women in the third trimester of pregnancy were recruited at 2 German level-1 perinatal centers and affiliated outpatient facilities. Participants underwent either CS with PAP (single dose cefuroxime 1.5 g intravenously) or VD without antibiotics. Stool samples were self-collected at home and returned by mail at 3 predefined time points: late pregnancy (T0), 2 to 3 days post partum (T1), and 90±10 days post partum (T2). Primary outcomes are feasibility indicators, including recruitment rate, sample and questionnaire return rates at each time point, adherence to sampling windows, and participant retention across follow-up. Secondary outcomes are exploratory microbiome measures based on 16S rRNA gene sequencing (V3-V4), including alpha diversity indices, beta diversity metrics, and relative taxonomic abundances. Microbiome analyses are explicitly compositional and hypothesis-generating. Group comparisons and longitudinal within-individual changes will be assessed using nonparametric diversity metrics and multivariate distance–based methods. No confirmatory hypothesis testing is planned.

Results: Recruitment occurred between May 2022 and October 2023, with 37 women enrolled (25 CSs and 12 VDs). Follow-up was completed with receipt of the final stool sample in March 2024. DNA extraction and sequencing were completed in a single batch in October 2024. Bioinformatic processing and statistical analyses were initiated in June 2025 and are ongoing as of December 2025. Results from the exploratory microbiome analyses are expected to be published in 2026.

Conclusions: This protocol demonstrates the feasibility of conducting fully decentralized, longitudinal microbiome research in a peripartum population without requiring on-site visits. By integrating study procedures into maternal realities, the remote design reduces participation barriers and addresses a clinically relevant research gap that has remained largely unexamined despite routine use of PAP. While microbiome-related outcomes are exploratory, the methodological framework established here provides a scalable model for future maternal and postpartum research, supporting ethically grounded, participant-centered study designs and evidence-informed care strategies.

Trial Registration: German Clinical Trials Register DRKS00027305; https://drks.de/search/en/trial/DRKS00027305

International Registered Report Identifier (IRRID): DERR1-10.2196/84909

(JMIR Res Protoc 2026;15:e84909) doi:10.2196/84909

XSL•FO
RenderX

KEYWORDS

maternal health; peripartum health; cesarean section; CS; perioperative antibiotic prophylaxis; PAP; maternal gut microbiome; decentralized clinical research; remote observational study

Introduction

Background

Cesarean section (CS) accounts for a substantial and steadily increasing proportion of births worldwide (21% in 2021), representing one of the most common surgical interventions performed in otherwise healthy women [1]. Nevertheless, it carries the third-highest incidence of surgical site infections among all surgical interventions [2]. To reduce postoperative infectious morbidity, perioperative antibiotic prophylaxis (PAP) is universally recommended and constitutes a central component of obstetric surgical safety, as endorsed by major gynecological societies and supported by World Health Organization (WHO) guidelines [3-5]. Recommended prophylactic regimens typically involve a single preincisional intravenous dose of a first- or second-generation cephalosporin, although other antibiotic classes may be used based on local epidemiology and formulary structures. In Germany, cefuroxime is the only antibiotic explicitly approved for cesarean prophylaxis, providing reliable coverage of pathogens commonly associated with postcesarean infections [6]. This universal application of PAP represents an essential protective measure but simultaneously introduces a systemic antibiotic exposure during a physiologically dynamic period for postpartum women.

Short-term systemic antibiotic exposure is known to modulate gastrointestinal microbial communities. Even a single perioperative dose can induce compositional and diversity changes in gut microbial communities, affecting metabolic function and the relative abundance of commensal and opportunistic organisms [7,8]. During pregnancy, endocrine and immunometabolic adaptations reshape the gut microbiome toward lower diversity and increased inflammatory tone, resulting in reduced ecological stability. These pregnancy-associated shifts inherently diminish microbial resilience, potentially heightening vulnerability to additional perturbations such as perioperative antibiotics, postoperative dysmotility, or transient inflammatory responses following cesarean delivery [9,10].

Antibiotics administered intravenously during CS reach high systemic levels and can exert direct antimicrobial activity within the intestinal lumen. Pharmacologically active luminal concentrations have been demonstrated for parenterally administered cephalosporins via biliary excretion [11-13]. In addition, postoperative physiological changes, including transient gastrointestinal dysmotility, reduced oral intake, shifts in hormonal and immune activation, analgesic use, and surgical inflammation, further influence microbial composition [14,15]. Clinically, systemic antibiotic exposure is associated with postantibiotic diarrhea, which may involve overgrowth of opportunistic taxa such as *Clostridioides difficile*, illustrating that short, prophylactic antibiotic regimens can perturb microbial homeostasis even in otherwise healthy individuals [16,17]. Together, these mechanisms support the biological plausibility that PAP could induce measurable, short-term compositional changes in the maternal gut microbiome.

Despite this plausibility, empirical evidence specific to postpartum women is remarkably scarce. The maternal gut microbiome during the peripartum and immediate postpartum period has received far less research attention than neonatal colonization or vertical microbial transmission. Alterations in maternal microbial communities could influence metabolic recovery, immune reconstitution, gastrointestinal symptoms, inflammation, wound healing, and overall well-being. However, the postpartum period is underrepresented in microbiome research because traditional, visit-based study designs impose substantial logistical burdens on new mothers. A recent review underscored this gap, highlighting the near absence of high-quality studies evaluating maternal gut microbiome dynamics following CS and PAP [18].

Rationale for the Study Design

A central methodological consideration is the need for an appropriate comparator. PAP cannot be randomized for ethical reasons as withholding it would contradict the standard of care. Assessing only pre-post changes within CS patients would confound potential antibiotic effects with postoperative physiological responses. A vaginal birth comparison group therefore provides a meaningful counterfactual, capturing physiological postpartum microbial trajectories in the absence of surgical intervention and perioperative antibiotics. However, existing literature is an insufficient substitute for such a comparator, as microbiome datasets vary widely in sequencing region (eg, V1-V2 vs V3-V4), primer selection, read depth, and bioinformatic processing (operational taxonomic unit–based clustering vs amplicon sequence variant [ASV] inference) [19,20]. These methodological discrepancies substantially limit cross-study comparability and hinder valid inference from aggregated external controls. A contemporaneous comparator group is therefore essential for exploratory evaluation of whether postpartum microbial trajectories differ following CS with PAP.

Postpartum research also faces substantial feasibility challenges. New mothers often have limited mobility, competing care demands, and little capacity for additional clinical visits [21]. These barriers disproportionately exclude women from participating in research and limit the representativeness of available data. To address these constraints, we developed a fully remote, participant-centered protocol using at-home stool collection and postal submission. Remote sampling approaches have demonstrated strong feasibility in other areas of microbiome research [22,23], but they have not yet been systematically applied to postpartum maternal populations. Such designs reduce participation burden, enhance accessibility across geographic and socioeconomic contexts, and provide a pragmatic framework for capturing real-world postpartum microbial patterns.

Given the achieved sample size and group imbalance, this study is primarily framed as a feasibility investigation assessing

XSL•FO
RenderX

recruitment, adherence, and the performance of remote stool sampling in the early postpartum period. Secondary exploratory analyses aim to characterize maternal gut microbiome composition and diversity following CS with PAP compared with vaginal birth, to evaluate short-term microbial trajectories, and to identify hypothesis-generating taxonomic patterns. These analyses are not powered for confirmatory inference but will inform the design of future, adequately powered studies.

Although PAP is a universal component of CS, its potential impact on the maternal gut microbiome remains insufficiently understood. This protocol addresses this gap by implementing a remote, participant-centered study design that enables feasible, scalable, and ethically grounded investigation of postpartum microbiome dynamics.

Objectives

The primary objective of the MAMA (Microbiome Changes Due to Antibiotic Prophylaxis in Mothers at Birth) study is to assess the feasibility of a fully remote, nonsite-bound study design for longitudinal gut microbiome research in a peripartum population, including recruitment, sample return, and follow-up completion.

Secondary objectives are to explore compositional changes in the maternal gut microbiome associated with PAP using cefuroxime during CS, compared with vaginal delivery (VD), across the peripartum period.

In this context, the study aims to establish a scalable methodological framework for future remote microbiome studies in maternal health and other vulnerable populations.

Methods

Study Design and Setting

The MAMA study was conducted as a prospective, 2-arm observational cohort study comparing women undergoing CS with PAP to women delivering vaginally without intrapartum antibiotic exposure. Recruitment was integrated into routine clinical care, and participants were followed from the third trimester through 90 days postpartum. The study was carried out between May 2022 and March 2024 at 2 German level-1 perinatal centers (Municipal Hospitals of Cologne and St. Marien Hospital Bonn [GFO Kliniken Bonn]) as well as affiliated outpatient providers. An overview of the overall study design and workflow is provided in Figure 1.

Figure 1. Study design of the MAMA (Microbiome Changes Due to Antibiotic Prophylaxis in Mothers at Birth) trial: fully remote workflow and sample collection schedule. CS: cesarean section; PAP: perioperative antibiotic prophylaxis; VD: vaginal delivery.

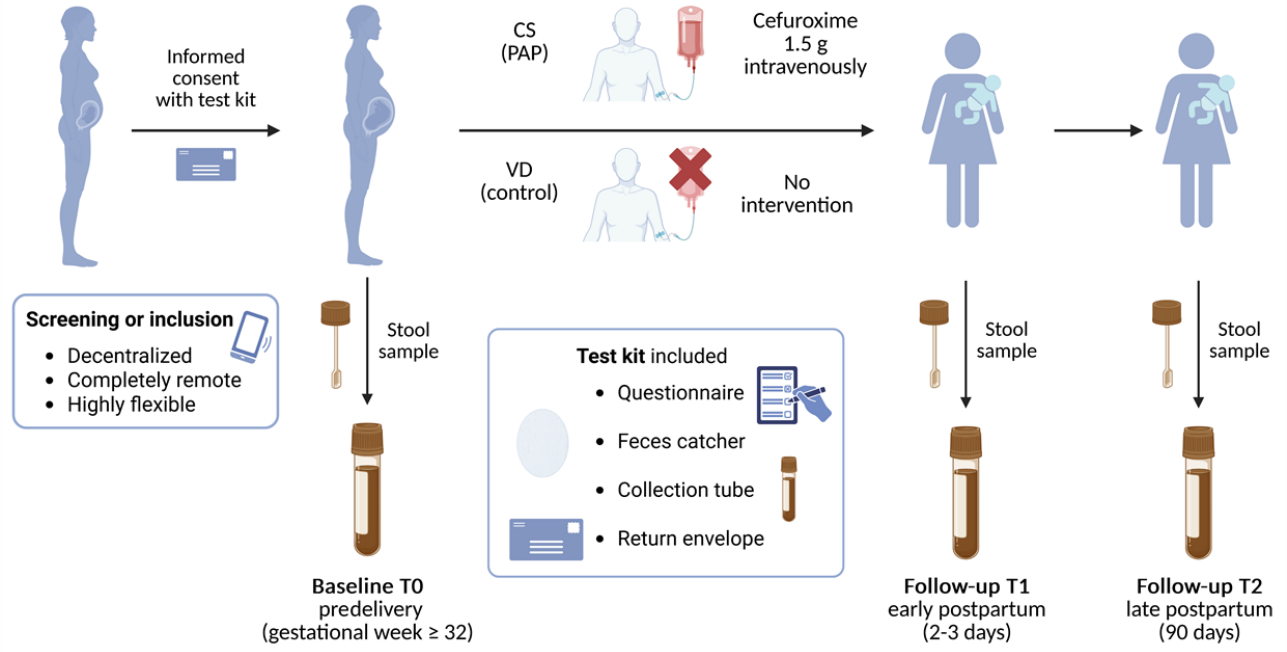

The study was originally conceptualized as a hypothesis-testing cohort design aimed at characterizing maternal gut microbiome trajectories in relation to delivery mode and PAP exposure. However, real-world recruitment constraints in this peripartum population resulted in a smaller and unevenly distributed sample size. In view of these conditions, the present protocol primarily serves as a feasibility investigation of a fully decentralized study framework for maternal microbiome research. Accordingly, subsequent microbiome analyses are defined as exploratory and hypothesis-generating rather than confirmatory.

A contemporaneous comparator group was required to interpret microbial patterns observed after CS. Randomization or withholding of the standard of care was ethically impermissible, and pre-post analyses within CS alone would have conflated antibiotic-associated effects with surgical and physiological consequences of operative delivery. VD therefore represented the most appropriate counterfactual, reflecting postpartum microbial trajectories in the absence of PAP and operative stress. External control datasets were not considered suitable due to widely recognized methodological heterogeneity across microbiome studies (eg, primer regions, sequencing workflows, and bioinformatic pipelines), which limits cross-study comparability.

XSL•FO
RenderX

The study was explicitly implemented as a fully remote, decentralized design to minimize participant burden during pregnancy and early postpartum. All study procedures, including screening, informed consent, questionnaire completion, and stool sampling time points, were completed by participants at home. Communication occurred via telephone or institutional email. Preassembled sampling kits were mailed to participants and included written instructions, a stool catcher, a DNA stabilization tube enabling reliable sample quality even when samples remained at home for several days before mailing, a structured paper–based metadata questionnaire, and a preaddressed return envelope. Participants received reminder notifications from the study coordinator regarding upcoming collection time points and missing documentation. The absence of mandatory on-site visits enabled flexible participation, allowing women to complete all study tasks according to their individual schedules and daily responsibilities. This design facilitated nationwide participation and broadened accessibility, particularly for individuals living in rural areas, with limited mobility, or with substantial childcare and occupational commitments.

The study was developed and implemented in close collaboration between microbiologists, clinical pharmacists, obstetricians, and allied health care professionals, ensuring both methodological rigor and practical feasibility. The reporting of this protocol follows the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) recommendations (Checklist 1) for observational cohort studies, as recommended by the EQUATOR Network [24].

The study was registered in the German Clinical Trials Register (DRKS00027305), which is part of the WHO International Clinical Trials Registry Platform (ICTRP).

Participants were screened and informed by telephone. Upon inclusion, a postal test kit was sent to eligible participants, which included the written informed consent form, a questionnaire, a feces catcher, a stool collection tube, and a preaddressed return envelope. The signed consent form was returned together with the baseline stool sample (T0). Based on the mode of delivery, participants were assigned to either the CS group with PAP (cefuroxime 1.5 g intravenously) or to the control group with VD (no antibiotic exposure). Follow-up samples were collected at 2 to 3 days postpartum (T1) and 90±10 days postpartum (T2). All procedures were conducted by the participants at home, without the need for on-site visits.

Ethical Considerations

The study was reviewed and approved by the ethics committee of Witten/Herdecke University (274/2021) in February 2022. A substantial amendment covering the addition of a second study center and modifications to the screening process was approved in September 2022. Recruitment commenced after ethical approval, with the first participant enrolled in May 2022, and the last participant in October 2023. Follow-up concluded in March 2024, with the return of the final stool sample. The remote consent procedure was explicitly reviewed and approved by the ethics committee. All participants provided verbal consent during the initial telephone screening and subsequently returned signed written consent forms together with their first stool sample. Study participation was voluntary, and participants could withdraw at any time without providing reasons. No incentives were provided for participation. All data were collected and processed in pseudonymized form in compliance with the General Data Protection Regulation (GDPR). No directly identifying information was accessible to the research team. The data protection concept was developed in collaboration with the institutional data protection officer and reviewed as part of the ethical application. There was no patient or public involvement in the study design or implementation. Pseudonymized raw data will be made available to qualified researchers upon reasonable request to the principal investigator and in accordance with institutional and GDPR requirements. The study was conducted as part of a doctoral project at the Faculty of Health, Witten/Herdecke University.

Eligibility and Recruitment

Participants were eligible for inclusion if they were aged at least 18 years, in the third trimester of pregnancy, and planned to deliver either by elective CS or VD without medical indication for peripartum antibiotics. Sufficient German language skills were required to complete the informed consent process and written study documentation. Exclusion criteria comprised participation in another clinical study, chronic inflammatory bowel disease (eg, Crohn disease or ulcerative colitis), hematologic or oncologic disease, chronic renal or hepatic disease, and chronic viral infections such as HIV, hepatitis B virus, or hepatitis C virus. Additional exclusions included systemic use of immunosuppressants, virostatics, cytostatics, proton pump inhibitors, or systemic antimicrobial agents (antibiotics or antifungals) within the past 3 months. Women with a foreseeable need for surgery or antimicrobial treatment other than cefuroxime within the study period (90 days postpartum) were also excluded.

Participants were recruited between May 2022 and October 2023 at 2 German level 1 perinatal centers and affiliated outpatient facilities. Eligible women were identified by attending obstetricians and midwives during routine clinical care. To support recruitment beyond the hospital setting and to address recruiting bias, awareness of the study was raised through posters in clinic waiting areas, printed flyers distributed during antenatal visits, and social media announcements via the official channels of the participating institutions. In addition to reaching a broader population, the social media strategy was specifically intended to include women with low-risk pregnancies who received all prenatal care in the outpatient sector and would otherwise have had limited contact with hospital-based recruitment. This approach aimed to reduce recruitment bias toward high-risk pregnancies, such as those with gestational diabetes, multiple gestation, or prior cesarean delivery. Pregnant women and new mothers represent a population with specific logistical and time-related needs, which informed the choice of a decentralized recruitment and study design.

Interested individuals were first approached by clinical staff and subsequently contacted by the study coordinator. Informed consent was obtained remotely in a 2-step process. Following an initial telephone-based eligibility screening and verbal explanation, participants confirmed their consent in writing by

XSL•FO
RenderX

mailing back the signed documents together with their first stool sample and completed questionnaire.

Assignment to study groups was based on the actual mode of delivery. Women initially recruited for the control group who ultimately required CS for medical reasons were reallocated to the intervention group. Recruitment concluded after inclusion of 25 women in the CS group and 12 in the vaginal birth group, without reaching the originally planned sample size.

Data Collection and Sample Handling

Sample collection was performed independently by participants at home as part of the fully remote study workflow. Final group assignment (CS vs VD) was based retrospectively on the actual mode of delivery at birth. Each participant was asked to provide 3 stool samples at predefined time points: during late pregnancy (T0, from gestational week 32 onward), within 2 to 3 days postpartum (T1), and at 90 (±10) days postpartum (T2). These time points were selected to capture baseline microbiome composition, immediate peripartum changes, and long-term recolonization dynamics. For each time point, participants received a sample collection kit that included a collection tube (DNA/RNA Shield Fecal Collection Tube, Zymo Research), a single-use stool catcher (Feces Catcher, Zymo Research), written instructions, and a prepaid return envelope. Paper-based questionnaires accompanied each sample kit and covered maternal health status, medication use, pregnancy and birth-related variables, and lifestyle factors such as nutrition and physical activity. Participants returned both the stool sample and completed questionnaire in the same shipment. Questionnaires were designed to be completed within approximately 10 minutes and spanned 2 to 3 pages, depending on the time point. Incomplete or missing forms were actively followed up by the study coordinator.

Samples were shipped by standard postal service and received by the coordinating study center. Although the exact time between sampling and mailing was not recorded for all participants, no sample exceeded the manufacturer's recommended stability window.

Upon arrival, samples were documented and immediately stored in a temperature-monitored, alarm-secured refrigerator at 4°C to 8 °C, where they remained until DNA extraction. The median storage duration was 434 days (range: 137 - 784 d), with no deviations from the defined storage conditions. Despite the stabilizing effect of DNA/RNA Shield, the wide range of storage periods may still cause subtle bias.

The nucleic acid stabilization buffer (DNA/RNA Shield) ensures preservation of microbial DNA for at least 24 months at ambient temperature (4° C - 25 °C) and indefinite stability when frozen (<−20 °C), according to the manufacturer. Accordingly, all returned samples were accepted for processing. No exclusions were made due to transport delays, damage, or temperature fluctuations.

DNA extraction was performed using the ZymoBIOMICS 96 MagBead DNA Kit (D4308, Zymo Research) on an automated pipetting workstation (epMotion 5075t, Eppendorf). A slightly modified protocol was applied: 200 µL of fecal sample were mixed with 450 µL of lysis buffer and subjected to mechanical disruption using a bead-beating step in a TissueLyzer (30 Hz; 5 min). DNA was eluted in 50 µL of nuclease-free water and stored at 4 °C to 8 °C until library preparation. For amplification of the V3-V4 region of the 16S rRNA gene, the Quick-16S Next Generation Sequencing Library Prep Kit (D6400, Zymo Research) was used with 341f/806r primers. The protocol was modified by using half reaction volumes and performing 30 polymerase chain reaction (PCR) cycles. No fluorometric quantification was performed prior to amplification; however, postamplification quality control confirmed consistent DNA concentrations across all samples, and no samples had to be excluded due to insufficient yield or quality. Sequencing was conducted on an Illumina MiSeq platform using the MiSeq Reagent Kit v3 (600-cycle, paired-end 301 bp; MS-102 - 3003, Illumina). Cluster generation occurred onboard the MiSeq instrument. Pooled libraries were denatured and diluted according to the manufacturer's protocol and loaded at a final concentration of 12 pM. A PhiX control library (15%) was spiked in to increase run diversity and monitor sequencing quality. Quality controls included 2 extraction blanks consisting of lysis buffer from the DNA extraction kit, one of which was carried through library preparation, and 1 PCR-negative control (nuclease-free water). No spike-in standards for absolute quantification were used. A positive control from the Quick-16S Next Generation Sequencing Library Prep Kit (ZymoBIOMICS Microbial Community DNA Standard) was processed alongside study samples. Mock community controls beyond this kit-derived standard were not included. All study samples and controls were processed and sequenced in a single batch to avoid inter-batch variation.

Microbiome Analysis and Bioinformatics Pipeline

Raw sequencing data will be processed in a QIIME2 (Quantitative Insights Into Microbial Ecology 2; version 2023.2; QIIME 2 Development Team) environment [25], executed within the Qiita platform (Knight Lab, University of California San Diego) [26]. Upon import into Qiita, FASTQ files will undergo automated metadata validation and quality filtering, including removal of low-quality reads, trimming of low-quality ends according to run-specific quality profiles. Adapter and primer sequences will be removed using cutadapt (version 4.2; Max Planck Institute for Developmental Biology) prior to downstream denoising. ASVs will be inferred using Deblur (Knight Lab, University of California San Diego) [27]. Deblur performs error correction and dereplication to generate high-resolution ASVs, while removing putative chimeras and low-abundance artefactual sequences as part of its standard denoising workflow. Because 16S amplicon sequencing provides only relative abundance estimates, absolute bacterial load is not quantified. This will be considered as a possible source of bias when interpreting results. Taxonomic assignment will be conducted using the pretrained Greengenes2 2024.9 full-length classifier, based on the corresponding Greengenes2 2024.9 reference taxonomy [28]. Sequences classified as chloroplast or mitochondria will be identified using Greengenes (version 13.8) and excluded [29]. Representative sequences will be generated within Qiita, and phylogenetic placement will rely on the Greengenes (version 13.8) 99% identity reference tree. Extraction blanks and PCR negative controls will be processed

XSL•FO
RenderX

alongside study samples to detect potential laboratory or reagent contamination. Mock community controls were not included. Samples will be rarefied to a minimum depth of 30,000 reads for alpha diversity and beta diversity analyses. Samples below this threshold will be excluded from diversity metrics but may be retained for descriptive taxonomic summaries.

Analyses will follow a 2-level structure. First, community-level analyses where ASV-based distance matrices (Bray-Curtis, and weighted and unweighted UniFrac) will be used to assess group differences via permutational multivariate ANOVA [30]. Longitudinal microbiome changes will be evaluated by comparing within-individual distance differences across time intervals using Mann-Whitney *U* tests, visualized via principal coordinates analysis. Second, taxonomic-level analyses, in which higher-rank taxonomic profiles (phylum, genus, and species, where available) are summarized descriptively. Hypothesis testing at the taxonomic level will only be conducted if community-level differences emerge, to avoid overinterpretation of sparse compositional data. Given the modest cohort size, all analyses are designed for hypothesis-generating, exploratory interpretation.

Alpha diversity (Shannon index, observed features, and Faith phylogenetic diversity) and beta diversity (Bray-Curtis, and weighted and unweighted UniFrac) will be calculated to describe within- and between-sample diversity. Group comparisons will be analyzed using the Wilcoxon rank-sum or Kruskal-Wallis tests for alpha diversity and permutational multivariate ANOVA for beta diversity. Differential-abundance analysis will be performed using the Analysis of Composition of Microbiomes II (ANCOM-II) [31], a method developed for sparse compositional microbiome data. Given the modest sample size, differential-abundance findings will be interpreted exploratorily. All downstream analyses and visualizations will be performed in QIIME2 and a Python-based environment (Python Software Foundation). Workflow steps, parameters, and software versions will be version-controlled to ensure reproducibility. In addition to taxonomic data, analyses will integrate metadata on parity, medication use, and dietary habits. Associations between microbiome composition and clinical, lifestyle, or sociodemographic variables will be explored through correlation analyses and, where applicable, multivariate approaches such as redundancy analysis or generalized linear models.

Although DNA/RNA Shield stabilizes microbial nucleic acids, prolonged refrigerated storage before extraction may introduce subtle compositional shifts. This potential source of bias will be considered during interpretation.

Outcomes

The primary outcomes of this protocol are feasibility indicators that assess the performance of the fully remote study design. These include the proportion of contacted women who enrolled; sample and questionnaire return rates at T0, T1, and T2; timeliness of sample returns relative to the predefined sampling windows; and overall participant retention across study time points. Together, these measures evaluate compliance, protocol adherence, and the suitability and scalability of the decentralized workflow for maternal microbiome research.

Microbiome-related outcomes are exploratory and focus on compositional rather than absolute changes in microbial communities, as 16S rRNA-sequencing yields relative abundances. These outcomes include alpha diversity metrics (eg, Shannon index and Chao1 richness), beta diversity metrics (eg, Bray-Curtis dissimilarity, and weighted and unweighted UniFrac distances), and relative abundances of bacterial taxa across taxonomic levels from phylum to genus with ASV-level feature tables used for technical analyses. Exploratory identification of marker taxa may be performed where feasible. No diagnostic or clinical interpretations of “dysbiosis” will be drawn.

In line with the observational design and modest, imbalanced sample size, all hypotheses are nonconfirmatory and serve to guide hypothesis-generating analyses. We hypothesize that women undergoing CS with PAP will exhibit a transient reduction in alpha diversity at T1 compared to VD; that compositional perturbations at T1 will at least partially resolve by T2; and that the relative abundance of *Clostridioides difficile* will be higher at T1 in the CS group than in VD. These hypotheses reflect biologically plausible patterns without implying confirmatory testing.

Exploratory analyses will consider associations between microbiome features and maternal metadata such as parity, medication intake, and dietary patterns, where feasible. Subgroup analyses may be limited by sample size constraints, and residual confounding is expected.

Outcomes will be described both cross-sectionally (CS vs VD at each time point) and longitudinally (across T0→T1→T2 within individuals as well as within groups). The statistical procedures for these analyses are outlined in the subsequent section.

Statistical Analysis Plan

The following procedures specify how the predefined outcomes will be analyzed. The analysis will follow an exploratory framework. Although simple, biologically plausible hypotheses have been specified, no formal power calculation was performed. Consequently, results will be interpreted as hypothesis-generating rather than confirmatory.

Feasibility outcomes will be summarized descriptively using proportions, medians, and ranges. Timing of sample returns relative to the predefined collection windows (T0, T1, and T2) will be evaluated by calculating time intervals (in d) between the target and actual return dates. These intervals will be summarized using medians and IQR and compared between delivery groups where appropriate. Reasons for loss to follow-up (eg, no sample returned) will be documented, and analyses will be performed on available cases without imputation of missing data. This analysis will help assess protocol adherence and identify potential logistical barriers in the home-based study model.

To address the previously defined microbiome outcomes, the following statistical approaches will be applied: group comparisons between CS and VD will be performed for each time point (T0, T1, and T2) and longitudinally within participants as well as within groups, using all available paired

XSL•FO
RenderX

samples (eg, T0-T1 or T1-T2), where applicable. Alpha diversity metrics (eg, Shannon index and Chao1 richness) will be compared using nonparametric tests (Wilcoxon rank-sum or Kruskal-Wallis test). Beta diversity will be evaluated using permutational multivariate ANOVA based on distance matrices (eg, Bray-Curtis and UniFrac). Visualization will include principal coordinates analysis.

Analyses will follow a 2-level structure, with primary community-level testing and secondary, exploratory taxonomic-level summaries.

Relative abundances of microbial taxa at different taxonomic levels will be summarized descriptively.

Exploratory differential-abundance analyses will be performed using ANCOM-II, a method suitable for sparse compositional data that incorporates its own procedures for controlling false discoveries. Taxonomic-level hypothesis testing will only be performed if community-level analyses indicate meaningful group differences.

Missing data will not be imputed. Samples with insufficient sequencing depth will be excluded from diversity analyses but may be retained for descriptive taxonomic summaries. Sensitivity analyses will be considered to evaluate the robustness of key findings. Sensitivity analyses may include rerunning key comparisons with alternative distance metrics or differential-abundance methods. If imbalances in relevant maternal variables (eg, prevalence of gestational diabetes) are observed between study groups, multivariate statistical models (eg, generalized linear models and redundancy analysis) will be applied to adjust for potential confounding effects. Stratified analyses or subgroup comparisons may also be conducted if sample size permits.

All statistical analyses will be performed in Python, using QIIME 2-compatible plugins (eg, ANCOM-II) and standard scientific Python libraries. Software versions and analysis scripts will be documented for reproducibility.

Dissemination Plan

The findings of this study will be disseminated through open-access, peer-reviewed publications and presentations at clinical and scientific meetings. This protocol represents the first of 2 planned manuscripts. The main results paper is anticipated for submission in the first half of 2026. Upon publication of the results, raw sequencing data (FASTQ files) will be deposited in the European Nucleotide Archive. The bioinformatic workflow (Python/QIIME 2 scripts) will be made available in a version-controlled repository or provided upon reasonable request to support transparency and reproducibility. Pseudonymized metadata will be shared with qualified researchers upon request in compliance with GDPR and institutional requirements.

Interested participants may request a plain-language summary of the study results in German. Study registration in DRKS/WHO ICTRP ensures continuous transparency, and future updates to the study outputs will be linked to the registry entry. The decentralized design and methodological framework developed here are intended to inform larger maternal health studies and may be adapted for broader applications in remote microbiome research.

Results

The recruitment and sample collection phase of the study has been completed. Between May 2022 and October 2023, a total of 37 participants were enrolled, including 25 women in the CS group and 12 in the VD group. Recruitment concluded below the originally planned sample size, reflecting real-world constraints in this peripartum population. Follow-up was completed with receipt of the final stool sample in March 2024. An overview of baseline characteristics of participants who returned at least one stool sample is presented in Table 1. DNA extraction and 16S rRNA gene sequencing were completed in a single batch in October 2024. All samples passed quality control and were retained for downstream analyses. An overview of included and excluded samples by group and time point is provided in Figure 2. Bioinformatic processing and statistical analyses were initiated in June 2025 and were ongoing as of December 2025.

XSL•FO
RenderX

Table . Baseline characteristics of participants who submitted at least 1 sample (N=31)[a].

Characteristic	Total (N=31)	CS[b] with PAP[c] (n=20)	VD[d] without PAP (n=11)
Continuous variables			
Age (y), median (IQR)	37 (33-39)	37 (33-39)	36 (33-38)
Categorical variables, n (%)			
Delivery mode			
CS	20 (65)	20 (100)	—[e]
VD	11 (35)	—	11 (100)
Multiple births			
Yes	2 (6)	2 (10)	0 (0)
No	29 (94)	18 (90)	11 (100)
Parity			
0	12 (39)	9 (45)	3 (27)
1	13 (42)	5 (25)	8 (73)
2	5 (16)	5 (25)	0 (0)
3	0 (0)	0 (0)	0 (0)
4	0 (0)	0 (0)	0 (0)
5	1 (3)	1 (5)	0 (0)
>1	19 (61)	11 (55)	8 (73)

[a]Only participants in the CS group received PAP; no antibiotics were administered in the VD group.

[b]CS: cesarean section.

[c]PAP: perioperative antibiotic prophylaxis.

[d]VD: vaginal delivery.

[e]Not applicable.

XSL•FO
RenderX

Figure 2. Flowchart of participant enrollment, allocation, follow-up, and analysis in the MAMA (Microbiome Changes Due to Antibiotic Prophylaxis in Mothers at Birth) study, in line with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) recommendations for observational studies. CS: cesarean section; VD: vaginal delivery.

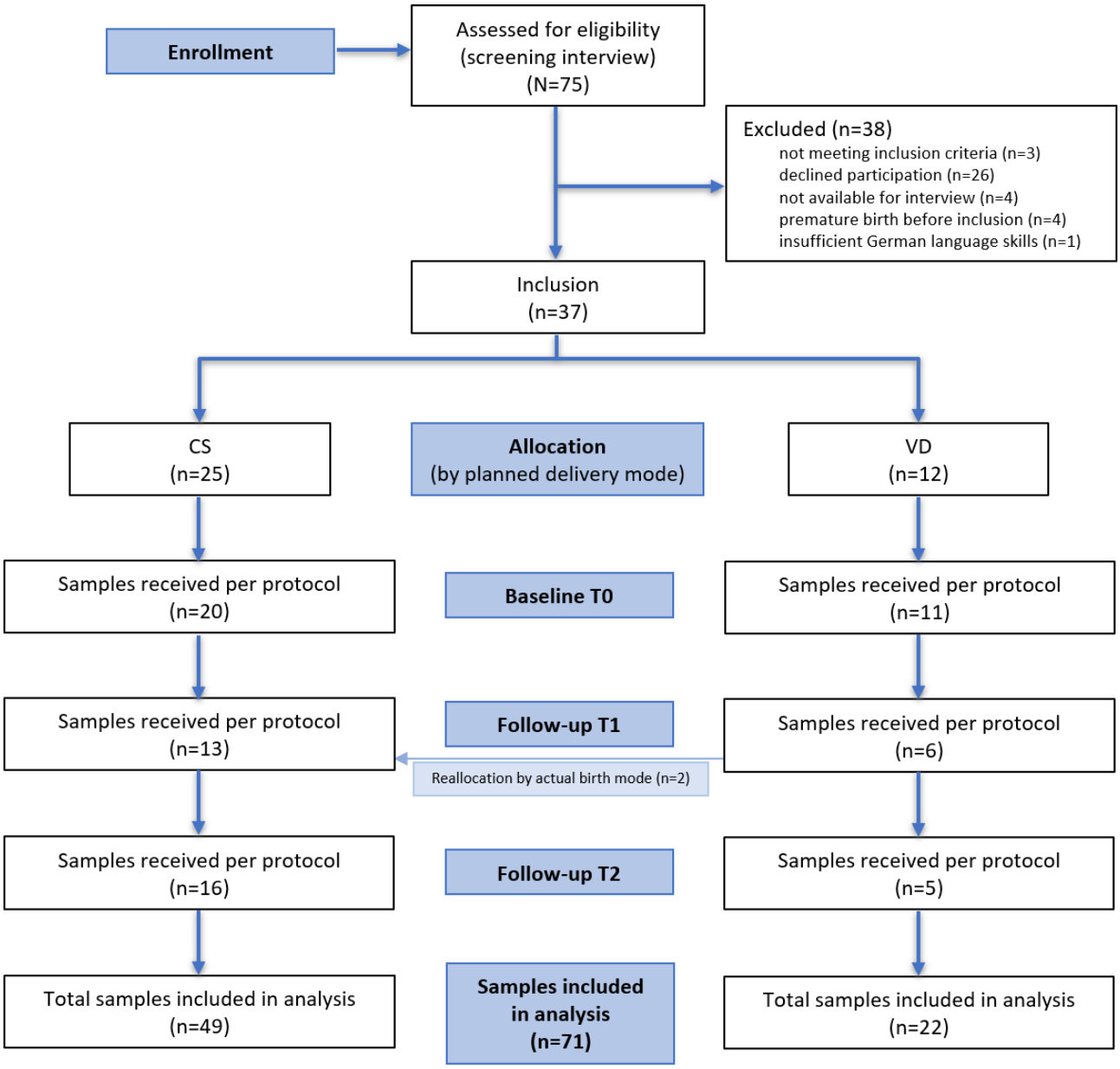

A per-protocol sample was defined as a stool sample received within the prespecified time window and passing quality control. Protocol deviation was defined as exclusion from the per protocol set due to intake of antimicrobial agents (antibiotic or antifungal) other than the study medication or as sample collection outside the defined protocol windows and could not be assigned to a valid time point. Two women, who were initially allocated to the vaginal delivery group, delivered by CS and were therefore analyzed in the CS group; these women are counted under vaginal delivery at baseline (T0) and under CS at follow-up (T1 and T2). Numbers reported per group reflect participants analyzed at each time point; totals across groups were: T0 (n=31), T1 (n=19), T2 (n=21), and final analysis (N=71).

Discussion

Principal Findings

This study evaluates the feasibility of a fully decentralized study model in a peripartum population, addressing practical barriers to participation in longitudinal microbiome research. Within this feasible framework, we anticipate exploratory, compositional alterations of the maternal gut microbiome following CS with PAP, most pronounced shortly after birth and attenuating over time. Given the limited and imbalanced sample size, these microbiome analyses are considered hypothesis-generating rather than confirmatory. In this context, the MAMA study exemplifies how to address practical barriers frequently encountered in peripartum research, including logistical challenges, limited time resources, and competing family responsibilities, which have made participation in

XSL•FO
RenderX

longitudinal microbiome studies particularly difficult for pregnant women and new mothers.

We addressed these limitations by implementing a fully off-site study model that enables microbiome research during pregnancy and early motherhood—without requiring any in-person study visits. This approach not only facilitates participation across geographic and social boundaries but also integrates scientific inquiry into the realities of maternal life. The fully remote design, with telephone-based recruitment and postal sample return, was logistically feasible and well-accepted by participants. Particularly for pregnant women with limited mobility, young children, or those living far from clinical centers, the ability to participate from home appeared to have lowered the barrier to enrollment. Moreover, the participant-based approach allowed for maximum flexibility in scheduling the informed consent process. Participants were able to choose individually convenient times, including evenings and weekends, and rescheduling was possible at any time. In light of the exploratory nature and limited sample size, the primary contribution of this protocol lies in evaluating the feasibility of a fully remote study model in a peripartum population, while microbiome-related analyses are planned as secondary, exploratory outcomes.

Although digital data collection might be considered more efficient, a fully digital questionnaire system was not implemented due to institutional constraints and concerns regarding digital accessibility. Instead, paper-based questionnaires were included in each sample kit, aligning with the postal return of stool samples. This low-threshold approach ensured that participants could complete the forms directly when packaging their sample, without requiring technical equipment or online access. To maximize usability, all materials were written in plain language, following the principle of one idea per sentence, with no more than 15 words per sentence. Where possible, checkbox-style response options were used to reduce cognitive load and completion time.

Beyond logistical feasibility, participant acceptance also emerged as a relevant aspect of the remote study design, as several participants actively expressed interest in the study's scientific goals and asked to be informed of the results. This feedback underscores the importance of building research models that are not only scientifically robust but also accessible, respectful of participants' realities, and responsive to their curiosity and health concerns.

Despite growing awareness of the microbiome's role in health, none of the participating women reported using probiotic supplements during the study period. This observation highlights a potential disconnect between perceived relevance and actionable guidance. The present study may provide a foundation for future research exploring the efficacy and safety of probiotic or dietary interventions during pregnancy and post partum.

At the same time, the implementation of a fully non–site-bound model presented specific challenges. In the VD group, it was difficult to anticipate the exact timing of delivery and thus ensure timely sample collection at T0 and T1. Participants were encouraged to prepare the sampling kit in advance and bring it to the hospital in their birth bag, but adherence to this instruction was not always possible. Additionally, logistical issues were encountered early in the study due to the use of time-limited digital postage codes. As these codes expired before participants mailed their samples, significant mailing delays occurred. The issue was resolved by switching to traditional postage stamps, which ensured reliable sample dispatch without further complications. Although this did not result in sample loss, it created confusion and required additional follow-up communication. These experiences highlight the need for practical guidance and reliable infrastructure when conducting off-site clinical research.

Recruitment was also affected by varying levels of institutional engagement across study sites. While the home-based design reduced participant burden, it still required initial support from clinical staff to initiate contact or distribute study materials. Staff at one study center demonstrated strong interdisciplinary cooperation, with committed medical professionals actively facilitating participant outreach. This differed significantly in the other study center, where limited personnel resources and skepticism about participants' willingness to provide samples postpartum led to reluctance in supporting recruitment. This ultimately contributed to the premature closure of enrollment before the initially planned sample size was reached. This limitation in institutional support directly impacted the final sample size, which remained below target and unevenly distributed between the 2 study arms. In addition, several participants had to be excluded during the study due to protocol violations, such as postpartum antibiotic use identified retrospectively through questionnaire data, or delayed sample return beyond the predefined protocol windows. Despite these constraints, the sampling scheme captured three critical time points—before delivery, shortly after birth, and during postpartum recovery—enabling the identification of both acute and long-term changes in microbial composition.

While the observational design does not aim to establish causal effects, the resulting dataset is expected to yield meaningful insights into temporal microbiome dynamics in the peripartum period. Due to the exploratory nature of this observational study, no formal measures to control for bias resulting from group imbalance or participant dropouts were applied at the design stage. Potential confounding will be addressed during statistical analysis through stratification and adjustment, as appropriate. To this end, several relevant variables were prospectively recorded, including parity, medication or supplement use, diet, and allergies. Where numbers allow, these factors will be considered in stratified or adjusted analyses; however, residual confounding is likely given the modest and imbalanced sample size. Limitations in group comparability will be transparently addressed in the planned results publication in 2026.

Limitations

Importantly, the modest and uneven sample sizes between study groups substantially limit statistical power and preclude confirmatory inference. As microbiome analyses are based on relative abundances derived from 16S rRNA sequencing, absolute bacterial load was not quantified; consequently, the lack of spike-in controls or quantitative PCR may affect the interpretation of apparent taxonomic enrichments or depletions.

XSL•FO
RenderX

Consequently, microbiome-related analyses must be interpreted with caution, as they are susceptible to random variation and cannot reliably detect small or moderate effect sizes. This limitation directly informed the prioritization of feasibility outcomes as primary end points, while microbiome outcomes are framed as secondary and hypothesis-generating.

In addition, the unbalanced group sizes and incomplete longitudinal follow-up further restrict generalizability and limit the ability to perform robust subgroup or multivariable analyses. Consequently, the present study is not intended to provide definitive estimates of effect magnitude, but rather to delineate plausible patterns and inform the design of future, adequately powered investigations.

Comparison With Prior Work

While several longitudinal studies have investigated peripartum changes in the maternal gut microbiome [9,32], existing evidence does not address the specific impact of CS and, in particular, PAP nor their potential association with dysbiosis in the postpartum period. This represents a relevant gap in current research [18], given the widespread use of antibiotic prophylaxis during CS. Microbiome dysbiosis is commonly characterized by compositional shifts, including reduced diversity and altered relative abundances of immunologically and metabolically relevant taxa. Such patterns have been associated with adverse pregnancy outcomes, for example, in preeclampsia, underscoring the potential relevance of maternal gut microbiome alterations during and after pregnancy [33].

Perturbations of the maternal gut microbiome induced by CS with PAP may have implications that extend beyond the immediate postpartum period. Pregnancy is characterized by profound metabolic and immunological adaptations that facilitate fetal tolerance while supporting increased maternal energy demands, processes that are closely intertwined with gut microbial composition [10]. While these adaptations are physiologically beneficial during pregnancy, alterations of the maternal microbiome in the postpartum period may be relevant for subsequent pregnancies. Evidence from a longitudinal cohort study suggests that specific gut microbiome patterns can precede the clinical diagnosis of gestational diabetes, indicating a potential link between microbial composition and later metabolic dysregulation [34]. In addition, dysbiosis in other maternal microbial niches, such as the vaginal microbiome, has been associated with pregnancy-related complications, including gestational diabetes, preterm birth, preeclampsia, and miscarriage [35]. Although these associations are exploratory and do not imply causality, they underscore the importance of longitudinal maternal microbiome research that extends beyond a single pregnancy.

Future studies may build on this protocol by recruiting larger and more balanced cohorts to enable adequately powered analyses. Incorporating absolute quantification approaches, such as quantitative PCR or spike-in standards, and extending longitudinal follow-up across multiple pregnancies would allow more robust assessment of maternal microbiome dynamics and their potential clinical relevance. The remote, participant-centered design evaluated here provides a feasible framework for such future larger-scale investigations.

Conclusions

The MAMA study demonstrates the feasibility of conducting fully decentralized microbiome research in a peripartum population without requiring on-site study visits. By integrating recruitment, informed consent, and longitudinal sample collection into participants' everyday lives through flexible, patient-oriented logistics, this protocol provides a practical and acceptable framework for maternal microbiome research during pregnancy and early post partum.

Given the exploratory design, modest sample size, and group imbalance, this study is not intended to support confirmatory conclusions regarding microbiome alterations following CS with perioperative antibiotic prophylaxis. Accordingly, microbiome-related outcomes should be interpreted as hypothesis-generating, while the primary contribution of this work lies in evaluating and documenting the feasibility of a remote study model in a clinically and logistically sensitive population.

By deliberately centering the maternal perspective and prioritizing maternal health, the MAMA study addresses an underrepresented aspect of perinatal research and illustrates that methodological rigor can be achieved alongside high accessibility for pregnant women and new mothers through an interdisciplinary approach. This framework enables future, larger-scale investigations aimed at robustly characterizing maternal microbiome dynamics and their potential relevance across and beyond pregnancy.

Acknowledgments

The authors would like to thank all study participants for their trust, time, and commitment, as well as the clinical and laboratory collaborators who made the implementation of this study possible. Special thanks go to Ralf Möller (Radiation Biology Department, Institute of Aerospace Medicine, German Aerospace Center), who originally conceptualized the study idea and supported its initiation. The authors gratefully acknowledge his vision and encouragement. The authors further thank Martin Sieber (Institute for Functional Gene Analytics, Bonn-Rhein-Sieg University of Applied Sciences [Hochschule Bonn-Rhein-Sieg]) for his contribution to the initial study protocol, and Ingo Winterfeld and his team at the Institute of Hygiene, Municipal Hospitals of Cologne (Kliniken der Stadt Köln), for their support in sample reception and storage. At the German Aerospace Center, DNA extraction and sequencing were performed by Denise Engel under the supervision of Katharina Siems and Stefan Leuko. Stefan Janssen (Chair of Algorithmic Bioinformatics, Justus Liebig University Giessen) provided statistical supervision, supported by Anna Rehm. The authors also thank Susanne Peter (Department of Obstetrics, St. Marien Hospital Bonn, GFO Kliniken Bonn) for her dedicated efforts in participant recruitment. The study was originally developed in collaboration with the Department of

XSL•FO
RenderX

Obstetrics, Municipal Hospitals of Cologne, under the direction of Werner Neuhaus. No generative artificial intelligence tools were used in the preparation of this manuscript.

Funding

This study did not receive external grant funding. Personnel and material resources for study coordination and implementation were provided by the Institute of Hygiene, Municipal Hospitals of Cologne (Kliniken der Stadt Köln). All laboratory work, including DNA extraction and sequencing, was supported in kind by the German Aerospace Center, which provided personnel, laboratory infrastructure, equipment, and consumables. Statistical analysis was supported by the Chair of Algorithmic Bioinformatics at Justus Liebig University Giessen, which provided personnel time and computing infrastructure. This study was conducted as part of a doctoral project at Witten/Herdecke University.

Data Availability

This manuscript describes a study protocol only. No study data were generated or analyzed at this stage. Data arising from this study will be made available in future publications in accordance with ethical and legal requirements.

Authors' Contributions

Conceptualization: EF, FM
Data curation: EF
Methodology: EF, FM
Investigation: EF
Project administration: EF
Supervision: FM
Visualization: EF
Writing – original draft: EF
Writing – review & editing: EF, FM

Conflicts of Interest

None declared.

Checklist 1
Revised STROBE checklist for the MAMA study protocol on perioperative antibiotic prophylaxis in cesarean section.
[PDF File, 181 KB - resprot_v15i1e84909_app1.pdf]

References

1. Caesarean section rates continue to rise, amid growing inequalities in access: rising rates suggest increasing numbers of medically unnecessary, potentially harmful procedures. World Health Organization. 2021 Jun 16. URL: https://www.who.int/news/item/16-06-2021-caesarean-section-rates-continue-to-rise-amid-growing-inequalities-in-access [accessed 2026-01-31]
2. Global guidelines for the prevention of surgical site infection. : World Health Organization; 2018 URL: https://iris.who.int/server/api/core/bitstreams/5c422323-5c27-4e6a-812e-dbd76023ba48/content [accessed 2026-01-31]
3. Committee on Practice Bulletins-Obstetrics. ACOG Practice Bulletin No. 199: use of prophylactic antibiotics in labor and delivery. Obstet Gynecol 2018 Sep;132(3):e103-e119. [doi: 10.1097/AOG.0000000000002833] [Medline: 30134425]
4. S3-leitlinie die sectio caesarea [Report in German]. : AWMF – arbeitsgemeinschaft der wissenschaftlichen medizinischen fachgesellschaften URL: https://register.awmf.org/assets/guidelines/015-084m_S3_Sectio-caesarea_2020-06_1_02.pdf [accessed 2026-01-31]
5. WHO recommendation on prophylactic antibiotics for women undergoing caesarean section. : World Health Organization; 2021 URL: https://iris.who.int/server/api/core/bitstreams/314369b7-1c7f-4b59-a448-4770ddd20d85/content [accessed 2026-01-31]
6. Cefuroxim-saar® 750 mg / 1500 mg [Article in German]. Rote Liste. URL: https://www.fachinfo.de/fi/detail/020332/cefuroxim-saar-r-750-mg-1500-mg [accessed 2026-01-31]
7. Li JKM, Wang LL, Lau BSY, et al. Oral antibiotics perturbation on gut microbiota after prostate biopsy. Front Cell Infect Microbiol 2022;12:959903. [doi: 10.3389/fcimb.2022.959903] [Medline: 36051239]
8. Doan T, Liu Z, Sié A, et al. Gut microbiome diversity and antimicrobial resistance after a single dose of oral azithromycin in children: a randomized placebo-controlled trial. Am J Trop Med Hyg 2024;110(2):291-294. [doi: 10.4269/ajtmh.23-0651] [Medline: 38227963]
9. Koren O, Goodrich JK, Cullender TC, et al. Host remodeling of the gut microbiome and metabolic changes during pregnancy. Cell 2012 Aug 3;150(3):470-480. [doi: 10.1016/j.cell.2012.07.008] [Medline: 22863002]

XSL•FO
RenderX

10. Huang T, Liang X, Bao H, et al. Multi-omics analysis reveals the associations between altered gut microbiota, metabolites, and cytokines during pregnancy. mSystems 2024 Mar 19;9(3):e0125223. [doi: 10.1128/msystems.01252-23] [Medline: 38323818]
11. Torvorapanit P, Kawang K, Chariyavilaskul P, Kerr SJ, Chatsuwan T, Nilaratanakul V. The in vitro efficacy of activated charcoal in fecal ceftriaxone adsorption among patients who received intravenous ceftriaxone. Antibiotics (Basel) 2023 Jan 9;12(1):127. [doi: 10.3390/antibiotics12010127] [Medline: 36671328]
12. Nilsson-Ehle I, Nord CE, Ursing B. Ceftriaxone: pharmacokinetics and effect on the intestinal microflora in patients with acute bacterial infections. Scand J Infect Dis 1985;17(1):77-82. [doi: 10.3109/00365548509070424] [Medline: 3992209]
13. Sales JE, Rimmer DM. Biliary tract excretion of cefuroxime. Proc R Soc Med 1977;70(Suppl 9):95-97. [doi: 10.1177/00359157770700S920] [Medline: 20919428]
14. Liu Z, Li N, Dang Q, et al. Exploring the roles of intestinal flora in enhanced recovery after surgery. iScience 2023;26(2):105959. [doi: 10.1016/j.isci.2023.105959] [Medline: 36756379]
15. Lederer AK, Chikhladze S, Kohnert E, Huber R, Müller A. Current insights: the impact of gut microbiota on postoperative complications in visceral surgery-a narrative review. Diagnostics (Basel) 2021 Nov 13;11(11):2099. [doi: 10.3390/diagnostics11112099] [Medline: 34829446]
16. Mullish BH, Williams HR. Clostridium difficile infection and antibiotic-associated diarrhoea. Clin Med (Lond) 2018 Jun;18(3):237-241. [doi: 10.7861/clinmedicine.18-3-237] [Medline: 29858434]
17. Mukhtar S, Shaker H, Basarab A, Byrne JP. Prophylactic antibiotics and Clostridium difficile infection. J Hosp Infect 2006 Sep;64(1):93-94. [doi: 10.1016/j.jhin.2006.05.006] [Medline: 16860438]
18. Feles EA, Neidhöfer C, Wessels C, Gruber R, Mattner F. Impact of perioperative antibiotic prophylaxis in caesarean section on the maternal gut microbiome: a systematic review. J Clin Med 2025 Jul 18;14(14):5104. [doi: 10.3390/jcm14145104] [Medline: 40725797]
19. Abellan-Schneyder I, Matchado MS, Reitmeier S, et al. Primer, pipelines, parameters: issues in 16S rRNA gene sequencing. mSphere 2021 Feb 24;6(1):e01202-20. [doi: 10.1128/mSphere.01202-20] [Medline: 33627512]
20. Chiarello M, McCauley M, Villéger S, Jackson CR. Ranking the biases: the choice of OTUs vs. ASVs in 16S rRNA amplicon data analysis has stronger effects on diversity measures than rarefaction and OTU identity threshold. PLoS ONE 2022;17(2):e0264443. [doi: 10.1371/journal.pone.0264443] [Medline: 35202411]
21. Rokicki S, Gobburu A, Weidner M, et al. Barriers and strategies for recruitment of pregnant women in contemporary longitudinal birth cohort studies. BMC Med Res Methodol 2025 Apr 28;25(1):117. [doi: 10.1186/s12874-025-02570-w] [Medline: 40295914]
22. Feigelson HS, Bischoff K, Ardini MAE, et al. Feasibility of self-collection of fecal specimens by randomly sampled women for health-related studies of the gut microbiome. BMC Res Notes 2014 Apr 1;7:204. [doi: 10.1186/1756-0500-7-204] [Medline: 24690120]
23. Hogue SR, Gomez MF, da Silva WV, Pierce CM. A customized at-home stool collection protocol for use in microbiome studies conducted in cancer patient populations. Microb Ecol 2019 Nov;78(4):1030-1034. [doi: 10.1007/s00248-019-01346-2] [Medline: 30929045]
24. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med 2007 Oct 16;4(10):e297. [doi: 10.1371/journal.pmed.0040297] [Medline: 17941715]
25. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 2019 Aug;37(8):852-857. [doi: 10.1038/s41587-019-0209-9] [Medline: 31341288]
26. Gonzalez A, Navas-Molina JA, Kosciolek T, et al. Qiita: rapid, web-enabled microbiome meta-analysis. Nat Methods 2018 Oct;15(10):796-798. [doi: 10.1038/s41592-018-0141-9] [Medline: 30275573]
27. Amir A, McDonald D, Navas-Molina JA, et al. Deblur rapidly resolves single-nucleotide community sequence patterns. mSystems 2017;2(2):e00191-16. [doi: 10.1128/mSystems.00191-16] [Medline: 28289731]
28. McDonald D, Jiang Y, Balaban M, et al. Greengenes2 unifies microbial data in a single reference tree. Nat Biotechnol 2024 May;42(5):715-718. [doi: 10.1038/s41587-023-01845-1] [Medline: 37500913]
29. McDonald D, Price MN, Goodrich J, et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J 2012 Mar;6(3):610-618. [doi: 10.1038/ismej.2011.139] [Medline: 22134646]
30. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol 2001;26(1):32-46. [doi: 10.1046/j.1442-9993.2001.01070.x]
31. Lin H, Peddada SD. Analysis of compositions of microbiomes with bias correction. Nat Commun 2020 Jul 14;11(1):3514. [doi: 10.1038/s41467-020-17041-7] [Medline: 32665548]
32. Gorczyca K, Obuchowska A, Kimber-Trojnar Ż, Wierzchowska-Opoka M, Leszczyńska-Gorzelak B. Changes in the gut microbiome and pathologies in pregnancy. Int J Environ Res Public Health 2022 Aug 12;19(16):9961. [doi: 10.3390/ijerph19169961] [Medline: 36011603]
33. Jin J, Gao L, Zou X, et al. Gut dysbiosis promotes preeclampsia by regulating macrophages and trophoblasts. Circ Res 2022 Sep 2;131(6):492-506. [doi: 10.1161/CIRCRESAHA.122.320771] [Medline: 35950704]

XSL•FO
RenderX

34. Mokkala K, Houttu N, Vahlberg T, Munukka E, Rönnemaa T, Laitinen K. Gut microbiota aberrations precede diagnosis of gestational diabetes mellitus. Acta Diabetol 2017 Dec;54(12):1147-1149. [doi: 10.1007/s00592-017-1056-0] [Medline: 28980079]
35. Gerede A, Nikolettos K, Vavoulidis E, et al. Vaginal microbiome and pregnancy complications: a review. J Clin Med 2024 Jun 30;13(13):3875. [doi: 10.3390/jcm13133875] [Medline: 38999442]

Abbreviations

ANCOM-II: Analysis of Composition of Microbiomes II
ASV: amplicon sequence variant
CS: cesarean section
DRKS: German Clinical Trials Register (Deutsches Register Klinischer Studien)
GDPR: General Data Protection Regulation
ICTRP: International Clinical Trials Registry Platform
MAMA: Microbiome Changes Due to Antibiotic Prophylaxis in Mothers at Birth
PAP: perioperative antibiotic prophylaxis
PCR: polymerase chain reaction
QIIME 2: Quantitative Insights Into Microbial Ecology 2
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
VD: vaginal delivery
WHO: World Health Organization

Edited by A Schwartz; submitted 30.Sep.2025; peer-reviewed by C Neidhofer, N Murgueitio; accepted 09.Jan.2026; published 22.Apr.2026.

Please cite as:
Feles EAL, Mattner F
Perioperative Antibiotic Prophylaxis in Cesarean Section and the Maternal Gut Microbiome: Protocol for a Remote Observational Cohort Study
JMIR Res Protoc 2026;15:e84909
URL: https://www.researchprotocols.org/2026/1/e84909
doi:10.2196/84909

© Elisabeth AL Feles, Frauke Mattner. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Children's Attitudes Toward Dementia: Protocol for a Cross-Sectional Study

Gema Somoza-Fernández[1], MD; Javier Ortiz-Alonso[1,2,3,4], MD; José Antonio Serra-Rexach[1,2,3,4], MD, PhD; María Teresa Vidán[1,2,3,4], MD, PhD

[1]Geriatrics Department, Hospital General Universitario Gregorio Marañón, Calle Doctor Esquerdo 46, Madrid, Spain

[2]Biopathology of Aging Group, Instituto de Investigacion Sanitaria Hospital General Universitario Gregorio Marañón, Madrid, Spain

[3]Biomedical Research Networking Center on Frailty and Healthy Aging (CIBERFES), Instituto de Salud Carlos III, Madrid, Spain

[4]Faculty of Medicine, Universidad Complutense de Madrid, Madrid, Spain

Corresponding Author:
Gema Somoza-Fernández, MD
Geriatrics Department, Hospital General Universitario Gregorio Marañón, Calle Doctor Esquerdo 46, Madrid, Spain

Abstract

Background: Research on dementia has primarily focused on adults, leaving a gap in understanding related to children's attitudes toward dementia. This gap is relevant given the growing societal impact of dementia and the potential role of early education in shaping attitudes.

Objective: The primary objective of this study is to assess attitudes toward dementia in a sample of Spanish primary school children. Secondary objectives include (1) examining the psychometric properties of the questionnaire developed for this study, (2) identifying factors associated with differing attitudes toward dementia, and (3) evaluating children's social desirability and its correlation with questionnaire responses.

Methods: A 23-item, tripartite Likert-type questionnaire (Cuestionario de Actitudes de los Niños Españoles Hacia la Demencia, or "Spanish Children's Attitudes Toward Dementia Questionnaire") was developed through a comprehensive literature review, a pilot study with a convenience sample of 10 children, and expert panel consultation. The study sample will comprise children aged 8 to 11 years enrolled in the fourth and fifth grades at 6 selected schools in the Community of Madrid, Spain. Participants will complete the questionnaire alongside a social desirability scale. Additional variables (eg, current or previous cohabitation with grandparents) will be collected to explore factors associated with differing attitudes toward dementia.

Results: This study is not funded. Data collection began in January 2025 and is projected to conclude in January 2026. As of July 2025, a total of 164 participants from 3 selected schools have been recruited. Data analysis is expected to begin in January 2026, and results are planned for publication in 2027.

Conclusions: This study represents a feasible and novel initiative that may enhance understanding of children's attitudes toward dementia. The findings could inform the development of targeted educational interventions and support the inclusion of dementia-related content in school-based education.

International Registered Report Identifier (IRRID): DERR1-10.2196/81576

(JMIR Res Protoc 2026;15:e81576) doi:10.2196/81576

KEYWORDS

attitudes; knowledge; dementia; children; surveys and questionnaires

Introduction

Background

Dementia is a leading cause of disability among older individuals worldwide [1]. Current estimates suggest that approximately 55 million individuals are living with dementia worldwide, a figure projected to increase to 150.8 million by 2050 [2]. Dementia affects not only the quality of life of those diagnosed but also that of their caregivers and families, as caregiving often entails personal, psychological, and financial challenges [3].

To address this global health issue, the *Global Action Plan on the Public Health Response to Dementia 2017-2025* was unanimously adopted by all 194 member states of the World Health Organization at the 70th World Health Assembly. The objective of this initiative is to foster a society in which individuals with dementia and their caregivers can realize their full potential with dignity, respect, and equality [4].

XSL•FO
RenderX

Despite these objectives, significant gaps remain in knowledge and awareness of dementia. This is evident in the widespread misconception of the disease as either reversible [5] or an unavoidable consequence of aging that cannot be prevented [6]. Furthermore, it is widely recognized that the general public often experiences discomfort and shame in the presence of individuals with dementia, contributing to the stigma surrounding the condition [7]. For instance, a 2012 survey involving 127 individuals with dementia and 1716 caregivers revealed that 25% and 11% of respondents, respectively, had concealed the diagnosis from friends or family members due to stigma [8].

While research on dementia perception has primarily focused on adults, there remains a significant gap in understanding how children perceive this condition. This gap is critical given that children represent the future of a society increasingly affected by the growing prevalence of dementia and related challenges. Considering that childhood is a critical period for instilling knowledge and values [9], addressing this gap is imperative.

Research on children's feelings and behaviors regarding dementia has mainly focused on the impact of the disease on grandparent-grandchild relationships [10-12]. Regarding knowledge about the disease, most children and adolescents obtain information about dementia from media sources [13], as parents often avoid discussing dementia-related topics with their children [14]. However, several educational interventions over the past decade have successfully improved children's understanding and knowledge of dementia [15-18]. Similarly, positive outcomes have been reported from "intergenerational interventions" [19-21], which involve direct interactions between children and individuals living with dementia.

Study Hypothesis

The central hypothesis of this study is that children's feelings and behaviors regarding dementia are largely unknown and that their level of disease-related knowledge is generally low. These 3 aspects—affective, behavioral, and cognitive responses toward a person, situation, or object—constitute the widely accepted tripartite model of "attitude" [22]. Accordingly, this study aims to examine children's attitudes toward dementia, as well as the 3 domains of this construct, in a targeted sample of primary school children.

Study Objectives

The primary objective of this study is to assess children's overall attitude toward dementia and its 3 domains—feelings, behaviors, and knowledge about the disease—among fourth- and fifth-grade primary school students from 6 selected schools in the Community of Madrid, Spain.

Secondary objectives include the following:

- Examining the psychometric properties of the questionnaire developed for the main objective
- Identifying factors associated with differing dementia-related attitudes, both overall and across the 3 domains (feelings, behaviors, and knowledge about the disease)
- Determining children's social desirability and its correlation with the questionnaire responses

Methods

Study Design

This is a quantitative, cross-sectional study that will be conducted between January 2025 and January 2026.

Study Participants and Setting

As mentioned previously, the study sample will consist of children enrolled in fourth and fifth grade of primary education (aged 8-11 years) in 6 selected schools in the Community of Madrid.

Schools will be purposively selected to reflect diversity in socioeconomic context, educational management, and religious orientation. As a first step, regions in the Community of Madrid will be ranked according to area-level socioeconomic status [23] and divided into 2 strata representing the lower and upper halves of the distribution. Subsequently, within each stratum, 1 public school, 1 secular private or concerted school, and 1 private or concerted school with religious education will be selected.

Schools will be contacted via email and provided with an information sheet describing the study aims and procedures; those that agree to participate will then be scheduled for data collection visits. If a school declines participation, a replacement school from the same stratum with the same educational management and religious orientation will be invited.

Sample Size

As will be described below, the main primary outcome of this study is the total score on a 23-item, Likert-type questionnaire (5 response options per item; maximum total score=115) measuring attitude toward dementia. Therefore, the sample size was calculated using the Cochran formula for estimating a mean (continuous outcome) [24,25]:

$$no = Z^2 \cdot SD^2 d^2$$

In this formula, Z is 1.96 (95% confidence), SD is 26 (a conservative estimate of the SD of the total questionnaire score), and d is 3.45 points (3% of the maximum score of 115). This yields 219 evaluable children (n).

Finally, an explicit adjustment for nonevaluable questionnaires was incorporated. We anticipated 15% to 20% nonevaluable questionnaires (eg, parental consent not returned or declined, absence on the survey day, or child nonassent) and inflated the sample size using the standard approach [26]:

$$N = no(1-r)$$

In this formula, N is the target number to recruit, n is the minimum required number of evaluable participants, and r is the anticipated proportion of nonevaluable cases (0.15-0.20). This yields 260 to 270 children.

Inclusion Criteria

All children currently attending fourth or fifth grade of primary education in the selected schools will be eligible to participate

XSL•FO
RenderX

in the study provided they meet the following criteria: (1) provision of written consent from their parents or legal guardians (opt-in procedure) in accordance with Spanish Law 26/2015 of July 28 on the modification of the system of protection for childhood and adolescence and (2) provision of verbal assent immediately prior to participation and only if parental or guardian consent has been obtained.

Primary Outcomes

Overview

Given the absence of a validated scale in Spain to assess attitudes toward dementia in children, a questionnaire was developed. As described above, the construct of "attitude" comprises an individual's cognitive, behavioral, and affective response to an object, person, or situation. This tripartite model served as the foundation for the questionnaire developed for this study, named the Cuestionario de Actitudes de los Niños Españoles Hacia la Demencia (CANDem; "Spanish Children's Attitudes Toward Dementia Questionnaire").

The CANDem comprises 23 items on a 5-point Likert scale related to the 3 domains of the tripartite model: 12 cognitive items (eg, "All people develop dementia as they age"), 6 affective items (eg, "It would be a very boring experience to spend an entire afternoon with a person with dementia"), and 5 behavioral items (eg, "If a person with dementia forgot something important, I would get a little angry so that they remember it better next time"). The maximum total score is 115, with higher scores indicating more positive attitudes toward dementia.

The main primary outcome will be the total score on the CANDem, which reflects overall attitudes toward dementia. The separate scores on each domain (cognitive, affective, and behavioral) will be considered additional primary outcomes.

The CANDem was developed through the following stages: (1) content development, (2) measure design, (3) expert panel review, (4) pilot-testing, and (5) re-evaluation by the expert panel.

Phase 1: Content Development

This phase involved identifying relevant domains for the questionnaire, informed by a comprehensive review of prior studies, expert opinions, and methodological guidelines for attitude questionnaires [27].

Previous Studies

Item selection was based on previously validated dementia attitude scales:

- Dementia Attitudes Scale—a scale comprising 20 items on a 7-point Likert scale measuring attitudes toward Alzheimer disease among American college students and care workers, demonstrating strong reliability and a 2-factor structure ("dementia knowledge" and "social comfort") [28]
- Escala de Actitudes Hacia la Demencia ("Attitudes Toward Dementia Scale")—a scale comprising 20 items on a 7-point Likert scale validated among Argentine psychology, medical, and nursing students showing good reliability and a 3-factor structure ("rejection," "affection," and "experience") [29]
- Kids Insight Into Dementia Survey—a scale comprising 14 items on a 5-point Likert scale developed in Australia to assess children's perceptions of dementia, with good reliability and a 3-factor structure ("personhood," "stigma," and "dementia understanding") [30]
- Dementia Community Attitudes Questionnaire—a 10-item Likert-type scale assessing attitudes toward dementia in the general population comprising 3 factors ("engagement," "challenges," and "decision-making") [31]

Additionally, findings from a qualitative study were considered [32]. In this study, 27 participants with personal experience of dementia—including 12 primary caregivers, 4 nonprimary caregivers, 6 grandchildren of people with dementia, and 5 individuals with early-stage dementia—were asked what they considered most important for children to learn about the disease.

Item Formulation

Items were designed to be brief, clear, concise, and relevant.

Number of Items

An initial pool of items exceeding the expected final number was generated to allow for item reduction during subsequent stages.

In consideration of the aforementioned recommendations, a preliminary 24-item questionnaire was constructed. Items addressed themes such as the dignity of individuals with dementia (eg, that they can still enjoy life and feel the love of their family and friends); the importance of nonpharmacological interventions; lifestyle habits potentially influencing dementia risk; and common misconceptions, including the beliefs that dementia is inevitable, contagious, or reversible.

Phase 2: Measure Design

Given the preference for Likert scales over dichotomous items in attitude measurement [33], a 24-item questionnaire with a Likert-type response format was selected. Regarding the number of response categories, there is evidence suggesting that scales comprising 2-, 3-, and 4-point options tend to perform suboptimally compared to those with more response options [34]. Therefore, each item was rated using a 5-point Likert-type scale including an "I don't know" option. Affective and behavioral items were rated as follows: "strongly agree" (5), "agree" (4), "I don't know" (3), "disagree" (2), and "strongly disagree" (1). For cognitive (knowledge) items, the same 5-point response options were used; however, "I don't know" responses will be treated as missing rather than as the midpoint as they indicate absence of knowledge rather than a neutral opinion. The proportion of "I don't know" responses will be reported given that they provide relevant information about children's understanding.

The inclusion of the "I don't know" option aimed to reduce arbitrary guessing and social desirability bias [35]. Additionally, to mitigate acquiescence bias, which refers to the tendency to select positive responses, 13 items are reverse scored [36].

XSL•FO
RenderX

Phase 3: Review by a Panel of Experts

A panel of 7 dementia experts evaluated the 24 questionnaire items for clarity and content validity using a structured evaluation framework [37]. The panel included 6 geriatricians with academic, clinical, and research expertise, as well as 1 individual with lived experience as a long-term caregiver for a family member with Alzheimer disease. In addition, to ensure that the questionnaire was age appropriate, 3 primary school teachers and 2 pediatricians contributed to the adaptation of item wording into child-friendly language.

The panel reached a consensus on the final set of 24 items. Personal data were protected in accordance with Spain's Organic Law 3/2018 on the Protection of Personal Data and Guarantee of Digital Rights.

Phase 4: Pilot-Testing

The aim of this phase was not to evaluate participants' responses but rather to identify potential issues related to item wording and comprehension. Accordingly, a relatively small sample size was considered appropriate in line with previous methodological recommendations [27]. A convenience sample of 10 children aged 8 to 11 years was recruited. Prior to participation, parents or legal guardians were provided with an information sheet and asked to sign a consent form.

The questionnaire was administered in paper format by one of the study authors at the participants' homes. Each child provided verbal assent immediately before starting the questionnaire in accordance with ethical procedures. The questionnaire was completed in a quiet setting in a one-to-one manner with the researcher to ensure privacy and minimize external influence. Parents or guardians were asked not to be present during completion. The researcher was available throughout to clarify any questions or explain unfamiliar terms.

Before starting, children were asked whether they were familiar with the terms "dementia" or "Alzheimer's." If not, the researcher provided a brief, neutral explanation to support basic understanding. Care was taken to avoid influencing responses to the questionnaire items.

Children described the questionnaire as "simple" and "easy to understand." Comprehension difficulties were noted in items 10, 18, and 23. The estimated completion time was 20 minutes. After finishing, the children expressed a desire to learn more about dementia and asked questions about the disease, including the possibility of a cure, the impact of dementia on an individual's ability to bathe or dress, and the most effective treatment for an anxious person with dementia.

Phase 5: Re-Evaluation by the Panel of Experts

On the basis of the results of the pilot study, this phase aimed to finalize the CANDem. The expert panel reviewed the pilot findings and made the necessary modifications. In accordance with previous studies [30], items with more than 50% of extreme responses (ie, "strongly agree" or "strongly disagree") were excluded to avoid floor or ceiling effects.

Items 10, 18, and 23 were revised to enhance clarity and comprehension. Item 24 ("I find it is sadder to have dementia than to fall and break an arm") was removed as 80% of the children selected the "strongly agree" option.

Given that 4 out of 10 children in the pilot study demonstrated difficulty understanding the term "dementia," the expert panel decided to introduce the final version of the CANDem with selected vignettes from the Spanish children's comic "You, Me, and Alzheimer's Disease," which depicts a person experiencing temporal disorientation and short-term memory loss [38]. From the comic, vignettes were chosen that did not provide clues or otherwise influence children's responses to the questionnaire, thereby minimizing potential framing effects.

Textbox 1 presents the final version of the CANDem. To facilitate comprehension of this manuscript, the questionnaire has been translated into English from its original Spanish format. For this purpose, a formal translation and back translation procedure was applied. First, a bilingual researcher translated the Spanish version into English. Second, a graduate in translation and interpreting independently back translated the English version into Spanish. Finally, the research team compared the back translation with the original Spanish version and reconciled any discrepancies to ensure equivalence of meaning.

XSL•FO
RenderX

Textbox 1. Final version of the CANDem (Cuestionario de Actitudes de los Niños Españoles Hacia la Demencia, or "Spanish Children's Attitudes Toward Dementia Questionnaire").

- Item 1: "All people develop dementia as they age" (cognitive item; reverse scored)
- Item 2: "It would be a very boring experience to spend an entire afternoon with a person with dementia" (affective item; reverse scored)
- Item 3: "There are no medications that can cure dementia" (cognitive item)
- Item 4: "If I were introduced to a person with dementia, I would try to spend time with them to get to know them better" (behavioral item)
- Item 5: "I would not feel upset if a family member or friend with dementia forgot my birthday" (affective item)
- Item 6: "A person with dementia may transmit it to another person" (cognitive item; reverse scored)
- Item 7: "It is possible to adopt certain habits (such as not smoking or exercising) that can help to prevent dementia" (cognitive item)
- Item 8: "If I met someone with dementia, I would speak to them as if they were little children to facilitate their understanding" (behavioral item; reverse scored)
- Item 9: "People with dementia are sad almost all of the time" (cognitive item; reverse scored)
- Item 10: "I would not be scared if I saw someone with dementia being nervous and shouting" (affective item)
- Item 11: "If a person with dementia forgot something important, I would get a little angry so that they remember it better next time" (behavioral item; reverse scored)
- Item 12: "Dementia can impair a person's ability to perform activities such as showering, dressing, or walking" (cognitive item)
- Item 13: "If someone with dementia kept asking me the same question all the time (for example, 'What day of the week is it today?'), I would start laughing" (behavioral item; reverse scored)
- Item 14: "People with dementia often do not perceive displays of affection from their family or friends" (cognitive item; reverse scored)
- Item 15: "If I were walking down the street in a hurry and saw someone with dementia who was lost, I would stop to help them" (behavioral item)
- Item 16: "People with dementia often have poor hearing, and that's why it's very important to speak to them loudly" (cognitive item; reverse scored)
- Item 17: "If someone with dementia gets angry or nervous, we must give them a pill to calm them down and then talk to them" (cognitive item; reverse scored)
- Item 18: "Puzzles and music are activities that can help people with dementia" (cognitive item)
- Item 19: "People with dementia can recover from the disease and go back to being themselves" (cognitive item; reverse scored)
- Item 20: "If I had a family member with dementia, I would not be embarrassed to tell my friends" (affective item)
- Item 21: "I wouldn't like to play a game with someone with dementia because they would forget the rules all the time" (affective item; reverse scored)
- Item 22: "I would get a little nervous if a person with dementia smiled at me" (affective item; reverse scored)
- Item 23: "People with dementia can have hobbies or enjoy their free time" (cognitive item)

Secondary Outcomes

Secondary outcomes will provide additional information about children's attitudes toward dementia and the psychometric performance of the CANDem. The first secondary outcome is psychometric properties of the CANDem, including internal consistency, item discrimination, and construct and content validity. Additional details on the statistical evaluation of these properties are provided in the Statistical Analysis section. The second outcome is children's social desirability, defined as the tendency to respond in culturally appropriate or socially approved ways. This will be measured using an 8-item dichotomous ("yes" or "no") scale previously validated for use in Spanish-speaking children, with higher scores indicating greater social desirability [39]. This scale will be administered together with the CANDem in paper format. Details on how associations between social desirability and CANDem scores will be examined are provided in the Statistical Analysis section. The third outcome is factors associated with attitudes toward dementia. The following variables will be collected: age, sex, country and city of birth, parents' country of birth, prior familiarity with the term "dementia" (ie, whether the child had heard of the term before the survey session), prior knowledge of its meaning (ie, whether the child knew the meaning of the term before the survey session), presence of a family member or acquaintance with dementia, previous cohabitation with a person with dementia, and previous cohabitation with grandparents or other older adults. These variables will be completed by the child after finishing both the CANDem and the social desirability scale, and their relationships with differing dementia-related attitudes—both overall and across the 3 domains (feelings, behaviors, and knowledge about the disease)—will be assessed using regression models (see the Statistical Analysis section).

XSL•FO
RenderX

Ethical Considerations

This study was reviewed and approved by the ethics committee of Gregorio Marañón University Hospital (approval Act 18/2023; October 9, 2023).

A few days before the researcher visits a school to administer the questionnaire, the parents or legal guardians of all fourth- and fifth-grade students at the corresponding school will receive a participant information sheet and the informed consent form for study participation. On the day of the survey, immediately before questionnaire administration, children whose parents or legal guardians have given written consent will receive age-appropriate information and be asked to provide verbal assent. Only children with both written parental or legal guardian consent and their own verbal assent will participate in the study following an opt-in procedure. No compensation will be provided to participants.

All data will be collected, processed, and stored in accordance with the General Data Protection Regulation. Study data will be anonymized prior to analysis to ensure participant confidentiality. The study will be conducted in accordance with the principles of the Declaration of Helsinki and will follow JMIR [40] and Committee on Publication Ethics (COPE) ethical guidelines [41].

Data Collection Procedure

Data collection will take place during prescheduled visits coordinated with each school to minimize disruption to normal academic activities. Only children for whom written parental or legal guardian consent has been obtained and who provide verbal assent immediately prior to participation will receive an individual paper-based booklet containing (1) the vignettes presenting a person with dementia, (2) the CANDem, (3) the social desirability scale, and (4) the section covering sociodemographic and dementia-related exposure variables. At the end of the booklet, children will respond to 2 brief acceptability items—"Was the questionnaire easy to complete?" and "Would you like to learn more about dementia?"—both with dichotomous ("yes" or "no") response options.

Before beginning the survey, children will be reminded that their participation is anonymous, that they should answer honestly, and that they may use the "I don't know" option whenever they are unsure. The classroom teacher will remain present throughout the session, and the researcher will be available to clarify procedural questions without influencing responses.

As demonstrated in the pilot study, completing the CANDem will take approximately 20 minutes. As the additional components of the booklet—reading the vignettes, completing the social desirability scale, and responding to the sociodemographic and dementia-related exposure questions—were not timed during the pilot study, each classroom session is planned to last approximately 1 hour. This includes distribution of materials, completion of all components, and a brief 10-minute discussion at the end during which children may ask questions and provide feedback on any difficulties encountered with the questionnaire.

A child distress protocol will be implemented to safeguard participants' well-being. Should any child show signs of discomfort (eg, if the content evokes personal experiences, such as having a family member with dementia), the researcher will provide supportive listening, immediately inform the classroom teacher, and pause the questionnaire for that child.

Statistical Analysis

All statistical analyses will be conducted using SPSS Statistics (version 25; IBM Corp) or other equivalent statistical software. A 2-sided significance level at a P value of .05 will be applied for prespecified analyses.

Clustering and School Effects

Because children are nested within classrooms and schools, regression-based inferential analyses will account for clustering. Given that only 6 schools are included, multilevel models with random effects cannot be reliably estimated [42,43]; therefore, school will be included as a fixed-effect variable, and within-classroom dependency will be addressed using robust SEs clustered at the classroom level. For planning purposes, an intraclass correlation coefficient of 0.05 will be assumed, consistent with typical values in educational research [44].

Primary Outcomes

As noted, the primary outcomes of this study are the CANDem total score and the scores on its 3 domains (cognitive, affective, and behavioral). Descriptive statistics, including means, SDs, medians, IQRs, and observed ranges, will be calculated for the total score and each domain.

Secondary Outcomes

As stated above, secondary outcomes include the psychometric properties of the CANDem, children's social desirability, and factors associated with differing attitudes toward dementia.

Psychometric Properties of the CANDem

Internal consistency will be evaluated using the Cronbach α for the total score and each domain, with values above 0.70 considered acceptable [36,45]. Item discrimination will be assessed using the corrected homogeneity index, calculated as the Pearson correlation between each item and (1) its corresponding domain score and (2) the total score, excluding the item under evaluation; items with a homogeneity index of 0.20 or lower will be considered to show poor discrimination [46]. Construct validity will be examined through exploratory factor analysis (EFA) using principal axis factoring after verifying sampling adequacy. Content validity has already been ensured through a rigorous development process, including an exhaustive literature review, consultation with an expert panel, and pilot-testing.

As the CANDem has been translated into English to facilitate understanding of this manuscript, it is important to note that all psychometric analyses will be conducted exclusively on the Spanish version of the instrument, which is the language in which it will be administered to participants.

Children's Social Desirability

Scores on the 8-item dichotomous scale will be analyzed. Associations between social desirability and CANDem scores

XSL•FO
RenderX

will be described using Pearson correlation coefficients. For inferential purposes, associations will also be examined using linear regression models, with school included as a fixed effect and robust SEs clustered at the classroom level.

Factors Associated With Children's Attitudes Toward Dementia

The sociodemographic and contextual variables described in the Secondary Outcomes section (eg, age, sex, previous cohabitation with grandparents or other older adults, and prior exposure to dementia) will first be described using appropriate descriptive statistics.

Subsequently, to explore whether these variables are associated with differing attitudes toward dementia, the CANDem total score and its 3 domain scores (cognitive, affective, and behavioral) will be analyzed as continuous outcomes using linear regression models accounting for school effects and clustering, as described above. In addition, the total score and domain scores will be categorized into quartiles, and associations with the same predictors will be examined using ordinal logistic regression models (proportional odds), likewise accounting for school effects and clustering. The proportional odds assumption will be assessed.

Linearity (where applicable), homoscedasticity, and collinearity will be assessed for all regression models, and appropriate adjustments will be applied if necessary.

Exploratory Analyses

Exploratory analyses will compare the CANDem total score and domain scores across prespecified subgroups (age group, sex, child's country and city of birth, parents' country of birth, prior familiarity with the term "dementia," prior knowledge of its meaning, presence of a family member or acquaintance with dementia, previous cohabitation with a person with dementia, previous cohabitation with grandparents or other older adults, and school type), with corresponding *P* values reported.

In addition, item-level response patterns will be described by reporting the percentage distribution across the 5 Likert response options for each CANDem item. Item-level responses will be treated as ordinal variables, and differences across subgroups will be explored by comparing response distributions. All exploratory analyses will be clearly labeled as such. When multiple hypothesis tests are performed within exploratory analyses, *P* values will be adjusted to control the false discovery rate at 0.05.

Missing Data

The extent of missing data will be reported for all study variables. As described above, "I don't know" responses to cognitive (knowledge) items will be treated as missing, and their frequency will be reported.

Descriptive analyses will generally use available-case data for each variable so that participants with missing responses on a given measure can still contribute to analyses of other variables. For regression-based analyses, complete-case analyses will be performed for the variables included in each model. Missingness below 5% will be considered negligible. If more than 10% of values are missing for any key variable, patterns of missingness will be examined, and results will be interpreted cautiously.

Results

This study is not funded. Data collection began in January 2025 and is projected to conclude in January 2026. As of July 2025, a total of 164 participants from 3 selected schools have been recruited. The study has been well received, as reflected by the fact that all children with parental consent have provided verbal assent and participated. There has been no need to activate the child distress protocol. Data analysis is expected to begin in January 2026, and results are planned for publication in 2027.

Discussion

Study Rationale and Limitations

This study is among the first in Spain to evaluate attitudes toward dementia among schoolchildren. The proposed design and the pilot phase demonstrate the feasibility of the project.

Investigating this issue in a country such as Spain—one of the nations with the highest aging rates worldwide—is critically important. Over the past 5 decades, societal changes have led to a decline in intergenerational cohabitation, with most older adults no longer living with younger family members. As a result, children today are increasingly unfamiliar with aging and age-related diseases such as dementia.

As previously noted, due to the lack of a validated Spanish-language scale to assess children's attitudes toward dementia, a new questionnaire was developed. One prior study in Spain explored the emotional experiences and relational changes between grandchildren and grandparents with dementia, resulting in the development of a specific tool for that context [10]. However, its focus differed significantly as it addressed personal emotions such as guilt, jealousy, and satisfaction derived from helping without examining other domains of attitudes. Additionally, that study targeted adolescents and young adults aged 14 to 21 years, making it not directly comparable to this study.

The CANDem comprises cognitive, affective, and behavioral items. Compared with other dementia attitude scales validated in children and adolescents [15,47,48], this questionnaire includes a greater number of items assessing children's knowledge about the disease (eg, "All people develop dementia as they get older" or "There are no drugs that can cure dementia"). The decision to prioritize the assessment of knowledge over the other domains of attitude was informed by the well-established link between understanding dementia and subsequent feelings and behaviors toward people living with the condition [7]. Furthermore, knowledge is arguably the most modifiable component through school-based interventions and plays a critical role in shaping children's values and attitudes.

Participants in the pilot study completed the questionnaire within 20 minutes and described it as clear and easy to understand, supporting its feasibility for use with the selected age group. The children's strong interest in learning more about dementia aligns with findings from a British survey in which 62% of

XSL•FO
RenderX

children expressed a desire to help people with dementia but reported that a lack of understanding limited their ability to do so. Additionally, more than half of the surveyed children believed that life for people with dementia would improve if the general public had more knowledge about the condition [49].

Although the psychometric evaluation of the questionnaire is not the primary focus of this study, we considered it important to assess its internal consistency and validity.

Several limitations should be acknowledged; one of the most important ones is that some items in the CANDem may reflect common stereotypes about individuals with dementia (eg, "If someone with dementia gets angry or nervous, we must give them a pill to calm them down" or "People with dementia are sad almost all of the time"). These items were included to capture children's real attitudes, including misconceptions, rather than to reinforce or endorse such beliefs. To minimize potential bias, all items are presented neutrally, and care was taken during pilot-testing to ensure that they were age appropriate and clearly understood. Furthermore, after completing the survey, a 10-minute discussion session will be held in each classroom to allow children to ask questions and reflect on the items. While these precautions have been implemented, the inclusion of potentially stereotype-related items may have influenced responses in some children and may affect the interpretation of the findings.

Methodological limitations should also be noted. First, the cross-sectional design precludes causal inference, so relationships between attitudes and other variables cannot be interpreted as causal. Second, social desirability bias may influence children's responses. Despite efforts to present items neutrally and ensure a safe and age-appropriate survey environment, participants may answer in ways they perceive as socially acceptable rather than fully reflecting their true attitudes. To address this, the CANDem will be administered together with a child social desirability scale to assess the potential influence of this bias. Third, the psychometric evaluation of the CANDem will rely on EFA to identify its underlying structure. While EFA is appropriate for initial validation, future studies should conduct confirmatory factor analysis to test the stability and robustness of the factor structure derived from this work. Finally, the generalizability of the findings may be limited as the study is being conducted in schools within the Community of Madrid, and the CANDem may require adaptation for use in cultural contexts outside Spain considering differences in educational systems and family structures.

Conclusions and Implications

The proposed study appears to be a viable project that will contribute to a deeper understanding of children's attitudes toward dementia in Spain. By identifying existing attitudes and associated factors, it may inform the development of targeted interventions within the Spanish educational system aimed at increasing awareness and empathy among younger populations. These findings could also guide educational policy by supporting the inclusion of dementia-related content in school curricula, thereby fostering more inclusive and compassionate communities and potentially improving the care and quality of life of individuals living with dementia.

Acknowledgments

The authors extend their gratitude to the 10 children who took part in the pilot study and their parents, as well as the schools that have agreed to participate in the final survey. The authors confirm that no generative artificial intelligence tools were used in the preparation of this manuscript.

Funding

This study is ongoing and has received no specific grants from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed in this protocol.

Authors' Contributions

GS-F, MTV, JO-A, and JAS-R contributed to the conception and design of the work. GS-F and MTV contributed to the acquisition, analysis, and interpretation of the pilot study data. GS-F and MTV drafted the manuscript. GS-F, MTV, JO-A, and JAS-R read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Harwood RH, Sayer AA, Hirschfeld M. Current and future worldwide prevalence of dependency, its relationship to total population, and dependency ratios. Bull World Health Organ 2004 Apr;82(4):251-258. [Medline: 15259253]

XSL•FO
RenderX

2. GBD 2019 Dementia Forecasting Collaborators. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. Lancet Public Health 2022 Feb;7(2):e105-e125. [doi: 10.1016/S2468-2667(21)00249-8] [Medline: 34998485]
3. Arons AM, Krabbe PF, Schölzel-Dorenbos CJ, van der Wilt GJ, Rikkert M. Quality of life in dementia: a study on proxy bias. BMC Med Res Methodol 2013 Sep 6;13:110. [doi: 10.1186/1471-2288-13-110] [Medline: 24011428]
4. World Health Organization. Global action plan on the public health response to dementia 2017-2025. 2017. URL: https://apps.who.int/iris/bitstream/handle/10665/259615/9789241513487-eng.pdf [accessed 2024-01-22]
5. Kafadar AH, Barrett C, Cheung KL. Knowledge and perceptions of Alzheimer's disease in three ethnic groups of younger adults in the United Kingdom. BMC Public Health 2021 Jun 12;21(1):1124. [doi: 10.1186/s12889-021-11231-8] [Medline: 34118921]
6. Glynn RW, Shelley E, Lawlor BA. Public knowledge and understanding of dementia-evidence from a national survey in Ireland. Age Ageing 2017 Sep 1;46(5):865-869. [doi: 10.1093/ageing/afx082] [Medline: 28531240]
7. Chang CY, Hsu HC. Relationship between knowledge and types of attitudes towards people living with dementia. Int J Environ Res Public Health 2020 May 26;17(11):3777. [doi: 10.3390/ijerph17113777] [Medline: 32466533]
8. Batsch NL, Mittelman MS. World Alzheimer Report 2012: overcoming the stigma of dementia. : Alzheimer's Disease International; 2012 URL: https://www.alzint.org/u/WorldAlzheimerReport2012.pdf [accessed 2024-01-20]
9. Alfieri T, Ruble DN, Higgins ET. Gender stereotypes during adolescence: developmental changes and the transition to junior high school. Dev Psychol 1996;32(6):1129-1137. [doi: 10.1037//0012-1649.32.6.1129]
10. Celdrán M, Triadó C, Villar F. "My grandparent has dementia": how adolescents perceive their relationship with grandparents with a cognitive impairment. J Appl Gerontol 2011;30(3):332-352. [doi: 10.1177/0733464810368402]
11. Celdrán M, Triadó C, Villar F. Learning from the disease: lessons drawn from adolescents having a grandparent suffering dementia. Int J Aging Hum Dev 2009;68(3):243-259. [doi: 10.2190/AG.68.3.d] [Medline: 19522181]
12. Celdrán M, Villar F, Triadó C. Thinking about my grandparent: how dementia influences adolescent grandchildren's perceptions of their grandparents. J Aging Stud 2014 Apr;29:1-8. [doi: 10.1016/j.jaging.2013.12.004] [Medline: 24655668]
13. Farina N, Hughes LJ, Griffiths AW, Parveen S. Adolescents' experiences and perceptions of dementia. Aging Ment Health 2020 Jul;24(7):1175-1181. [doi: 10.1080/13607863.2019.1613343] [Medline: 31081351]
14. Chalhoub TM, Mackenzie E, Siette J. "Establishing healthy habits and lifestyles early is very important": parental views of brain health literacy on dementia prevention in preschool and primary school children. Front Public Health 2024;12:1401806. [doi: 10.3389/fpubh.2024.1401806] [Medline: 39165778]
15. Baker JR, Goodenough B, Jeon YH, Bryden C, Hutchinson K, Low LF. The Kids4Dementia education program is effective in improving children's attitudes towards dementia. Dementia (London) 2019 Jul;18(5):1777-1789. [doi: 10.1177/1471301217731385] [Medline: 31319740]
16. Burns P, Green C, Eady M, et al. Dementia knowledge, art, research and education - teaching primary school children about dementia. Health Promot J Austr 2021 Oct;32 Suppl 2:167-177. [doi: 10.1002/hpja.424] [Medline: 32966642]
17. Pawlicka N, Eady MJ, Green CA, et al. In their own words: an exploration of primary children's participation in a dementia education program. Health Promot J Austr 2024 Jan;35(1):251-259. [doi: 10.1002/hpja.739] [Medline: 37323033]
18. Noble JM, Hedmann MG, Williams O. Improving dementia health literacy using the FLOW mnemonic: pilot findings from the Old SCHOOL hip-hop program. Health Educ Behav 2015 Feb;42(1):73-83. [doi: 10.1177/1090198114537063] [Medline: 24893605]
19. Di Bona L, Kennedy S, Mountain G. Adopt a care home: an intergenerational initiative bringing children into care homes. Dementia (London) 2019 Jul;18(5):1679-1694. [doi: 10.1177/1471301217725420] [Medline: 28845688]
20. Atkinson T, Bray J. Dementia Awareness & Intergenerational Exchange in Schools: A Pioneer Project Supporting Dementia Friendly Communities. : University of Worcester; 2013 URL: https://www.worcester.ac.uk/documents/Schools-Evaluation-Final-Report.pdf [accessed 2024-01-22]
21. Smith AE, Kamm GL, Lai S, et al. A RE-AIM analysis of an intergenerational dementia education program. Front Public Health 2020;8:248. [doi: 10.3389/fpubh.2020.00248] [Medline: 32719762]
22. Breckler SJ. Empirical validation of affect, behavior, and cognition as distinct components of attitude. J Pers Soc Psychol 1984;47(6):1191-1205. [doi: 10.1037/0022-3514.47.6.1191] [Medline: 6527214]
23. Gandarillas A, Domínguez Berjón MF, Duque I, et al. III Atlas de Mortalidad y Desigualdades Socioeconómicas en la Comunidad de Madrid, 2008-2015. Consejería de Sanidad, Comunidad de Madrid. 2021 URL: https://www.madrid.org/bvirtual/BVCM050460.pdf [accessed 2026-04-16]
24. Cochran WG. Sampling Techniques: John Wiley & Sons; 1977.
25. Bartlett JE II, Kotrlik JW, Higgins CC. Organizational research: determining appropriate sample size in survey research. Inf Technol Learn Perform J 2001;19(1):43-50 [FREE Full text]
26. Charan J, Kaur R, Bhardwaj P, Singh K, Ambwani SR, Misra S. Sample size calculation in medical research: a primer. Ann Natl Acad Med Sci 2021;57(2):74-80. [doi: 10.1055/s-0040-1722104]
27. Morales Vallejo P. Guía Para Construir Cuestionarios y Escalas de Actitudes: Universidad Rafael Landívar; 2011.
28. O'Connor ML, McFadden SH. Development and psychometric validation of the Dementia Attitudes Scale. Int J Alzheimers Dis 2010;2010:1-10. [doi: 10.4061/2010/454218]

XSL•FO
RenderX

29. Serrani D. Construction and validation of a scale to assess attitudes toward patients with Alzheimer's disease and related disorders [Article in Spanish]. Rev Esp Geriatr Gerontol 2012;47(1):5-14. [doi: 10.1016/j.regg.2011.09.005] [Medline: 22209500]
30. Baker JR, Low LF, Goodenough B, et al. The KIDS Insight into Dementia Survey (KIDS): development and preliminary psychometric properties. Aging Ment Health 2018 Aug;22(8):947-953. [doi: 10.1080/13607863.2017.1320703] [Medline: 28481643]
31. Read ST, Wynaden D, Albrecht MA, Toye C. Development of the Dementia Community Attitudes Questionnaire. Dementia (London) 2021 Aug;20(6):1940-1957. [doi: 10.1177/1471301220977649] [Medline: 33307762]
32. Baker JR, Jeon YH, Goodenough B, et al. What do children need to know about dementia? The perspectives of children and people with personal experience of dementia. Int Psychogeriatr 2018 May;30(5):673-684. [doi: 10.1017/S1041610217002022]
33. Likert R. A technique for the measurement of attitudes. Arch Psychol 1932;140(22):1-55 [FREE Full text]
34. Preston CC, Colman AM. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. Acta Psychol (Amst) 2000 Mar;104(1):1-15. [doi: 10.1016/s0001-6918(99)00050-5] [Medline: 10769936]
35. O'Muircheartaigh C, Krosnick JA, Helic A. Middle alternatives, acquiescence, and the quality of questionnaire data. : Harris School of Public Policy Studies, University of Chicago; 2000. URL: https://ideas.repec.org/p/har/wpaper/0103.html [accessed 2026-04-10]
36. Nunnally JC, Bernstein IH. Psychometric Theory: McGraw-Hill Companies; 1994.
37. Mastaglia B, Toye C, Kristjanson LJ. Ensuring content validity in instrument development: challenges and innovative approaches. Contemp Nurse 2003 Jun;14(3):281-291. [doi: 10.5172/conu.14.3.281] [Medline: 12868667]
38. Blasco A. Tú, yo y el Alzheimer. Kern Pharma. 2016. URL: https://www.kernpharma.com/sites/default/files/documentos/28_COMIC%20ALZHEIMER.pdf [accessed 2024-04-20]
39. Lemos V. Construcción y validación de una escala para la evaluación de la deseabilidad social infantil (EDESI) [construction and validation of a scale to assess social desirability in children (EDESI)]. Interdisciplinaria 2005;22:77-96 [FREE Full text]
40. Institutional research board / research ethics board and informed consent. JMIR Publications. 2025. URL: https://support.jmir.org/hc/en-us/articles/360048970851--for-authors-Institutional-Research-Board-Research-Ethics-Board-and-Informed-Consent [accessed 2026-04-16]
41. Principles of Transparency and Best Practice in Scholarly Publishing.: Committee on Publication Ethics, Directory of Open Access Journals, Open Access Scholarly Publishing Association, World Association of Medical Editors; 2022 URL: https://publicationethics.org/guidance/guideline/principles-transparency-and-best-practice-scholarly-publishing [accessed 2026-04-16]
42. Maas CJ, Hox J. Sufficient sample sizes for multilevel modeling. Methodol Eur J Res Methods Behav Soc Sci 2005 Jan;1(3):86-92. [doi: 10.1027/1614-2241.1.3.86]
43. McNeish DM, Stapleton LM. Modeling clustered data with very few clusters. Multivariate Behav Res 2016;51(4):495-518. [doi: 10.1080/00273171.2016.1167008]
44. Hedges LV, Hedberg EC. Intraclass correlation values for planning group-randomized trials in education. Educ Eval Policy Anal 2007 Mar;29(1):60-87. [doi: 10.3102/0162373707299706]
45. Robinson JP, Shaver PR, Wrightsman LS. Measures of Personality and Social Psychological Attitudes: Gulf Professional Publishing; 1991.
46. Streiner DL, Norman GR, Cairney J. Health Measurement Scales: A Practical Guide to Their Development and Use, 5th edition: Oxford University Press; 2015.
47. Griffiths AW, Parveen S, Shafiq S, Oyebode JR. Development of the Adolescent Attitudes towards Dementia Scale (A-ADS). Int J Geriatr Psychiatry 2018 Aug;33(8):1139-1145. [doi: 10.1002/gps.4907] [Medline: 29851166]
48. Farina N, Griffiths AW, Hughes LJ, Parveen S. Measuring adolescent attitudes towards dementia: the revalidation and refinement of the A-ADS. J Health Psychol 2022 Feb;27(2):374-385. [doi: 10.1177/1359105320953479] [Medline: 32909455]
49. 2.3 million under 18s know someone with dementia but lack understanding. Alzheimer's Society. 2015 Mar 5. URL: https://www.dementiafriends.org.uk/WEBNewsStory?storyId=a0B2000000cL4hsEAC [accessed 2025-06-30]

Abbreviations

CANDem: Cuestionario de Actitudes de los Niños Españoles Hacia la Demencia (Spanish Children's Attitudes Toward Dementia Questionnaire)
COPE: Committee on Publication Ethics
EFA: exploratory factor analysis

XSL•FO
RenderX

Edited by J Sarvestan; submitted 30.Jul.2025; peer-reviewed by V Rashedi; revised version received 27.Dec.2025; accepted 27.Dec.2025; published 20.Apr.2026.

Please cite as:
Somoza-Fernández G, Ortiz-Alonso J, Serra-Rexach JA, Vidán MT
Children's Attitudes Toward Dementia: Protocol for a Cross-Sectional Study
JMIR Res Protoc 2026;15:e81576
URL: https://www.researchprotocols.org/2026/1/e81576
doi:10.2196/81576

© Gema Somoza-Fernández, Javier Ortiz-Alonso, José Antonio Serra-Rexach, María Teresa Vidán. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Impact of Short-Notice Accreditation Assessment on Hospitals' Patient Safety and Quality Culture: Protocol for a Mixed Methods Study

Robyn Scanlan[1*], BN, GradDipMid, MAdvNsgPrac, MBA, MPM; Tracy Flenady[2*], BN, PhD; Amy-Louise Byrne[2], BN, PhD; Jenni Judd[1], DHSc, MPH, MEd, DipHPE

[1]Reseach Division, Central Queensland University, University Drive, Bundaberg, Australia

[2]Central Queensland University, Rockhampton, Australia

[*]these authors contributed equally

Corresponding Author:
Robyn Scanlan, BN, GradDipMid, MAdvNsgPrac, MBA, MPM
Reseach Division, Central Queensland University, University Drive, Bundaberg, Australia

Abstract

Background: Accreditation programs are used by hospitals and health services to be independently reviewed against established safety and quality standards and have been a feature of global health systems for over a century. While evidence that accreditation programs directly improve patient safety and quality outcomes exists, the findings of various researchers remain mixed. Inefficiencies and a culture of "gaming" the system have also been observed, raising questions about the overall effectiveness of accreditation programs and assessment processes. Consequently, exploration of other formats of accreditation assessment, such as short-notice accreditation assessment, has arisen. From July 1, 2023, the Australian Commission on Safety and Quality in Healthcare mandated that Australian public and private hospitals must engage in short-notice accreditation assessment.

Objective: This study aims to explore the impact of short-notice accreditation assessment on hospitals, both in terms of safety and quality indicators, and organizational culture. A mixed methods design will be used to investigate these impacts.

Methods: Quantitative safety and quality indicators will be drawn from a regional health service prior to and following its first short-notice accreditation assessment cycle. From the same site, staff will be invited to complete the Patient Safety Culture Survey and participate in semistructured interviews. Using Schein's Culture Framework as an organizational culture model, the study will examine observable outcomes (artifacts, behaviors, and indicators) alongside staff perceptions and experiences (norms and values) to form an understanding of underlying assumptions and beliefs about short-notice accreditation assessment processes. Quantitative data will be analyzed through cross-tabulation, trend analysis, and other statistical techniques, while qualitative data will be synthesized to provide a comprehensive understanding.

Results: This protocol outlines the planned evaluation of short-notice accreditation assessment and its influence on patient safety and quality culture within a regional health service. Data collection is underway, with preintervention surveys being completed, and recruitment open for postintervention interviews. The study is expected to generate new knowledge on how this accreditation assessment process affects patient safety and quality culture of a regional and a rural hospital.

Conclusions: The findings will inform health policy on the suitability and long-term viability of short-notice accreditation assessment as an approach to ensuring safe, high-quality health care.

International Registered Report Identifier (IRRID): PRR1-10.2196/76945

(JMIR Res Protoc 2026;15:e76945) doi:10.2196/76945

KEYWORDS

mixed methods; health care quality; health care standards; staff perceptions; organizational culture

Introduction

Accreditation programs are the process by which hospitals and health services are assessed against contemporary and evidence-based quality and safety standards [1,2] to ensure minimum standards of care delivery are embedded in the organizational functioning, and that such outcomes are received by the recipients of care [3,4].

Accreditation programs in health care originated in the early 20th century as part of efforts to standardize hospital practices and improve patient care [3,5]. Initially, these processes were voluntary and focused on basic structural and procedural standards, such as sanitation and record-keeping. Over time,

XSL•FO
RenderX

accreditation programs evolved into a comprehensive system of external evaluation, incorporating evidence-based standards and performance indicators to assess hospitals' compliance with best practices [5,6]. Today, achievement of accreditation is widely recognized as a cornerstone of health care quality assurance globally, with most health service organizations (HSOs) viewing it as essential for maintaining public trust, meeting regulatory requirements, and driving continuous improvement [3]. In Australian hospitals, engagement in accreditation programs is now a mandated process for public and private services [7], and failure to meet specified standards can lead to regulatory consequences.

Despite accreditation's longevity in health care, debates persist about whether accreditation truly enhances patient outcomes or primarily promotes compliance behaviors rather than cultural transformation within organizations [6,8-10].

The strong assumption that accreditation programs hold HSOs to the expected level of care delivery and continually improve quality is supported at some level by research and practice [3,6,8,11], with some comparative studies being inconclusive [6,9,10]. In their Evidence Brief, Deeble Institute for Health Policy Research [9] identifies that research into HSO accreditation is lacking, and, subsequently, the value and impact on quality and safety outcomes are ambiguous [9]. Interestingly, Hussein et al [8] identify that accreditation may support improvement in performance measures, but that there is inconclusive evidence to link accreditation with an improved patient safety culture.

Many studies also demonstrate the unintended consequences of the established assurance processes [12-15]. Significantly, the resource investment into preparing and complying with an accreditation assessment has been reported to create additional burden and costs in engaging in accreditation [15-17]. Additionally, as accreditation outcomes may be publicly available, organizational actions that may pursue a positive result could be prioritized [18] over processes to achieve continuous patient care improvement [17-19]. This can result in inefficiencies in the standard accreditation process [6,19,20], a heavy investment in resources to achieve a successful outcome [18,20], and focus on compliance rather than quality [18,19].

Indeed, contemporary literature increasingly critiques the adequacy of hospital accreditation in capturing the true culture and outcomes of HSOs. While accreditation is widely accepted as the mechanism for quality assurance, its effectiveness in driving meaningful cultural change remains contested. Lewis and Hinchcliff [21] argue that the body of research on accreditation is largely atheoretical and fails to explain how or why accreditation programs influence quality improvement, noting that most studies focus on compliance rather than cultural transformation. Similarly, Husseinet al [8] found that although accreditation programs may positively affect safety culture and process-related performance measures, their impact on patient outcomes such as satisfaction and readmission rates is inconsistent and often negligible. These findings suggest that attainment of accreditation may not be a reliable proxy for assessing the deeper, more nuanced aspects of organizational culture.

Further, a growing concern in the literature is the phenomenon of "accreditation readiness" or "gaming the system," where health services prepare intensively for scheduled assessments, potentially masking the everyday realities of care delivery. The concept of "Accreditation Theater," as described by the Accreditation Commission for Health Care [22], highlights how punitive and performative assessment practices can create an illusion of rigor without advancing actual care quality. Tactics used during the accreditation assessment may include, but are not limited to, the distraction of accreditors to divert attention away from potential areas of noncompliance; rostering practices to ensure the most compliant staff are present during the assessment review; cognitive bias creation through relationship development with surveyors; and planning and timetabling ensuring areas of noncompliance are potentially hidden during accreditation reviews [19]. Brubakk, Vist [23] further emphasize that the push for accreditation persists despite scant evidence of its effectiveness, and that most studies fail to account for the context and implementation of accreditation processes, which are critical to understanding their true impact. These critiques underscore the need for accreditation models that move beyond episodic compliance checks and instead foster continuous cultural and quality improvement embedded in everyday practice.

Nonetheless, HSOs generally support the value of accreditation programs, which are seen as essential in driving engagement with and developing collective quality improvement [8,24-26] and improvements in organizational cultures generally [27]. However, inefficiencies in the standard accreditation assessment process have raised questions about the value and effectiveness of a planned and cyclic accreditation procedure [8,28].

To drive a continuous state of accreditation readiness, short-notice or unannounced formats of accreditation assessment have been introduced [28,29]. A short-notice accreditation assessment process provides the HSO with a short period of time (eg 48-hours' notice) that surveyors will be onsite to undertake an assessment; and an unannounced accreditation assessment process provides no notice to the HSO that surveyors will be onsite to undertake an assessment. Notably, the American health care accreditation body, the Joint Commission, introduced unannounced visits in 2006 with the goal of ensuring quality care at all times, built on the premise that it was the "right thing to do" [30]. More recently, the Danish Institute for Quality and Accreditation in Healthcare undertook a trial of unannounced accreditation surveys [19]. Ehlers et al [19] found that unannounced accreditation surveys were not more effective in detecting quality issues in hospitals; however, the study did not examine the impact of an unannounced model on the safety and quality culture of HSOs. Subsequently, an Australian-first trial of a short-notice accreditation assessment occurred in 2017, resulting in changes to the Australian Health Service Safety and Quality Accreditation Scheme (AHSSQAS) mandating short-notice accreditation assessment processes for all Australian public and private hospitals [7,17,20].

Those Australian HSOs engaged in short-notice accreditation assessment trial argue this type of process ensures assessment results reflect the reality of the HSO safety and quality environment, and embeds a culture of quality and safety

XSL•FO
RenderX

functioning within the daily practice, thus improving patient care [20], experience [31], and outcomes [20]. Importantly, short-notice accreditation assessment offers an alternative to traditional accreditation assessment processes, which may reduce opportunities for gaming. Indeed, short-notice or unannounced assessment processes of accreditation have been proposed to minimize the resource investment in preparing for accreditation [7,20], support the reliability of the accreditation outcomes [17,20,28], and reposition the accreditation program as a tool to support HSOs' quality and safety agenda [20,29].

Despite the change to a new format of accreditation assessment, there is a paucity of evidence that supports the use of a short-notice assessment process [19,28], and little research that explores the impact that this process has on the culture of quality and safety for a health service. Scanlan et al [32] concluded that there is limited evidence reporting on the effectiveness of short-notice accreditation assessment models used by hospitals, and that there are no studies undertaken which seek to understand the impact a short-notice accreditation assessment model can have on patient safety and quality cultures. As such, this study protocol details a mixed methods study which uses Schein's Culture Framework [33] as a theoretical model within a postpositivism paradigm to evaluate the impact of a short-notice accreditation assessment.

Methods

Research Aims

This research aims to understand the impact of short-notice accreditation assessment processes on HSO patient safety and quality culture, and to understand the impact of a short-notice accreditation assessment process from a safety and quality indicator perspective.

Research Objectives

The research objectives are as follows:

1. Measure and compare HSOs' safety and quality key performance indicators before and after the short-notice accreditation assessment.
2. Evaluate the effectiveness of short-notice accreditation in increasing individual accountability and translating safety and quality standards into clinical practice.
3. Assess the extent to which short-notice accreditation embeds a sustainable safety and quality culture within the organization.

Theoretical Framework

This study adopts a postpositivist orientation consistent with its mixed methods design and focus on empirical evaluation of organizational outcomes. Postpositivism supports the use of quantitative indicators to examine patterns and relationships in safety and quality performance, while also recognizing the value of qualitative inquiry to explore staff perceptions and experiences that cannot be fully captured through measurement alone.

Schein's Culture Framework [33] provides a pragmatic structure for integrating these data sources by linking observable organizational artifacts and behaviors with reported norms, values, and assumptions. This approach enables a comprehensive examination of how short-notice accreditation processes may influence patient safety and quality culture without requiring commitment to a single explanatory theory of organizational reality.

Measuring organizational culture is notoriously difficult and complex [34], with agreed indicators lacking [35]. In simplest terms, culture is the pattern of shared basic assumptions within a group [33]. Typically, HSOs focus on the observable and easily measured outcomes to infer the underlying culture of their organization [36,37]. Research into determinants of positive patient safety and quality health outcomes indicates culture is a significant driver [38-42]; however, limited information in triangulating accreditation, health service outcomes, and culture exists [43-46].

No standard tool for measuring organizational culture of complex systems such as HSOs exists; however, the Schein Cultural Framework [33] remains constant in published approaches [47-49]. Cacciattolo [50] synthesized contemporary knowledge and research into applying models for understanding and analyzing organizational culture, finding that 2 main approaches to understanding organizational culture exist: interpretive and structural [50]. The Schein Culture Framework [33] fits within the structural approach [50] and provides a valuable model for understanding a culture's influence on HSO indicators. Hogan and Coote [51] and Cacciattolo [50] have identified the Framework's wide acceptance and practical application for understanding and influencing organizational cultural change. While Hogan and Coote [51] have recognized the limited research validating Schein's Culture Framework [33], they have also demonstrated the value of using the model to build a positive organizational culture [51]. Understanding the determinants of a positive patient safety and quality culture within the health care setting is central to this study, and Schein's Culture Framework [33] was determined to be beneficial in devising this research approach.

Schein's Culture Framework [33] is particularly well-suited for evaluating organizational culture in health care settings due to its multilayered approach that captures both visible and invisible dimensions of culture [49]. Compared to other models, Schein's [33] framework is uniquely positioned to explore organizational-level culture in complex systems like hospitals. Indeed, Schein's Cultural Framework [33] excels in diagnosing and influencing culture within specific institutions by acknowledging the interplay between structure, meaning, and behavior. Moreover, its compatibility with critical realism, as discussed by Khaddour [52], supports a nuanced understanding of culture that avoids oversimplification and embraces complexity, making it ideal for mixed methods research in health care. This theoretical robustness and practical applicability justify its use in studies aiming to assess the cultural impact of short-notice accreditation processes.

Schein [33] proposes a culture framework where understanding the norms, values, and underlying assumptions can be linked to the observable outcomes of organizational indicators (Figure 1).

XSL•FO
RenderX

Figure 1. Schein's [33] culture framework.

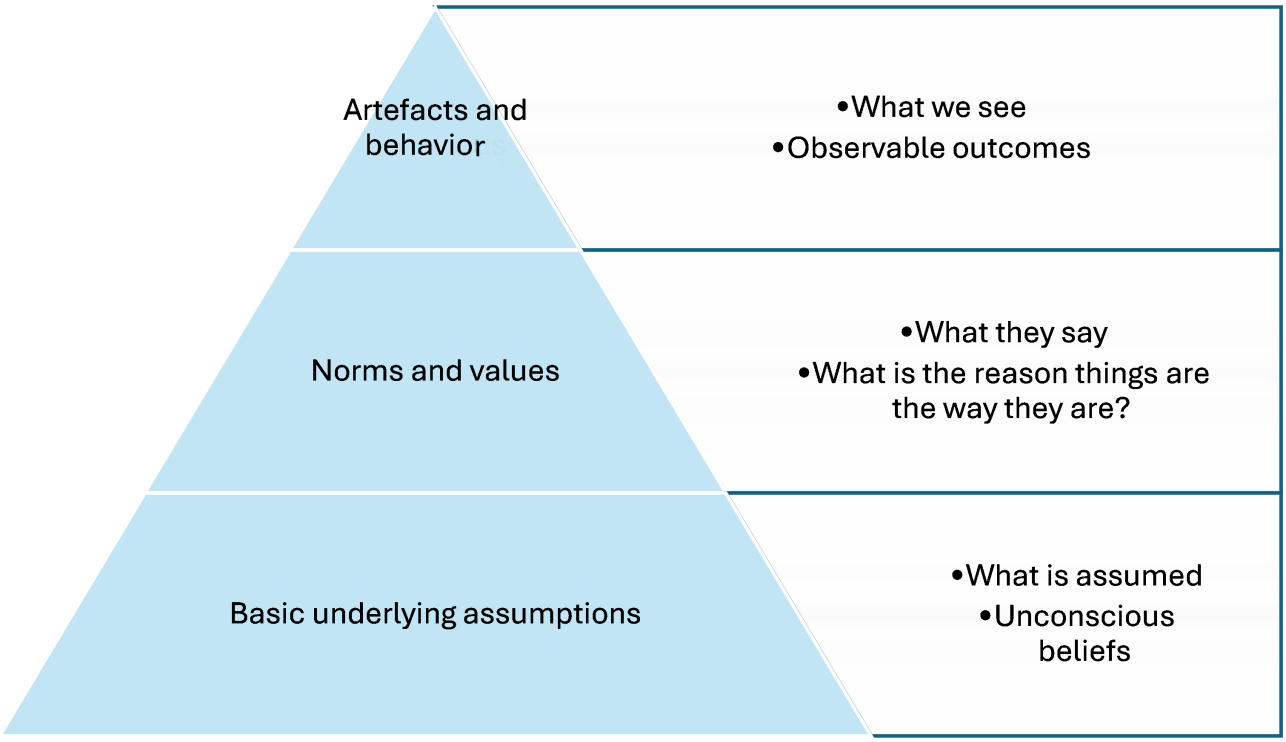

Using the Schein [33] Culture Framework as a model, the research will consider the effectiveness of a short-notice accreditation assessment process to embed a culture of safety and quality, demonstrated through observable outcomes in organizational indicators. The research will test the underlying but inconclusive belief that accreditation and, subsequently, short-notice accreditation assessment improves patient safety and quality outcomes.

Design

Considering both organizational culture and impact on quantitative and qualitative indicators lends this research to a mixed methods approach and triangulation of the resultant data through a convergent design process [53]. A mixed methods approach to the research will be used with reporting adhering to the GRAMMS (Good Reporting of A Mixed Methods Study) [54] guidelines.

Using a convergent mixed methods design, the research methods and data collection will progress as per the basic flowchart in Figure 2.

XSL•FO
RenderX

Figure 2. Mixed methods design.

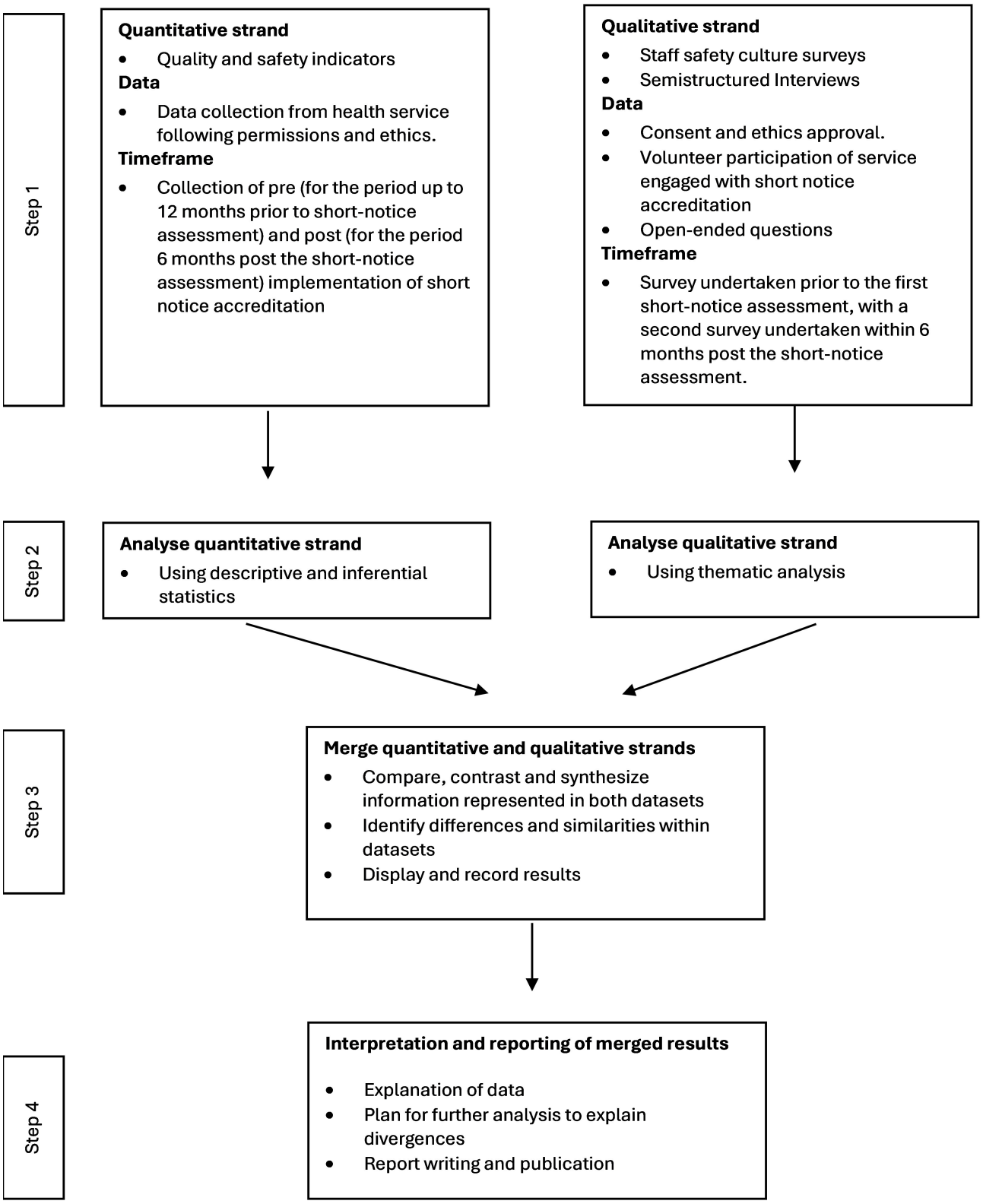

(Adapted from Creswell [53]). Hogden et al [55] identified that a mixed methods approach is advocated when evaluating health care safety culture [55]. Their literature review identified that no individual tool could evaluate the safety culture and recommended that a mixed methods assessment of culture as one of the approaches that could be taken [55].

The quantitative data strand will use descriptive and inferential analysis methods. Descriptive statistical methods summarize, describe, and draw conclusions on the population or datasets, while inferential statistical methods use the sample to draw conclusions and make predictions [56]. For the qualitative branch, coding and thematic analysis will be used. Both data strands will be merged to compare, contrast, and synthesize the information, leading to interpretation and reporting, as per Figure 2.

XSL•FO
RenderX

Table 1 illustrates how the anticipated research measures map to Schein's Culture Framework levels [33].

Table . Anticipated research measures versus Schein's [33] culture framework levels.

Schein's levels	Definition	Anticipated research measures
Artifacts and behaviors	What we see Observable outcomes	Established health service safety and quality indicators
Norms and values	What they say What is the reason why things are the way they are	Staff perceptions and experience
Basic underlying assumptions	What is assumed Unconscious beliefs	Accreditation improves patient safety (not a measure, but an underlying belief or assumption)

Research Environment

The research will be conducted in one Queensland (Australia) public HSO. The research will include two hospitals: a 331-bed regional hospital and a 36-bed rural hospital. Both hospitals provide various medical services, including emergency, inpatient, outpatient, and allied health services. While the short-notice accreditation format of assessment has been mandated since July 1, 2023, this HSO has yet to engage in their first short-notice accreditation assessment. Including both hospitals from this HSO offers an opportunity to evaluate the applicability of a short-notice accreditation assessment process across various contexts. Importantly, the research environment is shaped not only by the characteristics of the hospitals themselves but also by the external factors such as policy or funding shifts, which may influence organizational culture. By selecting hospitals within the same service area, the researcher can better account for these contextual influences. Interview questions will further explore organizational norms, values, and underlying assumptions about accreditation processes that underpin a patient safety and quality culture.

Data

Overview

A scoping review [32] conducted by the research team identified gaps in understanding the impact of short-notice accreditation. Based on these findings, the study will focus on three components of data collection: (1) safety and quality indicators, (2) patient safety culture surveys, and (3) staff interviews.

All safety and quality indicators, as well as participants for the surveys and interviews, will be drawn from the regional HSO in which the study is being conducted. Figure 3 provides a graphical representation of the data collection timeline.

Figure 3. Data collection timeline.

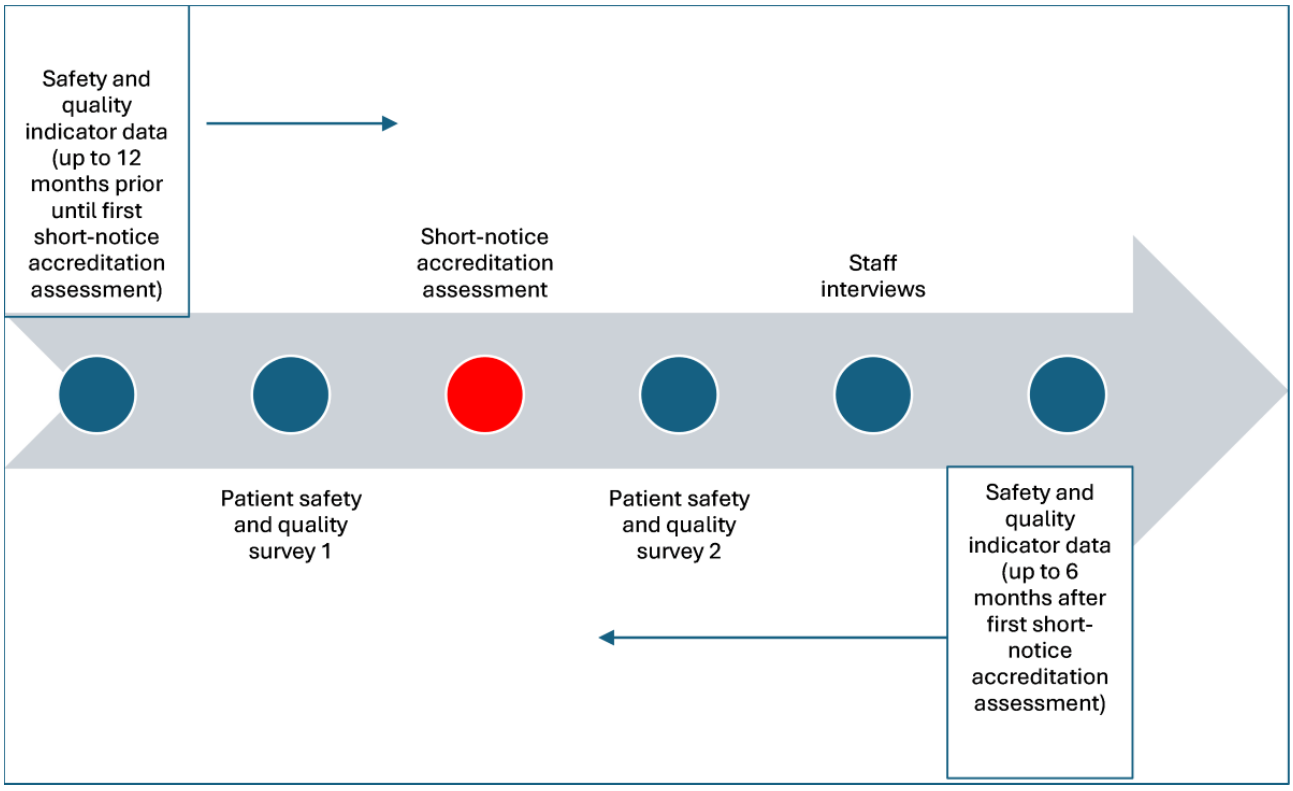

Data from all domains will be collected longitudinally. Safety and quality indicators will be examined for the 12 months preceding the HSO's first short-notice accreditation assessment to establish baseline trends. Because the HSO will not know

XSL•FO
RenderX

the exact timing of their assessment, these 12 months of preassessment data will serve as the baseline. Postassessment data will then be collected for 6 months following the first short-notice accreditation assessment, enabling comparison to determine whether the assessment influenced safety and quality indicators. While shorter than the pre-period, the 6-month time frame postassessment is adequate to identify trends pre- and postaccreditation assessment and to inform researchers of future research direction.

Patient safety culture surveys will be administered both before and after the first short-notice accreditation assessment. Staff interviews will be conducted after the accreditation assessment, allowing participants to reflect on and compare their experiences of previous accreditation processes with the new short-notice format.

Safety and Quality Indicators

HSOs routinely collect quality and safety indicators as part of their routine governance [37,57]. The indicators are widely recognized as being linked to hospital performance, safety culture, and patient outcomes [36]. Numerous studies have also used clinical quality and safety indicators to evaluate the impact of accreditation processes [19,29,57,58].

In Australia, HSOs establish safety and quality indicators in alignment with the National Safety and Quality Standards [1] which are reliably collected against agreed data specifications. For this study, standard indicators aligned to the National Safety and Quality Standards [59] will be used.

Table 2 presents the indicator population used within the research study.

Table . Safety and quality indicators.

National safety and quality standard	Process and outcome indicator set	Balancing or linking indicators
Standard 1 – Clinical Governance	• Hospital standardized mortality ratio • Deaths in low-risk diagnostic related groups Registered quality improvements • Severity assessment code 1 clinical incident rates	• Readmission rates • Unplanned return to theater • Length of stay • Relative stay index
Standard 2 – Partnering with Consumers	• Patient reported experience measures	• Readmission rates • Unplanned return to theater • Length of stay • Relative stay index
Standard 3 – Infection Control	• Hospital acquired infections	• Readmission rates • Unplanned return to theater • Length of stay • Relative stay index
Standard 4 – Medication Safety	• Reported medication incidents – severity assessment code 1 and 2 • Medication complications	• Readmission rates • Unplanned return to theater • Length of stay • Relative stay index
Standard 5 – Comprehensive Care	• Pressure injuries, falls, malnutrition, delirium	• Readmission rates • Unplanned return to theater • Length of stay • Relative stay index
Standard 6 – Communicating for Safety	• Discharge summary completion rates	• Readmission rates • Unplanned return to theater • Length of stay • Relative stay index
Standard 7 – Blood Management	• Hemovigilance incidents – severity assessment code 1 and 2	• Readmission rates • Unplanned return to theater • Length of stay • Relative stay index
Standard 8 – Recognizing and Responding to Acute Deterioration	• Deterioration scoring • Response escalation • Rapid response calls • Deterioration patient incidents - severity assessment code 1 and 2 • Patient or family escalation of care	• Readmission rates • Unplanned return to theater • Length of stay • Relative stay index

XSL•FO
RenderX

This quantitative data will be sourced from health service-specific repositories accessible to the research team under the ethics approval protocol. Indicators will be collected for a period beginning 12 months prior to the HSO's first short-notice accreditation assessment and extending to 6 months post assessment. This timeframe provides both a preassessment baseline and a postassessment follow-up. In contrast to the traditional 3 to 4-year accreditation cycle, this design allows a timely assessment of the potential influence of the short-notice accreditation process on organizational safety and quality indicators.

Patient Safety Culture Surveys

The assessment of a HSO's safety culture is mandated by Australian governing bodies [59]. To support this, the Australian Commission on Safety and Quality in Health Care [60] recently released the Australian Hospital Survey on Patient Safety Culture 2.0 (A-HSOPS 2.0) under a Creative Commons license. The validated tool, specifically designed for use within Australian hospitals, will be used to capture staff perceptions of safety culture within the participating HSO. The survey can be securely administered online, allowing staff to provide responses anonymously.

Participation will be voluntary, with the target sample drawn from the population of staff of the HSO undergoing short-notice accreditation. Based on a 90% CI, the required sample size is 66 clinical staff [61]. As the questionnaire records respondents' professional role, alignment between the response profile and the overall staffing profile will further support the validity of findings.

To ensure confidentiality, the survey will not collect identifying data beyond: (1) role in the organization, (2) department of employment, (3) years of service in the organization, and (4) gender.

These demographic variables will enable comparison of survey respondents with the overall staff population, while aggregated reporting will minimize any risk of re-identification.

Staff Interviews

To supplement the quantitative findings from the HSO quality and safety indicators and safety culture surveys, semistructured staff interviews will be undertaken to provide clarification and contextualization. This qualitative component will help validate and deepen the interpretation of the quantitative data. Semistructured interviews, using open-ended questions, will elicit greater depth and understanding of the subjects' experience [61]. Participants will be drawn from across the HSO undergoing short-notice accreditation, with representation sought from all organizational levels, including executive, governance, clinical, support, and operational staff. Participation will be voluntary.

Interview guides will be informed by analysis of indicators and survey data with questions designed to explore areas where findings may be limited, unclear, or conflicting (Multimedia Appendix 1). Further, interviews will focus on staff's perceptions of short-notice accreditation, their experience of organizational culture, and the extent to which these domains may intersect.

Data Analysis

Organizational Quality and Safety Indicators

Data from organizational safety and quality indicators will be reviewed in relation to accreditation processes. Cross-tabulation, trend analysis, and statistical analysis techniques will be applied. Both descriptive and inferential statistical methodologies will be used to provide a comprehensive understanding of indicator trends and their relationship to short-notice accreditation [61].

Staff Safety Culture Surveys

Survey responses will generate statistical data reflecting staff's perceptions of safety culture. Cross-trend analysis and inferential techniques, including ANOVA, will be applied to compare differences across staff groups and over time. As with the indicator data, descriptive and inferential statistics will be used to build a comprehensive understanding of patterns and changes [61]. Data will be exported from Qualtrics for a pre- and poststatistical analysis through time-series comparison.

Staff Interviews

Interview transcripts will be analyzed using a qualitative data analysis approach, with coding to identify patterns, themes, and relationships between short-notice accreditation processes and organizational culture. Braun and Clarke [62] 6-phase thematic analysis framework will guide the analysis, ensuring a rigorous and systematic interpretation of staff experiences and perspectives.

Data Compilation and Report Writing

Data compilation and report writing will occur during steps 3 and 4 of the convergent mixed methods process.

Table 3 provides an overview of data sources aligned with Schein's [33] organizational culture levels and illustrates how each research artifact addresses the research questions.

XSL•FO
RenderX

Table . Overview of data compilation.

Schein level	Research question	Research artifact	Population	Analysis
Artifacts and behaviors	What impact does a short-notice accreditation process have on a health service organization's core safety and quality indicators?	Quantitative analysis of organizational quality and safety indicators	Listed in indicators in Table 1 from a single Hospital and Health Service	Statistical analysis Trend over time-correlated to short-notice accreditation cycles and historical accreditation processes
Norms and values	What impact do accreditation processes have on the perceptions of health service organizational safety and quality culture?	Use of mixed methods • Staff Safety Culture Surveys • Qualitative Staff Interviews	Representative staff at all levels (ward to executive) Nonclinical, clinical, management, executive from a single Hospital and Health Service	Using mixed methods protocols: • Survey statistical analysis • Thematic analysis
Basic underlying assumption	What impact does a short-notice accreditation process have on a health service's organizational safety and quality? How does a short-notice accreditation process support a health service organization's ability to embed a safety and quality culture?	What impact does a short-notice accreditation process have on a health service's organizational safety and quality? How does a short-notice accreditation process support a health service organization's ability to embed a safety and quality culture?	What impact does a short-notice accreditation process have on a health service's organizational safety and quality? How does a short-notice accreditation process support a health service organization's ability to embed a safety and quality culture?	Merging of qualitative and quantitative data identifying differences and similarities Interpretation and reporting of merged results to answer the question

Table 3 identifies the primary research question for this project. When the datasets are converged, joint displays and triangulation will occur during interpretation, thus contributing to addressing the primary research question.

The qualitative and quantitative strands of data will then be merged for comparative analysis. This process will enable triangulation of findings, allowing discrepancies and inconsistencies to be interrogated and similarities and differences to be explored. The convergent design ensures that outcomes are interpreted in alignment with Schein's Culture Framework [33], strengthening the validity of the findings. Where divergences are evident, further analysis will be undertaken to provide an increased understanding of the synthesized results.

Ethical Considerations

This project has been approved by Queensland Health Human Research Ethics Committee, approval number LNR/2022/QGC/84035 (August 4, 2022), and CQUniversity Human Research Ethics Committee, approval number 0000023861 (September 14, 2022).

Results

The study is currently collecting the preintervention of patient safety and quality survey data prior to the HSO engaging in their first short-notice accreditation process (Figure 3). Through the survey data collection, the study is also enrolling interested participants for postintervention interviews. Final data and results are expected to be completed 6 months after the short-notice accreditation process. The date of this process is not known to the HSO nor the researchers.

Discussion

Anticipated Findings

This study protocol outlines research designed to evaluate the impact of a short-notice accreditation process on the patient safety and quality culture of a health care organization. Existing research in health service accreditation provides limited insight into short-notice accreditation influences organizational culture, and this study aims to contribute to addressing that gap.

In their scoping review, Scanlan et al [32] identified limited and sometimes conflicting evidence on the value of short-notice accreditation processes. The reviewed literature, findings were inconsistent in demonstrating whether accreditation improved safety and quality performance [18,19,27,63]. This lack of consensus highlights the importance of further research to provide robust evidence on the cultural and organizational impacts of short-notice accreditation.

With the recent mandate requiring short-notice accreditation within the Australian accreditation scheme [17], there is a clear imperative to understand the effectiveness of such a format. Grounding this research in Schein's Culture Framework [33] provides a structured lens through which to examine organizational culture across artifacts and behaviors, norms and values, and basic underlying assumptions. When combined with a mixed methods design, this approach allows for a comprehensive exploration of how short-notice accreditation may shape patient safety and quality culture. Findings from this study have the potential to inform both national and international health policy, while also extending the validation of Schein's Culture Framework [33] in health care settings.

Limitations

This study is limited to 2 hospitals within a single regional HSO. While findings will provide valuable insights into the impact

XSL•FO
RenderX

of short-notice accreditation in this context, the applicability of results to larger, metropolitan, or differently structured services can only be inferred and will not be directly examined within the scope of this protocol.

An important limitation of this study is the potential influence of external factors on both the organizational culture and the safety and quality indicators being measured. The unpredictable timing of short-notice accreditation assessments means that data collection may coincide with broader systemic changes such as policy reforms, funding adjustments, seasonal fluctuations in patient demand, or workforce transitions. These factors can independently affect staff perceptions, organizational performance, and cultural dynamics, making it difficult to isolate the impact of accreditation alone. For example, staffing shortages or leadership changes during the data collection period may skew survey responses or interview narratives, while external policy shifts could influence indicator trends unrelated to accreditation. Although the study design attempts to mitigate these risks by selecting hospitals within the same HSO, the complexity of health care environments means that such confounding variables cannot be fully controlled. Future research may benefit from multi-site studies or longitudinal designs that better account for these contextual influences.

The staff safety culture survey and interviews also introduce methodological constraints. Survey participation is voluntary and may be subject to low response rates or response bias, with some staff groups under-represented. Interview participants, while purposively sampled to represent all organizational levels, may not reflect the full diversity of staff experiences. In addition, reliance on self-report raises the possibility of social desirability or recall bias. Finally, the study applies Schein's Culture Framework as its theoretical lens. While this provides a structured approach to examining organizational culture, it may not capture all relevant dimensions that influence how accreditation processes shape safety and quality culture.

Conclusions

The short-notice accreditation format of assessment has been introduced into the AHSSQAS to strengthen and improve accreditation processes. However, evidence on whether this format enhances patient safety and quality culture remains limited. By applying Schein's Culture Framework [33] within a mixed methods design, this study seeks to generate new knowledge on the organizational impact of short-notice accreditation. Findings will contribute to understanding how accreditation influences the artifacts, values, and underlying assumptions that shape patient safety and quality culture in HSOs. In doing so, the research aims to inform both policy and practice, nationally and internationally, on the effectiveness of short-notice accreditation as a strategy for improving safety and quality in health care.

Acknowledgments

This protocol is the original work of the authors. The first author made substantial contributions to the conception and design of the protocol. All authors have made substantial contributions to the drafting of the protocol. The second and third authors made substantial contributions to the critical revision of the protocol. All authors approve the final manuscript. All authors prepared and submitted ethics applications.

All authors attested that generative artificial intelligence has not been used in any portion of the manuscript generation.

Funding

Research Higher Degree (RHD) candidate funding has been received. The authors acknowledge the support of the Australian Government.

Authors' Contributions

Conceptualization, formal analysis, investigation, writing – original draft, writing – review and editing, and visualization: RLS
Conceptualization, formal analysis, writing – review and editing, and supervision: TF
Conceptualization, formal analysis, writing – review and editing, and supervision: JJ
Conceptualization, writing – review and editing, and supervision: ALB

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample interview questions.
[DOCX File, 14 KB - resprot_v15i1e76945_app1.docx]

References

1. McPhail R, Avery M, Fisher R, Fitzgerald A, Fulop L. The changing face of healthcare accreditation in Australia. Asia Pacific Journal of Health Management 2015;10(2):58-64 [FREE Full text]

XSL•FO
RenderX

2. Scrivens E. Accreditation and the regulation of quality in health services. In: Regulating Entrepreneurial Behaviour in European Health Care Systems: Open University Press; 2002:91-105 URL: https://www.researchgate.net/profile/Ana-Rico-3/publication/282858573_What_can_we_learn_from_the_regulation_of_public_utilities/links/561f97d108aea35f267dee94/What-can-we-learn-from-the-regulation-of-public-utilities.pdf#page=107 [accessed 2026-04-14]
3. World Health Organization. Health care accreditation and quality of care: exploring the role of accreditation and external evalutation of health care facilities and organizations. 2022 URL: https://www.who.int/publications/i/item/9789240055230 [accessed 2026-04-09]
4. Pomey MP, Lemieux-Charles L, Champagne F, Angus D, Shabah A, Contandriopoulos AP. Does accreditation stimulate change? A study of the impact of the accreditation process on Canadian healthcare organizations. Implement Sci 2010 Apr 26;5(1):31. [doi: 10.1186/1748-5908-5-31] [Medline: 20420685]
5. Alkhenizan A, Shaw C. Impact of accreditation on the quality of healthcare services: a systematic review of the literature. Ann Saudi Med 2011;31(4):407-416. [doi: 10.4103/0256-4947.83204] [Medline: 21808119]
6. Araujo CAS, Siqueira MM, Malik AM. Hospital accreditation impact on healthcare quality dimensions: a systematic review. Int J Qual Health Care 2020 Nov 9;32(8):531-544. [doi: 10.1093/intqhc/mzaa090] [Medline: 32780858]
7. Fact sheet 17: short notice accreditation assessment. : Australian Commission on Safety and Quality in Health Care; 2024 URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/fact-sheet-17-short-notice-accreditation-assessment [accessed 2026-04-09]
8. Hussein M, Pavlova M, Ghalwash M, Groot W. The impact of hospital accreditation on the quality of healthcare: a systematic literature review. BMC Health Serv Res 2021 Oct 6;21(1):1057. [doi: 10.1186/s12913-021-07097-6] [Medline: 34610823]
9. Deeble Institute for Health Policy Research. Evidence brief 18: assessing the value of accreditation to health systems and organisations. : Deeble Institute; 2019 URL: https://ahha.asn.au/wp-content/uploads/2024/04/Evidence-Brief-No.18-Assessing-the-value-of-accreditation.pdf [accessed 2026-04-09]
10. Van Wilder A, Bruyneel L, De Ridder D, et al. Is a hospital quality policy based on a triad of accreditation, public reporting and inspection evidence-based? A narrative review. Int J Qual Health Care 2021 May 28;33(2):mzab085. [doi: 10.1093/intqhc/mzab085] [Medline: 34013956]
11. Hinchcliff R, Greenfield D, Moldovan M, et al. Narrative synthesis of health service accreditation literature. BMJ Qual Saf 2012 Dec;21(12):979-991. [doi: 10.1136/bmjqs-2012-000852] [Medline: 23038406]
12. Engel C, Andersen HB. Side effects of overdoing it: lessons from a comprehensive hospital accreditation programme. In: Researching Patient Safety and Quality in Healthcare: CRC Press; 2016:115-127 URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781315605609-10/side-effects-overdoing-carsten-engel-henning-boje-andersen [accessed 2026-04-14] [doi: 10.1201/9781315605609-10]
13. Ho MJ, Chang HH, Chiu YT, Norris JL. Effects of hospital accreditation on medical students: a national qualitative study in Taiwan. Acad Med 2014 Nov;89(11):1533-1539. [doi: 10.1097/ACM.0000000000000481] [Medline: 25250745]
14. Katoue MG, Somerville SG, Barake R, Scott M. The perceptions of healthcare professionals about accreditation and its impact on quality of healthcare in Kuwait: a qualitative study. J Eval Clin Pract 2021 Dec;27(6):1310-1320. [doi: 10.1111/jep.13557] [Medline: 33749091]
15. Mosadeghrad AM. Hospital accreditation: the good, the bad, and the ugly. Int J Healthc Manag 2021 Oct 2;14(4):1597-1601. [doi: 10.1080/20479700.2020.1762052]
16. Mumford V, Forde K, Greenfield D, Hinchcliff R, Braithwaite J. Health services accreditation: what is the evidence that the benefits justify the costs? Int J Qual Health Care 2013 Oct;25(5):606-620. [doi: 10.1093/intqhc/mzt059] [Medline: 23942825]
17. Review of the Australian health service safety and quality accreditation scheme: improving the reliability of health service organisation accreditation processes. : Australian Commission on Safety and Quality in Health Care; 2018 URL: https://www.safetyandquality.gov.au/sites/default/files/2019-06/review-of-the-australian-health-service-safety-and-quality-accreditation-scheme-improving-the-reliability-of-health-service-organisation-accreditation-processes-0.pdf [accessed 2026-04-09]
18. Aghaei Hashjin A, Delgoshaei B, Kringos DS, Tabibi SJ, Manouchehri J, Klazinga NS. Implementing hospital quality assurance policies in Iran: balancing licensing, annual evaluation, inspections and quality management systems. Int J Health Care Qual Assur 2015;28(4):343-355. [doi: 10.1108/IJHCQA-03-2014-0034] [Medline: 25982635]
19. Ehlers LH, Simonsen KB, Jensen MB, Rasmussen GS, Olesen AV. Unannounced versus announced hospital surveys: a nationwide cluster-randomized controlled trial. Int J Qual Health Care 2017 Jun 1;29(3):406-411. [doi: 10.1093/intqhc/mzx039]
20. Uren H, Vidakovic B, Daly M, Sosnowski K, Matus V. Short-notice (48 hours) ACCREDITATION trial in Australia: stakeholder perception of assessment thoroughness, resource requirements and workforce engagement. BMJ Open Qual 2019;8(3):e000713. [doi: 10.1136/bmjoq-2019-000713] [Medline: 31637325]
21. Lewis K, Hinchcliff R. Hospital accreditation: an umbrella review. Int J Qual Health Care 2023 Feb 24;35(1):mzad007. [doi: 10.1093/intqhc/mzad007] [Medline: 36738157]
22. Accreditation theater: the illusion of rigor in healthcare accreditation. : Accreditation Commission for Health Care; 2025 URL: https://achc.org/accreditation-theater-the-illusion-of-rigor-in-healthcare-accreditation/ [accessed 2026-04-09]

XSL•FO
RenderX

23. Brubakk K, Vist GE, Bukholm G, Barach P, Tjomsland O. A systematic review of hospital accreditation: the challenges of measuring complex intervention effects. BMC Health Serv Res 2015 Jul 23;15:280. [doi: 10.1186/s12913-015-0933-x] [Medline: 26202068]
24. Desveaux L, Mitchell JI, Shaw J, Ivers NM. Understanding the impact of accreditation on quality in healthcare: a grounded theory approach. Int J Qual Health Care 2017 Nov 1;29(7):941-947. [doi: 10.1093/intqhc/mzx136] [Medline: 29045664]
25. Mitchell JI, Graham ID, Nicklin W. The unrecognized power of health services accreditation: more than external evaluation. Int J Qual Health Care 2020 Sep 23;32(7):445-455. [doi: 10.1093/intqhc/mzaa063] [Medline: 32514539]
26. Jha AK. Accreditation, quality, and making hospital care better. JAMA 2018 Dec 18;320(23):2410-2411. [doi: 10.1001/jama.2018.18810] [Medline: 30561469]
27. Devkaran S, O'Farrell PN, Ellahham S, Arcangel R. Impact of repeated hospital accreditation surveys on quality and reliability, an 8-year interrupted time series analysis. BMJ Open 2019 Feb 15;9(2):e024514. [doi: 10.1136/bmjopen-2018-024514] [Medline: 30772852]
28. Hinchcliff R, Glennie M, Travaglia J, Carter D, Billington L, Debono D. Short notice and unannounced survey methods: literature review. : Australian Commission on Safety and Quality in Health Care; 2017 URL: https://www.safetyandquality.gov.au/sites/default/files/migrated/Short-Notice-and-Unannounced-Survey-Methods-Literature-review.pdf [accessed 2026-04-09]
29. Barnett ML, Olenski AR, Jena AB. Patient mortality during unannounced accreditation surveys at US hospitals. JAMA Intern Med 2017 May 1;177(5):693-700. [doi: 10.1001/jamainternmed.2016.9685] [Medline: 28319229]
30. Hall A. State of the surveys: 18 months of joint commission unannounced visits. Biomed Instrum Technol 2007 Jul;41(4):309-310. [doi: 10.2345/0899-8205(2007)41[309:SOTSMO]2.0.CO;2]
31. Bergholt MD, Falstie-Jensen AM, Brink Valentin J, et al. Patients experience more support, information and involvement after first-time hospital accreditation: a before and after study in the Faroe Islands. Int J Qual Health Care 2021 Nov 12;33(4):mzab149. [doi: 10.1093/intqhc/mzab149] [Medline: 34698825]
32. Scanlan R, Flenady T, Judd J. Impact of short-notice accreditation assessments on hospitals' patient safety and quality culture-a scoping review. J Adv Nurs 2024 Oct;80(10):3965-3976. [doi: 10.1111/jan.16169] [Medline: 38553879]
33. Schein EH. Organizational Culture and Leadership: John Wiley and Sons; 2017.
34. Allen S, Chiarella M, Homer CSE. Lessons learned from measuring safety culture: an Australian case study. Midwifery 2010 Oct;26(5):497-503. [doi: 10.1016/j.midw.2010.07.002] [Medline: 20692077]
35. Alsalem G, Bowie P, Morrison J. Assessing safety climate in acute hospital settings: a systematic review of the adequacy of the psychometric properties of survey measurement tools. BMC Health Serv Res 2018 May 10;18(1):353. [doi: 10.1186/s12913-018-3167-x] [Medline: 29747612]
36. Brown DS, Wolosin R. Safety culture relationships with hospital nursing sensitive metrics. J Healthc Qual 2013;35(4):61-74. [doi: 10.1111/jhq.12016] [Medline: 23777363]
37. Kristensen S, Mainz J, Bartels P, Kristensen S, Mainz J, Bartels P. Selection of indicators for continuous monitoring of patient safety: recommendations of the project "safety improvement for patients in Europe". Int J Qual Health Care 2009 Jun;21(3):169-175. [doi: 10.1093/intqhc/mzp015] [Medline: 19359329]
38. Chatburn E, Macrae C, Carthey J, Vincent C. Measurement and monitoring of safety: impact and challenges of putting a conceptual framework into practice. BMJ Qual Saf 2018 Oct;27(10):818-826. [doi: 10.1136/bmjqs-2017-007175] [Medline: 29511091]
39. Ginsburg L, Gilin Oore D. Patient safety climate strength: a concept that requires more attention. BMJ Qual Saf 2016 Sep;25(9):680-687. [doi: 10.1136/bmjqs-2015-004150] [Medline: 26453636]
40. Kalteh HO, Mortazavi SB, Mohammadi E, Salesi M. The relationship between safety culture and safety climate and safety performance: a systematic review. Int J Occup Saf Ergon 2021 Mar;27(1):206-216. [doi: 10.1080/10803548.2018.1556976] [Medline: 30526393]
41. Steyrer J, Schiffinger M, Huber C, Valentin A, Strunk G. Attitude is everything? The impact of workload, safety climate, and safety tools on medical errors: a study of intensive care units. Health Care Manage Rev 2013;38(4):306-316. [doi: 10.1097/HMR.0b013e318272935a] [Medline: 23085639]
42. DiCuccio MH. The relationship between patient safety culture and patient outcomes: a systematic review. J Patient Saf 2015 Sep;11(3):135-142. [doi: 10.1097/PTS.0000000000000058] [Medline: 24583952]
43. Sexton JB, Berenholtz SM, Goeschel CA, et al. Assessing and improving safety climate in a large cohort of intensive care units. Crit Care Med 2011 May;39(5):934-939. [doi: 10.1097/CCM.0b013e318206d26c] [Medline: 21297460]
44. Caris MG, Kamphuis PGA, Dekker M, de Bruijne MC, van Agtmael MA, Vandenbroucke-Grauls C. Patient safety culture and the ability to improve: a proof of concept study on hand hygiene. Infect Control Hosp Epidemiol 2017 Nov;38(11):1277-1283. [doi: 10.1017/ice.2017.209] [Medline: 28967350]
45. Carrothers KM, Barr J, Spurlock B, Ridgely MS, Damberg CL, Ely EW. Contextual issues influencing implementation and outcomes associated with an integrated approach to managing pain, agitation, and delirium in adult ICUs. Crit Care Med 2013 Sep;41(9 Suppl 1):S128-S135. [doi: 10.1097/CCM.0b013e3182a2c2b1] [Medline: 23989090]

XSL•FO
RenderX

46. Jones L, Pomeroy L, Robert G, Burnett S, Anderson JE, Fulop NJ. How do hospital boards govern for quality improvement? A mixed methods study of 15 organisations in England. BMJ Qual Saf 2017 Dec;26(12):978-986. [doi: 10.1136/bmjqs-2016-006433] [Medline: 28689191]
47. Simpson D, Hamilton S, McSherry R, McIntosh R. Measuring and assessing healthcare organisational culture in the England's National Health Service: a snapshot of current tools and tool use. Healthcare (Basel) 2019 Nov 1;7(4):127. [doi: 10.3390/healthcare7040127] [Medline: 31683839]
48. Mannion R, Davies H. Understanding organisational culture for healthcare quality improvement. BMJ 2018 Nov 28;363:k4907. [doi: 10.1136/bmj.k4907] [Medline: 30487286]
49. Feng X, Bobay K, Weiss M. Patient safety culture in nursing: a dimensional concept analysis. J Adv Nurs 2008 Aug;63(3):310-319. [doi: 10.1111/j.1365-2648.2008.04728.x] [Medline: 18702778]
50. Cacciattolo K. Understanding organisational cultures. Eur Sci J 2014;2(1):1-7 [FREE Full text]
51. Hogan SJ, Coote LV. Organizational culture, innovation, and performance: a test of Schein's model. J Bus Res 2014 Aug;67(8):1609-1621. [doi: 10.1016/j.jbusres.2013.09.007]
52. Khaddour WN. Critical realism and organizational culture—framing Schein's model of studying organizational culture by using critical realism philosophy. OJPP 2025;15(1):77-97. [doi: 10.4236/ojpp.2025.151006]
53. Creswell JW. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches: Sage Publications; 2009.
54. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
55. Hogden A, Ellis LA, Churruca K, Bierbaum M. Safety Culture Assessment in Health Care: A Review of the Literature on Safety Culture Assessment Modes: Australian Commission on Safety and Quality in Health Care; 2017. URL: https://www.safetyandquality.gov.au/sites/default/files/migrated/Safety-Culture-Assessment-in-Health-Care-A-review-of-the-literature-on-safety-culture-assessment-modes.pdf [accessed 2026-04-14]
56. Sutanapong C, Louangrath PI. Descriptive and inferential statistics. International Journal of Research & Methodology in Social Science 2015;1(1):14 [FREE Full text]
57. Collopy BT. Clinical indicators in accreditation: an effective stimulus to improve patient care. Int J Qual Health Care 2000 Jun;12(3):211-216. [doi: 10.1093/intqhc/12.3.211] [Medline: 10894192]
58. Mumford V, Reeve R, Greenfield D, Forde K, Westbrook J, Braithwaite J. Is accreditation linked to hospital infection rates? A 4-year, data linkage study of Staphylococcus aureus rates and accreditation scores in 77 Australian acute hospitals. Int J Qual Health Care 2015 Dec;27(6):479-485. [doi: 10.1093/intqhc/mzv078] [Medline: 26424700]
59. National Safety and Quality Standards 2nd Edition: User Guide for Governing Bodies: Australian Commision on Safety and Quality in Health Care; 2019. URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/nsqhs-standards-user-guide-governing-bodies [accessed 2026-04-14]
60. A-HSOPS 20 survey form. Australian Commission on Safety and Quality in Health Care. URL: https://www.safetyandquality.gov.au/publications-and-resources/resource-library/hsops-20-survey-form [accessed 2026-04-14]
61. O'Leary Z, Tennent E. The Essential Guide to Doing Your Research Project, 3rd edition: Sage Publications Ltd; 2017.
62. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
63. Greenfield D, Moldovan M, Westbrook M, et al. An empirical test of short notice surveys in two accreditation programmes. Int J Qual Health Care 2012 Feb;24(1):65-71. [doi: 10.1093/intqhc/mzr074] [Medline: 22140191]

Abbreviations

A-HSOPS 2.0: Australian Hospital Survey on Patient Safety Culture (version 2.0)
AHSSQAS: Australian Health Service Safety and Quality Accreditation Scheme
GRAMMS: Good Reporting of A Mixed Methods Study
HSO: health service organization

Edited by J Sarvestan; submitted 05.May.2025; peer-reviewed by D Greenfield, MA Zoubi; revised version received 24.Jan.2026; accepted 29.Jan.2026; published 22.Apr.2026.

Please cite as:
Scanlan R, Flenady T, Byrne AL, Judd J
Evaluating the Impact of Short-Notice Accreditation Assessment on Hospitals' Patient Safety and Quality Culture: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e76945
URL: https://www.researchprotocols.org/2026/1/e76945
doi:10.2196/76945

XSL•FO
RenderX

© Robyn Scanlan, Tracy Flenady, Amy-Louise Byrne, Jenni Judd. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Development of a Modified Intimate Wash From Yoniprakshalan Kwatha (YOSHA) With Comparative Antimicrobial and Quality Assessment: Protocol for an Experimental Study

Prachi Vandan Jadhav, UG, MD; Bharat Jagdishji Rathi, UG, PG, PhD

Department of Rasashashtra and Bhaishajya Kalpana, Mahatma Gandhi Ayurved College and Research Centre, Datta Meghe Institute of Higher Education and Research Centre, Paloti Road, Sawangi Meghe, Wardha, India

Corresponding Author:
Prachi Vandan Jadhav, UG, MD
Department of Rasashashtra and Bhaishajya Kalpana, Mahatma Gandhi Ayurved College and Research Centre, Datta Meghe Institute of Higher Education and Research Centre, Paloti Road, Sawangi Meghe, Wardha, India

Abstract

Background: Vaginal itching is a frequent gynecological issue impacting women of various age groups. Although commonly linked to nonserious conditions, it may also signal infections, sexually transmitted diseases, prolonged use of certain medications, such as oral contraceptives, hormonal fluctuations, or suboptimal hygiene. Standard pharmaceutical treatments, despite being widely used, often produce undesirable effects, such as irritation of the vaginal lining, alterations in vaginal pH, and disruption of healthy microbial flora. These challenges highlight the need for safer, natural alternatives rooted in traditional systems of medicine, such as Ayurveda.

Objective: This study aims to prepare *yoniprakshalan kwatha* (intimate wash decoction) as per the classical reference of *Yogratnakar* and to develop its modified dosage form, YOSHA. Additionally, this study aims to assess the quality control parameters of both yoniprakshalan kwatha and YOSHA to ensure their safety and efficacy. Furthermore, the antimicrobial effects of both formulations will be investigated, and a comparative analysis will be conducted to evaluate and contrast the antimicrobial activity of YOSHA and yoniprakshalan kwatha.

Methods: Medicinal herbs were sourced from the institute's *rasashala* and authenticated by taxonomists. The traditional decoction and YOSHA will be prepared using Ayurvedic and modern pharmaceutical methods. Quality tests will include physicochemical and microbial analyses. The disc diffusion method will be used to evaluate antimicrobial activity against *Candida albicans*, *Escherichia coli*, and *Staphylococcus aureus*. Standard antibiotics and sterile discs will serve as controls. Tests will be repeated over 3 days. Statistical analysis will be performed using one-way ANOVA in SPSS (version 17), with $P<.05$ considered significant.

Results: This study commenced in July 2025, with data collection scheduled from August 2025 to September 2025. By December 2025, it is expected that laboratory testing and antimicrobial assessments will have been completed, along with the finalization of quality control evaluations. Preliminary findings are anticipated to show that YOSHA will exhibit strong antimicrobial activity, comparable to or exceeding that of the original yoniprakshalan kwatha. Significant inhibition zones are likely to be observed against *C albicans* and *E coli*. The modified formulation is projected to meet pharmaceutical standards for consistency, safety, and shelf stability. Data analysis will be completed, and the full results are expected to be published in December 2025.

Conclusions: The study will show that YOSHA, a modern herbal product based on the traditional Ayurvedic yoniprakshalan kwatha, has strong in vitro antimicrobial effects against common vaginal pathogens and meets pharmaceutical quality and safety standards. Unlike raw decoctions, it offers a stable, user-friendly formulation that blends traditional knowledge with scientific validation. Strengths include classical preparation methods, solid quality control, and reliable results. Future work should involve clinical trials, broader testing, and user-focused evaluations. Findings will be disseminated through journals, conferences, and practitioner outreach to promote YOSHA as a natural option for women's hygiene.

(JMIR Res Protoc 2026;15:e73239) doi:10.2196/73239

KEYWORDS

yoniprakshalan kwatha; yoniwash; YOSHA; antimicrobial; Ayurveda

XSL•FO
RenderX

Introduction

Background

Vaginal itching is a common and typically harmless condition that most women experience at some point in their lives. It can be triggered by a variety of factors, including infections, hormonal changes, sexually transmitted diseases, prolonged use of medications such as birth control pills, or poor vaginal hygiene. Additionally, certain skin conditions, such as psoriasis, eczema, or a vaginal yeast infection, can contribute to the irritation. In rare cases, it may be a symptom of a more serious underlying condition, such as vulvar cancer. If left untreated, vaginal itching can potentially lead to serious complications or, in extreme cases, become life-threatening [1].

Treating these conditions with oral medications is as important as topical therapies [2]. Female genital disorders are treated using antibiotics, antifungals, antiprotozoals, and other medications administered both orally and topically. The *sthanik chikitsa* (local treatment) recommended by the ancient Acharyas has shown excellent and satisfactory results in the management of *stree roga* (gynecological disorders). These localized therapies are particularly effective in treating *tryavarta yoni* disease [3]. According to the World Health Organization, an estimated 333 million new cases of treatable vulvovaginal infections occur globally each year. A study conducted in India reported a 37% prevalence of reproductive tract infections based on symptoms and a 36.7% prevalence based on laboratory findings. Among these cases, 31% were due to candidiasis, 3% to gonorrhea, 2% to trichomoniasis, and 45% to bacterial vaginosis [4].

Using intimate wash is important for maintaining vaginal health. It helps preserve the natural pH balance of the vagina, which is essential for preventing infections and discomfort. Practicing proper intimate hygiene can significantly reduce the risk of bacterial and fungal infections, including yeast infections and bacterial vaginosis [5]. Using a gentle, specially formulated wash for the intimate area can help eliminate odor-causing bacteria and keep the area feeling clean and refreshed [6]. Additionally, regular use of an intimate wash can enhance overall comfort and confidence, helping women feel clean and hygienic throughout the day. In essence, a female intimate wash plays a vital role in supporting women's reproductive health and overall well-being [7].

Yoniprakshalana kwatha (intimate wash decoction) [8] is the formulation described in the *Yogratnakar* for the treatment of *yonikandu* and includes *triphala*, *gudhuchi*, and *dantimul*. In this study, this traditional *kwatha* will be adapted into a ready-to-use surfactant-based vaginal wash, named YOSHA. The term "YOSHA" signifies "woman" (*stree*) [9]. The name "YOSHA" is derived from the words *yoni* ("YO") and *wash* ("SH"), symbolizing a product designed specifically for cleansing the *yoni* (vagina) and intended for women. This modified version of a medicated intimate wash, enriched with the benefits of traditional herbs, offers a more convenient and travel-friendly alternative to the freshly prepared *kwatha* (decoction). Compared with the original decoction, the improved formulation ensures easier use, thoroughly cleanses the vaginal area, and provides a refreshing, hygienic effect. This intimate wash is based on the classical formulation for *yoniprakshalan kwatha* described in ancient Ayurvedic texts, particularly the *Yogratnakar*. This study aims to formulate a modification of yoniprakshalan kwatha as an intimate wash (YOSHA) and to conduct a comparative quality control assessment and in vitro evaluation of antimicrobial activity.

Need for the Study

Maintaining vaginal hygiene is essential, especially in today's fast-paced world, where working women and school-going girls are frequently exposed to public restrooms. While there are several intimate wash products available on the market, many contain harsh chemicals that can lead to side effects such as dryness and itching.

To address this, we have developed a simplified and user-friendly version of *yonidhavan kwatha*. Unlike the traditional formulation, which has a short shelf life of about 24 hours, YOSHA is designed in a surfactant-based form that is both effective and convenient. It helps maintain healthy vaginal pH and ensures thorough daily cleansing.

YOSHA is travel friendly, easy to handle, and simple to use, eliminating the need for complex preparation. By offering this product, we support women in maintaining their intimate health and well-being with comfort and confidence.

This study aims to formulate a modification of yoniprakshalan kwatha as an intimate wash (YOSHA) and to conduct a comparative quality control assessment and in vitro evaluation of antimicrobial activity.

Objectives

The objective of this study is to prepare yoniprakshalan kwatha as per the classical reference of *Yogratnakar* and to develop its modified dosage form, YOSHA. Additionally, the study aims to assess the quality control parameters of both yoniprakshalan kwatha and YOSHA to ensure their safety and efficacy. Furthermore, the antimicrobial effects of both formulations will be investigated, with a comparative analysis conducted to evaluate and contrast the antimicrobial activity of YOSHA and yoniprakshalan kwatha.

Methods

Drug Collection and Authentication

An in vitro experimental study will be conducted at the Central Microbiology Laboratory of Jawaharlal Nehru Medical College, a constituent of the Datta Meghe Institute of Higher Education and Research, located in Sawangi (Meghe), Wardha, Maharashtra. The study will be conducted between July 2025 and December 2025.

The research drugs will be procured from Vedpoorna Ayurveda Rasashala, Mahatma Gandhi Ayurved College Hospital and Research Centre (MGACHRC), Salod (H), Wardha. The Dravyaguna Department at MGACHRC will be responsible for the authentication and verification of the raw drugs. These raw materials will be standardized according to the Ayurvedic Pharmacopoeia of India guidelines. Necessary studies will be

XSL•FO
RenderX

conducted at certified or recognized institutes, organizations, or laboratories of national repute, as endorsed or recommended by the Datta Meghe Institute of Higher Education and Research. An experimental study will be conducted at the Department of Rasashastra and Bhaishajya Kalpana, MGACHRC, from July 2025 to December 2025.

Data Handling

The data management plan will follow strict protocols for accurate entry, coding, security, and storage to ensure data integrity. Standardized methods, including double entry for key data and range checks, will minimize errors. Data will be coded according to established guidelines to support analysis. Security measures such as restricted access, encryption, and secure storage (both electronic and physical) will be implemented. Regular backups will prevent data loss, and all procedures will be detailed in the study protocol for transparency.

Drug Review

Yoniprakshalan kwatha (intimate wash decoction) is an important herbo-mineral formulation described in a book by Vaidya Lakshmpati Sastri, published by Chaukhamba Prakashan, Varanasi, with a Hindi commentary by Bhisagratna and Brahmasankar Sastri. This formulation holds significant therapeutic value and contains ingredients such as *guduchi*, *amalaki*, *haritaki*, *bibhitaki*, and (Table 1) *dantimul*. It is primarily indicated for the treatment of *yoni kandu* (vaginal itching).

Table . Ingredients of *yoniprakshalan kwatha* along with their properties.

Ingredient	Botanical name	Family	Property and description
Amalaki [10]	*Emblica officinalis* Gaertn	Phyllanthaceae	• Rasa: Pancharasa (lavanarasa rahita), amlaka pradhana • Virya: sheeta • Vipak: madhur • Part used: fruit • Quantity: 1 part
Haritaki [11]	*Terminalia chebula* Retz	Combretaceae	• Rasa: Pancharasa (lavan varjita), kashaya pradhana • Virya: ushna • Vipak: madhur • Part used: fruit • Quantity: 1 part
Bibhitak [11]	*Terminalia bellirica* Roxb	Combretaceae	• Rasa: kashaya • Virya: ushna • Vipak: madhur • Part used: fruit • Quantity: 1 part
Guduchi [11]	*Tinospora cordifolia* Miers	Menispermaceae	• Rasa: tikta, kashaya • Virya: ushna • Vipak: madhur • Part used: stem • Quantity: 1 part
Danti [11]	*Baliospermum montanum* Muell-Arg	Euphorbiaceae	• Rasa: katu • Virya: ushna • Vipak: root • Quantity: 1 part

Detail of Drug Preparation

In the first step, yoniprakshalan kwatha will be prepared according to the textual reference and in the second step, its modified dosage form, YOSHA, will be prepared as mentioned in the following sections (Table 2).

XSL•FO
RenderX

Table . Ingredients and compositions of yoniprakshalan kwatha (intimate wash decoction).

Ingredients	Botanical name	Part used	Proportion
Amalaki	*Emblica officinalis* Gaertn.	Fruit	1 part
Haritaki	*Terminalia chebula* Retz	Fruit	1 part
Bibhitaki	*Terminalia bellirica* Roxb	Fruit	1 part
Guduchi	*Tinospora cordifolia* Miers	Stem	1 part
Dantimoola	*Baliospermum montanum* Muell-Arg.	Root	1 part
Water	—[a]	—	16 parts of total drugs

[a]Not available.

Preparation of Yoniprakshalan Kwatha

Yoniprakshalan kwatha will be prepared according to the classical method described in the *Yogratnakar* [12], using authenticated raw materials and standard *kwatha* preparation procedures (Textbox 1).

Textbox 1. Preparation step of yoniprakshalan kwatha.

Step 1: inspection of raw drugs

- Remove deformed pieces or contaminants

Step 2: powdering and mixing

- Take all raw drugs in coarse powder form
- Use a clean stainless-steel vessel

Step 3: soaking

- Add 16 parts of water to the mixture
- Leave undisturbed for 12 hours

Step 4: boiling

- Boil the mixture
- Reduce it to 1/8th of the original volume

Step 5: filtration

- Filter the decoction using a clean cloth

Step 6: final product

- The filtrate is yoniprakshalana kwatha

Preparation of YOSHA

The modified dosage form, YOSHA, will be prepared from yoniprakshalan kwatha using a surfactant-based formulation approach to obtain a stable, ready-to-use intimate wash (Table 3 and Textbox 2).

XSL•FO
RenderX

Table . Composition of YOSHA.

Ingredients	Percentage	Weight
Yoniprakshalan kwatha (intimate wash decoction)	w/w 40	w/w 40 mL
Methylparaben (preservative) [13]	w/w 4	w/w 4 gm
Sodium metabisulfite (antioxidant)	w/w 4	w/w 4 gm
Hydroxypropyl methylcellulose [14]	w/w 8	w/w 8 gm
Tartaric acid [15]	w/w 4	w/w 4 gm
Distilled water	w/w 40	w/w 40 mL
Total	100	100 ml w/w YOSHA

Textbox 2. Preparation steps of YOSHA.

Step 1: base decoction preparation

- Take yoniprakshalana kwatha in a clean vessel

Step 2: pH balancing

- Add tartaric acid to adjust pH

Step 3: preservation and antioxidation

- Add methylparaben (preservative)
- Add sodium metabisulfite (antioxidant)

Step 4: viscosity adjustment

- Dissolve hydroxypropyl methylcellulose in distilled water
- Add the hydroxypropyl methylcellulose gel base to the decoction

Step 5: final mixing and storage

- Stir thoroughly to ensure uniformity
- Store in an airtight glass container

Antimicrobial Study

This study will evaluate the antimicrobial activity of yoniprakshalan kwatha and YOSHA using the agar well diffusion method [16].

Principle of Antimicrobial Study

A total of 9 agar plates per treatment will be used, and the antimicrobial test will be repeated 3 times. The antimicrobial constituents in yoniprakshalan kwatha and YOSHA will be allowed to diffuse into the agar medium and interact with a plate freshly inoculated with the test organism. The resulting inhibition zone will form a uniform circle, as the growth will spread continuously. The diameter of the inhibition zone can be measured in millimeters. The agar well will be 10 mm in diameter and will hold around 0.5 mL—this volume fills the well evenly, ensuring full contact with the agar surface and avoiding overflows that could interfere with diffusion. Yoniprakshalan kwatha and YOSHA will be introduced into the wells, and ofloxacin 10 µg/mL will be used as the standard for the study. The petri dishes will then be incubated at 37 °C for 24 hours, and the zones of inhibition will be measured [17].

Selection of Test Microorganisms

The bacterial strains for the experiment will be obtained from the microbiology department of a reputable institute [18]. The antimicrobial activity of the formulations will be assessed against selected gram-positive bacteria, gram-negative bacteria, and fungal strains commonly associated with vaginal infections. The gram-positive bacteria tested will include *Streptococcus aureus* (National Commission for Indian Systems of Medicine [NCISM] 2079) [19] and *Streptococcus epidermidis* (NCISM 2493) [20]. The gram-negative bacteria evaluated will include *Escherichia coli* (NCISM 2931) [21] and *Proteus vulgaris* (NCISM 2027) [22]. The fungal strain tested will be *Candida albicans* [23].

Positive Control Group Drugs

Standard antimicrobial agents will be used as positive controls to validate the antimicrobial assay. For gram-positive bacteria, azithromycin (50 µg/6 µL) will be used. For gram-negative bacteria, ciprofloxacin (30 µg/6 µL) will be used. For *C albicans*, clotrimazole (50 µg/6 µL) will be used.

XSL•FO
RenderX

Analytical Study of Drugs

Analytical Parameters of Yoniprakshalan Kwatha

The analytical evaluation of yoniprakshalan kwatha will include the assessment of organoleptic and physicochemical parameters. Organoleptic characteristics, including touch, appearance, taste, and odor, will be examined [24]. Physicochemical parameters will include alcohol content, pH value, specific gravity at 25 °C, total solids, water-soluble extractive values, alcohol-soluble extractive values, and high-performance thin-layer chromatography [25].

Analytic Parameter for Yoni Wash (YOSHA)

YOSHA will be evaluated using organoleptic and physicochemical parameters. Organoleptic characteristics, including touch, appearance, and odor, will be assessed [26]. Physicochemical evaluation will include pH, water-soluble extractive values or alcohol-soluble extractive values, and tests for specific pathogens [27].

Data Analysis

The data will be analyzed to determine the effects of the test compounds on the assessed parameters. This will typically involve statistical analysis to compare the treated groups to the control group.

Summary

Pharmaceutical and analytical studies of yoniprakshalan kwatha and its modified form as intimate wash will provide the standard parameters.

Study Guidelines

This manuscript will adhere to the Checklist for Reporting In-vitro Studies (CRIS) guidelines [28] to ensure methodological transparency and comprehensive reporting of in vitro research. The CRIS guidelines were followed during the preparation and reporting of this study.

Statistical Analysis

Data analysis will be conducted using SPSS (version 17; IBM Corp). Three groups—YOSHA, yoniprakshalan kwatha, and a standard antimicrobial control—will be compared, with a minimum of 9 replicates per group. Descriptive statistics will include means (SD) for continuous variables and percentages for categorical data. Normality and homogeneity of variance will be assessed using the Shapiro-Wilk test and the Levene test, respectively. One-way ANOVA will be applied to compare antimicrobial activity (zone of inhibition), followed by the Tukey honestly significant difference post hoc test if significant differences are found (P<.05). If ANOVA assumptions are violated, nonparametric alternatives such as the Kruskal-Wallis test will be used. Regression analysis will be performed if dose-response data are available. This approach ensures statistical validity and reliable comparisons between formulations.

Ethical Considerations

Ethical approval for the research has been granted by the Institutional Ethics Committee (MGACHRC/IEC/Jun-2024/854) of the Datta Meghe Institute of Higher Education & Research, Salod (H), Wardha-442004, Maharashtra.

Results

This study commenced in July 2025, with data collection scheduled from August 2025 to September 2025. By December 2025, it is expected that laboratory testing and antimicrobial assessments will have been completed, along with the finalization of quality control evaluations. Preliminary findings are anticipated to show that YOSHA will exhibit strong antimicrobial activity, comparable to or exceeding that of the original yoniprakshalan kwatha. Significant inhibition zones are likely to be observed against *C albicans* and *E coli*. The modified formulation is projected to meet pharmaceutical standards for consistency, safety, and shelf stability. Data analysis will be completed, and full results are expected to be published in December 2025.

Discussion

Anticipated Findings

This study is expected to demonstrate that YOSHA, a modified formulation of yoniprakshalan kwatha, will demonstrate strong in vitro antimicrobial activity, particularly against *C albicans* and *E coli*. Designed as a modern, surfactant-based intimate wash, YOSHA is expected to meet the pharmaceutical standards of stability, safety, and usability while preserving Ayurvedic therapeutic principles. The aim is to convert a traditional decoction into a stable, ready-to-use product with improved shelf life and consistent quality. YOSHA may offer a safer, more convenient alternative to commercial feminine hygiene products, many of which can disrupt vaginal pH or flora [29].

While Ayurvedic treatments, such as yoniprakshalan kwatha, are traditionally used for vaginal conditions such as *yoni kandu* [30], their practical application is limited by preparation complexity and short shelf life [31]. YOSHA addresses these gaps through standardized pharmaceutical development. Although the individual herbs used (*triphala*, *guduchi*, and *dantimul*) have known antimicrobial properties [32], their combined and stabilized use in topical form remains underexplored. The strengths of this study lie in its grounding in classical texts, the use of validated preparation methods, and laboratory testing. However, as an in vitro study, it lacks clinical data. In addition, user acceptability, long-term stability, and postmarket outcomes also need further evaluation. Future research directions include clinical trials, enhanced formulations, user-focused adaptations, and sustainable manufacturing practices. These steps will help to validate the efficacy of YOSHA and support its integration into modern female hygiene care.

Dissemination Plan

This protocol will be published in the form of a thesis to share our research on Ayurvedic intimate wash. The thesis will provide a thorough overview of the study's design, methodology, data gathering methods, data analysis plan, and ethical considerations. We hope that by making this protocol

XSL•FO
RenderX

public, we will contribute to expanding knowledge in the field and promoting future research efforts.

Conclusions

The study will show that YOSHA, a modern herbal product based on the traditional Ayurvedic yoniprakshalan kwatha, has strong in vitro antimicrobial effects against common vaginal pathogens and meets the pharmaceutical quality and safety standards. Unlike raw decoctions, it offers a stable, user-friendly formulation that blends traditional knowledge with scientific validation. Strengths include classical preparation methods, solid quality control, and reliable results. Future work should involve clinical trials, broader testing, and user-focused evaluations. Findings will be shared through journals, conferences, and practitioner outreach to promote YOSHA as a natural option for women's hygiene.

Acknowledgments

The authors express their gratitude to the researchers whose works are cited in the text for their valuable contributions and to the management for providing the necessary facilities. During the preparation of the manuscript, the authors used Grammarly for grammar and language checks.

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: PVJ
Formal analysis: BJR
Methodology: PVJ
Validation: BJR
Writing – original draft: PVJ
Writing – review and editing: BJR

Conflicts of Interest

None declared.

References

1. Patil VS, Pargunde VS. Literature review of contents of Yonikanduhara Lepa (Rasatarngini). J Ayurveda Integr Med Sci 2020;5(2):241-245. [doi: 10.21760/jaims.v5i02.894]
2. Palmeira-de-Oliveira R, Palmeira-de-Oliveira A, Martinez-de-Oliveira J. New strategies for local treatment of vaginal infections. Adv Drug Deliv Rev 2015 Sep 15;92:105-122. [doi: 10.1016/j.addr.2015.06.008] [Medline: 26144995]
3. Patil S. Topical dosage forms (lepa kalpana) of Ayurveda: an unexplored treasure. Int J Green Pharm 2015;9:S1-S7 [FREE Full text] [doi: 10.22377/ijgp.v9i4.579]
4. Puri KJ, Madan A, Bajaj K. Incidence of various causes of vaginal discharge among sexually active females in age group 20-40 years. Indian J Dermatol Venereol Leprol 2003;69(2):122-125 [FREE Full text] [Medline: 17642853]
5. Rajamanoharan S, Low N, Jones SB, Pozniak AL. Bacterial vaginosis, ethnicity, and the use of genital cleaning agents: a case control study. Sex Transm Dis 1999 Aug;26(7):404-409. [doi: 10.1097/00007435-199908000-00008] [Medline: 10458635]
6. Holdcroft AM, Ireland DJ, Payne MS. The vaginal microbiome in health and disease-what role do common intimate hygiene practices play? Microorganisms 2023 Jan 23;11(2):298. [doi: 10.3390/microorganisms11020298] [Medline: 36838262]
7. Bruning E, Chen Y, McCue KA, Rubino JR, Wilkinson JE, Brown ADG. A 28 day clinical assessment of a lactic acid-containing antimicrobial intimate gel wash formulation on skin tolerance and impact on the vulvar microbiome. Antibiotics (Basel) 2020 Feb 1;9(2):55. [doi: 10.3390/antibiotics9020055] [Medline: 32024047]
8. Shastri V. Yogratnakar, 1st edition: Chaukhamba Prakashan; 2008:405-406.
9. Anwar S, Hameed A. Chronicles of Social Status of Females in Ancient Society of the Present-Day Pakistan: Methodology (Gott); 2003.

XSL•FO
RenderX

10. Sharma PV. Dravyaguna Vigyan: Chaukhamba Bharti Academy; 2017, Vol. 2.
11. Vaidya BG. Nighantu Adarsha: Chaukhamba Bharti Academy; 2019.
12. A brief study on Ayurvedic Kwath: A review. IAR J Med Surg Res 2024 Sep 21;5(4). [doi: 10.70818/iarjmsr.2024.V05i04.02]
13. Scientific committee on consumer safety (SCCS). European Commission. 2019. URL: https://health.ec.europa.eu/scientific-committees/scientific-committee-consumer-safety-sccs_en [accessed 2025-05-25]
14. Burdock GA. Safety assessment of hydroxypropyl methylcellulose as a food ingredient. Food Chem Toxicol 2007 Dec;45(12):2341-2351. [doi: 10.1016/j.fct.2007.07.011] [Medline: 17723258]
15. Lang WR. Vaginal acidity and pH: a review. Obstet Gynecol Surv 1955 Aug;10(4):546-560. [doi: 10.1097/00006254-195508000-00009] [Medline: 13244967]
16. Ananthanarayan R. Ananthanarayan and Paniker's Textbook of Microbiology: Orient Blackswan; 2006.
17. Balouiri M, Sadiki M, Ibnsouda SK. Methods for in vitro evaluating antimicrobial activity: a review. J Pharm Anal 2016 Apr;6(2):71-79. [doi: 10.1016/j.jpha.2015.11.005] [Medline: 29403965]
18. Acharya S, Dahal P. Bacterial and fungal infections in the vaginal microflora. J Microbiol Exp 2020;8(4):187-192.
19. Deng L, Schilcher K, Burcham LR, et al. Identification of key determinants of Staphylococcus aureus vaginal colonization. MBio 2019 Dec 24;10(6):e02321-19. [doi: 10.1128/mBio.02321-19] [Medline: 31874913]
20. Dover SE, Aroutcheva AA, Faro S, Chikindas ML. Natural antimicrobials and their role in vaginal health: a short review. Int J Probiotics Prebiotics 2008;3(4):219-230. [Medline: 20657710]
21. McDonald HM, O'Loughlin JA, Jolley P, Vigneswaran R, McDonald PJ. Vaginal infection and preterm labour. Br J Obstet Gynaecol 1991 May;98(5):427-435. [doi: 10.1111/j.1471-0528.1991.tb10335.x] [Medline: 2059587]
22. Bilal S, Anam S, Mahmood T, et al. Antimicrobial profiling and molecular characterization of antibiotic resistant genes of Proteus vulgaris isolated from tertiary care hospital, Islamabad, Pakistan. Pak J Pharm Sci 2019 Nov;32(6(Supplementary):2887-2891. [Medline: 32024629]
23. Sobel JD. Recurrent vulvovaginal candidiasis. Am J Obstet Gynecol 2016 Jan;214(1):15-21. [doi: 10.1016/j.ajog.2015.06.067] [Medline: 26164695]
24. A brief study on ayurvedic kwath: a review. IAR J Med Surg Res 2024;5(4). [doi: 10.70818/iarjmsr.2024.V05i04.02]
25. Bhatt N, Deshpande M, Chaskar S. A critical review of standardization of ayurvedic dosage form kwatha - part - II: approaches and outcome. Int J Ayurvedic Med 2020;11(3):378-386. [doi: 10.47552/ijam.v11i3.1552]
26. Baranova II, Bezpala YO, Shpychak OS, Havrysh NB. Optimisation of physical and chemical characteristics of a modern intimate hygiene agent. J Pharm Pharmacol 2017;69(3):387-394 [FREE Full text]
27. Gedik G. Formulation and evaluation of intimate wash containing manuka oil and salix bark extract. jrp 2022;26(5)(26(5):1252-1260. [doi: 10.29228/jrp.217]
28. Faggion Jr CM Jr. Guidelines for reporting pre-clinical in vitro studies on dental materials. J Evid Based Dent Pract 2012 Dec;12(4):182-189. [doi: 10.1016/j.jebdp.2012.10.001]
29. Fashemi B, Delaney ML, Onderdonk AB, Fichorova RN. Effects of feminine hygiene products on the vaginal mucosal biome. Microb Ecol Health Dis 2013;24:19703. [doi: 10.3402/mehd.v24i0.19703] [Medline: 24009546]
30. Thakur C, Jain A. A clinical study to evaluate the effect of triphala kwatha with gomutra yoni prakshalana followed by dhatakyadi taila yoni pichu in kaphaja yonivyapad wsr to vulvovaginal candidiasis. World J Pharm Res 2025;14(7) [FREE Full text]
31. Sharma U, Verma A, Yadavendra Yadav DK. Review of kwatha kalpana and its modified forms. Int J Ayurvedic Herb Med 2020;10(6):3911-3916 [FREE Full text] [doi: 10.31142/ijahm/v10i6.06]
32. Peterson CT, Denniston K, Chopra D. Therapeutic uses of triphala in ayurvedic medicine. J Altern Complement Med 2017 Aug;23(8):607-614. [doi: 10.1089/acm.2017.0083] [Medline: 28696777]

Abbreviations:

CRIS: Checklist for Reporting In-Vitro Studies
MGACHRC: Mahatma Gandhi Ayurved College, Hospital and Research Centre
NCISM: National Commission for Indian Systems of Medicine

Edited by J Sarvestan; submitted 28.Feb.2025; peer-reviewed by BS Das, B Lucke-Wold, M Umar; revised version received 25.Jul.2025; accepted 25.Jul.2025; published 17.Apr.2026.

Please cite as:
Jadhav PV, Rathi BJ
Development of a Modified Intimate Wash From Yoniprakshalan Kwatha (YOSHA) With Comparative Antimicrobial and Quality Assessment: Protocol for an Experimental Study
JMIR Res Protoc 2026;15:e73239
URL: https://www.researchprotocols.org/2026/1/e73239
doi:10.2196/73239

XSL•FO
RenderX

© Prachi Jadhav, Bharat Jagdishji Rathi. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Impact of Initiating Endocrine Therapy on Metabolic Health in Early Breast Cancer: Protocol for the Prospective Follow-Up EMETA-Study

Lærke Nissen[1,2], BSc; Jonas Busk Holm[1,2], MD; Signe Borgquist[1,2], MD, PhD

[1]Department of Clinical Medicine, Aarhus University, Palle Juul-Jensens Boulevard 11, Aarhus N, Denmark

[2]Department of Oncology, Aarhus University Hospital, Aarhus N, Denmark

Corresponding Author:
Lærke Nissen, BSc
Department of Clinical Medicine, Aarhus University, Palle Juul-Jensens Boulevard 11, Aarhus N, Denmark

Abstract

Background: Adjuvant endocrine therapy is a cornerstone in managing estrogen receptor–positive early breast cancer but may adversely affect metabolic health, including weight gain, insulin resistance, and dyslipidemia. These changes increase the risk of cardiovascular disease and may influence breast cancer outcomes. However, the timing and magnitude of early metabolic changes following endocrine therapy initiation remain poorly characterized. Conventional definitions such as metabolic syndrome rely on dichotomous thresholds and may lack sensitivity to detect early treatment-related metabolic changes, highlighting the need for refined assessment approaches.

Objective: This prospective follow-up study aims to investigate early metabolic effects of initiating adjuvant endocrine therapy in women with estrogen receptor–positive early breast cancer and to compare conventional and expanded approaches to metabolic health classification.

Methods: This single-center, prospective observational study was conducted at Aarhus University Hospital, Denmark. Women aged≥18 years with early-stage estrogen receptor–positive breast cancer initiating adjuvant endocrine therapy and without pre-existing diabetes were eligible. Metabolic health was assessed at baseline and after 3 months using biometric measurements (weight, waist and hip circumference, waist-to-hip ratio, and blood pressure) and non-fasting blood samples (plasma glucose; hemoglobin $A_{1c,}$ (HbA_{1c}); lipid profile; and estradiol). The 3-month follow-up was selected to capture early metabolic changes while aligning with routine clinical care to minimize additional visits and reduce selection bias. Metabolic health will be evaluated using two conventional measures and two extended, exploratory measures. Conventional measures are metabolic syndrome (MetS), defined as meeting ≥3 of 5 established criteria (blood pressure ≥130/85 mmHg, triglycerides >2 mmol/l, high-density lipoprotein cholesterol <1.295 mmol/l, waist circumference >88 cm, and plasma glucose >7.8 mmol/l), and the Metabolic Syndrome z score (MetS-Z), a continuous standardized composite of the MetS components. Additional extended measures are exploratory: the extended MetS, which expands the standard MetS definition by incorporating low-density lipoprotein cholesterol (>3 mmol/l), body mass index (≥30 kg/m²), waist-to-hip ratio (>0.85), and HbA_{1c} (≥42 mmol/mol), and the EMETA score, a standardized composite of the extended MetS components calculated using the same approach as the MetS-*z* score.

Results: The study was funded in July 2024. Recruitment occurred between November 2024 and April 2025, and follow-up was completed in September 2025. Statistical analyses are planned for February 2026, with results expected to be published in summer 2026.

Conclusions: This study is expected to provide insights into early metabolic changes following initiation of adjuvant endocrine therapy and evaluate different approaches to classifying metabolic health. The aim to inform future research by helping to identify patients at increased risk of cardiometabolic complications and adverse breast cancer outcomes, warranting confirmation and validation of expanded metabolic measures in longer-term, larger cohorts.

Trial Registration: Clinicaltrials.gov NCT06623903; https://clinicaltrials.gov/study/NCT06623903

International Registered Report Identifier (IRRID): DERR1-10.2196/78589

(JMIR Res Protoc 2026;15:e78589) doi:10.2196/78589

KEYWORDS

metabolic syndrome; breast neoplasms; antineoplastic agents; cardiometabolic risk factors; weight gain; hormonal

XSL•FO
RenderX

Introduction

Breast cancer is both the most frequent type of cancer and the leading cause of cancer-related mortality among women [1]. As survival rates improve, currently up to 90% in developed countries, there is increasing attention to the long-term health of breast cancer survivors [1,2]. Estrogen receptor-positive (ER+) early (stage I-III) breast cancer (EBC) represents the most common subtype, and adjuvant endocrine therapy (ET) is a cornerstone of its treatment [3,4]. While ET has significantly reduced recurrence and mortality, its metabolic side effects have emerged as a growing concern in survivorship care [4-6].

ET may influence metabolic health through estrogen deprivation [5,7]. Estrogen plays an important role in metabolic homeostasis by promoting insulin sensitivity, regulating lipid metabolism, and supporting vascular function [8,9]. By blocking estrogen receptors (tamoxifen) or suppressing aromatase activity (aromatase inhibitors, AI), ET may disrupt normal metabolic regulation, promoting weight gain, insulin resistance, and dyslipidemia [5,10-13]. These changes are risk factors for cardiovascular disease (CVD), which represents a major noncancer cause of morbidity and mortality among breast cancer survivors [14]. Furthermore, obesity and metabolic syndrome (MetS) at diagnosis have been associated with increased risk of breast cancer recurrence and mortality [15-18].

Although ET is typically administered for several years, routine clinical follow-up is often concentrated around treatment initiation. This limits systematic long-term monitoring of metabolic health in routine care, which highlights the importance of characterizing early metabolic changes. A short-term follow-up allows evaluation of initial changes in weight, lipid metabolism, and glycemic markers that may signal early metabolic vulnerability. A 3-month timeframe balances biological relevance with clinical feasibility, aligns with standard follow-up practice at the study site, and supports broad inclusion while minimizing selection bias.

Metabolic health is traditionally assessed using MetS, a dichotomous classification that identifies individuals as metabolically unhealthy when at least three of five criteria are met: hypertension (blood pressure ≥130/85), hypertriglyceridemia (triglycerides >2 mmol/l), low HDL (HDL <1.295 mmol/l), central obesity (waist circumference >88 cm), or hyperglycemia (plasma glucose >7.8 mmol/l) [19]. However, this binary approach may overlook subtle or early metabolic changes that do not meet clinical thresholds. [19,20]. An established alternative is the Metabolic Syndrome *z* score (MetS-*z* score), a continuous composite measure of standardized MetS components and widely used in research to assess metabolic health along a continuum [21-24]. However, both MetS and MetS-Z omit several cardiometabolic indicators relevant to breast cancer survivorship, including low-density lipoprotein (LDL), hemoglobin A_{1c} (HbA_{1c}), BMI, and waist-to-hip ratio (WHR) [13,25,26]. To address this limitation, this study introduces two expanded measures of metabolic health, for exploratory and hypothesis-generating purposes: An extended MetS definition and the EMETA score, a dichotomous classification and a composite *z* score using the same approach as the MetS-*z* score, both incorporating the extended metabolic variables.

The aim of this study is to evaluate the early impact of initiating adjuvant endocrine therapy on metabolic health in women with ER+ EBC and to compare conventional and expanded approaches to classify metabolic health.

Methods

Study Population

This prospective, single-center observational study enrolled 112 patients with ER+ EBC who initiated adjuvant ET. Over a six-month period from November 2024, participants were recruited from the Department of Oncology, Aarhus University Hospital (AUH), Denmark. Patients were invited to participate during the consultation ET was prescribed. Participants received usual care throughout the study.

Eligibility Criteria

Patients were eligible if they met the following criteria: age ≥18 years, diagnosis of invasive ER+ EBC, and initiation of adjuvant ET. Up to three days of ET treatment were allowed, to allow for the timeframe of informed consent.

Patients were excluded for enrollment in case of pregnancy or lactation, pre-existing diabetes (type I or II), more than three dispenses of ET, or any psychological, familial, sociological, or geographical conditions potentially hampering compliance with the study protocol and follow-up schedule. These conditions were discussed with the patient before trial registration. Patients with pre-existing diabetes were excluded because their chronic metabolic dysregulation and glucose-lowering treatments could confound assessment of early metabolic changes. Patients were excluded if metastatic breast cancer was diagnosed in the period between study inclusion and the 3-month follow-up. As the study did not involve any intervention administered by clinical staff beyond routine care, there were no additional eligibility criteria for sites or individuals delivering interventions.

Study Design

Participants underwent two metabolic health screenings: one at baseline and one at 3-month follow-up (Figure 1). The baseline screening was performed on the same day as study inclusion. The 3-month follow-up screening was performed on the same day as the 3-month ET follow-up. This design minimized additional study-related visits and supported participant retention.

XSL•FO
RenderX

Figure 1. EMETA study design and participation created in BioRender [27] is licensed under CC BY 4.0 [28]. AUH: Aarhus University Hospital.

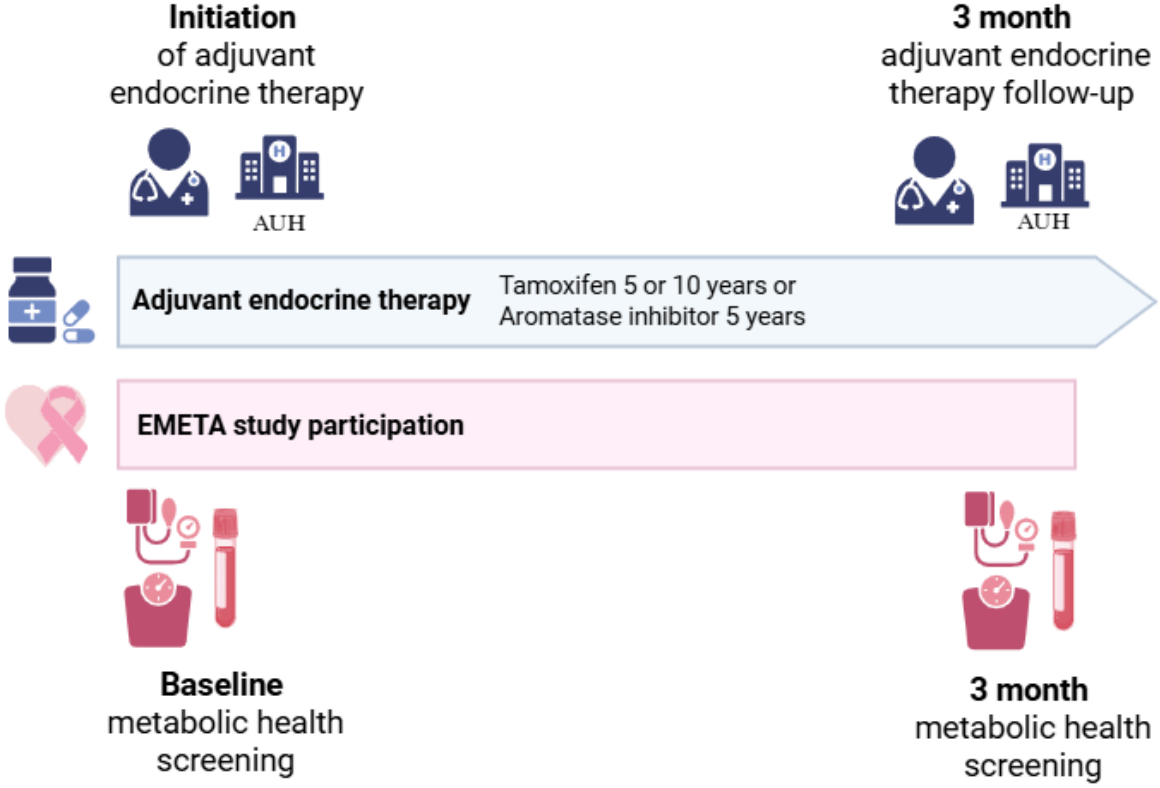

Exposure

The exposure of interest was initiation of ET over a 3-month period. ET consisted of either tamoxifen (20 mg daily) or an AI (letrozole 2.5 mg daily or exemestane 25 mg daily), prescribed according to standard clinical practice.

Outcomes

The primary outcomes were selected to address whether initiation of ET is associated with early changes in overall metabolic health. Given the short-term follow-up and the aim to capture early, potentially subtle metabolic alterations, the primary outcomes include an anthropometric and a composite metabolic measure.

The primary outcomes are:

1. Change in mean BMI from baseline to 3-month follow-up
2. Change in mean MetS-Z from baseline to 3-month follow-up

BMI was chosen because weight changes are a well-known sideeffect of ET and may occur early following ET initiation, representing a clinically meaningful marker of metabolic health. Although short-term changes may be modest, even small increases may indicate emerging metabolic vulnerability. As a continuous composite measure, the widely used MetS-*z* score was selected to capture multidimensional metabolic changes that may not be reflected by individual metabolic parameters alone [21-24,29].

Secondary outcomes were selected to further characterize early metabolic changes and to compare expanded and conventional approaches to metabolic health classification.

The secondary outcomes include:

1. Change in the prevalence of MetS from baseline to 3-month follow-up
2. Comparison of the prevalence of MetS and extended MetS at 3-month follow-up
3. Change in mean EMETA score from baseline to 3-month follow-up
4. Change in mean values of individual metabolic parameters from baseline to 3-month follow-up, including waist circumference, hip circumference, WHR, blood pressure, and circulating levels of HbA_{1c}, total cholesterol, triglycerides, LDL cholesterol, HDL cholesterol, and estradiol

The extended MetS and EMETA score are included as exploratory outcomes to assess whether broader composite measures capture early metabolic changes not identified by conventional definitions. Analysis of individual metabolic parameters provides insight into which metabolic domains are most affected shortly after ET initiation.

Measurements

The screening procedures included biometric measurements and blood sample collections. The biometric measurements were height, weight, hip and waist circumference, and blood pressure. Non-fasting venous blood samples were analyzed for HbA_{1c}, plasma glucose, hemoglobin, total cholesterol, triglycerides, LDL, HDL, and circulating estradiol.

Biometric Measurements

The biometric measurements were measured through the following protocols. Weight was measured using a standardized

XSL•FO
RenderX

scale with participants wearing light clothing and no shoes. Height was recorded with participants standing straight against a wall-mounted measuring tape without shoes. Waist circumference was measured 5 cm above the center of the navel and hip circumference at the widest part of the hip, both using a stretch-resistant tape, ensuring the tape was parallel to the floor and snug but not compressing the skin. Blood pressure was measured using a calibrated cuff on the upper arm after participants had rested for 5 minutes, with three consecutive readings taken with the participant seated both feet on the floor and both arms on their legs.

Blood Samples

Nonfasting venous blood samples were analyzed by the Department of Clinical Biochemistry, AUH for levels of HbA_{1c}, glucose, total cholesterol, triglycerides, LDL, HDL, estradiol, and sex-hormone binding globulin. The blood samples were destroyed immediately after analysis. Results outside standard values were reviewed by consultant physicians and patients, and their general practitioners were notified with follow-up recommendations when necessary.

Data Collection and Management

Information on patient demographics, treatment modalities, and cancer-specific characteristics were collected through a manual search of electronic medical journals, including pathological records. All data were securely stored in a RedCap database hosted by Aarhus University.

To promote participant retention and ensure complete follow-up, the 3-month appointments were booked by hospital secretaries as part of routine care. If a participant was not scheduled for a follow-up appointment, trial personnel proactively contacted the participant to arrange the study visit. If participants discontinued ET, outcome data were still be collected at the 3-month follow-up visit, unless consent was withdrawn.

Cohort Size

The study aimed to include 112 participants over the six-month enrollment period. The target was based on an annual referral rate of approximately 360 patients to the Department of Oncology, AUH, for adjuvant systemic therapy, of which approximately 90% were eligible for ET due to ER+ EBC, and assuming a 70% participation rate among eligible patients.

Statistical power calculations supported the sample size. Based on a comparable study with a similar population [30], where the standard deviation of the change in BMI over 3 - 5 years after diagnosis was 3, a cohort of 112 participants provides >80% power to detect a true change in BMI of 0.8 kg/m^2 at a 5% significance level. Furthermore, other studies have found larger BMI changes over shorter periods. For instance, Heideman et al observed a mean BMI change of 2.0 kg/m^2 (± 4.9) after one year, with 26% of patients gaining ≥5 kg following adjuvant treatment [31].

Statistical Analysis

Definition of Analytic Values

BMI will be calculated as weight (kg)/$height^2$ (m). WHR will be calculated as waist circumference divided by hip circumference. Blood pressure will be calculated as the average of three measurements. Middle arterial pressure will be calculated as 2 times diastolic blood pressure added systolic blood pressure divided by 3.

MetS will be defined as meeting at least three of five criteria: hypertension (blood pressure ≥130/85), hypertriglyceridemia (triglycerides >2 mmol/l), low HDL (HDL <1.295 mmol/l), central obesity (waist circumference >88 cm), or hyperglycemia (plasma glucose >7.8 mmol/l) (Figure 2).

Extended MetS will be defined as meeting at least 3 of the five criteria: hypertension (blood pressure ≥130/85), hypertriglycemia (triglycerides >2 mmol/l or LDL >3 mmol/l), low HDL (HDL <1.295 mmol/l), adiposity (waist circumference >88 cm, BMI ≥30, or WHR >0.85), or hyperglycemia (plasma glucose >7.8 mmol/l or HbA_{1c} ≥42 mmol/mol).

XSL•FO
RenderX

Figure 2. Comparison of MetS, extended MetS, MetS-Z, and EMETA scores. Red text: Variables specific to the extended MetS and EMETA score. HDL: High-density lipoprotein, LDL: Low-density lipoprotein. HbA_{1c}: Hemoglobin A_{1c}. Created in BioRender. Busk Holm, J. (2025) https://BioRender.com/

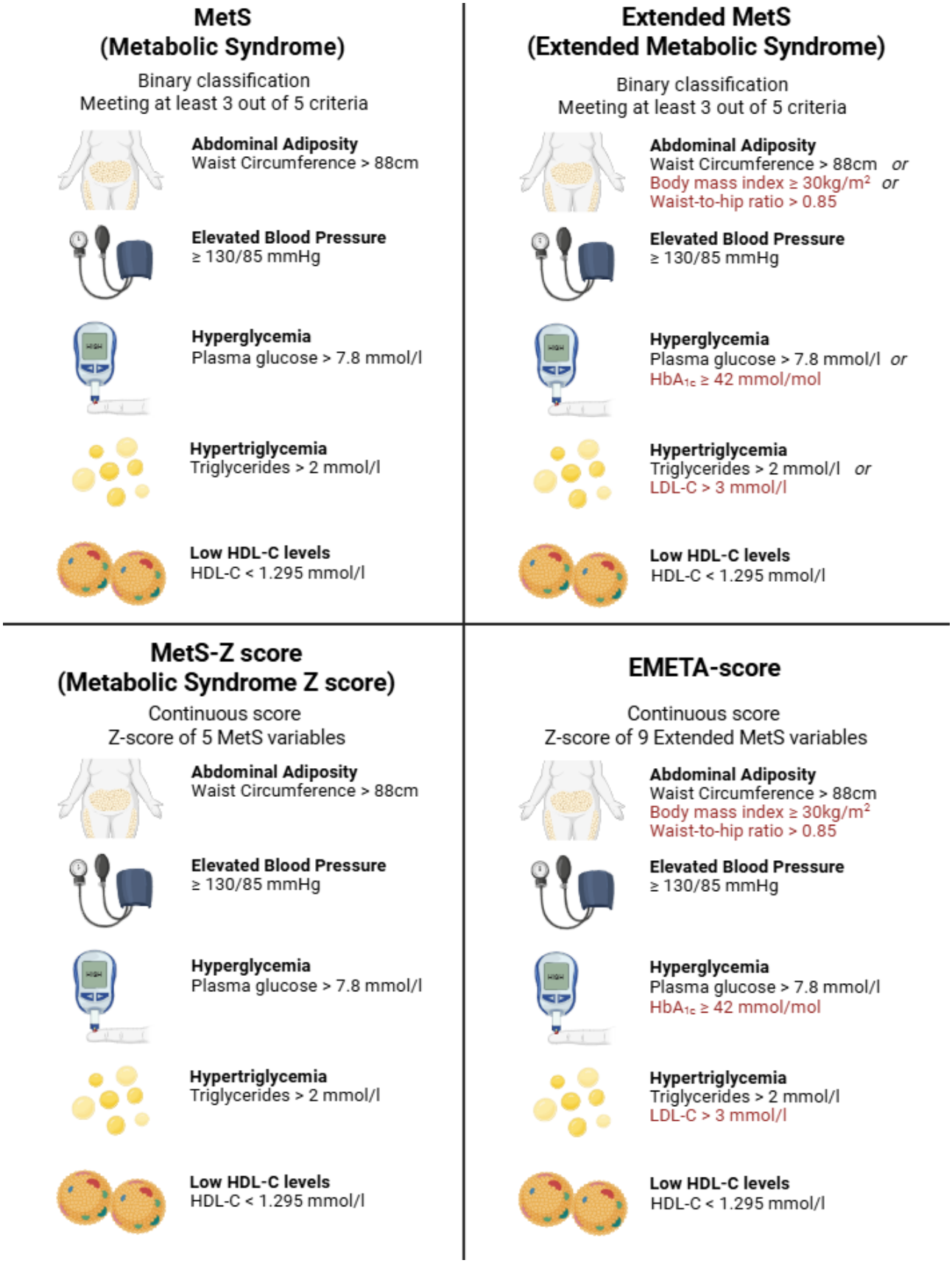

XSL•FO
RenderX

z Score Calculations

The MetS-*z* score will be calculated as a standardized composite score of the five MetS components. Standardization will be achieved by subtracting ATP III criteria from each individual's value of waist circumference (WC), triglycerides (TG), middle arterial pressure (MAP), plasma glucose (PG), and inverted HDL, dividing by the baseline cohort standard deviation (σ).

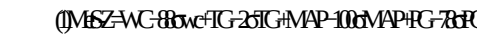

The EMETA score is a exploratory composite *z* score that will be calculated using the same approach as MetS-*z* score, based on the extended MetS components.



Covariables

Information on the following co-variables has been collected from medical records and will be included in adjusted analyses. These include patient demographics: age, weight, height, Carlson Comorbidity score (0, 1‐2, 3‐4, ≥5), current comorbidities (none, cardiopulmonary disease, neurological, dyslipidemia, musculoskeletal, other cancers, abdominal, urogenital, endocrine, dermatological, others), smoking status (never smoker, former smoker, active smoker), civil status (living alone, living with a partner), and area of living (rural/city). Cancer-specific characteristics: date of breast cancer diagnosis, disease stage (T+N), human epidermal growth factor receptor 2 status, histological type, grade of malignancy, and previous breast cancer history. Treatment modalities included duration and type of neo-adjuvant and adjuvant therapy (including supportive prednisolone treatment), date of breast cancer surgery, and radiotherapy.

Analytical Approach

Statistical analyses will be conducted using Stata (version 18.5), with a 5% significance level for hypothesis testing. Changes in continuous primary outcomes (BMI and MetS-*z* score) will be analyzed using linear regression or linear mixed-effects models, as appropriate, to estimate mean within-person changes from baseline to 3-month follow-up. Changes in secondary continuous outcomes, including the EMETA score and individual metabolic parameters, will be analyzed using similar regression-based approaches. Changes in categorical outcomes (MetS and Extended MetS prevalence) will be analyzed using logistic regression models.

To account for exposure heterogeneity and treatment-related confounding, type of ET (tamoxifen vs aromatase inhibitor) and prior cancer treatments with potential metabolic impact, including neoadjuvant or adjuvant chemotherapy and supportive corticosteroid use, will be included as covariates in all adjusted analyses.

To address the secondary research question of whether certain patients are more vulnerable to early metabolic health deterioration following the initiation of ET, exploratory subgroup analyses will be conducted across both primary and secondary outcomes. These analyses will examine whether metabolic changes differ according to baseline and treatment-related characteristics, including age, BMI category, baseline metabolic status, menopausal status, type of ET, prior chemotherapy, smoking status, and comorbidity score. These subgroup analyses will be considered hypothesis-generating. Missing data will be assessed for randomness and extent. If data are missing at random, appropriate methods such as multiple imputation or mixed-effects models will be applied. Sensitivity analyses will be conducted to evaluate the potential impact of missing data on the study's primary and secondary outcomes.

Ethical Considerations

This study was conducted in agreement with the Declaration of Helsinki and the laws and regulations of Denmark, whichever provided the highest level of patient protection. The study protocol was approved by the Scientific Ethics Committee for the Central Denmark Region (case nr: 1-10-72-122-24). Any important modifications to the protocol (eg, changes to eligibility criteria, outcomes, or analysis methods) were communicated to the relevant ethics committee and trial registry (ClinicalTrials.gov ID NCT06623903).

Informed Consent

Written informed consent was obtained from all participants before study entry, and any study-specific procedures were performed. This process was conducted in accordance with national and local regulatory requirements. Each consent form had to be signed and personally dated by the participant.

Risks and Patient Inconvenience

The risk of adverse events associated with study participation was minimal. Rarely, minor bleeding or inflammation could occur at the blood sample puncture site. No long-term adverse events were anticipated. Participation required two additional blood samples and an approximate extension of 30 minutes during two oncology department visits (at study inclusion and the three-mo follow-up). Participants were covered under the standard patient compensation system in Denmark for any unexpected harm related to study procedures.

Data Protection and Privacy

Personal information collected during the study was securely stored in the RedCap database hosted at Aarhus University. Data were pseudonymized to enable linkage with clinical records, with access to identifiable information restricted to authorized study personnel. All data were stored on secure, password-protected servers with two-factor authentication. All data were processed in accordance with the Danish Data Protection Act and the European General Data Protection Regulation (GDPR). The project was registered in the internal directory of the Central Denmark Region.

Dissemination of Results

Study findings will be disseminated in international peer-reviewed journals and shared at local journal clubs, scientific meetings, and at least one international congress. Study participants who consented at study inclusion to receive study results will be provided with a lay summary of the findings, supporting transparency and engagement in survivorship care. Patients or members of the public were not involved in the design, conduct, or reporting of this study.

XSL•FO
RenderX

Monitoring

A formal Data Monitoring Committee was not established for this study, as this was a noninterventional, observational trial involving minimal risk. Instead, the study was overseen by the study personnel (the authors) who were responsible for the daily conduct and integrity of the study. The study personnel completed and monitored all data collection to ensure protocol adherence, data quality, and participant safety. As the study sponsor (AUH) did not influence design, data analysis, or reporting, oversight by the study team was considered both appropriate and independent in practice. No interim analyses or formal stopping guidelines were planned, given the short follow-up period and non-interventional design. All data will be analyzed following the completion of the follow-up period.

Results

The study was funded in July 2024. Participant recruitment occurred from November 2024 through April 2025, during which 112 of 166 eligible patients were enrolled. Among the 54 eligible patients who were not enrolled, 18 declined participation, 26 were not informed of the study by their treating practitioner, and 10 elected not to initiate adjuvant ET. All enrolled participants completed baseline assessments, and follow-up data collection was completed in September 2025. No deviations from the approved study protocol occurred, and no protocol amendments were made following ethical approval. Data cleaning and statistical analyses are scheduled to commence in February 2026, with dissemination of the final study results anticipated in July 2026.

Discussion

This study is anticipated to indicate that initiation of ET in women with ER+ EBC will be associated with modest early metabolic changes over the first three months of treatment. These changes may include small increases in BMI and selected lipid-related parameters, while markers with longer biological response times, such as HbA_{1c}, may show minimal or no short-term change. Extended and continuous measures of metabolic health may capture early alterations not identified by conventional MetS definitions. Importantly, these analyses are exploratory and intended to evaluate sensitivity and feasibility rather than to establish clinical thresholds or predictive value.

Previous studies have linked ET with weight gain, insulin resistance, and dyslipidemia, all of which increase cardiovascular risk and may impact breast cancer outcomes [5,10-12,14,32]. However, most prior work has focused on longer-term outcomes or relied on dichotomous classifications. By evaluating both conventional and expanded measures within a short-term framework, this study contributes to understanding how early metabolic changes may be characterized following ET initiation.

Key strengths of this study include its prospective design with paired baseline and 3-month follow-up data, allowing within-subject comparison of early metabolic changes. The use of both conventional and expanded metabolic measures enables comparison of different classification approaches and addresses known limitations of binary definitions. The short follow-up period and alignment with routine clinical care minimize additional participant burden, supporting broad inclusion and reducing the risk of selection bias related to study participation.

The study is conducted at a single center, which may limit generalizability beyond similar healthcare settings. Exclusion of patients with pre-existing diabetes, while necessary to reduce baseline metabolic confounding, restricts applicability to a population that commonly receives endocrine therapy. In addition, although efforts were made to minimize attrition, informative dropout cannot be excluded, particularly if early metabolic or treatment-related symptoms influenced follow-up participation. The modest sample size limits power for subgroup analyses, which are therefore considered exploratory.

Overall, this study aims to inform future research and clinical practice by clarifying early metabolic patterns following ET initiation and by evaluating alternative approaches to metabolic health assessment. Larger and longer-term studies are needed to determine whether early metabolic changes are associated with subsequent cardiovascular risk or breast cancer outcomes, and to validate expanded metabolic measures before any potential clinical application.

Acknowledgments

The authors would like to express their gratitude to and acknowledge the patients who participated in the study, and the scientific and medical staff from the Department of Oncology, AUH, who participated in conducting the study. The sponsor of this study is AUH (Contact: AarhusUniversitetshospital@auh.rm.dk).

Funding

This work was supported by The Danish Cancer Society grant number R382-A22956, Fru Astrid Thaysens Legat for Lægevidenskabelig Grundforskning grant number ATL 24/06, and Department of Oncology Research Fund grant number 09-2271.

Disclaimer

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: Idea generation & Proofreading and editing. The GAI tool used was: ChatGPT 5. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are not publicly available due to patient privacy and confidentiality concerns but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualiztion: LN (lead), SB (equal), JBH (equal), and SB research group (supporting).
Data curation, funding acqusition, investigation, methology, project administration, visualization: LN (lead), SB (supporting), JBH (supporting).
Supervision: SB (Lead), JBH (supporting).
Writing-original draft, review & editing: LN (lead), JBH (Supporting), SB (supporting).

Conflicts of Interest

None declared.

References

1. Miller KD, Nogueira L, Devasia T, et al. Cancer treatment and survivorship statistics, 2022. CA Cancer J Clin 2022 Sep;72(5):409-436. [doi: 10.3322/caac.21731] [Medline: 35736631]
2. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]
3. Galuia M, Fedorova J, McHayleh W, Mamounas E, Ahmad S, Pavri S. Perioperative drug management of systemic therapies in breast cancer: a literature review and treatment recommendations. Curr Oncol 2025 Mar 9;32(3):154. [doi: 10.3390/curroncol32030154] [Medline: 40136358]
4. Franzoi MA, Agostinetto E, Perachino M, et al. Evidence-based approaches for the management of side-effects of adjuvant endocrine therapy in patients with breast cancer. Lancet Oncol 2021 Jul;22(7):e303-e313. [doi: 10.1016/S1470-2045(20)30666-5] [Medline: 33891888]
5. Thomas NS, Scalzo RL, Wellberg EA. Diabetes mellitus in breast cancer survivors: metabolic effects of endocrine therapy. Nat Rev Endocrinol 2024 Jan;20(1):16-26. [doi: 10.1038/s41574-023-00899-0] [Medline: 37783846]
6. Nardin S, Mora E, Varughese FM, et al. Breast cancer survivorship, quality of life, and late toxicities. Front Oncol 2020;10:864. [doi: 10.3389/fonc.2020.00864] [Medline: 32612947]
7. Li Y, Deng Z, Wang Y, Shen S. Lipid changes during endocrine therapy in early-stage breast cancer patients: a real-world study. Lipids Health Dis 2024 Jan 8;23(1):9. [doi: 10.1186/s12944-024-02002-6] [Medline: 38191454]
8. Tramunt B, Smati S, Grandgeorge N, et al. Sex differences in metabolic regulation and diabetes susceptibility. Diabetologia 2020 Mar;63(3):453-461. [doi: 10.1007/s00125-019-05040-3] [Medline: 31754750]
9. Sánchez-García M, León-Wu K, de Miguel-Ibáñez R, et al. Metabolic changes in patients with premature ovarian insufficiency: adipose tissue focus-a narrative review. Metabolites 2025 Apr 2;15(4):242. [doi: 10.3390/metabo15040242] [Medline: 40278371]
10. Kwan ML, Cheng RK, Iribarren C, et al. Risk of cardiometabolic risk factors in women with and without a history of breast cancer: The Pathways Heart Study. J Clin Oncol 2022 May 20;40(15):1635-1646. [doi: 10.1200/JCO.21.01738] [Medline: 35025627]
11. Rillamas-Sun E, Kwan ML, Iribarren C, et al. Development of cardiometabolic risk factors following endocrine therapy in women with breast cancer. Breast Cancer Res Treat 2023 Aug;201(1):117-126. [doi: 10.1007/s10549-023-06997-x] [Medline: 37326764]
12. Franchi M, Tritto R, Tarantini L, Navazio A, Corrao G. Adjuvant hormonotherapy and cardiovascular risk in post-menopausal women with breast cancer: a large population-based cohort study. Cancers (Basel) 2021 May 8;13(9):2254. [doi: 10.3390/cancers13092254] [Medline: 34066685]
13. Mehta LS, Watson KE, Barac A, et al. Cardiovascular disease and breast cancer: where these entities intersect: a scientific statement from the American Heart Association. Circulation 2018 Feb 20;137(8):e30-e66. [doi: 10.1161/CIR.0000000000000556] [Medline: 29437116]
14. Bradshaw PT, Stevens J, Khankari N, Teitelbaum SL, Neugut AI, Gammon MD. Cardiovascular disease mortality among breast cancer survivors. Epidemiology 2016 Jan;27(1):6-13. [doi: 10.1097/EDE.0000000000000394] [Medline: 26414938]
15. Harborg S, Cronin-Fenton D, Jensen MBR, Ahern TP, Ewertz M, Borgquist S. Obesity and risk of recurrence in patients with breast cancer treated with aromatase inhibitors. JAMA Netw Open 2023 Oct 2;6(10):e2337780. [doi: 10.1001/jamanetworkopen.2023.37780] [Medline: 37831449]
16. Playdon MC, Bracken MB, Sanft TB, Ligibel JA, Harrigan M, Irwin ML. Weight gain after breast cancer diagnosis and all-cause mortality: systematic review and meta-analysis. J Natl Cancer Inst 2015 Dec;107(12):djv275. [doi: 10.1093/jnci/djv275] [Medline: 26424778]

XSL•FO
RenderX

17. Wang X, Hui TL, Wang MQ, Liu H, Li RY, Song ZC. Body mass index at diagnosis as a prognostic factor for early-stage invasive breast cancer after surgical resection. Oncol Res Treat 2019;42(4):195-201. [doi: 10.1159/000496548] [Medline: 30852575]
18. Harborg S, Larsen HB, Elsgaard S, Borgquist S. Metabolic syndrome is associated with breast cancer mortality: a systematic review and meta-analysis. J Intern Med 2025 Mar;297(3):262-275. [doi: 10.1111/joim.20052] [Medline: 39775978]
19. Neeland IJ, Lim S, Tchernof A, et al. Metabolic syndrome. Nat Rev Dis Primers 2024 Oct 17;10(1):77. [doi: 10.1038/s41572-024-00563-5] [Medline: 39420195]
20. DeBoer MD, Filipp SL, Gurka MJ. Use of a Metabolic syndrome severity z score to track risk during treatment of prediabetes: an analysis of the Diabetes Prevention Program. Diabetes Care 2018 Nov;41(11):2421-2430. [doi: 10.2337/dc18-1079] [Medline: 30275282]
21. Johnson JL, Slentz CA, Houmard JA, et al. Exercise training amount and intensity effects on metabolic syndrome (from Studies of a Targeted Risk Reduction Intervention through Defined Exercise). Am J Cardiol 2007 Dec 15;100(12):1759-1766. [doi: 10.1016/j.amjcard.2007.07.027] [Medline: 18082522]
22. Hettchen M, von Stengel S, Kohl M, et al. Changes in menopausal risk factors in early postmenopausal osteopenic women after 13 months of high-intensity exercise: the randomized controlled ACTLIFE-RCT. Clin Interv Aging 2021;16:83-96. [doi: 10.2147/CIA.S283177] [Medline: 33469276]
23. Morales-Palomo F, Moreno-Cabañas A, Alvarez-Jimenez L, Mora-Gonzalez D, Mora-Rodriguez R. Long-term effects of high-intensity aerobic training on metabolic syndrome: an 8-year follow-up randomized clinical trial. J Cachexia Sarcopenia Muscle 2025 Apr;16(2):e13780. [doi: 10.1002/jcsm.13780] [Medline: 40171886]
24. Gonzalo-Encabo P, Christopher CN, Lee K, et al. High-intensity interval training improves metabolic syndrome in women with breast cancer receiving Anthracyclines. Scand J Med Sci Sports 2023 Apr;33(4):475-484. [doi: 10.1111/sms.14280] [Medline: 36427275]
25. Grundy SM, Cleeman JI, Daniels SR, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation 2005 Oct 25;112(17):2735-2752. [doi: 10.1161/CIRCULATIONAHA.105.169404] [Medline: 16157765]
26. Sinning C, Makarova N, Völzke H, et al. Association of glycated hemoglobin A1c levels with cardiovascular outcomes in the general population: results from the BiomarCaRE (Biomarker for Cardiovascular Risk Assessment in Europe) consortium. Cardiovasc Diabetol 2021 Nov 15;20(1):223. [doi: 10.1186/s12933-021-01413-4] [Medline: 34781939]
27. EMETA study design and participation. BioRender. URL: https://BioRender.com/d5aglhl [accessed 2026-03-13]
28. Attribution 4.0 international (CC BY 4.0). Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2026-03-13]
29. Dieli-Conwright CM, Courneya KS, Demark-Wahnefried W, et al. Effects of aerobic and resistance exercise on metabolic syndrome, sarcopenic obesity, and circulating biomarkers in overweight or obese survivors of breast cancer: a randomized controlled trial. J Clin Oncol 2018 Mar 20;36(9):875-883. [doi: 10.1200/JCO.2017.75.7526] [Medline: 29356607]
30. Goyal A, Milner GE, Cimino-Mathews A, et al. Weight gain after hormone receptor-positive breast cancer. Curr Oncol 2022 Jun 6;29(6):4090-4103. [doi: 10.3390/curroncol29060326] [Medline: 35735435]
31. Heideman WH, Russell NS, Gundy C, Rookus MA, Voskuil DW. The frequency, magnitude and timing of post-diagnosis body weight gain in Dutch breast cancer survivors. Eur J Cancer 2009 Jan;45(1):119-126. [doi: 10.1016/j.ejca.2008.09.003] [Medline: 18930387]
32. Bergman R, Berko YA, Sanchez V, et al. Obesity and metabolic syndrome are associated with short-term endocrine therapy resistance in early ER + breast cancer. Breast Cancer Res Treat 2023 Jan;197(2):307-317. [doi: 10.1007/s10549-022-06794-y] [Medline: 36396775]

Abbreviations

AI: aromatase inhibitor
AUH: Aarhus University Hospital
CVD: cardiovascular disease
EBC: early breast cancer
ER+: estrogen receptor positive
ET: adjuvant endocrine therapy
HbA_{1c}: hemoglobin A_{1c}
HDL: high-density lipoprotein
LDL: low-density lipoprotein
MetS: metabolic syndrome
MetS-Z score: Metabolic Syndrome z score
WHR: waist-to-hip ratio

XSL•FO
RenderX

Edited by A Schwartz; submitted 16.Jul.2025; peer-reviewed by C Koebnick; revised version received 14.Jan.2026; accepted 28.Jan.2026; published 25.Mar.2026.

Please cite as:
Nissen L, Busk Holm J, Borgquist S
Exploring the Impact of Initiating Endocrine Therapy on Metabolic Health in Early Breast Cancer: Protocol for the Prospective Follow-Up EMETA-Study
JMIR Res Protoc 2026;15:e78589
URL: https://www.researchprotocols.org/2026/1/e78589
doi:10.2196/78589

© Lærke Nissen, Jonas Busk Holm, Signe Borgquist. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 25.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Community Health Worker–Integrated Model for Sickle Cell Disease Management: Protocol for a Feasibility Study

Tilicia Mayo-Gamble[1], MA, MPH, PhD, CHES; Doreen Ugwu[1], MPH, MBBCh; Tobi Oloyede[1], MA; Joseph Telfair[1], MSW, DrPH; Abdullah Kutlar[2], MD

[1]Health Policy and Community Health, Jiann-Ping Hsu College of Public Health, Georgia Southern University, 501 Forest Dr, Statesboro, GA, United States

[2]Center for Blood Disorders, Medical College of Georgia, Augusta University, Augusta, GA, United States

Corresponding Author:
Tilicia Mayo-Gamble, MA, MPH, PhD, CHES
Health Policy and Community Health, Jiann-Ping Hsu College of Public Health, Georgia Southern University, 501 Forest Dr, Statesboro, GA, United States

Abstract

Background: Sickle cell disease (SCD) is a chronic inherited blood disorder that disproportionately affects African Americans. Managing SCD is complex, often requiring coordination of care among multiple health care professionals. These care challenges result in fragmented service delivery and poor health outcomes. Integrating community health workers (CHWs) into patient-centered coordinated care may enhance disease management and patient satisfaction with the care received.

Objective: The proposed study aims to test the feasibility and acceptability of a CHW-integrated model for SCD management.

Methods: This protocol is for a multimethod study to evaluate the feasibility of an integrated CHW intervention. The intervention includes CHW-led patient education and care coordination. Adults with SCD (aged 18 - 45 years) will be recruited from the Augusta University Sickle Cell Center's satellite clinic in Albany, Georgia. Quantitative outcomes include self-care behaviors and patient satisfaction. Qualitative data from semistructured interviews will be used to assess the acceptability of the intervention.

Results: The proposed study will provide preliminary data on the impact of a CHW-integrated model on care coordination and self-management among adults with SCD. As of March 2026, we have enrolled 16 participants.

Conclusions: Despite national calls for improved care coordination and disease management in this population, few published protocols detail community-based, culturally aligned approaches that center on the lived experiences and structural barriers faced by adults with SCD. Our study addresses a gap in the literature by highlighting a theory-informed, multimethod study protocol for evaluating the feasibility, effectiveness, and acceptability of integrating CHWs into a patient-centered coordinated care model for adults with SCD. This protocol outlines a scalable, community-embedded strategy that may inform future clinical trials and implementation efforts.

International Registered Report Identifier (IRRID): DERR1-10.2196/82663

(JMIR Res Protoc 2026;15:e82663) doi:10.2196/82663

KEYWORDS

sickle cell disease; community health worker; patient-centered care; self-care; mixed methods; research protocol

Introduction

Sickle cell disease (SCD) is an inherited hemoglobin disorder that affects 1 in 365 African Americans [1,2]. Disparities in receiving health care among African Americans and other racial/ethnic minority groups in the United States are well documented and are directly related to poor health outcomes associated with SCD. Individuals with SCD experience physical, psychological, and social comorbidities that impact effective disease management [3,4]. The challenges of providing care to this population with complex needs have led to fragmented care, poor patient experiences, and inadequate implementation of evidence-based SCD management [5,6]. In 2014, the National Heart, Lung, and Blood Institute released a report on "Evidence-Based Management of Sickle Cell Disease" [7], calling for strategies to improve disease management by patients and providers [8]. Patient-centered coordinated care (PCCC) is an evidence-based strategy that can potentially reduce care fragmentation and improve patient experiences. However, studies report deficiencies in care coordination and patient satisfaction among individuals with SCD [9,10]. These deficiencies contribute to the inability of African Americans with SCD to obtain consistent, preventive care, resulting in poorer physical and psychological health outcomes.

Georgia ranks among the top 3 states in the prevalence of SCD, affecting more than 10,000 individuals, including a significant

XSL•FO
RenderX

adult population, across nearly every county [11]. Access to specialized care constitutes a significant barrier for these individuals, as data from the Sickle Cell Data Collection program reveal that 10% of babies born with SCD between 2004 and 2016 lived more than 1 hour away from the nearest specialty care center, a situation that contributes to delays in treatment and increased health care use that can persist into adulthood [11]. Many adults with SCD experience challenges related to treatment adherence and health care access, which underscores the urgent need for community-based interventions aimed at enhancing support for this population. To address these barriers, the Georgia Senate Study Committee on Sickle Cell Disease recommended expanding community-based solutions, such as leveraging Federally Qualified Health Centers, mobile health units, and telemedicine—approaches that align with the integration of community health workers (CHWs) to support disease management [12].

A community-based approach could bridge gaps in health care services to improve patient care for SCD. CHWs have demonstrated promise in improving health behaviors and outcomes among racial and ethnic minority populations with chronic diseases who lack access to adequate health care [13,14]. CHWs can be integrated into a PCCC team to enhance care coordination and facilitate disease management among adults with SCD. This evidence-based strategy is supported by the Sickle Cell Disease Association of America, a national advocacy organization for the SCD community [15]. Studies have demonstrated that both PCCC and CHW interventions are effective models for supporting disease management in chronic conditions [16,17]. However, there are few models for integrating CHWs into a PCCC team to enhance disease management and patient experiences [17,18]. Furthermore, no published studies have used such an integrative model among African American adults with SCD.

This research will build on our understanding of the role of CHWs in the evidence-based management of SCD and will identify challenges associated with integrating a CHW into a clinical care team. To realize the tenets of this hypothesis, three specific aims will be pursued: (1) assessment of the feasibility of integrating a CHW into a PCCC team for SCD, (2) evaluation of the effectiveness of enhanced care coordination using an integrated PCCC-CHW model on self-care behaviors and patient satisfaction, and (3) determination of the acceptability of the intervention and protocol among patients and care team members through semistructured interviews. Our study will evaluate the hypotheses that by integrating a CHW into a PCCC team, barriers to disease management can be overcome while simultaneously improving the quality of care for adults with SCD.

Methods

Overview

The chronic care model (CCM) is a comprehensive framework designed to shift the management of chronic diseases from a reactive approach to a more proactive, planned strategy that emphasizes continuous patient-provider relationships, individualized care, and evidence-based practices across health care systems [19]. It identifies elements of a health care system that encourage high-quality chronic disease care. The CCM posits that productive interactions should occur between patients and providers [19]. However, the framework is limited in describing processes for facilitating these interactions. The proposed study will examine the effectiveness of integrating a CHW as a liaison between patients and providers. This integration is expected to enhance care coordination and support chronic disease management, particularly among adults affected by SCD [13,14].

The theory of self-care management for SCD is adapted from the broader theory of self-care management for vulnerable populations, which describes variables that influence self-care management, health status, and quality of life among populations that experience or are at risk for health disparities [20]. This theory underscores the significance of self-care practices in managing chronic disease, particularly by helping individuals navigate the demands of their condition. In this research, the self-care construct from the theory of self-care management for SCD will be applied as the self-management support element of the CCM, thereby fostering self-care behaviors to improve health outcomes among individuals with SCD.

The goal of this research protocol is to evaluate the feasibility and effectiveness of integrating a CHW into a PCCC team to enhance self-care for adults with SCD. Adopting the CCM and the theory of self-care management for SCD, we hypothesize that a CHW can be successfully integrated into a PCCC team to overcome barriers to disease management and improve the quality of care for adults with SCD (Figure 1).

XSL•FO
RenderX

Figure 1. Adapted version of the chronic care model.

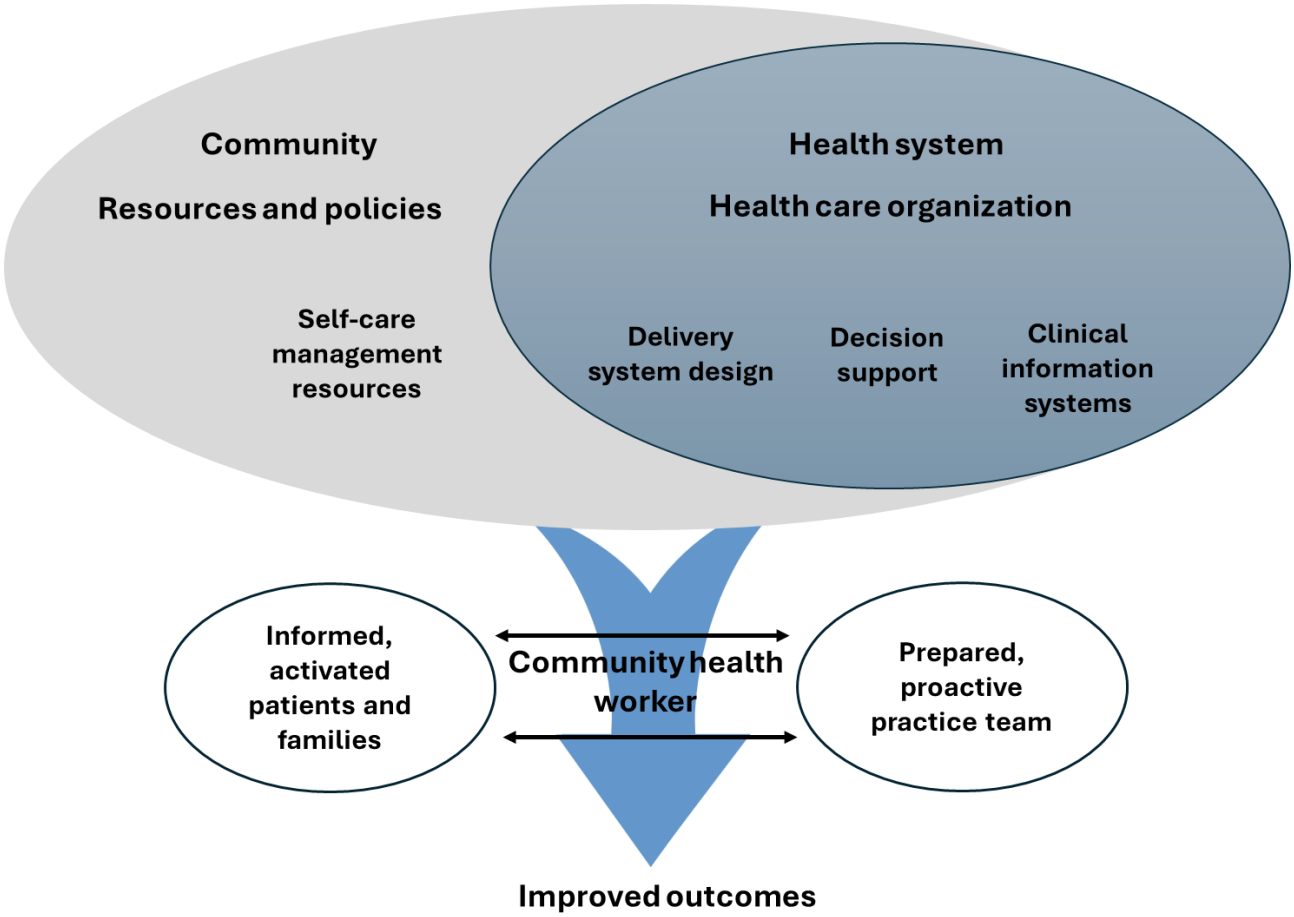

Figure 1 illustrates the proposed integrated model. The CCM identifies elements of a health care system that encourage high-quality chronic disease care and emphasizes productive interactions between informed, activated patients and prepared, proactive care teams [19]. For the proposed model, we will assess the effectiveness of a CHW as a liaison to facilitate productive interactions between patients and providers [16-18]. To strengthen theoretical alignment with SCD-specific care needs, the self-management support component of the CCM is operationalized using the theory of self-care management for SCD, which specifies individual-level processes that influence self-care behaviors and health outcomes in this population [20]. Within this integrated framework, CHWs will support self-care skill development, reinforce patient engagement, and facilitate coordination among patients, providers, and community resources. The proposed study will use a CHW to offer social support and implement evidence-based guidelines through virtual or telephone-based sessions. This approach will allow patients to remain in their home setting, where they experience the daily challenges of self-management.

Ethical Considerations

This protocol will be conducted in accordance with the Declaration of Helsinki. Ethics approval has been granted on an expedited basis by the institutional review board of Georgia Southern University (H21406) for surveys and interviews conducted with human subjects. Informed consent will be obtained from all participants, who will then receive a link to the baseline questionnaire to be completed via Qualtrics XM (Silver Lake and CPP Investments). Each participant will receive a US $20 gift card after completing the questionnaire.

Health Insurance Portability and Accountability Act (HIPAA)–protected information will not be discussed during calls. Information collected during calls will be stored on a cloud server that is only accessible to the research team. To protect the confidentiality of participants, each of them will be assigned a unique participant code.

CHW Training

In a prior study, the lead investigator conducted a patient-centered outcomes project that trained CHWs on evidence-based management for SCD and the facilitation of education on SCD management for patients and caregivers [21]. On the basis of the study's findings, we anticipate that participants will perceive CHWs as an acceptable method for delivering SCD education. CHWs (n=2) will be recruited from My Three Sicklers Sickle Cell Foundation Inc (MTS), a Georgia-based community organization that focuses on SCD outreach, management, and policy.

For training, CHWs will complete a "Remote Community Health Worker Certificate" via HomeTown Health, a network of rural hospitals that provides continuing education and training for health care providers. This training provides CHWs with foundational knowledge on the role of CHWs and the delivery of education in remote settings. Once completed, the principal investigator (PI) will provide project onboarding and training

XSL•FO
RenderX

for SCD management. MTS will provide ongoing support and supervision of CHWs throughout the study duration.

Study Design, Setting, and Sample

This study will use a pre-post, nonrandomized, 6-month longitudinal intervention design, implemented using questionnaires and semistructured interviews (Figure 2). Enrollment will occur through the Augusta University Sickle Cell Center's satellite clinic, Phoebe Putney Cancer Center, in Albany, Georgia. This site has not previously worked with a CHW, either directly or through a community-based organization.

Eligible participants will be adults with SCD aged 18 - 45 years. Given the target sample size (n=40), at 80% power and α=.05, we can detect small to moderate effect sizes (r^2=0.4 to 0.6) using a 2-tailed paired *t* test [22]. This sample includes an expected 20% attrition rate. There are currently 97 eligible individuals seeking SCD care at this location. This is more than double the expected enrollment of 30 participants. Using a rolling recruitment process, the PI will work with MTS to enroll additional participants if an enrolled participant is lost to follow-up. The assumption of a small to moderate effect is conservative, given that studies designed to improve self-care in SCD report a medium effect size (*d*=0.71) [8,23].

Figure 2. Sickle cell disease community health worker intervention flowchart.

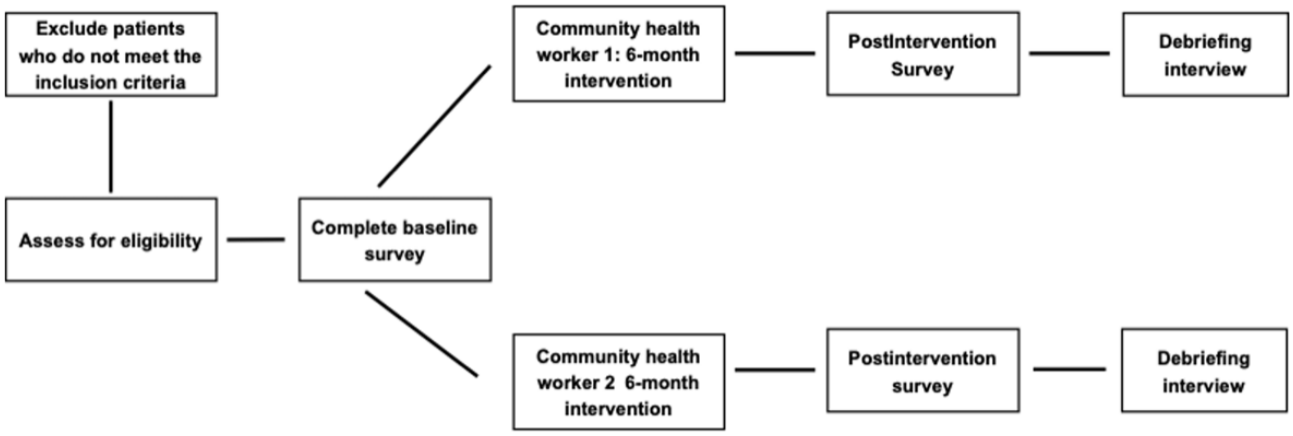

Procedures and Data Collection

The primary outcomes for this study are patient satisfaction and self-care behaviors (including appointment keeping, tracking health issues, and managing medications). These will be assessed using validated instruments at baseline and after intervention (following the 6-month CHW-integrated PCCC model). Feasibility will be assessed through (1) enrollment numbers within a 2-year time frame, (2) retention, (3) intervention completion, and (4) acceptability of the intervention among participants.

Study Procedures

Adults with SCD (aged 18 - 45 years) will be recruited from the Augusta University Sickle Cell Center's satellite clinic in Albany, Georgia. Clinic staff will identify participants who meet the inclusion criteria. For eligible patients, clinic staff will describe the study and offer participation during the patient's regularly scheduled clinic visit. During the visit, interested patients will provide written informed consent and the link to the baseline questionnaire to be completed via Qualtrics. Each participant will receive a $20 gift card after completing the questionnaires. Following enrollment, participants will be referred to MTS for assignment to a CHW.

The CHW will schedule an initial session (60 minutes) with participants, followed by 5 subsequent monthly sessions (virtual or telephone-based) scheduled for 30 to 60 minutes, for a maximum of 6 patient contact hours over 6 consecutive months. During each monthly session, the CHW will educate participants on SCD management (ie, preventive services, management of acute complications, and management of chronic complications). The CHW will use a subjective, objective, assessment, and plan (SOAP) format for reporting via Qualtrics. The SOAP form is a standardized documentation method used by health care providers to systematically record patient information, thereby enhancing communication about health care experiences [24]. By organizing patient encounters into 4 distinct sections, the SOAP note serves as a clear framework for evaluating patient data and outlining management strategies [24]. This report will allow the CHW to document observations from patient interactions.

After completion of the sixth and final session, the CHW will provide a link to the postintervention questionnaire to be completed via Qualtrics. The postintervention questionnaire contains a question of interest about participating in debriefing interviews. We will conduct semistructured debriefing interviews with patients (n=35 approximately) to determine the perceived feasibility and acceptability of the study procedures. A member of the research team will conduct the interviews using a semistructured interview guide. Interviews are expected to last 45 to 60 minutes, and compensation will be a US $25 gift card. All incentives will be distributed by the PI per the institutional review board–approved protocol for research involving human subjects.

Measures

Overview

Eligible patients will be referred to MTS for assignment to a CHW. At the initial session, participants will complete a baseline survey assessing patient satisfaction and self-report self-care behaviors. We will use a demographics questionnaire, the

XSL•FO
RenderX

Medical Outcomes Study Patient Satisfaction Questionnaire [25], and the Jenerette Self-Care Assessment Tool [20] (refer to Multimedia Appendix 1 for details). The CHW will complete a SOAP report at each session, documenting observations of the participant's condition and responses related to self-care behaviors. After each visit, the CHW will report the results to the PI. These sessions will allow the PI to assess study fidelity and address any challenges with implementation or potential participant dropout. It will also allow the research team to determine whether modifications to the program are needed to ensure participant retention. During the final CHW session, the participant will complete a postintervention survey on patient satisfaction and self-report self-care behaviors.

Patient satisfaction with medical care will be assessed using the Medical Outcomes Study's Patient Satisfaction Questionnaire Short Form (PSQ-18) [25]. The PSQ-18 is an 18-item self-report instrument designed to evaluate patients' satisfaction with health care services across multiple domains, including general satisfaction, technical quality, interpersonal manner, communication, financial aspects, time spent with providers, and accessibility and convenience. Items are rated on a 5-point Likert scale ranging from strongly disagree to strongly agree. Subscale scores are computed by averaging item responses, with higher scores indicating greater satisfaction with care. The PSQ-18 has demonstrated good reliability and validity across diverse patient populations and is widely used in health services research [25].

SCD-specific self-care behaviors will be assessed using the Jenerette Self-Care Assessment Tool [20]. This instrument was developed to evaluate self-care management behaviors among individuals with SCD, including symptom monitoring, pain management, medication adherence, and engagement in health-promoting behaviors. The tool consists of multiple self-report items rated on a Likert-type scale, with higher scores indicating greater engagement in effective self-care behaviors. The Jenerette Self-Care Assessment Tool has demonstrated acceptable reliability and validity and has been used in previous studies examining self-care management among adults with SCD [20].

Acceptability of the Intervention

To determine the acceptability of the protocol, we will use semistructured interviews to determine patient satisfaction with the intervention protocol and specific components (clarity, ease of use, relevance, helpfulness, and new information), actual time required, trust or rapport with the CHW, influence on self-care, and suggestions for improvement. Additionally, research staff will conduct interviews with the CHWs (n=2) and clinic staff (n=3). Interviews are expected to last 30 minutes.

Data Analysis

SPSS (version 25; IBM Corp) will be used to analyze quantitative data collected from the survey instruments. Categorical data will be described using frequencies and percentages. Continuous measures will be described using means and SDs. Cronbach α will be used to calculate the internal consistency for the instruments [26]. To examine relationships between measures, χ^2 tests will be used for categorical variables, paired *t* tests will be used to compare pre-post measures, and ordinary least squares multiple regression analyses will be conducted to examine associations between demographic variables and outcome measures. Missing data will be minimized using stop measures including using Qualtrics to prevent skipping questions and applying case deletion for surveys with missing entries that are important to answer the primary research question.

NVivo (version 15; Lumivero) will be used to analyze qualitative data collected through semistructured interviews. Data will be transcribed, deidentified, and coded for analysis [27]. To ensure rigor, we will work with another member of the research team associated with the parent grant to independently analyze the transcripts. Findings from the 2 independent analyses will be reviewed and refined until agreement is reached [28]. Findings will be reviewed for accuracy of data interpretation, themes generated, and conclusions drawn.

Results

Quantitative and qualitative results will be triangulated. The qualitative debriefing interviews will allow us to expand on implementation and practical challenges. We will also identify areas for improvement regarding the intervention logistics and patient recommendations in preparation for a larger study. A challenge of this novel approach is the limited precedent for integrating a CHW to supplement SCD specialty care. However, this study is expected to yield important preliminary data for developing an optimally effective intervention to be tested in a future feasibility study. As of March 2026, we have enrolled 16 participants.

Discussion

This research protocol outlines a novel approach to enhancing care coordination and self-care management for adults with SCD through the integration of CHWs into a PCCC team. The methodology draws inspiration from established models for chronic disease management and adapts them to the specific needs of adults with SCD in Georgia, a state with a high prevalence of SCD [11].

The proposed intervention leverages the known effectiveness of CHW support in managing chronic conditions, particularly among racial and ethnic minority groups who experience health care disparities [13]. Similar to studies with a focus on improving adherence to hydroxyurea among adolescents with SCD [29,30], our protocol recognizes the value of community-based approaches to address barriers to care. The use of a CHW as a liaison between patients and providers is a key element of this approach, aiming to bridge gaps in fragmented care and foster productive interactions, as highlighted by the CCM [19].

While previous studies have demonstrated the effectiveness of CHWs in various chronic disease contexts, including preliminary work in SCD [16,29], there remains a need for empirical evaluation of integrated PCCC-CHW models specifically for African American adults with SCD. The inclusion of semistructured interviews will offer qualitative insights into the

XSL•FO
RenderX

acceptability of the intervention from the perspectives of patients and the care team, which is a vital step toward developing sustainable and patient-centered programs. Additionally, this study supports the standardization of CHW training and certification for SCD management.

This study acknowledges potential limitations, including the feasibility of the sample size, which, while appropriate for evaluating the protocol's implementation, may limit statistical power to detect significant differences in outcomes. Larger studies using similar protocols would provide concrete data to demonstrate that CHWs are essential contributors to interdisciplinary teams that care for patients with a broad spectrum of chronic diseases. However, the multimethod approach to data collection, incorporating both quantitative measures and qualitative feedback, is a strength that will provide a comprehensive understanding of the intervention's impact and inform necessary refinements for future research. The focus on adults is particularly relevant given the unique challenges this population faces in accessing consistent and specialized care, especially in areas with geographic disparities [31].

Funding

This research protocol was funded by the National Heart, Lung, and Blood Institute as part of a diversity supplement (3U01HL134004-06S1).

Disclaimer

The authors declare that artificial intelligence was not used in the preparation of this manuscript.

Authors' Contributions

Visualization: TM-G, JT, DU, TO
Conceptualization: TM-G, JT, AK
Funding acquisition: TM-G, JT, AK
Methodology: TM-G, JT, AK
Project administration: TM-G
Writing—original draft: TM-G, JT, DU, TO
Writing—review and editing: TM-G, JT, DU, TO

Conflicts of Interest

None declared.

Multimedia Appendix 1
Integrated intervention measures.
[DOCX File, 11 KB - resprot_v15i1e82663_app1.docx]

References

1. Brousseau DC, Panepinto JA, Nimmer M, Hoffmann RG. The number of people with sickle-cell disease in the United States: national and state estimates. Am J Hematol 2010 Jan;85(1):77-78. [doi: 10.1002/ajh.21570] [Medline: 20029951]
2. Data and statistics on sickle cell disease. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/sickle-cell/data/index.html [accessed 2026-03-11]
3. Asnani MR, Fraser R, Lewis NA, Reid ME. Depression and loneliness in Jamaicans with sickle cell disease. BMC Psychiatry 2010 Jun 7;10:40. [doi: 10.1186/1471-244X-10-40] [Medline: 20529274]
4. Brousseau DC, Owens PL, Mosso AL, Panepinto JA, Steiner CA. Acute care utilization and rehospitalizations for sickle cell disease. JAMA 2010 Apr 7;303(13):1288-1294. [doi: 10.1001/jama.2010.378] [Medline: 20371788]
5. Crosby LE, Quinn CT, Kalinyak KA. A biopsychosocial model for the management of patients with sickle-cell disease transitioning to adult medical care. Adv Ther 2015 Apr;32(4):293-305. [doi: 10.1007/s12325-015-0197-1] [Medline: 25832469]
6. Cronin RM, Mayo-Gamble TL, Stimpson SJ, et al. Adapting medical guidelines to be patient-centered using a patient-driven process for individuals with sickle cell disease and their caregivers. BMC Hematol 2018;18:12. [doi: 10.1186/s12878-018-0106-3] [Medline: 29977566]
7. Evidence-based management of sickle cell disease. National Heart, Lung, and Blood Institute. 2014. URL: https://www.nhlbi.nih.gov/health-topics/evidence-based-management-sickle-cell-disease [accessed 2026-03-23]
8. Crosby LE, Hood A, Kidwell K, et al. Improving self-management in adolescents with sickle cell disease. Pediatr Blood Cancer 2020 Oct;67(10):e28492. [doi: 10.1002/pbc.28492] [Medline: 32697889]
9. Raphael JL, Rattler TL, Kowalkowski MA, Brousseau DC, Mueller BU, Giordano TP. Association of care in a medical home and health care utilization among children with sickle cell disease. J Natl Med Assoc 2013;105(2):157-165. [doi: 10.1016/s0027-9684(15)30109-7] [Medline: 24079216]

XSL•FO
RenderX

10. Rattler TL, Walder AM, Feng H, Raphael JL. Care coordination for children with sickle cell disease: a longitudinal study of parent perspectives and acute care utilization. Am J Prev Med 2016 Jul;51(1 Suppl 1):S55-S61. [doi: 10.1016/j.amepre.2016.01.023] [Medline: 27320467]
11. Snyder AB, Lakshmanan S, Hulihan MM, et al. Surveillance for sickle cell disease - sickle cell data collection program, two states, 2004-2018. MMWR Surveill Summ 2022 Oct 7;71(9):1-18. [doi: 10.15585/mmwr.ss7109a1] [Medline: 36201430]
12. Georgia Senate Study Committee on Sickle Cell Anemia. The final report of the senate study committee on sickle cell anemia. 2021 URL: https://www.senate.ga.gov/committees/Documents/sicklecell_finalreport21.pdf [accessed 2026-03-23]
13. Viswanathan M, Kraschnewski J, Nishikawa B, et al. Outcomes of community health worker interventions. Evid Rep Technol Assess (Full Rep) 2009 Jun(181):1-144 [FREE Full text] [Medline: 20804230]
14. Pullen-Smith B, Carter-Edwards L, Leathers KH. Community health ambassadors: a model for engaging community leaders to promote better health in North Carolina. J Public Health Manag Pract 2008 Nov;14 Suppl:S73-S81. [doi: 10.1097/01.PHH.0000338391.90059.16] [Medline: 18843243]
15. National community health worker training. Sickle Cell Disease Association of America. 2026. URL: https://chwtraining.sicklecelldisease.org [accessed 2026-03-11]
16. Hsu LL, Green NS, Donnell Ivy E, et al. Community health workers as support for sickle cell care. Am J Prev Med 2016 Jul;51(1 Suppl 1):S87-S98. [doi: 10.1016/j.amepre.2016.01.016] [Medline: 27320471]
17. Friedman A, Howard J, Shaw EK, Cohen DJ, Shahidi L, Ferrante JM. Facilitators and barriers to care coordination in patient-centered medical homes (PCMHs) from coordinators' perspectives. J Am Board Fam Med 2016;29(1):90-101. [doi: 10.3122/jabfm.2016.01.150175] [Medline: 26769881]
18. Rogers EA, Manser ST, Cleary J, Joseph AM, Harwood EM, Call KT. Integrating community health workers into medical homes. Ann Fam Med 2018 Jan;16(1):14-20. [doi: 10.1370/afm.2171] [Medline: 29311170]
19. Garland-Baird L, Fraser K. Conceptualization of the chronic care model: implications for home care case manager practice. Home Healthc Now 2018;36(6):379-385. [doi: 10.1097/NHH.0000000000000699] [Medline: 30383597]
20. Jenerette CM, Murdaugh C. Testing the theory of self-care management for sickle cell disease. Res Nurs Health 2008 Aug;31(4):355-369. [doi: 10.1002/nur.20261] [Medline: 18247376]
21. Mayo-Gamble TL, Murry VM, Cunningham-Erves J, et al. Engaging individuals with sickle cell disease in patient-centered outcomes research: a community health ambassador training model. J Health Care Poor Underserved 2020;31(1):353-369. [doi: 10.1353/hpu.2020.0027] [Medline: 32037336]
22. Thalheimer W, Cook S. How to Calculate Effect Sizes From Published Research: a Simplified Methodology. : Work-Learning Research; 2009 URL: https://www.paulogentil.com/pdf/How%20to%20calculate%20effect%20sizes%20from%20published%20research%20-%20a%20simplified%20methodology.pdf [accessed 2026-03-19]
23. Quasie-Woode DP, Cunningham-Erves J, Mayo-Gamble TL. The use of religion in the management of depression in sickle cell disease: a systematic review. J Relig Health 2020 Dec;59(6):3110-3125. [doi: 10.1007/s10943-020-01039-y] [Medline: 32447507]
24. Podder V, Lew V, Ghassemzadeh S. SOAP notes. In: StatPearls: StatPearls Publishing LLC; 2025.
25. Marshall GN, Hays RD. The Patient Satisfaction Questionnaire Short Form (PSQ-18). : RAND Corporation; 1994 URL: https://www.rand.org/content/dam/rand/pubs/papers/2006/P7865.pdf [accessed 2026-03-23]
26. Sousa VD, Zauszniewski JA, Bergquist-Beringer S, Musil CM, Neese JB, Jaber AF. Reliability, validity and factor structure of the Appraisal of Self-Care Agency Scale-Revised (ASAS-R). J Eval Clin Pract 2010 Dec;16(6):1031-1040. [doi: 10.1111/j.1365-2753.2009.01242.x] [Medline: 20626539]
27. Santos JA. Cronbach's alpha: a tool for assessing the reliability of scales. J Ext 1999;37(2):1-5 [FREE Full text]
28. Saldan˜a J. The Coding Manual for Qualitative Researchers / Johnny Saldan˜a, 3rd edition: London: SAGE; 2016.
29. Green NS, Manwani D, Matos S, et al. Randomized feasibility trial to improve hydroxyurea adherence in youth ages 10-18 years through community health workers: the HABIT study. Pediatr Blood Cancer 2017 Dec;64(12):12. [doi: 10.1002/pbc.26689] [Medline: 28643377]
30. Buser JM, Bakari A, Moyer CA. Feasibility of establishing a network of community health workers to support care of people with sickle cell disease in Kumasi, Ghana. J Community Genet 2021 Jan;12(1):155-161. [doi: 10.1007/s12687-020-00501-4] [Medline: 33403556]
31. Telfair J, Haque A, Etienne M, Tang S, Strasser S. Rural/urban differences in access to and utilization of services among people in Alabama with sickle cell disease. Public Health Rep 2003;118(1):27-36. [doi: 10.1093/phr/118.1.27] [Medline: 12604762]

Abbreviations

CCM: chronic care model
CHW: community health worker
HIPAA: Health Insurance Portability and Accountability Act
MTS: My Three Sicklers Sickle Cell Foundation Inc

XSL•FO
RenderX

PCCC: patient-centered coordinated care
PI: principal investigator
PSQ-18: Patient Satisfaction Questionnaire Short Form
SCD: sickle cell disease
SOAP: subjective, objective, assessment, and plan

Edited by J Sarvestan; submitted 19.Aug.2025; peer-reviewed by J Azai; revised version received 08.Feb.2026; accepted 17.Feb.2026; published 26.Mar.2026.

Please cite as:
Mayo-Gamble T, Ugwu D, Oloyede T, Telfair J, Kutlar A
Community Health Worker–Integrated Model for Sickle Cell Disease Management: Protocol for a Feasibility Study
JMIR Res Protoc 2026;15:e82663
URL: https://www.researchprotocols.org/2026/1/e82663
doi:10.2196/82663

© Tilicia Mayo-Gamble, Doreen Ugwu, Tobi Oloyede, Joseph Telfair, Abdullah Kutlar. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Measuring Accuracy (Classification Probabilities, Positive, and Negative Predictive Values) of Executive Function Electroencephalogram Metrics in Attention-Deficit/Hyperactivity Disorder Diagnosis: Protocol for and Perspectives From the SINCRONIA Study

Hilario Blasco-Fontecilla[1,2,3], MD, PhD; Javier Sánchez-Cerezo[4], MD, PhD; Irene Gómez[5], MSc; Georgelina Abreu-Fernández[5], PhD; Sandra Ortiz[5], MSc; Jesús F Villoria[6], MD; Miguel Blanco[5], MSc; Ana García[5], MSc; Julia Ballesteros[5], MSc; Roldán Martínez[5], MSc; Gerardo Gálvez[5], PhD; Fernando Maestú[7], PhD; Álvaro López-Medrano[5], MSc

[1]Department of Psychiatry, Emooti – Institute of Mental Health and Wellbeing, Calle José Abascal 51, Madrid, Spain

[2]Instituto de Transferencia e Investigación (ITEI), Universidad Internacional De La Rioja (UNIR), Logroño, Spain

[3]Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, Spain

[4]Department of Psychiatry, Puerta de Hierro University Hospital, Majadahonda, Spain

[5]Bitsphi Diagnosis, Madrid, Spain

[6]Medicxact, Alpedrete, Spain

[7]Department of Experimental Psychology, Complutense University of Madrid, Madrid, Spain

Corresponding Author:
Hilario Blasco-Fontecilla, MD, PhD
Department of Psychiatry, Emooti – Institute of Mental Health and Wellbeing, Calle José Abascal 51, Madrid, Spain

Abstract

Background: Attention deficit/hyperactivity disorder (ADHD) is the most prevalent neurodevelopmental disorder worldwide, affecting approximately 5% - 7% of school-aged children and 2% - 5% of adults worldwide. However, there is still no reliable diagnostic tool for it. The lack of specific biomarkers further complicates the accurate diagnosis of ADHD.

Objective: The SINCRONIA study seeks to develop and optimize an electroencephalogram (EEG)-based ADHD diagnostic classification algorithm by identifying biomarkers that provide optimal diagnostic performance.

Methods: This protocol introduces a single-center, case-control study involving at least 165 participants, aged between 7 and 12 years, that is being conducted at the Puerta de Hierro University Hospital in Madrid, Spain. Participants will be allocated to 3 groups, including ADHD predominantly inattentive, ADHD predominantly combined or hyperactive/impulsive, and a control group, according to the best estimated diagnosis based on clinical interviews and a neuropsychological assessment that includes the Conners Continuous Performance Test. In addition, an EEG recording will be conducted separately, and functional connectivity metrics will be used to characterize brain networks associated with inhibitory control. The index test is expected to match or improve the clinical diagnosis of ADHD in children aged between 7 and 12 years and provide a set of eventual biomarkers that maximize diagnostic performance and provide pathophysiological clues.

Results: The SINCRONIA study began screening and recruitment in March 2023. Recruitment ended on December 11, 2024. A total of 165 eligible participants were enrolled.

Conclusions: The SINCRONIA project is a high-quality, large-scale, unicenter study devoted to improving the objective diagnosis of ADHD by using EEG biomarkers. The EEG-based ADHD diagnosis is expected to have greater sensitivity and specificity than the Conners Continuous Performance Test.

Trial Registration: ISRCTN Registry ISRCTN12110752; https://doi.org/10.1186/ISRCTN12110752

International Registered Report Identifier (IRRID): DERR1-10.2196/79150

(JMIR Res Protoc 2026;15:e79150) doi:10.2196/79150

XSL•FO
RenderX

KEYWORDS

attention deficit hyperactivity disorder; ADHD; electroencephalogram; EEG; functional connectivity; machine learning; neural network; neurocognitive assessment; diagnosis; ADHD comorbidities.

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder characterized by inappropriate levels of inattention, hyperactivity, or impulsivity, typically emerging in childhood with a relatively high prevalence [1]. Reliable estimates place the prevalence above 5% in children and adolescents and 2.5% in adults [2-5]. However, these rates vary by country, with the prevalence reaching 11.4% in American children aged 3-17 years in 2022, according to the Centers for Disease Control and Prevention [6].

ADHD is linked to neurocognitive deficits and impairments in psychosocial and vocational functioning, including increased mortality and suicide rates [7-12]. Neurocognitive deficits span a broad range of functions, including working memory, inhibitory control, and vigilance [13], as well as imbalances between immediate and delayed rewards [14], processing speed [15], and motor activation and control [16]. Furthermore, ADHD often co-occurs with other psychiatric disorders such as generalized anxiety, major depression, bipolar disorder, substance use, and conduct disorders; or with other neurodevelopmental disorders [17,18].

Accurate diagnosis is crucial for effective treatment planning and improving clinical research. Efforts to refine psychiatric nosology by exploring connections between neurobiology and psychopathology are underway [19-21] with a key advancement being the identification of biomarkers for psychiatric disorders [22,23]. Advances in understanding the pathophysiology of ADHD have highlighted alterations in noradrenergic signaling within posterior attentional circuits [24], disruptions in dopaminergic pathways across various systems [25], dysfunction of the default mode network [26], and delayed development of the frontal cortex. Despite these insights, little progress has been made in classification and diagnosis, as reflected by the fact that neither the *DSM-5* (*Diagnostic and Statistical Manual of Mental Disorders* [Fifth Edition]) nor the *ICD-11* (*International Classification of Diseases, 11th Revision)* grants a significant role to neurobiology.

Unfortunately, ADHD diagnosis remains inaccurate, partly due to its multifactorial etiology, psychiatric comorbidities, and neurocognitive and brain alterations. Although diagnosis is based on a comprehensive evaluation of current and past symptoms and functional performance, overlapping symptoms with other disorders—such as mood or learning disorders—and frequent co-occurrence with conditions like autism spectrum disorder (ASD), learning disabilities, personality disorders, or substance use further complicate differential diagnosis [27-31]. While some authors contend that ADHD has been overdiagnosed [32-35], others dispute this view [29,36]. Additionally, assessing the widespread impact of ADHD on various life areas with limited consultation time presents a challenge, requiring evaluation of multiple contexts [30].

ADHD is a dynamic condition where slowly developing neurobiological deficits are rapidly influenced by maturation and compensatory mechanisms, resulting in significant heterogeneity both across individuals and within the same individual over time, thereby complicating both diagnosis and treatment [37]. This heterogeneity has driven increased neuroimaging research and the search for biomarkers [38]. Functional magnetic resonance imaging (fMRI) studies have revealed delayed cortical development—especially in the frontal, temporal, and parietal regions—and disruptions in connectivity and default mode network functionality [39,40]. Moreover, deficits in working memory and inhibitory control have been linked to underactivation in frontoparietal and ventral attention circuits, which play key roles in executive function and attention reorientation [41]. Additionally, reduced activation in the ventral striatum has been associated with altered reward processing [42]. Research also indicates a loss of the typical counter-regulation between cognitive control and default mode networks during rest [36], with some studies suggesting compensatory hyperactivation in motor and visual areas [43]. Importantly, much of the observed heterogeneity may stem from the fact that ADHD subtypes or syndrome-level distinctions are often collapsed into single-group analyses, obscuring meaningful neurophysiological differences. Studies that stratify participants by ADHD subtype have begun to reveal distinct neural signatures and developmental trajectories, which may prove crucial for biomarker discovery and personalized intervention design [44].

These findings underscore the complexity of diagnosing and understanding attention disorders, highlighting the need for more precise and scalable diagnostic tools. In this context, electroencephalogram (EEG)-based techniques have gained increasing interest due to their ability to track dynamic interactions across large-scale brain networks with millisecond temporal resolution [45,46]. Unlike fMRI, EEG can directly measure neural activity and is better suited to capture transient cognitive processes, such as inhibition and attentional shifts, key domains affected in ADHD. Source-level EEG analyses further enhance spatial resolution, making it a practical and cost-effective alternative to higher-end techniques like magnetoencephalography in clinical research [47]. When applied to stratified samples reflecting distinct ADHD subtypes, EEG may help reveal subtype-specific connectivity patterns that are obscured in traditional binary comparisons. This supports its growing potential not only as a complementary biomarker platform but also as a tool for advancing precision diagnostics in neurodevelopmental disorders.

The SINCRONIA study builds on the 2 attention-control networks model developed by Corbetta et al [48]. The dorsal attention network (DAN) exerts top-down control by maintaining focus on task-relevant information while filtering out distractions, thus ensuring sustained attention. In contrast, the ventral attention network (VAN) evaluates each incoming stimulus based on its novelty and importance. When a stimulus

XSL•FO
RenderX

is deemed high priority, the VAN activates—interrupting the DAN's filtering process—and sends signals to trigger executive control, reorienting attention and prompting the appropriate response [49-51]. Conversely, if the stimulus is considered of low priority, the VAN deactivates, and the DAN's inhibitory influence is reinforced, preventing unnecessary shifts in attention [48,52,53].

While structural brain alterations in ADHD are acknowledged [54], functional changes—reflecting the interactions within neural circuits—hold greater diagnostic relevance [30]. Although EEG's diagnostic utility at the time of this writing is limited, its high temporal resolution, accessibility, and low cost support its potential in developing tools and identifying biomarkers [55,56]. Capturing transient neural dynamics has shown promise [57], and using multivariate methods with machine learning (ML) significantly enhances accuracy, especially when incorporating signal synchronization metrics [55]. However, existing EEG studies for ADHD detection remain suboptimal as diagnostic tools [58,59]. For all these reasons, the SINCRONIA study aims to develop an ML-based classification algorithm using EEG brain connectivity metrics and to determine the biomarkers that maximize the diagnostic performance on the Bit task (hereafter referred to as the index test).

The primary objectives of our study are:

1. To demonstrate that the diagnostic performance of the index test (EEG-based)—determined against the final diagnostic judgment of the study's principal investigator (PI) following a comprehensive clinical assessment—exceeds a predefined minimum threshold of clinical acceptability.
2. To identify the biomarkers that optimize the diagnostic accuracy of the index test.

Secondary objectives are the following:

1. To develop and optimize a diagnostic classification algorithm based on selected biomarkers—identified in accordance with the primary objectives—to differentiate between controls, ADHD predominantly inattentive (without hyperactivity), and ADHD combined or hyperactive/impulsive (with hyperactivity) through classification models.
2. Investigation of factors influencing diagnostic performance, including patient demographics and clinical characteristics of ADHD, as well as health care provider factors (eg, specialty and experience).
3. To evaluate and compare the diagnostic performance of the index test—determined against the final diagnostic judgment of the study's PI following a comprehensive clinical assessment—with that of the Continuous Performance Test Third Edition (CPT-3).
4. Assessment of the clinical utility and value of the index test.

Tertiary objectives are the following:

1. To investigate the construct validity of sluggish cognitive tempo (SCT) in comparison to ADHD diagnosis.
2. To optimize the preprocessing automatic system for artifact removal in EEG signals.
3. To explore Bayesian mathematical models and their application to brain functional connectivity patterns.

Methods

Study Design

This is a noninterventional, case-control, cross-sectional, and assessor-blinded study, which is part of the development of a medical diagnostic test. This manuscript describes the latest version of the study protocol (HUPH-BP-BMCCEEG-CC22 version 4.0, approved on June 7, 2024, as Amendment 4), which was reviewed and approved by the appropriate ethics committee (see below). In order to allow for more complete and detailed reporting, we adhere to 2 acceptable reporting guidelines that match the study design: the Standards for Reporting Diagnostic Accuracy Studies (STARD; Checklist 2) and the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2025 checklist (Checklist 1) of items to address in a randomized trial protocol, though, *sensu stricto*, our study is not a randomized trial. Following the terminology of the STARD, the EEG and the algorithms used to analyze the EEG data will be regarded as the index test, and the executive control task during which brain connectivity metrics will be pursued will be regarded as the index task. The best estimated diagnosis based on clinical interviews and neuropsychological assessment will be regarded as the reference test. To minimize bias, the PI establishing the reference standard will not perform EEG analyses or access EEG results. Blinding to index test results will be maintained, as no diagnostic or therapeutic decisions will be based on these findings. Consequently, no clinical decision will be influenced by information from the index test. Likewise, the team in charge of EEG processing will not have access to the clinical diagnosis data during data collection for the study. Participants will be categorized into 3 groups, including ADHD predominantly inattentive, ADHD combined or hyperactive/impulsive, and a control group, based on clinical evaluation including the reference CPT-3.

Participants

A total of at least 165 participants will be recruited (Multimedia Appendix 1). The target population will consist of children who exhibit symptoms and signs suggesting ADHD who contact the health system and who may or may not have the disorder. It will be attempted to represent this population through a source population consisting of (1) children aged 7-12 years who either received a prior diagnosis of ADHD (cases; predominantly hyperactive/impulsive or combined presentation or predominantly inattentive) or (2) children attending routine pediatric check-ups at the Puerta de Hierro University Hospital in Majadahonda (HUPHM) who do not meet criteria for any significant psychiatric disorder (controls). Recruitment ended in December 2024. Accordingly, our study was a convenience series of participants.

A data monitoring committee was not considered necessary due to the low-risk nature of the study; oversight is ensured through standard monitoring procedures.

XSL•FO
RenderX

Instruments

Tanner Scale

The Tanner sexual maturity scale will be administered to assess the degree of development and physiological status in participants [60].

Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ) is a brief screening questionnaire that helps detect pediatric psychopathology in individuals between 4 and 16 years old, with measures of social, emotional, and behavioral functioning. It consists of 5 scales (emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior) and provides a total score of difficulties. Due to its use in scientific literature and the demonstration of its good psychometric properties, this scale has been used to rule out relevant psychiatric pathology that could interfere with the study [61,62].

Child and Adolescent Behavior Inventory

The Child and Adolescent Behavior Inventory (CABI) consists of 9 subscales that allow measuring symptoms of inattention, hyperactivity, SCT, anxiety, depression, oppositional defiant disorder, prosocial attitude, and academic and social difficulties. The parent version will be used for this study. Each scale is made up of several items, and it must be specified whether the absence or presence of such symptoms has an impact on an academic or social level (no difficulty, mild difficulty, moderate difficulty, or severe difficulty). We will use the most recent normative data that have been published for Spanish youth aged 5 - 16 years [63].

ADHD Rating Scale IV

The validated Spanish version of the Attention Deficit/Hyperactivity Disorder Rating Scale IV (ADHD-RS-IV) will be used to check the severity of ADHD, considering the total score of the scale—composed of 18 items that are scored according to a 4-point Likert scale (0=never or rarely, 1=sometimes, 2=often, or 3=very often)—and the scores of the 2 subscales of the test, inattention and hyperactivity/impulsivity [64].

Wechsler Intelligence Scale for Children, Fifth Edition

The Wechsler Intelligence Scale for Children, Fifth Edition (WISC-V) allows for the assessment of the total IQ and consists of 5 main indices (verbal, visuospatial, fluid reasoning, working memory, and processing speed) that are measured from 10 tests. Secondary indices (quantitative, auditory working memory, nonverbal, general ability, and cognitive competence) will also be included [65].

Continuous Performance Test

The CPT-3 assesses attention, impulsiveness, and processing speed through a go/no-go interference task where the participant must press the space bar on the keyboard to all letters that appear on a computer screen, except for the letter X (nontarget). The letter presentation interval is 1, 2, and 4 seconds, and the test measures omission errors (unanswered targets), commission errors (answered nontargets), hit reaction time, and response speed variability. The test lasts for 14 minutes [66]. The test positivity cutoff is automatically generated by the program depending on age and sex of the participant.

Neuropsychological Evaluation of Executive Functions

The Neuropsychological Evaluation of Executive Functions (ENFEN) is an assessment battery applicable to children between 6 and 12 years old that consists of 4 tests (fluency, paths, rings, and interference) that measure executive functions. The full battery will be used in this study [67].

Test for Detecting Dyslexia in Children

For this study, a screening battery consisting of 12 tests known as PROLEXIA (Emooti Neurotech SL) will be used. These tests are applicable from the age of 7 years and are designed to detect symptoms of dyslexia, particularly those related to the phonological component, aiding in differential diagnosis [68].

Child Concentration Inventory-Version 2

The Child Concentration Inventory-Version 2 (CCI-2) is a test developed to detect SCT. It is designed to be self-administered by children. It is composed of 16 items, rated on a 4-point Likert scale (0=never, 1=sometimes, 2=often, or 3=always) [69].

Index Test

The index test (also known as the Bit task) is a go/no-go interference test consisting of 3 parts, preceded by a 4-minute resting-state EEG recording. Each part of the test contains 400 trials, with an interstimulus interval between 450 and 550 milliseconds long. The ratio of go and no-go trials changes across parts, including 25:75 (first part), 50:50 (second part), and 75:25 (third part). If participants make mistakes, a warning will be displayed on the screen reminding them of the instructions. This task will be executed using PsychoPy software (version 2022.1.3; Open Science Tools Ltd).

EEG Recording System

We will use the Waveguard net (ANT Neuro) HD-C68 Tyco with Ag rings, 2 HD-C68pin connectors, and 64 semidry electrodes arranged in an equidistant layout, compatible with eego processing software (version 1.9.2 or higher; ANT Neuro). The EEG eego adapter will be used for 6 type A channels, along with the eego amplifier EE-225 (ANT Neuro BV). The reference signal is placed at the 5Z position, while the ground signal is positioned at the 0Z coordinate.

Participant Recruitment

The study will be conducted in the Department of Psychiatry at the HUPHM. Participants will be recruited from 2 distinct sources. First, cases will be identified by reviewing medical records from the hospital to locate participants with an existing diagnosis of predominantly inattentive, predominantly hyperactive/impulsive, or combined ADHD prior to the commencement of study procedures. Additionally, cases may be prospectively selected from patients routinely seen at the study site who meet the specified selection criteria. These diagnoses will be confirmed through clinical evaluation during the study reference test. Control participants will be recruited from individuals attending Pediatric Services at HUPHM.

XSL•FO
RenderX

Selected candidates will be contacted and invited to participate, with an in-person appointment scheduled to complete the study procedures. ADHD cases will be instructed to temporarily discontinue their medication for the disorder prior to the first visit, with a 24-hour cessation for those using methylphenidate or atomoxetine and 48-hour for those using lisdexamfetamine.

Study Visits

First Visit

During the first visit, the study's purpose will be thoroughly explained to participants. Following this, informed consent will be obtained from parents or guardians, and assent will be sought from the child participants. No data will be collected if the participant withdraws consent from the study. The eligibility criteria (Table 1) will be reviewed to confirm participants' suitability. Additionally, participants will be instructed to abstain from stimulant substances, such as caffeine, xanthines, or medications, 24-48 hours prior to the visit, depending on the type of treatment. If this time frame has not been adhered to, the study evaluation will be rescheduled for the following days.

Table . Eligibility criteria for participants in the SINCRONIA study.

Participant group	Inclusion		Exclusion	
	Cases (ADHD[a])	Pediatric controls (non-ADHD)	Cases (ADHD)	Pediatric controls (non-ADHD)
Agreement to participate, with written informed consent obtained from a legal representative together with their own assent	Yes	Yes	No	No
Age (years)	7-12	7-12	<7->13	<7->13
Right-handedness (recruitment of up to 10% of left-handed children is permitted)	Yes	Yes	No	No
Time schedule	ADHD diagnosis within the past 3 years or at the time of the study	Pediatric attendance within the past 3 years or at the time of the study	ADHD diagnosis more than 3 years before the time of the study	Pediatric attendance more than 3 years before the time of the study
Type of ADHD	Predominantly hyperactive/impulsive, or combined presentation, or predominantly inattentive presentation	—[b]	ADHD presentations different from those considered in the inclusion criteria	—
Neuropsychiatric comorbidities	Mild/moderate comorbid conditions (such as anxiety, dyslexia, or SCT[c]) that do not compromise participation	Mild/moderate comorbid conditions (such as anxiety, dyslexia, or SCT) that do not compromise participation	Severe comorbid conditions (such as autism or epilepsy) that may compromise participation	Severe comorbid conditions (such as autism or epilepsy) that may compromise participation
Color blindness or any other condition that may cause instrumental difficulties for sensorineural perception	No	No	Yes	Yes
Severe cognitive or behavioral problems or insufficient capacity to maintain a minimum level of relaxation and rest	No	No	Yes	Yes
Regular intake of medications other than the treatment of ADHD that may alter attentional or executive capacity	No	—	Yes	—

[a]ADHD: attention-deficit/hyperactivity disorder.

[b]Not applicable.

[c]SCT; sluggish cognitive tempo.

Sociodemographic and clinical data will be collected from eligible participants, including a physical assessment to document their maturation status (Tanner stage) and a record of all prior treatments at the time of this writing, including those for ADHD. Questionnaires will be administered to screen for relevant psychiatric disorders (SDQ) and comorbid conditions

XSL•FO
RenderX

(CCI-2 and CABI). The WISC-V will be conducted as part of the neuropsychological evaluation. If the assessment is not completed during the first visit, an additional visit will be scheduled to complete the evaluation.

Second Visit

During the second visit, the suspension of medication and abstinence from stimulant substances such as caffeine and xanthines will be confirmed with the participants. Once both have been confirmed, the child's head circumference will be measured to select the appropriate EEG cap size (S, M, or L). The cap will then be immersed in saline solution for 10 minutes before being placed on the child's head. The impedance measured in ohms will be assessed at this point.

The EEG recording will begin with a trigger check and a 4-minute eyes-closed resting phase to record brain activity at baseline, followed by the task. The instructions for the task will be introduced, with example trials before full implementation of each part of the task. Throughout the task, breaks will be provided for brief play activities and impedance checks. The EEG session will conclude with the creation of a 3D model of cap placement while the child is wearing the cap.

Third Visit

During the third visit, the suspension of medication and abstinence from stimulant substances such as caffeine and xanthines will be again verified. Once both are confirmed, the neuropsychological evaluation will proceed with the administration of the CPT-3, the ENFEN, and additional tests to screen for comorbid conditions, such as PROLEXIA.

Although the sample size calculation indicates a total of 120 participants (40 per group), an additional 10 participants per group will be recruited to explore factors influencing diagnostic performance and SCT construct validity (Figure 1). Furthermore, an extra 10% will be added to account for nonassessable data, resulting in a total of 162 participants. This total may be slightly increased to ensure 54 participants per group after accounting for screening failures and protocol deviations. The maximum time interval between the index test and the reference standard is 72 hours; however, whenever feasible, a 24-hour interval will be pursued.

Figure 1. Procedure of the study. ADHD: attention deficit/hyperactivity disorder; CPT-3: Continuous Performance Test, Third Edition; EEG: Electroencephalogram; SDQ: Strengths and Difficulties Questionnaire.

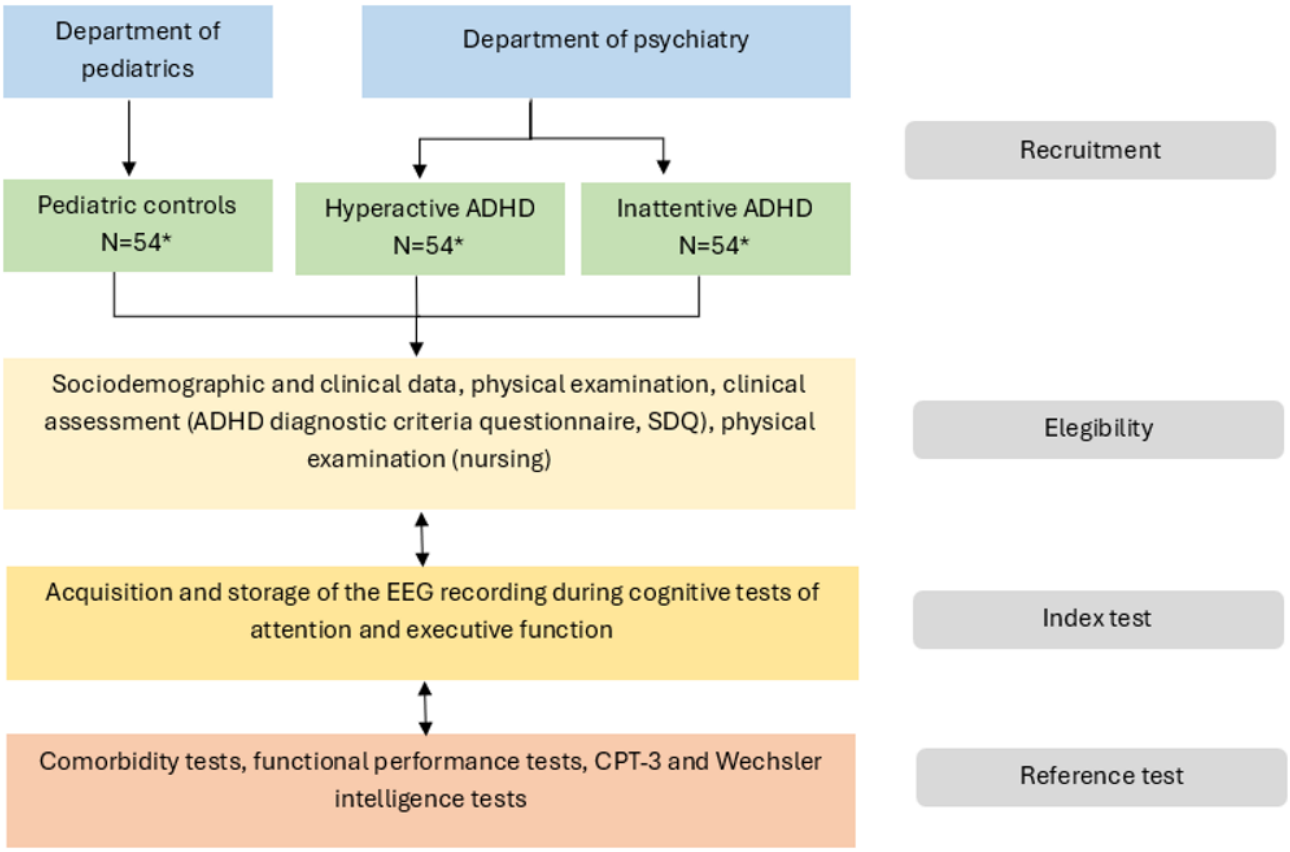

Planned Data Analysis

This clinical study focuses on the characterization and initial validation of an experimental medical diagnostic test under development. To avoid bias, the physicians involved in diagnosing and selecting the study groups will not participate in the analysis. Furthermore, missing data will not be imputed. In addition, indeterminate index test or reference standard results will not be allowed, as a positive or negative diagnosis will be required according to the previously defined cutoffs. Analyses will be performed on complete cases within each analysis set. The planned sample size (Multimedia Appendix 1) includes a margin for nonevaluable data (10% added for nonassessable data-complete-case analysis). No interim analyses or stopping rules are planned.

XSL•FO
RenderX

Hypotheses

The primary hypothesis, though involving 3 participant groups, focuses on demonstrating a minimum diagnostic performance to distinguish between the ADHD group as a whole and the pediatric control group. Regarding the true positive fraction (true positive fraction [TPF] or sensitivity), due to its high acceptability threshold (TPF_0=0.90), a noninferiority test will be used for this dimension, with a noninferiority margin of δT_0=−0.10. Since the primary objective is to determine whether the index test yields classification probabilities—TPF and false positive fraction (false positive fraction [FPF] or 1-specificity)—that are minimally acceptable (TPF_0 and FPF_0), the corresponding hypothesis, considering the noninferiority assumption for sensitivity, can be formally stated as:

(1)H0:{TPF−TPF0≤δ ∪0TFPF≥FPF0}

If H_0 can be rejected, TPF−TPF_0 > δT0 and FPF<FPF_0; therefore, the index test would be noninferior to a hypothetical test with 90% sensitivity and superior to a hypothetical test with a specificity of 1−FPF_0.

Two additional pairwise comparisons will be conducted: one between each ADHD group and pediatric controls, and another between the ADHD groups themselves, with hypotheses formulated as previously described. The Hochberg procedure will be used to control Type I error [70].

Biomarker identification will use dimensionality reduction methods, which are descriptive and do not require hypothesis testing or multiplicity adjustments.

Statistical Analysis

The primary objective will be to assess the diagnostic accuracy of the index test in terms of classification probabilities, specifically sensitivity and specificity. Sample size calculation ensures a minimum performance threshold within the target population.

This objective is interdependent with another primary aim, identifying the minimal set of biomarkers that optimize the diagnostic performance of the index test. As this objective does not involve hypothesis testing, adjustments for multiplicity are unnecessary.

The analyses of the secondary objectives will be exploratory, except for the first, which shares the hypothesis-free nature of biomarker identification. Tertiary objectives will also be exploratory, focusing on the validity of the SCT construct and EEG signal preprocessing systems. These analyses may be expanded in future studies to identify EEG-based clusters associated with ADHD symptomatology.

Analysis of Primary Parameters

The analysis of diagnostic accuracy requires preliminary analyses to identify biomarkers that optimize the diagnostic performance and to develop a classification algorithm based on biomarker values across the three study groups defined by the clinical team, serving as the reference test. The first secondary parameter, essential for evaluating the primary parameter, will also be analyzed in a confirmatory fashion without explicit hypothesis testing.

Diagnostic Performance of the Index Test

The primary parameters are the classification probabilities of the index test relative to the reference test. In addition to these classification probabilities, positive and negative predictive values (PPV and NPV, respectively) and diagnostic likelihood ratios will be calculated. Inferences will be made by calculating 95% asymmetric rectangular confidence regions for each pair of interest values (TPF and FPF), (PPV and NPV), and (DLR^+ and DLR^-). Confidence regions will be calculated from the cross-product of the 97.5% confidence intervals of each component in each pair. Sensitivity values will be assessed for noninferiority (and potentially superiority), while specificity will be evaluated for strict superiority. In addition to naive predictive value estimates, which will be biased by design, other estimates will be adjusted using external prevalence data and diagnostic likelihood ratios or Bayesian multiplication factors to mediate prior and posterior probabilities:

, (2)logit PPV=logit ρ+log DLR+

, (3)logit NPV=−logit ρ + −log DLR−

where ρ is the prevalence in the target population (7.0% [71]). Clinical information and reference standard results will not be available to the performers of the index test. Furthermore, the reference standard will be performed after the index test. Moreover, performers of the index test will neither be involved in the technological processing nor the statistical analyses of either the index or the reference standard tests. Finally, cross-tabulation of the index test results (or their distribution) by the results of the reference standard will be considered.

Ethical Considerations

This study was approved by the Research Ethics Committee of the HUPHM on February 13, 2023 (PI 207/22). Informed consent was obtained from legal guardians and assent from minors, in accordance with the provisions of Organic Law 3/2018 of December 5 [72], on the Protection of Personal Data and Guarantee of Digital Rights. All procedures involving human participants were conducted in accordance with the ethical standards of the responsible institutional and national research committee and with the 1964 Declaration of Helsinki and its subsequent amendments. Monitoring of the clinical trial was carried out as required by applicable clinical research regulations. Privacy was always guaranteed. All participants had a secret code which was dissociated and saved separately from personal identification data. All participants entered into the study without compensation. All data generated or analyzed will be fully anonymized to ensure participant confidentiality (ISRCTN Registry ISRCTN12110752).

Risks and Adverse Events

Any adverse events from the EEG recordings, index task, and neuropsychological tests will be recorded according to the clinical guidelines established in the legal framework. Adverse events will be recorded from the first visit to the last. The researcher (HB-F) will assess the association between the events and the assessment and report this to the research ethics committee.

XSL•FO
RenderX

Data Management

Data will be collected using a pre-established standardized variable form and entered in a password-protected electronic database with role-based access control. Data quality will be ensured through predefined regular monitoring and periodic audits to identify missing or inconsistent entries. Access to the final dataset will be restricted to authorized members of the research team.

Results

The SINCRONIA study began screening and recruitment in March 2023. Recruitment ended on December 11, 2024. A total of 165 eligible participants were enrolled. This research was funded by the sponsor, Bitsphi Diagnosis. The SINCRONIA study was registered in the ISRCTN Registry, a primary clinical study registry recognized by the World Health Organization (Trial registration: ISRCTN12110752; February 20, 2025).

Discussion

Comparison With Prior Work

EEG has been extensively employed to explore various neurophysiological metrics in ADHD, such as power spectral densities and event-related potentials. For instance, Clarke et al [73] demonstrated that children with ADHD exhibit increased theta band activity, decreased alpha band activity, and an elevated theta/beta ratio (TBR) compared to controls. Furthermore, these differences persisted into adulthood, with individuals showing lower beta levels in specific regions and a reduction in the ratio in ADHD compared to controls in recordings with eyes closed [74]. Similarly, Jarret et al [75] found increased relative theta power in the T5 and F7 regions in individuals with ADHD, aligning with findings from Chutko et al [76], who observed elevated TBR in children with both ADHD and SCT. Despite these promising findings, the utility of metrics like TBR as a robust diagnostic tool has been questioned due to inconsistent results [77,78], although in recent years researchers have increasingly favored the use of TBR in combination with Bayesian Gaussian models [79].

Although EEG power metrics offer valuable insights, recent studies have shifted focus toward functional connectivity analyses, which may serve as more stable and reliable biomarkers for ADHD. McNorgan et al [80] identified impairments in functional connectivity within the default mode network, reinforcing the hypothesis that ADHD involves disrupted brain networks underlying attention and executive control. Michelini et al [81] further highlighted abnormalities in the corpus callosum and connectivity between the default mode network and frontoparietal/salience networks, suggesting that ADHD is fundamentally a disorder of brain connectivity and providing critical insights by examining brain connectivity across controls, individuals with persistent ADHD, and those in remission. Their findings revealed that individuals with ADHD exhibited less pronounced increases in theta band connectivity following cognitive control tasks, indicating deficits in modulating connectivity during transitions from rest to task-oriented states. Additionally, during the Flanker task, those with ADHD demonstrated global hyperconnectivity in theta, alpha, and beta bands when the task's prestimulus cues were incongruent and only showed hyperconnectivity in the beta band during the processing of target stimuli when the stimuli were congruent, which correlated with poorer cognitive control and dysfunctional behavior [81]. These findings are consistent with previous research showing enhanced connectivity in the alpha and beta bands during attention tasks [36,82].

Despite recent advancements, significant limitations persist in the literature. Many studies rely on small sample sizes and predominantly male samples or fail to report sex distribution, limiting the generalizability of their findings [81,83,84]. Furthermore, individual biomarkers often explain only a small fraction of the variance in treatment responses, highlighting the need for multifactorial and integrative approaches.

In addition, several methodological issues remain unaddressed in previous studies. Most existing tasks used for biomarker discovery are not grounded in theoretical models of attentional functioning and do not incorporate inhibition as a core element of the pathology, despite its central role in ADHD-related executive dysfunction [7]. These studies also tend to overlook functional connectivity at the source level and instead rely on scalp-level analyses, which limits the spatial precision and neural interpretability of their findings [45]. Moreover, EEG-based investigations are rarely validated against higher spatial resolution techniques such as magnetoencephalography, which could serve as a useful benchmark for signal source validation [47]. Finally, many studies use binary group comparisons (ADHD vs controls), neglecting the added value of including subtypes or differential diagnoses such as SCT and dyslexia in a 3-group design, which would allow for more clinically informative models [44,85].

Our study targets to address these gaps by using functional connectivity patterns, derived from EEG signals in the frequency domain, as biomarkers for ADHD. In this study, which will include at least 165 participants, we seek to improve the diagnostic accuracy achieved in the clinic at the time of this writing and to differentiate ADHD subtypes or other comorbid conditions that might be potentially confused with ADHD, using EEG-based ML algorithms applied to a go/no-go task. This study's strength lies in its comprehensive design and robust sample size, enabling the identification of ADHD with high accuracy.

By integrating EEG connectivity metrics with ML algorithms, we aim to uncover reliable biomarkers for ADHD, advancing the field and contributing to a deeper understanding of this neurodevelopmental disorder.

Strengths and Limitations

The study included primarily right-handed children (10% left-handed) aged 7 - 12 years who could suspend medication and tolerate EEG recording sessions. The inclusion of left-handed children, up to 10% of eligible participants (Table 1), partly addressed this limitation within the study. Despite excluding children with ASD or other comorbidities, the sample of this study is fairly representative of most ADHD populations. However, the team is planning future studies to examine

XSL•FO
RenderX

biomarker performance in more heterogeneous, clinically representative ADHD populations, including children with common comorbidities, such as ASD.

Another limitation was that the final clinical diagnosis made by the PI was used as the reference test, as well as to define the participant grouping for algorithm training, thus introducing some circularity that could inflate diagnostic accuracy. However, several measures were introduced in order to mitigate this issue: first, the PI established the diagnosis independently, following established clinical criteria and without knowledge of any algorithmic or EEG outputs, thereby reducing the risk of circularity; second, several patients were previously diagnosed by other professionals, and the PI only confirmed the ADHD diagnosis; and third, a contract research organization was hired and an external neuropsychologist independently evaluated all patients.

Despite the aforementioned limitations, the study also presents notable strengths. First, it benefits from a large sample size, which enhances the reliability of the findings and supports the generalization of the results and allows for comparisons between ADHD subtype groups. Additionally, it incorporates a comprehensive neuropsychological assessment that enables a detailed examination of participants' cognitive profiles. This approach allows for a more accurate identification of the specific neurocognitive processes affected in children with ADHD, providing a robust dataset that can inform future research and support the exploration of other clinical and neurodevelopmental aspects of the disorder across different ages.

Acknowledgments

The authors would like to express their gratitude to Ricardo Bruña, PhD, for his insights and contributions to electroencephalography (EEG) signal analysis. Additionally, the authors acknowledge the financial support that has made this research possible. Generative artificial intelligence (AI) was not used for the creation of this manuscript.

Funding

Bitsphi Diagnosis funded this study as sponsor. Bitsphi Diagnosis can be contacted through the CTO, Miguel Blanco (email: info@bitsphi.com).

Data Availability

The datasets generated and analyzed during this study are not publicly available due to data protection regulations. The full study protocol may be shared upon request.

Authors' Contributions

Conceptualization: AL-M, MB, FM, HB-F, SO.
Funding Acquisition: AL-M, MB.
Investigation: HB-F, JS-C, IGG, GA-F.
Methodology: JFV, AL-M, MB, SO, FM, RM, and JB.
Project administration: AL-M, MB, and AG.
Resources: AL-M and MB.
Software: MB, GG, RM, and JB
Supervision: HB-F, JS-C, SO, and AG.
Writing—original draft: SO, IG, and GA-F.
Writing—review and editing: all authors.
All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

GA-F, SO, IGG, MB, AG, JB, RM, GG, and ALP are either employers or owners of the company, which may be considered a competing interest. HBF, GA-F, and IGG received compensation from the study sponsor for conducting the Clinical Trial. HBF, GA-F, and IGG received, or entered into an agreement to receive, payments, grants, and/or equipment from sponsors that have a cumulative monetary value of US $25,000 or more, during the time they were carrying out the Clinical Trial and for one year after completion of the Clinical Trial. HBF has received lecture fees from Takeda, BIAL, Rubio Laboratories, and Rovi laboratoriesLaboratories. He has also been granted three prizes regarding the development of a serious videogame video game for treating ADHD (The Secret Trail of Moon): the Shibuya Prize by Takeda; the first prize of the college of psychologists of Madrid; and a prize to the best innovative health initiative within the healthstart prize. He was Principal Investigator (PI) of an iPFIS research contract ([85]; IFI16/00039), co-PI of a MINECO research grant (RTI2018-101857-B-I00), and PI of the SINCRONIA study funded by Bitsphi Diagnosis (this study). Bitsphi Diagnosis has two publicly available patent applications: one in Europe (EP4335370A1) and one in the United States (US20240081719A1). Both have been published and are pending at the time of this writing. Corresponding applications to the European filing have also been submitted in China (CN 202380064709.5), Japan (JP 2025- ‐ 511499), and India (IN 202547020286).

XSL•FO
RenderX

Multimedia Appendix 1
Sample size determination.
[DOCX File, 17 KB - resprot_v15i1e79150_app1.docx]

Checklist 1
SPIRIT checklist.
[PDF File, 135 KB - resprot_v15i1e79150_app2.pdf]

Checklist 2
STARD checklist.
[PDF File, 756 KB - resprot_v15i1e79150_app3.pdf]

References

1. Global Burden of Disease Child and Adolescent Health Collaboration, Kassebaum N, Kyu HH, et al. Child and adolescent health from 1990 to 2015: findings from the global burden of diseases, injuries, and risk factors 2015 study. JAMA Pediatr 2017 Jun 1;171(6):573-592. [doi: 10.1001/jamapediatrics.2017.0250] [Medline: 28384795]
2. Michielsen M, Semeijn E, Comijs HC, et al. Prevalence of attention-deficit hyperactivity disorder in older adults in The Netherlands. Br J Psychiatry 2012 Oct;201(4):298-305. [doi: 10.1192/bjp.bp.111.101196] [Medline: 22878132]
3. Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA. The worldwide prevalence of ADHD: a systematic review and metaregression analysis. Am J Psychiatry 2007 Jun;164(6):942-948. [doi: 10.1176/ajp.2007.164.6.942] [Medline: 17541055]
4. Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. Annual research review: a meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. J Child Psychol Psychiatry 2015 Mar;56(3):345-365. [doi: 10.1111/jcpp.12381] [Medline: 25649325]
5. Simon V, Czobor P, Bálint S, Mészáros A, Bitter I. Prevalence and correlates of adult attention-deficit hyperactivity disorder: meta-analysis. Br J Psychiatry 2009 Mar;194(3):204-211. [doi: 10.1192/bjp.bp.107.048827] [Medline: 19252145]
6. Attention-Deficit / Hyperactivity Disorder (ADHD) Atten-Deficit Hyperact Disord ADHD. CDC. 2024. URL: https://www.cdc.gov/adhd/index.html [accessed 2024-12-18]
7. Barkley RA. Major life activity and health outcomes associated with attention-deficit/hyperactivity disorder. J Clin Psychiatry 2002;63 Suppl 12:10-15. [Medline: 12562056]
8. Barkley RA, Dawson G. Higher risk of mortality for individuals diagnosed with autism spectrum disorder or attention-deficit/hyperactivity disorder demands a public health prevention strategy. JAMA Pediatr 2022 Apr 1;176(4):e216398. [doi: 10.1001/jamapediatrics.2021.6398] [Medline: 35157011]
9. Catalá-López F, Hutton B, Page MJ, et al. Mortality in persons with autism spectrum disorder or attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. JAMA Pediatr 2022 Apr 1;176(4):e216401. [doi: 10.1001/jamapediatrics.2021.6401] [Medline: 35157020]
10. Erskine HE, Norman RE, Ferrari AJ, et al. Long-term outcomes of attention-deficit/hyperactivity disorder and conduct disorder: a systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry 2016 Oct;55(10):841-850. [doi: 10.1016/j.jaac.2016.06.016]
11. Impey M, Heun R. Completed suicide, ideation and attempt in attention deficit hyperactivity disorder. Acta Psychiatr Scand 2012 Feb;125(2):93-102. [doi: 10.1111/j.1600-0447.2011.01798.x] [Medline: 22118301]
12. Nourredine M, Gering A, Fourneret P, et al. Association of attention-deficit/hyperactivity disorder in childhood and adolescence with the risk of subsequent psychotic disorder: a systematic review and meta-analysis. JAMA Psychiatry 2021 May 1;78(5):519-529. [doi: 10.1001/jamapsychiatry.2020.4799] [Medline: 33625499]
13. Willcutt EG, Nigg JT, Pennington BF, et al. Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. J Abnorm Psychol 2012 Nov;121(4):991-1010. [doi: 10.1037/a0027347] [Medline: 22612200]
14. Scheres A, Lee A, Sumiya M. Temporal reward discounting and ADHD: task and symptom specific effects. J Neural Transm (Vienna) 2008;115(2):221-226. [doi: 10.1007/s00702-007-0813-6] [Medline: 17876680]
15. Toplak ME, Tannock R. Time perception: modality and duration effects in attention-deficit/hyperactivity disorder (ADHD). J Abnorm Child Psychol 2005 Oct;33(5):639-654. [doi: 10.1007/s10802-005-6743-6] [Medline: 16195956]
16. Fliers EA, Franke B, Lambregts‐Rommelse NNJ, et al. Undertreatment of motor problems in children with ADHD. Child Adoles Ment Health 2010 May;15(2):85-90 [FREE Full text] [doi: 10.1111/j.1475-3588.2009.00538.x]
17. DuPaul GJ, Gormley MJ, Laracy SD. Comorbidity of LD and ADHD: implications of DSM-5 for assessment and treatment. J Learn Disabil 2013;46(1):43-51. [doi: 10.1177/0022219412464351] [Medline: 23144063]
18. Mak ADP, Lee S, Sampson NA, et al. ADHD comorbidity structure and impairment: results of the WHO world mental health surveys international college student project (WMH-ICS). J Atten Disord 2022 Jun;26(8):1078-1096. [doi: 10.1177/10870547211057275] [Medline: 34753324]

XSL•FO
RenderX

19. Gaebel W, Stricker J, Kerst A. Changes from ICD-10 to ICD-11 and future directions in psychiatric classification. Dialogues Clin Neurosci 2020 Mar;22(1):7-15. [doi: 10.31887/DCNS.2020.22.1/wgaebel] [Medline: 32699501]
20. Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry 2010 Jul;167(7):748-751. [doi: 10.1176/appi.ajp.2010.09091379] [Medline: 20595427]
21. Webb SJ, Naples AJ, Levin AR, et al. The autism biomarkers consortium for clinical trials: initial evaluation of a battery of candidate EEG biomarkers. Am J Psychiatry 2023 Jan 1;180(1):41-49. [doi: 10.1176/appi.ajp.21050485] [Medline: 36000217]
22. Lozupone M, Montagna M, D'Urso F, et al. The role of biomarkers in psychiatry. In: Guest PC, editor. Rev Biomark Stud Psychiatr Neurodegener Disord: Springer International Publishing; 2019:135-162. [doi: 10.1007/978-3-030-05542-4_7]
23. Mahmood T. Biomarkers in psychiatry: a clinician's viewpoint. Br Med Bull 2020 Oct 14;135(1):23-27. [doi: 10.1093/bmb/ldaa019] [Medline: 32676652]
24. Biederman J, Spencer T. Attention-deficit/hyperactivity disorder (ADHD) as a noradrenergic disorder. Biol Psychiatry 1999 Nov 1;46(9):1234-1242. [doi: 10.1016/s0006-3223(99)00192-4] [Medline: 10560028]
25. Castellanos FX, Sonuga-Barke EJS, Milham MP, Tannock R. Characterizing cognition in ADHD: beyond executive dysfunction. Trends Cogn Sci 2006 Mar;10(3):117-123. [doi: 10.1016/j.tics.2006.01.011] [Medline: 16460990]
26. Sonuga-Barke EJS, Castellanos FX. Spontaneous attentional fluctuations in impaired states and pathological conditions: a neurobiological hypothesis. Neurosci Biobehav Rev 2007;31(7):977-986. [doi: 10.1016/j.neubiorev.2007.02.005] [Medline: 17445893]
27. Asherson P, Buitelaar J, Faraone SV, Rohde LA. Adult attention-deficit hyperactivity disorder: key conceptual issues. Lancet Psychiatry 2016 Jun;3(6):568-578. [doi: 10.1016/S2215-0366(16)30032-3] [Medline: 27183901]
28. Mayes SD, Calhoun SL, Crowell EW. Learning disabilities and ADHD: overlapping spectrumn disorders. J Learn Disabil 2000;33(5):417-424. [doi: 10.1177/002221940003300502] [Medline: 15495544]
29. Pliszka SR. Treating ADHD and Comorbid Disorders: Psychosocial and Psychopharmacological Interventions: Guilford Press; 2009.
30. Posner J, Polanczyk GV, Sonuga-Barke E. Attention-deficit hyperactivity disorder. The Lancet 2020 Feb;395(10222):450-462. [doi: 10.1016/S0140-6736(19)33004-1]
31. Thapar A, Cooper M, Rutter M. Neurodevelopmental disorders. Lancet Psychiatry 2017 Apr;4(4):339-346. [doi: 10.1016/S2215-0366(16)30376-5] [Medline: 27979720]
32. Gascon A, Gamache D, St-Laurent D, Stipanicic A. Do we over-diagnose ADHD in North America? A critical review and clinical recommendations. J Clin Psychol 2022 Dec;78(12):2363-2380. [doi: 10.1002/jclp.23348] [Medline: 35322417]
33. Jj GP, Peñas JJG. Existe un sobrediagnóstico del trastorno de déficit de atención e hiperactividad (TDAH)? : Evidencias en pediatría; 2012.
34. Kazda L, Bell K, Thomas R, McGeechan K, Sims R, Barratt A. Overdiagnosis of attention-deficit/hyperactivity disorder in children and adolescents: a systematic scoping review. JAMA Netw Open 2021 Apr 1;4(4):e215335. [doi: 10.1001/jamanetworkopen.2021.5335] [Medline: 33843998]
35. Trujillo DS, López LCT. Sobrediagnóstico del Trastorno por Déficit de Atención e Hiperactividad: producto contextual o afinidad clínica. Rev Electrónica Psyconex 2018;10(16):1-15 [FREE Full text]
36. Posner J, Park C, Wang Z. Connecting the dots: a review of resting connectivity MRI studies in attention-deficit/hyperactivity disorder. Neuropsychol Rev 2014 Mar;24(1):3-15. [doi: 10.1007/s11065-014-9251-z] [Medline: 24496902]
37. Faraone SV, Asherson P, Banaschewski T, et al. Attention-deficit/hyperactivity disorder. Nat Rev Dis Primers 2015 Aug 6;1(1):15020. [doi: 10.1038/nrdp.2015.20] [Medline: 27189265]
38. Uddin LQ, Dajani DR, Voorhies W, Bednarz H, Kana RK. Progress and roadblocks in the search for brain-based biomarkers of autism and attention-deficit/hyperactivity disorder. Transl Psychiatry 2017 Aug 22;7(8):e1218-e1218. [doi: 10.1038/tp.2017.164] [Medline: 28892073]
39. Gülhan PG, Özmen G. The use of fMRI regional analysis to automatically detect ADHD through a 3D CNN-based approach. J Imaging Inform Med 2025 Feb;38(1):203-216. [doi: 10.1007/s10278-024-01189-5] [Medline: 39028358]
40. Rubia K. Cognitive neuroscience of attention deficit hyperactivity disorder (ADHD) and its clinical translation. Front Hum Neurosci 2018;12:100. [doi: 10.3389/fnhum.2018.00100] [Medline: 29651240]
41. Cortese S, Kelly C, Chabernaud C, et al. Toward systems neuroscience of ADHD: a meta-analysis of 55 fMRI studies. AJP 2012 Oct;169(10):1038-1055. [doi: 10.1176/appi.ajp.2012.11101521]
42. Plichta MM, Scheres A. Ventral–striatal responsiveness during reward anticipation in ADHD and its relation to trait impulsivity in the healthy population: a meta-analytic review of the fMRI literature. Neurosci Biobehav Rev 2014 Jan;38:125-134. [doi: 10.1016/j.neubiorev.2013.07.012]
43. Fassbender C, Schweitzer JB. Is there evidence for neural compensation in attention deficit hyperactivity disorder? A review of the functional neuroimaging literature. Clin Psychol Rev 2006 Aug;26(4):445-465. [doi: 10.1016/j.cpr.2006.01.003] [Medline: 16500007]
44. Fair DA, Posner J, Nagel BJ, et al. Atypical default network connectivity in youth with attention-deficit/hyperactivity disorder. Biol Psychiatry 2010 Dec 15;68(12):1084-1091. [doi: 10.1016/j.biopsych.2010.07.003] [Medline: 20728873]

XSL•FO
RenderX

45. Michel CM, Murray MM. Towards the utilization of EEG as a brain imaging tool. Neuroimage 2012 Jun;61(2):371-385. [doi: 10.1016/j.neuroimage.2011.12.039] [Medline: 22227136]
46. Clarke AR, Barry RJ, Johnstone S. Resting state EEG power research in attention-deficit/hyperactivity disorder: a review update. Clin Neurophysiol 2020 Jul;131(7):1463-1479. [doi: 10.1016/j.clinph.2020.03.029] [Medline: 32387965]
47. Hari R, Puce A. EEG Primer: Oxford University Press; 2023.
48. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci 2002 Mar;3(3):201-215. [doi: 10.1038/nrn755] [Medline: 11994752]
49. Katsuki F, Constantinidis C. Bottom-up and top-down attention: different processes and overlapping neural systems. Neuroscientist 2014 Oct;20(5):509-521. [doi: 10.1177/1073858413514136] [Medline: 24362813]
50. Koechlin E, Summerfield C. An information theoretical approach to prefrontal executive function. Trends Cogn Sci 2007 Jun;11(6):229-235. [doi: 10.1016/j.tics.2007.04.005] [Medline: 17475536]
51. Ptak R. The frontoparietal attention network of the human brain: action, saliency, and a priority map of the environment. Neuroscientist 2012 Oct;18(5):502-515. [doi: 10.1177/1073858411409051] [Medline: 21636849]
52. Corbetta M, Patel G, Shulman GL. The reorienting system of the human brain: from environment to theory of mind. Neuron 2008 May 8;58(3):306-324. [doi: 10.1016/j.neuron.2008.04.017] [Medline: 18466742]
53. DiQuattro NE, Geng JJ. Contextual knowledge configures attentional control networks. J Neurosci 2011 Dec 7;31(49):18026-18035. [doi: 10.1523/JNEUROSCI.4040-11.2011] [Medline: 22159116]
54. Hoogman M, Bralten J, Hibar DP, et al. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. Lancet Psychiatry 2017 Apr;4(4):310-319. [doi: 10.1016/S2215-0366(17)30049-4] [Medline: 28219628]
55. Lenartowicz A, Loo SK. Use of EEG to diagnose ADHD. Curr Psychiatry Rep 2014 Nov;16(11):498. [doi: 10.1007/s11920-014-0498-0] [Medline: 25234074]
56. Loo SK, Makeig S. Clinical utility of EEG in attention-deficit/hyperactivity disorder: a research update. Neurotherapeutics 2012 Jul;9(3):569-587. [doi: 10.1007/s13311-012-0131-z] [Medline: 22814935]
57. Johnstone SJ, Barry RJ, Clarke AR. Ten years on: a follow-up review of ERP research in attention-deficit/hyperactivity disorder. Clin Neurophysiol 2013 Apr;124(4):644-657. [doi: 10.1016/j.clinph.2012.09.006] [Medline: 23063669]
58. Catherine Joy R, Thomas George S, Albert Rajan A, Subathra MSP. Detection of ADHD from EEG signals using different entropy measures and ANN. Clin EEG Neurosci 2022 Jan;53(1):12-23. [doi: 10.1177/15500594211036788] [Medline: 34424101]
59. Peterson BS, Trampush J, Brown M, et al. Tools for the diagnosis of ADHD in children and adolescents: a systematic review. Pediatrics 2024 Apr 1;153(4):e2024065854. [doi: 10.1542/peds.2024-065854] [Medline: 38523599]
60. Emmanuel M, Bokor BR. Tanner Stages: StatPearls Publishing; 2024. [Medline: 29262142]
61. Goodman R. Psychometric poperties of the strengths and difficulties questionnaire. J Am Acad Child Adolesc Psychiatry 2001 Nov;40(11):1337-1345. [doi: 10.1097/00004583-200111000-00015]
62. Vostanis P. Strengths and difficulties questionnaire: research and clinical applications. Curr Opin Psychiatry 2006 Jul;19(4):367-372. [doi: 10.1097/01.yco.0000228755.72366.05] [Medline: 16721165]
63. Burns GL, Lee S, Servera M, McBurnett K, Becker SP. Child and adolescent behavior inventory-parent version 10. 2015. URL: https://scholar.google.com/scholar?hl=es&as_sdt=0%2C5&q=Burns%2C+G.+L.%2C+S.+Lee%2C+M.+Servera%2C+K.+McBurnett+and+S.+P.+Becker+%282018%29.+Child+and+Adolescent+Behavior+Inventory+-+Parent+Version+1.1.+Pullman%2C+WA%2C+USA%2C+Author.&btnG= [accessed 2024-12-17]
64. DuPaul GJ, Power TJ, Anastopoulos AD, Reid R. ADHD Rating Scale—IV: Checklists, Norms, and Clinical Interpretation: The Guilford Press; 1998.
65. Wechsler D. WISC-V: Technical and Interpretive Manual: NCS Pearson, Incorporated; 2014.
66. Conners CK. Conners Continuous Performance Test 3rd Edition (Conners CPT 3) & Connors Continuous Auditory Test of Attention (Conners CATA): Technical Manual: MHS; 2014.
67. Portellano JA, Martínez-Arias R, Zumárraga LENFEN. Evaluación Neuropsicológica de Las Funciones Ejecutivas En Niños: TEA ediciones; 2009.
68. Cuetos F, Arribas-Aguila D, Suárez-Coalla P, Martínez-García C. P. Diagnóstico y Detección Temprana de La Dislexia 2020.
69. Sáez B, Servera M, Burns GL, Becker SP. Advancing the multi-informant assessment of sluggish cognitive tempo: child self-report in relation to parent and teacher ratings of SCT and impairment. J Abnorm Child Psychol 2019 Jan;47(1):35-46. [doi: 10.1007/s10802-018-0436-4] [Medline: 29700714]
70. Alosh M, Bretz F, Huque M. Advanced multiplicity adjustment methods in clinical trials. Stat Med 2014 Feb 20;33(4):693-713. [doi: 10.1002/sim.5974] [Medline: 24105821]
71. Sayal K, Prasad V, Daley D, Ford T, Coghill D. ADHD in children and young people: prevalence, care pathways, and service provision. Lancet Psychiatry 2018 Feb;5(2):175-186. [doi: 10.1016/S2215-0366(17)30167-0] [Medline: 29033005]
72. Legislación consolidada. Agencia Estatal Boletín Oficial del Estado. URL: https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673 [accessed 2026-02-11]

XSL•FO
RenderX

73. Clarke AR, Barry RJ, Dupuy FE, McCarthy R, Selikowitz M, Heaven PCL. Childhood EEG as a predictor of adult attention-deficit/hyperactivity disorder. Clin Neurophysiol 2011 Jan;122(1):73-80. [doi: 10.1016/j.clinph.2010.05.032] [Medline: 20598939]
74. Loo SK, Cho A, Hale TS, McGough J, McCracken J, Smalley SL. Characterization of the theta to beta ratio in ADHD: identifying potential sources of heterogeneity. J Atten Disord 2013 Jul;17(5):384-392. [doi: 10.1177/1087054712468050] [Medline: 23264365]
75. Jarrett MA, Gable PA, Rondon AT, Neal LB, Price HF, Hilton DC. An EEG study of children with and without ADHD symptoms: between-group differences and associations with sluggish cognitive tempo symptoms. J Atten Disord 2020 May;24(7):1002-1010. [doi: 10.1177/1087054717723986] [Medline: 28800715]
76. Chutko LS, Yakovenko EA, Surushkina S, et al. Clinical and neurophysiological manifestations of sluggish cognitive tempo in children. SS Korsakov J Neuro Psych 2024 Jun 17;124(5):120. [doi: 10.17116/jnevro2024124051120]
77. Chang MY, Ouyang CS, Chiang CT, et al. A new method of diagnosing attention-deficit hyperactivity disorder in male patients by quantitative EEG analysis. Clin EEG Neurosci 2019 Sep;50(5):339-347. [doi: 10.1177/1550059419859164] [Medline: 31321994]
78. Rezaeezadeh M, Shamekhi S, Shamsi M. Attention Deficit Hyperactivity Disorder Diagnosis using non-linear univariate and multivariate EEG measurements: a preliminary study. Phys Eng Sci Med 2020 Jun;43(2):577-592. [doi: 10.1007/s13246-020-00858-3] [Medline: 32524443]
79. Bussalb A, Collin S, Barthélemy Q, et al. Is there a cluster of high theta-beta ratio patients in attention deficit hyperactivity disorder? Clin Neurophysiol 2019 Aug;130(8):1387-1396. [doi: 10.1016/j.clinph.2019.02.021] [Medline: 31176621]
80. McNorgan C, Judson C, Handzlik D, Holden JG. Linking ADHD and behavioral assessment through identification of shared diagnostic task-based functional connections. Front Physiol 2020;11:583005. [doi: 10.3389/fphys.2020.583005] [Medline: 33391011]
81. Michelini G, Jurgiel J, Bakolis I, et al. Atypical functional connectivity in adolescents and adults with persistent and remitted ADHD during a cognitive control task. Transl Psychiatry 2019 Apr 12;9(1):137. [doi: 10.1038/s41398-019-0469-7] [Medline: 30979865]
82. Jensen O, Mazaheri A. Shaping functional architecture by oscillatory alpha activity: gating by inhibition. Front Hum Neurosci 2010;4:186. [doi: 10.3389/fnhum.2010.00186] [Medline: 21119777]
83. Abibullaev B, An J. Decision support algorithm for diagnosis of ADHD using electroencephalograms. J Med Syst 2012 Aug;36(4):2675-2688. [doi: 10.1007/s10916-011-9742-x] [Medline: 21671069]
84. Khoshnoud S, Nazari MA, Shamsi M. Functional brain dynamic analysis of ADHD and control children using nonlinear dynamical features of EEG signals. JIN 2018 Jan 1;17(1):17-30. [doi: 10.3233/JIN-170033]
85. Castellanos FX, Proal E. Large-scale brain systems in ADHD: beyond the prefrontal-striatal model. Trends Cogn Sci 2012 Jan;16(1):17-26. [doi: 10.1016/j.tics.2011.11.007] [Medline: 22169776]

Abbreviations

ADHD: attention deficit/hyperactivity disorder
ADHD-RS-IV: Attention Deficit/Hyperactivity Disorder Rating Scale IV
ASD: autism spectrum disorder
CABI: Child and Adolescent Behavior Inventory
CCI-2: Child Concentration Inventory-Version 2
CPT-3: Continuous Performance Test, Third Edition
DAN: dorsal attention network
***DSM-5*:** *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition*
EEG: electroencephalogram
ENFEN: Neuropsychological Evaluation of the Executive Functions
fMRI: functional magnetic resonance imaging
FPF: false positive fraction
HUPHM: Puerta de Hierro University Hospital in Majadahonda
***ICD-11*:** *International Classification of Diseases, 11th Revision*
ML: machine learning
NPV: negative predictive value
PI: principal investigator
PPV: positive predictive value
SCT: sluggish cognitive tempo
SDQ: Strengths and Difficulties Questionnaire
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
STARD: Standards for Reporting Diagnostic Accuracy Studies
TBR: theta/beta ratio

XSL•FO
RenderX

TPF: true positive fraction
VAN: ventral attention network
WISC-V: Wechsler Intelligence Scale for Children, Fifth Edition

Edited by J Sarvestan; submitted 16.Jun.2025; peer-reviewed by R Garotti; revised version received 29.Dec.2025; accepted 30.Dec.2025; published 27.Mar.2026.

Please cite as:
Blasco-Fontecilla H, Sánchez-Cerezo J, Gómez I, Abreu-Fernández G, Ortiz S, Villoria JF, Blanco M, García A, Ballesteros J, Martínez R, Gálvez G, Maestú F, López-Medrano Á
Measuring Accuracy (Classification Probabilities, Positive, and Negative Predictive Values) of Executive Function Electroencephalogram Metrics in Attention-Deficit/Hyperactivity Disorder Diagnosis: Protocol for and Perspectives From the SINCRONIA Study
JMIR Res Protoc 2026;15:e79150
URL: https://www.researchprotocols.org/2026/1/e79150
doi:10.2196/79150

© Hilario Blasco-Fontecilla, Javier Sánchez-Cerezo, Irene Gómez, Georgelina Abreu-Fernández, Sandra Ortiz, Jesús F Villoria, Miguel Blanco, Ana García, Julia Ballesteros, Roldán Martínez, Gerardo Gálvez, Fernando Maestú, Álvaro López-Medrano. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Competency and Role Development in Advanced Nursing Practice Following a Swiss Master's Cohort Through Education and Early Clinical Practice: Protocol for a Longitudinal Convergent Mixed Methods Study

Astrid Braun[1,2*], MSc; André Fringer[1*], Prof Dr; Heidrun Gattinger[3*], Prof Dr; Christian Eissler[4*], Prof Dr; Maya Zumstein-Shaha[4*], Prof Dr; Thomas Beer[3*], Prof Dr; Martina Roes[2*], Prof Dr

[1]School of Health Sciences, ZHAW Zurich University of Applied Sciences, Katharina-Sulzer-Platz 9, Winterthur, Switzerland

[2]Department of Nursing Science, Witten/Herdecke University, Witten, Germany

[3]Institute of Health Sciences, Ostschweizer Fachhochschule OST, St. Gallen, Switzerland

[4]School of Health Professions, Bern University of Applied Sciences, Bern, Switzerland

[*]all authors contributed equally

Corresponding Author:
Astrid Braun, MSc
School of Health Sciences, ZHAW Zurich University of Applied Sciences, Katharina-Sulzer-Platz 9, Winterthur, Switzerland

Abstract

Background: Master's degree programs in advanced nursing practice have been established in Switzerland for approximately 2 decades, and the role of advanced practice nurses (APNs) is increasingly embedded within the Swiss health care system. Despite their growing presence, there remains a lack of clarity and consistency regarding this progress; persistent ambiguity in role definitions, expectations, and competencies continues to hinder consistent implementation and contributes to uncertainty for both APNs and employers. Existing research has not sufficiently examined how APN competencies and role perceptions develop during graduate education and the early transition to practice. This substudy of the *EDUCate* research initiative, *EDUCate*–Competence and Role Development (*EDUCate C-ARE*), focuses exclusively on a defined cohort of master's students and follows their competence and role development longitudinally.

Objective: The *EDUCate C-ARE* substudy aims to provide a comprehensive understanding of how APN competencies and role perceptions change during a master's program and in the first year following graduation (research question [RQ] 1), identify factors that facilitate or hinder their development (RQ 2), and explore challenges that APNs report during this process (RQ 3).

Methods: A longitudinal convergent mixed methods design will be applied. Quantitative data on APN competencies are collected using the Advanced Practice Nursing Competency Assessment Instrument at 4 time points (T0-T3), and qualitative data are generated through semistructured interviews at 3 time points (I1-I3). The quantitative and qualitative strands are analyzed independently and subsequently integrated using joint displays and narrative weaving to highlight convergence, complementarity, or divergence. Data collection spans 4 years, beginning in September 2023.

Results: Baseline (T0) and midprogram (T1) quantitative data have been collected, with 141 students responding at T0 (full time: n=25, 17.7%; part time: n=56, 39.7%) and T1 (full time: n=17, 12.1%; part time: n=43, 30.5%). Data collection at T2 is ongoing. Initial analyses have been completed for T0 and T1. Qualitative interviews at I1 have been conducted and preliminarily analyzed (full time: n=6; part time: n=20). Data collection for full-time students is expected to conclude by May 2026, with preliminary findings anticipated by August 2026; for part-time students, data collection will continue until August 2027, with initial findings expected by October 2027. Integrated study results are anticipated by late 2026, with final analyses following completion of all data collection phases.

Conclusions: This study will generate longitudinal evidence on APN competence and role development in a Swiss master's cohort. By developing a conceptual, explanatory model of competence acquisition, role perception, and influencing factors, the study is expected to inform curriculum development, strengthen role clarity, and support the evidence-based implementation of APN roles in the Swiss health care system. Findings will contribute to national and international discussions on APN competency frameworks and professional role development.

International Registered Report Identifier (IRRID): DERR1-10.2196/85773

(JMIR Res Protoc 2026;15:e85773) doi:10.2196/85773

XSL•FO
RenderX

KEYWORDS

advanced practice nurse; advanced nursing practice; competence development; role development; mixed methods; longitudinal design; educational research

Introduction

Background

Since the 1960s, advanced nursing practice (ANP) has been recognized as a key strategy for addressing global health care challenges and advancing health system development, with ANP programs implemented in more than 70 countries [1,2]. In particular, the integration of ANP in primary care has been promoted to improve access to care and respond to evolving population health needs [1,2]. The role of advanced practice nurses (APNs) was conceptualized to meet increasing demands for health care services while supporting professional development and career progression within nursing [2].

In Switzerland, universities of applied sciences and academic institutions have offered master's degree programs in ANP for over 2 decades [3,4], awarding the Master of Science in Nursing degree [5]. A structured nursing development framework was established in collaboration with graduates of level I and II professional training programs [4,6,7]. In 2010, a total of 3 universities of applied sciences (Bern University of Applied Sciences [BFH], Eastern Switzerland University of Applied Sciences [OST], and Zurich University of Applied Sciences [ZHAW]) jointly developed a master's program in nursing [8,9]. Since 2019, these institutions have independently offered the program with a variety of specializations.

Conceptual Background and Role Expectations

The conceptual foundation of the APN role is rooted in the ANP framework. APNs are expected to apply the theoretical and evidence-based knowledge acquired during academic training; role development and implementation are shaped by national health care system structures and population needs [7,10]. Academic preparation generally involves a Master of Science in Nursing or an equivalent graduate qualification. The ANP framework includes several established advanced nursing roles, such as APNs, clinical nurse specialists, and nurse practitioners. These roles share a common foundation in advanced clinical expertise, evidence-based practice, and expanded professional responsibility while allowing for context-specific differentiation and adaptation [10]. They are characterized by competencies such as expert coaching and guidance, counseling and consultation, ethical decision-making, interdisciplinary collaboration, clinical and professional leadership, and evidence-based practice [10]. International efforts, including the 2020 International Council of Nurses APN guidelines, aim to strengthen coherence and clarity in the conceptualization and implementation of ANP and expand APN roles to meet evolving health care needs [11-17].

Swiss Context: Opportunities and Ongoing Role Ambiguity

In German-speaking Switzerland, the APN role has become increasingly integrated into existing health care structures. APN competencies are recognized as a sustainable and essential contribution to current and future system challenges [18]. However, persistent role ambiguity, including variability in expectations within and across professional groups and context-dependent perceptions of the defined and self-defined role, has led to role conflicts, excessive demands, and insecurity, which have been identified as potential contributors to early attrition from the profession [1,8,13,14,19-26]. Studies demonstrate that insufficient role clarity and variability in role enactment contribute to role conflict, limited autonomy, and barriers to effective implementation [22,26].

Regulatory efforts by the Swiss association Advanced Practice Nursing Switzerland focus on minimum standards for certification and consistent training quality, including criteria governing access to the APN title to ensure that qualifications match the expanded scope of practice [27].

In parallel, the objectives defined by the Swiss Professional Association of Nurses provide a central reference point for educational institutions to align curricula with evolving system needs and anticipated employment opportunities, ensuring that graduates can be strategically integrated into the health care workforce [28,29]. Taken together, these developments underscore the importance of role clarity and structured competency development to support effective APN implementation in Switzerland. International research underscores that the transition from graduate education to early clinical practice is a critical phase for APNs, marked by shifting expectations, increasing responsibility, and role adaptation challenges [30,31]. Evidence indicates that targeted support during this transition is essential for role enactment, professional identity formation, and sustainable APN integration.

Gap in the Literature

Despite progress in ANP education and role establishment, research on competence-oriented master's programs in nursing remains limited, and systematic analyses that identify and evaluate the competencies acquired within such programs are lacking [32-37]. This gap constrains evidence-informed curriculum development and the strategic deployment of APNs within the Swiss health care system.

Rationale and Scope of This Substudy

In 2021, BFH, OST, and ZHAW launched the overarching *EDUCate* initiative in German-speaking Switzerland.

This substudy of EDUCate–Competence and Role Development (*EDUCate C-ARE*) focuses exclusively on a defined cohort of master's students to ensure feasibility within the 4-year study period.

It addresses the identified evidence gap by examining how higher education influences the development of competencies and professional roles among master's students in ANP, with the aim of informing the strategic deployment of APNs within the Swiss health care system [8,38-41]

XSL•FO
RenderX

Conceptual Definitions

In this study, 3 related but distinct concepts are used:

- "Role development" refers to the broader process through which APN roles evolve at the system and organizational level, such as through policy changes, workforce planning, and institutional role implementation. This concept is not the focus of this substudy.
- "Role acquisition" describes how individual nurses adopt and internalize the advanced practice role during and after their education. It reflects the process of gaining competencies and integrating the APN role into professional identity.
- "Role perception" refers to how nurses understand and interpret their own role and responsibilities within clinical practice at a given point in time. It captures subjective experiences and expectations related to the APN role.

These definitions guide the interpretation of findings and ensure conceptual consistency throughout the manuscript. They are grounded in established conceptual frameworks for APN roles [10,20].

Study Aim and Objectives

The aim of the substudy is to longitudinally examine how APN competencies and role perceptions evolve among a defined cohort of master's students during their education and into the first year of clinical practice in Switzerland.

To achieve this aim, the study pursues the following objectives:

1. To assess changes in APN competencies and role perceptions at 4 measurement points (T0-T3) during and after master's education
2. To identify factors that facilitate or hinder competency acquisition and role development throughout the educational and early practice trajectory
3. To explore challenges experienced by APNs during this process as described by master's students and alumni
4. To integrate quantitative and qualitative findings to provide a comprehensive understanding of APN development in the Swiss context

To address these objectives, the study is guided by the following research questions (RQs):

- RQ 1—in what ways do APN competencies and role perceptions change during the master's program and in the first year following graduation?
- RQ 2—which factors facilitate or hinder the development of APN competencies and roles?
- RQ 3—what challenges do APNs report in relation to their professional development during this process?

Methods

Study Design

This study uses a longitudinal convergent mixed methods approach in which the quantitative and qualitative components are conducted concurrently and independently [42,43]. Data from both strands are collected and analyzed separately to reduce the risk of mutual influence and preserve the distinct strength of each methodological approach. Figure 1 provides an overview of this parallel structure.

XSL•FO
RenderX

Figure 1. Longitudinal convergent mixed methods design with quantitative (QUAN) and qualitative (QUAL) measurement time points and subsequent data integration (own illustration). I1: midprogram; I2: program completion; I3: 1 year after graduation; T0: program start; T1: midprogram; T2: program completion; T3: after graduation.

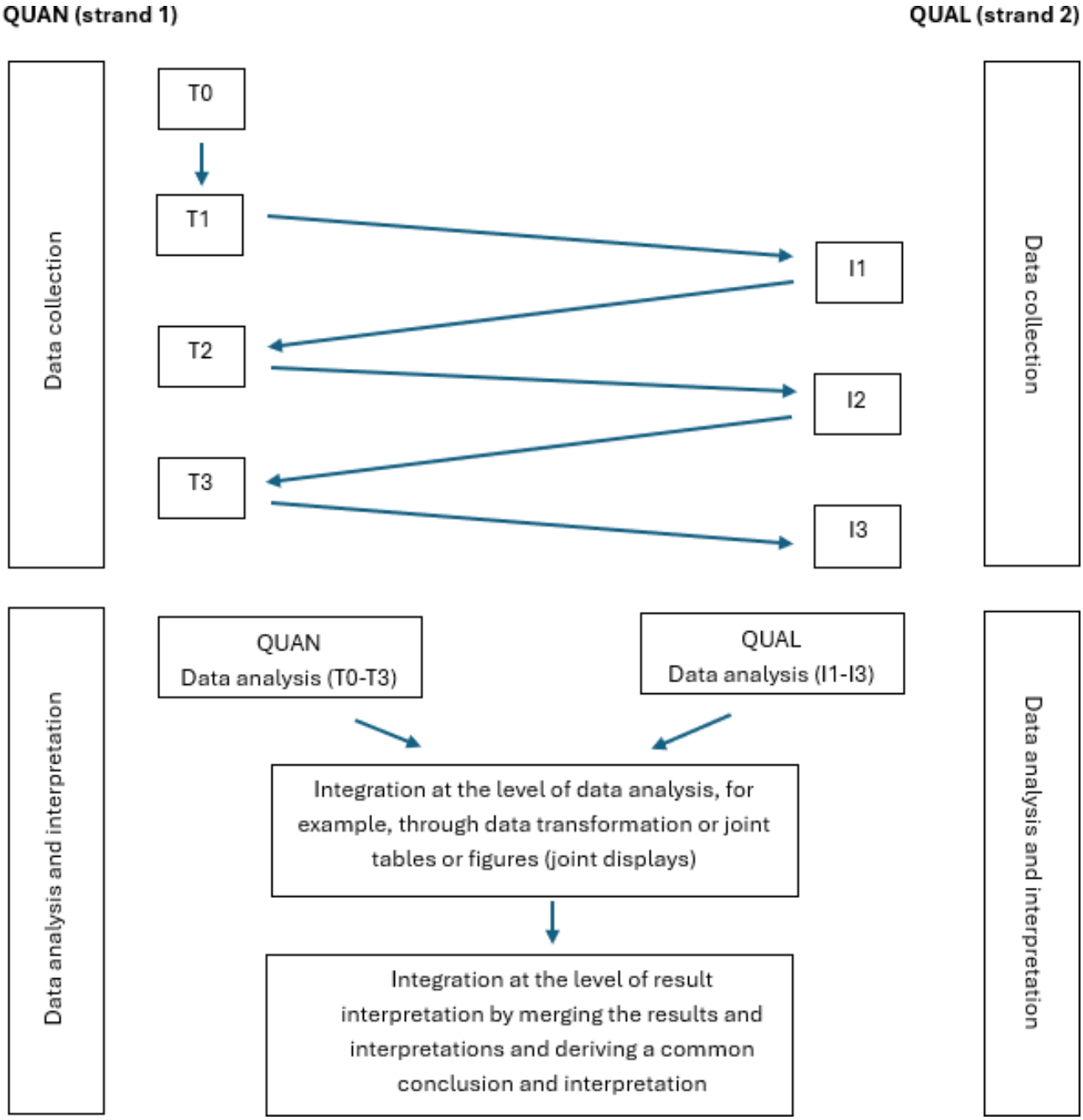

A defined cohort of master's students (start year: 2023) will be followed across 4 quantitative measurement time points (T0: program start; T1: midprogram; T2: program completion; T3: after graduation) and 3 qualitative interview points (I1: midprogram; I2: program completion; I3: 1 year after graduation).

Quantitative data are collected using the Advanced Practice Nursing Competency Assessment Instrument (APNCAI) [44], whereas qualitative data are generated through semistructured interviews. This design allows for the examination of APN competency acquisition and role perception trajectories within the same cohort over time. After separate analyses, data from both strands are integrated to provide a comprehensive picture of APN development during education and into early clinical practice.

Data collection spans 4 years beginning in 2023, covering both the academic training phase and the transition to professional practice. This extended longitudinal period supports valid and meaningful analysis. Given the complexity of the research topic, the longitudinal convergent mixed methods design is particularly suited to capturing both broad patterns and in-depth experiences [42,45-48].

Methodological Framework and Integration

The qualitative component follows the constructivist grounded theory approach by Charmaz [49], emphasizing iterative coding, constant comparison, and theory generation grounded in

XSL•FO
RenderX

participants' accounts. Quantitative and qualitative analyses are conducted independently. Integration occurs subsequently through joint displays and narrative weaving to identify convergence and complementarity.

The combination of methodological strands offers a complementary perspective: quantitative data provide measurable trends in competence development, whereas qualitative data offer detailed insights into individual experiences, contextual factors, and perceived challenges. The mixed methods design is particularly appropriate because the study involves diverse data types that capture both measurable trends (quantitative) and rich experiential accounts (qualitative). This combination allows for deeper explanatory insights, ensuring that quantitative patterns can be interpreted and contextualized through qualitative perspectives. Together, these approaches capture both measurable progression and the lived experience.

In line with Kuckartz [50] and international mixed methods scholarship [42,43,47], mixed methods are understood as the purposeful combination and integration of quantitative and qualitative approaches. Data from both strands are collected within the same overarching design, and integration typically occurs after completion of separate analyses.

Methodological Rationale

To address the study's complex RQs, a well-grounded methodological approach is essential [51]. The phenomenon under investigation is embedded in the understanding that education and learning processes occur continuously across the life span [52]. Individuals acquire knowledge and competencies at different stages, making a longitudinal mixed methods design appropriate.

The conceptualization by Gräsel [53] of empirical educational research provides a relevant framework for this study. Educational research examines learning processes, competence acquisition, influencing factors, and the effects of education on individuals and institutions. Its problem orientation, interdisciplinary character, and use of empirical methods align closely with the aims of investigating competency and role development in advanced nursing education.

Setting and Participants

The study will be conducted at 3 universities of applied sciences in Switzerland: OST, BFH, and ZHAW. The target population comprises all students enrolled in the 2023 master's cohort of APN programs at these sites. Cohort sizes vary by institution (approximately 12 in OST, 60 in BFH, and 60 in ZHAW). Students are enrolled in either full-time (FT) or part-time (PT) study modes. Measurement time points are mode adjusted: T1 takes place at the program midpoint (FT: 9 months; PT: 18 months), and T2 takes place at program completion (FT: 18 months; PT: 36 months). The inclusion criteria for participation in this study are membership in the 2023 cohort and provision of informed consent. Exclusion criteria include instances of nonconsent.

Recruitment and Sampling

All eligible students in the 2023 master's cohort are invited at program entry (T0) to participate in the quantitative surveys and to be recontacted at subsequent time points (T1-T3). For the qualitative component, interview participants are drawn from the same cohort at I1 to I3. Sampling for interviews aims to capture variation by site (OST, BFH, and ZHAW) and study mode (FT or PT). Participation is voluntary, with the option to withdraw at any time.

As of autumn of 2023, approximately 120 students had commenced their studies across the participating universities as part of the overarching *EDUCate* project. Recruitment of both master's students and clinically active APNs will be coordinated through the heads of the respective degree programs and supported by a dedicated project website.

For the quantitative component, a target response of 15% to 20% is anticipated. The qualitative component follows the principles of grounded theory methodology, in which theoretical sampling, a core feature of this approach, guides the number and selection of interviews. Data collection continues until theoretical saturation is reached [54].

Data Collection

Overview

In the *EDUCate C-ARE* study, data collection follows a longitudinal convergent mixed methods design in which quantitative and qualitative components are conducted sequentially across multiple time points (T0, T1, I1, T2, I2, T3, and I3), as illustrated in Figure 2. This design enables a comprehensive mapping of the development of APN competencies and role development over time.

XSL•FO
RenderX

Figure 2. Overview of the timeline for data collection (DC), data analysis (DA), and integration. FT: full time; I1: midprogram; I2: program completion; I3: 1 year after graduation; PT: part time; QUAL: qualitative; QUAN: quantitative; T0: program start; T1: midprogram; T2: program completion; T3: after graduation.

XSL•FO
RenderX

Quantitative and qualitative data are collected and analyzed separately to ensure methodological rigor. The design provides both cross-sectional and longitudinal insights. When appropriate, comparisons with relevant subgroups (eg, PT students or other cohorts) may be conducted to identify patterns of convergence, divergence, or overlap.

Quantitative Data Collection

Quantitative data are collected using the APNCAI [44]. This 44-item instrument assesses core competencies of APNs across 8 dimensions: research and evidence-based practice, clinical and professional leadership, interprofessional relationships, mentoring, professional autonomy, quality and care management, professional education and teaching, and health promotion. Data are gathered from the selected student cohort at the 4 predefined time points: T0 (at the commencement of the master's program), T1 (midway through the program; after 12 months for FT students or 16 months for PT students), T2 (at program completion), and T3 (12 months after graduation). Mode (FT or PT) and site (OST, BFH, or ZHAW) are recorded to enable subgroup summaries and comparisons.

The quantitative component (strand 1) of the study follows a longitudinal design with repeated measurement across 4 time points [42,55], enabling the assessment of changes in participants' self-assessed competencies and role perceptions over time. This approach supports the identification of temporal developments and potential long-term effects of the educational program on participants' professional growth and role understanding.

Qualitative Data Collection

Qualitative data are collected via semistructured interviews at 3 predefined phases: I1 (midprogram), I2 (program completion), and I3 (1 year after graduation). The interview guide is aligned with APN competency domains and focuses on role acquisition and role perception, as well as perceived facilitators, barriers, and challenges during education and transition to practice. Interviews are audio recorded and transcribed verbatim in accordance with data protection regulations and the transcription standards of Dresing and Pehl [56], including low-level transcription details [50]. Pretests of the interview guide were conducted with former student cohorts to ensure clarity and relevance. Transcription is performed by an internal transcription service of the university to ensure accuracy and data security. Interview guides are provided in Multimedia Appendix 1 [42,50,57-59].

The qualitative component (strand 2) is designed as a descriptive study using grounded theory methodology [54,60]. Semistructured interviews are conducted with a purposive sample of students and practicing APNs to explore their subjective experience of competence development and role formation. The interviews aim to elicit participants' perceptions and expectations related to their academic and professional roles, explore how they experience their everyday work and study life, and identify factors that influence role performance and potential solutions to address challenges [57].

Participants are selected using a maximum variation sampling strategy (minimum and maximum contrast) to ensure heterogeneity within each cohort. Approximately 15% to 20% of students per cohort will be invited to participate [58]. To reach theoretical saturation, further recruitment may be necessary, with a target distribution of 50%, 30%, and 10% of interviews at I1, I2, and I3, respectively. Recruitment is coordinated by the respective program directors at the participating universities of applied sciences. All interviews are conducted online, recorded with participants' consent, and stored securely on a protected network drive. Access to the data will be restricted to the first author and designated project staff.

Outcomes

The primary quantitative outcome is the APNCAI total score over time (T0-T3). The secondary quantitative outcomes are APNCAI domain-level scores summarized overall and by study mode (FT or PT) and site (OST, BFH, or ZHAW). Primary qualitative outcomes are themes describing role acquisition and role perception across I1 to I3. Secondary qualitative outcomes are themes on facilitators, barriers, and reported challenges during the educational trajectory and early clinical practice.

Quantitative Data Analysis

Quantitative data will be analyzed using R (version 4.5.0; R Foundation for Statistical Computing). Descriptive summaries (means, SDs, medians, and IQRs) are provided for total and domain scores at each time point. Visualizations such as box plots depict distributions and changes over time, including stratifications by study mode (FT or PT) and site (BFH, OST, or ZHAW). Where repeated measurements per individual are available, within-cohort change is examined using appropriate paired or repeated-measure approaches. Where only cross-sectional snapshots are available, between-group comparisons at each time point use appropriate parametric or nonparametric tests. Effect sizes and 95% CIs are reported where applicable.

Statistical procedures include descriptive statistics; ANOVA; Spearman rank correlation; and, where appropriate, linear mixed-effects models to account for repeated measures and individual-level variation over time. All statistical assumptions will be tested prior to analysis, including normality (eg, Shapiro-Wilks test), homogeneity of variances (Leven test), and sphericity (Mauchly test) for repeated-measure ANOVA. Residual diagnostics will be conducted for mixed models to ensure model validity.

Group comparisons explore differences between FT and PT students, as well as between cohorts from the 3 participating universities of applied sciences: BFH and ZHAW (each with approximately 55 students) and OST (with approximately 10 students). Due to unequal group sizes, appropriate statistical techniques are applied to ensure robustness. All participants hold a Bachelor of Science in Nursing, ensuring a consistent educational baseline across cohorts.

To uphold methodological rigor, the principles of objectivity, reliability, and validity are maintained throughout all stages of data collection and analysis. Objectivity is ensured through standardized procedures independent of the individual administering the survey. Reliability and validity are supported by implementing the instrument across multiple institutions.

XSL•FO
RenderX

Pretests with earlier student cohorts were conducted to optimize the instrument and ensure its applicability. Data are securely stored and accessible only to authorized project staff.

Qualitative Data Analysis

Qualitative data are analyzed using constructivist grounded theory (Charmaz) following iterative coding and categorization to develop a theoretical understanding of the studied phenomenon. The analysis includes initial coding, focused coding, constant comparison across cases and time points (I1-I3), and analytic memoing to trace category development. Classic coding steps (open coding, axial coding, and selective coding [61]) are applied. The MAXQDA 2025 software (VERBI GmbH) [50,62] is used to support data management, coding, and analysis.

The analysis seeks to build a conceptual understanding of role acquisition and role perception and explicate facilitators, barriers, and challenges across the educational and transition-to-practice trajectory. An audit trail and reflexive memos are maintained to enhance transparency and trustworthiness.

To ensure methodological rigor, the analysis will adhere to established qualitative quality criteria such as transparency, intersubjectivity, and comprehensiveness [63]. In addition, the study will follow the trustworthiness criteria proposed by Lincoln und Guba [64], including credibility, transferability, dependability, and confirmability, to further strengthen the robustness and transparency of the qualitative component. Transparency is ensured through detailed documentation of all analytical steps and the use of software-supported coding. Intersubjective comprehensibility is promoted through regular team discussions and reflective engagement with the data. The aim is to develop a theoretically grounded and context-sensitive understanding of APN role development from the participants' perspectives.

Integration of Findings and Data Integration

Integration of findings is conducted after the completion of both quantitative and qualitative analyses. A sequential approach is applied in which results from each strand are first analyzed independently and then brought together during the interpretation phase. Joint displays and narrative weaving are used to relate quantitative patterns in APNCAI scores (T0-T3) to qualitative themes (I1-I3) on role acquisition, role perception, and contextual factors. Integration is guided by the study objectives and RQs, highlighting convergence, complementarity, or divergence between strands.

Data integration represents a central component of the mixed methods approach. In accordance with best practices, quantitative and qualitative data will be analyzed separately using methodologies appropriate to each paradigm. Following the completion of data collection and initial analyses, the findings will be brought together through a process of integration by triangulation. This approach involves systematic comparison and synthesis of results from both strands to enhance the validity, complementarity, and depth of interpretation.

Visual representations such as joint displays and matrices are used to facilitate the integration process and support interpretation and reporting. The integration strategy is grounded in the principle of methodological triangulation, which leverages the strengths of both quantitative and qualitative approaches to provide a more comprehensive understanding of the research problem [65]. This process also serves to mitigate the limitations inherent to each individual method by enabling cross-validation and complementary insights [66].

Integration focuses on identifying areas of convergence, divergence, and expansion between the 2 data types. Quantitative results contribute to generalizability, whereas qualitative findings will offer a contextualized, in-depth understanding.

Following the framework proposed by Creswell and Plano Clark [42], this integrative process culminates in the development of meta-inferences, which are overarching conclusions that emerge from the synthesis of both data strands. These meta-inferences aim to generate insights that are more robust, nuanced, and meaningful than those derived from a single method, thereby strengthening the overall explanatory power of the study.

Reporting Standards and Methodological Quality

The reporting of this protocol and the subsequent mixed methods study is guided by internationally recognized standards to ensure methodological rigor, transparency, and clarity. Reporting follows the GRAMMS (Good Reporting of a Mixed Methods Study) [47,52] framework for mixed methods research. For the qualitative components, the SRQR (Standards for Reporting Qualitative Research) guidelines will be applied [67]. For the observational quantitative elements, the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines will be used [68].

To support methodological coherence and quality across both strands, the Mixed Methods Appraisal Tool [69] and the GRAMMS guidelines [47,52] are used throughout the design, analysis, and integration phases. These tools were selected because they are widely accepted in health research and are recommended by the EQUATOR (Enhancing the Quality and Transparency of Health Research) Network [70], which provides central guidance for best practices in research reporting.

Methodological quality is further ensured through strategies such as maintaining an audit trail, reflexive memoing, and peer debriefing during qualitative analysis. Data collection and analysis procedures are standardized across participating sites to enhance consistency and reliability.

Ethical Considerations

The overarching *EDUCate* study has been reviewed by the Ethics Committee of the Canton of Zurich, Switzerland, and was deemed exempt from formal authorization in accordance with the Swiss Human Research Act (Business Administration System for Ethics Committees Req-2019-00875). This substudy is conducted within the approved *EDUCate* framework.

Participation in both the survey and interview components is voluntary, and individuals who choose not to participate will not experience any disadvantages or negative consequences.

XSL•FO
RenderX

Written informed consent is obtained from all participants prior to data collection.

The study is being conducted in full compliance with the Swiss Human Research Act [71], applicable Swiss legislation on data protection and research involving human participants [72], the Declaration of Helsinki [73], and the principles of good clinical practice [74]. All procedures related to confidentiality, data protection, and participant rights are strictly adhered to throughout the study. No financial or material compensation is provided to participants for their involvement in this study.

Results

Overview

The *EDUCate C-ARE* study is internally funded by ZHAW in June 2023. Data collection began in September 2023 with the cohort starting the master's program in ANP in the same year.

Quantitative data collection at baseline (T0) and midprogram (T1) has been conducted for the 2023 cohort. In total, 81 students completed T0 (FT: n=25, 30.9%; PT: n=56, 69.1%), and 60 students completed T1 (FT: n=17, 28.3%; PT: n=43, 71.7%).

Data collection at T2 (program completion) is ongoing, with 15 responses received to date (FT: n=15, 100%); PT and additional FT responses are pending. Final numbers will be reported once data collection is complete.

Quantitative Data Collection Timeline

For FT students, T0 is September 2023, T1 is September 2024, T2 is January 2025, and T3 is January 2026. For PT students, T1 is April 2025, T2 is May 2026, and T3 is May 2027.

Qualitative Interview Timeline

For FT students, I1 is October 2024, I2 is February 2025, and I3 is February 2026. For PT students, I1 is May 2025, I2 is June 2026, and I3 is June 2027.

Preliminary qualitative analyses have been completed for the I1 interviews.

Overall Study Progress and Preliminary Findings

Initial quantitative analyses have been completed for T0 and T1. Data collection for FT students is expected to be completed by May 2026, with preliminary results anticipated by August 2026. For PT students, data collection is projected to conclude by August 2027, with initial findings expected by October 2027. Final analysis and integrated results will follow the completion of all data collection phases. The first overarching study results are planned to be available by the end of 2026.

Discussion

Principal Findings

The primary objective of the *EDUCate C-ARE* substudy is to document and analyze the development and application of competencies, as well as the evolving roles of master's students and clinically active APNs. Guided by RQs 1 to 3, the study is expected to yield novel insights into the representation of professional competencies, role expectations and enactment, and the perceived benefits and challenges associated with APN roles over the program and the first year in practice. Particular attention will be paid to the clarity or ambiguity of role definitions and the implications for professional identity and practice. The study further aims to contribute to the international comparability of APN competencies and role profiles [75].

A central anticipated outcome is the formulation of a conceptual, explanatory model that outlines barriers to competency and role development and proposes feasible strategies to address them. In doing so, the study seeks to deepen understanding of APN competencies and role trajectories and underscore the importance of role clarity and professional development for the advancement of nursing practice and the effectiveness of educational programs. These insights are expected to inform academic curricula and clinical practice and support the strategic implementation of APN roles in health care systems.

Comparison With Prior Work

The study's objectives and anticipated contributions are situated within established frameworks and evidence. Alignment with the model of ANP by Hamric [76] provides a conceptual basis for examining advanced competencies, role expectations, and professional identity. The use of the APNCAI enables a connection with existing validation studies and supports international comparability of competence profiles [44]. In the Swiss context, where empirical evidence on APN role development remains limited, this study addresses a documented gap by offering longitudinal, mixed methods evidence on competency acquisition, role perception, and influencing factors. In combination, these anchors situate the expected findings within the broader international APN discourse while responding to national needs for role clarification and evidence-informed implementation.

Strengths and Limitations

The study's convergent mixed methods design, applied longitudinally across 4 quantitative time points and 3 qualitative interview phases, allows for the examination of change over time and the integration of breadth and depth of evidence. The combination of a standardized competence assessment (APNCAI) with constructivist grounded theory supports cross-method triangulation and enhances credibility, validity, and depth of interpretation. Data collection across multiple universities of applied sciences increases heterogeneity and supports external validity. The explicit integration strategy (eg, joint displays and narrative weaving) strengthens meta-inference and theory development and constitutes a methodological contribution for the study of APN role and competence development.

Several limitations must be acknowledged. The mixed methods design is resource intensive and constrained by time and participant availability. The sample is limited to German-speaking regions of Switzerland, which may affect the generalizability to other linguistic or cultural contexts. Additional constraints include potential language and professional barriers, limited financial and temporal resources, and the risk of selection bias as individuals with a preexisting

XSL•FO
RenderX

interest in APN role development may have been more likely to participate.

Conclusions

The *EDUCate C-ARE* study is expected to provide relevant insights into the development of competencies and roles among master's students and practicing APNs in Switzerland. Through methodological triangulation and the development of a conceptual explanatory model, the study aims to contribute to both theoretical advancement and practical innovation in the ANP field. By identifying key barriers and proposing actionable solutions, it seeks to support targeted improvements in education, clinical practice, and professional development and reinforce the critical contribution of APNs to the evolving health care landscape.

Dissemination

The findings of the *EDUCate C-ARE* study will be disseminated through publications in national and international peer-reviewed journals. In addition, the validated APNCAI assessment instrument will be translated following established, evidence-based guidelines for the translation of research and clinical instruments. A dedicated project website has been developed to provide background information and will be regularly updated with new research findings and outputs [77].

Acknowledgments

The authors would like to express their sincere gratitude to all students who participated in the study.

Funding

This study was funded by internal resources from Zurich University of Applied Sciences. No external funding was received.

Data Availability

The datasets generated or analyzed during this study will be made available by the first author (AB) on reasonable request.

Authors' Contributions

AB led the design and implementation of the EDUCate C-ARE (EDUCate–Competence and Role Development) study. AB prepared the initial manuscript draft. AF and MR reviewed the manuscript and provided critical feedback. AB also prepared Figures 1 and 2. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[PDF File, 476 KB - resprot_v15i1e85773_app1.pdf]

References

1. Bryant-Lukosius D, Martin-Misener R. Advanced practice nursing: an essential component of country level human resources for health. : International Council of Nurses; 2016.
2. Schober M. Introduction to Advanced Nursing Practice: An International Focus: Springer; 2016.
3. Schwendimann R, Fierz K, Spichiger E, Marcus B, De Geest S. A master of nursing science curriculum revision for the 21st century - a progress report. BMC Med Educ 2019 May 8;19(1):135. [doi: 10.1186/s12909-019-1588-9] [Medline: 31068167]
4. Spirig R. 10 years advanced nursing practice in Switzerland: retrospect and prospects. Pflege 2010 Dec;23(6):363-366. [doi: 10.1024/1012-5302/a000075] [Medline: 21280251]
5. Definition and characteristics for nurse practitioner/advanced practice nursing roles. : International Council of Nurses (ICN); 2009.
6. Schubert M, Herrmann L, Spichiger E. Akademisierung Der Pflege–Evidenz Und Wirksamkeitsforschung Akademisch Ausgebildetes Pflegefachpersonal: Entwicklung Und Chancen: Springer; 2017:85-100. [doi: 10.1007/978-3-662-54887-5_7]
7. Ulrich A, Hellstern P, Kressig RW, Eze G, Spirig R. Advanced Nursing Practice (ANP) im direkten Pflegealltag: Die pflegerische Praxisentwicklung eines akutgeriatrischen ANP-Teams. Pflege 2010 Dec 1;23(6):403-410. [doi: 10.1024/1012-5302/a000079]
8. Gysin S, Sottas B, Odermatt M, Essig S. Advanced practice nurses' and general practitioners' first experiences with introducing the advanced practice nurse role to Swiss primary care: a qualitative study. BMC Fam Pract 2019 Nov 27;20(1):163. [doi: 10.1186/s12875-019-1055-z] [Medline: 31775651]
9. Mahrer-Imhof R, Eicher M, Frauenfelder F, Oulevey Bachmann A, Ulrich A, Expertenbericht APN. Verein für pflegewissenschaft [vfp]. : Swiss Association for Nursing Science; 2012.

XSL•FO
RenderX

10. Tracy MF, O'Grady ET, Philips SJ. Hamric & Hansons's Advanced Practice Nursing-E-Book: An Integrative Approach, 7th edition: Elsevier Health Sciences; 2022.
11. Ordóñez-Piedra J, Ponce-Blandón JA, Robles-Romero JM, Gómez-Salgado J, Jiménez-Picón N, Romero-Martín M. Effectiveness of the advanced practice nursing interventions in the patient with heart failure: a systematic review. Nurs Open 2021 Jul;8(4):1879-1891. [doi: 10.1002/nop2.847] [Medline: 33689229]
12. Schober M, Lehwaldt D, Rogers M, et al. Guidelines on advanced practice nursing. : International Council of Nurses; 2020 URL: https://www.icn.ch/system/files/documents/2020-04/ICN_APN%20Report_EN_WEB.pdf [accessed 2026-03-20]
13. Schönenberger N, Sottas B, Merlo C, Essig S, Gysin S. Patients' experiences with the advanced practice nurse role in Swiss family practices: a qualitative study. BMC Nurs 2020;19(1):90. [doi: 10.1186/s12912-020-00482-2] [Medline: 32982581]
14. Spichiger E, Zumstein-Shaha M, Schubert M, Herrmann L. Gezielte Entwicklung von Advanced Practice Nurse-Rollen für spezifische Patient(inn)engruppen in einem Schweizer Universitätsspital. Pflege 2018 Feb;31(1):41-50. [doi: 10.1024/1012-5302/a000594]
15. Swan M, Ferguson S, Chang A, Larson E, Smaldone A. Quality of primary care by advanced practice nurses: a systematic review. Int J Qual Health Care 2015 Oct;27(5):396-404. [doi: 10.1093/intqhc/mzv054] [Medline: 26239474]
16. State of the world's nursing 2020: investing in education, jobs and leadership. : World Health Organization; 2020.
17. Woo BFY, Lee JXY, Tam WWS. The impact of the advanced practice nursing role on quality of care, clinical outcomes, patient satisfaction, and cost in the emergency and critical care settings: a systematic review. Hum Resour Health 2017 Sep 11;15(1):63. [doi: 10.1186/s12960-017-0237-9] [Medline: 28893270]
18. Keinath E, Dirksen A, Lehwaldt D, Glarcher M, Essl-Maurer R, Dach C, et al. Nurse practitioner development in German-speaking countries: Germany, Austria, and Switzerland. In: Thomas SL, Rowles JS, editors. Nurse Practitioners and Nurse Anesthetists: The Evolution of the Global Roles: Springer International Publishing; 2023:207-223. [doi: 10.1007/978-3-031-20762-4_15]
19. Bryant-Lukosius D, Dicenso A, Browne G, Pinelli J. Advanced practice nursing roles: development, implementation and evaluation. J Adv Nurs 2004 Dec;48(5):519-529. [doi: 10.1111/j.1365-2648.2004.03234.x] [Medline: 15533090]
20. Bryant-Lukosius D, Spichiger E, Martin J, et al. Framework for evaluating the impact of advanced practice nursing roles. J Nurs Scholarsh 2016 Mar;48(2):201-209. [doi: 10.1111/jnu.12199] [Medline: 26869323]
21. Josi R, Bianchi M, Brandt SK. Advanced practice nurses in primary care in Switzerland: an analysis of interprofessional collaboration. BMC Nurs 2020;19(1):1. [doi: 10.1186/s12912-019-0393-4] [Medline: 31908597]
22. Mackavey C, Henderson C, van Leeuwen EDZ, Maas L, Ladd A. The advanced practice nurse role's development and identity: an international review. International Journal for Advancing Practice 2024 Jan 2;2(1):36-44. [doi: 10.12968/ijap.2024.2.1.36]
23. Newall F, Twomey B, Lima S. Advanced practice nursing–promoting organisation clarity and connectedness: a mixed methods approach. Collegian 2018 Feb;25(1):97-103. [doi: 10.1016/j.colegn.2017.04.003]
24. Serena A, Dwyer AA, Peters S, Eicher M. Acceptance of the advanced practice nurse in lung cancer role by healthcare professionals and patients: a qualitative exploration. J Nurs Scholarsh 2018 Sep;50(5):540-548. [doi: 10.1111/jnu.12411] [Medline: 30033614]
25. Stürmer N, Frei IA, Nicca D. Erfahrungen von Advanced Practice Nurses mit einem klinischen Weiterbildungsstudiengang im Kontext ihrer beruflichen Entwicklung. Pflege 2021 Aug;34(4):213-220. [doi: 10.1024/1012-5302/a000822]
26. Van Hecke A, Vlerick I, Akhayad S, Daem M, Decoene E, Kinnaer LM. Dynamics and processes influencing role integration of advanced practice nurses and nurse navigators in oncology teams. Eur J Oncol Nurs 2023 Feb;62:102257. [doi: 10.1016/j.ejon.2022.102257] [Medline: 36634592]
27. Green Paper Zu Den Rollen Und Handlungsfeldern Der Pflegeexpertinnen / Pflegeexperten APN - Stand Der Reglementierung Online: APN-CH: Organisation der Reglementierung; 2020.
28. De Geest S, Moons P, Callens B, Gut C, Lindpaintner L, Spirig R. Introducing Advanced Practice Nurses / Nurse Practitioners in health care systems: a framework for reflection and analysis. Swiss Med Wkly 2008 Nov 1;138(43-44):621-628. [doi: 10.4414/smw.2008.12293] [Medline: 19005867]
29. Pflege 2030 positionspapier des schweizer berufsverbandes der pflegefachfrauen und pflegefachmänner SBK. : Schweizer Berufsverbandes der Pflegefachfrauen und Pflegefachmänner [SBK-ASI]; 2021.
30. Li Y, Wang C, Tan W, Jiang Y. The transition to advanced practice nursing: a systematic review of qualitative studies. Int J Nurs Stud 2023 Aug;144:104525. [doi: 10.1016/j.ijnurstu.2023.104525] [Medline: 37263057]
31. Swenson C, Ogilvie S, Richardson S. Mentorship and the advanced practice nurse: a concept analysis. J Am Assoc Nurse Pract 2025 Dec 1;37(12):683-691. [doi: 10.1097/JXX.0000000000001135] [Medline: 40359187]
32. Deutscher berufsverband für pflegeberufe (dbfk) bundesverband e.v. et al. In: Advanced Nursing Practice in Deutschland, Österreich Und Der Schweiz Positionspapier DBfK, ÖGKV Und SBK: Eine Positionierung des DBfK, ÖGKV und SBK; 2013.
33. Zumstein-Shaha M. Pflegeexpertin/-e APN in der Schweiz Rechtliche Grundlagen und Schwierigkeiten bei der Abrechnung. Pflegerecht 2022;3(3):151-157. [doi: 10.24451/arbor.17820]
34. Stephanow V. Advanced Practice Nurses: Pionierarbeit am Patienten. Pflegez 2020 Apr;73(4):10-13. [doi: 10.1007/s41906-020-0680-4]

XSL•FO
RenderX

35. Zumstein-Shaha M, von Dach C, Moramba R, et al. Neue Rollen der nicht-ärztlichen Berufe in der Schweizer Grundversorgung. Prim Hosp Care (de) 2022. [doi: 10.4414/phc-d.2022.10398]
36. Eissler C, Zumstein-Shaha M. Kompetenzniveaus neuer Rollen in der Schweizer Gesundheitsversorgung: eine Literaturrecherche. Präv Gesundheitsf 2023 Aug;18(3):308-315. [doi: 10.1007/s11553-022-00958-w]
37. Schirmer AM, Golz C, Eissler C. Advanced Nursing Practice im europäischen Raum. Präv Gesundheitsf 2025 May;20(2):202-211. [doi: 10.1007/s11553-024-01116-0]
38. Swiss Research Agenda for Nursing SRAN 2019-2029. : Swiss Association for Nursing Science; 2019 URL: https://www.vfp-apsi.ch/pflegewissenschaft/forschung [accessed 2026-03-13]
39. Die gesundheitspolitische strategie des bundesrates 2020–2030. In: Gesundheit: Bundesamt für Gesundheit [BAG]; 2019.
40. Künzi K, Jäggi J, Dutoit L. Aktueller Stand Der Schweizerischen Diskussion Über Den Einbezug von Hoch Ausgebildeten Nichtärztlichen Berufsleuten in Der Medizinischen Grundversorgung: Bundesamt für Gesundheit [BAG]; 2014.
41. Steinbrüchel-Boesch C. Advanced Practice Nurses in der Hausarztpraxis. Der informierte Arzt 2019;9(3):24-26 [FREE Full text]
42. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research, 3rd edition: SAGE Publications; 2017.
43. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
44. Sastre-Fullana P, Morales-Asencio JM, Sesé-Abad A, Bennasar-Veny M, Fernández-Domínguez JC, De Pedro-Gómez J. Advanced Practice Nursing Competency Assessment Instrument (APNCAI): clinimetric validation. BMJ Open 2017 Feb 23;7(2):e013659. [doi: 10.1136/bmjopen-2016-013659] [Medline: 28235968]
45. Hagenauer G, Gläser-Zikuda M. Mixed methods. In: Harring M, Rohlfs C, Gläser-Zikuda M, editors. Handbuch Schulpädagogik: Waxmann; 2019:801-812.
46. Creswell JW. A Concise Introduction to Mixed Methods: SAGE Publications; 2014.
47. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008 Apr;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
48. Plano Clark VL, Ivankova NV. Mixed Methods Research: A Guide to The Fielf: Sage Publications; 2016.
49. Charmaz K. Constructing Grounded Theory, 3rd edition: Sage; 2025.
50. Kuckartz U. Mixed Methods: Methodologie, Forschungsdesigns Und Analyseverfahren: Springer-Verlag; 2014.
51. Cohen L, Manion L, Morrison K. Research Methods in Education, 8th edition: Routledge/Taylor & Francis Group; 2017:944.
52. Mejeh M, Hagenauer G, Gläser-Zikuda M. Mixed methods research on learning and instruction—meeting the challenges of multiple perspectives and levels within a complex field. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research 2023;24(1). [doi: 10.17169/fqs-24.1.3989]
53. Gräsel C. Was ist empirische bildungsforschung? In: Reinders H, Ditton H, Gräsel C, Gniewosz BH, editors. Empirische Bildungsforschung: Strukturen Und Methoden: Springer Science & Business Media; 2010:195.
54. Creamer EG. Advancing Grounded Theory with Mixed Methods: Routledge; 2022.
55. Kuckartz U, Rädiker S. Qualitative Content Analysis: Methods, Practice and Software, 2nd edition: SAGE Publications; 2023.
56. Dresing T, Pehl T. Praxisbuch Interview, Transkription & Analyse: Anleitungen Und Regelsysteme Für Qualitativ Forschende, 8th edition: Dr. Dresing und Pehl; 2018. [doi: 10.1007/978-3-658-18387-5_56-2]
57. Helfferich C. Die Qualität Qualitativer Daten Manual Für Die Durchführung Qualitativer Interviews: VS Verlag für Sozialwissenschaften; 2014:739.
58. Kelle U, Kluge SK. Vom Einzelfall Zum Typus Fallvergleich Und Fallkontrastierung in Der Qualitativen Sozialforschung: VS Verlag für Sozialwissenschaften; 2010. [doi: 10.1007/978-3-531-92366-6]
59. Hamric AB, Hanson CM, Tracy MF, O'Grady ET. Advanced Practice Nursing-EBook: An Integrative Approach, 5th edition: Elsevier Health Sciences; 2013.
60. Sandelowski M. Theory unmasked: the uses and guises of theory in qualitative research. Res Nurs Health 1993 Jun;16(3):213-218. [doi: 10.1002/nur.4770160308] [Medline: 8497673]
61. Strauss AL, Corbin JM. Grounded Theory: Grundlagen Qualitativer Sozialforschung: Beltz: Psychologie-Verlag-Union; 1996.
62. MAXQDA 2025 [computer software]. : VERBI Software; 2025.
63. Flick U. Gütekriterien qualitativer sozialforschung. In: Baur N, Blasius J, editors. Handbuch Methoden Der Empirischen Sozialforschung: Springer Fachmedien Wiesbaden; 2019:473-488. [doi: 10.1007/978-3-658-21308-4_33]
64. Lincoln YS, Guba EG. Naturalistic: Sage; 1985. [doi: 10.1016/0147-1767(85)90062-8]
65. Kuckartz U. Datenanalyse in der Mixed-Methods-Forschung. Köln Z Soziol 2017 Oct;69(S2):157-183. [doi: 10.1007/s11577-017-0456-z]
66. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258. [doi: 10.1136/bmj.h1258] [Medline: 25791983]
67. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]

XSL•FO
RenderX

68. Cuschieri S. The STROBE guidelines. Saudi J Anaesth 2019 Apr;13(Suppl 1):S31-S34. [doi: 10.4103/sja.SJA_543_18] [Medline: 30930717]
69. Hong QN, Fàbregues S, Bartlett G, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. EFI 2018;34(4):285-291. [doi: 10.3233/EFI-180221]
70. EQUATOR Network. URL: https://www.equator-network.org [accessed 2026-03-12]
71. Federal Act on Research Involving Human Beings (Human Research Act, HRA). Fedlex. URL: https://www.fedlex.admin.ch/eli/cc/2013/617/en [accessed 2026-09-19]
72. Federal Acton Data Protection (Data Protection Act, FADP). Fedlex. URL: https://www.fedlex.admin.ch/eli/cc/2022/491/en [accessed 2026-03-19]
73. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
74. ICH Harmonised Guideline: Integrated Addendum to ICH E6(R1): Guideline for Good Clinical Practice E6(R2). : International Council for Harmonisation; 2016 URL: https://database.ich.org/sites/default/files/E6_R2_Addendum.pdf [accessed 2026-03-20]
75. Turnbull J, Lathlean J. The research process in nursing. In: Mixed Methods: Wiley-Blackwell; 2015:371-383.
76. Hamric AB, Hanson CM, Tracy MF, O'Grady ET. Hamric and Hanson's Advanced Practice Nursing: An Integrative Approach, 7th edition: Elsevier Health Sciences; 2022.
77. EDUCATE – edukationsbegleitforschung master of science in pflege. Zürcher Hochschule Für Angewandte Wissenschaften (ZHAW). URL: https://www.zhaw.ch/de/gesundheit/forschung/pflege/projekte/educate-edukationsbegleitforschung-master-of-science-in-pflege [accessed 2026-03-12]

Abbreviations

ANP: advanced nursing practice
APN: advanced practice nurse
APNCAI: Advanced Practice Nursing Competency Assessment Instrument
BFH: Bern University of Applied Sciences
EDUCate C-ARE: EDUCate–Competence and Role Development
EQUATOR: Enhancing the Quality and Transparency of Health Research
FT: full time
GRAMMS: Good Reporting of a Mixed Methods Study
OST: Eastern Switzerland University of Applied Sciences
PT: part time
RQ: research question
SRQR: Standards for Reporting Qualitative Research
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
ZHAW: Zurich University of Applied Sciences

Edited by A Schwartz; submitted 14.Oct.2025; peer-reviewed by C Jackson, DC Gray; accepted 24.Feb.2026; published 30.Mar.2026.

Please cite as:
Braun A, Fringer A, Gattinger H, Eissler C, Zumstein-Shaha M, Beer T, Roes M
Competency and Role Development in Advanced Nursing Practice Following a Swiss Master's Cohort Through Education and Early Clinical Practice: Protocol for a Longitudinal Convergent Mixed Methods Study
JMIR Res Protoc 2026;15:e85773
URL: https://www.researchprotocols.org/2026/1/e85773
doi:10.2196/85773

© Astrid Braun, André Fringer, Heidrun Gattinger, Christian Eissler, Maya Zumstein-Shaha, Thomas Beer, Martina Roes. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Multidisciplinary Treatment With Hepatic Arterial Infusion Chemotherapy, Radiotherapy, and Immunotherapy for Advanced Hepatocellular Carcinoma With Major Vascular Invasion: Prospective Registry Protocol

Yoshiko Doi[1*], MD, PhD; Hiroshi Aikata[2*], MD, PhD; Yumi Kosaka[2], MD, PhD; Takashi Nakahara[2], MD, PhD; Michiyo Kodama[3], MD, PhD; Masashi Hieda[4], MD, PhD; Masakazu Hashimoto[5], MD, PhD; Hideki Nakahara[5], MD, PhD; Ippei Takahashi[6], MD, PhD; Hideaki Kakizawa[7], MD, PhD; Nami Mori[8], MD, PhD; Keiji Tsuji[8], MD, PhD; Nobuki Imano[9], MD, PhD; Yuji Murakami[9], MD, PhD; Saki Sueda[10], MD; Tomokazu Kawaoka[10], MD, PhD; Masataka Tsuge[11], MD, PhD; Shiro Oka[10], MD, PhD

[1]Department of Radiation Oncology, Hiroshima Prefectural Hospital, Hiroshima, Japan

[10]Department of Gastroenterology, Hiroshima University Hospital, Hiroshima, Japan

[11]Liver Center, Hiroshima University Hospital, Hiroshima, Japan

[2]Department of Gastroenterology and Hepatology, Hiroshima Prefectural Hospital, 1-5-54, Ujinakanda, Minami-ku, Hiroshima, Japan

[3]Department of Medical Oncology, Hiroshima Prefectural Hospital, Hiroshima, Japan

[4]Department of Diagnostic Radiology, Hiroshima Prefectural Hospital, Hiroshima, Japan

[5]Department of Gastroenterological Surgery, Hiroshima Prefectural Hospital, Hiroshima, Japan

[6]Department of Radiation Oncology, Hiroshima Red Cross Hospital and Atomic-bomb Survivors Hospital, Hiroshima, Japan

[7]Department of Diagnostic Radiology, Hiroshima Red Cross Hospital and Atomic-bomb Survivors Hospital, Hiroshima, Japan

[8]Department of Gastroenterology and Hepatology, Hiroshima Red Cross Hospital and Atomic-bomb Survivors Hospital, Hiroshima, Japan

[9]Department of Radiation Oncology, Hiroshima University Hospital, Hiroshima, Japan

[*]these authors contributed equally

Corresponding Author:
Hiroshi Aikata, MD, PhD
Department of Gastroenterology and Hepatology, Hiroshima Prefectural Hospital, 1-5-54, Ujinakanda, Minami-ku, Hiroshima, Japan

Abstract

Background: Systemic therapy, including immune checkpoint inhibitors, has improved survival in advanced hepatocellular carcinoma (HCC); however, its efficacy remains limited in patients with macroscopic vascular invasion (MVI), a subgroup with an extremely poor prognosis. Although combining immunotherapy with local treatments such as hepatic arterial infusion chemotherapy (HAIC) and radiation therapy (RT) is considered a promising approach, robust supportive evidence from routine clinical practice is lacking.

Objective: This study aims to evaluate the safety and therapeutic effectiveness of a multidisciplinary treatment strategy involving RT after HAIC, followed by immunotherapy, in patients with MVI-positive HCC, using real-world clinical data from Japan.

Methods: This is a prospective, multicenter registry study conducted at 3 hospitals in Hiroshima Prefecture, Japan. Eligible patients will have unresectable MVI-positive HCC confirmed by dynamic computed tomography. The treatment protocol follows a standardized sequence: 1 session of HAIC (cisplatin), RT targeting the MVI site (25 Gy in 5 fractions), and subsequent systemic immunotherapy. The primary end point is safety, which will be evaluated using the Common Terminology Criteria for Adverse Events (version 5.0). The secondary end points include progression-free survival at 12 and 24 weeks, tumor response, median progression-free survival, overall survival, and objective response rate at 12 and 24 weeks, assessed according to the Response Evaluation Criteria in Solid Tumors criteria. Data will be collected prospectively and analyzed according to the intention-to-treat principle.

Results: Patient enrollment began in March 2025, and data collection and analysis are ongoing as participants continue to be followed.

Conclusions: This prospective registry study will generate real-world evidence on the safety and effectiveness of a multidisciplinary strategy combining HAIC, RT, and immunotherapy in patients with MVI-positive HCC. Given that all components

XSL•FO
RenderX

are covered under Japan's national health insurance, this approach could be readily implemented in clinical practice and may inform future treatment guidelines for MVI-positive HCC.

International Registered Report Identifier (IRRID): DERR1-10.2196/82992

(JMIR Res Protoc 2026;15:e82992) doi:10.2196/82992

KEYWORDS

hepatocellular carcinoma; macroscopic vascular invasion; hepatic arterial infusion chemotherapy; radiotherapy; immunotherapy; prospective registry study

Introduction

In the treatment of unresectable, locally advanced hepatocellular carcinoma (HCC), multikinase inhibitors such as sorafenib and lenvatinib were long regarded as standard first-line therapies, as established by the SHARP (Sorafenib HCC Assessment Randomized Protocol) and REFLECT (Study of Lenvatinib vs Sorafenib in First-Line Treatment of Patients With Unresectable Hepatocellular Carcinoma) trials [1,2]. However, recent landmark phase III trials—IMbrave150 (atezolizumab plus bevacizumab) [3], HIMALAYA (durvalumab plus tremelimumab) [4], and CheckMate 9DW (nivolumab plus ipilimumab) [5]—have demonstrated that immune checkpoint inhibitor–based combination regimens significantly improve overall survival (OS), resulting in a paradigm shift in the management of advanced HCC [6,7].

Despite these advances, immune checkpoint inhibitor–based regimens have shown limited benefit in patients with macroscopic vascular invasion (MVI), one of the most prognostically unfavorable subgroups. In IMbrave150, atezolizumab plus bevacizumab significantly improved OS in high-risk patients—including those with the main portal vein tumor thrombus, bile duct invasion, or ≥50% liver involvement—yet the median OS in this subgroup was only 7.6 months [8]. Moreover, the HIMALAYA and CheckMate 9DW trials excluded patients with the main portal vein tumor thrombus, leaving a lack of robust clinical evidence for this population.

Before the advent of immunotherapy, several studies demonstrated that combining locoregional therapies—such as hepatic arterial infusion chemotherapy (HAIC) and radiation therapy (RT)—with systemic agents could improve outcomes in MVI-positive HCC [9-13]. In a phase III trial, He et al [9] showed that HAIC using the FOLFOX regimen (folinic acid, fluorouracil, and oxaliplatin) plus sorafenib yielded a markedly higher overall response rate (ORR; 40.8% vs 2.46%; $P<.001$) and longer progression-free survival (PFS; 7.03 vs 2.6 months; $P<.001$) than sorafenib monotherapy for HCC with portal vein invasion. A Japanese prospective study also demonstrated improved survival with HAIC using the cisplatin (CDDP) regimen plus sorafenib vs sorafenib alone in patients with major portal vein tumor thrombus (median OS 11.4 vs 6.5 months) [10]. Furthermore, in a large-scale multicenter retrospective study conducted in Japan using propensity score matching, HAIC using the CDDP was associated with better survival outcomes than sorafenib monotherapy in MVI-positive patients without extrahepatic lesions (median OS 11.4 vs 6.5 months; $P=.02$) [11]. RT also has been shown to improve survival in patients with MVI-positive HCC, as demonstrated in a randomized clinical trial [12]. In particular, Kosaka et al [13] reported an ORR of 19.6%, a thrombus response rate of 51.0%, and a median OS of 19.8 months for the combination therapy of HAIC using the CDDP and RT in cases of MVI-positive HCC.

Since the advent of immunotherapy, combination treatments of HAIC with immunotherapy [14-17] or RT with immunotherapy [18] have been increasingly implemented for MVI-positive HCC, particularly in Asia. In both approaches, the addition of immunotherapy has tended to enhance therapeutic efficacy. This may be attributable to the induction of immunogenic cell death by HAIC or radiotherapy through tumor necrosis, thereby potentiating the effects of immunotherapy [19,20]. Furthermore, combining all 3 modalities—HAIC, RT, and immunotherapy—may also offer promising therapeutic potential. In patients with MVI-positive HCC, RT may contribute to local tumor control within the portal vein, thereby helping to preserve hepatic blood flow and liver function. Short-course hypofractionated RT can facilitate timely initiation of subsequent systemic therapy while minimizing treatment-related delays. A case series by Yamaoka et al [21] described 6 patients treated with atezolizumab plus bevacizumab following HAIC and RT, with 4 achieving survivals exceeding the previously reported median OS values. These findings suggest that multimodal strategies targeting both intrahepatic tumor burden and vascular invasion may offer meaningful clinical benefits for this challenging patient population.

Methods

Study Design and Setting

This study is a multicenter, prospective observational registry designed to collect real-world data on patients with MVI-positive HCC who are treated according to routine clinical practice at participating institutions in Hiroshima Prefecture, Japan (Hiroshima Prefectural Hospital, Hiroshima University Hospital, and Hiroshima Red Cross Hospital & Atomic-bomb Survivors Hospital). The treatment sequence of HAIC, RT, and subsequent immunotherapy reflects standard clinical practice at these institutions and is not mandated by the study protocol. The protocol does not direct or assign specific treatments; rather, it prospectively records treatment decisions made by the treating physicians. The described timing between treatment modalities represents typical clinical practice and target intervals, allowing for physician discretion based on individual patient conditions. Patient enrollment began in March 2025. On the basis of the sample size estimation, a total of 38 consecutive eligible patients

XSL•FO
RenderX

will be enrolled, with longitudinal data collection until treatment discontinuation or death. No randomization or intervention assignment will be performed, and all treatment decisions will follow routine clinical practice. An overview of the study design is shown in Figure 1.

Figure 1. Study flow of the prospective observational registry. Patients with hepatocellular carcinoma (HCC) with a tumor thrombus in the first-order branch or main trunk of the portal vein (Vp3/4) and/or a tumor thrombus in the inferior vena cava or hepatic veins (Vv2/3) are assessed for eligibility. Eligible patients receive an explanation of study participation and provide informed consent. Enrolled patients then undergo multidisciplinary treatment consisting of hepatic arterial infusion chemotherapy (HAIC), radiation therapy (RT), and subsequent immunotherapy according to routine clinical practice. Patients are followed longitudinally for safety and efficacy outcomes.

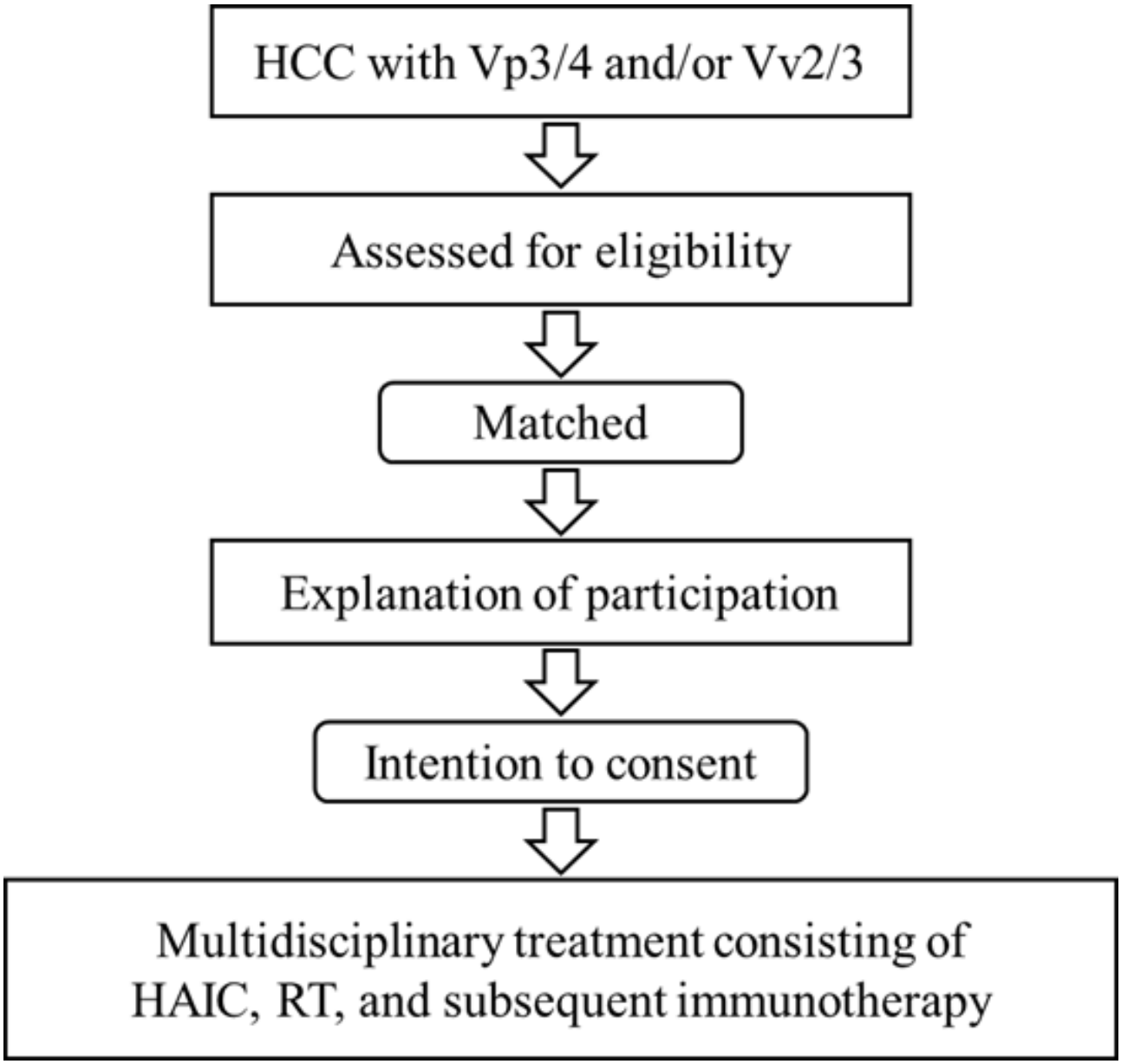

Ethical Considerations

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki and was approved by the institutional review boards (IRBs) of all participating institutions. The study protocol, including data collection and informed consent procedures, was reviewed and approved by the Hiroshima Prefectural Hospital IRB (approval 202503 - 1), the Hiroshima University Hospital IRB (approval e2024-0174-02), and the Hiroshima Red Cross Hospital & Atomic-bomb Survivors Hospital IRB (approval P7-10-1). Written informed consent will be obtained from all participants prior to enrollment. Participants will not receive any financial compensation for participation in this study.

All study data will be managed using institution-managed secure storage systems at each participating site in accordance with institutional policies and applicable data protection regulations. Patient data will be collected as part of routine clinical practice and deidentified prior to analysis. Access to study data will be restricted to authorized study personnel based on their roles, and data handling, storage, and transfer between participating institutions will be conducted using secure, institution-approved methods in compliance with ethical and legal requirements.

XSL•FO
RenderX

All monitoring activities will be conducted in compliance with institutional policies, relevant regulations, and the principles of Good Clinical Practice, ensuring that participant safety and data integrity are maintained throughout the study period.

Patient Population

Eligible participants are patients with unresectable, locally advanced HCC with MVI who are scheduled to receive multidisciplinary treatment consisting of HAIC, RT, and subsequent immunotherapy. MVI is defined radiologically as a tumor thrombus in the first-order branch or main trunk of the portal vein (Vp3/4) or in the inferior vena cava or hepatic veins (Vv2/3) based on dynamic contrast-enhanced computed tomography (CT). Key inclusion and exclusion criteria are summarized in Textbox 1.

Textbox 1. Key inclusion and exclusion criteria.

Inclusion criteria

- Hepatocellular carcinoma with macroscopic vascular invasion diagnosed by dynamic computed tomography
  - Vp3/Vp4: portal vein (first-order branches or main trunk)
  - Vv2/Vv3: tumor thrombus in inferior vena cava or hepatic veins (left, middle, or right)
- Age ≥18 years
- Eastern Cooperative Oncology Group performance status: 0‐2
- Chronic hepatitis or liver cirrhosis (Child-Pugh class A or B with score≤7)
- Written informed consent obtained

Exclusion criteria

- History of prior radiotherapy overlapping the planned field
- Unsuitable for immunotherapy
- Uncontrolled ascites or hepatic encephalopathy
- Active gastric or duodenal ulcers or untreated esophageal or gastric varices or hemorrhagic ulcers
- Severe infections or comorbidities (excluding chronic hepatitis or liver cirrhosis)
- Psychiatric disorders interfering with study participation (per investigator judgment)
- Active double cancers (concurrent malignancies)
- Contraindication for computed tomography imaging (eg, allergy to iodinated contrast)
- No informed consent provided
- Any other reason deemed inappropriate by the attending physician

Treatment Protocol

Overview

A schematic of the study flow and treatment sequence is shown in Figure 2. All eligible patients will receive 1 session of HAIC followed by RT targeting the MVI site and then systemic immunotherapy. The interval between treatments will generally be 1 to 2 weeks. Modifications to the sequence or addition of other therapies may be made at the treating physician's discretion based on therapeutic response and clinical condition.

XSL•FO
RenderX

Figure 2. Outcome assessment during follow-up. Patients are followed longitudinally after multidisciplinary treatment to evaluate safety and efficacy outcomes. Safety is assessed throughout the treatment and follow-up period, while efficacy outcomes include objective response rate (ORR), progression-free survival (PFS), and overall survival (OS). CDDPia: cisplatin intra-arterial infusion; HAIC: hepatic arterial infusion chemotherapy; RT: radiation therapy.

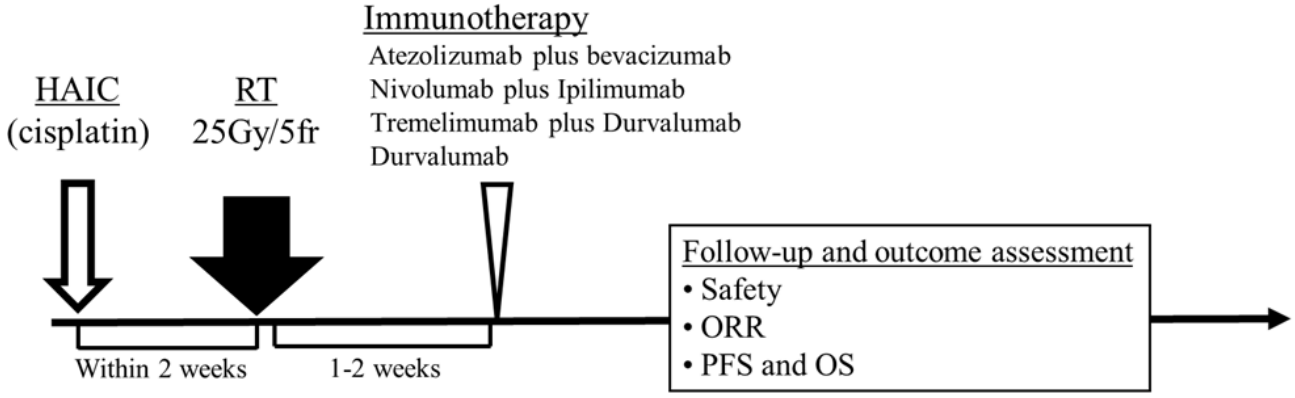

HAIC Procedure

HAIC will be performed promptly after diagnosis in clinically appropriate patients with adequate liver function (Child-Pugh class A or B7), acceptable renal function, and no contraindications to CDDP-based chemotherapy. Single-dose CDDP was selected for HAIC because it is the only agent approved and reimbursed for this indication within the Japanese public health insurance system and has demonstrated favorable efficacy and safety in patients with HCC with MVI in a previous study [22].

Catheter insertion will generally be performed via the femoral artery using the Seldinger technique. Abdominal angiography will confirm tumor staining and feeding arteries. A catheter will then be inserted into the feeding artery, and CDDP (65 mg/m²) will be infused over 20 to 40 minutes, with earlier termination permitted if sufficient tumor perfusion is achieved. If the estimated glomerular filtration rate is <60 mL/min/1.73 m², the CDDP dose will be reduced by 75%. When CDDP is infused into only part of the liver, the dose will be adjusted according to the proportion of tumor volume in the liver based on CT imaging.

To prevent CDDP-induced nephrotoxicity, patients will receive adequate intravenous hydration before and after infusion. Patients will be hospitalized for observation for at least 24 hours after infusion, with monitoring of vital signs, urine output, and laboratory tests to assess renal and hepatic function. Prophylactic antibiotics will be administered according to institutional protocols.

Radiotherapy

RT will be initiated preferably within 2 weeks after completion of HAIC. While specific technical approaches will be left to each institution's policy, two points will be standardized across all sites: (1) the irradiated target volume should primarily encompass the intravascular portion of the tumor while including, whenever feasible, the adjacent intrahepatic lesion; and (2) the prescribed dose should be 25 Gy in 5 fractions.

Under the supervision of a radiation oncologist at each institution, an appropriate margin of 5 to 20 mm will be added to the target volume to account for respiratory motion and setup error, aiming for a homogeneous dose distribution within the target while minimizing exposure to surrounding organs at risk. Dose constraints for organs at risk will follow institutional protocols, with typical limits including a mean dose to the normal liver (whole liver excluding the tumor) of <13 Gy, a maximum dose to the stomach or duodenum of <30 Gy, and a maximum spinal cord dose of <25 Gy.

Patients will be immobilized using customized vacuum cushions or equivalent devices to ensure reproducibility, with respiratory motion management (eg, breath-hold) used throughout treatment with each facility's policies. Either 3D conformal radiotherapy or intensity-modulated radiotherapy may be used. Image-guided radiotherapy, typically with daily cone-beam CT, is recommended for treatment delivery.

All treatment plans will be generated using a treatment planning system capable of Monte Carlo–equivalent dose calculation. Pretreatment plan quality assurance (QA) will be performed according to institutional standards, and daily machine QA will be conducted in accordance with international recommendations. Adaptive replanning will be considered if significant anatomical changes or tumor regression are observed during treatment.

Immunotherapy

Systemic immunotherapy will be initiated preferably within 2 weeks after the completion of RT, provided that adequate hematologic, hepatic, and renal function is confirmed. Regimens may include atezolizumab (1200 mg IV) plus bevacizumab (15 mg/kg IV) every 3 weeks; durvalumab (1500 mg IV) plus tremelimumab (300 mg IV once), followed by durvalumab monotherapy every 4 weeks (STRIDE [single tremelimumab regular interval durvalumab] regimen); or nivolumab (80 mg IV) plus ipilimumab (3 mg/kg IV) every 3 weeks, followed by nivolumab (240 mg IV) every 2 weeks or nivolumab (480 mg IV) every 4 weeks.

The choice of regimen will be determined at the discretion of the treating physician, taking into account patient comorbidities, previous treatment history, baseline organ function, and potential contraindications. In patients for whom combination immunotherapy is considered unsuitable, durvalumab monotherapy (1500 mg IV every 4 weeks) may be considered as an alternative option.

XSL•FO
RenderX

All regimens will adhere to recommendations outlined in established guidelines for systemic therapy in HCC, including the American Society of Clinical Oncology guideline for advanced HCC [23] and the Japan Society of Hepatology's 2021 clinical practice guidelines [6], as well as the European Society for Medical Oncology Clinical Practice Guideline [24].

Treatment will continue until radiologically confirmed disease progression, unacceptable toxicity, or loss of clinical benefit, whichever occurs first. Temporary dose delays, reductions, or regimen modifications will be permitted according to protocol-specific criteria and clinical judgment. All modifications, interruptions, or discontinuations will be documented in detail, including the reason and subsequent management.

Patients will undergo routine laboratory monitoring, including complete blood count, liver and renal function tests, and coagulation parameters, at each treatment cycle. Imaging assessments (dynamic contrast-enhanced CT or magnetic resonance imaging [MRI]) will be performed in accordance with the study's assessment schedule to evaluate tumor response and detect potential treatment-related complications, such as immune-related adverse events (irAEs). Prompt initiation of appropriate management strategies (eg, corticosteroids or immunosuppressants) will be undertaken for clinically significant irAEs, following current consensus recommendations [25-27].

Data Collection

Collected data will include the following:

- Patient background—age, sex, Eastern Cooperative Oncology Group performance status, comorbidities, prior cancer history, and prior HCC treatments
- Baseline liver function—etiology of liver disease, presence of ascites or hepatic encephalopathy, Child-Pugh class, and relevant laboratory values
- HAIC—procedure date, infusion duration, and CDDP dose
- RT—start and completion dates, technique, treatment planning system, dose calculation algorithm, respiratory motion management, dose parameters, and dose-volume histogram metrics
- Immunotherapy—regimen, dosage, initiation date, and treatment duration
- Laboratory trends—serial laboratory test results during the treatment course
- Additional therapies—details of any subsequent locoregional or systemic treatments administered during follow-up

Assessment Schedule

Clinical and laboratory assessments will be performed at baseline, at HAIC administration, at the start of RT, and at the initiation of immunotherapy. Adverse events (AEs) will be graded according to the Common Terminology Criteria for Adverse Events (CTCAE; version 5.0) and assessed at each visit or whenever clinically indicated. Specific laboratory monitoring for irAEs, including liver function, renal function, thyroid function, and blood glucose, will be conducted regularly. Imaging (contrast-enhanced CT, MRI, or ultrasound) will be performed at baseline, at 6 weeks and 12 weeks after starting immunotherapy, and every 2 months thereafter (+1 or −1 week), with tumor response evaluated according to Response Evaluation Criteria in Solid Tumors (RECIST; version 1.1) and/or modified RECIST (mRECIST) criteria.

Safety Assessment

Safety outcomes will be evaluated descriptively for the overall study population. AEs will be summarized according to immunotherapy regimen where feasible. As this study is a prospective observational registry reflecting real-world clinical practice, no formal comparative analyses between immunotherapy regimens are planned.

Sample Size Estimation

The primary objective of this study is to evaluate the safety of a multidisciplinary treatment strategy. Accordingly, the planned sample size was determined based on feasibility considerations and the desired precision of descriptive estimates rather than formal hypothesis testing. With an anticipated enrollment of 38 patients, the incidence of major AEs and the ORR can be estimated with an acceptable level of precision, as reflected by the width of the corresponding CIs. No formal hypothesis testing is planned, and all analyses will be descriptive and exploratory in nature.

Statistical Analysis

All analyses will be performed using descriptive statistics to summarize patient characteristics, treatment delivery, and outcomes. Continuous variables will be presented as medians with IQRs or means with SDs, as appropriate. Categorical variables will be expressed as counts and percentages.

Time-to-event end points (PFS and OS) will be estimated using the Kaplan-Meier method, with median survival times and 95% CIs reported. Exploratory subgroup analyses may be performed using the log-rank test and Cox proportional hazards regression models, adjusting for baseline covariates such as liver function, tumor burden, and treatment regimen. Hazard ratios with 95% CIs will be calculated. In multicenter analyses, participating institutions may be included as stratification factors.

ORR will be calculated with exact binomial 95% CIs. Safety will be summarized by incidence, severity (CTCAE version 5.0), and relationship to study treatment, with results presented overall and stratified by treatment phase (HAIC, RT, and immunotherapy). The timing of onset, resolution, and reasons for treatment discontinuation due to AEs will also be reported.

Missing data will not be imputed; patients without an event at data cutoff will be censored at the date of last contact. Reasons for missing data will be recorded, and sensitivity analyses may be conducted under different missing data assumptions (eg, worst-case imputation for progression events).

Given the observational nature of the study, all analyses will be exploratory and hypothesis generating. No formal sample size calculation has been performed, and *P* values from statistical tests will be interpreted descriptively without adjustment for multiplicity. Statistical analyses will be conducted using the EZR (Saitama Medical Center, Jichi

XSL•FO
RenderX

Medical University), a graphical user interface for R (version 3.4.1; The R Foundation for Statistical Computing).

Monitoring and QA

No formal coordinating center has been established for this study. Instead, the principal investigator and participating investigators will jointly oversee all aspects of study conduct. Regular meetings (either in person or via teleconference) will be held to review study progress, verify the accuracy and completeness of collected data, and ensure adherence to the study protocol.

Any protocol deviations, data inconsistencies, or other study-related issues identified during these reviews will be documented in meeting minutes. Corrective actions will be determined through consensus among the investigators, and follow-up will be performed to confirm resolution.

Protocol Deviations

All deviations from the approved protocol will be documented by the responsible investigator at each participating site, including the date, details, and reasons for the deviation. Given the investigator-driven nature of this study, no formal deviation classification system will be applied; however, any deviations judged by the investigators to potentially impact patient safety or data reliability will be discussed collectively among the study team.

Corrective actions will be decided through investigator consensus and implemented promptly. Records of all deviations and their resolution will be maintained securely by each investigator and shared within the study group as necessary. As this is not a formally regulated clinical trial, deviations will not be reported to regulatory authorities but will be transparently described in the final study report or publication when relevant.

Dissemination Plan

The findings from this study will be disseminated through publication in peer-reviewed journals and presentations at national and international scientific meetings. While the raw dataset will not be made publicly available, anonymized summary data may be shared upon reasonable request and with investigator approval.

Results

End Points

Primary End Point

The primary end point is safety, defined as the incidence of treatment-related AEs graded according to the CTCAE (version 5.0), occurring during the treatment period after the last dose of any study treatment. Treatment-related AEs are those considered by the investigator to be possibly, probably, or definitely related to any component of the study regimen. AEs will be summarized descriptively for the overall cohort and, where feasible, stratified by immunotherapy regimen.

Secondary End Points

The secondary end point is PFS at 12 and 24 weeks, overall median PFS, OS, and ORR at 12 and 24 weeks. The ORR will be defined as the proportion of patients achieving a complete response or partial response as their best overall response, assessed according to the RECIST (version 1.1) and/or mRECIST for HCC, in which viable tumor is defined as uptake in the arterial phase of dynamic imaging. Tumor response will be evaluated by dynamic contrast-enhanced CT or MRI at baseline, at 6 and 12 weeks after initiation of immunotherapy, and every 2 months thereafter. In patients with measurable lesions at baseline, all target and nontarget lesions will be assessed, and any unmeasurable but evaluable lesions (eg, tumor thrombus) will be included in accordance with the RECIST (version 1.1) and/or mRECIST guidelines. PFS is defined as the time from initiation of treatment to the first documented disease progression per the RECIST (version 1.1) and/or mRECIST or death from any cause, whichever occurs first. OS is defined as the time from initiation of treatment to death from any cause. Patients without an event at the time of analysis will be censored at the date of the last disease assessment. Patients lost to follow-up will be censored at the date they were last known to be alive. Given the limited sample size, analyses of efficacy outcomes, including PFS and OS, will be descriptive and exploratory in nature, and no formal hypothesis testing is planned.

Trial Status

As of March 2026, patient enrollment had already commenced, and data collection was actively ongoing. Participating investigators continue to recruit eligible patients and record study data in accordance with the protocol schedule. Patient enrollment began in March 2025, and data collection and analysis are ongoing as of March 2026.

Discussion

The primary objective of this prospective registry study is to evaluate the safety and therapeutic effectiveness of a multidisciplinary treatment approach that combines locoregional therapies and immunotherapy in patients with MVI-positive HCC.

In Japan, both the incidence and mortality rates of HCC have declined over recent decades, largely owing to the decreasing prevalence of viral hepatitis. According to the 24th Nationwide Follow-up Survey Report on Primary Liver Cancer, the numbers of patients with Vp3/4 and Vv2/3 tumor thrombi in 2016 to 2017 were approximately 980 and 570, respectively, underscoring the relative rarity of these high-risk cases. Furthermore, the prevalence of MVI-positive HCC has continued to decrease since that period, making such cases even less common in current clinical practice. Because of this low and declining incidence, conducting adequately powered prospective randomized controlled trials in this population is challenging, and a real-world, prospective registry design was, therefore, considered the most feasible and appropriate approach to capture treatment patterns and outcomes in a sufficiently large cohort.

In this multidisciplinary treatment strategy, RT was intended to achieve local tumor control within the portal vein and to preserve portal venous blood flow, thereby maintaining liver

XSL•FO
RenderX

function and eligibility for subsequent systemic therapy. Even modest tumor shrinkage may be clinically meaningful in patients with advanced HCC and major vascular invasion. A hypofractionated regimen of 25 Gy in 5 fractions in this study was selected to minimize treatment duration and allow timely initiation of immunotherapy, as longer radiotherapy courses or treatment-related toxicity could delay systemic therapy. This regimen represents a pragmatic balance between local disease control, safety, and uninterrupted transition to systemic therapy in real-world clinical practice.

A critical consideration is that MVI-positive HCC is associated with extremely poor prognosis, making the prompt initiation of aggressive treatment after diagnosis essential to maximize the potential for improved outcomes. Consequently, treatment strategies should prioritize modalities that are widely applicable and readily implementable in routine clinical settings. All components of this protocol—HAIC, RT, and immunotherapy—are established treatment methods in daily clinical practice in Japan, with minimal variation in availability or technical feasibility across institutions. Therefore, if this protocol demonstrates both safety and favorable efficacy, it could be rapidly adopted across many treatment centers, providing substantial clinical benefit for patients with difficult-to-treat MVI-positive HCC in Japan.

Recent reports suggest that multidisciplinary, team-based treatment may facilitate conversion to resectable status and improve long-term survival rates in patients with unresectable, MVI-positive HCC [28]. The sequential multidisciplinary collaboration strategy used in this study seeks to enhance therapeutic efficacy by leveraging the synergistic effects of each treatment modality. Furthermore, we anticipate that this approach will increase resectability and ultimately lead to better long-term outcomes.

Several limitations of this study should be acknowledged. As multiple immunotherapy regimens with different mechanisms of action and safety profiles are included, this study is not designed to distinguish regimen-specific safety signals, particularly given the limited sample size. This heterogeneity reflects real-world clinical practice in Japan, where several approved and reimbursed immunotherapy regimens with comparable efficacy are available and treatment selection is left to physician discretion. Therefore, safety findings should be interpreted descriptively and with caution.

In conclusion, this prospective registry study aims to generate valuable real-world evidence to inform treatment decisions for MVI-positive HCC and to contribute to the development of more effective therapeutic strategies in clinical practice.

Acknowledgments

The authors would like to thank all participating patients and medical staff involved in this study.

Funding

This research was funded in part by a grant from the Research Support Program of Hiroshima Prefectural Hospital (Kenritsu Hiroshima Byouin Kenkyu Shien Jigyo; April 2025 to March 2026).

Data Availability

The datasets generated and/or analyzed during this study are not publicly available due to patient privacy restrictions but are available from the corresponding author on reasonable request.

Authors' Contributions

Concept and design: HA
Funding acquisition: HA (lead), YD (supporting)
Investigation: HA (lead)
Writing—original draft: YD (lead), HA (equal)
Writing—review and editing: all authors
Approval of final manuscript: all authors

Conflicts of Interest

HA has received lecture fees from Eisai Co, Ltd and Chugai Pharmaceutical Co, Ltd. TK has received lecture fees from Eisai Co, Ltd, AstraZeneca, and Chugai Pharmaceutical Co, Ltd. MT has received lecture fees from Gilead Sciences and AbbVie. All other authors declare no other conflicts of interest.

References

1. Llovet JM, Ricci S, Mazzaferro V, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med 2008 Jul 24;359(4):378-390. [doi: 10.1056/NEJMoa0708857] [Medline: 18650514]
2. Kudo M, Finn RS, Qin S, et al. Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. Lancet 2018 Mar 24;391(10126):1163-1173. [doi: 10.1016/S0140-6736(18)30207-1] [Medline: 29433850]

XSL•FO
RenderX

3. Finn RS, Qin S, Ikeda M, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. N Engl J Med 2020 May 14;382(20):1894-1905. [doi: 10.1056/NEJMoa1915745] [Medline: 32402160]
4. Abou-Alfa GK, Lau G, Kudo M, et al. Tremelimumab plus durvalumab in unresectable hepatocellular carcinoma. NEJM Evid 2022 Aug;1(8):EVIDoa2100070. [doi: 10.1056/EVIDoa2100070] [Medline: 38319892]
5. Yau T, Galle PR, Decaens T, et al. Nivolumab plus ipilimumab versus lenvatinib or sorafenib as first-line treatment for unresectable hepatocellular carcinoma (CheckMate 9DW): an open-label, randomised, phase 3 trial. Lancet 2025 May 24;405(10492):1851-1864. [doi: 10.1016/S0140-6736(25)00403-9] [Medline: 40349714]
6. Hasegawa K, Takemura N, Yamashita T, et al. Clinical practice guidelines for hepatocellular carcinoma: the Japan Society of Hepatology 2021 version (5th JSH-HCC guidelines). Hepatol Res 2023 May;53(5):383-390. [doi: 10.1111/hepr.13892] [Medline: 36826411]
7. Reig M, Forner A, Rimola J, et al. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. J Hepatol 2022 Mar;76(3):681-693. [doi: 10.1016/j.jhep.2021.11.018] [Medline: 34801630]
8. Cheng AL, Qin S, Ikeda M, et al. Updated efficacy and safety data from IMbrave150: atezolizumab plus bevacizumab vs. sorafenib for unresectable hepatocellular carcinoma. J Hepatol 2022 Apr;76(4):862-873. [doi: 10.1016/j.jhep.2021.11.030] [Medline: 34902530]
9. He M, Li Q, Zou R, et al. Sorafenib plus hepatic arterial infusion of oxaliplatin, fluorouracil, and leucovorin vs sorafenib alone for hepatocellular carcinoma with portal vein invasion: a randomized clinical trial. JAMA Oncol 2019 Jul 1;5(7):953-960. [doi: 10.1001/jamaoncol.2019.0250] [Medline: 31070690]
10. Kudo M, Ueshima K, Yokosuka O, et al. Sorafenib plus low-dose cisplatin and fluorouracil hepatic arterial infusion chemotherapy versus sorafenib alone in patients with advanced hepatocellular carcinoma (SILIUS): a randomised, open label, phase 3 trial. Lancet Gastroenterol Hepatol 2018 Jun;3(6):424-432. [doi: 10.1016/S2468-1253(18)30078-5] [Medline: 29631810]
11. Ueshima K, Ogasawara S, Ikeda M, et al. Hepatic arterial infusion chemotherapy versus sorafenib in patients with advanced hepatocellular carcinoma. Liver Cancer 2020 Sep;9(5):583-595. [doi: 10.1159/000508724] [Medline: 33083282]
12. Wei X, Jiang Y, Zhang X, et al. Neoadjuvant three-dimensional conformal radiotherapy for resectable hepatocellular carcinoma with portal vein tumor thrombus: a randomized, open-label, multicenter controlled study. J Clin Oncol 2019 Aug 20;37(24):2141-2151. [doi: 10.1200/JCO.18.02184] [Medline: 31283409]
13. Kosaka Y, Kimura T, Kawaoka T, et al. Hepatic arterial infusion chemotherapy combined with radiation therapy for advanced hepatocellular carcinoma with tumor thrombosis of the main trunk or bilobar of the portal vein. Liver Cancer 2021 Apr;10(2):151-160. [doi: 10.1159/000513706] [Medline: 33977091]
14. Wu JS, Hong TC, Wu HT, et al. Hepatic arterial infusion chemotherapy and immune checkpoint inhibitors, alone or in combination, in advanced hepatocellular carcinoma with macrovascular invasion: a single-centre experience in Taiwan. J Gastrointest Oncol 2023 Apr 29;14(2):849-862. [doi: 10.21037/jgo-22-858] [Medline: 37201085]
15. Fu Y, Peng W, Zhang W, et al. Induction therapy with hepatic arterial infusion chemotherapy enhances the efficacy of lenvatinib and pd1 inhibitors in treating hepatocellular carcinoma patients with portal vein tumor thrombosis. J Gastroenterol 2023 Apr;58(4):413-424. [doi: 10.1007/s00535-023-01976-x] [Medline: 36894804]
16. Lou Y, Zhang X, Sun P, Chang X. Hepatic arterial infusion chemotherapy enhances the efficacy of lenvatinib plus PD-1 inhibitors in hepatocellular carcinoma patients with tumor thrombosis in the inferior vena cava and/or right atrium. Acad Radiol 2025 Feb;32(2):787-797. [doi: 10.1016/j.acra.2024.08.058] [Medline: 39278760]
17. Akyildiz A, Yavuz BG, Ismayilov R, et al. Comprehensive review of transarterial chemoembolization plus systemic therapy in advanced hepatocellular carcinoma. J Gastrointest Cancer 2025 Sep 27;56(1):194. [doi: 10.1007/s12029-025-01306-9] [Medline: 41015600]
18. Su CW, Teng W, Shen EYL, et al. Concurrent atezolizumab plus bevacizumab and high-dose external beam radiotherapy for highly advanced hepatocellular carcinoma. Oncologist 2024 Jul 5;29(7):e922-e931. [doi: 10.1093/oncolo/oyae048] [Medline: 38530254]
19. Kim HJ, Park S, Kim KJ, Seong J. Clinical significance of soluble programmed cell death ligand-1 (sPD-L1) in hepatocellular carcinoma patients treated with radiotherapy. Radiother Oncol 2018 Oct;129(1):130-135. [doi: 10.1016/j.radonc.2017.11.027] [Medline: 29366520]
20. Tojjari A, Yu J, Saeed A. Immunotherapy and radiation therapy combinatorial approaches in hepatocellular carcinoma. Cancers (Basel) 2024 Mar 5;16(5):1058. [doi: 10.3390/cancers16051058] [Medline: 38473415]
21. Yamaoka K, Kawaoka T, Fujii Y, et al. Multidisciplinary treatment for patients with advanced hepatocellular carcinoma complicated by Vp4 portal vein tumor thrombosis: combination of atezolizumab and bevacizumab after hepatic arterial infusion chemotherapy and radiotherapy: a case series. Medicine (Abingdon) 2024;5(6):e00325. [doi: 10.1097/MD9.0000000000000325]
22. Naruto K, Kawaoka T, Kodama K, et al. Efficacy and safety of chemoradiation therapy using one-shot cisplatin via hepatic arterial infusion for advanced hepatocellular carcinoma with major macrovascular invasion: a single-arm retrospective cohort study. BMC Gastroenterol 2022 Jun 2;22(1):275. [doi: 10.1186/s12876-022-02359-x] [Medline: 35655156]
23. Gordan JD, Kennedy EB, Abou-Alfa GK, et al. Systemic therapy for advanced hepatocellular carcinoma: ASCO guideline update. J Clin Oncol 2024 May 20;42(15):1830-1850 [FREE Full text] [doi: 10.1200/JCO.23.02745] [Medline: 38502889]

XSL•FO
RenderX

24. Vogel A, Chan SL, Dawson LA, et al. Hepatocellular carcinoma: ESMO clinical practice guideline for diagnosis, treatment and follow-up. Ann Oncol 2025 May;36(5):491-506. [doi: 10.1016/j.annonc.2025.02.006] [Medline: 39986353]
25. Schneider BJ, Naidoo J, Santomasso BD, et al. Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: ASCO guideline update. J Clin Oncol 2021 Dec 20;39(36):4073-4126. [doi: 10.1200/JCO.21.01440] [Medline: 34724392]
26. Management of immunotherapy-related toxicities (version 1). National Comprehensive Cancer Network. 2025. URL: https://www.nccn.org/guidelines/guidelines-detail?category=3&id=1486 [accessed 2025-08-18]
27. Haanen J, Obeid M, Spain L, et al. Management of toxicities from immunotherapy: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Ann Oncol 2022 Dec;33(12):1217-1238. [doi: 10.1016/j.annonc.2022.10.001] [Medline: 36270461]
28. Yuan T, Liu J, Lv X, Zhang E, Huang Z. Liver resection for hepatocellular carcinoma with vascular invasion: current insights and future perspectives. Oncol Transl Med 2025. [doi: 10.1097/ot9.0000000000000085]

Abbreviations

AE: adverse event
CDDP: cisplatin
CT: computed tomography
CTCAE: Common Terminology Criteria for Adverse Events
FOLFOX: folinic acid, fluorouracil, and oxaliplatin
HAIC: hepatic arterial infusion chemotherapy
HCC: hepatocellular carcinoma
IRB: institutional review board
mRECIST: modified Response Evaluation Criteria in Solid Tumors
MRI: magnetic resonance imaging
MVI: macroscopic vascular invasion
ORR: objective response rate
OS: overall survival
PFS: progression-free survival
QA: quality assurance
RECIST: Response Evaluation Criteria in Solid Tumors
REFLECT: Study of Lenvatinib vs Sorafenib in First-Line Treatment of Patients With Unresectable Hepatocellular Carcinoma
RT: radiation therapy
SHARP: Sorafenib HCC Assessment Randomized Protocol
STRIDE: Single Tremelimumab Regular Interval Durvalumab
Vp3/4: a tumor thrombus in the first-order branch or main trunk of the portal vein
Vv2/3: a tumor thrombus in the inferior vena cava or hepatic veins

Edited by J Sarvestan; submitted 06.Sep.2025; peer-reviewed by R Ismayilov; accepted 18.Feb.2026; published 31.Mar.2026.

Please cite as:
Doi Y, Aikata H, Kosaka Y, Nakahara T, Kodama M, Hieda M, Hashimoto M, Nakahara H, Takahashi I, Kakizawa H, Mori N, Tsuji K, Imano N, Murakami Y, Sueda S, Kawaoka T, Tsuge M, Oka S
Multidisciplinary Treatment With Hepatic Arterial Infusion Chemotherapy, Radiotherapy, and Immunotherapy for Advanced Hepatocellular Carcinoma With Major Vascular Invasion: Prospective Registry Protocol
JMIR Res Protoc 2026;15:e82992
URL: https://www.researchprotocols.org/2026/1/e82992
doi:10.2196/82992

© Yoshiko Doi, Hiroshi Aikata, Yumi Kosaka, Takashi Nakahara, Michiyo Kodama, Masashi Hieda, Masakazu Hashimoto, Hideki Nakahara, Ippei Takahashi, Hideaki Kakizawa, Nami Mori, Keiji Tsuji, Nobuki Imano, Yuji Murakami, Saki Sueda, Tomokazu Kawaoka, Masataka Tsuge, Shiro Oka. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Occupational Exposure to Resorcinol and Thyroid-Disrupting Effects: Protocol for an Exploratory Field Study in French Hairdressers

Anca Radauceanu[1], MD; Amandine Cambrai-Erb[2], PhD; Benedicte Adet[1], ADN; Marie-Isabelle Nioule[1], AS; Flavien Denis[2], AS; Romain Pons[1], PhD

[1]Department of Occupational Epidemiology, French Research and Safety Institute for the Prevention of Occupational Accidents and Diseases (INRS), 1 rue du Morvan, Vandoeuvre-lès-Nancy Cedex, France

[2]Department of Toxicology and Biomonitoring, French Research and Safety Institute for the Prevention of Occupational Accidents and Diseases (INRS), Vandoeuvre-lès-Nancy Cedex, France

Corresponding Author:
Anca Radauceanu, MD
Department of Occupational Epidemiology, French Research and Safety Institute for the Prevention of Occupational Accidents and Diseases (INRS), 1 rue du Morvan, Vandoeuvre-lès-Nancy Cedex, France

Abstract

Background: All around the world, the hairdressing sector constitutes a major occupational group, including about 90% women, most of whom are of reproductive age. Hairdressers are continuously exposed to numerous chemicals used in hair products, including endocrine-disrupting compounds such as resorcinol, parabens, phthalates, and UV filters. Few biomonitoring studies have explored occupational exposure to endocrine disruptors in hairdressers, and no data were found on their impact on the thyroid hormone system. Resorcinol is an oxidative hair dye with thyroid-disrupting properties that decrease thyroid hormone synthesis and could alter neurodevelopmental functions during fetal and perinatal stages in case of maternal exposure.

Objective: This study aims to assess the occupational exposure to resorcinol in French hairdressers and analyze the relationship with biological thyroid parameters, taking into account the occupational exposure to other potential thyroid disruptors (parabens and UV filters like benzophenone and cinnamates).

Methods: An exposed-unexposed cross-sectional study is proposed involving female hairdressers aged 18 to 45 years (working in hair salons) compared to occupationally unexposed controls (employed in office activities), who are recruited within 14 French occupational health centers. The hairdressers are followed during a 5-day working week to assess exposure data at both the individual level and the salon level. Urinary samples for the measurement of thyroid disruptors (resorcinol, parabens, metabolites of ethylhexyl methoxycinnamate, and benzophenone-3) are collected at 6 time points (before the day 1 shift, before and after the day 3 and day 4 shifts, and before the day 5 shift). Daily work tasks and use of hair products are self-reported within the workplace, and a complete inventory of hair products within the salon is carried out. Thyroid disruption effects are assessed by measuring blood thyroid parameters: triiodothyronine, thyroxine, thyroid-stimulating hormone, thyroglobulin, thyroperoxidase, and thyroglobulin antibodies. To assess nonoccupational exposure to thyroid disruptors and other confounding factors, information on sociodemographic data, place of residence, food and tobacco consumption, personal use of care products, professional career, and medical history is collected through questionnaires. Regarding statistical analysis, urinary samples from hairdressers and controls will be compared, and adjusted multivariable models will be used to analyze health outcomes.

Results: The study duration extends from 2022 to 2027. As of December 2025, 9 occupational health centers have enrolled 66 hairdressers (employed in 54 hair salons) and 30 occupationally unexposed participants.

Conclusions: The results will represent the first data on occupational exposure to resorcinol in France and its relationship with thyroid hormones in hairdressers. Following a multidisciplinary approach that includes biomonitoring, epidemiology, and exposure data collection at both the hairdressers and salon levels, this study enables an in-depth assessment of exposure to the thyroid disruptors in the workplace. Together with the inventory of hair products, these results may enhance the tools for chemical risk assessment and prevention in hair salons.

International Registered Report Identifier (IRRID): DERR1-10.2196/65833

(JMIR Res Protoc 2026;15:e65833) doi:10.2196/65833

XSL•FO
RenderX

KEYWORDS

resorcinol; hairdresser; endocrine-disrupting chemicals; occupational exposure; thyroid gland; biomonitoring; cosmetics; hair products; thyroid disruptors; disruptors of thyroid hormones

Introduction

Principles of Cosmetic Product Regulation in the European Union: Highlight on Endocrine-Disrupting Chemicals

In the European Union, the legal framework for cosmetic products relies on the Cosmetic Products Regulation Number 1223/2009 [1]. The main purpose of the cosmetic product legislation is to ensure the safety of cosmetics for consumers through rigorous science-based regulation. More specifically, ingredient regulation is based on a "risk-based" approach, implemented through the establishment of positive and negative lists of ingredients [2]. The Scientific Committee on Consumer Safety (SCCS) provides guidance on the safety of cosmetic ingredients for consumers only, but not for professional users. In the European Union, there is no specific management of endocrine-disrupting chemicals (EDCs) in cosmetics. These substances, defined by the World Health Organization as "exogenous substances or mixtures that alter the function(s) of the endocrine system and consequently cause adverse effects in an intact organism, or its progeny, or (sub)populations" [3], are evaluated and regulated on a compound-by-compound basis by the SCCS, which provides safety assessments of cosmetic products at certain concentrations [4]. A number of restrictions can be applied when the substances are also classified as carcinogenic, mutagenic, or toxic for reproduction or considered substances of very high concern (SVHC), defined as substances that may have serious and often irreversible effects on human health and the environment. The European Union recently introduced new hazard classes to reinforce the classification, labeling, and packaging (CLP) of substances and mixtures regulation, which serves as the central piece of the chemical legislation addressing endocrine disruptors for human health and the environment [5]. Thus, 2 hazard classes and corresponding statements for endocrine disruptors have been created based on their scientific strength of evidence: known or presumed endocrine disruptors (category 1) and suspected endocrine disruptors (category 2) for both human health and the environment.

Within the European Union, 6 national competent authorities compile information about the current status of substances identified as EDCs or those under evaluation for endocrine-disrupting properties through 3 lists, which are updated at least biannually and hosted on a dedicated website [6]. The 3 endocrine disruptor lists consist of the following: list 1, substances identified as endocrine disruptors that have undergone the full evaluation of endocrine-disrupting properties; list 2, substances that are currently under evaluation; and list 3, substances considered endocrine disruptors at the national level in one of the participating Member States.

EDCs in Hair Products: Focus on Thyroid Hormone System–Disrupting Chemicals

Overview

Cosmetic hair products belong to a wide spectrum of items ranging from shampoo, conditioner, and hair dyes to fragrances, solutions, or sprays for hair bleaching, perming, straightening, styling, and modeling. The endocrine disruptors in hair products are compounds that serve various functions (preservatives, hair dyeing, perfuming, UV protection, etc), and the most commonly reported endocrine disruptors in the literature are parabens, phenols, bisphenols, phthalates, cinnamates, and benzyl salicylate [7-9].

Brief Inventory of EDCs in Professional Hair Products

In 2020, we conducted a brief inventory of EDCs identified among the compounds of 1157 hair cosmetic products available for sale to hairdressers on professional websites [10].

In this inventory, the content of hair products revealed the presence of resorcinol in 40% of hair dyes, parabens in around 20% of styling waxes as well as gels and lacquers for men, salicylic acid in 20% of shampoos, cyclopentasiloxane in 15% of hair care products and conditioners, and butylated hydroxytoluene in 10% of hair gels or sprays. The most frequently used UV filters and absorbers were benzyl salicylate (25% of hair products), benzophenones (15% - 20% of shampoos for colored hair and styling products), and ethylhexyl methoxycinnamate (10% - 12% of hair gels, lacquers, sprays, hair care, and conditioners).

Thyroid Hormone System and Thyroid Dysfunction

Thyroid hormones are involved in numerous physiological processes as regulators of metabolism, bone remodeling, cardiac function, fetal neurodevelopment, and mental status [11,12]. Thyroid function is regulated through the hypothalamus-pituitary-thyroid axis in which the hypothalamus releases thyrotropin-releasing hormone to stimulate the anterior pituitary gland to produce thyroid-stimulating hormone (TSH), which in turn stimulates the production and secretion of thyroid hormones (thyroxine and triiodothyronine) from the thyroglobulin stored in the thyroid follicles.

The incidence of thyroid dysfunction has recently increased worldwide, particularly among women, since the ratio of thyroid dysfunction among women of reproductive age compared to men is 4:1 [13]. Hypothyroidism is the most frequent, followed by hyperthyroidism, which is also associated with a 4.5-fold increased risk of thyroid cancer [14]. Overall, 5% to 9% of adults have subclinical thyroid disease, and 0.8% to 7.5% of adults have overt clinical thyroid disease [13,15]. A 2014 meta-analysis of older European studies reported a mean prevalence of total thyroid dysfunction of 3.82% (0.75% for hyperthyroidism and 3.05% for hypothyroidism), while the prevalence of undiagnosed thyroid dysfunction was almost 7% [16].

XSL•FO
RenderX

In France, 3 studies estimated the prevalence of hyperthyroidism and hypothyroidism. The first study was conducted from 1994 to 1995, based on the measurement of serum thyrotropin (TSH), and estimated a prevalence of 5.7% for hyperthyroidism in participants aged between 35 and 60 years [17]. A recent study, conducted in 2020, estimated prevalences of treated hyperthyroidism and hypothyroidism, according to national health care data, at 0.17 and 4.45 per 100 inhabitants, respectively [18]. Prevalence was higher for women than for men and increased with age. Interestingly, French data on treated acquired hypothyroidism indicate a decreased incidence for both female and male participants from 2014 to 2019, for all age groups over 15 years of age [19].

Due to the serious health consequences and financial burden associated with thyroid diseases, the identification of risk factors, especially those that are modifiable, is of utmost public health interest. There is compelling evidence from in vivo, in vitro, and epidemiological studies that a growing number of chemicals can disrupt thyroid function and cause adverse health effects [20,21]. Thyroid-disrupting health effects from exposure to care and other consumer products have been reported, and their continuous consumption nowadays negatively affects susceptible populations, such as pregnant women and their offspring, with long-term adverse consequences, such as attention or deficit or hyperactivity disorders, autism spectrum disorders, and developmental delays in the offspring [22].

Thyroid-Disrupting Chemicals in Hair Products: Focus on Resorcinol

Among the compounds in hair products, parabens have been shown to have a disruptive effect on the thyroid in both experimental and human studies [21,23], whereas for benzophenone UV filters, the relevance of human exposure needs to be studied further [24]. However, increased urinary benzophenone-3 has been associated with decreased total triiodothyronine in pregnant women [25]. Regarding cinnamate UV filters, only experimental studies have indicated interferences with thyroid function [9]. Resorcinol is an oxidative hair dye (maximum allowed concentration up to 1.25%) and also an ingredient in hair lotions and shampoos (maximum allowed concentration up to 0.5%) [26]. Additionally, resorcinol is used in other industrial fields for its properties (antioxidant, antiseptic, enhancing strength of rubber and wood products, UV protection, flame retardant, colorant, and food additive). It is also found in the environment as a result of anthropogenic activities, as a metabolite of tannic acid frequently found in foods and beverages and as a component of tobacco smoke. The global resorcinol market stood at nearly 70,000 tons in 2022 and is expected to grow at a compound annual growth rate of 3.94% over the forecast period until 2032 [27].

In the professional use of resorcinol or resorcinol-containing products, the main route of exposure is skin contact, as resorcinol is a molecule of low volatility. Nevertheless, inhalation may occur if resorcinol is aerosolized, as well as oral exposure through hand-to-mouth contact.

After absorption, resorcinol is mainly excreted in urine, and urinary resorcinol serves as a valuable marker of occupational exposure [28]. Since resorcinol has a short half-life and no tissue bioaccumulation, it is expected that nearly all absorbed resorcinol will be excreted in urine within 24 hours after exposure.

In 2020, the French Agency for Food, Environmental and Occupational Health & Safety submitted a proposal for identifying resorcinol as an SVHC based on the criteria set out in the European regulation REACH (Registration, Evaluation, Authorization, and Restriction of Chemicals). This report was based on the scientific evidence of probable serious effects on human health in relation to its thyroid-disrupting potential, as shown in numerous in vitro and in vivo studies [29]. Resorcinol acts principally via the inhibition, in a dose-dependent manner, of the key enzyme involved in thyroid hormone synthesis, thyroperoxidase, resulting in the blockage of thyroid hormone synthesis and decreased serum thyroxine level. Consistent with thyroperoxidase inhibition, resorcinol inhibits intrathyroid iodine uptake and organification in humans, rodents, and more recently, fishes and lizards, leading to clinical hypothyroidism [29,30]. Low thyroid hormone levels trigger a negative feedback mechanism mediated by the hypothalamic thyrotropin-releasing hormone, which stimulates the synthesis and release of TSH in the pituitary gland to compensate for insufficient levels of thyroid hormones (Figure 1). Compensatory mechanisms within the thyroid gland, in response to the increase in TSH (thyroid hyperplasia, thyroglobulin depletion, increased iodine absorption), are unable to restore normal thyroid function because of the direct inhibition of thyroperoxidase.

Since resorcinol has been used as a dermally applied ointment to treat mainly skin ulcers and, more rarely, acne, clinical hypothyroidism has been reported in several human cases after repeated application, as well as the reversal of symptoms and goiter after cessation of the exposure [29]. In animals, effects on the thyroid were consistent with human effects, together with histopathological changes in the thyroid and probable reproductive and neurodevelopmental toxicity. For humans, the relationship between the inhibition of thyroperoxidase, the decrease in maternal circulating thyroid hormones, and prenatal and postnatal neurodevelopmental alterations in offspring is considered to be established with a high level of evidence. Finally, even though resorcinol is a chemical with clear thyroid-disrupting properties in humans, its thyroid-disrupting effects have not been explored in the context of occupational exposure.

XSL•FO
RenderX

Figure 1. Thyroid-disrupting mechanism of resorcinol involving thyroperoxidase inhibition leading to impaired iodination of thyroglobulin in the thyroid gland, reduced thyroid hormones (triiodothyronine and thyroxine) synthesis, and compensatory increase in thyroid-stimulating hormone (TSH) levels.

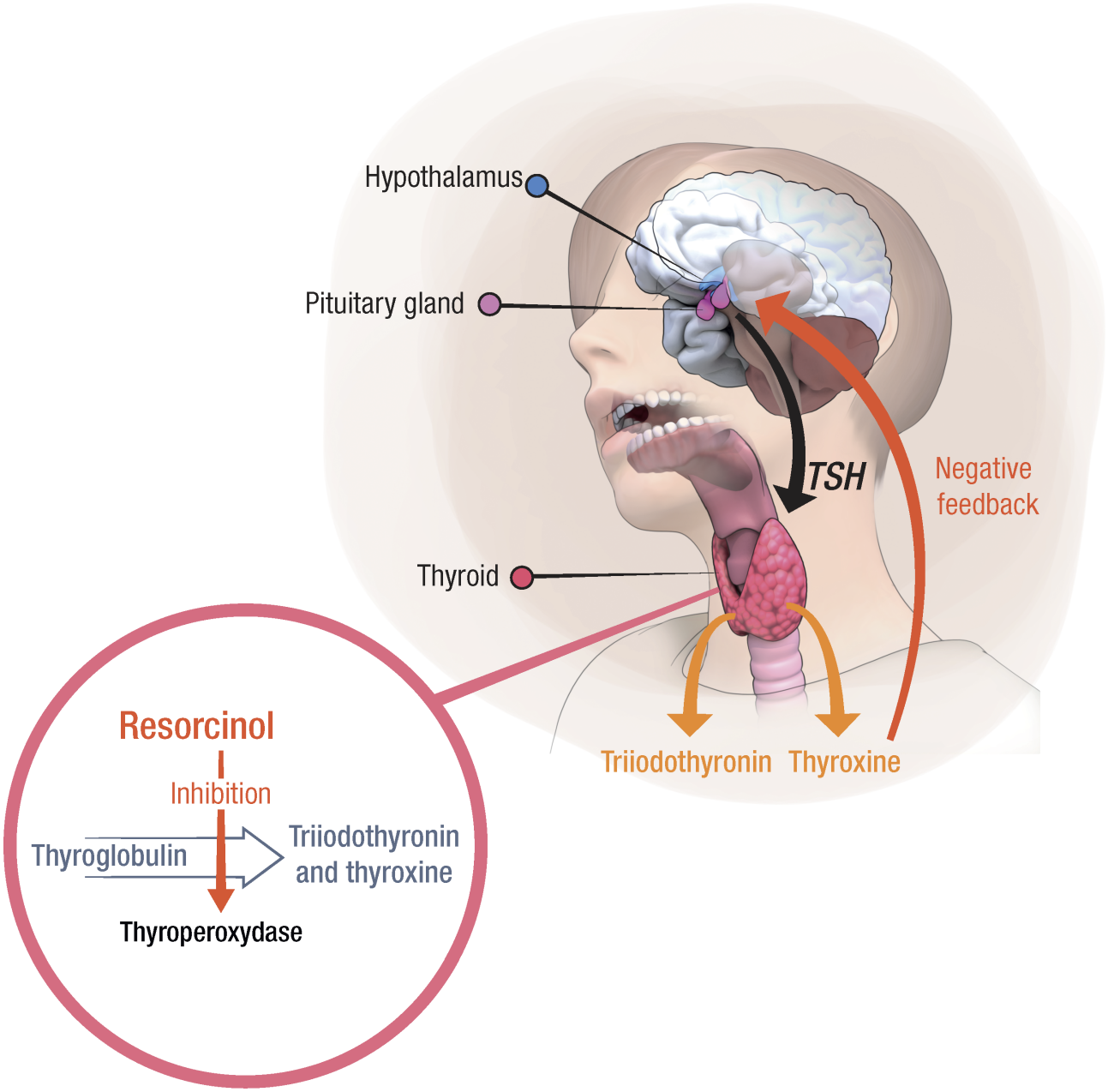

Hairdressing, Occupational Exposure to Endocrine-Disrupting Chemicals, and Health End Points

Around the world, the hairdressing sector constitutes a major occupational group employing a majority of women, many of whom are of reproductive age [31]. In France, in 2022, there were 180,000 hairdressers, and about 90% of them were women [32]. This occupational group is continuously exposed to numerous chemical ingredients found in hairdressing products, as well as to cleaning products used in the workplace [33]. In contrast to most customers who use cosmetics only for a few minutes each day, hairdressers may be exposed to these products for up to 8 hours per day and 5 or 6 days per week, throughout their careers. The average frequency of use by consumers, therefore, severely underestimates the exposure of hairdressers to daily skin contact with hairdressing products [34], and the safety assessment for consumer use differs from that for professional use. For instance, in an SCCS report in the Netherlands undertaken on 56 hair dyes, more than 37% of the products were considered safe for consumers but exhibited margins of safety that were considered as risky for professionals [35]. The routes of exposure differ depending on the type and use of the products. For example, during the hair coloring stages, powder may be inhaled during mixing, skin contact may occur when the mixtures are applied, and gases or vapor could be inhaled due to evaporation [36]. Workplace characteristics (ventilation and size of the salon), working techniques (haircutting after applying product mixtures and use of protective equipment), hairdressers' information, and prevention education are also relevant factors influencing the exposure of hairdressers [36].

XSL•FO
RenderX

The occupational hazards of cutaneous and respiratory allergic diseases in hairdressing are well described, but differences have been reported between information declared in surveys and work observations, particularly regarding the wearing of gloves, suggesting their use was not systematic during the steps of hair dyeing following application, such as washing or rinsing and haircutting [36]. Resorcinol was therefore found in most handwash samples from hairdressers after the application of hair dye, as well as after the haircut that follows hair dye, once gloves had been removed [37,38].

Overall, the inappropriate use of gloves during tasks (sharing and reusing, wearing jewelry, unsuitable cheap gloves), inadequate skin protection, working in wet environments, likely present to a large extent in hairdressers, impaired epidermal barrier function, aggregated exposure (exposure to the same ingredient via different products) among others, and facilitated the penetration of hazardous substances into the skin, consequently increasing the risk of local and possibly systemic adverse effects [34]. Recent studies assessing chemical exposure among hairdressers are limited and have mainly focused on air contaminants in hairdressing salons, such as volatile organic compounds, phthalates, or particulate matter [31,39]. Furthermore, there are few urinary biomonitoring studies that enable an evaluation of overall exposure, including the cutaneous route of exposure to hair products [31,34]. Among these studies, few have explored occupational exposure to endocrine disruptors in hairdressers. In 2019, Quiros-Alcala et al [31] reviewed 2 studies, one focusing on urinary resorcinol among hairdressers and the other on urinary phthalate among hairdressing apprentices. In the first study, which aimed to quantify environmental and occupational exposure to resorcinol in urine in Finland, no difference was observed between hairdressers and nonoccupationally exposed women [40]. Nevertheless, concentration levels of resorcinol were higher in postshift samples compared to the morning sample taken after at least 1 day off from work. In the second study, urinary phthalate biomarker levels were higher in Slovakian hairdressing apprentices compared to controls, and the latter were associated with negative outcomes in pulmonary function [41]. More recently, Kolena et al [42] published a second Slovakian study that confirmed significant exposure to phthalates related to hairdressing among apprentice hairdressers and found a relationship between lung function tests and anthropometric parameters adjusted for occupational exposure to phthalates. The authors suggested that exposure to phthalates could affect the body composition related to fat and fat-free mass index, which could secondarily affect lung function. In comparison to controls, US hairdressers serving an ethnically diverse clientele had concentrations of urinary phthalate metabolites 10 times higher after their shift [43]. These hairdressers were also exposed to 24 additional chemicals with median peak areas twice as high, including methylparaben, ethylparaben, propylparaben, and 2-naphthol [44].

As the hairdressing workforce includes a large majority of women of childbearing age, most of the research has focused on reproductive effects. Chemicals in hair products have changed significantly over time, but an updated review including epidemiological studies published in the last 2 decades still reported scarce adverse reproductive outcomes (poor neonatal health indicators and perinatal adverse events) [45]. A recent Brazilian study highlighted a higher number of premature deaths due to reproductive system cancers in female hairdressers [46]. Notably, the association of the hairdressing occupation with cancer and reproductive diseases has been consistently reported [33,45], leading the International Agency for Research on Cancer (IARC) in 2010 to classify the occupational exposure of hairdressers as probably carcinogenic (IARC group 2A).

In summary, the state of the art on occupational exposure to endocrine disruptors in hairdressers highlights the small number of recent studies, the rare reports on the type and magnitude of exposure, the mixture effects of all pollutants, the limitations in establishing the respective contributions of various chemicals, and the causal relationships that do not enable proper assessment of the endocrine-disrupting effects of occupational exposures. Conducting epidemiological studies on this topic is challenging, since endocrine disruptors are ubiquitous compounds, the health effects could be delayed, mixtures need to be considered to assess the endocrine disruption of ingredients in the final product, and a low enough magnitude of exposure, such as dermal exposure to cosmetics, seems not as hazardous as dietary intake [7]. In an attempt to address these gaps in the existing literature and increase current knowledge regarding thyroid-disrupting substances in this understudied population, we propose an exploratory field study in French hairdressers to investigate occupational exposures to the main thyroid disruptors in hair products and their relationship with blood thyroid parameters.

Study Objectives

Based on the presumed physiopathological mechanism presented in Figure 2, this study aims to address 2 objectives:

- The first objective is to assess occupational exposure to resorcinol, based on urinary measurements in hairdressers working in hair salons.
- The second objective is to explore the association between occupational exposure to resorcinol in hairdressers and biological thyroid parameters, taking into account occupational exposure to other endocrine disruptors, such as parabens and UV filters.

XSL•FO
RenderX

Figure 2. Thyroid-disrupting mode of action of resorcinol and subsequent study objectives. EDC: endocrine-disrupting chemical; TSH: thyroid-stimulating hormone.

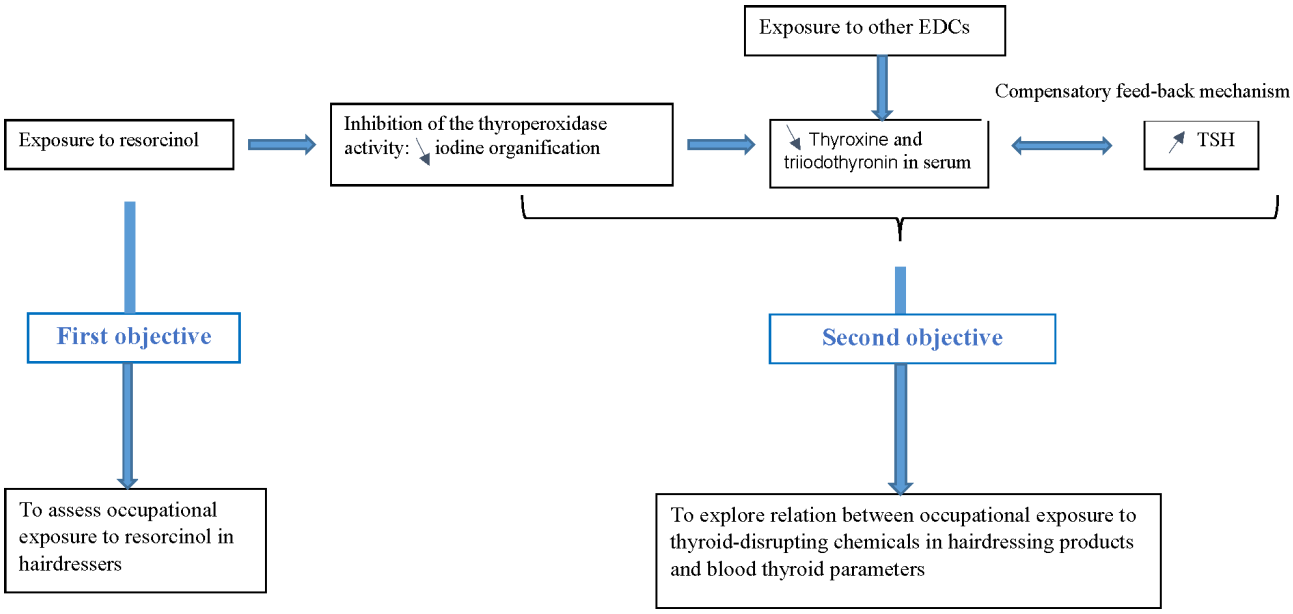

Methods

Study Design and Protocol

We designed an occupationally exposed-unexposed cross-sectional study in which hairdressers are followed up during a 5-day working week after at least 1 day of rest. During the working week, urine samples are collected before “day 1” shift (after at least 1 day off), before and after “day 3” and “day 4” shifts, and before “day 5” shift (Multimedia Appendix 1). This sampling strategy relies on the toxicokinetics of resorcinol and the occupational biomonitoring approach tailored for work shifts [40,47].

Few data on resorcinol toxicokinetics are available [29]. In humans, a biological half-life of 4.65 hours was predicted through modeling, and a urinary excretion half-life of 31 hours was reported after percutaneous penetration. For chemicals with short half-lives, within-subject biological concentrations vary strongly over time, and this temporal variability could bias the chemical levels during the toxicologically relevant exposure window [48]. Collecting several samples for each participant and performing intraworker repeated exposure measurements could provide a less biased design by reducing variability in exposure assessment.

Thus, the first sample collected after at least 1 day off work represents the individual background level of the hairdresser, while the following time points of urine collection are based on the anticipated exposure time throughout the work week to properly assess occupational exposure. Finally, we also considered the feasibility of the follow-up within hair salons and designed a sampling strategy that could provide a sufficient number of samples to overcome the issues of within-subject exposure variability. Daily tasks are self-reported by each hairdresser during the working week (daily work sheet), and a visit to the hairdressing salon is conducted by the research team. For unexposed women, only 2 urinary samples are collected: one at the beginning and one at the end of the working week.

At the inclusion medical visit, a blood sample is collected, a medical questionnaire is administered, and a self-administered questionnaire is provided to all participants. At the end of the follow-up, all daily work sheets and the self-administered questionnaire are collected by the research team. The medical questionnaire, self-administered questionnaire, and self-reported daily work sheet are provided in Multimedia Appendix 2 (available only in French). Finally, this protocol was designed to adhere to applicable guidelines for reporting observational studies (STROBE [Strengthening the Reporting of Observational Studies in Epidemiology] statement; Checklist 1) [49]. The completed checklist of items that should be included in reports of observational studies is presented in Multimedia Appendix 2.

Study Population

Inclusion and Exclusion Criteria

The occupationally exposed population consists of female hairdressers aged 18 to 45 years, not being apprentices, working in traditional hairdressing salons, and having a seniority of at least 1 year. Home-dressing and salons with other concomitant activities (nail salon, beauty salon, etc) practiced in the same room are not included.

The occupationally unexposed population (controls) consists of female participants aged 18 to 45 years who are not occupationally exposed to hair products or other chemicals and are employed in office activities, such as secretary, computer specialist, and accounting. Exclusion criteria for both exposed and unexposed groups include a history of thyroid disorders, ongoing or past pregnancy within the last 12 months, current breastfeeding, and a history of severe or chronic diseases. All participants are asked about the date of their last personal hair dyeing (not allowed within 7 days before or during the working week follow-up) and the use of iodine products for medical purposes (imaging or dermal products not allowed within 1 month before inclusion or during the working week follow-up).

XSL•FO
RenderX

In case of recent personal hair dyeing or use of iodine products, inclusion should be postponed.

Recruitment and Inclusion of Participants

The recruitment of the study population is carried out by occupational health professionals belonging to occupational health centers spread across different French regions. It is important to note that the health monitoring of employees within the occupational health centers is required by law in France, and compulsory medical examinations are regularly scheduled throughout an employee's working life, starting from the moment of hiring. A specific partnership covering this research was implemented with volunteer health centers. Fourteen occupational health centers (6 from the beginning of the study and 8 additional centers since 2024) volunteered to participate in this study because of their interest in EDCs and work. The recruitment of hairdressers is carried out by the occupational health centers based on a three-step procedure: (1) identification in the medical records of eligible hairdressers who meet the inclusion criteria of age, professional activity, and medical history; (2) identification of the salons where eligible hairdressers are working to obtain verbal approval from the salon manager to participate in the research; and (3) in the case of the manager's approval, hairdressers are informed about the study and then asked if they are willing to volunteer to participate. Voluntary hairdressers are invited to attend the medical inclusion visit within the occupational health centers. In cases of refusal to participate, nonidentifying data on the reasons for the refusal are collected, if possible, from the salon's manager and the hairdresser. The recruitment of unexposed participants is carried out during the compulsory medical examination within the occupational health centers, and each eligible woman is invited to participate in the study.

Ethical Considerations

This study was approved by the French Ethics Committee (Comité de Protection des Personnes Nord Ouest III, reference number: 2022 - 88). The processing of the study data was the subject of a declaration to the French Data Protection Authority in September 2021 (Commission Nationale de l'Informatique et des Libertés, registration number 2223653 v 0, September 21, 2021). Participation is voluntary, and each participant is given an information letter to review before volunteering. Each participant provides written, informed consent to participate in the study. The right to privacy and confidentiality of participants is upheld throughout the research. No compensation is provided to participants.

Data Collection

Once consent is obtained, data collection takes place at different time points—during the inclusion visit within the occupational health center and during a 5-day work week in the hairdressing salon. Data are collected at both individual and salon levels: at the individual level, relating to sociodemographics, medical history, environmental and professional exposure, and blood thyroid parameters; and at the hairdressing salon level, concerning the assessment of exposure throughout a salon visit and the inventory of chemical products.

During the inclusion visit, a medical questionnaire is administered face-to-face by the occupational health professional to collect data about participants' age, body mass, family history of thyroid disorders, personal use of current medicines, reproductive history, dermal disorders, work characteristics and schedules, and the date of the last personal hair dyeing. Then, a self-administered questionnaire is given to the participant to be completed at home and returned to the study team at the end of the 5-day follow-up. The self-administered questionnaire collects information on sociodemographic factors (educational level and marital status), housing data (address, housing type and age, number of persons, surroundings, etc), tobacco status, current food, water and alcohol consumption, current use of personal care and household products, and career history.

One blood sample is collected at the inclusion visit to measure thyroid hormone–related parameters, including free and total triiodothyronine and thyroxine, TSH, thyroglobulin, thyroperoxidase antibodies, and thyroglobulin antibodies. Because of the circadian rhythms reported in thyroid hormones [50], samples are collected for all participants in the morning during the same time slot (8 AM-12 AM). Serum TSH, free triiodothyronine, and free thyroxine are measured by Access Immunoassay Systems with a DXI Beckman Coulter, Inc, analyzer (reference B63284 Access TSH 3rd IS; reference A13422 Access Free T3; reference 33880 Access Free T4) [51]. Serum thyroglobulin and antibodies to thyroperoxidase and thyroglobulin are measured by chemiluminescence electromagnetic immunoassays with a Cobas e Roche Diagnostics immunoassay analyzer (reference 08906564190 Elecsys Tg II; reference 07026935190 Elecsys Anti-TPO; reference 09005021190 Elecsys Anti-Tg) [52]. Finally, total triiodothyronine and total thyroxine in serum are quantified by a chemiluminescent microparticle immunoassay on the Alinity i analyzer, a recently developed automated immunoassay analyzer [53].

Finally, the protocol for a 5-day work week follow-up is implemented to collect urine samples throughout the week and to gather self-reported data on daily job tasks by means of daily work sheets to be filled out at the workplace.

The exposure assessment is based on a multidisciplinary strategy, including urinary biomonitoring, self-reported work organization and daily job tasks, and a visit to the hair salon, including an inventory of the chemical products.

Individual Exposure Level

Urinary Measures

The urinary biomonitoring relates to the thyroid-disrupting chemicals found in hair products, namely resorcinol, parabens, and UV filters, such as ethylhexyl methoxycinnamate and benzophenone. Considering the issues of toxicokinetics, the different sources and routes of exposure, as well as the individual variability in endocrine disruptor measurements [54], urinary samples are collected at different time points. Hairdressers are asked to collect 6 urine samples during the working week: day 1 preshift, days 3 and 4 preshift and postshift, and day 5 preshift (Multimedia Appendix 1). For controls, 2 urine samples are collected during the working week

XSL•FO
RenderX

(day 1 preshift and day 5 preshift). All samples are stored at 4°C for a maximum of 3 hours after the collection time. Each sample is divided into 6 aliquots, which are finally stored at −20 °C until the urinary measurements of the endocrine disruptors are conducted.

Urinary measurements of resorcinol are assessed using a high-performance liquid chromatography-high resolution mass spectrometry method developed by our research team [28]. The lower limit of quantification for resorcinol is 0.3 µg/L, and the method has been fully validated for accuracy, precision, carryover, linearity, recovery, and stability (raw samples and postprocessing stabilities). The urinary measurements of parabens (methyl, ethyl, and propyl-paraben), ethylhexyl methoxycinnamate's metabolites (4'-methoxyacetophenone, 4-methoxycinnamic acid), and benzophenone-3 will be performed in the same laboratory.

Work Organization and Job Tasks

Information on work organization and daily job tasks is collected through questionnaires during the inclusion visit, then during a 5-day working week by means of self-reported daily work sheets. Data on work organization relate to the seniority in hairdressing activity, other concomitant activities, type of employment contract (permanent, fixed-term, temporary), weekly schedules (number of working days, working hours, and consecutive working days), and the date of the last working day. Information on daily job tasks relates to activities during the 5-day follow-up week and includes the number of customers served by the hairdresser and the salon per day, the type and number of hairdressing tasks (shampooing, haircutting, dyeing, bleaching, perming, straightening, styling, modeling), and the type and number of cleaning tasks performed at the workplace (tools, salon, furniture). For each hairdressing and cleaning task, additional questions are asked regarding the use of personal protective equipment (gloves, facial mask), the names of products used, the wearing of wrist or hand or finger jewelry, and smoking during the workday.

Salon Characteristics and Inventory

A hair salon visit is conducted by 1 to 2 members of the study team in every salon where at least 1 hairdresser is included in the study. The visit starts with an interview with the salon manager, followed by a brief observation of the salon and a complete inventory of hair and chemical products within the salon. Information on work organization (staff, schedules, customers, hairdressing tasks) and salon characteristics (opening year, total surface, number and usage of rooms, customer seats, shampoo bins, ventilation systems, personal protective equipment, cleaning tasks of tools or salon or furniture, care and hair products, cleaning products) is collected. A full inventory of products in the salon is therefore undertaken by means of dedicated numeric tools—bar code scanner devices, mobile phones, and a specific database maintained by our research team. The labels of all hair care and cleaning products are first scanned; then, photos of the name and composition are taken for each product. These data are sent to the database to feed 4 subdatabases. The first subdatabase contains all the products inventoried in the study; the second includes all the products per salon; the third contains all the compounds per product, as mentioned in the label; and the fourth contains a short toxicological information of each compound based on the Chemical Abstracts Service number where available. This includes IARC and CLP classifications for carcinogenic, mutagenic, or toxic for reproduction [55,56] and CLP classification for sensitizers [56], and 3 reference sources for endocrine disruptors. Along with endocrine disruptor lists presented above [6], 2 other sources are consulted: the DEDuCT version 2 (database of EDCs and their toxicity profiles), which compiles information on potential EDCs with supporting evidence from research papers published unto January 2020 [57,58], and a list of substances of interest because of their potential endocrine action published by French Agency for Food, Environmental & Occupational Health & Safety in 2021 [59], based on the DEDuCT database, enriched with more chemicals, such as coformulants in plant protection products (phyto or pesticides) or biocidal products and phyto or biocidal active substances.

Sample Size and Statistical Methods

The sample size was determined based on the mean serum levels of TSH (2.1 µUI/mL, SD 1.0) and free thyroxine (10.9 ng/L, SD 1.2) in healthy French women aged 35 to 60 years who were included in the Su.Vi.Max study [60]. So far, no data are available to estimate the differences in serum levels between hairdressers and the unexposed group, but we expect fairly small differences for free thyroxine (3% - 4%) and larger ones for TSH levels (13% - 15%).

For a ratio of 2 hairdressers to 1 unexposed participant, the population needed to detect these differences between hairdressers and unexposed participants consists of 300 hairdressers and 150 unexposed participants and corresponds to an 80% power at a 5% significance level. Descriptive statistics will be used to describe all population and separately analyze hairdressers and unexposed subjects. To meet the study objectives, different statistical methods will be carried out according to the conceptual framework presented in Figure 3.

For assessing the occupational exposure to resorcinol among hairdressers, urinary concentrations of EDCs at the beginning of the working day will be compared to those at the end of the day, as well as to the concentrations measured at the beginning and end of the working week, taking into account intraindividual and interindividual variability (related to the repetition of samples over 2 days). The mean concentrations will be calculated with all samples collected during the week for exposed and unexposed populations (up to 6 for exposed and 2 for nonoccupationally exposed) and compared between the 2 populations. An in-depth exposure assessment incorporating data from the hairdressing product inventory and daily work sheets will be carried out depending on the quality of the data. Nonoccupational exposures to resorcinol described in the literature will be considered by using data related to personal use of cosmetics, diet, alcohol consumption, and smoking.

For analyzing thyroid-disrupting effects in hairdressers, the relationship with thyroid parameters will be explored using 2 approaches. First, the association between the hairdressing profession and thyroid hormone concentration will be analyzed using linear regression models (path 1), taking into account

XSL•FO
RenderX

potential confounding factors collected via questionnaire and linked to participants' residence, use of personal cosmetic products, consumption of food and drink, smoking, cleaning product use, and occupational history. The different indicators of thyroid function will be treated either independently of each other or simultaneously using multivariate statistical models. Second, the relationship between thyroid parameters and urinary impregnation with resorcinol and other EDCs will be analyzed for all urine samples combined, adjusted for confounders (path 2), allowing us to assess an "average" exposure but also the concentrations measured in the end-of-shift (postshift) samples, under the hypothesis that the latter reflect a "maximum" exposure.

If possible, other individual exposure indices will be determined based on the daily work tasks crossed with the database of salon products and endocrine-disrupting substances. These indices will be used to evaluate occupational exposures and explore associations with thyroid parameters.

All analyses will be performed using Stata software (Stata Corp LP).

Figure 3. Conceptual framework following the study objectives.

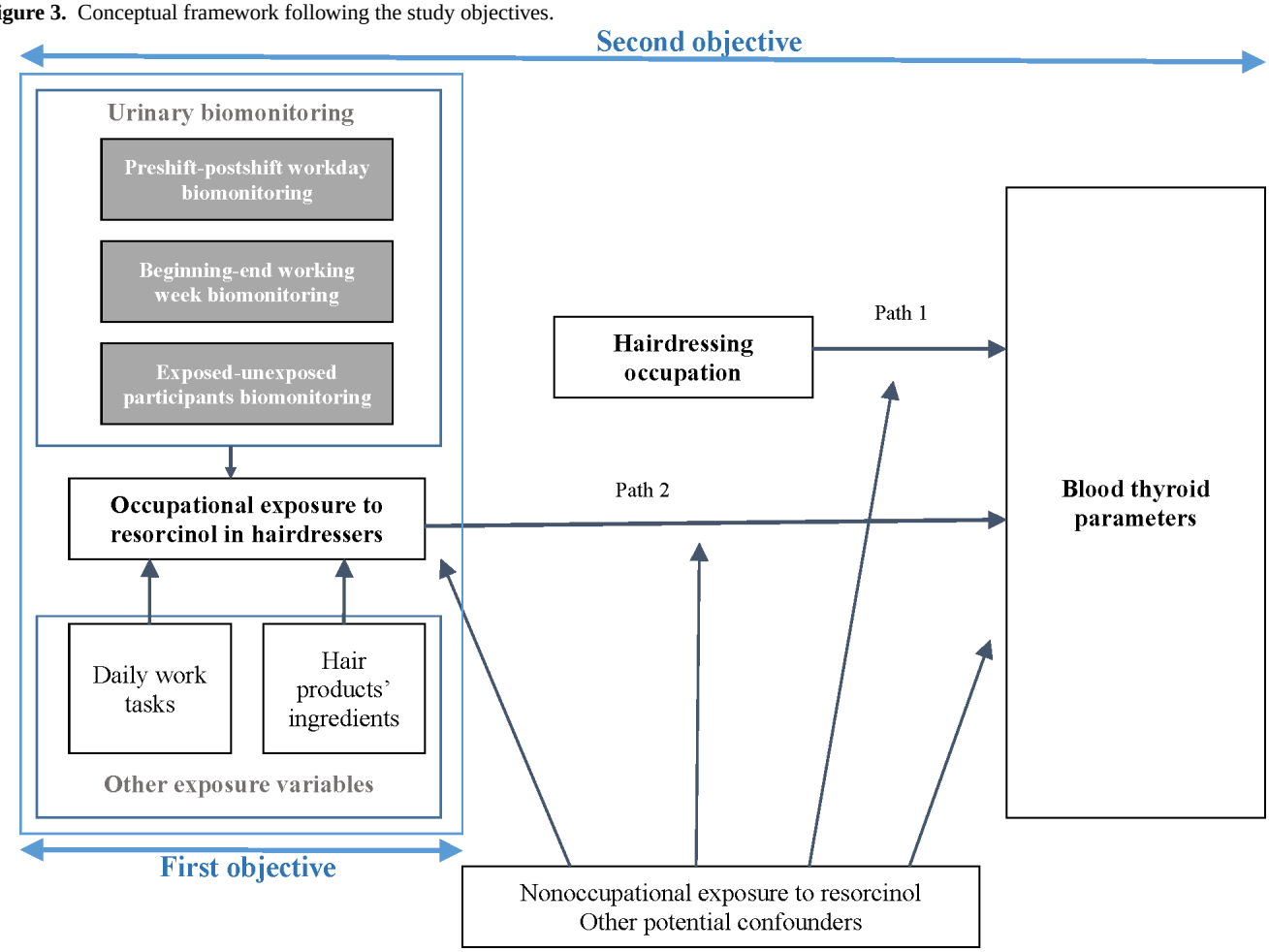

Results

Anticipated Schedule

The duration of the study spans from early 2022 to the end of 2027. The first year was dedicated to obtaining regulatory committees' approvals, developing analysis methods for resorcinol and other considered chemicals, preparing for the implementation of the partnership with occupational health centers, achieving data collection tools, and drafting the operating mode for health professionals to follow.

The data collection started in early 2023 and will run until the end of 2026. The final year (2027) will be dedicated to data management, data analysis, and the writing of the study report.

Update on the Implementation Progress and Participants' Inclusion

As of December 31, 2025, 66 hairdressers and 30 unexposed participants have been enrolled in the study across 9 occupational health centers. The hairdressers are employed in 54 hair salons where a full inventory of hair and chemical products is also conducted. The implementation of the study in the remaining 5 occupational health centers will continue throughout 2026.

Discussion

Expected Findings

This study is the first assessment of occupational exposure to resorcinol in France, a thyroid disruptor that is relevant to human health but not addressed, to the best of our knowledge, in

XSL•FO
RenderX

occupational settings. Considering the wide distribution of resorcinol in the environment and the significant increase in the global resorcinol market over the forecast decade, we propose a timely biomonitoring and epidemiological study to analyze its thyroid-disrupting effects related to occupational exposure to hair products.

Part of the "hairdressing exposome," thyroid disruptors are widely used as ingredients in hair products, and we attempt to take into account, beyond resorcinol, the most common compounds, such as parabens or UV filters.

Comparison to Prior Work

There is compelling evidence from in vivo, in vitro, and epidemiological studies that a growing number of chemicals, at environmentally relevant concentrations, can disrupt the thyroid function and cause adverse health effects.

Recent French epidemiological data showed a decrease in the incidence of acquired hypothyroidism, whereas the incidence of congenital hypothyroidism increased among female participants from 2014 to 2019. However, wide regional disparities within France were highlighted, suggesting a potential impact of environmental factors [61]. The effect of potential environmental factors underlying these disparities was examined with regard to the well-known thyroid disruptors found in outdoor air and tap water [61,62]. In some French regions, prenatal exposure to perchlorate and nitrate ions in tap water was significantly associated with increased neonatal TSH concentrations, and exposure to nitrate ions was associated with an elevated risk of congenital hypothyroidism. Also, prenatal exposure to airborne particulate matter $PM_{2.5}$ and PM_{10} during the third trimester of pregnancy was significantly associated with increased neonatal TSH concentrations. Notably, tap-water monitoring database and outdoor air quality mapping are open-access in France, allowing, where applicable, the integration of these environmental pollutants into the statistical modeling of adverse health effects.

Thyroid-disrupting health effects of exposure to care and other consumer products have been reported [20,21], and their consumption continues today, negatively affecting more susceptible populations, such as pregnant women and their offspring. The neurodevelopmental impact and long-term consequences at a neurocognitive level following in utero exposure to environmental thyroid disruptors were recently reviewed [12]. Massively distributed throughout consumer products, phthalates, bisphenol A, parabens, pesticides, persistent organic pollutants, and perchlorates represent the main maternal exposures to thyroid disruptors, which seem to be associated with long-term consequences, such as a risk for attention or deficit or hyperactivity disorders, autism spectrum disorders, and developmental delay in the offspring.

Regarding occupational exposure to thyroid-disrupting chemicals, most studies have been conducted within specific professional sectors, such as the agricultural and industrial sectors [63], making comparisons between the hairdressing sector and the existing literature limited.

Impact of the Current Regulation on the Risk Assessment of Endocrine Disruptors

Since there is no specific management of endocrine disruptors in cosmetic products, companies must identify and manage the risks linked to the substances they manufacture and market in the European Union, demonstrate how the substance can be safely used, and communicate the risk management measures to the users [1]. Further, cosmetic products are exempt from the legislation requiring the provision of a safety data sheet.

The identification of endocrine disruptors as SVHC enables the reduction of their use and ultimately their replacement with safer alternatives [64]. Nevertheless, the available regulatory guidelines and toxicological evaluations were not designed to assess endocrine-disrupting properties in themselves and need to be reconsidered to accurately assess thyroid disruption [65]. Relying almost exclusively on animal studies that may not be suitable for assessing thyroid disruption, the existing tools for regulatory evaluations of thyroid disruptors can lead to opposing opinions between regulatory instances, as was the case for resorcinol within the European Union [29]. Thus, the proposal to identify resorcinol as an SVHC was not adopted by EU instances, mostly because of some discrepancies in the animal experimental data required by the regulatory framework for EDC evaluation [65]. Moreover, the reversibility after exposure to clinical hypothyroidism in adults can raise questions, within the regulatory area, about whether such health effects on thyroid function are indeed adverse. It is noteworthy that developmental thyroid disruption can induce long-term changes in neural functions, behavior, and learning and memory in offspring. Therefore, although thyroid markers may return to normal after exposure in adults, some adverse effects may not be reversible in the context of in utero transient developmental exposure.

At the workplace level, the fact that cosmetic products do not require a safety data sheet makes it difficult to track occupational exposure to endocrine-disrupting substances contained in hairdressing products. In an attempt to support occupational health professionals in identifying endocrine disruptors in the workplace, the French Research and Safety Institute for the Prevention of Occupational Accidents and Diseases proposed in 2025 a digital tool for tracking and inventorying endocrine-disrupting chemicals [66].

Issues of Potential Generalizability and Dissemination Plan

The choice to recruit female participants aged 18 to 45 years having at least 1 year's seniority in hairdressing and working in traditional hairdressing salons relies on the use of building relatively homogeneous and stable groups concerning both exposure and thyroid parameters. In fact, the majority of hairdressers worldwide are women of childbearing age. Furthermore [1], we can expect that occupational and nonoccupational exposures to cosmetic products differ between young apprentices and experienced hairdressers [2]. The data obtained from the 2007-2008 National Health and Nutrition Examination Survey showed that the relationships between thyroid hormones and endocrine disruptors were different between women and men [67] and between adults and adolescents [68]. Moreover, the prevalence and pattern of

XSL•FO
RenderX

thyroid disorders differ between the 2 genders [69], as well as the occupational exposure between traditional hairdressing salons and nail or beauty salons [31]. Finally, although the construction of the study population took into account issues of statistical power and feasibility, the results of the study can serve the whole occupational group, including male hairdressers. Furthermore, considering ethnically similar clientele and comparable type and use of hairdressing products, the results of the study could potentially be generalized beyond France. Also, the urinary sampling framework at different time points throughout a working week (background, pre- and postshift, beginning and end of the week levels) could be generalized to assess occupational exposure to other substances with a short half-time and urinary excretion within 24 hours.

A dissemination plan related to this research has been drawn up under many axes: submission of scientific papers in international peer-reviewed journals related to urinary biomonitoring (analytical development for the quantification of considered endocrine disruptors), the study protocol, and the final results; inventory data dissemination as part of open science (the deposit of dataset in an appropriate external repository) and as scientific papers and communications; written feedback to occupational health center partners (biomonitoring report of proper data and the final report of the study); and French-written scientific papers and communications, as well as the dissemination of the final results to French occupational health and safety professionals.

Strengths and Limitations

A strength of this study is its multilevel and multidisciplinary approach, including epidemiology, occupational health, biomonitoring, and endocrinology, to set up the study methodology. Furthermore, the collection of exposure data at the hairdresser level and salon level allows for an extensive exposure assessment to thyroid-disrupting chemicals, combining biomonitoring, industrial hygiene, and inventory data exploitation. A specific partnership with occupational health centers is implemented for the recruitment process and data collection, and it is likely that their participation could improve the hairdressers' participation and the quality of data.

The study has a number of weaknesses. First, our study relies on salon managers' and hairdressers' willingness to participate in the study, so that the recruitment is very work-intensive, and it may be difficult to achieve the expected number of participants. To note, the participation rate ranges between 10% and 20% for hair salons and between 40% and 50% for hairdressers belonging to voluntary hair salons. The most common barriers reported by hair salon managers to participation in the study are the work overload, problems with the labor force, lack of interest in the research, concerns about the use of results (ie, hairdressers' health, job suitability, stock of hairdressing products), study procedures judged as too complicated, and hairdressers' refusal. Furthermore, noninclusion criteria related to recent reproductive events or thyroid disorders are quite frequently reported by hairdressers.

To improve the participation rate, different tools for better communication with each stakeholder are implemented, such as flyers and posters to inform hairdressers and controls about the study, a newsletter to share information with occupational health centers, and informal discussions with salon managers within the occupational health centers. A second drawback, related to the first, is that this population might not be considered representative of the total French hairdressers. Also, the occupational health centers are volunteers to participate, and we should consider this bias for the issue of representativity as well. Nevertheless, we can point out a wide participation of health centers scattered throughout the French regions. Finally, as this study was not built to deeply collect and analyze environmental pollutants, these factors will not be taken into account in the statistical modeling.

Conclusions and Perspectives

This research, based on a multidisciplinary approach including biomonitoring and epidemiology, will provide the first insight into occupational exposure to resorcinol in France and will advance our understanding of its thyroid-disrupting effects. Together with the inventory of hair products, these results may strengthen the tools for chemical risk diagnosis and prevention in hairdressing salons. Considering the limitations mentioned above, this study remains exploratory, and the results must be interpreted in the extents of selection bias and statistical power. There is a need for increasing the body of knowledge concerning the hairdressing sector and endocrine-related health effects. Future studies should, among others, be specifically designed to accurately examine cause-effect associations, completely measure hairdressers' exposome, consider toxicologically relevant mixtures, refine the endocrine-related pathways of adverse outcomes, and broadly adjust for coexisting potential confounders, such as environmental pollutants.

Acknowledgments

We would like to thank the staff of occupational health centers that are participating in the study: APST 18, MIST Normandie, PST 51, SISTEL, CIAMT, GIMS 13, AST 08, PROMEOM, AMETRA 06, AISMT 13, AISMT 04, SPST 19- ‐ 24, ASTPB, and ALSMT. We are very grateful to the managers of hair salons and all volunteers included in the study. Many thanks to our colleagues: Véronique Martin, Mathieu Dziurla, and Julie Voison-Christophe for their valuable contribution to data collection and data management. We thank Myriam Ricaud for her expertise and major support for partnerships. Finally, we appreciate the valuable feedback concerning the reasons for hairdressers' refusal or volunteering to participate in the study reported by APST 18 (Ms Julie Vonarx), MIST Normandie (Dr Mathilde Boulanger and Ms Caroline Paul), and PST 51 (Drs Jocelyn Castanon and Nana Dotsenko, Ms Korydwen Fruit).

The authors certify that no generative artificial intelligence was used in protocol conception and implementation, biomonitoring validation, data collection, manuscript generation, and revision process.

XSL•FO
RenderX

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this study.

Authors' Contributions

AR led the manuscript writing and contributed to the implementation of the study within occupational health centers. AC-E supervised the development and the validation of the urinary assessment of endocrine-disrupting products and substantially contributed to the conception and the design of the study. BA and MIN substantially contributed to protocol conception and implementation, data collection, data management, and validation. FD was in charge of developing assays and validation of the urinary measurement of endocrine disruptor chemicals. RP is the project coordinator and critically revised the final draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study design and data collection time points for exposed and unexposed participants.
[DOCX File, 14 KB - resprot_v15i1e65833_app1.docx]

Multimedia Appendix 2
Main questionnaires used for data collection throughout the study.
[PDF File, 8490 KB - resprot_v15i1e65833_app2.pdf]

Checklist 1
STROBE checklist.
[PDF File, 334 KB - resprot_v15i1e65833_app3.pdf]

References

1. Regulation (EC) no 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products. : Publications Office of the European Union; 2009 URL: https://www.legislation.gov.uk/eur/2009/1223/pdfs/eur_20091223_2019-06-17_en.pdf [accessed 2024-11-30]
2. Ferreira M, Matos A, Couras A, Marto J, Ribeiro H. Overview of cosmetic regulatory frameworks around the world. Cosmetics 2022 Aug;9(4):72. [doi: 10.3390/cosmetics9040072]
3. Global assessment on the state of the science of endocrine disruptors. : World Health Organization; 2002 URL: https://iris.who.int/server/api/core/bitstreams/1227bb2e-099e-4d2f-a1de-8c0239c1b650/content [accessed 2025-01-24]
4. Bouwmeester MC, Eliesen GAM, Pennings JLA, Wijnhoven SWP, Hessel EVS. Inventory of possible endocrine disrupting chemicals used in cosmetic products. Crit Rev Toxicol 2025;55(7):693-706. [doi: 10.1080/10408444.2025.2539186] [Medline: 40931777]
5. Kättström D, Beronius A, Boije Af Gennäs U, Rudén C, Ågerstrand M. Impact of the new hazard classes in the CLP regulation on EU chemicals legislation. Environ Sci Eur 2025;37(1):19. [doi: 10.1186/s12302-025-01054-4] [Medline: 39895675]
6. Endocrine Disruptor Lists. URL: https://edlists.org [accessed 2024-12-20]
7. Chantelouve M, Ripoll L. Endocrine disruptors in cosmetics: a review. : Université du Québec à Chicoutimi (UQAC); 2022 URL: https://constellation.uqac.ca/id/eprint/8464/1/Review_Endocrine_disruptors_in_cosmetics_preprint.pdf [accessed 2026-03-13]
8. He L, Michailidou F, Gahlon HL, Zeng W. Hair dye ingredients and potential health risks from exposure to hair dyeing. Chem Res Toxicol 2022 Jun 20;35(6):901-915. [doi: 10.1021/acs.chemrestox.1c00427] [Medline: 35666914]
9. Gunia-Krzyżak A, Słoczyńska K, Popiół J, Koczurkiewicz P, Marona H, Pękala E. Cinnamic acid derivatives in cosmetics: current use and future prospects. Int J Cosmet Sci 2018 Aug;40(4):356-366. [doi: 10.1111/ics.12471] [Medline: 29870052]
10. Pons R, Radauceanu A, Erb A. L'exposition professionnelle aux perturbateurs endocriniens dans les salons de coiffure [Article in French]. Hygièn Séc Trav 2023;271:36-41 [FREE Full text]
11. Boas M, Feldt-Rasmussen U, Main KM. Thyroid effects of endocrine disrupting chemicals. Mol Cell Endocrinol 2012 May 22;355(2):240-248. [doi: 10.1016/j.mce.2011.09.005] [Medline: 21939731]
12. Salazar P, Villaseca P, Cisternas P, Inestrosa NC. Neurodevelopmental impact of the offspring by thyroid hormone system-disrupting environmental chemicals during pregnancy. Environ Res 2021 Sep;200:111345. [doi: 10.1016/j.envres.2021.111345] [Medline: 34087190]

XSL•FO
RenderX

13. Vanderpump MPJ. Epidemiology of thyroid disorders. In: Luster M, Duntas LH, Wartofsky L, editors. The Thyroid and Its Diseases: A Comprehensive Guide for the Clinician: Springer International Publishing:75-85. [doi: 10.1007/978-3-319-72102-6_6]
14. Tran TVT, Kitahara CM, de Vathaire F, Boutron-Ruault MC, Journy N. Thyroid dysfunction and cancer incidence: a systematic review and meta-analysis. Endocr Relat Cancer 2020 Apr;27(4):245-259. [doi: 10.1530/ERC-19-0417] [Medline: 32045361]
15. Canaris GJ, Manowitz NR, Mayor G, Ridgway EC. The Colorado thyroid disease prevalence study. Arch Intern Med 2000 Feb 28;160(4):526-534. [doi: 10.1001/archinte.160.4.526] [Medline: 10695693]
16. Garmendia Madariaga A, Santos Palacios S, Guillén-Grima F, Galofré JC. The incidence and prevalence of thyroid dysfunction in Europe: a meta-analysis. J Clin Endocrinol Metab 2014 Mar;99(3):923-931. [doi: 10.1210/jc.2013-2409] [Medline: 24423323]
17. Valeix P, Dos Santos C, Castetbon K, Bertrais S, Cousty C, Hercberg S. Thyroid hormone levels and thyroid dysfunction of French adults participating in the SU.VI.MAX study [Article in French]. Ann Endocrinol (Paris) 2004 Dec;65(6):477-486. [doi: 10.1016/s0003-4266(04)95956-2] [Medline: 15731735]
18. Coste J, Mandereau-Bruno L, Bertagna X, Wémeau JL. Prevalence of treated hyper- and hypo-thyroidism and sociodemographic and geographic disparities in France in 2020. Eur Thyroid J 2025 Jun 20;14(3):e250041. [doi: 10.1530/ETJ-25-0041] [Medline: 40478662]
19. Chamot S, Al-Salameh A, Balcaen T, et al. Congenital and acquired hypothyroidism: temporal and spatial trends in France from 2014 to 2019. Ann Epidemiol 2024 Oct;98:18-24. [doi: 10.1016/j.annepidem.2024.07.091] [Medline: 39094977]
20. Khalid M, Abdollahi M. Environmental distribution of personal care products and their effects on human health. Iran J Pharm Res 2021;20(1):216-253. [doi: 10.22037/ijpr.2021.114891.15088] [Medline: 34400954]
21. Azeredo DBC, de Sousa Anselmo D, Soares P, Graceli JB, Magliano DC, Miranda-Alves L. Environmental endocrinology: parabens hazardous effects on hypothalamic-pituitary-thyroid axis. Int J Mol Sci 2023 Oct 17;24(20):15246. [doi: 10.3390/ijms242015246] [Medline: 37894927]
22. Xu G, Strathearn L, Liu B, Yang B, Bao W. Twenty-year trends in diagnosed attention-deficit/hyperactivity disorder among US children and adolescents, 1997-2016. JAMA Netw Open 2018 Aug 3;1(4):e181471. [doi: 10.1001/jamanetworkopen.2018.1471] [Medline: 30646132]
23. Wu NX, Deng LJ, Xiong F, et al. Risk of thyroid cancer and benign nodules associated with exposure to parabens among Chinese adults in Wuhan, China. Environ Sci Pollut Res Int 2022 Oct;29(46):70125-70134. [doi: 10.1007/s11356-022-20741-w] [Medline: 35581467]
24. Mustieles V, Balogh RK, Axelstad M, et al. Benzophenone-3: comprehensive review of the toxicological and human evidence with meta-analysis of human biomonitoring studies. Environ Int 2023 Mar;173:107739. [doi: 10.1016/j.envint.2023.107739] [Medline: 36805158]
25. Aker AM, Johns L, McElrath TF, Cantonwine DE, Mukherjee B, Meeker JD. Associations between maternal phenol and paraben urinary biomarkers and maternal hormones during pregnancy: a repeated measures study. Environ Int 2018 Apr;113:341-349. [doi: 10.1016/j.envint.2018.01.006] [Medline: 29366524]
26. Opinion on resorcinol (CAS no 108-46-3, EC no 203-585-2). : Scientific Committee on Consumer Safety (SCCS), European Commission; 2021 URL: https://health.ec.europa.eu/system/files/2022-08/sccs_o_241.pdf [accessed 2024-09-20]
27. Resorcinol market analysis 2015-2032. : ChemAnalyst; 2026 URL: https://www.chemanalyst.com/industry-report/resorcinol-market-608 [accessed 2024-10-10]
28. Cambrai-Erb A, Denis F, Pons R, Radauceanu A, Ndaw S, Grova N. Development of a novel HPLC-HRMS method for quantitative analysis of resorcinol in urine: application to hairdressers' occupational exposure. J Chromatogr B Analyt Technol Biomed Life Sci 2025 Mar 1;1253:124472. [doi: 10.1016/j.jchromb.2025.124472] [Medline: 39862513]
29. Proposal for identification of a substance of very high concern on the basis of the criteria set out in REACH article 57: resorcinol, EC no 203-585-2, CAS no 108-46-3. : European Chemicals Agency (ECHA); 2020 URL: https://echa.europa.eu/documents/10162/53d2eb0e-b0e8-fabb-b4b0-a56c246cb0a3 [accessed 2023-06-15]
30. Sciarrillo R, Falzarano A, Gallicchio V, et al. Resorcinol as "endocrine disrupting chemical": are thyroid-related adverse effects adequately documented in reptiles? In vivo experimentation in lizard Podarcis siculus. Chemosphere 2024 Sep;364:143009. [doi: 10.1016/j.chemosphere.2024.143009] [Medline: 39127192]
31. Quiros-Alcala L, Pollack AZ, Tchangalova N, DeSantiago M, Kavi LKA. Occupational exposures among hair and nail salon workers: a scoping review. Curr Environ Health Rep 2019 Dec;6(4):269-285. [doi: 10.1007/s40572-019-00247-3] [Medline: 31541357]
32. Les chiffres clés [Web page in French]. UNEC – Union Nationale des Entreprises de Coiffure. 2022. URL: https://unec.fr/chiffres-cles-de-la-coiffure [accessed 2026-03-04]
33. Henrotin JB, Picot C, Bouslama M, et al. Reproductive disorders in hairdressers and cosmetologists: a meta-analytical approach. J Occup Health 2015;57(6):485-496. [doi: 10.1539/joh.15-0068-RA] [Medline: 26269279]
34. Symanzik C, Johansen JD, Weinert P, et al. Differences between hairdressers and consumers in skin exposure to hair cosmetic products: a review. Contact Derm 2022 May;86(5):333-343. [doi: 10.1111/cod.14055] [Medline: 35088418]

XSL•FO
RenderX

35. Bremmer HJ, Lodder LCH, Engelen JGM. Cosmetics fact sheet: to assess the risks for the consumer—updated version for consexpo 4. : National Institute for Public Health and the Environment (RIVM); 2006 URL: https://www.rivm.nl/bibliotheek/rapporten/320104001.pdf [accessed 2026-03-04]
36. Ramirez-Martinez A, Granda-Torres P, Wesolek N, Ficheux AS, Roudot AC. Exposure of hairdressers to the main cosmetics used in hairdressing salons in France: a preliminary study. Arch Environ Occup Health 2016 Sep 2;71(5):247-258. [doi: 10.1080/19338244.2015.1024390] [Medline: 26010657]
37. Lind ML, Boman A, Sollenberg J, Johnsson S, Hagelthorn G, Meding B. Occupational dermal exposure to permanent hair dyes among hairdressers. Ann Occup Hyg 2005 Aug;49(6):473-480. [doi: 10.1093/annhyg/mei030] [Medline: 16081465]
38. Lind ML, Johnsson S, Meding B, Boman A. Permeability of hair dye compounds p-phenylenediamine, toluene-2,5-diaminesulfate and resorcinol through protective gloves in hairdressing. Ann Occup Hyg 2007 Jul;51(5):479-485. [doi: 10.1093/annhyg/mem028] [Medline: 17595170]
39. Kezic S, Nunez R, Babić Ž, et al. Occupational exposure of hairdressers to airborne hazardous chemicals: a scoping review. Int J Environ Res Public Health 2022 Mar 31;19(7):4176. [doi: 10.3390/ijerph19074176] [Medline: 35409860]
40. Porras SP, Hartonen M, Ylinen K, Tornaeus J, Tuomi T, Santonen T. Environmental and occupational exposure to resorcinol in Finland. Toxicol Lett 2018 Dec 1;298:125-133. [doi: 10.1016/j.toxlet.2018.03.027] [Medline: 29596886]
41. Kolena B, Petrovičová I, Šidlovská M, et al. Occupational phthalate exposure and health outcomes among hairdressing apprentices. Hum Exp Toxicol 2017 Oct;36(10):1100-1112. [doi: 10.1177/0960327116678295] [Medline: 27852937]
42. Kolena B, Petrovicova I, Sidlovska M, et al. Phthalates exposure and occupational symptoms among Slovakian hairdressing apprentices. Appl Sci (Basel) 2019;9(16):3321. [doi: 10.3390/app9163321]
43. Boyle MD, Kavi LK, Louis LM, et al. Occupational exposures to phthalates among Black and Latina U.S. hairdressers serving an ethnically diverse clientele: a pilot study. Environ Sci Technol 2021 Jun 15;55(12):8128-8138. [doi: 10.1021/acs.est.1c00427] [Medline: 34078083]
44. Newmeyer MN, Quirós-Alcalá L, Kavi LK, Louis LM, Prasse C. Implementing a suspect screening method to assess occupational chemical exposures among US-based hairdressers serving an ethnically diverse clientele: a pilot study. J Expo Sci Environ Epidemiol 2023 Jul;33(4):566-574. [doi: 10.1038/s41370-023-00519-z] [Medline: 36693958]
45. Babić Ž, Macan M, Franić Z, et al. Association of hairdressing with cancer and reproductive diseases: a systematic review. J Occup Health 2022 Jan;64(1):e12351. [doi: 10.1002/1348-9585.12351] [Medline: 36017574]
46. Buralli R, Ribeiro A, Cremonese C, Vineis P, Meyer A. Cancer mortality and premature deaths among hairdressers in Brazil. Environ Res 2024 Mar 1;244:117942. [doi: 10.1016/j.envres.2023.117942] [Medline: 38113991]
47. Hopf NB, Rousselle C, Poddalgoda D, et al. A harmonized occupational biomonitoring approach. Environ Int 2024 Sep;191:108990. [doi: 10.1016/j.envint.2024.108990] [Medline: 39244955]
48. Perrier F, Giorgis-Allemand L, Slama R, Philippat C. Within-subject pooling of biological samples to reduce exposure misclassification in biomarker-based studies. Epidemiology 2016 May;27(3):378-388. [doi: 10.1097/EDE.0000000000000460] [Medline: 27035688]
49. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
50. Ehrenkranz J, Bach PR, Snow GL, et al. Circadian and circannual rhythms in thyroid hormones: determining the TSH and free T4 reference intervals based upon time of day, age, and sex. Thyroid 2015 Aug;25(8):954-961. [doi: 10.1089/thy.2014.0589] [Medline: 26061389]
51. Winston-McPherson GN, Samraj AN, Poster K, et al. Performance characteristics of the Beckman Coulter UniCel DxI 800 TSH (3rd IS) assay. Clin Chim Acta 2018 Mar;478:90-100. [doi: 10.1016/j.cca.2017.12.029] [Medline: 29275143]
52. Giovanella L, D'Aurizio F, Algeciras-Schimnich A, et al. Thyroglobulin and thyroglobulin antibody: an updated clinical and laboratory expert consensus. Eur J Endocrinol 2023 Aug 2;189(2):R11-R27. [doi: 10.1093/ejendo/lvad109] [Medline: 37625447]
53. Ahn S, Cho Y, Lee SG, Kim JH. Performance evaluation of thyroid function tests on a novel chemiluminescent microparticle immunoassay system. J Appl Lab Med 2021 Mar 1;6(2):367-377. [doi: 10.1093/jalm/jfaa091] [Medline: 33005941]
54. Bousoumah R, Leso V, Iavicoli I, et al. Biomonitoring of occupational exposure to bisphenol A, bisphenol S and bisphenol F: a systematic review. Sci Total Environ 2021 Aug 20;783:146905. [doi: 10.1016/j.scitotenv.2021.146905] [Medline: 33865140]
55. Agents classified by the IARC monographs, volumes 1-140. : International Agency for Research on Cancer (IARC), World Health Organization (WHO); 2024 URL: https://monographs.iarc.who.int/list-of-classifications [accessed 2026-03-04]
56. European Chemicals Agency (ECHA) Chemicals Database. URL: https://chem.echa.europa.eu [accessed 2026-03-04]
57. Karthikeyan BS, Ravichandran J, Mohanraj K, Vivek-Ananth RP, Samal A. A curated knowledgebase on endocrine disrupting chemicals and their biological systems-level perturbations. Sci Total Environ 2019 Nov 20;692:281-296. [doi: 10.1016/j.scitotenv.2019.07.225] [Medline: 31349169]
58. Karthikeyan BS, Ravichandran J, Aparna SR, Samal A. DEDuCT 2.0: an updated knowledgebase and an exploration of the current regulations and guidelines from the perspective of endocrine disrupting chemicals. Chemosphere 2021 Mar;267:128898. [doi: 10.1016/j.chemosphere.2020.128898] [Medline: 33190914]

XSL•FO
RenderX

59. Elaboration d'une liste de substances chimiques d'intérêt en raison de leur activité endocrine potentielle méthode d'identification et stratégie de priorisation pour l'évaluation [Report in French]. : ANSES – Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail; 2021 URL: https://www.anses.fr/fr/system/files/REACH2019SA0179Ra-1.pdf [accessed 2023-03-30]
60. Estaquio C, Valeix P, Leenhardt L, Modigliani E, Boutron-Ruault M, Chérié-Challine L, et al. Incidence assessment of benign thyroid diseases in the SUVIMAX study, 1994-2002 [Article in French]. Bull Epidemiol Hebd 2009:445-448 [FREE Full text]
61. Chamot S, Al-Salameh A, Petit P, et al. Does prenatal exposure to multiple airborne and tap-water pollutants increase neonatal thyroid-stimulating hormone concentrations? Data from the Picardy region, France. Sci Total Environ 2023 Dec 20;905:167089. [doi: 10.1016/j.scitotenv.2023.167089] [Medline: 37717745]
62. Chamot S, Petit P, Al-Salameh A, et al. Environmental pollution and the risk of congenital hypothyroidism: insights from a French nationwide retrospective ecological cohort study. J Hazard Mater Adv 2025 Feb;17:100560. [doi: 10.1016/j.hazadv.2024.100560]
63. Singh V, Cortes-Ramirez J, Toms LM, Sooriyagoda T, Karatela S. Effects of polybrominated diphenyl ethers on hormonal and reproductive health in e-waste-exposed population: a systematic review. Int J Environ Res Public Health 2022 Jun 25;19(13):7820. [doi: 10.3390/ijerph19137820] [Medline: 35805479]
64. Substances of very high concern identification. European Chemicals Agency (ECHA). URL: https://echa.europa.eu/substances-of-very-high-concern-identification-explained [accessed 2026-03-04]
65. Pasquier E, Viguié C, Fini JB, Mhaouty-Kodja S, Michel-Caillet C. Limits of the regulatory evaluation of resorcinol as a thyroid disruptor: when limited experimental data challenge established effects in humans. Environ Res 2023 Apr 1;222:115330. [doi: 10.1016/j.envres.2023.115330] [Medline: 36693459]
66. Repérage des perturbateurs endocriniens en entreprise – outil d'aide à l'évaluation du risque chimique [Web page in French]. INRS – Institut national de recherche et de sécurité. 2025. URL: https://www.inrs.fr/media.html?refINRS=outil16 [accessed 2026-03-04]
67. Koeppe ES, Ferguson KK, Colacino JA, Meeker JD. Relationship between urinary triclosan and paraben concentrations and serum thyroid measures in NHANES 2007-2008. Sci Total Environ 2013 Feb 15;445-446:299-305. [doi: 10.1016/j.scitotenv.2012.12.052] [Medline: 23340023]
68. Meeker JD, Ferguson KK. Relationship between urinary phthalate and bisphenol A concentrations and serum thyroid measures in U.S. adults and adolescents from the National Health and Nutrition Examination Survey (NHANES) 2007-2008. Environ Health Perspect 2011 Oct;119(10):1396-1402. [doi: 10.1289/ehp.1103582] [Medline: 21749963]
69. Shinkov A, Borissova AM, Vlahov J, Dakovska L, Blajeva E. Male gender differences in the thyroid ultrasound features, thyroid peroxidase antibodies and thyroid hormone levels: a large population-based study. J Endocrinol Invest 2014 Mar;37(3):269-276. [doi: 10.1007/s40618-013-0024-y] [Medline: 24615363]

Abbreviations

CLP: classification, labeling, and packaging
EDC: endocrine-disrupting chemical
IARC: International Agency for Research on Cancer
REACH: Registration, Evaluation, Authorization, and Restriction of Chemicals
SCCS: Scientific Committee on Consumer Safety
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
SVHC: substance of very high concern
TSH: thyroid-stimulating hormone

Edited by A Schwartz; submitted 14.Mar.2025; peer-reviewed by S Chamot; accepted 30.Jan.2026; published 31.Mar.2026.

Please cite as:
Radauceanu A, Cambrai-Erb A, Adet B, Nioule MI, Denis F, Pons R
Occupational Exposure to Resorcinol and Thyroid-Disrupting Effects: Protocol for an Exploratory Field Study in French Hairdressers
JMIR Res Protoc 2026;15:e65833
URL: https://www.researchprotocols.org/2026/1/e65833
doi:10.2196/65833

© Anca Radauceanu, Amandine Cambrai-Erb, Benedicte Adet, Marie-Isabelle Nioule, Flavien Denis, Romain Pons. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits

XSL•FO
RenderX

unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Impact and Economic Evaluation of the Patient-Provider Support Agency Model Under India's National Tuberculosis Elimination Program: Protocol for a Cohort Study

Nikita Raula[1], MPH; Ingeborg Maria van der Putten[1], PhD; Nicole H T M Dukers-Muijrers[2,3], PhD; Shankar Dapkekar[4], MPH; Priyakanta Nayak[5], MPH; Silvia M A A Evers[1,6], PhD

[1]Department of Health Services Research, CAPHRI Care and Public Health Research Institute, Maastricht University, Minderbroedersberg 4-6, 6211 LH, Maastricht, The Netherlands

[2]Department of Health Promotion, Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands

[3]Department of Sexual Health, Infectious Diseases and Environmental Health, Public Health Service South Limburg, Heerlen, The Netherlands

[4]Public Health Department, Disha Foundation, Nashik, India

[5]World Health Organisation Health Emergency Programme (WHE), WHO Country Office, Ankara, Turkey

[6]Centre of Economic Evaluation, Trimbos, Netherlands Institute of Mental Health and Addiction, Utrecht, The Netherlands

Corresponding Author:
Nikita Raula, MPH
Department of Health Services Research, CAPHRI Care and Public Health Research Institute, Maastricht University, Minderbroedersberg 4-6, 6211 LH, Maastricht, The Netherlands

Abstract

Background: India is set to eliminate tuberculosis (TB) from the country by 2025, 5 years ahead of the global target. Optimizing TB care pathways is challenged by the predominance of private sector care-seeking among symptomatic patients (60%-70%), a setting associated with significant under-notification. To improve the notification of TB among private providers and improve patient management in the private sector, the National Tuberculosis Elimination Program (NTEP), India, has implemented the Patient-Provider Support Agency (PPSA). The PPSA model enables state and district TB units to engage third-party agencies to involve private providers in TB care, providing comprehensive services including diagnosis, notification, and treatment support.

Objective: The objective of this study is to assess the cost-effectiveness and feasibility of introducing the PPSA in the National TB elimination program to improve patient management by improving TB case notification, drug-resistant TB case notification, and treatment outcome and reducing loss to follow-up compared to baseline ongoing programs.

Methods: A cohort study will be conducted in 2 parallel groups in the private sector to assess the impact of the PPSA intervention on the primary and secondary outcomes of the NTEP. Eligible patients with TB from the intervention and control district, notified within a defined time frame, will be included. The intervention group will receive all the facilities and benefits under the PPSA model under the NTEP. The control group will receive care under usual conditions. The study will consist of three parts: (1) impact assessment study, (2) cost-effectiveness study, and (3) feasibility and accessibility study. The primary outcomes shall be TB case notification, drug-resistant TB case notification, universal drug susceptibility testing coverage, and treatment outcome. Secondary outcomes shall include the time between the onset of symptoms and treatment initiation, treatment adherence, adverse drug reaction management, improvement in quality of life, and societal costs. Outcome assessments will be done using questionnaires at baseline, after 3 months, at the completion of treatment, and 6 months after the completion of treatment.

Results: The study received Institutional ethical approval from Jodhpur School of Public Health Institutional Review Board, India, in accordance with the compliance of Section 4, National Ethical Guidelines for Biomedical and Health Research Involving Human Participants, ICMR (2017), with IRB Number: 10032/IRB/20-21. Data collection is underway.

Conclusions: Our work will add to other implementation studies capturing the role of the private sector in TB care in India. A key strength of this work is its use of detailed patient pathway data from the study site, suggesting that private sector engagement can support TB elimination while improving the quality and coordination of TB services across India's fragmented health care system.

International Registered Report Identifier (IRRID): DERR1-10.2196/76302

(JMIR Res Protoc 2026;15:e76302) doi:10.2196/76302

XSL•FO
RenderX

KEYWORDS

cost-effectiveness analysis; economic evaluation; National Tuberculosis Elimination Program; Patient-Provider Support Agency; private sector engagement; quality of life; tuberculosis

Introduction

Background

Tuberculosis (TB) is caused by a bacterium called *Mycobacterium tuberculosis* and is the 13th leading cause of death and the second leading infectious cause of death after COVID-19, followed by HIV or AIDS [1-4]. About one-fourth of the world's population is infected with TB. This can affect anyone anywhere and often affects people in their most productive years of life [5]. In 2021, 10.6 million people were estimated to be infected by TB, and 1.6 million died due to TB. India represents the highest TB burden globally, accounting for approximately 24.4% of the global TB burden [6]. The World Health Organization declared the sustainable development goals to eliminate TB globally by 2030. Universal access to TB care is key to reaching the global target of TB elimination [5]. However, not all patients receive adequate care for TB, mainly because of substantial underdiagnosis and underreporting of the disease [5]. In high-burden countries, a notable disparity exists between the number of individuals afflicted with TB and those who have been officially diagnosed and reported to the national TB programs. This incongruity highlights the need for greater vigilance and enhanced reporting mechanisms to ensure prompt diagnosis and treatment of TB cases. Failure to address this disparity could potentially transmit TB to unaffected populations and exacerbate the global burden of TB [5].

Aligning with the global TB elimination goal, India launched the End TB strategy under the National Tuberculosis Elimination Program (NTEP) in 2018. The ambition is to eliminate TB from the country by 2025, 5 years ahead of the global target [6,7]. In India, 2.59 million TB cases were reported to the NTEP (incidence ~26 lakh cases) in 2021 [6]. To reduce underdiagnoses and underreporting, the NTEP has decentralized the TB services to facilitate access to TB care in the country. An overview of the organization of the TB care in India can be found in Figure 1.

Figure 1. Organization of public health care response to tuberculosis (TB) in India to enable universal access to quality care. DR: drug-resistant.

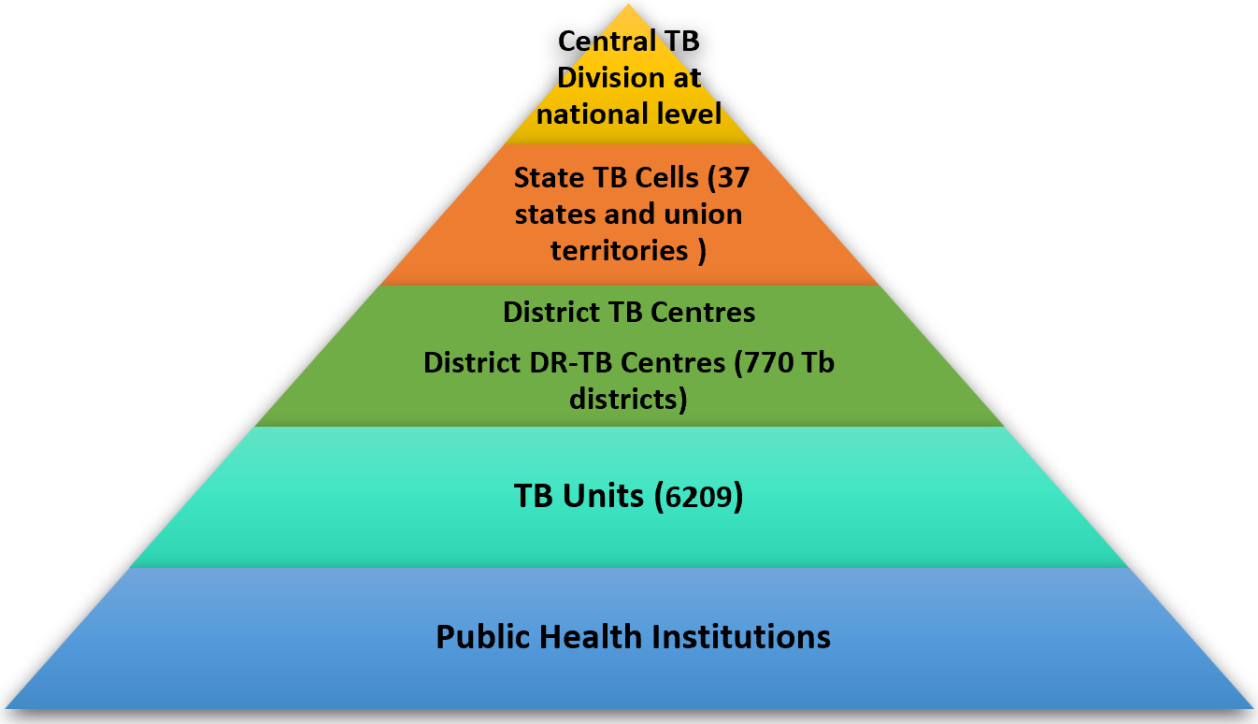

A complicating factor of optimizing the pathway to managing TB care is that about 60% - 70% of the patients with TB symptoms seek care in the private health care sector [8-10]. This vast sector is highly unregulated and leading to a high rate of under-notification of cases to the national registry [11,12]. To improve the notification of TB cases among private providers, several interventions are implemented within the NTEP, such as (1) TB case notification through call centers under NTEP, (2) mapping private providers in the program's web portal for more accessible notification, and (3) penalizing non-notification of TB. These interventions led to an increase in the notification of TB by private providers in recent years. However, private sector engagement in the program is still lagging behind, and many TB cases remain unreported from the private sector each year [5,6]. A contributing factor to the issue at hand could be insufficient awareness surrounding TB and the corresponding national TB program among private health care providers. Consequently, this has led to cases of underdiagnosis, delayed diagnoses, improper treatment of TB, as well as substandard care and support for patients with TB within the country, resulting in the emergence of drug resistance and poor treatment outcomes [13,14]. Enhancing private sector

XSL•FO
RenderX

engagement in the NTEP program would be instrumental in attaining the goal of TB elimination.

To address the loopholes in the NTEP, the national government has implemented the National Strategic Plan for 2017-2025 to extend the spectrum of high-quality TB care and control, suggesting an inclusion of private sector patients in the national program (NTEP) [7]. This emphasizes the need for public-private engagement, a vital component to eliminate TB from the country. Later on, the Central TB Division conceptualized the Patient-Provider Support Agency (PPSA) in the National Strategic Plan [7]. The PPSA acts as an interface between the public and private sector care provider.

Under the PPSA model, a third-party agency is appointed by a state, city, or district NTEP unit to collaborate with private-sector physicians in providing treatment to patients suffering from TB. This model offers comprehensive services, including diagnosis, notification, patient adherence and support, and treatment linkages, thereby facilitating end-to-end care for patients with TB [15]. An agency engaged under the public-private partnership for TB care plays a pivotal role, and their responsibilities encompass a multifaceted approach to combating TB, addressing various aspects of public health and health care service coordination.

Engagement of Private Health Care Providers

The agency's first role involves engaging with the private health care providers, which includes mapping and prioritizing different types of private health care facilities. These encompass hospitals, clinics, charitable or trust-based institutions, as well as various other health service providers. The agency is expected to foster strong and lasting relationships with these private entities through in-person visits, workshops, and other interactions. This engagement forms a crucial link in expanding the reach of the NTEP.

TB Case Notification

The agency is entrusted with the task of enhancing TB case notification. This involves registering private health care providers on NIKSHAY, a specialized TB case management system. The agency facilitates the generation of health establishment IDs for these providers and encourages them to directly notify TB cases. They also take the responsibility for training and supporting private providers in this notification process, ensuring that all TB cases are promptly reported.

Sample Collection and Transportation

Effective specimen management is a critical component of the TB elimination program. The agency is responsible for the collection of specimens from private providers and their secure transportation to NTEP laboratories, whether public or private. This includes strict adherence to bio-safety precautions and ensuring proper labeling, record-keeping, and the timely delivery of test reports back to the private health care providers.

Linkages for Free Diagnostics and Drugs

The agency is instrumental in establishing essential connections with public sector laboratories and radiology centers. This ensures that private sector patients with TB have access to free diagnostic services and radiology support. Moreover, the agency facilitates the supply of free anti-TB drugs to private patients, managing drug inventory and ensuring accurate records even in cases where patients choose to purchase drugs from the open market.

Public Health Actions

The agency's role expands to include various public health initiatives. This encompasses comorbidity testing for patients with TB, such as HIV and diabetes screening. It also includes universal drug susceptibility testing services, drug-resistant TB management, screening of close contacts, facilitating tuberculosis preventive treatment linkages, providing patient counseling, addressing stigma associated with TB diagnosis and treatment, and providing treatment adherence support. These actions ensure that patients with TB receive comprehensive care and support for their specific health care needs [16].

The active participation of the private sector holds the promise of notably boosting TB case reporting, aligning with national strategic plans. A study analyzing the accomplishments of PPSA in 6 states, backed by local funding, served as a compelling example of a robust and enduring collaboration between the public and private sectors contributing to increased notification of TB cases in these states. Another study involving private sector patients under the PPSA showed that free drugs increased engagement with treatment coordinators, leading to more opportunities to discuss treatment concerns, offer information, and connect with social support services. This enhanced social support likely contributes to increased treatment success [17]. These achievements provide valuable lessons for future initiatives aimed at integrating the private sector into national health programs.

As economies continue to grow and TB incidence declines, it will be increasingly challenging to maintain interest and funding for TB-only programs [18]. One possible solution to promote collaboration between individual health care providers and form networks and systems in the broader health system is the PPSA. Furthermore, the PPSA can create opportunities to include TB in other cross-cutting initiatives on health care consolidation, rationalization, and organization [18]. Introducing the PPSA can also highlight some difficult areas in the implementation of the national program and patient care that need to be addressed to develop better health systems service delivery [7]. One of these issues is dual practice, where health care workers work for both the government sector and hold a private practice. In many low-income settings in Asia, public sector salaries for health care workers are insufficient to sustain a family, leading to the need for dual practice [19].

In this research, we will assess the cost-effectiveness and feasibility of introducing the PPSA to improve patient management by enhancing TB case notification, drug-resistant TB case notification, treatment outcome, and reducing loss to follow-up under the NTEP of the Nanded Maharashtra district where the PPSA has been implemented by contracting a local agency, in comparison to care as usual under NTEP, provided in the Beed district of Maharashtra.

Previous attempts to quantify the economic and health impact of drug-sensitive TB with a shortened first-line regimen cost

XSL•FO
RenderX

have been limited. These efforts have focused on general analyses that are not specific to any setting. While such studies can identify key drivers of cost-effectiveness at a general level, further work is required to characterize the cost-effectiveness of drug-sensitive TB regimens at the district level. Incorporating good empirical data on health service costs is critical [20]. One of the central aims of introducing new shortened first-line regimens is to reduce the health system burden from TB treatment [7].

Furthermore, while previous studies have included estimates of provider costs, none of these studies have considered the potential cost savings for patients. The potential benefits of patient costs are crucial in light of the global TB target of ensuring that no one suffers catastrophic expenditures from TB under universal health coverage. Through this study, we aim to focus only on drug-sensitive TB with shortened first-line regimens to include in the study group, which constitutes the majority of the TB burden. Excluding drug-resistant TB avoids methodological heterogeneity, as its care pathways are distinct and programmatically more complex.

Private sector engagement is key for TB elimination in India. Pilot studies like the Mumbai Private Provider Interface Agency and state-level PPSA implementations in Jharkhand and Odisha improved case notifications, diagnostic quality, and treatment linkage [21-24]. Despite these gains, questions remain about the long-term sustainability and economic efficiency of the PPSA. Costing studies suggest that its expenses are broadly comparable to public sector interventions, highlighting the need for further evaluation. Therefore, we will conduct a study to research the primary and secondary outcomes, the economic consequences, and the feasibility and acceptability of the patient-provider support agency intervention in the Nanded district of India.

Objectives

To reach this overall goal, 3 studies will be conducted with the following research questions (Textbox 1).

Textbox 1. Study framework assessing the impact, cost-effectiveness, feasibility, and acceptability of the Patient-Provider Support Agency (PPSA) under the National Tuberculosis Elimination Program (NTEP), India.

An Impact Study of the PPSA intervention in the district of Nanded, India

- How does the PPSA intervention affect the primary outcomes in the NTEP in Nanded district, India?
- How does the PPSA intervention affect the secondary outcomes among the patients with tuberculosis (TB) in Nanded district, India?

Economic evaluation of the PPSA intervention in the district of Nanded, India

- Do the benefits of the PPSA intervention outweigh the related costs in Nanded district, India?

Feasibility and acceptability study of the PPSA intervention in the district of Nanded, India

- To assess the feasibility and acceptability of the PPSA intervention among the patients with TB in Nanded district, India.
- To assess the feasibility and acceptability of the PPSA intervention among the providers in Nanded district, India.

Methods

Design

The evaluation research will have 3 interlinked components: a cohort study, economic evaluation, and a qualitative study. The 3 studies are interlinked and together provide a broader picture of the PPSA intervention under the NTEP. The cohort study will provide the outcome data, which in turn inform the incremental cost analysis in the economic evaluation. The economic evaluation will assess the PPSA intervention from the perspectives of the health system, providers, and patients. It will include direct program costs, such as staff time, training, digital tools, monitoring, and private sector engagement, as well as indirect costs like patient time, travel, and lost income. These costs will be compared with outcomes to estimate cost-effectiveness. In line with the Medical Research Council framework, a broader cost-consequence approach will also be applied to capture a wider range of benefits, including health improvements, financial protection, and system efficiencies [25]. This will show whether the PPSA offers value for money and highlight trade-offs across sectors to guide resource allocation. The qualitative study backs the contextualization of the finding from the other 2 studies by exploring how costs and outcomes are shaped by implementation realities, stakeholder perspectives, and local health system dynamics. This triangulation will enhance the interpretation of the quantitative findings, ensuring that economic results are not viewed in isolation but in light of acceptability, feasibility, and sustainability. They form a complementary approach where the quantitative data are strengthened through the qualitative lens (Table 1).

XSL•FO
RenderX

Table . Overview of study methodology and workflow.

Category	Methodological choices
Study design	Three interlinked components: (1) cohort study on programmatic outcomes; (2) economic evaluation of cost-effectiveness; (3) qualitative study guided by Proctor's and Fleuren's frameworks to assess feasibility, acceptability, and sustainability of the PPSA[a]
Research settings	Conducted in Nanded (intervention) and Beed (control) districts, Maharashtra—comparable in demography and TB[b] burden
Primary listing	Data sources: Nikshay notification registers (December 2023), private provider lists, and PPSA staff rosters
Eligibility screening	Inclusion: ≥18 years, drug-sensitive TB, residing in study districts, consent given. Exclusion: Drug-resistant TB, transfer-outs, duplicates, nonconsent
Exclusion	Noneligible, resistant cases, transfer-outs, duplicates, or refusals
Participant enrollment	Eligible participants invited, consent obtained, enrolled in database. Providers, PPSA staff, and NTEP[c] officials included for qualitative interviews
Baseline assessment	Extraction of demographic and clinical data from Nikshay and facility records. Mapping of diagnostic and treatment pathways
Intervention	PPSA services: diagnostics, counseling, adherence support, private provider coordination, DBT[d] benefits, and so forth
Data collection	Quantitative: baseline, 6- and 12-month follow-up surveys; cost data at 6 months quantitative interviews with the patients Qualitative: semistructured interviews (patients, providers, PPSA/NTEP staff) using Proctor's and Fleuren's domains
Analysis	Quantitative: compare key outcomes between sites. Qualitative: thematic analysis using Proctor (implementation outcomes) and Fleuren (determinants). Economic: incremental cost-effectiveness. Triangulation: integration of all data streams
Correlation	Strong PPSA linkage is expected to improve timeliness, adherence, and treatment success. Higher feasibility, acceptability, and enabling determinants (per Proctor and Fleuren) should correlate with sustained adoption and better outcomes

[a]PPSA: Patient-Provider Support Agency.

[b]TB: tuberculosis.

[c]NTEP: National Tuberculosis Elimination Program

[d]DBT: direct benefit transfer.

To answer research question 1, a cohort study will be conducted to assess the impact of the PPSA intervention on primary and secondary outcomes of the NTEP. The cohort will consist of all the notified cases of drug-susceptible TB from December 1 to December 31, 2023, in the private sector in the intervention district and the control district. A prestructured questionnaire will be implemented at the baseline, 3 months after initiation of the treatment, and after completion of the treatment. A follow-up will be done after 6 months of completion of the patients' treatment.

To answer research question 2, an economic evaluation will be conducted to compare the cost-effectiveness of the PPSA in the NTEP in the intervention district and the control district with and without the intervention, respectively. Cost-effectiveness will give an overview of an increment cost to alleviate the health-related state of patients with TB getting treated at the private sector. The analysis will incorporate programmatic, intervention-specific, and patient-incurred costs, using an integrated health care and societal perspective (see Textbox 2). Quality of life (QOL) and years of life saved due to the intervention will be considered to calculate the benefit of the intervention. Each participant will be interviewed using a prestructured questionnaire.

XSL•FO
RenderX

Textbox 2. Overview of primary and secondary outcomes (description of the outcomes is attached in Multimedia Appendix 1).

Primary outcome

The primary outcome relates to 3 key elements: diagnosis, notification, and quality of care. Outcomes thus include the following:

- TB case notification
- Drug-resistant tuberculosis (TB) notification
- Universal drug susceptibility testing coverage
- Direct benefit transfer to patients
- Treatment outcome the treatment success rate among drug-sensitive TB cases

Secondary outcome

- The time between the onset of symptoms and treatment initiation
- Treatment adherence
- Adverse drug reaction management
- Contact tracing and isoniazid preventive therapy (IPT)
- Number of patients with TB screened for HIV
- Number of patients with TB screened for diabetes
- Improvement in quality of life of patients under treatment

To answer research question 3, a qualitative study using semistructured interviews will be conducted to assess the feasibility of the intervention and how that has influenced the access to TB care among the patients from the private sector and implementation challenges of the PPSA under the NTEP in Nanded district, Maharashtra, India. The schematic representation of the study framework depicting inclusion of drug-sensitive tuberculosis patients, intervention and control arms, scheduled outcome assessments at predefined time points, and administration of quality-of-life, cost, and qualitative questionnaires has been elaborated in Figure 2.

Figure 2. Schematic overview of the study. TB: tuberculosis.

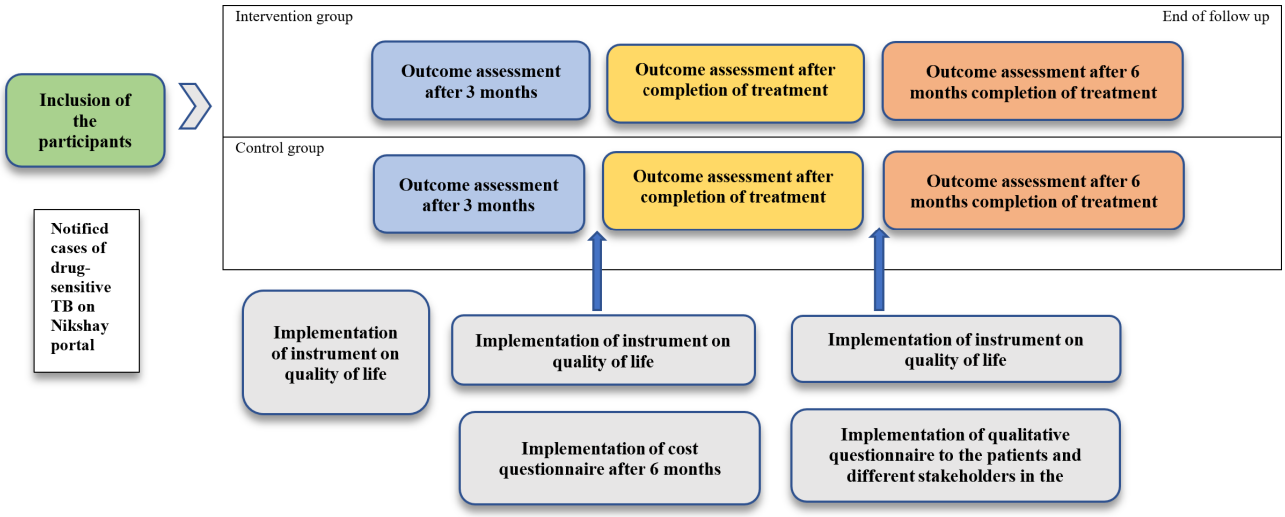

Research Setting

Maharashtra is a large state of India with a TB notification of 2,70,000 (public sector notification 95,642 and private sector notification 64,021) in the year 2021, which is 9.4% of the total TB burden of the country [6]. The study will be conducted in Nanded district, Maharashtra, India, as the intervention district, with an area of 10,528 km^2 and a population of 33,61,292 (rural: 2,447,394; urban: 913,898) [26]. A total of 2158 cases of TB from the public sector and 2308 cases from the private sector were notified during the year 2021 [27]. Therefore, it has many private practitioners in the urban setting who might be holding a significant proportion of patients with TB in the district. In addition, the PPSA was implemented in the district in February 2021 by the NTEP. The study will also include the Beed district of Maharashtra as a control district with comparable demography and disease burden to the intervention district. The Beed district expands to an area of 10,693 km^2 with a population of 2,585,049 (rural: 2,070,751; urban: 514,298) [28]. A total of 1034 cases of TB from the public sector and 862 cases from the private sector were notified during the year 2021, which is comparable to the intervention district [27].

XSL•FO
RenderX

Study Population

All adult patients with pulmonary or extrapulmonary TB, diagnosed with microscopic slide testing or sputum culture or cartridge-based nucleic acid amplification test or clinically, who will be notified between December 1 and December 31, 2023, in the private sector in the intervention and control district, and are willing to participate in the study, will be included in the study. We will exclude the patients with confirmed drug-resistant TB, keeping in view the timeline of the study duration.

This timeframe was chosen after reviewing TB notification trends from previous months and years, which showed that notifications in December were consistent with routine patterns. Besides, the completeness of reporting and data availability with the PPSA implementation rollout was also considered to ensure the reliability of the selected cohort. While the use of a single-month sample does not capture potential seasonal variation, the aim of this study is to assess the feasibility and implementation outcomes of the PPSA model within a defined cohort. This limitation has been acknowledged, and future studies should extend to multimonth cohorts to strengthen representativeness and generalizability.

Impact of the PPSA

In this intervention, the private providers continue managing their patients starting from diagnosis to treatment completion, along with some incentive and technical support from the program instead of transferring their patients to the public sector. Within the PPSA program, the patient receives subsidies for cartridge-based nucleic acid amplification test (Gene Xpert) test. Once diagnosed with TB, they avail the free drugs from the public sector. To help patients with treatment adherence, patients are also linked to the call center.

Inclusion Procedures

All the patients with drug-sensitive TB notified between December 1 and December 31, 2023, in the private sector in the intervention and control districts who will be willing to participate in the study will be included in this study. After acquiring written consent, a structured questionnaire will be administered at baseline, 3 months, the completion of treatment, and 6 months after the completion of treatment.

Measurements

The questionnaire will include demographic data of the patient and various primary and secondary outcomes listed below, along with the QOL of the patients. For the quantitative assessment, the World Health Organization *Tuberculosis Patient Cost Surveys: A Handbook* (2017) will be adapted [29]. To assess the QOL, we will be implementing a combined questionnaire of EQ-5D (EuroQol 5-Dimension Questionnaire) and SF-36 (Short Form Health Survey [36 items]) instrument, which describes various aspects of health, that is, mobility, self-care, usual activities, pain or discomfort, and anxiety or depression. The tool will be pretested among the study population to ensure appropriateness and understanding.

Economic Evaluation

With the introduction of the new "Guidance Document on Partnerships" in late 2019, the state of Maharashtra has decided to follow output-based contracting from January 2021. According to this, PPSA implementing agencies will be paid for the program output such as notifications, the number of tests, and quality of care, instead of inputs such as incremental costs. The states must fix the cost for each service delivery according to market-based competition and technical competence of the service providers [15]. This economic evaluation will consider both health care system and societal perspectives by accounting for various costs and relating them to the incremental health-related benefits among patients with tuberculosis receiving treatment in the private sector. The intervention costs will be calculated based on the existing contract between the National Health Mission of the state and the implementing agency. A structured questionnaire will be implemented to both the intervention group and the control group to collect information on various costs incurred at the patient side, starting from the appearance of disease symptoms to completion of the treatment. Finally, the cost-effectiveness ratio will be calculated and compared between the 2 groups.

Qualitative Study

This qualitative study will be conducted in the intervention district (Nanded district, Maharashtra) using semistructured interviews including patients with TB, private health care providers, and program-implementing personnel. The study will be based on Proctor's implementation outcomes framework, which provides a structure to assess implementation success through outcomes such as feasibility, acceptability, appropriateness, fidelity, penetration, and sustainability and Fleuren's determinants of innovation for the designing and analysis to understand multilevel factors influencing implementation, such as organizational context, intervention characteristics, user-related factors, and sociopolitical influences.

The study will adopt a combination of purposive, convenience, and snowball sampling. A purposive sampling approach will be used to include key PPSA staff and relevant NTEP officials to capture context-specific insights. For the private sector, convenience sampling will be applied. Snowball sampling will be used to identify patients through practitioner referrals. Interview guides will be developed in English, translated into local languages, and will be piloted to ensure cultural appropriateness.

Analysis

Impact Analysis

Secondary data will be collected from the web portal of the NTEP and treatment registers. Baseline data will be analyzed and described. Statistical procedures will be conducted using SPSS Statistics 22.0 (IBM Corporation). Missing data will be handled using SPSS missing value analysis on item level. To calculate the impact of the PPSA, outcomes and their changes from baseline will be assessed during and after treatment completion in the intervention district and compared with those observed in the control district. The primary outcomes focus on critical aspects of TB care that reflect how well the

XSL•FO
RenderX

intervention is working in real programmatic settings. These include the notification of TB and drug-resistant TB cases, and the coverage of universal drug susceptibility testing, which together indicate adherence to national standards for diagnosis and reporting. The study will also examine direct benefit transfers provided to patients through the Nikshay Poshan Yojana, as well as treatment outcomes such as cure, treatment completion, treatment failure, loss to follow-up, regimen changes, and mortality, including overall treatment success for drug-sensitive cases. Secondary outcomes will capture the essence of patient-centric care, including the time between symptom onset and treatment initiation, patient adherence as recorded in the Nikshay platform, management of adverse drug reactions, contact tracing and provision of isoniazid preventive therapy for high-risk contacts, and screening of patients with TB for HIV and diabetes. Patient-centered outcomes, such as improvements in physical, psychological, and social well-being, will also be assessed to understand the broader impact of PPSA-supported interventions on QOL. Taken together, these outcomes provide a comprehensive view of the effectiveness, coverage, and quality of TB care, allowing for a meaningful comparison between intervention and control districts under real-world program conditions. Primary and secondary outcome measures for evaluation of the PPSA intervention under the NTEP are elaborated in Textbox 2.

Baseline differences will be corrected by the inclusion of covariates in the analyses. A 2-sided significance level of .05 will be used as a threshold to determine whether differences are statistically significant. A probabilistic sensitivity analysis is planned to be conducted to examine how parameter uncertainty may impact the results of the analysis. Additionally, the combined effect of uncertainty in all parameters will be assessed [30,31]. A multivariate regression shall be used to predict missing values depending on the existing data. The precise approach will depend on the nature of the missing data: (1) missing completely at random, with no relation to the value of any other factors in the study population; (2) missing at random, but correlated in an observable way with the mechanism that generates the outcome; and (3) not missing at random, but dependent on unobserved variables [32].

Economic Evaluation

The economic evaluation will be performed, which will consist of a cost-effectiveness analysis [33]. Statistical procedures will be conducted using SPSS Statistics 22.0 (IBM Corporation). We intend to perform a comprehensive cost analysis considering both the health system and societal perspectives. Cost calculation will be done based on different costs incurred in the implementation of the intervention and costs from the patient perspective [34]. Intervention cost will be calculated based on the contract between the state National Health Mission and the delivering agency and partnership guidelines in Revised National Tuberculosis Control Program, India [35]. Health care and patient costs will be estimated using a questionnaire regarding health care utilization and productivity losses during the course of disease. The cost estimated in terms of rupees will be converted to US dollars [36]. The costs will be discounted over the years to make all included costs comparable. The outcome used in the cost-effectiveness analysis is the different programmatic indicators such as total diagnosed cases and treatment completion. Generic QOL is considered the main outcome, which will be calculated by comparing both EQ-5D-5L (EuroQol 5-Dimension 5-Level Questionnaire) and SF-36. Resource use and outcomes are measured at the same time points mentioned in the effectiveness study at baseline and 6 months or at the completion of the treatment in the intervention group as well as the control group. The tool to measure the QOL will be implemented at baseline, after 6 months, and after 12 months from the date of enrollment. A comparison will be made in terms of incremental costs and incremental effects. Costs and benefits included in the economic evaluation are summarized in Textbox 3.

The economic evaluation will follow the same analytical procedures as outlined for the impact analysis, including adjustment for baseline differences through covariates, application of a 2-sided significance level of .05, probabilistic sensitivity analysis to assess parameter uncertainty, and multivariate regression methods for handling missing data.

XSL•FO
RenderX

Textbox 3. Costs and benefits included in economic evaluation.

- **A. Cost of intervention**

1. Costs of engaging private providers in the program include
   - Cost for provider training
   - Payment to the public-private mix co-ordinator
   - Subsidies to the private providers
   - Information, education, and communication material
2. Cost toward the patients
   - Subsidies to the patients for tuberculosis (TB) diagnostics or diagnostic cost (microscopy, cartridge-based nucleic acid amplification test [CBNAAT', True NAAT line probe assay, culture, chest X-ray)
   - Cost of nutrition support to the patient
   - Cost of treatment counseling for adherence
   - Cost for the call center support mechanism
   - Cost of TB drugs
   - Sample transportation cost
   - Cost of patient house visits
   - Incentive toward transportation of the patient (max=750 Indian rupees [approximately US $8])
   - Cost of hospital stay per day

- **B. Household or cost from the patient side**

Household costs reported by the patient and the relative or guardian of the patient will be categorized into the following:

- The cost paid to the private provider
- Cost of drugs
- Cost of diagnosis
- Cost of hospital stay
- Cost of transportation to the private provider
- Cost of transportation to the hospital
- Cost of transportation of the accompanying person
- Loss of wage due to hospital visit or hospital stay
- Loss of remuneration of the relative or accompanying person
- Expenditure on extra food during the illness

- **C. Benefits of the intervention**
   - Years of life gained
   - Quality of life

Qualitative Analysis

The analysis will follow these steps: (1) summarizing all the transcripts; (2) identifying units of meaning that represent different factors that are relevant to the feasibility and acceptability of the intervention among the patients and various stakeholders, and coding for these, thematically using ATLAS.ti, applying both deductive and inductive coding; (3) condensing the content of each of the coded groups; (4) integrating the insights from the condensed meaning units into generalized descriptions that reflect apparently significant factors. The Atlas software shall be used to analyze qualitative data [37]. Trustworthiness will be maintained through triangulation of data sources, member checking, reflexive journaling, and adherence to COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting standards [38,39].

Ethical Considerations

The study received Institutional ethical approval from Jodhpur School of Public Health Institutional Review Board, India, in

XSL•FO
RenderX

accordance with the compliance of Section 4, National Ethical Guidelines for Biomedical and Health Research Involving Human Participants, ICMR (2017; IRB number: 10032/IRB/20 - 21). Ethical written informed consent will be obtained from all the study participants. The data collection will be anonymized to ensure confidentiality. No participant compensation is planned to be provided.

Results

The study is self-funded. As of March 2024, participant enrollment has been completed across the selected districts, and data collection is currently underway and includes patients, private healthcare providers, PPSA field staff, and program officials, using structured interviews, qualitative questionnaires, and secondary data analysis. The study aims to ensure comprehensive stakeholder representation across provider categories and implementation settings, with adequate coverage of PPSA functions, service delivery mechanisms, and associated cost components. No participant compensation is planned nor provided. Participant privacy will be protected through strict data anonymization and de-identification procedures. No personally identifiable information such as names, addresses, or contact details will be collected in the study dataset. Any reporting or publication of study findings will present aggregated results to ensure that individual participants cannot be identified.

Discussion

Engaging with India's vast, fragmented private health care sector is crucial in enhancing TB control in India. Our work will add to other implementation studies capturing the role of the private sector in TB care in India. A strength of this work is that it is informed by unique, detailed patient pathway data from the study site, suggesting that engaging the private sector will meet the country's aspirations for TB elimination. Such measures would also lay the foundations for TB control by maximizing the quality and coordination of essential TB services across India's vast and fragmented health care system.

The World Health Organization defines health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity." TB is considered one of the illnesses that can seriously undermine QOL [40]. TB adversely affects all dimensions of QOL, such as general health perception, corporal sense, psychological health, mental peace, and functionality of physical and social roles [41-44]. The factors in patients affected with TB, including long-term treatment; multidrug therapy, toxic reactions and side effects of medications, adherence to drug regimen, social impacts; social support, social acceptance of the illness, family, changes in lifestyle, patients' marital status, extent of access to health care services, socioeconomic status; patients' and their family's knowledge of the illness and treatments, as well as complications of TB, highly influence the QOL [22,45,46]. Still, many national TB programs across the globe universally evaluate the program success in terms of the number of people treated and cured, but not the impact of the program on well-being. Populations' well-being needs to be focused on to improve overall productivity and economy and advocate its use as an indicator of societal progress [47]. Tools measuring QOL are most appropriate for assessing subjective well-being because they assess satisfaction and functioning that respond to life events and are relatively stable compared to substantially different, fleeting feelings of euphoria or happiness. This study will be conducted in a low- and middle-income country setting; we will use QOL for population health assessments, priority setting, and program evaluation.

There are several generic instruments that can be used across a spectrum of conditions, and Short Form 36 is commonly used in TB research, which provides the in-depth assessment of QOL among the patients with TB along 8 different dimensions [48-50]. However, while evaluating a health intervention, utility scores cannot be directly calculated from the SF-36 without taking into account the EQ-5D. The reason is that the SF-36 is not based on individual preferences, whereas the EQ-5D uses a preference-based scoring algorithm [51]. It is the most commonly used QOL assessment tool in economical evaluation and is also the most validated instrument across many countries [52].

The qualitative study adopts a multistakeholder perspective, which enhances contextual understanding. Purposive, convenience, and snowball sampling will be employed to capture diverse stakeholder perspectives, particularly from private providers and patients who are often difficult to reach through formal channels. While this approach may introduce selection bias, several strategies will be incorporated to enhance methodological rigor [53,54]. The data will be triangulated across stakeholder groups, efforts will be made to include both highly engaged and less involved providers, and reflexive documentation will be maintained throughout data collection and analysis. These measures are expected to strengthen the credibility and transferability of the findings within the operational context of the PPSA (Checklist 1) [53].

Funding

This study is self-funded.

Authors' Contributions

Conceptualization: IMP, NHTMDM, NR, SMAAE
Data curation: IMP, NHTMDM, NR, SD, SMAAE
Formal analysis: IMP (lead), NHTMDM (supporting), NR (lead), SD (supporting), SMAAE (supporting)
Funding acquisition: NR

XSL•FO
RenderX

Investigation: IMP, NHTMDM, NR, PN, SD, SMAAE (equal)
Methodology: IMP, NHTMDM, NR, SMAAE
Project administration: IMP, NR, PN, SD (equal)
Resources: IMP, NR, PN, SD, SMAAE
Software: IMP, NR, SD
Supervision: IMP (supporting), NHTMDM (supporting), PN (supporting), SD (supporting), SMAAE (lead)
Validation: IMP, NHTMDM, NR, PN, SD, SMA (equal)
Visualization: NR, NS (equal)
Writing – original draft: IMP (supporting), NHTMDM, NR (lead), SMAAE
Writing – review and editing: IMP (lead), NHTMDM, PN, SD, SMAAE

Conflicts of Interest

None declared.

Multimedia Appendix 1
Primary and secondary outcomes [55-64].
[DOCX File, 18 KB - resprot_v15i1e76302_app1.docx]

Checklist 1
SPIRIT checklist.
[DOCX File, 34 KB - resprot_v15i1e76302_app2.docx]

References

1. Ravimohan S, Kornfeld H, Weissman D, Bisson GP. Tuberculosis and lung damage: from epidemiology to pathophysiology. Eur Respir Rev 2018 Mar 31;27(147):170077. [doi: 10.1183/16000617.0077-2017] [Medline: 29491034]
2. Churchyard G, Kim P, Shah NS, et al. What we know about tuberculosis transmission: an overview. J Infect Dis 2017 Nov 3;216(suppl_6):S629-S635. [doi: 10.1093/infdis/jix362] [Medline: 29112747]
3. Schluger NW. The pathogenesis of tuberculosis: the first one hundred (and twenty-three) years. Am J Respir Cell Mol Biol 2005 Apr;32(4):251-256. [doi: 10.1165/rcmb.F293] [Medline: 15778414]
4. Tuberculosis. World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/tuberculosis [accessed 2025-12-25]
5. Global Tuberculosis Report 2022. : World Health Organization; 2022 URL: https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2022 [accessed 2025-12-25]
6. Tuberculosis India: annual report of the Revised National Tuberculosis Control Programme. : Central Tuberculosis Division, Ministry of Health and Family Welfare, Government of India; 2022 URL: https://tbcindia.mohfw.gov.in/wp-content/uploads/2023/05/TBAnnaulReport2022.pdf [accessed 2025-12-25]
7. Revised National Tuberculosis Control Programme: National Strategic Plan for Tuberculosis Elimination 2017-2025. : Central TB Division, Directorate General of Health Services, Ministry of Health and Family Welfare, Government of India; 2017 URL: https://tbcindia.mohfw.gov.in/wp-content/uploads/2023/05/National-Strategic-Plan-2017-25.pdf [accessed 2025-12-25]
8. Kwan A, Daniels B, Saria V, et al. Variations in the quality of tuberculosis care in urban India: a cross-sectional, standardized patient study in two cities. PLoS Med 2018 Sep;15(9):e1002653. [doi: 10.1371/journal.pmed.1002653] [Medline: 30252849]
9. Veesa KS, John KR, Moonan PK, et al. Diagnostic pathways and direct medical costs incurred by new adult pulmonary tuberculosis patients prior to anti-tuberculosis treatment—Tamil Nadu, India. PLoS ONE 2018;13(2):e0191591. [doi: 10.1371/journal.pone.0191591] [Medline: 29414980]
10. Sreeramareddy CT, Qin ZZ, Satyanarayana S, Subbaraman R, Pai M. Delays in diagnosis and treatment of pulmonary tuberculosis in India: a systematic review. Int J Tuberc Lung Dis 2014 Mar;18(3):255-266. [doi: 10.5588/ijtld.13.0585] [Medline: 24670558]
11. Udwadia ZF, Amale RA, Ajbani KK, Rodrigues C. Totally drug-resistant tuberculosis in India. Clin Infect Dis 2012 Feb 15;54(4):579-581. [doi: 10.1093/cid/cir889] [Medline: 22190562]
12. Satyanarayana S, Nair SA, Chadha SS, et al. From where are tuberculosis patients accessing treatment in India? Results from a cross-sectional community based survey of 30 districts. PLoS One 2011;6(9):e24160. [doi: 10.1371/journal.pone.0024160] [Medline: 21912669]
13. Pai M, Dewan P. Testing and treating the missing millions with tuberculosis. PLoS Med 2015 Mar;12(3):e1001805. [doi: 10.1371/journal.pmed.1001805] [Medline: 25803483]
14. Pai M, Daftary A, Satyanarayana S. TB control: challenges and opportunities for India. Trans R Soc Trop Med Hyg 2016 Mar;110(3):158-160. [doi: 10.1093/trstmh/trw003] [Medline: 26884494]

XSL•FO
RenderX

15. Notice inviting tender for selection of service providers for working as a patient provider support agency (PPSA) for providing tuberculosis (TB) related services & drug logistic & support chain: Tender reference No. E-49/DHS/Proc.Cell/PPSA/DL&SC/2019-20. : Directorate of Health Services, Government of Maharashtra; 2018.
16. Standard document for selection of PPSA. : Ministry of Health & Family Welfare-Government of India; 2019 URL: https://tbcindia.mohfw.gov.in/wp-content/uploads/2023/05/4862716111Standard-Document-for-Selection-of-PPSA-revised.pdf [accessed 2025-12-25]
17. Sodhi R, Penkunas MJ, Pal A. Free drug provision for tuberculosis increases patient follow-ups and successful treatment outcomes in the Indian private sector: a quasi experimental study using propensity score matching. BMC Infect Dis 2023 Jun 21;23(1):421. [doi: 10.1186/s12879-023-08396-5] [Medline: 37344775]
18. Wells WA, Uplekar M, Pai M. Achieving systemic and scalable private sector engagement in tuberculosis care and prevention in Asia. PLoS Med 2015 Jun;12(6):e1001842. [doi: 10.1371/journal.pmed.1001842] [Medline: 26103555]
19. Moghri J, Rashidian A, Arab M, Akbari Sari A. Implications of dual practice among health workers: a systematic review. Iran J Public Health 2017 Feb;46(2):153-164. [Medline: 28451549]
20. Gomez GB, Dowdy DW, Bastos ML, et al. Cost and cost-effectiveness of tuberculosis treatment shortening: a model-based analysis. BMC Infect Dis 2016 Dec 1;16(1):726. [doi: 10.1186/s12879-016-2064-3] [Medline: 27905897]
21. Samal J, Preetha GS, Mondal A. Assessing the operationalization of patient-provider support agencies supporting the private-sector tuberculosis patients in an eastern Indian state. J Datta Meghe Inst Med Sci Univ 2025;20(2):345-352. [doi: 10.4103/jdmimsu.jdmimsu_777_24]
22. Narain S, Sagar V, Thankamma A, Mitra A. Impact of patient provider support agency on private sector engagement in tuberculosis care in eastern India. Int J Community Med Public Health 2023;10(9):3352-3357. [doi: 10.18203/2394-6040.ijcmph20232703]
23. Satyanarayana S, Subbaraman R, Shete P, et al. Quality of tuberculosis care in India: a systematic review. Int J Tuberc Lung Dis 2015 Jul;19(7):751-763. [doi: 10.5588/ijtld.15.0186] [Medline: 26056098]
24. Deo S, Jindal P, Gupta D, et al. What would it cost to scale-up private sector engagement efforts for tuberculosis care? Evidence from three pilot programs in India. PLoS One 2019;14(6):e0214928. [doi: 10.1371/journal.pone.0214928] [Medline: 31166942]
25. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
26. Nanded district—population 2011-2025. Population Census. 2011. URL: https://www.census2011.co.in/census/district/349-nanded.html [accessed 2025-12-25]
27. National Tuberculosis Elimination Programme. Central Tuberculosis Division. 2021. URL: https://tbcindia.mohfw.gov.in/ [accessed 2025-12-25]
28. Bid (Beed) district—population 2011-2025. Population Census. 2011. URL: https://www.census2011.co.in/census/district/361-bid.html [accessed 2025-12-25]
29. Tuberculosis patient cost surveys: a handbook.: World Health Organization; 2017 URL: https://iris.who.int/server/api/core/bitstreams/e2ed2136-7bdf-48c5-bfb8-68f074e43b45/content [accessed 2026-01-28]
30. Briggs AH, Weinstein MC, Fenwick EAL, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making 2012;32(5):722-732. [doi: 10.1177/0272989X12458348] [Medline: 22990087]
31. Claxton K. Exploring uncertainty in cost-effectiveness analysis. Pharmacoeconomics 2008;26(9):781-798. [doi: 10.2165/00019053-200826090-00008] [Medline: 18767898]
32. Beyad Y, Maeder M. Multivariate linear regression with missing values. Anal Chim Acta 2013 Sep 24;796:38-41. [doi: 10.1016/j.aca.2013.08.027] [Medline: 24016580]
33. Husereau D, Drummond M, Augustovski F, et al. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) statement: updated reporting guidance for health economic evaluations. MDM Policy & Practice 2022;7(1). [doi: 10.1177/23814683211061097]
34. Drummond MF, Torrance GW, Stoddart GL, Claxton K, Sculpher MJ. Methods for the Economic Evaluation of Health Care Programmes: Oxford University Press; 2015. URL: https://books.google.co.uk/books?id=lvWACgAAQBAJ [accessed 2026-01-01]
35. Guidance Document on Partnerships 2019. : Government of India, Ministry of Health and Family Welfare, Central TB Division; 2019 URL: https://tbcindia.mohfw.gov.in/wp-content/uploads/2023/05/9531588006Guidance-Doucument-on-Partnerships-RNTCP-2019.pdf [accessed 2025-12-25]
36. Foreign exchange rates: U.S. dollar to Indian rupee. Board of Governors of the Federal Reserve System. URL: https://www.federalreserve.gov/releases/h10/current/ [accessed 2025-12-25]
37. A complete guide on qualitative data analysis. ATLAS.ti. 2023. URL: https://atlasti.com/guides/qualitative-research-guide-part-2/qualitative-data-analysis [accessed 2025-12-25]
38. Korstjens I, Moser A. Series: practical guidance to qualitative research. Part 4: trustworthiness and publishing. Eur J Gen Pract 2018 Dec;24(1):120-124. [doi: 10.1080/13814788.2017.1375092] [Medline: 29202616]

XSL•FO
RenderX

39. Loh J. Inquiry into Issues of trustworthiness and quality in narrative studies: a perspective. Qual Rep ;18(33):1-15. [doi: 10.46743/2160-3715/2013.1477]
40. Nakane Y, Tazaki M, Miyaoka E. Distribution of QOL scores in the general population: using WHOQOL. Medicine and Society 1999;9(1):123-131. [doi: 10.4091/iken1991.9.1_123]
41. Aggarwal AN. Quality of life with tuberculosis. J Clin Tuberc Other Mycobact Dis 2019;17:100121. [doi: 10.1016/j.jctube.2019.100121] [Medline: 31788563]
42. Mohammed S, Nagla S, Morten S, Asma E, Arja A. Illness perceptions and quality of life among tuberculosis patients in Gezira, Sudan. Afr Health Sci 2015 Jun;15(2):385-393. [doi: 10.4314/ahs.v15i2.11] [Medline: 26124783]
43. Brown J, Capocci S, Smith C, Morris S, Abubakar I, Lipman M. Health status and quality of life in tuberculosis. Int J Infect Dis 2015 Mar;32:68-75. [doi: 10.1016/j.ijid.2014.12.045] [Medline: 25809759]
44. Unalan D, Soyuer F, Ceyhan O, Basturk M, Ozturk A. Is the quality of life different in patients with active and inactive tuberculosis? Indian J Tuberc 2008 Jul;55(3):127-137. [Medline: 18807744]
45. Marra CA, Marra F, Colley L, Moadebi S, Elwood RK, Fitzgerald JM. Health-related quality of life trajectories among adults with tuberculosis: differences between latent and active infection. Chest 2008 Feb;133(2):396-403. [doi: 10.1378/chest.07-1494] [Medline: 18198260]
46. Masumoto S, Yamamoto T, Ohkado A, Yoshimatsu S, Querri AG, Kamiya Y. Factors associated with health-related quality of life among pulmonary tuberculosis patients in Manila, the Philippines. Qual Life Res 2014 Jun;23(5):1523-1533. [doi: 10.1007/s11136-013-0571-x] [Medline: 24264802]
47. The end TB strategy. : World Health Organization; 2015 URL: https://iris.who.int/server/api/core/bitstreams/d128ca35-55c7-4ec8-b2f3-987ca112a4a2/content [accessed 2025-12-25]
48. Patel PG, Ramanuj V, Bala DV. Assessment quality of life (QoL) of TB patients registered in tuberculosis units of Ahmedabad municipal corporation area by using WHO short form-36 (SF-36) questionnaire. Sch J Appl Med Sci 2014;2:3303-3306. [doi: 10.36347/sjams.2014.v02i06.089]
49. Dujaili JA, Sulaiman SAS, Hassali MA, Awaisu A, Blebil AQ, Bredle JM. Health-related quality of life as a predictor of tuberculosis treatment outcomes in Iraq. Int J Infect Dis 2015 Feb;31:4-8. [doi: 10.1016/j.ijid.2014.12.004] [Medline: 25486011]
50. Jaber AAS, Khan AH, Syed Sulaiman SA, Ahmad N, Anaam MS. Evaluation of Health-related quality of life among tuberculosis patients in two cities in Yemen. PLoS One 2016;11(6):e0156258. [doi: 10.1371/journal.pone.0156258] [Medline: 27257990]
51. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002 Mar;21(2):271-292. [doi: 10.1016/s0167-6296(01)00130-8] [Medline: 11939242]
52. Kennedy-Martin M, Slaap B, Herdman M, et al. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. Eur J Health Econ 2020 Nov;21(8):1245-1257. [doi: 10.1007/s10198-020-01195-8] [Medline: 32514643]
53. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544. [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
54. Sadler GR, Lee HC, Lim RSH, Fullerton J. Recruitment of hard-to-reach population subgroups via adaptations of the snowball sampling strategy. Nurs Health Sci 2010 Sep 1;12(3):369-374. [doi: 10.1111/j.1442-2018.2010.00541.x] [Medline: 20727089]
55. Mandatory notification of TB diagnosis. Knowledge Base for the National TB Elimination Program - NTEP. URL: https://ntep.in/node/361/CP-mandatory-notification-tb-diagnosis [accessed 2025-12-25]
56. Universal drug-susceptibility testing [UDST] for TB. Knowledge Base for the National Tuberculosis Elimination Programme. URL: https://ntep.in/node/1385/CP-universal-drug-susceptibility-testing-udst-tb [accessed 2025-12-25]
57. Nutritional support to TB patients (Nikshay Poshan Yojana). : Government of India, Ministry of Health & Family Welfare (Central TB Division); 2018 URL: https://tbcindia.mohfw.gov.in/wp-content/uploads/2023/05/6851513623Nutrition-support-DBT-Scheme-details.pdf [accessed 2025-12-25]
58. Diagnostic and treatment delay in tuberculosis: an in-depth analysis of the health-seeking behaviour of patients and health system response in seven countries of the eastern mediterranean region. : World Health Organization; 2006 URL: https://applications.emro.who.int/dsaf/dsa710.pdf [accessed 2025-12-25]
59. Adherence to long-term therapies: evidence for action. : World Health Organization; 2003 URL: https://iris.who.int/server/api/core/bitstreams/121c6b73-8651-442f-9560-e3b1d9c8a75c/content [accessed 2025-12-25]
60. Treatment of tuberculosis: guidelines. : World Health Organization; 2010 URL: https://iris.who.int/server/api/core/bitstreams/05c9871e-852d-493c-99a2-c7b419c90098/content [accessed 2025-12-25]
61. Latent tuberculosis infection: updated and consolidated guidelines for programmatic management. : World Health Organization; 2018 URL: https://iris.who.int/server/api/core/bitstreams/d68456c8-2a25-419b-83eb-1dc9df1b838f/content [accessed 2025-12-25]

XSL•FO
RenderX

62. WHO policy on collaborative TB/HIV activities: guidelines for national programmes and other stakeholders. : World Health Organization; 2012 URL: https://iris.who.int/server/api/core/bitstreams/d38d86fb-a877-471b-88a6-9f62fba35790/content [accessed 2025-12-25]
63. The International Union Against Tuberculosis and Lung Disease. Collaborative framework for care and control of tuberculosis and diabetes. : World Health Organization; 2011 URL: https://iris.who.int/server/api/core/bitstreams/60acd5ab-701f-41a7-bea4-e24c72eb1806/content [accessed 2025-12-25]
64. Dhuria M, Sharma N, Ingle G. Impact of tuberculosis on the quality of life. Indian J Community Med 2008 Jan;33(1):58-59. [doi: 10.4103/0970-0218.39249] [Medline: 19967002]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
EQ-5D: EuroQol 5-Dimension Questionnaire
EQ-5D-5L: EuroQol 5-Dimension 5-Level Questionnaire
NTEP: National Tuberculosis Elimination Program
PPSA: Patient-Provider Support Agency
QOL: quality of life
SF-36: Short Form Health Survey (36 items)
TB: tuberculosis

Edited by J Sarvestan; submitted 22.Apr.2025; peer-reviewed by S Vijayan, S Chambers; revised version received 10.Oct.2025; accepted 13.Oct.2025; published 06.Apr.2026.

Please cite as:
Raula N, van der Putten IM, Dukers-Muijrers NHTM, Dapkekar S, Nayak P, Evers SMAA
Impact and Economic Evaluation of the Patient-Provider Support Agency Model Under India's National Tuberculosis Elimination Program: Protocol for a Cohort Study
JMIR Res Protoc 2026;15:e76302
URL: https://www.researchprotocols.org/2026/1/e76302
doi:10.2196/76302

© Nikita Raula, Ingeborg Maria van der Putten, Nicole H T M Dukers-Muijrers, Shankar Dapkekar, Priyakanta Nayak, Silvia M A A Evers. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Personal Support Worker Education, Recruitment, and Retention in Long-Term Care Homes in Ontario, Canada: Protocol for a Mixed Methods Study

Alison Kernoghan[1], MPH; Danielle Just[1], PhD; Catherine E Tong[1,2], PhD; Linda Sheiban Taucar[1,2], RN, MSc; D'Arcy Kirkwood[1], MBS; Benjamin Jay Smith[1], MA; Living Classroom Team[1,3,4]

[1]Research and Knowledge Mobilization, Schlegel-UW Research Institute for Aging, Building 4, 250 Laurelwood Drive, Waterloo, ON, Canada

[2]Centre for Health Care Research & Innovation, Conestoga College, Kitchener, ON, Canada

[3]See Acknowledgments

[4]

Corresponding Author:
Alison Kernoghan, MPH
Research and Knowledge Mobilization, Schlegel-UW Research Institute for Aging, Building 4, 250 Laurelwood Drive, Waterloo, ON, Canada

Abstract

Background: Long-term care (LTC) homes across Canada are facing a crisis in worker recruitment and retention. In the province of Ontario, personal support workers (PSWs) make up 58% of the LTC workforce and play a key role in addressing staffing shortages. Despite the provincial training and education standards implemented in 2014, PSWs working in LTC consistently report feeling unprepared and undersupported and a lack of confidence, leading to high turnover rates. Integrated learning approaches such as the Living Classroom may ease the transition from education to practice, supporting recruitment and retention of PSWs into LTC.

Objective: This study aims to explore potential relationships between Living Classroom models of PSW training and their recruitment and retention in LTC homes in Ontario, Canada, as compared to traditional, or non–Living Classroom, models and how the experience of the Living Classroom impacts the recruitment and retention of these workers.

Methods: We will use a convergent mixed methods study design in which quantitative survey data and qualitative interview data will be collected concurrently. The quantitative arm of the study will recruit new PSW graduates from both Living Classroom and non–Living Classroom programs, delivering online surveys at 1, 6, and 12 months after program completion to assess demographics, employment status and placement, intent to leave, and the changes in those metrics over time. The qualitative arm of the study will recruit PSW students within 8 weeks of program completion for 60-minute semistructured interviews conducted online via Zoom, over the telephone, or in person. Interview transcripts will be analyzed using thematic clustering and line-by-line emergent coding techniques as well as team-based analysis meetings. In addition to an emergent thematic analysis, we will use a priori codes to capture concepts from the quantitative measures, including recruitment into and intent to leave health care and LTC. Data collection and analysis from both arms will be presented in a side-by-side joint display.

Results: This research is funded by the Ontario Ministry of Long-Term Care from April 2024 to March 2026. Enrollment began on March 3, 2025, and is estimated to be completed by December 31, 2025. As of June 16, 2025, we had enrolled 35 participants for interviews and 132 participants in the surveys. Data collection and analysis are currently underway, and the first results are expected to be submitted for publication in 2026.

Conclusions: The results of this study will inform our understanding of the Living Classroom model for PSW education and its impact on recruitment and retention of PSWs into LTC in Ontario. This evidence will enable educators, LTC administrators, and policymakers to further develop their approaches to PSW training and education as one of many vectors to address the staffing shortages faced by the LTC sector.

International Registered Report Identifier (IRRID): DERR1-10.2196/79164

(JMIR Res Protoc 2026;15:e79164) doi:10.2196/79164

KEYWORDS

personal support worker; long-term care; recruitment; retention; education; integrated learning

XSL•FO
RenderX

Introduction

Background

Like many of their international equivalents, long-term care (LTC) homes across Canada are facing concurrent challenges, including a health human resource crisis, rising waitlists, and an aging population increasingly in need of the 24-hour care and congregate living that the sector provides. Worker recruitment and retention is a long-standing issue in LTC, and this was further exacerbated by the COVID-19 pandemic [1]. In the Canadian province of Ontario, where our research is situated, the LTC sector continues to face persistent and sometimes severe staffing shortages [2]. The staff in LTC includes administrators; nurses; allied health professionals; dietary, food, and housekeeping workers; recreational staff; physicians; and others. The largest proportion of LTC staff members are personal support workers (PSWs), who in other jurisdictions are referred to as direct care workers, nurse aides, care aides, and other titles. These support workers are often referred to as "the eyes and ears" of the LTC system [3] given their direct and regular contact with residents.

PSWs are unregulated frontline health care workers who provide direct care to support activities of daily living for individuals across the health care continuum, with most working in LTC, home, and community care settings [4-7]. PSWs are crucial members of the interdisciplinary team in the LTC sector as they are responsible for delivering over 80% of direct care to residents [6]. In Ontario, PSWs make up 58% of the LTC workforce, with most PSWs employed in LTC being female (88.5%) and the average age being 41.2 years [6,7].

Within the Canadian context, escalating pressures to manage the increasingly complex needs of an aging population have significantly evolved the role of PSWs, whose responsibilities have broadened in scope and depth across various clinical settings, including LTC [5,8-11]. Recognizing this expanding scope and the increasing acuity of the residents and patients that PSWs work with [12], in 2014, Ontario introduced education and training standards for the preparation of PSWs [13]. PSW training and certification are provided by community colleges, private career colleges, and continuing education school boards [14]. These programs generally range from 5 to 12 months for full-time students and can extend up to 2 years for part-time students. Courses typically include theory courses, laboratories and practical demonstrations, and a practicum placement in a health care organization (eg, a home care agency or an LTC home). After the pandemic, there has been a notable increase in hybrid PSW training programs in which students complete the bulk of their theory courses online as opposed to classroom work at the school. This broad array of training programs are known as "traditional" PSW education models; however, for the purpose of this study, the term "non–Living Classroom" will be used.

Despite these training and education standards, PSWs working in LTC have consistently reported feelings of being unprepared and undersupported and having a lack of confidence, leading to high turnover rates in LTC [15-22]. In addition to feeling underprepared and lacking confidence, PSWs working in LTC in Ontario also report insufficient staff-to-resident ratios, overworked team members, undertrained colleagues, and high turnover rates that prevent them from providing holistic care [23]. It is in this context that an estimated 40% of PSWs leave the LTC workforce within the first year of graduation and an additional 25% leave after 24 months of work experience [6]. However, Ontario needs more PSWs, not fewer. Researchers project that the Ontario LTC sector will need an additional 11,000 PSWs by 2035 [24], and the government itself projects that the province will need 50,000 more PSWs by 2032 across the continuum of care [25].

Recognizing these challenges, the government of Ontario introduced several initiatives and incentives to attract and retain PSWs in LTC [26]. However, there is a critical need to address PSWs' experience of feeling unprepared and undersupported and having a lack of confidence in the LTC setting. Innovative education models that better prepare PSWs for the unique demands of the LTC sector may be a solution to the high turnover rates experienced within the first year of employment in LTC [20,27]. The Living Classroom model is one such approach, offering an integrated, hands-on learning experience designed to enhance confidence and competence and longer-term retention in the field.

The Living Classroom Model for LTC

In September 2009, Conestoga College; the Schlegel-UW Research Institute for Aging; and Schlegel Villages, an LTC home owner and operator, partnered to open their first Living Classroom, where they welcomed a class of 30 PSW students at the Village of Riverside Glen, a retirement community in Guelph, Ontario [28]. They then went on to create a second Living Classroom at the Village at University Gates, which today provides year-round education for PSWs and practical nursing students. Conceptually rooted in the Teaching Nursing Home movement in the United States [29,30], the purpose of this model is to provide PSW education in the context of an LTC home so that students are actively participating in the home while they complete their education. The Living Classroom model takes an interprofessional approach where faculty, students, LTC teams, residents, and families work together within a culture of integrated learning [28]. The goal is to ease the transition from education to practice in LTC and support recruitment and retention. Preliminary research from this model suggests that students in these programs have increased intention to seek employment in LTC [31] and that this approach has supported new PSW graduates in their transition to practice in LTC [32] and increased their gerontological knowledge [33].

Expansion of the Living Classroom Model Through the Living Classroom Program

In 2023, the Ontario Ministry of Long-Term Care made a significant investment to expand the Living Classroom model in Ontario over 3 years (2023-2026), with the aim to support the training of 1300 new PSWs [34]. The expanded Living Classroom program is delivered on behalf of the province by the Schlegel-UW Research Institute for Aging in collaboration with the Ontario Association of Adult and Continuing Education School Board Administrators. The Living Classroom program had a significant uptake in its first year (2024-2025), doubling

XSL•FO
RenderX

the number of Living Classroom PSW education programs in multiple communities across Ontario, including in northern communities and rural and urban settings. Reflecting the diversity of the province, there are francophone and bilingual Living Classrooms (Canada's 2 official languages are French and English) and Indigenous learning institutes. Reflecting the range of educational offerings across the province, there are Living Classrooms currently associated with public colleges, career colleges, and high school and continuing education school boards that offer PSW training. The Living Classroom models implemented at each site are slightly different and tailored to the students and the community they are supporting.

A program evaluation is being completed to investigate the recruitment and retention of the PSW graduates in LTC through the lens of the students, educational institutions, and LTC homes. The student arm of the evaluation follows students from the beginning of their Living Classroom PSW education to 6 months after their graduation using both surveys and interviews to collect information about their experiences, confidence, and desire to work in LTC. The educational institutions and LTC homes are required to complete a survey at the beginning of their funding period to describe several details, including the delivery model of their Living Classroom, demographic information, educational cohorts, and PSW turnover rates.

Objectives

The purpose of this study is to explore potential relationships between Living Classroom models of PSW training and PSW graduate recruitment and retention in LTC homes in Ontario, Canada. In addition, we are interested in understanding the experience of the Living Classroom educational model on PSW graduates and how it impacts the recruitment and retention of these workers in the LTC sector.

This study will address the following 2 research questions (RQs):

1. How do PSW graduates from Living Classroom programs compare to PSW graduates from non–Living Classroom programs with respect to their characteristics, recruitment to the LTC sector, and retention within the LTC sector?
2. What are the experiences of PSW graduates from Living Classroom programs, and how does their experience impact their intention to work in LTC and their intent to stay in the LTC setting?

Terminology and Concepts

This study distinguishes between recruitment, a PSW graduate's intention to enter the profession, and retention, the longevity of an employee's term within their job or the LTC sector. To address PSW retention, we consider 2 distinct but related concepts: intent to stay and intent to leave.

These are not opposite measures, and while overlapping variables exist, they are unique concepts [35,36]. Intent to stay is strongly correlated with factors of job satisfaction and is indirectly associated with turnover rates, whereas intent to leave is a stronger and direct predictor of future turnover rates [37].

Intent to leave was selected as the theoretical prospective measure of retention of PSW graduates from Living Classroom programs compared to PSW graduates from non–Living Classroom programs. The focus on intent to leave aligns with the theoretical underpinnings of the Living Classroom model. By providing integrated learning and immersive exposure to the LTC environment during training, the model aims to bridge the gap between education and practice to reduce a graduate's likelihood of leaving the sector when entering the workforce.

In contrast, other interventions, such as financial incentive programs, are theorized to increase intent to stay by encouraging graduates to commit to working in LTC for a defined period. Together, these concepts clarify how pathways of education and the workplace influence a graduate's decision to leave or stay within the LTC sector.

Methods

Study Design

We will use a convergent mixed methods study design as there is strong congruence between the proposed RQs and the fundamental elements of this study design [38]. The qualitative and quantitative arms of this study will be conducted simultaneously, with both using distinct methods to explore the potential relationships between training PSWs in a Living Classroom and their recruitment and retention in LTC. The data sources from the quantitative and qualitative arms will be analyzed in tandem, allowing for comparison and supporting our interpretation [39]; however, for participants who complete both arms of the study, their datasets will not be linked.

Setting, Context, and Research Team

Data will be collected in Ontario, Canada, from both PSW students from a Living Classroom program and students who completed their training in a non–Living Classroom program. Data collection for both the qualitative and quantitative arms of the study commenced in March 2025. Qualitative data collection will continue until August 2025, and quantitative data will be collected throughout 2025 and until December 2026, with our team distributing surveys at 1, 6, and 12 months after graduation.

This protocol was developed collaboratively with members of our interdisciplinary research team. This includes early- and late-career research scientists as well as research staff with experience and expertise in mixed methods approaches [40], PSW education and work [41-44], LTC [32,33], and the Living Classroom model itself [33,45]. Our study also draws on the expertise and guidance of a Living Classroom advisory board, which includes representatives from the Ontario Association of Adult and Continuing Education School Board Administrators and LTC homes. This advisory board meets quarterly; they have advised on study design and data collection instruments and, in the future, will support reviewing initial analyses and knowledge dissemination efforts (eg, ensuring readability of plain-language summary reports).

Participants and Recruitment

Quantitative Arm

This study focuses on students in PSW programs in public colleges, school boards, and career colleges in Ontario, who

XSL•FO
RenderX

will complete the surveys. Participants will be eligible for the study if they fulfill the following criteria:

- Being currently enrolled in or having recently graduated from a PSW program in Ontario
- Ability to provide informed consent
- Ability to comprehend and respond to questions in English or French
- Access to digital devices to participate in the online survey

We will exclude individuals without access to a computer or cell phone and individuals who did not graduate from a PSW program in Ontario.

We will use a pragmatic, nonprobability sampling approach [46]. We aim to recruit a sample of 250 PSWs, with approximately 125 from a Living Classroom and 125 from a non–Living Classroom program, by emailing administrators to distribute invitations and QR codes to students 1 month before graduation. Interested students will complete an online contact form, after which letters of information, consent forms, and surveys will be digitally distributed. The study information and consent documents will be available in English and French.

We are mindful of the fact that attrition tends to be higher among research participants who have a lower income, are precariously employed, and face other systemic disadvantages [47], and these characteristics reflect some of our anticipated participants.

Qualitative Arm

We will use purposive recruitment to obtain a maximum variation sample of 30 to 35 PSW graduates from Living Classroom programs across the province, including from all 5 Ontario Health Regions. The sample ensures representation of geographic spread, delivery format (ie, online, in person, and hybrid), educational partner (ie, private and public colleges and school board programs), and LTC partner (ie, for profit, nonprofit, and municipal). Participants will be drawn from the quantitative arm among those who expressed interest in a follow-up interview. The eligibility criteria for the qualitative component of the study include the following:

- Being currently enrolled in or having recently graduated from a Living Classroom program in Ontario
- Ability to provide informed consent
- Ability to complete an interview in English
- Access to an email account to receive study information

Interviews will be conducted within 8 weeks of graduation. Given resource limitations, interviews will only be available in English. Our sample size is informed by principles of data adequacy, ensuring the capture of rich, diverse perspectives [48].

Data Collection

Quantitative Arm

A survey instrument was developed to collect demographic, predictor, and outcome data using LimeSurvey Cloud (version 6.12.0; LimeSurvey GmbH). Surveys are available in English and French following translation and bilingual review. Each survey takes approximately 10 minutes to complete.

To match responses over time while maintaining anonymity, we use self-generated IDs [49]. Data are collected at 1, 6, and 12 months after graduation to track longitudinal outcomes.

Qualitative Arm

Interviews (approximately 60 minutes) will be conducted via videoconferencing, over the telephone, or in person based on participant preference. Interviews will be led by trained interviewers, including 2 PhD-trained qualitative health researchers with prior experience interviewing PSWs and a Master's-level researcher who is a registered nurse and professor in a college PSW education program. All 3 interviewers are women. A semistructured interview guide will be used (see Textbox 1 for sample questions). The interview will focus on participants' experiences with the Living Classroom and their professional plans (ie, intent to leave and intent to work in LTC) and job satisfaction. Interviews will be digitally recorded using an external audio recording device; processed through transcription software (MAXQDA; VERBI GmbH); and then cleaned, reviewed, and anonymized by a trained research assistant using established protocols. Audio recordings and raw and cleaned transcripts will be retained on our secure research drive for future analyses.

Textbox 1. Sample interview guide questions.

- "Tell me about your experience in your Living Classroom PSW program."
- "What parts of the program do you feel have been helpful to prepare you for work as a PSW?"
- "Are there aspects of being a PSW that you feel you need more training on or practice with? Please describe."
- "Do you foresee PSW work as a long-term career for you? Why or why not?"
- "What might lead you to consider leaving work as a PSW?"
- "What do you plan to do when you graduate?" OR "If you have graduated already, are you working currently?"
- "If employed, how did you decide to work in your current job? (e.g., was it where you did your education placement?)"
- "Tell me about your current job satisfaction; what do you like and dislike?"
- "Are you working more than one job? (If yes, probe: the location and sector of other jobs & why they are working multiple jobs)"
- "Was there anything about your education experience that contributed to your plans for work?"

XSL•FO
RenderX

Data Analysis

Quantitative Analysis

Primary outcomes include (1) employment status (binary), (2) employment sector (single selection), and (3) intent to leave (4-point Likert scale). All outcomes will be measured at each time point, and descriptive statistics will be calculated on Living Classroom and non–Living Classroom outcomes to compare employment statuses and sector frequencies, as well as means and SDs of likelihood scale outcomes. Moreover, a McNemar test will be conducted to test for differences in employment statuses between the 2 groups, and a generalized estimating equations model with a logit or probit link function will be performed to test for employment status and temporal interaction effects. A Friedman test will be used to compare between the groups, and a generalized estimating equations model with an ordinal logit link function will be used to compare intent to leave outcomes. All statistical analyses will be performed using Microsoft Excel (version 2511) and R (version 4.5.1; R Foundation for Statistical Computing).

To address missing data due to nonresponses, we will use a combination of listwise or pairwise deletion, multiple imputation (MI), inverse probability weighting, and sensitivity analyses depending on the amount, type, and patterns of missing data as generalized by Mirzaei et al [50]. First, the proportions and patterns of missing data will be analyzed using descriptive statistics, visualization methods, and the Little test to determine the amount, type, and classification of missing data. Second, we will perform listwise deletion under four conditions: (1) if the missing data are demographic outcomes that prevent us from performing paired or subgroup analyses, (2) if there is more than 40% missingness, (3) if it is not possible to address the RQ due to missing data (eg, the participant did not respond to crucial items related to intent to leave or working status), and (4) if missingness prevents a temporal analysis. Third, in similar fashion to the second step, a pairwise deletion will be conducted if a certain item has more than 40% missingness. Fourth, to address nonresponses, MI (eg, MI by chained equations, full information maximum likelihood, and k-nearest neighbor) and inverse probability weighting will be conducted, wherein the specific method will be influenced by the proportions and patterns from step 2. A sensitivity analysis will be conducted to test for robustness given missing not at random data using techniques such as, though not limited to, pattern mixture models and scenario analyses.

Qualitative Analysis

We will upload transcripts and field notes or debriefing notes into the MAXQDA software. Following processes that our team has used elsewhere [51], our team-based, thematic qualitative analysis will include 4 steps. The first step is coding and theming the data. Researchers will read through a portion of the dataset (eg, 3-4 transcripts) and create an initial set of codes and definitions. We will then identify themes using a clustering technique [52]. Each cluster will have a proposed name, brief description, illustrative quotations from the data, and a list of codes that support it. In the second step, trained qualitative analysts who were also involved in the interviews will use line-by-line emergent coding techniques as outlined by Saldaña [53] to code the dataset in its entirety. In the third step, extensive memoing and team-based analysis meetings will take place throughout data collection and the theming and interpretation phase. In the fourth step, narrative descriptions will be prepared, with quotes that summarize and illustrate high-level themes. In addition to an emergent thematic analysis, in step 2 (line-by-line coding), we will also include a series of a priori codes and questions to code for to specifically capture concepts that are prominent in the quantitative measures (eg, recruitment to health care and LTC, intent to leave health care and LTC, and satisfaction with the Living Classroom experience). This alignment of core concepts in the quantitative and qualitative datasets will support our mixed methods analysis, which is further detailed below. Throughout this process, we will keep track of how our coding structure and definitions evolve, for example, if an initial code is expanded upon, collapsed, or even deleted. Final reports will use verbatim quotes, in which grammar, natural language, and repeated phrases, among other aspects, will be retained, staying true to participants' voices [54].

In our analyses, we will report gender when using participant quotes and data. Where appropriate or relevant, we will report findings for the total sample and disaggregate them by gender in accordance with the Sex and Gender Equity in Research guidelines [55]; we will achieve this by specifically coding for gender and the intersection of gender with other statuses (eg, international student status), adopting an intersectional lens [56].

Strategies for Rigor

Our strategies for rigor will include retaining a record of methodological and analytical decisions in an audit trail [57], peer reviewing and debriefing [58], clarifying researcher biases [58], and the careful examination of outliers [53]. Our team-based analysis process will benefit from the presence of an interdisciplinary team (PhD-trained qualitative health researchers, a public health professional, and a registered nurse who has taught in PSW programs), each bringing a different lens to the data.

Concurrent Mixed Methods Analysis

The quantitative outcome measures and qualitative interview questions examine overlapping concepts such as the intent to leave, recruitment, and retention. Following data collection and analysis from both arms, the results will be presented in a side-by-side joint display [59]. Our parallel analyses and presentation of the quantitative and qualitative findings in tandem will facilitate further analysis as a result of confirmation, expansion, or discordance between the 2 forms of data [38].

Ethical Considerations

This study has been reviewed by and received approval from the Conestoga College Institute of Technology and Advanced Learning Research Ethics Board (586). Informed consent will be gathered as the first item on each questionnaire. Those who do not wish to consent will not continue on to complete the remainder of the questionnaire. Participants in the qualitative arm will have the opportunity to ask any questions and provide verbal or written consent at the start of the interviews after

XSL•FO
RenderX

having reviewed the study information letter ahead of the interview. This research focuses on the experiences and perspectives of new PSW graduates, many of whom are international students with precarious status in the country (ie, student visas that will expire when schooling is complete) and uncertain pathways to employment. Many PSWs are racialized women and/or individuals who may have a lower socioeconomic status. As such, we take into consideration the guidance provided by van den Hoonaard and van den Hoonaard [60] on conducting qualitative research with potentially vulnerable populations and the specific guidance by Liamputtong [61] on working with racialized participants and in cross-cultural contexts. For example, our team has protocols and plans in place to respond to a participant who is distressed or expresses an unmet need, and we are mindful of power differentials (eg, not overtly identifying ourselves as PhD-trained scientists), creating a safe space for confidential and authentic conversations (eg, reiterating that any question can be skipped) and using plain language wherever possible. Furthermore, identifying information is not collected through questionnaires or interviews. Where identifying details such as email addresses or names are necessary to contact interview or survey participants, that information will be stored securely and separately from the study data. To support retention in the study, participants receive a CAD $20 (US $15) honorarium via a digital gift card for each of the 3 survey time points completed. Participants in the qualitative arm of the study will receive a CAD $50 (US $36) digital gift card.

Dissemination

We will present the findings from our research at national, well-attended gerontology conferences (eg, the Canadian Association on Gerontology), as well as relevant industry conferences (eg, Together We Care, This Is Long Term Care, and/or AdvantAge Ontario, annual meetings that together represent nearly all of the LTC operators in Ontario). We will prepare a knowledge brief for decision-makers and a final research report for the program funder. A plain-language summary will be available for study participants. Dissemination will also include peer-reviewed manuscripts in open access journals and accompanying social media posts to highlight new manuscripts as they become publicly available.

Results

Funding for this research was provided by the Ontario Ministry of Long-Term Care from April 1, 2024, to March 31, 2026; enrollment began on March 3, 2025, and is estimated to be completed by December 31, 2025. As of June 16, 2025, we had enrolled 35 participants for interviews and 132 participants in the surveys. Data collection and analysis are currently underway, and the first results are expected to be submitted for publication in 2026.

Discussion

Expected Findings

We anticipate that this study will demonstrate a positive correlation between the Living Classroom model and improved recruitment and retention of PSWs in LTC. Specifically, we anticipate that Living Classroom graduates may report greater confidence and readiness to work in LTC. The immersive nature of the Living Classroom, where students are integrated into the LTC home early and often, is designed to support the socialization of PSW students and may increase their intent to stay.

Existing literature highlights the benefits of the Living Classroom model for student learning [31,62], but there is a gap in longitudinal data that consider employment outcomes and longer-term retention of PSWs. This study builds on previous descriptive work using a convergent mixed methods design to compare graduates from Living Classroom and non–Living Classroom programs. Participants are followed for 12 months after graduation, extending the findings of previous studies to understand the transition of these students to the workforce.

Strengths and Limitations

A strength of this study is the longitudinal, mixed methods approach, allowing us to quantify recruitment and retention rates while also capturing the nuanced descriptions of the students' experiences through interviews. The interdisciplinary nature of the research team and the involvement of an advisory board further ensure that the findings are grounded in the practical realities of the LTC sector in Ontario. This work will represent the largest sample to date in any study specific to PSWs in Living Classrooms [31,32,62] and one of the largest samples to date of care aides in the Teaching Nursing Home initiative, where evaluations have focused on resident outcomes, nurses, and schools [63].

We recognize the current limitations of this work, and additional limitations may become apparent as we conduct this research. Due to resource limitations, we are only conducting in-depth interviews in English; as such, the experiences of students in francophone Living Classroom programs may not be represented in the qualitative arm of the study. Second, while we seek to understand students' intent to stay and retention in the sector, our time-limited study does not allow for longer-term follow-up beyond 12 months. In addition, participant attrition is expected, which may limit our sample sizes and, consequently, the statistical analyses that we are able to perform. Our pragmatic, nonprobability sampling approach means that our statistical analyses are not necessarily generalizable but also reflects known issues with accessing this population [64].

Future Directions

We are working to secure additional research funding, which would allow us to continue following Living Classroom students for a longer period. Current participants have the option to provide their contact information for potential inclusion in future research. Many of the Living Classrooms that we have recruited from are newly formed and newly funded and, as such, do not yet represent a fully mature Living Classroom program. We anticipate future research and evaluation initiatives to continue assessing the development and longer-term impact of Living Classrooms in Ontario and beyond.

XSL•FO
RenderX

Conclusions

The results of this study will inform our understanding of the Living Classroom model for PSW education and its impact on recruitment and retention of PSWs into LTC in Ontario. This evidence will enable educators, LTC administrators, and policymakers to further develop their approaches to PSW training and education as one of many vectors to address the staffing shortages faced by the LTC sector.

Acknowledgments

The authors wish to thank Amanda Nova, Hiroyoshi Hiratsuka, and Katrina Hitchman for their contributions to the conceptualization and refinement of the research protocol. Furthermore, they would also like to acknowledge everyone who has contributed to the establishment of Living Classrooms across Ontario. The collaborators for the Living Classroom team are Angela Butt, Anita Plunkett, Brittany Howlett, Carolyn Triemstra, Dana Kirkbride, Danica Bui, Hilda Smith, Jennifer Beaudin, Katrina Street, Patty Attwell, Paul Cox, and Tammy Cantin.

Funding

This work is funded by the government of Ontario as part of the Living Classroom program. The Living Classroom program is led by the Schlegel-UW Research Institute for Aging through the Ontario Centres for Learning, Research, and Innovation in Long-Term Care in partnership with the Ontario Association of Adult and Continuing Education School Board Administrators. The funder was not involved in the development of the research questions and objectives; data collection, analysis, and interpretation of the data; and/or the decision to submit the manuscript for publication, although they are aware of our intent to publish.

Data Availability

The datasets generated or analyzed during this study are not publicly available as they contain information that could compromise the privacy of the research participants but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: AK, CET, LST, DK, DJ
Data curation: BJS, DK
Formal analysis: CET, LST
Funding acquisition: AK
Investigation: CET, LST, DJ
Methodology: CET, LST, DJ
Project administration: AK, BJS, DK
Supervision: AK, CET, LST
Validation: CET
Writing—original draft: CET, LST, DJ
Writing—review and editing: AK, BJS, CET, LST, DJ

Conflicts of Interest

None declared.

References

1. Estabrooks CA, Ewa V, Keefe J, Straus SE. The predictable crisis of Covid-19 in Canada's long term care homes. BMJ 2023 Jul 24;382:e075148. [doi: 10.1136/bmj-2023-075148] [Medline: 37487622]
2. Berta WB, Stewart C, Baumann A. What is old is new again: global issues influencing workers and their work in long-term care. Healthc Policy 2022 Jun;17(SP):66-90. [doi: 10.12927/hcpol.2022.26853] [Medline: 35848557]
3. Kaasalainen S, Brazil K, Kelley ML. Building capacity in palliative care for personal support workers in long-term care through experiential learning. Int J Older People Nurs 2014 Jun;9(2):151-158. [doi: 10.1111/opn.12008] [Medline: 23051621]
4. Recommendations for advancing pan-Canadian data capture for personal support workers. Canadian Institute for Health Information. 2023. URL: https://web.archive.org/web/20250623205935/https://www.cihi.ca/sites/default/files/document/recommendations-advancing-pan-canadian-data-capture-psws-report-en.pdf [accessed 2025-06-06]
5. Just D. Personal support workers' role in end-of-life care in long-term care [Dissertation]. : University of Toronto; 2024 URL: https://www.proquest.com/openview/eb11db789a9547b135a54ff0ea56adda/1?cbl=18750&diss=y&pq-origsite=gscholar [accessed 2025-04-04]
6. Long-term care staffing study. Government of Ontario. 2020. URL: https://files.ontario.ca/mltc-long-term-care-staffing-study-en-2020-07-31.pdf [accessed 2024-09-17]
7. Zagrodney KAP, Deber R, Saks M, Laporte A. Personal support worker socio-demographic differences across care sectors in Canada. J Appl Gerontol 2023 Apr;42(4):670-679. [doi: 10.1177/07334648221142301] [Medline: 36464973]

XSL•FO
RenderX

8. Afzal A, Stolee P, Heckman GA, Boscart VM, Sanyal C. The role of unregulated care providers in Canada-a scoping review. Int J Older People Nurs 2018 Sep;13(3):e12190. [doi: 10.1111/opn.12190] [Medline: 29575512]
9. Berta W, Laporte A, Deber R, Baumann A, Gamble B. The evolving role of health care aides in the long-term care and home and community care sectors in Canada. Hum Resour Health 2013 Jun 14;11:25. [doi: 10.1186/1478-4491-11-25] [Medline: 23768158]
10. Kalles EA, King EC, Holyoke P. How high-performing personal support workers set and maintain boundaries when providing care: a case study in Ontario, Canada. J Appl Gerontol 2025 Jan;44(1):83-94. [doi: 10.1177/07334648241263500] [Medline: 39636657]
11. Saari M, Patterson E, Kelly S, Tourangeau AE. The evolving role of the personal support worker in home care in Ontario, Canada. Health Soc Care Community 2018 Mar;26(2):240-249. [doi: 10.1111/hsc.12514] [Medline: 29119635]
12. Ng R, Lane N, Tanuseputro P, et al. Increasing complexity of new nursing home residents in Ontario, Canada: a serial cross-sectional study. J Am Geriatr Soc 2020 Jun;68(6):1293-1300. [doi: 10.1111/jgs.16394] [Medline: 32119121]
13. Kelly C. Exploring experiences of personal support worker education in Ontario, Canada. Health Soc Care Community 2017 Jul;25(4):1430-1438. [doi: 10.1111/hsc.12443] [Medline: 28294449]
14. Kelly C, Bourgeault I. Developing a common education standard for personal support workers in Ontario. Health Reform Obs 2015;3(1). [doi: 10.13162/hro-ors.v3i1.169]
15. Alam S, Drakos A. Staffing crisis: addressing the high turnover rate of personal support workers in Ontario's long-term care sector. Meducator 2022;1(39). [doi: 10.15173/m.v1i39.3301]
16. Awosoga OA, Odole AC, Onyeso OK, et al. Perceived strategies for reducing staff-turnover and improving well-being and retention among professional caregivers in Alberta's continuing-care facilities: a qualitative study. Home Health Care Serv Q 2023;42(3):193-215. [doi: 10.1080/01621424.2023.2166889] [Medline: 36646111]
17. Boamah SA, Weldrick R, Havaei F, Irshad A, Hutchinson A. Experiences of healthcare workers in long-term care during COVID-19: a scoping review. J Appl Gerontol 2023 May;42(5):1118-1136. [doi: 10.1177/07334648221146252] [Medline: 36541274]
18. Keefe JM, Duynisveld A, Stevens S, Estabrooks C. Quality of work life and mental well-being for long-term care staff in Nova Scotia. Healthc Policy 2024 Aug;19(4):32-41. [doi: 10.12927/hcpol.2024.27348] [Medline: 39229661]
19. Li X, Dorstyn D, Mpofu E, et al. Nursing assistants and resident satisfaction in long-term care: a systematic review. Geriatr Nurs 2021;42(6):1323-1331. [doi: 10.1016/j.gerinurse.2021.08.006] [Medline: 34560527]
20. Rivett E, Hammond L, West J. What influences self-perceived competence and confidence in dementia care home staff? A systematic review. Psychogeriatrics 2019 Sep;19(5):440-456. [doi: 10.1111/psyg.12422] [Medline: 30809893]
21. Savundranayagam MY, Docherty-Skippen SM, Basque SR. Qualitative insights into the working conditions of personal support workers in long-term care in the context of a person-centered communication training intervention. Res Gerontol Nurs 2021;14(5):245-253. [doi: 10.3928/19404921-20210708-01] [Medline: 34292804]
22. Scales K. Transforming direct care jobs, reimagining long-term services and supports. J Am Med Dir Assoc 2022 Feb;23(2):207-213. [doi: 10.1016/j.jamda.2021.12.005] [Medline: 34973168]
23. PSW perspectives on the staffing challenge in long-term care. Centers for Learning, Research & Innovation in Long-Term Care. 2021. URL: https://clri-ltc.ca/wp-content/uploads/2021/02/PSW_Perspectives_FinalReport_Feb25_Accessible.pdf [accessed 2025-06-06]
24. Dass AR, Deber R, Laporte A. Forecasting staffing needs for Ontario's long-term care sector. Healthc Policy 2022 Jun;17(SP):91-106. [doi: 10.12927/hcpol.2022.26852] [Medline: 35848558]
25. Jones A. Ontario will need tens of thousands of new nurses, PSWs by 2032. The Canadian Press. 2024 May 13. URL: https://www.cbc.ca/news/canada/toronto/nurses-psws-ontario-foi-document-1.7202282 [accessed 2025-04-24]
26. Attracting new PSW graduates to long-term care and home and community care sectors. December. 2024. URL: https://web.archive.org/web/20250623020635/https://www.ontariohealth.ca/system-planning/funding-opportunities/attracting-new-psw-graduates [accessed 2025-04-04]
27. Newbould L, Samsi K, Wilberforce M. Developing effective workforce training to support the long-term care of older adults: a review of reviews. Health Soc Care Community 2022 Nov;30(6):2202-2217. [doi: 10.1111/hsc.13897] [Medline: 35791508]
28. Shared learning environments for PSW education. The Living Classroom. 2019. URL: https://livingclassroom.ca/wp-content/uploads/2022/06/RIA-Living-Classroom-Guide-2019-revised.pdf [accessed 2025-04-04]
29. Chilvers JR, Jones D. The teaching nursing homes innovation: a literature review. J Adv Nurs 1997 Sep;26(3):463-469. [doi: 10.1046/j.1365-2648.1997.t01-3-00999.x] [Medline: 9378864]
30. Mezey MD, Mitty EL, Bottrell M. The Teaching Nursing Home Program: enduring educational outcomes. Nurs Outlook 1997;45(3):133-140. [doi: 10.1016/s0029-6554(97)90044-0] [Medline: 9210161]
31. Boscart VM, d'Avernas J, Brown P, Raasok M. Changing the impact of nursing assistants' education in seniors' care: the living classroom in long-term care. Can Geriatr J 2017 Mar 31;20(1):15-21. [doi: 10.5770/cgj.20.238] [Medline: 28396705]
32. Boscart V, d'Avernas J, van der Horst ML, Garbutt K, Sheiban Taucar L, Raasok M. Shared learning environments for unregulated care provider education in long-term care: innovative approaches and key considerations. Gerontol Geriatr Educ 2022;43(1):75-83. [doi: 10.1080/02701960.2019.1645015] [Medline: 31317832]

XSL•FO
RenderX

33. Boscart V, Crutchlow LE, Sheiban Taucar L, et al. Chronic disease management models in nursing homes: a scoping review. BMJ Open 2020 Feb 5;10(2):e032316. [doi: 10.1136/bmjopen-2019-032316] [Medline: 32029485]
34. Ontario training thousands of new long-term care staff. Government of Ontario. 2024 Jan 31. URL: https://news.ontario.ca/en/release/1004135/ontario-training-thousands-of-new-long-term-care-staff [accessed 2025-06-04]
35. Nancarrow S, Bradbury J, Pit SW, Ariss S. Intention to stay and intention to leave: are they two sides of the same coin? A cross-sectional structural equation modelling study among health and social care workers. J Occup Health 2014;56(4):292-300. [doi: 10.1539/joh.14-0027-oa] [Medline: 24953092]
36. Rosen J, Stiehl EM, Mittal V, Leana CR. Stayers, leavers, and switchers among certified nursing assistants in nursing homes: a longitudinal investigation of turnover intent, staff retention, and turnover. Gerontologist 2011 Oct;51(5):597-609. [doi: 10.1093/geront/gnr025] [Medline: 21498629]
37. Cho S, Johanson MM, Guchait P. Employees intent to leave: a comparison of determinants of intent to leave versus intent to stay. Int J Hosp Manag 2009 Sep;28(3):374-381. [doi: 10.1016/j.ijhm.2008.10.007]
38. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 5th edition: SAGE Publications; 2018.
39. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156. [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
40. Tong C, Kernoghan A, Lemmon K, et al. Lessons and reflections from an extended co-design process developing an mHealth app with and for older adults: multiphase, mixed methods study. JMIR Aging 2022 Oct 28;5(4):e39189. [doi: 10.2196/39189] [Medline: 36306166]
41. Craven C, Byrne K, Sims-Gould J, Martin-Matthews A. Types and patterns of safety concerns in home care: staff perspectives. Int J Qual Health Care 2012 Oct;24(5):525-531. [doi: 10.1093/intqhc/mzs047] [Medline: 22864107]
42. Just DT, O'Rourke HM, Berta WB, Variath C, Cranley LA. Expanding the concept of end-of-life care in long-term care: a scoping review exploring the role of healthcare assistants. Int J Older People Nurs 2021 Mar;16(2):e12353. [doi: 10.1111/opn.12353] [Medline: 33124160]
43. Just DT, O'Rourke HM, Berta WB, Cranley LA. Navigating the virtual landscape: methodological considerations for qualitative research in long-term care. Int J Qual Methods 2024;23:1-12. [doi: 10.1177/16094069241244859]
44. Sims-Gould J, Byrne K, Craven C, Martin-Matthews A, Keefe J. Why I became a home support worker: recruitment in the home health sector. Home Health Care Serv Q 2010 Oct;29(4):171-194. [doi: 10.1080/01621424.2010.534047] [Medline: 21153997]
45. Boscart V, Davey M, Crutchlow L, et al. Effective chronic disease interventions in nursing homes: a scoping review based on the knowledge-to-action framework. Clin Gerontol 2022;45(5):1073-1086. [doi: 10.1080/07317115.2019.1707339] [Medline: 31902314]
46. Ahmed SK. How to choose a sampling technique and determine sample size for research: a simplified guide for researchers. Oral Oncol Rep 2024 Dec;12:100662. [doi: 10.1016/j.oor.2024.100662]
47. Young AF, Powers JR, Bell SL. Attrition in longitudinal studies: who do you lose? Aust N Z J Public Health 2006 Aug;30(4):353-361. [doi: 10.1111/j.1467-842x.2006.tb00849.x] [Medline: 16956166]
48. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Nov 21;18(1):148. [doi: 10.1186/s12874-018-0594-7] [Medline: 30463515]
49. Yurek LA, Vasey J, Sullivan Havens D. The use of self-generated identification codes in longitudinal research. Eval Rev 2008 Oct;32(5):435-452. [doi: 10.1177/0193841X08316676] [Medline: 18477737]
50. Mirzaei A, Carter SR, Patanwala AE, Schneider CR. Missing data in surveys: key concepts, approaches, and applications. Res Social Adm Pharm 2022 Feb;18(2):2308-2316. [doi: 10.1016/j.sapharm.2021.03.009] [Medline: 33775556]
51. Tong CE, Lopez KJ, Chowdhury D, et al. Understanding racialised older adults' experiences of the Canadian healthcare system, and codesigning solutions: protocol for a qualitative study in nine languages. BMJ Open 2022 Oct 10;12(10):e068013. [doi: 10.1136/bmjopen-2022-068013] [Medline: 36216419]
52. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
53. Saldaña J. The Coding Manual for Qualitative Researchers, 4th edition: SAGE Publications; 2021.
54. Corden A, Sainsbury R. Exploring 'quality': research participants' perspectives on verbatim quotations. Int J Soc Res Methodol 2006;9(2):97-110. [doi: 10.1080/13645570600595264]
55. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. Res Integr Peer Rev 2016 May 3;1:2. [doi: 10.1186/s41073-016-0007-6] [Medline: 29451543]
56. Abrams JA, Tabaac A, Jung S, Else-Quest NM. Considerations for employing intersectionality in qualitative health research. Soc Sci Med 2020 Aug;258:113138. [doi: 10.1016/j.socscimed.2020.113138] [Medline: 32574889]
57. Cutcliffe JR, McKenna HP. Expert qualitative researchers and the use of audit trails. J Adv Nurs 2004 Jan;45(2):126-133. [doi: 10.1046/j.1365-2648.2003.02874.x] [Medline: 14705996]
58. Morse JM. Critical analysis of strategies for determining rigor in qualitative inquiry. Qual Health Res 2015 Sep;25(9):1212-1222. [doi: 10.1177/1049732315588501] [Medline: 26184336]

XSL•FO
RenderX

59. Guetterman TC, Fàbregues S, Sakakibara R. Visuals in joint displays to represent integration in mixed methods research: a methodological review. Methods Psychol 2021 Dec;5:100080. [doi: 10.1016/j.metip.2021.100080]
60. van den Hoonaard WC, van den Hoonaard DK. Essentials of Thinking Ethically in Qualitative Research, 1st edition: Routledge; 2013. [doi: 10.4324/9781315429175]
61. Liamputtong P. Performing Qualitative Cross-Cultural Research, 1st edition: Cambridge University Press; 2010.
62. Boscart VM, McCleary L, Stolee P, et al. Enhancing nursing home care for seniors: impact of a living classroom on nursing assistant's education. Educ Gerontol 2020;46(8):461-472. [doi: 10.1080/03601277.2020.1774842]
63. Barnett K. Exploring the Teaching Nursing Home Model: Literature Review to Inform the National Evaluation of the TRACS Program. : Australian Workplace Innovation and Social Research Centre, The University of Adelaide; 2014. [doi: 10.13140/RG.2.1.2677.3921]
64. Zagrodney KA, King EC, McKay SM. Two gateways to understanding the health workforce: insights from working with organizational and population-level data. Healthc Manage Forum 2025 Nov;38(1_suppl):S24-S27. [doi: 10.1177/08404704251361919] [Medline: 41151075]

Abbreviations

LTC: long-term care
MI: multiple imputation
PSW: personal support worker
RQ: research question

Edited by J Sarvestan; submitted 16.Jun.2025; peer-reviewed by V Fiset; revised version received 24.Feb.2026; accepted 26.Feb.2026; published 15.Apr.2026.

Please cite as:
Kernoghan A, Just D, Tong CE, Sheiban Taucar L, Kirkwood D, Smith BJ, Living Classroom Team
Exploring Personal Support Worker Education, Recruitment, and Retention in Long-Term Care Homes in Ontario, Canada: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e79164
URL: https://www.researchprotocols.org/2026/1/e79164
doi:10.2196/79164

© Alison Kernoghan, Danielle Just, Catherine E Tong, Linda Sheiban Taucar, D'Arcy Kirkwood, Benjamin Jay Smith, Living Classroom Team. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessment of Programmed Cell Death Ligand-1 Expression in Oral Potentially Malignant Disorders and Tumor-Free Surgical Margins of Oral Squamous Cell Carcinoma: Protocol for a Retrospective Cross-Sectional Study

Archana Sonone, BDS, MDS; Alka Hande, BDS, MDS, PhD; Swati Patil, BDS, MDS; Aayushi Pakhale, BDS, MDS; Preethi Sharma, BDS, MDS, PhD

Department of Oral Pathology and Microbiology, Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Corresponding Author:
Archana Sonone, BDS, MDS
Department of Oral Pathology and Microbiology, Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research, Sawangi (Meghe), Wardha, Maharashtra, India

Abstract

Background: Oral squamous cell carcinoma (OSCC) is a highly prevalent and aggressive malignancy of the oral cavity, frequently preceded by oral potentially malignant disorder (OPMD). Despite therapeutic advances, survival rates remain unsatisfactory, primarily due to late diagnosis, recurrence, and molecular alterations in histologically tumor-free surgical margins. Programmed cell death ligand-1 (PD-L1), an immune checkpoint molecule, contributes to tumor immune evasion and has been implicated in cancer progression. Its expression in OPMDs and OSCC surgical margins may serve as an early indicator of malignant transformation and recurrence risk.

Objective: This study aims to assess PD-L1 expression in OPMDs and histologically negative surgical margins of OSCC and evaluate their association with 3-year survival outcomes.

Methods: This retrospective cross-sectional study will be conducted over 12 months at a tertiary care hospital in Sawangi Meghe, Wardha, India. Archived formalin-fixed, paraffin-embedded samples of OPMDs and tumor-free surgical margins from OSCC cases (2018 - 2020) will be retrieved. Immunohistochemistry for PD-L1 will be performed using the SP263 clone, and expression will be evaluated using the combined positive score. Demographic, clinical, and survival data will be collected from patient records. Statistical analysis will determine correlations among PD-L1 expression, clinicopathological variables, and 3-year survival.

Results: The study is expected to provide insights into the role of PD-L1 as a biomarker for early detection, prognostication, and risk stratification in patients with OPMDs and OSCC. Data collection and immunohistochemical analysis have not yet commenced at the time of submission.

Conclusions: By evaluating PD-L1 expression in premalignant lesions and histologically negative margins, this study aims to identify molecular predictors of OSCC progression and survival. The findings may help establish PD-L1 as a prognostic biomarker and support its integration into precision oncology and immunotherapy strategies.

International Registered Report Identifier (IRRID): PRR1-10.2196/82568

(JMIR Res Protoc 2026;15:e82568) doi:10.2196/82568

KEYWORDS

oral squamous cell carcinoma; oral potentially malignant disorders; surgical margins; programmed cell death ligand-1; PD-L1; immunohistochemistry; prognosis; survival

Introduction

Background

The most prevalent cancer that affects the oral and maxillofacial areas is oral squamous cell carcinoma (OSCC), which still has a major negative impact on global health. Head and neck cancers rank eighth worldwide according to GLOBOCAN 2020, with OSCC alone responsible for over 440,000 annual deaths and approximately 880,000 new cases [1,2]. Oral potentially malignant disorders (OPMDs), which impact approximately 2.5% of the general population, are the cause of a significant percentage of OSCCs [3]. Oral leukoplakia, oral submucous fibrosis, and oral lichen planus are the most common OPMDs;

XSL•FO
RenderX

they all have varying but noteworthy potential for malignant transformation [3].

Overall 5-year survival rates for OSCC have remained at roughly 50%, indicating that the prognosis has not improved much in recent decades despite advancements in surgery, radiotherapy, and chemotherapy [2]. Although most patients present at advanced stages, the prognosis is better in early, well-differentiated tumors. Local recurrence is frequent even in cases with histologically tumor-free margins, which suggests that conventional histopathology frequently misses minimal residual cancer cells [4,5].

Immunohistochemistry (IHC) has become a useful supplementary technique for identifying molecular alterations in tissues that seem histologically normal to get around this restriction. IHC analysis of negative surgical margins may be useful in identifying patients who are more likely to experience a recurrence because genetic and molecular changes occur before phenotypic transformation [5]. Nevertheless, the effectiveness of these methods mainly depends on trustworthy biomarkers that can differentiate cancerous from healthy epithelial cells and offer prognostic information [6].

The programmed cell death ligand-1 (PD-L1) is one of the markers that has been studied the most in recent years. The type I transmembrane protein PD-L1, also referred to as CD274 or B7-H1, is a member of the B7 family and consists of an intracellular cytoplasmic domain, a transmembrane segment, and extracellular immunoglobulin V– and immunoglobulin C–like domains [7,8]. By binding to the programmed cell death receptor-1 (PD-1) expressed on T cells, PD-L1 plays a crucial role in immune checkpoint regulation under physiological conditions. This inhibits T-cell proliferation, cytokine production, and cytotoxic activity [8,9]. The prevention of autoimmunity and preservation of immunological tolerance depend on this pathway. Nevertheless, abnormal PD-L1 expression in the tumor microenvironment allows cancerous cells to elude immune monitoring, which advances the course of the disease [10,11].

It is still unclear how PD-L1 functions in the early phases of oral carcinogenesis and in predicting OSCC recurrence despite the fact that it is now a recognized biomarker for directing immune checkpoint inhibitor treatments in a number of cancers [9,12]. According to recent data, PD-L1 may be expressed in both tumor-free margins of OSCC and OPMDs, suggesting that it plays a role in residual disease as well as malignant transformation [3,10,12]. The clinical significance of this pathway has also been highlighted by clinical trials using immune checkpoint inhibitors such as pembrolizumab and nivolumab, which have shown antitumor efficacy in recurrent and metastatic head and neck squamous cell carcinoma (HNSCC) [11,13].

PD-L1 expression is characterized by significant biological and technical heterogeneity, influenced by tumor microenvironment, inflammatory infiltrate, assay, platforms, antibody clone, and scoring algorithms. In HNSCC, variability in PD-L1 expression across tumor compartments and disease stages has been widely reported, complicating the interpretation of PD-L1 as a stand-alone biomarker by clinicians and researchers. Furthermore, discordance among primary tumor, surgical margins, and premalignant lesions highlights the dynamic nature of PD-L1 expression during carcinogenesis. Evidence from other epithelial malignancies, including lung and breast cancer, underscores the prognostic relevance of PD-L1 while simultaneously emphasizing challenges related to cutoff selections, interobserver variability, and the temporal instability of expression. Recognizing these nuances is essential when evaluating PD-L1 as a screening or prognostic biomarker in OPMDs and OSCC.

Several studies have demonstrated that PD-L1 expression is not restricted to advanced disease but may be detectable during early oral carcinogenesis. Yoshida et al [14] reported significant PD-L1 expression in tongue squamous cell carcinoma, suggesting its involvement in tumor development and progression. Furthermore, Schneider et al [15] demonstrated that PD-1 and PD-L1 expression in primary HNSCC and corresponding lymph node metastases had a significant impact on clinical outcomes. These findings support the hypothesis that PD-L1 expression may serve as an early molecular indicator of aggressive biological behavior in oral malignancies [14,15].

Biomarker-driven strategies that support risk stratification, early detection, and prognosis prediction are desperately needed, especially in light of the increasing prevalence of OPMDs and OSCC. In light of this, this study intends to assess PD-L1 expression in OSCC tumor-free surgical margins and OPMDs, as well as investigate its relationship to clinicopathological characteristics and patient survival. This study aims to produce evidence that might support the use of PD-L1 as a prognostic marker in the clinical management of OSCC by examining the continuum from premalignant lesions to recurrence.

Study Objectives

Primary Objective

The primary objective is to compare PD-L1 combined positive score (CPS) expression between OPMDs and tumor-free surgical margins of OSCC.

Secondary Objectives

The secondary objectives are to (1) evaluate PD-L1 CPS expression in the normal oral mucosa as a reference group and (2) assess the association between PD-L1 CPS expression in tumor-free surgical margins and 3-year overall survival (OS) and disease-free survival (DFS) in patients with OSCC.

Study Outcomes

Primary Outcome

The primary outcome is PD-L1 expression, measured using the CPS in OPMDs and tumor-free surgical margins of OSCC.

Secondary Outcomes

The secondary outcomes are as follows: (1) PD-L1 CPS expression in the normal oral mucosa, (2) 3-year OS and DFS in patients with OSCC, and (3) the correlation between PD-L1 CPS expression in tumor-free margins and survival outcomes.

XSL•FO
RenderX

Methods

Overview

This retrospective analysis will be conducted between June 2025 and June 2026 using archival samples collected from 2018 to 2020. All analyses will be performed on existing clinical and survival data available in patient records. This is a retrospective study conducted in the Department of Oral Pathology and Microbiology at Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research (Deemed to be University), located in Sawangi (M), Wardha, Maharashtra, India. Archived formalin-fixed, paraffin-embedded tissue samples and the corresponding clinical records of patients diagnosed with OPMD and surgically treated OSCC between January 2018 and December 2020 will be retrieved from departmental archives. All clinical outcomes, including survival status, recurrence, and follow-up details are based on retrospective review of medical records, with a minimum follow-up duration of 3 years where available. The study will be conducted over a 12-month period (June 2025 to June 2026), which will include retrieval of archival samples, immunohistochemical analysis, data extraction, statistical analysis, and manuscript preparation. No prospective data collection or follow-up is planned. The eligibility criteria, including detailed inclusion and exclusion parameters, are summarized in Table 1. This study protocol has been prepared in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for retrospective studies.

Table . Inclusion and exclusion criteria.

Category	Description
Inclusion criteria	Patients with a primary diagnosis of OPMDs[a] or OSCC[b], tumor sites restricted to the oral cavity, and histologically tumor-free surgical margins
Exclusion criteria	Recurrent tumors, patients lost to follow-up, variants of squamous cell carcinoma, and patients with histologically positive surgical margins

[a]OPMD: oral potentially malignant disorder.

[b]OSCC: oral squamous cell carcinoma.

The sample will be divided into three groups: (1) histologically diagnosed OPMD cases (20 samples; group A), (2) histologically tumor-free surgical margins of OSCC cases (20 samples; group B), and (3) normal oral mucosa (20 samples; group C).

The departmental archives will be searched for tumor-free surgical margins of OSCC cases and histopathologically diagnosed instances of OPMDs. The samples will be taken from the 2018 to 2020 time frame in the archives. At the beginning of the trial, the patients' current state of survival will be recorded. Each patient will have an additional 3-year follow-up.

Parameters That Will Be Studied

Both histopathological and clinical outcomes will be evaluated. They will be classified into 2 major groups: OPMDs and OSCC surgical margin cases.

Histopathological Parameters

For OSCC tumor-free surgical margins, the following parameters will be evaluated:

- Margin status (distance from tumor in millimeters; clear vs close vs involved)
- Presence of dysplasia (mild, moderate, or severe)
- Lymphovascular invasion
- Perineural invasion
- Pattern of invasion (cohesive or noncohesive)
- Histological grade of the primary tumor (well, moderately, or poorly differentiated)
- Depth of invasion
- Extranodal extension if lymph node metastasis is present

For OPMDs, the following parameters will be evaluated:

- Type of lesion (eg, leukoplakia, oral submucous fibrosis, erythroplakia, or lichen planus)
- Degree of epithelial dysplasia (World Health Organization grading: none, mild, moderate, severe, or carcinoma in situ)
- Architectural and cytological alterations

The assessment of both OS and DFS or malignant transformation–free survival (for OPMD) will be conducted based on the data available.

Clinical Parameters

The following clinical parameters will be evaluated:

- Age and gender
- Risk habits (tobacco, betel quid, or alcohol use)
- Site of lesion (eg, tongue, buccal mucosa, or floor of mouth)
- Date of diagnosis or surgery
- Date of last follow-up or death

Survival Analysis Plan

Overview

Kaplan-Meier survival curves will be used to visualize OS, DFS, or malignant transformation–free survival. The log-rank test will be used to compare survival between groups (eg, positive vs negative margins). The Cox proportional hazard model will be used to identify independent predictors of survival (eg, margin status, depth of invasion, and dysplasia grade).

IHC Analysis

The assessment of PD-L1 antigen will be performed via IHC. The Universal Immuno-enzyme Polymer method will be used for this assessment. After deparaffinization with xylene, tissue sections will be immersed into grades of alcohol for rehydration. The microwave heat antigen retrieval technique will be used.

XSL•FO
RenderX

Under this technique, sections will be heated for 10 minutes in 0.01 M sodium citrate buffer (pH 6.0). They will then be cooled at room temperature for approximately 20 minutes. This procedure will be carried out repeatedly until enough antigen is retrieved. After this, sections will be washed with the help of phosphate-buffered saline (PBS) and treated with 3% hydrogen peroxide in methanol for 30 minutes to block endogenous peroxidase activity. For nonspecific binding of the antigen, the section will be rinsed with PBS 3 times. The primary PD-L1 antibody (rabbit monoclonal anti–PD-L1 antibody [SP263 clone; Roche Diagnostics]) will then be applied to the tissue section and incubated for 90 minutes at room temperature in a humidifying chamber. Again, the section will be washed for 3 to 10 minutes with PBS and polymer antimouse antibody conjugated with horseradish peroxidase for 30 minutes. The antigen-antibody reaction will be observed using 3,3'-diaminobenzidine chromogen in a buffer solution. Mayer's hematoxylin counterstaining will be used. PD-L1 immunoexpression will be assessed by 2 independent observers in a blinded manner. Interobserver variability will be evaluated using appropriate statistical methods such as the Cohen κ for categorical data or the intraclass correlation coefficient for continuous data.

PD-L1 expression will be assessed using the CPS under light microscopy. Two pathologists will conduct the evaluation blinded without access to the clinical details.

The CPS is calculated by dividing the number of PD-L1 immunoexpression-positive cells, including cells of the epithelium, macrophages, and lymphocytes, by the total number of viable neoplastic cells and then multiplying the result by 100 [13], as follows: CPS = (number of PD-L1–staining cells [tumor cells, lymphocytes, and macrophages]/total number of viable tumor cells) × 100.

The CPS was selected as it is the most clinically validated scoring algorithm for PD-L1 assessment across multiple tumor types. Variability in PD-L1 interpretation due to differences in antibody clones, platforms, and scoring systems has been widely documented, highlighting the need for standardized evaluation methods [16].

Statistical Plan

One-way ANOVA or Kruskal-Wallis tests will be used to compare the CPS across the 3 study groups, followed by post hoc pairwise comparisons where appropriate. To assess the correlation between CPS in surgical margins and 3-year survival, the Spearman rank correlation or logistic regression may be used. If survival data are available as time to event, Kaplan-Meier survival analysis with the log-rank test is recommended.

Ethical Considerations

Institutional Ethics Committee approval has been obtained from Datta Meghe Institute of Higher Education and Research (DMIHER(DU)/IEC/2025/42). Every participant in the study has given their informed consent. Participant confidentiality will be maintained by anonymizing data and removing identifiable information. Data will be securely stored and accessed only by the research team. No monetary or nonmonetary compensation will be provided.

Results

This manuscript describes the protocol for a retrospective cross-sectional study. Data collection and immunohistochemical analysis are scheduled to commence in June 2025 and continue until June 2026. Results will be reported in a subsequent publication upon completion of the analysis, expected by late 2026 or early 2027.

The planned analyses include descriptive statistics of baseline clinical and demographic characteristics; assessment of PD-L1 expression in OPMDs, tumor-free surgical margins in OSCC, and normal oral mucosa; comparative analyses of PD-L1 expression between study groups; and Kaplan-Meier survival analysis and Cox proportional hazard regression to evaluate the association between PD-L1 expression and 3-year OS and DFS. Institutional ethics committee approval was obtained in January 2025 (DMIHER(DU)/IEC/2025/42). No external funding was received. Archival retrieval, IHC staining, and data extraction are planned from June 2025 to June 2026. As this is a retrospective study using cases from 2018 to 2020, no prospective recruitment is required. Data analysis is scheduled for mid-2026, and results are expected to be published by late 2026 or early 2027.

Final results will be reported in accordance with the study objectives upon completion of the analysis.

Discussion

Expected Findings

OSCC remains a major global health burden, with over 377,000 new cases and 177,000 deaths annually, ranking among the leading causes of cancer-related mortality worldwide [2]. A significant proportion of OSCCs arise from OPMDs, highlighting the importance of early biomarkers that can predict malignant transformation and guide patient management [3]. Despite advancements in surgical and adjuvant therapies, disease recurrence and poor survival remain key challenges, particularly due to residual tumor cells in histologically tumor-free surgical margins [4,6].

PD-L1 has emerged as a promising immune biomarker across multiple malignancies owing to its central role in tumor immune evasion [8,9]. Its overexpression has been reported in OPMDs as well as in invasive OSCC, suggesting a potential role in both early carcinogenesis and disease progression [3,10,12]. Studies such as those by Dave et al [10] and Pachpande et al [3] demonstrate that PD-L1 immunopositivity increases progressively from OPMDs to frank carcinoma, reinforcing its utility as a screening marker. A recent systematic review further supports the prevalence of PD-L1 expression in precancerous lesions of the head and neck, although with significant heterogeneity in cutoff values and detection methods [12].

In the context of surgical margins, molecular and immunohistochemical analyses provide deeper insights than conventional histopathology alone [4,6]. While histologically

XSL•FO
RenderX

clear margins are the current standard of care, residual molecular alterations can still be predisposed to recurrence. Braakhuis et al [4] emphasized that molecular mapping of margins reveals "field cancerization" phenomena, and Ranka et al [6] specifically highlighted PD-L1 as one of the immunohistochemical markers with prognostic relevance. Thus, assessing PD-L1 in tumor-free surgical margins may help identify patients at higher risk of recurrence despite negative histopathology.

PD-L1 expression has also been linked to patient prognosis and survival outcomes in other cancers, including lung [17] and breast cancers [18], as well as in advanced solid tumors treated with immune checkpoint inhibitors [11]. In head and neck cancer, immunotherapy with anti–PD-1 and PD-L1 agents (eg, nivolumab and pembrolizumab) has shown promising activity, particularly in PD-L1–positive patients [13]. These findings underscore the translational value of PD-L1 as both a prognostic and predictive biomarker in OSCC [19].

PD-L1 expression has also been shown to correlate with key molecular and behavioral risk factors. Tojyo et al [20] and Ahmadi et al [21] reported significant associations between PD-L1 expression and p53 status, smoking habits, and sex, suggesting that PD-L1 upregulation may reflect underlying genomic instability and carcinogen exposure. These associations further support the biological plausibility of PD-L1 as a prognostic and risk stratification biomarker in OSCC [20,21].

The clinical relevance of PD-L1 expression is further supported by therapeutic response data. Tang et al [22] reported a favorable response to nivolumab combined with radiotherapy and nimotuzumab in a metastatic patient with OSCC with strong PD-L1 expression, highlighting the predictive potential of this biomarker. Additionally, activation of the PD-1 and PD-L1 pathway has been linked to cancer-related pain, underscoring its broader biological and clinical significance beyond immune evasion alone [22,23].

At the same time, interpretation of PD-L1 expression remains challenging due to considerable biological and technical variability. PD-1 expression is influenced by factors such as the tumor microenvironment, inflammatory infiltrate, assay platforms, antibody clones, and scoring methods. In HNSCC, wide variations in PD-L1 expression across tumor sites and disease stage have been reported, limiting its use as a stand-alone biomarker. Differences in expression among primary tumors, tumor-free surgical margins, and premalignant lesions further indicate that PD-L1 expression may change during the process of carcinogenesis. Evidence from other epithelial malignancies such as lung and breast cancers similarly supports the prognostic relevance of PD-L1 while emphasizing the persistence of challenges related to cutoff selection, interobserver variability, and the temporal instability of expression. Recognizing these nuances is critical when evaluating PD-L1 as a potential screening or prognostic biomarker in OPMD and OSCC.

Given this evidence, our study aims to comprehensively evaluate PD-L1 expression through the CPS in OPMDs and OSCC surgical margins and correlate its expression with 3-year survival outcomes. By integrating biomarker assessment with clinical follow-up, this study will contribute to bridging the gap between molecular diagnostics and prognostic modeling in OSCC.

Study Limitations

This study has certain limitations inherent to its retrospective design. Although the sample size is sufficient to address the primary objective, the relatively small number of cases per group may limit the detection of subtle associations and restrict the feasibility of detailed subgroup analyses. Additionally, survival analyses are dependent on the completeness and accuracy of archival follow-up records. These limitations will be carefully considered during the interpretation of the results and highlight the need for larger, prospective studies to validate the findings.

Acknowledgments

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this article.

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Conflicts of Interest

None declared.

References

1. Ghapanchi J, Ranjbar Z, Mokhtari MJ, et al. The LncRNA H19 rs217727 polymorphism is associated with oral squamous cell carcinoma susceptibility in Iranian population. Biomed Res Int 2020;2020:1634252. [doi: 10.1155/2020/1634252] [Medline: 32337223]
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021 May;71(3):209-249. [doi: 10.3322/caac.21660] [Medline: 33538338]

XSL•FO
RenderX

3. Pachpande PS, Mandale MS, Bhavthankar JD, Humbe JG, Zanwar P. Assessment of programmed cell death ligand-1 (PD-L1) expression in oral potentially malignant disorders and oral squamous cell carcinoma - an immunohistochemical study. IP Arch Cytol Histopathol Res 2023;8(3):180-188. [doi: 10.18231/j.achr.2023.040]
4. Braakhuis BJ, Bloemena E, Leemans CR, Brakenhoff RH. Molecular analysis of surgical margins in head and neck cancer: more than a marginal issue. Oral Oncol 2010 Jul;46(7):485-491. [doi: 10.1016/j.oraloncology.2010.01.019] [Medline: 20189442]
5. Satgunaseelan L, Gupta R, Madore J, et al. Programmed cell death-ligand 1 expression in oral squamous cell carcinoma is associated with an inflammatory phenotype. Pathology 2016 Oct;48(6):574-580. [doi: 10.1016/j.pathol.2016.07.003] [Medline: 27590194]
6. Ranka R, Chaudhary M, Sharma P, Gawande M, Hande A, Patil S. Role of immunohistochemical markers in surgical margins of patients with head and neck carcinoma – a systematic review. Int J Curr Res Rev 2020;12(6):09-17. [doi: 10.31782/IJCRR.2020.12062]
7. Mildner F, Sopper S, Amann A, et al. Systematic review: soluble immunological biomarkers in advanced non-small-cell lung cancer (NSCLC). Crit Rev Oncol Hematol 2020 Sep;153:102948. [doi: 10.1016/j.critrevonc.2020.102948] [Medline: 32645684]
8. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012 Mar 22;12(4):252-264. [doi: 10.1038/nrc3239] [Medline: 22437870]
9. He X, Xu C. Immune checkpoint signaling and cancer immunotherapy. Cell Res 2020 Aug;30:660-669. [doi: 10.1038/s41422-020-0343-4] [Medline: 32467592]
10. Dave K, Ali A, Magalhaes M. Increased expression of PD-1 and PD-L1 in oral lesions progressing to oral squamous cell carcinoma: a pilot study. Sci Rep 2020 Jun 16;10(1):9705. [doi: 10.1038/s41598-020-66257-6] [Medline: 32546692]
11. Grosso J, Horak CE, Inzunza D, et al. Association of tumor PD-L1 expression and immune biomarkers with clinical activity in patients (pts) with advanced solid tumors treated with nivolumab (anti-PD-1; BMS-936558; ONO-4538). J Clin Oncol 2013 May 20;31(15_suppl):3016. [doi: 10.1200/jco.2013.31.15_suppl.3016]
12. Girolami I, Pantanowitz L, Munari E, et al. Prevalence of PD-L1 expression in head and neck squamous precancerous lesions: a systematic review and meta-analysis. Head Neck 2020 Oct;42(10):3018-3030. [doi: 10.1002/hed.26339] [Medline: 32567746]
13. Chow LQ, Haddad R, Gupta S, et al. Antitumor activity of pembrolizumab in biomarker-unselected patients with recurrent and/or metastatic head and neck squamous cell carcinoma: results from the phase Ib KEYNOTE-012 expansion cohort. J Clin Oncol 2016 Nov 10;34(32):3838-3845. [doi: 10.1200/JCO.2016.68.1478] [Medline: 27646946]
14. Yoshida S, Nagatsuka H, Nakano K, et al. Significance of PD-L1 expression in tongue cancer development. Int J Med Sci 2018;15(14):1723-1730. [doi: 10.7150/ijms.27860] [Medline: 30588196]
15. Schneider S, Kadletz L, Wiebringhaus R, et al. PD-1 and PD-L1 expression in HNSCC primary cancer and related lymph node metastasis - impact on clinical outcome. Histopathology 2018 Oct;73(4):573-584. [doi: 10.1111/his.13646] [Medline: 29742291]
16. Udall M, Rizzo M, Kenny J, et al. PD-L1 diagnostic tests: a systematic literature review of scoring algorithms and test-validation metrics. Diagn Pathol 2018 Feb 9;13(1):12. [doi: 10.1186/s13000-018-0689-9] [Medline: 29426340]
17. Srivastava P, Anand N, Agarwal D, Upadhyay R, Shukla S, Sharma V. Evaluation of programmed death-ligand 1 (PD-L1) expression and its correlation with clinicopathological parameters in oral squamous cell carcinoma: a tertiary care center study. Cureus 2025 Jun;17(6):e85186. [doi: 10.7759/cureus.85186] [Medline: 40600093]
18. Roncella S, Ferro P, Bacigalupo B, et al. Human mammaglobin mRNA is a reliable molecular marker for detecting occult breast cancer cells in peripheral blood. J Exp Clin Cancer Res 2005 Jun;24(2):265-271. [Medline: 16110760]
19. Girolami I, Marletta S, Fiorentino V, et al. Effect of radio-chemotherapy on PD-L1 immunohistochemical expression in head and neck squamous cell carcinoma. J Pers Med 2023 Feb 18;13(2):363. [doi: 10.3390/jpm13020363] [Medline: 36836595]
20. Tojyo I, Shintani Y, Nakanishi T, et al. PD-L1 expression correlated with p53 expression in oral squamous cell carcinoma. Maxillofac Plast Reconstr Surg 2019 Dec;41(1):56. [doi: 10.1186/s40902-019-0239-8] [Medline: 31857991]
21. Ahmadi N, Gao K, Chia N, et al. Association of PD-L1 expression in oral squamous cell carcinoma with smoking, sex, and p53 expression. Oral Surg Oral Med Oral Pathol Oral Radiol 2019 Dec;128(6):631-638. [doi: 10.1016/j.oooo.2019.07.008] [Medline: 31495715]
22. Tang X, Chen S, Sui Q, et al. Response to nivolumab combining radiotherapy and nimotuzumab in metastatic oral squamous cell carcinoma patient with strong PD-L1 expression: a case report. Ann Transl Med 2020 Mar;8(6):402. [doi: 10.21037/atm.2020.02.96] [Medline: 32355846]
23. Zhang J, Zhang H, Luo Y. Association between activation of the programmed cell death-1 (PD-1)/programmed death-ligand 1 (PD-L1) pathway and pain in patients with cancer. Med Sci Monit 2019 Feb 16;25:1275-1282. [doi: 10.12659/MSM.912632] [Medline: 30771277]

XSL•FO
RenderX

Abbreviations

CPS: combined positive score
DFS: disease-free survival
HNSCC: head and neck squamous cell carcinoma
IHC: immunohistochemistry
OPMD: oral potentially malignant disorder
OS: overall survival
OSCC: oral squamous cell carcinoma
PBS: phosphate-buffered saline
PD-1: programmed cell death receptor-1
PD-L1: programmed cell death ligand-1
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by J Sarvestan; submitted 18.Aug.2025; peer-reviewed by A Eccher, V Fiorentino; revised version received 17.Jan.2026; accepted 19.Jan.2026; published 16.Apr.2026.

Please cite as:
Sonone A, Hande A, Patil S, Pakhale A, Sharma P
Assessment of Programmed Cell Death Ligand-1 Expression in Oral Potentially Malignant Disorders and Tumor-Free Surgical Margins of Oral Squamous Cell Carcinoma: Protocol for a Retrospective Cross-Sectional Study
JMIR Res Protoc 2026;15:e82568
URL: https://www.researchprotocols.org/2026/1/e82568
doi:10.2196/82568

© Archana Sonone, Alka Hande, Swati Patil, Aayushi Pakhale, Preethi Sharma. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Self-Management Support After Burns: Protocol for a Multicenter, Stepped-Wedge Hybrid Type II Effectiveness-Implementation Study

Sharon L Blok[1,2,3], MSc; Marianne K Nieuwenhuis[1,3,4], PhD; Sonja M H J Scholten-Jaegers[1], MD, PhD; Irma Visser[5]; Raaba S M Thambithurai[6,7], MSc; Monique A C M de Craen-de Moor[6], BSc, RN; Kiran C Baran[8], MD; Anuschka S Niemeijer[1,3], PhD; Eelke Bosma[1,9], MD, PhD; Corry K van der Sluis[2], Prof Dr, MD; Sven J G Geelen[1,3], PhD; National Burn Care, Education & Research Group, The Netherlands[10,12]

[1]Alliance of Dutch Burn Care (ADBC), Burn Centre, Martini Hospital, van Swietenlaan 1, Groningen, The Netherlands

[10]See Acknowledgments

[12]

[2]University of Groningen, University Medical Center Groningen, Department of Rehabilitation Medicine, Groningen, The Netherlands

[3]Research Group Healthy Ageing, Allied Health Care and Nursing, Hanze University of Applied Sciences, Groningen, The Netherlands

[4]University of Groningen, University Medical Center Groningen, Department for Human Movement Sciences, Groningen, The Netherlands

[5]Dutch Association of Burn Survivors, Beverwijk, The Netherlands

[6]Alliance of Dutch Burn Care (ADBC), Burn Centre, Maasstad Hospital, Rotterdam, The Netherlands

[7]Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands

[8]Alliance of Dutch Burn Care (ADBC), Burn Centre, Red Cross Hospital, Beverwijk, The Netherlands

[9]Martini Hospital, Department of Surgery, Groningen, Groningen, The Netherlands

Corresponding Author:
Sharon L Blok, MSc
Alliance of Dutch Burn Care (ADBC), Burn Centre, Martini Hospital, van Swietenlaan 1, Groningen, The Netherlands

Abstract

Background: After a burn injury, the survivors have to manage and integrate the physical, psychological, and social consequences of their injury into their daily lives, such as functional limitations, aesthetic complaints, and fatigue. How successful survivors of burn injuries are at this depends on their self-management skills. Health care professionals play an important role in supporting the self-management of survivors of burn injuries. Currently, there are no burn-specific self-management support interventions. Therefore, we developed a self-management support intervention for survivors of burn injuries, called BreeZe (Brandwonden en Zelfmanagement).

Objective: This study aimed to describe a study protocol to implement and evaluate the BreeZe intervention.

Methods: This multicenter study in the Netherlands is an implementation-effectiveness hybrid type 2 study, with a nonrandomized stepped-wedge design. Starting April 2024, 3 phases have been sequentially rolled out across the 3 specialized Dutch burn centers over a period of 20 weeks—the preimplementation phase (usual care), implementation phase, and postimplementation phase. To identify barriers and facilitators of implementation, the Consolidated Framework for Implementation Research (CFIR) will be used. For evaluation, the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) evaluation framework is used. The coprimary outcomes are (1) self-management skills and (2) the implementation outcomes are reach, adoption, implementation, and maintenance. Secondary effectiveness outcomes are self-regulation, participation, dependency, patient-centeredness for survivors of burn injuries, self-management support skills for health care professionals, and cost-effectiveness. Data collection for survivors of burn injuries occurs at 2 weeks, 6 months, and 12 months post discharge, using questionnaires. Data collection for health care professionals occurs before training and 3, 6, and 12 months post implementation, using questionnaires, video observations, and interviews. Data analysis will include both quantitative and qualitative methods for comprehensive evaluation.

Results: Participant recruitment ended on June 30, 2025. Follow-up data collection ends in July 2026.

Conclusions: This study will evaluate both the effectiveness and implementation of the BreeZe self-management support intervention for survivors of burn injuries using a hybrid effectiveness–implementation design. By applying the CFIR and RE-AIM frameworks within a stepped-wedge design embedded in routine burn aftercare, this study aims to generate robust and practice-relevant evidence on how self-management support can be effectively implemented in burn care. The findings are expected to inform both clinical practice and future implementation efforts in burn aftercare settings.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT06782126; https://clinicaltrials.gov/study/NCT06782126
International Registered Report Identifier (IRRID): DERR1-10.2196/86671

(JMIR Res Protoc 2026;15:e86671) doi:10.2196/86671

KEYWORDS

burns; burn survivor; self-management; self-management support; stepped-wedge study; RE-AIM; hybrid type II; implementation science; Reach, Effectiveness, Adoption, Implementation, and Maintenance

Introduction

Background and Rationale

After a burn injury, the survivors have to integrate numerous short- and long-term consequences of their injury in their daily lives. These consequences extend beyond the acute phase of care and may persist for many years after the initial injury, resulting in ongoing physical, social, and emotional challenges. For example, scar maturation of burn scars is a slow process with multiple phases that can take up to 2 years or longer [1]. Other common long-term consequences include chronic pain, pruritus, reduced physical functioning, and enduring psychosocial difficulties, such as emotional distress and challenges related to altered appearance and social participation. All these long-term consequences frequently require ongoing monitoring, adaptation, and behavioral management by the survivor of burn injuries. Due to the persistent nature of these consequences and the associated long-term care demands, burn injuries have increasingly been conceptualized as a chronic condition rather than a purely acute event [2]. In order to manage the long-term and often persistent consequences of a burn injury and integrate these into daily life, survivors of burn injuries need to actively engage in self-management [3-6].

Self-management is a comprehensive, multifaceted process, which Richard and Shea [7] defined as "the ability of the individual in conjunction with family, community, and health care professionals to manage symptoms, treatments, lifestyle changes, and psychosocial, cultural, and spiritual consequences of health conditions." This definition emphasizes that self-management is not restricted to the medical domain but encompasses all aspects of life. For survivors of burn injuries, self-management therefore involves managing a range of ongoing demands, including persistent physical symptoms (such as pain, itch, or fatigue), adherence to often long-term wound and scar care regimens, psychosocial adjustment to changes in appearance and functioning, and the integration of these demands into everyday routines and social participation [8]. Although the relative importance and manifestation of these demands differ between individuals, survivors of burn injuries are commonly required to simultaneously manage these consequences over an extended period of time. Given the multimodality of these consequences over a long period of time, self-managing the consequences of burns is a complex and challenging process for survivors of burn injuries [8-11].

Although self-management is primarily a personal process, survivors of burn injuries can be supported in this process by health care professionals. Self-management support can be defined as the systematic provision of education and supportive interventions to improve patients' skills and confidence in managing their health problems, including regular assessment of progress and problems, goal setting, and problem-solving support [12]. In burn aftercare, such support is particularly relevant because responsibility for managing ongoing physical, treatment-related, and psychosocial consequences shifts increasingly to the survivor of burn injuries after discharge, when professional contact becomes less frequent. Previous research has shown that interventions focused on self-management support may lead to improved patient self-management skills and related outcomes, such as enhanced quality of life [5,13]. These interventions are typically composed of several elements, such as education, goal-setting, and behavioral coaching [14,15]. Although many components of self-management support interventions are not disease-specific, tailoring interventions to the specific demands and context of the target population is considered a core principle of effective self-management support [16].

Based on our review of the current literature on burn aftercare and self-management support interventions, we identified no published interventions that were specifically developed for survivors of burn injuries during their aftercare trajectory (SL Blok, unpublished data). Given the diverse yet recurring self-management demands faced by survivors of burn injuries during aftercare, we developed a self-management support intervention for Dutch survivors of burn injuries called BreeZe (Brandwonden en Zelfmanagement) [17]. The overall goal of the BreeZe intervention is to enhance the self-management skills of survivors of burn injuries in order to integrate treatment and life goals and subsequently optimize the quality of life of survivors of burn injuries and health-related outcomes. The BreeZe intervention is based on the content of the ZENN (Zelfmanagement Na Niertransplantatie) intervention, an evidence-based self-management support intervention for kidney transplant recipients, which we adapted to the context of burn care [18]. The ZENN intervention was based on the theoretical framework of the self-regulation theory [19]. The main intervention strategies of ZENN, and therewith BreeZe, are based on evidence-based techniques, namely goal setting and pursuit, solution-focused brief therapy, and motivational interviewing [20-22]. In practice, this means that the intervention focuses on a positive approach in order to enhance the intrinsic motivation and self-efficacy of survivors of burn injuries to encourage sustainable behavior change regarding self-management in burns aftercare. To achieve this, we will implement the BreeZe intervention using various implementation strategies. These will include educating health care professionals on self-management support, training them in communication skills (eg, motivational interviewing), and

XSL•FO
RenderX

providing supportive materials, such as a decision aid tool (ie, self-management web) and a workbook tailored to survivors of burn injuries [17,18].

Objectives

In this paper, we provide insight into the study protocol that we aim to use to evaluate the BreeZe intervention in Dutch burn care. For more detailed information about the development of the BreeZe intervention, including preliminary feedback regarding acceptability and feasibility, we refer readers to a separate article by Geelen et al [17].

The coprimary objectives of the evaluation study are (1) the effectiveness of the BreeZe intervention in improving self-management skills in survivors of burn injuries, and (2) the effects of our implementation approach on the implementation outcomes reach, adoption, implementation, and maintenance. The secondary objectives will be to evaluate the effectiveness of the BreeZe intervention on more latent health outcomes (eg, return to work or school) and to assess the cost-effectiveness of the BreeZe intervention from a societal perspective.

Methods

Trial Design

This multicenter stepped-wedge hybrid type II effectiveness-implementation study is designed to test the effectiveness of the BreeZe intervention while also testing the effect of our implementation approach [23,24]. The study is structured within a superiority framework, meaning we will systematically compare outcomes between the control and intervention groups, as well as assess changes from baseline to follow-up.

All 3 specialized Dutch burn centers will participate in this study—Martini Hospital Groningen (center 1), Red Cross Hospital Beverwijk (center 2), and Maasstad Hospital Rotterdam (center 3).

This study has 3 phases and uses in all 3 phases the implementation science frameworks “Consolidated Framework for Implementation Research” (CFIR) and the “Reach, Effectiveness, Adoption, Implementation, and Maintenance” (RE-AIM) [25-27].

In the first phase, the preimplementation phase, we will use the CFIR to provide insight into the context-specific determinants for implementing the BreeZe intervention, thereby informing the development of a context-specific implementation approach. Meetings will be held with burn physicians, nurse practitioners, aftercare nurses, and burn center managers. The findings will build on the preliminary feedback regarding acceptability and feasibility received during the development of the BreeZe intervention [17]. In addition, during the preimplementation phase, we will recruit patients to form a control group to evaluate effectiveness. This preimplementation phase starts for all 3 specialized burn centers in April 2024. Although protocol registration was finalized after the start of data collection in the preimplementation phase, we confirm that no modifications to predefined outcomes or analyses were made based on data collected before registration.

The second phase, the implementation phase, introduces and implements the BreeZe intervention for which we will use a stepped-wedge approach, visualized in Figure 1. We will begin by introducing the BreeZe intervention using the context-specific implementation approach developed in the preimplementation phase at center 1, in June 2024. Next, we will introduce the BreeZe intervention in center 2 in September 2024 and in center 3 in November 2024, respectively. Based on our development of the BreeZe intervention, we recognize that this implementation phase will require, at a minimum, education and training of health care professionals on how to apply the BreeZe intervention in routine clinical care. We anticipate that this implementation phase, consisting of training and intervention initiation, will take approximately 1 week; the intervention will then be delivered continuously.

Figure 1. Overview of the stepped-wedge design (yellow represents preimplementation phase, blue represents implementation phase, green represents postimplementation phase, and red represents end of postimplementation phase). MH: Martini Hospital; MSH: Maasstad Hospital; RCH: Red Cross Hospital.

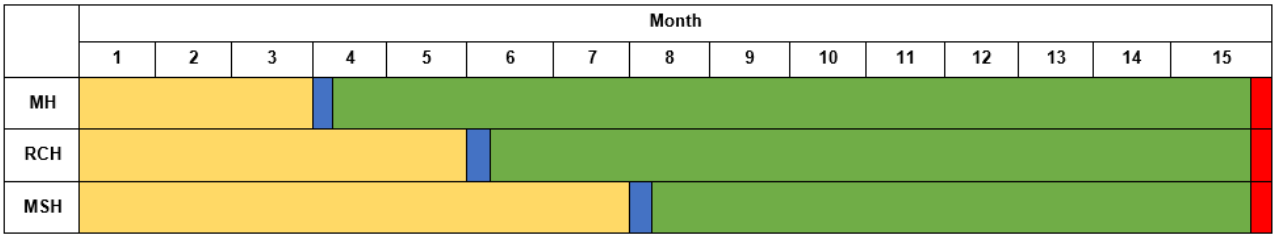

In the third and final phase, each burn center transitions to the postimplementation phase. Health care professionals are asked to apply the BreeZe intervention in routine clinical care. In addition, we will recruit patients during the postimplementation phase for the evaluation of the effectiveness, and we will collect data to assess the effect of our implementation approach using the RE-AIM. This will be our intervention group. Recruitment of patients ended on July 1, 2025.

Study Setting and Participants

Burn Centers

All 3 specialized burn centers in the Netherlands are characterized by their highly specialized multidisciplinary teams and are involved in the acute care of survivors of burn injuries, as well as in the aftercare. These centers, along with specialized trauma centers and other nonburn facilities, collectively constitute a well-organized network for burn care. Burn aftercare involves regular outpatient appointments, focused on wound

XSL•FO
RenderX

and scar management, minimizing complications, and long-term rehabilitation. The 3 centers are similar in most factors, with a few minor differences. There is a small difference in the geographical catchment area, with Groningen covering a larger but less densely populated region. All centers provide both inpatient and outpatient burn care, but there is a difference in the organizational structure of burn aftercare, particularly with respect to the professional roles involved and the level of autonomy in outpatient follow-up. Nevertheless, the characteristics of the patient populations between the centers are generally comparable, as the Netherlands is a small country.

Patients With Burns

The Emergency Management of Severe Burns referral criteria guide the decision-making process to refer patients to a specialized burn center [28]. Criteria for survivors of burn injuries to be treated in one of the specialized burn centers are (1) a total burned body surface area of >10%; (2) >5% full-thickness burns; (3) circular burns around the neck, thorax, or extremities; (4) burns around functional areas; (5) burns accompanied by other trauma; (6) chemical burns; (7) preexisting disease; or (8) burns due to electricity. Patients meeting the Emergency Management of Severe Burns criteria are either directly admitted to a burn center or subsequently transferred. In the Netherlands, around 1000 patients with burns are admitted to one of the 3 specialized burn centers per year (*Dutch Burn Repository R3*). Minor burn injuries can be managed in one of the burn centers by outpatient consultations, but can also be managed in specialized tertiary trauma centers or other nonburn facilities.

Survivors of burn injuries will be recruited continuously throughout each stepped-wedge study period. Only newly referred survivors of burn injuries entering burn aftercare at the time of study initiation may be approached for participation, provided they meet the eligibility criteria.

Health Care Professionals

The health care professionals involved in burn aftercare are surgeons, burn physicians, nurse practitioners, aftercare nurses, burn care nurses, physical therapists, occupational therapists, social workers, and psychologists.

Eligibility Criteria

Survivors of Burn Injuries

Eligible for inclusion are patients with burns who receive burn aftercare at 1 of the 3 specialized burn centers, who (1) are 18 years of age or older, (2) have been admitted for more than 24 hours or undergoing a wound-management procedure in theater, and (3) have proficiency in the Dutch language, as judged by the treating clinician. Exclusion criteria are cognitive limitations in such a way that self-management skills cannot be learned or executed, an acute psychiatric illness, or being discharged to a different health care institution. Cognitive limitations and acute psychiatric illness will be assessed by the treating clinician based on routine clinical judgment.

Health Care Professionals

Depending on the burn center and the complexity of the patient, a burn physician, nurse practitioner, or aftercare nurse is assigned the role of care coordinator in burn aftercare, putting them in the position to systematically provide self-management support. Therefore, all nurse practitioners, aftercare nurses, and burn physicians will be invited to participate and receive training.

Intervention

Explanation for the Choice of Comparators

The selection of comparators in this study is based on the need to comprehensively evaluate whether the BreeZe intervention is effective. Usual care serves as an appropriate comparator, as it reflects current burn aftercare practices, and it provides a baseline for assessing the added value of the BreeZe intervention.

Intervention Description

BreeZe—Overall Goal

The overall goal of BreeZe is to enhance the self-management and self-regulation skills of survivors of burn injuries, allowing them to integrate treatment and life goals and subsequently optimize their quality of life, participation in society, and health-related outcomes. The BreeZe intervention is based on the content of the ZENN intervention, an evidence-based self-management support intervention for recipients of a kidney transplant [18], which we adapted to the context of burn care [17]. The ZENN intervention was based on the theoretical framework of the self-regulation theory [19]. The main intervention strategies of ZENN, and therewith BreeZe, are based on evidence-based techniques, namely goal setting and pursuit, Solution-Focused Brief-Therapy, and motivational interviewing [20-22,29]. In practice, this means that the BreeZe intervention focuses on a positive approach in order to enhance the intrinsic motivation and self-efficacy of survivors of burn injuries to encourage sustainable behavior change regarding self-management.

BreeZe—Practical Application at Patient Level

The start of BreeZe depends on the health care trajectory of the survivor of burn injuries. For survivors of burn injuries who were admitted for a 1-day stay at a burn center for a debridement or skin graft surgery, BreeZe will commence ±7 to 14 days after discharge, and the steps of BreeZe will be divided into 4 sessions. For survivors of burn injuries who were admitted for more than 24 hours at a burn center, an extra session is added, which will take place around the day of discharge. It is important to note that the number of outpatient consultations needed to progress through these sessions depends on the needs for self-management support of the survivor of burn injuries and the logistical constraints of the burn center. Hence, a fixed intervention timeline cannot be specified. In case one or more informal caregivers are involved, they will also be offered the opportunity to be engaged from the start.

In the initial session, a holistic review of how the patient is doing in 14 life areas takes place, as outlined in the communication aid used, the self-management web. The self-management web is not a measurement tool, but is intended to structure the conversation, to open the view of both survivors of burn injuries and health care professionals to discuss possible

XSL•FO
RenderX

topics, and to make this visual for the patient. This web can also be revisited in subsequent sessions to make progress on personal goals, visual and tangible. This tool is also suitable for patients who have lower health literacy or language skills [18]. After completing the self-management web, the health care professional stimulates the survivor of burn injuries to prioritize a life area and to address and set a SMART (Specific, Measurable, Attainable, Relevant, and Time-based) goal. A global plan of action for goal attainment will be agreed upon. In addition, motivation for change and self-efficacy of survivors of burn injuries in relation to the goal will be discussed using visual analog scales. To support this process, a booklet has been developed that visually represents the overall BreeZe process and provides space to document all goals and progress. The second and third sessions are used to evaluate the progress of goal attainment. Facilitators and barriers are explored, and if necessary, strengths are emphasized, progress is complimented, and the action plan may be revised. In addition, motivation and self-efficacy will be evaluated and encouraged, alongside a discussion of internal versus external attribution of successes. This means that the survivor of burn injuries will be stimulated to look at his or her own role in the success, to promote internal attribution of that success. The aim of the fourth session is to evaluate and discuss the goal attainment, relapse prevention, and generalization of the learned skills to other situations.

For survivors of burn injuries who are discharged home after burn center admission and thereafter receive outpatient burn care, a session will be added to BreeZe to ensure they have a positive hospital-to-home experience. Adding this session to Breeze is pertinent to its success as survivors of burn injuries and their informal caregivers may feel overwhelmed, unsupported, and insecure during the hospital-to-home transition [30], which in turn strongly affects the ability to self-manage post discharge [18,31,32]. To counter possible negative perceptions, survivors of burn injuries and their informal caregivers have to feel empowered, encounter empathy and expertise, and have their expectations managed during the transition from hospital to home [18,30]. For these survivors of burn injuries, BreeZe will commence before discharge from the burn center. In this first session, the health care professional working in the burn aftercare visits the survivor of burn injuries in the burn center to discuss and prepare for the first 14 days post discharge. Instead of merely providing information to ensure understanding, motivation, and skills of survivors of burn injuries, these health care professionals will be trained to reach a common ground of understanding with the patient on what is needed during the first 14 days post discharge. Once this common ground has been reached, the health care professional uses solution-focused communication techniques to discuss the desired outcome at 14 days post discharge, discuss self-efficacy, and agree on a global plan of action for the first 14 days post discharge. A follow-up call ±3 days post discharge will allow the health care professional to explore the patient's actions, questions, and misunderstandings, including discrepancies in the discharge plan to be identified and addressed, as well as any concerns from (informal) caregivers.

The BreeZe sessions will primarily be delivered in-person during routine outpatient burn aftercare appointments at the burn center. Where clinically appropriate and in line with routine care practices, sessions may also be delivered by telephone. The intervention does not include fully digital delivery, but supporting materials (eg, workbook and decision aids) are provided in paper format to facilitate continuity between sessions. This flexible delivery mode is intended to support the feasibility and integration of the intervention within existing burn aftercare workflows.

For a detailed description of the BreeZe intervention components and the aims of each session, we refer to the intervention development study by Geelen et al [17].

Criteria for Discontinuing or Modifying Allocated Interventions

Participants can withdraw their permission to participate at any given time without providing a reason. Health care professionals will ensure that withdrawal of a participant does not influence their treatment quality.

Strategies to Improve Adherence to Interventions: The BreeZe Implementation Approach

We will use the CFIR to develop a context-specific implementation approach for the BreeZe intervention and guide its application in each burn center [25]. Our main implementation strategies include education and training of health care professionals, enabling them to deliver the BreeZe intervention. This education and training will feature a blended approach that incorporates an electronic learning (e-learning) followed by an in-person training session.

The e-learning is divided into 3 parts, blending theory with interactive questions and videos. Part 1 features 2 short webinars. The first covers the key components and theoretical basis of self-management support, while the second focuses on the solution-focused conversation method and the introduction of the self-management web tool. Part 2 explains the structure and phases of providing self-management support through the BreeZe intervention with example videos. Finally, part 3 presents a practical application through a challenging case study of a patient struggling with a chronic condition, such as burns.

The content of the in-person training session builds on the e-learning. The in-person training session focuses on the practical application of the BreeZe intervention in routine clinical care. Burn physicians, aftercare nurses, and nurse practitioners will learn to activate survivors of burn injuries to generate their own solutions rather than focusing on the problems. They will also learn how to support survivors of burn injuries in achieving progression toward their goals in a short space of time. The training will consist of 3 blocks of 2 hours given on the same day and will be provided by a psychologist (Emma Massey, PhD) together with a training actor. Dr Massey was involved in the development and implementation of the ZENN intervention on which BreeZe is based [18]. In addition, there will be an online booster session after 2 months with the aim of discussing what the health professionals have encountered in their practice and optimizing standardization between health care professionals.

XSL•FO
RenderX

In addition to education and training, we will hold frequent meetings in the preimplementation phase with burn physicians, nurse practitioners, aftercare nurses, and burn center managers to identify context-specific determinants for implementing the BreeZe intervention in each burn center. The CFIR-ERIC (expert recommendations for implementing change) Matching Tool will help us translate these determinants into implementation strategies supporting the education and training [25,33]. Together, they form the implementation approach, which will be put into practice during the implementation phase, that will improve adherence to the BreeZe intervention.

Relevant Concomitant Care Permitted or Prohibited During the Trial

During the preimplementation phase, health care professionals provide aftercare as usual. No other concomitant care or interventions are prohibited during the study.

Outcomes

Overview

To translate our intervention into practice, we adhere to the RE-AIM framework [26,27]. The RE-AIM framework consists of 5 dimensions—reach, effectiveness, adoption, implementation, and maintenance. We have presented each outcome measure under its RE-AIM dimension in Table 1. Detailed information on the definitions of the RE-AIM dimensions and their related measures can be found in Multimedia Appendix 1.

XSL•FO
RenderX

Table . Overview of the outcome measures categorized by each RE-AIM[a] dimension.

Outcome	Instrument	Number of questions	Number of domains	Answer scale	Minimum-maximum score
Coprimary outcome effectiveness measure					
Self-management skills	Partners in Health scale [34]	12	2	• 9-point Likert scale	0 - 96
Secondary outcome effectiveness measures – survivor of burn injuries level					
Self-regulation	Self-Regulation Assessment [35]	22	4	• 5-point Likert scale	0 - 110
Participation	Measure of Experiential Aspects of Participation [36]	4	4	• 5 answer options	__[b]
Daily activities	EuroQol-5D-5L [37]	1	—	• 5-point Likert scale	1 - 5
Self-care	EuroQol-5D-5L [37]	1	—	• 5-point Likert scale	1 - 5
Dependency	New formulated question	1	—	• 5-point Likert scale	1 - 5
Return to work or school	Adjusted International Consortium for Health Outcomes Measurement [38]	3	—	• 5 answer options	—
Patient centeredness	American Consumer Assessment of Health Plan Survey [39]	5	—	• 4-point Likert scale	0 - 20
Secondary outcome effectiveness measures – health care professional level					
Self-management support skills	Self-Efficacy and Performance in Self-management Support [40]	36	6	• 5-point Likert scale	0 - 180
Need-supportive counseling	Coding and Observing Need-Supportive Counseling in Chronic Care Encounters [41]	44	9	• 5-point Likert scale	0 - 24
	Nurse-Child Interaction Taxonomy [42]	16	3	• Percentages	0% - 100%
Secondary outcome effectiveness measures – setting level					
Cost-effectiveness	EuroQol-5D-5L [37]	6	5	• 5-point Likert scale • Visual Analog Scale	5 - 25 0 - 10
Coprimary outcome implementation approach measures					
Reach	Dutch Burn Repository R3	—	—	• Frequencies and percentages	—
Reach	Project coordinator records	—	—	—	—
Reach	Screening form	—	—	—	—
Reach	CFIR[c]-informed interviews	—	—	—	—
Adoption (staff level)	CFIR-informed interviews	—	—	—	—

XSL•FO
RenderX

Outcome	Instrument	Number of questions	Number of domains	Answer scale	Minimum-maximum score
Adoption (setting level)	Project coordinator records	—	—	—	—
Intervention fidelity	Electronic patient records	—	—	• Frequencies and percentages	—
Intervention fidelity	Project coordinator records	—	—	—	—
Intervention fidelity	CFIR-informed interviews	—	—	—	—
Implementation fidelity	Completion of implementation plan	—	—	• Frequencies and percentages	—
Implementation adaptations	Framework for Reporting Adaptations and Modifications-Enhanced [43]	—	—	—	—
Maintenance (staff level)	Provider Report of Sustainment Scale [44]	3	—	• Frequencies and percentages	—
Maintenance (staff level)	Self-Efficacy and Performance in Self-management Support [40]	36	6	• 5-point Likert scale	0 - 180
Maintenance (staff level)	CFIR-informed interviews	—	—	—	—
Maintenance (setting level)	Clinical Sustainability Assessment Tool [45]	35	7	• 7-point Likert scale	0 - 245
Maintenance (setting level)	CFIR-informed interviews	—	—	—	—

[a]RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance.

[b]Not applicable.

[c]CFIR: Consolidated Framework for Implementation Research.

RE-AIM: Reach

To determine the representativeness of survivors of burn injuries, participant demographics will be compared with those of nonparticipants using the Dutch Burn Repository R3. Nonparticipants will include both survivors of burn injuries who did not meet the eligibility criteria and eligible survivors of burn injuries who declined participation. We will compare some characteristics, such as sex, age, total body surface area burned, operation, and length of stay. Comparisons will be descriptive in nature and will be used to assess the representativeness of the study sample and to explore potential selection effects, rather than to draw causal inferences.

RE-AIM: Effectiveness

Effectiveness on Health Outcomes of Survivors of Burn Injuries – Coprimary Outcome Measure

Self-management skills: The primary outcome measure is self-management skills, which will be assessed with the revised version of the Partners in Health scale (PiH) [34]. The PiH measures self-management behavior. In the Dutch PiH, two subscales are identified, (1) knowledge and coping, and (2) recognition and management of symptoms and adherence to treatment [46]. Higher scores indicate better self-management skills. Psychometric properties of the PiH scale show good construct validity and reliability. Cronbach α to measure internal consistency varies from 0.82 to 0.86 [34,46,47]. The intraclass correlation coefficient (ICC) to measure test-retest reliability varies from 0.77 in the French version PiH to 0.818 in the Chinese version of the PiH, indicating a good test-retest reliability [48,49]. Although no data have been found on responsiveness, the PiH has shown to provide valid and consistent evidence of self-management skills of patients with chronic conditions [47].

Effectiveness on Health Outcomes of Survivors of Burn Injuries – Secondary Outcome Measures

Self-regulation: Currently available self-report measures of self-management among patients with chronic conditions do not address all elements incorporated into the intervention and are generic. Therefore, Mol et al [35] developed a new self-report instrument (Self-Regulation Assessment [SeRA]) based on a model of self-regulation, expert and patient opinions, and cognitive interviews. The SeRA assesses four domains of self-regulation: (1) insight into one's own health condition, (2) insight into own capabilities, (3) trust in self and apply self-regulation, and (4) organization of help. A higher score on the scale indicates a higher level of self-regulation skills in the

XSL•FO
RenderX

context of an organ transplantation. Mol et al [35] showed that the SeRA exhibits high internal consistency (Cronbach α=0.93) and that it is a valid and reliable instrument to assess the necessary self-regulation skills among rehabilitation patients.

Participation: To assess participation, we will use a shortened version of the Measure of Experiential Aspects of Participation [36]. Participants will be asked if they participate in four different life domains, (1) employment, (2) mobility, (3) sport, and (4) exercise. The response options will be used to assess both participation in the activity and the individual's satisfaction with the level of participation.

Daily activities: Participants will be asked if they can perform their daily activities. This question is part of the EQ-5D-5L questionnaire, which has been shown to have good known-group validity [37,50-52].

Self-care: In addition to daily activities, 1 item of the EQ-5D-5L also assesses self-care [37]. Participants will be asked if they are capable of performing self-care activities.

Dependency: Participants will be asked if they are dependent on other people in their daily lives. This is a newly formulated question.

Return to work or school: This will be assessed by the adjusted International Consortium for Health Outcomes Measurement return to work or school instrument at 6 and 12 months after discharge [38]. Participants will be asked about their daily occupation, with the following response options, (1) "I work 36 hours or more per week," (2) "I work 35 hours or less per week," (3) "I am job searching," (4) "I do not have a paid job (I go to school or am studying)," (5) "I do not have a paid job (retired, volunteer, stay-at-home occupation)," (6) "I cannot work because of my burn injury," and (7) "I cannot work due to other health issues." Furthermore, they will be asked if their job has been adjusted since their burn injury (eg, physical adjustments or reduced working hours), with "yes" or "no" as response options. Finally, they will be asked, if relevant, how long after their burn injury they have started working or studying again.

Patient-centeredness: The overall experience of care will be assessed using the Dutch translation of the subscale "patient-centeredness" of the American Consumer Assessment of Health Plan Survey, which consists of 5 questions [39,53]. This scale has been validated for use in the Dutch population and shows good internal consistency (Cronbach α=0.90) [38]. Participants will be asked if they feel listened to by their health care professional, if sufficient time is dedicated to their concerns, if they feel health care professionals take them seriously, if the health care professionals are able to appropriately explain things, and their level of trust in the health care professional's competence.

Health Care Professional Level Outcomes – Secondary Outcome Measures

Self-management support skills: The self-management support skills of the participating health care professionals will be assessed using the Self-Efficacy and Performance in Self-management Support (SEPSS) instrument [40]. The aim of the instrument is to assess self-efficacy and performance in self-management support. The six subscales of the SEPSS are based on the 5 A's model, (1) assess, (2) advise, (3) agree, (4) assist, (5) arrange, and (6) overall competencies. This questionnaire shows good internal consistency (Cronbach α=0.96) as well as a good test-retest ICC (ICC=0.95).

Need-supportive counseling: The self-management support skills of the participating health care professionals will be objectively evaluated in the preimplementation and postimplementation phases, via video observations using the Coding and Observing Need-Supportive Counseling in Chronic Care Encounters (COUNSEL-CCE) [41]. The survivors of burn injuries and health care professionals who will be filmed will be asked to give consent. Each health care professional will be filmed at least once during the preimplementation and postimplementation phases. The aim of the instrument is to observe need-supportive and need-thwarting counseling in chronic care encounters. The COUNSEL-CCE consists of 44 items divided into nine scales of behavioral approaches, namely (1) participative, (2) attuning, (3) guiding, (4) clarifying, (5) demanding, (6) domineering, (7) abandoning, (8) awaiting, and (9) relatedness supportive reciprocity. The procedure for the observation is the observation of the interaction between the professional and the patient during a 5-minute interval, coding of the observed interval, and observing the same interval for a second time to code any aspects missed. Before scoring the video recordings, the researchers will be trained in the use of the manual, which is provided by the developers of the instrument.

In addition to the COUNSEL-CCE, the video recordings will also be evaluated using the Nurse-Child Interaction Taxonomy (NCIT) [42]. Although this taxonomy was originally developed to improve nurse-child communication, it can also be used to systematically document the interaction between nurses and adult patients. The NCIT consists of 16 elements divided into three main patterns, (1) being considerate, (2) attuning oneself, and (3) procedural intervention.

Cost-Effectiveness

The costs will be calculated from a societal perspective. The total costs of burn care will be calculated separately for each survivor of burn injuries in the preimplementation and postimplementation phases. The included cost categories will be health care costs and patients' costs. Health care costs consist of burn center stay costs, treatment costs, clinical burn consultation costs, and outpatient burn care costs. Patient costs include travel costs, informal care costs, and productivity costs.

To evaluate the cost-effectiveness of the intervention, the complete EQ-5D-5L questionnaire will be used to compare the effectiveness in the control and intervention groups. The EQ-5D-5L questionnaire consists of 5 dimensions (mobility, self-care, daily activities, pain or discomfort, and anxiety or depression) and the EQ visual analog scale [37,50,54].

RE-AIM: Adoption

Adoption of the intervention will be measured at the staff and setting level. At the staff level, we will assess whether the intervention is adopted by conducting CFIR-informed interviews

XSL•FO
RenderX

with burn physicians, nurse practitioners, and aftercare nurses, by asking if they use the BreeZe intervention in their daily work. These interviews will provide working hypotheses to explain the degree of individual intervention adoption. At the setting level, we will assess whether the BreeZe intervention is used within an individual burn center.

RE-AIM: Implementation

The implementation of BreeZe will be evaluated by assessing the intervention's fidelity to the core components of BreeZe (eg, goal setting, action planning, and motivational interviewing), implementation fidelity to the implementation strategies (eg, completing the e-learning, attending the training, and being exposed to the additional ERIC implementation strategies), and the adaptations made to BreeZe or to the burn center (eg, workflows).

To gain insight into the intervention fidelity to the core components, health care professionals will be asked to document in patient records when they have used certain components. A convenience sample will be taken by assessing patient records to investigate how often the core components were used.

To gain insight into the implementation fidelity to the implementation strategies, each burn center will be asked to keep track of the completion of the implementation plan, including the staff attendance to the training plan and their e-learning completion rate.

To gain insight into the adaptations made to BreeZe or to the burn center (eg, workflows), adaptations will be documented using the Framework for Reporting Adaptations and Modifications-Enhanced (FRAME) during the postimplementation phase [43].

In addition to assessing the degree of adoption, the CFIR-informed interviews with burn physicians, nurse practitioners, aftercare nurses, and burn center managers will also be used to provide working hypotheses to explain implementation success or failure.

RE-AIM: Maintenance

Maintenance is assessed at the staff level and setting level [26,27]. At the staff level, maintenance will be assessed by using the health care professionals' self-report of continued use using the Provider Report of Sustainment Scale (PRESS) at 6 and 12 months post implementation [44]. The PRESS measure is a brief, reliable, and valid 3-item measure that is both pragmatic and usable across different evidence-based practices, provider types, and settings. The PRESS captures health care professionals' reports of continued use of an intervention in daily practice. In addition, health care professional level maintenance will also be assessed by using the health care professionals' self-efficacy and performance using the SEPSS-36 at 12 months post implementation [40].

At the setting level, maintenance will be assessed via the Clinical Sustainability Assessment Tool (CSAT) [45]. The CSAT is a brief and reliable instrument consisting of 35 items within 7 domains to assess an institution's capacity for maintaining a clinical practice. These domains are (1) engaged staff and leadership, (2) engaged stakeholders, (3) organizational readiness, (4) workflow integration, (5) implementation and training, (6) monitoring and evaluation, and (7) outcomes and effectiveness. CSAT development and testing demonstrated excellent internal consistency and several trends toward discriminant validity [45].

In addition to the questionnaires, the CFIR-informed interviews with burn physicians, nurse practitioners, aftercare nurses, and burn center managers will provide working hypotheses to explain the degree of individual (health care professional) and setting-level maintenance.

Participant Timeline

The participant timelines illustrated in Table 2 (for survivors of burn injuries) and Table 3 (for health care professionals) outline the different assessment time points of the study.

XSL•FO
RenderX

Table . Timeline for participating survivors of burn injuries: enrollment, interventions, and assessments.

	Study period – survivors of burn injuries				
	Enrollment	Allocation	Postallocation		Closeout
	Before discharge	At discharge	T1 (2 weeks postdischarge)	T2 (6 months postdischarge)	T3 (12 months postdischarge)
Enrollment					
Eligibility screen	✓				
Informed consent	✓				
Allocation		✓			
Interventions					
Control (usual care)[a]			✓	✓	✓
BreeZe aftercare[a]			✓	✓	✓
Assessments					
Demographics		✓			
Self-management skills (PiH[b]) [34]			✓	✓	✓
Self-regulation (SeRA[c]) [35]			✓	✓	✓
Participation (MeEAP[d]) [36]			✓	✓	
Daily activities (EQ-5D-5L) [37]			✓	✓	✓
Self-care (EQ-5D-5L) [37]			✓	✓	✓
Dependency			✓	✓	✓
Return to work or school (ICHOM[e]) [38]				✓	✓
Patient centeredness (CAHPS[f]) [39,53]				✓	
Quality of life (EQ-5D-5L) [37]			✓	✓	✓

[a]Duration of usual or BreeZe aftercare depends on the personal trajectory of the survivor of burn injuries.

[b]PiH: Partners in Health scale.

[c]SeRA: Self-Regulation Assessment.

[d]MeEAP: Measure of Experiential Aspects of Participation.

[e]ICHOM: Adjusted International Consortium for Health Outcomes Measurement.

[f]CAHPS: Consumer Assessment of Health Plan Survey.

XSL•FO
RenderX

Table . Timeline for participating health care professionals: enrollment, interventions, and assessments.

	Study period – health care professionals				
	Enrollment	Allocation	Postallocation		Closeout
	Preimplementation phase	Implementation phase	T1 (3 months postimplementation phase)	T2 (6 months post implementation phase)	T3 (12 months postimplementation phase)
Enrollment					
Eligibility screen	✓				
Informed consent	✓				
Allocation		✓			
Interventions					
Training		✓			
Assessments					
Demographics	✓				
Self-management support skills (SEPSS-36[a]) [40]	✓		✓	✓	✓
Need-supportive counseling (COUNSEL-CCE[b] and NCIT[c]) [41,42]	✓		✓		✓
Maintenance (PRESS[d]) [44]				✓	✓
Maintenance (CSAT[e]) [45]			✓		✓
CFIR[f]-informed interviews				✓	

[a]SEPSS-36: Self-Efficacy and Performance in Self-Management Support Questionnaire.

[b]COUNSEL-CCE: Coding and Observing Need-Supportive Counselling in Chronic Care Encounters.

[c]NCIT: Nurse-Child Interaction Taxonomy.

[d]PRESS: Provider Report of Sustainment Scale.

[e]CSAT: Clinical Sustainability Assessment Tool.

[f]CFIR: Consolidated Framework for Implementation Research.

Sample Size

To estimate sample size, we used a power calculation based on our primary outcome, which is self-management skills measured by the PiH questionnaire. As no minimum clinically important difference is established in the literature, 2 burn care nurse practitioners examined the PiH and considered an improvement of 8 points as clinically relevant. Based on this face-validity score and a previous dataset with PiH data from survivors of burn injuries collected with the Burn Centers Outcomes Registry the Netherlands (BORN), a mean difference of 8 points between the control and intervention group is feasible. As there is no validated stepped-wedge–specific sample size calculator for this outcome, we used a pragmatic approximation based on this predefined clinically relevant difference we established. In addition, as there is no data available on potential ICCs and we have a small number of burn centers, we will take center difference later on into account by adding 2 dummy variables (with center 1 being the reference center) in our analyses [55]. For the sample size calculation, we used the online sample size calculator of the University of British Columbia [56]. We aim to obtain 80% power with a 2-sided significance level of 0.05. Due to the stepped-wedge design of our study, our allocation rate for the control group:intervention group will be 0.625:1. With an adjustment of 20% for dropouts, missing data, and fixed time effects taken into account, we calculated that 57 participants in the control group and 89 participants in the intervention group will be sufficient. This amount (n=146) is deemed sufficient for (multilevel) regression analyses using the MIwiN statistical program, version 3.10 [57].

Recruitment

As of July 15, 2022, all survivors of burn injuries receiving burn care in one of the 3 centers are asked by a nurse practitioner or aftercare nurse to complete the routinely collected value-based health care burns core set for adult survivors of burn injuries as part of the Dutch Burn Repository R3 and the BORN [58]. In addition to the routinely collected core set data, survivors of burn injuries who participate during the pre- and postimplementation phases of this study will be asked to

XSL•FO
RenderX

complete additional questionnaires. A nurse practitioner or aftercare nurse will introduce the study to eligible patients and provide a patient information letter. After an appropriate consideration time (2 d), a member of the center's research team will answer any questions about the study and take informed consent. By involving the nurse practitioners and aftercare nurses, all eligible survivors of burn injuries will be reached.

Assignment of Interventions

Sequence Generation

We believe that a burn center's enthusiasm to participate in the intervention serves as a preliminary indicator of its organizational readiness. Organizational readiness is a crucial determinant that influences the implementation success or failure [59]. Hence, we decided to determine the order of commencement based on our interpretation of the burn centers' organizational readiness.

Patients will be recruited in consecutive order in the preimplementation phase as well as in the postimplementation phase. In principle, all health care professionals involved in burns aftercare will be recruited.

Blinding, Concealment Mechanism, and Implementation

Blinding is not feasible due to the nature of the intervention. As blinding is not feasible, no procedures for unblinding are necessary. Furthermore, due to the lack of blinding and randomization, no concealment mechanism is necessary. The sequence allocation of the centers has been chosen in deliberation with the BreeZe project team. In each center, 2 or 3 specialized researchers are responsible for enrollment and data collection.

Data Collection and Management

Data Collection Methods

Survivors of Burn Injuries

Assessment time points are at discharge, 2 weeks after discharge, and 6 and 12 months after discharge, as pictured in Table 2. Questionnaires at 2 weeks after discharge and 12 months are part of the routinely collected value-based health care burns core set, from the BORN [58]. Participants receive the questionnaires by email. In cases where patients do not have an email address or access to a computer, the questionnaire can be completed by telephone interview or on paper. Using the Dutch Burn Repository R3, the characteristics sex, age, and total burned body surface area will be collected from all eligible survivors of burn injuries during the study period.

Health Care Professionals

Assessment time points with health care professionals are preimplementation, and at 3, 6, and 12 months post implementation, as pictured in Table 3. They will receive the questionnaires via email or on paper, distributed by the responsible researcher.

Consultations that are eligible for video observations will be discussed and selected by the research team and health care professionals. Consultations will be eligible if they qualify as an aftercare consultation. Survivors of burn injuries will be informed and asked by a member of the research team 1 to 5 days before their appointment if they agree that their encounter is filmed. Before the start of the consultation, the researcher sets up the video camera and starts the recording. Then, the researcher leaves the room, and the health care professional and the survivor of burn injuries will start the consultation. Both the health care professional and survivor of burn injuries will be visible in the recording. At the end of the consultation and when the health care professional and survivor of burn injuries have left the room, the researcher will stop the recording and transfer the recording directly from the camera to the specialized research file. After the transfer, the recording will be directly deleted from the video camera.

The interviews will be carried out by the main researcher. Beforehand, a CFIR-informed topic guide will be developed. The interviews will be recorded with a tape recorder and will last approximately 45 minutes.

Plans to Promote Participant Retention and Complete Follow-Up

Survivors of Burn Injuries

To promote completion of the follow-up questionnaires, a researcher will remind participants a maximum of 2 times if they have not completed the questionnaires. These reminders can be an email or a phone call. We will adhere to an intention-to-treat approach. If treatment at the burn center is concluded before the survivor of burn injuries has been through all the BreeZe phases, the questionnaires will continue to be sent, except when a survivor of burn injuries withdraws their consent to participate.

Health Care Professionals

Health care professionals will be reminded by the responsible researchers to complete questionnaires, to select consultations for the video observations, and to set a date for the interviews.

Data Management

The responsible researcher of each center will save data in password-protected Microsoft Excel files, which are stored on a protected working drive. Passwords of the files are only known by the involved researchers. Data will be stored for 15 years according to the Dutch Personal Data Protection Law.

Survivors of Burn Injuries

Eligible survivors of burn injuries will be identified by the nurse practitioners or aftercare nurses, which will be communicated with the research team. For entry in the password-protected eligibility file, the characteristics documented are name, patient number, date of admission, date of surgery, date of discharge, and participation (yes or no, with reason if no).

Informed consent will be taken on paper and stored in a locked cupboard at the research office of each burn center. After the informed consent has been taken, each survivor of burn injuries will be assigned a unique study code which will be entered in a password-protected key data file in Excel containing key, name, and patient number.

XSL•FO
RenderX

Questionnaires will be sent out via Lighthouse QuestManager. Data will be extracted in an Excel file and then transported into IBM SPSS (version 25 or higher versions).

Health Care Professionals

Information on eligible health care professionals will be included in a password-protected eligibility file in Excel, which will be stored on a protected working drive. The password of the eligibility file is only known by the involved researchers. For entry in the eligibility file, the characteristics of eligible health care professionals that are documented are name, years of experience within the profession (divided into categories), and participation (yes or no).

Informed consent will be taken on paper and stored in a locked cupboard at the research office of each burn center. After taking informed consent, each health care professional will be assigned a unique study code in a password-protected key data file in Excel.

Questionnaires will be sent out on paper or via Castor [60]. Data from paper questionnaires will be imported into an Excel file and then transported into SPSS or MLwiN (Center for Multilevel Modelling, University of Bristol) [57]. Video recording files will be directly transferred from the video camera to the research working drive. Interview recordings will be directly transferred from the tape recorder to the research working drive, and then transported into Atlas.TI.

Statistical Methods

Statistical Methods for Primary and Secondary Outcomes

Overview

All quantitative analyses will be conducted using MLwiN 3.10 for nested data or IBM SPSS Statistics [57]. A P value <.05 will be considered statistically significant. For all outcomes, descriptive statistics will be used to present demographics and outcomes, using frequencies (%) for categorical variables, and median and IQR (if nonnormally distributed) or mean and SD (if normally distributed) for continuous variables.

RE-AIM: Reach

To examine if participating survivors of burn injuries are representative of the general survivor of burn injuries population, baseline demographics of participating survivors of burn injuries and nonparticipating survivors of burn injuries will be compared using chi-square tests for categorical data, and Mann-Whitney or independent *t* tests for numerical data.

RE-AIM: Effectiveness

Effectiveness on Health Outcomes of Survivors of Burn Injuries – Coprimary Outcome Measure

To examine if participants in the control and intervention groups are comparable on baseline demographics and clinical outcomes, chi-square tests will be used for categorical data, and Mann-Whitney or independent *t* tests will be used for numerical data. To test the effectiveness of BreeZe on health outcomes of survivors of burn injuries, we will conduct (multilevel) regression analyses for repeated measurements, with measurement occasions nested in patients. Variables that might explain differences between measurement occasions (eg, time since injury, time since discharge, season, and time since intervention started) or patient (eg, age and sex) will be tested. We expect that the variance between measurement occasions will be largely explained at the level of patients. To account for the difference between centers, we will add 2 dummy variables (with center 1 being the reference center) to these analyses [55].

In addition, as control and intervention participants are enrolled in different calendar periods in different centers due to the stepped-wedge design, we will test in the (multilevel) regression analyses whether time can best be modeled continuously (since discharge) or as a categorical variable. The latter is interpreted more easily and is more often used in clinical settings, but does increase the degrees of freedom in the model. We will also document major changes in burn care organization, staffing, or parallel initiatives during the study and consider these in sensitivity analyses and in the interpretation of our findings.

The primary data analyses will be conducted on the data of the PiH questionnaire. In the multilevel analyses, the patient's PiH score at a certain moment is the dependent variable. In addition, we will use three different approaches to provide a comprehensive evaluation of the intervention's impact. First, for each group, the number and percentage of participants who have improved by at least 8 points on their self-management skills will be calculated, and the difference between groups will be tested with a chi-square test. We consider an improvement of 8 points a clinically relevant score, based on a face validity assessment by burn care nurse practitioners.

Second, to examine whether greater improvement is achieved by individuals with low compared with high baseline scores, we will create 2 groups based on the median score. Besides a baseline effect and an intervention effect, an interaction effect will be tested.

Third, for each participant, we will tally the number of PiH questions where they score 6 or higher, as such a score will be considered clinically satisfactory. This sum score will range from 0 to 12 and will also be used as a dependent variable in a regression analysis with group and time as possible predictors. Our hypothesis is that there will be an increase in this sum score, which will be greater in the intervention group compared to the control group.

Effectiveness on Health Outcomes of Survivors of Burn Injuries – Secondary Outcome Measures

For the other outcome measures, we will perform similar multilevel analyses, testing the effects of baseline scores and intervention. In addition, we will calculate correlation coefficients to explore the association between the PiH and the SeRa. Correlation coefficients will be interpreted as negligible correlation (0 - 0.3), low correlation (0.3 - 0.5), moderate correlation (0.5 - 0.7), high correlation (0.7 - 0.9), or very high correlation (0.9 - 1.0) [61]. This method will allow us to explore the potential correlation between the constructs of self-management skills and self-regulation skills.

XSL•FO
RenderX

Health Care Professionals Level Outcomes – Secondary Outcome Measures

To evaluate the effect of the training on self-management support skills, we will use the data from the SEPSS-36 in a 2-level regression model with measurement occasions within a health care professional.

For the video observations, the video recordings will be broken down into 5-minute intervals. Each interval will be coded independently by 2 researchers according to the NCIT and COUNCEL-CCE coding manuals. Results will be presented with descriptive statistics. Differences in observations through time will be tested with (multilevel) regression analyses for repeated measures.

Setting Level Outcomes – Cost-Effectiveness

Burn care costs will be calculated by multiplying the volumes of health care used by the corresponding unit prices. The unit prices were derived from the Dutch guidelines [62] and previous studies [63-65]. The unit prices of previous studies will be updated by the inflation correction of the current financial year. Substantive additions, for instance, patients' costs will be based on previous studies and current care.

Utility values will be obtained from each patient with and without the BreeZe intervention by using the EQ-5D-5L questionnaire. The utility scores per domain range between 0 and 1, whereas 0 is death and 1 is perfect health. The total utility score per patient will be calculated. Quality-adjusted life years will be calculated by multiplying the length of time spent in a particular health state by its associated utility score.

RE-AIM: Adoption

At the staff level, adoption will be assessed with the results of the PRESS at 6 months post implementation, presented using descriptive statistics. At the setting level, adoption will be assessed by conducting CFIR-informed interviews with all participating health care professionals. These interviews will be transcribed verbatim and analyzed using a thematic content analysis using ATLAS.ti. Furthermore, 2 researchers will independently code each interview. Then, the researchers organize codes into categories and identify key themes that reflect adoption-related data.

RE-AIM: Implementation

To assess intervention fidelity, health care professionals are asked to document in patient records if, when, and which components of the BreeZe intervention they have used (eg, the Self-management web, goal setting, and action planning). This documentation will be used to assess how often certain components are used.

To assess implementation fidelity to the implementation strategies, the results of the implementation plan will be analyzed and presented with descriptive statistics (eg, completing the e-learning, attending the training, and being exposed to the additional ERIC implementation strategies).

Furthermore, to gain insight into the adaptations made to BreeZe or to the burn center (eg, workflows), we analyze the adaptations that have been documented during the study. These results will be schematically presented and summarized in text.

RE-AIM: Maintenance

At the individual health care professional level, maintenance will be assessed by using the results of the SEPSS-36 at 12 months post implementation (refer to Health Care Professional Level Outcomes – Secondary Outcome Measures). Furthermore, the PRESS outcomes at 12 months post implementation will be presented using descriptive statistics.

At the setting level, maintenance will be assessed by using the results of the CSAT questionnaire. Results of each center will be separately presented. First, descriptive statistics will be presented for both the total score as well as the scores of each subdomain. Second, an independent *t* test or a Mann-Whitney test will be used to assess the difference in results over time.

Interim Analyses

No interim analyses have been defined, due to the low anticipated incidence of adverse events in both the control and intervention group.

Methods for Additional Analyses

Our analyses, as described above, are based on an intention-to-treat approach. If survivors of burn injuries did start with the BreeZe intervention but did not complete it as intended, this will be documented. In case this occurs, we will perform additional analyses conforming to this protocol using only the subgroup of survivors of burn injuries who received and/or used at least 2 BreeZe components (eg, workbook and self-management web).

Methods in Analysis to Handle Protocol Nonadherence and Any Statistical Methods to Handle Missing Data

All analyses will be conducted based on the intention-to-treat approach. This means that all included survivors of burn injuries will be included in the analysis according to their assigned groups, regardless of their treatment adherence. This approach requires complete data for all participating survivors of burn injuries, and we will strive to minimize missing data throughout the study by keeping track of the completion of the questionnaires.

Missing data will be replaced with the value 999. To establish whether missing data is missing completely at random, missing at random, or missing not at random, we will use descriptive tables, frequencies, and correlation plots to identify potential patterns. If data are missing randomly, a maximum likelihood estimation model will be used to estimate parameters based on all available data. An advantage of performing multilevel analyses for nested data in patients is that these can be performed even if a whole measurement occasion is missing. However, then the missingness should be assumed to be random. If data are not missing randomly, we will conduct sensitivity analyses to examine how different assumptions about the missing data affect the results and evaluate the potential impact of the missing data on our conclusions. Finally, we will document the extent of missing data and the reasons for missingness.

XSL•FO
RenderX

Oversight and Monitoring

Composition of the Coordinating Center and Trial Steering Committee

This study will be coordinated by the research department of the burn center at the Martini Hospital. The primary research group will consist of the study principal investigator (MKN), a burn physician (SMHJSJ), a postdoctoral researcher (SJGG), and a PhD student (SLB). The latter will be responsible for running the overall study-related activities and the day-to-day organizational support at the Martini Hospital. At the other 2 burn centers, a local researcher will be the study coordinator and responsible for day-to-day organizational support and data management. These local study coordinators work in close collaboration with the coordinating group from the Martini Hospital, meeting at least once a month.

Data Monitoring Committee

This study does not establish a data and safety monitoring board, as the BreeZe intervention is regarded as low-risk and does not present any significant safety concerns.

Adverse Event Reporting and Harms

As the BreeZe intervention is regarded as low-risk, we do not anticipate any adverse events. Any unanticipated effects of the intervention will be documented and reported to the local Medical Ethics Committees of the Martini Hospital, Red Cross Hospital, and the Maasstad Hospital.

Trial Monitoring

This study will be supervised by the local Medical Ethics Committees of the Martini Hospital, Red Cross Hospital, and the Maasstad Hospital.

Patient and Public Involvement

The development of the BreeZe intervention was guided by a strong commitment to patient and public involvement. The research team included 2 patient experts as integral members of the team. One patient expert was engaged from the beginning of the project in 2022, and the second patient expert joined the research team in February 2023. They participated in monthly development sessions, which were facilitated by the lead researcher (SJGG) using a structured consensus-building approach as described by Leising and colleagues [66], emphasizing partnership-based collaboration. Within this process, the patient experts were positioned not merely as consultees but as cocreators, encouraged to actively contribute their experiential knowledge and perspectives.

Following the completion of the initial draft of the intervention, the research team, including the patient experts, continued to meet on a monthly basis. The focus of these sessions shifted toward implementation planning and the design of the evaluation study. The initial methodological framework was drafted by a postdoctoral researcher (SJGG) and a PhD candidate (SLB); however, systematic input and iterative feedback were sought from the patient experts to ensure that the design was responsive to patient needs and priorities. The selection of outcomes was primarily informed by the value-based health care burns core set, which itself was developed with substantive patient involvement to ensure the inclusion of outcomes and quality indicators most relevant to survivors of burn injuries [58].

Beyond development and evaluation, the role of the patient experts extends into dissemination. In collaboration with the Dutch Burn Survivors Association, they will contribute to knowledge translation activities, including advising on the accessibility and relevance of dissemination outputs and facilitating reach through their established networks.

Ethical Considerations

This study follows the procedures in accordance with the World Medical Association's Declaration of Helsinki and has been approved by the local Medical Ethical Committees of the Martini Hospital (reference 2024 - 033 and 2024 - 48), the Red Cross Hospital (references 24.010 and 24.018), and Maasstad Hospital (reference L2024046 and L2024086). All participating survivors of burn injuries and health care professionals will provide written informed consent. Participants are reassured that participation is voluntary and that withdrawal is possible at any time. No identifying information will be published. No compensation is available for participants.

Protocol Amendments

In the event of any protocol amendments, the local Medical Ethics Committee will be notified directly and asked for approval. After obtaining approval, the amendments will be communicated by the research team to all involved parties.

Consent

Survivors of Burn Injuries

As of July 15, 2022, survivors of burn injuries receiving burn care in 1 of 3 centers are asked by their health care professional to participate and complete the routinely collected value-based health care burns core set for adult survivors of burn injuries, as part of the Dutch Burn Repository R3 and the BORN [58]. In addition to these routinely collected data, a member of the research team will ask eligible survivors of burn injuries to complete additional questionnaires. This researcher will take informed consent if a survivor of burn injuries is willing to participate. Informed consent will be taken on paper and stored in a locked cupboard at the research office of each burn center. After the informed consent has been taken, each survivor of burn injuries will be assigned a unique study code, which will be entered in a password-protected key data file in Excel containing key, name, and patient number. Participants are reassured that participation is voluntary and that withdrawal is possible at any time.

Health Care Professionals

Informed consent from eligible health care professionals will be taken by a member of the research team of each burn center. Informed consent will be taken on paper and stored in a locked cupboard at the research office of each burn center. After taking informed consent, each health care professional will be assigned a unique study code in a password-protected key data file in Excel. Participants are reassured that participation is voluntary and that withdrawal is possible at any time.

XSL•FO
RenderX

Additional Consent

Enrolled health care professionals will participate in video observations of aftercare consultations to collect outcome data. Survivors of burn injuries will be informed and asked beforehand by a member of the research team if they agree to be filmed. If they agree, these survivors of burn injuries will sign an additional informed consent. Participants are reassured that participation is voluntary and that withdrawal is possible at any time.

All consent forms used in this study are available on request to the corresponding author. Every author has approved the final manuscript in its submitted form and is willing to take responsibility for all aspects of the research.

Confidentiality

Survivors of Burn Injuries

Eligible survivors of burn injuries will be identified by their practicing health care professional, which will be communicated with the main researchers. Then, researchers can enter patient records to identify demographic characteristics for the eligibility file. After enrolling in the study, participants are assigned an individual study code (key). For further data collection, only the study code will be assigned to the data, thereby maintaining confidentiality. Only the main researchers have access to the key document with participants' names. No identifying information will be published.

Health Care Professionals

Due to the small number of health care professionals in our study, we will adjust the data collection of the demographics. To maintain confidentiality, age and years of experience will be collected in categories, with 5-year steps. In addition, all health care professionals will be assigned a unique study code (key). For further data collection, only the study code will be assigned to the data. Only the main researchers have access to the key document with participants' names. No identifying information will be published.

Ancillary and Posttrial Care

The aim of the implementation of BreeZe during the implementation phase is to integrate BreeZe into daily practice. Necessary materials, such as laminated self-management webs and workbooks, are provided by the researchers to each burn center, which are allowed to keep after the implementation phase. No other forms of compensation are available for participants. Plans to ensure continuity of the BreeZe intervention are currently in progress.

Dissemination and Data Availability

The results of the study will be disseminated in scientific articles for publication in peer-reviewed journals. Both positive and negative results will be reported. Furthermore, we aim to present the results at national and international scientific conferences and other relevant meetings with stakeholders, including patient associations. Finally, the results will be disseminated via social and published media.

Results

This research is part of a larger project, which was funded in January 2022. Participant recruitment ended on June 30, 2025. In total, 42 participants are included in the control group, and 53 participants are included in the intervention group. Follow-up data collection is currently ongoing and will end in July 2026. Data analysis will start in August 2026, and the results are expected to be published in May 2027.

Discussion

Principal Findings

The aim of this study protocol was to describe the study that evaluates the effectiveness and implementation of the BreeZe intervention, a self-management support intervention that aims at improving self-management skills of survivors of burn injuries.

Methodological Considerations

One of the key strengths of this study protocol is the use of the well-established implementation science determinant framework, CFIR, in our implementation approach [25]. In addition to basing our intervention strategies on evidence-based techniques (eg, motivational interviewing), the implementation approach plays a crucial role in determining intervention success or failure [67]. The CFIR is a determinant framework designed to facilitate understanding and explanation of determinants (ie, facilitators and barriers) influencing implementation outcomes [68]. By prospectively identifying these determinants, we can use the appropriate implementation strategies to optimize the implementation process and ensure that all relevant factors, such as context, stakeholders, and resources, are taken into account [69,70].

Another key strength is the use of the implementation science evaluation framework, RE-AIM, in our outcome assessment. RE-AIM is an evaluation framework that helps to specify and assess aspects of implementation to determine overall success in terms of both intervention effectiveness and implementation [26,27]. Without specifying and assessing aspects of implementation, it is difficult to determine whether an intervention's success or failure is due to the intervention itself or how it was implemented [71].

The use of the RE-AIM framework also signifies that we are conducting a hybrid effectiveness-implementation study. Hybrid studies assess both the effectiveness of an intervention on health outcomes of a population and assess the effect of an implementation approach on implementation outcomes [23,24]. These studies are well-advocated as they can distinguish whether findings originate from the implementation process or the intervention itself, and they have the potential to accelerate the translation of research into clinical practice [23,72-74]. However, the combination of both effectiveness and implementation measures makes these hybrid studies highly extensive, which can pose challenges in execution due to resource constraints, logistical considerations, and time limitations [74,75]. Given the resources and time available, we have designed the most rigorous and robust study possible,

XSL•FO
RenderX

ensuring that we included sufficient outcomes and outcome measures.

A last methodological consideration in designing this study is our aim to strike a careful balance between internal and external validity, ensuring that the study is both generalizable to different burn care settings and sufficiently rigorous and robust to provide reliable answers to our research questions. While a parallel cluster randomized design is often considered the gold standard for minimizing bias and ensuring strong internal validity, we opted for a nonrandomized stepped-wedge design [24,76,77]. A parallel cluster randomized design, in which all clusters are randomly assigned to either the intervention or control condition, is often challenging in real-world settings, as it requires immediate and simultaneous implementation across all centers. Implementation success strongly influences the outcomes of an intervention, and whether you achieve implementation success is strongly dependent on organizational readiness [59]. Also, a parallel cluster randomized design is potentially less acceptable for stakeholders, as at least 1 cluster acts as the control group, meaning they are not exposed to the potential benefits of the intervention [24,78]. In contrast, a nonrandomized stepped-wedge design allows for a staggered implementation, where all centers eventually receive the intervention but at different time points and is logistically feasible [24]. Hence, to increase stakeholder acceptability and organizational readiness, we made the choice not to randomize. However, a nonrandomized stepped-wedge design carries a greater risk of bias, particularly due to lack of randomization, time-related confounders, and potential differences between early and late adopters [79,80]. To control for these, we took fixed-time effects into account in our sample size calculation [81].

Furthermore, while ensuring high internal validity is essential for accurately assessing the effectiveness of the intervention and effects of the implementation approach, it may limit the generalizability of the findings to real-world settings [82-84]. To enhance external validity, we decided to apply the same inclusion criteria for this study as those used for the patient-reported outcomes measures of the BORN, ensuring that the BreeZe intervention potentially reaches a large proportion of survivors of burn injuries in the Netherlands, and that the results are more generalizable to survivors of burn injuries populations in other countries. Finally, by integrating the BreeZe intervention into routine clinical practice, our study reflects real-world conditions, further enhancing its external validity.

Acknowledgments

We would like to thank ZonMW for funding this research. In addition, we would like to thank the Alliance of Dutch Burn Care, Red Cross Hospital Beverwijk, Maasstad Hospital Rotterdam, and Martini Hospital Groningen for their support. The authors declare the use of generative AI (ChatGPT-4) for refinement, correction, editing, or formatting the manuscript to improve clarity of language. Responsibility for the final manuscript lies entirely with the authors.

The members of the *National Burn Care, Education & Research Group, The Netherlands* are as follows:

Margriet E van Baar[6,11], Kiran C Baran[8], Sharon L Blok[1,2,3], Anna S van den Bosch[8,12,13], Eelke Bosma[1,9], Lotte van Dammen[1,5,9], Merit E van Eck[1], Sven J G Geelen[1,3], Rob van Gemert[5], Ymke Lucas[6], Esther Middelkoop[8,12,13], Marianne K Nieuwenhuis[1,3,4], Anouk Pijpe[8,12,13], Roos F C Salemans[6,14], Carine M H van Schie[15], Sonja M H J Scholten-Jaegers[1], Raaba S M Thambithurai[6,7], Denise van Uden[6], Irma Visser[5], Marjolein van der Vlegel[6,8], Gert Versluis[5,6], Robin A F Verwilligen[8,12,13], Cornelis H van der Vlies[6,14,16], Annebeth Meij-de Vries[8,17,18], Hendriet Wanders[5], Paul P M van Zuijlen[8,12,13,17]

[11]Erasmus MC, University Medical Centre Rotterdam, Department of Public Health, Rotterdam, the Netherlands

[12]Amsterdam UMC location VUmc, Department of Plastic, Reconstructive and Hand Surgery, Amsterdam, the Netherlands

[13]Amsterdam Movement Sciences (AMS), Tissue Function and Regeneration, Amsterdam UMC, Amsterdam, the Netherlands

[14]Trauma Research Unit Department of Surgery, Erasmus MC, University Medical Centre Rotterdam, Rotterdam, the Netherlands

[15]Alliance of Dutch Burn Care (ADBC), Dutch Burns Foundation, Beverwijk, the Netherlands

[16]Department of Trauma and Burn Surgery, Maasstad Hospital, Rotterdam, the Netherlands

[17]Department of Surgery, Red Cross Hospital, Beverwijk, the Netherlands

[18]Amsterdam UMC location University of Amsterdam, Pediatric Surgical Centre, Emma Children's Hospital, Amsterdam, the Netherlands.

Funding

This work was supported by ZonMW: The Netherlands Organisation for Health Research and Development, as part of the project "Optimizing top specialized burn care in the Netherlands," grant 10070022010003. The study sponsors were not involved, nor

XSL•FO
RenderX

had any influence in the study design, methodology choices, writing of the manuscript, or in the decision to submit the manuscript for publication, nor do they have ultimate authority over any of these activities. Contact info ZonMW: Postbus 93 2452509 AE Den Haag+31 70 515 03 97set@zonmw.nl.

Data Availability

This study is registered at the ClinicalTrials register (approved on January 6, 2025), which is accessible by the public. The final data from the study will be available upon publication of the study results. Due to the small number of health care professionals in the dataset, individual-level data will not be made publicly available. However, summary statistics and aggregated findings will be provided upon request, where appropriate.

Authors' Contributions

Editing and revising the article: SLB, SJGG, MKN, SMHJSJ, IV, MACMdCdM, KCB, EB, RSMT, ASN, CKvdS
Study design: SLB, SJGG, MKN, SMHJSJ, ASN, CKvdS
Writing original draft: SLB, SJGG, ASN, RSMT

Conflicts of Interest

All authors declare no conflicts of interest.

Multimedia Appendix 1
Definitions of the Reach, Effectiveness, Adoption, Implementation, and Maintenance dimensions, its related research questions, outcome measures, and data sources.
[DOCX File, 23 KB - resprot_v15i1e86671_app1.docx]

Checklist 1
SPIRIT checklist.
[PDF File, 403 KB - resprot_v15i1e86671_app2.pdf]

References

1. Serghiou MA, Ott S, Cowan A, Kemp-Offenberg J, Suman OE. Burn rehabilitation along the continuum of care. In: Herndon DN, editor. Total Burn Care, 5th edition: Elsevier; 2018:476-508. [doi: 10.1016/B978-0-323-47661-4.00046-0]
2. Barrett LW, Fear VS, Waithman JC, Wood FM, Fear MW. Understanding acute burn injury as a chronic disease. Burns Trauma 2019;7:23. [doi: 10.1186/s41038-019-0163-2] [Medline: 31534977]
3. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [doi: 10.1001/jama.288.19.2469] [Medline: 12435261]
4. Bringsvor HB, Skaug K, Langeland E, et al. Symptom burden and self-management in persons with chronic obstructive pulmonary disease. Int J Chron Obstruct Pulmon Dis 2018;13:365-373. [doi: 10.2147/COPD.S151428] [Medline: 29416327]
5. Allegrante JP, Wells MT, Peterson JC. Interventions to support behavioral self-management of chronic diseases. Annu Rev Public Health 2019 Apr 1;40:127-146. [doi: 10.1146/annurev-publhealth-040218-044008] [Medline: 30601717]
6. Cheng C, Inder K, Chan SWC. Patients' experiences of coping with multiple chronic conditions: a meta-ethnography of qualitative work. Int J Ment Health Nurs 2019 Feb;28(1):54-70. [doi: 10.1111/inm.12544] [Medline: 30232822]
7. Richard AA, Shea K. Delineation of self-care and associated concepts. J Nurs Scholarsh 2011 Sep;43(3):255-264. [doi: 10.1111/j.1547-5069.2011.01404.x] [Medline: 21884371]
8. Litchfield I, Jones LL, Moiemen N, Andrews N, Greenfield S, Mathers J. The role of self-management in burns aftercare: a qualitative research study. Burns 2019 Jun;45(4):825-834. [doi: 10.1016/j.burns.2018.10.020] [Medline: 30545694]
9. Dukes K, Baldwin S, Hagedorn J, et al. "More than Scabs and Stitches": an interview study of burn survivors' perspectives on treatment and recovery. J Burn Care Res 2022 Jan 5;43(1):214-218. [doi: 10.1093/jbcr/irab062] [Medline: 33895838]
10. Williams NR, Davey M, Klock-Powell K. Rising from the ashes: stories of recovery, adaptation and resiliency in burn survivors. Soc Work Health Care 2003;36(4):53-77. [doi: 10.1300/J010v36n04_04] [Medline: 12836780]
11. Kim KJ, Boo S, Oh H. Burn survivors' experiences of the ongoing challenges after discharge in South Korea: a qualitative study. Adv Skin Wound Care 2021 May 1;34(5):1-6. [doi: 10.1097/01.ASW.0000734380.80661.cc] [Medline: 33797424]
12. Adams K, Greiner AC, Corrigan JM, editors. Committee on the crossing the quality chasm: next steps toward a new health care system. In: The 1st Annual Crossing the Quality Chasm Summit: A Focus on Communities: National Academies Press (US); 2004.
13. Kim S, Park M, Song R. Effects of self-management programs on behavioral modification among individuals with chronic disease: a systematic review and meta-analysis of randomized trials. PLoS ONE 2021;16(7):e0254995. [doi: 10.1371/journal.pone.0254995]
14. Ryan P, Sawin KJ. The Individual and Family Self-Management Theory: background and perspectives on context, process, and outcomes. Nurs Outlook 2009;57(4):217-225. [doi: 10.1016/j.outlook.2008.10.004] [Medline: 19631064]

XSL•FO
RenderX

15. Dineen-Griffin S, Garcia-Cardenas V, Williams K, Benrimoj SI. Helping patients help themselves: a systematic review of self-management support strategies in primary health care practice. PLoS ONE 2019;14(8):e0220116. [doi: 10.1371/journal.pone.0220116] [Medline: 31369582]
16. Taylor SJC, Pinnock H, Epiphaniou E, Pearce G, Parke HL, Schwappach A, et al. A Rapid Synthesis of the Evidence on Interventions Supporting Self-Management for People with Long-Term Conditions: PRISMS – Practical Systematic Review of Self-Management Support for Long-Term Conditions: NIHR Journals Library; 2014. [doi: 10.3310/hsdr02530]
17. Geelen SJG, Blok SL, Niemeijer AS, et al. Development of BreeZe: a self-management support intervention for burn survivors informed by evidence, theory, and stakeholder co-creation. BMC Health Serv Res 2025 Dec 26;26(1):157. [doi: 10.1186/s12913-025-13849-5] [Medline: 41454276]
18. Beck D, Been-Dahmen J, Peeters M, et al. A nurse-led self-management support intervention (ZENN) for kidney transplant recipients using intervention mapping: protocol for a mixed-methods feasibility study. JMIR Res Protoc 2019 Mar 1;8(3):e11856. [doi: 10.2196/11856] [Medline: 30821694]
19. Maes S, Karoly P. Self‐regulation assessment and intervention in physical health and illness: a review. Applied Psychology 2005 Apr;54(2):267-299 [FREE Full text] [doi: 10.1111/j.1464-0597.2005.00210.x]
20. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implementation Sci 2011 Dec;6(1):42. [doi: 10.1186/1748-5908-6-42]
21. Ratner H, George E, Iveson C. Solution Focused Brief Therapy: 100 Key Points and Techniques: Routledge/Taylor & Francis Group; 2012.
22. Zhang A, Franklin C, Currin-McCulloch J, Park S, Kim J. The effectiveness of strength-based, solution-focused brief therapy in medical settings: a systematic review and meta-analysis of randomized controlled trials. J Behav Med 2018 Apr;41(2):139-151. [doi: 10.1007/s10865-017-9888-1] [Medline: 28975531]
23. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226. [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
24. Wolfenden L, Foy R, Presseau J, et al. Designing and undertaking randomised implementation trials: guide for researchers. BMJ 2021 Jan 18;372:m3721. [doi: 10.1136/bmj.m3721] [Medline: 33461967]
25. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
26. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]
27. Glasgow RE, Harden SM, Gaglio B, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64. [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
28. The Education Committee of the Australian and New Zealand Burn Association. Emergency management of severe burns (EMSB) course manual. dutch version ed. : Dutch Burn Foundation; 2009.
29. Miller WR, Rollnick S. Motivational Interviewing: Preparing People to Change Addictive Behavior: The Guilford Press; 1991.
30. Major ME, van Nes F, Ramaekers S, Engelbert RHH, van der Schaaf M. Survivors of critical illness and their relatives. A qualitative study on hospital discharge experience. Ann Am Thorac Soc 2019 Nov;16(11):1405-1413. [doi: 10.1513/AnnalsATS.201902-156OC] [Medline: 31394924]
31. Flink M, Ekstedt M. Planning for the discharge, not for patient self-management at home - an observational and interview study of hospital discharge. Int J Integr Care 2017 Nov 13;17(6):1. [doi: 10.5334/ijic.3003] [Medline: 29588634]
32. Brandberg C, Ekstedt M, Flink M. Self-management challenges following hospital discharge for patients with multimorbidity: a longitudinal qualitative study of a motivational interviewing intervention. BMJ Open 2021 Jul 30;11(7):e046896. [doi: 10.1136/bmjopen-2020-046896] [Medline: 34330857]
33. Powell BJ, Waltz TJ, Chinman MJ, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implementation Sci 2015 Dec;10(1):21. [doi: 10.1186/s13012-015-0209-1]
34. W Battersby M, Ask A, M Reece M, J Markwick M, P Collins J. The Partners in Health scale: the development and psychometric properties of a generic assessment scale for chronic condition self-management. Aust J Prim Health 2003 Aug 1;9(3):41-52. [doi: 10.1071/PY03022]
35. Mol TI, van Bennekom CAM, Scholten EWM, et al. The Self-Regulation Assessment (SeRA) questionnaire: development and exploratory analyses of a new patient-reported outcome measure for rehabilitation. Disabil Rehabil 2023 Jun;45(12):2038-2045. [doi: 10.1080/09638288.2022.2080289] [Medline: 35672153]
36. Caron JG, Martin Ginis KA, Rocchi M, Sweet SN. Development of the measure of experiential aspects of participation for people with physical disabilities. Arch Phys Med Rehabil 2019 Jan;100(1):67-77. [doi: 10.1016/j.apmr.2018.08.183] [Medline: 30268805]
37. Group E. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9]

XSL•FO
RenderX

38. ICHOM. International consortium for health outcome measurement. 2022.
39. Delnoij DMJ, ten Asbroek G, Arah OA, et al. Made in the USA: the import of American Consumer Assessment of Health Plan Surveys (CAHPS) into the Dutch social insurance system. Eur J Public Health 2006 Dec;16(6):652-659. [doi: 10.1093/eurpub/ckl023] [Medline: 16524940]
40. Duprez V, Van Hooft SM, Dwarswaard J, van Staa A, Van Hecke A, Strating MMH. The development and psychometric validation of the self-efficacy and performance in self-management support (SEPSS) instrument. J Adv Nurs 2016 Jun;72(6):1381-1395. [doi: 10.1111/jan.12918] [Medline: 26913585]
41. Duprez V, Haerens L, Wuyts D, et al. Self-Determination Theory to observe healthcare professionals' counselling in chronic care encounters: development of the COUNSEL-CCE tool. Patient Educ Couns 2021 Jul;104(7):1773-1780. [doi: 10.1016/j.pec.2020.12.002] [Medline: 33342579]
42. van Ingen Schenau-Veldman ISA, Niemeijer AS, Zuiker JK, Scholten-Jaegers S, Lamberts KF, Nieuwenhuis MK. A taxonomy to assess the interaction between nurses and children: development and reliability. J Clin Nurs 2020 Jun;29(11-12):2004-2010. [doi: 10.1111/jocn.15147] [Medline: 31856418]
43. Wiltsey Stirman S, Baumann AA, Miller CJ. The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. Implement Sci 2019 Jun 6;14(1):58. [doi: 10.1186/s13012-019-0898-y] [Medline: 31171014]
44. Moullin JC, Sklar M, Ehrhart MG, Green A, Aarons GA. Provider Report of Sustainment Scale (PRESS): development and validation of a brief measure of inner context sustainment. Implementation Sci 2021 Dec;16(1):86. [doi: 10.1186/s13012-021-01152-w]
45. Malone S, Prewitt K, Hackett R, et al. The Clinical Sustainability Assessment Tool: measuring organizational capacity to promote sustainability in healthcare. Implement Sci Commun 2021 Jul 17;2(1):77. [doi: 10.1186/s43058-021-00181-2] [Medline: 34274004]
46. Lenferink A, Effing T, Harvey P, et al. Construct validity of the Dutch version of the 12-item Partners in Health Scale: measuring patient self-management behaviour and knowledge in patients with chronic obstructive pulmonary disease. PLoS ONE 2016;11(8):e0161595. [doi: 10.1371/journal.pone.0161595] [Medline: 27564410]
47. Petkov J, Harvey P, Battersby M. The internal consistency and construct validity of the partners in health scale: validation of a patient rated chronic condition self-management measure. Qual Life Res 2010 Sep;19(7):1079-1085. [doi: 10.1007/s11136-010-9661-1] [Medline: 20437206]
48. Hudon É, Chouinard MC, Krieg C, et al. The French adaptation and validation of the Partners in Health (PIH) scale among patients with chronic conditions seen in primary care. PLoS ONE 2019;14(10):e0224191. [doi: 10.1371/journal.pone.0224191] [Medline: 31644561]
49. Chiu TML, Tam KTW, Siu CF, Chau PWP, Battersby M. Validation study of a Chinese version of Partners in Health in Hong Kong (C-PIH HK). Qual Life Res 2017 Jan;26(1):199-203. [doi: 10.1007/s11136-016-1315-5] [Medline: 27216940]
50. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736. [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
51. Feng YS, Kohlmann T, Janssen MF, Buchholz I. Psychometric properties of the EQ-5D-5L: a systematic review of the literature. Qual Life Res 2021 Mar;30(3):647-673. [doi: 10.1007/s11136-020-02688-y] [Medline: 33284428]
52. Szende A, Janssen B, Cabases J, editors. The EQ-5D Health-Related Quality of Life Questionnaire: Concepts and Methods: Springer; 2014.
53. Arah OA, ten Asbroek AHA, Delnoij DMJ, et al. Psychometric properties of the Dutch version of the Hospital-level Consumer Assessment of Health Plans Survey instrument. Health Serv Res 2006 Feb;41(1):284-301. [doi: 10.1111/j.1475-6773.2005.00462.x] [Medline: 16430612]
54. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. Value Health 2016 Jun;19(4):343-352. [doi: 10.1016/j.jval.2016.01.003] [Medline: 27325326]
55. McNeish DM, Stapleton LM. The effect of small sample size on two-level model estimates: a review and illustration. Educ Psychol Rev 2016 Jun;28(2):295-314. [doi: 10.1007/s10648-014-9287-x]
56. Brant R. Sample size calculators for designing clinical research. University of British Columbia. URL: https://www.stat.ubc.ca/~rollin/stats/ssize/n1.html [accessed 2026-02-18]
57. Charlton C, Rasbash J, Browne WJ, Healy M, Cameron B. MLwiN Version 310: Centre for Multilevel Modelling, University of Bristol; 2024.
58. Spronk I, van Uden D, Lansdorp CA, et al. Development of a value-based healthcare burns core set for adult burn care. Burns 2024 Sep;50(7):1925-1934. [doi: 10.1016/j.burns.2024.03.038] [Medline: 38902131]
59. Weiner BJ. A theory of organizational readiness for change. Implement Sci 2009 Oct 19;4:67. [doi: 10.1186/1748-5908-4-67] [Medline: 19840381]
60. Castor. URL: https://www.castoredc.com [accessed 2026-02-12]
61. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. Malawi Med J 2012 Sep;24(3):69-71. [Medline: 23638278]
62. Hakkaart-van Roijen L, Peeters S, Kanters T. Kostenhandleiding voor economische evaluaties in de gezondheidszorg: methodologie en referentieprijzen. herziene versie 2024 [in dutch]. : The National Health Care Institute (Zorginstituut

XSL•FO
RenderX

Nederland); 2024 URL: https://www.zorginstituutnederland.nl/site/binaries/site-content/collections/documents/2024/01/16/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg/Verdiepingsmodule+Kostenhandleiding+%28versie+2024%29.pdf [accessed 2026-02-12]
63. Hop MJ, Polinder S, van der Vlies CH, Middelkoop E, van Baar ME. Costs of burn care: a systematic review. Wound Repair Regen 2014;22(4):436-450. [doi: 10.1111/wrr.12189] [Medline: 25041616]
64. Hop MJ, Wijnen BFM, Nieuwenhuis MK, et al. Economic burden of burn injuries in the Netherlands: a 3 months follow-up study. Injury 2016 Jan;47(1):203-210. [doi: 10.1016/j.injury.2015.09.009] [Medline: 26454627]
65. Goei H, Hop MJ, van der Vlies CH, et al. Return to work after specialised burn care: a two-year prospective follow-up study of the prevalence, predictors and related costs. Injury 2016 Sep;47(9):1975-1982. [doi: 10.1016/j.injury.2016.03.031] [Medline: 27085837]
66. Leising D, Liesefeld H, Buecker S, Glöckner A, Lortsch S. A tentative roadmap for consensus building processes. Personality Science 2024 Apr;5. [doi: 10.1177/27000710241298610]
67. Peters DH, Tran NT, editors. In: Adam T, editor. Implementation Research in Health: A Practical Guide: World Health Organization; 2013.
68. Nilsen P. Making sense of implementation theories, models and frameworks. Implement Sci 2015 Apr 21;10:53. [doi: 10.1186/s13012-015-0242-0] [Medline: 25895742]
69. Damschroder LJ. Clarity out of chaos: use of theory in implementation research. Psychiatry Res 2020 Jan;283:112461. [doi: 10.1016/j.psychres.2019.06.036] [Medline: 31257020]
70. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. Implement Sci 2013 Dec 1;8:139. [doi: 10.1186/1748-5908-8-139] [Medline: 24289295]
71. Proctor E, Silmere H, Raghavan R, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76. [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
72. Wolfenden L, Williams CM, Wiggers J, Nathan N, Yoong SL. Improving the translation of health promotion interventions using effectiveness-implementation hybrid designs in program evaluations. Health Promot J Austr 2016 Feb;27(3):204-207. [doi: 10.1071/HE16056] [Medline: 29241482]
73. Landes SJ, McBain SA, Curran GM. Reprint of: an introduction to effectiveness-implementation hybrid designs. Psychiatry Res 2020 Jan;283:112630. [doi: 10.1016/j.psychres.2019.112630] [Medline: 31722790]
74. Ullman AJ, Beidas RS, Bonafide CP. Methodological progress note: hybrid effectiveness-implementation clinical trials. J Hosp Med 2022 Nov;17(11):912-916. [doi: 10.1002/jhm.12936] [Medline: 35934981]
75. Curran GM, Landes SJ, McBain SA, et al. Reflections on 10 years of effectiveness-implementation hybrid studies. FrontHealth Serv 2022 Dec 8;2:1053496. [doi: 10.3389/frhs.2022.1053496]
76. Lane-Fall MB, Curran GM, Beidas RS. Scoping implementation science for the beginner: locating yourself on the “subway line” of translational research. BMC Med Res Methodol 2019 Jun 28;19(1):133. [doi: 10.1186/s12874-019-0783-z] [Medline: 31253099]
77. Wolfenden L, Reilly K, Kingsland M, et al. Identifying opportunities to develop the science of implementation for community-based non-communicable disease prevention: a review of implementation trials. Prev Med 2019 Jan;118:279-285. [doi: 10.1016/j.ypmed.2018.11.014] [Medline: 30468794]
78. Brown CA, Lilford RJ. The stepped wedge trial design: a systematic review. BMC Med Res Methodol 2006 Nov 8;6(1):54. [doi: 10.1186/1471-2288-6-54] [Medline: 17092344]
79. Hemming K, Taljaard M. Reflection on modern methods: when is a stepped-wedge cluster randomized trial a good study design choice? Int J Epidemiol 2020 Jun 1;49(3):1043-1052. [doi: 10.1093/ije/dyaa077] [Medline: 32386407]
80. Hughes JP, Granston TS, Heagerty PJ. Current issues in the design and analysis of stepped wedge trials. Contemp Clin Trials 2015 Nov;45(Pt A):55-60. [doi: 10.1016/j.cct.2015.07.006] [Medline: 26247569]
81. Li F, Hughes JP, Hemming K, Taljaard M, Melnick ER, Heagerty PJ. Mixed-effects models for the design and analysis of stepped wedge cluster randomized trials: an overview. Stat Methods Med Res 2021 Feb;30(2):612-639. [doi: 10.1177/0962280220932962] [Medline: 32631142]
82. Huebschmann AG, Leavitt IM, Glasgow RE. Making health research matter: a call to increase attention to external validity. Annu Rev Public Health 2019 Apr 1;40(1):45-63. [doi: 10.1146/annurev-publhealth-040218-043945] [Medline: 30664836]
83. Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. Eval Health Prof 2006 Mar;29(1):126-153. [doi: 10.1177/0163278705284445] [Medline: 16510882]
84. Godwin M, Ruhland L, Casson I, et al. Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. BMC Med Res Methodol 2003 Dec 22;3(1):28. [doi: 10.1186/1471-2288-3-28] [Medline: 14690550]

Abbreviations

BORN: Burn centers Outcome Registry Netherlands
BreeZe: Brandwonden en Zelfmanagement ("Burns and Self-management")

XSL•FO
RenderX

CFIR: Consolidated Framework for Implementation Research
COUNSEL-CCE: Coding and Observing Need-Supportive Counseling in Chronic Care Encounters
CSAT: Clinical Sustainability Assessment Tool
ERIC: expert recommendations for implementing change
FRAME: Framework for Reporting Adaptations and Modifications-Enhanced
ICC: intraclass correlation coefficient
NCIT: Nurse-Child Interaction Taxonomy
PiH: Partners in Health
PRESS: Provider Report of Sustainment Scale
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
SEPSS: Self-Efficacy and Performance in Self-management Support
SeRA: Self-Regulation Assessment
SMART: Specific, Measurable, Attainable, Relevant, and Time-based
ZENN: ZElfmanagement Na Niertransplantatie (‘Self-management after kidney transplantation’)

Edited by J Sarvestan; submitted 28.Oct.2025; peer-reviewed by T Folkertsma; revised version received 18.Dec.2025; accepted 19.Dec.2025; published 16.Mar.2026.

Please cite as:
Blok SL, Nieuwenhuis MK, Scholten-Jaegers SMHJ, Visser I, Thambithurai RSM, de Craen-de Moor MACM, Baran KC, Niemeijer AS, Bosma E, van der Sluis CK, Geelen SJG, National Burn Care, Education & Research Group, The Netherlands
Self-Management Support After Burns: Protocol for a Multicenter, Stepped-Wedge Hybrid Type II Effectiveness-Implementation Study
JMIR Res Protoc 2026;15:e86671
URL: https://www.researchprotocols.org/2026/1/e86671
doi:10.2196/86671

© Sharon L Blok, Marianne K Nieuwenhuis, Sonja M H J Scholten-Jaegers, Irma Visser, Raaba S M Thambiturai, Monique A C M de Craen-de Moor, Kiran C Baran, Anuschka S Niemeijer, Eelke Bosma, Corry K van der Sluis, Sven J G Geelen, National Burn Care, Education & Research Group, The Netherlands. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Molecular and Pharmacogenetic Marker Evaluation in Relation to the Toxicity and Clinical Response of Acute Lymphoblastic Leukemia Treatment in Indian Children (MPGx-INDALL): Protocol for a Prospective Observational Cohort Study

Swetambri Sharma[1*], MSc; Shuvadeep Ganguly[2*], DM; Kamali Murugadoss[2], MSc; Smita Kayal[2], DM; Swaminathan Keerthivasagam[2], DM; Jaikumar Ramamoorthy[3], DM; Archna Singh[4], PhD; Deepam Pushpam[1], DM; Jayanthi Mathaiyan[5], MD; Yvonne Gloor[6], PhD; Frederic Baleydier[6], MD, PhD; Marc Ansari[6*], MD; Sameer Bakhshi[1*], MD; Biswajit Dubashi[2*], DM; Chakradhara Rao S Uppugunduri[2,6*], PhD

[1]Department of Medical Oncology, All India Institute of Medical Sciences, New Delhi, India

[2]Department of Medical Oncology, Jawaharlal Institute of Postgraduate Medical Education and Research, Dhanvantri Nagar, Puducherry, India

[3]Department of Pediatrics, Jawaharlal Institute of Post Graduate Medical Education and Research, Puducherry, India

[4]Department of Biochemistry, All India Institute of Medical Sciences, New Delhi, India

[5]Department of Pharmacology, Jawaharlal Institute of Post Graduate Medical Education and Research, Puducherry, India

[6]Department of Pediatrics, Gynecology and Obstetrics, CANSEARCH Research Platform in Pediatric Oncology and Haematology, University of Geneva, Geneva, Switzerland

[*]these authors contributed equally

Corresponding Author:
Chakradhara Rao S Uppugunduri, PhD
Department of Medical Oncology, Jawaharlal Institute of Postgraduate Medical Education and Research, Dhanvantri Nagar, Puducherry, India

Abstract

Background: Understanding interindividual variability in treatment response and toxicity is essential for optimizing outcomes in pediatric acute lymphoblastic leukemia (ALL). Molecular and pharmacogenetic markers hold promise in predicting treatment efficacy and adverse effects, particularly in genetically diverse populations. This protocol outlines the methodology for a prospective, nonrandomized observational cohort designed to evaluate molecular and pharmacogenetic factors associated with treatment response and toxicity in Indian children diagnosed with ALL.

Objective: The primary objective is to identify genetic markers associated with treatment-related toxicity and therapeutic response. Secondary objectives include evaluating associations between the occurrence of early toxicities and quality of life during active ALL treatment, specific pharmacogenetic variants, and survival outcomes along with generating data to support the future implementation of personalized treatment strategies in Indian children with ALL.

Methods: In this prospective, observational cohort, 556 children (≤18 years of age) with newly diagnosed ALL treated under the Indian Childhood Collaborative Leukemia–Acute Lymphoblastic Leukemia 2014 (ICiCLe-ALL-14) protocol at two Indian centers will be enrolled, aiming for a minimum of 500 evaluable children. Eligible participants will be enrolled prior to the initiation of chemotherapy and followed longitudinally throughout treatment. Clinical and laboratory data (demographics, nutritional assessment, quality of life, comorbidities, treatment regimen, toxicity graded by Common Terminology Criteria for Adverse Events v5.0, remission status, and survival) will be collected at predefined intervals up to day 100 of the maintenance phase. Germline and somatic DNA will be sampled at diagnosis and remission. The first phase will use whole-exome sequencing to discover candidate variants by implementing a candidate gene prioritization strategy. The second phase will genotype the top candidates in the full cohort using array technology. Associations with early treatment–related toxicities, steroid response, and survival will be tested by multivariable regression and Cox models. A machine learning approach with pharmacogenetic predictors as classifiers will be implemented further with cross-validation and sensitivity analyses.

Results: Ethical committees approved the protocol version 1.0 in 2020: IEC-1167/06.11.2020 (All India Institute of Medical Sciences, New Delhi), JIP/IEC/2020/201 (Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry), and AO_2021-00048 (UNIGE, Geneva). Funding was received from Swiss National Science Foundation, Switzerland; Department of Biotechnology, India; and CANSEARCH Foundation, Switzerland. Recruitment began in December 2022 and is likely to conclude by 2027. A comprehensive analysis of the complete study cohort is anticipated to be completed by 2027.

XSL•FO
RenderX

Conclusions: The MPGx-INDALL (Molecular and Pharmacogenetic Marker Evaluation in Relation to the Toxicity and Clinical Response of Acute Lymphoblastic Leukemia Treatment in Indian Children) study will generate actionable insights for individualized ALL therapy in India via systematically evaluating germline and somatic markers in a large ethnically distinct cohort.

Trial Registration: ClinicalTrials.gov NCT05512169; https://clinicaltrials.gov/study/NCT05512169

International Registered Report Identifier (IRRID): DERR1-10.2196/79865

(JMIR Res Protoc 2026;15:e79865) doi:10.2196/79865

KEYWORDS

pharmacogenetics; germline variants; somatic variants; toxicity; quality of life; treatment-related toxicity; ICiCLe; acute lymphoblastic leukemia; pediatric ALL; sequencing

Introduction

Background

Acute lymphoblastic leukemia (ALL) is the most common childhood malignancy accounting for about 25% of all cancers in children aged 2 - 15 years [1]. Globally, the odds of surviving ALL have improved dramatically, with 5-year overall survival (OS) reaching up to 90% in high-income countries (HICs) [2,3]. The proportion of patient deaths due to treatment-related toxicity (TRT) during treatment remains significant, ranging from 2% to 24% [4]. There is clearly a survival gap between low-income countries and HICs. Over recent years, the treatment protocols in ALL treatment have been focused on minimizing TRT and long-term adverse effects while preserving favorable survival outcomes. The protocols developed by BFM (Berlin-Frankfurt-Münster), the Children's Oncology Group, United Kingdom ALL group, and the French protocol for the treatment of ALL in children and adolescents are widely used in HICs [5].

Most of the oncology centers in India recently implemented these protocols, or modified versions of them, for the treatment of childhood ALL, especially the more recently implemented Indian Childhood Collaborative Leukemia–Acute Lymphoblastic Leukemia 2014 (ICiCLe-ALL-14) protocol [6-11]. The ICiCLe-ALL-14 protocol study on the Indian pediatric population reported a 4-year OS rate at 74%, and the event-free survival (EFS) rate was 62% [12]. Risk stratification clearly marked differences in the outcomes as well, with 4-year EFS and OS at 76% (95% CI 72% - 79%) and 88% (95% CI 85% - 90%) in the standard risk group, 70% (95% CI 66% - 74%) and 80% (95% CI 77% - 83%) in the intermediate risk group, 61% (95% CI 51% - 64%) and 73% (95% CI 70% - 76%) in the high risk groups, and 69% (95% CI 62% - 75%) and 77% (95% CI 70% - 83%) in T-ALL cases [12]. The cumulative incidence of treatment deaths and relapses at centers varied from 2% (95% CI 1% - 5%) to 13% (95% CI 10% - 17%; $P \leq .001$) and 21% (95% CI 17%-26%) to 45% (95% CI 39% - 51%), respectively. About 39% of the relapsed cases were early relapses, and steroid poor response indicated worse EFS in univariate analyses having 53% of EFS in poor responders [12]. These results indicate addressing the issue of improving survival rates, but when it comes to the prevention of toxicities, we have no prospectively collected data from this geographical region. Further, delayed presentation at diagnosis, treatment abandonment [13-15], treatment-associated infections, adverse effects [16,17], higher steroid nonresponse rates (up to 18%-40%) [6,13], and higher incidence of relapse simultaneously account for the death of children with ALL undergoing treatment [14,18]. ALL is a genetically heterogeneous disease. Specific genetic variations have been identified as biomarkers to predict the toxicity and efficacy of drug therapies used to manage ALL [19], for example, mutations in the IL7R signaling components *JAK1* and *KRAS* with steroid resistance [20]. However, baseline prognostic stratification using genetic screening and risk-adapted protocols is yet to be uniformly accessible across low-resource settings and even in high-resource settings. The use of pharmacogenetic markers to guide the dosing of chemotherapeutic drugs is also nonexistent in low or low-middle income settings. Geographic and ethnic variations in childhood cancers [21] and variations in supportive care therapy further increase the need for population-specific, preemptive markers for optimized treatment management.

Indian children may harbor distinct genetic mutations [20] and risk profiles not well captured in Western studies. As Gogoi et al [12] emphasize, scaling up genetic profiling is both significant and transformative. Incorporating genetic testing into the management of childhood ALL—especially in India's diverse and resource-limited setting—might help in the reduction of treatment-related mortality. Building on these insights, we would like to use this opportunity to combine this current standardized treatment protocol and the unique ethnic population of India, which also has high rates of inbreeding practices and consanguinity in certain parts of the country [22]. This prospective observational cohort study aims to evaluate the association of genetic variants (both germline and somatic) with TRTs, steroid response, and clinical outcomes among patients with childhood ALL being treated with the ICiCLe protocol.

Objectives

The primary objectives of the study are as follows:

1. To study the associations of static germline genetic variants with early chemotherapy–related toxicities (treatment-related toxicities) in children with ALL undergoing the ICiCLe treatment protocol.
2. To investigate the somatic and germline genetic markers associated with the efficacy and toxicity of glucocorticoid response, respectively.
3. To biobank biological samples and clinical data for future association analyses to develop biomarkers predicting the efficacy and toxicity of the treatment protocol.

The secondary objectives of the study are as follows:

XSL•FO
RenderX

1. To study the impact of the occurrence of early toxicities on quality of life (QoL) during active ALL treatment (as estimated by PedsQL Cancer module tool).
2. To evaluate the association of genetic variants (somatic and germline) with OS, nonrelapse mortality, relapse-free survival (RFS), and EFS.

Methods

Study Design

The proposed study is a prospective observational multicentric cohort study, where patient recruitment and clinical follow-up are being done at Regional Cancer Centre in Jawaharlal Institute of Post-Graduate Medical Education and Research, Puducherry in south India and at Dr. B.R.A. Institute Rotary Cancer Hospital, All India Institute of Medical Sciences, New Delhi in north India. The study is being carried out in 2 phases. The first phase includes the identification of germline genetic variants associated with TRTs and somatic mutations with that of steroid response using whole-exome sequencing (WES; n=100) for selecting candidate gene variants. WES combined with array technologies can shed light on many important coding region and regulatory region variants of interest. Through this exercise, we may also be able to identify any novel variants at a reasonable frequency in this ethnically distinct population. Further, selected variants from the first phase analysis will be genotyped in the rest of the cohort (n=400). Final association analyses will be performed using the entire cohort. First phase analyses will be conducted at the CANSEARCH Research platform of pediatric oncology and hematology, University of Geneva, Switzerland. Second phase analyses will be carried out using array technology at the two centers in India in collaboration with the University of Geneva, Switzerland. The study flow, including various phases, is shown in Figure 1.

Figure 1. Study flow.

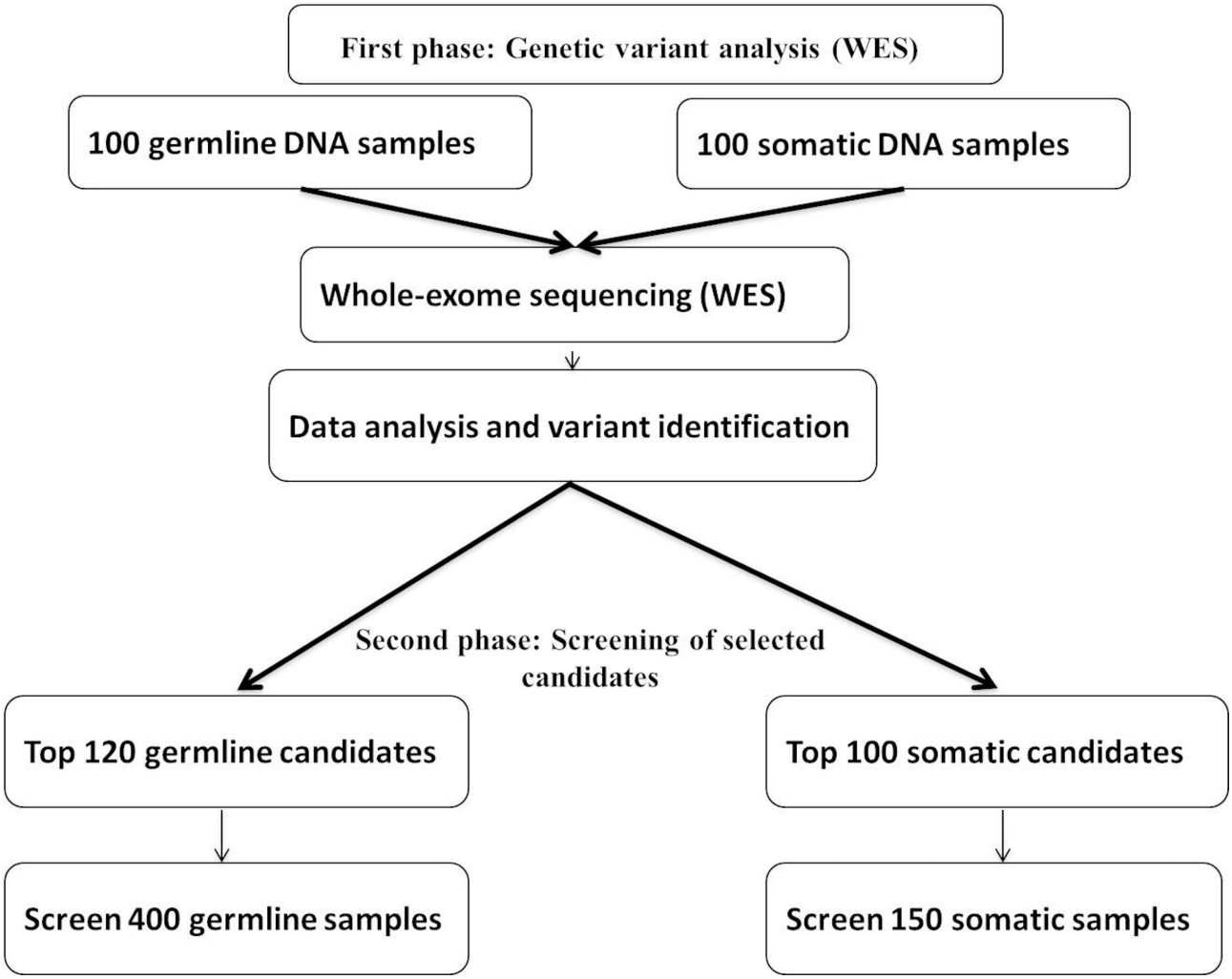

Patient Recruitment

Consecutive patients of pediatric ALL presenting to any of the treatment centers and being treated with the ICiCLe-ALL-14 protocol will be screened for the study.

The inclusion criteria of the study include patients (1) that are older than 1 year and younger than 18 years at enrollment; (2) that are previously untreated except for patients who received treatment up to 7 days of only steroid with documented baseline steroid dose details; (3) with an ALL diagnosis confirmed by morphology and immunophenotyping by flow cytometry; (4) with self-declared Indian origins; (5) who fulfill ICiCLe treatment protocol inclusion criteria and receive treatment as per the protocol; and (6) who provide written informed consent to participate in the study, signed by the participant, parent, or

XSL•FO
RenderX

guardian. The exclusion criteria of the study include (1) previously treated patients and (2) patients with Down syndrome.

Biological Sampling

The schedule of biological sampling planned for this study is mentioned in Multimedia Appendix 1. In brief, saliva or buccal swabs at the time of diagnosis or saliva or buccal swabs or blood at complete remission will be collected for germline DNA. For somatic DNA extraction, bone marrow samples will be collected at the time of diagnosis and at relapse (if any). In children with an inadequate amount of bone marrow, a peripheral blood sample will be considered for somatic DNA extraction, provided the percentage of blasts in peripheral blood is 80% or above. All saliva or buccal swabs and white blood cell pellets obtained from blood or bone marrow samples will be stored at 4 and −80 °C, respectively, until extraction. Plasma samples (4 mL whole-blood samples) will also be collected at the time of diagnosis (for drug level measurements), at remission, and at relapse (if any). Plasma separation will be performed immediately and stored at −80 °C. All biological sample aliquots remaining after addressing the study objectives will be stored in biobanks at respective centers for future investigations. Plasma or whole-blood samples will be collected at predefined time points. The details of the timelines for sample collection to estimate drug levels and biobanking are given in Multimedia Appendix 2.

DNA Preparation Protocol, Samples Storage, and Shipment

DNA extraction will be performed in batches using QIAamp DNA Mini Kit. DNA quality and quantity will be determined using Nanodrop 260/280 ratio and Qubit [1] (Qubit 4 Fluorometer by Thermo Fisher Scientific). Integrity will be measured using Tapestation, with genomic quality number values greater than 7.5 being used for phase 1 analyses. An aliquot of 1 µg DNA will be shared with the Geneva team for phase 1 sequencing analyses and storage in the BaHOP (Biobank of the Pediatric Hematology and Oncology Unit, IRB number: PB_2017 - 00533). The remaining aliquots will be stored in their respective biobanks in properly labeled vials at –80 °C temperature. DNA QC thresholds were as follows: minimum A260/280 ratio of 1.8 - 2.0, Qubit-measured concentration 20 ng/µL or higher, and Tapestation genomic quality number less than 7.5 for phase 1 sequencing. At least 1 µg of germline DNA and 1 µg of somatic DNA per sample are required for shipment, with 2 or more additional aliquots stored at –80 °C at the center's biobank for future investigations. DNA samples will be shipped on dry ice, with a maximum acceptable transit time of 48 hours and temperature logs maintained. For plasma sample separation, whole blood must be centrifuged within 30 minutes of collection, and separated plasma shall be aliquoted into prelabeled cryovials and immediately stored at −80 °C until shipment on dry ice. All the harmonized laboratory and sampling procedures have been finalized and are provided in the laboratory manual (Multimedia Appendix 2) to support consistent implementation across centers.

Clinical Data Collection, Patient Follow-Up, and Recording of Outcomes

The baseline sociodemographic details, including socioeconomic status by the modified Kuppuswamy socioeconomic scale and clinical and laboratory details (immunophenotype, karyotyping, and molecular tests) during disease presentation, will be collected [23]. All patients will undergo risk stratification and receive treatment as per ICiCLe-ALL-14 protocol and use uniform case record forms implemented in REDCap (Research Electronic Data Capture). The data will be anonymized by each center before the sharing of the clinical data and genetic data with other parties and in the public domain.

Patients will be followed up at each outpatient visit and any unscheduled visit due to toxicity. At each visit, all the available details of drug dose, modifications, or omissions due to toxicities will be recorded. This will aid in assessing the intensities of the treatment in relation to the incidence of toxicities and other clinical outcomes. All drug-related toxicities as per common terminology criteria for adverse events (version 5.0) [24] occurring from the first day of induction until day 100 of maintenance therapy will be recorded. For objective 1 (primary end point analyses), a predefined set of clinically significant early toxicities will be evaluated, including grade 3 or higher hematological, hepatotoxicity, pancreatitis, neurotoxicity, and serious infections. These events are explicitly mapped to the case record forms to ensure harmonized data capture across study sites. Post day 100 of maintenance, the patient will continue to be monitored for long-term outcomes including relapses (if any), need for hematopoietic stem cell transplantation, and mortality.

The data on steroid response at the end of 7 days (1 week of steroid prophase), bone marrow morphological remission status, and minimal residual disease using flow cytometry status at the end of induction or consolidation (if applicable) will be systematically recorded and analyzed in relation to somatic genetic variants.

Other clinical outcomes to be collected are nonrelapse mortality (time from study enrollment to death due to cause other than relapse of disease), incidence of relapse of disease (ie, the duration between the day of complete remission and the day of occurrence of relapse), RFS (time from study enrollment to relapse for patients attaining complete remission), OS (the duration from enrollment in the study to death due to any cause), and EFS (from treatment initiation to the first induction failure, nonresponse, or the progression of the disease, death from any cause).

Additionally, QoL will be assessed by the interview of children and their parents or caregivers using the age-specific PedsQL Cancer Module v3.0 questionnaire [25]. QoL will be assessed longitudinally initially at baseline, then at end of induction, consolidation, interim maintenance, and at day 100 of maintenance phases of the ICiCLe-ALL-14 protocol. We have obtained a license to use the questionnaires (work order 312005), and translated versions are validated as per the guidelines of the Mapi research trust and are made available at the Mapi research trust. Nutritional status will be evaluated by anthropometric assessment at baseline and serially along with

XSL•FO
RenderX

the recording of the implementation of any nutritional recommendations. The case record form template used for clinical data collection and XML files of the REDCap database template are available as Multimedia Appendices 3 and 4, respectively.

Schedule of Patient Participation in the Study

The patient follow-up data will be collected for the study objectives from the day of induction treatment until day 100 in the maintenance phase of the ICiCLe treatment protocol for ALL. Each patient will be scheduled for data collection and laboratory tests as a part of routine clinical care for 37 to 42 weeks from the day of the initiation of the treatment to evaluate the primary objectives and one of the secondary objectives related to QoL. However, the survival outcomes would require having follow-up data up to 1 year post treatment. The details of the planned procedures for patients and timelines are shown in Multimedia Appendix 2.

Detection and Analysis of Genetic Variants

First Phase Genetic Variant Analysis

Whole-Exome Sequencing

WES will be performed in 100 germline DNA samples and 100 DNA samples from leukemic cells (same patients). WES will use the TWIST Biosciences TWIST Comprehensive Exome workflow according to manufacturers' protocols: (1) library preparation; (2) sequencing using an Illumina NovaSeq6000 system with a mean coverage of 70× for germline DNA samples and 150× for somatic DNA samples; and (3) raw data integration, storage, and analyses. The sequencing will be done at the Health 2030 Genome Center, Campus Biotech, Geneva, Switzerland, and the sequences will be stored on a secured server of the University of Geneva until analyses and at the corresponding centers until the end of the study and archival.

Genetic Data Analysis—First Phase

For genetic data analysis, whole exomes and flanking regions (n=100, paired somatic and germline DNA) captured will be mapped to the GRCh38 reference and the variants called using the DRAGEN v.4.0.3 analysis platform. Plink 1.9 will be used to perform the quality control of the resulting sequences including call rate, Hardy-Weinberg equilibrium, sex mismatch, relatedness, heterogeneity, and ethnicity. Quality control investigations on cryptic relatedness and concordance between the reported and genetic sex will be included for patient selection. Principal component analyses will be an integral part of the analyses due to the possibility of unique clusters. Thus, the population structure is also accounted for while performing association analyses. When it comes to pharmacogenetic association (objective 1), frequencies might be impacted by the population structure, but for somatic genetic association, we expect no impact of population structures.

We will perform association analysis via (1) a candidate gene approach with filtered variants or mutations to the selected genes using candidate prioritization strategies from a systematic review (PROSPERO CRD420251112557 and CRD42021229748) and candidates associated with phenotypic measures used for toxicity definitions or steroid sensitivity from public datasets, for example, UK Biobank (application 91515), and (2) a hypothesis-free, exome-wide, and association analysis. The predicted effects of missense variants on protein function will be assessed in silico using SIFT and PolyPhen2 incorporated into variant effect predictor tools [26]. Variant filtering will be performed based on the 1000 Genomes, UK Biobank, IndiGen variation [22], and the NHLBI GO Exome Sequencing projects [27]. Fisher exact test (allelic association) and the Cochran-Armitage test for trends will be implemented in PLINK [28] to search for associations between clinical outcomes and genetic variants. Association analysis for quantitative and binary data will be analyzed using general linearized models in PLINK, with a *P* value significance threshold of .05 after adjustments for multiple testing that will be performed using the Benjamini-Hochberg false discovery rate method [29]. Candidate prioritization strategies for phase 1 analyses will be communicated in a separate report. Phase 1 analyses will thus enable the creation of a leukemia sequencing database of a single ethnicity, along with its clinical and follow-up data, all obtained from a single study. This will serve as a valuable resource for researchers investigating questions related to ALL treatment in children.

Second Phase Genotyping for Selected Candidates From Phase 1 Analysis

Somatic short variants will be called using GATK Mutect2 with matched tumor (bone marrow) and normal (buccal swab) samples. To control for germline mutations, we will provide a panel of normals generated from the buccal swab samples (GATK CreateSomaticPanelOfNormals) and use the gnomAD af-only resource (af-only-gnomad.vcf.gz; reference set of germline mutations) as our germline resource. Cross-sample contamination will be estimated (GATK GetPileupSummaries+CalculateContamination) and used during filtering. The resulting calls will be filtered and annotated for downstream analyses (GATK FilterMutectCalls) "purity thresholds." Tumor purity and copy number will be inferred using PureCN, as well as computing cancer cell fraction for each somatic variant. Samples with estimated tumor purity *p* less than 20% will be excluded from primary discovery analyses (but retained in QC tables). This threshold may be refined after inspecting the purity distribution. The rationale behind choosing this threshold is as follows: for a heterozygous clonal mutation (c=1) in a diploid region, the expected variant allele frequency (VAF) of a clonal mutation would be *p*/2 (eg, when *p*=1.0, VAF=0.50, when *p*=0.20, VAF=0.10). A mutation present in only 10% of the tumor cells (subclonal, cancer cell fraction=0.1), and tumor purity *p*=0.20, would correspond to a VAF of 0.01, which corresponds to 1.5 alternate reads at 150×. This is below a reliable detection level that can be confounded with sequencing noise. Additionally, to reduce false positives from sequencing noise, we will require a minimum of 5 alternate reads for somatic calls to pass filters for downstream analyses.

Phase 2 analyses will include screening 400 germline DNA samples for approximate top candidates identified in phase 1's germline sequencing association analysis with TRT. Somatic DNA analysis includes the screening of 150 samples from the top 100 candidates from phase 1 analysis. The final TRT association analyses will be conducted for 500 patients for the

XSL•FO
RenderX

top candidates and for steroid response (steroid or prednisone response rate on day 8; good responders: peripheral blood blast count <1000/μL; poor responders: peripheral blast count ≥1000/μL) [7] in 250 patients before meta-analysis with phase 1 participants. Phase 2 genotyping will be realized by genotyping arrays. We further plan to develop a project for cost-effective screening for implementation in a clinic, for example, custom open array.

Withdrawal and Discontinuation

The expected dropout rate is low for the entire period of the study. However, it is expected that approximately 5% of the recruited patients may discontinue or drop the study, and this will be managed by recruiting new patients into the study, if the data are not available for use to evaluate primary objectives.

Statistics and Analyses Methodology

SPSS (version 25; IBM Corp) or R statistical software (packages: survival, cmprsk, glmnet, and mbmdr [30]) will be used for statistical analysis. The appropriate genetic models (dominant, recessive, or additive) for specific genotypes or haplotypes will subsequently be derived based on the results of those analyses. Nonparametric or chi-square tests will be used to test differences in sex, age, and other categorical variables between genotype groups. The allele genotype frequencies, Hardy-Weinberg equilibriums, and haplotypes will be analyzed using Plink v1.9, the Haploview software (Broad Institute), and PHASE [31]. The frequency of all incidents and toxic events will be compared between genotype or haplotype or demographic- and disease characteristic–based groups using chi-square tests. The cumulative incidences of clinical outcomes, such as TRTs, will be estimated in relation to genotype or haplotype groups using cumulative incidence curves (cmprsk package in R). Survival curves for OS, nonrelapse mortality, RFS, and EFS will be calculated using Kaplan-Meier estimations, and log-rank tests will be used to compare the differences between genotype groups in univariate analysis. Univariate Cox regression analysis will be used to estimate hazard ratios with 95% CI; multivariate Cox regression analysis will then be used to estimate the impact of genotypes or haplotypes on clinical outcomes in the presence of other covariates. The influence of multiple genes on clinical parameters will be assessed in gene-gene interaction epistasis models incorporating various genetic and clinical confounders using multivariate regression analysis and multifactor-dimensionality reduction analysis [32]. All primary genetic association models will be adjusted for a predefined set of covariates but not limited to age, sex, immunophenotype, baseline white blood cell count, measurable residual disease, risk category, nutritional status category, socioeconomic status (modified Kuppuswamy), and treatment center.

We will use appropriate statistical tests for less frequent outcomes, for example, Least Absolute Shrinkage and Selection Operator–penalized regression [33], thus avoiding the overfitting of the data and biasing results due to the collinearity of the factors included. Confounders, such as socioeconomic and nutritional status, will also be included in the analyses. Reduced QoL due to the occurrence of toxicity will also be evaluated using logistic regression (after defining the impaired scores for both physical and emotional components). PedsQL scores are collected longitudinally across treatment phases; hence, we will attempt mixed-effects modeling to appropriately account for within-patient correlation and temporal trends. The objective of the final joint analyses is to determine the sensitivity and specificity of gene variants combined with other patient-specific characteristics to predict the clinical outcomes. This will be analyzed by decision curve analysis using all the categorical variables (some continuous variables would be converted into categorical variables using the X-tile program from Rimm lab, Yale School of Medicine), including genotypes along with reclassification improvement and integrated discrimination improvement in R using the “rms” package. The concordance index would be used to test a new model in comparison to the previous model. The performance of this strategy would be compared with that of receiver operating characteristic curve analysis. We will be applying suitable models only after the exploration of the data based on the collinearity matrix of the variables. However, we hereby propose possibilities for the analyses a priori. Appropriate statistical tools will be implemented upon discussions with peers to derive meaningful conclusions from the data collected in secondary analyses.

Additionally, for multiplicity control, the 2-stage discovery-validation design incorporates explicit control for multiple comparisons. In phase 1 (WES discovery), variant associations will undergo Benjamini-Hochberg false discovery rate (FDR) control (target FDR=0.05) with the reporting of both *q*-values and unadjusted *p*-values. Only the top ≈120 predefined candidates meeting prespecified criteria (biological relevance, effect size, and FDR ranking) will advance to phase 2 validation. In phase 2, confirmation analyses will apply a Bonferroni-adjusted significance threshold (α = .05/number of candidate variants). Joint (meta-analytic) models combining phase 1 and phase 2 will apply the same validation-level α threshold. This strategy balances discovery sensitivity with stringent clinical-grade validation.

Power Considerations for Primary Objectives

For germline pharmacogenetic variant associations, the sample size (n=500) was calculated using a power calculator for genome-wide association studies using a 2-stage design [34]. In this 2-stage design, 20% of the patients are genotyped for all the candidate markers, and then the remaining 80% are genotyped for the top 120 candidates (0.02% of markers analyzed in phase 1). This design has sufficient power for the phase 1 analyses: replicating and joint analyses. Criteria used for estimation are 5% to 15% minor allele frequency; dominant mode of inheritance; α value of .00001 (0.05/5000; 2-sided); power ($1 - \beta$) of 80%; early toxicities prevalence of 40%; and relative risk of 3.5 to estimate for the gene variant or variants. We used a stringent α value of .00001 (instead of .0004), that is, adequate for joint analysis (.05/120 = .0004) as required when developing models for clinical use. We acknowledge that this approach may lack smaller but clinically important associations in resource-limited settings. This design reduces the total amount of genotyping required by 80% (and thus associated costs), and there is a 90.1% probability that the associated markers will be followed in stage 2 analysis. Considering 5% of patients with loss to follow-up and an

XSL•FO
RenderX

additional 5% of samples being not analyzable due to laboratory issues, a total of 556 patients need to be recruited to get 500 evaluable patients for the primary objective 1. For secondary outcomes (OS, EFS, QoL, and steroid response), the power considerations are detailed in Multimedia Appendix 5.

Handling of Missing Data

This study expects to have minimal missing data, as a dry run for collecting data will be implemented, and documented issues will be fixed before initiating the recruitment of the patients. Moreover, the study includes participants who are under routine clinical care settings (in-patient settings and out-patient settings) for about 2.5 years, and our primary objective's data collection is limited to a maximum of 40 weeks; hence, we foresee dropouts (5%) from the study. Most importantly, the social worker coordinating the ALL treatment at both centers will ensure that patients receive this standard treatment. For secondary survival outcomes, we will collect the data retrospectively through forms that are currently filled out by the participating centers. The few dropouts that may occur shall be replaced by the recruitment of new participants. Patients whose germline samples could not be collected or analyzed will be excluded from the study and will not be analyzed. Furthermore, patients requesting the withdrawal of their informed consent will not be analyzed by the end of the study.

Ethical Considerations

The ethical approval to conduct the study was obtained from the Institute Ethics Committee, All India Institute of Medical Sciences, New Delhi, India (IEC-1167/06.11.2020); Jawaharlal Institute of Postgraduate Medical Education and Research, Puducherry (JIP/IEC/2020/201); and a positive stance from the ethics committee of Canton of Geneva (AO_2021 - 00048). All study procedures will be conducted in accordance with the principles of the Helsinki Declaration and in accordance with national and local regulations. All consecutive eligible patients will be recruited in the study after informed consent from a legal guardian and patient assent (if applicable). The participant or their legal representatives may withdraw consent to participate in the study at any time without justification, and without affecting the quality of routine standard medical care. For any withdrawal of consent, the data collected prior to that point will be used unless the participant or his or her legal representative does not wish to do so in writing.

Results

As this is a protocol, no clinical or genetic results are presented here. This section summarizes the current study status. Ethical approval has been obtained from all participating centers, and recruitment began in December 2022. A total of 346 participants were recruited for the study, and recruitment is ongoing. Phase 1 exome sequencing for both germline and somatic DNA has been completed. As per the most recent internal review, recruitment is projected to conclude by 2027. Data cleaning, harmonization, and preliminary genetic analyses for phase 1 are planned to begin in early 2026, with full cohort analyses expected in 2027.

Discussion

Principal Findings

To date, there have been no comprehensive studies for the treatment of childhood ALL across India's unique ethnic population in a standardized treatment setting with predefined methodology. We would like to use this opportunity to initiate this comprehensive project to serve as a benchmark pharmacogenetic study and to provide valuable resources (biological material and clinical data) for future investigations aimed at personalizing ALL treatment in Indian children. The novelty of the study lies in the fact that the 2-stage design with candidate prioritization strategies is implemented in a prospective cohort that exclusively covers the ICiCLe protocol that is uniformly used to treat patients with pediatric ALL across the Indian population. Moreover, harmonized treatment protocols in India have increasingly demonstrated benefits to survival progressively. This study design will enable us to identify both common and specific variants as markers of TRTs and clinical outcomes of ALL treatment in Indian children.

The study's results will be useful in the future to apply therapies to children and adolescents from India and elsewhere being treated for ALL. The concept of avoiding lost-to-follow-up patients by preventing the development of toxicities through personalized dosing is innovative. Thus, identifying such patients earlier using genetic markers would enable the physician to plan a proper management of their therapies well in advance, which is advantageous in limited-resource settings. Glucocorticoids are the principal components of pediatric ALL therapy protocols, including ICiCLe. It has been established through previous studies that there is interindividual variability in the treatment response to glucocorticoids as well as in susceptibilities to their toxicity [19,20]. Steroid nonresponders are usually at a higher risk of treatment failure, but response assessment is done after 8 days of treatment in the current treatment regimen. Therefore, with the early identification of responders, this study will help in reducing deaths related to steroid nonresponders. As the treatment protocol involves several chemotherapeutic drugs, there is an inadvertent need to identify genetic markers involved in drug metabolization, transport, and pharmacodynamics. So far, studies have been focused on selecting single genetic markers, but there has been emerging evidence suggesting associations between single-nucleotide polymorphisms and the development of toxicity or resistance or relapse to ALL therapies. Various single-nucleotide polymorphisms (target genes or metabolic enzymes) have been identified that are associated with increased susceptibility to drugs used in treatments, such as thiopurine methyltransferase, nudix hydrolase 15 for 6-mercaptopurine, methylenetetrahydrofolate reductase for methotrexate, and asparagine synthetase for L-asparaginase. This study will assess the relevance of these findings in addition to shedding light on novel variants, leading to a more comprehensive genetic profile for each patient that may aid in developing precision medicine strategies tailored specifically for the Indian population.

Therefore, it is our expectation that this study will provide direction on the utility of using genetic biomarkers for patient

XSL•FO
RenderX

stratification. If the proposed genetic stratification criteria could be clinically justified, it would be possible to involve preemptive genetic testing in routine clinical practice. As the study spans 2 centers and 2 laboratory environments, we prospectively established harmonized standard operating procedures and predetermined covariate sets. Batch-effect control measures, including technical replicates and platform-specific QC pipelines, have been incorporated to minimize systematic bias. The study will include whole-exome discovery with downstream genotyping; integration of clinical, QoL, and biological data; and biobanking for future work. This study design is not devoid of any limitations; for example, the design includes the evaluation of the most frequent variations and lacks power to quantify the effect of rare variants. The 2-phase design of the study optimizes the costs of genotyping but compromises study power to quantify the effect of low-frequency variations that may have impact together in a gene-gene interaction model. Moreover, the sample size of the study is calculated considering objective 1, that is, germline genetic predisposition. Within this study, we do not plan to collect late toxicities due to follow-up limited to day 100 of maintenance for primary objectives and reduced power for detecting modest genetic effects. However, patients will be followed up to obtain the retrospective data collection on toxicities and other clinical outcomes to assess long-term clinical outcomes in relation to germline and somatic genetic variants. Another major limitation includes the inclusion of patients from 2 geographical locations ie, North and South India, that are distinct in terms of ethnic origins. However, there is no clear distinction that was observed in terms of pharmacogene variant distribution between these 2 distinct groups [14]. This limitation, indeed, is an advantage to identify the variants that may have clinical utility for both populations, so that countrywide policies can be made efficient for translating and implementing the findings in clinics. Another limitation is the long recruitment time that may have an impact on the institute's policies and practices in treating the disease using the ICiCLe protocol. This can be resolved by considering treatment year as a confounder in the association analyses. This protocol study is reported in accordance with the SPIROS (Standardized Protocol Items Recommendations for Observational Studies) and STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for observational studies and the TRIPOD-AI (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis plus Artificial Intelligence) guideline for studies involving the development of machine learning–based prediction models.

Conclusion

To conclude, this study will prospectively evaluate the association of germline genetic variants with TRT and clinical outcomes in patients with pediatric ALL of Indian origin. The study will also explore somatic variant associations with clinical outcomes in the background of socioeconomic and nutritional status. In addition, the impact of toxicity occurrence on QoL will also be assessed as well, along with storing biological material and clinical data for future studies.

Acknowledgments

We acknowledge the support received from the clinical research center, University Hospitals of Geneva, especially for building the REDCap database. We also acknowledge Resonance INC (Dr Scott Howard) for monitoring the study during the initial phase. We thank Dr. Aurore Britan-Wood for critical feedback and support on the database and the management of the study. We thank Mme Zoe Giorgetta for her administrative support. Assistance by Mme Valerie Steck admin and Mme Thanh Mai is also acknowledged. We thank Florence Schmidt from the CANSEARCH foundation for their unending support. We thank the entire research team members (past and present) of Jawaharlal Institute of Postgraduate Medical Education and Research (Dr Charanraj Goud, Dr Yadav Nisha, Dr Govindan Krishnamoorthy, and Ms Sindhuja Jayamurthi) and team members from AIIMS (Ms Farhana Jahana, Ms Himanshi Sharma, Ms Mamta, Ms Muskan David, and Mr Chakshu Sharma) for their support during various stages of the study. We thank the NOCI India team (Dr Prasanth Ganesan) for their support in monitoring the study. We thank Dr Youssef Daali, Dr Caroline Samer, Prof Jules Desmeules, and the late Prof Hans Peter Wagner for supporting the idea of a pharmacogenomic study in acute lymphoblastic leukemia. Continued mentoring support from Prof Maja Krajinovic is highly appreciated.

Funding

The study is supported by the Swiss National Science Foundation (Switzerland; grant IZLIZ3_200243), the Department of Biotechnology (India; grant IC-12044(12)/3/2022-ICD-DBT), and the CANSEARCH Foundation (Switzerland).

Authors' Contributions

MA, SB, BD, and CRSU conceptualized the study. CRSU developed the methodological framework. SS, SG, MA, AS, SB, BD, and CRSU contributed to the design and optimization of experimental protocols. Original manuscript, tables, and figures were prepared by SS, SG, and CRSU. Reviewing and editing the manuscript were performed by SS, SG, KM, SB, BD, YG, and CRSU. Project supervision, recruitment, and administration are overseen by SG, S Kayal, S Keerthivasagam, AS, JM, JR, YG, FB, MA, SB, DP, BD, and CRSU. Funding was secured by SB, BD, and CRSU. All authors have read and approved the final manuscript. SG and SS contributed equally and held joint first authorship. CRSU, SB, BD, and MA held joint last authorship.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details of biological sampling.
[DOCX File, 18 KB - resprot_v15i1e79865_app1.docx]

Multimedia Appendix 2
Laboratory manual.
[PDF File, 2069 KB - resprot_v15i1e79865_app2.pdf]

Multimedia Appendix 3
MPX-IND ALL Case Record Form Version 1.3.
[PDF File, 831 KB - resprot_v15i1e79865_app3.pdf]

Multimedia Appendix 4
Electronic case record form template (REDCap).
[RAR File, 71 KB - resprot_v15i1e79865_app4.rar]

Multimedia Appendix 5
Objectives 2 and 3.
[DOCX File, 14 KB - resprot_v15i1e79865_app5.docx]

References

1. Ward E, DeSantis C, Robbins A, Kohler B, Jemal A. Childhood and adolescent cancer statistics, 2014. CA Cancer J Clin 2014;64(2):83-103. [doi: 10.3322/caac.21219] [Medline: 24488779]
2. Hunger SP, Lu X, Devidas M, et al. Improved survival for children and adolescents with acute lymphoblastic leukemia between 1990 and 2005: a report from the children's oncology group. J Clin Oncol 2012 May 10;30(14):1663-1669. [doi: 10.1200/JCO.2011.37.8018] [Medline: 22412151]
3. Gaynon PS. Childhood acute lymphoblastic leukaemia and relapse. Br J Haematol 2005 Dec;131(5):579-587. [doi: 10.1111/j.1365-2141.2005.05773.x] [Medline: 16351633]
4. Dharia NV, Kugener G, Guenther LM, et al. A first-generation pediatric cancer dependency map. Nat Genet 2021 Apr;53(4):529-538. [doi: 10.1038/s41588-021-00819-w] [Medline: 33753930]
5. Reismüller B, Steiner M, Pichler H, et al. High hyperdiploid acute lymphoblastic leukemia (ALL)—a 25-year population-based survey of the Austrian ALL-BFM (Berlin-Frankfurt-Münster) Study Group. Pediatr Blood Cancer 2017 Jun;64(6):e26327. [doi: 10.1002/pbc.26327] [Medline: 27804199]
6. Roy P, Narula G, Arora B, et al. Implementation of risk adapted therapeutic strategy for childhood acute lymphoblastic leukaemia—interim report of the pilot InPOG-ALL-15-01 study. Pediatric Hematology Oncology Journal 2018;3(3):S19. [doi: 10.1016/j.phoj.2018.11.052]
7. Das N, Banavali S, Bakhshi S, et al. Protocol for ICiCLe-ALL-14 (InPOG-ALL-15-01): a prospective, risk stratified, randomised, multicentre, open label, controlled therapeutic trial for newly diagnosed childhood acute lymphoblastic leukaemia in India. Trials 2022 Jan 31;23(1):102. [doi: 10.1186/s13063-022-06033-1] [Medline: 35101099]
8. Pui CH, Yang JJ, Hunger SP, et al. Childhood acute lymphoblastic leukemia: progress through collaboration. J Clin Oncol 2015 Sep 20;33(27):2938-2948. [doi: 10.1200/JCO.2014.59.1636]
9. Pui CH, Robison LL, Look AT. Acute lymphoblastic leukaemia. The Lancet 2008 Mar;371(9617):1030-1043. [doi: 10.1016/S0140-6736(08)60457-2]
10. Karrman K, Forestier E, Heyman M, et al. Clinical and cytogenetic features of a population-based consecutive series of 285 pediatric T-cell acute lymphoblastic leukemias: rare T-cell receptor gene rearrangements are associated with poor outcome. Genes Chromosomes Cancer 2009 Sep;48(9):795-805. [doi: 10.1002/gcc.20684] [Medline: 19530250]
11. Gupta SK, Singh M, Chandrashekar PH, et al. Clinical and prognostic impact of copy number alterations and associated risk profiles in a cohort of pediatric B-cell precursor acute lymphoblastic leukemia cases treated under ICiCLe protocol. Hemasphere 2022 Oct;6(10):e782. [doi: 10.1097/HS9.0000000000000782] [Medline: 36204689]
12. Gogoi MP, Das P, Das N, et al. Risk stratified treatment for childhood acute lymphoblastic leukaemia: a multicentre observational study from India. Lancet Reg Health Southeast Asia 2025 Jun;37:100593. [doi: 10.1016/j.lansea.2025.100593] [Medline: 40487914]

XSL•FO
RenderX

13. Pandey A, Ahlawat S, Singh A, Singh S, Murari K, Aryan R. Outcomes and impact of minimal residual disease (MRD) in pediatric, adolescent and young adults (AYA) with acute lymphoblastic leukemia treated with modified MCP 841 protocol. Cancer Res Stat Treat 2020;3(2):183-191. [doi: 10.4103/CRST.CRST_85_20]
14. Kodidela S, Dorababu P, Thakkar DN, et al. Association of NUDT15*3 and FPGS 2572C>T variants with the risk of early hematologic toxicity during 6-MP and low-dose methotrexate-based maintenance therapy in Indian patients with acute lymphoblastic leukemia. Genes (Basel) 2020 May 28;11(6):594. [doi: 10.3390/genes11060594] [Medline: 32481505]
15. Ganesan P, Jain H, Bagal B, et al. Outcomes in adolescent and young adult (AYA) acute lymphoblastic leukemia (ALL): a report from the Indian acute leukemia research database (INwARD) of the hematology cancer consortium (HCC). Blood 2019 Nov 13;134(Supplement_1):1306. [doi: 10.1182/blood-2019-127091]
16. Faruqui N, Joshi R, Lowe J, et al. Exploring barriers to accessing care for childhood cancers In India: a qualitative study (Inpog-Acc-16-03). Pediatr Hematol Oncol J 2018;3(3):S1-S2. [doi: 10.1016/j.phoj.2018.11.004]
17. Ahuja S, Sharma J, Gupta S, et al. Patient tracking during treatment of children with cancer in India—an exploratory study. Cancer Rep 2022 Jun;5(6):e1359 [FREE Full text] [doi: 10.1002/cnr2.1359]
18. Arya LS, Kotikanyadanam SP, Bhargava M, et al. Pattern of relapse in childhood ALL: challenges and lessons from a uniform treatment protocol. J Pediatr Hematol Oncol 2010 Jul;32(5):370-375. [doi: 10.1097/MPH.0b013e3181d7ae0d] [Medline: 20463606]
19. Khera S, Trehan A, Bhatia P, Singh M, Bansal D, Varma N. Prevalence of TPMT, ITPA and NUDT 15 genetic polymorphisms and their relation to 6MP toxicity in north Indian children with acute lymphoblastic leukemia. Cancer Chemother Pharmacol 2019 Feb;83(2):341-348. [doi: 10.1007/s00280-018-3732-3] [Medline: 30474703]
20. Li Y, Buijs-Gladdines JGCAM, Canté-Barrett K, et al. IL-7 receptor mutations and steroid resistance in pediatric T cell acute lymphoblastic leukemia: a genome sequencing study. PLoS Med 2016 Dec;13(12):e1002200. [doi: 10.1371/journal.pmed.1002200] [Medline: 27997540]
21. Stiller CA, Parkin DM. Geographic and ethnic variations in the incidence of childhood cancer. Br Med Bull 1996 Oct;52(4):682-703. [doi: 10.1093/oxfordjournals.bmb.a011577] [Medline: 9039726]
22. Sahana S, Bhoyar RC, Sivadas A, et al. Pharmacogenomic landscape of Indian population using whole genomes. Clin Transl Sci 2022 Apr;15(4):866-877. [doi: 10.1111/cts.13153] [Medline: 35338580]
23. Saleem SM, Jan SS. Modified Kuppuswamy socioeconomic scale updated for the year 2021. Indian J Forensic Community Med 2021 Apr 28;8(1):1-3. [doi: 10.18231/j.ijfcm.2021.001]
24. Common Terminology Criteria for Adverse Events (CTCAE): Version 5.0. : US Department of Health and Human Services; 2017 Nov 27 URL: https://dctd.cancer.gov/research/ctep-trials/for-sites/adverse-events/ctcae-v5-5x7.pdf [accessed 2026-02-04]
25. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in pediatric cancer: reliability and validity of the pediatric quality of life inventory generic core scales, multidimensional fatigue scale, and cancer module. Cancer 2002 Apr 1;94(7):2090-2106. [doi: 10.1002/cncr.10428] [Medline: 11932914]
26. Reich D, Thangaraj K, Patterson N, Price AL, Singh L. Reconstructing Indian population history. Nature 2009 Sep 24;461(7263):489-494. [doi: 10.1038/nature08365] [Medline: 19779445]
27. The 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature New Biol 2010 Oct 28;467(7319):1061-1073. [doi: 10.1038/nature09534]
28. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015;4:7. [doi: 10.1186/s13742-015-0047-8] [Medline: 25722852]
29. Storey JD, Taylor JE, Siegmund D. Strong control, conservative point estimation and simultaneous conservative consistency of false discovery rates: a unified approach. J R Stat Soc B 2004 Feb 1;66(1):187-205. [doi: 10.1111/j.1467-9868.2004.00439.x]
30. R Project. URL: https://www.r-project.org/ [accessed 2026-02-09]
31. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001 Apr;68(4):978-989. [doi: 10.1086/319501] [Medline: 11254454]
32. Hsieh AR, Hsiao CL, Chang SW, Wang HM, Fann CSJ. On the use of multifactor dimensionality reduction (MDR) and classification and regression tree (CART) to identify haplotype-haplotype interactions in genetic studies. Genomics 2011 Feb;97(2):77-85. [doi: 10.1016/j.ygeno.2010.11.003] [Medline: 21111805]
33. Li Z, Sillanpää MJ. Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection. Theor Appl Genet 2012 Aug;125(3):419-435. [doi: 10.1007/s00122-012-1892-9] [Medline: 22622521]
34. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet 2006 Feb;38(2):209-213. [doi: 10.1038/ng1706] [Medline: 16415888]

Abbreviations

ALL: acute lymphoblastic leukemia
EFS: event-free survival
FDR: false discovery rate

XSL•FO
RenderX

HIC: high-income country
ICiCLe-ALL-14: Indian Childhood Collaborative Leukemia—Acute Lymphoblastic Leukemia 2014
MPGx-INDALL: Molecular and Pharmacogenetic Marker Evaluation in Relation to the Toxicity and Clinical Response of Acute Lymphoblastic Leukemia Treatment in Indian Children
OS: overall survival
PIROS: Standardized Protocol Items Recommendations for Observational Studies
QoL: quality of life
REDCap: Research Electronic Data Capture
RFS: relapse-free survival
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
TRIPOD-AI: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis plus Artificial Intelligence
TRT: treatment-related toxicity
VAF: variant allele frequency
WES: whole-exome sequencing

Edited by J Sarvestan; submitted 23.Aug.2025; peer-reviewed by M Jani, X Ding; accepted 24.Dec.2025; published 17.Mar.2026.

Please cite as:
Sharma S, Ganguly S, Murugadoss K, Kayal S, Keerthivasagam S, Ramamoorthy J, Singh A, Pushpam D, Mathaiyan J, Gloor Y, Baleydier F, Ansari M, Bakhshi S, Dubashi B, Uppugunduri CRS
Molecular and Pharmacogenetic Marker Evaluation in Relation to the Toxicity and Clinical Response of Acute Lymphoblastic Leukemia Treatment in Indian Children (MPGx-INDALL): Protocol for a Prospective Observational Cohort Study
JMIR Res Protoc 2026;15:e79865
URL: https://www.researchprotocols.org/2026/1/e79865
doi:10.2196/79865

© Swetambri Sharma, Shuvadeep Ganguly, Kamali Murugadoss, Smita Kayal, Swaminathan Keerthivasagam, Jaikumar Ramamoorthy, Archna Singh, Deepam Pushpam, Jayanthi Mathaiyan, Yvonne Gloor, Frederic Baleydier, Marc Ansari, Sameer Bakhshi, Biswajit Dubashi, Chakradhara Rao S Uppugunduri. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Associated With Loss to Follow-Up Among People Living With HIV in a National Tertiary Care Hospital: Protocol and Baseline Analysis of a Prospective Cohort Study

Luis Eduardo Del Moral Trinidad[1], MPH; Jaime Federico Andrade Villanueva[2], IDP; Luis Alberto Ruíz Mora[3], MD; Carlos Valentino García y Nuño[2], MD; Maria Fernanda Perez Quintero[2], MD; Brian Eduardo Apodaca Escalante[2], MD; Jocelyn Graciela Torres Arias[2], MD; Juan Pablo Martínez Herrera[2], MD; Melva Guadalupe Herrera Godina[1], PhD; Luz Alicia González Hernández[2,4], PhB

[1]Doctorado en Ciencias de la Salud Pública, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, Guadalajajara, Mexico

[2]Universidad de Guadalajara, Centro Universitario de Ciencias de la Salud, Unidad de VIH, Hospital Civil de Guadalajara "Fray Antonio Alcalde", Guadalajara, Mexico

[3]Infectología, Hospital Civil de Guadalajara “Juan I. Menchaca”, Guadalajara, Mexico

[4]Departamento de Clínicas Médicas, Centro Universitario de Ciencias de la Salud, Universidad de Guadalajara, Hospital 278, Guadalajara, Mexico

Corresponding Author:
Luz Alicia González Hernández, PhB
Universidad de Guadalajara, Centro Universitario de Ciencias de la Salud, Unidad de VIH, Hospital Civil de Guadalajara "Fray Antonio Alcalde", Guadalajara, Mexico

Abstract

Background: Advances in antiretroviral therapy (ART) have significantly improved the life expectancy of people living with HIV. However, maintaining retention in care—defined as ongoing engagement with medical services from diagnosis through regular follow-up—is essential for optimal clinical outcomes. Loss to follow-up (LTFU), commonly defined as the absence of ART prescription refills or medical visits for more than 90 days, has been associated with increased mortality, treatment failure, and continued community transmission. Although multiple individual and structural factors have been linked to LTFU, evidence from the Mexican context remains limited.

Objective: This protocol describes a prospective cohort study designed to identify factors associated with LTFU among recently diagnosed people living with HIV in Mexico.

Methods: We conducted a prospective cohort study at a national tertiary care hospital in Guadalajara, Jalisco, Mexico. Eligible participants were adults (≥18 years) who had initiated ART within 6 months prior to enrollment. Data on sociodemographic, clinical, HIV-related, and psychosocial variables were obtained from electronic medical records, pharmacy dispensing logs, and validated questionnaires (Simplified Medication Adherence Questionnaire, Berger HIV Stigma Scale, and the Medical Outcomes Study HIV Health Survey). The primary outcome is LTFU, defined as ≥90 consecutive days without a medical visit or ART refill, ascertained through institutional records and national ART dispensation systems. Participants will be followed for 24 months. Planned analyses include descriptive statistics, Kaplan-Meier curves for time to LTFU, and multivariable Cox and logistic regression models to identify factors independently associated with disengagement from care.

Results: Recruitment took place between December 2023 and March 2024, yielding 164 enrolled participants who completed all baseline assessments. The 24-month follow-up period for this cohort extends from April 2024 through March 2026, with primary analyses and dissemination of results planned for the second half of 2026. Baseline data indicate that the cohort is characterized by substantial socioeconomic vulnerability, a high prevalence of late presentation, and notable levels of perceived stigma and reduced health-related quality of life.

Conclusions: This protocol outlines a prospective cohort study to evaluate factors associated with LTFU among people living with HIV in a Mexican tertiary care setting. The baseline findings highlight substantial socioeconomic, clinical, and psychosocial vulnerabilities that may compromise long-term retention in care. The longitudinal follow-up of this cohort will provide context-specific evidence to inform targeted interventions aimed at improving engagement in care and reducing LTFU in similar populations.

International Registered Report Identifier (IRRID): DERR1-10.2196/76470

(JMIR Res Protoc 2026;15:e76470) doi:10.2196/76470

XSL•FO
RenderX

KEYWORDS

HIV; cohort studies; antiretroviral therapy; risk factors; Mexico; loss to follow-up

Introduction

In recent years, advancements in the treatment of HIV infection have allowed patients to achieve a life expectancy comparable to that of the HIV-negative population. This progress is largely attributable to the availability of new and more effective antiretroviral therapy (ART) regimens that also come with fewer side effects [1,2].

To achieve this, patients must comply with their prescribed pharmacological treatment and remain in regular medical care, a characteristic known as *retention in care* [3]. This refers to a patient's ongoing commitment, from the time of diagnosis through routine follow-up, to attend appointments and maintain engagement with treatment at their health care center [4]. In contrast, when a patient fails to fulfill this commitment, the phenomenon known as *loss to follow-up* (LTFU) occurs, which happens when the health care service user has not attended a medical review or received medication dispensing for 90 days or more [5-8].

Of the patients who discontinue treatment, it is known that a proportion of them will return to care after a variable period, while others will continue without treatment, resulting in various consequences [5]. These include a mortality rate that is twice as high as that of retained patients, greater difficulty in achieving viral suppression with first-line regimens [9], and the continuation of the transmission chain at the community level. Patients not retained in care are estimated to infect, on average, 5 people due to their lack of viral control [10]. Furthermore, from a health economics perspective, achieving patient retention in care can reduce morbidity, mortality, and the resources required to provide specialized medical care for disease progression [11].

The literature has reported that sociodemographic factors, such as advanced age and male sex, have been associated with a higher risk of treatment discontinuation compared with younger age groups and female sex [12]. Moreover, low educational levels, lack of access to basic household services, limited social support, and ethnic background have been associated with a higher risk of treatment discontinuation [5,13]. Likewise, perceived stigma and discrimination have been identified as additional risk factors [14]. These variables have been incorporated into the current protocol to explore their relationship with care disengagement in our population.

Furthermore, economic and clinical factors have also been shown to play a critical role in treatment discontinuation. Lack of financial resources and the associated costs of transportation and medical care represent significant barriers [15,16]. Additionally, the presence of advanced disease, a low CD4+ T cell count, or opportunistic infections increases disease burden, raises pill burden, and heightens the likelihood of LTFU [17-19]. These clinical variables are especially relevant in our setting, where many patients present with advanced disease at the time of diagnosis [20,21].

In addition, geographical factors such as living in a rural area, insufficient availability of public transportation, and the distance to the health care center also influence the retention of people living with HIV [22]. Likewise, certain regions have a higher prevalence of the disease, which is often associated with greater economic and social inequalities. This can represent an obstacle to continuous access to medication and increase the risk of LTFU [23,24].

Despite the extensive literature on factors influencing retention in HIV care from high-income and some middle-income countries, evidence from Mexico remains scarce and fragmented. National reports tend to focus on overall cascade indicators, yet fail to explore in depth the sociodemographic, clinical, and geographic barriers that affect patients' ability to remain engaged in care, particularly during the early stages following diagnosis [20,21,25,26]. Given the country's regional disparities in health service access and the sociostructural vulnerabilities faced by key populations, it is crucial to understand the context-specific factors that hinder retention.

This study is designed to address this gap by establishing a prospective cohort of recently diagnosed people living with HIV who are receiving ART at a major Mexican hospital. In this protocol, we present a baseline analysis describing their sociodemographic, clinical, geographic, and psychosocial characteristics as a first step to identify variables that will be examined as potential factors associated with LTFU. We hypothesize that sociodemographic vulnerability, clinical severity, and geographic barriers will be associated with a higher risk of attrition from care. By gaining a deeper understanding of the obstacles patients encounter in achieving retention in care, the findings can inform the development of targeted interventions to improve care engagement and address this critical public health challenge.

Methods

Study Design

This protocol was developed and reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines for cohort studies. This is a prospective cohort study designed to evaluate factors associated with ART discontinuation among people living with HIV. The exposure of interest is the ART. The current report focuses on a baseline cross-sectional analysis to describe the sociodemographic, clinical, HIV-related, and psychological characteristics of participants at enrollment.

Study Setting

The study was conducted at a specialized unit for the care of people living with HIV in the city of Guadalajara, Jalisco, Mexico. According to 2023 data, Mexico was estimated to have a prevalence of 360,000 people living with HIV, of whom 240,000 were receiving ART. In the state of Jalisco, approximately 9000 cases were reported to be in treatment that same year.

XSL•FO
RenderX

Regarding newly diagnosed patients, in the same year, a total of 19,000 new infections were registered, of which 1200 corresponded to the state of Jalisco, representing 6.3% of the new cases nationwide [27].

For this study, we selected the Hospital Civil de Guadalajara "Fray Antonio Alcalde" for recruitment and follow-up, as it treats a large number of patients annually and has a robust multidisciplinary team.

Primary Outcomes

The main outcome variable will be LTFU, defined as the absence of medical consultations or antiretroviral prescription refills for 90 consecutive days or more. This operational definition aligns with widely accepted criteria used in previous studies [21]. LTFU was selected due to its critical role in evaluating retention in HIV care, as it reflects patient disengagement and is strongly associated with increased morbidity, mortality, and community-level transmission risks [8].

The variable will be obtained from institutional electronic medical records and pharmacy dispensing logs and cross-verified by the medical care team at the HIV clinic.

Covariates

The selection of covariates—including sociodemographic, clinical, and HIV-related characteristics—was guided by prior international and national literature identifying these factors as potential predictors of disengagement from care in similar populations [18,28-32]:

Sociodemographic variables included age, marital status, educational level, number of dependent children, mode of transportation, estimated monthly income in Mexican pesos, type of employment (formal or informal), number of economic dependents, mode of transportation, and distance from home to the clinic. Clinical variables included alcohol consumption, tobacco consumption within the last 30 days, use of illicit drugs, and BMI classification. HIV-related variables included CD4+ T lymphocyte count (cells/µL), HIV-1 viral load (copies/mL), opportunistic infections at diagnosis, prophylactic treatment for opportunistic infections, disease stage, ART regimen, HIV-related hospitalizations within the last 6 months, late presenter status (<200 cells/mm^3), diagnosis during pregnancy, time on ART at recruitment, time between diagnosis and initiation of treatment, adherence to treatment, perceived stigma, and quality of life

Instruments

- Stigma—it was measured using the HIV Stigma Scale developed by Berger et al [25], consisting of 40 items validated for the Mexican population (Ω=0.86). This scale evaluates 4 dimensions: perceived stigma, social stigma, self-stigma, and concerns about disclosure. Scores are obtained through a Likert scale (1-4), with higher scores indicating greater levels of stigma. Scores can be categorized as high or moderate (>52 points) or low (<52 points) [33-35].
- Quality of life—quality of life was assessed using the Medical Outcomes Study HIV Health Survey, validated in Mexican populations (Cronbach α=0.79). This 35-item instrument measures 11 dimensions, such as physical functioning, social functioning, mental health, and general health perception. Results are summarized into 2 global indices: physical health and mental health, with scores ranging from 0 to 100, where higher values indicate better quality of life [33,36].

Adherence was evaluated using the Simplified Medication Adherence Questionnaire, validated for the Mexican population in a previous study [37]. The questionnaire consists of 6 items addressing different dimensions of adherence: intentional (eg, "Do you stop taking your medication when feeling unwell?"), unintentional (eg, "Do you forget to take your medication?"), and frequency-related adherence (eg, "How many full days have you missed your medication in the last three months?"). Nonadherence is defined as any affirmative response to questions 1, 3, 4, or 6, or a negative response to question 2 [38].

Stigma was measured using the HIV Stigma Scale developed by Berger et al [25], consisting of 40 items validated for the Mexican population (Ω=0.86). This scale evaluates 4 dimensions: perceived stigma, social stigma, self-stigma, and concerns about disclosure. Scores are obtained through a Likert scale (1-4), with higher scores indicating greater levels of stigma. Scores can be categorized as high or moderate (>52 points) or low (<52 points) [33-35].

Quality of life was assessed using the Medical Outcomes Study HIV Health Survey, validated in Mexican populations (Cronbach α=0.79). This 35-item instrument measures 11 dimensions, such as physical functioning, social functioning, mental health, and general health perception. Results are summarized into 2 global indices: physical health and mental health, with scores ranging from 0 to 100, where higher values indicate better quality of life [35,38].

Recruitment and Follow-Up

The study follows each participant for 24 months after enrollment, starting with a baseline evaluation at the time of inclusion. The initial follow-up started in December 2023 and will continue until the last patient completes the 2-year follow-up period. Patients were recruited by reviewing daily patient lists at both consultation sites, identifying first-time patients, and cross-checking their eligibility against the inclusion criteria. Eligible candidates were invited to participate during their next appointment, where the study objectives were explained in detail, questions were addressed, and informed consent was obtained following ethical guidelines. After consent, participants completed a data collection questionnaire along with the study instruments (Figure 1).

XSL•FO
RenderX

Figure 1. Study procedures and follow-up schedule. Flowchart summarizing the study procedures for baseline recruitment, informed consent, sociodemographic and psychological assessments (Medical Outcomes Study HIV Health Survey [MOSS HIV], Berger Stigma Scale, and Simplified Medication Adherence Questionnaire [SMAQ] adherence), and follow-up for detection of loss to follow-up (LTFU) through medical records, pharmacy logs, phone contact, and national databases. Only 1 baseline visit is conducted, and follow-up continues passively for up to 24 months. ART: antiretroviral therapy; EMR: electronic medical record.

XSL•FO
RenderX

When a patient experiences the event of interest LTFU, their participation in the study will be considered complete. However, we will conduct an intentional search to locate the patient. First, we will make 3 phone calls (1 per day). If there is no response, we will attempt to call again 7 days later. If this effort is unsuccessful, we will contact the social worker to arrange a home visit or explore other alternatives to reach the patient.

To ensure the patient is still active in the study, we will check the national registration database to verify if their citizen registry remains active. We will also review the national ART dispensation system to confirm that they have not received therapy at another center. Additionally, we will contact the addiction program to verify that they are not receiving treatment at a specialized addiction site. If patients are in prisons or other closed settings, we will verify their inclusion in the specialized program for individuals in such facilities. For any of these reasons, patients will be excluded from our study.

Eligibility Criteria

- Exclusion criteria—patients receiving temporary treatment at the hospital who do not reside in the state of Jalisco or are expected to be referred to another health care unit, people currently in prisons or other closed settings, patients receiving ART at facilities specialized in addiction treatment, and patients with severe cognitive limitations that prevent them from completing study questionnaires

Inclusion criteria were being aged 18 years or older, having initiated ART within 6 months before recruitment, and having provided written informed consent to participate in the study

Exclusion criteria were receiving temporary treatment at the hospital, not residing in the state of Jalisco, being expected to be referred to another health care unit, being currently in prison or another closed setting, receiving ART at facilities specialized in addiction treatment, and having severe cognitive limitations that prevent completing study questionnaires.

The exclusion of people in prisons or other closed settings or those receiving ART at facilities specialized in addiction treatment was motivated by the fact that these patients belong to special programs that deliver medication directly to their facilities or have individuals who monitor their drug refills. Thus, their retention and adherence could act as confounding factors for this study. People living with HIV with cognitive limitations were considered for exclusion, as the instruments required self-response and reliable answers needed to be ensured.

Sample Size

The sample size was calculated using a software based on the following assumptions: a 2-tailed significance level of 95%, a 1:1 ratio between patients with CD4+ T lymphocyte counts greater than or less than 200 cells/mm^3, a proportion of patients with lower CD4+ T lymphocyte levels experiencing LTFU of 15%, and an expected risk ratio of 3.68, as reported in previous studies by Gerver et al [2]. Thus, the minimum required sample size is 120 participants, with 60 in each group.

While the cohort was not stratified by CD4 count upon recruitment, this factor is known to substantially impact long-term retention in care [2]. This distribution provides sufficient power to detect differences in retention outcomes across CD4 levels.

Risk of Bias

To minimize selection bias, we recruited all consecutive eligible patients starting ART at the HIV clinic during the study period, applying predefined inclusion and exclusion criteria. Information bias was reduced by using standardized case report forms; extracting clinical data from electronic medical records; and administering validated instruments for adherence, stigma, and quality of life that have been adapted for Mexican populations. Research staff received specific training on questionnaire administration and data entry, and the database was checked for inconsistencies and duplicate entries. To limit misclassification of LTFU, we cross-verified appointments and ART refills using electronic records, pharmacy dispensing logs, the national ART registry, and records from specialized programs (eg, addiction or prison-based care) before classifying a participant as LTFU. In the planned analyses, we will account for confounding by including relevant sociodemographic, clinical, and psychosocial variables in multivariable models and performing sensitivity analyses for key assumptions.

Statistical Analysis

Descriptive analyses were used to summarize the sociodemographic, clinical, HIV-related, and psychological characteristics of the study participants. Categorical variables are presented as absolute numbers and percentages, while continuous variables are presented as means and SDs for normally distributed data or medians and IQRs for nonnormally distributed data.

For the longitudinal cohort analysis, retention and LTFU outcomes will be assessed over a 2-year period. Kaplan-Meier survival curves will be used to estimate time to LTFU, and log-rank tests will compare survival distributions between groups stratified by CD4 count. Multivariable Cox proportional hazards models will be constructed to identify independent predictors of LTFU, adjusting for sociodemographic, clinical, and psychological factors.

Ethical Considerations

This study adheres to the International Ethical Guidelines for Health-Related Research Involving Humans established by the Council for International Organizations of Medical Sciences [1]. The protocol was reviewed and approved by the institutional ethics committee (Comité de Ética en Investigación del Hospital Civil de Guadalajara “Fray Antonio Alcalde,” approval number CEI 211/23). Written informed consent was obtained from all participants prior to enrollment, and the process was conducted in the presence of an independent witness to ensure full understanding of the study procedures. To protect privacy and confidentiality, participants were assigned unique alphanumeric codes, and all identifying information was stored separately from the analytic database, which is kept in a secure repository accessible only to authorized members of the research team. Data used for analysis are deidentified, and results will be reported in aggregate form only. Participants did not receive any financial incentives or compensation for taking part in the

XSL•FO
RenderX

study, and their decision to participate or not had no impact on the clinical care they received.

Results

Recruitment began in December 2023 and was completed in March 2024, reaching a total of 164 participants. All eligible individuals provided written informed consent and completed the baseline data collection instruments, which included sociodemographic and clinical variables, and validated scales for ART adherence, HIV-related stigma, and health-related quality of life.

As of April 2024, all participants had entered the follow-up phase, which consists of passive monitoring for LTFU over a 2-year period. The follow-up strategy includes linkage to national databases to verify ART refill continuity, transfers to other care programs, or exclusion based on predefined criteria (eg, incarceration or entry into addiction treatment programs).

The follow-up phase spans 24 months per participant and uses passive monitoring strategies based on linkage to institutional clinical records and national databases to ascertain treatment continuity, ART refills, program transfers, or LTFU events.

Key milestones are as follows:

- December 2023 to March 2024—participant recruitment and baseline data collection
- April 2024 to March 2026—passive follow-up of participants (monitoring for LTFU, deaths, or exclusions)
- April 2026 to June 2026—final data cleaning, classification of outcomes, and dataset closure
- July 2026 to October 2026—statistical analysis and manuscript drafting for registered report 2
- Late 2026—expected submission of results manuscript for peer review

At baseline, the cohort was predominantly composed of young adult men and showed high levels of socioeconomic vulnerability, with frequent late presentation to care and important psychosocial challenges, including perceived stigma and moderate health-related quality of life. Detailed descriptive statistics for baseline sociodemographic, clinical, HIV-related, and psychosocial variables are presented in Multimedia Appendices 1-4 for reference and transparency.

Discussion

Anticipated Findings

This prospective cohort study is designed to identify factors associated with LTFU among people living with HIV receiving ART in a tertiary-level hospital in Mexico. While no inferential analyses have yet been conducted, the baseline characteristics collected provide an important contextual framework for understanding the population under follow-up and formulating hypotheses regarding future disengagement from care.

On the basis of prior studies conducted in Mexico and other Latin American settings, it is anticipated that younger age groups and male sex may be associated with a higher risk of LTFU. These patterns have been consistently reported in regional literature and may reflect structural, occupational, and social vulnerabilities affecting continuity of care [2,3]. Understanding how these sociodemographic characteristics relate to retention over time will be critical for designing age- and gender-sensitive interventions aimed at improving long-term engagement in care [4].

Economic instability and informal employment are highly prevalent among people living with HIV in Mexico and have been described as potential barriers to sustained engagement in care. In this cohort, socioeconomic variables were intentionally included to explore their prospective association with LTFU, given their relevance in settings where out-of-pocket costs, transportation expenses, and job insecurity may interfere with regular clinic attendance [5]. The longitudinal follow-up of this cohort will allow evaluation of whether these economic factors translate into a higher risk of disengagement from care.

Substance use, nutritional status, and clinical severity at diagnosis were included as covariates based on strong evidence linking these factors to HIV treatment outcomes in other populations [6]. Although baseline prevalence varies across settings, their role as potential predictors of LTFU will be assessed prospectively in this cohort.

Similarly, late presentation, advanced disease stage, and the presence of opportunistic infections are hypothesized to increase the risk of disengagement from care due to greater clinical complexity, pill burden, and psychosocial stressors, as described in a previous study [7].

Psychosocial factors, such as perceived stigma and health-related quality of life, were included given their documented association with adherence and retention in HIV care [7-9]. This study will allow exploration of whether these dimensions independently or jointly influence LTFU over time in the Mexican context.

The prospective design will allow for time-to-event and multivariable modeling of retention trajectories. Findings may help inform tailored interventions, such as economic support, transportation aid, or stigma reduction programs, aimed at reducing LTFU in similar contexts.

Dissemination Plan

We plan to disseminate findings through peer-reviewed publications, national HIV program briefings, and stakeholder engagement with clinical and community-based organizations involved in HIV care delivery in Mexico. Additionally, summarized results will be shared with patient advocacy groups and incorporated into local quality improvement initiatives.

Limitations

A major strength of this study is its prospective design and the comprehensive assessment of sociodemographic, clinical, geographic, and psychosocial variables. However, as a single-center study, generalizability may be limited. The findings from this cohort will inform future multicenter studies and targeted interventions aimed at improving retention in HIV care. Results from the completed follow-up will be presented in a subsequent manuscript.

XSL•FO
RenderX

Conclusions

This study protocol describes the design and methodological foundations of a prospective cohort study aimed at identifying factors associated with LTFU in people living with HIV receiving care in a tertiary-level hospital in Mexico. Although results are not yet available, this work seeks to address critical evidence gaps in the Mexican context, where longitudinal data on HIV care retention remain limited. By systematically following a clinical cohort and incorporating both sociodemographic and clinical variables, this study aims to provide useful insights for strengthening patient retention strategies.

The findings are expected to inform future public health interventions and improve continuity of care for people living with HIV, with potential applications in similar settings across low- and middle-income countries.

Acknowledgments

The authors are deeply grateful to the patients participating in this study, whose contribution is invaluable and essential to the development of this research. They also extend their thanks to the clinic staff for their dedication and support throughout every stage of the project. Artificial intelligence–based tools were not used to generate scientific content, analyze data, or draw conclusions in this manuscript. A language editing tool (Grammarly), which uses artificial intelligence–assisted algorithms, was used only to improve grammar, spelling, and style after the authors drafted the text.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The study was conducted using institutional resources from the participating hospital and university.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

LEDMT conceived the study, led the protocol design and methodology, coordinated data collection, performed preliminary data management, and drafted the initial version of the manuscript. JFAV contributed to the study conception, provided clinical expertise on HIV care, and critically reviewed and edited the manuscript. LARM contributed to the development of the study instruments, coordinated participant recruitment, and assisted with data collection and manuscript revision. CVNGN contributed to data curation, quality control of the database, and review of the clinical content. MFPQ supported participant recruitment and questionnaire administration and provided input on the interpretation of psychosocial measures. BEAE assisted with data entry, verified clinical records, and contributed to manuscript editing. JGTA contributed to the operational aspects of the study, including linkage with clinical services and follow-up procedures, and reviewed the manuscript for intellectual content. JPMH collaborated in refining the analytical plan, contributed to the interpretation of the study design, and critically revised the manuscript. MGHG provided senior supervision of the project and contributed to the conceptualization and overall study design. LAGH ensured methodological and ethical rigor and critically reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sociodemographic characteristics of participants at baseline.
[DOCX File, 20 KB - resprot_v15i1e76470_app1.docx]

Multimedia Appendix 2
Baseline clinical characteristics of the participants.
[DOCX File, 17 KB - resprot_v15i1e76470_app2.docx]

Multimedia Appendix 3
Baseline HIV-related characteristics.
[DOCX File, 18 KB - resprot_v15i1e76470_app3.docx]

Multimedia Appendix 4
Baseline adherence, quality of life, and stigma characteristics.
[DOCX File, 17 KB - resprot_v15i1e76470_app4.docx]

XSL•FO
RenderX

Checklist 1
STROBE checklist.
[PDF File, 118 KB - resprot_v15i1e76470_app5.pdf]

References

1. Ibiloye O, Jwanle P, Masquillier C, et al. Long-term retention and predictors of attrition for key populations receiving antiretroviral treatment through community-based ART in Benue State Nigeria: a retrospective cohort study. PLoS ONE 2021 Nov 30;16(11):e0260557. [doi: 10.1371/journal.pone.0260557] [Medline: 34847194]
2. Gerver SM, Chadborn TR, Ibrahim F, Vatsa B, Delpech VC, Easterbrook PJ. High rate of loss to clinical follow up among African HIV-infected patients attending a London clinic: a retrospective analysis of a clinical cohort. J Int AIDS Soc 2010 Aug 4;13:29. [doi: 10.1186/1758-2652-13-29] [Medline: 20684760]
3. Beres LK, Mody A, Sikombe K, et al. The effect of tracer contact on return to care among adult, “lost to follow-up” patients living with HIV in Zambia: an instrumental variable analysis. J Int AIDS Soc 2021 Dec;24(12):e25853. [doi: 10.1002/jia2.25853] [Medline: 34921515]
4. Odafe S, Idoko O, Badru T, et al. Patients’ demographic and clinical characteristics and level of care associated with lost to follow-up and mortality in adult patients on first-line ART in Nigerian hospitals. J Int AIDS Soc 2012 Sep 18;15(2):17424. [doi: 10.7448/IAS.15.2.17424] [Medline: 23010378]
5. Chi BH, Yiannoutsos CT, Westfall AO, et al. Universal definition of loss to follow-up in HIV treatment programs: a statistical analysis of 111 facilities in Africa, Asia, and Latin America. PLoS Med 2011 Oct;8(10):e1001111. [doi: 10.1371/journal.pmed.1001111] [Medline: 22039357]
6. Mirzazadeh A, Eshun-Wilson I, Thompson RR, et al. Interventions to reengage people living with HIV who are lost to follow-up from HIV treatment programs: a systematic review and meta-analysis. PLoS Med 2022 Mar 15;19(3):e1003940. [doi: 10.1371/journal.pmed.1003940] [Medline: 35290369]
7. Mugavero MJ, Westfall AO, Cole SR, et al. Beyond core indicators of retention in HIV care: missed clinic visits are independently associated with all-cause mortality. Clin Infect Dis 2014 Nov 15;59(10):1471-1479. [doi: 10.1093/cid/ciu603] [Medline: 25091306]
8. Skarbinski J, Rosenberg E, Paz-Bailey G, et al. Human immunodeficiency virus transmission at each step of the care continuum in the United States. JAMA Intern Med 2015 Apr;175(4):588-596. [doi: 10.1001/jamainternmed.2014.8180] [Medline: 25706928]
9. Maulsby C, Jain KM, Weir BW, et al. The cost and threshold analysis of Retention in Care (RiC): a multi-site national HIV care program. AIDS Behav 2017 Mar;21(3):643-649. [doi: 10.1007/s10461-016-1623-3] [Medline: 27873083]
10. Mody A, Sikombe K, Beres LK, et al. Profiles of HIV care disruptions among adult patients lost to follow-up in Zambia: a latent class analysis. J Acquir Immune Defic Syndr 2021 Jan 1;86(1):62-72. [doi: 10.1097/QAI.0000000000002530] [Medline: 33105396]
11. Tobias CR, Cunningham W, Cabral HD, et al. Living with HIV but without medical care: barriers to engagement. AIDS Patient Care STDS 2007 Jun;21(6):426-434. [doi: 10.1089/apc.2006.0138] [Medline: 17594252]
12. Weiser SD, Fernandes KA, Brandson EK, et al. The association between food insecurity and mortality among HIV-infected individuals on HAART. J Acquir Immune Defic Syndr 2009 Nov 1;52(3):342-349. [doi: 10.1097/QAI.0b013e3181b627c2] [Medline: 19675463]
13. Saberi P, Neilands TB, Vittinghoff E, Johnson MO, Chesney M, Cohn SE. Barriers to antiretroviral therapy adherence and plasma HIV RNA suppression among AIDS clinical trials group study participants. AIDS Patient Care STDS 2015 Mar;29(3):111-116. [doi: 10.1089/apc.2014.0255] [Medline: 25615029]
14. Ndege RC, Okuma J, Kalinjuma AV, et al. Failure to return pillbox is a predictor of being lost to follow-up among people living with HIV on antiretroviral therapy in rural Tanzania. HIV Med 2022 Jul;23(6):661-672. [doi: 10.1111/hiv.13223] [Medline: 34964236]
15. Chammartin F, Zürcher K, Keiser O, et al. Outcomes of patients lost to follow-up in African antiretroviral therapy programs: individual patient data meta-analysis. Clin Infect Dis 2018 Nov 13;67(11):1643-1652. [doi: 10.1093/cid/ciy347] [Medline: 29889240]
16. Akullian AN, Mukose A, Levine GA, Babigumira JB. People living with HIV travel farther to access healthcare: a population-based geographic analysis from rural Uganda. J Int AIDS Soc 2016 Feb 10;19(1):1-8. [doi: 10.7448/IAS.19.1.20171] [Medline: 26869359]
17. Hailu BA, Tadese F, Bogale GG, Molla A, Miheretu BA, Beyene J. Spatial patterns and associated factors of HIV seropositivity among adults in Ethiopia from EDHS 2016: a spatial and multilevel analysis. BMC Infect Dis 2020 Oct 14;20(1):751. [doi: 10.1186/s12879-020-05456-y] [Medline: 33054788]
18. Goswami ND, Schmitz MM, Sanchez T, et al. Understanding local spatial variation along the care continuum: the potential impact of transportation vulnerability on HIV linkage to care and viral suppression in high-poverty areas, Atlanta, Georgia. J Acquir Immune Defic Syndr 2016 May 1;72(1):65-72. [doi: 10.1097/QAI.0000000000000914] [Medline: 26630673]
19. Gwitira I, Murwira A, Mberikunashe J, Masocha M. Spatial overlaps in the distribution of HIV/AIDS and malaria in Zimbabwe. BMC Infect Dis 2018 Nov 27;18(1):598. [doi: 10.1186/s12879-018-3513-y] [Medline: 30482166]

XSL•FO
RenderX

20. González R, Augusto OJ, Munguambe K, et al. HIV incidence and spatial clustering in a rural area of southern Mozambique. PLoS ONE 2015;10(7):e0132053. [doi: 10.1371/journal.pone.0132053] [Medline: 26147473]
21. Boletín de atención integral de personas con VIH. Gobierno de Mexico. 2025. URL: https://www.gob.mx/censida/documentos/boletin-de-atencion-integral-de-personas-con-vih [accessed 2026-03-01]
22. Del Moral Trinidad LE, González Hernández LA, Andrade Villanueva JF, et al. Simplified Medication Adherence Questionnaire (SMAQ) for people living with HIV in a national hospital in Mexico: instrument validation study. Interact J Med Res 2025 Jan 7;14:e59562. [doi: 10.2196/59562] [Medline: 39773472]
23. Agala CB, Fried BJ, Thomas JC, et al. Reliability, validity and measurement invariance of the Simplified Medication Adherence Questionnaire (SMAQ) among HIV-positive women in Ethiopia: a quasi-experimental study. BMC Public Health 2020 Apr 28;20(1):567. [doi: 10.1186/s12889-020-08585-w] [Medline: 32345253]
24. Peña de León E, Aguilar Gaytán SS, Suárez Mendoza AA, Reyes Terán G. Validación mexicana de la escala MOS-HIV de calidad de vida en pacientes infectados por el VIH. Rev Panam Salud Publica 2007;21(5):313-319. [doi: 10.1590/S1020-49892007000400007]
25. Berger BE, Ferrans CE, Lashley FR. Measuring stigma in people with HIV: psychometric assessment of the HIV stigma scale. Res Nurs Health 2001 Dec;24(6):518-529. [doi: 10.1002/nur.10011] [Medline: 11746080]
26. López A, Rafful C, Orozco R, Contreras-Valdez JA, Jiménez-Rivagorza L, Morales M. HIV stigma mechanisms scale: factor structure, reliability, and validity in Mexican adults. AIDS Behav 2023 Apr;27(4):1321-1328. [doi: 10.1007/s10461-022-03868-2] [Medline: 36287344]
27. Oguntibeju OO. Quality of life of people living with HIV and AIDS and antiretroviral therapy. HIV AIDS (Auckl) 2012;4:117-124. [doi: 10.2147/HIV.S32321] [Medline: 22893751]
28. Pautas éticas internacionales para la investigación relacionada con la salud con seres humanos. : Consejo de Organizaciones Internacionales de las Ciencias Médicas; 2017 URL: https://cioms.ch/wp-content/uploads/2017/12/CIOMS-EthicalGuideline_SP_INTERIOR-FINAL.pdf [accessed 2026-03-01]
29. Azamar-Alonso A, Bautista-Arredondo SA, Smaill F, Mbuagbaw L, Costa AP, Tarride JE. Patient characteristics and determinants of CD4 at diagnosis of HIV in Mexico from 2008 to 2017: a 10-year population-based study. AIDS Res Ther 2021 Nov 13;18(1):84. [doi: 10.1186/s12981-021-00409-0] [Medline: 34774077]
30. Haris M, Abbas R. Four decades of HIV: global trends, testing assays, treatment, and challenges. Zoonoses 2024;4(1):e997. [doi: 10.15212/ZOONOSES-2023-0039]
31. Maulsby CH, Ratnayake A, Hesson D, Mugavero MJ, Latkin CA. A scoping review of employment and HIV. AIDS Behav 2020 Oct;24(10):2942-2955. [doi: 10.1007/s10461-020-02845-x] [Medline: 32246357]
32. Swann M. Economic strengthening for retention in HIV care and adherence to antiretroviral therapy: a review of the evidence. AIDS Care 2018 Jul 8;30(sup3):99-125. [doi: 10.1080/09540121.2018.1479030]
33. Amirkhanian YA, Kelly JA, DiFranceisco WJ, et al. Predictors of HIV care engagement, antiretroviral medication adherence, and viral suppression among people living with HIV infection in St. Petersburg, Russia. AIDS Behav 2018 Mar;22(3):791-799. [doi: 10.1007/s10461-016-1638-9] [Medline: 27990579]
34. Bhatia R, Hartman C, Kallen MA, Graham J, Giordano TP. Persons newly diagnosed with HIV infection are at high risk for depression and poor linkage to care: results from the Steps Study. AIDS Behav 2011 Aug;15(6):1161-1170. [doi: 10.1007/s10461-010-9778-9] [Medline: 20711651]
35. Uebelacker LA, Weisberg RB, Herman DS, Bailey GL, Pinkston-Camp MM, Stein MD. Chronic pain in HIV-infected patients: relationship to depression, substance use, and mental health and pain treatment. Pain Med 2015 Oct;16(10):1870-1881. [doi: 10.1111/pme.12799] [Medline: 26119642]
36. Gonzalez A, Barinas J, O'Cleirigh C. Substance use: impact on adherence and HIV medical treatment. Curr HIV/AIDS Rep 2011 Dec;8(4):223-234. [doi: 10.1007/s11904-011-0093-5] [Medline: 21858414]
37. Terzian AS, Younes N, Greenberg AE, et al. Identifying spatial variation along the HIV care continuum: the role of distance to care on retention and viral suppression. AIDS Behav 2018 Sep;22(9):3009-3023. [doi: 10.1007/s10461-018-2103-8] [Medline: 29603112]
38. Bilinski A, Birru E, Peckarsky M, et al. Distance to care, enrollment and loss to follow-up of HIV patients during decentralization of antiretroviral therapy in Neno District, Malawi: a retrospective cohort study. PLoS ONE 2017 Oct 3;12(10):e0185699. [doi: 10.1371/journal.pone.0185699] [Medline: 28973035]

Abbreviations

ART: antiretroviral therapy
LTFU: loss to follow-up
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

XSL•FO
RenderX

Edited by A Schwartz; submitted 23.Apr.2025; peer-reviewed by A AL-Asadi; accepted 17.Feb.2026; published 18.Mar.2026.

Please cite as:
Del Moral Trinidad LE, Andrade Villanueva JF, Ruíz Mora LA, García y Nuño CV, Perez Quintero MF, Apodaca Escalante BE, Torres Arias JG, Martínez Herrera JP, Herrera Godina MG, González Hernández LA
Factors Associated With Loss to Follow-Up Among People Living With HIV in a National Tertiary Care Hospital: Protocol and Baseline Analysis of a Prospective Cohort Study
JMIR Res Protoc 2026;15:e76470
URL: https://www.researchprotocols.org/2026/1/e76470
doi:10.2196/76470

© Luis Eduardo Del Moral Trinidad, Jaime Federico Andrade Villanueva, Luis Alberto Ruíz Mora, Carlo Valentino García y Nuño, Maria Fernanda Perez Quintero, Brian Eduardo Apodaca Escalante, Jocelyn Graciela Torres Arias, Juan Pablo Martínez Herrera, Melva Guadalupe Herrera Godina, Luz Alicia González Hernández. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 18.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Influencing the Transition From Physiological Hyperopia to Myopia in Children: Protocol for a Prospective Nested Case-Control Study

Huaying Xu[1*], MD; Jianing Pu[2*], MD; Shimeng Bian[1], MD; Xuejing Mi[1], MD; Zhen Zhou[3], MD; Wei Chen[2], MD; Yonghong Jiao[1], MD

[1]Department of Ophthalmology, Beijing Tongren Eye Center, Capital Medical University, Beijing Tongren Hospital, 2 West Ring South Road, Beijing Economic and Technological Development Zone, Beijing, China

[2]Department of Ophthalmology, Maternal and Child Health Hospital of Haidian District, Beijing, China

[3]School of Biomedical Engineering, Capital Medical University, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Yonghong Jiao, MD
Department of Ophthalmology, Beijing Tongren Eye Center, Capital Medical University, Beijing Tongren Hospital, 2 West Ring South Road, Beijing Economic and Technological Development Zone, Beijing, China

Abstract

Background: The early onset of myopia in children has become a critical public health issue that requires urgent attention. Notably, high myopia-related retinal diseases have emerged as the leading cause of irreversible blindness in adults in certain regions of China. Physiological hyperopia, as a protective factor and one of the strongest predictors of myopia development, plays a key role in delaying the progression of early-onset myopia and reducing the risk of high myopia in adulthood. However, the dynamic changes, critical turning points, and factors contributing to the rapid regression of physiological hyperopia during childhood remain unclear.

Objective: This study aims to explore the key influencing factors for children's physiological hyperopia fading and myopia onset, as well as the protective mechanisms of physiological hyperopia against myopia.

Methods: This was a prospective nested case-control study. Our research team previously established a prospective cohort of 2109 preschool children, aged 3 to 6 years, through cluster sampling in 22 kindergartens in Haidian District, Beijing, and completed a 3-year follow-up. Building on this cohort, this study adopts a prospective nested case-control design. We will continue to follow up this cohort with a 3-year longitudinal study, tracking children aged 6 to 9 years. During the 3-year follow-up period, participants will undergo annual eye examinations and complete questionnaires regarding their living habits and environment. The primary outcome is incident myopia, while secondary outcomes include the prevalence of myopia and the changes in various ocular biological parameters. The study received ethics approval on April 7, 2024, from the Ethics Committee of Beijing Tongren Eye Center, Beijing Tongren Hospital (TREC2024-KY034) and was registered on July 5, 2024.

Results: Participant recruitment began on July 10, 2024, and is expected to be completed by December 31, 2026. As of February 2026, recruitment is ongoing, and the final results are expected by December 2026.

Conclusions: This prospective nested casecontrol study investigates the dynamic changes and regression patterns of physiological hyperopia in school-aged children and evaluates the associations between ocular biometric parameters and the onset of myopia. The findings are expected to support standardized myopia screening, inform clinical interventions, and provide evidence-based guidance for government policies on myopia prevention and control in young children.

Trial Registration: ClinicalTrials.gov NCT06498947; https://clinicaltrials.gov/study/NCT06498947

International Registered Report Identifier (IRRID): DERR1-10.2196/84888

(JMIR Res Protoc 2026;15:e84888) doi:10.2196/84888

KEYWORDS

physiological hyperopia; myopia; children; risk factor; nested case-control study

XSL•FO
RenderX

Introduction

Myopia is a major global public health problem and is projected to affect approximately 50% of the world's population by 2050 [1]. It is the most common ocular condition in many Asian countries, with prevalence rates substantially higher in regions such as China, Singapore, Hong Kong, and Taiwan than in Western countries [2,3]. A 2025 nationwide analysis of more than 5 million Chinese children and adolescents reported a persistently high prevalence of myopia, with a steep age-related increase and projections indicating a continued rise in high myopia, highlighting the long-term public health burden in China [4].

Myopia adversely affects the physical and psychological well-being of children and adolescents and may lead to serious ocular complications, including retinal detachment, macular degeneration, glaucoma, and cataract [5-7]. In some regions of China, high myopia-associated retinopathy has become the leading cause of irreversible blindness in adults [8], resulting in a marked reduction in quality of life. Myopia is an irreversible ocular condition, and treatment options are limited to controlling its progression and preventing complications after onset [9]. Therefore, early prevention and intervention of myopia in children in China are urgent public health priorities.

Physiologically, newborns typically present with hyperopia, with an average refractive error of approximately +2.50 to +3.00 diopters (D). With increasing age, hyperopia gradually decreases and progresses toward emmetropia (−0.50 to +0.50 D) [5]. However, environmental and behavioral changes in modern society—including increased exposure to artificial lighting, widespread use of electronic devices, reduced outdoor activity, and intensified early educational demands—have substantially altered this developmental process [10-13]. In some children, these factors lead to an early onset of emmetropia, which often persists only briefly before progressing to myopia. This premyopic hyperopic state, referred to as physiological hyperopia, may serve as an important predictor of early myopia onset [14]. From an epidemiological perspective, reducing disease incidence is central to effective prevention. In childhood myopia, this strategy translates into preventing myopia among children who are not yet affected. Accordingly, the trajectory of physiological hyperopia during early childhood represents a critical target for myopia prevention. However, the mechanisms underlying physiological hyperopia changes, the timing of critical transition points, and the factors driving rapid physiological hyperopia regression and subsequent myopia development remain poorly understood.

In China, childhood myopia screening and prevention are guided by national public health policies. The Myopia Prevention and Control Guidelines (2024 edition) issued by the National Health Commission recommend regular vision screening in kindergarten and primary school settings and emphasize early identification of children at risk for myopia. These national recommendations are implemented locally through school-based vision screening programs in cities such as Beijing. However, evidence on the longitudinal trajectory of physiological hyperopia and its transition to myopia during the critical preschool to school-age period remains limited.

Therefore, this study conducted a nested case-control investigation among school-aged children in the Haidian District of Beijing to establish a cohort of children with physiological hyperopia. The baseline values of refractive error, visual acuity, axial length (AL), anterior chamber depth (ACD), corneal radius (CR), and other ocular biometric parameters were recorded following cycloplegia. Through longitudinal follow-up, this study aims to identify key factors influencing physiological hyperopia regression and myopia onset, as well as to explore the protective role of physiological hyperopia in myopia development. The findings are expected to provide evidence to inform population-level strategies for early myopia prevention and control.

Methods

Objective

This study aims to explore the key influencing factors for children's physiological hyperopia fading and myopia onset, as well as the protective mechanisms of physiological hyperopia against myopia.

Study Design

This study is a prospective observational study based on an established cohort, with planned nested case-control analyses embedded within the longitudinal follow-up framework. All participants are followed annually using standardized ophthalmic examinations and questionnaires. Incident myopia cases identified during follow-up are used for the primary inferential analyses through risk-set matched nested case-control sampling, while cohort-level and cross-sectional analyses are conducted to describe longitudinal refractive and biometric changes. The primary inferential analyses are conducted using risk-set matched nested case-control sampling focused on incident myopia. The study period spans from January 1, 2024, to December 31, 2026, and the protocol is registered at ClinicalTrials.gov (NCT06498947).

Our research team previously established a prospective cohort of 2109 preschool children, aged 3 to 6 years, through cluster sampling in 22 kindergartens in Haidian District, Beijing, and completed a 3-year follow-up. Building on this established cohort, this study continues prospective follow-up as children transition into school age (6 - 9 years), with annual ophthalmic examinations and questionnaire-based assessments.

Although all participants are followed longitudinally, the primary scientific objective of this study is not to model population-average trajectories alone, but to identify factors associated with the transition from physiological hyperopia to incident myopia during this critical developmental window. For this purpose, a prospective nested case-control design was selected as the main inferential strategy within the cohort framework.

The nested case-control study is an innovative research method that combines elements of traditional case-control and cohort studies, and its validity has been demonstrated in multiple

XSL•FO
RenderX

large-scale studies [15-17]. Nested case-control sampling enables efficient and focused comparison between children who develop myopia and those who remain nonmyopic at the same follow-up time point, thereby preserving the temporal sequence between exposure and outcome while minimizing selection bias. In addition, key exposures of interest—such as baseline hyperopic reserve and the rate of physiological hyperopia regression—are age dependent and time varying. Risk-set matching within the cohort facilitates appropriate control for age and follow-up time and supports robust multivariable modeling of detailed exposure profiles.

Case-control sampling is performed only when new cases of myopia are identified during follow-up. Each incident myopia case is matched in a 1:1 ratio with a control selected from cohort members who remain at risk at the same time point. Cross-sectional analyses conducted at each follow-up visit are intended solely for descriptive and exploratory purposes and do not constitute part of the primary inferential analyses.

Study Setting and Participants

This study focuses on school-aged children in Haidian District, Beijing, and includes a 3-year prospective follow-up. As the educational and technological hub of northwest Beijing, Haidian District has a well-established education management system, facilitating research coordination. Furthermore, the high level of parental compliance with myopia prevention measures in the region ensures the successful implementation of long-term follow-up.

Infants typically exhibit a hyperopia of +3.0 to +4.0 D, which transitions into a period of rapid development before the age of 3 years, followed by a slower developmental phase. By approximately 12 years, the process of emmetropization is gradually completed. On the basis of this developmental pattern, this study focuses on school-aged children as the target population. For this study, children aged 6 to 9 years who met the eligibility criteria were invited to continue annual follow-up examinations for an additional 3 years. Figure 1 illustrates the overall participant flow, including cohort enrollment, eligibility assessment, annual follow-up visits, and identification of incident myopia cases for nested case-control analyses. Textbox 1 lists the inclusion and exclusion criteria.

Figure 1. Flowchart of the study.

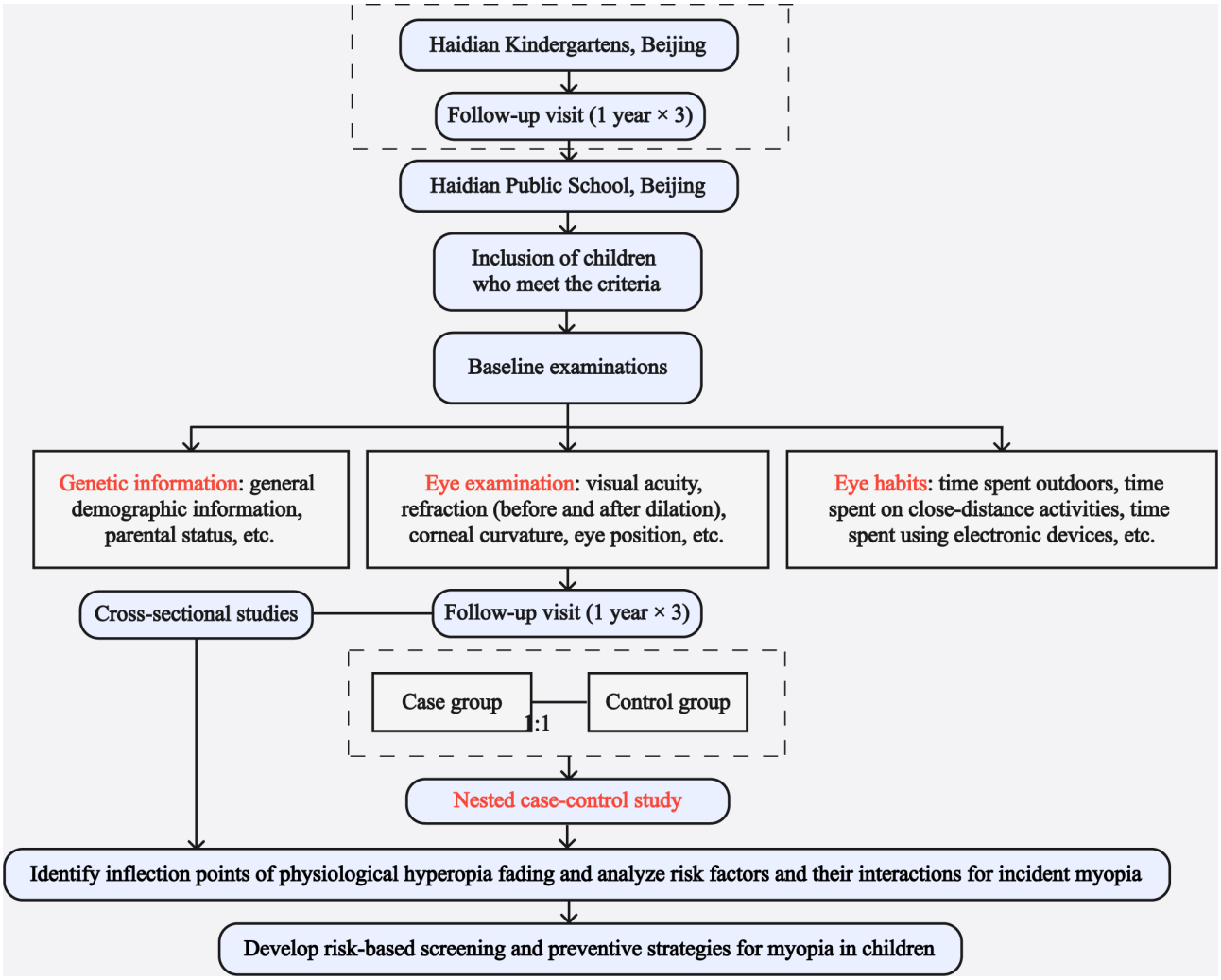

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Children aged 6 to 9 years, regardless of sex
- Ability to cooperate during examinations
- Parental consent obtained through signed informed consent

Exclusion criteria

- Children with a history of drug allergies
- Children with heart disease, cranial trauma, epilepsy, Down syndrome, or glaucoma
- Children with a history of myopia prevention interventions (eg, atropine or peripheral defocus corrective lenses)

Sample Size Calculation

The sample size was estimated based on the planned analyses of incident myopia over the 3-year follow-up in the prospective cohort, with prespecified 1:1 matching in the nested case-control component. Published epidemiological studies indicate that myopia prevalence is relatively low in preschool-aged children but increases rapidly after school entry. For example, Lan et al [18] reported that the prevalence of myopia among 3-year-old children was <2.5%, whereas He et al [19] reported a prevalence of 5.9% among 6-year-old children and an average prevalence of approximately 36% among primary school students in Guangzhou.

On the basis of these data and the anticipated number of incident myopia cases during follow-up, the minimum required cohort size was estimated to be 915 participants to support multivariable analyses and the prespecified nested case-control comparisons. Allowing for an anticipated 10% loss to follow-up, the final target sample size was increased to 1006 participants. This cohort size is expected to generate a sufficient number of incident myopia cases during follow-up to allow for 1:1 risk-set matching and adjustment for key confounders.

Outcome Measures

Primary Outcome

The primary outcome is incident myopia, defined as the development of myopia during follow-up among participants who were nonmyopic at baseline. Myopia is defined as a spherical equivalent refraction (sphere +1/2 cylinder) of ≤−0.50 D in either eye under cycloplegia [20].

Secondary Outcomes

The secondary outcomes assessed in this study include the following:

- Prevalence of myopia at baseline and each annual follow-up visit [4]
- Annual change in AL [21]
- Annual change in ACD [22]
- Annual change in CR [23]
- Annual change in the AL to CR ratio [23]

All outcomes are assessed annually throughout the 3-year follow-up period using standardized protocols.

Examinations

Parameters measured included cycloplegic refraction, recording of refractive changes, visual acuity, postdilated refraction, AL, ACD, CR, eye position, fundus, height, and weight.

All the abovementioned metrics were collected at each follow-up visit, and data were collected annually for a specified period.

Questionnaires

General Information Questionnaire

Participants' general information was collected, including name, sex, gestational age at birth, mode of delivery, birth weight, whether the parents were myopic, the degree of myopia, and the health of the mother during pregnancy.

Myopia-Related Genetic and Environmental Factors Questionnaire

These data were collected using Children's Myopia-Related Factors Questionnaire, which parents completed by scanning the QR code; the questionnaire includes past history, family history, close-distance activities, outdoor activity time, reading habits, and education.

Data Collection and Follow-Up

This study does not involve any intervention. All procedures are observational and consist of standardized ophthalmic examinations and questionnaire-based assessments conducted at baseline and at each annual follow-up visit. The following biological parameters were measured annually in children in the cohort: visual acuity, cycloplegic refraction, AL, ACD, CR, ocular position, fundus, height, weight, and other parameters.

First, visual acuity was assessed using the Early Treatment Diabetic Retinopathy Study visual acuity chart at a distance of 4 m, with the 1.0 line at eye level. Visual acuity was assessed monocularly, with one eye covered without applying pressure while the other eye was tested. Measurements were performed in a right-to-left sequence. The number of correctly identified letters was determined by reading optotypes from top to bottom.

Second, cycloplegic refraction was performed using 1% cyclopentolate hydrochloride. Initially, 1 drop of 0.4% oxybuprocaine hydrochloride (as topical anesthetic) was instilled. Following a 3-minute interval, 1% cyclopentolate hydrochloride eye drops were administered at 5- to 10-minute intervals for 3 consecutive doses. Pupillary diameter and light

XSL•FO
RenderX

reflex were assessed 30 minutes after the final instillation. Complete cycloplegia was defined as pupillary dilation ≥6.0 mm and absence of pupillary light reflex (grade 0 on the standardized light reaction scale). Participants meeting these criteria underwent automated refraction using a calibrated autorefractor (Nidek ARK-1), with triplicate measurements obtained under mesopic conditions (3 cd/m²).

Third, ocular biometric parameters, including AL, CR, and ACD, are measured using an optical interferometric AL meter (AL-Scan, Nidek Co, Ltd). The AL to CR ratio is calculated as AL-to-CR=AL/mean CR. The device automatically obtained multiple measurements and reported averaged values.

Fourth, eye position was assessed using masking-demasking and alternating masking methods to screen for strabismus. The direction of ocular movement was assessed to determine whether the participant had strabismus or occult strabismus.

Data collection was completed annually within the specified time frame, with a total follow-up period of 3 years. To ensure patient adherence and maintain a high long-term follow-up rate, the following strategies were implemented: (1) collaborating institutions had extensive screening experience and a dedicated professional examination team; (2) a personalized refractive profile was established for each child in the cohort, along with a feedback mechanism for results; (3) volunteer certificates were awarded to children participating in the cohort study to enhance their engagement; and (4) an additional clinical coordinator was appointed to conduct detailed telephone communication with families before each annual follow-up, aiming to reduce lost follow-up rates. These measures also support the feasibility and acceptability of long-term follow-up in school-aged children and their families.

Quality Control Measures and Data Management

The research team established standard operating procedures for the study and organized training programs, with records maintained in the project management files. To ensure that study personnel are adequately qualified to perform specific clinical research tasks, all involved staff must complete the following training before initiation of the project: (1) study protocol; (2) ophthalmic examination standard operating procedures, including protocols for cycloplegic refraction, visual acuity assessment, AL measurement, ACD, and other biometric parameter measurements; and (3) electronic database use. Training related to the project will be ongoing throughout the study, with the principal investigator and project coordinator adding content as needed.

The purpose of this study is to propose data management methods based on the Smart Vision Information OPERA-MC Adolescent Vision Screening and Monitoring Management System. Through an electronic questionnaire, we collected basic information about patients, their families, and their environment. Automated refraction (Nidek ARK-1) and noncontact tonometry (Canon TX-20P; Canon Medical Systems) were digitally interfaced via the standardized HL7 protocol through networked medical device ports. Digital fundus photography images (Topcon TRC-50DX; Topcon Healthcare, Inc) were automatically transmitted and archived in digital imaging and communications in medicine format using picture archiving and communication system integration. Examination results were transmitted electronically for direct collection and storage without manual data entry. This process improved data accuracy and timeliness while reducing the likelihood of missing data. We will check the accuracy and completeness of the database storage results by sampling the original data and comparing them with the stored data. The collected data will be recorded in case report forms and uploaded to the electronic data capture system for centralized management.

Statistical Analysis Plan

In this project, the basic statistical design plan was a prospective nested case-control study, as shown in Figure 2. Statistical analyses will be conducted according to prespecified research questions and outcomes. Analyses will include descriptive analyses, primary inferential analyses for incident myopia, and secondary analyses for longitudinal changes in ocular biometric parameters.

Descriptive analyses will summarize participants' baseline characteristics and annual follow-up measurements, including refractive status and ocular biometric parameters. Continuous variables will be reported as means (SDs) or medians (IQRs), as appropriate, and categorical variables will be summarized as frequencies and proportions. The distribution of continuous variables will be assessed using graphical methods (eg, histograms and Q-Q plots) and the Shapiro-Wilk test.

The primary inferential analysis will evaluate the association between physiological hyperopia and incident myopia. A nested case-control approach will be applied, in which children who develop myopia during follow-up will be identified as cases and matched in a 1:1 ratio with controls selected from cohort members who remain nonmyopic at the same follow-up time point. Conditional logistic regression models will be used to estimate odds ratios and 95% CIs, adjusting for potential confounders, including age, sex, parental myopia, time spent in outdoor activities, and time spent on near-work activities.

XSL•FO
RenderX

Figure 2. The basic statistical design plan of the study.

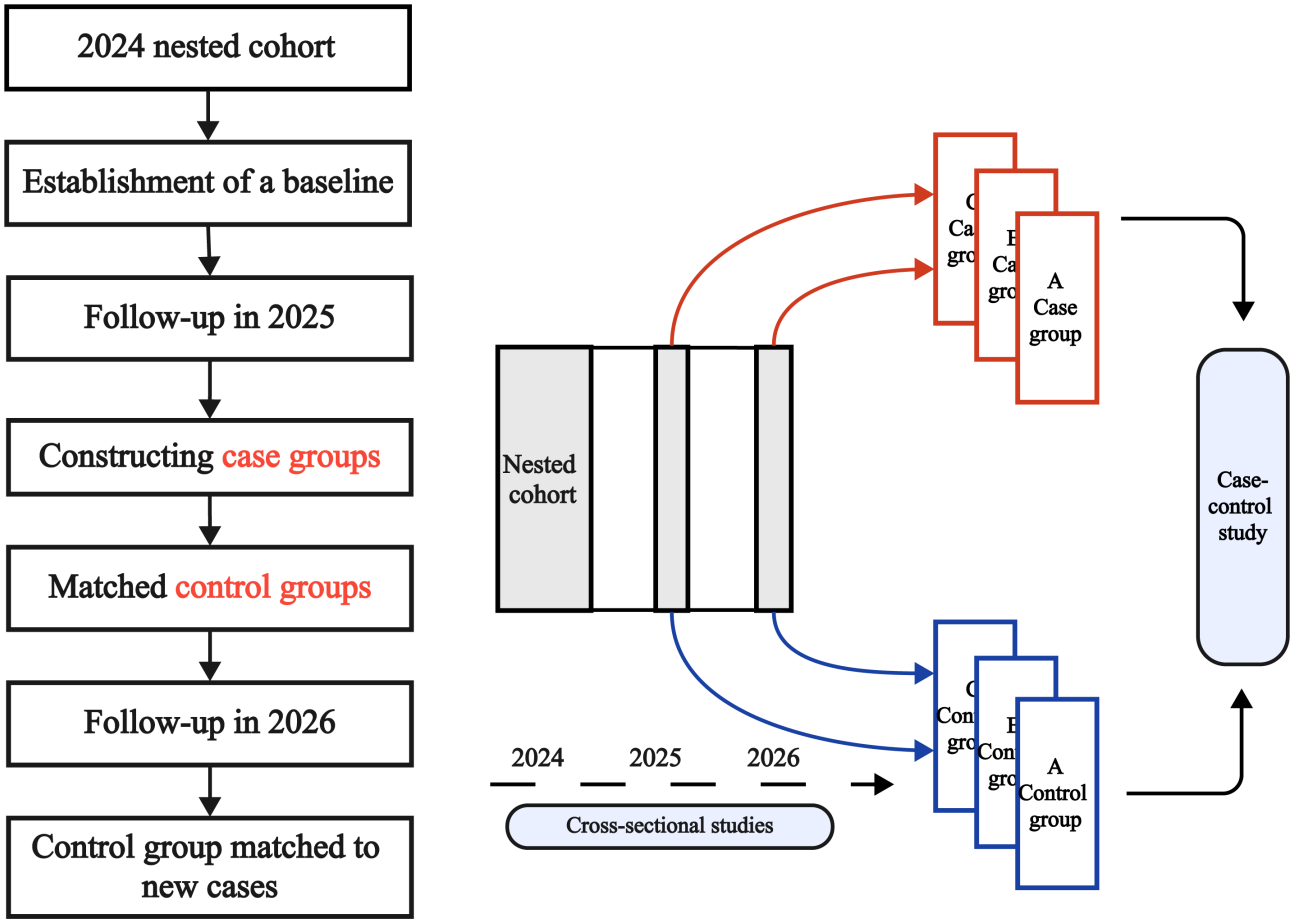

Secondary analyses will examine longitudinal changes in ocular biometric parameters, including AL, ACD, CR, and the AL to CR ratio. These analyses will be performed using linear mixed-effects models to account for within-subject repeated measurements over time. Participants will be grouped according to baseline levels of physiological hyperopia, and sensitivity analyses will be conducted accordingly.

Sensitivity analyses will be performed by stratifying participants according to baseline physiological hyperopia levels (eg, age-specific categories) and by evaluating alternative analytic-eye definitions (right eye only vs both eyes using appropriate correlation structures).

All statistical analyses will be performed using SPSS (version 26.0; IBM Corp), RStudio (version 1.4; Posit PBC), and MATLAB (version 2021a; MathWorks Inc). A 2-sided $P<.05$ will be considered statistically significant.

Ethical Considerations

Throughout the study, all procedures were in accordance with the Declaration of Helsinki.

The study received ethics approval on April 7, 2024, from the Ethics Committee of Beijing Tongren Eye Center, Beijing Tongren Hospital, (TREC2024-KY034), and was registered on July 5, 2024. In addition to explaining the purpose and methods of the study, the researchers also explained possible side effects and precautions online before each test. Informed consent was obtained before using cyclopentolate eye drops. No financial or material compensation was provided to participants for their participation in this study.

Medical records and datasets will be maintained at Beijing Tongren Hospital. No personal information will be revealed in any public results report. Data will be examined for authenticity before being given to the study statistician. Data will be kept strictly confidential and evaluated by study staff and authorized personnel. All data files will be password protected and stored on a secure server for a minimum of 10 years and disposed of if they are no longer required. If changes are made to the project, they will be published in ClinicalTrials.gov.

Results

Participant recruitment began on July 10, 2024, and is expected to be completed by December 31, 2026. As of February 2026, recruitment is ongoing. Data analysis will commence after the completion of data collection, and the final results are expected to be published by December 31, 2026.

Discussion

Anticipated Findings

This study has several key advantages. First, it incorporates the regression trajectory of physiological hyperopia into the research framework, using a prospective cohort design to track its association with the onset and progression of myopia. This approach systematically reveals the dynamic changes in the

XSL•FO
RenderX

emmetropization process in children, providing vital scientific evidence for the development of myopia prevention strategies for young children. Second, the study uses internationally standardized cycloplegic refraction methods to accurately measure physiological hyperopia in children, along with multiple ocular biometric parameters. The data obtained help fill a gap in the understanding of basic ocular development parameters in young children. Third, this study integrates longitudinal cohort follow-up with prospectively planned nested case-control analyses, allowing focused investigation of risk factors associated with the transition from physiological hyperopia to incident myopia during a critical developmental period. This approach provides a solid methodological foundation for establishing causal relationships in changes in physiological hyperopia.

Several interventions are currently used to address childhood myopia, including behavioral strategies (such as increasing outdoor activity and optimizing near-work habits), optical interventions (eg, orthokeratology and spectacle or contact lenses designed to induce peripheral myopic defocus), and pharmacological approaches, most commonly low-dose atropine eye drops [24-27]. These interventions have demonstrated efficacy in slowing myopia progression after onset. However, they are typically implemented once myopia has already developed and therefore do not directly address the identification of children at high risk before disease onset. This study complements existing intervention strategies by focusing on early refractive development and the role of physiological hyperopia as a potential marker for early myopia risk stratification.

Physiological hyperopia has a protective effect on the development of myopia in children. In a 13-year longitudinal study of kindergarten and school children aged 3 to 17 years, conducted between 1984 and 1996, Matsumura et al [28] found a significant increase in the incidence of myopia among children aged 7 years or older. Myopia development was most prominent in the −1 and −2 D groups, while the change in refractive status was minimal in the +1 D group. This suggests that a certain degree of physiological hyperopia may protect against the onset of myopia. Li et al [29] conducted a 5-year annual follow-up study involving 2835 first-grade students from Anyang City to identify potential risk factors for myopia. They found that the 5-year incidence of myopia was lowest in children with a baseline spherical equivalent refraction greater than +2.00 D (4.4%), and it increased to nearly 92% in children with a baseline spherical equivalent refraction between 0.00 and −0.50 D. These findings highlight the importance of monitoring hyperopic refractive changes in children before they enter elementary school, especially in the first grade.

Previous evidence suggests that the lower the level of physiological hyperopia, the earlier the onset of myopia. Population-based evidence from the Refractive Error Study in Children conducted by Morgan et al [3] in Australia suggests that mild hyperopia (+0.5 to +2.0 D), the typical end point of refractive development in children, may still pose a risk of progressing to myopia if emmetropia develops during childhood. Zadnik et al [30] in the United States conducted a 21-year longitudinal study of 4512 ethnically diverse nonmyopic schoolchildren in grades 1 through 8 to assess risk factors for myopia onset. They found that lower baseline hyperopia was associated with a higher risk of developing myopia, suggesting that future myopia prevention trials should target children with low levels of hyperopia, that is, physiological hyperopia deficiency. Therefore, during the critical period of refractive transition from preschool to school age—before myopia actually develops—further research is needed to identify children at high risk of myopia based on relevant risk factors and to determine effective strategies to preserve physiological hyperopia to delay or prevent the onset of myopia.

Limitations

This study has several limitations that should be considered when interpreting the findings. First, although the cohort is relatively large and well characterized, participants were recruited from a single district in Beijing. Therefore, the findings may not be fully generalizable to children in other regions with different socioeconomic, educational, or environmental contexts. Second, this study is observational in nature. While the prospective design and nested case-control analyses strengthen temporal inference, causal relationships between physiological hyperopia–related exposures and myopia onset cannot be definitively established, and residual confounding may persist despite multivariable adjustment. Third, physiological hyperopia and related environmental exposures were assessed at discrete annual time points. More frequent measurements might better capture short-term fluctuations in refractive development and behavioral factors. In addition, some exposure information was collected via parent-reported questionnaires, which may be subject to recall or reporting bias. Finally, although standardized cycloplegic refraction and ocular biometry were used, analytic decisions such as eye selection and categorization of physiological hyperopia may influence effect estimates. However, sensitivity analyses were conducted to assess the robustness of the findings to these analytic choices.

Conclusions

This prospective nested case-control study investigates the dynamic changes and regression patterns of physiological hyperopia in school-aged children and evaluates the associations between ocular biometric parameters and the onset of myopia. The findings are expected to support standardized myopia screening, inform clinical interventions, and provide evidence-based guidance for government policies on myopia prevention and control in young children.

Acknowledgments

The authors thank all the staff members for their contributions to this research.

XSL•FO
RenderX

Funding

This study was supported by the Capital's Funds for Health Improvement and Research (grant CFH2024-2G-2057).

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: YJ
Formal analysis: HX (lead), JP (equal), ZZ (supporting)
Investigation: HX, JP
Supervision: WC
Writing—original draft: HX (lead), JP (equal)
Writing—review and editing: HX, JP, ZZ, YJ, WC, SB, XM

Conflicts of Interest

None declared.

Checklist 1
STROBE checklist.
[PDF File, 199 KB - resprot_v15i1e84888_app1.pdf]

References

1. Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. Ophthalmology 2016 May;123(5):1036-1042. [doi: 10.1016/j.ophtha.2016.01.006] [Medline: 26875007]
2. Pan CW, Ramamurthy D, Saw SM. Worldwide prevalence and risk factors for myopia. Ophthalmic Physiol Opt 2012 Jan;32(1):3-16. [doi: 10.1111/j.1475-1313.2011.00884.x] [Medline: 22150586]
3. Morgan IG, French AN, Ashby RS, et al. The epidemics of myopia: aetiology and prevention. Prog Retin Eye Res 2018 Jan;62:134-149. [doi: 10.1016/j.preteyeres.2017.09.004] [Medline: 28951126]
4. Pan Z, Xian H, Li F, et al. Myopia and high myopia trends in Chinese children and adolescents over 25 years: a nationwide study with projections to 2050. Lancet Reg Health West Pac 2025 Jun 11;59:101577. [doi: 10.1016/j.lanwpc.2025.101577] [Medline: 40568343]
5. Flitcroft DI. The complex interactions of retinal, optical and environmental factors in myopia aetiology. Prog Retin Eye Res 2012 Nov;31(6):622-660. [doi: 10.1016/j.preteyeres.2012.06.004] [Medline: 22772022]
6. Wong TY, Ferreira A, Hughes R, Carter G, Mitchell P. Epidemiology and disease burden of pathologic myopia and myopic choroidal neovascularization: an evidence-based systematic review. Am J Ophthalmol 2014 Jan;157(1):9-25.e12. [doi: 10.1016/j.ajo.2013.08.010] [Medline: 24099276]
7. Liang YB, Friedman DS, Wong TY, et al. Prevalence and causes of low vision and blindness in a rural Chinese adult population: the Handan Eye Study. Ophthalmology 2008 Nov;115(11):1965-1972. [doi: 10.1016/j.ophtha.2008.05.030] [Medline: 18684506]
8. Iwase A, Araie M, Tomidokoro A, et al. Prevalence and causes of low vision and blindness in a Japanese adult population: the Tajimi Study. Ophthalmology 2006 Aug;113(8):1354-1362. [doi: 10.1016/j.ophtha.2006.04.022] [Medline: 16877074]
9. Chen M, Shu Q, Li F, Li L, Fan X. The whole life cycle myopia management. Asia Pac J Ophthalmol (Phila) 2025;14(6):100161. [doi: 10.1016/j.apjo.2025.100161] [Medline: 39875085]
10. Hinterlong JE, Holton VL, Chiang CC, Tsai CY, Liou YM. Association of multimedia teaching with myopia: a national study of school children. J Adv Nurs 2019 Dec;75(12):3643-3653. [doi: 10.1111/jan.14206] [Medline: 31566786]
11. Tsai TH, Liu YL, Ma IH, et al. Evolution of the prevalence of myopia among Taiwanese schoolchildren: a review of survey data from 1983 through 2017. Ophthalmology 2021 Feb;128(2):290-301. [doi: 10.1016/j.ophtha.2020.07.017] [Medline: 32679159]
12. Wang D, Xiong R, Zhang J, et al. Effect of extracurricular after-school physical activities on academic performance of schoolchildren: a cluster randomized clinical trial. JAMA Pediatr 2023 Nov 1;177(11):1141-1148. [doi: 10.1001/jamapediatrics.2023.3615] [Medline: 37721735]
13. Wang CY, Hsu NW, Yang YC, Chen YL, Shyong MP, Tsai DC. Premyopia at preschool age: population-based evidence of prevalence and risk factors from a serial survey in Taiwan. Ophthalmology 2022 Aug;129(8):880-889. [doi: 10.1016/j.ophtha.2022.03.017] [Medline: 35331752]
14. Bian S, Pu J, Chen W, et al. Preschool children with insufficient physiological hyperopia experience accelerated myopic drift. Optom Vis Sci 2025 May 1;102(5):271-279. [doi: 10.1097/OPX.0000000000002235] [Medline: 39977000]

XSL•FO
RenderX

15. Wong ES, Zhang XJ, Yuan N, et al. Association of optical coherence tomography angiography metrics with detection of impaired macular microvasculature and decreased vision in amblyopic eyes: the Hong Kong Children Eye Study. JAMA Ophthalmol 2020 Aug 1;138(8):858-865. [doi: 10.1001/jamaophthalmol.2020.2220] [Medline: 32584368]
16. Tsui JC, Yu Y, VanderBeek BL. Association of treatment type and loss to follow-up with tractional retinal detachment in proliferative diabetic retinopathy. JAMA Ophthalmol 2023 Jan 1;141(1):40-46. [doi: 10.1001/jamaophthalmol.2022.4942] [Medline: 36454552]
17. Niazi S, Gnesin F, Thein AS, et al. Association between glucagon-like peptide-1 receptor agonists and the risk of glaucoma in individuals with type 2 diabetes. Ophthalmology 2024 Sep;131(9):1056-1063. [doi: 10.1016/j.ophtha.2024.03.004] [Medline: 38490274]
18. Lan W, Zhao F, Lin L, et al. Refractive errors in 3-6 year-old Chinese children: a very low prevalence of myopia? PLoS ONE 2013;8(10):e78003. [doi: 10.1371/journal.pone.0078003] [Medline: 24205064]
19. He M, Zeng J, Liu Y, Xu J, Pokharel GP, Ellwein LB. Refractive error and visual impairment in urban children in southern China. Invest Ophthalmol Vis Sci 2004 Mar;45(3):793-799. [doi: 10.1167/iovs.03-1051] [Medline: 14985292]
20. Morgan IG, Ohno-Matsui K, Saw SM. Myopia. Lancet 2012 May 5;379(9827):1739-1748. [doi: 10.1016/S0140-6736(12)60272-4] [Medline: 22559900]
21. Tideman JWL, Polling JR, Vingerling JR, et al. Axial length growth and the risk of developing myopia in European children. Acta Ophthalmol 2018 May;96(3):301-309. [doi: 10.1111/aos.13603] [Medline: 29265742]
22. Mutti DO, Sinnott LT, Lynn Mitchell G, et al. Ocular component development during infancy and early childhood. Optom Vis Sci 2018 Nov;95(11):976-985. [doi: 10.1097/OPX.0000000000001296] [Medline: 30339640]
23. He X, Sankaridurg P, Naduvilath T, et al. Normative data and percentile curves for axial length and axial length/corneal curvature in Chinese children and adolescents aged 4-18 years. Br J Ophthalmol 2023 Feb;107(2):167-175. [doi: 10.1136/bjophthalmol-2021-319431] [Medline: 34531198]
24. Wildsoet CF, Chia A, Cho P, et al. IMI - Interventions Myopia Institute: interventions for controlling myopia onset and progression report. Invest Ophthalmol Vis Sci 2019 Feb 28;60(3):M106-M131. [doi: 10.1167/iovs.18-25958] [Medline: 30817829]
25. He M, Xiang F, Zeng Y, et al. Effect of time spent outdoors at school on the development of myopia among children in China: a randomized clinical trial. JAMA 2015 Sep 15;314(11):1142-1148. [doi: 10.1001/jama.2015.10803] [Medline: 26372583]
26. Lam CSY, Tang WC, Tse DYY, et al. Defocus Incorporated Multiple Segments (DIMS) spectacle lenses slow myopia progression: a 2-year randomised clinical trial. Br J Ophthalmol 2020 Mar;104(3):363-368. [doi: 10.1136/bjophthalmol-2018-313739] [Medline: 31142465]
27. Yam JC, Jiang Y, Tang SM, et al. Low-concentration atropine for myopia progression (LAMP) study: a randomized, double-blinded, placebo-controlled trial of 0.05%, 0.025%, and 0.01% atropine eye drops in myopia control. Ophthalmology 2019 Jan;126(1):113-124. [doi: 10.1016/j.ophtha.2018.05.029] [Medline: 30514630]
28. Matsumura H, Hirai H. Prevalence of myopia and refractive changes in students from 3 to 17 years of age. Surv Ophthalmol 1999 Oct;44 Suppl 1:S109-S115. [doi: 10.1016/s0039-6257(99)00094-6] [Medline: 10548123]
29. Li SM, Wei S, Atchison DA, et al. Annual incidences and progressions of myopia and high myopia in Chinese schoolchildren based on a 5-year cohort study. Invest Ophthalmol Vis Sci 2022 Jan 3;63(1):8. [doi: 10.1167/iovs.63.1.8] [Medline: 34989760]
30. Zadnik K, Jones LA, Irvin BC, et al. Myopia and ambient night-time lighting. CLEERE study group. Collaborative longitudinal evaluation of ethnicity and refractive error. Nature New Biol 2000 Mar 9;404(6774):143-144. [doi: 10.1038/35004661] [Medline: 10724157]

Abbreviations

ACD: anterior chamber depth
AL: axial length
CR: corneal radius
D: diopter
SER: spherical equivalent refractive

XSL•FO
RenderX

Edited by J Sarvestan; submitted 26.Sep.2025; peer-reviewed by X Li; revised version received 28.Dec.2025; accepted 06.Jan.2026; published 24.Mar.2026.

Please cite as:
Xu H, Pu J, Bian S, Mi X, Zhou Z, Chen W, Jiao Y
Factors Influencing the Transition From Physiological Hyperopia to Myopia in Children: Protocol for a Prospective Nested Case-Control Study
JMIR Res Protoc 2026;15:e84888
URL: https://www.researchprotocols.org/2026/1/e84888
doi:10.2196/84888

© Huaying Xu, Jianing Pu, Shimeng Bian, Xuejing Mi, Zhen Zhou, Wei Chen, Yonghong Jiao. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation of the Long-Term Safety of Avacopan in Antineutrophil Cytoplasmic Antibody–Associated Vasculitis in the Real World (AvacoStar): Protocol for a Noninterventional Prospective Cohort Study

David R W Jayne[1], MD, MBChB; Raashid Luqmani[2], MD, MBBS; Benjamin Terrier[3], MD, PhD; Achim Obergfell[4], MD, MBA, PhD; Charlotte Pollet[4], MSc; Marie Boff[5], MSc; Monica Balcells-Oliver[4], MSc; Bernhard Hellmich[6], MD

[1]Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge, United Kingdom

[2]Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom

[3]Service de Médecine Interne, Hôpital Cochin, Paris, France

[4]Department of Global Medical Affairs, CSL Vifor, Glattbrugg, Switzerland

[5]Department of Biostatistics, CSL Vifor, Glattbrugg, Switzerland

[6]Medius Kliniken, University of Tübingen, Kirchheim unter Teck, Germany

Corresponding Author:
David R W Jayne, MD, MBChB
Department of Medicine, University of Cambridge, Addenbrooke's Hospital, Hills Road, Cambridge, United Kingdom

Abstract

Background: Established treatments for granulomatosis with polyangiitis (GPA) or microscopic polyangiitis (MPA) include the use of immunosuppressive agents for remission induction followed by maintenance therapy. However, patients continue to experience disease progression, organ damage, and adverse events related to current therapies. Avacopan, an oral selective C5a receptor antagonist, was approved by the European Commission in January 2022 for the treatment of adult patients with severe, active GPA or MPA in combination with rituximab (RTX) or cyclophosphamide (CYC). In the pivotal phase 3 ADVOCATE (Avacopan Development in Vasculitis to Obtain Corticosteroid Elimination and Therapeutic Efficacy) study, avacopan was noninferior to prednisone taper in achieving remission at week 26 and superior in sustaining remission at week 52; furthermore, a greater improvement in estimated glomerular filtration rate with avacopan was also observed at week 52. The AvacoStar study will generate data on the benefit and risk and safety profile of avacopan in patients in a real-world context, including in those where treatment may potentially continue beyond 1 year.

Objective: The primary objective of AvacoStar is to evaluate the incidence of defined medical events of special interest in patients with antineutrophil cytoplasmic antibody–associated vasculitis commencing avacopan. These include liver injury, cardiac safety, serious infections, and malignancy.

Methods: AvacoStar is a noninterventional, multinational, prospective postauthorization safety study. It will enroll up to 500 patients in Germany and the United Kingdom in 2 groups of approximately 250 participants each: those treated with avacopan (plus a standard of care at local investigators' discretion; usually an RTX- or CYC-based regimen) and a second cohort treated with a CYC- or RTX-based induction regimen without avacopan. Avacopan and the standard of care will be prescribed in the usual manner in accordance with the corresponding summary of product characteristics under the sole decision of the investigator. The treatment decision will fall within current established practice. Eligible participants will be aged ≥18 years with severe, active GPA or MPA as determined by the investigator at the time of commencing avacopan or non-avacopan standard-of-care induction therapy. Patients will be followed for up to 7 years.

Results: This study has been enrolling patients since September 11, 2023. The final report is expected in the second half of 2031; interim reports are planned every 24 months after the first patient first visit.

Conclusions: The AvacoStar study will be the largest European prospective real-world evidence comparative study conducted to date that evaluates the long-term safety of avacopan in severe, active GPA or MPA. This study is expected to yield important insights on the use of avacopan in severe, active GPA or MPA in a real-world setting.

Trial Registration: ClinicalTrials.gov NCT05897684; https://clinicaltrials.gov/study/NCT05897684

International Registered Report Identifier (IRRID): DERR1-10.2196/81415

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e81415) doi:10.2196/81415

KEYWORDS

antineutrophil cytoplasmic antibody–associated vasculitis; ANCA-associated vasculitis; AAV; avacopan; C5a receptor; Avacopan Development in Vasculitis to Obtain Corticosteroid Elimination and Therapeutic Efficacy; ADVOCATE; granulomatosis with polyangiitis; GPA; microscopic polyangiitis; MPA; real-world evidence

Introduction

Antineutrophil cytoplasmic antibody (ANCA)–associated vasculitis (AAV) comprises a group of rare small- and medium-sized vasculitides [1]. Granulomatosis with polyangiitis (GPA) and microscopic polyangiitis (MPA) are 2 clinicopathological forms of AAV characterized by inappropriate activation of neutrophils [2]. A third form of AAV, eosinophilic GPA, has a distinct clinical presentation and management pathway outside the scope of this study [3]. GPA and MPA are associated with the circulating ANCA against the neutrophil-expressed antigens proteinase 3 and myeloperoxidase [2].

AAV can affect several organs, but the kidneys, respiratory tract, eyes, skin, and nervous system are most commonly affected [2,3]. Despite improvements in the treatment strategies for AAV, it has been shown that long-term mortality and morbidity remain high [4]. Patients with GPA or MPA have a 1-year mortality probability of 11.1% [5]. The main causes of first-year mortality in patients with GPA or MPA are infection [5,6] and active vasculitis [5], whereas after the first year, the major causes of death are cardiovascular disease, malignancy, and infection [7,8].

Established treatments for GPA and MPA are based on the European Alliance of Associations for Rheumatology 2022, the American College of Rheumatology 2021, and the Kidney Disease: Improving Global Outcomes 2024 recommendations and involve the use of biologic and nonbiologic immunosuppressive agents for remission induction and maintenance therapy [9-11]. This consists of rituximab (RTX) and/or cyclophosphamide (CYC) with long-term use of glucocorticoids (GCs) [10,12]. Despite refinements in the dosing regimens of GCs and CYC and the introduction of RTX, patients with AAV can struggle to control their disease, with relapses occurring in 30% to 50% of patients over 5 years [13,14]. The pathology of AAV itself can lead to organ damage and chronic morbidity and mortality [15], whereas medication toxicity can also result in serious adverse effects, including increased risk of infection, malignancy, and GC-induced diabetes [13,16]. Furthermore, GCs are major contributors to impaired quality of life [17,18]. However, despite the adverse events (AEs) associated with GC therapy, GCs remain central to current induction of remission and maintenance treatment strategies [10,12].

Growing evidence has identified the role of the alternative complement cascade in the pathogenesis of AAV through activation of the C5a receptor (C5aR) [2]. Avacopan is an orally administered, small-molecule C5aR antagonist that selectively blocks the effects of C5a through the C5aR, including blocking neutrophil chemoattraction and activation [17]. Avacopan does not interfere with the production of C5b or the membrane attack complex and does not block C5a binding to a second receptor, C5L2, which is thought to have anti-inflammatory properties [19].

Avacopan was approved for the treatment of adult patients with severe, active GPA or MPA in combination with RTX or CYC by the European Commission in January 2022 [20]. The assessment of the European Union marketing authorization application for avacopan was based on data available from the clinical development program. In the phase 2 CLEAR (Complement 5a Receptor Inhibitor Levels of Efficacy and Assessment of Safety in Relapsing or Newly Diagnosed ANCA-Associated Vasculitis) study, patients treated with avacopan both with and without a reduced-dose prednisone taper achieved higher clinical response rates at 12 weeks than the control group. Both avacopan regimens demonstrated noninferiority, with AE rates remaining similar across all groups [21]. In the pivotal phase 3 ADVOCATE (Avacopan Development in Vasculitis to Obtain Corticosteroid Elimination and Therapeutic Efficacy) study, the avacopan-based regimen was noninferior to prednisone taper with the achievement of remission at week 26, was superior in sustaining remission at week 52, and was associated with improved glomerular filtration rate from baseline in patients with GPA or MPA [17]. In the ADVOCATE study, the incidences of serious AEs (SAEs), life-threatening AEs, and death were recorded. The number of SAEs (excluding events of worsening vasculitis) was 33% higher in the prednisone taper group than in the avacopan group after 52 weeks of treatment, consistent with the higher exposure to GCs in that group. In a safety analysis of combined data from the CLEAR, CLASSIC (Clinical ANCA Vasculitis Safety and Efficacy Study of Inhibitor of C5aR), and ADVOCATE clinical trials of GPA or MPA, use of avacopan was associated with fewer AEs, SAEs, white blood cell count reductions, and infections than those yielded by the non-avacopan treatment [22].

There are currently some real-world evidence studies of avacopan use that confirm ADVOCATE data in terms of remission and safety; however, cohorts are small [23,24]. In addition, the European Alliance of Associations for Rheumatology 2022 guidelines mention that long-term use of avacopan cannot be recommended at this time because there are no data on use beyond 1 year [10]. Limited data from patients treated with an avacopan-based regimen as part of an early-access program confirm the manageable safety profile of avacopan and suggest that it has the potential to be effective beyond 52 weeks in sustaining disease control [25]. Therefore, it is of interest to observe the drug in real-life use and follow up on safety in the long-term, including beyond the 12-month period of the phase 3 ADVOCATE study.

XSL•FO
RenderX

AvacoStar (NCT05897684) is a postauthorization safety study (PASS) designed to evaluate the incidence of defined medical events of special interest (MESIs) during long-term follow-up in a real-world cohort of participants commencing avacopan for severe, active GPA or MPA.

Methods

Study Participants and Sites

AvacoStar is a noninterventional, multinational, prospective cohort study enrolling approximately 500 patients with newly diagnosed or relapsing severe, active GPA or MPA from Germany and the United Kingdom who are receiving treatment as part of routine clinical practice. Inclusion and exclusion criteria are outlined in Textbox 1.

Textbox 1. Study inclusion and exclusion criteria.

Inclusion criteria

- Diagnosis of antineutrophil cytoplasmic antibody–associated vasculitis (AAV; granulomatosis with polyangiitis or microscopic polyangiitis) as determined by the investigator according to their usual practice
- Severe, active AAV at the time of commencing avacopan or non-avacopan standard-of-care induction therapy in the opinion of the investigator
- Participants of either sex and aged ≥18 years at the time of commencing avacopan or non-avacopan standard-of-care induction therapy
- Provision of written informed consent
- Commencement within the previous 6 months of or plans to commence avacopan, cyclophosphamide, or rituximab for the treatment of severe, active AAV outside of an interventional clinical study

Exclusion criteria

- Concurrent participation in an interventional study unless prospectively discussed and agreed upon with the medical monitor

Study Design and Treatment

Patients are enrolled into 1 of the 2 groups: those treated with avacopan plus a standard of care at local investigators' discretion (usually an RTX- or CYC-based regimen; the avacopan group) and a second cohort treated with a CYC- or RTX-based induction regimen without avacopan (the non-avacopan group). Avacopan and the standard of care are prescribed in the usual manner in accordance with the corresponding summary of product characteristics (SmPC) under the sole decision of the investigator. As per the SmPC, avacopan in combination with an RTX or CYC regimen is indicated for the treatment of adult patients with severe, active GPA or MPA [26]. Participants may start or stop avacopan during the study in line with usual practice. The baseline visit is defined as the day on which induction treatment (avacopan or non-avacopan standard of care with CYC or RTX) is started for active AAV. Patients are being enrolled prospectively, but up to 6 months of data may be collected retrospectively if necessary. The overall study duration is anticipated to be up to 7 years, including a recruitment period of approximately 3 years. Follow-up of enrolled patients will continue until the last patient last visit milestone, 4 years after the last participant is enrolled (Figure 1). Participants in the non-avacopan group who initiate avacopan during follow-up due to disease flare will be withdrawn from the study at that time but may cross over and re-enroll in the avacopan group if eligibility criteria are met and recruitment is still ongoing. Re-enrollment will require a new informed consent form (ICF) and assignment of a new patient identification number. Participants in the avacopan group who discontinue avacopan treatment during follow-up will remain in the study, and data will be collected as part of the avacopan group. All decisions on therapeutic or diagnostic procedures, treatments, disease management, timing of visits, and resource use are made following the investigator's usual clinical practice. Individual participant follow-up data will be collected periodically at routine clinic visits from enrollment until the last patient last visit.

XSL•FO
RenderX

Figure 1. AvacoStar study design. The schematic outlines participant enrollment; treatment decision; follow-up visits; and assessment time points, including the measurement of medical events of special interest from first patient first visit (FPFV) to last patient last visit (LPLV). GPA: granulomatosis with polyangiitis; LPFV: last patient first visit; MPA: microscopic polyangiitis.

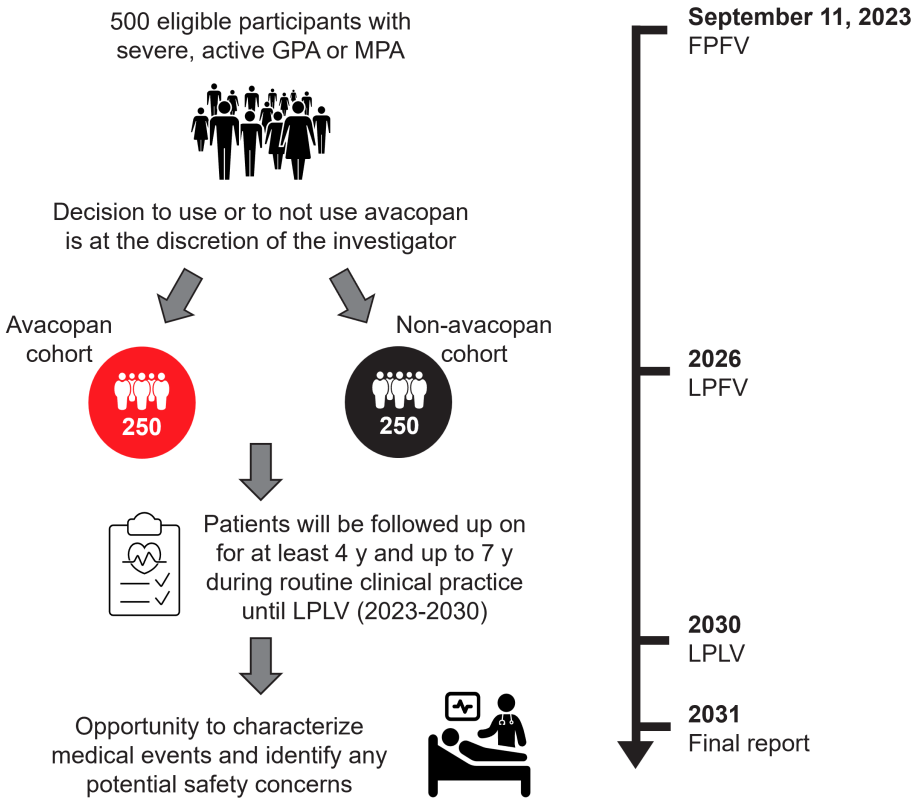

Study Objectives

The primary objective of this study is to evaluate the incidence of defined MESIs in patients with AAV commencing avacopan. This includes the incidence (number and percentage of patients with events and total number of events), incidence per patient year, and exposure-adjusted incidence rate (EAIR) per patient year with 95% CIs of defined MESIs (liver injury, cardiac safety, serious infection, and malignancy). The secondary objectives of this study are outlined in Textbox 2. This protocol fulfills the requirements of the European Medicines Agency to conduct a PASS as a postmarketing commitment.

Textbox 2. Secondary objectives of the AvacoStar study.

1. To evaluate the incidence of adverse events (AEs), AEs leading to discontinuation of therapy, serious AEs (SAEs), adverse drug reactions (ADRs), serious ADRs, laboratory abnormalities, disease flares as measured using the Birmingham Vasculitis Activity Score (BVAS), and organ damage as measured using the Vasculitis Damage Index (VDI) in patients with antineutrophil cytoplasmic antibody–associated vasculitis (AAV) commencing avacopan
2. To evaluate the background incidence of AEs, medical events of special interest (MESIs), SAEs, laboratory abnormalities, disease flares as measured using the BVAS, and organ damage as measured using the VDI in a similar population of patients with severe, active AAV but not receiving avacopan
3. To compare SAEs and MESIs in patients with AAV between patients treated with and without avacopan
4. To describe patterns of immunosuppression and glucocorticoid use in a real-world cohort
5. To describe avacopan use in a real-world cohort

Statistical Analysis

Statistical analyses will be exploratory in nature, performed using the SAS statistical software (version 9.4 or later; SAS Institute). All variables will be analyzed descriptively using appropriate statistical methods: categorical variables will be analyzed via frequency tables (absolute and relative frequencies), and continuous variables will be analyzed via descriptive statistics (ie, number of patients, mean, SD, minimum, median, quartiles, and maximum). Continuous variables will be summarized via absolute value and changes from baseline per analysis time point if applicable. Where

XSL•FO
RenderX

appropriate, 95% CIs will be provided. Comparisons between treatment groups will be performed using appropriate methods to adjust for potential bias. A propensity score will be constructed by estimating the probability of exposure by regressing the treatment assignment on baseline characteristics using logistic regression. Treatment group comparisons will be conducted using 2 methods: adjustment (the propensity score is included in the regression model) and inverse probability weighting. The proportion of patients with at least 1 event will be analyzed using logistic regression (using the Firth correction if the event is rare), and results will be reported as odds ratios with 95% CIs and 2-sided *P* values, along with estimated event proportions (least squares mean) and their 95% CIs. Incidence rates (IRs; unique IR, EAIR, and unique EAIR) per patient year will be analyzed using negative binomial regression (SAS GENMOD), with results reported as rate ratios with 95% CIs and 2-sided *P* values, along with the corresponding estimated rates per patient year and their 95% CIs. Outcomes of statistical comparisons will be interpreted with caution.

Sample Size and Power Calculations

The sample size of 250 participants in the avacopan group was considered reasonable; given an expected attrition rate of 15% or 10% per year, the power of the study to detect a single event will be 77% or 81%, respectively, for an AE with an IR per year of 0.2% for a follow-up duration of 4 years. Malignancy related to immunosuppression is the most infrequent form of the expected AEs, with the literature reporting an incidence of 1.4% per year. A sample size sufficient to detect an AE with an IR of 0.2% per year is well placed to detect a malignancy event related to immunosuppression. Skin cancer is the most common malignancy related to immunosuppression after organ transplantation, with a US study reporting a 1.4% annual incidence, well within the detection limit of this study [27].

Analysis of Primary End Point

For MESIs, the incidence (number and percentage of patients with events and total number of events), as well as the incidence per patient year, will be provided for the avacopan group. For each MESI, both the IR and EAIR will be calculated with CIs. The EAIR will account for individual differences in patient follow-ups. The EAIR per patient year will be calculated as (number of MESIs)/(total patient years of exposure). Total patient years of exposure is calculated as the sum of each patient's individual treatment duration measured from treatment initiation to the end of the exposure period. The incidence of MESIs will be presented by year as well as cumulatively over the full period of observation. To determine whether MESIs occurring after avacopan has been discontinued should be reported as part of the avacopan group, appropriate risk windows will be used. Malignancy events will be analyzed using an indefinite risk window irrespective of last avacopan dose, whereas other events (liver injury, cardiac safety, and serious infections) will be analyzed using a risk window of 9 weeks at either the end of follow-up or last avacopan dose.

Analysis of Secondary End Points

IRs of AEs leading to discontinuation of therapy, serious adverse drug reactions, and adverse drug reactions for the avacopan group will be presented. For AEs, SAEs, and MESIs, the incidence as well as the incidence per patient year will be provided by treatment group, overall, and by Medical Dictionary for Regulatory Activities system organ classes and preferred terms. All AEs and treatment-emergent AEs will be reported separately. For SAEs and MESIs, treatment groups will be compared for the proportion of patients with at least 1 event and the number of events per year, and for the IRs per patient year, the unique IR, EAIR, and unique EAIR will be reported. Descriptive statistics will also be provided by treatment group for change in Vasculitis Damage Index (VDI), change in laboratory assessments over time, proportion of GC-free patients over time, concomitant and cumulative immunosuppression over time, incidence of disease flares, and time to first flare as assessed using the Birmingham Vasculitis Activity Score (BVAS) for patients achieving remission. The duration of treatment with avacopan by reason for treatment discontinuation will be summarized descriptively.

Assessments

Baseline data being collected include demographics; smoking history; MPA or GPA diagnosis; medical history and comorbidities; prior treatment with immunosuppressive drugs; treatment group and exposure; investigator's reason for choosing the treatment group; ANCA titers; and laboratory values, BVAS, and VDI. Variables for each of the treatment groups being collected during the follow-up period include information on the use of immunosuppressives alongside avacopan- and non–avacopan-based regimens, ANCA titers, laboratory values, BVAS, VDI, AEs, MESIs, SAEs, selected safety laboratory tests, disease-related damage, flares, and use of concomitant immunosuppression and GCs. Data on AEs observed from baseline, including the seriousness, timing of onset, duration, causal relationship to avacopan, action taken with regard to avacopan administration and dosing in response to the event, and outcome of the event will also be collected. The expected timings for the collection of each variable are presented in Table 1. The study team will ensure data quality through targeted on-site verification of critical variables to confirm accuracy against source documents. Remote monitoring will enable continuous oversight, timely query resolution, and early detection of inconsistencies and deviations from the protocol. Periodic data review and ongoing data cleaning will further ensure that the final dataset is complete, reliable, and fit for analysis.

XSL•FO
RenderX

Table . Planned study schedule and expected timing for the collection of each variable in the AvacoStar study.

Frequency of measurements	Specific measurement variable
At enrollment	• Signed informed consent form[a] • Inclusion and exclusion criteria • Year of birth, sex, ethnicity, and race
At baseline[b,c]	• Height and body weight • Smoking history and current status • MPA[d] or GPA[e] diagnosis[f] • Medical history and comorbidities[g] • Prior treatment with immunosuppressive drugs[h] • Investigator's reason for choosing the treatment group • ANCA[i] binding levels[j]
At baseline, every 3 months for the first year[c], and every 6 months after the first year	• Treatment group and exposure[k]
Every 3 months for the first year, every 6 months after the first year, and at study exit or end of study	• ANCA positivity[l] • Treatment discontinuation and reasons
At baseline, every 3 months for the first year, every 6 months after the first year, and at study exit or end of study	• Laboratory tests[m] • BVAS[n] (version 3.0)[o] • VDI[p] • Concomitant medication[q] • Recording of AEs[r], SAEs[s], ADRs[t], SADRs[u], and MESIs[v] (liver injury, serious infections, malignancy, and serious cardiac events)[w]
At study exit or end of study	• Study discontinuation and reasons

[a]Informed consent includes the approval to collect retrospective data.

[b]Baseline visit corresponds to avacopan or non-avacopan induction therapy (cyclophosphamide or rituximab) start date.

[c]Patients who started avacopan or standard of care within 6 months of study start may have these data collected retrospectively.

[d]MPA: microscopic polyangiitis.

[e]GPA: granulomatosis with polyangiitis.

[f]Including antineutrophil cytoplasmic antibody–associated vasculitis phenotype, year of diagnosis, pattern of organ involvement, year of last relapse, number of relapses in the past 5 years, history of positive antineutrophil cytoplasmic antibody titer, and specificity.

[g]Medical history includes etiology of underlying disease and history of major related health issues and procedures (eg, surgery due to underlying disease).

[h]Including dose, duration before start of treatment, and reason for change in dose or medication.

[i]ANCA: antineutrophil cytoplasmic antibody.

[j]At baseline, if the patient is ANCA positive, proteinase 3 and/or myeloperoxidase binding levels will be collected (if available). The highest proteinase 3 and myeloperoxidase titers in the patient medical history will be collected (if available).

[k]Including start date; stop date (if applicable); daily dose or frequency; route of administration (intravenous or oral); reason for treatment initiation or switch; and rationale for starting, stopping, or changing the dose (completion of planned course, adverse event, investigator's or patient's decision, withdrawal of consent, administrative problems, death, or other).

[l]During follow-up, proteinase 3 or myeloperoxidase positivity will be collected.

[m]Includes creatinine, estimated glomerular filtration rate, immunoglobulin G, creatine phosphokinase, alanine transaminase, aspartate transaminase, bilirubin, white blood cell count, and albumin if collected per the site's routine practice.

[n]BVAS: Birmingham Vasculitis Activity Score.

[o]During the follow-up, the BVAS will be collected at each flare episode.

[p]VDI: Vasculitis Damage Index. At baseline, the VDI is modified to permit a baseline assessment that captures prior (non–vasculitis-related) damage, including infections.

[q]Including start date; stop date (if applicable); daily dose or frequency; route of administration (intravenous or oral); reason for treatment initiation or switch; and rationale for starting, stopping, or changing the dose (completion of planned course, adverse event, investigator's or patient's decision, withdrawal of consent, administrative problems, or other). Limited to glucocorticoids and immunosuppressants.

[r]AE: adverse event.

[s]SAE: serious AE.

[t]ADR: adverse drug reaction.

[u]SADR: serious ADR.

XSL•FO
RenderX

[v]MESI: medical event of special interest.

[w]AEs, SAEs, ADRs, or MESIs will be collected from baseline up to 9 weeks after the end of follow-up. Collection of safety events includes the seriousness, duration, causal relationship to avacopan, action taken with avacopan in response to the event, and outcome of the event.

Ethical Considerations

This study is being conducted within an approved indication in accordance with the Declaration of Helsinki, the International Society for Pharmacoepidemiology Guidelines for Good Publication Practice, the European Medicines Agency guideline on good pharmacovigilance practices module VI and VIII for PASSs, and local laws and regulations as applicable in each participating country [28-30]. The PASS will comply with Revision 6 of the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance methodological standards for study protocols, the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance Checklist for Study Protocols, and the guidelines for good publication practice of the International Society for Pharmacoepidemiology, as well as the data vendor's (IQVIA) quality management system [31-33]. Additionally, guidelines of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use concerning good clinical practice are being followed whenever possible [34].

Each study site has received approval from their regulatory authority or ethics committee before study initiation. A full list of the ethics committees and application IDs is available in Multimedia Appendix 1.

Written informed consent must be obtained from the patient (or their legally authorized representative) before participation in the study and in accordance with the local requirements. The informed consent process and signed ICF must be documented and included in their medical file, and a copy of the signed ICF (with certified translation if applicable) must be provided to the patient or representative. All signed and dated ICFs must be retained in the study file and be available for monitoring at any time. Patients may withdraw consent at any time without it affecting their medical care or access to treatment. Participants in the non-avacopan group who initiate avacopan during follow-up may re-enroll in the avacopan group if eligibility criteria are met and recruitment remains open, and they must resign the ICF at re-enrollment. Patients will not receive any compensation for their participation in this study.

Patient data are collected in a pseudonymized or deidentified form. Any new information that might influence the evaluation of the benefit and risk balance of the product shall be immediately communicated to competent authorities of member states. The analysis will be performed on the pseudonymized or deidentified patient-level dataset. The data system will be maintained and secured as required by applicable laws and regulations. Processes assuring data security will be used during data extraction, storage, and backup. The progress reports and final analysis will be reviewed and evaluated by a medical monitor and/or pharmacovigilance delegate.

Results

This study has been enrolling patients since study start on September 11, 2023, with 35 sites actively participating in enrolling patients. The final report is expected in the second half of 2031. Interim reports are planned every 2 years from the start of enrollment and will provide information about the progress of the study. All enrollment has now been completed; data analysis is under way, and the first interim report will be submitted to the European Medicines Agency in March 2026.

Discussion

Overview

Despite current treatments for AAV, disease burden still has a high impact on the quality of life of individuals with GPA or MPA [17,18,35], and many do not achieve or sustain remission [36,37]. Existing treatments cannot induce sustained disease control in all patients, and immunosuppressive agents, including GCs, are associated with frequent clinically relevant AEs and comorbidities, both resulting in impaired quality of life [13,15]. There is an unmet need for alternative, more targeted, and less toxic therapeutic options [38].

The avacopan-based regimen in combination with RTX or CYC may be considered for induction of remission in AAV as part of a strategy to substantially reduce exposure to a tapering prednisone regimen [10,11,17,39]. However, the safety of avacopan beyond 52 weeks has not been assessed. Therefore, it is of interest to monitor the drug in daily use as per clinical practice and generate long-term safety data in real-world settings. This information will be gathered in the context of this study. The main rationale for this PASS is to further understand the identified and potential risks of avacopan described above by studying the use of avacopan among patients in a real-world context, where treatment may extend beyond 1 year. The AvacoStar study is designed to assess long-term safety and observe avacopan in daily use in a real-world context.

Study Design Features

Given the heterogeneous clinical presentation of patients with AAV, the planned multicenter, multinational large sample size will provide an adequate representative number of patients. The data collected in this study will be frequently monitored during the study lifetime. Data collection will ensure that applicable variables of interest and critical data elements are routinely collected as part of real-world clinical care and are available via medical charts and physician reporting.

Strengths of the Study Design

To our knowledge, this is the largest European prospective real-world evidence study of avacopan in patients with GPA or MPA. This study evaluates long-term safety data on the use of avacopan and is the longest study to date. This study will also generate data that reflect real-world management of AAV. Although we are recruiting from a European population, the

XSL•FO
RenderX

multicenter, multinational study design will facilitate the generalizability of the study results.

Limitations of the Research Methods

As this is a noninterventional study, there are several potential confounding factors to be considered. Although several steps were taken to minimize bias, it is likely that some residual bias remains and, therefore, the results of any direct comparisons between the avacopan and non-avacopan groups will need to be interpreted with this in mind. For example, avacopan may be used by clinicians preferentially in patient subgroups where randomized controlled trial data have shown the greatest benefit (eg, severe, active renal disease or presence of GC-related comorbidities). Eligibility criteria have been selected to ensure that baseline characteristics in the non-avacopan group are similar to those of the avacopan group, and all participants enrolled in the study will be potentially eligible to receive avacopan according to the European SmPC even if they are enrolled into the non-avacopan group. Additionally, detailed information will be collected at baseline on potential confounding factors, including those related to disease severity, diagnosis, prior immunosuppression exposure, and MESI-related risk factors. To account for the risk of bias resulting from confounding factors, treatment groups will be compared using propensity scoring methodology.

Conclusions

The AvacoStar study will be a large, prospective real-world evidence study that evaluates the long-term safety of avacopan in GPA or MPA. The study is expected to yield important insights on the use of avacopan in GPA or MPA in a real-world setting, including treatment beyond 12 months.

Acknowledgments

The authors are grateful to all patients, study coordinators, and investigators who are participating in the AvacoStar study.

Funding

The AvacoStar study is funded by Vifor Fresenius Medical Care Renal Pharma Ltd. Medical writing for this manuscript was provided by Donna Kennedy (Obsidian Healthcare Group Ltd) and funded by Vifor Fresenius Medical Care Renal Pharma Ltd.

Conflicts of Interest

DRWJ has received consulting fees from Alentis Therapeutics, Amgen, AstraZeneca, Aurinia Pharmaceuticals, BMS, Boehringer Ingelheim, Chinook, GSK, Novartis, Otsuka Pharmaceutical, Roche/Genentech, Takeda Pharmaceuticals, and CSL Vifor and lecture fees from Amgen, CSL Vifor, and Otsuka Pharmaceutical and holds stocks with Alentis Therapeutics and Aurinia Pharmaceuticals. RL has received grants from BMS, Celgene, and CSL Vifor; consulting fees from GSK, Roche, and CSL Vifor; lecture fees from GSK and CSL Vifor; and advisory board fees from GSK. BT has received advisory board fees from AstraZeneca, GSK, Boehringer Ingelheim, CSL Vifor, and Novartis and lecture fees from AstraZeneca, GSK, Boehringer Ingelheim, CSL Vifor, and Novartis. AO, CP, and MB are employees of CSL Vifor. MB-O is an employee of CSL Vifor and owns stocks in Amgen and CSL Vifor. BH has received fees for consulting and/or lectures from AbbVie, AstraZeneca, BMS, GSK, Janssen, InflaRx, Novartis, Pfizer, Recordati, Roche, and CSL Vifor.

Multimedia Appendix 1
Full list of ethics committees and application IDs.
[DOCX File, 29 KB - resprot_v15i1e81415_app1.docx]

References

1. Redondo-Rodriguez R, Mena-Vázquez N, Cabezas-Lucena AM, Manrique-Arija S, Mucientes A, Fernández-Nebro A. Systematic review and metaanalysis of worldwide incidence and prevalence of antineutrophil cytoplasmic antibody (ANCA) associated vasculitis. J Clin Med 2022 May 4;11(9):2573. [doi: 10.3390/jcm11092573] [Medline: 35566698]
2. Kronbichler A, Lee KH, Denicolò S, et al. Immunopathogenesis of ANCA-associated vasculitis. Int J Mol Sci 2020 Oct 3;21(19):7319. [doi: 10.3390/ijms21197319] [Medline: 33023023]
3. Wallace ZS, Miloslavsky EM. Management of ANCA associated vasculitis. BMJ 2020 Mar 18;368:m421. [doi: 10.1136/bmj.m421] [Medline: 32188597]
4. Sánchez Álamo B, Moi L, Bajema I, et al. Long-term outcomes and prognostic factors for survival of patients with ANCA-associated vasculitis. Nephrol Dial Transplant 2023 Jun 30;38(7):1655-1665. [doi: 10.1093/ndt/gfac320] [Medline: 36617233]
5. Little MA, Nightingale P, Verburgh CA, et al. Early mortality in systemic vasculitis: relative contribution of adverse events and active vasculitis. Ann Rheum Dis 2010 Jun;69(6):1036-1043. [doi: 10.1136/ard.2009.109389] [Medline: 19574233]
6. Sachez-Alamo B, Moi L, Bajema I, et al. Long-term outcome of kidney function in patients with ANCA-associated vasculitis. Nephrol Dial Transplant 2024 Aug 30;39(9):1483-1493. [doi: 10.1093/ndt/gfae018] [Medline: 38268409]
7. Kronbichler A, Bajema IM, Bruchfeld A, Mastroianni Kirsztajn G, Stone JH. Diagnosis and management of ANCA-associated vasculitis. Lancet 2024 Feb 17;403(10427):683-698. [doi: 10.1016/S0140-6736(23)01736-1] [Medline: 38368016]

XSL•FO
RenderX

8. Quartuccio L, Treppo E, Valent F, De Vita S. Healthcare and economic burden of ANCA-associated vasculitis in Italy: an integrated analysis from clinical and administrative databases. Intern Emerg Med 2021 Apr;16(3):581-589. [doi: 10.1007/s11739-020-02431-y] [Medline: 32666174]
9. Chung SA, Langford CA, Maz M, et al. 2021 American College of Rheumatology/Vasculitis Foundation Guideline for the management of antineutrophil cytoplasmic antibody-associated vasculitis. Arthritis Rheumatol 2021 Aug;73(8):1366-1383. [doi: 10.1002/art.41773] [Medline: 34235894]
10. Hellmich B, Sanchez-Alamo B, Schirmer JH, et al. EULAR recommendations for the management of ANCA-associated vasculitis: 2022 update. Ann Rheum Dis 2024 Jan 2;83(1):30-47. [doi: 10.1136/ard-2022-223764] [Medline: 36927642]
11. Kidney Disease: Improving Global Outcomes (KDIGO) ANCA Vasculitis Work Group. KDIGO 2024 clinical practice guideline for the management of antineutrophil cytoplasmic antibody (ANCA)-associated vasculitis. Kidney Int 2024 Mar;105(3S):S71-S116. [doi: 10.1016/j.kint.2023.10.008] [Medline: 38388102]
12. Floege J, Jayne DRW, Sanders JSF, et al. Executive summary of the KDIGO 2024 clinical practice guideline for the management of ANCA-associated vasculitis. Kidney Int 2024 Mar;105(3):447-449. [doi: 10.1016/j.kint.2023.10.009] [Medline: 38388147]
13. Smith R. Complications of therapy for ANCA-associated vasculitis. Rheumatology (Oxford) 2020 May 1;59(Suppl 3):iii74-iii78. [doi: 10.1093/rheumatology/kez618] [Medline: 31967652]
14. King C, Druce KL, Nightingale P, et al. Predicting relapse in anti-neutrophil cytoplasmic antibody-associated vasculitis: a systematic review and meta-analysis. Rheumatol Adv Pract 2021 Mar 9;5(3):rkab018. [doi: 10.1093/rap/rkab018] [Medline: 34476335]
15. Alvarez L, Kambham N, Su R. Renal improvement and remission in a patient with refractory ANCA-associated vasculitis treated with avacopan. J Nephrol 2023 Nov;36(8):2365-2370. [doi: 10.1007/s40620-023-01614-y] [Medline: 37036661]
16. Scherbacher PJ, Hellmich B, Feng YS, Löffler C. Prospective study of complications and sequelae of glucocorticoid therapy in ANCA-associated vasculitis. RMD Open 2024 Feb 29;10(1):e003956. [doi: 10.1136/rmdopen-2023-003956] [Medline: 38428978]
17. Jayne DRW, Merkel PA, Schall TJ, Bekker P, ADVOCATE Study Group. Avacopan for the treatment of ANCA-associated vasculitis. N Engl J Med 2021 Feb 18;384(7):599-609. [doi: 10.1056/NEJMoa2023386] [Medline: 33596356]
18. Furuta S, Nakagomi D, Kobayashi Y, et al. Effect of reduced-dose vs high-dose glucocorticoids added to rituximab on remission induction in ANCA-associated vasculitis: a randomized clinical trial. JAMA 2021 Jun 1;325(21):2178-2187. [doi: 10.1001/jama.2021.6615] [Medline: 34061144]
19. Merkel PA, Jayne DR, Wang C, Hillson J, Bekker P. Evaluation of the safety and efficacy of avacopan, a C5a receptor inhibitor, in patients with antineutrophil cytoplasmic antibody-associated vasculitis treated concomitantly with rituximab or cyclophosphamide/azathioprine: protocol for a randomized, double-blind, active-controlled, phase 3 trial. JMIR Res Protoc 2020 Apr 7;9(4):e16664. [doi: 10.2196/16664] [Medline: 32088663]
20. Tavneos. European Medicines Agency. URL: https://www.ema.europa.eu/en/medicines/human/EPAR/tavneos [accessed 2026-02-28]
21. Jayne DRW, Bruchfeld AN, Harper L, et al. Randomized trial of C5a receptor inhibitor avacopan in ANCA-associated vasculitis. J Am Soc Nephrol 2017 Sep;28(9):2756-2767. [doi: 10.1681/ASN.2016111179] [Medline: 28400446]
22. Merkel PA, George MD, Yue H, et al. Safety of avacopan for the treatment of antineutrophil cytoplasmic antibody-associated vasculitis: combined data from three clinical trials. ACR Open Rheumatol 2025 Apr;7(4):e70001. [doi: 10.1002/acr2.70001] [Medline: 40192160]
23. Zonozi R, Aqeel F, Le D, et al. Real-world experience with avacopan in antineutrophil cytoplasmic autoantibody-associated vasculitis. Kidney Int Rep 2024 Mar 26;9(6):1783-1791. [doi: 10.1016/j.ekir.2024.03.022] [Medline: 38899183]
24. Zimmermann J, Sonnemann J, Jabs WJ, et al. Avacopan in anti-neutrophil cytoplasmic autoantibodies-associated vasculitis in a real-world setting. Kidney Int Rep 2024 Jul 6;9(9):2803-2808. [doi: 10.1016/j.ekir.2024.07.007] [Medline: 39291204]
25. Alberici F, Salvarani C, Morales E, Maxime S, Obergfell A, Popov T. #2256 Safety and effectiveness of avacopan beyond 52 weeks: experience to date in the Early Access Program (EAP). Nephrol Dial Transplant 2024 May;39(Supplement_1). [doi: 10.1093/ndt/gfae069.1249]
26. Avacopan Vifor 10 mg hard capsules. Electronic Medicines Compendium. 2025. URL: https://www.medicines.org.uk/emc/product/13744/smpc#gref [accessed 2026-02-28]
27. Garrett GL, Blanc PD, Boscardin J, et al. Incidence of and risk factors for skin cancer in organ transplant recipients in the United States. JAMA Dermatol 2017 Mar 1;153(3):296-303. [doi: 10.1001/jamadermatol.2016.4920] [Medline: 28097368]
28. Guidelines for Good Pharmacoepidemiology Practices (GPP). International Society for Pharmacoepidemiology. 2015. URL: https://www.pharmacoepi.org/resources/policies/guidelines-08027/ [accessed 2026-02-05]
29. Guideline on good pharmacovigilance practices (GVP): module VI – collection, management and submission of reports of suspected adverse reactions to medicinal products (rev 2). European Medicines Agency. 2017. URL: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/guideline-good-pharmacovigilance-practices-gvp-module-vi-collection-management-submission-reports-suspected-adverse-reactions-medicinal-products-rev-2_en.pdf [accessed 2026-02-05]

XSL•FO
RenderX

30. Guideline on good pharmacovigilance practices (GVP): module VIII – post-authorisation safety studies (rev 3). European Medicines Agency. 2017. URL: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-good-pharmacovigilance-practices-gvp-module-viii-post-authorisation-safety-studies-rev-3_en.pdf [accessed 2026-02-05]
31. Guide on methodological standards in pharmacoepidemiology (revision 11). The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). 2010. URL: https://encepp.europa.eu/document/download/f6e403a6-8033-4c22-a5ff-195ba3666299_en?filename=01.ENCePPMethodsGuideRev.11.pdf [accessed 2026-03-01]
32. ENCePP checklist for study protocols. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. URL: https://encepp.europa.eu/encepp-toolkit/encepp-checklist-study-protocols_en [accessed 2026-02-05]
33. Clinical data management. Merative. URL: https://www.merative.com/clinical-development [accessed 2026-02-05]
34. ICH harmonised guideline: guideline for good clinical practice E6(R3). International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. 2025. URL: https://database.ich.org/sites/default/files/ICH_E6%28R3%29_Step4_FinalGuideline_2025_0106.pdf [accessed 2026-02-05]
35. Basu N, McClean A, Harper L, et al. The characterisation and determinants of quality of life in ANCA associated vasculitis. Ann Rheum Dis 2014 Jan;73(1):207-211. [doi: 10.1136/annrheumdis-2012-202750] [Medline: 23355077]
36. Stone JH, Merkel PA, Spiera R, et al. Rituximab versus cyclophosphamide for ANCA-associated vasculitis. N Engl J Med 2010 Jul 15;363(3):221-232. [doi: 10.1056/NEJMoa0909905] [Medline: 20647199]
37. Specks U, Merkel PA, Seo P, et al. Efficacy of remission-induction regimens for ANCA-associated vasculitis. N Engl J Med 2013 Aug 1;369(5):417-427. [doi: 10.1056/NEJMoa1213277] [Medline: 23902481]
38. Nozaki Y. New insights into novel therapeutic targets in ANCA-associated vasculitis. Front Immunol 2021 Apr 1;12:631055. [doi: 10.3389/fimmu.2021.631055] [Medline: 33868250]
39. Stone JH, McDowell PJ, Jayne DRW, et al. The glucocorticoid toxicity index: measuring change in glucocorticoid toxicity over time. Semin Arthritis Rheum 2022 Aug;55:152010. [doi: 10.1016/j.semarthrit.2022.152010] [Medline: 35486995]

Abbreviations

AAV: antineutrophil cytoplasmic antibody–associated vasculitis
ADVOCATE: Avacopan Development in Vasculitis to Obtain Corticosteroid Elimination and Therapeutic Efficacy
AE: adverse event
ANCA: antineutrophil cytoplasmic antibody
BVAS: Birmingham Vasculitis Activity Score
C5aR: C5a receptor
CLASSIC: Clinical Antineutrophil Cytoplasmic Antibody Vasculitis Safety and Efficacy Study of Inhibitor of C5a Receptor
CLEAR: Complement 5a Receptor Inhibitor Levels of Efficacy and Assessment of Safety in Relapsing or Newly Diagnosed Antineutrophil Cytoplasmic Antibody–Associated Vasculitis
CYC: cyclophosphamide
EAIR: exposure-adjusted incidence rate
GC: glucocorticoid
GPA: granulomatosis with polyangiitis
ICF: informed consent form
IR: incidence rate
MESI: medical event of special interest
MPA: microscopic polyangiitis
PASS: postauthorization safety study
RTX: rituximab
SAE: serious adverse event
SmPC: summary of product characteristics
VDI: Vasculitis Damage Index

XSL•FO
RenderX

Edited by J Sarvestan; submitted 22.Sep.2025; peer-reviewed by K Wojcik, SE Sattui; revised version received 24.Jan.2026; accepted 28.Jan.2026; published 24.Mar.2026.

Please cite as:
Jayne DRW, Luqmani R, Terrier B, Obergfell A, Pollet C, Boff M, Balcells-Oliver M, Hellmich B
Evaluation of the Long-Term Safety of Avacopan in Antineutrophil Cytoplasmic Antibody–Associated Vasculitis in the Real World (AvacoStar): Protocol for a Noninterventional Prospective Cohort Study
JMIR Res Protoc 2026;15:e81415
URL: https://www.researchprotocols.org/2026/1/e81415
doi:10.2196/81415

© David R W Jayne, Raashid Luqmani, Benjamin Terrier, Achim Obergfell, Charlotte Pollet, Marie Boff, Monica Balcells-Oliver, Bernhard Hellmich. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors That Influence Patients' Decisions About Repetitive Transcranial Magnetic Stimulation as a Treatment Option for Treatment-Resistant Depression: Protocol for a Prospective Mixed Methods Cohort Study

Alexandra Godinho[1], MScCH; Sanvitti Dalmia[1], MDes; Steven Selchen[1], MD, MST; Daniel Tziatis[1], MBA; Pete Wegier[1,2,3], PhD

[1]Research Institute, Humber River Health, 1235 Wilson Avenue, North York, ON, Canada

[2]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada

[3]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Alexandra Godinho, MScCH
Research Institute, Humber River Health, 1235 Wilson Avenue, North York, ON, Canada

Abstract

Background: Treatment-resistant depression (TRD), affecting approximately 20% to 30% of individuals with major depressive disorder, is associated with significant disability, reduced quality of life, and an increased risk of hospitalization and suicide. Repetitive transcranial magnetic stimulation (rTMS), a noninvasive neuromodulation therapy, has demonstrated strong efficacy for TRD but is typically limited to research contexts or private clinics. Existing research on patient perspectives on rTMS is limited and largely retrospective, focusing on individuals who have already undergone treatment. As a result, little is known about the factors that influence patients' decisions to accept or decline rTMS at the time of referral, particularly within real-world clinical settings.

Objective: This study aims to address the gap in the literature by prospectively examining decision-making processes surrounding rTMS in a community hospital outpatient clinic.

Methods: This prospective mixed methods cohort study will recruit 30 adults with TRD referred to a public rTMS clinic. Participants will be stratified based on their decision to opt in or out of treatment. Data collection will include hybrid card sorting interviews, self-report questionnaires (assessing depression, well-being, cognitive flexibility, decisional conflict, and health literacy), and medical chart reviews. Each participant will complete a baseline and 6-month follow-up interview and survey. Qualitative data will be analyzed using constant comparative analysis, informed by bounded rationality and prospect theory. Quantitative data will be analyzed using bivariate statistics and hierarchical cluster analysis to identify patterns in decision-making factors.

Results: This study is being funded by a charitable donation from Jack and Pat Kay to Humber River Health, which is also supporting the establishment of the rTMS clinic, committed to in April 2024. Recruitment commenced in December 2025 and is expected to conclude in December 2026; no participants have been enrolled as of February 2026.

Conclusions: To the best of our knowledge, this is the first study to prospectively examine decision-making regarding rTMS in a real-world, publicly funded clinic including both individuals who initiate and those who decline treatment. The findings may inform the development of patient educational and engagement materials and highlight gaps in patient-physician communication during the rTMS decision-making process.

International Registered Report Identifier (IRRID): DERR1-10.2196/82491

(JMIR Res Protoc 2026;15:e82491) doi:10.2196/82491

KEYWORDS

treatment-resistant depression; repetitive transcranial magnetic stimulation; mental health decision-making; shared decision-making; card sorting; mixed methods research

XSL•FO
RenderX

Introduction

Treatment-resistant depression (TRD) is a significant public health challenge, affecting approximately 20% to 30% of individuals with major depressive disorder (MDD) worldwide [1]. In Canada, more than 20% of primary care patients with MDD meet the minimum criteria for TRD, defined as the failure to respond to at least 2 adequate trials of antidepressant medication [2]. Furthermore, up to 55% of individuals with MDD may not achieve remission and could be diagnosed with TRD during their lifetime [3]. Individuals with TRD experience disproportionately high levels of disability, reduced quality of life, and an increased risk of hospitalization and suicide [4,5]. As a result, patients with TRD visit outpatient clinics 58% more frequently and have 3 times more psychiatric consultations compared with individuals with non-TRD MDD, placing a greater demand on health care resources [6]. These challenges underscore the need for improved access to more effective treatment strategies beyond traditional approaches such as medication switching or augmentation with psychotherapy [3]. One promising alternative is repetitive transcranial magnetic stimulation (rTMS), a noninvasive neuromodulatory intervention that has demonstrated efficacy in individuals with TRD [7,8].

Approved by multiple health care regulatory bodies such as the US Food and Drug Administration, European Medicines Agency, and Health Canada, rTMS uses focused magnetic field pulses to induce lasting changes in neural activity in brain regions associated with the regulation of thoughts, emotions, and behavior [9]. While conventional rTMS typically involves delivering high-frequency stimulation in 40-minute sessions over 4 to 6 weeks, a modified form known as intermittent theta-burst stimulation can be delivered in 3 to 8 minutes with comparable efficacy [10]. Despite these advancements, which improve accessibility through shorter treatment durations and increased capacity and cost-effectiveness, rTMS remains largely inaccessible to most individuals with TRD [11]. In Canada, rTMS is predominately offered through private clinics and research studies; however, the last 5 years have seen a growing number of public clinics offering rTMS through philanthropic donations.

To improve access to rTMS and ensure that new clinics understand the needs of patients, it is first crucial to understand patients' perceptions, willingness, and likelihood of engaging with this treatment. The decision of individuals with TRD to pursue rTMS may be influenced by multiple factors, including knowledge of rTMS, family involvement, perceived efficacy, financial considerations, and trust in health care providers. A recent systematic review identified 4 studies examining psychiatric patient perceptions of noninvasive brain stimulation [12]. Overall, the review concluded that patients generally viewed rTMS positively, considering it to be as safe as antidepressant medication, recommending it to others, and reporting low levels of treatment-related fear. However, these studies offer limited insight into the decision-making process for rTMS. First, all 4 studies focused on patients after treatment, offering perspectives on the treatment process (eg, outcomes, side effects, and lasting changes) rather than perceptions of rTMS at the time of referral [13-16]. Second, only patients who had elected to receive rTMS were surveyed, leaving gaps in understanding the beliefs and/or barriers of those who declined treatment. Third, it is unclear whether participants in these studies were presented with a choice to pursue rTMS or whether it was the prescribed course of treatment for an acute inpatient group. Finally, all the studies relied solely on survey methods, which restrict a deeper exploration of the personal factors that influence patients' decisions to accept or refuse rTMS treatment.

In the absence of studies directly examining treatment decision-making for rTMS, research on patient choices for receiving electroconvulsive therapy (ECT) offers a valuable comparative framework. This is particularly relevant given the limited public awareness and understanding of rTMS [17], with one study revealing that many individuals mistakenly associate rTMS with procedures and side effects more characteristic of ECT [17]. Among the few studies that have been conducted on patients receiving ECT, several key factors have been identified as influencing decisions, including trust in health care providers, the severity of depression, prior experience with and knowledge of ECT, feelings of helplessness, family support, and perceived stigma [18-21]. While similar factors may shape the decision-making process for individuals referred for rTMS, no studies to date have specifically investigated this.

This study aims to prospectively examine the factors influencing individual decision-making on opting in or out of rTMS for TRD in a community hospital clinic. Using a mixed methods approach (ie, card sorting interviews, surveys, and chart reviews), we will explore patients' decision-making process at the time of referral and identify the key factors that influence their decisions. Additionally, we will examine how perceptions of the decision-making process evolve over time (ie, 6 months) and investigate individual-level characteristics such as demographic and clinical characteristics, cognitive flexibility, and general decision-making ability that may be associated with treatment decisions. The following research questions will be examined:

1. What factors (eg, trust, personal attitudes, practical barriers and accessibility, knowledge, past experiences, family and social relationships, culture, and religion) influence rTMS decision-making in patients diagnosed with TRD?
2. Do patients' perceptions of key factors influencing rTMS decision-making change 6 months after their initial decision?
3. How do individual differences in cognitive flexibility, past depression treatment use, decisional conflict, and health literacy impact the decision to opt in or out of rTMS treatment?
4. How do patients perceive their role in the rTMS referral process, and what are their preferences regarding involvement and shared decision-making?

As this study is exploratory in nature, no a priori hypotheses have been made.

XSL•FO
RenderX

Methods

Participants

Outpatients at the Humber River Health (HRH) mental health and addictions clinic who have been referred for rTMS consultation by their treating psychiatrist will be invited to take part in a study examining factors influencing rTMS decision-making, including patient perceptions of involvement and shared decision-making. Eligibility criteria include being aged ≥18 years, having a diagnosis of TRD, receiving an rTMS referral from their treating psychiatrist, and being English speaking. Potential participants with contraindications for rTMS (eg, history of epilepsy or other seizure disorder, ferrous metallic implants in the skull, history of drug abuse within the previous 3 to 6 months, any unstable medical condition, or pregnancy or attempting to become pregnant) or cognitive impairments that prevent them from providing informed consent for treatment will be excluded. Exclusion criteria will be determined by the treating psychiatrist based on clinical judgment. It should be noted that the eligibility criteria for this study are the same as the eligibility criteria to receive rTMS treatment at HRH, with the exception of being English speaking, which will be determined by research staff upon initial contact.

Study Design and Rationale

This project will be a prospective longitudinal cohort study using a concurrent mixed methods design to investigate key factors influencing rTMS decision-making in patients diagnosed with TRD. It will incorporate hybrid card sorting qualitative interviews, self-report surveys using validated scales, and medical chart reviews with a 6-month follow-up. Hybrid card sorting qualitative interviews were chosen as they are an interactive research methodology that encourages structured and accessible engagement with researchers [22]. The main analysis will include a quantitative component where cards will be grouped into clusters based on measurable attributes (eg, pile sorting and card rankings) and a qualitative constant comparative analysis (CCA) to explore the meaning that participants attach to each card when choosing whether to engage in rTMS. This dual approach will strengthen the depth and validity of the findings. Additionally, triangulating medical chart reviews will provide clinical context, offering further insights into how medical histories and treatment influence decision-making and vice versa. Finally, integrating quantitative survey data into the qualitative findings will allow for greater replicability, generalizability, and clinical utility by examining individual differences (eg, demographics and cognitive measures) that influence rTMS decision-making. This comprehensive approach will provide a more nuanced understanding of factors guiding patient decision-making for unfamiliar treatments such as rTMS.

Recruitment and Study Procedures

Patients with TRD attending HRH's outpatient mental health and addictions clinic may be referred to rTMS by their treating psychiatrist. rTMS will be available to clinically eligible patients free of cost as a result of a charitable donation to establish the new clinic; participants who do not enroll in the study will not be required to pay for treatment. The rTMS referral process will involve three steps: (1) discussing rTMS as an option with their psychiatrist, (2) receiving a formal referral to the HRH rTMS clinic, and (3) attending an initial consultation at the rTMS clinic. Patients can opt out during steps 1 and 3 by declining rTMS or by not attending or canceling their initial consultation. Opting in will be defined as attending the initial consultation and scheduling their first rTMS session. A total of 30 participants will be recruited: 15 who opt in and 15 who opt out of rTMS.

Treating psychiatrists who identify patients as appropriate for and recommend rTMS (ie, informal discussion or prescription referral) will ask patients for consent to be contacted by the research institute regarding a study on treatment decision-making, regardless of their rTMS choice. This information will be documented on a file that is shared between clinicians and research staff on HRH's institutional OneDrive (Microsoft Corp). This Microsoft Excel file will contain patient names and telephone numbers and will be destroyed upon recruitment. Research staff will contact consenting patients via telephone and invite them to participate in a study examining factors influencing rTMS decisions. For individuals who agree to an initial consultation, informed consent and study activities will be conducted subsequent to this appointment; as individuals may decline rTMS during the first consultation, this will ensure that an equal number of participants are recruited for the opt-in and opt-out groups.

Potential participants, regardless of rTMS decision, will receive an initial call from research staff to discuss the study and provide an opportunity for questions. Patients who would like more information can be sent an electronic copy of the consent form for further information via email. Informed consent will be scheduled to be obtained in person using teach-back strategies to ensure that potential participants make an informed choice to take part and understand that research activities are not part of the care they may receive while attending the clinic. Potential participants will be provided with up to 1 hour to discuss the contents of the consent form, and this discussion will be scheduled at the preference of the participant (ie, before or subsequent to a clinical appointment or on a separate occasion). Participants may decline to take part at any stage of the study and request to have their data withdrawn; however, only participants who consent to all study activities will be enrolled. In addition to obtaining informed consent to participate in the study, the consent process will include (1) optional consent to be contacted for future studies and (2) the option to be informed of study results at the close of the study. For clarity, consent to be contacted for future study interest will be collected separately. The results of the study will be distributed via email along with an optional survey on Microsoft Forms. This survey will consist of 4 questions designed to assess the accessibility, usefulness, and overall design of the summary. The questions will be derived from a modified version of the Information Assessment Method, which was originally developed to evaluate online health information for parents [23]. The collected feedback will be instrumental in refining and improving future participant-facing research materials, enhancing the overall communication strategy of our research efforts.

XSL•FO
RenderX

Participants who provide consent will be asked to complete a 1-hour in-person baseline card sorting interview and 30-minute survey. Six months after baseline, participants will be asked to schedule an additional in-person appointment to complete another 1-hour follow-up card sorting interview and 30-minute survey. Baseline and follow-up study activities can be completed over 1 or 2 appointments to accommodate participant schedules and minimize burden. Figure 1 provides the CONSORT (Consolidated Standards of Reporting Trials) diagram of the study.

Figure 1. Overview of the proposed cohort study. rTMS: repetitive transcranial magnetic stimulation.

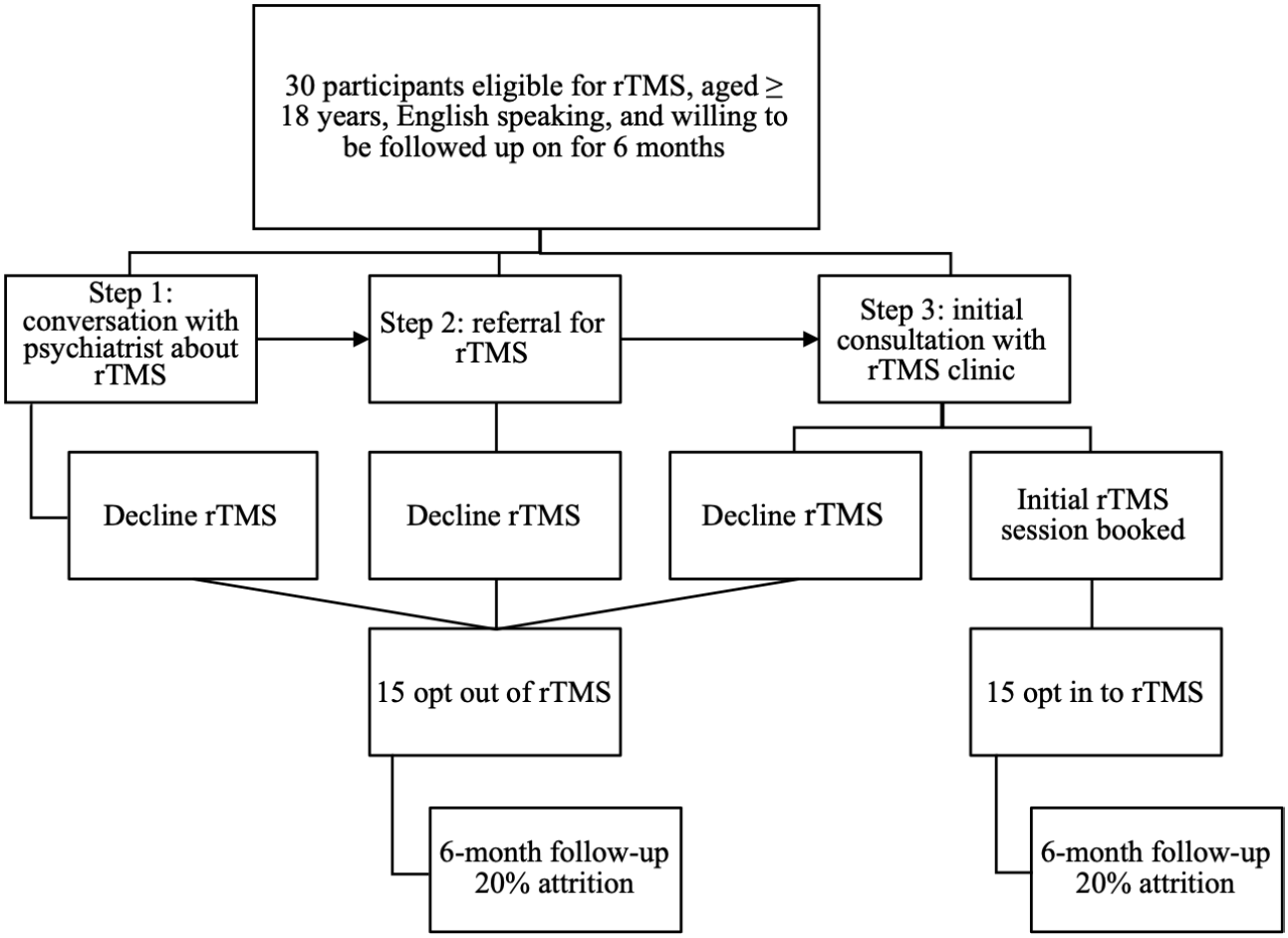

Ethical Considerations

This study was approved by the institutional research ethics board of our team's hospital, HRH (25-0003), and all participants will provide informed written consent before being enrolled in the study. Participation in this survey is voluntary, and responses will not be linked to the participant study data or any other identifying information. All participants who complete the baseline card sorting interview and survey will be compensated with a CAD $35 (US $25.57) Walmart gift card, and those who complete the 6-month follow-up card sorting interview and survey will be compensated with an additional CAD $45 (US $32.88) Walmart gift card for a potential total of CAD $80 (US $58.46) for completing all study activities. Although the time commitment for each visit is the same, the increased honorarium at follow-up is intended to promote retention and reduce attrition, consistent with common practice in longitudinal research [24,25], and compensation amounts are in accordance with published guidelines for fair compensation of research participants (ie, CAD $25 [US $18.27] per hour) [26].

Measures and Qualitative Methods

Hybrid Card Sorting Within Qualitative Interviews

A closed card sorting methodology incorporating "think-aloud" techniques [22,27,28] will be used to explore the factors influencing patient decision-making regarding rTMS treatment. The goal of the card sorting task will be to systematically identify and prioritize criteria that patients weigh when deciding to opt in or out of rTMS. A standardized deck of 24 cards has been developed with key factors relevant to decision-making for unfamiliar depression treatments such as rTMS. The content of the cards has been informed by an ongoing scoping review on individual-level factors that influence decision-making for depression treatment [29], including factors found relevant to ECT treatment choices [18-21], and by consultations with clinical staff of the HRH mental health and addictions clinic.

Participants will be provided with a briefing sheet outlining the study procedure and a set of standardized instructions before beginning the task. Once participants have reviewed the instructions with study staff, they will be presented with the deck of cards, asked to review them, and sort them into three predefined categories: (1) "important to my decision," (2)

XSL•FO
RenderX

"somewhat important to my decision," and (3) "not important at all to my decision." Within each category, participants will be asked to rank cards in descending order of importance. To ensure that a comprehensive list of factors influencing treatment decisions is captured, participants will also be provided with blank cards and markers, allowing them to include additional factors that they perceive as relevant to their decision-making that were not included in the initial deck (Figure 2). Throughout the card sorting activity, participants will be asked to verbalize their thought process to provide "real-time" insights into how they evaluate, prioritize, and consider factors when making a decision regarding rTMS.

Figure 2. Card sorting task.

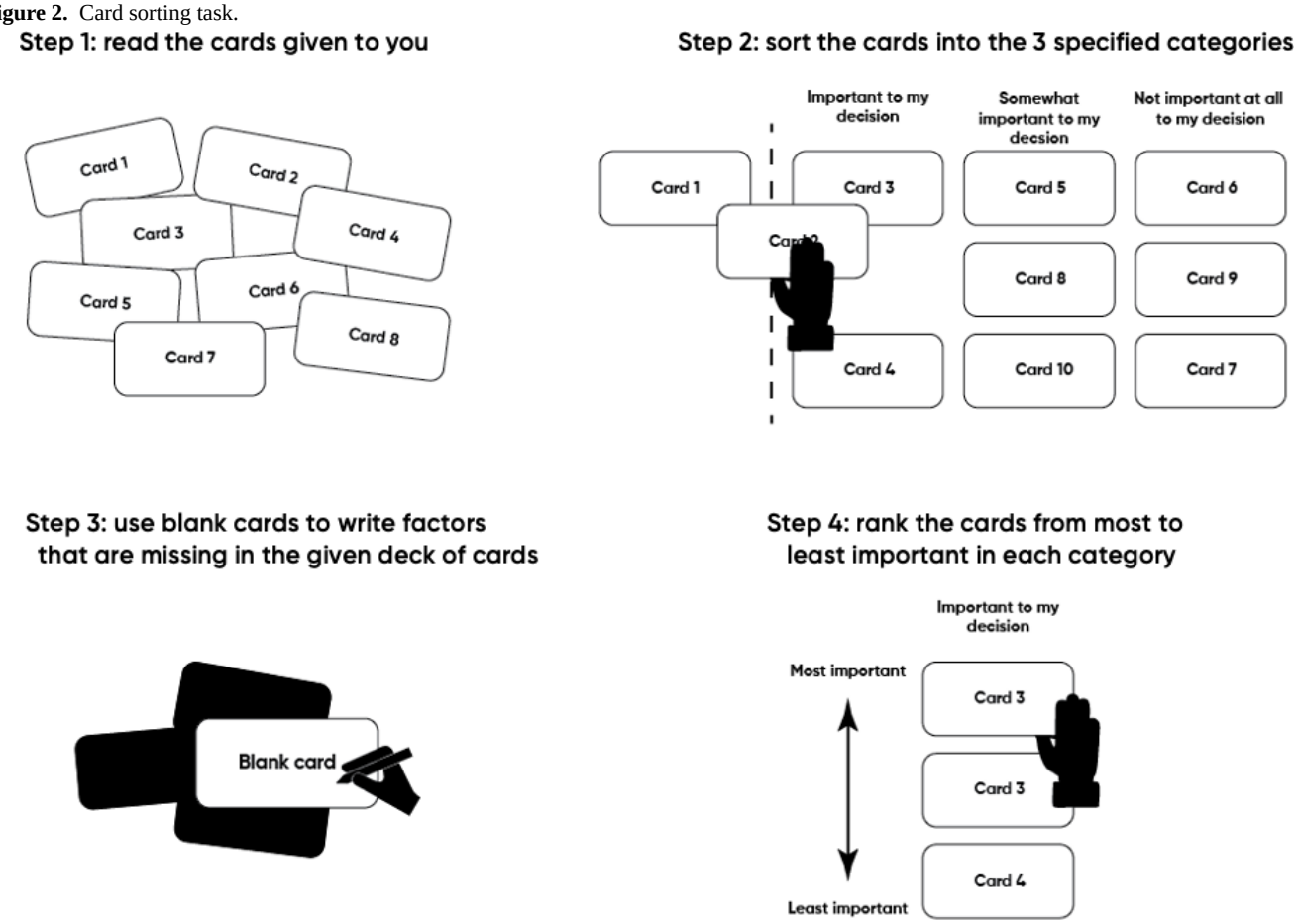

Following completion of the card sorting task, a semistructured interview will be conducted to explore the rationale behind their categorizations and rankings, allowing participants to reflect on their decision-making process. The entire session, including the card sorting task and semistructured interview, will be audio recorded for subsequent transcription and analysis, and final card categorizations and rankings will be photographed. Photographs will be taken using research institute iPads (Apple Inc), and no personal identifying information will be captured on audio recordings or photographs (eg, participants' hands and faces).

Patient Surveys

Baseline Survey

The baseline survey will assess demographic characteristics (ie, age, sex at birth, gender, educational level, relationship status, employment status, household income, country of birth, race, and religion), clinical characteristics (ie, depressive symptoms, mental well-being, and past treatment use), cognitive flexibility, decisional conflict, and health literacy.

Depression

The severity of depressive symptoms will be assessed via self-report using the 10-item version of the Center for Epidemiologic Studies Depression Scale. This brief version of the full 20-item scale has demonstrated good reliability in general (Cronbach α=0.85) and psychiatric (Cronbach α=0.90) samples [30,31]. Individuals rate the frequency of experiencing feelings such as worthlessness and low energy and sleep disturbances over the previous week on a 4-point Likert scale ranging from 0="rarely or none of the time" to 4="most of the time." Total scores range from 0 to 30, with scores greater than 10 indicating significant depressive symptoms.

Well-Being

Psychological well-being will be measured using the World Health Organization–Five Well-Being Index [32]. This 5-item scale asks individuals to rate how often, within the previous 2 weeks, they experienced positive mental health states (eg, "I have felt cheerful in good spirits") on a 6-point Likert scale ranging from 0="at no time" to 5="all of the time." Scores range from 0 to 100 and are calculated by summing the item responses and multiplying the total by 4. Scores below 50 are often used to screen for depression, whereas changes of 10 points across

XSL•FO
RenderX

administrations indicate clinically significant pretest-posttest treatment differences (ie, treatment sensitivity) [33].

Past Treatment Use

Previous depression treatments will be assessed via a self-report checklist developed for this study, based on American Psychological Association practice guidelines for treating MDD [34]. The checklist includes pharmacotherapy (eg, antidepressants), psychotherapy (eg, cognitive behavioral therapy, interpersonal psychotherapy, psychodynamic psychotherapy, problem-solving therapy, marital and family therapy, and group therapy), herbal supplements, light therapy, acupuncture, ECT, transcranial magnetic stimulation, and vagus nerve stimulation. Participants may also report other treatments not listed.

Cognitive Flexibility

Cognitive flexibility, operationally defined as the ability to consider multiple alternatives, adapt to different situations, and think flexibly, will be measured using the Cognitive Flexibility Scale [35]. This scale has demonstrated good internal consistency in general (Cronbach α=0.76-0.77) and clinical anxiety and depression (Cronbach α=0.81) samples [35,36]. Individuals rate their agreement with 12 items on a 6-point Likert scale ranging from 1="strongly disagree" to 6="strongly agree." Total scores are computed by summing responses, with higher scores indicating greater cognitive flexibility.

Decisional Conflict

Five dimensions of decision-making will be assessed using the 16-item Decisional Conflict Scale: feeling uncertain, uninformed, unclear about values, unsupported, and ineffective in decision-making. This scale has been widely used as an outcome measure for health decision interventions [37] and has shown good test-retest reliability (Cronbach α=0.78) [38]. Individuals rate their agreement with 16 statements about their treatment decision (eg, opting in or out of rTMS treatment) on a 5-point Likert scale ranging from 0="strongly disagree" to 4="strongly agree." Total scores (0-100) are calculated by summing responses, dividing by 16, and multiplying by 25, with lower scores indicating lower decisional conflict. Subscores for each dimension are calculated similarly by summing the respective subscale items, dividing by the number of items, and multiplying by 25.

Health Literacy

The Brief Health Literacy Screening Tool will be used to screen for comprehension and literacy challenges in understanding medical information. It assesses difficulty understanding hospital materials and written and verbal information and confidence in filling out medical forms on a 5-point Likert scale from 1="always" to 5="never." Total scores (4-20) categorize individuals as having inadequate (4-12), marginal (13-16), or adequate (17-20) comprehension. The Brief Health Literacy Screening Tool has shown a sensitivity of 0.79 for detecting inadequate skills and a moderate correlation with other validated literacy tools (r=0.40 with the Rapid Estimate of Adult Literacy in Medicine and r=0.42 with the Short Test of Functional Health Literacy in Adults), despite consisting of only 4 items [39].

Follow-Up Survey

The 6-month follow-up survey will use the same measures as the baseline survey (listed above), with the exception of demographics and past treatment use and the addition of a decisional regret measure.

Decisional Regret

The level of distress or remorse following the decision to opt in or out of rTMS treatment will be assessed using the Decision Regret Scale [40]. This validated 5-item measure evaluates regret associated with the effects of a past decision by having individuals reflect on their choice and rate their agreement with statements on a 5-point Likert scale ranging from 1="strongly agree" to 5="strongly disagree." Scores are summed and converted to a scale from 0 to 100 (subtracting 1 from each item and then multiplying by 25), with higher scores indicating greater regret. The tool has demonstrated good internal consistency across clinical samples (Cronbach α=0.81 - 0.92) and strong correlations with decision satisfaction (r=–0.40 to –0.60) and decisional conflict (r=0.31-0.52) [41].

Chart Review

To understand the clinical implications of patients' decision-making on rTMS treatments and how past medical treatment and depression history impact decision-making and vice versa, a chart review for each patient will be conducted. Electronic medical records, including consultation notes, progress notes, and patient self-reported outcomes, in Meditech and Greenspace will be accessed to complete this review. Variables of interest include (1) initial decision on rTMS referral, discussions with clinicians, and changes in decision; (2) rTMS treatment and adherence for those who opt in, such as number of appointments attended, number of cancellations or missed appointments, number of rTMS sessions, and length of rTMS sessions; (3) other ongoing treatment for depression, such as current antidepressant use and psychotherapy; and (4) other clinical details, such as hospital admissions and length of patient-provider relationship at HRH.

Sample Size

As the main aim of this study is to determine what factors are important for patients when making decisions about rTMS as a treatment option in a community setting, sample size calculations were determined based on the minimum number to conduct the card sorting analysis, while considering the feasibility of recruitment (ie, number of referrals based on a treatment capacity of 35 per year). Previous literature has identified that between 10 and 15 participants per group is optimal, with larger card sorting decks requiring larger sample sizes. A total of 24 participants, 12 per group (ie, deciding to opt in vs opt out of rTMS treatment), are required for analyzing the data from a 24-card deck. Previous experience from our research team indicates that a 20% attrition rate is expected at the 6-month follow-up in a sample of individuals with depression [42]. Therefore, a total of 30 participants will be recruited.

XSL•FO
RenderX

Data Analysis

Factors Influencing Patients' Decisions on rTMS Treatment at Baseline

Hierarchical agglomerative clustering will be conducted independently for each study cohort (ie, those who opt in vs out of rTMS treatments) to identify patterns in how participants prioritize the factors that influence their decision-making at baseline. The analysis will use 2 data points collected for each card in the 24-card deck plus any blank cards on which participants write personal factors: the category (ie, the pile that the card is placed in: important=1, neutral=2, and not important=3) and the rank of the card within the category (with 1 being the most important, followed by 2, 3, and so on). A maximum of 52 variables will be included in the cluster analysis assuming that each participant adds only 2 personalized factor cards. The analysis will use the squared Euclidean distance to calculate the distance between observations, and the Ward method will be applied to cluster card rankings and groupings, minimizing variance within clusters. The resulting dendrograms will be examined separately for those who choose to opt in vs opt out of rTMS treatments. The cophenetic correlation coefficient will be used to assess how well each dendrogram preserves the original squared Euclidean distances, as suggested in the literature [43]. To determine the optimal number of clusters for each group, the dendrogram will be cut at different levels, and silhouette scores will be calculated for k-cluster solutions. The number of clusters with the highest average silhouette score will be selected, and the final dendrogram will be presented for visualization.

Qualitative data from the card sorting baseline interviews will be analyzed to provide context and deeper insights into the clustering results. A CCA will be used to identify relationships between key factors identified as important, neutral, or not important for making decisions about rTMS treatment. Qualitative data will be mapped onto the cluster groupings to both interpret clusters and conceptually validate groupings. Specifically, participants' explanations, reasoning, and thought process during card grouping and prioritizations will be used to define clusters, refine cluster interpretations, and support the selection of the final number of clusters. While the analysis will be conducted using CCA methodology [44], it will be guided by 2 key frameworks—bounded rationality and prospect theory—rather than grounded theory [45]. Bounded rationality will inform the understanding of how participants make decisions within cognitive and informational limits [46], whereas prospect theory will provide insights into how patients weight potential gains and losses [47] from the reference point of being diagnosed with TRD. The analysis will follow a three-step process of analyzing transcripts, with memoing throughout the process providing further context: (1) open coding of each interview and internal comparisons to develop categories and understand the core decision-making process; (2) axial coding to examine relationships among codes and identify properties and dimensions within categories, with a focus on differences and similarities between those who opt in vs opt out of rTMS treatment; and (3) triangulation of data across interview groups to validate and refine themes and concepts that have been identified.

Decisional Changes, Conflict, and Regret

To examine changes in participants' priorities for decision-making on rTMS treatment, 2 analyses will be conducted. First, the hierarchical clustering analysis performed at baseline will be repeated using data from participants' 6-month card sorting task. This will allow for the visual comparison of dendrograms created for the baseline and 6-month card sorting activities, allowing for the assessment of how participants' prioritization of decision-making factors evolves over time. Differences in clustering patterns between baseline and the 6-month follow-up will be examined to determine whether shifts occur in the grouping of decision-making factors. Additionally, subgroup analyses will be conducted for those who opt in to rTMS to account for participants' treatment status (ie, whether they have started, completed, or are on a waitlist for rTMS treatment), ensuring that observed changes are contextualized within each individual's phase of treatment. Second, changes in participants' card rankings (ie, sorted by importance category and rank within each category) from baseline to the 6-month follow-up will be analyzed qualitatively using a longitudinal adaptation of the CCA. This approach will build on the baseline CCA by comparing thematic patterns across time points. Specifically, the three-step CCA process will be extended as follows: (1) open coding of 6-month interviews to identify emergent themes and changes in decision-making perspectives; (2) axial coding to examine relationships among new and preexisting codes, focusing on whether shifts occur within or between importance categories and rank orderings; and (3) triangulation across baseline and 6-month interview data to validate and refine evolving themes, ensuring consistency and reliability in interpretations of decisional change.

In addition to evaluating how factors influencing rTMS decisions change over time, group differences in individual-level variables between those who opt in vs out of rTMS will be examined. Exploratory analyses could include the use of chi-square tests and *t* tests to examine differences in demographic characteristics, clinical characteristics, and psychological constructs (ie, cognitive flexibility and decisional conflict) between those who opt in vs out of rTMS treatment. In addition, correlations between baseline cognitive flexibility and decisional conflict and decisional regret at 6 months will be explored. Finally, for those who opt in for rTMS treatment, changes in cognitive flexibility and decisional conflict from baseline to 6 months will be examined to understand how treatment impacts psychological constructs or vice versa. Due to the small sample size, exploratory quantitative analyses will primarily serve to guide future research hypotheses.

All quantitative statistical analyses will be conducted using SPSS (version 29.0; IBM Corp), and qualitative data analyses will be performed in NVivo (version 14; Lumivero).

Data Management and Governance

All collected data (ie, survey responses, interview transcripts, audio recordings and files, photographs of sorted cards, and

XSL•FO
RenderX

qualitative memos) will be stored in the hospital's secure OneDrive (ie, no data will be stored on local computers or outside of the institutionally secure OneDrive).

Separate (1) master linking logs containing patients' study IDs and personal identifying information and (2) contact datasets containing participant-provided and consented contact information for future research will be kept as password-protected files and stored on HRH's institutionally secure OneDrive. These files will be stored in different folders and will only be accessible to authorized research team members. Upon completion of the study, all master linking logs will be deleted. Contact datasets will only be retained for 7 years, after which they will be destroyed by the study team in accordance with the HRH data retention policy. Contact information will only be used for recontacting participants for future studies that have received research ethics board approval.

Results

This study was funded by a charitable donation from Jack and Pat Kay to Humber River Health, which is also supporting the establishment of the rTMS clinic, committed to in April 2024. Recruitment commenced in December 2025 and is expected to conclude in December 2026; no participants have been enrolled as of February 2026.

Discussion

This will be the first study, to our knowledge, that aims to explore factors influencing patient decision-making regarding rTMS treatment in a community hospital clinic. While previous research has focused on posttreatment perceptions, this study will prospectively examine patients' decision-making processes at the time of referral, providing a real-time perspective on how individuals evaluate rTMS before undergoing treatment. Additionally, by exploring patients' reflections on their decisions 6 months later, this study will also illustrate what factors contribute to decisional regret over time for patients with TRD. By considering both those who opt in for and those who decline rTMS treatment, this study will provide a comprehensive understanding of the motivations and barriers that shape treatment choices among individuals with TRD for unfamiliar treatment options such as rTMS.

Given the growing interest among health care providers in understanding the effectiveness and accessibility of rTMS at the community level, it is important to explore factors that drive patients' decision-making. Such insights could inform the development of patient-centered education and engagement strategies, which can help enhance both patient and public knowledge of rTMS. Furthermore, this study will contribute to existing models of health decision-making for depression by investigating the role of individual differences in cognitive flexibility, health literacy, and decision-making ability on treatment choice. The findings will inform patient-provider communication strategies, enhance shared decision-making models, and better equip clinicians to support patients in making informed decisions about rTMS.

A key strength of this study is the mixed methods design, which combines quantitative (ie, surveys and chart reviews) and qualitative (ie, card sorting interviews) data to capture a comprehensive view of what personal factors influence rTMS decision-making. This approach will allow us to have a more complete understanding of the facilitators of and barriers to rTMS treatment, potential associations between different factors, and how individual-level differences all play a role in the treatment decision-making process for rTMS. By examining decision-making at the time of referral and following patients over time, this study will offer novel insights into how patients evaluate rTMS initially and how their perceptions of it change over time. In addition, the focus on both those who opt in and out of rTMS provides a more balanced view of the decision-making process for patients overall.

Despite the strengths and novelty of this study, there are a few limitations to consider. First, as rTMS is not publicly funded across Canada and public awareness of the treatment remains low, patient decisions may be strongly influenced by the amount and type of information provided about rTMS by the referring physician or clinic rather than by independent preexisting attitudes. Second, this study is being conducted in a newly established clinic, which may introduce variability into patient experience. For example, factors such as evolving clinic procedures, staff expertise, and ongoing improvements to the clinic workflow and procedures could influence perceptions and decision-making. Finally, while the study focuses on patient-level factors, broader health care system–level influences such as referral pathways, institutional policies, and other system-level factors that can influence patient decision-making indirectly will not be examined.

Acknowledgments

The authors attest that there was no use of generative artificial intelligence technology in the generation of text, figures, or other informational content of this manuscript.

Funding

The Jack and Pat Kay charitable donation to Humber River Health, which is supporting the establishment of the rTMS clinic, is also being used to support this research study.

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the sensitive and identifiable nature of the interview transcript data but are available from the corresponding author on reasonable request and pending institutional ethics approval.

Authors' Contributions

Conceptualization: AG, SD, SS, DT, PW
Funding acquisition: SS, DT
Methodology: AG, SD, SS, DT, PW
Project administration: AG, SD
Resources: SS, DT
Supervision: AG, SD, PW
Validation: AG, SD, PW
Visualization: AG, SD
Writing—original draft: AG, SD
Writing—review and editing: AG, SD, SS, DT, PW

Conflicts of Interest

None declared.

References

1. Howes OD, Thase ME, Pillinger T. Treatment resistance in psychiatry: state of the art and new directions. Mol Psychiatry 2022 Jan;27(1):58-72. [doi: 10.1038/s41380-021-01200-3] [Medline: 34257409]
2. Rizvi SJ, Grima E, Tan M, et al. Treatment-resistant depression in primary care across Canada. Can J Psychiatry 2014 Jul;59(7):349-357. [doi: 10.1177/070674371405900702] [Medline: 25007419]
3. McIntyre RS, Alsuwaidan M, Baune BT, et al. Treatment-resistant depression: definition, prevalence, detection, management, and investigational interventions. World Psychiatry 2023 Oct;22(3):394-412. [doi: 10.1002/wps.21120] [Medline: 37713549]
4. Johnston KM, Powell LC, Anderson IM, Szabo S, Cline S. The burden of treatment-resistant depression: a systematic review of the economic and quality of life literature. J Affect Disord 2019 Jan 1;242:195-210. [doi: 10.1016/j.jad.2018.06.045] [Medline: 30195173]
5. Lundberg J, Cars T, Lööv S, et al. Association of treatment-resistant depression with patient outcomes and health care resource utilization in a population-wide study. JAMA Psychiatry 2023 Feb 1;80(2):167-175. [doi: 10.1001/jamapsychiatry.2022.3860] [Medline: 36515938]
6. McIntyre RS, Millson B, Power GS. Burden of Treatment Resistant Depression (TRD) in patients with major depressive disorder in Ontario using Institute for Clinical Evaluative Sciences (ICES) databases: economic burden and healthcare resource utilization. J Affect Disord 2020 Dec 1;277:30-38. [doi: 10.1016/j.jad.2020.07.045] [Medline: 32791390]
7. Gaynes BN, Lloyd SW, Lux L, et al. Repetitive transcranial magnetic stimulation for treatment-resistant depression: a systematic review and meta-analysis. J Clin Psychiatry 2014 May;75(5):477-489. [doi: 10.4088/JCP.13r08815] [Medline: 24922485]
8. Vida RG, Sághy E, Bella R, et al. Efficacy of repetitive transcranial magnetic stimulation (rTMS) adjunctive therapy for major depressive disorder (MDD) after two antidepressant treatment failures: meta-analysis of randomized sham-controlled trials. BMC Psychiatry 2023 Jul 27;23(1):545. [doi: 10.1186/s12888-023-05033-y] [Medline: 37501135]
9. Blumberger DM, Vila-Rodriguez F, Wang W, et al. A randomized sham controlled comparison of once vs twice-daily intermittent theta burst stimulation in depression: a Canadian rTMS treatment and biomarker network in depression (CARTBIND) study. Brain Stimul 2021;14(6):1447-1455. [doi: 10.1016/j.brs.2021.09.003] [Medline: 34560319]
10. Kishi T, Sakuma K, Matsuda Y, Kito S, Iwata N. Intermittent theta burst stimulation vs. high-frequency repetitive transcranial magnetic stimulation for major depressive disorder: a systematic review and meta-analysis. Psychiatry Res 2023 Oct;328:115452. [doi: 10.1016/j.psychres.2023.115452] [Medline: 37657200]
11. Basiru T, Onyeaka H, Oladunjoye AF, et al. Trend and geo-availability of somatic therapies for treatment resistant depression in the US. Psychiatry Res Commun 2024 Mar;4(1):100157. [doi: 10.1016/j.psycom.2024.100157]
12. Stillianesis G, Cavaleri R, Tang CY, Summers SJ. Exploring patient perceptions of noninvasive brain stimulation: a systematic review. Neuromodulation 2022 Jun;25(4):487-493. [doi: 10.1111/ner.13461] [Medline: 35667767]
13. Mayer G, Faivel N, Aviram S, Walter G, Bloch Y. Repetitive transcranial magnetic stimulation in depressed adolescents: experience, knowledge, and attitudes of recipients and their parents. J ECT 2012 Jun;28(2):104-107. [doi: 10.1097/YCT.0b013e318250058c] [Medline: 22513510]
14. Walter G, Martin J, Kirkby K, Pridmore S. Transcranial magnetic stimulation: experience, knowledge and attitudes of recipients. Aust N Z J Psychiatry 2001 Feb;35(1):58-61. [doi: 10.1046/j.1440-1614.2001.00852.x] [Medline: 11270457]

XSL•FO
RenderX

15. Singh SM, Sharma M, Aggarwal A, Avasthi A. The knowledge, experience and attitudes of recipients of repetitive transcranial magnetic stimulation: a study from North India. Asian J Psychiatr 2018 Jan;31:102-106. [doi: 10.1016/j.ajp.2018.01.014] [Medline: 29459341]
16. Lonergan B, Nguyen E, Lembo C, et al. Patient- and technician-oriented attitudes toward transcranial magnetic stimulation devices. J Neuropsychiatry Clin Neurosci 2018;30(3):242-245. [doi: 10.1176/appi.neuropsych.17070146] [Medline: 29366375]
17. Morrison AS, Uusberg A, Ryan J, Goldenberg A, Etkin A, Gross JJ. Assessing and improving public mental health literacy concerning rTMS. BMC Psychiatry 2022 Apr 8;22(1):249. [doi: 10.1186/s12888-022-03880-9] [Medline: 35395758]
18. Rosenquist PB, Dunn A, Rapp S, Gaba A, McCall WV. What predicts patients' expressed likelihood of choosing electroconvulsive therapy as a future treatment option? J ECT 2006 Mar;22(1):33-37. [doi: 10.1097/00124509-200603000-00007] [Medline: 16633204]
19. Smith M, Vogler J, Zarrouf F, Sheaves C, Jesse J. Electroconvulsive therapy: the struggles in the decision-making process and the aftermath of treatment. Issues Ment Health Nurs 2009 Sep;30(9):554-559. [doi: 10.1080/01612840902807947] [Medline: 19657869]
20. Amazon J, McNeely E, Lehr S, Marquardt MG. The decision making process of older adults who elect to receive ECT. J Psychosoc Nurs Ment Health Serv 2008 May;46(5):45-52. [doi: 10.3928/02793695-20080501-01] [Medline: 18548870]
21. Wells K, Scanlan JN, Gomez L, et al. Decision making and support available to individuals considering and undertaking electroconvulsive therapy (ECT): a qualitative, consumer-led study. BMC Psychiatry 2018 Jul 24;18(1):236. [doi: 10.1186/s12888-018-1813-9] [Medline: 30041602]
22. Conrad LY, Tucker VM. Making it tangible: hybrid card sorting within qualitative interviews. J Doc 2019 Feb 19;75(2):397-416. [doi: 10.1108/JD-06-2018-0091]
23. Pluye P, Loignon C, Lagarde F, et al. Assessing and improving the use of online information about child development, education, health, and well-being in low-education, low-income parents: protocol for a mixed-methods multiphase study. JMIR Res Protoc 2018 Nov 20;7(11):e186. [doi: 10.2196/resprot.9996] [Medline: 30459143]
24. Laurie H, Lynn P. The use of respondent incentives on longitudinal surveys. In: Lynn P, editor. Methodology of Longitudinal Surveys: John Wiley & Sons; 2009:205-233. [doi: 10.1002/9780470743874]
25. Cunningham JA, Hendershot CS, Kay-Lambkin F, et al. Does providing a brief internet intervention for hazardous alcohol use to people seeking online help for depression reduce both alcohol use and depression symptoms among participants with these co-occurring disorders? Study protocol for a randomised controlled trial. BMJ Open 2018 Jul 19;8(7):e022412. [doi: 10.1136/bmjopen-2018-022412] [Medline: 30030322]
26. Fair & inclusive compensation for research participants: a guideline. Wellesley Institute. 2018. URL: https://www.wellesleyinstitute.com/wp-content/uploads/2018/07/Fair-compensation-Guideline-.pdf [accessed 2025-02-04]
27. Conrad LY, Demasson A, Gorichanaz T, VanScoy A. Exploring card sort methods: interaction and implementation for research, education, and practice. Proc Assoc Inf Sci Technol 2019;56(1):525-528. [doi: 10.1002/pra2.81]
28. Eccles DW, Arsal G. The think aloud method: what is it and how do I use it? Qual Res Sport Exerc Health 2017;9(4):514-531. [doi: 10.1080/2159676X.2017.1331501]
29. Factors influencing depression patient's choice for mental health treatment in real-world settings: a scoping review. OSF. URL: https://osf.io/gka29 [accessed 2025-02-05]
30. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994;10(2):77-84. [Medline: 8037935]
31. Björgvinsson T, Kertz SJ, Bigda-Peyton JS, McCoy KL, Aderka IM. Psychometric properties of the CES-D-10 in a psychiatric sample. Assessment 2013 Aug;20(4):429-436. [doi: 10.1177/1073191113481998] [Medline: 23513010]
32. Wellbeing measures in primary health care/the DepCare project: report on a WHO meeting. : World Health Organization; 1998 URL: https://iris.who.int/server/api/core/bitstreams/8af98105-30ec-4da4-9fe9-097f7459e6da/content [accessed 2026-02-22]
33. Topp CW, Østergaard SD, Søndergaard S, Bech P. The WHO-5 Well-Being Index: a systematic review of the literature. Psychother Psychosom 2015;84(3):167-176. [doi: 10.1159/000376585] [Medline: 25831962]
34. Practice guideline for the treatment of patients with major depressive disorder (revision). American Psychiatric Association. Am J Psychiatry 2000 Apr;157(4 Suppl):1-45. [Medline: 10767867]
35. Martin MM, Rubin RB. A new measure of cognitive flexibility. Psychol Rep 1995;76(2):623-626. [doi: 10.2466/pr0.1995.76.2.623]
36. Martin MM, Anderson CM. The cognitive flexibility scale: three validity studies. Commun Rep 1998;11(1):1-9. [doi: 10.1080/08934219809367680]
37. Garvelink MM, Boland L, Klein K, et al. Decisional conflict scale use over 20 years: the anniversary review. Med Decis Making 2019 May;39(4):301-314. [doi: 10.1177/0272989X19851345] [Medline: 31142194]
38. User manual - Decisional Conflict Scale. Ottawa Hospital Research Institute. 1993. URL: https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_decisional_conflict.pdf [accessed 2026-02-22]

XSL•FO
RenderX

39. Haun J, Noland-Dodd V, Varnes J, Graham-Pole J, Rienzo B, Donaldson P. Testing the BRIEF health literacy screening tool health literacy. : MDEdge; 2017 URL: https://cdn.mdedge.com/files/s3fs-public/Document/September-2017/026120024.pdf [accessed 2026-02-22]
40. O'Connor AM. User manual - Decision Regret Scale. Ottawa Hospital Research Institute. 1996. URL: https://decisionaid.ohri.ca/docs/develop/User_manuals/UM_Regret_Scale.pdf [accessed 2026-02-22]
41. Brehaut JC, O'Connor AM, Wood TJ, et al. Validation of a decision regret scale. Med Decis Making 2003;23(4):281-292. [doi: 10.1177/0272989X03256005] [Medline: 12926578]
42. Cunningham JA, Godinho A, Hendershot CS, et al. Randomized controlled trial of online interventions for co-occurring depression and hazardous alcohol consumption: primary outcome results. Internet Interv 2021 Nov 2;26:100477. [doi: 10.1016/j.invent.2021.100477] [Medline: 34786350]
43. Macías JA. Enhancing card sorting dendrograms through the holistic analysis of distance methods and linkage criteria. J User Exp 2021;16(2):73-90 [FREE Full text]
44. Boeije H. A purposeful approach to the constant comparative method in the analysis of qualitative interviews. Qual Quant 2002;36:391-409. [doi: 10.1023/A:1020909529486]
45. Fram SM. The constant comparative analysis method outside of grounded theory. Qual Rep 2013;18(1):1-25. [doi: 10.46743/2160-3715/2013.1569]
46. Simon HA. Bounded rationality. In: Eatwell J, Milgate M, Newman P, editors. Utility and Probability: Palgrave Macmillan; 1990:15-18. [doi: 10.1007/978-1-349-20568-4_5]
47. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica 1979 Mar;47(2):263. [doi: 10.2307/1914185]

Abbreviations

CCA: constant comparative analysis
CONSORT: Consolidated Standards of Reporting Trials
ECT: electroconvulsive therapy
HRH: Humber River Health
MDD: major depressive disorder
rTMS: repetitive transcranial magnetic stimulation
TRD: treatment-resistant depression

Edited by J Sarvestan; submitted 15.Aug.2025; peer-reviewed by A AL-Asadi; revised version received 08.Dec.2025; accepted 09.Dec.2025; published 06.Mar.2026.

Please cite as:
Godinho A, Dalmia S, Selchen S, Tziatis D, Wegier P
Factors That Influence Patients' Decisions About Repetitive Transcranial Magnetic Stimulation as a Treatment Option for Treatment-Resistant Depression: Protocol for a Prospective Mixed Methods Cohort Study
JMIR Res Protoc 2026;15:e82491
URL: https://www.researchprotocols.org/2026/1/e82491
doi:10.2196/82491

© Alexandra Godinho, Sanvitti Dalmia, Steven Selchen, Daniel Tziatis, Pete Wegier. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Disease Progression and Outcomes in Patients With Benign Prostatic Hyperplasia: Protocol for a Multicenter Retrospective Cohort Study

Zexuan Lv[1*], MM; Jiaxing Liu[2*], BA; Yin Lu[1], MM; Xu Zhang[1], MD, PhD; Fan Gao[1], MM; Bingyang Guo[1], MM; Hongyu Zhang[1], MM; Hongzhao Li[1], MD; Chao Lv[1,3], MD

[1]Senior Department of Urology, The Third Medical Center of PLA General Hospital, 69 Yongding Road, Haidian District, Beijing, China

[2]School of Medicine, Nankai University, Tianjin, China

[3]Air Force Special Medical Center, Beijing, China

[*]these authors contributed equally

Corresponding Author:
Chao Lv, MD
Senior Department of Urology, The Third Medical Center of PLA General Hospital, 69 Yongding Road, Haidian District, Beijing, China

Abstract

Background: Benign prostatic hyperplasia (BPH) is highly prevalent among aging men and may lead to progressive lower urinary tract symptoms, surgical intervention, and serious outcomes such as acute urinary retention and renal impairment. In routine clinical practice, the long-term spectrum, timing, and determinants of multiple key outcome events after BPH diagnosis remain insufficiently characterized. This evidence gap is particularly relevant in China, where care pathways, follow-up patterns, and comorbidity profiles may differ from those reported in other populations. A comprehensive real-world description of long-term outcome patterns and determinants is needed to inform risk-stratified follow-up intensity and intervention timing, where current evidence remains limited.

Objective: This study aims to characterize BPH progression trajectories and associated risk factors, quantify long-term risks and time-to-event profiles of major clinical outcomes, and develop and validate prognostic models using multicenter retrospective data from the People's Liberation Army General Hospital consortium.

Methods: We will establish a multicenter retrospective cohort using routinely collected clinical data from the People's Liberation Army General Hospital consortium (the First, Third, Fourth, Fifth, Sixth, Seventh, and Eighth Medical Centers) spanning 2001 to 2021. The index date is defined as the first recorded BPH diagnosis, and we will include patients with confirmed BPH for whom long-term follow-up data are available. To approximate incident (first-diagnosis) BPH, we applied a 2-year washout period before the index date and cleaned the cohort during assembly by excluding patients with any BPH-related records within this window. Prespecified outcomes include BPH-related surgery, urinary retention, inguinal hernia, chronic kidney disease, urothelial carcinoma, and all-cause mortality. Cox proportional hazard regression will be applied for all-cause mortality, and Fine-Gray competing risk models will be used for nonmortality end points where death may act as a competing event. Model development and validation will be performed across centers, with performance assessed using discrimination, calibration, and clinical utility metrics.

Results: By mid-January 2026, data collection for the multicenter retrospective dataset (2001 - 2021) had been completed, and data cleaning and harmonization had commenced. As this is a retrospective electronic health record–based study, no participant recruitment is involved; the final analytic sample size will be confirmed after harmonization and cleaning. The baseline characteristic table for cohort 1 (First Medical Center) has been completed as part of data quality control. Completion of data harmonization and the prespecified analyses (including model development and validation) is anticipated by May 2026, followed by dissemination and translation to support routine clinical decision-making.

Conclusions: This protocol describes a large, multicenter, real-world cohort study designed to provide robust long-term evidence on BPH progression and clinically important outcomes in China and establish interpretable and validated prognostic models to inform risk-stratified follow-up and intervention decisions.

International Registered Report Identifier (IRRID): PRR1-10.2196/84965

(JMIR Res Protoc 2026;15:e84965) doi:10.2196/84965

XSL•FO
RenderX

KEYWORDS

benign prostatic hyperplasia; BPH; Chinese patients; disease progression; clinical outcomes; multicenter retrospective cohort study

Introduction

Benign prostatic hyperplasia (BPH) is one of the most common urological diseases in older men, and its prevalence increases markedly with age. Epidemiological data indicate that prevalence exceeds 50% among men aged >50 years and may approach 80% in those aged >80 years [1-3]. Clinically, BPH predominantly presents with lower urinary tract symptoms (LUTS) during the storage, voiding, and terminal phases; imaging typically demonstrates prostatic enlargement, and pathological findings are characterized by glandular and stromal hyperplasia [4-7]. With disease progression, bladder outlet obstruction may worsen and lead to dysfunction of the bladder and upper urinary tract [8,9]. Notably, BPH progression is often insidious and heterogeneous: some patients experience rapid deterioration of LUTS, whereas others develop detrusor dysfunction due to long-standing obstruction, resulting in complications such as overactive bladder, urinary retention, and recurrent urinary tract infections [10-14]. These complications not only impair quality of life but also impose substantial individual and societal burdens; however, in routine clinical practice, the long-term spectrum, timing, and determinants of different outcome events following a diagnosis of BPH have not been systematically characterized, which limits refined decision-making regarding follow-up intensity and the optimal timing of intervention [15-18].

In other countries and populations, cohort studies and clinical trials have explored progression-related outcomes and associated factors in BPH and LUTS, suggesting that age, prostate volume and related indicators, baseline symptom burden, and certain metabolic comorbidities may be associated with the risks of acute urinary retention and surgical intervention [19-23]. Evidence also supports that different treatment strategies improve symptoms and, in some settings, reduce the risk of progression-related outcomes [24-26]. Nevertheless, these findings are influenced by differences in population sources, follow-up duration, and care pathways. Many studies have relatively limited follow-up, enroll patients from single centers or under restrictive eligibility criteria, and apply nonuniform outcome definitions and assessment frameworks, which complicates estimation of long-term risks and comparison across studies [27-30]. In addition, prior analyses often focus on a single outcome or short-term end points, and fewer studies have systematically evaluated multiple key outcomes within a unified long-term framework [31]. More importantly, although individualized risk prediction tools have been proposed, limitations remain in external or multicenter validation, coverage of comorbidity profiles and low-incidence outcomes, and the handling of methodological issues such as competing risks, thereby constraining reliability and generalizability for decision support in routine practice [29,32]. Therefore, real-world studies with long-term follow-up that incorporate multiple clinically important outcomes and apply rigorous methodology are needed to strengthen the evidence chain from risk factor identification to risk stratification and management decisions [29,33,34].

Overall, existing evidence on BPH remains insufficient for a systematic assessment of disease progression, prognostic determinants, and long-term clinical outcomes, and large-scale longitudinal cohorts spanning the full "progression-outcomes-prediction" continuum are scarce [19,29,35]. In China, this evidence gap is more pronounced; rapid population aging, regional variation in health care accessibility and care pathways, and differences in real-world follow-up and treatment patterns may influence outcomes and management effectiveness, creating uncertainty regarding direct extrapolation of evidence from other countries [33,36,37]. Therefore, this study aims to leverage long-term multicenter follow-up data from the People's Liberation Army (PLA) General Hospital consortium to quantify long-term risks of key outcomes and develop and validate interpretable prediction models to inform follow-up and intervention decisions for BPH management in China.

Methods

Study Population and Data Source

This multicenter retrospective cohort study will use routinely collected clinical data from the PLA General Hospital consortium (the First, Third, Fourth, Fifth, Sixth, Seventh, and Eighth Medical Centers) spanning 2001 to 2021. Patients will be divided into 2 cohorts: cohort 1 (First Medical Center) and cohort 2 (the remaining medical centers). Eligible patients include those with a confirmed diagnosis of BPH and long-term follow-up, with the date of first diagnosis serving as the baseline. Patients with prostate cancer, urethral stricture, neurogenic bladder, or missing key clinical data will be excluded.

Routinely collected data will be extracted from each center's electronic health record (EHR) system, including outpatient, inpatient, and health examination records. Diagnoses and end point events will be identified using prespecified *International Classification of Diseases, 9th and 10th Revision*, codes and procedure code lists (Table 1), and relevant laboratory and clinical measurements will be retrieved from structured records. Deidentified datasets will then be transferred to the coordinating center for harmonization before analysis (Figure 1).

XSL•FO
RenderX

Table . *International Classification of Diseases, 10th Revision* (*ICD-10*), and *International Classification of Diseases, 9th Revision* (*ICD-9*), diagnostic codes for benign prostatic hyperplasia (BPH)–related comorbidities and BPH-related outcome events.[a]

	ICD-10	ICD-9
Outcomes		
Surgical treatment for BPH (any type)	Codes beginning with 0VT0, 0VT5, 0VT7, and 0VT8 (resection, incision, destruction, or enucleation of prostate via open or transurethral approach)	60.21, 60.29, 60.3 - 60.5, 60.69, and 60.97
AUR[b]	R33.x	788.2
Inguinal hernia	K40.x	550.x
CKD[c]	N18.x	585.x
Urothelial carcinoma	C65, C66, and C67.x	188.x, 189.0, and 189.1
All-cause mortality	Not *ICD*[d] based	Not *ICD* based
Baseline diseases		
BPH	N40.0 and N40.1	600.00 and 600.01
Urothelial carcinoma (renal pelvis, ureter, and bladder)	C65, C66, and C67.x	188.x, 189.0, and 189.1
Prostate cancer	C61	185
Neurogenic bladder	N31.x	344.61
Urethral stricture	N35.x	598.x
History of pelvic radiotherapy	Z92.3	V15.3
Hypertension	I10-I15	401.x
Diabetes mellitus	E10-E14	250.x
Coronary artery disease	I20-I25	410-414
Psychiatric disorders (any)	F01-F99	290-319
Insomnia (any type)	F51.0 and G47.0	780.52 and 307.42
COPD[e]	J44.x	491 - 492 and 496
Asthma	J45.x	493
Chronic sinusitis	J32.x	473
Constipation	K59.0	564

[a]Surgical treatment was defined as any open or transurethral procedure performed for BPH, including transurethral resection, transurethral incision, laser vaporization or enucleation (eg, holmium laser enucleation of the prostate), and open simple prostatectomy. *ICD-10 Procedure Coding System* codes were identified by procedure root operations corresponding to resection, incision, destruction, or excision of the prostate (body part value: prostate). Radical prostatectomy and procedures performed for prostate malignancy (*ICD-10* C61 or *ICD-9* 185) were excluded.

[b]AUR: acute urinary retention.

[c]CKD: chronic kidney disease.

[d]*ICD*: *International Classification of Diseases*.

[e]COPD: chronic obstructive pulmonary disease.

XSL•FO
RenderX

Figure 1. Study workflow, data sources, eligibility criteria, outcomes, follow-up, and analysis plan. The figure summarizes the data sources, inclusion and exclusion criteria and index date definition, prespecified outcomes, follow-up and censoring rules, and the overall development and validation workflow across cohorts. AUC: area under the curve; AUR: acute urinary retention; BPH: benign prostatic hyperplasia; CHARLS: China Health and Retirement Longitudinal Study; CKD: chronic kidney disease; DCA: decision curve analysis; IBS: integrated Brier score; LASSO: least absolute shrinkage and selection operator; PLAGH: People's Liberation Army General Hospital.

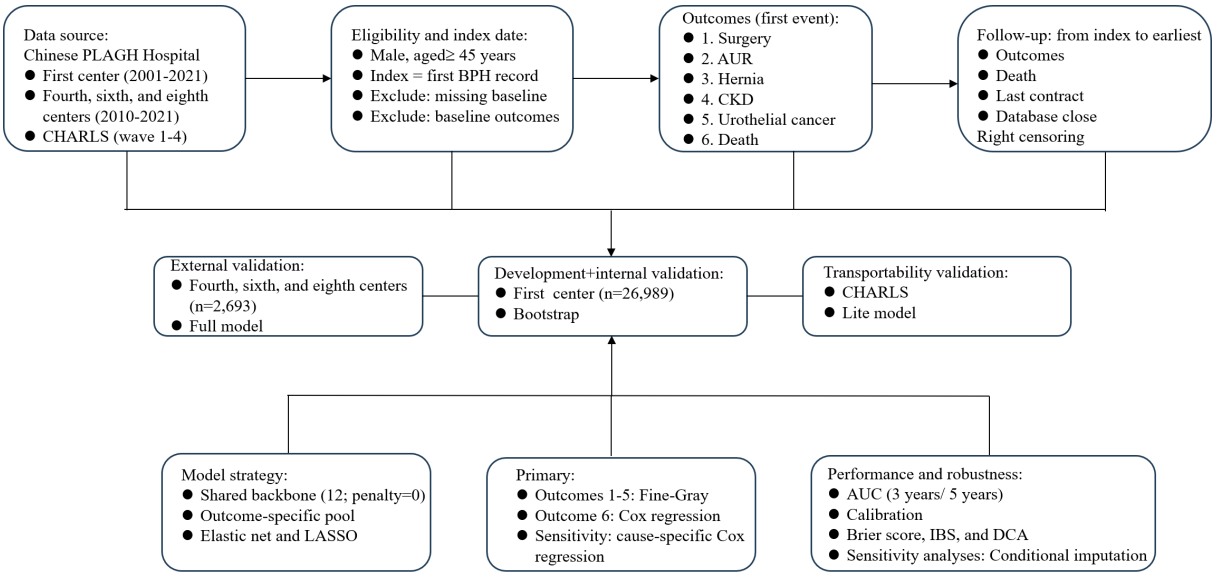

To evaluate model transportability beyond hospital-based EHRs, we will use the China Health and Retirement Longitudinal Study (CHARLS), a nationally representative longitudinal cohort, as an independent external validation dataset. Variables and end points will be harmonized between the PLA General Hospital EHR cohort and the CHARLS using prespecified definitions; external validation will be conducted for predictors and end points that can be consistently operationalized in the CHARLS.

Outcomes

The outcomes of this study are (1) BPH-related surgery (transurethral resection of the prostate and bipolar or laser procedures), (2) urinary retention, (3) inguinal hernia, (4) chronic kidney disease (CKD; estimated glomerular filtration rate <60 mL per minute per 1.73 m^2 persisting for ≥90 days), (5) urothelial carcinoma, and (6) all-cause mortality (Table 1).

Eligibility Criteria

The inclusion and exclusion criteria are outlined in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Cohort construction period: January 1, 2001, to December 31, 2021.
- Benign prostatic hyperplasia (BPH) definition: BPH will be identified by either an International Prostate Symptom Score of ≥7 or a clinician-confirmed diagnosis based on digital rectal examination, imaging (ultrasound or magnetic resonance imaging), and/or pathological examination.
- Data completeness: patients must have sufficiently complete medical records for the prespecified analyses.
- Washout requirement: to approximate incident (first-diagnosis) BPH, patients must have no in-hospital BPH diagnosis record, no α-blocker or 5α-reductase inhibitor prescription, and no history of BPH-related surgery within 2 years before cohort entry.

Exclusion criteria

- Competing or urologic conditions: diagnosis of prostate cancer, urethral stricture, neurogenic bladder, or a history of pelvic radiotherapy before cohort entry or within 1 year after cohort entry.
- Prior outcome events: any study outcome event occurring at or before cohort entry.

XSL•FO
RenderX

Baseline Variables for Analysis

Baseline variables used for outcome analyses and risk stratification modeling will include residence, age (years), BMI (kg/m^2), smoking status (yes or no), alcohol consumption (yes or no), comorbidities (hypertension, diabetes mellitus, coronary artery disease, psychiatric disorders, sleep disorders, respiratory diseases, and constipation), and laboratory or clinical measures (serum creatinine, serum uric acid, triglyceride-to–high-density lipoprotein cholesterol ratio, neutrophil-to-lymphocyte ratio, alanine aminotransferase, D-dimer, total prostate-specific antigen, free prostate-specific antigen, and prostate volume).

Data Security and Quality Assurance

Rigorous measures will be implemented to guarantee data quality, integrity, confidentiality, and statistical validity:

Data Quality

A multilevel verification process will be established, with clinical physicians and data managers jointly reviewing patient eligibility, baseline characteristics, and end point events. Cases not meeting criteria will be excluded to ensure authenticity.

Data Integrity

Baseline variables such as BMI, serum creatinine, and prostate volume will be defined as the mean values recorded at the first visit or within 6 months before or after. Missing data will be assessed using statistical methods to determine missingness patterns; only variables with low and nonbiased missing rates will be retained to avoid compromising validity.

Data Confidentiality

All data will be encrypted and deidentified before analysis, with personally identifiable information (eg, names and ID numbers) removed; data will be stored on secure hospital servers with access restricted to authorized team members and transferred via encrypted channels.

Statistical Oversight

Professional statisticians will oversee the entire process, ensuring appropriate application of Cox regression and Fine-Gray competing risk models as well as standardized procedures for Kaplan-Meier curve generation, cumulative incidence, and median time-to-event calculations. All analytic steps will be documented in a version-controlled workflow to support auditability and reproducibility.

Data Traceability (Audit Trail)

To ensure traceability across the workflow (data extraction, transfer, cleaning, and analysis), each dataset will be labeled with a center code and extraction time stamp, and raw files will be archived in a read-only format. All data transformations (eg, recoding, outlier handling, and missing data processing) will be recorded in a versioned data processing log, with pseudonymized unique study identifiers to track records across processing steps. Access to raw and processed datasets will be role based and logged.

Sample Size Estimation

On the basis of Cox regression and Fine-Gray competing risk models, at least 300 to 490 events are required to analyze approximately 25 covariates [38-40]. Assuming a 10% incidence of surgery at 5 years, approximately 3334 patients will be needed. For rarer outcomes such as CKD (3% incidence), approximately 10,000 patients are required. The final sample size will be refined according to empirical event rates [40].

If, after data cleaning and cohort assembly, the effective sample size or event numbers for any prespecified end point are insufficient, we will implement the following prespecified strategies: (1) prioritization of end points with adequate events for multivariable modeling while reporting end points with sparse events descriptively and/or as exploratory analyses; (2) simplification of model specification by restricting covariates to a clinically justified core set, combining sparse categories when appropriate and using shrinkage or penalized methods to improve stability and minimize overfitting; and (3) use of alternative time-to-event approaches where needed, including penalized or stratified Cox models or time-varying effects for Cox analyses and cause-specific hazard models (or cumulative incidence estimation without multivariable regression) as alternatives when Fine-Gray models are unstable due to sparse subdistribution events. Any deviations from the initial plan and their rationale will be transparently reported.

Statistical Analysis

To address the study objectives, analyses will comprise (1) long-term outcome profiling, (2) association analyses, and (3) prognostic risk modeling. Time-to-event methods will be used to summarize incidence, cumulative incidence, and median time to event, as appropriate. For association analyses, Cox proportional hazard regression will be applied for all-cause mortality, whereas Fine-Gray competing risk models will be used for nonmortality end points where death may act as a competing event, reporting hazard ratios or subdistribution hazard ratios with 95% CIs. Baseline variables for analysis and modeling are specified in the Baseline Variables for Analysis section. For prognostic risk modeling, multivariable prediction models will be developed and validated to support risk stratification, with performance assessed using discrimination (eg, concordance index or area under the curve), calibration, and clinical utility metrics (eg, decision curve analysis), as applicable.

Cox proportional hazard regression is appropriate for this study because the primary end points are time-to-event outcomes with right censoring and the Cox model provides an efficient and interpretable framework to account for censoring while modeling risk over follow-up. In this protocol, Cox regression is used within a prognostic (risk stratification) framework rather than solely for etiologic inference. Cox models will be applied to all-cause mortality (event 6). For nonmortality end points where death may act as a competing event, Fine-Gray competing risk models will be used to support risk estimation based on cumulative incidence.

If supported by the empirical data structure and transition processes (eg, sufficient event counts and identifiable transition pathways), we will further explore and construct multistate models to characterize dynamic transitions between clinical states (Figure 2).

XSL•FO
RenderX

Figure 2. Schematic diagram of the multistate model. The final state definitions and transition pathways to be modeled will be determined based on data availability and observed event counts. AUR: acute urinary retention; BPH: benign prostatic hyperplasia; CKD: chronic kidney disease.

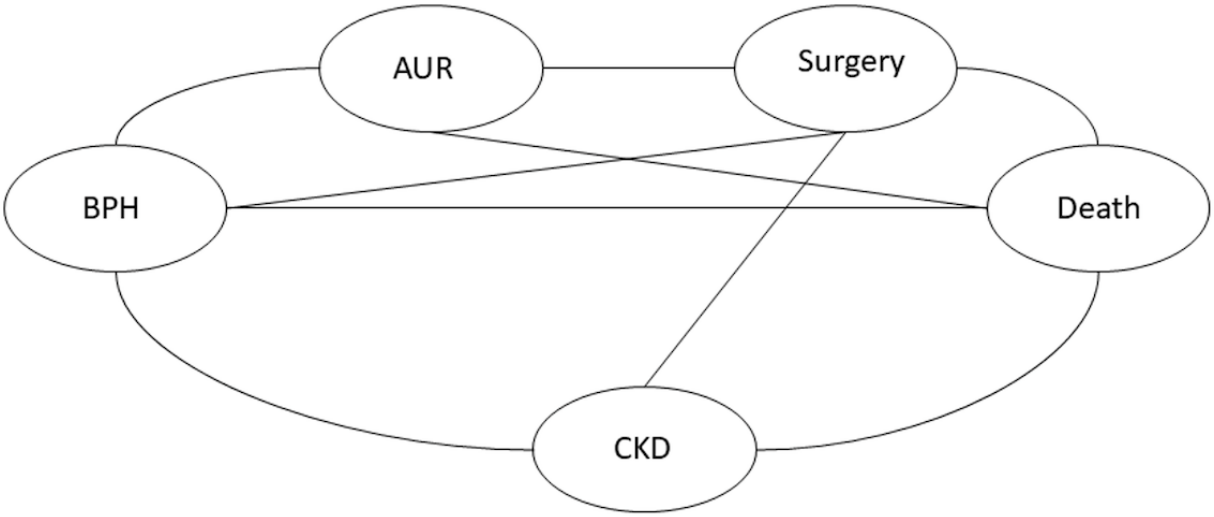

Standardized case report forms will be extracted by trained staff at each center using a unified procedure and securely transferred to the coordinating center via encrypted channels. The coordinating center will perform logic checks and randomly audit 10% of records to monitor data consistency (target inconsistency rate of ≤3%). Outliers will be identified during data cleaning and handled according to prespecified rules. Missing data will be handled using multiple imputation for all variables with missing values, and an overall sensitivity analysis will be conducted using conditional imputation. Preliminary checks indicate that most variables have <20% missingness (many with none), and all variables have <30% missingness.

Baseline characteristics will be summarized as means and SDs, medians and IQRs, or counts and percentages (Table 2).

XSL•FO
RenderX

Table . Baseline characteristic table.

Variable	Type of variable	Summary statistics
Demographic characteristics		
Residence	Categorical	Frequencies and percentages
Age (years)	Continuous	Mean and SD or median and IQR
BMI (kg/m^2)	Continuous	Mean and SD or median and IQR
Smoking status	Binary	Frequencies and percentages
Alcohol consumption	Binary	Frequencies and percentages
Comorbidities		
Hypertension	Binary	Frequencies and percentages
Diabetes mellitus	Binary	Frequencies and percentages
Coronary artery disease	Binary	Frequencies and percentages
Psychiatric disorders	Binary	Frequencies and percentages
Sleep disorders	Binary	Frequencies and percentages
Respiratory diseases	Binary	Frequencies and percentages
Constipation	Binary	Frequencies and percentages
Laboratory measurements		
Serum creatinine (µmol/L)	Continuous	Mean and SD or median and IQR
Serum uric acid (µmol/L)	Continuous	Mean and SD or median and IQR
TG/HDL[a]	Continuous	Mean and SD or median and IQR
NLR[b]	Continuous	Mean and SD or median and IQR
ALT[c] (U/L)	Continuous	Mean and SD or median and IQR
D-dimer (mg/L)	Continuous	Mean and SD or median and IQR
Total PSA[d] (ng/mL)	Continuous	Mean and SD or median and IQR
Free PSA (ng/mL)	Continuous	Mean and SD or median and IQR
Prostate volume (mL)	Continuous	Mean and SD or median and IQR
Clinical outcomes		
Event 1	Categorical	Frequencies and percentages
Event 2	Categorical	Frequencies and percentages
Event 3	Categorical	Frequencies and percentages
Event 4	Categorical	Frequencies and percentages
Event 5	Categorical	Frequencies and percentages
Event 6	Categorical	Frequencies and percentages

[a]TG/HDL: triglyceride-to–high-density lipoprotein cholesterol ratio.

[b]NLR: neutrophil-to-lymphocyte ratio.

[c]ALT: alanine aminotransferase.

[d]PSA: prostate-specific antigen.

Time-to-event distributions will be summarized using appropriate survival and competing risk methods, with effect estimates from univariable and multivariable models reported as hazard ratios or subdistribution hazard ratios with 95% CIs (Multimedia Appendix 1).

Prediction models (eg, nomograms) will be constructed from multivariable results. Internal validation will use bootstrapping. External validation will include internal-external cross-validation between the 2 cohorts and an independent transportability assessment on the CHARLS (for predictors and end points that can be consistently harmonized). If miscalibration is observed in the CHARLS, we will report recalibration approaches (eg, intercept or slope updating) and present both original and recalibrated performance. Model performance will be assessed using concordance index, receiver operating characteristic curves, area under the curve, calibration plots, decision curves,

XSL•FO
RenderX

and reclassification metrics (net reclassification improvement and integrated discrimination improvement).

Statistical significance will be defined as 2-sided tests, with $P<.10$ (univariate analysis) and $P<.05$ (multivariate analysis) considered significant.

Patient and Public Involvement

As this study is a retrospective cohort analysis based on routinely collected follow-up data, patients and the public were not directly involved in formulating the research question, designing the study, or conducting the analyses. Nevertheless, we place strong emphasis on the clinical relevance and accessibility of the findings. Upon completion, results will be disseminated through peer-reviewed publications and scientific conferences and will also be shared with urology-related patient organizations and professional communities in China to facilitate translation into routine clinical decision-making and better reflect patient priorities. Building on these findings, we plan to explore prospective studies that incorporate patient and public involvement to further strengthen clinical relevance and societal impact.

Ethical Considerations

The study protocol has been approved by the ethics committee of the PLA General Hospital (approval S2022-700-01). The entire study will be conducted in accordance with the principles of the Declaration of Helsinki and relevant ethical guidelines for medical research.

As this study is a retrospective secondary analysis of routinely collected electronic health record data, the Ethics Committee of the PLA General Hospital granted a waiver of informed consent because the analysis uses deidentified data, involves minimal risk to participants, and no direct contact with participants will occur. No participant compensation will be provided.

The study will use routinely collected clinical data for research purposes only. Access to the dataset will be restricted to authorized team members, thereby minimizing the risk of privacy breaches. Data protection measures (including encryption, deidentification, and secure storage and transfer) are described elsewhere in the Methods section.

Dissemination and Data Availability

Findings will be disseminated through publications in peer-reviewed international and domestic journals and presented at scientific conferences, targeting clinicians, researchers, and the broader public to share key insights into the long-term prognosis of patients with BPH. Authorship will comply with the recommendations of the International Committee of Medical Journal Editors to ensure academic integrity and recognition of contributions.

Upon study completion, deidentified follow-up datasets will be archived in the Clinical Research Data Platform of the PLA General Hospital. Internal team members may conduct further subgroup analyses. External researchers may apply for access by submitting a written request detailing study aims and protocol. Access will be granted only after review and approval by the ethics committee and data platform and upon signing a data use agreement specifying scope and confidentiality obligations. Data will be retained for 20 years after study completion, after which they will be permanently destroyed according to institutional policy.

Results

As of mid-January 2026, data collection for the multicenter retrospective dataset from the PLA General Hospital network spanning 2001 to 2021 has been completed, and data cleaning and harmonization has commenced. The baseline characteristic table for cohort 1 has been completed as part of data quality control (Table 3). In parallel, we have summarized the distributions of the core variables and outcome-specific variables, which will inform subsequent harmonization and modeling (Textbox 2). We anticipate completing data harmonization and the prespecified analyses (including model development and validation) by May 2026, followed by translation of the findings to support routine clinical decision-making.

XSL•FO
RenderX

Table . Baseline characteristic table. Continuous variables are presented as means and SDs when normally distributed and as medians and IQRs when nonnormally distributed; categorical variables are expressed as frequencies and percentages (n=26,989).

Variable	Values
Age (years), mean (SD)	61.9 (12.8)
BMI (kg/m^2), mean (SD)	25.3 (3.5)
Smoking history, n (%)	
Yes	15,173 (56.2)
No	11,815 (43.8)
Alcohol drinking history, n (%)	
Yes	16,240 (60.2)
No	10,748 (39.8)
Hypertension, n (%)	
Yes	7757 (28.7)
No	19,232 (71.3)
Diabetes mellitus, n (%)	
Yes	4505 (16.7)
No	22,484 (83.3)
Coronary artery disease, n (%)	
Yes	5058 (18.7)
No	21,931 (81.3)
Psychiatric disorders, n (%)	
Yes	345 (1.3)
No	26,644 (98.7)
Sleep disorders, n (%)	
Yes	473 (1.8)
No	26,516 (98.2)
Respiratory diseases, n (%)	
Yes	964 (3.6)
No	26,025 (96.4)
Constipation, n (%)	
Yes	76 (0.3)
No	26,913 (99.7)
Geographic distribution, n (%)	
North China	13,726 (50.9)
Northeast China	3573 (13.2)
East China	5255 (19.5)
Central China	2683 (9.9)
South China	222 (0.8)
Southwest China	536 (2.0)
Northwest China	994 (3.7)
Serum creatinine (µmol/L), median (IQR)	79.6 (70.6-90.8)
Serum uric acid (µmol/L), median (IQR)	344.2 (293.0-400.5)
ALT[a] (U/L), median (IQR)	20.0 (14.5-28.7)
Plasma D-dimer (mg/L), median (IQR)	0.4 (0.2-0.8)

XSL•FO
RenderX

Variable	Values
tPSA[b] (ng/mL), median (IQR)	1.2 (0.6-2.8)
fPSA[c] (ng/mL), median (IQR)	0.3 (0.2-0.6)
NLR[d], median (IQR)	0.5 (0.4-0.6)
TG/HDL-C[e], median (IQR)	1.2 (0.8-1.7)
Prostate volume (mL), median (IQR)	31.3 (22.4-45.4)

[a]ALT: alanine aminotransferase.

[b]tPSA: total prostate-specific antigen.

[c]fPSA: free prostate-specific antigen.

[d]NLR: neutrophil-to-lymphocyte ratio.

[e]TG/HDL-C: triglyceride-to–high-density lipoprotein cholesterol ratio.

XSL•FO
RenderX

Textbox 2. Configuration of core variables and outcome-specific variables across outcomes. This textbox summarizes the planned variable configuration framework for outcome analyses and model development. Core variables were initially considered for all outcomes, whereas outcome-specific variables were additionally considered for the corresponding outcome. The final variable set will be determined based on data availability, missingness, and model stability.

Core variables

- Age
- BMI
- Smoking history
- Alcohol drinking history
- Hypertension
- Diabetes mellitus
- Coronary artery disease
- Psychiatric disorders
- Sleep disorders
- Respiratory diseases
- Geographic distribution
- Serum creatinine

Outcome 1 (benign prostatic hyperplasia–related surgery)

- Prostate volume
- Alanine aminotransferase (ALT)
- Plasma D-dimer
- Total prostate-specific antigen (tPSA)
- Free prostate-specific antigen (fPSA)
- Neutrophil-to-lymphocyte ratio (NLR)
- Triglyceride-to–high-density lipoprotein cholesterol ratio (TG/HDL-C)

Outcome 2 (urinary retention)

- Prostate volume
- Plasma D-dimer
- tPSA
- fPSA
- NLR
- Constipation

Outcome 3 (inguinal hernia)

- Prostate volume
- Constipation
- tPSA
- fPSA
- Plasma D-dimer

Outcome 4 (chronic kidney disease)

- Serum uric acid
- ALT
- Plasma D-dimer
- tPSA

XSL•FO
RenderX

- fPSA
- NLR
- TG/HDL-C
- Prostate volume

Outcome 5 (urothelial carcinoma [bladder or upper tract])

- NLR
- TG/HDL-C
- ALT
- Serum uric acid
- Constipation

Outcome 6 (all-cause mortality)

- NLR
- ALT
- Plasma D-dimer
- TG/HDL-C
- Serum uric acid
- Constipation

Discussion

This study aims to establish a large-scale, long-term, multicenter retrospective cohort to systematically elucidate the full disease continuum of BPH—from progression to outcomes to prediction—and address the current knowledge gaps regarding the natural history and prognostic determinants of BPH in the Chinese population.

Limitations

Research on the clinical course and outcomes of BPH has several limitations. First, although there is existing evidence suggesting that frailty phenotype and larger prostate volume are established risk factors for disease progression [41,42], robust quantitative data on mid- to long-term risks of key outcomes, such as surgical intervention and acute urinary retention, remain limited. Second, most studies have relatively short follow-up durations, lacking long-term risk estimates over 10 years, particularly for older adult and frail populations [43-45]. Third, many prior studies have small sample sizes or single-center designs, limiting statistical power for infrequent outcomes such as CKD [46,47]. More importantly, inadequate adjustment for competing risks in some analyses may lead to overestimation of disease-specific outcomes [48,49]. These limitations highlight the urgent need for large-sample, long-term, multicenter studies using robust statistical frameworks.

Strengths of This Study

Compared with previous studies, this work offers several methodological advantages. The follow-up duration is extended, and leveraging nearly 2 decades of clinical data, this study can delineate long-term trajectories of BPH outcomes, overcoming the constraints of short- to midterm studies. The study will also use a multicenter, large-scale cohort of patients across various disease stages and treatment pathways, enhancing representativeness and external validity, with a sample size estimation that ensures sufficient power to detect low-incidence outcomes. The outcome assessment is comprehensive—beyond conventional end points such as surgery and urinary retention, the study incorporates inguinal hernia, CKD, urothelial carcinoma, and all-cause mortality, providing a more holistic evaluation of disease burden, thereby yielding further insights into the pathogenesis and progression of BPH and enhancing the robustness of statistical evidence. The statistical methodology is rigorous—depending on outcome characteristics, both Cox proportional hazard and Fine-Gray competing risk models will be used, with multivariable adjustments to minimize confounding and maximize robustness. Finally, the study will use a predictive modeling framework, building on identified risk factors to construct visual predictive tools (eg, nomograms) with internal validation and cross-validation across cohorts and quantify predictive performance using C-statistics. This addresses the lack of reliable prognostic tools for BPH outcomes.

Expected Contributions

This study is expected to generate several important insights. First, it may identify novel risk factors beyond age and prostate volume, including metabolic parameters, comorbidity profiles, and initial treatment strategies, thereby improving early risk stratification. Second, it will provide authoritative, large-scale estimates of cumulative incidence and median time to event for major outcomes in the Chinese population, offering a robust evidence base for prognosis and health care planning.

XSL•FO
RenderX

Future Perspectives

While retrospective cohort studies inherently face challenges [50-54]—such as potential data incompleteness from historical records and inherent biases—these limitations point to clear directions for future advancement. Looking ahead, 2 key priorities will drive follow-up efforts.

The first is expanding and enhancing data quality. Future work will integrate multisource data (including community-based patient records, primary care data, and long-term follow-up supplements) to fill gaps in existing datasets, paired with advanced data cleaning techniques to reduce measurement bias. This expansion will improve data accuracy, completeness, and external validity, making findings more applicable to diverse populations beyond hospital-based cohorts.

The second priority is transitioning to prospective research. Building on this study's theoretical framework and preliminary insights, a dedicated prospective cohort will be designed to collect real-time clinical data—capturing dynamic BPH symptom changes, treatment responses, and long-term outcomes. Notably, leveraging long-term follow-up, multicenter design, and rigorous methodology, this prospective effort is expected to achieve significant breakthroughs in the study of BPH's natural history and prognosis. Its results are poised to provide high-quality evidence-based support for optimizing patient management, implementing precision interventions, and formulating health care policies.

Acknowledgments

The authors would like to express their gratitude to the patients whose data were included in the cohort for their contribution of clinical information to support the statistical analyses conducted in this study. They would also like to acknowledge the Big Data Center of the People's Liberation Army General Hospital for providing platform support for data retrieval.

Funding

This study was supported by the National Key Research and Development Program of China (grant 2022YFC3602901).

Authors' Contributions

HL and XZ contributed to conception and design. XZ and CL contributed to overall academic supervision and methodological guidance. ZL, JL, YL, FG, and CL contributed to planned data acquisition. ZL, JL, BG, HZ, and CL contributed to analysis and interpretation of the data. ZL, JL, and CL contributed to drafting the manuscript. HL and CL contributed to critical revision of the manuscript for important intellectual content. All authors are accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Univariable associations with time-to-event outcomes and multivariable models for events 1 to 6.
[DOCX File, 21 KB - resprot_v15i1e84965_app1.docx]

Checklist 1
STROBE checklist.
[PDF File, 163 KB - resprot_v15i1e84965_app2.pdf]

References

1. Bschleipfer T. Case-based presentation of the S2e guideline on the diagnosis and treatment of benign prostatic hyperplasia. Urologie 2024 May;63(5):431-438. [doi: 10.1007/s00120-024-02327-1] [Medline: 38597947]
2. Daryanto B, Suryanullah WS, Putra PYP. Holmium laser enucleation of the prostate versus transurethral resection of the prostate in treatment of benign prostatic hyperplasia: a meta-analysis of 13 randomized control trials. Curr Urol 2025 Jan;19(1):6-16. [doi: 10.1097/CU9.0000000000000257] [Medline: 40313420]
3. Cavanaugh D, Urbanucci A, Mohamed NE, Tewari AK, Figueiro M, Kyprianou N. Link between circadian rhythm and benign prostatic hyperplasia (BPH)/lower urinary tract symptoms (LUTS). Prostate 2024 Apr;84(5):417-425. [doi: 10.1002/pros.24656] [Medline: 38193363]
4. Tamalunas A, Keller P, Schott M, et al. Benign prostate hyperplasia - current medical therapy, new developments, and side effects. Ther Umsch 2023 Apr;80(3):113-122. [doi: 10.1024/0040-5930/a001423] [Medline: 36975030]
5. Villard N, Tsoumakidou G, Moldovan PC, et al. Giant prostatic hyperplasia in a 54-years old patient treated by prostate artery embolization: case report. Front Urol 2024;4:1446650. [doi: 10.3389/fruro.2024.1446650] [Medline: 40777098]

XSL•FO
RenderX

6. Ahmed R, Hamdy O, Awad RM. Diagnostic efficacy of systemic immune-inflammation biomarkers in benign prostatic hyperplasia using receiver operating characteristic and artificial neural network. Sci Rep 2023 Sep 8;13(1):14801. [doi: 10.1038/s41598-023-41781-3] [Medline: 37684320]
7. Hata J, Harigane Y, Matsuoka K, et al. Mechanism of androgen-independent stromal proliferation in benign prostatic hyperplasia. Int J Mol Sci 2023 Jul 19;24(14):11634. [doi: 10.3390/ijms241411634] [Medline: 37511400]
8. Koudonas A, Anastasiadis A, Tsiakaras S, et al. Overview of current pharmacotherapeutic options in benign prostatic hyperplasia. Expert Opin Pharmacother 2023;24(14):1609-1622. [doi: 10.1080/14656566.2023.2237406] [Medline: 37448198]
9. Tang X, Liu Z, Ren J, et al. Comparative RNA-sequencing analysis of the prostate in a mouse model of benign prostatic hyperplasia with bladder outlet obstruction. Mol Cell Biochem 2023 Dec;478(12):2721-2737. [doi: 10.1007/s11010-023-04695-2] [Medline: 36920576]
10. Popovics P, Skalitzky KO, Schroeder E, et al. Steroid hormone imbalance drives macrophage infiltration and Spp1/osteopontin+ foam cell differentiation in the prostate. J Pathol 2023 Jun;260(2):177-189. [doi: 10.1002/path.6074] [Medline: 36825524]
11. Baran C. The effect of acute urinary retention on the results of transurethral resection of the prostate. Urologia 2023 Nov;90(4):642-646. [doi: 10.1177/03915603231189627] [Medline: 37491943]
12. Mohamed AG, Badawy HF, Said AS, et al. Additive value of pyridostigmine to silodosin in the management of acute urinary retention secondary to benign prostatic hyperplasia: a randomized controlled trial. J Clin Med 2025 Jan 21;14(3):674. [doi: 10.3390/jcm14030674] [Medline: 39941344]
13. Wu ST, Juan TJ, Lu JH, et al. A novel approach to bladder dysfunction induced by bladder outlet obstruction: Low-intensity extracorporeal shock wave therapy promotes bladder regeneration. Life Sci 2025 Oct 1;378:123799. [doi: 10.1016/j.lfs.2025.123799] [Medline: 40505832]
14. Wei JT, Dauw CA, Brodsky CN. Lower urinary tract symptoms in men: a review. JAMA 2025 Sep 2;334(9):809-821. [doi: 10.1001/jama.2025.7045] [Medline: 40658396]
15. Camelo ALM, de Oliveira AM, Zamora-Obando HR, et al. Self-organizing maps to aid prognostic and diagnostic biomarker identification in exploratory metabolomics of benign prostatic hyperplasia. Metabolomics (Los Angel) 2025 Jun 19;21(4):83. [doi: 10.1007/s11306-025-02290-8] [Medline: 40537581]
16. Kuang WW, Cindolo L, Alsaody T, Chughtai BI. Tackling the progression of benign prostatic hyperplasia/benign prostatic obstruction progression: deobstructing within the “window of curability” (a hypothesis-generating review). Curr Urol 2025 Nov;19(6):388-395. [doi: 10.1097/CU9.0000000000000299] [Medline: 41058768]
17. Lama J, Winograd J, Codelia-Anjum A, et al. AI for BPH surgical decision-making: cost effectiveness and outcomes. Curr Urol Rep 2025 Dec;26(1):4. [doi: 10.1007/s11934-024-01240-6] [Medline: 39312102]
18. Sandhu JS, Bixler BR, Dahm P, et al. Management of lower urinary tract symptoms attributed to benign prostatic hyperplasia (BPH): AUA guideline amendment 2023. J Urol 2024 Jan;211(1):11-19. [doi: 10.1097/JU.0000000000003698] [Medline: 37706750]
19. Sapoval MR, Bhatia S, Déan C, et al. Two-year outcomes of prostatic artery embolization for symptomatic benign prostatic hyperplasia: an international, multicenter, prospective study. Cardiovasc Intervent Radiol 2024 Nov;47(11):1515-1524. [doi: 10.1007/s00270-024-03802-0] [Medline: 39230672]
20. Ahmed MM, Prasannakumary R, Ravichandran A, et al. Retrospective assessment of post-transurethral resection of the prostate complications in benign prostatic hyperplasia patients: impact of acute urinary retention. Cureus 2024 Oct;16(10):e72223. [doi: 10.7759/cureus.72223] [Medline: 39583458]
21. Khasanah N, Djaali W, Viventius Y. Electroacupuncture therapy for urinary retention in an elderly patient. Med Acupunct 2024 Apr 1;36(2):108-112. [doi: 10.1089/acu.2023.0081] [Medline: 38665926]
22. Jeh SU, Park C, Choi MS, et al. Development of a new diagnostic tool (LUTS/BPH Score) for male lower urinary tract symptoms due to benign prostatic hyperplasia. World J Mens Health 2026 Jan;44(1):129-135. [doi: 10.5534/wjmh.240306] [Medline: 40583029]
23. Fu X, Wang Y, Lu Y, Liu J, Li H. Association between metabolic syndrome and benign prostatic hyperplasia: the underlying molecular connection. Life Sci 2024 Dec 1;358:123192. [doi: 10.1016/j.lfs.2024.123192] [Medline: 39488266]
24. Ganesan V, Agarwal D. Medical advancements in benign prostatic hyperplasia treatments. Curr Urol Rep 2024 May;25(5):93-98. [doi: 10.1007/s11934-024-01199-4] [Medline: 38448685]
25. Shin BNH, Qu L, Rhee H, Chung E. Systematic review and network meta-analysis of re-intervention rates of new surgical interventions for benign prostatic hyperplasia. BJU Int 2024 Aug;134(2):155-165. [doi: 10.1111/bju.16304] [Medline: 38600763]
26. Te AE, Ramaswamy A, Kaplan SA. Medical therapy for benign prostatic hyperplasia and lower urinary tract symptoms. Urol Clin North Am 2025 Nov;52(4):595-604. [doi: 10.1016/j.ucl.2025.08.001] [Medline: 41106988]
27. Marchi A, Pellerin O, Querub C, et al. Long-term outcomes of prostate artery embolization for acute urinary retention: an analysis of 88 patients. Diagn Interv Imaging 2023 Jun;104(6):292-296. [doi: 10.1016/j.diii.2023.01.012] [Medline: 36801097]

XSL•FO
RenderX

28. Tiruneh T, Getachew T, Getahun F, Gebreyesus T, Meskele S. Magnitude and factors associated with benign prostatic hyperplasia among adult male patients visiting Wolaita Sodo University comprehensive specialized hospital southern Ethiopia. Sci Rep 2024 Dec 30;14(1):31556. [doi: 10.1038/s41598-024-83483-4] [Medline: 39738533]
29. Gravas S, Manuel-Palacios J, Chavan C, et al. Modeling study of the effect of placebo and medical therapy on storage and voiding symptoms, nocturia, and quality of life in men with prostate enlargement at risk for progression. Prostate Cancer Prostatic Dis 2024 Sep;27(3):469-477. [doi: 10.1038/s41391-023-00731-w] [Medline: 37794168]
30. Ivanuta M, Puia D, Pricop C. Elements for trial without catheter (TWOC) success in benign prostatic hyperplasia patients: lessons we have learned. Cureus 2023 Dec;15(12):e50980. [doi: 10.7759/cureus.50980] [Medline: 38259407]
31. Creta M, Russo GI, Bhojani N, et al. Bladder outlet obstruction relief and symptom improvement following medical and surgical therapies for lower urinary tract symptoms suggestive of benign prostatic hyperplasia: a systematic review. Eur Urol 2024 Oct;86(4):315-326. [doi: 10.1016/j.eururo.2024.04.031] [Medline: 38749852]
32. Ji T, Huang K, He B, Wang H. The predictive value of prostate spherical volume ratio in lower urinary tract symptoms and clinical progression of benign prostatic hyperplasia: a retrospective cohort study. Int Urol Nephrol 2025 Jul;57(7):2023-2032. [doi: 10.1007/s11255-024-04355-4] [Medline: 39812965]
33. Lin L, Wang W, Shao Y, Li X, Zhou L. National prevalence and incidence of benign prostatic hyperplasia/lower urinary tract symptoms and validated risk factors pattern. Aging Male 2025 Dec;28(1):2478875. [doi: 10.1080/13685538.2025.2478875] [Medline: 40084396]
34. Bschleipfer T, Abt SD, Becher KF, et al. Conservative and pharmacological treatment of benign prostatic hyperplasia: the German S2e-guideline 2023-part2. Urologie 2023 Oct;62(10):1048-1056. [doi: 10.1007/s00120-023-02183-5] [Medline: 37796278]
35. Pina-Vaz T, Silva AC, Silva C, Miranda I, Cruz F, Silva JA. The urinary and prostatic microbiome in non-neurogenic male LUTS/BPH: a systematic review. Neurourol Urodyn 2024 Nov;43(8):1938-1947. [doi: 10.1002/nau.25530] [Medline: 38934484]
36. An Z, Zhou Q, Shao J, et al. Life disturbance and hospital visit experiences among Chinese patients with benign prostatic hyperplasia: a qualitative study. BMC Prim Care 2024 May 3;25(1):149. [doi: 10.1186/s12875-024-02378-5] [Medline: 38702603]
37. Liu D, Li C, Li Y, et al. Benign prostatic hyperplasia burden comparison between China and United States based on the global burden of disease study 2019. World J Urol 2023 Dec;41(12):3629-3634. [doi: 10.1007/s00345-023-04658-8] [Medline: 37831157]
38. Schoenfeld DA. Sample-size formula for the proportional-hazards regression model. Biometrics 1983 Jun;39(2):499-503. [Medline: 6354290]
39. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 1995 Dec;48(12):1503-1510. [doi: 10.1016/0895-4356(95)00048-8] [Medline: 8543964]
40. Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med 2019 Mar 30;38(7):1276-1296. [doi: 10.1002/sim.7992] [Medline: 30357870]
41. Bauer SR, Walter LC, Ensrud KE, et al. Assessment of frailty and association with progression of benign prostatic hyperplasia symptoms and serious adverse events among men using drug therapy. JAMA Netw Open 2021 Nov 1;4(11):e2134427. [doi: 10.1001/jamanetworkopen.2021.34427] [Medline: 34817584]
42. Manfredi C, Napolitano L, Ditonno F, et al. Long-term functional outcomes and surgical retreatment after thulium laser enucleation of the prostate: a 10-year follow-up study. Int Braz J Urol 2024;50(3):309-318. [doi: 10.1590/S1677-5538.IBJU.2024.0039] [Medline: 38446905]
43. Adil Mehmood M, Jehan S, Khan I, Ilyas M, Khan U, Ali W. The assessment of patient-reported outcomes in the medical management of patients with benign prostatic hyperplasia. Cureus 2024 Aug;16(8):e67027. [doi: 10.7759/cureus.67027] [Medline: 39280479]
44. Rowaiee R, Akhras A, Lakshmanan J, Sikafi Z, Janahi F. Rezum therapy for benign prostatic hyperplasia: Dubai's initial experience. Cureus 2021 Sep;13(9):e18083. [doi: 10.7759/cureus.18083] [Medline: 34692298]
45. Codelia-Anjum AJ, Berjaoui MB, Khondker A, et al. Procedural intervention for benign prostatic hyperplasia in men ≥ age 70 years - a review of published literature. Clin Interv Aging 2023;18:1705-1717. [doi: 10.2147/CIA.S414799] [Medline: 37849957]
46. Lee SWH, Chan EMC, Lai YK. The global burden of lower urinary tract symptoms suggestive of benign prostatic hyperplasia: a systematic review and meta-analysis. Sci Rep 2017 Aug 11;7(1):7984. [doi: 10.1038/s41598-017-06628-8] [Medline: 28801563]
47. Ouyang C, Zhou L, Liu J, Wang L, Lu Y. Benign prostatic hyperplasia increases long-term chronic kidney disease risk. Mayo Clin Proc 2025 May;100(5):790-800. [doi: 10.1016/j.mayocp.2024.11.024] [Medline: 40196966]
48. Hamada T, Oyama H, Tange S, et al. The revised Kyoto criteria and risk of malignancy among patients with intraductal papillary mucinous neoplasms. Clin Gastroenterol Hepatol 2024 Dec;22(12):2413-2423. [doi: 10.1016/j.cgh.2024.05.043] [Medline: 38880125]

XSL•FO
RenderX

49. Hamada T, Oyama H, Nakai Y, et al. Clinical outcomes of intraductal papillary mucinous neoplasms with dilatation of the main pancreatic duct. Clin Gastroenterol Hepatol 2023 Jul;21(7):1792-1801. [doi: 10.1016/j.cgh.2023.01.032] [Medline: 36787835]
50. Yang CY, Chen GM, Wu YX, et al. Clinical efficacy and complications of transurethral resection of the prostate versus plasmakinetic enucleation of the prostate. Eur J Med Res 2023 Feb 18;28(1):83. [doi: 10.1186/s40001-023-00989-9] [Medline: 36805825]
51. Wang M, Xi Y, Qiu Z, Huang N, Zhang L, Liu J. Comparative efficacy and safety of 180 W XPS vs. 120 W HPS GreenLight laser therapy for benign prostatic hyperplasia: a systematic review and meta-analysis. PeerJ 2024;12:e18615. [doi: 10.7717/peerj.18615] [Medline: 39619178]
52. Deng T, Duan X, He Z, Zhao Z, Zeng G. Association between 5-alpha reductase inhibitor use and the risk of depression: a meta-analysis. Urol J 2020 Aug 23;18(2):144-150. [doi: 10.22037/uj.v16i7.5866] [Medline: 32869255]
53. Lee SU, Lee SH, So AH, et al. Association between benign prostatic hyperplasia and suicide in South Korea: a nationwide retrospective cohort study. PLoS ONE 2022;17(3):e0265060. [doi: 10.1371/journal.pone.0265060] [Medline: 35271681]
54. Cakir SS, Polat EC, Ozcan L, Besiroglu H, Ötunctemur A, Ozbek E. The effect of prostatic inflammation on clinical outcomes in patients with benign prostate hyperplasia. Prostate Int 2018 Jun;6(2):71-74. [doi: 10.1016/j.prnil.2017.12.003] [Medline: 29922636]

Abbreviations

BPH: benign prostatic hyperplasia
CHARLS: China Health and Retirement Longitudinal Study
CKD: chronic kidney disease
EHR: electronic health record
LUTS: lower urinary tract symptoms
PLA: People's Liberation Army
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Schwartz; submitted 30.Sep.2025; peer-reviewed by O Olowokere; accepted 28.Jan.2026; published 02.Mar.2026.

Please cite as:
Lv Z, Liu J, Lu Y, Zhang X, Gao F, Guo B, Zhang H, Li H, Lv C
Disease Progression and Outcomes in Patients With Benign Prostatic Hyperplasia: Protocol for a Multicenter Retrospective Cohort Study
JMIR Res Protoc 2026;15:e84965
URL: https://www.researchprotocols.org/2026/1/e84965
doi:10.2196/84965

© Zexuan Lv, Jiaxing Liu, Yin Lu, Xu Zhang, Fan Gao, Bingyang Guo, Hongyu Zhang, Hongzhao Li, Chao Lv. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Safety of Administration of Vasopressors Through Peripheral Compared to Central Venous Catheters in a Rural Kenyan Hospital: Protocol for a Prospective Observational Cohort Study

Grace Hezi Gao[1*], MD; Brian Kimani[2*], ECCCO; Linette Lepore[2], MD; Zahra Aghababa[3], MA; Hannah Gitura[2], ECCCO; Michael Kinuthia[4], PECCCO; Jack Musau[4], PECCCO; Judy Muthoki[2], ECCCO; Alice Mwonge[2], ECCCO; Fredrick Ndibaru[2], ECCCO; Ann Viola Ndubi[4], MD; Juliah Ndungu[2], RN; Kinzi Nixon[4], PECCCO; Moses Odhiambo[5], BSc; Winslet Okari[2], ECCCO; Benjamin Thairu[2], ECCCO; Lilian Vihenda[2], CCN; Rachel M Wojcik[6], MD; B Jason Brotherton[1,7], MD; Kristina E Rudd[3], MD

[1]Department of Critical Care Medicine, University of Pittsburgh Medical Center, 3520 Fifth Ave., Suite 100, Pittsburgh, PA, United States

[2]Department of Medicine, AIC Kijabe Hospital, Kijabe, Kenya

[3]Department of Critical Care Medicine, Center for Research, Investigation, and Systems Modeling of Acute Illness (CRISMA), University of Pittsburgh, Pittsburgh, PA, United States

[4]Department of Pediatrics, AIC Kijabe Hospital, Kijabe, Kenya

[5]Department of Research, AIC Kijabe Hospital, Kijabe, Kenya

[6]Department of Medicine, Division of Pulmonary, Allergy, and Critical Care Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[7]Department of Border and Global Health, University of Arizona, Tucson, AZ, United States

[*]these authors contributed equally

Corresponding Author:
Grace Hezi Gao, MD
Department of Critical Care Medicine, University of Pittsburgh Medical Center, 3520 Fifth Ave., Suite 100, Pittsburgh, PA, United States

Abstract

Background: The infusion of vasopressors is a standard treatment for shock, and international guidelines recommend administering these medications through central venous catheters (CVCs) due to concerns about potential extravasation and local tissue injury with peripheral intravenous (PIV) administration. However, CVCs are often unavailable in resource-variable settings due to lack of human and material resources. Previous studies have assessed the safety of vasopressor infusion through PIV catheters but have considered only limited patient populations or a short infusion time or have used retrospective designs that may have failed to capture mild complications.

Objective: The primary objective of this study is to observe and describe the incidence of complications among patients receiving vasopressor infusion via PIV administration. The secondary objective is to assess whether the safety of PIV vasopressor administration is noninferior to the safety of administration via CVC.

Methods: This prospective observational study is being conducted at African Inland Church (AIC) Kijabe Hospital, a 360-bed tertiary care teaching hospital in rural Kenya. All patients (adult, obstetric, and pediatric) receiving intravenous vasopressor infusions who are admitted to the intensive care unit or high-dependency unit will be included. Patients will be followed up on twice daily from the start of vasopressor infusion to 72 hours after vasopressor discontinuation or death, whichever occurs first. Demographic, physiological, laboratory, therapeutic, and outcome data will be collected.

Results: Consecutive enrollment began in October 2023 and is ongoing. As of July 2025, we have enrolled 190 patients. We anticipate that the time to enroll the number of patients required to reach our power goal will be 24 months.

Conclusions: This study in a resource-variable setting will allow for more accurate and comprehensive data collection on vasopressor administration and potential complications as they arise, whereas most previous studies have been retrospective in nature. In addition, this is the first study of its kind to include both adult and pediatric patients within a mixed intensive care unit population with broad etiologies of shock, which could improve generalizability.

International Registered Report Identifier (IRRID): DERR1-10.2196/81794

(JMIR Res Protoc 2026;15:e81794) doi:10.2196/81794

XSL•FO
RenderX

KEYWORDS

vasopressors; intensive care; critical care; global health; limited resources; shock

Introduction

Shock is a life-threatening condition characterized by inadequate global oxygen delivery relative to demand [1]. It is a common problem in hospitalized patients, can result from various etiologies, most frequently septic shock, and is a leading cause of mortality worldwide [2]. The infusion of vasopressors is a critical component of shock management, with recommendations not to delay use of vasopressor administration until fluid resuscitation is completed but to initiate vasopressors early to achieve target mean arterial pressure [3]. While international guidelines recommend rapid initial administration of these medications through peripheral intravenous (PIV) catheters for a short period if central venous catheters (CVCs) are unavailable, they recommend eventual transition to CVCs for continued administration due to concerns about extravasation and local tissue injury with PIV administration [4].

However, CVCs are often unavailable or challenging to use in many resource-variable settings due to supply chain issues, limited clinician training, and infection control concerns [5]. Recognizing these challenges, several studies have been conducted to assess the safety of vasopressor infusion through PIV catheters as an alternative to CVCs by evaluating the frequency of adverse events of peripherally administered vasopressors [6,7]. A 2019 retrospective single-center cohort study by Lewis et al [8] evaluated 202 patients who received peripheral vasopressors, primarily norepinephrine, with a median infusion duration of 7.5 hours. The incidence of extravasation events was 4%, all managed conservatively. An Australian cohort study of 202 patients reported a 23.2% complication rate with peripheral vasopressor administration, mostly involving phenylephrine infusions for a median duration of 10 hours [9]. In a cancer hospital in India with resources similar to those of the proposed study setting, 122 adults received peripheral vasopressors for a median duration of 11 hours, with only 1 case of extravasation-related tissue injury [10]. A 2022 prospective study in a Pakistani pediatric intensive care unit (ICU) reported a 2.2% extravasation injury rate among 369 patients receiving peripheral vasopressors, primarily epinephrine, for a median duration of 24 hours [11]. A recent prospective cohort study conducted across multiple hospitals in Addis Ababa, Ethiopia, enrolling adults with circulatory shock also found low rates of extravasation injury (3 in 250 patients who received peripheral vasopressors). None of the extravasation injuries occurred in patients who received peripheral vasopressors for less than 5 days [12].

When available, CVCs are not without complications [13]. A study from a high-volume facility in South Africa found a complication rate of 18% among more than 1000 patients receiving a CVC [14]. A recent meta-analysis of 130 studies found a 3% rate of major complications associated with CVC placement—including arterial cannulation or puncture, pneumothorax, infection, and deep vein thrombosis [15]. In the pediatric population, as many as 1 in 4 CVCs fail before completion of therapy [16]. Placement of a CVC is the main risk factor for development of deep vein thrombosis in children, with one study finding an incidence of CVC-related thrombosis of 30.3% [17,18].

Even in well-resourced settings in which CVCs are readily available, there may still be a role for peripheral vasopressor administration. Studies have shown that early initiation of vasopressors in patients with septic shock is associated with increased survival [19]. It generally takes more time to place CVCs as compared to PIV catheters, and depending on the location of the CVC, radiographic confirmation of placement must be obtained before it can be used, further delaying initiation of vasopressors. Thus, multiple previous studies have been conducted to describe and assess the prevalence of complications of peripheral vasopressor administration. A 2020 systematic review and meta-analysis by Tran et al [20], which included 8 observational studies and 1 randomized controlled study with a total of 1835 adult patients who received peripheral vasopressor administration, reported 7% adverse event occurrence, with only 4% of those being major events. Similarly, another systematic review and meta-analysis by Owen et al [21] in 2021, which included both adults and children, reported a 1.8% complication rate for PIV catheters. A 2015 systematic review by Loubani and Green [22], which included primarily case reports, described published reports on local tissue injury or extravasation events. The average duration of peripheral vasopressor administration was 55.9 hours before local tissue injury and 35.2 hours before extravasation [22]. Most recently, a secondary analysis of the Crystalloid Liberal or Vasopressors Early Resuscitation in Sepsis trial evaluated the route of vasopressor initiation and continuation of peripheral vasopressors beyond 6 hours and found that most patients had vasopressors initiated peripherally and continued past 6 hours, with rare low-grade complications [23,24]. Importantly, none of these reports were from the African continent.

Collectively, these studies suggest that short-term peripheral vasopressor administration may be a reasonable alternative when CVCs are unavailable. However, most studies have evaluated relatively limited patient populations and short infusion times. Additionally, most studies have been retrospective, limiting the ability to capture administration and complication details comprehensively. To date, few studies have thoroughly evaluated the safety profile of prolonged peripheral vasopressor infusions regarding duration or concentration, particularly in resource-variable settings in which PIV catheter use is common. Finally, to our knowledge, no study has directly compared the safety of vasopressor administration via PIV catheter vs CVC.

At AIC Kijabe Hospital (KH) in Kenya, it is standard practice to use PIV catheters for the duration of a patient's vasopressor requirement, often lasting several days, regardless of the agent used or concentration and dose required. There is no maximum concentration or dose for which it is required or standard to switch to a CVC. This is because CVCs are infrequently used due to cost, supply chain issues, and limited clinician training.

XSL•FO
RenderX

Therefore, this study aims to address knowledge gaps by documenting the complication rates among patients receiving vasopressor infusion through PIV catheters.

Methods

Objectives

The primary objective of this study is to observe and describe the incidence of complications among patients receiving vasopressor infusion via PIV catheter. The secondary objective is to assess whether the safety of vasopressor administration via PIV catheter is noninferior to administration via CVC. This study aims to fill the gap in data on the safety of prolonged peripheral vasopressor administration in resource-variable settings.

Setting

This is a prospective observational single-center cohort study to be conducted at KH. KH is a rural 350-bed tertiary care teaching hospital that receives referrals from all over Kenya and East Africa. There are a number of graduate medical education programs at KH, including residency programs for physicians, postgraduate clinical officer training in emergency and critical care, and nursing education. Services available at KH include adult medicine, pediatric medicine, obstetrics and gynecology, surgical services, and outpatient clinics. KH has a 5-bed adult ICU, 10 adult high-dependence unit (HDU) beds, a 5-bed maternity HDU (obstetric patients who are critically ill are admitted to the adult ICU), and an 8-bed combined pediatric ICU+HDU. At KH, the HDUs are capable of caring for patients who are critically ill and do not require invasive mechanical ventilation, including those receiving vasopressor infusions. All hospitalized patients receiving vasopressor infusions at KH are admitted to an HDU or ICU and are not admitted to general hospital wards.

Eligibility Criteria

All adult, maternity, and pediatric patients receiving intravenous (IV) vasopressor (eg, norepinephrine, epinephrine, vasopressin, or dopamine) infusion for less than 24 hours admitted to the ICU or HDU and receiving their infusion via a PIV catheter or CVC placed at KH will be included regardless of the etiology of shock. Post–cardiac arrest patients achieving return of spontaneous circulation and requiring vasopressor infusion to maintain acceptable blood pressure will be included. If a patient experiences cardiac arrest while at KH and was not on vasopressors previously, they will be included if return of spontaneous circulation is achieved and ongoing vasopressor support is required.

Patients not receiving vasopressor infusions will be excluded. Additionally, patients requiring vasopressor support solely while in the operating theater (operating room) or the casualty (accident and emergency department) or dialysis unit will be excluded due to the inability of the study team to maintain adequate data tracking. Individuals transferred to KH with a CVC in situ will also be excluded due to the inability to fully track complications related to CVC placement. Patients in the general maternity, adult, or pediatric acute care wards or the neonatal ICU or those who have received vasopressor infusion for >24 hours before enrollment will also be excluded. Neonates receiving vasopressors via umbilical vein catheter will be excluded. Patients receiving vasopressors via intraosseous access will not be included. However, if they start to receive vasopressors via a PIV catheter or CVC, they will then be eligible for this study. Textbox 1 provides full details of the inclusion and exclusion criteria.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Any adult, pediatric, or maternity patient admitted to an intensive care unit (ICU) or high-dependence unit (HDU) receiving vasopressor (epinephrine, norepinephrine, vasopressin, dopamine, or phenylephrine) infusion
- Vasopressors being administered via peripheral intravenous catheter placed at any facility or via central venous catheter (CVC) placed at African Inland Church Kijabe Hospital
- Vasopressor infusion for <24 hours before screening for study enrollment

Exclusion criteria

- Patients not receiving vasopressor infusion
- Patients requiring vasopressor support only while in the operating theater or casualty (eg, accident and emergency department) or dialysis unit
- Patients with cardiac arrest without return of spontaneous circulation
- Patients in maternity, adult, or pediatric wards (outside of ICUs or HDUs)
- Patients in nursery (neonatal ICU)
- Patients with CVC placed at another facility
- Patients receiving only an inotrope (eg, dobutamine) infusion
- Vasopressor infusion for >24 hours before screening for study enrollment

XSL•FO
RenderX

Patient Enrollment and Data Collection

Eligible patients will be identified and consecutively enrolled by study staff through review of all patients admitted to an HDU or ICU twice every day using the inclusion and exclusion criteria listed in Textbox 1.

A case report form has been created and is available for review (Multimedia Appendix 1). At the time of enrollment, demographic identifiers such as age and gender will be recorded. The types of shock reflected in the electronic medical record on admission will be recorded as the etiology for vasopressor administration. Patient location at the time of enrollment will be recorded, as well as the primary medical service caring for the patient (eg, medicine or surgery). Clinical information such as vital signs and laboratory values will be recorded as the highest and lowest values within 12 hours of enrollment for each variable. Oxygen support (eg, liters per minute) will be recorded, including whether a patient is receiving invasive mechanical ventilation. Finally, patient status at the time of discontinuation of vasopressor (ie, alive vs dead) will be recorded.

Enrolled patients will be followed up on from the start of vasopressor infusion to 72 hours after discontinuation or death, whichever occurs first. Patients will be evaluated twice daily for the occurrence of complications, and any new complications will be documented by study team members. Information on complications will be recorded in the case report form along with how the complication was addressed (ie, phentolamine vesicant injection, surgical intervention, change of cannula, or no intervention needed). All vasopressors are infused through a syringe pump.

Detailed data on the type, rate, and route of vasopressor infusion will be recorded twice daily on a bedside data collection form (Multimedia Appendix 2). Data obtained from the bedside will include site and route of IV vasopressor infusion, size of PIV cannula, number and name of vasopressors infused, and highest rate of vasopressor infusion during each 12-hour shift. Each vasopressor at KH has a standard concentration regardless of peripheral or central route of administration, which will be used to determine the maximum dose used during vasopressor infusion. The standard concentrations can be found in Table 1. A study team clinician that is participating in the care of each patient will record the maximum rate and IV access information (eg, infusion complications, loss of IV catheter, hypotension, and transition to CVC) for the previous 12 hours. All peripheral cannulas at KH are of standard length—there are no long peripheral cannulas. In addition, all CVCs are standard—no midline catheters or peripherally inserted CVCs are used at KH. All patient care will be delivered according to standard clinical practices at KH and will in no way be altered for study patients.

Table . Standard concentrations of vasopressors at African Inland Church Kijabe Hospital.

Vasopressor	Concentration
Norepinephrine	8 mg per 50 mL
Epinephrine	8 mg per 50 mL
Vasopressin	20 units per 50 mL
Dopamine	500 mg per 50 mL

Other variables of interest will be hospital mortality, duration of vasopressor infusion, length of hospital stay, and overall disposition. Severity of illness using the Modified Sequential Organ Failure Assessment, Modified Early Warning Score, and Pediatric Logistic Organ Dysfunction score will be calculated. We will collect data regarding CVC placement conditions, including environmental conditions of placement (eg, sterile vs nonsterile and ward vs operating theater), use of ultrasound guidance, and medical service placing the CVC. Duration of CVC placement will also be tracked.

In the event of missing data, we will review the electronic medical record to attempt to obtain these data. All collected data will be entered into a secure electronic database using REDCap (Research Electronic Data Capture; Vanderbilt University). Quality checks will be conducted regularly to ensure the accuracy and completeness of the data. Access to the database will be restricted to authorized personnel only.

Outcomes

Our primary end point of interest is a composite of any complication of peripheral vasopressor infusion (Textbox 2). Our secondary end point of interest is the comparison of rates of complications among patients receiving vasopressor infusion via PIV catheters vs CVCs (Textbox 3). All complications will be recorded and grouped into mild, moderate, or severe categories for analysis.

XSL•FO
RenderX

Textbox 2. Adverse events of peripheral vasopressor administration.

Mild

- Extravasation or infiltration
- Thrombophlebitis
- Loss of line
- Blistering

Moderate

- Cellulitis
- Skin necrosis or ulcer
- Hypotension due to loss of line

Severe

- Gangrene
- Digital ischemia or necrosis

Textbox 3. Adverse events of central vasopressor administration.

Mild

- Difficult line placement documented by the clinical team
- Arterial puncture
- Unplanned line removal (eg, fell out or patient pulled it out)
- Thrombosis of line
- Thrombophlebitis

Moderate

- Line-associated deep venous thrombosis
- Hypotension due to loss of line

Severe

- Pneumothorax
- Associated bloodstream infection

Sample Size and Power Calculations

Assuming approximately 4:1 enrollment of patients receiving vasopressor infusions via PIV catheters vs CVCs (consistent with baseline rates at KH), the necessary sample size to meet all study objectives is 143. Our primary objective is descriptive in nature, so power calculations are not applicable. For our secondary objective (test of noninferiority), we assumed that the complication rates of CVC and PIV vasopressor administration are 25% and 10%, respectively, based on clinical experience at KH and literature review. With an absolute tolerable difference of up to 10%, we calculated that we need a sample size of 29 patients in the CVC group and 114 in the PIV group to assess noninferiority with 80% power and a one-sided α of .025.

Data Analysis Plan

Statistical analyses will be performed on all patients enrolled using the Stata analytical software (version 16; StataCorp). Categorical data will be described using percentages, whereas continuous data will be described using median values and IQRs as we anticipate potential nonnormal distributions due to the nature of the variables. Chi-square tests will be conducted for categorical data, and Wilcoxon rank sum tests will be conducted for continuous data. Statistical significance will be determined at $P<.05$. There are planned, prespecified subgroup analyses for adults (aged ≥16 years) and children (aged ≤15 years). Models will be adjusted for baseline characteristics where appropriate.

Ethical Considerations

This study has received ethics approval from the KH and University of Pittsburgh institutional review boards (approval numbers: KH/ISERC/02718/0060/2023; STUDY23080034). It will be conducted under a waiver of consent as it is a noninterventional study with minimal risk that does not deviate from the current standard of care. There are no planned interventions, nor are there plans to change standard practices

XSL•FO
RenderX

at KH. This study does not plan to prohibit the insertion of CVCs in patients with shock. The decision to do so will be made at the judgment of the treating clinician. Privacy and confidentiality of research participant data and identity will be maintained throughout enrollment and data collection.

Dissemination

In accordance with international research standards, we will share the results of this study as widely as possible, with particular attention to disseminating the findings with the clinical staff at KH in each unit where we conduct the study. We will also share our results with the international medical community by submitting at least one full-length manuscript for publication in a peer-reviewed open access journal and submitting abstracts to relevant scientific conferences. When possible, we will give priority to conferences in Kenya or sub-Saharan Africa, with special attention to AIC Kijabe Research Day.

Results

Consecutive enrollment began in October 2023 and is currently ongoing. We anticipate that the time to enroll the number of patients required to reach our power goal will be 24 months. Once this goal has been reached, we estimate that reviewing, analyzing, and filling in the missing data will take 6 additional months. During this time, there will be periodic review of the data collected to ensure accuracy, as well as periodic analysis to assess for the occurrence of our primary outcome.

Discussion

This is a prospective, investigator-initiated, single-center observational cohort study that plans to enroll and follow up on patients receiving vasopressor infusions via PIV catheters and CVCs to describe and compare the incidence of complications. Our hypothesis is that there will be few severe complications of administration of vasopressors via PIV catheter in this setting and that the safety of vasopressor administration via PIV catheter will be noninferior to that of administration via CVC.

This study is being conducted at a rural Kenyan hospital in a resource-variable setting. KH was chosen as the study site for a variety of reasons. Most importantly, the current usual practice at KH is the use of PIV catheters for administration of vasopressors regardless of agent, duration, or rate used. CVCs are not commonly placed primarily for vasopressor administration but, rather, are more commonly placed for other indications, such as total parenteral nutrition, hemodialysis, or chemotherapy, and are sometimes subsequently used for vasopressors if available [25]. This is in contrast with current standard practice in many high-resource settings, where vasopressor administration is frequently a primary indication for CVC placement. In addition, the duration of vasopressor administration via PIV catheter at KH can be multiple days. This differs from current international guidelines that recommend insertion of a CVC to continue vasopressor administration for prolonged infusion times (>6 hours) [4]. Thus, KH provides a unique setting for this study to help provide evidence on the adequacy and safety of peripheral vasopressor administration in resource-variable settings in which CVC placement may not be feasible.

Compared to previous studies, this is the first study to include both adult and pediatric patients, resulting in a relatively large sample size. In addition, patients are enrolled not just from the ICU but also from the HDU, and they include a mixed ICU population with broad etiologies of shock not limited to septic shock, which improves the generalizability of the results. As mentioned previously, many patients at KH have a prolonged duration of vasopressor infusion via PIV catheter. This contrasts with previous studies that have mostly looked at short infusion times, which not only provides an ideal setting for this study but also reflects the reality of resource-variable setting practice. Finally, this is a prospective observational study with enrollment ongoing in real time, allowing for more accurate and comprehensive data collection on vasopressor administration and potential complications as they arise.

There are several limitations to this study, the most important of which is that, as this is a prospective observational study without randomization, there will be residual confounding variables and differences in baseline between the PIV and CVC groups. As patient populations receiving CVCs generally have higher severity of illness and mortality than those without CVCs in critical care studies, we will not compare the PIV and CVC groups based on these clinical outcomes. In addition, as there is no blinding, a heightened awareness of potential complications may result in a lower threshold to detect a complication and change the IV site, thereby overestimating the complication rate. We have categorized the complications as mild, moderate, and severe to better account for this. Finally, this study is being conducted in a resource-limited hospital with a limited electronic medical record, with potential for limited or missing data.

This study aims to address the critical knowledge gap in understanding vasopressor administration safety in resource-variable settings in which central venous access may be limited or unavailable. By studying vasopressor delivery in a setting in which PIV catheter use is the primary method of delivery rather than the exception, we hope to provide insights into alternative approaches to critical care medication delivery in resource-variable settings.

Acknowledgments

The authors thank the patients and families at African Inland Church (AIC) Kijabe Hospital for their participation in this project. They also thank the many AIC Kijabe Hospital clinical staff members involved in data collection in the adult and pediatric departments for their dedication and contribution to the success of this project. The authors thank Dr Immaculate Barasa, Dr Peter Halestrap, and Anna Nakayama for their contributions to the early conception of this project.

XSL•FO
RenderX

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for profit sectors. KER reports funding from National Institutes of Health and National Institute of General Medical Sciences (grant 1K23GM141463).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: GHG, BK, LL, BT, HW, MO, AM, BJB, KER
Investigation: GHG, BK, LL, HG, MK, J Musau, J Muthoki, AM, FN, AVN, KN, DO, WO, BT, LV, JW, RMW
Formal analysis: GHG, BK, ZA, MO, BJB, KER
Writing - original draft: GHG, BK, LL, ZA, BJB, KER
Writing - review and editing: GHG, BK, LL, ZA, HG, MK, J Musau, J Muthoki, AM, FN, AVN, KN, MO, DO, WO, BT, LV, JW, RMW, BJB, KER
Supervision: GHG, BK, LL, BJB, KER

Conflicts of Interest

None declared.

Multimedia Appendix 1
Case report form.
[DOCX File, 27 KB - resprot_v15i1e81794_app1.docx]

Multimedia Appendix 2
Bedside data collection sheet.
[DOCX File, 22 KB - resprot_v15i1e81794_app2.docx]

References

1. Soong JT, Soni N. Circulatory shock. Medicine 2013 Feb;41(2):64-69. [doi: 10.1016/j.mpmed.2012.11.012]
2. Chiu C, Legrand M. Epidemiology of sepsis and septic shock. Curr Opin Anaesthesiol 2021 Apr 1;34(2):71-76. [doi: 10.1097/ACO.0000000000000958] [Medline: 33492864]
3. Scheeren TW, Bakker J, De Backer D, et al. Current use of vasopressors in septic shock. Ann Intensive Care 2019 Jan 30;9(1):20. [doi: 10.1186/s13613-019-0498-7] [Medline: 30701448]
4. Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Med 2021 Nov;47(11):1181-1247. [doi: 10.1007/s00134-021-06506-y] [Medline: 34599691]
5. Baelani I, Jochberger S, Laimer T, et al. Availability of critical care resources to treat patients with severe sepsis or septic shock in Africa: a self-reported, continent-wide survey of anaesthesia providers. Crit Care 2011;15(1):R10. [doi: 10.1186/cc9410] [Medline: 21219619]
6. Cardenas-Garcia J, Schaub KF, Belchikov YG, Narasimhan M, Koenig SJ, Mayo PH. Safety of peripheral intravenous administration of vasoactive medication. J Hosp Med 2015 Sep;10(9):581-585. [doi: 10.1002/jhm.2394] [Medline: 26014852]
7. Medlej K, Kazzi AA, El Hajj Chehade A, et al. Complications from administration of vasopressors through peripheral venous catheters: an observational study. J Emerg Med 2018 Jan;54(1):47-53. [doi: 10.1016/j.jemermed.2017.09.007] [Medline: 29110979]
8. Lewis T, Merchan C, Altshuler D, Papadopoulos J. Safety of the peripheral administration of vasopressor agents. J Intensive Care Med 2019 Jan;34(1):26-33. [doi: 10.1177/0885066616686035] [Medline: 28073314]
9. Stolz A, Efendy R, Apte Y, Craswell A, Lin F, Ramanan M. Safety and efficacy of peripheral versus centrally administered vasopressor infusion: a single-centre retrospective observational study. Aust Crit Care 2022 Sep;35(5):506-511. [doi: 10.1016/j.aucc.2021.08.005] [Medline: 34600834]
10. Padmanaban A, Venkataraman R, Rajagopal S, Devaprasad D, Ramakrishnan N. Feasibility and safety of peripheral intravenous administration of vasopressor agents in resource-limited settings. J Crit Care Med 2020 Nov 7;6(4):210-216. [doi: 10.2478/jccm-2020-0030] [Medline: 33200091]
11. Abrar S, Abbas Q, Inam M, Khan I, Khalid F, Raza S. Safety of vasopressor medications through peripheral line in pediatric patients in PICU in a resource-limited setting. Crit Care Res Pract 2022 Mar 31;2022:6160563. [doi: 10.1155/2022/6160563] [Medline: 35402044]

XSL•FO
RenderX

12. Petros A, Melkie A, Kotiso KS, et al. Peripheral line for vasopressor administration: prospective multicenter observational cohort study for survival and safety. PLoS ONE 2025 Oct 13;20(10):e0333275. [doi: 10.1371/journal.pone.0333275] [Medline: 41082535]
13. Parienti JJ, Mongardon N, Mégarbane B, et al. Intravascular complications of central venous catheterization by insertion site. N Engl J Med 2015 Sep 24;373(13):1220-1229. [doi: 10.1056/NEJMoa1500964] [Medline: 26398070]
14. Odendaal J, Kong VY, Sartorius B, Liu TY, Liu YY, Clarke DL. Mechanical complications of central venous catheterisation in trauma patients. Ann R Coll Surg Engl 2017 May;99(5):390-393. [doi: 10.1308/rcsann.2017.0022] [Medline: 28462650]
15. Teja B, Bosch NA, Diep C, et al. Complication rates of central venous catheters: a systematic review and meta-analysis. JAMA Intern Med 2024 May 1;184(5):474-482. [doi: 10.1001/jamainternmed.2023.8232] [Medline: 38436976]
16. Ullman AJ, Marsh N, Mihala G, Cooke M, Rickard CM. Complications of central venous access devices: a systematic review. Pediatrics 2015 Nov;136(5):e1331-e1344. [doi: 10.1542/peds.2015-1507] [Medline: 26459655]
17. Higgerson RA, Lawson KA, Christie LM, et al. Incidence and risk factors associated with venous thrombotic events in pediatric intensive care unit patients. Pediatr Crit Care Med 2011 Nov;12(6):628-634. [doi: 10.1097/PCC.0b013e318207124a] [Medline: 22067813]
18. Östlund Å, Fläring U, Norberg Å, et al. Incidence of and risk factors for venous thrombosis in children with percutaneous non-tunnelled central venous catheters. Br J Anaesth 2019 Sep;123(3):316-324. [doi: 10.1016/j.bja.2019.04.055] [Medline: 31176448]
19. Bai X, Yu W, Ji W, et al. Early versus delayed administration of norepinephrine in patients with septic shock. Crit Care 2014 Oct 3;18(5):532. [doi: 10.1186/s13054-014-0532-y] [Medline: 25277635]
20. Tran QK, Mester G, Bzhilyanskaya V, et al. Complication of vasopressor infusion through peripheral venous catheter: a systematic review and meta-analysis. Am J Emerg Med 2020 Nov;38(11):2434-2443. [doi: 10.1016/j.ajem.2020.09.047] [Medline: 33039229]
21. Owen VS, Rosgen BK, Cherak SJ, et al. Adverse events associated with administration of vasopressor medications through a peripheral intravenous catheter: a systematic review and meta-analysis. Crit Care 2021 Apr 16;25(1):146. [doi: 10.1186/s13054-021-03553-1] [Medline: 33863361]
22. Loubani OM, Green RS. A systematic review of extravasation and local tissue injury from administration of vasopressors through peripheral intravenous catheters and central venous catheters. J Crit Care 2015 Jun;30(3):653.e9. [doi: 10.1016/j.jcrc.2015.01.014] [Medline: 25669592]
23. Bartoli A, D'Angelo A, Ippolito D, Delgado F, Colombo G. The Crystalloid Liberal or Vasopressors Early Resuscitation in Sepsis (CLOVERS) randomized clinical trial. Intern Emerg Med 2023 Nov;18(8):2419-2421. [doi: 10.1007/s11739-023-03419-0] [Medline: 37697214]
24. Munroe ES, Co IN, Douglas I, et al. Peripheral vasopressor use in early sepsis-induced hypotension. JAMA Netw Open 2025 Aug 1;8(8):e2529148. [doi: 10.1001/jamanetworkopen.2025.29148] [Medline: 40864467]
25. Milford K, von Delft D, Majola N, Cox S. Long-term vascular access in differently resourced settings: a review of indications, devices, techniques, and complications. Pediatr Surg Int 2020 May;36(5):551-562. [doi: 10.1007/s00383-020-04640-0] [Medline: 32200406]

Abbreviations

AIC: African Inland Church
CVC: central venous catheter
HDU: high-dependence unit
ICU: intensive care unit
IV: intravenous
KH: African Inland Church Kijabe Hospital
PIV: peripheral intravenous
REDCap: Research Electronic Data Capture

Edited by J Sarvestan; submitted 03.Aug.2025; peer-reviewed by AC Faust, J Waterson; accepted 15.Jan.2026; published 05.Mar.2026.

Please cite as:
Gao GH, Kimani B, Lepore L, Aghababa Z, Gitura H, Kinuthia M, Musau J, Muthoki J, Mwonge A, Ndibaru F, Ndubi AV, Ndungu J, Nixon K, Odhiambo M, Okari W, Thairu B, Vihenda L, Wojcik RM, Brotherton BJ, Rudd KE
Safety of Administration of Vasopressors Through Peripheral Compared to Central Venous Catheters in a Rural Kenyan Hospital: Protocol for a Prospective Observational Cohort Study
JMIR Res Protoc 2026;15:e81794
URL: https://www.researchprotocols.org/2026/1/e81794
doi:10.2196/81794

XSL•FO
RenderX

© Grace Hezi Gao, Brian Kimani, Linette Lepore, Zahra Aghababa, Hannah Gitura, Michael Kinuthia, Jack Musau, Judy Muthoki, Alice Mwonge, Fredrick Ndibaru, Ann Viola Ndubi, Juliah Ndungu, Kinzi Nixon, Moses Odhiambo, Winslet Okari, Benjamin Thairu, Lilian Vihenda, Rachel M Wojcik, B Jason Brotherton, Kristina E Rudd. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Novel Social Network Approach to Measure Intersectional Stigma Among Latino Men Who Have Sex With Men in San Diego, California (NEXUS): Protocol for a Longitudinal Cohort Study

Laramie R Smith[1], PhD; Angel B Algarin[2], MPH, PhD; Eileen V Pitpitan[3], PhD; Heather A Pines[4,5], MPH, PhD; Nicole K Kelly[1], MPH, PhD; Carl Latkin[6], PhD; Aaron Gutierrez[1], BA; Maryam Hussain[1], PhD; Francisco Soto[1], MSW; Ricardo Vazquez Jr[7], BA; Stevie Juarez[8], AA; Albert Genegaling[1], BA; Rosalinda Rodriguez[1], BA; Anthony Cirilo[9], BA; Juan Esparza[8], BA; Saul Cruz[1], BS; Erick Jimenez[8]; Ulises Reyes[1], BA; Veronica Moore[7], MPA; Malek Guerbaoui[8], MPH; Beth Davenport[8], MBA, LCSW, DSW; Jeannette Aldous[7], MD; Katherine Penniga[7], LCSW; Kenyatta Parker[8], BA; Britt Skaathun[1], PhD

[1]Department of Medicine, University of California, 9500 Gilman Drive, #0507, San Diego, CA, United States

[2]Center for Health Promotion and Disease Prevention, Arizona State University, Tempe, AZ, United States

[3]School of Social Work, San Diego State University, San Diego, CA, United States

[4]Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, CA, United States

[5]School of Public Health, San Diego State University, San Diego, CA, United States

[6]Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

[7]San Ysidro Health, San Diego, CA, United States

[8]The San Diego LGBT Community Center, San Diego, CA, United States

[9]Rollins School of Public Health, Emory University, Atlanta, GA, United States

Corresponding Author:
Laramie R Smith, PhD
Department of Medicine, University of California, 9500 Gilman Drive, #0507, San Diego, CA, United States

Abstract

Background: Latino men who have sex with men (LMSM) account for a disproportionate and growing number of HIV diagnoses in the United States. Intersectional stigma remains a key driver of HIV inequities; however, most quantitative intersectional stigma measures are limited and do not consider the larger social context.

Objective: NEXUS is a longitudinal cohort study that will use social network methods and theory to rigorously measure intersectional stigma among LMSM and quantify the longitudinal association between intersectional stigma and HIV prevention outcomes.

Methods: We will prospectively enroll 500 HIV-negative LMSM in San Diego, California, and follow participants over 1 year. At baseline and every 6 months thereafter (Month 0, Month 6, and Month 12), participants will complete an interviewer-administered social network inventory and a self-administered survey to collect information on their social networks (alter types, size, and characteristics) and HIV prevention engagement (HIV testing and pre-exposure prophylaxis use), respectively. Information on HIV prevention engagement will also be abstracted from medical records. Intersectional stigma will be operationalized as a multilevel latent variable comprised of observed measures of anticipated and enacted stigma experienced by a participant from an alter toward the participant's Latino, masculine, and sexual identities. Multilevel structural equation modeling will be used to estimate the longitudinal association between intersectional stigma, HIV testing, and pre-exposure prophylaxis use, considering potential mediators and moderators.

Results: NEXUS recruitment began in June 2021, and as of March 11, 2025, a total of 482 participants had been enrolled. Enrollment is planned to end by May 2025, with baseline results expected late 2025 and through the following year. Data collection for our prospective study aims is expected to be complete in June 2026, with data analysis and expected results published later that year.

Conclusions: NEXUS will advance quantitative intersectional stigma measurement using a novel social network approach. This study will identify intervention targets to reduce HIV inequities among LMSM and mitigate the harms of intersectional stigma in this population.

International Registered Report Identifier (IRRID): DERR1-10.2196/72334

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e72334) doi:10.2196/72334

KEYWORDS

intersectional stigma; social networks; latino MSM; HIV prevention; HIV stigma; PrEP use; HIV testing; men who have sex with men; pre-exposure prophylaxis

Introduction

Inequities in HIV acquisition among Latino men who have sex with men (LMSM) are increasing in the United States [1-3]. From 2018 to 2022, the number of HIV diagnoses among US LMSM rose dramatically (22% increase), despite declining or remaining stable among men who have sex with men of other ethno-racial groups (non-Hispanic White: 5% decrease, multiracial: 32% decrease, and non-Hispanic Black: no change) [4]; underscoring the need for targeted prevention efforts in this multiply marginalized group. The US HIV epidemic among LMSM is primarily localized in 4 states—California, Texas, Florida, and New York—with the largest number of new HIV diagnoses occurring in California [2]. San Diego County, California, in particular, has been named one of 50 priority jurisdictions by the US Ending the HIV Epidemic initiative [5]. Together, these jurisdictions account for over 50% of new HIV diagnoses in the country [5]. Despite biomedical advances in HIV prevention (eg, pre-exposure prophylaxis [PrEP]) [6], substantial barriers remain and reinforce existing inequities in prevention for LMSM, especially in high-incidence areas such as San Diego County [7,8]. Stigma is one such barrier that greatly undermines efforts to reduce HIV incidence through many different pathways [9], such as by impairing mental health and influencing substance use. While this is true of all populations, the effects of stigma on HIV are even more deleterious among multiply marginalized groups, such as LMSM [8,10-12].

Stigma is a social process that maintains health inequities, often experienced as anticipated (expecting unjust treatment), enacted (experiencing unjust acts perpetrated by others), and internalized (applying negative stereotypes, beliefs, and feelings toward oneself) stigma mechanisms [9,13,14]. However, stigma related to a marginalized component of one's identity or characteristics rarely occurs in isolation. Instead, people often experience multiple, overlapping forms of stigma (ie, intersectional stigma) via stigmatizing interactions [15,16]. Applied to HIV prevention with LMSM, intersectional stigma reflects how stigma toward multiple stigmatized identities—ethnicity, masculinity, and sexuality—may interact with HIV stigma in diverse social contexts (eg, social relationship types and identity expression) [15-17]. For example, LMSM may contend with racism from non-Latino sexual partners and homophobia and machismo (ie, historical and cultural expectations of men taking pride in and performing hypermasculine gender roles) [18] from Latino family members, increasing fear of further rejection if one acquires HIV. LMSM also vary in the degree to which they present these identities to others (eg, acculturation, outness, and masculinity), influencing the degree to which they might experience intersectional stigma. Despite recognition that the social context of intersectional stigma matters, these contexts are rarely measured when evaluating stigma as a barrier to HIV prevention [19].

Despite the substantial focus that HIV stigma has received in the literature, when an intersectionality lens is applied, it is too often used to describe HIV stigma experiences within historically marginalized populations. This work typically does not quantitatively assess or contextualize the effects of intersecting stigmas on the health of these populations. It often relies on analytical approaches that are not grounded in intersectionality theory (eg, using interaction terms in regression models) [20-22]. Experts emphasize that intersectional stigma has been primarily studied via qualitative methods, which are important for illustrating the lived experience of intersectional identities, but point to the promise that quantitative methods have in intersectionality research [19,23]. While there is no single best practice to measure and analyze intersectional stigma, methods that can assess the magnitude and effects of intersectional stigma on health behaviors, such as HIV testing and PrEP use, are needed to advance the development of intersectional stigma-responsive interventions and improve HIV outcomes [19,23].

Social network methods offer a promising and innovative approach to measuring intersectional stigma within the larger social context and provide an opportunity to shift the focus of stigma from the individual level toward upstream social factors. In particular, social network methods measure social contexts by examining how individuals (ie, LMSM) interact with and are affected by the people in their social networks (ie, "alters," such as family, friends, sexual partners, and providers) [24]. Within social networks, stigma toward multiple identities is produced and experienced in relation to others (social network intersectional stigma exposure) [13]. Social network intersectional stigma mechanisms are important intervention targets, since alters' social response to intersectional identities can (1) stigmatize LMSM (ie, intersectional stigma sources) and (2) foster LMSM's resilience, reducing the negative impact of stigma (eg, identity affirmation and social support) [25]. Social network methods can help us learn about the social contexts where stigma toward single or multiple stigmatized identities interacts to affect health behaviors by provoking a deleterious or protective response [25].

Quantitatively measuring intersectional stigma is a well-documented challenge, and most methods either largely ignore the broader social context or fail to capture its nuances. However, stigma is fundamentally a social process, so incorporating social context is critical to understanding the complex interplay of these dynamics and to developing effective stigma-reduction interventions. Yet, to our knowledge, no studies have used a social network approach to measure intersectional stigma. We are therefore conducting a theory-informed, longitudinal cohort study of LMSM in San Diego County, California, called NEXUS—named to reflect

XSL•FO
RenderX

the central point at which these social processes and contexts intersect. NEXUS aims to (1) identify how social network intersectional stigma manifests among LMSM by alter type (eg, family and sexual partner) and the degree to which LMSM express their marginalized identities (eg, acculturation, outness, and masculinity); (2) examine the longitudinal associations between social network intersectional stigma, HIV testing, and PrEP use; (3) explore network- and individual-level factors that may mediate (eg, social support and identity salience) and moderate (eg, network density and age) the association between social network intersectional stigma exposure on HIV testing and PrEP use; and (4) explore whether social network intersectional stigma exposure influences network stability (Aim 4A) and if changes in network stability are associated with increases in HIV testing and PrEP use (Aim 4B; Figure 1).

Figure 1. Visual representation of the NEXUS study aims 1 - 4. Green color denotes affirming alter and red denotes stigmatizing alter. PrEP: pre-exposure prophylaxis; SN: social network.

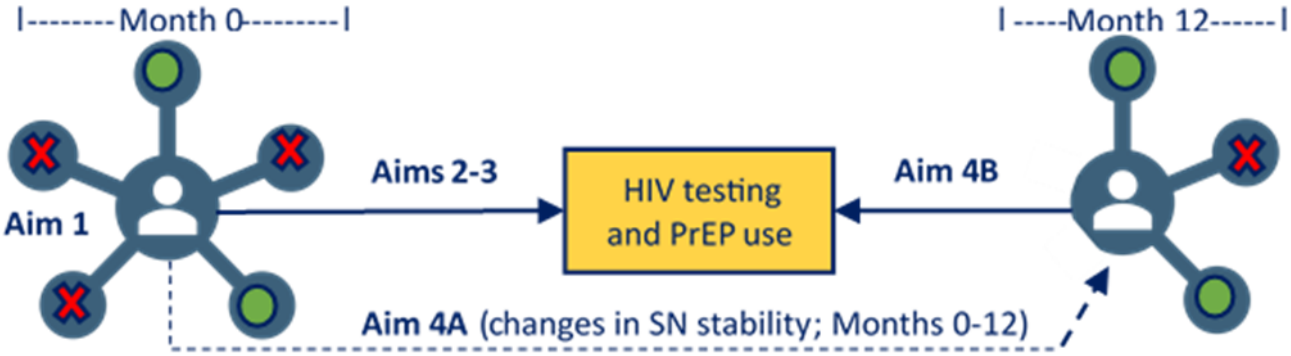

Methods

Study Design

NEXUS is a longitudinal, prospective cohort study that will aim to enroll 500 Latino gay, bisexual, and other men who have sex with other men (ie, LMSM) in San Diego, California, who are HIV negative (Figure 2). Our community-based recruitment will extend to urban lesbian, gay, bisexual, and transgender, and queer (LGBTQ) communities and exurban, predominantly Latino communities. Following a baseline rapid HIV test, HIV-negative participants will complete 3 study visits: baseline (Month 0), Month 6, and Month 12 postbaseline. A certified HIV tester and counselor will conduct baseline HIV tests as part of routine HIV prevention services offered in San Diego County. A trained bilingual (English-Spanish) and bicultural interviewer will lead study visits, and data will be collected electronically. At each study visit, we will collect self-report data captured through network- and individual-level survey methods (ie, a social network inventory [SNI] and Qualtrics survey [Qualtrics Software]) to study the longitudinal associations between intersectional stigma within egocentric social networks. HIV testing and PrEP use outcomes are obtained via self-report at each study visit and via medical records abstraction for the time between Month 0 and Month 12. As described in the “Data Management and Analysis” section below, data will be analyzed prospectively using multilevel structural equation modeling (MSEM). All study measures and protocols were first vetted by our HIV service community partners and bilingual LMSM who are also part of our study team and community advisory board (CAB), before piloting with 40 LMSM, to ensure they were acceptable and feasible to implement with the study cohort. Our prospective cohort study design will establish the degree to which LMSM are exposed to intersectional stigma within their social networks and how that exposure varies across different social relationships and contexts to affect HIV prevention behaviors over time.

XSL•FO
RenderX

Figure 2. NEXUS study protocol. MSM: men who have sex with men; SNI: social network inventory.

XSL•FO
RenderX

Study Population

Eligible participants are LMSM in San Diego County, California, who meet the following criteria: (1) are at least 18 years of age, (2) identify as cisgender men, (3) identify as Latino or Hispanic, (4) are fluent in English or Spanish, (5) report sexual activity with men in the past 12 months, and (6) are HIV negative. We will aim to oversample LMSM aged 18‐44 years, among whom most incident HIV acquisitions are diagnosed [1]. The study will collaborate with 2 community partners who provide bilingual HIV services (the San Diego LGBT Community Center and a local HIV service provider) to enroll those living in geographically diverse regions of San Diego County (Figure 3). The San Diego LGBT Community Center is a large community-based organization located in Central San Diego County, a predominantly lesbian, gay, bisexual, transgender, and queer–identified urban area. The local HIV service provider is a large federally qualified health center with clinical sites primarily located in South and Southeast San Diego County. These sites are in proximity to the US-Mexico border and primarily serve patients in exurban Latino communities.

Figure 3. Map of San Diego County, California Health and Human Services Agency regions.

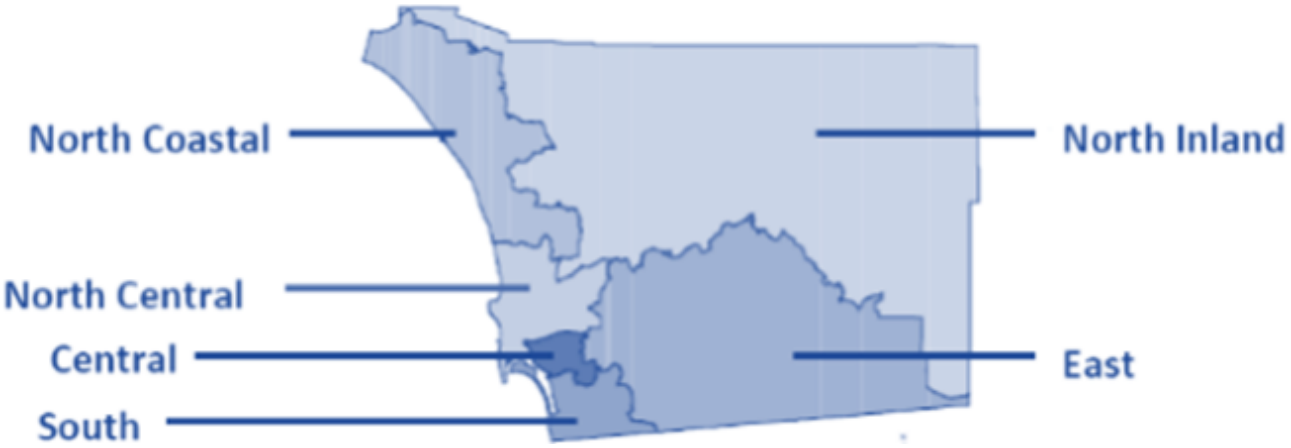

An Integrated Theory-Informed Social Network Approach to Intersectional Stigma

Overview

In conceptualizing the NEXUS Study, we first articulated an integrated theory-informed social network approach to intersectional stigma, building on work by Pescosolido and Manago [26], who used contact theory [27] to articulate how stigma is likely experienced within social networks via 3 types of social processes (identity, influence, and affiliation). Social stigma [9,13,28], intersectionality [15,16,29], and social network [24,25] theories are inherently related, yet the linkages between them have not been well articulated. By articulating the integrated social stigma, social network, and intersectionality theories across these 3 social processes, we lay the foundation for NEXUS, a novel approach to assessing intersectional stigma as a socially networked phenomenon.

Identity Process

In brief, the identity process will specify from whom in the social network (intersectional stigma sources) anticipated and enacted stigma or affirmation can be experienced toward LMSM. The method will involve asking about each stigma (or affirmation) source (eg, family member, sexual partner) for each intersecting “marked” or socially devalued status (ie, ethnicity, masculinity, and sexuality; Figure 4, bottom half in blue) or “unmarked” stigmatized status they could acquire such as HIV (Figure 4, top half in green). Internalized stigma reflects the degree to which LMSM might turn the social devaluation process inward and self-stigmatize their ethnic, masculine, or sexual identities [9,26]. Stigma theory further specifies that the degree to which marked characteristics are visible (eg, skin tone and spoken language) or potentially concealable (eg, sexuality and gender expression) further influences how stigmatized statuses are experienced in social interactions [30-32]. NEXUS will focus on 3 specific intersectional identities (Latino, masculinity, and sexuality). These identities were selected because the scientific evidence linking ethnic-, gender-, and sexuality-related stigma to poorer HIV prevention outcomes among LMSM is strongest [33-38]. Additionally, this decision balances concerns of participant burden, as the addition of each identity exponentially increases the number of survey items, while selecting identities that are maximally inclusive of LMSM. However, an expanded intersectional identity focus is important for future programmatic research, once precedent for our novel social network approach is established.

XSL•FO
RenderX

Figure 4. An integrated theory-informed social network approach measuring intersectional stigma among Latino men who have sex with men (LMSM) in the NEXUS study. ISS: intersectional stigma; PrEP: pre-exposure prophylaxis; SN: social network.

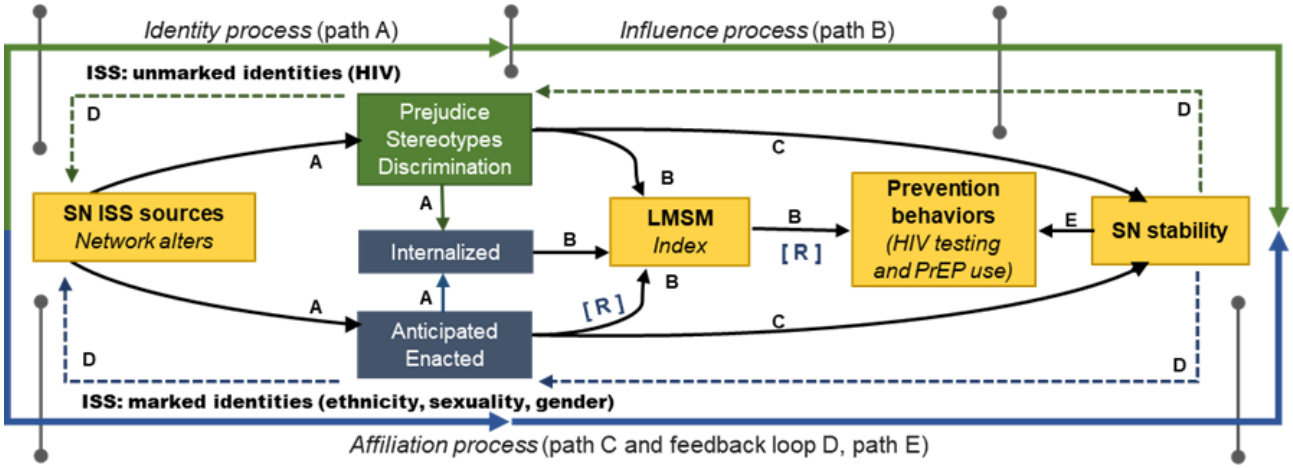

Influence Process

The influence process will specify the association between anticipated and enacted intersectional stigma (via social interactions with network members) with LMSMs' HIV prevention behaviors (ie, HIV testing and PrEP use) over time (Figure 4, path B) [26,30]. In the context of HIV research, an intersectionality lens is often applied to ensure co-occurring social statuses are observed at their distinct intersections and not in isolation from each other [15,39,40]; yet, approaches to this work vary [19]. Following guidance by Bowleg [41], many approach intersectionality as a blended identity (eg, stigma experienced as a gay-Latino-man). In this approach, the contributions of being Latino cannot be parsed out because it fundamentally shapes how life is experienced as a male sexual minority [42]. Yet, research on social identity centrality and salience suggests individuals may not experience these interlocking stigmatized identities equally in all social relationships or contexts [43].

Adaptive Process

Social network methods will enable NEXUS to model intersectional stigma exposure simultaneously across co-occurring social statuses to identify from whom and in what contexts intersectional stigma exposure affects future HIV prevention behaviors, while attending to both network-level factors (eg, social support from alters for HIV testing and network density) and individual-level factors (eg, identity salience and age) that may affect how (mediation) and for whom (moderation) exposure to social network intersectional stigma affects our outcomes of interest [26,44]. Relatedly, we articulate resilience as an adaptive (mediation) process [45] that may be experienced through interactions with network members who affirm and support traditionally stigmatized statuses or through an individual's ability to rebound from and resist adversity. As such, we specify that the resilience process may facilitate LMSMs' engagement in HIV testing and PrEP use despite high levels of social network exposure to intersectional stigma (Figure 4, path [R]) [46-48].

Affiliation Process

Relatedly, the affiliation process illustrates a dynamic social phenomenon through which we will articulate that social network exposure to intersectional stigma may affect the stability, density, and composition of the social network over time (Figure 4, path C) through the addition or removal of alters that affirm or reject LMSMs' socially stigmatized statuses and thereby influence future social network exposure (Figure 4, feedback loop D) [26,44]. It will also reflect whether such changes in network stability are associated with increased HIV testing and PrEP use (Figure 4, path E) [26,49,50].

Participant Procedures

Recruitment

We will use multiple bilingual nonrandom purposive sampling strategies to enroll the NEXUS cohort, including in-reach among clients and patients seeking services through our 2 community partners in San Diego County, California. Outreach strategies will include tabling at local community events and bars frequented by LMSM and using various social media platforms (eg, Instagram [Meta] and Facebook [Meta]) and dating apps (eg, Grindr [Grindr Inc] and Scruff [Perry Street Software]) with profiles tied to the NEXUS study and community partners.

Enrollment

Bilingual study staff will screen interested individuals to determine eligibility and, if needed, provide access to existing HIV testing services to confirm HIV seronegative status up to 14 days before the first study visit. Written informed consent, locator information to facilitate retention, and signed consent for a HIPAA (Health Insurance Portability and Accountability Act) authorization waiver and a medical release of information form will be collected at enrollment.

Study Visits

NEXUS participants will complete up to 3 study visits at Month 0 (baseline), Month 6, and Month 12, during which they will complete an interviewer-administered SNI via Network Canvas (Complex Data Collective) [51] and a self-administered individual-level survey via Qualtrics (Figure 2). While these visits were intended to be in person to allow for participants and interviewers to construct an interactive visual of the participants' social network via Network Canvas on a large touch screen, NEXUS launched during the COVID-19 pandemic, necessitating the development of a virtual study visit

XSL•FO
RenderX

protocol via Zoom (Zoom Communications, Inc). As physical distancing restrictions eased, participants had the option of in-person study visits. Baseline study visits typically last 2.5 - 3 hours, while follow-up study visits last 1 - 2 hours, depending on the participants' social network size. Participants will be compensated as follows: Month 0: US $100, Month 3 (retention call; detailed below in the "Cohort Maintenance Strategies" subheading): US $10, Month 6: US $115, Month 9 (retention call): US $10, and Month 12: US $125.

Data Collection Procedures

Overview

NEXUS will collect 3 primary types of data, detailed below in the "Social Network Inventory, Individual-level Survey, and Medical Records" sections, from cohort participants as outlined in Table 1. This includes an interviewer-administered SNI, a self-administered individual-level Qualtrics survey, and medical records data to document participants' prospective HIV testing and PrEP use behaviors.

Table . NEXUS data collection sources and measures.

Source	Measure
Social network inventory	• Social network name generators (alters) • Network characteristics (density, size, and stability) • Alter sociodemographics (age, gender, race, ethnicity, and sexuality) • Alter relationship history (type, duration, frequency, and locations) • Alter sexual history (as applicable) • Identity awareness (ethnicity, masculinity, and sexuality) • Anticipated intersectional stigma (ethnicity, masculinity, and sexuality) • Avoidance of alters due to intersectional stigma (ethnicity, masculinity, and sexuality) • Enacted intersectional stigma (ethnicity, masculinity, and sexuality) • Other intersectional stigma exposure • Intersectional affirmation and support from alters (ethnicity, masculinity, and sexuality) • HIV testing and pre-exposure prophylaxis (PrEP) support • COVID-19–related changes in social interactions
Individual-level Qualtrics survey	• Sociodemographic information • Latino identity (pride, centrality, salience, expression, and internalized stigma) • Masculine identity (pride, centrality, salience, expression, and internalized stigma) • Sexuality identity (pride, centrality, salience, expression, and internalized stigma) • HIV prevention behaviors • HIV and PrEP stigma • Structural stigma • Substance use behaviors • Physical health • Mental health • Social desirability
Medical records	• HIV testing • PrEP visits

Piloting

In February 2021, before piloting data collection protocols and study measures, a community forum with simultaneous Spanish translation was held via Zoom, inviting potential participants and HIV prevention service providers to provide feedback on typical social network protocols to optimize the confidentiality of the network alters, while still allowing participants to identify and recall individual alters in the NEXUS SNI throughout each study visit and across follow-up study visits (described below in the "Social Network Inventory" section). From March to May 2021, we piloted our NEXUS study protocols, measures, and recruitment materials to iteratively optimize comprehension, reduce participant burden, minimize social desirability bias, and refine the linguistic and cultural relevance among 40 LMSM (21 English and 19 Spanish) using cognitive interviewing techniques where participants "think aloud" their interpretation of study materials and response process [52]. This allowed interviewers to document where revisions to study materials needed to be made. This one-time interview lasted up to 3 hours, and we alternated, starting with social network and individual-level measures to determine the best process for delivering the study protocol and eliciting network alters that are sources of anticipated and enacted intersectional stigma. From these procedures, we determined that the SNI should precede the individual-level survey. NEXUS pilot participants were remunerated US $100 for their participation.

XSL•FO
RenderX

Social Network Inventory

Trained bilingual interviewers will administer the SNI first using an interactive open-source software, Network Canvas [51]. Social networks will be enumerated using an egocentric network approach that ascertains information from participants on each member of their social network (ie, alters) [53]. We developed 7 name generators (ie, questions that probe participants to name people in their network) to identify alters that participants have interacted with in the past 6 months related to (1) social, (2) sexual, and (3) substance use interactions, as well as (4) emotional and (5) informational support that may increase exposure to HIV acquisition and HIV prevention, and (6 and 7) conflictual ties to capture interactions with alters who may not otherwise be named due to stigma (Table 2). We developed probes for each name generator to enhance the recall of alters. Interviewers were further trained to ask key questions if participants were having trouble naming anyone or listed fewer than 5 people for any given name generation, for example, asking, "Who else can I add?" and "What about people you interact with in a typical week?" Using the protocol developed in our community forum, which is typical of social network protocols and refined through the pilot phase, each alter will be identified by a first name or nickname that remains visible in Network Canvas for the rest of the interview, aiding recall of alters across study visits. To help distinguish between alters with similar first names and further support recall at follow-up visits, the first 4 letters of alters' last names will be entered, but will not be visible unless that information needs to be accessed to aid in recall. Participants can also choose to add descriptors next to names that will be noted in parentheses (eg, coworker 1). If a participant does not know or declines to provide an alter's first name or first 4 letters of the last name, NEXUS interviewers will be trained to enter "(9999)" and to have a brief discussion about how to refer to this alter in conversations and to input this information in the interview's notes section programmed into Network Canvas. Information on the NEXUS alter generation protocol, including additional probes and Spanish translations, is available in Multimedia Appendix 1.

Table . Alter domains and name generator prompts used in the social network inventory among Latino men who have sex with men in the NEXUS study.

No	Alter domain	Name generator prompt
1	Social	"With whom did you hang out with or socialize regularly in person or through technology?"
2	Sexual	"With whom did you have oral, anal, or vaginal sex that was 18 years old or older?"
3	Substance use	"With whom did you drink alcohol with or use drugs with more than once?"
4	Emotional support	"With whom did you talk to about the MOST private and personal things in your life?"
5	Informational support	"From whom did you get advice about your health, HIV, or PrEP[a]?"
6	Conflictual tie: rejection	"Who makes comments (even as a joke) that make you think or feel like they don't accept you for who you are?"
7	Conflictual tie: avoidance	"Who might you want to avoid because they say or do things that make you feel uncomfortable or unwelcome?"

[a]PrEP: pre-exposure prophylaxis.

Participants will then complete a sociogram to identify which alters know each other to characterize network density (ie, connectivity) [53]. If participants name more than 5 alters for a domain during the name generator process, we will ask them to identify the top 5 people who are most important to them for each domain. For the top 5 alters for each name generator, we will elicit key social network characteristics, including: sociodemographics (age, gender, race, and ethnicity), sexual orientation, relationship type (family, friend, sexual partner, health care provider, and other), relationship strength (very close, somewhat close, no longer close, and never close), proximity (resides in the San Diego-Tijuana border region), and perceived or known HIV and PrEP status of alter (Table 1). We will then assess the characteristics of social interactions with each alter by eliciting the structure (typical activities), frequency, and locations or venues of typical social interactions. Finally, we will assess the context of social interactions with alters related to anticipated and enacted intersectional stigma related to their ethnicity, masculinity, and sexuality, as well as other social reactions related to these socially stigmatized statuses (eg, alters affirming LMSM's socially stigmatized statuses as an important part of who they are and LMSM avoiding alters because of their stigmatizing behavior), and social support for HIV testing and PrEP behaviors.

At follow-up visits, we will assess social network stability over time by collecting information for each alter named at baseline who was not named at subsequent visits, and each new alter added at follow-up visits to identify reasons for (1) not naming previous alters at follow-up and (2) not interacting with these alters at follow-up. For the top 5 alters named for each domain at follow-up, we will again assess social network characteristics, as well as social interaction characteristics and context as

XSL•FO
RenderX

described above. Refer to Multimedia Appendix 2 for items assessing social network stability. Network Canvas is not designed, at this time, to support longitudinal social network data collection by retaining alters across interview protocols. As such, we developed a NEXUS longitudinal social network data integration protocol with multiple verification checks to generate alter reports from previous surveys to be input into Network Canvas by the study team ahead of subsequent study visits. In brief, this protocol is completed in three steps: (1) use SNI data from the previous interview to generate an individual alter report for each participant (ie, using Month 0 SNI data or Month 6 SNI data), (2) enter data from the alter report ahead of the participants' upcoming scheduled interview (ie, Month 6 interview or Month 12 interview), and (3) at each follow-up visit with participants, assess what alters left or entered the network over time. Detailed instructions and best practices learned through the development and implementation of the NEXUS longitudinal social network data integration protocol, including sample syntax for running the alter reports, are provided in Multimedia Appendix 2 and can be easily adapted by other researchers using Network Canvas in longitudinal study designs to ensure data integrity while tracking alters across multiple study visits.

Individual-Level Survey

The individual-level survey will be conducted through Qualtrics and will collect a range of information, including sociodemographic factors, identity measures (Latino, masculinity, and sexuality), health behaviors, internalized stigma, and substance use. A full list of survey domains is provided in Table 1 and described in detail in the "Measures" subheading below.

Medical Records

Signed HIPAA authorization waiver and release of medical information forms specifically indicate participants agree to and request the release of HIV testing and PrEP-related information as part of their study participation. NEXUS participants will be able to enroll in the study if they decline to sign these forms, but, based on experience, we anticipate that few (<2% of participants) will opt out. These forms will facilitate access to HIV testing and PrEP visit data abstracted from participants' medical records for 12 months following the NEXUS baseline visit via secure fax or encrypted email per the medical providers' protocols. If no HIV testing or PrEP data are available or the participants' records cannot be located by the provider, the provider will send documentation indicating these results, and we will record this response as "no medical records data available." To offset the risks of missing medical records data, HIV testing and PrEP data will also be collected via self-report at each study interaction (Figure 2: Months 0, 3, 6, 9, and 12). Details on each of these outcomes are provided in the "Measures" subheading, but in brief, information on HIV testing date, HIV testing results, PrEP prescription coverage (dates of injection and dose provided or dates of pills prescribed, number of refills for pills, and number of pills per refill) and PrEP clinical care (visit dates with a PrEP-prescribing provider, laboratory tests performed, and laboratory test results) will be abstracted. The medical records data will be immediately deidentified, abstracted into a secure database, and linked to the participants' unique study ID before analysis.

Cohort Maintenance Strategies

To facilitate retention, study staff will contact participants by phone at Month 3 and Month 9, three months after their first and second study visit, respectively (Figure 2). Retention calls will update participant locator information and assess whether HIV testing and PrEP use occurred and where those services were accessed since the last study visit.

Measures

Overview

All measures will be captured at Month 0 (baseline), Month 6, and Month 12 study visits unless otherwise specified (Table 3). All measures will be available in both English and Spanish. When a validated Spanish version of a measure is not available, the measure will be translated from English to Spanish by a bilingual, bicultural Spanish speaker and back-translated by a second bilingual, bicultural member of the study team who is native to Mexico. Discrepancies will be discussed and resolved via consensus with a third bilingual Spanish speaker. The Spanish translation will then be reviewed by multiple bilingual individuals from both community partners, and refinements will be made through a working group meeting. Items related to stigmatized identities have been refined as needed to be culturally responsive to LMSM via feedback and discussions with LMSM from our study CAB, community partners, and the NEXUS study team. When needed, adaptations were made to ensure clarity or cultural responsiveness. Any adaptations made to these measures are noted below under the "Measures" subheading. Unless otherwise specified, measures will be available upon reasonable request from the corresponding author. Once measures are validated in the NEXUS sample, the English and Spanish versions will be published in the corresponding paper.

XSL•FO
RenderX

Table . Analytical measures and data sources by the NEXUS study's primary aims.

Measure type and name	Source	Aims
Exposures		
Intersectional stigma	Social network inventory	1 - 4
Alter type	Social network inventory	1
Identity expression: L[a], M[a], and S[a]	Qualtrics survey	1
Outcomes		
HIV testing	Medical records and Qualtrics	2 - 4
PrEP[b] use	Medical records and Qualtrics	2 - 4
Social network stability	Social network inventory	4
Mediators[c]		
Resilience	Qualtrics survey	3
Medical mistrust	Qualtrics survey	3
Identity pride: L, M, and S	Qualtrics survey	3
Identity centrality: L, M, and S	Qualtrics survey	3
Identity salience: L, M, and S	Qualtrics survey	3
Internalized stigma: L, M, and S	Qualtrics survey	3
HIV stigma	Qualtrics survey	3
PrEP stigma	Qualtrics survey	3
Moderators[c]		
Network size	Social network inventory	3
Network density	Social network inventory	3
Network composition	Social network inventory	3
Nativity	Qualtrics survey	3
Local community affiliation	Qualtrics survey	3
Age group	Qualtrics survey	3
Covariates[d]		
Age	Qualtrics survey	1 - 4
Race	Qualtrics survey	1 - 4
Education	Qualtrics survey	1 - 4
Income	Qualtrics survey	1 - 4
Years living in San Diego County	Qualtrics survey	1 - 4
Alcohol use	Qualtrics survey	1 - 4
Substance use	Qualtrics survey	1 - 4
Depression	Qualtrics survey	1 - 4
Anxiety	Qualtrics survey	1 - 4
HIV knowledge	Qualtrics survey	1 - 4
PrEP willingness	Qualtrics survey	1 - 4
HIV testing frequency	Qualtrics survey	1 - 4
PrEP use history	Qualtrics survey	1 - 4

[a]L: Latino, M: masculinity, and S: sexuality.

[b]PrEP: pre-exposure prophylaxis.

[c]Mediators and moderators as hypothesized.

XSL•FO
RenderX

[d]Covariates for each aim are measured at the individual level and will be identified from a directed acyclic graph, theory, or the extant literature.

Primary Exposures

Intersectional Stigma

Social network intersectional stigma will be operationalized as a multilevel latent variable comprising 6 observed scores that are specific to anticipated and enacted stigma mechanisms experienced from an alter toward a participant's Latino, masculine, and sexual identities. Each observed score will reflect the mean response (at baseline) or individual response (at follow-up) given for each alter in response to the 6 mechanism-specific stigma exposures. By modeling observed scores from each alter, we can maximize variance in exposure to stigma across all alters, rather than truncating this variance into a composite score averaged across all alter types (eg, a mean score referencing all "family and friends"), as is common in traditional survey research. This approach allows us to capture how the magnitude, significance, and direction of association for these 6 mechanism-specific stigma exposures vary when simultaneously modeled in relation to different stigma sources (ie, family, friends, sexual partners, providers, and other alters), as well as participants' self-reported identity expression (ie, the degree to which LMSM report their interests, attitudes, behaviors, and appearances as being typically Latino, masculine, or gay). Social network intersectional stigma will be measured using the SNI data. The NEXUS social network intersectional stigma items, response options, and Spanish translations are provided in Multimedia Appendix 3.

At baseline, anticipated intersectional stigma exposure will be assessed by asking, "In the future, how likely it is that each alter will 'stereotype you,' 'judge you or look down on you' (prejudice), and 'treat you poorly or avoid you' (discrimination) for each intersectional stigma identity" (3 items each and 9 items total). Responses will be recorded on a 5-point Likert-type scale (1=very unlikely, 2=unlikely, 3=it's possible, 4=likely, and 5=very unlikely). Baseline enacted intersectional stigma exposure will be assessed by asking, "In the past, how often has each alter 'stereotyped you' 'judged you or looked down on you' (prejudice), and 'treated you poorly or avoided you' (discrimination) for each identity" (3 items each and 9 items total). Responses will be recorded on a 5-point Likert-type scale (1=never, 2=not often, 3=somewhat often, 4=often, and 5=very often).

For Month 6 and Month 12, we refined the measures to ask a single anticipated stigma item for each stigmatized identity (3 items total) and a single enacted stigma item for each stigmatized identity (3 items total). The single item asks how likely, in the future, each alter (anticipated), and how often, in the past, has each alter (enacted) "stereotyped you, judged you, treated you poorly, or avoided you" for each marginalized identity. No changes were made to the response options. While this change will limit our ability to compare the same social network intersectional stigma items across time, it was deemed necessary to reduce the SNI response burden following the baseline assessment and ensure sufficient retention at study follow-up. Specifically, each SNI item is asked for each alter; a difference of 18 total items per alter at baseline vs 6 total items per alter at follow-up can dramatically affect response burden, especially for larger networks. A direct comparison of social network intersectional stigma items across assessment points would be further limited by changes in social network composition over time, as individual alters may leave or enter the network at various time points. Notably, this change does not affect the primary study aims, which focus on the prospective associations between baseline social network exposure to intersectional stigma among LMSM and their future HIV testing and PrEP use. However, this change will allow us to compare the precision with which social network intersectional stigma exposure can be measured with multiple items per intersectional stigma mechanism compared with a more general single item per intersectional stigma mechanism in cross-sectional analyses. Future results can better inform under which circumstances greater measurement precision (and participant response burden) is recommended, enabling future studies to make an informed decision when selecting the multi-item or single-item social network intersectional stigma measurement approach that can be compared across time.

Stigma Sources

We will examine alter type to assess how social network intersectional stigma varies across diverse social relationships. Alter type will be modeled as a continuous observed variable for the proportion of alters identified as a specific alter type (eg, the number of family members out of the total network size) in the SNI. Alter characteristics will be assessed in the SNI. Alters will be categorized as (1) family, (2) friends, (3) sexual partners, or (4) other. Family members will be further categorized as parent, sibling, aunt, uncle, cousin, grandparent, nonblood family (including chosen family and family by marriage), or other. Sexual partners will be further categorized as someone in a serious relationship, someone you dated, a casual sexual partner (including "fuckbuddy" and "friends with benefits"), a one-night stand, or anonymous. Alters, first described as "other," will be further categorized as a medical provider, therapist, neighbor, coworker, religious group member, or other.

Identity Expression

We will collect information on identity expression to assess how social network intersectional stigma varies in relation to LMSM's expression of their intersectional identities. Identity expression will be modeled as continuous variables (mean scores) derived from Qualtrics responses related to the degree to which LMSM report their (1) self-view, (2) ideal-self, (3) interests, (4) attitudes, (5) behaviors, and (6) appearance as being typically Latino (ethnicity), masculine, and gay (sexuality). These items were adapted for ethnicity and sexuality from the 6-item Traditional Masculinity-Femininity scale [54]. A seventh item was added from Rider et al [55]: "Other people would describe me as…." to capture how LMSM perceive how others typically view them. Responses will be scored on a single bipolar 7-point Likert scale, ranging from 1 (very Latino) to 7 (not at all Latino), 1 (very masculine) to 7 (very feminine), and 1 (very gay) to 7 (very straight), respectively.

XSL•FO
RenderX

We added an exploratory measure of Latino identity expression using a single item from the Short Acculturation Scale for Hispanic people language subscale, "In which language(s) do you usually think?" with response options including 1 (only Spanish), 2 (more Spanish than English), 3 (both equally), 4 (more English than Spanish), and 5 (only English) [56], as a proxy for acculturation (ethnicity). We will measure expression or endorsement of hypermasculinity (masculine identity) using the 10-item machismo subscale from the existing Machismo and Caballarismo Scale [57]. Minor edits to language were made to make the questions more applicable to a nonheteronormative context (eg, replacing "wife" with "partner"). Finally, we will measure sexual identity expression as a single-item indicator of sexual orientation outness [58] on a 5-point Likert scale (range: 1=not at all open or out to 5=open or out to almost all people I know). The item was modified to reflect same-sex behaviors vs same-sex identity or attraction: "How open (out) are you as someone who is a man who has sex with other men?"

Primary Outcomes

The primary outcomes will be HIV testing and PrEP use at Month 6 and Month 12. These outcomes will be ascertained from both the medical records (primary outcome) and self-reports during the Qualtrics survey (exploratory outcome). We will define HIV testing as an HIV test date occurring in the past 12 months. We will define PrEP use (1) as the proportion of the past 12 months during which daily oral or injectable PrEP was prescribed, and (2) as a binary variable indicating whether any PrEP was prescribed in the past 12 months, to better account for participants who may have opted for on-demand oral PrEP dosing. Our decision not to assess on-demand PrEP oral dosing (ie, 2-1-1) was based on the fact that this dosing schedule was not approved by the Centers for Disease Control and Prevention or the US Food and Drug Administration, though this decision may potentially lead to an undercount of PrEP use. However, we will be able to conduct a sensitivity analysis with participants who are prescribed any oral PrEP during the assessment period and who self-report that they are "currently taking PrEP" without specifying their dosing schedule at Months 0, 3, 6, 9, and 12. We will also assess social network stability as an outcome (exploratory), which will be modeled as a latent variable using the proportion of alters named at Month 0 in the SNI that were identified as either an intersectional stigma source or identity-affirming source, but either (1) not named at Month 6 or 12, or (2) named as a new alter at Month 6 or 12.

Mediators

Identity-Related Measures

In line with Quinn and Earnshaw's [59] construction of the stigmatized identity model, we will assess how LMSM construct the individual-level meaning of their ethnic, masculine, and sexual identities in terms of their magnitude (centrality and salience) and valence (pride and internalized stigma) in Qualtrics. Ethnicity, masculinity, and sexuality-related identity centrality will be measured using the 6-item Identity Centrality Scale [60]. Items will be articulated to specify participants' "Latino or Hispanic," "masculine," and "sexual" identities, respectively. Responses will be provided on a 7-point Likert-type scale (1=strongly disagree to 7=strongly agree). We will measure identity salience (for Latino, masculine, and sexual identities) using Quinn and Chaudoir's [61] single-item measure, and responses will be provided on a 7-point Likert-type scale (1=almost never, 2=several times a year, 3=once a month, 4=once a week, 5=a few times a week, 6=once a day, and 7=many times a day). Ethnicity-, sexuality-, and masculinity-related pride will be used as a measure of identity affirmation (positive valence). A single-item continuous measure will ask participants, "On a scale of 0-100, how much pride do you take in your (Latino, masculine, or sexual) identity?" The definition of ethnic pride by Upadhyayula et al [62] was used to inform the structure of these items.

Internalized Stigma

Internalized Latino-, masculinity-, and sexuality-related stigma will be used as a measure of identity self-rejection or devaluation (negative valence) and will be collected during the Qualtrics survey. For Latino-related stigma, we adapted the Devaluation of Own Group subscale from the Appropriated Racial Oppression Scale [63] to be more appropriate for Latino or Hispanic participants (8 items). After consulting with community partners, we created 3 additional questions that aimed to reflect emotional responses to one's Latino or Hispanic ethnicity that were deemed to have a lower potential for social desirability bias than the scale's original items. Responses will be given on a 7-point Likert-type scale (1=strongly disagree to 7=strongly agree).

After failing to find an internalized masculinity stigma scale that did not conflate sexuality and gender, we created a new masculinity-related internalized stigma measure for this study. Item content was informed through a 2-step process: first, we used the 5-item structure of the Revised Internalized Homophobia Scale (IHP-R) by Herek et al [64], but changed "sexuality" to "masculinity" or "femininity." Next, we reworded items to mirror the framing of items from existing validated internalized stigma subscales [65]. Responses will be given on a 5-point Likert-type scale (1=strongly disagree to 5=strongly agree).

For sexuality-related stigma, we also adapted the IHP-R [64]. After consulting with community members, 2 items were revised using language from the 4-item internalized sexuality stigma measure of Earnshaw et al [66]. We made several other language revisions to ensure that questions better reflected the changes in social conventions against conversion therapy and to better reflect the language our participants might use. An additional sixth item was added based on community partner concerns that the "get help" to change one's sexuality item might perform poorly due to lack of relevance, and that "feeling guilty" about one's sexuality would be more relevant in our population. Responses will be given on a 5-point Likert-type scale (1=strongly disagree to 5=strongly agree).

Resilience

Potential mediators may also reflect social responses that may influence how social network intersectional stigma affects HIV testing and PrEP use (modeled as latent variables) in Qualtrics. We measure resilience using the 10-item Connor-Davidson

XSL•FO
RenderX

Resilience Scale (CD-RISC 10) [67] to reflect an individual-level response to intersectional stigma on our outcomes of interest. However, an error occurred during survey creation: respondents were asked to indicate "how much you agree or disagree with the following statements about how you generally respond to challenging situations," and response options were recorded as strength of agreement (1=strongly disagree to 5=strongly agree) instead of frequency (1=not true at all to 5=nearly all of the time). We will note this error and consider its implications in all future analyses using the CD-RISC 10 measure.

Medical Mistrust

Medical mistrust reflects another means of social response to structural manifestations of intersectional stigma that may influence our outcomes of interest. Medical mistrust toward health care systems (hospitals) will be measured in the Qualtrics survey at Month 0 and Month 6 using the 5-item Medical Mistrust Scale from LaVeist et al [68]. In line with this measure, 5 additional items were created in an attempt to capture a more intersectional LMSM experience related to ethnicity, language, masculinity, sexuality, and its disclosure. All questions will be asked on a 4-point scale (1=strongly disagree to 4=strongly agree).

Sexual Behaviors and HIV Prevention Support

We will assess sexual behaviors in the SNI as a relational process mediating the association between intersectional stigma and HIV testing and PrEP use (primary outcomes). For each sexual partner, we will assess contextual factors (eg, relationship type and how they met) and behaviors that may affect HIV-related transmission risk (eg, frequency of drinking or using drugs before sex, past 6-month anal and vaginal sexual partners, and anal or vaginal sex without a condom). HIV prevention norms will be assessed for all alters in terms of expressing favorable or unfavorable opinions for HIV testing and PrEP use, perceived social support from alters for HIV testing and PrEP use (informational, tangible, and emotional support), and factors that may affect the perceived credibility of the HIV prevention norms (alters' perceived HIV, ART, and PrEP status).

HIV and PrEP Stigma

We will assess stigma toward HIV and PrEP in Qualtrics at Month 0 and Month 12 as a sociocognitive process mediating the relationship between intersectional stigma with HIV testing and PrEP use. We will measure PrEP stigma using Siegler's 12-item PrEP Stigma Scale [69] and HIV stigma using the 13-item HIV Stigma Mechanisms Scale for people without HIV [70]. We adapted the HIV stigma scale in response to feedback from community partners and CAB members. Based on this feedback, we eliminated the need for reverse-coded items in the prejudice subscale and replaced 3 of the 4 discrimination subscale items that community members felt were outdated or potentially stigmatizing. Three new discrimination items were developed with our community partners to better reflect contemporary contexts in which LMSM may express HIV-related bias toward people with HIV (eg, dating apps, in-person dates, and HIV status disclosure in the dating context). One item in the stereotypes subscale was adapted to be more relevant to the LMSM context. Responses to both stigma scales will be given on a 5-point Likert-type scale (1=strongly disagree to 5=strongly agree).

Moderators

Potential moderators will include social network and LMSM characteristics, such as network size (number of alters), network density (connections between alters), and network composition (alter characteristics; all measured in the SNI). Those measured in Qualtrics will include nativity (born in the United States), years lived in the United States, local community affiliation (urban men who have sex with men [MSM] community or exurban Latino community based on zip code), and age group (18 - 24, 25 - 34, 35 - 44, and ≥45 years).

Covariates

Covariates will be collected during the baseline Qualtrics survey and will include sociodemographic characteristics, such as age, race, skin tone [71], self-described sexual orientation, education, income, and years living in San Diego County. We will measure alcohol use with the Alcohol Use Disorders Identification Test-Concise (AUDIT-C) [72] and substance use with the World Health Organization's Alcohol, Smoking and Substance Involvement Screening Test (ASSIST; version 3.0) [73]. Intimate partner violence will be assessed at baseline via 3 items adapted for this study [74-76]. Information on mental health (collected at Month 0 and Month 6) will include measuring symptoms of depression (using the 10-item Center for Epidemiologic Studies Depression [CESD-10] scale [77]), anxiety (using the Generalized Anxiety Disorder-7 Scale [GAD-7] [78]), and post-traumatic stress disorder using the Primary Care-Post-Traumatic Stress Disorder scale [79]. Social desirability will be assessed via RAND's 5-item social desirability response set [80].

We adapted the SHIPP HIV Knowledge Scale [69] based on piloting feedback to measure HIV knowledge at Month 0, removing 3 items with limited variability reflecting (1) knowledge that people with HIV can look healthy, and misinformation that (2) there is an HIV vaccine, and (3) that a shower after sex reduces HIV risk. A new item was developed with community partners to measure knowledge of Undetectable=Untransmittable. PrEP knowledge will be assessed at Month 0 using a 13-item subscale from Walsh's [81] PrEP Information-Motivation-Behavioral Skills model measure. Items were modified to add the word "pills" after the word PrEP to reduce conflating responses between oral vs injectable PrEP. Perceived HIV risk will be measured on a scale from 0 to 100 (0=no chance of getting HIV to 100=definitely). Other potential covariates may include HIV testing frequency, HIV testing barriers [82], PrEP barriers [83], and engagement in the PrEP care cascade [84,85]. Items related to COVID-19 (testing, vaccine status, vaccine willingness, and COVID-19–related fear [86]) will be assessed at Month 0.

XSL•FO
RenderX

Data Management and Analysis

Data Structure and Software

This study will include multilevel data, where level 1 data are variables related to alters in a participant's social network (eg, alter characteristics and intersectional stigma exposure from an alter). Alters will be nested within participants (level 2 data). Level 2 variables are related to LMSM (eg, participant characteristics) and aggregate network characteristics. SPSS (IBM SPSS Statistics) will be used to run descriptive statistics and univariate analyses to examine frequencies and distributions of the exposures, outcomes, mediators, moderators, and exploratory variables of interest. Appropriate nonparametric alternatives will be considered if parametric assumptions fail. Significance tests will be 2-sided ($P \leq .05$). For adapted or newly created measures (outlined above under the "Measures" subheading), we will test validity and internal reliability before including measures in analyses for the study's primary aims.

All study aims (1-4) will be analyzed in Mplus (Muthén and Muthén; version 8.1) using MSEM, a robust tool for modeling egocentric social network data, accounting for the dependence between level 1 data (eg, intersectional stigma exposure from alters) nested within level 2 data (eg, LMSM HIV testing and PrEP use) [87-89]. We will evaluate model fit using standard fit indices for all MSEMs (root-mean-square error of approximation ≤0.05, comparative fit index >0.95, and Tucker-Lewis index >0.95) [90]. Models may be adjusted to balance model fit, parsimony, and theory. Potential confounders for each aim will be identified from a directed acyclic graph [91] or identified via theory and the extant literature. If theoretically meaningful or statistically significant differences are observed across participants by recruitment site or geographic regions in San Diego County, these factors will be included as covariates in all final models.

Aim 1: Intersectional Stigma Measurement and Experiences

In the first aim, we will identify how social network intersectional stigma manifests at baseline (Month 0) by alter type and the degree to which LMSM express their intersectional identities (Figure 1). First, intraclass correlation coefficients (ICCs) will be examined for all level 1 variables to confirm sufficient variance for MSEM (ICC≥0.05). If the ICC is low (<0.05), indicating that the data are independent, level 1 data will be aggregated and treated as level 2 data (eg, group-mean centered) [92,93]. We will then construct a multilevel latent variable measurement model for social network intersectional stigma, composed of all observed scores (level 1) for anticipated and enacted stigma experienced for each intersectional stigma identity (Latino ethnicity, masculinity, and sexuality) at baseline (Figure 5).

Figure 5. Hypothesized multilevel structural equation model for study aims 1 and 2. A: anticipated stigma; E: enacted stigma; SN: social network; social network; ISS: intersectional stigma; PrEP: pre-exposure prophylaxis; L: Latino identity; M: masculine identity; S: sexuality identity.

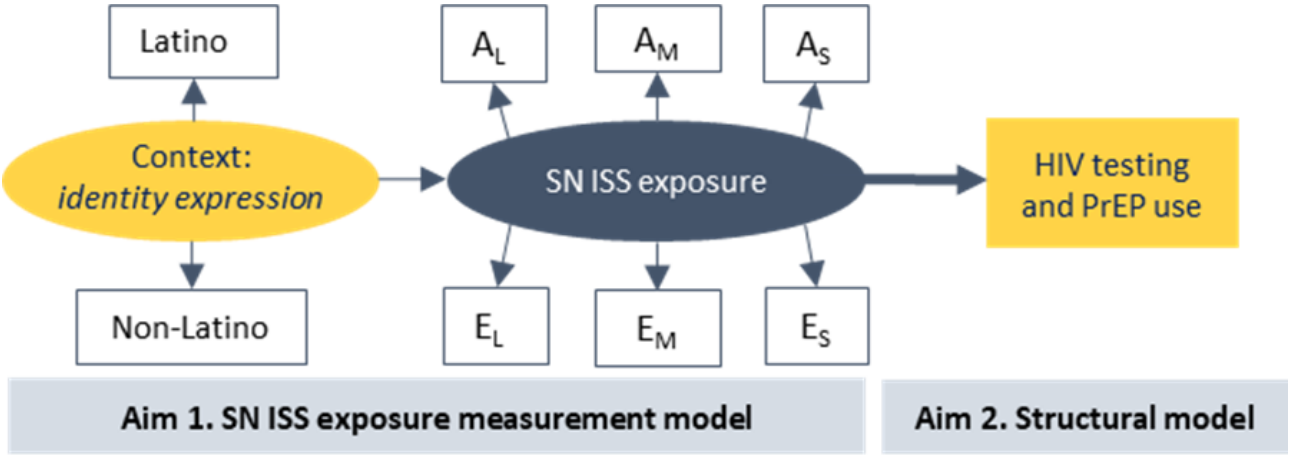

After ensuring measurement model fit, we will fit separate structural models to examine the associations between the social network intersectional stigma latent variable and eight different level 2 contextual variables by alter type (family, friend, sexual partner, provider, other) and the degree to which LMSM express each identity (ethnicity, masculinity, and sexuality). For each a priori theory-informed structural model by partner type and by identity expression, we will control for Type I errors via a modified Bonferroni procedure for MSEM [94]. We will use the previously specified fit statistics, trim nonsignificant paths, and adjust the model (if necessary) using modification indices that align with a priori hypotheses.

Aim 2: Relationship Between Intersectional Stigma, HIV Testing, and PrEP Use

In aim 2, we will examine the longitudinal association between the social network intersectional stigma latent variable at baseline (Month 0) and HIV testing and PrEP use at the final study visit (Month 12; Figure 1). A separate structural model for each primary outcome of interest (HIV testing [binary] and PrEP use [continuous]; level 2) will be assessed for each of the 8 models tested in aim 1 (16 a priori models total) to examine which contexts (5 intersectional stigma sources and 3 intersectional identity expressions) best predict future HIV testing and PrEP use (Figure 5). This will be determined by model fit comparison (as specified above under the Data Structure and Software subheading), and the magnitude, direction, and significance of the outcome regressed on the social network intersectional stigma latent variable.

XSL•FO
RenderX

Aim 3: Mediation and Moderation

In aim 3, we will explore factors that may mediate and moderate the association between social network intersectional stigma with HIV testing and PrEP use (Figure 1). We will use Preacher's [93] framework for assessing MSEM mediation models on level 2 outcomes (HIV testing and PrEP use). For each structural model developed in aim 2, we will fit our mediators of interest as a random effect separately for each social network (latent variables) and LMSM (observed score) response to stigmatized identities, HIV testing, and PrEP use. We will estimate the direct, indirect (level 2), and total effects for the proposed mediated processes. If there are multiple statistically significant mediators for an a priori structural model, each mediator will be entered into a single MSEM multiple mediation model to assess their independent effects.

For each structural model developed in aim 2, we will fit our moderators of interest by separately specifying an interaction term between social network intersectional stigma and each moderator. For models that demonstrate statistically significant moderation, we will conduct post hoc analyses to explore the level (by quartiles for continuous variables) at which the social network characteristic moderates the observed relationship between social network intersectional stigma, HIV testing, and PrEP use.

Aim 4: Social Network Stability

In aim 4, we will explore whether social network intersectional stigma exposure at baseline (Month 0) predicts changes in network stability across all study visits (Months 0 - 12) and whether this change in network stability is associated with increased HIV testing and PrEP use at the final study visit (Month 12; Figure 1). A structural model will be assessed for each of the 8 social network intersectional stigma measurement models that test the path between social network intersectional stigma exposure at Month 0 on LMSMs' social network stability latent variable at Month 12. To control for these associations, a separate (noncausal) path will be specified in which the social network stability variable at Month 12 is regressed on a latent variable comprised of observed scores for HIV testing and PrEP use at Month 12. Model fit, the magnitude, direction, and significance of the model paths will be examined.

Missing Data

Since NEXUS preferred sites comprise the majority of HIV service providers in Central and South San Diego County, we expect that the amount of missing medical records data for our primary outcomes will be minimal (<10%). Substantial efforts will be made to ensure Health Insurance Portability and Accountability Act authorization waivers and medical records release of information documents are updated and that participants complete follow-up. We will investigate the content (eg, HIV test dates and PrEP visits) and the type of missing data observed (eg, missing completely at random). However, a distinct advantage of using MSEM is that parameters are estimated with maximum likelihood (ML) estimation (for normal distributions) or ML with robust SEs, which is ML with robust SEs for nonnormal or ordinal data. ML is a rigorous method to handle missing data that yields unbiased parameter estimates, is appropriate for handling data missing completely at random or missing at random, and requires fewer assumptions about the data than multiple imputation [95]. In sensitivity analyses, we will (1) explore the impact of data missing not at random (eg, declining consent to obtain medical records), and (2) assess whether using HIV testing and PrEP use self-report data at Month 6 and 12 to supplement cases with null medical records may bias our findings.

Statistical Power

The following power calculations were conducted for aim 2, as aims 3 and 4 are exploratory, and aim 1 is a measurement model. Our target sample size of 500 LMSM will provide adequate statistical power (2-sided α=.05, 1-$\beta \geq$.81) to detect the effect of social network intersectional stigma on HIV testing and PrEP use. As most intersectional stigma research with MSM is qualitative, we used published effect size estimates for race-related [33,34,96], sexuality-related [33,34,96,97], and gender-related stigma [98,99] among ethno-racially diverse MSM on HIV testing and sexual risk (σ=0.30-0.60). We were unable to locate estimates on PrEP use. We modeled a range for the ICC (0.50-0.90), assuming a high degree of shared norms toward intersectional stigma identities is observed by alter type (eg, sexual vs family alters) clustered within LMSM. A sample size of 500 will yield sufficient power for effect sizes greater than 0.40, or an effect size of 0.30 with a moderate (0.50) ICC. For observed ICC values <0.50, we may be underpowered to detect an effect. In response, we will model our 2-sided α (α=.05) for the number of participants enrolled (eg, N=500), along with our observed ICC, and assess our power to detect the observed effect to aid in the interpretation of study results.

Data Management

The program manager will review data monthly to track study enrollment and monitor data quality, allowing for timely adjustments if needed. To ensure high-quality data and fidelity to SNI administration, the first 10 baseline SNIs per study interviewer and a random sample of 5% of all SNI surveys thereafter will be audio-recorded for fidelity monitoring (ie, elicitation of alters from the index and completeness of SNI delivery).

Ethical Considerations

The NEXUS study protocol was approved by the University of California, San Diego's Institutional Review Board (protocol no 200712). Study protocols were solicited and modified in response to feedback from our 2 community partners and our study-specific CAB across all study phases. Our CAB comprises 4 bilingual self-identified LMSM who are diverse in terms of length of residence in San Diego County and knowledge of local HIV services. CAB members received a biannual honorarium of US $250 for their time and expertise. Per protocol, written informed consent was obtained from all enrolled NEXUS participants in their preferred language (English or Spanish), and all data will be de-identified to maximize participants' privacy and confidentiality. Enrolled participants could earn up to US $360 for completing all NEXUS study activities.

XSL•FO
RenderX

Results

The NEXUS study received funding in July 2020, with an expected end date of April 2025. Protocols were adapted to meet the changing COVID-19 response landscape. From March to May 2021, we piloted our NEXUS study protocols, measures, and recruitment materials. Cohort recruitment began in June 2021, and our anticipated enrollment rate was impacted by COVID-19. As of March 11, 2025, a total of 482 of 500 (96%) LMSM had enrolled in NEXUS and completed the baseline visit, 345 of 412 (84%) had completed the Month 6 visit, and 317 of 351 (90%) had completed the Month 12 visit. Enrollment is expected to end by May 2025, and data analysis for Aim 1 and other baseline characteristics is expected to begin in June 2025. Data collection for our prospective study aims is expected to be completed in June 2026, with data analysis and publication of expected results later that year.

Discussion

Anticipated Principal Findings

We anticipate that the NEXUS study will advance our understanding of how and from whom experiences of intersectional stigma vary in relation to LMSM social networks to affect HIV prevention outcomes. For example, from aim 1, we expect that we may observe the strongest associations with enacted masculinity and sexuality stigma when looking at network members who are family, but stronger associations with anticipated Latino and masculinity stigma when looking at network members who are sexual partners. Subsequently, aim 2 should inform us from whom these exposures to intersectional stigma are most strongly associated with PrEP use at Month 12, hypothesizing that we may observe stronger negative associations between PrEP use and intersectional stigma exposure from family than sexual partners. Aim 3 results should provide greater insights into what mediation processes may strengthen (ie, internalized masculinity stigma) or attenuate (ie, resilience) these negative associations with PrEP use, and for whom those associations are strongest (ie, LMSM with smaller social networks and LMSM who are younger). In contrast, aim 4 results should provide insights into the role network stability plays in HIV prevention, for example, greater stability among friend networks who are welcoming of LMSM may be associated with greater HIV testing at Month 12, compared to less stability among sexual partners who are stigmatizing.

These findings will be highly relevant to the US HIV epidemic, as LMSM experience a disproportionate risk of HIV acquisition, and intersectional stigma is widely recognized as a key barrier to eliminating this inequity [14,21,100,101]. Current approaches to quantitative intersectional stigma measurement are limited, and experts continually underscore that this remains a key challenge in the field [19,21]. Most existing measures of stigma toward intersectional identities are rarely measured in relation to one another or in relation to the social contexts that affect how and from whom intersectional stigma is experienced (ie, alter types and expression of stigmatized identities) [19,21,23]. The NEXUS study will use a social network approach to studying intersectional stigma. This methodology is highly innovative because it reflects a substantial departure from how intersectional stigma has been studied and potentially opens new opportunities to end the HIV epidemic. Specifically, NEXUS will allow us to identify the social contexts that affect how intersectional stigma is experienced by LMSM to affect HIV testing and PrEP use. It will also help elucidate the mechanisms that mediate or moderate the association between intersectional stigma exposure, HIV testing, and PrEP use. We will use results from NEXUS to develop novel intersectional stigma intervention targets for future piloting to reduce LMSM's HIV incidence. In doing so, we also aim to promote resilience and reduce the harmful effects of intersectional stigma. Without understanding the larger social context of intersectional stigma and using this information to guide future work, HIV inequities will persist, and we will struggle to meet the national goals to end the HIV epidemic by 2030 [5].

Limitations

However, this approach is not without its limitations, including the added response burden associated with a social network assessment of intersectional stigma. We hope that the ability to analyze and compare the cross-sectional associations between our multi-item mechanism-specific approach assessed at baseline (Month 0) and the more general single-item mechanism-specific approach assessed at follow-up (Months 6 and 12) can inform when or if this response burden can be attenuated while still capitalizing on the rich nature of social network data. Similarly, the COVID-19 pandemic coincided with the implementation of the NEXUS study, which required innovation in our intended recruitment and retention strategies and the development of more time-intensive virtual protocols. These strategies may have affected the types of LMSM we were able to reach with fewer in-person recruitment opportunities earlier on, but may have also increased the range of LMSM who agreed to participate, given the potential for greater perceived privacy and flexibility of our virtual protocols. These shifts may have further impacted study retention; however, our retention protocols and twice-monthly study team meetings with real-time retention monitoring and implementation of solutions-oriented retention strategies have allowed us to maintain an 84% and 90% retention rate at Month 6 and Month 12 follow-up visits based on the available data. Finally, medical mistrust may play a role in who agreed to enroll in the study and which enrolled participants consented to have their HIV testing and PrEP use data abstracted from their medical records and deidentified for analysis. Medical mistrust, along with variation in how medical records data are recorded across providers, is a limitation to be expected in any HIV-related study, especially those among ethno-racially marginalized groups. As such, the findings should be as generalizable as those of other published studies, and limitations in generalizability to the broader LMSM community will be noted in subsequent publications to guide our interpretation of the study results. Relatedly, we will assess if enrolled participants who declined to have their medical records abstracted, withdrew from the study for any reason, or were lost to follow-up varied in relation to their baseline medical mistrust scores, and if observed associations with HIV testing and PrEP

XSL•FO
RenderX

use outcomes were moderated by medical mistrust at baseline to help guide our interpretation of study results.

Conclusions

As NEXUS study results emerge, all findings will be interpreted in light of the extant literature, theory, and lived experiences of NEXUS study community partners, interviewers, and CAB members. In addition to traditional dissemination models (ie, peer-reviewed publications), we aim to disseminate study results through bilingual, in-person, and virtual forums with our community partners, hosted by our NEXUS interviewers, who worked hard to build and maintain trust and rapport with NEXUS participants. Additionally, we aim to develop brief infographics that summarize NEXUS study findings and their implications for supporting LMSM's health and well-being, which can be disseminated via channels used for study recruitment, including our study website and local social media channels, tabling at local health fairs and bar-based HIV testing, and more broadly through our established relationships with scientific and community organizations in California and across the United States (eg, the California Center for HIV Syndemic Policy Research, the San Diego Center for AIDS Research, National Minority AIDS Council, the National Hispanic Science Network, and the National Latinx Conference on HIV, hepatitis C virus, and substance use disorder).

Acknowledgments

The authors thank all study participants and the NEXUS study team, including our community partners and members of our community advisory board, for their time and dedication to ensure all aspects of the study were culturally and linguistically appropriate. We also wish to thank the HIV testers and counselors, as well as community partner service providers, including Esteban Duarte and his organization VIDA, who pivoted in response to the COVID-19 pandemic to ensure HIV prevention services remained accessible to Latino men who have sex with men (LMSM) in San Diego County. Generative AI was not used to support any aspect of this manuscript or the NEXUS study.

Funding

This study was funded as a 5-year grant from the National Institutes of Health (NIH; R01MH123282) starting July 1, 2020, for a total of US $3,464,530. Additional scientific support was provided through 3 NIH training mechanisms (T32 DA023356, T32 AI007384 and K01 DA055521). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This research also received methodological support from the San Diego Center for AIDS Research (SD CFAR) Biostatistical Modeling Core and Socoiobehavioral Science Core, an NIH-funded program (P30 AI036214), which is supported by the following NIH Institutes and Centers: NIAID, NCI, NHLBI, NIA, NICHD, NIDA, NIDCR, NIDDK, NIMH, NIMHD, NINR, FIC, and OAR.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request, which may require securing institutional data-use agreements and having a proposed analysis plan approved by the corresponding author.

Conflicts of Interest

None declared.

Multimedia Appendix 1
NEXUS social network name generator protocol.
[DOCX File, 722 KB - resprot_v15i1e72334_app1.docx]

Multimedia Appendix 2
NEXUS longitudinal social network data integration protocol.
[DOCX File, 247 KB - resprot_v15i1e72334_app2.docx]

Multimedia Appendix 3
NEXUS social network intersectional stigma–related items.
[DOCX File, 51 KB - resprot_v15i1e72334_app3.docx]

References

1. HIV Surveillance Supplemental Report: Estimated HIV Incidence and Prevalence in the United States, 2018-2022. : Centers for Disease Control and Prevention (CDC) Stacks; 2024 URL: https://stacks.cdc.gov/view/cdc/156513 [accessed 2026-02-06]
2. AtlasPlus - overview. Centers for Disease Control and Prevention (CDC). URL: https://gis.cdc.gov/grasp/nchhstpatlas/main.html [accessed 2024-10-24]

XSL•FO
RenderX

3. Guilamo-Ramos V, Thimm-Kaiser M, Benzekri A, et al. The invisible US Hispanic/Latino HIV crisis: addressing gaps in the national response. Am J Public Health 2020 Jan;110(1):27-31. [doi: 10.2105/AJPH.2019.305309] [Medline: 31725313]
4. HIV Surveillance Report: Diagnoses, Deaths, and Prevalence of HIV in the United States and 6 Territories and Freely Associated States, 2022. : Centers for Disease Control and Prevention (CDC) Stacks; 2022 URL: https://stacks.cdc.gov/view/cdc/156509 [accessed 2026-02-06]
5. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. JAMA 2019 Mar 5;321(9):844-845. [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
6. Fonner VA, Dalglish SL, Kennedy CE, et al. Effectiveness and safety of oral HIV preexposure prophylaxis for all populations. AIDS 2016 Jul 31;30(12):1973-1983. [doi: 10.1097/QAD.0000000000001145] [Medline: 27149090]
7. Monitoring Selected National HIV Prevention and Care Objectives by Using HIV Surveillance Data United States and 6 Dependent Areas, 2020. : Centers for Disease Control and Prevention; 2022 URL: https://stacks.cdc.gov/view/cdc/117803 [accessed 2026-02-06]
8. Matos LA, Janek SE, Holt L, Ledbetter L, Gonzalez-Guarda RM. Barriers and facilitators along the PrEP continuum of care among Latinx sexual minoritized men and transgender women: a systematic review. AIDS Behav 2024 Nov;28(11):3666-3709. [doi: 10.1007/s10461-024-04434-8] [Medline: 39083153]
9. Earnshaw VA, Chaudoir SR. From conceptualizing to measuring HIV stigma: a review of HIV stigma mechanism measures. AIDS Behav 2009 Dec;13(6):1160-1177. [doi: 10.1007/s10461-009-9593-3] [Medline: 19636699]
10. Babel RA, Wang P, Alessi EJ, Raymond HF, Wei C. Stigma, HIV risk, and access to HIV prevention and treatment services among men who have sex with men (MSM) in the United States: a scoping review. AIDS Behav 2021 Nov;25(11):3574-3604. [doi: 10.1007/s10461-021-03262-4] [Medline: 33866444]
11. Andrade EA, Stoukides G, Santoro AF, Karasz A, Arnsten J, Patel VV. Individual and health system factors for uptake of pre-exposure prophylaxis among young Black and Latino gay men. J Gen Intern Med 2023 Sep;38(12):2768-2774. [doi: 10.1007/s11606-023-08274-5] [Medline: 37429976]
12. Brooks RA, Nieto O, Rosenberg-Carlson E, et al. Barriers and facilitators to accessing PrEP and other sexual health services among immigrant Latino men who have sex with men in Los Angeles county. Arch Sex Behav 2024 Oct;53(9):3673-3685. [doi: 10.1007/s10508-024-02928-z] [Medline: 38977534]
13. Stigma GE. Notes on the Management of Spoiled Identity: Touchstone (Nashv); 1963.
14. Hatzenbuehler ML, Phelan JC, Link BG. Stigma as a fundamental cause of population health inequalities. Am J Public Health 2013 May;103(5):813-821. [doi: 10.2105/AJPH.2012.301069] [Medline: 23488505]
15. Crenshaw K. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics: University of Chicago Legal Forum, Vol. 1989. URL: https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=1052&context=uclf [accessed 2026-02-06]
16. Berger MT. Workable Sisterhood: The Political Journey of Stigmatized Women with HIV/AIDS: Princeton University Press; 2010.
17. Relf MV, L Holzemer W, Holt L, Nyblade L, Ellis Caiola C. A review of the state of the science of HIV and stigma: context, conceptualization, measurement, interventions, gaps, and future priorities. J Assoc Nurses AIDS Care 2021;32(3):392-407. [doi: 10.1097/JNC.0000000000000237] [Medline: 33654005]
18. Pérez OFR, Machismo MA. The Wiley Encyclopedia of Personality and Individual Differences: John Wiley & Sons, Ltd; 2020:243-246. [doi: 10.1002/9781119547181]
19. Turan JM, Elafros MA, Logie CH, et al. Challenges and opportunities in examining and addressing intersectional stigma and health. BMC Med 2019 Dec;17(1). [doi: 10.1186/s12916-018-1246-9]
20. Earnshaw VA, Jonathon Rendina H, Bauer GR, et al. Methods in HIV-related intersectional stigma research: core elements and opportunities. Am J Public Health 2022 Jun;112(S4):S413-S419. [doi: 10.2105/AJPH.2021.306710] [Medline: 35763749]
21. Karver TS, Atkins K, Fonner VA, et al. HIV-related intersectional stigma and discrimination measurement: state of the science. Am J Public Health 2022 Jun;112(S4):S420-S432. [doi: 10.2105/AJPH.2021.306639] [Medline: 35763725]
22. Lett E. Crossing lines does not equal intersectionality. J Behav Med 2022 Dec;45(6):983-984. [doi: 10.1007/s10865-022-00375-6] [Medline: 36355228]
23. Bauer GR. Incorporating intersectionality theory into population health research methodology: challenges and the potential to advance health equity. Soc Sci Med 2014 Jun;110:10-17. [doi: 10.1016/j.socscimed.2014.03.022] [Medline: 24704889]
24. Valente TW. Social Networks and Health: Models, Methods, and Applications: Oxford University Press; 2010. URL: http://archive.org/details/socialnetworkshe0000vale [accessed 2024-10-01]
25. Latkin CA, Knowlton AR. Social network assessments and interventions for health behavior change: a critical review. Behav Med 2015;41(3):90-97. [doi: 10.1080/08964289.2015.1034645] [Medline: 26332926]
26. Pescosolido BA, Manago B. Getting underneath the power of “contact”: revisiting the fundamental lever of stigma as a social network phenomenon. In: Major B, Dovidio JF, Link BG, editors. The Oxford Handbook of Stigma, Discrimination, and Health: Oxford University Press; 2018. [doi: 10.1093/oxfordhb/9780190243470.013.16]
27. Pettigrew TF. Intergroup contact theory. Annu Rev Psychol 1998;49(1):65-85. [doi: 10.1146/annurev.psych.49.1.65] [Medline: 15012467]

XSL•FO
RenderX

28. Stangl AL, Earnshaw VA, Logie CH, et al. The Health Stigma and Discrimination Framework: a global, crosscutting framework to inform research, intervention development, and policy on health-related stigmas. BMC Med 2019 Dec;17(1). [doi: 10.1186/s12916-019-1271-3]
29. Collins PH. Intersectionality as Critical Social Theory: Duke University Press; 2019.
30. Chaudoir SR, Earnshaw VA, Andel S. "Discredited" versus "Discreditable": understanding how shared and unique stigma mechanisms affect psychological and physical health disparities. Basic Appl Soc Psych 2013 Jan 1;35(1):75-87. [doi: 10.1080/01973533.2012.746612] [Medline: 23729948]
31. Pachankis JE. The psychological implications of concealing a stigma: a cognitive-affective-behavioral model. Psychol Bull 2007 Mar;133(2):328-345. [doi: 10.1037/0033-2909.133.2.328] [Medline: 17338603]
32. Quinn DM, Earnshaw VA. Understanding concealable stigmatized identities: the role of identity in psychological, physical, and behavioral outcomes. Social Issues Policy Review 2011 Dec;5(1):160-190 [FREE Full text] [doi: 10.1111/j.1751-2409.2011.01029.x]
33. Frye V, Nandi V, Egan J, et al. Sexual orientation- and race-based discrimination and sexual HIV risk behavior among urban MSM. AIDS Behav 2015 Feb;19(2):257-269. [doi: 10.1007/s10461-014-0937-2] [Medline: 25381561]
34. Mizuno Y, Borkowf C, Millett GA, Bingham T, Ayala G, Stueve A. Homophobia and racism experienced by Latino men who have sex with men in the United States: correlates of exposure and associations with HIV risk behaviors. AIDS Behav 2012 Apr;16(3):724-735. [doi: 10.1007/s10461-011-9967-1] [Medline: 21630014]
35. Brooks RA, Landrian A, Nieto O, Fehrenbacher A. Experiences of anticipated and enacted pre-exposure prophylaxis (PrEP) stigma among Latino MSM in Los Angeles. AIDS Behav 2019 Jul;23(7):1964-1973. [doi: 10.1007/s10461-019-02397-9] [Medline: 30649635]
36. Choi KH, Han CS, Paul J, Ayala G. Strategies for managing racism and homophobia among U.S. ethnic and racial minority men who have sex with men. AIDS Educ Prev 2011 Apr;23(2):145-158. [doi: 10.1521/aeap.2011.23.2.145] [Medline: 21517663]
37. Sears JL. Machismo as a determinant for HIV/STD risk behavior among Latino MSM. : Virginia Commonwealth University; 2006. [doi: 10.25772/8MR1-WJ21]
38. Jarama SL, Kennamer JD, Poppen PJ, Hendricks M, Bradford J. Psychosocial, behavioral, and cultural predictors of sexual risk for HIV infection among Latino men who have sex with men. AIDS Behav 2005 Dec;9(4):513-523. [doi: 10.1007/s10461-005-9022-1] [Medline: 16328712]
39. Watkins-Hayes C. Intersectionality and the sociology of HIV/AIDS: past, present, and future research directions. Annu Rev Sociol 2014 Jul 30;40(1):431-457 [FREE Full text] [doi: 10.1146/annurev-soc-071312-145621]
40. Cho S, Crenshaw KW, McCall L. Toward a field of intersectionality studies: theory, applications, and praxis. Signs: Journal of Women in Culture and Society 2013 Jun;38(4):785-810. [doi: 10.1086/669608]
41. Bowleg L. When Black + lesbian + woman ≠ Black lesbian woman: the methodological challenges of qualitative and quantitative intersectionality research. Sex Roles 2008 Sep;59(5-6):312-325. [doi: 10.1007/s11199-008-9400-z]
42. Bowleg L. The problem with the phrase women and minorities: intersectionality-an important theoretical framework for public health. Am J Public Health 2012 Jul;102(7):1267-1273. [doi: 10.2105/AJPH.2012.300750] [Medline: 22594719]
43. Morris RC. Identity salience and identity importance in identity theory. Curr Res Soc Psychol ;21(8) [FREE Full text]
44. Perry BL. The labeling paradox: stigma, the sick role, and social networks in mental illness. J Health Soc Behav 2011 Dec;52(4):460-477. [doi: 10.1177/0022146511408913] [Medline: 22144733]
45. Luthar SS, Cicchetti D, Becker B. The construct of resilience: a critical evaluation and guidelines for future work. Child Dev 2000;71(3):543-562. [doi: 10.1111/1467-8624.00164] [Medline: 10953923]
46. Earnshaw VA, Bogart LM, Dovidio JF, Williams DR. Stigma and racial/ethnic HIV disparities: moving toward resilience. Am Psychol 2013;68(4):225-236. [doi: 10.1037/a0032705] [Medline: 23688090]
47. Fletcher D, Sarkar M. Psychological resilience: a review and critique of definitions, concepts, and theory. Eur Psychol 2013;18(1):12-23. [doi: 10.1027/1016-9040/a000124]
48. Meyer IH. Resilience in the study of minority stress and health of sexual and gender minorities. Psychol Sex Orientat Gend Divers 2015;2(3):209-213. [doi: 10.1037/sgd0000132]
49. Crawford ND, Ford C, Rudolph A, Kim B, Lewis CM. Drug use discrimination predicts formation of high-risk social networks: examining social pathways of discrimination. AIDS Behav 2017 Sep;21(9):2659-2669. [doi: 10.1007/s10461-016-1639-8] [Medline: 28025736]
50. Phelan JC, Link BG, Dovidio JF. Stigma and prejudice: one animal or two? Soc Sci Med 2008 Aug;67(3):358-367. [doi: 10.1016/j.socscimed.2008.03.022] [Medline: 18524444]
51. Janulis P, Phillips G, Melville J, et al. Network canvas: an open-source tool for capturing social and contact network data. Int J Epidemiol 2023 Aug 2;52(4):1286-1291. [doi: 10.1093/ije/dyad036] [Medline: 36944105]
52. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. Public Opin Q 2007 Jan 1;71(2):287-311. [doi: 10.1093/poq/nfm006]
53. Wasserman S, Faust K. Social Network Analysis: Methods and Applications: Cambridge University Press; 1994.
54. Kachel S, Steffens MC, Niedlich C. Traditional masculinity and femininity: validation of a new scale assessing gender roles. Front Psychol 2016;7(956):956. [doi: 10.3389/fpsyg.2016.00956] [Medline: 27458394]

XSL•FO
RenderX

55. Rider GN, McMorris BJ, Gower AL, Coleman E, Eisenberg ME. Health and care utilization of transgender and gender nonconforming youth: a population-based study. Pediatrics 2018 Mar;141(3):e20171683. [doi: 10.1542/peds.2017-1683] [Medline: 29437861]
56. Marin G, Gamba RJ. A new measurement of acculturation for Hispanics: the Bidimensional Acculturation Scale for Hispanics (BAS). Hisp J Behav Sci 1996 Aug;18(3):297-316. [doi: 10.1177/07399863960183002]
57. Arciniega GM, Anderson TC, Tovar-Blank ZG, Tracey TJG. Toward a fuller conception of Machismo: development of a traditional Machismo and Caballerismo Scale. J Couns Psychol 2008;55(1):19-33. [doi: 10.1037/0022-0167.55.1.19]
58. Wilkerson JM, Noor SW, Galos DL, Rosser BRS. Correlates of a single-item indicator versus a multi-item scale of outness about same-sex attraction. Arch Sex Behav 2016 Jul;45(5):1269-1277. [doi: 10.1007/s10508-015-0605-2] [Medline: 26292840]
59. Quinn DM, Earnshaw VA. Concealable stigmatized identities and psychological well-being. Soc Personal Psychol Compass 2013 Jan;7(1):40-51. [doi: 10.1111/spc3.12005] [Medline: 23730326]
60. Quinn DM, Williams MK, Quintana F, et al. Examining effects of anticipated stigma, centrality, salience, internalization, and outness on psychological distress for people with concealable stigmatized identities. PLoS ONE 2014;9(5):e96977. [doi: 10.1371/journal.pone.0096977] [Medline: 24817189]
61. Quinn DM, Chaudoir SR. Living with a concealable stigmatized identity: the impact of anticipated stigma, centrality, salience, and cultural stigma on psychological distress and health. J Pers Soc Psychol 2009 Oct;97(4):634-651. [doi: 10.1037/a0015815] [Medline: 19785483]
62. Upadhyayula S, Ramaswamy M, Chalise P, Daniels J, Freudenberg N. The association of ethnic pride with health and social outcomes among young Black and Latino men after release from jail. Youth Soc 2017 Nov;49(8):1057-1076. [doi: 10.1177/0044118X15576736] [Medline: 29176915]
63. Campón RR, Carter RT. The Appropriated Racial Oppression Scale: development and preliminary validation. Cultur Divers Ethnic Minor Psychol 2015 Oct;21(4):497-506. [doi: 10.1037/cdp0000037] [Medline: 25894835]
64. Herek GM, Gillis JR, Cogan JC. Internalized stigma among sexual minority adults: insights from a social psychological perspective. J Couns Psychol 2009;56(1):32-43. [doi: 10.1037/a0014672]
65. Smith LR, Earnshaw VA, Copenhaver MM, Cunningham CO. Substance use stigma: reliability and validity of a theory-based scale for substance-using populations. Drug Alcohol Depend 2016 May 1;162:34-43. [doi: 10.1016/j.drugalcdep.2016.02.019] [Medline: 26972790]
66. Earnshaw VA, Reed NM, Watson RJ, Maksut JL, Allen AM, Eaton LA. Intersectional internalized stigma among Black gay and bisexual men: a longitudinal analysis spanning HIV/sexually transmitted infection diagnosis. J Health Psychol 2021 Mar;26(3):465-476. [doi: 10.1177/1359105318820101] [Medline: 30599761]
67. Campbell-Sills L, Stein MB. Psychometric analysis and refinement of the Connor-Davidson Resilience Scale (CD-RISC): validation of a 10-item measure of resilience. J Trauma Stress 2007 Dec;20(6):1019-1028. [doi: 10.1002/jts.20271] [Medline: 18157881]
68. LaVeist TA, Nickerson KJ, Bowie JV. Attitudes about racism, medical mistrust, and satisfaction with care among African American and White cardiac patients. Med Care Res Rev 2000;57 Suppl 1:146-161. [doi: 10.1177/1077558700057001S07] [Medline: 11092161]
69. Siegler AJ, Wiatrek S, Mouhanna F, et al. Validation of the HIV Pre-exposure Prophylaxis Stigma Scale: performance of Likert and semantic differential scale versions. AIDS Behav 2020 Sep;24(9):2637-2649. [doi: 10.1007/s10461-020-02820-6] [Medline: 32157490]
70. Earnshaw VA, Smith LR, Chaudoir SR, Lee IC, Copenhaver MM. Stereotypes about people living with HIV: implications for perceptions of HIV risk and testing frequency among at-risk populations. AIDS Educ Prev 2012 Dec;24(6):574-581. [doi: 10.1521/aeap.2012.24.6.574] [Medline: 23206205]
71. Sheehan DM, Cano M, Trepka MJ. Measuring race among Latinos to address racism, discrimination, and HIV health inequities: comparing self-reported race and Fitzpatrick skin phototype. J Assoc Nurses AIDS Care 2021;32(4):522-525. [doi: 10.1097/JNC.0000000000000260] [Medline: 33859078]
72. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
73. Humeniuk R, Henry-Edwards S, Ali R, Poznyak V, Monteiro MG, World Health Organization. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Manual for Use in Primary Care: World Health Organization; 2010:978-992 URL: https://iris.who.int/handle/10665/44320 [accessed 2024-10-22]
74. Ford-Gilboe M, Wathen CN, Varcoe C, et al. Development of a brief measure of intimate partner violence experiences: the Composite Abuse Scale (Revised)—Short Form (CAS R -SF). BMJ Open 2016 Dec;6(12):e012824. [doi: 10.1136/bmjopen-2016-012824]
75. Straus MA, Hamby SL, Boney-McCoy S, Sugarman DB. The revised Conflict Tactics Scales (CTS2): development and preliminary psychometric data. J Fam Issues 1996;17(3) [FREE Full text]

XSL•FO
RenderX

76. Smith LR, Yore J, Triplett DP, et al. Impact of sexual violence across the lifespan on HIV risk behaviors among transgender women and cisgender people living with HIV. J Acquired Immune Defic Syndr 2017;75(4):408-416. [doi: 10.1097/QAI.0000000000001423]
77. Lewinsohn PM, Seeley JR, Roberts RE, Allen NB. Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. Psychol Aging 1997;12(2):277-287. [doi: 10.1037/0882-7974.12.2.277]
78. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
79. Prins A, Ouimette P, Kimerling R, et al. The primary care PTSD screen (PC–PTSD): development and operating characteristics. Prim Care Psych 2004 Jan 1;9(1):9-14. [doi: 10.1185/135525703125002360]
80. Hays RD, Hayashi T, Stewart AL. A five-item measure of socially desirable response set. Educ Psychol Meas 1989 Sep;49(3):629-636. [doi: 10.1177/001316448904900315]
81. Walsh JL. Applying the Information-Motivation-Behavioral Skills model to understand PrEP intentions and use among men who have sex with men. AIDS Behav 2019 Jul;23(7):1904-1916. [doi: 10.1007/s10461-018-2371-3] [Medline: 30554396]
82. Pines HA, Goodman-Meza D, Pitpitan EV, Torres K, Semple SJ, Patterson TL. HIV testing among men who have sex with men in Tijuana, Mexico: a cross-sectional study. BMJ Open 2016 Feb 4;6(2):e010388. [doi: 10.1136/bmjopen-2015-010388] [Medline: 26846899]
83. Patrick R, Jain J, Harvey-Vera A, et al. Perceived barriers to pre-exposure prophylaxis use among HIV-negative men who have sex with men in Tijuana, Mexico: a latent class analysis. PLOS ONE 2019;14(8):e0221558. [doi: 10.1371/journal.pone.0221558] [Medline: 31437243]
84. Blashill AJ, Brady JP, Rooney BM, et al. Syndemics and the PrEP cascade: results from a sample of young Latino men who have sex with men. Arch Sex Behav 2020 Jan;49(1):125-135. [doi: 10.1007/s10508-019-01470-7] [Medline: 31591666]
85. Parsons JT, Rendina HJ, Lassiter JM, Whitfield THF, Starks TJ, Grov C. Uptake of HIV pre-exposure prophylaxis (PrEP) in a National Cohort of Gay and Bisexual Men in the United States. J Acquir Immune Defic Syndr 2017 Mar 1;74(3):285-292. [doi: 10.1097/QAI.0000000000001251] [Medline: 28187084]
86. Fitzpatrick KM, Harris C, Drawve G. Living in the midst of fear: depressive symptomatology among US adults during the COVID-19 pandemic. Depress Anxiety 2020 Oct;37(10):957-964. [doi: 10.1002/da.23080] [Medline: 32667117]
87. Heck RH, Thomas SL. An Introduction to Multilevel Modeling Techniques: MLM and SEM Approaches, 4th edition: Routledge; 2020.
88. Muthén B, Asparouhov T, Sobel M. Applications of causally defined direct and indirect effects in mediation analysis using SEM in Mplus. Semantic Scholar. 2011. URL: https://www.semanticscholar.org/paper/Applications-of-Causally-Defined-Direct-and-Effects-Muth%C3%A9n-Asparouhov/776d2df784e67ab691bbffc280d3f4a480740300?p2df [accessed 2024-10-02]
89. Muthén LK, Muthén B. Mplus short courses, topic 7. Mplus.
90. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
91. Pearl J. Causal diagrams for empirical research. Biometrika 1995;82(4):669-688. [doi: 10.1093/biomet/82.4.669]
92. Geiser C. Data Analysis with Mplus: Guilford Press; 2012.
93. Preacher KJ, Zyphur MJ, Zhang Z. A general multilevel SEM framework for assessing multilevel mediation. Psychol Methods 2010 Sep;15(3):209-233. [doi: 10.1037/a0020141] [Medline: 20822249]
94. Smith CE, Cribbie RA. Multiplicity control in structural equation modeling: incorporating parameter dependencies. Struct Equ Modeling 2013 Jan 1;20(1):79-85. [doi: 10.1080/10705511.2013.742385]
95. Allison PD. Missing data techniques for structural equation modeling. J Abnorm Psychol 2003 Nov;112(4):545-557. [doi: 10.1037/0021-843X.112.4.545] [Medline: 14674868]
96. Scott HM, Pollack L, Rebchook GM, Huebner DM, Peterson J, Kegeles SM. Peer social support is associated with recent HIV testing among young Black men who have sex with men. AIDS Behav 2014 May;18(5):913-920. [doi: 10.1007/s10461-013-0608-8] [Medline: 24065436]
97. Jeffries WL, Marks G, Lauby J, Murrill CS, Millett GA. Homophobia is associated with sexual behavior that increases risk of acquiring and transmitting HIV infection among Black men who have sex with men. AIDS Behav 2013 May;17(4):1442-1453. [doi: 10.1007/s10461-012-0189-y] [Medline: 22569942]
98. Brennan DJ, Souleymanov R, George C, et al. Masculinity, muscularity, and HIV sexual risk among gay and bisexual men of color. Psychol Men Masc 2015;16(4):393-403. [doi: 10.1037/a0038725]
99. Parent MC, Torrey C, Michaels MS. “HIV testing is so gay”: the role of masculine gender role conformity in HIV testing among men who have sex with men. J Couns Psychol 2012 Jul;59(3):465-470. [doi: 10.1037/a0028067] [Medline: 22774868]
100. Rodriguez-Hart C, Boone CA, del Río-González AM, et al. Monitoring intersectional stigma: a key strategy to ending the HIV epidemic in the United States. Am J Public Health 2022 Jun;112(S4):S350-S355. [doi: 10.2105/AJPH.2022.306733] [Medline: 35763745]

XSL•FO
RenderX

101. García M, Harris AL. PrEP awareness and decision-making for Latino MSM in San Antonio, Texas. PLoS ONE 2017;12(9):e0184014. [doi: 10.1371/journal.pone.0184014]

Abbreviations

ASSIST: Alcohol, Smoking, and Substance Involvement Screening Test
AUDIT-C: Alcohol Use Disorders Identification Test-Concise
CAB: community advisory board
CD-RISC 10: 10-item Connor-Davidson Resilience Scale
CESD-10: 10-item Center for Epidemiologic Studies Depression
GAD-7: Generalized Anxiety Disorder-7 Scale
HIPAA: Health Insurance Portability and Accountability Act
ICC: intraclass correlation coefficient
IHP-R: Revised Internalized Homophobia Scale
LGBTQ: lesbian, gay, bisexual, and transgender, and queer
LMSM: Latino men who have sex with men
ML: maximum likelihood
MSEM: multilevel structural equation modeling
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
SNI: social network inventory

Edited by J Sarvestan; submitted 12.Mar.2025; peer-reviewed by M Hibbert; revised version received 29.Dec.2025; accepted 30.Dec.2025; published 27.Feb.2026.

Please cite as:
Smith LR, Algarin AB, Pitpitan EV, Pines HA, Kelly NK, Latkin C, Gutierrez A, Hussain M, Soto F, Vazquez Jr R, Juarez S, Genegaling A, Rodriguez R, Cirilo A, Esparza J, Cruz S, Jimenez E, Reyes U, Moore V, Guerbaoui M, Davenport B, Aldous J, Penniga K, Parker K, Skaathun B
A Novel Social Network Approach to Measure Intersectional Stigma Among Latino Men Who Have Sex With Men in San Diego, California (NEXUS): Protocol for a Longitudinal Cohort Study
JMIR Res Protoc 2026;15:e72334
URL: https://www.researchprotocols.org/2026/1/e72334
doi:10.2196/72334

© Laramie R Smith, Angel B Algarin, Eileen V Pitpitan, Heather A Pines, Nicole K Kelly, Carl Latkin, Aaron Gutierrez, Maryam Hussain, Francisco Soto, Ricardo Vazquez Jr, Stevie Juarez, Albert Genegaling, Rosalinda Rodriguez, Anthony Cirilo, Juan Esparza, Saul Cruz, Erick Jimenez, Ulises Reyes, Veronica Moore, Malek Guerbaoui, Beth Davenport, Jeannette Aldous, Katherine Penniga, Kenyatta Parker, Britt Skaathun. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Addressing Gaps in the Hypertension and Diabetes Care Continuum in Rural Bangladesh Through Digital Technology Supported Decentralized Primary Care: Study Protocol and Baseline Results for a Hybrid Effectiveness-Implementation Trial

Wubin Xie[1], DrPH; Sabrina Ahmed[2,3], PhD; Ali Ahsan[4], MPH; Ananya Gupta[1], MPH; Anais Masako Keenan[5], MPH; Tanmoy Sarker[4], MPH; Fahmida Akter[4,6], MSc; Aysha Anan[4], MBBS; Md Mokbul Hossain[4], MSc; Zahidul Quayyum[7], PhD; AHM Enayet Hussain[8], MBBS; Robed Amin[8], MBBS; Imran Ahmed Chowdhury[9], MPH; Mithila Faruque[10], MPH; Sohel Reza Choudhury[11], PhD; Ian Y Goon[12,13,14], PhD; Fred Hersch[15], MPH; Lora L Sabin[16], PhD; Brian Oldenburg[12,17], PhD; John Chambers[1,18], PhD; Malay Kanti Mridha[4], PhD

[1]Population and Global Health, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[10]Department of Noncommunicable Diseases, Bangladesh University of Health Sciences, Dhaka, Bangladesh

[11]Department of Epidemiology and Research, National Heart Foundation Hospital and Research Institute, Dhaka, Bangladesh

[12]Baker Heart and Diabetes Institute, Melbourne, Australia

[13]Tyree Foundation Institute of Health Engineering, UNSW Sydney, Sydney, Australia

[14]Sprightly Pte Ltd, Singapore, Singapore

[15]Google Health, Palo Alto, CA, United States

[16]Department of Global Health, School of Public Health, Boston University, Boston, MA, United States

[17]School of Psychology and Public Health, La Trobe University, Melbourne, Australia

[18]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom

[2]Department of Public Health, North South University, Dhaka, Bangladesh

[3]NCD Epidemiology Research Center, Shiga University of Medical Science, Otsu, Shiga, Japan

[4]Centre for Non-communicable Diseases and Nutrition, BRAC James P Grant School of Public Health, BRAC University, 65, Bir Uttam AK Khandakar Road, Mohakhali, Dhaka, Bangladesh

[5]Australian Institute of Family Studies, Melbourne, Australia

[6]Department of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

[7]BRAC James P Grant School of Public Health, BRAC University, Dhaka, Bangladesh

[8]Non-Communicable Disease Control, Directorate General of Health Services, Dhaka, Bangladesh

[9]BRAC Health Programme, BRAC, Dhaka, Bangladesh

Corresponding Author:
Malay Kanti Mridha, PhD
Centre for Non-communicable Diseases and Nutrition, BRAC James P Grant School of Public Health, BRAC University, 65, Bir Uttam AK Khandakar Road, Mohakhali, Dhaka, Bangladesh

Abstract

Background: Hypertension and diabetes are very common, interrelated chronic conditions. Awareness, diagnosis, treatment, and control rates of these conditions remain low, and access to quality care—particularly in rural areas—is a persistent challenge in many low- and middle-income countries. Strengthening primary health care, including the use of digital tools, is important to improve management of these chronic conditions.

Objective: This study aims to assess the implementation and effectiveness of a multicomponent, decentralized primary care model in comparison with a digital health–only intervention and usual care in rural Bangladesh.

Methods: The study applies a type 2 hybrid effectiveness-implementation design, using a 3-arm quasi-experimental approach, comprising 2 intervention arms and 1 usual care comparison arm. The study is being conducted across 3 subdistricts in the Dinajpur district, Rangpur division, northern Bangladesh. Primary outcomes include blood pressure and blood glucose control rates, assessed by population-based repeated cross-sectional surveys with independent samples, supplemented by facility-based prospective cohort data. Additionally, a mixed methods process evaluation is being conducted to capture the quantity, fidelity, adaptations, reach, and context of the interventions.

XSL•FO
RenderX

Results: The baseline community survey was conducted between January and March 2024, enrolling 6849 participants distributed across 3 arms: 2262 in usual care, 2287 in the digital-only arm, and 2300 in the multicomponent intervention arm. Participants had a mean age of 55.9 (SD 10.6) years with equal sex distribution (female: 3432/6849, 50.1%). Educational attainment was low, with 39.5% (2704/6849) of participants having no formal schooling and only 12.1% (917/6849) attaining secondary or higher education. The majority (6316/6849, 92.2%) reported being either self-employed or homemakers. The age-standardized baseline blood pressure control rate among all participants with hypertension was 10.2% overall, while the glycemic control rate among those with diabetes was 14.9%. Awareness and treatment rates for hypertension were 35.3% and 23.0%, respectively, compared to 60.7% and 34.5% for diabetes.

Conclusions: The study findings will provide critical evidence on scalable models for decentralized noncommunicable disease care and will have important implications for improving the management of hypertension and diabetes in Bangladesh and similar low-resource settings globally.

Trial Registration: ClinicalTrials.gov NCT06258473; https://clinicaltrials.gov/study/NCT06258473

International Registered Report Identifier (IRRID): DERR1-10.2196/71696

(JMIR Res Protoc 2026;15:e71696) doi:10.2196/71696

KEYWORDS

hypertension; diabetes; decentralization; task shifting; Bangladesh; Simple app

Introduction

Hypertension is the leading modifiable risk factor for cardiovascular diseases (CVDs) [1,2], affecting approximately 1 in 5 adult women and 1 in 4 adult men globally [3]. Between 1990 and 2019, the number of adults with elevated blood pressure (BP) has doubled from 650 million to 1.3 billion, largely driven by the increasing prevalence in low- and middle-income countries (LMICs) [4]. Diabetes affects 10.5% of adults worldwide, with nearly half of cases undiagnosed [1]. Poorly controlled hypertension and hyperglycemia contribute substantially to complications including heart disease, stroke, chronic kidney disease, and vision loss [1,5], as well as premature mortality [6]. Despite these health risks, the rates of awareness, treatment, and control of hypertension and diabetes remain alarmingly low in LMICs [7-9], where health systems, especially at the primary care level [10], struggle to meet the rising burden of noncommunicable diseases (NCDs) [11-13]. In 2022, the 75th World Health Assembly established the first global coverage targets for diabetes, aiming for 80% diagnosis rates among people with diabetes and 80% glycemic control among diagnosed cases by 2030. According to the World Health Organization (WHO) 2023 Global Report on Hypertension, achieving the 80-80-80 targets could prevent approximately 79 million nonfatal myocardial infarctions and 76 million cardiovascular deaths worldwide [14].

Access to NCD care remains severely constrained in the primary care systems of most LMICs, particularly in rural settings where health facilities often lack capability and capacity for NCD management [15,16]. Once diagnosed, patient retention poses a tremendous challenge for hypertension and diabetes care [17,18]. Studies in LMIC report dropout rates of over 50% of enrolled patients 6 months following treatment initiation [19,20]. Decentralizing care by task shifting (or task sharing) has been shown to be a feasible and acceptable strategy for scaling up antiretroviral therapy for HIV and AIDS care in resource-constrained settings [21-23]. This approach, which involves transferring routine management of stable patients from physician-managed clinics to peripheral facilities staffed by nonphysician health workers, enhances geographic accessibility, reduces patient costs, and enables the primary care system to serve a much larger patient population [21]. Despite these demonstrated benefits and the potential to address health care workforce shortages and improve access to HIV and AIDS care, few studies have evaluated this strategy for NCDs. Several small-scale studies have experimented with this approach [24-26], including the India Hypertension Control Initiative and the integrated tracking, referral, electronic decision support, and care coordination program for the management of hypertension and diabetes in India [22]. However, broader implementation and effectiveness remain understudied, and the impact on awareness and treatment remains largely unknown.

While digital tools for improving hypertension control and diabetes management are increasingly being tested in LMICs, the evidence for their effectiveness is mixed [27-29], and it is unclear whether standalone digital solutions can significantly improve NCD care delivery [30]. In contrast, multicomponent interventions in primary care have been shown to improve treatment outcomes in several large-scale randomized controlled trials [28,31-33]. These interventions typically combine digital tools with health care provider training, task shifting, community health worker (CHW)-led home-based BP monitoring, counseling, and referral. While showing effectiveness as a whole, partitioning the effect of individual intervention components was not possible in these studies. Nevertheless, evaluating the components individually, early evidence from LMICs shows that task shifting, when accompanied by health system restructuring, is a potentially effective strategy for improving access to NCD care [34], and that using CHWs in health programs may be effective in BP and diabetes control [35].

Among emerging digital solutions, the Simple app—developed and maintained by Resolve to Save Lives—has been increasingly used in South Asia and other LMICs [36]. Currently deployed in 7245 public health facilities in Bangladesh, Ethiopia, India, and Sri Lanka, the platform has registered 3.9 million people with hypertension, diabetes, or both conditions

XSL•FO
RenderX

[37,38]. The app addresses two critical determinants of treatment success: (1) timely medication titration and (2) therapy adherence and continuity [39]. Its design intentionally limits data collection to essential variables (eg, prescriptions, follow-up visits, and BP or blood glucose [BG] measurements) to optimize usability in busy primary care settings. However, rigorous evaluation of its impact on clinical and behavioral outcomes remains pending [37].

As part of a UK National Institute for Health Research–funded program, we previously conducted a proof-of-concept trial demonstrating the feasibility of a digital tool to support decentralized care for hypertension and diabetes management in rural Bangladesh [26]. Building on these promising findings, we designed a 3-arm hybrid implementation-effectiveness trial to evaluate the impact of the Simple app within a multicomponent intervention framework, addressing critical gaps in understanding both effectiveness in clinical outcomes and implementation challenges. The paper describes the study methods and baseline findings, serving as a reference framework for subsequent publications from the Dinajpur study.

Methods

Study Setting and Populations

Bangladesh, an LMIC in South Asia with a total population of 171.5 million, has undergone major demographic and epidemiologic transitions in recent decades, and NCDs now account for approximately 70% of all deaths nationally [40]. The prevalence of hypertension and diabetes among Bangladeshi adults is estimated to be 27.4% and 9.8%, respectively [41,42]. The awareness, treatment, and control rates among people living with these conditions remain alarmingly low [41,42]. Furthermore, tobacco use, insufficient fruit or vegetable intake, and overweight are highly prevalent [43]. Over the past decade, the Government of Bangladesh has tried to implement national multisectoral actions to strengthen NCD care, most notably through the establishment of dedicated NCD care delivery points (known as “NCD Corners”) in subdistrict hospitals (“Upazila health complex” in local language) since 2011. The initiative’s planned decentralization to village-level primary care facilities (community clinics [CCs]) has not materialized [41]. Consequently, access to public NCD services remains severely limited in rural Bangladesh [41]. The Simple app has been progressively implemented across primary NCD care facilities in Bangladesh since 2020. As of current reporting, more than 180 facilities use the platform to manage over 350,000 hypertension and diabetes patients. The app’s integration with the national health management information system enables real-time performance monitoring and feedback for health facilities.

Study Design

This hybrid effectiveness-implementation trial simultaneously evaluates intervention effectiveness and implementation strategies [44], using a 3-arm, quasi-experimental design. The aims of this trial are (1) to evaluate the effectiveness of a multicomponent, decentralized care model for hypertension and diabetes management within the public primary care systems, compared to both usual care and a standalone digital health intervention (Simple app); (2) to examine implementation processes, including explanatory factors influencing intervention effectiveness and barriers to and facilitators of delivery and sustainability; and (3) to undertake an economic evaluation. We hypothesize that compared with usual care, the multicomponent decentralized primary care—supported by the Simple app—will improve all steps along the hypertension and diabetes care continuum. Conversely, we hypothesize that the mobile health intervention alone (Simple app) may improve BP and glycemic control compared with usual care but will have a limited impact on rates of screening, diagnosis, and treatment; multicomponent integrated care will lead to a higher treatment success rate compared to the mobile health intervention alone.

The study is being conducted across 3 subdistricts in the Dinajpur district, Rangpur division, northern Bangladesh (Figure 1 and Table 1). The multicomponent intervention is being implemented in the Chirirbandar subdistrict, while the digital health–only intervention is being implemented in the Parbatipur subdistrict. The Biral subdistrict serves as the reference site. Primary outcomes include BP and glycemic control rates, assessed by population-based repeated cross-sectional surveys with independent samples, supplemented by facility-based prospective cohort data. Additionally, a mixed methods process evaluation is being conducted to capture the quantity, fidelity, adaptations, reach, and context of the interventions. The study duration is 36 months, including an intervention period of 24 months.

XSL•FO
RenderX

Figure 1. Study sites: Dinajpur district, Rangpur division, Bangladesh. Multicomponent interventions are implemented in the Chirirbandar subdistrict, a digital-only intervention is implemented in the Parbatipur subdistrict, and the Biral subdistrict is the usual care arm. CC: community clinic; UHC: Upazila health complex.

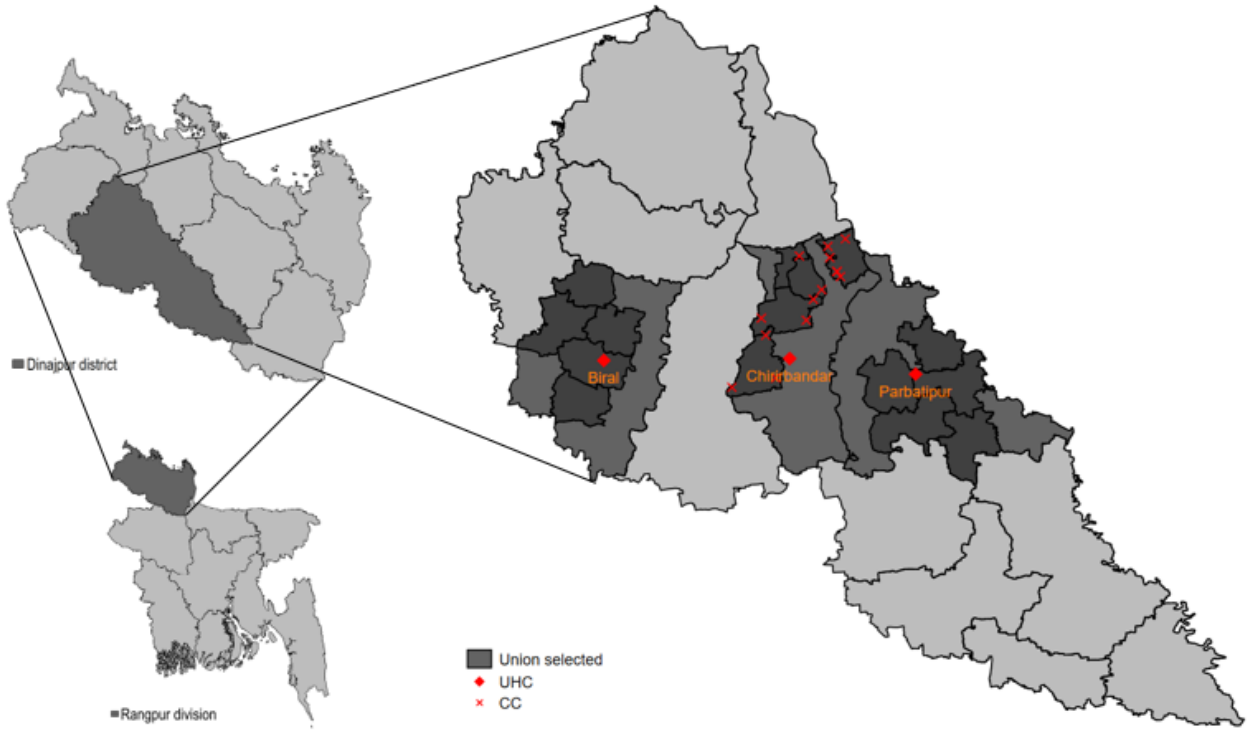

Table . Population and health facility statistics in study areas.[a]

Characteristics	Study sites		
	Chirirbanda (multicomponent)	Parbatipur (digital health only)	Biral (usual care)
Area (square km)	312.7	395.0	353.4
Population, n	292,500	365,103	257,925
Adults aged 40+ years, %	25.6	25.4	25.3
Persons aged 60+ years, %	7.3	7.2	7.3
Muslim, %	76.3	85.5	72.0
Urban, %	3.0	11.0	3.5
Unions, n	12	10	10
Villages (wards), n	141	229	238
Health facilities, n			
UHC[b]	1	1	1
Community clinics	35	40	34
Literacy rate, %	52.9	53.9	47.3

[a]Data source: 2011 Bangladesh Census, Bangladesh Ministry of Health and Family Welfare facility registry [45,46].

[b]UHC: Upazila health complex.

Ethical Considerations

This study was approved by the Institutional Review Board of James P Grant School of Public Health, BRAC University (reference number: IRB-16 November-23 - 041). Written informed consent was obtained from all participants. Identifying information associated with the study participants will be kept confidential through unique identifying numbers on all paper forms and computer-based files to protect privacy and ensure confidentiality of data being collected. Participants received no compensation for participation.

XSL•FO
RenderX

Interventions

Development of the Interventions

Building on an assessment of hypertension and diabetes management barriers at the patient, provider, and health system levels [47-50], along with evidence from previous intervention studies and the updated UK Medical Research Council (MRC) guidance on developing and evaluating complex interventions [31-33,51], a multicomponent intervention package has been designed to increase access to primary care and to improve care quality and patient retention (Figure 2).

Figure 2. NCD service delivery hierarchy and decentralized care (adapted from Xie et al [26], which is published under Creative Commons Attribution 4.0 International License). CC: community clinic; CHCP: community health care provider; CHW: community health worker; NCD: noncommunicable disease; Tx: treatment; UHC: Upazila health complex.

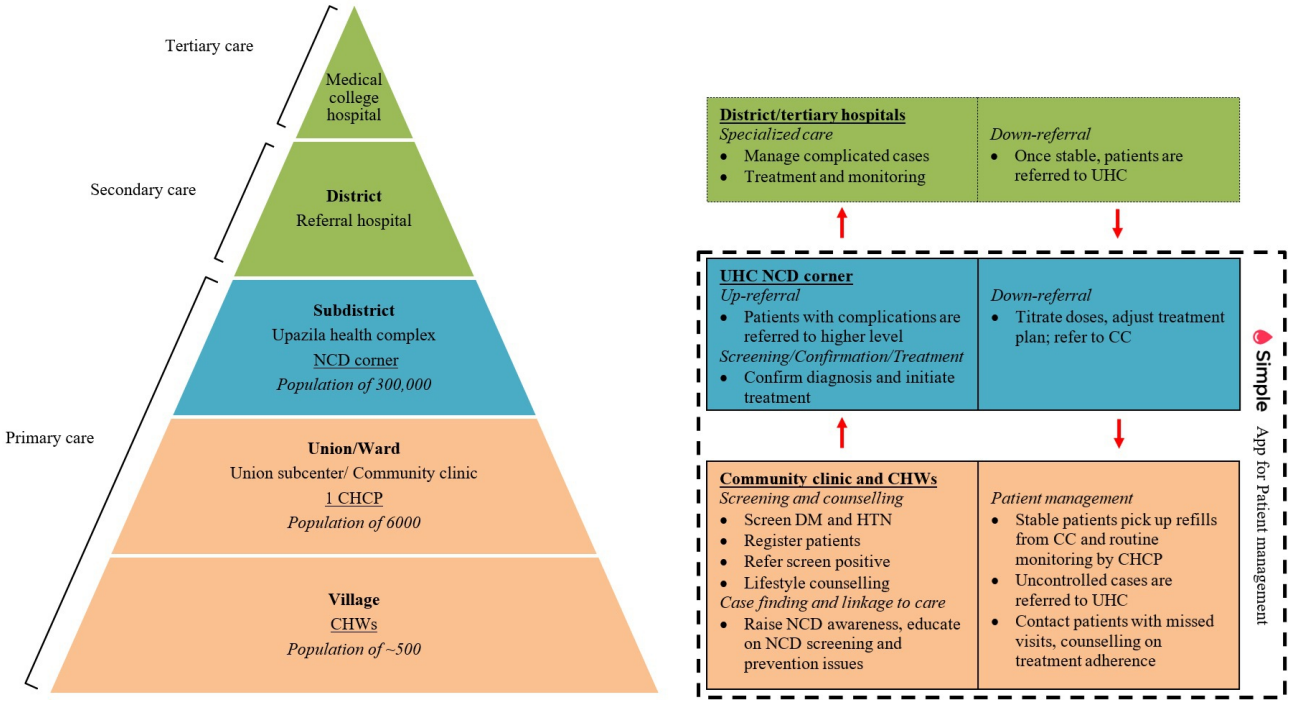

Multicomponent Decentralized Care

Digital Health for Patient Management

In alignment with the Bangladesh government recommendations, the Simple app is being used in primary care facilities for coordinated hypertension and diabetes management. The app supports prescription and continuity of care through unique patient IDs and enables functions including: (1) monitoring patient longitudinal BP and BG changes, prescription history, and prompting titration when indicated; (2) flagging overdue patients for action; and (3) generating performance reports on indicators such as patient enrollment and BP and BG control rates. A dedicated project assistant supports health workers with data entry and follow-up with overdue patients.

Decentralization With Task Sharing

Supported by the subdistrict NCD corner, 15 CCs were equipped to conduct screening, follow-ups, and medication refills. Implementation includes several core components: (1) providing regular trainings (every 6 months) for health care providers from the subdistrict NCD corner and CCs on team-based NCD care, (2) equipping CCs for routine care, and (3) providing regular supportive supervision by health officials. CCs received medical devices, Android tablets for record keeping, and essential medications.

Involvement of Community Health Workers

The intervention incorporates hypertension and diabetes management into community-based care through trained CHWs, leveraging evidence from successful LMIC programs [31-33,51-54]. CHWs perform 4 core functions: community screening and case identification, lifestyle counseling and health education, referral coordination, and patient follow-up. Capacity building includes an initial 6-day training conducted in June 2024, supplemented by biannual refresher trainings.

Supportive Monitoring and Supervision and Team-Based Care

The intervention incorporates a multitiered supportive supervision system to ensure quality service delivery. At the subdistrict level, NCD corners receive regular oversight from senior health administrators and medical professionals who address clinical management challenges, medication supply issues, and operational barriers. These supervisory visits are data-driven, using performance metrics generated by the Simple app dashboard to identify priority areas for improvement.

A cascading supervision framework extends to CCs, where subdistrict health officials conduct routine monitoring visits. Community health care providers (CHCPs) similarly provide ongoing supervision to CHWs in their catchment areas. To strengthen health system coordination, monthly meetings convene providers across levels to enhance collaboration, troubleshoot challenges, and optimize care processes.

XSL•FO
RenderX

Throughout the intervention period, dedicated technical support teams maintain system functionality by providing continuous assistance with the Simple app operations and troubleshooting.

Care Continuum Within the Multicomponent Decentralized Care

CHWs conduct targeted health promotion activities to increase NCD and their risk factors, with particular emphasis on encouraging adults aged ≥40 years to undergo screening at CCs. Adults detected with elevated BP or BG at CCs are referred to the subdistrict NCD corner for confirmation and care plan initiation if diagnosed. Patients with known hypertension and diabetes are referred to the NCD Corners for care plan initiation or resumption (for those who discontinued treatment).

At the subdistrict NCD corner, a structured clinical workflow ensures comprehensive patient management. Nurses perform detailed clinical assessments and maintain up-to-date patient records in the digital system. Physicians initiate care plans for new patients and decide if existing patients can be referred to CCs for routine follow-up and management according to the national guidelines (ie, having a controlled status in the three most recent consecutive visits). Prescriptions are guided by the Simple app BP/BG record and prescription history; treatment titration may be done when indicated.

At CC, CHCPs deliver essential NCD services for stable patients, comprising (1) routine clinical monitoring, (2) prescribed medication dispensing, and (3) lifestyle modification counseling. Patients maintaining treatment targets continue community-based management through scheduled follow-ups. For cases demonstrating suboptimal control or presenting with acute symptoms, CHCPs initiate immediate referral to the subdistrict NCD corner for physician evaluation and therapeutic change.

Digital Health–Only Intervention

In the subdistrict (Parbatipur) with digital health–only intervention, initial training on the national protocols for hypertension and diabetes management and the Simple app in the NCD corner was conducted in July 2024. Biannual refresher training will be done throughout the implementation period. A project assistant is hired to help with the Simple app data entry and to follow up with overdue patients. CCs and CHWs are essentially not involved in care provision. Patient pathways remain the same as in usual care.

Usual Care

The usual care subdistrict (Biral) receives biannual training on national protocols for hypertension and diabetes management. Existing usual care is being provided by the subdistrict NCD corner, including screening, treatment initiation, drug refill, and routine follow-up. CCs and CHWs are not involved in NCD care provision. The local government agreed to delay the rollout of the Simple app in the comparison subdistricts until the end of the study to avoid contamination.

Effectiveness and Implementation Outcomes and Assessment

The primary outcome is the proportion of treated patients achieving or maintaining disease control according to national standards and WHO PEN (WHO Package of Essential Noncommunicable Diseases Interventions) protocols, with hypertension control defined as systolic/diastolic BP <140/90 mm Hg and glycemic control as fasting capillary glucose < 7.0 mmol/L or random capillary glucose <11.1 mmol/L. Effectiveness is being assessed primarily using data from repeated community-based surveys. Secondary outcomes of this study include hypertension and diabetes care cascade (ie, percentage of patients ever screened, percentage aware of their condition, and percentage on treatment), changes in health behaviors (eg, smoking cessation among hypertension and diabetes patients, and meeting recommended weekly physical activity level). To complement assessment of population-level changes using repeated community-based surveys, facility-level data will be extracted to evaluate changes in care quality at primary care facilities.

Specific aim 2 is evaluated by using the RE-AIM (reach, effectiveness, adoption, implementation, and maintenance) framework (Table 2) with patient assessments and stakeholder interviews, informed by previous studies [25,55]. Implementation fidelity and process evaluation are guided by the MRC guidelines on process evaluation of complex interventions [56] and the WHO's NCD Facility-Based Monitoring Guidance [57]. The program theory is depicted in Figure 3. Data sources include training reports, prescription practice captured by the Simple app, facility records, and patient registries. The Simple app captures essential data related to patient background, prescriptions and titration, dates of follow-up visits, and longitudinal BP and BG records. Baseline qualitative data were collected from patients, health care providers, CHWs, and public health managers to inform the evaluation of barriers and enablers of implementing the interventions in the primary care system for improved NCD outcomes. Another round of qualitative data collection will be conducted upon completion of the interventions.

XSL•FO
RenderX

Table . Data collection plan for evaluation indicators.

RE-AIM[a] domains or indicators	Operational definition	Data source	Instrument	Timeline, at months
Effectiveness				
Primary outcomes				
Percentage of patients with HTN[b] who achieved BP[c] control	BP <140/90 mm Hg	Community surveys	Omron HEM-7120	0, 12, 24
Percentage of patients with DM[d] who achieved glycemic control	Fasting capillary glucose <7.0 mmol/L or random capillary glucose <11.1 mmol/L	Community surveys	Accu-Chek Instant	0, 12, 24
Secondary outcomes				
Percentage of adults with HTN aware of their condition	Ever been told by a health care provider that they have raised BP or HTN	Community surveys	Questionnaire	0, 12, 24
Percentage of adults with DM aware of their condition	Ever been told by a health care provider that they have diabetes	Community surveys	Questionnaire	0, 12, 24
Percentage of adults with diagnosed HTN on treatment	Currently taking medication for HTN, not including herbal or traditional remedy	Community surveys	Questionnaire	0, 12, 24
Percentage of adults with diagnosed DM on treatment	Taking medication for DM, not including herbal or traditional remedy	Community surveys	Questionnaire	0, 12, 24
Percentage of HTN and DM patients who quit smoking	Quit smoking is defined as not having smoked, even 1 or 2 puffs, during the past 6 months	Community surveys	Tobacco Use Questionnaire	0, 12, 24
Percentage of HTN and DM patients who met the recommended PA[e] level	≥150 minutes MVPA [f] per week	Community surveys	WHO STEPS[g]	0, 12, 24
Reach				
Percentage of adults with HTN screened for HTN	Had a BP measurement in past 12 months	Community surveys	Questionnaire	0, 12, 24
Percentage of adults with DM screened for DM	Had a BG[h] measurement in past 12 months	Community surveys	Questionnaire	0, 12, 24
Percentage of patients with HTN and DM receiving treatment from public PHC[i] facilities	Currently receiving care from government PHC facilities	Community surveys	Questionnaire	0, 12, 24
Adoption				
Percentage of patients with HTN and DM registered in the Simple app	Information recorded in the Simple app, separated for CC[j] and NCD[k] corner	Simple app	Simple app	Quarterly
Percentage of CCs who conducted NCD screening	Screening ≥50 adults/month	Simple app	Simple app	Quarterly
Percentage of CCs who performed routine care for patients with NCD	Following up with ≥25 patients/month	Simple app	Simple app	Quarterly

XSL•FO
RenderX

RE-AIM[a] domains or indicators	Operational definition	Data source	Instrument	Timeline, at months
Percentage of CHWs[l] who conducted home visits	Home visits for ≥5 patients/month for lifestyle counseling and adherence support	CHW report	CHW report	Quarterly
Implementation				
Percentage of providers who participated in training and refresher	Participated in training on NCD management	Training report	Training registry	Biannually
Percentage of CCs with functional essential equipment and supply	BP/BG machines are calibrated, strips, and IEC[m] materials available	Facility records	Facility report	Quarterly
Percentage of planned supportive supervision conducted	8 planned supervisions, separated by levels of PHC and CHW	Supervision report	Checklist	Quarterly
Maintenance				
Percentage of newly enrolled patients retained in care	LTFU[n]: ≥3 months late for the last scheduled visit, 6/12 months	Patient registry	Patient registry	12, 24, 30
CCs who performed routine care for NCD patients	Following up with ≥25 patients/month	Patient registry	Patient registry	30
NCD corner that used the Simple app for patient management	≥80% patient encounters recorded in the Simple app	Patient registry	Patient registry	30

[a]RE-AIM: reach, effectiveness, adoption, implementation, and maintenance framework.

[b]HTN: hypertension.

[c]BP: blood pressure.

[d]DM: diabetes mellitus.

[e]PA: physical activity.

[f]MVPA: moderate to vigorous physical activity.

[g]WHO STEPS: World Health Organization's STEPwise approach to NCD risk factor surveillance.

[h]BG: blood glucose.

[i]PHC: primary health care.

[j]CC: community clinic.

[k]NCD: noncommunicable disease.

[l]CHW: community health worker.

[m]IEC: information, education, and communication.

[n]LTFU: lost to follow-up.

XSL•FO
RenderX

Figure 3. Program theory. (A)HI: assistant health inspector; BG: blood glucose; BP: blood pressure; CC: community clinic; CHCP: community health care provider; CHW: community health worker; FU: follow-up; HCP: health care provider; IEC: information, education, communication; NCD: noncommunicable disease; PPP: public-private partnership; SAC: stakeholder advisory committee; SES: socioeconomic status; Tx: treatment.

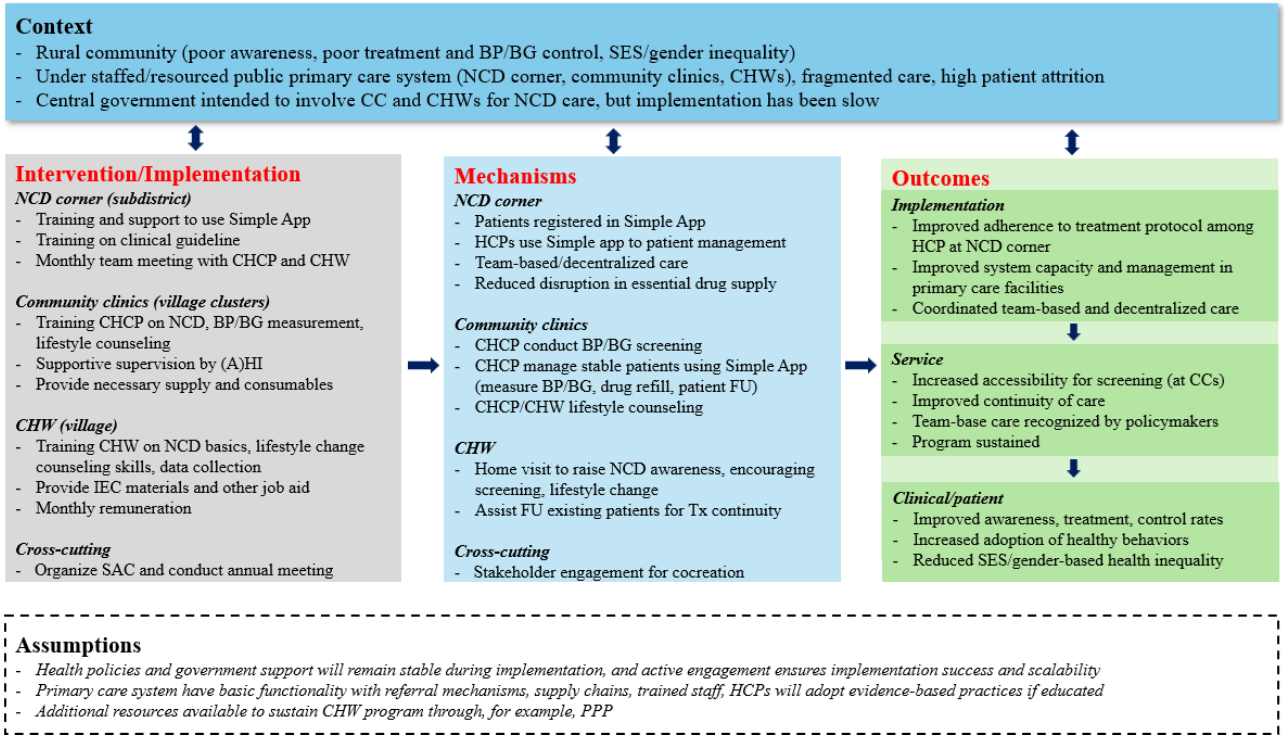

We anticipated some minor practical refinements and adaptations of the intervention components during the implementation process. Any refinements and the rationale for these are being documented for transparent reporting, guided by the program theory and updated MRC guidance.

For specific aim 3, the primary cost-effectiveness measure is the incremental cost per 1-percentage-point increase in the proportion of participants achieving control status over the 24-month intervention period. Detailed costs by input categories and steps in intervention components will be estimated. Both health system costs and patient costs will be compared between the intervention and usual care arms.

Sampling Strategy and Power Calculation

Three community-based surveys are undertaken at months 0, 12, and 24. For each evaluation survey, an independent random sample of adults aged 40 years and above who are community residents in 3 study subdistricts is randomly selected through a multistage cluster sampling approach. We opted for a quasi-experimental design with repeated cross-sectional surveys, instead of a traditional randomized trial, to capture changes in the entire care continuum at the population level and to generate real-world evidence on effectiveness and implementation strategies for greater external validity [58,59]. A buffer zone between the 2 intervention subdistricts (ie, Chirirbandar and Parbatipur) was created to mitigate potential spillover effects between contiguous communities. Given the available resources and time constraints, we randomly select a total of 15 villages from each subdistrict. The villages are divided into segments of approximately 250 households without disrupting the boundaries of the villages. One segment per village is then randomly selected. Subsequently, a list of adults aged 40 years of age within each village of the selected segment is obtained. From each segment (cluster), an equal number of women and men ≥40 years of age are randomly selected by using systematic random sampling with the condition that no more than 1 adult male and 1 adult female is included from the same household.

The planned sample size was determined to be 6750 participants for each community survey, with an equal number of clusters per subdistrict (15 clusters per subdistrict, or 45 clusters total) and similar cluster sizes (150 participants: 75 females and 75 males per cluster). Assuming a prevalence of diagnosed hypertension of 40%, and 15% for diabetes among adults aged 40 years and above, based on South Asia Biobank [60] data (unpublished findings), we expected to collect information from 2700 individuals with hypertension and 1000 individuals with diabetes. Assuming a conservative intraclass correlation coefficient of 0.02 [25,61], and a 2-sided type I error rate of 5%, the survey has 80% power to detect a difference as small as 7 percentage points between the intervention and reference groups for BP control, and a 10 percentage point difference for glycemic control. Power calculations were performed using Stata 17 (Stata Inc.) power analysis for a clustered 2-sample proportions test [62]. The planned sample size for binary outcomes ensures sufficient power for continuous specifications of the outcomes [25,61,63].

Data Collection, Management, Quality, and Security

Baseline community surveys used interviewer-administered questionnaires and physical examinations. The baseline questionnaire, an adapted version of the WHO STEPS (World Health Organization's STEPwise approach to NCD risk factor surveillance) questionnaire, covers sociodemographic characteristics, comorbidities, medication use, and health

XSL•FO
RenderX

behaviors. The questionnaire was translated into Bengali and piloted for clarity and digitized using Kobo Toolbox for data collection. BP was measured by research staff using Omron HEM-7120. Capillary BG after at least 8 hours of fasting was measured using Accu-Chek Instant (Roche Diabetes Care).

Facility-based data collection is being conducted via the Simple app by on-site staff. Demographic and NCD history, medication use, BP/BG measurements, and treatment dosages are being recorded. In usual care subdistricts, the same data are being collected through patient registry reviews. Additional facility-level data, such as medication availability, device functionality, and treatment success rates, are being collected for process evaluation. An acceptability and utility survey of the Simple app for NCD management will be administered to health care providers in intervention subdistricts at months 12 and 24. All providers involved in hypertension or diabetes care at NCD corners and CCs will be invited to participate.

All health care providers in primary care facilities who are directly involved in NCD care were approached for qualitative data collection at baseline. This included 15 doctors and nurses from 3 subdistrict NCD corners and 15 CHCPs from CCs. Additionally, 6 public health officials (2 from each subdistrict) were approached for an in-depth interview (IDI). Furthermore, 9 focus group discussions (FGDs) with patients living with hypertension, diabetes, or both conditions were conducted. For each FGD, 8 - 12 patients were included. Participants were purposively selected to maximize the diversity of the sample on sociodemographic characteristics (eg, age, sex, religion, socioeconomic status, and geographic distance to health facilities) and NCD history (new and experienced). Qualitative data collection will be repeated by the end of the study with the same group of public health officials and NCD care providers. In cases of staff turnover, the ones who are in position at the time of data collection will be approached. Patient FGDs will be repeated with the same groups of participants. Moreover, all 15 CHWs will be approached for an FGD session. FGD/IDI guides were developed based on the WHO's 7-domain framework of health care delivery [64], chronic care model [10], and prespecified program theory. Interviews will be audio-recorded and are expected to each take 30 - 45 minutes to complete. Qualitative data collection will be performed at baseline and endline surveys.

The qualitative data collected from health administrators, clinicians, and CHWs in intervention subdistricts provide qualitative insights on barriers and facilitators to NCD care delivery, experiences with the Simple app, study participation, workflow restructuring, and teamwork with clinics and CHWs. Data collection at endline allows participants to reflect on changes in these aspects over time. Patient FGDs explore barriers and facilitators to care accessibility and perceptions of care quality, including specific intervention components like drug refills, CHW home visits, and CHCP follow-ups.

Hypertension- and diabetes-related treatment costs were collected for economic evaluation, including direct and indirect costs (eg, transportation, food, childcare, and lost work time) at baseline. Health service delivery costs, including staff, transportation, laboratory, training, utilities, and overheads, are being assessed using microcosting to track time and resources spent on activities [65]. Intervention costs account for the setting, target population, resources, and consumables such as medical supplies and overheads.

To protect privacy and ensure confidentiality of data being collected, identifying information associated with the study participants is being kept confidential through unique identifying numbers on all paper forms and computer-based files. This file linking names and study numbers is password-protected, only accessible to authorized study personnel. All electronic data are encrypted, password protected, and stored in secure computer networks. All study personnel were trained to follow standard protocols.

Adverse Event Monitoring

Serious adverse events, including death, hospitalization, and other conditions that result in persistent or significant disability, were handled by medical professionals at the subdistrict NCD corner as per standard protocol. Health care providers at CC were required to send patients back to the subdistrict NCD corner for evaluation when persistent poor control of BP and BG is observed among patients managed at CCs. Adverse events were recorded by study personnel.

Statistical Analysis Plan

Baseline characteristics of participants, including sociodemographic characteristics, medical history, anthropometrics, and lifestyle factors, will be compared between the 2 intervention and 1 control subdistricts using 1-way ANOVA or Rao-Scott χ^2 tests. Analyses will be stratified by hypertension and diabetes status. A difference-in-difference estimate for hypertension and diabetes will be implemented with multivariate logistic regression analyses, which take the following form:

$$\log(p/1-p) = \beta_0 + \beta_1 * \text{Time} + \beta_2 * \text{Intervention} + \beta_3 * [\text{Time} * \text{Intervention}] + \beta_4 * \text{Covariates} + \varepsilon$$

The β_3 coefficient captures the difference in change over time. Clustered standard errors at the village segment level will be specified to allow for intragroup correlation, relaxing the usual requirement that the observations within a cluster are independent [66]. Several individual- and area-level covariates (β_4) will be included in the analytic models to mitigate confounding by potential different sample composition. These covariates will include age, sex, education, wealth, tobacco use, physical activity, self-reported fruit and vegetable intake, self-reported history of chronic disease (eg, CVD, chronic respiratory disease, cancer, and chronic kidney disease), and rural or urban status. Given that the evaluation follows a multistage process whereby village segment clusters are randomly sampled from each subdistrict, and individuals are sampled randomly from the village segment clusters, a clustering adjustment for standard error is necessary to avoid inflating the precision of the estimated intervention effect [67,68]. Marginal effects will be calculated to illustrate how the predicted probability (ie, the proportion of patients with controlled conditions) changes over time among the three groups. Potential heterogeneity of effects will be explored in subgroup analyses, defined by age, sex, religion, and socioeconomic status. Several sensitivity analyses will be performed to check the robustness

XSL•FO
RenderX

of the findings to critical model assumptions and missing data. While the "parallel" assumption should be tested with repeated measures prior to the intervention, this was not feasible given the budget and other limitations. Instead, we are testing the assumption with a negative control in treatment or outcome. Second, we will test the robustness of the results to different modeling approaches to handle geographical clustering of participants and over time. Third, missing data patterns and potential mechanisms (ie, whether missing at random can be assumed) will be assessed. If missing data only affect a small proportion (ie, <5%) of the sample, listwise deletion will be done; otherwise, missing data will be handled with multiple imputation with chained equations if missing at random assumption is reasonable.

The facility-based cohort data collected longitudinally will capture patients' treatment trajectory over the study period. Changes in the proportion of patients with a controlled condition will be analyzed with generalized estimating equation Poisson regression with robust variance [69,70]. Continuous changes in systolic BP and diastolic BP from baseline will be assessed with linear mixed effects models.

For qualitative data, dual Bengali-English language speakers on the study team will transcribe and translate the audio-recorded IDIs and FGDs into English. The English language version of the transcripts will then be coded and analyzed thematically using NVivo, a qualitative data software program developed by Lumivero. Qualitative data will allow the identification of specific barriers to and facilitators of interventions by exploring the experiences of patients and clinicians engaged in different models of NCD care delivery. Qualitative data analysis will be guided by prespecified program theory, the RE-AIM framework, and standard grounded theory to identify themes, build and apply codebooks, and describe thematic characteristics, patterns, and relationships [71,72].

For the economic evaluation, we will compare the costs and effects of intervention with the usual care group, both from financial and economic perspectives [73]. Financial cost per unit outcome will be calculated by dividing the total cost by the quantifiable unit of outcome (screening, treatment, retention in care, and control) for each of the three groups. Incremental cost-effectiveness ratios (ICERs) will be computed to compare the additional costs and effects of each intervention with the usual care. A nonparametric bootstrap with 1000 replications will be used to estimate 95% CIs around the point estimate of ICERs. All local currency (Bangladeshi Taka) costs will be converted to US dollars using the prevailing exchange rates and will be adjusted for inflation rates, discounted at 3%, and expressed in the 2025 US dollar present value terms. Sensitivity analysis will explore the robustness of the results to alternative probability distributions, time horizon, and uncertainty of key variables.

Descriptive statistics were reported for baseline characteristics of the sample stratified by subdistricts. Percentage estimates for key primary and secondary outcomes were age-standardized to the 2023 Bangladeshi population using age groups 40 - 49, 50 - 59, and ≥60 years, derived from the United Nations' Population Division of the Department of Economic and Social Affairs World Population Prospects. We used a direct standardization approach using the STDIZE and STDWEIGHT Stata commands. The preliminary analyses were performed using Stata 18.0 (StataCorp Inc).

Stakeholder Engagement

To ensure that the strategies developed are feasible, contextually appropriate, and sustainable in real-world settings, stakeholder groups (including government agencies, health care providers, and community representatives) have been actively engaged to enable cocreation. At the study design phase, input from relevant government agencies, major nongovernmental organizations, and community representatives was sought. During implementation, monthly meetings with subdistrict-level public health officials and CHWs were convened to facilitate bidirectional exchange of knowledge. As a member of the Bangladesh national committee for the revision of the national hypertension and diabetes protocol, the site principal investigator has shared findings with national policymakers. A stakeholder advisory group has been convened, with its first annual meeting planned for the end of 2025.

Results

A baseline community survey was conducted between January and March 2024. The baseline data consist of a sample of 6849 adults aged 40 years and above randomly selected from 45 villages from 3 subdistricts (Biral, Parbatipur, and Chirirbandar) in Dinajpur district. Exclusion criteria included pregnancy, adults with terminal illness who have difficulty completing the survey or taking anthropometric measurements and physical examinations, and people who are unable to give consent. The sample was drawn with a 3-stage cluster sampling scheme. The first stage included a selection of 45 villages, which were selected stratified by subdistrict and proportional to population size. Stage 2 included the selection of one segment from each village, with each segment consisting of ~250 households. Household listing was done to create a sampling frame of male and female adults aged 40 years and above currently residing in the selected villages. A simple random sample of 88 male adults and 88 female adults was selected with the condition that not more than 1 male or female could be selected from the same household. The response rate was 86%.

The social and demographic characteristics of participants involved in the baseline survey are presented in Table 3. Overall, the mean age of the participants was 55.9 (SD 10.6) years; almost exactly 50.1% (3432/6849) were female. Forty percent of participants (2704/6849) had no formal schooling, and only 12.1% (829/6849) had secondary or higher education. The majority (6316/6849, 92.2%) of the sample was self-employed or worked as a homemaker. The majority of the sample was Muslim (5313/6849, 77.6%) and was currently married (5932/6849, 86.6%). Other than religion, employment, and wealth score, there was no statistical difference among the 3 subdistricts (arms) based on these sociodemographic characteristics.

XSL•FO
RenderX

Table . Sample characteristics of baseline community survey.

Variables	Total	Comparison groups			
		Usual care	Digital only	Multicomponent	*P* value
Participants, n (%)	6849 (100)	2262 (33.0)	2287 (33.6)	2300 (33.4)	
Age in years, mean (SD)	55.9 (10.6)	56.0 (10.4)	56.1 (10.9)	55.7 (10.6)	.37
Sex, n (%)					.96
Male	3417 (49.9)	1134 (50.1)	1143 (49.7)	1140 (49.8)	
Female	3432 (50.1)	1128 (49.9)	1157 (50.3)	1147 (50.2)	
Education attainment, n (%)					.73
No formal schooling	2704 (39.5)	916 (40.5)	897 (39.0)	891 (39.0)	
Primary	3319 (48.4)	1075 (47.5)	1116 (48.5)	1125 (49.2)	
Secondary or higher	829 (12.1)	271 (12.0)	287 (12.5)	271 (11.8)	
Marital status, n (%)					.58
Currently married	5932 (86.6)	1954 (86.4)	1983 (86.2)	1995 (87.2)	
Widowed or other	917 (13.4)	308 (13.6)	317 (13.8)	292 (12.8)	
Religion, n (%)					<.001
Muslim	5313 (77.6)	1741 (77.0)	1868 (81.2)	1704 (74.5)	
Non-Muslim	1536 (22.4)	521 (23.0)	432 (18.8)	583 (25.5)	
Employment, n (%)					<.001
Employed or retired	212 (3.1)	69 (3.1)	72 (3.1)	71 (3.1)	
Self-employed or homemaker	6316 (92.2)	2093 (92.5)	2082 (90.6)	2141 (93.6)	
Unemployed	319 (4.7)	100 (4.1)	144 (6.3)	75 (3.3)	
Wealth score, n (%)					.004
Low	2261 (33.0)	736 (32.5)	761 (33.1)	764 (33.4)	
Medium	2270 (33.1)	703 (31.1)	768 (33.4)	799 (34.9)	
High	2318 (33.8)	823 (36.4)	771 (33.5)	724 (31.7)	
BMI (kg/m^2) category, n (%)					.30
<18.5	911 (13.3)	306 (13.5)	321 (14.0)	284 (12.4)	
18.5 - 22.9	3067 (44.8)	1041 (46.0)	1024 (44.5)	1002 (43.8)	
23.0 - 27.4	2185 (31.9)	701 (31.0)	731 (31.8)	753 (32.9)	
≥27.5	686 (10.0)	214 (9.5)	224 (9.7)	248 (10.8)	
Elevated WC[a], n (%)	2583 (37.7)	827 (36.6)	878 (38.2)	878 (38.4)	.45
Mental health[b], n (%)					
Depression	666 (9.7)	252 (11.1)	191 (8.3)	223 (9.8)	.005
Anxiety	361 (5.3)	161 (7.1)	80 (3.5)	120 (5.2)	<.001
Family history, n (%)					
Premature CVD[c]	702 (10.2)	249 (11.0)	252 (11.0)	201 (8.8)	.02
Diabetes	1155 (16.9)	399 (17.6)	355 (15.4)	401 (17.5)	.09
Self-reported medical history, n (%)					
CVD	202 (2.9)	54 (2.4)	70 (3.0)	78 (3.4)	.12

XSL•FO
RenderX

Variables	Total	Comparison groups			
		Usual care	Digital only	Multicomponent	*P* value
COPD[d]/asthma	307 (4.5)	111 (4.9)	125 (5.4)	71 (3.1)	<.001
Other NCD[e]	231 (3.4)	75 (3.3)	74 (3.2)	82 (3.6)	.76
Hypertension, n (%)	2690 (39.3)	876 (38.7)	959 (41.7)	855 (37.4)	.01
Diabetes, n (%)	967 (14.1)	329 (14.5)	298 (13.0)	340 (14.9)	.14

[a]WC: waist circumference.

[b]Depression and anxiety were assessed using PHQ-2 and GAD-2, respectively, with a cutoff point of ≥3 as indicating these mental health symptoms.

[c]CVD: cardiovascular disease.

[d]COPD: chronic obstructive pulmonary disease.

[e]NCD: noncommunicable disease.

Among all participants, the prevalence of hypertension, diabetes, overweight and obesity, and self-reported CVD were 39.3% (2690/6849), 14.1% (967/6849), 41.9% (2871/6849), and 2.9% (202/6849), respectively. Ten percent (702/6849) of the participants had a family history of premature CVD, and 16.9% (1155/6849) had a family history of diabetes. Across the 3 arms (subdistricts), the digital intervention arm appeared to have a higher prevalence of hypertension compared with the multicomponent and the usual care arms (855/2300, 41.7%; 959/2287, 37.4%; and 876/2262, 38.7%; respectively).

The baseline age-standardized BP control rate among participants with hypertension was 10.2%, while baseline glycemic control among participants with diabetes was 14.9% (Table 4). There were slight differences across the 3 subdistricts. For secondary outcomes, age-standardized awareness and treatment rates for hypertension were 35.3% and 23.0%, respectively, and 60.7% and 33.9% for diabetes, respectively. Among participants living with hypertension, diabetes, or both conditions, 47.1% met the physical activity level recommended by WHO, and 48.5% had quit smoking.

XSL•FO
RenderX

Table . Outcome measures at baseline.[a]

Outcomes	Total, n (%)	Comparison groups, n (%)			
		Usual care	Digital only	Multicomponent	*P* value
Primary outcomes					
Achieved blood pressure control[b]	263 (10.2)	67 (8.5)	80 (8.8)	116 (13.4)	.003
Achieved glycemic control[c]	154 (14.9)	72 (20.0)	36 (11.0)	46 (13.5)	.001
Secondary outcomes					
Awareness of hypertension status[b]	942 (35.3)	287 (33.4)	324 (34.2)	331 (38.6)	.06
On treatment for hypertension[b]	635 (23.0)	166 (19.1)	212 (21.3)	257 (29.0)	<.001
Awareness of diabetes status[c]	591 (60.7)	215 (65.3)	169 (56.5)	207 (60.2)	.09
On treatment for diabetes	334 (33.9)	96 (27.3)	105 (34.5)	133 (39.1)	.03
Patients with hypertension/diabetes who met physical activity target[d]	1461 (47.1)	469 (45.9)	441 (41.2)	551 (54.6)	<.001
Patients with hypertension/diabetes who quit smoking[e]	579 (48.5)	189 (48.2)	205 (52.7)	185 (45.4)	.05

[a]Percentage estimates are age-standardized to the 2023 Bangladeshi population using age groups 40 - 49, 50 - 59, and ≥60 years, derived from the United Nations' Population Division of the Department of Economic and Social Affairs World Population Prospects.

[b]The number of participants with hypertension was 2690 overall (867, 959, and 855 for usual care, digital-only, and multicomponent arms, respectively).

[c]The number of participants with diabetes was 967 overall (329, 298, and 340 for usual care, digital-only, and multicomponent arms, respectively).

[d]The number of participants with hypertension, diabetes, or both conditions was 3137 overall (1021, 1101, and 1015 for usual care, digital-only, and multicomponent arms, respectively).

[e]The number of participants with hypertension, diabetes, or both conditions who ever smoked was 957 overall (328, 309, and 320 for usual care, digital-only, and multicomponent arms, respectively).

Discussion

Driven by population aging, rapid urbanization, and the globalization of unhealthy lifestyles, the burden of NCDs is rapidly increasing in LMICs [2,74,75]. However, access to quality care remains limited in LMICs, particularly in rural areas. Awareness, diagnosis, treatment, and control rates of hypertension and diabetes are low. Strengthening primary care to address NCDs requires novel approaches to address complex barriers at the health system, provider, and patient levels [12]. Several features of NCDs, including shared lifestyle risk factors, comorbidity, and chronicity, necessitate a shifting of care organization from episodic care toward a long-term integrated approach to prevention, diagnosis, and treatment across overlapping conditions [12]. A well-coordinated team-based model of care is critical for the care continuity necessary to achieve sustained control of chronic conditions [29,76]. Digital technologies offer the potential to tackle health system challenges in LMICs, improve access to and quality of health care, and reduce health system costs [77]. The emergence of the COVID-19 pandemic has intensified the need for digital technologies to support decentralized NCD care [78,79]. However, few digital interventions have demonstrated effectiveness in NCD care [28,30,80]. Progress toward successful implementation and scaling up of digital innovations for primary care delivery in LMIC settings remains limited and requires collaborative research and development efforts between health system stakeholders and technology communities to address holistically.

This study is among the first to evaluate the effectiveness, cost-effectiveness, and implementation strategy of a digital technology-supported decentralized primary care model for integrated hypertension and diabetes management in an LMIC context [24,25]. With a combination of repeated cross-sectional community surveys, routinely collected facility-based longitudinal data, qualitative data collection, and a cost-effectiveness evaluation, this study is expected to provide rich implementation and effectiveness data on a multicomponent digital technology-supported decentralized intervention for integrated hypertension and diabetes care. The findings from the study will provide valuable insights for the Simple app developers and the larger global health technology community,

XSL•FO
RenderX

supporting their efforts to develop effective digital solutions to address NCD challenges.

The baseline community survey has been completed with a high response rate and excellent data completion rates. At the time of writing this paper, initial training of health care providers and managers has been completed, the Simple app has been successfully deployed, and CCs have been equipped to provide hypertension and diabetes screening, referral, and management. Facility-based data collection has been set up. Data analysis on the baseline community survey and facility-based survey is ongoing. Stakeholders have been actively engaged, and initial findings have been shared with relevant government agencies.

Designed to generate rich process and implementation data, the study is limited in the number of primary care facilities involved and geographical representativeness. Thus, the results will need to be interpreted in relation to the specific regional socioeconomic context. Nevertheless, our implementation assessment results will shed light on the transportability of the interventions in different contexts, which might include density of primary facilities, density and information technology knowledge of health care providers, availability of CHWs, internet connection, etc. Furthermore, unlike a randomized controlled trial, the quasi-experimental approach relies on a strong "common trends" assumption to establish a counterfactual. Without randomization, causal inferences are difficult to establish. Nevertheless, we have collected extensive data on potential confounding variables and planned extensive sensitivity analyses to evaluate potential biases as previously discussed.

As the government of Bangladesh is taking more steps to mitigate the increasing burden of NCDs and to achieve the Sustainable Development Goals, the evidence generated from the proposed study will be directly relevant for policymaking and programmatic efforts for NCD prevention and management in Bangladesh. The implementation and cost-effectiveness data may be particularly important to inform the scalability and sustainability of the interventions. The findings will likely be internationally relevant, as many LMICs share similar challenges as Bangladesh regarding NCD prevention and control.

Funding

This research was funded by the UK National Institute for Health Research (NIHR) (16/136/68 and 132960) using UK aid from the UK Government to support global health research. The views expressed in this publication are those of the authors and not necessarily those of the NIHR or the UK Department of Health and Social Care.

Authors' Contributions

JC, MKM, and BO conceived the study. JC, MKM, BO, and WX designed the initial study protocol. WX wrote the initial draft of the manuscript, and all the authors were involved in preparing this manuscript and contributed to the revision of the manuscript. JC acquired funding. MKM and JC supervised the implementation of the study. JC is the co-corresponding author and can be reached via email at john.chambers@imperial.ac.uk.

Conflicts of Interest

None declared.

References

1. Sun H, Saeedi P, Karuranga S, et al. IDF Diabetes Atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract 2022 Jan;183:109119. [doi: 10.1016/j.diabres.2021.109119] [Medline: 34879977]
2. Roth GA, Mensah GA, Johnson CO, et al. Global burden of cardiovascular diseases and risk factors, 1990-2019. J Am Coll Cardiol 2020 Dec;76(25):2982-3021. [doi: 10.1016/j.jacc.2020.11.010]
3. Zhou B, Bentham J, Di Cesare M, et al. Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19·1 million participants. Lancet 2017 Jan;389(10064):37-55. [doi: 10.1016/S0140-6736(16)31919-5]
4. Zhou B, Carrillo-Larco RM, Danaei G, et al. Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. Lancet 2021 Sep;398(10304):957-980. [doi: 10.1016/S0140-6736(21)01330-1]
5. Gu Q, Dillon CF, Burt VL, Gillum RF. Association of hypertension treatment and control with all-cause and cardiovascular disease mortality among US adults with hypertension. Am J Hypertens 2010 Jan;23(1):38-45. [doi: 10.1038/ajh.2009.191] [Medline: 19851295]
6. Zhou D, Xi B, Zhao M, Wang L, Veeranki SP. Uncontrolled hypertension increases risk of all-cause and cardiovascular disease mortality in US adults: the NHANES III Linked Mortality Study. Sci Rep 2018 Jun 20;8(1):9418. [doi: 10.1038/s41598-018-27377-2] [Medline: 29925884]
7. Manne-Goehler J, Geldsetzer P, Agoudavi K, et al. Health system performance for people with diabetes in 28 low- and middle-income countries: a cross-sectional study of nationally representative surveys. PLoS Med 2019 Mar;16(3):e1002751. [doi: 10.1371/journal.pmed.1002751] [Medline: 30822339]

XSL•FO
RenderX

8. Geldsetzer P, Manne-Goehler J, Marcus ME, et al. The state of hypertension care in 44 low-income and middle-income countries: a cross-sectional study of nationally representative individual-level data from 1·1 million adults. Lancet 2019 Aug;394(10199):652-662. [doi: 10.1016/S0140-6736(19)30955-9]
9. Zhou B, Perel P, Mensah GA, Ezzati M. Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. Nat Rev Cardiol 2021 Nov;18(11):785-802. [doi: 10.1038/s41569-021-00559-8] [Medline: 34050340]
10. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. JAMA 2002 Oct 16;288(15):1909-1914. [doi: 10.1001/jama.288.15.1909] [Medline: 12377092]
11. Kabir A, Karim MN, Islam RM, Romero L, Billah B. Health system readiness for non-communicable diseases at the primary care level: a systematic review. BMJ Open 2022 Feb;12(2):e060387. [doi: 10.1136/bmjopen-2021-060387]
12. Kruk ME, Nigenda G, Knaul FM. Redesigning primary care to tackle the global epidemic of noncommunicable disease. Am J Public Health 2015 Mar;105(3):431-437. [doi: 10.2105/AJPH.2014.302392]
13. Kruk ME, Gage AD, Arsenault C, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. Lancet Glob Health 2018 Nov;6(11):e1196-e1252. [doi: 10.1016/S2214-109X(18)30386-3]
14. Global report on hypertension: the race against a silent killer. : World Health Organization; 2023 Sep 19 URL: https://www.who.int/publications/i/item/9789240081062 [accessed 2025-12-12]
15. Chowdhury HA, Paromita P, Mayaboti CA, et al. Assessing service availability and readiness of healthcare facilities to manage diabetes mellitus in Bangladesh: findings from a nationwide survey. PLoS One 2022;17(2):e0263259. [doi: 10.1371/journal.pone.0263259] [Medline: 35171912]
16. Islam K, Huque R, Saif-Ur-Rahman KM, Ehtesham Kabir ANM, Enayet Hussain AHM. Implementation status of non-communicable disease control program at primary health care level in Bangladesh: findings from a qualitative research. Public Health Pract (Oxf) 2022 Jun;3:100271. [doi: 10.1016/j.puhip.2022.100271] [Medline: 36101774]
17. Correia JC, Lachat S, Lagger G, et al. Interventions targeting hypertension and diabetes mellitus at community and primary healthcare level in low- and middle-income countries: a scoping review. BMC Public Health 2019 Nov 21;19(1):1542. [doi: 10.1186/s12889-019-7842-6] [Medline: 31752801]
18. Ye J, Orji I, Baldridge AS, et al. Characteristics and patterns of retention in hypertension care in primary care settings from the Hypertension Treatment in Nigeria Program. Circulation 2022 Mar;145(Suppl_1):e2230025. [doi: 10.1001/jamanetworkopen.2022.30025]
19. Labhardt ND, Balo JR, Ndam M, Grimm JJ, Manga E. Task shifting to non-physician clinicians for integrated management of hypertension and diabetes in rural Cameroon: a programme assessment at two years. BMC Health Serv Res 2010 Dec 14;10(1):1-10. [doi: 10.1186/1472-6963-10-339] [Medline: 21144064]
20. Deo S, Singh P. Community health worker-led, technology-enabled private sector intervention for diabetes and hypertension management among urban poor: a retrospective cohort study from large Indian metropolitan city. BMJ Open 2021 Aug 12;11(8):e045246. [doi: 10.1136/bmjopen-2020-045246] [Medline: 34385229]
21. Brennan AT, Long L, Maskew M, et al. Outcomes of stable HIV-positive patients down-referred from a doctor-managed antiretroviral therapy clinic to a nurse-managed primary health clinic for monitoring and treatment. AIDS 2011 Oct 23;25(16):2027-2036. [doi: 10.1097/QAD.0b013e32834b6480] [Medline: 21997488]
22. Sanne I, Orrell C, Fox MP, et al. Nurse versus doctor management of HIV-infected patients receiving antiretroviral therapy (CIPRA-SA): a randomised non-inferiority trial. Lancet 2010 Jul;376(9734):33-40. [doi: 10.1016/S0140-6736(10)60894-X]
23. Lehmann U, Van Damme W, Barten F, Sanders D. Task shifting: the answer to the human resources crisis in Africa? Hum Resour Health 2009 Jun 21;7(1):1-4. [doi: 10.1186/1478-4491-7-49] [Medline: 19545398]
24. Kaur P, Kunwar A, Sharma M, et al. The India Hypertension Control Initiative—early outcomes in 26 districts across five states of India, 2018-2020. J Hum Hypertens 2023 Jul;37(7):560-567. [doi: 10.1038/s41371-022-00742-5] [Medline: 35945426]
25. Patel SA, Sharma H, Mohan S, et al. The Integrated Tracking, Referral, and Electronic Decision Support, and Care Coordination (I-TREC) program: scalable strategies for the management of hypertension and diabetes within the government healthcare system of India. BMC Health Serv Res 2020 Nov 9;20(1):1022. [doi: 10.1186/s12913-020-05851-w] [Medline: 33168004]
26. Xie W, Paul RR, Goon IY, et al. Enhancing care quality and accessibility through digital technology-supported decentralisation of hypertension and diabetes management: a proof-of-concept study in rural Bangladesh. BMJ Open 2023 Nov;13(11):e073743. [doi: 10.1136/bmjopen-2023-073743]
27. Anchala R, Kaptoge S, Pant H, Di Angelantonio E, Franco OH, Prabhakaran D. Evaluation of effectiveness and cost-effectiveness of a clinical decision support system in managing hypertension in resource constrained primary health care settings: results from a cluster randomized trial. J Am Heart Assoc 2015 Jan 5;4(1):e001213. [doi: 10.1161/JAHA.114.001213] [Medline: 25559011]
28. Prabhakaran D, Jha D, Prieto-Merino D, et al. Effectiveness of an mHealth-based electronic decision support system for integrated management of chronic conditions in primary care. Circulation 2019 Jan 15;139(3):380-391. [doi: 10.1161/CIRCULATIONAHA.118.038192]

XSL•FO
RenderX

29. Kumar A, Schwarz D, Acharya B, et al. Designing and implementing an integrated non-communicable disease primary care intervention in rural Nepal. BMJ Glob Health 2019;4(2):e001343. [doi: 10.1136/bmjgh-2018-001343] [Medline: 31139453]
30. Xiong S, Lu H, Peoples N, et al. Digital health interventions for non-communicable disease management in primary health care in low- and middle-income countries. NPJ Digit Med 2023 Feb 1;6(1):12. [doi: 10.1038/s41746-023-00764-4] [Medline: 36725977]
31. He J, Irazola V, Mills KT, et al. Effect of a community health worker-led multicomponent intervention on blood pressure control in low-income patients in Argentina: a randomized clinical trial. JAMA 2017 Sep 19;318(11):1016-1025. [doi: 10.1001/jama.2017.11358] [Medline: 28975305]
32. Jafar TH, Gandhi M, de Silva HA, et al. A community-based intervention for managing hypertension in rural South Asia. N Engl J Med 2020 Feb 20;382(8):717-726. [doi: 10.1056/NEJMoa1911965]
33. Schwalm JD, McCready T, Lopez-Jaramillo P, et al. A community-based comprehensive intervention to reduce cardiovascular risk in hypertension (HOPE 4): a cluster-randomised controlled trial. Lancet 2019 Oct;394(10205):1231-1242. [doi: 10.1016/S0140-6736(19)31949-X]
34. Joshi R, Alim M, Kengne AP, et al. Task shifting for non-communicable disease management in low and middle income countries – a systematic review. PLoS One 2014;9(8):e103754. [doi: 10.1371/journal.pone.0103754]
35. Jeet G, Thakur JS, Prinja S, Singh M. Community health workers for non-communicable diseases prevention and control in developing countries: evidence and implications. PLoS One 2017;12(7):e0180640. [doi: 10.1371/journal.pone.0180640]
36. Simple. URL: https://www.simple.org/ [accessed 2025-12-15]
37. Burka D, Gupta R, Moran AE, et al. Keep it simple: designing a user-centred digital information system to support chronic disease management in low/middle-income countries. BMJ Health Care Inform 2023 Jan;30(1):e100641. [doi: 10.1136/bmjhci-2022-100641] [Medline: 36639189]
38. Ganeshkumar P, Bhatnagar A, Burka D, et al. Discovery, development, and deployment of a user-centered point-of-care digital information system to treat and track hypertension and diabetes patients under India Hypertension Control Initiative 2019-2022, India. Digit Health 2024;10. [doi: 10.1177/20552076241250153] [Medline: 38715975]
39. Burnier M, Egan BM. Adherence in hypertension: a review of prevalence, risk factors, impact, and management. Circ Res 2019;124(7):1124-1140. [doi: 10.1161/CIRCRESAHA.118.313220]
40. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct;396(10258):1204-1222. [doi: 10.1016/S0140-6736(20)30925-9]
41. Khan M, Oldroyd JC, Chowdhury EK, et al. Prevalence, awareness, treatment, and control of hypertension in Bangladesh: findings from National Demographic and Health Survey, 2017-2018. J Clin Hypertens 2021 Oct;23(10):1830-1842. [doi: 10.1111/jch.14363]
42. Khan M, Oldroyd JC, Hossain MB, Islam RM. Awareness, treatment, and control of diabetes in Bangladesh: evidence from the Bangladesh Demographic and Health Survey 2017/18. . 2022 p. 8349160. [doi: 10.1155/2022/8349160]
43. Riaz BK, Islam MZ, Islam A, et al. Risk factors for non-communicable diseases in Bangladesh: findings of the population-based cross-sectional national survey 2018. BMJ Open 2020 Nov 27;10(11):e041334. [doi: 10.1136/bmjopen-2020-041334] [Medline: 33247026]
44. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226. [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
45. Population and Housing Census [Web page in Bangla]. Bangladesh Bureau of Statistics. 2011. URL: https://bbs.gov.bd/pages/static-pages/6922e0ff933eb65569e297dc [accessed 2023-11-10]
46. Facility registry. Directorate General of Health Services. URL: https://hrm.dghs.gov.bd/public/facility-registry [accessed 2023-11-10]
47. Schwalm JD, McKee M, Huffman MD, Yusuf S. Resource effective strategies to prevent and treat cardiovascular disease. Circulation 2016 Feb 23;133(8):742-755. [doi: 10.1161/CIRCULATIONAHA.115.008721] [Medline: 26903017]
48. Naheed A, Haldane V, Jafar TH, Chakma N, Legido-Quigley H. Patient pathways and perceptions of hypertension treatment, management, and control in rural Bangladesh: a qualitative study. Patient Prefer Adherence 2018;12:1437-1449. [doi: 10.2147/PPA.S163385] [Medline: 30147302]
49. Legido-Quigley H, Naheed A, de Silva HA, et al. Patients' experiences on accessing health care services for management of hypertension in rural Bangladesh, Pakistan and Sri Lanka: a qualitative study. PLoS One 2019;14(1):e0211100. [doi: 10.1371/journal.pone.0211100]
50. Nam S, Chesla C, Stotts NA, Kroon L, Janson SL. Barriers to diabetes management: patient and provider factors. Diabetes Res Clin Pract 2011 Jul;93(1):1-9. [doi: 10.1016/j.diabres.2011.02.002] [Medline: 21382643]
51. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]

XSL•FO
RenderX

52. Abrahams-Gessel S, Denman CA, Montano CM, et al. The training and fieldwork experiences of community health workers conducting population-based, noninvasive screening for CVD in LMIC. Glob Heart 2015;10(1):45-54. [doi: 10.1016/j.gheart.2014.12.008]
53. Alaofè H, Asaolu I, Ehiri J, et al. Community health workers in diabetes prevention and management in developing countries. Ann Glob Health 2017;83(3-4):661-675. [doi: 10.1016/j.aogh.2017.10.009] [Medline: 29221543]
54. Anand TN, Joseph LM, Geetha AV, Prabhakaran D, Jeemon P. Task sharing with non-physician health-care workers for management of blood pressure in low-income and middle-income countries: a systematic review and meta-analysis. Lancet Glob Health 2019 Jun;7(6):e761-e771. [doi: 10.1016/S2214-109X(19)30077-4]
55. Schwarz D, Dhungana S, Kumar A, et al. An integrated intervention for chronic care management in rural Nepal: protocol of a type 2 hybrid effectiveness-implementation study. Trials 2020 Jan 29;21(1):119. [doi: 10.1186/s13063-020-4063-3] [Medline: 31996250]
56. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ 2015 Mar 19;350:h1258. [doi: 10.1136/bmj.h1258] [Medline: 25791983]
57. Noncommunicable disease facility-based monitoring guidance: framework, indicators, and application. : World Health Organization; 2022 Nov 11 URL: https://www.who.int/publications/i/item/9789240057067 [accessed 2025-12-12]
58. Xia T, Zhao F, Nianogo RA. Interventions in hypertension: systematic review and meta-analysis of natural and quasi-experiments. Clin Hypertens 2022 May 1;28(1):13. [doi: 10.1186/s40885-022-00198-2] [Medline: 35490246]
59. Miller CJ, Smith SN, Pugatch M. Experimental and quasi-experimental designs in implementation research. Psychiatry Res 2020 Jan;283:112452. [doi: 10.1016/j.psychres.2019.06.027] [Medline: 31255320]
60. Song P, Gupta A, Goon IY, et al. Data resource profile: understanding the patterns and determinants of health in South Asians—the South Asia Biobank. Int J Epidemiol 2021 Jul 9;50(3):717-718e. [doi: 10.1093/ije/dyab029] [Medline: 34143882]
61. Gandhi M, Assam PN, Turner EL, et al. Statistical analysis plan for the control of blood pressure and risk attenuation-rural Bangladesh, Pakistan, Sri Lanka (COBRA-BPS) trial: a cluster randomized trial for a multicomponent intervention versus usual care in hypertensive patients. Trials 2018 Nov 29;19(1):658. [doi: 10.1186/s13063-018-3022-8] [Medline: 30486858]
62. Ahn C, Heo M, Zhang S. Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research: CRC Press; 2014.
63. Campbell MJ, Julious SA, Altman DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. BMJ 1995 Oct 28;311(7013):1145-1148. [doi: 10.1136/bmj.311.7013.1145] [Medline: 7580713]
64. Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies: World Health Organization; 2010.
65. Xu X, Lazar CM, Ruger JP. Micro-costing in health and medicine: a critical appraisal. Health Econ Rev 2021 Jan 6;11:1. [doi: 10.1186/s13561-020-00298-5] [Medline: 33404857]
66. White H. Maximum likelihood estimation of misspecified models. Econometrica 1982 Jan;50(1):1. [doi: 10.2307/1912526]
67. Bertrand M, Duflo E, Mullainathan S. How much should we trust differences-in-differences estimates? Q J Econ 2004 Feb 1;119(1):249-275. [doi: 10.1162/003355304772839588]
68. Abadie A, Athey S, Imbens GW, Wooldridge J. When should you adjust standard errors for clustering? : National Bureau of Economic Research; 2017 Nov. [doi: 10.3386/w24003]
69. Zou G, Donner A. Extension of the modified Poisson regression model to prospective studies with correlated binary data. Stat Methods Med Res 2013 Dec;22(6):661-670. [doi: 10.1177/0962280211427759]
70. Yelland LN, Salter AB, Ryan P. Performance of the modified Poisson regression approach for estimating relative risks from clustered prospective data. Am J Epidemiol 2011 Oct 15;174(8):984-992. [doi: 10.1093/aje/kwr183] [Medline: 21841157]
71. Strauss A, Corbin J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory, 2nd edition: Sage Publications; 1998.
72. Ryan GW, Bernard HR. Techniques to identify themes. Field Methods 2003 Feb;15(1):85-109. [doi: 10.1177/1525822X02239569]
73. Sanders GD, Neumann PJ, Basu A, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. JAMA 2016 Sep 13;316(10):1093-1103. [doi: 10.1001/jama.2016.12195] [Medline: 27623463]
74. Noncommunicable diseases. World Health Organization. 2022 Sep 25. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2025-12-12]
75. Bloom DE, Cafiero E, Jané-Llopis E, Abrahams-Gessel S, Bloom LR, Fathima S, et al. The global economic burden of noncommunicable diseases. : World Economic Forum; 2011 Sep URL: https://www3.weforum.org/docs/WEF_Harvard_HE_GlobalEconomicBurdenNonCommunicableDiseases_2011.pdf [accessed 2025-12-12]
76. Starfield B. Is primary care essential? Lancet 1994 Oct;344(8930):1129-1133. [doi: 10.1016/S0140-6736(94)90634-3]
77. Labrique AB, Wadhwani C, Williams KA, et al. Best practices in scaling digital health in low and middle income countries. Global Health 2018 Dec;14(1):1-8. [doi: 10.1186/s12992-018-0424-z]

XSL•FO
RenderX

78. Kluge HHP, Wickramasinghe K, Rippin HL, et al. Prevention and control of non-communicable diseases in the COVID-19 response. Lancet 2020 May;395(10238):1678-1680. [doi: 10.1016/S0140-6736(20)31067-9]
79. Reddy SK, Kunwar A, Durgad K, et al. Decentralization of India Hypertension Control Initiative services to maintain continuum of care for hypertensive patients during COVID-19 pandemic in Telangana. WHO South East Asia J Public Health 2021;10(Suppl 1):S49-S58. [doi: 10.4103/2224-3151.309873]
80. Peiris D, Praveen D, Mogulluru K, et al. SMARThealth India: a stepped-wedge, cluster randomised controlled trial of a community health worker managed mobile health intervention for people assessed at high cardiovascular disease risk in rural India. PLoS One 2019;14(3):e0213708. [doi: 10.1371/journal.pone.0213708]

Abbreviations

BG: blood glucose
BP: blood pressure
CC: community clinic
CHCP: community health care provider
CHW: community health worker
CVD: cardiovascular disease
FGD: focus group discussion
ICER: incremental cost-effectiveness ratio
IDI: in-depth interview
LMIC: low- and middle-income country
MRC: UK Medical Research Council
NCD: noncommunicable disease
RE-AIM: reach, effectiveness, adoption, implementation, and maintenance framework
WHO: World Health Organization
WHO PEN: WHO Package of Essential Noncommunicable Diseases Interventions
WHO STEPS: World Health Organization's STEPwise approach to NCD risk factor surveillance

Edited by J Sarvestan; submitted 17.Jul.2025; peer-reviewed by N Mor, Y Shen; revised version received 03.Nov.2025; accepted 19.Nov.2025; published 02.Feb.2026.

Please cite as:
Xie W, Ahmed S, Ahsan A, Gupta A, Keenan AM, Sarker T, Akter F, Anan A, Hossain MM, Quayyum Z, Hussain AHME, Amin R, Chowdhury IA, Faruque M, Choudhury SR, Goon IY, Hersch F, Sabin LL, Oldenburg B, Chambers J, Mridha MK
Addressing Gaps in the Hypertension and Diabetes Care Continuum in Rural Bangladesh Through Digital Technology Supported Decentralized Primary Care: Study Protocol and Baseline Results for a Hybrid Effectiveness-Implementation Trial
JMIR Res Protoc 2026;15:e71696
URL: https://www.researchprotocols.org/2026/1/e71696
doi:10.2196/71696

© Wubin Xie, Sabrina Ahmed, Ali Ahsan, Ananya Gupta, Anais Masako Keenan, Tanmoy Sarker, Fahmida Akter, Aysha Anan, Md Mokbul Hossain, Zahidul Quayyum, AHM Enayet Hussain, Robed Amin, Imran Ahmed Chowdhury, Mithila Faruque, Sohel Reza Choudhury, Ian Y Goon, Fred Hersch, Lora L Sabin, Brian Oldenburg, John Chambers, Malay Kanti Mridha. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Optimizing Antiemetic Support in Anthracycline-Based Chemotherapy for Early Breast Cancer: Protocol for a Prospective Observational Study of Four-Drug Antiemetic Therapy Including Fosnetupitant and Olanzapine

Ayako Higuchi[1], MSc; Yumiko Koi[2], MD, PhD; Tomoaki Eto[1,3], MSc; Yuka Maeda[1], BSc; Wakako Tajiri[2], MD; Junji Kawasaki[2], MD; Sayuri Akiyoshi[2], MD, PhD; Hideki Ijichi[2], MD, PhD; Yoshiaki Nakamura[2], MD, PhD; Chinami Koga[2], MD; Mototsugu Shimokawa[4], PhD; Hiroaki Shimizu[1,5], BSc; Toshihiro Matsumoto[1], BSc; Eriko Tokunaga[2], MD, PhD

[1]Department of Pharmacy, National Hospital Organization Kyushu Cancer Center, Fukuoka, Japan

[2]Department of Breast Oncology, National Hospital Organization Kyushu Cancer Center, 3-1-1 Notame Minami-ku, Fukuoka, Japan

[3]Department of Pharmacy, National Hospital Organization Kokura Medical Center, Kitakyushu, Japan

[4]Department of Biostatistics, Yamaguchi University, Yamaguchi, Japan

[5]Department of Pharmacy, National Sanatorium Kikuchi Keifuen, Koshi, Japan

Corresponding Author:
Eriko Tokunaga, MD, PhD
Department of Breast Oncology, National Hospital Organization Kyushu Cancer Center, 3-1-1 Notame Minami-ku, Fukuoka, Japan

Abstract

Background: Prophylaxis for chemotherapy-induced nausea and vomiting (CINV) in patients receiving highly emetogenic chemotherapy (HEC) is essential. Four-drug antiemetic therapy, consisting of a neurokinin-1 receptor antagonist (NK1RA), a 5-hydroxytryptamine type 3 receptor antagonist (5-HT3RA), dexamethasone (DEX), and olanzapine (OLZ), is currently recommended for HEC. However, the efficacy, optimal dosing schedule, and appropriate dosage of OLZ remain unclear when combined with a highly selective NK1RA, fosnetupitant (FosNTP).

Objective: This study aimed to evaluate the efficacy and safety of a four-drug antiemetic regimen including FosNTP and OLZ for prophylaxis of CINV in patients receiving anthracycline-based (neo)adjuvant chemotherapy for early breast cancer (EBC) and to explore outcomes with OLZ dose adjustments in the second cycle.

Methods: This single-institution, prospective observational study will enroll 100 patients with EBC who undergo HEC. All patients will receive a four-drug antiemetic regimen consisting of FosNTP, 5-HT3RA, DEX, and the guideline-recommended dose of OLZ administered orally from days –1 to 4. During the second cycle of treatment, OLZ dosing may be adjusted based on tolerability and patient preference. The primary endpoint is the proportion of patients who experienced no nausea during the overall phase (0 - 120 h) of the first cycle, as assessed using the daily visual analog scale. A sample size of 86 was calculated to assess the efficacy of the four-drug antiemetic regimen, including OLZ, assuming an expected no nausea rate of 42%, a threshold no nausea rate of 27% based on historical data, 90% power, and a 1-sided significance level of 5.0%. A single-sample z-test with a normal approximation will be used for the analyses.

Results: This study was approved by the Ethics Committee of the National Hospital Organization Kyushu Cancer Center, and participant recruitment and data collection commenced in May 2024. As of September 2025, approximately 70 participants have been recruited. Data collection is expected to continue until April 2026, and both data collection and analyses are anticipated to be completed in 2027.

Conclusions: This study will provide real-world evidence on the effectiveness and safety of a four-drug antiemetic regimen, including FosNTP and OLZ, in patients receiving anthracycline-based HEC and may inform optimal OLZ dosing strategies.

Trial Registration: UMIN-CTR UMIN000054250; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000061951

International Registered Report Identifier (IRRID): DERR1-10.2196/85648

(JMIR Res Protoc 2026;15:e85648) doi:10.2196/85648

XSL•FO
RenderX

KEYWORDS

antiemetics; olanzapine; Neurokinin-1 Receptor Antagonists; patient reported outcome measures; breast neoplasms

Introduction

Chemotherapy-induced nausea and vomiting (CINV) is one of the most common and distressing side effects of cancer treatment, significantly impairing patients' quality of life (QOL), especially due to nausea [1-3]. For highly emetogenic chemotherapy (HEC), international and Japanese guidelines currently recommend a four-drug antiemetic therapy consisting of a neurokinin-1 receptor antagonist (NK1RA), 5-hydroxytryptamine type 3 receptor antagonist (5-HT3RA), dexamethasone (DEX), and olanzapine (OLZ) [4-7]. OLZ is a multireceptor antipsychotic that exerts antiemetic effects partly via 5-HT2B/2C receptor antagonism and enhancement of ghrelin signaling [8]. However, owing to its broad receptor activity, OLZ frequently causes daily somnolence, particularly at the standard dose of 10 mg, with reported rates as high as 73.0%. Across guidelines and clinical studies, OLZ dosing for HEC prophylaxis generally ranges from 5 to 10 mg/day, reflecting a balance between efficacy and tolerability [5]. Based on previous studies demonstrating comparable efficacy and favorable tolerability of 5 mg OLZ relative to the standard 10 mg dose [9-12], the recommended initial dose in Japan is 5 mg/day [7]. Nevertheless, even at a dose of 5 mg, 20.3% of patients required dose reduction and 6.2% discontinued OLZ because of adverse effects [3,13]. In addition, a phase III randomized trial demonstrated that a 2.5 mg dose of OLZ was noninferior to 10 mg for CINV prevention, with significantly less somnolence [14].

Fosnetupitant (FosNTP) is a relatively new NK1RA that has demonstrated noninferiority to fosaprepitant in two phase III trials conducted in Japan (CONSOLE [15] and CONSOLE-BC [16]). In routine clinical practice, fosaprepitant has practical limitations, including admixture incompatibility with PALO and the risk of injection site reactions. FosNTP addresses several of these limitations and is therefore increasingly used in clinical practice [16]. However, evidence regarding the combined use of FosNTP with OLZ for CINV prophylaxis remains limited [17].

This study will evaluate the real-world efficacy and safety of a four-drug regimen (FosNTP, palonosetron [PALO], DEX, and OLZ) for CINV prophylaxis in patients receiving (neo)adjuvant anthracycline-based chemotherapy (HEC) for early breast cancer. The primary objective is to estimate the proportion of patients who experience no nausea during the overall phase (0 - 120 h) in the first chemotherapy cycle. A secondary objective is to explore clinical outcomes among patients who require OLZ dose adjustments after the first cycle.

Methods

Study Setting

This prospective, single-institution, observational cohort study will be conducted at the National Hospital Organization Kyushu Cancer Center, Japan. At our institution, FosNTP-based antiemetic prophylaxis is used in routine care for anthracycline-based HEC, and the use of alternative NK1RAs (eg, fosaprepitant) is limited due to practical concerns such as admixture incompatibility with PALO and injection site reactions. As a result, establishing a concurrent comparison group receiving another NK1RA is not feasible; therefore, we selected a prospective observational cohort design to evaluate the real-world efficacy and safety of the FosNTP-containing four-drug regimen and to explore OLZ dose adjustments. This study was designed as a prospective observational cohort study, and all antiemetic treatments, including OLZ dose adjustments, were provided as part of routine clinical practice at the discretion of the treating physicians; the study protocol did not assign or mandate any intervention. We aim to enroll 100 participants over 2 years from May 2024 to April 2026. The first participant was enrolled on May 7, 2024.

Study Participants

Inclusion criteria

- Patients eligible for this study must meet all the following criteria:

1. Female patients 20 - 74 years old who were able to understand and provide their written informed consent
2. A diagnosis of early-stage or locally advanced breast cancer (stages I–III) in a patient scheduled to receive (neo)adjuvant chemotherapy
3. No prior chemotherapy
4. A plan was made to receive either EC (epirubicin 90 mg/m^2 and cyclophosphamide 600 mg/m^2 every 3 weeks) or dose-dense EC (epirubicin 90 mg/m^2 and cyclophosphamide 600 mg/m^2 ever weeks)
5. An Eastern Cooperative Oncology Group performance status (ECOG PS) of 0 or 1
6. A preserved major organ function within 8 days prior to enrollment (defined as AST<100 IU/L, ALT<100 IU/L, and total bilirubin <2.0 mg/dL [34.2 µmol/L])
7. Able to understand and report patient-reported outcomes (PROs)
8. Clinical data available for analyses

Exclusion Criteria

- Patients excluded for this study must meet any the following criteria:

1. A history of allergy to any of the study drugs or related compounds
2. Vomiting or nausea within one day prior to enrollment
3. Use of antiemetic drugs within two days prior to enrollment
4. A history of, or scheduled to undergo surgery or radiation therapy within seven days prior to enrollment or during the observation period
5. Regular use of medications that may affect antiemetic efficacy, including selective serotonin reuptake inhibitors (SSRIs), serotonin-norepinephrine reuptake inhibitors (SNRIs), serotonin-dopamine antagonists,

XSL•FO
RenderX

multireceptor-targeted antipsychotics, corticosteroids, histamine H_1 receptor antagonists, phenothiazine antipsychotics, dopamine receptor antagonists, and benzodiazepines
6. Initiation of opioid analgesics within 14 days prior to enrollment
7. Seizure disorders requiring anticonvulsant treatment
8. Gastrointestinal obstruction, including pyloric stenosis or intestinal obstruction
9. Pregnant or lactating women or women of childbearing potential unwilling to use effective contraception
10. A history of diabetes requiring treatment with insulin or oral hypoglycemic agents
11. Concurrent participation in another interventional clinical trial
12. Any condition that, in the opinion of the attending physician, would render the patient unsuitable for participation in the study

Procedure and Treatment

The study consisted of a screening period prior to the initiation of EC, followed by a treatment period of 7 days (168 h), beginning at the start of EC administration. Written informed consent will be obtained prior to enrollment after confirming that the patients meet all the inclusion criteria and do not meet any exclusion criteria.

All patient information will be handled in a reidentifiable anonymized format, with the linkage key strictly controlled and never shared externally. Data will be securely managed within the institution and securely stored for at least five years after study completion or three years after publication, whichever is longer.

Treatment

All enrolled patients will receive PALO (day 1: 0.75 mg intravenously, 30 min before chemotherapy), FosNTP (day 1: 235 mg intravenously, 30 min before chemotherapy), DEX (day 1: 9.9 mg intravenously, 30 min before chemotherapy), and OLZ 5 mg. The initial dose of FosNTP, PALO, and DEX follows the approved dosing in Japan and the recommended antiemetic regimen for HEC in current guidelines [4-7]. OLZ was initiated at 5 mg/day as the recommended initial dose in Japan, considering the balance between antiemetic efficacy and tolerability [7,9-13]. In this study, OLZ will be administered orally and post-prandially from day –1 to day 4, rather than initiating administration on the evening of day 1 as is conventional [9]. This earlier initiation was adopted to ensure adequate plasma concentrations at the start of treatment based on its oral pharmacokinetics, with a time to maximum plasma concentration of approximately 6 hours and an elimination half-time of about 30 hours [18]. Given the higher incidence of acute-phase CINV on days 1‐2 reported for HEC [19] , initiating OLZ on the evening of day –1 was intended to provide sufficient and sustained drug exposure at the onset of treatment and throughout the early high-risk period for optimal antiemetic control.

For the second cycle, the OLZ dose may be adjusted based on patient preference after the physician reviews tolerability with attention to clinically relevant adverse events such as somnolence in the first cycle. Based on this assessment, patients may choose to continue with 5 mg/day, reduce the dose to 2.5 mg/day, increase it to 10 mg/day, or discontinue OLZ. The reason for the second cycle OLZ dose decision is documented. The selected OLZ dose will be confirmed directly with the patient before the start of the second cycle. Any OLZ dose reduction or discontinuation during a cycle will be determined by the physician in consultation with the patient, considering patient preference and clinical tolerability.

If nausea or vomiting occurs and the physician deems it necessary, or if the patient requests additional antiemetic therapy, rescue medications not included in the study protocol may be administered. These include metoclopramide, domperidone, alprazolam, and other drugs, as appropriate.

Follow-Up

During the seven-day treatment period, all patients will record symptoms in a diary, including nausea, vomiting, use of rescue medications (additional antiemetics), daytime somnolence, and its impact on daily activities. Nausea, somnolence, and the impact of somnolence on daily activities will be assessed using a visual analog scale (VAS). The impact of somnolence on daily activities is defined as interference with usual daily activities.

At the end of the observation period, patients will retrospectively assess the adverse events they experienced over a seven-day period using the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) [20]. They will also assess their satisfaction with antiemetic therapy using a seven-point Likert scale: very satisfied, satisfied, somewhat satisfied, slightly satisfied, slightly dissatisfied, dissatisfied, or very dissatisfied. Figure 1 illustrates the timeline for participant enrolment, intervention, and assessment.

XSL•FO
RenderX

Figure 1. The schedule of enrollment, intervention, and assessments. EC: epirubicin plus cyclophosphamide; ECOG PS: Eastern Cooperative Oncology Group performance status; OLZ: Olanzapine; PRO-CTCAE: Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events. X indicates the timing of planned procedures, interventions, or assessments. Repeated Xs or arrows across consecutive days indicate repeated daily administration/assessment during the indicated period.

	Trial period																
	Enrollment		Post-enrollment														Close-out
			Antiemetic therapies in cycle 1							Antiemetic therapies in cycle 2							
Timepoint	day -30 to -1	day -1	day 1	day 2	day 3	day 4	day 5	day 6	day 7	day -1	day 1	day 2	day 3	day 4	day 5	day 6	day 7
EC treatment			X								X						
Enrollment																	
Eligibility screen	X																
History and Physical Examination	X																
ECOG PS	X																
Laboratory studies	X									X							
Informed consent	X																
Enrollment		X															
Intervention																	
Palonosetron			X								X						
Dexamethasone			X								X						
Fosnetupitant			X								X						
Olanzapine		X	→	→	→	X				X	→	→	→	X			
Assessments																	
Patient diaries		X	X	X	X	X	X	X	X		X	X	X	X	X	X	X
OLZ dose selection										X							
PRO-CTCAE									X								X
Patient satisfaction									X								X

Endpoints

- Primary endpoint: The primary endpoint is no nausea during the overall phase (0 - 120 h after the start of epirubicin administration) in the first cycle. This binary endpoint was selected to reflect the clinically meaningful goal of nausea control, consistent with endpoint definitions commonly used in clinical trials evaluating antiemetic efficacy in the CINV/HEC setting [5,21]. Nausea is assessed using a 100-mm VAS, and no nausea is operationally defined as a VAS score of <5 mm, as described in the Definitions of key endpoints section. Continuous VAS scores will also be summarized as supportive information.
- Secondary endpoints: The secondary endpoints include no nausea during phases other than the overall phase of the first cycle, complete response (CR) rate, complete control (CC) rate, total control (TC) rate, time to treatment failure, degree of daytime somnolence, impact of somnolence on daily activities, no nausea during the second cycle, rate of OLZ dose modification, incidence of adverse events, and patient satisfaction with antiemetic therapy.

Observation Periods

The observation periods will be classified into five phases: acute (0 - 24 h after the start of epirubicin administration), delayed (24 - 120 h), overall (0 - 120 h), extended delayed (24 - 168 h), and extended overall (0 - 168 h).

Definitions of Key Endpoints

In this study, no nausea is defined as a VAS score of <5 mm. VAS is widely used to assess nausea and allows a more granular evaluation of nausea severity and changes over time compared with categorical rating scales [22]. Although no universally accepted VAS cutoff for nausea has been established, this threshold was selected based on previously published CINV studies involving anthracycline-treated breast cancer patients [23-25]. The CR rate is defined as not an emetic event, and no rescue medication will be administered. The CC rate showed no emetic event, no rescue medication, and no significant nausea (VAS score <25 mm). The TC rate is defined as no emetic events, rescue medication, or nausea. Time-to-treatment failure was defined as the time from the start of epirubicin administration to the earliest occurrence of either the first episode of vomiting or the first use of rescue medication.

Patient and Public Involvement

The patients and public were not involved in the development of the research question, study design, outcome measures, or recruitment plans.

Statistical Considerations

Sample Size Justification

Previous studies have reported varying rates of no nausea during the overall phase depending on the antiemetic therapy used. A 3-drug combination of FosNTP, PALO, and DEX reportedly achieved a no nausea rate of 27.5% in patients receiving HEC [26], and this value was used as the historical control threshold in the current study. In addition, a phase 3 randomized controlled trial by Navari et al showed that adding 10 mg of OLZ to a 3-drug combination of aprepitant, PALO (5-HT3RA), and DEX improved the rate of nausea by 15.4% compared to the antiemetic regimen without OLZ [27]. Based on this evidence, a threshold no-nausea rate of 27% was defined for this study, especially in the context of using a lower and potentially more tolerable dose of OLZ (5 mg). The required sample size was calculated using an expected a no-nausea rate of 42%, a minimum acceptance rate of 27%, a one-sided significance level of 5.0%, and a power of 90%. A one-sided significance level was selected because the primary objective is to demonstrate that the no nausea rate exceeds a prespecified historical threshold, and only an improvement over the threshold is clinically meaningful. Based on these assumptions, the required number of participants was 86. The target sample size was set at 100 patients, considering the anticipated consent and valid response rates.

XSL•FO
RenderX

Statistical Analyses

The primary analysis will be conducted on the full analysis set, defined as patients who have undergone at least one efficacy assessment. Patients who were retrospectively deemed ineligible or registered in violation of the protocol will be excluded from this set in accordance with the predefined criteria. The primary endpoint is a proportion, and a two-sided 90% confidence interval will be estimated using a normal approximation. A z-test based on the normal approximation will be performed. If the result is statistically significant, the primary objective will be considered met. We will conduct prespecified exploratory stratified analyses by OLZ dose modification in the second cycle and established CINV risk factors. These analyses will be descriptive and hypothesis-generating. Missing diary entries and incomplete PRO-CTCAE responses will not be imputed. Analyses will be conducted using available data, and the extent of missingness will be reported. For the primary endpoint, participants without sufficient diary or VAS data to determine "no nausea" during the overall phase will be classified as not having achieved no nausea. All statistical analyses will be performed using R version 4.4.2 or later.

Ethical Considerations

This study was approved by the Ethics Committee of the National Hospital Organization Kyushu Cancer Center (approval number: [2023 - 41]). Written informed consent will be obtained from all the participants prior to enrollment, after the treating physician has explained the study and a study pharmacist has provided additional explanations. The study will be conducted in accordance with the Declaration of Helsinki and applicable national regulations. To protect participants' privacy and confidentiality, all study data will be deidentified before analysis. Personal identifiers will be removed and replaced with unique study IDs. The linkage file will be stored separately on a secure, access-restricted system and will be accessible only to authorized study personnel. Where feasible, data will be reported in aggregate to minimize reidentification risk. No compensation will be provided to participants. The results will be disseminated through presentations at academic conferences and publications in peer-reviewed journals. No personally identifiable information is disclosed.

This study has been registered with the University Hospital Medical Information Network Clinical Trials Registry (UMIN-CTR) under the identifier UMIN000054250.

No ancillary or post-trial care is required, as the study involves routine chemotherapy and antiemetic treatment without additional risks beyond usual care.

Results

This study was approved by the Ethics Committee in March 2024, and participant recruitment and data collection commenced in May 2024. As of September 2025, approximately 70 participants had been recruited. Data collection is expected to continue until April 2026, and both data collection and analyses are anticipated to be completed in 2027.

Discussion

This study is expected to provide real-world evidence that a four-drug antiemetic regimen including FosNTP and OLZ offers clinically meaningful primary prophylaxis of CINV with acceptable tolerability in patients receiving perioperative epirubicin–cyclophosphamide chemotherapy for breast cancer. It also aims to inform the feasibility of OLZ dose reduction or discontinuation in subsequent cycles while maintaining CINV control. This protocol incorporates PROs as a direct measure of symptom burden. Despite adherence to current guideline-recommended regimens, approximately 70% of patients with breast cancer continue to experience insufficient control of nausea [5,28], underscoring a substantial unmet need in supportive care. Effective CINV prevention is essential for maintaining treatment adherence and enabling successful completion of (neo)adjuvant chemotherapy. Among patients receiving HEC, anthracycline-based chemotherapy is associated with an increased incidence of acute-phase CINV on days 1 and 2 when OLZ is not included in antiemetic therapy, highlighting the importance of earlier intervention [19]. In the present study, OLZ was administered from day –1 to day 4, a novel schedule that differs from the conventional initiation on the evening of day 1 and, to our knowledge, has not been previously evaluated in this population.

By assessing the real-world effectiveness of a four-drug antiemetic therapy incorporating FosNTP and OLZ, this study will generate clinically relevant data beyond those available from conventional approaches. Prior phase III trials conducted in Japan established the efficacy and safety of FosNTP-based prophylaxis in combination with PALO and DEX [15,16]; however, evidence on FosNTP in combination with OLZ remains limited. Furthermore, evaluating the tolerability of guideline-recommended OLZ dosing and exploring dose adjustments in subsequent cycles may inform individualized antiemetic strategies that balance efficacy and safety.

A key strength of this study is the prospective collection of PROs, enabling a detailed assessment of nausea control and tolerability in routine practice.

This study has several limitations. As a single-arm observational study, it is inherently limited in its ability to support causal inference, and residual confounding cannot be excluded. In particular, patient characteristics and clinical decision-making may have influenced both OLZ dose adjustments and outcomes. Accordingly, the results should be interpreted descriptively as real-world associations rather than as evidence of causal effects.

These findings are expected to refine antiemetic practice by providing real-world evidence on optimized four-drug therapy, improving supportive care standards, and enhancing the quality of life and treatment adherence for patients with early breast cancer receiving anthracycline-based chemotherapy. Future studies in broader settings will be valuable to confirm generalizability and further refine OLZ dosing strategies.

XSL•FO
RenderX

Acknowledgments

Artificial intelligence (AI)-based language editing tools (ChatGPT, OpenAI) was used to enhance linguistic quality. No AI tool was used in the generation of scientific content, data analysis, or interpretation.

Funding

This study received no specific grants from any funding agency in the public, commercial, or not-for-profit sector.

Data Availability

No data are publicly available for this study protocol. Patient-reported outcome data constitute original source documents completed directly by participants and cannot be shared, even in deidentified form, due to ethical and privacy considerations related to participant consent. No data monitoring committee is planned because this is a noninterventional observational study with minimal risk; study oversight will be provided by the principal investigator and the institutional ethics committee in accordance with institutional policies.

Authors' Contributions

All the authors contributed to the conception and design of the study protocol. The protocol was initially drafted by AH, YK, TE and ET. MS provided statistical expertise. AH and YK wrote the manuscript. YM, WT, JK, SA, HI, YN, CK, MS, HS and ET critically reviewed the manuscript. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT 2025 checklist.
[PDF File, 568 KB - resprot_v15i1e85648_app1.pdf]

References

1. Glaus A, Knipping C, Morant R, et al. Chemotherapy-induced nausea and vomiting in routine practice: a European perspective. Support Care Cancer 2004 Oct;12(10):708-715. [doi: 10.1007/s00520-004-0662-x] [Medline: 15278682]
2. Yeo W, Mo FKF, Yip CCH, et al. Quality of life associated with nausea and vomiting from anthracycline-based chemotherapy: a pooled data analysis from three prospective trials. Oncologist 2021 Dec;26(12):e2288-e2296. [doi: 10.1002/onco.13978] [Medline: 34516038]
3. Clemons M, Dranitsaris G, Sienkiewicz M, et al. A randomized trial of individualized versus standard of care antiemetic therapy for breast cancer patients at high risk for chemotherapy-induced nausea and vomiting. The Breast 2020 Dec;54:278-285. [doi: 10.1016/j.breast.2020.11.002]
4. NCCN clinical practice guidelines in oncology (NCCN guidelines): Antiemetics. version 2.2025. National Comprehensive Cancer Network. URL: https://www.nccn.org/guidelines/guidelines-detail?category= 3&id=1415 [accessed 2026-01-20]
5. Herrstedt J, Clark-Snow R, Ruhlmann CH, et al. MASCC and ESMO guideline update for the prevention of chemotherapy- and radiotherapy-induced nausea and vomiting. ESMO Open 2024 Feb;9(2):102195. [doi: 10.1016/j.esmoop.2023.102195] [Medline: 38458657]
6. Hesketh PJ, Kris MG, Basch E, et al. Antiemetics: ASCO Guideline Update. J Clin Oncol 2020 Aug 20;38(24):2782-2797. [doi: 10.1200/JCO.20.01296] [Medline: 32658626]
7. Iihara H, Abe M, Wada M, et al. 2023 Japan Society of clinical oncology clinical practice guidelines update for antiemesis. Int J Clin Oncol 2024 Jul;29(7):873-888. [doi: 10.1007/s10147-024-02535-x]
8. Y I. A possible role for ghrelinergic stimulation through blockade of 5-HT2B/5-HT2C receptors in antiemetic action of olanzapine. J Transl Sci 2019;5(6). [doi: 10.15761/JTS.1000300]
9. Hashimoto H, Abe M, Tokuyama O, et al. Olanzapine 5 mg plus standard antiemetic therapy for the prevention of chemotherapy-induced nausea and vomiting (J-FORCE): a multicentre, randomised, double-blind, placebo-controlled, phase 3 trial. Lancet Oncol 2020 Feb;21(2):242-249. [doi: 10.1016/S1470-2045(19)30678-3] [Medline: 31838011]
10. Yanai T, Iwasa S, Hashimoto H, et al. A double-blind randomized phase II dose-finding study of olanzapine 10 mg or 5 mg for the prophylaxis of emesis induced by highly emetogenic cisplatin-based chemotherapy. Int J Clin Oncol 2018 Apr;23(2):382-388. [doi: 10.1007/s10147-017-1200-4] [Medline: 29039073]
11. Nakashima K, Murakami H, Yokoyama K, et al. A phase II study of palonosetron, aprepitant, dexamethasone and olanzapine for the prevention of cisplatin-based chemotherapy-induced nausea and vomiting in patients with thoracic malignancy. Jpn J Clin Oncol 2017 Sep 1;47(9):840-843. [doi: 10.1093/jjco/hyx084] [Medline: 28633504]
12. Abe M, Hirashima Y, Kasamatsu Y, et al. Efficacy and safety of olanzapine combined with aprepitant, palonosetron, and dexamethasone for preventing nausea and vomiting induced by cisplatin-based chemotherapy in gynecological cancer:

XSL•FO
RenderX

KCOG-G1301 phase II trial. Support Care Cancer 2016 Feb;24(2):675-682. [doi: 10.1007/s00520-015-2829-z] [Medline: 26130365]
13. Tan L, Liu J, Liu X, et al. Clinical research of olanzapine for prevention of chemotherapy-induced nausea and vomiting. J Exp Clin Cancer Res 2009 Dec;28(1). [doi: 10.1186/1756-9966-28-131]
14. Bajpai J, Kapu V, Rath S, et al. Low-dose versus standard-dose olanzapine with triple antiemetic therapy for prevention of highly emetogenic chemotherapy-induced nausea and vomiting in patients with solid tumours: a single-centre, open-label, non-inferiority, randomised, controlled, phase 3 trial. Lancet Oncol 2024 Feb;25(2):246-254. [doi: 10.1016/S1470-2045(23)00628-9] [Medline: 38224701]
15. Hata A, Okamoto I, Inui N, et al. Randomized, double-blind, phase iii study of fosnetupitant versus fosaprepitant for prevention of highly emetogenic chemotherapy-induced nausea and vomiting: CONSOLE. JCO 2022 Jan 10;40(2):180-188. [doi: 10.1200/JCO.21.01315]
16. Matsuura K, Tsurutani J, Inoue K, et al. A phase 3 safety study of fosnetupitant as an antiemetic in patients receiving anthracycline and cyclophosphamide: CONSOLE - BC. Cancer 2022 Apr 15;128(8):1692-1698. [doi: 10.1002/cncr.34088]
17. Abe M, Iihara H, Aogi K. Fosnetupitant for the prevention of chemotherapy-induced nausea and vomiting: a short review and clinical perspective. Adv Ther 2023 May;40(5):1913-1925. [doi: 10.1007/s12325-023-02474-5]
18. Keepers GA, Fochtmann LJ, Anzia JM, et al. The American Psychiatric Association Practice Guideline for the treatment of patients with schizophrenia. AJP 2020 Sep 1;177(9):868-872. [doi: 10.1176/appi.ajp.2020.177901]
19. Tamura K, Aiba K, Saeki T, et al. Testing the effectiveness of antiemetic guidelines: results of a prospective registry by the CINV Study Group of Japan. Int J Clin Oncol 2015 Oct;20(5):855-865. [doi: 10.1007/s10147-015-0786-7] [Medline: 25681876]
20. Basch E, Reeve BB, Mitchell SA, et al. Development of the National Cancer Institute's Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE). JNCI Journal of the National Cancer Institute 2014 Sep 1;106(9):dju244-dju244. [doi: 10.1093/jnci/dju244]
21. Navari RM, Rapoport BL, Powers D, Arora S, Clark - Snow R. Rolapitant for the prevention of nausea in patients receiving highly or moderately emetogenic chemotherapy. Cancer Med 2018 Jul;7(7):2943-2950. [doi: 10.1002/cam4.1560]
22. Börjeson S, Hursti TJ, Peterson C, et al. Similarities and differences in assessing nausea on a verbal category scale and a visual analogue scale. Cancer Nurs 1997 Aug;20(4):260-266. [doi: 10.1097/00002820-199708000-00005] [Medline: 9265812]
23. Bloechl-Daum B, Deuson RR, Mavros P, Hansen M, Herrstedt J. Delayed nausea and vomiting continue to reduce patients' quality of life after highly and moderately emetogenic chemotherapy despite antiemetic treatment. JCO 2006 Sep 20;24(27):4472-4478. [doi: 10.1200/JCO.2006.05.6382]
24. Schwartzberg L, Navari R, Clark-Snow R, et al. Phase IIIb Safety and Efficacy of Intravenous NEPA for Prevention of Chemotherapy-Induced Nausea and Vomiting (CINV) in patients with breast cancer receiving initial and repeat cycles of anthracycline and cyclophosphamide (AC) chemotherapy. Oncologist 2020 Mar 1;25(3):e589-e597. [doi: 10.1634/theoncologist.2019-0527]
25. Yeo W, Lau TKH, Li L, et al. A randomized study of olanzapine-containing versus standard antiemetic regimens for the prevention of chemotherapy-induced nausea and vomiting in Chinese breast cancer patients. The Breast 2020 Apr;50:30-38. [doi: 10.1016/j.breast.2020.01.005]
26. Summary technical documentation for Arokaris (in Japanese) [Technical documentation describing the formulation, pharmacological background, and regulatory information of arokaris relevant to its clinical use]. PMDA. URL: https://www.pmda.go.jp/drugs/2022/P20220325004/index.html [accessed 2026-01-20]
27. Navari RM, Qin R, Ruddy KJ, et al. Olanzapine for the prevention of chemotherapy-induced nausea and vomiting. N Engl J Med 2016 Jul 14;375(2):134-142. [doi: 10.1056/NEJMoa1515725] [Medline: 27410922]
28. Ng TL, Hutton B, Clemons M. Chemotherapy-induced nausea and vomiting: time for more emphasis on nausea? Oncologist 2015 Jun;20(6):576-583. [doi: 10.1634/theoncologist.2014-0438] [Medline: 25948677]

Abbreviations

5-HT3RA: 5-hydroxytryptamine type 3 receptor antagonist
CC: complete control
CINV: chemotherapy-induced nausea and vomiting
CR: complete response
DEX: dexamethasone
EBC: early breast cancer
EC: epirubicin and cyclophosphamide
PS: Eastern Cooperative Oncology Group performance status
HEC: highly emetogenic chemotherapy
NK1RA: neurokinin-1 receptor antagonist
OLZ: olanzapine

XSL•FO
RenderX

PALO: palonosetron
PRO: patient-reported outcomes
PRO-CTCAE: Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events
QOL: quality of life
TC: total control
VAS: visual analog scale

Edited by A Schwartz; submitted 14.Oct.2025; peer-reviewed by K Kobayashi, K Matsuura; accepted 30.Dec.2025; published 03.Feb.2026.

Please cite as:
Higuchi A, Koi Y, Eto T, Maeda Y, Tajiri W, Kawasaki J, Akiyoshi S, Ijichi H, Nakamura Y, Koga C, Shimokawa M, Shimizu H, Matsumoto T, Tokunaga E
Optimizing Antiemetic Support in Anthracycline-Based Chemotherapy for Early Breast Cancer: Protocol for a Prospective Observational Study of Four-Drug Antiemetic Therapy Including Fosnetupitant and Olanzapine
JMIR Res Protoc 2026;15:e85648
URL: https://www.researchprotocols.org/2026/1/e85648
doi:10.2196/85648

© Ayako Higuchi, Yumiko Koi, Tomoaki Eto, Yuka Maeda, Wakako Tajiri, Junji Kawasaki, Sayuri Akiyoshi, Hideki Ijichi, Yoshiaki Nakamura, Chinami Koga, Mototsugu Shimokawa, Hiroaki Shimizu, Toshihiro Matsumoto, Eriko Tokunaga. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 3.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Pilot Implementation of a Post–Hypertensive Disorders of Pregnancy Education and Follow-Up Package for Health Care Providers: Protocol for a Mixed Methods Pilot Study

Jennifer Elizabeth Green[1,2], BMid, MPH, MHLM; Heike Roth[1,2], BMid, PhD; Ben Harris-Roxas[3,4], BSW, MPASR, PhD; Kathleen Baird[1], BMid, SFHEA, MA Ed, PhD; Caroline Guirgis[5,6,7], MBBS, DCH, FRACGP, AFRACMA; Zarin Gundevia[6,8], BMedSc, MBChB, DFSRH, MRCOG, MRCGP, FRACGP; Amanda Henry[2,4,9], MPH, PhD, FRANZCOG

[1]Collective for Midwifery, Child and Family Health, Faculty of Health, School of Nursing and Midwifery, University of Technology Sydney, Building 10, Level 11, 235 Jones Street, Sydney, Australia

[2]School of Clinical Medicine, Faculty of Medicine & Health, UNSW Sydney, Sydney, Australia

[3]School of Population Health, Faculty of Medicine & Health, UNSW Sydney, Sydney, Australia

[4]Faculty of Health, University of Technology Sydney, Sydney, Australia

[5]Myhealth Newtown, Sydney, Australia

[6]Royal North Shore Hospital, Sydney, Australia

[7]Sutherland Hospital, Sydney, Australia

[8]Randwick Doctors Medical Centre, Sydney, Australia

[9]The George Institute for Global Health, Sydney, Australia

Corresponding Author:
Jennifer Elizabeth Green, BMid, MPH, MHLM
Collective for Midwifery, Child and Family Health, Faculty of Health, School of Nursing and Midwifery, University of Technology Sydney, Building 10, Level 11, 235 Jones Street, Sydney, Australia

Abstract

Background: Medical complications of pregnancy provide a window into a woman's future health risk. Hypertensive disorders of pregnancy (HDP) affect 1 in 10 pregnant women and elevate the risk for women of experiencing long-term health complications within 5 years of the affected pregnancy, continuing lifelong. These risks include a doubled to tripled risk of developing cardiovascular disease, a doubled risk of developing type 2 diabetes, and a 5- to 10-fold risk of developing chronic kidney disease. Early assessment and intervention following HDP are therefore crucial to improving women's life-course health trajectory, as well as outcomes for any subsequent pregnancies. However, previous research has shown that Australian women and their primary health care providers are largely unaware of ongoing health risks and the necessary follow-up screening and assessments. Primary care providers also receive inadequate hospital-to-community handover and support to promote preventive health measures to women following pregnancy complications. Consequently, post-HDP care remains insufficient for optimizing long-term health.

Objective: This study aims to (1) explore whether a post-HDP education and follow-up service package can be designed, developed, and implemented among targeted general practitioners (GPs) and maternity hospitals across Sydney, Australia, and (2) evaluate whether the post-HDP education and follow-up package can address knowledge gaps among health care providers regarding the long-term health risks after HDP and build capacity among GPs to implement evidence-based care.

Methods: This pilot study will design, develop, and implement a post-HDP education and follow-up package ("the package") adopting a collaborative and implementation methodological approach. The package, designed by expert health care providers and informed by prior evidence-based research, will include educational materials, improved hospital-to-community handover, and a funded 6-month postpartum visit.

Results: Data collection occurred over an 18-month implementation and follow-up period between April 2024 and October 2025. Sixteen GPs across the Central and Eastern Primary Health Network (CESPHN) were recruited, along with their antenatal shared care (ANSC) affiliated tertiary referral hospitals in Sydney. Postimplementation data collection and analysis is planned for completion throughout 2026.

Conclusions: Mixed methods evaluation will assess the efficacy, acceptability, and utility of the post-HDP package among health care providers and inform its suitability for deployment at scale.

International Registered Report Identifier (IRRID): DERR1-10.2196/81069

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e81069) doi:10.2196/81069

KEYWORDS

hypertensive disorders of pregnancy; transitions of care; implementation; women's health; long-term health; cardiovascular; primary care; primary health care; health services administration; metabolic diseases

Introduction

Background

Pregnancy-induced complications provide an opportunity to facilitate women's engagement with primary health care services to promote health and reduce the burden of chronic disease in later life [1,2]. However, strategies for preventive interventions are lacking [3,4]. Supporting primary care services to enable this is a vital part of such health promotion endeavors [5,6]. Promoting population health and well-being throughout the life course is a critical global strategic objective, calling for an integrative, multisectoral approach [7]. Interdisciplinary postnatal care is acknowledged to influence short- and long-term health [1,8]. However, although the desirability of integrated health care is widely accepted throughout health policy, effective and sustainable implementation remains challenging [7,9,10].

Previous mixed methods Australian research [4,11-13] identified knowledge gaps of the long-term health risks following hypertensive disorders of pregnancy (HDP) among both women and their health care providers (HCPs). Among HCPs, knowledge gaps were identified in the classification, signs, and symptoms of various HDP and, specifically, the link between all HDP and increased cardiovascular disease and metabolic disorders in women and their children [4,13]. These findings cemented the necessity of addressing these gaps to improve long-term health outcomes for women following HDP [12,13].

The Society of Obstetric Medicine in Australia and New Zealand (SOMANZ) [14] has published recommendations to guide short- and long-term postpartum care after HDP. Addressing identified knowledge gaps and preferences of HCPs following HDP through a collaborative design of a post-HDP education and follow-up package (hereafter referred to as "the package") may also contribute to improved early intervention for modifiable health risks following HDP. Addressing HCP preferences for content and delivery seeks to enhance clinical integration between tertiary maternity providers and primary health care practitioners, as well as facilitate ongoing counseling and management of health risks after HDP. This study provides a unique opportunity to translate research findings into clinical practice using implementation methodology. This evidence-based approach supports optimal health service design and improved preventive and primary health outcomes. Upon completion of this research, findings will further inform adjustments to existing resources to prepare the package for broader scale-up. This may also include insights for future development of such an approach for all pregnancies, particularly those affected by pregnancy complications with associated long-term health risks.

Through a collaborative design process, using earlier research findings, this protocol describes the design, development, and implementation of a targeted post-HDP education and follow-up package for HCPs. The TIDieR (Template for Intervention Description and Replication) and StaRI (Standards for Reporting Implementation Studies) checklist will be used to support replicability for health service quality improvement [15,16].

Separate complementary research is currently underway focused on addressing the educational preferences and follow-up needs of women after HDP [17]. However, we acknowledge that the women receiving care from participating general practitioners (GPs) and maternity care providers may also, by association, benefit from improved knowledge and follow-up care facilitated throughout the study. While not specifically assessed in this study, increased HCP knowledge has the potential to improve outcomes for women at risk and reduce the distress for a woman, her family, and the community in the event of an adverse cardiovascular event.

Aims

This study aims to (1) explore whether a post-HDP education and follow-up service package can be designed, developed, and implemented among targeted GPs and maternity hospitals across Sydney, Australia, and (2) evaluate whether the post-HDP education and follow-up package can address knowledge gaps among health care providers regarding the long-term health risk after HDP and build capacity among GPs to implement evidence-based care.

Objectives

The objectives of this study are to design, develop, implement, and assess the efficacy, acceptability, and utility for health care providers of an evidence-based post-HDP education and follow-up package for women in the first 12 months following a hypertensive pregnancy (Table 1).

XSL•FO
RenderX

Table . Objectives and evaluation mapping.

Objective	Tool	Data	Evaluation	Time
Primary objective(s)				
Development of package	Evidence-based research: • Academic research • Policy and guidelines Co-design: • Steering committee consultation • Identification of platforms in consultation with steering committee	• Literature review [18] • Prior evidence reporting on HCP[a] knowledge gaps and education preferences regarding long-term health risks after HDP[b] [4,19] • SOMANZ[c] guidelines [14]	• Steering committee approval • Process and postimplementation interview data with hospital-based health care providers to describe the design and implementation of the package	• Preimplementation • Context-specific iterations throughout implementation period in consultation with site champions/practices (implementation science iterations)
Assess implementation success and sustainability	• Online surveys (GP[d]) • Semistructured interviews (HCP and GP)	• Qualitative and quantitative survey data • Qualitative interview data	• Aligning with CFiR[e] and RE-AIM[f] implementation methodology assessing fidelity and the enablers and barriers to implementation.	• Collected before, midway, and after implementation • Collected during and after implementation
Efficacy of education package	• Website data • HealthPathways data	• Deidentified quantitative data	• Analyzing website and HealthPathways access and traffic	• Collected after implementation
Efficacy of follow-up package	• Discharge referral communication between hospital and GP • 6-month follow-up appointments with participating GP	• Deidentified data relating to women receiving care from participating ANSC[g] GP. Collected from participating GP and associated maternity service.	• Quantitative analysis monitoring women discharged from maternity service with those followed up at participating GP	• Collected mid-way and after implementation
Acceptability and utility of the package (GP)	• Online surveys	• Qualitative and quantitative survey data	• Quantitative survey analysis will be undertaken using SPSS	• Collected before, midway, and after implementation
Acceptability and utility of package (GP and other HCPs)	• Semistructured interviews	• Qualitative data	• Thematic analysis	• Collected after implementation
Secondary objective(s)				
Changes in HCP confidence in transmission of information	• Online surveys • Semistructured interviews	• Qualitative and quantitative survey data • Qualitative interview data	• Comparing baseline confidence with postimplementation analysis	• Collected before, midway, and after implementation
Changes in HCP capability to provide evidence-based follow-up	• Online surveys • Interviews	• Qualitative and quantitative survey data • Qualitative interview data	• Assessing follow-up package including, discharge communication and adherence to 6-month follow-up appointment	• Collected before, midway, and after implementation
Collection of data informing potential scale-up of the package	• Online surveys • Semistructured interviews	• Qualitative and quantitative survey data • Qualitative interview data	• Quantitative survey data and qualitative free-text analysis. • Thematic analysis from interview data collection.	• Collected before, midway, and after implementation

[a]HCP: health care provider.

[b]HDP: hypertensive disorders of pregnancy.

XSL•FO
RenderX

[c]SOMANZ: Society of Obstetric Medicine of Australia and New Zealand.
[d]GP: general practitioner.
[e]CFiR: The Consolidated Framework for Implementation Research.
[f]RE-AIM: Reach, Effectiveness, Adoption, Implementation, Maintenance.
[g]ANSC: antenatal shared care.

Methods

Study Design

This study will utilize a multisite hybrid type 2 implementation study design [20,21]. A hybrid implementation study design acknowledges a dual focus to research to determine both clinical effectiveness and implementation strategies promoting accelerated research translation. This hybrid type 2 implementation study design aims to determine the effectiveness of an intervention pertaining to improved HCP knowledge and delivery of care after HDP, while also assessing the acceptability and utility of the implementation strategy ("the package") [20,21]. Evaluation will be undertaken using mixed methods. Mixed methods evaluation is supported as an appropriate method for pragmatic research and implementation science. It facilitates a multidimensional lens to measure the drivers of implementation success and sustainability [22-24].

1. Assessment of acceptability and utility of the post-HDP package will require health care provider insight. This will be achieved quantitatively by survey data collection before, during, and after implementation, and qualitatively through semistructured interviews after implementation.
2. Assessment of the efficacy of system change lends itself to a quantitative analysis, which will be achieved by monitoring the number of women followed up by their GP throughout the study period.
3. Changes in HCP knowledge and confidence in transmitting evidence-based information to women regarding long-term health risks following HDP will be assessed quantitatively before, during, and after implementation via survey data collection.

The study will be conducted in 3 phases: phase 1 (preimplementation), phase 2 (implementation), and phase 3 (postimplementation) (Figure 1).

Figure 1. Study diagram. CESPHN: Central and Eastern Primary Health Network; EOI: expression of interest; GP: general practitioner; HCP: health care provider; HDP: hypertensive disorders of pregnancy.

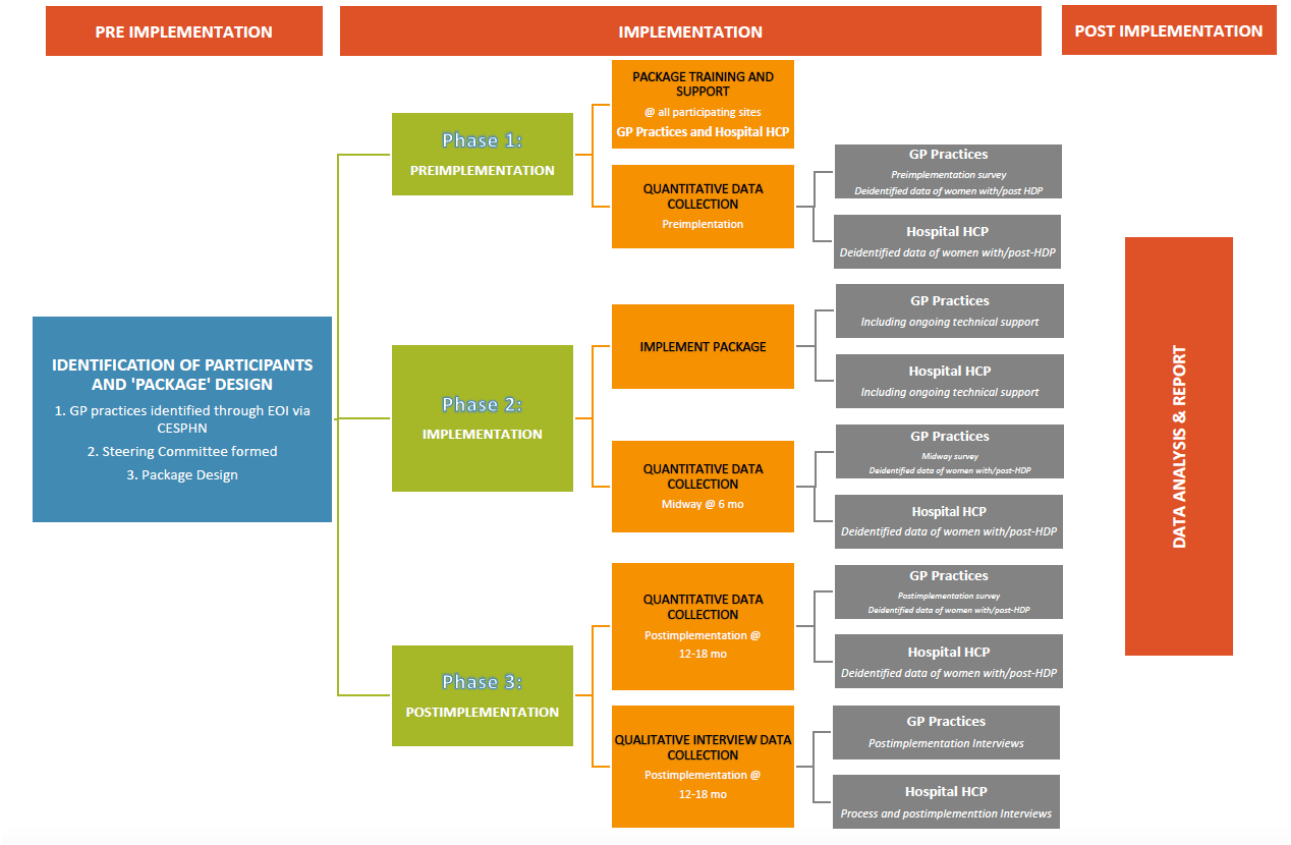

Phase 1: Preimplementation

The preimplementation phase includes:

- Recruitment,
- The package co-design, and
- Preimplementation survey data collection with participating GPs.

Phase 2: Implementation

The implementation phase includes:

XSL•FO
RenderX

- Implementation of the package at participating GP practices and affiliated hospitals over an 18-month period,
- Training and implementation support regarding the package at participating practices and hospitals, and
- Midway implementation survey data collection with participating GPs 6 months following implementation.

Phase 3: Postimplementation

The postimplementation phase includes:

- Postimplementation survey data collection with participating GPs at the conclusion of the implementation period and
- Semistructured qualitative interviews with health care providers (GPs and hospital-based staff) within 6 months of the implementation period concluding.

Sample Size

The planned sample size consists of at least 10 GPs (n=10) employed in participating GP practices where implementation and evaluation of the package will occur. The planned sample size aims to provide sufficient numbers for testing the acceptability and utility of the package, and the potential for scale-up. In keeping with the pilot and implementation science nature of the study, no formal power calculation was done.

Key midwifery and obstetric staff at the linked maternity sites involved in care of women with HDP and/or GP antenatal shared care (ANSC) will also be interviewed (anticipated n=5 - 10).

Setting and Participants

The study will specifically seek to include ANSC GPs from practices within the Central Eastern Sydney Primary Health Network (CESPHN) and their affiliated maternity care providers. With a reported 19,356 births within CESPHN in 2021, CESPHN covers a large and diverse population, providing services across approximately 590 km^2 [25]. The inclusion of ANSC GPs will promote a cohort engaged with women's health throughout the perinatal period with already established relationships with maternity care providers. Furthermore, the diversity of the population combined with the locality of tertiary level maternity services within CESPHN is appropriate in the context of providing complex maternity care such as HDP.

Eligibility Criteria

The following eligibility criteria will be used for this study:

- ANSC GPs within the participating CESPHN practices caring for women who experience HDP
- Maternity services within CESPHN collaborating with identified ANSC GP/GP practices participating in study providing care to women experiencing HDP

Recruitment

Expression of Interest process for CESPHN GP practices, including recruitment of GP representatives for the package co-design and study Steering Committee.

Recruitment of affiliated hospitals via targeted emails and meetings with relevant maternity service providers.

Ethical Considerations

Ethical approval has been obtained from the appropriate Human Research Ethics Committees across 3 hospital sites in Sydney (2023/ETH00812 | X23-0281 & 2023/STE01377 | 2023/STE01378 | 2023/STE01379). Ethical ratification was also granted by the University of Technology Sydney (ETH23-8511) and University of New South Wales.

The study will be conducted as outlined throughout the approved study protocols associated with ethical approval. Written informed consent for all aspects of the study will be obtained from GP participants and from hospital providers for their participation in interviews. Women being cared for will not provide informed consent as post-HDP follow-up is recommended standard of care [14].

The Post-HDP Package

Design

The package concept was primarily informed by prior research and evidence-based practice guidelines [4,14,26-31]. To develop a feasible, appropriate, and acceptable post-HDP package for HCPs, an expert project steering committee was formed. This collaborative approach aligns with an experience-based co-design process aimed at improving knowledge and capacity to promote evidence-based integrated health care. This approach specifically addresses the needs of the HCP end users [32,33].

The project steering committee consists of the research team, site-specific hospital-based maternity care providers, GPs, and representation from the CESPHN. The multidisciplinary team reflects targeted expertise across primary care, midwifery, obstetrics, care related to HDP and post-HDP follow-up, integrated care, and research design.

The project objectives specifically outline the design, implementation, and evaluation of a post-HDP package targeted at HCPs. Therefore, women were not included in the steering committee or co-design. Separate complementary research is currently underway focused on addressing educational preferences and follow-up needs of women after HDP [17].

A steering committee–led, co-designed post-HDP package comprising the following components will be implemented.

Post-HDP Education Package

The education package will include:

1. A dedicated, password-protected website for HCPs providing tailored post-HDP education regarding long-term health risks.
2. A prerecorded webinar series for health care providers—made available through the dedicated website—outlining long-term health risks following HDP, recommended follow-up care, and case study examples.
3. A dedicated post-HDP "HealthPathway" [34], which is an online health information tool designed for GPs and HCPs, aims to provide context-specific health education, assessment, management, and referral information to support clinical decision-making, providing recommended education and follow-up for women following HDP.

XSL•FO
RenderX

4. Hard-copy and electronic materials provided to participating GPs to facilitate information transmission of the long-term health risks borne by women following HDP, including recommended follow-up care.

Post-HDP Follow-Up Package

Through site-specific specialist consultation, the package will be adapted to each site-specific needs to align with variations in local operating procedures.

1. *Post-HDP hospital discharge pathway*: Dedicated post-HDP discharge communication will include post-HDP–specific risk information and defined follow-up care.
2. *Reminder system for GPs to recall women after HDP for recommended follow-up*: A reminder system will be enabled utilizing GP practice management systems to recall women following HDP for recommended follow-up. This process will be managed by the GP practices individually, is reflected in GP practice remuneration, and will be agreed on prior to participation.
3. *Postnatal follow-up appointment at 6 months for women post HDP*: Aligning with evidence-based guidelines [14], a 6-month post-HDP visit will be offered to women by the participating GP. Costs will be covered by allocated research funds and reflected in the GP practice's agreed remuneration to eliminate any out-of-pocket costs for the participating GP practices or women attending the visit.

Implementation

Extensive consultation will occur across each participating site (GPs and maternity hospitals) to ensure development of a context-appropriate intervention within the constraints of local operating procedures.

This flexible approach, adopting context-specific design, aligns with a realist approach to implementation science to promote sustainable and effective implementation [22]. This further aligns with the prior research undertaken by Roth et al [4] exploring health care provider preferences regarding education on long-term health after HDP.

Implementation Support

The researcher (JG), along with site "champions" (nominated by site-specific representatives and principal investigators), will facilitate implementation support, including the development of educational resources and quick reference guide dissemination, in-service education, and face-to-face training with staff tasked with implementing the post-HDP package.

Sustained support will continue throughout the study period. This may include, but is not limited to, adapting implementation strategies, which will be documented according to the hybrid implementation study design.

Data Collection and Analysis

Quantitative Data Collection

Survey data will be collected using the REDCap (Research Electronic Data Capture) database (University of Technology Sydney, administered and secured).

Deidentified quantitative data on educational website access and general traffic will be collected for analysis to evaluate health care provider acceptability and utility. This will include the number of clicks on each page and/or their embedded items and length of time spent on each page.

Deidentified data relating to women receiving care from participating GPs with post-HDP or referred to participating GP practices will be collected from participating maternity hospitals and GP practices. These data will be sourced by reports managed by health informatics personnel, as agreed within the participating sites within the study period.

Qualitative Data Collection

Semistructured interviews with all participating GPs will occur within 6 months of completion of the implementation study period. Hospital HCPs participating in the study will receive an invitation to attend a postimplementation interview via email. Semistructured interviews with all consenting hospital HCPs will occur within 6 months of the completion of the study period. Interview data will be collected either in person or by telephone or videoconference, depending on participant preference, by an investigator/researcher (JEG).

Data Analysis

Survey data will be analyzed descriptively for each time point and comparatively across preimplementation, mid-way, and postimplementation periods. Interview data will be analyzed using thematic analysis as described by Braun and Clarke [35].

Deidentified data relating to women receiving care from participating GPs and maternity care providers will be analyzed quantitatively to determine the number of women discharged from maternity care providers and those followed up with by participating GPs. This analysis will determine the efficacy of hospital-to-GP communication.

Categorical data will be presented as numbers and proportions. Quantitative survey analyses will be undertaken using SPSS Statistics for Windows, version 29 (IBM Corp).

Results

Data collection occurred over an 18-month implementation and follow-up period between April 2024 and October 2025. A total of 16 GPs across the CESPHN were recruited, along with their ANSC affiliated tertiary referral hospital in Sydney. Postimplementation data collection and analysis is planned for completion throughout 2026. This study will provide mixed methods data on the efficacy, acceptability, and utility of a co-designed, targeted, and evidence-based post-HDP education and follow-up package for HCPs. The results will also describe the co-design of the package, highlighting the barriers and enablers to effective implementation with a view to informing broader future scale-up.

Discussion

Principal Purpose

Given the strong association between HDP and increased risk of chronic disease [30,36-39], the care provided throughout the

XSL•FO
RenderX

perinatal period is vital for optimizing harm minimization through targeted education and primary health promotive strategies [30,39-41]. As highlighted previously, complementary research is currently underway focused on addressing educational preferences and follow-up needs of women after HDP [17]. To our knowledge, this is the first implementation of a co-designed targeted post-HDP education and follow-up package in Australia aimed at addressing knowledge gaps and enhancing capacity among HCPs to support care after HDP.

A recent publication by Slater et al [42] provides evidence to support our multidisciplinary co-design approach for targeted interventions to improve the uptake of evidence-based care after HDP. It also demonstrates demand for this research in the context of improving primary care management post HDP more broadly. Additionally, although not HDP-specific, MacKay et al [43] developed a complex health system intervention for improving systems of care in the context of hyperglycemia. MacKay et al [43] used implementation frameworks to support sustainable implementation in the Northern Territory and Far North Queensland. This research further supports our methodological approach to implementing a multifaceted post-HDP education and follow-up package in Australia's complex health system.

Current Australian postnatal transitions of care following pregnancy are inadequate, highlighting the lack of appropriate handover between services as a barrier to the uptake of recommended follow-up care [18]. This research is therefore timely in the context of the integrated care landscape in Australia.

Strengths and Limitations

The research aims to bridge an identified practice gap by enhancing knowledge and improving health care systems to promote uptake of contemporary evidence-based care guidelines. Adopting a collaborative and evidence-informed approach in consultation with health care providers (the end-users) is a key strength of the study. Although the focus on metropolitan Sydney may limit the generalizability, the multisite design will provide valuable context-specific insights for future scale-up.

The planned sample size of GPs (n=10) and HCPs (n=5 - 10) is appropriate for the pilot study design. However, the small sample size limits statistical power and generalizability. Furthermore, the targeted recruitment of ANSC GPs likely favors early adopters. However, the complexities of health system design provide challenges for implementing meaningful system change. Therefore, meaningful and sustained engagement is essential for providing insight to inform evidence-based strategies for future scale-up.

Furthermore, the study period may limit the ability to influence permanent digital health solutions (follow-up intervention) due to lengthy approval processes. However, the use of evidence-based methodological implementation and reporting frameworks will promote replicable and robust results to influence policy and practice.

Conclusions

This pilot implementation study will determine whether a post-HDP education and follow-up service package can be designed, developed, and implemented among GPs and maternity hospitals across Sydney. The findings of this research will provide an opportunity to influence contemporary development of effective post-HDP communication strategies, thereby improving the transition of care between tertiary and primary care services, and informing future implementation at scale.

Funding

AH is supported by a Cardiovascular Early-Mid Career Researcher Grant from NSW Health (Preventing Heartbreak After Hypertensive Pregnancy, https://medicalresearch.nsw.gov.au/grants/projects/2022/preventing-heartbreak-after-hypertensive-pregnancy). JEG currently receives an Australian Government Research Training Program Fees Offset Scholarship for her PhD which this study forms part of.

Data Availability

The full ethics-approved protocol will be available on reasonable request once the study data collection has concluded and findings have been published.

Authors' Contributions

Conceptualization: JEG, HR, BH-R, AH
Methodology: JEG
Investigation: JEG
Formal analysis: JEG
Supervision: HR, BH-R, KB, AH
Project administration: AH
Project administration: AH
Funding acquisition: AH
Steering committee collaboration: CG, ZG
Writing – original draft: JEG
Writing – review & editing: JEG, HR, BH-R, KB, CG, ZG, AH

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Sheiner E, Kapur A, Retnakaran R, et al. FIGO (International Federation of Gynecology and Obstetrics) postpregnancy initiative: long-term maternal implications of pregnancy complications-follow-up considerations. Int J Gynaecol Obstet 2019 Sep;147 Suppl 1(Suppl 1):1-31. [doi: 10.1002/ijgo.12926] [Medline: 32323876]
2. Hill B, Skouteris H, Boyle JA, et al. Health in preconception, pregnancy and postpartum global alliance: international network pregnancy priorities for the prevention of maternal obesity and related pregnancy and long-term complications. J Clin Med 2020 Mar 18;9(3):822. [doi: 10.3390/jcm9030822] [Medline: 32197374]
3. Roth H, Homer CSE, LeMarquand G, et al. Assessing Australian women's knowledge and knowledge preferences about long-term health after hypertensive disorders of pregnancy: a survey study. BMJ Open 2020 Dec 17;10(12):e042920. [doi: 10.1136/bmjopen-2020-042920] [Medline: 33334841]
4. Roth H, Morcos V, Roberts LM, Hanley L, Homer CSE, Henry A. Preferences of Australian healthcare providers regarding education on long-term health after hypertensive disorders of pregnancy: a qualitative study. BMJ Open 2022 May 26;12(5):e055674. [doi: 10.1136/bmjopen-2021-055674] [Medline: 35618327]
5. Rossiter C, Henry A, Roberts L, et al. Optimising mothers' health behaviour after hypertensive disorders of pregnancy: a qualitative study of a postnatal intervention. BMC Public Health 2022 Jun 27;22(1):1259. [doi: 10.1186/s12889-022-13590-2] [Medline: 35761317]
6. Henry A, Mangos G, Roberts LM, et al. Preeclampsia-associated cardiovascular risk factors 6 months and 2 years after pregnancy: the p4 study. Hypertension 2024 Apr;81(4):851-860. [doi: 10.1161/HYPERTENSIONAHA.123.21890] [Medline: 38288610]
7. Australian Health Ministers' Advisory Council. National strategic framework for chronic conditions. : Australian Government; 2017 URL: https://www.health.gov.au/sites/default/files/documents/2019/09/national-strategic-framework-for-chronic-conditions.pdf [accessed 2026-01-21]
8. Haran C, van Driel M, Mitchell BL, Brodribb WE. Clinical guidelines for postpartum women and infants in primary care-a systematic review. BMC Pregnancy Childbirth 2014 Jan 29;14:51. [doi: 10.1186/1471-2393-14-51] [Medline: 24475888]
9. Amri M, Chatur A, O'Campo P. Intersectoral and multisectoral approaches to health policy: an umbrella review protocol. Health Res Policy Syst 2022 Feb 15;20(1):21. [doi: 10.1186/s12961-022-00826-1] [Medline: 35168597]
10. Trankle SA, Usherwood T, Abbott P, et al. Key stakeholder experiences of an integrated healthcare pilot in Australia: a thematic analysis. BMC Health Serv Res 2020 Oct 7;20(1):925. [doi: 10.1186/s12913-020-05794-2] [Medline: 33028299]
11. Roth H, Henry A, Roberts L, Hanley L, Homer CSE. Exploring education preferences of Australian women regarding long-term health after hypertensive disorders of pregnancy: a qualitative perspective. BMC Womens Health 2021 Nov 1;21(1):384. [doi: 10.1186/s12905-021-01524-w] [Medline: 34724948]
12. Roth H, LeMarquand G, Henry A, Homer C. Assessing knowledge gaps of women and healthcare providers concerning cardiovascular risk after hypertensive disorders of pregnancy-a scoping review. Front Cardiovasc Med 2019;6:178. [doi: 10.3389/fcvm.2019.00178] [Medline: 31850374]
13. Roth H, Homer CSE, Arnott C, Roberts L, Brown M, Henry A. Assessing knowledge of healthcare providers concerning cardiovascular risk after hypertensive disorders of pregnancy: an Australian national survey. BMC Pregnancy Childbirth 2020 Nov 23;20(1):717. [doi: 10.1186/s12884-020-03418-5] [Medline: 33228597]
14. SOMANZ hypertension in pregnancy guideline 2023. : SOMANZ - Society of Obstetric Medicine of Australia and New Zealand; 2024 Feb 5 URL: https://somanz.org/content/uploads/2024/01/SOMANZ_Hypertension_in_Pregnancy_Guideline_2023.pdf [accessed 2026-01-21]
15. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014 Mar 7;348:g1687. [doi: 10.1136/bmj.g1687] [Medline: 24609605]
16. Pinnock H, Epiphaniou E, Sheikh A, et al. Developing standards for reporting implementation studies of complex interventions (StaRI): a systematic review and e-Delphi. Implement Sci 2015 Mar 30;10:42. [doi: 10.1186/s13012-015-0235-z] [Medline: 25888928]
17. McNally S, Roth H, Henry A. Closing knowledge gaps for women after hypertensive disorders of pregnancy: an educational pilot study. J Paediatr Child Health 2023;59(S1):55-134. [doi: 10.1111/jpc.16357]
18. Green JE, Ang N, Harris-Roxas B, Baird K, Roth H, Henry A. Exploring Australian knowledge and practice for maternal postnatal transition of care between hospital and primary care: a scoping review. Women Birth 2025 Jan;38(1):101852. [doi: 10.1016/j.wombi.2024.101852] [Medline: 39752774]
19. Roth H. Postnatal cardiovascular risk after hypertensive disorder of pregnancy: identifying knowledge needs and education recommendations for women and healthcare providers [Master's Thesis]. : University of Technology Sydney; 2021 URL: https://opus.lib.uts.edu.au/handle/10453/154739 [accessed 2026-01-21]

XSL•FO
RenderX

20. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226. [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
21. Curran GM, Landes SJ, McBain SA, et al. Reflections on 10 years of effectiveness-implementation hybrid studies. Front Health Serv 2022;2:1053496. [doi: 10.3389/frhs.2022.1053496] [Medline: 36925811]
22. Sarkies MN, Francis-Auton E, Long JC, Pomare C, Hardwick R, Braithwaite J. Making implementation science more real. BMC Med Res Methodol 2022 Jun 25;22(1):178. [doi: 10.1186/s12874-022-01661-2] [Medline: 35752754]
23. Bauer MS, Damschroder L, Hagedorn H, Smith J, Kilbourne AM. An introduction to implementation science for the non-specialist. BMC Psychol 2015 Sep 16;3(1):32. [doi: 10.1186/s40359-015-0089-9] [Medline: 26376626]
24. Long KM, McDermott F, Meadows GN. Being pragmatic about healthcare complexity: our experiences applying complexity theory and pragmatism to health services research. BMC Med 2018 Jun 20;16(1):94. [doi: 10.1186/s12916-018-1087-6] [Medline: 29921277]
25. GP antenatal shared care program. CESPHN - Central and Eastern Sydney Primary health Network. 2021. URL: https://cesphn.org.au/general-practice/help-my-patients-with/child-and-maternal-health [accessed 2026-01-21]
26. Magee LA, Brown MA, Hall DR, et al. The 2021 International Society for the Study of Hypertension in Pregnancy classification, diagnosis & management recommendations for international practice. Pregnancy Hypertens 2022 Mar;27:148-169. [doi: 10.1016/j.preghy.2021.09.008] [Medline: 35066406]
27. Lowe SA, Bowyer L, Lust K, et al. The SOMANZ guidelines for the management of hypertensive disorders of pregnancy 2014. Aust N Z J Obstet Gynaecol 2015 Feb;55(1):11-16. [doi: 10.1111/ajo.12253] [Medline: 25308532]
28. Communication at clinical handover. Australian Commission on Safety and Quality in Health Care. 2023. URL: https://www.safetyandquality.gov.au/standards/nsqhs-standards/communicating-safety-standard/communication-clinical-handover [accessed 2026-01-21]
29. Kilgour C, Bogossian FE, Callaway L, Gallois C. Postnatal gestational diabetes mellitus follow-up: perspectives of Australian hospital clinicians and general practitioners. Women Birth 2019 Feb;32(1):e24-e33. [doi: 10.1016/j.wombi.2018.04.011] [Medline: 29735392]
30. Davis GK, Henry A, Arnott C, Brown MA. The long-term cardiovascular impact of hypertension in pregnancy - a missed opportunity. Aust N Z J Obstet Gynaecol 2021 Jun;61(3):474-477. [doi: 10.1111/ajo.13335] [Medline: 33734427]
31. Psaila K, Fowler C, Kruske S, Schmied V. A qualitative study of innovations implemented to improve transition of care from maternity to child and family health (CFH) services in Australia. Women Birth 2014 Dec;27(4):e51-e60. [doi: 10.1016/j.wombi.2014.08.004] [Medline: 25257376]
32. Vargas C, Whelan J, Brimblecombe J, Allender S. Co-creation, co-design, co-production for public health - a perspective on definition and distinctions. Public Health Res Pract 2022 Jun 15;32(2):3222211. [doi: 10.17061/phrp3222211] [Medline: 35702744]
33. Fylan B, Tomlinson J, Raynor DK, Silcock J. Using experience-based co-design with patients, carers and healthcare professionals to develop theory-based interventions for safer medicines use. Res Social Adm Pharm 2021 Dec;17(12):2127-2135. [doi: 10.1016/j.sapharm.2021.06.004] [Medline: 34187746]
34. What is HealthPathways? HealthPathways. 2025. URL: https://www.healthpathwayscommunity.org/About [accessed 2026-01-21]
35. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021 Jul 3;18(3):328-352. [doi: 10.1080/14780887.2020.1769238]
36. Brown MA, Roberts L, Hoffman A, et al. Recognizing cardiovascular risk after preeclampsia: the P4 study. J Am Heart Assoc 2020 Nov 17;9(22):e018604. [doi: 10.1161/JAHA.120.018604] [Medline: 33170079]
37. Roberts JM, Hubel CA. Pregnancy: a screening test for later life cardiovascular disease. Womens Health Issues 2010 Sep;20(5):304-307. [doi: 10.1016/j.whi.2010.05.004] [Medline: 20800765]
38. Roberts L, Davis GK, Homer CSE. Depression, anxiety, and post-traumatic stress disorder following a hypertensive disorder of pregnancy: a narrative literature review. Front Cardiovasc Med 2019;6(147):147. [doi: 10.3389/fcvm.2019.00147] [Medline: 31649935]
39. Arnott C, Nelson M, Alfaro Ramirez M, et al. Maternal cardiovascular risk after hypertensive disorder of pregnancy. Heart 2020 Dec;106(24):1927-1933. [doi: 10.1136/heartjnl-2020-316541] [Medline: 32404402]
40. Lui NA, Jeyaram G, Henry A. Postpartum interventions to reduce long-term cardiovascular disease risk in women after hypertensive disorders of pregnancy: a systematic review. Front Cardiovasc Med 2019;6:160. [doi: 10.3389/fcvm.2019.00160] [Medline: 31803757]
41. Henry A, Arnott C, Makris A, et al. Blood pressure postpartum (BP2) RCT protocol: follow-up and lifestyle behaviour change strategies in the first 12 months after hypertensive pregnancy. Pregnancy Hypertens 2020 Oct;22:1-6. [doi: 10.1016/j.preghy.2020.07.001] [Medline: 32679537]
42. Slater K, Taylor R, McLaughlin K, et al. Co-designing an intervention for cardiovascular disease risk assessment and management after hypertensive disorders of pregnancy in primary care. Health Res Policy Syst 2025 Feb 20;23(1):23. [doi: 10.1186/s12961-024-01269-6] [Medline: 39979977]

XSL•FO
RenderX

43. MacKay D, Kirkham R, Freeman N, et al. Improving systems of care during and after a pregnancy complicated by hyperglycaemia: a protocol for a complex health systems intervention. BMC Health Serv Res 2020 Sep 1;20(1):814. [doi: 10.1186/s12913-020-05680-x] [Medline: 32867837]

Abbreviations

ANSC: antenatal shared care
CESPHN: Central Eastern Sydney Primary Health Network
GP: general practitioner
HCP: health care provider
HDP: hypertensive disorders of pregnancy
REDCap: Research Electronic Data Capture
StaRI: Standards for Reporting Implementation Studies
TIDieR: Template for Intervention Description and Replication

Edited by J Sarvestan; submitted 22.Jul.2025; peer-reviewed by O Sari; revised version received 30.Nov.2025; accepted 24.Dec.2025; published 10.Feb.2026.

Please cite as:
Green JE, Roth H, Harris-Roxas B, Baird K, Guirgis C, Gundevia Z, Henry A
Pilot Implementation of a Post–Hypertensive Disorders of Pregnancy Education and Follow-Up Package for Health Care Providers: Protocol for a Mixed Methods Pilot Study
JMIR Res Protoc 2026;15:e81069
URL: https://www.researchprotocols.org/2026/1/e81069
doi:10.2196/81069

© Jennifer Elizabeth Green, Heike Roth, Ben Harris-Roxas, Kathleen Baird, Caroline Guirgis, Zarin Gundevia, Amanda Henry. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Establishing Reference Intervals for Fat-Soluble Vitamins in Healthy Chinese Adults: Study Protocol for a Cross-Sectional, Multicenter Study

Zhong Liu, PhD; Chenbing Liu, MA; Nan Li, MS; Chao Shen, MS; Lihong Qiu, PhD; Di Sheng, BA

Department of Health Management Center, The First Affiliated Hospital of Zhejiang University School of Medicine, 79 Qingchun Rd, Hangzhou, Zhejiang Province, China

Corresponding Author:
Zhong Liu, PhD
Department of Health Management Center, The First Affiliated Hospital of Zhejiang University School of Medicine, 79 Qingchun Rd, Hangzhou, Zhejiang Province, China

Abstract

Background: Fat-soluble vitamins (FSVs)—vitamins A, D, E, and K—are essential micronutrients involved in key physiological processes. Both deficiency and excess can influence nutritional assessment and disease risk. In China, clinical evaluation of FSVs often relies on reference intervals (RIs) derived from Western populations, and no large-scale study has comprehensively evaluated all 4 FSVs in healthy Chinese adults.

Objective: The aim of this study is to establish population-specific RIs for vitamins A, D, E, and K in healthy Chinese adults, thereby supporting more accurate clinical assessment of vitamin status and the prevention of related diseases.

Methods: This ongoing multicenter cross-sectional study aims to recruit 100,000 adults (≥18 y) from 20 hospitals across China. Participants undergo a standardized questionnaire, physical examination, and blood sample collection. FSV concentrations are quantified using liquid chromatography-tandem mass spectrometry. Group differences are assessed using the chi-square test or Fisher exact test. Nested ANOVA evaluates variation across subgroups. RIs will be established using parametric or nonparametric percentile methods following rigorous outlier removal to accurately determine the 2.5th and 97.5th percentile limits.

Results: As of October 2025, 13,545 adults have been enrolled, and 1690 participants have met the inclusion criteria. Recruitment began on 1 July 2024 and is expected to conclude by 30 June 2026.

Conclusions: This study will generate the first large-scale, population-specific RIs for FSVs in healthy Chinese adults. The findings are expected to be published in 2027 and will provide an important evidence base for clinical nutrition assessment and disease prevention in China.

International Registered Report Identifier (IRRID): DERR1-10.2196/80425

(JMIR Res Protoc 2026;15:e80425) doi:10.2196/80425

KEYWORDS

fat-soluble vitamins; reference intervals; Chinese population; physical examination; multicenter

Introduction

Fat-soluble vitamins (FSVs), including A, D, E, and K, are micronutrients that play essential roles in maintaining human health. Unlike water-soluble vitamins, these nutrients are stored in the adipose tissues, liver, and muscles, making them prone to accumulation and potential toxicity [1]. Vitamin A is crucial for vision, immune function, and cellular communication [2]. Its deficiency can lead to impaired vision, particularly night blindness, and an increased risk of infections [3]. Vitamin D is well-known for regulating calcium homeostasis and bone metabolism. Deficiency in vitamin D can result in different pathologies such as diabetes, certain cancers, cardiovascular diseases, fertility, and many other conditions [4]. Vitamin E, primarily an antioxidant, is essential for nervous system development, growth, and immunity [5]. Vitamin E deficiency is associated with increased infection, anemia, nervous system abnormalities, and stunted growth [6]. Vitamin K is essential for blood coagulation and bone health, while insufficient vitamin K leads to higher risk of bone fracture and cardiovascular disease [7]. Because of these essential functions, maintaining appropriate levels of FSVs is crucial; both deficiency and excess can influence nutritional assessment and clinical decision-making.

Reference intervals (RIs) serve as the basis of laboratory testing and aid the physician in differentiating between the healthy and diseased patients. Standard methods for determining the RIs are to define and obtain a healthy population of at least 120 individuals and use nonparametric estimates of the 95% RI [8]. In China, the need for population-specific RIs is particularly

XSL•FO
RenderX

pressing due to the country's vast genetic diversity and unique environmental and lifestyle factors. Many clinical laboratories in China rely on guideline- or manufacturer-provided RIs, which may not fully reflect the Chinese population and can contribute to misclassification and inappropriate clinical decisions [9,10]. Therefore, developing and periodically updating RIs tailored to the Chinese population is crucial for effective health care delivery and public health management.

Research on FSVs often focuses on specific vitamins, particular subgroups, or limited regions, leading to a lack of comprehensive data tailored to the healthy Chinese adult population. For example, global research on vitamin D emphasizes widespread deficiency among diverse populations and its correlation with chronic diseases [11]. A recent study from a southwestern province reported RIs for vitamins A and E, but its relatively small sample size limits its generalizability to the broader Chinese population [12]. Previous studies on vitamin E status have focused on elderly cohorts [13], infants [14], or specific regional populations [12], and lack conclusive data [15], underscoring the need for broader research that reflects the overall healthy adult population. Similarly, vitamin K research often focuses on its deficiency in osteoporosis, cardiovascular contexts, or other related diseases, neglecting its broader nutritional status in populations [16-18]. Therefore, this gap is especially notable in China, as no large-scale study has analyzed all 4 FSVs in a healthy adult population. Nonetheless, current clinical practice often relies on RIs based on Western studies, which may not fully capture the nutritional needs of Chinese individuals. Factors such as dietary patterns, geographic diversity, and cultural practices can all influence FSV levels. Given the substantial differences in lifestyle and dietary patterns between Chinese and Western populations, establishing population-specific RIs is essential for accurate nutritional assessment and public health guidance.

To address these issues, this study presents a protocol for establishing population-specific RIs for FSVs in a healthy Chinese adult population. Using a cross-sectional study design, we will collect a representative sample of individuals across diverse regions and demographics to account for variations in lifestyle and environmental exposure. This protocol outlines a detailed methodology for conducting the study, including participant selection, laboratory methods for vitamin quantification, and statistical techniques for establishing the RIs. To the best of our knowledge, this is the first cross-sectional study to assess the RIs for FSVs in this specific population. These RIs will offer clinicians more accurate tools for evaluating vitamin status and supporting public health efforts, ultimately helping prevent diseases related to both deficiency and excess.

Methods

Study Design

This is a multicenter, cross-sectional study. Participants will be recruited from 20 participating hospitals in China. The First Affiliated Hospital of Zhejiang University designed and initiated the study. All participating hospitals will be required to conduct the study in strict accordance with the established rules. Figure 1 presents a summary of the study process.

XSL•FO
RenderX

Figure 1. A flowchart of the study plan.

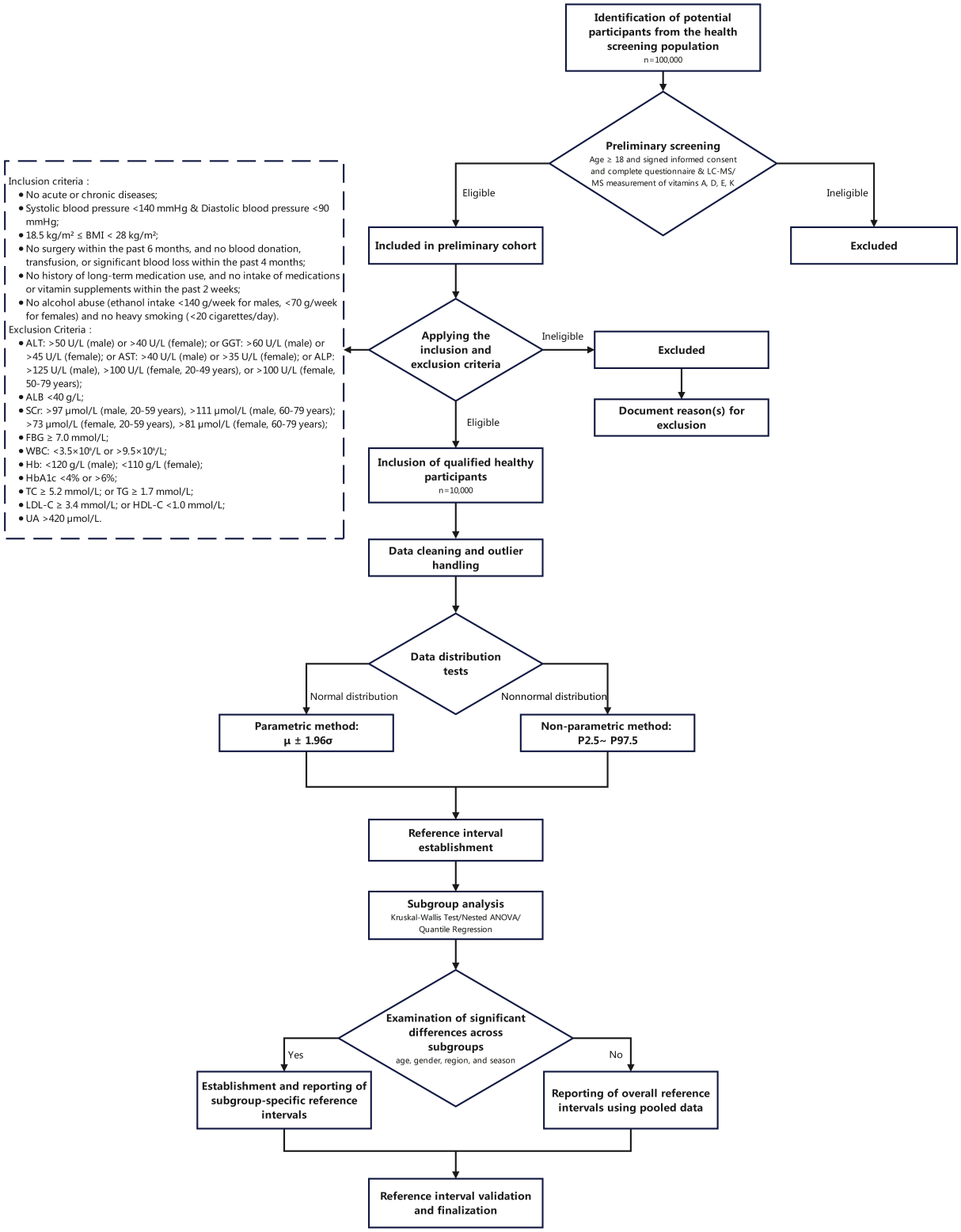

Patient and Public Involvement

Neither participants nor the public were involved in the design, implementation, reporting, or dissemination planning of this research.

Data Collection

Data collection consists of a standardized questionnaire and a comprehensive physical examination with laboratory testing. Participants are individuals undergoing health check-ups at participating hospital centers. Before the check-up begins, trained nurses confirm the selected examination package and

XSL•FO
RenderX

invite participants to complete the questionnaire, which captures demographic characteristics, medical and family history, smoking and alcohol use, physical activity, and medication or supplement use (details in Table 1). Afterward, participants undergo standardized physical examinations, including measurements of height, weight, body composition, and blood pressure. Laboratory evaluations include a comprehensive metabolic panel, lipid panel, liver function tests, and FSVs level test. All test instruments are calibrated following the manufacturer's guidelines. Blood samples are collected following an overnight fasting of 12 hours.

Table . The core components of the questionnaire.

Questionnaire outline	Description of content
Basic personal details	ID number, contact information, age, gender, place of residence, and ethnicity
Medical history	Severe cardiovascular and cerebrovascular diseases, severe liver and kidney disease, malignant tumors, acute and chronic infections, diabetes, and autoimmune diseases
Supplement use	Use of any supplements
Family medical history	Cardiovascular or cerebrovascular diseases, severe liver and kidney disease, malignant tumors, diabetes, and autoimmune diseases
Alcohol consumption	History of alcohol consumption, frequency, and type of alcohol consumed
Smoking	Number of cigarettes per day
Physical activity	Current and past participation in physical activities, weekly exercise duration, and intensity
Sleeping habit	Average sleeping hours, sleep quality, history of insomnia, or other sleep disorders

Laboratory results generated during the health check-up are automatically uploaded into the hospital information system and subsequently extracted for statistical analysis, with all data stored in a secure, access-controlled database. Since study initiation, the data collection workflow has proceeded smoothly, indicating high feasibility and acceptability of the study procedures within the health check-up setting. The inclusion and exclusion criteria are mentioned in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

The eligibility criteria for inclusion are as follows:

1. Age ≥18 years old, both male and female.
2. Undergo a health examination in participating hospitals.
3. Agree to participate in the study and sign the informed consent form.
4. BMI <28 kg/m^2 and ≥18.5 kg/m^2 [19].

The criteria for exclusion are as follows:

1. Admission to the hospital as an in-patient during the previous 4 weeks or surgery performed within 6 months.
2. Individuals receiving vitamin supplements or drugs that can affect FSV metabolism or absorption (bile acid sequestrants, statins, antimicrobials, antiepileptic drugs, corticosteroids, etc) [20].
3. Women who are pregnant or breastfeeding.
4. Alcohol consumption was recorded as grams of pure alcohol per week (g/week). Alcohol consumption >140 g/week for males and >70 g/week for females and smoking (>20 cigarettes/day).
5. Previous history of severe cardiovascular and cerebrovascular diseases, severe liver and kidney disease, malignant tumors, acute and chronic infections, diabetes, and autoimmune diseases.
6. Individuals with glomerular filtration rate, alanine transaminase, gamma-glutamyl transferase, albumin, fasting glucose, hemoglobin A1c, uric acid, total cholesterol, triglyceride, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, and platelet count values falling outside the common reference [21].
7. Subtract replicates for the same individual (only the first value was considered).

XSL•FO
RenderX

Data Quality Control

To ensure reliability, consistency, and accuracy of the study data, a multilevel data quality control (QC) strategy will be implemented, covering questionnaire surveys, sample collection and processing, laboratory testing, and data management.

Questionnaire Data QC

Training and Standardization

All research staff involved in the survey will receive standardized training to ensure consistent interpretation and uniform administration of the questionnaire.

Recall Bias Verification

To assess and control recall bias, 100 participants will be randomly selected to complete the questionnaire a second time, approximately 4 weeks after the initial administration. The second questionnaire will be identical to the first. To ensure independence, the follow-up survey will be conducted either through a self-administered electronic questionnaire or a telephone interview conducted by a different research assistant. Consistency of key variables (eg, supplement use and smoking or alcohol history) will be evaluated using the kappa coefficient or intraclass correlation coefficient (ICC). Kappa or ICC values >0.75 will be interpreted as good reliability, 0.60 - 0.75 as moderate, and <0.60 as poor.

Logical Checks

During data entry, logical skip patterns and range checks will be implemented to automatically flag inconsistent or contradictory responses, which will then be reviewed and verified by research staff.

Electronic Data Collection

Electronic questionnaire systems will be used whenever possible to minimize manual data entry errors.

Standardization of Sample Collection, Transport, and Storage

Unified Standard Operating Procedures

All participating centers will follow the same standard operating procedures for sample collection, processing, storage, and transport.

Sample Labeling and Tracking

A barcode-based system will be used to label and track samples, ensuring full traceability from collection to analysis.

Transport Monitoring

Temperature monitoring devices will be used to record temperature conditions during transport, ensuring that samples remain within acceptable ranges.

Laboratory Testing QC

Methodological Standardization

All participating laboratories will use the same liquid chromatography-tandem mass spectrometry (LC-MS/MS) platform, reagent brands, and testing protocols.

Internal QC

Each analytical batch will include commercial QC materials at low, medium, and high concentrations, as well as pooled human serum controls. Results must meet predefined acceptance criteria within ±2 SD.

External Quality Assessment

Laboratories will regularly participate in national or international proficiency testing programs for FSVs to ensure analytical accuracy and interlaboratory comparability.

Central Laboratory Retesting

Each month, 2% (2/100) of samples from each participating center will be randomly selected and sent to the central laboratory for repeat testing. Intercenter consistency will be assessed using Bland-Altman analysis or ICCs, with corrective actions implemented if necessary.

Data Management QC

Dual Independent Data Entry

Paper-based questionnaire data will be entered electronically by 2 independent operators. Any discrepancies will be resolved by a third reviewer with reference to the original records.

Data Cleaning and Review

Regular data cleaning procedures will be performed, including outlier detection, identification of missing values, and logical consistency checks.

Preanalysis Data Verification

Before statistical analysis, a statistician will review the final dataset for completeness, distribution characteristics, and potential sources of biases.

Documentation of QC Activities

All QC procedures, results, and records related to exception handling will be documented in detail and archived to ensure methodological transparency and to support future audits and manuscript preparation.

FSVs Quantification

LC-MS/MS-Based Quantification of FSVs

This study will employ LC-MS/MS to quantitatively analyze FSVs in human blood samples. LC-MS/MS is widely utilized for the quantitative analysis of small molecules in complex biological matrices, such as blood, plasma, and urine, due to its high sensitivity and specificity [22,23]. For vitamin A analysis, recent studies have developed and validated LC-MS/MS methods for the simultaneous quantification of retinol and its metabolites in human serum, demonstrating high sensitivity and specificity suitable for clinical applications [24]. In the analysis of vitamin E, validated LC-MS/MS assays allow simultaneous determination of tocopherols and their oxidative metabolites in plasma and serum, providing excellent sensitivity, minimal sample preparation, and high-throughput capabilities [25-27]. For vitamin D quantification, LC-MS/MS has been widely applied to accurately measure serum 25-hydroxyvitamin D_2 and 25-hydroxyvitamin D_3 concentrations, offering superior

XSL•FO
RenderX

specificity and precision compared to traditional immunoassays [28]. Furthermore, recent advancements, such as the development of derivatization techniques, have enabled the simultaneous quantification of multiple vitamin D metabolites with enhanced sensitivity, particularly in patients with diabetes and hyperlipidemia [29]. Additionally, recent LC-MS/MS protocols have been developed for the determination of different vitamin K forms in human plasma, which are essential for clinical evaluation [30]. In this study, the analytical method is based on established protocols with minor modifications to support high-throughput analysis. Briefly, samples will undergo protein precipitation followed by liquid-liquid extraction using organic solvents. Chromatographic separation will be performed on a reversed-phase C18 column, and detection will be carried out in multiple reaction monitoring mode. Quantification will be achieved using calibration curves constructed with isotope-labeled internal standards for each analyte, allowing correction for matrix effects and variability in extraction recovery.

Method Validation Parameters

Prior to study initiation, the LC-MS/MS method was fully validated in the central laboratory in accordance with International Council for Harmonisation (ICH) M10 guidelines [31] to ensure reliability. The key validation parameters are summarized below.

Precision

Intraday precision (repeatability) and interday precision (intermediate precision) were assessed by analyzing QC samples at 3 concentration levels (low, medium, and high) across multiple runs. Coefficients of variation were ≤8% for intraday precision and ≤12% for interday precision for all analytes.

Accuracy

Accuracy was assessed using spike-and-recovery experiments and by analyzing certified reference materials, where available. Mean recovery rates ranged from 92% to 108% across the validated concentration ranges for all vitamins.

Linearity

The method demonstrated linearity over clinically relevant concentration ranges (eg, vitamin D: 5 - 250 nmol/L) with ICCs (R^2)>0.995 for all calibration curves.

Limits of Detection and Limits of Quantification

The limits of detection and limits of quantification (LOQ) were determined based on signal-to-noise ratios of 3:1 and 10:1, respectively. The LOQs were established below the lower limit of the expected physiological ranges for each vitamin.

Specificity and Selectivity

No significant interference from endogenous compounds or commonly used medications was observed at the retention times of the target analytes or their corresponding internal standards.

Intersite Harmonization and Quality Assurance

A rigorous, multilayered strategy is implemented to ensure the comparability of results across all 20 participating hospitals.

Standardization of Preanalytical and Analytical Procedures

All sites adhere to identical, detailed standard operating procedures for sample collection (fasting status, tube type), processing (centrifugation speed and duration), storage (−80 °C), and shipment (on dry ice with continuous temperature monitoring).

Unified Analytical Platform and QC Materials

All participating laboratories use the same model of LC-MS/MS instrument, identical reagent suppliers, and a standardized analytical protocol. A common batch of commercial, multilevel QC materials (low, medium, and high) is provided to each site. Additionally, a centrally prepared pooled human serum QC sample is distributed monthly.

Mandatory QC Runs

Each analytical batch must include all 3 levels of QC samples. A batch is accepted only if all QC results fall within predefined limits (±2 SD of the established mean).

External Quality Assessment

All laboratories are required to participate and successfully complete relevant national or international external quality assessment programs on a biannual basis.

Central Laboratory Retesting Program

On a monthly basis, 2% of samples from each center, randomly selected, are shipped to the central coordinating laboratory for blind reanalysis. Intercenter agreement is assessed using Bland-Altman analysis and ICC.

Data Harmonization and Corrective Action

An ICC ≥0.90 is considered indicative of excellent agreement. If systematic bias is identified, an investigation is triggered, and corrective actions (such as instrument recalibration or personnel retraining) are implemented before the site resumes sample analysis. This process ensures continuous longitudinal harmonization across the study network.

Sample Size Estimation

According to the Chinese Health Industry Standard for Establishing RIs for Clinical Laboratory Test Items [32], to ensure the reliability of the results, this study plans to recruit 100,000 participants to establish RIs for FSVs A, D, E, and K in the adult Chinese population. The sample size estimation for establishing RIs was conducted based on a multicenter, cross-sectional study design. The goal was to recruit a total of 100,000 adult individuals from the physical examination cohorts of 20 participating hospitals across China. Based on existing literature and pilot data, it was estimated that approximately 10% of the screened population would meet the predefined health criteria for RI establishment. Therefore, the final analytical cohort was projected to include approximately 10,000 eligible and healthy adults.

The stratification strategy prioritized core biological variables. Participants were stratified by sex (2 strata) and age (5 strata: 18 - 30, 30 - 40, 40 - 50, 50 - 60, and >60 y), resulting in 10 primary strata. With an expected sample of 10,000 healthy

XSL•FO
RenderX

individuals, the average sample size per stratum would be 1000, which substantially exceeds the Clinical and Laboratory Standards Institute (CLSI) EP28-A3 recommendation of at least 120 subjects per stratum for nonparametric RI estimation [33]. This ensures sufficient power and precision for estimating percentile-based RIs.

Environmental factors, including geographical region (North/South) and season (spring, summer, autumn, or winter), were not included as independent stratification variables during initial sample size calculation to avoid excessive fragmentation and insufficient sample sizes within multiple substrata. Instead, these variables will be evaluated as covariates during statistical analysis (eg, using analysis of covariance or quantile regression). If region or season demonstrates a statistically and clinically meaningful influence on vitamin concentrations, separate RIs will be considered. In that scenario, post hoc stratification (eg, region×sex×age=20 strata) would still result in an average of approximately 500 individuals per stratum, exceeding the CLSI minimum requirement.

This sampling framework ensures that the study is adequately powered to establish precise sex- and age-specific RIs for the Chinese population, while maintaining sufficient statistical robustness to explore potential effects of region and season.

Statistical Data Analysis

Descriptive and Comparative Analyses

All statistical analyses will be performed using SPSS (version 28.0.1.1), GraphPad Prism (version 9.5.0), and R software (version 4.2.1). Continuous variables will be assessed for normality using the Kolmogorov-Smirnov test and visual inspection of Q-Q plots. Normally distributed data will be presented as mean (SD) and nonnormally distributed data as median (IQR). Categorical variables will be expressed as frequencies and percentages. Group comparisons for continuous variables will be conducted using independent-samples *t* tests for normally distributed data or Mann-Whitney *U* tests for nonnormally distributed data. For comparisons involving more than 2 groups, one-way ANOVA or Kruskal-Wallis tests will be applied, as appropriate. Comparisons of categorical variables will be performed using the chi-square test or Fisher exact test. To address multiple testing in planned subgroup analyses (eg, by age strata, region, and season), the false discovery rate will be controlled using the Benjamini-Hochberg procedure where applicable, with a significance threshold of $P<.05$.

RI Estimation Procedure

Overview

The establishment of RIs will follow a standardized, multistep process. A conservative consensus approach will be applied using 3 independent methods—Tukey boxplot method (1.5× IQR rule), the standard *z* score ($|z|>3$), and the modified *z* score ($|M|>3.5$). A data point will be excluded only if identified as an outlier by at least 2 of these 3 methods.

Distribution Assessment and Method Selection

For each vitamin within each predefined or partitioned subgroup, normality of the cleaned data will be reassessed. If the data follow a normal distribution, the central 95% RI will be calculated using a parametric approach (mean ±1.96× SD). If the data are nonnormally distributed, a nonparametric approach will be used, with the 2.5th and 97.5th percentiles directly defining the lower and upper reference limits.

CI Estimation for Reference Limits

To assess the precision of the estimated reference limits, 90% CIs will be calculated for both the lower (2.5th percentile) and upper (97.5th percentile) limits. For nonparametric estimates, bootstrap resampling with 1000 iterations will be used to generate percentile-based CIs. For parametric estimates, standard formulas based on the standard errors of the mean and SD will be applied. Reporting these 90% CIs provides an important measure of the reliability of the established RIs.

Rounding and Final Presentation

The final reference limits and their corresponding CIs will be rounded according to clinically relevant units (eg, to the nearest 0.1 nmol/L for vitamin D) to produce the final, publishable RIs.

Subgroup Analysis

To determine whether separate RIs should be established for different subgroups, we will systematically evaluate the influence of 4 key factors: age, sex, geographic region, and season. The analysis will proceed in 2 sequential stages.

Stage 1—Assessment of Primary Stratification Factors (Age and Sex)

As age and sex are predefined stratification variables in the sampling design, we will first assess the necessity of partitioning RIs based on these factors. Vitamin concentrations will be compared across 5 age strata (18 - 30, 30 - 40, 40 - 50, 50 - 60, and >60 y) and between sexes using the Kruskal-Wallis test or ANOVA, as appropriate. If statistically significant differences ($P<.05$) with a standardized effect size greater than 0.30 are observed, age- and/or sex-specific RIs will be established.

Stage 2—Exploration of Environmental and Contextual Factors (Region and Season)

For factors not included in the initial stratification—geographic region (North vs South) and season (spring, summer, autumn, and winter)—a nested modeling approach will be used to evaluate their independent effects after accounting for age and sex. A nested ANOVA model will be constructed with FSV concentration as the dependent variable, age group and sex as fixed effects, and study center treated as a random effect nested within geographic region. A separate model will be used to assess the effect of season. If region or season explains a statistically significant proportion of variance ($P<.05$), its clinical relevance will be further evaluated by calculating age- and sex-adjusted standardized differences between groups.

The decision to establish separate RIs for region or season will follow a predefined 2-step criterion: (1) statistical significance ($P<.05$) and (2) clinical relevance, defined as a standardized difference greater than 0.30, indicating a small-to-moderate effect size.

XSL•FO
RenderX

Quantile Regression for RI Estimation

For any subgroup requiring separate RIs, the 2.5th and 97.5th percentiles will be estimated using quantile regression. This approach enables direct estimation of reference limits while adjusting for relevant covariates (eg, age and sex when establishing region-specific RIs), thereby improving the stability and precision of the estimates, even in subgroups with relatively smaller sample sizes.

Robustness Assurance

The study is adequately powered to support this level of partitioning. The planned sample size ensures that, even if RIs are stratified simultaneously by age, sex, region, and season, each resulting subgroup will include approximately 500 individuals, which exceeds the CLSI EP28-A3c minimum requirement of 120. If any partitioned subgroup fails to meet the predefined sample size threshold or does not satisfy the clinical relevance criterion (standardized difference ≤0.30), it will be merged with the most adjacent or clinically relevant group (eg, neighboring age strata or combined seasons) to preserve statistical robustness and clinical use. For all reported RIs, 90% CIs for the reference limits will be provided to ensure transparency regarding estimate precision.

Ethical Considerations

This study complies with the Declaration of Helsinki and obtained ethical approval from the Clinical Research Ethics Committee of the First Affiliated Hospital, Zhejiang University School of Medicine (2024 - 0728). Written informed consent will be obtained from all participants prior to enrollment. Participants' privacy and confidentiality will be strictly protected; all personal data will be deidentified, securely stored, and be accessible only to authorized research personnel. This is an observational study and does not involve any treatment; therefore, participants will not receive any financial compensation. The findings of this study will be disseminated through conference presentations and peer-reviewed publications. Findings will be disseminated through publication in a peer-reviewed journal and presentations at conferences.

Future Plans

On completion of subject recruitment, we will proceed with the systematic collection of relevant data, including (1) standardized questionnaires, (2) comprehensive physical examinations, and (3) blood samples for laboratory analysis. RIs will be established using parametric methods for normally distributed data and nonparametric methods otherwise, with values appropriately rounded. Nested ANOVA will be applied to assess the need for partitioning by gender, age, season, and region; groups without significant differences will be combined. To address potential sampling bias arising from the hospital-based health check-up population, future studies will prioritize validating the established RIs in external cohorts from more diverse socioeconomic and rural populations. Additionally, we plan to conduct longitudinal follow-up studies to track changes in FSV levels over time and examine their relationships with clinical outcomes.

Results

As of October 2025, 13,545 adults have been enrolled, and 1690 participants have met the inclusion criteria. Recruitment began on 1 July 2024 and is expected to conclude by 30 June 2026. Data analysis is scheduled to begin on July 2026 and to complete by no later than December 2026. The projected end date of this study is December 2026, and results are expected to be published in May 2027.

Discussion

This multicenter, cross-sectional study is designed to establish population-specific RIs for FSVs in a healthy Chinese adult population. Prior research on FSVs in China has been limited in scope and population coverage. Existing studies have typically focused on specific subgroups rather than establishing comprehensive national RIs: Yin et al [13] reported RIs for vitamins A and E in Chinese older adults ≥64 years, based on LC-MS/MS measurements; Nie et al [34] established vitamin K–related reference values in healthy women of childbearing age; and Yu et al [35] examined seasonal and age-specific vitamin D RIs in children. Although these studies contribute foundational data, they are constrained by age range, population specificity, or single-center design. In contrast, Western research has begun generating broader adult reference datasets. In European adults, Rigo-Bonnin et al [36] derived RIs for vitamins A and E using indirect data-mining approaches from over 2000 samples, and Caballero et al [37] later verified updated RIs using a direct approach in adults aged 18 - 65 years. For vitamin D, population-based surveillance in Germany showed mean 25(OH)D levels around 45.6 nmol/L and a high prevalence of insufficient status [38]. No study has established comprehensive RIs across multiple FSVs in a large, diverse adult Chinese population, highlighting the need and significance of the current multicenter study. The RIs derived from this study are expected to serve as national laboratory benchmarks for healthy adults. The findings will provide valuable insights for clinicians to improve the assessment and management of FSVs and their related diseases.

The generalizability of the established RIs requires careful consideration because FSV concentrations are influenced by multiple population-level factors that vary across China. Dietary patterns, sunlight exposure, supplement use, and socioeconomic conditions differ substantially between urban and rural regions, and these differences may shift baseline vitamin distributions in ways not fully captured by our predominantly urban, health-check cohort. As a result, the current RIs may best reflect individuals with relatively stable health status and higher health awareness, rather than the broader nutritional and environmental diversity of the national population. In settings involving rural or nutritionally at-risk populations, these RIs should be interpreted with consideration of population-specific differences. Future validation in geographically and socioeconomically diverse cohorts will be essential to ensure that these RIs accurately represent the full spectrum of biological variability across the Chinese adult population.

XSL•FO
RenderX

The study has several strengths. First, the large-scale, multicenter design ensures that the sample is representative of the diverse demographic and geographic characteristics of China. This will allow the study to capture the regional variability in aging, lifestyle, and environmental exposures that influence FSV levels. Second, the standardized data collection process, which includes detailed questionnaires, physical examinations, and calibrated laboratory instruments, ensures the reliability and consistency of the collected data. Third, the inclusion of comprehensive physical and biochemical assessments, such as metabolic panels, lipid panels, and liver function tests, provides a comprehensive dataset for evaluating the factors that may influence FSV levels. Despite these strengths, the study has several limitations. First, this study uses a cross-sectional design with a single time-point measurement of FSV levels. Although nested ANOVA is applied to assess whether seasonal or regional partitioning of RIs is warranted, the design cannot capture within-individual variation in FSV levels related to seasonal dietary patterns, lifestyle fluctuations, or metabolic adaptation. As a result, the derived RIs may reflect average population distributions rather than the full physiologic variability experienced across the year. Second, the reliance on self-reported data for the questionnaire components introduces the potential for recall bias, particularly for lifestyle-related variables that influence vitamin status. A validation or reliability check was not performed, thus misclassification may have occurred and could affect subgroup comparisons. Third, our strict exclusion criteria likely yielded a “super healthy” cohort, potentially narrowing the estimated RIs compared with those from a more heterogeneous population. Therefore, caution is needed when applying these RIs in general clinical practice, and validation in more diverse cohorts is warranted. Finally, the study population consists of individuals undergoing routine health check-ups in economically developed urban regions. These individuals generally have better access to health care and may maintain healthier nutritional status, which may limit the generalizability of the derived RIs to adults from rural or lower-income areas. Nevertheless, this cohort reflects the population most likely to receive preventive laboratory testing in current clinical practice, making the resulting RIs highly relevant for routine use. Future comparative research involving more diverse geographic and socioeconomic groups, including relatively impoverished areas, is needed to validate the findings and enhance the generalizability.

This study holds critical importance for public health and clinical practice in China. Existing RIs for FSVs are mostly based on data from Western populations, which may not fully reflect the dietary habits, genetic traits, or environmental exposures of Chinese individuals. By establishing population-specific RIs, this research will improve the diagnostic accuracy of FSV assessments, enabling clinicians to better identify and manage both deficiencies and toxicities. Furthermore, the comprehensive data collected in this study will support the development of evidence-based public health strategies to prevent malnutrition-related health conditions. These findings are expected to contribute significantly to advancing precision nutrition and improving health outcomes for the Chinese population.

Acknowledgments

The authors thank the entire department of health management for supporting this research.

Data Availability

The datasets generated and analyzed during this study contain identifiable clinical information and therefore cannot be publicly shared due to patient privacy and institutional ethical restrictions but are available from the corresponding author upon reasonable request.

Authors' Contributions

ZL and CL designed the research. CL registered the study. CL wrote the draft of the study protocol. NL and CS provided the statistical analysis. LQ and DS performed a critical revision of the manuscript. LQ and CS drafted tables and figures. All authors read and approved the manuscript and agreed to be accountable for all aspects of the research.

Conflicts of Interest

None declared.

References

1. Zhang Y, Bala V, Mao Z, Chhonker YS, Murry DJ. Quantification of fat-soluble vitamins and their metabolites in biological matrices: an updated review. Bioanalysis 2020 May;12(9):625-640. [doi: 10.4155/bio-2020-0069] [Medline: 32412347]
2. Wiseman EM, Bar-El Dadon S, Reifen R. The vicious cycle of vitamin A deficiency: a review. Crit Rev Food Sci Nutr 2017 Nov 22;57(17):3703-3714. [doi: 10.1080/10408398.2016.1160362] [Medline: 27128154]
3. Hombali AS, Solon JA, Venkatesh BT, Nair NS, Peña-Rosas JP. Fortification of staple foods with vitamin A for vitamin A deficiency. Cochrane Database Syst Rev 2019 May 10;5(5):CD010068. [doi: 10.1002/14651858.CD010068.pub2] [Medline: 31074495]

XSL•FO
RenderX

4. Nikolac Gabaj N, Unic A, Miler M, et al. In sickness and in health: pivotal role of vitamin D. Biochem Med (Zagreb) 2020 Jun 15;30(2):202-214. [doi: 10.11613/BM.2020.020501] [Medline: 32550812]
5. Browne D, McGuinness B, Woodside JV, McKay GJ. Vitamin E and Alzheimer's disease: what do we know so far? Clin Interv Aging 2019;14:1303-1317. [doi: 10.2147/CIA.S186760] [Medline: 31409980]
6. de Castro Lobo LM, Hadler MCCM. Vitamin E deficiency in childhood: a narrative review. Nutr Res Rev 2023 Dec;36(2):392-405. [doi: 10.1017/S0954422422000142] [Medline: 35929460]
7. Boegh Andersen I, Brasen CL, Schmedes A, Brandslund I, Madsen JS. In search of normality for vitamin K1: establishing age-dependent reference intervals in the Danish population. J Appl Lab Med 2020 May 1;5(3):531-543. [doi: 10.1093/jalm/jfaa017] [Medline: 32445363]
8. Horn PS, Pesce AJ. Reference intervals: an update. Clin Chim Acta 2003 Aug;334(1-2):5-23. [doi: 10.1016/s0009-8981(03)00133-5] [Medline: 12867273]
9. Man S, Gao Y, Lv J, et al. Establishment of reference intervals of ten commonly used clinical chemistry analytes: a real-world study in China. Biomark Med 2021 Aug;15(11):797-806. [doi: 10.2217/bmm-2021-0233] [Medline: 33955784]
10. Wu X, Zhao M, Pan B, et al. Complete blood count reference intervals for healthy Han Chinese adults. PLoS ONE 2015;10(3):e0119669. [doi: 10.1371/journal.pone.0119669]
11. Giustina A, Lazaretti-Castro M, Martineau AR, Mason RS, Rosen CJ, Schoenmakers I. A view on vitamin D: a pleiotropic factor? Nat Rev Endocrinol 2024 Apr;20(4):202-208. [doi: 10.1038/s41574-023-00942-0] [Medline: 38253860]
12. Ye Q, Zhong Q, Huang G, Zhang W. Establishment of reference range for serum concentration of vitamin A and vitamin E in Southern Sichuan area of China. J Clin Lab Anal 2024 May;38(10):e25074. [doi: 10.1002/jcla.25074] [Medline: 38847175]
13. Yin Y, Wang D, Ma C, et al. Establishing reference intervals for vitamins A and E in Chinese elderly people using liquid chromatography-tandem mass spectrometry. J Clin Lab Anal 2021 Apr;35(4):e23726. [doi: 10.1002/jcla.23726] [Medline: 33559895]
14. Hanson C, Lyden E, Furtado J, et al. Vitamin E status and associations in maternal-infant Dyads in the Midwestern United States. Clin Nutr 2019 Apr;38(2):934-939. [doi: 10.1016/j.clnu.2018.02.003] [Medline: 29496275]
15. Malik A, Eggersdorfer M, Trilok-Kumar G. Vitamin E status in healthy population in Asia: a review of current literature. Int J Vitam Nutr Res 2021 Jun;91(3-4):356-369. [doi: 10.1024/0300-9831/a000590] [Medline: 31124407]
16. Paulus MC, Drent M, Kouw IWK, Balvers MGJ, Bast A, van Zanten ARH. Vitamin K: a potential missing link in critical illness-a scoping review. Crit Care 2024 Jul 1;28(1):212. [doi: 10.1186/s13054-024-05001-2] [Medline: 38956732]
17. Skalny AV, Aschner M, Tsatsakis A, et al. Role of vitamins beyond vitamin D3 in bone health and osteoporosis (review). Int J Mol Med 2024 Jan;53(1):9. [doi: 10.3892/ijmm.2023.5333] [Medline: 38063255]
18. Zhao QY, Li Q, Hasan Rashedi M, Sohouli M, Rohani P, Velu P. The effect of vitamin K supplementation on cardiovascular risk factors: a systematic review and meta-analysis. J Nutr Sci 2024;13:e3. [doi: 10.1017/jns.2023.106] [Medline: 38282652]
19. Zhou BF, Cooperative Meta-Analysis Group of the Working Group on Obesity in China. Predictive values of body mass index and waist circumference for risk factors of certain related diseases in Chinese adults--study on optimal cut-off points of body mass index and waist circumference in Chinese adults. Biomed Environ Sci 2002 Mar;15(1):83-96. [Medline: 12046553]
20. Robien K, Oppeneer SJ, Kelly JA, Hamilton-Reeves JM. Drug-vitamin D interactions: a systematic review of the literature. Nutr Clin Pract 2013 Apr;28(2):194-208. [doi: 10.1177/0884533612467824] [Medline: 23307906]
21. Miyamoto H, Kawakami D, Hanafusa N, et al. Determination of a serum 25-hydroxyvitamin D reference ranges in Japanese adults using fully automated liquid chromatography–tandem mass spectrometry. J Nutr 2023 Apr;153(4):1253-1264. [doi: 10.1016/j.tjnut.2023.01.036] [Medline: 36806449]
22. Van Eeckhaut A, Lanckmans K, Sarre S, Smolders I, Michotte Y. Validation of bioanalytical LC-MS/MS assays: evaluation of matrix effects. J Chromatogr B Analyt Technol Biomed Life Sci 2009 Aug 1;877(23):2198-2207. [doi: 10.1016/j.jchromb.2009.01.003] [Medline: 19179125]
23. Xu RN, Polzin J, Kranz M, et al. Strategies for developing sensitive and automated LC-MS/MS assays of a pharmaceutical compound and its metabolite from whole blood matrix. Pharmaceutics 2010 Apr 30;2(2):159-170. [doi: 10.3390/pharmaceutics2020159] [Medline: 27721348]
24. Jia Y, Liu X, Liu C, Chen Y, Li W. Simultaneous determination of Vitamins A and E and their generated metabolites in human Serum by LC-MS/MS. Biomed Chromatogr 2025 Jan;39(1):e6064. [doi: 10.1002/bmc.6064] [Medline: 39732515]
25. Torquato P, Ripa O, Giusepponi D, et al. Analytical strategies to assess the functional metabolome of vitamin E. J Pharm Biomed Anal 2016 May 30;124:399-412. [doi: 10.1016/j.jpba.2016.01.056] [Medline: 26947319]
26. Giusepponi D, Torquato P, Bartolini D, et al. Determination of tocopherols and their metabolites by liquid-chromatography coupled with tandem mass spectrometry in human plasma and serum. Talanta 2017 Aug 1;170:552-561. [doi: 10.1016/j.talanta.2017.04.030] [Medline: 28501210]
27. Giusepponi D, Galarini R, Barola C, et al. LC-MS/MS assay for the simultaneous determination of tocopherols, polyunsaturated fatty acids and their metabolites in human plasma and serum. Free Radic Biol Med 2019 Nov 20;144:134-143. [doi: 10.1016/j.freeradbiomed.2019.04.017] [Medline: 31009660]

XSL•FO
RenderX

28. Saenger AK, Laha TJ, Bremner DE, Sadrzadeh SMH. Quantification of serum 25-hydroxyvitamin D(2) and D(3) using HPLC-tandem mass spectrometry and examination of reference intervals for diagnosis of vitamin D deficiency. Am J Clin Pathol 2006 Jun;125(6):914-920. [doi: 10.1309/J32U-F7GT-QPWN-25AP] [Medline: 16690491]
29. Wang X, Qin Q, Li F, Fu Y, Liu N. A novel LC-MS/MS method combined with derivatization for simultaneous quantification of vitamin D metabolites in human serum with diabetes as well as hyperlipidemia. RSC Adv 2023 Nov 16;13(48):34157-34166. [doi: 10.1039/D3RA05700C]
30. Fusaro M, Gallieni M, Rizzo MA, et al. Vitamin K plasma levels determination in human health. Clin Chem Lab Med 2017 May 1;55(6):789-799. [doi: 10.1515/cclm-2016-0783] [Medline: 27732556]
31. ICH harmonised guideline M10: bioanalytical method validation and study sample analysis. : International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH); 2022 URL: https://database.ich.org/sites/default/files/M10_Guideline_Step4_2022_0524.pdf [accessed 2026-01-20]
32. WS/t 402-2024: defining, establishing, and verifying reference intervals of quantitative analytes in the clinical laboratory. China National Health Commission. 2024. URL: https://www.nhc.gov.cn/wjw/s9492/202407/c69784a165424889bedb36d854e61554/files/1739781687925_45929.pdf [accessed 2026-01-20]
33. EP28-a3c defining, establishing, and verifying reference intervals in the clinical laboratory; approved guideline—third edition. : Clinical and Laboratory Standards Institute; 2010 URL: https://img.antpedia.com/standard/files/pdfs_ora/20230614/CLSI/CLSI%20EP28-A3C-2010.pdf [accessed 2026-01-20]
34. Nie S, Yang L, Feng J, et al. Reference range of vitamin K evaluating indicators in Chinese childbearing women. Nutrients 2023 Apr 19;15(8):1977. [doi: 10.3390/nu15081977] [Medline: 37111196]
35. Yu J, He X, Huang S. The establishment of serum 25-hydroxyvitamin D reference intervals in children aged 0-14 years in Zigong area, China. J Physiol Anthropol 2021 Oct 6;40(1):14. [doi: 10.1186/s40101-021-00265-x] [Medline: 34615548]
36. Rigo-Bonnin R, Aliart-Fernández I, Escalante-Vilanova A, Brunet M, Parra-Robert M, Morales-Ruiz M. Calculation of reference intervals for the concentrations of α-tocopherol and retinol in serum using indirect data-mining procedures. Clin Chim Acta 2024 Jul 15;561:119822. [doi: 10.1016/j.cca.2024.119822] [Medline: 38908772]
37. Caballero A, Gonzalez-Silva G, Gabriel-Medina P, et al. Optimizing vitamin supplementation via reference interval update of vitamins A, E, B1, and B6 using HPLC. J Clin Biochem Nutr 2025 Mar;76(2):148-155. [doi: 10.3164/jcbn.24-155] [Medline: 40151408]
38. Rabenberg M, Scheidt-Nave C, Busch MA, Rieckmann N, Hintzpeter B, Mensink GBM. Vitamin D status among adults in Germany–results from the German health interview and examination survey for adults (DEGS1). BMC Public Health 2015 Jul 11;15:641. [doi: 10.1186/s12889-015-2016-7] [Medline: 26162848]

Abbreviations

CLSI: Clinical and Laboratory Standards Institute
FSV: fat-soluble vitamin
ICC: intraclass correlation coefficient
LC-MS/MS: liquid chromatography-tandem mass spectrometry
LOQ: limits of quantification
QC: quality control
RI: reference interval

Edited by J Sarvestan; submitted 10.Jul.2025; peer-reviewed by S Bakrim; revised version received 06.Jan.2026; accepted 07.Jan.2026; published 17.Feb.2026.

Please cite as:
Liu Z, Liu C, Li N, Shen C, Qiu L, Sheng D
Establishing Reference Intervals for Fat-Soluble Vitamins in Healthy Chinese Adults: Study Protocol for a Cross-Sectional, Multicenter Study
JMIR Res Protoc 2026;15:e80425
URL: https://www.researchprotocols.org/2026/1/e80425
doi:10.2196/80425

© Zhong Liu, Chenbing Liu, Nan Li, Chao Shen, Lihong Qiu, Di Sheng. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

In Vitro Comparison of the Dimensional Accuracy of Implant Impressions Using Custom Acrylic Trays and Prefabricated Self-Perforating Trays: Protocol for a Comparative Evaluation Study

Arti Agrawal[1], BDS; Sharayu Nimonkar[1], MDS, PhD; Surekha Godbole[1], MDS, PhD; Rushabh Parakh[2], BDS; Vikram Belkhode[1], MDS, PhD; Namita Zilpilwar[1], BDS

[1]Department of Prosthodontics and Crown & Bridge, Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, 202, SPDC building, DMIHER, Sawangi (Meghe), Wardha, Maharashtra, India

[2]Bharati Vidyapeeth Deemed University, Sangli, Maharashtra, India

Corresponding Author:
Arti Agrawal, BDS
Department of Prosthodontics and Crown & Bridge, Sharad Pawar Dental College, Datta Meghe Institute of Higher Education and Research, 202, SPDC building, DMIHER, Sawangi (Meghe), Wardha, Maharashtra, India

Abstract

Background: The accurate transfer of implant positions from the patient's mouth to the laboratory is crucial for the successful fabrication of prostheses in implant-based prosthodontics. Making an implant impression is a crucial part of this procedure as it replicates the intraoral implant position and transfers it to the cast to be used in the laboratory for fabricating a passive fit prosthesis. The purpose of this study is to compare and assess the accuracy of implant impressions made with prefabricated self-perforating trays and custom acrylic trays.

Objective: The primary objective is to compare the dimensional accuracy of implant impressions made using custom acrylic trays versus prefabricated self-perforating trays in an in vitro setting. The secondary objective is to determine whether the impressions made with prefabricated self-perforating trays achieve clinically acceptable accuracy.

Methods: A partially edentulous mandibular model with implants placed at the second premolar and first molar region will function as a master model. Impressions will be made using the custom acrylic tray and prefabricated self-perforating implant impression tray. The cast obtained will be scanned, and an STL file will be generated to evaluate the accuracy of implant positions using HyperMesh software. Statistical analysis will be done with IBM SPSS Statistics at a 95% CI and 80% power.

Results: The study aims to evaluate whether there is a statistically significant difference in the accuracy of implant casts obtained using prefabricated self-perforating trays versus custom acrylic trays. As of July 2025, the models have been prepared for further analysis. It is projected that data collection will be completed in May 2026, with results to be published in early 2027.

Conclusions: For the success of implant dentistry, the precise fit of the final prosthesis has become a necessity. A mismatched framework may overload the implant, endangering its durability. To achieve a precise fit of the implant framework, a precise definitive cast is necessary, which depends on several factors, one of them being the implant impression tray selection. Traditionally, custom acrylic trays have been used. However, they are impractical to use in routine clinical practice due to certain disadvantages. A newer range of trays for the open tray impression technique is being marketed, and one would be able to reduce the additional expense and laboratory time associated with using custom acrylic trays by using this newer variety of self-perforating implant impression tray.

International Registered Report Identifier (IRRID): PRR1-10.2196/79452

(JMIR Res Protoc 2026;15:e79452) doi:10.2196/79452

KEYWORDS

implant impressions; custom-made trays; prefabricated trays; dimensional accuracy; prosthodontics; dentistry

XSL•FO
RenderX

Introduction

Background

The accurate transfer of implant positions from the oral environment to a working cast is paramount for the long-term success of implant-supported prostheses. Any discrepancies in this transfer can lead to misfits, resulting in mechanical complications such as screw loosening, screw fractures, and biological issues like peri-implantitis. Among the various factors influencing the accuracy of implant impressions, the choice of impression technique and tray type plays a significant role [1].

The open tray (direct) impression technique is the gold standard for implant impressions. This technique involves the use of impression copings that protrude through openings in the tray, allowing them to be unscrewed and removed along with the impression, thereby minimizing repositioning errors [1]. In contrast, the closed tray (indirect) technique requires the removal and repositioning of impression copings, which can introduce inaccuracies, particularly in complex implant scenarios [2].

Custom trays, typically fabricated from autopolymerizing acrylic resin, are designed to fit an individual patient's anatomy, ensuring uniform thickness of the impression material and enhanced rigidity. These characteristics contribute to the high accuracy of impressions made using custom trays. However, the fabrication process is time-consuming and may not be feasible in all clinical situations [3].

Prefabricated self-perforating trays can be used as an alternative to these custom acrylic trays, as they do not require laboratory fabrication and are readily available. These trays consist of a thin film on the occlusal surface that is perforated by the transfer during tray positioning. This helps in achieving a clean and precise impression without excessive impression material. Although they offer practical advantages, concerns remain regarding their stability, their adaptation to the arch form, and the resulting accuracy of the impressions.

Previous in vitro studies have explored the accuracy of various impression techniques and tray types. Patil et al [4] evaluated the precision of different splinting materials in open tray impressions and found that all tested materials produced casts that had measurements very similar to the reference model, with prefabricated pattern resin bars showing the highest accuracy. However, there is a paucity of data specifically comparing the accuracy of impressions made using custom acrylic trays versus prefabricated self-perforating trays.

Therefore, we designed this in vitro study that aims to perform a comparative evaluation of the accuracy of implant impressions made using custom acrylic trays and prefabricated self-perforating trays. By controlling variables such as impression material, implant position, and environmental conditions, this study seeks to provide objective data to guide clinical decision-making regarding tray selection in implant prosthodontics.

Aim

The primary aim is to evaluate the accuracy of implant impressions made with custom acrylic trays and prefabricated self-perforating trays.

The secondary aim is to assess if there is a clinically acceptable accuracy of impressions made by prefabricated self-perforating trays compared to the master model.

Methods

Study Design

The in vitro study will be conducted at the Department of Prosthodontics and Crown & Bridge of the Datta Meghe Institute of Medical Sciences.

Sample Size Calculation

The sample size was calculated using OpenEpi software (version 3.0) for comparison of two independent means at a 95% confidence level and 80% power, based on previously published data. A minimum of 8 samples per group was determined, resulting in a total sample size of 16 (Group 1: mean 73.31, SD 26.01; variance: 676.52; Group 2: mean 154.41, SD 74.64; variance: 5571.13).

This was done with reference to the study conducted by Bohner et al [5].

The formula for the same is:

$$N=(\sigma 12+\sigma 22/\kappa)(Z1-\alpha/2+Z1-\beta)2\Delta 2$$

The notations for the formula are:

N=sample size

σ1=standard deviation of Group 1

σ2=standard deviation of Group 2

Δ=difference in group means

κ=ratio=1

Z1 – α /2 = two-sided Z value (eg, Z=1.96 for 95% CI).

Z1 – β = power

$N=(26.01 \times 26.01) + (74.64 \times 74.64)(1.96+0.84)^2 / (-81.1 \times -81.1)=8$

The total sample size is 16, with 8 in each group.

The study procedure is shown in Figure 1.

XSL•FO
RenderX

Figure 1. Schematic representation of the study procedure.

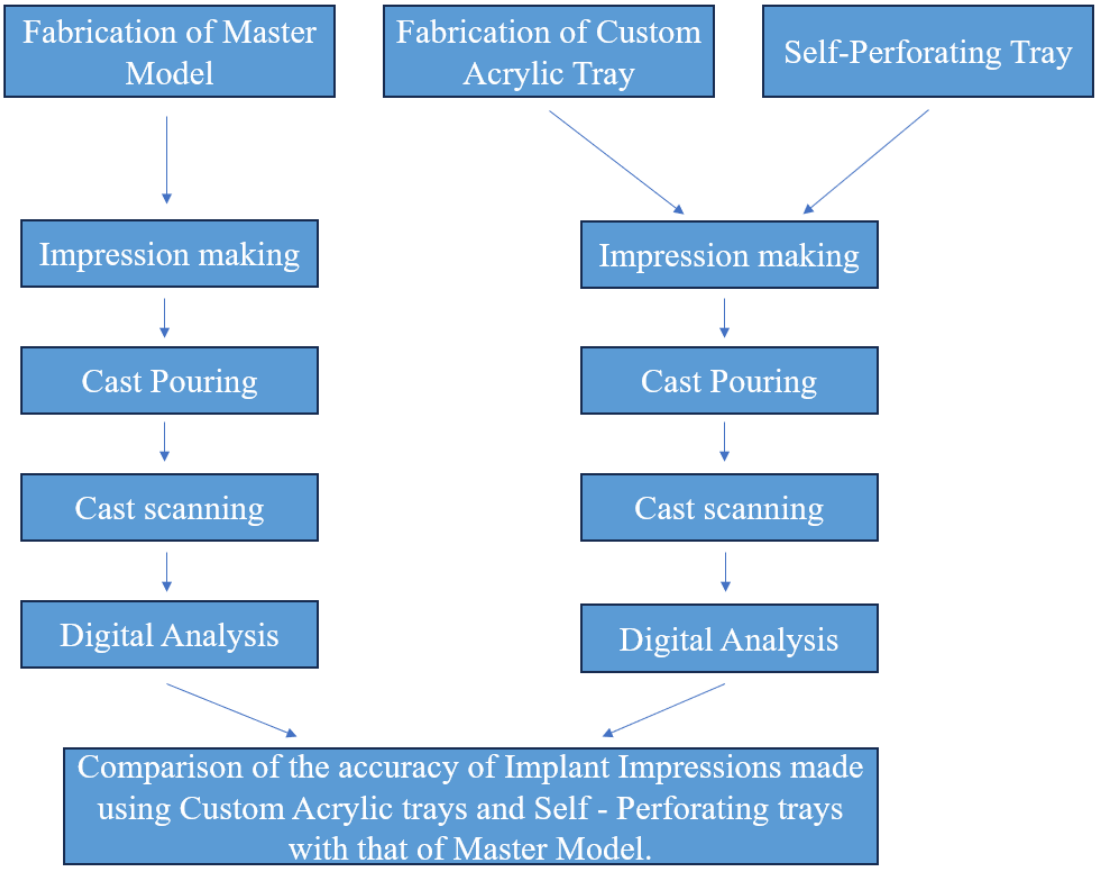

Fabrication of the Master Model

A heat-cured acrylic resin model will be fabricated as the master model. This master model fabrication will be done by first creating a wax-up of the partially edentulous arch and positioning implant analogs using a dental surveyor. The assembly will be flasked, the wax will be boiled out, and heat-cured autopolymerizing polymethyl methacrylate (PMMA) resin will be packed into the mold. After curing in a water bath, the model will be deflasked, finished, and polished. This model will accurately replicate implant positions and serve as the reference for evaluating impression accuracy. Two implants will be placed parallel to each other in the second premolar and first molar region (Figure 2).

XSL•FO
RenderX

Figure 2. Fabrication of the master model.

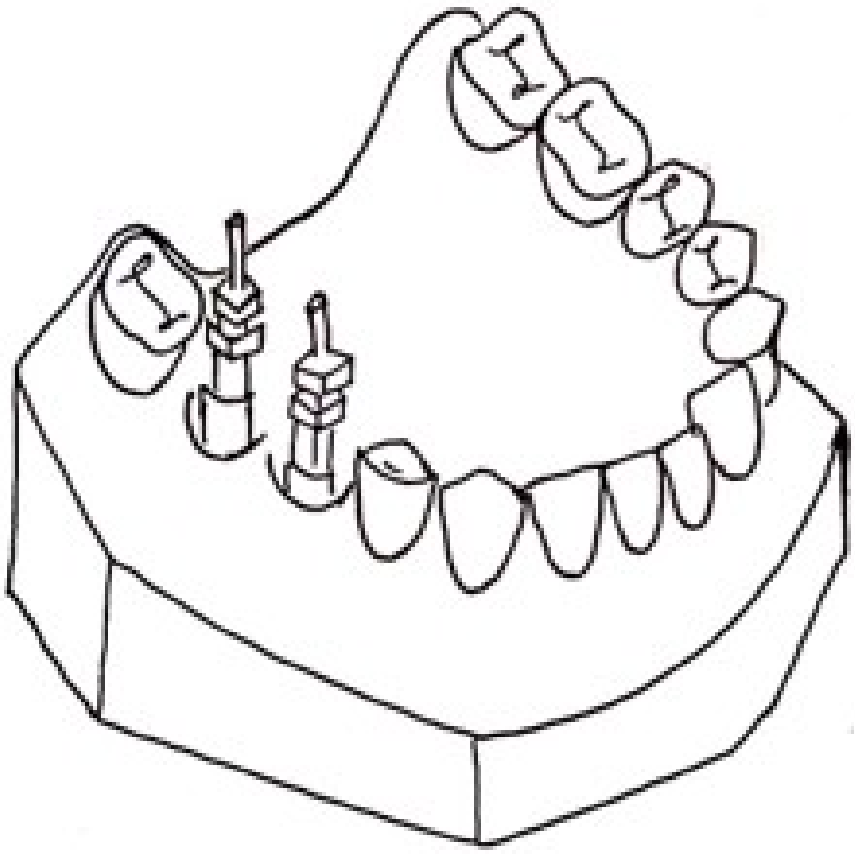

Fabrication of Custom Acrylic Trays

To fabricate the custom acrylic trays from PMMA, which offers high rigidity and dimensional stability, a standardized procedure will be followed.

A denture flask will be used to create a mold. An impression will be made using irreversible hydrocolloid to obtain a space cast, and 4 mm of modeling wax will be applied over the master model to allow room for the impression material in order to build all of the custom trays with the same spacer thickness. To standardize the orientation of the custom trays on the model, the wax spacer will incorporate three tissue stops: two in the posterior region and one in the anterior region. A silicone putty mold will be made in the denture flask, and a tray of wax of consistent thickness will be applied over this space cast. The wax will then be removed by boiling it out for 7 to 10 minutes. Both halves of the flask will be covered with separating material. This flask mold will be used to create custom trays of the same thickness after PMMA resin has been mixed according to the manufacturer's instructions. Once the polymerization is done, the trays will be taken out and will be bench cured for a full day before being used. A round bur will be used to create retention holes in each of the custom trays at evenly spaced intervals (Figure 3).

XSL•FO
RenderX

Figure 3. Fabrication of custom tray.

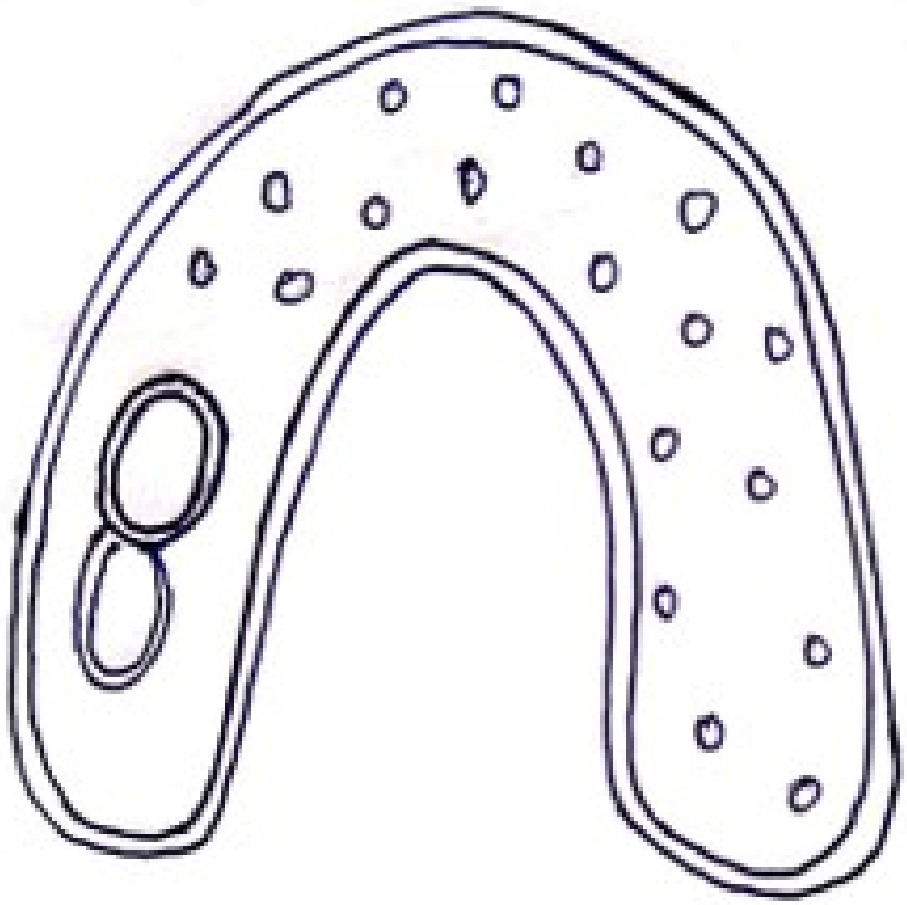

Prefabricated Self-Perforating Trays

Prefabricated self-perforating trays available under the brand name Cotisen Trays, which are made of rigid medical-grade polymer plastic, will be used in the test group. The occlusal surface of the prefabricated self-perforating tray is covered in patented plastic foil. After loading the tray with polyvinyl siloxane impression material, it will be positioned intraorally over the open tray impression heads and forced down crestally until the top of the impression pins can be seen through the clear foil. The impression pins will be pressed up against the tray until they pierce the foil and are visibly sticking out of it. After the impression material has set, an impression will be retrieved (Figure 4).

XSL•FO
RenderX

Figure 4. Prefabricated self-perforating tray.



Impression Making

The sample will be divided into two groups (A and B).

A total of 8 impressions will be made in each group—one using custom acrylic trays and the other using prefabricated self-perforating trays.

Polyvinyl siloxane medium-body impression material will be used to make impressions in both groups due to its excellent dimensional stability, high tear resistance, and superior elastic recovery—key factors for accurately capturing implant positions. Its minimal shrinkage and proven reliability make it the gold standard for implant-level impressions, especially with the open tray technique.

Impressions will be poured using type IV stone, and casts will be obtained.

Scanning of the Obtained Casts

The casts will be scanned using a lab scanner (inEos X5), and an STL file representing the implant placement sites will be generated. This file will be imported into Altair HyperMesh 2021 (Altair Engineering Inc). HyperMesh will serve as the primary platform for refining the model, isolating the implants, and preparing the geometry for further analysis. The Circle Fit Tool in HyperMesh will be used to analyze the cross-sectional profiles of the implants, approximating their base geometry to ensure accurate center identification. Using the Node Tool, the geometric center node of each implant will be determined by fitting a perfect circle to the cross-section of the cylindrical body. This method will ensure precise identification of the implant's central point.

Measuring the Accuracy

After identifying the center nodes for both implants, the Measure Distance tool in HyperMesh will be used to calculate the distance between the two nodes. This measurement will

XSL•FO
RenderX

represent the center-to-center distance between the implants. The process will be repeated multiple times to confirm consistency, and the results will be averaged to reduce the impact of potential variability or measurement noise (Figure 5).

Figure 5. Measuring the distance between two implants.

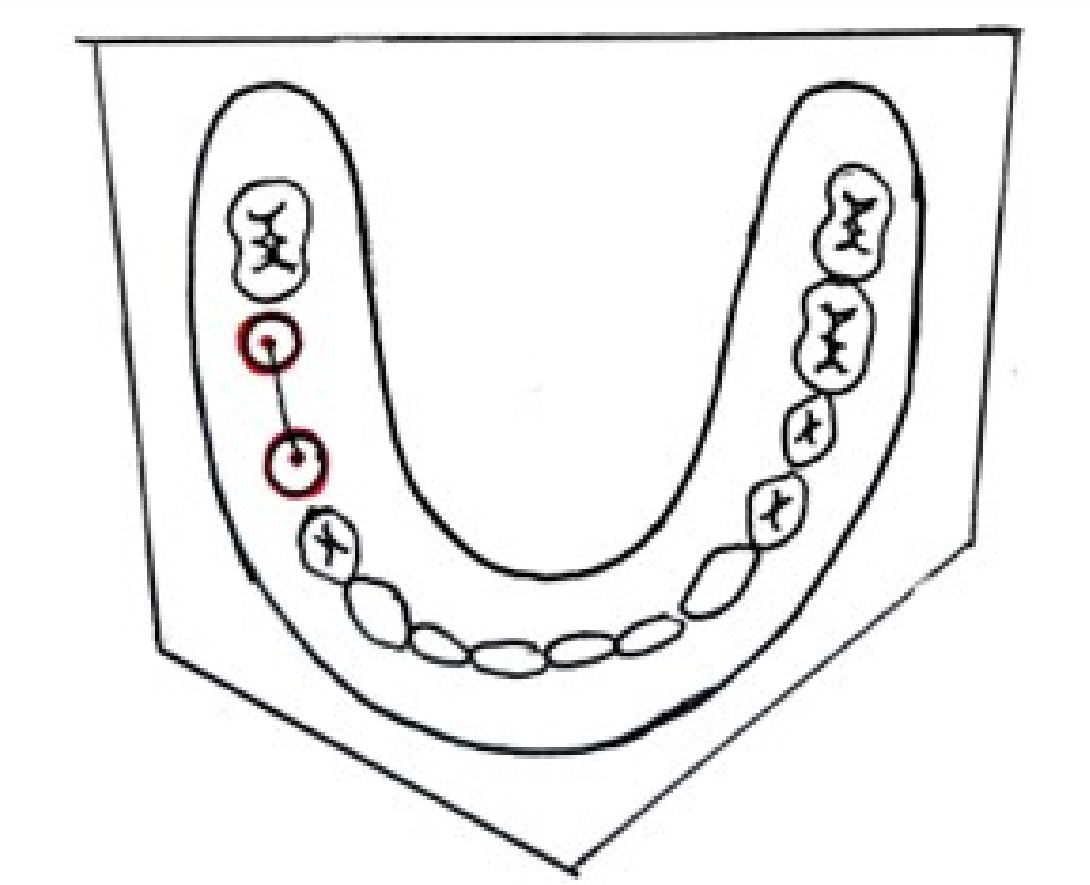

Evaluation of Accuracy

The accuracy of the implant impression tray will be evaluated for both groups. The experimental casts obtained from two distinct implant impression trays will be compared to the interimplant (analog) distance on the master model.

All impression-making, cast pouring, and scanning procedures will be performed by a single calibrated operator to ensure standardization and eliminate interoperator variability.

Primary Outcome

The deviation in interimplant distance between experimental casts obtained from each tray type and the master model will be the primary outcome measure.

Statistical Analysis

All analyses will be done using IBM SPSS Statistics for Windows (version 21.0; IBM Corp). The normality of data will be checked using the Shapiro-Wilk test as it is suitable for small sample sizes. In addition to that, Q-Q plots will be used for visual inspection of normality.

Intergroup comparisons will be performed using an independent samples *t* test if data are normally distributed; otherwise, the Mann-Whitney *U* test will be applied. Descriptive statistics will include mean and standard deviation or median and interquartile range depending on the underlying distribution of data.

Ethical Considerations

The study has been approved by the institutional ethical committee of the Datta Meghe Institute Of Higher Education and Research (reference number DMIHER(DU)/IEC/2025/552).

Results

The interimplant distance measured from casts obtained using custom acrylic trays and prefabricated self-perforating trays will be quantitatively obtained in this study and compared with the corresponding interimplant distance on the master model to determine the dimensional deviation for each tray type.

XSL•FO
RenderX

No funding was obtained for this study. Ethical approval was granted in February 2025. As of July 2025, the models have been prepared for further analysis. It is projected that data collection will be completed in May 2026, and results will be published early in 2027.

Discussion

The accurate replication of intraoral implant positions is critical for the passive fit and long-term success of implant-supported prostheses [6]. The implant-bone connection necessitates a precise fit. A natural tooth's periodontal ligament can move up to 100 µm, allowing for some misfitting to a fixed partial denture, although the maximum movement of an osseointegrated implant is 10 µm [7].

Misfit between the implant framework and underlying structures may result in internal stresses leading to both mechanical and biological complications such as screw loosening, screw fracture, or peri-implantitis [3,8]. Therefore, to create a successful implant-supported prosthesis, the first step is to accurately transfer 3D implant position and angulation from the mouth to the master cast using an impression [9].

One pivotal factor influencing this accuracy is the choice of impression tray—either custom-made or prefabricated self-perforating trays—used during the open tray impression technique.

Custom trays have long been considered the gold standard due to their ability to provide a uniform thickness of impression material and enhanced rigidity, reducing deformation during the impression procedure [10]. However, custom tray fabrication is time-consuming and may not be feasible in all clinical scenarios, especially where rapid impression-taking is necessary [11]. Prefabricated self-perforating trays offer a convenient alternative, allowing for immediate use without the need for laboratory fabrication. These trays are designed to facilitate direct access to the impression copings, essential for open-tray techniques. Despite their practicality, concerns remain about their stability, their adaptation to the arch form, and the resulting accuracy. This practical limitation has led to the need for a reliable alternative to custom-made implant impression trays [3].

Kwon et al [12] compared the dimensional accuracy of implant casts made with and without impression copings using an acrylic resin maxillary model with three implants. Two impression techniques were used to fabricate definitive casts. Centroid and long axis measurements of implant replicas were analyzed. Results showed significantly greater linear distortion in the no-coping group, especially at the first and second molar sites. The study concluded that impression copings improve cast accuracy, particularly when implants have greater interabutment distances.

Gupta et al [10] compared stock trays (plastic and metal) with custom trays and concluded that, when used with medium-viscosity impression materials, stiff nonperforated stock trays made of metal or plastic may serve as a substitute for custom trays for multi-implant impressions. Similarly, Pastoret et al [13] conducted a study to compare the dimensional accuracy of three impression techniques—a separating foil impression, a custom tray impression, and a stock tray impression. They concluded that the separation foil technique is a simple alternative to the custom tray technique for single tooth restorations, while limitations may exist for extended restorations with multiple abutment teeth [13]. Similarly, Goel et al [3] conducted an in vitro study comparing the accuracy of implant impressions made using custom trays (autopolymerizing and light-cured composite) versus specialized aluminum stock trays. Although statistically insignificant differences were observed among the groups, the light-cured custom trays demonstrated the highest accuracy, followed by autopolymerizing and stock trays. This reinforces the notion that custom trays can provide superior dimensional stability while also hinting at the potential clinical viability of stock trays under controlled conditions.

A review by Yasar et al [14] stressed that tray design must be evaluated alongside other variables such as implant angulation, number, and depth, which significantly impact impression accuracy.

Madhan et al [15] concluded that, despite advancements in materials and methodologies, some degree of dimensional distortion persisted, necessitating more studies to evaluate the transfer accuracy of implant impressions. The preliminary study by Spector et al [16] suggests that more research is required to identify methods that can reliably yield precise endosseous implant position recording.

In previous studies, when comparing casts made from stock versus custom trays to the master model, there was a slight statistical difference [17-19]. Although custom trays continue to demonstrate superior dimensional accuracy, prefabricated self-perforating trays offer a promising, efficient, and cost-effective alternative. Their clinical performance can be optimized through the use of appropriate impression materials and techniques. This study will further validate these findings and contribute to establishing evidence-based guidelines for tray selection in implant prosthodontics.

Acknowledgments

The authors are grateful to all of the teaching and nonteaching staff at Sharad Pawar Dental College and Hospital's Department of Prosthodontics for their assistance and support.

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

XSL•FO
RenderX

Authors' Contributions

AA, SN, and SG contributed to the concept and design of the study and drafting of the manuscript. RP, VB, and NZ interpreted the data for the work and drafted the manuscript. All the authors critically revised the manuscript for important intellectual content.

Conflicts of Interest

None declared.

References

1. Mpikos P, Kafantaris N, Tortopidis D, Galanis C, Kaisarlis G, Koidis P. The effect of impression technique and implant angulation on the impression accuracy of external- and internal-connection implants. Int J Oral Maxillofac Implants 2012;27(6):1422-1428. [Medline: 23189292]
2. Bhakta S, Vere J, Calder I, Patel R. Impressions in implant dentistry. Br Dent J 2011 Oct 21;211(8):361-367. [doi: 10.1038/sj.bdj.2011.862] [Medline: 22015512]
3. Goel M, Dhawan P, Tandan P, Madhukar P. Need for a reliable alternative to custom-made Implant Impression trays: an in vitro study comparing accuracy of custom trays versus specialized aluminum stock tray. J Indian Prosthodont Soc 2018;18(3):271-276. [doi: 10.4103/jips.jips_232_17] [Medline: 30111917]
4. Patil P, Madhav VNV, Alshadidi AAF, et al. Comparative evaluation of open tray impression technique: investigating the precision of four splinting materials in multiple implants. BMC Oral Health 2023 Nov 8;23(1):844. [doi: 10.1186/s12903-023-03583-x] [Medline: 37940890]
5. Bohner L, Marques A, Keedi C, Steagall W, Tortamano P. Accuracy of self-perforating impression tray for maxillary dental implant. Clinical and Laboratorial Research in Dentistry 2018. [doi: 10.11606/issn.2357-8041.clrd.2017.135626]
6. Arikan H, Muhtarogullari M, Uzel SM, et al. Accuracy of digital impressions for implant-supported complete-arch prosthesis when using an auxiliary geometry device. J Dent Sci 2023 Apr;18(2):808-813. [doi: 10.1016/j.jds.2023.01.012] [Medline: 37021239]
7. Joseph TM, Ravichandran R, Lylajam S, Viswabharan P, Janardhanan K, Rajeev S. Evaluation of positional accuracy in multiple implants using four different splinting materials: an in vitro study. J Indian Prosthodont Soc 2018;18(3):239-247. [doi: 10.4103/jips.jips_27_18] [Medline: 30111913]
8. Buzayan MM, Yunus NB. Passive fit in screw retained multi-unit implant prosthesis understanding and achieving: a review of the literature. J Indian Prosthodont Soc 2014 Mar;14(1):16-23. [doi: 10.1007/s13191-013-0343-x] [Medline: 24604993]
9. Wee AG, Aquilino SA, Schneider RL. Strategies to achieve fit in implant prosthodontics: a review of the literature. Int J Prosthodont 1999;12(2):167-178. [Medline: 10371920]
10. Gupta S, Narayan AI, Balakrishnan D. In vitro comparative evaluation of different types of impression trays and impression materials on the accuracy of open tray implant impressions: a pilot study. Int J Dent 2017;2017:6306530. [doi: 10.1155/2017/6306530] [Medline: 28348595]
11. Mahesh L, Kurtzman G. Advances in dental impressions. Impladent. 2015. URL: https://www.centrefordentalimplant.com/wp-content/uploads/2016/11/ADVANCES-IN-IMPLANT-IMPRESSIONS.pdf [accessed 2026-01-27]
12. Kwon JH, Son YH, Han CH, Kim S. Accuracy of implant impressions without impression copings: a three-dimensional analysis. J Prosthet Dent 2011 Jun;105(6):367-373. [doi: 10.1016/S0022-3913(11)60075-9] [Medline: 21640237]
13. Pastoret MH, Krastl G, Bühler J, Weiger R, Zitzmann NU. Accuracy of a separating foil impression using a novel polyolefin foil compared to a custom tray and a stock tray technique. J Adv Prosthodont 2017 Aug;9(4):287-293. [doi: 10.4047/jap.2017.9.4.287] [Medline: 28874996]
14. Yasar MN, Cetinsahin C, Bayar O, Ozer HY. Implant impression techniques using different materials and methods: a review. JCDR 2022;16. [doi: 10.7860/JCDR/2022/53057.16014]
15. Madhan R, Nayar S, Annapoorani H. Comparative evaluation of accuracy of six different implant impression techniques: an in vitro study. J Indian Prosthodont Soc 2006;6(4):185. [doi: 10.4103/0972-4052.30693]
16. Spector MR, Donovan TE, Nicholls JI. An evaluation of impression techniques for osseointegrated implants. J Prosthet Dent 1990 Apr;63(4):444-447. [doi: 10.1016/0022-3913(90)90235-5] [Medline: 2184229]
17. Rueda LJ, Sy-Muñoz JT, Naylor WP, Goodacre CJ, Swartz ML. The effect of using custom or stock trays on the accuracy of gypsum casts. Int J Prosthodont 1996;9(4):367-373. [Medline: 8957875]
18. Mallick R, Pisulkar SK, Reddy SG. Assessment of outcomes of immediately loaded dental implants in orofacial cleft patients: protocol for a single-arm clinical trial. JMIR Res Protoc 2021 May 5;10(5):e25244. [doi: 10.2196/25244] [Medline: 33949960]
19. Kumar PS, Ks SK, Grandhi VV, Gupta V. The effects of titanium implant surface topography on osseointegration: literature review. JMIR Biomed Eng 2019 Jun 11;4(1):e13237. [doi: 10.2196/13237]

XSL•FO
RenderX

Edited by GRML Rosa; submitted 21.Jun.2025; peer-reviewed by MM ElSheikh, S Chaturvedi; revised version received 29.Jul.2025; accepted 30.Jul.2025; published 19.Feb.2026.

Please cite as:
Agrawal A, Nimonkar S, Godbole S, Parakh R, Belkhode V, Zilpilwar N
In Vitro Comparison of the Dimensional Accuracy of Implant Impressions Using Custom Acrylic Trays and Prefabricated Self-Perforating Trays: Protocol for a Comparative Evaluation Study
JMIR Res Protoc 2026;15:e79452
URL: https://www.researchprotocols.org/2026/1/e79452
doi:10.2196/79452

© Arti Agrawal, Sharayu Nimonkar, Surekha Godbole, Rushabh Parakh, Vikram Belkhode, Namita Zilpilwar. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Analysis of Associated Factors and Construction of Risk Prediction Models for Frailty in Hospitalized Older Adults Living With HIV: Protocol for a Prospective Observational Study

Fan Li[1,2,3], MMSc; Chaoying Xie[2], MMSc; Fang Xiang[2], MMSc

[1]School of Nursing, University of South China, Hengyang, China

[2]The First Hospital of Changsha, Changsha First Hospital, Kaifu District, Changsha, China

[3]Nursing Department, The Affiliated Changsha Hospital of Xiangya School of Medicine, Central South University, China

Corresponding Author:
Chaoying Xie, MMSc
The First Hospital of Changsha, Changsha First Hospital, Kaifu District, Changsha, China

Abstract

Background: The aging trend of people living with HIV or AIDS in China is increasing day by day. Frailty is a common condition among older adults living with HIV or AIDS and represents a significant cause of poor prognosis, including falls, decreased quality of life, increased mortality, and potentially prolonged hospital stays. Consequently, early frailty screening in this population holds important clinical significance.

Objective: This study aims to describe the theoretical basis, research objectives, and implementation plan of a prospective observational study. It will focus on investigating the current status of frailty syndrome in hospitalized older adults living with HIV or AIDS, while simultaneously exploring the development of a clinically applicable risk prediction model.

Methods: This study is an ongoing single-center prospective observational study, with a plan to recruit at least 556 hospitalized older adults living with HIV or AIDS (n=445 for development and n=111 for validation). According to the theory of unpleasant symptoms, candidate predictors are categorized into physiological factors (including sociodemographic factors, disease-related influencing factors, sleep, nutrition, and neurocognitive function), psychological factors (including anxiety and depression status), and environmental factors (including social support status). Potential predictors are screened using univariate analysis and least absolute shrinkage and selection operator regression to identify variables for final model inclusion. Model construction and validation employ 3 standard machine learning algorithms: logistic regression, random forest, and support vector machine. Model performance will be evaluated by reporting accuracy, precision, sensitivity, specificity, and the area under the curve.

Results: This study is conducted at a designated infectious disease hospital in Changsha, Hunan Province, China. Participant recruitment commenced on December 22, 2024, and as of December 5, 2025, a total of 603 patients have been enrolled. The primary study findings are anticipated to be published in August 2026.

Conclusions: The findings of this study are expected to provide clinicians in the department of infectious diseases with a convenient tool for frailty risk prediction, thereby enabling early intervention and ultimately improving the long-term health status and quality of life of people living with HIV.

Trial Registration: Chinese Clinical Trial Registry ChiCTR2500103387; https://www.chictr.org.cn/showproj.html?proj=266120

International Registered Report Identifier (IRRID): DERR1-10.2196/84271

(JMIR Res Protoc 2026;15:e84271) doi:10.2196/84271

KEYWORDS

people living with HIV; frailty; risk factors; predictive models; theory of unpleasant symptoms

Introduction

With the promotion and application of highly active antiretroviral therapy, the treatment prospect of AIDS has been significantly improved. Highly active antiretroviral therapy effectively suppresses viral load, controls disease progression in people living with HIV, prolongs survival, and enhances quality of life. The increased survival rate has led to a growing population of people living with HIV or AIDS aged ≥50 years [1]. Consequently, the Centers for Disease Control and Prevention in the United States defines individuals aged ≥50 years with HIV as the “older adult population” rather than the commonly defined population aged ≥60 years [2]. Older adults living with HIV or AIDS face not only age-related disease risks but also chronic HIV-associated conditions and synergistic interactions between HIV and non-HIV comorbidities [2].

XSL•FO
RenderX

Chronic immune damage and inflammation contribute to declining physical function and resilience, predisposing them to frailty—a dynamic geriatric syndrome. The weakened state can lead to a higher risk of adverse events for people living with HIV or AIDS, which is an independent risk factor for new chronic diseases, falls, cognitive decline, multidrug treatment, hospitalization, loss of independence, and increased mortality [3]. Thus, early identification of frailty and implementation of targeted preventive measures are critically important.

Frailty represents a complex state. Strain's study analyzed neuroimaging markers (including diffusion tensor imaging and cerebral blood flow), revealing that frailty in people living with HIV or AIDS correlates with structural and functional brain alterations, specifically impaired psychomotor speed and executive function performance [4]. HIV-infected individuals who are in a frailty state exhibit atrophy in the anterior cingulate cortex and cingulate cortex regions, which is associated with a decrease in their quality of life. A multidimensional study of HIV-infected adults at the University of California also confirmed this result, showing that frailty has a significant negative impact on the health-related quality of life of people living with HIV or AIDS, and the impact on people living with HIV or AIDS is significantly stronger than that on uninfected adults [5]. In addition, statistical analysis of clinical events following baseline frailty assessment elevated cardiovascular disease risk in frail older adults living with HIV or AIDS (incidence rate ratio of 3.83), followed by new-onset diabetes and orthopedic disorders [6].

The impact of frailty on cognitive function in older adults living with HIV or AIDS is increasingly recognized. HIV-associated neurocognitive disorders (HAND) are one of the most common complications of HIV-related nervous system diseases. According to the diagnostic criteria of the US HIV Neurobehavioral Research Center in 2007, HAND can be divided into asymptomatic neurocognitive impairment, HIV-related mild neurocognitive disorder, and HIV-related dementia [7]. Research has shown that the incidence of HAND in people living with HIV or AIDS over 50 years old is as high as 40.5% to 54.8% [8,9], and there is a synergistic effect between cognitive impairment and frailty [10], that is, frailty can lead to a decline in cognitive function and a higher risk of cognitive impairment. Conversely, impaired cognitive function can also exacerbate the process of frailty. Through a 1-year follow-up study of 60 people living with HIV or AIDS, half of the participants experienced varying degrees of cognitive impairment after the follow-up, with higher levels of frailty indicating a greater risk of cognitive impairment [11]. Notably, both prefrailty and frailty status significantly correlate with HAND, especially mild neurocognitive disorder, further confirming that frailty is an important cause of HAND occurrence [12].

Currently, unidimensional frailty assessment remains predominant in research involving people living with HIV or AIDS, while research on multidimensional frailty status is still insufficient, and the factors leading to multidimensional frailty in hospitalized older adults living with HIV or AIDS are not yet clear. Therefore, we will explore multidimensional frailty as the main outcome. The theory of unpleasant symptoms (TOUS), as a mid-level theory, is commonly used to establish methodological methods and analyze results, or as a theoretical framework for analyzing influencing factors [13,14]. The theory suggests that factors that affect symptoms can be divided into 3 parts: physiological, psychological, and environmental. The three interact and influence each other, jointly affecting disease symptoms and influencing the outcome of symptoms. Guided by TOUS, we categorized potential frailty factors in hospitalized older adults living with HIV or AIDS into physiological, psychological, and environmental domains based on literature review and clinical practice. Ultimately, all included potential influencing factors were identified and finalized through online expert consultation.

Early-stage frailty in older adults living with HIV or AIDS is often relatively insidious. However, to our knowledge, there is relatively little research and development in China on frailty prediction tools for hospitalized population. To address these gaps, we aim to develop and validate a prediction model using 3 machine learning algorithms. By comparing their performance, we seek to enhance risk prediction accuracy, ultimately improving patient prognosis and health outcomes.

This prospective observational study aims to (1) determine the prevalence of multidimensional frailty among hospitalized people living with HIV or AIDS aged ≥50 years, (2) identify factors associated with frailty guided by the TOUS, and (3) develop and validate risk prediction models.

Methods

Ethical Considerations

This prospective observational study protocol has been approved by the Medical Ethics Office of Changsha First Hospital (approval number: Ethics Committee: [2025] Express Review [Clinical Documents] No.12) and the Medical Ethics Committee of University of South China (approval number: 2024HLSC025) and registered on ClinicalTrials.gov (ChiCTR2500103387) to ensure transparency and compliance with best practices.

Written informed consent from all participants is required before participating in the study. During the research, all personal data collected (such as identity information, date of birth, medical payment, health information, etc) will be encoded and replaced to protect the identity of participants and be protected under applicable laws. Only researchers and authorized personnel can access these materials. The research results will be reported and disseminated through open and transparent peer-reviewed journals and conference reports.

This study is a noncommercial academic research study and does not provide any direct financial compensation to the participants. All participants receive professional personalized health guidance and counseling based on their examination results, which is the main benefit of their participation in this study.

Participants and Qualifications

The inclusion and exclusion criteria are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for a prospective observational study.

Inclusion Criteria

1. People living with HIV or AIDS, who are diagnosed according to the HIV/AIDS virus infection diagnosis industry standard (WS 293-2019) [15]
2. Age ≥50 years
3. Clear awareness and ability to cooperate in completing investigations and testing
4. Voluntary participation and provision of informed consent

Exclusion Criteria

1. Physical functional impairments such as severe language communication barriers, hearing impairments, or visual impairments that hinder normal assessment
2. Hemiplegia
3. Refusal to sign the informed consent form

Patient Recruitment

All eligible patients will be recruited within 24 to 48 hours after admission. Trained researchers will determine potential participants based on inclusion and exclusion criteria and explain this observational study to patients and their families. After patients confirm and sign informed consent forms, they will be included in this study. The examination results and related data of all participants will be subject to strict confidentiality management to ensure the privacy and security of patients. Participants have the right to withdraw from the study in any form at any stage. We plan to recruit at least 556 hospitalized older adults living with HIV or AIDS (n=445 for development and n=111 for validation). The research evaluation will be completed during the patient's hospitalization period.

Sample Size

Frailty occurrence in people living with HIV or AIDS represents a binary outcome. Previous studies reported a frailty incidence of 59.3% in hospitalized older adults living with HIV or AIDS, which was assessed using the FRIDE frailty scale [16]. Due to the current lack of a frailty prediction model for HIV/AIDS patients, it is not possible to determine R^2 through previous models. Based on the proportion of outcome events, the maximum R^2 (R^2 cs_max) is calculated to be 0.7416. Assuming that the new model can explain 20% of the variability, the expected R^2 is 0.1483 (=0.7416×20%). We expect the predictive model to include 5 to 10 predictor variables, based on "pmsampsize(type='b', csrsquared=0.1483, parameters=10 , prevalence=0.593)." In this formula, csrsquared represents the ability of the expected prediction model to explain outcome variability; parameters represent the total number of candidate variables expected to be included; and prevalence refers to the prevalence of frailty events in older people living with HIV or AIDS in China based on previous research. The calculation indicates that a minimum sample size of 556 is required. A total of 330 events were expected, corresponding to 33 events per predictor variable. The above results are calculated using the pmsampsize package in R Statistical Software (version 4.4.0) [17].

Definition of Outcomes

At present, the frailty assessment methods adopted in various domestic studies are inconsistent, and there are significant differences in research results. Most studies use a single physiological dimension frailty scale (such as frailty phenotype). We believe that frailty phenotype, as a universal frailty assessment tool including core indicators such as weight loss and perceived fatigue, may be affected by HIV-specific factors. For example, a decline in immune system function leads to the occurrence of opportunistic infections, as well as nausea, vomiting, diarrhea, and nutrient absorption disorders caused by the side effects of most antiviral drugs, which makes people living with HIV or AIDS more prone to weight loss. In addition, most antiretroviral drugs cause patients to experience fatigue and exhaustion, which may exacerbate physical inactivity. Current research on multidimensional frailty is expanding. Global studies have confirmed the strong applicability of the Tilburg Frailty Indicator in people living with HIV or AIDS [18,19]. Its comprehensive advantage in assessing frailty across physical, psychological, and social domains may render it more suitable for the objectives of this study than unidimensional tools.

The main outcome of this study is the presence of frailty among people living with HIV or AIDS. We use the Tilburg Frailty Indicator for reporting, developed by Gobbens et al [20] and localized by Li et al [21]. The scale consists of 15 items, with physiological dimensions including 8 items: physical health, natural weight loss, walking difficulties, balance, vision, hearing, grip strength, and fatigue; the psychological dimension includes 4 items: memory, depression, anxiety, and coping ability; and the social dimension includes 3 items: living alone, social relationships, and social support. If the patient's current condition meets the description of 5 or more items, it indicates that the patient has frailty. The final assessment results of the scale will be confirmed by 2 specialists in infection and immunology.

Prediction Variables

Determination of Candidate Variables

The selection of predictive variables for this prediction model is based on a review of literature and consensus opinions by an expert group, taking into account its practicality in clinical

XSL•FO
RenderX

settings. Possible clinically relevant variables will be selected as candidate predictors of frailty and evaluated after the patient's admission examination is completed. Through 3 group discussions and 1 online expert consultation meeting, a total of 31 candidate predictor variables were identified and classified into physiological, psychological, and environmental factors based on TOUS. The variables to be explored are detailed in the following subsections.

Physiological Factors

The sociodemographic data assessed included age, gender, marital status, place of residence, education level, work situation, average monthly income, medical payment method, height, weight, BMI, smoking history, alcohol consumption history, and exercise status [22-25].

Disease-related conditions include time of HIV diagnosis, start time of antiviral treatment, route of infection, and the types of diseases that have been diagnosed at present and other associated complications (heart failure, hypertension, cerebrovascular disease, diabetes, moderate or severe liver disease, moderate or severe kidney disease, chronic obstructive pulmonary disease, and tumors) [26-29].

The Pittsburgh Sleep Quality Index (PSQI) developed by Buysse is used for collecting sleep data [30,31]. It includes 7 dimensions: sleep quality, bedtime, sleep duration, sleep efficiency, sleep disorders, hypnotic drugs, and daytime dysfunction, totaling 18 items. Each dimension is scored from 0 to 3 points, with 0 indicating the absence of sleep disorders, 1 indicating mild disorders, 2 indicating moderate disorders, and 3 indicating the presence of severe disorders. The total score ranges from 0 to 21, with a PSQI >7 indicating sleep problems.

Nutritional status is assessed using the Mini Nutritional Assessment-Short Form scale developed by Rubenstein and consists of 6 items, including changes in appetite, weight, mobility, activity, mental health status, and BMI index [32,33]. The total score is 0 to 14 points, with <8 points indicating malnutrition, 8 to 11 points indicating a risk of malnutrition, and ≥12 points indicating good nutritional status. Research has shown that this scale has good sensitivity and specificity in China [34].

Neurocognitive function is evaluated using the International HIV Dementia Scale, a rapid screening tool developed by Sacktor et al [35,36,37]. It can screen for psychomotor speed, attention, working memory, executive function, and language. Including 3 tests with a total score of 12, patients with a score of ≤10 should be further evaluated for possible dementia. Studies have shown that the International HIV Dementia Scale has better efficiency in detecting the neurocognitive performance of patients using antiretroviral drugs [38].

Psychological Factors

The Hospital Anxiety and Depression Scale is used for anxiety and depression [39,40]. Mainly used for screening anxiety and depression status of patients in hospitals. The Hospital Anxiety and Depression Scale is divided into 7 anxiety items and 7 depression items. Each item is scored from 0 to 3 points, with higher scores indicating more severe anxiety and depression. This scale has good reliability and validity in the application of people living with HIV or AIDS in China [41,42].

Social Factors

The Chinese version of the Social Support Rating Scale developed by Xiao [43] is used to measure an individual's social support status, including three dimensions: objective support, subjective support, and personal utilization of social support [44,45]. There are a total of 10 items, with a total score of 12 to 66 points. The higher the score, the higher the level of social support. A general total score of ≥45 indicates a high level of support, 23 to 44 indicates a moderate level of support, and ≤22 indicates a low level of support. Social Support Rating Scale is widely used in people living with HIV or AIDS [46].

Laboratory examination indicators were HIV load, CD^4 T lymphocyte count, albumin, C-reactive protein (CRP), interleukin-6 (IL-6), and soluble tumor necrosis factor receptors (sTNFR1 and sTNFR2) [47-50]. It is worth noting that although inflammatory indicators such as IL-6 and sTNFR1/2 are potential predictive factors for hospitalized older adults living with HIV or AIDS, these items do not have clinical applicability due to their noninfectious and immunological routine clinical testing content, and relevant test results cannot be obtained.

Data Acquisition

Data will be acquired through questionnaires and hospital medical record systems. We use face-to-face, on-site distribution, and questionnaire collection to obtain general patient information (demographic and sociological information, lifestyle habits, and disease-related conditions), frailty scores, PSQI, anxiety and depression scores, nutrition scores, and social support scores. Physiological measurements will be taken on-site. For participants with lower educational levels or who are unable to complete the survey independently, researchers will fill in the questionnaire on their behalf based on their answers and verify them one by one. To ensure that the natural heterogeneity of the population is not disturbed by human factors, the questionnaire collection process will strictly follow the pre-established standard operating procedures.

The relevant laboratory tests will be conducted within 24 to 48 hours after admission. The collection of venous blood samples is strictly carried out by specialized nurses in the field of infection and immunology, following standard operating procedures. This step is included in the routine admission examination, and we do not charge any additional examination fees for it. We will inform patients before conducting the study that we need to use some personal biochemical test data and enter it into the dataset after obtaining consent. Obtain laboratory test indicators, including HIV load, CD^{4+} T lymphocyte count, albumin, and CRP, through the test records in the case record system of Changsha First Hospital. However, this study did not collect inflammatory indicators such as IL-6 and sTNFR1/2, which may affect the in-depth analysis of the mechanism of frailty inflammation. It is recommended to prioritize supplementary verification in future research.

Neurocognitive function assessments will be administered face-to-face by 2 researchers within 48 to 72 hours of admission.

XSL•FO
RenderX

The examination is conducted in a separate interview room to avoid different patients in the same ward being informed of the test content in advance, which may affect the results of the neurocognitive function examination.

Statistical Analysis

SPSS software (version 26.0; IBM Corp) and R Statistical Software (version 4.4.0; R Foundation for Statistical Computing) are used for data analysis. Categorical variables are presented as frequencies and percentages. Normally distributed continuous data are expressed as mean (SD), while nonnormally distributed continuous or ordinal variables are summarized as median (IQR). Group comparisons employed *t* test or one-way ANOVA for normally distributed continuous variables, Mann-Whitney *U* test (2 groups) or Kruskal-Wallis *H* (multiple groups) test for non-normal/ordinal data, and chi-square test for categorical variables. For baseline data, a *P* value <.05 is considered statistically significant, indicating a significant difference in baseline characteristics between the group with frailty and the group without frailty. For missing data, considering the instability of missing values, this study used multiple imputation method (imputation tool: "mice" package in R Statistical Software [version 4.4.0]) to fill in the missing parts. Participants will be excluded if there are 3 or more candidate variables missing. After imputation, sensitivity analysis will be used to validate the plausibility of the Missing at Random assumption and the robustness of the imputation results. By comparing the results of multiple imputation datasets with complete case analysis (ie, removing missing data), we further validate the robustness of our main conclusions.

Model Development and Validation

The sample size of this study is relatively small. This study uses logistic regression (LR) as the benchmark for performance comparison. LR has certain advantages in handling limited sample sizes and can demonstrate feature coefficients and significance levels. Although random forest (RF) has the risk of overfitting in small sample situations, it has certain advantages in handling complex nonlinear relationships. We will minimize the risk by adjusting decision tree parameters and cross-validation. Support vector machine (SVM) is particularly suitable for handling situations with high feature dimensions but relatively small sample sizes, and its regularization properties help control the risk of overfitting. As one of the tree ensemble algorithms, XGBoost also has the risk of overfitting and relatively complex parameter tuning. Neural networks have high requirements for sample size and poor clinical interpretability for small sample studies. We hope to ensure the robustness of the model to a large extent in this small sample study. For future larger-scale research, more complex machine learning methods will be used.

Randomly split the training set and validation set in a 7:3 ratio, and use the following machine learning algorithms for modeling: LR model, RF, and SVM. Univariate analysis combined with least absolute shrinkage and selection operator regression will be used to screen predictive factors. To address nonlinearity, we will analyze the relationships between relevant factors and HIV frailty risk using restricted cubic splines and incorporate them as spline terms in the model. Restricted cubic splines can also indirectly capture potential interactions between features. Interactions and complex nonlinear decision boundaries will be handled by selecting nonlinear kernel functions in SVM. The primary model performance evaluation metric is accuracy, which represents the proportion of correctly predicted samples in classification tasks. This will be considered alongside precision, recall, F_1-score, and area under the receiver operating characteristic curve for comprehensive assessment. Through integrated comparison, the algorithm achieving optimal balance across these metrics will be selected as the final model for predicting frailty status in people with HIV.

To prevent overfitting, the nonlinear modeling effects of all models will be validated via cross-validation. A 3×10-fold cross-validation scheme will be adopted, with each split performed independently, yielding 30 model performance estimates (eg, area under the curve [AUC], F_1-score). Overfitting will be monitored by ensuring that the performance difference between the training set and the cross-validation set remains below a threshold (AUC difference <0.05). During cross-validation, hyperparameters (including max_depth for RF, and C and γ for SVM) will be jointly optimized to select the parameter combination that achieves the best trade-off between fitting efficacy and generalization ability, thereby further reducing overfitting risk. The final model selected through cross-validation will then be evaluated on an independent test set for ultimate performance assessment. For internal validation, we performed bootstrap resampling (n=1000 iterations) to generate training sets matching the original cohort size. Retraining prediction models with these training sets and evaluating performance using the above indicators, we will repeat this process to obtain internal validation performance. External validation will be implemented in prospective observational studies, using external period validation to analyze samples from September 1, 2025 to December 31, 2025. In future research, we plan to collaborate with other institutions to continue external validation. The sample size required for external validation in this study is 1/4 of the modeling sample size, and the minimum sample size for external validation is calculated to be 139 cases (Figure 1).

XSL•FO
RenderX

Figure 1. Flow chart [15]. SVM: support vector machine.

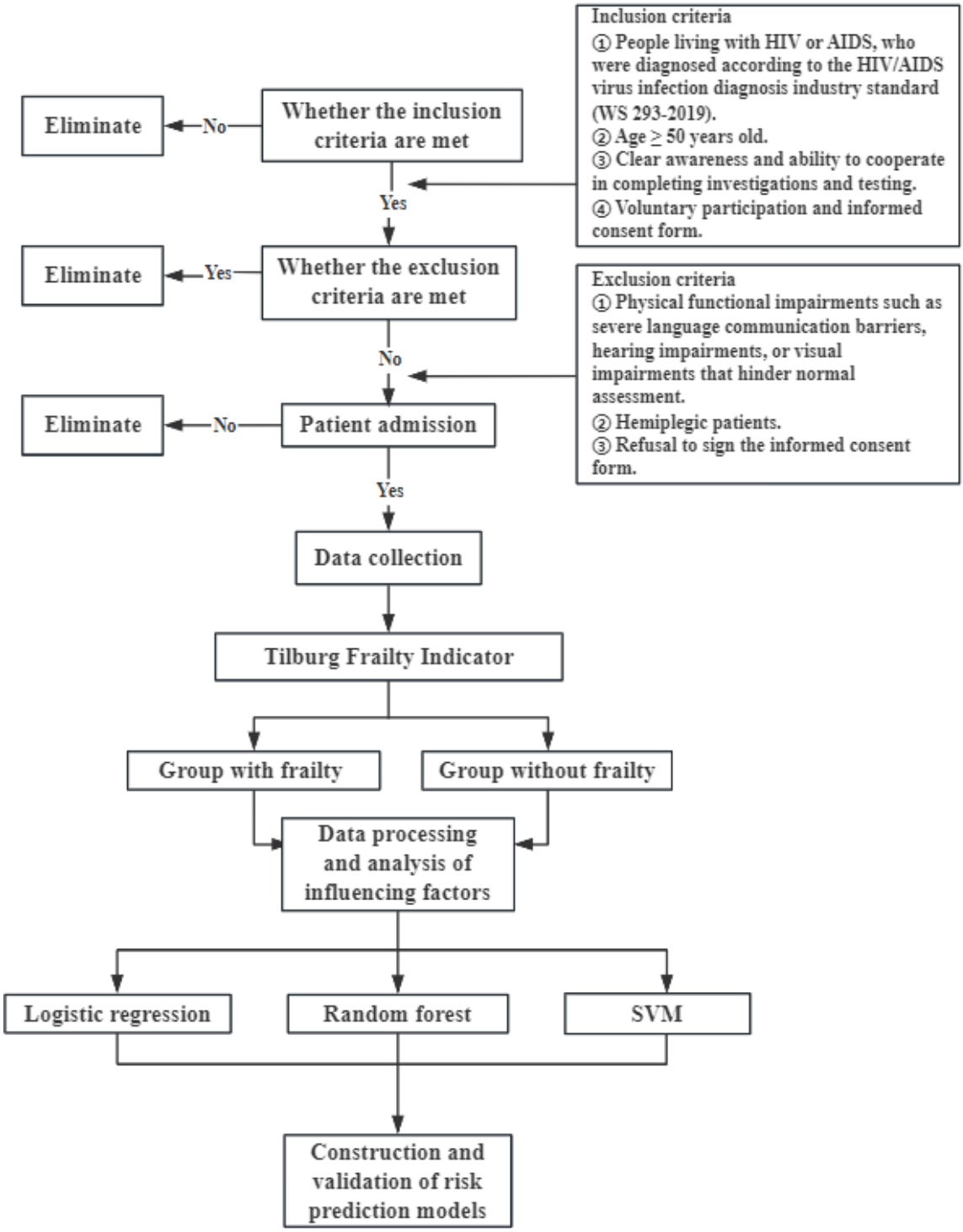

Results

This study is conducted at a designated infectious disease hospital, the First Hospital (Infectious Diseases Hospital) of Changsha City, Hunan Province, China, which is one of the professional designated treatment centers for HIV/AIDS in China. It is supported by the Traditional Chinese Medicine Research Program of the Health Commission of Changsha, Hunan Province (no. B202318) and the Hunan Provincial Natural Science Foundation of China (no. 2023JJ60388). The study was planned to be conducted from December 2024 to August 31, 2025, with external validation conducted from September 1, 2025, to December 31, 2025. Recruitment for the study began on December 22, 2024. As of December 5, 2025, a total of 603 patients have been recruited. Currently, no data analysis has been conducted. The primary study results are expected to be announced in August 2026.

XSL•FO
RenderX

Discussion

Studies indicate a frailty prevalence of up to 12% among people living with HIV or AIDS in China [51], with higher rates observed in hospitalized older adults [16]. With the increasing aging phenomenon of people living with HIV or AIDS, researchers need to further focus on the complex health challenges it brings. The PWH Primary Care Guidelines suggest that it is necessary to conduct frailty screening for individuals aged 50 years and above [52]. This study introduces the principles and design of a registry-based observational study. We hope that the use of this predictive model tool can accurately evaluate whether hospitalized people living with HIV or AIDS have weakened, and we hope that the identification of predictive factors will help health care workers provide preventive care.

Previous studies in China have mostly used frailty assessment tools that focus on physiological dimensions [53,54]. Recent evidence suggests that social vulnerability is particularly prominent in people living with HIV or AIDS [55], and frailty is increasingly reflected in psychological distress and high levels of social isolation [56]. This suggests that using a single physiological dimension evaluation tool may have certain limitations, and the necessity of conducting multidimensional evaluations is evident [18,19]. In addition, current evidence on frailty is mainly derived from cross-sectional studies. Given that frailty is a dynamic evolutionary process (meaning that there are often transitions between different stages of frailty: non-frailty, pre-frailty, and frailty), more cohort studies may be needed in China to validate the risk factors and their dynamic impacts across various dimensions [48].

Considering clinical applicability, some potential relevant biomarker data (IL-6, sTNFR1, sTNFR2) missing from the electronic health data system in this study were not included in the analysis, which may affect the estimation of predictor effects. Studies have shown that inflammatory markers IL-6 and TNFR1 are strongly associated with an increased risk of frailty, with each one-standard-deviation increase in $\log_{10}$-IL-6 (OR 1.67, 95% CI 1.20 - 2.32) and $\log_{10}$ TNFR1 (OR 1.91, 95% CI 1.40 - 2.61) [57]. Based on the predictive value of inflammatory markers for frailty in existing research, the absence of such indicators may lead to a decrease in the predictive power (AUC) of the model and result in the estimation of the effects of identified risk factors deviating from the true values. However, the CRP included in this study serves as a systemic inflammatory marker for routine clinical testing, and its expression is induced by IL-6, thus being somewhat correlated with IL-6 levels [58]. This allows CRP to partially compensate for the impact of inflammatory cytokine deficiency on outcomes.

As a regional AIDS treatment center, the First Hospital of Changsha covers many places in Hunan Province and surrounding provinces and cities. Our research is based on this single center design, which may limit its universality and applicability: different regions have certain differences in cultural backgrounds, accessibility of medical care resources, and social support systems [59]. However, it is worth noting that a drug resistance study conducted HIV-1 subtype analysis on the people living with HIV or AIDS of our center [60], showing that the most common subtypes in the region include CRF01_AE and CRF07BC, which is similar to the results reported in multiple other regions of China (Shanxi Province, Guangdong Province, Jiangsu Province, and Shanghai City) [61-64]. This indicates that the virological background of patients in the region may be comparable to a wider population nationwide. In addition, this study is a real-world hospital observational cross-sectional study and cannot establish causal relationships between various related factors and multidimensional frailty. In the future, we plan to conduct multicenter prospective studies to expand the sample size and unify the collection and detection of more potential related biomarkers, including IL-6 and sTNFR1/2. By comparing the performance of the new model with the original model on an independent validation set, we hope to further evaluate the incremental predictive value provided by these biomarkers and verify the causal relationship of related factors.

Funding

This work was supported by the Traditional Chinese Medicine Research Program of Changsha Health Commission in Hunan Province, China (B202318) and the Natural Science Foundation of Hunan Province (2023JJ60388).

Data Availability

This protocol does not include data sharing as data collection and analysis are still ongoing.

Authors' Contributions

FL contributed to study concept and design, as well as original draft writing. CX participated in project management, funding acquisition and project administration. FX participated in formal analysis. All authors participated in the review and editing of the manuscript.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Krentz HB, Gill MJ. Increased costs of HIV care associated with aging in an HIV-infected population. HIV Med 2015 Jan;16(1):38-47. [doi: 10.1111/hiv.12176] [Medline: 25105798]
2. Rodriguez-Penney AT, Iudicello JE, Riggs PK, et al. Co-morbidities in persons infected with HIV: increased burden with older age and negative effects on health-related quality of life. AIDS Patient Care STDS 2013 Jan;27(1):5-16. [doi: 10.1089/apc.2012.0329] [Medline: 23305257]
3. Eke UA, Mohanty K, Gruber-Baldini AL, Ryan AS. Frailty and aging in HIV: status post 13 years of national awareness. J Frailty Aging 2023;12(1):49-58. [doi: 10.14283/jfa.2022.45] [Medline: 36629084]
4. Strain JF, Cooley S, Kilgore C, et al. The structural and functional correlates of frailty in persons with human immunodeficiency virus. Clin Infect Dis 2022 Nov 14;75(10):1740-1746. [doi: 10.1093/cid/ciac271] [Medline: 35404408]
5. Rubtsova AA, Sabbag S, Sundermann E, et al. Frailty and neurocognitive impairment: impacts on quality of life in HIV. J Assoc Nurses AIDS Care 2020;31(3):290-300. [doi: 10.1097/JNC.0000000000000142] [Medline: 31789684]
6. Kelly SG, Wu K, Tassiopoulos K, Erlandson KM, Koletar SL, Palella FJ. Frailty is an independent risk factor for mortality, cardiovascular disease, bone disease, and diabetes among aging adults with human immunodeficiency virus. Clin Infect Dis 2019 Sep 27;69(8):1370-1376. [doi: 10.1093/cid/ciy1101] [Medline: 30590451]
7. McArthur JC, Brew BJ, Nath A. Neurological complications of HIV infection. Lancet Neurol 2005 Sep;4(9):543-555. [doi: 10.1016/S1474-4422(05)70165-4] [Medline: 16109361]
8. Metral M, Nadin I, Locatelli I, et al. How helpful are the European AIDS Clinical society cognitive screening questions in predicting cognitive impairment in an aging, well-treated HIV-positive population? HIV Med 2020 May;21(5):342-348. [doi: 10.1111/hiv.12828] [Medline: 31883203]
9. Salis F, Belfiori M, Bellisai A, et al. Cognitive impairment in people living with HIV and the impact of mood: results from a cross-sectional study. J Clin Med 2024 Mar 13;13(6):1631. [doi: 10.3390/jcm13061631] [Medline: 38541857]
10. Malmstrom TK, Morley JE. Frailty and cognition: linking two common syndromes in older persons. J Nutr Health Aging 2013 Sep;17(9):723-725. [doi: 10.1007/s12603-013-0395-y] [Medline: 24154641]
11. Ruiz M, Cefalu C. Characteristics of frail patients in a geriatric-HIV program: the experience of an urban academic center at one year follow-up. J Int Assoc Physicians AIDS Care (Chic) 2011;10(3):138-143. [doi: 10.1177/1545109711399658] [Medline: 21502439]
12. Zamudio-Rodríguez A, Belaunzarán-Zamudio PF, Sierra-Madero JG, et al. Association between frailty and HIV-associated neurodegenerative disorders among older adults living with HIV. AIDS Res Hum Retroviruses 2018 May;34(5):449-455. [doi: 10.1089/AID.2017.0100] [Medline: 29385834]
13. Lenz ER, Suppe F, Gift AG, Pugh LC, Milligan RA. Collaborative development of middle-range nursing theories: toward a theory of unpleasant symptoms. ANS Adv Nurs Sci 1995 Mar;17(3):1-13. [doi: 10.1097/00012272-199503000-00003] [Medline: 7778887]
14. Lenz ER, Pugh LC, Milligan RA, Gift A, Suppe F. The middle-range theory of unpleasant symptoms: an update. ANS Adv Nurs Sci 1997 Mar;19(3):14-27. [doi: 10.1097/00012272-199703000-00003] [Medline: 9055027]
15. WS 293--2019: diagnosis for HIV/AIDS [Report in Chinese]. : National Health Commission of the People's Republic of China; 2019 URL: http://www.nhc.gov.cn/wjw/s9491/201901/7b1e6a7242084a67993f01cc14f8e960/files/1739781460870_69990.pdf [accessed 2026-01-13]
16. Lü CR, Zhao X, Wu XL, et al. Frailty and cognitive function status in elderly inpatients with AIDS. Journal of Wannan Medical College 2019;38(6):589-592 [FREE Full text]
17. The R Project for Statistical Computing. URL: https://www.R-project.org [accessed 2026-01-13]
18. Ye Q, Lu QL, Hu PW, et al. Frailty status and influencing factors in people living with human immunodeficiency virus infection/acquired immunodeficiency syndrome. Chin J Infect Control 2024;4:522-526. [doi: 10.12138/j.issn.1671-9638.20245347]
19. Iriarte E, Cianelli R, De Santis JP, et al. Multidimensional frailty, quality of life and self-management in aging Hispanics living with HIV. J Appl Gerontol 2024 Jul;43(7):899-909. [doi: 10.1177/07334648231211743] [Medline: 38173356]
20. Gobbens RJJ, van Assen MALM, Luijkx KG, Wijnen-Sponselee MT, Schols JMGA. The Tilburg Frailty Indicator: psychometric properties. J Am Med Dir Assoc 2010 Jun;11(5):344-355. [doi: 10.1016/j.jamda.2009.11.003] [Medline: 20511102]
21. Li F, Liu HS, Zha LX, et al. Reliability and validity testing of the Chinese version of Tiburg Frailty Indicator. J Nurs Sci 2017;32(7):18-20 [FREE Full text]
22. Krankowska DC, Załęski A, Wiercińska-Drapało A. Frailty and prefrailty in people living with HIV, with focus on women living with HIV. Int J STD AIDS 2022 Nov;33(13):1106-1110. [doi: 10.1177/09564624221127744] [Medline: 36217985]
23. Ruderman SA, Odden MC, Webel AR, et al. Tobacco smoking and pack-years are associated with frailty among people with HIV. J Acquir Immune Defic Syndr 2023 Oct 1;94(2):135-142. [doi: 10.1097/QAI.0000000000003242] [Medline: 37368939]
24. Gustafson DR, Shi Q, Thurn M, et al. Frailty and constellations of factors in aging HIV-infected and uninfected women—The Women's Interagency HIV Study. J Frailty Aging 2016;5(1):43-48. [doi: 10.14283/jfa.2016.79] [Medline: 26980368]
25. Bellelli F, Consorti E, Hettiarachchige TMK, et al. Relationship among age, education and frailty in older persons. J Frailty Aging 2023 Oct;12(4):326-328. [doi: 10.14283/jfa.2023.39] [Medline: 38008985]

XSL•FO
RenderX

26. Melo GC, Carvalho ACA, Mendes MLT, et al. Association between frailty phenotype, quantification of plasma HIV-1 RNA, CD4 cell count and HAART in HIV-positive subjects: a systematic review and meta-analysis of observational studies. AIDS Care 2022 Sep;34(9):1159-1168. [doi: 10.1080/09540121.2021.1956414] [Medline: 34292108]
27. Iriarte E, Cianelli R, De Santis JP, et al. Factors related to multidimensional frailty among Hispanic people living with HIV aged 50 years and above: a cross-sectional study. J Assoc Nurses AIDS Care 2023;34(3):259-269. [doi: 10.1097/JNC.0000000000000398] [Medline: 36917646]
28. Gustafson DR, Shi Q, Thurn M, et al. Frailty-related factors among women living with and without HIV aged 40 years and older: the Women's Interagency HIV Study. J Frailty Aging 2024;13(1):40-49. [doi: 10.14283/jfa.2023.41] [Medline: 38305442]
29. Woldesemayat EM, St Clair-Sullivan N, Kassa A, et al. Frailty status and associated factors among older PLHIV in Southern Ethiopia. PLoS ONE 2023;18(4):e0284376. [doi: 10.1371/journal.pone.0284376] [Medline: 37093810]
30. Mwangala PN, Nasambu C, Wagner RG, Newton CR, Abubakar A. Prevalence and factors associated with frailty among older adults living with HIV compared to their uninfected peers from the Kenyan coast. Int J Public Health 2024;69:1606284. [doi: 10.3389/ijph.2024.1606284] [Medline: 38426187]
31. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213. [doi: 10.1016/0165-1781(89)90047-4] [Medline: 2748771]
32. Inceer M, Mayo N. Health-related quality of life measures provide information on the contributors, components, and consequences of frailty in HIV: a systematic mapping review. Qual Life Res 2024 Jul;33(7):1735-1751. [doi: 10.1007/s11136-024-03613-3] [Medline: 38462582]
33. Rubenstein LZ, Harker JO, Salvà A, Guigoz Y, Vellas B. Screening for undernutrition in geriatric practice: developing the short-form mini-nutritional assessment (MNA-SF). J Gerontol A Biol Sci Med Sci 2001 Jun;56(6):M366-M372. [doi: 10.1093/gerona/56.6.m366] [Medline: 11382797]
34. He YL, Jian ZJ. The application and evaluation of the elderly malnutrition assessment methods [Article in Chinese]. Chin J Geriatr 2005;4:37-40 [FREE Full text]
35. Masters MC, Perez J, Wu K, et al. Baseline neurocognitive impairment (NCI) is associated with incident frailty but baseline frailty does not predict incident NCI in older persons with human immunodeficiency virus (HIV). Clin Infect Dis 2021 Aug 16;73(4):680-688. [doi: 10.1093/cid/ciab122] [Medline: 34398957]
36. Erlandson KM, Wu K, Koletar SL, et al. Association between frailty and components of the frailty phenotype with modifiable risk factors and antiretroviral therapy. J Infect Dis 2017 Mar 15;215(6):933-937. [doi: 10.1093/infdis/jix063] [Medline: 28453849]
37. Sacktor NC, Wong M, Nakasujja N, et al. The International HIV Dementia Scale: a new rapid screening test for HIV dementia. AIDS 2005 Sep 2;19(13):1367-1374. [Medline: 16103767]
38. Skinner S, Adewale AJ, DeBlock L, Gill MJ, Power C. Neurocognitive screening tools in HIV/AIDS: comparative performance among patients exposed to antiretroviral therapy. HIV Med 2009 Apr;10(4):246-252. [doi: 10.1111/j.1468-1293.2008.00679.x] [Medline: 19187172]
39. Brañas F, Jiménez Z, Sánchez-Conde M, et al. Frailty and physical function in older HIV-infected adults. Age Ageing 2017 May 1;46(3):522-526. [doi: 10.1093/ageing/afx013] [Medline: 28203694]
40. Li MJ, Chen R, Li AL, et al. Analysis on weakness status and influencing factors of HIV/AIDS among elderly individuals in Luzhou city [Article in Chinese]. Prev Med Tribune 2024;30(3):213-218 [FREE Full text]
41. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [doi: 10.1111/j.1600-0447.1983.tb09716.x] [Medline: 6880820]
42. Xie NH, Yan H, Ding J, Wang X. Reliability and validity analysis of the hospital anxiety and depression scale among HIV/AIDS patients [Article in Chinese]. Chin J AIDS STD 2020;26(12):1328-1331 [FREE Full text]
43. Xiao SY. Theoretical basis and research application of the Social Support Rating Scale [Article in Chinese]. J Clin Psychiatry 1994;2:98-100 [FREE Full text]
44. Rueda S, Law S, Rourke SB. Psychosocial, mental health, and behavioral issues of aging with HIV. Curr Opin HIV AIDS 2014 Jul;9(4):325-331. [doi: 10.1097/COH.0000000000000071] [Medline: 24824890]
45. Rubtsova AA, Marquine MJ, Depp C, et al. Psychosocial correlates of frailty among HIV-infected and HIV-uninfected adults. Behav Med 2019;45(3):210-220. [doi: 10.1080/08964289.2018.1509053] [Medline: 30431403]
46. Huang H, Tu L, Zhang X, et al. Social support, self-efficacy, self-esteem, and self-management behaviors among people living with HIV/AIDS in China: a structural equation modeling analysis. BMC Public Health 2024 Nov 6;24(1):3070. [doi: 10.1186/s12889-024-20541-6] [Medline: 39508244]
47. Cybularz M, Wydra S, Berndt K, et al. Frailty is associated with chronic inflammation and pro-inflammatory monocyte subpopulations. Exp Gerontol 2021 Jul 1;149:111317. [doi: 10.1016/j.exger.2021.111317] [Medline: 33744391]
48. Li N, Liu G, Gao H, et al. Geriatric syndromes, chronic inflammation, and advances in the management of frailty: a review with new insights. Biosci Trends 2023 Sep 15;17(4):262-270. [doi: 10.5582/bst.2023.01184] [Medline: 37612125]
49. Heinze-Milne SD, Banga S, Howlett SE. Frailty and cytokines in preclinical models: comparisons with humans. Mech Ageing Dev 2022 Sep;206:111706. [doi: 10.1016/j.mad.2022.111706] [Medline: 35835224]

XSL•FO
RenderX

50. Mailliez A, Guilbaud A, Puisieux F, Dauchet L, Boulanger É. Circulating biomarkers characterizing physical frailty: CRP, hemoglobin, albumin, 25OHD and free testosterone as best biomarkers. Results of a meta-analysis. Exp Gerontol 2020 Oct 1;139:111014. [doi: 10.1016/j.exger.2020.111014] [Medline: 32599147]
51. Lü KY, Liao XL, Xiong WQ, et al. Incidence and influencing factors of frailty in HIV/AIDS patients: a meta-analysis [Article in Chinese]. Chin J AIDS STD 2023;29(8):948-953 [FREE Full text]
52. Horberg M, Thompson M, Agwu A, et al. Primary care guidance for providers of care for persons with human immunodeficiency virus: 2024 update by the HIV Medicine Association of the Infectious Diseases Society of America. Clin Infect Dis 2024 Oct 12:ciae479. [doi: 10.1093/cid/ciae479] [Medline: 39393187]
53. Wang B, Peng X, Lu Y, et al. Association between frailty and low sexual function among sexually active older adults living with HIV in China: a multi-centre cross-sectional study in China. AIDS Care 2025 Sep;37(9):1539-1550. [doi: 10.1080/09540121.2025.2480151] [Medline: 40153538]
54. Li J, Chen X, Lin H, et al. Associations between HIV infection and frailty status and its individual components: are frailty components disproportionally affected? HIV Med 2023 May;24(5):533-543. [doi: 10.1111/hiv.13429] [Medline: 36288971]
55. Sun W, Chen WT, Pan Y, Lu X, Zhang L, Arbing R. Social isolation and frailty among older people living with HIV in China: mediation effects of psychological distress. J Gerontol Nurs 2025 Nov;51(11):48-57. [doi: 10.3928/00989134-20250905-03] [Medline: 40953374]
56. Li M, Xu Y, Zou J, et al. Analysis of risk factors of social frailty in older adults living with HIV/AIDS. Sci Rep 2025 Apr 14;15(1):12769. [doi: 10.1038/s41598-025-97466-6] [Medline: 40229403]
57. Han WM, Wu K, Tassiopoulos K, et al. Inflammation in frailty, cognitive impairment, clinical events, and mortality among older adults with HIV in the ACTG HAILO cohort. AIDS 2025 Dec 1;39(15):2218-2228. [doi: 10.1097/QAD.0000000000004324] [Medline: 40811075]
58. Ngwa DN, Pathak A, Agrawal A. IL-6 regulates induction of C-reactive protein gene expression by activating STAT3 isoforms. Mol Immunol 2022 Jun;146:50-56. [doi: 10.1016/j.molimm.2022.04.003] [Medline: 35430542]
59. Li Z, Morano JP, Khoshnood K, Hsieh E, Sheng Y. HIV-related stigma among people living with HIV/AIDS in rural Central China. BMC Health Serv Res 2018 Jun 15;18(1):453. [doi: 10.1186/s12913-018-3245-0] [Medline: 29903006]
60. Long S, Yao R, Tang Y, et al. Pre-treatment and acquired HIV drug resistance in Hunan: a cross-sectional analysis of a growing challenge. BMC Infect Dis 2025 Oct 23;25(1):1393. [doi: 10.1186/s12879-025-11656-1] [Medline: 41131459]
61. Xia H, Jin J, Ba H, et al. Genetic diversity and characteristics of drug resistance among treatment-naive people living with HIV in Xi'an, China. Drug Des Devel Ther 2023;17:1485-1494. [doi: 10.2147/DDDT.S406255] [Medline: 37220545]
62. Lan Y, Li L, He X, et al. Transmitted drug resistance and transmission clusters among HIV-1 treatment-naïve patients in Guangdong, China: a cross-sectional study. Virol J 2021 Sep 6;18(1):181. [doi: 10.1186/s12985-021-01653-6] [Medline: 34488793]
63. Zhang M, Ma Y, Wang Z, et al. Prevalence and transmission of pretreatment drug resistance in people living with HIV-1 in Shanghai China, 2017-2021. Virulence 2024 Dec;15(1):2373105. [doi: 10.1080/21505594.2024.2373105] [Medline: 38934465]
64. Xu Y, Shi H, Dong X, et al. Transmitted drug resistance and transmission clusters among ART-naïve HIV-1-infected individuals from 2019 to 2021 in Nanjing, China. Front Public Health 2023;11:1179568. [doi: 10.3389/fpubh.2023.1179568] [Medline: 37674678]

Abbreviations

AUC: area under the curve
CRP: C-reactive protein
HAND: HIV-associated neurocognitive disorders
IL-6: interleukin-6
LR: logistic regression
PSQI: Pittsburgh Sleep Quality Index
RF: random forest
SVM: support vector machine
TOUS: theory of unpleasant symptoms

XSL•FO
RenderX

Edited by J Sarvestan; submitted 17.Sep.2025; peer-reviewed by L Liao; revised version received 19.Dec.2025; accepted 22.Dec.2025; published 19.Feb.2026.

Please cite as:
Li F, Xie C, Xiang F
Analysis of Associated Factors and Construction of Risk Prediction Models for Frailty in Hospitalized Older Adults Living With HIV: Protocol for a Prospective Observational Study
JMIR Res Protoc 2026;15:e84271
URL: https://www.researchprotocols.org/2026/1/e84271
doi:10.2196/84271

© Fan Li, Chaoying Xie, Fang Xiang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Stereotactic Adaptive Radiation Therapy for Borderline Resectable or Locally Advanced Pancreatic Cancer to Minimize Gastrointestinal Toxicity (ARTIA-Pancreas): Protocol for a Single-Arm Prospective Trial

Lauren E Henke[1], MS, MD; Hyun Kim[2], MD; Eric Laugeman[2], MS; Steven Kohlmyer[3], MS; Claire McCann[3], PhD; Kate Pietrovito[3], MS; Kenneth Russell[3], MD; Jennifer Woo[3], BSc; Alex T Price[1], PhD, DABR

[1]Department of Radiation Oncology, Seidman Cancer Center, University Hospitals, 10900 Euclid Avenue, Cleveland, OH, United States

[2]Department of Radiation Oncology, Siteman Cancer Center, Washington University School of Medicine, St. Louis, MO, United States

[3]Varian Medical Systems, LLC, A Siemens Healthineers Company, Palo Alto, CA, United States

Corresponding Author:
Alex T Price, PhD, DABR
Department of Radiation Oncology, Seidman Cancer Center, University Hospitals, 10900 Euclid Avenue, Cleveland, OH, United States

Abstract

Background: Computed tomography (CT)–guided online stereotactic adaptive radiotherapy (CT-STAR) allows for ablative radiation doses to be delivered to selected patients with borderline resectable (BR) or locally advanced pancreatic cancer (LAPC) or unresectable pancreatic cancer. However, the use of CT-STAR to deliver an ablative dose to the pancreas while minimizing gastrointestinal (GI) side effects to reduce acute and late toxicity rates compared to historic controls has yet to be prospectively evaluated.

Objective: The primary objective of this prospective, single-arm, multicenter phase 2 clinical trial (Adaptive Radiation Therapy Individualized Approach [ARTIA]–Pancreas) is to evaluate the rate of acute grade 3+ GI toxicity in patients with BR/LAPC treated with ablatively dosed CT-STAR compared to historical controls.

Methods: Patients with histologically or cytologically confirmed BR, locally advanced, or medically inoperable pancreatic adenocarcinoma are eligible for participation. Consenting and eligible patients will be treated with CT-STAR, delivering 5 fractions over 1 to 2 weeks with daily adaptation based on anatomic changes observed with onboard cone beam CT imaging. The primary end point of this trial is the rate of acute patient-reported grade 3+ treatment-related GI toxicities, assessed at 90 days post CT-STAR and compared to a historical control rate of 20%. The key powered secondary end point is the rate of long-term patient-reported grade 3+ treatment-related GI toxicities, evaluated at 12 months post CT-STAR and compared to a historical control rate of 25%. Additional secondary end points include overall survival, local (in-field) control rates, and distant progression-free survival at 1 and 2 years post CT-STAR.

Results: Study completion is anticipated in February 2029, and the final study results will be published upon completion of the study.

Conclusions: ARTIA-Pancreas represents the first prospective phase 2 clinical trial to evaluate whether CT-STAR can reduce the rate of acute patient-reported GI toxicities in patients with BR/LAPC compared to historical controls. Findings from this clinical trial will provide evidence for safely and effectively incorporating ablatively dosed adaptive radiotherapy into treatment regimens for this population.

Trial Registration: ClinicalTrials.gov NCT05764720; https://clinicaltrials.gov/study/NCT05764720

International Registered Report Identifier (IRRID): PRR1-10.2196/84607

(JMIR Res Protoc 2026;15:e84607) doi:10.2196/84607

KEYWORDS

adaptive radiation; radiation; pancreatic cancer; pancreas; dose escalation

XSL•FO
RenderX

Introduction

The only known curative intervention for patients with pancreatic cancer is surgical resection [1]. However, only approximately 20% of patients are clearly eligible for curative intent resection at the time of diagnosis, while 30% to 40% of patients present with either borderline resectable or locally advanced disease [2]. In patients with borderline resectable (BR) or locally advanced pancreatic cancer (LAPC), neoadjuvant chemotherapy and/or chemoradiotherapy are typically used for the purpose of tumor downstaging and affording the opportunity for curative resection [3]. Despite neoadjuvant treatment, only 20% to 30% of patients with BR/LAPC are typically converted to resectability, highlighting the need for improvements in these chemo- and radiotherapies to contribute to or provide definitive therapy [3]. Computed tomography (CT)–guided stereotactic adaptive radiotherapy (CT-STAR), a recent technological advancement in the radiotherapeutic space, may overcome the shortcomings of conventional radiotherapy to improve care for BR/LAPC.

Historically, conventional radiotherapy has shown only modest use in patients with pancreatic cancer, with randomized phase 3 clinical trials demonstrating a lack of survival benefit and only modest local control gains in patients treated with conventional chemoradiotherapy versus chemotherapy alone [4,5]. Dose-escalated conventional radiotherapy may convert 20% to 30% of unresectable patients to resectability [3] but imparts significant side effects, with rates of acute grade 3+ nonhematologic and gastrointestinal (GI) toxicity as high as 20% to 30% [3,4,6,7].

Stereotactic body radiotherapy (SBRT), which leverages fewer treatment days and higher dose conformality compared to conventional radiotherapy, has emerged as an alternative approach for achieving local disease control in patients with pancreatic cancer. However, ablative CT-guided pancreatic SBRT has been shown to induce similarly (>20%) concerning rates of grade 3+ GI toxicity [8] when target coverage is prioritized over organ-at-risk (OAR) protection. In patients with unresectable pancreatic cancer, studies suggest that reducing treatment margins and building in constrained planning-organ-at-risk volumes may partly mitigate acute GI toxicity following ablative SBRT [9-11]. Even so, these approaches do not consistently lead to meaningful reductions in late GI toxicity [12] and/or result in intentional limitation of tumor coverage. Therefore, historical use of ablative SBRT has been limited to near-palliative dosing regimens [13] or to use in patients with highly selective anatomic presentations [9,10,14].

Online adaptive radiotherapy (ART), a form of radiotherapy in which interfraction changes to the radiation dosimetry plan are made in response to observed changes in tumor and OAR anatomy, has emerged as a feasible and safe mechanism to deliver ablative SBRT to patients with abdominal cancer [15,16]. In patients with tumor locations with direct OAR abutment or invasion that would traditionally preclude treatment with ablative SBRT, online adaptive planning and delivery may enable safe delivery of ablative SBRT [17]. Findings from both a small, multi-institutional retrospective analysis and a recent phase 2 clinical trial suggest that dose escalation enabled by ART may improve 2-year overall survival while reducing grade 3+ GI toxicity in patients with BR/LAPC [18-20]. Furthermore, single-institution series and prospective analyses indicate that daily ART with magnetic resonance (MR) guidance may allow for safe incorporation of elective nodal irradiation (ENI) volumes with ablative doses [21,22], which is desirable per current treatment guidelines [23], but historically limited due to concerns of toxicity [24,25].

The delivery of ART is challenged by the need for high-quality daily on-board imaging, which is not achievable through conventional CT and has historically required MR-guided adaptive platforms. MR-guided adaptive platforms are costly and require specialized resources, restricting clinic access to ART from both financial and resource perspectives [26]. Recent technological improvements to conventional cone beam CT (CBCT) now permit CT guidance of ART using a ring-gantry linac (Ethos, Varian Medical Systems) [27]. Specifically, images acquired via an advanced onboard kV-CBCT technology have been found to approach the quality of diagnostic CT simulation images, allowing high-quality visualization and delineation of targets and OARs, providing a more accessible alternative to MR platforms [28].

Currently, there are limited data on the potential of CT-STAR to deliver ablative SBRT to the pancreas and ENI volumes. Washington University prospectively evaluated, in silico, the feasibility of using the kV-CBCT-based STAR system for delivery of ART to patients with unresectable pancreatic cancer [29]. CT-STAR was found to be feasible, with successful workflow and creation of online adaptive pancreas SBRT plans that met strict OAR constraints in 97.5% of fractions [29]. CT-STAR was also found to be dosimetrically advantageous, and subsequently, the first pilot study (N=8 participants) of CT-STAR for ablative treatment of pancreatic cancer was reported [30].

The aim of the Adaptive Radiation Therapy Individualized Approach (ARTIA)-Pancreas is to conduct a large (N=134 participants), prospective, single-arm, multicenter phase 2 clinical trial to evaluate the efficacy of CT-STAR in delivering ablative radiotherapy dose for BR/LAPC while reducing the rate of acute grade 3+ GI toxicity compared to historical controls (NCT05764720; *A Prospective Trial of Stereotactic Adaptive Radiation Therapy for Borderline Resectable/Locally Advanced Pancreatic Cancer: An Individualized Approach to Minimizing Gastrointestinal Toxicity, ARTIA-Pancreas*).

Methods

Study Objectives and Design Overview

The primary objective of ARTIA-Pancreas is to evaluate the rate of acute (90-d) grade 3+ GI toxicity resulting from ablatively dosed CT-STAR administered to patients with BR/LAPC compared to a historical rate of 20%. It is hypothesized that CT-STAR will reduce acute grade 3+ GI toxicity in this patient population, with an effect size of 10%.

XSL•FO
RenderX

ARTIA-Pancreas is also hierarchically powered to evaluate the key secondary objective of rate of late (12 mo) grade 3+ GI toxicity from ablatively dosed CT-STAR administered to patients with BR/LAPC compared to historical rates. We hypothesize that CT-STAR will reduce late grade 3+ GI toxicity, with a 15% effect size, from a 25% historical rate. Other secondary end points that will be evaluated include (1) overall survival at 12 and 24 months post completion of CT-STAR, (2) local (in-field) control rates at 12 and 24 months post completion of CT-STAR, and (3) distant progression-free survival at 12 and 24 months post completion of CT-STAR.

Exploratory objectives, along with associated end points, of ARTIA-Pancreas are listed in Table S1 in Multimedia Appendix 1.

The design of this prospective, single-arm phase 2 clinical trial is depicted in Figure 1. Briefly, potential participants are screened for eligibility according to inclusion and exclusion criteria (see "Participant Eligibility and Enrollment" section), and informed consent is obtained from eligible participants. Chemotherapy is suspended 1 to 2 weeks prior to CT-STAR treatment, and baseline assessments are conducted by the treatment team (Table 1). During the course of ablatively dosed CT-STAR administration (50 Gy in 5 fractions), physicians review the daily ART plans and perform interim assessments (Table 1) of GI toxicity.

One to two weeks after the completion of CT-STAR, chemotherapy may be reinitiated, and follow-up assessments are conducted in accordance with study end points (Table 1).

Figure 1. Adaptive Radiation Therapy Individualized Approach (ARTIA)–Pancreas study schema. CTCAE: Common Terminology Criteria for Adverse Events; EORTC: European Organization for Research and Treatment of Cancer; SBRT: stereotactic body radiotherapy; QLQ C30: Core Quality of Life Questionnaire; PAN-26: Pancreatic Cancer Module for Assessing Health Related Quality of Life.

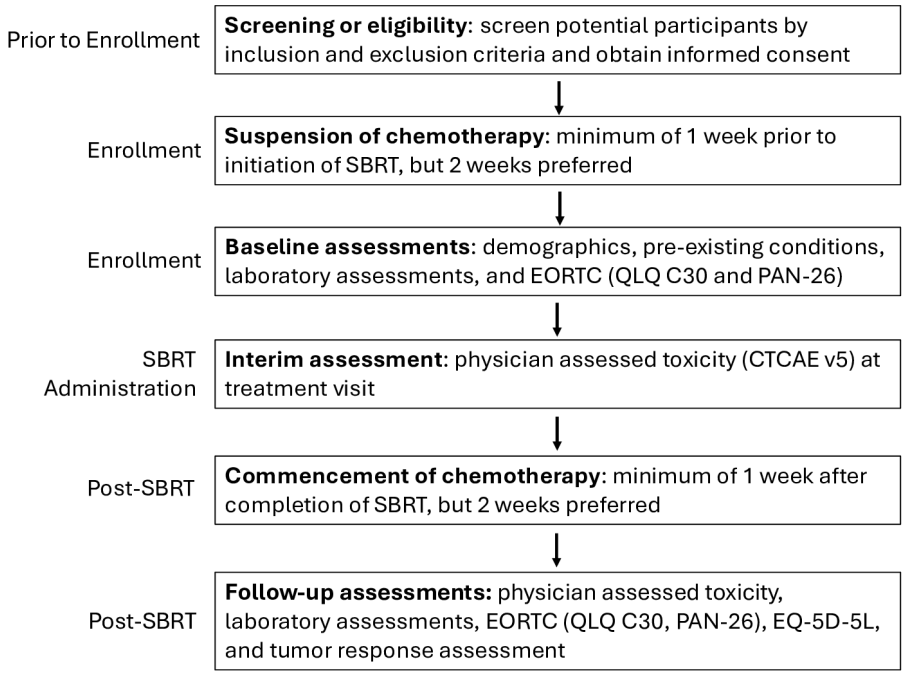

XSL•FO
RenderX

Table . Study assessment timeline and frequency.

Data required	0	Pre	RT[a] 1 - 2	Post 1 week	Post 6 week	Post 3 months	Post 6 months	Post 12 months	Post 18 months	Post 24 months
Consent	✓									
Patient demographics	✓									
PMH[b]	✓									
Volume imaging	✓					✓	✓	✓	✓	✓
RECIST[c]	✓					✓	✓	✓	✓	✓
CBC[d]	✓		✓	✓	✓	✓	✓	✓		
CMP[e]	✓		✓	✓	✓	✓	✓	✓		
CA[m] 19 - 9	✓		✓	✓	✓	✓	✓	✓		
Pregnancy	✓									
RT planning		✓								
CT-STAR[g]			✓							
MD review[h]			Each fx							
AE[i] assessment			OTV[j]		✓	✓	✓	✓	✓	✓
EORTC[k]		✓			✓	✓	✓	✓		
EQ-5D-5L		✓			✓	✓	✓	✓		

[a]RT: radiation therapy.

[b]PMH: past medical history.

[c]RECIST: Response Evaluation Criteria for Solid Tumors.

[d]CBC: complete blood count.

[e]CMP: comprehensive metabolic panel.

[f]CA: carbohydrate antigen.

[g]CT-STAR: computed tomography–guided online stereotactic adaptive radiotherapy.

[h]MD review: real-time physician review of adapted plan.

[i]AE: adverse event.

[j]OTV: on treatment visit.

[k]EORTC: European Organization for Research and Treatment of Cancer.

Site Credentialing and Quality Assurance

To maintain study rigor and quality across participating institutions, sites are credentialed prior to activation and required to upload planning images for the first two patients for central peer review. As a part of site credentialing, 2 benchmark cases (pancreas head and pancreatic tail tumor locations) consisting of a simulation CT and an Ethos CBCT are provided to the site, which will then prepare and submit the reference dosimetry plan as well as an adapted plan for approval by a centralized review committee for both clinical presentations.

After trial site activation, sites upload the simulation CT and pretreatment plan for the first 2 enrolled participants for central peer review 4 days prior to the planned start of external beam radiation therapy. Within 1 business day of completion of the participants' first CT-STAR fraction, sites upload the deidentified daily fraction digital imaging and communications in medicine (DICOM)-RT data for clinical accuracy and protocol adherence review by the centralized committee. Central peer review is repeated every 5 subsequent participants or if a site has not enrolled a participant in 6 months.

Participant Eligibility and Enrollment

Patients aged 18 years or older with histologically or cytologically confirmed BR, locally advanced, or medically inoperable (based on the National Comprehensive Cancer Network criteria) pancreatic adenocarcinoma are eligible for participation in this phase 2 clinical trial. Key exclusion criteria include a history of radiotherapy within the projected treatment field or prior, intercurrent, or planned treatment with an investigational agent related to pancreatic cancer diagnosis

XSL•FO
RenderX

within 90 days of CT-STAR. Patients of all genders, races, and ethnic groups are eligible for ARTIA-Pancreas. Additional inclusion and exclusion criteria are listed in Textbox 1.

Textbox 1. Additional participant inclusion and exclusion criteria.

Inclusion criteria

- Eastern Cooperative Oncology Group performance status 0‐1.
- Receipt of at least 2 months of lead-in chemotherapy (regimen at the discretion of the treating physician) prior to planned initiation of computed tomography–guided online stereotactic adaptive radiotherapy (CT-STAR).
- Limited regional lymphadenopathy. Node+ patients are restricted to those with up to 3 clinically involved (≥1 cm on cross-sectional imaging or pathologically proven) nodes provided that the lymph nodes are adjacent to the primary tumor.
- Able to interrupt systemic therapy at least 1 week prior to planned start of CT-STAR (2 wk preferred) lasting for the duration of CT-STAR and continuing for at least 1 week following end of CT-STAR (2 wk preferred).
- Capable of single end-exhale breath-hold of at least 20 seconds in duration and of repeated end-exhale or deep inspiratory breath-hold of at least 10 seconds in duration upon verbal instruction.
- Anatomy of target and adjacent organs-at-risk adequately visualized on Ethos simulation imaging, as determined by treating and study physicians.
- Able to understand and willing to sign an institutional review board–approved written informed consent document (or that of legally authorized representative, if applicable).

Exclusion criteria

- Any additional cancer diagnosis made and/or treated within the preceding year (other than nonmelanoma skin cancer).
- Any history of prior malignancy (independent of time frame), with evidence of persistent or progressive disease (other than nonmelanoma skin cancer).
- Uncontrolled intercurrent illness, including but not limited to ongoing or active infection, symptomatic congestive heart failure, unstable angina pectoris, cardiac arrhythmia, or psychiatric illness or social situations that would limit compliance with study requirements.
- Pregnant and/or breastfeeding (patients of childbearing potential must have a negative pregnancy test within 14 d of study entry).
- Gross tumor invasion of the stomach or duodenum (defined either radiographically or endoscopically).

Ethical Considerations

The protocol was approved by the University Hospitals Institutional Review Board (IRB) of the University Hospitals Cleveland Medical Center (protocol STUDY20240615). All participants are required to give written informed consent, indicated by a dated signature of the participant or legal representative, in compliance with International Organization for Standardization (ISO) 14155:2020 (WCG protocol 20225669). Personnel delegated by the principal investigator to conduct the informed consent process must have documented training in both good clinical practice and the study protocol. Patients are compensated ($1,000 - $1,300 USD) for trial participation, which is dependent on the site of enrollment and visit completion. Authorization to collect research data from participants' medical records is obtained at the time of informed consent. Only deidentified data are entered into the electronic database where data are stored, and participants will be represented by a unique identifier in the database. All efforts are made to remove participant-identifying information and to deidentify data before entrance into the study database or transmitting the data to the sponsor. All reasonable efforts are made to protect the privacy of participants.

Baseline Assessments

Baseline assessments include patient demographics, past medical history, complete blood count, pregnancy test, and comprehensive metabolic panel. In addition, physicians evaluate the patient's oncologic status through carbohydrate antigen 19‐9 testing, Response Evaluation Criteria for Solid Tumors scoring, and volumetric imaging of tumors. Quality of life is also evaluated through the patient-reported European Organization for Research and Treatment of Cancer and EQ-5D-3L questionnaires.

Pretreatment Anatomy Assessment and Radiotherapy Planning

To assess whether participants' anatomy can be visualized adequately with CBCT adaptive treatment, all patients undergo both a CT simulation consisting of an end-exhale CT and a 4D CT, as well as an Ethos CBCT. Participants will be in the supine position, and all imaging extends from the superior aspect of the dome of the liver to the iliac crest. Participants with adequately visualized anatomy on CBCT then proceed to dosimetric treatment planning.

Investigators are provided with a contour atlas, which shows example contours for primary tumors in the head or neck and tail of the pancreas as well as for accompanying ENI. Target and normal structure contouring guidelines were developed to meet the clinical goals of ARTIA-Pancreas and are described for gross tumor volume_5000, clinical target volume_3300, planning target volume (PTV)_5000, and PTV_3300 as well as for the stomach, duodenum, small bowel, large bowel, spinal cord, left and right kidney, and liver.

Dose-volume histogram information for the target volumes and surrounding critical structures is mandatory. Minimum PTV coverage goals are 95% of the volumes to be covered by 25 Gy,

XSL•FO
RenderX

with a reduction in PTV dose allowed in areas of OAR and target overlap. The initial treatment plan is created such that the plan will normalize to all priority 1 OAR (eg, stomach, duodenum, small and large bowel, spinal cord, and solitary kidney) objectives (Table 2). As such, OAR dose constraints function as strict constraints in treatment planning for both initial and adaptive (CT-STAR) plans, and gross tumor volume, PTV_5000, and PTV_3300 coverage is sacrificed, if necessary, to meet the OAR dose constraints.

Table . Organ-at-risk and target planning objectives, goals, and suggested planning weight priorities for the Ethos 2.0 treatment planning system optimization algorithm.

Name of structure	Suggested planning priority	Dosimetric parameter	Per protocol	Variation acceptable[a]
Stomach	• 1 • 4	• V3300 cGy • V3000 cGy	• ≤0.5 cc • ≤5.0 cc	• None[b]
Duodenum	• 1 • 4	• V3300 cGy • V3000 cGy	• ≤0.5 cc • ≤5.0 cc	• None
Small bowel	• 1	• V3300 cGy	• ≤0.5 cc	• None
Large bowel	• 1	• V3300 cGy	• ≤0.5 cc	• None
Spinal canal	• 1 • 3	• V2500 cGy • V1500 cGy	• ≤0.5 cc • ≤0.5 cc	• None • None
Kidneys	• 3 and 1	• Mean dose	• ≤1500 cGy	• ≤1800 cGy
Solitary kidney (if solitary kidney present)	• — • 1	• V1000 cGy • Mean dose	• ≤10% • ≤1000 cGy	• 45% • None
Kidney_L	• 3	• Mean dose	• ≤1500 cGy	• None
Kidney_R	• 3	• Mean dose	• ≤1500 cGy	• None
Uninvolved liver	• 2 • R • 1	• Mean dose • V2000 cGy • Critical volume of 700 cc	• ≤1500 cGy • ≤33% • ≤1500 cGy	• None • None • None

[a]Parameters not meeting the variation acceptable limits will lead to rescheduling of treatment for the following day.

[b]"None" means no acceptable variation permitted.

Dose Prescription and Delivery Considerations

Radiotherapy treatment consists of SBRT delivered in 5 fractions once daily or once every other day in 1 to 2 weeks. Participants receive a planned dose of 50 Gy in 5 fractions to the PTV with 33 Gy in 5 fractions to the ENI volume, with strict OAR sparing of luminal GI OARs prioritized over target coverage.

An exhale breath-hold approach using ancillary, commercially available respiratory management systems is used for respiratory gating. At the time of treatment, an end-exhale breath-hold CBCT daily alignment image is acquired to establish a new reference threshold based on the daily anatomy. The participant's respiratory motion is monitored using a respiratory trace generated and displayed to the treatment team for manually triggered breath-hold respiratory gating of radiation delivery. To monitor intrafraction participant motion, CBCT can be repeated a maximum of 10 times per fraction in the breath-hold position.

Daily Treatment Workflow and Assessments

For daily adaptation of SBRT, participants are imaged using CBCT at an end-exhale breath-hold position. The ART team uses clinical discretion to determine if the image quality is appropriate for target alignment and OAR delineation. If on 3 consecutive attempts prior to delivery of the first fraction, the CBCT image quality is poor, precluding organ visualization and adaptive planning, then the participant is removed from the study and pursues institutional standard-of-care treatment. Once adaptive planning optimization is complete, the adapted plan is compared to the scheduled plan, and any adapted plan or scheduled plan that does not meet the SBRT dose constraints is not delivered. If the adaptive plan meets all objectives and the scheduled plan does not, the adaptive plan is selected for treatment. If both plans meet SBRT dose constraints, the adapted plan is selected over the scheduled plan only if the PTV coverage goals increase by more than 5%.

In addition to the review of the adaptive CT-STAR plan at each fraction, adverse events are evaluated during treatment visits. Specifically, grade 2+ adverse events that are attributed by the physician as possibly, probably, or definitely related to ART

XSL•FO
RenderX

are recorded and classified per the Common Terminology Criteria for Adverse Events (CTCAE) v5.0.

Follow-Up Assessments

Participants are assessed at 6 weeks and at 3, 6, 12, 18, and 24 months after completion of CT-STAR treatment. CTCAE v5.0-evaluated adverse events are reported at all of these time points, while patient-reported quality of life is ascertained via the European Organization for Research and Treatment of Cancer and EQ-5D-3L at 6 weeks and 3-, 6-, and 12-months after CT-STAR. Volumetric imaging and tumor response, both having occurred at baseline, are repeated at 3, 6, 12, 18, and 24 months posttreatment to determine the patients' response to radiotherapy.

Study End Points

The primary end point for ARTIA-Pancreas is the rate of acute (90 d) grade 3+ treatment-related GI toxicities as defined by the CTCAE v5.0 and compared to the historical rate of 20%. The trial's key powered secondary end point is the rate of late (12-month) grade 3 or higher treatment-related GI toxicities as defined by the CTCAE v5.0 criteria and compared to the historical rate of 25%. Additional secondary end points include (1) a Kaplan-Meier estimate of the rates of overall survival at 12 and 24 months post completion of CT-STAR; (2) local (in-field) control rates at 12 and 24 months post completion of CT-STAR, defined as stable disease, partial response, or complete response by the Response Evaluation Criteria in Solid Tumors criteria; and (3) the Kaplan-Meier estimates of distant progression-free survival at 12 and 24 months post completion of CT-STAR.

Exploratory end points of ARTIA-Pancreas are listed in Table S1 in Multimedia Appendix 1.

Statistical Analysis

Primary End Point

ARTIA-Pancreas is powered to the primary end point. It is hypothesized that CT-STAR will reduce the rate of CTCAE v5.0 grade 3+ GI toxicity compared to non-ART, where the historical acute toxicity rate is 20% or higher. It is hypothesized that using CT-STAR, the acute toxicity rate will be <10% for an effect size of 10% or more. A total of 107 evaluable participants (ie, participants with nonmissing data for the primary end point) are targeted. There must be ≤13/107 (12.1%; 95% CI 7.2% - 19.7%) participants with acute grade 3+ GI toxicity in the study. With this rule, the study's 1-sided false positive error rate (α) is 0.023, and the power is 0.82. Given the possibility that participants could become eligible for surgical resection at the treated site following radiation, which would preclude toxicity assessment in those participants, we estimate a 20% attrition rate, and 134 participants will be targeted for enrollment.

CTCAE v5.0 grade 3+ GI toxicity compared to a historical rate of 20% will be analyzed by constructing a Wilson score 95% CI for the observed acute grade 3+ GI toxicity incidence rate. The CI's upper limit being <20% will indicate rejection of the historical rate of 20% at a 1-sided $P<.025$ significance level (ie, success for the study's primary end point).

Key Secondary End Point

ARTIA-Pancreas is also powered to analyze the key secondary end point, contingent upon a positive result for the primary end point. For the key secondary end point, it is hypothesized that CT-STAR will reduce late GI toxicity from a historical rate of 25% or greater to less than 10%. Given an evaluable 107 participants, there must be 17 (15.9%) participants or fewer with late grade 3+ GI toxicity in the study to reject a 25% or higher CT-STAR toxicity. With this rule, the key secondary end point's 1-sided false-positive error rate is 0.016 when the true late toxicity rate is 25%, and the power is 0.98 when the true late toxicity rate is 10%.

If the key secondary end point is indeed evaluated, a Wilson score 95% CI will be constructed for the study's observed 12-month grade 3+ GI toxicity incidence rate. The CI's upper limit being less than 25% will indicate rejection of the historical rate of 25% at a 1-sided P value less than .025 significance level (ie, success for the study's key secondary end point).

Secondary End Points

Twelve- and 24-month overall survival as well as distant progression-free survival will be evaluated by Kaplan-Meier analysis. All other secondary end points will be descriptively summarized.

Stopping Criteria

The sequential probability ratio test is being used to guard against an unacceptable toxicity rate during the study.

An expected toxicity rate of 10% and an unacceptable toxicity rate of 20% will be used with the total sample size of 134 participants. The probability of stopping early under the expected toxicity rate of 10% is set to .025, and the probability of stopping early when the true toxicity rate is 20% is set to .80.

Data Collection, Management, and Monitoring

Data collection includes obtaining data in the form of case report forms (electronic and/or paper), patient-reported outcomes (electronic and/or paper), DICOM images from CT simulation, DICOM-RT data from the pretreatment, daily fraction DICOM and DICOM-RT data, investigator feedback (email and/or written), observations from sponsor representatives (email or written), responses to data queries, and copies of pertinent records (ie, participant charts and laboratory data). All data are recorded on case report forms that are supported by source documentation. While no imputation methods will be used for missing outcomes, sensitivity analyses will be conducted evaluating results for different assumptions of toxicity levels among participants who are withdrawn from the study.

To monitor safety and efficacy signals in real time, study staff are required to report serious adverse events to the IRB according to IRB reporting requirements and to the study sponsor within 72 hours of becoming aware of the event. Likewise, radiotherapy device issues that occur during study performance and protocol deviations are documented and reported to the IRB and sponsor per reporting requirements. Per the stopping criteria defined above, the study can be suspended for review by the sponsor and lead investigators if unacceptable toxicity is suspected.

XSL•FO
RenderX

Results

The ARTIA-Pancreas trial opened for enrollment in May 2023. Primary completion is anticipated in February 2026, and study completion is anticipated in February 2029. The final study results will be published upon completion of the study.

Discussion

Overview

ARTIA-Pancreas is the first multicenter phase 2 clinical trial to evaluate the use of CT-STAR to minimize acute GI toxicity and result in favorable outcomes in patients with BR/LAPC. The study's primary objective is to evaluate the rate of high-grade GI toxicity at 90 days following a 5-fraction course of CT-STAR, with the hypothesis that adaptive SBRT will reduce acute grade 3+ toxicity from the historical control rate of 20% to 10%. Importantly, ARTIA-Pancreas has been strategically designed to hierarchically power the evaluation of the key secondary end point, 12-month treatment-related CTCAE v5.0 grade 3+ GI toxicity incidence rate, compared to a historical control rate of 25%.

Anticipated Results

Enrollment for ARTIA-Pancreas began in May 2023 and is anticipated to yield clinical findings shortly after completion in February 2028. The trial will provide evidence as to whether ablatively dosed radiotherapy that is adapted to the daily anatomy of each patient with BR/LAPC will reduce GI side effects. If, as is hypothesized, CT-STAR can reduce rates of treatment-related GI toxicity in this population, it may provide a rationale for the safe and effective incorporation of ablatively dosed CT-guided ART as a standard-of-care therapy.

Limitations

Historically, delivery of ablatively dosed SBRT has been difficult to safely achieve in the upper abdomen due to the proximity and sensitivity of adjacent OARs as well as inter- and intrafraction organ motion. The ARTIA-Pancreas trial seeks to overcome this limitation by treating patients with BR/LAPC using an adaptive SBRT strategy that enables safe dose escalation by reshaping the target dose in response to changes in OAR proximity and movement.

ARTIA-Pancreas has the same limitations as any trial using historical comparators, including potential time confounding related to changes in clinical practice or affected populations. Evidence generated by ARTIA-Pancreas may lay the foundation for potential future randomized trials directly comparing the safety and efficacy of CT-guided ART platforms to MR-guided ART platforms and standard-of-care in patients with BR/LAPC.

Should ARTIA-Pancreas yield positive findings concerning the use of CT-STAR in patients with BR/LAPC, there are several noteworthy considerations for generalization of these data. First, findings from ARTIA-Pancreas may not be generalizable to all patients with locally advanced pancreatic cancer due to the study's exclusion of participants with defined comorbidities as well as more advanced regional adenopathy. Second, despite being less resource-intensive than MR-guided ART, CT-guided ART still requires significant infrastructural, personnel, and financial commitments to successfully implement. While daily changes to patients' radiotherapy plans may allow for reduced treatment-related side effects, they often require additional staffing and resource utilization [31]. To address these barriers to CT-STAR implementation, logistical and workflow learnings from large-scale, multicenter trials such as ARTIA-Pancreas can be disseminated across the clinical community.

The success of ARTIA-Pancreas depends on the adequate imaging of tumor targets and OARs, which could be affected by variability in imaging quality across centers. Elements of the study that address this include assessment of patient anatomy visualization on CBCT for the determination of final eligibility, site credentialing, and central quality assurance reviews of participant imaging and ART planning at scheduled intervals.

Comparison With Prior Work

The historical framework for the use of radiotherapy in BR/LAPC is summarized in the "Introduction" section. The current clinical paradigm favors the use of ablative doses of SBRT with a strategy focused on minimizing adverse effects to adjacent normal tissues using daily adaptive techniques. In the recent phase I and retrospective studies, MR-based ART has been reported to improve the therapeutic index of SBRT for BR/LAPC, with encouraging local control rates of 70% to 80% at 2 years following treatment and a favorable toxicity profile [32]. While MR provides high-quality radiotherapy planning images, MR-guidance platforms require special construction, zoning, training, and staff presence that may impede widespread access to ART. In addition, patients may have MR imaging–incompatible medical devices or surgical implants that adversely impact the ability to deliver MR-guided ART.

ARTIA-Pancreas evaluates CT-, rather than MR-, guided ART [21,22] in an appropriately powered prospective clinical trial. The commercially available, advanced kV-CBCT imaging technology used in ARTIA-Pancreas can perform large field-of-view scans in the time of a single breath-hold and has demonstrated clinically acceptable target delineation and OAR daily anatomy visualization. Should the primary hypothesis of ARTIA-Pancreas be met, CT-guided ART may become a solution that achieves both clinical performance and potentially more widespread availability.

Conclusions

The results of ARTIA-Pancreas will provide clinical evidence to support health care professionals in decision-making regarding the use of daily adaptive CT-STAR in the treatment of BR/LAPC.

XSL•FO
RenderX

Acknowledgments

The authors gratefully acknowledge the contributions of Ping-Yu Liu, who contributed to the statistical plan for this clinical trial, and Lauren Pedersen, who contributed to the writing of this manuscript. Generative artificial intelligence was not used to write any section of this protocol.

Authors' Contributions

Conceptualization – LEH, ATP
Data curation – LEH, HK, EL, ATP
Funding acquisition – SK, CM, KP, KR, JW
Investigation – LEH, HK, EL, ATP
Methodology – LEH, ATP
Project administration – SK, CM, KP, KR, JW, LEH
Resources – LEH, HK
Supervision – LEH, KR, JW, ATP
Visualization – ATP
Writing – original draft – ATP, LEH
Writing – review & editing – ATP, LEH, EL, HK

Conflicts of Interest

Adaptive Radiation Therapy Individualized Approach (ARTIA)–Pancreas is funded by Varian Medical Systems, Inc, Palo Alto, California.
LEH reports grant funding to their institution from Varian Medical Systems/Siemens Healthineers, paid to the university; speaker honoraria and travel support for speaking; and consulting fees for clinical trial design, paid to self from Varian Medical Systems. LEH is also a clinical advisory board member for Elekta for adaptive radiotherapy, with fees paid to self.
EL reports grant funding from Varian Medical Systems, paid to the university, and speaker honoraria and travel funding from Varian Medical Systems, paid to self.
SK is an employee of Siemens Healthineers.
KP and JW are employees of Varian, a Siemens Healthineers Company.
ATP reports honorarium from Radformation, Inc., and stock options from Aitrium not related to this work.
HK reports research, travel, consulting, and speaker fees from Varian, as well as travel and consulting from Elekta.

Multimedia Appendix 1
Adaptive Radiation Therapy Individualized Approach (ARTIA)–Pancreas exploratory objectives and end points.
[DOCX File, 16 KB - resprot_v15i1e84607_app1.docx]

References

1. Doi R, Imamura M, Hosotani R, et al. Surgery versus radiochemotherapy for resectable locally invasive pancreatic cancer: final results of a randomized multi-institutional trial. Surg Today 2008;38(11):1021-1028. [doi: 10.1007/s00595-007-3745-8] [Medline: 18958561]
2. Hidalgo M. Pancreatic cancer. N Engl J Med 2010 Apr 29;362(17):1605-1617. [doi: 10.1056/NEJMra0901557] [Medline: 20427809]
3. Ben-Josef E, Schipper M, Francis IR, et al. A phase I/II trial of intensity modulated radiation (IMRT) dose escalation with concurrent fixed-dose rate gemcitabine (FDR-G) in patients with unresectable pancreatic cancer. Int J Radiat Oncol Biol Phys 2012 Dec 1;84(5):1166-1171. [doi: 10.1016/j.ijrobp.2012.02.051] [Medline: 22543215]
4. Hammel P, Huguet F, van Laethem JL, et al. Effect of chemoradiotherapy vs chemotherapy on survival in patients with locally advanced pancreatic cancer controlled after 4 months of gemcitabine with or without erlotinib: the LAP07 randomized clinical trial. JAMA 2016 May 3;315(17):1844-1853. [doi: 10.1001/jama.2016.4324] [Medline: 27139057]
5. Boone BA, Steve J, Krasinskas AM, et al. Outcomes with FOLFIRINOX for borderline resectable and locally unresectable pancreatic cancer. J Surg Oncol 2013 Sep;108(4):236-241. [doi: 10.1002/jso.23392] [Medline: 23955427]
6. Badiyan SN, Olsen JR, Lee AY, et al. Induction chemotherapy followed by concurrent full-dose gemcitabine and intensity-modulated radiation therapy for borderline resectable and locally advanced pancreatic adenocarcinoma. Am J Clin Oncol 2016 Feb;39(1):1-7. [doi: 10.1097/COC.0000000000000003] [Medline: 26132367]
7. Kim EJ, Ben-Josef E, Herman JM, et al. A multi-institutional phase 2 study of neoadjuvant gemcitabine and oxaliplatin with radiation therapy in patients with pancreatic cancer. Cancer 2013 Aug 1;119(15):2692-2700. [doi: 10.1002/cncr.28117] [Medline: 23720019]
8. Hoyer M, Roed H, Sengelov L, et al. Phase-II study on stereotactic radiotherapy of locally advanced pancreatic carcinoma. Radiother Oncol 2005 Jul;76(1):48-53. [doi: 10.1016/j.radonc.2004.12.022] [Medline: 15990186]

XSL•FO
RenderX

9. Krishnan S, Chadha AS, Suh Y, et al. Focal radiation therapy dose escalation improves overall survival in locally advanced pancreatic cancer patients receiving induction chemotherapy and consolidative chemoradiation. Int J Radiat Oncol Biol Phys 2016 Mar 15;94(4):755-765. [doi: 10.1016/j.ijrobp.2015.12.003] [Medline: 26972648]
10. Reyngold M, O'Reilly EM, Varghese AM, et al. Association of ablative radiation therapy with survival among patients with inoperable pancreatic cancer. JAMA Oncol 2021 May 1;7(5):735-738. [doi: 10.1001/jamaoncol.2021.0057] [Medline: 33704353]
11. Brunner TB, Nestle U, Grosu AL, Partridge M. SBRT in pancreatic cancer: what is the therapeutic window? Radiother Oncol 2015 Jan;114(1):109-116. [doi: 10.1016/j.radonc.2014.10.015] [Medline: 25466369]
12. Hoffe S, Frakes JM, Aguilera TA, et al. Randomized, double-blinded, placebo-controlled multicenter adaptive phase 1-2 trial of GC 4419, a dismutase mimetic, in combination with high dose stereotactic body radiation therapy (SBRT) in locally advanced pancreatic cancer (PC). Int J Radiat Oncol Biol Phys 2020 Dec 1;108(5):1399-1400. [doi: 10.1016/j.ijrobp.2020.09.022] [Medline: 33427657]
13. Crane CH. Hypofractionated ablative radiotherapy for locally advanced pancreatic cancer. J Radiat Res 2016 Aug;57(Suppl 1):i53-i57. [doi: 10.1093/jrr/rrw016] [Medline: 27029741]
14. Jolissaint JS, Reyngold M, Bassmann J, et al. Local control and survival after induction chemotherapy and ablative radiation versus resection for pancreatic ductal adenocarcinoma with vascular involvement. Ann Surg 2021 Dec 1;274(6):894-901. [doi: 10.1097/SLA.0000000000005080] [Medline: 34269717]
15. Green OL, Henke LE, Hugo GD. Practical clinical workflows for online and offline adaptive radiation therapy. Semin Radiat Oncol 2019 Jul;29(3):219-227. [doi: 10.1016/j.semradonc.2019.02.004] [Medline: 31027639]
16. Henke L, Kashani R, Robinson C, et al. Phase I trial of stereotactic MR-guided online adaptive radiation therapy (SMART) for the treatment of oligometastatic or unresectable primary malignancies of the abdomen. Radiother Oncol 2018 Mar;126(3):519-526. [doi: 10.1016/j.radonc.2017.11.032] [Medline: 29277446]
17. D'Souza A, Kang KH, Lattin JE, et al. Feasibility of stereotactic body radiation therapy for pancreatic tumors abutting organs at risk using magnetic resonance guided adaptive radiation therapy. Int J Radiat Oncol Biol Phys 2025 Jul 1;122(3):668-677. [doi: 10.1016/j.ijrobp.2025.03.049] [Medline: 40185211]
18. Rudra S, Jiang N, Rosenberg SA, et al. Using adaptive magnetic resonance image-guided radiation therapy for treatment of inoperable pancreatic cancer. Cancer Med 2019 May;8(5):2123-2132. [doi: 10.1002/cam4.2100] [Medline: 30932367]
19. Kim H, Olsen JR, Green OL, et al. MR-guided radiation therapy with concurrent gemcitabine/nab-paclitaxel chemotherapy in inoperable pancreatic cancer: a TITE-CRM phase I trial. Int J Radiat Oncol Biol Phys 2023 Jan 1;115(1):214-223. [doi: 10.1016/j.ijrobp.2022.07.015] [Medline: 35878713]
20. Parikh PJ, Lee P, Low DA, et al. A multi-institutional phase 2 trial of ablative 5-fraction stereotactic magnetic resonance-guided on-table adaptive radiation therapy for borderline resectable and locally advanced pancreatic cancer. Int J Radiat Oncol Biol Phys 2023 Nov 15;117(4):799-808. [doi: 10.1016/j.ijrobp.2023.05.023] [Medline: 37210048]
21. Chuong MD, Bryant J, Mittauer KE, et al. Ablative 5-fraction stereotactic magnetic resonance–guided radiation therapy with on-table adaptive replanning and elective nodal irradiation for inoperable pancreas cancer. Pract Radiat Oncol 2021;11(2):134-147. [doi: 10.1016/j.prro.2020.09.005] [Medline: 32947042]
22. Chuong MD, Lee P, Low DA, et al. Stereotactic MR-guided on-table adaptive radiation therapy (SMART) for borderline resectable and locally advanced pancreatic cancer: a multi-center, open-label phase 2 study. Radiother Oncol 2024 Feb;191:110064. [doi: 10.1016/j.radonc.2023.110064] [Medline: 38135187]
23. Sanford NN, Narang AK, Aguilera TA, et al. NRG oncology international consensus contouring atlas on target volumes and dosing strategies for dose-escalated pancreatic cancer radiation therapy. Int J Radiat Oncol Biol Phys 2025 Mar 15;121(4):918-929. [doi: 10.1016/j.ijrobp.2024.10.026] [Medline: 39510320]
24. Herman JM, Chang DT, Goodman KA, et al. Phase 2 multi-institutional trial evaluating gemcitabine and stereotactic body radiotherapy for patients with locally advanced unresectable pancreatic adenocarcinoma. Cancer 2015 Apr 1;121(7):1128-1137. [doi: 10.1002/cncr.29161] [Medline: 25538019]
25. Palta M, Godfrey D, Goodman KA, et al. Radiation therapy for pancreatic cancer: executive summary of an ASTRO clinical practice guideline. Pract Radiat Oncol 2019;9(5):322-332. [doi: 10.1016/j.prro.2019.06.016] [Medline: 31474330]
26. Dona Lemus OM, Cao M, Cai B, Cummings M, Zheng D. Adaptive radiotherapy: next-generation radiotherapy. Cancers (Basel) 2024 Mar 19;16(6):1206. [doi: 10.3390/cancers16061206] [Medline: 38539540]
27. Parikh NR, Lee PP, Raman SS, et al. Time-driven activity-based costing comparison of CT-guided versus MR-guided SBRT. JCO Oncol Pract 2020 Nov;16(11):e1378-e1385. [doi: 10.1200/JOP.19.00605] [Medline: 32539652]
28. Henke LE, Fischer-Valuck BW, Rudra S, et al. Prospective imaging comparison of anatomic delineation with rapid kV cone beam CT on a novel ring gantry radiotherapy device. Radiother Oncol 2023 Jan;178:109428. [doi: 10.1016/j.radonc.2022.11.017] [Medline: 36455686]
29. Schiff JP, Price AT, Stowe HB, et al. Simulated computed tomography-guided stereotactic adaptive radiotherapy (CT-STAR) for the treatment of locally advanced pancreatic cancer. Radiother Oncol 2022 Oct;175:144-151. [doi: 10.1016/j.radonc.2022.08.026] [Medline: 36063981]

XSL•FO
RenderX

30. Kim M, Schiff JP, Price A, et al. The first reported case of a patient with pancreatic cancer treated with cone beam computed tomography–guided stereotactic adaptive radiotherapy (CT-STAR). Radiat Oncol 2022 Sep 13;17(1):157. [doi: 10.1186/s13014-022-02125-z] [Medline: 36100866]
31. McComas KN, Yock A, Darrow K, Shinohara ET. Online adaptive radiation therapy and opportunity cost. Adv Radiat Oncol 2023;8(3):101034. [doi: 10.1016/j.adro.2022.101034] [Medline: 37273924]
32. Hassanzadeh C, Rudra S, Bommireddy A, et al. Ablative five-fraction stereotactic body radiation therapy for inoperable pancreatic cancer using online MR-guided adaptation. Adv Radiat Oncol 2021;6(1):100506. [doi: 10.1016/j.adro.2020.06.010] [Medline: 33665480]

Abbreviations

ART: adaptive radiotherapy
ARTIA: Adaptive Radiation Therapy Individualized Approach
BR: borderline resectable
CBCT: cone beam computed tomography
CT: computed tomography
CT-STAR: computed tomography–guided online stereotactic adaptive radiotherapy
CTCAE: Common Terminology Criteria for Adverse Events
ENI: elective nodal irradiation
GI: gastrointestinal
IRB: institutional review board
ISO: International Organization for Standardization
LAPC: locally advanced pancreatic cancer
MR: magnetic resonance
OAR: organ-at-risk
PTV: planning target volume
SBRT: stereotactic body radiotherapy

Edited by J Sarvestan; submitted 22.Sep.2025; peer-reviewed by J Schiff, MD Chuong; revised version received 19.Nov.2025; accepted 25.Nov.2025; published 23.Feb.2026.

Please cite as:
Henke LE, Kim H, Laugeman E, Kohlmyer S, McCann C, Pietrovito K, Russell K, Woo J, Price AT
Stereotactic Adaptive Radiation Therapy for Borderline Resectable or Locally Advanced Pancreatic Cancer to Minimize Gastrointestinal Toxicity (ARTIA-Pancreas): Protocol for a Single-Arm Prospective Trial
JMIR Res Protoc 2026;15:e84607
URL: https://www.researchprotocols.org/2026/1/e84607
doi:10.2196/84607

© Lauren E Henke, Hyun Kim, Eric Laugeman, Steven Kohlmyer, Claire McCann, Kate Pietrovito, Kenneth Russell, Jennifer Woo, Alex T Price. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Implementation of a Personalized Medicine Approach in Patients With Type 2 Diabetes Mellitus Receiving Multiple Daily Insulin Injections (POMA Project): Protocol for a Before-and-After Intervention Study

Rosa Giné-Balcells[1], RN, MSc; Bogdan Vlacho[2,3], PharmD, MSc, PhD; Olga Carmen Vidal-Pérez[4], RN; Albert Lecube[2,5,6], MD, PhD; Josep Franch-Nadal[2,3,7], MD, PhD; Dídac Mauricio[2,3,8,9,10], MD, PhD; Àngels Molló-Iniesta[1,3*], MD, PhD; Marta Hernández[4,11*], MD, PhD; POMA Study Group[4,12]

[1]Primary Health Care Center Tàrrega, Gerència d'Atenció Primària Lleida, Institut Català de la Salut, Lleida, Spain

[10]Faculty of Medicine, University of Vic - Central University of Catalonia (UVIC/UCC), Vic, Spain

[11]Institut de Recerca Biomèdica de Lleida Fundació Dr. Pifarré, University of Lleida, Lleida, Spain

[12]Primary Health Care Center, Gerència d'Atenció Primària Lleida, Institut Català de la Salut, Lleida, Spain

[2]CIBER de Diabetes y Enfermedades Metabólicas Asociadas, Instituto de Salud Carlos III, Barcelona, Spain

[3]DAP-Cat group, Unitat de Suport a la Recerca Barcelona, Fundació Institut Universitari d'Investigació en Atenció Primària Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain

[4]Department of Endocrinology and Nutrition, University Hospital Arnau de Vilanova, Avda. Alcalde Rovira Roure, 80, Lleida, Spain

[5]Endocrinology Department, Vall d'Hebron University Hospital, Autonomous University Barcelona, Barcelona, Spain

[6]Diabetes and Metabolism Research Group, Vall d'Hebron Research Institute (VHIR), Barcelona, Spain

[7]Primary Health Care Center, Gerència d'Àmbit d'Atenció Primària Barcelona Ciutat, Institut Català de la Salut, Barcelona, Spain

[8]Research Group in Endocrinology, Diabetes and Nutrition, Institut de Recerca Sant Pau, Barcelona, Spain

[9]Department of Endocrinology and Nutrition, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[*]these authors contributed equally

Corresponding Author:
Marta Hernández, MD, PhD
Department of Endocrinology and Nutrition, University Hospital Arnau de Vilanova, Avda. Alcalde Rovira Roure, 80, Lleida, Spain

Related Article:

This is a corrected version. See correction statement: https://www.researchprotocols.org/2026/1/e95267

Abstract

Background: The management of type 2 diabetes mellitus (T2DM) remains a complex clinical challenge, particularly for patients requiring multiple daily insulin injections (MDI). Advances in precision medicine and continuous glucose monitoring (CGM) have created opportunities to personalize treatment and potentially reduce the therapeutic burden on people with T2DM. Assessing β cell function and autoimmunity could help identify patients with T2DM eligible for simplified regimens without compromising glycemic control.

Objective: The aim of this study is to test a simple personalized medicine protocol in routine clinical practice for people with T2DM treated with MDI. The intervention is based on evaluating C-peptide and glutamic acid decarboxylase autoantibody status with the goal of improving diagnostic accuracy and optimizing treatment.

Methods: This is a pragmatic before-and-after intervention study involving people with T2DM currently receiving MDI across primary care centers and a referral hospital in the Lleida health care region in Catalonia (Spain). Eligible participants will undergo clinical and laboratory assessment, including C-peptide and glutamic acid decarboxylase autoantibody testing, and wear a CGM device. On the basis of a predefined algorithm, patients may either continue or discontinue prandial insulin. The primary outcome is the proportion of patients in whom prandial insulin is discontinued and remains discontinued over 6 months. Secondary outcomes include changes in hemoglobin A_{1c}, CGM metric variables, quality of life, adherence, and treatment satisfaction.

Results: Recruitment was completed on March 31, 2025. The follow-up phase is ongoing and expected to conclude by September 30, 2025. Data analysis will begin thereafter.

XSL•FO
RenderX

Conclusions: This study will evaluate the feasibility and impact of implementing a personalized therapeutic approach for persons with T2DM receiving MDI in real-world clinical settings. If effective, this strategy could contribute to safer, simpler, and more individualized diabetes care.

Trial Registration: ClinicalTrials.gov NCT06148376; https://clinicaltrials.gov/ct2/show/NCT06148376

International Registered Report Identifier (IRRID): DERR1-10.2196/85375

(JMIR Res Protoc 2026;15:e85375) doi:10.2196/85375

KEYWORDS

type 2 diabetes mellitus; multiple insulin doses; continuous glucose monitoring; precision medicine; anti-GAD65 autoantibody; C-peptide

Introduction

Type 2 Diabetes Mellitus and Continuous Glucose Monitoring

Diabetes mellitus (DM) is considered a global pandemic with increasing incidence and prevalence and is expected to affect 853 million people worldwide by 2050 [1]. Type 2 DM (T2DM) is a chronic and progressive disease characterized by peripheral insulin resistance and a decline in β cell function [2,3]. The degree of insulin resistance and secretion, the rate of deterioration, the associated metabolic phenotype, the response to treatment, and the occurrence of chronic complications vary across individuals [4].

Although T2DM accounts for over 90% of diabetes cases worldwide [1], a significant proportion of patients diagnosed with T2DM may actually have other forms of diabetes. Among these, latent autoimmune diabetes in adults (LADA) is noteworthy. LADA, a slowly progressive form of autoimmune diabetes, is characterized by the presence of pancreatic autoantibodies and a faster progression to insulin dependence than is typically observed in T2DM [5]. It is often misdiagnosed due to reliance on the clinical phenotype alone [6]. The prevalence of LADA varies by population and geography, ranging from 3% to 12% of all persons with DM [7]. Diagnosis requires detection of at least one pancreatic autoantibody, with the most common being glutamic acid decarboxylase autoantibody (GADAb) [8].

Managing T2DM remains a significant clinical challenge as its progressive nature often increases treatment complexity over time. While approximately 10% of patients eventually require complex insulin regimens [9], newer pharmacological agents and individualized treatment strategies may allow for treatment simplification, reduction of hypoglycemia risk, and improvements in quality of life when aligned with personalized glycemic targets [10-13]. C-peptide measurement can be used to assess endogenous insulin production and serve as a proxy for residual β cell function. This information may help determine whether prandial insulin is necessary or could be safely discontinued in favor of alternative therapies [14].

A randomized controlled trial by Rosenstock et al [15] demonstrated that, in people with T2DM, discontinuing prandial insulin and initiating a once-weekly glucagonlike peptide-1 receptor agonist led to improved glycemic control approaching normoglycemia, reduced hypoglycemia, decreased body weight, and improved quality of life. Similarly, the BEYOND trial by Giugliano et al [16] evaluated switching from basal bolus insulin regimens to a glucagonlike peptide-1 receptor agonist or a sodium-glucose cotransporter type 2 inhibitor. Comparable glycemic control was achieved with lower total insulin doses, fewer daily injections, and reduced hypoglycemia, along with improved weight management and patient satisfaction. A comprehensive review by Giugliano et al [17] including these and other studies confirmed that insulin therapy can be simplified in persons with T2DM without compromising glycemic control while reducing treatment complexity, hypoglycemia risk, and polypharmacy.

Continuous glucose monitoring (CGM) has been evaluated in both randomized trials and real-world studies involving people with T2DM treated with multiple daily insulin injections (MDI), basal insulin, or noninsulin regimens [18,19]. In Catalonia (northeast Spain), the public implementation of CGM for people with T2DM on MDI was approved in November 2023. This required collaboration and coordination between different care providers, constituting an opportunity to develop and implement this protocol in routine clinical practice [20].

This project proposes implementing a precision medicine algorithm for persons with T2DM treated with MDI in real-world clinical practice. This algorithm incorporates assessment of pancreatic autoimmunity and residual β cell function (pancreatic reserve) into the decision-making workup.

Hypothesis and Objectives

We hypothesize that approximately 25% of people with T2DM on MDI have sufficient β cell secretory capacity to safely discontinue prandial insulin. In these cases, treatment could be optimized by introducing agents with proven cardio-renal or weight-lowering benefits according to current clinical guidelines.

The main objective is to evaluate a simple personalized medicine protocol in routine clinical practice for people with T2DM on MDI. The protocol, based on clinical and analytical parameters, aims to improve diagnostic accuracy and potentially simplify treatment in this population.

Secondary objectives include describing patient characteristics; conducting detailed phenotyping of pancreatic function and autoimmunity status; and evaluating the impact of the intervention on glycemic control, treatment adherence, and quality of life.

XSL•FO
RenderX

Methods

This paper reports a study protocol developed in accordance with the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines.

Study Design and Settings

We designed a nonrandomized pragmatic intervention study targeting people diagnosed with T2DM who were treated with MDI. This before-and-after intervention was designed to assess the real-world applicability of a precision medicine algorithm. Rather than using a control group, this study uses each participant as their own control, comparing clinical outcomes before and after the intervention. The intervention was adapted to be integrated into routine clinical workflows across both primary and specialized care settings.

This study is being conducted within 21 primary health care centers of the Catalan Health Institute (Institut Català de la Salut) in the Lleida health care region (Catalonia, Spain; Multimedia Appendix 1) together with the Department of Endocrinology and Nutrition, University Hospital Arnau de Vilanova, as the reference tertiary center.

Eligibility Criteria

We included adults aged ≥18 years with a confirmed diagnosis of T2DM. To be enrolled in the study, participants must be on an MDI regimen, defined as receiving at least 2 doses of prandial insulin per day (either regular insulin or a rapid-acting insulin analogue) administered separately or as part of a fixed-dose combination.

We excluded individuals with a known diagnosis of diabetes other than T2DM, pregnant women, and those with severe cognitive impairment or any condition that prevented the use of CGM. We also excluded participants with current drug or alcohol abuse, those receiving systemic glucocorticoids for acute conditions at the time of inclusion, and individuals undergoing temporary insulin therapy expected to last less than 6 months.

Recruitment and Sampling Strategy

The participant selection sampling strategy was a convenient approach based on the availability and willingness to participate of the patients who fulfilled the study criteria [21].

Participating centers were selected based on geographic distribution and availability of primary health care teams. Teams received virtual training sessions. Sessions included information related to the background, objectives, protocol procedures, and data collection methods before participant recruitment. The study recruitment period lasted 6 months.

Study Procedures and Timeline

Potential participants were identified by local study physicians and nurses using the primary care eCAP (Catalan Health Institute) electronic clinical record software. Each participant underwent 4 study visits. Initially, a preselection visit was conducted through electronic health record review without direct patient contact. Eligible individuals were then contacted and provided with detailed study information. Those who agreed to participate were invited to the inclusion visit (visit 0), during which written informed consent was obtained by research staff. This visit involved a clinical interview, a physical examination, laboratory testing (including C-peptide and GADAb measurements), and the placement of a CGM device. All participants used CGM for 10 to 15 days prior to the intervention. Two weeks later, the first follow-up visit (visit 1) was conducted. At this visit, therapeutic decisions were made based on C-peptide and GADAb results, and treatment was adjusted according to the predefined therapeutic intervention algorithm (Figure 1).

XSL•FO
RenderX

Figure 1. Therapeutic intervention algorithm. CGM: continuous glucose monitoring; GADAb: glutamic acid decarboxylase autoantibody; GLP-1: glucagonlike peptide-1; SGLT2: sodium-glucose cotransporter type 2.

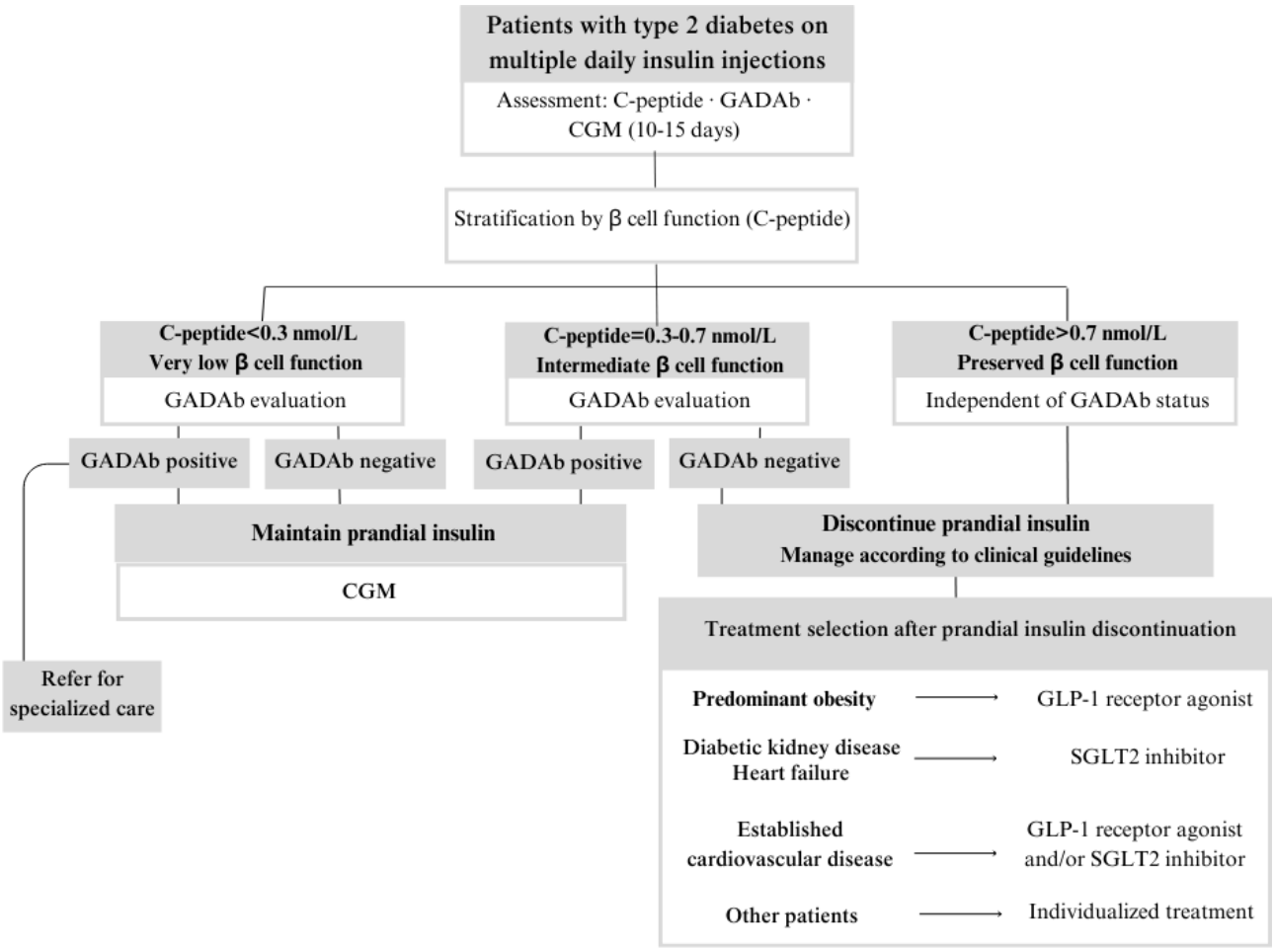

The second follow-up and final in-person study visit (visit 2) took place at month 6, during which outcomes were assessed. Glycemic control was evaluated through hemoglobin A_{1c} (HbA_{1c}) testing, and other clinically relevant follow-up variables were collected. Throughout the study period, participants received additional follow-up and support as needed via telephone or virtual consultations from their primary health care team.

Figure 2 illustrates the planned participant flow and study procedures across the different study visits.

XSL•FO
RenderX

Figure 2. Study flowchart. eCAP: electronic clinical record software of primary care (Catalonia); MDI: multiple daily insulin injection; T2DM: type 2 diabetes mellitus.

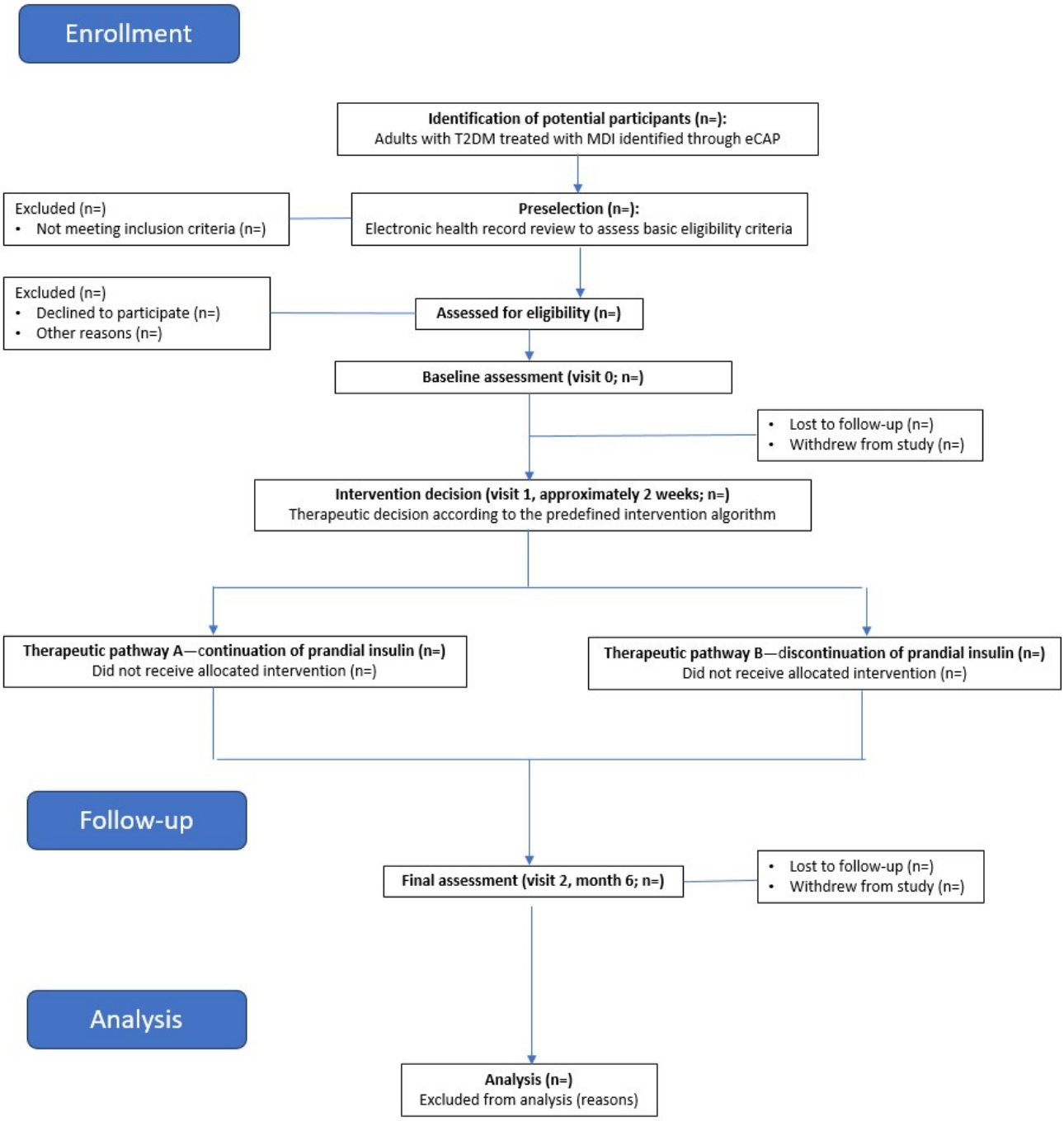

Data Collection and Management

Overview

Data for this study are collected from electronic medical records and patient-reported questionnaires during study visits. Clinical and demographic information, including age, sex, educational level, BMI, diabetes duration, comorbidities, complications, and current pharmacological treatment, is obtained from the eCAP (electronic medical record) software and supplemented by medical interviews. Laboratory variables, including C-peptide, GADAb, HbA_{1c}, and estimated glomerular filtration rate, are extracted from recent laboratory reports or requested when unavailable. CGM data, including metrics such as time in range, hypoglycemia episodes, and glucose variability, are downloaded directly from the devices used by participants throughout the study. Patient-reported outcomes, including quality of life, treatment satisfaction, medication adherence, and physical activity, are assessed using validated questionnaires. All study variables and corresponding time points are summarized in Table 1. In addition, detailed information on the study variables, including their definitions and data sources, is provided in Multimedia Appendix 2.

XSL•FO
RenderX

Table . Study timeline, variables, and procedures.

Data collected	Evaluation time point			
	Preselection	Visit 0—inclusion	Visit 1—intervention (2 weeks)	Visit 2—final follow-up (6 months)
Inclusion and exclusion criteria	✓	✓		
Sociodemographic data[a]		✓		
Anthropometric data[b]		✓		✓
Clinical data[c]		✓		✓
CGM[d] data			✓	✓
Laboratory data[e]		✓		✓
Smoking status		✓		
Pharmacological treatment[f]		✓		✓
Severe hypoglycemia[g]		✓		✓
Questionnaires[h]		✓		✓

[a]Age, gender, primary care center, current occupational status, level of education, number of children, date of menopause, and whether the participant is the sole caregiver of another person.

[b]Weight, height, BMI, and abdominal circumference.

[c]Arterial hypertension, dyslipidemia, micro- and macrovascular complications, heart failure, and renal disease of any etiology (estimated glomerular filtration rate <60 mL per minute per 1.73 m^2).

[d]CGM: continuous glucose monitoring.

[e]Inclusion visit: hemoglobin A_{1c} (HbA_{1c}), glutamic acid decarboxylase autoantibody, glycemia, and C-peptide; final visit: HbA_{1c}.

[f]Prescription and dosage of diabetes drugs.

[g]Inclusion visit: last 12 months; final visit: 6 months of study.

[h]International Physical Activity Questionnaire, Diabetes Treatment Satisfaction Questionnaire–Status Version, Diabetes Treatment Satisfaction Questionnaire–Change Version (final visit only), Adherence to Refills and Medications Scale–Spanish Version, and Spanish version of the Diabetes Quality of Life Questionnaire.

All data are systematically recorded using an electronic case report form (eCRF) specifically developed for this study. Local investigators at each participating center are responsible for ensuring accurate, complete, and timely entry of study data into the eCRF. We perform quality control procedures to verify informed consent and participant eligibility prior to data entry. In addition, the coordinating research team regularly monitors the eCRF data entries to ensure high data quality and compliance with good clinical practice standards throughout the study.

Questionnaires

Several validated questionnaires are used to assess patient-reported outcomes related to quality of life, treatment satisfaction, medication adherence, and physical activity. These questionnaires are administered at both the baseline and final study visits (6 months), as outlined in the study timeline (Table 1).

A detailed description of each questionnaire, including constructs assessed, number of items, scoring, and references, is provided in Table 2.

XSL•FO
RenderX

Table . Summary of patient-reported outcome questionnaires used in this study.

Questionnaire	Construct measured	Description and domains	Number of items	Response scale	Scoring and interpretation	Time points	Reference
ESDQoL[a]	Diabetes-related quality of life	Assesses multiple domains of diabetes-specific quality of life, with domain impact weighted by patient-reported importance	43	5-point Likert scale	Weighted average score; higher scores indicate poorer quality of life	Baseline and 6 months	[22]
DTSQs[b]	Diabetes treatment satisfaction	Measures satisfaction with current diabetes treatment; includes separate items on perceived frequency of hyperglycemia and hypoglycemia	8	7-point Likert scale (0 to 6)	Sum of 6 items (range 0 to 36); higher scores indicate greater satisfaction; 2 items analyzed separately	Baseline and 6 months	[23]
DTSQc[c]	Change in treatment satisfaction	Assesses perceived change in satisfaction compared with previous treatment	8	7-point Likert scale (–3 to +3)	Positive scores indicate improvement; negative scores indicate deterioration	6 months	[23]
ARMS-e[d]	Medication adherence	Evaluates adherence behaviors and barriers related to medication use	12	4-point Likert scale (“never” to “always”)	Lower total scores indicate better adherence	Baseline and 6 months	[24]
IPAQ[e]	Physical activity	Assesses physical activity in the previous 7 days, including walking, moderate and vigorous activity, and sedentary behavior	7	Frequency (d per wk) and duration (min per d)	Physical activity expressed as MET[f] min per week; higher values indicate higher activity levels	Baseline and 6 months	[25]

[a]ESDQoL: Spanish Diabetes Quality of Life Questionnaire.

[b]DTSQs: Diabetes Treatment Satisfaction Questionnaire–Status Version.

[c]DTSQc: Diabetes Treatment Satisfaction Questionnaire–Change Version.

[d]ARMS-e: Adherence to Refills and Medications Scale–Spanish Version.

[e]IPAQ: International Physical Activity Questionnaire–Short Form (Spanish version).

[f]MET: metabolic equivalent of task.

Intervention Algorithm

Therapeutic decisions in the study are guided by an algorithm based on C-peptide and GADAb levels (Figure 1). Patients with a C-peptide level below 0.3 nmol/L and positive GADAb are considered suspected type 1 diabetes cases; they continue MDI and CGM and are referred for specialized care. Those with similarly low C-peptide levels but negative GADAb also continue MDI and CGM with no changes in treatment.

For patients with intermediate C-peptide values (0.3 - 0.7 nmol/L), the approach varies based on GADAb status. If GADAb is positive, treatment remains unchanged with continued MDI and CGM. However, if GADAb is negative, clinicians consider discontinuing prandial insulin and introducing noninsulin hypoglycemic agents as appropriate.

Participants with C-peptide levels above 0.7 nmol/L regardless of GADAb status are also considered for discontinuation of prandial insulin and initiation of noninsulin therapies. In individuals who discontinue prandial insulin, CGM use is typically discontinued as well as they do not meet the criteria for public funding [26].

Study Outcomes

We defined the primary study outcome as the proportion of participants in whom prandial insulin is successfully discontinued after the application of the therapeutic algorithm

XSL•FO
RenderX

and who remain without prandial insulin at the final visit (6 months).

Secondary outcomes encompass several domains. Glycemic control following therapeutic optimization is evaluated through changes in HbA_{1c} and CGM-derived metrics, including time in range and time below range. Patient-reported outcomes are also assessed by comparing changes from baseline to 6 months after the intervention, including changes in quality of life (using the Diabetes Quality of Life Questionnaire), treatment satisfaction (using the Diabetes Treatment Satisfaction Questionnaire, both status and change versions), medication adherence (measured using the Adherence to Refills and Medications Scale), and physical activity levels (evaluated using the Spanish version of the International Physical Activity Questionnaire).

Safety outcomes include the number of severe hypoglycemic episodes at the final visit 6 months after the intervention compared with the 12-month period before enrollment. The study also examines the predictive value of CGM metrics by analyzing their association with pancreatic reserve and insulin resistance considering both continuous and categorical variables.

Finally, among patients who continue prandial insulin and CGM use for 6 months, the study assesses the impact of CGM on glycemic control (HbA_{1c} and CGM metrics), as well as changes in quality of life, treatment satisfaction, and medication adherence.

Sample Size and Statistical Analysis

At the time of the study planning, data from the Lleida health care region had identified 1388 people with T2DM treated with MDI. On the basis of prior clinical experience, approximately 25% of these individuals may have sufficient residual β cell function to discontinue prandial insulin. To estimate this proportion with a 95% CI and a margin of error of –5% to +5%, a sample size of 239 participants was estimated.

The study population will be characterized at baseline with respect to age, BMI, duration of diabetes, presence of chronic diabetes-related complications (microvascular and macrovascular), and cardiovascular risk factors. In addition, in-depth phenotyping will be performed using the homeostatic model assessment of β cell function and homeostatic model assessment of insulin resistance indexes [27] and GADAb measurements.

For the primary outcome, the proportion of participants who discontinue prandial insulin after the application of the diagnostic algorithm and sustain this change over a 6-month follow-up period will be calculated, along with its corresponding 95% CI.

Secondary outcomes will include measures of glycemic control, patient-reported outcomes, and safety outcomes. Glycemic control will be assessed using HbA_{1c} and CGM-derived metrics, including time in range (70 - 180 mg/dL), time below range, time above range, mean glucose, and glycemic variability indexes. Patient-reported outcomes will include validated scores for treatment satisfaction, adherence, and quality of life analyzed according to each instrument's scoring manual. Safety outcomes will include the incidence of severe hypoglycemia.

To evaluate changes over time (eg, from baseline to follow-up visits), paired 2-tailed *t* tests or Wilcoxon signed-rank tests will be used depending on data distribution. Between-group comparisons (eg, discontinuation vs continuation of prandial insulin) will be conducted using Student *t* tests or Mann-Whitney *U* tests as appropriate, as well as using linear mixed-effects models to account for within-subject correlation.

For categorical secondary outcomes, including severe hypoglycemia, comparisons between groups will be performed using chi-square or Fisher exact tests as appropriate. For outcomes measured repeatedly across follow-up visits, analyses will be conducted using linear mixed-effects models.

CGM variables will also be examined as potential predictors of pancreatic reserve and insulin resistance. Depending on whether outcomes are continuous or categorical, linear or logistic regression models will be applied. Results will be expressed as regression coefficients or odds ratios with corresponding 95% CIs.

Loss to follow-up will be quantified and reported as the proportion of participants without outcome data at each follow-up time point. When available, reasons for loss to follow-up will be documented. Baseline characteristics of participants with complete follow-up and those lost to follow-up will be compared to assess potential differential attrition. A complete-case analysis will be used as the primary analytical approach when the proportion of missing data is low, and missingness will be considered compatible with a missing completely at random mechanism. To account for potentially informative missingness, multiple imputation by chained equations will be performed under a missing-at-random assumption, including all variables used in the main analyses and auxiliary variables related to missingness.

All statistical analyses will be performed using the R software (version 4.1.2; R Foundation for Statistical Computing), with a 2-sided *P* value of <.05 considered statistically significant.

Ethical Considerations

Overview

This study will be conducted in accordance with the Declaration of Helsinki and the principles of good clinical practice. The study protocol received approval from the ethics committee of Institut Universitari d'Investigació en Atenció Primària Jordi Gol i Gurina, the research institute for primary care in Catalonia. The approved protocol is registered under reference 23/193-EOm. Participants did not receive any financial or material compensation for taking part in the study.

The trial is registered at ClinicalTrials.gov with the identifier NCT06148376.

Informed Consent

All participants received verbal and written information about the study and provided signed informed consent before enrollment. When participants had limited capacity to consent, their legal representatives may provide authorization.

XSL•FO
RenderX

Data Protection and Confidentiality

All personal data will be pseudonymized and processed in compliance with the General Data Protection Regulation (2016/679) and Spanish Organic Law 3/2018 on the Protection of Personal Data and Guarantee of Digital Rights. Individual patient information will be encrypted to ensure anonymity. Only authorized researchers and study monitors will have access to the data.

Publication Commitment

The results of the study will be disseminated regardless of the outcome and will be submitted for peer-reviewed publication. Participant confidentiality will be maintained in all presentations and publications.

Results

Recruitment for the study began in January 2024 and was completed on March 31, 2025.

As of August 31, 2025, the study is in its final follow-up phase, scheduled to conclude on September 30, 2025. Formal data analysis will begin after completion of follow-up and final data collection.

Discussion

Expected Findings

According to data from the Catalan Agency for Health Quality and Assessment and the eCAP electronic medical record system, the Lleida health care region has approximately 27,769 individuals diagnosed with T2DM, representing 8.95% of the adult population. Among these patients, 5% are currently treated with MDI.

This study aims to improve the available evidence in precision medicine for T2DM, specifically targeting patients treated with MDI. Using the simple precision medicine algorithm (Figure 1) proposed in this study based on assessment of β cell function (C-peptide) and autoimmunity (GADAb), we will evaluate whether prandial insulin can be safely removed or replaced by another treatment in real-world clinical practice.

The study's main outcome is the proportion of patients in whom prandial insulin is successfully discontinued and remains unnecessary at 6 months. Secondary outcomes include clinical measures such as HbA_{1c} levels; CGM-derived metrics related to insulin resistance and pancreatic function; incidence of severe hypoglycemia; and several patient-reported outcomes—including quality of life, treatment satisfaction, medication adherence, and physical activity. Among patients who continue prandial insulin and CGM, the study also examines the added value of CGM use on both metabolic control and patient experience.

The algorithm is implemented by a multidisciplinary research team—including general practitioners, endocrinologists, and nurses—who integrate biomarker results with patient-specific clinical characteristics to guide medical decision-making. The strategy proposed by this study is intended for inclusion in clinical practice. If successful, this approach could be adopted by health care professionals managing T2DM to achieve simpler, safer, and more effective treatment strategies, particularly those that offer cardiovascular, renal, and weight management benefits.

This personalized strategy aligns with a broader innovation trend in diabetes care. Similar to the study by Dennis et al [28], which developed and validated a predictive model to individualize glucose-lowering therapies in T2DM, our project aims to move beyond conventional treatment algorithms by incorporating patient-specific pathophysiological characteristics.

Adopting a simple, precision-based approach to T2DM represents both a challenge and an opportunity. It can foster improved coordination between primary and specialized diabetes care teams and enhance understanding of diabetes heterogeneity. This study exemplifies how precision strategies can be tested and implemented in standard care settings. By enabling more individualized treatment decisions, the intervention may reduce overtreatment, minimize hypoglycemia, enhance quality of life, empower patients, and strengthen the use of digital health tools and telemedicine in routine care.

Limitations

This study has several limitations. First, as with many observational studies involving voluntary participation, there is potential selection bias. Individuals who choose to participate may differ systematically from those who do not, for example, by being more motivated, having greater health literacy, or exhibiting more stable diabetes management. Such characteristics could affect adherence, clinical outcomes, and openness to treatment simplification, thereby potentially limiting the generalizability of the findings [21].

Second, although the study includes all persons with a diagnosis of T2DM treated with MDI within the participating centers, recruitment depended on the initiative of local primary care teams, which may introduce variability in enrollment rates and patient selection. Third, the intervention algorithm offers flexibility in its clinical application, potentially leading to deviations from the proposed protocol. However, given the pragmatic nature of the study, this flexibility is considered necessary to ensure patient safety and adapt the intervention to the heterogeneous clinical profiles of T2DM, a condition influenced by multiple factors affecting treatment response. Lastly, the duration of diabetes may affect the reliability of autoantibody testing. In long-standing cases, GADAb titers may decline over time, leading to false-negative results. Although longitudinal data are limited, up to 40% of GADAb-positive patients may become seronegative 10 years after diagnosis [29]. Despite this, GADAb was selected over other autoantibodies due to its higher prevalence and persistence in patients with LADA [5].

In summary, these limitations—including the use of convenience-based recruitment, potential selection bias, variability in recruitment across centers, and the pragmatic flexibility of the intervention algorithm—should be considered when interpreting the findings. These factors may limit causal inference and applicability, but they are consistent with the study's feasibility-focused, real-world nature.

XSL•FO
RenderX

Conclusions

More accurate phenotyping of persons with T2DM on MDI may improve diagnostic accuracy and enable treatment optimization by identifying candidates for prandial insulin withdrawal or for the introduction of agents with better safety profiles and added cardiovascular benefits [30].

Simplifying therapeutic strategies in T2DM management can be effective provided that personalized care is prioritized, patient safety is preserved, and glycemic control is maintained.

Furthermore, implementing CGM represents both a challenge and an opportunity to enhance diabetes care across primary care and endocrinology services. Strengthening collaboration across care levels and promoting the use of digital health tools may empower patients and improve the management of this highly prevalent and impactful chronic disease.

Acknowledgments

The POMA Study Group members are as follows: Magda Mateu Amorós, Anaïs Arqué Badia, Núria Balsells Bailón, Marta Bueno Díez, Marta Carrasco Marín-Blázquez, Alba Dalmau Vila, Carolina López Cano, Raquel Ruano Esteban, Ana Gloria Soler Beunza, Javier Suárez Balaguer, Gabriela Monroy Rodríguez, Josep León Mengíbar, Ferran Rius Riu, Maribel Martí Sotus, Mª José Rosauro Gómez, Laia Naté López, Josefa Tomàs Anastasi, Rocío Quintillà Pueo, Sara Areny Rodrigo, Fernando Chilangua Sepúlveda, Mercè Oliver García, Ruben Perales Sánchez, Anna Serra Peruchet, Jaume Martí Montesa, Irene Gómez Companys, Janette Sandrà Medina, Silvia Guiu Vendrell, Roxana Lenuta Mocanu, David José De la Rica Escuin, Mª Teresa Guasch Clapés, Eva Ribalta Calvet, Rosa M Aran Padulles, Encarna López Esteve, Meritxell Cunillera Batlle, Laura Camats Pascual, Esther Call Camps, Laura García Esteve, Teresa Ibañez Plana, Laia Llubes Arria, Irene Muñoz del Campo, Agustí Jové Subirada, Silvia Tierz Puyuelo, Teresa Vidal Bertran, Maria Mesalles Rodríguez, Natalia Reñé Puigfel, Asunción Garrido Díaz, Sandra Maria Millera, Mª Antònia Lafarga Giribets, Marta Ferminán Miró, Mònica Solanes Cabús, Mª Abel Cucurull Llobet, Laura Roure Chiquero, Salvador Salanova Montané, Mª Fe Diaz Fernandez, Arantxa Morales Valle, Cristina Cebrián Aiguadé, Anna Pagès Cónsul, Neus Miró Vallvé, Raquel Rodríguez Aguilar, Maite Puig Solé, Montserrat Torra Solé, Úrsula Belén Flix Pau, Sara Porta Acosta, and Marta Tolosa Fortuny.

The authors would like to express their sincere thanks to all collaborating researchers, general practitioners, and nurses from the participating primary care centers, as well as to the endocrinologists and nurses at University Hospital Arnau de Vilanova in Lleida, Spain. All of them are part of the POMA (Implementation of a Personalized Medicine Approach in Patients With Type 2 Diabetes Mellitus on Multiple Daily Insulin Injections) study group.

The authors also acknowledge that the Diabetes Treatment Satisfaction Questionnaire (DTSQ) is a copyrighted instrument developed by Professor Clare Bradley. Permission to use the DTSQ was obtained from Health Psychology Research Ltd (HPR), UK.

Funding

This project received financial support from the Fundación Red de Grupos de Estudio de la Diabetes en Atención Primaria de la Salud through the seventh call for the José Luis Torres Research Support Grant, totaling €8000 (US $9445.84). In addition, an unrestricted grant from Abbott Laboratories was awarded to support protected time for the first author (RG-B) to coordinate the implementation of continuous glucose monitoring in the health care region, which aligns with the scope of this study. The sponsor had no role in the design, conduct, analysis, or reporting of the study.

Authors' Contributions

RG-B, BV, DM, JF-N, AM-I, and MH contributed to the conception, study design, and methodological planning of the project. They serve as the scientific coordinators and were responsible for developing the intervention model. RG-B contributed to study oversight and operational execution. RG-B, BV, AM-I, and MH drafted the manuscript. All authors reviewed and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of study participating centers.
[DOCX File, 31 KB - resprot_v15i1e85375_app1.docx]

Multimedia Appendix 2
Study variables.

XSL•FO
RenderX

[DOCX File, 31 KB - resprot_v15i1e85375_app2.docx]

Checklist 1
SPIRIT checklist.
[DOCX File, 42 KB - resprot_v15i1e85375_app3.docx]

References

1. IDF diabetes atlas. International Diabetes Federation. 2025. URL: https://diabetesatlas.org/ [accessed 2025-05-25]
2. U.K. prospective diabetes study 16. Overview of 6 years' therapy of type II diabetes: a progressive disease. U.K. prospective diabetes study group. Diabetes 1995 Nov;44(11):1249-1258. [doi: 10.2337/diabetes.44.11.1249] [Medline: 7589820]
3. ElSayed NA, McCoy RG, Aleppo G. Diagnosis and classification of diabetes: standards of care in diabetes-2025. Diabetes Care 2025 Jan 1;48(1 Suppl 1):S27-S49. [doi: 10.2337/dc25-S002] [Medline: 39651986]
4. Redondo MJ, Hagopian WA, Oram R, et al. The clinical consequences of heterogeneity within and between different diabetes types. Diabetologia 2020 Oct;63(10):2040-2048. [doi: 10.1007/s00125-020-05211-7] [Medline: 32894314]
5. Laugesen E, Østergaard JA, Leslie RDG, Danish Diabetes Academy Workshop and Workshop Speakers. Latent autoimmune diabetes of the adult: current knowledge and uncertainty. Diabet Med 2015 Jul;32(7):843-852. [doi: 10.1111/dme.12700] [Medline: 25601320]
6. Hernández M, Mauricio D. Latent autoimmune diabetes in adults: a review of clinically relevant issues. Adv Exp Med Biol 2021;1307:29-41. [doi: 10.1007/5584_2020_533] [Medline: 32424495]
7. Vich-Pérez P, Abánades-Herranz JC, Mora-Navarro G, et al. Development and validation of a clinical score for identifying patients with high risk of latent autoimmune adult diabetes (LADA): the LADA primary care-protocol study. PLoS ONE 2023;18(2):e0281657. [doi: 10.1371/journal.pone.0281657] [Medline: 36758065]
8. Pieralice S, Pozzilli P. Latent autoimmune diabetes in adults: a review on clinical implications and management. Diabetes Metab J 2018 Dec;42(6):451-464. [doi: 10.4093/dmj.2018.0190] [Medline: 30565440]
9. Vinagre I, Mata-Cases M, Hermosilla E, et al. Control of glycemia and cardiovascular risk factors in patients with type 2 diabetes in primary care in Catalonia (Spain). Diabetes Care 2012 Apr;35(4):774-779. [doi: 10.2337/dc11-1679] [Medline: 22344609]
10. Artasensi A, Pedretti A, Vistoli G, Fumagalli L. Type 2 diabetes mellitus: a review of multi-target drugs. Molecules 2020 Apr 23;25(8):1987. [doi: 10.3390/molecules25081987] [Medline: 32340373]
11. Jude EB, Malecki MT, Gomez Huelgas R, et al. Expert panel guidance and narrative review of treatment simplification of complex insulin regimens to improve outcomes in type 2 diabetes. Diabetes Ther 2022 Apr;13(4):619-634. [doi: 10.1007/s13300-022-01222-2] [Medline: 35274219]
12. Mehta R, Billings LK, Liebl A, Vilsbøll T. Transitioning from basal-bolus or premix insulin therapy to a combination of basal insulin and glucagon-like peptide-1 receptor agonist in people with type 2 diabetes. Diabet Med 2022 Sep;39(9):e14901. [doi: 10.1111/dme.14901] [Medline: 35708737]
13. Davies MJ, Aroda VR, Collins BS, et al. Management of hyperglycaemia in type 2 diabetes, 2022. A consensus report by the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetologia 2022 Dec;65(12):1925-1966. [doi: 10.1007/s00125-022-05787-2] [Medline: 36151309]
14. Maddaloni E, Bolli GB, Frier BM, et al. C-peptide determination in the diagnosis of type of diabetes and its management: a clinical perspective. Diabetes Obes Metab 2022 Oct;24(10):1912-1926. [doi: 10.1111/dom.14785] [Medline: 35676794]
15. Rosenstock J, Nino A, Soffer J, et al. Impact of a weekly glucagon-like peptide 1 receptor agonist, albiglutide, on glycemic control and on reducing prandial insulin use in type 2 diabetes inadequately controlled on multiple insulin therapy: a randomized trial. Diabetes Care 2020 Oct;43(10):2509-2518. [doi: 10.2337/dc19-2316] [Medline: 32694215]
16. Giugliano D, Longo M, Caruso P, et al. Feasibility of simplification from a basal-bolus insulin regimen to a fixed-ratio formulation of basal insulin plus a GLP-1RA or to basal insulin plus an SGLT2 inhibitor: BEYOND, a randomized, pragmatic trial. Diabetes Care 2021 Jun;44(6):1353-1360. [doi: 10.2337/dc20-2623] [Medline: 33883195]
17. Giugliano D, Scappaticcio L, Longo M, et al. Simplification of complex insulin therapy: a story of dogma and therapeutic resignation. Diabetes Res Clin Pract 2021 Aug;178:108958. [doi: 10.1016/j.diabres.2021.108958] [Medline: 34280468]
18. American Diabetes Association Professional Practice Committee. 7. Diabetes technology: standards of care in diabetes-2025. Diabetes Care 2025 Jan 1;48(1 Suppl 1):S146-S166. [doi: 10.2337/dc25-S007] [Medline: 39651978]
19. Ajjan RA, Battelino T, Cos X, et al. Continuous glucose monitoring for the routine care of type 2 diabetes mellitus. Nat Rev Endocrinol 2024 Jul;20(7):426-440. [doi: 10.1038/s41574-024-00973-1] [Medline: 38589493]
20. Model organitzatiu per a la prestació dels sistemes de monitoratge continu de glucosa a pacients amb diabetis mellitus de tipus 2 al SISCAT. CatSalut: Servei Català de la Salut. URL: https://scientiasalut.gencat.cat/handle/11351/9917 [accessed 2025-05-25]
21. Convenience sampling method: how and when to use it? Qualtrics. URL: https://www.qualtrics.com/experience-management/research/convenience-sampling/ [accessed 2025-05-25]

XSL•FO
RenderX

22. Millan M. Quality-of-life questionnaire designed for diabetes mellitus (EsDQOL). Aten Primaria 2002 May 15;29(8):517-521. [doi: 10.1016/s0212-6567(02)70623-9] [Medline: 12031227]
23. Bradley C. Diabetes treatment satisfaction questionnaire: (DTSQ). In: Bradley C, editor. Handbook of Psychology and Diabetes: A Guide to Psychological Measurement in Diabetes Research and Practice: Routledge; 1994.
24. Kripalani S, Risser J, Gatti ME, Jacobson TA. Development and evaluation of the Adherence to Refills and Medications Scale (ARMS) among low-literacy patients with chronic disease. Value Health 2009;12(1):118-123. [doi: 10.1111/j.1524-4733.2008.00400.x] [Medline: 19911444]
25. Craig CL, Marshall AL, Sjöström M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
26. Resolución de 7 de abril de 2022, de la dirección general de cartera común de servicios del sistema nacional de salud y farmacia, por la que se hace público el acuerdo de la comisión de prestaciones, aseguramiento y financiación de 2 de marzo de 2022 sobre sistema de monitorización de glucosa para pacientes con diabetes mellitus tipo 2 en la cartera común de servicios del sistema nacional de salud. : Ministerio de Sanidad; 2022 URL: https://www.sanidad.gob.es/ca/profesionales/prestacionesSanitarias/CarteraDeServicios/ContenidoCS/docs/Resoluc_DG__DM2_DEF.pdf [accessed 2026-02-20]
27. Diabetes trials unit: HOMA calculator. University of Oxford. URL: https://www.dtu.ox.ac.uk/homacalculator/ [accessed 2025-05-25]
28. Dennis JM, Young KG, Cardoso P, et al. A five-drug class model using routinely available clinical features to optimise prescribing in type 2 diabetes: a prediction model development and validation study. The Lancet 2025 Mar;405(10480):701-714. [doi: 10.1016/S0140-6736(24)02617-5]
29. Tuomi T, Santoro N, Caprio S, Cai M, Weng J, Groop L. The many faces of diabetes: a disease with increasing heterogeneity. The Lancet 2014 Mar;383(9922):1084-1094. [doi: 10.1016/S0140-6736(13)62219-9]
30. Fritsche A, Heni M, Peter A, et al. Considering insulin secretory capacity as measured by a fasting C-peptide/glucose ratio in selecting glucose-lowering medications. Exp Clin Endocrinol Diabetes 2022 Mar;130(3):200-204. [doi: 10.1055/a-1242-9809] [Medline: 32947641]

Abbreviations

CGM: continuous glucose monitoring
DM: diabetes mellitus
eCRF: electronic case report form
GADAb: glutamic acid decarboxylase autoantibody
HbA_{1c}: hemoglobin A_{1c}
LADA: latent autoimmune diabetes in adults
MDI: multiple daily insulin injection
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
T2DM: type 2 diabetes mellitus

Edited by J Sarvestan; submitted 16.Oct.2025; peer-reviewed by O Ajayi; revised version received 18.Jan.2026; accepted 19.Jan.2026; published 24.Feb.2026.

Please cite as:
Giné-Balcells R, Vlacho B, Vidal-Pérez OC, Lecube A, Franch-Nadal J, Mauricio D, Molló-Iniesta À, Hernández M, POMA Study Group
Implementation of a Personalized Medicine Approach in Patients With Type 2 Diabetes Mellitus Receiving Multiple Daily Insulin Injections (POMA Project): Protocol for a Before-and-After Intervention Study
JMIR Res Protoc 2026;15:e85375
URL: https://www.researchprotocols.org/2026/1/e85375
doi:10.2196/85375

© Rosa Giné-Balcells, Bogdan Vlacho, Olga Carmen Vidal Perez, Albert Lecube, Josep Franch-Nadal, Dídac Mauricio, Àngels Molló-Iniesta, Marta Hernández, POMA Study Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Using Textural Analysis of Thermal Imaging to Predict Healing Status of Diabetes-Related Neuropathic Foot Ulcers: Protocol for a Co-Design and Longitudinal Study

Kate Waller[1], BPod; Rajna Ogrin[2,3], PhD; Quoc Cuong Ngo[4], PhD; Barbara Polus[4], PhD; Samantha Hanna[5], BPod; Aye Nyein Tint[5,6], MBBS, MSc; Richard J MacIsaac[1,3,7,8], MBBS, PhD; Anna Galligan[1,7], MBBS, MMed; Elif Ilhan Ekinci[3,6,8], MBBS, PhD; Dinesh Kumar[3,4], PhD

[1]Department of Podiatry and High Risk Foot Service, St Vincent's Hospital Melbourne, Melbourne, Victoria, Australia

[2]Support Services Office, Bolton Clarke, Forest Hill, Victoria, Australia

[3]Australian Centre for Accelerating Diabetes Innovations (ACADI), Melbourne Medical School, University of Melbourne, Melbourne, Victoria, Australia

[4]School of Engineering, RMIT University, 124 La Trobe St, Melbourne, Victoria, Australia

[5]Department of Podiatry and High Risk Foot Service, Austin Hospital, Heidelberg, Victoria, Australia

[6]Department of Endocrinology, Austin Health, Heidelberg, Victoria, Australia

[7]Department of Endocrinology and Diabetes, St Vincent's Hospital Melbourne, Melbourne, Victoria, Australia

[8]Department of Medicine, Melbourne Medical School, University of Melbourne, Melbourne, Victoria, Australia

Corresponding Author:
Quoc Cuong Ngo, PhD
School of Engineering, RMIT University, 124 La Trobe St, Melbourne, Victoria, Australia

Abstract

Background: Diabetes-related foot ulcers (DFUs) are common in people with diabetes and a major cause of poor quality of life and disability. If not treated in a timely and appropriate way, DFUs may result in prolonged hospitalization and amputation. Currently, methods to predict the healing trajectory of DFUs lack accuracy. Thermal imaging has been proposed to overcome these limitations but has been unable to accurately predict delayed healing of DFUs in the early stages of ulcer management. This project aims to ascertain whether textural analysis of a thermal image can predict the healing trajectory of DFUs.

Objective: The study aims (1) to co-design an accurate, fast, easy-to-use, computer-aided, nontouch test to predict DFU healing trajectory using texture analysis of thermal imaging that is fit for purpose and acceptable to both, those being tested and users of the device and (2) to validate whether textural analysis of thermal images can accurately predict healing of DFUs at week 12 from an image taken at week 1.

Methods: This project will be undertaken in 2 phases: Phase 1, co-design and development of the software prototype; and Phase 2, technology validation. Phase 1 requires a participatory action, co-design approach, engaging clinicians and biomedical engineers in a facilitated focus group and email communication. Interviews with adults living with diabetes and a DFU or who have a history of a DFU will be undertaken to understand their information needs about the device and its findings. Phase 2 will be a longitudinal observational study of 120 adults living with a DFU over a 12-week period. Demographic and other data that have been shown to impact wound healing will be collected at baseline, including participant age, gender, wound size, wound duration, and biomedical markers. Thermal and standard red, green, and blue images will be taken at weeks 1, 2, and 12. Wound textural features will be entered into a Bayesian neural network to identify the healing trajectory.

Results: Phase 1 has been completed with biomedical engineers and clinicians from High Risk Foot Services at 2 hospitals in Melbourne, Australia. Nine clinicians participated in a co-design focus group, and 6 clinicians communicated via email. Four themes were identified: insights for sector use, how to document and use the device, legal implications, and workflow requirements. Interviews with 4 people living with diabetes were undertaken, with data to be analyzed thematically. Phase 2 data collection will be completed by April 2026.

Conclusions: The study aims to co-design, test, and validate an accurate, feasible, and acceptable device to predict the healing of DFUs at week 12 from an image taken at week 1. This has the potential to assist clinicians in making informed and timely decisions for instigating adjuvant therapies, thereby improving healing and preventing lower extremity amputations.

International Registered Report Identifier (IRRID): DERR1-10.2196/69793

(JMIR Res Protoc 2026;15:e69793) doi:10.2196/69793

XSL•FO
RenderX

KEYWORDS

diabetic foot; foot ulcer; wound healing; thermal imaging; textural analysis; lower extremity

Introduction

Background

Diabetes affects an estimated 463 million people worldwide, a number expected to increase to over 700 million by 2045 [1]. An estimated 131 million people (1.8% of the global population) had a diabetes-related lower extremity complication (neuropathy, foot ulcer, and amputation) in 2016 [2], and the prevalence of diabetes-related foot ulcers (DFUs) worldwide is approximately 6.3% [3], with estimated costs between US $3999 to US $26,000, with high variability due to heterogeneity of research methodologies [4].

Each year in Australia, diabetes-related foot disease causes an estimated 28,000 hospital admissions, 4500 lower extremity amputations, and 1700 deaths, and costs the Australian health system AU $1.6 billion (US $1,073,696) annually [5-7]. People with DFUs have high rates of delayed wound healing and increased rates of disability, cardiovascular disease, lower extremity amputation, and mortality [8].

Reviewing real-world data, as many as 55% to 70% of all DFUs exhibit delayed healing [9]. Predictive knowledge regarding whether an ulcer is healing at the expected trajectory is crucial for timely clinical decisions, as early intervention in ulcers that are not healing can prevent lower extremity amputation [10,11]. It has been suggested that implementing adjuvant therapies at an early stage in wounds that exhibit delayed healing may reduce healing times and other complications of wounds, including amputations [12].

Imaging modalities have been proposed for the assessment of wounds and have the advantage of being noninvasive, noncontact, and providing real-time digital assessment [13]. Conventional imaging methods, such as magnetic resonance imaging, computed tomography, and positron emission tomography, are used to detect infection or bone involvement in DFUs [14]. However, these techniques are not well-suited for routine outpatient care or for predicting wound healing trajectories, as they are constrained by factors such as radiation exposure (tomography computed and positron emission tomography), prolonged acquisition times, and limited accessibility [15-18]. Recent advances in noninvasive imaging technologies have broadened the range of approaches available for DFU assessment and monitoring, with growing interest in the use of standard red, green, and blue (RGB) imaging, thermal imaging, and hyperspectral imaging to provide practical, safe, and scalable solutions for evaluating wound status and healing potential.

Standard RGB images are used to analyze the area [19] or the shape of an ulcer [20,21]. While these have the advantage of using cameras that are located in currently available digital devices, such as a smartphone, the color and texture of the skin can affect the outcomes of such a technique [22]. The advancements in artificial intelligence (AI) and deep learning have overcome some of the limitations, but there are large differences between the performance of different imaging modalities in assessing wounds in different studies [23].

Hyperspectral imaging is a modality for the assessment of the spectrum of reflected light and has been proposed for the assessment of DFUs and other wounds [24]. However, the performance is dependent on factors such as lighting conditions and the skin color of the person who has the wound [24]. Further, the cost of hyperspectral imaging may be prohibitive, estimated at AU $50,000 (US $33,553) per unit. Furthermore, the process is slow and is computationally expensive, which can be challenging in a clinical setting [24].

In recent years, thermal imaging has been investigated to predict the healing trajectory in DFUs [13]. Thermal imaging enables the detection of temperature differences, quantifying sensitive changes in skin temperature that are associated with pathological changes such as soft tissue inflammation of the skin, subsequent breakdown, and infection of ulcers [25]. Isothermal maps of thermal images to assess changes in the area of DFUs have been used to successfully predict healing at week 4 [13]. The results, however, were highly sensitive to ambient conditions such as the temperature [13,26]. There is currently no dedicated research focused on identifying delayed healing DFUs solely at the first presentation.

Research into venous leg ulcers has shown that textural features of thermal images can provide a suitable binary classification of wounds to determine healing versus delayed healing wounds [27]. Textural features of thermal images using machine learning and AI-based technology were investigated [27,28]. This approach involves machine learning, training the model from wounds recorded and analyzed with known labels of healing within a 12-week period and was successfully proven in venous leg ulcers, overcoming the environmental limitations [27]. The advantage of this technique is that it is not sensitive to the absolute values of the temperature and thus not affected by ambient conditions or skin color [27]. Further, it has been shown to predict the healing of venous leg ulcers with 1 image, with a sensitivity of 78% [27]. This is relevant for the device to have global usability, regardless of large climatic variations and access to specialist care being limited.

Study Aims and Objectives

The aims of this research, based on previous co-design workshops, are (1) to co-design and develop an inexpensive, fast, AI-based computer-aided, noninvasive test to predict DFU healing trajectory using a thermal imaging device enabling textural analysis that is fit for purpose and acceptable to both those being tested and users of the device; (2) to validate whether textural analysis of thermal images can accurately predict healing of DFUs at week 12 from an image taken at week 1.

XSL•FO
RenderX

Methods

Study Overview

This study will be undertaken in 2 phases: Phase 1 is the co-design and development of the prototype, a software-based medical device; and Phase 2 is technology validation. The study reporting is guided by the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) checklist for cohort studies [29] (Checklist 1).

Setting

This research will be undertaken in Melbourne, Australia, at High Risk Foot Services of 2 tertiary hospitals. These high-risk foot services provide multidisciplinary care for patients with diabetes-related foot complications. The prevention and treatment of diabetes-related foot disease depends upon a well-organized team of integrated clinicians from various disciplines who use a holistic approach to the assessment and treatment of a foot ulcer and its contribution to multiorgan disease [30].

Phase 1: Co-Design and Improvement of Prototype

Study Design

Co-design entails the active engagement of all stakeholders, drawing on their unique experiences to shape the design of tools, products, or programs [31,32].

Phase 1 used a co-design approach with a combination of the Design Council's Double Diamond framework [33] and work conducted in New Zealand [31] (Figure 1) to explore participants' views, knowledge, and perceptions regarding the use of imaging AI technology to collect data to enable the prediction of DFU healing and develop mitigation strategies to address these.

- *Discover:* Undertaking literature searches of (a) the variables that impact the ability of DFUs to heal and (b) devices used to predict the healing of DFUs.
- *Define:* Refining the literature findings to generate a near-final list to include in the device regarding device features and variables to include in the validation of the device.
- *Develop:* Generate a first prototype and then work with key stakeholders to refine the device features and variables to be included as part of the device.
- *Deliver:* Test and validate in the pilot.

Figure 1. A co-design framework for this study.

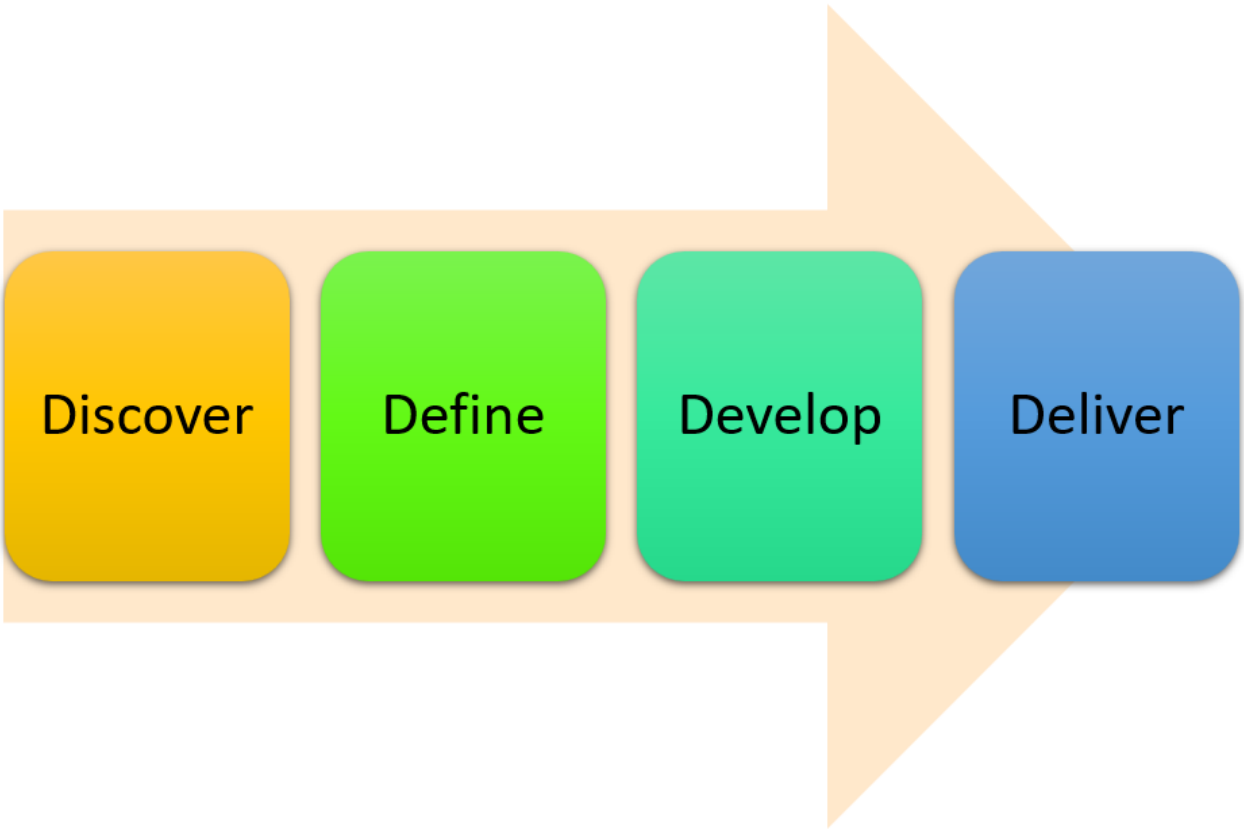

Participants

Inclusion criteria for key stakeholders are: (i) clinicians working with people who have DFUs, (ii) adults living with diabetes who have a history of or a current foot ulcer, and (iii) biomedical engineer research team members.

XSL•FO
RenderX

Recruitment

Clinicians of participating High Risk Foot Services were identified by the research team and invited either in person, via email, or by telephone to participate in the focus group sessions. Existing adult patients of the participating High Risk Foot Services who have diabetes and current or history of a foot ulcer were informed of the research by clinicians working in these services and directed to the advertisement at participating hospitals. All the biomedical engineers from the research team participated in the study. All participants provided informed consent to participate in the study.

Data Collection

Facilitator-guided (RO, DK) face-to-face group discussion was conducted with clinicians working with people who have DFUs (n=9) and biomedical engineer research team members (n=2). The session was held in a meeting room at a participating hospital during a time that suited participants, and a working lunch was provided. The device was explained, and a prototype of the software-based medical device was shown to the participants. The meeting was conversational, and participants were asked their opinion on the device concept and what would be needed to use the device at their workplace. Notes were taken of the discussion by 2 scribes (BP and QCN). Clinicians were also asked about their employment position, level and type of education, duration in the field, and duration with the current health service. Follow-up information was obtained through emails between the biomedical engineers and clinicians.

Individual interviews have also been undertaken by a research team member with experience in qualitative research (RO) with 4 adults living with diabetes who had a history of (n=1) or a current foot ulcer (n=3). Interviews were undertaken in a private area within the hospital. The participants were informed about the purpose and design of the device, and participants were asked to share their thoughts on the design of the device and the format and information of the report from the device. They were also asked for their thoughts on the process of collecting the image and the presentation of the information obtained by the device. The interviews were recorded on Microsoft Teams, transcribed using the embedded AI within the program, and validated by a researcher (RO). Participants were also asked about their education level, duration of diabetes, gender, and history of any current or previous lower limb wound.

Co-Design Data Analysis

The research team (RO, DK, BP, QCN, and EIE) reviewed the data generated from the co-design to become familiar with the content and key ideas that were raised during the co-design session. These team members met as a group to discuss what concepts they identified from their respective perspectives, coming to a consensus on core themes, as well as additional questions to be raised subsequently for future research. A structured summary was developed by 1 member of the team (RO), with input from the other research team members (DK, BP, and QCN), where the final draft was shared with participants for their input. Further discussion with the research team followed, considering the additional input from participants, and a final summary including next steps was generated, then shared with the participants.

Decisions to incorporate co-design findings in developing the prototype device for use in Phase 2 were made by the biomedical engineers, considering whether the recommendations were (1) a priority to the participants, (2) would make a significant improvement for data collection, and (3) feasible to incorporate, that is not overly costly or require considerable resources, such as engineer time.

Interview Data Analysis

The researchers will use an interpretive paradigm, utilizing a relativist approach to ontology and a constructivist approach to epistemology. Reflexive thematic analysis, both deductive and inductive approaches, will be used [34]. Deductive codes are based on usability, accessibility, clinical integration, and support needs. Inductive codes emerge through open coding of the transcripts, which capture concerns and suggestions. To generate key themes and concepts, a researcher will review the interview transcripts to become familiar with the content and consider what key ideas were raised during the interviews. The data will also consider specifics around device use and how it can be best explained by health care providers.

Improvement of the Prototype

The current version of the prototype of the thermal imaging device with the smartphone app has been named Thermul. The Thermul will be shown to and used by the stakeholders during the co-design workshops. The feedback will be obtained from these workshops and will be used to modify the device such that it is better accepted by the users and people living with the disease. The major factors that require modification will be identified, and the modified device will be ready for the next discussion with the stakeholders.

Phase 2: Validation of the Technology

Study Design

Participants will be enrolled in this longitudinal observational study for 12 weeks to test whether differences between textural features of thermal images of DFUs can predict ulcer healing at 12 weeks from an image taken at week 1.

Participants

Inclusion criteria (Textbox 1) will be adults with diabetes and peripheral neuropathy with a loss of protective sensation and a foot ulcer of any duration with Grade 0 WIfI Wound Infection and Grade 0 WIfI Ischemia classifications [35]. Loss of protective sensation will be determined by the detection of the 10 g (5.07 Semmes-Weinstein) monofilament, according to the sensory foot examination described by Schaper et al [30]. People with signs or symptoms of foot ulcer infection (WIfI infection grade 1 - 3), and peripheral arterial disease (WIfI ischemia grades 1 - 3) [35] will be excluded. Participants must have sufficient English and cognition (Rowland Universal Dementia Assessment Scale score >22 [36]) to understand the study information and be available to participate over the study period of 12 weeks.

XSL•FO
RenderX

Textbox 1. Inclusion criteria of the study are as follows:

- Diabetes diagnosis
- Peripheral neuropathy with loss of protective sensation
- Foot ulcer of any duration
- Wound, Ischemia, foot Infection (WIfI) Wound Infection grade 0 (uninfected)
- WIfI Ischemia grade 0 (Ankle-Brachial Index ≥0.8 or ankle systolic pressure >100 mmHg or toe pressure ≥60 mmHg).

Recruitment

Patients attending the participating High Risk Foot Services who fulfill the study inclusion criteria will be invited to participate in the study.

Sample Size Calculation

Based on published epidemiological data [37-41] and supported by other pilot work [13], the expected population proportion of true positive ratings for wound healing within 12 weeks is 0.7. Using the goodness-of-fit approach [42], recruiting 120 patients would yield 0.8 power to reject the null hypothesis (substantial agreement, Kappa=0.7) in favor of the alternative hypothesis of almost perfect agreement (Kappa=0.9) using a 2-sided significance threshold of 0.05. For the primary analysis, the concordance between the model and the true status of wound healing at 12 weeks will be evaluated using Cohen Kappa measure for statistical agreement and reported with a respective 95% CI. In addition, sensitivity, specificity, area under the receiver operating characteristic curve, and diagnostic odds ratio will be reported to evaluate the device's performance as a diagnostic test for determining the true wound-healing status at 12 weeks.

Participants will not be matched but will rather act as their own control. Areas of normal skin from the same image will be compared with areas of ulceration in the digital images. The differences between these 2 sites and changes in these sites over time will enable us to determine if a wound is improving or deteriorating with time.

Thermal Imaging Device

The imaging methodology is based on the work of Ngo et al [27], with the base camera specifications shown in Table 1. The device, named Thermul, AU $1400 (US $939.48), takes a thermal and standard RGB image simultaneously in a physically fixed configuration. This ensures that the 2 images do not display measurable parallax. The FLIR One device (thermal camera) is calibrated before each use using the company-provided calibration function. The RGB image with higher resolution is taken from the built-in camera of the DOOGEE (Android) phone. Calibration is undertaken by placing a fixed-size marker next to the wound. Image registration is performed by software that tests the registration of the 2 modalities based on the wound edges. While previous works have used temperature as a feature, our approach uses texture, and this approach has been designed to overcome the ambient temperature limitation present when using thermal imaging of the foot.

Photos are being taken in normal consulting clinic rooms. The equipment has been designed to require minimal training from the user, allowing it to be used much like a regular smartphone. Each staff member is trained to use the device using a minimum of 2 examples, to ensure they are confident to record images independently.

Table . Base camera specifications prior to co-design.

Camera specifications (minimum)	Thermal camera	Color camera (standard RGB)[a]
Brand options	FLIR or similar	DOOGEE
Model number options	One Pro or similar	S96Pro or similar
Image size (Pixels)	160×120 (thermal resolution) or higher	4000×3000 or higher
Physical dimension (L×W×H) options	68×34×14 mm	167×81.4×15.5
Spectral range	8-14 µm	RGB
Weight	≤40 g	310 g
Power	Battery-USB[b] data/charging port	Battery-USB data/charging port
Mode of action	Snapshot	Snapshot

[a]RGB; red, blue, and green.

[b]USB: Universal Serial Bus.

Data Collection

Participant demographic information includes age, gender, socioeconomic factors, and years of formal education. Factors related to the healing of DFUs will also be collected, the specifics of which will be informed on the finalization of the findings of Phase 1, and may include: body mass index, smoking history, retinopathy, types of diabetes treatment, types of DFU treatment, ulcer duration, area of ulcer, depth of ulcer, location of ulcer, infection, blood glucose level, and estimated glomerular

XSL•FO
RenderX

filtration rate [43-48]. This information may be revised at a later stage and, if appropriate, will be added to the prediction model.

The following data are collected at baseline, 1-, 2-, and 12-week follow-up visits for all participants (Figure 2):

1. Any signs and symptoms of foot ulcer infection as per WIfI Classification [35].
2. After the ulcer dressing has been removed and prior to any wound hygiene, cleansing, or debridement, the following images will be taken (detailed specifications in Table 1):
   a. 160×120 pixel thermal images using the FLIR One Pro.
   b. 4000×3000 pixel standard (RGB) using a DOOGEE S96Pro camera.
   c. Wound dimensions (length, width, and depth in millimeters).

Figure 2. Study timeline.

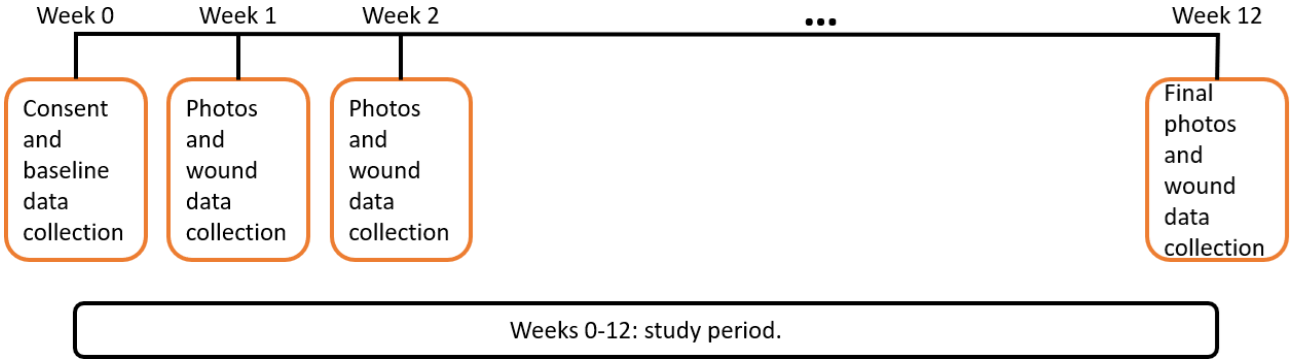

Images are taken by trained research and clinical staff, using the following method:

1. The participant is sitting or lying in the bed in a position to ensure that the ulcer is clearly visible.
2. As the 2 cameras are fixed together, the angle of capture of both cameras is the same, with the focus on the ulcer (Figure 3).
   a. Both cameras are snapshot cameras, so the images are captured in a single shot, virtually instantaneously.
   b. In cases of multiple ulcers in 1 foot or in both feet, each ulcer is considered as an independent ulcer.
   c. The imaging uses a fixed-size marker that allows for recording independent images for each ulcer .
3. The images are taken twice while the participant is barefoot.
   a. The first image is taken immediately after the wound dressing is removed, and the second approximately 15 minutes later, following all wound hygiene and debridement, as required and part of the usual wound care.

XSL•FO
RenderX

Figure 3. Thermal data collection setup.

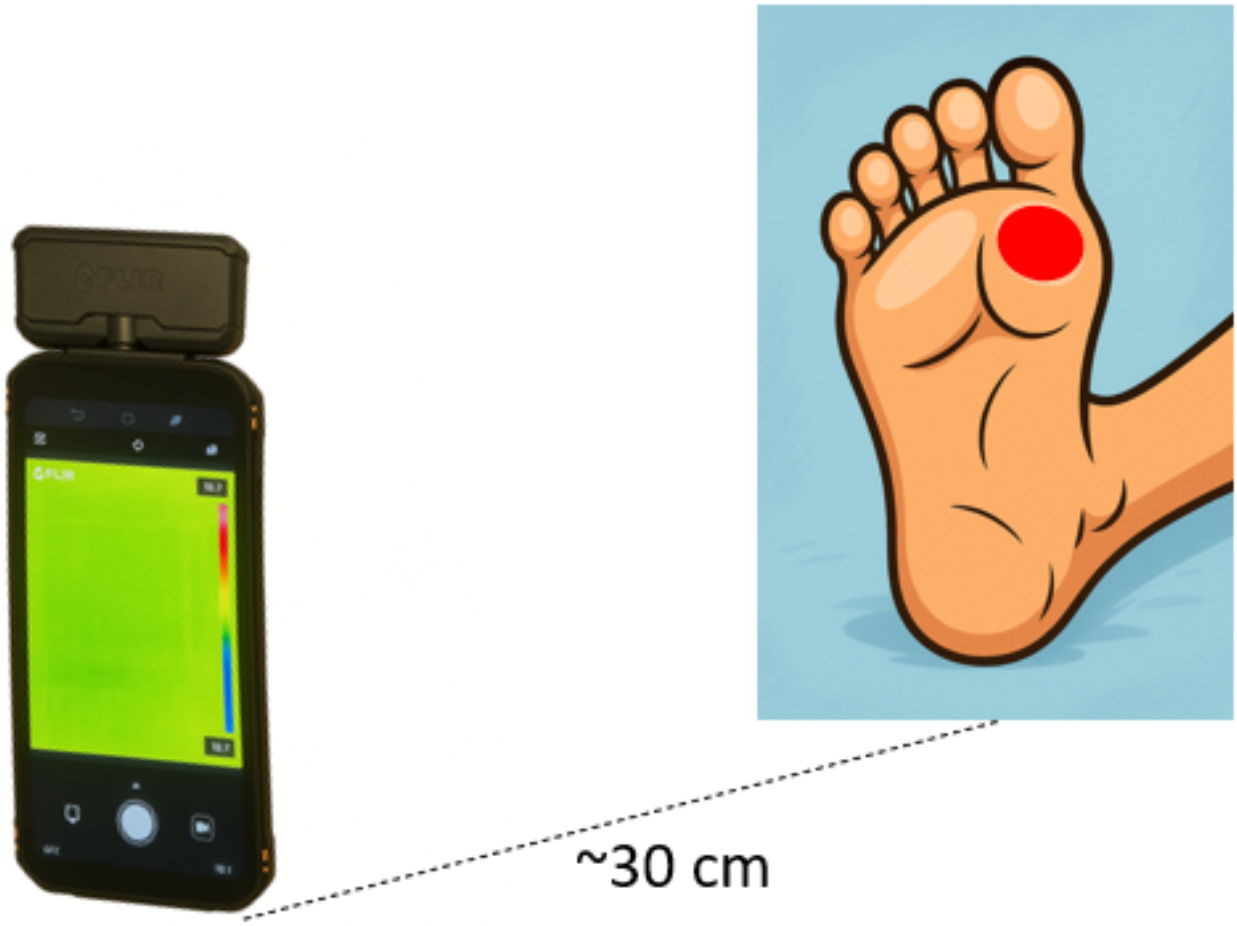

Statistical Analysis

Thermal images will first be analyzed to obtain textural features using the gray-level co-occurrence matrix. This method helps to extract clinically relevant information from thermal images of wounds by quantifying the spatial distribution of temperature variations rather than relying solely on mean temperature values. Wound tissue often exhibits heterogeneous heat patterns due to differences in vascularization, inflammation, infection, or necrosis, and these subtle variations are best captured through gray-level co-occurrence matrix-derived features. For example, contrast highlights sharp thermal gradients at wound borders, homogeneity reflects the degree of uniform healing, entropy indicates irregularity or randomness associated with pathological processes, and energy captures the smoothness of well-healed tissue [49]. Following this, textural features will be used to identify the healing trajectory of the wound. The textural features will first be reduced using principal component analysis, and the selected features will be classified using different machine learning models, such as logistic regression and Bayesian neural networks, to obtain the probability of healing. The machine learning models will be trained on outcome labels (healed or unhealed) determined by podiatrists at the 12-week clinical evaluation. Two cross-validation approaches will be used: leave-one-out and *k*-fold. The receiver operating characteristic curve will determine the threshold for optimizing the training of the model for the prediction of unhealed wounds. The hypothesis is that factors such as size and infection history should not alter the texture of the thermal image of the wound, and hence should not play a role in the assessment by this modality. The patient demographic information will be analyzed using multivariate logistic regression.

Ethical Considerations

This study follows the Declaration of Helsinki for human participants, which considers respect for the individual, his or her right to self-determination, and the right to make informed decisions regarding participation in research [50]. Human Research Ethics project approval has been obtained from the Austin Health Human Research Ethics Committee (HREC/92993/Austin-2023) and from St Vincent's Hospital Melbourne (2024/PID00243). Written consent is obtained from all participants.

Participant privacy and confidentiality were strictly maintained throughout the study. All data were deidentified prior to analysis, stored on secure, password-protected servers at RMIT University, and accessed only by authorized members of the research team. No personally identifiable information is included in any publications.

XSL•FO
RenderX

Results

Data collection for Phase 1 commenced in February 2024, and co-design workshops with clinicians, biomedical engineers, and clinical researchers concluded in June 2025. Co-design data collection with adults living with diabetes who have a history or a current foot ulcer continues and is expected to conclude in April 2026. Phase 2 commenced in May 2024 and is also expected to conclude in April 2026.

Co-design was undertaken with 9 clinicians (podiatrists, prosthetist-orthotists, and an endocrinologist), 2 biomedical engineers, and 2 clinician researchers, with follow-up information obtained from 6 clinicians.

Participants identified the practical aspects to support device use in 4 key themes: (1) insights for sector use; (2) how to document and use the device; (3) legal implications; and (4) workflow requirements. These are outlined in Textbox 2.

Textbox 2. Practical aspects to support device use.

Insights for sector use

- Device more useful for primary care/community health and regional/rural areas where expertise is limited
- Suggest data from the device can supplement existing diagnostics in high-risk foot clinics, not replace them

How to document and use device

- Percentage prediction of healing—refer/not to expert care
- To inform the person regarding the need for referral, if required

Legal implications

- What if the device suggests favorable results and the opposite happens?
- What if the device is used by untrained professionals?

Workflow requirements

- Battery charging duration; more robust device; multiple wounds
- How to ensure the taken image is viable—visual feedback
- Report generated by device—content needs further discussion

Discussion

Principal Findings

This study aims to co-design, test, and validate a feasible and acceptable device that can accurately predict the healing of neuropathic DFUs. Clinicians provided guidance on practical aspects to enable optimal device use in the clinical setting. The next phase involves collecting images, and a computer model will analyze those wounds that have healed within a 12-week period. The device will use textural analysis and machine learning to accurately predict the healing of DFUs at week 12 from an image taken at the first visit.

Many health care technologies have been developed, yet very few are used in routine clinical care due to limitations in the device's usability and technologies not catering to the pressing needs of end users [51]. This study is strengthened by engaging with key stakeholders in the early part of prototype development. We anticipate this early engagement will lead to a device that is fit for purpose and more likely to be used by the target clinicians upon validation and release to market [51].

Co-design approaches are used in health care innovation to harness the ideas and insights of key stakeholders, including consumers, to ensure that what is co-designed is more likely to meet the needs of the end user(s) and be feasible to use in target context(s) [52]. Co-design involves working collaboratively in partnership with key stakeholders, based on an equal and reciprocal relationship, to develop ways of doing things that truly meet people's needs and preferences [53]. As such, it extends beyond traditional approaches to stakeholder engagement, such as informing, which involves sharing information about an intended service or intervention with no feedback sought; and consultation, which involves seeking feedback on an intended service or intervention with the intention of making changes based on responses [54].

This study will develop a software-based nontouch device to detect the healing status of the wound in the first presentation. The identification of variables that can be used in ulcers that have a greater likelihood of healing will enable the generation of a machine learning algorithm. Once this algorithm is developed, further research can be undertaken on ulcer types that have a limited capacity to heal to refine the algorithm for a broader range of DFUs, including ischemic and neuro-ischemic ulcers.

This technology has overcome a number of limitations identified in previous studies [13,24,26-28], including faster speed of image capture, calculation of image data, and reduced cost compared to other diagnostic devices. Ambient temperature, lighting, and skin color and texture also do not influence thermal image quality. This technology is also suitable for investigating the ulcers of any duration or stage of healing.

XSL•FO
RenderX

Conclusions

The findings of this study, a combined co-design and longitudinal study, will develop a device aimed at predicting the healing of DFUs. The study will determine if textural characteristics from thermal images of neuropathic DFUs can identify which ulcers will not heal by 12 weeks. If successful, the device may be used to guide treatment approaches and initiate a more timely escalation of care for those ulcers not expected to heal in 12 weeks.

Funding

This work was undertaken as part of the Australian Centre for Accelerating Diabetes Innovations (ACADI), which was established through Medical Research Future Fund (MRFF) funding from the Australian Government's Targeted Translation Research Accelerator program, delivered by MTPConnect.

Authors' Contributions

DK – Co-guarantor, conceptualization, protocol development, writing – review and editing.
EIE – Co-guarantor, conceptualization, protocol development, writing – review and editing.
RO – Protocol development, writing – review and editing.
QCN – Protocol development, writing – review and editing.
BP – Writing – review and editing.
SH – Protocol development, writing – review and editing.
KW – Protocol development, writing (lead) - review and editing.
RJM – Protocol development, writing – review and editing.
AG – Protocol development, writing – review and editing.
ANT – Protocol development, writing – review and editing.

Conflicts of Interest

EIE has received research funding from Eli Lilly, Novo Nordisk, Boehringer Ingelheim, Novartis, and Versanis for unrelated projects.

Checklist 1
STROBE checklist.
[DOCX File, 34 KB - resprot_v15i1e69793_app1.docx]

References

1. Saeedi P, Petersohn I, Salpea P, et al. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: results from the International Diabetes Federation Diabetes Atlas, 9th edition. Diabetes Res Clin Pract 2019 Nov;157:107843. [doi: 10.1016/j.diabres.2019.107843] [Medline: 31518657]
2. Zhang Y, Lazzarini PA, McPhail SM, van Netten JJ, Armstrong DG, Pacella RE. Global disability burdens of diabetes-related lower-extremity complications in 1990 and 2016. Diabetes Care 2020 May;43(5):964-974. [doi: 10.2337/dc19-1614] [Medline: 32139380]
3. Zhang P, Lu J, Jing Y, Tang S, Zhu D, Bi Y. Global epidemiology of diabetic foot ulceration: a systematic review and meta-analysis. Ann Med 2017 Mar;49(2):106-116. [doi: 10.1080/07853890.2016.1231932] [Medline: 27585063]
4. Jodheea-Jutton A, Hindocha S, Bhaw-Luximon A. Health economics of diabetic foot ulcer and recent trends to accelerate treatment. Foot (Edinb) 2022 Sep;52:101909. [doi: 10.1016/j.foot.2022.101909] [Medline: 36049265]
5. Lazzarini PA, Gurr JM, Rogers JR, Schox A, Bergin SM. Diabetes foot disease: the Cinderella of Australian diabetes management? J Foot Ankle Res 2012 Oct 1;5(1):24. [doi: 10.1186/1757-1146-5-24] [Medline: 23021818]
6. Lazzarini PA, Hurn SE, Kuys SS, et al. Direct inpatient burden caused by foot-related conditions: a multisite point-prevalence study. BMJ Open 2016 Jun 20;6(6):e010811. [doi: 10.1136/bmjopen-2015-010811] [Medline: 27324710]
7. Lazzarini PA, Hurn SE, Kuys SS, et al. The silent overall burden of foot disease in a representative hospitalised population. Int Wound J 2017 Aug;14(4):716-728. [doi: 10.1111/iwj.12683] [Medline: 27696693]
8. McDermott K, Fang M, Boulton AJM, Selvin E, Hicks CW. Etiology, epidemiology, and disparities in the burden of diabetic foot ulcers. Diabetes Care 2023 Jan 1;46(1):209-221. [doi: 10.2337/dci22-0043] [Medline: 36548709]
9. Fife CE, Eckert KA, Carter MJ. Publicly reported wound healing rates: the fantasy and the reality. Adv Wound Care (New Rochelle) 2018 Mar 1;7(3):77-94. [doi: 10.1089/wound.2017.0743] [Medline: 29644145]
10. Fife CE, Horn SD, Smout RJ, Barrett RS, Thomson B. A predictive model for diabetic foot ulcer outcome: the wound healing index. Adv Wound Care (New Rochelle) 2016 Jul 1;5(7):279-287. [doi: 10.1089/wound.2015.0668] [Medline: 27366589]

XSL•FO
RenderX

11. Boulton AJM, Vileikyte L, Ragnarson-Tennvall G, Apelqvist J. The global burden of diabetic foot disease. The Lancet 2005 Nov;366(9498):1719-1724. [doi: 10.1016/S0140-6736(05)67698-2]
12. Frykberg RG, Banks J. Challenges in the treatment of chronic wounds. Adv Wound Care (New Rochelle) 2015 Sep 1;4(9):560-582. [doi: 10.1089/wound.2015.0635] [Medline: 26339534]
13. Aliahmad B, Tint AN, Poosapadi Arjunan S, et al. Is thermal imaging a useful predictor of the healing status of diabetes-related foot ulcers? A pilot study. J Diabetes Sci Technol 2019 May;13(3):561-567. [doi: 10.1177/1932296818803115] [Medline: 30255722]
14. Godavarty A, Leiva K, Amadi N, Klonoff DC, Armstrong DG. Diabetic foot ulcer imaging: an overview and future directions. J Diabetes Sci Technol 2023 Nov;17(6):1662-1675. [doi: 10.1177/19322968231187660] [Medline: 37594136]
15. Rogers LC, Frykberg RG, Armstrong DG, et al. The Charcot foot in diabetes. Diabetes Care 2011 Sep;34(9):2123-2129. [doi: 10.2337/dc11-0844] [Medline: 21868781]
16. Arnon-Sheleg E, Keidar Z. Diabetic foot infection: the role of PET/CT imaging. Curr Pharm Des 2018;24(12):1277-1286. [doi: 10.2174/1381612824666180227095439] [Medline: 29484984]
17. Nawaz A, Torigian DA, Siegelman ES, Basu S, Chryssikos T, Alavi A. Diagnostic performance of FDG-PET, MRI, and plain film radiography (PFR) for the diagnosis of osteomyelitis in the diabetic foot. Mol Imaging Biol 2010 Jun;12(3):335-342. [doi: 10.1007/s11307-009-0268-2] [Medline: 19816744]
18. Rubitschung K, Sherwood A, Crisologo AP, et al. Pathophysiology and molecular imaging of diabetic foot infections. Int J Mol Sci 2021 Oct 26;22(21):11552. [doi: 10.3390/ijms222111552] [Medline: 34768982]
19. Jørgensen LB, Sørensen JA, Jemec GB, Yderstraede KB. Methods to assess area and volume of wounds - a systematic review. Int Wound J 2016 Aug;13(4):540-553. [doi: 10.1111/iwj.12472] [Medline: 26250714]
20. Rani P, Aliahmad B, Kumar DK. The association of temperature of diabetic foot ulcers with chronic kidney disorder. Presented at: 2019 41st Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC); Jul 23-27, 2019; Berlin, Germany p. 2817-2820. [doi: 10.1109/EMBC.2019.8856401]
21. Rani P, Aliahmad B, Kumar DK. Curve irregularity index for quantification of roughness in non-fractal curves. Phys. A 2019 Aug;528:121435. [doi: 10.1016/j.physa.2019.121435]
22. Scebba G, Zhang J, Catanzaro S, et al. Detect-and-segment: a deep learning approach to automate wound image segmentation. Informatics in Medicine Unlocked 2022;29:100884. [doi: 10.1016/j.imu.2022.100884]
23. Marijanović D, Filko D. A systematic overview of recent methods for non-contact chronic wound analysis. Appl Sci (Basel) 2020;10(21):7613. [doi: 10.3390/app10217613]
24. Poosapadi Arjunan S, Tint AN, Aliahmad B, et al. High-resolution spectral analysis accurately identifies the bacterial signature in infected chronic foot ulcers in people with diabetes. Int J Low Extrem Wounds 2018 Jun;17(2):78-86. [doi: 10.1177/1534734618785844] [Medline: 30012069]
25. Barone S, Paoli A, Razionale AV. Assessment of chronic wounds by three-dimensional optical imaging based on integrating geometrical, chromatic, and thermal data. Proc Inst Mech Eng H 2011 Feb;225(2):181-193. [doi: 10.1243/09544119JEIM705] [Medline: 21428152]
26. Ogrin R, Motin MA, Aliahmad B, Elder K, Anderson J, Kumar DK. Can thermal imaging technique be used to predict the healing status of a venous leg ulcer? Int J Low Extrem Wounds 2023 Mar;22(1):85-92. [doi: 10.1177/15347346211002351] [Medline: 33856237]
27. Ngo QC, Ogrin R, Kumar DK. Computerised prediction of healing for venous leg ulcers. Sci Rep 2022 Oct 26;12(1):17962. [doi: 10.1038/s41598-022-20835-y] [Medline: 36289299]
28. Monshipouri M, Aliahmad B, Ogrin R, et al. Thermal imaging potential and limitations to predict healing of venous leg ulcers. Sci Rep 2021 Jun 24;11(1):13239. [doi: 10.1038/s41598-021-92828-2] [Medline: 34168251]
29. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. J Clin Epidemiol 2008 Apr;61(4):344-349. [doi: 10.1016/j.jclinepi.2007.11.008] [Medline: 18313558]
30. Schaper NC, van Netten JJ, Apelqvist J, et al. Practical guidelines on the prevention and management of diabetes-related foot disease (IWGDF 2023 update). Diabetes Metab Res Rev 2024 Mar;40(3):e3657. [doi: 10.1002/dmrr.3657] [Medline: 37243927]
31. Boyd H, McKernon S, Mullin B, Old A. Improving healthcare through the use of co-design. N Z Med J 2012 Jun 29;125(1357):76-87. [Medline: 22854362]
32. De Weger E, Van Vooren N, Luijkx KG, Baan CA, Drewes HW. Achieving successful community engagement: a rapid realist review. BMC Health Serv Res 2018 Apr 13;18(1):285. [doi: 10.1186/s12913-018-3090-1] [Medline: 29653537]
33. Design methods for developing services. Design Council. 2015. URL: https://www.designcouncil.org.uk/fileadmin/uploads/dc/Documents/DesignCouncil_Design%2520methods%2520for%2520developing%2520services.pdf [accessed 2026-01-13]
34. Braun V, Clarke V. Conceptual and design thinking for thematic analysis. Qualit Psychol 2022;9(1):3-26. [doi: 10.1037/qup0000196]
35. Mills JL Sr, Conte MS, Armstrong DG, et al. The society for vascular surgery lower extremity threatened limb classification system: risk stratification based on Wound, Ischemia, and foot Infection (WIfI). J Vasc Surg 2014 Jan;59(1):220-234. [doi: 10.1016/j.jvs.2013.08.003] [Medline: 24126108]

XSL•FO
RenderX

36. Storey JE, Rowland JTJ, Basic D, Conforti DA, Dickson HG. The Rowland Universal Dementia Assessment Scale (RUDAS): a multicultural cognitive assessment scale. Int Psychogeriatr 2004 Mar;16(1):13-31. [doi: 10.1017/s1041610204000043] [Medline: 15190994]
37. Snyder RJ, Cardinal M, Dauphinée DM, Stavosky J. A post-hoc analysis of reduction in diabetic foot ulcer size at 4 weeks as a predictor of healing by 12 weeks. Ostomy Wound Manage 2010 Mar 1;56(3):44-50. [Medline: 20368673]
38. Cardinal M, Eisenbud DE, Phillips T, Harding K. Early healing rates and wound area measurements are reliable predictors of later complete wound closure. Wound Repair Regen 2008;16(1):19-22. [doi: 10.1111/j.1524-475X.2007.00328.x] [Medline: 18211575]
39. Lavery LA, Barnes SA, Keith MS, Seaman JW Jr, Armstrong DG. Prediction of healing for postoperative diabetic foot wounds based on early wound area progression. Diabetes Care 2008 Jan;31(1):26-29. [doi: 10.2337/dc07-1300] [Medline: 17934156]
40. Sheehan P, Jones P, Caselli A, Giurini JM, Veves A. Percent change in wound area of diabetic foot ulcers over a 4-week period is a robust predictor of complete healing in a 12-week prospective trial. Diabetes Care 2003 Jun;26(6):1879-1882. [doi: 10.2337/diacare.26.6.1879] [Medline: 12766127]
41. Rajput D, Wang WJ, Chen CC. Evaluation of a decided sample size in machine learning applications. BMC Bioinformatics 2023 Feb 14;24(1):48. [doi: 10.1186/s12859-023-05156-9] [Medline: 36788550]
42. Donner A, Eliasziw M. A goodness-of-fit approach to inference procedures for the kappa statistic: confidence interval construction, significance-testing and sample size estimation. Stat Med 1992 Aug;11(11):1511-1519. [doi: 10.1002/sim.4780111109] [Medline: 1410963]
43. Bundó M, Vlacho B, Llussà J, et al. Prediction of outcomes in subjects with type 2 diabetes and diabetic foot ulcers in Catalonian primary care centers: a multicenter observational study. J Foot Ankle Res 2023 Feb 28;16(1):8. [doi: 10.1186/s13047-023-00602-6] [Medline: 36849888]
44. Jalilian M, Ahmadi Sarbarzeh P, Oubari S. Factors related to severity of diabetic foot ulcer: a systematic review. Diabetes Metab Syndr Obes 2020;13:1835-1842. [doi: 10.2147/DMSO.S256243] [Medline: 32547145]
45. Tehan PE, Burrows T, Hawes MB, et al. Factors influencing diabetes-related foot ulcer healing in Australian adults: a prospective cohort study. Diabet Med 2023 Jan;40(1):e14951. [doi: 10.1111/dme.14951] [Medline: 36054775]
46. Wang S, Xia C, Zheng Q, Wang A, Tan Q. Machine learning models for predicting the risk of hard-to-heal diabetic foot ulcers in a Chinese population. Diabetes Metab Syndr Obes 2022;15:3347-3359. [doi: 10.2147/DMSO.S383960] [Medline: 36341229]
47. Zhang Y, Cramb S, McPhail SM, et al. Multiple factors predict longer and shorter time-to-ulcer-free in people with diabetes-related foot ulcers: survival analyses of a large prospective cohort followed-up for 24-months. Diabetes Res Clin Pract 2022 Mar;185:109239. [doi: 10.1016/j.diabres.2022.109239] [Medline: 35131379]
48. Zhang Y, Cramb S, McPhail SM, et al. Factors associated with healing of diabetes-related foot ulcers: observations from a large prospective real-world cohort. Diabetes Care 2021 Jul;44(7):e143-e145. [doi: 10.2337/dc20-3120] [Medline: 34074651]
49. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst, Man, Cybern 1973;SMC-3(6):610-621. [doi: 10.1109/TSMC.1973.4309314]
50. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human participants. JAMA 2025 Jan 7;333(1):71-74. [doi: 10.1001/jama.2024.21972] [Medline: 39425955]
51. Whitelaw S, Pellegrini DM, Mamas MA, Cowie M, Van Spall HGC. Barriers and facilitators of the uptake of digital health technology in cardiovascular care: a systematic scoping review. Eur Heart J Digit Health 2021 Mar;2(1):62-74. [doi: 10.1093/ehjdh/ztab005] [Medline: 34048508]
52. Greenhalgh T, Jackson C, Shaw S, Janamian T. Achieving research impact through co-creation in community-based health services: literature review and case study. Milbank Q 2016 Jun;94(2):392-429. [doi: 10.1111/1468-0009.12197] [Medline: 27265562]
53. Palmer VJ, Weavell W, Callander R, et al. The Participatory Zeitgeist: an explanatory theoretical model of change in an era of coproduction and codesign in healthcare improvement. Med Humanities 2019 Sep;45(3):247-257. [doi: 10.1136/medhum-2017-011398]
54. Bird M, McGillion M, Chambers EM, et al. A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. Res Involv Engagem 2021 Mar 1;7(1):12. [doi: 10.1186/s40900-021-00252-7] [Medline: 33648588]

Abbreviations

AI: artificial intelligence
DFUs: diabetes-related foot ulcers
RGB: red, blue, and green
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WIfI: Wound, Ischemia, foot Infection

XSL•FO
RenderX

Edited by J Sarvestan; submitted 08.Dec.2024; peer-reviewed by A Ghosh, D Leizaola, I Julian-Rochina; revised version received 02.Dec.2025; accepted 03.Dec.2025; published 26.Feb.2026.

Please cite as:
Waller K, Ogrin R, Ngo QC, Polus B, Hanna S, Tint AN, MacIsaac RJ, Galligan A, Ekinci EI, Kumar D
Using Textural Analysis of Thermal Imaging to Predict Healing Status of Diabetes-Related Neuropathic Foot Ulcers: Protocol for a Co-Design and Longitudinal Study
JMIR Res Protoc 2026;15:e69793
URL: https://www.researchprotocols.org/2026/1/e69793
doi:10.2196/69793

© Kate Waller, Rajna Ogrin, Quoc Cuong Ngo, Barbara Polus, Samantha Hanna, Aye Nyein Tint, Richard J MacIsaac, Anna Galligan, Elif Ilhan Ekinci, Dinesh Kumar. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Educational Intervention to Improve Sexual Health and Quality of Life in Survivors of Breast and Gynecological Cancer: Protocol for a Mixed Methods Feasibility Study

Nathalia Salazar Falla[1,2*], MEd, MD; María de los Ángeles Navas Jojoa[3*], MD; Juan David Lalinde Riviño[4*], MD; Auramaría Escandón Bucheli[5*], MD; Juan Sebastián Galindo-Sánchez[2,6*], MSc, MD

[1]Department of Internal Medicine, Fundación Valle del Lili, Cra 98 No. 18 - 49, Cali, Valle del Cauca, Colombia

[2]Facultad de Ciencias de la Salud, Universidad Icesi, Cali, Valle del Cauca, Colombia

[3]Department of Gynecology, Consorcio Clínica Nueva Rafael Uribe Uribe, Cali, Valle del Cauca, Colombia

[4]Department of Gynecologic Oncology, Fundación Valle del Lili, Cali, Valle del Cauca, Colombia

[5]Department of Psychiatry, Fundación Valle del Lili, Cali, Valle del Cauca, Colombia

[6]Centro de Investigaciones Clínicas, Fundación Valle del Lili, Cali, Valle del Cauca, Colombia

[*]all authors contributed equally

Corresponding Author:
Nathalia Salazar Falla, MEd, MD
Department of Internal Medicine, Fundación Valle del Lili, Cra 98 No. 18 - 49, Cali, Valle del Cauca, Colombia

Abstract

Background: Sexual health is often underaddressed in cancer care, especially among survivors of breast and gynecological cancer. In Colombia, treatment side effects, cultural stigma, and limited training among health care providers affect well-being. Although international guidelines promote patient-centered and culturally sensitive approaches, few interventions have been implemented in low- and middle-income settings.

Objective: This study aimed to describe the feasibility and acceptability of an educational psychosexual intervention for survivors of breast and gynecological cancer, including design, integration of qualitative and quantitative methods, core components, measurement time points, and a prespecified analysis plan.

Methods: This is a prospective, single-group feasibility study at a university hospital in Colombia. Participants are adult women who completed treatment for breast or gynecological cancer at least 2 years before enrollment. The intervention combines tiered educational workshops, semistructured group interviews, and optional individualized counseling. Qualitative data will explore perceptions, barriers, and facilitators related to sexual health and well-being. Quantitative outcomes include the Female Sexual Function Index and European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 before and after the intervention to estimate change. Pre- and postintervention scores will be compared with paired tests and CIs. Qualitative data will undergo reflexive thematic analysis. Mixed findings will inform feasibility, acceptability, and practical recommendations. The intervention was co-designed with survivors and a multidisciplinary team, refined by expert review and a manual, and pilot-tested to optimize acceptability, logistics, and fidelity.

Results: The study was funded in December 2024 and approved by the Comité de Ética en Investigación Biomédica of Fundación Valle del Lili. Recruitment began in March 2025 and is expected to end in September 2025. As of July 2025, 13 participants had enrolled (n=2, 15.4% survivors of breast cancer; n=11, 84.6% survivors of gynecological cancer). Based on pretrial piloting, shorter session blocks and optional hybrid delivery were made available to support adherence while preserving core content and objectives. Final analysis is planned for October 2025, with results expected in December 2025.

Conclusions: This protocol evaluates the feasibility and acceptability of a codesigned, context-adapted psychosexual education model. Findings will guide implementation refinements and the design of a future comparative study of effectiveness and longer-term outcomes and may inform sustainable integration of sexual health into survivorship care in low- and middle-income settings. The model is intended to be replicable and scalable across oncology and other chronic disease contexts.

International Registered Report Identifier (IRRID): DERR1-10.2196/80567

(JMIR Res Protoc 2026;15:e80567) doi:10.2196/80567

KEYWORDS

female genital neoplasms; breast neoplasms; survivors; health education; sexual health; quality of life

XSL•FO
RenderX

Introduction

Background

Among patients with gynecological cancer, research has traditionally focused on oncological treatments aimed at prolonging overall survival and progression-free survival. However, quality of life is increasingly recognized as crucial for women who survive these diseases. Sexual health is a fundamental component of overall well-being and quality of life, yet it remains one of the most overlooked aspects of oncological care, especially among women who have survived gynecological cancers. A systematic review by Sousa Rodrigues Guedes et al [1] reported that 30% to 80% of women with cancer experience sexual dysfunction and that the likelihood of this condition is 2.7 to 3.5 times higher than in women without cancer.

In gynecological cancers, where organs directly involved in sexual function may be affected, the burden of sexual dysfunction is particularly high. For example, Firmeza et al [2] reported a 5% to 16% prevalence of dyspareunia after laparoscopic radical hysterectomy, and Jensen et al [3] found that among patients with cervical cancer receiving radiotherapy, 55% reported mild to severe dyspareunia and 30% were dissatisfied with their sexual life. In Colombia, survivors face additional barriers to sexual well-being, including physical treatment effects, psychological distress, cultural taboos, and limited training among health care providers [4-7]. Across gynecological cancers, surgery, pelvic radiotherapy, chemotherapy, and endocrine therapies contribute to decreased desire and arousal, vaginal dryness, dyspareunia, pelvic floor pain, body image disturbance, and relationship strain; recent reviews highlight multifactorial mechanisms (anatomical, neurovascular, hormonal, and psychosocial) and underscore the need for proactive assessment during survivorship care [8].

Beyond anatomical or physiological measures, such as vaginal dilators or estrogen gels, the psychosexual component is central to women's sexual pleasure and recovery. Trials report benefits of psychosocial approaches: telephone counseling improved depressive symptoms and gynecological or cancer-specific concerns [9], and a nurse-led sexual rehabilitation program improved sexual function from 1 to 12 months after treatment [10].

Building on this, brief, structured clinician- and internet-delivered interventions have improved sexual concerns and communication, including a clinician communication pilot [11], documented patient preferences for routine discussion [12], and web-based psychoeducational programs (Fex-Can) under randomized evaluation [13,14]. Counseling frameworks such as PLISSIT (permission, limited information, specific suggestion, intensive therapy) and BETTER (bring up, explain, tell, time, education, record) provide practical, stepwise scripts for time-limited visits [15] and show comparative advantages [16], although the evidence base specific to survivors of gynecological cancer remains limited [9,10].

Worldwide, organizations such as the World Health Organization (WHO) [17]; the United Nations Educational, Scientific, and Cultural Organization [18]; and the World Association for Sexual Health (WAS) [19] advocate for integrated, patient-centered approaches that recognize sexual health as a fundamental right and an essential element of postcancer recovery. Despite these calls, translation into routine practice remains challenging in many low- and middle-income countries, where issues of access and cultural relevance persist [17,18]. Moreover, the evidence base is derived largely from high-income settings [20], motivating context-specific, culturally adapted interventions such as the one evaluated in this study.

This study is primarily designed to evaluate the feasibility and acceptability of implementing an educational psychosexual intervention for survivors of breast and gynecological cancer. As a secondary objective, it will estimate the preliminary effectiveness of the intervention. It is important to note that this is not a definitive effectiveness trial but an initial step to determine whether the intervention can be successfully implemented and is acceptable to participants.

Protocol Objectives

This protocol aims to describe the feasibility and acceptability of implementing EDUSEXONCO—an educational psychosexual intervention for survivors of breast and gynecological cancer—including its design, mixed methods integration, intervention components, measurement time points, and prespecified analysis plan.

Prespecified Study Aims

The primary aim is to describe the feasibility and acceptability of implementing an educational strategy to promote sexual health and quality of life in a cohort of Colombian patients (EDUSEXONCO project).

The secondary aims are to (1) estimate within-group change in sexual function (Female Sexual Function Index; FSFI) and health-related quality of life (European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30; EORTC QLQ-C30) from baseline (before session 1) to after the intervention (end of session 3) and (2) explore the relationship between quantitative changes in FSFI and EORTC QLQ-C30 scores and qualitative findings on sexual well-being, communication, agency, and body image from baseline (before session 1) to after the intervention (end of session 3).

Methods

Design

This is a prospective, single-center, single-group (pretest-posttest) mixed methods feasibility study designed to evaluate the implementation and acceptability of an educational psychosexual intervention for survivors of gynecological cancer. Estimating preliminary effectiveness is a secondary objective. This is not a definitive effectiveness trial but an initial step to determine whether the intervention can be delivered successfully and is acceptable to participants.

This study uses an embedded convergent mixed methods design, giving equal priority to quantitative and qualitative strands. Quantitative data (FSFI and EORTC QLQ-C30 scores) will be collected at baseline (before session 1) and after the intervention

XSL•FO
RenderX

(end of session 3). Qualitative data (focus groups, semistructured interviews, participant observation field notes, and anonymous diaries) will be gathered after each session. Integration will occur during interpretation through meta-inferences and joint displays that juxtapose pretest-posttest quantitative changes with qualitative themes related to sexual well-being, communication, agency, and body image. We will seek convergence, complementarity, and expansion; discrepancies will be examined to refine inferences and inform practical recommendations.

Recruitment

Overview

We will use consecutive sampling from survivorship and gynecological oncology outpatient clinics. Procedures include (1) weekly electronic medical record prescreening (diagnosis, treatment completion of ≥2 years, age of ≥18 years, and active follow-up), (2) clinician referral during routine visits via a standardized slip, (3) approach in the clinic or telephone or email contact within 7 days, (4) brief scripted eligibility confirmation, and (5) provision of written informed consent (in Spanish). A recruitment log (approached, eligible, provided consent, and ineligible; reasons for decline; and dates) will quantify feasibility. Retention supports include flexible scheduling, reminders 48 to 72 hours before sessions, and transportation provided for participants who require it (funded by the study grant and arranged by the research team in coordination with institutional services).

Recruitment Setting

The study will take place at Fundación Valle del Lili (FVL), a university-affiliated tertiary hospital in Cali, Colombia, recognized as a leading center for high-complexity care, including oncology services, in Latin America. FVL provides coverage across southwestern Colombia, serving a population with significant ethnic diversity, and maintains agreements with multiple health plans. Since 2017, the institution has operated a functional cancer unit dedicated to delivering comprehensive care for patients with oncological diseases. This unit offers facilities and procedures for initiation, follow-up, and monitoring of cancer treatment supported by an experienced multidisciplinary team [21]. The study will be conducted from 2024 to 2025 within this setting.

Study Population

Participants will be adult women (aged ≥18 years) who have survived breast or gynecological cancer (cervical, uterine, vaginal, vulvar, or ovarian), completed oncological treatment at least 2 years prior to enrollment, and remain under follow-up at FVL.

Eligibility Criteria

The eligibility criteria are shown in Textbox 1.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Women aged ≥18 years
- History of breast or gynecological cancer (cervical, uterine, vaginal, vulvar, or ovarian)
- Completion of primary oncological treatment ≥24 months before enrollment (ongoing endocrine therapy or topical vaginal estrogen is allowed)
- Being disease free or having stable disease not requiring urgent oncological treatment as per the most recent clinic note
- Active follow-up at Fundación Valle del Lili
- Being Spanish speaking and able to complete study procedures
- Ability and willingness to provide written informed consent and attend the 3 intervention sessions (in person or hybrid)

Exclusion criteria

- Cognitive impairment that prevents informed participation (clinician judgment or Mini-Cog <3 if uncertain)
- Medical or psychiatric instability that would compromise safe participation (eg, uncontrolled pain crisis, acute severe depression with suicidality, or psychosis) per the treating clinician
- Active cancer recurrence requiring immediate treatment at screening
- Any condition that, in the investigator's judgment, precludes group participation (eg, major morbidity or infectious isolation requirements) or completion of study assessments
- Concurrent enrollment in another sexual health behavioral intervention study

A brief operational screening checklist is provided in Multimedia Appendix 1.

Intervention

EDUSEXONCO is a structured psychoeducational program to enhance sexual health and quality of life in survivors of gynecological cancer. The intervention targets modifiable psychosexual determinants—knowledge, communication, agency, body image, dyspareunia self-management, and pelvic floor function—through evidence-based education, brief skill practice, and optional individualized counseling to support recovery of sexual well-being.

XSL•FO
RenderX

Materials include a facilitator manual (objectives and scripts for sensitive topics), slides and handouts, 3D anatomical models, fidelity checklists, attendance and adherence logs, and brief take-home practice sheets. Delivery comprises three components: (1) large-group theoretical-practical workshops organized in 3 progressive levels (basic, intermediate, and advanced), (2) semistructured group interviews (6‐8 participants) to elicit lived experience and contextual barriers and facilitators, and (3) optional individual counseling when chosen. Brief skills (communication drills, breathing and relaxation, and pelvic floor basics) are embedded in each session.

EDUSEXONCO is delivered as a sequence of 3 progressive theoretical-practical workshop levels (basic, intermediate, and advanced) to support gradual learning and skill acquisition. Each level includes two complementary strands: (1) a medical-sexual strand focused on anatomy and sexual response, treatment-related sexual changes, symptom self-management (eg, lubrication, pain, and dyspareunia), and safe use of supportive resources; and (2) a sociopsychological strand focused on stigma reduction, communication, consent and boundaries, self-compassion and body image, emotional regulation, and social reintegration. Each workshop incorporates brief guided practice (eg, “anatomy tour” using 3D models, communication scripts, sensate focus–inspired exercises, paced breathing, pelvic floor downtraining, and gentle yoga and relaxation), a take-home task to reinforce learning between sessions, and fidelity markers (eg, coverage of key items, completion of role-plays or practice, and delivery of handouts) to standardize delivery across facilitators. The full-day schedule is provided in Multimedia Appendix 2. A detailed topic-by-level syllabus including objectives, key content, practices, materials, leads, take-home tasks, fidelity markers, and qualitative prompts is provided in Multimedia Appendix 3. Examples of the semistructured interview guide are provided in Multimedia Appendix 4.

In-person group delivery takes place in education or rehabilitation rooms with acoustic privacy at FVL comprising 3 sequential full-day sessions (approximately 8 hours each) over approximately 3 weeks (roughly 24 hours in total). The intervention includes optional home practice for 10 to 15 minutes per day.

Language and examples are culturally adapted to Colombia, alternatives are provided for participants without a partner, activities are adapted for pain or mobility limitations, and Spanish-language materials target approximately B1 to B2 readability.

Fidelity is monitored using per-session checklists covering objectives, core activities, materials, timing, inclusivity, skill practice, and end-of-session outputs (anonymous diaries). Attendance and adherence logs capture session-level participation (sessions 1‐3), completion of pretest and posttest questionnaires, participation in the group interview and optional 1:1 counseling, and take-home practice reporting.

Usual care is allowed (gynecology, psycho-oncology, pelvic floor therapy, and analgesia). Cointerventions (eg, topical estrogen, dilators, and sex therapy) are documented. Adverse events (eg, pelvic pain flare or marked emotional distress) are recorded and managed via institutional pathways; urgent cases are referred to psycho-oncology immediately.

A multidisciplinary team will deliver and evaluate the intervention. All facilitators will complete a 2-hour orientation on the facilitator manual, stigma-free and rights-based communication, trauma-informed practices, confidentiality, and use of fidelity tools (presession and per-session checklists). Interviewers will receive additional coaching on qualitative interviewing and reflexivity memoing. Each professional involved in the study will assume 2 primary roles: participation in the theoretical-practical workshops according to their field of expertise and moderation or support during semistructured interviews focusing on topics related to their professional area. The team comprises the following members:

- Principal investigator and study clinician (internist or gynecological oncologist)—oversees protocol fidelity and participant safety, addresses medical queries arising from workshops or interviews, and adjudicates eligibility or adverse event decisions as needed.
- Sexual health physician and sexologist—leads medical-sexual content on postcancer sexual physiology and function, pleasure, autoeroticism, and nonpharmacologic pain or dyspareunia strategies and provides brief one-to-one counseling when indicated.
- Psychologist—facilitates sessions on emotions, anxiety management, self-esteem, body image, and stigma; moderates group interviews; and manages emotional crises and referral to psycho-oncology per institutional pathways.
- Social worker—addresses sociocultural and socioeconomic barriers, family and social support networks, and reintegration and identifies cases requiring individualized follow-up or referral.
- Pelvic floor physical therapist—teaches basic pelvic floor contraction and relaxation, breathing, and posture for pain modulation; clarifies rehabilitation questions; and links exercises to sexual function.
- Nutritionist—provides guidance on nutrition to support recovery, energy, and quality of life; answers participant questions; and aligns advice with medical care.
- Yoga and relaxation instructor—leads brief, adapted breathing and posture practices to support emotional and physical well-being and offers alternatives for pain or mobility limitations.
- Sex educator—delivers comprehensive sexual education (communication skills, sexual aids, and diverse expressions of sexual health) using inclusive, culturally appropriate language.
- Data collection specialist and qualitative research assistant—manages audio recordings, field notes, deidentification, and secure transfer for transcription; maintains recruitment and fidelity logs; and supports postsession reflexivity memos.

Role Boundaries and Safety

Companions may attend large-group workshops when desired by the participant but do not join interviews or individual counseling. Only trained team members deliver content in their

XSL•FO
RenderX

scope; urgent clinical or emotional issues are escalated to the study clinician and psycho-oncology. All adverse events are documented and handled per institutional policy.

Measures and Outcomes

The primary objective is feasibility and acceptability. We will quantify feasibility (eligibility, invitation, consent, recruitment rate, retention, attendance and adherence, fidelity, and data completeness) and acceptability (postsession satisfaction, perceived usefulness, and comfort; recommendation item scored from 0 - 10). A priori thresholds are as follows: retention of ≥80%, attendance of ≥80% per session, fidelity of ≥80%, data completeness of ≥85%, no serious adverse events, acceptability mean of ≥4.0/5, and net promoter score of ≥30.

As a secondary objective, this study assesses quantitative outcomes using validated psychometric tools and qualitative dimensions using reflexive thematic analysis. Mixed methods integration will offer a comprehensive understanding of EDUSEXONCO's effects on sexual rehabilitation and psychosocial recovery.

Quantitative Component

Sexual Function

The FSFI is a validated 19-item instrument that assesses 6 domains of sexual function: desire, arousal, lubrication, orgasm, satisfaction, and pain. Each domain generates an individual score, and the total FSFI score is the sum of all domain scores, with a maximum of 36 points. A total score below 26.55 indicates sexual dysfunction. Authorization for the use of the FSFI has been obtained [22].

Health-Related Quality of Life

The EORTC QLQ-C30 version 3.0 is a validated 30-item instrument designed to assess health-related quality of life across 3 scales: functional, global quality of life, and symptoms. The functional and global scales range from 0 to 100, with higher scores indicating better functioning or overall quality of life. The symptom scale (eg, fatigue, pain, and insomnia) also ranges from 0 to 100, where lower scores reflect fewer symptoms and, therefore, better quality of life. Authorization for the use of the EORTC QLQ-C30 has been obtained [23].

Qualitative Component

Qualitative insights on barriers, facilitators, and perceptions of sexual health and well-being will be derived from focus groups, semistructured interviews, participant observation field notes, and anonymous postsession diaries.

Sociodemographic and Clinical Characteristics

At session 1 (before the intervention), participants will complete a brief self-administered questionnaire capturing age, educational level, employment status, relationship status, parity, menopausal status, time since diagnosis and completion of primary treatment, cancer type and stage (self-reported), prior and current treatments (including topical vaginal estrogen), relevant comorbidities and medications, pelvic floor therapy history, and sexual activity and partner status. Sensitive items are optional and include a "prefer not to answer" choice.

Participants will complete the sociodemographic questionnaire and baseline FSFI and EORTC QLQ-C30 at session 1 and repeat the FSFI and EORTC QLQ-C30 at session 3. Focus groups and semistructured interviews will be audio recorded and transcribed verbatim; participant observation field notes and anonymous diaries will be collected after each session. Optional one-to-one counseling is not recorded; clinicians may keep brief field notes. Trained staff will enter data into REDCap (Research Electronic Data Capture; Vanderbilt University).

Before each in-person session, participants receive a telephone call inviting them to attend the session. During the call, transportation needs are assessed; if necessary, transportation is arranged by the research team. Participants are informed that meals and refreshments are provided free of charge. If desired, participants may bring an accompanying adult family member who may attend the group sessions but will not participate in interviews or individual counseling to maintain privacy and confidentiality. To uphold participants' autonomy and freedom, all individuals are explicitly informed that voluntary withdrawal is allowed at any time without penalty. This ensures full participant control over involvement and minimizes the risk of perceived coercion.

Personal, religious, or moral values will not be imposed. No judgments about patients' sexual lives that could generate stigma, guilt, or shame are permitted. Inappropriate or nonconsensual language or materials are prohibited. Explicit or mocking sexual content is prohibited unless educationally justified and previously consented to by participants. Private information will not be disclosed without authorization. It will be strictly forbidden to deny or ignore the needs of lesbian, gay, bisexual, transgender, and queer survivors. Heteronormative assumptions are not permitted; all content must be inclusive and adapted to participants with nonbinary identities or same-sex partners. Unqualified or clinically unsupported interventions will not be allowed. All professionals must present credentials accrediting their expertise and complete training in cultural sensitivity and sexual ethics to prevent stigma and bias.

Sample Size

The quantitative pretest-posttest sample size was calculated based on mean differences and an adjusted SD of 0.69, informed by prior studies such as Starting the Conversation, a pilot trial using the FSFI to assess sexual function in survivors of gynecological cancer [11]. Assuming a moderate effect size (d=0.5), a 95% confidence level, and 80% power, the estimated sample size is 60 participants. Calculations were performed using G*Power (Heinrich Heine University Düsseldorf), a standard tool for sample size estimation in mean difference studies.

For the qualitative strand, sample adequacy is guided by the information power principle—considering study aim, sample specificity, quality of dialogue, and analytic strategy—rather than numeric saturation targets [24]. We purposefully seek heterogeneity in survivorship stage and symptom profiles to capture a range of experiences relevant to sexual health education.

XSL•FO
RenderX

Analysis Plan

Quantitative Analysis

Continuous variables will be summarized as means with SDs or medians with IQRs depending on the distribution, assessed using the Shapiro-Wilk test. Categorical variables will be presented as absolute and relative frequencies. Primary analyses will evaluate within-group changes (session 1 vs session 3) in FSFI and EORTC QLQ-C30 scores. For continuous outcomes, paired 2-tailed *t* tests will be used when normality assumptions are met; otherwise, Wilcoxon signed rank tests will be applied. A 2-sided significance level of α=.05 will be adopted. Effect sizes will be reported (Cohen *dz* for paired *t* tests; *r* for the Wilcoxon signed rank test), along with 95% CIs and graphical representations of response distributions. Missing data will be described; if they constitute >5% and are plausibly missing at random, an exploratory sensitivity analysis using multiple imputation will be performed.

Qualitative Analysis

Reflexive thematic analysis will be used following the framework by Braun and Clarke [25], prioritizing researcher reflexivity and active theme development over codebook reliability metrics. The analytic process will include (1) familiarization with transcripts and field notes, (2) inductive semantic and latent coding, (3) initial theme generation, (4) theme review against extracts and the full dataset, (5) theme definition and naming, and (6) construction of an analytic narrative linking themes to research questions and quantitative findings. Two researchers will independently code an initial subset to capture diverse interpretations, followed by reflexive dialogue. Subsequent coding will be led by the primary analyst with iterative peer debriefs focused on interpretive rigor rather than forced consensus. Reflexivity memos and an audit trail (decision logs and code or theme evolution) will be maintained. Credibility will be supported through data triangulation and thick description, dependability and confirmability will be supported through the audit trail, and transferability will be supported through detailed contextual reporting.

Reflexivity and Trustworthiness

The qualitative team (clinicians and educators experienced in sexual health and oncology) will maintain a reflexive stance through positionality statements, analytic memos, and documented decision trails. Credibility will be obtained via prolonged engagement across sessions, reflexive peer debriefs, and participant "member reflections" focused on emergent interpretations (not transcript verification). Dependability will be ensured based on an audit trail (versioned codebooks, memos, and decision logs) and clear documentation of analytic moves. Confirmability will be derived from reflexive journaling and an evidentiary chain linking raw data to themes and interpretations. Transferability will be dependent on "thick" descriptions of the setting, participants, and delivery to support readers' judgments of applicability to similar contexts.

Mixed Methods Integration

Integration will occur at the interpretation stage through meta-inferences and joint displays juxtaposing quantitative pretest-posttest changes (FSFI and EORTC QLQ-C30 scores) with qualitative themes (eg, sexual well-being, communication, agency, and body image). We will seek convergence, complementarity, and expansion; discrepancies will be explicitly examined to refine interpretations and inform practical recommendations.

Pilot Test

EDUSEXONCO was developed through a structured, user-centered process with three phases: (1) formative co-design, (2) content finalization and manualization, and (3) pretrial piloting and feasibility testing.

Formative Co-Design

We held participatory workshops with survivors (breast and gynecological cancer), clinicians (gynecological oncology, internal medicine, and psychiatry), and allied professionals (psychology, social work, pelvic floor physiotherapy, nutrition, yoga and relaxation, and sexual education). Workshops identified priority needs, culturally appropriate language, and acceptable delivery formats. Preliminary materials (slides, handouts, and 3D educational models) were iteratively refined using survivor feedback and brief "think-aloud" walk-throughs to optimize clarity, tone, and inclusiveness.

Patient and Public Involvement

Survivors of breast and gynecological cancers contributed to co-design (prioritizing topics, refining language, shaping levels, and informing 3D educational models) and will continue to provide iterative feedback on content and delivery during the study. They will also be invited to participate in dissemination and the codevelopment of implementation recommendations for scale-up.

Content Finalization and Manualization

A multidisciplinary panel reviewed all modules against international guidance (WHO; United Nations Educational, Scientific, and Cultural Organization; and WAS) and local practice to ensure accuracy, cultural relevance, and inclusivity. We developed a facilitator manual (learning objectives, session flow, time allocations, materials, and scripts for sensitive topics); fidelity tools (checklists and attendance and adherence trackers); and brief staff training on stigma-free, rights-based communication. Minor wording and sequencing changes enhanced accessibility for diverse literacy levels.

Pretrial Piloting and Feasibility Testing

Core components were piloted with a convenience sample of eligible survivors (n=15) to assess feasibility (recruitment and retention, session length, and logistics), acceptability (satisfaction, perceived usefulness, and comfort discussing sexual health), and fidelity (adherence to the manual). During pretrial piloting, full-day sessions were identified as a potential feasibility barrier; consequently, contingency adaptations (shorter session blocks, optional hybrid delivery, and flexible scheduling) were prepared while preserving the protocol's objectives and core content. In this study, brief participant and facilitator feedback will be reviewed after each session to inform minor, prespecified logistical refinements (eg, scheduling and session length and pace) for subsequent groups without modifying core content or learning objectives. These findings

XSL•FO
RenderX

guided the final structure of the sessions and the implementation plan to be used.

Ethical Considerations

The protocol was approved by the Research Ethics Committee of FVL (protocol 2024.248). All participants will provide written informed consent before enrollment. The informed consent form will outline study activities and the use of data for analyses aligned with the study objectives. Participation is voluntary, and patients may decline without any impact on the provision of usual clinical care. Participants will not receive financial compensation or additional services; however, the cost of transportation from their residence to FVL and a meal voucher will be covered.

Only trained team members will deliver content within their scope of practice. Any clinical, psychological, or emotional issues arising from the implementation of the EDUSEXONCO intervention are escalated to the study clinician and the psycho-oncology team, who are professionally qualified to manage such situations. All adverse events are documented and handled according to the institutional policy of FVL. If necessary, participants may be referred to the institution's emergency department, accompanied by the psycho-oncologist or a study representative, to initiate the required standard clinical care.

Ethical safeguards include respect for autonomy, confidentiality, and the right to withdraw at any time without affecting clinical care. Procedures are in place to minimize risk when addressing sensitive topics, including trauma-informed facilitation, private spaces for optional one-on-one counseling, and rapid referral pathways to psycho-oncology when indicated. All data will be deidentified and stored in secure, password-protected systems accessible only to authorized team members. Audio files and transcripts will be stored on secure institutional servers at the Clinical Research Center of FVL; identifiers will be removed during transcription, and a linkage file will be kept separately under restricted access. Paper materials (consent forms and checklists) will be stored in locked cabinets within restricted-access offices. Data retention will comply with institutional and national regulations. No photographs or video recordings of participants will be taken.

Dissemination and Implementation Plan

In line with open science and transparency principles, we plan to disseminate study materials and outputs through (1) open access educational resources (eg, facilitator guides and modules) where appropriate, (2) interdisciplinary training workshops for health care providers to support culturally competent sexual health counseling in survivorship care, (3) presentations at relevant national and international scientific meetings and patient support networks, and (4) partnerships with academic and community stakeholders to support integration into clinical education and survivorship pathways. These activities aim to facilitate future replication and support sustainable incorporation of sexual health education into oncology survivorship care.

This protocol describes the rationale, intervention components, and mixed methods feasibility and acceptability evaluation of EDUSEXONCO, a co-designed, context-adapted educational psychosexual program for survivors of breast and gynecological cancer in Colombia. This study will assess implementation outcomes—including recruitment, retention, session attendance, fidelity to the manualized content, and completeness of quantitative and qualitative data collection—alongside participants' acceptability and perceived usefulness of the intervention.

The findings will inform refinement of the EDUSEXONCO curriculum and delivery format (including prespecified contingency options) and will provide context-specific implementation knowledge to support integration of sexual health education into survivorship care pathways in low- and middle-income settings. Results from this feasibility study will also guide the design of a subsequent comparative, multisite study with longer follow-up to evaluate effectiveness and scalability.

Results

This study is funded by the Women's World Banking Colombia Foundation, awarded in December 2024, and received ethics approval from FVL. Recruitment began in March 2025 and is planned until September 2025. As of July 2025, a total of 13 participants have been enrolled (n=2, 15.4% with breast cancer and n=11, 84.6% with gynecological cancer). During feasibility activities, full-day sessions were identified as a potential barrier to attendance; accordingly, contingency adaptations (shorter session blocks and optional hybrid delivery) were suggested and will be considered as needed without altering protocol objectives or core content. Quantitative and qualitative data analysis is scheduled for October 2025; final results are expected by December 2025.

Discussion

Anticipated Findings and Contribution

This feasibility and acceptability study is designed to determine whether a co-designed, multidisciplinary educational psychosexual intervention (EDUSEXONCO) can be delivered as planned and is acceptable to survivors of gynecological and breast cancer in a Colombian university hospital setting. We anticipate generating actionable implementation evidence on recruitment and retention, session attendance, fidelity to the manualized content, and completeness of quantitative and qualitative data collection. In addition, we expect to identify participants' perceived usefulness of the sessions, barriers to participation, and opportunities to refine content and delivery to better address sexual health needs during survivorship care.

Comparison With Prior Work

Prior research has consistently documented that sexual health concerns are common among survivors of cancer and that health care professionals frequently report limited training, time constraints, and discomfort in addressing sexuality in routine oncology care [12,26]. Digital and blended interventions such as Fex-Can and Fex-Can 2.0 have shown that structured psychoeducational content, coupled with opportunities for reflection and support, can be acceptable and useful for survivors [13,14]. EDUSEXONCO is conceptually aligned with this body

XSL•FO
RenderX

of work by combining evidence-informed education with multidisciplinary counseling and a supportive group environment to facilitate communication, normalize sexual concerns, and address survivorship-related changes in body image and intimacy. This approach also aligns with frameworks from the WHO and the WAS, which position sexual health as an integral component of health and a human right within survivorship care [17,19]. In Latin American contexts where stigma and sociocultural norms may discourage disclosure, group-based formats may offer a particularly valuable space for peer learning and validation; this feasibility study will help clarify how such delivery functions in the Colombian setting.

Strengths and Limitations

Strengths of this protocol include the mixed methods design, the participant-centered co-design process, and the manualized intervention structure (facilitator guides, checklists, and fidelity tools) that supports reproducibility. The use of validated instruments (FSFI and EORTC QLQ-C30) enables comparability with prior survivorship research and provides a structured approach to capturing patient-reported outcomes. Qualitative methods will complement questionnaire data by exploring survivors' lived experiences and perceived relevance of the content and the contextual factors influencing participation.

As a single-center feasibility study using a single-group pretest-posttest design and short follow-up, this study is not intended to provide definitive causal inference or effectiveness estimates. Potential limitations include self-selection and social desirability bias, which may affect acceptability assessments and self-reported outcomes. Feasibility challenges such as scheduling constraints and sustained participation may arise; therefore, prespecified contingency options (eg, shorter session blocks or optional hybrid delivery) are included to support adherence while preserving core intervention components.

Future Directions

Building on this feasibility study, the next phase of EDUSEXONCO will be designed to evaluate effectiveness and implementation outcomes using a multicenter approach with larger samples and longer follow-up to assess durability and scalability. Hybrid delivery models (combining in-person and digital modules) may be explored to improve accessibility and continuity, particularly for participants in remote or resource-limited settings. Standardized facilitator training and fidelity monitoring will remain essential to ensure consistency across sites. Future studies should also incorporate economic evaluations to assess cost-effectiveness and sustainability. If feasible, the intervention framework could be adapted and tested in other clinical populations affected by sexual health concerns, contributing to a broader biopsychosocial rehabilitation model.

The findings will inform refinement of the EDUSEXONCO curriculum and delivery format (including prespecified contingency options) and will provide context-specific implementation knowledge to support integration of sexual health education into survivorship care pathways in low- and middle-income settings. Results from this feasibility study will also guide the design of a subsequent comparative, multisite study with longer follow-up to evaluate effectiveness and scalability.

Acknowledgments

The authors thank the health care professionals who are part of the EDUSEXONCO project, including the nutritionist, social worker, pelvic floor physiotherapist, psycho-oncologist, sexual wellness advisor, complementary practices consultant, and yoga instructor. Special acknowledgment goes to the patients who participated in the pilot test and placed their trust in this project.

Funding

This study is funded by the Fondo Fundación WWB Colombia para la investigación aimed at supporting the activities of the project as part of a fieldwork grant awarded for NSF's doctoral research. The funder was not involved in the study design; data collection, analysis, and interpretation; or writing of the manuscript. The funding does not cover the costs of publications derived from the project.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: NSF
Funding acquisition: NSF
Investigation: NSF
Methodology: NSF, MANJ, JSG-S
Project administration: NSF
Resources: NSF
Supervision: JDLR, AEB, JSG-S
Writing—original draft: NSF, MANJ
Writing—review and editing: JDLR, AEB, JSG-S
All authors reviewed and approved the final version of the manuscript.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Operational screening checklist.
[DOCX File, 19 KB - resprot_v15i1e80567_app1.docx]

Multimedia Appendix 2
Intervention schedule (per cohort).
[DOCX File, 20 KB - resprot_v15i1e80567_app2.docx]

Multimedia Appendix 3
Detailed topic-by-level syllabus (TIDieR [Template for Intervention Description and Replication] compatible).
[DOCX File, 23 KB - resprot_v15i1e80567_app3.docx]

Multimedia Appendix 4
Example of semistructured group interview guide (sexual health physician session).
[DOCX File, 22 KB - resprot_v15i1e80567_app4.docx]

References

1. Sousa Rodrigues Guedes T, Barbosa Otoni Gonçalves Guedes M, de Castro Santana R, et al. Sexual dysfunction in women with cancer: a systematic review of longitudinal studies. Int J Environ Res Public Health 2022 Sep 21;19(19):11921. [doi: 10.3390/ijerph191911921] [Medline: 36231221]
2. Firmeza MA, Vasconcelos CTM, Vasconcelos Neto JA, et al. The effects of hysterectomy on urinary and sexual functions of women with cervical cancer: a systematic review. Rev Bras Ginecol Obstet 2022 Aug;44(8):790-796. [doi: 10.1055/s-0042-1748972] [Medline: 36075225]
3. Jensen PT, Groenvold M, Klee MC, Thranov I, Petersen MA, Machin D. Longitudinal study of sexual function and vaginal changes after radiotherapy for cervical cancer. Int J Radiat Oncol Biol Phys 2003 Jul 15;56(4):937-949. [doi: 10.1016/s0360-3016(03)00362-6] [Medline: 12829128]
4. Rodríguez-Burbano AY, Galván-Canchila DM, de Diego-Cordero R. Knowledge, attitudes and practices towards sexual and reproductive health and rights of girls among Colombian healthcare professionals. Int J Environ Res Public Health 2022 Sep 28;19(19):12295. [doi: 10.3390/ijerph191912295] [Medline: 36231597]
5. Ruiz-Vallejo F, Bruhl K, Guerrero G. Percepciones de proveedores de servicios de salud sobre la salud sexual de los adultos mayores en Colombia. Revista Salud Uis 2023;55(1). [doi: 10.18273/saluduis.55.e:23030]
6. Bermúdez PC, Arrivillaga M, Torres Poveda K, Castrillón Libreros DM, Castillo Castillo LE, Neira Acevedo D. Barriers to adherence to cytology exam: a case study in low-income Colombian women. BMC Health Serv Res 2023 Jul 25;23(1):796. [doi: 10.1186/s12913-023-09700-4] [Medline: 37491282]
7. Ochoa E, Carrillo GM, Sanabria D. Finding myself as a cervical cancer survivor: a qualitative study. Eur J Oncol Nurs 2019 Aug;41:143-148. [doi: 10.1016/j.ejon.2019.06.008] [Medline: 31358247]
8. Shams M, Coman C, Fatone F, et al. The impact of gynecologic cancer on female sexuality in Europe and MENA (Middle East and North Africa): a literature review. Sex Med Rev 2024 Sep 25;12(4):587-599. [doi: 10.1093/sxmrev/qeae039] [Medline: 38850565]
9. Wenzel L, Osann K, Hsieh S, Tucker JA, Monk BJ, Nelson EL. Psychosocial telephone counseling for survivors of cervical cancer: results of a randomized biobehavioral trial. J Clin Oncol 2015 Apr 1;33(10):1171-1179. [doi: 10.1200/JCO.2014.57.4079] [Medline: 25713429]
10. Bakker RM, Mens JWM, de Groot HE, et al. A nurse-led sexual rehabilitation intervention after radiotherapy for gynecological cancer. Support Care Cancer 2017 Mar;25(3):729-737. [doi: 10.1007/s00520-016-3453-2] [Medline: 27787681]
11. Reese JB, Bober SL, Sorice KA, et al. Starting the conversation: randomized pilot trial of an intervention to promote effective clinical communication about sexual health for gynecologic cancer survivors. J Cancer Surviv 2024 Jun;18(3):800-809. [doi: 10.1007/s11764-022-01327-4] [Medline: 36604391]
12. Hay CM, Donovan HS, Hartnett EG, et al. Sexual health as part of gynecologic cancer care: what do patients want? Int J Gynecol Cancer 2018 Nov;28(9):1737-1742. [doi: 10.1097/IGC.0000000000001376] [Medline: 30358703]
13. Lampic C, Ljungman L, Micaux Obol C, Eriksson LE, Wettergren L. A web-based psycho-educational intervention (Fex-Can) targeting sexual dysfunction and fertility-related distress in young adults with cancer: study protocol of a randomized controlled trial. BMC Cancer 2019 Apr 11;19(1):344. [doi: 10.1186/s12885-019-5518-3] [Medline: 30975116]
14. Skog R, Olsson EMG, Gorman JR, Bober SL, Lampic C, Wettergren L. An internet-delivered psychoeducational intervention (Fex-Can 2.0) targeting fertility-related distress and sexual dysfunction in young adults diagnosed with cancer: study

XSL•FO
RenderX

protocol of a randomized controlled trial with an internal pilot phase. PLoS ONE 2025;20(4):e0322368. [doi: 10.1371/journal.pone.0322368] [Medline: 40300010]
15. Agrawal LS, O'Riordan L, Natale C, Jenkins LC. Enhancing sexual health for cancer survivors. Am Soc Clin Oncol Educ Book 2025 Jun;45(3):e472856. [doi: 10.1200/EDBK-25-472856] [Medline: 40173399]
16. Shalamzari KH, Asgharipour N, Jamali J, Kermani AT, Babazadeh R. Comparing the effect of sexual counseling based on BETTER and PLISSIT models on sexual assertiveness in women with breast cancer after mastectomy. J Educ Health Promot 2024 Feb 26;13(1):55. [doi: 10.4103/jehp.jehp_1384_22] [Medline: 38549654]
17. Sexual health and its linkages to reproductive health: an operational approach. World Health Organization. 2017 Sep 20. URL: https://www.who.int/publications/i/item/978924151288 [accessed 2025-07-11]
18. International Technical Guidance on Sexuality Education: An Evidence-Informed Approach: UNESCO; 2018. URL: https://unesdoc.unesco.org/ark:/48223/pf0000260770 [accessed 2026-02-24] [doi: 10.54675/UQRM6395]
19. Declaration of sexual rights. World Association for Sexual Health. 2014. URL: https://www.worldsexualhealth.net/_files/ugd/793f03_779f193815894895b9d1480bde676c37.pdf?index=true [accessed 2025-11-16]
20. Okpo E, Othieno R, Eleje GU, Oragwu CI, Eke AC. Interventions to improve psychosexual function in women treated for gynaecological cancers. Cochrane Database Syst Rev 2022 Aug 22;2022(8):CD014674. [doi: 10.1002/14651858.CD014674]
21. Inicio - Fundación Valle del Lili. Fundación Valle del Lili. URL: https://valledellili.org/ [accessed 2026-02-15]
22. Sánchez-Sánchez B, Navarro-Brazález B, Arranz-Martín B, Sánchez-Méndez Ó, de la Rosa-Díaz I, Torres-Lacomba M. The Female Sexual Function Index: transculturally adaptation and psychometric validation in Spanish women. Int J Environ Res Public Health 2020 Feb 5;17(3):994. [doi: 10.3390/ijerph17030994] [Medline: 32033334]
23. Sánchez Pedraza R, Valderamar A. Validación para utilización en Colombia de la escala EORTC QLQ C-30 para evaluación de la calidad de vida de los pacientes con cáncer. Rev Colomb Cancerol 2020;24(2):61-71. [doi: 10.35509/01239015.128]
24. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
25. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2022 Oct 25;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
26. Sopfe J, Pettigrew J, Afghahi A, Appiah LC, Coons HL. Interventions to improve sexual health in women living with and surviving cancer: review and recommendations. Cancers (Basel) 2021 Jun 24;13(13):3153. [doi: 10.3390/cancers13133153] [Medline: 34202477]

Abbreviations

BETTER: bring up, explain, tell, time, education, record
EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
FSFI: Female Sexual Function Index
FVL: Fundación Valle del Lili
PLISSIT: permission, limited information, specific suggestion, intensive therapy
REDCap: Research Electronic Data Capture
WAS: World Association for Sexual Health
WHO: World Health Organization

Edited by A Schwartz; submitted 12.Jul.2025; peer-reviewed by A Bakare; accepted 07.Jan.2026; published 27.Feb.2026.

Please cite as:
Salazar Falla N, Navas Jojoa MDLÁ, Lalinde Riviño JD, Escandón Bucheli A, Galindo-Sánchez JS
Educational Intervention to Improve Sexual Health and Quality of Life in Survivors of Breast and Gynecological Cancer: Protocol for a Mixed Methods Feasibility Study
JMIR Res Protoc 2026;15:e80567
URL: https://www.researchprotocols.org/2026/1/e80567
doi:10.2196/80567

© Nathalia Salazar Falla, María de los Ángeles Nava Jojoa, Juan David Lalinde Riviño, Auramaría Escandón Bucheli, Juan Sebastián Galindo-Sánchez. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information,

XSL•FO
RenderX

a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

GORE VIABAHN VBX Balloon-Expandable Endoprostheses as a Bridging Stent With Branched and Fenestrated Endografts in the Endovascular Treatment of Aortic Aneurysms: Protocol for a Retrospective Multicenter Registry (EMBRACE Study)

Pietro Dioni[1*], MD; Anders Wanhainen[2], MD; Jacob Budtz-Lilly[3], MD; Nuno Dias[4], MD; Tab Bonny[5], MD; Luca Bertoglio[1], MD; EMBRACE study investigator[5,6*]

[1]Division of Vascular Surgery, Department of Clinical and Experimental Sciences (DSCS), University and ASST Spedali Civili Hospital of Brescia, Piazzale Spedali Civili 1, Brescia, Italy

[2]Department of Surgical Sciences, Vascular Surgery, Uppsala University, Uppsala, Sweden

[3]Division of Vascular Surgery, Department of Cardiothoracic and Vascular Surgery, Aarhus University Hospital, Aarhus, Denmark

[4]Vascular Center Malmö, Skåne University Hospital, Malmö, Sweden

[5]W. L. Gore & Associates (United States), Flagstaff, AZ, United States

6

[*]these authors contributed equally

Corresponding Author:
Luca Bertoglio, MD
Division of Vascular Surgery, Department of Clinical and Experimental Sciences (DSCS), University and ASST Spedali Civili Hospital of Brescia, Piazzale Spedali Civili 1, Brescia, Italy

Abstract

Background: Covered stents commercially available are frequently used off-label in conjunction with fenestrated and branched aortic stent grafts, but there is a lack of dedicated devices.

Objective: This study aims to assess the safety and mid-term clinical performance of a new dedicated covered stent, the GORE VIABAHN VBX Balloon Expandable Endoprosthesis (VBX stent graft), when used as a bridging stent with branched and fenestrated aortic endografts in treating complex abdominal aortic and thoraco-abdominal aneurysms.

Methods: A retrospective, multicenter, single-arm study in the European Union (ClinicalTrials.gov NCT05143138) enrolled patients treated with the VBX stent graft as a bridging stent in branched endovascular repair (BEVAR) and fenestrated endovascular repair (FEVAR) to allow endovascular aneurysm repair between January 2017 and December 2021. Up to 15 sites in Europe were required to enroll a minimum of 220 patients. Patients' medical records were reviewed by the investigator, and specific data were collected ambispectively for up to 5 years of follow-up from the index procedure. The primary end point is target vessel patency (patient level) through 12 months. The registry was designed to statistically test target vessel patency at 12 months in both FEVAR and BEVAR populations. The hypothesis will be tested separately in the 2 cohorts (fenestrated or branched endovascular repair), using patients with core laboratory imaging results available annually through 5 years. The binomial exact test will be used with a 1-sided 2.5% level of significance to test the null hypothesis.

Results: In total, 259 patients were retrospectively enrolled for a prospective follow-up of 5 years: 136 patients (n=99, 72.8% male; mean age 73, SD 8.9 y) in the BEVAR cohort, 92 patients (n=80, 87.0% male, mean age 72.7, SD 8.1 y) in the FEVAR cohort, and 31 patients (n=17, 54.8% male, mean age 70.9, SD 9.4 y) in the mixed fenestrated or branched endovascular repair cohort. Overall, 662 target vessels were stented with the investigational covered stents: 163 (24.6%) celiac trunk, 192 (29.0%) superior mesenteric artery, and 307 (46.4%) renal arteries. The VBX stent grafts were paired with branches in 451 (68.1%) cases or fenestrations in 211 (31.9%) cases among all Cook Medical stent graft cases. The 1-year results will be published in the fourth quarter of 2025, and the 5-year follow-up results will be analyzed by mid-year 2028.

Conclusions: This study will investigate the VBX stent graft performance in combination with fenestrated and branched aortic grafts to corroborate its use in complex aortic endovascular procedures and support the modification of current device instructions for use.

Trial Registration: ClinicalTrials.gov NCT05143138; https://clinicaltrials.gov/study/NCT05143138

International Registered Report Identifier (IRRID): DERR1-10.2196/78970

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e78970) doi:10.2196/78970

KEYWORDS

aorta; fenestrated; branched; covered; stent

Introduction

The renal-mesenteric aortic vessels, which include the celiac trunk, superior and inferior mesenteric arteries, and renal arteries, are variably involved in thoraco-abdominal and complex abdominal aortic aneurysms [1]. Extensive aortic pathologies are currently managed mainly by using fenestrated or branched endovascular repair (F/B-EVAR) with high rates of technical success. It is widely recognized as the technique of choice, offering acceptable mid-term results in terms of mortality and morbidity guidelines [2,3] and supported by international societies [4,5]. In order to both preserve perfusion of visceral vessels and successfully exclude the aortic aneurysm, the fenestrations or branches of the main body of the aortic stent graft are bridged to visceral aortic vessels with covered stents (CSs). However, more than 2 decades after the introduction of the first aortic endoprosthesis, the lack of standardization and limited availability of dedicated ancillary materials for F/BEVAR have left this technique highly dependent on individual physician preferences and choices and the evolution of these ancillary devices [6]. Therefore, different CSs have been used during F/B-EVAR, but none of them were specifically approved for use as bridging stents in combination with the different aortic stent grafts available on the market at the time this study was designed [6]. The performance of different CSs has been investigated in single-center or retrospective registries, and only 1 European CS manufacturer, Bentley (Bentley InnoMed), obtained approval for BeGraft Peripheral in combination with fenestrated endovascular repair (FEVAR; NCT03987035) and is testing BeGraft Peripheral PLUS, in combination with branched endovascular repair (BEVAR; NCT03982940) and Cook Medical aortic stent grafts to modify their instructions for use.

The GORE-sponsored EMBRACE study is being conducted to confirm the clinical performance and safety of the GORE VIABAHN VBX Balloon Expandable Endoprosthesis (VBX stent graft), specifically when used as a bridging stent in combination with F/B-EVAR aortic stent grafts to support and corroborate the change of instruction for use and therefore an on-label use.

Methods

Generalities

The EMBRACE study (ClinicalTrials.gov NCT05143138) is a retrospective, multicenter, single-arm registry in the European Union that retrospectively enrolled patients treated with the VBX stent graft as a bridging stent in BEVAR and FEVAR between January 2017 and December 2021, with prospective data collection and follow-up. Patients' medical records were reviewed by the investigator, and specific data were collected retrospectively and prospectively for up to 5 years of follow-up from the index procedure. Considering consistent device design, documented procedures, available clinical outcomes, and operational approaches, a retrospective and prospective registry was deemed an appropriate observational study to address the research question. Retrospective enrollment was feasible in this case because pre- and postimplant computed tomography scans were routinely performed as part of standard clinical practice. Central imaging analysis of multicenter results was intended to support regulatory approval. To treat complex and thoraco-abdominal aortic aneurysms involving the renal-mesenteric arteries, up to 15 sites (Table 1) in Europe (Denmark, Italy, the Netherlands, Norway, Spain, and Sweden) must enroll a minimum of 220 patients who have received treatment with the VBX stent graft as a bridging stent in F/B-EVAR. Each site cannot contribute more than 10% of the total patient number for each cohort without sponsor approval, with a maximum of 30% of total enrollment for a single site. This observational study was reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) statement guidelines.

XSL•FO
RenderX

Table . Information on the 15 European sites involved in the study.

Institution	Principal investigator	City or country	Enrollment, n (%)
Amsterdam UMC, Vascular Surgery	Arjan Hoksbergen	Amsterdam	32 (12.4)
Hospital Clinico San Carlos	Isaac Martínez	Madrid	27 (10.4)
Policlinico Umberto I	Wassim Mansour	Roma	23 (8.9)
Rigshospitalet, Dep. Vascular Surgery	Timothy Resch	Copenhagen	23 (8.9)
Skane University Hospital, SUS, Malmo	Nuno Dias	Malmo	22 (8.5)
Aarhus University, Department of Thoracic and Cardiovascular Surgery	Jacob Budtz-Lilly	Aarhus	21 (8.1)
Azienda Ospedaliera Complesso Ospedaliero San Giovanni—Addolorata	Rocco Giudice	Roma	21 (8.1)
IRCCS Ospedale San Raffaele, Chirurgia Vascolare	Enrico Rinaldi	Milano	20 (7.7)
IRCCS Ospedale Policlinico San Martino	Giovanni Pratesi	Genova	19 (7.3)
UNN Tromsø	Kare Nordhus	Tromsø	19 (7.3)
Uppsala University Hospital	Anders Wanhainen	Uppsala	16 (6.2)
Haukeland Universitetssjukehus	Antje Butter	Bergen	10 (3.9)
Hospital Clínico Universitario San Cecilio	Luis Miguel Salmerón	Granada	5 (1.9)
St. Olavs Hospital	Frode Manstad Hulaas	Trondheim	1 (0.4)

Ethical Considerations

Due to the retrospective nature of this observational study, review by the institutional review board or ethics committee at each of the study sites was ad hoc required in accordance with local regulations and institutional policies regarding minimal-risk observational studies; however, the clinical protocol and all amendments were submitted to the ethics committee for each clinical site [7]. An informed consent form signed by the patient or next of kin or legal representative (for deceased patients at registry entry, unless a waiver was granted), according to local regulations, was required for inclusion in the study. For contacting the potential patient, the investigator must use reasonable methods, including confirming the patient's survival, looking through medical records, calling readily available numbers, or acquiring contact details from demographic and/or health care registers. All data were deidentified prior to analysis to ensure privacy and confidentiality. The study will be conducted according to the Standards of Good Clinical Practice, the Declaration of Helsinki guidelines, and in compliance with the ISO 14155 standard. No compensation was provided to participants. No identifiable images of participants are included in this study. The study (design illustrated in Figure 1) is registered with ClinicalTrials.gov (NCT05143138).

XSL•FO
RenderX

Figure 1. Study design schema. BEVAR: branched endovascular repair; EDC: electronic data capture; FEVAR: fenestrated endovascular repair.

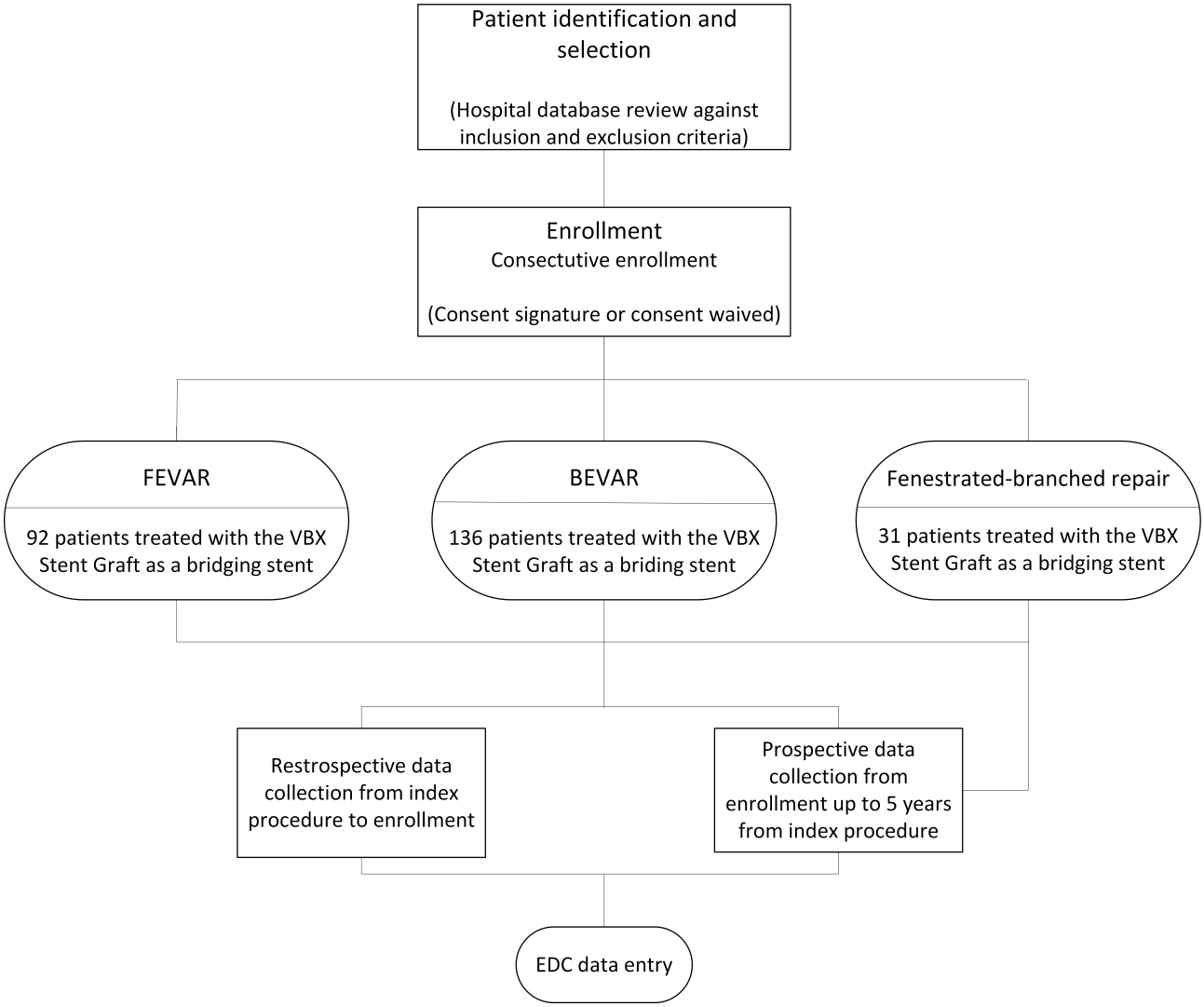

Study Population

The inclusion criteria (Textbox 1) reflect the potential broad application of the device as evaluated by the medical judgment of the implanting physician. The inclusion criteria will be assessed during the initial identification, which is part of the retrospective device-use search that considers the patient's health status at the time of the index procedure. Patients were treated with the VBX stent graft as a bridging stent in conjunction with an F/B-EVAR from December 31, 2021, until January 1, 2017.

XSL•FO
RenderX

Textbox 1. Inclusion or exclusion criteria.

Inclusion criteria

- Patients treated with the VBX stent graft as a bridging stent in conjunction with a branched or fenestrated stent graft to facilitate endovascular aneurysm repair from December 31, 2021, until January 1, 2017
- Age 18 years or older at the time of implant
- An informed consent form signed by participant or next of kin or legal representative (for deceased patients at registry entry, unless a waiver was granted), according to local regulations

Exclusion criteria

- Patient treated for ruptured aneurysm or who were otherwise hemodynamically unstable at the time of the procedure
- Patient treated for acute or subacute dissection, less than 90 days from onset of symptoms
- Patient treated using physician-modified endovascular grafts
- Patient intended to be treated with chimney, periscope, octopus, and sandwich technique per the pre-treatment case plan
- At the time of treatment, patient had known coagulation disorders, including hypercoagulability, that were not amenable to treatment
- Patient was pregnant at the time of treatment
- Participation in another drug or device investigational study within 1 year of device implant, which can confound the registry end points
- Patient had known or suspected systemic infection (including treatment for mycotic aneurysm) at the time of implant

Study Devices

The GORE VIABAHN VBX Balloon Expandable Endoprosthesis (hereafter, VBX stent graft; W. L. Gore & Associates, Inc) is made of a balloon-expandable stent graft mounted on a balloon catheter delivery system. The VBX stent graft is made of independent stainless-steel rings that are fully enveloped by expanded polytetrafluoroethylene and fluorinated ethylene propylene film (Figure 2). The delivery system has two radiopaque balloon markers embedded in the shaft to allow accurate delivery of the balloon-expandable stent graft. The delivery system, compatible with 0.035″ (0.89 mm) guide wires, can be used for initial stent placement and subsequent poststent dilatation. The VBX stent graft is available in lengths ranging from 15 to 79 mm and in a variety of diameters, ranging from 5 to 11 mm. The delivery system is also offered in two shaft lengths, 80 and 135 cm. This product is supplied sterile.

Figure 2. VBX stent graft picture with detailed diameter and length ranges.



Nominal diameter (mm)	Introducer sheath size (Fr)	Max postdilated stent diameter (mm)	Crimped stent length (mm)
5	7	8	15, 19, 29, 39, 59, 79
6	7	8	15, 19, 29, 39, 59, 79
7	7	11	15, 19, 29, 39, 59, 79
8	7	11	29, 39, 59
	8		79
8L	7	16	29, 39
	8		59, 79
9	8	13	29, 39, 59, 79
10	8	13	29, 39, 59, 79
11	8	16	29, 39, 59, 79

Study Procedure

The sites were expected to perform a search of patients treated with the VBX stent graft as a bridging stent in F/B-EVAR within the hospital database. This search was based on the information already present in the patient's medical record. Screening was conducted in reverse order, starting with patients treated from December 31, 2021, back to January 1, 2017, or their registry device start date. The research results were recorded in the prescreening list by patients' index procedure date, arranged from the newest to the oldest patient treated. All patients included in this registry were retrospectively enrolled from participating centers. No prospective enrollment was performed.

Bias

The registry is meant to retrospectively enroll and prospectively collect data for assessing "real world" clinical experience and patient outcomes during treatment and throughout follow-up up to 5 years. Physicians determine patient selection, diagnostic imaging, and treatment interventions based on this protocol and clinical practice standards. Bias was avoided through strict adherence to the registry protocol, with site compliance monitored, including patient eligibility criteria, per ethics committee regulations. To minimize selection bias, it was mandatory that the enrollment of patients be consecutive in nature based on the date of the index procedure. It is well known that excluding treated patients who are deceased or not reachable

XSL•FO
RenderX

at the time of inclusion due to lack of informed consent can generate an important selection bias in a retrospective study and is likely to prevent the achievement of the research objectives. Therefore, for patients who were deceased or not reachable at the time of inclusion, the consent could be waived or signed by next of kin or legal representative, according to local regulations. Additionally, all patients present in the prescreening list who underwent inclusion or exclusion criteria assessment were recorded in the electronic data capture, regardless of enrollment, until the site reached the maximum cap of enrolled patients planned or until the cohort was completed. Potential bias related to the study sponsor's role in design and oversight will be addressed by strict adherence to the registry protocol. Patient selection, diagnostic imaging, and treatment interventions will be determined by the treating physician based on this protocol and clinical practice standards, independent of the sponsor.

End Points

The primary and secondary end points are defined in Table 2. Those end points were defined according to the Society of Vascular Surgery reporting standard created for fenestrated and branched endovascular repair [1]. The primary end point is 12-month target vessel patency at the patient's level. This is defined as uninterrupted patency with no occlusion or procedure performed to maintain patency on the VBX stent graft or native target vessel. Interventions intended to treat endoleak or stent disconnection do not count as loss of primary patency. Patient-level primary patency will be considered lost if any VBX-treated artery loses patency, regardless of the patency status of other treated vessels. End point-related clinical events will be reviewed and adjudicated by the independent clinical events committee (CEC). Secondary and additional end points include all-cause reintervention, VBX-treated artery reinterventions, target vessel technical success, primary technical success, target vessel instability, target vessel patency by both patient-level and vessel-level analysis, aneurysm-related mortality, and major adverse events at 30 days. All study end points will be evaluated annually for up to 5 years following the index procedure. End point analyses will include only arteries treated with a VBX stent graft. Arteries treated during the procedure without a VBX stent graft will be excluded from the analysis.

XSL•FO
RenderX

Table . Study end point definitions.

Outcome	Definition of SVS[a] reporting standards
Technical success–related outcomes	
Target vessel technical success	• Successful catheterization and VBX stent graft placement in all intended target vessels. For this registry, technical success will be reported for the overall procedure and specifically for the vessels targeted for treatment with the VBX stent graft
Primary technical success (total endovascular procedure)	A modified technical success definition, requiring the following: • Successful side branch catheterization and placement of bridging stents with restoration and maintenance of flow in all intended target vessels • Patency of all aortic modular stent graft components and intended side branch components • Absence of type I or type III endoleaks at completion angiography that extends beyond 30 days by confirmatory imaging (computed tomography angiography, magnetic resonance angiography, or duplex ultrasound)
Vessel patency–related outcomes	
Primary patency (primary end point)	• Uninterrupted patency with no occlusion or procedure performed to maintain patency on the VBX stent graft or native target vessel. Interventions intended to treat endoleak or stent disconnection do not count as loss of primary patency
Primary assisted patency	• Endovascular intervention performed to maintain patency in the presence of a stenosis before occlusion
Secondary patency	• Endovascular restoration of patency after occlusion of the side branch, stent, or stent graft has already occurred. Conversion to bypass or inability to treat by endovascular means defines loss of secondary patency
Occlusion	• Objective documentation by angiography, computed tomography, or ultrasound of complete stent occlusion with or without minimal flow into a targeted vessel
Stenosis	• Objective documentation by angiography, computed tomography, or ultrasound of stenosis intrastent or into a targeted vessel
Kink	• Objective documentation by angiography, computed tomography, or ultrasound of kink in the stented or native segment of a targeted vessel
Other outcomes	
Intraprocedural complications	• Any vessel perforation, dissection, or occlusion during target vessel stenting
Reintervention	• Any repeated vascular or nonvascular procedure related to the index procedure • Reintervention will be adjudicated as major or minor based on the following: Major: deployment of proximal or distal aortic or iliac extensions, removal of the device, use of thrombectomy or thrombolysis, and any major open surgical procedure. Minor: endovascular procedures (percutaneous transluminal angioplasty, atherectomy, stenting) without thrombectomy or thrombolysis, interventions to treat branch vessel stenosis, interventions to treat type II endoleak or branch-related endoleaks, and minor surgical revisions (patch angioplasty) of the access vessel. • Each reintervention will be adjudicated as related to a VBX-20–treated branch vessel component, related to a non–VBX–treated branch vessel component, or related to main body components, if possible

XSL•FO
RenderX

Outcome	Definition of SVS[a] reporting standards
Target vessel instability	• Death or rupture related to side branch complication (eg, endoleak) or reintervention to treat a branch-related complication, including endoleak, disconnection, kink, stenosis, occlusion, or rupture
Aneurysm-related mortality	• Any death that occurs within the first 30 days or any death that results from aneurysm rupture, aorta-related complications (eg, infection, occlusion, dissection, hematoma), or a complication of a secondary intervention
Major adverse events	• All-cause mortality • Myocardial infarction: Myocardial infarction resulting in severe hemodynamic dysfunction necessitating resuscitation, cardiac arrest, or fatal outcome • Respiratory failure requiring prolonged (less than 24 hours from anticipated) mechanical ventilation or reintubation from anticipated) mechanical ventilation or reintubation • Any renal function deterioration according to the RIFLE[b] classification system • Bowel ischemia requiring surgical resection or not resolving with medical therapy • Permanent paraplegia (any grade 3 A-C spinal cord injury) in a patient who is nonambulatory • Any major stroke defined according to National Institutes of Health Stroke Scale or equivalent

[a]SVG: Society of Vascular Surgery.

[b]RIFLE: risk, injury, failure, loss of kidney function, and end-stage kidney disease.

Statistical Method

No randomization or blinding is to be performed due to the confirmatory nature of the study. The registry is designed to statistically test the hypothesis that primary patency of target vessels (primary end point) in F/B-EVAR at 12 months is greater than 77% [8]. The 12-month imaging results will be analyzed, and the hypothesis will be tested separately in the 2 independently powered cohorts (F/B-EVAR). The binomial exact test will be used with a 1-sided 2.5% level of significance to test the null hypothesis (α).

The statistical hypothesis is specified as follows:

- H0: $P< 77\%$
- HA: $P>77\%$
- $\alpha=.025$ (1-sided)

However, *P* is the probability of maintaining primary patency at 12 months. The intent-to-treat population will be used to perform this test. A separate analysis of the subgroup of patients treated with mixed endografts (fenestrations and directional branches) as well as other clinical uses not described as only F/B-EVAR will be assessed with no formal hypothesis.

Sample Size Considerations

This study is designed with the assumptions used for the calculation of sample size, including the anticipated performance of the VBX stent graft in F/B-EVAR, the acceptance criteria determined from the literature, and the derived performance goal. The preliminary clinical evaluation suggested that the acceptable patency of the VBX stent graft used in target vessels at 12 months would be 87%. Given the complexity of the procedure, a noninferiority margin of 10% is considered appropriate. The literature determined that the anticipated 12-month primary patency is assumed to be 89% and 88% in the FEVAR and BEVAR cohorts, respectively [7]. Considering the acceptance criteria of 87% with a noninferiority margin of 10%, the resulting performance goal of 77% for both cohorts, and the expected patency of 89% and 88%, the sample size is determined to be 82 and 98 patients in the F/B-EVAR cohorts, respectively. The calculated sample size in the 2 cohorts will provide 80% power (1-beta) to test the primary hypothesis using a binomial exact test with a 1-sided 2.5% level of significance (alpha). The derived sample size is further inflated by about 18% for the lost-to-follow and missed imaging or visit at 12 months, thereby resulting in 100 and 120 patients to be enrolled in the F/B-EVAR cohorts, respectively.

Trial Organization and Oversight

An independent CEC will review end point–related clinical events to ensure the consistency of end points reported by the site. The EMBRACE CEC features an interdisciplinary team of 3 members with pertinent expertise who are not directly involved in the conduct of the registry. Sponsors will not participate in any CEC activity and may only tutor the CEC members on the protocol and registry device initially. The core laboratory used is Gore Imaging Sciences. Gore Imaging Sciences will assess migrations, patency, endoleak, and anatomical characterization using objective processes and standard procedures to ensure accuracy, consistency, and objectivity in the results by trained and qualified personnel. Medical imagery will be transmitted to the core laboratory using secure, web portal functionality and will be stripped of personal health information and replaced with numerical patient identifiers in an automated, systemic fashion. Once completed,

XSL•FO
RenderX

measurements will be reported directly in the electronic data capture system and considered source data. Study definitions and measurement criteria were standardized prior to image transfer to the core laboratory. The retrospective phase consisted entirely of collecting data from source documents available at each site, covering the period from the time of treatment until the enrollment date, while in the prospective phase, follow-up data were collected prospectively. Diagnostic imaging was performed at the discretion of the physician according to clinical practice standards, and all imaging analyses were conducted exclusively by the core laboratory. Local sites did not participate in the interpretation or adjudication of imaging findings.

Results

Site enrollment is complete (Table 1), and the data are currently being reviewed by the core laboratory and the CEC. The end point data are anticipated to be available in 2025.

In total, 259 patients were retrospectively enrolled for a prospective follow-up of 5 years. In the BEVAR cohort, 136 patients were enrolled, 99 (72.8%) were male, and the mean age was 73 (SD 8.9) years. In the FEVAR cohort, 92 patients were enrolled, 80 (87.0%) were male, and the mean age was 72.7 (SD 8.1) years. The remaining 31 had both BEVAR and FEVAR procedures (n=17, 54.8% male, mean age 70.9, SD 9.4 y). Of the total patient population, 202 (79.6%) have a history of smoking, 210 (81.0%) present with hypertension, 132 (56.2%) with hypercholesterolemia, and 85 (33.1%) with chronic obstructive pulmonary disease. A total of 662 vessels were treated with a VBX stent graft, including 451 branched and 211 fenestrated vessels. In the BEVAR cohort, 130 (28.8%) of the vessels treated were the superior mesenteric artery, while the remaining vessels were nearly evenly distributed among the celiac trunk (n=116, 25.7%) and left (n=99, 22.0%) and right (n=106, 23.5%) renal arteries. For the FEVAR cohort, the distribution across those vessels was 62 (29.4%), 47 (22.2%), 53 (25.1%), and 49 (23.3%) for the superior mesenteric artery, celiac trunk, and left and right renal arteries, respectively. Of the 14 sites, total enrollment rates ranged from 1 (0.4%) to 32 (12.4%). The range of the enrollment rates for the 3 cohorts was 0 (0%) to 21 (15.4%) for BEVAR, 0 (0%) to 15 (16.3%) for FEVAR, and 0 (0%) to 7 (22.6%) for mixed F/B-EVAR. The 1-year results will be published in the fourth quarter of 2025, and the 5-year follow-up results will be analyzed by mid-year 2028. In the meantime, an intermediate analysis at 3 years will be available in the first and second quarters of 2026, allowing subset analysis of specific populations (ie, connective tissue disorders) or special anatomical analyses (ie, inner vs outer branches)

Discussion

Principal Findings

The VBX stent graft is a third-generation stent graft specifically designed for F/BEVAR procedures, and the EMBRACE study is the first controlled study specifically designed to evaluate the outcomes of a covered stent with homogeneous dedicated reporting standards. Not only does it investigate the specific 5-year performance of the VBX stent graft, but also the large real-world patient cohort will provide a dataset available for subset analysis targeted on, for example, the type of vessel, the underlying pathology, and postoperative medications.

The ideal bridging covered stent for use with FEVAR and BEVAR must exhibit a sophisticated equilibrium between high longitudinal flexibility and robust radial strength. This ensures that the device can track through tortuous anatomy and conform to angulated visceral vessels without kinking, while simultaneously resisting compressive forces from the aortic graft or vessel wall. Crucially, the stent design should feature a low-profile delivery system for precise maneuverability and possess predictable flareability at the proximal end, facilitating a secure, seal-tight mating with the fenestration or directional branch to mitigate the risk of type III endoleaks. Finally, the integration of highly radiopaque markers and a biocompatible, thrombosis-resistant encapsulation (typically ePTFE or similar polymers) is essential to ensure accurate deployment and long-term patency. To date, covered stent design and choice have been deeply influenced by the aortic stent graft design. While sealing within reinforced fenestrations is obtained by flaring the proximal edge of the bridging stent, sealing between a branch and its covered stent is based on their overlap. For this reason, stent choice was traditionally guided by the design of the endograft, with balloon-expandable stents (BESGs) preferred for reinforced fenestrations and either self-expanding stents (SESGs) or BESGs for directional branches. Several bridging stents with similar structural features have been used in combination with fenestrated or branched endografts, yielding debated results and pitfalls due to, for example, insufficient radial force to maintain sealing after flaring [6]. The apposition of the bridging stent to the graft at the level of the fenestration is the most frequent site of type III endoleaks in FEVAR [8]. During follow-up, flared stents may become misaligned or rupture at the fenestration, making the selection of the bridging stent and procedural planning critical for long-term freedom from reintervention. To address this issue, several authors have assessed the durability of different bridging stents by simulating aortic physiology in bench tests. These studies demonstrated that the crimped VBX stent graft requires greater force to bend but offers superior flexibility when fully deployed when compared to the Atrium stent (MAQUET Holding B.V. & Co. KG) [9-11]. The data of the FEVAR cohort of the EMBRACE study will be analyzed separately, and postdoc and core laboratory analysis will investigate the conformability and flaring properties of this new stent graft.

Prior research demonstrated the suitability of SESGs in directional branches [12]. However, some endovascular surgeons prefer dedicated BESGs in BEVAR and believe that they may provide clinical value. First, BESGs provide the possibility of performing postdilatation (in situ diameter adjustment) at the distal edge of the stent and inside the branch cuff, which provides flexibility in different anatomical scenarios. Second, the deployment of BESGs is more accurate than SESGs because of the releasing balloon embedded in the stent graft, which allows gradual release and adjustments in case of stent migration. In contrast, SESGs have less precise releasing systems, which may cause displacement from the intended deployment spot, especially in the presence of short sealing

XSL•FO
RenderX

zones. Third, trackability remains a topic of debate between the 2 stent graft types. For stents of the same diameter, BESGs are compatible with smaller introducer sheaths compared to SESGs. In cases of narrow endografts and/or aortic lumens, the lower profile of BESGs ensures improved navigability and facilitates catheterization of target vessels [13].

Some limitations associated with BESGs stem from their increased rigidity compared to SESGs; this can lead to kinking of the bridged vessel at the distal edge of the stent. In such cases, a self-expanding bare metal stent is deployed within the BESG to smooth the angle created by the stent graft and restore lumen patency [14]. Stent graft rigidity can also compromise patency in small vessels, especially renal arteries [15,16]. Additionally, renal arteries are more prone to loss of patency compared to splanchnic vessels in branched configurations, regardless of the stent type used for bridging [5,17]. This increased susceptibility may be explained by the kidney's motion driven by diaphragm movement during breathing [18]. Another limitation of BESG is restricted length availability; for example, the most widely used stent globally, the Atrium (MAQUET Holding B.V. & Co. KG), has a maximum length of 59 mm. This feature may necessitate the use of 2 imbricated BESGs to bridge long distances, thereby increasing the risk of type III endoleak due to component disconnection [19]. In contrast, the VBX stent graft offers lengths ranging from 15 to 79 mm, making it a suitable option for BEVARs that require longer bridging distances.

The VBX stent graft is composed of independent stainless-steel rings that enhance strength and fluoroscopic visibility [20]. Interestingly, independent rings, which are held within a fluoropolymer sleeve that serves as the only linkage for the rings, make the stent less rigid compared to other BESGs, offering an alternative for tortuous vessels. The enhanced flexibility of the stent graft may avoid the use of self-expandable bare metal stents to prevent kinking at the distal sealing zone [21]. As reported in the literature, the performance of the VBX stent graft used as a bridging stent in F/B-EVAR seems reassuring and comparable to other stents available on the market [21-23]. Single-center studies evaluating the performance of the VBX stent graft in F/B-EVAR are essential to better define its indications for use. Previous studies' results were inconclusive, sparking debate on the application of the VBX stent graft in branched aortic endografts [24,25], where some authors argue that SESGs are preferable [26]. All these considerations justify the effort of the EMBRACE registry to analyze the performance of the VBX stent graft in BEVAR implants with the aim of assessing how its unique features may alter daily practices with a specific focus on renal artery outcomes.

Study Limitations

The retrospective nature of patient enrollment may introduce selection bias; however, precautions were taken to minimize this risk through standardized inclusion criteria. Variability in the number of patients enrolled in each cohort across sites may help mitigate site-specific biases, enhancing the generalizability of the results, particularly given the inclusion caps. Additionally, due to both the retrospective and confirmatory nature of the study, randomization and blinding were not possible, which may introduce additional potential biases.

Conclusion

With this study, the authors aim to clarify the role of balloon-expandable covered stents in F/B-EVAR and provide stronger evidence for their use in fenestrated and branched configurations. Once the study end point data are complete, we hope that the findings of this study will refine the indications for endoprostheses in the treatment of abdominal or thoracoabdominal aneurysms, with the inclusion of the VBX stent graft as a dedicated bridging stent in fenestrated or branched aortic endografts.

Acknowledgments

EMBRACE Investigators (in alphabetical order) are as follows: Luca Bertoglio, MD (luca.bertoglio@unibs.it; Division of Vascular Surgery, Department of Clinical and Experimental Sciences, University and ASST Spedali Civili Hospital of Brescia, Brescia, Italy); Jacob Budtz-Lilly, MD (jacoblilly@clin.au.dk; Department of Thoracic and Cardiovascular Surgery, Aarhus University, Aarhus, Denmark); Nuno Dias, MD, PhD (nuno.dias@med.lu.se; Vascular Center Malmö, Skånes University Hospital, Malmö, Sweden); Rocco Giudice, MD (rgiudice@hsangiovanni.roma.it; Vascular Surgery Unit, San Giovanni-Addolorata Hospital, Rome, Italy); Arjan Hoksbergen, MD (a.hoksbergen@amsterdamumc.nl; Department of Surgery, Amsterdam University Medical Centers, Location Vrije Universiteit, Amsterdam, The Netherlands); Bergros Johannesdottir, MD (bergros.kristin.johannesdottir@helse-bergen.no; Department of Vascular Surgery, Haukeland University Hospital, Bergen, Norway); Wassim Mansour, MD (wassim.mansour@uniroma1.it; Vascular Surgery Division, Department of Surgery

"Paride Stefanini," Policlinico Umberto I-"La Sapienza" University of Rome, Italy); Frode Manstad-Hulaas, MD (frode.manstad-hulaas@stolav.no; Department of Surgery, St. Olavs University Hospital, Trondheim, Norway); Isaac Martinez, MD (isaacml20@hotmail.com; Department of Angiology and Vascular Surgery, Hospital Clinico San Carlos, Madrid, Spain); Kare Nordhus, MD (kare.chr.nordhus@unn.no; Department of Radiology, University Hospital of North Norway, Tromsø, Norway); Giovanni Pratesi, MD (Giovanni.pratesi@unige.it; Clinic of Vascular and Vascular Surgery, IRCCS Ospedale Policlinico San Martino, University of Genoa, Genoa, Italy); Timothy Resch, MD (timothy.andrew.resch@regionh.dk; Department of Vascular Surgery, Rigshospitalet, Copenhagen, Denmark); Enrico Rinaldi, MD (rinaldi.enrico@hsr.it; Division of Vascular Surgery, IRCCS San Raffaele Scientific Institute, Vita-Salute San Raffaele University, Milan, Italy); Luis Miguel Salmeron, MD (lmsalmeron95@gmail.com; Hospital Universitario San Cecilio, Granada, Spain); and Anders Wanhainen, MD

XSL•FO
RenderX

(anders.wanhainen@surgsci.uu.se; Department of Surgical Sciences, Vascular Surgery, Uppsala University, Uppsala, Sweden). Paper preparation assistance was provided by Chrystal Redding and Kayla Stewart, employees of W. L. Gore & Associates.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article. This work was conceded as a part of a W. L. Gore & Associates clinical trial. W. L. Gore & Associates was involved in the approval of the study design, analysis, and interpretation of data in consultation with the authors. None of the authors or coauthors were provided financial compensation for the writing of the manuscript.

Conflicts of Interest

LB is a clinical investigator and consultant for Gore Medical and Cook Medical. AW, JB-L, and ND are clinical investigators for Gore Medical. TB is an employee of W. L. Gore & Associates, Inc, Flagstaff, Arizona.

References

1. Oderich GS, Forbes TL, Chaer R, et al. Reporting standards for endovascular aortic repair of aneurysms involving the renal-mesenteric arteries. J Vasc Surg 2021 Jan;73(1):4S-52S. [doi: 10.1016/j.jvs.2020.06.011]
2. Bertoglio L, Katsarou M, Loschi D, et al. Elective multistaged endovascular repair of thoraco-abdominal aneurysms with fenestrated and branched endografts to mitigate spinal cord ischaemia. Eur J Vasc Endovasc Surg 2020 Apr;59(4):565-576. [doi: 10.1016/j.ejvs.2019.10.003] [Medline: 31870689]
3. Tenorio ER, Schanzer A, Timaran CH, et al. Mid-term renal and mesenteric artery outcomes during fenestrated and branched endovascular aortic repair for complex abdominal and thoracoabdominal aortic aneurysms in the United States aortic research consortium. Ann Surg 2023 Oct 1;278(4):e893-e902. [doi: 10.1097/SLA.0000000000005859] [Medline: 37051912]
4. Mazzolai L, Teixido-Tura G, Lanzi S, et al. 2024 ESC Guidelines for the management of peripheral arterial and aortic diseases. Eur Heart J 2024 Sep 29;45(36):3538-3700. [doi: 10.1093/eurheartj/ehae179] [Medline: 39210722]
5. Abdelhalim MA, Tenorio ER, Oderich GS, et al. Multicenter trans-Atlantic experience with fenestrated-branched endovascular aortic repair of chronic post-dissection thoracoabdominal aortic aneurysms. J Vasc Surg 2023 Oct;78(4):854-862. [doi: 10.1016/j.jvs.2023.05.053] [Medline: 37321524]
6. Bertoglio L, Loschi D, Cambiaghi T, et al. Preliminary outcomes of the LifeStream balloon-expandable covered stent in fenestrated and branched thoracoabdominal endovascular repairs. J Endovasc Ther 2018 Apr;25(2):230-236. [doi: 10.1177/1526602817752449] [Medline: 29313455]
7. International ethical guidelines for health-related research involving humans. : Council for International Organizations of Medical Sciences (CIOMS); 2016 URL: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf [accessed 2026-01-29]
8. Martin-Gonzalez T, Mastracci T, Carrell T, et al. Mid-term outcomes of renal branches versus renal fenestrations for thoraco-abdominal aneurysm repair. Eur J Vasc Endovasc Surg 2016 Aug;52(2):141-148. [doi: 10.1016/j.ejvs.2016.03.018] [Medline: 27105550]
9. Torsello GF, Herten M, Müller M, Frank A, Torsello GB, Austermann M. In vitro evaluation of the Gore Viabahn balloon-expandable stent-graft for fenestrated endovascular aortic repair. J Endovasc Ther 2019 Jun;26(3):361-368. [doi: 10.1177/1526602819842569] [Medline: 30975017]
10. Torsello G, Müller M, Litterscheid S, Berekoven B, Austermann M, Torsello GF. Fatigue resistance of the Advanta V12/iCast and Viabahn balloon-expandable stent-graft as bridging stents in experimental fenestrated endografting. J Endovasc Ther 2021 Jun;28(3):463-468. [doi: 10.1177/1526602821996724] [Medline: 33629597]
11. Blanco Amil CL, Mestres G, Barahona F, Raventós C, Alomar X, Riambau V. Comparison of biomechanical properties of the iCover and VBX stent grafts when used as bridge stents in fenestrated endografts: an in vitro study. J Endovasc Ther 2024 Oct 18:15266028241289291. [doi: 10.1177/15266028241289291] [Medline: 39422262]
12. Nana P, Spanos K, Brodis A, et al. Meta-analysis of comparative studies between self- and balloon-expandable bridging stent grafts in branched endovascular aneurysm repair. J Endovasc Ther 2023 Jun;30(3):336-346. [doi: 10.1177/15266028221083458] [Medline: 35293261]
13. Bertoglio L, Grandi A, Veraldi GF, et al. Midterm results on a new self-expandable covered stent combined with branched stent grafts: insights from a multicenter Italian registry. J Vasc Surg 2023 Jun;77(6):1598-1606. [doi: 10.1016/j.jvs.2023.02.007] [Medline: 36822256]
14. Torsello GB, Pitoulias A, Litterscheid S, et al. Performance of the Gore VBX balloon expandable endoprosthesis as bridging stent-graft in branched endovascular aortic repair for thoracoabdominal aneurysms. J Endovasc Ther 2021 Aug;28(4):549-554. [doi: 10.1177/15266028211010455] [Medline: 33908821]
15. Tenorio ER, Kärkkäinen JM, Mendes BC, et al. Outcomes of directional branches using self-expandable or balloon-expandable stent grafts during endovascular repair of thoracoabdominal aortic aneurysms. J Vasc Surg 2020 May;71(5):1489-1502. [doi: 10.1016/j.jvs.2019.07.079] [Medline: 31611106]

XSL•FO
RenderX

16. Kärkkäinen JM, Tenorio ER, Pather K, et al. Outcomes of small renal artery targets in patients treated by fenestrated-branched endovascular aortic repair. Eur J Vasc Endovasc Surg 2020 Jun;59(6):910-917. [doi: 10.1016/j.ejvs.2020.02.015] [Medline: 32197996]
17. Mastracci TM, Greenberg RK, Eagleton MJ, Hernandez AV. Durability of branches in branched and fenestrated endografts. J Vasc Surg 2013 Apr;57(4):926-933. [doi: 10.1016/j.jvs.2012.09.071] [Medline: 23433817]
18. Tran K, Suh GY, Mougin J, Haulon S, Cheng C. Respiratory-induced changes in renovisceral branch vessel morphology after fenestrated thoracoabdominal aneurysm repair with the BeGraft balloon-expandable covered stent. J Vasc Surg 2021 Aug;74(2):396-403. [doi: 10.1016/j.jvs.2020.12.093] [Medline: 33548438]
19. Chait J, Tenorio ER, Mendes BC, et al. Impact of gap distance between fenestration and aortic wall on target artery instability following fenestrated-branched endovascular aortic repair. J Vasc Surg 2022 Jul;76(1):79-87. [doi: 10.1016/j.jvs.2022.01.135] [Medline: 35181519]
20. Mafeld S, Annamalai G, Lindsay TF, et al. Initial experience with Viabahn VBX as the bridging stent graft for branched and fenestrated endovascular aneurysm repair. Vasc Endovascular Surg 2019 Jul;53(5):395-400. [doi: 10.1177/1538574419840880] [Medline: 31018828]
21. Motta F, Parodi FE, Knowles M, et al. Performance of Viabahn balloon-expandable stent compared with self-expandable covered stents for branched endovascular aortic repair. J Vasc Surg 2021 Feb;73(2):410-416. [doi: 10.1016/j.jvs.2020.05.028] [Medline: 32473341]
22. Usai MV, Gargiulo M, Haulon S, et al. One-year results of a balloon expandable endoprosthesis as a bridging stent for branched endovascular aortic repair. J Vasc Surg 2023 Dec;78(6):1376-1382. [doi: 10.1016/j.jvs.2023.07.061] [Medline: 37572891]
23. Fouad F, Saleem BR, Tielliu IFJ, et al. Three-dimensional geometric analysis of Viabahn VBX bridging stent grafts in fenestrated endovascular aortic repair: a multicenter, retrospective cohort study. J Endovasc Ther 2026 Feb;33(1):267-275. [doi: 10.1177/15266028241248600] [Medline: 38708986]
24. Kappe KO, van Knippenberg SEM, Tran BL, et al. Mid-term outcomes of the Viabahn balloon-expandable endoprosthesis as bridging stent graft for fenestrated and branched endovascular aortic repair. J Endovasc Ther 2024 Nov 22:15266028241300005. [doi: 10.1177/15266028241300005] [Medline: 39575702]
25. Usai MV, Gargiulo M, Haulon S, et al. A three-year experience with the balloon expandable GORE VIABAHN VBX in the treatment of thoraco-abdominal aortic aneurysms within the EXPAND trial. J Vasc Surg 2025 Feb;81(2):319-323. [doi: 10.1016/j.jvs.2024.10.002] [Medline: 39389214]
26. Migliari M, Leone N, Veraldi GF, et al. Comparison of bridging stent grafts in branched endovascular aortic repair. J Vasc Surg 2024 May;79(5):1026-1033. [doi: 10.1016/j.jvs.2023.12.037] [Medline: 38154606]

List of Abbreviations

BESG: balloon-expandable stent graft
BEVAR: branched endovascular repair
CEC: clinical events committee
CS: covered stent
FEVAR: fenestrated endovascular repair
SESG: self-expanding stent graft
STROBE : Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Schwartz; submitted 13.Jun.2025; peer-reviewed by A Ali, G Piffaretti; revised version received 22.Dec.2025; accepted 29.Dec.2025; published 27.Feb.2026.

Please cite as:
Dioni P, Wanhainen A, Budtz-Lilly J, Dias N, Bonny T, Bertoglio L, EMBRACE study investigator
GORE VIABAHN VBX Balloon-Expandable Endoprostheses as a Bridging Stent With Branched and Fenestrated Endografts in the Endovascular Treatment of Aortic Aneurysms: Protocol for a Retrospective Multicenter Registry (EMBRACE Study)
JMIR Res Protoc 2026;15:e78970
URL: https://www.researchprotocols.org/2026/1/e78970
doi:10.2196/78970

© Pietro Dioni, Anders Wanhainen, Jacob Budtz-Lilly, Nuno Dias, Tab Bonny, Luca Bertoglio, EMBRACE study investigator. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR

XSL•FO
RenderX

Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Interventions for Students' Well-Being at the University of Helsinki (INSIGHT): Protocol and Preliminary Descriptive Results for a Quasi-Experimental Controlled Trial of a Social Identity Intervention and Two Active Comparators

Silja Martikainen[1,2*], PhD; Oona Karhunen[1*], MA; Iiris Mankki[1], BA; Vera Gergov[1], PhD; Päivi Berg[3], PhD; Satu Venäläinen[3], PhD; Elisa Rissanen[4], PhD; Johanna Lammintakanen[4], PhD; Erkki Heinonen[5,6], PhD; Anu Lehtinen[7], MA, MS; Janna Pöntinen[7], MA; Henna Asikainen[8], PhD; Nina Katajavuori[8], PhD; Outi Linnaranta[6], MD, PhD; Milla Räisänen[9], PhD; Samuli Salminen[1], BA, BS; Jari Lahti[1,10], PhD

[1]Department of Psychology, Faculty of Medicine, University of Helsinki, Haartmaninkatu 3, Helsinki, Finland

[10]Folkhälsan Research Centre, Helsinki, Finland

[2]Department of Education, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

[3]Inclusion and Equality Research Group, Faculty of Social Sciences, University of Helsinki, Helsinki, Finland

[4]Department of Health and Social Management, University of Eastern Finland, Kuopio, Finland

[5]Department of Psychology, University of Oslo, Oslo, Norway

[6]The Department of Healthcare and Social Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland

[7]University Services, Sustainable Wellbeing Unit, University of Helsinki, Helsinki, Finland

[8]HYPE Centre for University Teaching and Learning, Wells Research Project, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

[9]HYPE Centre for University Teaching and Learning, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

[*]these authors contributed equally

Corresponding Author:
Silja Martikainen, PhD
Department of Psychology, Faculty of Medicine, University of Helsinki, Haartmaninkatu 3, Helsinki, Finland

Abstract

Background: University students' mental health problems are prevalent globally, which underlines the need for accessible and cost-effective mental health services in universities. Loneliness is a key risk factor for mental health problems, and it disproportionately affects students from minority backgrounds. Therefore, addressing loneliness and fostering inclusion and equality can be crucial strategies for enhancing students' well-being.

Objective: The aim of this study is to investigate a social-identity group intervention called Groups 4 Health (G4H) for university students' well-being using both quantitative and qualitative methods. Here, we present the research protocol and report preliminary descriptive findings from the study cohort.

Methods: The quantitative part of the study is a 4 parallel-arm nonrandomized controlled trial aiming to recruit 600 student participants from the University of Helsinki. The experimental group, which receives the G4H intervention, includes 5 group meetings held over a 7-week period. The experimental group will be compared with 2 active comparators: groups organized by the University of Helsinki study psychologists and a 7-week online intervention course focused on well-being and study skills, and to a no-intervention control group. The primary quantitative outcomes of the study are loneliness and depression; secondary outcomes include several measures of students' well-being, academic performance, and cost-effectiveness of the intervention. Quantitative data are collected before the intervention, during the intervention (at week 3), immediately post intervention (at week 7 after baseline), and at 1- and 3-month follow-ups. The qualitative part of the study explores the challenges and opportunities related to inclusion and equality identified in the G4H intervention using observations, interviews, and focus group discussions.

Results: In the preliminary findings based on the first data freeze in March 2025, we observed differences in the background characteristics between the trial arms, highlighting the need to address group selection bias. First results from the study are expected in 2026.

Conclusions: If proven effective, these interventions have significant potential to improve students' well-being in both short and long term, fostering mental health and supporting academic success and future career paths.

Trial Registration: ClinicalTrials.gov NCT06542029; https://clinicaltrials.gov/study/NCT06542029

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/79319

(JMIR Res Protoc 2026;15:e79319) doi:10.2196/79319

KEYWORDS

university students; group intervention; effectiveness; cost-effectiveness; wellbeing; social inclusion; equality; multimethod research; controlled study; groups 4 health

Introduction

Background

University students' mental health problems are highly prevalent globally [1]. In Finland, 36% of university students reported high levels of psychological distress and 24% reported increased loneliness in the *2021 Finnish Student Health and Well-Being Survey* [2]. Anxiety and depressive symptoms at the start of university studies are often persistent and related to poorer academic performance and lower engagement [3]. The need for accessible, cost-effective mental health services in universities is substantial [4].

Loneliness (the discrepancy between one's actual and desired social relationships) has been recognized as one of the key risk factors for numerous adverse psychological and somatic health outcomes among students and youth [5] and in various population-based samples [6,7].

Having positive relationships and feeling socially connected is also essential for academic performance. Loneliness, social isolation, and low social status impact intellectual achievement [8], whereas engaging in the academic environment reinforces individuals' learning activities [9].

Mental health problems are more prevalent among many minority groups, who experience psychological symptoms more often than the rest of the population [10-12]. Several studies have also found that belonging to more than one minority group increases the risk of stressful life situations, mental health problems, and thus the need to use mental health services [13,14]. Belonging to minorities may also increase the risk of social isolation. According to the School Health Promotion study [12], young people belonging to gender and sexual minorities show elevated levels of loneliness, exclusion, and difficulties in receiving help in Finland. In addition, young people with immigrant backgrounds and young people with disabilities face considerable risks of loneliness [15]. These difficulties are largely caused by minority stress, which in turn has been linked with experiences of discrimination and unequal treatment [12,16].

With this background, interventions targeting loneliness and supporting social connectedness, inclusion, and equality could be a significant way to approach the student mental health crisis. The Groups 4 Health (G4H) intervention studied in this protocol aims at improving well-being by supporting individuals' social group memberships that are compatible with their social identity [17].

G4H builds on the social identity approach to health (SIAH) [18,19] and the social identity model of identity change (SIMIC) [20,21]. The SIAH [18,19] comprises 2 related theories, the social identity theory [22] and self-categorization theory [23], claiming that individuals' internalized group memberships have a significant role in their sense of self and that shared identities are important for productive social interaction. Therefore, belonging to multiple social groups that align with the individual's identity is crucial for a sense of purpose and belonging, as well as for receiving support from others [18]. The role of group memberships is further emphasized during transitional life events such as moving to another country or entering a university, which may be stressful and may also affect an individual's access to their social groups. The SIMIC further states that during stressful and transitional life situations, well-being can be enhanced by both managing and strengthening the individual's social group memberships before the life change (the social identity continuity pathway) and by becoming a member of new social groups that align with the individual's identity (the social identity gain pathway) [20,21].

The fact that G4H addresses major life changes and social identities makes it particularly relevant in the university context. Many students experience major changes, such as relocating to a new country or city, integrating into new communities, and leaving familiar ones behind. For minority students, finding new social groups that align with their identity can be even more challenging, emphasizing their need for support.

This protocol provides a unique opportunity to study the G4H in the university context with a multimethod approach that integrates both quantitative and qualitative methods. By implementing this approach, we aim to capture not only the effectiveness of the G4H intervention and the participant- or facilitator-related factors that moderate the effectiveness, but also the perspectives of minority students participating in the intervention, which is an aspect that has been overlooked in previous research. The inclusion of qualitative methods is crucial, as it allows us to gain deeper insights into the experiences and potential challenges faced by these students, enriching our understanding of the intervention's impact.

Additionally, neither the cost-effectiveness of the intervention nor its effectiveness on academic performance has been explored in previous studies. Both of these analyses are essential for assessing its suitability and adaptability in promoting student well-being. Furthermore, to our knowledge, the effectiveness of the G4H has been assessed only in Australia in studies organized by the developers of the intervention. These represent clear gaps in the existing literature that this protocol aims to address.

Aim and Objectives

The aim of this study is to investigate the effectiveness of the G4H intervention for university students' well-being using both quantitative and qualitative methods. Our study includes 2 active

XSL•FO
RenderX

comparators: intervention groups facilitated by the University of Helsinki (UH) study psychologists and an online intervention course focused on well-being and study skills. A control group of students who are not participating in any active well-being services at UH is also included.

Our first objective is to investigate the effectiveness and cost-effectiveness of the G4H intervention in (1) supporting well-being (most importantly by reducing loneliness and depression) and (2) supporting academic performance among university students. The second objective is to explore which factors moderate or mediate the effectiveness of the interventions. The third objective is to focus on the G4H intervention's inclusiveness and the capacity of the G4H intervention to promote equality via qualitative methods.

Our research questions (RQs) are as follows:

- RQ1: How effective is the G4H intervention in reducing loneliness and depression and in supporting university students' well-being compared with the 2 active comparators and the control group? (objective 1)
- RQ2: Does the G4H intervention support academic performance compared with the active comparators and the control group? (objective 1)
- RQ3: Is G4H cost-effective compared with the 2 active comparators and the control group? (objective 1)
- RQ4: How do factors related to the participant, facilitator, or intervention fidelity mediate or moderate the effectiveness of the intervention? (objective 2)
- RQ5: What are the challenges and opportunities related to inclusion and equality in the G4H intervention? (objective 3)

Compared with the no-intervention control group, we hypothesize that all active interventions reduce students' loneliness and depression (RQ1) and improve students' academic performance (RQ2). However, participants in the G4H intervention will experience a more long-term reduction in loneliness than those in the other active intervention groups [24]. We hypothesize that treatment alliance [25,26] and higher fidelity to the G4H manual mediate the interventions' outcomes (RQ4). Regarding facilitators' characteristics, we hypothesize that their greater self-rated skills and enjoyment in therapeutic work as well as agreeableness, openness, and engagement in personality traits and interpersonal relating will be positively associated with treatment alliances and intervention outcomes; and, conversely, greater difficulties and negative feelings in therapy work will be negatively associated with treatment alliances and outcomes [27-29]. Regarding motivation (RQ4), we set our hypotheses based on the self-determination theory [30] and the integral model of treatment motivation [31]. We hypothesize that autonomous motivation to participate in the intervention predicts greater engagement and increased effectiveness, whereas controlled motivation predicts lower engagement and effectiveness. We further hypothesize that engagement partially mediates the relationship between motivation and intervention effectiveness. No preliminary hypotheses are posed regarding RQs 3 and 5.

Methods

Trial Design

We conduct a parallel arm nonrandomized controlled trial with a 1:1 allocation ratio. We compare 3 intervention arms to a no-intervention control group as follows:

- Arm 1: The experimental group, G4H.
- Arm 2: Active comparator group 1, the study psychologists' group interventions (SPGs)
- Arm 3: Active comparator group 2, toward better well-being and studying intervention (TBWS).
- Arm 4: No-intervention control group.

The study design and timeline are presented in Figure 1. The study follows the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) guidelines.

Figure 1. Trial arms, intervention schedule, and points of assessment. The dashed vertical lines indicate points of assessment. A circle with a dashed outline indicates an optional meeting. △: online meeting; ◯: live or online meeting; X: Groups 4 Health facilitators' assessments.

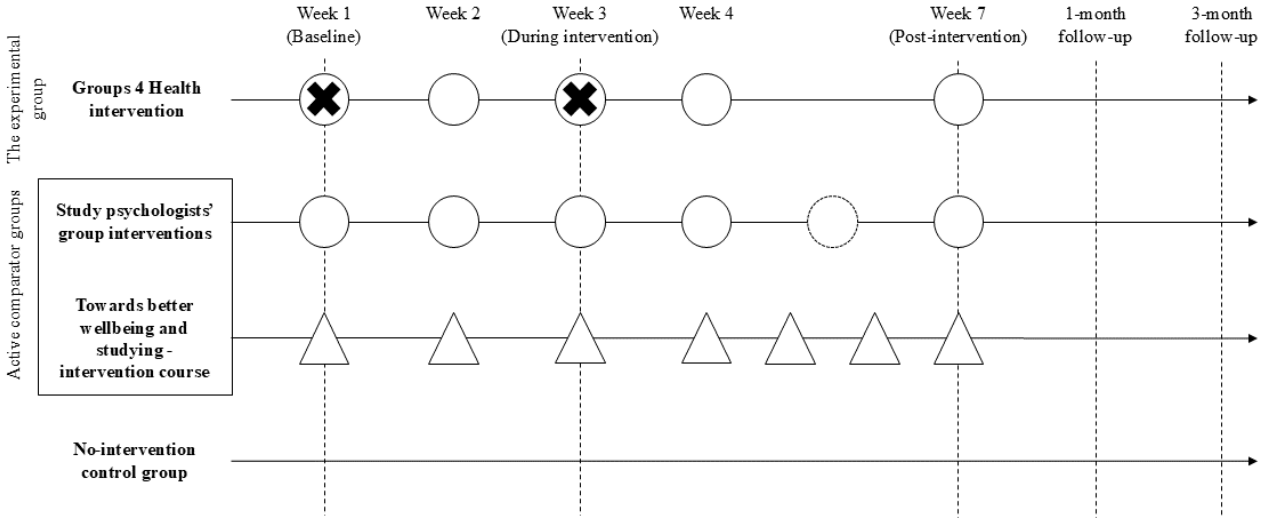

XSL•FO
RenderX

Interventions

Overview

The interventions in each active trial arm represent courses provided by UH for students who are seeking help to improve their overall well-being and study skills. Additionally, a control group consisting of students receiving no intervention is included. Participation in the study and in the interventions is voluntary, and participants can withdraw from participating at any point without any negative consequences.

Arm 1: G4H (The Experimental Group)

G4H is a manualized group intervention that targets social group memberships with the aim of improving well-being [17], as discussed in the introduction.

The G4H intervention consists of five 90-minute sessions that aim to provide participants with the knowledge and skills they need to effectively manage their social group memberships and identities. Each of the 5 intervention sessions pertains to 1 module. Each module integrates exercises and discussions from the G4H workbook targeting various aspects of group dynamics as outlined in the SIAH and SIMIC frameworks. In brief, the first module "schooling" provides psychoeducation about the importance of social group memberships for health; the second module "scoping" supports the participants in exploring their social identities and networks; the third module "sourcing" aims at strengthening the social identities that are valued by the participant; the fourth module "scaffolding" supports making a social action plan to establish new valued group memberships; and finally, the fifth module "sustaining" aims at revisiting the intervention, discussing any potential difficulties, and encouraging the maintenance of group memberships [17]. A detailed description of the modules can be found in the G4H manual [32].

At the UH, the G4H group size can vary between 4 and 9 university students. The groups are mostly organized face-to-face, although some groups meet entirely remotely. The students earn 2 European Credit Transfer and Accumulation System (ECTS) credits for completing the G4H intervention.

G4H groups at the UH are facilitated by 2 psychology students carefully trained for the intervention and mentored by experienced clinical psychologists. Psychology students interested in facilitating the G4H intervention can sign up for G4H facilitator training. During the 2-month training course, facilitators are introduced to the intervention literature (approximately 15 h) and attend lectures on the background theory, structure, and content of the intervention (approximately 6 h). In the more practical part of the training, the facilitators practice how to deliver the intervention and gain first-hand experience of participating in the intervention by conducting the intervention with each other. Finally, after the practice meetings, each pair facilitates one 5-session G4H intervention for a group of UH students. Additional reflection meetings are held with a clinically experienced psychologist (J Lahti), where the facilitators share their experiences and any problems or questions they may have encountered. During the training, facilitators keep a learning diary, which supports further introspection and reflection. The student facilitators earn 2.5 ECTS credits for completing the intervention training.

The effectiveness of G4H has been studied earlier in Australia. In a nonrandomized controlled study, G4H (n=83) significantly improved well-being and social connectedness of young adults after the intervention and in 6-month follow-up relative to the control group (n=75). In a randomized controlled trial by Haslam and colleagues [33], G4H (n=66) decreased loneliness and social anxiety and increased group memberships significantly more than treatment as usual (n=54) [17]. In a randomized noninferiority trial by Cruwys and colleagues [24], G4H (n=84) was compared with cognitive behavioral therapy (CBT; n=90). Both interventions were equally effective in reducing loneliness by the end of the intervention period. However, the G4H group demonstrated a maintained reduction in loneliness during the follow-up. Both interventions significantly reduced depressive symptoms, with no significant differences between G4H and CBT [24]. In all 3 controlled studies on the effects of G4H, the inclusion criteria were low mood and loneliness.

Arm 2: SPGs (Active Comparator 1)

The UH study psychology team organizes group interventions for various challenges related to studying for bachelor's and master's students. The SPGs offer client-centered peer support facilitated mostly by 2 study psychologists (some individual sessions may have only 1 facilitator). The recurring themes in the SPGs include challenges with attention, independent studying, scientific writing, and support for re-entering the studies after a break. The structure of the different thematic groups is similar—short presentations, group discussions, practices, and homework. These groups are implemented mostly as face-to-face meetings; however, in some groups, part of the meetings are conducted face-to-face and part as remote meetings. One group typically lasts for a study period and comprises five to six 90 - 105 minute sessions. The typical group size ranges from 10 to 12 students. The interventions have potential in supporting study skills and the challenges of initiating study tasks, identifying and normalizing unpleasant thoughts and emotions related to studying, and acknowledging the role of one's individual situation and well-being in study planning. The study psychologists' work is bound by the legal and ethical requirements for licensed psychologists in Finland, overseen by the National Supervisory Authority for Welfare and Health (Valvira).

Arm 3: TBWS (Active Comparator 2)

TBWS is an optional, 7-week online intervention course available for all UH students that aims to foster well-being and study skills. The course protocol has been described in detail earlier by Asikainen and Katajavuori [34]. The course is based on the principles of acceptance and commitment therapy [35] and aims to promote psychological flexibility. In addition, it aims to support study skills such as time management. The course takes place on an online learning platform (Moodle), and the students earn 3 ECTS credits for completing it. This course progresses weekly and has 6 central themes, including an introduction to the themes, individual work in practicing psychological flexibility and study skills, and weekly peer group discussions. The peer group work is an important part of the

XSL•FO
RenderX

intervention and lasts throughout the course. In the peer groups, there are about 5 students in each group, and the groups remain the same throughout the course. Acceptance and commitment therapy–based interventions have been shown to support well-being among university students [36], and previous studies have suggested that this course can improve students' well-being and studying in several ways [37,38]. Thus, participating students have the potential to benefit from this intervention in terms of increased well-being.

Arm 4: The No-Intervention Control Group

The control group consists of students who are currently not taking part in any group interventions for study skills or psychological well-being at the UH. The participants in the control group may, however, take part in other social and mental health services than those provided by the UH, and this is not restricted during their participation in the study. The recruitment and data collection among the control group are scheduled to take place simultaneously with the active intervention groups: the participants are recruited at the same time, and they answer the questionnaires at the same pace as the active interventions progress during the academic year.

The interventions studied in this protocol have been developed for educational contexts and thus are expected to be well-suited for the university context. As discussed earlier, the G4H builds on SIAH and SIMIC, particularly aiming at supporting well-being during transitions in life, such as entering a university [21]. In addition, both the SPGs and the TBWS courses have been specifically developed by experts within the UH for supporting university students' well-being and academic ability.

Participants

Setting and Timeline

The study is conducted in an academic setting among UH students in Finland. The recruitment has progressed as follows: for the RQ1-RQ3, the recruitment for the first data freeze ended at the end of 2024, and for RQ5, the recruitment ends by the end of the year 2025. For the more explorative objective 2 (RQ4) on the factors that mediate or moderate intervention effectiveness, we assume that a larger sample is required, and recruitment continues longer (current ethical approval allows data collection until the end of 2028). We completed the first data freeze in March 2025 for the RQ1 to RQ3 and will complete the data collection of the qualitative data for the RQ5 by December 2026.

The quantitative data collection started in March 2022 for the G4H and the SPGs. From October 2022, the data collection from all 3 interventions and the no-intervention control group has taken place simultaneously. The intervention schedule and points of assessment are illustrated in Figure 1. It should be noted that data on health care and social service use, as well as the participants' minority status and readiness to attend the intervention, were gathered only from students who participated starting from March 2023 for the no-intervention control group and from September 2023 for the active intervention arms. The research ethics committee was informed about these added outcome measures, and a new ethical approval was obtained from the committee before data collection started.

In the collection of qualitative data, participant observation is conducted in the meetings of the G4H group and interviews with facilitators (n=9) and intervention participants (n=11) at the end of the intervention period. In addition, the follow-up interviews with intervention participants will be conducted 12 months after the first interview (2024 - 2026). Additionally, focus groups with young people belonging to minorities (n=19 participants) are arranged to discuss G4H in terms of equality and inclusion during the data collection period (2023 - 2025).

Sample Size

For RQ1-RQ3, the target minimum sample for the quantitative study is 600 participants, distributed evenly across 4 groups (150 participants per group). Since comparable previous studies with a similar group of participants using the same outcome measures are not available, it was not possible to make exact a priori sample size calculations. However, based on an estimation with the G*Power software (version 3.1) [39] on a repeated measures analysis of variance design with between-groups factors, this target sample size would yield at least 95% power for the detection of small effect sizes (f=0.20) as defined by Cohen [40] when considering multiple comparisons and a 35% attrition rate. Previous studies with different designs investigating the G4H intervention have demonstrated significant treatment effects with smaller samples ranging from 66 [33] to 84 [24] participants in the G4H group, supporting the expectation that our larger sample will be sufficient to detect meaningful treatment effects, or lack thereof.

Recruitment

The UH student well-being services are available to all students at the university, and there are no exclusion criteria for the study. All the students willing to participate are included regardless of their level of loneliness, state of psychological well-being, or stage of studies. Students are recruited for the active intervention groups and no-intervention group by email, website, and social media advertisements circulated by the UH media and communications services, communications services at the faculties, student health services, and student organizations. The students are not randomly assigned to the intervention groups and can sign up for a course of their choice. The facilitators of the G4H intervention are also invited to participate in the study during the G4H facilitator training. All facilitators willing to participate are included in the study. Taking part in relevant concomitant care is not prohibited during the intervention study. The participating students or facilitators are not blinded to the study protocol.

Quantitative Methods

To collect the quantitative data, we use paper-and-pencil and electronic versions of the questionnaires. Those in the active intervention groups fill out the questionnaires before or after the intervention sessions if the questionnaires are provided electronically and during the session if the questionnaires are provided on paper. Those in the no-intervention group fill out the questionnaires based on the email reminders of the electronic data collection system (time interval between reminders is the same as the interval between sessions in the active intervention groups).

XSL•FO
RenderX

Electronic data collection is implemented in the Research Electronic Data Capture (REDCap) system (Vanderbilt University) [41], provided by the UH for collecting sensitive research data electronically. In the case of the TBWS intervention, data collection from time points before and immediately after the intervention is partly integrated into the online course platform, and the students can respond either via the course platform or via the REDCap system, while data collection from week 3 and the 1- and 3-month follow-ups is implemented only in the REDCap.

Before the analysis phase, the data validity is confirmed via range checks, missing value checks, and outlier detection. Potentially erroneous data are checked against the paper and electronic forms, and other necessary sources, and corrected if needed. All modifications are documented.

Participant retention is promoted by reminding the participants during the group meetings to fill out the questionnaires. All participants receive up to 5 reminder messages from the electronic data collection system if they do not complete the necessary questionnaires on time. Retention is also promoted by the possibility of receiving written feedback on the questionnaire results after participation. Participating students also receive movie tickets as compensation for completing all follow-up questionnaires. Adherence to the G4H intervention is monitored by documenting the number of sessions in which each student participated.

All the quantitative methods used are presented in Table 1, which also includes the measurement time points and study arms from which the data are collected. Multimedia Appendix 1 includes a table that describes the unabbreviated titles of each scale, scale details, and information on their reliability and validity. Most measures have demonstrated at least acceptable internal consistency (Cronbach $\alpha \geq 0.70$) and convergent validity with other similar measures. The few exceptions with lower Cronbach α are either very short scales (4 items) or include a specific subscale that performs poorly.

XSL•FO
RenderX

Table . The quantitative assessment design including areas of assessment, used measures, and an assessment schedule

Assessment	Scales, forms, and measures	Time points and arms (1-4)[a] where data are collected						
		Week 1	Week 2	Week 3	Week 4	Week 7	1-month follow-up	3-month follow-up
Primary outcome measures								
Loneliness	ULS-8[b] [42]	1 - 4	—	—	—	1 - 4	1 - 4	1 - 4
Loneliness	ULS-4[c] [43]	1 - 4	—	1 - 4	—	1 - 4	1 - 4	1 - 4
Depression	PHQ-9[d] [44]	1 - 4	—	—	—	1 - 4	1 - 4	1 - 4
Secondary outcome measures								
Active group memberships	EXITS[e] [45]	1 - 4	—	1 - 4	—	1 - 4	1 - 4	1 - 4
Mental well-being	SWEMWBS[f] [46]	1 - 4	—	—	—	1 - 4	1 - 4	1 - 4
General well-being	GP-CORE[g] [47]	1 - 4	—	—	—	1 - 4	1 - 4	1 - 4
Anxiety	GAD-7[h] [48]	1 - 4	—	—	—	1 - 4	1 - 4	1 - 4
Social anxiety	Mini-SPIN[i] [49]	1 - 4	—	—	—	1 - 4	1 - 4	1 - 4
Quality of life	EQ-5D-5L[j] [50]	1 - 4	—	—	—	1 - 4	—	1 - 4
Use of social and health care services	Health service usage self-report (medications, health care appointments, hospitalization, and counseling)	1 - 4	—	—	—	1 - 4	—	1 - 4
Time used by facilitators	Time used at planning, executing, and assessing the sessions	1	1	1	1	1	—	—
Study credits	UH[k] registries, 1 year before and 1 year after the intervention (arms 1 - 4)	—	—	—	—	—	—	—
Grade point average	UH registries, 1 year before and 1 year after the intervention (arms 1 - 4)	—	—	—	—	—	—	—
Weighted grade point average	UH registries, 1 year before and 1 year after the intervention (arms 1 - 4)	—	—	—	—	—	—	—
Other outcome measures								
Identification with the intervention group	SISI[l] [51]	—	—	1 - 3	—	—	—	—

XSL•FO
RenderX

Assessment	Scales, forms, and measures	Time points and arms (1-4)[a] where data are collected						
		Week 1	Week 2	Week 3	Week 4	Week 7	1-month follow-up	3-month follow-up
Participants' working alliance	WAI-SR[m] [52]	—	—	1 - 2	—	—	—	—
Facilitators' working alliance	WAI-SRT[n] [53]	—	—	1	—	—	—	—
Mood	Analog mood scale (0 - 7)	1 - 4		1 - 4		1 - 4	1 - 4	1 - 4
Fidelity to the G4H manual	G4H[o] manual fidelity checklist	1	1	1	1	1	—	—
Participant attendance	Number of sessions participated	1	1	1	1	1	—	—
Background variables								
Participant background								
Sociodemographic characteristics	Gender, birth date, mother tongue, marital status, household composition, residential situation, employment status, education, income, subjective financial situation, subjective health, subjective loneliness, and subjective minority status	1-4	—	—	—	—	—	—
Social support	SSQ3[p] [54]	1 - 4	—	—	—	—	—	—
Treatment motivation	SIMS[q] [55], tailored motivation rulers [56]	1 - 3	—	—	—	—	—	—
Organized studying	4 items from a HowULearn questionnaire [57]	1 - 4	—	—	—	—	—	—
Facilitators' assessment								
Sociodemographic characteristics	Gender, birth date, mother tongue, marital status, household composition, residential situation, employment status, education, and income	—	—	—	—	—	—	—

XSL•FO
RenderX

Assessment	Scales, forms, and measures	Time points and arms (1-4)[a] where data are collected						
		Week 1	Week 2	Week 3	Week 4	Week 7	1-month follow-up	3-month follow-up
Professional and individual characteristics	XS5[r] [58], ECR-S[s] [59], TBIF-PI[t] [60]	1	—	—	—	—	—	—
Self-assessed skillfulness	TWIS[u] [61]	—	—	1	—	—	—	—

[a]Arm 1: Groups 4 Health intervention; arm 2: study psychologists' group interventions; arm 3: toward better well-being and studying intervention; and arm 4: mo-intervention control group.

[b]ULS-8: 8-item University of California Los Angeles Loneliness Scale.

[c]ULS-4: 4-item University of California Los Angeles Loneliness Scale.

[d]PHQ-9: Patient Health Questionnaire.

[e]EXITS: Exeter Identity Transition Scales

[f]SWEMWBS: Short Warwick-Edinburgh Mental Well-being Scale.

[g]GP-CORE: General Population version of the Clinical Outcomes in Routine Evaluation –scale

[h]GAD-7: Generalized Anxiety Disorder -questionnaire.

[i]Mini-SPIN: Mini Social Phobia Inventory.

[j]EQ-5D-5L: EuroQol - 5 dimension - 5 level questionnaire.

[k]UH: University of Helsinki.

[l]SISI: Single-Item Social Identification measure.

[m]WAI-SR: Working Alliance Inventory Short and Revised.

[n]WAI-SRT: Working Alliance Inventory Short and Revised Therapist version.

[o]G4H: Groups 4 Health.

[p]SSQ3: Social Support Questionnaire (three-item version).

[q]SIMS: Situational Motivation Scale.

[r]XS5: Extra-Short Five.

[s]ECR-S: Experiences in Close Relationships Short form.

[t]TBIF-PI: Personal Identity scale of the Trainee Background Information Form.

[u]TWIS: Therapist Work Involvement Scales.

Primary Outcomes

The primary outcomes are loneliness and depression (refer to Table 1 and Multimedia Appendix 1 for details). Based on previous evidence, the G4H intervention has at least equal potential to promote well-being in these areas as CBT [24].

Secondary Outcomes

The secondary outcomes are active group memberships, mental well-being, general well-being, anxiety, social anxiety, academic performance, health-related quality of life, use of social and health care services, intervention costs, and cost-effectiveness of the intervention (refer to Table 1 and Multimedia Appendix 1 for details).

To assess participants' academic performance, data will be retrieved from the UH student register (Sisu student information system). Academic performance is determined by 3 outcome variables: overall accumulated study credits, the grade point average (GPA) of the study credits and the weighted GPA. When determining the weighted GPA, only completed studies that have been assessed on a numeric scale (from 1 to 5) are considered, as some courses accumulate study credits but are graded only as passed or failed. Weighted GPA is determined by the grades of each course weighted by the number of credits gained for it. The academic performance data are collected from 2 time periods: a 1-year period before and a 1-year period after starting the assigned intervention (baseline, week 1).

Cost-effectiveness analyses will apply a societal perspective by including intervention costs as well as university students' health and social services costs. For the cost-effectiveness analyses, the facilitators will provide information on the time used for planning, executing, and assessing the sessions to evaluate the costs of the G4H intervention. Furthermore, the students in all trial arms fill out 2 questionnaires before the intervention, immediately after the intervention, and at 3-month follow-up (Table 1). The questionnaires measure health-related quality of life and self-reported use of health and social services (Table 1). The use of health and social services will be monetarized using Finnish national unit costs [62].

Other Outcomes

Other outcome measures include the participant's identification with the intervention group, attendance, and mood, as well as facilitators' fidelity to the intervention manual, and both the participant's and facilitator's evaluations of the working alliance (Table 1 and Multimedia Appendix 1). These measures are used to examine their potential in moderating or mediating the effectiveness of the intervention. Fidelity to the intervention

XSL•FO
RenderX

manual is measured with a questionnaire detailing how well the facilitators covered the core themes of each intervention session. Each facilitator pair fills in the questionnaire immediately after each intervention session. This questionnaire is the same as used by the previous study that assessed program adherence to the G4H [33] .

Background Variables

In addition to the outcome measures described above, survey data on the participants' background are gathered to account for them in the analyses and to examine their potential in moderating or mediating the effectiveness of the intervention (Table 1 and Multimedia Appendix 1). The participant background questionnaire includes questions on sociodemographic characteristics (refer to Table 1 for details), experienced social support, organized studying, and situational intrinsic and extrinsic motivation. Readiness to attend the intervention is enquired by a set of readiness rulers adapted from Miller and Rollnick [56]. Furthermore, data are gathered from the UH registries, including information on study rights (faculty, degree programs, and fields of study) and their starting years, active and passive registration, and completed degrees.

Facilitators' Characteristics

To investigate how facilitators' individual professional and personal characteristics may predict the intervention process and effectiveness, G4H facilitators complete self-report measures both before and during the interventions (Table 1 and Multimedia Appendix 1). Before initiating the intervention, facilitators complete questionnaires related to personality traits, attachment style, and self-experiences in close personal relationships. After the third session, the facilitators report experiences of their current skillfulness, difficulties, and emotions in group work.

Statistical Analyses

We will conduct our analysis following the intention-to-treat principle, including all participants regardless of whether they completed the intervention.

The change in the outcome variables over time will be analyzed with linear mixed-model analyses due to the hierarchical data structure. Information criteria are used to assess model fit and determine the appropriate model structure. Mixed effects models can also be fitted for participants with missing data from one or more time points. If needed, other appropriate imputation methods, such as multivariate imputation by chained equations [63], will be used. For this reason, the number of participants included in the analyses is likely to be higher than the number of participants providing data at the last follow-up.

We will explore with logistic regression analyses whether background characteristics and baseline psychiatric symptoms predict study dropout, and necessary means to account for any systematically missing data will be taken. We aim to show that the missing observations of the outcome variable, due to dropout, are conditionally independent from the observed values of the outcome variable given our model for dropping out from the study. If needed, inverse propensity score weighting is carried out so that the treatment effects are conditionally independent from the treatment group assignments and from differing dropout rates given the propensity scores. We consider individuals' selection into different study groups (arms) under the stable unit treatment value assumption based on their background characteristics, given that we have access to an extensive dataset. The selection of individuals into groups is accounted for by making the estimators double robust, which is a well-known approach to adjust for selection into groups based on background characteristics in model-based estimates [64,65]. This ensures that our results are unbiased and consistent estimators of the treatment effects. Potential effects of the COVID-19 pandemic on the stability and change in students' well-being will be inspected and considered in the analyses if necessary.

All models will be inspected for normality of residuals, homoscedasticity, and multicollinearity. The outcome variables will be transformed to attain normality if needed. Appropriate interaction terms are added for further subgroup analyses. Potential mediator (eg, therapeutic alliance, motivation, and model fidelity) and moderator (eg, participant and facilitators' background) effects will be inspected in additional mediation and moderation analyses.

In the case of missing questionnaire items, within-person imputation is carried out for responses when less than 50% of the answers are missing per scale (the within-person average of the existing items is imputed to the missing responses). We acknowledge that this procedure may decrease the within-person variation in the data. However, as we discuss later, the prevalence of item nonresponse in the dataset is low.

We will take appropriate measures to balance between Type I and Type II error in the reporting of analyses results. If needed, we will use methods to account for multiple testing such as false discovery rate or new eigenvalue-based methods that provide a more accurate adjustment for nonindependence among tests [66].

Cost-Effectiveness

We will conduct 3-month follow-up cost-utility and cost-effectiveness analyses. The cost-utility analysis will use EQ-5D-5L–based quality-adjusted life-years as the effectiveness outcome measure [50,67]. The cost-effectiveness analysis will use both primary outcomes of the effectiveness study, namely loneliness measured with the 8-item University of California Los Angeles Loneliness Scale [42] and depression measured with the Patient Health Questionnaire [44] (Table 1 and Multimedia Appendix 1) as the effectiveness outcome measure. The analyses will use a societal perspective by including the intervention costs and costs of students' health and social service use. The analyses will be conducted with the intention-to-treat principle. The missing outcome measures will be imputed with multiple imputation methods [68]. We will estimate incremental net health benefits between interventions. The estimation of the incremental net health benefits and 95% CIs will be used for statistical methods (eg, ordinary least squares regressions) [69]. We will perform multiple 1-way sensitivity analyses to test the robustness of the economic evaluation results. The results of sensitivity analyses will be presented in a tornado diagram. If interventions show effectiveness, we will conduct

XSL•FO
RenderX

decision-analytical modeling of long-term cost-utility and cost-effectiveness beyond 3-month follow-up data.

Qualitative Methods

We explore inclusion and equality, with respect to G4H using three qualitative data collection methods: (1) participant observation, (2) interviews with G4H participants and facilitators, and (3) focus groups with young people who belong to minorities. The aim is to examine challenges and possibilities in the G4H intervention regarding social inclusion and equality and whether G4H promotes long-term social inclusion and equality among the participants. The focus groups target young people belonging to sexual and gender minorities, young people with immigrant backgrounds, and young people with disabilities. These groups represent youth minorities with a considerable risk of loneliness [15]. To date, research on their experiences with different services is lacking [12,70-72]. These groups are also underrepresented in higher education [2,73,74].

The participant observation focuses on the interaction and the practices undertaken in the groups to gain knowledge about how the groups work, what kinds of issues they deal with, and especially how inclusion and equality are accomplished in the groups (are there minority participants, can participants influence the themes dealt with in the groups, or is inclusion and equality taken into account in some other ways in the groups). Following the ethnographic research tradition, the researcher takes part in the activities of the groups instead of observing them from the outside. Depending on their preferences, the study participants (both facilitators and participants in the groups) will be interviewed in a group, in pairs, or individually, in facilities located outside the university (eg, in a study room of a local library).

In the focus groups, young people belonging to minorities are asked to comment on the design or description of G4H in terms of equality and inclusion, with the goal of enhancing inclusivity by listening to their views on the interventions and accumulating knowledge on the fit between the interventions and their specific needs. All the interviews will be tape-recorded and transcribed for analysis.

To ensure the validity of qualitative findings, data triangulation is used via collecting 3 different types of data: focus group interviews, intervention participant interviews, and facilitator interviews. Setting these different data into a dialogue enables a multifaceted understanding of intervention effectiveness, inclusion, and equity. The analyses of qualitative data draw on minority stress theory [75] and intersectional theory [76,77].

The analyses are conducted by following reflexive thematic analysis [78,79]. Reflexive thematic analysis is conducted by following 6 steps that move from familiarization with data, coding, formulating, and refining themes to writing the report in alignment with the key principles of the approach, highlighting reflexivity and in-depth attention to nuance in participants' meaning-making. If several analysts are involved in the analysis, the goal is to engage in discussion throughout each step of the analysis, to refine the analysis with the help of various analysts' viewpoints. This approach, therefore, acknowledges that analysts' interpretations of the data may differ, which is why transparency and a close consideration of the subjectivity of interpretation are highlighted. The first steps of the analysis may be conducted with the help of software (such as ATLAS.ti 24; Lumivero), but the actual analysis requires repeated rounds of reading of the whole dataset by the analysts. The emphasis on researcher reflexivity in this analytical approach further enhances the validity of interpretations and helps ensure that the findings reflect the multiplicity of participants' viewpoints and accurately portray repetitive patterns in the data.

Ethical Considerations

Overview

The study protocol was approved by the Research Ethics Committee in the Humanities and Social and Behavioural Sciences at the UH, Finland (Statements: 15/2022, 32/2023, 10/2025). The ethical board will be informed in case of important protocol modifications. Written informed consent to participate in the study is obtained from all participants and facilitators. The participants are informed about the study via an information sheet and a data protection notice detailing the research protocol, data management, and participants' rights following the General Data Protection Regulation (GDPR) of the European Union. The consent form can be signed electronically or on paper. Electronic consent forms are delivered via an email from the REDCap system before the first intervention session or at a corresponding time point for the control group. Paper consent forms are delivered during the first meeting of each active intervention. The participants are also told that they may be contacted about potential ancillary studies. The participants are informed that they can withdraw from the study at any point without any consequences. The participants receive movie tickets as compensation for participating. A data monitoring committee has not been appointed as there is no suspicion of harm caused by the interventions studied in this trial to the participants.

Privacy and Confidentiality

The access to the data collection systems is restricted to assigned data managers. The questionnaire data collected on paper are transferred to REDCap by a data manager. Paper data are stored in a secure locked facility with restricted access provided by the Faculty of Medicine at the UH. All the data are downloaded from the data collection system and backed up on secure UH servers with restricted and password-protected access. Upon download, the data are pseudonymized via ID codes, and data that include personal identifiers are stored separately. All identifiers enabling linkage to the university records are removed from the data during the pseudonymization process. No such data that could risk reidentification of the participants is shared outside the research group.

The participants' identifiers are not shared outside the research group at the UH. Other pseudonymized research data may be shared with selected research partners for further research. A research collaboration agreement and a data sharing and processing agreement, which specifies the secure management and use of the data, will be made.

XSL•FO
RenderX

The consent form including identifiers will be destroyed 10 years after the end of the research project. Pseudonymized data will be kept for later, compatible scientific research in accordance with the requirements of the GDPR. The data will be stored at a secure UH server and kept protected so that only those who need the data can access them. Before the data can be used for later, compatible scientific research, the controller will ensure that the new use of data is compatible with the initial purpose in accordance with the requirements of the GDPR. Participants will receive a new data protection notice on the new use of the research data, unless the controller can no longer identify the participants from the data.

Participants' safety

We will do everything we can to ensure the safety of the participants. We do recognize that some topics discussed during the sessions, such as challenging group memberships and difficulties related to academic stress, may cause distress (by activating existing anxieties). The group sessions are designed to create a supportive environment where students feel secure and comfortable to express concerns to the facilitators. Each intervention has clear protocols on providing further support to the students if needed, and the training of the facilitators or course organizers includes plans to mitigate risks to the safety of the participants. Moreover, anonymous feedback is collected after all the interventions to tackle potential issues for the subsequent courses.

In the G4H course, potential challenges are addressed during the facilitator training and the mentoring sessions. Facilitators are instructed to consult with the course supervisor, J. Lahti, a clinical psychologist and licensed trainer-level psychotherapist, if they have any concerns about a participant's well-being. Additionally, they are provided with resources for further support, including contacts within the university's counseling services and public health care, which can be shared with participants if needed.

In the TBWS course, students are also provided resources for further support, including contacts within the university's counseling services and public health care. After assessing their well-being in the beginning of the course, all students get feedback and are instructed to seek this additional support if they get worried about their well-being. In addition, students get detailed instructions for peer group discussions instructing them to focus on the course themes and how the learned aspects can support their studying rather than discussing their personal health issues. Students are also requested to contact course teachers if they become concerned about their own or a peer's well-being.

In the SPGs, psychological safety is promoted by discussing the group's rules during the initial meeting. Students have the option to anonymously inform the facilitators, via a note, about any specific concerns or needs they would like the facilitators to be aware of. Study psychologists guide group discussions and intervene in potential challenging situations. If concerns arise, the psychologist offers discussion support to the student. Additionally, students have the opportunity for individual guidance from the study psychologist, who can assist them in contacting health care services if necessary. Students are encouraged to provide anonymous feedback to identify any potential weaknesses of the intervention.

We are aware that the young people invited to participate in the study about inclusion and equality might be in vulnerable positions. Our research team has extensive experience in conducting qualitative interviews with young people in vulnerable positions or who belong to minorities, such as LGBTIQ+ young people (lesbian, gay, bisexual, transgender, intersex, queer, and other), young people with diverse ethnic backgrounds, and young people not in employment, education, or training. The interviewers aim at sensitivity in their interaction with the interviewees and, for instance, provide the interviewees with room for regulating the depth and detail in which they discuss their experiences and emotions. The interviews and focus groups have been carefully designed to avoid placing excessive mental strain on participants. Participants are free to decide for themselves how much detail they wish to share about their own experiences and what they wish to reveal about their minority identities and related experiences. In addition, the use of peer facilitators in the focus groups promotes a sense of security and confidentiality, as well as peer support.

The opportunity to talk about one's experiences with the G4H intervention and the opportunity to be respectfully heard can be empowering for the interviewed young people. In addition, sharing one's thoughts and experiences of observed practices can alleviate stress and provide a sense of social support. On the other hand, talking about potential negative aspects, or remembering and describing possible experiences and observations of discrimination, or talking about experiencing limitations related to gender identity, sexual orientation, ethnicity, or ability can cause minor, short-term discomfort. However, according to our evaluation, this is highly unlikely to exceed the discomfort caused by the challenges and small annoyances of normal everyday life. The participants will be offered an opportunity to talk about the feelings caused by the interview afterwards with the interviewer or the researcher responsible for the entire research project. They will also be provided with contact information for services that provide support for young people, and if necessary, the interviewer will help them find relevant services in their hometown. If the young person wishes to, they will be given a chance to read the transcript of the interview, that is, the transcribed text of the audio recording.

Results

Study Status

Figure 2 illustrates the number of participants from the start of the study in March 2022 until the first data freeze in March 2025.

XSL•FO
RenderX

Figure 2. Study recruitment flow chart. Duplicate participants had already participated in the study (either in the same trial arm or in another arm). Controls: no-intervention control group; G4H: Groups 4 Health; SPG: study psychologists' group interventions; TBWS: toward better well-being and studying intervention; UH: University of Helsinki.

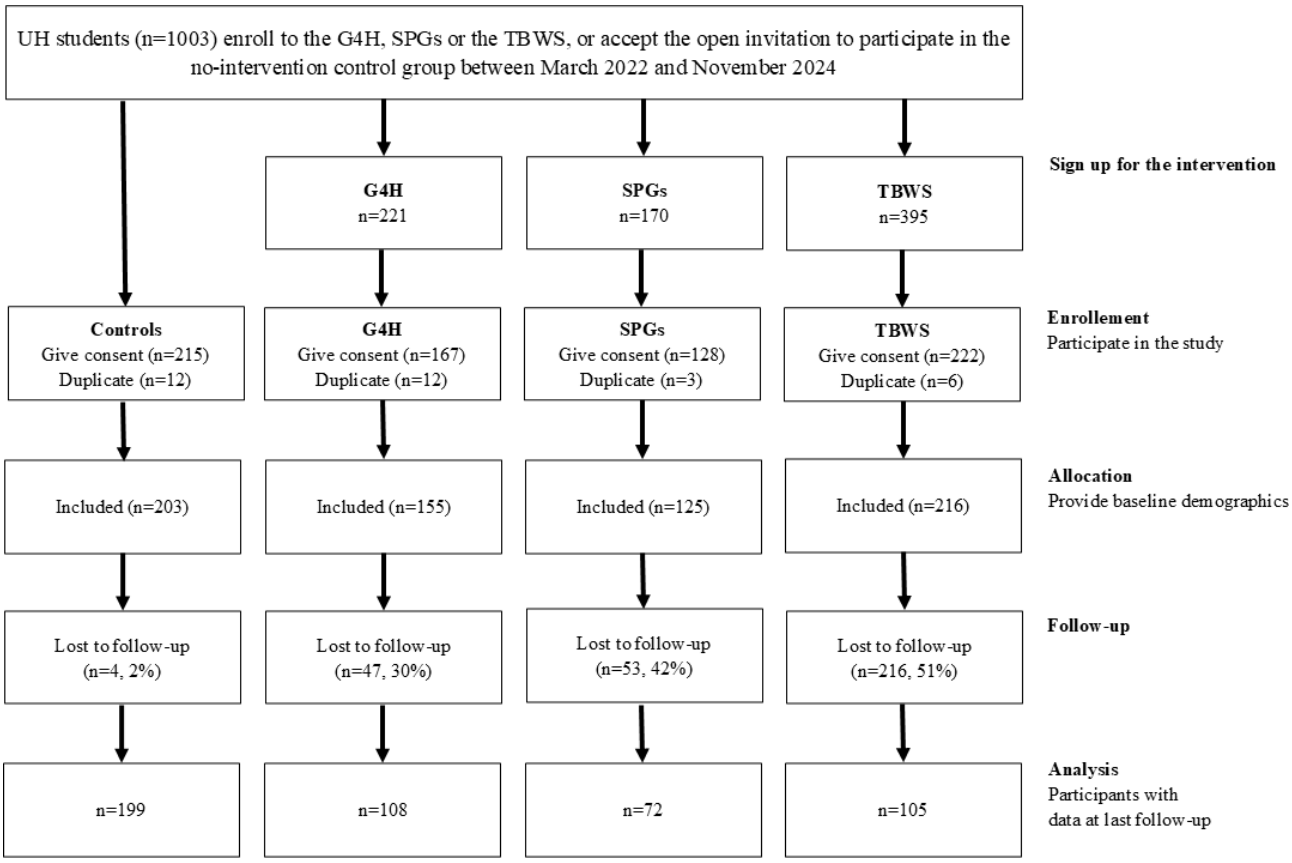

A high proportion of the university students enrolled in the interventions decided to participate in the study. The minimum target sample size (n=150) was attained for the G4H (n=155), TBWS (n=216), and control group (n=203). For the SPGs, the target sample size was not attained (n=125).

The anticipated study completion date for all data collection is by the end of the year 2028. Analyses for RQ1 and RQ3 are anticipated to take place in years 2025 and 2026, analyses for RQ2 in years 2026 to 2027, and analyses for RQ4 and RQ5 in years 2026 to 2028. First results from the study are anticipated to be published in the year 2026.

Attrition

Figure 3 illustrates the attrition rates by trial arms by the first data freeze in March 2025. The percentages do not accurately indicate the number of study dropouts, as all surveys are sent to participants regardless of nonresponse to previous surveys, allowing them to re-enter the study. The attrition rate has exceeded the anticipated 35% in the SPGs and TBWS. For the TBWS, the bars with black outlines represent the responses to the REDCap survey. The low response rate at the week 3 assessment (during intervention) is due to the survey only being available in REDCap and not integrated into the online course platform as in week 1 and week 7. Based on the attrition rate, we took action to increase retention, including adding incentives for the participants (a movie ticket for completing the surveys).

XSL•FO
RenderX

Figure 3. Attrition rates by trial arms. For TBWS, the smaller percentages at W1 and W7 denote the percentage of answers collected via REDCap. FU1: 1-month follow-up; FU3: 3-month follow-up; G4H: Groups 4 Health, SPG: Study psychologists' group intervention; TBWS: toward better well-being and studying intervention, Controls: no-intervention control group, UH: University of Helsinki; W1: week 1; W3: week 3; W7: week 7.

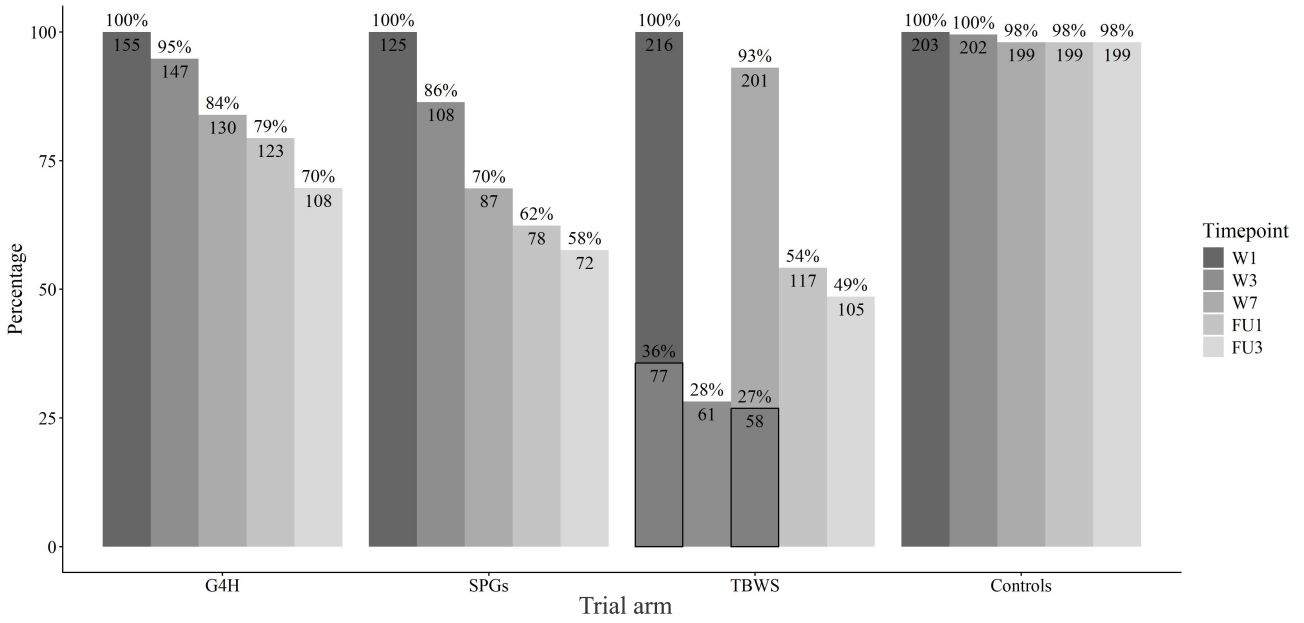

Participant Descriptive Statistics

Table 2 illustrates the participants' characteristics of the current sample at baseline by trial arm. In all groups, most participants are women (87/125, 70% - 175/203, 86%), not in a relationship (119/214, 56% - 107/154, 69%), and childless (175/208, 84% - 101/113, 89%). The age medians are lower than the age means ($median_{G4H}$ 25, IQR 22-29 y, $median_{SPGs}$ 27, IQR 25-31 y, $median_{TBWS}$ 24, IQR 22-31 y, and $median_{Control}$ 23 y, IQR 21-27 y).

XSL•FO
RenderX

Table . Characteristics of the study sample (n= 699) recruited between March 2022 and November 2024.

Characteristics	Trial arm				
	G4H[a] (n=155)	SPGs[b] (n=125)	TBWS[c] (n=216)	Controls[d] (n=203)	*P* value[e]
Age, mean (SD; range)	27.4 (8.0; 19-59)	28.9 (7.1; 20-53)	26.8 (7.0; 19-59)	25.1 (6.6; 18-58)	<.001
Missing, n	1	0	0	1	
Gender, n (%)					<.001
Woman	110 (71)	87 (70)	172 (80)	175 (86)	
Man	35 (23)	28 (22)	42 (19)	21 (10)	
Other	9 (6)	10 (8)	2 (1)	7 (3)	
Missing, n	1	0	0	0	
Relationship status, n (%)					.01
Single	107 (69)	78 (63)	119 (56)	130 (64)	
In a relationship	41 (27)	39 (31)	92 (43)	68 (33)	
Divorced or widowed	6 (4)	7 (6)	3 (1)	5 (2)	
Missing, n	1	1	2	0	
Living situation, n (%)					.05
Lives alone	79 (51)	57 (46)	86 (40)	95 (47)	
Lives with family	52 (34)	52 (42)	105 (49)	92 (46)	
Lives with others	23 (15)	16 (13)	23 (11)	15 (7)	
Missing, n	1	0	2	1	
Number of children, n (%)					.54
0	98 (88)	101 (89)	175 (84)	132 (88)	
≥1	14 (13)	12 (11)	33 (16)	18 (12)	
Missing, n	43	12	8	53	
Employment, n (%)					.08
Unemployed	87 (57)	58 (46)	113 (53)	90 (45)	
Employed	66 (43)	67 (54)	99 (47)	112 (55)	
Missing, n	2	0	4	1	
Gross income, n (%)					.07
<9 k €/year[f]	85 (56)	57 (46)	113 (55)	125 (62)	
9– - 18 k €/year[f]	40 (26)	46 (37)	51 (25)	49 (24)	
>18 k €/year[f]	28 (18)	21 (17)	43 (21)	29 (14)	
Missing, n	2	1	9	0	
Study years in current degree, mean (SD; range)	1.9 (1.6; 0 - 8)	4.6 (3.7; 0 - 23)	1.6 (2.3; 0 - 19)	1.8 (1.8; 0 - 10)	<.001
Missing, n	6	2	11	0	

[a]G4H: Groups 4 Health.

[b]SPGs: study psychologists' group interventions.

[c]TBWS: toward better well-being and studying intervention.

[d]Controls: no-intervention control group.

[e]*P* values are from tests comparing the trial arms with each other, using Kruskal-Wallis tests with continuous variables and χ^2 tests with categorical variables.

[f]1 €=US $1.08 in March 2025.

XSL•FO
RenderX

To determine whether the participants in the trial arms differed from one another in terms of background characteristics, we performed chi-square tests for categorical variables and Kruskal-Wallis tests for continuous variables (based on visual inspection, neither followed a normal distribution) (Table 2). Trial arms differed in age (Kruskal-Wallis χ^2_3=40.7, *P*<.001), gender (χ^2_6=5.4, *P*<.001), relationship status (χ^2_6=15.8, *P*=.02), living situation (χ^2_6=12.6, *P*=.05), and study years in current degree (Kruskal-Wallis χ^2_3=125.7, *P*<.001). These observations highlight the need to address group selection bias.

Missing Values Analyses

Item nonresponse (ie, a participant does not respond to an item within a questionnaire) was observed with missing data pattern plots across all groups. First, we examined the outcome variables which are assessed at all time points except week 3: loneliness, depression, mental well-being, general well-being, anxiety, social anxiety, multiple group memberships, and mood. Structural missingness (not all questionnaires were administered to all TBWS participants) was not included. Across all time points (excluding week 3), 94% (2174/2317) of survey responses were complete, 5% (105/2317) included a single missing item, and 2% (38/2317) included multiple missing items. Regarding the shorter survey taken at week 3 that includes the assessment of group memberships, mood, identification with the intervention group, loneliness, and working alliance, 96% (499/518) of responses were complete, 3% (16/518) included a single missing item, and 1% (3/518) included multiple missing items (0 - 7 missing items per survey). The level of missingness within given survey responses appears acceptable, especially when considering the length of the longer surveys (55 items, approximately 15 - 23 min).

Missingness was more prevalent for the selected background variables (Table 2). Of all survey responses of these selected variables, 79% (581/732) of responses were complete, 19% (141/732) included a single missing item, and 1% (10/732) included multiple missing items. Most often, participants left their number of children unreported. This may be due to a misunderstanding where some childless participants skipped this question instead of reporting zero. However, as the overall missingness in the selected background variables is 3% (172/5856), it is still within acceptable limits (<5%).

Discussion

Overview

Psychological ill-being and loneliness are highly prevalent among university students, often leading to long-lasting negative effects on mental health, social connectedness, and academic performance. To address this growing issue, accessible and preventive services at universities are needed. We investigate the effectiveness of the G4H intervention for university students' well-being using both quantitative and qualitative methods including 2 active comparators.

Principal Findings

We hypothesize that all the active interventions will improve students' well-being and that the decrease in loneliness will be more pronounced for the participants of the G4H group, in line with earlier studies [17,24,33,37]. We also anticipate valuable qualitative insights into the experiences of minority students. The multimethod approach will provide a comprehensive understanding of the intervention's effectiveness and the contextual factors influencing its impact.

Strengths and Limitations

A key strength of this study is its multifaceted approach, which allows for a comprehensive exploration of existing student well-being interventions at the UH. By involving multiple collaborators, each with unique research interests and expertise, the study can address a wide range of critical questions related to the interventions' impact. This collaborative framework enables the examination of various mediators and moderators, as well as the integration of both qualitative and quantitative methods. This perspective not only enriches the understanding of students' well-being but also ensures that these phenomena are not oversimplified. The multimethod design of the study allows participants to express their thoughts and experiences in more depth, enhancing participant engagement and potentially leading to richer data.

Furthermore, the qualitative insights gained from this approach will provide a better understanding of the intervention's effectiveness—or lack thereof—from the perspective of minority students. This insight is crucial for further developing the G4H intervention in terms of inclusion and equality. Moreover, the study is particularly relevant from an implementation point of view, as it will ensure the suitability and effectiveness of the method in Finland, addressing the significant need for early-stage support within the service system.

This study also has some limitations. The comparator arms (SPGs and TBWS) differ from the G4H slightly in terms of intensity (G4H: 5 sessions, TBWS: 6 sessions, and SPGs: 5 - 7 sessions) and delivery methods. This variability complicates direct comparisons between these groups and the G4H intervention, potentially introducing confounding factors that could influence the study's outcomes. Additionally, the varying data collection methods may introduce mode effects or measurement bias, as well as structural missingness, impacting the reliability and comparability of the data. To mitigate these concerns, we will conduct sensitivity analyses and discuss the implications of these methodological differences in our results.

Furthermore, the lack of exclusion criteria and the nonrandomized design of the protocol may introduce potential biases regarding group composition. However, allowing students to choose which intervention they participate in respects their autonomy and aligns with ethical research practices. This nonrandomized design increases the ecological validity of our findings by capturing the experiences of students in a real-world educational environment, which is important for making informed decisions that can improve services at both individual and population levels [80]. Ecological validity was further increased by including all students in the interventions and in

XSL•FO
RenderX

the study. We will implement strategies to address potential selection bias, including collecting detailed demographic and baseline data and using inverse propensity score weighting using double robust estimators, if necessary.

Due to the absence of comparable previous research with the same study population, a priori sample size calculations were not conducted. This represents a limitation, as underpowered samples may fail to detect treatment effects, while overpowered samples may lead to unnecessary use of resources. Nonetheless, the existing literature indicates that significant treatment effects can be observed with smaller sample sizes in different populations [17,24]. Given the complexity of our analysis plan, which includes mixed model analyses and potential (double robust) inverse propensity score weighting, we believe it was justified to aim for a larger target sample size in this study.

We did not appoint an independent data monitoring committee, as this is a naturalistic trial without randomization, and it investigates interventions that focus on psychosocial well-being rather than alleviating psychiatric pathologies. Moreover, there were no clinical inclusion criteria, and as the participants were university students, most showed rather low levels of psychological distress. Therefore, we anticipated minimal risk of harm from the interventions, a decision that was also accepted by the ethical review board of the study. However, we recognize that the presence of such a committee could have enhanced the integrity and trustworthiness of the data collection process.

The study protocol was registered retrospectively. This lack of prospective registration may affect the perceived transparency and credibility of the work. However, in this protocol paper, we have aimed to enhance transparency by providing a comprehensive and detailed description of the study design and methodology before analyzing and publishing the results.

Future Directions

These results can be expected to be generalizable across Finnish higher education, but caution is needed when extending the results beyond the higher education context or to institutions in other countries. Future studies should investigate the implementation and effectiveness of these interventions in different contexts, including other educational institutions in Finland and internationally, to enhance the understanding of their applicability and impact.

Dissemination

The trial results will be published in peer-reviewed scientific journals and scientific conferences. We follow the authorship eligibility guidelines of the International Committee of Medical Journal Editors [81]. We do not intend to use professional writers.

Conclusions

If proven effective, the interventions outlined in this protocol could be implemented in universities nationwide and have a significant impact in improving university students' well-being in both the short and long term, fostering positive mental health and supporting academic success and future career paths. These models may also be applicable in other settings, such as basic and secondary education, and in workplace environments. Ultimately, this research aims to support inclusion and equality in youth well-being services.

Acknowledgments

We thank our research assistant Hanna Kis for the invaluable help with the data collection and management. We thank all the students and intervention facilitators at the University of Helsinki for participating and investing their time in this study.

Funding

Primary sponsor and funder: The Strategic Research Council established within the Research Council of Finland, years 2022 - 2028 (grants Lahti 352700 and 372317, and Lammintakanen 352002 and 372321).

Funding

Source(s) of Monetary or Material Support: University of Helsinki, Finland.

Funding

A peer review of the protocol was not part of the funding process. The study sponsors and funders do not have any role or authority in the following: study design; collection, management, analysis, and interpretation of data; writing of the report; and the decision to submit the report for publication.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to protection of individual privacy but are available from the corresponding author on reasonable request.

Authors' Contributions

AL, EH, J Lahti, JP, MR, NK, OL, and VG contributed to the conceptualization and the design of the study. ER and J. Lammintakanen were responsible for conceptualization, study design, methodology, and writing, review, and editing of text relating to the cost-effectiveness study. PB and SV were responsible for the qualitative study design, qualitative data collection and analyses, as well as the original draft about all parts of the text relating to qualitative methods, data, and ethics. IM, J Lahti,

XSL•FO
RenderX

OK, SM, and VG participated in the original draft preparation, and SM coordinated the draft preparation and finalized the draft. IM, J. Lahti, OK, SM, and SS planned the quantitative analysis methods. AL, EH, HA, IM, J Lahti, JP, MR, OK, SM, and VG contributed to the quantitative data collection. IM, J Lahti, OK, SM, and VG are responsible for curation of the quantitative data. OK and SM planned and prepared the tables and figures. OK conducted all the descriptive analyses. IM, J. Lahti, and OK coordinated the implementation of the G4H intervention and facilitator training. J. Lahti and J. Lammintakanen contributed to funding acquisition. J Lahti is the principal investigator of the study. All authors participated in reviewing the work critically for important intellectual content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Details of all questionnaires and scales used in the study, including information on number of items, subscales, scoring, reliability, and validity.
[XLSX File, 23 KB - resprot_v15i1e79319_app1.xlsx]

Checklist 1
SPIRIT checklist.
[PDF File, 351 KB - resprot_v15i1e79319_app2.pdf]

References

1. Sheldon E, Simmonds-Buckley M, Bone C, et al. Prevalence and risk factors for mental health problems in university undergraduate students: a systematic review with meta-analysis. J Affect Disord 2021 May 15;287:282-292. [doi: 10.1016/j.jad.2021.03.054] [Medline: 33812241]
2. Parikka S, Holm N, Koskela T, Ikonen J, Kilpeläinen H. Health and well-being survey of higher education students 2021: research implementation and methods [Article in Finnish]. : National Institute for Health and Welfare THL; 2022 URL: https://www.julkari.fi/server/api/core/bitstreams/1086b6c3-2eab-4b20-a020-22942305f9c2/content [accessed 2025-12-31]
3. Duffy A, Keown-Stoneman C, Goodday S, et al. Predictors of mental health and academic outcomes in first-year university students: identifying prevention and early-intervention targets. BJPsych Open 2020 May 8;6(3):e46. [doi: 10.1192/bjo.2020.24] [Medline: 32381150]
4. Auerbach RP, Mortier P, Bruffaerts R, et al. WHO World Mental Health Surveys International College Student Project: prevalence and distribution of mental disorders. J Abnorm Psychol 2018 Oct;127(7):623-638. [doi: 10.1037/abn0000362] [Medline: 30211576]
5. Goosby BJ, Bellatorre A, Walsemann KM, Cheadle JE. Adolescent loneliness and health in early adulthood. Sociol Inq 2013 Nov 1;83(4):505-536. [doi: 10.1111/soin.12018] [Medline: 24187387]
6. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. Ann Behav Med 2010 Oct;40(2):218-227. [doi: 10.1007/s12160-010-9210-8] [Medline: 20652462]
7. Holt-Lunstad J, Smith TB, Layton JB. Social relationships and mortality risk: a meta-analytic review. PLoS Med 2010 Jul 27;7(7):e1000316. [doi: 10.1371/journal.pmed.1000316] [Medline: 20668659]
8. Walton GM, Cohen GL. A brief social-belonging intervention improves academic and health outcomes of minority students. Science 2011 Mar 18;331(6023):1447-1451. [doi: 10.1126/science.1198364] [Medline: 21415354]
9. Robbins SB, Oh IS, Le H, Button C. Intervention effects on college performance and retention as mediated by motivational, emotional, and social control factors: integrated meta-analytic path analyses. J Appl Psychol 2009 Sep;94(5):1163-1184. [doi: 10.1037/a0015738] [Medline: 19702363]
10. French B, Daley D, Groom M, Cassidy S. Risks associated with undiagnosed ADHD and/or autism: a mixed-method systematic review. J Atten Disord 2023 Oct;27(12):1393-1410. [doi: 10.1177/10870547231176862] [Medline: 37341291]
11. Kieseppä V, Holm M, Jokela M, Suvisaari J, Gissler M, Lehti V. Depression and anxiety disorders among immigrants living in Finland: comorbidity and mental health service use. J Affect Disord 2021 May 15;287:334-340. [doi: 10.1016/j.jad.2021.03.049] [Medline: 33813253]
12. Lehtonen J, Majlander S, Sares-Jäske L, Jehkoi A, Luopa P. Well-being of 8th and 9th graders belonging to gender and sexual minorities – school health survey 2019–2023 [Article in Finnish]. : THL; 2024 URL: https://www.julkari.fi/server/api/core/bitstreams/d4d0fbb3-e6c9-494d-a86a-4f524409a6c8/content [accessed 2025-12-31]
13. Settles IH, Buchanan NT. Multiple groups, multiple identities, and intersectionality. In: The Oxford Handbook of Multicultural Identity: Oxford University Press; 2014, Vol. 1:160-180.
14. Vargas SM, Huey SJ, Miranda J. A critical review of current evidence on multiple types of discrimination and mental health. Am J Orthopsychiatry 2020;90(3):374-390. [doi: 10.1037/ort0000441] [Medline: 31999138]

XSL•FO
RenderX

15. Peltola M, Moisio J. Voices and silence in service fields: a review of information on children and young people's service experiences [Article in Finnish]. : Finnish Youth Research Society; 2017 URL: https://edition.fi/nuorisotutkimusseura/catalog/view/300/236/814 [accessed 2025-12-31]
16. Green AE, Price MN, Dorison SH. Cumulative minority stress and suicide risk among LGBTQ youth. Am J Community Psychol 2022 Mar;69(1-2):157-168. [doi: 10.1002/ajcp.12553] [Medline: 34534356]
17. Haslam C, Cruwys T, Haslam SA, Dingle G, Chang MXL. Groups 4 Health: evidence that a social-identity intervention that builds and strengthens social group membership improves mental health. J Affect Disord 2016 Apr;194:188-195. [doi: 10.1016/j.jad.2016.01.010] [Medline: 26828756]
18. Haslam C, Jetten J, Cruwys T, Dingle GA, Haslam SA. The New Psychology of Health, 1st edition: Routledge; 2018. [doi: 10.4324/9781315648569]
19. Haslam SA, Jetten J, Postmes T, Haslam C. Social identity, health and well‐being: an emerging agenda for applied psychology. Applied Psychology 2009 Jan;58(1):1-23 [FREE Full text] [doi: 10.1111/j.1464-0597.2008.00379.x]
20. Iyer A, Jetten J, Tsivrikos D. Torn between identities: predictors of adjustment to identity change. In: Self Continuity: Individual and Collective Perspectives: Psychology Press; 2008:187-197.
21. Haslam C, Haslam SA, Jetten J, Cruwys T, Steffens NK. Life change, social identity, and health. Annu Rev Psychol 2021 Jan 4;72:635-661. [doi: 10.1146/annurev-psych-060120-111721] [Medline: 32886584]
22. Tajfel H, Tirner J. An integrative theory of intergroup conflict. In: Worchel S, Austin WG, editors. The Social Psychology of Intergroup Relations: Brooks/Cole; 1979:33-47.
23. Turner JC. Rediscovering the Social Group: A Self-Categorization Theory: Basil Blackwell; 1987.
24. Cruwys T, Haslam C, Rathbone JA, Williams E, Haslam SA, Walter ZC. Groups 4 Health versus cognitive-behavioural therapy for depression and loneliness in young people: randomised phase 3 non-inferiority trial with 12-month follow-up. Br J Psychiatry 2022 Mar;220(3):140-147. [doi: 10.1192/bjp.2021.128] [Medline: 35049477]
25. Baier AL, Kline AC, Feeny NC. Therapeutic alliance as a mediator of change: a systematic review and evaluation of research. Clin Psychol Rev 2020 Dec;82:101921. [doi: 10.1016/j.cpr.2020.101921] [Medline: 33069096]
26. Gergov V, Marttunen M, Lindberg N, Lipsanen J, Lahti J. Therapeutic alliance: a comparison study between adolescent patients and their therapists. Int J Environ Res Public Health 2021 Oct 26;18(21):11238. [doi: 10.3390/ijerph182111238] [Medline: 34769766]
27. Delgadillo J, Branson A, Kellett S, Myles-Hooton P, Hardy GE, Shafran R. Therapist personality traits as predictors of psychological treatment outcomes. Psychother Res 2020 Sep;30(7):857-870. [doi: 10.1080/10503307.2020.1731927] [Medline: 32089100]
28. Heinonen E, Lindfors O, Laaksonen MA, Knekt P. Therapists' professional and personal characteristics as predictors of outcome in short- and long-term psychotherapy. J Affect Disord 2012 May;138(3):301-312. [doi: 10.1016/j.jad.2012.01.023] [Medline: 22335889]
29. Heinonen E, Lindfors O, Härkänen T, Virtala E, Jääskeläinen T, Knekt P. Therapists' professional and personal characteristics as predictors of working alliance in short-term and long-term psychotherapies. Clin Psychol Psychother 2014;21(6):475-494. [doi: 10.1002/cpp.1852] [Medline: 23813617]
30. Ryan RM, Deci EL. Self-Determination Theory: Basic Psychological Needs in Motivation, Development, and Wellness: Guilford Publications; 2017.
31. Drieschner KH, Lammers SMM, van der Staak CPF. Treatment motivation: an attempt for clarification of an ambiguous concept. Clin Psychol Rev 2004 Jan;23(8):1115-1137. [doi: 10.1016/j.cpr.2003.09.003] [Medline: 14729425]
32. Haslam C, Cruwys T, Bentley S, Haslam AS, Dingle GA, Jetten J. GROUPS 4 HEALTH Facilitator Manual (Version 40): The University of Queensland; 2021.
33. Haslam C, Cruwys T, Chang MXL, et al. GROUPS 4 HEALTH reduces loneliness and social anxiety in adults with psychological distress: findings from a randomized controlled trial. J Consult Clin Psychol 2019 Sep;87(9):787-801. [doi: 10.1037/ccp0000427] [Medline: 31403815]
34. Asikainen H, Katajavuori N. Development of a web-based intervention course to promote students' well-being and studying in universities: protocol for an experimental study design. JMIR Res Protoc 2021 Mar 9;10(3):e23613. [doi: 10.2196/23613] [Medline: 33687336]
35. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006 Jan;44(1):1-25. [doi: 10.1016/j.brat.2005.06.006] [Medline: 16300724]
36. Howell AJ, Passmore HA. Acceptance and commitment training (ACT) as a positive psychological intervention: a systematic review and initial meta-analysis regarding ACT's role in well-being promotion among university students. J Happiness Stud 2019 Aug;20(6):1995-2010. [doi: 10.1007/s10902-018-0027-7]
37. Katajavuori N, Vehkalahti K, Asikainen H. Promoting university students' well-being and studying with an acceptance and commitment therapy (ACT)-based intervention. Curr Psychol 2023 Feb;42(6):4900-4912. [doi: 10.1007/s12144-021-01837-x]
38. Räihä K, Katajavuori N, Vehkalahti K, Huotilainen M, Asikainen H. University students' stress and burnout risk: results of an ACT-based online-course using self-assessments and HRV-measurements. Curr Psychol 2024 Jun;43(22):20182-20195. [doi: 10.1007/s12144-024-05800-4]

XSL•FO
RenderX

39. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
40. Cohen J. Statistical Power Analysis for the Behavioral Sciences, 2nd edition: Lawrence Erlbaum Associates, Inc; 1988. [doi: 10.4135/9781071812082]
41. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
42. Hays RD, DiMatteo MR. A short-form measure of loneliness. J Pers Assess 1987;51(1):69-81. [doi: 10.1207/s15327752jpa5101_6] [Medline: 3572711]
43. Russell D, Peplau LA, Cutrona CE. The revised UCLA Loneliness Scale: concurrent and discriminant validity evidence. J Pers Soc Psychol 1980 Sep;39(3):472-480. [doi: 10.1037//0022-3514.39.3.472] [Medline: 7431205]
44. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
45. Haslam C, Holme A, Haslam SA, Iyer A, Jetten J, Williams WH. Maintaining group memberships: social identity continuity predicts well-being after stroke. Neuropsychol Rehabil 2008;18(5-6):671-691. [doi: 10.1080/09602010701643449] [Medline: 18924001]
46. Shah N, Cader M, Andrews B, McCabe R, Stewart-Brown SL. Short Warwick-Edinburgh Mental Well-being Scale (SWEMWBS): performance in a clinical sample in relation to PHQ-9 and GAD-7. Health Qual Life Outcomes 2021 Nov 24;19(1):260. [doi: 10.1186/s12955-021-01882-x] [Medline: 34819104]
47. Evans C, Connell J, Audin K, Sinclair A, Barkham M. Rationale and development of a general population well-being measure: psychometric status of the GP-CORE in a student sample. Br J Guid Counc 2005 May;33(2):153-173. [doi: 10.1080/03069880500132581]
48. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
49. Connor KM, Kobak KA, Churchill LE, Katzelnick D, Davidson JRT. Mini-SPIN: a brief screening assessment for generalized social anxiety disorder. Depress Anxiety 2001;14(2):137-140. [doi: 10.1002/da.1055] [Medline: 11668666]
50. The EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [doi: 10.1016/0168-8510(90)90421-9]
51. Postmes T, Haslam SA, Jans L. A single‐item measure of social identification: reliability, validity, and utility. British J Social Psychol 2013 Dec;52(4):597-617 [FREE Full text] [doi: 10.1111/bjso.12006] [Medline: 23421003]
52. Hatcher RL, Gillaspy JA. Development and validation of a revised short version of the Working Alliance Inventory. Psychother Res 2006 Jan;16(1):12-25. [doi: 10.1080/10503300500352500]
53. Hatcher RL, Lindqvist K, Falkenström F. Psychometric evaluation of the Working Alliance Inventory-Therapist version: current and new short forms. Psychother Res 2020 Jul;30(6):706-717. [doi: 10.1080/10503307.2019.1677964] [Medline: 31621525]
54. Sarason IG, Sarason BR, Shearin EN, Pierce GR. A brief measure of social support: practical and theoretical implications. J Soc Pers Relat 1987 Nov;4(4):497-510. [doi: 10.1177/0265407587044007]
55. Guay F, Vallerand RJ, Blanchard C. On the assessment of situational intrinsic and extrinsic motivation: the Situational Motivation Scale (SIMS). Motiv Emot 2000 Sep;24(3):175-213. [doi: 10.1023/A:1005614228250]
56. Miller WR, Rollnick S. Motivational Interviewing: Helping People Change, 3rd edition: Guilford Publications; 2012.
57. Parpala A, Lindblom-Ylänne S. Using a research instrument for developing quality at the university. Quality in Higher Education 2012 Nov;18(3):313-328. [doi: 10.1080/13538322.2012.733493]
58. Konstabel K, Lönnqvist JE, Leikas S, et al. Measuring single constructs by single items: constructing an even shorter version of the “Short Five” personality inventory. PLoS ONE 2017;12(8):e0182714. [doi: 10.1371/journal.pone.0182714] [Medline: 28800630]
59. Wei M, Russell DW, Mallinckrodt B, Vogel DL. The Experiences in Close Relationship Scale (ECR)-short form: reliability, validity, and factor structure. J Pers Assess 2007 Apr;88(2):187-204. [doi: 10.1080/00223890701268041] [Medline: 17437384]
60. Heinonen E, Orlinsky DE, Willutzki U, et al. Psychotherapist trainees’ quality of life: patterns and correlates. Front Psychol 2022;13:864691. [doi: 10.3389/fpsyg.2022.864691] [Medline: 35401345]
61. Hartmann A, Orlinsky DE, Rønnestad MH, Willutzki U, Schröder TA, Heinonen E. Measuring psychotherapist functioning with the Therapist Work Involvement Scales (TWIS): reliability, factor structure, and measurement invariance. Psychother Res 2025 Nov;35(8):1287-1299. [doi: 10.1080/10503307.2024.2389857] [Medline: 39159177]
62. Mäklin S, Kokko P. Unit costs of health and social care in Finland in 2017 [Article in Finnish]. : THL; 2021 URL: https://www.julkari.fi/server/api/core/bitstreams/5268c80a-cf96-473b-b525-679e7bf02905/content [accessed 2025-12-31]
63. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. J Stat Softw 2011;45(3):1-67. [doi: 10.18637/jss.v045.i03]
64. Kurz CF. Augmented inverse probability weighting and the double robustness property. Med Decis Making 2022 Feb;42(2):156-167. [doi: 10.1177/0272989X211027181] [Medline: 34225519]

XSL•FO
RenderX

65. Wager S. Doubly robust methods. In: Causal Inference: A Statistical Learning Approach: Stanford University; 2025.
66. Galwey NW. A new measure of the effective number of tests, a practical tool for comparing families of non-independent significance tests. Genet Epidemiol 2009 Nov;33(7):559-568 [FREE Full text] [doi: 10.1002/gepi.20408] [Medline: 19217024]
67. Sun S, Chuang LH, Sahlén KG, Lindholm L, Norström F. Estimating a social value set for EQ-5D-5L in Sweden. Health Qual Life Outcomes 2022 Dec 23;20(1):167. [doi: 10.1186/s12955-022-02083-w] [Medline: 36564844]
68. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. Pharmacoeconomics 2014 Dec;32(12):1157-1170. [doi: 10.1007/s40273-014-0193-3] [Medline: 25069632]
69. Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. Health Econ 2002 Jul;11(5):415-430. [doi: 10.1002/hec.678] [Medline: 12112491]
70. Fraser G, Brady A, Wilson MS. Mental health support experiences of Rainbow Rangatahi youth in Aotearoa New Zealand: results from a co-designed online survey. J R Soc N Z 2022;52(4):472-489. [doi: 10.1080/03036758.2022.2061019] [Medline: 39440320]
71. McDermott E, Eastham R, Hughes E, et al. “What Works” to support LGBTQ+ young people’s mental health: an intersectional youth rights approach. Int J Soc Determinants Health Health Serv 2024 Apr;54(2):108-120. [doi: 10.1177/27551938241230766] [Medline: 38385356]
72. Mailander S, Lehtonen J, Luopa P, et al. Rainbow youth in times of emergency. In: Kekkonen M, Gissler M, Känkänen P, Isola AM, editors. Exceptional Youth in the Time of Corona: Youth Living Conditions Yearbook 2022 [Book in Finnish]: Youth Research Society Youth Research Network; 2022:179-196.
73. Nori H, Lyytinen A, Juusola H, Kohtamäki V, Kivistö J. Marginal groups in higher education: applying to study, study experiences and future plans [Article in Finnish]. : Ministry of Education and Culture; 2021 URL: https://julkaisut.valtioneuvosto.fi/server/api/core/bitstreams/b73fe053-dd66-4c57-968d-908b38a6708b/content [accessed 2025-12-31]
74. Values and attitudes of University of Applied Students 2022 [Article in Finnish]. : SAMOK; 2022 URL: https://samok.fi/wp-content/uploads/2022/11/samok_arvot-ja-asenteet.pdf [accessed 2025-12-31]
75. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychol Bull 2003 Sep;129(5):674-697. [doi: 10.1037/0033-2909.129.5.674] [Medline: 12956539]
76. Collins PH. Intersectionality as Critical Social Theory: Duke University Press; 2019. [doi: 10.2307/j.ctv11hpkdj]
77. Moradi B, Grzanka PR. Using intersectionality responsibly: toward critical epistemology, structural analysis, and social justice activism. J Couns Psychol 2017 Oct;64(5):500-513. [doi: 10.1037/cou0000203] [Medline: 29048196]
78. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
79. Braun V, Clarke V. Thematic analysis: a practical guide 2021. [doi: 10.1007/978-3-319-69909-7_3470-2]
80. Singal AG, Higgins PDR, Waljee AK. A primer on effectiveness and efficacy trials. Clin Transl Gastroenterol 2014 Jan 2;5(1):e45. [doi: 10.1038/ctg.2013.13] [Medline: 24384867]
81. Defining the role of authors and contributors. International Committee of Medical Journal Editors. URL: https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html [accessed 2025-12-31]

Abbreviations

CBT: cognitive behavioral therapy
ECTS: European Credit Transfer and Accumulation System
EQ-5D-5L: EuroQol - 5 dimension - 5 level questionnaire
G4H: Groups 4 Health
GDPR: General Data Protection Regulation
GPA: grade point average
LGBTIQ+: lesbian, gay, bisexual, transgender, intersex, queer, and other
REDCap: Research Electronic Data Capture
RQ: research question
SIAH: social identity approach to health
SIMIC: social identity model of identity change
SPG: study psychologists’ group intervention
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TBWS: toward better well-being and studying intervention
UH: University of Helsinki

XSL•FO
RenderX

Edited by J Sarvestan; submitted 24.Jun.2025; peer-reviewed by AL Sachs, R Chigevenga; revised version received 11.Nov.2025; accepted 13.Nov.2025; published 23.Jan.2026.

Please cite as:
Martikainen S, Karhunen O, Mankki I, Gergov V, Berg P, Venäläinen S, Rissanen E, Lammintakanen J, Heinonen E, Lehtinen A, Pöntinen J, Asikainen H, Katajavuori N, Linnaranta O, Räisänen M, Salminen S, Lahti J
Interventions for Students' Well-Being at the University of Helsinki (INSIGHT): Protocol and Preliminary Descriptive Results for a Quasi-Experimental Controlled Trial of a Social Identity Intervention and Two Active Comparators
JMIR Res Protoc 2026;15:e79319
URL: https://www.researchprotocols.org/2026/1/e79319
doi:10.2196/79319

© Silja Martikainen, Oona Karhunen, Iiris Mankki, Vera Gergov, Päivi Berg, Satu Venäläinen, Elisa Rissanen, Johanna Lammintakanen, Erkki Heinonen, Anu Lehtinen, Janna Pöntinen, Henna Asikainen, Nina Katajavuori, Outi Linnaranta, Milla Räisänen, Samuli Salminen, Jari Lahti. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Investigating the Role of the Environment on Physical Activity Interventions (the InSPACE Project): Protocol for a Pooled Secondary Analysis of Randomized Controlled Trials

Amy J Youngbloom[1], MPH; Maya Rowland[2], MPH; Stephen J Mooney[1], PhD; Kaitlin Singer[1], MS; Adam Szpiro[3], PhD; Brian E Saelens[2,4], PhD

[1]Department of Epidemiology, School of Public Health, University of Washington, 3980 15th Ave NE, Seattle, WA, United States

[2]Center for Child Health, Behavior and Development, Seattle Children's Research Institute, Seattle, WA, United States

[3]Department of Biostatistics, School of Public Health, University of Washington, Seattle, WA, United States

[4]Department of Pediatrics, School of Medicine, University of Washington, Seattle, WA, United States

Corresponding Author:
Amy J Youngbloom, MPH
Department of Epidemiology, School of Public Health, University of Washington, 3980 15th Ave NE, Seattle, WA, United States

Abstract

Background: Physical activity (PA) interventions can increase levels of PA to help participants meet recommended levels. The impact of PA interventions may be affected by an individual's neighborhood environment, including attributes such as walkability, crime rates, or greenspace availability, but research to date has lacked the power and geographic spread to adequately assess the role of the environment.

Objective: The Interventions Supporting Physical Activity and the Environment (InSPACE) study used the Automatic Context Measurement Tool (ACMT) to gather environmental measures for participants around their home address in completed lifestyle intervention trials across the United States, then pooled and harmonized demographic and device-based activity data, creating a dataset for use in assessing the moderation effect of neighborhood attributes on interventions to increase PA.

Methods: PA intervention trials were recruited from across the United States, and trialists were instructed in the use of the ACMT to geocode and collect prespecified environmental measures. The InSPACE research team gathered deidentified data from trialists, including demographics, raw accelerometry data, and ACMT-generated environmental measures and harmonized data to create a pooled dataset of PA intervention trial participants.

Results: As of August 2025, a total of 39 PA intervention trials have been recruited and data from 31 of these trials have been processed and harmonized, creating a current pooled dataset of 4471 participants with any harmonized data, of whom 4360 (97.5%) have linked environmental data and 2208 (49.4%) have specified (3 days of at least 8 hours per day) accelerometry data for both baseline and postintervention. Results from the primary analysis for the InSPACE project are expected to be published in late 2026.

Conclusion: InSPACE will contribute to understanding the role of the environment in moderating the effect of interventions to increase PA. The protocols and processes of InSPACE can inform future projects in pooled data harmonization and analysis.

International Registered Report Identifier (IRRID): DERR1-10.2196/83151

(JMIR Res Protoc 2026;15:e83151) doi:10.2196/83151

KEYWORDS

physical activity intervention; built environment; pooled analysis; physical activity; interventions; social environment; accelerometry

Introduction

Background

Physical activity (PA) is linked to metabolic health, diabetes, cardiovascular disease, and some types of cancer [1-4]. Getting adequate daily PA may also improve cognitive function and bone strength, reduce depression, and improve overall well-being [2,5,6]. Most adults and youth in the United States do not meet the recommended guidelines for PA [7,8].

PA interventions have the potential to increase PA in youth and adults. Such interventions can include goal setting, supervised instructor-led exercise, educational sessions and information sharing, or peer support [9-17]. Meta-analyses have shown a positive mean effect of PA interventions, but the effect size varies significantly between types of interventions and population attributes [18-20]. Researchers have examined some

XSL•FO
RenderX

potential modifiers of PA intervention effectiveness, including participant demographics and intervention type, but there has yet to be an extensive examination of environmental attributes as a moderator of PA intervention effectiveness, in part due to the lack of adequate data for such a study.

An individual's neighborhood environment may affect their level of PA [21-23]. For example, a more walkable area around one's home, as measured by land use, street connectivity and density, and greater access to sidewalks around one's home is associated with more time spent in PA, specifically walking [24,25]. Living near a park or other urban green space can also promote more PA, especially for higher-income residents [26-28]. Physical disorder, such as vacant homes, litter, or vandalism, may reduce PA levels among residents [29]. Despite existing research on neighborhood environment and PA, and the growing prevalence of ecological models of behavior change that consider multiple levels of influence, there remains a gap in research examining how the neighborhood environment may influence the effectiveness of PA intervention outcomes.

A few studies have considered the moderating role of neighborhood attributes on PA intervention effectiveness, but these have been limited to perceptions of walkability and the National Walkability score and have shown mixed results [30-32]. These studies were significantly limited by small sample sizes, limited access to objective neighborhood attribute measures, and lack of geographic and demographic heterogeneity. The lack of substantial research in this area emphasizes the need for a large, diverse pooled dataset of PA intervention trials with standardized environmental measures that can be more rigorously examined for their potential modifying effect. The goal of the Interventions Supporting Physical Activity and the Environment (InSPACE) project was to help fill this gap and create opportunities for research on the modifying effect of the built environment on PA intervention effectiveness.

Objective

The InSPACE study seeks to assess the modifying effect of the neighborhood environment on the impact of interventions in increasing PA by pooling and analyzing data from completed PA intervention trials of varied geographies, contexts, and populations. The InSPACE team provided and instructed trialists to use the Automatic Context Measurement Tool (ACMT) to geocode participant address data and generate environmental data specific to the neighborhood around the home address of pooled trial participants [33]. This paper describes the protocols used to gather and harmonize demographic and accelerometry data across trials in the InSPACE study and the use of the ACMT to estimate residential environment measures for participants in InSPACE trials. These protocols may not only guide future research on environmental effect modification of PA intervention effectiveness, but also more broadly inform future pooling studies examining environmental, individual demographic, and device-based measurement of PA.

Methods

Scientific Advisory Council

The InSPACE study team is advised by a scientific advisory council (SAC) comprised of 7 PA intervention research experts who provided ongoing feedback on trial inclusion criteria, trial recruitment strategies, and data harmonization processes. SAC members also shared expertise regarding intervention research (in PA as well as other health behaviors) in diverse populations, including across income, race and ethnicity, urban and rural geography, and among youth and older adults. Finally, SAC members assisted with recruitment by reaching out to potential trialists in their professional networks and informed individual-level demographic variable harmonization.

Inclusion Criteria

Initial trial inclusion criteria required trials to have at least 50 free-living participants 6 years of age and older, and have targeted increasing PA as part of an intervention. For data collection and harmonization purposes, trials must have collected PA data using accelerometers or pedometers, retained individual participant address data, linked participant demographic and PA data, and have participants in the United States. We included only trials conducted fully or mostly between 2010 and 2020 to align better with the timing of the available environment data and to avoid potential impacts from the COVID-19 pandemic. The intervention must have had an anticipated effect on PA, although PA did not necessarily have to be the primary or only target of change. For example, if a trialist hypothesized that an increase in PA would improve mental health and therefore developed a trial intervention that aimed to increase participants' PA to improve their mental health, this trial would be eligible for InSPACE, provided all other inclusion criteria were met. Trials must have also had a control condition, expected at the time of trial design to be less effective in promoting PA than the intervention. The InSPACE investigator team examined the plausibility that the intervention could be affected by the neighborhood environment before confirming the trial eligibility for inclusion. All trials must have collected PA data for at least two time points: (1) at least once prior to the intervention and (2) at least once after the intervention was complete (postintervention), such that any activity occurring as part of the intervention (eg, an exercise class facilitated by the research team to which participants were encouraged to attend) was not included in either activity collection. Initial criteria excluded trials with any PA intervention facilitated by the intervention team or that took place at a prescribed location (eg, in a fitness center).

After the initial trial recruitment and in consultation with the SAC, eligibility criteria were expanded to include interventions with intervention team–facilitated PA interventions at prescribed locations, to reflect an ecological theory of behavior change that any PA may increase other self-initiated community-based PA as well, particularly if the community environment supports PA [34]. Additionally, criteria were expanded to include trials with control groups that were not inert but were different from the intervention in that the trialists hypothesized there would be differences in PA between intervention and controls at

XSL•FO
RenderX

postintervention. Ultimately, if the trialist hypothesized that the intervention group would have a greater increase in PA than the control group, the trial was eligible regardless of the amount of intervention team–facilitated PA within the intervention, or the content or intensity of the control group. With the inclusion of intervention-team or instructor-led PA trials, it was important to ensure that postintervention accelerometry data did not overlap with any intervention-team or instructor-led PA.

Our final trial inclusion change, adopted to broaden our trial population after an initial trial recruitment pass exhausted our trial list, was to lower the minimum number of participants per trial from 50 to 30. The full list of the final inclusion and exclusion criteria, as well as changes that were made in inclusion or exclusion criteria, is listed in Table 1. As the inclusion criteria evolved, literature was rereviewed and investigator teams from newly eligible trials were contacted.

XSL•FO
RenderX

Table . Interventions Supporting Physical Activity and the Environment trial inclusion and exclusion criteria.

Inclusion criteria category	Final inclusion and exclusion criteria details	Changes to inclusion and exclusion criteria
Population	• Inclusion: Free-living humans in the contiguous United States with an available geocodable home address who are 6 years of age or older. • Exclusion: Animal trials or populations that reside outside of the contiguous United States.	• Original criteria did not specify the contiguous United States. This was changed as we realized some of our national measures do not include data for AK[a] and HI[b].
Sample size	• Inclusion: Trial must have enrolled 30 or more participants (or 30 or more dyads). • Exclusion: Trials with fewer than 30 participants (or fewer than 30 dyads) enrolled.	• Original criteria only included trials with 50 or more participants (or 50 or more dyads). After exhausting the list of trials with 50 or more, we adjusted our criteria to include smaller trials.
Interventions	• Inclusion: (1) PA[c] is targeted for increases, regardless of whether PA is one of the primary aims of the trial. (2) Focus is on manipulating PA through variation in intervention delivery, content, and other factors, including those that target other issues or health outcomes and include PA. (3) Interventions can include some intervention-team or instructor-led PA that participants are able to, or are expected to attend. • Exclusion: Intervention is very specific to single systems or parts of the body.	• Trials that included any investigative team or instructor-led PA were not originally included and were added at the encouragement of our SAC[d].
Trial design	• Inclusion: (1) There should be an expectation of superiority for the intervention arm and NOT just noninferiority. (2) There is an expectation that the delta in PA between study arms could be influenced by the neighborhood environment. (3) Trial must be a randomized controlled trial. (4) Some control arm that is not only a different dose of prescribed physical activity.	• Original criteria limited to inert control groups (waitlists, no-contact, etc).
Physical activity measurement timing	• Inclusion: Free-living PA both before and after the intervention, and in any follow-ups after the intervention. Okay if PA measurement overlaps with some intervention time.[e] • Inclusion: (1) PA outcome is measured outside of intervention team–facilitated PA time both at baseline, post, and other follow-ups.[f] (2) Measurement captures self-initiated activity that might be in the community.[f] • Exclusion: PA is measured only within specific intervention settings.[f]	• Trials with instructor-led PA were not initially included. They were later included at the encouragement of our SAC only if the timing of their PA measurement met criteria.
Data collected and available	• Inclusion: (1) Device measured PA- Accelerometer measured PA OR Pedometer measured PA or steps if collected data per minute (eg, not by week or day) AND in adults 18 years or older. (2) Address data. (3) Minimum demographic data (age, race, and sex). • Exclusion: Self-report PA only. (2) Trials using pedometer measurement in those younger than 18 years or pedometer trials with steps aggregated by day or larger unit. Data are no longer available.	• Initially excluded trials that assessed PA via pedometer.

XSL•FO
RenderX

Inclusion criteria category	Final inclusion and exclusion criteria details	Changes to inclusion and exclusion criteria
Trial dates	• Inclusion: Intervention began in 2010 or later and mostly completed pre-COVID-19. • Exclusion: (1) Intervention began prior to 2010. (2) Participants were still receiving the intervention during COVID-19. Follow-ups after March 2020 allowed.	• Did not change.

[a]AK: Alaska.

[b]HI: Hawaii.

[c]PA: physical activity.

[d]SAC: scientific advisory council.

[e]Outcomes for trials without an instructor-led component.

[f]Outcomes for trials with an instructor-led component.

Trial Recruitment

The InSPACE research team used the National Institutes of Health (NIH) eReporter and ClinicalTrials.gov websites, as well as word-of-mouth, to identify potentially eligible intervention trials. ClinicalTrials.gov searches were conducted by Seattle Children's librarians and reviewed by the InSPACE team. The principal investigators or leads of trials considered potentially eligible by the InSPACE team were contacted by the InSPACE principal investigator by email. The initial email included a one-page document describing the study and a brief introduction. If other members of the investigator team or SAC had previously collaborated with the trialist, these colleagues were CC'd on the introduction email and asked to "bump" the initial contact by responding and prompting the trialist to consider learning more about and engaging in InSPACE. If trialists responded with interest, basic eligibility questions were asked via email (eg, whether the participant's home address and accelerometry data were still available). If still eligible, a conference call was organized to further describe InSPACE participation and further screen for eligibility and interest. Thereafter, trial information was brought back to the investigative team for review and to confirm eligibility. After confirmation, trialists were reapproached with a participation inclusion offer.

Offers included a contract to be established between the trialist's institution and Seattle Children's Research Institute. These contracts typically established a 6-month contract period and paid US $6000 in direct costs plus facilities and administrative costs for the first trial and US $2000 in direct costs for each additional trial for teams who contributed data from multiple trials. Contracts were milestone-based (eg, contract in place, Institutional Review Board [IRB] approval, ACMT training completed, and data transferred). Because of the unusual nature of these small subawards, the Seattle Children's contracts team offered meetings with partner institutions early in the contracting process to answer questions ahead of time and prevent avoidable delays.

Measures

Measures of Environmental Context

Participating trialists worked with the InSPACE team to generate residential environmental measures for each trial participant. This measurement collaboration leveraged the ACMT, an open-source tool for linking environmental data to any point in the United States (eg, participant addresses for InSPACE) that was provided to each trial team. The ACMT includes an RShiny-driven user-friendly interface that does not require prior geospatial software experience to link environmental measures. The ACMT's geocoding and linkage occur locally such that trial teams can generate and share the measures without sharing personally identifiable participant data with the InSPACE team or a third-party geocoding site [35]. The ACMT is described more fully elsewhere [33].

To run the ACMT, a user must have administrative privileges on their computer. In some cases, our trialist partners were not administrators on their computers or encountered institutional firewalls that prevented them from installing and running the ACMT without IT assistance. Additionally, because geocoding and data linkage involve large geographic files, ACMT-driven linkage often requires considerable disk space and memory resources. While requirements depended on the number of state files that needed to be downloaded, in most cases, 100 GB of available disk space and 8 GB of memory were adequate for the ACMT to function properly on trialists' computers. Despite challenges in working through firewalls and administrator privileges, none of the recruited trials dropped out of the InSPACE study due to technical issues with the ACMT.

Trialists who had not already used their own geocoding systems to identify participants' home locations (this was rare) used the geocoder embedded in the ACMT to generate the latitude and longitude values for each participant's address. Trialists who used the ACMT geocoder were asked to update addresses that received geocode ratings higher than 0 (the PostGIS geocoder embedded in the ACMT uses 0 to indicate the highest level of confidence in the accuracy of the geocode; Corti et al [36]). To improve the geocoding accuracy, trialists used the ACMT to edit and improve the address text, including removing apartment numbers, fixing obvious typos, and spelling out street names and cardinal directions. If these initial edits did not improve the geocode to a rating of 0, the ACMT mapping tool provided the user a way to view geocoded locations on an OpenStreet map (embedded within the ACMT so as not to reveal addresses to third parties) and to document a note as to whether the geocode appeared accurate (mapped to the correct street and city) or not. This geocode checking process helped improve the quality of

XSL•FO
RenderX

the geocodes and deterred trialists from using web-based third-party mapping tools to verify addresses and geocodes.

Next, trialists used the ACMT to link environmental data from the datasets listed in the table below (Table 2) to the geocoded home addresses. For most datasets, only 1 year of data was available. Notable exceptions include American Community Survey (ACS) 5-year estimates, for which data are provided in 5-year increments, starting with 2005‐2009. The National Land Cover Database (NLCD) includes data every 2‐3 years from 2004 through 2019, regional price parity was available for all years between 2008 and 2018, and gentrification was available for 2000 (indicating gentrification from 1990 through 2000) and 2010 (indicating gentrification status from 2000 to 2010). For datasets with multiple years of data, we chose 2015 as the a priori estimated midpoint of enrollment years for all recruited trials and linked environmental data for the year of available data closest to 2015 (ie, default year for NLCD data was 2016—the closest to 2015 of the available years of data; for ACS this meant pulling the 5-year span of data with a midpoint of 2015 [2013‐2017]). In addition to the default year(s) of data, trialists were asked to pull environmental data that most closely aligned with the trial's first year of enrollment. For example, if a trial enrolled participants from 2018to 2019, they would pull data for ACS for both the 5-year span of data with a midpoint of 2015 (ACS 5-year data for 2014‐2018; ie, the default year for environmental data), as well as the 5-year span whose midpoint most closely aligned with their enrollment (ACS 5-year data for 2016‐2020; ie, ACS data with a midpoint of 2018). The same trial would also pull NLCD data for both 2016 and 2019 (data for 2018 are not available for NLCD), CDC (Centers for Disease Control and Prevention) Places data for 2017 (closest available year to the default year of 2015) and 2018, and Regional Price Parity data for 2015 and 2018.

XSL•FO
RenderX

Table . Environmental measure datasets pulled for Interventions Supporting Physical Activity and the Environment trials.

Dataset	Description	Years available	Data summary method
American Community Survey [37]	Five-year pooled sociodemographic census data.	2009 through 2021 (default 2013 - 2017)	Radial buffer interpolation (500 m, 1000 m, 5000 m)
Modified Retail Food Environment Index [38]	Census tract estimates of the ratio of healthy food outlets to total food outlets.	2011[a]	Radial buffer interpolation (500 m, 1000 m, 5000 m)
National Walkability Index [39]	Block group level ranking (1 to 20) of walkability, based on street intersection density, proximity to transit stops, and diversity of land use.	2019[a]	Radial buffer interpolation (500 m, 1000 m, 5000 m)
CDC Places [40]	Census tract-level estimates of population behaviors taken from survey responses to BRFSS.[b]	2017[a], 2018	Radial buffer interpolation (500 m, 1000 m, 5000 m)
National Land Cover Database [41]	Satellite imagery is used to create pixel-by-pixel maps of land cover features, including forest, developed areas, open space, etc.	2004, 2006, 2008, 2011, 2013, 2016[a], 2019	Radial buffer interpolation (500 m, 1000 m, 5000 m)
ParkServe [42]	Geographical information of parks throughout the United States. Used to calculate the distance to the nearest park (nonradial buffer measure), and the proportion of buffers that is park land.	2022	Radial buffer interpolation (500 m, 1000 m, 5000 m) for proportion of park land measure
CrimeRisk [43]	From Applied Geographic Systems, block group level risk ratio of various types of crime.	2022	Radial buffer interpolation (500 m, 1000 m, 5000 m)
Sidewalk and Crosswalk Availability [44]	National sidewalk and crosswalk presence dataset, created by applying machine learning to Google Street View images to create census tract-level estimates of the prevalence of sidewalks.	Latest imagery as of 2017	Radial buffer interpolation (500 m, 1000 m, 5000 m)
Regional Price Parity [45]	Price indices measuring geographic differences in the cost of housing, services, and more for a given metropolitan or state-wide area.	2008 - 2018 (default year: 2015[a])	Linked via MSA[c] GEOID[d] or state GEOID if participant resided outside of an MSA.
Gentrification [46]	Developed by the Urban Health Collaborative at Drexel University, this dataset categorizes each census tract as ineligible to gentrify, not gentrified, gentrified, intensely gentrified.	2000[a], 2010[a]	Linked via census tract

[a]Default year, pulled by all trialists, regardless of enrollment year.

[b]BRFSS: Behavioral Risk Factor Surveillance System.

[c]MSA: metropolitan statistical area.

[d]GEOID: geographic identifier.

Due to the limited availability of many of the environmental datasets, there is not perfect temporal alignment between the year of participant enrollment and the year of environmental measures. We do not expect that this temporal misalignment will create significant bias, particularly given the evidence that environmental attributes, such as sidewalk availability, greenspace, and gentrification change very slowly [47-49]. Study metadata, including year of enrollment for participants, length of intervention, and year of follow-up, will also be available in the pooled data to allow for more precise temporal alignment so that future sensitivity analyses may assess potential bias due to temporal misalignment.

For most environmental data linkages, radial buffers were used to define a participant's neighborhood, and environmental measures were computed using area-weighted averaging for the buffered area. For example, to calculate the walkability index for a buffer around a participant's house that was comprised 50% of Block group A, 20% of Block group B, and 30% of Block group C, the ACMT calculates the weighted mean of the walkability score for Block groups A, B, and C using the

XSL•FO
RenderX

proportion of the buffer that is comprised each Block group as that tract's weight.

Across the lifespan, a substantial amount of physical activity occurs or originates near or around one's home or neighborhood, with walking in particular most commonly occurring within 500 to 1000 m of the home [50,51]. Larger buffers may be necessary to more accurately capture neighborhood attributes in more suburban or rural areas, particularly for attributes such as green space and park access [52-54]. No buffer size fully captures the bandwidth at which neighborhood factors might influence PA; the most appropriate buffer to use may vary by age, urbanicity, and by the neighborhood environment factor [24,55-57]. Because all linkage was performed at one time, but future analyses might want to consider different buffer sizes, we selected 3 sizes for all area-buffered environmental data: 500, 1000, and 5000 m. Notable exceptions to the buffering approach include gentrification and Regional Price Parity (Table 2).

Demographic Harmonization

Trialists provided individual-level participant demographic data. While we anticipated that the available demographic variables would differ across trials, we requested that trials share baseline demographic data, including gender or sex, age, race, ethnicity, household size, or other variables that could help us determine household size, any measure or indicator of income and socioeconomic status, including income, educational level, insurance status, etc, and a variable to indicate which group the participant was randomized to. For trials where more than one household member was included (ie, parent-child dyads), we requested that a household ID be included.

A set of principles was developed to inform the harmonization of race and ethnic identity categories across trials. Race harmonization categories were based in part on review of trialists' collection of race and ethnicity data and were further informed by racial harmonization methods used in prior research [58]. Race values were not assumed or assigned unless a participant was specifically asked and self-selected or affirmed in the race or ethnicity category. For example, in cases where a trial asked about race and ethnicity in a single question, participants' ethnicity was marked as missing if they selected a race or ethnicity category other than Hispanic, and participants' race was marked as missing if they selected only Hispanic. Participants who selected more than one race category were categorized as multiracial, and participants who selected "Other" for their race but did not specify any specific races or specified a race other than the designated categories were marked as Other for their race. Ultimately, the harmonized racial categories included: American Indian or Alaska Native, Asian, Black, Native Hawaiian or Other Pacific Islander, White, Multiracial, Other, or Missing (Table 3). Ethnicity categories were Hispanic or Latino or not Hispanic or Latino. In trials where a specific demographic population was recruited, and only individuals from that specific demographic group were eligible for the study, trialists typically did not record this information in their demographic dataset. For such cases, we imputed the relevant data (eg, for a trial whose inclusion criteria required participants to self-identify as Black women, we imputed their race as Black and their gender as female).

XSL•FO
RenderX

Table . Harmonized race values.

Harmonized race value	Raw race values
American Indian Alaska Native	• Native American • American Indian or Alaskan Native
Asian	• South Asian • Asian (having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent, including Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam) • Asian Indian • Chinese • Filipino • Japanese • Korean • Vietnamese • Other Asian
Black	• African American • Black
Native Hawaiian or Other Pacific Islander	• Native Hawaiian or Other Pacific Islander (having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands) • Pacific or Hawaiian • Native Hawaiian or Pacific Islander • Guamanian or Chamorro • Samoan • Other Pacific Islander
White	• Caucasian • White • Non-Hispanic White
Multiple	• More than one race • Multiple • Multiple selected race categories • Multiracial
Other	• Other race or origin • Other

Education harmonization was informed by the education categories used by trials in the pool. Education level for adults was combined and collapsed as needed to create six categories: (1) no schooling; (2) less than high school degree; (3) high school degree or General Educational Development; (4) some college, including technical school or associate's degree; (5) bachelor's degree; (6) some postgraduate education or more (Table 4). For trials that recruited children and collected data on parent education, we harmonized parent education according to these same 6 categories.

XSL•FO
RenderX

Table . Education harmonization mapping.

Harmonized education value	Raw education values from trial data
No schooling	No schooling
Less than high school degree	• Less than 7th grade • 8th grade or less • Less than 8th grade • Grades 0 - 8 • Grades 1 - 8 • Less than high school • Some high school • Junior high or middle school • 8th grade or more, but less than high school • Grades 9 - 11 • 11th grade or less • Partial high school
High school degree or GED	• 12 - 13 years • High school degree or GED[a] • Completed high school • High school degree • High school diploma or GED • High school or equivalent (9-12) • High school graduate • High school
Some college, technical, or associate school	• Some college • 14 - 15 years • Trade or technical school • Some college or vocational training • 1 - 3 years of college • Associate's degree • Technical school certificate • Some college or technical school • 2 years of college, AA[b] or technical school • College (13-16) • College or vocational or technical school • 2-year college
Bachelor's degree	• College graduate • 16 - 17 years • Completed college or university • Bachelor's degree • College (BS[c] or BA[d]) • College degree • College or university • Graduated from college • 4-year college
Some postgraduate education	• Graduate degree • 18 or more years • Postgraduate training • Completed graduate degree • Postgraduate or professional degree • Graduate work or higher • Master's degree • Doctoral degree • Postgraduate degree • PhD or equivalent • Postgraduate or advanced degree

[a]GED: General Educational Development.

[b]AA: Associate of Arts.

[c]BS: Bachelor of Science.

[d]BA: Bachelor of Arts.

XSL•FO
RenderX

Harmonizing participant-reported income data was challenging. Preharmonization tables provided no clear options for harmonizing income by collapsing or combining categories (as having categories was the common practice, such as <US $20,000, US $20,001-US $40,000, etc) due to vastly different ranges of income categories across trials. It was particularly problematic to determine how to harmonize across the highest income category option, given the variability across trials and the open-ended nature of the highest category (eg, >US $75,000 vs >US $300,000). Furthermore, because the purchasing power of income varies across regions, the same income level might represent substantially different economic resources for participants in trials in different geographic regions. Due to these complexities, we chose to use not to harmonize income at all for primary aims analyses, but rather to use a nonharmonized within-trial categorical income variable for confounding adjustment in analyses.

Physical Activity Harmonization

Each InSPACE trial and the associated intervention took place over its own distinct amount of time (eg, some trials lasted weeks and some months or years). Trialists provided individual-level accelerometry files for each data collection period. We defined baseline accelerometry data to include all accelerometry data collected prior to the start of the intervention. We defined the postintervention accelerometry data to be the first tranche of data collected after the end of the intervention period. Some trialists shared additional follow-up accelerometry data (eg, after noncontact follow-up periods), which we harmonized for use in future analyses. We processed the data using established definitions of valid accelerometry data, based on existing literature. Our primary definition of valid accelerometry data requires 8 hours of wear time per day for a given day to be valid, and 3 days of valid wear time for a given time period of data collection to be considered valid [59]. Baseline, postintervention, and any additional follow-up after postintervention accelerometry data were processed using R (R Core Team) as detailed below.

We requested accelerometry data in the rawest form available. For example, in most cases, Actigraph accelerometer (most common device used in pooled trials) files were shared in ActiLife's native file format (agd), or as a csv file aggregated at no more than the 1-minute epoch level to ensure that the InSPACE team could implement consistent wear time and conversion to physical activity minutes processing protocols across all participants from different trials. Most (31/39, 79%) of trialists used hip-worn Actigraph accelerometers (models 7146, GT3X, or GT9X). However, data from wrist-worn Actigraph, thigh-worn activPAL3, and wrist-worn Actical and GENEactiv accelerometers were also included in the pooled accelerometry data (see below for differences in processing dependent on wear site).

While trials with only day-level accelerometry data were not considered eligible during screening, there were 2 trials that were recruited and entered into a contract believing that they had minute-level accelerometry data but ultimately were only able to locate day-level accelerometry data. In these cases, we confirmed with the trialists' study team the specific wear time algorithm used as well as the activity count thresholds to determine whether these could rigorously be combined with the other harmonized and processed accelerometry data.

Due to varying accelerometer types, wear locations, and file types, we developed a process to derive minute- and day-level accelerometry summaries for each participant in each study at baseline and postintervention. First, we summarized axis counts for each individual at each data collection time point by 1-minute epochs and derived a wear-time indicator variable using a previously validated wear-time algorithm [60]. Next, an indicator variable was derived that categorized each minute as moderate-to-vigorous physical activity (MVPA) based on the axis count threshold that aligned with the accelerometer type, wear location, and age of the participant [61-66]. Because we did not consistently have triaxial data across all trials, we chose to use only the vertical axis counts to determine wear time and MVPA.

One trial used wrist-worn GENEActiv devices and provided the raw accelerometry data in gravitational units. For this study, wear time was determined using an SD-based approach developed by van Hees et al [67,68]. This approach uses a 60-minute time window with 15-minute overlapping increments. Physical activity thresholds for this study were based on the mean acceleration per epoch rather than counts per minute to be consistent with available validation studies for GENEActiv accelerometry data [69].

To establish MVPA thresholds used for examining physical activity, we conducted a literature review of MVPA threshold validation studies for all types of accelerometers across all participant ages represented in participating trials, as well as thresholds used by participating trialists. Using this information, we chose appropriate MVPA thresholds that (1) considered vertical axis counts only (as opposed to a triaxial vector magnitude measure), (2) were most frequently used by InSPACE trialists, and (3) were well established and commonly used in other physical activity studies (Table 5). Once a threshold was chosen for each accelerometer type, wear location, and age group combination represented the pooled dataset, we applied the chosen thresholds to the pooled accelerometry data to create harmonized PA measures across trials. In cases where age was missing for an individual, study population information was used to categorize the individual as adult, child, or older adult to choose an appropriate MVPA cut point. The current process has documented the accelerometer types, wear location, and age group combinations that are present across all currently recruited trials. If additional trials were to be recruited, the InSPACE investigator team would conduct a similar process for establishing the appropriate MVPA threshold, using existing validation studies that are relevant for the accelerometer type, wear location, and age group as well as the MVPA threshold used by the trial to inform the decision-making process. Raw accelerometry data for all studies will also be available to researchers to allow for different processing and harmonization approaches or to conduct sensitivity analyses.

XSL•FO
RenderX

Table . Moderate-to-vigorous physical activity thresholds by accelerometer type, wear location, and participant age.

Accelerometer	MVPA[a] threshold
Actigraph	
Wrist	
Adults (18 - 65)	4730 cpm[b]
Hip	
Children (<18)	2296 cpm[c]
Adults (18-65)	1952 cpm[d]
Older adults (>65)	1041 cpm[e]
Actical	
Hip	
Adults (≥18)	1535 cpm[f]
Children (<18)	1600 cpm[f]
ActivPAL	
Thigh	
Adults (18-65)	12,028 cpm[g]
GeneActiv	
Wrist	
Adults (18-65)	93.2 mg[h]
Children (<18)	191.6 mg[h]

[a]MVPA: moderate-to-vigorous physical activity.

[b]Counts per minute; weighted average of calibrated thresholds [64,66].

[c]Calibrated on uniaxial counts on children 5 - 8 years old [62].

[d]Calibration of uniaxial counts [63].

[e]Calibration of uniaxial counts [61].

[f]Calibration of uniaxial counts [70].

[g]Based on 3007 counts per 15-second epoch [65].

[h]Threshold based on average gravitational units (mg) [69,71].

Day-level summaries for each participant provided the total wear minutes per day, which were used to determine whether a given day was valid. Accelerometry days were considered valid if the day included 8 or more hours of wear time [59]. Of the included trials, 95% (n=37) trials requested that participants wear the accelerometer for 7 days at each assessment time point. One trial had participants wear accelerometers continuously throughout the intervention and into the postintervention period, so we extracted 7 days postintervention from the continuous data collection. Participants with fewer than 3 days of valid accelerometry data at baseline or postintervention were excluded. An exception was made for data from one trial that collected only 2 days of accelerometry data at each time point. For this trial, any participant with at least 2 valid days with 8 or more hours of wear time for baseline and postintervention was included. A final wide-format InSPACE accelerometry dataset was created that included a variable for the mean minutes of MVPA per valid baseline day, mean minutes of MVPA per valid postintervention day, the type of accelerometer and the specific threshold used to define MVPA, total valid days at baseline, and total valid days at postintervention. Figure 1 shows the full cycle of trial recruitment and data collection and processing.

XSL•FO
RenderX

Figure 1. Interventions Supporting Physical Activity and the Environment data workflow diagram. ACMT: Automatic Context Measurement Tool; InSPACE: Interventions Supporting Physical Activity and the Environment; IRB: institutional review board; MVPA: moderate-to-vigorous physical activity.

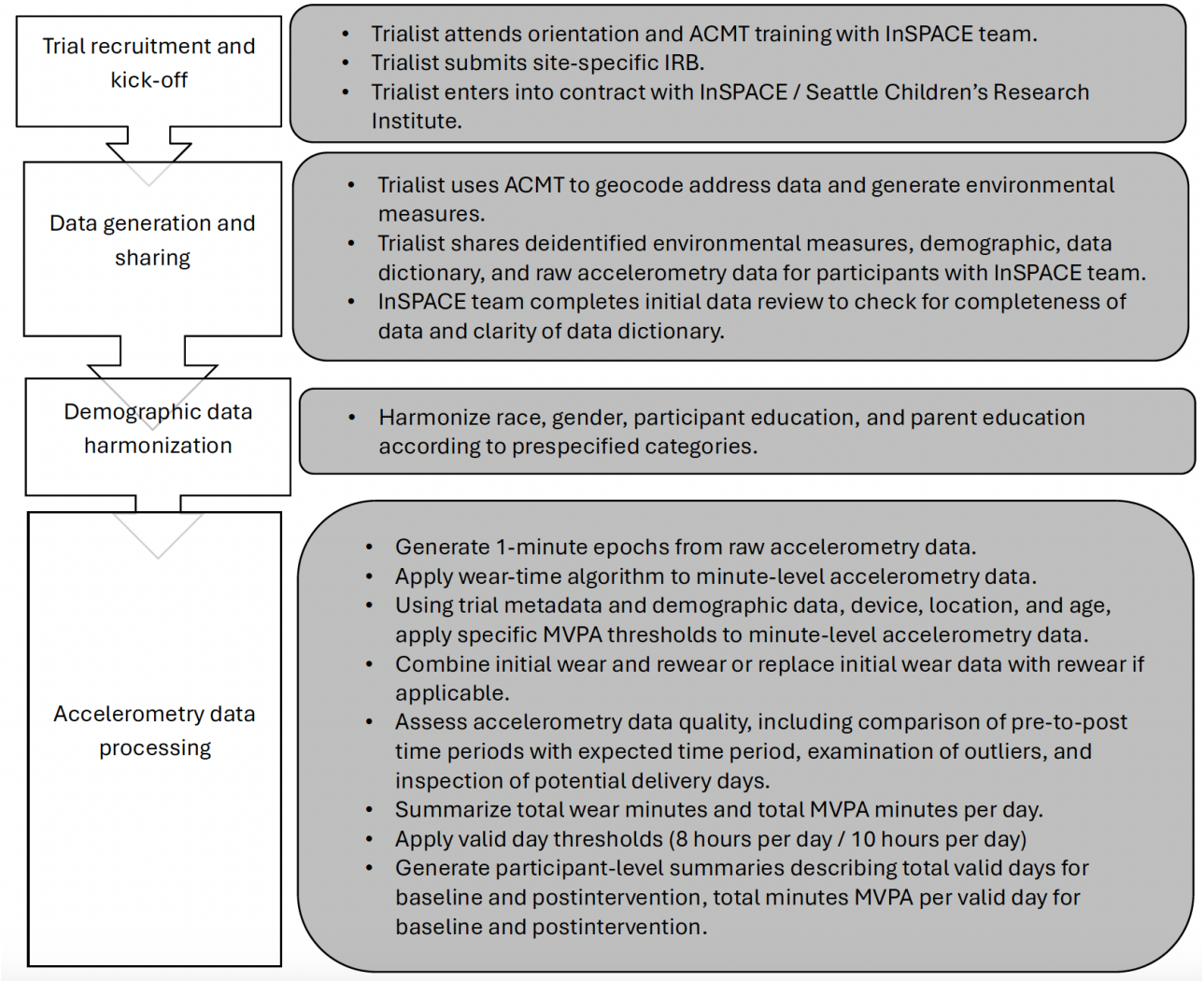

Accelerometry Rewear Data

In some cases, trialists recontacted individuals whose initial accelerometry data did not provide adequate data (ie, not enough valid wear days) and asked participants to wear the accelerometer again. For these "rewear" scenarios, we applied the following logic: if valid rewear data were available and initial wear data did not have any valid days, only the rewear data were used. If both the initial wear data and rewear data contained valid wear days, we used the initial wear data up to the first valid rewear day and combined the initial wear data with all valid rewear days.

Visual Inspection of Delivery Days

Several quality assurance reports were generated to help us validate accelerometry data. We did not have complete data from trialists about when and if accelerometers were delivered or returned via mail. Some trials gave accelerometers to participants in person, and some mailed them, and at least 1 trial removed delivery days prior to sharing their accelerometry with the InSPACE team. While previous studies have used machine learning to develop algorithms for identifying delivery days (Moore et al [72]), the inconsistencies in expected delivery days and the lack of triaxial data across all InSPACE trials required us to use a combination of algorithm and manual review to identify potential delivery days.

In reviewing the pooled accelerometry data, we were primarily concerned with days that were marked as a valid wear day (ie, 480 or more minutes of wear time) but were capturing movement of the accelerometer in transit. Days that were already marked as invalid due to insufficient wear time were not reviewed as potential delivery days, given that they were already excluded from analysis. In a manual review of data from 2 trials, we identified patterns that resembled potential delivery days. Participants were expected to wear devices for 7 consecutive days, so we identified deviations from this pattern, such as 1 or 2 days marked as a valid wear day, followed by or preceded by 2 or more invalid days (Figure 2), potentially indicating the device was in transit. When this pattern was identified for individuals who had 4 or more consecutive days of valid wear time, we flagged these individuals for manual review. Flagged days were visually inspected using a plot of accelerometer counts (refer to Figure 2 for a visual of marked possible delivery days) and confirmed as likely delivery (nonwear) days.

XSL•FO
RenderX

Figure 2. Plot of potential delivery days for an individual baseline data collection period with wear minutes per day at the top.

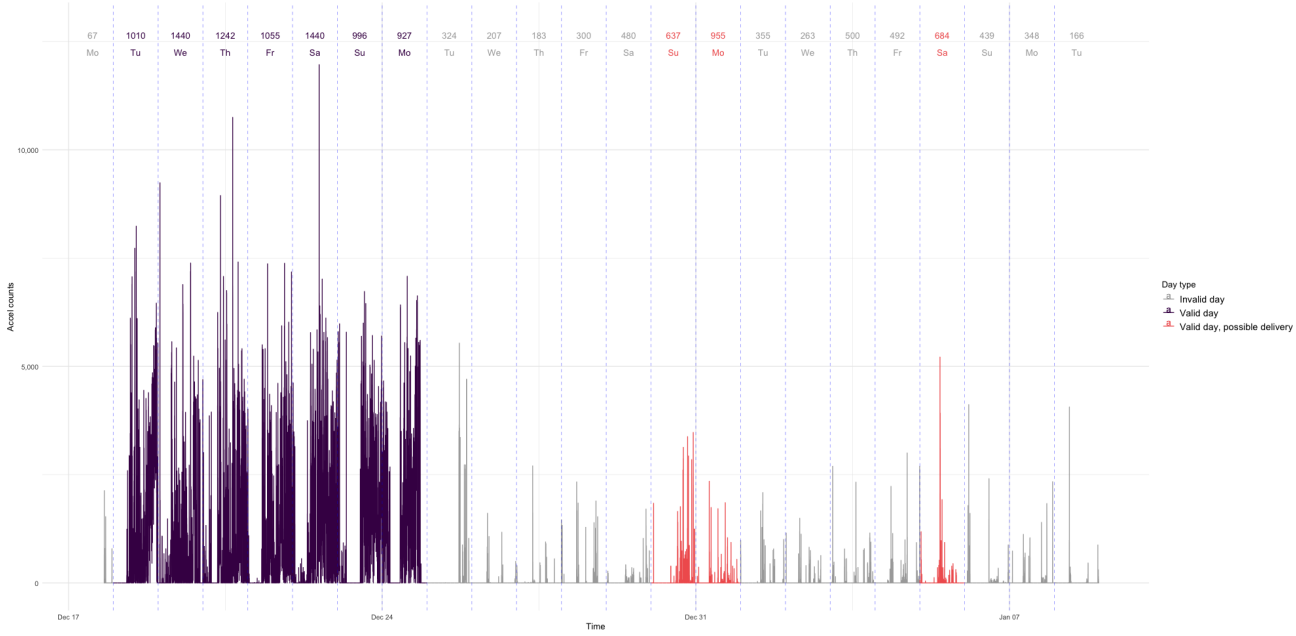

Accelerometer Data Collection Dates

Additional QA reports were created to examine the time elapsed between the baseline accelerometry data collection and postintervention data collection. Time between data collection periods was compared to expected time elapsed based on trial metadata (eg, the planned length of time of the intervention) to ensure that accelerometry data files were properly sorted into baseline and postintervention files, and that there were no accelerometry data that were potentially overlapping in time with the trial intervention period for the participant, particularly for interventions that included an intervention team-facilitated physical activity component. In cases where baseline and postintervention accelerometry data overlapped, accelerometry start dates were compared to expected start dates for accelerometry data collection for baseline and postintervention, and only one set of accelerometry data was retained, while the other was omitted due to overlapping data collection period. If no other accelerometry data were available for the individual for the missing data collection period, the individual was omitted from the analysis.

Ethical Considerations

The InSPACE project was approved by Seattle Children's IRB (STUDY00002955). Because of the nature of InSPACE and the fact that the trials whose data were contributed to the data pool were already covered under the IRB of the trialist's institution, this study was determined not to require a single-site IRB. The InSPACE study team provided boilerplate IRB language to study teams and offered individual consultation regarding approaches for seeking IRB approval at the trial site. These template documents were updated throughout the recruitment period in response to patterns and common questions that emerged from working with many teams over time. Given that all recruited trials chose to have their own trial team geocode and generate environmental measures, only deidentified data were shared with the InSPACE team, allowing for a much simpler IRB process at the pooled level.

IRB review and approval for participating trials generally followed one of four patterns: (1) Most commonly, the site IRB considered the InSPACE work exempt, and no modifications were required. (2) The trial had existing data sharing language in its protocol or consent documents and had to create a modification and get IRB approval for the new use of the data as a part of InSPACE. (3) The site did not have data sharing language already in their protocol or consent documents and had to create and get IRB approval for a modification requesting a waiver of reconsent and adding data sharing to the protocol. (4) The IRB did not approve the use of the participants' home address data. This fourth scenario arose most often when the trial's investigator had changed institutions and was not allowed to transfer identifiable data as part of the move. In some cases, trials were closed with their IRB and needed to be reopened before modification requests and approvals could be made.

Results

As part of the InSPACE project, we recruited 39 physical activity intervention trials across the United States as of August 2025. Trial sample sizes ranged from 32 participants to over 500 participants. Nine of the trials explicitly recruited children (<18 years old), 2 of which recruited parent-child dyads. Trials with parent-child or 2-adult (couple-dyads) are identified by household or family IDs to account for their identical environmental exposures. Trial recruitment is completed as of August 2025, and analysis of the additional 8 recruited trials will be completed in 2026, with results expected to be published in late 2026.

One initially recruited trial was ultimately unable to locate accelerometry data and was thus omitted. Two participating trials discovered after entering into a contract with InSPACE that they only had accelerometry data available at the day level, summarized using trial-specific wear-time validation and MVPA thresholds. These data were retained in the InSPACE pool but flagged due to the potential lack of consistency in PA

XSL•FO
RenderX

summarization. Additionally, we omitted participants from one trial whose postintervention accelerometry data were collected after March 2020, due to the potential influence of COVID-19 restrictions on physical activity.

All recruited trials collected data on age and either collected data on sex/gender or explicitly recruited only participants of a particular sex/gender. In 94% of the trials, race was either explicitly collected or recruitment was restricted to a particular racial group. The availability of additional demographic variables across the recruited trials is shown in Table 6.

Of the 39 recruited trials as of August 2025, we have completed data processing and harmonization for 31 trials. Some trials provided some data for all potential participants, including those who were not ultimately randomized, while others only included randomized participants. A total of 4471 participants who had been randomized to an intervention were included in the pooled data. Of those randomized, 97.5% (4360) of participants had addressed that InSPACE trialists were able to geocode and link to environmental measures, and 76.4% (3417) of randomized participants had at least 1 valid day of accelerometry data, with 49.4% (2208) of the randomized participants having complete, valid baseline and postintervention data.

Table . Availability of demographic variables across trials (N=31).

Demographic variable	Trials that collected data, n (%)
Age	31 (100)
Race[a]	29 (94)
Ethnicity[a]	29 (94)
Gender[a]	31 (100)
Education (adult trials)	22 (71)
Parent education (trials with children)	31 (100)
Income	24 (77)
Household size	15 (48)

[a]Includes trials that did not explicitly collect the demographic variable, but whose study population was selected based on a specific demographic characteristic.

Discussion

Data Pooling

To date, research on whether the built environment can modify the effectiveness of physical activity interventions has been limited by small sample sizes and the limited geographic spread of participants. The InSPACE project is creating a pooled dataset of physical activity intervention trial participants along with derived individual participant home neighborhood environmental attributes to be used in studying the role of the environment in influencing physical activity promotion. We have also documented our processes of data collection and harmonization for use in future work.

Strengths and Challenges

There are numerous strengths of the pooled InSPACE dataset. We used standardized methods to analyze all accelerometry data to ensure consistency in defining wear time and operationalizing MVPA. We also linked standardized environmental measures using our own standardized interpolation methods to ensure consistency across all pooled trials, rather than harmonizing any trial-specific measures. Trialists themselves used the ACMT to geocode and generate environmental measures, so no sharing of address or other identifying data was required, greatly streamlining the IRB process for the InSPACE team and for participating trialists. We included numerous environmental measures of interest, including attributes of the social, economic, and built environment allowing us to explore several potential pathways through which the environment may affect physical activity changes within intervention trials.

The InSPACE study is further strengthened by the input of numerous experts in the field of physical intervention trials who made up the SAC. These individuals assisted us in identifying and recruiting trials across the country to help ensure the diversity of participants in the pooled trial and a variety of intervention types.

We also faced several challenges both in recruiting trials and in harmonizing trial data. Due to the challenges of harmonizing income data across trials and the considerably different income categories captured across trials, we do not have a consistent measure of income or socioeconomic status for participants in the pooled data but rather maintain the study-specific income variables. This approach reduces potential bias that may be introduced in harmonizing vastly different income groups but also makes comparisons by individual income levels more complex. Despite these limitations in harmonizing income, we do have harmonized education level for participants themselves (for adults) or for the parent of the participant (for children) in 94% of the pooled trials, which may provide a proxy for socioeconomic status.

Additionally, while we created standard measures of MVPA across trials using existing validation studies for the specific accelerometry type and wear location, there is potential for error in using standardized uniaxial count-based measures of physical activity, and there remains a lack of consensus on the accurate count threshold for hip-worn accelerometers among older adults

XSL•FO
RenderX

and for wrist-worn devices among all ages. Different thresholds or approaches to measuring MVPA with accelerometry data can result in vastly different measures of physical activity [73,74]. However, uniaxial count-based measures are a convenient and well-studied method of measuring physical activity that can be applied to numerous types of devices, including both uniaxial and triaxial.

Next Steps

Since we anticipate that the InSPACE data pool might be usable to address questions outside of the primary aims of InSPACE, we developed a process for the use of pooled data outside of the scope of InSPACE. Language regarding this process was included in the data use agreement with consortium trials. Proposals from the InSPACE consortium trialists or from anyone outside the InSPACE Investigator team are first reviewed by the InSPACE Investigator team where they determine whether the proposed research question (1) provides an important contribution to the field and is addressable using InSPACE data, (2) is outside of the original InSPACE aims, and (3) is within the scope of InSPACE (ie, focuses on environmental factors related to physical activity). Once approval from the InSPACE investigators' team is obtained, the proposal is brought to the SAC for final approval. For proposals that fall outside the scope of InSPACE, additional consent from all participating trialists is required for access to their data. Proposals brought from the InSPACE Investigators team do not require SAC approval if the proposed research question falls within the scope of the InSPACE. We are hopeful that the InSPACE data pool will provide opportunities to address additional interesting questions around physical activity and neighborhood environment.

Conclusion

Despite the challenges in gathering and harmonizing data across InSPACE trials, the resulting pooled dataset provides a unique opportunity to examine the potential effect of the environment on physical activity changes in the context of interventions seeking to increase this behavior.

Acknowledgments

Map data are copyrighted by OpenStreetMap contributors and available from [75].

Thank you to all the members of the Scientific Advisory Committee for their contributions to our evolving trial eligibility criteria and the development of our data harmonization protocols. Members include Jim Sallis, PhD (SAC Chair); Marc Adams, PhD; Monica Baskin, PhD; Ross Brownson, PhD; Abby King, PhD; David X. Marquez, PhD; Tiffany Powell-Wiley, PhD.

Funding

Research reported in this publication was supported in part by the National Heart Lung and Blood Institute of the National Institutes of Health (NIH) under award number R01HL157188. It is subject to the NIH Public Access Policy. Through acceptance of this federal funding, NIH has been given a right to make this manuscript publicly available in PubMed Central upon the Official Date of Publication, as defined by NIH. The content is solely the responsibility of the authors and does not necessarily represent official views of the NIH.

Data Availability

The datasets generated during this study are available from the corresponding author upon reasonable request and approval (see Next Steps section for further description of the data request process).

Authors' Contributions

Funding acquisition: BES, AS, SJM
Project administration: MR, BES
Data curation and software: AJY, SJM, KS
Writing – original draft: MR, AJY
Writing – review & editing: BES, MR, SJM, AS, KS, AJY

Conflicts of Interest

None declared.

References

1. Abell B, Glasziou P, Hoffmann T. The contribution of individual exercise training components to clinical outcomes in randomised controlled trials of cardiac rehabilitation: a systematic review and meta-regression. Sports Med Open 2017 Dec;3(1):19. [doi: 10.1186/s40798-017-0086-z] [Medline: 28477308]
2. Anderson D, Seib C, Rasmussen L. Can physical activity prevent physical and cognitive decline in postmenopausal women? A systematic review of the literature. Maturitas 2014 Sep;79(1):14-33. [doi: 10.1016/j.maturitas.2014.06.010] [Medline: 25008420]

XSL•FO
RenderX

3. Kyu HH, Bachman VF, Alexander LT, et al. Physical activity and risk of breast cancer, colon cancer, diabetes, ischemic heart disease, and ischemic stroke events: systematic review and dose-response meta-analysis for the Global Burden of Disease Study 2013. BMJ 2016 Aug 9;354:i3857. [doi: 10.1136/bmj.i3857] [Medline: 27510511]
4. Warburton DER, Bredin SSD. Health benefits of physical activity: a systematic review of current systematic reviews. Curr Opin Cardiol 2017 Sep;32(5):541-556. [doi: 10.1097/HCO.0000000000000437] [Medline: 28708630]
5. Abdin S, Welch RK, Byron-Daniel J, Meyrick J. The effectiveness of physical activity interventions in improving well-being across office-based workplace settings: a systematic review. Public Health (Fairfax) 2018 Jul;160:70-76. [doi: 10.1016/j.puhe.2018.03.029] [Medline: 29751224]
6. Adamson BC, Ensari I, Motl RW. Effect of exercise on depressive symptoms in adults with neurologic disorders: a systematic review and meta-analysis. Arch Phys Med Rehabil 2015 Jul;96(7):1329-1338. [doi: 10.1016/j.apmr.2015.01.005] [Medline: 25596001]
7. Elgaddal N, Kramarow EA, Reuben C. Physical activity among adults aged 18 and over: United States 2020. Centers for Disease Control and Prevention. 2022. URL: https://stacks.cdc.gov/view/cdc/120213 [accessed 2025-03-14]
8. Piercy KL, Troiano RP, Ballard RM, et al. The physical activity guidelines for Americans. JAMA 2018 Nov 20;320(19):2020-2028. [doi: 10.1001/jama.2018.14854] [Medline: 30418471]
9. Adams MA, Hurley JC, Phillips CB, et al. Rationale, design, and baseline characteristics of WalkIT Arizona: a factorial randomized trial testing adaptive goals and financial reinforcement to increase walking across higher and lower walkable neighborhoods. Contemp Clin Trials 2019 Jun;81:87-101. [doi: 10.1016/j.cct.2019.05.001] [Medline: 31063868]
10. Chasan-Taber L, Marcus BH, Rosal MC, et al. Proyecto Mamá: a lifestyle intervention in overweight and obese Hispanic women: a randomised controlled trial--study protocol. BMC Pregnancy Childbirth 2015 Jul 30;15(1):157. [doi: 10.1186/s12884-015-0575-3] [Medline: 26223246]
11. Crist K, Full KM, Linke S, et al. Health effects and cost-effectiveness of a multilevel physical activity intervention in low-income older adults; results from the PEP4PA cluster randomized controlled trial. Int J Behav Nutr Phys Act 2022 Jun 27;19(1):75. [doi: 10.1186/s12966-022-01309-w] [Medline: 35761363]
12. Gell NM, Grover KW, Savard L, Dittus K. Outcomes of a text message, Fitbit, and coaching intervention on physical activity maintenance among cancer survivors: a randomized control pilot trial. J Cancer Surviv 2020 Feb;14(1):80-88. [doi: 10.1007/s11764-019-00831-4] [Medline: 31776849]
13. Jake-Schoffman DE, Turner-McGrievy G, Wilcox S, Moore JB, Hussey JR, Kaczynski AT. The mFIT (Motivating Families with Interactive Technology) Study: a randomized pilot to promote physical activity and healthy eating through mobile technology. J technol behav sci 2018 Sep;3(3):179-189. [doi: 10.1007/s41347-018-0052-8]
14. Keadle SK, Meuter L, Phelan S, Phillips SM. Charity-based incentives motivate young adult cancer survivors to increase physical activity: a pilot randomized clinical trial. J Behav Med 2021 Oct;44(5):682-693. [doi: 10.1007/s10865-021-00218-w] [Medline: 33825070]
15. King AC, Campero I, Sheats JL, et al. Testing the effectiveness of physical activity advice delivered via text messaging vs. human phone advisors in a Latino population: the On The Move randomized controlled trial design and methods. Contemp Clin Trials 2020 Aug;95:106084. [doi: 10.1016/j.cct.2020.106084] [Medline: 32659437]
16. Mendoza JA, Baker KS, Moreno MA, et al. A Fitbit and Facebook mHealth intervention for promoting physical activity among adolescent and young adult childhood cancer survivors: a pilot study. Pediatr Blood Cancer 2017 Dec;64(12):e26660. [doi: 10.1002/pbc.26660] [Medline: 28618158]
17. Pekmezi D, Ainsworth C, Joseph R, et al. Rationale, design, and baseline findings from HIPP: a randomized controlled trial testing a home-based, individually-tailored physical activity print intervention for African American women in the Deep South. Contemp Clin Trials 2016 Mar;47:340-348. [doi: 10.1016/j.cct.2016.02.009] [Medline: 26944022]
18. Conn VS, Hafdahl AR, Mehr DR. Interventions to increase physical activity among healthy adults: meta-analysis of outcomes. Am J Public Health 2011 Apr;101(4):751-758. [doi: 10.2105/AJPH.2010.194381] [Medline: 21330590]
19. Gasana J, O'Keeffe T, Withers TM, Greaves CJ. A systematic review and meta-analysis of the long-term effects of physical activity interventions on objectively measured outcomes. BMC Public Health 2023 Sep 2;23(1):1697. [doi: 10.1186/s12889-023-16541-7] [Medline: 37660119]
20. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF. Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis. Soc Sci Med 2017 Nov;192:125-133. [doi: 10.1016/j.socscimed.2017.09.021] [Medline: 28965003]
21. Akwo EA, Kabagambe EK, Harrell FE Jr, et al. Neighborhood deprivation predicts heart failure risk in a low-income population of Blacks and Whites in the Southeastern United States. Circ Cardiovasc Qual Outcomes 2018 Jan;11(1):e004052. [doi: 10.1161/CIRCOUTCOMES.117.004052] [Medline: 29317456]
22. Frank LD, Iroz-Elardo N, MacLeod KE, Hong A. Pathways from built environment to health: a conceptual framework linking behavior and exposure-based impacts. J Transp Health 2019 Mar;12:319-335. [doi: 10.1016/j.jth.2018.11.008]
23. Laddu D, Paluch AE, LaMonte MJ. The role of the built environment in promoting movement and physical activity across the lifespan: implications for public health. Prog Cardiovasc Dis 2021;64:33-40. [doi: 10.1016/j.pcad.2020.12.009] [Medline: 33428966]

XSL•FO
RenderX

24. Villanueva K, Knuiman M, Nathan A, et al. The impact of neighborhood walkability on walking: does it differ across adult life stage and does neighborhood buffer size matter? Health Place 2014 Jan;25:43-46. [doi: 10.1016/j.healthplace.2013.10.005] [Medline: 24239702]
25. Wei J, Wu Y, Zheng J, Nie P, Jia P, Wang Y. Neighborhood sidewalk access and childhood obesity. Obes Rev 2021 Feb;22 Suppl 1(Suppl 1):e13057. [doi: 10.1111/obr.13057] [Medline: 32638457]
26. Cohen DA, Han B, Nagel CJ, et al. The first national study of neighborhood parks: implications for physical activity. Am J Prev Med 2016 Oct;51(4):419-426 [FREE Full text] [doi: 10.1016/j.amepre.2016.03.021] [Medline: 27209496]
27. Cohen DA, Leuschner KJ. How can neighborhood parks be used to increase physical activity? Rand Health Q 2019 May;8(3):4. [Medline: 31205804]
28. Schipperijn J, Bentsen P, Troelsen J, Toftager M, Stigsdotter UK. Associations between physical activity and characteristics of urban green space. Urban For Urban Green 2013 Jan;12(1):109-116. [doi: 10.1016/j.ufug.2012.12.002]
29. Douglas JA, Briones MD, Bauer EZ, Trujillo M, Lopez M, Subica AM. Social and environmental determinants of physical activity in urban parks: testing a neighborhood disorder model. Prev Med 2018 Apr;109:119-124. [doi: 10.1016/j.ypmed.2018.01.013] [Medline: 29366818]
30. Kerr J, Norman GJ, Adams MA, et al. Do neighborhood environments moderate the effect of physical activity lifestyle interventions in adults? Health Place 2010 Sep;16(5):903-908. [doi: 10.1016/j.healthplace.2010.05.002] [Medline: 20510642]
31. King AC, Toobert D, Ahn D, et al. Perceived environments as physical activity correlates and moderators of intervention in five studies. Am J Health Promot 2006;21(1):24-35. [doi: 10.1177/089011710602100106] [Medline: 16977910]
32. Sallis JF, Cervero RB, Ascher W, Henderson KA, Kraft MK, Kerr J. An ecological approach to creating active living communities. Annu Rev Public Health 2006;27:297-322. [doi: 10.1146/annurev.publhealth.27.021405.102100] [Medline: 16533119]
33. Zhou W, Youngbloom A, Ren X, Saelens BE, Mooney SD, Mooney SJ. The Automatic Context Measurement Tool (ACMT) to compile participant-specific built and social environment measures for health research: development and usability study. JMIR Form Res 2024 Oct 4;8:e56510. [doi: 10.2196/56510] [Medline: 39365663]
34. Glanz K, Rimer BK, Viswanath K. Health Behavior and Health Education: Theory, Research, and Practice, 4th edition: Jossey-Bass/Wiley; 2008.
35. Rundle AG, Bader MDM, Mooney SJ. The disclosure of personally identifiable information in studies of neighborhood contexts and patient outcomes. J Med Internet Res 2022 Mar 17;24(3):e30619. [doi: 10.2196/30619] [Medline: 35103610]
36. Corti P, Kraft TJ, Mather SV, Park B. PostGIS Cookbook: Over 80 Task-Based Recipes to Store, Organize, Manipulate, and Analyze Spatial Data in a PostGIS Database Reprint: Packt Publishing; 2014.
37. American Community Survey (ACS). US Census Bureau. URL: https://www.census.gov/programs-surveys/acs [accessed 2025-03-27]
38. Census tract level state maps of the modified food environment index (mRFEI). Centers for Disease Control and Prevention.: National Center for Chronic Disease Prevention and Health Promotion (U.S. Division of Nutrition, Physical Activity, and Obesity); 2012 Apr. URL: https://stacks.cdc.gov/view/cdc/61367 [accessed 2026-01-06]
39. National walkability index. US Environmental Protection Agency. 2021 Jun. URL: https://www.epa.gov/smartgrowth/smart-location-mapping#walkability [accessed 2026-01-06]
40. About PLACES: local data for better health. Centers for Disease Control and Prevention. 2024. URL: https://www.cdc.gov/places/about/index.html [accessed 2025-03-27]
41. About annual NLCD. USGS. URL: https://www.usgs.gov/centers/eros/science/about-annual-nlcd [accessed 2025-03-27]
42. Park data downloads. Trust for Public Land. URL: https://www.tpl.org/parkserve/downloads [accessed 2023-02-21]
43. AGS CrimeRisk methdology. Applied Geographic Solutions. 2022. URL: https://appliedgeographic.com/crimerisk/ [accessed 2023-03-21]
44. Nguyen QC, Keralis JM, Dwivedi P, et al. Leveraging 31 million Google Street View images to characterize built environments and examine county health outcomes. Public Health Rep 2021;136(2):201-211. [doi: 10.1177/0033354920968799] [Medline: 33211991]
45. Methodology for regional price parities, real personal consumption expenditures, and real personal income. Beureau of Economic Analysis. 2023 Apr. URL: https://www.bea.gov/data/prices-inflation/regional-price-parities-state-and-metro-area [accessed 2025-03-27]
46. Hirsch JA, Schinasi LH. A measure of gentrification for use in longitudinal public health studies in the US. Drexel University Urban Health Collaborative. 2019 Aug. URL: https://drexel.edu/uhc/resources/briefs/Measure-of-Gentrification-for-Use-in-Longitudinal-Public-Health-Studies-in-the-US [accessed 2025-12-04]
47. Finio N. Measurement and definition of gentrification in urban studies and planning. J Plan Lit 2022 May;37(2):249-264. [doi: 10.1177/08854122211051603]
48. Homer C, Dewitz J, Jin S, et al. Conterminous United States land cover change patterns 2001-2016 from the 2016 National Land Cover Database. ISPRS J Photogramm Remote Sens 2020 Apr;162:184-199. [doi: 10.1016/j.isprsjprs.2020.02.019] [Medline: 35746921]

XSL•FO
RenderX

49. Mooney SJ, Smith CM, Spalt EW, et al. Built environment change over time using Google Street View assessments of Hispanic Community Health Study/Study of Latinos (HCHS/SOL) cities. J Urban Health 2025 Jun;102(3):670-679. [doi: 10.1007/s11524-025-00980-1] [Medline: 40467920]
50. Hurvitz PM, Moudon AV, Kang B, Fesinmeyer MD, Saelens BE. How far from home? The locations of physical activity in an urban U.S. setting. Prev Med 2014 Dec;69:181-186. [doi: 10.1016/j.ypmed.2014.08.034] [Medline: 25285750]
51. Kang B, Moudon AV, Hurvitz PM, Reichley L, Saelens BE. Walking objectively measured: classifying accelerometer data with GPS and travel diaries. Med Sci Sports Exerc 2013 Jul;45(7):1419-1428. [doi: 10.1249/MSS.0b013e318285f202] [Medline: 23439414]
52. Addy CL, Wilson DK, Kirtland KA, Ainsworth BE, Sharpe P, Kimsey D. Associations of perceived social and physical environmental supports with physical activity and walking behavior. Am J Public Health 2004 Mar;94(3):440-443. [doi: 10.2105/ajph.94.3.440] [Medline: 14998810]
53. Banda JA, Wilcox S, Colabianchi N, Hooker SP, Kaczynski AT, Hussey J. The associations between park environments and park use in southern US communities. J Rural Health 2014;30(4):369-378. [doi: 10.1111/jrh.12071] [Medline: 24717017]
54. Wilson DK, Kirtland KA, Ainsworth BE, Addy CL. Socioeconomic status and perceptions of access and safety for physical activity. Ann Behav Med 2004 Aug;28(1):20-28. [doi: 10.1207/s15324796abm2801_4] [Medline: 15249256]
55. Brownson RC, Hoehner CM, Day K, Forsyth A, Sallis JF. Measuring the built environment for physical activity: state of the science. Am J Prev Med 2009 Apr;36(4 Suppl):S99-123. [doi: 10.1016/j.amepre.2009.01.005] [Medline: 19285216]
56. Diez Roux AV, Evenson KR, McGinn AP, et al. Availability of recreational resources and physical activity in adults. Am J Public Health 2007 Mar;97(3):493-499. [doi: 10.2105/AJPH.2006.087734] [Medline: 17267710]
57. Fotheringham AS, Wong DWS. The modifiable areal unit problem in multivariate statistical analysis. Environ Plan A 1991 Jul;23(7):1025-1044. [doi: 10.1068/a231025]
58. Lyren A, Haines E, Fanta M, et al. Racial and ethnic disparities in common inpatient safety outcomes in a children's hospital cohort. BMJ Qual Saf 2024 Jan 19;33(2):86-97. [doi: 10.1136/bmjqs-2022-015786] [Medline: 37460119]
59. Evenson KR, Wen F, Herring AH. Associations of accelerometry-assessed and self-reported physical activity and sedentary behavior with all-cause and cardiovascular mortality among US adults. Am J Epidemiol 2016 Nov 1;184(9):621-632. [doi: 10.1093/aje/kww070] [Medline: 27760774]
60. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364. [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]
61. Copeland JL, Esliger DW. Accelerometer assessment of physical activity in active, healthy older adults. J Aging Phys Act 2009 Jan;17(1):17-30. [doi: 10.1123/japa.17.1.17] [Medline: 19299836]
62. Evenson KR, Catellier DJ, Gill K, Ondrak KS, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci 2008 Dec;26(14):1557-1565. [doi: 10.1080/02640410802334196] [Medline: 18949660]
63. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [doi: 10.1097/00005768-199805000-00021] [Medline: 9588623]
64. Lee P, Tse CY. Calibration of wrist-worn ActiWatch 2 and ActiGraph wGT3X for assessment of physical activity in young adults. Gait Posture 2019 Feb;68:141-149. [doi: 10.1016/j.gaitpost.2018.11.023] [Medline: 30476691]
65. Powell C, Carson BP, Dowd KP, Donnelly AE. Simultaneous validation of five activity monitors for use in adult populations. Scand J Med Sci Sports 2017 Dec;27(12):1881-1892. [doi: 10.1111/sms.12813] [Medline: 27905148]
66. Rhudy MB, Dreisbach SB, Moran MD, Ruggiero MJ, Veerabhadrappa P. Cut points of the Actigraph GT9X for moderate and vigorous intensity physical activity at four different wear locations. J Sports Sci 2020 Mar;38(5):503-510. [doi: 10.1080/02640414.2019.1707956] [Medline: 31865845]
67. van Hees VT, Gorzelniak L, Dean León EC, et al. Separating movement and gravity components in an acceleration signal and implications for the assessment of human daily physical activity. PLoS One 2013;8(4):e61691. [doi: 10.1371/journal.pone.0061691] [Medline: 23626718]
68. van Hees VT, Renström F, Wright A, et al. Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. PLoS One 2011;6(7):e22922. [doi: 10.1371/journal.pone.0022922] [Medline: 21829556]
69. Hildebrand M, VAN Hees VT, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. Med Sci Sports Exerc 2014 Sep;46(9):1816-1824. [doi: 10.1249/MSS.0000000000000289] [Medline: 24887173]
70. Colley RC, Tremblay MS. Moderate and vigorous physical activity intensity cut-points for the Actical accelerometer. J Sports Sci 2011 May;29(8):783-789. [doi: 10.1080/02640414.2011.557744] [Medline: 21424979]
71. Hildebrand M, Hansen BH, van Hees VT, Ekelund U. Evaluation of raw acceleration sedentary thresholds in children and adults. Scand J Med Sci Sports 2017 Dec;27(12):1814-1823. [doi: 10.1111/sms.12795] [Medline: 27878845]
72. Moore R, Archer KR, Choi L. Statistical and machine learning models for classification of human wear and delivery days in accelerometry data. Sensors (Basel) 2021 Apr 13;21(8):2726. [doi: 10.3390/s21082726] [Medline: 33924388]
73. Migueles JH, Cadenas-Sanchez C, Ekelund U, et al. Accelerometer data collection and processing criteria to assess physical activity and other outcomes: a systematic review and practical considerations. Sports Med 2017 Sep;47(9):1821-1845. [doi: 10.1007/s40279-017-0716-0] [Medline: 28303543]

XSL•FO
RenderX

74. Migueles JH, Cadenas-Sanchez C, Tudor-Locke C, et al. Comparability of published cut-points for the assessment of physical activity: implications for data harmonization. Scand J Med Sci Sports 2019 Apr;29(4):566-574. [doi: 10.1111/sms.13356] [Medline: 30548545]
75. Welcome to OpenStreetMap!. OpenStreetMap. URL: https://www.openstreetmap.org [accessed 2025-12-19]

Abbreviations

ACMT: Automatic Context Measurement Tool
ACS: American Community Survey
CDC: Centers for Disease Control and Prevention
InSPACE : Interventions Supporting Physical Activity and the Environment
IRB: institutional review board
MVPA: moderate-to-vigorous physical activity
NIH: National Institutes of Health
NLCD: National Land Cover Database
PA: physical activity
SAC: scientific advisory council

Edited by J Sarvestan; submitted 28.Aug.2025; peer-reviewed by RQ Wang, T Prochnow; revised version received 04.Dec.2025; accepted 05.Dec.2025; published 22.Jan.2026.

Please cite as:
Youngbloom AJ, Rowland M, Mooney SJ, Singer K, Szpiro A, Saelens BE
Investigating the Role of the Environment on Physical Activity Interventions (the InSPACE Project): Protocol for a Pooled Secondary Analysis of Randomized Controlled Trials
JMIR Res Protoc 2026;15:e83151
URL: https://www.researchprotocols.org/2026/1/e83151
doi:10.2196/83151

© Amy J Youngbloom, Maya Rowland, Stephen J Mooney, Kaitlin Singer, Adam Szpiro, Brian E Saelens. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 22.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Norwegian Adult Mental Health Registry for Quality Control in Specialized Mental Health Services: Protocol for a Nationwide Naturalistic Study

Tore Tjora[1,2], PhD; Kent Jensen[3]; Ingrid Dieset[4,5], MD, PhD; Katrine Kveli Fjukstad[6,7], MD, PhD; Per Arne Holman[8], MA; Martin Tesli[5,9,10], PhD; Inge Joa[1,11], PhD

[1]TIPS - Network for Clinical Research in Psychosis, Stavanger University Hospital, Gerd-Ragna Bloch Thorsens Gate 8, Stavanger, Norway

[10]Department of Mental Health, Norwegian Institute of Public Health, Oslo, Norway

[11]Faculty of Health, University of Stavanger, Stavanger, Norway

[2]Department of Social Studies, Faculty of Social Sciences, University of Stavanger, Stavanger, Norway

[3]Universitetssykehuset i Nord Norge, Tromsø, Norway

[4]Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

[5]Department of Adult Psychiatry, Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[6]Department of Psychiatry, Levanger Hospital, Nord-Trøndelag Hospital Trust, Levanger, Norway

[7]Department of Mental Health, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway

[8]Department of Patient Safety and Research, Lovisenberg Diakonale Sykehus AS, Oslo, Norway

[9]Division of Mental Health and Substance Abuse, Diakonhjemmet Hospital, Oslo, Norway

Corresponding Author:
Tore Tjora, PhD
TIPS - Network for Clinical Research in Psychosis, Stavanger University Hospital, Gerd-Ragna Bloch Thorsens Gate 8, Stavanger, Norway

Abstract

Background: Mental disorders are highly prevalent, and they significantly impact individuals and society. Patients experiencing long-term, severe mental disorders with functional impairment and reduced quality of life often have a history of adolescent onset anxiety and depressive disorders. Despite the long-term costs to both patients and society, studies examining treatment effects over time and across diagnoses are scarce.

Objective: The Norwegian Adult Mental Health Registry (NAMHR) aims to systematically reuse health data to monitor and improve treatment outcomes, patient safety, health service quality, and research. The registry addresses the need for comprehensive data on the effects and utility of mental health services, interventions, and therapy variants in specialized mental health care.

Methods: The NAMHR is a nationwide naturalistic registry, including all Norwegian adults eligible for treatment in specialized mental health care services who have not opted out. Patients are automatically enrolled when treated in these services. The population includes patients treated in public specialized services and those treated in private services having a contract with public health services. The registry is based on secondary data from the Norwegian Patient Registry (NPR), patient-reported outcome measures (PROMs), patient-reported experience measures (PREMs), the Norwegian Registry for Primary Health Care (KPR), and several other sources, including electronic health records (EHRs). Data linkage uses unique national identity numbers, ensuring high-quality data. The registry collects information on diagnoses, treatments, medication, and patient-reported outcomes, providing a holistic approach to mental health care. Statistical analyses will be defined in each project.

Results: As of December 2025, the NAMHR is approved and is being constructed. The registry anticipates enrolling up to 170,000 participants, with a new incidence rate of around 10,000 patients per year. Key predictors and outcomes include PROMs and PREMs, and automatically reported measures involving a wide range of data, including EHR data from inpatient and outpatient treatments, data from primary health care, data on job and education status, and data on cause of death. Enrollment is planned to start in 2026, initially by adding journal data and patient-reported data. Other sources will be included. The NAMHR has no planned end date. Results will be made available for internal quality improvement purposes, and data for research are expected to be available around mid-2026 for approved projects.

Conclusions: The NAMHR will promote quality improvement initiatives and research, including registry-based randomized clinical trials. It will also be possible to link the NAMHR to a similar registry for children and adolescents, making it possible to follow patients from birth to death and supporting the monitoring of diagnostic drift. The NAMHR will inform health policy

XSL•FO
RenderX

decisions at local, regional, national, and international levels, contributing to the evaluation and development of clinical guidelines and enhancing personalized treatment approaches.

Trial Registration: ClinicalTrials.gov NCT06115200; https://clinicaltrials.gov/study/NCT06115200

International Registered Report Identifier (IRRID): PRR1-10.2196/82696

(JMIR Res Protoc 2026;15:e82696) doi:10.2196/82696

KEYWORDS

mental illness; global burden of disease; national quality registry; adult patient; quality improvement; quality indicators; patient-reported outcome measures; data reuse

Introduction

Although mental disorders are highly prevalent and impact the lives of affected individuals and their next of kin, transdiagnostic nationwide quality-improvement programs are scarce. Two-thirds of mental disorders occur before the age of 25 years [1,2]. In both high-risk populations and the general population, mental disorders that develop in adolescence persist into adulthood, cause functional impairment, and reduce quality of life [3,4]. Mental disorders now account for at least 45% of the overall disease burden in those aged 10 - 24 years [5,6] and represent a substantial burden at the societal level [7]. Early onset is associated with a longer duration of illness and additional comorbid disorders [8]. In particular, anxiety and depressive disorders rank high with regard to nonfatal health issues [9].

A quality registry is a systematic collection of data used to monitor and improve treatment outcomes, patient safety, and health service quality. In Norwegian national quality registers, the term “quality” especially refers to data accuracy, reliability, and relevance. To ensure a high level of data quality, the data must be complete, consistent, and up to date. A registry with good data, relevant measures, and relevant and accessible reports offers opportunities for policy makers, service managers, and researchers. Good quality also entails the active use of registries in health service improvement and innovation. The Norwegian Adult Mental Health Registry (NAMHR) will merge nationwide data sources, including patient-reported data, and is expected to provide new and unique data for both quality improvement and research. The data from the registry will contribute to the identification of variations in practice and the assessment of whether health services meet established quality criteria.

Established quality criteria include high effectiveness and safety, user involvement, resource coordination and utilization, high accessibility, and fair distribution. The NAMHR will provide clinicians and leaders with updated and accessible information on the effects and utility of health services, including specific interventions, therapy variants, and compulsory treatment use. This is important as specialized mental health care faces several challenges, and mental disorders are already the most expensive category of disorders in Norway [10,11].

The extent to which evidence-based clinical guidelines are implemented in daily clinical practice for specialized mental health care is unknown, and thus, it is difficult to monitor unwarranted variation [12]. Currently available data in Norway report on patient rates, admission rates, bed days, number of consultations, variation in staff specialities, rejected referrals, and waiting times. There is little information on patients’ perspectives. Although the volume, expertise, and availability of treatment are increasing, little is known about the effects of these increases on patients’ outcomes in the short and long term. There is a national need to follow patients’ pathways and progress; demonstrate and document treatment outcomes and effectiveness; perform quality control evaluations of treatment outcomes and effectiveness; and determine how treatments lead to changes in patients’ symptoms, functioning, and quality of life [13]. Furthermore, there is a need to uncover unwarranted variance across units, hospitals, and health regions. For example, a report from 2023 on attention-deficit/hyperactivity disorder (ADHD) diagnosis revealed large between-clinic variation in the incidence rate of ADHD [14].

Further adding to the importance of quality control is the evidence of increasing rates of mental disorders, including anxiety, depressive symptoms, psychological distress, self-harm, and suicide, among young people since early 2010 [15-21]. A recent Australian study of mental health and well-being (2020 - 2022) revealed a 50% increase in mental disorders at the diagnosis level among people aged 16 - 24 years since 2007, with an annual prevalence rate of 39% from 2020 to 2022 [16]. Similar but weaker tendencies are found in Norway, as reported in the SHOT Study (Students’ Health and Wellbeing Study) [22] and HUNT Study (Trøndelag Health Study) [23]. The presence of mental disorders among young people is a largely ignored risk factor for age-related medical illnesses [24]. The lifetime risk in a population that meets the criteria for a diagnosable mental disorder is estimated to be 50% [25]. If these disorders are not treated, they can become major causes of reduced quality of life, high costs to the family, and high disability-adjusted life years [26]. Suicide remains the fourth most common cause of death among people aged 15 - 29 years globally [5], and mental disorders are some of the most frequent reasons for visits to general practitioners. Patients with mental disorders have 5 - 15 life-years lost (LYL) [27], and patients with serious mental illnesses, such as schizophrenia, have 9 - 12 LYL [28]. The most common causes of death overall are natural causes, such as cardiovascular diseases and cancer, many of which are preventable and treatable [29].

In real-life clinical psychiatric settings, it is relatively common for diagnoses to change over time. An individual receiving an anxiety diagnosis in early adolescence might be assessed as having depression in late adolescence before being diagnosed with schizophrenia in their mid-20s. Adding to this, the diagnostic criteria themselves have changed, including the

XSL•FO
RenderX

Diagnostic and Statistical Manual of Mental Disorders (*DSM*) [30] and the *International Statistical Classification of Diseases and Related Health Problems* (*ICD*) [31]. The individual diagnostic "drift" and criteria "drift" highlight the importance of transdiagnostic quality registers. Diagnostic drift also highlights the importance of incorporating new models or frameworks in specialized mental health quality control, and a suitable model is clinical staging. The clinical staging model [32-36] (Figure 1) presents a new and important longitudinal perspective. Clinical staging is congruent and synergistic with many developmental and related conceptual frameworks and allows for exploring mechanisms underpinning the onset and course of mental disorders. The clinical staging model places people along a multidimensional continuum from health to illness in order to capture elements of risk, onset, course, and prognosis. This transdiagnostic model [32,36-38] recognizes the fluid nature of mental disorders and distinguishes between early clinical stages (with low rates of progression to more severe disorders) and later stages (with higher rates of persistence and impairment).

Figure 1. Heuristic clinical staging model for mental disorders (reproduced with permission from McGorry and Mei [36]).

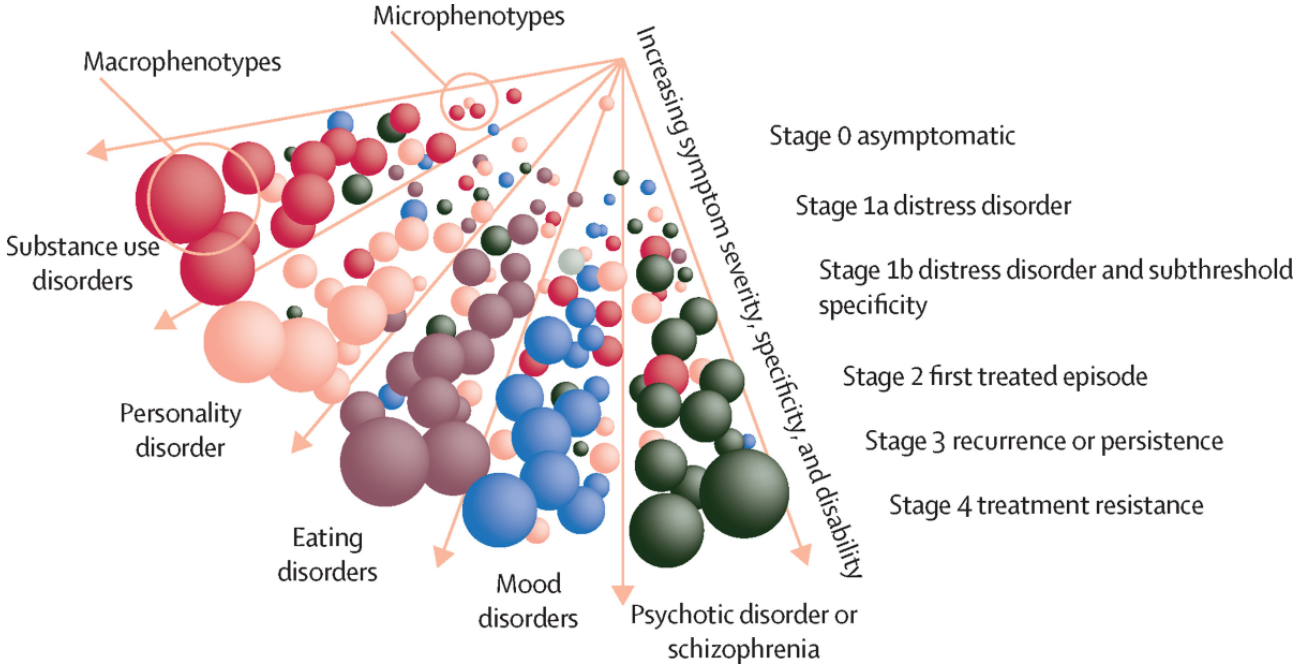

The model illustrates clusters of early symptoms (microphenotypes) and their potential progression into clear and often comorbid syndromes (macrophenotypes). This progression across stages is characterized by increasing symptom severity, specificity, and disability (represented by the sphere size; the spheres and colors represent phenotypes).

The clinical staging model is suitable as an overarching framework for the NAMHR to guide the ongoing selection of unwarranted variations to monitor. Currently, even established diagnostic-specific Norwegian care pathways [39] lack implementation studies. These care pathways specify time frames for assessment, treatment, and clinical evaluation and are based on previous findings that show positive prognostic effects of early diagnosis and treatment [40]. In line with clinical staging, the purpose of care pathways is to improve services across functions, activities, and units, as well as the wider health and welfare service system. The care pathway packages follow the same principles as those previously implemented for cancer [41]. Collecting data on care pathways enables the NAMHR to observe continuous chains of treatment and follow-up across service providers and to provide data on unnecessary delays in assessment, treatment, care, and follow-up in order to contribute to equal quality of care across geographical regions. Findings from several care pathways may provide data for a more overarching model like the clinical staging model. For example, poor coordination of medical health care has been associated with early mortality and a lack of physical health checks, prescriptions, and medical procedures [42-44]. More importantly, it is unknown whether and to what extent care pathways improve mental, medical, and functional outcomes. Additionally, there are no current national quality indicators that measure patients' mental health treatment outcomes, and such indicators are not linked to the degree of compliance with care pathways.

Patient-reported measures are important in health care in general, but patients' own experiences are especially important in mental health care as mental disorders manifest as challenges and disturbances in the experience of oneself and the world. In research settings, standardized and validated self-report measures often define outcomes. In mental health care, other outcomes, such as diagnoses and waiting times, are more frequent. The systematic collection of patient-reported outcome measures (PROMs) [45] and patient-reported experience measures (PREMs) [46,47] is increasing. The PROMs subgroup of routine outcome monitoring (ROM) involves clinical data primarily intended as part of a medical journal for individual documentation and treatment. Great Britain's public health services have used such ROM data to document treatment effects [48,49] at the individual level. However, aggregated ROM data may be used as process measures in quality registries to detect

XSL•FO
RenderX

unwarranted variations. Overall, the patient perspective is emphasized as an important aspect of quality in health services. To enhance the patient perspective, the National Mental Health Service Organization (Mental Helse Norge [50]) participated in the development of the NAMHR and has 2 members on the advisory board. The patient perspective includes the patient's satisfaction with treatment and provides subjective measures of effect and change. In mental health care, several PROMs are at a level of precision to be accepted as outcome variables in a quality registry [48], but are not sufficiently used at the system level. In Norway, this need is highlighted in the government's (Stortinget's) report on quality and patient safety [51].

The NAMHR's purpose is to provide a platform for transdiagnostic, nationwide quality improvement. One such issue is increased knowledge of "personalized treatment" in adults (eg, what works, when, where, in what way, and for whom). This is primarily fulfilled by facilitating the statutory duty for quality improvement that specialist health care providers have, and secondarily by delivering data for research. Specialized mental health care services in Norway are divided into child and adolescent services and adult services, resulting in 2 separate registers. Given the abovementioned importance of early interventions and longitudinal perspectives, this is suboptimal. However, as both registers have a common system architecture, ethically approved research studies may use merged data, following individuals from birth to death.

Methods

Study Design

The sample will include individuals who have been referred to specialized mental health care services and have not opted out. As individuals may be included at any time and included individuals may file reservations (opt out) at any time, both the population size and sample size will vary. The total population was approximately 170,000 participants as of 2023 [52]. The population includes patients treated at public specialized mental health care services and similar private services having a contract with public health services (contract specialists). Mental health services funded outside public health services (eg, privately or through insurance) are not considered. According to the South-Eastern Norway Regional Health Authority (the largest of the Norwegian health authorities), only 1% of licensed health personnel were working in private specialized health services in 2025 [53], which indicates that the vast majority of specialized health care services are funded publicly in Norway. According to the Norwegian Patient Registry (NPR) [54], the yearly treated incidence rate is estimated at approximately 10,000 patients. The start of data collection has been planned for autumn 2025. Patients treated prior to the start will be considered as new cases, resulting in an initial overestimation of incidence rates. The degree of registry coverage is defined as the number of nonreserved participants divided by the number of eligible participants. The coverage is currently unknown as neither data collection nor reservation has started. The NAMHR anticipates similar coverage as that of the reservation-based Norwegian Diabetes Register for Adults, which reported a nationwide coverage of 88% in 2023 [55].

The NAMHR was approved in late 2023 and is designed as a national naturalistic follow-along study based on all Norwegian adults eligible for treatment in specialized mental health care services, who have not opted out of participation (detailed inclusion and exclusion criteria are provided below). This protocol is compliant with the standards outlined in the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 Statement [56]. The SPIRIT checklist is provided in Checklist 1. The trial is registered at ClinicalTrials.gov (NCT06115200) [57,58].

Patient and Public Involvement

The NAMHR is run by Stavanger University Hospital and managed by an advisory board. The advisory board consists of clinicians and researchers from all 4 regional health trust organizations in Norway, representatives from work organizations (psychiatrists, psychologists, and nurses), and representatives from large Norwegian organizations for users and their relatives. The user (patient) representatives have been actively involved from the start and have contributed to many areas, especially quality topic selection and prioritization. The interests of the public are maintained through representatives from the 4 regional health trusts.

Population

Inclusion Criteria

All individuals referred for treatment to specialized mental health care services, who have not already opted out, will be included. Patients will be informed about the registry as part of the general information received by all patients with the right to undergo treatment at specialized mental health care services. To allow patients time to opt out, data from new eligible patients will not be used for the first 30 days after inclusion. If new eligible patients do not opt out within 30 days, their data, including data from the first 30 days, will be included in the registry.

Exclusion Criteria

Patient Reservation

Patients who opt out (reserve themselves) will be excluded. Patients may remove or reserve themselves at any time, including prior to any contact with specialized mental health services and possible registry inclusion. Patients may remove their reservation at any time. Patients who remove their reservation will be included in the registry again; however, data from the reservation period will not be regenerated.

Patients Already Included in Other Registries

Patients between 18 and 21 years of age who are already included in the KVABUP registry (Child and Adolescent Mental Health Registry) will be excluded. Moreover, patients older than 65 years who are included in the KVALAP registry (Quality Register for Old Age Psychiatry Patients) will be excluded.

Recruitment and Follow-Up Procedures

As the registry is based on the secondary use of data, participants will be recruited and followed up as part of their treatment. Treatment will not be affected in any way by inclusion or

XSL•FO
RenderX

exclusion in the registry, and the registry will not initiate contact or follow-up, with 2 exceptions. First, if the registry discovers data that are clearly incorrect (eg, a height of more than 3 meters), the registry will contact the data owner for correction. Second, the registry will follow up with patients who contact the registry directly. The registry has established procedures to handle requests for correcting, deleting, and viewing data of individual participants and has introduced procedures for patients to remove themselves or other patients under their care from inclusion.

Data Sources

The NAMHR uses previously collected data (secondary use) from mental health care services and combines data from other sources, such as government records and registries (Figure 2). The main information source for the NAMHR is the NPR [54]. The NPR contains comprehensive information about individuals referred to or treated by specialized health care providers like hospitals, outpatient clinics, and contract specialists. The NPR includes personal details (eg, age, sex, and address), claim data (eg, provider information, episodes of care, patient rights, time of death, and registry consent), medical information (eg, medical specialty, diagnoses, procedures, timestamps on episodes of care, compulsory treatment, and substance use), and social information (eg, housing conditions and care situations). The data in the NPR stem from mandatory monthly reports from all the public health trusts, private and nonprofit organizations under the public health trusts (contract), and private contract specialists [59]. The reports are automatically generated from the electronic health records (EHRs) and reflect the interdisciplinary collaboration among medical doctors, psychologists, nurses, and individuals from other disciplines entering data in the EHRs.

The NAMHR will enrich the NPR data by merging multiple other sources. To enhance patient perspectives and experiences, PROMs and PREMs are included. PROMs and PREMs are currently collected for individual treatments. The NAMHR plans to integrate all platforms of PROMs and PREMs used in Norway.

The NAMHR will also merge data from hospital EHRs and data from the Norwegian Registry for Primary Health Care (KPR) [60]. The latter provides information on the direct and indirect contact of general practitioners with patients. Further, the NAMHR has been authorized to merge other sources (details are provided in Figure 2).

As the NAMHR is primarily a tool for quality improvement in specialized mental health care in Norway, both exposure and outcome variables may change as the need for quality indicators changes over time. The list of exposure and outcome variables should be considered tentative.

XSL•FO
RenderX

Figure 2. Norwegian Adult Mental Health Registry data flow diagram. PREMs: patient-reported experience measures; PROMs: patient-reported outcome measures.

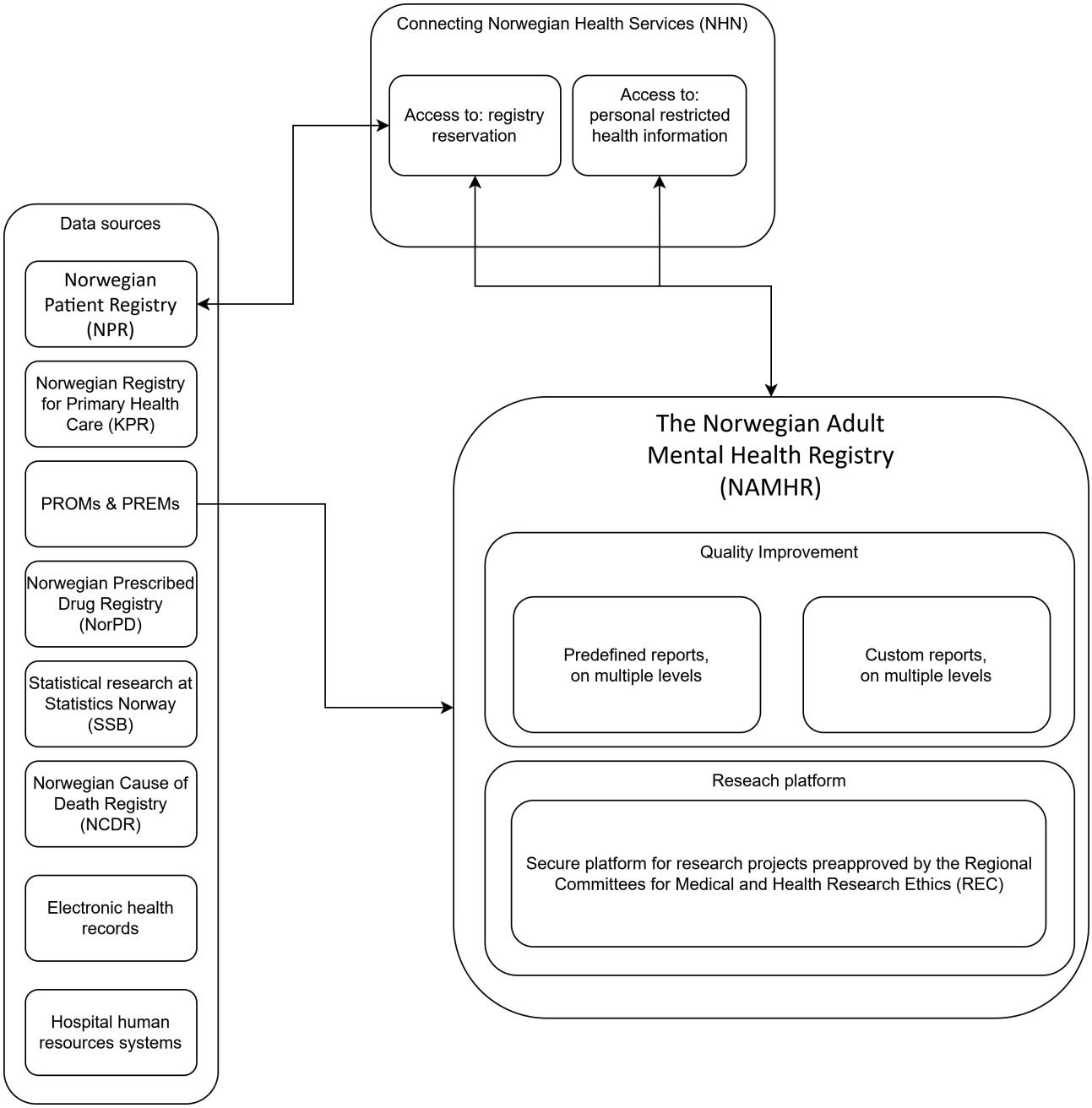

Outcomes: Quality Indicators

The purpose of the national quality indicator system [61] is to ensure equal access to good-quality health services for the population based on the framework of the Health Care Quality and Outcomes program of the Organisation for Economic Co-operation and Development [62-64]. The national quality indicators serve several purposes: (1) they provide central health authorities with a basis for monitoring, prioritization, and management; (2) they provide managers at all levels with a basis for quality improvement; and (3) they contribute to public transparency on quality and variation in services.

The NAMHR's quality indicators are dynamic, as they aim to reflect quality challenges in mental health care for adults at any given time. The register's proposed quality indicators (Table 1) are based on and comply with recommendations in national guidelines. As of 2025, in Norway, there are five national guidelines in the field of mental health (adults): (1) National professional guidelines for ADHD/hyperkinetic disorder [65]; (2) National professional guidelines for the prevention of suicide in mental health care [66]; (3) National professional guidelines for the early detection, assessment, and treatment of eating disorders [67]; (4) National professional guidelines on the use of electroconvulsive therapy [68]; and (5) Guidelines on psychotic disorders and medications [69].

XSL•FO
RenderX

Table . Initial Norwegian Adult Mental Health Registry quality indicators.

Overarching theme and quality indicator (measure)	Type	Expected level for the measure
Examination and diagnostics		
Proportion of incident (new) patients diagnosed with structured diagnostic instruments (eg, SCID-CV[a] and MINI[b] 7.0.2)	Process	≥90%
Somatic health and cardiovascular risk		
Proportion of incident (new) patients whose BMI is measured	Process	≥90%
Proportion of incident (new) patients whose lipid and HbA_{1c} levels are measured	Process and results	≥90%
Proportion of incident (new) patients whose blood pressure is measured	Process and results	≥90%
Outcome/results		
Proportion of patients who use the mental health feedback tool for adults	Process and results	≥90%
Similar results for comparable groups (CORE-OM[c]) [70]	Results	Group difference in improvement >5 points
Improved quality of life (WHODAS[d]) [71]	Results	Share of participants who report improved quality of life

[a]SCID-CV: Structured Clinical Interview for *DSM-5*, Clinician Version.

[b]MINI: Mini International Neuropsychiatric Interview.

[c]CORE-OM: Clinical Outcomes in Routine Evaluation – Outcome Measure.

[d]WHODAS: World Health Organization Disability Assessment Schedule.

Measures

Individual and Structural Variables (Predictors/Covariates)

Individual variables reported at baseline include age, gender, tentative diagnosis, and reason for referral. Structural variables consist of system-level variables such as the name of the institution, health region, and county. System-level variables collected from hospital record systems include staff size, personnel turnover, tenure, the percentage of doctors who are psychiatrists, the percentage of psychologists with relevant specializations, and the number of available hospital rooms. The list of variables is tentative, may vary by health region, and may change over time.

Process Variables (Mediators/Moderators)

Data on processes will be collected at the individual and system levels. Individual-level data include the type and length of treatment, and structured data from EHRs, including BMI, blood pressure, and the type and dosage of medication. Additionally, PREMs and PROMs will be collected. Examples are detailed in Textbox 1. The process variables also include time-variant variables at the system level. For example, case-load indicators may be defined as the number of consultations registered divided by the number of employees.

XSL•FO
RenderX

Textbox 1. Example structural, process, and outcome variables in the Norwegian Adult Mental Health Registry.

Predictor variables (independent); individual and structural levels

- Symptoms and function at baseline; patient-reported outcome measures (PROMs) (Clinical Outcomes in Routine Evaluation–Outcome Measure [CORE-OM], Alcohol Use Disorders Identification Test [AUDIT] [72], and Drug Use Disorders Identification Test [DUDIT] [73])
- Age (Norwegian Patient Registry [NPR]) and gender (NPR)
- Referral date (NPR)
- Assessment date (NPR)
- Start date, both outpatient and inpatient (NPR)
- Unit name and description (NPR)
- Prior referral/treatment (NPR)
- Prior episode of compulsory mental health care (NPR)
- Suicide assessment (NPR)
- System-level variables, such as turnover, tenure, percentage of doctors who are psychiatrists, percentage of psychologists with relevant specializations, and number of available hospital rooms

Process variables; individual level

- Symptoms and function at multiple times during treatment; PROMs (CORE-OM, AUDIT, and DUDIT)
- Hours spent on outpatient treatment
- End date (NPR)
- Duration of treatment (NPR)
- Use of structured symptom and function assessment instruments (NPR)
- Structured codes for diagnoses and procedures from the national patient registry (NPR)
- Quality of life assessment (World Health Organization Disability Assessment Schedule [WHODAS])
- Ongoing compulsory mental health care (NPR)
- Suicide assessment
- Type of given treatment, standardized procedure codes (NPR)
- Somatic: blood pressure, height, weight, and other relevant tests (electronic health record [EHR])
- Medication: type, administration, dosage (Norwegian Prescribed Drug Registry [NorPD])
- System-level variables, such as turnover, tenure, percentage of doctors who are psychiatrists, percentage of psychologists with relevant specializations, and number of available hospital rooms

Result variables (outcome); individual and structural levels

- Symptoms and function at discharge; PROMs (CORE-OM, AUDIT, and DUDIT)
- PROMs, quality of life change (WHODAS)
- Patient-reported experience measures; overall satisfaction with mental health care (Pasient Opplevelser - patients' satisfaction with treatment [PASopp])
- Mortality rate and cause of death (Norwegian Cause of Death Registry [NCDR])
- Outcome measures regarding primary care follow-up, such as individual treatment plan, general practitioner assessment, municipality coordinator, type of work, and ongoing education (Norwegian Registry for Primary Health Care [KPR])
- Compulsory mental care as part of aftercare (NPR)
- Suicide assessment
- Medication type, administration, and dosage in epicrisis on discharge (EHR and NorPD)

Outcome Variables

PROMs are registered at baseline, multiple times during the intervention, at postintervention, at discharge, and 6 months after discharge. The PROM battery may vary across health regions, and a reliable change index score [74] may be used for comparison. Outcome variables at discharge include discharge diagnosis, treatment type, treatment length, treatment frequency, region ID, and facility ID. The list of outcome variables is tentative, may vary from one health region to another, and may

XSL•FO
RenderX

change over time. However, the abovementioned variables are expected to be stable over time and place.

Statistical Analysis

Statistical analyses will vary from project to project, and for both internal quality improvement projects and published research papers, a wide range of statistical approaches may be used. "Statistical process control" [75] may be used, as it allows close monitoring of longitudinal changes in complex systems. Classic approaches, ranging from descriptive cross-sectional studies (case-control and cohort studies) to randomized controlled trials (RCTs) based on registry data may also be performed. Further, the vast amount of data, especially many predictors from different sources, may suggest a more multilevel approach. For example, exploratory factor analysis may be used to explore factor structures, which, in combination with structural equation modeling, may be used to explore relationships between observed and unobserved (latent) variables. Latent class models and latent growth models [76] may be used to explore the difference in developmental trajectories. Exploratory studies and quality improvement projects may use forecast models based on time-series data from the registry, both to describe ongoing trends and plan the future of health care. The registry may also form the basis for machine learning approaches, such as neural networks, to both model very complex, longitudinal data in an understandable way and establish models for more precise prediction of mental disorders.

Future Registry Expansion

The registry has been approved for the inclusion of data from other national health-related registries, such as the Norwegian Prescribed Drug Registry (NorPD) [77], the Norwegian Cause of Death Registry (NCDR) [78], and the KPR [60]. The NorPD includes data on drugs prescribed by doctors in primary health care and drugs prescribed and administered as part of treatment in specialized health care. The NAMHR has also been approved for the inclusion of data from other national general registries, such as income, tax, and social benefits.

Data Linkage and Storage

The data will come from multiple sources (detailed above) and will be linked using the unique national identity number. Each patient will be given a random number for internal use. The data will be stored within the same infrastructure as most of the data sources that are included, owned, and run by Connecting Norwegian Health Services (Norsk helsenett) [79].

Risk and vulnerability assessments and data protection impact assessments [80] have been performed to ensure compliance with the General Data Protection Regulation (GDPR) [81] and will be revised in accordance with procedures in Helse Vest Health Care Trust.

Trial Duration

The registry has no planned end date.

Ethical Considerations

In Norway, legislation and regulation regarding quality improvement and research are handled separately. The registry is primarily part of the ongoing quality improvement program imposed by the Norwegian law on specialized health care, section 3.4a [82]. All national medical quality registries are appointed by the Norwegian Directorate of Health. The registry will be run by a board, as detailed earlier in the *Methods* section. Data collection and usage for quality improvement are regulated by the abovementioned laws and according to GDPR regulations and standards [83].

The participants will not receive compensation for any inconvenience or time spent. The NAMHR will reuse health data, and the participants do not have to take any additional effort. The participants have the right to opt out (reserve themselves), correct their data, and delete their data. The NAMHR will follow the European Union's data privacy law (GDPR), which will ensure privacy and confidentiality.

The data will be made available for use in research. In Norway, "all medical and health research projects must be preapproved by the *Regional Committees for Medical and Health Research Ethics*" [84]. Ethical approval must be obtained before the beginning of the project and data delivery from the NAMHR. The full study protocol and relevant statistical code may be provided on request.

The results of the research projects will be published in peer-reviewed journals and at research conferences. Authorship will follow the "Vancouver Recommendations." Open-access journals will be preferred. Important protocol modifications will be reported at ClinicalTrials.gov [57].

Results

The NAMHR was approved as a mandatory quality instrument in late 2023. As of December 2025, the NAMHR is approved and is being constructed. The study anticipates enrolling up to 170,000 participants, with a new incidence rate of around 10,000 patients per year. Key predictors and outcomes include both PROMs and PREMs, as well as automatically reported measures on a wide range of data, including EHR data from inpatient and outpatient treatments, data from primary health care, data on job and education status, and data on cause of death. The NAMHR has no planned end date. Data for both quality improvement and research projects are planned to be available from mid-2026.

Discussion

Overview

This paper outlines our study protocol for the NAMHR. The primary goal of the NAMHR is to monitor and develop future mental health care, and it will provide important knowledge that benefits all patients and the general population. The NAMHR provides a unique opportunity by gathering large-scale, high-quality, and representative data. As Norwegians have free access to public health insurance, the vast majority of Norwegians use public health care, resulting in highly representative data. Further, public-delivered or commissioned services require public and private providers to deliver standardized claim data to the NPR.

XSL•FO
RenderX

Hypothesized Main Findings

The registry may be used to investigate the effects of various treatments and as a nationwide control for RCTs. The NAMHR combines data from all major EHRs in Norway, ranging from data on medication and medication administration from general practitioners to high-temporal-resolution drug monitoring systems for inpatient treatment. Hence, the NAMHR will be one of the first nationwide data sources for monitoring longitudinal medication usage for individual patients and combining in- and outpatient prescriptions and psychological treatments provided. This may be done in combination with service provision variables and PROMs or PREMs, hence providing a unique opportunity for quality improvement. In addition, it will provide the opportunity to perform large-scale research on the long-term effects of psychological and drug treatments. The NAMHR will also be a platform for registry-based randomized clinical trials, which are prospective randomized trials that use a clinical registry for one or several major functions to conduct trials and report outcomes. The advisory board has already decided that involuntary treatment in mental health services is one of the first quality indicators that will be targeted for quality improvement and research.

Comparison to Prior Work

In contrast to the other Nordic mental health registries, the NAMHR is organized as a transdiagnostic mental health registry, which makes the NAMHR more suitable for investigating the fluid nature of diagnostics across the lifespan [38]. Further, it provides an opportunity to detect and end mental health interventions with low returns on investment [85]. The combination of a transdiagnostic approach and a large sample size provides a possibility of monitoring compliance with a wide range of clinical guidelines. The registry also represents an opportunity to inform new clinical guidelines, including guidelines involving more than one diagnostic category, a possibility few mental health registries possess. When unwarranted variation is revealed, the advisory board and each hospital trust may use this information to develop new quality indicators. These indicators and the assessment of their effects on unwarranted variation may result in new clinical guidelines based on large-scale empirical data. This is not new; however, automated data collection and merging may result in a much cheaper and faster process. When new guidelines are introduced, the system for monitoring guideline implementation and compliance is already up and running. Traditionally, discharged patients who were not seen again were assumed to have had a good course of treatment. However, the use of more objective data, such as sick leave and taxes paid, which are included in the NAMHR, can provide more accurate information.

Hence, data delivered from the NAMHR will be highly relevant for health policy decision-making at the local, regional, national, and international levels. For example, care pathway packages are implemented with limited evidence of their efficacy or effectiveness in mental health. Thus, it is relevant for policymakers to evaluate the effects on treatment delivery and outcomes. The wide variation in health service delivery is a cause for concern among policymakers at all levels [14], but traditional RCTs fall short of defining the best practices. Some important questions are as follows: Can high rates of involuntary treatment be justified by reduced suicide and violence rates, or does this reflect the unnecessary use of coercive measures? Do early intervention policies produce improved outcomes for all patients in the system or only for those selected for early intervention at the expense of other patients [86]? Can patient-generated data with personalized psychometrics enhance treatment processes, shared decision-making, and outcomes?

Strengths and Limitations

The NAMHR provides a relatively high temporal resolution with a longitudinal perspective and may be used to describe various trajectories or typical pathways through health services. In addition, the registry may be used to calculate which of the abovementioned trajectories are associated with preferable outcomes, ranging from subjective outcomes, such as PREMs and PROMs, to more objective outcomes, such as work status and taxes paid. The registry will be linked to the upcoming national quality registry for child and adolescent specialized mental health services for research purposes after the development of trajectories. As mentioned, the NAMHR is organized as a transdiagnostic mental health registry, which is also considered a strength.

The reuse of data is one of the greatest strengths of the NAMHR, but it can also be considered a major limitation. As the registry only reuses data collected for other purposes, the registry has no direct influence on what is collected. Another possible limitation is the unknown degree of reservation. If the percentage of patients who chose to reserve themselves (opt out) from the registry is large and/or associated with key predictors or outcomes, it will affect to what degree the findings are generalizable to other populations.

Dissemination Plan

Findings from the registry will have 2 main outlets. First, quality improvements will be monitored and reported through a dashboard available for health services and reports. Second, as mentioned above, data will be made available for projects with ethical approval from the Regional Committees for Medical and Health Research Ethics (REC) [84]. The NAMHR does not decide the publication channel of ethically approved research projects. However, a proper plan for dissemination is necessary for each project to obtain ethical approval from the REC. The NAMHR will strongly suggest using open-access journals. The NAMHR will publish new papers on ClinicalTrials.gov [57].

Conclusion

The NAMHR, in addition to other Nordic mental health registries and similar registries in other countries, may play an important role in psychiatric health care worldwide. The registry will deliver data for both quality improvement and research with real-world outcomes, including patient symptoms, functioning, and quality of life.

XSL•FO
RenderX

Acknowledgments

We extend our gratitude to the participants on the board, especially representatives from organizations for users and their relatives. The members of the board in 2025 were Ketil Ødegaard, Kenneth Lindberget, Ragnar Nesvåg, Fredrik Skarderud, Kenneth Sandin, Hilde B Steen, Karl Olaf Sundfør, and Marianne Lyche. The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy, the following tasks were delegated to GAI tools under full human supervision: summarizing text. The GAI tool used was Microsoft Copilot (GPT-4.5; August 2025). The responsibility for the final manuscript lies entirely with the authors. The GAI tool is not listed as an author and does not bear responsibility for the final outcomes. The authors take collective responsibility. Moreover, Microsoft Copilot (GPT-4.5) was used to generate an abstract draft.

Funding

The Norwegian Adult Mental Health Registry (NAMHR) is funded as part of public specialist mental health care through the regional Health Vest Trust; hence, the research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The NAMHR is administratively subject to the clinic director of adult mental health care at Stavanger University Hospital, Norway. The CEO of Stavanger University Hospital, Norway, is the responsible data controller (General Data Protection Regulation).

Data Availability

Data will be made available for research projects with approval from the Regional Committees for Medical and Health Research Ethics (REC). Strict regulations related to data management, storage, and processing will be enforced, including the possibility to only allow analysis and storage at Connecting Norwegian Health Services [79].

Authors' Contributions

The first version of the manuscript was drafted by TT and IJ. KJ, ID, KKF, PAH, and MT reviewed and commented on the manuscript and provided significant suggestions for all major parts of the article. All authors approved the final version.

Conflicts of Interest

None declared.

Checklist 1
SPIRIT checklist.
[PDF File, 146 KB - resprot_v15i1e82696_app1.pdf]

References

1. Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustün TB. Age of onset of mental disorders: a review of recent literature. Curr Opin Psychiatry 2007 Jul;20(4):359-364. [doi: 10.1097/YCO.0b013e32816ebc8c] [Medline: 17551351]
2. Solmi M, Radua J, Olivola M, et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol Psychiatry 2022 Jan;27(1):281-295. [doi: 10.1038/s41380-021-01161-7] [Medline: 34079068]
3. Mulraney M, Coghill D, Bishop C, et al. A systematic review of the persistence of childhood mental health problems into adulthood. Neuroscience & Biobehavioral Reviews 2021 Oct;129:182-205. [doi: 10.1016/j.neubiorev.2021.07.030]
4. Sivertsen B, O'Connor RC, Nilsen SA, et al. Mental health problems and suicidal behavior from adolescence to young adulthood in college: linking two population-based studies. Eur Child Adolesc Psychiatry 2024 Feb;33(2):421-429. [doi: 10.1007/s00787-023-02167-y] [Medline: 36843045]
5. Suicide. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/suicide [accessed 2026-01-08]
6. Patton GC, Sawyer SM, Santelli JS, et al. Our future: a Lancet commission on adolescent health and wellbeing. The Lancet 2016 Jun;387(10036):2423-2478. [doi: 10.1016/S0140-6736(16)00579-1]
7. Christensen MK, Lim CCW, Saha S, et al. The cost of mental disorders: a systematic review. Epidemiol Psychiatr Sci 2020 Aug 18;29:e161. [doi: 10.1017/S204579602000075X] [Medline: 32807256]
8. Caspi A, Houts RM, Ambler A, et al. Longitudinal assessment of mental health disorders and comorbidities across 4 decades among participants in the Dunedin Birth Cohort Study. JAMA Netw Open 2020 Apr 1;3(4):e203221. [doi: 10.1001/jamanetworkopen.2020.3221] [Medline: 32315069]
9. Tollånes MC, Knudsen AK, Vollset SE, Kinge JM, Skirbekk V, Øverland S. Disease burden in Norway in 2016. Tidsskr Nor Laegeforen 2018 Oct 2;138(15). [doi: 10.4045/tidsskr.18.0274] [Medline: 30277038]
10. Kinge JM, Dieleman JL, Karlstad Ø, et al. Disease-specific health spending by age, sex, and type of care in Norway: a national health registry study. BMC Med 2023 Jun 6;21(1):201. [doi: 10.1186/s12916-023-02896-6] [Medline: 37277874]
11. Kinge JM, Sælensminde K, Dieleman J, Vollset SE, Norheim OF. Economic losses and burden of disease by medical conditions in Norway. Health Policy 2017 Jun;121(6):691-698. [doi: 10.1016/j.healthpol.2017.03.020] [Medline: 28461038]

XSL•FO
RenderX

12. Sutherland K, Levesque JF. Unwarranted clinical variation in health care: definitions and proposal of an analytic framework. J Eval Clin Pract 2020 Jun;26(3):687-696. [doi: 10.1111/jep.13181] [Medline: 31136047]
13. Riksrevisjonens undersøkelse av psykiske helsetjenester [Article in Norwegian]. : Riksrevisjonen; 2021 URL: https://www.stortinget.no/no/Saker-og-publikasjoner/Publikasjoner/Dokumentserien/2020-2021/dok3-202021/dok3-202021-013/ [accessed 2026-01-08]
14. Widding-Havneraas T, Markussen S, Elwert F, et al. Geographical variation in ADHD: do diagnoses reflect symptom levels? Eur Child Adolesc Psychiatry 2023 Sep;32(9):1795-1803. [doi: 10.1007/s00787-022-01996-7] [Medline: 35585272]
15. Office of the Surgeon General (OSG). Protecting Youth Mental Health: The US Surgeon General'S Advisory: US Department of Health and Human Services; 2021. URL: https://www.ncbi.nlm.nih.gov/books/NBK575984/ [accessed 2026-01-08]
16. National study of mental health and wellbeing: 2020-2022. Australian Bureau of Statistics. 2023. URL: https://www.abs.gov.au/statistics/health/mental-health/national-study-mental-health-and-wellbeing/latest-release [accessed 2026-01-08]
17. Jensen HAR, Davidsen M, Møller SR, Román JEI, Kragelund K. Danskernes Sundhed – Den Nationale Sundhedsprofl 2021 [Article in Danish]: Sundhedsstyrelsens publikationer; 2022. URL: https://www.sst.dk/udgivelser/2022/danskernes-sundhed [accessed 2026-01-08]
18. Newlove-Delgado T, Marcheselli F, Williams T, Mandalia D, Davis J, McManus S, et al. Mental health of children and young people in England 2022. : NHS Digital; 2022 URL: https://openaccess.city.ac.uk/id/eprint/30558/1/Mental%20Health%20of%20Children%20and%20Young%20People%20in%20England%202022%20%20wave%203%20follow%20up%20to%20the%202017%20survey%20%20NDRS.pdf [accessed 2026-01-08]
19. Twenge JM, Joiner TE, Rogers ML, Martin GN. Increases in depressive symptoms, suicide-related outcomes, and suicide rates among U.S. adolescents after 2010 and links to increased new media screen time. Clin Psychol Sci 2018 Jan;6(1):3-17. [doi: 10.1177/2167702617723376]
20. Naghavi M, Global Burden of Disease Self-Harm Collaborators. Global, regional, and national burden of suicide mortality 1990 to 2016: systematic analysis for the Global Burden of Disease Study 2016. BMJ 2019 Feb 6;364:l94. [doi: 10.1136/bmj.l94] [Medline: 31339847]
21. Mental health of children and young people in England 2022 - wave 3 follow up to the 2017 survey. : NHS England; 2022 URL: https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2022-follow-up-to-the-2017-survey [accessed 2026-01-08]
22. Sivertsen B, Knudsen AKS, Kirkøen B, et al. Prevalence of mental disorders among Norwegian college and university students: a population-based cross-sectional analysis. Lancet Reg Health Eur 2023 Nov;34:100732. [doi: 10.1016/j.lanepe.2023.100732] [Medline: 37927428]
23. Krokstad S, Weiss DA, Krokstad MA, et al. Divergent decennial trends in mental health according to age reveal poorer mental health for young people: repeated cross-sectional population-based surveys from the HUNT Study, Norway. BMJ Open 2022 May 18;12(5):e057654. [doi: 10.1136/bmjopen-2021-057654] [Medline: 35584877]
24. Moffitt TE, Caspi A. Psychiatry's opportunity to prevent the rising burden of age-related disease. JAMA Psychiatry 2019 May 1;76(5):461-462. [doi: 10.1001/jamapsychiatry.2019.0037] [Medline: 30916735]
25. Mykletun A, Knudsen AK, Mathiesen KS. Psykiske lidelser i Norge: Et folkehelseperspektiv [Article in Norwegian]. : Nasjonalt folkehelseinstitutt; 2009 URL: https://www.fhi.no/publ/eldre/psykiske-lidelser-i-norge-et-folkeh/ [accessed 2026-01-08]
26. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine 2022 Dec;54:101675. [doi: 10.1016/j.eclinm.2022.101675] [Medline: 36193171]
27. Plana-Ripoll O, Pedersen CB, Agerbo E, et al. A comprehensive analysis of mortality-related health metrics associated with mental disorders: a nationwide, register-based cohort study. The Lancet 2019 Nov;394(10211):1827-1835. [doi: 10.1016/S0140-6736(19)32316-5]
28. Tesli M, Degerud E, Plana-Ripoll O, et al. Educational attainment and mortality in schizophrenia. Acta Psychiatr Scand 2022 May;145(5):481-493. [doi: 10.1111/acps.13407] [Medline: 35152418]
29. Gronholm PC, Chowdhary N, Barbui C, et al. Prevention and management of physical health conditions in adults with severe mental disorders: WHO recommendations. Int J Ment Health Syst 2021 Mar 3;15(1):22. [doi: 10.1186/s13033-021-00444-4] [Medline: 33658061]
30. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition: American Psychiatric Association Publishing; 2013. [doi: 10.1176/appi.books.9780890425596]
31. International Statistical Classification of Diseases and Related Health Problems 10th Revision. World Health Organization. 2019. URL: https://icd.who.int/browse10/2019/en [accessed 2026-01-08]
32. McGorry P, Nelson B. Why we need a transdiagnostic staging approach to emerging psychopathology, early diagnosis, and treatment. JAMA Psychiatry 2016 Mar;73(3):191-192. [doi: 10.1001/jamapsychiatry.2015.2868] [Medline: 26765254]
33. Iorfino F, Scott EM, Carpenter JS, et al. Clinical stage transitions in persons aged 12 to 25 years presenting to early intervention mental health services with anxiety, mood, and psychotic disorders. JAMA Psychiatry 2019 Nov 1;76(11):1167-1175. [doi: 10.1001/jamapsychiatry.2019.2360] [Medline: 31461129]
34. McGorry P, van Os J. Redeeming diagnosis in psychiatry: timing versus specificity. Lancet 2013 Jan 26;381(9863):343-345. [doi: 10.1016/S0140-6736(12)61268-9] [Medline: 23351805]

XSL•FO
RenderX

35. McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. Aust N Z J Psychiatry 2006 Aug;40(8):616-622. [doi: 10.1080/j.1440-1614.2006.01860.x] [Medline: 16866756]
36. McGorry PD, Mei C. Clinical staging for youth mental disorders: progress in reforming diagnosis and clinical care. Annu Rev Dev Psychol 2021 Dec 9;3(1):15-39. [doi: 10.1146/annurev-devpsych-050620-030405]
37. McGorry PD, Hartmann JA, Spooner R, Nelson B. Beyond the "at risk mental state" concept: transitioning to transdiagnostic psychiatry. World Psychiatry 2018 Jun;17(2):133-142. [doi: 10.1002/wps.20514] [Medline: 29856558]
38. McGorry PD, Mei C, Dalal N, et al. The Lancet Psychiatry Commission on youth mental health. Lancet Psychiatry 2024 Sep;11(9):731-774. [doi: 10.1016/S2215-0366(24)00163-9] [Medline: 39147461]
39. Ådnanes M, Høiseth JR, Magnussen M, Thaulow K, Kaspersen SL. Pakkeforløp for psykisk helse og rus – brukere, pårørende og fagfolks erfaringer [Article in Norwegian]. : SINTEF Digital; 2021 URL: https://www.sintef.no/publikasjoner/publikasjon/1961108/ [accessed 2026-01-08]
40. Forløpstid for utredning i psykisk helsevern for voksne [Article in Norwegian]. Helsedirektoratet. 2021. URL: https://www.helsedirektoratet.no/statistikk/kvalitetsindikatorer/psykisk-helse-for-voksne/forlopstid-for-utredning-i-psykisk-helsevern-for-voksne [accessed 2026-01-08]
41. Pakkeforløp for kreft (treatment packages for cancer) [Article in Norwegian]. Helsedirektoratet. URL: https://www.helsedirektoratet.no/search?searchquery=pakkeforl%C3%B8p+kreft&type=Nasjonalt+forl%C3%B8p [accessed 2026-01-08]
42. Haussleiter I, Emons B, Hoffmann K, Juckel G. The somatic care situation of people with mental illness. Health Sci Rep 2021 Mar;4(1):e226. [doi: 10.1002/hsr2.226] [Medline: 33364443]
43. Prior A, Vestergaard CH, Vedsted P, et al. Healthcare fragmentation, multimorbidity, potentially inappropriate medication, and mortality: a Danish nationwide cohort study. BMC Med 2023 Aug 15;21(1):305. [doi: 10.1186/s12916-023-03021-3] [Medline: 37580711]
44. Rodgers M, Dalton J, Harden M, Street A, Parker G, Eastwood A. Integrated care to address the physical health needs of people with severe mental illness: a mapping review of the recent evidence on barriers, facilitators and evaluations. Int J Integr Care 2018 Jan 25;18(1):9. [doi: 10.5334/ijic.2605] [Medline: 29588643]
45. Patient-reported outcome measures (PROMs). Canadian Institute for Health Information. URL: https://www.cihi.ca/en/patient-reported-outcome-measures-proms [accessed 2026-01-16]
46. Fernandes S, Fond G, Zendjidjian XY, et al. Measuring the patient experience of mental health care: a systematic and critical review of patient-reported experience measures. Patient Prefer Adherence 2020;14:2147-2161. [doi: 10.2147/PPA.S255264] [Medline: 33192054]
47. Olsen RV, Garratt AM, Iversen HH, Bjertnaes OA. Rasch analysis of the Psychiatric Out-Patient Experiences Questionnaire (POPEQ). BMC Health Serv Res 2010 Sep 28;10(1):282. [doi: 10.1186/1472-6963-10-282] [Medline: 20920164]
48. Gelkopf M, Mazor Y, Roe D. A systematic review of patient-reported outcome measurement (PROM) and provider assessment in mental health: goals, implementation, setting, measurement characteristics and barriers. Int J Qual Health Care 2022 Apr;34(Supplement_1):ii13-ii27. [doi: 10.1093/intqhc/mzz133]
49. Barkham M, De Jong K, Delgadillo J, Lutz W. Routine outcome monitoring (ROM) and feedback: research review and recommendations. Psychother Res 2023 Sep;33(7):841-855. [doi: 10.1080/10503307.2023.2181114] [Medline: 36931228]
50. Mental Helse Norge. URL: https://mentalhelse.no/ [accessed 2026-01-08]
51. Kvalitet og pasientsikkerhet 2016 [Article in Norwegian]. : Regjeringen - Helse og omsorgsdepartementet; 2016 URL: https://www.regjeringen.no/no/dokumenter/meld.-st.-6-20172018/id2581316/ [accessed 2026-01-08]
52. Aktivitetsdata for spesialisthelsetjenesten [Article in Norwegian]. : Norwegian Institute of Public Health; 2023 URL: https://www.fhi.no/globalassets/dokumenterfiler/rapporter/2024/aktivitet-phv-2023---spesialisthelsetjenesten.pdf [accessed 2026-01-08]
53. Forberedt på hardere konkurranse om ansatte [Article in Norwegian]. Helse Sør-Øst. 2025. URL: https://www.helse-sorost.no/nyheter/forberedt-pa-hardere-konkurranse-om-ansatte [accessed 2026-01-08]
54. Norwegian Patient Registry (NPR). The Norwegian Institute of Public Health. URL: https://helsedata.no/en/forvaltere/norwegian-institute-of-public-health/norwegian-patient-registry-npr [accessed 2026-01-08]
55. Norsk diabetesregister for voksne: Data fra diabetespoliklinikker - Diabetes type 1 - Årsrapport for 2023 [Article in Norwegian]. : Nasjonalt Servicemiljø for Medisinske Kvalitetsregistre; 2024 URL: https://www.kvalitetsregistre.no/49e637/siteassets/dokumenter/arsrapporter/diabetesregisteret/arsrapport-2023-norsk-diabetesregister-for-voksne-type-1.pdf [accessed 2026-01-08]
56. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 5;158(3):200-207. [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
57. Norwegian Adult Mental Health Registry (NAMHR). ClinicalTrials.gov. URL: https://www.clinicaltrials.gov/study/NCT06115200 [accessed 2026-01-08]
58. Norwegian Adult Mental Health Registry (NAMHR). WHO International Clinical Trials Registry Platform (ICTRP). URL: https://trialsearch.who.int/Trial2.aspx?TrialID=NCT06115200 [accessed 2026-01-08]

XSL•FO
RenderX

59. Bakken IJ, Ariansen AMS, Knudsen GP, Johansen KI, Vollset SE. The Norwegian Patient Registry and the Norwegian Registry for Primary Health Care: Research potential of two nationwide health-care registries. Scand J Public Health 2020 Feb;48(1):49-55. [doi: 10.1177/1403494819859737] [Medline: 31288711]
60. Norwegian Registry for Primary Health Care (KPR). The Norwegian Institute of Public Health. 2016. URL: https://helsedata.no/en/forvaltere/norwegian-institute-of-public-health/norwegian-registry-for-primary-health-care-kpr [accessed 2026-01-08]
61. National healthcare quality indicators. Helsedirektoratet. URL: https://www.helsedirektoratet.no/english/national-health-care-quality-indicators [accessed 2026-01-08]
62. Healthcare quality and outcomes. Organisation for Economic Co-operation and Development (OECD). URL: https://www.oecd.org/en/topics/policy-issues/healthcare-quality-and-outcomes.html [accessed 2026-01-08]
63. Arah OA, Klazinga NS, Delnoij DMJ, ten Asbroek AHA, Custers T. Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement. Int J Qual Health Care 2003 Oct;15(5):377-398. [doi: 10.1093/intqhc/mzg049] [Medline: 14527982]
64. Smith PM, Mossialos E, Papanicolas I, Leatherman S. Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects: Cambridge University Press; 2010. [doi: 10.1017/CBO9780511711800]
65. ADHD - Nasjonal faglig retningslinje [Article in Norwegian]. Helsedirektoratet. URL: https://www.helsedirektoratet.no/retningslinjer/adhd [accessed 2026-01-08]
66. Selvmordsforebygging i psykisk helsevern og tverrfaglig spesialisert rusbehandling (TSB) - Nasjonal faglig retningslinje [Article in Norwegian]. Helsedirektoratet. URL: https://www.helsedirektoratet.no/retningslinjer/selvmordsforebygging-i-psykisk-helsevern-og-tsb [accessed 2026-01-08]
67. Spiseforstyrrelser - Nasjonal faglig retningslinje [Article in Norwegian]. Helsedirektoratet. URL: https://www.helsedirektoratet.no/retningslinjer/spiseforstyrrelser [accessed 2026-01-08]
68. Nasjonal Faglig Retningslinje Om Bruk Avelektrokonvulsiv Behandling - ECT [Book in Norwegian]: Helsedirektoratet; 2017. URL: https://www.helsedirektoratet.no/retningslinjer/elektrokonvulsiv-behandling-ect/Elektrokonvulsiv%20behandling%20(ECT)%20%E2%80%93%20Nasjonal%20faglig%20retningslinje.pdf [accessed 2026-01-08]
69. Psykoselidelser – legemiddelbehandling - Nasjonal faglig retningslinje [Article in Norwegian]. Helsedirektoratet. URL: https://www.helsedirektoratet.no/retningslinjer/psykoselidelser-legemiddelbehandling [accessed 2026-01-08]
70. Evans C, Mellor-Clark J, Margison F, Barkham M, Audin K, Connell J, et al. CORE: clinical outcomes in routine evaluation. J Ment Health 2000 Jan;9(3):247-255. [doi: 10.1080/jmh.9.3.247.255]
71. Üstün TB, Kostanjsek N, Chatterji S, Rehm J. Measuring Health and Disability: Manual for WHO Disability Assessment Schedule (WHODAS 20) : World Health Organization; 2012. URL: https://www.who.int/publications/i/item/measuring-health-and-disability-manual-for-who-disability-assessment-schedule-(-whodas-2.0) [accessed 2026-01-08]
72. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT Alcohol Consumption Questions (AUDIT-C): an effective brief screening test for problem drinking. Arch Intern Med 1998;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
73. Berman AH, Bergman H, Palmstierna T, Schlyter F. Drug Use Disorders Identification Test Extended (DUDIT-E). The European Union Drugs Agency (EUDA). URL: https://www.euda.europa.eu/drugs-library/drug-use-disorders-identification-test-extended-dudit-e_en [accessed 2026-01-08]
74. Blampied NM. Reliable change and the reliable change index: still useful after all these years? tCBT 2022;15. [doi: 10.1017/S1754470X22000484]
75. Thor J, Lundberg J, Ask J, et al. Application of statistical process control in healthcare improvement: systematic review. Qual Saf Health Care 2007 Oct;16(5):387-399. [doi: 10.1136/qshc.2006.022194] [Medline: 17913782]
76. Duncan TE, Duncan SC, Strycker LA. An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Application, 2nd edition: Routledge; 2006. [doi: 10.4324/9780203879962]
77. Norwegian Prescribed Drug Registry. The Norwegian Institute of Public Health. URL: https://helsedata.no/en/forvaltere/norwegian-institute-of-public-health/norwegian-prescribed-drug-registry [accessed 2026-01-08]
78. Norwegian Cause of Death Registry. The Norwegian Institute of Public Health. URL: https://helsedata.no/en/forvaltere/norwegian-institute-of-public-health/norwegian-cause-of-death-registry [accessed 2026-01-08]
79. Norsk helsenett – Cconnecting Norwegian Health Services. Norsk helsenett. URL: https://www.nhn.no/om-oss/about-us-connecting-norwegian-health-services [accessed 2026-01-08]
80. Data protection impact assessment (DPIA). European Data Protection Supervisor. URL: https://www.edps.europa.eu/data-protection-impact-assessment-dpia_en [accessed 2026-01-08]
81. Lov om behandling av personopplysninger (personopplysningsloven) [Article in Norwegian]. Lovdata. 2018. URL: https://lovdata.no/dokument/NL/lov/2018-06-15-38 [accessed 2026-01-08]
82. Lov om spesialisthelsetjenesten m.m. (spesialisthelsetjenesteloven) [Article in Norwegian]. Lovdata. 2001. URL: https://lovdata.no/dokument/NL/lov/1999-07-02-61 [accessed 2026-01-08]
83. General Data Protection Regulation. URL: https://gdpr.eu/ [accessed 2026-01-08]
84. Regional committees for medical and health research ethics. National Research Ethics Comittees. 2019. URL: https://www.forskningsetikk.no/en/about-us/our-committees-and-commission/rek [accessed 2026-01-08]

XSL•FO
RenderX

85. Masters R, Anwar E, Collins B, Cookson R, Capewell S. Return on investment of public health interventions: a systematic review. J Epidemiol Community Health 2017 Aug;71(8):827-834. [doi: 10.1136/jech-2016-208141] [Medline: 28356325]
86. Gupta P, Gupta M, Koul N. Overdiagnosis and overtreatment; how to deal with too much medicine. J Family Med Prim Care 2020;9(8):3815. [doi: 10.4103/jfmpc.jfmpc_433_20]

Abbreviations

ADHD: attention-deficit/hyperactivity disorder
***DSM*:** *Diagnostic and Statistical Manual of Mental Disorders*
EHR: electronic health record
GDPR: General Data Protection Regulation
***ICD*:** *International Statistical Classification of Diseases and Related Health Problems*
KPR: Norwegian Registry for Primary Health Care
LYL: life-years lost
NAMHR: Norwegian Adult Mental Health Registry
NCDR: Norwegian Cause of Death Registry
NorPD: Norwegian Prescribed Drug Registry
NPR: Norwegian Patient Registry
PREM: patient-reported experience measure
PROM: patient-reported outcome measure
RCT: randomized controlled trial
REC: Regional Committees for Medical and Health Research Ethics
ROM: routine outcome monitoring
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by A Schwartz; submitted 25.Aug.2025; peer-reviewed by C Asuzu, N Younes; revised version received 01.Dec.2025; accepted 22.Dec.2025; published 23.Jan.2026.

Please cite as:
Tjora T, Jensen K, Dieset I, Fjukstad KK, Holman PA, Tesli M, Joa I
The Norwegian Adult Mental Health Registry for Quality Control in Specialized Mental Health Services: Protocol for a Nationwide Naturalistic Study
JMIR Res Protoc 2026;15:e82696
URL: https://www.researchprotocols.org/2026/1/e82696
doi:10.2196/82696

© Tore Tjora, Kent Jensen, Ingrid Dieset, Katrine Kveli Fjukstad, Per Arne Holman, Martin Tesli, Inge Joa. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Utility of RECELL for Traumatic Skin Defects: Protocol for a Prospective, Single-Arm Pilot Study

Naoki Maegawa, MD, PhD; Hironobu Konishi, MD; Keisuke Masuda, MD, PhD; Akinori Okuda, MD, PhD; Hideki Asai, MD, PhD; Keita Miyazaki, MD, PhD; Yasuyuki Kawai, MD, PhD; Hidetada Fukushima, MD, PhD

Department of Emergency and Critical Care Medicine, Nara Medical University, Shijyou-cho 840, Kashihara, Japan

Corresponding Author:
Naoki Maegawa, MD, PhD
Department of Emergency and Critical Care Medicine, Nara Medical University, Shijyou-cho 840, Kashihara, Japan

Abstract

Background: Skin grafting is currently the gold standard for managing traumatic skin defects. However, split-thickness skin grafting, the most common technique, is challenging because it yields a limited amount of skin and creates a new wound at the healthy donor site. RECELL (Avita Medical), a noncultured skin cell suspension technique, yields healing outcomes with comparable duration and quality to skin grafting for burn trauma. Although RECELL is widely used for burn trauma, its use for severe traumatic skin defects is not yet prevalent in Japan.

Objective: This study aims to evaluate the utility of the RECELL skin reconstruction technique in trauma cases involving significant skin defects. Specifically, the primary aim is to validate the efficacy of the RECELL skin reconstruction technique when combined with split-thickness mesh grafting in patients with significant posttraumatic skin defects, with a particular focus on wound closure and scar quality.

Methods: This prospective, single-center, open-label pilot study will enroll patients aged 16 years and older with traumatic limb skin defects ≥160 cm² and sufficient dermal-like tissue. Participants will undergo RECELL-assisted autologous skin grafting with mesh grafts under general anesthesia. Wound healing, scarring (Vancouver Scar Scale), and pain (visual analog scale) will be assessed up to 24 weeks postoperatively. The primary end point is ≥95% reepithelialization at postoperative day 14. Descriptive statistical analyses will be performed to evaluate outcomes and inform future randomized trials.

Results: Patient recruitment began in April 2024 and is currently ongoing in our hospital. As of April 10, 2025, 10 patients had been screened, of whom 4 were enrolled. No adverse events or protocol deviations have been noted.

Conclusions: This study aims to validate the utility of RECELL for skin defects resulting from trauma by assessing wound closure and scar formation, including adverse events, delayed healing, infection, and durability.

Trial Registration: Japan Registry of Clinical Trials jRCTs052230203; https://jrct.mhlw.go.jp/latest-detail/jRCTs052230203

International Registered Report Identifier (IRRID): DERR1-10.2196/74159

(JMIR Res Protoc 2026;15:e74159) doi:10.2196/74159

KEYWORDS

skin grafting; flap surgery; autograft; wound healing; extremity trauma; noncultured skin cell suspension technique; split-thickness skin graft; open fracture

Introduction

Skin grafting is currently the gold standard for managing skin defects [1]. Split-thickness skin grafting, the most common technique, is challenging because it yields a limited amount of skin and creates a new wound at a healthy donor site [2]. For extensive skin defects caused by severe trauma, such as open fractures or degloving injuries, the limited availability of harvested skin often necessitates extended coverage using mesh grafting. However, epithelialization of the meshed area requires a considerable amount of time, and the cosmetic outcome may be less than ideal [3-6]. RECELL (Avita Medical) is a regenerative technique that uses the skin cells of a patient to create a noncultured skin cell suspension for wound healing [7-9]. This suspension is applied to the wound, enabling closure without extensive skin grafting. Although RECELL is widely used for burn trauma, its use for severe traumatic skin defects is not yet prevalent in Japan. To our knowledge, only a few studies have investigated the use of RECELL in patients with extremity trauma [10]. Therefore, this study aimed to evaluate the efficacy of the RECELL skin reconstruction technique in patients with significant posttraumatic skin defects.

While numerous preceding studies have demonstrated the utility of RECELL in the treatment of burn wounds, there are critical differences between burn and traumatic wounds. Burn wounds are typically clean and are primarily characterized by thermally

XSL•FO
RenderX

induced vascular damage. In contrast, traumatic wounds, particularly those resulting from crush or degloving injuries, carry a higher risk of contamination and often involve deeper and more complex tissue damage, including damage to blood vessels and subcutaneous structures. However, in large traumatic defects, if a clean and well-vascularized wound bed can be successfully prepared, there is a potential to achieve a higher graft take rate compared to burns, as the peripheral microcirculation may be relatively preserved. Currently, robust data demonstrating the efficacy of RECELL in combination therapy for large traumatic skin defects are lacking. To address this knowledge gap, we sought to clarify the treatment outcomes of layered skin grafting incorporating the RECELL system for traumatic soft tissue defects. Currently, RECELL is not covered by Japanese insurance for the treatment of traumatic skin defects. The findings of this study will aid in the expansion of RECELL technology application in trauma cases. Therefore, the primary aim of this study is to assess the utility and safety of RECELL combined with split-thickness meshed skin grafting in the reconstruction of large posttraumatic skin defects.

Methods

Study Design, Participants, and Setting

This is a single-center, open-label, uncontrolled, single-arm, comparative prospective study.

The participants will be patients with open limb wounds transported to our center by ambulance for treatment, with good subcutaneous granulation tissue formation. Furthermore, considering that the area of the skin defect covered by one RECELL 1 kit is 160 cm^2, we targeted patients with epidermal defects measuring ≥160 cm^2. Given that the number of such cases transported to our hospital in the past was approximately 5 per year, we decided to target cases over a period of 2 years. The study period will therefore last from April 1, 2024, to March 31, 2026.

Ethical Considerations

This study will be conducted in compliance with the Declaration of Helsinki and the Clinical Trial Act of the Ministry of Health, Labour, and Welfare in Japan. The investigators will provide each participant with a sufficient explanation of the study and obtain written informed consent to participate. The study protocol has already been approved by the Nara Medical University Certified Review Board (CRB5200002). The trial has been registered on the Japan Registry of Clinical Trials (jRCTs052230203). This study followed the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines (Checklist 1). All study data will be handled confidentially and stored securely. Participants have the right to withdraw consent at any time without penalty, and their participation data will be handled according to the ethical approval guidelines. All adverse events and serious adverse events will be recorded, investigated, and promptly reported to the Nara Medical University Certified Review Board as required by the protocol. This study will be conducted as part of regular treatment and will not involve any compensation to the participants.

Inclusion and Exclusion Criteria

This study will include patients with skin defects due to trauma (excluding burns caused by flames, excessive heat, or skin exposure to steam or hot water) or skin defects due to skin flaps. Eligible patients must have a skin defect area of ≥160 cm^2, excluding the hands and face, and sufficient dermal-like tissue formed at the test site. Participants must be aged 16 years or older at the time of consent and provide written informed consent. For those younger than 18 years, written consent will be obtained from their parents or legal guardians. The area to be grafted will be defined as an area considered difficult to treat with standard skin grafting, based on the range of indications for burns in Japan.

The study will exclude patients younger than 16 years; those with hypersensitivity to sodium lactate (washing solution), anesthetics, adrenaline, povidone-iodine, or chlorhexidine; and those with allergies to ingredients derived from pigs. Furthermore, patients will be excluded if they are judged to be unsuitable for study by the principal investigator (or subinvestigator). The face was defined as an excluded affected region due to cosmetic and anatomical complexity that requires specialized reconstructive techniques beyond the scope of this protocol. The selected size of ≥160 cm^2 was chosen not only because it is the effective coverage area of one standard RECELL kit but also to address the efficacy gap identified in the literature regarding large posttraumatic defects. The exclusion criteria were chosen because, although this study considered skin loss due to trauma, treating limb injuries with skin grafts on the hands is highly likely to lead to significant functional impairment and is therefore generally not recommended. Furthermore, allergic reactions to the medications used were considered a contraindication, and therefore patients who exhibit such reactions will be excluded.

Surgical Treatment Protocol

The enrolled patients will undergo surgery according to the study protocol. Under general anesthesia, skin grafts will be harvested from healthy skin donor sites (the groin, thigh, or scalp). In addition to the RECELL skin grafts, the necessary amount of mesh skin grafts will be collected. A mesh graft will be applied to the defect site where it is to be placed, and a noncultured cell suspension, previously prepared using RECELL, will be sprayed onto the mesh-grafted area. A nonadherent gauze will be applied to the skin graft, and the gauze will be placed over it for bandage fixation.

Postsurgical Treatment

The bandaged treatment site will be uncovered 6 to 8 days postoperatively, and the wounds will be evaluated at 24 weeks postoperatively. The administration of antibiotics, anticoagulants, steroid ointments, and any type of dressing material will be permitted. However, the following combined treatments will not be allowed: full-layer skin grafts, negative pressure wound closure therapy used solely for skin graft fixation, and cytotoxic drugs (eg, sulfadiazine silver–containing preparations).

Treatment after completing the protocol may include washing, ointment therapy, and negative pressure wound closure therapy.

XSL•FO
RenderX

Treatment will be administered at the discretion of the principal investigator and research associate.

Study Data Collection

This study will involve a comprehensive collection of the participants' data. This will be initiated by recording the participants' background information, including age, sex, height, weight, blood pressure, pulse rate, and body temperature, along with a detailed survey of their current and previous medical histories, as well as any underlying medical conditions present at the time of inclusion in the study.

Subsequently, a series of blood tests, including white blood cell, red blood cell, hemoglobin, hematocrit, and platelet counts, along with coagulation indicators, such as the prothrombin time-international normalized ratio and activated partial thromboplastin time, will be performed. We will also measure the levels of fibrinogen, fibrin degradation products, and a range of biochemical markers, including aspartate aminotransferase and alanine aminotransferase, lactate dehydrogenase, creatine phosphokinase, blood urea nitrogen and creatinine, and C-reactive protein.

For the trauma data, the cause and mechanism of injury will be documented in detail. During the protocol surgery, we will evaluate the recipient site by measuring its long and short diameters and thickness and noting any previous treatments and the method of excision. We will record the magnification of the mesh skin graft and the dimensions of the skin donor site. Photographic records of the recipient sites will be obtained immediately before and after surgery.

Posttreatment follow-up will be performed at 1, 2, 4, 6, 8, 12, and 24 weeks. The tolerance levels applied to the follow-up examination dates are defined in Table 1. During the posttreatment follow-up, we will carefully observe the healing of the recipient site; assess scarring using the Vancouver Scar Scale (VSS) [11]; survey patient satisfaction with the treatment; and observe delayed healing, infections, or scarring requiring surgical intervention. For these evaluations, particularly the photographic records and VSS assessments, the same evaluator will assess each patient whenever possible.

Table . Schedule of follow-up visits and tolerance levels.

Time point	Tolerance level
Pretreatment	
−7 days up to treatment	—[a]
Day 0	
Treatment day	—
Posttreatment	
After 1 week	±1 day
After 2 weeks	±3 days
After 4 weeks	±5 days
After 6 weeks	±6 days
After 8 weeks	±7 days
After 12 weeks	±10 days
After 24 weeks	±14 days

[a]Not available.

Finally, pain and concomitant treatments will be evaluated. Pain at both the recipient and donor sites will be evaluated using a visual analog scale (VAS). We also plan to maintain a detailed record of any concurrent treatments, including the names, daily doses, and durations of antibiotics and anticoagulants, as well as the types of dressing materials and steroid ointments used. A CONSORT (Consolidated Standards of Reporting Trials)–style timeline of the planned patient enrollment and assessment is shown in Table 2.

XSL•FO
RenderX

Table . Timeline of the preoperative and postoperative assessment protocol for this study.

	Pretreatment	Treatment day	After 1 week	After 2 weeks	After 4 weeks	After 6 weeks	After 8 weeks	After 12 weeks	After 24 weeks
Obtaining consent	✓								
Photography: test site		✓	✓	✓	✓	✓	✓	✓	✓
Healing of the test site			✓	✓	✓	✓	✓	✓	✓
Scar assessment (VSS[a])								✓	✓
Subject satisfaction after treatment								✓	✓
Delayed healing of test site					✓	✓	✓	✓	✓
Presence of wound infection		✓	✓	✓	✓	✓	✓	✓	✓
Scars requiring surgical intervention			✓	✓	✓	✓	✓	✓	✓
Assessment of pain at the test site (VAS[b])			✓	✓	✓	✓	✓		
Diseases failure			✓	✓	✓	✓	✓		
Concomitant treatment survey			✓	✓	✓	✓	✓	✓	✓

[a]VSS: Vancouver Scar Scale.

[b]VAS: visual analog scale.

Five years after the end of the study, the information will be processed so that specific individuals cannot be identified unless the information is matched with other information, and the records will be supplemented.

Data Analysis

This study is designed as a prospective, single-arm pilot study. Given the lack of robust data on RECELL for large traumatic defects, a formal power calculation is not feasible for the primary end point. Therefore, we aim to enroll 10 patients over 2 years, which will provide the descriptive data necessary to inform the design and power calculation for a subsequent randomized controlled trial.

The primary end point is the achievement of ≥95% reepithelialization at the recipient site, assessed at postoperative day 14 (±3 days tolerance), corresponding to the final planned dressing change within the early healing phase (refer to Table 2). We determined that 14 days was appropriate based on our past case experience and several published studies [12-14]. The secondary end points, including the VSS and VAS for scar assessment, will be primarily evaluated at the final observation point of 24 weeks postoperatively (and also at 12 weeks), as detailed in Table 2.

Descriptive analyses for continuous variables will be presented as the mean and SD. Categorical variables will be presented as counts and percentages. Any missing data will be reported, and the analysis will use a complete case analysis approach, given the pilot nature and short follow-up period. Although this is a single-arm pilot study, exploratory comparisons of the primary outcome (reepithelialization rate) and VSS scores will be made against published historical controls to inform the sample size calculation for future randomized controlled trials. Descriptive analyses, including central tendency (mean and median) and variability, will be performed after data collection is completed.

Results

Patient recruitment began in April 2024 and is currently ongoing. As of April 10, 2025, 10 patients have been screened, of whom 4 were enrolled. No adverse events or protocol deviations were observed. We expect to complete patient enrollment by March 2026, with a follow-up period of 24 weeks

XSL•FO
RenderX

postoperatively. The final data analysis is scheduled to commence in mid-2026.

Discussion

Anticipated Findings

The hypothesized key finding of this study is that the combination of RECELL and 3:1 meshed skin grafting will demonstrate both favorable wound healing rates and superior long-term scar quality when compared with historical controls of standard skin grafting techniques for similar large traumatic defects. The management of extensive lower extremity soft tissue and skin loss is often difficult; indeed, several comorbidities should be considered and addressed in these patients [11,15]. The approach is multifactorial and requires commitment from both the surgeon and the patient. Several methods have been developed to address soft tissue and skin coverage of the limbs. For recontraction of large skin defects, the existing literature describes the evaluation, harvesting, transplantation, and management of skin grafting techniques for the lower extremities. Thus, soft tissue defects remain a challenge for orthopedic surgery. Such defects are commonly encountered following orthopedic injuries or infections, and the management of these wounds varies significantly [15,16]. In patients with burns, Elkady et al [17] reported that autologous skin cell suspension (ASCS) with or without split-thickness skin grafting reduced the number of surgical procedures, postoperative complications, and potential cost savings compared with split-thickness skin grafting alone. Carter et al [18] noted that these findings underscore the benefits of integrating ASCS with or without split-thickness skin grafting as a practical approach in burn wound management, providing substantial advantages to both patients and health care institutions. We believe that ASCS with or without split-thickness skin grafting is beneficial for extremity trauma because the wound bed is healthier than that of patients with burns [11,19,20]. However, the role of ASCS with or without split-thickness skin grafting in patients with traumatic skin defects remains unclear. There have been some reports on the use of RECELL in trauma cases [21-24]. The safety profiles were comparable, with similar frequencies of treatment-emergent adverse events of interest observed between the treatment areas. In prior case reports and case series of patients presenting with various nonthermal tissue defect types, ASCS has also shown favorable outcomes related to healing, donor site size, and aesthetic appearance [25-27]. A preliminary study of 10 posttraumatic cases treated with RECELL was conducted in 2020 [9]. This report showed that RECELL could be equivalent to a skin graft with complete healing at 30 days for smaller skin defects (<70 cm^2). However, the rate of wound healing decreased to below 13% for larger skin defects (>70 cm^2). The authors concluded that RECELL could not be the choice for skin reconstruction of larger traumatic defects and that the best reconstructive option remains traditional split-thickness mesh skin grafts. However, among patients with burns, Holmes et al [28] compared the healing outcomes of RECELL alone (mean area 168.2 cm^2) and 2:1 meshed skin grafting (mean area 165.0 cm^2) within-patient allocation of treatments and concluded that both treatment modalities were comparable among 87 patients at 4 weeks postoperatively. Although the target patients differ (trauma and burns) between these studies, the results show conflicting efficacy of RECELL; therefore, it seems premature to investigate the efficacy of RECELL alone in patients with trauma. Therefore, we referred to the prospective, multicenter, randomized trial by Gardien et al [29], who compared mesh grafting alone with grafting with cultured autologous proliferating epidermal cells in patients with burns, found that mesh grafting with cultured epidermal cells was more effective than grafting with mesh alone on postoperative days 5 to 7 in terms of wound closure and scar quality [18]. To investigate the impact of RECELL on early wound closure and scar quality, we designed a treatment protocol for 3:1 split-thickness mesh grafting in combination with RECELL [20,30].

Another key feature of this study is that the participants will not be enrolled unless a good wound bed is prepared. Controlling wound infection is a critical factor that must be addressed in patients with trauma. Kahn et al [31] conducted a randomized open-label trial of RECELL for contaminated diabetic foot gangrene, reporting epithelialization in 66.7% of patients. The key factor influencing this relatively high wound closure rate in diabetic foot gangrene was that the study included patients with skin defects with “adequately vascularized” wounds [31]. As the condition of the wound bed plays a major role in the success or failure of skin grafting, we believe that including this condition in patients with trauma will yield good epithelialization and wound closure.

In the proposed study, we will also investigate postoperative scar formation. However, the observation period will be relatively short (24 weeks). Holmes et al [28] recorded VSS scores for burn injuries with RECELL at 16, 24, and 52 weeks. The authors reported no significant difference between the RECELL and mesh skin grafts throughout the study period. However, we believe that a study period of 24 weeks is sufficient to observe scar formation.

We acknowledge several limitations to this study. First, it is a single-center, open-label pilot study with a small target sample size (n=10), limiting the generalizability of the findings. Second, the lack of a contemporaneous control group means all comparisons with standard care must rely on historical or published data, which introduces inherent bias. However, the data gathered will be essential for calculating the required sample size for a future, adequately powered randomized controlled trial.

Conclusions

We believe that using RECELL for traumatic skin defects in the extremities will yield better clinical results than using it for patients with burns. With this study protocol, we aim to validate the usefulness of RECELL for skin defects owing to trauma by observing wound closure and scar formation, including adverse events, delayed healing, infection, and durability.

XSL•FO
RenderX

Acknowledgments

The authors thank the institutional ethics committee. The authors acknowledge the experts whose works are cited and discussed in this manuscript. The Nara Medical University Clinical Research Center provided assistance with the research protocol and ethics application procedures. The authors also thank Editage for editing this manuscript.

Funding

This research was funded by Nara Medical University, Japan.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

AO: Writing, review, and editing. HA: Writing of the original draft. HF: Conceptualization, design, and writing of the original draft. HK: Writing, review, and editing. K Masuda: Writing, review, and editing. K Miyazaki: Writing, review, and editing. NM: Conceptualization, design, and writing of the original draft. YK: Writing, review, and editing. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

COSMOTEC Co, Ltd. is the Japanese distributor of the RECELL device and provided the devices used in this study, which constitutes a conflict of interest. However, crucially, the sponsor had no involvement in data collection, data access, data analysis, or interpretation of results, nor did it influence the manuscript preparation. Furthermore, the company had no role in study design, data collection, data analysis, or manuscript preparation.

Checklist 1
SPIRIT checklist.
[PDF File, 272 KB - resprot_v15i1e74159_app1.pdf]

References

1. Fuller DA. Split thickness skin graft to lower leg. J Orthop Trauma 2016 Aug;30 Suppl 2:S34. [doi: 10.1097/BOT.0000000000000605] [Medline: 27441935]
2. Cho EH, Shammas RL, Carney MJ, et al. Muscle versus fasciocutaneous free flaps in lower extremity traumatic reconstruction: a multicenter outcomes analysis. Plast Reconstr Surg 2018 Jan;141(1):191-199. [doi: 10.1097/PRS.0000000000003927] [Medline: 28938362]
3. Bonaparte JP, Corsten MJ, Odell M, Gupta M, Allen M, Tse D. Management of the radial forearm free flap donor site using a topically applied tissue expansion device. Oral Surg Oral Med Oral Pathol Oral Radiol 2013 Jul;116(1):28-34. [doi: 10.1016/j.oooo.2013.01.007] [Medline: 23523463]
4. Mosquera C, Weyh A, Malik M, Fernandes R, Bunnell A, Nedrud S. Comparison of the outcomes of split thickness skin graft versus thickness skin graft for closure of the radial forearm free flap donor site: a systematic review. Microsurgery 2024 Jan;44(1):e31126. [doi: 10.1002/micr.31126] [Medline: 37990820]
5. Blackburn JH 2nd, Boemi L, Hall WW, et al. Negative-pressure dressings as a bolster for skin grafts. Ann Plast Surg 1998 May;40(5):453-457. [doi: 10.1097/00000637-199805000-00001] [Medline: 9600426]
6. Legemate CM, Ooms PJ, Trommel N, et al. Course of scar quality of donor sites following split skin graft harvesting: comparison between patients and observers. Wound Repair Regen 2020 Sep;28(5):696-703. [doi: 10.1111/wrr.12840] [Medline: 32519395]
7. Instructions for use: RECELL® autologous cell harvesting device. AVITA Medical. 2023. URL: https://avitamedical.com/wp-content/uploads/2023/08/AW-IFU030-Instructions-for-Use-RECELL%C2%AE-Autologous-Cell-Harvesting-Device-Rev_9.pdf [accessed 2025-12-07]
8. Peirce SC, Carolan-Rees G. ReCell® spray-on skin system for treating skin loss, scarring and depigmentation after burn injury: a NICE medical technology guidance. Appl Health Econ Health Policy 2019 Apr;17(2):131-141. [doi: 10.1007/s40258-018-00457-0] [Medline: 30635844]
9. Moris V, Cristofari S, Stivala A, et al. Recell in post-traumatic cases: preliminary results. J Plast Reconstr Aesthet Surg 2020 Oct;73(10):1897-1916. [doi: 10.1016/j.bjps.2020.03.009] [Medline: 32265168]
10. Baryza MJ, Baryza GA. The Vancouver Scar Scale: an administration tool and its interrater reliability. J Burn Care Rehabil 1995;16(5):535-538. [doi: 10.1097/00004630-199509000-00013] [Medline: 8537427]
11. Khan AA, Khan IM, Nguyen PP, et al. Skin graft techniques. Clin Podiatr Med Surg 2020 Oct;37(4):821-835. [doi: 10.1016/j.cpm.2020.07.007] [Medline: 32919607]

XSL•FO
RenderX

12. Karlsson M, Elmasry M, Steinvall I, Sjöberg F, Olofsson P, Thorfinn J. Scarring at donor sites after split-thickness skin graft: a prospective, longitudinal, randomized trial. Adv Skin Wound Care 2018 Apr;31(4):183-188. [doi: 10.1097/01.ASW.0000530684.31491.5f] [Medline: 29561343]
13. Kaartinen IS, Kuokkanen HO. Suprathel(®) causes less bleeding and scarring than Mepilex(®) Transfer in the treatment of donor sites of split-thickness skin grafts. J Plast Surg Hand Surg 2011 Sep;45(4-5):200-203. [doi: 10.3109/2000656X.2011.583515] [Medline: 22150140]
14. Poh Yuen Wen A, Halim AS, Mat Saad AZ, Mohd Nor F, Wan Sulaiman WA. A prospective study evaluating wound healing with sea cucumber gel compared with hydrogel in treatment of skin graft donor sites. Complement Ther Med 2018 Dec;41:261-266. [doi: 10.1016/j.ctim.2018.10.006] [Medline: 30477850]
15. Taylor BC, Triplet JJ, Wells M. Split-thickness skin grafting: a primer for orthopaedic surgeons. J Am Acad Orthop Surg 2021 Oct 15;29(20):855-861. [doi: 10.5435/JAAOS-D-20-01389] [Medline: 34547758]
16. Kowal S, Kruger E, Bilir P, et al. Cost-effectiveness of the use of autologous cell harvesting device compared to standard of care for treatment of severe burns in the United States. Adv Ther 2019 Jul;36(7):1715-1729. [doi: 10.1007/s12325-019-00961-2] [Medline: 31065995]
17. Elkady D, Larson BM, Sharma S, et al. Effectiveness of autologous skin cell suspension in large total body surface area burns: analysis of clinical outcomes and patient charges. J Burn Care Res 2024 Nov 14;45(6):1489-1498. [doi: 10.1093/jbcr/irae133] [Medline: 38943671]
18. Carter JE, Carson JS, Hickerson WL, et al. Length of stay and costs with autologous skin cell suspension versus split-thickness skin grafts: burn care data from US centers. Adv Ther 2022 Nov;39(11):5191-5202. [doi: 10.1007/s12325-022-02306-y] [Medline: 36103088]
19. Foster K, Amani A, Carter D, et al. Evaluating health economic outcomes of autologous skin cell suspension (ASCS) for definitive closure in US burn care using contemporary real-world burn center data. J Curr Med Res Opin 2021;4(11):1042-1054. [doi: 10.52845/CMRO/2021/4-11-1]
20. Hu ZC, Chen D, Guo D, et al. Randomized clinical trial of autologous skin cell suspension combined with skin grafting for chronic wounds. Br J Surg 2015 Jan;102(2):e117-e123. [doi: 10.1002/bjs.9688] [Medline: 25627123]
21. Bashir MA, Fung V, Kernohan MD, Ragbir M, Ahmed OA. “Z-plasty” modification of ulnar-based fasciocutaneous flap for closure of the radial forearm flap donor defect. Ann Plast Surg 2010 Jan;64(1):22-23. [doi: 10.1097/SAP.0b013e3181a13d8d] [Medline: 20023451]
22. Wester JL, Pittman AL, Lindau RH, Wax MK. AlloDerm with split-thickness skin graft for coverage of the forearm free flap donor site. Otolaryngol Head Neck Surg 2014 Jan;150(1):47-52. [doi: 10.1177/0194599813513713] [Medline: 24270163]
23. Sinha UK, Shih C, Chang K, Rice DH. Use of AlloDerm for coverage of radial forearm free flap donor site. Laryngoscope 2002 Feb;112(2):230-234. [doi: 10.1097/00005537-200202000-00006] [Medline: 11889375]
24. Wirthmann A, Finke JC, Giovanoli P, Lindenblatt N. Long-term follow-up of donor site morbidity after defect coverage with Integra following radial forearm flap elevation. Eur J Plast Surg 2014;37(3):159-166. [doi: 10.1007/s00238-013-0918-0] [Medline: 24563583]
25. Byun SH, Ahn KM, Kim SM, Lee JH. Functional and cosmetic outcome after closure of radial forearm free flap donor defect with porcine collagen membrane. J Craniomaxillofac Surg 2016 May;44(5):527-532. [doi: 10.1016/j.jcms.2016.02.013] [Medline: 27021266]
26. Tane N, Aihara M, Inoue H. The use of artificial dermis on the donor defect of the free forearm flap. J Reconstr Microsurg 1996 May;12(4):215-220. [doi: 10.1055/s-2007-1006479] [Medline: 8723831]
27. Garcia E, Stone E, Chan LS, Van Vliet M, Garner WL. Donor-site preferences in women during autologous skin grafting. Plast Reconstr Surg 2014 Mar;133(3):378e-382e. [doi: 10.1097/PRS.0000000000000062] [Medline: 24572883]
28. Holmes Iv JH, Molnar JA, Carter JE, et al. A comparative study of the ReCell® device and autologous spit-thickness meshed skin graft in the treatment of acute burn injuries. J Burn Care Res 2018 Aug 17;39(5):694-702. [doi: 10.1093/jbcr/iry029] [Medline: 29800234]
29. Gardien KL, Marck RE, Bloemen MC, et al. Outcome of burns treated with autologous cultured proliferating epidermal cells: a prospective randomized multicenter intrapatient comparative trial. Cell Transplant 2016;25(3):437-448. [doi: 10.3727/096368915X689569] [Medline: 26419871]
30. Carson JS, Carter JE, Hickerson WL, et al. Analysis of real-world length of stay data and costs associated with use of autologous skin cell suspension for the treatment of small burns in U.S. centers. Burns 2023 May;49(3):607-614. [doi: 10.1016/j.burns.2022.11.007] [Medline: 36813602]
31. Kahn SA, Carter JE, Wilde S, Chamberlain A, Walsh TP, Sparks JA. Autologous skin cell suspension for full-thickness skin defect reconstruction: current evidence and health economic expectations. Adv Ther 2024 Mar;41(3):891-900. [doi: 10.1007/s12325-023-02777-7] [Medline: 38253788]

Abbreviations

ASCS: autologous skin cell suspension
CONSORT: Consolidated Standards of Reporting Trials

XSL•FO
RenderX

VAS: visual analog scale
VSS: Vancouver Scar Scale

Edited by A Schwartz; submitted 09.Apr.2025; peer-reviewed by FD Francesco; revised version received 13.Nov.2025; accepted 24.Nov.2025; published 09.Jan.2026.

Please cite as:
Maegawa N, Konishi H, Masuda K, Okuda A, Asai H, Miyazaki K, Kawai Y, Fukushima H
Utility of RECELL for Traumatic Skin Defects: Protocol for a Prospective, Single-Arm Pilot Study
JMIR Res Protoc 2026;15:e74159
URL: https://www.researchprotocols.org/2026/1/e74159
doi:10.2196/74159

© Naoki Maegawa, Hironobu Konishi, Keisuke Masuda, Akinori Okuda, Hideki Asai, Keita Miyazaki, Yasuyuki Kawai, Hidetada Fukushima. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correlation Between Prakriti (Body Constitution) and Severity of Structural Alterations in the Lungs of Patients With SARS-CoV-2: Protocol for a Retrospective Cross-Sectional Study

Rugaved Gudadhe, BAMS, MD; Gaurav Sawarkar, BAMS, MD, PhD; Amol Deshpande, BAMS, MD

Department of Rachana Sharir, Mahatma Gandhi Ayurved College Hospital & Research Centre, Datta Meghe Institute of Higher Education, Salod (H), Wardha, Maharashtra, India

Corresponding Author:
Rugaved Gudadhe, BAMS, MD
Department of Rachana Sharir, Mahatma Gandhi Ayurved College Hospital & Research Centre, Datta Meghe Institute of Higher Education, Salod (H), Wardha, Maharashtra, India

Abstract

Background: SARS-CoV-2, a novel coronavirus, initially appeared in Wuhan, China, at the end of 2019 and has infected more than 31 million people worldwide. Infection can range from asymptomatic to multiorgan failure requiring prompt treatment. Overall, 80% of patients with SARS-CoV-2 experienced mild to moderate illness, while 5% developed serious illness. Physicians can evaluate a patient's digestive system (*Koshtha*), digestive capacity (*Agni*), strength (*Bala*), and longevity (*Ayu*) using the body constitution (*Deha Prakriti*). It also helps doctors estimate a patient's illness risk, severity, disease activity scores, and hematological, pathological, and biochemical changes. The study investigated the association between body composition (*Deha Prakriti*) and severity, as shown by structural lung abnormalities in patients with SARS-CoV-2.

Objective: This study aims to study the correlation between the severity of *Prakriti* and structural alterations in the lungs of patients with SARS-CoV-2.

Methods: This is a retrospective cross-sectional study of patients with SARS-CoV-2. The research data will be retrospectively collected from hospital records between September 1, 2020, and May 11, 2021, a period during which India experienced the second wave of the pandemic. The data will come from the Acharya Vinoba Bhave Rural Hospital in Sawangi (Meghe), Wardha, Maharashtra, India. Patients will be contacted via telephone and encouraged to visit the outpatient department and inpatient department at these institutions or in rural and urban areas of Wardha city. A structured case pro forma and a *Prakriti* assessment questionnaire will be used to evaluate lung structural changes during the COVID-19–positive period.

Results: The study is not funded by any organization. The study was initiated on March 4, 2023, and as of July 1, 2024, a total of 265 patients have been recruited. Results will be recorded from the observations of subjective and objective parameters. The study's primary outcome is to establish a relationship between abnormalities in the structure of the lungs in patients with COVID-19 and body constitution (*Prakriti*). The study's secondary outcome will help identify which body constitution is most susceptible to structural changes and disease severity in patients with COVID-19 and will also offer insights into preventive medicine.

Conclusions: Statistical investigation will lead to the conclusion that there is a specific association between structural abnormalities in the lungs of patients with COVID-19 and their body constitution. We hypothesize that *Prakriti* will be identified as being more prone to structural changes and severity in patients with COVID-19, offering insights into preventive therapy.

International Registered Report Identifier (IRRID): PRR1-10.2196/63916

(JMIR Res Protoc 2026;15:e63916) doi:10.2196/63916

KEYWORDS

SARS-CoV-2; COVID-19; prakriti; agni; koshtha; structural alterations

Introduction

Background

The appearance and spread of the 2019 novel COVID-19 or SARS-CoV-2 has caused a new public health disaster that is putting the entire world at risk [1]. As the SARS-CoV-2 virus causes COVID-19, it is commonly referred to as a viral respiratory and vascular disorder [2]. COVID-19 causes pneumonia similar to Middle East respiratory syndrome and SARS [3,4]. The clinical signs range from fever, headaches, and myalgia symptoms, which are common with viral pneumonia, which can result in severe respiratory failure and death [5,6]. Ground glass opacities and consolidation, primarily in the lower lobe and peripheral regions, are common computed tomography (CT) findings of COVID-19 pneumonia. Other

XSL•FO
RenderX

findings include a crazy paving pattern, vascular enlargement, an air bubble sign, halo and reverse halo signs, airway changes, and airway alterations in bronchograms [7,8]. Atypical observations include the following traits: central distribution, peri-broncho-vascular dissemination, isolated involvement of the upper lobe, lobar and segmental consolidation, nodules, subpleural sparing, pleural and pericardial effusion, and white lung [9-11]. According to the findings, the most common lesion patterns on thorax CT scans were ground-glass opacity and consolidation. There were also unusual pavement patterns, airway changes, reverse halo signs, halo signs, and air bubble signs [12,13].

Patients' CT thorax examinations frequently reveal the following lesion patterns: (1) ground-glass opacity and (2) consolidation. Following COVID-19's acute phase, pulmonary fibrosis develops as the second serious respiratory consequence. Patients with post-acute COVID-19 syndrome frequently have dyspnea, cough, oxygen dependence, difficulty weaning from mechanical ventilation or noninvasive ventilation, changes to their fibrotic lungs, decreased diffusion capacity, and decreased endurance [14]. Regardless of viral status, the term "Long COVID" refers to the presence of a variety of symptoms weeks or months after SARS-CoV-2 infection. It is also known as "post-COVID syndrome." Its pattern can be chronic or relapsing-remitting [15,16]. People with "long COVID" frequently experience symptoms, such as severe exhaustion, breathlessness, coughing, chest pain, palpitations, headache, joint pain, myalgia, weakness, insomnia, diarrhea, rash, hair loss, impaired balance and gait, neurocognitive issues, such as memory and concentration issues, and a lower quality of life. "Long COVID" patients may have 1 or more symptoms [17]. Chronic cough, fibrotic lung disease (post–COVID-19 fibrosis or post-acute respiratory distress syndrome fibrosis), bronchiectasis, and pulmonary vascular disease are some of the pulmonary issues that can develop as a result of SARS-CoV-2 infection [15]. Chronic shortness of breath may result from persistent pulmonary involvement, which usually resolves over time. Unfortunately, CT scans of many COVID-19 asymptomatic patients show severe lung involvement. COVID-19–induced lung fibrosis may cause persistent dyspnea and necessitate additional oxygen [18]. People infected with SARS-CoV-2 face significant social and economic consequences if their chronic symptoms persist. As the illness spreads, more people may require medical attention, putting pressure on the health care system [19]. It also has an impact on one's quality of life; therefore, the concept of personalized medication may be more advantageous. Ayurvedic treatment will be based on body constitution (*Prakriti*) in a similar vein, as it outlines the treatment plan and will aid in effective management. Body constitution (*Prakriti*) (air entity [*Vata*], hot entity [*Pitta*], and phlegm [*Kapha*], the 3 functional constituents of the body) is the formation of distinctive qualities caused by the dominance of *Doshas* in this sense.

Rationale for the Study

Body constitution (*Prakriti*) refers to an individual's predominance of 1 or more *Doshas*. Other factors also have an impact on a body's composition [20]. The ancient books of Ayurveda also include recommendations for maintaining a lifestyle per one's body constitution (*Prakriti*) to continue healthy living in a personalized manner [21]. Body constitution (*Prakriti*) refers to specific physical and mental characteristics that are associated with disease susceptibility. We also anticipate that it will be possible to genotypically screen newborns to identify their body constitution (*Prakriti*), which will help us choose the type of healthy lifestyle that will allow these newborns to live long, disease-free lives in the future. This will be a groundbreaking development for personalized preventative medicine in humans [22]. Ayurveda has provided explicit instructions in the form of a daily routine (*Dinacharya*) and a seasonal daily routine (*Rutucharya*) to maintain the associated with *Dosha*, the natural functioning of a specific body constitution (*Prakriti*) [23]. According to Ayurveda, body constitution (*Prakriti*) plays an important role in disease susceptibility as well as the selection of preventative and therapeutic methods. Notably, a person's body constitution (*Prakriti*) influences how their immune system (*Bala*) functions [24,25]. The study's goal is to investigate the relationship between body constitution (*Prakriti*) and severity as evidenced by structural changes in the lungs of patients with SARS-CoV-2.

Objectives

This research aims to investigate structural changes in the lungs of patients with SARS-CoV-2, assess their body constitution (*Prakriti*), and correlate structural changes in the lungs of patients with SARS-CoV-2 with their body constitution (*Prakriti*).

Methods

Overview

This is a retrospective cross-sectional study of patients with SARS-CoV-2. The research data will be retrospectively collected from hospital records between September 1, 2020, and May 11, 2021, as India experiences the second wave of the pandemic during this period. The data will come from the Acharya Vinoba Bhave Rural Hospital in Sawangi (Meghe), Wardha, Maharashtra, India. Patients will be contacted via telephone and encouraged to visit the outpatient department and inpatient department at these institutions or in rural and urban areas of Wardha city. A structured case pro forma and a *Prakriti* assessment questionnaire will be used to evaluate lung structural changes during the COVID-19–positive period (Multimedia Appendix 1).

Sample Calculation

Sample calculation will be done according to the following formula:

$$n \geq Z1-a22 \times P \times (1-p)d2$$

Sample Size

The sample size is calculated as 246 based on the proportion or prevalence of ground glass opacity at 80% (0.8) and Z alpha at 5% level (1.96 constant), with an estimated margin of error of 5% (0.05) [26]. The formula for calculating the sample size was N=(1.96)2 × 0.8 × (1–0.2)/(0.05)2=246 minimum

XSL•FO
RenderX

Grouping

The study will include 2 groups, each consisting of 246 individuals. Group A comprises individuals exposed to SARS-CoV-2 with lung involvement, while Group B includes those exposed to the virus without lung involvement. This classification aims to assess the impact of SARS-CoV-2 on lung health and compare outcomes between affected and unaffected individuals. Understanding these differences can help in evaluating disease severity, treatment approaches, and long-term health implications for patients with and without pulmonary complications due to the virus. Group A will include 246 participants exposed to SARS-CoV-2 having lung involvement, and Group B will include 246 participants exposed to SARS-CoV-2 having no involvement of the lung.

Assessment Parameter

Reverse Transcription–Polymerase Chain Reaction

Reverse transcription–polymerase chain reaction (RT-PCR) is the standard test for detecting SARS-CoV-2. It is typically performed on a sample of respiratory or nasopharyngeal secretions. Although RT-PCR is thought to be quite specific, its sensitivity might vary from 60% - 70% to 95% - 97% [27].

Chest X-Ray

It is helpful in the assessment of the structural changes in the lung field and identification of potential side effects, including pneumothorax, subcutaneous emphysema, and pneumo-mediastinum, and ongoing tracking of the progression of the illness. The most frequent discoveries are opacities in the airspace, which might be either consolidations or, less frequently, opacities in the ground glass [27].

High-Resolution Computed Tomography—Chest

It is a quick, accessible test considered to be the most sensitive imaging test for detecting COVID-19, with a reported sensitivity of up to 97%. Typical findings include ground glass opacities, consolidation, peripheral reticulation, and a crazy paving pattern [27].

The patient's body constitution (*Prakriti*) will be assessed by the personal interviewing method. The standard pro forma for the assessment of *Prakriti* is attempted to develop according to the Central Council for Research in Ayurvedic Sciences' *Prakriti* assessment module to ensure physical constitution of the human body based on several views of the body constitution [28].

Inclusion Criteria

Nonvaccinated patients with SARS-CoV-2, regardless of age or gender, who tested positive for RT-PCR between September 1, 2020, and May 11, 2021, and were willing to provide informed consent will be included in the study.

Exclusion Criteria

The study will exclude patients with respiratory disorders, such as chronic obstructive pulmonary disease and bronchial asthma, and those with multiple organ disorders. For more than a year, patients who smoked more than 2 cigarettes per day and consumed more than 60 mL of alcohol per day were also excluded.

Data Sources and Management

Overview

Patient data will be screened from the medical record department and patients with post–COVID-19 (who tested positive within the specified time period) at Acharya Vinoba Bhave Rural Hospital Sawangi's outpatient department and inpatient department. Personal interviews and assessments will be conducted to determine the body's constitution (*Prakriti*).

Gradation and Calculation of Prakriti

The final percentage scores for *Vata*, *Pitta*, and *Kapha* are determined using the formula:



The collected data will be categorized and fractionalized using frequencies and percentages [29,30].

Statistical Analysis

The body constitution (*Prakriti*) will be evaluated, and the resulting data will be classified and fractionalized using frequency and percentages. Following that, the chi-square test will be used in the study.

Ethical Considerations

This research obtained approval from the Institutional Ethics Committee of Datta Meghe Institute of Medical Sciences (deemed to be a University) on August 1, 2022, with reference letter MGACHRC/IEC/Aug-2022/558. The purpose of the study will be explained to the participants and their guardians in the local language, with a pro forma designed for this purpose. After bringing all these to the informant's notice, data will be collected with written informed consent, keeping the information confidential by interviewing them afterwards. Participants' information will be deidentified, with personal identifiers stored separately from research data. Data will be retained for a specific duration and securely disposed of thereafter. Study procedures will be conducted privately, and no identifiable information will be disclosed without explicit consent. The study will comply with applicable data protection regulations, and participants will be informed of confidentiality measures through the consent process. Participants will not incur any financial costs for participating in this study. If applicable, compensation will be provided. In the unlikely event of any adverse effects or injury directly caused by the study, appropriate medical care and compensation will be provided in accordance with institutional and regulatory guidelines. Participants will be informed about their rights and the compensation process during the consent process.

Results

The study was initiated in March 2023, and as of July 2024, a total of 265 patients have been recruited. Results will be recorded from the observations of subjective and objective parameters. The study's primary outcome is to establish a relationship between abnormalities in the structure of the lungs

XSL•FO
RenderX

of patients with COVID-19 and body constitution. The study's secondary outcome will provide a better understanding of which body constitution is most susceptible to structural changes and disease severity in patients with COVID-19 and will also offer insights relevant to preventive medicine (Figure 1).

Figure 1. Study design. AVBRH: Acharya Vinoba Bhave Rural Hospital; COPD: chronic obstructive pulmonary disease; MGACH & RC: Mahatma Gandhi Ayurved College Hospital and Research Centre; RT-PCR: reverse transcription–polymerase chain reaction.

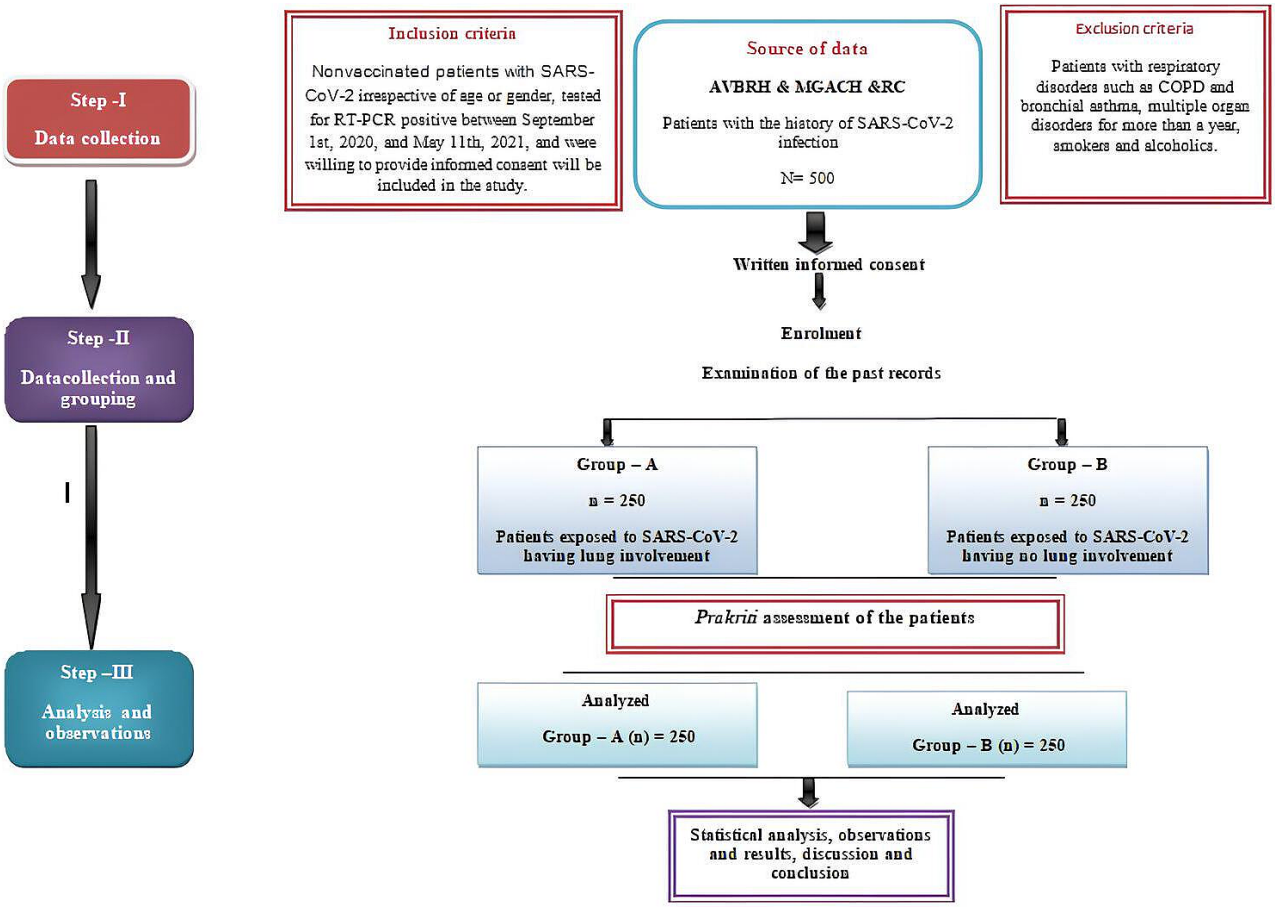

Discussion

Expected Findings

This study investigates the correlation between structural lung changes and different types of *Prakriti* in patients with SARS-CoV-2. If a significant relationship is found, it could support Ayurveda-based predictive models for disease severity. While prior research has focused on clinical and radiological factors, this study uniquely integrates traditional medicine with modern diagnostics. Strengths include its novel approach, though variability in *Prakriti* assessment may pose challenges. Future research should expand on these findings to enhance personalized health care strategies.

Importance of Body Constitution (*Deha Prakriti*) in Disease Susceptibility

A person's susceptibility to a specific illness can be determined by examining their body constitution (*Prakriti*), which is the key. According to Ayurvedic literature, the *Deha Prakriti* can affect the disease's susceptibility and severity. According to the literature, a person is more susceptible to a disease caused by the same *Dosha* in his body constitution (*Prakriti*) if the etiology is strong and comparable. Compared to people with other body constitutions (*Deha Prakriti*), those with a prominent *Kapha Dosha* are more quickly and easily impacted by powerful etiological variables that aggravate their *Kapha Dosha*. Similarly, if a person with *Vataja Prakriti* follows a diet and daily habits (*Ahara-Vihara*) that promote *Vata*, they are more likely to develop *Vataja* illness, which is difficult to treat due to the dominant *Dosha* of their body constitution (*Prakriti*). They experience this if they follow a *Vata*-inducing diet and daily habits (*Ahara-Vihara*). This belief holds that people with *Vataja Prakriti* are more susceptible to ailments caused by vitiated *Vata*. *Pittaja Prakriti* people are more likely to experience *pittaja* specialized difficulties, as are *Kaphaja* people. *Vata*, *Pitta*, and *Kapha* constitutions (*Prakritis*) have a lower risk of developing various ailments as their body immunity (*Bala*) and life span (*Ayu*) develop [31].

Importance of *Deha Prakriti* in Disease Prognosis

Using *Deha Prakriti* analysis, you can predict how a disease will develop. Unlike *Deha Prakriti*, the illness known as curable disease (*Sadhya Vyadhi*), which affects etiology (*Hetu*), premonitory symptoms (*Purvaroopa*), and signs and symptoms (*Rupa*), causes weakness (*Balakshaya*). When one of the body's constitutions (*Prakriti*), the vitiated part (*Dushya*), or period (*Kala*) resembles the *Vikarjanak Dosha*, and *Hetu*, *Purvaroopa*, and *Rupa* are moderate members, healing is difficult (*Kashtasadhya*) [31]. Only a small percentage of patients who

XSL•FO
RenderX

recover from COVID-19 experience the illness known as "long COVID," "Long Haulers," or "Post COVID syndrome," which is characterized by additional or persistent symptoms that last for weeks or months [15]. SARS-CoV-2 infection may cause bronchiectasis, post–COVID-19 fibrosis, chronic cough, and pulmonary vascular disease [17]. Unfortunately, many CT scans of asymptomatic patients with COVID-19 show severe lung disease. COVID-19 in lung fibrosis increases the risk of chronic dyspnea and necessitates the use of additional oxygen [32]. SARS-CoV-2 infection symptoms that last for an extended period of time have significant social and economic consequences. As the illness spreads, more people may require medical care in the near future, putting an unfair strain on the health care system. Providing medical professionals with specific instructions for managing chronic COVID-19 will help reduce confusion. More data will be obtained by tracking patients who have recovered from COVID-19 over time [33].

Conclusions

Statistical investigation will lead to the conclusion that there is a specific association between structural abnormalities in the lungs of patients with COVID-19 and their body constitution. We hypothesize that *Prakriti* will be identified as being more prone to structural changes and severity in patients with COVID-19, offering insights into preventive therapy.

Acknowledgments

This study contains sections generated with the assistance of artificial intelligence to enhance clarity and coherence. However, all artificial intelligence–generated content has been reviewed, edited, and validated by the authors to ensure accuracy, originality, and alignment with the study's objectives. The study reflects human oversight and intellectual contribution, adhering to ethical research and publication standards.

Funding

All authors declared that they had insufficient funding to support open access publication of this article, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided partial APF support for the publication of this article.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Case paper for *Prakriti* assessment.
[PDF File, 531 KB - resprot_v15i1e63916_app1.pdf]

References

1. Singhal T. A review of coronavirus disease-2019 (COVID-19). Indian J Pediatr 2020 Apr;87(4):281-286. [doi: 10.1007/s12098-020-03263-6] [Medline: 32166607]
2. Cascella M, Rajnik M, Aleem A, Dulebohn SC, Napoli R. Features, evaluation, and treatment of coronavirus (COVID-19). In: StatPearls: StatPearls Publishing; 2023. [Medline: 32150360]
3. Hosseiny M, Kooraki S, Gholamrezanezhad A, Reddy S, Myers L. Radiology perspective of coronavirus disease 2019 (COVID-19): lessons from severe acute respiratory syndrome and Middle East respiratory syndrome. AJR Am J Roentgenol 2020 May;214(5):1078-1082. [doi: 10.2214/AJR.20.22969] [Medline: 32108495]
4. Ajlan AM, Ahyad RA, Jamjoom LG, Alharthy A, Madani TA. Middle East respiratory syndrome coronavirus (MERS-CoV) infection: chest CT findings. AJR Am J Roentgenol 2014 Oct;203(4):782-787. [doi: 10.2214/AJR.14.13021] [Medline: 24918624]
5. Radwan NM, Mahmoud NE, Alfaifi AH, Alabdulkareem KI. Comorbidities and severity of coronavirus disease 2019 patients. Saudi Med J 2020 Nov;41(11):1165-1174. [doi: 10.15537/smj.2020.11.25454] [Medline: 33130835]
6. Hafez MAF. The mean severity score and its correlation with common computed tomography chest manifestations in Egyptian patients with COVID-2019 pneumonia. Egypt J Radiol Nucl Med 2020 Dec;51(1):1-9. [doi: 10.1186/s43055-020-00368-y]
7. Bayraktaroğlu S, Çinkooğlu A, Ceylan N, Savaş R. The novel coronavirus pneumonia (COVID-19): a pictorial review of chest CT features. Diagn Interv Radiol 2021 Mar;27(2):188-194. [doi: 10.5152/dir.2020.20304] [Medline: 32815523]
8. Caruso D, Polidori T, Guido G, et al. Typical and atypical COVID-19 computed tomography findings. World J Clin Cases 2020 Aug 6;8(15):3177-3187. [doi: 10.12998/wjcc.v8.i15.3177] [Medline: 32874972]

XSL•FO
RenderX

9. Ceylan N, Çinkooğlu A, Bayraktaroğlu S, Savaş R. Atypical chest CT findings of COVID-19 pneumonia: a pictorial review. Diagn Interv Radiol 2021 May;27(3):344-349. [doi: 10.5152/dir.2020.20355] [Medline: 33032981]
10. Emara DM, Naguib NN, Moustafa MA, Ali SM, El Abd AM. Typical and atypical CT chest imaging findings of novel coronavirus 19 (COVID-19) in correlation with clinical data: impact on the need to ICU admission, ventilation and mortality. Egypt J Radiol Nucl Med 2020 Dec;51(1). [doi: 10.1186/s43055-020-00339-3]
11. Emara DM, Naguib NN, Moustafa MA, Ali SM, El Abd AM. Typical and atypical CT chest imaging findings of novel coronavirus 19 (COVID-19) in correlation with clinical data: impact on the need to ICU admission, ventilation and mortality. Egypt J Radiol Nucl Med 2020 Dec;51(1):1-9. [doi: 10.1186/s43055-020-00339-3]
12. Korkmaz İ, Keleş F. COVID-19-related lung involvement at different time intervals: evaluation of computed tomography images with semiquantitative scoring system and COVID-19 reporting and data system scoring. Cureus 2021 Oct;13(10):e18554. [doi: 10.7759/cureus.18554] [Medline: 34765340]
13. Alarcón-Rodríguez J, Fernández-Velilla M, Ureña-Vacas A, et al. Radiological management and follow-up of post-COVID-19 patients. Radiologia (Engl Ed) 2021;63(3):258-269. [doi: 10.1016/j.rx.2021.02.003] [Medline: 33726915]
14. Chippa V, Aleem A, Anjum F. Post acute coronavirus (COVID-19) syndrome. In: StatPearls: StatPearls Publishing; 2024. [Medline: 34033370]
15. Raveendran AV, Jayadevan R, Sashidharan S. Long COVID: an overview. Diabetes Metab Syndr 2021;15(3):869-875. [doi: 10.1016/j.dsx.2021.04.007] [Medline: 33892403]
16. Nabavi N. Long covid: how to define it and how to manage it. BMJ 2020 Sep 7;370:m3489. [doi: 10.1136/bmj.m3489] [Medline: 32895219]
17. Fraser E. Long term respiratory complications of covid-19. BMJ 2020 Aug 3;370:m3001. [doi: 10.1136/bmj.m3001] [Medline: 32747332]
18. Evans RA, McAuley H, Harrison EM, et al. Physical, cognitive, and mental health impacts of COVID-19 after hospitalisation (PHOSP-COVID): a UK multicentre, prospective cohort study. Lancet Respir Med 2021 Nov;9(11):1275-1287. [doi: 10.1016/S2213-2600(21)00383-0] [Medline: 34627560]
19. Dwivedi L. Agnivesha Charaka Samhita, 4th edition: Chaukhamba Sanskrit Pratishthan; 2017, Vol. 2.
20. Hankey A. The scientific value of ayurveda. J Altern Complement Med 2005 Apr;11(2):221-225. [doi: 10.1089/acm.2005.11.221] [Medline: 15865485]
21. Bhalerao S, Deshpande T, Thatte U. Prakriti (ayurvedic concept of constitution) and variations in platelet aggregation. BMC Complement Altern Med 2012 Dec 10;12(1):1-5. [doi: 10.1186/1472-6882-12-248] [Medline: 23228069]
22. Kukade SV, Mekhale S, Agarwal M, Shahu D. Concept of Prakriti in Ayurveda and its significance in evading lifestyle disorders. Int Ayurvedic Med J 2016 Aug 12;4(6). [doi: 10.70066/jahm.v12i1.1218]
23. Sengupta N, Sengupta B, editors. Caraka Samhita of Agnivesa: Rashtriya Sanskrit Sansthan; 2002.
24. Acharya YT, editor. Susruta Samhita, 8th edition: Chaukambha Orientalia; 2005.
25. Goetz LH, Schork NJ. Personalized medicine: motivation, challenges, and progress. Fertil Steril 2018 Jun;109(6):952-963. [doi: 10.1016/j.fertnstert.2018.05.006] [Medline: 29935653]
26. Liu X, Zhou H, Zhou Y, et al. Temporal radiographic changes in COVID-19 patients: relationship to disease severity and viral clearance. Sci Rep 2020 Jun 24;10(1):10263. [doi: 10.1038/s41598-020-66895-w] [Medline: 32581324]
27. Martínez Chamorro E, Díez Tascón A, Ibáñez Sanz L, Ossaba Vélez S, Borruel Nacenta S. Radiologic diagnosis of patients with COVID-19. Radiologia (Engl Ed) 2021;63(1):56-73. [doi: 10.1016/j.rx.2020.11.001] [Medline: 33339622]
28. Manual of Standard Operative Procedures for Prakriti Assessment, 2nd edition: Central Council for Research in Ayurvedic Sciences, Ministry of Ayush; 2020.
29. Manual of Standard Operative Procedures for Prakriti Assessment, 1st edition: Central Council for Research in Ayurvedic Sciences, Ministry of Ayush; 2018.
30. Kumar M. Assesment of prakriti and its clinical relivence. Int J Emerg Technol Innov Res 2023;10(3) [FREE Full text]
31. Acharya YT, editor. Charaka Samhita of Agnivesha, Vimana Sthana: Chaukhambha Surbharati Prakashan; 2014.
32. Becker RC. Toward understanding the 2019 coronavirus and its impact on the heart. J Thromb Thrombolysis 2020 Jul;50(1):33-42. [doi: 10.1007/s11239-020-02107-6] [Medline: 32297133]
33. Moriguchi T, Harii N, Goto J, et al. A first case of meningitis/encephalitis associated with SARS-Coronavirus-2. Int J Infect Dis 2020 May;94:55-58. [doi: 10.1016/j.ijid.2020.03.062] [Medline: 32251791]

Abbreviations

CT: computed tomography
IPD: inpatient department
OPD: outpatient department
RT-PCR: reverse transcription–polymerase chain reaction

XSL•FO
RenderX

Edited by A Schwartz, J Sarvestan; submitted 03.Jul.2024; peer-reviewed by A Afzalian, VA Arumugam; revised version received 15.Mar.2025; accepted 17.Mar.2025; published 08.Jan.2026.

Please cite as:
Gudadhe R, Sawarkar G, Deshpande A
Correlation Between Prakriti (Body Constitution) and Severity of Structural Alterations in the Lungs of Patients With SARS-CoV-2: Protocol for a Retrospective Cross-Sectional Study
JMIR Res Protoc 2026;15:e63916
URL: https://www.researchprotocols.org/2026/1/e63916
doi:10.2196/63916

© Rugaved Gudadhe, Gaurav Sawarkar, Amol Deshpande. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 8.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Comparative Evaluation of the Antibacterial Efficacy of Metronidazole, Chlorhexidine, and Normal Saline as Laser- and Sonic-Activated Intracanal Irrigants Against Enterococcus faecalis: In Vitro Study Protocol

Sharvari Deshmukh, MDS; Aditya Patel, MDS, PhD

Department of Conservative Dentistry and Endodontics, Sharad Pawar Dental College & Hospital, DMIHER Campus, Sawangi-Meghe, Wardha, India

Corresponding Author:
Sharvari Deshmukh, MDS
Department of Conservative Dentistry and Endodontics, Sharad Pawar Dental College & Hospital, DMIHER Campus, Sawangi-Meghe, Wardha, India

Abstract

Background: The persistence of *Enterococcus faecalis* is a significant challenge in endodontic therapy, due to its ability to form biofilms and penetrate the dentinal tubules, frequently leading to treatment failures. Therefore, despite the proven antimicrobial activity of metronidazole (MTR) and chlorhexidine (CHX), the resilience of *E. faecalis* necessitates integrating activation techniques. Laser and sonic activation methods show promise in enhancing the antibacterial performance of irrigants. However, studies on the activation-enhanced efficacy of MTR, CHX, and saline remain limited.

Objective: This study aimed to compare the antibacterial efficacy of MTR, CHX, and normal saline as intracanal irrigants when activated by laser and sonic techniques against *E. faecalis* in extracted human mandibular premolars.

Methods: It is a laboratory-based in vitro experimental study involving extracted human teeth. Ninety freshly extracted single-rooted mandibular premolars will be obtained, decoronated, and biomechanically prepared. Canals will be inoculated with *E. faecalis* (ATCC 29212) and incubated for 7 days to allow biofilm formation. The teeth will then be randomly divided into 3 main groups ([MTR], [CHX], and saline), each subdivided based on the activation method used (laser or sonic). Irrigation protocols will be standardized, and pre- and postirrigation bacterial samples will be collected on paper points and cultured on brain heart infusion agar. Colony-forming units will be counted to evaluate antibacterial efficacy.

Results: As of June 2025, ethical approval for the study has been obtained. Sample collection of extracted human teeth commenced in July 2025, and 90 extracted teeth have been collected as of January 2026, corresponding to the planned sample size. Laboratory procedures and data analysis have not yet commenced.

Conclusions: This study aims to provide evidence-based insights into optimizing root canal disinfection protocols by evaluating and comparing the synergistic effects of antimicrobial agents and advanced activation techniques. The findings could contribute significantly to improving endodontic treatment outcomes and reducing the incidence of persistent infections caused by *E. faecalis*.

International Registered Report Identifier (IRRID): PRR1-10.2196/76783

(JMIR Res Protoc 2026;15:e76783) doi:10.2196/76783

KEYWORDS

Enterococcus faecalis; metronidazole; chlorhexidine; sonic activation; laser activation; root canal disinfection; intracanal irrigants; endodontic treatment

Introduction

The combined action of intracanal drugs and chemomechanical preparation reduces or eliminates bacteria during root canal therapy (RCT) [1]. The mechanical action of tools and the flow and backflow of irrigating solutions are used to remove bacteria and other irritants from the root canal's boundaries [2]. A common and successful endodontic treatment for pulpitis and apical periodontitis is conventional root canal therapy, which uses biomechanical equipment to remove affected tooth pulp and dentin tissue while using antibiotics and antibacterial materials to eradicate germs [3].

Enterococcus faecalis has been detected in approximately 4%-40% of primary endodontic infections, indicating its significant role in persistent and recurrent root canal infections, especially in cases of treatment-resistant apical periodontitis [4,5]. Despite root canal treatment, up to 90% of teeth may retain residual infection. In the twenty-first century, with rising advancements in RCT, the complete eradication of *E. faecalis* remains challenging due to its ability to survive under adverse

XSL•FO
RenderX

conditions, with 2 main pillars of its survival being penetrating dentinal tubules and forming biofilms that resist conventional antimicrobial strategies. Its biofilms are notably resilient to commonly used intracanal agents like sodium hypochlorite and calcium hydroxide [6].

Many medications, including metronidazole (MTR) and chlorhexidine (CHX), have been used to treat oral diseases because they are excellent at inhibiting bacteria and biofilms. A broad-spectrum antibiotic that targets both anaerobic bacteria and protozoa, MTR is a nitroimidazole molecule. In addition to Gram-positive and Gram-negative bacilli, it exhibits strong antibacterial activity against anaerobic cocci. Both topical and systemic applications have been made in the treatment of periodontal disease. The anatomical complexity of the RCT may, however, harbor germs and remaining pulp tissue following RCT, potentially leading to treatment failure [7].

CHX is a cationic biguanide with broad-spectrum coverage. It has activity against both Gram-positive and Gram-negative bacteria. At low concentrations with a range between 0.2% and 0.06%, CHX shows bacteriostatic properties. At high concentrations, more than 0.12%, it depicts bactericidal activity, antifungal properties, and some antiviral properties. Its effect is not hampered by biological fluids such as blood, and it works for up to 48 hours after application. It acts by penetrating the bacterial cell membrane, causing cytoplasmic leakage at lower concentrations and precipitating nucleic acid at higher concentrations [8].

Normal saline, although lacking intrinsic antimicrobial properties, is included in this study as a negative control to isolate and evaluate the specific contribution of activation techniques (laser and sonic) independent of chemical disinfection. Its inclusion enables assessment of whether activation alone can produce any measurable reduction in bacterial load.

The distribution method, in conjunction with the irrigant's characteristics, affects the irrigation efficacy [9]. Irrigant activation in endodontic procedures pertains to the use of various energy forms—mechanical, sonic, ultrasonic, or other physical modalities—to enhance the dynamic movement and distribution of irrigating solutions throughout the intricately structured root canal system. Evidence from systematic reviews and meta-analyses indicates that the use of activation, regardless of the specific device or mechanism employed, significantly improves the efficacy of debris and smear layer removal. Consequently, activation is recognized as an integral phase in the chemo-mechanical debridement process, contributing to the overall success of root canal disinfection [10]. In contrast to ultrasonic irrigation, sonic irrigation uses a lower frequency (1 - 10 kHz). As a result, shear wall stresses and fluid velocity are reduced. Conversely, sonic activation produces noticeably more amplitude (horizontal tip displacement) [11].

In recent years, nonsurgical RCT has made greater use of laser technology. When used in conjunction with chemical solutions, its high power and distinct coherent rays are helpful during the root canal cleaning and disinfection phase. It has been demonstrated that a laser can effectively activate a root canal irrigant by causing mechanical disruptions [12]. The 940 nm diode laser is frequently used in dentistry clinics due to its compact size, portability, affordability, and broad variety of oral medicine applications. Bacterial cell walls can be destroyed by the bactericidal photothermal disrupting impact of its radiation [13]. Since water and hydroxyapatite are the primary components of dentin, the 940 nm wavelength has low absorption coefficients in water (about $\mu a=0.04 - 0.05\ cm^{-1}$), which allows the laser beam to penetrate deeply into dentin (over 1000 µm) and kill distant bacteria [13,14].

It has been observed in previous studies that the use of CHX and MTR as intracanal irrigants showed optimal results. However, studies with MTR as an intracanal irrigant with different activation methods need to be explored. Although metronidazole has demonstrated antimicrobial efficacy as an intracanal irrigant, prior studies have primarily evaluated it in static or nonactivated forms. So, there is a need for research that compares the antibacterial effectiveness of MTR, CHX, and normal saline as intracanal irrigants, which are activated by laser and sonic activation methods, against *E. faecalis*. This comparison will help determine the most effective activation technique for enhancing the antibacterial effectiveness of these irrigants, ultimately contributing to more reliable and durable endodontic treatments, improving treatment success rates, and reducing the incidence of persistent infections associated with this resilient bacterium. This study might address this critical gap by evaluating the synergistic effects of activation-enhanced MTR, thereby expanding current understanding and potentially informing more effective endodontic disinfection protocols.

This in vitro study aims to compare the antibacterial efficacy of MTR, CHX, and normal saline as intracanal irrigants when activated using laser and sonic techniques against *E. faecalis* in extracted human mandibular premolars. The study evaluates the effectiveness of each irrigant with both laser and sonic activation techniques to determine their potential in eliminating bacterial contamination from the root canal system. It is hypothesized that laser and sonic activation will significantly enhance the antibacterial efficacy of these irrigants, with notable differences among the groups.

The null hypothesis for the study assumes that, "there will be no significant effect of laser and sonic activation on the antibacterial efficacy of MTR, CHX, and saline as intracanal irrigants in eradicating Enterococcus faecalis in root canal treatments" while the alternative hypothesis suggests, "laser and sonic activation will significantly enhance the antibacterial efficacy of MTR, CHX, and saline as intracanal irrigants in eradicating Enterococcus faecalis in root canal treatments."

Methods

Study Design

This study used a comparative in vitro study design.

Materials

The materials to be used in the current study are listed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Materials to be used for microbial inoculation, irrigation, activation, and root canal preparation.

- Microbiological sample preparation:
- 1. Strains of *Enterococcus faecalis*
- 2. Brain heart infusion (BHI) agar media
- 3. Incubator
- Irrigants:
- 1. Chlorhexidine (CHX)
- 2. Metronidazole (MTR)
- 3. Normal saline
- Activation devices:
- 1. Sonic activation device
- 2. Diode laser
- For sample preparation:
- 1. Protaper Universal treatment rotary files (for biomechanical preparation)
- 2. Paper points

Sample Size

The sample size was estimated based on previously published data [15], assuming a power of 80%, significance level of 5%, and a mean difference of 23 in colony-forming unit (CFU) between groups. Group SDs were 17.33 and 2.81. These assumptions resulted in a minimum of 15 samples per subgroup. The complete sample size formula, derivation, and calculation steps are provided in Multimedia Appendix 1. Although the reference study involved a different irrigant (BioPure MTAD) and microbial model, it was the most comparable source available and served as a reasonable baseline for estimating variability in CFU reduction in the current study.

Sample Selection

The inclusion and exclusion criteria are given in Textbox 2.

Textbox 2. Inclusion and exclusion criteria.

Inclusion Criteria

- Freshly extracted single-rooted permanent mandibular lower premolar teeth with closed apex, free from caries, fracture, and previous endodontic treatment.

Exclusion Criteria

- Teeth with visible cracks, abnormal morphology.
- Teeth with developmental anomalies.
- Teeth with external and internal resorption.
- Teeth with multiple roots.
- Teeth with calcified canals.

Sample Allocation

Distribution of sample size was 90. The specimens will be randomly divided into 3 groups and 6 subgroups (Table 1).

XSL•FO
RenderX

Table . Sample distribution across experimental groups based on irrigant type and activation methodology.

Groups and subgroups	Samples (n)
Group I: 5 ml of 0.5% MTR[a] as an irrigant for 3 min	
Subgroup IA: Laser activation for 30 sec	15
Subgroup IB: Sonic activation	15
Group II: 5 ml of 2% CHX[b] as an irrigant for 3 min	
Subgroup IIA: Laser activation	15
Subgroup IIB: Sonic activation	15
Group III: 5 ml of normal saline as an irrigant for 3 min	
Subgroup IIIA: Laser activation	15
Subgroup IIIB: Sonic activation	15

[a]MTR: metronidazole.

[b]CHX: chlorhexidine.

Groups

This study aims to evaluate the antibacterial efficacy of 3 irrigant solutions—MTR, CHX, and normal saline—against *E. faecalis* using 2 activation methods (laser and sonic). The study will involve a total sample size of 90, divided into 3 main groups based on the intracanal irrigants used, and will be further divided into subgroups based on the activation protocol employed.

Sample Preparation

90 extracted single-rooted premolar teeth will be obtained from healthy patients with informed consent. Extractions will be performed for reasons unrelated to dental pathology, such as orthodontic treatment, prosthetic rehabilitation (eg, to facilitate fixed or removable partial dentures), esthetic concerns, or to preserve adjacent vital structures. Teeth exhibiting caries, restorations, root resorption, calcifications, or a history of endodontic treatment will be excluded from the study. In case of sample loss due to contamination, fracture, or invalid measurements during any phase of the study, replacements will be made from the remaining pool of eligible specimens to maintain balanced group sizes. All such instances will be recorded and reported.

Tooth Decoronation and Canal Preparation

The teeth will then be cleaned and sterilized, removing any residual tissue. Following the removal of any remaining tissue, the teeth will be washed and sterilized before being kept at room temperature in a 0.1% thymol solution until the experimental phase. To achieve a uniform root length of 13 mm, the crowns will thereafter be decoronated using a diamond disc. To get the working length, 1 mm will be subtracted from the total root length. From the total root length to establish the working length, 1 mm will be deducted. Root canal shaping will be carried out up to the F2 instrument.

Inoculation of *E. faecalis*

The strain of *E. faecalis* (ATCC 29212) used in this study will be obtained from the Department of Microbiology, Jawaharlal Nehru Medical College, Datta Meghe Institute of Medical Sciences (Deemed to be University), Wardha, India. The strains will be cultured and inoculated into the prepared canals of all samples. The samples will be incubated for 7 days to allow biofilm formation. Before irrigation, sterile paper points will be inserted into each canal for 60 seconds to collect pretreatment samples. These paper points will then be transferred to BHI Agar media and incubated for 48 hours. The colonies will then be counted using the classic bacterial counting method. Preirrigation samples will be taken. The inoculated teeth will then be randomly assigned to their respective groups and subgroups.

Biosafety Measures

All procedures involving *E. faecalis* will be conducted in a biosafety level-2 laboratory, following institutional biosafety guidelines. Sterile gloves, masks, and lab coats will be worn during microbial handling. All contaminated instruments and waste will be autoclaved and discarded in biohazard disposal systems. Teeth will be stored in 0.1% thymol prior to use to maintain sterility.

Irrigation Protocol

Irrigation will involve 3% sodium hypochlorite (Shivam, India), followed by 3 ml of 17% ethylene diamine tetra acetic acid for 1 minute to remove the smear layer, and subsequent irrigation protocol of every group will then be followed by using 5 ml of respective irrigants of 0.5% MTR, 2% CHX, and saline for 3 minutes accompanied by laser sonic activation with the EndoActivator (Dentsply Maillefer, Switzerland) operating at 10,000 cycles per minute (167 Hz) using a medium-sized tip (25/04). Each canal will be agitated for 30 seconds continuously. Laser activation will be performed using a diode laser (Biolase Epic X, Biolase Inc., USA) operating at a wavelength of 980 nm and a power setting of 1.5 W in pulsed mode. A 200-µm flexible fiber tip will be introduced 1 mm short of the working length. Each canal will undergo 3 cycles of laser activation, with each cycle lasting 5 seconds, followed by a 10-second interval between cycles. Postirrigation samples will be collected with the help of paper points and deposited into BHI Agar media and incubated for 48 hours.

XSL•FO
RenderX

Outcome Measures

The colonies will then be counted using the classic bacterial counting method to evaluate bacterial reduction. The data collected will then be subjected to statistical analysis. Scanning electron microscopy can offer visual confirmation of biofilm formation. The present study will follow a validated 7-day *E. faecalis* incubation protocol, as established in previous literature. In phase I of the study, antimicrobial efficacy will be quantitatively assessed using CFU counts, which provide a reliable and clinically relevant measure.

The procedure to be followed is depicted in the following flowchart (Figure 1).

Figure 1. Experimental protocol workflow. CFU: colony-forming units; *E. faecalis*: *Enterococcus faecalis*.

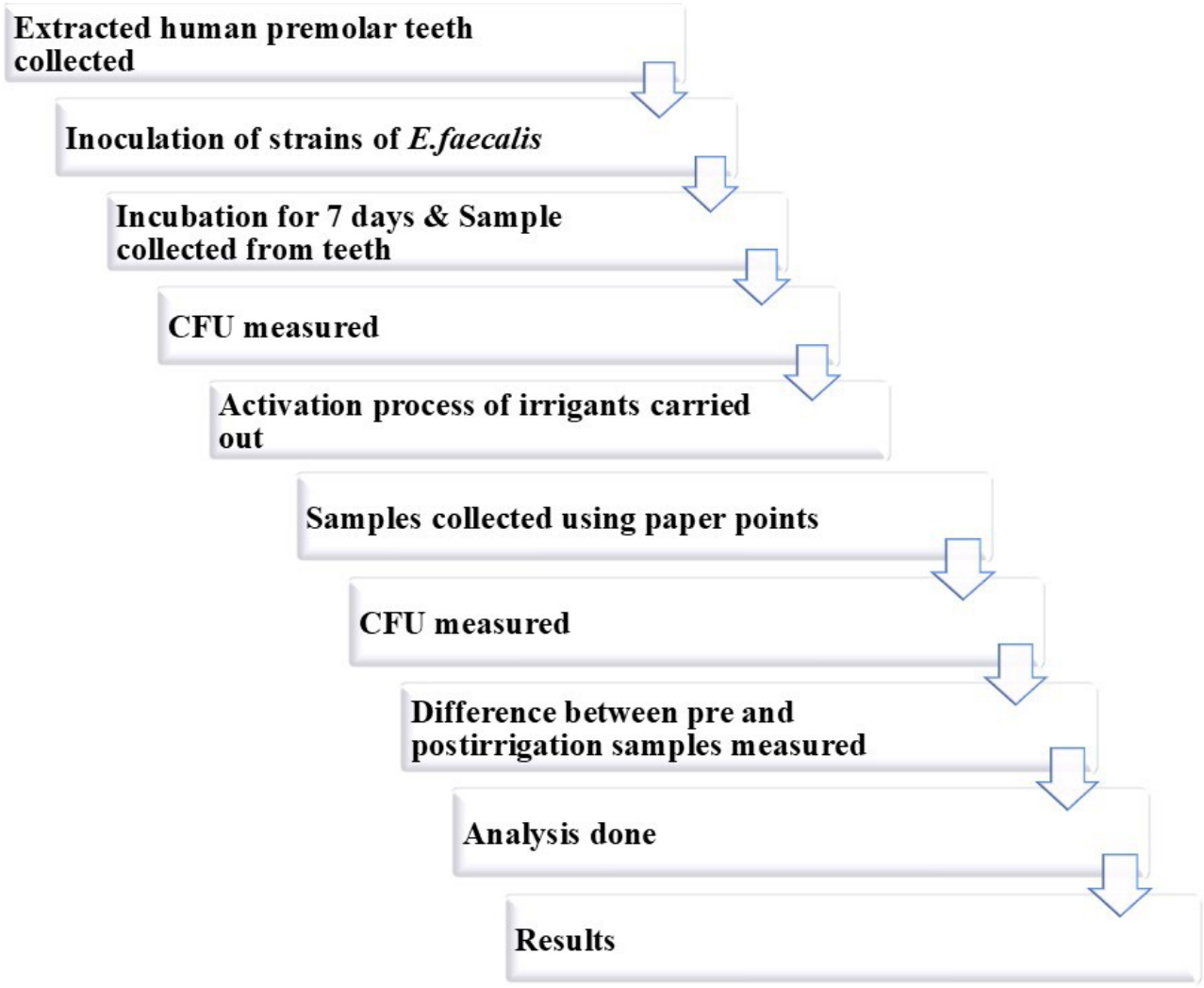

Statistical Analysis

The dataset will undergo both summarization and advanced statistical interpretation. Normality of data distribution will be assessed using the Shapiro-Wilk test. Analytical tools will include the ^{2} method, Student's *t* tests (applicable for both related and independent samples), as well as one-way ANOVA followed by Tukey's postanalysis comparisons. All statistical procedures and visual data representations will be executed using SPSS (version 27.0) and GraphPad Prism (version 7.0). A probability value $P<.05$ will be treated as indicative of a statistically meaningful result. The statistical methods to be applied are illustrated in Figure 2.

XSL•FO
RenderX

Figure 2. Stepwise statistical analysis workflow. CFU: colony-forming units; CHX: chlorhexidine; NS: normal saline.

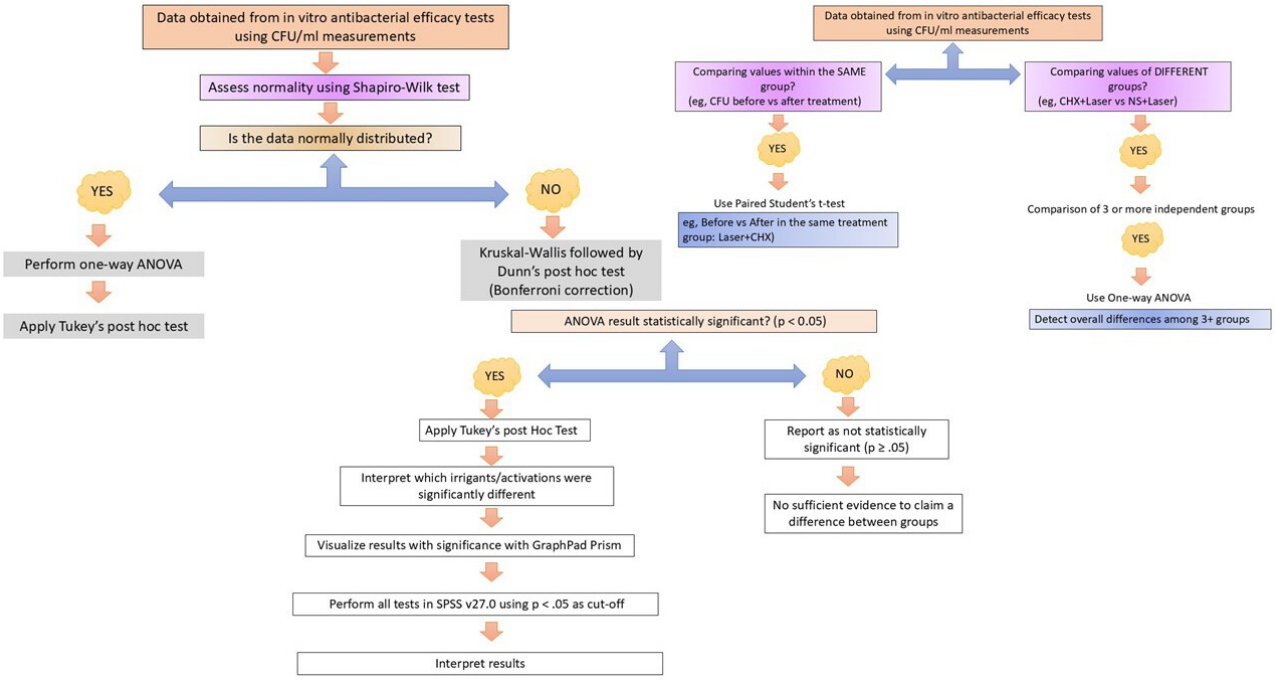

Ethical Considerations

Ethical approval for the study was obtained from the Institutional Ethical Committee of Datta Meghe Institute of Higher Education & Research, Sawangi, Wardha (Approval number: DMIHER(DU)/IEC/2025/545 dated: 07/02/2025). Written informed consent will be obtained from patients prior to extraction regarding the use of extracted teeth for research purposes in this study. All extracted teeth will be deidentified before inclusion in the experimental protocol to maintain complete anonymity of the donors. No financial or nonfinancial compensation was provided to the patients for participation in this study or for the use of extracted teeth for research purposes.

Results

Overview

Ethical approval was obtained in June 2025. Sample collection of extracted human teeth commenced in July 2025, and 90 extracted teeth have been collected as of January 2026, corresponding to the planned sample size. Laboratory procedures and data analysis have not yet commenced.

Primary Outcome Measure

The primary outcome measure was the evaluation and comparison of reduction in the CFU count of *E. faecalis* following irrigation by MTR with sonic and laser activation techniques with that of CHX and normal saline.

Secondary Outcome Measure

The secondary outcome measure was to determine the most effective combination of irrigant and activation technique to target and eliminate *E. faecalis* residing in the root canal space effectively.

Discussion

Achieving successful root canal outcomes requires the meticulous elimination of all microorganisms harbored within the complicated architecture of the root canal system. Despite advancements in irrigation protocols and antimicrobial strategies, *E. faecalis* remains a major causative agent in treatment failures due to its exceptional resistance and adaptability [1,2]. This facultative anaerobe is known for its ability to penetrate dentinal tubules, survive in nutrient-deprived environments, and resist common intracanal medicaments, including calcium hydroxide [3]. These characteristics underscore the need for not only potent antibacterial agents but also effective activation techniques to enhance their intracanal performance.

In this context, MTR and CHX have emerged as promising irrigants. MTR, a nitroimidazole derivative, is especially effective against anaerobic bacteria and has been used in both systemic and local applications in dental practice [4]. Although its antibacterial spectrum includes *E. faecalis*, the clinical efficacy of MTR may be compromised when used without adjunct activation techniques due to its limited ability to penetrate biofilms [5]. CHX, on the other hand, is a broad-spectrum cationic biguanide that retains its activity in the presence of organic matter and has substantivity, remaining active for prolonged durations postapplication [6]. It has shown superior performance in reducing microbial load within the canal system, especially when used at concentrations above 0.12% [6].

However, the delivery and penetration of these irrigants into canal irregularities are critically dependent on activation techniques. Sonic and laser activation have been shown to improve irrigant flow dynamics, resulting in deeper penetration and enhanced disinfection. Sonic activation systems such as the EndoActivator operate at lower frequencies (1 - 10 kHz),

XSL•FO
RenderX

creating hydrodynamic agitation that improves smear layer and debris removal. Their greater amplitude allows more effective lateral displacement of irrigants within the root canal [9].

Laser activation, particularly using diode lasers at 940 nm, represents a significant advancement in endodontic disinfection. The unique properties of laser energy—specifically its ability to induce photothermal and photoacoustic effects—result in deeper penetration into dentin and effective bacterial cell wall disruption [10,11]. Unlike conventional irrigants whose efficacy is often limited to the main canal, laser activation facilitates disinfection in otherwise inaccessible microanatomical structures, including lateral canals and apical ramifications [12]. Furthermore, diode lasers are advantageous due to their compact size, cost-effectiveness, and compatibility with clinical settings [11].

The present investigation is designed to explore an under-researched area by directly assessing the antimicrobial performance of MTR, CHX, and saline when activated via laser and sonic methods. While previous studies have highlighted the individual efficacy of these irrigants and activation methods, few have explored their synergistic impact, particularly with MTR as the primary irrigant. Given that metronidazole's efficacy is enhanced in anaerobic conditions and that *E. faecalis* can adapt to both aerobic and anaerobic environments, activation becomes a pivotal factor in optimizing its bactericidal action [4,5].

Moreover, CHX's substantivity and ability to disrupt microbial membranes at higher concentrations may complement the mechanical benefits of sonic and laser activation, potentially yielding superior outcomes compared to passive irrigation techniques [6,8]. In contrast, saline serves as a control, allowing for a clear demarcation of the role of activation in enhancing antibacterial performance.

Previous studies support the antibacterial potential of MTR as an endodontic irrigant. Amin et al [16] reported significant bacterial reduction with 5 ml of 0.5% MTR and 2% CHX in primary teeth, with saline showing no effect. Dubey et al [17] found MTR effective against *Bacteroides fragilis* and *Propionibacterium acnes*, while BioPure MTAD showed greater efficacy against *E. faecalis* and *Candida albicans*, highlighting MTR's limited spectrum. Shweta et al [15] demonstrated that the combination of 0.5% MTR and 2% glutaraldehyde exhibited enhanced antibacterial action over saline, with ultrasonic activation outperforming hand irrigation. These findings justify evaluating the combination of activated MTR and CHX in the present study to enhance efficacy against *E. faecalis*.

The incubation protocol is based on previously validated studies, such as those by Amin et al [16] and Shweta et al [15], which demonstrated reliable biofilm formation of *E. faecalis* following a 7-day incubation period under controlled conditions. These studies used similar in vitro infection models without requiring SEM confirmation to validate biofilm maturity, relying instead on established microbiological protocols and CFU quantification. As this study constitutes Phase I of a multiphase research design, our primary aim is to evaluate the antibacterial efficacy of different irrigants activated by laser and sonic techniques. While SEM is not incorporated in this phase due to resource constraints and focus on microbial reduction outcomes, Phase II is planned to include SEM analysis of selected samples to confirm dentinal tubule penetration and biofilm ultrastructure, thereby augmenting the qualitative depth of findings.

The exploration of enhanced antimicrobial strategies in endodontics, including the use of nanoparticle-modified sealers [18], reflects a broader effort to improve root canal disinfection outcomes. While the present study evaluates the in vitro efficacy of various irrigant–activation combinations, it also underscores the importance of translating such findings into clinical education and practice. As highlighted by Haupt and Kanzow [19], a persistent gap often exists between theoretical understanding and practical competence in endodontics, particularly during preclinical training. Addressing this gap may be critical for the successful clinical application of evidence-based disinfection protocols.

Overall, the findings of this study are anticipated to contribute significantly to clinical endodontics by providing evidence for more effective disinfection protocols. If MTR, when activated, proves comparable or superior to CHX, it may represent a cost-effective and biocompatible alternative for routine endodontic practice. This integrated approach may offer a promising direction toward reducing persistent infections and improving long-term treatment outcomes.

Acknowledgments

The author expresses sincere gratitude to the esteemed guide for constant support, expert guidance, and encouragement throughout the course of this research. The author is thankful to the Department of Conservative Dentistry and Endodontics, Sharad Pawar Dental College & Hospital, Datta Meghe Institute of Higher Education & Research, Sawangi, Wardha, for providing the necessary facilities and resources to carry out this study. The author is also grateful to the faculty members and laboratory technicians for their valuable suggestions and moral support. Special thanks are extended to the author's family for their unwavering love, patience, and motivation. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Sample size formula.
[DOCX File, 45 KB - resprot_v15i1e76783_app1.docx]

References

1. Kargül B, Tanboga I, Altinok B. Conventional endodontic treatment of primary molars using metronidazole as an intra-canal medicament: a pilot study. Eur Arch Paediatr Dent 2010 Aug;11(4):196-200. [doi: 10.1007/BF03262744] [Medline: 20840831]
2. Özsezer Demiryürek, E. Evaluation of microbial contamination of resilon and gutta-percha cones and their antimicrobial activities. Afr J Microbiol Res 2012;6(33):6275-6280 [FREE Full text] [doi: 10.5897/AJMR12.878]
3. Lee DK, Kim SV, Limansubroto AN, et al. Nanodiamond-Gutta Percha composite biomaterials for root canal therapy. ACS Nano 2015 Nov 24;9(11):11490-11501. [doi: 10.1021/acsnano.5b05718] [Medline: 26452304]
4. Rôças IN, Siqueira JF Jr, Santos KRN. Association of Enterococcus faecalis with different forms of periradicular diseases. J Endod 2004 May;30(5):315-320. [doi: 10.1097/00004770-200405000-00004] [Medline: 15107642]
5. Stuart CH, Schwartz SA, Beeson TJ, Owatz CB. Enterococcus faecalis: its role in root canal treatment failure and current concepts in retreatment. J Endod 2006 Feb;32(2):93-98. [doi: 10.1016/j.joen.2005.10.049] [Medline: 16427453]
6. Evans M, Davies JK, Sundqvist G, Figdor D. Mechanisms involved in the resistance of Enterococcus faecalis to calcium hydroxide. Int Endod J 2002 Mar;35(3):221-228. [doi: 10.1046/j.1365-2591.2002.00504.x] [Medline: 11985673]
7. Abou-Rass M, Bogen G. Microorganisms in closed periapical lesions. Int Endod J 1998 Jan;31(1):39-47. [Medline: 9823127]
8. Ghosh D, Bhardwaj S, Koyalada S, et al. Comparison of efficacy of ozonated water, normal saline, and povidone-iodine after surgical removal of impacted mandibular third molars: a cross-sectional study. J Family Med Prim Care 2020 Aug;9(8):4139-4144. [doi: 10.4103/jfmpc.jfmpc_534_20] [Medline: 33110822]
9. Susila A, Minu J. Activated irrigation vs. conventional non-activated irrigation in endodontics - a systematic review. Eur Endod J 2019;4(3):96-110. [doi: 10.14744/eej.2019.80774] [Medline: 32161895]
10. Virdee SS, Seymour DW, Farnell D, Bhamra G, Bhakta S. Efficacy of irrigant activation techniques in removing intracanal smear layer and debris from mature permanent teeth: a systematic review and meta-analysis. Int Endod J 2018 Jun;51(6):605-621. [doi: 10.1111/iej.12877] [Medline: 29178166]
11. Karade P, Chopade R, Patil S, et al. Efficiency of different endodontic irrigation and activation systems in removal of the smear layer: a scanning electron microscopy study. Iran Endod J 2017;12(4):414-418. [doi: 10.22037/iej.v12i4.9571] [Medline: 29225634]
12. Jurič IB, Anić I. The use of lasers in disinfection and cleaning of root canals: a review. Acta Stomatol Croat 2014 Mar;48(1):6-15. [doi: 10.15644/asc48/1/1] [Medline: 27688346]
13. Tilakchand M, Singh NN, Yeli MM, Naik BD. “Evaluation of the antibacterial efficacy of EZLASE diode LASER on the infected root canal system:” An in vivo study. J Conserv Dent 2018;21(3):306-310. [doi: 10.4103/JCD.JCD_14_18] [Medline: 29899635]
14. Martins MR, Franzen R, Depraet F, Gutknecht N. Rationale for using a double-wavelength (940 nm + 2780 nm) laser in endodontics: literature overview and proof-of-concept. Laser Dent Sci 2018 Mar;2(1):29-41. [doi: 10.1007/s41547-017-0017-9]
15. Jain S, Thakur S. Antibacterial efficacy of metronidazole and glutaraldehyde with hand and ultrasonic activation-an in vivo study. Unique Journal Of Medical and Dental Sciences 2013;01(1):35-59 [FREE Full text]
16. Barakat I, ElPatal MA, Abushanan A, Abd Elrady BEA. Antibacterial effect of metronidazole vs chlorhexidine solutions in treatment of root canals of primary anterior teeth. J Contemp Dent Pract 2020 Apr 1;21(4):396-399. [doi: 10.5005/jp-journals-10024-2719] [Medline: 32584275]
17. Dubey S, Saha SG, Rajkumar B, Dhole TK. Comparative antimicrobial efficacy of selected root canal irrigants on commonly isolated microorganisms in endodontic infection. Eur J Dent 2017;11(1):12-16. [doi: 10.4103/ejd.ejd_141_16] [Medline: 28435359]
18. Mohamed A, Ismail E, Adam RZ. A comparison of conventional root canal sealers with ones that use green synthesized nanoparticles for antimicrobial activity: protocol for a systematic review. JMIR Res Protoc 2024 Oct 11;13:e51351. [doi: 10.2196/51351] [Medline: 39393053]
19. Haupt F, Kanzow P. The relation between students’ theoretical knowledge and practical skills in endodontics: retrospective analysis. Interact J Med Res 2023 Apr 18;12:e46305. [doi: 10.2196/46305] [Medline: 36994987]

Abbreviations

BHI: brain heart infusion
CFU: colony-forming units
CHX: chlorhexidine
***E. faecalis*:** *Enterococcus faecalis*
MTR: metronidazole

XSL•FO
RenderX

RCT: root canal therapy

Edited by J Sarvestan; submitted 30.Apr.2025; peer-reviewed by RSG Mahmoud, T Amusa; revised version received 08.Aug.2025; accepted 26.Aug.2025; published 07.Jan.2026.

Please cite as:
Deshmukh S, Patel A
Comparative Evaluation of the Antibacterial Efficacy of Metronidazole, Chlorhexidine, and Normal Saline as Laser- and Sonic-Activated Intracanal Irrigants Against Enterococcus faecalis: In Vitro Study Protocol
JMIR Res Protoc 2026;15:e76783
URL: https://www.researchprotocols.org/2026/1/e76783
doi:10.2196/76783

© Sharvari Deshmukh, Aditya Patel. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 7.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

The Role of Iron-Overloaded Macrophages in Mesenchymal Stem Cell Senescence and Anemia in Myelodysplastic Syndromes: Protocol for an In Vitro Study

Lu Bai[1], MPH; Fan Wang[1], MPH; Yaxian Jiang[1], MPH; Hongmei Ouyang[1], MPH

Department of Clinical Laboratory, The First People's Hospital of Yunnan Province (Affiliated Hospital of Kunming University of Science and Technology), Kunming, Yunnan Province, China

Corresponding Author:
Hongmei Ouyang, MPH
Department of Clinical Laboratory
The First People's Hospital of Yunnan Province (Affiliated Hospital of Kunming University of Science and Technology)
157 Jinbi Road
Xishan District
Kunming, Yunnan Province
China
Phone: 86 15825291497
Email: ouyhmei@163.com

Abstract

Background: Myelodysplastic syndromes (MDSs) are a group of highly heterogeneous myeloid clonal diseases. Anemia is the most common clinical symptom, yet its pathogenesis remains incompletely understood. Preliminary evidence suggests an increase in macrophage infiltration and iron load in the bone marrow of patients with an MDS, alongside elevated interleukin-6 (IL-6) expression in bone marrow mesenchymal stem cells (BMSCs). The Kelch-like ECH-associated protein 1 (Keap1)–nuclear factor erythroid 2–related factor 2 (Nrf2)–antioxidant response element (ARE) pathway is a critical regulator of cellular antioxidant responses and inflammation, but its role in mediating the effects of iron overload in the microenvironment of patients with an MDS remains unclear.

Objective: This study aims to investigate the hypothesis that iron-overloaded bone marrow macrophages promote BMSC senescence and IL-6 secretion via the Keap1-Nrf2-ARE pathway, thereby impairing the survival and differentiation of hematopoietic stem cells and hematopoietic progenitor cells and contributing to anemia. Specific objectives are to (1) quantify the macrophage iron load across major MDS subtypes and correlate it with anemia severity and prognosis; (2) define the phenotype of erythroid island macrophages (CD68+, CD169+, and vascular cell adhesion molecule 1+) and the expression of endothelial microparticles in bone marrow tissues of patients with an MDS; and (3) experimentally test whether iron-overloaded macrophages and their exosomes regulate IL-6 secretion and senescence in BMSCs via the Keap1-Nrf2-ARE pathway.

Methods: Bone marrow samples will be collected from patients with key MDS subtypes (MDS with single-lineage dysplasia, refractory anemia with ring sideroblasts, MDS with multilineage dysplasia, and MDS with excess blasts; n=30 per subgroup) and control patients with iron-deficiency anemia (n=30). Methods will include histochemistry (Perls Prussian blue staining), immunohistochemistry and immunofluorescence (for macrophage and endothelial microparticle analysis), enzyme-linked immunosorbent assay, flow cytometry, quantitative polymerase chain reaction, and Western blotting. An in vitro model of iron-overloaded macrophages will be established using phorbol 12-myristate 13-acetate–differentiated THP-1 cells treated with ferric ammonium citrate. Exosomes will be isolated from these macrophages via ultracentrifugation. The effects of iron-overloaded macrophages and their exosomes on BMSC IL-6 secretion, senescence (senescence-associated β-galactosidase staining), and Keap1-Nrf2-ARE pathway activity will be assessed in coculture systems, with and without pharmacological inhibitors (5,6-dichloro-1-β-D-ribofuranosylbenzimidazole) or activators (dimethyl fumarate) of Nrf2.

Results: The study received ethics approval in December 2024. Patient recruitment and sample collection are in progress. As of December 2025, a total of 85 samples had been accrued. Pilot experiments to optimize macrophage differentiation and iron loading conditions have been completed. The full experimental workflow, including all sample analyses and in vitro experiments, is anticipated to be completed by May 2026.

XSL•FO
RenderX

Conclusions: This study is expected to elucidate a novel molecular mechanism linking iron overload in macrophages to BMSC dysfunction and anemia in MDS. The findings could identify the Keap1-Nrf2-ARE pathway as a potential therapeutic target for managing MDS-related anemia.

International Registered Report Identifier (IRRID): DERR1-10.2196/77936

(JMIR Res Protoc 2026;15:e77936) doi:10.2196/77936

KEYWORDS

myelodysplastic syndromes; mesenchymal stem cells; iron-overloaded macrophages; Keap1-Nrf2 pathway; anemia

Introduction

Myelodysplastic syndromes (MDS) are a group of highly heterogeneous myeloid clonal diseases, characterized by abnormal hematopoietic cell development in one or more myeloid lineages, manifesting as ineffective hematopoiesis and a high risk of transformation into acute myeloid leukemia. Anemia is a common clinical symptom and an important factor affecting the survival and quality of life of patients with MDS [1]. However, the current pathogenesis of anemia in MDS remains unclear.

About 80% to 85% of patients with MDS are anemic [2], and intracellular iron content is increased in erythrocytes. However, studies examining iron levels in bone marrow macrophages are limited. Anemia has been shown to be associated with abnormalities in the hematopoietic microenvironment [3]. The main hematopoietic cells in the hematopoietic microenvironment are macrophages, mesenchymal stem cells (MSCs), adipocytes (fat cells), sinusoidal endothelial cells, and fibroblasts [4,5]. Macrophages not only serve as the core components of the innate immune system but also play a role in phagocytosis, chemotaxis, the secretion of bioactive substances, and the participation and regulation of immune responses. In addition, macrophages are the main iron storage cells and key members of hematopoietic niches. We have completed the culture and identification of macrophages and MSCs and detected the increased infiltration density of CD68+, CD169+, and vascular cell adhesion molecule 1 (VCAM-1)+ macrophages in the bone marrow tissues of patients with MDS. However, the relationship between the intensity of iron staining in macrophages in MDS bone marrow tissue and anemia severity, disease subtype, and disease process requires further clarification. In addition, macrophages within erythroblastic islands provide iron to promote the differentiation and maturation of nucleated erythrocytes at all stages. Moreover, MSCs play an important role in the survival, differentiation, and maturation of hematopoietic stem cells (HSCs) or hematopoietic progenitor cells (HPCs). However, the correlation between iron-overloaded macrophages and their exosomes and the secretion of the inflammatory senescence-associated cytokine interleukin-6 (IL-6) by MSCs remains unclear, and the effects of this interaction on the survival, differentiation, and maturation of HSCs and HPCs remain to be determined.

The hematopoietic microenvironment is the site of survival, differentiation, and maturation of hematopoietic cells. It consists of reticular connective tissue and nonhematopoietic cells and is characterized by alterations in hematopoietic-matrix interactions, abnormal production of growth factors, and dysregulation of hematopoietic growth factor signaling [6-9]. The erythroid hematopoietic island is composed of 1 to 2 macrophages at the center and the surrounding erythrocytes at different stages of differentiation. Under physiological or pathological conditions, the functional state of macrophages can be regulated to affect erythropoiesis [10,11]. During erythroid maturation, macrophages promote the maturation of nucleated erythrocytes into reticulocytes by retaining immature erythrocytes in bone marrow and clearing old red blood cells. In the interaction between macrophages and erythroid nucleated cells, the anatomical structure and functional integrity of erythroid hematopoietic islands are necessary. Intercellular adhesion molecules are important components for maintaining the structural and functional integrity of erythroid hematopoietic islands [12]. Therefore, the phenotype of macrophages within erythroid islands in MDS bone marrow must be clarified, and the relationship between macrophage phenotype and the expression and differentiation of erythroid adhesion molecules requires further investigation.

Given the difficulty in obtaining macrophages from erythroid hematopoietic islands, research on their identification, function, and mechanism of action supporting erythroid development has been limited. Currently, the macrophage markers within erythroid hematopoietic islands include CD68, CD 169, and VCAM-1 [13,14]. CD68 is a canonical marker of human macrophages. CD 169, also known as sialoadhesin (Siglec-1), belongs to the sialic acid–binding immunoglobulin-like lectin family and is expressed in specific macrophage subsets. VCAM-1 is a vascular cell adhesion molecule that belongs to the immunoglobulin superfamily and assists in the homing of HSCs and HPCs. Li [14] demonstrated functional heterogeneity among erythroid island macrophages and identified at least 4 macrophage subgroups in the human fetal liver, including populations associated with early erythrocytes, late erythrocytes, and erythroblast nuclear phagocytosis [15,16]. However, the identification and functional analysis of macrophages within erythroid hematopoietic islands in the anemic bone marrow of patients with MDS are rare. On the basis of the existing clinical and laboratory resources, confocal immunofluorescence microscopy was proposed to analyze the expression of CD68+, CD 169+, and VCAM-1+ macrophages in the bone marrow tissues from included study samples and to preliminarily determine the phenotype and frequency of macrophages in the erythropoiesis island of bone marrow.

Molecules adhering to macrophages that interact with nucleated erythrocytes play an important role in promoting the differentiation and maturation of nucleated erythrocytes. According to Mao et al [15], endothelial microparticles (EMPs)

XSL•FO
RenderX

were closely associated with the formation of hematopoietic islands [17,18]. EMP antagonists can inhibit the proliferation, maturation, and enucleation of erythrocytes and greatly reduce the number of erythrocytes. EMP-knockout mice showed severe anemia. However, the expression of EMP in bone marrow tissue and its relationship with anemia and clinical characteristics remain unclear. Therefore, this study aims to detect the expression of EMP in bone marrow tissues and determine the clinical significance of EMP expression in anemia of MDS.

The exosome is the carrier of intercellular communication within the liquid microenvironment of the body and is a nanoscale extracellular vesicle secreted by living cells. It carries biological information molecules such as proteins, nucleic acids (mRNA, microRNA, DNA, etc) and lipids derived from parent cells, and it can regulate the physiological and pathological processes in target cells over short or long distances through body fluids [19]. Iron is not only an important raw material in erythropoiesis but also a key catalyst for redox reactions. It affects bone marrow hematopoiesis by modulating the level of oxidative stress in vivo [20-22]. However, the molecular mechanism by which the Kelch-like ECH-associated protein-1 (Keap1)–nuclear factor erythroid-2–related factor 2 (Nrf2)–antioxidant response element (ARE) signaling pathway mediates the effects of iron-overloaded macrophages and their exosomes on inflammatory cytokine IL-6 secretion and cellular senescence remains unclear.

The Keap1-Nrf2-ARE pathway plays a critical role in antioxidant responses, modulation of inflammation, and regulation of the immune balance of Th1/Th2 [23]. Nrf2 is the core transcription factor of an endogenous antioxidant response that regulates apoptosis, cell cycle, inflammatory cytokines, and immune response to induce tumor, inflammation, and aging-related diseases [24,25]. After the activation of the Keap1-Nrf2-ARE pathway, the expression of various downstream target proteins is upregulated, such as the antioxidant proteins and enzymes heme oxygenase 1, superoxide dismutase, glutathione peroxidase, and γ-glutamylcysteine synthetase. Iron overload was defined as serum ferritin levels greater than 1000 μg/L. Abnormal activation of the AMP-activated protein kinase-mitochondrial fission factor-dynamin-related protein 1, Wnt/β-Catenin, and reactive oxygen species-hypoxia-inducible factor 1α pathways has been reported to affect the secretion of cytokines and immunomodulatory functions of bone marrow MSCs (BMSCs) [19,26]. However, whether iron-overloaded macrophages and their exosomes regulate the oxidative stress and inflammatory state of BMSCs through the Keap1-Nrf2-ARE pathway remains unclear.

On the basis of the preliminary work of our team, we hypothesize that bone marrow iron-overloaded macrophages in patients with anemia regulate IL-6 secretion and cellular senescence through the Keap1-Nrf2-ARE pathway, thereby influencing the survival and differentiation of HSCs and HPCs and contributing to anemia [27-32]. We have completed the culture and identification of macrophages and MSCs, and we observed an increase in the infiltration density of macrophages and iron overload in bone marrow tissues [33]. MDS are highly heterogeneous diseases that are divided into 7 subtypes: MDS with single-lineage dysplasia, refractory anemia with ring sideroblasts (RARS), MDS with multilineage dysplasia, MDS with excess blasts, MDS unclassifiable, 5q syndrome, and refractory cytopenia of childhood. On the basis of the large sample size of MDS with subtypes including MDS with single-lineage dysplasia, RARS, MDS with multilineage dysplasia, and MDS with excess blasts, this study primarily aims to detect the iron load of bone marrow macrophages of these 4 MDS subtypes; analyze the correlation between iron load and MDS classification, anemia severity, prognosis, and other clinical and laboratory data; and determine the clinical significance of the macrophage iron overload of MDS bone marrow. In addition, the subtypes of erythropoietic macrophages and the expression of EMP in MDS bone marrow tissues will be determined. Furthermore, this study will investigate whether iron-overloaded macrophages and their exosomes regulate the expression of IL-6 and the senescence of MSCs through the Keap1-Nrf2-ARE pathway, thereby influencing the survival and differentiation of HSCs and HPCs and contributing to the occurrence and development of anemia.

This study aims to explore the clinical significance of iron overload in bone marrow macrophages by histochemistry, immunohistochemistry, enzyme-linked immunosorbent assay, flow cytometry, quantitative polymerase chain reaction, and Western blotting to determine the phenotype and function of erythroid hematopoietic island macrophages and to assess whether iron-overloaded macrophages upregulate the expression level of MSC-derived IL-6 via the Keap1-Nrf2-ARE pathway and promote cellular senescence, thereby affecting the survival and differentiation of HSCs and HPCs.

Methods

Design

The study design will be an in vitro cell biology study.

Time and Setting

This study will be performed at the First People's Hospital of Yunnan Province. The research is currently underway. Samples will be stored at −80 °C, and laboratory analyses will be performed every 2 weeks.

Ethical Considerations

This study was approved by the Clinical Research Ethics Committee of the First People's Hospital of Yunnan Province (KHLL2021-188). Written informed consent will be obtained from all participants before sample collection. All participant data and samples will be deidentified and coded to ensure confidentiality. Data will be stored on a secure, password-protected server. No participant compensation will be provided.

Participant Recruitment and Samples

Bone marrow biopsy specimens will be collected from patients diagnosed with MDS (including the following World Health Organization criteria subtypes: MDS with single-lineage dysplasia, RARS, MDS with multilineage dysplasia, and MDS with excess blasts) and control patients with iron-deficiency anemia (IDA). Inclusion criteria for the MDS group will be a

XSL•FO
RenderX

confirmed diagnosis by a hematologist, while exclusion criteria may include prior allogeneic stem cell transplantation or concurrent active malignancy. The control group will consist of patients with confirmed IDA. A target sample size of 30 per subgroup was estimated based on practical feasibility and preliminary data variability from previous similar studies. Clinically relevant information, including complete blood counts, serum ferritin, transfusion history, and cytogenetic data (eg, mutations in tet methylcytosine dioxygenase 2, additional sex combs like transcriptional regulator 1, and tumor protein p53), will be collected.

Experimental Procedure

Expression Intensity of Iron-Carrying Macrophages in Bone Marrow Tissues Determined by Prussian Blue Staining

Prussian blue staining is a classical histochemical reaction, used for the sensitive detection of ferric iron in tissue and is routinely performed in the laboratory. For semiquantitative analysis, the staining results will be independently evaluated by 2 experienced examiners.

Immunohistochemistry and Immunofluorescence Analysis

Immunofluorescence multiplex staining and immunohistochemistry were used to detect the expression of erythroid hematopoietic island CD68+, CD169 +, and VCAM-1+ macrophages and EMPs in the bone marrow tissues from the included study samples.

The paraffin-embedded sections of conventional tissue specimens will be prepared using standard processing procedures. The specific primary antibody and its corresponding secondary antibody will be selected to optimize the working concentration for multiple staining. Positive cell density will be quantified using image analysis software (eg, QuPath).

Establishment and Identification of Iron-Overloaded Macrophages In Vitro

Cell Culture

Resuscitated and expanded THP-1 cells will be cultured in RPMI-1640 complete medium containing 10% fetal bovine serum.

Induction of Macrophage Differentiation

THP-1 cells will be induced to differentiate into macrophages using 50 µg/mL of phorbol 12-myristate 13-acetate (PMA) for 24 hours and will be cultured in an RPMI-1640 complete medium without exosomes. Differentiation will be confirmed by flow cytometry analysis of surface markers CD11b and CD14.

Construction of Iron-Overloaded Macrophages

To construct iron-overloaded macrophages, multiple concentrations of ferric ammonium citrate (FAC; 10, 20, 40, 80, and 160 µmol/L) will be applied to PMA-induced THP-1 cells and cultured for 24 hours. The control group (THP-1+PMA) will receive equivalent volumes of phosphate-buffered saline (PBS). After the FAC treatment, cell viability will be assessed using the Cell Counting Kit-8 assay with absorbance measured at 450 nm. By comparing the optical density 450 values between the FAC treatment group and the control group, the effect of FAC concentration on cell viability will be analyzed, and the appropriate FAC concentration for constructing an iron overload model will be determined.

Model Validation

To verify the successful construction of the model, different concentrations of FAC (10, 20, 40, 80, and 160 µmol/L) will be added to PMA-induced M0 macrophages with THP-1 will be continued for 24 hours. After treatment, the Calcein-AM fluorescent probe will be added, and after incubation for a period, intracellular fluorescence intensity will be observed and recorded using a fluorescence microscope. Changes in fluorescence intensity will reflect alterations in the intracellular labile iron pool, confirming successful iron overload model establishment.

Extraction and Identification of Exosomes From Iron-Overloaded Macrophages and Culture and Identification of BMSCs

Our research group has established an experimental method for extracting exosomes by ultracentrifugation and using the SBI kit (ExoAb Antibody Kit, Shanghai Yanhui Biotechnology Co, Ltd). BMSCs will be extracted from bone marrow samples of patients with IDA based on our previous work [33].

BMSC Isolation and Primary Culture

Bone marrow mononuclear cells will be isolated from bone marrow aspirates obtained from patients with IDA (control group) using density gradient centrifugation. Briefly, bone marrow samples will be diluted with PBS and carefully layered over Ficoll-Paque PLUS (Cytiva). After centrifugation at 400 g for 30 minutes at room temperature, the bone marrow mononuclear cells layer at the interface will be collected, washed twice with PBS, and counted. The cells will be resuspended in complete culture medium consisting of α-minimum essential medium supplemented with 10% fetal bovine serum, 100 U/mL penicillin, and 100 µg/mL streptomycin. Cells will be seeded into culture flasks at a density of 1 to 2×10^5 cells/cm² and maintained at 37 °C in a humidified atmosphere of 5% CO_2. The medium will be changed every 3 to 4 days to remove nonadherent cells.

BMSC Identification by Flow Cytometry

The immunophenotype of the cultured BMSCs will be confirmed by flow cytometry. Briefly, passage 3 BMSCs will be harvested, washed, and incubated with fluorescein-conjugated antibodies against classic mesenchymal surface markers (eg, CD73, CD90, and CD105) and hematopoietic lineage markers (eg, CD34, CD45, and HLA-DR). Appropriate isotype controls will be used. A minimum of 10,000 events will be acquired using a flow cytometer (eg, BD FACSCanto II; BD Biosciences), and data will be analyzed using FlowJo software (BD Biosciences). BMSCs will be defined as positive for CD73, CD90, and CD105 (>95% positive) and negative for CD34, CD45, and HLA-DR (<2% positive).

XSL•FO
RenderX

Coculture of Iron-Overloaded Macrophages and Their Exosomes With BMSCs

A transwell coculture system will be used. The coculture model of iron-overloaded macrophages and their exosomes will be established with BMSCs, and the proportion of senescent cells among BMSCs will be detected by senescence-associated β-galactosidase staining. The expression of IL-6 will be detected by enzyme-linked immunosorbent assay or flow cytometry.

Gene and Protein Expression Analysis

Quantitative polymerase chain reaction will be used to detect the expression of Keap1, Nrf2 and ARE mRNA in BMSCs, as well as the intervention effect of the Nrf2 inhibitor 5,6-Dichloro-1-β-D-ribofuranosylbenzimidazole or activator dimethyl fumarate. A semiquantitative analysis of the β-actin gene will also be performed.

Western blot will be used to detect the expression of Nrf2 in BMSCs. The semiquantitative analysis will be performed using β-actin as the internal control protein. The Western blot experiment is a routine experimental process of the laboratory. The antibody will be determined by the protein expressed by the selected target gene. Proteins will be extracted using the radioimmunoprecipitation assay buffer kit and protein will be separated by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The experimental conditions will be optimized by changing the transmembrane conditions, the type and time of blocking reagent, and the concentration and incubation time of primary and secondary antibodies.

Isolation and Cultivation of HSCs and HPCs

HSC and HPC samples extracted from umbilical cord blood obtained from healthy term deliveries will be used to evaluate the effects of iron-overloaded macrophages on the survival and differentiation of HSCs and HPCs by flow cytometry and soft agar assays.

Main Observation Indicators

The basic morphology of bone marrow, the iron-staining intensity of macrophages, the phenotype of macrophages in erythroid hematopoietic islands, the expression of IL-6 in bone marrow, the expression of IL-6 in BMSCs, and the effects of iron-overloaded macrophages on the survival and differentiation of HSCs and HPCs will be observed in this study.

Statistical Analysis

Data analysis will be performed using SPSS software (version 26.0; IBM Corp). Descriptive data will be presented as mean (SD). Normality will be tested using the Shapiro-Wilk test. Comparisons between 2 groups will be performed using the Student *t* test (2-tailed; parametric) or Mann-Whitney *U* test (nonparametric). Comparisons among multiple groups will be performed using 1-way ANOVA with post hoc tests or the Kruskal-Wallis test. A *P* value <.05 will be considered statistically significant.

Results

Ethical approval was secured in December 2024. Patient recruitment and sample collection are underway. To date, pilot experiments have successfully established the protocols for THP-1 macrophage differentiation, FAC-induced iron loading, and BMSC culture. The main phase of sample processing and in vitro validation experiments is in progress. The timeline for the study anticipates completion of data collection by May 2026, followed by data analysis and manuscript preparation. A schematic timeline is presented in Table 1.

Table 1. Study timeline.

Activity	Start date	End date	Status
Ethics and study setup	December 2024	February 2025	Completed
Patient recruitment and sample collection	January 2025	February 2026	Ongoing
Pilot experiments (method optimization)	December 2024	March 2025	Completed
Main laboratory analysis	April 2025	May 2026	Ongoing
Data analysis	May 2026	June 2026	Planned
Manuscript preparation	June 2026	July 2026	Planned

Discussion

Anticipated Findings

We anticipate that this study will demonstrate a significant correlation between bone marrow macrophage iron load and the severity of anemia in MDS. Furthermore, we expect to show that iron-overloaded macrophages and their exosomes can induce cellular senescence and IL-6 secretion in BMSCs, and that this effect is mediated, at least in part, through the Keap1-Nrf2-ARE signaling pathway.

Comparison With Prior Work

Our previous work showed increased macrophage infiltration in MDS [33], consistent with other reports of microenvironment inflammation [3]. This study builds on that by focusing specifically on the iron content of these macrophages and its functional consequences. The link between iron overload and oxidative stress is well established [10], but its specific effect on the Keap1-Nrf2 axis in the context of BMSCs in MDS is novel. Our findings are expected to complement recent research on mitochondrial dysfunction in MSCs from patients with MDS [16].

XSL•FO
RenderX

Strengths and Limitations

In this study, the iron load of macrophages in the bone marrow microenvironment will be taken as the starting point to determine whether iron-overloaded macrophages promote the IL-6 expression in BMSCs and whether cellular senescence affects HSC survival and differentiation, thereby providing a new perspective for exploring the pathogenesis of different types of anemia. Elucidating the molecular mechanisms of MDS-related anemia and providing a theoretical basis for making diagnostic and treatment strategies are important to supplement and enrich the information of macrophage iron load in the bone marrow microenvironment, clarify the phenotype and frequency of macrophages in the erythroid hematopoietic islands of bone marrow, assess the effect of EMP expression on erythroid maturation and differentiation, determine the clinical significance of the iron load of macrophages, and confirm the molecular mechanism by which iron-overloaded macrophages promote IL-6 secretion and cellular senescence in BMSC.

A limitation, as noted by a reviewer, is the use of patients with IDA as the sole control group, which does not allow for the separation of the effects of anemia from those of iron status. While including an iron-deficient nonanemic control would be ideal, obtaining bone marrow samples from such individuals is ethically challenging. This limitation will be acknowledged, and the in vitro models will be critical for dissecting the specific role of iron.

In addition, this study has a limited sample size because of time and budget constraints. In the future, the sample size should be expanded to explore the diagnosis and treatment strategies of MDS-related anemia in different countries and races.

Future Directions and Dissemination

If our hypothesis is confirmed, future work will focus on validating these findings in animal models and exploring therapeutic interventions targeting the Keap1-Nrf2 pathway in MDS. The results of this study will be disseminated through publication in peer-reviewed scientific journals and presentations at international conferences.

Funding

This work was supported by Yunnan Provincial Department of Science and Technology-Kunming Medical University Applied Fundamental Research Joint Special Fund (grant 202201AY070001-237).

Data Availability

The datasets generated or analyzed during this study will be available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: LB
Data curation: FW, YJ
Formal analysis: FW, YJ
Funding acquisition: LB
Investigation: FW, YJ
Methodology: LB, FW, YJ
Supervision: HO
Writing—original draft: LB
Writing—review and editing: LB, HO

Conflicts of Interest

None declared.

References

1. Brownstein CG, Daguenet E, Guyotat D, Millet GY. Chronic fatigue in myelodysplastic syndromes: looking beyond anemia. Crit Rev Oncol Hematol 2020 Oct;154:103067. [doi: 10.1016/j.critrevonc.2020.103067] [Medline: 32739782]
2. Lewis R, Bewersdorf JP, Zeidan AM. Clinical management of anemia in patients with myelodysplastic syndromes: an update on emerging therapeutic options. Cancer Manag Res 2021 Jan;13:645-657. [doi: 10.2147/cmar.s240600]
3. Bewersdorf JP, Zeidan AM. Risk-adapted, individualized treatment strategies of myelodysplastic syndromes (MDS) and chronic myelomonocytic leukemia (CMML). Cancers (Basel) 2021 Mar 31;13(7):1610 [FREE Full text] [doi: 10.3390/cancers13071610] [Medline: 33807279]
4. Wu CH, Weng TF, Li JP, Wu KH. Biology and therapeutic properties of mesenchymal stem cells in leukemia. Int J Mol Sci 2024 Mar 21;25(5):645-657 [FREE Full text] [doi: 10.3390/ijms25052527] [Medline: 38473775]
5. Fenaux P, Kiladjian JJ, Platzbecker U. Luspatercept for the treatment of anemia in myelodysplastic syndromes and primary myelofibrosis. Blood 2019 Feb 21;133(8):790-794 [FREE Full text] [doi: 10.1182/blood-2018-11-876888] [Medline: 30602619]

XSL•FO
RenderX

6. Song D, Wu L. The role of macrophages in erythroid hematopoiesis. Chin J Exp Hematol 2017;25(4):1223-1226. [Medline: 28823297]
7. Dong PY, Huang LF, Sun HY. [Research progress of bone marrow microenvironment abnormalities in myelodysplastic syndrome]. Zhonghua Xue Ye Xue Za Zhi 2017 Jul 14;38(7):643-646 [FREE Full text] [doi: 10.3760/cma.j.issn.0253-2727.2017.07.022] [Medline: 28810341]
8. Galán-Díez M, Cuesta-Domínguez Á, Kousteni S. The bone marrow microenvironment in health and myeloid malignancy. Cold Spring Harb Perspect Med 2018 Jul 02;8(7):a031328 [FREE Full text] [doi: 10.1101/cshperspect.a031328] [Medline: 28963115]
9. Kokkaliaris KD. Dissecting the spatial bone marrow microenvironment of hematopoietic stem cells. Curr Opin Oncol 2020 Mar;32(2):154-161. [doi: 10.1097/CCO.0000000000000605] [Medline: 32022758]
10. Li W, Wang Y, Zhao H, Zhang H, Xu Y, Wang S, et al. Identification and transcriptome analysis of erythroblastic island macrophages. Blood 2019 Aug 01;134(5):480-491 [FREE Full text] [doi: 10.1182/blood.2019000430] [Medline: 31101625]
11. Seu KG, Papoin J, Fessler R, Hom J, Huang G, Mohandas N, et al. Unraveling macrophage heterogeneity in erythroblastic islands. Front Immunol 2017;8:1140 [FREE Full text] [doi: 10.3389/fimmu.2017.01140] [Medline: 28979259]
12. Javan GT, Can I, Yeboah F, Lee Y, Soni S. Novel interactions between erythroblast macrophage protein and cell migration. Blood Cells Mol Dis 2016 Sep;60:24-27. [doi: 10.1016/j.bcmd.2016.06.004] [Medline: 27519940]
13. Chow A, Huggins M, Ahmed J, Hashimoto D, Lucas D, Kunisaki Y, et al. CD169 macrophages provide a niche promoting erythropoiesis under homeostasis and stress. Nat Med 2013 Apr 17;19(4):429-436 [FREE Full text] [doi: 10.1038/nm.3057] [Medline: 23502962]
14. Li W. Identification and Biological Characteristics of Erythroid Hematopoietic Island Macrophages. Zhengzhou, China: Zhengzhou University; 2020.
15. Mao X, Shi X, Liu F, Li G, Hu L. Evaluation of erythroblast macrophage protein related to erythroblastic islands in patients with hematopoietic stem cell transplantation. Eur J Med Res 2013 Apr 08;18(1):9 [FREE Full text] [doi: 10.1186/2047-783X-18-9] [Medline: 23566571]
16. de Back DZ, Kostova EB, van Kraaij M, van den Berg TK, van Bruggen R. Of macrophages and red blood cells; a complex love story. Front Physiol 2014;5:9 [FREE Full text] [doi: 10.3389/fphys.2014.00009] [Medline: 24523696]
17. Chasis JA, Mohandas N. Erythroblastic islands: niches for erythropoiesis. Blood 2008 Aug 01;112(3):470-478 [FREE Full text] [doi: 10.1182/blood-2008-03-077883] [Medline: 18650462]
18. Li M, Li S, Du C, Zhang Y, Li Y, Chu L, et al. Exosomes from different cells: characteristics, modifications, and therapeutic applications. Eur J Med Chem 2020 Dec 01;207:112784. [doi: 10.1016/j.ejmech.2020.112784] [Medline: 33007722]
19. Zheng Q, Zhao Y, Guo J, Zhao S, Fei C, Xiao C, et al. Iron overload promotes mitochondrial fragmentation in mesenchymal stromal cells from myelodysplastic syndrome patients through activation of the AMPK/MFF/Drp1 pathway. Cell Death Dis 2018 May 01;9(5):515 [FREE Full text] [doi: 10.1038/s41419-018-0552-7] [Medline: 29725013]
20. Cilloni D, Ravera S, Calabrese C, Gaidano V, Niscola P, Balleari E, et al. Iron overload alters the energy metabolism in patients with myelodysplastic syndromes: results from the multicenter FISM BIOFER study. Sci Rep 2020 Jun 08;10(1):9156 [FREE Full text] [doi: 10.1038/s41598-020-66162-y] [Medline: 32514107]
21. Kim M, Hwang S, Park K, Kim SY, Lee YK, Lee DS. Increased expression of interferon signaling genes in the bone marrow microenvironment of myelodysplastic syndromes. PLoS One 2015 Mar 24;10(3):e0120602 [FREE Full text] [doi: 10.1371/journal.pone.0120602] [Medline: 25803272]
22. Hu J, Meng F, Hu X, Huang L, Liu H, Liu Z, et al. Iron overload regulate the cytokine of mesenchymal stromal cells through ROS/HIF-1α pathway in myelodysplastic syndromes. Leuk Res 2020 Jun;93:106354. [doi: 10.1016/j.leukres.2020.106354] [Medline: 32380365]
23. Zhou Q, Cen P, Chen Z, Jin J. Roles of the Keap1/Nrf2 pathway and mitophagy in liver diseases. J Zhejiang Univ Sci B 2025 Aug 28;26(10):972-994. [doi: 10.1631/jzus.B2400053] [Medline: 41116207]
24. Lin P, Ren Y, Yan X, Luo Y, Zhang H, Kesarwani M, et al. The high NRF2 expression confers chemotherapy resistance partly through up-regulated DUSP1 in myelodysplastic syndromes. Haematologica 2019 Mar 27;104(3):485-496 [FREE Full text] [doi: 10.3324/haematol.2018.197749] [Medline: 30262569]
25. Lu MC, Ji JA, Jiang ZY, You QD. The Keap1-Nrf2-ARE pathway as a potential preventive and therapeutic target: an update. Med Res Rev 2016 Sep 18;36(5):924-963. [doi: 10.1002/med.21396] [Medline: 27192495]
26. Falconi G, Fabiani E, Fianchi L, Criscuolo M, Raffaelli CS, Bellesi S, et al. Impairment of PI3K/AKT and WNT/β-catenin pathways in bone marrow mesenchymal stem cells isolated from patients with myelodysplastic syndromes. Exp Hematol 2016 Jan;44(1):75-83.e1 [FREE Full text] [doi: 10.1016/j.exphem.2015.10.005] [Medline: 26521017]
27. Liu Z, Lv X, Song E, Song Y. Fostered Nrf2 expression antagonizes iron overload and glutathione depletion to promote resistance of neuron-like cells to ferroptosis. Toxicol Appl Pharmacol 2020 Nov 15;407:115241. [doi: 10.1016/j.taap.2020.115241] [Medline: 32937103]
28. Han B, Li H, Chen F, Wu J. Causes and treatment strategies of anemia in myelodysplastic syndrome. Tianjin Med J 2018;46(8):794-798. [doi: 10.11958/20180561]
29. Buha A, Baralić K, Djukic-Cosic D, Bulat Z, Tinkov A, Panieri E, et al. The role of toxic metals and metalloids in Nrf2 signaling. Antioxidants (Basel) 2021 Apr 21;10(5):630 [FREE Full text] [doi: 10.3390/antiox10050630] [Medline: 33918986]

XSL•FO
RenderX

30. Gbotosho OT, Kapetanaki MG, Ross M, Ghosh S, Weidert F, Bullock GC, et al. Nrf2 deficiency in mice attenuates erythropoietic stress-related macrophage hypercellularity. Exp Hematol 2020 Apr;84:19-28.e4 [FREE Full text] [doi: 10.1016/j.exphem.2020.02.005] [Medline: 32151553]
31. Cao XL, Zhao M, Li D, Xing Y, Zhang Y, Chen J, et al. [Establishment of macrophage model of iron overload in vitro and the injury induced by oxidative stress on macrophage with iron overload]. Zhonghua Yi Xue Za Zhi 2016 Jan 12;96(2):129-133. [doi: 10.3760/cma.j.issn.0376-2491.2016.02.012] [Medline: 26792697]
32. Hao W, Song Y, Yao H, Du X, Hu K. Value of iron determination in bone marrow macrophages in differential diagnosis of anemia. J Qiqihar Med Coll 2007;28(10):1175-1178. [doi: 10.3969/j.issn.1002-1256.2007.10.010]
33. Song JX, OuYang HM, Dian ZJ, Wen Y. Study on the expression of tumor-associated macrophages in patients with myelodysplastic syndrome. Chong Med 2017;46(2):228-230. [doi: 10.3969/j.issn.1671-8348.2017.02.025]

Abbreviations

ARE: antioxidant response element
BMSC: bone marrow mesenchymal stem cell
EMP: endothelial microparticle
FAC: ferric ammonium citrate
HPC: hematopoietic progenitor cell
HSC: hematopoietic stem cell
IDA: iron-deficiency anemia
IL-6: interleukin-6
Keap1: Kelch-like ECH-associated protein 1
MDS: myelodysplastic syndromes
MSC: mesenchymal stem cell
Nrf2: nuclear factor erythroid 2–related factor 2
PBS: phosphate-buffered saline
PMA: phorbol 12-myristate 13-acetate
RARS: refractory anemia with ring sideroblasts
VCAM-1: vascular cell adhesion molecule 1

Edited by A Schwartz; submitted 22.May.2025; peer-reviewed by A Murakonda, N Brathaban; comments to author 16.Oct.2025; accepted 11.Dec.2025; published 13.Jan.2026.

Please cite as:
Bai L, Wang F, Jiang Y, Ouyang H
The Role of Iron-Overloaded Macrophages in Mesenchymal Stem Cell Senescence and Anemia in Myelodysplastic Syndromes: Protocol for an In Vitro Study
JMIR Res Protoc 2026;15:e77936
URL: https://www.researchprotocols.org/2026/1/e77936
doi:10.2196/77936
PMID:

©Lu Bai, Fan Wang, Yaxian Jiang, Hongmei Ouyang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Rapid Qualitative Approaches in Pandemic Research: Protocol for an Exploratory Qualitative Multimethod Study (VERDIQual) on Mpox in Italy, Nigeria, Thailand, and the United Kingdom

Chinye Osa-Afiana[1], MPH; Marthe Le Prevost[2], MSc; Emily Jay Nicholls[3], PhD; Davide Bilardi[4,5], DPhil; Thomas E Guadamuz[6], PhD; Esekwe E Soje-Amadosi[1], MPH; Grace I Adebisi[1], MPH; Chibueze Adirieje[1], MPH; Adeyosola A Adetunji[7], MD; Karima Yusufu[7], MD; T Charles Witzel[3,6], PhD; Worawalan Waratworawan[6,8], MA; Nattharat Samoh[6], MA; Tom May[9,10], PhD; Sarah Denford[9,10], PhD; Willian Gomes[11], MA; Uzodinma Adirieje[12], PhD; Shema Tariq[3,13*], MD, PhD; Nadia Adjoa Sam-Agudu[1,7,14,15*], MD

[1]International Research Center of Excellence, Institute of Human Virology Nigeria, Abuja, Nigeria

[2]MRC Clinical Trials Unit, University College London, London, United Kingdom

[3]Institute for Global Health, University College London, London, United Kingdom

[4]Fondazione Penta, Padua, Italy

[5]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, England, United Kingdom

[6]Mahidol Center for Health, Behavior and Society, Faculty of Tropical Medicine, Mahidol University, Nakorn Pathom, Thailand

[7]Pediatric and Adolescent ART Unit, Department of Prevention, Care and Treatment, Institute of Human Virology Nigeria, Abuja, Nigeria

[8]Faculty of Social Sciences and Humanities, Department of Society and Health, Mahidol University, Nakorn Pathom, Thailand

[9]NIHR Health Protection Research Unit in Evaluation and Behavioural Science, University of Bristol, Bristol, United Kingdom

[10]Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, United Kingdom

[11]European AIDS Treatment Group, Lisbon, Portugal

[12]Afrihealth Optonet Association, Abuja, Nigeria

[13]Mortimer Market Centre, Central and Northwest London, NHS Foundation Trust, London, England, United Kingdom

[14]Department of Paediatrics and Child Health, School of Medical Sciences, University of Cape Coast, Cape Coast, Central, Ghana

[15]Global Pediatrics Program and Division of Pediatric Infectious Diseases, Department of Pediatrics, University of Minnesota Medical School, Minneapolis, MN, United States

*these authors contributed equally

Corresponding Author:
Nadia Adjoa Sam-Agudu, MD
International Research Center of Excellence
Institute of Human Virology Nigeria
Plot 62, Emeritus Umaru Shehu Avenue
Cadastral Zone COO
Abuja
Nigeria
Phone: 234 9461 0342
Email: nsamagudu@ihvnigeria.org

Abstract

Background: Recent mpox outbreaks have underscored significant gaps in global preparedness for emerging and re-emerging infections. These outbreaks have disproportionately affected vulnerable and marginalized populations, exposing the weaknesses of health systems, particularly in resource-limited settings. The global spread of mpox beyond endemic African countries in 2022 and the emergence of a new Clade Ib in 2024 emphasize the pressing need for comprehensive and context-specific public health responses. We outline the protocol for an innovative multimethod qualitative study (VERDIQual [SARS-CoV-2 (and Mpox) Variants Evaluation in Pregnancy and Paediatrics Cohorts Qualitative Study]). This study is being conducted across 4 different countries and settings—Italy, Nigeria, Thailand, and the United Kingdom.

XSL•FO
RenderX

Objective: VERDIQual uses multiple qualitative methods to explore the lived experiences of different populations with mpox and frontline health care workers in endemic and nonendemic settings. With this approach, we aim to identify missed and new opportunities for effective public health messaging on prevention and treatment in this and future pandemics.

Methods: VERDIQual's flexible, multimethod approach integrates content analysis of news and social media, focus group discussions, semistructured interviews, and participatory photography. We apply intersectionality theory to capture perspectives from a diverse range of participants, including gay, bisexual, and other men who have sex with men, pregnant women, adolescents with mpox, and frontline health care workers providing mpox services. Data will be collected and analyzed using the rapid assessment procedure approach, enabling real-time data synthesis to provide timely and contextually relevant insights. We use a standardized approach to data integration at the interpretation stage, summarizing findings in a multimethod integration matrix to visualize and synthesize data by objective, method, and site.

Results: Data collection tools, including focus group discussions and semistructured interview topic guides, have been developed in collaboration with local community advisory boards. Data collection was completed by October 31, 2025, and analysis is ongoing. Teams across the 4 countries have identified media houses and social media platforms for inclusion in the news and social media analysis. Study results will be disseminated through peer-reviewed journals, scientific conferences, policy briefs, and stakeholder workshops and community events, with a focus on informing equitable and inclusive future public health responses to re-emerging infections.

Conclusions: We expect our principal findings to be applicable to a range of settings. Our use of intersectionality theory will also facilitate considerations for intersecting identities and characteristics in equity-centered pandemic responses. Ultimately, we expect VERDIQual to inform pandemic preparedness, inclusive of people with stigmatized and vulnerable characteristics or identities.

International Registered Report Identifier (IRRID): DERR1-10.2196/77321

(JMIR Res Protoc 2026;15:e77321) doi:10.2196/77321

KEYWORDS

pandemics; mpox; monkeypox; global health; qualitative research

Introduction

The recent COVID-19 and mpox pandemics have brought attention to global preparedness for emerging and re-emerging infectious disease outbreaks [1-3]. Certain populations have borne the brunt of shortfalls in health systems' infrastructure and preparedness. For example, children and pregnant women have experienced exclusion from intervention studies, delayed vaccination, delayed or inadequate treatment options, and poor inclusion in public health policy [4]. Furthermore, marginalized groups such as those who are racially or gender and sexual minoritized are often at greatest risk of infection and poorer clinical outcomes [5,6]. The needs of such key populations should be considered and promptly addressed at all stages of pandemics, including prepandemic planning and postpandemic surveillance.

Since the first reported case of human mpox in the Democratic Republic of the Congo in 1970 [7], mpox case reports have largely been limited to endemic countries in West and Central Africa [7]. The rare cases reported in nonendemic countries were linked to travel to or exposure to animals transported from an endemic African country [8]. This changed in May 2022, when a large number of mpox cases were first reported in European countries, later spreading outside Europe [8]. The World Health Organization (WHO) declared mpox a Public Health Emergency of International Concern in July 2022, in response to mpox being reported in more than 70 countries [8]. In May 2023, the WHO announced the end of mpox's status as a Public Health Emergency of International Concern due to sustained reduction in incident cases [9]. However, in 2024, a new outbreak of mpox was identified in Central Africa, attributed to the emergence of a distinct Clade Ib variant [9,10].

Over 85% of cases in the global 2022 mpox outbreak were among gay, bisexual, and other men who have sex with men; also, HIV coinfection was reported in over 50% of global mpox cases [9]. The ongoing Clade Ib outbreak has a different epidemiological profile, having emerged in eastern Democratic Republic of the Congo in 2024 [10]. It predominantly affects nonendemic areas within the Democratic Republic of the Congo and neighboring countries, with transmission occurring across different biological sexes (male and female) and sexual behavior. Infection is largely influenced by factors such as individual exposure risk and immune status, with children and pregnant women being at greatest risk due to lower immunity and increased vulnerability to poorer health outcomes [9].

Mpox epidemiology in African countries contrasts with global data. Reports from the WHO African region indicate a general population profile, with endemic West and Central African countries bearing over 99% of the case burden [8,9]. There is a paucity of robust data on mpox incidence and mortality in children [8,9], pregnant women, and groups at high risk (such as people living with HIV), all of whom are the most affected during infectious disease outbreaks [8].

Like other infectious disease outbreaks, the recent mpox outbreaks are complex biosocial phenomena, disproportionately impacting marginalized and vulnerable groups and exposing inadequacies in pandemic responses. The HIV [11] and COVID-19 pandemics [12] have highlighted the central importance of considering social contexts in planning and implementing outbreak responses.

XSL•FO
RenderX

Qualitative research plays a crucial role in understanding the interplay of human behavior, social dynamics, and public health responses. It helps uncover the social, political, and cultural factors that shape experiences and influence the implementation of health interventions. For example, qualitative data can provide insights into decision-making about vaccination and perceptions of public health messaging [13,14]. This allows for public health policy and interventions to be tailored to specific populations while minimizing potential harms such as stigma. Furthermore, such approaches generate rich insights into lived experiences, yielding contextual understanding of the multidimensional impacts of pandemics [15].

Epidemiological data indicate that children and young male adults have been the most affected in Africa [16]. Beyond intense bias against sexual and gender minorities, mpox stigma in these settings is shaped by fear of contagion, the visibility of skin lesions, misinformation, and association of infection with poverty or perceived promiscuity [17,18]. These drivers of stigma intersect with systemic issues such as social inequality and weak health infrastructure [19-21]. Studies conducted among gay, bisexual, and other men who have sex with men with mpox in Australia, China, and the United Kingdom have found that stigma was central to their experiences and was often exacerbated at health facilities [22-25]. Other studies have focused on the role of news and social media (eg, Twitter [subsequently rebranded X] and Reddit) in spreading or countering mis- or disinformation [26], with less data available to understand information-sharing on newer social media platforms such as TikTok and Instagram.

There has been limited qualitative research on mpox, especially in endemic African countries. Studies have predominantly focused on gay, bisexual, and other men who have sex with men and high-income settings, with a lack of data on experiences in endemic African regions and other resource-limited settings [14]. In addition, there is a relative neglect of how multiple experiences of marginalization intersect to shape experiences of mpox and prevention. Furthermore, limited community engagement in research design and implementation results in studies that insufficiently address community priorities and needs. We aim to directly address these gaps through our expanded geographic and population scope, intersectionality-informed analyses, strong patient and public involvement, and inclusive and innovative methodologies, drawing from data that have been underused in mpox research.

The European Union–funded VERDI (SARS-CoV-2 (and Mpox) Variants Evaluation in Pregnancy and Paediatrics Cohorts) consortium is a global partnership between 30 centers of excellence in Europe, the United States, Africa, the Caribbean, the Middle East, and Southeast Asia [27]. VERDI prioritizes children, pregnant women, and vulnerable groups in current and future pandemic research. The consortium also aims to inform preparedness and public health responses for future infectious disease outbreaks in these populations, drawing upon recent experiences with COVID-19 and mpox. One of VERDI's projects is a collaborative multicountry qualitative study on mpox (VERDIQual [SARS-CoV-2 (and Mpox) Variants Evaluation in Pregnancy and Paediatrics Cohorts Qualitative Study]), which we describe in this paper [27].

VERDIQual supports the broader goals of the VERDI consortium by providing a contextual understanding of the experiences and needs of diverse populations in different geographical settings during the 2022 global mpox outbreak. Our findings are expected to inform more equitable access to relevant interventions, reduce stigma, and improve public health messaging and the research response for ongoing and future pandemics. Specific objectives are (1) to understand the lived experiences of individuals who acquired or were otherwise affected by mpox, (2) to identify missed and new opportunities for effective public health messaging on prevention and treatment, and (3) to develop capacity and establish methods for collaborative rapid qualitative research in response to outbreaks of emerging and re-emerging infections.

Methods

Overall Study Design

This exploratory multimethod, multisite qualitative study comprises content analysis of news and social media, focus group discussions (FGDs), semistructured interviews (SSIs), and participatory photography and uses rapid qualitative analysis approaches. The study is being conducted across 4 countries: Italy, Nigeria, Thailand, and the United Kingdom. Study teams in each country have received rigorous training and are experienced in each of these methods. Study countries were purposely selected based on the number of reported mpox cases, geographical spread (Africa, Asia, and Europe), diversity of mpox epidemiology (sexual and gender minority vs general populations in nonendemic vs endemic settings), and resources (high-income [United Kingdom and Italy] vs low- and middle-income countries [Nigeria and Thailand]). The methods used in each country will vary according to resources and feasibility (Table 1). For example, participatory photography was considered feasible and acceptable for target study participants with recently acquired mpox in Thailand; however, this was a less familiar method in Nigeria, and there was a concern for exacerbation of stigma or self-stigma and low acceptability among people with mpox and their communities. The UK team could not conduct SSIs among people with recently acquired mpox due to the significant decline in cases by late 2022. The flexibility for different country teams to use different methods across reflects VERDIQual's commitment to country team autonomy and the accommodation of local realities and norms with respect to the feasibility and acceptability of the different research approaches. We present our study design using the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist [28] (Multimedia Appendix 1).

XSL•FO
RenderX

Table 1. VERDIQual[a] data collection methods by country.

	News media	Social media	Focus groups	Interviews	Photography
Italy	✓				
Nigeria	✓	✓	✓	✓	
Thailand		✓	✓	✓	✓
United Kingdom	✓	✓	✓		

[a]VERDIQual: SARS-CoV-2 (and Mpox) Variants Evaluation in Pregnancy and Paediatrics Cohorts Qualitative Study.

Theoretical Framework

This study is informed by intersectionality theory, a framework for understanding how multiple social categories, such as race, gender, sex, sexuality, and geographic residence, interact to shape unique experiences of privilege and oppression that cannot be captured by considering each category independently [29]. The principles of intersectionality theory will be embedded throughout the life cycle of the study. We are specifically interested in how multiple social identities and structural forces combine to influence risk and experiences of mpox, resisting reducing populations to homogenous groups. We will do this by sampling participants with diverse characteristics and identities in terms of race, gender (being a man or woman), sex (male or female), sexuality (heterosexual, lesbian, gay, bisexual, or transgender), and HIV status. For example, in the United Kingdom, we have collaborated with a charity advocating for the sexual health and well-being needs of racially minoritized gay, bisexual, and other men who have sex with men. In Nigeria, we have partnered with an organization with extensive nationwide experience in community-based research and community-level health promotion. In analyses, we will use coding frameworks that are sensitive to co-occurring and overlapping forms of disadvantage and privilege, critically reflecting on how multiple identities interact to produce mpox-related stigma. Finally, comparative analyses will be conducted across all sites to yield insights into geographic variation in experiences of and attitudes to mpox [29].

Study Settings and Populations

Italy

Between 2022 and 2023, Italy reported approximately 955 confirmed cases of mpox [30], as part of the larger global 2022 outbreak. The majority (98%) of these cases were gay, bisexual, and other men who have sex with men [30]. Italian health authorities, including the Istituto Superiore di Sanità, responded swiftly by implementing measures such as contact tracing, targeted vaccination campaigns, and public health advice. The vaccination campaign was targeting “all people at risk” and was provided for free with the provision of gratuity for the Italian health system [30]. By early 2023, the incidence of mpox had significantly declined in Italy, partly due to vaccination and behavioral change.

Nigeria

Nigeria has reported mpox since 1971 and had its largest outbreak in 2017, before the global 2022 pandemic. However, public awareness about mpox is low, particularly in rural areas—perpetuating stigma, underreporting of cases, and reluctance to seek medical care [31,32]. Vaccines are an effective prevention tool; however, they remain unavailable in most African countries [33,34]. A positive development is that Nigeria received a 10,000-dose Modified Vaccinia Ankara vaccine donation from the US government in August 2024 [35], and in September 2025, initiated a targeted hotspot vaccination campaign with WHO-prequalified Modified Vaccinia Ankara-Bavarian Nordic mpox vaccines from Gavi, the Vaccine Alliance [36].

Thailand

In Thailand, the mpox outbreak also predominantly affects male individuals (97%), particularly gay, bisexual, and other men who have sex with men, many of whom also live with HIV [37,38]. The urgency of addressing mpox is underscored by Thailand’s approval of an mpox vaccination program, which commenced in March 2024 in Bangkok. This vaccine program was in the beginning limited to 1 health care provider (Red Cross Anonymous Clinic) [38], but it is now more widely available; however, it is still mostly in Bangkok and is prohibitively expensive for individuals, at between US $125 and US $485 (with an average monthly salary of around US $600) [39].

The United Kingdom

Between May 2022 and September 2023, there were over 3700 cases of mpox identified in the United Kingdom, peaking in July 2022 [40]. In total, 99% of cases in the United Kingdom were male; 96% identified as gay, bisexual, and other men who have sex with men [41]. No mpox-related deaths have been reported in the United Kingdom [42]. The monthly number of Clade IIb cases in England was between 1 and 41 in 2023-2025, compared to the peak of 1139 cases in July 2022 [43]. Public health responses to mpox in the United Kingdom included increasing public awareness, advice to self-isolate for 21 days, and contact tracing [44]. The UK mpox vaccination program started in June 2022, providing free vaccines to gay, bisexual, and other men who have sex with men considered to be at higher risk of infection, contacts of people with mpox, and people with potential occupational exposure.

Reflecting the differences in epidemiology between endemic and nonendemic settings, study populations of interest differ by country. In Nigeria, the focus is on children and pregnant and postpartum women, whereas in Thailand and the United Kingdom, it is on gay, bisexual, and other men who have sex with men—the population most affected in nonendemic countries. Both Nigerian and Thai studies will also recruit health

XSL•FO
RenderX

care professionals. In Italy, methods are restricted to news media analysis, with no primary data collection.

Data Collection

News Media (Italy, Nigeria, and the United Kingdom)

A systematic search of published news articles will be carried out in the 3 countries to explore representations of mpox in news media. Searches will be conducted of both online and in-print articles using keywords including "mpox," "monkeypox," "monkey pox," and "MPX." Publications will be selected using criteria such as scope of circulation and political alignment. The following newspapers will be included in the country analyses: *La Repubblica* and *Corriere della Sera* in Italy; *the Vanguard*, *the Nation*, *the Guardian*, *theSun*, and *the Punch* in Nigeria; and *theGuardian*, *Sun*, *Times*, *Daily Mail*, and *Daily Mirror* in the United Kingdom.

Social Media (Nigeria, Thailand, and the United Kingdom)

We will conduct a systematic search of selected social media platforms to explore representations of mpox outbreaks and public discourse about mpox (including vaccination). Searches will use keywords such as "mpox," "monkeypox," "monkey pox," and "MPX." In Thailand, Thai-specific keywords such as [illegible] (*Phidard*), [illegible] (*Phidardling*), and [illegible] (*Phidardwanon*) will be added. In the United Kingdom and Thailand, the analysis will focus on TikTok, whereas in Nigeria, it will focus on Instagram and Facebook. For these analyses, as content is shown to users algorithmically, user accounts will be created using the anticipated demographics of the populations of interest, and these accounts will be used only for research purposes. These platforms were selected for each country by popularity, access to data, and (for TikTok) methodological novelty. Multimodal data (including text, images, video, and sound) will be captured using standardized forms and analyzed thematically.

FGDs (Nigeria, Thailand, and the United Kingdom)

FGDs will be conducted to gain insights from participants about lived experiences of mpox [45], providing an in-depth understanding of participants' knowledge and perceptions regarding the mpox outbreak. Topic guides will cover public perceptions, sources of information, attitudes toward transmission-reducing behaviors, public health messaging, and social and emotional impacts of mpox. FGD sampling will vary across settings depending on local mpox epidemiology (Table 2). The number of FGDs in each context will be determined by the number of different groups invited to participate, bearing in mind that 90% of themes are likely to be covered in 3-6 focus groups [46]. FGDs will comprise 3-10 participants each, depending on recruitment (which may be more or less challenging depending on stigma and other factors in local contexts).

Table 2. Populations of interest for VERDIQual[a] interviews and focus group discussions.

Country	Period of mpox outbreak	Caregivers of people with mpox	People with mpox			Frontline health care workers (aged ≥18 years)
		Adult caregivers of children (aged ≤11 years)	Adolescents (aged 12-17 years)	Pregnant women and nursing mothers (aged ≥15 years)	Adults (aged ≥18 years)	
Nigeria	September 1, 2017[b], to present	Yes	Yes	Yes	Yes, general population	Yes
Thailand	July 1, 2022, to present	No	No	No	Yes[c]	Yes
United Kingdom	May 1, 2022[d], to present	No	No	No	Yes[e]	No

[a]VERDIQual: SARS-CoV-2 (and Mpox) Variants Evaluation in Pregnancy and Paediatrics Cohorts Qualitative Study.

[b]This date signifies the year of the largest mpox outbreak in Nigeria.

[c]Sexual and gender minorities, including gay, bisexual, and other men who have sex with men, and other people at high risk of mpox, including sex workers, people who use drugs, and people living with HIV.

[d]This date signifies the time point immediately preceding the first global mpox reports in Europe.

[e]Gay, bisexual, and other men who have sex with men.

In Nigeria, 10-15 in-person FGDs will be conducted with individuals who have acquired mpox (children and adolescents aged ≤17 years [or their caregivers where appropriate], pregnant or nursing women aged ≥15 years, and adults aged ≥18 years) and individuals aged ≥18 years affected by mpox (caregivers and health care workers). Recruitment will be supported by a community advisory board drawn from a local organization (Afrihealth Optonet Association), which will facilitate the identification, selection, and engagement of communities where mpox cases have been identified.

In Thailand, up to 8 in-person FGDs will be conducted with individuals who perceive themselves to be at risk of acquiring mpox (people who self-identify as gay, bisexual, and other men who have sex with men, transgender women, male and female sex workers, people who use drugs, and people living with HIV) and health care providers who provide care to patients with mpox, recruited from public health agencies and nongovernmental organizations (such as the Thailand Ministry of Public Health through the Department of Disease Control [Bamrasnaradura Infectious Diseases Institute and Bureau of Epidemiology], the Institute of HIV Research and Innovation,

XSL•FO
RenderX

and the Hospital for Tropical Diseases). Recruitment will be supported by local nongovernmental organizations with strong community ties to the target populations. Specifically, male and female sex worker participants will be recruited with support from the Service Workers in Group Foundation; men who have sex with men who use drugs will be engaged through the Health and Opportunity Network; and men who have sex with men living with HIV will receive additional support from The Poz Home Thailand Foundation.

In the United Kingdom, 4-6 in-person or online [47] (participant choice) FGDs will be conducted with gay, bisexual, and other men who have sex with men who perceive themselves to have been at risk of mpox infection. Recruitment will be supported by voluntary sector organizations (George House Trust, NAZ Project London, and Positively UK) to ensure the inclusion of participants living with HIV, queer people of color, and people living outside of London.

SSIs (Nigeria and Thailand)

SSIs will be used to elicit rich information from key stakeholders' experiences, perceptions, and opinions on mpox. We will continue to recruit participants until data saturation has been reached.

In Nigeria, SSIs will explore public perceptions, information sources, attitudes toward transmission-reducing behaviors, public health messaging, social and emotional impacts of mpox with a primary focus on the policies, and interventions for the management of mpox cases. One systematic review found that studies comprising between 9 and 17 interviews were likely to reach saturation [48]; however, this will vary depending on the diversity of participants and experiences included. Up to 20 SSIs will be conducted with (1) staff and policymaking officials from the Nigeria Centre for Disease Control and Prevention and State Ministries of Health; these are agencies involved in infectious disease surveillance and management, including mpox outbreaks; (2) key informants who have been diagnosed with mpox; and (3) health care workers.

In Thailand, SSI participants (aged ≥18 years) will be recruited from the national mpox surveillance database. Interviews will cover attitudes and experiences regarding mpox infection and associated stigmas, as well as policy and public health practice suggestions. Interviews will be discontinued if it is felt that saturation has been reached. Up to 20 SSIs will be conducted with individuals who self-identify as gay, bisexual, and other men who have sex with men, transgender women or *kathoey*, and who have had an mpox diagnosis. SSIs will be conducted among individuals who have experienced mpox either historically (eg, since the start of the Clade IIb outbreak) or more recently (within the past 4 weeks).

Participatory Photography (Thailand)

In Thailand, up to 10 SSI participants aged 18 years and older who provide additional consent will be enrolled in a participatory photography study (adapted from the photovoice approach) [49], following completion of their SSIs. This sample size is consistent with participatory photography in other health fields [50-53]. This study will involve participants providing photos and accompanying brief descriptions reflecting their lived experience, needs, values, and priorities in response to an mpox diagnosis. Researchers will ask participants to respond to twice-weekly prompts through SMS text message or email (depending on participant preference). Participants will be asked to submit up to 5 photographs to a secure university database in response to each prompt; they will also be invited to join 2 researcher-facilitated online workshops to share and discuss their photographs with each other. Captions and notes taken during workshops will be analyzed thematically and triangulated with SSI data.

Data Management Plan

Data to be collected include qualitative text, images, and reports from news and social media; FGD and SSI audio and transcripts (with audio deleted after accuracy checking); rapid assessment procedure (RAP) sheets; and digital photographs (Thailand only). All data will be stored on a Penta Foundation password-protected shared drive in Italy (news media only), a password-protected project-based cloud storage database in Nigeria, and secure university shared drives in Thailand and the United Kingdom. Only VERDIQual study team members will have access to study data in each country; all staff working on this study have completed up-to-date training in research ethics and information governance.

Team-Based Collaboration

This is a complex study involving multiple qualitative methods across 4 countries. We have established structures and processes to support team-based collaboration, facilitating methodological consistency and rigor. All teams have received standardized training in methods and analysis and use standardized protocols and standard operating procedures in the design and conduct of studies at their site, adapting them for their unique contexts.

Data Analysis

Rapid Analysis

We will adopt an approach developed by the Rapid Research Evaluation and Appraisal Lab, which allows for simultaneous collection, synthesis, and reporting of large amounts of qualitative data in a short time using RAP sheets (Multimedia Appendix 2) [54]. The RAP sheet is a 2-column data collection and analysis form that (1) facilitates data standardization across study sites over time, (2) is adaptable to rapid changes in study designs and circumstances, (3) triangulates findings from different data sources, and (4) quickly synthesizes the most important findings [54]. Team members will capture key findings separately from SSIs, FGDs, news and social media analysis, and photovoice, with each row relating to a different SSI, FGD, or photovoice question or theme (in the case of news and social media analysis). There will be separate RAP sheets by method and site, as well as potentially by researcher. RAP sheets will record reflections per method across all participants (ie, no separate RAP sheets for each participant). RAP sheets will not replace in-depth thematic analysis and coding. They will allow us to generate early insights within and across countries, followed by in-depth analysis using an appropriate approach such as reflexive thematic analysis, while allowing us as researchers to provide feedback to key stakeholders at speed, an important consideration during an outbreak.

XSL•FO
RenderX

RAP sheets will also be used to guide collaborative analysis across countries. Through discussion and review of RAP sheets, we will reach consensus on cross-cutting themes as well as areas of difference. This will inform further in-depth coding and analysis of data from all countries. Initial findings from the rapid and thematic analyses of each dataset will be presented to community partners to allow for member checking.

News Media

News media articles will be analyzed via content analysis at each country level using NVivo qualitative data analysis software (version 14; QSR International Pty Ltd). A preliminary analysis of the Italian news media will be carried out on a selection of newspapers. All articles on mpox published in the selected newspapers between May 1, 2022, and December 31, 2022, will be included.

In Nigeria, all articles on mpox published during the peak periods of mpox outbreaks between May 2017 and December 2023 will be randomly selected from selected media houses. These articles will be reviewed by 2 independent reviewers to identify relevant themes and concepts central to the investigation and generate a codebook for analysis. Dual coding will be conducted until consensus is reached, and all coding will be conducted by 2 researchers.

The UK team will conduct a content analysis of selected articles that were published between May 1 and December 31, 2022 (the rise, peak, and decline of mpox in this context). The UK sample will comprise all relevant articles from the selected sources, published on a single day of the week, following a rotating weekday sampling schedule. Similar to Nigeria, a codebook will be developed by 2 researchers, and dual coding will be conducted until consensus is reached. The remainder of the coding will be carried out by an individual researcher (ST).

Themes identified from the analyses in Italy, Nigeria, and the United Kingdom will be reviewed by at least 2 researchers from each team to identify relevant cross-cutting themes of interest for the comparative analysis. Codes for these themes will be shared by the research teams and compared to identify commonalities and points of departure. The cross-country analysis will be led by the Italian team, with regular discussion involving the Nigerian and UK teams.

Social Media Analysis

Across the implementing countries, analysis of relevant mpox data from the selected social media platforms will be analyzed using NVivo (version 14) software to identify common themes. Three members of the Nigerian team (CO-A, EES-A, and GIA) will conduct discourse analysis of videos and posts on Facebook and Instagram using the hashtags “monkeypox,” “monkey pox,” and “mpox” in the search engine. Relevant mpox content posted from or on Nigeria from May 2017 to December 2023 will be extracted, and duplicates will be excluded. The structure and language of this content will be analyzed to explore representations of mpox, framing of responsibility, and perception of and attitudes to mpox vaccination.

At least 2 members from each of the UK and Thai teams will conduct an analysis of TikTok, a social media platform for sharing short videos. This will involve searching for videos with the search terms “monkeypox,” “monkey pox,” and “mpox” (Inline graphic 6. [*Phidard*], Inline graphic 7. [*Phidardling*], and Inline graphic 8. [*Phidardwanon*], respectively). Relevant TikToks will be selected, multimodal data extracted using a standardized form, and data analyzed thematically.

FGDs and SSIs

All FGDs and SSI audio will also be transcribed to facilitate subsequent in-depth thematic analysis using NVivo (version 14) software. Following accuracy checks, we will analyze data using the reflexive thematic analysis approach described by Braun and Clarke [55], which involves (1) data familiarization through repeated reading and note-taking; (2) coding sections of the transcript using a combination of a priori codes developed from literature review and review of RAP sheets, as well as inductive coding; (3) generating and refining themes to develop an overarching interpretation of the data; and (4) reflecting on researcher positionality throughout and how this may shape the analysis and interpretation. For all countries collecting these data, a proportion of FGD and SSI transcripts will be dual-coded, with findings compared and consensus reached for the codebook for each country.

Participatory Photography

This will only be undertaken in Thailand. Interpretation of visual materials and writing tasks will include a focus on the subject matter of separate photographs (eg, how many photographs are intended to document stigma or symptoms) as well as the visual narrative created by each participant (ie, how the visual data tells the story of each participant’s individual experience). Workshops will be cofacilitated by at least 2 researchers, allowing detailed notes to be kept and key themes to be identified between the researchers.

Collaborative Analysis

To harness the full potential of this rich dataset comprising data obtained from 4 countries using a variety of methods, we will use a standardized approach for data integration (Table 3). This will occur at the interpretation stage. Teams in each country will complete RAP sheets in English (the project’s working language). This will facilitate collaboration as described earlier. Further in-depth analyses will be conducted by native speakers in the original language, with summaries of findings and interpretations completed in English. These summaries will be recorded in a multimethod data integration matrix, which will allow visualization of data by objective, method, and site, aiding synthesis across datasets. We acknowledge that some meaning will be lost as a result of language differences, but this pragmatic approach, which will include native speakers at all stages, allows us to integrate data while preserving context-specific meaning. The use of NVivo (including researcher memos) and preservation of RAP sheets and data integration matrices will ensure transparency by creating an audit trail.

XSL•FO
RenderX

Table 3. Example of multimethod data integration matrix.

	Study 1 (eg, FGDs[a]) findings	Study 2 (eg, SSI[b]) findings	Study 3 (eg, photovoice) findings	Study 4 (eg, TikTok) findings	Study 5 (eg, news media) findings	Synthesis (by objective across sites and methods)
State objective 1	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	Record synthesis here
State objective 2	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	Record synthesis here
State objective n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	• Site 1 • Site 2 • Site n	Record synthesis here
Synthesis (by method across sites)	• Record synthesis here	• Record synthesis here	• Record synthesis here	• Record synthesis here	• Record synthesis here	Record synthesis here

[a]FGD: focus group discussion.

[b]SSI: semistructured interview.

Patient and Public Involvement

Each country team has identified community partners who will lead the formation of CABs, comprising representatives of diverse patients, gatekeepers, and community member groups. CAB members will review study documents, contribute to analyses and manuscripts, and inform dissemination and future pandemic preparedness research priorities. In addition to regular meetings with their individual country teams, the CABs from Italy, Nigeria, Thailand, and the United Kingdom will have quarterly international CAB meetings to discuss findings, challenges, and lessons learned.

Ethical Considerations

This study has received ethics approval from the following institutional review boards: in Nigeria, from the National Health Research Ethics Committee (NHREC/01/01/2007); in the United Kingdom, from University College London (6698/005); and in Thailand, from the Mahidol University, Faculty of Tropical Medicine Ethical Committee (MUTM 2025-001-01). Ethical review and approval are not indicated for the news media analysis in Italy. Participant consent is not applicable or required for data sourced from news and social media, as this information is considered public (if accounts are openly accessible). However, for social media analyses, usernames will not be included in reports, and quotes will only be used where they are not considered to be sensitive (ie, talking about traumatic experiences). Ethics around consent and anonymity in social media research is a constantly evolving issue, and special care will be taken [56]. Given the potentially delicate and emotive nature of the research, participants will be informed that they are not obliged to answer questions they find difficult. Written informed or verbal consent will be sought from individuals recruited for FGDs, SSIs, and participatory photography in each country where it applies and will be obtained by members of the research team. In addition to written or verbal consent, for participatory photography in Thailand, emphasis will be placed on visual anonymity, with participants advised to avoid capturing identifiable faces or settings that could compromise privacy or safety. To ensure that the consent process accommodates the needs of study participants (which include adolescents, people living with HIV, and gay, bisexual, and other men who have sex with men), consent procedures will prioritize privacy and autonomy. In Nigeria, written assent and parental permission will be sought for adolescent participants aged 12 to 17 years in accordance with national guidelines [57]. Country teams include experienced qualitative researchers with expertise in working with young people and in HIV and sexual health settings. Participants will be given sufficient time to review the participant's information sheet, ask questions, and discuss their concerns. All qualitative data will be anonymized; identifying information will be removed during transcription and analysis. For participatory photography in Thailand, participants will maintain ownership and control over their photographs, with the sole decision of which images to be shared, discussed, or included in the dissemination materials. All approved images will be reviewed collaboratively to ensure alignment with participants' intent while adhering to ethical standards of respect and confidentiality. All participants were informed that participation is voluntary and that refusal or withdrawal from the study will not affect access to any services, including those provided by community partners. Participant reimbursement followed local ethics guidelines. In Nigeria, participants received approximately US $4 for time and transport; this sum was approved by the National Health Research Ethics Committee. In Thailand, participants received US $15 as reimbursement for their time and transport; this sum was approved by the Mahidol University Faculty of Tropical Medicine Ethical Committee. In the United Kingdom, FGD participants received US $26 shopping vouchers, which was approved by the University College London Research Ethics Committee.

XSL•FO
RenderX

Results

Study Status

At the time of writing, all teams—Italy, Nigeria, the United Kingdom, and Thailand—have obtained the required national and institutional ethics approvals. The study was funded in September 2023, and initial institutional review board approvals were received in each country on varying dates: in Nigeria, January 2024; the United Kingdom, August 2023; and Thailand, January 2025. The development of data collection tools including FGD and SSI topic guides has been completed in collaboration with the CABs. Recruitment has commenced in all countries, with the completion of all qualitative data collection anticipated by October 31, 2025. The VERDIQual umbrella (master) protocol and multiple relevant study tools are available online at protocols.io [58]. At the time of publication, Italy has completed news media data collection and is currently conducting analysis; the Nigerian and Thai teams have completed data collection for all components of their qualitative studies, and analysis is ongoing; the United Kingdom team has finished all FGDs and is analyzing their data. The earliest results manuscripts are expected to be published between May and July 2026.

Results Dissemination Plan

A comprehensive, implementation science–guided dissemination plan [59] to share VERDIQual findings is being codeveloped and will be coimplemented with the international CAB to reach diverse stakeholders and maximize impact on pandemic preparedness in study countries and globally. Dissemination is planned for 3 main audiences:

- Funder: VERDIQual reports to the European Commission.
- Scientific: We will reach the scientific community through abstracts and invited presentations at major national, regional, and international conferences and publications in peer-reviewed journals.
- Nonscientific or community: The nonscientific audience has been segmented into patients or patient advocacy groups; the general public or communities; providers, including health care workers and health professional organizations; and policymakers such as government officials, Ministries of Health, and other relevant policymaking agencies.

The mode of dissemination for each nonscientific or community audience will be determined by each country team for their local setting. A range of options include social media messaging for patients; community town hall meetings for study communities; results communications through flyers at health facilities; and technical reports, policy briefs, and presentations at policymaking meetings for policymakers.

Discussion

Overview

The mpox outbreak continues to be a pressing public health concern. The experience of the 2019 SARS-CoV-2 pandemic allowed lessons to be learned and insights into missed opportunities for application to the 2022 mpox pandemic. Beyond the public health experience of past pandemics, rigorous data on the experiences of individuals affected by mpox will be critical to better inform public health interventions and messaging for mpox and future pandemics.

Expectations for Principal Findings

Given our inclusion of 3 global regions, mpox endemic and nonendemic countries, general and marginalized populations, and high- and low-resource settings, we expect our principal findings to be applicable to a range of settings. Furthermore, our application of intersectionality theory will also facilitate considerations for intersecting identities and characteristics in equity-centered pandemic responses. By this, we mean intentionality in pandemic preparedness, inclusive of people with 1 or more stigmatized and vulnerable characteristics or identities with respect to infection or access to health interventions, such as geographical location, gender, pregnancy, young age, sexual orientation, race, and class or economic status. Our approach to dissemination and potential translation of our findings into policy are described in our dissemination plan mentioned earlier.

Comparison to Prior Work

Several qualitative studies have documented the experiences of people with mpox, including people with HIV coinfection, during the 2022 pandemic [22-25,60-64]. However, almost all of these studies were conducted among gay, bisexual, and other men who have sex with men in nonendemic, high-income countries such as France, the United States, the United Kingdom, Australia, and China. Notable studies from Africa include 2 studies among men who have sex with men and the general population in Nigeria [65,66]. At the time of writing, there is a paucity of studies on the experiences of different patient populations with mpox in endemic African countries or across both endemic and nonendemic settings.

Anticipated Challenges and Mitigation Strategies

For a multicountry, multimethod study such as VERDIQual, implementation challenges may occur due to variations in public health infrastructure, case burden, and the range of populations disproportionately affected by mpox. Language differences, legal and ethical regulations, cultural norms, and stigma, particularly around age, gender identity, and sexual orientation, may impact data collection. To mitigate these, we have adopted a country-led approach to the collaboration, allowing for the selection of locally feasible qualitative research methods, local adaptation of study tools, and application of locally appropriate recruitment strategies using CABs while providing standardized training and requiring standardized reporting for all country teams to ensure methodological consistency across sites.

XSL•FO
RenderX

Acknowledgments

The authors express sincere thanks to all study participants for their time and effort in sharing their opinions and experiences for this study. The authors appreciate the valuable contributions of their community partners: the European AIDS Treatment Group (Portugal), NAZ Project London (United Kingdom), Afrihealth Optonet Association (Nigeria), and their community advisory board members: Aaron Choudhry, Abdullah Zafar, Asiton-a Ibama, Chiara Longhi, Esther Onimoto, Harun Tulunay, Jaime Sylla, Marco Stizoli, Matthew Hamer, Maxwell Ati, Oscar Patton Lyons, Parminder Sekhon, Rita Ifeyinwa Okonkwo, Ronnapoom Samakkeekarom, Sanjay Kumar, Simon Collins, and Uzodinma Adirieje. The authors also acknowledge the contributions of Alexandra David, Anya Maclaren, Grace Adebayo, Precious Omonkhegbele, Rati Taesombat, Sailesh Kandasamy, Siriya Lekkang, Thasaporn Damri, and Yomna Gharib. Finally, the authors are grateful to the entire SARS-CoV-2 (and Mpox Variants) Evaluation in Pregnancy and Paediatrics Cohorts consortium for their collegiate support.

Community advisory board for facilitating access to participants: Asiton-a Ibama (supporting), Esther Onimoto (supporting), Maxwell Ati (supporting), Rita Ifeyinwa Okonkwo (supporting), Aaron Choudhry (supporting), Abdullah Zafar (supporting), Harun Tulunay (supporting), Jaime Sylla (supporting), Marco Stizoli (supporting), Matthew Hamer (supporting), Oscar Patton Lyons (supporting), Simon Collins (supporting), Sanjay Kumar (supporting).

Acknowledgments for providing study materials: Rati Taesombat (supporting), Sailesh Kandasamy (supporting), Siriya Lekkang (supporting), Thasaporn Damri (supporting), Alexandra David (supporting), Anya Maclaren (supporting), Yomna Gharib (supporting), Grace Adebayo (supporting), Precious Omonkhegbele (supporting).

Funding

This work was supported by the European Commission (grant 101045989). TCW is supported by the Wellcome Trust (226345/Z/22/Z).

Authors' Contributions

Conceptualization: ST (lead) NAS-A (supporting)
Data curation: CO-A (lead), MLP (equal), EJN (equal), DB (equal), TEG (equal)
Funding acquisition: ST (lead), NAS-A (equal)
Investigation: CO-A (lead), MLP (equal), EJN (equal), TEG (equal), DB (equal), EES-A (supporting), GIA (supporting), CA (supporting), AAA (supporting), KY (supporting), TCW (supporting), WW (supporting), NS (supporting), TM (supporting), SD (supporting), WG (supporting), Grace Adebayo (supporting), Precious Omonkhegbele (supporting)
Methodology: DB (lead), EJN (equal), MLP (equal), ST (equal)
Project administration: NAS-A (lead), ST (equal), CO-A (supporting), EJN (supporting), MLP (supporting), TEG (supporting), DB (supporting)
Resources: ST (lead), NAS-A (equal), TEG (equal), DB (equal), UA (equal), Ronnapoom Samakkeekarom (equal), Chiara Longhi (equal), Parminder Sekhon (equal), CO-A (supporting), MLP (supporting), EJN (supporting), TEG (supporting), DB (supporting), EES-A (supporting), GIA (supporting), CA (supporting), AAA (supporting), KY (supporting), TCW (supporting), WW (supporting), NS (supporting), TM (supporting), SD (supporting), WG (supporting)
Supervision: NAS-A (lead), ST (equal)
Writing—original draft: CO-A (lead), NAS-A (equal)
Writing—review and editing: MLP (lead), EJN (equal), DB (equal), TEG (equal), EES-A (equal), GIA (equal), CA (equal), AAA (equal), KY (equal), TCW (equal), WW (supporting), NS (supporting), TM (supporting), SD (supporting), WG (supporting), UA (supporting), ST (equal), NAS-A (equal)

Conflicts of Interest

None declared.

Multimedia Appendix 1
COREQ checklist.
[PDF File (Adobe PDF File), 217 KB - resprot_v15i1e77321_app1.pdf]

Multimedia Appendix 2
Example of rapid assessment procedure sheet to be used for rapid qualitative research data collection.
[PDF File (Adobe PDF File), 224 KB - resprot_v15i1e77321_app2.pdf]

References

XSL•FO
RenderX

1. Sabino EC, Buss LF, Carvalho MPS, Prete CA, Crispim MAE, Fraiji NA, et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. Lancet 2021;397(10273):452-455 [FREE Full text] [doi: 10.1016/S0140-6736(21)00183-5] [Medline: 33515491]
2. Janik E, Niemcewicz M, Podogrocki M, Majsterek I, Bijak M. The emerging concern and interest SARS-CoV-2 variants. Pathogens 2021;10(6):633 [FREE Full text] [doi: 10.3390/pathogens10060633] [Medline: 34064143]
3. Koyama T, Platt D, Parida L. Variant analysis of SARS-CoV-2 genomes. Bull World Health Organ 2020;98(7):495-504 [FREE Full text] [doi: 10.2471/BLT.20.253591] [Medline: 32742035]
4. Krubiner CB, Faden RR, Karron RA, Little MO, Lyerly AD, Abramson JS, et al. Pregnant women & vaccines against emerging epidemic threats: ethics guidance for preparedness, research, and response. Vaccine 2021;39(1):85-120 [FREE Full text] [doi: 10.1016/j.vaccine.2019.01.011] [Medline: 31060949]
5. Zaccardi F, Tan PS, Shah BR, Everett K, Clift AK, Patone M, et al. Ethnic disparities in COVID-19 outcomes: a multinational cohort study of 20 million individuals from England and Canada. BMC Public Health 2023;23(1):399 [FREE Full text] [doi: 10.1186/s12889-023-15223-8] [Medline: 36849983]
6. Thornhill JP, Barkati S, Walmsley S, Rockstroh J, Antinori A, Harrison LB, et al. Monkeypox virus infection in humans across 16 countries - April-June 2022. N Engl J Med 2022 Aug 25;387(8):679-691. [doi: 10.1056/NEJMoa2207323] [Medline: 35866746]
7. Rimoin AW, Mulembakani PM, Johnston SC, Lloyd Smith JO, Kisalu NK, Kinkela TL, et al. Major increase in human monkeypox incidence 30 years after smallpox vaccination campaigns cease in the Democratic Republic of Congo. Proc Natl Acad Sci U S A 2010;107(37):16262-16267 [FREE Full text] [doi: 10.1073/pnas.1005769107] [Medline: 20805472]
8. 2022-23 mpox (monkeypox) outbreak: global trends. World Health Organization. 2025. URL: https://worldhealthorg.shinyapps.io/mpx_global/ [accessed 2025-10-20]
9. 2022-24 mpox (monkeypox) outbreak: global trends. World Health Organization. 2025. URL: https://worldhealthorg.shinyapps.io/mpx_global/ [accessed 2025-04-29]
10. Vakaniaki E, Kacita C, Kinganda-Lusamaki E, O'Toole Á, Wawina-Bokalanga T, Mukadi-Bamuleka D, et al. Sustained human outbreak of a new MPXV clade I lineage in eastern Democratic Republic of the Congo. Nat Med 2024;30(10):2791-2795 [FREE Full text] [doi: 10.1038/s41591-024-03130-3] [Medline: 38871006]
11. Gonsalves GS, Mayer K, Beyrer C. Déjà vu all over again? Emergent monkeypox, delayed responses, and stigmatized populations. J Urban Health 2022;99(4):603-606 [FREE Full text] [doi: 10.1007/s11524-022-00671-1] [Medline: 35916973]
12. Kartono R, Salahudin, Sihidi IT. Covid-19 stigmatization: a systematic literature review. J Public Health Res 2022;11(3):22799036221115780 [FREE Full text] [doi: 10.1177/22799036221115780] [Medline: 36090024]
13. Razai MS, Mansour R, Ravindran P, Freeman S, Mason-Apps C, Morris J, et al. Facilitators and barriers to vaccination uptake in pregnancy: a qualitative systematic review. PLoS One 2024;19(4):e0298407 [FREE Full text] [doi: 10.1371/journal.pone.0298407] [Medline: 38640190]
14. May T, Towler L, Smith LE, Horwood J, Denford S, Rubin GJ, et al. Mpox knowledge, behaviours and barriers to public health measures among gay, bisexual and other men who have sex with men in the UK: a qualitative study to inform public health guidance and messaging. BMC Public Health 2023;23(1):2265 [FREE Full text] [doi: 10.1186/s12889-023-17196-0] [Medline: 37978506]
15. Thomas C, Faghy MA, Owen R, Yates J, Ferraro F, Bewick T, et al. Lived experience of patients with Long COVID: a qualitative study in the UK. BMJ Open 2023;13(4):e068481 [FREE Full text] [doi: 10.1136/bmjopen-2022-068481] [Medline: 37185640]
16. Hossain A, Monem M, Rahman M, Raza R. Mpox (monkeypox): a comprehensive updated of current epidemic evidence. Sci One Health 2025;4:100100 [FREE Full text] [doi: 10.1016/j.soh.2024.100100] [Medline: 39872962]
17. Ogoina D, Iroezindu M, James H, Oladokun R, Yinka-Ogunleye A, Wakama P, et al. Clinical course and outcome of human monkeypox in Nigeria. Clin Infect Dis 2020;71(8):e210-e214. [doi: 10.1093/cid/ciaa143] [Medline: 32052029]
18. Foláyan MO. Mpox, stigma and the public health emergency of continental security declaration: addressing public health challenges in Africa. J Public Health Afr 2024;15(1):757. [doi: 10.4102/jphia.v15i1.757] [Medline: 39229344]
19. Kangbai J, Sesay U, Vickos U, Kagbanda F, Fallah M, Osborne A. Mpox in Africa: what we know and what is still lacking. PLoS Negl Trop Dis 2025;19(6):e0013148 [FREE Full text] [doi: 10.1371/journal.pntd.0013148] [Medline: 40504780]
20. Adigun O, Okesanya O, Ahmed M, Ukoaka B, Lucero-Prisno D, Onyeaghala E, et al. Syndemic challenges: addressing the resurgence of mpox amidst concurrent outbreaks in the DRC. Transbound Emerg Dis 2024;2024:1962224. [doi: 10.1155/tbed/1962224] [Medline: 40303062]
21. Edward M, shadrack O, Engmann S, Jubilate A. Global mpox transmission and stigma: addressing barriers in affected communities. Discov Public Health 2025;22(1):1-4 [FREE Full text] [doi: 10.1186/s12982-025-00470-4]
22. Witzel TC, Ghobrial A, Palich R, Charles H, Rodger AJ, Sabin C, et al. Experiences of mpox illness and case management among cis and trans gay, bisexual and other men who have sex with men in England: a qualitative study. EClinicalMedicine 2024;70:102522 [FREE Full text] [doi: 10.1016/j.eclinm.2024.102522] [Medline: 38685928]
23. Zhang W, Qi X, Yang L, Meng X, Xu G, Luo S, et al. Mpox patients' experience from infection to treatment and implications for prevention and control: a multicenter qualitative study in China. J Med Virol 2024;96(1):e29338. [doi: 10.1002/jmv.29338] [Medline: 38163280]

XSL•FO
RenderX

24. Smith A, Storer D, Lancaster K, Haire B, Newman C, Paparini S, et al. Mpox illness narratives: stigmatising care and recovery during and after an emergency outbreak. Qual Health Res 2024;34(12):1161-1174 [FREE Full text] [doi: 10.1177/10497323241234482] [Medline: 38462475]
25. Storer D, Holt M, Paparini S, Haire B, Cornelisse VJ, MacGibbon J, et al. Informed, but uncertain: managing transmission risk and isolation in the 2022 mpox outbreak among gay and bisexual men in Australia. Cult Health Sex 2025;27(1):16-31 [FREE Full text] [doi: 10.1080/13691058.2024.2346540] [Medline: 38756104]
26. Garcia-Iglesias J, May T, Pickersgill M, Williams J, Nagington M, Buijsen S, et al. Social media as a public health tool during the UK mpox outbreak: a qualitative study of stakeholders' experiences. BMJ Public Health 2023;1(1):e000407. [doi: 10.1136/bmjph-2023-000407] [Medline: 40017890]
27. VERDI project. VERDI Consortium. URL: https://verdiproject.org/ [accessed 2024-12-23]
28. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
29. Crenshaw K. Demarginalizing the intersection of race and sex: a Black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. Femin Legal Theor 2013:23-51 [FREE Full text] [doi: 10.4324/9780429500480-5]
30. Sexually transmitted infections: update of data from the two sentinel surveillance systems active in Italy as of December 31, 2023. Report of d'Instituto Superiore de Sanita (Research Institute of the Italian Ministry of Health). 2025 Aug. URL: https://www.iss.it/documents/20126/6683812/Vol.+38%2C+n.+7-8%2C+2025+%281%29.pdf/25bc01e1-f29c-2769-991c-2555977c2e47?version=1.0&t=1753688350186&download=true [accessed 2025-11-23]
31. León-Figueroa DA, Barboza JJ, Siddiq A, Sah R, Valladares-Garrido MJ, Rodriguez-Morales AJ. Knowledge and attitude towards mpox: systematic review and meta-analysis. PLoS One 2024;19(8):e0308478 [FREE Full text] [doi: 10.1371/journal.pone.0308478] [Medline: 39121048]
32. Olojede O, Bakare D, Sogbesan A, Salako J, Bakare A. Assessment of the level of awareness, knowledge, and risk perception of community members about mpox infection in Nigeria. Ann Ib Postgrad Med 2024;22(1):76. [doi: 10.38029/babcockuniv.med.j..v7i2.402]
33. Mitjà O, Ogoina D, Titanji BK, Galvan C, Muyembe JJ, Marks M, et al. Monkeypox. Lancet 2023;401(10370):60-74 [FREE Full text] [doi: 10.1016/S0140-6736(22)02075-X] [Medline: 36403582]
34. Zumla A, Rosenthal P, Sam-Agudu NA, Ogoina D, Mbala-Kingebeni P, Ntoumi F, et al. The 2024 Public Health Emergency of International Concern: a global failure to control mpox. Am J Trop Med Hyg 2024 Oct 15;112(1):17-20 [FREE Full text] [doi: 10.4269/ajtmh.24-0606] [Medline: 39406210]
35. First-ever delivery of mpox vaccines in Africa outside of clinical trials arrives in Nigeria. World Health Organization Regional Office for Africa. 2024. URL: https://www.afro.who.int/countries/nigeria/news/first-ever-delivery-mpox-vaccines-africa-outside-clinical-trials-arrives-nigeria [accessed 2025-10-20]
36. Nigeria scales up mpox prevention with targeted vaccination in high-burden states. World Health Organization. 2025. URL: https://www.afro.who.int/countries/nigeria/news/nigeria-scales-mpox-prevention-targeted-vaccination-high-burden-states [accessed 2025-10-20]
37. Monkeypox (mpox) situation report. Department of Disease, Thailand Ministry of Public Health. 2024. URL: https://ddc.moph.go.th/monkeypox/dashboard.php [accessed 2025-11-23]
38. Monkeypox vaccination services. AIDS and Infectious Diseases Research Center, Thai Red Cross Society. 2024. URL: https://th.trcarc.org/vaccine-mpox14-4-23/ [accessed 2025-11-23]
39. EC_RL_014_S2_BK Average wage classified by industry (ISIC Rev.4) in Bangkok 1/. Bank of Thailand. 2024. URL: https://app.bot.or.th/BTWS_STAT/statistics/ReportPage.aspx?reportID=738&language=eng [accessed 2025-10-22]
40. Mpox outbreak: epidemiological overview, June 2022 to December 2024. UK Government Web Archive. 2024 Dec. URL: https://webarchive.nationalarchives.gov.uk/ukgwa/20250129194727/https://www.gov.uk/government/publications/monkeypox-outbreak-epidemiological-overview [accessed 2025-11-23]
41. Investigation into monkeypox outbreak in England: technical briefing 1. GOV.UK. URL: https://www.gov.uk/government/publications/monkeypox-outbreak-technical-briefings/investigation-into-monkeypox-outbreak-in-england-technical-briefing-1 [accessed 2024-09-24]
42. Monkeypox cases confirmed in England—latest updates. UK Health Security Agency. 2022. URL: https://www.gov.uk/government/news/monkeypox-cases-confirmed-in-england-latest-updates [accessed 2024-05-17]
43. Mpox clade Ib and clade IIb outbreak: epidemiological overview. UK Health Security Agency. 2025. URL: https://www.gov.uk/government/publications/monkeypox-outbreak-epidemiological-overview [accessed 2025-04-29]
44. Ogaz D, Enayat Q, Brown JR, Phillips D, Wilkie R, Jayes D, et al. Mpox diagnosis, behavioral risk modification, and vaccination uptake among gay, bisexual, and other men who have sex with men, United Kingdom, 2022. Emerg Infect Dis 2024 May;30(5):916-925. [doi: 10.3201/eid3005.230676] [Medline: 38573160]
45. Khan NA, Abedin S. Focus Group Discussion. In: Islam MR, Khan NA, Baikady R, editors. Principles of Social Research Methodology. Singapore: Springer; 2022.
46. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. Field Methods 2016;29(1):3-22. [doi: 10.1177/1525822x16639015]

XSL•FO
RenderX

47. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. Int J Qual Methods 2019;18:1609406919874596. [doi: 10.1177/1609406919874596]
48. Hennink M, Kaiser B. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022;292:114523 [FREE Full text] [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
49. Sutton-Brown CA. Photovoice: a methodological guide. Photogr Cult 2015;7(2):169-185. [doi: 10.2752/175145214x13999922103165]
50. Shumba T, Moodley I. Part 1: A review of using photovoice as a disability research method: implications for eliciting the experiences of persons with disabilities on the Community Based Rehabilitation programme in Namibia. Afr J Disabil 2018;7:418 [FREE Full text] [doi: 10.4102/ajod.v7i0.418] [Medline: 30473999]
51. Kamila A, Widyawati W, Hasanbasri M, Hakimi M. Capturing the HIV-related social exclusion practices experienced by key populations through photovoice: an interpretative phenomenological study. Reprod Health 2024;21(1):107 [FREE Full text] [doi: 10.1186/s12978-024-01832-y] [Medline: 39004733]
52. Owens TK, Mizock L, Ormerod AJ, Nelson A, St Amand C, Paces-Wiles D, et al. "Invisible in the Most Tragic of Ways": exploring internalized transphobia and coping through photovoice. Health Promot Pract 2023;24(4):682-693. [doi: 10.1177/15248399221114340] [Medline: 35923148]
53. Kimera E, Vindevogel S. Photovoicing empowerment and social change for youth living with HIV/AIDS in Uganda. Qual Health Res 2022;32(12):1907-1914. [doi: 10.1177/10497323221123022] [Medline: 35998362]
54. Vindrola-Padros C, Chisnall G, Polanco N, San JN. Iterative cycles in qualitative research: introducing the RREAL sheet as an innovative process. Soc Sci Humanit Open 2022 Jul 14:1-17 http://dx.doi.org/10.2139/ssrn.4162797 [FREE Full text] [doi: 10.31219/osf.io/9dp2w]
55. Braun V, Clarke V. Thematic analysis. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology, Vol. 2. Washington, DC: American Psychological Association; 2012:57-71.
56. Ford E, Shepherd S, Jones K, Hassan L. Toward an ethical framework for the text mining of social media for health research: a systematic review. Front Digit Health 2020;2:592237 [FREE Full text] [doi: 10.3389/fdgth.2020.592237] [Medline: 34713062]
57. Policy statement regarding enrollment of children in research in Nigeria (PS2. 1016). In: National Health Research Ethics Committee (NHREC) of Nigeria. Abuja, Nigeria: Federal Ministry of Health Nigeria; 2014.
58. An umbrella protocol for qualitative research in infectious disease outbreaks. Protocols.io. 2025. URL: https://www.protocols.io/view/an-umbrella-protocol-for-qualitative-research-in-i-36wgqpee3vk5/v1 [accessed 2025-11-26]
59. Brownson R, Eyler A, Harris J, Moore J, Tabak R. Getting the word out: new approaches for disseminating public health science. J Public Health Manag Pract 2018;24(2):102-111 [FREE Full text] [doi: 10.1097/PHH.0000000000000673] [Medline: 28885319]
60. Dos Santos M, Salcedo M, Mora M, Siguier M, Velter A, Leclercq V, et al. Negotiating access to healthcare and experience of stigma among cisgender gay men diagnosed with mpox during the 2022 epidemic in France: a qualitative study. AIDS Behav 2025;29(9):2831-2840. [doi: 10.1007/s10461-025-04739-2] [Medline: 40289037]
61. Yellin H, Bornstein S, Balachandran M, Siegel M, Magnus M. Psychosocial and behavioral impacts of the mpox outbreak among people with and without HIV in the United States. AIDS Behav 2025;29(4):1294-1304. [doi: 10.1007/s10461-024-04603-9] [Medline: 39739285]
62. Brick M, Rudolph D, Dashler G, Varnauskas A, Qiao L, Hamill M, et al. When public health emergencies hit the front line: a qualitative study of the patient experience in the emergency department and outpatient infectious disease clinic during the 2022 Mpox clade IIb outbreak. BMC Infect Dis 2025;25(1):754 [FREE Full text] [doi: 10.1186/s12879-025-11124-w] [Medline: 40419995]
63. Hughes D, Ai J, Vazirnia P, McLeish T, Krajco C, Moraga R, et al. A qualitative study of Chicago gay men and the Mpox outbreak of 2022 in the context of HIV/AIDS, PrEP, and COVID-19. BMC Infect Dis 2024;24(1):1174 [FREE Full text] [doi: 10.1186/s12879-024-09491-x] [Medline: 39425023]
64. Takenaka B, Kirklewski S, Griffith F, Gibbs J, Lauckner C, Nicholson E, et al. "It's another gay disease": an intersectional qualitative approach contextualizing the lived experiences of young gay, bisexual, and other sexual minoritized men in the United States during the mpox outbreak. BMC Public Health 2024;24(1):1574 [FREE Full text] [doi: 10.1186/s12889-024-19062-z] [Medline: 38862933]
65. Kunnuji M, Schmidt-Sane M, Adegoke O, Abbas S, Shoyemi E, Lawanson AO, et al. Mpox and the men who have sex with men (MSM) community in Nigeria: exploratory insights from MSM and persons providing healthcare services to them. Glob Public Health 2025;20(1):2433725 [FREE Full text] [doi: 10.1080/17441692.2024.2433725] [Medline: 39833709]
66. Adegoke O, Schmidt-Sane M, Kunnuji M, Abbas S, Lawanson A, Jegede A, et al. Diagnosis, treatment, and management of Mpox in urban informal settlements in southwestern Nigeria: an ethnographic approach. BMC Public Health 2025;25(1):115 [FREE Full text] [doi: 10.1186/s12889-024-21267-1] [Medline: 39789521]

XSL•FO
RenderX

Abbreviations

CAB: community advisory board
COREQ: Consolidated Criteria for Reporting Qualitative Research
FGD: focus group discussion
RAP: rapid assessment procedure
SSI: semistructured interview
VERDI: SARS-CoV-2 (and Mpox) Variants Evaluation in Pregnancy and Paediatrics Cohorts
VERDIQual: SARS-CoV-2 (and Mpox) Variants Evaluation in Pregnancy and Paediatrics Cohorts Qualitative Study
WHO: World Health Organization

Edited by J Sarvestan; submitted 12.May.2025; peer-reviewed by C Obianyo, M Gasmi , A Girma; comments to author 05.Jun.2025; revised version received 25.Oct.2025; accepted 12.Nov.2025; published 15.Jan.2026.

Please cite as:
Osa-Afiana C, Le Prevost M, Nicholls EJ, Bilardi D, Guadamuz TE, Soje-Amadosi EE, Adebisi GI, Adirieje C, Adetunji AA, Yusufu K, Witzel TC, Waratworawan W, Samoh N, May T, Denford S, Gomes W, Adirieje U, Tariq S, Sam-Agudu NA
Rapid Qualitative Approaches in Pandemic Research: Protocol for an Exploratory Qualitative Multimethod Study (VERDIQual) on Mpox in Italy, Nigeria, Thailand, and the United Kingdom
JMIR Res Protoc 2026;15:e77321
URL: https://www.researchprotocols.org/2026/1/e77321
doi:10.2196/77321
PMID:41223310

©Chinye Osa-Afiana, Marthe Le Prevost, Emily Jay Nicholls, Davide Bilardi, Thomas E Guadamuz, Esekwe E Soje-Amadosi, Grace I Adebisi, Chibueze Adirieje, Adeyosola A Adetunji, Karima Yusufu, T Charles Witzel, Worawalan Waratworawan, Nattharat Samoh, Tom May, Sarah Denford, Willian Gomes, Uzodinma Adirieje, Shema Tariq, Nadia Adjoa Sam-Agudu. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Nutrition Education for Emerging Adults: Protocol for Program Evaluation

Stephanie Rogus[1], PhD, RDN; Katelin Alfaro Hudak[1], PhD, MID; Lilian O Ademu[1], PhD; Sumathi Venkatesh[2], PhD; Michael Laguros[2], RD; Elizabeth F Racine[1], DrPH, RD

[1]Texas A&M AgriLife Research Center, El Paso, TX, United States

[2]Texas A&M AgriLife Extension Service, College Station, TX, United States

Corresponding Author:
Stephanie Rogus, PhD, RDN
Texas A&M AgriLife Research Center
1380 A and M Circle
El Paso, TX, 79927
United States
Phone: 1 9157013088
Email: stephanie.rogus@ag.tamu.edu

Abstract

Background: Emerging adults (ie, those aged 18 to 25 years) in the United States exhibit the poorest diet quality among all adult age groups, contributing to adverse health and academic outcomes. Existing nutrition education programs often overlook this population, particularly those without children.

Objective: This study aims to present the development of a nutrition education curriculum for emerging adults, including its process and outcome evaluation strategies and measures.

Methods: The Fuel to Thrive program was adapted from an established state program for adults. The content was informed by reviewing the literature, conducting focus groups with health educators, and holding regular meetings with a curriculum committee. The final program consists of five 1-hour lessons incorporating nutrition education, recipe demonstrations, and physical activity. Process evaluation will involve focus groups with participants and interviews with educators to assess relevance and feasibility. Outcome evaluation will include surveys administered before and after the program to assess dietary behaviors, physical activity, food safety, and food resource management.

Results: The program is currently being piloted at a Texas university. Additional implementations are planned for fall 2025 and spring 2026, with the inclusion of the Short Healthy Eating Index Survey to better assess diet quality changes.

Conclusions: The Fuel to Thrive program addresses a critical gap in nutrition education for emerging adults by offering a tailored curriculum. Future iterations will refine the program based on participant and educator feedback and expand evaluation efforts. The long-term goal is national dissemination to improve dietary behaviors and health outcomes among emerging adults.

International Registered Report Identifier (IRRID): DERR1-10.2196/81647

(JMIR Res Protoc 2026;15:e81647) doi:10.2196/81647

KEYWORDS

emerging adults; nutrition education; curriculum development; program evaluation; extension

Introduction

Emerging adulthood (ages 18-25 years) is a pivotal developmental stage characterized by numerous life transitions, including entry into higher education or the workforce; increased financial and personal autonomy; and, for some, the assumption of parenting responsibilities [1,2]. This period also offers increased discretionary time and independence in decision-making, providing a critical window for establishing long-term health behaviors [1]. Despite its importance, this age group exhibits the poorest diet quality among all US adult populations, with average dietary patterns falling well below recommended nutritional guidelines [3].

Suboptimal diet quality among emerging adults has been associated with lower academic performance, poorer mental health outcomes, and low socioeconomic status [4-7]. These findings underscore the need to provide nutrition education and

XSL•FO
RenderX

implement age-appropriate, evidence-based nutrition interventions for this population. Nutrition education programs that teach food preparation skills, food budgeting, and nutrition knowledge can potentially help emerging adults establish lifelong healthy eating behaviors [8-10]. Such programs offer practical, hands-on learning that builds confidence in preparing balanced meals, selecting nutritious options on a limited budget, and understanding the long-term health implications of dietary choices [11]. Beyond imparting knowledge, these interventions foster critical thinking about food systems, cooking methods, and the importance of meal planning, helping to combat common barriers such as time constraints, cost, and lack of culinary skills.

While existing state and federal programs focus on various aspects of diet-related behavior change, few are tailored specifically to emerging adults. For example, the Expanded Food and Nutrition Education Program (EFNEP), a US Department of Agriculture-funded Extension (outreach) program, serves families by delivering community-based nutrition education [12]. The program has historically focused on parenting populations, thereby overlooking the specific needs of emerging adults without children. The US Department of Agriculture only recently granted state EFNEP programs the ability to serve emerging adults, including those without children. In 2023, approximately 40 emerging adults without children enrolled in Texas EFNEP. A preliminary analysis of their data from the Web-Based Nutrition Education Evaluation and Reporting System [13] indicated an improvement in their Healthy Eating Index scores from a baseline of 43 to 51 after the intervention. While this suggests positive dietary change, the score still falls within the lowest category according to the National Cancer Institute's recommended grading scale, reflecting poor overall diet quality [14]. Therefore, nutrition education programs that target specific needs and address persistent gaps in diet quality are warranted.

This paper presents the rationale, design, and proposed evaluation framework for a nutrition education intervention adapted from an existing EFNEP curriculum and tailored for emerging adults, that is, those aged 18 to 25 years. The program was developed as a 5-week, theory-informed intervention designed explicitly for this age group, regardless of educational or parental status. It retains EFNEP's core emphasis on practical skill building in 4 domains—diet quality and physical activity, food resource management, food safety, and food security—while tailoring content to align with the life stage and challenges of emerging adults. The program aims to enhance nutrition knowledge, promote sustainable behavior change, and ultimately improve overall dietary health in this vulnerable age group.

Methods

Curriculum Development

Given the gap in nutrition education programs focused on emerging adults, the Fuel to Thrive (F2T) program was developed for an emerging adult audience, with initial testing to be conducted in a university setting. It was also designed for use within EFNEP programming. University health educators recognized that students would benefit from nutrition education that met their unique needs and noted that many students who participate in their programs receive federal financial assistance (eg, Pell Grants). During program development, the curriculum development team considered several pieces of information on emerging adults' food choice and dietary habits, including (1) low diet quality and high prevalence of food insecurity; (2) newfound autonomy and key influences on food choice such as taste, convenience, and cost; (3) limited knowledge of basic nutrition, cooking, and food safety; and (4) low self-efficacy in food preparation. An Extension program specialist, a registered dietitian with experience in developing EFNEP programs, led the curriculum development, with input from other nutrition and health educators from the university and Extension, as described in the next paragraphs.

F2T is a modified version of Texas EFNEP's "Healthy Bites, Healthy Moves" (HBHM) curriculum. The HBHM program is a series of 8 nutrition education lessons, informed by adult learning theory, covering healthy eating (using the MyPlate guide), family meal planning, grocery budgeting, food safety, and the importance of physical activity [15]. HBHM was created by Texas EFNEP in 2018, implemented in 2019, and revised in 2024. It was written to reflect the Dietary Guidelines for Americans and is currently offered only in Texas.

Before modifying the HBHM curriculum, a literature review on factors influencing emerging adult dietary behaviors was conducted, and 3 focus groups were held with EFNEP agents and educators familiar with educating emerging adults. Eleven staff (7 agents and 4 educators) from 8 counties across the state participated in the focus groups. Participants were asked to attend 1 of 3 informal virtual focus groups that lasted between 1 and 1.5 hours. Focus group participants reviewed the concepts and activities in the HBHM curriculum and discussed ways to tailor the program to meet the needs of emerging adults. These discussions helped identify topics important to emerging adults, areas in which additional education is needed, and strategies for relaying information in ways that emerging adults would retain and apply to their lives.

A curriculum review committee was formed in August 2024 and met weekly until mid-January 2025 to review and modify the HBHM curriculum for emerging adults. The committee included the curriculum developer (program specialist), 3 EFNEP staff members (the EFNEP coordinator and 2 EFNEP agents who had participated in the focus groups), and 2 staff members from the university's health center. The curriculum was modified by reducing the number of lessons from 8 to 5. This was done, in part, because previous research suggests that short, focused curricula may improve program engagement, retention, and accessibility given the competing demands of emerging adults, including work, school, and social commitments [16-19]. The committee carefully considered each lesson and its content for relevance and redundancy to determine what could be removed without omitting important information. Examples of lessons that were removed include those focused on feeding children and the final summary lesson. Although EFNEP guidelines suggest that an 8-lesson program is standard [20], the number of lessons varies across programs [21-23], and EFNEP staff on the committee agreed that a 5-lesson series

XSL•FO
RenderX

would be appropriate to deliver the core material while keeping emerging adults engaged throughout. Additionally, language and activities were adapted to focus on emerging adults rather than families, and campus-specific resources, such as nutrition facts information from campus dining options, were included. Other materials, such as Dietary Guidelines for Americans statistics and MyPlate resources, were updated, and information on protein and fueling performance was added. These modifications were supported by behavioral theory and nutrition education research, indicating that interventions tailored to developmental stage and lived experience are more effective in improving diet-related behavior [24,25].

The modified program, titled Fuel to Thrive, consists of 5 lessons, each approximately 1 hour in duration. This duration was deemed appropriate by the committee to adequately cover the material. Each lesson includes a PowerPoint (version 2511; Microsoft Corp) presentation, a recipe demonstration with tasting, and a physical activity component. The curriculum toolkit includes suggested enhancement items to help reinforce participant learning in each lesson (Table 1). Low-cost enhancement items that encourage program engagement are allowable expenses when offered as part of an EFNEP class [26].

Table 1. Fuel to Thrive program details.

Lesson number	Title	Lesson content	Enhancement items
Lesson 1	MyPlate, my goals	• Goal setting • Introduction to MyPlate, the Dietary Guidelines for Americans, and mobile nutrition apps • Benefits of physical activity	Measuring cups
Lesson 2	Slice to serve	• Fruit and vegetable groups (sources, nutrients, and benefits) • Knife safety • Knife skills, including slicing, dicing, chopping, and julienne (hands-on)	Cutting boards
Lesson 3	Read it to eat it	• Dairy group (sources, nutrients, and benefits) • How to effectively use a nutrition facts label • Strategies to identify and decrease intake of saturated fats, sodium, and added sugars	Lacrosse balls
Lesson 4	Save money and eat better	• Grains group (sources, nutrients, and benefits) • Saving money at the store (budgeting, meal planning, shopping tips, and preparation tips)	Grocery list pad
Lesson 5	Protect your plate	• Protein group (sources, nutrients, and benefits) • Food safety (clean, separate, cook, and chill)	Food thermometer

Evaluation Tool

Overview

The evaluation toolkit includes evaluations before and after the intervention. They are intended to be administered via Qualtrics (Qualtrics, LLC) at the beginning of the first lesson and the end of the final lesson. The presurvey evaluation contains 60 items, while the postsurvey evaluation includes 46 items. The surveys include questions from the Adult EFNEP Questionnaire assessing diet, physical activity, food safety, and food resource management, in addition to participant demographics. The Adult EFNEP survey was used because F2T is a modified EFNEP program with similar objectives. A description of the survey measures is outlined in the following sections.

Diet

The validated dietary behavior measures include several questions on consumption of targeted food groups and 1 question on the frequency of preparing meals at home [27]. The 10-item food frequency questionnaire asks participants about how many times per day (or per week) they consume certain foods or beverages, including vegetables, fruit, milk and soymilk, beans and peas, and regular soda. The food preparation question asks participants how many days a week they cook dinner (or their main meal) at home. Response options range from “I rarely cook dinner at home” to “6 to 7 days a week.”

Physical Activity

The Adult EFNEP Questionnaire includes several questions regarding participants’ physical activity levels. One question asks participants how many days in the past week they exercised for at least 30 minutes. The next question asks how many days they did strength-related workouts. A third asks how often they make small changes to be more active.

Food Safety

Four survey questions assess food safety practices. Participants indicate how often they wash their hands before preparing food, wash all items and surfaces that came into contact with raw meat or seafood, thaw frozen food at room temperature, and use a meat thermometer to check the internal temperature when cooking meat. Response options range from “never”=1 to “always”=5.

Food Resource Management

Nine validated survey questions assess food resource management practices [28]. These include how often participants compare food prices to save money, how often they plan meals before grocery shopping, and how often they use a written

XSL•FO
RenderX

weekly or monthly food spending plan (refer to Multimedia Appendix 1 for the evaluation survey questions). Response options range from “never”=1 to “always”=5.

Demographic Variables

Questions are included to capture demographic characteristics, including year of birth, sex (male or female), race (American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Pacific Islander, White, or other), Hispanic ethnicity (yes or no), occupation status (employed or self-employed full time, employed or self-employed part time, unemployed or unable to work, or student only), and current relationship status (single, in a relationship, married or in a domestic partnership, divorced or separated, or widowed). To gain some understanding of the environment in which participants live and resources available to them, additional questions ask about current living situation (campus or university housing, a family member’s home, or off-campus or other nonuniversity housing), whether they are on a university meal plan (yes or no), whether they are an international student (yes or no), and programs from which they receive assistance (Pell Grants, food assistance programs [eg, the Supplemental Nutrition Assistance Program], and other programs [eg, Temporary Assistance for Needy Families]).

Curriculum Evaluation

The goal of the curriculum evaluation is to determine the program’s relevance, feasibility, and effectiveness. The specific objectives are (1) to assess the acceptability and relevance of the curriculum to emerging adults, (2) to determine the feasibility of implementing the program within a university setting, and (3) to examine the effectiveness of the program in improving participant knowledge and food-related behaviors.

The first 2 objectives are part of the process evaluation and will be addressed through focus group discussions with program participants and one-on-one, semistructured interviews with program educators. The third objective will be addressed by analyzing the quantitative data collected before and after the program and responses to behavioral change–related questions discussed in the focus groups.

Process Evaluation

The focus groups will be held in a separate room following the fifth lesson. A trained moderator will ask participants for their perspectives on the relevance, engagement, and usefulness of the course content and activities, along with questions to obtain feedback on improving the curriculum (refer to Multimedia Appendix 2 for the focus group questions). Video and audio of the focus groups will be recorded and later transcribed, and a member of the research team will be present to take notes. Focus groups will be conducted until data saturation.

Interviews with instructors will take place virtually 1 week after the final class. A trained interviewer will ask the instructors questions related to the overall flow and suitability of the curriculum, whether adjustments were made to the planned curriculum and activities, and barriers and facilitators to student recruitment and retention in the program (refer to Multimedia Appendix 3 for the interview questions). Interviews will be video- and audio-recorded and later transcribed.

The process evaluation will also include quantitative measures of the number of students who sign up for each class and the number of classes students attended. In the analysis, quantitative evaluation measures will be reported as means and SDs. The lead qualitative researcher will develop a codebook based on a priori descriptive categories, carefully read and code 1 interview and 1 focus group, and refine the codebook accordingly. Two other researchers will use this codebook to independently code 1 additional interview and 1 focus group, and then the 3 researchers will meet as a team to discuss discrepancies in coding decisions. The codebook will be revised based on the consensus reached during these discussions. Two researchers will then independently code the remaining interviews and focus groups using the updated codebook. Categories and supporting quotes will be compiled and presented.

Outcome Evaluation

Data collected before and after the program will be evaluated for program effectiveness. We will use a single-group, pretest-posttest design to assess outcomes before program initiation and at program completion. Two variables will be created for physical activity questions: (1) an indicator of whether participants meet current guidelines [29] (ie, engaging in any exercise on at least 5 days a week and resistance training on at least 2 days a week) and (2) how often the participants make small changes to be more physically active.

For diet behaviors, food safety, and food resource management, descriptive statistics will be used to summarize the presurvey and postsurvey responses to each question. For the main analysis, we will sum the responses to each question within a given category (ie, diet, physical activity, food safety, and food resource management) to form a scale in which higher numbers indicate greater use of each recommended practice or dietary behavior. This approach has been used in previous nutrition education research [30].

Participant demographic characteristics will be reported at baseline, and chi-square tests (categorical variables) and Wald tests (continuous variables) will be used to compare characteristics of those who completed the program with those who did not. The main analysis will be restricted to participants who completed the program and have both presurvey and postsurvey data. Outcome evaluation measures will be reported as proportions (categorical variables) or means and SDs (continuous variables) at baseline and at program completion. Generalized linear models with robust standard errors clustered at the participant level will be used to assess changes in outcomes from baseline (week 1) to program completion (week 5). Outcomes will also be examined by student participation in federal assistance programs to understand program effectiveness among low-income participants. The independent variable of interest will indicate whether the outcome was measured at baseline or after the intervention. Effect sizes and 95% CIs will be reported for each outcome. All quantitative analyses will be conducted using Stata (version 18.5; StataCorp LLC).

XSL•FO
RenderX

The focus group guide included a few outcome evaluation questions to assess participants' perceptions of how the program shaped their (1) diet-related knowledge and skills, (2) self-efficacy to eat healthier, and (3) other health behaviors (ie, food resource management, food preparation, and food safety). Findings from the qualitative analysis of these questions will be used to offer a comprehensive understanding of the program's outcomes.

Ethical Considerations

The focus group guide and in-depth interview protocols, which include informing participants that they could opt out of the study and obtaining informed consent, have been reviewed and approved by the Texas A&M University Institutional Review Board (STUDY2025-0270). To ensure privacy, no identifiable information will be included in any write-up or presentation of the results. Participants will not receive compensation for participation. A separate institutional review board approval will be obtained for the analysis of deidentified secondary data collected in program surveys.

Results

The F2T program is currently being pilot-tested by the university's health center. In January 2025, the curriculum developer trained health center staff to deliver the curriculum. This involved providing the curriculum to the 4 educators who would teach the series a few weeks before a 2-day, in-person training. The educators then participated in virtual teach-backs 1 week later. In the teach-backs, the educators taught the class and were given time to ask questions, provide suggestions, and receive constructive feedback.

The health center advertised the program via social media, the health center's online calendar of events, and a large sign displayed just inside the center. The 5 lessons were delivered in person at the health center over 4 weeks. Four educators were involved in program delivery, with one educator leading each lesson and another educator present to assist. Educators took turns leading the lessons so that each had the opportunity to lead at least 1 session.

Enhancement items were offered to all attendees at each lesson (Table 1), and participants were entered into a raffle to win 1 of 3 prizes, which were drawn and distributed at a meeting with participants following the fifth lesson. The 3 final prizes included an air fryer, a blender, and an electric skillet. Educators administered the evaluation tool described previously at the beginning and end of the program to measure program effectiveness.

The program will continue to be offered through fall 2025 and spring 2026. For these cohorts, we will include the Short Healthy Eating Index Survey to more accurately capture and quantify changes in participant diet quality [31]. This will allow comparison of effectiveness results with other program evaluations that examine diet quality while minimizing respondent and staff burden. We will also include select questions from the Young Adult Nutrition Literacy Tool to assess changes in confidence around healthy eating behaviors [32].

Discussion

The transition to adulthood is characterized by newfound autonomy and responsibility, which may impact dietary behaviors and health. Despite having the lowest diet quality of all adults, emerging adults have been understudied and have not been targeted for nutrition interventions as a population distinct from general population adults. A nutrition education program developed specifically for emerging adults can fill this important nutrition intervention gap. In creating, implementing, and evaluating this program, we expect to find improvements in participants' confidence around healthy eating behaviors and behaviors related to diet, food safety, physical activity, and food resource management at program completion. We also anticipate identifying areas for program improvement to better serve emerging adults and subsequently refining the program for future iterations.

The expected outcomes of the F2T program align with findings from nutrition education interventions that have been implemented among college student populations. Three such programs focused on nutrition education and cooking skills and were provided over 4 sessions, ranging from 15 minutes to 2 hours [33-35]. All 3 incorporated elements similar to those of the F2T program, including nutrition education lessons, cooking demonstrations, and tastings, along with other hands-on activities. The concepts covered, such as goal setting, cooking skills, healthy dietary patterns, and identifying healthy foods in the grocery store, also align with those covered in F2T; however, our program adds information on budgeting and physical activity. Each program reported positive impacts on self-efficacy related to cooking, fruit and vegetable consumption, or grocery shopping, or increased healthy food consumption following the program. Two interventions included follow-up assessments with participants either 4 or 6 months later and reported sustained increases in nutrition knowledge and healthy eating self-efficacy [33,35]. Together, these findings suggest that F2T may also be effective in improving confidence and dietary intake, and they underscore the importance of following up with participants in future research to better understand sustained changes and to revise the program as needed.

The strengths of this program and evaluation include modification of an existing EFNEP program using input from experts who have worked closely with emerging adult populations across multiple settings. We are also using an established EFNEP questionnaire in our evaluation, which will allow comparison of our results with those of other EFNEP programs. Additionally, the qualitative data we will collect will enable improvements to future program offerings and provide additional evidence on effectiveness. Limitations include students' self-selection into the program, which likely reflects greater interest in nutrition and health compared with other emerging adults. The program will also include college students from 1 university, limiting the generalizability of the findings. Finally, there will be no comparison group in the evaluation.

We plan to disseminate our evaluation findings to academic audiences through presentations at research conferences and publications in peer-reviewed journals. We will also share our

XSL•FO
RenderX

findings with Extension professionals through presentations at Extension-focused conferences and via a program brief. Future directions include considerations for program modifications (eg, additional sessions or engagement, different content, and tailored examples) based on the findings. We also plan to test the curriculum with emerging adult EFNEP participants without children in Texas. As part of this effort, we will consider the number of sessions needed for program graduation and compare outcomes from this program with those from longer EFNEP series. Additional efforts may include longer-term follow-up and the addition of a comparison group.

Our goal for this program is for it to be adopted by EFNEP and implemented at universities and in other settings with emerging adults across the United States following additional testing. We hope that this program will positively impact the dietary behaviors and health of emerging adults in the future.

Funding

No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Evaluation survey questions.
[DOCX File , 33 KB - resprot_v15i1e81647_app1.docx]

Multimedia Appendix 2
Focus group questions.
[DOCX File , 16 KB - resprot_v15i1e81647_app2.docx]

Multimedia Appendix 3
Interview questions.
[DOCX File , 19 KB - resprot_v15i1e81647_app3.docx]

References

1. Wood D, Crapnell T, Lau L, Bennett A, Lotstein D, Ferris M, et al. Emerging adulthood as a critical stage in the life course. In: Halfon N, Forrest CB, Lerner RM, Faustman EM, editors. Handbook of Life Course Health Development. Cham, Switzerland: Springer; 2018:123-143.
2. Corder K, Winpenny E, Love R, Brown HE, White M, Sluijs EV. Change in physical activity from adolescence to early adulthood: a systematic review and meta-analysis of longitudinal cohort studies. Br J Sports Med 2019 Apr 24;53(8):496-503 [FREE Full text] [doi: 10.1136/bjsports-2016-097330] [Medline: 28739834]
3. Cowan-Pyle AE, Bailey RL, Gao J, Hess JM, Ademu LO, Smith JL, et al. Dietary quality and diet-related factors among emerging adults (18-23 y) in the United States are a cause for concern: National Health and Nutrition Examination Survey 2015-2018. J Nutr 2024 Aug;154(8):2524-2533 [FREE Full text] [doi: 10.1016/j.tjnut.2024.06.015] [Medline: 38944345]
4. Whatnall MC, Patterson AJ, Burrows TL, Hutchesson MJ. Higher diet quality in university students is associated with higher academic achievement: a cross-sectional study. J Hum Nutr Diet 2019 Jun 27;32(3):321-328. [doi: 10.1111/jhn.12632] [Medline: 30810252]
5. Solomou S, Robinson H, Perez-Algorta G. The association of diet quality with the mental health of students during their transition to university. PLoS One 2024 Oct 15;19(10):e0312123 [FREE Full text] [doi: 10.1371/journal.pone.0312123] [Medline: 39405296]
6. Phelan JM, Rosenkranz RR, Logan C, Haub MD, Mailey EL, Ezzati A, et al. The association of overall dietary quality on stress, anxiety, depression, and academic performance in college students: a systematic review. J Am Coll Health 2025 Mar 08;73(3):977-988 [FREE Full text] [doi: 10.1080/07448481.2024.2418521] [Medline: 39514800]
7. Merhout F, Doyle J. Socioeconomic status and diet quality in college students. J Nutr Educ Behav 2019 Oct;51(9):1107-1112. [doi: 10.1016/j.jneb.2019.06.021] [Medline: 31345675]
8. Larson NI, Perry CL, Story M, Neumark-Sztainer D. Food preparation by young adults is associated with better diet quality. J Am Diet Assoc 2006 Dec;106(12):2001-2007. [doi: 10.1016/j.jada.2006.09.008] [Medline: 17126631]
9. Reicks M, Trofholz AC, Stang JS, Laska MN. Impact of cooking and home food preparation interventions among adults: outcomes and implications for future programs. J Nutr Educ Behav 2014;46(4):259-276 [FREE Full text] [doi: 10.1016/j.jneb.2014.02.001] [Medline: 24703245]

XSL•FO
RenderX

10. Utter J, Larson N, Laska MN, Winkler M, Neumark-Sztainer D. Self-perceived cooking skills in emerging adulthood predict better dietary behaviors and intake 10 years later: a longitudinal study. J Nutr Educ Behav 2018 May;50(5):494-500 [FREE Full text] [doi: 10.1016/j.jneb.2018.01.021] [Medline: 29525525]
11. Meiklejohn S, Ryan L, Palermo C. A systematic review of the impact of multi-strategy nutrition education programs on health and nutrition of adolescents. J Nutr Educ Behav 2016 Oct;48(9):631-46.e1. [doi: 10.1016/j.jneb.2016.07.015] [Medline: 27720105]
12. Expanded Food and Nutrition Education Program (EFNEP). National Institute of Food and Agriculture. URL: https://www.nifa.usda.gov/grants/programs/capacity-grants/efnep/expanded-food-nutrition-education-program [accessed 2025-06-20]
13. WebNEERS: EFNEP reporting system. National Institute of Food and Agriculture. URL: https://www.nifa.usda.gov/grants/programs/capacity-grants/expanded-food-nutrition-education-program-efnep/webneers [accessed 2025-12-11]
14. Krebs-Smith SM, Pannucci TE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, et al. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet 2018 Sep;118(9):1591-1602 [FREE Full text] [doi: 10.1016/j.jand.2018.05.021] [Medline: 30146071]
15. Feder M. Adult learning theory and the principles of andragogy. University of Phoenix. URL: https://tinyurl.com/28tp8be6 [accessed 2025-10-16]
16. Lanctôt J, Poulin F. Cumulative transitions in emerging adulthood: portents of adjustment problems or a window for opportunities? Emerg Adulthood 2024 Feb 26;12(3):275-289. [doi: 10.1177/21676968241234951]
17. Allman-Farinelli MA. Nutrition promotion to prevent obesity in young adults. Healthcare (Basel) 2015 Sep 07;3(3):809-821 [FREE Full text] [doi: 10.3390/healthcare3030809] [Medline: 27417798]
18. Plotnikoff RC, Costigan SA, Williams RL, Hutchesson MJ, Kennedy SG, Robards SL, et al. Effectiveness of interventions targeting physical activity, nutrition and healthy weight for university and college students: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2015 Apr 01;12:45 [FREE Full text] [doi: 10.1186/s12966-015-0203-7] [Medline: 25890337]
19. Whatnall MC, Patterson AJ, Chiu S, Oldmeadow C, Hutchesson MJ. Feasibility and preliminary efficacy of the eating advice to students (EATS) brief web-based nutrition intervention for young adult university students: a pilot randomized controlled trial. Nutrients 2019 Apr 23;11(4):905 [FREE Full text] [doi: 10.3390/nu11040905] [Medline: 31018565]
20. Expanded Food and Nutrition Education Program (EFNEP) guidelines for using technology in programming. National Institute of Food and Agriculture. 2022 Jun. URL: https://www.nifa.usda.gov/efnep-program-guidelines [accessed 2025-10-24]
21. What is youth EFNEP? Ohio State University Extension. URL: https://hamilton.osu.edu/program-areas/efnep/efnep-programs-youth [accessed 2025-10-24]
22. Rhode Island Expanded Food and Nutrition Education Program. University of Rhode Island. URL: https://web.uri.edu/community-nutrition/efnep/curriculum/ [accessed 2025-10-24]
23. About EFNEP. The University of Arizona Cooperative Extension. URL: https://extension.arizona.edu/programs/expanded-food-and-nutrition-education-program-efnep/about [accessed 2025-10-24]
24. Story M, Kaphingst KM, Robinson-O'Brien R, Glanz K. Creating healthy food and eating environments: policy and environmental approaches. Annu Rev Public Health 2008 Apr;29(1):253-272. [doi: 10.1146/annurev.publhealth.29.020907.090926] [Medline: 18031223]
25. Contento IR, Koch PA. Nutrition Education: Linking Research, Theory, and Practice. 4th edition. Burlington, MA: Jones & Bartlett Learning; 2020.
26. The Expanded Food and Nutrition Education Program policies. National Institute of Food and Agriculture. URL: https://www.nifa.usda.gov/sites/default/files/2025-04/EFNEP-Policy-2025-Update.pdf [accessed 2025-10-24]
27. Murray EK, Baker SS, Betts NM, Hess A, Auld G. Development of a National Dietary Behaviors questionnaire for EFNEP adult participants. J Nutr Educ Behav 2020 Dec;52(12):1088-1099. [doi: 10.1016/j.jneb.2020.06.003] [Medline: 32763052]
28. Barale K, Aragón MC, Yerxa K, Auld G, Hess A. Development of reliable and valid questions to assess food resource management behaviors in adults with limited income. J Nutr Educ Behav 2022 Apr;54(4):346-358. [doi: 10.1016/j.jneb.2021.11.004] [Medline: 35131185]
29. Physical activity guidelines for Americans, 2nd edition. U.S. Department of Health and Human Services. 2018. URL: https://odphp.health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf [accessed 2025-05-15]
30. Adedokun OA, Plonski P, Jenkins-Howard B, Cotterill DB, Vail A. Healthy choices for every body adult curriculum improves participants' food resource management skills and food safety practices. J Nutr Educ Behav 2018 Jun;50(6):638-644. [doi: 10.1016/j.jneb.2018.02.005] [Medline: 29625915]
31. Colby S, Zhou W, Allison C, Mathews AE, Olfert MD, Morrell JS, et al. Development and validation of the short healthy eating index survey with a college population to assess dietary quality and intake. Nutrients 2020 Aug 27;12(9):2611 [FREE Full text] [doi: 10.3390/nu12092611] [Medline: 32867172]
32. McNamara J, Kunicki ZJ, Neptune L, Parsons K, Byrd-Bredbenner C. Development and validation of the Young Adult Nutrition Literacy Tool. J Nutr Educ Behav 2022 Jul;54(7):691-701. [doi: 10.1016/j.jneb.2022.01.011] [Medline: 35469757]
33. Szczepanski JR, Litchfield RE, Beirman EA, Nolting LM, Genschel U. Effects of a Culinary Boot Camp intervention on food/nutrition aptitudes and dietary intake of college students. J Am Coll Health 2024 Jan;72(1):55-64. [doi: 10.1080/07448481.2021.2016771] [Medline: 35015961]

XSL•FO
RenderX

34. McMullen J, Ickes M. The influence of a campus-based culinary, nutrition education program, "College CHEF," on college students' self-efficacy with cooking skills and nutrition behaviors. Build Healthy Acad Communities J 2017 Nov 27;1(2):61-76. [doi: 10.18061/bhac.v1i2.5987]
35. Clifford D, Anderson J, Auld G, Champ J. Good Grubbin': impact of a TV cooking show for college students living off campus. J Nutr Educ Behav 2009 May;41(3):194-200. [doi: 10.1016/j.jneb.2008.01.006] [Medline: 19411053]

Abbreviations

EFNEP: Expanded Food and Nutrition Education Program
F2T: Fuel to Thrive
HBHM: Healthy Bites, Healthy Moves

Edited by J Sarvestan; submitted 31.Jul.2025; peer-reviewed by ST Henes, Y Qin; comments to author 17.Sep.2025; accepted 17.Dec.2025; published 15.Jan.2026.

Please cite as:
Rogus S, Alfaro Hudak K, Ademu LO, Venkatesh S, Laguros M, Racine EF
Nutrition Education for Emerging Adults: Protocol for Program Evaluation
JMIR Res Protoc 2026;15:e81647
URL: https://www.researchprotocols.org/2026/1/e81647
doi:10.2196/81647
PMID:

©Stephanie Rogus, Katelin Alfaro Hudak, Lilian O Ademu, Sumathi Venkatesh, Michael Laguros, Elizabeth F Racine. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Comparative Evaluation of Push-Out Bond Strength of Glass Fiber Posts and Carbon Fiber Posts in Root Canals Treated With Calcium Hydroxide–Based, Resin-Based, and Bioceramic-Based Root Canal Sealers: Protocol for an In Vitro Study

Vinus Shivlani[1], MDS; Manoj Chandak[1], MDS; Sanika Damle[1], MDS

Department of Conservative Dentistry and Endodontics, Faculty of Dentistry, Datta Meghe Institute of Medical Sciences, Wardha, Maharashtra, India

Corresponding Author:
Vinus Shivlani, MDS
Department of Conservative Dentistry and Endodontics
Faculty of Dentistry
Datta Meghe Institute of Medical Sciences
Wardha, Maharashtra
India
Phone: 91 8975020387
Email: vinus.shivlani@gmail.com

Abstract

Background: For root canal procedures to be successful, adequate bond strength between endodontic sealer and post material is necessary. For postendodontic restorations, glass fiber posts and carbon fiber posts are frequently used. Depending on the type of root canal sealer used, such posts may work differently. The chemical composition and characteristics of calcium hydroxide–based, resin-based, and bioceramic-based sealers vary, which may have an impact on the posts' binding strength. Therefore, optimizing rehabilitative results requires an understanding of these connections.

Objective: This study aims to evaluate and compare the push-out bond strength of glass fiber posts and carbon fiber posts in root canals treated with calcium hydroxide–based, resin-based, and bioceramic-based endodontic sealers.

Methods: A total of 60 extracted human premolars will be used in this in vitro investigation. After canal preparation, specimens will be separated into 3 groups according to the type of sealer used (bioceramic-based, resin-based, and calcium hydroxide–based). The type of post (carbon fiber or glass fiber) will be used to further split each group into 2 subgroups. A universal testing machine will be used to exert a compressive force on each post to test its push-out bond strength. Bond strength data will be recorded in megapascals and analyzed using ANOVA and post hoc tests.

Results: The results are expected to demonstrate significant differences in push-out bond strength among different post and sealer combinations. Glass fiber posts are expected to have higher bond strength values than carbon fiber posts across all sealer groups, with the highest bond strength anticipated in the bioceramic-based sealer group. Resin-based sealers are expected to exhibit intermediate bond strength values, whereas calcium hydroxide–based sealers are expected to show the lowest bond strength values across both types of post.

Conclusions: Glass fiber posts are expected to offer superior push-out bond strength in comparison to carbon fiber posts, especially when used with bioceramic-based sealers. The type of root canal sealer significantly affects bond strength, with bioceramic-based sealers providing the most reliable bond. Findings are expected that will suggest that careful selection of both post material and sealer type is necessary to enhance the long-term success of root canal restorations.

International Registered Report Identifier (IRRID): DERR1-10.2196/76621

(JMIR Res Protoc 2026;15:e76621) doi:10.2196/76621

KEYWORDS

bioceramic sealer; endodontic sealer; fiber post; carbon post; resin-based sealer; calcium hydroxide–based sealer

XSL•FO
RenderX

Introduction

Background

The resistance of a tooth to fracture should be enhanced by restoration. Posts are indicated to preserve the core that restores lost coronal structure in teeth with significant damage [1]. When the remaining cervical tooth tissue is unable to hold and retain the tooth, a post should be considered. In recent decades, the most popular use of metal cast posts has been for post and core restorations [2]. However, conventional cast posts have several drawbacks, including stiffness due to their high modulus of elasticity compared with that of dentin and the need for several fabrication sessions.

Concerns such as root fracture, risk of corrosion, and loss of retention of the post or crown associated with different metals prompted the search for posts with better load-bearing and retentive properties. According to clinical research [3], teeth rebuilt with fiber-reinforced posts have a success rate of 95% to 99%, with no root fractures occurring during the study period. The main benefits of fiber posts include high fatigue and tensile strength, a modulus of elasticity similar to that of dentin, and the ability to be cemented with an adhesive luting material, which reduces friction between the post and root canal walls and ensures uniform distribution of applied forces along the length of the post [3].

The milestone that altered some of the underlying concepts in the restoration of teeth that had undergone endodontic treatment was the introduction of the first fiber post in dentistry, known as carbon fiber–reinforced resin post (or carbon fiber post), by Duret et al [4]. These posts were the first widely accepted substitute for prefabricated metal and cast posts. Initially, unidirectional carbon or graphite fibers with a diameter of 8 µm were used to strengthen a matrix composed of 64% epoxy resin by weight in carbon fiber posts. Nevertheless, as carbon fiber posts are dark in color, they lack esthetic appeal. Their strong flexural resistance means that less dentin structure needs to be removed, which helps prevent fracture of both the post and the remaining tooth [4]. When subjected to intense forces, the post can follow tooth movement because its modulus of elasticity is equal to that of dentine, thereby providing weakened roots with greater resistance to breakage [4].

To improve the distribution of masticatory pressures and ensure more retention and support for restorative materials, glass fiber posts are recommended for the rehabilitation of severely damaged teeth. The benefits of glass fiber posts include high tensile strength, esthetic appeal, and a dentin-like modulus of elasticity, which distributes stress uniformly along the length of the post [5]. However, comparative research indicates that carbon fiber posts exhibit more flexural resistance than glass fiber posts [6]. Posts are frequently bonded to root canal walls using resin cements. Fiber posts adhere to dentin and resin cements with remarkable bond strength.

Although several types of prefabricated posts made of metal and resin are currently available in the market, carbon fiber posts and glass fiber posts are often advised because they strengthen the tooth structure and increase its overall integrity. Previous studies provide valuable insights into the performance of different post and sealer combinations under stress, mimicking the conditions encountered during functional loading [7]. Furthermore, the longevity and adherence of a post retainer to radicular dentin may be affected by the endodontic sealer, depending on its composition. Presence of eugenol in the cement's composition significantly reduced adhesion, according to research showing the effect of zinc oxide–eugenol–based cement on intraradicular post retention [7]. The effects of zinc oxide–eugenol cement, epoxy resin, and calcium hydroxide on the retention of prefabricated titanium dowels luted with resin cement were investigated in a different investigation. According to the results [8], the groups that did not use cement and those that used a sealant containing eugenol differed significantly. In a related investigation, Ngoh et al [9] found that eugenol decreased the resin cement's ability to adhere to dentin in the coronal third of the root.

In clinical practice, calcium hydroxide–based, resin-based, and bioceramic-based root canal sealers are commonly used, although a variety of sealers have been developed and are commercially available. Although resin-based sealers are well known for their superior bonding powers to dentin and root posts, calcium hydroxide is recognized for its antibacterial properties and its capacity to aid in the healing of periapical tissues. Bioactivity, biocompatibility, and sealing properties of bioceramic sealers, a more recent class of materials, are valued [10]. This emphasizes the necessity of assessing the effects of various fiber posts and popular sealers combinations on bond strength in endodontically treated teeth.

Due to varying properties such as surface chemistry and elasticity of carbon fiber posts and glass fiber posts, along with resin-based, bioceramic-based, and calcium hydroxide–based sealers used in endodontics, these materials leave behind a distinct residual impact on dentin, which ultimately affects the adhesion of resin cements. Hence, studying all these combinations would suggest a more reliable post and sealer pairing, maximizing push-out bond strength and guiding toward better clinical restorability. For restorations retaining the dental post to be successful over the long term, a good relationship between post materials and root canal sealers is essential. Few studies have examined the effects of different endodontic sealers on the bonding interface between the post, resin cement, and radicular dentin, although the mechanical compatibility of fiber posts with dentin has been extensively researched. It is well known that some sealers, especially those based on eugenol, inhibit the polymerization of resin cements, thereby compromising post retention [7-9].

Although bioceramic sealers are valued for their bioactivity and sealing ability and differ chemically from other sealers, resin-based sealers provide better adhesion to dentin and greater compatibility with resin cements [10]. The surface effects and residual chemical composition of these sealers on canal walls can differ greatly, which affects how well they adhere to glass fiber posts and carbon fiber posts. Nevertheless, there are currently no thorough analyses in the literature contrasting the push-out bond strength of these fiber post types with various widely used sealers.

XSL•FO
RenderX

There is a dearth of literature comparing the bond strength of carbon fiber posts and glass fiber posts in root canals treated with calcium hydroxide–based, resin-based, or bioceramic-based endodontic sealers [10].

Objectives

The primary objective of this study is to evaluate the push-out bond strength of glass fiber posts and carbon fiber posts in root canals treated with calcium hydroxide–based, resin-based, and bioceramic-based sealers.

Moreover, this study also compares the push-out bond strength of glass fiber posts and carbon fiber posts when used with the same root canal sealers.

Methods

Materials Required

A total of 60 extracted human premolars will be chosen for this investigation. Teeth that are removed for orthodontic reasons and are devoid of cavities, cracks, or previous endodontic procedures will be used. The teeth will be preserved in distilled water after being cleaned with ultrasonic scalers to remove any calculus and soft tissue remnants before the experiment. The teeth will be decoronated at 15 (SD 1) mm using water coolant and slow-speed disks.

A high-speed handpiece (KaVo) equipped with a round carbide bur (size #2) will be used to create a straight-line access to the canals. After that, a crown-down procedure will be used to instrument the root canals. To clean the canal and remove the debris, 3% sodium hypochlorite will be used.

The first cleaning will be done with a size 15 K-file (Dentsply Sirona). Using step-back instrumentation, the canal will be progressively enlarged up to a size 40 file. There will be extensive irrigation between each file to get rid of germs and debris. A size 40/04 rotary file (ProTaper Next, Dentsply Sirona) will be used for the final preparation once the canals reach the necessary size. This will guarantee that the canals are correctly tapered and well shaped. After removal of the smear layer with final irrigation using a solution of 3% sodium hypochlorite and 17% ethylenediaminetetraacetic acid, the canals will be rinsed with distilled water.

After preparation, the following 3 distinct types of root canal sealers will be used to obturate the canals:

1. Calcium hydroxide–based sealer (Sealapex, Kerr)
2. Resin-based sealer (Meta Adseal, Meta Biomed)
3. Bioceramic-based sealer (Bio C sealer, Angelus) The root canals will be ready for post insertion once the sealer has solidified, which should happen in 6 to 7 days. To make room for the placement of posts, a little quantity of the obturation material will be removed from the canals using a Gates-Glidden drill (Dentsply Sirona). Using a bespoke drill, a uniform post space of approximately 10 to 12 mm will be created.

Luting and Post Selection

This study will use carbon fiber posts (Reforpost, Angelus) and glass fiber posts (Reforpost, Angelus). To match the designated post space, the posts will be chosen according to their individual diameters. To improve bonding to the sealer, the posts will be coated with a silane coupling agent (Silane, 3M ESPE) before being placed.

Cementing

In accordance with the manufacturer's instructions, a self-etch resin cement (Dentsply, Calibra, Universal Self-Adhesive resin system) will be used to cement the posts within the post space. To ensure that the sealer is evenly distributed around the post, the posts will be carefully placed into the prepared canals. After removing any excess cement, the posts will be light-cured from various angles for 20 seconds each to ensure full polymerization. It is recommended to apply the standardized force of 20 to 30 N for cementation of the post to the root dentin. While light curing, the tip of the curing unit must be placed at 90° to the canal, ensuring maximum contact and light exposure.

Push-Out Bond Strength Test

To allow the cement to adequately set, these specimens will be kept in distilled water at 37 °C for 1 day following the post luting procedure. A diamond blade in a water-cooled cutting machine will next be used to segment the teeth transversely into slices that are 1 mm thick. Once the slices are placed in acrylic blocks, a universal testing machine will be used to measure the push-out bond strength at a crosshead speed of 1 mm per minute. The post will be compressed in an apical-coronal orientation using a cylindrical rod until dislodgement occurs.

The bond strength will be estimated by dividing the maximum force required for post dislodgement by the surface area of the post.

Inclusion Criteria

Only healthy human extracted teeth, free of caries, fractures, and any prior dental procedures, will be included in the study. Teeth must have enough coronal structure for post placement, a single straight root canal system, and a root length of approximately 14 mm. Teeth without previous endodontic treatment and extracted for nonpathological reasons, such as orthodontics, will be used.

Exclusion Criteria

Caries-laden, fractured, cracked, or structurally flawed teeth will not be included. Additionally, teeth with numerous roots or intricate canal systems, previous endodontic procedures, severe root resorption, or inadequate root length (<12 mm) will be excluded. Teeth with inadequate coronal structure that cannot sustain appropriate post implantation will be excluded.

Sample Size

The sample size formula is as follows:

$$n_1=\Delta 2(\sigma_1^2+\kappa\sigma_2^2)\ (Z_{1-\alpha/2}+Z_{1-\beta})2$$

Where n_1=sample size of group 1; n_2=sample size of group 2; σ_1=SD of group 1; σ_2=SD of group 2; Δ=difference in group means; κ=ratio of n_2 to n_1 (n_2/n_1); $Z_{1-\alpha/2}$=2-sided Z value (eg, Z=1.96 for 95% CI); and $Z_{1-\beta}$=power.

XSL•FO
RenderX

Specific values are as folows: mean bond strength in Endofill (Dentsply Sirona)+Variolink II=10.38; mean bond strength in Endofill+RetyXU200=11.75; σ_1=SD of bond strength in Endofill+Variolink II=1.8; and σ_2=SD of bond strength in Endofill+RetyXU200=2.0.

Thus, for detecting a mean difference of 1.37, that is, Δ=11.75−10.38=1.37, we need K=1; N=(1.8 × 1.8 + 2.0 × 2.0) $(1.9 + 0.7)^2$ 2.0 × 2.0 = 9.87 = 10 specimens in each group (6 × 10 = 60). The power of the test will be 80% and the level of significance will be 5% (95% CI).

Sample Allocation

A total of 60 samples will be divided into 2 main groups based on the type of post: 30 samples each for carbon fiber posts (group A) and glass fiber posts (group B). Each group will then be divided into 3 subgroups according to the type of root canal sealer used: calcium hydroxide–based, resin-based, or bioceramic-based.

All 60 samples will first be allocated to a specific post type (n=30 each). The segregation of samples according to different sealer types (n=10 each) will be done using a random sequence. This allocation will be sealed and sequentially numbered by an independent researcher. Outcome assessment of push-out bond strength will be performed by a blinded evaluator using neural-coded samples, thereby minimizing detection bias and ensuring unbiased allocation. This study will use a comparative design: an in vitro experimental study comparing multiple intervention groups under standardized conditions.

Ethical Considerations

Ethics approval for this study has been obtained from the institutional ethics committee of Datta Meghe Institute of Higher Education & Research (IEC/2025/544).

All extracted teeth will be collected from the department of oral surgery, ensuring that they are not at risk of harm due to the use of the teeth in research. This study will follow strict guidelines for ethical handling of human tissues. All study data will be handled in accordance with applicable data protection and privacy regulations. Participant confidentiality will be strictly maintained throughout the study. Personal identifiers will be removed and replaced with unique study codes prior to data analysis. The key linking participant identities to study codes will be stored separately and securely, accessible only to the principal investigator.

Electronic data will be stored on password-protected computers and/or encrypted institutional servers, while any physical records will be kept in locked cabinets within secure research facilities. Only authorized members of the research team will have access to the data. No individually identifiable information will be disclosed in any publications, presentations, or reports arising from this study.

Written informed consent will be obtained from all participants prior to enrollment in the study. Participants will be provided with detailed information regarding the study objectives, procedures, potential risks and benefits, data handling practices, and their right to withdraw from the study at any time without any impact on their care. For participants who are unable to provide written consent, consent will be obtained from a legally authorized representative in accordance with institutional and ethical guidelines

Primary Outcome

The primary finding of this study is expected to be that glass fiber posts will exhibit the strongest push-out bond, particularly when combined with resin-based and bioceramic-based sealers. Conversely, it is expected that carbon fiber posts will exhibit a lower push-out bond strength, especially when combined with calcium hydroxide–based sealers.

Statistical Analysis Plan

All push-out bond strength values will first be summarized using descriptive statistics (mean [SD] for each post and sealer combination). Data distribution normality and variance homogeneity will be assessed using the Shapiro-Wilk and Levene tests, respectively. If both assumptions are satisfied, group comparisons will be performed using one-way ANOVA, followed by Tukey post hoc tests for multiple comparisons and independent Student *t* tests (2-tailed) for paired contrasts. If variance homogeneity is violated, Welch ANOVA (or Brown-Forsythe test) will be used, whereas if normality assumptions are not met—especially in small or skewed samples—data transformation (eg, logarithmic or square root) or nonparametric alternatives will be used, including the Mann-Whitney U test for 2-group comparisons and the Kruskal-Wallis test for comparisons involving ≥3 groups, with appropriate pairwise follow-up testing (eg, Dunn test). All analyses will be conducted using SPSS (version 27.0; IBM Corp) and GraphPad Prism (version 7.0; GraphPad Software Inc) to ensure robustness in handling assumption violations while preserving statistical power.

Results

Luting and Post Selection

Nearly all posts (95%-98%) are expected to fit within +0.1 or −0.1 mm of the designated post spaces, ensuring appropriate seating. Under microscopy, mixed failures are anticipated in approximately 60% of specimens, with adhesive failures at the cement-dentin interface comprising approximately 30% and adhesive failures at the cement-post interface comprising approximately 10%.

Cementing

Posts are expected to seat fully under a standardized force of 20 to 30 N without delay. Light curing at a 90° angle is expected to yield appropriate polymerization beyond International Organization for Standardization (ISO) standards. Cement extrusion beyond the apex is anticipated in approximately 5% of specimens, occurring most frequently in the calcium hydroxide–based sealer (Sealapex) group.

Push-Out Bond Strength

Results are expected to demonstrate substantial differences in push-out bond strength among the different post and sealer combinations. Glass fiber posts are expected to have higher

XSL•FO
RenderX

bond strength values than carbon fiber posts across all sealer groups, with the highest bond strength observed in the bioceramic-based sealer group. Resin-based sealers are expected to show intermediate bond strengths, whereas calcium hydroxide–based sealers are expected to show the lowest bond strength across both types of posts.

Timeline

The expected results of this study will be published by January 2026. This research project is currently not funded. Data collection is underway, and analysis of the results began in August 2025.

Discussion

The restoration of endodontically treated teeth using posts has gained immense popularity and is proven to provide better clinical results in badly damaged teeth. Restoration of a tooth with severely damaged coronal structure has been a challenge for endodontists. Effective postendodontic restoration is frequently necessary for long-term success of endodontically treated teeth, and fiber-reinforced posts have become attractive options because of their biomechanical compatibility with dentin. The 2 most commonly used types, glass fiber posts and carbon fiber posts, have unique physical and chemical properties that affect their adhesion to the root dentin.

In root canal obturation, a variety of sealers with differing properties have been used. Biocompatibility and potent antimicrobial effects are the main criteria in the development of root canal sealers [10,11]. Additionally, the bond strength of posts to root dentin is considered to be better in nonfilled canals than that in filled canals [10]. Hence, the most appropriate sealer must be used in endodontic treatment because the type of sealer used can severely affect the bond strength of posts to root dentin. With regard to all these concepts, the effect of different sealer types affecting the push-out bond strength of glass fiber posts and carbon fiber posts to root dentin will be studied.

Teixeira et al [12] analyzed the effect of various endodontic sealers—EndoFill, Sealapex, and EndoREZ (Ultradent Products Inc)—on the bond strength of carbon fiber posts cemented with resin cement. A total of 30 extracted human premolars were segregated based on sealer type and post cementation timing (48 hours vs 7 days) [13]. The results showed that EndoREZ provided the highest bond strength, particularly in the coronal and middle thirds, whereas EndoFill showed the lowest. Mixed failure modes were expected to be the most common. The study emphasized that sealer type, especially the use of EndoREZ, significantly influences bond strength in post retained restorations [12].

A study by Forough Reyhani et al [14] assessed the effect of various endodontic sealers, including MTA Fillapex (Angelus), Dorifill, and AH Plus (Dentsply DeTrey GmbH), on the bond strength of fiber posts cemented with a self-etch adhesive. A total of 72 upper incisors were split up into 4 groups, each of which will be filled with gutta-percha and allocated to either AH Plus, Dorifill, MTA Fillapex, or a control group (no sealer). According to the results, the control group had the highest bond strength (mean 4.45, SD 0.09 MPa), whereas Dorifill had the lowest (mean 1.02, SD 0.03 MPa). Although MTA Fillapex and AH Plus demonstrated stronger bonds than Dorifill, their values remained lower than those of the control group. These findings suggest that the type of sealer significantly influences the bond strength of fiber posts, with MTA Fillapex exhibiting lower resistance to post dislodgment than other sealers [14].

Ruiz et al [13] investigated the influence of root canal sealers, storage duration, and cementation systems on bond strength by dividing 56 extracted human canines into 8 groups. Specimens were sectioned and subjected to bond strength testing after designated storage periods [13]. Their findings revealed that only the storage duration following obturation had a significant impact on bond strength, whereas the type of sealer and cementation system did not show a statistically significant effect ($P>.05$). Notably, glass fiber posts cemented 6 months after obturation demonstrated higher bond strength than those cemented after 1 week, suggesting that the timing of cementation plays a more critical role than the type of sealer used [13].

This study has certain limitations, including in vitro conditions that do not fully replicate the oral environment, the use of premolar teeth that may not represent all tooth types, and a short-term evaluation that does not account for long-term factors such as wear or cyclic loading. Material variability and the lack of clinical follow-up further limit the generalizability of the results. Additionally, the sample size may be insufficient to detect small differences between groups, especially when effect sizes are minimal.

The objective of this study is to assess the effects of various post types and root canal sealers on the push-out bond strength of endodontically treated teeth. The research will yield important information for enhancing restorative techniques and extending the durability of endodontic procedures.

Acknowledgments

The authors acknowledge the generous technical support provided by the research house at Datta Meghe Institute of Higher Education and Research and gratefully acknowledge their guide for support with the project.

Funding

All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article

XSL•FO
RenderX

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request, subject to ethical and data protection restrictions.

Authors' Contributions

All authors contributed to the drafting and preparation of the manuscript. The study was conceptualized and registered by VS and MC. The study design and preliminary manuscript preparation were completed by VS and SD. VS finalized the study design and completed the final manuscript. All authors made equal contributions to the work and reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Chappell RP, Cobb CM, Spencer P, Eick JD. Dentinal tubule anastomosis: a potential factor in adhesive bonding? J Prosthet Dent 1994 Aug;72(2):183-188. [doi: 10.1016/0022-3913(94)90078-7] [Medline: 7932266]
2. Mendoza DB, Eakle WS. Retention of posts cemented with various dentinal bonding cements. J Prosthet Dent 1994 Dec;72(6):591-594. [doi: 10.1016/0022-3913(94)90290-9] [Medline: 7853255]
3. Das AK, Muddugangadhar BC, Amarnath GS, Garg A, Kumar U, Rao TR. Comparative evaluation of push out bond strength of a fiber post system using four different resin cements: an in-vitro study. J Int Oral Health 2015;7(Suppl 1):62-67 [FREE Full text] [Medline: 26225108]
4. Duret B, Reynaud M, Duret F. New concept of coronoradicular reconstruction: the Composipost (1). Chir Dent Fr 1990 Dec 22;60(540):131-141. [Medline: 2272223]
5. Wang X, Shu X, Zhang Y, Yang B, Jian Y, Zhao K. Evaluation of fiber posts vs metal posts for restoring severely damaged endodontically treated teeth: a systematic review and meta-analysis. Quintessence Int 2019;50(1):8-20. [doi: 10.3290/j.qi.a41499] [Medline: 30600326]
6. Schwartz RS, Robbins JW. Post placement and restoration of endodontically treated teeth: a literature review. J Endod 2004 May;30(5):289-301. [doi: 10.1097/00004770-200405000-00001] [Medline: 15107639]
7. Tjan AH, Nemetz H. Effect of eugenol-containing endodontic sealer on retention of prefabricated posts luted with adhesive composite resin cement. Quintessence Int 1992 Dec;23(12):839-844. [Medline: 1305302]
8. Hagge MS, Wong RD, Lindemuth JS. Effect of three root canal sealers on the retentive strength of endodontic posts luted with a resin cement. Int Endod J 2002 May 19;35(4):372-378. [doi: 10.1046/j.0143-2885.2001.00493.x] [Medline: 12059939]
9. Ngoh EC, Pashley DH, Loushine RJ, Weller RN, Kimbrough WF. Effects of eugenol on resin bond strengths to root canal dentin. J Endod 2001 Jul;27(6):411-414. [doi: 10.1097/00004770-200106000-00011] [Medline: 11487137]
10. Nelson S, Albert JM, Milgrom P. Comparative effectiveness of two nonsurgical treatments to reduce oral health disparities from untreated tooth decay in older adults: protocol for a cluster randomized trial. JMIR Res Protoc 2020 Oct 08;9(9):e17840 [FREE Full text] [doi: 10.2196/17840] [Medline: 32897236]
11. Mohamed A, Ismail E, Adam RZ. A comparison of conventional root canal sealers with ones that use green synthesized nanoparticles for antimicrobial activity: protocol for a systematic review. JMIR Res Protoc 2024 Oct 11;13:e51351 [FREE Full text] [doi: 10.2196/51351] [Medline: 39393053]
12. Teixeira CS, Pasternak-Junior B, Borges AH, Paulino SM, Sousa-Neto MD. Influence of endodontic sealers on the bond strength of carbon fiber posts. J Biomed Mater Res B Appl Biomater 2008 Mar 26;84(2):430-435. [doi: 10.1002/jbm.b.30888] [Medline: 17595025]
13. Ruiz L, Mongruel Gomes G, Bittencourt B, Rutz da Silva F, Mongruel Gomes OM, Chidoski Filho JC, et al. Effect of root canal sealers on bond strength of fiber posts to root dentin cemented after one week or six months. Iran Endod J 2018;13(1):54-60 [FREE Full text] [doi: 10.22037/iej.v13i1.17998] [Medline: 29692836]
14. Forough Reyhani M, Ghasemi N, Rahimi S, Milani AS, Omrani E. Effect of different endodontic sealers on the push-out bond strength of fiber posts. Iran Endod J 2016;11(2):119-123 [FREE Full text] [doi: 10.7508/iej.2016.02.009] [Medline: 27141220]

Abbreviations

ISO: International Organization for Standardization

XSL•FO
RenderX

Edited by J Sarvestan, GRM La Rosa; submitted 27.Apr.2025; peer-reviewed by Y Noiri, N Ohkura; comments to author 20.May.2025; revised version received 24.Jun.2025; accepted 19.Jul.2025; published 26.Jan.2026.

Please cite as:
Shivlani V, Chandak M, Damle S
Comparative Evaluation of Push-Out Bond Strength of Glass Fiber Posts and Carbon Fiber Posts in Root Canals Treated With Calcium Hydroxide–Based, Resin-Based, and Bioceramic-Based Root Canal Sealers: Protocol for an In Vitro Study
JMIR Res Protoc 2026;15:e76621
URL: https://www.researchprotocols.org/2026/1/e76621
doi:10.2196/76621
PMID:

©Vinus Shivlani, Manoj Chandak, Sanika Damle. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Comparative Evaluation of Sonic and Ultrasonic Activation of 3 Post Space Irrigants on Push-Out Bond Strength of Fiber Posts Luted With Dual-Cure Resin Cement: Protocol for an In Vitro Study

Apurva Wamane[1]; Aditya Patel[2]

[1]Department Of Conservative Dentistry And Endodontics, Sharad Pawar Dental College And Hospital, Wardha, India

[2]Department of Conservative Dentistry and Endodontics, Datta Meghe Institute of Higher Education and Research, Sharad Pawar Dental College and Hospital, Wardha, India

Corresponding Author:
Apurva Wamane
Department Of Conservative Dentistry And Endodontics
Sharad Pawar Dental College And Hospital
Datta Meghe Institute of Higher Education And Research
Wardha, 442001
India
Phone: 91 8378053480
Email: apurva.wamane@gmail.com

Abstract

Background: For long-term longevity and proper function, endodontically treated teeth must be restored appropriately. Fiber posts luted with dual-cure resin cement are frequently used because of their improved biomechanical and adhesive properties. However, achieving a strong binding between the fiber post and root canal dentin is difficult due to the physical constraints of the root canal and the presence of a smear layer. Bond strength may be increased by post space irrigation with solutions, such as 5% calcium hypochlorite, 17% ethylenediaminetetraacetic acid (EDTA), and saline, that have been activated using sonic or ultrasonic techniques. This study compares the effects of several activation techniques on the push-out bond strength of fiber posts luted with dual-cure resin cement.

Objective: This study aims to evaluate how the push-out bond strength of dual-cure resin–cemented fiber posts in excised human mandibular single-rooted premolars is affected by the sonic and ultrasonic activation of 3 post space irrigants: 5% calcium hypochlorite, 17% EDTA, and saline.

Methods: A total of 120 single-rooted mandibular premolars will be used in this in vitro investigation. A standardized procedure will be used to decoronate, instrument, and obturate the teeth. One of the 3 solutions, activated either sonically or ultrasonically, will be used to prepare and irrigate post spaces. Dual-cure resin cement will then be used to lute the fiber posts. Using a universal testing machine, push-out bond strength will be measured. The Tukey post hoc test and ANOVA will be used for statistical analysis of the data.

Results: Ethics approval for this study was obtained from the institutional ethics committee of Datta Meghe Institute of Higher Education and Research. The study is self-funded. Sample collection of extracted human mandibular premolars began in June 2025, and 68 specimens have been collected as of submission. Specimen preparation is ongoing, with data collection projected to be completed by October 2025. Data analysis is scheduled for November 2025, and results are expected to be published in early 2026. It is anticipated that irrigant type and activation method will influence push-out bond strength, with 17% EDTA and 5% calcium hypochlorite, particularly under ultrasonic activation, potentially yielding higher values than saline.

Conclusions: On the basis of previous literature, we expect that EDTA and calcium hypochlorite with ultrasonic activation may exhibit higher push-out bond strength.

International Registered Report Identifier (IRRID): DERR1-10.2196/76907

(JMIR Res Protoc 2026;15:e76907) doi:10.2196/76907

XSL•FO
RenderX

KEYWORDS

fiber post adhesion; bond strength; post space irrigants; ultrasonic vs sonic activation; 5% calcium hypochlorite; 17% ethylenediaminetetraacetic acid; 17% EDTA; dual-cure resin; smear layer removal

Introduction

Background

Restoring endodontically treated teeth requires consideration of several factors, such as the position of the tooth in the dental arch, the type of restoration required, the amount of healthy tooth structure remaining, and the bond between the restoration and the natural tooth structure. When a significant amount of healthy tooth structure remains, the optimal approach often involves the use of a post and core followed by a final restoration [1,2].

In recent years, adhesive nonmetallic post systems, including glass fiber posts, fiber-reinforced posts, and polyethylene posts, have gained popularity because of their ability to preserve the remaining tooth structure during radicular preparation [3]. Glass fiber posts, which are clinically beneficial, offer a straightforward preparation process. When combined with adhesive cementation, this method produces a stronger bond, ultimately improving the clinical prognosis by enhancing biomechanical behavior under stress and reducing the likelihood of root fractures [4].

Following post space preparation, a smear layer containing abrasive debris, sealer residue, and gutta-percha can form on the dentin walls, which can prevent adequate bonding of the post to the tooth. Therefore, post space irrigation is essential to remove this debris and to provide benefits such as low toxicity, tissue dissolution, and antibacterial activity [5,6].

Achieving reliable bonding to root canal dentin can be challenging because of irregular canal geometry and differences in the physical and chemical properties of the luting materials. Additionally, any previous endodontic treatments may interfere with bonding during post cementation [7]. A strong bond between the fiber post and the root canal walls is critical because debonding is often the primary cause of fiber post failure in clinical settings [8]. The development of self-adhesive cements has simplified the adhesion process, improving bonding to radicular dentin and making it less technique sensitive [9].

Treatment irrigating solutions can alter the chemical and structural composition of dentin, potentially affecting its permeability and solubility [10]. These changes may enhance or hinder the ability of adhesive agents to bond effectively to the dentin [11]. Specifically, calcium hypochlorite has been shown to increase calcium content while decreasing carbon content, which may promote mineralization and lead to the formation of calcium phosphate and calcium carbonate crystals on the dentin surface, thereby strengthening the bond between the adhesive and dentin [12].

Ethylenediaminetetraacetic acid (EDTA), a bivalent cationic substance that chelates calcium ions, is frequently used in endodontics to remove the smear layer. This process removes noncollagenous proteins and hydroxyapatite, thereby facilitating better penetration of resinous agents and resulting in a thinner hybrid layer without denaturation of collagen [13].

The activation of irrigants plays a crucial role in enhancing penetration and effectiveness. Sonic irrigation, operating at frequencies ranging from 1 to 6 kHz (SuperEndo Activator; Eighteeth Healthcare), generates less shear stress and, when paired with soft polymer tips, can prevent unnecessary dentin loss [13].

Ultrasonic agitation (Woodpecker UDS-K ultrasonic unit; Guilin Woodpecker Medical Instrument), operating at approximately 25,000 vibrations per second, enhances irrigant penetration into dentinal tubules through cavitation and acoustic streaming. Martin and Cunningham [14] highlighted its superior biological and chemical effects on root canal dentin surfaces.

The push-out bond strength test was chosen for its accuracy and reliability over shear and microtensile tests, especially when evaluating different root regions [12]. Adhesion of fiber posts to root canal dentin is complex because of anatomical and material-related challenges [4], with most failures occurring at the cement-dentin interface [15]. Previous studies have shown that 17% EDTA [16] and 5% calcium hypochlorite [17] significantly improve the bond strength of dual-cure resin–luted fiber posts.

In recent years, calcium hypochlorite ($[Ca(OCl)_2]$) has emerged as a promising alternative to sodium hypochlorite for post space irrigation because of its higher release of available chlorine, improved antimicrobial action, and formation of calcium-rich precipitates on dentin surfaces. These precipitates enhance mineral deposition and may improve the chemical interaction between dentin and resin cement.

Ferreira et al [12] reported that calcium hypochlorite modifies the dentin surface by increasing calcium content and reducing carbon content, which can favor adhesive bond formation. In addition, Sheikh Ghahderijani et al [17] demonstrated that 5% calcium hypochlorite significantly improved push-out bond strength when used as a post space irrigant, making it a viable option for enhancing adhesion between fiber posts and root dentin.

Although several studies have evaluated the effect of different irrigants and activation methods on smear layer removal and post retention, the available evidence remains inconsistent. Some investigations have reported superior cleaning efficacy and bond strength with ultrasonic activation, while others found no statistically significant difference compared with sonic activation. These discrepancies may be attributed to variations in study design, irrigant concentration, activation duration, and tooth type. Moreover, many previous studies have focused on single irrigants or have not used standardized comparative designs, making it difficult to draw definitive clinical conclusions.

XSL•FO
RenderX

This study uses the push-out bond strength test, which has advantages over other laboratory methods, such as shear bond strength and microtensile testing. The push-out test applies a uniform stress along the post-dentin interface, more closely simulating the dislodgement forces experienced clinically. It also allows evaluation across different root regions, providing more clinically relevant data on the performance of irrigant-activation combinations. By addressing these methodological gaps, this study aims to provide a direct comparison of sonic and ultrasonic activation across multiple irrigants under standardized conditions.

However, the effects of sonic and ultrasonic activation of these irrigants have not been compared extensively in research. Therefore, to help increase the durability and efficacy of endodontic treatments, this study aims to assess and contrast the effects of sonic and ultrasonic activation of 5% calcium hypochlorite, 17% EDTA, and saline on the push-out bond strength of fiber posts luted with dual-cure resin cement.

Research Question

This study was designed to evaluate the influence of different irrigant activation techniques on the adhesion of fiber posts to root canal dentin. The research question was as follows: Will the use of sonic and ultrasonic activation of 3 different post space irrigating solutions, namely, 5% calcium hypochlorite, 17% EDTA, and saline, affect the push-out bond strength of dual-cure resin–luted fiber posts in extracted human mandibular single-rooted premolars?

Aim

The aim of this study is to compare the mean push-out bond strength (in megapascals [MPa]) of fiber posts luted with dual-cure resin cement after post space irrigation with 5% calcium hypochlorite, 17% EDTA, or saline when activated using either sonic or ultrasonic techniques in extracted human mandibular single-rooted premolars.

Objectives

The objectives of this study are as follows:

First, to measure and compare the push-out bond strength (in MPa) of fiber posts following post space irrigation with 5% calcium hypochlorite, 17% EDTA, or saline, using sonic activation.

Second, to measure and compare the push-out bond strength (in MPa) of fiber posts following post space irrigation with 5% calcium hypochlorite, 17% EDTA, or saline, using ultrasonic activation.

Third, to determine what type of irrigant, activation method, or combination produces the highest mean push-out bond strength (in MPa).

Null Hypothesis

The null hypothesis is that there is no statistically significant difference in mean push-out bond strength (in MPa) among any of the irrigant-activation combinations tested.

Alternative Hypothesis

The alternative hypothesis is that at least 1 irrigant-activation combination produces a statistically significant difference in mean push-out bond strength (in MPa) compared with the others.

Methods

Eligibility Criteria

The inclusion and exclusion criteria are presented in Textbox 1.

Textbox 1. Inclusion and exclusion criteria of the research participants.

Inclusion criteria

- Recently extracted permanent mandibular lower premolar teeth with a closed apex
- Teeth free of cavities, fractures, and previous endodontic treatment

Exclusion criteria

- Teeth with unusual shapes or visible cracks
- Teeth with abnormalities in their development
- Teeth with external and internal resorption

Sample Size Estimation

The sample size was estimated with the following formula:

$$n1=\{(\sigma 1^2+\sigma_2{}^2/\kappa)\ (Z_{1-\alpha/2}+Z_{1-\beta})^2\} \div \Delta 2$$

Where n_1=sample size of group 1; n_2=sample size of group 2; σ_1=SD of group 1; σ_2=SD of group 2; Δ=difference in group means; κ=ratio of n_2 to n_1 (n_2/n_1); $Z_{1-\alpha/2}$=2-sided *Z* value (eg, *Z*=1.96 for 95% CI); and $Z_{1-\beta}$=0.84 (80% power).

σ=pooled SD, calculated as follows:

$$\sigma = Ö(\sigma_1{}^2 + \sigma_2)/2=Ö[(10.19)^2 + (10.04)^2]/2 = 10.12$$

Where Δ=expected mean difference in push-out bond strength (49.08 − 39.82 = 9.2649.08 – 39.82 = 9.2649.08 − 39.82 = 9.26 MPa).

Substituting values into the equation, we obtain the following:

$$n = [2(1.96 = 0.84)^2\ (10.12)^2]\ 9.26\ ^2 = [2 \times (2.80)^2 \times 102.41] \div 85.71 \approx 18.7.$$

This equals 19 specimens per group. To account for a potential 10% specimen loss, the sample size was increased to 20 per group. With 6 groups (3 irrigants × 2 activation methods), the

XSL•FO
RenderX

total sample size is 120 specimens. Furthermore, if mean bond strength in the EDTA group=49.08, mean bond strength in the normal saline group=39.82, σ1=SD of bond strength in the EDTA group=10.19, and σ2=SD of bond strength in the normal saline group=10.04, for detecting a mean difference of 9.26, that is, Δ =49.08 − 39.82=9.26, we need K = 1, N = $(17.33 \times 17.33 + 2.81 \times 2.81)(1.96 + 0.84)^2$, 23 × 23 = 18.71 = 19 patients in each group.

On the basis of previously published data [16], the mean push-out bond strength for the EDTA group should be 49.08 MPa (SD 10.19) and for the normal saline group it should be 39.82 MPa (SD 10.04). Using these values, a power analysis was performed assuming a 5% significance level (α=.05), 80% power (β=.20), and a mean difference of 9.26 MPa between groups.

Group Allocation

The specimens were randomly allocated into 3 main groups based on the post space irrigant used, and further, each group was subdivided into 2 subgroups according to the activation method, as shown in Table 1.

Table 1. Sample size distribution.

Irrigant and activation method	Sample size, n
Group A: normal saline	
Subgroup A1: ultrasonic activation	20
Subgroup A2: sonic activation	20
Group B: 5% calcium hypochlorite	
Subgroup B1: ultrasonic activation	20
Subgroup B2: sonic activation	20
Group C: 17% EDTA[a]	
Subgroup C1: ultrasonic activation	20
Subgroup C2: with sonic activation	20

[a]EDTA: ethylenediaminetetraacetic acid.

Recruitment and Procedure

For this in vitro investigation, a total of 120 healthy, recently excised human mandibular premolars of comparable size will be chosen. The teeth will be stored in 0.1% thymol solution at 4 °C until use to prevent microbial growth and preserve dentin integrity. Each specimen will be standardized to a 14-mm root length after the crowns are removed with a water-cooled diamond disk.

To verify working length and canal patency, a #10 K-file (Mani Inc) will be used. Using the ProTaper Universal rotary system (Dentsply) up to the F3 finishing file, canal preparation will adhere to the crown-down procedure. Following each filing, canals will be irrigated with 2 mL of distilled water and 5% sodium hypochlorite during instrumentation. Upon completion of instrumentation, a final flush will be given using 5 mL of 17% EDTA.

The single-cone obturation technique will be used with Diadent Dia-ProSeal sealer. The coronal 10 mm of the canal will be ready for fiber post implantation, and an apical 4 mm of gutta-percha (Dentsply Sirona) will be kept in the canal to ensure a sufficient apical seal. The coronal gutta-percha will be extracted using Peeso reamers (Mani Inc; 28 mm).

Following that, specimens will be randomly split into 3 major groups according to the type of post space irrigant, as described in Table 1. Sodium hypochlorite will be used during instrumentation for its established cleaning properties, whereas calcium hypochlorite will be tested as a post space irrigant to assess its effect on bond strength during post cementation.

Sonic activation will be done using the SuperEndo Activator, operating at frequencies between 1 and 6 kHz and fitted with flexible polymer tips appropriate for post space irrigation. Ultrasonic activation will be done using the Woodpecker UDS-K ultrasonic unit, operating at a frequency of approximately 25 to 30 kHz, with a size #20 noncutting tapered ultrasonic file.

The sonic activation frequency (1-6 kHz) was chosen according to the manufacturer's specifications for the SuperEndo Activator device and previous studies showing the effectiveness of this range in agitating irrigants within post spaces. The ultrasonic activation frequency (approximately 25-30 kHz) reflects the operational parameters of the Woodpecker UDS-K ultrasonic unit, supported by literature reporting optimal cavitation and acoustic streaming effects at these frequencies.

Each post space will be irrigated with 5 mL of the allotted irrigant for 60 seconds. This activation time of 60 seconds was selected based on findings from Saber et al [16], Rödig et al [13], and Haupt et al [18], which demonstrated that this duration effectively removes the smear layer and enhances irrigant penetration without causing dentin erosion or excessive structural alteration. After irrigation, distilled water will be used to rinse the canals, and absorbent paper points will be used to dry them [16,18].

In compliance with the manufacturer's instructions, fiber posts (Dentsply RelyX Fiber Posts, size #2 [diameter 1.5 mm and length 20 mm]) will be luted using a dual-cure resin cement (Dentsply Calibra Universal). Standardization of canal diameter and post space preparation to a uniform size (#2 post, 1.5-mm diameter) will be implemented to control for variability in post

XSL•FO
RenderX

fit and bonded surface area, which could otherwise confound bond strength measurements. An etch-and-rinse technique will be used. The cement will be mixed and applied using a lentulo spiral. The post will be inserted with firm finger pressure for 40 seconds. Excess cement will be removed, followed by light curing for 40 seconds per side.

Three horizontal slices, each 1 mm thick, will be cut from each sample's midroot region. The use of 1-mm–thick slices for push-out testing follows recommendations by Ferreira et al [12], as this thickness minimizes bending stresses during testing and provides more accurate measurement of interfacial bond strength. Each section will undergo mechanical testing after being checked for consistency. Universal testing equipment will be used to assess the strength of the push-out bond at a speed of 0.5 mm per minute, wherein a 1-mm–diameter metal pin will be positioned to supply force in an apical-to-coronal direction.

The maximum load at which the post dislodges must be recorded in newtons (N). The bond strength in MPa is calculated by dividing the recorded load (N) by the bonded interface area (A). The bonded interface area will be calculated using the following formula: $A = \pi(R + r) \sqrt{[h^2 + (R - r)^2]}$, where R=coronal radius, r=apical radius, and h=slice thickness.

Blinding

To minimize operator and measurement bias, specimen preparation, irrigation, and bond strength testing will be performed by different trained investigators who are blinded to the group allocations. The group codes will be concealed during mechanical testing and data entry.

Interexaminer and Intraexaminer Reliability

To ensure measurement consistency and validity, 10% of the samples will be randomly selected and retested by the same examiner and by a second examiner. The intraexaminer and interexaminer reliability will be assessed using intraclass correlation coefficients, with values >0.80 considered acceptable.

Ethical Considerations

Ethics approval for this study was obtained from the institutional ethics committee of Datta Meghe Institute of Higher Education and Research (IEC/2025/543).

All extracted teeth were obtained from the department of oral surgery as biological waste tissue following routine extractions performed for therapeutic reasons unrelated to this study. The specimens were anonymized before collection, with no patient identifiers recorded, and were stored in accordance with institutional biosafety protocols. As the specimens constituted waste tissue, individual informed consent was waived by the ethics committee; however, all procedures complied with national and institutional guidelines for the ethical handling and disposal of human-derived materials.

Statistical Analysis

Statistical analysis will use both descriptive and inferential statistical methods, including the chi-square test, Student paired and unpaired *t* tests, 1-way ANOVA, and Tukey test. The analysis will be conducted using SPSS (version 27.0; IBM Corp) and GraphPad Prism (version 7.0; GraphPad Software Inc), with a significance level of $P<.05$. Descriptive statistics (including mean, SD, and 95% CI) will be calculated for each group. Differences in mean push-out bond strength (in MPa) among the 6 groups will be assessed using 1-way ANOVA. Post hoc comparisons will be performed using the Tukey honestly significant difference test to control for type I errors associated with multiple comparisons. A significance level of $P<.05$ will be considered statistically significant.

Outcome Measure

Primary Outcome Measure

The primary outcome of this study is to compare the effects of sonic and ultrasonic activation of 3 post space irrigants—5% calcium hypochlorite, 17% EDTA, and saline—on the push-out bond strength of fiber posts luted with dual-cure resin cement in extracted human mandibular single-rooted premolars.

Secondary Outcome Measures

The secondary outcomes of this study are as follows:

- To assess how the push-out bond strength of fiber posts luted with dual-cure resin cement is affected by the sonic activation of normal saline, 5% calcium hypochlorite, and 17% EDTA.
- To assess how the push-out bond strength of fiber posts luted with dual-cure resin cement is affected by the ultrasonic activation of normal saline, 5% calcium hypochlorite, and 17% EDTA.
- To contrast the push-out bond strength attained for each irrigant using sonic and ultrasonic activation.
- To ascertain which combination of irrigant and activation offers the strongest push-out bond.

Results

At the time of submission, ethics approval has been obtained (on February 7, 2025), and the study is self-funded. Sample collection commenced in June 2025, with 68 extracted mandibular premolars obtained to date. Specimen preparation is underway, and data collection is expected to be completed by March 2026. Data analysis is planned for November 2025, with results anticipated to be available for publication in early 2026.

Discussion

Anticipated Findings

In this in vitro study, the effects of sonic and ultrasonic activation of 3 post space irrigants—5% calcium hypochlorite, 17% EDTA, and saline—on the push-out bond strength of fiber posts luted with dual-cure resin cement will be investigated. The findings may indicate that both activation method and irrigant type affect adhesion to root canal dentin. It is hypothesized that the chelating action of EDTA may facilitate better bonding by efficiently eliminating the smear layer and revealing dentinal tubules [19]. If this is observed, such a finding would be consistent with the findings of Saber et al [16], who

XSL•FO
RenderX

reported that EDTA significantly enhanced bonding strength compared with saline.

Additionally, 5% calcium hypochlorite is expected to strengthen the connection. This might be because calcium-containing crystals develop, which enhance chemical reactivity and promote bonding [17]. Saline, which has no chemical reactivity, is expected to produce the lowest bond strength [17].

An important factor is irrigant activation. On the basis of previous reports, ultrasonic activation is expected to produce a stronger bond than sonic activation due to enhanced cavitation and acoustic streaming. This is explained by ultrasonic cavitation and sonic streaming effects, which improve irrigant penetration and smear layer removal [16]. Despite being less potent, sonic activation is expected to outperform nonactivated irrigation and, because of its conservative nature, might be better in some therapeutic situations. This would be consistent with the findings of Marchese et al [19], who observed that EDTA was superior to other irrigants for post space cleaning.

To maximize post retention, this study will highlight the significance of using efficient irrigants and applying ultrasonic stimulation. In teeth that have received endodontic treatment, such procedures may lessen debonding and enhance clinical results.

The results might not accurately reflect intraoral circumstances because this study was conducted in vitro. Long-term aging, fatigue loading, and advanced imaging should all be included in future research to assess the interfacial properties and durability of the bonded post-dentin complex. While statistically significant differences in MPa values may be observed, their clinical significance should be interpreted with caution. Literature suggests that even modest increases in bond strength can enhance post retention and reduce debonding, potentially improving clinical longevity. However, long-term survival data will be required to validate these effects in vivo.

Limitations

Although this study will provide controlled data on push-out bond strength, it is limited by the absence of aging simulations, such as thermocycling or mechanical fatigue loading. Future studies should incorporate cyclic loading and long-term water storage to better replicate intraoral conditions and evaluate the durability of the bonded interface over time.

A nonactivated control group for each irrigant was not included, as the focus was to compare activation techniques. Future studies may benefit from incorporating control groups to assess absolute efficacy compared with baseline irrigation. No failure mode analysis (cohesive, adhesive, or mixed) was performed; this has been proposed as a secondary analysis for future studies.

The study did not use scanning electron microscopy or confocal laser scanning microscopy to confirm smear layer removal, which limits direct correlation with bond strength outcomes. Such analyses are planned for inclusion in the secondary phase of the study. While calcium hypochlorite adds novelty, the exclusion of more recent irrigants, such as maleic acid or nanoparticle-based solutions, limits broader clinical relevance. These agents will be considered in subsequent research phases to enhance clinical relevance.

Acknowledgments

The authors acknowledge the generous technical support provided by the research facility and expresses gratitude to their guide for assistance with the project. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Authors' Contributions

The study was conceptualized and registered by AW and AP. The study design and preliminary paper preparation were completed by AP. AW completed the final study design and manuscript writing. Each author made an equal contribution to the work and reviewed the completed paper.

Conflicts of Interest

None declared.

References

1. Bateman G, Ricketts DN, Saunders WP. Fibre-based post systems: a review. Br Dent J 2003 Jul 12;195(1):43-8; discussion 37. [doi: 10.1038/sj.bdj.4810278] [Medline: 12856030]
2. Robbins JW. Guidelines for the restoration of endodontically treated teeth. J Am Dent Assoc 1990 May;120(5):558, 560, 562 passim. [doi: 10.14219/jada.archive.1990.0087] [Medline: 2186075]
3. Monteiro LC, Pecorari VG, Gontijo IG, Marchi GM, Lima DA, Aguiar FH. PEEK and fiberglass intra-radicular posts: influence of resin cement and mechanical cycling on push-out bond strength. Clin Oral Investig 2022 Dec 26;26(12):6907-6916. [doi: 10.1007/s00784-022-04645-0] [Medline: 35881239]
4. Ferrari M, Cagidiaco M, Grandini S, De Sanctis M, Goracci C. Post placement affects survival of endodontically treated premolars. J Dent Res 2007 Aug 01;86(8):729-734. [doi: 10.1177/154405910708600808] [Medline: 17652200]

XSL•FO
RenderX

5. Perdigão J, Gomes G, Augusto V. The effect of dowel space on the bond strengths of fiber posts. J Prosthodont 2007;16(3):154-164. [doi: 10.1111/j.1532-849X.2006.00166.x] [Medline: 17581176]
6. Okino LA, Siqueira EL, Santos M, Bombana AC, Figueiredo JA. Dissolution of pulp tissue by aqueous solution of chlorhexidine digluconate and chlorhexidine digluconate gel. Int Endod J 2004 Jan 13;37(1):38-41. [doi: 10.1111/j.1365-2591.2004.00749.x] [Medline: 14718055]
7. Muniz L, Mathias P. The influence of sodium hypochlorite and root canal sealers on post retention in different dentin regions. Oper Dent 2005;30(4):533-539. [Medline: 16130876]
8. Naumann M, Blankenstein F, Kiessling S, Dietrich T. Risk factors for failure of glass fiber-reinforced composite post restorations: a prospective observational clinical study. Eur J Oral Sci 2005 Dec;113(6):519-524. [doi: 10.1111/j.1600-0722.2005.00257.x] [Medline: 16324143]
9. Sarkis-Onofre R, Skupien JA, Cenci MS, Moraes RR, Pereira-Cenci T. The role of resin cement on bond strength of glass-fiber posts luted into root canals: a systematic review and meta-analysis of in vitro studies. Oper Dent 2014;39(1):E31-E44. [doi: 10.2341/13-070-LIT] [Medline: 23937401]
10. Tang W, Wu Y, Smales RJ. Identifying and reducing risks for potential fractures in endodontically treated teeth. J Endod 2010 Apr;36(4):609-617. [doi: 10.1016/j.joen.2009.12.002] [Medline: 20307732]
11. Dogan Buzoglu H, Calt S, Gümüsderelioglu M. Evaluation of the surface free energy on root canal dentine walls treated with chelating agents and NaOCl. Int Endod J 2007 Jan;40(1):18-24. [doi: 10.1111/j.1365-2591.2006.01169.x] [Medline: 17209828]
12. Ferreira MB, Carlini Júnior B, Galafassi D, Gobbi DL. Calcium hypochlorite as a dentin deproteinization agent: microleakage, scanning electron microscopy and elemental analysis. Microsc Res Tech 2015 Aug 02;78(8):676-681. [doi: 10.1002/jemt.22524] [Medline: 26031607]
13. Rödig T, Koberg C, Baxter S, Konietschke F, Wiegand A, Rizk M. Micro-CT evaluation of sonically and ultrasonically activated irrigation on the removal of hard-tissue debris from isthmus-containing mesial root canal systems of mandibular molars. Int Endod J 2019 Aug 04;52(8):1173-1181. [doi: 10.1111/iej.13100] [Medline: 30773661]
14. Martin H, Cunningham W. Endosonics--the ultrasonic synergistic system of endodontics. Endod Dent Traumatol 1985 Dec;1(6):201-206. [doi: 10.1111/j.1600-9657.1985.tb00582.x] [Medline: 3867504]
15. Andreani Y, Gad BT, Cocks TC, Harrison J, Keresztes ME, Pomfret JK, et al. Comparison of irrigant activation devices and conventional needle irrigation on smear layer and debris removal in curved canals. (Smear layer removal from irrigant activation using SEM). Aust Endod J 2021 Aug;47(2):143-149. [doi: 10.1111/aej.12482] [Medline: 33682268]
16. Saber SED, Hashem AAR. Efficacy of different final irrigation activation techniques on smear layer removal. J Endod 2011 Sep;37(9):1272-1275. [doi: 10.1016/j.joen.2011.06.007] [Medline: 21846546]
17. Sheikh Ghahderijani M, Khoroushi M, Feiz A. Effect of calcium hypochlorite and sodium hypochlorite as root canal irrigants on push-out bond strength of fiber post cemented with etch-and-rinse resin cement: an in vitro study. Front Dent 2021 May 25;18:18 [FREE Full text] [doi: 10.18502/fid.v18i18.6325] [Medline: 35965718]
18. Haupt F, Meinel M, Gunawardana A, Hülsmann M. Effectiveness of different activated irrigation techniques on debris and smear layer removal from curved root canals: a evaluation. Aust Endodontic J 2019 Mar 25;46(1):40-46. [doi: 10.1111/aej.12342]
19. Marchese M, Denise P, Ferrari Cagidiaco E, Iandolo A, Grandini S, Ferrari M. Endodontic irrigants and their activation efficacy on cleansing post-space root canal walls. Prosthesis 2021 Dec 01;3(4):406-414. [doi: 10.3390/prosthesis3040036]

Abbreviations

EDTA: ethylenediaminetetraacetic acid
MPa: megapascals

Edited by J Sarvestan; submitted 04.May.2025; peer-reviewed by N Taymour, WG Escalante-Otarola; comments to author 27.May.2025; revised version received 15.Jul.2025; accepted 27.Aug.2025; published 28.Jan.2026.

Please cite as:
Wamane A, Patel A
Comparative Evaluation of Sonic and Ultrasonic Activation of 3 Post Space Irrigants on Push-Out Bond Strength of Fiber Posts Luted With Dual-Cure Resin Cement: Protocol for an In Vitro Study
JMIR Res Protoc 2026;15:e76907
URL: https://www.researchprotocols.org/2026/1/e76907
doi:10.2196/76907
PMID:

XSL•FO
RenderX

©Apurva Wamane, Aditya Patel. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 28.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Developing a Program Theory on Ventilator Weaning in Adult Intensive Care: Protocol for a Multimethods Study

Fritz Sterr[1,2], MSc; Lydia Bauernfeind[2], PhD; Christian Rester[2], Prof Dr; Sabine Metzing[1], Prof Dr; Rebecca Palm[1,3], Prof Dr

[1]School of Nursing Sciences, Faculty of Health, Witten/Herdecke University, Witten, Germany

[2]Faculty of Applied Healthcare Sciences, Deggendorf Institute of Technology, Deggendorf, Germany

[3]Department of Health Services Research, School VI Medicine and Health Sciences, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

Corresponding Author:
Fritz Sterr, MSc
School of Nursing Sciences
Faculty of Health
Witten/Herdecke University
Alfred-Herrhausen-Straße 50
Witten, 58455
Germany
Phone: 49 99136158344
Email: fritz.sterr@uni-wh.de

Abstract

Background: Worldwide, mechanical ventilation and ventilator weaning have been widely researched. Nevertheless, rates of weaning failure remain high. According to the Medical Research Council framework, ventilator weaning is a complex intervention. While there are various guidelines on this, there is no abstract theoretical understanding that organizes the interventions, outcomes, and their contexts.

Objective: This study aims to explore the interconnectedness of interventions, outcomes, and context in ventilator weaning of adult intensive care patients.

Methods: Using the approach of Funnell and Rogers, we develop a program theory for ventilator weaning in a multimethod study comprising 2 main steps. First, 3 literature reviews on interventions and outcomes, predictors of weaning failure, and patients' experiences were triangulated with stakeholder conversations. Using abduction, we then developed an initial program theory. Second, the initial theory will be revised in an iterative process. To this end, semistructured group discussions and workshops will be conducted, followed by a deductive thematic analysis and adaptation of our theory. This process will be repeated until stakeholder statements and data analyses are congruent with the program theory.

Results: The initial program theory developed in step 1 is presented in this protocol and serves as the basis for review and refinement in step 2. The results of this iterative process and the final program theory are expected in 2026.

Conclusions: Following the Medical Research Council framework, a program theory on ventilator weaning will be developed in this study. This may enable a differentiated understanding of ventilator weaning and more sustainable and comprehensive research. The program theory emphasizes the interdisciplinary nature of ventilator weaning and supports health care professionals in combining interventions appropriately and evaluating relevant outcomes.

Trial Registration: Open Science Framework YGJ3T; https://doi.org/10.17605/OSF.IO/YGJ3T

International Registered Report Identifier (IRRID): DERR1-10.2196/83342

(JMIR Res Protoc 2026;15:e83342) doi:10.2196/83342

KEYWORDS

artificial respiration; critical care; intensive care; mechanical ventilation; program theory; study protocol; ventilator weaning

XSL•FO
RenderX

Introduction

Background

Worldwide, an increasing number of patients receive mechanical ventilation (MV) in an intensive care unit (ICU) [1-6]. The indication for MV is usually due to acute respiratory insufficiency [7-11]. Other reasons are acute shock, coma, or planned surgical procedures [11,12]. While MV is an important and life-saving intervention in these situations, it also leads to several negative consequences.

It can cause endotracheal tube–related complications (eg, displacement, laryngeal injury, and tracheomalacia) [11], physical consequences (eg, cardiovascular and respiratory collapse, inflammation, ventilator-induced lung injury, peripheral organ injury, and impaired cognition) [11,13-15], and psychological consequences (eg, anxiety, stress, indifference, depression, and posttraumatic stress disorder) [16,17].

It is of central importance to initiate ventilator weaning at an early stage. To this end, a broad body of evidence is already available to guide this process. Various risk factors [18-20], interventions [21-24], and the patient's experience of weaning [25-27] have been empirically investigated, and the nursing diagnosis "dysfunctional adult ventilatory weaning response" has been developed [28,29]. These findings are summarized in various international guidelines [8,30,31].

Problem Definition

The Medical Research Council (MRC) framework defines a complex intervention according to several criteria. These include "the number of components, the range of behaviors targeted; expertise and skills required by those delivering and receiving the intervention; the number of groups, settings, or levels targeted; or the permitted level of flexibility" [32]. Following this understanding and the information in the Background section, ventilator weaning is to be understood as a complex intervention.

For a complex intervention to be developed and effective in the long term, it requires a theory-based understanding of which interventions interact, which outcomes they affect, and which context they depend on [32]. Despite the fact that MV and ventilator weaning have been investigated extensively, treatment in ICUs has not yet reached its optimum. To date, less than two-thirds of patients can be successfully weaned from the respirator [12,33,34]. In the Weaning According to a New Definition (WIND) study, 24.3% never entered the weaning process [33], and in the Worldwide Assessment of Separation of Patients From Ventilatory Assistance (WEAN SAFE) study, 28.3% died under MV [12]. While these findings can be explained by various reasons, more than 50% of the patients previously declared unweanable could still be weaned after their ICU stay [35]. This is a problem in that the longer durations of MV are associated with a higher mortality rate [33].

In addition, most studies only examine singular aspects of the weaning process. While this is of central importance for drawing clear conclusions about the effects of interventions and consequently making recommendations for or against them, it does not represent the real embedding of these interventions in complex intensive care [36]. As described by the MRC framework, it is essential to understand the interplay of interventions, outcomes, and their context. Therefore, reducing ventilator weaning to isolated aspects is a cause for concern. It is not a simple but a multimodal and multilayered process in which a large number of factors must be considered [20]. Thus, further research into weaning methods is repeatedly demanded, for example, in the context of patients with neurological conditions [37,38].

Two major research gaps emerge. On the one hand, most research studies only answer the question of "whether," but not "how and why." Nevertheless, these questions need to be answered as required by the MRC framework [32]. The limited perspective makes it impossible to understand the real influences and causal relationships of weaning interventions, outcomes, and their context. This finding is reinforced by the fragmentary nature of the studies conducted to date; there is a lack of evidence linking the interventions. The required opening of the black box does not take place at this point [39-42].

By contrast, there is a substantial deficit in the theoretical understanding of ventilator weaning. While no grand theory could be identified, only 1 middle-range theory on the nursing diagnosis "dysfunctional adult ventilatory weaning response" was found [29]. However, this theory only covers "weaning failure" and its antecedents and clinical consequents. In addition, only practical guidelines for clinicians appear at this stage that lack abstraction and theoretical conceptualization. A theory covering the entire weaning process in the ICU is still missing. This results in the versatile use of the term "ventilator weaning," which has changed in its classification, definition, and understanding over the years [33,43] and is the basis of an inconsistent, incoherent research logic [44]. This is another reason why an international consensus in MV research and the necessary theory development have recently been requested [45].

Aim and Research Questions

To conceptualize, develop, and organize ventilator weaning as a complex program in the long term, this study is conducted in line with the recommendations of the MRC framework. The central aim of this study is to explore the interconnectedness of individual interventions in ventilator weaning with relevant outcomes and their context.

Specifically, ventilator weaning as a program, the interconnection of interventions, and the associated factors should be described comprehensively. This is followed by an elaboration of the central mechanisms in ventilator weaning. Particularly, the focus is on the contribution of the identified interventions to changing the individual endpoints, which should also be formulated. On the basis of these aims, three research questions were formulated: (1) How is ventilator weaning composed in adult intensive care patients? (2) What results can be achieved through ventilator weaning in adult intensive care patients? and (3) How can the mechanisms in the process of ventilator weaning be theoretically summarized and comprehensively depicted?

XSL•FO
RenderX

Methods

Overview

To answer the research questions, a program theory is to be developed in line with the method of Funnell and Rogers [46]. The development of a program theory is the method of choice in this project, as it is explicitly proposed as a core element for complex interventions by the MRC framework [32] and can find answers to the central questions of this study. This study is oriented toward the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guideline [47,48] and the relevant aspects herein, as there are currently no reporting guidelines for qualitative studies or program theory development. The study was registered in the Open Science Framework (YGJ3T) on May 14, 2025.

Philosophical Background of Program Theories

The ontological and epistemological classification of program theories is a complex process that requires various considerations. A current and differentiated discussion assigns program theories to ontological realism [49]. This assumes that real structures and processes exist and work on their own, regardless of what we know or experience. In addition, a clear distinction is to be made between the transitive domain (including social structures and cause-and-outcome relationships that can change over time) and the intransitive domain (understanding and interpreting reality at a deeper level).

Furthermore, this study [49] allocates program theories to epistemological relativism. This means that knowledge and truth are relative to specific cultural or conceptual frameworks or settings. They note that researchers need to consider "the multi-causal nature of social structures and grasp their complexity through causal mechanisms leading to outcomes" [49].

In this regard, program theories are mainly understood as technological theories that are assigned to evaluation research and find themselves in theory-driven or theory-based approaches [46,50,51].

Methodological Framework

Overview

This study builds on the methodological guidance of Funnell and Rogers [46]. To develop a program theory, they recommend an iterative process in which various data sources are included and triangulated with each other. In detail, they refer to the combination of three approaches: (1) deductive development (eg, existing literature and reviews), (2) inductive development (eg, group discussions and interviews), and (3) articulating program stakeholders' mental models (eg, stakeholder conversations and workshops) [46].

Following these recommendations, the approach of this study is structured along two central steps: (1) the development of an initial program theory and (2) the review and revision of the program theory. The first step is already complete, and the results obtained will be presented in this study together with the planned revision of the theory.

Figure 1 illustrates the overarching methodological approach of the study. It depicts the 2 phases and the interconnections of the various steps in this approach. These are described in detail subsequently.

Figure 1. Methodological approach to developing the program theory.

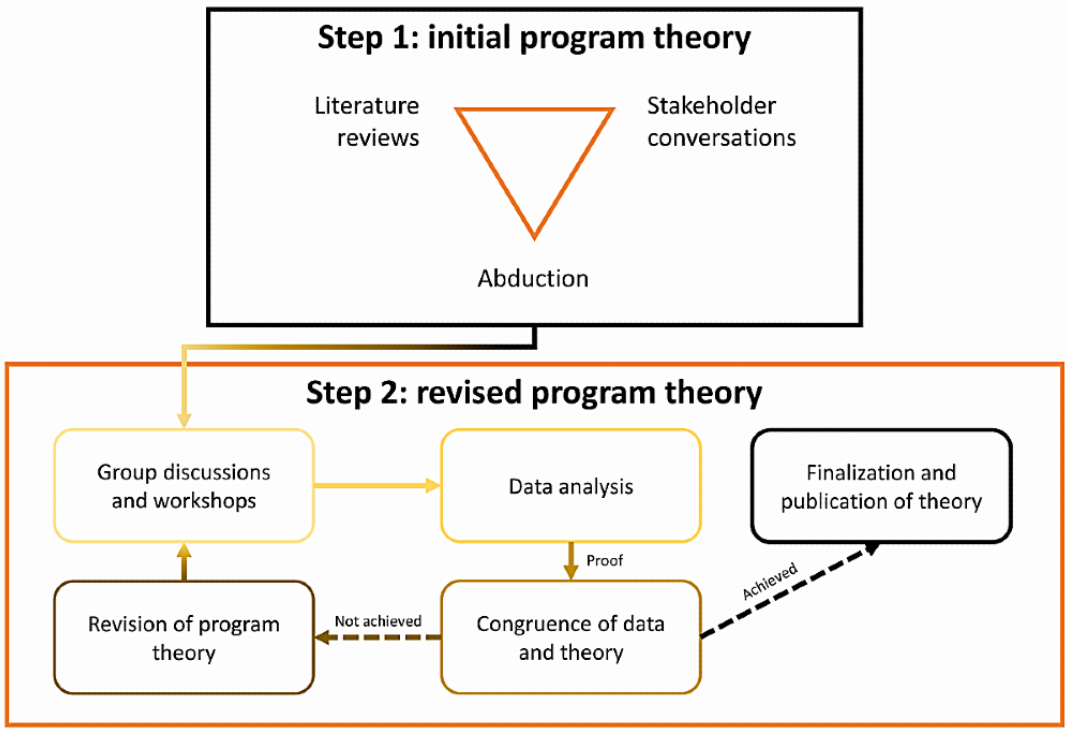

XSL•FO
RenderX

Step 1: Developing the Initial Program Theory

According to the underlying methodology [46], a program theory consists of 2 core components: the "theory of action" (ToA; related to research questions 1 and 3) and the "theory of change" (ToC; related to research questions 2 and 3). In addition, there is usually a logic model that narratively and/or graphically depicts how the ToA affects the ToC and induces the changes.

In detail, the ToC "refers to the central mechanism by which change comes about for individuals, groups, and communities" [46]. It consists of a comprehensive situation analysis, the focus and scope of the program, and the outcomes chain as "the centerpiece of the program theory" [46]. The ToA "explains how programs or other interventions are constructed to activate their theory of change" [46]. It contains desired attributes of outcomes, program features, and external factors affecting the outcomes and the activities of the program and external factors [46].

To develop a program theory, Funnell and Rogers [46] recommend starting with a comprehensive situation analysis. Researchers should identify what is already known about the program, what evidence base exists, and what still needs to be researched, and, above all, they are asked for "the nature and extent of the main problem that the program addresses" [46]. This step is followed by focusing and shaping the planned program and theory, setting clear boundaries. Subsequently, key outcomes and their attributes are derived, the interventions are identified, and, finally, the interrelationships in this program are reflected upon. The ToA and ToC can then be developed from this. In this study, the development of the initial program theory comprised the triangulation of data from (1) literature analysis, (2) stakeholder conversations, and (3) an abductive process.

As MV is widely researched, 3 systematic reviews were conducted in the literature phase. In a scoping review, interventions associated with ventilator weaning and related patient outcomes were identified [52]. In the next step, predictors of weaning failure were compiled in a systematic evidence map as observational criteria for the program theory [53]. Finally, the perspective of those affected was also examined. For this purpose, a meta-synthesis was carried out to gain an in-depth understanding of patients' experiences under MV [54].

Following the completion of the initial literature phase, 13 stakeholder conversations were held with various health care professionals (HCPs) between January 2025 and April 2025. These conversations were conducted using a semistructured interview guide and documented with field notes. According to the recommendations of Funnell and Rogers [46], these conversations were intended for stakeholders to articulate their initial mental models. In addition to a nurse and a physiotherapist, 2 physicians (an anesthesiologist and an intensivist), 2 critical care advanced practice nurses, 3 senior nursing researchers, and a paramedic were interviewed. The interview guide (Multimedia Appendix 1) included questions on the causality and indication of MV, its aim and associated problems, the initiation of weaning, relevant outcomes and their attributes in the process, key interventions, weaning failure, and influencing factors from outside. The triangulation of data from the literature phase and the stakeholder conversations was carried out in MAXQDA (version 24; VERBI Software GmbH), followed by a deductive analysis according to the key terminology of Funnell and Rogers [46], and resulted in an initial program theory.

In addition, the initial theory was complemented and expanded by the first author through an abductive process. This is defined as a subconscious process that occurs involuntarily and spontaneously as the only creative step in the entire research process. The abduction itself cannot be justified; only its results can be justified [55]. Therefore, we have identified additional connections in the intervention and outcome chain that are integrated into the initial program theory. It is important to note that this process alone may not result in accurate statements [56]. Therefore, results from the abductive process are embedded in a multimethod approach, and statements derived are critically examined in the revision of our theory.

To discuss the overall comprehensibility and consistency of the initial theory development, 2 social scientists and a sociologist were consulted, who are familiar with theory development but have no connection to intensive care medicine.

Step 2: Revision of the Program Theory

To evaluate and revise the initial theory, an iterative process is recommended that draws, in particular, on stakeholders and practitioners [46]. We have designed a cyclical process for this purpose, which is illustrated in Figure 1.

We will hold semistructured group discussions [57,58] and workshops [46] with various HCPs in 4-hour sessions every 3 to 4 weeks. These will focus on outcomes, interventions, the context, and their interrelationships. The interview guides for the group discussions and the workshops are provided in Multimedia Appendix 1. Each session will be audio recorded, transcribed verbatim, and then undergo a reflective thematic analysis based on a deductive approach [59-61].

Subsequently, the current program theory and the collected data are assessed for congruence. If no congruence is achieved, the program theory will be adapted to the data collected and reviewed again in group discussions and workshops with other study participants. This iterative process will be repeated until congruence of data and program theory has been achieved. If there is congruence, the program theory will then be finalized and published.

Study Setting

The study will be conducted at the Deggendorf Institute of Technology in Bavaria, Germany. Participants will meet with colleagues from other institutions independently of their own institution. This will create a space of openness and mutual inspiration.

Eligibility Criteria

People who have completed training in a health care profession (eg, medicine, nursing, physiotherapy, occupational therapy, speech therapy, or respiratory therapy), are aged 18 years or older, and have 12 months of experience in treating ICU patients receiving MV can participate in the study. In addition, they

XSL•FO
RenderX

must participate voluntarily in the study and be able to give informed consent.

Sample Size

To include different professions and perspectives and allow each participant to speak sufficiently, the sample size for each session is of crucial importance. Therefore, we recruit 6 to 10 people per group discussion so that not too many people are involved at the same time, and interaction between participants, follow-up questions, and an in-depth exploration of the subject matter are still possible.

Recruitment

Purposive sampling was used in this study for initial stakeholder conversations in step 1. To this end, 1 or 2 representatives from each relevant health care profession were requested for a conversation by the first author. Participants had to be experts in their field and hold at least a master's degree and have 5 years of professional experience. As this was an informal conversation and only field notes were taken, but no personal data were recorded, the participants had to declare their voluntary participation verbally; however, they did not have to sign an informed consent.

In step 2, purposive sampling will be used for the revision of the program theory. To this end, potential participants are informed about the study via various channels. They can then voluntarily contact the study coordinator if they wish to participate. As a first step, the professional networks of the study authors are used, and the information regarding the study is disseminated there. When potential participants respond, they are assessed by the researchers for their suitability. Following the ontological realism on which program theory is based, we aim to represent the actual staffing composition in ICUs as realistically as possible (unlike in step 1). Therefore, in the assessment, we look for a broad distribution of skills and degrees, tasks and roles, as well as professional experience. If participation is too low, the relevant professional and scientific societies in Germany will be contacted in the second step, and information regarding this study will be spread within them.

Researcher Reflexivity

Reflecting on the role and previous knowledge of the researchers themselves is of great importance in empirical studies [62]. All 5 researchers involved in this study are registered nurses and have at least several years of professional experience in the ICU. They all have actively cared for MV patients themselves and accompanied them through the process of ventilator weaning. Currently, only 1 author (LB) still works part time in the ICU. All 5 authors work at different universities—3 (CR, RP, and S Metzing) as professors and 2 (LB and FS) as research associates. One author is a specialist in health services research and the development of program theories (RP), and another is a specialist in qualitative research methods (S Metzing). One author deals primarily with professional development and nursing diagnostics (CR), while the other 2 authors mainly conduct research in acute and intensive care (LB and FS). Finally, 4 authors have a doctorate in nursing (CR, LB, RP, and S Metzing), and 1 author is currently working on his doctorate in this study (FS).

Ensuring the Quality of This Study

Compliance with quality criteria is central to research. On the basis of the primarily qualitative research design of this study, the relevant quality criteria must be identified. Nevertheless, "a single and specific set of quality criteria is neither feasible nor anticipated" [63] due to the diverse paradigms in qualitative research. Therefore, researchers must reflect on their own project and subject matter and determine suitable criteria along various aspects [63].

On the one hand, the rigor and quality of the study relate to the 3 central aspects of ontology, epistemology, and methodology [64]. In contrast, it also involves "the steps of designing, conducting, and reporting qualitative research in a step-wise approach" [65].

Therefore, in this study, we are guided by 4 criteria to enrich our own trustworthiness [63]: credibility, transferability, dependability, and confirmability. These can essentially be achieved through reflexivity, peer involvement, triangulation of methods and data, prolonged engagement, persistent observation, maximum variation, typical sampling, respondent validation, data collection until saturation, and thick description [63]. Information on all these criteria can be found throughout this study protocol. In addition, we will consider the quality criteria in the subsequent publications on the results of this study.

Therefore, we will adhere to established guidelines when reporting the results. As there are currently no specific reporting guidelines for the development of program theories, we will follow the "Standards for Reporting Qualitative Research" [66] due to the qualitative focus of this study.

Ethical Considerations

This study follows the ethical principles of the Declaration of Helsinki [67]. In this regard, an ethics application was submitted to the ethics committee of Witten/Herdecke University to ensure that the study was legally and ethically sound. We received a positive ethics vote from the committee on April 28, 2025 (S-109/2025).

As third parties are involved in this research project, and participants' and personal data are collected, a data privacy management process was carried out at the university conducting the project and approved in March 2025. The contents are publicly accessible to the participants via the Deggendorf Institute of Technology website. It was agreed with the data protection officers that only the personal data that are absolutely necessary for the research project will be collected.

The ethics application and data protection management stipulated that all HCPs must sign an informed consent before the study, confirming their voluntary participation. The agreement also assured them that, once the study had been completed, all personal information would be immediately pseudonymized by the researchers and anonymized for third parties. After completion of the study, but no later than December 2026, all personal information and data collected will be irrevocably deleted. During the study and the analysis, participants will have the opportunity to be informed about the

XSL•FO
RenderX

processing of their data at any time. In addition, they can withdraw from the study until the results are published and retract their statements without giving reasons.

HCPs will not receive compensation for travel expenses or lost work time for their participation. However, catering will be provided on-site during the workshop days.

Dissemination Policy

We plan to publish 2 open-access and peer-reviewed articles on the results. One article will contain the core results of the group discussions and workshops as a separate, empirical report. The second article will include the revised and finalized program theory developed in this study. We will also present the results at national and international scientific conferences. Finally, we intend to actively approach people in clinical care and disseminate the results to the ICUs there.

Results

The initial program theory, which was developed in the first step of this study from March 2023 to May 2025, will be presented subsequently. This will serve as the basis for the empirical review, revision, and refinement conducted in the second step of this study.

Defining the Core Problem

This program addresses ventilator dependency (Figure 2). While the starting point in this case is ongoing spontaneous breathing, reasons arise for an intubation or cannulation of patients. These causes are linked to a pronounced indication for MV. In this first period, patients are dependent on respiratory support to some degree, but the dependency increases along with the duration of MV, substance use, and the loss of muscles. Therefore, the core problem is not the ventilation itself, but the iatrogenic dependence on the ventilator.

Figure 2. The phenomenon of ventilator dependency.

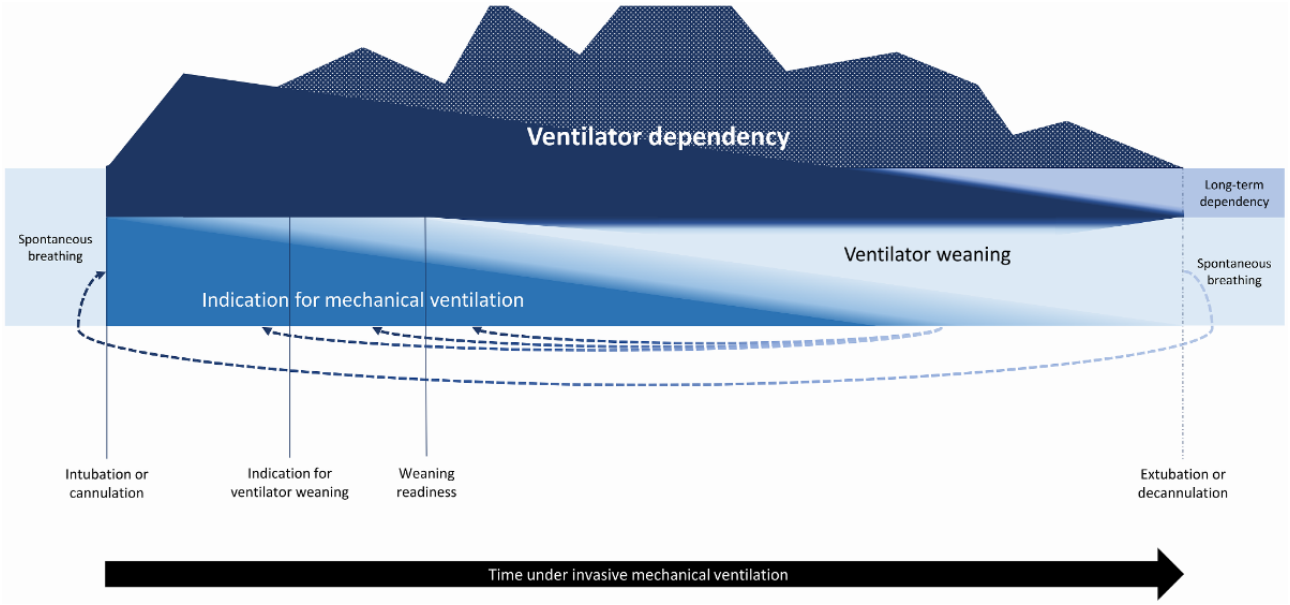

While the indication for MV decreases, the importance of required efforts for ventilator weaning increase over time. Their transition is fluid, and the processes can run in parallel. Nevertheless, an indication for ventilator weaning must be given initially, and weaning readiness must then be achieved so that the weaning process can be initiated.

In the meantime, the ventilator dependency ideally decreases continuously but can also change in the course of time. During the weaning process, patients may also fall behind in their progress, no longer be ready for weaning, or have their indication for weaning withdrawn.

Finally, patients are ideally no longer dependent on the ventilator, are extubated, and can breathe spontaneously again. However, ventilator dependency can also manifest itself and turn into long-term dependency. Ultimately, patients can also experience weaning failure after extubation and undergo reintubation. At this point, the process starts all over again.

Focusing and Scoping of the Program

The main interest of this program theory lies in facilitating ventilator liberation. Extubation or decannulation in conjunction with sufficient spontaneous breathing is its key goal. The setting of interest is the ICU in acute inpatient care. The main actors in this regard are HCPs who are responsible for the treatment and ventilator weaning of the affected patients (eg, intensivists, critical care nurses, and therapists). In addition, not only the patients themselves and their relatives and friends but also pastoral workers, cleaning staff, and hospital management play an important role.

Key Outcomes in Ventilator Weaning

Before the initiation of the program, preconditions need to be fulfilled (Figure 3). Patients need to be ready for a ventilator weaning and have no signs of possible weaning failure. In addition, there must be an indication for ventilator weaning.

XSL•FO
RenderX

Figure 3. Key outcomes of ventilator weaning. MV: mechanical ventilation.

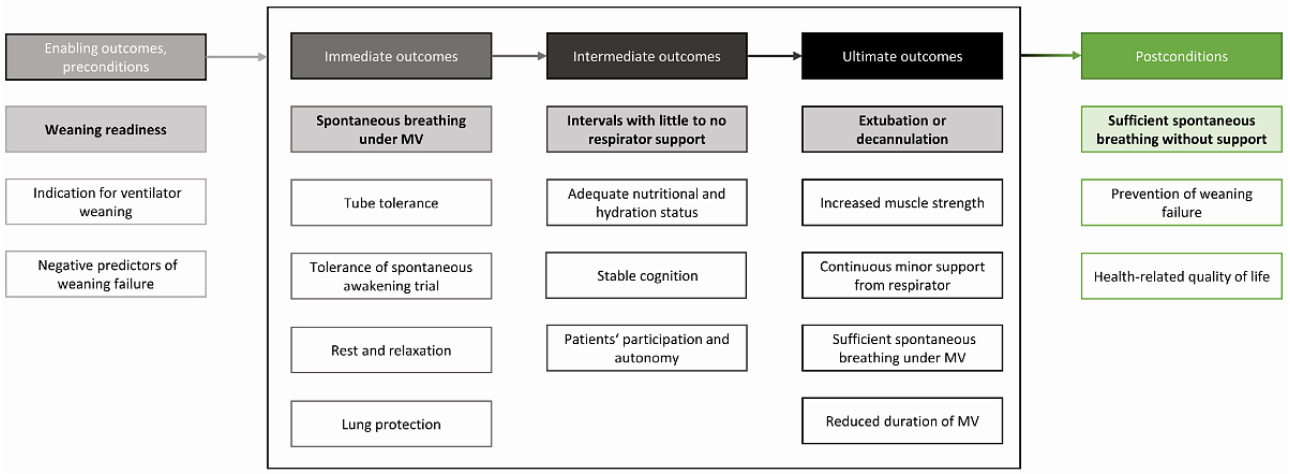

During ventilator weaning, immediate, intermediate, and ultimate outcomes manifest themselves consecutively. These should not only be achieved at the respective point in time but should also be continuously maintained from then on. The outcomes build on each other and are a prerequisite for the next level. For instance, patients must be able to breathe spontaneously under MV in the short term so that the duration of spontaneous breathing can become longer in the medium term, and they can be extubated in the long term.

After extubation and completion of the program, postconditions that are needed to maintain the patient's stable situation become relevant. This includes sufficient spontaneous breathing without support, the prevention of weaning failure, and a preserved health-related quality of life.

Key Interventions in Ventilator Weaning

This program incorporates primary interventions—such as spontaneous breathing trial, muscle training, or the reduction of MV support—that have a direct impact on the outcomes (Figure 4). The secondary interventions have an indirect impact on the outcomes. For instance, nutrition, promotion of perception, or pain management enable spontaneous breathing trials or muscle training and thus have an influence on the key outcomes but do not have a direct effect on the latter.

Figure 4. Key interventions of ventilator weaning. MV: mechanical ventilation.

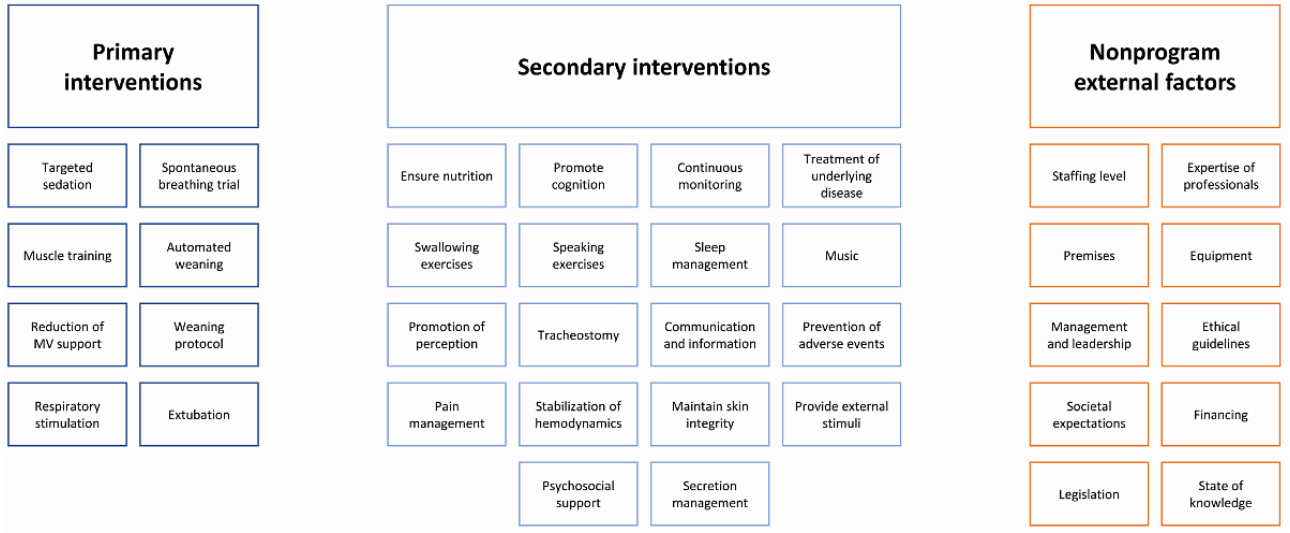

In addition to these interventions, nonprogram external factors also affect the outcomes but are beyond the control of the program managers. In addition to the staffing level or the equipment on the premises, these also include legislation and the funding system.

Interaction of Program and Outcomes

The complexity of ventilator weaning in adult intensive care patients unfolds in the compilation of the central outcomes and interventions (Figure 5). Individual interventions influence different outcomes at different points in time, which, in turn, are prerequisites for each other. However, not every outcome is targeted equally, and not every intervention has an impact on several endpoints.

XSL•FO
RenderX

Figure 5. Linking primary interventions and outcomes. MV: mechanical ventilation.

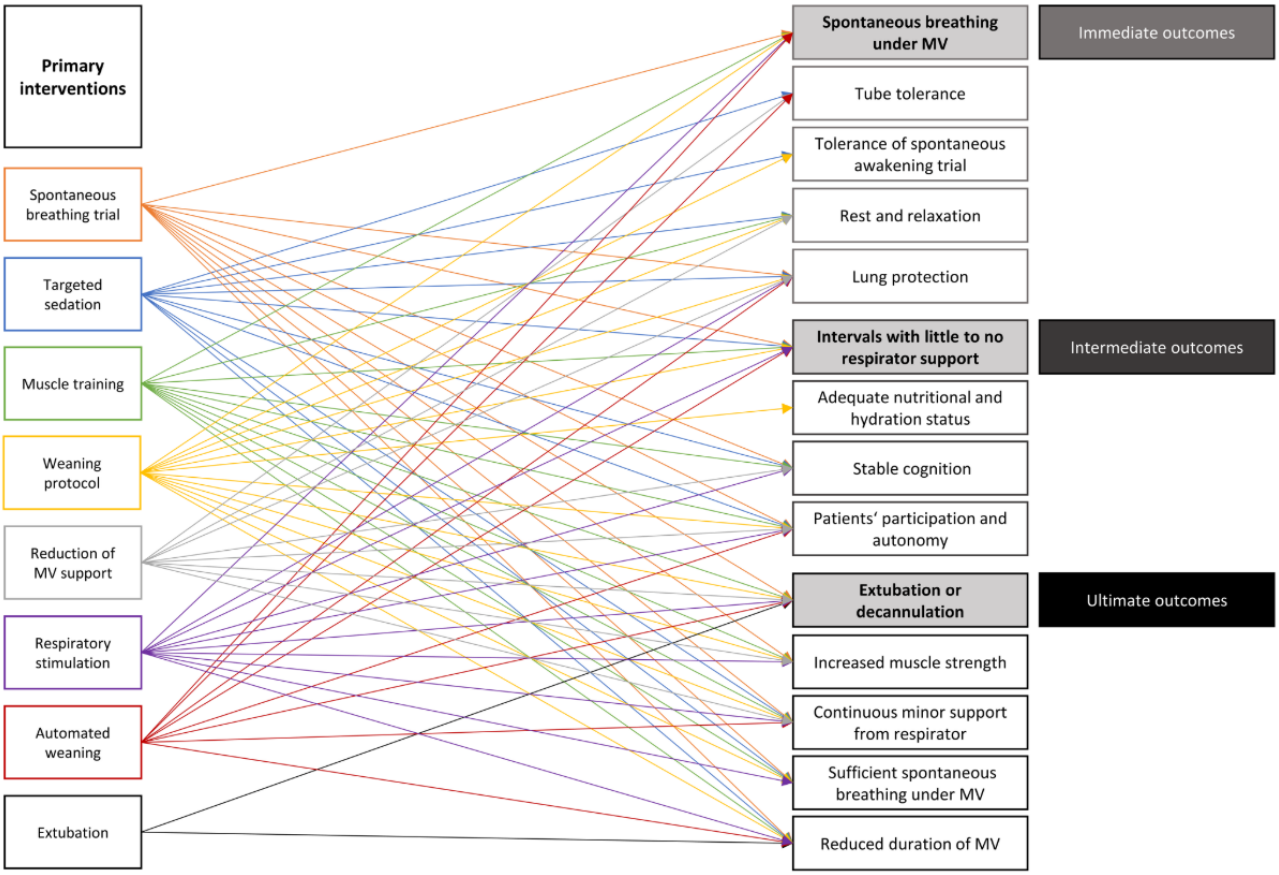

Expected Results From Step 2

In the second step of this study, the aforementioned initial program theory will be further developed and refined. The revised program theory will provide a complex theoretical understanding of ventilator dependency as the core problem and ventilator weaning as an appropriate response. To this end, the ToA and the ToC will be advanced and combined in a logic model. The necessary data are being collected in an iterative process and will be evaluated in 2026. The results and the final program theory are to be presented in 2 publications planned for the same year.

Discussion

Anticipated Findings

The aim of this study is to develop a program theory on ventilator weaning in adult intensive care patients. To this end, we have designed a comprehensive research program that is methodologically based on 2 steps and follows the current recommendations of the MRC framework [32,46]. The initial program theory presented in this paper serves as the basis for the empirical review and already differentiates between the core phenomenon, a variety of outcomes, multiple interventions, and the interconnection of these elements.

This procedure has not yet been described in this way and is unique in its combination of methods. This is also based on the fact that the development of program theories is still rare internationally. In addition, the methodological guidelines are often kept at an abstract and general level. Therefore, it is important to make methodological approaches transparent and to propose these procedures to the scientific community.

It can be assumed that the development and publication of this theory will lead to a differentiated understanding of ventilator weaning and sustainable research in science and clinical care. As mentioned by the MRC framework, program theories are not definitive and permanent. They should be further developed in the course of intervention development and testing as well as adapted to changes in health care over the years. In this regard, our revised program theory must undergo a critical examination in clinical practice. The next steps, along with the MRC framework, are piloting, evaluation, and implementation. At each stage, the program theory should be reviewed and refined as necessary [32]. In this way, program theory enables the development and manifestation of ventilator weaning as a complex intervention. We will pursue this aim in a subsequent research project.

Strengths

Several experts in critical care and social sciences are involved in the development of this theory, which is a key strength of the study. The careful consideration of the literature in step 1 provides the theory with a sound basis and makes it relevant to clinical practice and existing research. The iterative process in step 2 enables not only a one-off but repeated confirmation and refinement of the core ideas of this theory development. Moreover, all steps in the development of the theory were reported transparently and sequentially. This not only enables

XSL•FO
RenderX

traceability but also further development of the method and theory.

Limitations

This study also has limitations. First, participants are recruited exclusively from German-speaking countries. HCPs from other countries are not included in this study. Second, no quantitative methods are used to review the theory at this stage. This is due to the fact that Funnell and Rogers [46] primarily recommend stakeholder discussions and workshops. In addition, a broad body of literature with various statistical evaluations was incorporated into the theory development in the initial phase. Third, ICU patients receiving MV are diverse. They differ in terms of their primary diagnosis and comorbidities, age, and course of the disease as well as central treatment recommendations. What they all have in common is ventilator dependency as the central phenomenon. However, the extent to which the program theory can actually consider the heterogeneity of the patients or whether subgroups need to be formed will become clear in the course of the study. Fourth, no patients or relatives are to be included in the revision of this theory. This may reinforce professional perspectives in the theory and leave aspects unaddressed. However, this step was taken deliberately. As these patients constitute a vulnerable group [68], it should be carefully evaluated whether they should be studied (again) at all [69]. In our meta-synthesis [54], we found that there are already a large number of studies on the experience of MV and ventilator weaning whose data we already have access to. In addition, it turned out that the patients could not consciously distinguish between the processes of MV and ventilator weaning [27,70]. This could result in more bias in the study than potential benefit.

Conclusions

The initial program theory confirms ventilator weaning as a complex intervention and a multidisciplinary field in which nurses play a central role. After its revision, it can lead to a differentiated understanding that supports the HCPs in the ICUs in their daily work and creates sensitivity for relevant outcomes and appropriate interventions. These can be specifically combined, coordinated, and better evaluated. On the basis of this program theory, ventilator weaning can be developed, piloted, evaluated, and implemented as a complex intervention in the sense of the MRC framework.

Acknowledgments

The authors thank all stakeholders for their participation and intensive discussions during the development of the initial program theory. The authors attest that there was no use of generative artificial intelligence technology in the generation of text, figures, or other informational content of this manuscript.

Funding

FS receives a scholarship for his PhD project from the HBG Foundation [71]. The HBG Foundation only supports FS financially and has no influence on the approach and content of the dissertation and its publications, including this paper.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: FS (lead), LB, CR, S Metzing, RP
Data curation: FS
Formal analysis: FS (lead), LB
Funding acquisition: FS
Investigation: FS (lead), LB
Methodology: FS (lead), LB, CR, S Metzing, RP
Project administration: FS, S Metzing, RP
Resources: FS, LB
Supervision: FS (lead), S Metzing, RP
Validation: LB, CR, S Metzing, RP
Visualization: FS
Writing – original draft: FS (lead), LB, CR, S Metzing, RP
Writing – review & editing: FS (lead), LB, CR, S Metzing, RP

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guides for stakeholder conversations (step 1) as well as group discussions and workshops (step 2).
[PDF File (Adobe PDF File), 78 KB - resprot_v15i1e83342_app1.pdf]

XSL•FO
RenderX

References

1. Carson SS, Cox CE, Holmes GM, Howard A, Carey TS. The changing epidemiology of mechanical ventilation: a population-based study. J Intensive Care Med 2006 May 01;21(3):173-182. [doi: 10.1177/0885066605282784] [Medline: 16672639]
2. Jivraj NK, Hill AD, Shieh MS, Hua M, Gershengorn HB, Ferrando-Vivas P, et al. Use of mechanical ventilation across 3 countries. JAMA Intern Med 2023 Aug 01;183(8):824-831 [FREE Full text] [doi: 10.1001/jamainternmed.2023.2371] [Medline: 37358834]
3. Karagiannidis C, Krause F, Bentlage C, Wolff J, Bein T, Windisch W, et al. In-hospital mortality, comorbidities, and costs of one million mechanically ventilated patients in Germany: a nationwide observational study before, during, and after the COVID-19 pandemic. Lancet Reg Health Eur 2024 Jul;42:100954 [FREE Full text] [doi: 10.1016/j.lanepe.2024.100954] [Medline: 39070745]
4. Needham DM, Bronskill SE, Calinawan JR, Sibbald WJ, Pronovost PJ, Laupacis A. Projected incidence of mechanical ventilation in Ontario to 2026: preparing for the aging baby boomers. Crit Care Med 2005 Mar;33(3):574-579. [doi: 10.1097/01.ccm.0000155992.21174.31] [Medline: 15753749]
5. Peñuelas O, Frutos-Vivar F, Muriel A, Mancebo J, García-Jiménez A, de Pablo R, et al. Mechanical ventilation in Spain, 1998-2016: epidemiology and outcomes. Med Intensiva (Engl Ed) 2021;45(1):3-13. [doi: 10.1016/j.medin.2020.04.024] [Medline: 32723483]
6. Wunsch H, Linde-Zwirble WT, Angus DC, Hartman ME, Milbrandt EB, Kahn JM. The epidemiology of mechanical ventilation use in the United States. Crit Care Med 2010 Oct;38(10):1947-1953. [doi: 10.1097/CCM.0b013e3181ef4460] [Medline: 20639743]
7. Bellani G, Laffey JG, Pham T, Madotto F, Fan E, Brochard L, LUNG SAFE Investigators, ESICM Trials Group. Noninvasive ventilation of patients with acute respiratory distress syndrome. Insights from the LUNG SAFE study. Am J Respir Crit Care Med 2017 Jan 01;195(1):67-77 [FREE Full text] [doi: 10.1164/rccm.201606-1306OC] [Medline: 27753501]
8. Grasselli G, Calfee CS, Camporota L, Poole D, Amato MB, Antonelli M, European Society of Intensive Care Medicine Taskforce on ARDS. ESICM guidelines on acute respiratory distress syndrome: definition, phenotyping and respiratory support strategies. Intensive Care Med 2023 Jul;49(7):727-759 [FREE Full text] [doi: 10.1007/s00134-023-07050-7] [Medline: 37326646]
9. Karagiannidis C, Hentschker C, Westhoff M, Weber-Carstens S, Janssens U, Kluge S, et al. Observational study of changes in utilization and outcomes in mechanical ventilation in COVID-19. PLoS One 2022 Jan 14;17(1):e0262315 [FREE Full text] [doi: 10.1371/journal.pone.0262315] [Medline: 35030205]
10. Pitre T, Zeraatkar D, Kachkovski GV, Leung G, Shligold E, Dowhanik S, et al. Noninvasive oxygenation strategies in adult patients with acute hypoxemic respiratory failure: a systematic review and network meta-analysis. Chest 2023 Oct;164(4):913-928. [doi: 10.1016/j.chest.2023.04.022] [Medline: 37085046]
11. Theodoulou D, Alevrakis E, Iliopoulou M, Karampitsakos T, Katsaras M, Tsipilis S, et al. Invasive mechanical ventilation: when and to whom? Indications and complications of invasive mechanical ventilation. Pneumon 2020 Mar;33(1):13-24 [FREE Full text]
12. Pham T, Heunks L, Bellani G, Madotto F, Aragao I, Beduneau G, WEAN SAFE Investigators. Weaning from mechanical ventilation in intensive care units across 50 countries (WEAN SAFE): a multicentre, prospective, observational cohort study. Lancet Respir Med 2023 May;11(5):465-476. [doi: 10.1016/S2213-2600(22)00449-0] [Medline: 36693401]
13. Silva PL, Ball L, Rocco PR, Pelosi P. Physiological and pathophysiological consequences of mechanical ventilation. Semin Respir Crit Care Med 2022 Jun 19;43(3):321-334 [FREE Full text] [doi: 10.1055/s-0042-1744447] [Medline: 35439832]
14. Husain-Syed F, Slutsky AS, Ronco C. Lung-kidney cross-talk in the critically ill patient. Am J Respir Crit Care Med 2016 Aug 15;194(4):402-414. [doi: 10.1164/rccm.201602-0420CP] [Medline: 27337068]
15. Blanch L, Quintel M. Lung-brain cross talk in the critically ill. Intensive Care Med 2017 Apr 06;43(4):557-559. [doi: 10.1007/s00134-016-4583-1] [Medline: 27714405]
16. Li R, Zhou Y, Liu X, Huang J, Chen L, Zhang H, et al. Functional disability and post-traumatic stress disorder in survivors of mechanical ventilation: a cross-sectional study in Guangzhou, China. J Thorac Dis 2021 Mar;13(3):1564-1575 [FREE Full text] [doi: 10.21037/jtd-20-2622] [Medline: 33841948]
17. Jubran A, Grant BJ, Duffner LA, Collins EG, Lanuza DM, Hoffman LA, et al. Long-term outcome after prolonged mechanical ventilation. A long-term acute-care hospital study. Am J Respir Crit Care Med 2019 Jun 15;199(12):1508-1516 [FREE Full text] [doi: 10.1164/rccm.201806-1131OC] [Medline: 30624956]
18. Burns SM, Fisher C, Tribble SE, Lewis R, Merrel P, Conaway MR, et al. The relationship of 26 clinical factors to weaning outcome. Am J Crit Care 2012 Jan 31;21(1):52-59. [doi: 10.4037/ajcc2012425] [Medline: 22210700]
19. Ji Q, Chi L, Mei Y, Wang X, Feng J, Cai J, et al. Risk factors for late extubation after coronary artery bypass grafting. Heart Lung 2010 Jul;39(4):275-282. [doi: 10.1016/j.hrtlng.2009.09.002] [Medline: 20561839]
20. Mårtensson IE, Fridlund B. Factors influencing the patient during weaning from mechanical ventilation: a national survey. Intensive Crit Care Nurs 2002 Aug;18(4):219-229. [doi: 10.1016/s0964339702000630] [Medline: 12470012]

XSL•FO
RenderX

21. Blackwood B, Burns KE, Cardwell CR, O'Halloran P. Protocolized versus non-protocolized weaning for reducing the duration of mechanical ventilation in critically ill adult patients. Cochrane Database Syst Rev 2014 Nov 06;2014(11):CD006904 [FREE Full text] [doi: 10.1002/14651858.CD006904.pub3] [Medline: 25375085]
22. Elkins M, Dentice R. Inspiratory muscle training facilitates weaning from mechanical ventilation among patients in the intensive care unit: a systematic review. J Physiother 2015 Jul;61(3):125-134 [FREE Full text] [doi: 10.1016/j.jphys.2015.05.016] [Medline: 26092389]
23. Lo SC, Ma KS, Li YR, Li ZY, Lin CH, Lin HC, et al. Nutritional support for successful weaning in patients undergoing prolonged mechanical ventilation. Sci Rep 2022 Jul 14;12(1):12044 [FREE Full text] [doi: 10.1038/s41598-022-15917-w] [Medline: 35835785]
24. Rose L, Adhikari NK, Leasa D, Fergusson DA, McKim D. Cough augmentation techniques for extubation or weaning critically ill patients from mechanical ventilation. Cochrane Database Syst Rev 2017 Jan 11;1(1):CD011833 [FREE Full text] [doi: 10.1002/14651858.CD011833.pub2] [Medline: 28075489]
25. Crocker C, Scholes J. The importance of knowing the patient in weaning from mechanical ventilation. Nurs Crit Care 2009 Oct 13;14(6):289-296. [doi: 10.1111/j.1478-5153.2009.00355.x] [Medline: 19840275]
26. Khalafi A, Elahi N, Ahmadi F. Continuous care and patients' basic needs during weaning from mechanical ventilation: a qualitative study. Intensive Crit Care Nurs 2016 Dec;37:37-45. [doi: 10.1016/j.iccn.2016.05.005] [Medline: 27503771]
27. Schou L, Egerod I. A qualitative study into the lived experience of post-CABG patients during mechanical ventilator weaning. Intensive Crit Care Nurs 2008 Jun;24(3):171-179. [doi: 10.1016/j.iccn.2007.12.004] [Medline: 18280735]
28. Logan J, Jenny J. Deriving a new nursing diagnosis through qualitative research: dysfunctional ventilatory weaning response. Nurs Diagn 1990 Jun 28;1(1):37-43. [doi: 10.1111/j.1744-618x.1990.tb00231.x] [Medline: 2350526]
29. Lemos LA, Cavalcante LA, Lopes MV. Middle range theory for the nursing diagnosis of dysfunctional ventilatory weaning response. Int J Nurs Knowl 2020 Oct 17;31(4):253-259. [doi: 10.1111/2047-3095.12280] [Medline: 32181585]
30. Akella P, Voigt LP, Chawla S. To wean or not to wean: a practical patient focused guide to ventilator weaning. J Intensive Care Med 2022 Nov 11;37(11):1417-1425 [FREE Full text] [doi: 10.1177/08850666221095436] [Medline: 35815895]
31. Schönhofer B, Geiseler J, Dellweg D, Fuchs H, Moerer O, Weber-Carstens S, et al. Prolonged weaning: S2k guideline published by the German Respiratory Society. Respiration 2020 Dec 10;99(11):1-102. [doi: 10.1159/000510085] [Medline: 33302267]
32. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ 2021 Sep 30;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
33. Béduneau G, Pham T, Schortgen F, Piquilloud L, Zogheib E, Jonas M, WIND (Weaning according to a New Definition) Study Group and the REVA (Réseau Européen de Recherche en Ventilation Artificielle) Network. Epidemiology of weaning outcome according to a new definition. The WIND study. Am J Respir Crit Care Med 2017 Mar 15;195(6):772-783. [doi: 10.1164/rccm.201602-0320OC] [Medline: 27626706]
34. Windisch W, Dellweg D, Geiseler J, Westhoff M, Pfeifer M, Suchi S, et al. Prolonged weaning from mechanical ventilation. Dtsch Arztebl Int 2020 Mar 20;117(12):197-204 [FREE Full text] [doi: 10.3238/arztebl.2020.0197] [Medline: 32343653]
35. Dolinay T, Hsu L, Maller A, Walsh BC, Szűcs A, Jerng JS, et al. Ventilator weaning in prolonged mechanical ventilation-a narrative review. J Clin Med 2024 Mar 26;13(7):1909 [FREE Full text] [doi: 10.3390/jcm13071909] [Medline: 38610674]
36. Crocker C, Kinnear W. Weaning from ventilation: does a care bundle approach work? Intensive Crit Care Nurs 2008 Jun;24(3):180-186. [doi: 10.1016/j.iccn.2007.11.003] [Medline: 18155551]
37. Joseph A, McWilliams D, Rose L. Call to research in ventilator weaning and extubation in neurocritical care. Intensive Crit Care Nurs 2021 Aug;65:103071. [doi: 10.1016/j.iccn.2021.103071] [Medline: 34020857]
38. Robba C, Poole D, McNett M, Asehnoune K, Bösel J, Bruder N, et al. Mechanical ventilation in patients with acute brain injury: recommendations of the European Society of Intensive Care Medicine consensus. Intensive Care Med 2020 Dec 11;46(12):2397-2410 [FREE Full text] [doi: 10.1007/s00134-020-06283-0] [Medline: 33175276]
39. Astbury B, Leeuw FL. Unpacking black boxes: mechanisms and theory building in evaluation. Am J Eval 2010 Sep 13;31(3):363-381. [doi: 10.1177/1098214010371972]
40. Conn VS. Unpacking the black box: countering the problem of inadequate intervention descriptions in research reports. West J Nurs Res 2012 Jun 29;34(4):427-433. [doi: 10.1177/0193945911434627] [Medline: 22547207]
41. Harachi TW, Abbott RD, Catalano RF, Haggerty KP, Fleming CB. Opening the black box: using process evaluation measures to assess implementation and theory building. Am J Community Psychol 1999 Oct;27(5):711-731. [doi: 10.1023/A:1022194005511] [Medline: 10676545]
42. Nielsen K, Randall R. Opening the black box: presenting a model for evaluating organizational-level interventions. Eur J Work Organ Psychol 2013 Oct;22(5):601-617. [doi: 10.1080/1359432X.2012.690556]
43. Boles JM, Bion J, Connors A, Herridge M, Marsh B, Melot C, et al. Weaning from mechanical ventilation. Eur Respir J 2007 May 30;29(5):1033-1056 [FREE Full text] [doi: 10.1183/09031936.00010206] [Medline: 17470624]
44. Contentin L, Ehrmann S, Giraudeau B. Heterogeneity in the definition of mechanical ventilation duration and ventilator-free days. Am J Respir Crit Care Med 2014 Apr 15;189(8):998-1002. [doi: 10.1164/rccm.201308-1499LE] [Medline: 24735035]

XSL•FO
RenderX

45. Jaber S, De Jong A. Weaning from mechanical ventilation in intensive care units: a call for new international consensus guidelines. Lancet Respir Med 2023 May;11(5):398-400. [doi: 10.1016/S2213-2600(22)00502-1] [Medline: 36693399]
46. Funnell SC, Rogers PJ. Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. Hoboken, NJ: John Wiley & Sons; 2011.
47. Chan A, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 05;158(3):200-207 [FREE Full text] [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]
48. Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 08;346:e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
49. Brousselle A, Buregeya JM. Theory-based evaluations: framing the existence of a new theory in evaluation and the rise of the 5th generation. Eval 2018 Apr 16;24(2):153-168. [doi: 10.1177/1356389018765487]
50. Chen HT. Theory-driven evaluation: conceptual framework, application and advancement. In: Strobl R, Lobermeier O, Heitmeyer W, editors. Evaluation von Programmen und Projekten für eine demokratische Kultur. Wiesbaden, Germany: Springer VS; 2012.
51. Döring N, Bortz J. Evaluationsforschung. In: Döring N, Bortz J, editors. Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. 5th edition. Berlin, Heidelberg: Springer; 2015:975-1036.
52. Sterr F, Bauernfeind L, Knop M, Rester C, Metzing S, Palm R. Weaning-associated interventions for ventilated intensive care patients: a scoping review. Nurs Crit Care 2024 Nov 18;29(6):1564-1579. [doi: 10.1111/nicc.13143] [Medline: 39155350]
53. Sterr F, Reintke M, Bauernfeind L, Senyol V, Rester C, Metzing S, et al. Predictors of weaning failure in ventilated intensive care patients: a systematic evidence map. Crit Care 2024 Nov 12;28(1):366 [FREE Full text] [doi: 10.1186/s13054-024-05135-3] [Medline: 39533438]
54. Sterr F, Hechinger M, Bauernfeind L, Rester C, Palm R, Metzing S. Being an observer of one's own life-a meta-synthesis on the experience of mechanically ventilated patients in intensive care units. Crit Care 2025 Mar 08;29(1):105 [FREE Full text] [doi: 10.1186/s13054-025-05326-6] [Medline: 40057786]
55. Reichertz J. Die Abduktion in der qualitativen Sozialforschung: Über die Entdeckung des Neuen. 2nd edition. Wiesbaden, Germany: Springer VS; 2013.
56. Lipscomb M. Abductive reasoning and qualitative research. Nurs Philos 2012 Oct;13(4):244-256. [doi: 10.1111/j.1466-769X.2011.00532.x] [Medline: 22950728]
57. Bohnsack R, Przyborski A. Gruppendiskussionsverfahren und focus groups. In: Buber R, Holzmüller HH, editors. Qualitative Marktforschung. Wiesbaden, Germany: Gabler; 2007:491-506.
58. Przyborski A, Wohlrab-Sahr M. Qualitative Sozialforschung: Ein Arbeitsbuch. Munich, Germany: Oldenbourg Wissenschaftsverlag; 2008.
59. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008 Jul 21;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
60. Braun V, Clarke V. What can "thematic analysis" offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014 Oct 16;9:26152 [FREE Full text] [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
61. Braun V, Clarke V. Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: a review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). Palliat Med 2024 Jun 12;38(6):608-616 [FREE Full text] [doi: 10.1177/02692163241234800] [Medline: 38469804]
62. Berger R. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. Qual Res 2013 Jan 03;15(2):219-234. [doi: 10.1177/1468794112468475]
63. Yadav D. Criteria for good qualitative research: a comprehensive review. Asia Pacific Edu Res 2021 Sep 18;31(6):679-689. [doi: 10.1007/s40299-021-00619-0]
64. Healy M, Perry C. Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm. Qual Mark Res 2000;3(3):118-126. [doi: 10.1108/13522750010333861]
65. Johnson JL, Adkins D, Chauvin S. A review of the quality indicators of rigor in qualitative research. Am J Pharm Educ 2020 Jan;84(1):7120 [FREE Full text] [doi: 10.5688/ajpe7120] [Medline: 32292186]
66. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251 [FREE Full text] [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
67. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
68. Nickel PJ. Vulnerable populations in research: the case of the seriously ill. Theor Med Bioeth 2006 Jun 9;27(3):245-264. [doi: 10.1007/s11017-006-9000-2] [Medline: 16763881]
69. Gordon BG. Vulnerability in research: basic ethical concepts and general approach to review. Ochsner J 2020 Apr 03;20(1):34-38 [FREE Full text] [doi: 10.31486/toj.19.0079] [Medline: 32284680]

XSL•FO
RenderX

70. Tingsvik C, Hammarskjöld F, Mårtensson J, Henricson M. Patients' lived experience of intensive care when being on mechanical ventilation during the weaning process: a hermeneutic phenomenological study. Intensive Crit Care Nurs 2018 Aug;47:46-53. [doi: 10.1016/j.iccn.2018.03.004] [Medline: 29628332]
71. Scholarships and grants. HBG Foundation. URL: https://wissenschaft-der-pflege.de/ [accessed 2026-01-21]

Abbreviations

HCP: health care professional
ICU: intensive care unit
MRC: Medical Research Council
MV: mechanical ventilation
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
ToA: theory of action
ToC: theory of change
WIND: Weaning According to a New Definition
WEAN SAFE: Worldwide Assessment of Separation of Patients From Ventilatory Assistance

Edited by J Sarvestan; submitted 01.Sep.2025; peer-reviewed by P Nydahl; comments to author 01.Dec.2025; revised version received 08.Jan.2026; accepted 09.Jan.2026; published 30.Jan.2026.

Please cite as:
Sterr F, Bauernfeind L, Rester C, Metzing S, Palm R
Developing a Program Theory on Ventilator Weaning in Adult Intensive Care: Protocol for a Multimethods Study
JMIR Res Protoc 2026;15:e83342
URL: https://www.researchprotocols.org/2026/1/e83342
doi:10.2196/83342
PMID:

©Fritz Sterr, Lydia Bauernfeind, Christian Rester, Sabine Metzing, Rebecca Palm. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Establish a PrEP (Pre-Exposure Prophylaxis) Epidemiology, Modeling, and Surveillance (PREMISE) System to Analyze Trends in PrEP Uptake and the Impact of PrEP Programs and Policies: Protocol for a Natural Experiment and Modeling Study in the United States

Aaron J Siegler[1], PhD; Shi Hao Ernest Koh[1], MPH; Wenting Huang[1], PhD; Eric Hall[2], PhD; Jeb Jones[1], PhD; Courtney R Yarbrough[3], PhD; Xiao Zang[4], PhD; Bohdan Nosyk[5], PhD; Edwin E Corbin-Gutierrez[6], MA; Patrick S Sullivan[1], PhD

[1]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[2]OHSU-PSU School of Public Health, Oregon Health & Science University, Portland, OR, United States

[3]Department of Health Policy and Management, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[4]Division of Health Policy and Management, School of Public Health, University of Minnesota, Minneapolis, MN, United States

[5]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada

[6]National Alliance of State and Territorial AIDS Directors, Washington, DC, United States

Corresponding Author:
Aaron J Siegler, PhD
Department of Epidemiology
Rollins School of Public Health
Emory University
1518 Clifton Rd N E
Atlanta, GA, 30322
United States
Phone: 1 3178475142
Email: asiegle@emory.edu

Abstract

Background: Pre-exposure prophylaxis (PrEP) is highly effective in preventing HIV transmission; yet, many people who would benefit from PrEP are not currently using it. Numerous programs and policies, including those provided under the US Ending the HIV Epidemic effort, have been implemented to increase PrEP use. Programs vary enormously, ranging from telemedicine PrEP support to electronic medical record prompts to social marketing and messaging campaigns. However, limited evidence exists regarding their relative impact on PrEP uptake.

Objective: The aims of the PREMISE (PrEP Epidemiology, Modeling, and Surveillance) research program are to (1) provide context for PrEP scale-up in the United States, (2) assess the impact of different programs and policies on PrEP use, and (3) model the impact of PrEP-related programs and policies on population health.

Methods: The primary outcomes of PrEP use and PrEP-to-need ratios will be extracted from a national medical data aggregator database that represents a majority of PrEP users in the United States. These data will inform all proposed analyses of the project: the dataset will allow the exploration of longitudinal trends in PrEP use by modality as a cohort study, it will be the outcome data for assessing changes associated with particular PrEP programs and policies for quasi-experiments, and it will provide baseline information to inform modeling regarding future impacts of PrEP policies and programs. The implementation of policies will be assessed using legal coding at the state level, and the implementation of programs across health jurisdictions will be assessed using a jurisdiction survey conducted in collaboration with participating health departments. Guided by a legal implementation framework, we will use descriptive and regression analyses to contextualize PrEP scale-up and use quasi-experimental designs to inform causal assessments of the effect of programs and policies. Here, we provide, as preliminary data, our extraction of PrEP prescribing from the national dataset.

XSL•FO
RenderX

Results: This research was funded in August 2024. We obtained the national PrEP database and started data cleaning in March 2025. From 2016 to 2024, there were 20,394,619 claims for medications that were FDA-approved for PrEP, and we determined 13,644,979 claims to be PrEP prescriptions, representing over 1 million PrEP users. For medical benefit claims, there were 34,525 procedure claims for PrEP medications, and we determined 22,910 procedure claims to be for PrEP, representing over 6000 PrEP users.

Conclusions: To optimally use HIV prevention resources, it is critical to understand the effects of different programs and policies. Over 1 million people have started PrEP, and tracking how this scale-up has occurred by PrEP modality and user groups will inform future HIV prevention efforts. By collaborating with health jurisdictions, we will provide systematic data regarding the panoply of programs and policies that have been enacted to support PrEP use.

International Registered Report Identifier (IRRID): DERR1-10.2196/80911

(JMIR Res Protoc 2026;15:e80911) doi:10.2196/80911

KEYWORDS

HIV; prevention; pre-exposure prophylaxis; algorithm; modeling; policy

Introduction

Background

The US Ending the HIV Epidemic (EHE) launched in 2019 and is led by the US Department of Health and Human Services. EHE sought to reduce new HIV infections and scale up HIV prevention and treatment strategies to improve health outcomes for populations and communities most affected by HIV [1]. A total of 57 jurisdictions across 26 states, as well as Washington, DC, and San Juan, Puerto Rico, have been prioritized in the US EHE initiative [2]. Among these 26 states, 48 counties across 19 states were prioritized due to significant HIV burden, and 7 additional states were prioritized due to substantial occurrence of HIV in rural areas [2].

Pre-exposure prophylaxis (PrEP) is highly effective in preventing HIV transmission and remains the centerpiece of HIV prevention efforts in the United States, yet the scale-up of PrEP has led to stark disparities by race and ethnicity, limiting its impact [3]. PrEP is well tolerated, available in inexpensive generic forms, and prevents over 99% of HIV transmission when taken as directed. In numerous areas, including the United States [4] and New South Wales [5], scale-up of PrEP has been associated with reductions in new HIV transmission at the population level. Despite the spectacular promise of PrEP, less than half of persons in the United States indicated for PrEP were taking it in 2024 [6]. From 2012 to 2017, PrEP use increased by an average of 56% annually [7], followed by a more gradual annual growth of 18% from 2018 to 2021 [3]. The PrEP-to-need ratio (PnR) is a measure of PrEP use that compares the number of PrEP users in a group or geographic area (numerator) to the number of HIV diagnoses (denominator) [8]. This metric has been used to measure PrEP uptake among groups that would most benefit from it in numerous studies and has been used in other areas such as the US Centers for Disease Control and Prevention's (CDC) recent examination of mpox vaccine provision [9]. PnR assessments have revealed lower PrEP use among Black and Latinx persons, younger persons, women, and persons in the US South relative to epidemic burden [3,10]. Additionally, modeling suggests that the greatest overall HIV epidemic reduction occurs in scenarios with equal PrEP provision relative to HIV burden across racial groups (PnR=1) [11]. Thus, increasing uptake of PrEP in areas with greatest need, as defined by higher HIV incidence, maximizes the number of new infections averted for a given level of PrEP use. National disparities in PrEP provision by race, however, have increased over time [12]. There is an urgent need to identify policies and programs that improve PrEP uptake and PnR.

Despite substantial scale-up of PrEP attributable to past efforts and policies, legal and program environments in which PrEP-eligible people live are highly heterogeneous, and evidence regarding policy impact is lacking. The United States has made significant efforts to scale up HIV PrEP through EHE Programs, with over US $1.5 billion invested in HIV prevention from 2020 to 2024 [13]. EHE funding has been granted to the EHE priority jurisdictions through the Health Resources and Services Administration and CDC mechanisms that allowed state and local health departments and community health centers to implement HIV prevention and treatment initiatives. Diverse programs have been implemented with this funding—ranging from social media messaging to medical records improvements, and from clinician training to telemedicine PrEP prescribing [14-16]—yet, no research program has sought to determine the comparative effectiveness and efficiency of these PrEP interventions on a population basis to guide future investments. Outside of HIV EHE efforts, PrEP uptake may also be influenced by state-level policy decisions and the availability of CDC grant programs such as grants that were designed to support PrEP demonstration projects. Both Medicaid expansion and the use of Pre-Exposure Prophylaxis Drug Assistance Programs (PrEP DAP) offer promising approaches that are associated with increased PrEP use that is also more aligned with epidemic need [17]. It is critical to identify the program elements that have contributed most to successful outcomes within each PrEP service.

National prescription, medical, and diagnosis datasets using data from commercial entities are robust and capture a majority of all US prescriptions with line-level, deidentified individual data [17]. These comprehensive datasets enable accurate, nationwide data on PrEP prescribing, which can be used to develop valuable insights on HIV prevention efforts when analyzed in tandem with data regarding PrEP policies and programs provision. Moreover, such aggregated datasets allow for data to be rapidly assessed regarding the uptake of newer PrEP formulations and modalities (eg, long-acting injectable

XSL•FO
RenderX

[LAI] PrEP medications), providing information regarding the use and care characteristics of newer PrEP modalities relative to previously approved modalities.

Overview of Research Objectives

To support public health jurisdictions in reducing new HIV transmission, the consistent availability of national HIV PrEP data is essential. Such data can be used to describe patterns of PrEP use and to understand which programs and policies are effective at facilitating PrEP uptake, particularly among groups that will most benefit from PrEP use. The PREMISE (PrEP Epidemiology, Modeling, and Surveillance) research program aims to (1) provide context regarding the ways in which PrEP is being brought to scale in the United States, such as the use characteristics of different PrEP modalities; (2) describe the impact of health policies and programs on PrEP use and PnR in the United States; and (3) assess the characteristics of these policies and programs, their impact on population health, and their cost-effectiveness (Figure 1).

Figure 1. PrEP Epidemiology, Modeling, and Surveillance (PREMISE) research program objectives overview. Overview of the 3 main PREMISE research program objectives, their corresponding study designs, and examples of the planned analyses. PnR: PrEP-to-need ratio; PrEP: pre-exposure prophylaxis; PrEP DAP: Pre-Exposure Prophylaxis Drug Assistance Program.

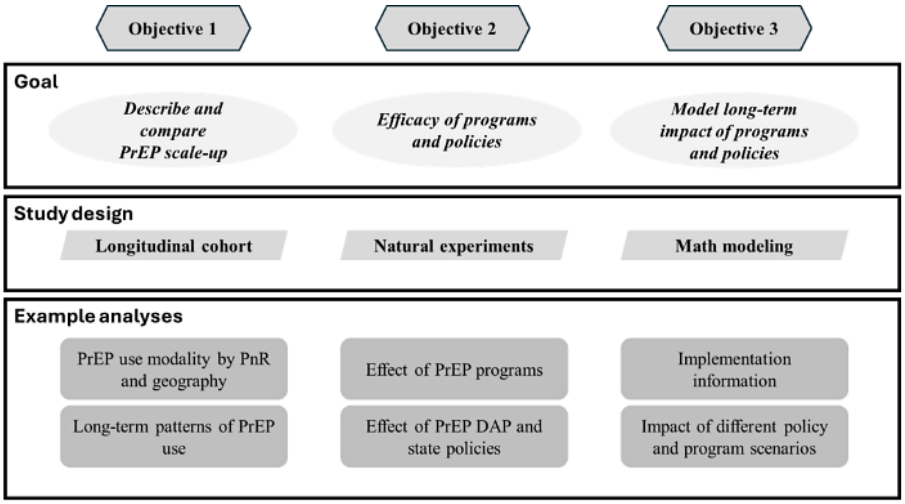

Methods

Research Setting and Population

The study population consists of patients who had claims between 2016 and 2024 for antiretroviral (ARV) medications approved by the Food and Drug Administration (FDA) for PrEP. Data on PrEP use are available from all 50 states, Washington, DC, and Puerto Rico. Objective 1 will use descriptive and exploratory assessments to contextualize PrEP scale-up. Objective 2 will compare trends in jurisdictions implementing a particular program versus a control group of jurisdictions not implementing that program. Selection of control jurisdictions will vary depending on the level of analysis: for instance, state-level analyses may include all states, whereas analyses at the county level may include counties that are matched by relevant characteristics such as HIV prevalence or county size. Data will be analyzed retrospectively from January 1, 2016, to December 31, 2024, and we will incorporate 4 future annual data updates for the years 2025 to 2028. This study period was selected because the PrEP dataset was considered by the data provider to be internally consistent in the way it was received from pharmacies over the time period, and allows for sufficient time before the launch of EHE programs in 2019. Objective 3 will model program effects for all EHE jurisdictions on future PrEP use based on an up to 20-year time horizon, expanding our team's existing models for 6 EHE jurisdictions [18,19].

Research Design Overview

Our analyses will use varied analytic approaches to answer our multiple research questions. To assess changes in PrEP provision over time (objective 1), we will use descriptive and modeling studies. To address gaps in understanding of what programs and policies will best facilitate PrEP scale-up and improved efficiency in the form of PnR (objective 2), we will use quasi-experimental approaches. Many policies and access programs influence PrEP access and uptake among states and over time. Therefore, we will use natural experiments to improve understanding of the impact on PrEP use by collecting information from jurisdictions regarding which programs were conducted and the dates of program operation. We will collect data regarding program characteristics and financial investment to improve our ability to assess causal relationships between different policies or programs and PrEP outcomes. Critical to this design is that, unlike randomized trials, it allows for assessment of health program outcomes with data from programs that were brought to scale outside of research settings by health jurisdictions. Such data will more accurately represent how health interventions are actualized in the United States. Objective 3 will use a mathematical modeling approach that is detailed below. For this paper, we followed the RECORD (Reporting of Studies Conducted Using Observational Routinely Collected Data) checklist (Multimedia Appendix 1) of items to be reported in observational studies using routinely collected health data [20].

XSL•FO
RenderX

Conceptual Framework

Our inquiry will be grounded in a legal implementation framework that we have previously proposed to facilitate the development of health policies to optimally address infectious diseases (Figure 2) [21]. In this framework, now adapted to this research program, we posit law as a causal factor influencing HIV transmission. Here, we define law broadly, as sets of rules and regulations that are enforced by the government [22]. As seen in the model, law is enacted through institutions and can influence health positively through constructs such as insurance coverage of an otherwise unaffordable service, or negatively through adverse changes in coverage programs such as Medicaid. Other mechanisms are indirect; for example, wraparound services may assist clients in gaining stable housing, which may in turn facilitate medication adherence. Given this theoretical framework, changes in laws affecting PrEP provision have the potential to have substantial and long-lasting effects on vulnerability to HIV acquisition, including the provision of health services relative to needs for those services. Laws can be either a public health tool (eg, PrEP DAP programs) or a public health threat (eg, judicial rulings that may prevent the US Preventive Service Taskforce recommendations on PrEP from being binding for insurance coverage without cost sharing). Our research acknowledges the more proximal causes of persons taking or adhering to PrEP (eg, an individual's perceived benefits of taking PrEP) and also focuses on programs and policies that shape and provide context for these decisions. We study programs and policies because (1) they are amenable to change, (2) they are measurable, and (3) there is plausibility that policies and programs can influence PrEP uptake. Relevant data sources will include legal coding and surveys to quantify programs and policies, census data to describe neighborhood factors, and PrEP claims databases to describe use and sociodemographic factors (eg, race, ethnicity, sex, and age). Due to the inherent limitations of our data sources, we cannot measure all components of the framework; instead, we will leverage it to inform the selection of data that are available.

Figure 2. PrEP Epidemiology, Modeling, and Surveillance (PREMISE) research program legal implementation framework. Overview of the legal implementation framework used to guide the PREMISE research program and assess the impact of legal and policy changes on PrEP provision and uptake, and the corresponding data sources and planned analyses. PH: public health; PrEP: pre-exposure prophylaxis.

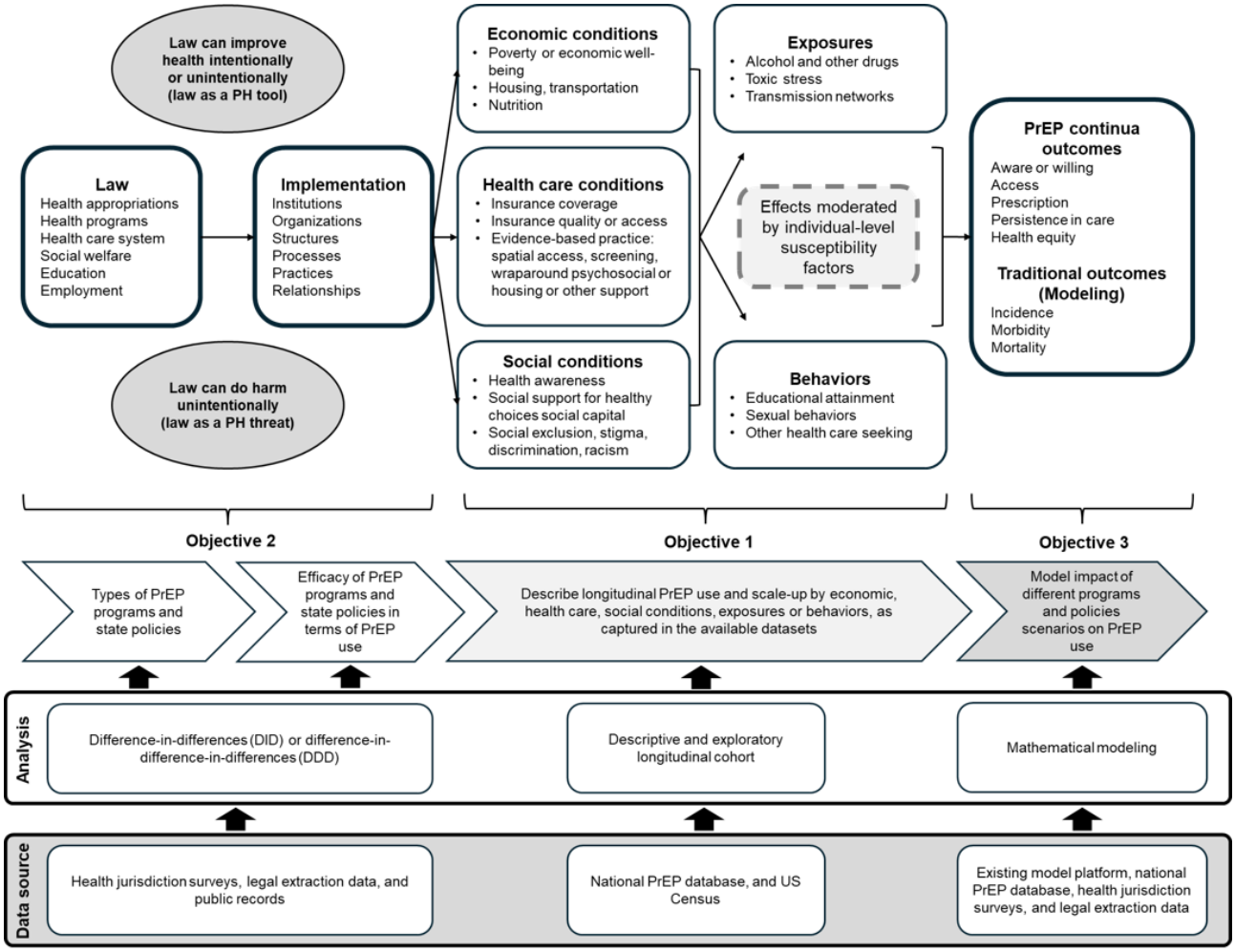

Data Sources

This program of research will use three main data sources (see Table 1 for a summary of the variables of interest and covariates): (1) data from surveys of health jurisdictions and legal data extracted by the National Alliance of State and Territorial AIDS Directors (NASTAD), (2) a national PrEP database from a third-party data provider (prescription,

XSL•FO
RenderX

procedure, and diagnosis claims datasets), and (3) US Census data. Other data sources as they become available will be included in analyses.

Table 1. PrEP (pre-exposure prophylaxis) Epidemiology, Modeling, and Surveillance (PREMISE) research program variables and covariates: definitions and data sources[a].

Construct	Levels or definition	Data source	Definition
Primary explanatory variables			
Program type	Programs supporting PrEP[b] initiation or retention	Health jurisdiction surveys and legal extraction data	Intervention program implemented
Funding spent	Dollars in staffing and direct costs expended to conduct each program	Health jurisdiction surveys	Surrogate for program magnitude
Primary outcomes			
PrEP starts	New PrEP users in each jurisdiction per time period	National PrEP database	PrEP initiation/population
PrEP start needs	New PrEP users, ratio-to-need in jurisdiction per time period	National PrEP database	PrEP initiation/HIV diagnoses
PrEP retention	User retained in care per time period	National PrEP database	Average retention time/population
Program covariates			
Funding	US Centers for Disease Control and Prevention, US Department of Health and Human Services, States	Health jurisdiction surveys and public records	Agency funding the program
Jurisdiction	County or state	Health jurisdiction surveys and legal extraction data	Jurisdiction operating the program
PrEP user and structural covariates			
Age (years)	18-24, 25-39, 40-54, 55 or older[c]	National PrEP database	Age of individual PrEP user
Race and ethnicity	Black, Latino or Hispanic, White, Asian, other[c]	National PrEP database	Race and ethnicity of individual PrEP user
Gender	Male and female[c]	National PrEP database	Self-identified gender of PrEP user
Copay	Average PrEP copay in dollars	National PrEP database	Copay of individual PrEP user
PrEP payer	Commercial, government, and other	National PrEP database	Insurer or payer of the PrEP prescription
Provider	National Provider Identifier	National PrEP database	Unique identification number of providers
Jurisdiction characteristics	% Bachelor's degree holder, % living in poverty, % uninsured, % Black residents, % Latinx residents	US Census	Proportion of county residents with a characteristic

[a]Summary of PREMISE research program constructs, including primary explanatory variables, primary outcomes, program covariates, and PrEP user and structural covariates, with corresponding levels or definitions. The table also lists data sources and how each variable is defined for analysis.

[b]PrEP: pre-exposure prophylaxis.

[c]Line-level data acquired from health exchanges. Granular data are not always available such as other identity and additional racial or age categories.

Health Jurisdiction Survey

We will conduct a survey of health jurisdictions to understand which types of programs were provided in each jurisdiction over time and the degree of resources committed to implementing each program to obtain program data; these data will inform objective 2 analyses. Table 2 details the primary program types that we will assess, representing an array of program settings and types ranging from clinician training to electronic medical records changes to direct financial coverage of PrEP services. We will also assess community change mechanisms such as messaging campaigns and community-based organization-led PrEP outreach. Each of these mechanisms might feasibly impact PrEP use, so generating evidence on the relative impact of each program will be critical to guide future investments. The full text of the jurisdictional survey can be seen in Multimedia Appendix 2. Using lists of health jurisdiction members of NASTAD, we sent the survey to program staff representatives for each of the 50 states and the 8 jurisdictions that receive EHE funding independently from their state. Our instructions clarify that completing the survey may require the knowledge of multiple jurisdictional team members, and the survey is set up to facilitate this with a link that can be accessed by multiple respondents. Data collection launched in May 2025, accompanied by a webinar offered to

XSL•FO
RenderX

all health jurisdictions to introduce the survey. We plan to complete data collection in the fall of 2025. To maximize the survey response rate, we will conduct outreach activities, including periodic email reminders, hosting open question and answer sessions, and individual support calls to engage jurisdictions to assist their survey completion.

Table 2. PrEP (pre-exposure prophylaxis) Epidemiology, Modeling, and Surveillance (PREMISE) research program health jurisdictions, program types, and details[a].

Setting and program type	Program details
In-clinic programs	
Training for clinicians and clinical staff	Training or organizational capacity building to increase PrEP[b] prescribing
EHR[c] optimization	Changes to EHR to ease PrEP delivery: this includes electronic reminders for providers to assess a patient for PrEP eligibility or to develop a list of PrEP candidates using EHR data
External clinical programs	
Safety-Net PrEP Clinic	Delivery of PrEP through an existing STI[d] clinic, community health center, pharmacy, or similar public health clinic, aimed at expanding PrEP access among key populations through reduced cost or free services
Telemedicine PrEP	Consists of full or hybrid telemedicine PrEP services that reduce or eliminate the need for in-person clinical visits
PrEP Assistance Program	Cover the cost of PrEP-related services, either directly or indirectly, usually by reimbursing providers for clinical visits and lab tests, or paying PrEP-related out-of-pocket costs to health plans
PrEP navigation services	Provide assistance to find a PrEP provider and address barriers to access PrEP, as well as support strategies to facilitate retention or reinitiation into PrEP care
Nonclinical programs	
Messaging campaigns	Digital or physical messaging, social marketing aimed at raising awareness about PrEP or sharing information about PrEP services available
PrEP outreach and education	Nonclinical outreach and education activities, typically delivered by staff or peers at community-based organizations focusing on raising awareness about PrEP, addressing misinformation, and sharing information about the availability of PrEP

[a]Health jurisdictions program settings (in-clinic, external clinic, and nonclinical) with corresponding program types and program details. The program types will be assessed to understand which programs were implemented in each jurisdiction over time and the level of resources committed to each to obtain program data.

[b]PrEP: pre-exposure prophylaxis.

[c]EHR: electronic health record.

[d]STI: sexually transmitted infection.

We will also gather information on the passage and implementation of policies regarding PrEP at the state level through a systematic review of legislative records in existing state public health statutes and the legal database Westlaw. Data will be extracted, if available, regarding program types, program start or end dates, funding agencies, jurisdictions implementing the program, the program implementers, and the program financial expenditures. To improve data quality, we will optimize interrater reliability through clear definitions of domains, use raters who are familiar with the content area, and provide thorough training for raters through practice-based learning sessions. Validation of the dataset will be accomplished through conducting parallel extraction of independent coders to determine intercoder agreement for a proportion of the dataset.

National PrEP Database (Prescription Claim, Procedure Claim, and Diagnosis Code Data)

We have obtained a national PrEP database through a contract with a third-party data provider—a health care company that aggregates national health data from clinics, providers, and pharmacies. This database includes prescription claim data, procedure claim data, and diagnosis claim data. PrEP medications can be billed through either pharmacy or medical benefits, depending on how the medication is dispensed (oral tablet or injection, pharmacy, hospital or clinic) and the requirements of the insurance plan. The diagnosis claim dataset contains medical conditions diagnosed using either the *ICD-9-CM* (*International Classification of Diseases, Ninth Edition, Clinical Modification*) or *ICD-10-CM* (*International Classification of Diseases, Tenth Edition, Clinical Modification*). For each deidentified person included in the database, the following variables are available: prescribed medication, date of each fill, number of pills or injections, sex, age, payer (Medicaid and private health insurance), and ZIP3 (the first 3 digits of a ZIP code). We will seek more specific user locations through both data crosswalks to the county level and providers' practice location addresses. On June 18, 2025, lenacapavir was approved by the FDA as a twice-yearly LAI PrEP [23], and during the research period, we expect additional medications to be approved by the FDA for PrEP. Therefore, our data provider contract includes access to prescription data for lenacapavir and other potential future modalities.

XSL•FO
RenderX

US Census Bureau Data

We will gather publicly available data about jurisdictions (eg, characteristics and demographic information) from the US Census Bureau website. Using the Census's search and filtering functions, we will extract relevant data at the appropriate geographic level (eg, state or county). These data will include distributions of educational attainment, poverty status, insurance coverage, and race and ethnicity composition, each of which will be used to examine jurisdictional differences and their associations with PrEP uptake.

Other Data Sources

Other potential sources of data may be incorporated into analyses as they become available over time.

Primary Outcome Measures

When possible, we will use PrEP outcomes that have been previously defined in prior literature [7,8,17,24-30]. Given a range of analyses across objectives 1-3 that include exploratory, quasi-experimental, and modeling that will require different time horizons and levels of aggregation, we anticipate a number of different PrEP outcomes across analyses for this research program. When appropriate, we will assess PrEP use in terms of starts and retention. PrEP starts will be defined as the number of persons starting PrEP per 100,000 population (PrEP starts per capita). For monthly assessments of this outcome, a PrEP start is defined as a person initiating PrEP for the first time or reinitiating PrEP after a lapse in prescription of a cut point, such as at least 1 month, calculated from the biweekly proportion of days covered (PDC) reaching 0.57 or lower (eg, less than 4 days per week). PDC measures the percentage of days that filled medication prescriptions could provide coverage for a period if the medication is taken as prescribed. The PDC metric has been commonly used in PrEP studies [24,25,27,28,31] and has been suggested as a good metric for studies focusing on PrEP use at specific time points [24]. We will conduct sensitivity analyses to assess the impact of different cut points for lapses in coverage. For retention in care, all PrEP fills not classified as PrEP starts will be considered prescriptions obtained while retained in PrEP care. PrEP retention will be defined as consecutive PrEP prescription refills based on their PrEP modalities, calculated as the time period in continuous PrEP care. For analyses that aggregate beyond the monthly level, such as yearly analyses, we will use definitions previously used in the literature and by the CDC. For yearly analyses, this involves counting any people with a prescribed day of PrEP in that calendar year, including persons with sufficient pills prescribed in the prior year to have days prescribed in the following year [3,8]. We will also analyze how PrEP is used relative to need by assessing the number of persons using PrEP per new HIV diagnosis (PnR). For models in objective 3, the primary outcomes will be HIV infections averted and cost-effectiveness.

Data Analyses

We plan to conduct analyses that explore trends in PrEP use over time, reveal the impact of programs and policies, and assess the long-term impact of policies and programs, including health and cost-effectiveness.

To understand PrEP use over time (objective 1), we will include current and new PrEP formulations to explore patterns regarding the growth of PrEP use, changes in retention in PrEP care over time, reinitiation of PrEP over time, and associations with changes in PrEP use. These exploratory assessments will be conducted largely with descriptive and regression analyses, and we anticipate new analyses will emerge as the prevention landscape develops. One area of potential analysis involves characterizing the long-term users of PrEP services, seeking to explore differences between those able to maintain long-term PrEP use versus those with shorter periods of PrEP use. This analysis would likely be largely descriptive and supplemented by regression models. Other analyses may be specific to the PrEP modality. For instance, analyses of injectable PrEP modalities may look at associations between their use and type of insurance coverage, socioeconomic characteristics of users, prior PrEP use patterns, and service provider locations.

To understand the impact of PrEP policies and programs (objective 2), we will leverage the legal coding and health jurisdiction program provision datasets described above to conduct quasi-experiments with outcomes of PrEP starts, PnR, and retention in care. PrEP programs have been adopted by local health authorities in the context of the disbursement of EHE funds as grants to applicants. Despite the lack of random assignment of jurisdictions to policies and programs, we propose a quasi-experimental design that allows us to have findings that reflect several of the characteristics of randomized studies, including temporality and controlling for time-invariant factors. Although quasi-experimental approaches are subject to some limitations inherent to nonrandomized designs such as challenges in controlling for time-varying characteristics, these designs minimize bias relative to other analysis approaches. Moreover, quasi-experimental approaches improve upon the external validity of randomized studies because the data stem from programs that have been enacted in practice. Because of the study design, results will be well positioned to inform real-world programs such as those that comprise EHE efforts. For quasi-experiments, we will conduct analyses using methods such as difference-in-differences (DID) that account for intervention and control jurisdiction trends before program implementation. Because interventions are implemented at different points in time, we will use the heterogeneous DID method proposed by Callaway and Sant'Anna [32] or a similar method to account for potential confounding from multiple implementation periods. Health jurisdictions implementing a program or policy will be the primary intervention unit, and controls will be jurisdictions not implementing a program. DID analyses will be accomplished with regressions that include a term for group (eg, telemedicine PrEP program or not), period (month), and group-by-period interactions that serve to estimate the effect of the intervention. Our approach provides a relatively large sample size for such analyses, with many of the 50 counties or local jurisdictions funded under the EHE program implementing programs under assessment, and other non-EHE jurisdictions also implementing some of the identified programs. Control jurisdictions will be counties or states in the United States that do not implement a particular program, using 9 years of monthly PrEP use data (2016-2024). Depending on characteristics of the intervention, such as statewide provision,

XSL•FO
RenderX

we may restrict the control group to jurisdictions meeting certain criteria in order to improve the match between intervention and control groups. For instance, we might match counties by HIV prevalence or jurisdiction size.

Although we have conducted substantial outreach to health jurisdictions to optimize engagement, we anticipate that there will be some missing data from the health jurisdiction survey. It is likely not a valid approach to impute data on whether a program was implemented in a particular jurisdiction. Therefore, analyses will be conducted only among jurisdictions providing data. To explore the impact of the missingness, we will assess descriptive differences between missing and nonmissing jurisdictions for the outcome over time and for key sociodemographic variables. Our outcome assessment will include the determination of policy impacts on PnR, an equity metric. Additionally, we will identify potential inequities in PrEP use based on key geographic, socioeconomic, and demographic characteristics. This may be conducted empirically with interaction terms between the policies of interest and these characteristics. DID models assume that trends in the intervention and control groups are parallel before the intervention. This is a critical consideration because control jurisdictions might have additional interventions to the one under consideration. We will seek to assess the validity of this premise, using techniques such as visual assessments of event-study plots and the Granger causality test to determine whether the design is sufficient to meet the parallel trends assumption. If parallel trends assumptions are not met, we will consider alternative approaches such as multivariable matching, lagged dependent variable regression, and synthetic control models (adapted for multiple treated units). The synthetic control method uses a matching algorithm to create a weighted dataset for the control (nonintervention) group, assigning a weight to each element in the group to eliminate prior differences in critical covariates and the outcome. Another option is to use difference-in-difference-in-differences (DDD) to assess the additional change from a specific policy (eg, telemedicine) above and beyond the overall EHE program effect for funded EHE jurisdictions. The nature of the DID and DDD designs controls for confounders that do not vary over time. Therefore, to optimize model efficiency, we will not control for variables expected to remain constant over time (eg, area-level poverty), only adding time-variant covariates. To further ascertain the validity of the DID approach, we will conduct placebo tests for policies that our models have found to have statistically significant effects on PrEP use. These tests assign earlier implementation dates selected at random (eg, false) to treated units to observe if we continue to reject the null hypothesis under these “false” implementation dates. Statistically insignificant results from the placebo tests will increase confidence that the model outcomes are valid. Further, we will assess the sensitivity of our models to the implementation time period by conducting the Bacon-Goodman [33] decomposition and reporting any disproportionate influence of certain implementation periods. We will seek to conduct additional sensitivity analyses that explore whether funding levels determine the impact of programs, with the program intervention variable being related to prevention budget expended rather than whether the program was enacted. Analyses will be conducted in SAS 9.4, STATA 19, or similar software. Given that there are multiple policies, it will be critical to address challenges with multiple comparisons. To accomplish this, we will, where possible, seek unified model approaches that simultaneously test the effects of multiple programs.

To assess the impact of scaling-up different PrEP programs and policies investigated in health jurisdictions (objective 3), we will use a previously published mathematical model leveraging the estimates of the effectiveness and cost of each policy and program from DID analyses to inform our models. The Localized Economic Modeling (LEM) platform is a dynamic compartmental HIV transmission model used to simulate changes in HIV epidemics. Populated with locally sourced data on demographic, epidemiological, and policy conditions for specific jurisdictions, the model has been previously applied to evaluate combination implementation strategies for EHE in 6 US cities: Atlanta (Georgia), Baltimore (Maryland), Los Angeles (California), Miami (Florida), New York City (New York), and Seattle (Washington) [18,19,34-37]. The model stratifies the study population by biological sex, race and ethnicity (non-Hispanic Black, Hispanic or Latino, non-Hispanic White, and other), HIV transmission risk group (men who have sex with men, people who inject drugs, and heterosexuals), and sexual risk behavior intensity (high- or low-risk). The adult population is followed from susceptibility to HIV infection through seroconversion, diagnosis, and treatment with antiretroviral therapy (ART), accounting for observed disparities in access to HIV health services, including HIV testing, ART, PrEP, and harm reduction services. The LEM has been applied to evaluate various public health interventions related to EHE and their health impacts [36,38,39]. As part of ongoing efforts, we are expanding and updating the underlying evidence while adapting and recalibrating the model for all 48 EHE priority counties and Washington, DC (not including San Juan, Puerto Rico, and the 7 EHE priority states with high HIV incidence in rural areas). The updated models will be leveraged to examine incident HIV infections averted, incremental costs, gains in quality-adjusted life years (a composite measure capturing HIV-related morbidity, mortality, and transmission) [40], cost-effectiveness, and health impacts (measured by disparities in HIV incidence) of the effective PrEP scale-up programs or policies identified in objective 2—compared to a “status quo” scenario assuming no further efforts to scale-up access to PrEP. We will perform the modeling analysis of PrEP programs and policies up to 20-year time horizon, evaluating them both individually and in combination to capture their long-term health and economic impacts in different jurisdictions, as well as potential synergistic effects between programs and policies. Analyses will be conducted according to best practices guidelines in simulation modeling [41] and cost-effectiveness analysis [42].

Power Calculation: Objective 2

The study design achieves 90% power to detect a minimum difference between intervention and control county PrEP rate means of 2.4 PrEP starts per 100,000 population. This is based on a sample size of 8,784,000 county-person-month observations, a figure obtained from a minimum estimate of 6 counties implementing a particular program and 238 control

XSL•FO
RenderX

counties not implementing the program. We assume that each county has an average of 300 individual PrEP users (a conservative estimate because the 2021 average and median county PrEP users are 1210 and 482), and we assume an average of 60 months of follow-up (a conservative estimate because the total follow-up period is 144 months). The standard deviation of PrEP use at the county level is 73, and we conservatively assumed a small (0.05) within-person correlation. The power assessment is based on a time-by-treatment interaction test using a mixed-model analysis with a significance of 0.05. We performed a similar analysis for state-level policies and found a minimum detectable difference of <3 additional PrEP starts. Power calculations were performed in PASS 2022.

Preliminary Findings Regarding PrEP Use Determination

We sought to determine levels of PrEP use in the third-party provider dataset (Symphony Health, an ICON plc Company, PatientSource®, January 1, 2016 to December 31, 2024). For prescriptions from aggregated datasets, a medication may sometimes be used for more than one purpose, so it is necessary to determine which prescriptions are intended for PrEP use. Informed by prior work in this area [43], we developed an algorithm to identify prescriptions for PrEP use (see Figure 3), with our algorithm foremost guided by published clinical guidance [44-50] and the availability of relevant data in the dataset. For this analysis, we extracted all prescription claims with approved PrEP medication names, that is, emtricitabine and tenofovir disoproxil fumarate (fluoro-thiacytidine [FTC]/TDF), emtricitabine and tenofovir alafenamide fumarate (FTC/TAF), and cabotegravir (CAB) (inclusion criteria 1), and then excluded the claims associated with either HIV treatment (ART) or postexposure prophylaxis (PEP). We first excluded any prescription claims that were part of a triple ARV regimen (exclusion criterion 1) and then excluded any prescription and procedure claims within 3 days of a PEP diagnosis (exclusion criterion 2). To further exclude prescription claims that were intended for HIV treatment, we incorporated relevant diagnosis data, a dataset of over 7 million HIV diagnosis claims and 60,000 PEP diagnosis claims. Medical conditions diagnosed by providers were provided in the dataset with *ICD-9-CM* or *ICD-10-CM* codes. While *ICD-10-CM* codes have replaced the *ICD-9-CM* codes since 2015 for billing purposes, some health systems may still be using *ICD-9-CM* codes. We extracted all codes related to HIV conditions (see Multimedia Appendix 3) and aggregated the number of HIV diagnoses based on the number of HIV condition-related codes for each patient. Given that for PrEP or PEP care HIV diagnoses can be mistakenly entered into the medical chart [51-53], we did not automatically exclude all persons with a single HIV diagnosis. Instead, we excluded (1) all claims for persons with two or more HIV diagnoses, (2) claims for persons with one HIV diagnosis and three or more periods of triple ARV regimen, and (3) claims for persons with no HIV diagnosis but with three or more periods of triple ARV regimen that indicate likely treatment (exclusion criterion 3). For persons with only a single HIV diagnosis and 2 or fewer triple ARV prescriptions, the triple ARV prescriptions could be for PEP, so nontriple therapy PrEP medication claims for these persons were eligible for inclusion. For individuals meeting exclusion criterion 3, the HIV diagnosis criteria, any claims for PrEP coming before the earlier of either the first HIV diagnosis or the first triple ARV regimen claim were not excluded (inclusion criteria 2). Last, we excluded the small number of claims that could represent entry errors or did not follow prescribing guidance (exclusion criterion 4): any claims for PrEP that did not reflect the standard drug strength, any claims for persons under the age of 12 years at the time of prescription, and claims for medications prior to their FDA approval—FTC/TAF prior to October 3, 2019, generic FTC/TDF prior to October 2, 2020, and CAB prior to December 20, 2021.

PrEP medications may be billed through either pharmacy or medical benefits, depending on how the medication is dispensed and the requirements of the insurance plan. Oral PrEP is typically filled at a pharmacy, and so it is captured under pharmacy claims within national aggregator datasets. In rare cases, oral PrEP can also be filled through medical benefits under the Healthcare Common Procedure Coding System (HCPCS) level two J codes, such as J0750 (FTC/TDF), J0751 (FTC/TAF), or unclassified drug codes (C9399 or J3490), with either FTC/TDF or FTC/TAF listed as the drug administered. LAI PrEP (ie, CAB) is administered in clinical settings and therefore reimbursed through medical benefits, which are captured as procedure claims using the HCPCS level two J codes such as J0739 (CAB), or unclassified drug codes (C9399 or J3490) with CAB listed as the drug administered. Depending on insurance coverage, LAI PrEP may be obtained via specialty pharmacies and covered under the pharmacy benefit. Although there is limited overlap of these datasets, we sought to ensure deduplication across PrEP claims across prescription and procedure claims datasets. For LAI PrEP, prescription and procedure claims occurring within 15 days of each other were considered as the same injection and deduplicated. A 15-day period was selected because it is less than the minimum clinical window of 1-month spacing for LAI PrEP [54]. For oral PrEP, prescription and procedure claims occurring on the same day or within a 3-day window were considered as the same prescribing event and deduplicated accordingly. We selected a 3-day window because some monthly patients can opt to refill their medications early, but we expect a small period of time to elapse before they refill.

XSL•FO
RenderX

Figure 3. PrEP Epidemiology, Modeling, and Surveillance (PREMISE) research program of PrEP use determination algorithm flow diagram. The flow diagram shows the algorithm developed to identify eligible PrEP prescription and procedure claims from the national PrEP database from a third-party data provider, including the inclusion and exclusion criteria applied at each decision point, leading to the final PrEP prescription and procedure datasets. ARV: antiretroviral; CAB: cabotegravir; FDA: Food and Drug Administration; FTC: fluoro-thiacytidine; ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification; ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification; PEP: postexposure prophylaxis; PrEP: pre-exposure prophylaxis; TAF: tenofovir alafenamide fumarate; TDF: tenofovir disoproxil fumarate.

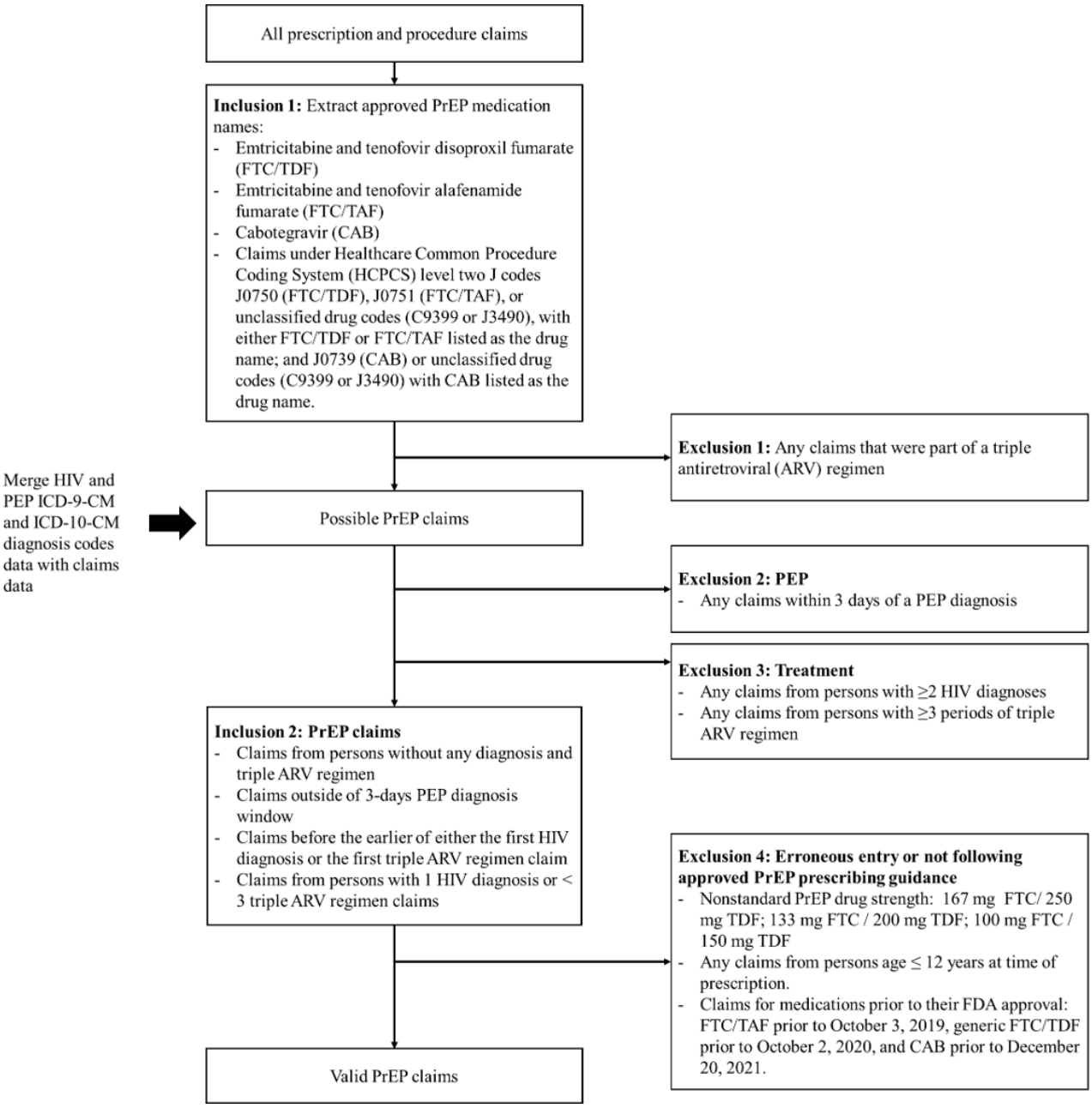

Ethical Considerations

This study involves secondary data analysis of fully deidentified prescription claims, procedure claims, and diagnosis code data, as well as a survey of US health jurisdictions to gather information on PrEP-related funding and program implementation. There was no human subjects enrollment or interaction, and no collection or use of any personal identifiable information. As such, this study does not constitute human subjects research or clinical investigation as defined in the federal regulations and was not subjected to institutional review board (IRB) review [55]. Investigators completed a worksheet provided by the Emory IRB, documenting that IRB review was not required. The study claims data are accessed by Emory under a contract and data use agreement with a third-party provider that specifies appropriate storage and data security procedures. As is normative for this type of claims data, the contract does not allow for external sharing of line-level data by Emory. However, other investigators can access the data directly through the data provider, ICON plc.

Results

Preliminary Findings

From 2016 to 2024, there were more than 20 million claims of medications approved by the FDA for PrEP: FTC/TDF,

XSL•FO
RenderX

FTC/TAF, and CAB (see Figure 4). From these prescriptions, we excluded about 6 million claims for coinciding periods that included triple ARV drug regimens (exclusion criterion 1). We then merged the diagnosis claims into the dataset to allow for further determination. We confirmed 11 million PrEP prescription claims that were connected to individuals who had no HIV diagnosis, no PEP diagnosis, and had never used a triple ARV regimen (Group 1 certified PrEP claims). For the remaining 3 million claims, we used clinical guidance for possible other uses of the drugs to determine whether a prescription was for PrEP (Group 2 certified PrEP claims) or for an alternative use. We determined 700,000 of these claims to be not for PrEP use: 6000 claims for PrEP medications that occurred within 3 days of a PEP diagnosis (exclusion criterion 2), 447,000 claims from individuals with 2 or more HIV diagnosis codes (exclusion criterion 3), 247,000 claims from individuals with only one HIV diagnosis, and 3 or more claims for triple ARV regimens (exclusion criterion 3). About 25,000 claims for PrEP medications occurred prior to FDA medication approval, 500 claims for individuals under the age of 12 years at the time of prescription, and 26,000 claims did not reflect standard PrEP drug strengths (exclusion criterion 4). This led to a final prescription dataset with over 1 million cumulative patients having initiated PrEP and over 13 million separate PrEP prescriptions.

Figure 4. PrEP Epidemiology, Modeling, and Surveillance (PREMISE) research program filtering process for identifying and extracting valid PrEP prescription claims. This figure illustrates the determination of PrEP prescriptions; after initial analysis of prescription claims, diagnosis claims inform exclusion, allowing us to identify the total number of PrEP prescription claims and patients. From 2016 to 2024, there were 20,394,619 prescription claims for FTC/TDF, FTC/TAF, and cabotegravir. We excluded 5,997,614 prescription claims identified to be part of the triple ARV regimen (exclusion 1), 6354 prescription claims occurring within 3 days of PEP diagnosis identified to be for PEP (exclusion 2), 694,379 prescription claims identified to be for treatment (exclusion 3), and 51,293 prescription claims identified to be erroneous entries or not following approved PrEP prescribing guidance (exclusion 4). The final prescription dataset contained 13,644,979 PrEP prescription claims representing 1,079,942 unique patients. ARV: antiretroviral; FTC: fluoro-thiacytidine; PEP: postexposure prophylaxis; PrEP: pre-exposure prophylaxis; TAF: tenofovir alafenamide fumarate; TDF: tenofovir disoproxil fumarate.

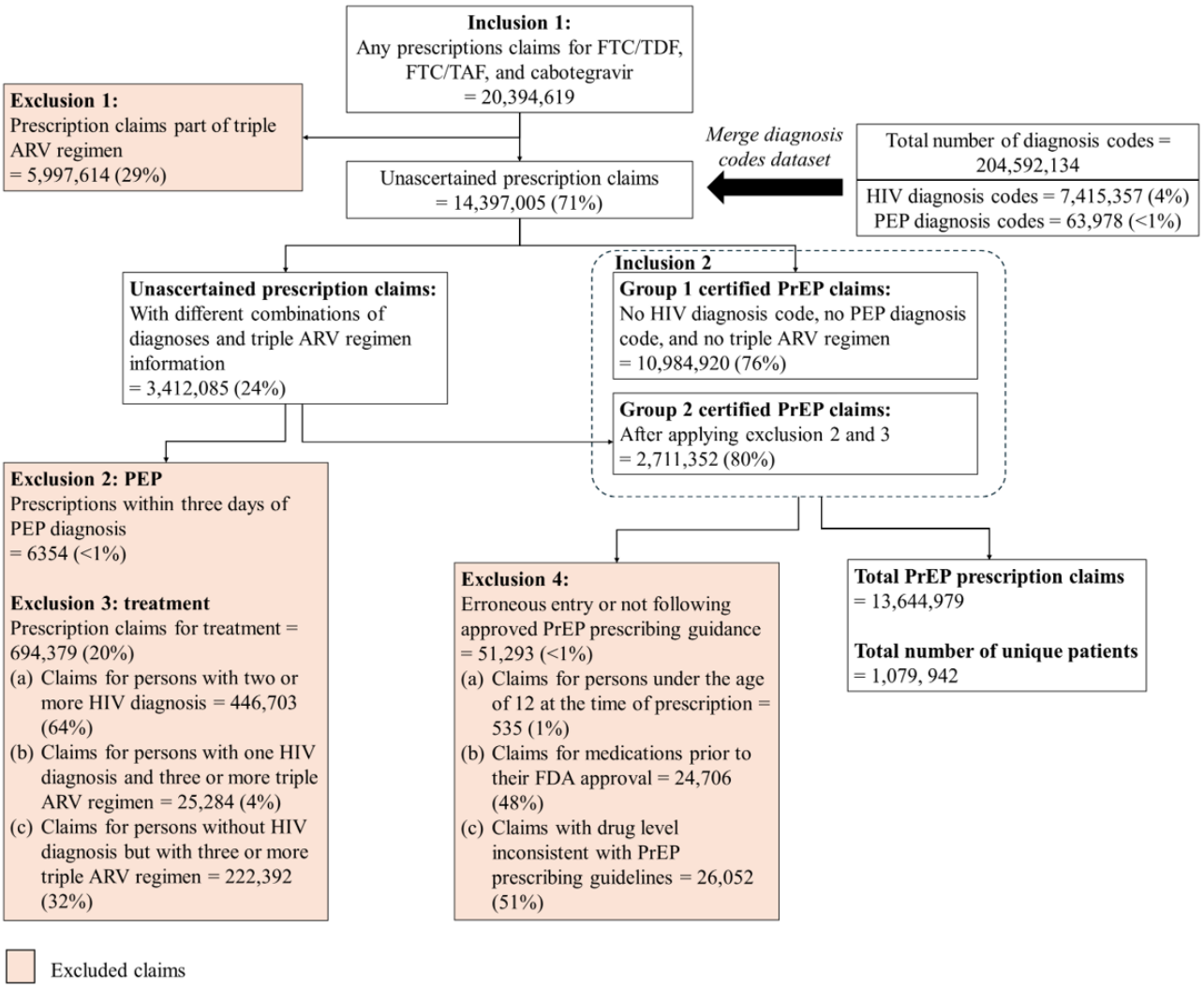

For medical benefits claims, there were about 34,000 procedure claims for PrEP provided in a clinical setting (see Figure 5). From these procedure claims, we excluded about 600 claims for coinciding periods that included triple ARV drug regimens (exclusion criterion 1). Using the same diagnosis claims dataset as described above, we confirmed 25,000 PrEP procedure claims for individuals who had no HIV diagnosis, had no PEP diagnosis, and had never used a triple ARV regimen (Group 1

XSL•FO
RenderX

certified PrEP claims). For the remaining 9000 claims, using the same clinical guidance as described above to determine whether a procedure was for PrEP (Group 2 certified PrEP claims), we determined 2200 of these claims to be not for PrEP use: 1670 claims from individuals with 2 or more HIV diagnosis codes (exclusion criterion 3), with the remaining claims from individuals with only 1 HIV diagnosis and 3 or more claims for triple ARV regimens (exclusion criterion 3), or PrEP medications that occurred within 3 days of a PEP diagnosis (exclusion criterion 2). About 9000 procedure claims for PrEP were duplicates. Under exclusion criterion 4, there was 1 claim for an individual under the age of 12 years at the time of procedure and about 100 claims for LAI PrEP with the drug name "cabotegravir" listed but an incorrect HCPCS level two J code. This led to a final procedure dataset with over 6000 cumulative patients having initiated PrEP through medical benefits with about 22,500 procedure claims for LAI PrEP and about 400 procedure claims for oral PrEP.

Figure 5. PrEP Epidemiology, Modeling, and Surveillance (PREMISE) research program filtering process for identifying and extracting valid PrEP procedure claims. This figure illustrates the process of determining PrEP procedures from the dataset, following the exclusion criteria as described in the main text. After initial analysis with procedure claims, we add diagnosis claims to further exclude procedures anticipated to not be for PrEP, showing the final number of PrEP procedure claims and patients retained. From 2016 to 2024, there were 34,525 procedure claims for FTC/TDF, FTC/TAF, and cabotegravir. We excluded 564 procedure claims identified to be part of the triple ARV regimen (exclusion 1), 2 procedure claims occurring within 3 days of PEP diagnosis identified to be for PEP (exclusion 2), 2244 procedure claims identified to be for treatment (exclusion 3), and 98 procedure claims identified to be erroneous entries or not following approved PrEP prescribing guidance (exclusion 4). The final procedure dataset contained 22,910 PrEP procedure claims representing 6196 unique patients. ARV: antiretroviral; FTC: fluoro-thiacytidine; HCPCS: Healthcare Common Procedure Coding System; PEP: postexposure prophylaxis; PrEP: pre-exposure prophylaxis; TAF: tenofovir alafenamide fumarate; TDF: tenofovir disoproxil fumarate.

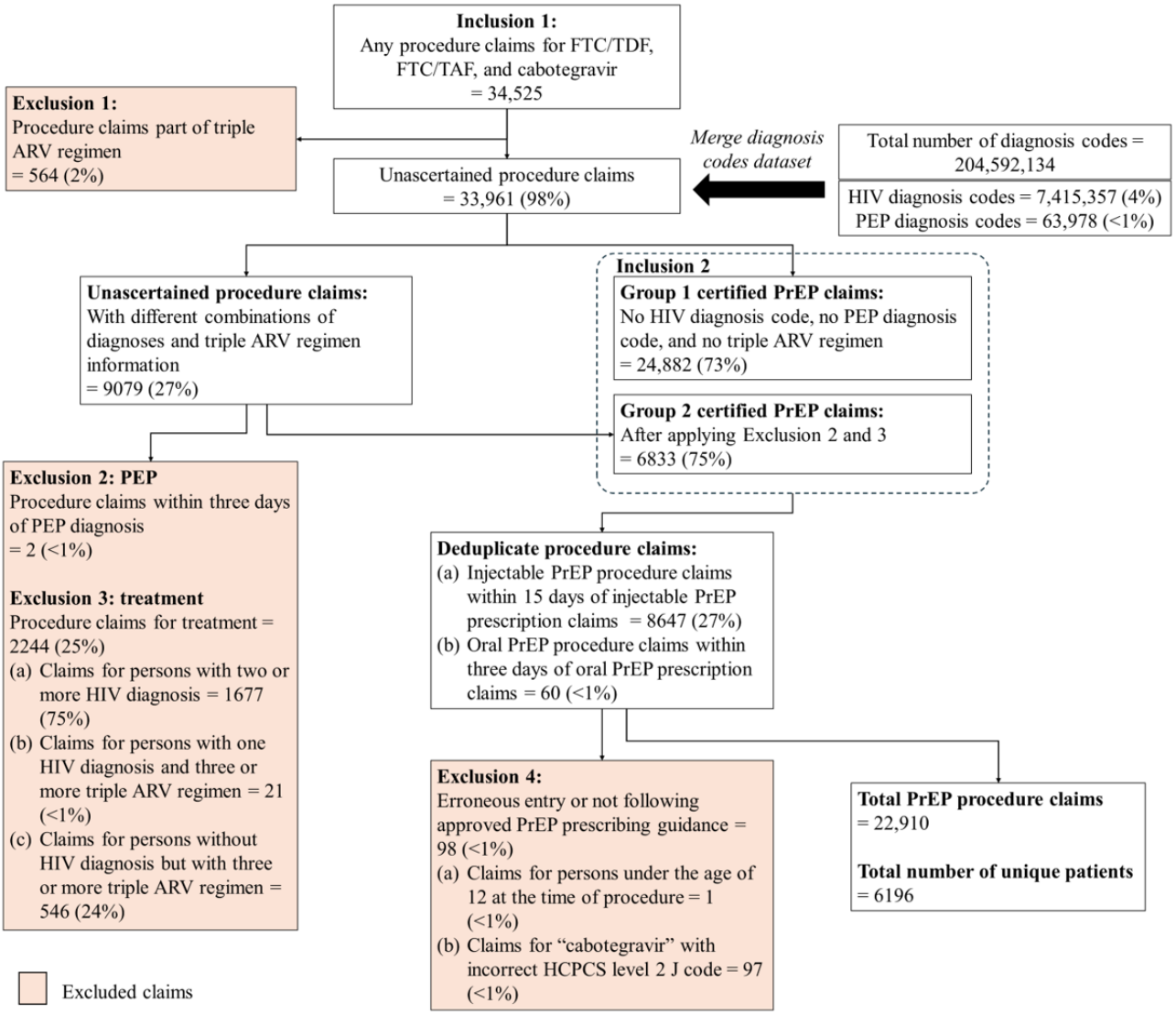

By developing an algorithm to determine PrEP use from other prescribing of ARV medication, this early analysis provides outcome data for all subsequent components of the PREMISE project. For objective 1, these data will be used to inform our descriptive and longitudinal analyses of changes in PrEP use over time. To accomplish objective 2, the data will be used as an outcome variable, allowing for the determination of the impact of PrEP policies and programs. To guide objective 3, these data will be used to inform the modeling of the future

XSL•FO
RenderX

impacts and cost-effectiveness of different programs and policies.

Anticipated Research Timeline

Extraction of PrEP prescription and procedure claims was completed as part of objective 1 in 2025. Following this critical component, the planned analyses listed under objectives 1-3 in Table 3 will be conducted. This timeline provides detailed information about the planned analyses for the grant, although we may add additional analyses as new products (eg, long-acting lenacapavir) come to market or new programs, policies, or their cessation occur over time.

Table 3. PrEP (pre-exposure prophylaxis) Epidemiology, Modeling, and Surveillance (PREMISE) anticipated timeline and research activities across the 5-year grant period[a].

Activities	Year					Data source	Planned analyses
	1	2	3	4	5		
Objective 1: Longitudinal PrEP[b] provision analyses							
Descriptive analyses of PrEP modality (TDF[c], TAF[d], and LAI[e]) by race	✓	✓				National PrEP database	Descriptive and exploratory longitudinal cohort
10-year longitudinal patterns of PrEP use		✓	✓			National PrEP database	Descriptive and exploratory longitudinal cohort
Analyses of LAI-PrEP users and retention in care		✓	✓			National PrEP database	Descriptive and exploratory longitudinal cohort
Objective 2: Quasi-experiments of PrEP policies and programs							
Develop program and policy dataset	✓	✓	✓			Health jurisdiction surveys, public records, and legal extraction data	DID[f] or DDD[g]
Conduct quasi-experiments on PrEP initiation: EHE[h], PrEP DAP[i], Medicaid Expansion, and pharmacist-provided PrEP		✓	✓	✓	✓	National PrEP database, program, and policy dataset	DID or DDD
Assess policy effects on PrEP initiation: PrEP DAP and Medicaid		✓	✓	✓		National PrEP database, program, and policy dataset	DID or DDD
Assess program and policy effects on retention in PrEP care		✓	✓	✓	✓	National PrEP database, program, and policy dataset	DID or DDD
Objective 3: Model epidemic impact and cost-effectiveness analyses							
Modeling program and policy long-term impact			✓	✓	✓	National PrEP database, program, and policy dataset	Mathematical modeling
Cost-effectiveness analyses				✓	✓	✓	National PrEP database, program, and policy dataset	Mathematical modeling
Dissemination phase							
AIDSVu blog posts and website infographics	✓	✓	✓	✓	✓	✓	✓	✓	✓	National PrEP database, program, and policy dataset	N/A[j]

[a]This table outlines the planned research activities for the 5-year grant period, detailing the research objectives, corresponding analyses, required data sources, and the anticipated year in which each activity will be conducted.

[b]PrEP: pre-exposure prophylaxis.

[c]TDF: tenofovir disoproxil fumarate.

[d]TAF: tenofovir alafenamide fumarate.

[e]LAI: long-acting injectable.

[f]DID: difference-in-differences.

[g]DDD: difference-in-difference-in-differences.

[h]EHE: Ending the HIV Epidemic.

[i]PrEP DAP: Pre-Exposure Prophylaxis Drug Assistance Program.

[j]N/A: not applicable.

XSL•FO
RenderX

Discussion

Overview

To reach EHE targets, it is crucial to have consistent national HIV PrEP data to describe patterns of use and to understand which programs and policies are effective at facilitating PrEP uptake. In PREMISE, we propose to build on our history of developing public PrEP data resources and innovative metrics of PrEP use. The research program comes from a collaborative and interdisciplinary group that will use national PrEP data, legal coding, and policy coding to inform the US HIV epidemic response. Prior studies internationally assessing PrEP policies have largely focused on whether PrEP is legally accessible and covered by insurance or public programs [56,57], which may not fully reflect the effectiveness of PrEP implementation. There is a need to understand the impact of policies and programs that support and promote PrEP as a public health good.

Limitations

The aggregated medical claims dataset that provides PrEP use outcome data for all study aims contains a majority of claims in the United States but is subject to a number of limitations. First, it does not include prescriptions from closed health care systems, including the Veterans Affairs. Data must therefore be interpreted with caution for areas with high numbers of persons enrolled in closed health care systems, such as Kaiser Permanente users. Second, PrEP use in clinical trials and other informal channels may not be captured by the claims dataset. A limitation of the quasi-experimental approach is the lack of randomization to determine the effects of programs and policies, a limitation inherent to such assessments. We will minimize the impact of this limitation by using analytic designs that carefully adjust for secular trends and will also explore in our findings other considerations for establishing a causal relationship, including effect size, dose-response, and plausibility of the causal mechanism (eg, some programs might be anticipated to have a higher impact on care initiation than retention in care). Another limitation of the quasi-experimental approach is that the COVID-19 pandemic impacted PrEP use during the epidemic period [58-60]; this could influence the outcomes of analyses. To address this challenge, we will conduct a series of visual and quantitative sensitivity analyses for each program or policy that considers the pandemic period and its impact on results. There is a dependency in identifying programs with positive effects and modeling of these effects. We anticipate that some of the funds invested in EHE will have a quantifiable impact on PrEP use and care retention, but if we find no impact from any programs, we will conduct modeling informed by other data, such as those from clinical trials and observational studies. Another limitation is that funding for EHE programs could be discontinued, but based on our power calculations, we still will have sufficient data to perform the program and policy analyses, and evidence of success may become critical to encouraging funding reauthorization. Another potential challenge is the possibility of new medication approvals; we have discussed this with our aggregated data provider, and the contract will include all current and future PrEP medication data. Lastly, while the general structure of the algorithm may be applicable to other datasets, it may not be fully transferable due to potential variations in data structure, coding practices, and available variables across different data providers.

Conclusions

HIV PrEP has the potential to contribute substantially to ending the HIV epidemic, and increasing PrEP coverage to persons at substantial risk of transmission is critical to deliver on the promise of the intervention. Diverse programs have been launched to support increased PrEP use and retention in care, and systematic data are needed to understand the impact of these programs in real-world implementation settings such as health departments. The analyses proposed in this protocol will fill this gap, providing critical information regarding program impact. We anticipate that such data will facilitate increased investment in prevention efforts by providing an evidence base to understand the relative and absolute impacts of PrEP promotion interventions.

Funding

This work was supported by the National Institute of Allergy and Infectious Diseases (1R01AI181799-01A1). The contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

Data Availability

Proprietary prescription, diagnosis, and procedure claims data were purchased through a contract agreement with a third-party data provider. Data for this publication are not available for public sharing.

Authors' Contributions

AJS led conceptualization, supervision, funding acquisition, methodology, and writing (original draft, review, and editing). SHEK contributed to project administration, data curation, formal analysis, and writing (original draft, review, and editing). WH contributed to formal analysis and writing (original draft, review, and editing). EH contributed to methodology and writing (review and editing). JJ contributed to methodology and writing (review and editing). EEC-G contributed to methodology and writing (review and editing). CRY contributed to methodology and writing (review and editing). XZ contributed to methodology and writing (review and editing). BN contributed to methodology and writing (review and editing). PSS contributed to conceptualization,

XSL•FO
RenderX

funding acquisition, methodology, and writing (review and editing). All authors had full access to all the data in this research program and had final responsibility for the decision to submit for publication.

Conflicts of Interest

AJS has received grant funding to Emory University from ViiV LLC. PSS has received grant funding and personal consulting fees from Gilead Sciences and Merck. EEC-G has received grant funding to NASTAD from Gilead Sciences and ViiV LLC, and speaker fees from ViiV LLC. All other authors have no conflicts to declare.

Multimedia Appendix 1
RECORD checklist.
[PDF File (Adobe PDF File), 134 KB - resprot_v15i1e80911_app1.pdf]

Multimedia Appendix 2
PrEP impact evaluation survey that was developed and used to survey health jurisdictions to understand the types of PrEP programs that were provided in each jurisdiction over time and the degree of resources committed to implementing each program to obtain program data. PrEP DAP: Pre-Exposure Prophylaxis Drug Assistance Program.
[DOC File , 410 KB - resprot_v15i1e80911_app2.doc]

Multimedia Appendix 3
ICD-9-CM and ICD-10-CM codes related to HIV conditions. ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification; ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification.
[DOCX File , 19 KB - resprot_v15i1e80911_app3.docx]

References

1. Ending the HIV epidemic in the US. CDC. 2024 Mar 20. URL: https://www.cdc.gov/ehe/php/about/index.html [accessed 2025-12-09]
2. EHE Overview. 2025. URL: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview [accessed 2025-06-09]
3. Sullivan PS, DuBose SN, Castel AD, Hoover KW, Juhasz M, Guest JL, et al. Equity of PrEP uptake by race, ethnicity, sex and region in the United States in the first decade of PrEP: a population-based analysis. Lancet Reg Health Am 2024;33:100738 [FREE Full text] [doi: 10.1016/j.lana.2024.100738] [Medline: 38659491]
4. Smith DK, Sullivan PS, Cadwell B, Waller LA, Siddiqi A, Mera-Giler R, et al. Evidence of an association of increases in pre-exposure prophylaxis coverage with decreases in human immunodeficiency virus diagnosis rates in the United States, 2012-2016. Clin Infect Dis 2020;71(12):3144-3151 [FREE Full text] [doi: 10.1093/cid/ciz1229] [Medline: 32097453]
5. Grulich AE, Guy R, Amin J, Jin F, Selvey C, Holden J, Expanded PrEP Implementation in Communities New South Wales (EPIC-NSW) research group. Population-level effectiveness of rapid, targeted, high-coverage roll-out of HIV pre-exposure prophylaxis in men who have sex with men: the EPIC-NSW prospective cohort study. Lancet HIV 2018;5(11):e629-e637. [doi: 10.1016/S2352-3018(18)30215-7] [Medline: 30343026]
6. Lyons SJ, Johnson AS, Hu X, Hou P, Helton B, Elenwa F, et al. Monitoring selected national HIV prevention and care objectives by using HIV surveillance data: United States and 6 dependent areas, 2020. CDC. 2022 May. URL: https://stacks.cdc.gov/view/cdc/117803 [accessed 2025-12-09]
7. Sullivan PS, Giler RM, Mouhanna F, Pembleton ES, Guest JL, Jones J, et al. Trends in the use of oral emtricitabine/tenofovir disoproxil fumarate for pre-exposure prophylaxis against HIV infection, United States, 2012-2017. Ann Epidemiol 2018;28(12):833-840 [FREE Full text] [doi: 10.1016/j.annepidem.2018.06.009] [Medline: 30037634]
8. Siegler AJ, Mouhanna F, Giler RM, Weiss K, Pembleton E, Guest J, et al. The prevalence of pre-exposure prophylaxis use and the pre-exposure prophylaxis-to-need ratio in the fourth quarter of 2017, United States. Ann Epidemiol 2018;28(12):841-849 [FREE Full text] [doi: 10.1016/j.annepidem.2018.06.005] [Medline: 29983236]
9. Kota KK, Hong J, Zelaya C, Riser AP, Rodriguez A, Weller DL, et al. Racial and ethnic disparities in mpox cases and vaccination among adult males - United States, May-December 2022. MMWR Morb Mortal Wkly Rep 2023;72(15):398-403 [FREE Full text] [doi: 10.15585/mmwr.mm7215a4] [Medline: 37053122]
10. Deeper look: PrEP. AIDSVu. 2025. URL: https://aidsvu.org/resources/deeper-look/prep/ [accessed 2025-12-09]
11. Goedel WC, Bessey S, Lurie MN, Biello KB, Sullivan PS, Nunn AS, et al. Projecting the impact of equity-based preexposure prophylaxis implementation on racial disparities in HIV incidence among MSM. AIDS 2020;34(10):1509-1517 [FREE Full text] [doi: 10.1097/QAD.0000000000002577] [Medline: 32384282]
12. Sullivan PS, Whitby S, Hipp P, Juhasz M, DuBose SN, McGuinness P, et al. Trends in PrEP inequity by race and census region, United States, 2012-2021. In: NATAP. 2022 Presented at: 24th International AIDS Conference; 2022 Jul 29 to 2022 Aug 2; Montreal URL: https://www.natap.org/2022/IAC/IAC_46.htm
13. Ending the HIV epidemic funding. HIV.gov. 2025 Mar 20. URL: https://www.hiv.gov/federal-response/ending-the-hiv-epidemic/funding [accessed 2025-12-09]

XSL•FO
RenderX

14. McGee KS. Overview of the US national HIV strategy and ending the HIV epidemic initiative. Nurs Clin North Am 2024;59(2):297-308. [doi: 10.1016/j.cnur.2023.12.003] [Medline: 38670696]
15. Li DH, Gaines Lanzi R, Pyra M. Progress and priorities to end the HIV epidemic by 2030: combined lessons from 120 US implementation research projects. J Acquir Immune Defic Syndr 2025;98(5S):e1-e9. [doi: 10.1097/QAI.0000000000003636] [Medline: 40163050]
16. Lanzi RG, Varley A, Ott C, Wolfner C, Balise R, Betancourt GS, et al. Implementation science methods, challenges, and associated solutions: transportable lessons learned and best practices from the Second National Meeting for Research and Community Collaboration Toward "Ending the HIV Epidemic in the United States". J Acquir Immune Defic Syndr 2025;98(5S):e68-e79. [doi: 10.1097/QAI.0000000000003637] [Medline: 40163057]
17. Siegler AJ, Mehta CC, Mouhanna F, Giler RM, Castel A, Pembleton E, et al. Policy- and county-level associations with HIV pre-exposure prophylaxis use, the United States, 2018. Ann Epidemiol 2020;45:24-31.e3 [FREE Full text] [doi: 10.1016/j.annepidem.2020.03.013] [Medline: 32336655]
18. Zang X, Krebs E, Min JE, Pandya A, Marshall BDL, Schackman BR, Localized HIV Modeling Study Group. Development and calibration of a dynamic HIV transmission model for 6 US cities. Med Decis Making 2020;40(1):3-16 [FREE Full text] [doi: 10.1177/0272989X19889356] [Medline: 31865849]
19. Nosyk B, Zang X, Krebs E, Enns B, Min JE, Behrends CN, Localized HIV Modeling Study Group. Ending the HIV epidemic in the USA: an economic modelling study in six cities. Lancet HIV 2020;7(7):e491-e503 [FREE Full text] [doi: 10.1016/S2352-3018(20)30033-3] [Medline: 32145760]
20. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, RECORD Working Committee. The Reporting of Studies Conducted Using Observational Routinely-Collected Health Data (RECORD) statement. PLoS Med 2015;12(10):e1001885 [FREE Full text] [doi: 10.1371/journal.pmed.1001885] [Medline: 26440803]
21. Siegler AJ, Komro KA, Wagenaar AC. Law everywhere: a causal framework for law and infectious disease. Public Health Rep 2020;135(1 Suppl):25S-31S [FREE Full text] [doi: 10.1177/0033354920912991] [Medline: 32735203]
22. Wagenaar AC, Burris S. Public Health Law Research: Theory and Methods. Rochester, NY: Social Science Research Network; 2013. URL: https://papers.ssrn.com/abstract=2273231 [accessed 2025-07-07]
23. NA. Yeztugo® (Lenacapavir) is now the first and only FDA-approved HIV prevention option offering 6 months of protection. Gilead Sciences, Inc. 2025 Jun 18. URL: https://www.gilead.com/news/news-details/2025/yeztugo-lenacapavir-is-now-the-first-and-only-fda-approved-hiv-prevention-option-offering-6-months-of-protection [accessed 2025-12-10]
24. Pyra M, Rusie L, Castro M, Keglovitz Baker K, McNulty M, Bohm N, et al. A taxonomy of pragmatic measures of HIV preexposure prophylaxis use. AIDS 2020;34(13):1951-1957 [FREE Full text] [doi: 10.1097/QAD.0000000000002618] [Medline: 33009011]
25. Townes A, Pyra M, Smith DK, Babu AS, Williams T, Wiener J, et al. PrEP use and adherence among transgender persons in Chicago, IL (SHIPP Study, 2014-2018, USA). AIDS Behav 2023;27(3):901-908 [FREE Full text] [doi: 10.1007/s10461-022-03826-y] [Medline: 36094640]
26. Jenness SM, Maloney KM, Smith DK, Hoover KW, Goodreau SM, Rosenberg ES, et al. Addressing gaps in HIV preexposure prophylaxis care to reduce racial disparities in HIV incidence in the United States. Am J Epidemiol 2019;188(4):743-752 [FREE Full text] [doi: 10.1093/aje/kwy230] [Medline: 30312365]
27. Tung EL, Thomas A, Eichner A, Shalit P. Implementation of a community pharmacy-based pre-exposure prophylaxis service: a novel model for pre-exposure prophylaxis care. Sex Health 2018;15(6):556-561. [doi: 10.1071/SH18084] [Medline: 30401342]
28. van Epps P, Maier M, Lund B, Howren MB, Beck B, Beste L, et al. Medication adherence in a nationwide cohort of veterans initiating pre-exposure prophylaxis (PrEP) to prevent HIV infection. J Acquir Immune Defic Syndr 2018;77(3):272-278. [doi: 10.1097/QAI.0000000000001598] [Medline: 29210835]
29. Chang LW, Serwadda D, Quinn TC, Wawer MJ, Gray RH, Reynolds SJ. Combination implementation for HIV prevention: moving from clinical trial evidence to population-level effects. Lancet Infect Dis 2013;13(1):65-76 [FREE Full text] [doi: 10.1016/S1473-3099(12)70273-6] [Medline: 23257232]
30. Siegler AJ, Bratcher A, Weiss KM. Geographic access to preexposure prophylaxis clinics among men who have sex with men in the United States. Am J Public Health 2019;109(9):1216-1223. [doi: 10.2105/AJPH.2019.305172] [Medline: 31318587]
31. Chang J, Do D, Delgado H, Kanimian N, Huynh A. A retrospective analysis of bone loss in tenofovir-emtricitabine therapy for HIV PrEP. Int J STD AIDS 2022;33(14):1183-1192. [doi: 10.1177/09564624221130129] [Medline: 36220789]
32. Callaway B, Sant'Anna PH. Difference-in-differences with multiple time periods. J Econom 2021;225(2):200-230. [doi: 10.1016/j.jeconom.2020.12.001]
33. Goodman-Bacon A. Difference-in-differences with variation in treatment timing. J Econom 2021;225(2):254-277. [doi: 10.1016/j.jeconom.2021.03.014]
34. Krebs E, Zang X, Enns B, Min JE, Behrends CN, Del Rio C, Localized Economic Modeling Study Group. The impact of localized implementation: determining the cost-effectiveness of HIV prevention and care interventions across six United States cities. AIDS 2020;34(3):447-458 [FREE Full text] [doi: 10.1097/QAD.0000000000002455] [Medline: 31794521]

XSL•FO
RenderX

35. Nosyk B, Zang X, Krebs E, Min JE, Behrends CN, Del Rio C, et al. Ending the epidemic in America will not happen if the status quo continues: modeled projections for human immunodeficiency virus incidence in 6 US Cities. Clin Infect Dis 2019;69(12):2195-2198 [FREE Full text] [doi: 10.1093/cid/ciz1015] [Medline: 31609446]
36. Zang X, Mah C, Quan AML, Min JE, Armstrong WS, Behrends CN, Localized HIV Modeling Study Group. Human immunodeficiency virus transmission by HIV risk group and along the HIV care continuum: a contrast of 6 US Cities. J Acquir Immune Defic Syndr 2022;89(2):143-150 [FREE Full text] [doi: 10.1097/QAI.0000000000002844] [Medline: 34723929]
37. Zang X, Krebs E, Chen S, Piske M, Armstrong WS, Behrends CN, Localized HIV Modeling Study. The potential epidemiological impact of coronavirus disease 2019 (COVID-19) on the human immunodeficiency virus (HIV) epidemic and the cost-effectiveness of linked, opt-out HIV testing: a modeling study in 6 US Cities. Clin Infect Dis 2021;72(11):e828-e834 [FREE Full text] [doi: 10.1093/cid/ciaa1547] [Medline: 33045723]
38. Zang X, Piske M, Humphrey L, Enns B, Sui Y, Marshall BDL, et al. Estimating the epidemiological impact of reaching the objectives of the Florida integrated HIV prevention and care plan in Miami-Dade County. Lancet Reg Health Am 2023;27:100623 [FREE Full text] [doi: 10.1016/j.lana.2023.100623] [Medline: 37928440]
39. Nosyk B, Krebs E, Zang X, Piske M, Enns B, Min JE, et al. "Ending the epidemic" will not happen without addressing racial/ethnic disparities in the United States human immunodeficiency virus epidemic. Clin Infect Dis 2020;71(11):2968-2971 [FREE Full text] [doi: 10.1093/cid/ciaa566] [Medline: 32424416]
40. Chapman RH, Berger M, Weinstein MC, Weeks JC, Goldie S, Neumann PJ. When does quality-adjusting life-years matter in cost-effectiveness analysis? Health Econ 2004;13(5):429-436. [doi: 10.1002/hec.853] [Medline: 15127423]
41. Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, ISPOR-SMDM Modeling Good Research Practices Task Force. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-5. Value Health 2012;15(6):828-834 [FREE Full text] [doi: 10.1016/j.jval.2012.06.011] [Medline: 22999132]
42. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. JAMA 2016;316(10):1093-1103. [doi: 10.1001/jama.2016.12195] [Medline: 27623463]
43. Sullivan PS, Mera-Giler RM, Bush S, Shvachko V, Sarkodie E, O'Farrell D, et al. Claims-based algorithm to identify pre-exposure prophylaxis indications for tenofovir disoproxil fumarate and emtricitabine prescriptions (2012-2014): validation study. JMIR Form Res 2024;8:e55614 [FREE Full text] [doi: 10.2196/55614] [Medline: 39141024]
44. DESCOVY (emtricitabine/tenofovir alafenamide) FOR PrEP®. Gilead Sciences, Inc. URL: https://www.descovyhcp.com/ [accessed 2025-12-10]
45. What is Truvada®? Gilead Sciences, Inc. URL: https://www.truvada.com/ [accessed 2025-12-10]
46. Home | See Boxed Warning | APRETUDE (cabotegravir). ViiV Healthcare. URL: https://apretudehcp.com/ [accessed 2025-12-10]
47. HIV/AIDS fact sheets. NIH. URL: https://hivinfo.nih.gov/understanding-hiv/fact-sheets [accessed 2025-12-10]
48. Clinical guidance for PrEP. CDC. 2025 Feb 10. URL: https://www.cdc.gov/hivnexus/hcp/prep/index.html [accessed 2025-12-10]
49. Clinical guidance for PEP. CDC. 2025 Feb 10. URL: https://www.cdc.gov/hivnexus/hcp/pep/index.html [accessed 2025-12-10]
50. What to start: initial combination antiretroviral regimens. NIH. 2024 Sep 12. URL: https://clinicalinfo.hiv.gov/en/guidelines/hiv-clinical-guidelines-adult-and-adolescent-arv/what-start-initial-combination [accessed 2025-12-10]
51. Pre-exposure prophylaxis (PrEP), post-exposure prophylaxis (PEP), and other HIV prevention strategies: billing and coding guide. NASTAD. 2023 Oct 12. URL: https://nastad.org/resources/billing-coding-guide-hiv-prevention [accessed 2025-12-10]
52. Burks K, Shields J, Evans J, Plumley J, Gerlach J, Flesher S. A systematic review of outpatient billing practices. SAGE Open Med 2022;10:20503121221099021 [FREE Full text] [doi: 10.1177/20503121221099021] [Medline: 35646364]
53. Pocobelli G, Oliver M, Albertson-Junkans L, Gundersen G, Kamineni A. Validation of human immunodeficiency virus diagnosis codes among women enrollees of a U.S. health plan. BMC Health Serv Res 2024;24(1):234 [FREE Full text] [doi: 10.1186/s12913-024-10685-x] [Medline: 38389066]
54. APRETUDE (cabotegravir 200 mg/mL) dosing. HCP Website. URL: https://apretudehcp.com/dosing-and-drug-interactions/ [accessed 2025-12-10]
55. Does My Project Need IRB Review? Emory University | Institutional Review Board. URL: https://irb.emory.edu/guidance/getting-started/review.html [accessed 2026-01-27]
56. Sang JM, McAllister K, Wang L, Barath J, Lal A, Parlette A, et al. Examining provincial PrEP coverage and characterizing PrEP awareness and use among gay, bisexual and other men who have sex with men in Vancouver, Toronto and Montreal, 2017-2020. J Int AIDS Soc 2022;25(10):e26017 [FREE Full text] [doi: 10.1002/jia2.26017] [Medline: 36306245]
57. Murphy L, Bowra A, Adams E, Cabello R, Clark JL, Konda K, et al. PrEP policy implementation gaps and opportunities in Latin America and the Caribbean: a scoping review. Ther Adv Infect Dis 2023;10:20499361231164030 [FREE Full text] [doi: 10.1177/20499361231164030] [Medline: 37114192]
58. Schmidt MA, Salas SB, Donald JL, Gift TL, Tao G. Impact of the early COVID-19 pandemic on the number of HIV preexposure prophylaxis uses and the proportion of preexposure prophylaxis users receiving sexually transmitted infection

XSL•FO
RenderX

testing services. Sex Transm Dis 2023;50(5):304-309 [FREE Full text] [doi: 10.1097/OLQ.0000000000001726] [Medline: 36730891]
59. Pampati S, Emrick K, Siegler AJ, Jones J. Changes in sexual behavior, PrEP adherence, and access to sexual health services because of the COVID-19 pandemic among a cohort of PrEP-using MSM in the South. J Acquir Immune Defic Syndr 2021;87(1):639-643 [FREE Full text] [doi: 10.1097/QAI.0000000000002640] [Medline: 33512848]
60. Hong C. Characterizing the impact of the COVID-19 pandemic on HIV PrEP care: a review and synthesis of the literature. AIDS Behav 2023;27(7):2089-2102 [FREE Full text] [doi: 10.1007/s10461-022-03941-w] [Medline: 36459263]

Abbreviations

ART: antiretroviral therapy
ARV: antiretroviral
CAB: cabotegravir
CDC: US Centers for Disease Control and Prevention
DDD: difference-in-difference-in-differences
DID: difference-in-differences
EHE: Ending the HIV Epidemic
FDA: Food and Drug Administration
FTC: fluoro-thiacytidine
HCPCS: Healthcare Common Procedure Coding System
ICD-9-CM: International Classification of Diseases, Ninth Revision, Clinical Modification
ICD-10-CM: International Classification of Diseases, Tenth Revision, Clinical Modification
IRB: institutional review board
LAI: long-acting injectable
LEM: Localized Economic Modeling
NASTAD: National Alliance of State and Territorial AIDS Directors
PDC: proportion of days covered
PEP: postexposure prophylaxis
PnR: PrEP-to-need ratio
PREMISE: PrEP Epidemiology, Modeling, and Surveillance
PrEP: pre-exposure prophylaxis
PrEP DAP: Pre-Exposure Prophylaxis Drug Assistance Program
RECORD: Reporting of Studies Conducted Using Observational Routinely Collected Data
TAF: tenofovir alafenamide fumarate
TDF: tenofovir disoproxil fumarate

Edited by A Schwartz; submitted 21.Jul.2025; peer-reviewed by O Igunma, NP Calaguas; comments to author 24.Oct.2025; accepted 30.Nov.2025; published 30.Jan.2026.

Please cite as:
Siegler AJ, Koh SHE, Huang W, Hall E, Jones J, Yarbrough CR, Zang X, Nosyk B, Corbin-Gutierrez EE, Sullivan PS
Establish a PrEP (Pre-Exposure Prophylaxis) Epidemiology, Modeling, and Surveillance (PREMISE) System to Analyze Trends in PrEP Uptake and the Impact of PrEP Programs and Policies: Protocol for a Natural Experiment and Modeling Study in the United States
JMIR Res Protoc 2026;15:e80911
URL: https://www.researchprotocols.org/2026/1/e80911
doi:10.2196/80911
PMID:

©Aaron J Siegler, Shi Hao Ernest Koh, Wenting Huang, Eric Hall, Jeb Jones, Courtney R Yarbrough, Xiao Zang, Bohdan Nosyk, Edwin E Corbin-Gutierrez, Patrick S Sullivan. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Differentiating the COVID-19 Infection and Vaccine Experiences of Patients With Systemic, Single Organ, and Overlap Immune-Mediated Inflammatory Disease: Protocol for a Secondary Analysis for Enhancing COVID-19 Vaccine Pharmacovigilance

Meredith Leston[1], BA, MPH; José M Ordóñez-Mena[1], PhD; Xinchun Gu[1], PhD; Mark Joy[1], PhD; F D Richard Hobbs[1,2], Prof Dr Med; Lennard Y W Lee[1], PhD; Ioannis Parodis[3], PhD; Laura Andreoli[4], PhD; Nelly Ziade[5], PhD; Latika Gupta[6], PhD; Vikas Agarwal[7], PhD; Richard Conway[8], Prof Dr Med; John Isaacs[9,10], MB, BS, PhD; Jeffrey Curtis[11], Prof Dr Med; Alessia Alunno[12], PhD; Stefan Siebert[13], Prof Dr Med; Ori Elkayam[14], MD; Karen Schreiber[15,16,17], PhD; Simon de Lusignan[18], Prof Dr Med

[1]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom
[2]Oxford Institute of Digital Health, University of Oxford, Oxford, United Kingdom
[3]Karolinska University Hospital, Stockholm, Sweden
[4]University of Brescia, Brescia, Italy
[5]St Joseph University, Beirut, Lebanon
[6]The Royal Wolverhampton NHS Trust, Wolverhampton, United Kingdom
[7]Sanjay Gandhi Post Graduate Institute of Medical Sciences, Lucknow, India
[8]St James' Hospital, Dublin, Ireland
[9]Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom
[10]Newcastle NIHR Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom
[11]University of Alabama at Birmingham, Birmingham, AL, United States
[12]Internal Medicine and Nephrology Division, MeSVA Department, University of L'Aquila, L'Aquila, Italy
[13]University of Glasgow, Glasgow, United Kingdom
[14]Department of Rheumatology, Tel Aviv Medical Center, Gray Faculty of Medical and Health Sciences, Tel Aviv University, Tel Aviv, Israel
[15]Danish Center for Expertise in Rheumatology (CeViG), Danish Hospital for Rheumatic Diseases, Sønderborg, Denmark
[16]Institute for Regional Health, Southern Danish University, Odense, Denmark
[17]Thrombosis and Haemostasis, Guy’s and St Thomas NHS Foundation Trust, London, United Kingdom
[18]University of Oxford, Oxford, United Kingdom

Corresponding Author:
Meredith Leston, BA, MPH
Nuffield Department of Primary Care Health Sciences
University of Oxford
Radcliffe Primary Care Building, Radcliffe Observatory Quarter
Woodstock Rd
Oxford, OX2 6GG
United Kingdom
Phone: 44 01865 617855
Email: meredith.leston@phc.ox.ac.uk

Abstract

Background: Patients with immune-mediated inflammatory disease (IMID), including autoimmunity, fared substantially worse than the general population during the COVID-19 pandemic, both in terms of infection outcomes and disruption to daily life. Despite this, COVID-19 vaccine uptake has not been universal in this population. The absence of patients with IMID from clinical trials and the subsequent lack of precision in vaccine safety profiling have resulted in vaccine hesitancy in this high-risk group.

XSL•FO
RenderX

Objective: This protocol sets out an investigation that aims to address this by enhancing COVID-19 vaccine pharmacovigilance for patients with IMID. Combining the international data and knowledge assets of the COVID-19 Vaccination in Autoimmune Diseases (COVAD) 1 study and the electronic Delphi Study to Define and Risk-Stratify Immunosuppression (DESTINIES), the objective is to differentiate patient-reported COVID-19 infection and vaccine outcomes between participants with systemic, single organ, and overlap IMID and general population controls.

Methods: The COVAD-1 study successfully collected anonymized data on the demographic, health, COVID-19 infection, and COVID-19 vaccination outcomes of a broad range of participants with IMID between March and December 2021. This protocol expands on this initial analysis by using IMID specialists within the DESTINIES Consortium to allocate survey respondents into single organ and systemic categories and thereby produce comparative vaccine benefit-risk profiles between these and general population controls. Because of the respondents' ability to self-report multiple diagnoses, an overlap group was introduced for those affected by both single organ and systemic disease. Descriptive statistics and both single and multivariable logistic regressions will be used to test for significant differences in COVID-19 infection rates, severity, duration, and vaccine side effects between these study groups and general population controls.

Results: A panel of 7 IMID experts successfully allocated COVAD-1 diagnoses into single organ and systemic categories; this also directed overlap category membership. Although this work is preliminary and highly exploratory, we anticipate that subsequent analysis will reveal disproportionate levels of severe COVID-19 infection outcomes (hospitalization with and without oxygen support) and vaccine side effects (mild and major) among participants with systemic manifestations of IMID, especially those that qualify for the overlap IMID category.

Conclusions: Advocating for direct-to-patient vaccine reporting pathways, this study intends to produce more precise vaccine safety profiles of patients with IMID. It seeks to resolve current gaps in pharmacovigilance and potentially remedy vaccine hesitancy in high-risk groups by doing so. The international nature of COVAD-1 data collection and the nuance of information made available through participant self-report are to the advantage of this protocol. However, the dependence of this study on participant recall, the small sample sizes handled, and the questionable relevance of these data in the contemporary Omicron era are to the detriment of this work.

International Registered Report Identifier (IRRID): DERR1-10.2196/68785

(JMIR Res Protoc 2026;15:e68785) doi:10.2196/68785

KEYWORDS

vaccines; vaccination; pharmacovigilance; surveillance; patient-centered research; adverse events of interest; side effects; digital health; COVID-19

Introduction

Background

The COVID-19 pandemic has had a disproportionate impact on the lives of those with immune-mediated inflammatory diseases (IMID), including autoimmunity [1]. By virtue of their immunosuppressed state, these patients were incorporated into shielding programs where possible and were prioritized for novel medical supplies, including COVID-19 antivirals and vaccines [2]. Such measures were necessary given the heightened rates of hospitalizations, respiratory complications, and mortality observed in patients with IMID who contracted COVID-19 [3,4].

Despite this, COVID-19 vaccine uptake was not universal among patients with IMID and the immunosuppressed more broadly [5]. The absence of these patients from vaccine development trials [6], coupled with the lack of observational studies on their specific vaccine tolerance [7], has generated notable hesitancy in this group [8]. This is visible in comparisons of vaccine refusal rates between key risk groups: work by Gaur et al [9] saw patients with autoimmune disease led chronic risk groups with 19% COVID-19 vaccine refusal versus 17.8% and 13.4% seen among patients with chronic lung disease and cancer, respectively. More precise assessments of vaccine risk in the IMID population are urgently needed to reassure these patients about the safety and suitability of these products [10].

One solution is to enhance pharmacovigilance surveillance. Current methods, including the analysis of computerized medical records or data from centralized reporting mechanisms (Yellow Card, Vaccine Adverse Event Reporting System, etc), are primarily clinician facing and select for major adverse events [11]. It is hypothesized that direct-to-patient vaccine side effect surveys may deliver the detailed and differentiable data that high-resolution pharmacovigilance requires.

The first iteration of the COVID-19 Vaccination in Autoimmune Diseases (COVAD) study, COVAD-1 [12], was established to test this hypothesis, initially identifying whether a patient COVID-19 vaccine survey could provide sufficient data to compare safety profiles between patients with a diverse range of IMID—with a special emphasis on autoimmunity—and general population controls. This protocol, however, describes a follow-on study that will enhance this initial analysis, assessing whether the COVAD dataset can also be used to test for internal heterogeneity in the COVID-19 infection and vaccine experience of the IMID population.

Study Design

To do so, this work will allocate COVAD respondents with IMID into categories aligned with the electronic Delphi Study

XSL•FO
RenderX

to Define and Risk-Stratify Immunosuppression (DESTINIES) phenotype [13]—a COVID-19 risk hierarchy produced and ratified over the course of a multistage international clinical consensus-building exercise of 64 global experts in immunology, vaccinology, and infectious disease (the DESTINIES Consortium). IMID data will be categorized as follows: single organ, systemic, and overlap IMID.

As per the DESTINIES evaluations, we are anticipating that, overall, respondents with IMID will have experienced worse COVID-19 infection and vaccine outcomes than the general population controls. However, internal heterogeneity is highly likely. Following the structure of the DESTINIES phenotype, we also hypothesize that patients with single organ IMID will report milder COVID-19 infections and fewer vaccine side effects than their systemic and overlap counterparts. We also expect that the multimorbidity of the overlap category will predispose this group to the worst outcomes of all those evaluated. However, this investigation is preliminary and highly exploratory; these hypotheses should therefore only be considered as evidence generating at this moment in time.

Methods

Overview

COVAD-1 study design, recruitment, data collection, and initial results have been described in detail elsewhere [12]. This multinational cross-sectional survey was active in over 90 countries and accumulated support from a collective of 106 rheumatologists, internists, neurologists, and immunologists. Recruitment used a diverse array of social media channels and online patient support groups globally [14].

Between March and December 2021, the COVAD-1 study collected the following information from vaccine surveys presented to an international cohort of patients with IMID and population-based controls: demographic details, IMID type, treatment details, current symptoms, COVID-19 infection history (inclusive of symptoms, duration, and complications), COVID-19 vaccine details, 7-day COVID-19 adverse events of interest, and specific patient-reported outcome measures (health, pain, activity, fatigue, and physical function status). All personally identifiable data were removed, anonymizing this asset for secondary analysis. Informed consent was obtained at the start of the survey, and no financial incentives were offered for completion.

In total, 18,882 patients with IMID or population-based controls responded to the survey [14]. Self-reported by survey respondents, the following IMIDs were represented in this study set: ankylosing spondylitis or psoriatic arthritis, antisynthetase syndrome, Crohn disease or ulcerative colitis (inflammatory bowel disease), dermatomyositis, hemolytic anemia or idiopathic thrombocytopenic purpura, inclusion body myositis, juvenile dermatomyositis, mixed connective tissue disease, multiple sclerosis, myasthenia gravis, necrotizing myositis, overlap myositis with lupus or Sjögren syndrome or systemic sclerosis or rheumatoid arthritis, pernicious anemia, polymyalgia rheumatica, polymyositis, rheumatoid arthritis, scleroderma, Sjögren syndrome, systemic lupus erythematosus, thyroid disease (hypothyroidism or hyperthyroidism), type 1 diabetes, and vasculitis. Respondents were able to indicate when they had multiple diagnoses.

All survey respondents were older than 18 years, ensuring congruence between this study set and the DESTINIES primary output, the DESTINIES phenotype [13]. Visualized in Figure 1 [13], the DESTINIES phenotype COVID-19 segments the medical diagnoses and procedures that clinical experts agreed confer immunosuppression into 10 risk levels. Here, IMIDs are categorized into single organ and systemic categories, the latter of which is considered at higher risk for severe COVID-19 outcomes. The term single site was updated to single organ to better align with prevailing taxonomies [15].

Figure 1. The Delphi Study to Define and Risk-Stratify Immunosuppression (DESTINIES) phenotype.

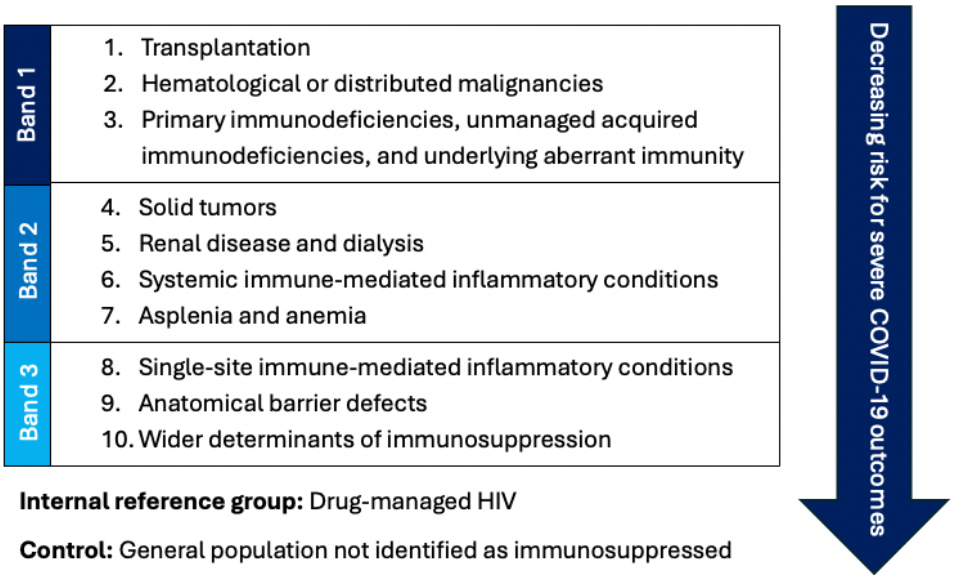

In our intended investigation, the conditions just cited will be allocated into these categories by relevant experts within the DESTINIES Consortium. Multiple rounds of consensus building will take place if necessary, with 75% agreement as the threshold for consensus in the first round (corresponding with the median consensus value of Diamond et al [16] review of more than 100

XSL•FO
RenderX

successful exercises) and, for any diagnoses that could not be allocated to this consensus threshold, a second round of majority-based allocations will be conducted. Determinations will be made anonymously to reduce risks of bias: this blinding minimizes the influence of panelists' reputations, personalities, or propensity to defer to group averages [17]. A third IMID category, overlap IMID, will be instituted for survey respondents who self-report both single organ and systemic diseases.

It is anticipated that small sample sizes may preclude detailed comparative analysis, particularly after 3-part IMID segmentation. Incomplete surveys will also be removed from the study set, including those in which respondents do not indicate their IMID status or provide full pharmacovigilance data. This includes respondents who have not received a vaccine dose.

However, statistical power calculations, including minimum sample sizing, were not performed because of the highly exploratory nature of this exercise. The diversity of peer-reviewed outputs that this dataset has produced—including those exclusively focused on subgroup experiences (eg, pregnant or breastfeeding women with autoimmunity) [18]—gives the authors confidence that this work will represent a meaningful contribution to the field.

Furthermore, efforts will be made to preserve usable data where possible, imputing "No response" if questions are skipped. Free-text data, such as that clarifying "Other" selections, will be removed. This includes immune-mediated conditions that fall outside of the 22 diagnoses specified. Google Translate (Alphabet Inc) will be used to ensure eligible diagnoses provided in other languages are not omitted.

To investigate our hypothesis that patients with overlap and systemic IMID experience worse COVID-19 infection and vaccine outcomes than their single organ and general population counterparts, the following analyses will be conducted.

Descriptive Analysis

Cohort profiling will tabulate and compare IMID categories and the control population by age, gender, ethnicity, country economic status, and self-reported health and physical function levels. Country economic status will be operationalized using World Bank Country classifications [19], whereby "Higher" will refer to countries demarcated as high-income, and "Lower" will refer to those classified as either low, lower-middle, or upper-middle income, accurate as of the 2025 fiscal year. IMID data will be compared with general population controls as an aggregate before being differentiated into systemic, single organ, and overlap categories. COVID-19 vaccine history, including dose count and type, will also be reported. Absolute numbers of COVID-19 infections, hospitalizations (with and without oxygen support), and average case duration and vaccine side effects (major vs mild) will also be tabulated with their supporting summary data. For completeness, efforts will be made to tabulate the frequencies of each specific COVID-19 symptom and vaccine side effect reported and their respective proportions of totals. For vaccine data, results will be presented by each vaccine type (mRNA, adenoviral, inactivated virus, protein subunit, mixed schedules, and unsure) to identify those associated with disproportionately high numbers of side effects or muted outcomes by comparison.

Anticipating small sample sizes, this study will preserve all usable data by not performing age-gender matching. However, as elaborated upon further below, multivariable logistic regression will assess the relative contributions of key demographic and medical variables to the likelihood of a participant reporting hospitalization for COVID-19 or vaccine side effects.

Simple Logistic Regression for COVID-19 Outcomes

Obtained via simple logistic regression, odds ratios of each outcome of interest will be estimated against the control group reference with their respective 95% CI to assess the influence of study groups on each event. Outcomes of interest include reporting a major vaccine side effect (any, then by each side effect type), reporting a minor vaccine side effect (any, then by each side effect type), COVID-19 infection, COVID-19 hospitalization without oxygen support, and COVID-19 hospitalization with oxygen support.

Anticipating the prospective contributions of vaccine uptake and wider medication and income variables to clinical outcomes, we will perform a variety of subgroup analyses. The above exercise will be repeated but with data segmented by vaccine dose (1 or 2); medication status (received steroid immunosuppressives, did not receive steroid immunosuppressives, received nonsteroid immunosuppressives, or did not receive nonsteroid immunosuppressives); and economic status (higher or lower). For immunosuppressive medication-based subgroup analyses, odds ratios will be calculated using single organ IMID data as the reference category.

Multivariable Logistic Regression for Assessing the Influence of Covariates

Multivariable logistic regressions will be used to quantify the relative contributions of demographic, health, and vaccination variables on the likelihood of participants self-reporting (1) severe COVID-19 (hospitalization or hospitalization with oxygen support) and (2) vaccine side effects (mild or major) for each population. Survey responses that omit information on vaccine side effects or hospitalization will be withheld from this analysis.

For COVID-19 hospitalization, the following models will be run at the whole cohort level and then separately for IMID (aggregated and categorized) and general population data.

- Model 1 (demographics): age, gender, ethnicity, and country economic status
- Model 2 (demographics+prior health): age, gender, ethnicity, country economic status, health status, and medication status
- Model 3 (demographics+prior health+vaccination): age, gender, ethnicity, country economic status, health status, medication status, and vaccine dose

For COVID-19 vaccine side effects, the following models will be run at the whole cohort level and then separately for IMID (aggregated and categorized) and general population data.

XSL•FO
RenderX

- Model 1 (demographics): age, gender, ethnicity, and country economic status
- Model 2 (demographics+prior health): age, gender, ethnicity, country economic status, health status, and medication status
- Model 3 (demographics+prior health+vaccination): age, gender, ethnicity, country economic status, health status, medication status, vaccine dose, prevaccine symptoms, and medication suspensions

Categorical variables (eg, ethnicity, gender, and self-reported ability to complete daily tasks) will have each of their levels' influence on COVID-19 vaccine side effect reporting or hospitalization analyzed in turn. This will be assessed by quantifying the impact of changing a category's reference level (eg, Caucasian [White] for ethnicity data) to test level (eg, Asian) on the likelihood of either event occurring. The reference levels for each categorical variable will be prespecified. The influence of study category on COVID-19 hospitalization and vaccine side effect outcomes will be estimated at the whole cohort and IMID (aggregated) levels, using general population data as the reference in the former instance and single organ IMID data as the reference in the latter instance. All coefficients will be exponentiated and reported as odds ratios with their associated 95% CIs. All findings will be interpreted with the caveats of selection and recall bias that cannot be controlled for in this secondary analysis.

Ethical Considerations

As a secondary analysis of COVAD-1 data, this protocol adheres to existing COVAD-1 ethical approvals including local institutional ethics committees [20]—Institutional Ethics Committee of Sanjay Gandhi Postgraduate Institute of Medical Sciences (IEC 202-143-IP-EXP-39) [19]—and the Checklist for Reporting Results of Internet E-Surveys [21], which can be accessed in the COVAD protocol [12]. It has the ongoing oversight, endorsement, and participation of the entire COVAD steering group. This protocol pertains to the secondary analysis of anonymized data. Informed consent was obtained before survey completion.

Results

As a collective, this analysis will identify key demographic differences between IMID categories and general population controls, quantify differences in the COVID-19 infection and vaccine experiences of these groups, and delineate the contributions of numerous demographic and health variables to each group's likelihood of being hospitalized for COVID-19 or experiencing a side effect after inoculation.

Over the course of the peer review process, this protocol has since been actioned in full; results are now being quality assured by the present authorship. Work began in August 2024, data analysis occurred between November 2024 and July 2025, and journal submission is intended for October 2025. However, this paper will elaborate on results initially presented in this protocol's preprint edition.

In November 2024, 7 IMID experts within the DESTINIES Consortium were enlisted to categorize COVAD 1 respondents' diagnoses into their respective study categories (single organ IMID, systemic IMID, or control). Two rounds of consensus building were necessary for this process. However, 7 diagnoses did not meet the initial 75% threshold specified and were entered into the second round of consensus building, where allocations were made by majority vote. The results of these 2 rounds and final categorization are provided in Tables 1 and 2.

Final allocations are provided in Textbox 1. This directed subsequent data cleaning and analysis as specified in the Methods section (active December 2024 to July 2025). Respondents who self-identified as having both single organ and systemic diagnoses according to these categorizations were then allocated into the overlap group.

XSL•FO
RenderX

Table 1. Round 1 of consensus building (N=7).

Diagnosis	Single organ, n (%)	Systemic, n (%)	Categorization
Ankylosing spondylitis or psoriatic arthritis	1 (14)	6 (86)	Systemic
Antisynthetase syndrome	0 (0)	7 (100)	Systemic
Crohn disease or ulcerative colitis (inflammatory bowel disease)	2 (29)	5 (71)	Contested
Dermatomyositis	0 (0)	7 (100)	Systemic
Hemolytic anemia or ITP[a]	4 (57)	3 (43)	Contested
Inclusion body myositis	4 (57)	3 (43)	Contested
Juvenile dermatomyositis	0 (0)	7 (100)	Systemic
Mixed connective tissue disease	0 (0)	7 (100)	Systemic
Multiple sclerosis	3 (43)	4 (57)	Contested
Myasthenia gravis	3 (43)	4 (57)	Contested
Necrotizing myositis	3 (43)	4 (57)	Contested
Overlap myositis with lupus or Sjögren syndrome or systemic sclerosis or rheumatoid arthritis	0 (0)	7 (100)	Systemic
Pernicious anemia	4 (57)	3 (43)	Contested
Polymyalgia rheumatica	0 (0)	7 (100)	Systemic
Polymyositis	0 (0)	7 (100)	Systemic
Rheumatoid arthritis	0 (0)	7 (100)	Systemic
Scleroderma	1 (14)	6 (86)	Systemic
Sjögren syndrome	0 (0)	7 (100)	Systemic
Systemic lupus erythematosus	0 (0)	7 (100)	Systemic
Thyroid (hypothyroid or hyperthyroid)	7 (100)	0 (0)	Single organ
Type 1 diabetes	7 (100)	0 (0)	Single organ
Vasculitis	0 (0)	7 (100)	Systemic

[a]ITP: idiopathic thrombocytopenic purpura.

Table 2. Round 2 of consensus building (N=7).

Diagnosis	Single organ, n (%)	Systemic, n (%)	Categorization
Crohn disease or ulcerative colitis (inflammatory bowel disease)	2 (29)	5 (71)	Systemic
Inclusion body myositis	4 (57)	3 (43)	Single organ
Multiple sclerosis	5 (71)	2 (29)	Single organ
Hemolytic anemia or ITP[a]	7 (100)	0 (0)	Single organ
Myasthenia gravis	6 (88)	1 (14)	Systemic
Necrotizing myositis	2 (29)	5 (71)	Systemic
Pernicious anemia	6 (86)	1 (14)	Single organ

[a]ITP: idiopathic thrombocytopenic purpura.

XSL•FO
RenderX

Textbox 1. Final single organ and systemic immune-mediated inflammatory disease allocations.

Single organ immune-mediated inflammatory disease

- Type 1 diabetes
- Thyroid (hypothyroid or hyperthyroid)
- Myasthenia gravis
- Multiple sclerosis
- Pernicious anemia
- Hemolytic anemia or idiopathic thrombocytopenic purpura
- Inclusion body myositis

Systemic immune-mediated inflammatory disease

- Scleroderma
- Rheumatoid arthritis
- Sjögren syndrome
- Systemic lupus erythematosus
- Overlap myositis with lupus, Sjögren syndrome, systemic sclerosis, or rheumatoid arthritis
- Polymyalgia rheumatica
- Ankylosing spondylitis or psoriatic arthritis
- Vasculitis
- Mixed connective tissue disease
- Antisynthetase syndrome
- Juvenile dermatomyositis
- Dermatomyositis
- Polymyositis
- Crohn disease or ulcerative colitis (inflammatory bowel disease)
- Necrotizing myositis

Discussion

Anticipated Findings

Irrespective of whether our data are consistent with the hypotheses specified, an analysis such as this provides valuable insights for clinicians, policymakers, and pharmaceutical firms. For example, any indications of elevated vaccine side effects among specific IMID types, even historical ones, warrant close examination. Alternatively, indications that the COVID-19 vaccine was universally well tolerated among participants with IMID should be made known to patients concerned with taking up this offering going forward. Such evidence would be especially pressing in a context of repeat vaccination, where the risks of experiencing adverse events can accumulate [22].

This direct-to-patient reporting pathway holds immense promise for pharmacovigilance going forward. The COVAD study group has already used this dataset to interrogate multiple longstanding questions in the vaccine benefit-risk remit for patients with IMIDs [23]. It is unique in amplifying the patient voice that is currently missing from pharmacovigilance; by enhancing the granularity of this data, as we intend, this patient survey offers intelligence that could not be procured through routine surveillance alone.

Limitations

We acknowledge the following limitations in this protocol. First, the COVID-19 vaccine landscape has changed considerably because the COVAD-1 data were collected. At the time of writing, vaccine uptake in certain immunosuppressed categories has since exceeded 10 doses, for example [24]. In comparison, 2 doses are the maximum within this COVAD-1 dataset. This discrepancy could call the relevance of our findings into question; however, their prospective value for enhancing pharmacovigilance and pandemic preparedness efforts counters this assertion.

Second, these data were not collected in real time with vaccination; an accurate picture of 7-day side effects therefore relies on respondents' recall. To prevent false recall bias, prospective data would have been preferable.

Third, despite the international nature of this work, the sample size is still relatively small—especially once incomplete surveys and the unvaccinated are excluded from the study set and the IMID population is segmented. The imbalanced allocations of

XSL•FO
RenderX

IMID diagnoses between single organ and systemic IMID categories are also of concern.

Furthermore, the authors acknowledge that this protocol does little to control for various forms of bias beyond the anonymization of the consensus exercises specified. For example, panelists may have privileged the risk profiling of the patients they are most familiar with, allocating them into the higher risk systemic IMID category—at least by the assessment of the DESTINIES Study [13]. In addition, there was no ability to control for the selection bias inherent to primarily recruiting participants from medical settings; those sick enough to require regular medical supervision may not be representative of the average patient in that population—especially when considered from a perspective of developing economies lens, where medical access is scarcer and can select for the more affluent [25].

Finally, this protocol does not offer any means of controlling for the potential influence of changing COVID-19 variants on infection outcomes. Surveys were completed between March and December 2021, coinciding with the emergence of the Omicron lineage that has since remained dominant [26]. Similarly, any changes in vaccine practices that occurred during this data collection period—such as local preferences for homologous or mixed scheduling [27]—cannot be controlled for. In either case, without survey completion or vaccine administration date data, we are unable to segment the analysis to assess the influence of these changes. We acknowledge that this is detrimental to the rigor of this work.

Conclusions

At present, patients with IMID are not provided with vaccine safety data that are relevant to their diagnosis. This opacity has been seen to create vaccine hesitancy despite the well-established vulnerability of these patients. This investigation seeks to enhance COVID-19 pharmacovigilance for the IMID population. If successful, this work will corroborate the value of direct-to-patient reporting pathways for vaccine benefit-risk surveillance in complex groups.

Acknowledgments

The authors are grateful to all respondents for completing the questionnaire. The authors also thank the Myositis Association, Myositis India, Myositis UK, Myositis Support and Understanding, the Myositis Global Network, Deutsche Gesellschaft für Muskelkranke e.V. (DGM), Dutch and Swedish Myositis patient support groups, Cure JM, Cure IBM, Sjögren's India Foundation, Patients Engage, Scleroderma India, Lupus UK, Lupus Sweden, Emirates Arthritis Foundation, EULAR PARE, ArLAR research group, AAAA patient group, Myositis Association of Australia, APLAR myositis special interest group, Thai Rheumatism association, PANLAR, AFLAR NRAS, Anti-Synthetase Syndrome support group, and various other patient support groups and organizations for their contribution to the dissemination of this survey. The authors wish to thank all members of the COVAD study group for their invaluable role in the data collection.

COVAD Study Group Authors: Steering committee: Ai Lyn Tan, Samuel Katsuyuki Shinjo, Tsvetelina Velikova, Abraham Edgar Gracia, Masataka Kuwana, Guochun Wang, Johannes Knitza, Carlos Enrique Toro Gutiérrez, Carlo Vinicio Caballero-Uribe, Jessica Day, Francis Guillemin, Nicola Dalbeth, Gerd-Rüdiger Burmester, Lorenzo Cavagna. Elena Nikiphorou.

CoVAD Study group members: Sreoshy Saha, Jasmine Parihar, Kunal Chandwar, Akanksha Ghodke, Mahabaleshwar Mamadapur, Prithvi Sanjeevkumar Gaur, Arunkumar R Pande, Anuj Shukla, Keerthi Talari Bommakanti, Ajay Asranna, Padnamabha Shenoy, Ashish Jacob Mathew, Abhra Chandra Chowdhury, Rudra Prosad Goswami, Sanat Phatak, Sajal Ajmani, Pradeepta Sekhar Patro, Chengappa G Kavadichanda, Vikas Gupta, Manesh Manoj M, Aman Sharma, Rakesh Kumar Pilania, Sapan C Pandya, Avinash Jain, Rajiv Ranjan, Yogesh Preet Singh, Bhupen Barman, Naveen R, Armen Yuri Gasparyan, Arvind Nune, James B Lilleker, John D Pauling, Chris Wincup, Minchul Kim, Ashima Makol, Tulika Chatterjee, Aarat Patel, Döndü Üsküdar Cansu, Babur Salim, Zoha Zahid Fazal, Lisa S Traboco, Syahrul Sazliyana Shaharir, Raquel Aranega, Jesús Loarce-Martos, Sergio Prieto-González, Albert Gil-Vila, Albert Selva O'Callaghan, Leonardo Santos Hoff, Marcin Milchert, Margarita Aleksandrovna Gromova, Melinda Nagy-Vincze, Zoltán Griger, Dzifa Dey, Marie Hudson, Océane Landon-Cardinal, Russka Shumnalieva, Manuel Francisco Ugarte-Gil, Lyn Chinchay, José Proaño Bernaola, Victorio Pimentel, A.T.M. Tanveer Hasan, Binit Vaidya, Tamer A Gheita, Hanan Mohamed Fathi, Reem Hamdy A Mohammed, Miguel A Saavedra, Erick Adrian Zamora Tehozol, Jorge Rojas Serrano, Ignacio García-De La Torre,, Iris J. Colunga-Pedraza, Javier Merayo-Chalico, Ghita Harifi, Lina El Kibbi, Hussein Halabi, Phonpen Akawatcharangura Goo, Wanruchada Katchamart, Yurilís Fuentes-Silva, Lilith Stange Nunez, Cristian Vergara M, Jossiell Then Báez, Rodrigo García Salinas, Wendy Calapaqui, Ivonne Quezada, Stylianos Tomaras, Suryo Anggoro Kusumo Wibowo, Mr Or Aharonov, Nicoletta Del Papa, Gianluca Sambataro, Atzeni Fabiola, Marcello Govoni, Simone Parisi, Elena Bartoloni Bocci, Gian Domenico Sebastiani, Enrico Fusaro, Marco Sebastiani, Luca Quartuccio, Franco Franceschini, Pier Paolo Sainaghi, Giovanni Orsolini, Rossella De Angelis, Maria Giovanna Danielli, Vincenzo Venerito, Ran Nakashima, Shinji Sato, Naoki Kimura, Yuko Kaneko, Daman Langguth, Vidya Limaye, Merrilee Needham, Dr Nilesh Srivastav, Oliver Distler.

Other CoVAD investigators: Margherita Giannini, François Maurier, Alain Meyer, Uyi Ima-Edomwonyi, Uyi Ima-Edomwonyi, Emorinken Airenakho, Nwankwo Henry Madu, Abubakar Yerima, Hakeem Olaosebikan, Celestine Chibuzo Okwara, Becky A, Ouma Devi Koussougbo, Javier Cajas, Álvaro Arbeláez, Wilmer Gerardo Rojas Zuleta. Andrea Bran Ordóñez, Nilmo Chávez, Alejandro Quiñónez Obiols, Antonio Cachafeiro-Vilar, Generoso Guerra Bautista, Enrique Julio Giraldo Ho, Ihsane Hmamouchi, Pr Imane El bouchti, Zineb Baba, Dina Arrieta, Eduardo Romero Hidalgo, Ricardo Saenz, José António Pereira Silva, João Eurico Fonseca, Karoll Cabriza, Jonathan Losanto, Nelly Colaman, Hugo Alonzo, Carlos Benito Santiago Pastelin.

XSL•FO
RenderX

Finally, the authors wish to acknowledge the role of the DESTINIES Consortium in producing the IMID categorisations featured in this work. DESTINIES Consortium members that wished to be named include: Agustin Portela, Alessia Alunno, Camille Kotton, Carl Goodyear, Carol Coupland, Christopher Griffiths, David Lowe, David Lye Chien Boon, Emma Thomson, Féline Kroon, Francois Raffi, Hana M. El Sahly, Iain Buchan, John Isaacs, Kevin Winthrop, Lakshmi Panagiotakopoulos, Lisa Spencer, Martin Underwood, Michael Hoerger, Michelle Willicombe, Monique Andersson, Richard Conway, Robert Landewe, Stanley Perlman, Stefan Siebert, Thushan de Silva, Wilbur Chen, Yasmin Maor.

Funding

This secondary analysis was supported through an Industrial Cooperative Awards in Science and Engineering (iCASE) Doctoral Fellowship between the UK Medical Research Council and Optum UK (formerly EMIS Health).

Data Availability

COVID-19 Vaccination in Autoimmune Diseases (COVAD) datasets are not available for public access. Those interested in conducting their own secondary investigations must apply to do so. The COVAD steering group handles these applications on a case-by-case basis.

Authors' Contributions

Conceptualization: ML, LG
Methodology: ML, JMOM, XG, MJ, FDRH, SdL, RC, JI, JC, AA, SS, OE, KS
Writing – original draft: ML
Writing – review & editing: all authors

Conflicts of Interest

LA has received consultancy fees from Abbott, Pfizer, UCB. IP has received research funding and/or honoraria from Amgen, AstraZeneca, Aurinia, BMS, Elli Lilly, Gilead, GSK, Janssen, Novartis, Otsuka, and Roche. SS has received institutional research funding from Eli Lilly, GSK, Janssen, Pfizer and UCB; has received consulting fees for AbbVie, Amgen, AstraZeneca, Janssen, Syncona and Teijin Pharma; has been an unpaid advisory member of the UK McInnes COVID-19 Neutralising Monoclonal Antibodies (nMABs) and Antivirals Access Independent Advisory Group 2021–2023. ML's academic contribution to this work was provided while she was affiliated with the University of Oxford. SdL has received funding for vaccine related work from AstraZeneca, GSK, Moderna, Pfizer, Sanofi, and Seqirus; has been a member of advisory boards for GSK, Sanofi, and Seqirus; has been funded for conference travel or received a speaking fee from AstraZeneca and Moderna. JC has received consulting fees and/or research grants from Pfizer, GSK, and Moderna. All other authors declare no conflicts of interest.

References

1. Tan EH, Sena AG, Prats-Uribe A, You SC, Ahmed WU, Kostka K, et al. COVID-19 in patients with autoimmune diseases: characteristics and outcomes in a multinational network of cohorts across three countries. Rheumatology 2021 Oct;60(SI):SI37-SI50 [FREE Full text] [doi: 10.1093/rheumatology/keab250] [Medline: 33725121]
2. COVID-19: guidance for people whose immune system means they are at higher risk. Government of UK. URL: https://tinyurl.com/3aus7seh [accessed 2025-09-16]
3. Altuwairqi A, Ali AH, Alariefy AA, Bahlas S, AlZahrani SK, Zarei EW, et al. Assessment of COVID-19 morbidity and mortality among patients with autoimmune diseases at King Abdulaziz University Hospital. Cureus 2024 Jan;16(1):e52492 [FREE Full text] [doi: 10.7759/cureus.52492] [Medline: 38370992]
4. MacKenna B, Kennedy NA, Mehrkar A, Rowan A, Galloway J, Matthewman J, et al. Risk of severe COVID-19 outcomes associated with immune-mediated inflammatory diseases and immune-modifying therapies: a nationwide cohort study in the OpenSAFELY platform. Lancet Rheumatol 2022 Jul;4(7):e490-e506 [FREE Full text] [doi: 10.1016/S2665-9913(22)00098-4] [Medline: 35698725]
5. Felten R, Dubois M, Ugarte-Gil MF, Chaudier A, Kawka L, Bergier H, et al. Vaccination against COVID-19: expectations and concerns of patients with autoimmune and rheumatic diseases. Lancet Rheumatol 2021 Apr;3(4):e243-e245 [FREE Full text] [doi: 10.1016/S2665-9913(21)00039-4] [Medline: 33655219]
6. Bukan K, Pearce-Slade T, Eiberg M, Tinelli M, Yahav D, Tuells J, et al. Exclusion of older adults and immunocompromised individuals in influenza, pneumococcal and COVID-19 vaccine trials before and after the COVID-19 pandemic. Aging Clin Exp Res 2023 May;35(5):917-923 [FREE Full text] [doi: 10.1007/s40520-023-02380-4] [Medline: 37027085]
7. Trøseid M, Hentzien M, Ader F, Cardoso SW, Arribas JR, Molina J, Eu Response, COMBINE. Immunocompromised patients have been neglected in COVID-19 trials: a call for action. Clin Microbiol Infect 2022 Sep;28(9):1182-1183 [FREE Full text] [doi: 10.1016/j.cmi.2022.05.005] [Medline: 35623577]

XSL•FO
RenderX

8. Aberumand B, Ayoub Goulstone W, Betschel S. Understanding attitudes and obstacles to vaccination against COVID-19 in patients with primary immunodeficiency. Allergy Asthma Clin Immunol 2022 May 09;18(1):38 [FREE Full text] [doi: 10.1186/s13223-022-00679-x] [Medline: 35534860]
9. Gaur P, Agrawat H, Shukla A. COVID-19 vaccine hesitancy in patients with systemic autoimmune rheumatic disease: an interview-based survey. Rheumatol Int 2021 Sep;41(9):1601-1605 [FREE Full text] [doi: 10.1007/s00296-021-04938-9] [Medline: 34213580]
10. Leston M, de Lusignan S, Hobbs FD. A targeted approach to vaccine hesitancy. Oxf Open Immunol 2023;4(1):iqad007 [FREE Full text] [doi: 10.1093/oxfimm/iqad007] [Medline: 38089637]
11. Avery AJ, Anderson C, Bond CM, Fortnum H, Gifford A, Hannaford PC, et al. Evaluation of patient reporting of adverse drug reactions to the UK 'Yellow Card Scheme': literature review, descriptive and qualitative analyses, and questionnaire surveys. Health Technol Assess 2011 May;15(20):1-iv [FREE Full text] [doi: 10.3310/hta15200] [Medline: 21545758]
12. Sen P, Gupta L, Lilleker JB, Aggarwal V, Kardes S, Milchert M, COVAD Study Group. COVID-19 vaccination in autoimmune disease (COVAD) survey protocol. Rheumatol Int 2022 Jan;42(1):23-29 [FREE Full text] [doi: 10.1007/s00296-021-05046-4] [Medline: 34779868]
13. Leston M, Ordóñez-Mena JM, Joy M, Hobbs FD, de Lusignan S, Teh BW, DESTINIES Consortium. The DESTINIES Study: an online Delphi study to build international consensus on the medical conditions and procedures that confer immunosuppression and their respective COVID-19 risk profiles. EClinicalMedicine 2025 May;83:103239 [FREE Full text] [doi: 10.1016/j.eclinm.2025.103239] [Medline: 40453534]
14. Naveen R, Parodis I, Joshi M, Sen P, Lindblom J, Agarwal V, COVAD Study Group. COVID-19 vaccination in autoimmune diseases (COVAD) study: vaccine safety and tolerance in rheumatoid arthritis. Rheumatology 2023 Jul 05;62(7):2366-2376 [FREE Full text] [doi: 10.1093/rheumatology/keac624] [Medline: 36315075]
15. Ramos-Casals M, Brito-Zerón P, Mariette X. Systemic and organ-specific immune-related manifestations of COVID-19. Nat Rev Rheumatol 2021 Jun;17(6):315-332 [FREE Full text] [doi: 10.1038/s41584-021-00608-z] [Medline: 33903743]
16. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 2014 Apr;67(4):401-409. [doi: 10.1016/j.jclinepi.2013.12.002] [Medline: 24581294]
17. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. World J Methodol 2021 Jul 20;11(4):116-129 [FREE Full text] [doi: 10.5662/wjm.v11.i4.116] [Medline: 34322364]
18. Andreoli L, Lini D, Schreiber K, Parodis I, Sen P, Ravichandran N, COVAD Study Group. COVID-19 vaccine safety during pregnancy and breastfeeding in women with autoimmune diseases: results from the COVAD study. Rheumatology (Oxford) 2024 May 02;63(5):1341-1351 [FREE Full text] [doi: 10.1093/rheumatology/kead382] [Medline: 37505460]
19. World Bank country and lending groups. World Bank. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [accessed 2024-10-16]
20. Mathur R. National Ethical Guidelines for Biomedical and Health Research Involving Human Participants. New Delhi, India: Indian Council of Medical Research; Oct 2017.
21. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
22. Frasca L, Ocone G, Palazzo R. Safety of COVID-19 vaccines in patients with autoimmune diseases, in patients with cardiac issues, and in the healthy population. Pathogens 2023 Feb 02;12(2):233 [FREE Full text] [doi: 10.3390/pathogens12020233] [Medline: 36839505]
23. COVAD study group's scientific contributions. ResearchGate. URL: https://www.researchgate.net/scientific-contributions/COVAD-Study-Group-2226638050 [accessed 2024-10-13]
24. Autumn/winter COVID-19 vaccination program 2025. Immunodeficiency UK. URL: https://www.immunodeficiencyuk.org/immunodeficiency/management/covid-19/covid19vaccinesandimmnodeficiency/ [accessed 2025-09-16]
25. Ailshire JA, Herrera CA, Choi E, Osuna M, Suzuki E. Cross-national differences in wealth inequality in health services and caregiving used near the end of life. EClinicalMedicine 2023 Apr;58:101911 [FREE Full text] [doi: 10.1016/j.eclinm.2023.101911] [Medline: 36969343]
26. SARS-CoV-2 variants of concern as of 6 May 2021. European Centre for Disease Prevention and Control. URL: https://www.ecdc.europa.eu/en/covid-19/variants-concern [accessed 2025-09-16]
27. Chiu NC, Chi H, Tu YK, Huang YN, Tai YL, Weng SL, et al. To mix or not to mix? A rapid systematic review of heterologous prime-boost COVID-19 vaccination. Expert Rev Vaccines 2021 Oct;20(10):1211-1220 [FREE Full text] [doi: 10.1080/14760584.2021.1971522] [Medline: 34415818]

Abbreviations

COVAD: COVID-19 Vaccination in Autoimmune Diseases
DESTINIES: Delphi Study to Define and Risk-Stratify Immunosuppression
IMID: immune-mediated inflammatory disease

XSL•FO
RenderX

Edited by J Sarvestan; submitted 13.Nov.2024; peer-reviewed by GS Pileggi, YM Hwang, A Bhat; comments to author 30.Mar.2025; revised version received 28.Aug.2025; accepted 15.Sep.2025; published 27.Feb.2026.

Please cite as:
Leston M, Ordóñez-Mena JM, Gu X, Joy M, Hobbs FDR, Y W Lee L, Parodis I, Andreoli L, Ziade N, Gupta L, Agarwal V, Conway R, Isaacs J, Curtis J, Alunno A, Siebert S, Elkayam O, Schreiber K, de Lusignan S
Differentiating the COVID-19 Infection and Vaccine Experiences of Patients With Systemic, Single Organ, and Overlap Immune-Mediated Inflammatory Disease: Protocol for a Secondary Analysis for Enhancing COVID-19 Vaccine Pharmacovigilance
JMIR Res Protoc 2026;15:e68785
URL: https://www.researchprotocols.org/2026/1/e68785
doi:10.2196/68785
PMID:

©Meredith Leston, José M Ordóñez-Mena, Xinchun Gu, Mark Joy, F D Richard Hobbs, Lennard Y W Lee, Ioannis Parodis, Laura Andreoli, Nelly Ziade, Latika Gupta, Vikas Agarwal, Richard Conway, John Isaacs, Jeffrey Curtis, Alessia Alunno, Stefan Siebert, Ori Elkayam, Karen Schreiber, Simon de Lusignan. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Evaluation of the Antimicrobial Efficacy of Cerium Oxide Nanoparticles and Chlorhexidine Gluconate Against Enterococcus faecalis in Root Canal Systems: Protocol for an In Vitro Study

Bhuvaneshwari Shrikant Khanadali[1], BDS; Anuja Ikhar[1], BDS, MDS, PhD

Department of Conservative Dentistry and Endodontics, Sharad Pawar Dental College, Datta Meghe Institute of Medical Sciences, Sawangi (Meghe), Maharashtra, India

Corresponding Author:
Bhuvaneshwari Shrikant Khanadali, BDS
Department of Conservative Dentistry and Endodontics
Sharad Pawar Dental College
Datta Meghe Institute of Medical Sciences
DMIHER Campus
Sawangi (Meghe), Maharashtra, 442001
India
Phone: 91 06363941284
Email: khnadalibhuvana@gmail.com

Abstract

Background: *Enterococcus faecalis* is a facultative anaerobic bacterium frequently associated with persistent root canal infections and endodontic treatment failures. Its resistance is primarily attributed to its ability to form robust biofilms, survive under harsh conditions, and penetrate deep into dentinal tubules. Chlorhexidine gluconate (CHX), commonly used as an intracanal medicament, exhibits broad-spectrum antimicrobial activity; however, its limited ability to eliminate biofilms, potential cytotoxicity, and restricted dentinal penetration pose significant drawbacks. In recent years, nanotechnology has introduced promising alternatives such as cerium oxide nanoparticles (CeO_2-NPs), which exhibit unique redox properties, reactive oxygen species scavenging, and antimicrobial action due to their nanoscale size and surface chemistry.

Objective: This study aims to evaluate and compare the antimicrobial efficacy of 0.05% CeO_2-NPs, 2% CHX, and their combination against *E. faecalis* in root canal systems.

Methods: Forty-eight extracted human single-rooted premolars will be decoronated, instrumented, and sterilized. The root canals will be inoculated with *E. faecalis* and incubated for 21 days to allow biofilm formation. The specimens will then be divided into four groups (n=12 per group), each receiving one of the following treatments: (1) 0.05% CeO_2-NPs, (2) 2% CHX, (3) a combination of CeO_2-NPs and CHX, and (4) sterile water as the negative control. After the medicaments are applied for 5 days, microbial sampling will be conducted, and the remaining viable bacteria will be quantified using colony-forming unit (CFU) counts on selective agar.

Results: Ethics approval was obtained from the institutional ethics committee of Datta Meghe Institute of Higher Education and Research, Wardha (approval DMIHER(DU)/IEC/2025/546) in January 2025. As of January 2026, laboratory procedures, including specimen preparation, microbial inoculation, and data collection, have not yet commenced. Data collection is planned to begin in February 2026 and is expected to be completed by April 2026. Data analysis will follow immediately thereafter, with results anticipated to be submitted for publication in mid-2026. As this is an in vitro study, no participant recruitment, funding allocation, or clinical trial registration is applicable. It is expected that the combination of CeO_2-NPs and CHX will result in the greatest reduction in bacterial load, followed by CeO_2-NPs alone, CHX alone, and sterile water. These expected outcomes will be assessed based on comparative CFU counts for each treatment group.

Conclusions: This study aims to provide insight into the antimicrobial potential of CeO_2-NPs either alone or in combination with CHX against *E. faecalis*. If proven effective, CeO_2-NPs could represent a novel and biocompatible adjunct or alternative to conventional intracanal medicaments, contributing to improved disinfection protocols and better clinical outcomes in endodontic therapy.

International Registered Report Identifier (IRRID): PRR1-10.2196/77998

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e77998) doi:10.2196/77998

KEYWORDS

nanoparticles; intracanal medicament; chlorhexidine gluconate; Enterococcus faecalis; biofilm; cerium oxide nanoparticles; single-rooted premolar

Introduction

Background

Root canal treatment success depends on the complete elimination of microorganisms from the root canal system and the prevention of reinfection. Despite advancements in endodontic techniques, microorganisms such as *Enterococcus faecalis* frequently persist in the root canal, contributing to treatment failure [1]. *E. faecalis* is particularly resilient due to its ability to survive in nutrient-depleted environments, penetrate deeply into dentinal tubules, and form biofilms [2]. These biofilms protect the bacteria from antimicrobials, making eradication extremely challenging [3].

Chlorhexidine gluconate (CHX) has been a gold standard intracanal medicament in endodontics, valued for its broad-spectrum antimicrobial properties and ability to bind to dentin, providing long-lasting antibacterial effects [4]. However, CHX is not without limitations. Its cytotoxicity, potential to cause discoloration, and inability to penetrate deep into dentinal tubules have driven researchers to explore alternative options [5]. Cerium oxide nanoparticles (CeO_2-NPs) offer unique characteristics that distinguish them from conventional antimicrobial agents. These nanoparticles have dual oxidation states (Ce^{3+} and Ce^{4+}) that enable them to generate reactive oxygen species, disrupt bacterial biofilms, and exhibit potent antimicrobial activity [6]. Furthermore, their nanoscale size allows for enhanced penetration into dentinal tubules, which is crucial for addressing *E. faecalis* located in deeper layers of the tooth substrate [7]. Additionally, CeO_2-NPs demonstrate excellent biocompatibility, making them a promising candidate for clinical use [8].

This study addresses the persistent challenge of eliminating *E. faecalis*, a key contributor to root canal failure due to its resistance, dentinal invasion, and survival capabilities. Although CHX is a widely used intracanal medicament, its limitations, including cytotoxicity, discoloration, and limited penetration into dentinal tubules, necessitate alternative solutions [9].

CeO_2-NPs offer promising advantages, including reactive oxygen species generation, biofilm disruption, nanoscale penetration, and superior biocompatibility. A comparative evaluation of CeO_2-NPs and CHX is essential to determine their relative efficacy in targeting *E. faecalis* and improving clinical outcomes in endodontic treatments. Therefore, the aim of this study is to evaluate and compare the antimicrobial efficacy of 0.05% CeO_2-NPs, 2% CHX, and their combination against *E. faecalis* in an in vitro root canal infection model (extracted human teeth).

Objectives

The objectives of our study are as follows: (1) to assess the antimicrobial efficacy of 2% CHX as an intracanal medicament by measuring its ability to reduce the colony-forming units (CFUs) of *E. faecalis* in dentinal tubules, (2) to assess the antimicrobial efficacy of 0.05% CeO_2-NPs as an intracanal medicament by measuring its ability to reduce CFUs of *E. faecalis* in dentinal tubules, and (3) to assess the antimicrobial efficacy of 0.05% CeO_2-NPs and 2% CHX as an intracanal medicament by measuring its ability to reduce CFUs of *E. faecalis* in dentinal tubules.

Methods

Study Design and Settings

This study adopts an in vitro design to simulate and control the root canal environment for standardized and reproducible evaluation of antimicrobial efficacy. This design is appropriate, as it eliminates confounding variables present in clinical conditions and enables precise comparison among test groups. The selection of this design is based on feasibility, ethical considerations (no human/animal subjects involved), and the objective to isolate the effects of the medicaments on *E. faecalis*. This study was conducted in the Department of Conservative Dentistry and Endodontics, Sharad Pawar Dental College & Hospital, Sawangi (Meghe), Wardha.

Ethical Considerations

Ethical clearance for this in vitro study was obtained from the institutional ethics committee of Datta Meghe Institute of Higher Education and Research, Wardha, India (approval DMIHER(DU)/IEC/2025/546). All extracted teeth will be collected from the department of oral surgery, ensuring that they are not at risk of harm due to the use of teeth in research. This study will follow strict guidelines for ethical handling of human tissues. As this is an in vitro study, no participant recruitment, funding allocation, or clinical trial registration is applicable. Data will be stored on a password-protected computer accessible only to the principal investigator and supervisor. Data privacy and confidentiality will be maintained in accordance with institutional research policies.

Study Items

In this study, we use 48 freshly extracted, single-rooted human premolar teeth extracted for orthodontic or periodontal reasons. Single-rooted premolars provide a uniform root canal system ideal for standardizing canal instrumentation and medicament application. Extracted human teeth offer realistic dentinal tubule structures necessary for valid medicament assessment.

Inclusion Criteria

The inclusion criteria are as follows: (1) single-rooted human premolars with fully developed apices; (2) teeth with a single

XSL•FO
RenderX

canal confirmed radiographically; (3) teeth free from caries, restorations, cracks, fractures, or resorption; and (4) no prior endodontic treatment.

Exclusion Criteria

The exclusion criteria are as follows: (1) multi-rooted teeth or teeth with complex canal morphology, (2) teeth with calcified canals or accessory canals, and (3) teeth with structural defects (eg, resorption, fractures).

Sample Size

The sample size is calculated based on a previous study by Pandey et al [10] by using mean CFU values and SDs, as shown in Textbox 1.

Textbox 1. Sample size calculation.

Mean colony forming unit (CFU) values and SDs

Mean CFU in control = 40.545

Mean CFU in test group = 28.091

Standard deviation ($\sigma 1$) = 11.021

Standard deviation ($\sigma 2$) = 10.270

Desired difference (Δ) = 12.454

Power=80%

α=.05

Sample size calculation formula

$N = (Z_{1-\alpha/2} + Z_{1-\beta})^2(\sigma_1^2 + \sigma_2^2) / (\mu_1-\mu_2)^2$

The notation for the formulae are as follows:

- n_1 = sample size of group 1
- n_2 = sample size of group 2
- σ_1 = standard deviation of group 1
- σ_2 = standard deviation of group 2
- Δ = difference in group means
- κ = ratio = n_2 / n_1
- Z_1-$\alpha/_2$= 2-sided Z value (eg, Z = 1.96 for 95% CI)
- Z_1-β= power

For detecting the mean difference of 12.454, that is, Δ = 40.545-28.091 = 12.454

K = 1

$N = (11.021*11.021 + 28.091*28.091)(1.96+0.84)^2 / 12.454*12.454$

= 11.47 = 12 samples needed in each group (12*4=48)

Result: 12 samples per group × 4 groups = 48 total samples

Level of significance: 5% (95% CI)

Sampling Method and Group Allocation

Simple random sampling was performed from the pool of the eligible extracted teeth. The intervention for each group is allocated as shown in Table 1.

Table 1. Intervention for each group of eligible extracted teeth.

Group	Intervention	Samples, n
Group 1	2% chlorhexidine gluconate (CHX)	12
Group 2	0.05% cerium oxide nanoparticles (CeO_2-NPs)	12
Group 3	0.05% CeO_2-NPs + 2% CHX	12
Group 4	Sterile water (control)	12

XSL•FO
RenderX

Medicaments

The following medications are used.

1. CHX (2%)
2. CeO_2-NPs (0.05%): <25 nm particle size, 0.5 mg/mL dispersion in sterile distilled water, prepared using ultrasonication for 1 minute for uniform dispersion, and then a thickening agent methyl cellulose is added.
3. Combination group: equal volume of 0.5 mg/mL CeO_2-NPs and 2% CHX is mixed and ultrasonicated for 1 minute for uniform dispersion, and then a thickening agent methyl cellulose is added.

Tooth Preparation

The eligible extracted teeth are prepared as follows.

1. Decoronation: standardized to 6-mm root segments using a diamond disc (Mani Inc).
2. Canal preparation: Gates-Glidden drills (#3 and #4) for standardization of canal diameter.
3. Smear layer removal: 2 mL of 17% ethylenediaminetetraacetic acid (5 min) followed by 5 mL of 3% sodium hypochlorite (5 min) and then rinsed with 5 mL of sterile distilled water.
4. Sterilization: double autoclave cycle, including immersion in Mueller-Hinton broth for the second cycle.

Root Canal Infection

The root canals will be inoculated with *E. faecalis* and incubated for 21 days to allow biofilm formation. To induce root canal infection, the following agents are used: (1) *E. faecalis* (ATCC 29212) cultured in Mueller-Hinton broth and (2) dentin blocks incubated with 1 mL of 0.5 McFarland bacterial suspension for 21 days at 37 °C, with fresh inoculum every 48 hours.

Application of Medicaments

Medications are applied by introducing 100 µL of the medicine into the samples by using a 27-gauge syringe, and then the samples are sealed with sticky wax to simulate clinical conditions. Thereafter, the samples are incubated under anaerobic conditions at 37 °C for 5 days.

Posttreatment Sampling

Ater 5 days of incubation, dentin shavings will be collected at 400 µm depth by using Gates-Glidden drill size #5. The shavings will be suspended in Mueller-Hinton broth and serially diluted. Approximately 50 µL will be plated on Mueller-Hinton agar and incubated for 24 hours at 37 °C.

Colony Counting

CFUs are counted manually or using a digital colony counter (Remi Lab Instruments).

Outcome Measures

The primary outcome is quantitative comparison of the antimicrobial efficacy measured by CFU counts. The secondary outcome is the comparative effectiveness of the combination group vs individual agents and sterile water.

Data Management

All data collected during the study will be manually recorded in prestructured data collection sheets and later transferred into a secure digital database. Double data entry will be employed to minimize errors. Data will be stored on a password-protected computer accessible only to the principal investigator and supervisor. Data privacy and confidentiality will be maintained in accordance with institutional research policies.

Descriptive Statistics

The mean, standard deviation, and range will be calculated for CFU counts in each group. The graphical representation (bar graphs or box plots) will be used to visualize the microbial reduction across groups.

Inferential Statistics

To evaluate the null hypothesis of no significant difference among groups, inferential statistical analysis will be performed. The Shapiro-Wilk test will be used to assess the normality of the CFU data. If the data are nonnormally distributed, appropriate transformations or nonparametric alternatives will be applied. One-way ANOVA will be employed to compare mean CFU counts across the 4 groups, followed by Tukey Honest Significant Difference test for pairwise comparisons. Levene test will be used to assess the homogeneity of the variances, and if ANOVA assumptions are violated, the Kruskal-Wallis test will be considered. A *P* value of less than .05 will be considered statistically significant. All analyses will be conducted using SPSS Statistics for Windows (version 27.0; IBM Corp) and GraphPad Prism (version 7.0), which will also be used for data visualization and assumption verification.

Results

This study is in the protocol development phase, and no experimental results are available at this time. Following completion of the laboratory phase, the primary outcomes will include quantitative assessment of bacterial reduction based on CFU counts for each experimental group. Ethical approval was obtained from the institutional ethics committee of Datta Meghe Institute of Higher Education and Research, Wardha (approval DMIHER(DU)/IEC/2025/546) in January 2025. As of January 2026, laboratory procedures, including specimen preparation, microbial inoculation, and data collection, have not yet commenced. Data collection is planned to begin in February 2026 and is expected to be completed by April 2026. Data analysis will follow immediately thereafter, with results anticipated to be submitted for publication in mid-2026. The results will be analyzed to compare the antimicrobial efficacy of 2% CHX, CeO_2-NPs, their combination, and sterile water (control). Outcome data will be presented in both tabular and graphical formats, including groupwise mean CFU counts, SDs, and *P* values derived from intergroup comparisons.

Discussion

Persistent root canal infections, particularly those involving *E. faecalis*, remain a significant challenge in endodontic therapy. This is primarily attributed to the microbe's deep dentinal

XSL•FO
RenderX

penetration and resistance to standard antimicrobial agents. Despite advancements in mechanical instrumentation and irrigation protocols, complete disinfection of the root canal system is difficult to achieve, necessitating the exploration of more effective intracanal medicaments. A 21-day incubation period was selected to allow sufficient time for *E. faecalis* to colonize and penetrate the dentinal tubules, achieving a mature and resilient intraradicular infection model. Previous studies have demonstrated that *E. faecalis* requires extended incubation periods—typically ranging from 14 to 21 days—to reach substantial colonization and simulate persistent root canal infections [1,11]. Sanju et al [11] specifically used a 21-day incubation model to assess the antimicrobial effects of intracanal medicaments at deeper dentin levels, affirming that this duration provides consistent and clinically relevant bacterial invasion. This time frame enables the formation of bacterial populations that are more difficult to eradicate, thus providing a rigorous test for evaluating antimicrobial interventions.

Several recent studies have emphasized the potential of CeO_2-NPs as a promising antimicrobial agent. CeO_2-NPs have demonstrated significantly higher efficacy than calcium hydroxide, achieving a 66.9% reduction in *E. faecalis* CFUs at 400-µm dentin depth [11]. This is particularly relevant in endodontics, where bacteria often penetrate deeply into dentinal tubules. The enhanced antimicrobial action of CeO_2-NPs is attributed to their nanoscale size and surface reactivity, which allow for better penetration and biofilm disruption [11]. This study builds upon this foundational work by directly comparing the antimicrobial efficacy of 0.05% CeO_2-NPs, 2% CHX, their combination, and sterile water within an in vitro root canal biofilm model. The deep dentin effectiveness of 2% CHX gel has been previously confirmed, underscoring its value as a standard agent in endodontic disinfection [12]. However, [12] also highlighted its limitations, particularly in eliminating biofilm-residing bacteria. CeO_2-NPs have demonstrated broad-spectrum antimicrobial potential, attributed to their ability to generate reactive oxygen species and induce oxidative stress in microbial cells [13]. Their biofilm-disruption capacity has also been reported to be higher than that of CHX, prompting recommendations for further investigation into their combined use [14].

The potential value of combining CHX and CeO_2-NPs lies in their distinct yet potentially complementary antimicrobial mechanisms. CHX exerts its antimicrobial effect by disrupting bacterial cytoplasmic membranes, leading to leakage of intracellular components and cell death while also binding to dentin to provide prolonged antibacterial activity [4,5]. CeO_2-NPs, in contrast, operate via redox cycling between Ce^{3+} and Ce^{4+} oxidation states, facilitating the generation of reactive oxygen species and inducing oxidative stress within microbial cells [6,13]. Moreover, the nanoscale size of CeO_2-NPs enables deeper penetration into dentinal tubules, enhancing access to areas where conventional agents may be less effective [7,8]. When used in combination, these agents may exert additive or synergistic effects, CHX compromising membrane integrity and CeO_2-NPs amplifying oxidative damage thereby offering a multi-targeted antimicrobial approach. Clinically, such a combination could significantly improve root canal disinfection, especially against persistent organisms such as *E. faecalis* that inhabit deeper layers of dentin [1,2]. This approach may overcome the limitations of monotherapies and enhance long-term treatment outcomes. Consequently, evaluating this combination is both scientifically and clinically justified.

The anticipated outcomes of this study are expected to support and build upon previous findings in the field. By evaluating both individual and combined effects of CeO_2-NPs and CHX, this study explores possible synergistic interactions, a topic previously suggested but not thoroughly investigated in the context of dentin-embedded *E. faecalis* [14].

A key strength of this study is the use of extracted human teeth, offering a realistic anatomical model that closely resembles clinical conditions. This enhances the translational value of the findings. The controlled in vitro design minimizes external variables and allows precise comparison between the treatment groups. Furthermore, focusing on *E. faecalis*, a pathogen well-documented in endodontic failures, increases the clinical relevance of the study outcomes [12]. The investigation of 0.05% CeO_2-NPs as a novel nanotherapeutic agent adds innovation to the study, offering insights into future directions for endodontic disinfection strategies [11,13]. It is hypothesized that the CHX and CeO_2-NP combination could improve antimicrobial efficacy, justifying further experimental exploration [14]. Although some constraints are inherent to in vitro designs, such models provide a controlled environment ideal for evaluating antimicrobial efficacy in a reproducible and standardized manner. This foundational research supports the potential of CeO_2-NPs—alone or in combination with CHX—as effective agents against resistant endodontic pathogens and sets the stage for future in vivo investigations.

Although this study employs a controlled and standardized in vitro design to evaluate the antimicrobial efficacy of CeO_2-NPs and CHX, it does not fully replicate the complex biological conditions present within the human oral cavity. The absence of host immune responses, salivary enzymes, and polymicrobial communities may influence bacterial behavior differently in vivo. Furthermore, factors such as biofilm heterogeneity, fluid dynamics, and interactions with periapical tissues are not represented in this model. Although the use of extracted human teeth enhances anatomical relevance, clinical extrapolation of the findings should be approached with caution. Future in vivo studies and randomized clinical trials are necessary to validate the translational potential of these agents under physiological conditions. This in vitro study aims to explore the antimicrobial potential of CeO_2-NPs, alone and in combination with CHX, in the disinfection of root canal systems. By addressing the limitations of current medicaments and leveraging nanotechnology, this protocol may contribute valuable insights into future endodontic disinfection strategies. The findings, once available, will support the development of enhanced treatment protocols for persistent root canal infections caused by *E. faecalis*.

XSL•FO
RenderX

Acknowledgments

We acknowledge the generous technical support from the research house in Datta Meghe Institute of Higher Education and Research and our guide in supporting this project. All authors declared that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this article.

Authors' Contributions

BSK and AI conceptualized and registered this study. BSK completed the study design and preliminary paper preparation. BSK will complete the final study design and manuscript writing.

Conflicts of Interest

None declared.

References

1. Yang S, Meng X, Zhen Y, Baima Q, Wang Y, Jiang X, et al. Strategies and mechanisms targeting biofilms associated with endodontic infections: a comprehensive review. Front Cell Infect Microbiol 2024;14:1433313 [FREE Full text] [doi: 10.3389/fcimb.2024.1433313] [Medline: 39091674]
2. Alghamdi F, Shakir M. The influence of Enterococcus faecalis as a dental root canal pathogen on endodontic treatment: a systematic review. Cureus 2020 Mar 13;12(3):e7257 [FREE Full text] [doi: 10.7759/cureus.7257] [Medline: 32292671]
3. Love RM. Enterococcus faecalis--a mechanism for its role in endodontic failure. Int Endod J 2001 Jul;34(5):399-405. [doi: 10.1046/j.1365-2591.2001.00437.x] [Medline: 11482724]
4. Yin IX, Udduttulla A, Xu VW, Chen KJ, Zhang MY, Chu CH. Use of antimicrobial nanoparticles for the management of dental diseases. Nanomaterials (Basel) 2025 Jan 28;15(3):11-14 [FREE Full text] [doi: 10.3390/nano15030209] [Medline: 39940185]
5. Sy K, Chevalier C, Maton M, Mokbel I, Mahieux S, Houcke I, et al. Therapeutic potential of chlorhexidine-loaded calcium hydroxide-based intracanal medications in endo-periodontal lesions: an ex vivo and in vitro study. Antibiotics (Basel) 2023 Sep 07;12(9):1416 [FREE Full text] [doi: 10.3390/antibiotics12091416] [Medline: 37760713]
6. Silina EV, Ivanova OS, Manturova NE, Medvedeva OA, Shevchenko AV, Vorsina ES, et al. Antimicrobial activity of citrate-coated cerium oxide nanoparticles. Nanomaterials (Basel) 2024 Feb 13;14(4):41-42 [FREE Full text] [doi: 10.3390/nano14040354] [Medline: 38392727]
7. Mohamed A, Ismail E, Adam RZ. A comparison of conventional root canal sealers with ones that use green synthesized nanoparticles for antimicrobial activity: protocol for a systematic review. JMIR Res Protoc 2024 Oct 11;13:e51351 [FREE Full text] [doi: 10.2196/51351] [Medline: 39393053]
8. Xue Y, Yang F, Wu L, Xia D, Liu Y. CeO2 nanoparticles to promote wound healing: a systematic review. Adv Healthc Mater 2024 Mar;13(6):e2302858. [doi: 10.1002/adhm.202302858] [Medline: 37947125]
9. Wang H, Wu Y, Dai M, Zhu T, Wu D, Leng D. PB/CeO nanoparticles regulating reactive oxygen species for the control of Enterococcus faecalis infection in root canals. J Biomed Mater Res B Appl Biomater 2025 Sep;113(9):e35646-e35721. [doi: 10.1002/jbm.b.35646] [Medline: 40877036]
10. Pandey A, Bhushan J, Joshi RK, Uppal AS, Angrup A, Kansal S. Comparative evaluation of antimicrobial efficacy of chitosan nanoparticles and calcium hydroxide against endodontic biofilm of Enterococcus faecalis: An in vitro study. J Conserv Dent Endod 2024 Jul;27(7):750-754. [doi: 10.4103/JCDE.JCDE_219_24] [Medline: 39262597]
11. Sanju D, Kumari VA, Thomas T, Thomas JT. Comparative evaluation of cerium oxide nanoparticles and calcium hydroxide as intracanal medicaments against Enterococcus faecalis on tooth substrate: an in vitro study. Endodontology 2022 Jul 01;34(3):162-167. [doi: 10.4103/endo.endo_25_22]
12. Vaghela D, Kandaswamy D, Venkateshbabu N, Jamini N, Ganesh A. Disinfection of dentinal tubules with two different formulations of calcium hydroxide as compared to 2% chlorhexidine: as intracanal medicaments against Enterococcus faecalis and Candida albicans: an in vitro study. J Conserv Dent 2011 Apr;14(2):182-186 [FREE Full text] [doi: 10.4103/0972-0707.82625] [Medline: 21814363]
13. Dar M, Gul R, Karuppiah P, Al-Dhabi N, Alfadda A. Antibacterial activity of cerium oxide nanoparticles against ESKAPE pathogens. Crystals 2022 Jan 26;12(2):179. [doi: 10.3390/cryst12020179]
14. Shrestha A, Kishen A. Antibacterial nanoparticles in endodontics: a review. J Endod 2016 Oct;42(10):1417-1426. [doi: 10.1016/j.joen.2016.05.021] [Medline: 27520408]

Abbreviations

CeO_2-NP: cerium oxide nanoparticle

XSL•FO
RenderX

CFU: colony-forming unit
CHX: chlorhexidine gluconate

Edited by T Leung, GRM La Rosa; submitted 23.May.2025; peer-reviewed by WM Nageeb, G Patri; comments to author 02.Jul.2025; revised version received 03.Aug.2025; accepted 04.Aug.2025; published 02.Mar.2026.

Please cite as:
Khanadali BS, Ikhar A
Evaluation of the Antimicrobial Efficacy of Cerium Oxide Nanoparticles and Chlorhexidine Gluconate Against Enterococcus faecalis in Root Canal Systems: Protocol for an In Vitro Study
JMIR Res Protoc 2026;15:e77998
URL: https://www.researchprotocols.org/2026/1/e77998
doi:10.2196/77998
PMID:

©Bhuvaneshwari Shrikant Khanadali, Anuja Ikhar. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 02.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Effects of Time-Restricted Eating on Ketone Metabolism and Immunoregulation in Premenopausal Women: Protocol for a Pilot, Prospective, Single-Arm Dietary Intervention Study

Geethika P Thota[1*], MD; Natalie A Macheret[1*], BS; Sophia B Glaros[1], BS; Ila N Kacker[1*], BA; Samson L Cantor[1*], BA; Lilian Mabundo[1*], RN, MSN, CDCES; Noemi Malandrino[1*], MD, PhD; Kong Y Chen[1*], PhD; Amber B Courville[1*], RD, PhD; Sara Turner[1*], MPH, RD; Shanna Yang[1*], MPH, RD; Andrea Krenek[1*], RD, PhD; Thomas C Recupero[2*], BA; Rachael J Klein[2*], MD; Kim Han[2*], PhD; Patrycja Puchalska[3*], PhD; Peter Crawford[3*], MD, PhD; Michael N Sack[2*], MD, PhD; Stephanie T Chung[1*], MBBS

[1]Diabetes Endocrinology and Obesity Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Rockville, MD, United States

[2]Laboratory of Mitochondrial Biology and Metabolism, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, MD, United States

[3]Clinical and Translational Science Institute, University of Minnesota, Minneapolis, MN, United States

[*]these authors contributed equally

Corresponding Author:
Stephanie T Chung, MBBS
Diabetes Endocrinology and Obesity Branch
National Institute of Diabetes and Digestive and Kidney Diseases
National Institutes of Health
10 Center Drive, Building 10 RM-59542
Rockville, MD, 20892
United States
Phone: 1 3014961211
Email: stephanie.chung@nih.gov

Abstract

Background: Intermittent fasting interventions, such as time-restricted eating (TRE) without calorie restriction, may offer diverse cardiometabolic health benefits, including reductions in inflammation. However, the underlying metabolic mechanisms are poorly understood.

Objective: This study is designed to evaluate ketone metabolism and immunoregulation during TRE (6-hour feeding and 18-hour fasting) without caloric restriction compared to a conventional dietary regimen (12-hour feeding and 12-hour fasting) in women.

Methods: This is a pilot, prospective, single-arm dietary intervention study in premenopausal women classified as lean and obese, with intraindividual and stratified comparisons. Ketone turnover, immunologic responses, and metabolic profiles will be compared between a conventional dietary regimen (12-hour feeding and 12-hour fasting) at admission and after 3 days of TRE (6-hour feeding and 18-hour fasting).

Results: The trial commenced in August 2024. As of February 2026, participant enrollment is ongoing. Complete enrollment is expected in the December 2026. Integrated analyses will be reported once data analysis is completed, which is expected in the winter of 2027.

Conclusions: This clinical protocol will evaluate whole-body ketone metabolism and immunoregulatory changes in women during TRE using state-of-the-art stable isotope and immunophenotyping techniques. This study will determine the effects of a time-limited dietary pattern, without caloric restriction, on dynamic measures of immunity and metabolism.

Trial Registration: ClinicalTrials.gov NCT06169137; https://clinicaltrials.gov/study/NCT06169137

International Registered Report Identifier (IRRID): DERR1-10.2196/81063

(JMIR Res Protoc 2026;15:e81063) doi:10.2196/81063

XSL•FO
RenderX

KEYWORDS

time-restricted eating; intermittent fasting; ketones; immune cells; immunology

Introduction

Background

Time-restricted eating (TRE) is a popular form of intermittent fasting, an eating pattern that alternates between eating and fasting within defined daily eating windows (4-10 h). Early TRE dietary patterns confine eating windows to the daylight waking hours with or without caloric or macronutrient restriction [1]. TRE has been linked to multiple metabolic and immune benefits, including antiaging and longevity benefits [2], healthy body weight, improved insulin sensitivity, and reduced glucose excursions in healthy and overweight individuals [3-11]. TRE has also been associated with changes in thyroid function, menstrual regularity, fatty acid metabolism, and neuroendocrine stress responses, particularly when associated with energy deprivation [12-15].

The underlying mechanisms mediating these metabolic benefits of TRE have not yet been elucidated. Chronic fasting may extend longevity and reduce inflammation [13,14,16]; however, whether shorter fasting durations during TRE modulate immune activation to promote healthy metabolism remains unclear. Fasting-induced ketogenesis during intermittent fasting regimens may be a key mediator of the anti-inflammatory response associated with TRE [17]. Physiological ketosis is protective and homeostatic in contrast to pathological states such as diabetic ketoacidosis [18]. Canonically, ketone bodies are a critical energy source for the brain during periods of diminished glucose availability during prolonged fasting [19]. However, contemporary data suggest that elevated circulating ketone bodies are also key signaling molecules and epigenetic or posttranslational modifiers [19].

Ketone bodies (acetoacetate and β-hydroxybutyrate [βOHB]) interact with various histone deacetylases and other proteins to induce posttranslational modifications, reduce oxidative stress, and inhibit the activation of the nucleotide-binding oligomerization domain–like receptor family pyrin domain containing 3 inflammasome [18,20-22]. Notably, βOHB activates transcription factors and adenosine monophosphate–activated protein kinase signaling, which plays a crucial role in regulating inflammatory pathways [23-25]. During a 24-hour fast in healthy individuals, we demonstrated that fasting blunted inflammatory markers and immune activation in monocytes or macrophages and $CD4^+$ T cells both in healthy individuals and in those with mild inflammatory disease (steroid-naive asthma) [26,27]. In a small pilot study of men and women with multiple sclerosis, alternate-day intermittent fasting with caloric restriction was associated with gut microbiome remodeling and immune changes suggestive of reduced interleukin-17 (IL-17)–producing T cells and increased numbers of regulatory T cells [28].

Although prolonged fasting durations have been associated with T cell differentiation, whether the 18-hour fasting component of a TRE dietary pattern would be sufficient to induce changes in ketone metabolism and T cell activation is unexplored. An early 6-hour TRE was associated with increased fat oxidation [5] and increased βOHB concentrations by 0.03 (SE of mean 0.01) mM using a crossover 4-day design [4]; however, no studies have linked ketogenesis during TRE with T cell immunophenotyping. Additionally, variations in the metabolic effects of TRE may be confounded by other biological factors; the ketogenic response to TRE may vary according to biological sex and obesity status. Individuals with obesity, compared to lean individuals, have lower rates of ketogenesis, and this may potentially be influenced by nonalcoholic fatty liver disease [29-31]. Moreover, some studies report higher ketogenesis in women than in men [32-35].

Research Gap and Significance

No studies have examined the metabolic mechanisms of ketone body metabolism and the anti-inflammatory effects of a 6-hour TRE regimen stratified by sex and obesity status. Given the limited research and discrepant findings regarding the effects of biological sex and obesity on ketone turnover, it is crucial to understand their role in response to dietary interventions. This clinical translational study uses state-of-the-art stable isotope and immune-metabolic phenotyping techniques to determine the impact of shifts in nutrient access and availability on dynamic changes in whole-body and cellular ketone turnover and immune activation in healthy women stratified by obesity status. This protocol complements the previous study conducted in men (NCT04728165) and will be used to explore TRE-mediated immune response in men and women.

Study Aims

Primary Aim 1

The first primary aim is to quantify whole-body βOHB turnover after short-term early 6-hour TRE compared to a conventional 12-hour dietary regimen in women.

Primary Aim 2

The second primary aim is to quantify $CD4^+$ T cell responses after short-term early 6-hour TRE compared to a conventional 12-hour dietary regimen in women.

Secondary Aim 1

The first secondary aim is to quantify changes in postabsorptive glucose rate of appearance (R_a), measured using (6,6-$^2H_2$) glucose and βOHB R_a (intravenous), after 3 days of 6-hour TRE.

Secondary Aim 2

The second secondary aim is to quantify $CD4^+$ T cell responsiveness (T helper 17 [Th17] polarization) in lean versus obese women after 3 days of 6-hour TRE.

XSL•FO
RenderX

Methods

Study Setup

This single-site 5-day inpatient study is a pilot, prospective, single-arm intervention trial in premenopausal women classified as lean or obese, with intraindividual and stratified comparisons. Ketone turnover and immunologic responses will be compared between a conventional dietary regimen (12-hour feeding and 12-hour fasting) at admission and after 3 days of TRE (6-hour feeding and 18-hour fasting; Figure 1). Each participant will serve as their own control, with analyses stratified by obesity classification. We chose a 3-day TRE intervention to determine the acute effects of this dietary pattern on ketone metabolism and immune function. Our previous publications showed that a 3-day TRE in men was feasible and improved glycemic variability and hepatic insulin resistance [36,37]. This study was designed to replicate and validate the 3-day TRE and extend these findings by evaluating glycemia and immunometabolism in women and examining ketone-mediated mechanisms. Secondary measures of glucose metabolism will be compared between groups at the same time points. To explore potential effects on hormonal mediators, circadian rhythms, and cardiometabolic markers, these measures will be evaluated at admission and after intervention. This study will be conducted in the Metabolic Clinical Research Unit at the National Institutes of Health (NIH) Clinical Center, Bethesda, Maryland.

Figure 1. Study schema and timeline. Overview of the study design and timeline across screening, conventional diet, and time-restricted eating phases.

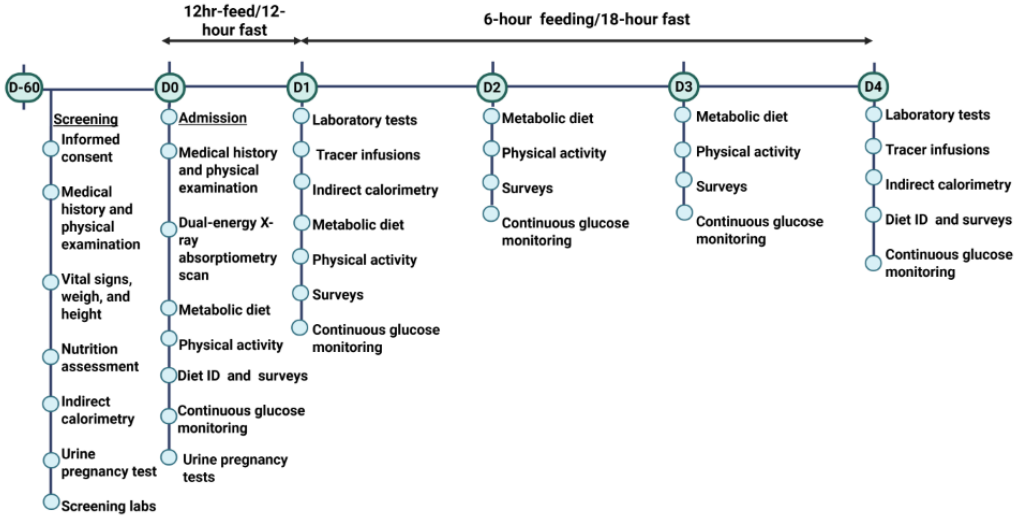

Participants

Healthy premenopausal women aged 18 to 50 years of all races and ethnicities will be recruited. A full overview of inclusion and exclusion criteria is provided in Textbox 1. An overview of recruitment strategies is provided in Textbox 2. Interested participants will complete an online interest questionnaire and telephone interview before screening. (Figure 2).

XSL•FO
RenderX

Textbox 1. Study inclusion and exclusion criteria.

Inclusion criteria

- Willingness to comply and be available for study procedures
- Women aged between 18 and 50 years in their premenopausal phase (regular menses and follicle-stimulating hormone is below the upper limit of normal)
- BMI of 18 to 24.9 kg/m^2 or 30 kg/m^2 or greater
- In good health per medical history and screening laboratory evaluation
- Agreement to adhere to lifestyle considerations throughout the study

Exclusion criteria

- Immune, inflammatory, or metabolic conditions that could affect study outcomes, except chronic, controlled hypothyroidism
- Current use of antihyperglycemic medications; systemic steroids; adrenergic stimulants; or drugs affecting sleep, circadian rhythms, or metabolism
- Caffeine intake more than 300 mg daily (approximately three 8-oz cups of coffee)
- Factors affecting circadian rhythms, including shift work, irregular sleep or eating schedules, or fasting for more than 15 hours per day
- History of an eating disorder
- Food allergies, intolerances, or dietary patterns that restrict consumption of the metabolic diet
- Inability to provide informed consent
- Pregnancy or lactation
- Unstable body weight, defined as more than 5% change in the past 3 months
- Competitive sports training or unwilling to follow exercise protocol
- Alcohol intake of more than 3 servings per day
- Current smoker or regular tobacco, vaping, or nicotine use within the past 3 months

Textbox 2. Recruitment strategies.

Recruitment strategies

- ClinicalTrials.gov
- Clinical Center Research Studies website
- Twitter posts and chats with study investigators
- Social media posts in accordance with National Institutes of Health (NIH) Social Media Guidelines
- Use of the NIH Clinical Center Office of Patient Recruitment services, including the creation and distribution of study flyers and dissemination of information through preexisting recruitment avenues, such as the NIH recruitment listserve

XSL•FO
RenderX

Figure 2. Study flow diagram. Participant recruitment, screening, and study completion. NIH: National Institutes of Health.

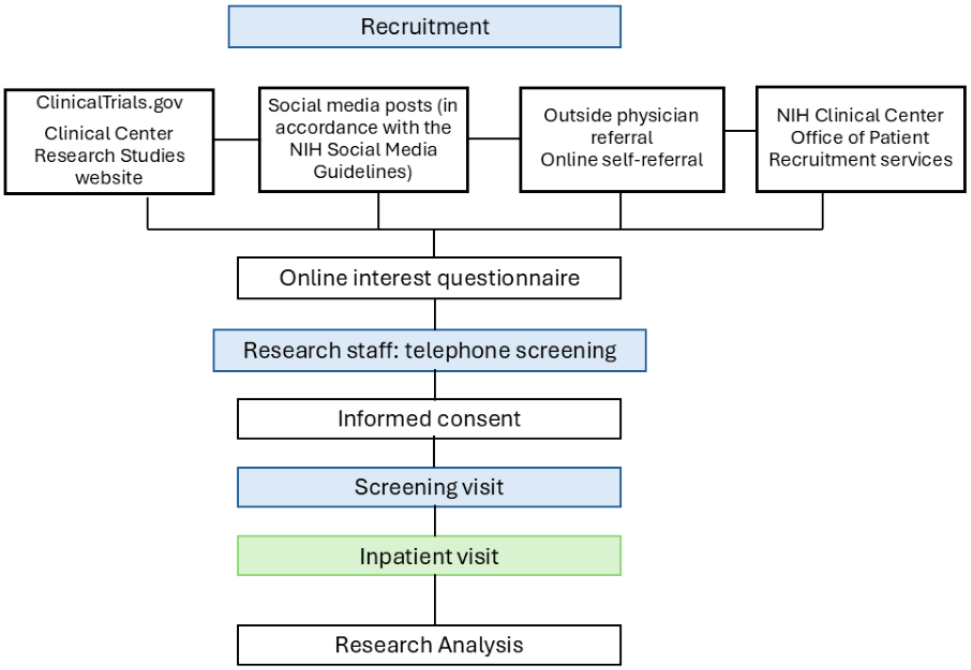

Screening Visit

Study participants will be screened after a 10- to 12-hour overnight fast to confirm eligibility. Screening evaluation will occur within 60 days of the inpatient visit and include obtaining a medical history, physical examination, fasting blood sampling, urine pregnancy testing, anthropometric measurements, resting energy expenditure measured by indirect calorimetry using the ventilated hood technique (Parvo Medics TrueOne 2400), and a nutrition assessment with nutrition research staff. Participants will be asked to complete 3 to 7 days of food logs before their inpatient visit.

The continuous glucose monitoring device (Dexcom, Inc) will either be provided at the screening visit (if eligibility is confirmed) or mailed to the participant approximately 1 week before admission. During the screening visit, participants will receive instructions on how to insert the US Food and Drug Administration–approved continuous glucose monitoring. Device placement instructions will also be reviewed during an in-person or virtual (telehealth) session with a clinician or qualified designee, including guidance on inserting the subcutaneous sensor and attaching the removable transmitter used to measure interstitial glucose levels. Participants will be asked to place the device at home 12 to 24 hours before admission and to wear it throughout their inpatient stay.

Inpatient Study Visit

Figure 1 outlines the study procedures. On admission to the metabolic unit on day 0, participants will be provided a conventional dietary regimen (09:30 AM to 09:30 PM). A dual-energy X-ray absorptiometry scan (Lunar iDXA; GE HealthCare) will be performed on day 0 to measure fat mass and lean body mass (LBM). Point-of-care ketone levels will be checked every morning using the Precision Xtra Ketone Meter (Abbott Diabetes Care Inc). Surveys will be administered to assess outcomes, including hunger and fullness, appetite, stress, and self-reported positive affect or well-being [38-46]. Participants will have 30 minutes of scheduled light-to-moderate physical activity, involving treadmill walking at a speed calculated to target 35% to 45% of heart rate reserve, using their age-predicted maximum heart rate and resting heart rate obtained during their screening visit.

TRE Intervention Design

The TRE intervention is a 6-hour early eating window from 09:30 AM to 02:30 PM and will be initiated for 3 consecutive days (days 1-3). Participants will have a 60-minute window to consume each meal (Figure 1). Participants will be asked to consume meals in their entirety, and any foods or beverages not consumed will be reweighed and recorded for data purposes. Eucaloric metabolic breakfast, lunch, and dinner will be provided; each meal is estimated to provide 33% of an individual's estimated energy needs and a macronutrient distribution of 50% carbohydrate, 35% fat, and 15% protein (Figure 1). Standard breakfast, lunch, and dinner options will be offered from a menu designed by nutrition research team members. Sample menu options are shown in Figure 3.

XSL•FO
RenderX

Figure 3. Illustration of a sample menu with available options. (A) Conventional dietary regimen (day 0). (B) Time-restricted eating dietary regimen (days 1-3).

Options	Diet type	Breakfast	Lunch	Dinner	Snack
1 ☐	Standard American	Oatmeal with walnuts and raisins Bacon Multigrain toast with butter and jelly Skim milk	Grilled chicken breast with wild rice and buttered green beans Oreo cookies Lemonade	Turkey sandwich on multigrain bread with lettuce, tomato and mayo Carrots and hummus Shortbread cookies Apples	2% Milk Granola bar
2 ☐	Gluten Free	Scrambled egg with peppers and onions Roasted potatoes Peach yogurt Apples	Baked lemon garlic tilapia Grits with butter Buttered broccoli Peaches	Entrée salad with grilled chicken, lettuce, tomato, cucumber, onion and Italian dressing Carrots and Hummus Chocolate pudding Grapes	Celery and peanut butter with raisins 2% Milk

Options	Diet type	Breakfast	Lunch	Dinner
1 ☐	Standard American	English muffin with butter Turkey bacon Peach Yogurt Granola with raisins and walnuts	Turkey & Provolone wrap with lettuce, tomato and mayo Carrots and cucumber with ranch Applesauce	Beef tender roast with baked potato and buttered green beans Oreo cookies
2 ☐	Gluten Free	Scrambled egg Turkey bacon Rice Chex cereal Whole milk Bananas	Tuna salad with mayo on a bed of lettuce with onion, carrots and tomato with fat-free Ranch Baked potato chips Oranges Lemonade	Beef tender roast with mashed potatoes and steamed, buttered broccoli Apples Vanilla pudding

Intervention Preparation

Metabolic meals will be prepared at the NIH Clinical Center in the Metabolic Kitchen. Diets will be designed by nutrition research team members using ProNutra diet planning software (Viocare, Inc) to align with a specified macronutrient distribution and total energy provision, informed by indirect calorimetry and an activity factor of 1.4.

Adherence Monitoring

As an inpatient intervention protocol, this study will monitor and confirm dietary compliance through food weigh backs (as applicable) and detailed case report forms completed by the clinical research study team and clinicians.

Tracer Administration

On days 1 and 4, participants will receive a fasting 3-hour peripheral infusion of a stable isotope-labeled ketone body and glucose with blood draws obtained at scheduled intervals to trace ketone and glucose metabolism (Figure 4). Ketone body and glucose turnover will be measured using stable isotope tracers. Glucose and βOHB R_a will be measured using a primed continuous 3-hour infusion. The (6,6-$^2H_2$) glucose dose will be 0.49 µmol/kg LBM per minute for 180 minutes, with a priming dose of 29.4 µmol/kg LBM, and the (U-$^{13}C_4$) βOHB will be administered as 1.8 µmol/kg LBM per minute priming dose (10-minute infusion), followed by 0.18 µmol/kg LBM per minute for 170 minutes. Blood samples will be collected before stable isotope administration and during the near–steady-state period (last 30 minutes of the infusion). Whole blood will be centrifuged, and plasma will be transferred and frozen at –80 °C until analysis. Blood samples will be collected for the measurement of glucose, insulin, C-peptide, and ketone bodies. Intravenous stable isotopes will meet NIH minimum standards for stable isotope administration and will be obtained through an NIH-approved pharmacy compounding vendor.

XSL•FO
RenderX

Figure 4. Tracer study protocol. Schematic overview of tracer infusion, blood sampling, and timing relative to the time-restricted eating protocol. βOHB: β-hydroxybutyrate.

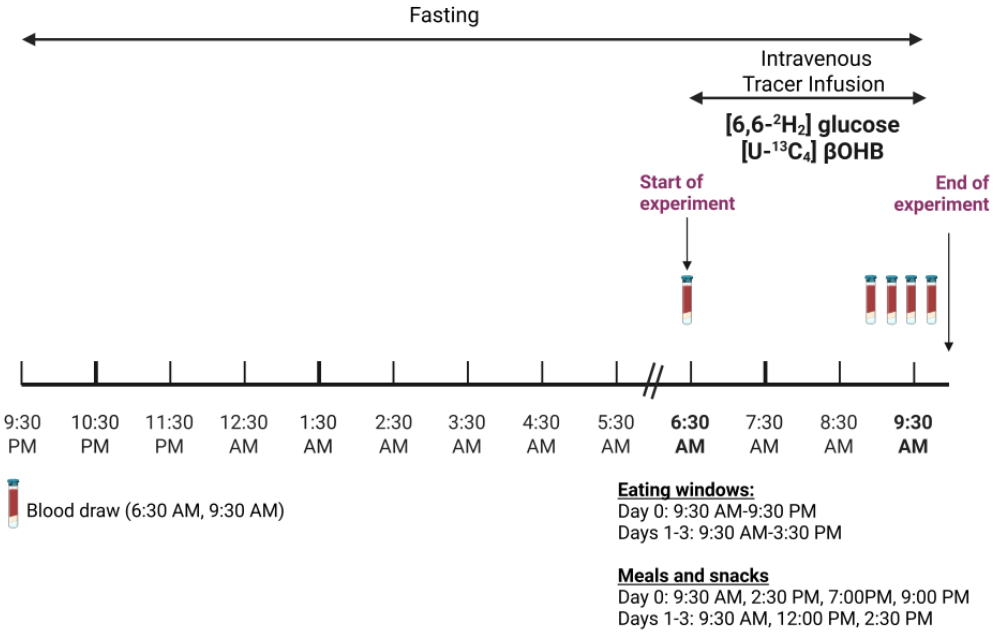

Sample Size Justification

Sample size determination is based on the ability to detect meaningful differences in ketone-body turnover and $CD4^+$ T cell responsiveness and determine effect sizes for short-term TRE. Because calculations are derived from limited past data, variances are anticipated, and the sample size calculations are feasibility estimates.

The primary outcomes are as follows: (1) change in βOHB R_a to determine the effect of short-term TRE on ketone-body metabolism (2) change in $CD4^+$ T cell responsiveness (T helper type 1, T helper type 2, Th17, and Treg cytokines)

These outcomes will be evaluated using a 2-sided paired Student *t* test (2-tailed) with an α of .05. For primary outcome 1, based on unpublished data (Crawford, M, personal communication, March 2025) in studies of human participants (βOHB whole-body R_a: 10.4, SD 3.16 µmol/kg LBM/min). A sample of 15 participants has 80% power to detect a difference of 2.5 µmol/kg LBM per minute (25%), assuming a 2-sided α of .05, an SD of 3.16, and a correlation of 0.50.

For primary outcome 2, based on a study of 13 healthy volunteers conducted at the NIH Clinical Center (ClinicalTrials.gov NCT02719899) in which isolated $CD4^+$ T cells were investigated, the mean paired refed–fasting differences in IL-17 release were 187.1 (SD 398.79) pg/mL, respectively [21]. Using a 2-tailed paired *t* test with an alpha of .05 (n=15), we will have 80% power, assuming an SD of the difference of 400 pg/mL to detect a clinically important TRE-induced change from admission in IL-17 of 311 pg/mL. We will attempt to recruit and screen in person up to 150 women and enroll 30 women to complete the study, including 15 lean women and 15 women with obesity.

Data Collection and Statistical Analysis

Data Collection Methods

Clinical data (including adverse events, concomitant medications, and expected adverse reactions) and clinical laboratory data will be entered into the Clinical Research Information System, a 21 Code of Federal Regulations Part 11–compliant data capture system provided by the NIH as well as REDCap (Research Electronic Data Capture; Vanderbilt University), a data management system provided by the National Institute of Diabetes and Digestive and Kidney Diseases. These data systems include password protection and controlled access.

Outcome Analysis

Data will be analyzed with a statistical software package (eg, Stata [version 18; StataCorp]), with descriptive statistics presented as means (95% CIs), medians (IQRs), or percentages. Due to the pilot nature of this study, no prespecified covariates are assigned. Normality of data will be checked using the Shapiro-Wilk test and visually inspected using histograms. If severe violations of data normality are observed, transformations such as logarithmic or square-root transformations will be applied. If such transformations fail to generate normal data, we will pursue statistical analyses using nonparametric tests, which are conservative but less powerful.

Analysis of glucose levels, glycemic responses, and metabolic rates will be compared using paired 2-tailed Student *t* tests for within-group comparisons or unpaired 2-tailed Student *t* tests for between-group comparisons (lean women versus women with obesity). Statistical significance will be tested at an α level of .05. Quality control procedures will be implemented beginning with the data entry system, and data quality control checks will be run on the database. Any missing data or anomalies will be communicated to the relevant departments for clarification or resolution.

XSL•FO
RenderX

Study procedures will be subject to audits and monitoring visits to ensure compliance with the protocol and applicable regulatory requirements, consistent with the quality assurance program plan. Audit or monitoring visit results will be reported to the principal investigator for further reporting, as appropriate. Study documents and pertinent hospital or clinical records will be reviewed to verify that the conduct of the study is consistent with the protocol plan.

Circumstances that may warrant study termination or suspension include, but are not limited to, the following:

- Identification of unexpected, significant, or unacceptable risks to participants
- Demonstration of efficacy that would warrant stopping
- Insufficient compliance with protocol requirements
- Data that are insufficiently complete and evaluable
- Determination that the primary end point has been met
- Determination of futility

If the study is halted, it may resume once concerns about safety, protocol compliance, and data quality are addressed and approval is obtained from the institutional review board and study sponsor. The study will comply with the NIH Data Management and Sharing Policy, which applies to all new and ongoing NIH-funded research with a clinical protocol that undergoes scientific review.

Isotopic Calculations

Rates of glucose and βOHB appearance (R_a) will be calculated under near–steady-state conditions using the average enrichment of (6, 6-$^2H_2$) glucose or (U-$^{13}C_4$) βOHB in the systemic circulation using conventional isotope dilution calculations [22].



(1)

Under fasting and steady-state conditions, R_a of glucose reflects hepatic glucose production, and R_a of βOHB reflects whole-body βOHB turnover.

Glucose production rate is the sum of the rates of gluconeogenesis and glycogenolysis and is calculated as follows:

Glucose production rate = R_a (total) – rate of infusion of labeled glucose

Immuno-Metabolomic Profiling and Study Design

The ex vivo metabolic substrate analysis in $CD4^+$ T cells will help corroborate the mechanistic insights gained from the clinical study. For cell culture and ex vivo analyses, blood will be drawn immediately before the infusion of isotope tracers on days 1 and 3 after TRE on day 4 (Figure 4). $CD4^+$ T cells will be extracted from peripheral blood mononuclear cells (PBMCs) by negative selection using a $CD4^+$ T cell Isolation Kit (Miltenyi Biotech) and divided into cells for analysis using mass spectrometry (stable isotope tracing using [U-$^{13}C_4$] acetoacetate and D-[U-$^{13}C_4$] βOHB incubations in cell culture and measurement of energy adenylates), immunologic assays (Th17 responsiveness and flow cytometry), and for RNA and protein interrogation of ketolysis and Th17 regulatory pathways.

Patient and Public Involvement

Patients and members of the public were not involved in the development of the research question, study design, outcome measures, or recruitment plans for this study protocol.

Ethical Considerations

This pilot study is registered on ClinicalTrials.gov (NCT06169137) and has received approval from the institutional review board of the National Institutes of Health (IRB001746) at the study site. Informed consent will be obtained from eligible and interested participants before enrollment and the start of study procedures. All study data will be deidentified by removal of direct identifiers, including names, addresses, dates of birth, phone numbers, social security numbers, email addresses, medical record numbers, and institutional identification numbers. Each participant will be assigned a unique study identification code based on order of enrollment. Participants will be informed that participation is voluntary and that they may withdraw at any time without impact on their clinical care at the institution. Because there is no direct clinical benefit, participants will receive financial monetary compensation for their time and the inconvenience associated with study procedures.

Publication Plans

Limited data sharing will be available at the conclusion of the study. Findings from this study will be shared through presentations at national and international scientific conferences and published in peer-reviewed journals.

Results

Regulatory board approval was obtained in April 2024. Participant screening and enrollment commenced in September 2024 and are ongoing. As of February 2026, 20 participants have been recruited. Data analysis will commence after all participants have been recruited, and results are expected to be published by the winter of 2027.

Discussion

Anticipated Findings

To our knowledge, studies have not explored the effects of TRE on the inflammatory and ketone pathways to determine if these factors contribute to improved metabolic health. TRE and intermittent fasting are postulated to induce stress-resistance pathways at integrated metabolic and canonical signaling levels that may be immunomodulatory [2]. Chronic fasting extends longevity and reduces inflammation [13,14,16]; however, whether this occurs with shorter fasting durations during TRE remains unclear. We have previously shown that a prolonged 24-hour fast blunted inflammatory markers and immune activation in monocytes or macrophages and $CD4^+$ T cells, both in healthy individuals and in those with a mild inflammatory disease [26,27]. Furthermore, we showed that fasting reduced $CD4^+$ T helper type 1 and Th17 polarization and diminished

XSL•FO
RenderX

cytokine responsiveness in PBMCs following T cell receptor activation. These findings suggest reprogramming of the PBMCs and differential gene expression changes that were most pronounced in the 24-hour fast versus refed state compared to the 12-hour fast versus refed state [47].

To further determine whether a shorter fasting duration (18 h) will alter T cell activation and metabolism, we conducted a pilot clinical 6-hour TRE trial in men (NCT04728165) at the NIH Clinical Center Metabolic Clinical Research Unit. This study evaluated the effect of a 6-hour TRE versus a conventional 12-hour fasting regimen on immunologic signatures in patients with mild to moderate psoriasis compared to a matched healthy control population. Preliminary data in that trial suggested that 6-hour TRE is sufficient to blunt $CD4^+$ T inflammatory responses. This clinical protocol represents the next step in exploring potential mediators of the immunomodulatory response. Key pathways that may influence immunometabolism include fasting-induced ketogenesis, hormonal shifts (fasting-induced low insulin-to-glucagon ratio, increased peptide YY [a satiety hormone], and increased corticosteroid signaling), alterations in microbiome-derived metabolism of short-chain fatty acids, or upregulation of autophagy [48].

Strengths and Limitations

This translational study integrates clinical and basic scientific research applications. Using a combination of specialized in vivo and ex vivo metabolic deep phenotyping techniques, including stable isotope tracer studies, will elucidate the immunologic signatures of TRE and explore whether obesity is a confounding factor in women. Stable isotope tracer administration will facilitate quantification of ketone production (hepatic) as well as ketone uptake and use by extrahepatic organs and tissues under conditions relevant to the 6-hour TRE intervention. Evaluating the ketogenic response to TRE in women is critical and clinically relevant for understanding the role of biological sex–specific differences in dietary intervention response.

Study limitations include the single-center pilot design, inclusion of women only, and the short study duration, which may limit the generalizability of findings. This is a significant constraint when considering the generalizability of findings to men, older adults, or individuals with metabolic disorders, underscoring the need to address this in a future, larger study with longer-term follow-up. Given the limited sample size and the exploratory nature of the study with intraindividual and stratified comparisons, the results should be interpreted with caution. Findings are intended to be hypothesis generating, which can be further explored in larger and more targeted studies. Future studies should include parallel-arm control design or randomized crossover trials with a conventional 12:12-hour eating schedule.

Conclusions

This pilot study will be the first to yield generalizable knowledge on whole-body ketone metabolism and immunoregulatory changes in women during early 6-hour TRE. The mechanistic findings will help determine the importance of fasting duration, ketone body production, and immunoregulation to the salutary outcomes of TRE. This analysis will provide much-needed foundational data to inform the design of larger, multicenter randomized trials.

Funding

This research was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases within the National Institutes of Health (NIH) through the Bench-to-Bedside Program Funds Award (ID 994837). Contributions of the NIH authors were made as part of their official duties as NIH federal employees, are in compliance with agency policy requirements, and are considered works of the US Government. However, the findings and conclusions presented in this paper are those of the authors and do not necessarily reflect the views of the NIH or the US Department of Health and Human Services.

Data Availability

The anonymized dataset and statistical code supporting this study will be made available in a public repository upon study completion and publication.

Authors' Contributions

GPT, MNS, and STC contributed to conceptualization, investigation, methodology, writing the original draft, and reviewing the manuscript. NAM, SBG, INK, SLC, LM, NM, KYC, ABC, ST, SY, AK, TCR, RJK, KH, PP, PC, MNS, and STC reviewed and edited all sections of this paper. STC was also principally responsible for developing the intervention content and reviewed all sections of this paper. STC, MNS, PP, and PC contributed to formal analysis, funding acquisition, investigation, methodology, supervision, and final editing. STC was the principal investigator responsible for the conception, overall design, and conduct of this trial. All authors approved the final version of the manuscript.

Conflicts of Interest

None declared.

References

XSL•FO
RenderX

1. Parr EB, Devlin BL, Hawley JA. Perspective: time-restricted eating-integrating the what with the when. Adv Nutr 2022 Jun 01;13(3):699-711 [FREE Full text] [doi: 10.1093/advances/nmac015] [Medline: 35170718]
2. de Cabo R, Mattson MP. Effects of intermittent fasting on health, aging, and disease. N Engl J Med 2019 Dec 26;381(26):2541-2551. [doi: 10.1056/NEJMra1905136] [Medline: 31881139]
3. Sutton EF, Beyl R, Early KS, Cefalu WT, Ravussin E, Peterson CM. Early time-restricted feeding improves insulin sensitivity, blood pressure, and oxidative stress even without weight loss in men with prediabetes. Cell Metab 2018 Jun 05;27(6):1212-21.e3 [FREE Full text] [doi: 10.1016/j.cmet.2018.04.010] [Medline: 29754952]
4. Jamshed H, Beyl RA, Della Manna DL, Yang ES, Ravussin E, Peterson CM. Early time-restricted feeding improves 24-hour glucose levels and affects markers of the circadian clock, aging, and autophagy in humans. Nutrients 2019 May 30;11(6):1234 [FREE Full text] [doi: 10.3390/nu11061234] [Medline: 31151228]
5. Ravussin E, Beyl RA, Poggiogalle E, Hsia DS, Peterson CM. Early time-restricted feeding reduces appetite and increases fat oxidation but does not affect energy expenditure in humans. Obesity (Silver Spring) 2019 Aug 24;27(8):1244-1254 [FREE Full text] [doi: 10.1002/oby.22518] [Medline: 31339000]
6. Hutchison AT, Regmi P, Manoogian EN, Fleischer JG, Wittert GA, Panda S, et al. Time-restricted feeding improves glucose tolerance in men at risk for type 2 diabetes: a randomized crossover trial. Obesity (Silver Spring) 2019 May 19;27(5):724-732. [doi: 10.1002/oby.22449] [Medline: 31002478]
7. Ezzati A, Rosenkranz SK, Phelan J, Logan C. The effects of isocaloric intermittent fasting vs daily caloric restriction on weight loss and metabolic risk factors for noncommunicable chronic diseases: a systematic review of randomized controlled or comparative trials. J Acad Nutr Diet 2023 Feb;123(2):318-29.e1. [doi: 10.1016/j.jand.2022.09.013] [Medline: 36126910]
8. Cioffi I, Evangelista A, Ponzo V, Ciccone G, Soldati L, Santarpia L, et al. Intermittent versus continuous energy restriction on weight loss and cardiometabolic outcomes: a systematic review and meta-analysis of randomized controlled trials. J Transl Med 2018 Dec 24;16(1):371 [FREE Full text] [doi: 10.1186/s12967-018-1748-4] [Medline: 30583725]
9. Liu D, Huang Y, Huang C, Yang S, Wei X, Zhang P, et al. Calorie restriction with or without time-restricted eating in weight loss. N Engl J Med 2022 Apr 21;386(16):1495-1504. [doi: 10.1056/nejmoa2114833]
10. Zhao D, Guallar E, Woolf TB, Martin L, Lehmann H, Coughlin J, et al. Association of eating and sleeping intervals with weight change over time: the Daily24 cohort. J Am Heart Assoc 2023 Feb 07;12(3):e026484 [FREE Full text] [doi: 10.1161/JAHA.122.026484] [Medline: 36651320]
11. Varady KA, Cienfuegos S, Ezpeleta M, Gabel K. Clinical application of intermittent fasting for weight loss: progress and future directions. Nat Rev Endocrinol 2022 May 22;18(5):309-321. [doi: 10.1038/s41574-022-00638-x] [Medline: 35194176]
12. Ikhsan M, Siregar MF, Muharam R. The relationship between Ramadan fasting with menstrual cycle pattern changes in teenagers. Middle East Fertil Soc J 2017 Mar;22(1):43-47. [doi: 10.1016/j.mefs.2016.08.004]
13. Salgin B, Marcovecchio ML, Humphreys SM, Hill N, Chassin LJ, Lunn DJ, et al. Effects of prolonged fasting and sustained lipolysis on insulin secretion and insulin sensitivity in normal subjects. Am J Physiol Endocrinol Metab 2009 Mar;296(3):E454-E461 [FREE Full text] [doi: 10.1152/ajpendo.90613.2008] [Medline: 19106250]
14. Cienfuegos S, Corapi S, Gabel K, Ezpeleta M, Kalam F, Lin S, et al. Effect of intermittent fasting on reproductive hormone levels in females and males: a review of human trials. Nutrients 2022 Jun 03;14(11):2343 [FREE Full text] [doi: 10.3390/nu14112343] [Medline: 35684143]
15. Fahrenholtz IL, Sjödin A, Benardot D, Tornberg Å, Skouby S, Faber J, et al. Within-day energy deficiency and reproductive function in female endurance athletes. Scand J Med Sci Sports 2018 Mar 05;28(3):1139-1146. [doi: 10.1111/sms.13030] [Medline: 29205517]
16. Longo VD, Mattson MP. Fasting: molecular mechanisms and clinical applications. Cell Metab 2014 Feb 04;19(2):181-192 [FREE Full text] [doi: 10.1016/j.cmet.2013.12.008] [Medline: 24440038]
17. Neudorf H, Little JP. Impact of fasting and ketogenic interventions on the NLRP3 inflammasome: a narrative review. Biomed J 2024 Feb;47(1):100677 [FREE Full text] [doi: 10.1016/j.bj.2023.100677] [Medline: 37940045]
18. Laffel L. Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes. Diabetes Metab Res Rev 1999;15(6):412-426. [doi: 10.1002/(sici)1520-7560(199911/12)15:6<412::aid-dmrr72>3.0.co;2-8] [Medline: 10634967]
19. Puchalska P, Crawford PA. Metabolic and signaling roles of ketone bodies in health and disease. Annu Rev Nutr 2021 Oct 11;41(1):49-77 [FREE Full text] [doi: 10.1146/annurev-nutr-111120-111518] [Medline: 34633859]
20. Fann DY, Santro T, Manzanero S, Widiapradja A, Cheng YL, Lee SY, et al. Intermittent fasting attenuates inflammasome activity in ischemic stroke. Exp Neurol 2014 Jul;257:114-119. [doi: 10.1016/j.expneurol.2014.04.017] [Medline: 24805069]
21. Han K, Singh K, Rodman MJ, Hassanzadeh S, Wu K, Nguyen A, et al. Fasting-induced FOXO4 blunts human CD4+ T helper cell responsiveness. Nat Metab 2021 Mar;3(3):318-326 [FREE Full text] [doi: 10.1038/s42255-021-00356-0] [Medline: 33723462]
22. Sunehag AL, Treuth MS, Toffolo G, Butte NF, Cobelli C, Bier DM, et al. Glucose production, gluconeogenesis, and insulin sensitivity in children and adolescents: an evaluation of their reproducibility. Pediatr Res 2001 Jul;50(1):115-123. [doi: 10.1203/00006450-200107000-00021] [Medline: 11420428]

XSL•FO
RenderX

23. Kim DH, Park MH, Ha S, Bang EJ, Lee Y, Lee AK, et al. Anti-inflammatory action of β-hydroxybutyrate via modulation of PGC-1α and FoxO1, mimicking calorie restriction. Aging (Albany NY) 2019 Feb 27;11(4):1283-1304 [FREE Full text] [doi: 10.18632/aging.101838] [Medline: 30811347]
24. Chriett S, Dąbek A, Wojtala M, Vidal H, Balcerczyk A, Pirola L. Prominent action of butyrate over β-hydroxybutyrate as histone deacetylase inhibitor, transcriptional modulator and anti-inflammatory molecule. Sci Rep 2019 Jan 24;9(1):742 [FREE Full text] [doi: 10.1038/s41598-018-36941-9] [Medline: 30679586]
25. Bae HR, Kim DH, Park MH, Lee B, Kim MJ, Lee EK, et al. β-Hydroxybutyrate suppresses inflammasome formation by ameliorating endoplasmic reticulum stress via AMPK activation. Oncotarget 2016 Oct 11;7(41):66444-66454 [FREE Full text] [doi: 10.18632/oncotarget.12119] [Medline: 27661104]
26. Traba J, Kwarteng-Siaw M, Okoli TC, Li J, Huffstutler RD, Bray A, et al. Fasting and refeeding differentially regulate NLRP3 inflammasome activation in human subjects. J Clin Invest 2015 Nov 03;125(12):4592-4600 [FREE Full text] [doi: 10.1172/JCI83260] [Medline: 26529255]
27. Han K, Nguyen A, Traba J, Yao X, Kaler M, Huffstutler RD, et al. A pilot study to investigate the immune-modulatory effects of fasting in steroid-naive mild asthmatics. J Immunol 2018 Sep 01;201(5):1382-1388 [FREE Full text] [doi: 10.4049/jimmunol.1800585] [Medline: 30021766]
28. Cignarella F, Cantoni C, Ghezzi L, Salter A, Dorsett Y, Chen L, et al. Intermittent fasting confers protection in CNS autoimmunity by altering the gut microbiota. Cell Metab 2018 Jun 05;27(6):1222-35.e6 [FREE Full text] [doi: 10.1016/j.cmet.2018.05.006] [Medline: 29874567]
29. Vice E, Privette JD, Hickner RC, Barakat HA. Ketone body metabolism in lean and obese women. Metabolism 2005 Nov;54(11):1542-1545. [doi: 10.1016/j.metabol.2005.05.023] [Medline: 16253646]
30. Mey JT, Erickson ML, Axelrod CL, King WT, Flask CA, McCullough AJ, et al. β-Hydroxybutyrate is reduced in humans with obesity-related NAFLD and displays a dose-dependent effect on skeletal muscle mitochondrial respiration in vitro. Am J Physiol Endocrinol Metab 2020 Jul 01;319(1):E187-E195 [FREE Full text] [doi: 10.1152/ajpendo.00058.2020] [Medline: 32396388]
31. Soeters MR, Sauerwein HP, Faas L, Smeenge M, Duran M, Wanders RJ, et al. Effects of insulin on ketogenesis following fasting in lean and obese men. Obesity (Silver Spring) 2009 Jul 06;17(7):1326-1331 [FREE Full text] [doi: 10.1038/oby.2008.678] [Medline: 19369940]
32. Haymond M, Karl IE, Clarke WL, Pagliara AS, Santiago JV. Differences in circulating gluconeogenic substrates during short-term fasting in men, women, and children. Metabolism 1982 Jan;31(1):33-42. [Medline: 7043160]
33. Halkes CJ, van Dijk H, Verseyden C, de Jaegere PP, Plokker HW, Meijssen S, et al. Gender differences in postprandial ketone bodies in normolipidemic subjects and in untreated patients with familial combined hyperlipidemia. Arterioscler Thromb Vasc Biol 2003 Oct;23(10):1875-1880. [doi: 10.1161/01.atv.0000092326.00725.ed]
34. Cahill GFJ. Fuel metabolism in starvation. Annu Rev Nutr 2006 Aug 01;26(1):1-22. [doi: 10.1146/annurev.nutr.26.061505.111258] [Medline: 16848698]
35. Browning JD, Baxter J, Satapati S, Burgess SC. The effect of short-term fasting on liver and skeletal muscle lipid, glucose, and energy metabolism in healthy women and men. J Lipid Res 2012 Mar;53(3):577-586 [FREE Full text] [doi: 10.1194/jlr.P020867] [Medline: 22140269]
36. Macheret NT, Thota GP, Glaros SB, Kacker IN, Cantor SL, Malandrino N, et al. Impact of time-restricted eating on glycemic variability and metabolic flexibility in healthy and psoriatic subjects. Diabetes 2025 Jun 13:1648-P [FREE Full text] [doi: 10.2337/db25-1648-p]
37. Thota GP, Macheret NA, Glaros SB, Kacker IN, Cantor SL, Malandrino N, et al. Improved hepatic insulin sensitivity and insulin secretion in short-term early time-restricted eating in men with and without Psoriasis. Diabetes 2025;74(Supplement_1):1649-P [FREE Full text] [doi: 10.2337/db25-1649-P]
38. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037/0022-3514.54.6.1063]
39. Salsman JM, Butt Z, Pilkonis PA, Cyranowski JM, Zill N, Hendrie HC, et al. Emotion assessment using the NIH Toolbox. Neurology 2013 Mar 12;80(11_supplement_3):S76-S86. [doi: 10.1212/wnl.0b013e3182872e11]
40. NIH Toolbox® for assessment of neurological and behavioral function. National Institutes of Health and Northwestern University. URL: https://www.healthmeasures.net/images/nihtoolbox/Training-Admin-Scoring_Manuals/NIH_Toolbox_App_Administrators_Manual_v1.17.pdf [accessed 2026-02-03]
41. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396 [FREE Full text] [doi: 10.2307/2136404]
42. Svedlund J, Sjödin I, Dotevall G. GSRS--a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci 1988 Feb;33(2):129-134. [doi: 10.1007/BF01535722] [Medline: 3123181]
43. Beke M, Burns AM, Weir S, Solch RJ, Judkins TC, Nieves C, et al. Validation of a novel quality of life questionnaire: the Digestion-associated Quality of Life Questionnaire (DQLQ). Health Qual Life Outcomes 2022 Mar 28;20(1):53 [FREE Full text] [doi: 10.1186/s12955-022-01956-4] [Medline: 35346230]
44. Meule A, Hermann T, Kübler A. A short version of the Food Cravings Questionnaire-Trait: the FCQ-T-reduced. Front Psychol 2014 Mar 04;5:190 [FREE Full text] [doi: 10.3389/fpsyg.2014.00190] [Medline: 24624116]

XSL•FO
RenderX

45. Hammond L, Morello O, Kucab M, Totosy de Zepetnek JO, Lee JJ, Doheny T, et al. Predictive validity of image-based motivation-to-eat visual analogue scales in normal weight children and adolescents aged 9-14 years. Nutrients 2022 Feb 01;14(3):636 [FREE Full text] [doi: 10.3390/nu14030636] [Medline: 35276995]
46. Lesage FX, Berjot S, Deschamps F. Clinical stress assessment using a visual analogue scale. Occup Med (Lond) 2012 Dec 10;62(8):600-605. [doi: 10.1093/occmed/kqs140] [Medline: 22965867]
47. Han K, Singh K, Rodman MJ, Hassanzadeh S, Baumer Y, Huffstutler RD, et al. Identification and validation of nutrient state-dependent serum protein mediators of human CD4 T cell responsiveness. Nutrients 2021 Apr 28;13(5):1492 [FREE Full text] [doi: 10.3390/nu13051492] [Medline: 33924911]
48. Newman JC, Verdin E. β-Hydroxybutyrate: a signaling metabolite. Annu Rev Nutr 2017 Aug 21;37(1):51-76 [FREE Full text] [doi: 10.1146/annurev-nutr-071816-064916] [Medline: 28826372]

Abbreviations

βOHB: β-hydroxybutyrate
IL-17: interleukin-17
LBM: lean body mass
NIH: National Institutes of Health
PBMC: peripheral blood mononuclear cell
Ra: rate of appearance
REDCap: Research Electronic Data Capture
TRE: time-restricted eating

Edited by J Sarvestan; submitted 23.Jul.2025; peer-reviewed by D Duan; comments to author 28.Oct.2025; revised version received 12.Jan.2026; accepted 15.Jan.2026; published 06.Mar.2026.

Please cite as:
Thota GP, Macheret NA, Glaros SB, Kacker IN, Cantor SL, Mabundo L, Malandrino N, Chen KY, Courville AB, Turner S, Yang S, Krenek A, Recupero TC, Klein RJ, Han K, Puchalska P, Crawford P, Sack MN, Chung ST
Effects of Time-Restricted Eating on Ketone Metabolism and Immunoregulation in Premenopausal Women: Protocol for a Pilot, Prospective, Single-Arm Dietary Intervention Study
JMIR Res Protoc 2026;15:e81063
URL: https://www.researchprotocols.org/2026/1/e81063
doi:10.2196/81063
PMID:

©Geethika P Thota, Natalie A Macheret, Sophia B Glaros, Ila N Kacker, Samson L Cantor, Lilian Mabundo, Noemi Malandrino, Kong Y Chen, Amber B Courville, Sara Turner, Shanna Yang, Andrea Krenek, Thomas C Recupero, Rachael J Klein, Kim Han, Patrycja Puchalska, Peter Crawford, Michael N Sack, Stephanie T Chung. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 06.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

3D Comparative Evaluation of Condylar Morphology Between Chronic Areca Nut Chewers and Nonchewers: Protocol for a Case-Control Study

Vaishnavi Tarpe[1], BDS; Suwarna Dangore Khasbage[1], BDS, MDS

Department of Oral Medicine and Radiology, Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research, Wardha, India

Corresponding Author:
Vaishnavi Tarpe, BDS
Department of Oral Medicine and Radiology
Sharad Pawar Dental College and Hospital
Datta Meghe Institute of Higher Education and Research
Sawangi
Wardha, 442001
India
Phone: 91 7758840081
Email: vaishnavitarpe97@gmail.com

Abstract

Background: Areca nut (AN) is a commonly consumed psychoactive substance, especially in South and Southeast Asia. Chronic chewing of AN has been linked to multiple health problems, including temporomandibular joint (TMJ) disorders. Excessive strain on TMJ during chronic AN chewing can lead to repetitive injury, resulting in microtrauma and macrotrauma to both the TMJ and the surrounding masticatory structures. Previous studies have reported the long-term impact of AN chewing on TMJ by using conventional 2D imaging.

Objective: This study aims to evaluate and compare condylar morphology in chronic AN chewers and nonchewers by using 3D imaging.

Methods: This study will include 90 patients who will be divided into 2 groups: chronic AN chewers (n=45, 50%) and nonchewers (n=45, 50%). The study will be undertaken after obtaining institutional ethics committee approval and written informed consent from each patient. A detailed habit history of all the participants will be recorded. Each patient will undergo a clinical examination and radiographic evaluation of condylar morphology. Condylar morphology will be evaluated using cone beam computed tomography scans in both sagittal and coronal planes. All the findings will be recorded and then examined for statistical significance.

Results: On comparison of condylar morphology between chronic AN chewers and nonchewers by using cone beam computed tomography, statistical variations relevant to structural and pathological alterations such as osteophytes, surface flattening, and erosions are likely to occur.

Conclusions: This study aims to overcome the limitations of conventional 2D radiography and provide a more accurate assessment of condylar morphology. The findings should fill an existing gap in the literature by providing useful insights on the effects of chronic AN chewing on condylar structure by using 3D imaging. This research may help to improve the diagnosis, prevention, and management of TMJ disorders.

Trial Registration: Clinical Trials Registry-India CTRI/2025/06/088238; https://ctri.nic.in/Clinicaltrials/pmaindet2.php?EncHid=MTMzNzQz&Enc=&userName=

International Registered Report Identifier (IRRID): PRR1-10.2196/84038

(JMIR Res Protoc 2026;15:e84038) doi:10.2196/84038

KEYWORDS

areca nut; cone beam computed tomography; CBCT; condyle; condylar morphology; temporomandibular joint

XSL•FO
RenderX

Introduction

Areca nut (AN) has a long history of use. A number of factors contribute to the popularity of AN chewing addiction, such as social acceptance, religious convictions, and perceived health advantages. AN is an ingredient in a variety of items, including *khaini*, *pan masala*, *mawa*, *paan*, and *gutkha*. AN (generally referred to as *Areca catechu* L and betel nut) is the fourth most frequently used psychoactive drug all over the world, after alcohol, nicotine, and caffeine [1].

AN is chewed for various reasons, such as pain treatment, euphoria, refreshment, and increased salivation. Although AN chewing is a highly common habit in Southeast Asia, its use has expanded to North America and Europe. The fruit of the *Catechu* tree produces AN as an endosperm. It is composed of tannins and alkaloids, which include arecoline, arecolidine, guvacine, guvacoline, and arecaidine. Regular consumption of AN may be harmful to one's health. The risk of oral cancer increases with the frequency and duration of AN chewing, highlighting AN as a significant carcinogen. As a psychoactive substance, chronic AN use is categorized by the World Health Organization as a Group 1 human carcinogen, as regular AN chewing is associated with an increased risk of oral squamous cell carcinoma estimated at 2.3%-7.6%. The associated mortality risk is substantial, making this behavior clinically noteworthy [2,3].

Its principal alkaloid, arecoline, is responsible for systemic effects, including cardiovascular disturbances, metabolic disorders such as type II diabetes, immune dysregulation, neurotoxicity, respiratory impairment, and reproductive complications. An association between AN chewing and an increased risk of liver disease has been reported, with several studies demonstrating an additive effect on individuals who are positive for hepatitis B or C. Given its extensive adverse health consequences, AN use poses a significant public health concern [4,5].

Chronic chewing of AN, often kept in the mouth for hours, has been associated with various oral health problems. These include lesions of the mucosa, staining of teeth, dry mouth, periodontal diseases, and plasma cell stomatitis. Furthermore, it contributes to masticatory muscle pain and temporomandibular joint disorders (TMDs), which are common conditions affecting the craniofacial area. The overuse of the temporomandibular joint (TMJ) due to prolonged AN chewing is likely to cause microtrauma and macrotrauma to the TMJ and the masticatory apparatus [6].

Repetitive and prolonged chewing habits, such as AN, betel quid, and tobacco use, subject the masticatory muscles to sustained excessive forces, leading to masseter muscle hypertrophy, a benign condition characterized by an increase in the size of the masseter muscle, which may be unilateral or bilateral. With continued exposure, this mechanical overload results in increased muscle prominence and altered echogenicity of the masseter muscle on imaging [7,8].

Although there is some debate over its relationship with TMD, tooth wear is linked to cumulative parafunctional and oral functions. While modern studies show no clear relationship between tooth loss and TMJ changes, a correlation has been observed in Aboriginal and prehistoric populations. Some studies have claimed that parafunctional habits, including bruxism, clenching, and chewing gum or qat, may be linked to TMD [9].

One study examined the association between occlusal tooth wear and TMD by using the Smith and Knight Tooth Wear Index. Tooth wear was more common in patients with TMD, particularly grade 1 wear in the 26- to 40-year age group and in females, with posterior teeth most frequently affected. TMD is common in young adults and is significantly associated with functional and static occlusal parameters. These findings highlight the importance of evaluating occlusal factors for early diagnosis of TMD. On the basis of the TMD diagnostic criteria, previous studies have hypothesized that excessive tooth wear caused by chronic AN chewing is linked to TMD [9-11].

Mandibular condyle is a crucial bony structure that is often affected by TMD. It forms an articulation with the glenoid fossa, facilitating the movement of the mandible. The mandibular condyle is also the primary growth site in the development of the mandible. Normally, the condyle has a broadly ovoid form and measures approximately 8 mm to 10 mm anteroposteriorly and 15 mm to 20 mm mediolaterally. However, its morphology significantly fluctuates among different individuals and across different age groups. The shape and morphology of the condyle can change due to numerous factors such as normal developmental variability, developmental diseases, tooth wear, parafunctional habits, malocclusion, condylar remodeling, trauma, and disorders of the TMJ [12]. Tooth wear, regardless of its cause, disrupts the normal balance of force distribution, leading to structural adaptation within the condyle. These changes might not show any clear clinical symptoms [9].

A complete evaluation of structural changes in the condyle requires 2 steps: clinical examination and radiographic assessment to confirm the diagnostic findings. Orthopantomograms and transcranial views are some conventional radiographic techniques that have been used to examine the TMJ. However, these 2D images have limitations such as the superimposition of surrounding structures. In contrast, 3D imaging modalities, such as computed tomography (CT), magnetic resonance imaging (MRI), and cone beam CT (CBCT), offer a clearer view of condylar morphology, allowing a more accurate assessment of the TMJ [13].

CT imaging provides high-resolution, detailed images of bony structures. It effectively identifies bone abnormalities such as fractures, osteoarthritis, or degenerative changes in the TMJ. However, CT involves a high-exposure dose of ionizing radiation. It also has drawbacks in showing soft tissues, such as the articular disc, muscles, and ligaments, which are important for diagnosing TMD. Furthermore, CT can be expensive and may be impacted by artifacts from dental hardware, reducing image clarity [14].

MRI is a valuable tool for visualizing the TMJ, as it can produce clear images of soft connective tissues. However, MRI is not the best choice for checking bony abnormalities or fractures in the TMJ, as it is less effective for imaging bones compared to

XSL•FO
RenderX

CBCT. Additionally, patients with metal implants, such as certain dental appliances or prostheses, may not be able to undergo an MRI due to dangers associated with the magnetic field [15].

CBCT is a reliable and effective imaging method for viewing craniofacial structures. This 3D imaging technique provides detailed structural information without overlapping or distortion. It is also preferred because it uses a relatively low radiation dose compared to conventional CT, making it safer for patients. Moreover, CBCT scans are more affordable than conventional CT scanners, making them a cost-effective choice for many dental and orthopedic practices. This combination of detailed imaging, safety, and efficiency makes CBCT an excellent tool for evaluating the shape of the condylar head [16,17].

Numerous studies have investigated the effect of chronic AN chewing and showed significant effects on oral tissues, including the oral mucosa and hard tissues [2]. However, a literature search revealed that few studies have focused on changes in condylar morphology among chronic AN chewers. Therefore, this study aims to use CBCT to compare the condylar morphological changes between chronic AN chewers and nonchewers, addressing the paucity of research on the condylar morphological changes associated with chronic AN chewers.

Methods

A case-control study comprising 90 patients (chronic AN chewer: n=45, 50%) and nonchewer: n=45, 50%) will be conducted between September 2025 and May 2027.

Ethical Considerations

The study will be carried out in Oral Medicine and Radiology department at Sharad Pawar Dental College and Hospital, after receiving approval from the "Institutional Ethics Committee (IEC)" of Datta Meghe Institute of Higher Education and Research (Deemed to be University), Sawangi (Meghe), Wardha, India [Approval No. DMIHER (DU)/IEC/2025/694)] and will be conducted in accordance with the Declaration of Helsinki. The patients, who are willing to participate and provide written informed consent will be included for the study. Participation is voluntary, and all participants may withdraw from the study at any time without any effect on their clinical care. All collected data will be fully de-identified prior to analysis and stored on password-protected, access-restricted institutional systems to ensure confidentiality and data security. No financial compensation will be provided to study participants. CBCT imaging in prospectively enrolled AN chewers will be performed only when clinically indicated and not solely for research purposes, in accordance with the ALARA (As Low As Reasonably Achievable) principle. No additional radiation exposure will be incurred for the control group, as their CBCT scans will be retrospectively analysed. All exposure parameters will be standardized and documented [18-20].

Eligibility Criteria

Inclusion Criteria

Patients aged 18 to 40 years with a habit of chronic AN chewing for more than 1 year and a complete set of permanent dentition, excluding third molars, will be included in the study. Patients showing clinical evidence of generalized attrition and having a normal skeletal relationship will be considered for inclusion. Patients in the control group will have all permanent teeth except wisdom teeth, no history of AN chewing, and no evidence of attrition or reduced vertical height. Tooth wear will be assessed as a covariate by using the Smith and Knight Tooth Wear Index and will not be considered a criterion for group allocation [21].

Exclusion Criteria

Individuals who have undergone TMJ surgery or experienced a fracture, those with congenital TMJ anomalies, patients with a history of trauma from occlusion, those with systemic diseases affecting bone health, patients taking medications that impact bone health, and those with pathological lesions in the jaw or developmental tooth anomalies will be excluded.

Study Procedure

The first step will be a detailed assessment of habit history. This will include chewing frequency per day, average duration of each chewing episode, total years of use, chewing-side preference, and type of AN product consumed. A cumulative "chewing-load index" will be calculated by combining frequency, daily duration, and total years of chewing to create a standardized measure of exposure. Additionally, an analysis of a subgroup within the AN chewer group will compare condylar morphology between the predominant chewing side and the nonchewing side. This will help to evaluate the mechanical effects related to each side. This method improves the measurement of exposure and boosts the study's analytical power.

A complete medical history will also be gathered to rule out any systemic diseases affecting joints, medications influencing joint structure, and any history of trauma or surgery that impacts the TMJ.

A comprehensive clinical examination will be performed for each patient. The TMJ will be examined to identify any disorders. The assessment will include the following:

- Joint movement inspection—evaluation of joint movement was done during mouth opening and closing to detect restricted mouth opening, pain, deviation, or deflection.
- Palpation—palpation of the joint will be performed to assess clicking sounds, crepitus, and tenderness during jaw movements.

Thorough intraoral examination will be done for patients with decreased vertical height due to AN chewing and loss of tooth structure. Tooth structure loss will be recorded based on the Smith and Knight Tooth Wear Index [22].

Each chronic AN chewer will then be evaluated for condylar morphology by performing CBCT. Imaging will be performed using a Planmeca ProMax 3D unit (Planmeca Oy; equipment ID: G-XR-109125), with the following parameters: 90 kV, 6.3 mA current, exposure time of 12.056 seconds, field of view of 5×5 cm, dose-area product of 447 mGy·cm², 251 slices, voxel size of 200 µm, rotation angle of 360°, pixel size of 200 µm, and software version 3.9.6.152. All AN nonchewers will be

XSL•FO
RenderX

evaluated retrospectively, and their dental and medical history, clinical records, and radiographic data (CBCT) will be available.

Condylar morphology will be evaluated using CBCT scans in both sagittal and coronal planes. To calculate the length of the condyle, a line will be drawn connecting the anterior mandibular condyle point and the posterior mandibular condyle point. These points will be situated 4 mm below the superior mandibular condyle point on either side of the condyle [6]. Width of the condylar will be measured in the coronal plane as the linear distance between the medial and lateral mandible poles [6]. In contrast, height of the condyle will be determined as a perpendicular linear distance from the superior mandible condyle to a line drawn between the most inferior point of the sigmoid notch and the tangent of the posterior surface of the ramus in the sagittal plane [6], as shown in Figure 1.

Figure 1. Cone beam computed tomography images images showing condylar morphology measurements: (A) condylar length between ACo–PCo, (B) condylar width between MCo–LCo, and (C) condylar height from SCo to the sigmoid notch. ACo: anterior condyle point; PCo: posterior condyle point; MCo: medial condyle poles; LCo: lateral condyle poles: SCo: superior condyle point.

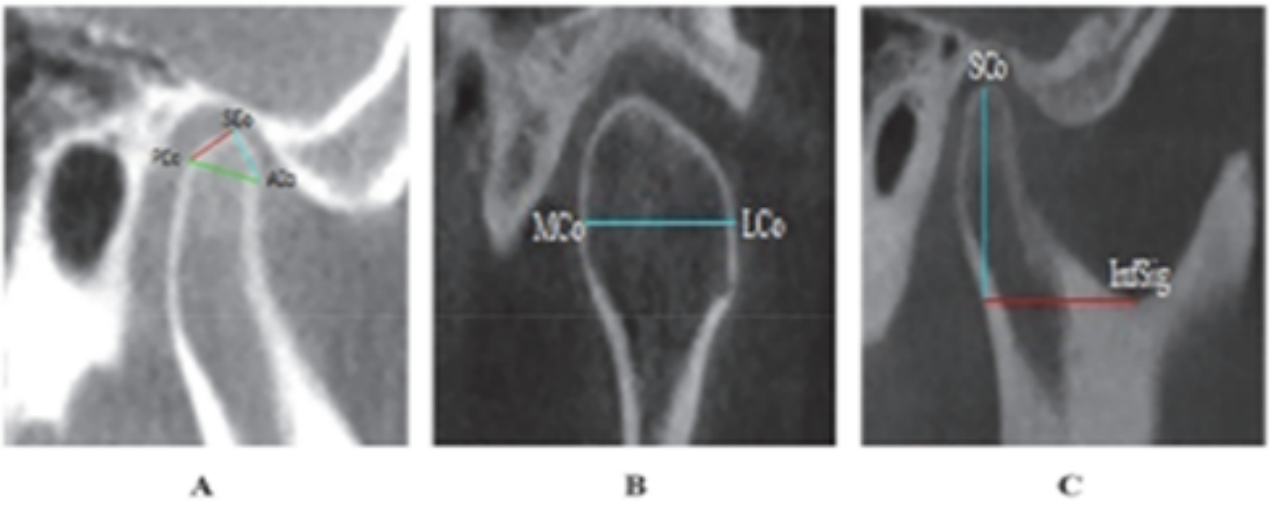

Bias Mitigation

To reduce possible sources of bias, several steps will be taken during the study. Controls will be matched to cases based on age and sex to limit confounding. All participants will receive CBCT imaging with the same machine settings, whenever possible. Retrospective scans will be included only if they were obtained using the same equipment and parameters and fulfill the inclusion criteria. All scan details, such as machine type, 90 kilovoltage (kV), tube 6.3 current (mA), voxel size, field of view, and scan date, will be recorded. Image evaluation will be done by 2 radiologists who are blinded to the participants' habit history. They will use standardized scoring sheets and repeated measurements to ensure consistency between raters. In addition to matching, multivariable analyses will adjust for factors that may differ between groups, such as age, sex, degree of attrition, and duration of the chewing habit. Sensitivity analyses will also be done to check if the findings remain consistent when we focus on scan acquisition parameters. Finally, we will clearly describe the study design, potential biases, and limitations linked to mixed data sources to guarantee transparent reporting.

Sample Size

The sample size for this investigation was determined based on the expected difference in continuous condylar morphometric measurements between groups. The entire sample size estimation procedure, including the full statistical formula and parameter assumptions, is described in Multimedia Appendix 1 [6,23]. On the basis of these predictions, at least 45 participants will be needed per group (chronic AN chewers and nonchewers).

Statistical Methods

This study will compare chronic AN chewers and nonchewers by using random sample selection and strong statistical methods, including the Student 1-tailed *t* test, 1-way ANOVA, and chi-square test. SPSS (version 27.0; IBM Corp) and GraphPad Prism (version 7.0; GraphPad Software) will be used for data analysis, and the Nawaz et al formula will be applied to determine sample size [24]. To ensure scientific reliability, the analytical strategy was adapted from the study by Almashraqi et al [6].

Reliability Assessment

A reliability plan has been included to ensure measurement accuracy. Two radiologists who had attended calibration sessions standardized landmark identification and measurement techniques. This will be according to the principles used in the study by Almashraqi et al [6]. They will assess all CBCT scans independently by using a standardized protocol. Intraobserver and interobserver reliability will be evaluated on 20% of randomly selected scans by using intraclass correlation coefficients, with values ≥0.80 considered acceptable. Any discrepancies will be reviewed and resolved by consensus to ensure consistent and reproducible measurements.

XSL•FO
RenderX

Data Collection, Management, and Analysis

Plan of Data Collection

Patients who report to the Department of Oral Medicine and Radiology will have their AN chewing habit assessed. Following diagnosis, the patients will provide written informed consent and receive a comprehensive explanation of the intervention process. After consent has been obtained, CBCT will be used to analyze the AN chewer prospectively. This will create the following 2 distinct groups: AN chewers (45/90, 50%) and nonchewers (45/90, 50%), based on the patient's habit history and comprehensive clinical examination. To improve transparency and reproducibility, the manuscript includes clearer and more organized reporting of the study design, participant recruitment, and flow. It also defines outcome measures. Additional details specify how CBCT imaging will be done, how clinical and habit-related data will be collected, and how all measurements will be recorded and analyzed. The procedures for radiologists, who were blinded to chewing status and attended calibration sessions, reliability testing, and statistical analysis, have also been described in more detail. This enhances methodological clarity and ensures that the study protocol is fully reproducible. All the results will be documented in a standardized tabular manner, as shown in Multimedia Appendix 2.

Data Collection and Retention

Proper counseling regarding the disease will increase the patient's participation in the study.

Data Management

A master data sheet will be created to include the CBCT image data of the AN chewers and nonchewers, along with the difference in condylar morphology and pathological alterations in each group. All collected data will be assigned coded identifiers, and no personal information will be stored with the clinical or CBCT records. Digital files, including scan data and measurement sheets, will be stored on password-protected institutional servers. Access will be limited to authorized study investigators only. Data will be backed up regularly on secure, encrypted systems to prevent loss or unauthorized access. These measures ensure confidentiality, data integrity, and adherence to institutional ethical standards.

Data Monitoring

Oversight of data accuracy, protocol compliance, and ethical adherence will be maintained by the Department of Oral Medicine and Radiology under the supervision of the guide and the IEC. No interim analyses or stopping guidelines are applicable to this observational study. Data will be analyzed only after completion of enrollment and imaging.

Results

Only individuals fulfilling the inclusion and exclusion criteria will be considered as study subjects. This study is a registered, unfunded protocol study (CTRI/2025/06/088238) and is currently in the pre-recruitment phase. Participant enrollment and data collection are scheduled to be conducted from September 2025 to May 2027. As of the time of manuscript submission, no participants have been recruited and data analysis has not yet commenced. The primary outcome will evaluate condylar morphological changes associated with chronic areca nut (AN) chewing, while secondary outcomes will assess pathological alterations of the condyle, including osteophyte formation with subcortical sclerosis, articular surface flattening of the superior condylar surface with subcortical sclerosis, abnormal condylar shape, and surface erosion. The results are expected to provide insights into the functional impact of chronic AN chewing on the temporomandibular joint and may aid in early identification of pathological condylar changes. Study results are anticipated to be published after completion of data analysis in 2027.

Discussion

Chronic AN chewing has been extensively associated with oral mucosal lesions. However, its impact on the TMJ, particularly on condylar morphology, remains insufficiently studied using CBCT. Almashraqi et al [6] investigated CBCT findings in the TMJ of individuals with a history of chronic qat chewing (QC). They examined TMJ dimensions and osteoarthritic alterations on CBCT. According to the qualitative CBCT results, QC had substantially more osteoarthritic alterations than non qat chewers, including osteophyte, subcortical sclerosis, articular surface flatness, and subcortical cysts. Quantitative assessments of condylar dimensions showed no significant difference among both categories. Notably, the chewing side of individuals with QC displayed more alterations than the nonchewing side. The findings show that chewing qat has negative effects on the TMJ, primarily in the form of osteoarthritic alterations [6].

Liu et al [9] performed a study to investigate the relation between severe tooth wear caused by betel nut chewing and the development of TMD. Age, sex, and severely worn dentition due to betel nut chewing were significant factors contributing to the overall occurrence of TMD. Multivariable analysis revealed a significant, dose-dependent association between intra-articular TMD and badly worn dentition resulting from betel nut chewing, indicating a direct relationship between the severity of tooth wear and the likelihood of developing intra-articular TMD [9].

Al-koshab et al [22] investigated the morphology of the condyle and glenoid fossa by using CBCT in South-East Asians. Condylar length and roof of the glenoid fossa (RGF) thickness did not differ significantly by sex; however, male participants had significantly larger condylar volume, width, height, and joint gaps. While the mean left condylar height and RGF thickness were higher, the mean right TMJ condylar volume, width, and length were significantly higher when comparing the 2 TMJs. With the exception of condylar height, which is larger in Chinese participants, we could not find any significant differences between the 2 ethnic groups' condylar measures and RGF thickness [22].

Khalid Nawaz et al [24] conducted a clinical study on TMJ dysfunction associated with betel nut chewing and reported that long-term chewing of betel nuts can lead in TMJ dysfunction syndrome, as the masticatory forces are transferred to the TMJ. Chewing betel nuts for an extended period can result in TMJ

XSL•FO
RenderX

dysfunction syndrome. The consistency of the betel nut is another important factor in TMJ dysfunction syndrome. Clicking and preauricular pain appeared to be the most common symptoms across all age groups. They concluded that the incidence in female participants was greater than that in male participants [23].

Larheim et al [25] conducted a study on TMJ diagnostics by using CBCT. This study provides an updated perspective on the role of CBCT in evaluating TMJ conditions. The study highlights the diagnostic accuracy and value of CBCT compared to other imaging methods. CBCT is a cost-effective and dose-efficient tool for assessing different TMJ issues, such as osteoarthritis, juvenile osteoarthritis, and rheumatoid arthritis. CBCT offers better images of bone structures than conventional radiographs and MRI scans [25].

Acknowledgments

All authors declare that they had insufficient funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee support for the publication of this article.

Funding

This research did not receive external funding. All study activities, including data collection, CBCT imaging, and analysis, are being conducted using institutional resources, without financial support from funding agencies, sponsors, or commercial entities.

Data Availability

Deidentified datasets created during this study will be deposited in a publicly accessible research data repository after completion of data collection and analysis. Before public release, the data will be available from the corresponding author on reasonable request for research purposes, subject to institutional approval and data-sharing agreements. Any additional materials, including coding sheets or measurement templates, will be provided as multimedia appendices. No identifiable participant information will be shared.

Authors' Contributions

All authors contributed to the development of this study protocol. VT conceptualized and designed the study. VT and SDK collaboratively developed the methodology and planned data collection procedures. They are responsible for cone beam computed tomography conducted data curation and prepared measurement templates. VT wrote the first draft of the manuscript, and SDK critically reviewed, edited, and refined the final version. SDK also supervised the project and managed administration.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample size calculation.
[DOCX File , 15 KB - resprot_v15i1e84038_app1.docx]

Multimedia Appendix 2
Data on the morphological and pathological changes of the condyle.
[DOCX File , 19 KB - resprot_v15i1e84038_app2.docx]

Multimedia Appendix 3
SPIRIT 2025 checklist.
[DOCX File , 42 KB - resprot_v15i1e84038_app3.docx]

References

1. Athukorala IA, Tilakaratne WM, Jayasinghe RD. Areca nut chewing: initiation, addiction, and harmful effects emphasizing the barriers and importance of cessation. J Addict 2021;2021:9967097 [FREE Full text] [doi: 10.1155/2021/9967097] [Medline: 34123457]
2. Shirzaiy M, Neshat F. Effect of areca nut on oral health: a review. J Res Dent Maxillofac Sci 2020 Aug 01;5(3):1-6. [doi: 10.29252/jrdms.5.3.1]
3. Warnakulasuriya S, Trivedy C, Peters TJ. Areca nut use: an independent risk factor for oral cancer. BMJ 2002 Apr 06;324(7341):799-800 [FREE Full text] [doi: 10.1136/bmj.324.7341.799] [Medline: 11934759]

XSL•FO
RenderX

4. Garg A, Chaturvedi P, Gupta PC. A review of the systemic adverse effects of areca nut or betel nut. Indian J Med Paediatr Oncol 2014 Jan 19;35(1):3-9 [FREE Full text] [doi: 10.4103/0971-5851.133702] [Medline: 25006276]
5. Khasbage SB, Bhowate RR, Khatib N. Risk of liver disease in areca nut habitual: a systematic review. J Oral Maxillofac Pathol 2022;26(1):128-129 [FREE Full text] [doi: 10.4103/jomfp.jomfp_345_21] [Medline: 35571298]
6. Almashraqi AA, Halboub ES, Al-Maweri SA, Barngkgei I, Al-Wesabi MA, Al-Kamel A, et al. Cone beam computed tomography findings in temporomandibular joint of chronic qat chewers: dimensional and osteoarthritic changes. J Oral Rehabil 2020 Dec 30;47(12):1538-1549. [doi: 10.1111/joor.13092] [Medline: 32955738]
7. Anandan P, Saravanan K, Raj R.R M. Evaluation of internal echogenic pattern and thickness of masseter muscle with ultrasonography in chronic areca nut/ tobacco chewers and oral submucous fibrosis patients in a tertiary care hospital Chengalpattu District, Tamilnadu. J Neonatal Surg 2025 Jun 09;14(27S):1153-1158. [doi: 10.63682/jns.v14i27s.7229]
8. Raghunandan Iyengar A, Patil S, Guddannanavar Karibasappa G, Beloor Vasudev S, Kumar Joshi R. Evaluation of internal echogenic pattern of masseter in subjects with myofascial pain/ myositis, oral submucous fibrosis, chewers, bruxers and healthy individuals- a preliminary ultrasonographic study. J Dent (Shiraz) 2016 Dec;17(4):361-366 [FREE Full text] [Medline: 27942553]
9. Liu Y, Yin T, He M, Fang C, Peng S. The association of severely worn dentition resulting from betel nut chewing with temporomandibular disorders: a cross-sectional study. BMC Oral Health 2023 Jul 07;23(1):459 [FREE Full text] [doi: 10.1186/s12903-023-03160-2] [Medline: 37420243]
10. Fathima F, Dharman S, Senthil Murugan P. Association of occlusal wear facets in patients with temporomandibular disorders. Bioinformation 2020 Dec 31;16(12):1060-1068 [FREE Full text] [doi: 10.6026/973206300161060] [Medline: 34938006]
11. Trehan M, Sharma P, Tiwari A, Kumar A, Anamika, Bamal R. Prevalence of temporomandibular disorders in young adults and their association with functional and static occlusal parameters: a cross-sectional study. J Indian Orthod Soc 2025 Sep 17;59(4):325-333 [FREE Full text] [doi: 10.1177/03015742251374388]
12. Sahithi D, Reddy S, Divya Teja D, Koneru J, Sai Praveen KN, Sruthi R. Reveal the concealed – morphological variations of the coronoid process, condyle and sigmoid notch in personal identification. Egypt J Forensic Sci 2016 Jun;6(2):108-113. [doi: 10.1016/j.ejfs.2015.11.003]
13. Ahmad M, Hollender L, Anderson Q, Kartha K, Ohrbach R, Truelove EL, et al. Research diagnostic criteria for temporomandibular disorders (RDC/TMD): development of image analysis criteria and examiner reliability for image analysis. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 2009 Jun;107(6):844-860. [doi: 10.1016/j.tripleo.2009.02.023] [Medline: 19464658]
14. Scarfe WC, Farman AG, Sukovic P. Clinical applications of cone-beam computed tomography in dental practice. J Can Dent Assoc 2006 Feb;72(1):75-80. [Medline: 16480609]
15. Bag AK, Gaddikeri S, Singhal A, Hardin S, Tran BD, Medina JA, et al. Imaging of the temporomandibular joint: an update. World J Radiol 2014 Aug 28;6(8):567-582 [FREE Full text] [doi: 10.4329/wjr.v6.i8.567] [Medline: 25170394]
16. Dhabale GS, Bhowate R. Cone-beam computed tomography for temporomandibular joint imaging. Cureus 2022 Nov;14(11):e31515. [doi: 10.7759/cureus.31515] [Medline: 36532912]
17. Pauwels R, Araki K, Siewerdsen JH, Thongvigitmanee SS. Technical aspects of dental CBCT: state of the art. Dentomaxillofac Radiol 2015 Jan;44(1):20140224 [FREE Full text] [doi: 10.1259/dmfr.20140224] [Medline: 25263643]
18. Pauwels R. Cone beam CT for dental and maxillofacial imaging: dose matters. Radiat Prot Dosimetry 2015 Jul;165(1-4):156-161. [doi: 10.1093/rpd/ncv057] [Medline: 25805884]
19. Scarfe WC, Farman AG. What is cone-beam CT and how does it work? Dent Clin North Am 2008 Oct;52(4):707-730. [doi: 10.1016/j.cden.2008.05.005] [Medline: 18805225]
20. Oakley PA, Harrison DE. Death of the ALARA radiation protection principle as used in the medical sector. Dose Response 2020 Apr 29;18(2):1559325820921641 [FREE Full text] [doi: 10.1177/1559325820921641] [Medline: 32425724]
21. López-Frías FJ, Castellanos-Cosano L, Martín-González J, Llamas-Carreras J, Segura-Egea J. Clinical measurement of tooth wear: tooth wear indices. J Clin Exp Dent 2012 Feb;4(1):e48-e53 [FREE Full text] [doi: 10.4317/jced.50592] [Medline: 24558525]
22. Al-koshab M, Nambiar P, John J. Assessment of condyle and glenoid fossa morphology using CBCT in South-East Asians. PLoS One 2015 Mar 24;10(3):e0121682 [FREE Full text] [doi: 10.1371/journal.pone.0121682] [Medline: 25803868]
23. Wang H, Chow SC. Sample size calculation for comparing proportions. In: D'Agostino RB, Sullivan L, Massaro J, editors. Wiley Encyclopedia of Clinical Trials. Hoboken, NJ: John Wiley & Sons; 2007.
24. Khalid Nawaz MK, Sivaraman G, Krishnamoorthy S, Balaji S. Temporomandibular joint dysfunction syndrome associated with Betel Nut Chewing: a clinical study. J Orofac Res 2015;5:142-145. [doi: 10.5005/jp-journals-10026-1199]
25. Larheim TA, Abrahamsson A, Kristensen M, Arvidsson LZ. Temporomandibular joint diagnostics using CBCT. Dentomaxillofac Radiol 2015 Jan;44(1):20140235 [FREE Full text] [doi: 10.1259/dmfr.20140235] [Medline: 25369205]

Abbreviations

AN: areca nut
CBCT: cone beam computed tomography

XSL•FO
RenderX

CT: computed tomography
IEC: institutional ethics committee
MRI: magnetic resonance imaging
RGF: roof of the glenoid fossa
TMD: temporomandibular disorder
TMJ: temporomandibular joint
QC: qat chewing

Edited by J Sarvestan; submitted 13.Sep.2025; peer-reviewed by M Musa, MM Kostik; comments to author 14.Nov.2025; accepted 27.Dec.2025; published 06.Mar.2026.

Please cite as:
Tarpe V, Khasbage SD
3D Comparative Evaluation of Condylar Morphology Between Chronic Areca Nut Chewers and Nonchewers: Protocol for a Case-Control Study
JMIR Res Protoc 2026;15:e84038
URL: https://www.researchprotocols.org/2026/1/e84038
doi:10.2196/84038
PMID:

©Vaishnavi Tarpe, Suwarna Dangore Khasbage. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 06.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Perceptions of Abortion and Sexual and Reproductive Health in Chilean Medical and Midwifery Education: Protocol for a Mixed Methods Study

Lidia Casas-Becerra[1], PhD; Alejandra Ramm[2], PhD; Beatriz Pérez[3], PhD; Heidy Kaune[4,5], PhD; Mercedes Carrasco-Portiño[6], PhD; Adela Montero[7], MSc; Cecilia Bustos-Ibarra[8], PhD; Henry Castro-Arias[9], MD; Claudio López-Labarca[10], MSc; Leonardo Reyes-Torres[10], MSc

[1]Law School, Universidad Diego Portales, Santiago, Metropolitana, Chile

[2]Escuela de Sociologia, Facultad de Ciencias Sociales, Universidad de Valparaíso, Valparaíso, Chile

[3]Departamento de Psicología, Universidad de la Frontera, Temuco, Chile

[4]Programa de Ética y Políticas Públicas en Reproducción Humana, Facultad de Medicina, Universidad Diego Portales, Santiago, Chile

[5]Centro de Investigación Biomédica, Facultad de Medicina, Universidad Diego Portales, Santiago, Chile

[6]Department of Obstetrics and Childcare, School of Medicine, Universidad de Concepción, Concepción, Chile

[7]Center for Reproductive Medicine and Integral Development of Adolescence, Faculty of Medicine, Universidad de Chile, Santiago, Chile

[8]Department of Social Work, Faculty of Social Sciences, Universidad de Concepción, Concepción, Chile

[9]Department of Obstetrics and Gynecology, School of Medicine, Universidad de Concepción, Concepción, Chile

[10]Department of Obstetrics and Childcare, Universidad de Atacama, Copiapó, Chile

Corresponding Author:
Lidia Casas-Becerra, PhD
Law School
Universidad Diego Portales
República 112
Santiago, Metropolitana, 7750171
Chile
Phone: 56 226768849 ext 8849
Email: lidia.casas@udp.cl

Abstract

Background: University education for physicians and midwives in sexual and reproductive health (SRH), particularly regarding abortion, is shaped by broader societal debates, which are often polarized. Teaching at faith-based universities might limit the scope and quality of SRH education. The study is contextualized within Chile's shifting legal and social landscape following the landmark 2017 reform under President Michelle Bachelet. This reform partially lifted Chile's absolute ban on abortion, permitting it under 3 circumstances: risk to the mother's life, fatal fetal anomaly, and a pregnancy resulting from rape. The protocol presented here intends to examine how the enacted legislation plays out in the curricula and how abortion is taught in secular and faith-based universities, which often take conflicting stances on abortion.

Objective: This study aims to improve our understanding of how SRH, particularly abortion, is being approached by both students and faculty of medicine and midwifery in Chile. Specifically, it provides comparative evidence on attitudes, perceptions, and experiences related to abortion among medical and midwifery students and faculty in Chile.

Methods: The study uses a cross-sectional and mixed methods design, with the participation of students and faculty within Chilean medical and midwifery programs. An interdisciplinary team developed and validated 2 instruments through an iterative process: a quantitative survey for undergraduate students to assess knowledge, attitudes, and competencies regarding abortion, and a semistructured interview guide for faculty to explore curricular decision-making. The instruments underwent expert content validation and pilot testing to ensure relevance to the local legal and educational context. Quantitative data will be analyzed using descriptive and inferential statistics, while qualitative findings will undergo thematic analysis to triangulate the current state of abortion training.

Results: The funding for this study was awarded by the Diego Portales University (CG 1110689019), November 2022. Data collection for both the quantitative and qualitative components was completed as of July 2025. Recruitment finalized with 2795

XSL•FO
RenderX

students for the online survey and 57 faculty members for the semistructured interviews. Data analysis is in progress, and we expect to publish the findings from June 2026 onwards.

Conclusions: This protocol examines the intersection of law, education, and public health by evaluating abortion training for Chilean medical and midwifery students. Transcending ideological debates to prioritize professional competency and health equity, the findings will guide educators and policymakers in preparing technically and ethically competent health care professionals.

International Registered Report Identifier (IRRID): DERR1-10.2196/81427

(JMIR Res Protoc 2026;15:e81427) doi:10.2196/81427

KEYWORDS

ethics; public health; medical education; reproductive health; health services accessibility

Introduction

Abortion rights have long dominated the agenda of Chilean feminist organizations. In 2017, under the leadership of then-President Michelle Bachelet, abortion laws evolved from a blanket ban to partial decriminalization [1]. Law 21.030 allows abortion when a woman's life is at risk, when the fetus is fatally impaired, and when pregnancy is the result of rape [2]. As expected, the proposition to allow abortion on these grounds was fiercely resisted by conservative, religious, and right-wing groups. Chilean universities and academics became strongly engaged in the debate. While many scholars from non–faith-based universities argued in favor of the proposed legislation, their peers from faith-based universities argued forcefully against it, exposing deep fractures between faith-based and secular universities [1].

To be sure, Catholic opposition was led not by clergy but by the president of the Pontifical Catholic University, a trained physician who argued that institutions, including private universities, clinics and hospitals, should also be allowed to claim conscientious objector status [1,3,4]. Once the new legislation was passed, in a bid to prevent its enactment, conservative groups asked the Constitutional Court to review its constitutionality [5]. The Court ruled to confirm, but agreed to allow conscientious objector status to be extended to institutions, and to the rest of health care professionals, including technicians, who work in the surgical ward during the procedure. Faith-based universities operating clinics and teaching hospitals swiftly moved to do so.

A majority of Chileans do approve of abortion under certain circumstances, with surveys confirming widespread support, especially on the currently allowed 3 grounds [6,7]. A recent survey conducted by the respected Centre for Public Studies shows that support for abortion is actually growing. Indeed, asked in 1999 if "Women should always be able to choose an abortion," 10% of respondents agreed; asked again in 2024, 34% agreed [7]. Likewise, support for abortion under "special circumstances" rose from 35% in 1999 to 50% in 2024. In other words, about 84% of survey respondents supported abortion in certain cases. Agreement with the view that "Abortion should always be banned" slipped from 55% in 1999 to 15% in 2024. Young people aged 18 to 24 years were especially supportive. In 2024, 45% agreed that "Women should always be able to choose an abortion" (in contrast to 27% agreement among respondents aged 55 years and older). Respondents with higher education also showed greater support for voluntary pregnancy termination [7].

Structural reforms spearheaded by the Pinochet dictatorship (1973-1990) laid the foundation for significant private involvement in providing public services such as retirement pensions, health care, and education. At the time, institutions of higher education numbered eight, 6 private and 2 public [8], all of which received State funding and charged no tuition. Along with promoting private involvement in higher education, the Pinochet regime also sought to cut public spending by transforming the 14 regional campuses of the large 2 State universities into as many freestanding universities. Today, both private and State universities are funded mostly through tuition fees. Since tuition is expensive, subsequent governments augmented funding support, and Bachelet introduced in 2018 free tuition for lower-income students.

Ministry of Education data indicates that in 2024, there were 58 universities, including 12 public and 46 privately-owned. At present, private universities account for more than 70% of undergraduate enrollment. Most universities are secular (44), while 14 are faith-based (all Catholic except for 1 Adventist) [9]. Universities' core values and mission statements might limit or veto curricular contents while shaping training delivery in particular ways. This has specific implications for the medical and midwifery professions, and more generally for the integration of young practitioners into the labor market and for the proper implementation of laws and public policy that liberalize abortion or guarantee health services.

Previous research surveyed the views and attitudes of medical and midwifery students and faculty members at the time of enactment of Law 21.030 across universities based in the Metropolitan region, where the capital city of Santiago is located [10-15]. Although the Metropolitan region is where most universities are based, previous research did not account for subnational differences. Findings revealed that the core values, whether secular or faith-based, of universities shape the content, methodology, and scope of teaching. Likewise, another study also showed that students' attitudes toward sexual and reproductive health (SRH) differ according to their home regions. A multicenter study of SRH literacy in Chile found that students of medicine and health sciences have higher levels of knowledge compared to students from social sciences programs, but this varies by region [16]. Specific educational

XSL•FO
RenderX

approaches further shape the attitudes and views of future health care practitioners. Against this backdrop, our study intends to delve into the attitudes, perceptions and experiences of Chile's medical and midwifery students and faculty members regarding the SRH training they receive or deliver, with a special focus on abortion.

The role of universities in medical education and in the provision of abortion services has been studied in different settings, both in more liberal contexts, such as the United Kingdom [17], Germany [18], Canada [19], and Australia [20], and in more restrictive contexts, such as Brazil and Chile [21,22], with a focus on stigma, conscientious objection (CO), and the availability of services. In the United States, some studies share a concern about the impact of the Supreme Court's decision in Dobbs (2022) on medical training and access to abortion services [23,24].

These studies reveal significant shortcomings in the training of medical students regarding abortion and the care of those who require these health services. Although students may support laws permitting abortion, their training is often deficient. A survey of medical students in Germany showed strong support for existing legislation and the belief that the law should be further liberalized. Students also expressed willingness to receive training in the future; however, 71% reported dissatisfaction with the curriculum at their faculty, and 75.5% of upper-year students indicated that they had not received any clinical training in abortion care [18].

A literature review also examined the extent and quality of education received by midwifery and nursing students, who are often considered more supportive professionals in relation to women's reproductive health. The review found that, regardless of the legal context, abortion training is not prioritized and is difficult to access due to institutional barriers [25].

In Chile, there are no regulatory guidelines for training in Medicine or Midwifery regarding abortion. The Association of Medical Faculties of Chilean universities establishes a graduate profile outlining the competencies students must acquire; yet, it does not specify content regarding SRH [26]. However, the National Examination of Medical Knowledge, which is mandatory for medical graduates, qualifying them to practice in public health services and apply for medical specialties, includes content on spontaneous and induced abortion, as well as family planning. Nevertheless, medical schools retain the discretion to decide whether to teach these contents [27].

In the case of Obstetrics, Article 117 of the Sanitary Code establishes that midwifery professionals are responsible for SRH care and for carrying out procedures derived from medical diagnosis and treatment [28].

The international context described above prompts us to ask about the current state of abortion-related education received by medical and midwifery students in Chile, 7 years after the implementation of the law that partially decriminalized abortion in the country. As universities train the health care practitioners of the future, there is a need to assess whether curricular changes have been implemented to address cultural resistance by practitioners and instructors who either explicitly or implicitly convey negative attitudes toward abortion and reproductive autonomy. After the law entered into force, Robledo et al [29] found that the Health Ministry had to update physicians on vacuum aspiration techniques not taught in medical schools. Much criticism has emerged ever since regarding the poor knowledge, training, and awareness of health care practitioners about the new law [30], all factors that exacerbate the stigma surrounding abortion [30,31]. Key barriers to access abortion services identified include pervasive CO claims, particularly in abortion due to rape, revealing important regional differences [32,33], religiousness, and the stigma associated with performing an abortion [34,35]. Additional challenges include structural public health issues, such as underfunding, accessibility for the population in rural communities, insufficient mental health personnel, and lack of information campaigns about the law [32], given the connection between sexuality, reproduction, and abortion [36,37].

In this context, this study seeks to gain a deeper understanding of the views, perceptions, attitudes, and experiences of medical and midwifery students and faculty at Chilean universities regarding SRH education, with a particular focus on abortion. Specifically, the study aims to: (1) examine how SRH, and abortion in particular, is taught in medical and midwifery programs, including faculty members' approaches to teaching and both students' and faculty members' perceptions of the training received; (2) ascertain the views and attitudes of medical and midwifery students and faculty members toward SRH, particularly in relation to abortion; (3) examine the perceptions and attitudes of both faculty members and students regarding patient confidentiality, CO, and the reporting of both lawful and unlawful abortions; and (4) explore students and faculty willingness to undergo training on abortion procedures and their views on facilitating access to legal abortion services.

We hypothesize that cultural changes may influence overall favorable opinions toward abortion among students in both degree programs, and that students from faith-based universities may report more stigmatizing attitudes toward abortion, more favorable views of CO, greater acceptance of reporting women for unlawful abortion, and more limited technical training in abortion care compared to students in secular institutions. We hypothesize that midwifery students may exhibit higher levels of empathy toward women compared to medical students. In addition, we anticipate identifying regional disparities, whereby students may encounter educational barriers associated with local cultural conservatism or limited access to clinical training. By examining the relationship between faculty attitudes and students' perceptions, this study aims to explore how hidden curricula may shape the willingness of future professionals to provide abortion care, with potential implications for the implementation of the law and the protection of reproductive rights in Chile.

This study aims to improve our understanding of how SRH, particularly abortion, is being approached by both students and faculty of medicine and midwifery in Chile. Specifically, it provides comparative evidence on attitudes, perceptions, and experiences related to abortion among medical and midwifery students and faculty in Chile. Furthermore, its findings will provide evidence-based insights relevant for curriculum

XSL•FO
RenderX

development, faculty training, and institutional guidelines, with potential implications for strengthening abortion-related training and improving the implementation of lawful abortion services. Ultimately, this study seeks to contribute to a more equitable access to abortion services and quality care.

Methods

Convergent Mixed Methods Design

A convergent mixed methods design is applied [38], which aims to achieve an enhanced understanding of a research problem through comparing and merging quantitative and qualitative research. As Creswell asserts, quantitative and qualitative data "provide different insights, and their combination contributes to seeing the problem from multiple angles and multiple perspectives [...] their combination adds up to not only more data but also a more complete understanding of the problem" [38]. In this kind of design, data gathering and analysis are done separately—usually in parallel—and then merged. Comparison is made possible through addressing similar topics and questions through both types of research (quantitative and qualitative).

In the case of the current research, we defined 2 distinctive yet complementary populations, medical and midwifery students and faculty, and we aimed to identify their perceptions, experiences, and attitudes toward abortion. Data gathering instruments—online survey questionnaires for students and semistructured interviews for faculty—address the same topics and also some specific ones for each population. After both types of results are obtained—quantitative and qualitative—inferences and interpretations from them are brought together in a discussion, to make meta-inferences or conclusions. We will first report the quantitative results to provide a "big picture" of tendencies toward abortion among students of medicine and midwifery. This will be followed by the qualitative findings on the same issues, based on interviews with faculty members, which will offer more detailed, contextualized, and in-depth insights. We will then conduct a follow-up discussion comparing results from both populations, students and faculty, seeking points of convergence as well as divergence. From this comparison, we will draw broader inferences, conclusions, and interpretations, which will in turn be contrasted with the existing literature on the subject. In doing so, we, and other researchers, will be able to assess whether our conclusions provide a valuable explanation of the proposed research problem. It should be highlighted that pragmatism has been frequently linked to mixed methods research [39-41], and hence this research draws upon pragmatism, applying a "pragmatic approach" [42,43]. Table 1 shows the overall study workflow, the estimated duration of each phase, and the participants involved in each of them.

XSL•FO
RenderX

Table 1. Study phases, duration, and participants.

Study phases	Duration (months)	Experts and participants
1. Preliminary phase	8	Research team (11 academic experts in law, medicine, midwifery, sociology, and psychology)
1.1 Quantitative: instrument development and content validity through expert judgement		
1.2 Qualitative: development and expert review of the semistructured interview guide		
2. Pilot study	3	
2.1 Quantitative: comprehension of instructions and preliminary psychometric evaluation		50 Chilean medical and midwifery students
2.2 Qualitative: comprehension and adjustment of the semistructured guide		5 Chilean academics in medicine and midwifery
3. Sampling phase	12	
3.1 Quantitative sample selection and recruitment		1788 Chilean medical and midwifery students
3.2 Qualitative sample selection and recruitment		52 Chilean academics in medicine and midwifery
4. Data analysis	12	—[a]
4.1 Quantitative: psychometric study of the instruments and multivariate data analysis		
4.2 Qualitative: coding and deductive and inductive analysis		
5. Dissemination of results		
5.1 Congresses and seminars in different regions of the country	24	11 research team members
5.2 Scientific manuscripts	24	11 research team members
5.3 Presentation to university authorities	12	11 research team members

[a]Not applicable.

Participant Selection and Recruitment

In accordance with the convergent mixed methods design, and in order to enable comparison between the 2 populations (students and faculty), participant selection criteria are similar for both the student survey and faculty semistructured interview. These criteria are (1) type of degree program (medicine or midwifery), (2) type of university (secular or religious), (3) geographical area (Northern, Central, Metropolitan, and Southern), and (4) sex (female and male). With regard to gender, it should be noted that students and faculty members who self-identify as lesbian, gay, bisexual, transgender, and queer or questioning, and others are not excluded from the study.

The planned quantitative study involves applying a multivariate correlational, cross-sectional, observational design to a nationwide sample of medical and midwifery students using a range of strategies. The study is designed to be conducted in each of the country's 3 macroregions: in the North, in one of its 5 regions; in the center, in one of its 3 regions; and in the South, in 3 of its 7 regions, because of its larger geographical area compared to the other macro-regions. In addition, the Metropolitan area will be considered independently, as it corresponds to a region within the Central zone where the nation's capital and the highest concentration of population and universities are located; to protect participant identity, specific regions will not be named.

In 2022, total enrollment at medical and midwifery schools was 30,862, distributed across Northern Chile (n=1686, 5.46%), Central Chile (n=10,972, 35.55%), Southern Chile (n=4138, 13.40%), and Metropolitan Santiago (n=14,066, 45.57%) [44]. Regions were treated as independent populations based on accessibility and enrollment in the target fields. For sample size calculation, a stratified sampling design by geographic region was used. Stratum-specific sample sizes were calculated to estimate proportions with a 95% confidence level and ±5% precision, assuming P=0.5 and applying finite population correction. Based on 2022 enrollment figures, the resulting target sample sizes were Metropolitan Santiago (P=14,066; n=375), Northern Chile (P=433; n=205), Central Chile 1 (P=2803; n=339), Central Chile 2 (P=4124; n=281), Southern Chile 1 (P=1497; n=307), and Southern Chile 2 (P=1034; n=281). The total sample size corresponds to the sum of stratum-specific targets (n=1788), with adjustment for anticipated nonresponse.

Once the target sample size was defined for each region, the distribution and selection of participants were conducted using a nonprobabilistic stratified sampling method, which consists of dividing the population into homogeneous subgroups (strata)

XSL•FO
RenderX

and selecting participants by convenience within each one. This stratification is based on three key variables due to their correlation with attitudes toward abortion: (1) type of degree program, (2) type of university (secular or religious), and (3) gender. The sample was designed to include equal proportions of students from medicine (n=894, 50%) and midwifery (n=894, 50%), as well as from secular (n=894, 50%) and religious (n=894, 50%) universities, where both types of institutions were available. In some selected regions (one in the northern zone and one in the southern zone), there are no religious universities offering these programs; only secular universities were included in those areas. Gender stratification was applied within each field of study to reflect the national enrollment distribution (medicine: n=392, 43.8% men and n=502, 56.07% women; midwifery: n=40, 4.46% men and n=854, 95.53% women), while also ensuring adequate participation of men, given their typically lower response rate. The survey was distributed, and students were recruited through a variety of methods, notably (1) institutional contacts, (2) personal team contacts, and (3) social media invitations (Facebook [Meta], Instagram [Meta], and WhatsApp [Meta]).

Participant selection of faculty for the semistructured interviews followed the same criteria used to stratify the quantitative sample: type of degree program, type of university, geographical area, and gender. In accordance with the study's pragmatic approach, the total number of faculty and their distribution across the sampling variables were determined based on considerations of information power, that is, the sufficiency of the number and diversity of cases [45], as well as practical considerations related to the feasibility of conducting the proposed interviews. A sample of 52 faculty was designed. Table 2 shows the distribution of the qualitative sample.

Of the total sample, half of the cases (n=26) are from the medical faculty and the remaining half (n=26) from the midwifery faculty. A similar distribution is planned in terms of gender, with equal numbers of women and men. Given the smaller number of religious universities compared to secular institutions, and the greater reluctance of faculty from religious universities to participate in interviews on SRH issues, particularly abortion, the sample is estimated to include 20 faculty members from religious universities and 32 from secular universities. Regarding subnational diversity, the number of cases in each geographical area is determined based on the number and type (secular or religious) of universities offering medical and midwifery training in each region. Accordingly, 16 interviews are allocated to the Metropolitan area and 16 to the Southern region, as both geographical areas have the highest concentration and diversity of universities nationwide. In contrast, 8 interviews are allocated to the Northern region, reflecting the lower number of universities and the absence of religious institutions. An intermediate number of 12 interviews is allocated to the Central region, given the limited presence of religious universities.

Table 2. Qualitative sample distribution[a].

Geographical zone and gender	Confessional universities		Secular universities		Total
	Medicine	Midwifery	Medicine	Midwifery	
Northern					
Female	0	0	2	2	4
Male	0	0	2	2	4
Central					
Female	1	1	2	2	6
Male	1	1	2	2	6
Metropolitan					
Female	2	2	2	2	8
Male	2	2	2	2	8
Southern					
Female	2	2	2	2	8
Male	2	2	2	2	8
Total	10	10	16	16	52

[a]In the Northern area, there are no confessional universities with schools of medicine and/or midwifery. While in the Central zone, only one of its component regions has universities with schools of medicine and/or midwifery.

Recruitment is designed to proceed through formal and informal team contacts, and respondents will be asked to provide leads for additional interviews. As stated, it is anticipated that securing interviews with academic staff from faith-based universities will pose a challenge. Unless prospects prefer an in-person session, interviews will be conducted remotely on video conferencing platforms such as Zoom. Respondents will be asked to be in a location that guarantees privacy. Conducting interviews remotely is considered more appropriate due to participants' time constraints. In addition to teaching duties, most medical and midwifery lecturers are health care providers and have little free time. Sessions are designed to last a maximum of 30 to 45 minutes.

XSL•FO
RenderX

Data Collection Instruments

A student survey and a faculty interview guide were developed through a rigorous, iterative process carried out collaboratively by an interdisciplinary research team comprising law scholars, physicians, midwives, sociologists, and psychologists. All team members are academics with expertise in SRH, health law, bioethics, and/or medical education in Chile. Both the faculty interview guide and the student survey were developed based on previous empirical and theoretical research in the field [9,11-14]. The research team met on a weekly basis over an 8-month period to develop the student survey and the faculty interview guide.

In relation to the student survey, the research team defined the conceptual and theoretical framework underlying each construct, reviewed existing validated instruments, and adapted or developed items to ensure contextual relevance to the Chilean legal, educational, and health care settings. This process provided expert-based content validation, allowing for the refinement of item wording, response formats, and scale structure prior to pilot testing.

The student survey comprises a variety of instruments, including a questionnaire designed to collect sociodemographic data, such as age, gender, ethnicity, political and religious affiliation, church attendance, type of school attended, nationality, type of degree program, and type of university attended (public or private; secular or faith-based; region). Other instruments are listed in Table 3. Instruments using a 5-point Likert scale have response options consistently labeled, ranging from 1=Strongly disagree to 5=Strongly agree. Higher scores indicate a greater presence of the construct being measured.

XSL•FO
RenderX

Table 3. Student survey scales.

Factors and definition of constructs	Items	Scale and comments
1. Scale on the Centrality of Political affinity in Self-Concept [31]		
Unidimensional. Prominence and subjective importance that participants attach to identification with a right-wing political orientation in the construction of their self-concept	3	5-point Likert scale. In addition, it presents two additional, independent items on the left and center political orientation
2. Scale on Conscientious Objection to Abortion: Positioning and Reasons ad hoc		
4 factors. Factors 1, 2, and 3 capture the motivations for declaring CO[a]: (1) Value, ethical, or religious reasons; (2) fear of consequences; (3) professional reasons; and (4) intention to declare conscientious objection	17	5-point Likert scale. These factors can be combined to detect pseudo-CO, ie, the declaration of CO for reasons other than those legally supported: fear of consequences and professional reasons
3. Conscientious Objection to Abortion Knowledge Scale ad hoc		
Unidimensional. Knowledge of the rights and obligations of CO to abortion, as required by Chilean law	4	3 response options: “Yes,” “No,” or “Don't know”
4. Support for Individual and Institutional Conscientious Objection ad hoc		
Nonpsychometric. Support for the legality of individual and institutional CO and associated competencies	7	Various response scales
5. Intentionality to Promote Abortion Law Compliance Scale ad hoc		
2 factors. Participants' intentionality to facilitate or hinder access to legalized abortion: (1) promote compliance with the law and (2) hinder compliance with the law	6	5-point Likert scale
6. Abortion Attitudes Scale: Criminalization, Acceptance, and Entitlement ad hoc		
3 factors: (1) criminalization, beliefs about the conditions under which an abortion should be punishable by imprisonment; (2) acceptability, beliefs about the conditions under which an abortion is acceptable, and finally; and (3) right, belief in abortion as a right	19	5-point Likert scale
7. Abortion Attitudes Scale: Reporting and Confidentiality in Professional Practice ad hoc		
3 factors: (1) willingness to report. This factor assesses the belief in the professional duty to report abortion situations not covered by the law; (2) risks associated with not reporting. This dimension measures beliefs about the risks of not reporting abortion situations not protected by law; and (3) willingness to maintain confidentiality. Preference for Confidentiality, based on the belief that professional secrecy should prevail over reporting in abortion situations not protected by law	20	5-point Likert scale
8. Abortion Attitudes Scale: Best Friend Support ad hoc		
Unidimensional. Intention to help a woman with whom one is a friend who seeks an illegal abortion	7	5-point Likert scale
9. School sex education and university training in SRH[a] ad hoc		
Nonpsychometric. Type and quality of sexuality education received at school and university, and the opinion about university education on abortion issues	6	Various response scales

[a]CO: Conscientious Objection.

[a]SRH: sexual and reproductive health.

Pilot Study

The pilot study helped to assess the comprehension of instructions and items and provided a preliminary psychometric evaluation of the instruments. After the exclusion of 5 participants due to incomplete responses, the final sample consisted of 50 students, exceeding recommended sample size guidelines for this type of study based on statistical considerations and decision-making requirements for early-stage instrument development [46,47].

Of the total participants, 66% (33/50) were women. Overall, 54% (27/50) were enrolled in medicine, while the remaining 46% (23/50) were studying midwifery. Regarding level of training, most participants were in advanced years of their programs (first to fourth year: 20/50, 40%; fifth to seventh year: 30/50, 60%). In addition, 36% (18/50) were enrolled in universities with a religious orientation. The sample also included representation from all macro-regions of the country (Northern: 5/50, 9.1%; Central: 22/50, 40%; Metropolitan: 12/50, 21.8%; Southern: 11/50, 20%). Overall, this diversity contributed to meeting the representativeness recommendations proposed by Johanson and Brooks [47] for pilot studies.

Decisions regarding scale refinement were based on analyses of inter item correlations (bivariate correlations between items),

XSL•FO
RenderX

internal consistency, and corrected item–total correlations, both for the full scales and for their theoretical factors. These criteria were applied systematically in order to identify redundant items or items showing weak associations with the overall scale, prioritizing parsimonious and conceptually coherent versions of the instruments. In addition, participants' subjective evaluations of the clarity and comprehensibility of instructions and items were considered:

1. Scale on the Centrality of Political Affinity in Self-Concept (CPolA): the research team explored the possibility of expanding this previously published scale [48] by adding 6 items intended to measure the centrality of centrist and left-wing political orientations. However, analyses indicated conceptual redundancy. Consequently, only 2 additional items, one for each political orientation, were retained as a complement to the original scale.
2. Scale on Conscientious Objection to Abortion: Positioning and Reasons (COA-PR), ad hoc. Initially composed of 19 items, 1 item was removed due to low association, while another item corresponding to the theoretical factor "Intention to Object" was transferred to the SOC-Ind/Ins scale based on theoretical and statistical criteria. The final version of the COA-PR scale, therefore, consisted of 17 items.
3. Conscientious Objection to Abortion Knowledge Scale (Kno-COA), ad hoc. A difficulty was identified with the response format, leading to a change from a 5-point Likert-type scale to a trichotomous nominal response scale (yes, no, or don't know). The number of items (four) and their wording remained unchanged.
4. Support for Individual and Institutional Conscientious Objection (SOC-Ind/Ins): this nonpsychometric scale, composed of 7 items, required wording modifications in 4 items.
5. Intentionality to Promote Abortion Law Compliance Scale (InC-AL). Of its 7 items, 1 was removed due to low association, and another required wording modification.
6. Abortion Attitudes Scale: Criminalization, Acceptance and Entitlement (AA-C/A/E). Of the original 23 items, 3 were removed due to low association (2 from the "Acceptance" dimension and one from the "Entitlement" dimension). In addition, 2 highly similar items within the "Acceptance" dimension showed high interitem correlation and were therefore merged into a single item. Finally, the wording of one item in the "Criminalization" dimension was modified.
7. Abortion Attitudes Scale: Reporting and Confidentiality in Professional Practice (AA-R/C). One item from the "Willingness to Report" dimension was removed due to low association, resulting in a final scale of 20 items. The wording of 2 additional items was modified.
8. Abortion Attitudes Scale: Best Friend Support. In this 7-item scale, only 2 items required wording modifications.
9. School Sex Education and University Training in SRH: this 6-item nonpsychometric scale did not undergo any modifications.

The faculty semistructured interview guide was designed to address the same topics as the student survey, which is central to a convergent mixed methods design. The first part of the interview guide delves into three key themes: (1) delivery of SRH education, with a focus on abortion; (2) ethical dilemmas associated with the ongoing criminalization of abortion beyond the permitted grounds; and (3) perceptions on the prevalence of abortion in Chile and on CO issues, while the second part involves a close-ended sociodemographic questionnaire designed to gather information on age, gender, level of education, political affinity, religious affiliation, years teaching SRH, and type of university where employed, among others.

The interview guide was tested through 5 pilot interviews involving 3 men and 2 women; 3 midwives and 2 physicians; 3 participants from secular universities and 2 from religious universities; and participants from different geographical areas (2 from the Central area, 1 from the Metropolitan area, and 1 from the Southern area). Testing of the interview guide resulted in the addition of one question regarding faculty members' own experiences as students in relation to the SRH training they received, particularly concerning abortion, as well as improvements to question wording. The final semistructured interview guide is provided in Multimedia Appendix 1.

Data Analysis

Two levels of analysis are considered for the students' survey data. First, a large part of the results will be approached as theoretical constructs, as per classic test theory. This will help develop psychometric instruments with potential for replicability, comparison and guaranteed reliability and validity. This will also help categorize and differentiate participants in the resulting theoretical constructs. Results will then be approached item by item in order to discern the scope of observed variables in detail.

Analysis of the psychometric properties of designed scales involves several steps: (1) analysis of the discriminative ability of items: skewness, kurtosis, item-total correlation, and bivariate correlation will be checked; (2) study of the factorial structure by means of cross-validation, applying exploratory and confirmatory factor analysis in independent subsamples; (3) analysis of the internal consistency of participant scores. McDonald omega coefficient or KR20 will be used as appropriate; and (4) analysis of measurement validity based on links to other variables. Assuming parametric assumptions are met, Pearson correlations and mean differences between 2 or more groups (Student *t* test or one-factor ANOVA) will be used. For analysis of results as theoretical constructs or independent observed variables, various strategies will be used: (1) Descriptive statistics of central tendency and frequency; (2) Bivariate statistics, such as Pearson correlations; (3) Means comparison between 2 or more groups (student *t* test or one-factor ANOVA or MANOVA); and (4) Multivariate analysis techniques intended to understand the relative importance of various variables to predicting or explaining the phenomena under review, for example, multivariate multiple and binary logistic regression, factor ANOVA and multivariate analysis of variance, or structural equation modelling [48-51].

Qualitative coding is going to be carried out through "structural coding," which consists of using the questions from the interview guide to create a priori codes [49]. As stated by

XSL•FO
RenderX

Saldaña, structural coding is particularly appropriate for studies involving multiple participants and semi-structured interviews. Structuring coding allows large fragments of the interviews to be grouped into a small number of broad codes, which identify the main topics addressed in the interviews. In this way, it helps to initially categorize the data corpus, identifying similarities, differences and relationships among different interview transcripts. Thereafter, similarly coded segments -that have been coded together- are further analyzed and recorded, producing a more detailed coding scheme for each main topic [49]. Five structural codes were defined a priori: (1) teaching SRH and abortion (approach based on institutional values, teaching methodologies, academic freedom, and hidden curriculum), (2) knowledge of legal abortion protocols, (3) medical ethics, patient confidentiality and reporting of illegal abortions, (4) conscientious objector status, and (5) expanding abortion grounds and the right to choose. In turn, these predefined codes will be revised, modified, and expanded, involving inductive analysis too. ATLAS.ti v. 24.2.0 (Lumivero) qualitative data analysis software will be used to ensure systematic, consistent, and transparent analysis. Codes will also be analyzed per interviewee characteristics in order to identify differences according to variables such as type of university (faith-based or secular); field of endeavor (medicine or midwifery); gender, teaching experience, and geographical area, among others.

Drawing on the pragmatic approach endorsed by this research [43,52] and the established qualitative quality principles, the study will address credibility through systematic and iterative analysis, dependability through transparent documentation of analytic decisions, and reflexivity through critical examination of researchers' assumptions. Likewise, transferability becomes a question about how the findings from a specific context might be meaningful to another. These criteria will be applied pragmatically, with an emphasis on ensuring that the qualitative findings are robust, interpretable, and meaningfully integrated with the quantitative results.

Ethical Considerations

The research protocol was reviewed and approved by the research ethics committee of Diego Portales University (No 022-2022). Pursuant to Laws No 20.120 [53] and 19.628 [54], and article 10 of the Decree 114 of the Ministry of Health (November 19, 2011) [55], for a multicenter study, the ethics approval is required by a single accredited ethics committee on scientific research. The University of Concepción and the Austral University of Chile approved the feasibility of the study in their local communities. The ethics committees of La Frontera University, the Faculty of Medicine of the University of Valparaíso, the Faculty of Medicine of the University of Chile, and the University of Atacama waived approval, given the ethics authorization by the Diego Portales University research ethics committee.

All faculty participants signed an informed consent. For students, the informed consent was displayed electronically, and they had to consent before accessing the online survey. Informed consent forms provided information on research objectives and methods (questionnaires or interviews, as applicable), confidentiality of personal data, the voluntary nature of participation, the right to withdraw at any time, and the contact information of the investigator responsible for answering queries. During recruitment, analysis and results dissemination, all personal data will be treated as confidential. Codes have been used to protect the interviewee's identity. As a token of appreciation for their participation, 70 gift cards equivalent to US $23 were raffled among all students who completed the survey. Multimedia Appendix 2 SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2025 checklist.

Results

The funding for this study was awarded by the Diego Portales University (CG 1110689019) in November 2022. Data collection was completed in July 2025. As of July 2025, a total of 2795 medical and midwifery students were enrolled in the quantitative phase of the study and answered the online survey, and 57 faculty members were recruited for the qualitative stage and were interviewed. We are currently conducting the data analysis. The results are expected to be published from June 2026 onwards.

Discussion

Anticipated Findings

This study represents the first nationwide effort to comprehensively assess the interaction between institutional ideologies and SRH training in Chile following the 2017 legal reform. We anticipate that our findings will reveal a significant gap between the competencies required by Law 21.030 and the actual training provided by medical and midwifery schools.

Current literature on abortion education highlights a global deficit in medical training [17-22,25]. Shifting political contexts can lead to improved abortion training, or conversely restrict it, as appears to be occurring in the United States in the Supreme Court's decision in Dobbs [23,24]. Studies in various countries reveal that even where there are liberal abortion laws, medical curricula often lack practical training and fail to address the social aspects of SRH in the UK [17] and Germany [18]. In addition, prior evidence suggests that training gaps may be more pronounced in faith-based or otherwise institutionally restrictive settings [56,57]. Our study will contribute to this body of work by examining a context of "partial decriminalization," a scenario common in Latin America but less studied in terms of its educational impact. The results should also help provide an empirical basis for developing more effective teaching strategies adapted to the socio-cultural realities of students, thus contributing to educational policies that promote up-to-date, sound, and more inclusive SRH education.

A major strength of this study is its mixed methods design, which allows integration between quantitative (survey) and qualitative data (interviews). The inclusion of both medicine and midwifery students is key, as midwives in Chile play a central role in women's health, providing most of the health care throughout their life course and often holding more progressive views on SRH than physicians. Additionally, the national scope will allow us to enrich the analysis by

XSL•FO
RenderX

understanding the diverse realities of teaching across the country and reducing the bias of concentrating solely on the capital city.

However, our study has limitations given the still-controversial nature of abortion in Chilean society. The primary challenge is the potential self-selection bias; students and faculty with strong pro- or antiabortion views may be more willing to participate than those with neutral positions. Furthermore, recruitment of faculty from conservative faith-based institutions was a significant challenge due to fear of institutional reprisal or stigma. To mitigate this, we implemented strict anonymity protocols and used a recruiting strategy through reliable professional networks to build trust with possible participants. Finally, the cross-sectional nature of this study captures a specific moment in time, and longitudinal studies would be needed to assess how educational interventions could change attitudes over the long term.

The evidence generated by this study will provide guidance to faculty members seeking to address controversial issues in the classroom and will help to define minimum curricular standards for SRH training in Chile. The study should also help provide an empirical basis for developing more effective teaching strategies adapted to the sociocultural realities of students, thus contributing to educational policies that promote up-to-date and more inclusive SRH education. If our hypothesis regarding the training gap is confirmed, the results will support the need for the Ministry of Health and the Ministry of Education to collaborate on standardized guidelines that ensure all health graduates possess the competencies to comply with the law, regardless of the university where they obtained their degree. Future studies should explore the effectiveness of specific educational interventions in reducing abortion stigma among students.

To maximize impact, findings will be disseminated through peer-reviewed publications in journals focused on medical education, public health, and reproductive rights. We plan to present results at national medical and midwifery conferences to directly reach educators. Additionally, policy briefs will be developed for university deans and government officials, highlighting recommendations to improve SRH curricula.

Conclusions

This research protocol addresses a critical intersection between law, education, and public health. By evaluating how future physicians and midwives are taught about abortion in a changing legal frame, the study will surpass ideological debates to focus on professional competency and promote health care equality for women and pregnant people. The findings will provide a route for educators and policymakers to ensure that the next generation of Chilean health care providers is technically competent and ethically prepared to fulfil the needs of a diverse population.

Funding

This research protocol was funded by the Diego Portales University (CG 1110689019). The study has been funded for a 3-year period, starting in November 2022.

Data Availability

All data generated or analyzed during this study will be included in the published articles and their supplementary information files.

Authors' Contributions

Conceptualization, methodology, validation, investigation, writing – original draft, writing – review and editing, supervision, project administration, funding acquisition: LC-B
Conceptualization, methodology, validation, investigation, data curation, writing – original draft, writing – review and editing, supervision: AR
Methodology, validation, investigation, data curation, writing original draft, writing – review and editing: BP-S
Methodology, validation, investigation, writing – original draft, writing – review and editing, supervision: HK
Methodology, validation, investigation, writing – original draft, writing – review and editing, supervision: MC-P
Methodology, validation, investigation, writing – review and editing: AM-V
Methodology, validation, investigation, data curation, writing – review and editing: CB-I
Methodology, validation, investigation, writing – review and editing: HC-A
Investigation, writing – review and editing: CL-L
Investigation, writing – review and editing: LR-T

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 17 KB - resprot_v15i1e81427_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
SPIRIT 2025 checklist.
[DOCX File , 23 KB - resprot_v15i1e81427_app2.docx]

References

1. Casas-Becerra L. Between the secular and the religious: the role of academia in the abortion debate in Chile. In: Abortion and Democracy Contentious Body Politics in Argentina, Chile, and Uruguay. New York: Routledge; 2021:207-225.
2. Regula la despenalización de la interrupción voluntaria del embarazo en tres causales [Web page in Spanish]. Biblioteca del Congreso Nacional de Chile – Ley Chile. 2017 Sep 23. URL: https://www.bcn.cl/leychile/navegar?idNorma=1108237 [accessed 2026-02-14]
3. Montero A, Villarroel R. A critical review of conscientious objection and decriminalisation of abortion in Chile. J Med Ethics 2018;44(4):279-283. [doi: 10.1136/medethics-2017-104281] [Medline: 29306873]
4. Montero VA, Vergara VJ, Ríos HM, Villarroel SR. La objeción de conciencia en el debate sobre la despenalización del aborto por tres causales en Chile [Article in Spanish]. Rev chil obstet ginecol 2017;82(4):350-360 [FREE Full text] [doi: 10.4067/s0717-75262017000400350]
5. Sentencia Rol Nº 3729 (3751-17). August 28, 2017. [Web page in Spanish]. Tribunal Constitucional de Chile. 2017 Aug 20. URL: https://clacaidigital.info/handle/123456789/1772 [accessed 2026-02-20]
6. Universidad Diego Portales. Encuesta Feedback-UDP. Feedback Research. 2024. URL: https://www.feedbackresearch.cl [accessed 2026-02-14]
7. Encuesta CEP N° 92, Agosto-Septiembre 2024 [Web page in Spanish]. Centro de Estudios Públicos. 2014. URL: https://www.cepchile.cl/encuesta/encuesta-cep-n-92/ [accessed 2026-02-14]
8. Bernasconi A. Does the affiliation of universities to external organizations foster diversity in private higher education? Chile in comparative perspective. High Educ 2006;52(2):303-342. [doi: 10.1007/s10734-004-5326-8]
9. MiFuturo.cl – Portal de Información de Educación Superior (Mineduc) [Web page in Spanish]. Ministerio de Educación (Subsecretaría de Educación). 2024. URL: https://www.mifuturo.cl/ [accessed 2026-02-14]
10. Biggs MA, Casas L, Ramm A, Baba CF, Correa SP. Medical and midwifery students' views on the use of conscientious objection in abortion care, following legal reform in Chile: a cross-sectional study. BMC Med Ethics 2020;21(1):42 [FREE Full text] [doi: 10.1186/s12910-020-00484-4] [Medline: 32448300]
11. Casas L, Freedman L, Ramm A, Correa S, Baba CF, Biggs MA. Chilean medical and midwifery faculty's views on conscientious objection for abortion services. Int Perspect Sex Reprod Health 2020;46(Suppl 1):25-34. [doi: 10.1363/46e0620] [Medline: 33326397]
12. Biggs MA, Casas L, Ramm A, Baba CF, Correa SV, Grossman D. Future health providers' willingness to provide abortion services following decriminalisation of abortion in Chile: a cross-sectional survey. BMJ Open 2019;9(10):e030797 [FREE Full text] [doi: 10.1136/bmjopen-2019-030797] [Medline: 31666266]
13. Ramm A, Casas L, Correa S, Baba CF, Biggs MA. "Obviously there is a conflict between confidentiality and what you are required to do by law": Chilean university faculty and student perspectives on reporting unlawful abortions. Soc Sci Med 2020;261:113220. [doi: 10.1016/j.socscimed.2020.113220] [Medline: 32736097]
14. Baba CF, Casas L, Ramm A, Correa S, Biggs MA. Medical and midwifery student attitudes toward moral acceptability and legality of abortion, following decriminalization of abortion in Chile. Sex Reprod Healthc 2020;24:100502. [doi: 10.1016/j.srhc.2020.100502] [Medline: 32234666]
15. Casas L, Correa S, Novoa T, Ramm A. Evaluations of university-level sex education by medical and midwifery students in Chile. In: Beltrán y Puga A, Celorio R, editors. Building Bridges: Contemporary Perspectives on Gender, Sexuality and International Human Rights. Colombia: Universidad del Rosario; 2024:43-64.
16. Guerrero Núñez J, Jofré Olivares D, Guillén-Grima F, Aguinaga-Ontoso I, Bacho-Tapia A, Araya-Moraga L. Alfabetización en salud sexual y reproductiva así como variables asociadas: estudio multicéntrico. Rev Esp Salud Pública 2024;98:e202405035 [FREE Full text]
17. Horan C, Zadeh PG, Rennison C, Hoggart L, Kavanagh J. A qualitative analysis of medical students' attitudes to abortion education in UK medical schools. BMJ Sex Reprod Health 2022;48(3):205-209 [FREE Full text] [doi: 10.1136/bmjsrh-2021-201385] [Medline: 35102002]
18. Rees C, Kaifie A. Evaluation of the reproductive health curriculum at medical schools in Germany: an insight into medical students' knowledge and opinion towards emergency contraception and abortion - a cross-sectional study. BMC Public Health 2025;25(1):3056 [FREE Full text] [doi: 10.1186/s12889-025-24492-4] [Medline: 40993615]
19. Koyama A, Williams R. Abortion in medical school curricula. MJM 2020;8(2) [FREE Full text] [doi: 10.26443/MJM.V8I2.551]
20. Millar E. Abortion stigma, abortion exceptionalism, and medical curricula. Health Sociol Rev 2023;32(3):261-276 [FREE Full text] [doi: 10.1080/14461242.2023.2184272] [Medline: 36916481]
21. Saldanha Lehnen AJ, Rabello M, Martins BG, Oliveira Cardoso FF. O aborto legal en casos de gravidez decorrente de violence sexual: percepcoes e vivencias de médicas e médicos obstetras [Article in Portuguese]]. Cad Saùde Pública 2024;40(5):e00124423 [FREE Full text] [doi: 10.1590/0102-311xpt124423]

XSL•FO
RenderX

22. Ferruzzi Sacchetin L, Pereira de Souza A, Santina Murgo C. Objeção de consciência médica em casos de abortamento [Article in Portuguese]. Scientia Medica 2024;33(1):e45000. [doi: 10.15448/1980-6108.2023.1.45000]
23. French VA, Hou MY. Abortion education for medical students in an era of increased abortion restrictions. Clin Obstet Gynecol 2024;67(3):539-554 [FREE Full text] [doi: 10.1097/GRF.0000000000000874] [Medline: 38813914]
24. Beasley AD, Olatunde A, Cahill EP, Shaw KA. New gaps and urgent needs in graduate medical education and training in abortion. Acad Med 2023;98(4):436-439. [doi: 10.1097/ACM.0000000000005154] [Medline: 36656271]
25. Nicholson M, Hoggart L. Abortion learning mechanisms for nurses and midwives: a scoping review of evidence. Sex Reprod Health Matters 2025;33(1):2518672 [FREE Full text] [doi: 10.1080/26410397.2025.2518672] [Medline: 40576098]
26. Competencias comunes para los egresados de las escuelas de medicina [Web page in Spanish]. Asociación de Facultades de Medicina de Chile (ASOFAMECH). 2024. URL: https://www.asofamech.cl/documentos/ [accessed 2026-02-14]
27. Perfil de conocimientos EUNACOM (versión 2.0, mayo 2010) [Web page in Spanish]. EUNACOM - Examen Único Nacional de Conocimientos de Medicina. 2010. URL: https://www.eunacom.cl/contenidos/PerfilNew.pdf [accessed 2026-02-14]
28. Decreto con Fuerza de Ley Nº725. 1968. URL: https://bcn.cl/24us4 [accessed 2026-02-14]
29. Robledo P, Valdés T, Viera Y. El proceso de despenalización de la interrupción del embarazo en tres causales de Chile. In: Casas L, Maira M, editors. Aborto en Tres Causales en Chile. Lectura del Proceso de Despenalización. Santiago: Universidad Diego Portales, Law School and Human Rights Centre; 2019:27-47.
30. Casas L, Vivaldi L, Montero A, Bozo N, Álvarez JJ, Babul J. Primary care and abortion legislation in Chile: a failed point of entry. Dev World Bioeth 2023;23(2):154-165. [doi: 10.1111/dewb.12377] [Medline: 36206359]
31. Casas L. La legalización del aborto en tres causales en Chile y la reducción del estigma. In: Gauché X, Barría M, editors. Claves Interdisciplinarias Para el Acceso a La Justicia con Enfoque de Género y Diversidad Sexual. Santiago: Thomson Reuters; 2023.
32. Montero A, Ramírez-Pereira M, Robledo P, Casas L, Vivaldi L, González D. Main barriers to services linked to voluntary pregnancy termination on three grounds in Chile. Front Public Health 2023;11:1164049 [FREE Full text] [doi: 10.3389/fpubh.2023.1164049] [Medline: 37457269]
33. VI informe sobre los derechos sexuales y reproductivos en Chile [Report in Spanish]. Miles. 2024. URL: https://mileschile.cl/wp-content/uploads/2024/07/VI-Informe-1.pdf [accessed 2026-02-14]
34. Montero A, Ramírez-Pereira M, Robledo P, Casas L, Vivaldi L, Gonzalez D. Conscientious objection as structural violence in the voluntary termination of pregnancy in Chile. Front Psychol 2022;13:1007025 [FREE Full text] [doi: 10.3389/fpsyg.2022.1007025] [Medline: 36405205]
35. Informe objeción de conciencia en Chile: desafíos en la implementación de la ley sobre interrupción voluntaria del embarazo en tres causales a seis años de su dictación [Web page in Spanish]. Corporación Humanas. 2023. URL: https://tinyurl.com/3peuffac [accessed 2026-02-14]
36. Suárez-Baquero DFM, Dzuba IG, Romero M, Baba CF, Biggs MA. Failure of the law to grant access to legal abortion in Chile. Health Equity 2024;8(1):189-197 [FREE Full text] [doi: 10.1089/heq.2023.0050] [Medline: 38559845]
37. Alveal-Álamos C, Pérez B, Obando A, Carte L, Jara L. La objeción de conciencia frente a la interrupción voluntaria del embarazo: motivaciones que traspasan las creencias morales y religiosas en profesionales de la salud chilenos [Article in Spanish]. Rev Punto Género 2022;17:307-344. [doi: 10.5354/2735-7473.2022.67663]
38. Creswell JW. A Concise Introduction to Mixed Methods Research. Thousand Oaks, CA: SAGE Publications; 2022.
39. Biesta G. Pragmatism and the philosophical foundations of mixed methods research. In: Handbook of Mixed Methods in Social & Behavioral Research. Thousand Oaks, CA: SAGE Publications; 2015:95-118.
40. Johnson RB, Onwuegbuzie AJ, Turner LA. Toward a definition of mixed methods research. J Mix Methods Res 2007;1(2):112-133. [doi: 10.1177/1558689806298224]
41. Tashakkori A, Johnson RB, Teddlie C. Foundations of Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences. Thousand Oaks, CA: SAGE Publications; 2021.
42. Baert P. Philosophy of the Social Sciences: Towards Pragmatism. Cambridge: Polity Press; 2005.
43. Morgan DL. Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. J Mix Methods Res 2007;1(1):48-76 [FREE Full text] [doi: 10.1177/2345678906292462]
44. Consejo NDE. Índices de Matrícula. Base Índices Matrícula - 2022 (Excel version) Internet. 2022. URL: https://www.cned.cl/bases-de-datos [accessed 2026-02-20]
45. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
46. García-García JA, Reding-Bernal A, López-Alvarenga JC. Cálculo del tamaño de la muestra en investigación en educación médica [Article in Spanish]. Investigación en educación médica 2013;2(8):217-224 [FREE Full text] [doi: 10.1016/s2007-5057(13)72715-7]
47. Johanson GA, Brooks GP. Initial scale development: sample size for pilot studies. Educational and Psychological Measurement 2009;70(3):394-400 [FREE Full text] [doi: 10.1177/0013164409355692]
48. Pérez B, Jara Sepúlveda L, Saiz JL, Alveal-Álamos C. Influence of woman's circumstances and ideology on abortion acceptance. Curr Psychol 2025;44(10):9473-9486 [FREE Full text] [doi: 10.1007/s12144-025-07709-y]

XSL•FO
RenderX

49. Saldaña J. The Coding Manual for Qualitative Researchers. London: Sage Publications, Inc; 2016.
50. Elosua P, Zumbo B. Coeficientes de fiabilidad para escalas de respuesta categórica ordenada [Article in Spanish]. Psicothema 2008;20(4):896-901 [FREE Full text]
51. Martínez Arias MR, Hernández Lloreda MV, Hernández Lloreda MJ. Psicometría. Madrid: Alianza Editorial; 2014.
52. Morgan DL. Pragmatism as a paradigm for social research. Qualitative Inquiry 2014;20(8):1045-1053 [FREE Full text] [doi: 10.1177/1077800413513733]
53. Ley 20120 Sobre la investigación científica en el ser humano, su genoma, y prohíbe la clonación humana [Web page in Spanish]. Biblioteca del Congreso Nacional de Chile – Ley Chile. 2006 Sep 22. URL: https://www.bcn.cl/leychile/navegar?idNorma=253478 [accessed 2026-02-14]
54. Ley N° 19.628, sobre protección de la vida privada [Web page in Spanish]. Biblioteca de Regulación – Gobierno Digital Chile. 1999 Aug 28. URL: https://digital.gob.cl/biblioteca/regulacion/ley-n-19628-sobre-proteccion-de-la-vida-privada/ [accessed 2026-02-14]
55. Decreto 1114 Aprueba reglamento de la ley Nº 20.120, sobre la investigación científica en el ser humano, su genoma, y prohíbe la clonación humana [Web page in Spanish]. Biblioteca del Congreso Nacional de Chile – Ley Chile. 2011 Nov 19. URL: https://www.bcn.cl/leychile/navegar?idNorma=1032919 [accessed 2026-02-14]
56. Guiahi M, Teal S, Kenton K, DeCesare J, Steinauer J. Family planning training at Catholic and other religious hospitals: a national survey. Am J Obstet Gynecol 2020;222(3):273.e1-273.e9. [doi: 10.1016/j.ajog.2019.09.012] [Medline: 31526788]
57. Feltman RN, Lewis SR, Thompson NE. Family planning competency following medical school Ob/Gyn clerkships at faith-based and secular sites. Sci Rep 2024 Feb 14;14(1):3667 [FREE Full text] [doi: 10.1038/s41598-024-54304-5] [Medline: 38351035]

Abbreviations

AA-C/A/E: Abortion Attitudes Scale: Criminalization, Acceptance and Entitlement
AA-R/C: Abortion Attitudes Scale: Reporting and Confidentiality in Professional Practice
CPolA: Scale on the Centrality of Political Affinity in Self-Concept
CO: conscientious objection
COA-PR: Scale on Conscientious Objection to Abortion: Positioning and Reasons
InC-AL: Intentionality to Promote Abortion Law Compliance Scale
Kno-COA: Conscientious Objection to Abortion Knowledge Scale
SOC-Ind/Ins: Support for Individual and Institutional Conscientious Objection
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SRH: sexual and reproductive health

Edited by A Schwartz; submitted 28.Jul.2025; peer-reviewed by AA Villa-Rueda, H Pratomo; comments to author 23.Dec.2025; revised version received 03.Feb.2026; accepted 11.Feb.2026; published 06.Mar.2026.

Please cite as:
Casas-Becerra L, Ramm A, Pérez B, Kaune H, Carrasco-Portiño M, Montero A, Bustos-Ibarra C, Castro-Arias H, López-Labarca C, Reyes-Torres L
Perceptions of Abortion and Sexual and Reproductive Health in Chilean Medical and Midwifery Education: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e81427
URL: https://www.researchprotocols.org/2026/1/e81427
doi:10.2196/81427
PMID:

©Lidia Casas-Becerra, Alejandra Ramm, Beatriz Pérez, Heidy Kaune, Mercedes Carrasco-Portiño, Adela Montero, Cecilia Bustos-Ibarra, Henry Castro-Arias, Claudio López-Labarca, Leonardo Reyes-Torres. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 06.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Real-World Effectiveness and Safety of Liuwei Dihuang Pill for Menopausal Syndrome: Protocol for a Prospective, Observational, Multicenter Cohort Study

Shuoshuo Wei[1], MM; Hongyan Zhang[1], MMSc; Mengmeng Wang[1], MMSc; Yuhui Sun[2], MD; Tao Song[2], MD; Weifang Zheng[3], MBBS; Yunmei Ye[3], MBBS; Liran Ma[4], MM; Yunshuang Yang[5], PhD; Tong Wei[6], MM; Lan Cheng[7], MM; Chunquan Sun[8], MD; Xingting Ding[9], MM; Shichu Zhao[7*], PhD; Qingqiao Song[10*], MD; Lianxin Wang[1*], PhD

[1]Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
[2]First Affiliated Hospital of Harbin Medical University, Heilongjiang, China
[3]Lanxi Hospital of Traditional Chinese Medicine, Zhejiang, China
[4]Beijing Gulou Hospital of Traditional Chinese Medicine, Beijing, China
[5]Beijing Longfu Hospital, Beijing, China
[6]Dongcheng District, Hepingli Community Health Service Center, Beijing, China
[7]Beijing Puren Hospital, Beijing, China
[8]Tsinghua University Yuquan Hospital (Tsinghua University Hospital of Integrative Traditional Chinese and Western Medicine), Beijing, China
[9]Dongcheng District, Chaoyangmen Community Health Service Center, Beijing, China
[10]Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Lianxin Wang, PhD
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences
No. 16, Dongzhimen South Street, Dongcheng District
Beijing, 100700
China
Phone: 86 13521781839
Email: wlxing@126.com

Related Article:

This is a corrected version. See correction statement: https://www.researchprotocols.org/2026/1/e84867

Abstract

Background: Menopausal syndrome (MPS) results from declining ovarian function and estrogen fluctuations during the menopausal transition, typically presenting with vasomotor, psychological, and metabolic symptoms that impair quality of life. Current treatments, including hormone replacement therapy and nonhormonal medications, are limited by safety concerns and adverse effects. The Liuwei Dihuang (LWDH) pill, a classical traditional Chinese medicine formula, is widely used for kidney-yin deficiency and menopausal symptoms such as hot flashes, insomnia, depression, and anxiety. Preliminary studies suggest favorable efficacy and safety; however, robust real-world evidence, particularly in combination with conventional Western therapy, remains insufficient.

Objective: This study aims to evaluate the real-world effectiveness and safety of LWDH in combination with conventional Western therapy for MPS.

Methods: This prospective, multicenter, real-world, observational cohort study will be conducted at 8 centers in China from March 9, 2025, to December 2026. A total of 1000 patients with MPS in a real-time clinical setting will be allocated to either the exposure group (conventional therapy plus LWDH) or the nonexposure group (conventional therapy alone). The primary outcome is the change in the modified Kupperman Index. The secondary outcomes encompass changes in the Menopausal Quality of Life Scale, serum sex hormone levels, traditional Chinese medicine syndrome scores, and metabolic parameters. Metabolomics analyses

XSL•FO
RenderX

will be conducted to explore potential mechanisms. Safety will be assessed by monitoring adverse events and clinically significant abnormalities in vital signs or laboratory parameters. Assessments will be conducted at baseline and at 2 and 4 weeks, with a subsequent 2-week follow-up. Propensity score–based inverse probability of treatment weighting will be applied to balance baseline covariates, and weighted regression models will be used for comparative analyses.

Results: Participant recruitment and enrollment began in March 2025. It is expected that by December 2026, a total of 1000 participants will be recruited. At present, patient recruitment is ongoing, and preliminary organization and analysis of medical record data are underway. Blood sample collection for metabolomics analysis is also in progress. Final data collection is expected to be completed in March 2027. We expect to generate evidence to support clinical decision-making, with results anticipated to be published in esteemed, peer-reviewed health research journals by December 2027.

Conclusions: This study will assess the clinical effectiveness, safety, and potential mechanisms of LWDH in the treatment of MPS in real-world settings, providing a scientific basis for its use in clinical practice.

Trial Registration: ClinicalTrials.gov NCT06874738; https://clinicaltrials.gov/study/NCT06874738

International Registered Report Identifier (IRRID): DERR1-10.2196/84803

(JMIR Res Protoc 2026;15:e84803) doi:10.2196/84803

KEYWORDS

Liuwei Dihuang pill; real world; menopausal syndrome; protocol; effectiveness; safety

Introduction

Menopausal syndrome (MPS) arises from declining ovarian function and estrogen fluctuations during the perimenopausal transition, primarily manifesting as vasomotor symptoms (hot flashes and night sweats), psychological disturbances (anxiety and depression), and metabolic disorders (dysglycemia and dyslipidemia) [1,2]. It is estimated that by 2030, more than 1.2 billion women worldwide will be in the menopausal or postmenopausal stage, with up to 85% experiencing menopausal symptoms to varying degrees [3,4]. Hormonal alterations, particularly involving the hypothalamic-pituitary-ovarian axis, may furthermore increase the risk of glycolipid metabolic disorders, cardiovascular diseases, osteoporosis, and neurocognitive disorders, underscoring the substantial public health burden of MPS [5-7].

Current management primarily includes hormone therapy and nonhormonal pharmacological agents. Hormone therapy mitigates symptoms resulting from diminished estrogen levels by providing estrogen and progesterone; however, its prolonged use correlates with an elevated risk of breast cancer and adverse cardiovascular events [8,9]. Antidepressants, gabapentin, and clonidine are nonhormonal medications that can ameliorate symptoms such as hot flashes and mood fluctuations, but they may also cause side effects such as nausea, insomnia, and dry mouth [10,11]. In addition, owing to individual variability among patients, symptoms may recur after the discontinuation of these treatments. Therefore, safer and more sustainable therapeutic strategies are still required.

Traditional Chinese medicine (TCM) has a long history and rich experience in treating MPS, with its advantages lying in the emphasis on holistic management, which offers it unique benefits and broad prospects for application in this domain. In ancient China, MPS was viewed as a category of “depression syndrome, insomnia, and fatigue” based on its characteristic symptoms and signs. The principal causes of this condition were thought to be deficiencies in the liver and kidneys [12]. Treatment methods included acupuncture, *tuina* (Chinese therapeutic massage), moxibustion, dietary therapy, and herbal medicine, all aimed at nourishing and tonifying these organs [13,14]. According to literature survey, herbal medicines can successfully alleviate symptoms and enhance quality of life, exhibit few side effects, and show high patient compliance, rendering it suitable for long-term use [15,16]. It was found that many formulas used in the treatment of MPS due to kidney-yin deficiency are based on the foundational recipe of Liuwei Dihuang (LWDH), which shows significant therapeutic effects [17,18].

LWDH originates from the medical text “Xiao Er Yao Zheng Zhi Jue,” written by the renowned physician Qian Yi during the Song Dynasty. This formula is composed of 6 traditional Chinese herbs: *Rehmannia glutinosa* (Shudihuang), *Cornus officinalis* (Shanzhuyu), *Dioscorea opposita* (Shanyao), Poria (Fuling), *Alisma orientale* (Zexie), and Moutan Cortex (Mudanpi). Each ingredient contributes to its therapeutic effects, targeting nourishing yin and tonifying the kidneys [18]. Modern pharmacological studies indicate that LWDH exhibits a range of effects, including antioxidant, anti-inflammatory, and antitumor properties, along with antihyperglycemic and lipid-lowering effects [19-21]. Although prior clinical trials have suggested potential effectiveness of LWDH for MPS [22,23], these studies were conducted under highly controlled conditions with stringent inclusion and exclusion criteria, limited treatment flexibility, and relatively homogeneous patient populations, which may constrain external validity and generalizability to routine clinical practice. In real-world settings, treatment decisions for MPS—especially regarding TCM—are influenced by patient preferences, symptom burden, prior treatment history, and shared decision-making between clinicians and patients. Imposing random allocation in this context may reduce patient willingness to participate and interfere with authentic prescribing behaviors, thereby limiting the study’s ability to capture real-world treatment patterns. Accordingly, this multicenter real-world study will enroll 1000 patients with MPS to evaluate the effectiveness and safety of LWDH in routine clinical settings, aiming to provide pragmatic evidence for its integration into MPS management.

XSL•FO
RenderX

Methods

Research Design

This is a multicenter, prospective, real-world, nonrandomized controlled study (registered at ClinicalTrials.gov; NCT06874738) designed to evaluate the effectiveness and safety of LWDH in patients with MPS associated with kidney-yin deficiency.

A total of 1000 patients with MPS associated with kidney-yin deficiency will be recruited from 8 centers across China and divided into 2 cohorts: exposure group (patients receiving conventional Western medical therapy plus LWDH) and nonexposure group (patients receiving conventional Western medical therapy alone). Physicians will determine whether to prescribe LWDH based on clinical assessment, guideline recommendations, and patient preference. Assessments will be conducted at baseline and at 2, 4, and 6 weeks after enrollment, with the aim of systematically evaluating the real-world effectiveness of LWDH and further exploring its potential mechanisms of action.

The study protocol adheres to the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) statement [24]. The completed SPIRIT checklist is provided as Multimedia Appendix 1. The study flowchart and key elements of the protocol are presented in Figures 1 and 2.

Figure 1. Flow diagram of the study. EKG: electrocardiography; TCM: traditional Chinese medicine.

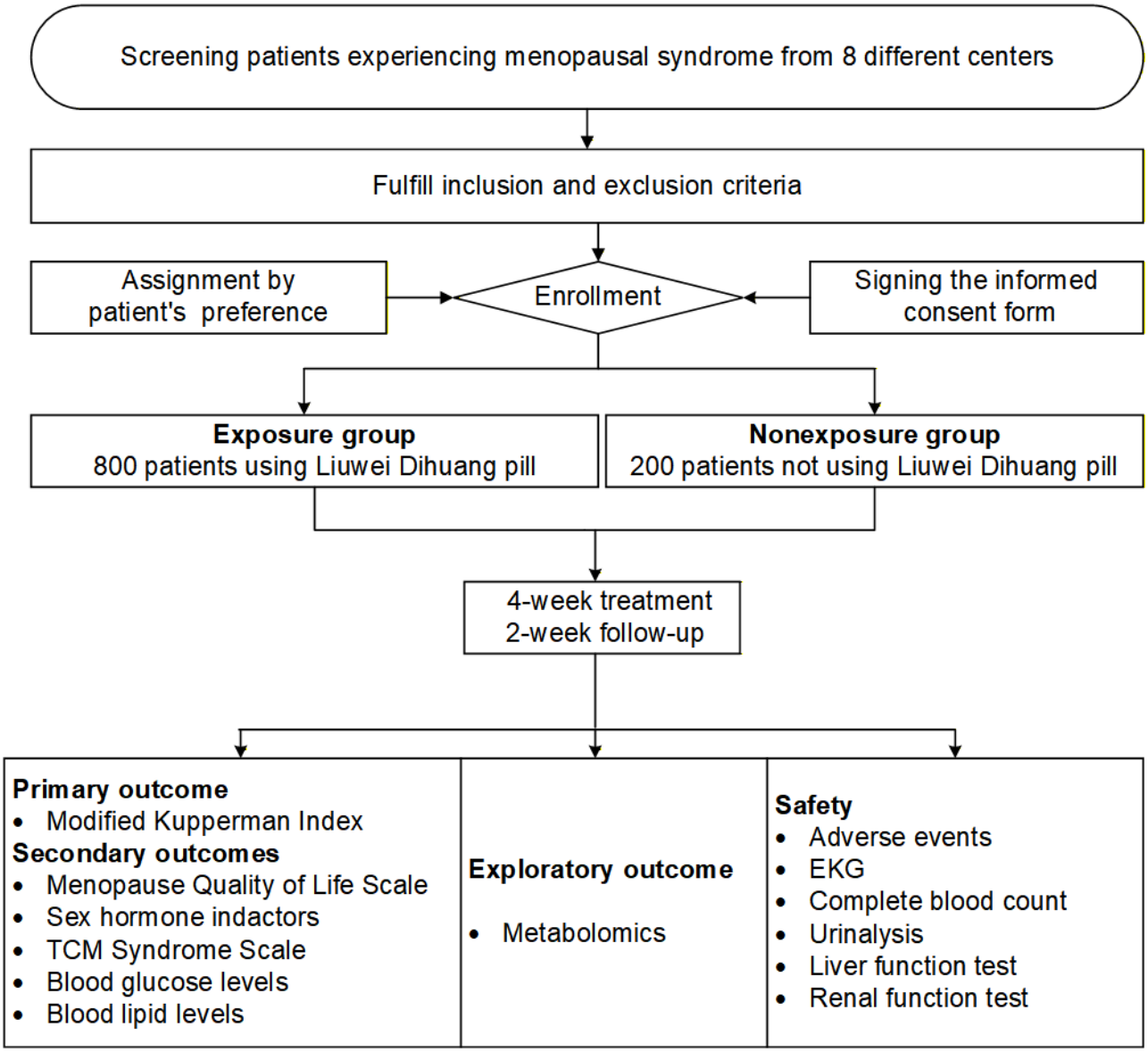

XSL•FO
RenderX

Figure 2. Summary of the protocol.

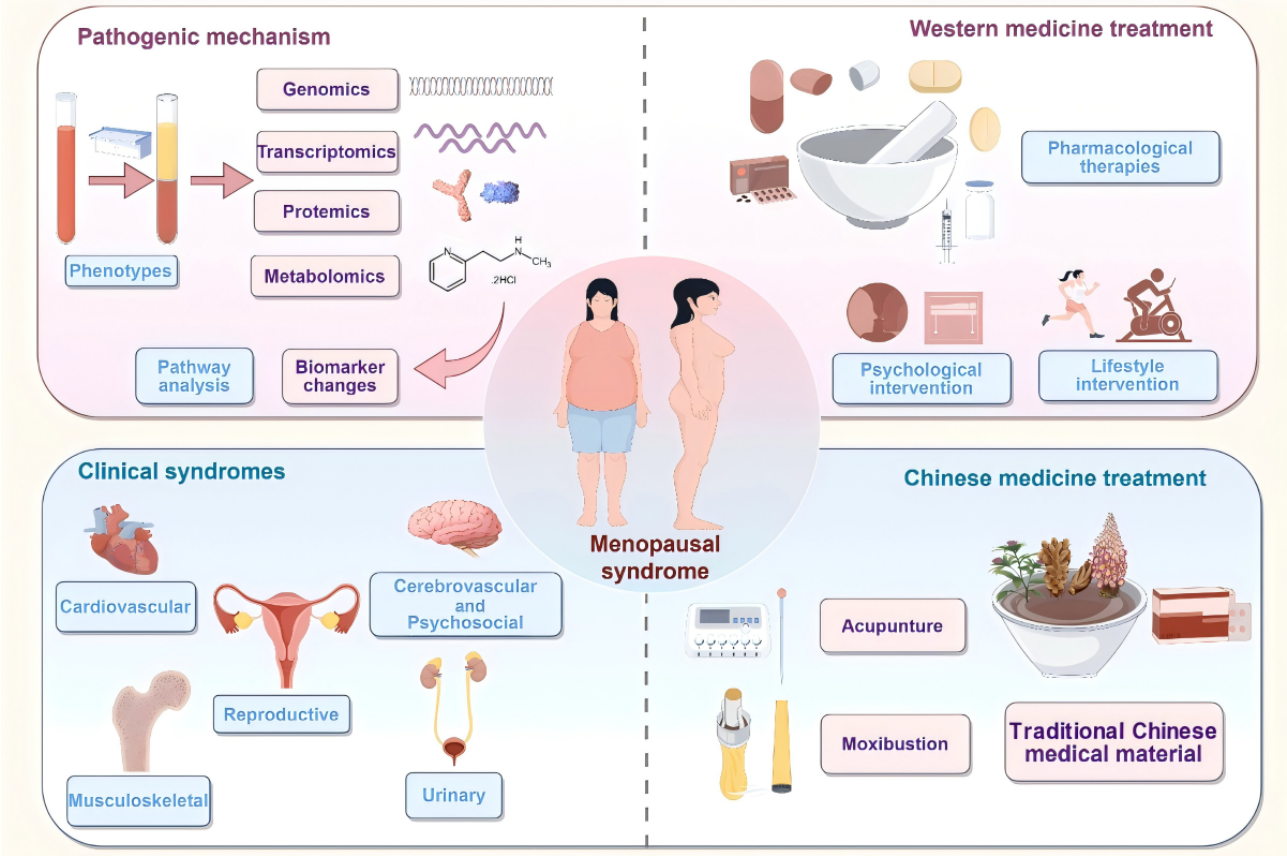

Allocation Strategy and Blinding of Outcome Assessment

As this is a real-world observational study, treatment allocation is determined by routine clinical decision-making and patient preference; therefore, randomization and participant blinding are not applicable.

To minimize potential analytical bias, statisticians and outcome assessors involved in data analysis will remain blinded to group identification during statistical evaluation.

Sample Size Calculation

The primary objective of this trial is to evaluate the therapeutic effectiveness of LWDH in patients with MPS. Sample size estimation was based on the assumption that the exposure group would demonstrate superior improvement in the modified Kupperman Index compared with the nonexposure group. On the basis of prior clinical data [25], we assumed that the mean difference in the change of the modified Kupperman Index between the 2 groups (δ) would be 0.6, with a common SD (σ) of 2.

The sample size calculation was performed using a 2-sample independent *t* test, with a 2-sided significance level (α) of .05 and a statistical power (1−β) of 0.90. Participants in the exposure and nonexposure groups will be included in a 4:1 ratio (K) [26]. The required sample size per group was calculated using the following formula for a superiority trial:



On the basis of these assumptions, the estimated sample size was 640 participants in the exposure group and 160 participants in the nonexposure group. Considering an anticipated dropout rate of 20%, the final sample size was adjusted accordingly. Therefore, the planned sample sizes were set at 800 participants in the exposure group and 200 participants in the nonexposure group, yielding a total sample size of 1000 participants.

Study Participants

This study focuses on MPS in patients diagnosed with a TCM pattern of kidney-yin deficiency. These patients receive treatment in the outpatient departments of 8 institutions: Beijing Longfu Hospital, Beijing Gulou Traditional Chinese Medicine Hospital, Lanxi Traditional Chinese Medicine Hospital, The First Affiliated Hospital of Harbin Medical University, Beijing Dongcheng District Hepingli Community Health Service Center, Beijing Dongcheng District Chaoyangmen Community Health Service Center, Beijing Puren Hospital, and Tsinghua University Yuquan Hospital (Tsinghua University Hospital of Integrative Traditional Chinese and Western Medicine).

Inclusion Criteria

Participants are deemed to be in perimenopause according to Stages of Reproductive Aging Workshop classification [27] and meet all the following criteria [28-30]: (1) women aged 45 to 55 years (inclusive), (2) diagnosis of MPS according to Western medical diagnostic criteria, (3) presence of menstrual disturbances or amenorrhea lasting more than 3 months, (4) diagnosis of kidney-yin deficiency syndrome according to TCM syndrome differentiation criteria, (5) a modified Kupperman Index score of 6 or higher, (6) serum follicle-stimulating

XSL•FO
RenderX

hormone level higher than 10 IU/L, and (7) voluntary provision of written informed consent.

Exclusion Criteria

Participants meeting any of the following criteria will be excluded from the trial: (1) those with known allergies to the ingredients of LWDH or contraindications to Chinese herbal preparations; (2) women who are pregnant, lactating, or intending to conceive; (3) patients with severe primary disorders of the liver, kidney, and hematopoietic system; and (4) patients with comorbidities that significantly impact assessment, including severe cognitive impairment and aphasia.

Treatments

Eligible individuals who provide informed consent will be assigned to either the exposure group (conventional Western therapy plus LWDH) or the nonexposure group (conventional Western therapy alone) based on their preferences and whether they are taking LWDH in real clinical practice. All participants will receive guideline-based conventional Western medical therapy tailored to their clinical condition [13,31,32]. In addition, patients in the exposure group will receive LWDH at a dose of 8 pills (3 g) orally, 3 times daily for 4 weeks.

Concomitant Medications

During the entire monitoring phase, the use of combination therapy and other treatments is permitted, whereas the use of traditional Chinese herbal formulas with yin-nourishing and kidney-tonifying effects is strictly prohibited.

Outcomes

Primary Outcome

The primary outcome is as follows: the modified Kupperman Index will be used to quantify the overall severity of MPS. This index comprises 13 items covering vasomotor, somatic, and psychological symptoms, with weighted scores ranging from 0 to 63; higher scores indicate greater symptom severity.

Secondary Outcomes

The secondary outcomes are as follows:

- The Menopause Quality of Life scale will be assessed as a key patient-reported secondary outcome. The scale includes 29 items across 4 domains (ie, vasomotor, psychosocial, physical, and sexual), with lower scores indicating better quality of life.
- The 6 sex hormone indicators will be used to characterize endocrine levels during the perimenopausal period, including follicle-stimulating hormone, luteinizing hormone, estradiol, progesterone, testosterone, and prolactin.
- TCM syndrome scores for kidney-yin deficiency will be evaluated using a standardized TCM Syndrome Scale. This scale evaluates 3 dimensions: physical sensations, psychological feelings, and urogenital symptoms, with higher scores indicating more severe symptoms.
- Blood glucose levels will be assessed using fasting blood glucose, 2-hour postprandial blood glucose, and glycated albumin to provide insights into the state of glucose metabolism.
- Blood lipid levels will be assessed using total cholesterol, triglycerides, high-density lipoprotein cholesterol, and low-density lipoprotein cholesterol to help evaluate the status of lipid metabolism.

Exploratory Outcomes

Metabolomics analysis will be conducted as an exploratory secondary outcome, focusing on the preliminary exploration of molecular signals potentially related to treatment response and metabolic changes.

All efficacy outcomes will be assessed at baseline and at 4 weeks after treatment initiation.

Safety Outcomes

Safety assessments will include monitoring adverse events, vital signs (blood pressure and heart rate), routine laboratory tests (hematology, urinalysis, and hepatic and renal function), and electrocardiography. Adverse events will be recorded throughout the study and followed until resolution or stabilization.

Data Collection and Study Procedures

The data collection and follow-up schedule is presented in Table 1.

XSL•FO
RenderX

Table 1. Time schedule of participant enrollment and follow-up.

	Screening and enrollment	Treatment phase		Follow-up
	Baseline	2 weeks +7 to −7 days	4 weeks +7 to −7 days	6 weeks +7 to −7 days
Enrollment				
Confirm eligibility	✓			
Informed consent	✓			
Demographics	✓			
Medical history-treatment history-medication history	✓			
Treatment				
LWDH[a]		✓	✓	
Assessments				
Vital signs	✓	✓	✓	
Hematological examination	✓	✓	✓	
Questionnaire surveys	✓	✓	✓	
EKG[b]	✓	✓	✓	
Prescription and distribution of LWDH	✓	✓		
Evaluation of medication compliance		✓	✓	
Combined medication	✓	✓	✓	
Adverse events		✓	✓	✓
Changes in medical history, treatment, or medication		✓	✓	✓
Conclusion of completion of the case				✓

[a]LWDH: Liuwei Dihuang.

[b]EKG: electrocardiography.

Screening and Enrollment Phase

Participants will be screened to determine their eligibility based on inclusion and exclusion criteria. Before administering the medication, data will be collected on general demographic information (eg, age, gender, ethnicity, occupation, height, and weight), vital signs (eg, temperature, blood pressure, heart rate, and respiratory rate), medical history, treatment history, and medication history. Patients enrolled in the study will undergo hematological tests (sex hormone test, complete blood count, urinalysis, blood lipid test, blood glucose test, liver function test, and renal function test), electrocardiography, and questionnaire assessments (modified Kupperman Index, Menopausal Quality of Life Scale, TCM Syndrome Scale). The collected information will be recorded in the screening and enrollment section of the case report forms (CRFs).

Treatment Phase

Follow-up visits will occur at 2 weeks (±7 days) and 4 weeks (±7 days) after enrollment. At each visit, vital signs, laboratory parameters, questionnaire assessments, adverse events, medical history, medication history, and treatment history since enrollment will be recorded. Furthermore, information regarding the prescription and distribution of LWDH will also be collected at screening and enrollment and the first follow-up visit. When LWDH is newly prescribed, its use will be assessed at the second follow-up visit and the final follow-up visit. Relevant information will be written in the treatment section of CRFs.

Follow-Up Visit

A telephone follow-up will be conducted at 6 weeks (+7 to −7 days) after enrollment to monitor adverse events and changes in clinical status or medication use. Study completion status will be documented accordingly.

Serum Sample Collection and Preparation

All blood tests will be conducted after an 8-hour fasting period and a minimum of 30 minutes of rest before collecting venous blood. For menopausal participants, blood samples will be taken before treatment and on the last day of treatment, while for menstruating participants, blood samples will be collected on the second day of their menstrual period. Investigators will collect morning fasting elbow venous blood samples (2-3 mL) from the 2 cohorts at 2 time points (before and after treatment). After 40 minutes at room temperature, the samples will be centrifuged at 3000 rpm at 4 °C for 10 minutes. The upper serum will be aliquoted into four 1.5 mL Eppendorf tubes (500 µL per tube), labeled, and stored at –80 °C for metabolomics analysis.

XSL•FO
RenderX

Data Entry and Quality Control

Data Management

The following procedures will be implemented to ensure data management:

- Study data will be initially recorded in CRFs and subsequently entered into a secure electronic data capture (EDC [33]) system. All data entry personnel will receive standardized training before study initiation. Double data entry with cross-verification will be implemented to ensure accuracy.
- The EDC system incorporates automated logic checks to identify missing values, inconsistencies, and outliers. Key variables, including reproductive hormone measurements, will undergo additional verification. Source documents (eg, informed consent forms, laboratory reports, and electrocardiography) will be retained for quality review.
- Data cleaning will include consistency checks, missing data assessment, outlier detection, and duplicate record screening. The final database will be locked after joint review and confirmation by the principal investigator and the statistician.

Quality Control Measures

All study procedures will be conducted in accordance with standardized operating procedures. Quality assurance measures will be implemented throughout the study to ensure data integrity and protocol compliance.

Investigator Training

Investigators at each center will receive protocol-specific training before study initiation. Ongoing updates related to study procedures will be communicated as necessary throughout the trial.

Participant Adherence and Safety Monitoring

Enhancing Participant Adherence

To enhance adherence, participants will receive standardized instructions regarding medication use and study procedures at enrollment. Reminder strategies (eg, telephone calls or text messages) will be implemented to promote follow-up attendance and treatment compliance. Medication use and symptom changes will be recorded at each visit.

Adverse Event Monitoring

Adverse events will be collected through scheduled assessments and spontaneous reports. Investigators will evaluate their severity and potential relationship to study treatment and manage them according to routine clinical practice. Serious adverse events will be reported to the ethics committees in accordance with regulatory requirements.

Statistical Analysis

Statistical Software and Significance Level

Statistical analyses will be performed using SPSS Statistics (version 26.0; IBM Corp) and R software (version 4.5.1; R Foundation for Statistical Computing). All statistical tests will be 2-sided, and a significance level (α) of .05 will be applied.

Descriptive Analysis

Descriptive analyses will be conducted to summarize the demographic and baseline characteristics of the participants. Continuous variables will be expressed as means (SDs) or medians (IQRs), as appropriate, while categorical variables will be presented as frequencies and percentages.

Primary Analysis and Control of Confounding

Given the nonrandomized, real-world design, the primary comparative analysis will be conducted using inverse probability of treatment weighting (IPTW) based on propensity scores to control for potential confounding and selection bias. Propensity scores will be estimated using a multivariable logistic regression model including a priori–selected baseline covariates, such as age, BMI, duration of MPS, comorbidities, and concomitant hormonal therapy and other pharmacological therapies (including the type of hormonal therapy, if applicable). The distribution and overlap of propensity scores between groups will be examined to assess the positivity assumption, and covariate balance after weighting will be assessed using standardized mean differences and other balance diagnostics. Univariable and multivariable analyses will be conducted to explore associations between baseline covariates and outcomes and to inform model specification; however, the primary comparative analyses will be based on IPTW-adjusted models. Covariate balance before and after weighting will be evaluated using standardized mean differences. A standardized mean difference of less than 0.1 will be considered indicative of adequate balance. A covariate balance table summarizing baseline characteristics before and after IPTW will be presented in the final study report to demonstrate the effectiveness of the weighting procedure. In addition, graphical diagnostics will be used to visually assess balance across covariates.

Outcome-Specific Analytical Methods

Analytical methods will be selected according to the type of outcome variables. Continuous outcomes, including scale scores, sex hormone levels, and metabolic indicators, will be analyzed using weighted linear regression models or analysis of covariance, with adjustment for baseline values when applicable. Categorical outcomes will be analyzed using weighted chi-square tests or weighted logistic regression models. Secondary unweighted analyses will be performed for descriptive and exploratory purposes. Normality of continuous variables will be assessed using the Shapiro-Wilk test. For normally distributed variables, between-group comparisons will be conducted using independent-samples 2-tailed *t* tests and within-group comparisons will be performed using paired-samples 2-tailed *t* tests. For nonnormally distributed variables, the Mann-Whitney *U* test will be applied.

Situations and Handling of Missing Values

Data missingness is described as variables explicitly stated as missing or unobtainable, unreported variables (ie, blank observations), reported data that are ambiguous, or data that are inconsistent or beyond acceptable ranges.

Missing data will be handled using multiple imputation under the assumption of missing at random for the primary analysis. As a sensitivity analysis, the last observation carried forward

XSL•FO
RenderX

method will also be applied, and results obtained using different imputation methods will be compared to assess robustness.

Sensitivity Analyses

To further evaluate the robustness of the primary findings, sensitivity analyses will be conducted using alternative propensity score approaches, including propensity score matching and stratification. Additional sensitivity analyses will be performed by excluding participants receiving concomitant hormonal therapy and by restricting analyses to patients with full treatment adherence.

Consideration of Center Effects

Given the multicenter prospective design, potential center effects will be considered in all adjusted analyses. Study center will be included as a covariate or treated as a stratification factor, as appropriate, to account for intercenter heterogeneity. Stratified analyses across study centers may be conducted when the sample size permits.

Reporting of Results

Effect estimates will be reported with corresponding 95% CIs and *P* values.

Serum Metabolomics Analysis

Serum samples stored at −80 °C will be thawed and sorted on ice according to their respective groups. Subsequently, 30 µL of serum will be transferred into an Eppendorf tube, and 120 µL of prechilled 80% methanol containing internal standard I will be added for extraction. The samples will then be centrifuged at 4 °C at 18,000×*g* for 30 minutes, and the supernatant will be collected. EDC and 3-nitrophenylhydrazine will be added to the supernatant for derivatization for 60 minutes. The reaction will be terminated by adding an equal volume of ice-cold 50% methanol, followed by centrifugation at 4 °C at 4000×*g* for 30 minutes. The resulting supernatant will be collected, internal standard II will be added, and the mixture will be prepared for injection. Analysis will be performed on an ultra-high performance liquid chromatography coupled to a quadrupole or time-of-flight high-resolution mass spectrometer (positive and negative ion modes). Raw metabolomics data will be preprocessed with MassLynx 4.1 (Waters Corporation) and imported into SIMCA-*P* 14.1 (MKS Umetrics) for Orthogonal Partial Least Squares Discriminant Analysis. Differential metabolites will be identified against the Human Metabolite Database (University of Alberta). Finally, pathway enrichment of differential metabolites will be performed using KEGG (Kanehisa Laboratories) via MetaboAnalyst 5.0 (McGill University).

Ethical Considerations

This study has been approved by the Ethics Committee of Beijing Longfu Hospital (LFYYLL-2025-02) and will be conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines (Multimedia Appendix 2). Written informed consent will be obtained from all participants before enrollment (Multimedia Appendix 3). Participants may withdraw from the study at any time without penalty. No financial or material compensation will be provided for participation. Venous blood sampling may cause temporary discomfort, bruising, or mild dizziness. The study will not provide participants with any medications. To safeguard data security, direct identifiers are removed at collection and replaced with coded IDs; the coding key is stored separately, with access limited to authorized staff. An ethics-approved, comprehensive data storage and management plan has been implemented to minimize any risk of breach.

Results

The study protocol (version 1.0) was finalized in February 2025. Participant recruitment began in March 2025 across 8 participating centers and is expected to be completed by December 2026. Final data collection, including all follow-up assessments, is anticipated to be completed in March 2027. As of February 2026, a total of 500 participants have been enrolled, with 420 (84%) in the LWDH exposure group and 80 (16%) in the nonexposure group. Among them, 230 (46%) participants have completed follow-up assessments (n=200, 87% in the exposure group and n=30, 13% in the nonexposure group). In addition, pretreatment and posttreatment blood samples have been collected from 20 (4%) participants for metabolomics analysis (n=15, 75% in the exposure group and n=5, 25% in the nonexposure group). We expect the future results of this study to be published in esteemed, peer-reviewed health research journals by December 2027.

Discussion

Expected Findings

We hypothesize that patients receiving LWDH in routine clinical practice may experience greater improvement in menopausal symptoms, hormone indices, and quality of life compared with those receiving standard Western medical care alone. In addition, exploratory serum metabolomics analysis is expected to provide preliminary insights into metabolic pathways potentially involved in symptom improvement.

Comparison to Prior Work

Previous studies have suggested that LWDH may alleviate menopausal symptoms through modulation of endocrine function and metabolic homeostasis [34]; however, most available evidence is derived from randomized controlled trials conducted under controlled conditions. In contrast, this study adopts a prospective, real-world design with preference-based treatment allocation, aiming to complement existing evidence by assessing treatment effectiveness under routine clinical conditions. By integrating clinical outcomes with exploratory metabolomics analysis, this study seeks to bridge clinical effectiveness research and mechanistic hypothesis generation.

Strengths and Limitations

The strengths of this study include its prospective real-world design; comprehensive collection of clinical and laboratory outcomes; and the application of advanced statistical methods, such as propensity score–based IPTW, to mitigate confounding associated with nonrandomized group allocation.

Several limitations should also be acknowledged. First, despite statistical adjustment, residual confounding due to unmeasured

XSL•FO
RenderX

factors cannot be fully excluded, and causal inference should therefore be interpreted with caution. Second, key outcomes rely on self-reported scales, which may introduce subjective reporting bias. Third, the relatively short follow-up period focuses on short-term effectiveness and safety and does not allow evaluation of long-term outcomes or delayed adverse effects.

Future Directions

Future studies with longer follow-up durations, larger sample sizes, and more rigorous designs are warranted to confirm the sustainability of treatment effects and to further investigate the mechanistic pathways suggested by the exploratory metabolomics findings. These results may inform the design of subsequent randomized trials or longitudinal cohort studies integrating clinical outcomes with molecular profiling.

Dissemination Plan

The findings of this study will be disseminated through peer-reviewed publications and academic conferences focusing on integrative medicine, menopausal health, and real-world evidence research. The results may also inform clinical decision-making and future guideline development for the management of MPS.

Conclusions

This prospective, multicenter, real-world cohort study will provide the first large-scale clinical evidence regarding the effectiveness and safety of LWDH in women with MPS due to kidney-yin deficiency. By integrating clinical outcomes with exploratory serum metabolomics analysis, the study aims to generate real-world comparative effectiveness evidence and preliminary hypotheses regarding potential metabolic pathways associated with treatment response. The findings will offer robust evidence to support clinical practice, promote precision medicine, and facilitate the development of integrative strategies for improving women's health during the menopausal transitions.

Acknowledgments

The authors gratefully acknowledge the contributions of the following hospitals: Beijing Longfu Hospital, Beijing Gulou Traditional Chinese Medicine Hospital, Lanxi Traditional Chinese Medicine Hospital, The First Affiliated Hospital of Harbin Medical University, Beijing Dongcheng District Hepingli Community Health Service Center, Beijing Dongcheng District Chaoyangmen Community Health Service Center, Beijing Puren Hospital, and Tsinghua University Yuquan Hospital (Tsinghua University Hospital of Integrative Traditional Chinese and Western Medicine).

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Funding

This work is supported by National Key Research and Development Program of China (2022YFC3502004).

Authors' Contributions

LW, SW, QS, and SZ conceived the research topic. LW, SW, and HZ designed the study. SW drafted the manuscript. MW and LW critically revised the manuscript. YS, TS, WZ, Y Ye, LM, Y Yang, TW, LC, CS, and XD served as local principal investigators at subcenters and were responsible for participant recruitment and data collection. All authors approved the final manuscript and take responsibility for the decision to submit it for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
SPIRIT checklist.
[PDF File (Adobe PDF File), 108 KB - resprot_v15i1e84803_app1.pdf]

Multimedia Appendix 2
Ethics statement.
[DOCX File , 11 KB - resprot_v15i1e84803_app2.docx]

Multimedia Appendix 3
Informed consent form.
[DOC File , 61 KB - resprot_v15i1e84803_app3.doc]

References

XSL•FO
RenderX

1. Monteleone P, Mascagni G, Giannini A, Genazzani AR, Simoncini T. Symptoms of menopause - global prevalence, physiology and implications. Nat Rev Endocrinol 2018 Apr;14(4):199-215. [doi: 10.1038/nrendo.2017.180] [Medline: 29393299]
2. Carr MC. The emergence of the metabolic syndrome with menopause. J Clin Endocrinol Metab 2003 Jun;88(6):2404-2411. [doi: 10.1210/jc.2003-030242] [Medline: 12788835]
3. Hill K. The demography of menopause. Maturitas 1996 Mar;23(2):113-127 [FREE Full text] [doi: 10.1016/0378-5122(95)00968-x] [Medline: 8735350]
4. Fang Y, Liu F, Zhang X, Chen L, Liu Y, Yang L, et al. Mapping global prevalence of menopausal symptoms among middle-aged women: a systematic review and meta-analysis. BMC Public Health 2024 Jul 02;24(1):1767 [FREE Full text] [doi: 10.1186/s12889-024-19280-5] [Medline: 38956480]
5. Mou H, Zhang J, Guo Y, Xu L, Luo X. Publisher correction: effects of key physiological parameters on cardiovascular disease and osteoporosis risk in perimenopausal and postmenopausal women. Sci Rep 2025 Mar 25;15(1):10251 [FREE Full text] [doi: 10.1038/s41598-025-95430-y] [Medline: 40133641]
6. Klisic A, Ahmad R, Daka B, Sindhu S. Editorial: cardiometabolic diseases in postmenopausal women. Front Endocrinol (Lausanne) 2024;15:1514913 [FREE Full text] [doi: 10.3389/fendo.2024.1514913] [Medline: 39574957]
7. Mosconi L, Berti V, Dyke J, Schelbaum E, Jett S, Loughlin L, et al. Menopause impacts human brain structure, connectivity, energy metabolism, and amyloid-beta deposition. Sci Rep 2021 Jun 09;11(1):10867 [FREE Full text] [doi: 10.1038/s41598-021-90084-y] [Medline: 34108509]
8. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, et al. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. JAMA 2002 Jul 17;288(3):321-333 [FREE Full text] [doi: 10.1001/jama.288.3.321] [Medline: 12117397]
9. Chlebowski RT, Hendrix SL, Langer RD, Stefanick ML, Gass M, Lane D, et al. Influence of estrogen plus progestin on breast cancer and mammography in healthy postmenopausal women: the Women's Health Initiative randomized trial. JAMA 2003 Jun 25;289(24):3243-3253. [doi: 10.1001/jama.289.24.3243] [Medline: 12824205]
10. Butt DA, Lock M, Lewis JE, Ross S, Moineddin R. Gabapentin for the treatment of menopausal hot flashes: a randomized controlled trial. Menopause 2008;15(2):310-318. [doi: 10.1097/gme.0b013e3180dca175] [Medline: 17917611]
11. Sicat BL, Brokaw DK. Nonhormonal alternatives for the treatment of hot flashes. Pharmacotherapy 2004 Jan;24(1):79-93. [doi: 10.1592/phco.24.1.79.34812] [Medline: 14740790]
12. Wang YP, Yu Q. The treatment of menopausal symptoms by traditional Chinese medicine in Asian countries. Climacteric 2021 Feb;24(1):64-67. [doi: 10.1080/13697137.2020.1832461] [Medline: 33094658]
13. Jin Z, Huang J, Cai Z. Guidelines for diagnosis and treatment of menopausal syndrome with integrated traditional Chinese and western medicine (2023 edition)Article in Chinese. Chinese Journal of Practical Gynecology and Obstetrics, 2023 2023 Aug;39(8):799-808 [FREE Full text] [doi: 10.19538/j.fk2023080109]
14. Smith-Francis MJ. Complementary and alternative medicine for menopause. Nurs Clin North Am 2024 Dec;59(4):551-562. [doi: 10.1016/j.cnur.2024.08.001] [Medline: 39477563]
15. Kargozar R, Azizi H, Salari R. A review of effective herbal medicines in controlling menopausal symptoms. Electron Physician 2017 Nov;9(11):5826-5833 [FREE Full text] [doi: 10.19082/5826] [Medline: 29403626]
16. Fu SF, Zhao YQ, Ren M, Zhang JH, Wang YF, Han LF, et al. A randomized, double-blind, placebo-controlled trial of Chinese herbal medicine granules for the treatment of menopausal symptoms by stages. Menopause 2016 Mar;23(3):311-323. [doi: 10.1097/GME.0000000000000534] [Medline: 26671188]
17. Scheid V. Traditional Chinese medicine--what are we investigating? The case of menopause. Complement Ther Med 2007 Mar;15(1):54-68 [FREE Full text] [doi: 10.1016/j.ctim.2005.12.002] [Medline: 17352972]
18. Chen Q, Zhang Y, Meng Q, Wang S, Yu X, Cai D, et al. Liuwei Dihuang prevents postmenopausal atherosclerosis and endothelial cell apoptosis via inhibiting DNMT1-medicated ERα methylation. J Ethnopharmacol 2020 Apr 24;252:112531. [doi: 10.1016/j.jep.2019.112531] [Medline: 31926314]
19. Wang X, Zhang F, Zhang J, Du Y, Zhang T, Liu S, et al. Liuwei Dihuang Formula ameliorates glycolipid metabolism in type 2 diabetic rats through an integrated analysis of gut microbiota and bile acids. J Pharm Biomed Anal 2026 Jan 15;268:117181. [doi: 10.1016/j.jpba.2025.117181] [Medline: 41092499]
20. Zhan X, Li H, Jin J, Ju X, Gao J, Chen X, et al. Network pharmacology and experimental validation to explore the role and potential mechanism of Liuwei Dihuang Decoction in prostate cancer. BMC Complement Med Ther 2024 Jul 26;24(1):284 [FREE Full text] [doi: 10.1186/s12906-024-04572-5] [Medline: 39061044]
21. Perry B, Zhang J, Saleh T, Wang Y. Liuwei Dihuang, a traditional Chinese herbal formula, suppresses chronic inflammation and oxidative stress in obese rats. J Integr Med 2014 Sep;12(5):447-454. [doi: 10.1016/S2095-4964(14)60044-3] [Medline: 25292344]
22. Zhou J, Gao W, Dong Y, Yu J, Hu X, Zhang L. Effects of hormone replacement therapy combined with Liuwei Dihuang pills on serum estradiol and prolactin levels and clinical efficacy in female menopausal syndrome patients. Prog Mod Biomed 2016;16(32):6302 [FREE Full text] [doi: 10.13241/j.cnki.pmb.2016.32.024]

XSL•FO
RenderX

23. Zhong D. Clinical study on Liuwei Dihuang Pill combined with Jiawei Xiaoyao Pill in the treatment of perimenopausal syndrome Article in Chinese. New Chinese Medicine 2019;51(6):79-83 [FREE Full text] [doi: 10.13457/j.cnki.jncm.2019.06.024]
24. Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 08;346:e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
25. Chen J. Clinical effect of modified Liuwei Dihuang Wan in treating kidney Yin deficiency type perimenopausal syndrome Article in Chinese. J Clin Med Res Prac 2020;5(33):152-153. [doi: 10.19347/j.cnki.2096-1413.202033056]
26. Bertke S, Hein M, Schubauer-Berigan M, Deddens J. A simulation study of relative efficiency and bias in the nested case-control study design. Epidemiol Methods 2013 Sep;2(1):85-93 [FREE Full text] [doi: 10.1515/em-2013-0007] [Medline: 26345580]
27. Harlow SD, Gass M, Hall JE, Lobo R, Maki P, Rebar RW, et al. Executive summary of the Stages of Reproductive Aging Workshop + 10: addressing the unfinished agenda of staging reproductive aging. Menopause 2012 Apr;19(4):387-395 [FREE Full text] [doi: 10.1097/gme.0b013e31824d8f40] [Medline: 22343510]
28. Lumsden MA, Davies M, Sarri G, Guideline Development Group for Menopause: Diagnosis and Management (NICE Clinical Guideline No. 23). Diagnosis and management of menopause: the National Institute of Health and Care Excellence (NICE) guideline. JAMA Intern Med 2016 Aug 01;176(8):1205-1206. [doi: 10.1001/jamainternmed.2016.2761] [Medline: 27322881]
29. Tao M, Shao H, Li C, Teng Y. Correlation between the modified Kupperman Index and the Menopause Rating Scale in Chinese women. Patient Prefer Adherence 2013;7:223-229 [FREE Full text] [doi: 10.2147/PPA.S42852] [Medline: 23569361]
30. Wu X, Zhou L, Dong H, Tian M, Liu S, Xu X. Efficacy, safety and mechanism of Honghua Xiaoyao Pill in the treatment of peri-menopausal syndrome: a study protocol for a randomized controlled trial. Front Pharmacol 2022;13:1001228 [FREE Full text] [doi: 10.3389/fphar.2022.1001228] [Medline: 36569324]
31. Academic Committee of the Korean Society of Menopause, Lee SR, Cho MK, Cho YJ, Chun S, Hong SH, et al. The 2020 menopausal hormone therapy guidelines. J Menopausal Med 2020 Aug;26(2):69-98 [FREE Full text] [doi: 10.6118/jmm.20000] [Medline: 32893509]
32. Taylor HS, Manson JE. Update in hormone therapy use in menopause. J Clin Endocrinol Metab 2011 Feb;96(2):255-264. [doi: 10.1210/jc.2010-0536] [Medline: 21296989]
33. Electronic Data Capture System. URL: https://www.yiankb.com/edc/ [accessed 2026-03-20]
34. Huang L, Shu Y, Zhang H, Zhang X. Effects of Liuwei Dihuang pills combined with low-dose Estrogen-Progesterone replacement therapy on sex hormones, bone mineral density and blood lipids in patients with perimenopausal syndrome. Prog Mod Biomed 2021;21(16):3120-3124 [FREE Full text] [doi: 10.13241/j.cnki.pmb.2021.16.025]

Abbreviations

CRF: case report form
EDC: electronic data capture
IPTW: inverse probability of treatment weighting
LWDH: Liuwei Dihuang
MPS: menopausal syndrome
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
TCM: traditional Chinese medicine

Edited by A Schwartz; submitted 25.Sep.2025; peer-reviewed by R Chen, X Wang; comments to author 19.Jan.2026; accepted 11.Mar.2026; published 25.Mar.2026.

Please cite as:
Wei S, Zhang H, Wang M, Sun Y, Song T, Zheng W, Ye Y, Ma L, Yang Y, Wei T, Cheng L, Sun C, Ding X, Zhao S, Song Q, Wang L Real-World Effectiveness and Safety of Liuwei Dihuang Pill for Menopausal Syndrome: Protocol for a Prospective, Observational, Multicenter Cohort Study
JMIR Res Protoc 2026;15:e84803
URL: https://www.researchprotocols.org/2026/1/e84803
doi:10.2196/84803
PMID:

©Shuoshuo Wei, Hongyan Zhang, Mengmeng Wang, Yuhui Sun, Tao Song, Weifang Zheng, Yunmei Ye, Liran Ma, Yunshuang Yang, Tong Wei, Lan Cheng, Chunquan Sun, Xingting Ding, Shichu Zhao, Qingqiao Song, Lianxin Wang. Originally published

XSL•FO
RenderX

in JMIR Research Protocols (https://www.researchprotocols.org), 25.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Safety of Postpartum Herbal Medicine in Primary Care Korean Medicine Clinics: Protocol for a Prospective, Multicenter, Registry-Based Observational Study

Anna Kim[1*], PhD; Eunbyul Cho[2*], PhD; Young-Eun Kim[1], MSc; Sumin Seo[1], PhD; Sungha Kim[1], PhD; Jae-Ahn Shin[3], BSc; Eunhee Lee[4], PhD; Mi Ju Son[1], PhD

[1]Korea Institute of Oriental Medicine, Daejeon, Republic of Korea

[2]Wonkwang University, Iksan, Republic of Korea

[3]Jayoon Korean Medicine Clinic, Bucheon, Republic of Korea

[4]Woosuk University Korean Medicine Hospital, Jeonju, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Mi Ju Son, PhD
Korea Institute of Oriental Medicine
1672 Yuseong-daero
Daejeon, 34054
Republic of Korea
Phone: 82 42 868 9446
Email: mj714@kiom.re.kr

Abstract

Background: Postpartum conditions significantly affect maternal and infant health. In South Korea, herbal medicine is commonly used for postpartum care in Korean medicine, but safety data remain limited. This study outlines a multicenter Korean medicine clinic registry designed to investigate treatment patterns and adverse events (AEs) associated with postpartum herbal medicine use. This primary care, practice-based research network registry will inform clinical decision-making and pharmacovigilance in postpartum care, supporting the quality and safety of care and guidance for comedication monitoring across community clinics.

Objective: This study aims to establish a multicenter registry to investigate AEs and treatment patterns associated with postpartum herbal medicine use in Korean medicine clinics.

Methods: This prospective observational registry will enroll postpartum patients receiving herbal medicine treatment from 15 Korean medicine clinics between July 2024 and June 2027, with a target sample size of 1000 participants. Patients within 6 months postpartum receiving herbal medicine will be included. Data will be collected for up to 8 weeks, including demographic information, medical history, herbal medicine prescriptions, AEs, and concomitant medication use. AEs are self-reported by patients, and causality is assessed using the World Health Organization Uppsala Monitoring Centre system and the Naranjo algorithm. Data collection and entry will follow standard operating procedures.

Results: The registry was registered with the Clinical Research Information Service in June 2024 (KCT0009552). Recruitment began in July 2024 across 15 Korean medicine clinics and is planned to continue until June 2027 with a target enrollment of 1000 postpartum patients.

Conclusions: This study represents the first large-scale, multicenter registry examining herbal medicine use during the postpartum period in Korean medicine clinics. By systematically collecting real-world data on AEs associated with postpartum herbal medicine use, we aim to establish a comprehensive safety database for individual herbal medicine and constituent herbs. The findings may support pharmacovigilance of herbal medicine and improve postpartum care outcomes for mothers and infants.

Trial Registration: WHO ICTRP KCT0009552; https://trialsearch.who.int/?TrialID=KCT0009552

International Registered Report Identifier (IRRID): DERR1-10.2196/87543

(JMIR Res Protoc 2026;15:e87543) doi:10.2196/87543

XSL•FO
RenderX

KEYWORDS

postpartum period; herbal medicine; safety; adverse events; registry

Introduction

Background

Childbirth, although a natural and universal process, imposes considerable stress on the mother's body, often resulting in a debilitated condition and a low quality of life [1,2]. Studies show that 49% to 94% of women experience at least 1 health issue up to 8 weeks post partum [3,4]. These issues include fatigue [5-9], headache [10-12], musculoskeletal pain [8,13,14], urinary incontinence [15-17], and depression [18-21], often persisting beyond the first year post partum. These prevalent and prolonged health problems significantly affect the well-being of both mothers and their children [22]. This situation highlights the need for safe and effective postpartum care strategies, leading many to consider herbal medicine as a treatment option [23,24].

Herbal medicine offers various potential benefits for postpartum women. Studies have shown that can mitigate postpartum pain [25]. Additionally, when used in conjunction with conventional treatments, herbal medicine may alleviate postpartum depression [26]. In East Asian countries, where herbal medicine is integrated into national health care systems, its use in postpartum care is particularly prevalent, with high satisfaction among consumers. Nevertheless, the lack of safety data on herbal medicine for both mothers and infants remains a considerable concern, possibly deterring some from using herbal medicine [23,27]. For instance, a study in Taiwan found that breastfeeding mothers frequently used HM; however, those consuming traditional Chinese herbs had elevated lead levels in their breast milk, emphasizing the importance of safety monitoring [28]. This global context underscores the urgent need for rigorous safety studies on postpartum herbal medicine use. In Korean medicine practice, postpartum herbal treatment often includes individualized prescriptions derived from traditional formulations such as *Boheo-tang*, *Ssanghwa-tang*, *Gunggwijohyeol-eum*, and *Dangguijakyak-san*, which are commonly used to address postpartum symptoms, including fatigue, musculoskeletal pain, and circulatory disturbances. Recent studies have reported that herbal medicine interventions may improve pain, quality of life, and depressive symptoms in postpartum women while demonstrating acceptable safety profiles [25,29,30].

In South Korea, Korean medicine plays an active role in addressing postpartum symptoms, with herbal medicine commonly used alongside acupuncture as the standard treatment method. According to a study on the satisfaction of mothers who received postpartum treatment with Korean medicine during the early postpartum period, 78.68% indicated a willingness to use Korean medicine again in future childbirths [31], while satisfaction with herbal medicine treatment was notably high at 72.13% [32]. Numerous local governments in South Korea support the use of herbal medicine after childbirth by subsidizing expenses, further indicating its perceived value in postpartum care [32].

However, South Korea's drug use review service does not provide adverse event (AE) information related to herbal medicine. While some national statistics provide estimates on the incidence rate of AEs related to herbal medicine, these figures are based on patient reports or survey data, lacking expert case evaluations and detailed information on specific herbal medicine treatments [33]. Moreover, some reported AEs may not be directly attributable to herbal medicine, underscoring the importance of assessing a clear causal relationship [34]. This lack of comprehensive safety data creates a significant gap in our understanding of the risks associated with postpartum herbal medicine use.

To address these concerns, especially the lack of safety information from official sources, this study aims to establish a registry for those receiving herbal medicine treatment at primary care institutions, specifically Korean medicine clinics. Because Korean medicine clinics account for approximately 96% of Korean medicine institutions in South Korea [35], research based on primary care Korean medicine clinics is crucial for collecting representative clinical data on Korean medicine. Through the systematic collection of data on the constituents of herbal medicine, details of AEs, and causality assessment, we seek to provide evidence for the safety monitoring of herbal medicine administration during the postpartum period.

The primary research questions of this study are as follows: (1) What is the incidence rate of AEs following the use of postpartum herbal medicine? (2) Which herbal medicines and constituent herbs are involved in reported AE cases? (3) What is the causality classification of the observed AEs?

This study represents the first nationwide multicenter registry examining the use of herbal medicine during the postpartum period in Korean medicine clinics across South Korea. By providing comprehensive safety data, this research aims to contribute to evidence-based clinical practice in South Korea and ultimately enhance patient safety for both mothers and infants.

Study Objective

The objective of the Safety of Herbal Medicine Registry in Korean medicine clinics (SAFEHERE-KM) is to systematically collect and analyze clinical data on AEs related to herbal medicine treatment in postpartum patients in South Korea. This web-based registry will collect patient characteristics, details of herbal medicine prescriptions, and AE information, including a causality assessment of any AEs that arise during herbal medicine treatment using the World Health Organization Uppsala Monitoring Centre causality assessment [36] and the Naranjo algorithm score [37,38]. By establishing AE monitoring databases, this registry further aims to contribute to the development of an active pharmacovigilance system for herbal medicine in South Korea.

XSL•FO
RenderX

Methods

Study Design

The SAFEHERE-KM registry is a multicenter, prospective, observational study designed to collect data on routine clinical care and associated AEs in postpartum women receiving herbal medicine treatment at Korean medicine clinics. This study aims to enroll at least 1000 eligible patients between July 2024 and June 2027 based on a previous study suggesting that this sample size rapidly decreases variation when reporting AE incidence [39]. Using empirical observations of postpartum patient visits to participating Korean medicine clinics, we anticipate enrolling approximately 400 participants annually (approximately 27 patients per clinic per year).

The participants are postpartum patients who voluntarily consent to the collection of their clinical data. This study does not introduce specific interventions. Instead, the patients receive individualized herbal medicine treatments as part of their routine postpartum care. Korean medicine doctors prescribe herbal medicine at their discretion without restrictions. This registry focuses on AEs experienced by postpartum women receiving herbal medicine, and infant AEs are not systematically assessed because infants are not enrolled as study participants.

Throughout the treatment period, which may extend up to 8 weeks depending on the duration of prescribed herbal medicine use, we collect data on demographics, medical history, Korean medicine treatment details, herbal medicine prescriptions, AEs, and concomitant medication use. AEs are self-reported by patients via an online survey and recorded in online databases. AEs, including the absence of an AE, will be recorded through scheduled assessments approximately every 2 weeks during the herbal medicine treatment period, with follow-up continuing for up to 8 weeks. The electronic case report form includes predefined fields to record AEs and to indicate whether an event meets the criteria for a serious AE. To ensure consistency in data collection and entry, Korean medicine doctors follow standard operating procedures specifically designed for this registry. All reported AEs, including serious AEs, are reviewed and validated by licensed Korean medicine doctors.

Patients will be recruited from 15 Jayoon Korean medicine clinics, a network specializing in gynecological disorders, located across South Korea. This network provides wide geographical coverage, including 5 clinics each in Seoul and Gyeonggi-do, as well as clinics in Jeju Province, Busan, Incheon, Daejeon, and Daegu (Multimedia Appendix 1).

The registry was registered in the Clinical Research Information Service (KCT0009552). Recruitment started on July 7, 2024.

Eligibility Criteria

This registry enrolls patients diagnosed with complications predominantly related to the puerperium (O85-O92 of the International Statistical Classification of Diseases-10) or puerperal wind disorder (U32.7 of the Korean Standard Classification of Diseases-8). Puerperal wind disorder indicates a set of symptoms including pain and autonomic disturbance caused by inadequate body care after childbirth [1]. Women who visited the primary Korean medicine clinics within 6 months after giving birth and have been prescribed herbal medicine for postpartum recovery are included. The inclusion and exclusion criteria are presented in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Women aged ≥19 years
- Have visited primary Korean medicine clinics within 6 months post partum (calculated from the date of delivery) and are undergoing herbal medicine treatment for postpartum recovery

Exclusion criteria

- Did not consent to participate
- Deemed inappropriate for participation in this study

Recruitment

Participants will be enrolled in 15 primary Korean medicine clinics that offer postpartum care. Recruitment posters will be placed at the entrance of each clinic to encourage voluntary participation. Health care professionals will explain the study's purpose and procedures—including data use, confidentiality, and data sharing for analysis—and obtain written informed consent from eligible participants before data collection. Participants may withdraw from the study at any time without any impact on their clinical care. Recruitment began in July 2024 and is currently ongoing across all 15 Korean medicine clinics.

Study Outcomes

For participants who meet the inclusion criteria and voluntarily sign the written informed consent form, we will assign a study registration number. Subsequently, we will collect data for up to 8 weeks while they consume herbal medicines (Textbox 2).

Table 1 presents the details of the data collection and the study schedule.

XSL•FO
RenderX

Textbox 2. Data collected for up to 8 weeks while postpartum patients consume herbal medicines.

- Demographics: date of birth and age
- Anthropometric measurements: height and weight
- Medical history (maternal health and delivery details): delivery information (date and mode of delivery and gestational age); obstetric history (number of pregnancies, full-term births, preterm births, miscarriages, and surviving children); high-risk pregnancy factors (advanced maternal age [≥35 years], in vitro fertilization, gestational hypertension, and gestational diabetes); and newborn information (demographics, birth details, and feeding methods)
- Details of herbal medicine prescription: name of prescribed herbal medicine, start date of administration, composition (herbs and doses), total and 1-day dose of herbal medicine, and duration of herbal medicine administration
- Other Korean medicine treatments: concurrent treatments, such as acupuncture, moxibustion, cupping, Chuna manipulation, and Korean medicine physical therapy
- Details of adverse events (AEs): the occurrence of AEs is monitored at each visit during herbal medicine intake for up to 8 weeks. If an AE occurs, the collected data include the AE start date, symptoms, outcome, severity, actions taken, recurrence after the reintroduction of the suspected herbal medicine, treatment methods for the AE, and concurrent medications. Follow-up telephone interviews may be conducted for detailed information and causality assessments.

Table 1. Study schedule.

Time points	Visit 1: baseline[a]	Subsequent visits at 2, 4, 6, and 8 weeks after taking herbal medicine[b,c]	End point: follow-up completion[a,d]
Informed consent	✓		
Demographics	✓		
Diagnosis	✓		
Eligibility screening	✓		
Anthropometrics	✓		
Medical history	✓		
Concomitant medication	✓	✓	✓
Korean medicine treatment information			
Details of the herbal medicine prescription			✓
Other Korean medicine treatments			✓
Check for adverse events			
Occurrence status		✓	✓
Details of adverse events[e]		✓	✓

[a]Data collected from the Korean medicine doctor.

[b]Visits 3, 4, and 5 may be conducted if herbal medicine intake continues, with the process occurring at 2-week intervals for up to 8 weeks.

[c]Data collected from the patient.

[d]Follow-up completion is conducted once data collection of adverse events is finalized or when herbal medicine administration is completed.

[e]Data are collected only in the presence of adverse events. Adverse events represent the primary outcome of this registry.

Data Management and Quality Control

For data collection, we will use the myTrial Electronic Data Capture system (NIKOM), an electronic case report form system validated by the National Agency for Korean Medicine Innovative Technologies Development. Each clinic's Korean medicine doctor can access the data entry form via the web server from a remote computer with internet access by logging in using the credentials provided by the website administrator [40]. Each case entered into the registry will be reviewed regularly, and data queries will be generated according to the data monitoring manual. Monitors regularly visit Korean medicine clinics to check the original registry records and research files, and they discuss and resolve any issues with the coinvestigators and principal investigators at each institution. All data collected for this study will be securely stored on a dedicated server with strict access-control permissions.

Statistical Analysis and Causality Assessment

Statistical analyses will be performed using R (version 4.5.0; R Foundation for Statistical Computing). For participants who withdraw consent for continued follow-up, the data collected until withdrawal will be used in the analysis. For continuous variables, the means and SDs will be presented. Categorical variables will be reported as frequencies and percentages. Herbal medicine exposure will be categorized based on high-frequency

XSL•FO
RenderX

prescriptions, major prescription groups, and functional classifications. Patterns of herbal medicine use in AE cases will be explored descriptively. The incidence of AEs will be presented as the proportion of participants experiencing at least 1 AE, and all AEs will be summarized using descriptive statistics. In addition, for cases in which AEs occur, the corresponding prescribed herbal medicine and their constituent herbs will be summarized using frequencies and proportions to describe observed patterns.

For any reported AEs, severity will be determined based on established criteria for assessing AE severity [41]. These criteria are commonly used in clinical safety studies to evaluate the severity of AEs and are applicable to general clinical symptoms regardless of treatment modality. The causality of AEs will be evaluated using both the World Health Organization Uppsala Monitoring Centre causality assessment [36] and the Naranjo algorithm score [37,38], based on the approach used in similar previous studies [42,43]. This evaluation will be carried out by an independent Adverse Reaction Assessment Committee affiliated with the Regional Drug Safety Center at the Korea Institute of Drug Safety and Risk Management.

Patient and Public Involvement

Patients will not be involved in the design, conduct, and dissemination plans of this research, as this registry study aims to collect real-world data on routine clinical practice.

Safety Considerations

While AE data are collected for research purposes, the treatment and management of any AEs are conducted by Korean medicine doctors as part of routine clinical care, ensuring that the participants receive appropriate care throughout the study period.

Ethical Considerations

This registry protocol received approval (version 1.0; April 24, 2024) from the institutional review board of Woosuk Korean medicine Hospital, Woosuk University, Jeonju, Republic of Korea (H2404-01-01). Following amendments to the research plan, the updated protocol (version 1.1) was subsequently approved on June 24, 2024. Any protocol amendments will be approved by the institutional review board and communicated to all investigators. The findings will be shared through publication in peer-reviewed journals and presentations at academic conferences. Written informed consent will be obtained from all participants prior to enrollment. Participant data will be recorded in case report forms and stored in secure databases with restricted access to ensure privacy and confidentiality. Participants will not receive financial compensation for participation in this registry.

Results

The SAFEHERE-KM registry was registered with the Clinical Research Information Service on June 20, 2024 (KCT0009552) [44]. Recruitment began on July 7, 2024, across 15 participating Korean medicine clinics. Enrollment is planned to continue until June 2027 with a target sample size of 1000 postpartum patients.

Discussion

The SAFEHERE-KM registry is expected to generate comprehensive real-world safety data on herbal medicine use during the postpartum period in Korean medicine clinics. By systematically collecting AE information and treatment details, this registry aims to provide an evidence base for monitoring the safety of postpartum herbal medicine use and supporting pharmacovigilance in primary care settings.

The SAFEHERE-KM study is the first large-scale, multicenter registry specifically designed to assess the safety of herbal medicine use during the postpartum period in primary Korean medicine clinics. This unique approach allows for the collection of real-world data on herbal medicine use in a setting where it is most commonly prescribed, potentially providing more representative and clinically relevant information than our previous study conducted in a Korean medicine hospital [45].

The puerperal period commonly extends beyond the initial 6 to 8 weeks when reproductive organ recovery is mostly complete. Many women continue to experience physical and mental burden because of childcare, household responsibilities, and stress. Insufficient care during this extended period can lead to various complications [1]. A previous survey has shown that mothers typically seek herbal medicine treatment at Korean medicine institutions approximately 13 to 14 weeks post partum. Recognizing the need for prolonged health management, many local governments in Korea provide support for Korean medicine treatments for up to 6 months post partum [32]. Therefore, we include women within 6 months of delivery in our registry. The decision to monitor AEs for up to 8 weeks after the initiation of herbal medicine treatment was based on clinical experience from participating Korean medicine doctors, who reported that most patients complete their herbal medicine treatments within this period. This approach allows us to maximize participant enrollment within our budget constraints while capturing the most relevant period of herbal medicine use in postpartum care.

The anticipated results of this study have several important clinical implications. By systematically collecting data on AEs associated with postpartum herbal medicine use, we aim to establish a comprehensive safety database for individual herbal medicine and constituent herbs. This dataset will include signs, symptoms, frequency, and severity of AEs, thereby providing valuable insights for both practitioners and patients. Furthermore, the causality assessment of reported AEs will strengthen the evidence for herbal medicine safety, promoting the development of guidelines for herbal medicine monitoring and decision-making processes in postpartum care. A dataset on causal herbs, their dosages, and concomitant drugs will enable the advanced analysis of herb-herb and herb-drug interactions [46].

Methodologically, the SAFEHERE-KM study has several strengths. The use of standardized data collection procedures, including detailed standard operating procedures and an electronic case report form, ensures consistency and quality in data gathering across multiple sites. This study collects closely monitored, patient-reported AEs to prevent underreporting and

XSL•FO
RenderX

collect patients' viewpoints [47]. We will recruit at least 1000 patients to gain an in-depth understanding of the AEs associated with herbal medicine. Patients are being recruited from 15 geographically diverse Korean medicine clinics across South Korea to minimize selection bias. Moreover, through data collection in real-world clinical settings across multiple primary care Korean medicine clinics, our results may achieve high external validity, making the findings more applicable to clinical practice [48].

Nevertheless, some limitations of this study should be acknowledged. The recruitment of participants from a single network of Korean medicine clinics that specialize in gynecological disorders may limit the diversity of herbal medicine prescriptions. While Korean medicine doctors in this network prescribe personalized herbal medicine for individual patients instead of uniform prescriptions, this may not capture the full range of herbal medicine practices across South Korea. Additionally, reliance on patient-reported AEs may underestimate unfavorable clinical events [49], although this is mitigated by standardized reporting procedures and visits to Korean medicine clinics during routine medical treatment. Because this registry primarily focuses on AEs associated with individual prescriptions, cumulative dose-response analyses based on total herbal medicine exposure may not be feasible.

Future research should build upon the foundation established by the SAFEHERE-KM study. The inclusion of a wider range of primary care institutions could provide more comprehensive evidence of herbal medicine use in postpartum care across various practice settings. Furthermore, long-term follow-up studies could offer insights into the long-term safety profile of postpartum herbal medicine use, which is beyond the scope of this study.

In conclusion, the SAFEHERE-KM registry is expected to provide important real-world evidence on the safety of postpartum herbal medicine use in Korean medicine clinics. This registry may contribute to improved pharmacovigilance and a better understanding of the safety profile of postpartum herbal medicine use. The findings of this study will contribute not only to the field of Korean medicine but also to global integrative approaches to postpartum care.

Acknowledgments

The authors thank the members of the Jayoon Korean medicine Clinic Network for their collaboration: Hee Eun Jung, Min Jae Gwak, Hwa Seon Ko, Ji Hyun Kim, Jun Hak Lee, Seo Kyung Kim, Hwang Jin Park, Ey Ryung Jung, Won Joon Cho, Chang Yun Lee, Jun Young Jo, Hyeon Ju Kim, Jong Soon Beak, Seul Ah Jang, Yu Jeong Choi, and Eun Ji Choi.

Funding

This work was supported by the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health and Welfare, Republic of Korea (grant RS-2023-KH138688). The funding source had no input regarding the design or publication of the study.

Data Availability

Data sharing is not applicable to this study as no datasets were generated or analyzed during this study.

Authors' Contributions

Conceptualization: MJS, AK, YEK, SK, and EC
Methodology: SK, SS, and YEK
Supervision: MJS
Project administration: MJS
Writing – original draft: AK and EC
Writing – review and editing: AK, EC, YEK, SS, SK, JAS, EL, and MJS

Conflicts of Interest

None declared.

Multimedia Appendix 1
Locations of Korean medicine clinics participating in the registry.
[PPTX File , 2482 KB - resprot_v15i1e87543_app1.pptx]

References

1. The Society of Korean Medicine Obstetrics and Gynecology. Korean Medicine Obstetrics & Gynecology. Seoul: Euisungdang; 2021:462.

XSL•FO
RenderX

2. Martin E, Fisher O, Tone J, Suldsuren N, Kularatna S, Beckmann M, et al. Health-related quality of life and utility of maternity health states amongst post-partum Australians. PLoS One 2024 Oct 7;19(10):e0310913 [FREE Full text] [doi: 10.1371/journal.pone.0310913] [Medline: 39374261]
3. Cooklin AR, Amir LH, Jarman J, Cullinane M, Donath SM, CASTLE Study Team. Maternal physical health symptoms in the first 8 weeks postpartum among primiparous Australian women. Birth 2015 Sep;42(3):254-260. [doi: 10.1111/birt.12168] [Medline: 26088503]
4. Meyling MM, Frieling ME, Vervoort JP, Feijen-de Jong EI, Jansen DE. Health problems experienced by women during the first year postpartum: a systematic review. Eur J Midwifery 2023 Dec 18;7:42 [FREE Full text] [doi: 10.18332/ejm/173417] [Medline: 38111746]
5. Saurel-Cubizolles MJ, Romito P, Lelong N, Ancel P. Women's health after childbirth: a longitudinal study in France and Italy. BJOG 2000 Oct 12;107(10):1202-1209. [doi: 10.1111/j.1471-0528.2000.tb11608.x] [Medline: 11028569]
6. Henderson J, Alderdice F, Redshaw M. Factors associated with maternal postpartum fatigue: an observational study. BMJ Open 2019 Jul 27;9(7):e025927 [FREE Full text] [doi: 10.1136/bmjopen-2018-025927] [Medline: 31352411]
7. Glazener CM, Abdalla M, Stroud P, Naji S, Templeton A, Russell IT. Postnatal maternal morbidity: extent, causes, prevention and treatment. Br J Obstet Gynaecol 1995 Apr 19;102(4):282-287. [doi: 10.1111/j.1471-0528.1995.tb09132.x] [Medline: 7612509]
8. Thompson JF, Roberts CL, Currie M, Ellwood DA. Prevalence and persistence of health problems after childbirth: associations with parity and method of birth. Birth 2002 Jun;29(2):83-94. [doi: 10.1046/j.1523-536x.2002.00167.x] [Medline: 12051189]
9. Bakker M, van der Beek AJ, Hendriksen IJ, Bruinvels DJ, van Poppel MN. Predictive factors of postpartum fatigue: a prospective cohort study among working women. J Psychosom Res 2014 Nov;77(5):385-390. [doi: 10.1016/j.jpsychores.2014.08.013] [Medline: 25214040]
10. Goldszmidt E, Kern R, Chaput A, Macarthur A. The incidence and etiology of postpartum headaches: a prospective cohort study. Can J Anaesth 2005 Nov;52(9):971-977. [doi: 10.1007/BF03022061] [Medline: 16251565]
11. Klein AM, Loder E. Postpartum headache. Int J Obstet Anesth 2010 Oct;19(4):422-430. [doi: 10.1016/j.ijoa.2010.07.009] [Medline: 20833030]
12. Anzola GP, Brighenti R, Cobelli M, Giossi A, Mazzucco S, Olivato S, et al. Postpartum headache: a prospective study. J Oral Facial Pain Headache 2017;31(4):346-352. [doi: 10.11607/ofph.1869] [Medline: 29073663]
13. Dunn G, Egger MJ, Shaw JM, Yang J, Bardsley T, Powers E, et al. Trajectories of lower back, upper back, and pelvic girdle pain during pregnancy and early postpartum in primiparous women. Womens Health (Lond) 2019;15:1745506519842757 [FREE Full text] [doi: 10.1177/1745506519842757] [Medline: 30991911]
14. Lee S, Hwang D, Lee J, Lee C, Jang J. A review of the effects of Korean medicine treatment on postpartum back pain. J Korean Obstet Gynecol 2023;36(1):73-89. [doi: 10.15204/jkobgy.2023.36.1.073]
15. Lee S, Kim S. The efficacy of acupuncture on postpartum urinary incontinence: a systemic review and meta-analysis. J Korean Obstet Gynecol 2022;35(3):37-55. [doi: 10.15204/jkobgy.2022.35.3.037]
16. Handa VL, Zyczynski HM, Burgio KL, Fitzgerald MP, Borello-France D, Janz NK, et al. The impact of fecal and urinary incontinence on quality of life 6 months after childbirth. Am J Obstet Gynecol 2007 Dec;197(6):636.e1-636.e6 [FREE Full text] [doi: 10.1016/j.ajog.2007.08.020] [Medline: 18060960]
17. Gao J, Liu X, Zuo Y, Li X. Risk factors of postpartum stress urinary incontinence in primiparas: what should we care. Medicine (Baltimore) 2021 May 21;100(20):e25796 [FREE Full text] [doi: 10.1097/MD.0000000000025796] [Medline: 34011042]
18. Kroska EB, Stowe ZN. Postpartum depression: identification and treatment in the clinic setting. Obstet Gynecol Clin North Am 2020 Sep;47(3):409-419. [doi: 10.1016/j.ogc.2020.05.001] [Medline: 32762926]
19. Liu X, Wang S, Wang G. Prevalence and risk factors of postpartum depression in women: a systematic review and meta-analysis. J Clin Nurs 2022 Oct;31(19-20):2665-2677. [doi: 10.1111/jocn.16121] [Medline: 34750904]
20. Moore Simas TA, Whelan A, Byatt N. Postpartum depression-new screening recommendations and treatments. JAMA 2023 Dec 19;330(23):2295-2296. [doi: 10.1001/jama.2023.21311] [Medline: 38010647]
21. Fan B, Zhao Y, Chai J, Bing B, Wang W. Effectiveness of acupuncture in postpartum depression: a protocol for an overview of systematic reviews. Medicine (Baltimore) 2022 Aug 12;101(32):e28678 [FREE Full text] [doi: 10.1097/MD.0000000000028678] [Medline: 35960108]
22. Ahmad K, Kabir E, Keramat SA, Khanam R. Maternal health and health-related behaviours and their associations with child health: evidence from an Australian birth cohort. PLoS One 2021 Sep 13;16(9):e0257188 [FREE Full text] [doi: 10.1371/journal.pone.0257188] [Medline: 34516561]
23. Tang L, Lee AH, Binns CW, Hui YV, Yau KK. Consumption of Chinese herbal medicines during pregnancy and postpartum: a prospective cohort study in China. Midwifery 2016 Mar;34:205-210. [doi: 10.1016/j.midw.2015.11.010] [Medline: 26628355]
24. Chuang CH, Chang PJ, Hsieh WS, Tsai YJ, Lin SJ, Chen PC. Chinese herbal medicine use in Taiwan during pregnancy and the postpartum period: a population-based cohort study. Int J Nurs Stud 2009 Jun;46(6):787-795. [doi: 10.1016/j.ijnurstu.2008.12.015] [Medline: 19193377]

XSL•FO
RenderX

25. Kwon NY, Lee HY, Hwang SI, Sung SH, Cho SJ, Yoon YJ, et al. Herbal medicine for postpartum pain: a systematic review of puerperal wind syndrome (Sanhupung). Healthcare (Basel) 2023 Oct 16;11(20):2743 [FREE Full text] [doi: 10.3390/healthcare11202743] [Medline: 37893817]
26. Li Y, Chen Z, Yu N, Yao K, Che Y, Xi Y, et al. Chinese herbal medicine for postpartum depression: a systematic review of randomized controlled trials. Evid Based Complement Alternat Med 2016;2016:5284234 [FREE Full text] [doi: 10.1155/2016/5284234] [Medline: 27774110]
27. Birdee GS, Kemper KJ, Rothman R, Gardiner P. Use of complementary and alternative medicine during pregnancy and the postpartum period: an analysis of the National Health Interview Survey. J Womens Health (Larchmt) 2014 Oct;23(10):824-829 [FREE Full text] [doi: 10.1089/jwh.2013.4568] [Medline: 25268759]
28. Chien LC, Yeh CY, Lee HC, Jasmine Chao H, Shieh MJ, Han BC. Effect of the mother's consumption of traditional Chinese herbs on estimated infant daily intake of lead from breast milk. Sci Total Environ 2006 Mar 01;354(2-3):120-126. [doi: 10.1016/j.scitotenv.2005.01.033] [Medline: 16398988]
29. Liu J, Rong A, Wang F. Meta-analysis of Xiaoyao formula as an adjuvant therapy for treating postpartum depression. Front Psychiatry 2025 Mar 24;16:1558505 [FREE Full text] [doi: 10.3389/fpsyt.2025.1558505] [Medline: 40195969]
30. Hyun MK, Park JK, Oh SY. Satisfaction and perceived effectiveness on herbal decoctions for postpartum care: a cross-sectional survey of mother's experience. J Pharmacopuncture 2023 Jun 30;26(2):175-183 [FREE Full text] [doi: 10.3831/KPI.2023.26.2.175] [Medline: 37405116]
31. Bak JY, Park JK, Park KY, Kim TH. A study about the satisfaction of oriental medical postpartum care. J Korean Obstet Gynecol 2009;22(4):173-186 [FREE Full text]
32. Jeong SY, Lee EH. A retrospective study of mothers who received herbal medicine treatment using the Jeollabuk-do postpartum health support project at one korean medicine hospital. J Korean Obstet Gynecol 2021;34(3):49-64 [FREE Full text]
33. Woo Y. Estimation of adverse events scale relating herbal medicine in Korea. Journal of Society of Preventive Korean Medicine 2020;24(1):27-35. [doi: 10.25153/spkom.2020.24.1.003]
34. Kim M, Han CH. Analysis of herbal-drug-associated adverse drug reactions using data from spontaneous reporting system in electronic medical records. J Korean Med 2015 Mar;36(1):45-60. [doi: 10.13048/jkm.15005]
35. Korean Committee for the Publication of the Yearbook of Korean Medicine. 2022 Yearbook of Traditional Korean Medicine. Daejeon: The Association of Korean Medicine, National Institute for Korean Medicine Development, Korea Institute of Oriental Medicine; 2024:462.
36. The use of the WHO-UMC system for standardised case causality assessment. Uppsala Monitoring Centre. URL: https://who-umc.org/media/164200/who-umc-causality-assessment_new-logo.pdf [accessed 2026-03-09]
37. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz I, Roberts EA, et al. A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther 1981 Aug;30(2):239-245. [doi: 10.1038/clpt.1981.154] [Medline: 7249508]
38. Acharya TA, Trivedi MD, Joshi KJ, Chhaiya SB, Mehta DS. A study of agreement between WHO-UMC causality assessment system and the Naranjo algorithm for causality assessment of adverse drug reactions observed in medical ICU of a tertiary care teaching hospital. Biomed Pharmacol J 2020 Mar;13(1):79-83 [FREE Full text] [doi: 10.13005/bpj/1863]
39. Lessing C, Schmitz A, Albers B, Schrappe M. Impact of sample size on variation of adverse events and preventable adverse events: systematic review on epidemiology and contributing factors. Qual Saf Health Care 2010 Dec;19(6):e24 [FREE Full text] [doi: 10.1136/qshc.2008.031435] [Medline: 20679137]
40. myTrial. NIKOM. URL: https://www.nikom.or.kr/ecrf/service/Login [accessed 2026-03-09]
41. Kim Y, Kim SE, Kim AJ, Kim M, Cho E, Suh S, et al. Incidence and risk factors of adverse drug reactions related with antimicrobials: a single center study. J Pharmacoepidemiol Risk Manage 2021 Mar;13(1):30-37. [doi: 10.56142/2021.13.1.30]
42. Ng CY, Zhao Y, Wang N, Chia KL, Teo CH, Peh W, et al. A multi-center cross-sectional study of Chinese Herbal Medicine-Drug adverse reactions using active surveillance in Singapore's Traditional Chinese Medicine clinics. Chin Med 2024 Mar 07;19(1):44 [FREE Full text] [doi: 10.1186/s13020-024-00915-z] [Medline: 38454483]
43. Lee SH, Song BW, Choi HJ, Kim EY. The analysis of herbal medicine-associated adverse drug reactions spontaneously reported in a Korean medicine hospital. J Korean Soc Health Syst Pharm 2017;34(2):183-199. [doi: 10.32429/jkshp.2017.34.2.005]
44. Safety of decoction forms of herbal medicine in postpartum period: a registry study of the Korean medicine clinics, SAFEHERE-KM. World Health Organization. 2024. URL: https://trialsearch.who.int/?TrialID=KCT0009552 [accessed 2026-03-09]
45. Kim A, Kim YE, Seo S, Kim S, Cho E, Kang YJ, et al. Safety of herbal medicine in the postpartum period of a Korean medicine hospital and postpartum care centre: protocol of a registry study (SAFEHERE-PC). BMJ Open 2024 Aug 24;14(8):e086645 [FREE Full text] [doi: 10.1136/bmjopen-2024-086645] [Medline: 39181559]
46. Choudhury A, Singh PA, Bajwa N, Dash S, Bisht P. Pharmacovigilance of herbal medicines: concerns and future prospects. J Ethnopharmacol 2023 Jun 12;309:116383. [doi: 10.1016/j.jep.2023.116383] [Medline: 36918049]
47. Banerjee AK, Okun S, Edwards IR, Wicks P, Smith MY, Mayall SJ, et al. Patient-reported outcome measures in safety event reporting: PROSPER consortium guidance. Drug Saf 2013 Dec;36(12):1129-1149 [FREE Full text] [doi: 10.1007/s40264-013-0113-z] [Medline: 24092596]

XSL•FO
RenderX

48. Gross AJ, Pisano CE, Khunsriraksakul C, Spratt DE, Park HS, Sun Y, et al. Real-world data: applications and relevance to cancer clinical trials. Semin Radiat Oncol 2023 Oct;33(4):374-385. [doi: 10.1016/j.semradonc.2023.06.003] [Medline: 37684067]
49. Basch E, Jia X, Heller G, Barz A, Sit L, Fruscione M, et al. Adverse symptom event reporting by patients vs clinicians: relationships with clinical outcomes. J Natl Cancer Inst 2009 Dec 02;101(23):1624-1632 [FREE Full text] [doi: 10.1093/jnci/djp386] [Medline: 19920223]

Abbreviations

AE: adverse event
SAFEHERE-KM: Safety of Herbal Medicine Registry in Korean medicine clinics

Edited by A Schwartz; submitted 11.Nov.2025; peer-reviewed by MY Park, Y Choi, HW Suh; comments to author 23.Feb.2026; revised version received 09.Mar.2026; accepted 26.Mar.2026; published 17.Apr.2026.

Please cite as:
Kim A, Cho E, Kim YE, Seo S, Kim S, Shin JA, Lee E, Son MJ
Safety of Postpartum Herbal Medicine in Primary Care Korean Medicine Clinics: Protocol for a Prospective, Multicenter, Registry-Based Observational Study
JMIR Res Protoc 2026;15:e87543
URL: https://www.researchprotocols.org/2026/1/e87543
doi:10.2196/87543
PMID:19503364

©Anna Kim, Eunbyul Cho, Young-Eun Kim, Sumin Seo, Sungha Kim, Jae-Ahn Shin, Eunhee Lee, Mi Ju Son. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Integrating Sperm Microscopy, Environmental Exposures, and Lifestyle Factors for Male Fertility Analysis: Protocol for the Nippon Semen and Environmental Exposure Database (N-SEED) Cross-Sectional Study

Tomoko Oguri[1*], PhD; Kosuke Kojo[2,3,4*], MD, PhD; Yoshiaki Endo[5], PhD; Takaaki Matsuda[3,6], MD, PhD; Haruhiko Midorikawa[7], MD, PhD; Ayumi Nakazono[4], MSc; Kaoru Yanagida[4], MD, PhD; Atsushi Ikeda[2], MD, PhD; Hiroyuki Nishiyama[2], MD, PhD; Isamu Ogura[1], PhD

[1]Research Institute of Science for Safety and Sustainability, National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki, Japan
[2]Department of Urology, Institute of Medicine, University of Tsukuba, Tsukuba, Ibaraki, Japan
[3]Tsukuba Clinical Research & Development Organization, University of Tsukuba, Tsukuba, Ibaraki, Japan
[4]Center for Human Reproduction, International University of Health and Welfare Hospital, Nasushiobara, Tochigi, Japan
[5]Department of Physical Therapy, School of Health Science, International University of Health and Welfare, Otawara, Tochigi, Japan
[6]Department of Endocrinology and Metabolism, University of Tsukuba Hospital, Tsukuba, Ibaraki, Japan
[7]Department of Psychiatry, University of Tsukuba Hospital, Tsukuba, Ibaraki, Japan
[*]these authors contributed equally

Corresponding Author:
Kosuke Kojo, MD, PhD
Department of Urology
Institute of Medicine
University of Tsukuba
1-1-1, Tennodai
Tsukuba, Ibaraki, 305-8575
Japan
Phone: 81 29 853 3223
Fax: 81 29 853 8854
Email: kojou-tuk@md.tsukuba.ac.jp

Abstract

Background: Conventional semen analysis does not fully capture male reproductive potential. The sperm DNA fragmentation index (DFI) may detect latent infertility, although the sperm chromatin structure assay (SCSA) is costly and technically demanding. Image-based analysis of semen microscopy, including artificial intelligence (AI), may enable lower-cost noninvasive assessment. However, progress is limited by a lack of standardized multimodal datasets linking sperm images with the DFI and relevant covariates.

Objective: This protocol describes the initial phase of the Nippon Semen and Environmental Exposure Database (N-SEED) initiative. The study aims to (1) establish standardized acquisition and quality control procedures for sperm microscopy/video and DFI measurement; (2) evaluate the feasibility and magnitude of associations between predefined image-/video-derived variables and the DFI; and (3) characterize clinical, environmental, physical, and lifestyle factors that may act as candidate correlates or confounders for subsequent image-based fertility research.

Methods: This multi-institutional cross-sectional study centralized clinical assessment and specimen collection at a single site in Japan. The prespecified group-specific enrollment targets were 25 for group 1 (a relatively homogeneous low-risk cohort for DFI quality control) and 100 for group 2 (apparently healthy community-based adult volunteers for exploratory association analyses). Microscopic sperm videos, automated semen parameters, and sperm oxidative stress data were obtained at collection. SCSA-based DFI assays are currently being performed sequentially. Group 2 participants additionally provided fasting blood and early-morning urine samples for endocrine, metabolic, environmental chemical, and elemental assays. They also underwent physical measurements and completed structured questionnaires/interviews on lifestyle, nutrition, sleep, psychological status, sexual function, and collection conditions. The primary outcome measure is the continuous DFI assessed with SCSA. Group 1

XSL•FO
RenderX

data will be used to evaluate intraday, interday, and interfacility variability and to fix flow cytometry gating settings. In groups 1 and 2, predefined image-/video-derived summary variables after standardized preprocessing will be analyzed against the continuous DFI using correlation analyses and simple linear regression to assess whether sperm microscopy data contain information relevant to sperm DNA integrity. Exploratory categorical DFI analyses will use simple logistic regression. Secondary exploratory analyses will examine candidate covariates and data completeness, including sensitivity analyses.

Results: The study was registered in the University Hospital Medical Information Network Clinical Trials Registry on January 18, 2026. Between January 18 and February 21, 2026, 124 unique participants were recruited (group 1: n=25; group 2: n=103; both groups: n=4). Sample and data collection were completed on February 21, 2026. Biospecimen assays and primary association analyses are expected to be completed by March 2028.

Conclusions: This initial N-SEED phase will deliver a standardized, quality-controlled multimodal resource linking sperm microscopy/video data with the DFI and broader physiological and environmental measures. Its immediate contribution is not to provide an immediately deployable clinical AI tool but rather to support feasibility assessment, confounder identification, and planning for subsequent database expansion and external validation of image-based male fertility assessment methods.

Trial Registration: UMIN Clinical Trials Registry UMIN000060395; https://tinyurl.com/mryfymw5

International Registered Report Identifier (IRRID): DERR1-10.2196/93803

(JMIR Res Protoc 2026;15:e93803) doi:10.2196/93803

KEYWORDS

computer-assisted image analysis; endocrine disruptors; environmental exposure; erectile dysfunction; healthy volunteers; male infertility; oxidative stress; sleep; sperm motility; testosterone

Introduction

Background

Concern about declining semen quality has grown because meta-analytic studies have suggested substantial reductions in sperm concentration and total sperm count in some populations over recent decades [1,2]. At the same time, temporal inferences remain sensitive to differences in population selection, laboratory methodology, geography, and study design [3,4]. Thus, although adverse trends in semen quality represent an important public health concern, their magnitude and generalizability remain debated. Several interacting explanations have been proposed, including aging and genetic susceptibility, as well as potentially modifiable factors, such as obesity, smoking, alcohol use, psychological stress, reduced physical activity, heat exposure, and environmental chemical exposures, particularly to endocrine-disrupting compounds [5,6]. The public health relevance extends beyond delayed conception alone, because impaired semen quality may also correlate with future cardiometabolic disease, malignancy, hospitalization, and reduced longevity [7-9].

However, population-level concerns about semen quality trends should be distinguished from the clinical problem of assessing fertility potential in individual men. Low sperm concentrations or other abnormal semen parameters increase the probability of subfertility at the population level, but except for severe phenotypes, such as azoospermia, conventional semen analysis alone does not fully capture individual reproductive potential [10,11]. This limitation has increased interest in sperm DNA fragmentation (SDF), a form of sperm genomic damage, and in the DNA fragmentation index (DFI), a quantitative indicator of sperm chromatin integrity. The DFI is commonly defined as the proportion of sperm with fragmented or acid-denaturable DNA [12,13]. It is typically quantified using the sperm chromatin structure assay (SCSA), which involves acid-induced DNA denaturation, acridine orange staining, and flow cytometry [14,15]. A higher DFI has been associated with prolonged time to pregnancy and less favorable assisted reproductive technology outcomes in some settings, although thresholds and clinical indications remain context dependent [15-17].

Despite its potential clinical relevance, SCSA-based DFI assessment is more technically demanding than routine semen analysis and has limited scalability for broader screening or repeated measurement [10,13,15]. This creates interest in lower-cost, noninvasive approaches that could extract information relevant to sperm DNA integrity from standard semen microscopy.

Prior Work

Recent artificial intelligence (AI) and machine learning (ML) applications in andrology have mainly focused on relatively narrow input domains, such as sperm images/videos, semen parameters, or single-modality clinical variables. Reviews consistently emphasize the need for larger standardized datasets, multimodal integration, and external validation before clinical adoption [18-20]. Progress is being made in attempts to predict the DFI using deep learning based on microscopic semen images and video data obtained through relatively inexpensive and noninvasive examination [21,22]. However, practical AI development remains constrained by the lack of high-quality datasets that directly link standardized sperm microscopy with the DFI.

The functional quality of sperm is influenced not only by testicular factors but also by multidimensional behavioral, physiological, and environmental influences. These include the abstinence period [23]; the ejaculation frequency [24]; collection-context factors, such as the duration of sexual arousal [25] and intensity of orgasm [26]; environmental chemical exposures [27,28]; and lifestyle habits [29]. Accordingly, the broader physiological and environmental context has been

XSL•FO
RenderX

underrepresented in many existing AI-oriented datasets and study designs.

Attempts have also been made to predict male fertility using AI analysis of large-scale databases of blood endocrine profiles related to testicular function, independent of semen findings [30]. Although such endogenous physiological factors may be informative, many existing large-scale datasets rely on retrospective observations in patients with infertility, where diverse pathologies limit adjustment for confounding. In contrast, prospective data collection in non-clinic-based community populations may reduce the influence of disease factors that directly impair fertility. However, clinical studies involving men are relatively rare [31], and when semen analysis is included, volunteer bias can become substantial [32]. Consequently, defining truly healthy reference values and exploring candidate risk factors remain difficult.

ML has been used to predict semen findings from lifestyle habits [33-35]. Similarly, AI models have been developed to automatically analyze semen parameters from sperm images [36-38]. In parallel, systematic reviews on the effects of environmental chemicals and air pollution on semen findings have been published in recent years [39,40]. However, to the best of our knowledge, no dataset integrating sperm images with comprehensive environmental, physiological, and lifestyle data to evaluate association with the DFI, and to support future predictive modeling, in a non-clinic-based general population where the influence of overt disease factors is minimized has been established.

Study Objectives

This protocol paper describes the initial phase of the Nippon Semen and Environmental Exposure Database (N-SEED) initiative as a multicenter cross-sectional study. N-SEED is designed as a long-term, scalable database infrastructure and is not intended to be completed in a single year. As a preliminary step, the 2026 project prospectively collected biospecimens (urine, blood, and semen) and related data (semen analysis, physical measurements, and questionnaire/interview data) from apparently healthy community-based adult men. These data will be used to statistically evaluate the associations between sperm microscopy image/video data and sperm functional assessment measures, including the DFI. The ultimate long-term goal is to support the development and validation of image-based male fertility assessment technologies. This initial phase focuses on data collection feasibility, measurement quality control, and analytical readiness. Accordingly, this initial phase is not intended to develop or externally validate a clinically deployable AI model. Rather, it is intended to establish standardized acquisition and quality control procedures; estimate the magnitude and variability of image-DFI associations; identify candidate confounders and features; and inform formal sample-size planning, model development, and independent validation strategies for later phases. Specifically, we will conduct a preliminary assessment focusing on three key aspects:

- First, we will evaluate interfacility variability and perform intrafacility quality control of the DFI testing method. In group 1, which is designed primarily for quality control of DFI testing, we will assess the reliability of the measurement system.
- Second, we will prospectively collect standardized microscopic sperm videos and related multimodal data. Initial association analyses between sperm image/video data and the DFI will then be conducted to evaluate feasibility and inform subsequent study design (groups 1 and 2).
- Finally, we will explore multilayered factors potentially related to the DFI and sperm image/video data, including standard semen analysis parameters, endocrine and metabolic biomarkers, environmental chemicals, physical measurements, and lifestyle factors, to support confounding adjustment and hypothesis generation in future phases (group 2).

Methods

Study Design

The study has been registered in the University Hospital Medical Information Network Clinical Trials Registry (UMIN-CTR; #UMIN000060395; registration date: January 18, 2026; disclosure date: January 19, 2026).

This is a multi-institutional cross-sectional study with clinical data collection centralized at a single site to ensure high-quality standardization. One of the aims is to construct a foundational database capable of supporting future methodological development (including potential AI approaches), external validation, and generation of new hypotheses. To this end, we prospectively and systematically collected multimodal information (semen analysis, physical measurements, and questionnaires) centered on microscopic sperm image/video data obtained from apparently healthy community-based adult men, together with biological samples (semen, blood, and urine).

A characteristic of this study is the use of two participant groups with different inclusion criteria according to the analytic purpose. Participants in group 1 were recruited to ensure the reproducibility and stability of the DFI testing method. By minimizing confounding factors that may adversely affect sperm quality, such as aging, obesity, underweight, and smoking, and by limiting the group to men with a high ejaculation frequency, we aimed to identify a population expected to maintain high spermatogenic function. In this group, only semen was collected as a biological sample. In contrast, participants in group 2 were included to ensure broader representation of potential risk factors associated with a worsening DFI in otherwise apparently healthy community-based volunteers. In this group, in addition to semen, blood and urine samples were collected to examine the effects of nutritional and endocrine profiles and environmental exposure.

Study Setting

The National Institute of Advanced Industrial Science and Technology (AIST, Tsukuba, Japan) serves as the central hub for data management and analysis in this study. AIST is responsible for planning research, obtaining ethics committee approval, receiving and storing samples and data, designing and constructing the database, overseeing clinical testing and

XSL•FO
RenderX

chemical analysis (including external outsourcing), and coordinating statistical analyses and subsequent methodological development as the database expands.

Research tasks, such as reviewing the research plan, recruiting participants, obtaining informed consent at the visit, collecting interview and questionnaire data and physical measurements, collecting biological samples, acquiring general semen analysis parameters (including sperm concentration and motility), and acquiring sperm microscopic videos, were conducted at the Osaka Pharmacology Clinical Research Hospital (Osaka, Japan). This hospital conducts clinical research as a single third-party institution following standardized procedures. In addition, expert supervision in reproductive medicine, endocrinology, psychiatry, analytical chemistry, physiotherapy, epidemiology, and hygiene, as well as specific measurement items (eg, environmental chemical measurements), is conducted in collaboration with multiple universities and research institutions.

The recruitment and registration period ran from January 18 to February 21, 2026. Data and biological sample collection was completed during visits in this period. No follow-up is planned because of the cross-sectional design. Biological sample analyses will be performed sequentially after collection and are scheduled to continue until March 2028, which marks the end of the study period.

Participants

Eligibility Criteria

The inclusion criteria for the participants are presented in Table 1. In group 1, the eligibility criteria were designed to minimize factors adversely affecting spermatogenic function for quality control of DFI measurement values. Specifically, group 1 included only men who were sexually active young adults (18-24 years old), had a high frequency of ejaculation (average of ≥10 per month), had a standard body type (BMI 18.5 to <25 kg/m²), and had never smoked. Group 2 had broader inclusion criteria to explore DFI risk factors in men from diverse backgrounds. The target population included men of general reproductive age (18-41 years) with minimal ejaculatory function (average of ≥1 time per month). However, to ensure that the blood and urine test results were accurate, compliance with visits during the designated time (8:00 to 11:00 A.M.), fasting the previous night, and a two-visit schedule was mandatory for this group.

Table 1. Inclusion criteria for groups 1 and 2.

Criteria	Group 1 (quality control)	Group 2 (exploratory)
Age (years)	18 to <25 years	18 to <42 years
BMI (kg/m²)	18.5 to <25.0	No specific restriction
Ejaculation frequency (times/month; average over the past 6 months)	≥10	≥1
Smoking status	Never smoked [41][a]	No specific restriction
Medical history	No history of congenital anomalies, male reproductive organ diseases/trauma, varicocele, malignant tumors, endocrine disorders, or autoimmune diseases	Same as group 1
Medication	No history of hormone medication use [42-45][b]	Same as group 1
Study compliance	Able to comply with 48-144 h of abstinence	Able to comply with 48-144 h of abstinence, visit between 8:00 and 11:00 A.M. (fasting condition), and attend two visits

[a]Lifetime consumption<100 cigarettes, no smoking in the past 6 months, and no smoking during the study period [41].

[b]Hormone medication and other drug exclusion criteria: participants who had used thyroid hormone preparations [42] or 5α-reductase inhibitors (eg, finasteride, dutasteride) [43] within the past 6 months were excluded. Minoxidil use was permitted [44]. Corticosteroids were generally permitted, except when used for self-reported autoimmune diseases [45]; however, topical or inhaled corticosteroids used for conditions such as atopic dermatitis or asthma were acceptable. Other medications were not exclusionary; current medication use (including that of drugs that may affect ejaculation or sexual function) was collected in detail via questionnaires/interviews and will be considered in sensitivity analyses, as appropriate.

The exclusion criteria were identical for both groups. We excluded individuals with a history of congenital anomalies or diseases of the male reproductive organs, varicocele, malignant tumors, endocrine diseases, or a history of hormone medication use. Additionally, individuals who might experience unjust disadvantages from participating, those who had difficulty providing consent because of cognitive impairment, or those judged by the study physician to be in poor physical condition on the day of the survey (eg., acute illness symptoms) were considered ineligible. If a participant in group 1 also met the requirements of group 2 and wished to participate, co-enrollment in both groups was permitted. Such cases were identifiable through study registration records. Because groups 1 and 2 were established for different purposes rather than as formal independent comparison arms, co-enrolled records were retained as group-specific records but were not treated as statistically independent in any cross-group descriptive comparison or pooled participant-level analysis.

Source Population and Recruitment

To reduce clinic-based sampling bias (ie, overrepresentation of men seeking fertility evaluation), we recruited men residing in Japan who were registered in a volunteer panel managed by InCROM Co, Ltd (Nishiwaki, Japan) rather than through

XSL•FO
RenderX

infertility clinics. After eligibility confirmation, participants were referred to the Osaka Pharmacology Clinical Research Hospital, a clinical research implementation facility, where written informed consent was obtained. This study concerns semen-based measures of biological male reproductive function; gender identity and sexual orientation were not assessed, and sex/gender was not analyzed as a separate covariate in this initial phase. Given that study participation required morning fasting visits and (for group 2) two visits, we anticipated potential overrepresentation of individuals with greater scheduling flexibility. To characterize this quantitatively, age and recruitment-related sociodemographic variables (eg, employment status, shift work, education, and income) were systematically collected. Their distributions, including "prefer not to answer" responses where applicable, will be reported.

Sample Size

Considering biological sample processing, assay capacity, and the budget in this initial phase, the target group-specific enrollment was set at 25 for group 1 and 100 for group 2. Because co-enrollment was permitted, these targets refer to group-specific enrollment records; therefore, the total number of unique individuals across both groups may be smaller than the sum of the two group-specific counts. The two groups address different objectives. Group 1 is intended for quality control and descriptive characterization of the DFI measurement system in a relatively homogeneous low-risk population, rather than formal hypothesis testing. Group 2 is intended for feasibility-oriented initial association analyses between sperm image/video data and the DFI and exploratory analyses of candidate factors. Therefore, the sample size was determined pragmatically rather than via a formal power calculation for a specific effect size. We acknowledge that this initial phase may be underpowered to detect small associations or to support multivariable models with many predictors. Accordingly, effect estimates, CIs, and consistency of direction will be emphasized, and the findings will be used to inform variable prioritization and formal sample size planning for subsequent phases.

Variables and Data Collection

A comprehensive dataset was collected to construct a multimodal database linking sperm microscopy image/video data with semen examination results, the DFI, and extensive clinical, environmental, and lifestyle covariates. The primary purpose of this initial phase is to establish standardized data collection and enable association analyses and hypothesis generation. The dataset will also provide a foundation for subsequent phases, including stepwise development and external validation of image-based analytical approaches (including AI) as the database expands.

Most participant-reported data (baseline characteristics, lifestyle factors, psychological status, sexual function, and collection-related subjective assessments) were obtained using paper-based structured self-administered questionnaires in Japanese. Study staff checked completeness with each participant at the visit. Physician-conducted interviews were limited to confirmation of eligibility (including poor physical condition on the day of the visit) and clarification of clinically complex medical history, reproductive/urological history, infertility treatment history, and medication use that may have been difficult to classify accurately through self-report alone. These interviews followed a prespecified oral checklist aligned with the eligibility criteria. For sensitive socioeconomic and sexual/reproductive items, participants could choose "prefer not to answer."

Baseline Characteristics and Environmental Exposures

Data on baseline characteristics were collected through these structured self-administered questionnaires and through physician-conducted interviews for clinically complex eligibility items. Demographic and socioeconomic factors included age, height, weight, educational attainment, marital status, and household income [46]. Regarding medical background, current and past medical histories were obtained. Regarding current medications, we verified the use of treatments for androgenetic alopecia, drugs for erectile dysfunction, and zinc supplementation. Furthermore, as part of the reproductive background, the partner's pregnancy history, history of infertility treatment [47,48], the participants' desire for children [49], and experience using home sperm-testing kits [50] were investigated.

Regarding environmental and occupational exposures, employment status, including shift work, was assessed [51]. Work in high-temperature environments and the frequency of laptop use on the lap were also assessed to evaluate potential thermal effects on spermatogenesis [52].

Lifestyle Factors, Psychological Status, and Sexual Function

Lifestyle factors, psychological status, and sexual function were evaluated multidimensionally using structured self-administered questionnaires, incorporating validated scales and specific self-developed questions. Where available, validated Japanese versions or established Japanese translations of standard instruments were used. Regarding lifestyle habits, data on smoking habits (Brinkman index; daily number of cigarettes × years) [53], alcohol consumption (Alcohol Use Disorders Identification Test–Consumption [AUDIT-C]) [54], and sauna/bathing habits (self-developed questionnaire) were collected. Nutritional intake was assessed using the Brief-type Self-Administered Diet History Questionnaire (BDHQ) [55]. Sleep quality was evaluated using the Japanese version of the Pittsburgh Sleep Quality Index (PSQI-J) [56], chronotype using the reduced Morningness-Eveningness Questionnaire (rMEQ) [57], and physical activity levels using the International Physical Activity Questionnaire (IPAQ) [58].

Psychological status was assessed using the Perceived Stress Scale (PSS) [59] and the Kessler Psychological Distress Scale (K6) [60]. Sexual function and urological symptoms were comprehensively evaluated using the Sexual Health Inventory for Men (SHIM) [61,62], a modified SHIM replacing "sexual intercourse" with "masturbation" in questions 3-5 [63]; the Erection Hardness Score (EHS) [64,65]; the Male Sexual Health Questionnaire–Ejaculatory Dysfunction–Short Form (MSHQ-EjD-SF) [66,67]; and a modified version of the first question of the MSHQ-EjD-SF, where "sexual intercourse" was replaced with "masturbation" (self-developed modification). The self-estimated intravaginal ejaculation latency time [68]

XSL•FO
RenderX

and the self-estimated ejaculation latency time during masturbation (self-developed question) were additionally assessed. The frequencies of sexual intercourse, masturbation, nocturnal erection, and feelings of sexual desire were evaluated using point scales similar to those used in the 2023 National Survey in Japan [69].

Furthermore, the Aging Males' Symptoms Scale (AMS) score [70-72] was used as an indicator of physical and psychological changes associated with androgen decline. The National Institutes of Health Chronic Prostatitis Symptom Index (NIH-CPSI) [73] was used to evaluate chronic prostatitis symptoms. These measures were included to support exploration of potential infertility risk factors and hypothesis generation.

Published studies support the reliability/validity or relative validity of the BDHQ [55], PSQI-J [56], rMEQ [57], IPAQ [58], AUDIT-C [54], PSS [59], K6 [60], AMS [71], and NIH-CPSI [73] in Japanese populations. A Japanese version of the SHIM [62] and MSHQ-EjD-SF [67] were used. The EHS is a single-item measure, for which internal consistency is not applicable. Modified or self-developed items were treated as exploratory measures.

Physical Measurements

For group 2 participants, we conducted detailed physical function evaluations and physical measurements based on standardized procedures. The weight, BMI, body fat percentage, muscle mass, and segmental muscle mass of the limbs and trunk were measured using a multifrequency body composition analyzer (MC-780A, TANITA, Tokyo, Japan). The skeletal muscle mass index [74], fat mass index [75,76], phase angle [77,78], and extracellular water–to–total body water ratio [79] were also measured. These indicators were collected to evaluate the detailed body composition, including regional characteristics and cellular status, in addition to indicators such as BMI and body fat percentage commonly used in previous studies [80,81]. Measurements were performed with participants barefoot and wearing light clothing, with corrections made for the weight of the clothing.

Subsequently, muscle strength was evaluated as a physical indicator of androgen status and whole-body muscle strength. In addition to grip strength [82,83], a widely used representative indicator of whole-body muscle strength and physical function, back muscle strength [84] was also measured for a more comprehensive evaluation of muscle strength status, including that of the trunk. Digital grip strength (TAKEI 5401, SANKA Co, Ltd, Health Equipment Division, Tokyo, Japan) and back muscle strength meters (TAKEI 54021, SANKA Co, Ltd) were used. Grip strength was measured twice for each hand alternately in the standing position, and the maximum value was adopted [85]. Back muscle strength was measured twice, with the upper body inclined forward at 30° and the knees extended (but not locked), and the maximum value was adopted [86].

Anatomical circumferences were measured based on standardized procedures using a nonstretchable measuring tape (SECA 201, Hamburg, Germany). Specifically, the waist circumference was measured at the midpoint between the lower rib margin and the iliac crest [87]. The maximum circumferences of the hip [88], thigh [89], and calf [90] were also measured. These circumferences were collected exploratorily for a simple evaluation of the regional body shape and fat/muscle distribution, which could not be captured using whole-body composition indicators. For quality control, each site was measured twice, and if the difference was ≥1 cm, a third measurement was performed. Sexual function may be associated with complex physical factors, such as systemic metabolic status, muscle mass/strength, fluid distribution, and physical activity. Thus, data on body composition, muscle strength, and morphological indicators were comprehensively collected to evaluate the whole-body physical status multidimensionally.

Semen Analysis and Subjective Assessment

Following an ejaculatory abstinence period of 48 to <144 h (verified upon arrival) [91], semen was collected via masturbation in a designated room at the clinical research facility and analyzed in an onsite laboratory. The samples were liquefied in a 37°C incubator or at room temperature, with a maximum liquefaction time of 60 minutes. Incomplete liquefaction was defined as a thread length of ≥2 cm during pipetting, in which case mechanical agitation using an 18-gauge needle was performed. Semen volume was calculated by weighing the collection container, assuming a semen density of 1.0 g/mL.

Semen analysis was performed using an automated semen quality analysis system (LensHooke X3 PRO, Bonraybio, Taichung, Taiwan) [92]. Sperm concentration and total motility were measured in accordance with the 6th Edition of the *World Health Organization (WHO) Manual for Human Semen Analysis* [93,94]. The following specific kinematic parameters were also analyzed: categorization of sperm motility (fast progressively motile, slow progressively motile, nonprogressively motile, and immotile), average path velocity, straight line velocity, curvilinear velocity, amplitude of lateral head displacement, beat cross frequency, linearity, straightness, and wobble [95]. In addition, the percentage of normal morphology, pH, and round cell concentration were measured using an automated image recognition system. To evaluate sperm oxidative stress, the static oxidation-reduction potential was measured using a sperm oxidative stress measurement system (MiOXSYS, Caerus Biotech, Vilnius, Lithuania) [96]. Microscopic videos generated by the system were stored in native output format, and file metadata (eg, frame rate and resolution) were retained for subsequent image-based analyses.

The participants additionally completed a Subjective Assessment at Collection questionnaire immediately after semen collection. The questionnaire recorded the abstinence period, presence of spillage, time taken for collection on the study day, subjective intensity of orgasm on the study day (4-point Likert scale) [97], subjective semen volume compared to usual (modified question 3 of the MSHQ-SF), subjective time taken for collection compared to usual (self-developed question), and stress level during collection (10-point Likert scale, self-developed question). This was intended to aid in assessing the influence of psychological and situational factors on sample quality.

Residual samples after analysis were preserved as smears for Diff-Quik staining. Residual semen for downstream exploratory analyses was aliquoted according to the volume available after

XSL•FO
RenderX

onsite testing and stored at −80°C before shipment to collaborating institutions or external laboratories. The smears are being morphologically evaluated based on Kruger's strict criteria [98]. Depending on residual sample availability, frozen semen aliquots are being used for exploratory analyses of the sperm DFI; high DNA stainability (HDS) using SCSA [99-101]; trace and major elements and chemical speciation elements in semen and seminal plasma; and additional biomarkers, such as inhibin B and insulin-like 3 (INSL3).

Blood and Urine Assays

Group 2 participants provided early-morning urine and fasting blood samples. Fasting was defined as no breakfast on the study day, and the participants were instructed to finish dinner by approximately 9:00 P.M. on the day before the visit. Water intake was permitted. To account for diurnal fluctuations in analytes, such as testosterone [102] and zinc [103] levels, blood samples were collected between 8:00 and 11:00 A.M. To avoid potential alterations in parameters such as postejaculation prolactin, blood sampling was performed before semen collection [104]. A portion of the blood was processed into plasma or serum and stored frozen at −80°C; these samples were shipped to collaborating institutions or external laboratories for exploratory analysis.

Urine assays include measurements of exposure to environmental chemicals (phthalate esters, parabens, and pesticides), soy isoflavones, trace and major elements, chemical speciation of elements, creatinine, specific gravity, and urea nitrogen. Blood assays include measurements of serum and plasma gonadotropins and sex steroid profiles (follicle-stimulating hormone, luteinizing hormone, testosterone, free testosterone, dehydroepiandrosterone sulfate, sex hormone–binding globulin, dihydrotestosterone, estradiol, prolactin, anti-Müllerian hormone, inhibin B, and INSL3), metabolic markers (serum albumin, fasting blood glucose, and insulin), and selected exposure substances (per- and polyfluoroalkyl substances, trace and major elements, and chemical speciation of elements) in whole blood, serum, and plasma.

Definition of Outcome Measures

The primary outcome measure is the continuous DFI. The DFI refers to the proportion of sperm with fragmented or acid-denaturable DNA. In this study, the DFI is measured using SCSA after acid treatment and acridine orange staining with flow-cytometric analysis [12-14]. The primary analysis will assess whether predefined microscopic sperm image-/video-derived summary variables are associated with a continuous DFI, as a feasibility-oriented proof-of-concept evaluation of whether sperm microscopy data contain information relevant to sperm DNA integrity. Secondary and exploratory outcome measures include a categorical DFI based on literature-informed or study-specific exploratory thresholds, HDS, and general semen analysis parameters. Clinical, environmental, physical, and lifestyle variables will be examined as candidate correlates or confounders in secondary exploratory analyses [105].

Statistical Analysis

Descriptive Statistics and Data Preprocessing

Continuous variables will be summarized as means (SDs) and/or medians (IQRs), depending on the distribution characteristics, and categorical variables will be summarized as frequencies (percentages) [106]. The DFI will be treated as a continuous variable. Categorization using literature-informed thresholds (eg, 15% and 27%) [107-110] will also be used in secondary exploratory analyses. In addition, if data characteristics permit, a study-specific candidate threshold will be explored using receiver operating characteristic (ROC) analysis against a prespecified binary semen quality classification based on the lower fifth percentile values for basic semen parameters described in the 6th Edition of the *WHO Manual for Human Semen Analysis* [94,111,112]. Any such data-driven threshold will be considered hypothesis generating only [113].

No formal between-group inferential comparison between groups 1 and 2 is a primary objective of this study. Any comparison of baseline characteristics between the two groups will therefore be descriptive/exploratory only. If participant overlap between groups is present, co-enrolled individuals will be counted only once in such cross-group summaries or in any pooled participant-level analysis requiring independence. Under the primary handling rule, the group 1 record will be retained and the corresponding group 2 record excluded, because group 1 had stricter eligibility criteria and a quality control purpose. Alternative handling rules, such as retaining the group 2 record or excluding co-enrolled participants entirely, may be examined as sensitivity analyses. This issue does not affect the main group-specific analyses, in which group 1 is used for DFI quality control and group 2 for exploratory association analyses.

All statistical tests will be two-sided and interpreted using a nominal significance level of $P<.05$. However, the secondary analyses are exploratory and involve multiple related comparisons and sensitivity analyses. Thus, emphasis will be placed on effect sizes, 95% CIs, and consistency of direction rather than on statistical significance alone, and no formal multiplicity adjustment is planned. As exposure indices (eg, environmental chemicals) may generally exhibit skewness, transformations (eg, logarithmic transformations) will be considered according to the distribution [114]. For urinary indices, a correction using urinary creatinine or specific gravity will be applied [115,116].

Quality Control Analysis

Data from group 1 will be used for quality control of the DFI measurement system. Specifically, using SCSA for DFI and HDS measurements, we will evaluate intraday, interday, and interfacility variability using quality control semen samples. Given that SCSA-derived DFI values can be influenced by the assay protocol, instrument settings, and gating strategy, this quality control step is required before the main association analyses [14,117,118]. After evaluating measurement validity, the flow cytometry gating settings will be fixed to ensure measurement reliability. Using data from group 1, we will confirm the distribution in young, healthy men; exclude outliers; and describe the reference value range.

XSL•FO
RenderX

Image Information Analysis

For the primary analysis, predefined microscopic sperm image-/video-derived summary variables will be assessed for association with the continuous DFI (primary outcome measure) as a preliminary proof-of-concept evaluation of whether image/video data contain information relevant to sperm DNA integrity. Associations with the continuous DFI will initially be assessed using correlation analyses and simple linear regression. When the DFI is categorized using prespecified literature-informed thresholds or a study-specific exploratory candidate threshold, simple logistic regression will be used as a secondary exploratory analysis. Any predictive modeling (including AI approaches) will be considered exploratory and will be specified in subsequent phases after database expansion. Exploratory image-DFI association analyses will be reported separately by group. If any additional pooled participant-level analysis across groups 1 and 2 is performed, the nonoverlapping dataset described earlier will be used.

Exploratory Analysis

For group 2 data, associations among the DFI, microscopic sperm image/video data, and clinical variables will be analyzed to identify candidate predictors and confounding factors for image-based analyses. Core analyses will prioritize standard semen parameters, questionnaire data, and physical measurements. The DFI will be analyzed primarily as a continuous outcome, with secondary analyses using thresholds [107-110]. Building on the reference values established in the quality control analysis, we may explore a study-specific DFI threshold against the WHO criteria [93] using ROC analysis. The area under the ROC curve and Youden index–derived cut points will be reported as hypothesis-generating results.

Candidate variables will first be examined using correlation analyses, group comparisons, or simple linear/logistic regression, as appropriate to variable type. Age and recruitment-related participant characteristics (eg, employment status, shift work, education, and income) will also be examined as candidate covariates or stratification factors to evaluate potential selection-related bias. The data of participants selecting “prefer not to answer” for education or income will be retained in descriptive analyses, and sensitivity analyses will compare alternative handling of these responses (eg, separate-category vs exclusion approaches). Given that a population-based sampling frame is not available, these analyses are intended to assess robustness rather than fully remove selection bias. Multivariable linear or logistic regression models will be considered only when data completeness and the number of informative observations is adequate. To reduce overfitting, the number of predictors simultaneously included will be restricted, with priority given to variables selected based on prior knowledge and univariable results. Multicollinearity will be assessed using the variance inflation factor, with values >3 considered to indicate potential concern. In such cases, we will preferentially retain representative variables or use dimension reduction, clustering, or mixed-exposure approaches as exploratory analyses.

Missing Data and Sensitivity Analysis

Participants with missing data for primary outcome measure analysis (DFI or images) will be excluded from the primary analysis, and the reasons for missingness will be described. Considering that the status of questionnaires or samples, such as self-reports of spilling all or part of the sample during collection [119] or recent infectious diseases (eg, a cold, influenza, or COVID-19) with or without a high fever (≥38°C) within the past 3 months [120,121], may affect the results, sensitivity analyses excluding such participants will be conducted, as necessary. If azoospermic samples are observed, sensitivity analyses will be performed both including and excluding them because their underlying etiologies cannot be fully characterized in this non-clinic-based setting [122].

Ethical Considerations

The study protocol underwent central ethical review and was approved by the Institutional Review Board (IRB) of the Osaka Pharmacology Clinical Research Hospital (IRB approval #R01114a; committee review date: December 18, 2025; notification date: December 19, 2025) in accordance with the Ethical Guidelines for Medical and Biological Research Involving Human Subjects in Japan. AIST, which leads protocol development and overall study governance, does not accept reliance on another institution’s review. Hence, independent approval was additionally obtained from the IRB of AIST (IRB approval #Life 2025-0618; review completion date: January 6, 2026; notification date: January 7, 2026). Based on the central review, study implementation or institutional approval was also obtained from the University of Tsukuba Hospital (approval #R07-275), the Tsukuba Gakuen Hospital (approval #25-23), the International University of Health and Welfare (approval #25-MB-18), the Nagoya University Graduate School of Medicine (approval #2025-0536; commission #2025-184), and the National Institute for Environmental Studies (approval #2025-024).

The study will be conducted in accordance with the Declaration of Helsinki and Ethical Guidelines for Medical and Biological Research Involving Human Subjects in Japan. Written informed consent was obtained from all participants after a full explanation of the study objectives, procedures, potential benefits, and risks. Questionnaire data and biospecimens were deidentified using a study identification code. The linkage key is retained only at the Osaka Pharmacology Clinical Research Hospital, and collaborating institutions and external laboratories receive coded data/samples only. Sensitive sexual and reproductive information was collected primarily through self-administered questionnaires, where participants were allowed to select a “prefer not to answer” option to ensure their autonomy and minimize psychological distress, consistent with a recent study [123]. The participants received reimbursement for study-related expenses to reduce the burden of transportation and time commitment rather than as an incentive for participation: JPY 10,000 (~US $63) for group 1 and JPY 15,000 (~US $94) for group 2 (JPY 5000, ~US $31, on the explanation day and JPY 10,000, ~US $63, on the study day). Individual research results, including incidental findings, are not routinely returned because the assays were conducted for research

XSL•FO
RenderX

purposes and may not be clinically interpretable at the individual level. This policy was explained during the informed consent process.

Results

Participant recruitment began on January 18, 2026, and recruitment and sample/data collection were completed on February 21, 2026. Final enrollment yielded 25 group 1 records and 103 group 2 records. In addition, 4 individuals contributed records to both groups; thus, the study comprises 124 unique participants overall. For any cross-group descriptive comparison or pooled participant-level analysis requiring independence, these co-enrolled cases will be handled using the prespecified de-duplication rule described in the *Statistical Analysis* section. The last participant visit occurred on February 21, 2026. Biological samples will be sequentially analyzed after collection, and exploratory assays are planned to continue throughout the study period until March 2028. The findings of this study will be disseminated through domestic and international academic conferences and journal publications. After completion of biospecimen assays, the primary association analysis (between sperm image/video data and the DFI) and exploratory analyses of secondary outcome measures will be conducted, with primary analyses planned to be completed by March 2028.

Discussion

Summary

The principal contribution of this initial phase of N-SEED is the establishment of a standardized methodology for constructing a prospective multimodal database that links sperm microscopy image/video data with the DFI and a wide range of environmental, physiological, and lifestyle covariates. Compared to much of the existing image analysis literature on male infertility, this protocol incorporates environmental chemicals, endocrine profiles, and lifestyle-related measures, alongside semen microscopy [18-20]. In this initial phase, the resulting resource is intended to support feasibility-oriented association analyses, confounder assessment, and hypothesis generation. It is not intended to provide a clinically deployable AI tool or definitive evidence regarding predictive performance.

Comparison With Prior Work

Most current AI applications in male infertility have focused on relatively narrow tasks, such as sperm morphology, motility, sperm retrieval, or in vitro fertilization-related prediction. Recent reviews consistently note methodological heterogeneity, limited multimodal integration, and insufficient multicenter external validation as major barriers to clinical implementation [18-20]. More broadly, the literature on semen quality has also been affected by heterogeneous source populations, retrospective designs, and differences in laboratory methods, which complicate interpretation across studies [2,4].

Overall, these prior studies indicate two relevant gaps. First, datasets directly linking standardized sperm microscopy image/video data with a functional marker of sperm DNA integrity, such as the DFI, are limited [18-20]. Second, environmental, endocrine, physical, and lifestyle variables that may confound or contextualize image-based associations are often incompletely represented in existing datasets [18-20].

The protocol is specifically designed to address these methodological gaps by prospectively collecting standardized sperm microscopy image/video data, together with the SCSA-based DFI and broader covariates in apparently healthy community volunteers. By recruiting outside infertility clinics, the study attempts to reduce, although not eliminate, the disease-related selection structure that characterizes many retrospective clinic-based datasets. Accordingly, the immediate contribution of this initial phase is methodological: to establish standardized acquisition procedures, characterize data completeness and measurement variability, and support feasibility-oriented analyses of associations between image-/video-derived variables and the DFI. Questions of model development, threshold validation, and external generalizability remain for later expanded phases.

Strengths and Limitations

The strength of this study lies in the establishment of two cohorts according to specific objectives. Group 1 was recruited to secure reliable reference measurements of the DFI through rigorous quality control and by minimizing major confounding factors. Group 2, which included participants from diverse backgrounds, was recruited to explore realistic risk factors. Another strength is the quantification of subjective ejaculation sensations and collection conditions to consider psychological and situational factors.

However, the study also has some limitations. The study has a cross-sectional design, which means that separate longitudinal studies are necessary to assess temporality and to validate any future image-based markers against fertility outcomes. In addition, single-point urine and blood samples may not fully capture long-term exposure trends. Another inherent limitation is the potential selection bias introduced due to volunteer panel recruitment and operational requirement for morning fasting visits (and two visits for group 2), which may preferentially include men with greater scheduling flexibility. Similar patterns have been reported in healthy-volunteer research, where participants can be disproportionately students or individuals with limited economic independence in some contexts [96,97]. Accordingly, this initial phase should be interpreted as a feasibility-oriented volunteer sample for association analyses and hypothesis generation rather than a population-representative prevalence study. Residual bias may remain because some socioeconomic variables permit nonresponse.

Moreover, the initial phase sample size was determined pragmatically to establish standardized data collection rather than to detect a prespecified small effect size. Thus, null or unstable findings, especially in subgroup or multivariable analyses, will require cautious interpretation and should primarily inform future sample size calculations and variable prioritization. Any study-specific DFI threshold derived from the dataset will also require validation in an independent dataset before it can be interpreted as a clinically useful decision threshold. Concurrently, the pragmatic scale of this initial phase and the centralized single-country collection design mean that

XSL•FO
RenderX

this cohort should be regarded as a high-fidelity pilot resource rather than a definitive etiologic or model validation dataset. Although it may identify candidate image-DFI relationships and relevant confounders, it cannot establish small exposure effects, definitive DFI thresholds, or broad algorithmic generalizability without later expansion and independent external validation [19].

However, the primary purpose of this study is to verify feasibility and conduct proof-of-concept association analyses, thereby serving as a foundation for future large-scale research. This protocol does not describe a tool ready for immediate clinical implementation and does not provide individual AI-derived fertility scores to participants. Any future implementation would require careful ethical oversight, appropriate counseling, and external validation.

Conclusion

This protocol describes the initial phase of the N-SEED initiative, which establishes a standardized and scalable framework integrating prospectively collected sperm microscopy image/video data, the DFI, and extensive environmental and physiological covariates in healthy adult men. The resulting dataset will enable association analyses to evaluate feasibility and generate hypotheses. It will also inform subsequent phases, including expansion of participant numbers, refinement of image acquisition procedures, and stepwise development and external validation of image-based analytical approaches (including AI/ML) against clinically relevant and, ideally, longitudinal fertility outcomes. The database structure is designed for long-term maintenance and iterative updates beyond the 2028 study end date, contingent on governance and resources. Thus, the immediate deliverable of the initial phase is a standardized, quality-controlled multimodal resource to support later hypothesis testing, model development, and external validation, rather than a finalized clinical AI tool.

Acknowledgments

We extend our thanks to Takazo Tanaka, Kazumitsu Yamasaki, Hideki Takeshita, Tatsuya Takayama, Naoki Saito, Jun Ueyama, and Hirokazu Nosato for their professional guidance. We are also grateful to Shiori Kushima, Maki Owada, and Masumi Chida for their assistance with laboratory management.

ChatGPT, Gemini, and Perplexity were used for language editing, proofreading, and literature research support during manuscript preparation. The authors have reviewed and revised the output, as necessary, and have accepted full responsibility for the final content of this manuscript.

Data Availability

Data sharing is not applicable to this protocol paper. The dataset generated by the Nippon Semen and Environmental Exposure Database initiative will be stored in a controlled-access manner, and secondary use will be considered upon reasonable request and subject to ethical approval and appropriate data-sharing agreements.

Funding

This project was supported by the Japan Science and Technology Agency (grant JPMJPF2017) and an internal grant for emergent cross-disciplinary research from the National Institute of Advanced Industrial Science and Technology.

Authors' Contributions

Conceptualization: TO and KK
Data curation: TO, KK, and YE
Funding acquisition: TO and HN
Methodology: TO, KK, YE, TM, HM, AN, and KY
Project administration: TO, KK, and AN
Resources: TO, KK, YE, AN, and KY
Software: TO and KK
Supervision: KY, AI, HN, and IO
Writing the original draft: TO, KK, and YE
Writing, review, and editing: TM, HM, AN, KY, AI, HN, and IO

Conflicts of Interest

None declared.

References

1. Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. Evidence for decreasing quality of semen during past 50 years. BMJ 1992 Sep 12;305(6854):609-613 [FREE Full text] [doi: 10.1136/bmj.305.6854.609] [Medline: 1393072]

XSL•FO
RenderX

2. Levine H, Jørgensen N, Martino-Andrade A, Mendiola J, Weksler-Derri D, Jolles M, et al. Temporal trends in sperm count: a systematic review and meta-regression analysis of samples collected globally in the 20th and 21st centuries. Hum Reprod Update 2023 Mar 01;29(2):157-176 [FREE Full text] [doi: 10.1093/humupd/dmac035] [Medline: 36377604]
3. Jørgensen N, Lamb DJ, Levine H, Pastuszak AW, Sigalos JT, Swan SH, et al. Are worldwide sperm counts declining? Fertil Steril 2021 Dec;116(6):1457-1463 [FREE Full text] [doi: 10.1016/j.fertnstert.2021.10.020] [Medline: 34836581]
4. Auger J, Eustache F, Chevrier C, Jégou B. Spatiotemporal trends in human semen quality. Nat Rev Urol 2022 Oct;19(10):597-626 [FREE Full text] [doi: 10.1038/s41585-022-00626-w] [Medline: 35978007]
5. Sciorio R, Tramontano L, Adel M, Fleming S. Decrease in sperm parameters in the 21st century: obesity, lifestyle, or environmental factors? An updated narrative review. J Pers Med 2024 Feb 11;14(2):198 [FREE Full text] [doi: 10.3390/jpm14020198] [Medline: 38392631]
6. Szabó A, Nyirády P, Kopa Z. Impact of lifestyle and environmental factors on fertility. Curr Opin Urol 2025 Nov 01;35(6):685-690 [FREE Full text] [doi: 10.1097/MOU.0000000000001339] [Medline: 40923126]
7. Zhao CC, Scott M, Eisenberg ML. Male fertility as a proxy for health. J Clin Med 2024 Sep 19;13(18):5559 [FREE Full text] [doi: 10.3390/jcm13185559] [Medline: 39337044]
8. Kaltsas A, Koumenis A, Stavropoulos M, Kratiras Z, Deligiannis D, Adamos K, et al. Male infertility and reduced life expectancy: epidemiology, mechanisms, and clinical implications. J Clin Med 2025 Jun 03;14(11):3930 [FREE Full text] [doi: 10.3390/jcm14113930] [Medline: 40507691]
9. Priskorn L, Lindahl-Jacobsen R, Jensen TK, Holmboe SA, Hansen LS, Kriegbaum M, et al. Semen quality and lifespan: a study of 78 284 men followed for up to 50 years. Hum Reprod 2025 Apr 01;40(4):730-738 [FREE Full text] [doi: 10.1093/humrep/deaf023] [Medline: 40037905]
10. Adler A, Roth B, Lundy SD, Takeshima T, Yumura Y, Kuroda S. Sperm DNA fragmentation testing in clinical management of reproductive medicine. Reprod Med Biol 2023;22(1):e12547 [FREE Full text] [doi: 10.1002/rmb2.12547] [Medline: 37915974]
11. Furini C, Costantino F, Granata AR, Spaggiari G, Santi D, Simoni M. Still counting sperm? Why novel, truly informative measurements of testis function in male infertility are urgently needed. Endocrine 2025 Dec;90(3):1067-1078 [FREE Full text] [doi: 10.1007/s12020-025-04453-y] [Medline: 41075128]
12. Agarwal A, Majzoub A, Baskaran S, Panner Selvam MK, Cho CL, Henkel R, et al. Sperm DNA fragmentation: a new guideline for clinicians. World J Mens Health 2020 Oct;38(4):412-471 [FREE Full text] [doi: 10.5534/wjmh.200128] [Medline: 32777871]
13. Esteves SC, Zini A, Coward RM, Evenson DP, Gosálvez J, Lewis SEM, et al. Sperm DNA fragmentation testing: summary evidence and clinical practice recommendations. Andrologia 2021 Mar;53(2):e13874 [FREE Full text] [doi: 10.1111/and.13874] [Medline: 33108829]
14. Evenson DP. The sperm chromatin structure assay (SCSA(®)) and other sperm DNA fragmentation tests for evaluation of sperm nuclear DNA integrity as related to fertility. Anim Reprod Sci 2016 Jun;169:56-75 [FREE Full text] [doi: 10.1016/j.anireprosci.2016.01.017] [Medline: 26919909]
15. Evenson DP. Sperm chromatin structure assay (SCSA®) for fertility assessment. Curr Protoc 2022 Aug;2(8):e508 [FREE Full text] [doi: 10.1002/cpz1.508] [Medline: 35926128]
16. Li F, Duan X, Li M, Ma X. Sperm DNA fragmentation index affect pregnancy outcomes and offspring safety in assisted reproductive technology. Sci Rep 2024 Jan 03;14(1):356 [FREE Full text] [doi: 10.1038/s41598-023-45091-6] [Medline: 38172506]
17. Wan B, Fu Y, Ma N, Zhou Z, Lu W. Sperm DNA fragmentation and assisted reproduction: an umbrella meta-analysis. Eur J Med Res 2026 Jan 07;31(1):218 [FREE Full text] [doi: 10.1186/s40001-025-03801-y] [Medline: 41501955]
18. Naik N, Roth B, Lundy SD. Artificial intelligence for clinical management of male infertility, a scoping review. Curr Urol Rep 2024 Nov 09;26(1):17 [FREE Full text] [doi: 10.1007/s11934-024-01239-z] [Medline: 39520645]
19. Qaderi K, Sharifipour F, Dabir M, Shams R, Behmanesh A. Artificial intelligence (AI) approaches to male infertility in IVF: a mapping review. Eur J Med Res 2025 Apr 05;30(1):246 [FREE Full text] [doi: 10.1186/s40001-025-02479-6] [Medline: 40186251]
20. Panner Selvam MK, Moharana AK, Baskaran S, Finelli R, Hudnall MC, Sikka SC. Current updates on involvement of artificial intelligence and machine learning in semen analysis. Medicina (Kaunas) 2024 Feb 06;60(2):279 [FREE Full text] [doi: 10.3390/medicina60020279] [Medline: 38399566]
21. McCallum C, Riordon J, Wang Y, Kong T, You JB, Sanner S, et al. Deep learning-based selection of human sperm with high DNA integrity. Commun Biol 2019;2:250 [FREE Full text] [doi: 10.1038/s42003-019-0491-6] [Medline: 31286067]
22. Shin T, Shin J, Kim J, Choi K, Kim J. P-064 prediction of human sperm DNA fragmentation using holotomography-based refractive index and artificial intelligence models. Hum Reprod 2025 Jun;40(Suppl 1):deaf097.373 [FREE Full text] [doi: 10.1093/humrep/deaf097.373]
23. Du C, Li Y, Yin C, Luo X, Pan X. Association of abstinence time with semen quality and fertility outcomes: a systematic review and dose-response meta-analysis. Andrology 2024 Sep;12(6):1224-1235 [FREE Full text] [doi: 10.1111/andr.13583] [Medline: 38197853]

XSL•FO
RenderX

24. Xi Q, Kong Q, Lv X, Dai R, Yu Y. Impact of ejaculation frequency on semen parameters and DNA fragmentation: a cross-sectional study. Reprod Biol Endocrinol 2025 Jul 16;23(1):100 [FREE Full text] [doi: 10.1186/s12958-025-01439-3] [Medline: 40671092]
25. Beyatlı M, Güngör HS, Uzel T, Sobay R, İnkaya A, Evci MU, et al. Longer arousal, better semen: effects of extended pre-ejaculatory arousal on semen parameters in infertile men. Diagnostics (Basel) 2025 Dec 13;15(24):3186 [FREE Full text] [doi: 10.3390/diagnostics15243186] [Medline: 41464187]
26. van Roijen JH, Slob AK, Gianotten WL, Dohle GR, van der Zon AT, Vreeburg JT, et al. Sexual arousal and the quality of semen produced by masturbation. Hum Reprod 1996 Jan;11(1):147-151 [FREE Full text] [doi: 10.1093/oxfordjournals.humrep.a019008] [Medline: 8671177]
27. Radke EG, Braun JM, Meeker JD, Cooper GS. Phthalate exposure and male reproductive outcomes: a systematic review of the human epidemiological evidence. Environ Int 2018 Dec;121(Pt 1):764-793 [FREE Full text] [doi: 10.1016/j.envint.2018.07.029] [Medline: 30336412]
28. Sciorio R, Greco PF, Greco E, Tramontano L, Elshaer FM, Fleming S. Potential effects of environmental toxicants on sperm quality and potential risk for fertility in humans. Front Endocrinol (Lausanne) 2025;16:1545593 [FREE Full text] [doi: 10.3389/fendo.2025.1545593] [Medline: 40469442]
29. Bocu K, Boeri L, Mahmutoglu AM, Vogiatzi P. Can lifestyle changes significantly improve male fertility: a narrative review? Arab J Urol 2025;23(3):190-200 [FREE Full text] [doi: 10.1080/20905998.2024.2421626] [Medline: 40747470]
30. Kobayashi H, Uetani M, Yamabe F, Mitsui Y, Nakajima K, Nagao K. A new model for determining risk of male infertility from serum hormone levels, without semen analysis. Sci Rep 2024 Jul 31;14(1):17079 [FREE Full text] [doi: 10.1038/s41598-024-67910-0] [Medline: 39085312]
31. Borg DJ, Haritopoulou-Sinanidou M, Gabrovska P, Tseng H, Honeyman D, Schweitzer D, et al. Barriers and facilitators for recruiting and retaining male participants into longitudinal health research: a systematic review. BMC Med Res Methodol 2024 Feb 22;24(1):46 [FREE Full text] [doi: 10.1186/s12874-024-02163-z] [Medline: 38389065]
32. Stewart TM, Liu DY, Garrett C, Brown EH, Baker HWG. Recruitment bias in studies of semen and other factors affecting pregnancy rates in fertile men. Hum Reprod 2009 Oct;24(10):2401-2408 [FREE Full text] [doi: 10.1093/humrep/dep215] [Medline: 19553241]
33. Huang H, Hsieh S, Chen M, Jhou M, Liu T, Shen H, et al. Machine learning predictive models for evaluating risk factors affecting sperm count: predictions based on health screening indicators. J Clin Med 2023 Feb 03;12(3):1220 [FREE Full text] [doi: 10.3390/jcm12031220] [Medline: 36769868]
34. Aykaç A, Kaya C, Çelik Ö, Aydın ME, Sungur M. The prediction of semen quality based on lifestyle behaviours by the machine learning based models. Reprod Biol Endocrinol 2024 Aug 29;22(1):112 [FREE Full text] [doi: 10.1186/s12958-024-01268-w] [Medline: 39210437]
35. Pan M, Zhang Y, Qin N, Xu Y, Ni S, Chen W, et al. Predictive model of sperm DNA fragmentation in infertile men based on lifestyle factors. Front Endocrinol (Lausanne) 2025;16:1675168 [FREE Full text] [doi: 10.3389/fendo.2025.1675168] [Medline: 41220579]
36. Ghayda RA, Cannarella R, Calogero AE, Shah R, Rambhatla A, Zohdy W, Global Andrology Forum. Artificial intelligence in andrology: from semen analysis to image diagnostics. World J Mens Health 2024 Jan;42(1):39-61 [FREE Full text] [doi: 10.5534/wjmh.230050] [Medline: 37382282]
37. Leung ETY, Mei X, Lee BKM, Lam KKW, Lee C, Li RHW, et al. Automatic identification of human spermatozoa with zona pellucida-binding capability using deep learning. Hum Reprod Open 2025;2025(3):hoaf024 [FREE Full text] [doi: 10.1093/hropen/hoaf024] [Medline: 40487847]
38. Abdelkefi O, Maalej R, Rebai T, Sellami A, Daoud S. Deep-learning based model for sperm morphology assessment using the SMD/MSS dataset. Future Sci OA 2025 Dec;11(1):2583010 [FREE Full text] [doi: 10.1080/20565623.2025.2583010] [Medline: 41203404]
39. Liu J, Dai Y, Li R, Yuan J, Wang Q, Wang L. Does air pollution exposure affect semen quality? Evidence from a systematic review and meta-analysis of 93,996 Chinese men. Front Public Health 2023;11:1219340 [FREE Full text] [doi: 10.3389/fpubh.2023.1219340] [Medline: 37601219]
40. Margiana R, Odhar HA, Prasad K, Oghenemaro EF, Kumawat R, Uthirapathy S, et al. Does outdoor air pollution cause poor semen quality? A systematic review and meta-analysis. BMC Urol 2025 Mar 13;25(1):50 [FREE Full text] [doi: 10.1186/s12894-025-01728-4] [Medline: 40082868]
41. Nagao T, Fukuta J, Hanioka T, Nakayama Y, Warnakulasuriya S, Sasaki T, Tobacco Cessation Intervention Study for Oral Diseases. A multicentre tobacco cessation intervention study in the dental setting in Japan. Int Dent J 2022 Feb;72(1):123-132 [FREE Full text] [doi: 10.1016/j.identj.2021.02.002] [Medline: 33743994]
42. Anelli V, Gatta E, Pirola I, Delbarba A, Rotondi M, Cappelli C. Thyroid impairment and male fertility: a narrative review of literature. Aging Male 2024 Dec;27(1):2310303 [FREE Full text] [doi: 10.1080/13685538.2024.2310303] [Medline: 38347677]
43. Fertig RM, Gamret AC, Darwin E, Gaudi S. Sexual side effects of 5-α-reductase inhibitors finasteride and dutasteride: a comprehensive review. Dermatol Online J 2017 Nov 11;23(11):13030/qt24k8q743 [FREE Full text] [Medline: 29447628]

XSL•FO
RenderX

44. Santana FDFV, Lozi AA, Gonçalves RV, Da Silva J, Da Matta SLP. Comparative effects of finasteride and minoxidil on the male reproductive organs: a systematic review of in vitro and in vivo evidence. Toxicol Appl Pharmacol 2023 Nov 01;478:116710 [FREE Full text] [doi: 10.1016/j.taap.2023.116710] [Medline: 37805090]
45. Drobnis EZ, Nangia AK. Immunosuppressants and male reproduction. Adv Exp Med Biol 2017;1034:179-210 [FREE Full text] [doi: 10.1007/978-3-319-69535-8_12] [Medline: 29256132]
46. Kapoor R, Panner Selvam MK, Sikka SC. Impact of socioeconomic status on male reproductive health: a mini review. Reprod Med 2025 Dec 10;6(4):44 [FREE Full text] [doi: 10.3390/reprodmed6040044]
47. Godart ES, Shin DH, Christensen E, Thompson ER, Turek PJ. A study of pregnancy rates in "cleared" male factor couples. Transl Androl Urol 2021 Feb;10(2):620-625 [FREE Full text] [doi: 10.21037/tau-20-1240] [Medline: 33718064]
48. Rimmer MP, Howie RA, Anderson RA, Barratt CLR, Barnhart KT, Beebeejaun Y, et al. A core outcome set for future male infertility research: development of an international consensus. Fertil Steril 2025 Jun;123(6):1017-1028 [FREE Full text] [doi: 10.1016/j.fertnstert.2025.03.009] [Medline: 40237684]
49. Anakwe A, BeLue R, Xian H, Xaverius P. Men's preconception health and fertility intentions: a latent class analysis approach. Am J Mens Health 2022;16(6):15579883221135764 [FREE Full text] [doi: 10.1177/15579883221135764] [Medline: 36373425]
50. Gonzalez D, Narasimman M, Best JC, Ory J, Ramasamy R. Clinical update on home testing for male fertility. World J Mens Health 2021 Oct;39(4):615-625 [FREE Full text] [doi: 10.5534/wjmh.200130] [Medline: 33474845]
51. Viramgami A, Balachandar R, Bagepally BS, Sheth A. Study on the association between night shift work and reproductive functions among male workers: a systematic review and meta-analysis. Endocrine 2025 May;88(2):410-419 [FREE Full text] [doi: 10.1007/s12020-025-04166-2] [Medline: 39833521]
52. McKinnon CJ, Joglekar DJ, Hatch EE, Rothman KJ, Wesselink AK, Willis MD, et al. Male personal heat exposures and fecundability: a preconception cohort study. Andrology 2022 Nov;10(8):1511-1521 [FREE Full text] [doi: 10.1111/andr.13242] [Medline: 35924639]
53. Komiya A, Kato T, Kawauchi Y, Watanabe A, Fuse H. Clinical factors associated with sperm DNA fragmentation in male patients with infertility. ScientificWorldJournal 2014;2014:868303 [FREE Full text] [doi: 10.1155/2014/868303] [Medline: 25165747]
54. Osaki Y, Ino A, Matsushita S, Higuchi S, Kondo Y, Kinjo A. Reliability and validity of the alcohol use disorders identification test - consumption in screening for adults with alcohol use disorders and risky drinking in Japan. Asian Pac J Cancer Prev 2014 Aug 30;15(16):6571-6574 [FREE Full text] [doi: 10.7314/apjcp.2014.15.16.6571] [Medline: 25169489]
55. Kobayashi S, Yuan X, Sasaki S, Osawa Y, Hirata T, Abe Y, et al. Relative validity of brief-type self-administered diet history questionnaire among very old Japanese aged 80 years or older. Public Health Nutr 2019 Feb;22(2):212-222 [FREE Full text] [doi: 10.1017/S1368980018002331] [Medline: 30274565]
56. Doi Y, Minowa M, Uchiyama M, Okawa M, Kim K, Shibui K, et al. Psychometric assessment of subjective sleep quality using the Japanese version of the Pittsburgh Sleep Quality Index (PSQI-J) in psychiatric disordered and control subjects. Psychiatry Res 2000 Dec 27;97(2-3):165-172 [FREE Full text] [doi: 10.1016/s0165-1781(00)00232-8] [Medline: 11166088]
57. Eto T, Nishimura Y, Ikeda H, Kunieda T, Enomoto M, Kubo T, et al. External validation of the Japanese version of the reduced Morningness-Eveningness Questionnaire (rMEQ) score using dim light melatonin onset and sleep-wake behavior in young adults. Chronobiol Int 2025 Sep;42(9):1137-1144 [FREE Full text] [doi: 10.1080/07420528.2025.2532795] [Medline: 40671574]
58. Kitamura N, Sato T, Kawagoshi A, Satake M, Shioya T. Evaluation of physical activity in the daily life of healthy young subjects with special reference to the reliability and validity of IPAQ as evaluated by a triaxial accelerometer. Rigakuryoho Kagaku 2010;25(5):767-771 [FREE Full text] [doi: 10.1589/rika.25.767]
59. Mimura C, Griffiths P. A Japanese version of the Perceived Stress Scale: cross-cultural translation and equivalence assessment. BMC Psychiatry 2008 Sep 30;8:85 [FREE Full text] [doi: 10.1186/1471-244X-8-85] [Medline: 18826581]
60. Furukawa TA, Kawakami N, Saitoh M, Ono Y, Nakane Y, Nakamura Y, et al. The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan. Int J Methods Psychiatr Res 2008;17(3):152-158 [FREE Full text] [doi: 10.1002/mpr.257] [Medline: 18763695]
61. Day D, Ambegaonkar A, Harriot K, McDaniel A. A new tool for predicting erectile dysfunction. Adv Ther 2001;18(3):131-139 [FREE Full text] [doi: 10.1007/BF02850301] [Medline: 11571825]
62. Minami H, Furukawa S, Sakai T, Niiya T, Miyaoka H, Miyake T, et al. Physical activity and prevalence of erectile dysfunction in Japanese patients with type 2 diabetes mellitus: the Dogo study. J Diabetes Investig 2018 Jan;9(1):193-198 [FREE Full text] [doi: 10.1111/jdi.12660] [Medline: 28371446]
63. Shiraishi K. Evaluation of sexual function after robot-assisted radical prostatectomy: a farewell to IIEF questionnaire. Int J Urol 2023 Nov;30(11):959-967 [FREE Full text] [doi: 10.1111/iju.15264] [Medline: 37526397]
64. Mulhall JP, Goldstein I, Bushmakin AG, Cappelleri JC, Hvidsten K. Validation of the erection hardness score. J Sex Med 2007 Nov;4(6):1626-1634 [FREE Full text] [doi: 10.1111/j.1743-6109.2007.00600.x] [Medline: 17888069]
65. Miyake H, Miyazaki A, Yao A, Hinata N, Fujisawa M. Significance of erection hardness score as a diagnostic tool to assess erectile function recovery in Japanese men after robot-assisted radical prostatectomy. J Robot Surg 2016 Sep;10(3):221-226 [FREE Full text] [doi: 10.1007/s11701-016-0571-2] [Medline: 26994775]

XSL•FO
RenderX

66. Rosen RC, Catania J, Pollack L, Althof S, O'Leary M, Seftel AD. Male Sexual Health Questionnaire (MSHQ): scale development and psychometric validation. Urology 2004 Oct;64(4):777-782 [FREE Full text] [doi: 10.1016/j.urology.2004.04.056] [Medline: 15491719]
67. Shirai M, Sano M, Anno Y, Ishikawa K, Taniguchi A, Kure A, et al. Efficacy of a new vacuum erection device (Vigor 2020) for erectile dysfunction: a retrospective study in Japan. Int J Urol 2024 Dec;31(12):1386-1392 [FREE Full text] [doi: 10.1111/iju.15574] [Medline: 39253956]
68. McNabney SM, Weseman CE, Hevesi K, Rowland DL. Are the criteria for the diagnosis of premature ejaculation applicable to gay men or sexual activities other than penile-vaginal intercourse? Sex Med 2022 Jun;10(3):100516 [FREE Full text] [doi: 10.1016/j.esxm.2022.100516] [Medline: 35477122]
69. Tsujimura A, Fukuhara S, Chiba K, Yoshizawa T, Tomoe H, Shirai M, Clinical Research Promotion Committee of the Japanese Society for Sexual Medicine. Erectile function and sexual activity are declining in the younger generation: results from a national survey in Japan. World J Mens Health 2025 Jan;43(1):239-248 [FREE Full text] [doi: 10.5534/wjmh.240137] [Medline: 39344114]
70. Ichioka K, Okubo K, Terai A, Itoh N, Nakayama T, Hatayama H, et al. [The Aging Male's Symptoms (AMS) scale in Japanese men attending an infertility clinic]. Hinyokika Kiyo 2006 Aug;52(8):615-618 [FREE Full text] [Medline: 16972623]
71. Kobayashi K, Hashimoto K, Kato R, Tanaka T, Hirose T, Masumori N, et al. The aging males' symptoms scale for Japanese men: reliability and applicability of the Japanese version. Int J Impot Res 2008;20(6):544-548 [FREE Full text] [doi: 10.1038/ijir.2008.27] [Medline: 18548079]
72. Tanaka T, Kojo K, Suetomi T, Nagumo Y, Midorikawa H, Matsuda T, et al. Distinct clusters of testosterone levels, symptoms, and serum trace elements in young men: a cross-sectional analysis. Nutrients 2025 Feb 28;17(5):867 [FREE Full text] [doi: 10.3390/nu17050867] [Medline: 40077734]
73. Takahashi S, Wada K, Kumon H, Masuda H, Suzuki Y, Yokoyama O, et al. [The Japanese version of the National Institutes of Health Chronic Prostatitis Symptom Index]. Nihon Hinyokika Gakkai Zasshi 2014 Apr;105(2):62-65 [FREE Full text] [doi: 10.5980/jpnjurol.105.62] [Medline: 24908819]
74. Wang Y, Luo D, Liu J, Song Y, Jiang B, Jiang H. Low skeletal muscle mass index and all-cause mortality risk in adults: a systematic review and meta-analysis of prospective cohort studies. PLoS One 2023;18(6):e0286745 [FREE Full text] [doi: 10.1371/journal.pone.0286745] [Medline: 37285331]
75. Ortega FB, Sui X, Lavie CJ, Blair SN. Body mass index, the most widely used but also widely criticized index: would a criterion standard measure of total body fat be a better predictor of cardiovascular disease mortality? Mayo Clin Proc 2016 Apr;91(4):443-455 [FREE Full text] [doi: 10.1016/j.mayocp.2016.01.008] [Medline: 26948431]
76. Ando F, Endo Y, Terauchi Y, Onoda K. Effect of body composition on supine pressure distribution in bed: a comparison between young and older adults. Cureus 2025 Jul;17(7):e88987 [FREE Full text] [doi: 10.7759/cureus.88987] [Medline: 40896022]
77. Gonzalez MC, Barbosa-Silva TG, Bielemann RM, Gallagher D, Heymsfield SB. Phase angle and its determinants in healthy subjects: influence of body composition. Am J Clin Nutr 2016 Mar;103(3):712-716 [FREE Full text] [doi: 10.3945/ajcn.115.116772] [Medline: 26843156]
78. Sato M, Ishizaka M, Endo Y, Ando F, Hirose M, Ohira K, et al. Comparison of body composition between patients admitted to convalescent rehabilitation and long-term care wards. J Phys Ther Sci 2025 Mar;37(3):141-144 [FREE Full text] [doi: 10.1589/jpts.37.141] [Medline: 40034555]
79. Zhang J, Zhang N, Du S, Liu S, Ma G. Effects of water restriction and water replenishment on the content of body water with bioelectrical impedance among young adults in Baoding, China: a randomized controlled trial (RCT). Nutrients 2021 Feb 08;13(2):553 [FREE Full text] [doi: 10.3390/nu13020553] [Medline: 33567589]
80. Ciardullo S, Zerbini F, Cannistraci R, Muraca E, Perra S, Oltolini A, et al. Differential association of sex hormones with metabolic parameters and body composition in men and women from the United States. J Clin Med 2023 Jul 19;12(14):4783 [FREE Full text] [doi: 10.3390/jcm12144783] [Medline: 37510898]
81. Viken AF, Siiak SP, Schlünssen V, Thorarinsdottir EH, Skulstad SM, Gyawali S, et al. Muscle strength and male sexual function. J Clin Med 2024 Jan 12;13(2):426 [FREE Full text] [doi: 10.3390/jcm13020426] [Medline: 38256560]
82. Beaudart C, Rolland Y, Cruz-Jentoft AJ, Bauer JM, Sieber C, Cooper C, et al. Assessment of muscle function and physical performance in daily clinical practice: a position paper endorsed by the European Society for Clinical and Economic Aspects of Osteoporosis, Osteoarthritis and Musculoskeletal Diseases (ESCEO). Calcif Tissue Int 2019 Jul;105(1):1-14 [FREE Full text] [doi: 10.1007/s00223-019-00545-w] [Medline: 30972475]
83. Wind AE, Takken T, Helders PJM, Engelbert RHH. Is grip strength a predictor for total muscle strength in healthy children, adolescents, and young adults? Eur J Pediatr 2010 Mar;169(3):281-287 [FREE Full text] [doi: 10.1007/s00431-009-1010-4] [Medline: 19526369]
84. Orange ST, Metcalfe JW, Liefeith A, Marshall P, Madden LA, Fewster CR, et al. Validity and reliability of a wearable inertial sensor to measure velocity and power in the back squat and bench press. J Strength Cond Res 2019 Sep;33(9):2398-2408 [FREE Full text] [doi: 10.1519/JSC.0000000000002574] [Medline: 29742745]

XSL•FO
RenderX

85. Roberts HC, Denison HJ, Martin HJ, Patel HP, Syddall H, Cooper C, et al. A review of the measurement of grip strength in clinical and epidemiological studies: towards a standardised approach. Age Ageing 2011 Jul;40(4):423-429 [FREE Full text] [doi: 10.1093/ageing/afr051] [Medline: 21624928]
86. Steele J, Bruce-Low S, Smith D, Jessop D, Osborne N. Isolated lumbar extension resistance training improves strength, pain, and disability, but not spinal height or shrinkage ("creep") in participants with chronic low back pain. Cartilage 2020 Apr;11(2):160-168 [FREE Full text] [doi: 10.1177/1947603517695614] [Medline: 29156985]
87. Ma W, Yang C, Shih S, Hsieh H, Hung CS, Chiu F, et al. Measurement of waist circumference: midabdominal or iliac crest? Diabetes Care 2013 Jun;36(6):1660-1666 [FREE Full text] [doi: 10.2337/dc12-1452] [Medline: 23275359]
88. Abedi Yekta AH, Khosravi S, Hassabi M, Poorsaid Esfahani M, Hassanmirzaei B, Asgari A. Comparing the accuracy of waist-hip ratio calculation by the BIA device versus the manual method. Ann Appl Sport Sci 2016 Jul 01;4(2):9-15 [FREE Full text] [doi: 10.18869/acadpub.aassjournal.4.2.9]
89. Jayedi A, Soltani S, Zargar MS, Khan TA, Shab-Bidar S. Central fatness and risk of all cause mortality: systematic review and dose-response meta-analysis of 72 prospective cohort studies. BMJ 2020 Sep 23;370:m3324 [FREE Full text] [doi: 10.1136/bmj.m3324] [Medline: 32967840]
90. Gonzalez MC, Mehrnezhad A, Razaviarab N, Barbosa-Silva TG, Heymsfield SB. Calf circumference: cutoff values from the NHANES 1999-2006. Am J Clin Nutr 2021 Jun 01;113(6):1679-1687 [FREE Full text] [doi: 10.1093/ajcn/nqab029] [Medline: 33742191]
91. Dupesh S, Pandiyan N, Pandiyan R, Kartheeswaran J, Prakash B. Ejaculatory abstinence in semen analysis: does it make any sense? Ther Adv Reprod Health 2020;14:2633494120906882 [FREE Full text] [doi: 10.1177/2633494120906882] [Medline: 32596668]
92. Finelli R, Leisegang K, Tumallapalli S, Henkel R, Agarwal A. The validity and reliability of computer-aided semen analyzers in performing semen analysis: a systematic review. Transl Androl Urol 2021 Jul;10(7):3069-3079 [FREE Full text] [doi: 10.21037/tau-21-276] [Medline: 34430409]
93. Boitrelle F, Shah R, Saleh R, Henkel R, Kandil H, Chung E, et al. The sixth edition of the WHO Manual for Human Semen Analysis: a critical review and SWOT analysis. Life (Basel) 2021 Dec 09;11(12):1368 [FREE Full text] [doi: 10.3390/life11121368] [Medline: 34947899]
94. Chawre S, Khatib MN, Rawekar A, Mahajan S, Jadhav R, More A. A review of semen analysis: updates from the WHO sixth edition manual and advances in male fertility assessment. Cureus 2024 Jun 29;16(6):e63485 [FREE Full text] [doi: 10.7759/cureus.63485] [Medline: 39081428]
95. Mortimer ST, Mortimer D. Kinematics of human spermatozoa incubated under capacitating conditions. J Androl 1990;11(3):195-203 [FREE Full text] [Medline: 2384341]
96. Yamasaki K, Uchida M, Watanabe N, Ihana T, Ishiguro Y, Kuroda S, et al. Effects of antioxidant co-supplementation therapy on spermatogenesis dysfunction in relation to the basal oxidation-reduction potential levels in spermatozoa: a pilot study. Reprod Med Biol 2022;21(1):e12450 [FREE Full text] [doi: 10.1002/rmb2.12450] [Medline: 35386378]
97. Kondoh N, Togo Y, Kajio K, Kaiduka Y, Yamamoto S. [Evaluation of orgasm using interview questionnaire for male patients with sexual dysfunction]. Jpn J Sex Med (Nihon Sei Kinō Gakkai zasshi) 2020;35(1):21-26 [FREE Full text]
98. Yamanaka M, Tomita K, Hashimoto S, Matsumoto H, Satoh M, Kato H, et al. Combination of density gradient centrifugation and swim-up methods effectively decreases morphologically abnormal sperms. J Reprod Dev 2016 Dec 20;62(6):599-606 [FREE Full text] [doi: 10.1262/jrd.2016-112] [Medline: 27616283]
99. Matsuura R, Takeuchi T, Yoshida A. Preparation and incubation conditions affect the DNA integrity of ejaculated human spermatozoa. Asian J Androl 2010 Sep;12(5):753-759 [FREE Full text] [doi: 10.1038/aja.2010.46] [Medline: 20562894]
100. Mohammadi Z, Tavalaee M, Gharagozloo P, Drevet JR, Nasr-Esfahani MH. Could high DNA stainability (HDS) be a valuable indicator of sperm nuclear integrity? Basic Clin Androl 2020;30:12 [FREE Full text] [doi: 10.1186/s12610-020-00110-8] [Medline: 32817794]
101. Kobayashi M, Yoshida A. Seishi DNA no sonshō to jusei. In: Shibahara H, Yanagida K, editors. Jusei to sono shōgai. Tokyo, Japan: Chūgai igakusha; 2022:29-36.
102. Brambilla DJ, Matsumoto AM, Araujo AB, McKinlay JB. The effect of diurnal variation on clinical measurement of serum testosterone and other sex hormone levels in men. J Clin Endocrinol Metab 2009 Mar;94(3):907-913 [FREE Full text] [doi: 10.1210/jc.2008-1902] [Medline: 19088162]
103. Couturier E, van Onderbergen A, Bosson D, Neve J. Circadian variations in plasma zinc and cortisol in man. J Trace Elem Electrolytes Health Dis 1988 Dec;2(4):245-249 [FREE Full text] [Medline: 2980824]
104. Exton MS, Krüger TH, Bursch N, Haake P, Knapp W, Schedlowski M, et al. Endocrine response to masturbation-induced orgasm in healthy men following a 3-week sexual abstinence. World J Urol 2001 Nov;19(5):377-382 [FREE Full text] [doi: 10.1007/s003450100222] [Medline: 11760788]
105. Pozzi E, Belladelli F, Corsini C, Boeri L, Capogrosso P, Fallara G, et al. Contemporary diagnostic work-up for male infertility: emphasizing comprehensive baseline assessment. World J Mens Health 2025 Apr;43(2):265-281 [FREE Full text] [doi: 10.5534/wjmh.240069] [Medline: 39028127]
106. Whitley E, Ball J. Statistics review 1: presenting and summarising data. Crit Care 2002 Feb;6(1):66-71 [FREE Full text] [doi: 10.1186/cc1455] [Medline: 11940268]

XSL•FO
RenderX

107. Zhang Z, Zhu L, Jiang H, Chen H, Chen Y, Dai Y. Sperm DNA fragmentation index and pregnancy outcome after IVF or ICSI: a meta-analysis. J Assist Reprod Genet 2015 Jan;32(1):17-26 [FREE Full text] [doi: 10.1007/s10815-014-0374-1] [Medline: 25392073]
108. Liu K, Mao X, Pan F, Chen Y, An R. Correlation analysis of sperm DNA fragmentation index with semen parameters and the effect of sperm DFI on outcomes of ART. Sci Rep 2023 Feb 15;13(1):2717 [FREE Full text] [doi: 10.1038/s41598-023-28765-z] [Medline: 36792684]
109. Chua SC, Yovich SJ, Hinchliffe PM, Yovich JL. The sperm DNA fragmentation assay with SDF level less than 15% provides a useful prediction for clinical pregnancy and live birth for women aged under 40 years. J Pers Med 2023 Jun 29;13(7):1079 [FREE Full text] [doi: 10.3390/jpm13071079] [Medline: 37511693]
110. Jiang H, Xia X, Luo Y, Pan H, Qu S, Xu J. Sperm DNA fragmentation index: limited effectiveness on predicting embryo quality in assisted reproduction technology treatments. Reprod Biol Endocrinol 2025 Jan 25;23(1):14 [FREE Full text] [doi: 10.1186/s12958-025-01345-8] [Medline: 39863907]
111. Boeri L, Fallara G, Pozzi E, Belladelli F, Corsini C, Raffo M, et al. The impact of different WHO reference criteria for semen analysis in clinical practice: who will benefit from the new 2021 thresholds for normal semen parameters? Andrology 2022 Sep;10(6):1134-1142 [FREE Full text] [doi: 10.1111/andr.13213] [Medline: 35726049]
112. Chung E, Atmoko W, Saleh R, Shah R, Agarwal A. Sixth edition of the World Health Organization Laboratory Manual of Semen Analysis: updates and essential take away for busy clinicians. Arab J Urol 2024;22(2):71-74 [FREE Full text] [doi: 10.1080/20905998.2023.2298048] [Medline: 38481407]
113. Wiweko B, Utami P. Predictive value of sperm deoxyribonucleic acid (DNA) fragmentation index in male infertility. Basic Clin Androl 2017;27:1 [FREE Full text] [doi: 10.1186/s12610-016-0046-3] [Medline: 28239474]
114. Andersson A. Mechanisms for log normal concentration distributions in the environment. Sci Rep 2021 Aug 12;11(1):16418 [FREE Full text] [doi: 10.1038/s41598-021-96010-6] [Medline: 34385549]
115. Barr DB, Wilder LC, Caudill SP, Gonzalez AJ, Needham LL, Pirkle JL. Urinary creatinine concentrations in the U.S. population: implications for urinary biologic monitoring measurements. Environ Health Perspect 2005 Feb;113(2):192-200 [FREE Full text] [doi: 10.1289/ehp.7337] [Medline: 15687057]
116. Carrieri M, Trevisan A, Bartolucci GB. Adjustment to concentration-dilution of spot urine samples: correlation between specific gravity and creatinine. Int Arch Occup Environ Health 2001 Jan;74(1):63-67 [FREE Full text] [doi: 10.1007/s004200000190] [Medline: 11196084]
117. Osaka A, Okada H, Onozuka S, Tanaka T, Iwahata T, Shimomura Y, et al. Evaluation of the sperm DNA fragmentation index in infertile Japanese men by in-house flow cytometric analysis. Asian J Androl 2022;24(1):40-44 [FREE Full text] [doi: 10.4103/aja.aja_49_21] [Medline: 34121749]
118. Zheng Y, Wu Y, Jia Y, Ying L, Yang T, Cheng Q, et al. External quality assessment scheme for sperm DNA fragmentation: a pilot study in China. Basic Clin Androl 2023 Nov 28;33(1):36 [FREE Full text] [doi: 10.1186/s12610-023-00211-0] [Medline: 38012601]
119. Björndahl L, Kvist U. Sequence of ejaculation affects the spermatozoon as a carrier and its message. Reprod Biomed Online 2003;7(4):440-448 [FREE Full text] [doi: 10.1016/s1472-6483(10)61888-3] [Medline: 14656406]
120. Andrade-Rocha FT. Temporary impairment of semen quality following recent acute fever. Ann Clin Lab Sci 2013;43(1):94-97 [FREE Full text] [Medline: 23462613]
121. Martinez MS, Ferreyra FN, Paira DA, Rivero VE, Olmedo JJ, Tissera AD, et al. COVID-19 associates with semen inflammation and sperm quality impairment that reverses in the short term after disease recovery. Front Physiol 2023;14:1220048 [FREE Full text] [doi: 10.3389/fphys.2023.1220048] [Medline: 37497433]
122. Aziz N. The importance of semen analysis in the context of azoospermia. Clinics (Sao Paulo) 2013;68 Suppl 1(Suppl 1):35-38 [FREE Full text] [doi: 10.6061/clinics/2013(sup01)05] [Medline: 23503953]
123. Endo Y, Tanaka T, Kojo K, Matsumoto C, Kurobe M, Nishiyama H, et al. Back muscle strength is associated with self-reported morning-erection frequency in apparently healthy Japanese male university students: a cross-sectional study. Healthcare (Basel) 2026 Mar 18;14(6):759 [FREE Full text] [doi: 10.3390/healthcare14060759] [Medline: 41897212]

Abbreviations

AI: artificial intelligence
AIST: National Institute of Advanced Industrial Science and Technology
AMS: Aging Males' Symptoms Scale
AUDIT-C: Alcohol Use Disorders Identification Test–Consumption
BDHQ: Brief-type Self-Administered Diet History Questionnaire
DFI: DNA fragmentation index
EHS: Erection Hardness Score
HDS: high DNA stainability
INSL3: insulin-like 3
IPAQ: International Physical Activity Questionnaire

XSL•FO
RenderX

IRB: Institutional Review Board
K6: Kessler Psychological Distress Scale
ML: machine learning
MSHQ-EjD-SF: Male Sexual Health Questionnaire–Ejaculatory Dysfunction–Short Form
NIH-CPSI: National Institutes of Health Chronic Prostatitis Symptom Index
N-SEED: Nippon Semen and Environmental Exposure Database
PSQI-J: Japanese version of the Pittsburgh Sleep Quality Index
PSS: Perceived Stress Scale
rMEQ: reduced Morningness-Eveningness Questionnaire
ROC: receiver operating characteristic
SCSA: sperm chromatin structure assay
SHIM: Sexual Health Inventory for Men
UMIN-CTR: UMIN Clinical Trials Registry

Edited by A Schwartz; submitted 25.Feb.2026; peer-reviewed by M Šteffl; comments to author 25.Mar.2026; revised version received 03.Apr.2026; accepted 07.Apr.2026; published 17.Apr.2026.

Please cite as:
Oguri T, Kojo K, Endo Y, Matsuda T, Midorikawa H, Nakazono A, Yanagida K, Ikeda A, Nishiyama H, Ogura I
Integrating Sperm Microscopy, Environmental Exposures, and Lifestyle Factors for Male Fertility Analysis: Protocol for the Nippon Semen and Environmental Exposure Database (N-SEED) Cross-Sectional Study
JMIR Res Protoc 2026;15:e93803
URL: https://www.researchprotocols.org/2026/1/e93803
doi:10.2196/93803
PMID:

©Tomoko Oguri, Kosuke Kojo, Yoshiaki Endo, Takaaki Matsuda, Haruhiko Midorikawa, Ayumi Nakazono, Kaoru Yanagida, Atsushi Ikeda, Hiroyuki Nishiyama, Isamu Ogura. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Building Capacity in Institutional Operational Research in Low-Resource Settings: Protocol for an Implementation Research Study

Sana Ali[1], MPH; G V S Murthy[1], MD; Kenneth L Bassett[2], PhD; Priya Reddy[3], MBA; Katie Judson[3], MPH; Suzanne Schwartz Gilbert[1], PhD; Abhilash Patra[1*], MPH; Anirudh Gaurang Gudlavalleti[1*], MA; Debangana Mahapatra[1*], MSc; Hira Pant[1*], MA; Parami Dakhwa[4*], MBS; Sarah Jameel[1*], MPH; Sarva Priya Pandey[1*], MPH; Shailaja Tetali[5*], PhD; Sirshendu Chaudhuri[5*], MD; Srinivas Imandi[1*], MPH; Suresh Kamalakannan[1,6*], PhD

[1]Pragyaan Sustainable Health Outcomes Foundation, Level 2, Kapil Kavuri Hub, Nanakramguda, Financial District, Hyderabad, India

[2]University of British Columbia, Vancouver, BC, Canada

[3]Seva Canada, Vancouver, BC, Canada

[4]Seva Canada, Kathmandu, Nepal

[5]Indian Institute of Public Health Hyderabad, Hyderabad, India

[6]Northumbria University, Newcastle, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Sana Ali, MPH
Pragyaan Sustainable Health Outcomes Foundation, Level 2, Kapil Kavuri Hub, Nanakramguda, Financial District, Hyderabad, India

Abstract

Background: Operational research (OR) in eye care within resource-constrained settings helps develop context-specific solutions to local challenges. Building OR capacity among eye care personnel enables them to independently generate evidence that drives improvements in eye care delivery.

Objective: This protocol describes the extended phase of the Institutional Operational Research Capacity Building (I-ORCB) program designed to train eye care professionals in the fundamentals and applications of OR.

Methods: The I-ORCB program will be conducted in collaboration with Seva Canada and the Pragyaan Sustainable Health Outcomes Foundation over 3 years (2024 - 2027). A total of 10 partner eye hospitals in India and Nepal will participate. The training is organized around 7 work packages covering quantitative and qualitative research methods, data management and analysis, community engagement, scientific writing, grant writing, and ethics. The program will be implemented in two sequential phases: (1) protocol development and ethics committee submission and (2) mentoring, data collection, data analysis, and manuscript preparation. Delivery will include workshops, ongoing mentorship, e-resources, and structured monitoring mechanisms. Evaluation will follow the Kirkpatrick model, and cost-effectiveness will be assessed through the change in knowledge among the participants brought about by the workshops.

Results: This program is currently in the implementation phase. As of March 2025, a total of 10 teams have been enrolled in the I-ORCB program. Data collection for program evaluation began in June 2025 and is expected to continue until late 2026. Expected results include enhanced OR competencies among hospital teams, institutional strengthening for research, development of peer-reviewed manuscripts, and improved capacity to apply OR to service delivery challenges. The results of this program are expected to be published in 2027.

Conclusions: The I-ORCB program applies a practical, learning-by-doing approach and is grounded in the Cooke framework for research capacity building. By integrating structured mentorship, institutional ownership, and flexible virtual learning components, the program is expected to strengthen OR culture in low- and middle-income countries and improve the delivery of eye care services in South Asia.

International Registered Report Identifier (IRRID): PRR1-10.2196/84284

(JMIR Res Protoc 2026;15:e84284) doi:10.2196/84284

KEYWORDS

operational research; delivery of health care; evidence-based practice; capacity building; program evaluation

XSL•FO
RenderX

Introduction

Background

Health care providers in low- and middle-income countries (LMICs) often face the challenge of delivering care to a large number of patients with limited financial and human resources [1]. Persistent challenges during routine practice—such as overcrowding, lack of awareness among patients about processes and treatment services, supply and data management issues, and poor communication between physicians and patients—affect the quality and efficiency of services, leading to inadequate care, patient attrition, poor treatment adherence, and suboptimal patient outcomes [2]. Operational research (OR) refers to the strategies, interventions, tools, or knowledge that can enhance the quality, coverage, effectiveness, or performance of the health system or programs in which the research is conducted [3]. OR in eye care, especially within resource-limited settings, plays a critical role in identifying targeted solutions to locally relevant challenges [4,5]. This makes OR a valuable approach for decision-makers seeking to enhance eye care services in different contexts, ultimately contributing to improved eye health outcomes [6]. Building capacity in OR is cost-effective in low-resource settings, where it can enable decision-makers to optimize resource allocation and evaluate interventions efficiently [7].

Recognizing this need, the Operational Research Capacity Building (ORCB) program was first launched in 2019 through a collaboration between the Seva Foundation; Seva Canada; and the Indian Institute of Public Health Hyderabad, a subsidiary of the Public Health Foundation of India. This program aimed to build OR capacity in eye care settings in India and Nepal [4]. Since then, 2 phases of the program have been implemented. During the first phase (2019-2021), eye care hospital staff, including clinicians and administrators, in India and Nepal were trained in OR concepts [4]. Building on the success of ORCB-1, the Evidence-Informed Practice program (2023 - 2024) introduced the plan-do-study-act framework focusing on health service quality improvement. Both programs promoted a culture of OR and aimed to generate and use evidence for improving eye health services in resource-poor settings.

However, these earlier phases also encountered challenges. ORCB-1 revealed the need for more targeted strategies to address health system–level and patient-centered challenges [8-11]. In ORCB-2, a major limitation was the lack of high-quality data, which could not be resolved during the program period. Participants also expressed the need for more comprehensive training and skilling and diverse learning formats. It was also observed that many of the staff members who were initially trained had left the institutions, and this hampered the institutions' adoption of OR more widely. A need for institutional strengthening rather than individual strengthening was perceived to be important for the future iterations of the ORCB program so that the institutions could take responsibility for embedding OR principles into their plans.

In response, Seva Canada, a charitable organization whose mission is to restore sight and prevent blindness in LMICs, and Pragyaan Sustainable Health Outcomes (PRASHO) Foundation, a not-for-profit organization that works to develop sustainable health care solutions through public health initiatives grounded in evidence-based practice, launched the Institutional ORCB (I-ORCB) program in October 2024. This phase marks a renewed focus on sustaining an OR environment in LMICs while addressing prior gaps through enhanced training, mentorship, and institutional engagement. A total of 10 partner hospitals—5 from Nepal and 5 from India—were enrolled in the I-ORCB training program, and a preliminary needs assessment was carried out to guide its development.

Prior Work

A baseline assessment was carried out in February 2025 to March 2025 to understand the OR culture across the 10 partner hospitals. This assessment involved a self-administered questionnaire and an on-site visit by mentors. Key areas of focus included the OR teams' familiarity with research, the leadership's interest and commitment to engage in OR activities, and the overall institutional environment to facilitate research. The findings of the needs assessment are summarized below.

All hospital leadership demonstrated a keen interest in building the research skills of their core teams and had a clear understanding of their institutions' OR priorities. The data records were well maintained in most hospitals. However, a few hospitals were in the initial stages of developing a hospital management information system or faced challenges integrating data across departments.

Although all teams had a foundational understanding of research, it was noted that research concepts such as study designs and developing an analysis plan remained weak. Scientific writing skills were mostly limited to senior faculty members, with core team members expressing a need for upskilling in this area. Structured training and hand-holding were identified as essential to strengthen these skills.

Most teams demonstrated limited proficiency in statistical analysis. While team members were familiar with the basics of Microsoft Excel, they lacked confidence in statistical applications. The mentors noted that time constraints posed a significant challenge, as teams had to allocate time for OR activities amid their clinical responsibilities. Some teams were exploring the option of recruiting dedicated research personnel to support OR efforts.

The needs assessment informed the design of a customized I-ORCB training program and contributed to the development of this protocol by identifying gaps, key challenges, and areas requiring targeted support.

Objectives

The objectives of the I-ORCB program are as follows:

- To train the team members of each partner hospital in the fundamentals and applications of OR over a 2-year period
- To provide constant mentorship support to each hospital team until the completion of an OR project over a period of 3 years
- To evaluate the cost-effectiveness of the I-ORCB program in the application of OR over a period of 3 years

XSL•FO
RenderX

Methods

Program Framework

This skilling program will be conducted in collaboration with Seva Canada and the PRASHO Foundation. It will run for a duration of 3 years, from July 2024 to July 2027. To guide the design and delivery of this program, the research capacity building framework by Cooke [12] will be applied. This framework provides a robust foundation for building sustainable and context-relevant research skills. It was recently adapted to support virtual and hybrid learning environments, incorporating tools such as online conceptual learning sessions, online skill-building workshops, e-mentorship, and e-resource repositories [13]. Such virtual platforms, particularly for mentoring, ease time and distance constraints in multicountry programs [13].

The model by Cooke [12] aligns well with our goal of building OR skills for individuals and organizations. The principles of this model will be specifically aligned with the I-ORCB program as follows:

- OR skills—OR capacity will be built by training the participants in appropriate skills to identify operational challenges and use evidence to generate actionable insights. Mentorship will ensure the application of these skills in real-world settings.
- Context relevant—all OR projects will be rooted in local operational challenges faced by the partner hospitals.
- Multidisciplinary team—each OR team will comprise professionals from both clinical and nonclinical backgrounds, each with their own set of skills and expertise.
- Knowledge sharing—findings from all OR projects will be shared across partner hospitals through online dissemination sessions.
- Self-sustainability—all teams will be supported to grow into future OR leaders and embed OR as a part of routine practice, ensuring a self-sustaining OR culture.
- Enabling environment—the program will help partner hospitals develop an environment conducive for OR, such as support for strengthening data management systems, setting up an ethics committee, and providing guidance in securing grants.
- Equity in training—the program will incorporate equity in its delivery by providing tailored mentorship support to reduce disparities in OR skills between the partner hospitals.
- Ownership—each OR project will be driven by a motivated hospital team with strong support from the institutional leadership. Teams will take full ownership and responsibility for allocating time, mobilizing resources, and leading the project from inception to completion.

Training Delivery

Overview

The training for the I-ORCB program will follow an output-oriented, learning-by-doing approach as recommended by the International Union Against Tuberculosis and Lung Disease and Médecins Sans Frontières, who advocate for building OR capacity over an extended period through structured training in protocol development, data management and analysis, and scientific writing [3,14,15]. In line with this recommendation, the I-ORCB training will be delivered in 2 phases: phase 1 will support protocol development for an OR topic for a study that each hospital will conduct and submission of the protocol to the ethics committee, and phase 2 will facilitate mentoring, data collection, data analysis, manuscript preparation, and submission to peer-reviewed journals.

Following consensus among the collaborating partners and in alignment with the needs of the partner hospitals, 7 work packages have been proposed, integrating key concepts of the ORCB program. The work packages and their expected outcomes are presented in Table 1. These work packages will provide a holistic framework for building OR capacity for eye care in low-resource settings.

The I-ORCB program will be delivered through 4 core components: skilling workshops, mentorship support, e-resources, and monitoring mechanisms.

Table . Work packages and their expected outcomes.

Work package	Expected outcome
Training in research methods	To provide participants with a comprehensive understanding of types of research, hypothesis generation, and research question formulation; various study designs appropriate to answer the research questions; types of bias; and confounders in research
Training in data management and analysis	To equip participants with the skills necessary for data collection, managing data and conducting primary and secondary data analysis
Training in qualitative research methods	To train participants in conducting qualitative research, including stakeholder interviews and focus group discussions at the population and eye facility levels
Community engagement for research	To orient the participants on the importance of and modalities for community engagement in eye research
Capacity-building workshops on scientific writing	To build participants' capacity in various aspects of scientific writing
Capacity building for grant writing	To build participants' capacity in various aspects of grant writing
Orientation on ethics in research	To orient the participants on the importance of the principles and practice of ethics in research

XSL•FO
RenderX

Skilling Workshops

The work packages will be delivered in 2 phases over a 3-year period. Each phase will have a series of interactive online skilling workshops conducted over 3 to 4 months. These workshops will be facilitated by experts from Seva Canada, the PRASHO Foundation, and the Indian Institute of Public Health Hyderabad, each bringing expertise from their respective areas.

The workshops will adopt diverse learning methodologies, such as virtual lectures, hands-on activities, and case-based discussions. Each session will last for approximately 2 hours and will be held once or twice a week. Sessions will be scheduled at convenient times to accommodate the busy schedules of the participants and will be adequately spaced to allow teams sufficient time to simultaneously apply the learnings to their particular OR project.

e-Mentorship Support

Each team will be assigned a dedicated mentor who will provide consistent guidance throughout both phases of the skilling workshops. While the lectures will cover the fundamental concepts of OR, the application of these learnings will be situation specific and will be supported by the mentors. The mentors will assist with problem-solving and offer hands-on support through regular interactions. The mentors will hold frequent one-on-one sessions with their teams to foster open communication, build trust, and ensure the timely progression of each project milestone.

e-Resource Repository

To reinforce learning and enhance accessibility, all training sessions will be recorded, and accompanying materials such as slide decks and reading resources will be uploaded to a digital platform. This repository will be accessible to all participants at any time, allowing them to refer back to key concepts when needed.

Monitoring Mechanism

A monitoring framework has been developed to track the progress toward achieving the program's objectives. Figure 1 presents the output-oriented logic model for the ORCB program, illustrating how the activities align with the intended outcomes. The corresponding measurement units for each indicator are detailed in Table 2.

XSL•FO
RenderX

Figure 1. Activity-based logic model. EC: ethics committee.

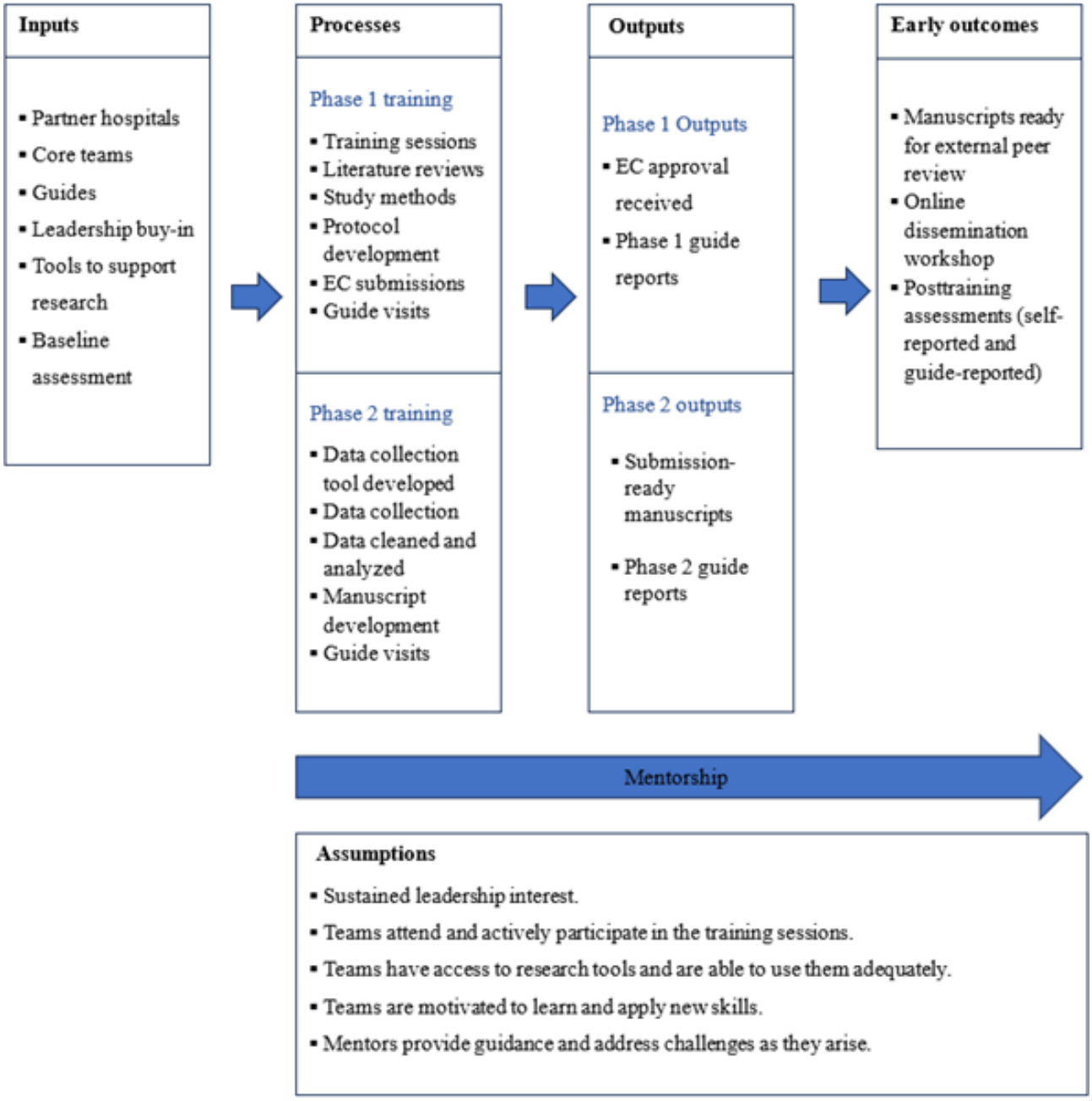

XSL•FO
RenderX

Table . Activity-based indicator matrix.

Measurement	Source
Process indicators: phase 1 training	
Number of research topics developed	Project lead
Number of training sessions conducted	Project in-charge
Number of literature reviews conducted	Hospital feedback form facilitated by the dedicated mentor
Number of study methodologies developed (including research methods, study designs, and ethical considerations)	Hospital feedback form facilitated by the dedicated mentor
Number of data management and analysis plans developed	Hospital feedback form facilitated by the dedicated mentor
Number of protocols developed	Hospital feedback form facilitated by the dedicated mentor
Number of protocols submitted to the EC[a]	Hospital feedback form facilitated by the dedicated mentor
Number of mentor visits completed	Hospital feedback form facilitated by the dedicated mentor
Number of e-mentorship sessions conducted	Project in-charge
Number of e-resources of the training program added to the repository (including PowerPoint presentations and study materials)	Project in-charge
Process indicators: phase 2 training	
Number of data collection tools finalized	Hospital feedback form facilitated by the dedicated mentor
Number of data collection processes completed	Hospital feedback form facilitated by the dedicated mentor
Number of datasets cleaned	Hospital feedback form facilitated by the dedicated mentor
Number of data analyses conducted	Reports submitted to the mentor
Number of mentor visits completed	Project in-charge
Number of manuscripts developed	Submissions to the mentor
Number of manuscripts submitted to journals	Follow-up by the mentor
Number of e-mentorship sessions conducted	Project in-charge
Number of e-resources added to the repository	Project in-charge
Output indicators	
Number of team members trained	Project in-charge
Number of protocols submitted to the EC or number of EC approvals received	Follow-up by the mentor and project in-charge
Number of manuscripts submitted to journals or number of manuscripts developed	Follow-up by the mentor and project in-charge
Phase 1 mentor reports	Phase 1 reports submitted to project in-charge
Phase 2 mentor reports	Phase 2 reports submitted to project in-charge
Early outcome indicators	
Number of publications or number of manuscripts submitted to journals	Project in-charge
Number of projects completed or number of projects disseminated in online workshops	Project in-charge
Posttraining assessment: number of self-reported questionnaires completed and returned	Project in-charge
Posttraining assessment: number of mentor assessments completed	Reports submitted to project in-charge

[a]EC: ethics committee.

I-ORCB Evaluation Plan

The I-ORCB program will use the Kirkpatrick model for evaluation of training [16]. This model has 4 levels of evaluation, and each level has an impact on the next level (Figure 2 [17]). Level 1 focuses on participants' satisfaction with the training they received and will be measured through feedback forms. Level 2 focuses on acquiring knowledge and skills and will be measured through pre- and posttraining tests. Level 3 focuses on transferring knowledge and skills to perform OR activities and will be measured through the achievement of activity-based milestones. Level 4 focuses on the main outcome

XSL•FO
RenderX

and the benefit to the organization as a result of OR training. This will be measured through completion of an OR project and a final mentor visit.

Figure 2. Kirkpatrick model for program evaluation [17]. OR: operational research.

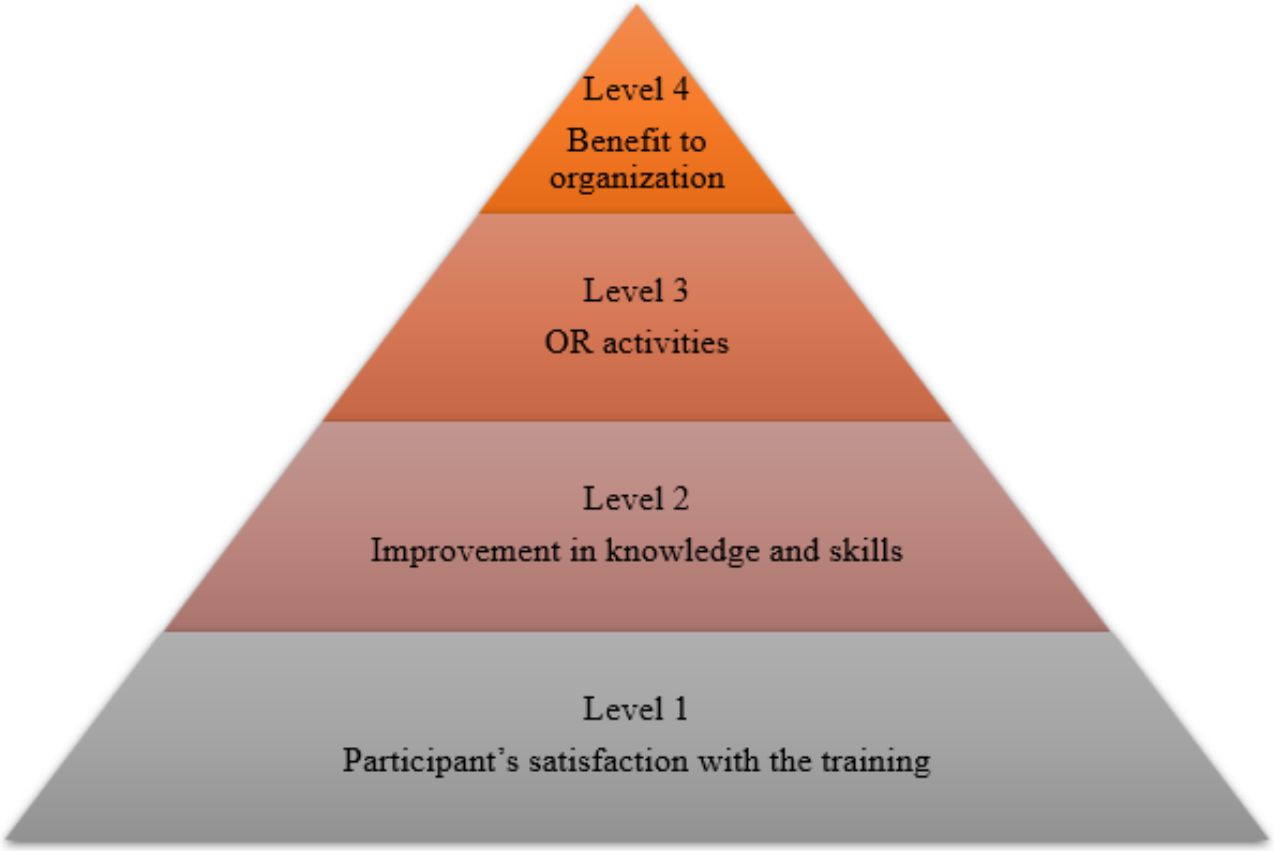

Evaluating the Cost-Effectiveness of the I-ORCB Program

The cost-effectiveness of the I-ORCB training program will be evaluated by estimating the cost of delivering the training and the resulting improvement in the OR knowledge of the ORCB teams [18]. The direct costs will include the training facilitators' fees, training materials, and logistics. The indirect costs will include opportunistic costs (eg, trainee time). The effectiveness of the training will be measured through the change in knowledge among the participants brought about by the workshops.

Expected Outcomes

On the basis of the gaps identified during the baseline assessment (Table 3), the I-ORCB program aims to achieve the following outcomes: (1) enhanced capacity to use and apply existing data for OR projects, (2) enhanced competencies and skills of the teams in conducting studies using OR approaches to address contextual service-related challenges, (3) partner hospitals becoming the hub for OR and expanding the OR environment by leading and training other individuals in the hospitals or partner eye institutions in OR, and (4) improved eye health services and eye health outcomes brought about through sustained OR capacity building.

XSL•FO
RenderX

Table . Baseline assessment for operational research (OR; n=10 hospitals).

OR components	Limited capacity, n (%)	Moderate capacity, n (%)	Strong capacity, n (%)
Leadership interest	0 (0)	0 (0)	10 (100)
Research-enabling environment	4 (40)	4 (40)	2 (20)
Data records	2 (20)	6 (60)	2 (20)
Data tools	4 (40)	3 (30)	3 (30)
Statistical proficiency	9 (90)	0 (0)	1 (10)
Data collection skills	3 (30)	1 (10)	6 (60)
Manuscript drafting skills	1 (10)	4 (40)	5 (50)
Basic research skills	0 (0)	5 (50)	5 (50)
Research dissemination skills	0 (0)	1 (10)	9 (90)
Total overall baseline capacity score	0 (0)	9 (90)	1 (10)

Recruitment

The partner hospitals included in the I-ORCB program were selected based on their prior participation in earlier phases of the ORCB program and interest in continuing their involvement. Additional eye hospitals were also invited, and those who were interested were subsequently enrolled. In total, 10 partner eye hospitals were recruited—5 from India and 5 from Nepal.

Each partner hospital was requested to form a core team comprising a hospital lead to provide leadership support and guide the selection of an OR project; a research lead to directly oversee the planning and implementation of an OR project; and a data lead to oversee data collection, management, and analysis. This core team then recruited supporting members from within the hospital to form an OR team.

Institutional and Participant Characteristics

The partner eye hospitals represent a mix of urban, semiurban, and rural settings across India and Nepal, serving both local and cross-border populations. These hospitals collectively provide a wide range of ophthalmic services, including cataract services and refractive error correction, as well as advanced subspecialties such as corneal and retinal surgeries, and community outreach activities. The hospitals operate through a network of vision centers, satellite clinics, and secondary hospitals, providing referral linkages and reaching both urban and rural populations.

Each hospital formed a multidisciplinary OR team comprising individuals from both clinical and nonclinical areas, such as ophthalmologists, outpatient department managers, research coordinators, optometrists, IT assistants, quality associates, fellows, pharmacists, cornea specialists, corporate social responsibility heads, and superintendents.

Data Collection Plan for Program Evaluation

Data will be collected to evaluate the 4 levels of the Kirkpatrick model.

For level 1, a structured postsession feedback form will be developed to assess participants' satisfaction with the training. This form will use a 5-point Likert scale to rate session content, delivery, discussions, and workshop logistics. Open-ended responses will capture participants' views on the usefulness of the session, need for additional resources or topics, and further suggestions.

For level 2, a pretest-posttest multiple-choice questionnaire (MCQ) will be designed for each session to measure knowledge gained. The questionnaires will be developed by a session facilitator and will be reviewed by team members to ensure content validity, clarity, and relevance.

Both level 1 and 2 tools will be developed in Google Forms. The MCQ form link will be circulated among the teams at least 1 hour prior to each session. Participants will be requested to provide the completed feedback form and postsession MCQs at the end of each session.

Level 3 will be evaluated through mentor reports that will be provided by the mentors of each team following completion of each training phase. These reports will document the extent to which the participants applied their learnings to OR activities.

Level 4 will be evaluated on completion at the end of phase 2. Each team will submit a final report describing how participation in the I-ORCB program supported the implementation of their OR project and contributed to organizational outcomes.

Data collection for levels 1, 2, and 3 will be conducted consistently across both training phases of the I-ORCB program.

Data Collection Plan for Evaluation of Program Cost-Effectiveness

The data to assess direct cost will be obtained from institutional records, such as invoices, facilitators' remuneration, and travel and logistics expenses related to the program. Indirect costs will be estimated based on the time spent by the facilitators in developing training materials and implementing the program (including time spent conducting online workshops and team meetings). Information on time use will be extracted from online session records, meeting logs, and activity schedules. The data to assess effectiveness will be obtained from participant responses to the baseline assessment compared against the end line assessment 6 months after completion of the 2 phases of workshops.

XSL•FO
RenderX

Data Analysis Plan for Program Evaluation

All quantitative data from levels 1 and 2 will be exported from Google Forms to Microsoft Excel. The dataset will be cleaned and coded prior to transfer into Stata (version 19; StataCorp) for statistical analysis. Descriptive statistics will be used to summarize participant feedback. The Likert scale responses will be reported as frequencies and percentages. For assessing knowledge, the responses will be scored dichotomously (correct=1; incorrect=0).

To compare pre- and posttest scores, a 2-tailed paired or unpaired *t* test will be applied if the data follow a normal distribution. For data not normally distributed, the Wilcoxon signed-rank test or Mann-Whitney *U* test will be used. The results will be reported with 95% CIs and *P* values. A *P* value <.05 will be considered statistically significant. Data will be checked for completeness and accuracy before analysis.

For levels 3 and 4, content analysis will be performed on the reports. The reports will be coded and grouped thematically to reflect learnings and applications of OR.

Data Analysis Plan for Evaluation of Program Cost-Effectiveness

Program cost-effectiveness will be assessed by calculating the cost-effectiveness ratio, that is, the ratio of total program cost to program effectiveness, using the formula cost-effectiveness ratio = total cost/effectiveness, where cost is the sum of direct and indirect costs and effectiveness is the mean difference between the baseline and end line assessment scores.

Ethical Considerations

The I-ORCB protocol has received ethics approval from the institutional ethics committee at Pushpagiri Vitreo Retina Institute, Hyderabad, (PVRI/IEC/2025/0009). Consent of participants is implied in the training program. The participants' right to privacy and confidentiality will be respected. All data will be anonymized prior to analysis and personal identifiers will not be disclosed while reporting the results. The data will be kept secure and will be accessible by authorized personnel only. The participants will receive training in operational research. There is no monetary compensation or financial incentive provided for participating in this training program.

Results

As of March 2025, a total of 10 teams have been enrolled in the I-ORCB program. This program is currently in the implementation phase, with teams undergoing training and ongoing mentorship. Data collection for program evaluation began in June 2025 and is expected to continue until late 2026. The results of this program are expected to be published in 2027. This study was funded in late 2024. Data analysis has not yet been initiated.

Discussion

Expected Findings

The I-ORCB program aims to enhance eye care services in low-resource settings by offering context-specific solutions to eye care providers, as well as to the ancillary health care staff. By equipping the participants with the knowledge and skills required to develop need-based solutions, the program supports evidence-based decision-making to improve eye health care delivery in LMICs. To the best of our knowledge, no other capacity-building initiative has focused exclusively on OR in eye care in the manner presented in this protocol.

In designing this program, we have drawn on a practical, output-based approach endorsed by the International Union Against Tuberculosis and Lung Disease and Médecins Sans Frontières, who have been the pioneers in building OR capacity among health professionals in public health programs in LMICs [3,14,15]. In addition, by adopting the research capacity building principles of the virtual model of the framework by Cooke [12], the I-ORCB program is grounded in proven methodologies and tailored to meet the challenges of delivering OR training across geographically diverse and resource-constrained settings. This design allows for both flexibility and inclusive learning opportunities [12,13].

OR has been widely applied to improve health care delivery across diverse clinical settings, such as dermatology, oncology, physiotherapy, and dialysis services [19-22]. These applications have demonstrated the value of OR in addressing common challenges related to service efficiency and resource use. Comprehensive eye care aims to provide a continuum of services, ranging from health promotion and prevention to treatment and rehabilitation [23]. Most of these services are delivered on an outpatient basis or as daycare procedures, which, in high volume settings, can lead to challenges in service delivery. OR applications have been shown to address these challenges through scheduling and optimizing patient flow; reducing waiting time; and, thus, improving patient satisfaction [24,25].

While earlier iterations of the ORCB program were promising, they also highlighted the need for deeper engagement. The I-ORCB program retains the core mission of building OR capacity in resource-limited settings and introduces new components to address the gaps identified in previous phases. The structured and spaced-out schedule of the I-ORCB workshops will enable participants to gain an in-depth understanding of key OR concepts while accommodating their demanding professional commitments. Long-term mentorship from dedicated professionals supports participants in real-world challenges in OR, and frequent one-on-one interactions maintain program momentum through the completion of an OR project. The e-resource repository gives participants the flexibility to access the training content at their own pace, accommodating different learning speeds and time constraints. By equipping eye personnel with OR skills, the I-ORCB program will contribute to the broader goal of improving eye health outcomes in LMICs.

The limitations of the I-ORCB program are, first, that the program heavily depends on an online mode for upskilling and mentorship. Therefore, the availability of stable internet connectivity is crucial and may affect participation and engagement. Second, although the workshops will be scheduled based on the convenience of the I-ORCB teams, participation

XSL•FO
RenderX

may still pose a challenge due to competing responsibilities. Third, the I-ORCB program will be implemented across diverse geographical and institutional settings, each with different levels of administrative support and resource availability. Due to these differences, the outcomes of the program may vary across sites.

Conclusions

The I-ORCB program offers a structured framework for building OR capacity in low-resource settings. By equipping eye care institutions with the skills needed to generate and use evidence, this program will contribute to the broader goal of improving eye health outcomes in LMICs.

Acknowledgments

The authors would like to acknowledge the commitment of all partner hospitals for their participation in the Institutional Operational Research Capacity Building program.

Funding

This training program will be funded by Seva Canada.

Authors' Contributions

All named authors have contributed sufficiently to the project to be included as authors. SA wrote the first draft of the manuscript. GVSM, KLB, PR, KJ, and SSG reviewed the manuscript.

Conflicts of Interest

None declared.

References

1. Zachariah R, Khogali M, Kumar AM, Harries AD, Reeder JC. Nationalizing operational research capacity building: necessity or luxury? Ann Glob Health 2020 Oct 20;86(1):136. [doi: 10.5334/aogh.3056] [Medline: 33134092]
2. Goel S. Prioritizing operational research in hospital settings: need of the hour. J Integr Med Public Health 2022;1(2):31-33. [doi: 10.4103/JIMPH.JIMPH_1_23]
3. Harries AD, Rusen ID, Reid T, et al. The Union and Médecins Sans Frontières approach to operational research. Int J Tuberc Lung Dis 2011 Feb;15(2):144-154. [Medline: 21219672]
4. Gilbert SS, Murthy G, Bassett KL, Operational Research Capacity Building Study Group bit.ly/ORCB-study. Building local capacity in operational research: a case study in Nepal and India. Community Eye Health 2022;35(117):14-16. [Medline: 37007828]
5. Zachariah R, Harries AD, Ishikawa N, et al. Operational research in low-income countries: what, why, and how? Lancet Infect Dis 2009 Nov;9(11):711-717. [doi: 10.1016/S1473-3099(09)70229-4] [Medline: 19850229]
6. Babu BS, Ravilla TD. Evidence-based management of eye care delivery. Community Eye Health 2018;31(102):S15-S16. [Medline: 30220820]
7. Bradley BD, Jung T, Tandon-Verma A, Khoury B, Chan TC, Cheng YL. Operations research in global health: a scoping review with a focus on the themes of health equity and impact. Health Res Policy Syst 2017 Apr 18;15(1):32. [doi: 10.1186/s12961-017-0187-7] [Medline: 28420381]
8. Shrestha M, Bhandari G, Kamalakannan S, et al. Evaluating the effectiveness of interventions to improve the follow-up rate for children with visual disabilities in an eye hospital in Nepal: nonrandomized study. JMIR Pediatr Parent 2023 Feb 23;6:e43814. [doi: 10.2196/43814] [Medline: 36821366]
9. Bhattarai B, Thapa HB, Bashyal S, et al. Structured Counselling and Regular Telephonic follow up to improve Referral flow and compliance in Nepal for Diabetic Retinopathy (SCREEN-D Study): a randomised controlled trial. BMC Health Serv Res 2024 Feb 10;24(1):188. [doi: 10.1186/s12913-024-10647-3] [Medline: 38336691]
10. Shrestha R, Singh P, Dhakwa P, et al. Augmenting the referral pathway for retinal services among diabetic patients at Reiyukai Eiko Masunaga Eye Hospital, Nepal: a non-randomized, pre-post intervention study. BMC Health Serv Res 2023 Feb 8;23(1):126. [doi: 10.1186/s12913-023-09105-3] [Medline: 36750897]
11. Thapa HB, Bajracharya K, Chaudhuri S, et al. Prevalence of refractive errors among under five year children attending in a tertiary eye care center of Nepal. J Coll Med Sci Nepal 2023;19(3):367-377. [doi: 10.3126/jcmsn.v19i3.55620]
12. Cooke J. A framework to evaluate research capacity building in health care. BMC Fam Pract 2005 Oct 27;6:44. [doi: 10.1186/1471-2296-6-44] [Medline: 16253133]
13. McGuire CM, Boskovic N, Fatusin BB, et al. Virtual health research capacity strengthening in low- and middle-income countries: a systematic integrative review. Ann Glob Health 2025;91(1):14. [doi: 10.5334/aogh.4543] [Medline: 40092966]
14. Bissell K, Harries AD, Reid AJ, et al. Operational research training: the course and beyond. Public Health Action 2012 Sep 21;2(3):92-97. [doi: 10.5588/pha.12.0022] [Medline: 26392960]

XSL•FO
RenderX

15. Ramsay A, Harries AD, Zachariah R, et al. The Structured Operational Research and Training Initiative for public health programmes. Public Health Action 2014 Jun 21;4(2):79-84. [doi: 10.5588/pha.14.0011] [Medline: 26399203]
16. Abdulghani HM, Shaik SA, Khamis N, et al. Research methodology workshops evaluation using the Kirkpatrick's model: translating theory into practice. Med Teach 2014 Apr;36 Suppl 1(sup1):S24-S29. [doi: 10.3109/0142159X.2014.886012] [Medline: 24617780]
17. Kirkpatrick DL, Kirkpatrick JD. Evaluating Training Programs: The Four Levels, 3rd edition: Berrett-Koehler Publishers; 2006. URL: https://books.google.co.in/books/about/Evaluating_Training_Programs.html?id=MiWMngEACAAJ&redir_esc=y [accessed 2026-03-31]
18. Dodani S, Songer T, Ahmed Z, Laporte RE. Building research capacity in developing countries: cost-effectiveness of an epidemiology course taught by traditional and video-teleconferencing methods in Pakistan. Telemed J E Health 2012 Oct;18(8):621-628. [doi: 10.1089/tmj.2011.0262] [Medline: 23061643]
19. Penneys NS. A comparison of hourly block appointments with sequential patient scheduling in a dermatology practice. J Am Acad Dermatol 2000 Nov;43(5 Pt 1):809-813. [doi: 10.1067/mjd.2000.107761] [Medline: 11050585]
20. Santibáñez P, Aristizabal R, Puterman ML, et al. Operations research methods improve chemotherapy patient appointment scheduling. Jt Comm J Qual Patient Saf 2012 Dec;38(12):541-553. [doi: 10.1016/s1553-7250(12)38071-9] [Medline: 23240262]
21. Harding KE, Bottrell J. Specific timely appointments for triage reduced waiting lists in an outpatient physiotherapy service. Physiotherapy 2016 Dec;102(4):345-350. [doi: 10.1016/j.physio.2015.10.011] [Medline: 26725373]
22. Fleming R, Gartner D, Padman R, James D. An analytical approach for improving patient-centric delivery of dialysis services. AMIA Annu Symp Proc 2019;2019:418-427. [Medline: 32308835]
23. Cicinelli MV, Marmamula S, Khanna RC. Comprehensive eye care - issues, challenges, and way forward. Indian J Ophthalmol 2020 Feb;68(2):316-323. [doi: 10.4103/ijo.IJO_17_19] [Medline: 31957719]
24. Abe TK, Beamon BM, Storch RL, Agus J. Operations research applications in hospital operations: part III. IIE Trans Healthc Syst Eng 2016 Jun;6(3):175-191. [doi: 10.1080/19488300.2016.1199613]
25. Zhang H, Ma WM, Zhu JJ, Wang L, Guo ZJ, Chen XT. How to adjust the expected waiting time to improve patient's satisfaction? BMC Health Serv Res 2023 May 8;23(1):455. [doi: 10.1186/s12913-023-09385-9] [Medline: 37158912]

Abbreviations

I-ORCB: Institutional Operational Research Capacity Building
LMIC: low- and middle-income country
MCQ: multiple-choice questionnaire
OR: operational research
ORCB: Operational Research Capacity Building
PRASHO: Pragyaan Sustainable Health Outcomes

Edited by S Khan; submitted 18.Sep.2025; peer-reviewed by P Dawson, S Brice; revised version received 04.Mar.2026; accepted 04.Mar.2026; published 13.Apr.2026.

Please cite as:
Ali S, Murthy GVS, Bassett KL, Reddy P, Judson K, Gilbert SS, Patra A, Gudlavalleti AG, Mahapatra D, Pant H, Dakhwa P, Jameel S, Pandey SP, Tetali S, Chaudhuri S, Imandi S, Kamalakannan S
Building Capacity in Institutional Operational Research in Low-Resource Settings: Protocol for an Implementation Research Study
JMIR Res Protoc 2026;15:e84284
URL: https://www.researchprotocols.org/2026/1/e84284
doi:10.2196/84284

© Sana Ali, G V S Murthy, Kenneth L Bassett, Priya Reddy, Katie Judson, Suzanne Schwartz Gilbert, Abhilash Patra, Anirudh Gaurang Gudlavalleti, Debangana Mahapatra, Hira Pant, Parami Dakhwa, Sarah Jameel, Sarva Priya Pandey, Shailaja Tetali, Sirshendu Chaudhuri, Srinivas Imandi, Suresh Kamalakannan. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Sickle Cell Disease at a Tertiary Care Center in the Vidarbha Region of India: Protocol for a Clinical and Observational Study

Tanushree Budhbaware, MSc; Jaishriram Rathored, PhD

Department of Central Research Laboratory and Molecular Diagnostics, Datta Meghe Institute of Higher Education and Research, Wardha, Maharashtra, India

Corresponding Author:
Jaishriram Rathored, PhD
Department of Central Research Laboratory and Molecular Diagnostics, Datta Meghe Institute of Higher Education and Research, Wardha, Maharashtra, India

Abstract

Background: More than 20 million individuals worldwide, especially in the Vidarbha region of India, are affected by sickle cell anemia (SCA), a hereditary condition that results in aberrant hemoglobin S and red blood cell distortion. The condition leads to anemia, organ complications, and recurrent pain crises, making region-specific data necessary for efficient therapy and public health initiatives.

Objective: The goal of the study is to examine the clinical characteristics and unusual manifestations of SCA in the Vidarbha region, with an emphasis on dietary practices, clinical presentations, demographic distribution, and lifestyle factors such as alcohol consumption and smoking.

Methods: This observational cross-sectional study with random sampling will be conducted at Acharya Vinoba Bhave Rural Hospital in Wardha for 3 months. We will recruit 131 individuals aged 18 to 50 years with dominant hemoglobin S and a positive sickling test. A standardized questionnaire addressing clinical symptoms, nutrition, substance use, inheritance patterns, and demographic information will be used to gather data. SPSS (version 17; IBM Corp) will be used for statistical analysis. Data will be summarized using descriptive statistics. Group differences will be evaluated using inferential tests such as 1-way ANOVA, independent 2-tailed *t* tests, and chi-square tests. Associations between symptoms and lifestyle variables will be investigated by correlation analysis. Statistical significance is defined as a 2-tailed *P* value <.05.

Results: The anticipated findings may support the need for targeted regional public health initiatives and underscore the importance of comprehensive screening, detailed patient history, and tailored care strategies for individuals with SCA. As of January 2026, this observational study has not received external funding. Participant recruitment and data collection commenced in January 2026 and are currently ongoing. Data analysis will be undertaken following completion of data collection, and the final results are expected to be submitted for publication in April 2026.

Conclusions: The findings will support the need for focused regional public health initiatives and emphasize the need for thorough screening, patient history, and customized care techniques for SCA.

International Registered Report Identifier (IRRID): PRR1-10.2196/80483

(JMIR Res Protoc 2026;15:e80483) doi:10.2196/80483

KEYWORDS

red blood cells disorder; genetic abnormality; hemoglobin S; sickle cell disease; sickle cell anaemia

Introduction

In 1910, the first clinical case of sickle cell anemia (SCA) in the United States was reported by Dr John Herrick [1]. SCA is a health condition affecting more than 20 million people worldwide [2].

The main cause of SCA is a genetic mutation that affects hemoglobin, the protein in red blood cells that carries oxygen [3]. A point mutation in the *HBB* gene on chromosome 11 is the underlying cause of SCA, as it leads to the production of aberrant hemoglobin, known as hemoglobin S (*HBS*) [4]. The particular mutation involves the replacement of a single nucleotide; namely, thymine replaces adenine. This results in the substitution of valine for glutamic acid at the sixth position of the beta-globin chain. This change alters hemoglobin structure and function, increasing its tendency to polymerize under conditions such as low oxygen tension or dehydration [5]. Due to the autosomal recessive inheritance pattern of SCA [6], an individual needs to inherit 2 copies of the mutated gene, 1 from each parent, to develop the disease. Individuals who inherit 1 mutated gene and 1 normal gene have the sickle cell trait and are usually asymptomatic but can transmit the trait to their offspring [7]. The molecular makeup of HbS is different from

XSL•FO
RenderX

that of normal adult hemoglobin A (HBA). Red blood cells deform and have less ability to pass through tiny blood vessels as a result of HbS, which makes them sticky and stiff [8]. Alterations in red blood cell morphology are associated with multiple SCA-related health complications [9]. Polymerization of HbS causes red blood cells to adopt the characteristic sickle shape, which is rigid and prone to hemolysis [10]. These sickled red blood cells clog blood vessels, leading to tissue ischemia, vaso-occlusion, and subsequent organ damage [10].

A significant health burden is caused by SCA, which is characterized by high morbidity and reduced life expectancy, in the Vidarbha region of India [11]. Research and intervention are crucial in this region, including Wardha, because of the high incidence of patients with SCA. The high frequency of SCA in the Vidarbha area enables the study of a large population and the generation of useful region-specific data. The genetic composition of the local population may offer special difficulties and insights into the development and treatment of the disease. The area has unique health care issues, such as restricted access to specialized care; therefore, creating customized health care plans is essential [10].

This study will be conducted to evaluate the observational and clinical profile of patients with SCA attending the Acharya Vinoba Bhave Rural Hospital, Wardha. The findings may underscore the importance of comprehensive screening and detailed patient history in accurately diagnosing and managing the disease. The clinical profiles, characterized by recurrent pain crises, anemia, and potential organ complications, may highlight the need for ongoing monitoring and tailored treatment strategies, with inclusion and exclusion criteria shown in Textbox 1. Observational data may reveal significant patterns in disease prevalence and severity, reinforcing the necessity for targeted public health interventions and continuous support for affected individuals and their families. The purpose of the study is to describe the clinical characteristics and unusual manifestations of SCA and to find associations with lifestyle and demographic factors among people in the Vidarbha area. Specifically, the study aims to examine the age, gender, and ethnic distribution of people with SCA; document and examine their clinical presentations; evaluate their dietary habits; and investigate the connection between clinical symptoms and substance abuse, especially alcohol consumption and tobacco use.

Textbox 1. Inclusion and exclusion criteria for the study.

Inclusion criteria

- Patients with a positive sickling test and dominant sickle hemoglobin (HbSS)
- Patients in the age group of 18 to 50 years

Exclusion criteria

- Age above 50 years
- Other genetic abnormalities
- Individuals with immunocompromising conditions, including HIV infection, organ transplantation, or ongoing chemotherapy.

Methods

Overview

This prospective observational cross-sectional study will be conducted for 3 months in the Unit of Sickle Cell Anemia at Acharya Vinoba Bhave Rural Hospital, Datta Meghe Institute of Higher Education and Research, Sawangi, Wardha. Data collection will involve administering a structured questionnaire to gather information on age, gender, area of residence, ethnicity, inheritance patterns, existing symptoms, dietary habits, and history of smoking and alcohol consumption. Objective clinical data, including hemoglobin levels, previous hospitalization records, and documented complications, will be confirmed by reviewing participant hospital case sheets and laboratory reports (as permitted by the consent form), despite the fact that the study does not involve blood or tissue sampling. This approach minimizes the possibility of bias from self-reported responses and ensures data correctness. The collected data will be analyzed to identify rare symptoms of SCA and correlate these findings with various demographic and lifestyle parameters (Figure 1).

XSL•FO
RenderX

Figure 1. Participant recruitment flowchart (N=131).

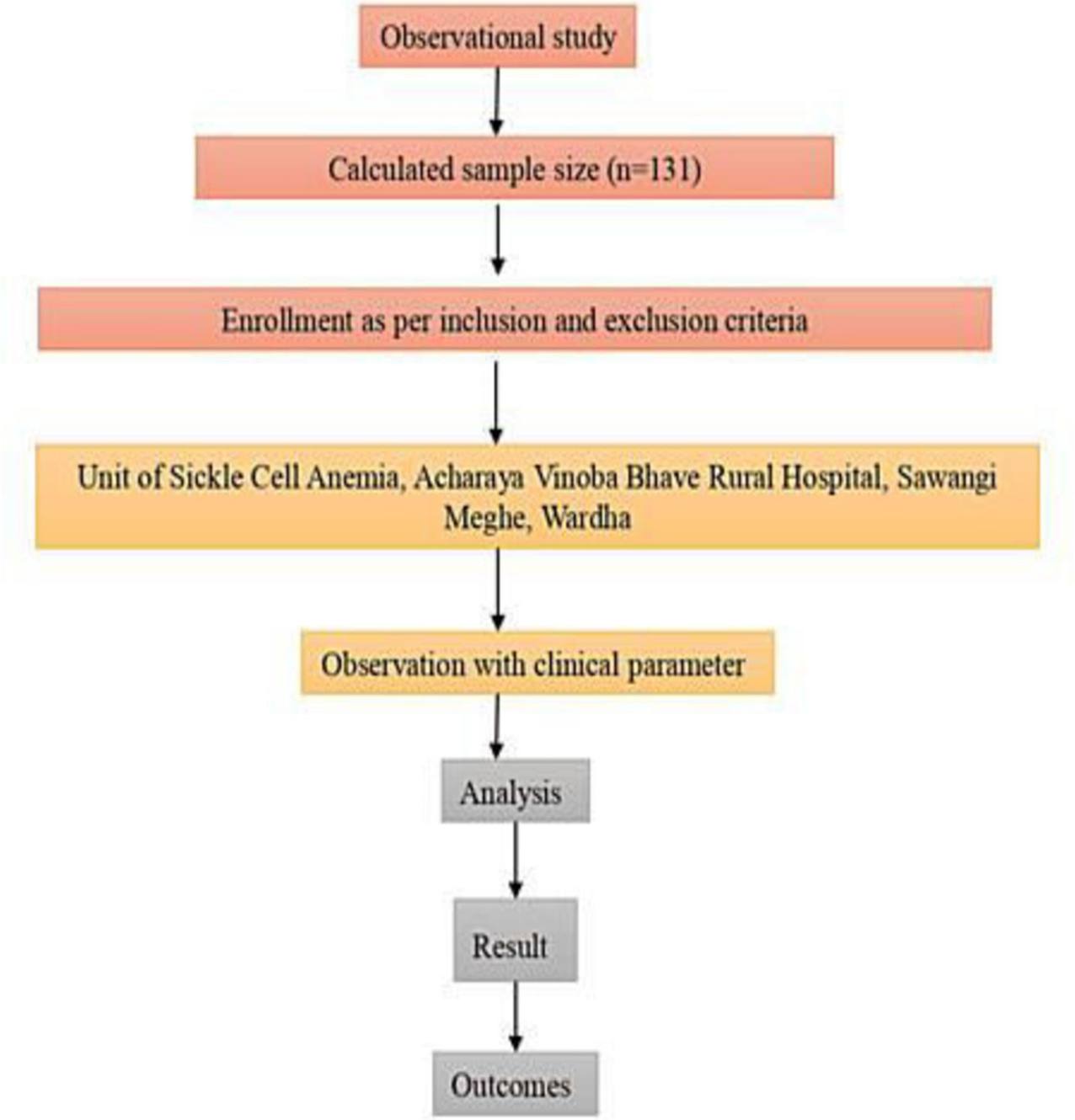

Participants

Participants will be consecutively recruited from the Sickle Cell Outpatient Department of Acharya Vinoba Bhave Rural Hospital for 3 months, using convenience sampling to include all eligible patients aged 18 to 50 years, with the aim of reducing age-related comorbidities. A sample size of 131 was determined based on a regional prevalence of 9.4%, although no separate power analysis for group comparisons was conducted, which is acknowledged as a limitation. Descriptive statistics will summarize demographic and clinical variables, while inferential tests (*t* tests, ANOVA, chi-square test, and correlation analyses) will be used to examine bivariate associations. Multivariable regression analysis will be conducted to adjust for confounders such as age, gender, and lifestyle factors. The inclusion criterion "dominant HbS" will be clarified to differentiate between homozygous (HbSS) and heterozygous (HbAS) genotypes, in line with the autosomal recessive inheritance of SCA.

Data Collection Procedure

After informed consent has been obtained and the study's goal has been explained, data will be gathered from patients with SCA who visit the Department of Sickle Cell. Trained interns will administer a structured questionnaire to document demographic and lifestyle factors, such as age, gender, location, ethnicity, inheritance patterns, current symptoms, dietary practices, and history of smoking and alcohol use. The questionnaire's content validity and clarity were pretested.

XSL•FO
RenderX

Cronbach α will be used to evaluate internal consistency (≥0.7 is considered acceptable), and test-retest reliability will be assessed through a 2-week follow-up with approximately 10% of participants. These metrics attest to the tool's validity and usefulness for clinical and observational sickle cell disease evaluation at the tertiary care facility in Vidarbha. Responses will be documented on paper, and each interview will last for approximately 15 minutes. Records will be stored securely, access will be limited, and they will be disposed of properly after study completion.

Sample Size

The sample size was determined using the following formula:

n = \frac{Z^{2} \times p \times q}{d^{2}}]

Where n = required sample size, Z = standard normal deviation at a 95% confidence level (1.96), p = estimated prevalence of the condition, q = 1−p1 - p1−p, and d = allowable error (precision).

Outcomes

Primary Outcome

Assessing the clinical features of patients with SCA will be the main goal of this research, with an emphasis on the frequency and severity of symptoms, including anemia, organ complications, and recurrent pain crises. This outcome will demonstrate the importance of careful screening and a thorough patient history in accurately diagnosing and treating the disease.

Secondary Outcome

To support affected individuals and their families, the secondary outcome will involve analyzing observational data to identify significant trends in disease prevalence and severity. This analysis will highlight the need for ongoing monitoring, customized treatment plans, and focused public health initiatives.

Statistical Methodology

SPSS Statistics for Windows (version 17.0; IBM Corp) will be used for statistical analysis. To summarize demographic and clinical features, descriptive statistics such as mean, SD, median, and frequency distributions will be computed. For inferential analyses, chi-square tests will be used for categorical variables and independent 2-tailed *t* tests or 1-way ANOVA will be used for continuous variables, as applicable. To assess associations between symptoms and nutritional or drug use habits, correlation analyses will be used. 95% CIs will be provided for a 2-tailed *P* value <.05, which is deemed statistically significant. Depending on the amount and type of missing values, missing or incomplete data will be handled using the proper statistical approaches, such as case-wise exclusion or imputation procedures.

Ethical Considerations

This is an observational study without collection of any type of blood or tissue samples. We received institutional ethics committee approval from Datta Meghe Institute of Higher Education and Research (DMIHER(DU)/IEC/2024/155). Informed consent is sought from all participants, and the participants are informed that they can opt out. All personal information is kept completely confidential, and coded IDs are used to record participant data to guarantee anonymity. To preserve participant anonymity and guarantee adherence to ethical guidelines, all gathered data will be safely kept on password-protected devices that will only be accessible by the study team.

Results

The study is expected to identify and characterize clinical features of sickle cell anemia, particularly among symptomatic individuals, with attention to severe or uncommon presentations. Recurrent episodes of pain, anemia, and possible organ damage, all suggestive of the disease's typical vaso-occlusive crises, may be observed. A thorough patient history that will address familial incidence, symptom frequency, and previous medical interventions may prove crucial for making an accurate diagnosis. This all-encompassing strategy will guarantee a complete comprehension and handling of SCA, underscoring the importance of clinical presentation and patient history in successful disease detection and therapy. As of January 2026, this observational study has not received external funding. Participant recruitment and data collection commenced in January 2026 and are currently ongoing. Data analysis will be undertaken following completion of data collection, and the final results are expected to be submitted for publication in April 2026.

Discussion

Anticipated Findings

Sickle cell anemia (SCA) affects over 20 million people worldwide [2], and was first described in 1910 by Dr. John Herrick [1]. It is caused by a mutation in the *HBB* gene on chromosome 11, involving a single nucleotide substitution in which thymine replaces adenine, leading to the replacement of glutamic acid with valine at the sixth position of the β-globin chain [4]. This structural alteration produces hemoglobin S (HbS), which polymerizes under conditions of low oxygen tension or dehydration, deforming red blood cells into a sickle shape and causing severe complications. SCA follows an autosomal recessive inheritance pattern; individuals with one mutated gene carry the sickle cell trait and are typically asymptomatic but can transmit the disorder [7]. The disease is highly prevalent in regions of Africa, India, the Middle East, and the Mediterranean [8]. In India, particularly in the Vidarbha region, including Wardha, SCA represents a major public health burden associated with high morbidity and reduced life expectancy [11], making it a critical area for research and intervention. The high prevalence of SCA in Vidarbha provides a substantial population for study and yields meaningful data. The genetic makeup of the population in this region may present unique challenges and insights into the disease's progression and management. The region faces specific health care challenges, including limited access to specialized care, making it vital to develop tailored health care strategies [10].

Sickle-shaped red blood cells can cause a range of problems. They can get trapped in small blood vessels, blocking the flow of blood and causing pain and organ damage. Normal red blood

XSL•FO
RenderX

cells survive for about 120 days, whereas sickled cells only last about 10 to 20 days. This leads to a chronic shortage of red blood cells, known as anemia [12]. Blocked blood flow can cause episodes of severe pain, known as pain crises or vaso-occlusive crises. These can occur without warning and may require hospitalization. People with SCA have a higher risk of infections because sickle cells can damage the spleen, an organ that helps fight infections. Some common clinical manifestations include fatigue, weakness, and shortness of breath due to a lack of healthy red blood cells; sudden, severe pain in different parts of the body, often in the bones, chest, and joints; and painful swelling, known as dactylitis, is often the first symptom in infants. Children with SCA may experience delayed growth compared with their peers and attain puberty at an older age. Sickle cells can block the tiny blood vessels that supply the eyes, leading to vision issues. Blocked blood flow to the brain can cause a stroke, a serious complication of SCA. Red blood cells with aberrant HbS change in morphology, becoming sticky and rigid, which can cause a number of health issues. The characteristic sickle-shaped deformation of red blood cells is associated with hemolysis and occlusion of small vessels, leading to tissue ischemia, organ damage, and vaso-occlusion [8]. Although there is no universal cure for SCA, several treatments can help manage symptoms and reduce complications. Pain relief is a primary treatment goal. Medications range from over-the-counter pain relievers to stronger prescription drugs. Folic acid (1 mg) and hydroxyurea (500 mg) help reduce the frequency of pain crises and the need for blood transfusions. It works by increasing the production of fetal hemoglobin, which prevents sickling of red blood cells [13]. Regular blood transfusions can help reduce the risk of stroke and other complications by increasing the proportion of normal red blood cells [14]. While hematopoietic stem cell transplantation offers the only potential for cure, its implementation is limited by the high risk of serious complications. This is currently the only potential cure for SCA. It involves replacing the patient's bone marrow with healthy marrow from a donor, but it is only suitable for some patients due to the risks involved. Comprehensive preventive care, including vaccination, prophylactic antibiotics, maintenance of adequate hydration, avoidance of extreme temperatures, and routine medical monitoring, is fundamental to improving patient outcomes [15]. Some promising areas of research for SCA include gene therapy approaches, new medications, and genome editing technologies, such as CRISPR-Cas9 [16]. The pathophysiological mechanisms underlying the clinical manifestations of SCA underscore the need for thorough screening, customized care, and ongoing monitoring to effectively manage the disease [17].

Limitations

It is important to acknowledge the limitations of this study. Recruiting individuals from a single institution may limit the generalizability of the findings, and the short study duration may make it more difficult to discover long-term effects. Furthermore, there may be selection bias because the research sample may not be representative of the entire eligible population. Recall bias resulting from self-reported eating habits may potentially impact the quality and reliability of the dietary data. These considerations should be taken into account by researchers when assessing the results. Furthermore, there is a chance of recall or reporting bias because the majority of the data collection will be done through self-reported questionnaires; although, attempts will be made to cross-check participant medical records and laboratory data to improve reliability and reduce the possibility of clinical outcomes being misclassified.

Conclusions

This investigation may offer crucial insights into the clinical spectrum of SCA and underscore the necessity for customized interventions in the Vidarbha region. The findings highlight the importance of comprehensive screening and a detailed patient history in accurately diagnosing and managing the disease. The clinical profiles, characterized by recurrent pain crises, anemia, and potential organ complications, emphasize the need for ongoing monitoring and tailored treatment strategies. Observational data may reveal significant patterns in disease prevalence and severity, reinforcing the necessity for targeted public health interventions and continuous support for affected individuals and their families.

Acknowledgments

The authors thank Dr Snehal Patil of Acharya Vinoba Bhave Hospital, and Dipali Joshi and Mithilesh Kumar of the central research laboratory in Sawangi, Wardha, Maharashtra, and all the staff for their guidance and support throughout the project.

All authors declared that they had insufficient or no funding to support open access publication of this manuscript, including from affiliated organizations or institutions, funding agencies, or other organizations. JMIR Publications provided article processing fee (APF) support for the publication of this paper.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

TB was responsible for conceptualization, data acquisition, and writing of the original draft. JR was responsible for conceptualization, methodology, and editing. Both authors agree to be accountable for all aspects of this study.

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Savitt TL, Goldberg MF. Herrick's 1910 case report of sickle cell anemia. The rest of the story. JAMA 1989 Jan 13;261(2):266-271. [Medline: 2642320]
2. Royal CDM, Babyak M, Shah N, et al. Sickle cell disease is a global prototype for integrative research and healthcare. Adv Genet (Hoboken) 2021 Mar;2(1):e10037. [doi: 10.1002/ggn2.10037] [Medline: 36618444]
3. Brandow AM, Liem RI. Advances in the diagnosis and treatment of sickle cell disease. J Hematol Oncol 2022 Mar 3;15(1):20. [doi: 10.1186/s13045-022-01237-z] [Medline: 35241123]
4. Inusa BPD, Hsu LL, Kohli N, et al. Sickle Cell Disease-Genetics, Pathophysiology, Clinical Presentation and Treatment. Int J Neonatal Screen 2019 Jun;5(2):20. [doi: 10.3390/ijns5020020] [Medline: 33072979]
5. Elendu C, Amaechi DC, Alakwe-Ojimba CE, et al. Understanding sickle cell disease: causes, symptoms, and treatment options. Medicine (Baltimore) 2023 Sep 22;102(38):e35237. [doi: 10.1097/MD.0000000000035237] [Medline: 37746969]
6. Tisdale JF, Thein SL, Eaton WA. Treating sickle cell anemia. Science 2020 Mar 13;367(6483):1198-1199. [doi: 10.1126/science.aba3827] [Medline: 32165573]
7. Ashley-Koch A, Yang Q, Olney RS. Sickle hemoglobin (HbS) allele and sickle cell disease: a HuGE review. Am J Epidemiol 2000 May 1;151(9):839-845. [doi: 10.1093/oxfordjournals.aje.a010288] [Medline: 10791557]
8. Diez-Silva M, Dao M, Han J, Lim CT, Suresh S. Shape and biomechanical characteristics of human red blood cells in health and disease. MRS Bull 2010 May;35(5):382-388. [doi: 10.1557/mrs2010.571] [Medline: 21151848]
9. Du E, Dao M. Faster sickling kinetics and sickle cell shape evolution during repeated deoxygenation and oxygenation cycles. Exp Mech 2019 Mar;59(3):319-325. [doi: 10.1007/s11340-018-00444-5] [Medline: 31178599]
10. Goetz LH, Schork NJ. Personalized medicine: motivation, challenges, and progress. Fertil Steril 2018 Jun;109(6):952-963. [doi: 10.1016/j.fertnstert.2018.05.006] [Medline: 29935653]
11. Andankar A, Ghodake A, Gujrathi A, Medicine C. MDS-952: Mapping the burden of sickle cell anemia in the tribal region of eastern Maharashtra, India: clinical profile, genotypic patterns, and spousal counseling insights from a 15-year (2009–2024) retrospective study. Clin Lymphoma Myeloma Leuk 2025 Sep;25:S624. [doi: 10.1016/S2152-2650(25)02047-6]
12. Sachdev V, Rosing DR, Thein SL. Cardiovascular complications of sickle cell disease. Trends Cardiovasc Med 2021 Apr;31(3):187-193. [doi: 10.1016/j.tcm.2020.02.002] [Medline: 32139143]
13. Díaz-Matallana M, Márquez-Benítez Y, Martínez-Lozano JC, Briceño-Balcázar I, Benavides-Benítez E, Bernal JE. Disease genotype, haplotypes, diagnosis and associated studies in sickle cell anemia. Rev Med Chil 2021 Sep;149(9):1322-1329. [doi: 10.4067/S0034-98872021000901322] [Medline: 35319686]
14. Fortin PM, Hopewell S, Estcourt LJ. Red blood cell transfusion to treat or prevent complications in sickle cell disease: an overview of Cochrane reviews. Cochrane Database Syst Rev 2018 Aug 1;8(8):CD012082. [doi: 10.1002/14651858.CD012082.pub2] [Medline: 30067867]
15. Parums DV. Editorial: First regulatory approvals for CRISPR-Cas9 therapeutic gene editing for sickle cell disease and transfusion-dependent β-thalassemia. Med Sci Monit 2024 Mar 1;30:e944204. [doi: 10.12659/MSM.944204] [Medline: 38425279]
16. Park SH, Bao G. CRISPR/Cas9 gene editing for curing sickle cell disease. Transfus Apher Sci 2021 Feb;60(1):103060. [doi: 10.1016/j.transci.2021.103060] [Medline: 33455878]
17. Galloway SJ, Harwood-Nuss AL. Sickle-cell anemia--a review. J Emerg Med 1988;6(3):213-226. [doi: 10.1016/0736-4679(88)90329-0] [Medline: 3049776]

Abbreviations

HbA: hemoglobin A
HbS: hemoglobin S
SCA: sickle cell anemia

Edited by J Sarvestan; submitted 11.Jul.2025; peer-reviewed by B Keikhaei, H Rajesh; revised version received 28.Oct.2025; accepted 30.Oct.2025; published 21.Jan.2026.

Please cite as:
Budhbaware T, Rathored J
Sickle Cell Disease at a Tertiary Care Center in the Vidarbha Region of India: Protocol for a Clinical and Observational Study
JMIR Res Protoc 2026;15:e80483
URL: https://www.researchprotocols.org/2026/1/e80483
doi:10.2196/80483

XSL•FO
RenderX

© Tanushree Budhbaware, Jaishriram Rathored. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Play and Supportive Environments Among Older Adults in the Context of Swedish Municipal Care: Protocol for the Development and Psychometric Evaluation of an Instrument to Measure Playfulness

Albert Westergren[1], Prof Dr; Anna Bergman[1], MSc; Sofi Knaust Lidén[1], MSc; Catarina Nahlén Bose[2], PhD; Anna Nivestam[1], PhD; Annika Norell[3], PhD; Petra Nilsson Lindström[1], PhD; Maria Haak[1], Prof Dr

[1]Faculty of Health Sciences, Kristianstad University, SE-291 88, Kristianstad, Sweden

[2]Department of Health Sciences, Swedish Red Cross University, Stockholm, Sweden

[3]School of Behavioural, Social and Legal Sciences, Örebro University, Örebro, Sweden

Corresponding Author:
Albert Westergren, Prof Dr
Faculty of Health Sciences, Kristianstad University, SE-291 88, Kristianstad, Sweden

Abstract

Background: Playfulness—being and acting playful—is often associated with childhood, yet evidence suggests that it remains a meaningful resource throughout life. In later life, playfulness may support social connectedness, emotional well-being, and a sense of agency, even in contexts of illness or institutional living. Playfulness encompasses not only observable playful activities but also an inner disposition, such as curiosity, humor, or spontaneity, which may be constrained by environmental barriers, aging, or functional limitations. Despite its potential relevance for health and person-centered care, playfulness remains underexplored in gerontological and caregiving research. No validated instrument currently exists to assess playfulness among older adults in Swedish municipal care. This research program addresses this gap by clarifying the concepts of play and playfulness and by developing and psychometrically evaluating a new instrument, Play and Supportive Environments (PLAY-SE).

Objective: The overall aim of the program is to clarify and operationalize the interrelated concepts of playfulness and playful activities among older adults receiving municipal care and to develop an instrument suitable for psychometric testing.

Methods: The program applies a hybrid model of concept development combined with an exploratory sequential mixed methods design. Phase 1 involves literature reviews and qualitative studies with older adults and staff to explore and define the lived meanings of playfulness. These findings inform item generation and refinement of the PLAY-SE instrument. Phase 2 includes content validation, cognitive interviews, pilot testing, and large-scale psychometric evaluation using both classical test theory and Rasch measurement theory.

Results: Two PhD students were recruited to the program in September 2024 and September 2025, and an expert group was established in autumn 2025. The PhD students are funded, for four years each, by Kristianstad University (from 2024) and Red Cross University College (from 2025). Ethical approval for the qualitative studies in phase 1 was granted by the Swedish Ethical Review Authority (2025-00211-01; decision date: February 3, 2025). Data collection for qualitative interviews with older adults in municipal care was conducted between February and April 2025. Fifteen older adults (aged 68 - 100 y) were interviewed in nursing home settings. The phenomenological findings from phase 1.1 have been published in March 2026. Additional qualitative interviews and focus groups with staff are scheduled for 2026 to 2027. Pilot testing of the first version of the PLAY-SE instrument is planned for autumn 2026, followed by large-scale psychometric validation between 2027 and 2029.

Conclusions: This program establishes a structured and theoretically grounded process for developing and validating an instrument to assess playfulness in later life. By integrating qualitative exploration with modern psychometric approaches, the PLAY-SE instrument is intended to support future research and contribute to the development of person-centered practices in municipal elder care.

International Registered Report Identifier (IRRID): PRR1-10.2196/86373

(JMIR Res Protoc 2026;15:e86373) doi:10.2196/86373

KEYWORDS

concept development; existential philosophy; instrument development; literature review; mixed methods; municipal care; older adults; person-centered care; playfulness; psychometrics; lifeworld theory

XSL•FO
RenderX

Introduction

Background

Playfulness (ie, being and acting playful) is increasingly recognized as important for older adults' health, well-being, and dignity, even in contexts of illness or institutional care. Yet, research and practice still lack validated instruments that capture this phenomenon from the perspective of older adults themselves. In this study protocol, we describe a research program organized into 2 interrelated phases (Figure 1). The first phase explores the lived experience and meanings of playfulness in later life, while the second focuses on the development and psychometric evaluation of a new instrument (Play and Supportive Environments [PLAY-SE]). To facilitate this process, the protocol provides conceptual clarifications of playfulness and play, as well as an initial conceptualization of the domains to be included in the instrument.

Figure 1. Overview of the Play and Supportive Environments program, the study protocol, and the 2-phase research design. Three cycles are applied iteratively to deepen both the conceptual understanding and the operationalization of playfulness.

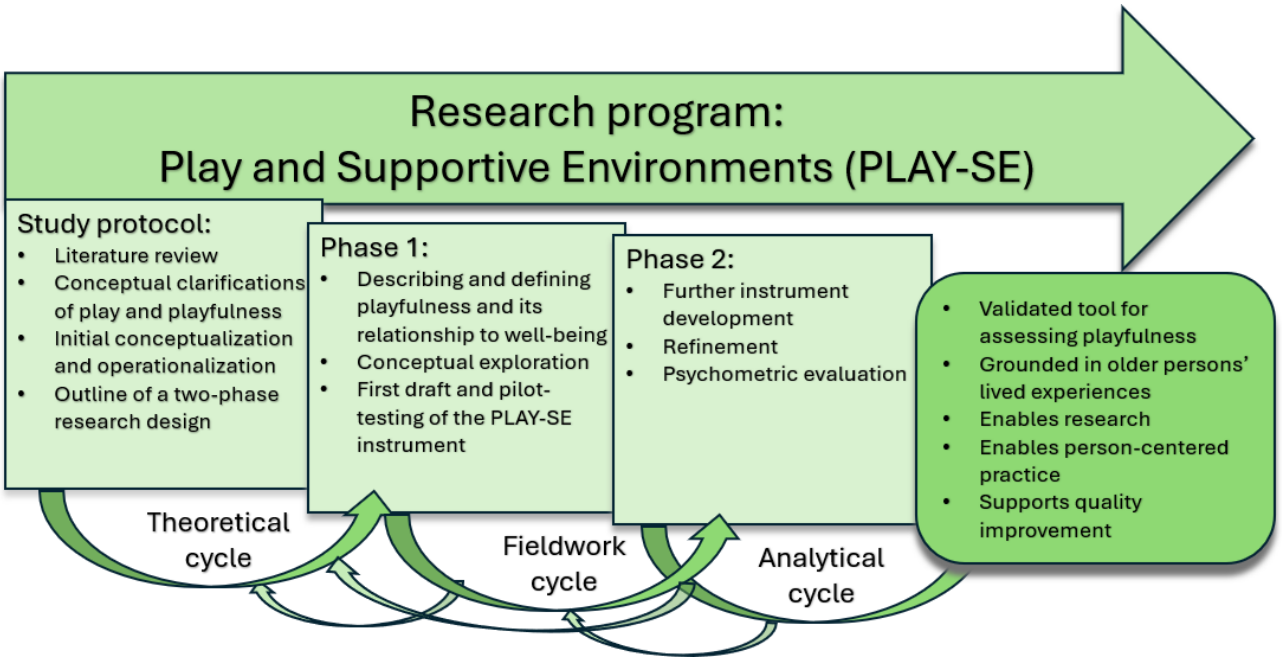

Although being playful and acting playfully are often associated with childhood, studies show that playfulness remains a stable trait across the lifespan, with older adults not inherently less interested in playful engagement [1-3]. Nevertheless, dominant life-course frameworks have traditionally linked playfulness to childhood, socialization to adolescence, work to adulthood, and passive leisure to old age, which has contributed to limited research attention on playfulness in later life [4]. Expanding this view is important, as playfulness in later life has been linked to well-being, creativity, cognitive health, and social connectedness [4,5].

Being and acting playfully are closely tied to a person's sense of identity, purpose, and relational connectedness [6], core components of person-centered care as articulated by McCormack and McCance [7]. Their framework emphasizes 4 modes of being: being with self, being in relation, being in a social world, and being in place. Engaging in playful activities or expressing a playful attitude may facilitate all these dimensions, thereby supporting the dignity, autonomy, and well-being of older adults in care. In contexts where daily routines are often shaped by institutional or structured living environments, playfulness can represent a moment of agency (ability to act independently, make choices) and self-expression [8,9].

Being and acting playful can also be understood through an existential lens. According to Merleau-Ponty [10], the body is the primary site of being in the world. In situations where mobility or independence is compromised, opportunities for unreflective participation in playful activities may diminish. Nonetheless, playfulness as an attitude or mode of being may remain accessible and vital, offering continuity and meaning even during times of change. This perspective reinforces the importance of enabling older adults to express their playfulness, even within the constraints of care environments.

In the context of municipal care, older adults living in long-term care facilities often report limited access to meaningful, self-chosen activities and social engagement. Both a Swedish report [11] and international studies [12,13] highlight widespread boredom, monotony, and passivity—conditions that stand in contrast to legal aims for an active and meaningful life in the community [14]. At the same time, international research suggests that playfulness may contribute to physical, emotional, and social well-being throughout the lifespan [15-17].

Literature Review

A structured exploratory literature review was conducted in May 2025 using the research platform Elicit [18], which applies large-scale paper search and screening based on predefined inclusion criteria. The guiding research question was: "What perspectives do older adults (65 years or older) have on playfulness and play, and what measures have been developed with a special focus on playfulness and play for older adults?

XSL•FO
RenderX

In addition, what dimensions of play and/or playfulness are captured by these measures?" The search strategy included both the terms *playfulness* and *play*, since playfulness is often described as a state characterized by elements of play [19].

The search was conducted within the Semantic Scholar corpus (>126 million academic papers) and retrieved the 50 most relevant records, which were screened according to population age, study focus, study design, and relevance. Ten studies were initially included from this screening process. In addition, 3 studies were identified through backward citation tracking. The final sample comprised qualitative interview and participatory design studies [8,9,20-22], quantitative survey-based investigations [2,3,23,24], instrument development studies [25-27], and 1 narrative review [28]. From these, 3 validated instruments assessing playfulness in older adults were identified: the Older Adult Playfulness Scale (OAP) [27], the Other-Directed, Light-Hearted, Intellectual, and Whimsical (OLIW) Playfulness Scale [23], and the Short Measure of Adult Playfulness (SMAP) [2]. The following sections summarize the findings from this review, beginning with perspectives on play and playfulness among older adults and continuing with validated measures used in this field.

Older adults describe play and playfulness as multidimensional constructs encompassing cognitive, emotional, social, and psychological aspects. These dimensions are associated with identity, continuity, and vitality in later life [2,8,20]. For example, in an interview study of older adults aged 62 to 89 (mean age 74 y), play was described as a source of enjoyment, structure, and personal meaning in everyday life [9].

Studies link playfulness to well-being, showing that it helps older adults reframe everyday challenges, maintain social ties, and support mental and physical health. In a study of persons aged 50 to 98 years, higher playfulness was associated with greater life satisfaction and flourishing [3]. Similarly, a survey study involving participants aged 60 years and older found that self-reported playfulness, especially in social contexts, was positively correlated with well-being [23].

Recent research has further strengthened this empirical foundation. For example, Brauer et al [3] demonstrated positive associations between playfulness and life satisfaction, character strengths, and flourishing in middle and older age, while Clifford et al [24], in a general adult sample, found that playfulness was linked to adaptive coping and reduced stress. Although not restricted to older adults, these findings reinforce the relevance of playfulness as a health-related resource across adulthood and into later life.

Qualitative studies also shed light on the significance of play. Burr et al [8], through in-depth interviews with older adults (aged 69 - 88 y, mean age 81 y, SD 6.06 y), emphasized play's role in shaping life narratives and creating meaning. Schutter and Abeele [20,21] used contextual inquiries and participatory design methods with participants aged 68 to 80 years and 50 to 72 years to explore how older adults perceive meaningful play, particularly in digital contexts. These studies found that play supports connectedness, nurtures the self and others, and enables meaningful societal engagement through shared experiences and intergenerational interaction.

Play in digital environments has also been explored through ethnographic and participatory design studies. Sayago et al [22] demonstrated that older adults (aged 60 - 85 y) actively engage in cocreating digital play experiences, possibly challenging ageist assumptions. In addition, Loos [28], in a literature review, identified social interaction, usability, and enjoyment as key motivators for engaging with exergames and digital play platforms.

Across both qualitative and quantitative studies, the social aspect of playfulness, especially the "other-directed" facet (enjoyment of others' company and using playfulness to create positive social interactions), is consistently highlighted. This includes using humor and lightheartedness to foster relationships and navigate social tensions [3,23]. Themes such as connectedness, continuity, and participation also emerged in qualitative narratives, highlighting play's role in enhancing older adults' sense of agency and well-being [9,20].

Three validated instruments were identified in the review as being used to assess playfulness in older adults:

- The OAP Scale, developed through mixed methods research and a classical test theory (CTT) approach [27], was designed specifically for older populations (exact age range not specified) and focuses on unique age-related expressions of playfulness. The OAP includes 15 items divided into 4 factors: psychologically upbeat, behaviorally impish, cognitively spontaneous, and amusing.
- The OLIW Playfulness Scale (used by [23]) assesses 4 key dimensions: other-directed, light-hearted, intellectual, and whimsical traits. The instrument was applied in a quantitative survey targeting older adults aged 60 years and older. The instrument was developed by Proyer [26] using a CTT approach. It focuses on attitudes and actions in relation to playfulness and contains 28 items divided into 4 traits (as described above).
- The SMAP [2] was validated in a cross-sectional study with participants aged 18 to 92 years (mean age 45), offering insight into the intensity, frequency, and ease of initiating playful behavior across the lifespan. The SMAP was developed by Proyer [25] using a CTT approach. It has 5 items: i1, I am a playful person; i2, Good friends would describe me as a playful person; i3, I frequently do playful things in my daily life; i4, It does not take much for me to change from a serious to a playful frame of mind; and i5, Sometimes, I completely forget about the time and am absorbed in playful activity.

Taken together, qualitative studies highlighting meaning, continuity, and connectedness [8,9,20,21] and quantitative research linking playfulness with well-being [2,3,23] suggest that playfulness in older age is complex, contextual, and relational. However, existing instruments primarily conceptualize playfulness as a personality trait and have been developed for general adult populations. They place less emphasis on how playfulness is enacted in everyday life within care contexts or on how environmental and functional constraints may shape this expression.

Valid and age-sensitive instruments, responsive to functional limitations and municipal care environments, are therefore

XSL•FO
RenderX

needed to capture these nuances in both research and practice. PLAY-SE will be grounded in older adults' lived experiences within Swedish municipal care and will integrate perspectives from caring science and person-centered care. It will be developed with explicit Rasch-based psychometric intentions to ensure appropriate targeting and invariant measurement in this specific population. For these reasons, adaptation of existing instruments was considered insufficient to meet the conceptual and methodological aims of this research program.

Theoretical Underpinnings

This program will be conducted from a person-centered perspective, emphasizing the ethical aim of a good life, lived with and for others, within fair and equitable institutions [29]. To explore playfulness in older adults living in long-term care facilities, we apply the 4 modes of being a person: being with self, being in relation, being in a social world, and being in place, as described by McCormack and McCance [7]. In addition, existential philosophy provides a lens through which human beings are not seen as predetermined. Although factual limitations shape our lives, they do not ultimately define who we are. Instead, we are always self-making beings, responsible for the meanings we give through our choices, the totality of which makes us who we are [30]. In caring science, this tension between determinism and freedom is understood as the interplay of vulnerability and freedom [31].

To achieve research depth that embraces lived experiences as meanings, a phenomenological reflective lifeworld research (RLR) approach will be adopted (in the first phase of the program), as outlined by Dahlberg et al [32]. In this way, the program will explore older adults' lived experiences of being playful while living in long-term care facilities. In RLR, the phenomenon is the central focus and is described through its essential meaning and constituents. The approach draws on the philosophical underpinnings of hermeneutics and phenomenology as described by Gadamer [33], Husserl [34], and Merleau-Ponty [10].

RLR provides methodological principles such as openness, flexibility, and bridling. Openness deals with the demand for sensitivity to the phenomenon being studied and entails having the capacity to be surprised and sensitive to the unpredictable and unexpected. Through this open attitude, the objects of the lifeworld may reveal themselves as different from how they were previously assumed to be [32]. To bridle means to openly return to the things themselves without making definitive what is indefinite, so as not to understand too quickly. In addition, within the bridling attitude, one moves beyond natural understanding and restrains one's preunderstanding, including personal beliefs, theories, and other assumptions, so that they do not mislead the understanding of meaning.

Conceptual Framework and Hierarchy

As part of the program, the PLAY-SE instrument will be developed at the end of phase 1 and refined and evaluated during phase 2 (Figure 1). The development of a psychometric instrument typically begins with qualitative conceptualization [35], where the construct of interest (eg, playfulness) is explored through interviews, observations, and/or literature reviews to ensure content is grounded in lived experience (phase 1). Based on these insights, developers generate domains and items that reflect different facets or levels of the construct, along with a response format suited to capture variation in intensity or frequency. To enable measurement on a hierarchical scale or "ruler," items should ideally vary in difficulty or endorsement level, allowing for ordering along a unidimensional latent trait. In this way, the qualitative and theoretical work contributes to defining an outcome space [36], that is, a structured representation of the range and organization of possible expressions of playfulness, ranging from lower to higher levels, from being playful to acting playfully. This is followed by pilot testing and psychometric evaluation using approaches such as Rasch model analysis, which assesses item fit, scale targeting, and unidimensionality, thereby supporting the creation of an interval-level measure. Final validation typically includes analyses of reliability, validity, and responsiveness across diverse samples [35,36].

Psychometrics is the scientific study of measurement instruments, such as rating scales [37]. Traditional approaches, often referred to as CTT, rely on correlation-based techniques such as Cronbach α, principal component analysis, and factor analysis [37,38]. Although widely used, CTT assumes normally distributed, continuous data, does not ensure invariant measurement across persons or contexts, and treats summed scores as interval-level data, which limits generalizability and validity [39]. Rasch measurement theory (RMT), by contrast, offers a probabilistic framework within modern test theory [40]. By modeling the probability of endorsing an item based on person ability and item difficulty, Rasch model analysis enables unidimensional, invariant, and interval-level measurement. Its grounding in measurement principles comparable to the physical sciences allows for alignment between individual- and group-level structures, leading some to regard the Rasch model as the ideal standard for scientific measurement [41-43]. This approach enables quality-assured comparisons both within individuals over time and between individuals [44]. Consequently, RMT provides a foundation for truly person-centered research, focusing on measurement with and for the individual person.

Conceptual Clarifications

In this program, a distinction is made between playful activities (sometimes referred to simply as "play") and playfulness, while recognizing that the terms are often used interchangeably in everyday conversation. Theoretically, however, they represent 2 related yet distinct constructs.

- Play refers to activities undertaken for the sheer enjoyment of the experience—activities that feel fun, creative, or relaxing. Play can be solitary or social, such as crafting, playing games, joking, dancing, painting, imagining, or trying something new. Importantly, play is not driven by external goals or outcomes; its value lies in the experience and the meaning it holds for the person.
- Playfulness, on the other hand, describes a personal orientation or way of being playful, characterized by the ability to experience joy, express curiosity, use imagination, and find amusement in everyday situations. As a personality

XSL•FO
RenderX

trait or state of mind, playfulness shapes how persons approach and interpret the world around them.

While playfulness can lead to playful activities, and playful activities may foster playfulness, they are not synonymous. In scientific research, it is valuable to distinguish between these concepts to avoid conceptual overlap and ensure analytical clarity. For instance, an older adult may retain a playful disposition (being playful) but be hindered from acting playfully due to situational and environmental circumstances, such as illness and institutional routines. In such circumstances, differentiating between inner orientation and outward expression becomes essential.

Moreover, "playful activities" may be understood as a broader construct than "play," since it includes not only traditional forms of play (eg, games, hobbies, creative pursuits) but also everyday situations infused with humor, curiosity, or imagination. Recognizing these distinctions between being playful and acting playfully and between play and playful activities allows for a more nuanced understanding of how playfulness contributes to well-being in later life. This perspective also reflects prior research highlighting the multidimensional and contextual character of playfulness among older adults [2,3,8,19].

Initial Conceptualization and Theoretical Hierarchy

As part of this review and study protocol, an initial conceptualization and theoretical hierarchy for the PLAY-SE instrument were developed. This process drew on insights from the literature, together with conceptual reasoning, providing a theoretically grounded basis for the forthcoming item generation and psychometric evaluation. The conceptualization and hierarchy will be further elaborated in forthcoming discussions within the research team, drawing on findings from phase 1 and serving as a springboard for phase 2 of the program. The PLAY-SE will be a self-administered instrument comprising approximately 25 statements across 6 preliminary conceptual domains: : "Across the Life Course," "Attitudes," "Environment and Enablers," "Emotional Meaning," "Health and Well-being," and "Playful Activity." While designed for independent completion, support from staff should be offered, if needed, to ensure that all participants can engage with the instrument on their own terms. Higher scores will reflect greater levels of playfulness or engagement in play among older adults receiving municipal care (Figure 2).

Figure 2. Hierarchical structure for the Play and Supportive Environments instrument: domains across increasing levels of playfulness.

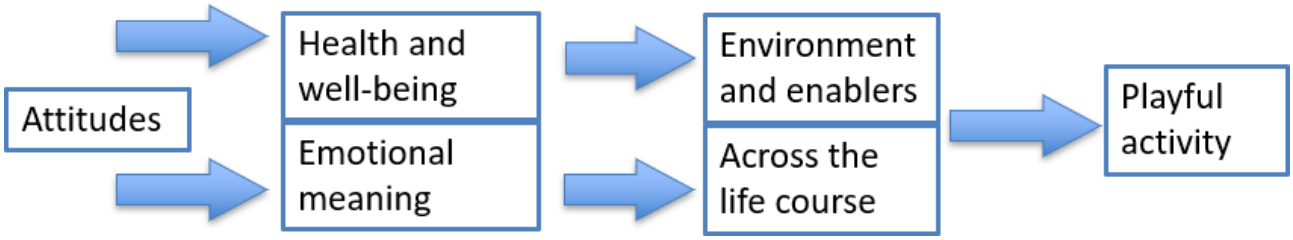

Each domain will reflect a distinct aspect of the phenomenon and contribute to a comprehensive understanding of how playfulness may be experienced and expressed in later life. These domains are not only conceptually defined but also hypothetically arranged along a continuum, ranging from abstract attitudes and internalized values to embodied, enacted expressions of playfulness through playful activities. This structure reflects the idea that playfulness can manifest at different levels of engagement, "from potential to practice" (as illustrated in Figure 2). Further to the left indicates that it is easiest to agree with a statement (a lower degree of playfulness is required). Further to the right indicates that it is hardest to agree with a statement (a higher degree of playfulness is required).

Compared to existing measures (OAP, OLIW, SMAP), the proposed domains extend beyond trait-level orientations by explicitly capturing both being playful and acting playfully, as well as environmental enablers and constraints. In this way, PLAY-SE aims to reflect the lived experience of older adults in municipal care more fully than previous instruments, thereby addressing important gaps in the assessment of playfulness in later life.

This structure reflects the idea that playfulness can manifest at different levels of engagement, "from potential to practice" (as further explained in Table 1).

XSL•FO
RenderX

Table . Domains, purpose, and arguments for hypothetical positions regarding playfulness and playful activity.

Domain	Purpose	Hypothetical position
1. Attitudes	Captures norms and self-perceptions related to play.	Considered the easiest domain to endorse, primarily involving the expression of general opinions (eg, "older adults can play") without requiring personal engagement.
2. Health and well-being	Captures the perceived impact of play on health and well-being.	Generally easy to agree with; many older adults recognize the positive effects of play, even if they do not actively engage in it themselves.
3. Emotional meaning	Assesses how important play is and what it feels like to lose access to it.	Requires a more personal and emotional connection to the role of play, though it can still be understood hypothetically.
4. Environment and enablers	Captures how the environment and the body influence the possibility of engaging in play.	Reflects a willingness to play, acknowledging that expression depends on external conditions such as physical ability, institutional rules, or support.
5. Across the life course	Measures self-image and the continuity of playfulness throughout the lifespan.	Requires a stronger personal anchoring, as identifying oneself as playful over time implies a deeper integration of playfulness into one's identity.
6. Playful activity	Captures behavioral expressions of play.	The most demanding domain reflects concrete engagement in playful actions and the creation of playful situations, indicating the highest level of expressed playfulness.

The 6 proposed domains for the PLAY-SE instrument are specifically designed to bridge older adults' lived experiences with complex theoretical perspectives (such as person-centered care, lifeworld theory, and existential philosophy) by translating abstract concepts into dimensions that are both measurable and meaningful to the individual in their context. The domains collectively ensure that the instrument does the following:

- Centers the person and their subjective meanings (person-centeredness). Mainly captured by the domains of *attitudes* (toward play) and *emotional meaning*, which reflect personal values, emotional relevance, and how individuals perceive the role of play in their own lives.
- Reflects the totality of human experience—body, time, space, emotion, and sociality (lifeworld theory). Mainly captured by *environment and enablers*, *across the life course*, and *health and well-being*, which together address the spatial, temporal, emotional, physical, and social or intersubjective aspects of being playful in everyday life.
- Accounts for existential challenges and choices, such as freedom, limitation, loss, and identity (existential philosophy). Mainly captured by *playful activity*, *emotional meaning*, and *across the life course*, which reflect the tension between agency and constraint, the significance of loss, and the continuity of identity through play.

Rationale and Aims

Overall Aim

The overall aim is to further clarify and operationalize the interrelated concepts of playfulness and playful activities among older adults receiving municipal care and to develop an instrument suitable for psychometric testing.

Phase 1—Conceptual Exploration

The overall aim of phase 1 is to deepen understanding of playfulness among older adults living in long-term care facilities. This includes exploring their lived experiences, engagement in playful activities, and the perspectives of care staff. A further objective is to examine how playfulness relates to well-being in this population.

Specific objectives are as follows:

1. To describe playfulness as experienced by older adults living in long-term care facilities
2. To explore older adults' experiences of engaging in playful activities and their relation to health and well-being
3. To investigate activity coordinators' and managers' experiences of older adults' possibilities and barriers to express their playfulness
4. To develop a first version of the PLAY-SE instrument, including initial item generation and pilot testing to explore its measurement properties and guide further refinement

Phase 2—Instrument Development and Psychometric Evaluation

The overall aim of phase 2 is to advance the development of the PLAY-SE instrument by exploring its theoretically developed domains against empirical data, refining its conceptual structure, and strengthening content validity [36]. This phase will also involve systematic psychometric evaluation, including assessments of usability, reliability, dimensionality, and validity, to ensure that the instrument provides a robust and nuanced measure of playfulness among older adults in the context of illness or institutional care.

Specific objectives are as follows:

XSL•FO
RenderX

1. To strengthen the theoretical and conceptual basis for the instrument
2. To assess content validity and usability through interviews
3. To conduct an initial psychometric evaluation
4. To validate a revised version using a larger, representative sample

This phase will also examine associations with the SMAP [25] and indicators of well-being among older adults receiving municipal care and service support.

Methods

Overall Design

Inspired by Huber et al [45] and the "hybrid model of concept development" [46], this program applies a hybrid model of concept development combined with an embedded mixed methods design. The hybrid model is particularly suited for refining and clarifying complex concepts by integrating theoretical insights with empirical observations. Concept development enhances conceptual clarity by identifying the essential attributes of a construct and translating them into measurable dimensions.

The hybrid model unfolds in 3 interrelated cycles (Figure 1):

1. Theoretical cycle: A literature review is conducted to examine how the concept has been defined and measured previously. In this protocol, this cycle has been initiated through a structured exploratory literature review using the Elicit platform and a synthesis of prior studies (see Literature Review). These findings inform conceptual clarifications of play and playfulness (see Conceptual Clarifications), tailored to the study's aims.
2. Fieldwork cycle: Empirical data are collected and analyzed to investigate how the concept is expressed in real-world contexts. In phase 1, phenomenological interviews with older adults in municipal care provide meaning units that expand, refine, or challenge the preliminary theoretical understanding.
3. Analytical cycle: Empirical findings are reintegrated with the theoretical groundwork to develop a refined conceptual description. This refined description directly informs item generation and domain structure for the PLAY-SE instrument in phase 2. Thus, qualitative findings are systematically translated into candidate items, which are subsequently evaluated and refined through cognitive interviews and psychometric testing.

The cycles are applied iteratively, meaning that quantitative findings from pilot testing and Rasch analysis may prompt further conceptual reflection to ensure alignment between empirical performance and theoretical meaning. Together, these cycles underpin the 2 phases of the PLAY-SE program: phase 1 (conceptual exploration) and phase 2 (instrument development and psychometric evaluation).

Setting and Participants

The program will be conducted in residential care facilities (ie, nursing homes) and in ordinary housing among persons receiving municipal care support in Sweden. Participants will include older adults aged 65 years and older, as well as activity coordinators and managers involved in municipal elder care.

Inclusion criteria (phase 1 and phase 2) were as follows:

- Older adults aged 65 years or older
- Receiving municipal care support, either in residential care facilities or in their own homes
- Able to provide informed consent
- Sufficient cognitive and communicative ability to participate in interviews, surveys, or focus groups

Exclusion criteria (phase 1 and phase 2) were as follows:

- Severe cognitive impairment or acute illness that prevents informed consent or meaningful participation
- Non-Swedish–speaking persons (due to the instrument being developed and tested in Swedish)

Preliminary sample size of phase 1 (conceptual exploration) was as follows:

- Approximately 15 to 25 older adults for qualitative interviews and 10 to 15 activity coordinators or managers for focus groups (guided by data saturation).

Preliminary sample size of phase 2 (instrument development and psychometric evaluation) was as follows:

- Content validity interviews with 10 to 15 older adults and 6 to 10 experts.
- A pilot survey including approximately 100 older adults to test the PLAY-SE instrument, version 1, and explore its initial measurement properties, as well as associations with SMAP [25] and well-being indicators.
- Large-scale validation study including 250 to 500 older adults to confirm psychometric properties (CTT and modern test theory).

Sampling Strategy

A purposive sampling strategy will be used in phase 1 to ensure variation in gender, age, and relevant care contexts within municipal elder care. The aim is not statistical representativeness but conceptual variation to inform domain development, with particular emphasis on older adults living in residential care facilities.

For phase 2 (psychometric testing), a stratified sampling approach will be applied across multiple participating municipalities, including both urban and rural contexts. Recruitment will focus primarily on older adults receiving municipal care support in ordinary housing while ensuring sufficient variation in demographic and contextual characteristics. Eligible individuals will be invited consecutively or through structured random selection, where feasible. This approach aims to enhance representativeness within the intended target population while acknowledging practical constraints within municipal services.

To reduce recruitment bias, initial contact will be facilitated by designated care staff, but final eligibility confirmation and informed consent will be handled directly by the research team. Recruitment procedures will be documented to allow transparent reporting of participation rates and sample characteristics.

XSL•FO
RenderX

Sample Size Considerations

The planned sample sizes are based on methodological considerations specific to qualitative research, CTT, and RMT.

For the qualitative interviews in phase 1 (approximately 15 - 25 older adults), the sample size is guided by the principle of conceptual saturation and information power, where the aim is depth and variation rather than statistical representativeness [32]. Recruitment will continue until thematic depth and conceptual clarity are achieved across variations in gender, care setting, and functional status.

For pilot testing (approximately 100 participants), the sample size is considered sufficient for preliminary evaluation under CTT, including the assessment of internal consistency, item-total correlations, floor and ceiling effects, and exploratory dimensionality [37,38,47]. Samples of around 100 participants are commonly recommended for early-stage scale refinement to identify poorly functioning items and evaluate reliability coefficients such as Cronbach α [38].

For large-scale validation (250 - 500 participants), the sample size is aligned with recommendations for RMT. Samples of approximately 250 participants allow reasonably stable estimation of item difficulty parameters, while samples approaching 500 participants improve the precision of item calibration and enable robust testing of differential item functioning (DIF) across subgroups [48-50]. Larger samples also strengthen the evaluation of unidimensionality and local independence and support more precise targeting of the instrument [51].

Together, these staged sample sizes are intended to ensure conceptual validity in phase 1 and rigorous psychometric evaluation in phase 2, in line with best practice recommendations for scale development.

Recruitment Procedures

Participants will be recruited from both residential care facilities and ordinary housing with municipal home care support. Recruitment procedures will be adapted to each setting while maintaining consistent inclusion criteria. In residential care, responsible nurses or care managers will identify potentially eligible participants based on documented eligibility criteria and provide initial study information. In home care settings, recruitment will occur through designated contact persons within municipal services who distribute study information to eligible individuals. In all settings, final eligibility and informed consent will be confirmed directly by the research team to ensure voluntariness and comprehension.

Assessment of Cognitive Ability

Eligibility requires sufficient cognitive and communicative capacity to understand study information and provide informed consent. Formal cognitive screening instruments will not be administered as part of the study. Instead, the assessment of cognitive capacity will be based on structured clinical judgment conducted in collaboration with responsible nurses or care staff, drawing on documented cognitive status in medical records and their ongoing knowledge of the individual's functioning.

The assessment will focus on the participant's ability to do the following:

1. Understand the purpose of the study
2. Comprehend what participation entails
3. Express a voluntary decision regarding participation

If uncertainty arises regarding decisional capacity, the individual will not be included. The basis for inclusion decisions will be documented to ensure procedural consistency.

Administration Procedures and Mode Effects

The PLAY-SE instrument is designed for self-administration. However, given the target population, some participants may require assistance (eg, due to visual impairment or motor limitations). To ensure procedural consistency, assistance will follow a standardized protocol. Support will be limited to reading items aloud and recording responses as expressed by the participant, without interpretation, rephrasing, or prompting. The mode of administration (independent completion vs assisted completion) will be documented for each participant. This variable will be included in subsequent psychometric analyses.

Potential mode effects will be examined analytically. In the Rasch analysis, DIF will be tested across administration modes to determine whether assisted completion systematically influences item responses. If significant DIF is identified, its impact on measurement invariance will be evaluated.

Item Generation and Content Development

Initial item development will be informed by the theoretical framework and the structured literature review described in this protocol, as well as by the research team's clinical and methodological expertise within gerontological care [52]. In addition, qualitative findings from phase 1 will contribute empirical grounding through a structured process in which meaning units are identified, clustered into conceptual domains, and translated into candidate items.

While older adults' lived experiences provide the empirical foundation for item content, item generation will not rely solely on participants' verbatim expressions. In line with the notion of epistemic dialogue in patient-centered measurement [52], experiential knowledge will be treated as a vital but not self-sufficient source of evidence. Participants' perspectives will inform item content and relevance, yet they will also be examined, conceptually refined, and integrated with theoretical frameworks (eg, person-centered care and lifeworld theory) and psychometric principles. In this way, the research team plays an active role in ensuring that items reflect both lived experience and a coherent construct suitable for rigorous measurement.

For each domain, an initial item pool (approximately 3 - 5 items per domain) will be generated to capture varying levels of endorsement along the hypothesized continuum from lower to higher expressions of playfulness. Items will be formulated in Swedish using clear, concrete, and contextually appropriate language.

Item wording and content will be reviewed in 3 stages:

1. Internal expert review focusing on conceptual coherence and redundancy

XSL•FO
RenderX

2. Cognitive interviews with older adults to assess clarity, relevance, and interpretability
3. Iterative refinement based on empirical feedback and measurement considerations

All decisions regarding item inclusion, modification, or exclusion will be documented to ensure transparency.

Documentation of Item Decisions

All stages of item development will be documented to ensure transparency and reproducibility. A structured decision log will be maintained throughout the development process, recording the origin of each item (eg, qualitative finding, theoretical refinement, or expert suggestion), revisions to wording, and the rationale for retention, modification, or exclusion.

During pilot testing and initial CTT analyses, the item pool will be retained in its entirety to preserve conceptual breadth. Early item removal will be avoided unless an item demonstrates clear redundancy or fundamental comprehension problems.

Subsequent Rasch analysis will provide a more rigorous evaluation of dimensionality, threshold functioning, item fit, and DIF. Decisions regarding item removal will be based on an integrated assessment of statistical performance and conceptual relevance. Items will not be removed solely on the basis of misfit statistics; instead, conceptual coverage, theoretical coherence, and construct validity will be weighed alongside empirical indicators.

This staged approach ensures that item reduction, if necessary, occurs only after a comprehensive psychometric evaluation and is supported by both measurement and conceptual considerations. The full decision process will be documented to provide a transparent audit trail linking qualitative findings, theoretical assumptions, and psychometric analyses to the final item selection.

Prespecified Psychometric Criteria

Prespecified psychometric criteria have been defined to guide the evaluation and scale refinement of PLAY-SE. These criteria, grounded in established methodological literature, specify statistical thresholds and decision rules under both CTT and RMT (Table 2).

XSL•FO
RenderX

Table . Prespecified criteria for psychometric evaluation and scale refinement of the Play and Supportive Environments instrument.

Measurement property	Method	Criterion	Decision rule	Reference
Content validity	I-CVI[a]	≥0.78	Item revised if <0.78	[53]
Content validity	S-CVI[b]	≥0.90	Overall revision if <0.90	[53]
Data quality	Missing data per item	<10%	Investigate content, wording, and administration if ≥10%	[47]
Floor/ceiling effects	% lowest/highest category	≤20%	Consider revision if >20%	[47]
Internal consistency (CTT)[c]	Cronbach α	≥0.70 (group level)	Review scale if <0.70	[47]
Item-total correlation (CTT)	Corrected item-total	≥0.30	Review if <0.30	[47]
Dimensionality (CTT)	EFA loadings	≥0.40 on intended factor	Review cross-loading items	[38]
Rasch model fit (RMT)[d]	Item fit residual	−2.5 to +2.5	Review outside range	[54]
Threshold ordering (RMT)	Category probability curves	Ordered thresholds	Consider collapsing categories if disordered	[51,54]
Local independence (RMT)	Residual correlations (Yen Q3)[e]	<0.20 above average	Consider subtests or revision	[54,55]
Reliability (RMT)	PSI[f]	≥0.70	Review if <0.70	[54]
DIF[g] (RMT)	2-way ANOVA on residuals	Nonsignificant after Bonferroni adjustment[h]	If significant, consider item split	[54,56]
Unidimensionality (RMT)	PCA[i] of residuals + *t* tests	≤5% significant	If >5%, explore substructure	[51,57,58]
Item removal principle	Integrated conceptual + statistical evaluation	Not solely data driven	Remove only after Rasch and conceptual review	[36,52]

[a]I-CVI: Item Content Validity Index.

[b]S-CVI: Scale Content Validity Index.

[c]CTT: classical test theory.

[d]RMT: Rasch measurement theory.

[e]Yen Q3: residual correlation index used to assess local dependence in Rasch analysis.

[f]PSI: Person Separation Index.

[g]DIF: differential item functioning.

[h]Bonferroni adjustment refers to correction for multiple testing by dividing the significance level by the number of comparisons.

[i]principal component analysis

Handling of Missing Data

Given the advanced age and varying functional capacity of participants, some degree of missing data is anticipated. The extent and pattern of missing responses will be examined descriptively. If item-level missingness exceeds 10%, this will be analyzed to determine whether it reflects systematic difficulties (eg, item wording, administration mode, or functional limitations). In CTT analyses, cases with excessive missing responses may be excluded from specific analyses. In RMT, the model accommodates some degree of missing data under the assumption of conditional independence; however, patterns of missingness will be evaluated to ensure that they do not reflect systematic bias.

Ethical Considerations

The PLAY-SE program is conducted in accordance with the ethical principles for research involving human subjects, as outlined by the Swedish Ethical Review Authority. Ethical considerations are made throughout the research program, in accordance with the Swedish Ethical Review Act [59] and the Declaration of Helsinki [60]. Ethical approval has been obtained for the qualitative work in phase 1 (registration number 2025-00211-01). A complementary application will be submitted before data collection for later milestones. All data will be handled in accordance with established ethical standards for privacy and confidentiality. No identifying information (eg, names, initials, or other personal identifiers) will be included in the manuscript or supplementary materials. No compensation will be offered to participants.

Participation is voluntary and based on informed consent. Only older adults capable of providing informed consent will be included, and participants may withdraw at any time without consequence. As participants are older adults receiving municipal care, they may constitute a potentially vulnerable group due to age-related functional limitations and dependency structures.

To minimize risks and safeguard voluntariness, several measures are implemented. Recruitment is facilitated by designated care

XSL•FO
RenderX

staff to identify eligible individuals; however, study information, eligibility confirmation, and informed consent are handled directly by the research team to emphasize independence from care provision. Participants are explicitly informed that participation or nonparticipation will not influence their care in any way.

Participants choose the time and location of interviews to ensure a familiar and secure setting. Focus groups are conducted separately for managers and staff to reduce power imbalances. During questionnaire administration, staff will not be present unless necessary for practical support. Any assistance provided follows a standardized protocol limited to reading items aloud and recording responses without interpretation. Given the risk of social desirability bias in care contexts, mode of administration (independent vs assisted completion) will be documented and examined analytically through DIF testing.

As discussions about playfulness may evoke emotional responses, participants are offered follow-up contact if needed. Confidentiality is emphasized throughout, and all research data are pseudonymized or anonymized as appropriate, securely stored, and managed in accordance with applicable data protection regulations and institutional policies.

A reference group of older adults has been involved from the outset and remains engaged throughout the research process. This participatory approach supports the ethical commitment to treating participants as situated knowers and strengthens the contextual sensitivity of instrument development.

Findings from the PLAY-SE program will be disseminated through publication in peer-reviewed journals and presentation at national and international conferences. Summaries will be prepared for practitioners and older adults and shared with participating care units and municipal stakeholders to facilitate knowledge translation into practice.

Results

Ethical approval for the qualitative component of phase 1 was granted by the Swedish Ethical Review Authority (2025-00211-01; decision date: February 3, 2025). Data collection for the first qualitative study was conducted between February and April 2025 and included 15 older adults aged 68 to 100 years residing in nursing homes. The phenomenological findings from phase 1.1 have been published in March 2026 in the *International Journal of Qualitative Studies on Health and Well-being* [61].

Two PhD students were recruited to the program in September 2024 and September 2025, respectively, and an expert group was established in autumn 2025. The PhD students are funded, for four years each, by Kristianstad University (from 2024) and Red Cross University College (from 2025). These structures support the ongoing conceptual and psychometric development of the PLAY-SE instrument.

A complementary ethical application will be submitted prior to subsequent milestones, including pilot testing and large-scale psychometric validation. Pilot testing of the first version of the PLAY-SE instrument (approximately 100 participants) is planned for autumn 2026. Large-scale validation (250 - 500 participants), including analyses based on CTT and RMT, is planned for 2027 to 2029. The overall project period extends to 2030.

Discussion

Principal Contribution and Implications

This study protocol outlines a structured and theory-informed approach to conceptualizing and measuring playfulness among older adults receiving municipal care. By integrating qualitative exploration with systematic psychometric evaluation, the program seeks to address the lack of a context-sensitive instrument grounded in the lived experiences of older adults receiving municipal care.

A central contribution of the program lies in its sequential design. The qualitative studies in phase 1 provide an empirical foundation for item generation, ensuring that the conceptualization of playfulness is informed by older adults' own narratives of being playful in everyday life. This integration of phenomenological inquiry and measurement development aligns with person-centered and lifeworld-oriented frameworks, in which individuals are understood as situated knowers whose experiences shape both meaning and practice. Grounding item development in qualitative findings is expected to strengthen content validity before proceeding to large-scale psychometric testing.

In contrast to existing measures of adult playfulness, which primarily assess trait-level dispositions, PLAY-SE is designed to capture both being playful and acting playfully within specific care contexts. The inclusion of environmental enablers and constraints reflects the theoretical assumption that playfulness in later life emerges in the interplay between personal orientation, bodily capacity, and contextual conditions. This contextual framing differentiates PLAY-SE from more general playfulness scales and may enhance its relevance for municipal care of older adults.

The protocol incorporates a staged psychometric evaluation plan. By applying both CTT and RMT, the study will examine dimensionality, reliability, targeting, and measurement invariance. Prespecified decision criteria for model fit, item retention, and DIF aim to enhance methodological transparency. Nevertheless, the ultimate adequacy of the instrument will depend on empirical performance in diverse samples.

Several methodological challenges must be acknowledged. Older adults receiving municipal care represent a heterogeneous population with varying cognitive and functional capacities. Ensuring accessibility while maintaining measurement precision requires standardized administration procedures and documentation of assistance during completion. Furthermore, as the instrument is initially developed within Swedish municipal contexts, its broader applicability will require subsequent cross-cultural validation.

If satisfactory measurement properties are established, PLAY-SE may support future research examining associations between playfulness, well-being, participation, and care quality. It may

XSL•FO
RenderX

also inform intervention studies exploring playful engagement as a potential resource for existential well-being in later life. However, such applications remain contingent upon empirical validation.

In summary, this protocol describes a theoretically grounded and methodologically transparent pathway toward measuring playfulness in later life. By combining qualitative depth with psychometric rigor, it aims to support the development of a context-sensitive instrument for person-centered municipal elder care.

Strengths and Limitations

The literature review was conducted using a structured approach supported by the research platform Elicit. Although this tool enables broad concept-based search across a large academic corpus, it does not replace systematic database searches conducted across multiple bibliographic databases. As the review was intended to provide a focused conceptual overview rather than a formal systematic review, it may not have captured all relevant studies. Future research could benefit from a more comprehensive systematic search strategy across databases such as PsycINFO, PubMed, and CINAHL to further strengthen the empirical foundation.

Acknowledgments

This study was conducted within the research environment STAGE (Societal Transformation, Ageing and Environments) and the Research Platform for Collaboration for Health, Faculty of Health Sciences, Kristianstad University, Sweden. Generative artificial intelligence (GPT-5.1, OpenAI) was used to assist with language refinement and clarity during manuscript preparation. The tool was not used to generate data, perform analyses, or interpret findings. All authors reviewed, revised, and take full responsibility for the final content of the manuscript. In addition, the research platform Elicit [18] was used to support the exploratory literature review conducted in May 2025. Elicit was employed to assist in identifying and screening potentially relevant peer-reviewed publications based on predefined inclusion criteria. Final study selection, interpretation, and synthesis were conducted exclusively by the authors.

Funding

The program is funded by Kristianstad University and Red Cross University College.

Authors' Contributions

AW serves as principal investigator and drafted the first version of this protocol, including the conceptual development of the Play and Supportive Environments instrument. AB and SKL will conduct the studies as part of their doctoral research. AB contributed, in particular, to the sections on the theoretical background as well as on ethics and dissemination. All authors contributed to the design of the overall research program and to the writing of this protocol.

Conflicts of Interest

None declared.

References

1. Proyer RT. The well-being of playful adults: adult playfulness, subjective well-being, physical well-being, and the pursuit of enjoyable activities. Eur J Humour Res 2013 Mar;1(1):84-98. [doi: 10.7592/EJHR2013.1.1.proyer]
2. Proyer RT. Playfulness over the lifespan and its relation to happiness: results from an online survey. Z Gerontol Geriatr 2014 Aug;47(6):508-512. [doi: 10.1007/s00391-013-0539-z] [Medline: 23982439]
3. Brauer K, Stumpf HSC, Proyer RT. Playfulness in middle- and older age: testing associations with life satisfaction, character strengths, and flourishing. Aging Ment Health 2024 Nov;28(11):1540-1549. [doi: 10.1080/13607863.2024.2372471] [Medline: 38940664]
4. Dillon P, Gan DRY, Trivic Z, et al. Playfulness for brain health in mid- and late-life: scoping review to support preventive non-pharmacological interventions in community settings. Arch Phys Med Rehabil 2022 Mar;103(3):e5. [doi: 10.1016/j.apmr.2022.01.012]
5. Golland Y, Ben-David BM, Mather M, Keisari S. Playful brains: a possible neurobiological pathway to cognitive health in aging. Front Hum Neurosci 2025;19:1490864. [doi: 10.3389/fnhum.2025.1490864] [Medline: 39989719]
6. Lubbers K, Cadwallader J, Lin Q, Clifford C, Frazier LD. Adult play and playfulness: a qualitative exploration of its meanings and importance. J Play Adulthood 2023;5(2):1-19. [doi: 10.5920/jpa.1258]
7. McCormack B, McCance T. Underpinning principles of person-centred practice. In: Person-Centred Practice in Nursing and Health Care: Theory and Practice: Wiley-Blackwell; 2017:13-34.
8. Burr B, Atkins L, Bertram AG, Sears K, McGinnis AN. “If you stop playing you get old”: investigating reflections of play in older adults. Educ Gerontol 2019 May 4;45(5):353-364. [doi: 10.1080/03601277.2019.1627058]
9. Hoppes S, Wilcox T, Graham G. Meanings of play for older adults. Phys Occup Ther Geriatr 2001 Jan;18(3):57-68. [doi: 10.1080/J148v18n03_04]

XSL•FO
RenderX

10. Merleau-Ponty M. Phenomenology of Perception, 2nd edition: Taylor and Francis; 2002.
11. Home is somewhere else: executive summary. : Swedish Institute for Human Rights; 2024 URL: https://mrinstitutet.se/download/18.12836cee194d834b690518e4/1740551595856/report-Home-is-somewhere-else_Executive-summary-2024.pdf [accessed 2026-03-24]
12. Neves BB, Sanders A, Kokanović R. "It's the worst bloody feeling in the world": experiences of loneliness and social isolation among older people living in care homes. J Aging Stud 2019 Jun;49:74-84. [doi: 10.1016/j.jaging.2019.100785] [Medline: 31229221]
13. Gebhard D, Frank JI. Everyday life and boredom of people living with dementia in residential long-term care: a merged methods study. BMC Geriatr 2024 Dec 31;24(1):1049. [doi: 10.1186/s12877-024-05641-7] [Medline: 39736528]
14. Socialtjänstlag (2025:400) [Swedish Social Services Act]. Swedish Code of Statutes. 2025. URL: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/socialtjanstlag-2025400_sfs-2025-400/ [accessed 2026-03-24]
15. Proyer RT, Gander F, Bertenshaw EJ, Brauer K. The positive relationships of playfulness with indicators of health, activity, and physical fitness. Front Psychol 2018;9:1440. [doi: 10.3389/fpsyg.2018.01440] [Medline: 30154749]
16. Van Vleet M, Feeney BC. Play behavior and playfulness in adulthood. Soc Pers Psychol Compass 2015 Nov;9(11):630-643. [doi: 10.1111/spc3.12205]
17. Waldman Levi A, Bar-Haim Erez A, Katz N. Healthy aging is reflected in well-being, participation, playfulness, and cognitive-emotional functioning. Healthy Aging Res 2015;4(8). [doi: 10.12715/har.2015.4.8]
18. Elicit. URL: https://elicit.com [accessed 2026-03-26]
19. Shen X. Constructing an interactionist framework for playfulness research: adding psychological situations and playful states. J Leis Res 2020 Oct 19;51(5):536-558. [doi: 10.1080/00222216.2020.1748551]
20. Schutter B, Vanden Abeele V. Meaningful play in elderly life. Presented at: 58th Annual Conference of the International Communication Association "Communicating for Social Impact"; May 22-26, 2008 URL: https://www.researchgate.net/publication/242691512_Meaningful_Play_in_Elderly_Life [accessed 2026-03-24]
21. De Schutter B, Vanden Abeele V. Designing meaningful play within the psycho-social context of older adults. Presented at: Fun and Games '10: Proceedings of the 3rd International Conference on Fun and Games; Sep 15-17, 2010; Leuven Belgium p. 84-93. [doi: 10.1145/1823818.1823827]
22. Sayago S, Rosales A, Righi V, Ferreira SM, Coleman GW, Blat J. On the conceptualization, design, and evaluation of appealing, meaningful, and playable digital games for older people. Games Cult 2016 Jan;11(1-2):53-80. [doi: 10.1177/1555412015597108]
23. Parker C, Kennedy-Behr A, Wright S, Brown T. Does the self-reported playfulness of older adults influence their wellbeing? An exploratory study. Scand J Occup Ther 2023 Jan;30(1):86-97. [doi: 10.1080/11038128.2022.2145993] [Medline: 36409561]
24. Clifford C, Paulk E, Lin Q, Cadwallader J, Lubbers K, Frazier LD. Relationships among adult playfulness, stress, and coping during the COVID-19 pandemic. Curr Psychol 2022 Feb 21;43(9):8403-8412. [doi: 10.1007/s12144-022-02870-0] [Medline: 35221637]
25. Proyer RT. Development and initial assessment of a short measure for adult playfulness: the SMAP. Pers Individ Dif 2012 Dec;53(8):989-994. [doi: 10.1016/j.paid.2012.07.018]
26. Proyer RT. A new structural model for the study of adult playfulness: assessment and exploration of an understudied individual differences variable. Pers Individ Dif 2017 Apr;108:113-122. [doi: 10.1016/j.paid.2016.12.011]
27. Yarnal C, Qian X. Older-adult playfulness: an innovative construct and measurement for healthy aging research. Am J Play 2011;4(1):52-79 [FREE Full text]
28. Loos E. Exergaming: meaningful play for older adults. In: Zhou J, Salvendy G, editors. Presented at: Human Aspects of IT for the Aged Population Applications, Services and Contexts; Jul 9-14, 2017. [doi: 10.1007/978-3-319-58536-9_21]
29. Ricœur P. Oneself as Another: University of Chicago Press; 1992.
30. Aho K. Existentialism: An Introduction: Polity Press; 2020.
31. Dahlberg K, Segesten K. Hälsa Och Vårdande: I Teori Och Praxis [Book in Swedish]: Natur & Kultur Läromedel och Akademi; 2010:344 URL: https://www.bokus.com/bok/9789127122116/halsa-och-vardande-i-teori-och-praxis/?srsltid=AfmBOopg5jMyQaTVTsAJNSG6jX4f2OSbbAITIuy2dn17x59hHhWn8bem [accessed 2026-03-24]
32. Dahlberg K, Dahlberg H, Nyström M. Reflective Lifeworld Research: Studentlitteratur; 2008. URL: https://www.researchgate.net/publication/265657817_Reflective_Life-World_Research [accessed 2026-03-24]
33. Gadamer HG. Truth and Method: A&C Black; 2013. URL: https://web.mit.edu/kaclark/www/gadamer_truth_and_method.pdf [accessed 2026-03-24]
34. Husserl E. The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy: Northwestern University Press; 1970. URL: https://archive.org/details/crisisofeuropean0000huss [accessed 2026-03-24]
35. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health 2018;6:149. [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]

XSL•FO
RenderX

36. Wilson M. Constructing Measures: An Item Response Modeling Approach, 2nd edition: Routledge; 2023. [doi: 10.4324/9781003286929]
37. Nunnally JC, Bernstein IH. Psychometric Theory: McGraw-Hill Companies; 1994.
38. Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use, 4th edition: OUP Oxford; 2008.
39. Stenner AJ, Stone MH, Fisher WP Jr. The unreasonable effectiveness of theory based instrument calibration in the natural sciences: what can the social sciences learn? J Phys: Conf Ser 2018 Jun 1;1044(1):012070. [doi: 10.1088/1742-6596/1044/1/012070]
40. Petrillo J, Cano SJ, McLeod LD, Coon CD. Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. Value Health 2015 Jan;18(1):25-34. [doi: 10.1016/j.jval.2014.10.005] [Medline: 25595231]
41. Andrich D. Rasch Models for Measurement: Sage Publications; 1988. [doi: 10.4135/9781412985598]
42. Mari L, Wilson M. An introduction to the Rasch measurement approach for metrologists. Measurement (Lond) 2014 May;51:315-327. [doi: 10.1016/j.measurement.2014.02.014]
43. Pendrill L. Man as a measurement instrument. NCSLI Measure 2014 Dec;9(4):24-35. [doi: 10.1080/19315775.2014.11721702]
44. Westergren A, Ahlström G, Persson M, Behm L. Next of kin participation in the care of older persons in nursing homes: a pre-post non-randomised educational evaluation, using within-group and individual person-level comparisons. PLoS One 2021;16(1):e0244600. [doi: 10.1371/journal.pone.0244600] [Medline: 33493153]
45. Huber E, Kleinknecht-Dolf M, Müller M, Kugler C, Spirig R. Mixed‐method research protocol: defining and operationalizing patient‐related complexity of nursing care in acute care hospitals. J Adv Nurs 2017 Jun;73(6):1491-1501. [doi: 10.1111/jan.13218] [Medline: 27878843]
46. Schwartz-Barcott D, Kim HS. An expansion and elaboration of the hybrid model of concept development. In: Rodgers BL, Knafl KA, editors. Concept Development in Nursing: Foundations, Techniques, and Applications, 2nd edition: Saunders; 2000. URL: https://www.researchgate.net/profile/Donna-Schwartz-Barcott/publication/285868499_An_expansion_and_elaboration_of_the_hybrid_model_of_concept_development/links/57fd770008ae49db47553bff/An-expansion-and-elaboration-of-the-hybrid-model-of-concept-development.pdf [accessed 2026-04-01]
47. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Improving the evaluation of therapeutic interventions in multiple sclerosis: development of a patient-based measure of outcome. Health Technol Assess 2004 Mar;8(9):1-48. [doi: 10.3310/hta8090] [Medline: 14982653]
48. Hagell P, Westergren A. Sample size and statistical conclusions from tests of fit to the Rasch model according to the Rasch unidimensional measurement model (RUMM) program in health outcome measurement. J Appl Meas 2016;17(4):416-431. [Medline: 28009589]
49. Tesio L, Caronni A, Simone A, Kumbhare D, Scarano S. Interpreting results from Rasch analysis 2. Advanced model applications and the data-model fit assessment. Disabil Rehabil 2024 Feb;46(3):604-617. [doi: 10.1080/09638288.2023.2169772] [Medline: 36744832]
50. Tesio L, Caronni A, Kumbhare D, Scarano S. Interpreting results from Rasch analysis 1. The “most likely” measures coming from the model. Disabil Rehabil 2024 Feb;46(3):591-603. [doi: 10.1080/09638288.2023.2169771] [Medline: 36740739]
51. Tennant A, Küçükdeveci AA. Application of the Rasch measurement model in rehabilitation research and practice: early developments, current practice, and future challenges. Front Rehabil Sci 2023;4:1208670. [doi: 10.3389/fresc.2023.1208670] [Medline: 37529206]
52. McClimans LM. Patient-Centered Measurement: Ethics, Epistemology, and Dialogue in Contemporary Medicine: Oxford University Press; 2024. [doi: 10.1093/oso/9780197572078.001.0001]
53. Polit DF, Beck CT. Nursing Research: Generating and Assessing Evidence for Nursing Practice, 11th edition: Wolters Kluwer; 2021.
54. Andrich D, Marais I. A Course in Rasch Measurement Theory: Measuring in the Educational, Social and Health Sciences: Springer; 2019.
55. Christensen KB, Makransky G, Horton M. Critical values for Yen’s Q3: identification of local dependence in the Rasch model using residual correlations. Appl Psychol Meas 2017 May;41(3):178-194. [doi: 10.1177/0146621616677520] [Medline: 29881087]
56. Hagquist C, Andrich D. Recent advances in analysis of differential item functioning in health research using the Rasch model. Health Qual Life Outcomes 2017 Sep 19;15(1):181. [doi: 10.1186/s12955-017-0755-0] [Medline: 28927468]
57. Tennant A, Conaghan PG. The rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum 2007 Dec 15;57(8):1358-1362. [doi: 10.1002/art.23108] [Medline: 18050173]
58. Hagell P. Testing rating scale unidimensionality using the principal component analysis (PCA)/t-test protocol with the Rasch model: the primacy of theory over statistics. OJS 2014;04(6):456-465. [doi: 10.4236/ojs.2014.46044]
59. Lagrummet. 2004. URL: https://www.lagrummet.se [accessed 2026-03-24]

XSL•FO
RenderX

60. WMA Declaration of Helsinki - ethical principles for medical research involving human participants. World Medical Association. 2024. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects [accessed 2026-03-24]
61. Bergman A, Haak M, Örmon K, Nivestam A, Westergren A, Nilsson Lindström P. Older persons' lived experiences of being playful in nursing home settings - a phenomenological reflective lifeworld research study. Int J Qual Stud Health Well-being 2026 Dec 31;21(1):2645255. [doi: 10.1080/17482631.2026.2645255] [Medline: 41830905]

Abbreviations

CTT: classical test theory
DIF: differential item functioning
OAP: Older Adult Playfulness
OLIW: Other-directed, Light-hearted, Intellectual, and Whimsical (Playfulness Scale)
PLAY-SE: Play and Supportive Environments
RLR: reflective lifeworld research
RMT: Rasch measurement theory
SMAP: Short Measure of Adult Playfulness

Edited by J Sarvestan; submitted 23.Oct.2025; peer-reviewed by A Hartanto, V Raj; revised version received 11.Mar.2026; accepted 16.Mar.2026; published 13.Apr.2026.

Please cite as:
Westergren A, Bergman A, Knaust Lidén S, Nahlén Bose C, Nivestam A, Norell A, Nilsson Lindström P, Haak M
Play and Supportive Environments Among Older Adults in the Context of Swedish Municipal Care: Protocol for the Development and Psychometric Evaluation of an Instrument to Measure Playfulness
JMIR Res Protoc 2026;15:e86373
URL: https://www.researchprotocols.org/2026/1/e86373
doi:10.2196/86373

© Albert Westergren, Anna Bergman, Sofi Knaust Lidén, Catarina Nahlén Bose, Anna Nivestam, Annika Norell, Petra Nilsson Lindström, Maria Haak. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Quality of Life of Patients With Multi-Organ Autoimmune Disease and Its Relationship to Patient-Centered Care: Protocol for a Longitudinal Cohort Study

Marianne Bayrhuber[1*], Dr; Christian Schlett[1*], Dr; Julia Slesaczeck[2], MSc; Sabrina Herbst[2], MA; Gabriele Mueller[2], Diplom-Ingenieur; Alexandra Nieters[3], Prof Dr; Bodo Grimbacher[3,4,5,6,7], Prof Dr; Erik Farin-Glattacker[1], Prof Dr

[1]Faculty of Medicine and Medical Center, Section of Health Care Research and Rehabilitation Research, Institute of Medical Biometry and Statistics, Albert-Ludwigs-University of Freiburg, Hugstetter Str. 49, Freiburg, Germany

[2]Center for Evidence-Based Healthcare, Faculty of Medicine and University Hospital Carl Gustav Carus, TUD Dresden University of Technology, Dresden, Germany

[3]Institute for Immunodeficiency, Center for Chronic Immunodeficiency (CCI), Medical Center, Faculty of Medicine, Albert-Ludwigs-University of Freiburg, Freiburg, Germany

[4]Clinic of Rheumatology and Clinical Immunology, Center for Chronic Immunodeficiency (CCI), Medical Center, Faculty of Medicine, Albert-Ludwigs-University of Freiburg, Freiburg, Germany

[5]DZIF – German Center for Infection Research, Satellite Center Freiburg, Freiburg, Germany

[6]RESIST – Cluster of Excellence 2155 to Hanover Medical School, Satellite Center Freiburg, Freiburg, Germany

[7]CIBSS – Centre for Integrative Biological Signalling Studies, Albert-Ludwigs University Freiburg, Freiburg, Germany

[*]these authors contributed equally

Corresponding Author:
Christian Schlett, Dr
Faculty of Medicine and Medical Center, Section of Health Care Research and Rehabilitation Research, Institute of Medical Biometry and Statistics, Albert-Ludwigs-University of Freiburg, Hugstetter Str. 49, Freiburg, Germany

Abstract

Background: Multi-organ autoimmunity belongs to a group of ultrarare diseases characterized by complex autoimmune or autoinflammatory processes affecting multiple organs. In addition to adequate medical care for patients with multi-organ autoimmunity, the understanding of patients' treatment preferences and the measurement of their health-related quality of life are essential for improving therapy for these patients.

Objective: This study has three goals: (1) to compile a modular questionnaire package using established instruments to assess health-related quality of life and health status; (2) to develop, through a participatory process, and conduct psychometric testing of a new self-report questionnaire assessing patients' treatment preferences; and (3) to evaluate quality of life and patients' preferences in a longitudinal study.

Methods: To address our goals, we are conducting a mixed methods study involving 300 adult patients undergoing treatment for multi-organ autoimmune diseases; the study includes qualitative interviews with patients and their physicians and a quantitative longitudinal study with 4 measurement time points (at study enrollment and at 3, 6, and 9 months later). We will address our first goal by conducting psychometric analyses to assess the applicability of the modularized questionnaire package to our target population. We will determine item characteristics and conduct exploratory factor analyses and internal consistency analyses with the scales used to assess generic and disease-specific quality of life, as well as disease-specific health status. To address our second goal, we will conduct (1) qualitative interviews with patients and their physicians to explore the preferences of our target population, (2) a 2-round Delphi study to select preferences that are highly relevant to most patients, (3) cognitive pretests to ensure the acceptability and comprehensibility of the scale items and instructions, and (4) psychometric analyses to guide item selection and evaluate the reliability and validity of the final scales.

Results: This paper describes the protocol for the longitudinal cohort study of the research project Qualy-GAIN (German Multi-Organ Auto-Immunity Network), which was funded from January 2023 to December 2025. Data were collected at 4 measurement points from July 2024 to June 2025. First results are expected to be submitted for publication in spring 2026.

Conclusions: A result of this study will be a package of questionnaires that can be used to assess health-related quality of life, disease-specific health status, and treatment preferences. Due to the voluntary nature of participation and the low burden associated with participation, we believe that the benefits of the study outweigh potential risks. In case the questionnaires evaluated in this study prove to be reliable, valid, and useful in practice, they can be transferred to the routine treatment of patients with multi-organ autoimmune diseases to improve treatment and promote patient-centered care.

XSL•FO
RenderX

Trial Registration: German Clinical Trials Register DRKS00033609; https://drks.de/search/de/trial/DRKS00033609/details
International Registered Report Identifier (IRRID): DERR1-10.2196/77786

(JMIR Res Protoc 2026;15:e77786) doi:10.2196/77786

KEYWORDS

health-related quality of life; multi-organ autoimmunity; patient preferences; rare disease; health status

Introduction

Patients with multi-organ autoimmunity belong to a group of patients with ultrarare diseases and are characterized by complex autoimmune or autoinflammatory processes affecting several organs. Multi-organ autoimmunity is often caused by rare genetic mutations that are categorized as congenital defects of immunity and can lead to various chronic disease manifestations [1]. However, multi-organ autoimmune diseases do not typically result in immediate mortality. Instead, affected individuals experience prolonged disease burden, extensive use of health care resources, and frequent consultations with multiple specialists ("physician hopping") while seeking optimal treatment.

Patient-reported outcome measures (PROMs) in health care can inform the treatment team about the patients' quality of life (QoL) at an individual level and can be used to monitor the treatment process and improve communication and shared decision-making [2]. At the system level, this information enables statements about the effectiveness and feasibility of care services, thus creating the prerequisite for improved adaptation processes to patients' needs. The QoL of patients with chronic diseases is not only influenced by pathophysiology and physical symptoms but also partly by the organization of treatment and the physician-patient relationship and interaction [3,4]. The dynamics of physician-patient communication, especially the decision-making process related to therapeutic choices, greatly influence patient adherence and satisfaction [5,6]. Patients with inflammatory rheumatic diseases in general prefer to have more information from their health care providers. However, lower-educated, older patients with low health literacy and high confidence in their health care providers prefer to leave treatment decisions to their physicians [3]. In addition to adequate medical care for patients with multi-organ autoimmunity, understanding patient preferences during interactions with physicians and the measurement of patients' QoL are essential for improving therapy [2,3]. Medical care for patients with autoimmune diseases can be improved when patient preferences and QoL are taken into account. By considering individual preferences, such as treatment goals and communication styles, health care providers can create more personalized and effective treatment plans [2]. This patient-centered approach can improve adherence to treatment and enhance overall patient satisfaction and long-term outcomes. However, most research focuses on clinical outcomes and health-related QoL, whereas patient preferences and satisfaction and their fulfillment are less studied [7].

The collection of patient-reported experience measures, that is, the fulfillment of patient preferences, also makes it possible to examine process indicators of medical care. Currently, no published research has examined the QoL of patients with multi-organ autoimmunity and its relationship to patient-centered care. In this understudied but important area, we will pursue 3 goals.

First, we will compile a modular questionnaire package of established instruments to assess QoL and disease-specific health status (eg, perceived symptoms, fatigue, and disease-related restrictions in activity and participation). The questionnaire package must meet the following requirements: (1) it should reflect the heterogeneity of different autoimmune diseases and their consequences for patients' health status in a meaningful way (a modular structure is planned for this purpose), (2) it should have a high methodological quality in our target group (reliability and validity), (3) it should be accepted by the patients, and (4) it should be informative for treatment practice.

Second, we will undertake the participatory development and psychometric testing of a new self-report questionnaire to assess (1) patients' preferences concerning patient-physician communication and interaction (eg, shared decision-making; microlevel), (2) preferences regarding the mode of cooperation between physicians (including diagnostic procedures and cocare; mesolevel), (3) preferences regarding the support system (including patient self-help; macrolevel), and (4) perceived fulfillment of these preferences. This instrument, which is to be newly developed according to psychometric criteria, is intended to capture a central aspect of patient-centered treatment, namely, the consideration of patient preferences in everyday care.

Third, we will use newly developed questionnaires in a patient sample of approximately 300 individuals in a prospective, longitudinal, epidemiological study with 4 measurement time points. In this study, we will test the following three hypotheses:

1. Fulfillment of preferences at an earlier time point is associated with better self-reported health status and improved health-related QoL at a subsequent time point.
2. There are plausible correlations between patients' self-reported health status and QoL, and pathophysiology and physical symptoms. Subjective health status can only partly be predicted by objective parameters, resulting in the need to consider it as an independent end point of treatment.
3. The feedback of the collected patient data on self-reported health status, QoL, and fulfillment of preferences to their physicians is considered useful by the treatment team to optimize treatment and, consequently, leads to more patient-centered care.

The identification of patient preferences and their relationship with patients' subjective health status and QoL offers a new

XSL•FO
RenderX

perspective on personalized care approaches for patients with multi-organ autoimmune diseases. The newly developed questionnaire can be used in routine care for patients with multi-organ autoimmune diseases to optimize their treatment. Furthermore, the procedure proposed here and the modular questionnaire set may be transferred to studies involving patients with other rare diseases.

Methods

Study Design

To address our goals, we will first conduct preliminary studies that include qualitative interviews with patients and physicians to explore patient treatment preferences, a 2-round Delphi study to select items that are highly relevant to most patients, and cognitive pretests to ensure the acceptability and comprehensibility of the questionnaire. Second, we will conduct an epidemiologic health services research cohort study with prospective longitudinal data collection at 4 measurement time points (at study enrollment and at 3, 6, and 9 months later). The data will be collected via online questionnaires. Patients' treatment preferences, QoL, and self-reported health status will be compiled into an individual-level report containing graphics for quickly assessing key values. This report will be sent to the patient's treating physician after the second measurement time point. Information about individual preferences and their fulfillment can help physicians to tailor treatment as closely as possible to patients' preferences. Information about the patient's health status and QoL can help physicians complement the impressions formed during consultations, providing them with a more comprehensive picture of the patients' well-being. To examine the usefulness of this feedback for treating physicians (third objective, hypothesis 3), we will conduct a qualitative interview study with them. This protocol is based on the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) reporting guidelines for study protocols [8] (Checklist 1). The quantitative longitudinal study will be embedded in the registry of the German Multi-Organ Auto-Immunity Network (GAIN)—a registry for individuals with congenital multi-organ autoimmune diseases, which has been documenting disease-specific patient data since 2020 [9]. The registry was designed to conduct embedded research studies whose output generates suggestions for improving patient care.

Sample and Setting

Our target population consists of adult patients with multi-organ autoimmune diseases enrolled in the GAIN registry. GAIN patients are defined as having at least 2 different autoimmune or autoinflammatory diseases or carrying a possibly pathogenic mutation in a gene that can cause a multi-organ autoimmune disease, including asymptomatic carriers. The GAIN registry includes approximately 440 adult patients documented by 8 university hospitals in Germany, 1 in Italy, and 1 in Portugal (status as of May 2024). Most of the adult patients have been documented at the centers in Freiburg, Hannover, and Berlin (375/440, 85.2%). This study includes patients enrolled in the GAIN registry that are treated at these 3 centers.

Inclusion and Exclusion Criteria

Inclusion criteria for participation in our study are that individuals are registered in the GAIN registry and have a multi-organ autoimmune disease as defined earlier. Furthermore, only individuals who have provided informed consent will be included in the study. Participants must be aged at least 18 years and have sufficient knowledge of the German language to answer the questionnaires. As the study asks about treatment preferences with respect to the documenting centers, participants must have been treated at the documenting centers at least once within the past 3 years.

Sample Size

We aim to include as many patients for the longitudinal study as possible. Therefore, the 3 participating medical centers invite all eligible patients enrolled in the GAIN registry to participate. As part of their enrollment in the GAIN registry, the patients have provided consent to be invited to participate in future studies. Moreover, they regularly visit their medical centers (that invited them to participate in our study) for treatment. Given these circumstances, which are likely to increase motivation to participate in our study, and based on the high participation rate of a comparable patient population in a previous study [10], we expect a participation rate of 80% and a dropout rate of 15%. Consequently, we expect a sample size of 240 at the first measurement point (t1) and 200 at the last (t4). Regarding the aim of our study, the largest sample size is required for the psychometric analyses of our first and second goals. Because these exploratory analyses depend on many factors that are not known in advance, the recommendations in the literature should be considered guidelines only. Although large samples (ie, n>250) are recommended for these analyses [11], samples in the range of 100 to 200 should also be sufficient [12], as we expect a small number of resulting factors and moderate item commonalities (h^2>0.50). Because the psychometric analyses will be conducted with the data collected at the first and second measurement time point, the expected sample size of ≥200 should provide a sufficiently large sample. This sample size should also be sufficiently large for testing hypothesis 1 of our third goal. A sample size of 200 allows for the detection of small-to-medium effects (Cohen f^2=0.02 to 0.15) with a power of 0.80 and an α of .05. Specifically, if considering the multilevel structure is not necessary, effects of f^2≥0.04 will be detectable with this sample size. However, if the examination of the intracluster correlation (ICC) indicates that the multilevel structure must be considered, then effects of f^2≥0.07 (ICC=1% [13]) or f^2≥0.12 (ICC=3% [14]) will be detectable.

Recruitment

Before recruitment began, the medical centers at the universities of Freiburg, Hanover, and Berlin assessed which of their patients enrolled in the GAIN registry met the study's inclusion criteria. Recruitment was then carried out by mail. To this end, the 3 centers sent letters from their leading physicians inviting their eligible patients to participate in the Qualy-GAIN study. These letters were accompanied by detailed study information; an informed consent form, which requested the patient's email

XSL•FO
RenderX

address because the longitudinal study is conducted using online questionnaires; and a prepaid envelope for the patients' replies. For patients who have provided informed consent, participation in the study begins with receipt of an email containing a link to the first online questionnaire. In total, the questionnaire study includes 4 surveys conducted over a 9-month period. At baseline (t1), 3 months (t2), 6 months (t3), and 9 months (t4), participants will receive a link to an online questionnaire that will take approximately 30 minutes to complete.

Measures

The focus of our study is on questionnaire development and exploratory research on patient preferences, QoL, and health status of patients with multi-organ autoimmune diseases. Relevant outcome measures for testing the hypotheses are disease-specific health status (self-reported), health-related QoL (generic and disease specific), as well as treatment preferences and their fulfillment.

Established questionnaires are used to measure subjective health status and health-related QoL with instruments from a pool of questionnaires used depending on the patient's symptoms. The questionnaire package is adaptable and modular, so that, for example, a patient with skin and bowel symptoms will receive the specific questionnaires for assessing health-related QoL of patients with dermatological and bowel diseases. We use questionnaires that have already been used in other studies on our target group or similar rare disease patient groups: Short Form 12 Health Survey [15], EQ-5D [16], Functional Assessment of Chronic Illness Therapy–Fatigue [17], Chronic Obstructive Pulmonary Disease Assessment Test [18], Short Inflammatory Bowel Disease Questionnaire [19,20], Dermatology Life Quality Index [21,22], and Extra Short Musculoskeletal Function Assessment Questionnaire [23].

As described earlier, the instruments used to assess treatment preferences and their fulfillment will be newly developed. Close patient involvement in the preliminary studies (eg, individual interviews, Delphi study, and cognitive pretests) will ensure relevance and applicability for the target population. Table 1 shows the measurement of the outcome variables at the 4 measurement points of the longitudinal Qualy-GAIN study.

XSL•FO
RenderX

Table . Outcome variables of the longitudinal Qualy-GAIN study.

Outcome	Instrument	Baseline (t1)	3 months (t2)	6 months (t3)	9 months (t4)
Subjective health status					
Diagnosis and affected organs[a]	Self-developed	✓	✓	✓	✓
Strength of impairment caused by the affected organs	Self-developed	✓	✓	✓	✓
Overall health status	EQ-5D (visual analog scale; 1 Item) [16]	✓	✓	✓	✓
Impairment due to autoimmune disease and side effects of medication	Self-developed	✓	✓	✓	✓
Number of days of sick leave in the past year	Self-developed	✓	✓	✓	✓
Health-related QoL[b] (generic)					
Physical and mental health–related QoL	Short Form 12 Health Survey [15]	✓	✓	✓	✓
Fatigue	Functional Assessment of Chronic Illness Therapy–Fatigue [17]	✓	✓	✓	✓
Health-related QoL (disease specific)[c]					
Affected organ: lung—QoL	COPD[d] Assessment Test [18]	✓	✓	✓	✓
Affected organ: bowel—QoL	Short Inflammatory Bowel Disease Questionnaire [19,20]	✓	✓	✓	✓
Affected organ: skin—QoL	Dermatology Life Quality Index [21]	✓	✓	✓	✓
Affected organ: musculoskeletal system—QoL	German Short Musculoskeletal Function Assessment Questionnaire [22,23]	✓	✓	✓	✓
Treatment preferences and their fulfillment					
Treatment preferences	Self-developed	✓[e]	✓[f]	✓	✓
Perception of current treatment	Self-developed		✓[f]	✓	✓
Fulfillment of preferences[a]	Self-developed		✓[f]	✓	✓

[a]Fulfillment of preferences is inferred from the difference between the treatment preferences and current treatment.

[b]QoL: quality of life.

[c]Disease-specific quality of life measures were assessed only if the corresponding organ is affected.

[d]COPD: chronic obstructive pulmonary disease.

[e]All k=42 items were presented and psychometrically analyzed to form the instrument used at time point t2.

[f]Items were psychometrically analyzed to form the final instruments to measure treatment preferences and the perceived current treatment at the time points t3 and t4.

Statistical and Qualitative Methods

Pseudonymized questionnaire data collected every 3 months will be used for data analysis. Psychometric analyses for our first goal include the determination of item characteristics, exploratory and confirmatory factor analyses, as well as internal consistency analyses. Our third goal comprises 3 hypotheses. To test hypothesis 1, linear mixed models will be used, in which

XSL•FO
RenderX

self-reported health status and health-related QoL at the later time point (t3) are each predicted by the fulfillment of treatment preferences recorded at the earlier time point (t2). To minimize confounding, control variables (eg, age and sex) are included in the models. In addition, we will examine the ICC to decide whether to include the center of treatment as a random effect (random intercept) in the model. To test hypothesis 2, correlation analyses will be performed to assess the relationship of QoL and the self-reported health status with objective medical parameters from the GAIN registry. Hypothesis 3 will be evaluated on the basis of the qualitative interviews with treating physicians. As with the qualitative interviews with patients and physicians in the preliminary studies, the qualitative interviews used to evaluate hypothesis 3 will be analyzed using structuring qualitative content analysis according to the approach described by Kuckartz and Rädiker [24].

Ethical Considerations

The study was approved by the Ethics Committee of the Albert-Ludwigs-University Freiburg (24-1064-S1; vote from May 23, 2024). Written informed consent is obtained from each participant (patient and physician) before participation. The privacy and confidentiality of participants' data and identity are assured through processes described in a detailed data protection concept approved by the responsible data protection officer. Respondents will receive a shopping voucher worth €50 (US $59) as compensation for their participation.

Results

The Qualy-GAIN study was funded from January 2023 to December 2025. Data collection for the prospective longitudinal study started in July 2024 and ended in June 2025. As of May 2025, data collection for 3 measurement points had been completed. Final data analyses began in summer 2025, with results expected to be submitted for publication in spring 2026.

Discussion

Anticipated Findings and Implications

The Qualy-GAIN study aims to contribute to improving the care and QoL of patients with multi-organ autoimmune diseases. This study uses a mixed methods approach that integrates quantitative and qualitative methods and is conducted by a multidisciplinary research team. Data from patients with multi-organ autoimmunity treated at 3 medical centers in Germany are collected in a longitudinal study with 4 measurement time points and will be linked to objective medical parameters from the GAIN registry.

Our first hypothesis (hypothesis 1) is that the fulfillment of preferences at an earlier stage is associated with a subsequently better self-reported health status and better health-related QoL. In line with previous studies [2,3] for other diseases, this is expected to improve our knowledge of the importance of patients' preferences. We expect hypothesis 1 to be supported based on our data analysis. Our longitudinal dataset covering a relatively large sample of patients with multi-organ autoimmune diseases will enable us to detect small-to-medium effects of preference fulfillment on patient-reported outcomes.

We assume in the second hypothesis (hypothesis 2) that there are plausible correlations between self-assessed health status and QoL, and pathophysiology and physical symptoms. Given the complex and heterogeneous phenotypes of multi-organ autoimmunity [1], we anticipate only moderate correlations between subjective health status and objective clinical parameters.

Furthermore, we expect in hypothesis 3 that providing physicians feedback from patients regarding their health status, QoL, and preferences will be perceived as valuable for optimizing treatment and patient-centered care. Such feedback may strengthen patient-physician communication and enhance patient-centered care.

No serious risks for the patients are to be expected from participation in the study. It requires approximately half an hour to complete the online questionnaire at 4 measurement points within 9 months. If participants are unable to take part in some of the 4 online surveys for health reasons, they can skip their participation at these measurement points. Although this is associated with a loss of data for the study, information provided by the participants up to that point can still be used for the study. As no sensitive questions are asked in the online surveys, it can be assumed that the participating patients will not be exposed to any risk of harm by completing the online questionnaire.

As a result of the study, a questionnaire package will become available that can be used to assess health-related QoL, disease-specific health status, treatment preferences, and the fulfillment of treatment preferences. In addition, data collected with these instruments will provide the basis for testing the formulated hypotheses. Due to the voluntary nature of participation and the minimal burden on participation, we consider the risk of participants experiencing disadvantages as a result of participating in the project to be extremely low. Overall, we therefore believe that the benefits of the study are expected to outweigh the potential risks.

Strengths and Limitations

A strength of the study is the prospective longitudinal design with multiple measurement points, which enables us to analyze the relationships between preference fulfillment, patient-reported outcomes, and clinical parameters. Another strength is that a relatively large sample of patients with an ultrarare disease who are treated in 3 medical centers can be examined in a longitudinal study.

Limitations of our study include the possibility of responder bias, as comparatively motivated and relatively healthy patients may respond disproportionately to the invitation to participate, leading to a selective sample of relatively healthy patients. Another limitation is that the patient-reported outcomes are collected 4 times at intervals of 3 months, whereas the measurement of objective medical parameters from the GAIN registry does not adhere to these time intervals. This may limit the comparability of the data and our ability to examine direct correlations between subjective and objective measures. To address these limitations, efforts were made to recruit representative patient populations through reminder strategies.

XSL•FO
RenderX

Dissemination

We will disseminate our research results in various ways, including abstracts, posters, and presentations at conferences, as well as articles in peer-reviewed journals. The newly developed patient preference and fulfillment questionnaire will be published in open-access journals, making it available for future research and practice. Furthermore, we will make our results available to the physicians at the 3 centers involved and will invite participating patients to discuss the results with us.

Conclusions

The procedure proposed here and the questionnaire set created may be transferred to other rare diseases to promote greater inclusion of patient-reported experience measures and PROMs in the everyday care of these patients. In case the questionnaires evaluated in this study prove to be reliable, valid, and practically useful, they can be transferred to the routine treatment of patients with multi-organ autoimmune diseases to improve their treatment and promote patient-centered care.

Acknowledgments

The authors would like to thank Gaby Müller, Michaela Ruf, PD Dr Georgios Sogkas, Dr Faranaz Atschekzei and Prof Dr Renate Krüger for their great support in inviting and recruiting patients for this study. The authors are grateful to Dr Gerhard Kindle and Jana Kopp for preparing the invitation of patients from the GAIN registry for inclusion in this study. The authors would like to thank Diana Druschke for her preliminary work on particularly relevant aspects of the treatment of patients with rare diseases. The authors acknowledge support by the Open Access Publication Fund of the University of Freiburg.

Funding

This work was supported by the Federal Ministry of Research, Technology and Space (BMFTR), Kapelle-Ufer 1, 10117 Berlin, Germany, through a grant to the German Auto-Immunity Network (GAIN), grant code 01GM2206A. The funder had no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript. In addition, BG receives funding from the Deutsche Forschungsgemeinschaft through the RESIST (Resolving Infection Susceptibility) Cluster of Excellence (EXC 2155; 390874280) and the Centre for Integrative Biological Signalling Studies Cluster of Excellence (EXC 2189).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

When informed consent is given by the patients, the data collected will be transferred to the German Multi-Organ Auto-Immunity Network register.

Authors' Contributions

Conceptualization: EF-G, GM, BG, AN, MB, CS, SH, JS
Data curation: CS, MB, SH, JS
Formal analysis: CS, MB, SH, JS
Funding acquisition: EF-G, GM, BG, AN
Investigation: CS, MB, SH, JS, EF-G, GM
Methodology: CS, MB, SH, JS, EF-G, GM
Project administration: CS, MB, SH, JS, EF-G, and GM
Supervision: EF-G, GM
Writing—original draft: MB, CS
Writing—review and editing: MB, CS, SH, JS, EF-G, GM, BG, AN

Conflicts of Interest

None declared.

Checklist 1
SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist.
[PDF File, 153 KB - resprot_v15i1e77786_app1.pdf]

References

1. Grimbacher B, Warnatz K, Yong PFK, Korganow AS, Peter HH. The crossroads of autoimmunity and immunodeficiency: lessons from polygenic traits and monogenic defects. J Allergy Clin Immunol 2016 Jan;137(1):3-17. [doi: 10.1016/j.jaci.2015.11.004] [Medline: 26768758]
2. Whittal A, Meregaglia M, Nicod E. The use of patient-reported outcome measures in rare diseases and implications for health technology assessment. Patient 2021 Sep;14(5):485-503. [doi: 10.1007/s40271-020-00493-w] [Medline: 33462774]

XSL•FO
RenderX

3. Mattukat K, Boehm P, Raberger K, Schaefer C, Keyszer G, Mau W. How much information and participation do patients with inflammatory rheumatic diseases prefer in interaction with physicians? Results of a participatory research project. Patient Prefer Adherence 2019 Dec 17;13:2145-2158. [doi: 10.2147/PPA.S209346] [Medline: 31908422]
4. Weeger S, Farin E. The effect of the patient-physician relationship on health-related quality of life after cardiac rehabilitation. Disabil Rehabil 2017 Mar;39(5):468-476. [doi: 10.3109/09638288.2016.1146360] [Medline: 26940035]
5. Lenderking WR, Anatchkova M, Pokrzywinski R, Skalicky A, Martin ML, Gelhorn H. Measuring health-related quality of life in patients with rare disease. J Patient Rep Outcomes 2021 Jul 20;5(1):61. [doi: 10.1186/s41687-021-00336-8] [Medline: 34283357]
6. Bril V, Lampe J, Cooper N, Kiessling P, Gardulf A. Patient-reported preferences for subcutaneous or intravenous administration of parenteral drug treatments in adults with immune disorders: a systematic review and meta-analysis. J Comp Eff Res 2024 Sep;13(9):e230171. [doi: 10.57264/cer-2023-0171] [Medline: 39115099]
7. Babac A, Damm K, Graf von der Schulenburg JM. Patient-reported data informing early benefit assessment of rare diseases in Germany: a systematic review. Health Econ Rev 2019 Dec 12;9(1):34. [doi: 10.1186/s13561-019-0251-9] [Medline: 31832812]
8. Chan AW, Tetzlaff JM, Gøtzsche PC, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013 Jan 8;346:e7586. [doi: 10.1136/bmj.e7586] [Medline: 23303884]
9. Stapornwongkul C, Nieters A, Staus P, et al. Research on rare diseases in Germany - the GAIN Registry: a registry for individuals with congenital multi-organ autoimmune diseases. J Health Monit 2023 Dec 13;8(4):24-30. [doi: 10.25646/11732] [Medline: 38235015]
10. Bayrhuber M, Tinsel I, Goldacker S, et al. Perceived health of patients with common variable immunodeficiency - a cluster analysis. Clin Exp Immunol 2019 Apr;196(1):76-85. [doi: 10.1111/cei.13252] [Medline: 30637715]
11. Bühner M. Einführung in Die Test- Und Fragebogenkonstruktion, 4th edition: Pearson; 2021. URL: https://elibrary.pearson.de/book/99.150005/9783863268138
12. MacCallum RC, Widaman KF, Zhang S, Hong S. Sample size in factor analysis. Psychol Methods 1999;4(1):84-99. [doi: 10.1037/1082-989X.4.1.84]
13. Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. J Clin Epidemiol 2004 Aug;57(8):785-794. [doi: 10.1016/j.jclinepi.2003.12.013] [Medline: 15485730]
14. Hemming K, Eldridge S, Forbes G, Weijer C, Taljaard M. How to design efficient cluster randomised trials. BMJ 2017 Jul 14;358:j3064. [doi: 10.1136/bmj.j3064] [Medline: 28710062]
15. Ware JJ, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
16. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001 Jul;33(5):337-343. [doi: 10.3109/07853890109002087] [Medline: 11491192]
17. Montan I, Löwe B, Cella D, Mehnert A, Hinz A. General population norms for the functional assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale. Value Health 2018 Nov;21(11):1313-1321. [doi: 10.1016/j.jval.2018.03.013] [Medline: 30442279]
18. Jones PW, Harding G, Berry P, Wiklund I, Chen WH, Kline Leidy N. Development and first validation of the COPD Assessment Test. Eur Respir J 2009 Sep;34(3):648-654. [doi: 10.1183/09031936.00102509] [Medline: 19720809]
19. Irvine EJ, Zhou Q, Thompson AK. The Short Inflammatory Bowel Disease Questionnaire: a quality of life instrument for community physicians managing inflammatory bowel disease. CCRPT investigators. Canadian Crohn's Relapse Prevention Trial. Am J Gastroenterol 1996 Aug;91(8):1571-1578. [Medline: 8759664]
20. Rose M, Fliege H, Hildebrandt M, et al. Validation of the new German translation version of the "Short Inflammatory Bowel Disease Questionnaire" (SIBDQ). Z Gastroenterol 2000 Apr;38(4):277-286. [doi: 10.1055/s-2000-14868] [Medline: 10820859]
21. Finlay AY, Khan GK. Dermatology Life Quality Index (DLQI)--a simple practical measure for routine clinical use. Clin Exp Dermatol 1994 May;19(3):210-216. [doi: 10.1111/j.1365-2230.1994.tb01167.x] [Medline: 8033378]
22. Swiontkowski MF, Engelberg R, Martin DP, Agel J. Short musculoskeletal function assessment questionnaire: validity, reliability, and responsiveness. J Bone Joint Surg Am 1999 Sep;81(9):1245-1260. [doi: 10.2106/00004623-199909000-00006] [Medline: 10505521]
23. Wollmerstedt N, Faller H, Ackermann H, et al. Evaluation of the Extra Short Musculoskeletal Function Assessment questionnaire XSMFA-D in patients with musculoskeletal disorders and surgical or medical in-patient treatment. Rehabilitation (Stuttg) 2006 Apr;45(2):78-87. [doi: 10.1055/s-2005-915337] [Medline: 16575713]
24. Kuckartz U, Rädiker S. Fokussierte Interviewanalyse Mit MAXQDA: Schritt Für Schritt: Springer; 2024. URL: https://link.springer.com/book/10.1007/978-3-658-40212-9

Abbreviations

GAIN: German Multi-Organ Auto-Immunity Network

XSL•FO
RenderX

ICC: intracluster correlation
PROM: patient-reported outcome measure
QoL: quality of life
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Edited by J Sarvestan; submitted 20.May.2025; peer-reviewed by A AL-Asadi; revised version received 27.Nov.2025; accepted 08.Dec.2025; published 09.Mar.2026.

Please cite as:
Bayrhuber M, Schlett C, Slesaczeck J, Herbst S, Mueller G, Nieters A, Grimbacher B, Farin-Glattacker E
Quality of Life of Patients With Multi-Organ Autoimmune Disease and Its Relationship to Patient-Centered Care: Protocol for a Longitudinal Cohort Study
JMIR Res Protoc 2026;15:e77786
URL: https://www.researchprotocols.org/2026/1/e77786
doi:10.2196/77786

© Marianne Bayrhuber, Christian Schlett, Julia Slesaczeck, Sabrina Herbst, Gabriele Mueller, Alexandra Nieters, Bodo Grimbacher, Erik Farin-Glattacker. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The Systematic Psychodynamic Multidimensional Organizational Assessment Questionnaire: Protocol for a Development and Validation Study

Anne-Maria Müller, MA; Sophia Sachs, MSc; Yannik Rieder, MSc; Claas Lahmann, Prof Dr Med

Department of Psychosomatic Medicine and Psychotherapy, Faculty of Medicine, University of Freiburg, University Medical Center Freiburg, Hauptstrasse 8, Freiburg, Germany

Corresponding Author:
Anne-Maria Müller, MA
Department of Psychosomatic Medicine and Psychotherapy, Faculty of Medicine, University of Freiburg, University Medical Center Freiburg, Hauptstrasse 8, Freiburg, Germany

Abstract

Background: Systems psychodynamics provide valuable insights into organizational development. However, to date, instruments that can reliably assess organizations based on systems psychodynamic theories are scarce. The Systematic Multidimensional Organizational Assessment (SyMOA) is a qualitative instrument that provides an in-depth systems psychodynamic analysis of organizational dynamics using a semistructured interview guide. To complement the method, a standardized, quantitative self-assessment questionnaire will be developed and validated.

Objective: The aim of this study is to develop and psychometrically validate an instrument for assessing organizational dynamics based on the SyMOA diagnostic system. The questionnaire is intended to provide a scientifically grounded yet practical diagnostic tool applicable in both research and corporate practice. The findings aim to contribute to the advancement of systems psychodynamic theory and to serve as a foundation for evidence-based interventions in organizational change processes.

Methods: The study follows a multistage development and validation process. First, the SyMOA instrument will be transformed into a questionnaire battery, and the items will be evaluated by experts (expert validity). The items will be tested through factor and item analyses by an online panel (n=150) and iteratively refined. Test-retest reliability will subsequently be assessed in a separate sample (n=150). Finally, factorial and discriminant validity will be examined in a larger validation sample (n=800).

Results: As of February 2026, a first draft of 158 items has been developed based on dimension 1 of the SyMOA framework. The draft has undergone an expert review process with 2 experts in psychodynamics, who provided feedback on content validity and conceptual alignment. Approximately 20% of the items have been revised to improve clarity and theoretical precision. Data collection using a panel started in May 2025 and concluded in late August 2025, with iterative item analysis conducted thereafter.

Conclusions: This study will evaluate the methodological robustness and usability of a questionnaire-based operationalization of systems psychodynamic organizational diagnostics. The anticipated findings are intended to guide further instrument development, normative calibration, and the responsible application of systems psychodynamic approaches in organizational research and practice.

Trial Registration: German Clinical Trials Register DRKS00039410; https://www.drks.de/search/de/trial/DRKS00039410/details

International Registered Report Identifier (IRRID): DERR1-10.2196/78530

(JMIR Res Protoc 2026;15:e78530) doi:10.2196/78530

KEYWORDS

questionnaire; organizational assessment; systems psychodynamics; psychodynamic theory; organizational health

Introduction

Background

The increasing complexity, volatility, and emotional demands of modern organizational life have heightened the need for diagnostic approaches that can access not only observable behaviors but also the latent, unconscious dynamics influencing individuals and systems. While conventional organizational diagnostics typically emphasize cognitive, behavioral, or structural dimensions, they often overlook the dynamics that drive behavior beneath the surface [1,2]. Systems psychodynamic diagnostics are a vital and underdeveloped avenue for capturing these phenomena in a scientifically grounded manner.

Rooted in the Tavistock tradition, the systems psychodynamic approach integrates psychoanalytic concepts with open systems

XSL•FO
RenderX

theory to analyze how unconscious processes manifest within individuals, teams, and organizational cultures [2-5]. Key concepts include social defense mechanisms, unconscious conflicts, and the projection of anxiety onto systems or structures. These ideas provide a theoretical foundation for understanding phenomena such as resistance to change, dysfunctional leadership dynamics, or behaviors not adequately explained by surface-level assessments.

Despite its conceptual richness, the systems psychodynamic paradigm has historically relied on qualitative methodologies such as organizational role analysis [6], reflective team consultations [7,8], and projective methods such as metaphor elicitation and storytelling [9-11]. These tools offer deep insight but often lack psychometric validation—rendering them less suitable for large-scale organizational research.

On the one hand, organizational practitioners and researchers acknowledge the importance of unconscious dynamics; on the other hand, they lack scalable diagnostic instruments that can reliably quantify such processes while respecting the depth and contextual sensitivity of the systems psychodynamic framework. The challenge lies in designing diagnostics that do not reduce complex unconscious phenomena to simplistic metrics but, rather, operationalize key constructs with theoretical fidelity and psychometric validity.

This study protocol describes the development and validation of a novel psychometric instrument designed to assess unconscious dynamics in organizational contexts: the Systematic Multidimensional Organizational Assessment–Questionnaire (SyMOA-Q). Drawing on both psychodynamic constructs and empirical approaches, the aim is to create a self-assessment questionnaire that is theoretically grounded, methodologically rigorous, and practically useful for both researchers and organizational consultants.

Prior Work

The SyMOA-Q builds on the authors' previous work, the Systematic Multidimensional Organizational Assessment (SyMOA; C Lahmann et al, unpublished report), a qualitative systems psychodynamic diagnostic method that assesses organizational dynamics based on 3 dimensions (Table 1). The SyMOA was developed using a theory-to-research approach based on psychodynamic and organizational theories and was refined through an iterative process of theory development, empirical application, and theoretical refinement. It combines the logic of clinical diagnostics based on the *Operationalized Psychodynamic Diagnosis* [12] and organizational research. The SyMOA is a comprehensive and in-depth instrument that captures systemic and psychodynamic aspects within an organization that, to the best of our knowledge, no existing instrument has been able to map so far. A follow-up study testing the quality of the instrument, in particular the intercoder reliability and construct validity, is in preparation [13].

XSL•FO
RenderX

Table . Systematic Multidimensional Organizational Assessment dimensions.

Dimension	Description
Dimension 1: sociotechnical integration and the organization's internal level of functioning	The first dimension describes the organization's understanding of current and past challenges, including its resources and barriers to change. Additionally, this dimension delves deeper beneath the surface to examine the organization's internal level of functioning. This includes how the organization as a system perceives itself and its environment; how it regulates itself both internally and externally in everyday situations as well as in critical moments; how internal and external communication functions; and how well it establishes bonds both internally among employees and externally with stakeholders such as customers, the market, and supply chains. Furthermore, this dimension captures the defense mechanisms typically used by the organization.
Dimension 2: internal relationship dynamics of the organization	The second dimension assesses the relationship dynamics between different groups within the organization. This includes relationship patterns in typical situations between individual employees, within teams, between teams, between management and stakeholders, and between management and employees.
Dimension 3: unconscious organizational conflicts	On the basis of psychodynamic theories of motivation, emotions, and unconscious conflicts, this dimension conceptualizes the internal conflicts within organizations. Typical conflicts include dependence vs autonomy, control vs subordination, intrinsic vs extrinsic organization-based self-worth conflicts, provision vs self-sufficiency, organizational exploitation vs prosocial engagement, market dominance, and identity conflicts. Conflicts are initially external, rooted in the early history of the company, but over time they become internalized and, thus, develop into unconscious internal conflicts. Conflict represents the clash of opposing positions, the collision of motives, desires, needs, values, and ideas. Conflicts between competing needs are a normal and everyday occurrence, such as between the need for self-determination in the workplace on the one hand and the desire for leadership and guidance on the other. If the underlying conflict is adequately resolved, individuals within the organization can shift flexibly between these needs depending on the situation. However, stress or dysfunction arises when the organization fails to develop an adequate strategy to resolve the fundamental conflict and persistently tries to rigidly implement one pole of these needs. These conflicts are repeatedly triggered by situational factors throughout the organization's development.

Research Objectives

Given the predominance of qualitative approaches in systems psychodynamic organizational research and the absence of validated quantitative instruments, this study seeks to develop and examine the psychometric properties of a questionnaire that operationalizes dimension 1 of the SyMOA framework. The objective is to establish an initial measurement model that captures core systems psychodynamic capacities—such as communication, attachment, awareness, and regulation—as perceived by organizational members. By providing a theoretically grounded and empirically testable tool, this study addresses a critical gap in the literature and lays the foundation for subsequent validation of the remaining SyMOA dimensions.

Methods

Aim, Design, and Setting

The qualitative methodology of the SyMOA will be converted into a standardized, quantitative self-assessment instrument. This study will first develop and validate a questionnaire for dimension 1 of the SyMOA diagnostics: sociotechnical integration and the organization's internal level of functioning. The reliability and validity of the newly developed instrument, particularly factorial and discriminant validity, will be determined. A panel will be used for participant recruitment and data collection. The methodological approach of this study follows established best practices in psychometric instrument development and validation, as demonstrated in prior research across diverse applied fields such as health psychology, behavioral assessment, and attitudinal measurement [14-16]. These studies inform the multistage design adopted in this study, including pilot testing, reliability assessment, and factorial validation. The study was initially registered in the Freiburg Register for Clinical Studies (FRKS005727), which is not publicly accessible. It was later additionally registered in the German Clinical Trials Register (DRKS00039410) to ensure public access of the trial record.

Study Procedure and Measures

The study follows a multistage development and validation process, which is scheduled in 4 phases (Figure 1). However, although data collection spans several months, each participant is involved only briefly, completing the questionnaire once or, at most, twice.

XSL•FO
RenderX

Figure 1. Study flow diagram. OPD: *Operationalized Psychodynamic Diagnosis*.

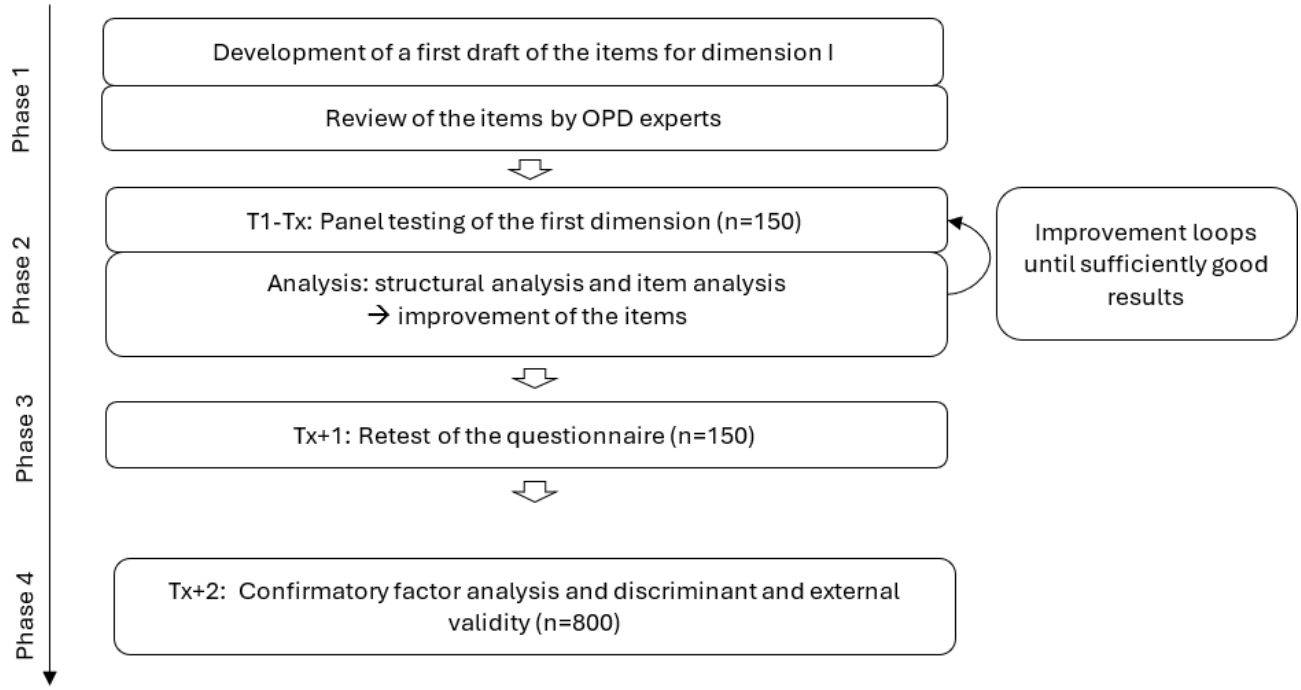

First, the SyMOA scale for dimension 1 will be converted into a set of items and evaluated by 2 experts (expert validity). To this end, a preliminary set of items for the dimension will be created based on the qualitative SyMOA instrument. The aim is to have at least 80 items, which will be answered on a 5-point Likert scale. The items will be presented in German and mostly formulated negatively, with 20% formulated positively. In a second step, the items will be tested using structural and item-level analyses by an online panel sample (n=150) and refined iteratively based on their psychometric performance. Data collection will be conducted via LimeSurvey (LimeSurvey GmbH), a secure web-based survey platform. The development process will be completed if no further adjustment of the items is necessary.

As a third step, the retest reliability will be checked by sending the questionnaire to the same sample as that in the previous round (n=150) after 2 weeks. In a fourth step, a confirmatory factor analysis (CFA) will be carried out to check the fit of the theoretically assumed factor structure based on the SyMOA construct (sample of n=800). In addition, the discriminant validity of the questionnaires will be assessed. For this purpose, existing questionnaires used in clinical practice, such as the *Operationalisierte Psychodynamische Diagnostik–Strukturfragebogen* (OPD-SF) [17], will be used. To be able to assume that the surveyed constructs relate to the organization rather than the person completing the questionnaire, there should be a maximum of moderate correlations. For the evaluation of external validity, employee satisfaction (*Inventar zur Erfassung der Arbeitszufriedenheit* [18]) and work-related quality of life (World Health Organization Quality of Life Scale–Brief [19]) will also be administered in the last round. Although neither scale directly measures the assumed construct of organizational health dynamics, it can be assumed that there is a strong correlation between organizational health and the satisfaction and quality of life exhibited by employees. In addition, sociodemographic data such as age, gender, educational qualifications, professional experience, employment relationship, position or management function, sector, size of the company (eg, number of employees: 30-50 or 50-150), and size of the work area (eg, number of employees: 1-5, 5-10, or 10-20) will be evaluated in relation to the results.

For the development of the questionnaire for T1 to Tx (ie, from the start of data collection to the final item analysis round), as well as for the verification of the test-retest reliability for Tx+1 and the confirmatory sample for Tx+2, the recorded variables and their operationalization are listed in Table 2.

XSL•FO
RenderX

Table . Data collection with instruments and time point of measurement.

Time point of measurement and measured variable	Operationalization
T1[a]-Tx[b]—development of questionnaire based on dimension 1	
Quality of the items of the scale	Item analysis: item difficulty, item selectivity, item variance, and distribution
Reliability of the scale	Internal consistency of the scale
Tx+1[c]—retest reliability	
Retest reliability	Agreement of the results of Tx+1 with those of Tx
Tx+2[d]—CFA[e] and validity	
Internal level of functioning of the organization	SyMOA-Q[f] based on dimension 1 of the SyMOA[g]
Factorial validity of the scales	CFA
Sociodemographic data	Age, gender, educational qualifications, professional experience, employment relationship, position or management function, sector, size of the company (30-50, 50-150, 150-250, 250 - 1000, or >1000), and size of work area (1-5, 5-10, 10-20, 20-50, and >50)
Discriminant validity via structural level of employees	OPD-SFK [17][h]
Employee satisfaction as an external criterion	IAZ[i] [18]
Work-related quality of life as an external criterion	WHOQOL-BREF [19][j]

[a]T1 marks the start of data collection.

[b]Tx refers to the final round of phase 2 (completion of item analysis).

[c]Tx+1 indicates the retest reliability assessment (phase 3).

[d]Tx+2 refers to the confirmatory factor analysis and validity testing (phase 4).

[e]CFA: confirmatory factor analysis.

[f]SyMOA-Q: Systematic Multidimensional Organizational Assessment–Questionnaire.

[g]SyMOA: Systematic Multidimensional Organizational Assessment.

[h]OPD-SFK: *Operationalisierte Psychodynamische Diagnostik–Strukturfragebogen Kurz*.

[i]IAZ: Inventar zur Erfassung der Arbeitszufriedenheit.

[j]WHOQOL-BREF: World Health Organization Quality of Life Scale–Brief.

Recruitment and Data Collection

The panel will be used for participant recruitment and data collection. Potential participants are registered users who meet the inclusion criteria. They must be aged ≥18 years, speak German, and be employed. Furthermore, the company with which they work must have at least 30 employees. Illiterate persons will be excluded. A message with the link to the study and the offer to participate will be sent to the database of potential participants until the required subsample of 150 people is reached for the first rounds. For the last round, a sample of 800 persons will be recruited.

Before the study, participants will receive written information about the procedure and any expected risks and side effects and the names of the people to contact. Informed consent will be obtained electronically in advance from the participants who agree to continue with the study.

Data Analysis and Measures

Power

A sample size of 150 will be used for the iterative development of the questionnaire to check the retest reliability, internal consistency, and item analyses. This is methodologically justified as item analyses [20], internal consistency [21], and retest reliability [22] require between 100 and 200 people for a robust estimate.

For the CFA, the aim is to achieve a sample size of at least 800 following the rule of thumb by Bühner [23], with medium communalities (h^2>0.50) and approximately 20 factors across all dimensions.

Statistical Analysis

The quantitative data will be analyzed using SPSS (IBM Corp), R (R Foundation for Statistical Computing), and JASP. The following statistical evaluations will be performed: a CFA (global fit using the chi-square test and goodness-of-fit indexes [root mean square error of approximation<0.05; Tucker-Lewis index and comparative fit index>0.9]), local fit (factor loadings>0.4), iterative item analyses (item distribution and variance, item difficulty [0.2-0.8], and discriminatory power [>0.3]), and verification of correlations and internal consistency of the scales (retest reliability [>0.7] and discriminant validity). The data from the confirmatory sample will also be checked for correlations between variables.

XSL•FO
RenderX

Process Validation

The entire research process will be documented in detail to ensure the reliability and transparency of the procedure. This includes data collection, analysis, and interpretation of the data. It also ensures that the questionnaire and the resulting findings fit into existing theoretical frameworks or contribute to the development of new theories. Theory-led validation underpins the relevance and depth of the system.

This study is part of a larger project. The remaining 2 SyMOA dimensions are also planned to be developed into questionnaires.

Ethical Considerations

This study was approved by Ethikkommission Universitätsklinikum Freiburg on May 9, 2025, under trial ID 25-1117-S2. This study is conducted in accordance with the local legislation and institutional requirements. The participants will provide their written informed consent to take part in this study. All data will be collected via the external online panel provider Bilendi. No personally identifiable information will be collected or stored by the research team, and all responses will be fully anonymized before analysis. As such, participant confidentiality is guaranteed throughout the study. Panel participants will be compensated by Bilendi in accordance with the platform's standard remuneration procedures.

Results

As of April 2025, a first draft of 158 items has been developed based on dimension 1 of the SyMOA framework. These items cover constructs such as organizational self-regulation, perception of challenges, modes of communication, and attachment capacity. The draft has undergone an expert review process with 2 OPD-trained experts, who provided feedback on content validity and conceptual alignment. On the basis of this feedback, approximately 20% of the items have been revised to enhance clarity and theoretical precision.

Data collection using the online panel is scheduled for the end of May 2025. Iterative item analyses, including evaluation of item distributions, reliability, and preliminary factor structure, will be conducted following data collection.

Discussion

Principal Findings and Comparison to Prior Work

Building on the authors' previous work, the SyMOA, a questionnaire will be developed. By formalizing systems psychodynamic diagnostics through psychometric means, we seek to enhance organizational self-awareness; improve leadership insight; and contribute to healthier, more emotionally sustainable workplaces. This work addresses a significant gap in the organizational diagnostics landscape and contributes to a broader agenda of integrating psychodynamic psychological insights into evidence-based organizational development.

Strengths and Limitations

The questionnaire can be used as a screening instrument or as a supplement to comprehensive multimodal diagnostics. In addition, the application of the questionnaire is more user-friendly and economical as, in contrast to the SyMOA interview, no one-on-one setup with a 1-hour time frame is required. Thus, the questionnaire enables a larger number of employees to be surveyed within a shorter time.

This study is firmly grounded in systems psychodynamics theory, translating the new qualitative SyMOA diagnostic tool into a quantitative self-assessment instrument. The multistage design—including expert validation, pilot testing, test-retest reliability, and CFA—reinforces the methodological rigor and psychometric robustness of the resulting questionnaire.

The instrument's development addresses all key aspects of validity: expert validity, factorial and discriminant validity, and criterion-related (external) validity. This comprehensive approach provides a well-rounded assessment of the questionnaire's reliability and practical utility.

Importantly, the study fills a well-recognized methodological gap in the field: the absence of empirically validated psychometric tools for systems psychodynamic organizational diagnostics. By offering a scalable, user-friendly questionnaire, this study has high practical relevance for organizational consultants, human resources professionals, and academic researchers alike.

Finally, this study acknowledges the interplay between organizational-level dynamics and intraindividual experiences and, thus, the possible influence of personal perceptions and character traits on collective organizational phenomena. By testing discriminant validity against clinical diagnostic tools (OPD-SF), this study is able to assess this interplay.

Future Directions

Future research should extend this work by developing and validating questionnaire-based measures for the remaining dimensions (2-3) of the SyMOA framework. This study focuses exclusively on the development and validation of a questionnaire based on dimension 1 of the SyMOA framework, which addresses sociotechnical integration and the internal level of organizational functioning. Accordingly, the validity evidence obtained in this study applies only to this dimension and does not yet support claims regarding the assessment of organizational dynamics in their entirety. Expanding beyond dimension 1 will enable a more comprehensive operationalization of systems psychodynamic organizational diagnostics and allow for an integrated assessment across multiple organizational domains.

A further essential step will be the normative calibration of the instrument. This study is explicitly focused on the development and initial psychometric validation of the SyMOA-Q, including the examination of factorial structure, internal consistency, and first evidence of construct validity. Establishing norms based on data from organizations across different industries, organizational sizes, and hierarchical levels is necessary to enhance the interpretation of scores. In this study, scoring is conceptualized at the level of factor scores for the subscales of dimension 1, with the option of examining higher-order factor structures in subsequent research.

XSL•FO
RenderX

In addition, multilevel research designs represent a key next step. This study is an initial step in the development and psychometric validation of the SyMOA-Q focusing on systems psychodynamic organizational dynamics as perceived by individuals. Future studies should use clustered sampling designs in which multiple respondents are nested within clearly identified organizations. This will allow for the examination of aggregation statistics (eg, intraclass correlations and within-group agreement indexes) to determine whether and under which conditions individual perceptions can be meaningfully aggregated to the organizational level. Building on these analyses, multilevel CFAs should be conducted to disentangle individual- and organization-level variance components and test the factorial structure of the SyMOA dimensions across levels of analysis.

Dissemination Plan

The results of this study will be disseminated through multiple channels targeting both academic and applied audiences. Scientific dissemination will include publication in peer-reviewed journals focusing on organizational psychology, psychometrics, and systems psychodynamics research, as well as presentations at national and international conferences in these fields.

To support practical implementation, the SyMOA-Q will be introduced to organizational consultants, human resources professionals, and leadership development practitioners through workshops, professional training programs, and practitioner-oriented publications. In addition, a standardized manual including theoretical background, administration guidelines, scoring procedures, normative data, and interpretation frameworks will be developed to ensure consistent and responsible use of the instrument.

All data collected in future studies, including clustered organizational samples for norming, will be pseudonymized to protect individual participants and organizations. Norms and interpretive benchmarks will be shared only at aggregated levels, preventing the identification of specific organizations and ensuring ethical and responsible dissemination.

Further dissemination will involve collaboration with organizations participating in the norming studies, allowing for feedback loops between research and practice. Finally, the instrument may be made available in digital format to facilitate future research collaborations.

Acknowledgments

The authors used DeepL (DeepL SE) and ChatGPT-5 (OpenAI) solely for refinement, correction, and editing of the manuscript to improve the quality and clarity of language.

Funding

The study is funded by the Department of Psychosomatic Medicine and Psychotherapy, Department of Mental Health at University Medical Center Freiburg.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AMM contributed to conceptualization, background, methodology, writing—original draft, writing—review and editing, and project administration. SS contributed substantially to the preparation for ethics approval, writing—original draft, writing—review and editing, and methodology. YR contributed to methodology and writing—review and editing. CL contributed to conceptualization and writing—review and editing. Each author has made a substantial contribution to the work. All authors have approved the final manuscript. They agree both to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any part of the work, even ones in which the authors were not personally involved, are appropriately investigated and resolved and the resolution is documented in the literature.

Conflicts of Interest

None declared.

References

1. Mikkelsen EN, Gray B, Petersen A. Unconscious processes of organizing: intergroup conflict in mental health care. J Manag Stud 2020;57(7):1355-1383. [doi: 10.1111/joms.12611]
2. McFillen JM, O'Neil DA, Balzer WK, Varney GH. Organizational diagnosis: an evidence-based approach. J Change Manag 2013 Jun 18;13(2):223-246. [doi: 10.1080/14697017.2012.679290]
3. Petriglieri G, Petriglieri JL. The return of the oppressed: a systems psychodynamic approach to organization studies. Acad Manag Ann 2020;14(1):411-449. [doi: 10.5465/annals.2017.0007]
4. Wolfe NS. The Living Organization: Transforming Business to Create Extraordinary Results: Quantum Leaders Publishing; 2011.

XSL•FO
RenderX

5. Stouten J, Rousseau DM, De Cremer D. Successful organizational change: integrating the management practice and scholarly literatures. Acad Manag Ann 2018;12(2):752-788. [doi: 10.5465/annals.2016.0095]
6. Long S, Newton J, Sievers B, editors. Coaching in Depth: The Organizational Role Analysis Approach: Routledge; 2018.
7. Obholzer A, Roberts VZ, editors. The Unconscious at Work: A Tavistock Approach to Making Sense of Organizational Life, 2nd edition: Routledge; 2019.
8. Long S. Socioanalytic Methods: Discovering the Hidden in Organisations and Social Systems: Routledge; 2018.
9. Morgan G. Images of Organization, th edition: SAGE Publications; 2006.
10. Gabriel Y. Storytelling in Organizations: Facts, Fictions, and Fantasies: Oxford University Press; 2000.
11. Barry D, Elmes M. Strategy retold: toward a narrative view of strategic discourse. Acad Manag Rev 1997;22:429-452. [doi: 10.2307/259329]
12. OPD-3: Operationalisierte Psychodynamische Diagnostik: Das Manual Für Diagnostik Und Therapieplanung: Hogrefe; 2023.
13. Müller AM, Sachs S, Rieder Y, Lahmann C. Systematic multidimensional organizational assessment: protocol for assessing the reliability and validity of a qualitative psychodynamic diagnostic instrument for organizations. in preparation. . Preprint posted online on 2025 URL: https://www.researchgate.net/publication/399035155_Systemic_Multidimensional_Organizational_Assessment_Questionnaire_Protocol_for_the_Development_and_Validation_of_a_Systemic-Psychodynamic_Organizational_Diagnostic_Questionnaire_Preprint
14. Shuja KH, Aqeel M, Khan KR. Psychometric development and validation of attitude rating scale towards women empowerment: across male and female university population in Pakistan. Int J Hum Rights Healthc 2020 Jul 4;13(5):405-420. [doi: 10.1108/IJHRH-12-2019-0090]
15. Nisar H, Aqeel M, Ahmad A. Indigenous need arise to protect human from self-harm behavior in Pakistan: translation and validation of inventory of statements about self-injury. Int J Hum Rights Healthc 2020 Jul 3;13(5):421-433. [doi: 10.1108/IJHRH-10-2019-0080]
16. Naz F, Shuja KH, Aqeel M, et al. A hindrance to proper health care: psychometric development and validation of opiophobia questionnaire among doctors in Pakistan. Int J Hum Rights Healthc 2021;16(3):221-239. [doi: 10.1108/IJHRH-12-2020-0127]
17. Ehrenthal JC, Dinger U, Schauenburg H, Horsch L, Dahlbender RW, Gierk B. Entwicklung einer Zwölf-Item-Version des OPD-Strukturfragebogens (OPD-SFK). Development of a 12-item version of the OPD-Structure Questionnaire (OPD-SQS). Z Psychosom Med Psychother 2015 Sep 1;61(3):262-274. [doi: 10.13109/zptm.2015.61.3.262]
18. Kanning UP. Inventar Zur Erfassung Der Arbeitszufriedenheit (IAZ): Hogrefe; 2021. URL: https://www.researchgate.net/publication/357401487_Kanning_U_P_2021_Inventar_zur_Erfassung_der_Arbeitszufriedenheit_IAZ_Hogrefe [accessed 2026-02-11]
19. Gunzelmann T, Brähler E. C. Angermeyer, R. Kilian & H. Matschinger: Deutschsprachige Version der WHO Instrumente zur Erfassung von Lebensqualität WHOQOL-100 und WHOQOL-BREFM. Z Med Psychol 2002 Feb;11(1):44-48. [doi: 10.3233/ZMP-2002-11_1_07]
20. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health 2018;6:149. [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
21. Clark LA, Watson D. Constructing validity: new developments in creating objective measuring instruments. Psychol Assess 2019 Dec;31(12):1412-1427. [doi: 10.1037/pas0000626] [Medline: 30896212]
22. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016 Jun;15(2):155-163. [doi: 10.1016/j.jcm.2016.02.012] [Medline: 27330520]
23. Bühner M. Einführung in Die Test- Und Fragebogenkonstruktion: Pearson; 2010. URL: https://elibrary.pearson.de/book/99.150005/9783863265700 [accessed 2026-02-11]

Abbreviations

CFA: confirmatory factor analysis
SyMOA: Systematic Multidimensional Organizational Assessment
SyMOA-Q: Systematic Multidimensional Organizational Assessment–Questionnaire

Edited by J Sarvestan; submitted 04.Jun.2025; peer-reviewed by JK Punzalan, M Aqeel; revised version received 19.Jan.2026; accepted 19.Jan.2026; published 02.Mar.2026.

Please cite as:
Müller AM, Sachs S, Rieder Y, Lahmann C
The Systematic Psychodynamic Multidimensional Organizational Assessment Questionnaire: Protocol for a Development and Validation Study
JMIR Res Protoc 2026;15:e78530
URL: https://www.researchprotocols.org/2026/1/e78530
doi:10.2196/78530

XSL•FO
RenderX

© Anne-Maria Müller, Sophia Sachs, Yannik Rieder, Claas Lahmann. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Integration of the Integrate, Design, Assess, and Share Framework in Developing the Environmental Health Literacy Toolkit Paraben-Free & Me: Protocol for a Randomized Controlled Trial

Graziella De Michino[1], MHSc; Satveer Dhillon[2], MSc; Caroline Barakat[1], MES, PhD

[1]Faculty of Health Sciences, Ontario Tech University, 2000 Simcoe Street North, Oshawa, ON, Canada

[2]Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON, Canada

Corresponding Author:
Graziella De Michino, MHSc
Faculty of Health Sciences, Ontario Tech University, 2000 Simcoe Street North, Oshawa, ON, Canada

Abstract

Background: Endocrine-disrupting chemicals, such as parabens, are commonly found in personal care products (PCPs). Exposure to parabens is linked to several significant health risks, such as reproductive disorders, breast cancer, infertility, and hormone imbalances. Women are particularly vulnerable to these effects due to their higher use of PCPs containing parabens. Despite these risks, Canada lacks regulatory frameworks for the use of parabens in PCPs, relying instead on consumer awareness for reducing exposure. Previous studies have highlighted that many women remain unaware of parabens, exhibiting low risk perception and limited knowledge, which restricts behavior change toward safer choices.

Objective: To address this gap, this project developed the Paraben-Free & Me educational toolkit using the integrate, design, assess, and share framework.

Methods: Toolkit development methodology involved empathizing with target users, defining specific behaviors, grounding the data in the health belief model, ideating implementation strategies, prototyping, gathering user feedback, and building a minimum viable product.

Results: The Paraben-Free & Me toolkit includes multimedia resources such as blog posts, interactive quizzes, videos, podcasts, and forums aimed at increasing knowledge, risk perceptions, and health beliefs and facilitating paraben-free behaviors.

Conclusions: This toolkit shows potential to inform women about endocrine-disrupting chemicals, reduce exposures, and improve health outcomes.

International Registered Report Identifier (IRRID): RR1-10.2196/85474

(JMIR Res Protoc 2026;15:e85474) doi:10.2196/85474

KEYWORDS

educational toolkit; behavior; endocrine disruptors; knowledge translation; parabens

Introduction

Human exposure to endocrine-disrupting chemicals (EDCs) is widespread, as they can be found in electronics, pesticides, cosmetics, personal care products (PCPs), plastics, and other everyday products [1]. Parabens, for example, are a type of EDC commonly used as both synthetic preservatives and fragrance ingredients [2]. Women are at greater risk of exposure to parabens than men due to relatively higher PCP use, which has the potential to negatively impact their health and well-being [3-5]. For example, several studies have underscored the multifaceted impact of exposure to parabens and other EDCs, including the increased risk of pathophysiological reproductive conditions such as polycystic ovarian syndrome and endometriosis, breast cancer, infertility, and abnormal sex steroid hormone levels [2,3,6-9].

Compounding the negative impact of parabens is the lack of governmental restrictions and regulations for parabens and other EDCs in Canada. Unlike other regions, such as Europe, where there is a more robust approach to regulating EDCs [10], Canada uses a more risk-based strategy with no explicit regulation of EDCs [11]. Only products registered with Health Canada as natural health products have regulated limits on paraben concentrations [12]. By not implementing stricter regulations on parabens, the government of Canada puts the onus on the consumer to make health risk decisions.

Placing the burden on the consumer is concerning as, while there is increasing awareness of parabens and the associated negative health impacts, a large majority of women do not have a sufficient amount of knowledge about them [13-16]. To highlight this, a study by Trifunovski et al [16] conducted in Ontario, Canada, observed that women had low risk perceptions

XSL•FO
RenderX

of parabens in PCPs, with 28% of participants having not previously heard of parabens. Additionally, increased knowledge, risk perceptions, and health beliefs were associated with an increase in women's avoidance behavior [16]. Hence, if women do not possess information about parabens, it is difficult for them to implement changes to reduce their risk [13]. These findings highlight the importance of enhancing women's knowledge and access to information, health beliefs, and risk perceptions to encourage paraben-free behaviors.

Hence, one such intervention that can be implemented to increase the awareness of parabens and, consequently, improve health outcomes is creating an educational toolkit. Educational toolkits are defined as a package of resources for knowledge translation and facilitating behavior change [17]. Toolkits typically include resources such as informational materials, checklists, guidelines, and interactive activities designed to educate and empower users. Studies have highlighted how toolkits can be used for behavior change [17,18]. A review by Barac et al [18] reported that, of the 31 included studies that evaluated toolkit effectiveness, 21 were found to be satisfactory or useful or resulted in an intention to change practice. However, there is no empirical evidence evaluating the effectiveness of a toolkit focused on parabens when used by women. To fill this gap, this paper aims to describe the methods for creating and developing an evidence-informed educational toolkit that facilitates paraben-free behavior.

Methods

Overview

The educational toolkit development and methodologies were guided by the integrate, design, assess, and share (IDEAS) framework for developing digital health behavior change interventions [19]. The IDEAS framework is made up of 10 stages: empathize, specify, ground, ideate, prototype, gather, build, pilot, evaluate, and share (Figure 1 [19]). The IDEAS framework aims to produce a digital intervention that integrates behavioral theory and user feedback [19]. This framework was selected due to its focus on human-centered design and iterative process. Therefore, by grounding the development of the educational toolkit in this framework, we can ensure that it is evidence-based and its efficacy can be evaluated.

The development of this toolkit began in November 2023 and included synthesizing previous research from members of our team to help inform the project, along with partnering with the Women's Healthy Environments Network to aid with the development of this toolkit.

Figure 1. Stages of the integrate, design, assess, and share framework for developing digital interventions (reproduced from Mummah et al [19], published under Creative Commons Attribution 4.0 International License [20]). RCT: randomized controlled trial.

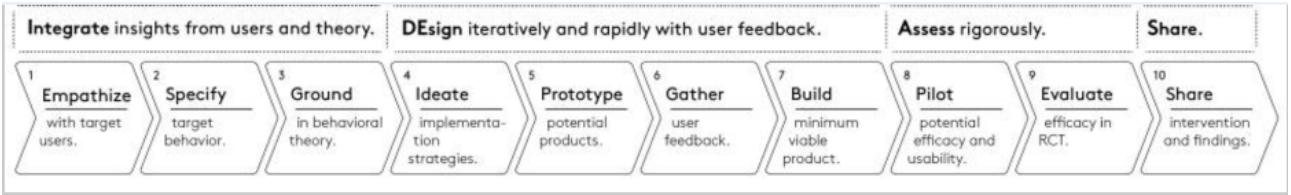

Step 1: Empathize With Target Users

The first step was to empathize with target users to understand their needs and preferences [19]. An earlier project by Trifunovski et al [16] identified 6 EDCs commonly found in personal care and household products: parabens, bisphenol A (BPA), phthalates, tetrachloroethylene, lead, and triclosan. The study aimed to understand women's current knowledge, risk perceptions, beliefs, and avoidance behavior related to these 6 harmful chemicals. A questionnaire was administered to 200 women to analyze associations between avoidance of parabens, BPA, phthalates, tetrachloroethylene, lead, and triclosan and knowledge, risk perceptions, and health beliefs [16]. Furthermore, 10 women living in Ontario, Canada, participated in virtual interviews to understand the factors that influence purchasing and avoidance behaviors.

Step 2: Specify the Target Behavior

The goal of step 2 was to define the target behavior to be specific and measurable [19]. Mean knowledge, health belief, risk perception, and avoidance behavior scores from a previous study by Trifunovski et al [16] were used to calculate the appropriate group mean difference between the control and intervention groups. Using a Likert scale, participants' knowledge and awareness, health beliefs, risk perceptions, and avoidance behaviors were evaluated through a composite score for each construct. For example, lower scores indicated lower knowledge and awareness. In this study, women scored an average of 21 (SD 6) points in the avoidance behavior section (scale from 6 to 30 in the original study). This step also included reviewing the current literature on educational toolkits to ensure that our toolkit was developed using best practices to elicit behavior change. Our research team conducted a comprehensive scoping review to answer the following research question: "What elements of educational toolkits contribute to improved risk perception and positive behavior change concerning the reduction of toxicant exposure from PCPs?" [21]. For this scoping review, we searched both academic studies and nonacademic literature, such as reviews of mobile apps that focus on PCPs, to ensure that we were able to analyze the literature surrounding educational toolkits that focus on increasing risk perception and promoting behavior change regarding harmful PCPs [21]. After the completion of this scoping review, and to further understand the aspects needed in an environmental health literacy toolkit, we also conducted a comprehensive systematic review to examine the current literature on environmental health literacy toolkits [22]. For this systematic review, we had two primary objectives:

1. To identify the characteristics of existing environmental health educational toolkits related to environmental health and behavior change

XSL•FO
RenderX

2. To identify the methods and data collection tools used to evaluate the effectiveness of these educational toolkits

For our systematic review, we searched peer-reviewed published articles and gray literature, which includes research and information produced outside of traditional academic publishing sources. This includes government reports, policy documents, health service guidelines, internal organizational reports, technical briefs, professional association publications, conference abstracts, dissertations, and nongovernmental organization or charity publications. This review included gray literature to capture practical and context-specific information that is often unavailable in academic journals but is directly relevant to the development and implementation of the toolkit.

Step 3: Ground in Behavioral Theory

The purpose of step 3 was to understand and use behavioral theories that could aid in the development of the educational toolkit [19]. The aim of this step was to ensure that each characteristic identified in steps 1 and 2 was grounded in theory. To achieve this, relevant behavior change theories were reviewed, and the health belief model (HBM) was selected as the primary framework due to its emphasis on individual perceptions, motivational determinants, and decision-making processes. The strategies and desired features found during steps 1 and 2 were evaluated for fit within the HBM, which focuses on explaining and predicting individual changes in health behavior.

Step 4: Ideate Implementation Strategies

The team synthesized data and information that were gathered in steps 1 to 3 and developed concrete strategies to ensure that the toolkit was effective. This included brainstorming potential online platforms to build the toolkit, possible modes of delivery, and important characteristics to incorporate. Additionally, we worked on identifying key themes and mapping content to behavior change objectives. In partnership with the Women's Healthy Environments Network, which is a nonprofit organization with the aim of educating the public and policymakers on environmental health, we ensured that our research was grounded in real-world insights.

Step 5: Prototype Potential Product

The iterative process continued as the research team discussed the feasibility of implementing specific design ideas that were generated during step 4. We ensured that the features of the toolkit aligned with the focus of the literature, the theory, and the capabilities of the web-based platform.

Step 6: Gather User Feedback on the Prototype

To gather user feedback, we gave our research team members (women aged 18-35 years) the opportunity to provide feedback on the first draft of our toolkit. The toolkit was emailed to the research team, and they were given 2 weeks to provide feedback.

Step 7: Build a Minimum Viable Product

The final step was to build a functional toolkit for pilot testing. Modifications were continuously reviewed to ensure accuracy and adherence to prior steps of the process. The functionality of the toolkit was tested by the research team, with the focus being on ensuring a smooth user experience, which included a simple consistent layout, interactivity for users, and bright and accessible colors.

Ethical Considerations

Ethics approval for this research was received from the Ontario Tech University Research Ethics Board on October 5, 2023 (17494). Informed consent to participate in this study was obtained from all participants. All participant-identifying details were omitted. No compensation was provided to the focus group, which consisted of members of the research team.

Results

Step 1: Empathize With Target Users

The findings suggest that women have low risk perceptions of EDCs, with 28% of participants in a previous study having not previously heard of parabens [16]. The authors of this study found positive associations between high knowledge, risk perceptions, and health beliefs and avoidance of lead, parabens, BPA, and phthalates. The qualitative paper (Trifunovski A et al, unpublished data, July 2024) concluded that social media, brand awareness, and price were all factors that impacted buying decisions for PCPs and household products. In terms of ingredients that may need to be avoided, participants avoided sulfates and fragrances. There was no mention of any other ingredients that may negatively impact health outcomes.

Step 2: Specify the Target Behavior

This was promoting women's paraben-free behaviors by increasing women's knowledge and access to information, health beliefs, and risk perceptions using an educational toolkit that has been informed by the literature. Our hypothesis was that women who had access to the Paraben-Free & Me educational toolkit would exhibit a 10% improvement in paraben-free behavior (3.5 points), knowledge and access to information (3.5 points), health beliefs (2.5 points), and risk perception (2.5 points) in group mean difference scores when compared with those who did not have access to the toolkit.

Our scoping review of academic sources and users' mobile app reviews revealed four themes of elements commonly observed in environmental health educational toolkits for facilitating behavior change: (1) toolkit accessibility and affordability, (2) simplicity of the presented information, (3) personalization of features, and (4) a clear focus on knowledge sharing [21]. This review highlighted not only the elements that are currently being featured in existing toolkits but also elements that users wish were included based on their reviews. Our systematic review highlighted the importance of toolkits being interactive, accessible, and personalized [22]. These factors were integrated into the development of our toolkit to ensure behavior change.

Step 3: Ground in Behavioral Theory

The themes emerging from previous studies and the literature were aligned with the HBM [23] (Figure 2 [24]). This model suggests that an individual's likelihood of adopting paraben-free behaviors is based on the individual's perceived severity of disease, perceived susceptibility to developing diseases related

XSL•FO
RenderX

to parabens, and perceived benefits of adopting the health behavior of using paraben-free PCPs [23].

This framework was selected as it emphasizes the psychological and social factors that shape health behaviors. It also provides guidance for designing educational health interventions as it highlights determinants of behavior change. Furthermore, research has underscored how there are positive associations between high knowledge, risk perceptions, and health beliefs and avoidance of lead, parabens, BPA, and phthalates [16].

Although a range of behavior change theories and frameworks, such as the capability, opportunity, and motivation–behavior framework; social cognitive theory; and the social ecological framework, offer valuable perspectives on the social and contextual determinants of behavior, this paper intentionally focuses on the HBM as the primary theoretical foundation for this project. This decision was guided by the aims of the toolkit, which center on addressing individual-level perceptions, decision-making processes, and motivational factors related to paraben-free behavior. Incorporating multiple frameworks risked diluting conceptual clarity and complicating the structure of the toolkit without clear added benefit for this specific application.

Figure 2. The health belief model of behavior change (reproduced from Fanwang0912 on Wikipedia [24] published under Creative Commons Attribution 4.0 International License [20]).

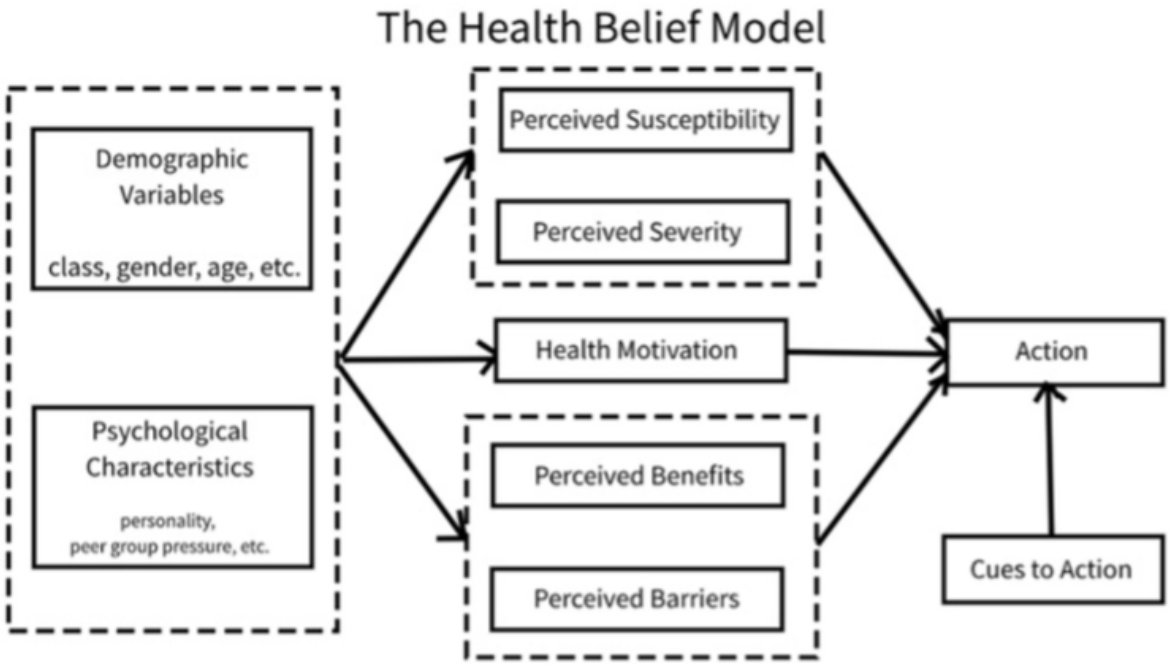

Steps 4 and 5: Ideate Implementation Strategies and Prototype Potential Product

The research team summarized the data from previous steps and developed solutions to improve the toolkit. On the basis of these findings, we identified Wix as a suitable platform that provides tools for creating websites and mobile apps. Through this platform, the Paraben-Free & Me educational toolkit was created in the format of a mobile app. This platform was selected because it could be accessed anywhere and contained blog posts, forums, question-and-answer sections, video and podcast recommendations, and quizzes.

Step 6: Gather User Feedback on the Prototype

Three users provided feedback to improve our toolkit. Feedback included design suggestions, such as including drop-down menus for blog posts and using smaller paragraphs. In terms of content, the users suggested that podcasts be included, as well as a larger number of videos. Suggestions were also provided on how often to update the toolkit, ranging from every 3 days to including a weekly summary message on Fridays.

Step 7: Build a Minimum Viable Product

Through this design process, we built the Paraben-Free & Me toolkit, which included synthesizing previous literature, conducted by members of our research team, and sharing the first version of the toolkit with our research team (Table 1). The final version of the toolkit provides individuals with an array of resources to support learning. These resources include blog posts, quizzes, videos, and podcasts. We ensured that the posts had simple and consistent messaging, features that were highlighted in our systematic review, and incorporated constructs from the HBM (Table 2 and Figure 3A). To ensure that users were able to comprehend the material, we created built-in quizzes for users to test their knowledge on parabens and see what topics they may need to do more reading on (Figure 3B). Moreover, to build a sense of community, we also included a forum section for users to share their knowledge and provide them with the opportunity to recommend other resources or products (Figure 3C). Furthermore, additional resources, such as YouTube videos, documentaries, and podcasts, were included to support personalized learning (Figure 3D).

XSL•FO
RenderX

Table . Identified characteristics, supportive feedback, and toolkit integration.

Theme	Feedback	Toolkit feature
Simple, consistent, and frequent messaging	Smaller paragraphs, buzzwords, and summarizing blog posts into short messages	Blog posts and weekly messages
Accessibility everywhere	—[a]	Mobile app
Health- and appearance-based motivators	—	Blog posts
Interactivity	Include checkpoints to interact with users	Forum page
Multimedia	Include audio and visual components	Blog posts
Personalization	Include audio and visual components and include additional resources to support different learning styles	YouTube videos, documentaries, and podcast recommendations
Incentives	—	Weekly quizzes

[a]Theme not discussed in the pilot-testing focus group.

Table . Paraben-Free & Me intervention design.

Focus of instruction	Learning goal	HBM[a] construct
"Parabens 101: What They Are and Why They're Used in Cosmetics?"	Knowledge	Perceived severity and susceptibility
"The Truth About Parabens: What's the Worst They Can Do?"	Risk perceptions	Perceived severity and susceptibility
"Can Switching to Paraben-Free Personal Care Products Improve Your Health?"	Health beliefs	Perceived benefits
"Are You Sure of What You're Putting on Your Skin? Learn How to Check the Ingredients of Your Personal Care Products"	Knowledge	Perceived barriers
"How to Decode and Understand Product Labels and Websites for Healthier Choices"	Knowledge	Perceived barriers
"How to Identify Paraben-Free Personal Care Products and Make the Switch: Tips and Tricks for Healthier Shopping"	Paraben-free PCP[b] behavior	Perceived barriers

[a]HBM: health belief model.

[b]PCP: personal care product.

XSL•FO
RenderX

Figure 3. (A) Example blog post featured in the Paraben-Free & Me toolkit to educate users on parabens, their health impacts, and how to avoid them. (B) Example of the weekly quizzes shared with users to test their knowledge and understanding of the blog post material. (C) Example forum page for users to share their knowledge and product recommendations. (D) One of the additional resources that users can access to support their learning.

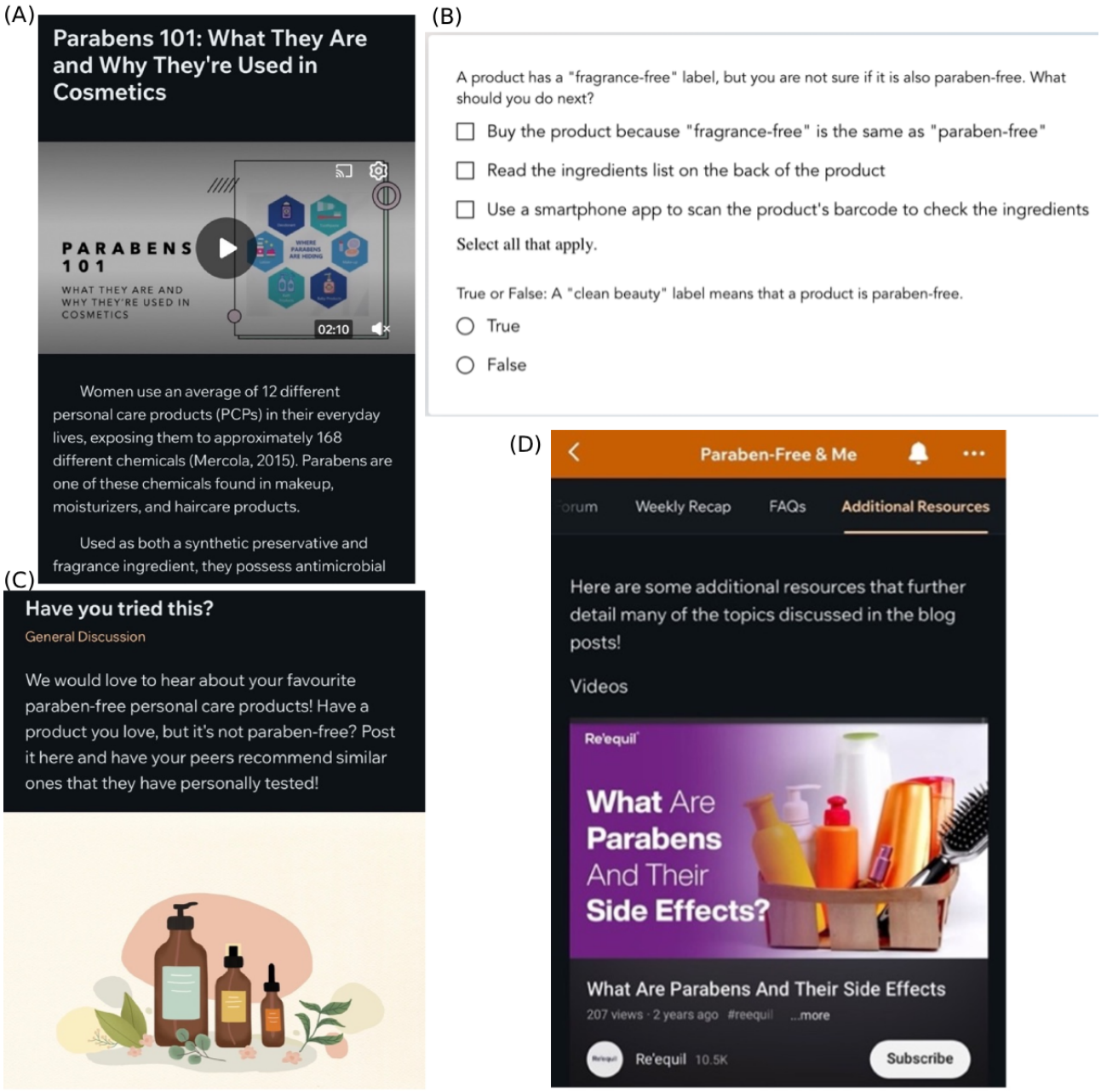

Discussion

Principal Findings

Overall, the main purpose of this manuscript was to describe a case study of the application of the IDEAS framework to guide the process of developing a new toolkit designed to promote paraben-free behavior and increase women's knowledge of parabens. Using the IDEAS framework allowed us to use an iterative design process that ensured that target users were able to provide feedback throughout the process. This ensured that the toolkit would work for the target population. The results from steps 1 and 6 provided important insights into the needs and preferences of the target population. These insights allowed the research team to make corrections and changes to the interactive toolkit during the design phase. In particular, gathering user feedback on the prototype (step 6) was extremely valuable as we were able to make modifications to ensure that the minimum viable product provided the target population with an array of resources.

Comparison With Prior Work

Through the use of the IDEAS framework, we designed a toolkit that includes several desired features that are consistent with previous research focusing on toolkits. For example, in the focus group conducted to pilot-test the app, the use of videos and audio recordings was suggested to supplement the material presented in the blog posts to provide users with multimedia tools that may assist them in personalizing their learning. This suggestion is consistent with previous literature in which multimedia approaches and personalization of learning were used to increase knowledge and facilitate behavior change

XSL•FO
RenderX

[21,22,25,26]. Additionally, the focus group highlighted the need for smaller paragraphs and summary messages at the end of the week. Simple, consistent, and frequent messaging is a common theme among educational toolkits in the literature [21,22,27].

Future Directions and Implications

Following the IDEAS framework, the next step for this project will be assessing the effectiveness of the Paraben-Free & Me toolkit using a randomized controlled approach. This includes piloting the toolkit for potential efficacy and usability (step 8), evaluating the efficacy through a randomized controlled trial (step 9), and sharing the intervention and findings (step 10). A single-site study at Ontario Tech University will be conducted to measure the effectiveness of the toolkit among female students between the ages of 18 and 35 years, with half of the participants in the control group and half in the intervention group. The primary outcome of assessment will be paraben-free behavior, and secondary outcomes will include knowledge and access to information, health beliefs, and risk perceptions. Both groups will complete the questionnaire that assesses these outcomes at baseline and after the intervention period. Group mean differences and 95% CIs will be analyzed to determine whether there were differences between baseline and follow-up scores. Future iterations of the toolkit may include developing simple games, replacing weekly quizzes with daily morning quotes, and potentially adding incentives through gift vouchers or promotional codes for paraben-free products.

Limitations

Despite the strengths of the project, there are several weaknesses as well. First, we only had 3 users, all of whom were highly educated, provide feedback on our prototype. If we had had a larger amount of feedback from a more diverse group of women (such as women who are immigrants or not proficient in using technology), our toolkit might have improved as we would have been able to incorporate feedback from a broader range of target users. Finally, pilot testing was not completed for usability and satisfaction. Conducting a small-scale evaluation of the potential efficacy and user satisfaction would have assisted in further refining the toolkit.

Conclusions

This manuscript described the use of the IDEAS framework to guide the development of the Paraben-Free & Me toolkit, an educational resource aimed at increasing women's knowledge and paraben-free behaviors. A scoping and systematic review was used to ensure that the appropriate characteristics were included, with feedback from the research team used to further support these findings and ensure their proper use within the development of the toolkit. The IDEAS framework provided a method for focusing on a human-centered and iterative development process to ensure that the toolkit was evidence-based and its efficacy could be evaluated. The Paraben-Free & Me toolkit has the potential to help overcome barriers faced by women from diverse backgrounds as it clearly presents evidence on the harms associated with parabens and provides practical guidance for identifying them in PCPs. By minimizing paraben exposure, the toolkit promotes better health outcomes and improved quality of life.

Acknowledgments

The authors would like to thank the Women's Healthy Environments Network for their ongoing consultations throughout this project.

Funding

This study was funded by the Social Sciences and Humanities Research Council Small Grants Program 2023 (Ignite Stream).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

CB conceptualized and supervised the study. GDM researched the literature and conceived the study. GDM and SD conducted result extraction and analysis and wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version.

Conflicts of Interest

None declared.

References

1. Hendryx M, Luo J. Association between exposure to parabens and total mortality in US adults. Environ Res 2022 Apr 1;205:112415. [doi: 10.1016/j.envres.2021.112415] [Medline: 34838567]
2. Chatterjee S, Adhikary S, Bhattacharya S, et al. Parabens as the double-edged sword: understanding the benefits and potential health risks. Sci Total Environ 2024 Dec 1;954:176547. [doi: 10.1016/j.scitotenv.2024.176547] [Medline: 39357765]

XSL•FO
RenderX

3. Hassan S, Thacharodi A, Priya A, et al. Endocrine disruptors: unravelling the link between chemical exposure and Women's reproductive health. Environ Res 2024 Jan 15;241:117385. [doi: 10.1016/j.envres.2023.117385] [Medline: 37838203]
4. Tricotteaux-Zarqaoui S, Lahimer M, Abou Diwan M, et al. Endocrine disruptor chemicals exposure and female fertility declining: from pathophysiology to epigenetic risks. Front Public Health 2024;12:1466967. [doi: 10.3389/fpubh.2024.1466967] [Medline: 39735741]
5. Swei H, Swanson S, Persellin K, Lacey A. Survey finds use of personal care products up since 2004 – what that means for your health. Environmental Working Group. 2023. URL: https://www.ewg.org/research/survey-finds-use-personal-care-products-2004-what-means-your-health [accessed 2025-05-25]
6. Kahn LG, Philippat C, Nakayama SF, Slama R, Trasande L. Endocrine-disrupting chemicals: implications for human health. Lancet Diabetes Endocrinol 2020 Aug;8(8):703-718. [doi: 10.1016/S2213-8587(20)30129-7] [Medline: 32707118]
7. Calaf GM, Ponce-Cusi R, Aguayo F, Muñoz JP, Bleak TC. Endocrine disruptors from the environment affecting breast cancer. Oncol Lett 2020 Jul;20(1):19-32. [doi: 10.3892/ol.2020.11566] [Medline: 32565930]
8. Liu H, Sun Y, Ran L, et al. Endocrine-disrupting chemicals and breast cancer: a meta-analysis. Front Oncol 2023;13. [doi: 10.3389/fonc.2023.1282651]
9. Patel S, Zhou C, Rattan S, Flaws JA. Effects of endocrine-disrupting chemicals on the ovary. Biol Reprod 2015 Jul;93(1):20. [doi: 10.1095/biolreprod.115.130336] [Medline: 26063868]
10. Kassotis CD, Vandenberg LN, Demeneix BA, Porta M, Slama R, Trasande L. Endocrine-disrupting chemicals: economic, regulatory, and policy implications. Lancet Diabetes Endocrinol 2020 Aug;8(8):719-730. [doi: 10.1016/S2213-8587(20)30128-5] [Medline: 32707119]
11. Malits J, Naidu M, Trasande L. Exposure to endocrine disrupting chemicals in Canada: population-based estimates of disease burden and economic costs. Toxics 2022 Mar 19;10(3):146. [doi: 10.3390/toxics10030146] [Medline: 35324771]
12. Risk management scope for parabens group. Health Canada. 2020. URL: https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/risk-management-scope-for-parabens-group-methylparaben-propylparaben-butylparaben-iso-butylparaben.html [accessed 2025-01-10]
13. Kelly M, Connolly L, Dean M. Public awareness and risk perceptions of endocrine disrupting chemicals: a qualitative study. Int J Environ Res Public Health 2020 Oct 23;17(21):7778. [doi: 10.3390/ijerph17217778] [Medline: 33114266]
14. Okman E, Yalçın SS. Awareness and knowledge of endocrine-disrupting chemicals among pregnant women and new mothers: a cross-sectional survey study. Toxics 2024 Dec 8;12(12):890. [doi: 10.3390/toxics12120890] [Medline: 39771105]
15. Rouillon S, Deshayes-Morgand C, Enjalbert L, et al. Endocrine disruptors and pregnancy: knowledge, attitudes and prevention behaviors of French women. Int J Environ Res Public Health 2017 Sep 6;14(9):1021. [doi: 10.3390/ijerph14091021] [Medline: 28878198]
16. Trifunovski A, Rotondi NK, Abbass-Dick J, Barakat C. Analysis of women's knowledge, health risk perceptions, beliefs and avoidance behaviour in relation to endocrine-disrupting chemicals in personal care and household products. Toxics 2025 May 21;13(5):414. [doi: 10.3390/toxics13050414] [Medline: 40423493]
17. Yamada J, Shorkey A, Barwick M, Widger K, Stevens BJ. The effectiveness of toolkits as knowledge translation strategies for integrating evidence into clinical care: a systematic review. BMJ Open 2015 Apr 13;5(4):e006808. [doi: 10.1136/bmjopen-2014-006808] [Medline: 25869686]
18. Barac R, Stein S, Bruce B, Barwick M. Scoping review of toolkits as a knowledge translation strategy in health. BMC Med Inform Decis Mak 2014 Dec 24;14:121. [doi: 10.1186/s12911-014-0121-7] [Medline: 25539950]
19. Mummah SA, Robinson TN, King AC, Gardner CD, Sutton S. IDEAS (integrate, design, assess, and share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317. [doi: 10.2196/jmir.5927] [Medline: 27986647]
20. Attribution 4.0 International (CC BY 4.0). Creative Commons. URL: Https://creativecommons.org/licenses/by/4.0/ [accessed 2026-2-2]
21. Diyab R, De Michino G, Yousufzai SJ, Barakat C. Risk perception of personal care products: a scoping review of knowledge translation strategies in environmental health. J Public Health Res 2025 Oct;14(4):22799036251395252. [doi: 10.1177/22799036251395252] [Medline: 41262156]
22. De Michino G, Diyab R, Barakat C. A systematic review of environmental health literacy toolkits: examining key components and data collection methods for effectiveness. Public Health Toxicol 2025;5(3):1-23. [doi: 10.18332/pht/210257]
23. Cummings KM, Jette AM, Rosenstock IM. Construct validation of the health belief model. Health Educ Monogr 1978;6(4):394-405. [doi: 10.1177/109019817800600406] [Medline: 299611]
24. Health belief model. Wikipedia. URL: https://en.wikipedia.org/wiki/Health_belief_model [accessed 2025-05-16]
25. Gosak L, Štiglic G, Budler LC, et al. Digital tools in behavior change support education in health and other students: a systematic review. Healthcare (Basel) 2022;10(1):1. [doi: 10.3390/healthcare10010001]
26. Hempel S, O'Hanlon C, Lim YW, Danz M, Larkin J, Rubenstein L. Spread tools: a systematic review of components, uptake, and effectiveness of quality improvement toolkits. Implement Sci 2019 Aug 19;14(1):83. [doi: 10.1186/s13012-019-0929-8] [Medline: 31426825]

XSL•FO
RenderX

27. Evans PH, Greaves C, Winder R, Fearn-Smith J, Campbell JL. Development of an educational “toolkit” for health professionals and their patients with prediabetes: the WAKEUP study (ways of addressing knowledge education and understanding in pre-diabetes). Diabet Med 2007 Jul;24(7):770-777. [doi: 10.1111/j.1464-5491.2007.02130.x] [Medline: 17403125]

Abbreviations

BPA: bisphenol A
EDC: endocrine-disrupting chemical
HBM: health belief model
IDEAS: integrate, design, assess, and share
PCP: personal care product

Edited by E Andrikopoulou; submitted 08.Oct.2025; peer-reviewed by K Favela, W Santosh; revised version received 11.Dec.2025; accepted 30.Dec.2025; published 13.Feb.2026.

Please cite as:
De Michino G, Dhillon S, Barakat C
Integration of the Integrate, Design, Assess, and Share Framework in Developing the Environmental Health Literacy Toolkit Paraben-Free & Me: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e85474
URL: https://www.researchprotocols.org/2026/1/e85474
doi:10.2196/85474

© Graziella De Michino, Satveer Dhillon, Caroline Barakat. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Developing and Validating Measures of Structural Ableism to Improve Health Outcomes for the Disability Community: Protocol for a Mixed Methods Study

Rupa S Valdez[1], PhD; Bonnielin K Swenor[2], PhD; Franz Castro[2], MD; Caroline Cerilli[2], BA; Bailey A Middleton[3], BA; Kara S Fitzgibbon[4], PhD; Ashley Shew[5], PhD; Varshini Varadaraj[2], MD; Hari Srinivasan[6], BA; Diana Cejas[7], MD

[1]Department of Systems and Information Engineering, Public Health Sciences, University of Virginia, Charlottesville, VA, United States

[2]Disability Health Research Center, Johns Hopkins University, Baltimore, MD, United States

[3]School of Nursing, Columbia University Irving Medical Center, New York, NY, United States

[4]Center for Survey Research, University of Virginia, Charlottesville, VA, United States

[5]College of Liberal Arts and Human Sciences, Virginia Tech, Blacksburg, VA, United States

[6]Brain Institute, Vanderbilt University, Nashville, TN, United States

[7]Department of Neurology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Rupa S Valdez, PhD
Department of Systems and Information Engineering, Public Health Sciences
University of Virginia
102A Olsson Hall
450 Whitehead Rd
Charlottesville, VA, 22903
United States
Phone: 1 4349822510
Email: rsv9d@virginia.edu

Abstract

Background: Structural ableism, defined as the processes, policies, and institutions that privilege able-bodied people over disabled people, is a root cause of health inequalities faced by the disability community. Unlike other forms of structural oppression, there are currently no adequate measurements for structural ableism and its impacts. Therefore, a necessary first step to addressing health inequities that impact the disability community is to create validated measures of structural ableism.

Objective: This paper outlines the methods of an ongoing project that aims to develop and validate measures of structural ableism. The resulting measures will facilitate the identification of relationships between structural ableism and health outcomes at both an individual and community level.

Methods: This project will take place across 3 phases. In Phase I, we will characterize the multiple factors that comprise the construct of structural ableism. We will begin by analyzing texts that discuss historical events, ideologies, and the lived experiences of disabled people to inform our understanding of contemporary dimensions of structural ableism. Simultaneously, key informant interviews with advocates and activists from the disability rights and disability justice movements will be conducted to further characterize the dimensions of structural ableism. In Phase II, the findings of Phase I will be used to inform the creation of an individual-level measure of structural ableism. Phase III will result in community-level measures of structural ableism, which will be developed using community-engagement studios with members of the disability community and publicly available datasets. This phase will build on the findings of the first 2 phases. Our methods purposefully include disabled people across all phases of this work, with a focus on maximizing the diversity of disability perspectives by including people across disability types and intersecting identities (eg, race and ethnicity, gender identity, geographic location, and other identities and demographics). Most importantly, our approach is deeply community-informed, drawing on multiple community partnerships from local and national organizations, a diverse advisory committee of disabled activists, advocates, and scholars, as well as researchers with expertise in developing measures of structural oppression, such as structural racism.

XSL•FO
RenderX

Results: This project was funded in August of 2024. As of October 2025, our team has read more than 50 texts as part of our historical and policy analysis of the factors that characterize structural ableism. We plan to complete our characterization of structural ableism in the spring of 2026, with individual-level measures of structural ableism being developed by the Winter of 2028 and community-level measures created by the Winter of 2029.

Conclusions: The measures developed by this work will lay the foundation for identifying and evaluating novel interventions aimed at dismantling structural ableism, which should be cocreated with the disability community.

International Registered Report Identifier (IRRID): DERR1-10.2196/86976

(JMIR Res Protoc 2026;15:e86976) doi:10.2196/86976

KEYWORDS

structural ableism; ableism; disability; health equity; health disparities; disability health; minority health

Introduction

The disability community was designated as a health disparity population by the National Institute on Minority Health and Health Disparities in 2023 [1], a recognition that disabled people face health disparities such as higher rates of mortality [2] and cancer [3], more severe outcomes of infectious diseases such as COVID-19 [4], poorer mental health [5], 4 times higher rates of victimization [6], and greater poverty [7,8] compared to nondisabled people. These disparities compound for disabled people who are part of other health disparity populations [7-10].

While research has examined how ableism within health care systems and from health professionals can lead to health disparities [11], there is a lack of knowledge on the greater impact of structural ableism on health outcomes. Structural ableism, which can be defined as "a complex system of hierarchical and discriminatory processes, policies, and institutions that privilege and prefer able-bodied people, fail to represent or meaningfully include disabled persons' voices, and are grounded in a network of ableist beliefs and practices that maintain and reproduce unequal outcomes for disabled people and benefit able-bodied people" [12], is a force embedded in our society that produces poor health outcomes for disabled people that often go unnoticed and understudied.

With a PubMed [13-25] search of "structural ableism" in April 2025 yielding only 13 papers, the conceptual clarity of structural ableism within the health sciences remains in its infancy, and measures remain inadequate. Achieving conceptual clarity is essential for identifying the factors that comprise structural ableism and is a foundational step for understanding and addressing the ways in which structural ableism shapes health outcomes for disabled people [26]. Validated measures are essential for analyzing and addressing the mechanisms and pathways linking structural ableism to health outcomes for disabled people [14,27].

An important aspect of this work is recognizing that there are certain disability types that historically have been excluded from conversations about disability. Although, for instance, the definition above is among the most current, we acknowledge that some of the terminology (such as "able-bodied") used to refer to people without disabilities may imply an overemphasis on physical disabilities. With this project, we will take steps to ensure that those with intellectual or developmental disabilities (IDD) and those with other nonphysical forms of disability are also included. As part of this effort, we have developed a working everyday language (the term preferred by members of the IDD community currently on our advisory committee, instead of "plain language") definition of structural ableism (Textbox 1), and note this definition is expected to evolve.

The definition is accompanied by 2 graphics (Figures 1 and 2), which both represent our early understanding of structural ableism, as well as an early attempt to visualize structural ableism in a way that is accessible to those who use everyday language.

XSL•FO
RenderX

Textbox 1. Working everyday language definition of structural ableism and examples.

Definition: Structural ableism means that the world is not designed to work for disabled people. It means that disabled people face more challenges than people without disabilities. For example, many places could be more accessible for people with disabilities, and many laws could better support disabled people. Disabled people are usually not asked about their opinions when decisions that matter to them are made. This means things stay unfair for them, and it is harder for them to live the life they want. Here are some examples:

- Workplace:

  Some jobs have policies that require many hours of work or don't let employees work from home. This makes it hard for employees with energy-related disabilities to succeed.

- Education:

  Some schools lack the resources needed to provide students with audiobooks or handouts in large print. This makes it harder for some students with vision-related disabilities to learn.

- Public spaces:

  Some public places, like restaurants and museums, were not built with a consideration for accessibility, so they lack ramps or elevators. This makes it difficult for people with mobility-related disabilities to enter them.

- Recreational spaces:

  Not all recreational spaces have adequate social support because many people are not taught the importance of community support for the success and well-being of people with intellectual disabilities. This makes it harder for individuals with intellectual disabilities to participate in activities like sports.

- Health care:

  Most speech-language pathologists have their job because they want to help people. Sometimes, without realizing it, they help someone communicate in the way they think is best instead of in a way that is most helpful for that person, because they are not trained on how to best center the patient's thoughts and opinions about their own care.

- Government policy:

  Medicaid was not made to consider all the needs of disabled people, and sets a very low savings limit. If disabled people save more than this, they lose coverage. This rule keeps people poor and stops them from saving for emergencies.

XSL•FO
RenderX

Figure 1. Early Visualization of Structural Ableism.

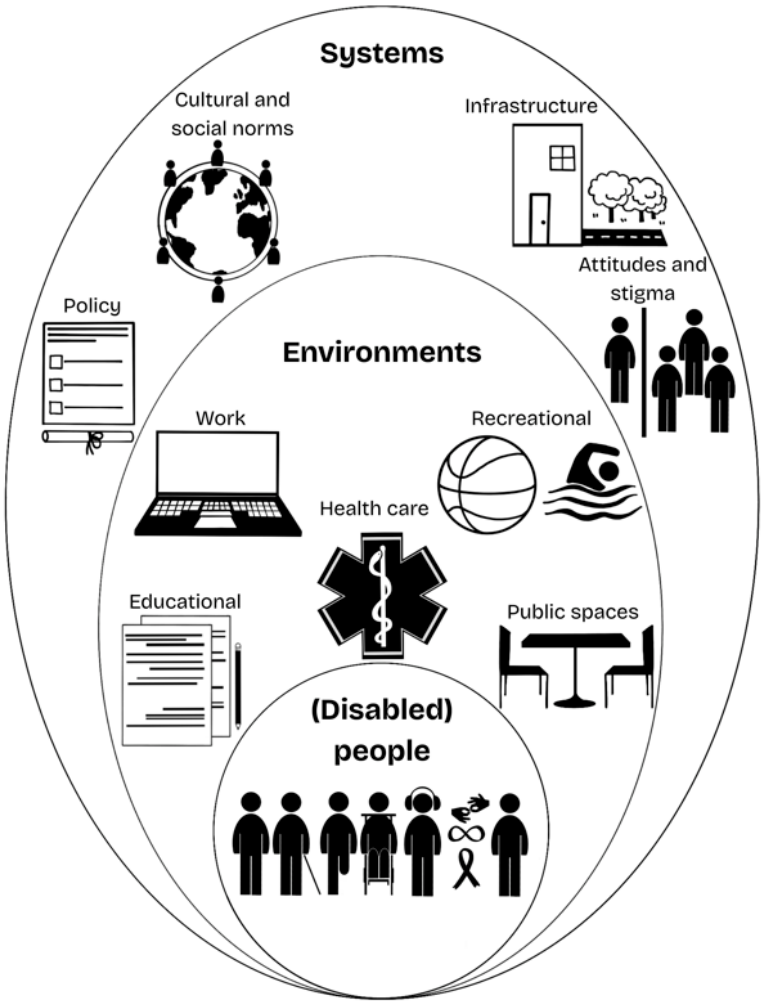

XSL•FO
RenderX

Figure 2. Early Visualization of Structural Ableism with Supports that Mitigate It.

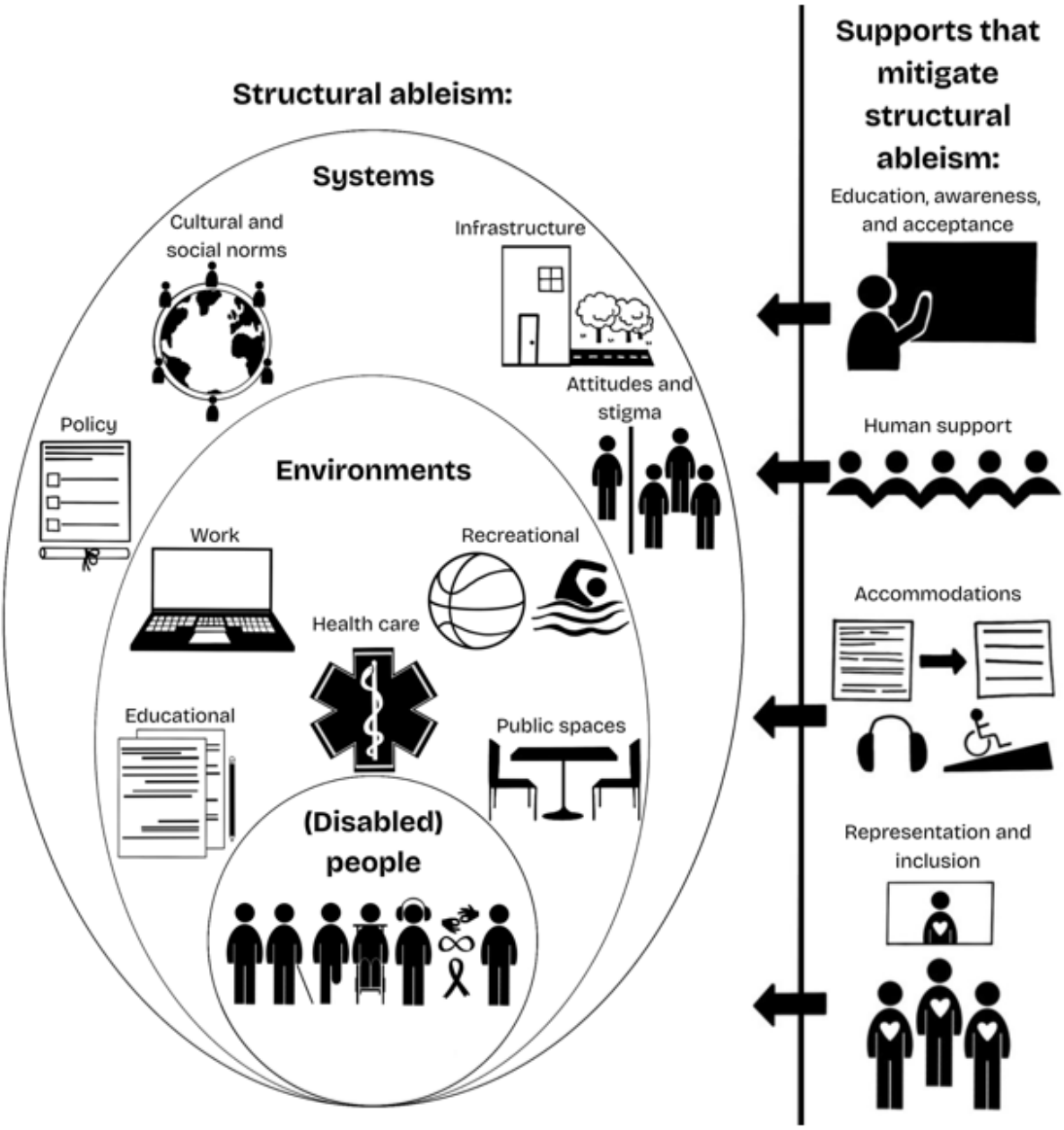

Learning from work about other forms of structural oppression [27-43], we understand that structural oppression is not limited to individual biases, and over time takes increasingly covert forms in social institutions while maintaining and reifying power imbalances [14,44]. For example, African American students with emotional, developmental, or learning disabilities are more likely than white students with the same disabilities to be suspended from school [45]. Such disparities deepen when considering variations in support needs and disability. This is demonstrated by people with autism spectrum disorder being less likely to receive vocational support than people with other disabilities [46]. We also note that, in this paper, we will be using the term oppression to express the experience of widespread unfair treatment, as it is the term used in other academic spaces. However, we acknowledge that, as identity and preference vary, this term may not resonate with all those in the disability community.

In addition to operating in conjunction with intersecting oppressed identities, structural ableism acts at both the individual and community levels. Examples at the individual level include: a patient who is deaf and not provided an American Sign Language (ASL) interpreter, a grandmother who uses a wheelchair and cannot visit her family because the building they live in does not have ramps, a college graduate with autism who cannot find a job due to employer biases, and a child with food allergies who struggles to find food options that are affordable and safe. The policies that uphold structural ableism create structural disparities that impact the disability community in all aspects of life: national shortages of ASL interpreters in hospital settings [47], higher rates of loneliness among older adults with disabilities than those without disabilities [48], higher unemployment rates among adults with autism [49], and an elevated percentage of children with food allergies who experience food insecurity [50]. Despite the knowledge that

XSL•FO
RenderX

individuals with disabilities experience disparities that impact their health, there is still a dearth of information about the extent of these impacts.

As defined by the Americans with Disabilities Act (ADA), disability can be described as the experience of "a person who has a physical or mental impairment that substantially limits one or more major life activities, a person who has a history or record of such an impairment, or a person who is perceived by others as having such an impairment" [51]. This definition encompasses a variety of experiences, including those with mental health–related, intellectual, developmental, cognitive, sensory, and physical disabilities. Disability can be acquired at any point during one's life, or be present since birth, can be both apparent or nonapparent to others [52], and varies depending on whether individuals are accommodated or not. Based on their environment and personal experiences, individuals may or may not identify as having a disability [53]. Structural ableism impacts all individuals with disabilities regardless of disability type or identity.

Measuring structural ableism is essential to creating interventions to slow and prevent it, ultimately improving health outcomes for disabled people by addressing the systems that create barriers to equitable health outcomes. Health inequities that impact disabled people are too often attributed to disabled people themselves and not the environment and contexts in which they live [54]. Yet, the experience of disability is multidimensional and contextual: it is social, cultural, and political, and is shaped by interactions with the environment [55-57]. Recent studies have shown the fallacy of attributing poor health outcomes solely to a disabled individual's underlying health condition, demonstrating instead that ableism still permeates society, including in education [58], policing [59], employment [60], health care [61], and many other domains. Structural ableism is embedded ableism that drives poor health outcomes for the disability community. Consequently, health outcomes for disabled people can only be improved through a structural approach. Therefore, an essential first step toward addressing the health inequities that impact disabled people is to create comprehensive measures of structural ableism [13,14]. Clear measurement of structural ableism transforms invisible problems into visible ones. Such measures allow us to identify where further work is needed and the degree to which interventions have resulted in successes or deepened divides. It is only when we have ways to measure structural ableism that research can consider and address its effects on health outcomes [40].

Measurement of structural ableism must focus on all relevant structures with which disabled people interact. If measures are only inclusive of some aspects of the broader experience of structural ableism (eg, those that focus only on employment or education), they will fail to capture the full system contributing to negative health outcomes for the disability community [62]. Efforts to measure the multitude of domains relevant to structural racism, structural sexism, structural ageism, and other forms of structural oppression have identified relevant structures (eg, policies, leadership demographics, and hierarchies) spanning housing, education, criminal justice, employment, health care, political participation, immigration, and others [28,41,63]. Many similar domains are likely to be relevant to structural ableism, but the ways in which these structures operate to create injustice for disabled people may be unique [14]. Moreover, additional domains such as dimensions of the built environment [14,20,64] (eg, curb cuts) and dimensions of communication structures, often not highlighted in other measures of structural oppression, may have particular relevance for measuring structural ableism.

In addition to considerations of multiple domains, there is also a need to consider multiple levels. As other measures of structural oppression, including Patricia Homan's work on structural sexism [40] and Adkins-Jackon et al [29] work on structural racism, account for both individual and community-level effects, measures of structural ableism must do the same [62]. Therefore, this work will focus on understanding and measuring the multidimensional and multilevel nature of structural ableism. This project will build on RSV's and BKS's *New England Journal of Medicine* paper, which outlined the need for measuring structural ableism [14], the team's lived experiences as disabled people, and previous research efforts related to barriers for the disability community, and established expertise in all proposed methods.

We describe our methodological approach to developing measures of structural ableism. Our work lays the foundation for creating intersectional measures of structural oppression that are inclusive of structural ableism. Moving toward a structural intersectionality approach is critical, as dismantling structural oppression for all people will require understanding the ways multiple forms of structural oppression, including structural ableism, interact to shape health outcomes for people belonging to multiple health disparity populations [60,65,66]. In the future, we aim to build off the work of this project in order to create intersectional approaches to addressing structural ableism. The goal is to promote better health outcomes for all disabled and oppressed people, rather than promote the reinforcement of normativity.

Methods

Overview and Objectives

This project has 3 phases. In the first phase of this work, we will use historical, policy, and qualitative approaches to characterize the dimensions of structural ableism. The second phase develops and validates a survey-based individual-level measure of structural ableism. The third phase develops a community-level measure of structural ableism using publicly available datasets. This study has been approved by the University of Virginia (UVA) Institutional Review Board (IRB) for the Social and Behavioral Sciences. For all phases of the project, informed consent will be obtained, and participants will be compensated for their participation.

Our approach recognizes that addressing structural ableism must center the disability community and involve people living with a wide range of disabilities [13,14]. Our work achieves this at 3 levels. First, we will engage the disability community through an advisory committee comprised of community members. Second, our research team is dominated by members of the disability community. Third, we will engage participants in our

XSL•FO
RenderX

research who are disabled. We will seek to have equal representation across all broad disability types: mobility and/or physical, cognitive, intellectual, and/or developmental, deaf and hard of hearing, communication disability, bodily difference, blind and visually impaired, mental health and psychological, and chronic illness [67,68]. More information on how we plan to engage with our advisory committee and research participants can be found in Table 1. We will prioritize accessibility in all aspects of this project. This includes ensuring all recruitment materials, data collection, tools, and deliverables are in plain language, 508 compliant, and Web Content Accessibility Guidelines (WCAG) 2.2 compliant [69-71]; 508 compliance indicates that all content complies with the accessibility standards for information and communication technology set by the United States Access Board, while WCAG 2.2 compliance indicates that materials comply with the international standard for accessibility of web content. We will additionally work with all participants to ensure accessibility needs are met. In order to meet the accessibility needs of each participant, interview questions are pasted into the chat on Zoom, which is particularly helpful for those who are deaf or hard of hearing or who have a disability related to language processing. A trained individual who is fluent in ASL leads interviews with participants who communicate with ASL. At times, there is also a team member speaking the interview questions so that closed captions are enabled and accessible during the interview. A trained individual with IDD leads interviews with participants who have IDD. Additionally, interview guides are provided to all participants ahead of time, and ample breaks are allowed throughout the process for those who need them. We anticipate that these methods will continue to be revised and expanded to meet the needs of our diverse participant population.

Table 1. Overview of disability community engagement.

Advisory committee	Research participants	Corresponding project phase
Advice on sourcing for historical and policy approach; review key informant interview questions	Informant interviews (n=30); targeted interviews and intake survey to select interviewees for diversity	Phase I
Mapping and review of the conceptual framework; review draft and validated individual-level measure of structural ableism	Semistructured cognitive interviews via Zoom (n=12); self-administered online survey (n=400+)	Phase II
Guide development of the conceptual framework; provide feedback on domains and indicators for community-level measures of structural ableism	Community engagement studio of disability advocates and activists (n=16); participating in the iterative design of a community-level measure of structural ableism	Phase III

Study Design

Phase I: Characterizing the Multiple Domains That Comprise the Construct of Structural Ableism

In Phase I, we aim to characterize structural ableism, offering a broad overview of the dimensions of structural ableism and helping us identify these elements within a societal context. At the conclusion of Phase I, we will have developed a final overarching codebook that represents the dimensions of structural ableism and potential links between dimensions of structural ableism and health outcomes.

To understand how structural ableism is experienced by disabled people, it is important to listen to what disabled people have to share, to give them epistemic authority. Too often, disabled voices and perspectives are not centered in research work that impacts disabled people. To guide our examination of contemporary domains of structural ableism, we analyzed a variety of life writings by disabled authors and editors, along with policy documents and sources on the social and historical context of disability. We began with a commitment to a minimum of 50 sources as a matter of feasibility, limiting sources mostly to the last 20 years, but we have gone beyond this quota. We selected our sources for a diversity of types of text, authorship, and disability. We aimed to read widely to consider how the type of disability, race, gender, sexuality, and other aspects, such as geography and class, influenced experiences of ableism.

This approach examines historical events, ideologies, and the lived experiences of disabled people to understand the production and development of disability as a category and structural ableism as an oppressive force. We adopt a dual approach combining this historical perspective [44] with a policy perspective [34,55]. Recognizing that structural ableism is intricately linked to legal frameworks, our analysis further extends to both disability laws and to broader policies shaping the lives of disabled people [72,73]. This approach entails examining how laws and policies are written and interpreted, influencing the understanding and processing of disability, and helps us identify the dimensions of structural ableism from a policy standpoint. This information will inform the development of a summary of the material, identifying emergent themes and categories related to the dimensions of structural ableism and their relationships to health.

A qualitative approach will then be used to further characterize the dimensions of structural ableism [74,75]. We will engage 30 advocates and activists (or more as needed to reach saturation) from the disability rights and disability justice movements in key informant interviews which will combine aspects of oral history interviews and interviews that rely on the critical incident technique [76-78], thus combining approaches that allow us to understand experiences across the life course and also understand particularly memorable experiences.

Recruitment for participation in key informant interviews will begin by electronic distribution [69] of a brief survey to

XSL•FO
RenderX

advocates. We will ask all who receive the survey to then send it on to other advocates and activists in the disability rights and disability justice movements with whom they are personally connected, and then ask those individuals to do the same. Our message accompanying this interest survey will contain the purpose of this study, a brief description of the purpose of the key informant interviews, eligibility criteria, and a note that not all who complete the survey will be invited to participate given our goal of including as diverse a group as possible in terms of disability type, advocacy type, gender identity, geographical location, race and ethnicity, and income status (maximum variance sampling). We will also share recruitment details via phone and in-person interactions to let advocates and activists in the disability community who may not be online know of the opportunity to participate. To be eligible, individuals must identify as an advocate or activist in the disability rights or disability justice movements, be 18 years of age or older, and live within the United States. They must further identify as disabled or have a close, personal relationship with a disabled individual (eg, caregiver or parent).

We expect each interview to last for approximately 90 minutes, with more time provided as needed. During the interviews, we will use a semistructured interview guide to explore how the individuals themselves and members of the disability community more broadly experience frustrations, barriers, and exclusion. The interview guides used in this study are available in Multimedia Appendices 1-5. We will begin by asking broad questions about these types of experiences across the life course, and then probe into experiences related to specific aspects of life, including education, employment, community spaces, transportation, criminal justice, and housing.

We will use an inductive approach to qualitative content analysis, in which themes and categories are inductively and iteratively derived from the data [79]. Our analysis will begin with investigators independently coding 3 transcripts from the qualitative interviews and then coming together to create an initial codebook through a consensus-building process. The remaining transcripts will then be divided among 4 research assistants for coding. All coding will take place within QSR NVivo.

At the completion of the historical and policy review and the key informant interviews, we will develop an overarching codebook that represents the dimensions of structural ableism and ways in which our 2 types of sources discuss links between dimensions of structural ableism and health outcomes. To establish rigor in our interpretive work, we will use source and analyst triangulation; peer debriefing; reflexive journaling; the presentation of quotes, both from texts and transcripts; and we will make data available upon request [80,81].

Phase II: Develop and Validate an Individual-Level Measure of Structural Ableism

In the second phase, we will qualitatively pretest and quantitatively assess the validity of an original survey instrument that will serve as an individual-level measure of structural ableism. This instrument will be grounded in Phase 1 findings and will provide a way to measure the relationship between individual experiences with structural ableism and health outcomes.

First, we will develop a detailed conceptual framework that identifies the various dimensions and corresponding specific items within each dimension of structural ableism that the survey questionnaire is intended to capture. All items will be developed in plain language. This model will also include the eligibility criteria and respondent characteristics to be measured, such as geographic location, age, gender, race and ethnicity, and socioeconomic background.

The research team will conduct a series of cognitive interviews to pretest the survey instrument with disabled individuals ahead of pilot survey administration [82-84]. This study proposes to conduct 12 interviews, or more as needed to reach saturation, to evaluate the instrument for clarity, accessibility, and face validity. Cognitive interviews are expected to be approximately 60 minutes. During the interview, the facilitator will verbally and/or visually review each question of the survey instrument. We will ask participants to articulate their interpretation of specific concepts and terms included in the survey, to rephrase the questions in their own words, to explain the thought processes they followed to come up with the answer that they did, and to explain their rationales for the responses they provided. Additionally, through the use of probing questions, the facilitator may ask the participant for additional clarification or explanation. Recruitment and sampling for the cognitive interview will parallel that described for the key informant interview in Phase I.

After the completion of each interview, the facilitator will synthesize the interview feedback into a detailed narrative summary report, identifying any questions or answers that posed challenges for the participant. Separately, a facilitator who did not participate in the interview will review the recording from the interview and prepare their own summary report of the feedback. At the conclusion of the interview data collection period, we will review all interview summaries and prepare a synthesized report of findings across interviews and the resulting recommendations for revisions to the survey instrument. The resulting report of recommendations will be shared with cognitive interview participants for member checking and will be revised in response to the feedback received. We will revise the survey questions and answer categories as needed based on the pretest interview feedback. We will then pilot test the revised instrument through online survey administration to determine usability.

Next, we will conduct pilot survey administration and validate measures through psychometric analyses. The pilot survey administration and recruitment will occur online and draw from multiple sampling strategies. The survey will be hosted by the UVA's Center for Survey Research on the University's license of Qualtrics. Currently, 15 question types within Qualtrics are WCAG 2.0 AA compliant.

All participants for the survey pilot will be adults living in the United States who have one or more disabilities. Eligibility will be based on the participant either self-identifying as having a disability and/or chronic illness or self-identifying as having a functional limitation. Screening questions for eligibility will

XSL•FO
RenderX

include the following: (1) Do you have a disability and/or chronic illness? (2) What types of disabilities do you have? Select all that apply. (3) Six questions from the American Community Survey, which inquire about difficulties across different domains of functioning (difficulties hearing, seeing, concentrating, or making decisions, walking or climbing stairs, dressing or bathing, and doing errands alone) [83]. An individual will be eligible if they self-select as identifying with at least one of these screening items. These screening measures include functional limitations as well as self-identifying as having a disability or chronic illness (and given the space to expand on that identification) in an effort to be comprehensive, recognizing that relying on just one or the other could result in a failure to capture the full range of respondents with relevant experiences. Eligible individuals will be 18 years of age or older. Given the specific eligibility criteria, nonprobability sampling methods are best suited to efficiently reach this population for the pilot. The self-administered survey is expected to take 20-30 minutes to complete, although we expect that time will vary across the disability community. The survey will be administered in English. We will offer survey support throughout data collection and provide both an email and phone number for participants to contact should they have questions or need assistance with or accommodations for completing the online questionnaire.

We will use two sources for survey participants: (1) an online access panel (Dynata [85], which includes over 70 million potential respondents and has a commitment to broad representation in their panelists across multiple identities including disability and chronic illness) and (2) indirect recruitment through the listservs and social media accounts of the project team and advisory committee's expansive network of disabled people, organizations and researchers. From the online panel, we will collect 400 responses. Due to the nature of the indirect recruitment through listservs and social media, it is unknown how many additional responses we may receive through these channels. The count of 400 achieved through the online panel will ensure we have the minimum number of completions needed to support the range of analyses planned to validate the instrument.

To assess the reliability and validity of the survey scale, we will conduct a series of psychometric analyses using pilot survey data, which will be prepared and analyzed in SPSS (IBM Corp). We will first calculate Cronbach to assess the reliability of the scale, specifically the internal consistency of the items measuring structural ableism. To investigate possible item reduction, we will run inter-item correlations and adjusted item-total correlations. Given that this is a newly developed scale, we will conduct exploratory factor analysis to assess the underlying factor structure within the data, followed by confirmatory factor analysis. To further assess construct validity, we will use correlation analysis, calculating the Pearson correlation coefficient (r) between the structural ableism measures and both theoretically related and unrelated measures to test for convergent validity and discriminant validity, respectively. As an additional assessment of construct validity, we will run a series of bivariate and multivariate regressions between the scale values and relevant health outcomes. If any of these analyses lead to substantial changes, we will administer a second, revised questionnaire, piloting the survey to a new, independent sample of 400 respondents from the online survey panel, controlling for the exclusion of any respondents who had participated in the prior survey. We will then rerun all analyses to test the psychometric properties of the revised instrument. These methods of psychometric analysis are in alignment with best practices for scale development [86].

Our survey will include 4 questions related to health outcomes: (1) having a usual health care provider, (2) not having any health insurance, (3) having an unmet health need due to cost, and (4) having a routine check-up in the past 12 months [87]. These outcome measures were selected because previous research has demonstrated health disparities between disabled and nondisabled people [88-90], and they are commonly used metrics across the public health and health care fields when creating medical surveys. To explore these relationships, we will run a series of nested model regressions of structural ableism across the 4 health outcomes of our individual-level measure of structural ableism. These analyses will test our hypothesis that higher levels of individual-level structural ableism are associated with greater barriers across these 4 indicators.

Phase II will result in a validated instrument to quantitatively measure individuals' experiences with structural ableism and will explore the relationship between this measure of structural ableism and health outcomes.

Phase III: Measure Structural Ableism at a Community Level Using Publicly Available Datasets

Phase III aims to create a community-level measure of structural ableism in order to explore its relationship with health outcomes.

We will use an iterative design approach to select potential indicators for the composite measure of structural ableism. This process will be informed by qualitative data collected from Phase I and feedback from our advisory committee and will help us outline content domains and potential indicators relevant to structural ableism at a community level (eg, food security, transportation, employment, etc). The resulting composite measure will be revised with advocates and activists in the disability rights and disability justice movements through 2 community engagement studios [91,92], an established methodology that facilitates bidirectional input between the research team and the disability community via virtual gatherings and facilitated discussion [93]. This approach ensures that the lived experiences of people with disabilities are prioritized in the selection of indicators and item development.

Community Engagement Studios

Sixteen advocates and activists in the disability rights and disability justice movements will participate in the community engagement studios. Eligible individuals will be 18 years of age or older, self-identify as an advocate or activist in the disability rights or disability justice movements, and live in the United States. We expect that about half of those engaged in the community engagement studios will have participated in the Phase I qualitative interviews.

XSL•FO
RenderX

Recruitment for participation in the community engagement studios will parallel the recruitment process described for the Phase I key informant interviews. We will host 2 community engagement studios: one at the start to gain further insight on the indicators of structural ableism identified in Phase I, and a second studio to include community feedback in the final selection of metrics. During each studio, we will pose 3 to 4 questions to obtain feedback on the range of dimensions of structural ableism represented, the particular datasets and questions to be used, and any additional overarching feedback. The community engagement studio sessions will last approximately 90 minutes, and material for discussion will be distributed beforehand in formats accessible to each participant.

After completion, the facilitator will synthesize the feedback into a detailed narrative summary report, identifying all points of feedback from the participants. Separately, another member of the research team who did not participate in the community engagement studio will listen to the audio recording and prepare their own summary report of the feedback. The consensus summary will be shared with community engagement studio participants for member checking and will be revised accordingly.

Development of Indicators

To develop indicators of structural ableism, we will use data primarily from the Behavioral Risk Factor Surveillance System, a state-based national survey of noninstitutionalized adults 18 years and older, to investigate the same health outcome measures stated in Phase II [87]. As a secondary step, we will explore linking indicators to data from the IPUMS Health Surveys, which are a comprehensive integration of health survey data from sources such as the National Health Interview Survey [94]. These datasets assess disability status using the American Community Survey Questions on Disability (ACS-6), assessing difficulties across different domains of functioning (difficulties hearing, seeing, concentrating or making decisions, walking or climbing stairs, dressing or bathing, and doing errands alone). This process will include developing a menu of indicators of structural ableism based on the themes emerging from Phase I and supplemented by input from our advisory committee and a review of pertinent literature. We will then examine existing variables available in the Behavioral Risk Factor Surveillance System, and available indicators will move forward to the next steps of the process.

Since our analysis is mostly focused on capturing gaps between disabled and nondisabled people, we propose to use the approach used by Dougherty et al [95]. A composite measure of structural ableism will be estimated using confirmatory factor analysis, which minimizes the effect of random measurement error by accounting for the variability across prespecified indicators in a summary factor. Unidimensional factor models will be produced using the composite measure as the single latent variable with a mean of 0 and SD of 1. Robust maximum likelihood estimation will be used to produce unbiased estimates, and model fit will be evaluated using a Tucker-Lewis index. Candidate models would at least contain one indicator per content domain, and a final model will be chosen, assessing the trade-off between model complexity and goodness of fit [96].

For each indicator, prevalence ratios will be computed, log-transformed, and normalized using robust min-max scaling relying on prespecified percentiles (2.5th and 97.5th). Values outside of that percentile range will be winsorized to those bounds, and the resulting bounded values will be linearly transformed to a 0-100 scale.

Phase III will result in the development of novel indicators of structural ableism identified by disabled people available from nationally representative surveys, and the development of a composite measure of structural ableism based on these measures.

Ethical Considerations

All aspects of this study will be conducted in a manner that is HIPAA (Health Insurance Portability and Accountability Act)-compliant and accountable under the responsible conduct of research protocols active at the UVA. A Human Subjects application has been approved by UVA's IRB for the Social and Behavioral Sciences (protocol number 6677), which operates in accordance with the Belmont Report, the Nuremberg Code, and the Declaration of Helsinki. The UVA IRB for the Social and Behavioral Sciences will serve as the single IRB for this study. This research falls under the classification of nonexempt human participants' research. The privacy of all participants will be maintained through the deidentification of all identifiable data, which will be stored on a secure server. Participant data will be linked to a participant ID number. All participants will receive a consent document in plain language, which will be kept on a secure server at UVA after being signed. For engagement in key informant interviews and community engagement studios, all participants will receive US $150 in compensation. For participation in the cognitive interviews, all participants will receive US $100 in compensation. No compensation will be provided directly to survey participants for Phase 2.

No prisoners or institutionalized individuals will be involved in the proposed research. We recognize that, as 66% of incarcerated individuals report having a disability [13], approximately 3% of disabled individuals live in institutional group quarters [97], and people with disabilities are more likely to experience state-sanctioned violence [13], excluding incarcerated and institutionalized individuals creates limitations on capturing the full experience of structural ableism. However, we will include historical narratives that explore the experience of incarceration and institutionalization, and will work to include disabled individuals who have previously experienced both incarceration and institutionalization.

Results

This project was funded in June of 2024. As of October 2025, our team has read over 50 texts as part of our historical and policy analysis of the factors that characterize structural ableism. We plan to complete our characterization of structural ableism in the spring of 2026, with individual-level measures of structural ableism being developed by the Winter of 2028 and community-level measures created by the Winter of 2029.

XSL•FO
RenderX

Discussion

Anticipated Findings

The disability community is in urgent need of systems-based solutions to improve health equity [14]. This project lays the essential groundwork for understanding the etiology of these disparities and the pathways that must be disrupted to achieve health equity for people with disabilities. We anticipate that the findings of our project will result in a multidimensional characterization of structural ableism that includes both individual and community-level measures. The Phase I deliverable of characterizing the factors that comprise structural ableism will be validated through analysis of texts and key informant interviews that will aim to capture the diverse ways individuals can be impacted by structural ableism. The Phase II deliverable of developing an individual-level measure of structural ableism will be validated through surveys conducted with members of the disability community. Finally, the Phase III deliverable of developing a community-level measure of structural ableism will be validated through the use of publicly available datasets. We hypothesize that the measures developed upon the completion of the 3 phases of this project will inform a new understanding of the relationship between health disparities and structural ableism in order to improve health outcomes of the disability community.

Across all 3 phases, we take an intersectional, interdisciplinary, and community-grounded approach [65]. We purposefully include disabled people across all steps and phases of this project, with a focus on maximizing diversity of disability perspectives by including people across disability types and with intersectional identities (eg, race and ethnicity, gender identity, and geographic location) [60,67,98]. Our interdisciplinary team has expertise in disability studies, public health, medicine, health policy, systems engineering, sociology, and cultural anthropology, and is led by 2 principal investigators with disabilities (RSV and BKS). Our approach is deeply community-informed, drawing on multiple community partnerships from local and national organizations in the context of our advisory committee and drawing on engagement directly with the disability community across all 3 aims and dissemination activities.

As we begin this project, we have made multiple decision points, primarily pertaining to how this work can be made as inclusive as possible. First, we have been working with the disability community to determine the best ways to have our work resonate with those who live with functional limitations or chronic illness but do not identify as part of the disability community. We want the measures we develop to reflect all who experience the barriers caused by disability, not just those who identify as disabled. Second, as pertaining to dissemination, we have made considerations around balancing ways to make the study as inclusive as possible with the funding and resources we have available. While we are working within these constraints, we are also searching for opportunities for additional funding to expand in directions that will make our work more inclusive and accessible. This includes a project funded by the WITH foundation in 2025 that focuses on using artificial intelligence to create plain language materials. Representing our work in plain language, incorporating artwork to visualize the concepts discussed, and finding ways to include the experiences of deaf individuals and ASL users are all areas being evaluated to improve dissemination and accessibility.

Conclusions

The measures developed by this work will create a foundation for identifying and evaluating novel interventions aimed at dismantling structural ableism. These interventions should be co-created with the disability community.

Acknowledgments

The authors would like to thank Ananya Sairaman for their graphical design work in producing the graphics in Figures 1 and 2.

Funding

The funding for this project is provided by the National Institutes of Health’s Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Data Availability

Data sharing is not applicable to this article as no data have yet been generated from the study.

Authors' Contributions

RSV and BKS conceptualized the project and the paper and were responsible for all final edits. AS and DC contributed to the drafting of Phase I, KSF contributed to the drafting of Phase II, and FC and VV contributed to the drafting of Phase III. CC contributed to the drafting of the introduction. HS led the drafting of all components of Figure 1. BAM led the drafting of the introduction, all components of Figure 1, and the editing of all other sections. All authors contributed to the final edits and review prior to manuscript submission.

XSL•FO
RenderX

Conflicts of Interest

Some reviewers were found to have undisclosed Conflicts of Interest with members of the author list under the terms of JMIR Publications Policy. The Editor-in-Chief determined that the content of the reviews did not unduly influence the decision to accept the article. The decision to accept the submission was based primarily on the remaining reviewer's comments and the Editor-in-Chief's contemporary appraisal of the manuscript.

Multimedia Appendix 1
Spanish language identify first Phase I interview guide.
[DOCX File , 481 KB - resprot_v15i1e86976_app1.docx]

Multimedia Appendix 2
Spanish language person first Phase I interview guide.
[DOCX File , 480 KB - resprot_v15i1e86976_app2.docx]

Multimedia Appendix 3
English language Phase I identify first interview guide.
[DOCX File , 2768 KB - resprot_v15i1e86976_app3.docx]

Multimedia Appendix 4
English language Phase I person first interview guide.
[DOCX File , 2768 KB - resprot_v15i1e86976_app4.docx]

Multimedia Appendix 5
Link to ASL Phase I interview guide video.
[DOCX File , 13 KB - resprot_v15i1e86976_app5.docx]

References

1. Pérez-Stable EJ, Valdez RO. Announcement of decision to designate people with disabilities as a population with health disparities. Natl. Inst. Minor. Health Health Disparities. 2023 Sep 26. URL: https://www.nih.gov/news-events/news-releases/nih-designates-people-disabilities-population-health-disparities [accessed 2023-11-19]
2. Kuper H, Rotenberg S, Azizatunnisa' L, Banks LM, Smythe T. The association between disability and mortality: a mixed-methods study. Lancet Public Health 2024;9(5):e306-e315 [FREE Full text] [doi: 10.1016/S2468-2667(24)00054-9] [Medline: 38702095]
3. Iezzoni LI, Rao SR, Agaronnik ND, El-Jawahri A. Associations between disability and breast or cervical cancers, accounting for screening disparities. Med Care 2021;59(2):139-147 [FREE Full text] [doi: 10.1097/MLR.0000000000001449] [Medline: 33201087]
4. Landes SD, Finan JM, Turk MA. COVID-19 mortality burden and comorbidity patterns among decedents with and without intellectual and developmental disability in the US. Disabil Health J 2022;15(4):101376 [FREE Full text] [doi: 10.1016/j.dhjo.2022.101376] [Medline: 36175298]
5. Cree RA, Okoro CA, Zack MM, Carbone E. Frequent mental distress among adults, by disability status, disability type, and selected characteristics - United States, 2018. MMWR Morb Mortal Wkly Rep 2020;69(36):1238-1243 [FREE Full text] [doi: 10.15585/mmwr.mm6936a2] [Medline: 32914770]
6. Harrell E. Crime against persons with disabilities, 2009-2019 – statistical tables. Bureau of Justice Statistics. 2021. URL: https://bjs.ojp.gov/content/pub/pdf/capd0919st.pdf [accessed 2026-01-23]
7. Mitra M, Long-Bellil L, Moura I, Miles A, Kaye HS. Advancing health equity and reducing health disparities for people with disabilities in the United States. Health Aff (Millwood) 2022;41(10):1379-1386. [doi: 10.1377/hlthaff.2022.00499] [Medline: 36190894]
8. Horner-Johnson W, Garg B, Caughey AB, Akobirshoev I, Mitra M. Adverse perinatal outcomes among black and hispanic women with physical disabilities. Am J Obstet Gynecol 2023;228(1):S411. [doi: 10.1016/j.ajog.2022.11.716]
9. Goodman N, Morris M, Boston K, Walton D. Financial inequality: disability, race and poverty in America. National Disability Institute. URL: https://www.nationaldisabilityinstitute.org/wp-content/uploads/2019/02/disability-race-poverty-in-america.pdf [accessed 2026-11-23]
10. Crankshaw K. Disability rates higher in rural areas than urban areas. United States Census Bureau. 2023. URL: https://tinyurl.com/5bsz2vtb [accessed 2026-11-23]
11. Iezzoni LI, Rao SR, Ressalam J, Bolcic-Jankovic D, Agaronnik ND, Donelan K, et al. Physicians' perceptions of people with disability and their health care. Health Aff (Millwood) 2021;40(2):297-306 [FREE Full text] [doi: 10.1377/hlthaff.2020.01452] [Medline: 33523739]

XSL•FO
RenderX

12. Keefe ES. From detractive to democratic: the duty of teacher education to disrupt structural ableism and reimagine disability. Teachers College Record: The Voice of Scholarship in Education 2022;124(3):115-147. [doi: 10.1177/01614681221086994]
13. Lundberg DJ, Chen JA. Structural ableism in public health and healthcare: a definition and conceptual framework. Lancet Reg Health Am 2024;30:100650 [FREE Full text] [doi: 10.1016/j.lana.2023.100650] [Medline: 38188095]
14. Valdez RS, Swenor BK. Structural ableism - essential steps for abolishing disability injustice. N Engl J Med 2023;388(20):1827-1829. [doi: 10.1056/NEJMp2302561] [Medline: 37183982]
15. Singh S. Structural ableism by the indian nursing council must be challenged. J Psychosoc Nurs Ment Health Serv 2024;62(9):4-4 [FREE Full text] [doi: 10.3928/02793695-20240815-01] [Medline: 39225672]
16. Iezzoni LI. Ableism and structural ableism in health care workplaces. JAMA Netw Open 2024;7(8):e2430315 [FREE Full text] [doi: 10.1001/jamanetworkopen.2024.30315] [Medline: 39196564]
17. Christakis DA, Iezzoni LI. Calling on the USPSTF to address ableism and structural ableism. JAMA 2023;330(14):1327-1328. [doi: 10.1001/jama.2023.17092] [Medline: 37747717]
18. Quon M. Addressing ableism in physician well-being planning. JAMA 2024;332(4):275-276. [doi: 10.1001/jama.2024.7736] [Medline: 38913395]
19. Poed S, Fox RA. Identifying and removing ableism from tier 1 school-wide positive behaviour support practices. Int J Dev Disabil 2023;69(1):83-94 [FREE Full text] [doi: 10.1080/20473869.2022.2111969] [Medline: 36743320]
20. Reber L, Kreschmer JM, James TG, Junior JD, DeShong GL, Parker S, et al. Ableism and contours of the attitudinal environment as identified by adults with long-term physical disabilities: a qualitative study. Int J Environ Res Public Health 2022;19(12):7469 [FREE Full text] [doi: 10.3390/ijerph19127469] [Medline: 35742718]
21. Mintz KT, Gammer B, Khan AJ, Shaub G, Levine S, Sisti D. Physical disability and psychedelic therapies: an agenda for inclusive research and practice. Front Psychiatry 2022;13:914458 [FREE Full text] [doi: 10.3389/fpsyt.2022.914458] [Medline: 35693959]
22. Keegan G, Rizzo J, Gonzalez CM, Joseph K. Reducing barriers through education: a scoping review calling for structured disability curricula in surgical training programs. Am J Surg 2025;239:116062. [doi: 10.1016/j.amjsurg.2024.116062] [Medline: 39504925]
23. Barton HJ, Valdez RS, Shew A, Swenor BK, Jolliff A, Claypool H, et al. A call for integrated approaches in digital technology design for aging and disability. Gerontologist 2025;65(6). [doi: 10.1093/geront/gnaf113] [Medline: 40123375]
24. Mintz KT. Serendipity and social justice: how someone with a physical disability succeeds in clinical bioethics. Narrat Inq Bioeth 2022;12(3):265-273 [FREE Full text] [doi: 10.1353/nib.2022.0061] [Medline: 37854450]
25. Iezzoni LI. What some physicians say about caring for patients with disability: responses to open-ended question to nationwide physician survey. Disabil Health J 2024;17(3):101635. [doi: 10.1016/j.dhjo.2024.101635] [Medline: 38688730]
26. Sloutsky VM, Sophia Deng W. Categories, concepts, and conceptual development. Lang Cogn Neurosci 2019;34(10):1284-1297 [FREE Full text] [doi: 10.1080/23273798.2017.1391398] [Medline: 32775486]
27. Brown TH, Homan P. Structural racism and health stratification: connecting theory to measurement. J Health Soc Behav 2024;65(1):141-160 [FREE Full text] [doi: 10.1177/00221465231222924] [Medline: 38308499]
28. Dean LT, Thorpe RJ. What structural racism is (or is not) and how to measure it: clarity for public health and medical researchers. Am J Epidemiol 2022;191(9):1521-1526 [FREE Full text] [doi: 10.1093/aje/kwac112] [Medline: 35792088]
29. Adkins-Jackson PB, Chantarat T, Bailey ZD, Ponce NA. Measuring structural racism: a guide for epidemiologists and other health researchers. Am J Epidemiol 2022;191(4):539-547 [FREE Full text] [doi: 10.1093/aje/kwab239] [Medline: 34564723]
30. Chantarat T, Van Riper DC, Hardeman RR. Multidimensional structural racism predicts birth outcomes for black and white Minnesotans. Health Serv Res 2022;57(3):448-457 [FREE Full text] [doi: 10.1111/1475-6773.13976] [Medline: 35468220]
31. Gee GC, Hicken MT. Structural racism: the rules and relations of inequity. Ethn Dis 2021;31(Suppl 1):293-300 [FREE Full text] [doi: 10.18865/ed.31.S1.293] [Medline: 34045831]
32. Groos M, Wallace M, Hardeman R, Theall KP. Measuring inequity: a systematic review of methods used to quantify structural racism. J Health Disparities Res Pract 2018;11(2) [FREE Full text]
33. Hardeman RR, Homan PA, Chantarat T, Davis BA, Brown TH. Improving the measurement of structural racism to achieve antiracist health policy. Health Aff (Millwood) 2022;41(2):179-186 [FREE Full text] [doi: 10.1377/hlthaff.2021.01489] [Medline: 35130062]
34. Hing AK, Hassan A, Hardeman RR. Advancing the measurement of structural racism through the lens of antiabortion policy. Am J Public Health 2022;112(11):1529-1531. [doi: 10.2105/AJPH.2022.307091] [Medline: 36223566]
35. Scambler G. Sociology, social class, health inequalities, and the avoidance of "Classism". Front Sociol 2019;4:56 [FREE Full text] [doi: 10.3389/fsoc.2019.00056] [Medline: 33869379]
36. Chang Y, Buerke M, Galfalvy H, Szanto K. Childhood trauma is associated with early-onset but not late-onset suicidal behavior in late-life depression. Int Psychogeriatr 2024;36(5):371-384 [FREE Full text] [doi: 10.1017/S1041610223000662] [Medline: 37642013]
37. Levy BR. The role of structural ageism in age beliefs and health of older persons. JAMA Netw Open 2022;5(2):e2147802 [FREE Full text] [doi: 10.1001/jamanetworkopen.2021.47802] [Medline: 35138405]
38. Rosales A, Fernández-Ardèvol M. Structural ageism in big data approaches. Nord Rev 2019;40:51-64. [doi: 10.2478/nor-2019-0013]

XSL•FO
RenderX

39. Beccia AL, Austin SB, Baek J, Agénor M, Forrester S, Ding EY, et al. Cumulative exposure to state-level structural sexism and risk of disordered eating: results from a 20-year prospective cohort study. Soc Sci Med 2022;301:114956 [FREE Full text] [doi: 10.1016/j.socscimed.2022.114956] [Medline: 35436662]
40. Homan P. Structural sexism and health in the United States: a new perspective on health inequality and the gender system. Am Sociol Rev 2019;84(3):486-516. [doi: 10.1177/0003122419848723]
41. Homan P, Burdette A. When religion hurts: structural sexism and health in religious congregations. Am Sociol Rev 2021;86(2):234-255. [doi: 10.1177/0003122421996686]
42. Krieger N. Measures of racism, sexism, heterosexism, and gender binarism for health equity research: from structural injustice to embodied harm-an ecosocial analysis. Annu Rev Public Health 2020;41:37-62 [FREE Full text] [doi: 10.1146/annurev-publhealth-040119-094017] [Medline: 31765272]
43. Hatzenbuehler ML, Keyes KM, Hasin DS. State-level policies and psychiatric morbidity in lesbian, gay, and bisexual populations. Am J Public Health 2009;99(12):2275-2281. [doi: 10.2105/AJPH.2008.153510] [Medline: 19833997]
44. Bonilla-Silva E. Rethinking racism: toward a structural interpretation. Am Sociol Rev 1997;62(3):465. [doi: 10.2307/2657316]
45. Miller CE, Meyers SA. Disparities in school discipline practices for students with emotional and learning disabilities and autism. JEHD 2015;4(1). [doi: 10.15640/jehd.v4n1a23]
46. Lawer L, Brusilovskiy E, Salzer MS, Mandell DS. Use of vocational rehabilitative services among adults with autism. J Autism Dev Disord 2009;39(3):487-494. [doi: 10.1007/s10803-008-0649-4] [Medline: 18810627]
47. Schniedewind E, Lindsay R, Snow S. Ask and ye shall not receive: interpreter-related access barriers reported by deaf users of American sign language. Disabil Health J 2020;13(4):100932. [doi: 10.1016/j.dhjo.2020.100932] [Medline: 32576507]
48. Kullgren J. Trends in loneliness among older adults from 2018-2023. Institute for Healthcare Policy & Innovation University of Michigan. URL: https://ihpi.umich.edu/national-poll-healthy-aging/national-findings/trends-loneliness-among-older-adults-2018-2023 [accessed 2023-03-13]
49. Ohl A, Grice Sheff M, Small S, Nguyen J, Paskor K, Zanjirian A. Predictors of employment status among adults with autism spectrum disorder. Work 2017;56(2):345-355. [doi: 10.3233/WOR-172492] [Medline: 28211841]
50. Johns CB, Savage JH. Access to health care and food in children with food allergy. J Allergy Clin Immunol 2014;133(2):582-585 [FREE Full text] [doi: 10.1016/j.jaci.2013.12.006] [Medline: 24636475]
51. Guide to disability rights laws - Americans with disabilities act (ADA). U.S. Department of Justice Civil Rights Division. 2020. URL: https://tinyurl.com/4d4tnh2w [accessed 2025-11-18]
52. Disability and health overview. Centers for Disease Control and Prevention (CDC). 2024. URL: https://www.cdc.gov/disability-and-health/about/index.html [accessed 2026-11-23]
53. Lee L, Mou F, Um Boock A, Fellinghauer C, Kohls M, Cieza A, et al. Identifying key environmental barriers experienced by persons with mild, moderate, or severe disability in Bankim Health District, Cameroon: a policy-targeted secondary analysis of data obtained with the World Bank and WHO model disability survey. Arch Public Health 2021;79(1):95 [FREE Full text] [doi: 10.1186/s13690-021-00619-y] [Medline: 34099049]
54. Janz H. MAID to die by medical and systemic ableism. In: Kotalik J, Shannon DW, editors. Medical Assistance in Dying (MAID) in Canada: Key Multidisciplinary Perspectives. Cham, Switzerland: Springer International Publishing; 2023:299-308.
55. Putnam M. Conceptualizing disability: developing a framework for political disability identity. J Disabil Policy Stud 2005;16(3):188-198. [doi: 10.1177/10442073050160030601]
56. Hahn HD, Belt TL. Disability identity and attitudes toward cure in a sample of disabled activists. J Health Soc Behav 2004;45(4):453-464. [doi: 10.1177/002214650404500407] [Medline: 15869116]
57. Brown RL. Structural ableism, health, and functioning among people with disabilities: an extension of the social determinants of health framework. J Health Soc Behav 2025:221465251379456. [doi: 10.1177/00221465251379456] [Medline: 41109850]
58. Timberlake M. Recognizing ableism in educational initiatives: reading between the lines. REPAM 2020;2(1):84-100. [doi: 10.46303/repam.02.01.5]
59. Rowe S, Dowse L, Baker M, Baldry E. Policing disability: alliance building, police divestment and community investment. Curr Issues Crim Justice 2022;34(2):171-187. [doi: 10.1080/10345329.2022.2029084]
60. Fuentes K, Hsu S, Patel S, Lindsay S. More than just double discrimination: a scoping review of the experiences and impact of ableism and racism in employment. Disabil Rehabil 2024;46(4):650-671. [doi: 10.1080/09638288.2023.2173315] [Medline: 36724368]
61. Iezzoni L. Disability Disparities and Ableism in Medicine, in ABC of equality, diversity and inclusion in healthcare, in ABC Series. Hoboken, New Jersey: Wiley Blackwell; 2023:61.
62. Krieger N. Discrimination and health inequities. Int J Health Serv 2014;44(4):643-710. [doi: 10.2190/HS.44.4.b] [Medline: 25626224]
63. Agénor M, Perkins C, Stamoulis C, Hall RD, Samnaliev M, Berland S, et al. Developing a database of structural racism-related state laws for health equity research and practice in the United States. Public Health Rep 2021;136(4):428-440 [FREE Full text] [doi: 10.1177/0033354920984168] [Medline: 33617383]
64. Imrie R, Kumar M. Focusing on disability and access in the built environment. Disability & Society 1998;13(3):357-374. [doi: 10.1080/09687599826687]

XSL•FO
RenderX

65. Wyatt TR, Johnson M, Zaidi Z. Intersectionality: a means for centering power and oppression in research. Adv Health Sci Educ Theory Pract 2022;27(3):863-875. [doi: 10.1007/s10459-022-10110-0] [Medline: 35366113]
66. Hudson STJ, Myer A, Berney EC. Stereotyping, prejudice, and discrimination at the intersection of race and gender: an intersectional theory primer. Soc Personal Compas 2024;18(2):e12939. [doi: 10.1111/spc3.12939]
67. Swenor B, Deal JA. Disability inclusion as a key component of research study diversity. N Engl J Med 2022;386(3):205-207. [doi: 10.1056/NEJMp2115475] [Medline: 35029848]
68. Epstein S, Campanile J, Cerilli C, Gajwani P, Varadaraj V, Swenor BK. New obstacles and widening gaps: a qualitative study of the effects of the COVID-19 pandemic on U.S. adults with disabilities. Disabil Health J 2021;14(3):101103 [FREE Full text] [doi: 10.1016/j.dhjo.2021.101103] [Medline: 33840617]
69. Web content accessibility guidelines (WCAG) 2.2. World Wide Web Consortium. 2023. URL: https://www.w3.org/TR/WCAG22/ [accessed 2023-11-27]
70. The Rehabilitation Act. GovInfo. URL: https://www.eeoc.gov/rehabilitation-act-1973 [accessed 2023-09-01]
71. Public law 111 - 274 - plain writing act of 2010. GovInfo. URL: https://www.govinfo.gov/app/details/PLAW-111publ274 [accessed 2023-11-27]
72. Nielsen K. A Disability History of the United States. Massachusetts: Beacon Press; Mar 24, 2013.
73. Schweik S. The Ugly Laws: Disability in Public. New York: New York University Press; 2009.
74. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
75. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
76. Ritchie D. Doing Oral History. Oxford, England: Oxford University Press; 2024.
77. Flanagan JC. The critical incident technique. Psychol Bull 1954;51(4):327-358. [doi: 10.1037/h0061470] [Medline: 13177800]
78. Hinkin TR. A review of scale development practices in the study of organizations. Journal of Management 1995;21(5):967-988. [doi: 10.1177/014920639502100509]
79. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
80. Lincoln YS, Guba EG, Pilotta JJ. Naturalistic Inquiry. Thousand Oaks, CA: SAGE Publ; 1985:438-439.
81. Marshall C, Rossman G, Blanco G. Designing Qualitative Research. In: Journal of Marketing Research. Thousand Oaks, CA: SAGE Publications; Aug 1990:370.
82. Balza JS, Cusatis R, McDonnell SM, Basir MA, Flynn KE. Effective questionnaire design: how to use cognitive interviews to refine questionnaire items. J Neonatal Perinatal Med 2022;15(2):345-349 [FREE Full text] [doi: 10.3233/NPM-210848] [Medline: 35034911]
83. How disability data are collected from the American community survey. US Census Bureau. URL: https://www.census.gov/topics/health/disability/guidance/data-collection-acs.html [accessed 2023-11-27]
84. Willis GS. Analysis of the Cognitive Interview in Questionnaire Design. England: Oxford University Press; 2015.
85. Dynata. URL: https://www.dynata.com/ [accessed 2023-11-28]
86. Boateng GO, Neilands TB, Frongillo EA, Melgar-Quiñonez HR, Young SL. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health 2018;6:149 [FREE Full text] [doi: 10.3389/fpubh.2018.00149] [Medline: 29942800]
87. Behavioral risk factor surveillance system 2021 data and documentation. Centers for Disease Control and Prevention. URL: https://www.cdc.gov/brfss/annual_data/annual_2021.html [accessed 2023-11-27]
88. Krahn GL, Walker DK, Correa-De-Araujo R. Persons with disabilities as an unrecognized health disparity population. Am J Public Health 2015;105 Suppl 2(Suppl 2):S198-S206. [doi: 10.2105/AJPH.2014.302182] [Medline: 25689212]
89. Akobirshoev I, Vetter M, Iezzoni LI, Rao SR, Mitra M. Delayed medical care and unmet care needs due to the COVID-19 pandemic among adults with disabilities in the US. Health Aff (Millwood) 2022;41(10):1505-1512. [doi: 10.1377/hlthaff.2022.00509] [Medline: 36190876]
90. Okoro CA, Hollis ND, Cyrus AC, Griffin-Blake S. Prevalence of disabilities and health care access by disability status and type among adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018;67(32):882-887 [FREE Full text] [doi: 10.15585/mmwr.mm6732a3] [Medline: 30114005]
91. Joosten YA, Israel TL, Williams NA, Boone LR, Schlundt DG, Mouton CP, et al. Community engagement studios: a structured approach to obtaining meaningful input from stakeholders to inform research. Acad Med 2015;90(12):1646-1650 [FREE Full text] [doi: 10.1097/ACM.0000000000000794] [Medline: 26107879]
92. Israel L, Farrow H, Joosten Y, Vaughn Y. Community Engagement Studio Toolkit 2.0. 2025. URL: https://heller.brandeis.edu/ibh/affiliates/mmhrc/pdfs/community-engagement-studio-toolkit.pdf [accessed 2026-01-23]
93. Joosten YA, Israel TL, Head A, Vaughn Y, Villalta Gil V, Mouton C, et al. Enhancing translational researchers' ability to collaborate with community stakeholders: lessons from the community engagement studio. J Clin Transl Sci 2018;2(4):201-207 [FREE Full text] [doi: 10.1017/cts.2018.323] [Medline: 30820357]

XSL•FO
RenderX

94. Blewett LA. National health interview survey, version 7.3. IPUMS Health Surveys. 2023. URL: https://www.ipums.org/projects/ipums-health-surveys/d070.V7.3 [accessed 2026-01-23]
95. Dougherty GB, Golden SH, Gross AL, Colantuoni E, Dean LT. Measuring structural racism and its association with BMI. Am J Prev Med 2020;59(4):530-537 [FREE Full text] [doi: 10.1016/j.amepre.2020.05.019] [Medline: 32863079]
96. DiStefano C, Hess B. Using confirmatory factor analysis for construct validation: an empirical review. J Psychoeduc Assess 2005;23(3):225-241. [doi: 10.1177/073428290502300303]
97. Annual report on people with disabilities in America: 2024. Rehabilitation Research and Training Center on Disability Statistics and and Demographics. 2024. URL: https://www.researchondisability.org/sites/default/files/media/2024-03/2024-annual-report-a11y-1.pdf [accessed 2026-01-23]
98. Ramsay NJ. Intersectionality: a model for addressing the complexity of oppression and privilege. Pastoral Psychol 2013;63(4):453-469. [doi: 10.1007/s11089-013-0570-4]

Abbreviations

ACS-6: American Community Survey Questions on Disability
ADA: Americans With Disabilities Act
ASL: American Sign Language
HIPAA: Health Insurance Portability and Accountability Act
IDD: intellectual and developmental disabilities
IRB: Institutional Review Board
UVA: University of Virginia
WCAG: Web Content Accessibility Guidelines

Edited by A Schwartz; submitted 02.Nov.2025; peer-reviewed by MM Wilcox, Anyonymous, M Kakara; comments to author 26.Nov.2025; revised version received 20.Dec.2025; accepted 20.Jan.2026; published 13.Mar.2026.

Please cite as:
Valdez RS, Swenor BK, Castro F, Cerilli C, Middleton BA, Fitzgibbon KS, Shew A, Varadaraj V, Srinivasan H, Cejas D
Developing and Validating Measures of Structural Ableism to Improve Health Outcomes for the Disability Community: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e86976
URL: https://www.researchprotocols.org/2026/1/e86976
doi:10.2196/86976
PMID:

©Rupa S Valdez, Bonnielin K Swenor, Franz Castro, Caroline Cerilli, Bailey A Middleton, Kara S Fitzgibbon, Ashley Shew, Varshini Varadaraj, Hari Srinivasan, Diana Cejas. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Machine Learning for Estimating Cardiorespiratory Fitness in Patients With Obesity: Protocol for a Retrospective and Prospective Multicenter Cohort Study

Jarle Berge[1,2,3], PhD; Vimala Nunavath[4], PhD; Rikke Aune Asbjørnsen[5,6], PhD; Heidi Borgeraas[1], PhD; Jens Kristoffer Hertel[1], PhD; Anita Dyb Linge[7,8], PhD; Inge Groote[9], PhD; Trine Stensrud[10], PhD; Espen Gjevestad[3], PhD; Linda Mathisen[1], BSc; Matthan WA Caan[11], PhD; Martin Paulson[12], MD; Eva Maria Støa[13], PhD; Jan - Michael Johansen[13], PhD; AI and VO2max Group[14,17]; Bjørn-Jostein Singstad[1,4,5,9,15,16], Phd

[1]Department of Endocrinology, Obesity and Nutrition, Vestfold Hospital Trust, Box 2168, Tonsberg, Norway
[10]Department of Sports Medicine, Norwegian School of Sports Sciences, Oslo, Norway
[11]Department of Radiology, University of Amsterdam, Amsterdam, The Netherlands
[12]Department of Oncology and Hematology, Vestfold Hospital Trust, Tonsberg, Norway
[13]Department of Sports, Physical Education and Outdoor Studies, University of South-Eastern Norway, Bø, Norway
[14]See Acknowledgments
[15]Medical Technology and E-health, Akershus University Hospital, Lørenskog, Norway
[16]Institute of Clinical Medicine, University of Oslo, Oslo, Norway
[17]
[2]Department of Neurology and Rehabilitation, Vestfold Hospital Trust, Stavern, Norway
[3]Department of Joint Operative Disciplines, Norwegian Police University College, Stavern, Norway
[4]Department of Science and Industry Systems, University of South-Eastern Norway, Kongsberg, Norway
[5]Department of Research and Innovation, Vestfold Hospital Trust, Tonsberg, Norway
[6]Departement of Digital Health Research, Oslo University Hospital, Oslo, Norway
[7]Muritunet Rehabilitation Centre, Valldal, Norway
[8]Department of Social Science and History, Volda University College, Volda, Norway
[9]Department of Radiology, Vestfold Hospital Trust, Tonsberg, Norway

Corresponding Author:
Jarle Berge, PhD
Department of Endocrinology, Obesity and Nutrition, Vestfold Hospital Trust, Box 2168, Tonsberg, Norway

Abstract

Background: Cardiorespiratory fitness (CRF) is a key predictor of cardiovascular and other health-related diseases in individuals with obesity. CRF is most accurately assessed through maximal exercise testing with advanced gas-analysis equipment (maximum volume of oxygen [VO_{2max}]); however, this approach is time-consuming, costly, and requires specialized expertise. Therefore, submaximal tests and self-reported physical activity levels have been used to develop predictive algorithms to estimate CRF, yet they often performed poorly in individuals with low CRF levels, such as patients with obesity, because they are predominantly developed using data from healthy populations. Studies using machine learning (ML) models based on VO_{2max} data from patients with obesity appear to be lacking in the literature. ML models based on routinely collected clinical measures may offer a more practical and potentially accurate way to estimate CRF, reducing time, costs, and clinical burden.

Objective: The primary aim of this study is to use multicenter, longitudinal, real-world clinical data from a uniquely characterized population with obesity to develop and validate a clinically relevant ML model for estimating CRF and to compare its performance with the gold standard of VO_{2max} testing.

Methods: A retrospective data set combining assessments of VO_{2max} tests and clinical parameters from adult patients with severe obesity BMI (≥40.0 kg/m^2 or 35.0 - 39.9 kg/m^2 with at least 1 obesity-related comorbidity) from Vestfold Hospital Trust, Muritunet Rehabilitation Institution, and Norwegian School of Sport Sciences will be the foundation for developing the ML model. The clinically relevant ML model for estimating CRF will be presented as a web application, allowing easy access and interaction. The model's estimations will be compared against direct VO_{2max} measurements obtained from medical equipment across institutions as part of a prospective validation. Ethical approval has been obtained for the use of 2 databases in the initial model development; approval for the remaining data and prospective phase is pending.

XSL•FO
RenderX

Results: Vestfold Hospital Trust, the Norwegian School of Sport Medicine, and Muritunet Rehabilitation Institution conducted more than 2623 $VO_{2\ max}$ tests and collected clinical parameters from 1279 adults with severe obesity during 2013-2025, both before, during, and after lifestyle interventions. The first scientific publication of the clinically relevant ML model is expected to be published in 2026. The results of the overall project are expected to be completed in 2028. The project was awarded salary funding in two rounds from Vestfold Hospital (October and December 2023) and salary funding from the Research Council of Norway (October 2024). In addition, the project received allocated supervision hours from InnoMed Norway (April 2024).

Conclusions: This project aims to develop a clinically relevant ML model, which serves as a cost-effective tool for CRF estimation in individuals with obesity, improving accessibility to this important health marker. To our knowledge, this is the first initiative in Norway to estimate CRF in individuals with obesity using ML, based on a unique clinical database. The project carries substantial societal value and holds national and international relevance for health care practice and patient outcomes.

Trial Registration: ClinicalTrials.gov NCT07011108; https://clinicaltrials.gov/study/NCT07011108

International Registered Report Identifier (IRRID): DERR1-10.2196/85069

(JMIR Res Protoc 2026;15:e85069) doi:10.2196/85069

KEYWORDS

VO2max; VO2max estimation; regression; obesity; cardiorespiratory fitness; cardiorespiratory fitness estimation; machine learning; web application

Introduction

Background

Obesity is a primary risk factor for more than 2.8 million preventable deaths worldwide [1] and over 3200 preventable deaths in Norway [2]. Obesity is also strongly linked to lost years of life [2], primarily due to cardiovascular disease and other obesity-related diseases [3-7]. Cardiovascular disease is closely related to cardiorespiratory fitness (CRF), defined as the body's ability to use oxygen [8]. Further, obesity is associated with poor physical ability and impaired CRF. Therefore, improvements in cardiorespiratory fitness reduce the risk of cardiovascular disease, all-cause mortality, and contribute to enhanced overall health among people with obesity [7,9]. CRF is thus an important health indicator for both the prevention and treatment of obesity [6,10,11].

The gold standard for measurement of CRF is a maximal test on a treadmill or bicycle connected to expensive medical equipment that analyzes respiratory gas (maximum volume of oxygen [VO_{2max}, [$mL*min^{-1}$]) [12,13]. Performing the test is a burden for the patients related to travel, time, absence from work or school, and costs (eg, parking and travel cost). The challenges for the clinic are financial costs, time investment, and the burden of administrating a direct physical measurement of the VO_{2max} test. In addition, the test requires experience and knowledge of a test leader [13]. Further, methods for indirect estimation of VO_{2max} or more precise assessment of CRF have been developed, such as the "submaximal tests" [14,15]. The submaximal assessment of CRF does not require patients to exercise to exhaustion but instead estimates CRF based on physiological responses, usually heart rate, speed, intensity, and duration during exercise on a treadmill, ergometer cycle, or step box [13,16]. A further simplification is self-reported questionnaires that include information on physical activity level, demographic characteristics, and clinical characteristics have also been used for estimating CRF alone or in combination with submaximal assessments [13,17]. Nevertheless, the challenges associated with both the direct maximal test and submaximal test have contributed to the development of different predictive algorithms to estimate CRF [16-18]. The algorithms are developed using data from direct physical measurements of VO_{2max} and commonly require information about gender, age, height, body weight, heart rate (maximal or resting values), and exercise habits to estimate CRF [16,17]. These algorithms are typically developed using conventional statistical models and are generally advantageous and convenient for large populations [16,17]. Furthermore, as algorithms are commonly trained on data from healthy volunteers, the lack of sufficient data from individuals with low VO_{2max} values (eg, patients with obesity) may lead to less reliable estimates [13,16,17,19,20].

In recent years, machine learning (ML) has demonstrated that advanced models can outperform conventional statistical models in clinical prediction tasks. A ML model is a mathematical construct trained to recognize patterns in data, enabling it to make predictions or informed decisions based on nonlinear combinations of input variables, without requiring explicit programming for each specific task. Therefore, ML may provide more accurate estimates of CRF compared to traditional methods. A digital accessibility solution incorporating a ML model with high accuracy for estimating CRF could enable routine assessment of this key health indicator in clinical settings for patients with obesity [18]. This, in turn, may facilitate the integration of CRF into broader clinical decision-making and evaluation of individuals' health.

Objectives

Studies using ML models based on VO_{2max} data from patients with overweight and obesity appear to be lacking in the literature [13]. Therefore, the primary objective of this study is to use multicenter, longitudinal, real-world clinical data from a uniquely characterized population with obesity to develop and validate a clinically relevant ML model for estimating CRF, and to compare its performance with the gold standard of VO_{2max} testing. Furthermore, the model will be implemented as a web-based tool to support real-world use in a European

XSL•FO
RenderX

clinical setting. Initial feedback and empirical insights into the usability and feasibility of the web application will be collected in health care contexts. The primary hypothesis is that a clinically relevant ML model can estimate CRF with acceptable accuracy. To accomplish this, the following subgoals have been established:

1. Model development and evaluation of retrospective data: Train and evaluate multiple advanced ML models using our retrospective dataset of VO_{2max} and other medical measurements (eg, body weight, age, and fat mass) to estimate CRF in individuals with obesity and select the top-performing model.
2. Web application development: Develop a web-based application that incorporates the optimal ML model identified in goal 1, to enable accessible VO_{2max} estimation.
3. Prospective validation and pilot study: Conduct a prospective validation study using the web application, developed in goal 2, to evaluate the performance of the ML model. This will involve comparing the model's predictions of CRF with direct physical measurements of VO_{2max}. Additionally, a pilot study will be carried out to gather initial feedback and insights on the usability and feasibility of the web application within the health care setting.
4. Model management and monitoring: Implement a system to manage and monitor the ML model, aligning with relevant regulatory frameworks like the European Union Artificial Intelligence Act [21]. This system will include mechanisms for detecting and addressing out-of-distribution data, as well as provisions for continuous model re-training using data routinely collected in the clinical setting.
5. Implementation in clinical practice*:* Implement the ML model estimation through the web application as a clinical tool.

Methods

Machine Learning Model Development and Prospective Validation

ML models differ from simpler statistical approaches by their ability to capture nonlinear relationships within data, unlike traditional statistical methods that focus on linear associations. The capacity of ML models to uncover nonlinear patterns enables them to approximate a wide range of mathematical functions, revealing relationships that would be difficult to discern using classical statistical techniques. Therefore, in the retrospective development phase, the model will be trained and internally evaluated using existing, previously collected measurements, both published and unpublished data. In the prospective validation phase, the final model will be evaluated in a prospective validation to compare its predictions to real-world, directly physically measured VO_{2max} values in a different health care setting. Concurrently with the prospective validation, a pilot study will be conducted to assess the feasibility, usability, and integration of the web application into clinical workflows.

Study Design

This protocol describes both a retrospective multicenter predictive modeling study and a prospective multicenter observational cohort study.

The study will be developed in accordance with the [22,22] extension, Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis plus Artificial Intelligence extension 2024 statement [23] and the CONSORT (Consolidated Standards of Reporting Trials)-EHEALTH [24]. The CONSORT-AI guideline ensures transparent reporting of artificial intelligence (AI)-based interventions in the trial, including algorithm version, input data handling, and human–AI interaction. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis-AI is used to guide reporting of the ML prediction model's development, validation, and performance. CONSORT-EHEALTH is applied to capture key digital intervention details such as delivery platform, user interaction, and data flow. Further, the study follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials)-AI checklist (Checklist 1) for clinical trial protocols involving AI [25]. Together, these guidelines support clarity, reproducibility, and critical appraisal of our AI-driven digital health evaluation.

Participants and Data Collection

Inclusion Criteria

The inclusion criteria for both the retrospective and prospective components of this study are patients (≥18 y) diagnosed with severe obesity, which corresponds to BMI greater than or equal to 40.0 kg/m^2, or 35.0 - 39.9 kg/m^2 with at least 1 obesity-related comorbidity (eg, type 2 diabetes, hypertension, dyslipidemia, or obstructive sleep apnea). Additionally, participants who completed 1 VO_{2max} test at either baseline, 9, 12, 17, 25 weeks, or at 1 to 2 years were included in the retrospective cohort study.

Exclusion Criteria

Patients will be excluded from the study if they are aged younger than 18 years or have a BMI below 35.0 kg/m^2. Patients who have not completed at least 1 VO_{2max} test at any of the specified assessment points (baseline, wk 9, 12, 17, 25, or at 1 to 2 y) will be ineligible. Further, patients with incomplete or insufficient medical measurements of VO_{2max} data will likewise be excluded.

Retrospective Study: Subgoal 1

In the retrospective study, an accurate estimation of CRF using ML models will be developed, specifically tailored to individuals with obesity. ML models will be trained on datasets combining results from direct physical assessments of VO_{2max} and relevant clinical parameters derived from patients with obesity. The retrospective study will use data collected at Vestfold Hospital Trust (VHT), Norwegian School of Sport Sciences (NIH), and Muritunet Rehabilitation Institution (MRI) to train a ML model.

VHT possesses a dataset comprising more than 700 adults with severe obesity. The data was collected in 3 different research

XSL•FO
RenderX

studies approved by the Regional Committees for Medical and Health Research Ethics South East, Norway.

The first study was a randomized controlled trial (REK ID: 2013/1849, Clinical Trials NCT02311738) [26]. Data from 73 patients seeking treatment at the Department of Endocrinology, Obesity, and Nutrition was collected between January 5, 2015, and June 9, 2017. A total of 243 VO_{2max} tests were performed at baseline and at approximately 9 weeks, 17 weeks, and 25 weeks (Table 1).

The second study was a retrospective cohort study (REK: 2016/1414, ClinicalTrials.gov: NCT03593798) [19]. Data from 180 patients seeking treatment at the Clinic of Physical Medicine and Rehabilitation were collected between November 1, 2013, and January 1, 2017. A total of 442 VO_{2max} tests were performed at baseline and at approximately 12-week and 1-year follow-ups (Table 1).

The third study is an ongoing retrospective cohort study (REK: 2017/88, Clinical Trials: NCT03593148) started in May 2017 and will end in June 2026. Data thus far have been collected from 453 patients seeking treatment at the Clinic of Physical Medicine and Rehabilitation. A total of1400 or more VO_{2max} tests were performed at baseline and approximately 12-week, 14-week, 24-week, 1-year, and 2-year follow-ups (Table 1).

Table . Test time of maximum volume of oxygen.

Test time	Baseline	9 wk	12 wk	17 wk	25 wk	1 y	2 y
Study 1 at VHT[a]	✓[b]	✓		✓	✓		
Study 2 at VHT	✓		✓			✓	
Study 3 at VHT	✓		✓			✓	✓
Study 4 at NIH[c]	✓		✓		✓		
Clinical registry 5 at MRI[d]	✓				✓	✓	

[a]VHT: Vestfold Hospital Trust.

[b]✓: data available at this measuring point.

[c]NIH: Norwegian School of Sport Sciences.

[d]MRI: Muritunet Rehabilitation Institution.

NIH has collected data from 43 patients with obesity through 2 projects.

NIH has collected data from patients with obesity in 2012 and 2013. The projects were part of a larger PhD project aiming to evaluate longitudinal lifestyle changes in obese patients (REK Sør-Øst C-2009/1699). A total of 101 VO_{2max} tests were performed at baseline to 12 and 25 weeks after the intervention (Table 1).

MRI in Møre and Romsdal, Norway, has collected data from more than 750 patients with obesity.

Muritunet has a clinical quality registry with consent for research (Norwegian Data Protection Authority 17/00433 - 2/SBO). The clinical quality registry started on May 16, 2017, and is still ongoing. More than 1300 VO_{2max} tests were performed at baseline to approximately 6- and 12-month follow-up (Table 1).

VHT, NIH, and MRI together provide a comprehensive dataset from more than 1500 adult patients with severe obesity and more than 3600 completed VO_{2max} tests. This diverse data source allows for the development of a robust ML model capable of accurately estimating VO_{2max} in patients with obesity.

Medical Measurements

Brief Summary of All Medical Measurements

The medical measures obtained from studies were maximal oxygen consumption, body weight, age, sex, BMI, height, waist and hip circumferences, fat-free mass and fat mass, diastolic and systolic blood pressure, heart rate maximum, respiratory quotient, resting oxygen saturation, oxygen saturation before test, speed or watt, incline and total time duration on treadmill, resting heart rate, and Borg Scale. Self-reported questionnaire of generic measure of health-related quality of life (Short Form Health Survey, SF-36) [27,28], obesity-specific measure of health-related quality of life [29], and a symptom measure of obesity (weight-related symptom measure of obesity) will also be used [30,31] (Table 2).

XSL•FO
RenderX

Table . Parameters collected from different studies or registries.

Parameters	Study 1 (VHT[a])	Study 2 (VHT)	Study 3 (VHT)	Study 4 (NIH[b])	Clinical registry 5 (MRI[c])
Sex (male/female)	✓[d]	✓	✓	✓	✓
Age (y)	✓	✓	✓	✓	✓
Height (cm)	✓	✓	✓	✓	✓
Body weight (kg)	✓	✓	✓	✓	✓
Maximal volume of oxygen (VO_{2max}, $mL•min^{-1}$)	✓	✓	✓	✓	✓
BMI (kg/m^2)	✓	✓	✓	✓	✓
Waist circumferences (cm)	✓	✓	✓	✓	✓
Hip circumferences (cm)	✓	✓	✓	✓	✓
Fat mass (kg)	✓	✓	✓	✓	✓
Fat-free mass (kg)	✓	✓	✓	✓	✓
Systolic blood pressure (mm Hg)	✓	✓	✓	✓	
Diastolic blood pressure (mm Hg)	✓	✓	✓	✓	
Resting heart rate (bpm)	✓	✓	✓	✓	
Heart rate maximum (bpm)	✓	✓	✓	✓	
Respiratory exchange ratio	✓	✓	✓	✓	
Oxygen saturation before test (%)				✓	
Ventilation maximum ($mL•min^{-1}$)				✓	
Speed treadmill (km/t)	✓	✓	✓		
Watt bike (W)					
Incline treadmill (%)	✓	✓	✓		
Speed or watt ergometer bike (w)		✓	✓		
Total time duration treadmill (min)	✓	✓	✓		
Total time bike (min)		✓			
SF-36[e]	✓		✓		
IWQOL-Lite[f]	✓		✓		
WRSM[g]	✓		✓		
Sociodemographic data	✓	✓	✓	✓	✓
Other relevant health parameters	✓	✓	✓	✓	✓

[a]VHT: Vestfold Hospital Trust.

[b]NIH: Norwegian School of Sport Sciences.

[c]MRI: Muritunet Rehabilitation Institution.

XSL•FO
RenderX

[d]✓: data available at this measuring point.

[e]SF-36: health-related quality of life was measured with the Short Form Health Survey.

[f]IWQOL-Lite: impact of weight on quality of life.

[g]WRSM: weight-related symptom measure of obesity.

Maximal Volume of Oxygen

The VO_{2max} was expressed in absolute values as liters per min (mL/min) and often presented as relative to body weight (mL/kg/min) as CRF [8]. The VO_{2max} was measured by the use of Metalyser, Cortex 2 (Biophysik) (study 1), Jaeger Oxycon Pro ergospirometry test system (JLAB 5.x) (study 2 and 3), and Vyntus CPX metabolic system (Vyaire Medical) (part of study 3 [gen. 3.20] and 4 [gen. 1]). Metalyser cortex 2 has an accuracy of 4% - 5%, Jaeger Oxycon Pro of 3%, and Vyntus CPX 3% compared to mechanical lung. Before each ergo test, the device was calibrated with room air (23 °C) at 20.93% O_2 and 0.03% CO_2 and a certified gas containing 15.80% O_2 and 4.98% CO_2. The calibration of ventilatory volume was conducted with a manual 3-liter syringe (Hans Rudolph, Shawnee, KS, USA). Barometric pressure was calibrated with an electronic barometer (GB3300 Greisinger electronic) or automatically by the system. A face mask (Hans Rudolph V2) with different sizes (petit, XS, S, M, L) was used to collect expired air during the tests. The Metalyser, Cortex 2, and Jaeger Oxycon Pro ergospirometry system was equipped with a mixing chamber, with O_2, and CO_2, and the ventilation was analyzed continuously. Vyntus CPX metabolic was equipped with breath-by-breath open system.

Maximal Volume of Oxygen Test

The test was performed as an individualized incremental treadmill test or modified Balke protocol on a Woodway PPS 55 plus or ELG 5, for patients with walking restrictions (n=8), as an incremental bicycle test (Lode Corival V3, Lode BV).

For the individualized incremental treadmill test at VHT, velocity was increased alternately by 0.5 km/h, or the inclination by 1% every 30 seconds until voluntary exhaustion. The modified Balke protocol at NIH starts at 3 minutes at 4.5 km/h and 4 % inclination, the next level was 4.8 km/h and 6% inclination, and thereafter 2% inclination every minute up to 20%. The modified Balke protocol at MRI starts at 2 minutes at 3.6 km/h and increases 0.6 km/h every 2 minutes until 6 minutes. Then the speed was standing at 5.4 km/h with an inclination of 4%, and thereafter 2% inclination every minute up to 20%. If the patient could bear to continue after the maximal incline, the speed was increased by 0.5/h every minute until exhaustion.

For the bicycle test, brake power was increased by 25 watts every 30 seconds until voluntary exhaustion. The duration of the tests ranged from 4 to 10 minutes.

Two or more of the following criteria must be met for a valid VO_{2max} test; voluntary exhaustion, respiratory exchange ratio greater than or equal to 1.05, heart rate greater than or equal to 95% of heart rate maximum, Borg scale greater than or equal to 17, and a flattening of the VO_2 curve (mL/min) were used to evaluate if VO_{2max} had been achieved. VO_{2max} was calculated as the highest of 3 consecutive 10 seconds measurements (total 30 s) or the highest of 6 consecutive 5 seconds measurements (total 30 s).

Heart Rate Maximum

Heart rate maximum was measured automatically every 5 seconds with Polar WearLink + H7 bluetooth or manually each minute (Polar Electro OY) and set as the highest observed value during the test.

Speed and Incline on Treadmill

Speed (km/h) and incline (%) were set as the maximum reported value during the test on treadmill.

Respiratory Exchange Ratio

Respiratory exchange ratio measures the ratio of carbon dioxide produced (VCO_2) to oxygen consumed (VO_2) during VO_{2max} test. The respiratory exchange ratio was measured with the same equipment as used to measure VO_{2max}.

Body Weight

Body weight (kg) was measured with patients wearing light clothing and no shoes on Scanvaegt DS-530 (study 2 and 3), or with the bioelectrical impedance analyzer Tanita (MC-780 or BC-418), or Inbody 720, (100 - 240 W).

Height

Height (cm) was measured to the nearest 0.1 cm using Soehnle professional stadiometer.

BMI Calculation

BMI was calculated as weight in kilograms divided by height in meters squared (kg/m^2).

Waist Circumference

Waist circumference (cm) was measured to the nearest 0.1 cm using a measuring tape placed horizontally at the midpoint between the lower margin of the lowest rib and the top of the iliac crest, with the abdomen relaxed.

Hip Circumference

Hip circumference (cm) was measured using a measuring tape placed around the widest part of the hips, with participants standing upright and with their feet together.

Fat Mass and Fat-Free Mass

Body composition as fat mass (kg) and fat-free mass (kg) was measured with patients wearing light clothing and no shoes or socks using the bioelectrical impedance analyzer Tanita BC-418 or Inbody 720, (100 - 240 W).

Blood Pressure

Blood pressure (mm Hg) was measured using an automatic sphygmomanometer (Dinamap V100). The person was seated in a relaxed position with the arm resting on the table. Three

XSL•FO
RenderX

measurements were collected, and the average of these 3 was used.

Resting Heart Rate

The resting heart rate was measured at the same time as the blood pressure, and the average of 3 measurements was used.

Self-Reported Measurements

Health-Related Quality of Life

The generic measure of health-related quality of life was measured with the Short Form Health Survey (SF-36, version 2.0), a 36-item instrument designed to assess generic health-related quality of life. Of these, 35 items are used to generate scores across 8 dimensions: physical functioning, role limitations due to physical problems, bodily pain, general health, vitality, social functioning, role limitations due to emotional problems, and mental health [27,28]. Each dimension is scored on a scale from 0 to 100, with higher scores indicating better health-related quality of life. The SF-36 has been translated into Norwegian [32] and has shown strong psychometric properties across various medical conditions [32-35], including among patients with severe obesity [36].

Obesity-Specific Measure of Health-Related Quality of Life

The obesity-specific measure of impact of weight on quality of life was measured with a 31-item questionnaire designed to assess weight-related quality of life [29]. Four domain scores were evaluated: physical function, self-esteem, sexual life, and public distress. Each domain is scored on a scale from 0 to 100, with lower scores reflecting greater impairment.

Symptom Measures of Obesity

The weight-related symptom measure of obesity was measured with a 20-item for the presence and distress of 20 weight-related symptoms [30,31]. The symptoms are scored on a 6-point Likert scale. The scale ranges from 0 (does not bother at all) to 6 (bothers a very great deal) with higher scores indicating worse symptom distress.

Web Application: Subgoal 2

The web application design will be developed by Natalia Kaur together with the "AI and VO_{2max}" group. An application programming interface that serves as the interface facilitating communication between a data system and an external service or system will be used in the web application. It enables the exchange of queries, whereby the external system processes incoming requests and provides a corresponding response. Figma will be used to develop a user-centric interface prototype in an agile manner. The finalized design will then be implemented as a .NET application and hosted on Microsoft Azure, a cloud-based platform that supports seamless deployment and scalable web applications. Our Python-based ML model will be hosted on a web server and made accessible to the web application through an application programming interface, enabling seamless communication with Microsoft Azure and our front-end application. Practitioners and patients will actively contribute to the design, development, testing, and feedback of the web application through an agile, interdisciplinary development team assembled for the project. This collaborative approach will ensure the implementation of a user-friendly web application (subgoal 2), continuing the focus on the user interface process.

Prospective Study: Subgoal 3

In the prospective validation, the results of the ML-based CRF estimation model in the web application will be compared with direct physical measurements of VO_{2max} obtained using cardiopulmonary exercise testing equipment for testing from different institutions. In each project participating institution, a principal investigator with expertise in physical measurements of direct VO_{2max} will be responsible for the coordination of the prospective study, together with the project leader (JB).

Patients who meet the eligibility criteria of overweight and obese will be given both verbal and written information of the prospective study. The participant will be given up to 1 week to consider their participation. If they agree to be included in the prospective study, written informed consent will be obtained from all patients.

Once enrolled, participants will undergo direct physical measurements of VO_{2max} with medical equipment, and simultaneously, the principal investigator will collect all necessary input parameters required for the ML-based model. The ML model will generate an estimated CRF with the web application, which will be recorded alongside the measured VO_{2max} result. All data, including participant details, model inputs, model output, and corresponding direct physical measurements of VO_{2max}, will be entered into a secure, study-specific database.

Testing and Iterative Improvement

The web application will be tested for user experience in both primary and secondary health care settings, focusing on usefulness, usability, and engagement to support iterative improvements. A study-specific questionnaire will be used (eg, technology acceptance model or system usability scale). The diverse project partners will actively contribute to its implementation across these health care sectors.

Model Management and Monitoring: Subgoal 4

To develop a robust ML model, we will continuously retrain it using new VO_{2max}, measurements, along with anthropometric and demographic data from the health care clinics. Each retrained model will be saved (version control) and tested immediately after retraining to assess the performance and control for unintentional performance drop. The history of model retraining and performance will be maintained and made accessible to system administrators via a dashboard.

Implementation: Subgoal 5

The ML model estimation through the web application will be implemented as a clinical tool, and thereby CRF may be used as a diagnostic and prognostic health indicator. The clinical tool will support health care professionals in hospitals, rehabilitation centers, primary care services, general practice, and in other settings. This clinical tool will enable faster, more targeted, and personalized treatment while reducing the workload on health

XSL•FO
RenderX

care services by eliminating the need for direct physical measurement of VO_{2max}.

Statistical Analysis

User Statistics

The statistical analysis will be performed using Python (3.12.12; Python Software Foundation). Categorical variables will be summarized as frequencies and percentages, and their comparisons will be conducted using the chi-square test. For continuous variables, analyses will be conducted using independent samples *t* tests or Mann-Whitney *U* tests, with results presented as mean (SD) for normally distributed data or median (IQR) for nonnormally distributed data. To estimate the variance and 95% CI of the model's performance on the test set, bootstrap resampling will be performed.

Sample Size Calculation

In traditional linear statistics, a common rule of thumb is that at least 10 - 20 observations per predictor are required to ensure stable model estimation [37]. However, in our study, we plan to apply more advanced ML models (such as random forest regression), which are likely to require substantially larger sample sizes to avoid overfitting and to achieve reliable generalization. In addition, we expect a considerable amount of missing data across the participating centers, which further increases the need for a larger dataset. For these reasons, we plan to include all available data from the study centers. This approach allows for a fair comparison between simpler regression models and more complex ML models, while also ensuring sufficient statistical power despite missing data and model complexity.

For the prospective study, we performed a sample size calculation using 1-sample mean test, shown in Equation 1, where Zβ represents the standard normal deviate corresponding to the desired statistical power. We selected 80% power, which is commonly used in clinical research, resulting in Zβ=0.84. The SD of VO_{2max} was used. In our population, estimated using data from 2 of the databases included in this study, this value was found to be 23% of the mean.

The parameter σ represents the acceptable clinical error and was set to 3%, based on the established measurement error of the gold standard method. Zα corresponds to the selected significance level. A 1-sided test was selected under the assumption that the model would outperform random prediction given the population distribution. Figure 1 shows the required sample size for different significance levels (α), while keeping the other parameters fixed as defined above.

Figure 1. Sample size (n) for different significance levels α.

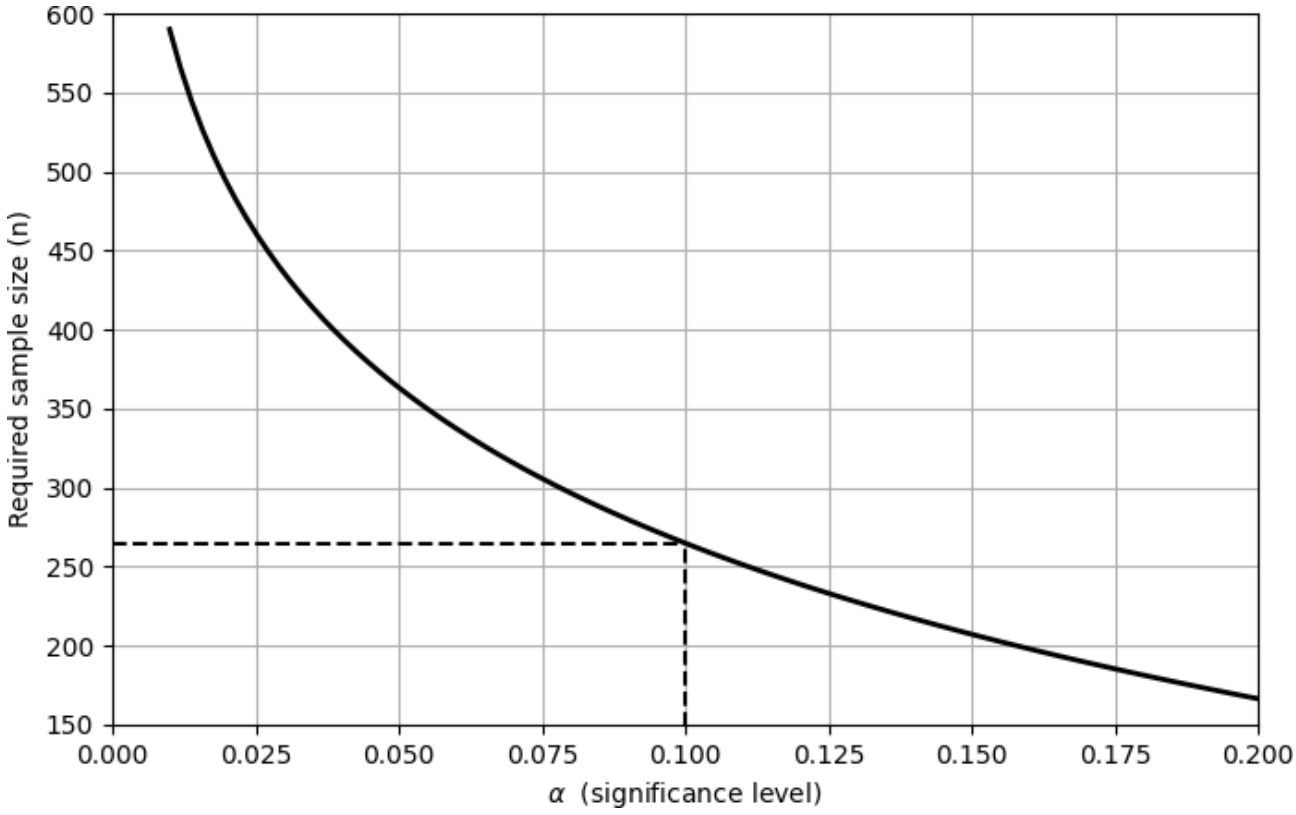

The selected significance level for our study is α=.10, which corresponds to a 1-sided 90% confidence level. Under these assumptions, the required sample size is 265 patients, enabling us to detect an improvement that exceeds the acceptable error threshold with 80% power.

(Eq.1)n=Zα+ZβσSD2

Model Architecture

Candidate algorithms were chosen to represent a broad range of linear, kernel-based, and tree-based regression approaches. These approaches included linear models (ordinary least squares linear regression, ridge regression, least absolute shrinkage and selection operator, and elastic net regression), Bayesian models (Bayesian ridge regression and automatic relevance

XSL•FO
RenderX

determination regression), support vector regression models (linear support vector regression and nu-support vector regression), and tree-based ensemble models (decision tree regressor, extremely randomized trees regression, random forest regression, and histogram-based gradient boosting regression).

Data Preprocessing

A sequential data preprocessing pipeline will be implemented, consisting of three primary components: (1) missing value imputation, (2) feature standardization, and (3) input to regression model. All preprocessing steps will be incorporated into a unified pipeline to ensure robustness and reproducibility.

To address incomplete data in model development and testing, missing values will be imputed using the k-nearest neighbor imputation method. The k-nearest neighbor algorithm will identify the set of k-nearest neighbors based on Euclidean distance metrics computed over the available (observed) feature values and impute missing values using the mean of these neighbors.

To ensure that features contribute proportionally to model training and to improve model convergence, feature standardization will be applied. Continuous predictors will be rescaled using *z*-score normalization. The estimated scaling parameters, trained on the training folds or the development set, will then be applied to the corresponding validation and test sets.

Model Development and Evaluation

Model development will be performed with Python (3.12.12) and packages such as scikit-learn (version 1.6.0; scikit-learn developers). Demographic and anthropometric parameters (Table 2) from the datasets obtained from VHT, MRI, and the NIH will be divided into a development set (80%) and a test set (20%), using a patient-level grouped split strategy to ensure that all measurements from a given patient are included in only 1 of the 2 sets. This approach will be used to prevent data leakage between the development and test sets.

The development set will be used to perform a structured hyperparameter search for the optimal configuration of the model. The final model will be selected based on cross-validation, and the best-performing model across the k-folds on the development set will be used for final testing on the test set. As with the development–test split, the *k*-fold cross-validation will be grouped at the patient level to ensure that all measurements from the same patient are included in only 1 fold, thereby preventing data leakage.

During the development phase, we will also experiment with different stratification strategies related to the center of origin of the dataset, to identify potential bias, including systematic differences arising from variations in measurement protocols between centers.

The performance on the test set will be reported in terms of metrics such as *R*-squared score, root mean squared error, mean absolute error and mean absolute percentage error, standard error of estimate (SEE), and SEE given in percentage. In addition to the CRF estimation model, a pipeline of preprocessing steps will be constructed, doing necessary things such as feature scaling and imputation. The ML model with the best performance (subgoal 1) will be implemented in the web application (subgoal 2) and further used in the prospective study (subgoal 4).

Continuous Monitoring and Model Updating

It is likely that our models, particularly those based on a larger number of input variables, do not reach their full performance plateau when trained on the described datasets. Consequently, retraining the model using data collected after system implementation is expected to further improve predictive performance. However, this process requires careful oversight, as the inclusion of erroneous or low-quality data could degrade rather than enhance model performance. To mitigate this risk, only observations and estimations that have a corresponding directly measured VO_{2max} value within plus or minus 7 days of the estimated value will be used for retraining.

To monitor and control data drift, we will retain 20% of all incoming data in a holdout monitoring set, in addition to the original test set from this study. This subset will be stratified by center's location to ensure geographic and operational diversity. This design allows us to detect both temporal drift and center-specific drift. Drift will be quantified using multiple performance metrics, including root mean squared error, mean absolute percentage error, SEE, and *R*-squared score.

Furthermore, we will cluster incoming observations and identify potential out-of-distribution samples. These cases will be flagged and manually reviewed before inclusion in either the training or evaluation datasets. Model performance will also be continuously assessed across clinically and demographically relevant subgroups, including sex, age groups, and BMI categories, to detect potential bias or subgroup-specific degradation in performance.

All monitoring procedures, performance metrics, and visualizations will be integrated into a dedicated dashboard accessible to system administrators.

Model retraining may present regulatory challenges in relation to the European Union Medical Device Regulation (MDR) and CE marking, as substantive changes to the model may require recertification. For this reason, retraining will initially be limited to scenarios where the application is marketed as a health or wellness application rather than a regulated clinical decision-support tool. The initial release will therefore be positioned as a health application, with the longer-term objective of potential reclassification as a clinical tool, subject to regulatory approval.

Ethical Considerations

First, this project relies on data from various research studies and a clinical registry, both of which have been approved for this purpose by the Regional Committees for Medical and Health Research Ethics South East, Norway (REK ID: 2013/1849, 2016/1414, 2017/88, and 2009/1699 [38]) or the Norwegian Data Protection Authority (ID: 17/00433 - 2/SBO [39]). Further, 2 of the studies are published in international research journals [19,26].

XSL•FO
RenderX

Second, all data from the included studies and registers are either pseudonymized or anonymized, and the welfare and integrity of participants will be safeguarded through the careful selection of variables that minimize the risk of reidentification. The variables do not include easily recognizable external background information, such as place of residence or occupation, which could increase the likelihood of retrospective identification. However, variables such as height, weight, and VO_{2max} may, in rare cases, allow for reidentification at extreme values. To address this, it may be necessary to exclude extreme values (outliers) within these categories.

Third, written informed consent was obtained from all patients, with the intention of using the data for research.

Fourth, a "motivated intruder" test will be conducted to assess the potential for data reidentification. While it is impossible to eliminate the risk of reidentification with absolute certainty, these measures are designed to significantly mitigate such risks. Furthermore, the project uses systems that have undergone rigorous risk assessment and are approved for handling sensitive data, ensuring adherence to the highest standards of data security and privacy.

Fifth, studies 1, 2, and 4 are fully anonymized, with the linking keys to patients' identities permanently deleted. As a result, the Norwegian Agency for Shared Services in Education and Research granted approval for the use of these data in the project. Studies 3 and 5 are pseudonymized; however, the linking keys to patients' identities have not been deleted. Consequently, the project applied to the Norwegian Directorate of Health for an exemption from the requirement to obtain patient consent for the use of the data. For Study 4, broad consent has been obtained from the patients, allowing the inclusion of these data in the project.

Sixth, to avoid recognition in the data material from the southeast region of Norway (VHT and NIH), this study included patients from the central region of Norway, Møre and Romsdal (MRI).

Seventh, to ensure rigorous adherence to ethical standards, the project applied to the Regional Committees for Medical and Health Research Ethics South East, Norway (REK ID 741918). They determined that the project falls outside the scope of the Health Research Act, cf Section II. However, clinical trials involving medical devices may fall under the purview of the European Union MDR 2017/745. The MDR explicitly states that applications, software, and devices incorporating AI can be classified as medical devices. Consequently, it may be necessary to apply following the project's completion to facilitate the implementation of the web application as a certified medical device at the end of this project.

This project is registered at ClinicalTrials.gov (identifier: NCT07011108). VHT is the coordinating center. Responsibility for data management lies with each center that is the data controller for its own data set. VHT will have access to the final data set. No financial compensation, incentives, or other forms of remuneration will be provided to participants.

Results

The project was awarded salary funding in two rounds from Vestfold Hospital (October and December 2023) and additional salary funding from the Research Council of Norway (October 2024). In addition, the project received allocated supervision hours from InnoMed Norway (April 2024). As of submission (February 2026), the dataset includes 1279 adults with severe obesity and 2623 VO_2 max tests collected before, during, and after lifestyle interventions conducted between 2013 and 2025.

For the retrospective component of the study, ethical approval has been obtained to use 2 of the 5 databases for initial model development. Ethics review for the remaining 3 databases is currently pending. Upon completion of the initial model development phase, ethical approval will be sought to conduct the prospective investigation.

Data analysis is ongoing. A web-based application has been developed, and the first and second rounds of formative user testing have been conducted with adult patients with overweight and obesity and health care personnel. The first scientific publication of the clinically relevant machine learning model is expected in 2026, and the overall project is expected to be completed in 2028.

Discussion

Principal Findings

Patients with obesity have low relative VO_{2max} compared to normal-weight people. Therefore, patients with obesity are outside the range of validity of the statistical models trained on normal-weight people [13,19,20], which results in an imprecise estimate of CRF for patients with obesity using statistical models trained on normal-weight people. To address this problem, a retrospective study will be conducted, using anonymized datasets from 4 research studies and 1 clinical quality registry.

Furthermore, to address the—often seen—performance drop by ML models on data gathered in a different context than training data, this study plans to validate the model from the previous step (retrospective part) prospectively.

Strengths and Limitations

The ML-based estimation will be implemented as a clinical tool in the form of a web application, designed to assist health care professionals across various settings, including hospitals, rehabilitation centers, primary care services, and general practice. This tool aims to facilitate faster, more precise, and personalized treatment while alleviating the burden on health care services by obviating the need for direct physical measurements of VO_{2max}. Furthermore, the integration of advanced technologies and optimized workflows has the potential to enhance efficiency, allowing health care professionals to allocate more time to patient-centered tasks or mitigate waiting times. Additionally, the web application seeks to empower both patients and the general population by fostering greater autonomy in managing their health, enhancing health literacy, and promoting engagement in preventive health measures.

XSL•FO
RenderX

Stakeholder Involvement

Further, the primary health care services play a pivotal role in strengthening health competencies at the municipal level and alleviating the demand for specialized health care services. Therefore, to ensure the successful implementation of this innovation and research initiative, the primary health care sector is actively engaged in its development, contributing to its relevance, feasibility, and integration into existing health care frameworks. Furthermore, the active involvement of patients throughout the entire study is considered to be essential for its successful development and implementation.

Conclusion

To conclude, the project members anticipate that using longitudinal, real-world clinical data from a uniquely characterized population with obesity will enable the development and validation of a clinically relevant ML model for estimating CRF that provides a more practical alternative to current assessment methods, thereby reducing time, costs, and procedural burden in clinical practice. This may enhance accessibility and reinforce CRF as a valuable diagnostic and prognostic assessment of cardiovascular health for individuals with obesity. Such a predictive tool may also facilitate more frequent CRF assessments, enabling more detailed monitoring of changes in CRF over time. The broader aim of this project is to achieve national and international significance, contributing to improved patient outcomes and supporting health care professionals globally.

To our knowledge, VHT, MRI, and NIH are likely the only institutions in Norway that systematically assess direct physical measurements of VO_{2max} as part of both research projects and clinical treatments for obese patients. This gives us a unique database for the development of ML models for CRF estimation for this specific patient population. To our knowledge, this study will be the first attempt to estimate CRF in patients with obesity, based on clinical parameters, using ML. The project members expect that the project will hold significant societal value by contributing to improved public health and advancing knowledge related to the management of severe obesity. Obesity represents a major public health challenge with substantial individual, societal, and economic consequences. By investigating physiological responses and cardiorespiratory fitness outcomes in patients with severe obesity, this study may provide evidence that informs clinical practice, supports individualized treatment strategies, and contributes to more effective health care resource allocation.

Communication

The project will be disseminated through professional hospital channels and to relevant stakeholders nationally and internationally. The Norwegian Association for People with Obesity will support accessible communication to the patient group through its established channels. Results will be disseminated through national and international media, presented at major scientific conferences, and published in peer-reviewed journals.

Acknowledgments

The project group thanks Vestfold Hospital Trust (VHT), Health South-East of Norway, Research Council of Norway, and InnoMed of Norway for funding the study.

In addition, sincere thanks are extended to the coauthors and the remaining members in the "AI and VO2max" group: Ida Husjord, VHT, Department of eHealth, Tønsberg, Norway; Einar Sudin, VHT, Department of eHealth, Tønsberg, Norway; Joacim Skarsholt Eldre, VHT, Department of Technology, Tønsberg, Norway; Tone Irene Baasland, Larvik Municipality, Larvikshjelpa, Larvik, Norway; Ingjerd Stulen, VHT, Department of Neurology and Rehabilitation, Stavern, Norway; Kathrine Vadøy Antonsen, the National Association for Women's Pelvic Girdle Health, Oslo, Norway; Mari-Mette Graff, the National Association for People with Obesity, Tromsø, Norway; Natalia Kaur, University of South-Eastern Norway, Horten, Norway.

Finally, the authors thank the Clinical Artificial Intelligence Research Group at VHT for the academic discussions.

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to generative artificial intelligence (GAI) tools under full human supervision: proofreading and editing, translation, and reformatting. The GAI tool used was ChatGPT 5.0. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by: collective responsibility.

Additional note: ChatGPT (OpenAI and GPT-5) was used to assist with language editing and improving clarity during manuscript preparation. The authors reviewed and revised all content and take full responsibility for the final version of the manuscript.

Funding

The project is partially funded by VHT and the Research Council of Norway.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available due to the inclusion of sensitive patient data and applicable data protection regulations. Data may be available from the corresponding author upon reasonable request

XSL•FO
RenderX

and with appropriate ethical and institutional approvals. The machine learning model developed in this study will be made publicly available at a later stage following validation and regulatory considerations.

Conflicts of Interest

MWAC is shareholder of Nico.lab International Ltd. All other authors declare no conflicts of interest.

Checklist 1
SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials)-artificial intelligence checklist.
[PDF File, 865 KB - resprot_v15i1e85069_app1.pdf]

References

1. Obesity. World Health Organization. 2021. URL: https://www.who.int/news-room/facts-in-pictures/detail/6-facts-on-obesity [accessed 2026-01-31]
2. Overvekt og fedme i norge: omfang, utvikling og samfunnskostnader [Report in Norwegian]. : Menon Economics; 2019 URL: https://menon.no/uploads/images/2019-09-Overvekt-og-fedme-i-Norge.pdf [accessed 2026-01-31]
3. Laukkanen JA, Lakka TA, Rauramaa R, et al. Cardiovascular fitness as a predictor of mortality in men. Arch Intern Med 2001 Mar 26;161(6):825-831. [doi: 10.1001/archinte.161.6.825] [Medline: 11268224]
4. Kodama S, Saito K, Tanaka S, et al. Cardiorespiratory fitness as a quantitative predictor of all-cause mortality and cardiovascular events in healthy men and women: a meta-analysis. JAMA 2009 May 20;301(19):2024-2035. [doi: 10.1001/jama.2009.681] [Medline: 19454641]
5. Wing RR, Jakicic J, Neiberg R, et al. Fitness, fatness, and cardiovascular risk factors in type 2 diabetes: look ahead study. Med Sci Sports Exerc 2007 Dec;39(12):2107-2116. [doi: 10.1249/mss.0b013e31815614cb] [Medline: 18046181]
6. Lee D, Artero EG, Sui X, Blair SN. Mortality trends in the general population: the importance of cardiorespiratory fitness. J Psychopharmacol 2010 Nov;24(4 Suppl):27-35. [doi: 10.1177/1359786810382057] [Medline: 20923918]
7. Kokkinos P, Faselis C, Franklin B, et al. Cardiorespiratory fitness, body mass index and heart failure incidence. Eur J Heart Fail 2019 Apr;21(4):436-444. [doi: 10.1002/ejhf.1433] [Medline: 30779281]
8. Kent M. The Oxford Dictionary of Sports Science & Medicine, 3rd edition: Oxford University Press; 2006:612 URL: https://www.oxfordreference.com/display/10.1093/acref/9780198568506.001.0001/acref-9780198568506 [accessed 2026-01-31]
9. Wammer F, Haberberger A, Linge AD, Myklebust T, Vemøy S, Hoff DAL. Lifestyle modification for weight loss: effects on cardiorespiratory capacity in patients with class II and class III obesity. Obes Sci Pract 2022 Feb;8(1):45-55. [doi: 10.1002/osp4.544] [Medline: 35127121]
10. Gaesser GA, Angadi SS. Obesity treatment: weight loss versus increasing fitness and physical activity for reducing health risks. iScience 2021 Oct 22;24(10):102995. [doi: 10.1016/j.isci.2021.102995] [Medline: 34755078]
11. Weeldreyer NR, De Guzman JC, Paterson C, Allen JD, Gaesser GA, Angadi SS. Cardiorespiratory fitness, body mass index and mortality: a systematic review and meta-analysis. Br J Sports Med 2025 Feb 20;59(5):339-346. [doi: 10.1136/bjsports-2024-108748] [Medline: 39537313]
12. McArdle WD, Katch FI, Katch VL. Exercise Physiology: Nutrition, Energy, and Human Performance, 7th edition: Lippincott Williams & Wilkins; 2010. URL: https://library.wur.nl/WebQuery/titel/1930622 [accessed 2026-01-31]
13. Ashfaq A, Cronin N, Müller P. Recent advances in machine learning for maximal oxygen uptake (VO2 max) prediction: a review. Inform Med Unlocked 2022;28:100863. [doi: 10.1016/j.imu.2022.100863]
14. Petelczyc M, Kotlewski M, Bruhn S, Weippert M. Maximal oxygen uptake prediction from submaximal bicycle ergometry using a differential model. Sci Rep 2023 Jul 12;13(1):11289. [doi: 10.1038/s41598-023-38089-7] [Medline: 37438405]
15. Cink RE, Thomas TR. Validity of the Astrand-Ryhming nomogram for predicting maximal oxygen intake. Br J Sports Med 1981 Sep;15(3):182-185. [doi: 10.1136/bjsm.15.3.182] [Medline: 7272663]
16. Helgerud J, Haglo H, Hoff J. Prediction of VO2max from submaximal exercise using the smartphone application Myworkout GO: validation study of a digital health method. JMIR Cardio 2022 Aug 4;6(2):e38570. [doi: 10.2196/38570] [Medline: 35925653]
17. Nes BM, Janszky I, Vatten LJ, Nilsen TIL, Aspenes ST, Wisløff U. Estimating V·O 2peak from a nonexercise prediction model: the HUNT study, Norway. Med Sci Sports Exerc 2011 Nov;43(11):2024-2030. [doi: 10.1249/MSS.0b013e31821d3f6f] [Medline: 21502897]
18. Alzamer H, Abuhmed T, Hamad K. A short review on the machine learning-guided oxygen uptake prediction for sport science applications. Electronics 2021;10(16):1956. [doi: 10.3390/electronics10161956]
19. Berge J, Støren Ø, Hertel JK, Gjevestad E, Småstuen MC, Hjelmesæth J. Associations between cardiorespiratory fitness and weight loss in patients with severe obesity undergoing an intensive lifestyle intervention program: retrospective cohort study. BMC Endocr Disord 2019 Jul 1;19(1):69. [doi: 10.1186/s12902-019-0394-z] [Medline: 31262301]
20. Nevill AM, Cooke CB. The dangers of estimating V˙O2max using linear, nonexercise prediction models. Med Sci Sports Exerc 2017 May;49(5):1036-1042. [doi: 10.1249/MSS.0000000000001178] [Medline: 27922463]

XSL•FO
RenderX

21. Rules for trustworthy artificial intelligence in the EU. EUR-Lex – Access to European Union law. 2024. URL: https://eur-lex.europa.eu/legal-content/EN/LSU/?uri=CELEX:32024R1689 [accessed 2026-01-31]
22. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Lancet Digit Health 2020 Oct;2(10):e537-e548. [doi: 10.1016/S2589-7500(20)30218-1] [Medline: 33328048]
23. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
24. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126. [doi: 10.2196/jmir.1923] [Medline: 22209829]
25. Rivera SC, Liu X, Chan AW, Denniston AK, Calvert MJ, SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. BMJ 2020 Sep 9;370:m3210. [doi: 10.1136/bmj.m3210] [Medline: 32907797]
26. Berge J, Hjelmesaeth J, Hertel JK, et al. Effect of aerobic exercise intensity on energy expenditure and weight loss in severe obesity: a randomized controlled trial. Obesity (Silver Spring) 2021 Feb;29(2):359-369. [doi: 10.1002/oby.23078] [Medline: 33491314]
27. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [doi: 10.1097/00005650-199206000-00002] [Medline: 1593914]
28. Ware JE Jr. SF-36 health survey update. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3130-3139. [doi: 10.1097/00007632-200012150-00008] [Medline: 11124729]
29. Kolotkin RL, Crosby RD, Kosloski KD, Williams GR. Development of a brief measure to assess quality of life in obesity. Obes Res 2001 Feb;9(2):102-111. [doi: 10.1038/oby.2001.13] [Medline: 11316344]
30. Patrick DL, Bushnell DM, Rothman M. Performance of two self-report measures for evaluating obesity and weight loss. Obes Res 2004 Jan;12(1):48-57. [doi: 10.1038/oby.2004.8] [Medline: 14742842]
31. Niero M, Martin M, Finger T, et al. A new approach to multicultural item generation in the development of two obesity-specific measures: the obesity and weight loss quality of life (OWLQOL) questionnaire and the weight-related symptom measure (WRSM). Clin Ther 2002 Apr;24(4):690-700. [doi: 10.1016/s0149-2918(02)85144-x] [Medline: 12017412]
32. Loge JH, Kaasa S, Hjermstad MJ, Kvien TK. Translation and performance of the Norwegian SF-36 Health Survey in patients with rheumatoid arthritis. I. Data quality, scaling assumptions, reliability, and construct validity. J Clin Epidemiol 1998 Nov;51(11):1069-1076. [doi: 10.1016/s0895-4356(98)00098-5] [Medline: 9817124]
33. Huppertz-Hauss G, Lie Høivik M, Jelsness-Jørgensen LP, et al. Health-related quality of life in patients with inflammatory bowel disease 20 years after diagnosis: results from the IBSEN study. Inflamm Bowel Dis 2016 Jul;22(7):1679-1687. [doi: 10.1097/MIB.0000000000000806] [Medline: 27206016]
34. Garratt AM, Stavem K. Measurement properties and normative data for the Norwegian SF-36: results from a general population survey. Health Qual Life Outcomes 2017 Mar 14;15(1):51. [doi: 10.1186/s12955-017-0625-9] [Medline: 28292292]
35. Loge JH, Kaasa S. Short form 36 (SF-36) health survey: normative data from the general Norwegian population. Scand J Soc Med 1998 Dec;26(4):250-258. [Medline: 9868748]
36. Karlsen TI, Tveitå EK, Natvig GK, Tonstad S, Hjelmesæth J. Validity of the SF-36 in patients with morbid obesity. Obes Facts 2011;4(5):346-351. [doi: 10.1159/000333406] [Medline: 22166753]
37. Riley RD, Snell KI, Ensor J, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med 2019 Mar 30;38(7):1276-1296. [doi: 10.1002/sim.7992] [Medline: 30357870]
38. REK. URL: https://rekportalen.no/en/ [accessed 2023-02-18]
39. Datatilsynet. URL: https://www.datatilsynet.no/en/ [accessed 2026-02-18]

Abbreviation

AI: artificial intelligence
CONSORT: Consolidated Standards of Reporting Trials
CRF: cardiorespiratory fitness
MDR: Medical Device Regulation
ML: machine learning
MRI: Muritunet Rehabilitation Institution
NIH: Norwegian School of Sport Sciences
SEE: standard error of estimate
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
VHT: Vestfold Hospital Trust
VO_{2max}: maximum volume of oxygen (mL/min)

XSL•FO
RenderX

Edited by J Sarvestan; submitted 04.Nov.2025; peer-reviewed by Z Xu; revised version received 16.Dec.2025; accepted 18.Dec.2025; published 02.Mar.2026.

Please cite as:
Berge J, Nunavath V, Asbjørnsen RA, Borgeraas H, Hertel JK, Linge AD, Groote I, Stensrud T, Gjevestad E, Mathisen L, Caan MWA, Paulson M, Støa EM, Johansen JM, AI and VO2max Group, Singstad BJ
Machine Learning for Estimating Cardiorespiratory Fitness in Patients With Obesity: Protocol for a Retrospective and Prospective Multicenter Cohort Study
JMIR Res Protoc 2026;15:e85069
URL: https://www.researchprotocols.org/2026/1/e85069
doi:10.2196/85069

© Jarle Berge, Vimala Nunavath, Rikke Aune Asbjørnsen, Heidi Borgeraas, Jens Kristoffer Hertel, Anita Dyb Linge, Inge Groote, Trine Stensrud, Espen Gjevestad, Linda Mathisen, Matthan WA Caan, Martin Paulson, Eva Maria Støa, Jan - Michael Johansen, AI and VO2max Group, Bjørn-Jostein Singstad. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assistive Robotic Arm to Support Activities of Daily Living in Individuals With Tetraplegia: Protocol for a Real-World Convergent Parallel Mixed Methods Feasibility Study

Vera Fosbrooke[1], MSc; Aline Christen[2], MSc; Barbara Catherine Wortmann[1], BSc; Iris Theodora Maria de Boer[1], BSc; Raphael Rätz[2], PhD; Julian Frederik Rösch[2], MSc; Gabriel Gruener[2], PhD; Anja M Raab[1], PhD

[1]Department of Health, School of Health Professions, Bern University of Applied Sciences, Murtenstrasse 10, Bern, Switzerland

[2]Department of Engineering and Information Technology, School of Engineering and Computer Science, Bern University of Applied Sciences, Biel, Switzerland

Corresponding Author:
Anja M Raab, PhD
Department of Health, School of Health Professions, Bern University of Applied Sciences, Murtenstrasse 10, Bern, Switzerland

Abstract

Background: Tetraplegia, often resulting from cervical spinal cord injury, may lead to significant motor and sensory loss, severely impacting independence and quality of life (QoL). Assistive technologies, such as wheelchair-mounted robotic arms (WMRAs), offer potential to enhance autonomy in daily living. However, adoption remains limited due to high costs, complex controls, and insufficient end user involvement. Robust evidence on their real-world effectiveness, particularly post hospitalization, is still lacking.

Objective: This study aims to explore the feasibility, usability, and user experience of a WMRA for individuals with tetraplegia in real-life posthospitalization settings. It aims to evaluate its support in activities of daily living and conduct a preliminary health economic analysis comparing cost-effectiveness and QoL outcomes with standard care.

Methods: This mixed methods feasibility study will be conducted in posthospitalization settings in Switzerland. Up to 15 participants with upper limb impairments (SCI C0–Th1, AIS A–D) using powered wheelchairs will be recruited. They will use the robotic arm for 6 consecutive days. An equal number of participants will be recruited for the economic analysis group. Quantitative data will be collected at baseline and postintervention via standardized questionnaires (Post-Study System Usability Questionnaire, National Aeronautics and Space Administration Task Load Index, European Quality of Life 5-Dimension 5-Level questionnaire, Visual Analogue Scale, adapted version of the Canadian Occupational Performance Measure, and Client Socio-Demographic and Service Receipt Inventory-European Version), while qualitative feedback will be gathered through an informal questionnaire and semistructured interviews. Feasibility will be assessed through task performance and health economic analysis. The latter will include quality-adjusted life years, which quantify quality and length of life, and modeling the Incremental Cost-Effectiveness Ratio, which compares the cost-effectiveness of the intervention based on cost per quality-adjusted life years gained. Findings will be integrated in line with the convergent parallel mixed methods design.

Results: Recruitment began in April 2025 and is ongoing as planned; full feasibility, usability, and economic results will be reported upon study completion. We expect the robotic system to reduce caregiver time and associated costs, while enhancing autonomy, QoL, and mental well-being. Potential technical and recruitment challenges have been identified and mitigation strategies planned. Evaluating real-life use of a WMRAs, this study could support the wider adoption of assistive robotic technologies.

Conclusions: This research offers key insights into the feasibility, usability, and economic value of robotic assistance for individuals with tetraplegia and will help inform future development and scale-up studies.

International Registered Report Identifier (IRRID): DERR1-10.2196/78339

(JMIR Res Protoc 2026;15:e78339) doi:10.2196/78339

KEYWORDS

WMRAs; robotic arm; tetraplegia; usability; health economic analysis; wheelchair-mounted robotic arms

XSL•FO
RenderX

Introduction

Tetraplegia is a severe form of spinal cord injury (SCI) resulting from damage to the cervical spine, leading to partial or complete loss of motor and/or sensory function in all 4 extremities and the trunk [1,2]. Globally, SCI affects approximately 15.4 million persons [3], with an annual incidence of 250,000 to 500,000 new cases according to the World Health Organization [2]. Individuals with tetraplegia face extensive challenges in daily life, including limitations in mobility, personal care, and social participation [2,4]. These restrictions significantly reduce autonomy and quality of life (QoL), while also placing a substantial economic burden on both affected individuals and health care systems. According to Pacheco Barzallo et al [5], it has been estimated that persons with SCI in Switzerland use health care services 11 times more (including physiotherapists, nursing services, general practitioners, and specialists) than the healthy population and 4 times more than persons with other chronic health conditions. Moreover, caregivers, especially family members providing unpaid care, often experience heightened psychological stress and are at increased risk of developing mental health conditions [4,6].

Assistive technologies (ATs) play a crucial role in mitigating the effects of physical impairments by enhancing independence and enabling participation in activities of daily living (ADLs) [3,7,8]. AT encompasses a wide range of tools, from adaptive cutlery to advanced robotic systems. Individuals with tetraplegia particularly benefit from wheelchair-mounted robotic arms (WMRAs), which enhance care and promote independence across different areas [9,10]. Examples include the Functional Robot with Dexterous Arm and User-Friendly Interface for Disabled People (FRIEND) system, a wheelchair-mounted robotic manipulator designed to assist users with tetraplegia in tasks, such as drinking and eating [11], or the Jaco 2 robotic arm (Kinova Inc), which has been applied in various contexts, such as adaptive feeding systems [12]. However, their adoption into daily life remains limited. Most devices are still in research or prototype phases, tested primarily in controlled experimental settings with participants without disabilities [7]. The high costs, lack of personalization, and the need for end users to be heavily involved in the development process contribute to the low acceptance and small market transfer of these devices [3,4,13]. The 3 robotic arm models that have reached the commercial market (the Exxomove Bateo, the iARM, and the JACO robotic arm) all lack robust scientific evidence demonstrating their long-term efficacy for individuals with tetraplegia [7,14,15]; furthermore, most published research focuses on technical feasibility only [4].

The current state of research, along with the lack of high-quality studies evaluating practical effectiveness of assistive robotic systems for individuals with tetraplegia in posthospitalization settings, led us to the following objectives of our study:

1. Evaluate the feasibility of a WMRA in supporting ADLs for individuals with tetraplegia.
2. Assess user satisfaction, usability, and perceived autonomy in ADLs involving the robotic arm.
3. Collect qualitative and quantitative data using a mixed methods approach to inform further development of user-centered robotic AT systems.
4. Conduct a health economic analysis to assess the cost-effectiveness of the robotic arm in everyday use compared to formal and informal care, considering both direct and indirect costs (eg, care time and productivity loss), and linking these to outcomes (health-related QoL and perceived independence).

Methods

Design and Setting

This study uses a mixed methods research design to evaluate the feasibility and usability of a WMRA for individuals with tetraplegia. The convergent parallel design facilitates a concurrent collection of both quantitative data from standardized and nonstandardized questionnaires and qualitative insights from individual semistructured interviews, allowing for a holistic understanding of system performance and user experiences [16]. The parallel data collection enables an alignment and integration of the findings as well as the use of traditional analysis methods of quantitative and qualitative data, relying on familiar, well-established techniques.

This study is conducted in a range of posthospitalization care environments including inpatient rehabilitation facilities, home settings (with or without support by informal caregivers and/or community-based services), and nursing homes or long-term care facilities in Switzerland. This aims at maximizing external validity and informing implementation across different care pathways, reflecting on real-world contexts in which the WMRA will be used.

The robotic arm system evaluated in this study is a prototype developed by the Bern University of Applied Sciences in Switzerland using a user-centered approach. The system is built around a Kinova Gen3 robotic arm with 7 degrees of freedom [17] equipped with a customized Robotiq 2 F-85 gripper (Robotiq) [18] and is designed to be mounted on a powered wheelchair. A 3D camera located near the gripper provides visual input, which is processed by machine learning algorithms to identify and localize objects. This enables the robotic arm to approach and grasp items semiautonomously, thereby improving efficiency in task execution. A graphical user interface (GUI) displayed as a webpage on a tablet near the user offers both live and processed camera feeds and control of the robot. The user can interact with the GUI in the same way a mobile phone can be controlled in everyday life, either via the touch input on the tablet's touch screen, via voice commands (Android Voice Access) [19] or through the wheelchair's joystick (eg, joystick or chin joystick). The robotic arm system was designed with a focus on ease of use and intelligent functionality. Importantly, its software architecture is hardware-independent, offering flexibility in implementation. The fact that the robotic system is still a prototype offers a significant advantage, enabling the incorporation of user feedback to make necessary adjustments and continuous improvements, supporting an iterative design approach.

XSL•FO
RenderX

Reporting Guidelines

The Good Reporting of a Mixed Methods Study (GRAMMS) is used as a reporting guideline to improve the quality and transparency of this mixed methods research [20]. The evaluated items of the GRAMMS checklist at this time point of the research are provided in Checklist 1.

Participants and Recruitment

Participants eligible for this study are adults (≥18 y) and live in a posthospitalization care environment (inpatient rehabilitation facilities, home settings with or without support by informal caregivers and/or community-based services, and nursing homes or long-term care facilities) in Switzerland. They must have an upper limb impairment due to a complete or incomplete cervical SCI (neurological levels C0 to Th1), classified according to the American Spinal Injury Association (ASIA) Impairment Scale (AIS) as grades A to D [21]. All participants are required to use a powered wheelchair, and their experience in using a WMRA is not considered as an exclusion criterion. They need to be able to understand and communicate in German, French, Italian, or English. Individuals are excluded from participation in this study if they are unable to operate AT for the upper extremities due to psychological or cognitive limitations. Further exclusion criteria include the presence of a progressive secondary condition that is expected to further impair upper limb function, a diagnosis of a functional neurological disorder, or current pregnancy.

Purposeful sampling [16] using a snowball approach is applied for recruitment, where study team members contact possible participants within their individual network and ask for interest in the study by phone or via email. A flyer containing information about the study and the recruitment is shared on social media (LinkedIn [Microsoft Corp], Facebook [Meta], and other websites) as well as with institutions connecting potential participants. Prospective participants receive detailed information about the study's purpose and their rights through an informed consent form, which can be signed by the participant or a legal representative. Participants are informed that participation is voluntary, that they may withdraw consent at any time without affecting their subsequent medical care or treatment, and that no compensation will be provided for participation in the study.

Study Endpoints

The main primary endpoint of this study is to assess the feasibility of using the robotic arm in supporting individuals with tetraplegia performing ADLs in a posthospitalization setting. Feasibility is evaluated through 3 predefined tasks: pickup, drink, and push a button, which serve to assess the usability and effectiveness of the robotic arm. A visual outline of the assessment of the 3 tasks is provided in the Multimedia Appendix 1. The robotic arm will be considered feasible if participants demonstrate the ability to complete at least 2 of the 3 predefined ADL tasks (pickup, drink, and push a button) independently in ≥70% of attempts, achieve a usability score equivalent to a System Usability Scale ≥68 (indicating good usability) [22], and report no excessive cognitive or physical workload as measured by the National Aeronautics and Space Administration-Task Load Index (NASA-TLX). Meeting these criteria collectively is considered an indication that the robotic arm is functionally effective, usable, reliable, and acceptable for supporting ADLs in a real-world setting.

The user-friendliness of the system is measured using the Post-Study System Usability Questionnaire (PSSUQ), a standardized 19-item instrument [23]. The NASA-TLX assesses the perceived cognitive and physical workload of the participants associated with controlling the robot arm when performing ADLs [24]. Additionally, the qualitative data collection aims to gain deeper insights into the physical relief in everyday life and the user experiences. For this, an informal questionnaire developed by the research team evaluates the feasibility, usability, and real-world impact of the robotic arm after use. Based on a semistructured guide, individual feedback is collected from every participant in an interview conducted via smartphone or online via Microsoft Teams. The semistructured interview guide was developed by a member of the research team (ITMdB) and has been reviewed by 2 physiotherapists (AMR and BW), 3 engineers (AC, RR, and JFR), and 1 anonymous individual with tetraplegia to ensure comprehensibility and relevance, and to improve its validity. An outline of the semistructured interview guide is provided in the Multimedia Appendix 2.

Secondary study endpoints include the QoL of the participants assessed using the EQ-5D-5L [25], as well as pain in the upper extremities determining the localization and quality, quantified using a Visual Analogue Scale (VAS) [26]. Self-perceived performance and satisfaction in ADLs are evaluated using an adapted version of the Canadian Occupational Performance Measure (aCOPM) [27], modified to focus on the 10 most important ADLs for individuals with tetraplegia as identified by Hutmacher et al [28]. Information on the sociodemographic background and the use of health and social care services by the study participants are collected using the Client Socio-Demographic and Service Receipt Inventory-European Version (CSSRI-EU) [29]. The inventory captures the type and amount of informal and formal care received (eg, h per wk), out-of-pocket expenses for assistive devices or services, perceived financial burden due to disability-related costs and employment status. The CSSRI-EU has been translated into German and adapted to the structure of the Swiss health and social care system by experts, including relevant terminology, service categories, and financing structures. In addition, it was extended by study-specific items to capture relevant health economic data eg, productivity losses of informal caregivers by the research team.

Other study endpoints include demographic and personal parameters of the participants (age, sex, address, phone number, and email) as well as medical parameters (time after injury, height of the lesion, completeness of injury, spasticity and tonus of upper extremities, and mobility of upper extremities), and technical parameters about the wheelchair and the living setting (brand and model of powered wheelchair, home setting, limited everyday activities, assistance in everyday life, current aids for performing everyday activities).

XSL•FO
RenderX

Participation Timeline

At least 10 days prior to the clinical phase, participants will receive detailed study information and an informed consent form via email. The signed consent form will be collected at least 7 days before the clinical phase. Upon the receipt of the informed consent form, the questionnaire on demographics, EQ-5D-5L, pain with the VAS, aCOPM, and the CSSRI-EU will be handed out to the participants via email.

All assessments will take place at the posthospital setting where the patient resides. First, baseline assessment of the 3 tasks (pickup, drink, and push a button) will be collected. The robotic arm will then be mounted on the personal wheelchair of the participant by an engineer of the research team. The system will be calibrated, and training is provided based on a strict manual given by the engineer, as well as individual testing of the system by the participant. The training lasts at least 45 minutes and only ends when the participant has no further questions. After training, participants must be able to issue robot commands via at least one supported input pathway (touch, voice, or wheelchair joystick or chin joystick). The robotic arm will then be steered by the participant him or herself and the preassessment of the 3 tasks (pickup, drink, and push a button) performed, as well as questionnaires (PSSUQ and NASA-TLX) filled in. The participant will now be using the robotic arm in ADLs for a total of 6 consecutive days at home. During the clinical test phase, daily log files of the robotic arm will automatically be sent to the responsible researcher through a mobile internet connection. Postassessment of the 3 tasks (pickup, drink, and push a button) as well as postquestionnaires (PSSUQ, NASA-TLX, EQ-5D-5L, as well as the informal questionnaire on the feasibility and usability) will be collected after the 6 consecutive days. The robotic arm will then be removed from the wheelchair by a member of the research team. Two days later, the aCOPM and the VAS will be distributed online to the participants via REDCap (Research Electronic Data Capture; Vanderbilt University). Around 2 to 5 days after the removal of the robot arm, qualitative data, based on the semistructured interview guide, will be collected individually either via telephone or online via Microsoft Teams. The interviews will be held by a member of the research team in German unless Swiss German is preferred and take about 30 minutes.

Underlining the importance of understanding individual variability and following comparable AT research, screening and recruitment will continue until the target sample size of a maximum of 15 study participants using the robotic arm is reached [30-32]. The target sample size follows recommendations of guidelines of comparable research where sample sizes of 30 participants and less are expected to reach saturation in quantitative studies, as well as practical considerations such as a realistically achievable number of participants meeting the inclusion criteria [33,34].

The data collected from these participants will also contribute to the health economic evaluation. To enhance the statistical power for detecting potential effects on QoL, an equal number of additional participants, who cannot be allocated to using the robotic arm due to individual reasons for technical incompatibility (eg, wheelchair incompatibility or environmental constraints), will be recruited from the same eligible study population [35]. Recruitment began in April 2025, with an anticipated enrollment period concluding in December 2025. A complete overview of the study procedures can be found in Figure 1.

XSL•FO
RenderX

Figure 1. Overview of study procedures. Displayed in the upper row is the intervention group, where screening and recruitment are followed by the baseline assessment, after which the robotic arm system is mounted on the participant's wheelchair and used to support activities of daily living for 6 days. Within 2 to 5 days postassessments and questionnaires, as well as individual interviews, are conducted. The bottom row shows the nonintervention group, where after screening and recruitment, no clinical testing phase is undertaken but only the completion of the prequestionnaires. aCOPM: Questionnaire on independence in everyday life adapted from Canadian Occupational Performance Measure; CSSRI-EU: Client Socio-Demographic and Service Receipt Inventory-European Version; NASA-TLX National Aeronautics and Space Administration-Task Load Index; PSSUQ: Post-Study System Usability Questionnaire; VAS: Visual Analogue Scale.

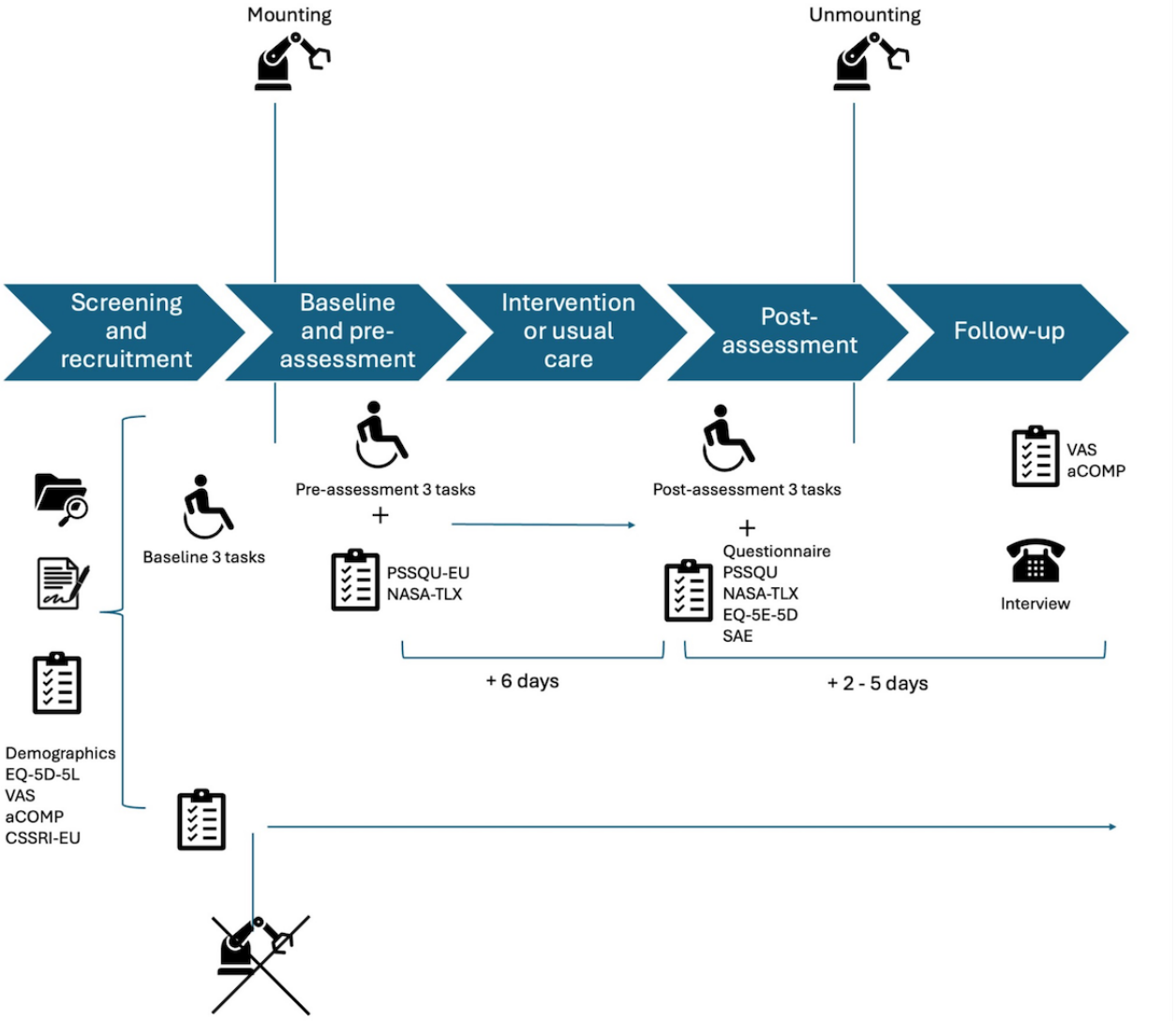

Outcomes

Primary and secondary outcomes will be assessed at 3 time points: first, at inclusion and preassessment; second, at the primary endpoint, following 6 days of use with the robotic arm mounted on the wheelchair; and third, at the secondary endpoint, 2 to 3 days after the robotic arm has been unmounted (Table 1).

XSL•FO
RenderX

Table . Schedule of study procedures and assessments.

Study contacts	Screening and recruitment	Inclusion	Baseline assessments	Preassessment	Intervention or usual care	Primary endpoint	Secondary endpoint
Time	10 days before clinical phase (without robotic arm)	7 days before clinical phase (without robotic arm)	(Without robotic arm)	After 45 minutes training (with robotic arm)	6 consecutive days	After 6 consecutive days (with robotic arm)	2-5 days after clinical phase (without robot arm)
Informed consent form	✓	—[a]	—	—	—	—	—
Informal questionnaire	—	—	—	—	—	✓	—
Demographics[b]	—	✓	—	—	—	—	—
Primary outcome: usability							
Assessment of 3 tasks (pick up, drink, and push a button)	—	—	✓	✓	—	✓	—
PSSUQ[c]	—	—	—	✓	—	✓	—
NASA-TLX[d]	—	—	—	✓	—	✓	—
Secondary outcome: increased QoL[e] with the use of the robotic arm							
EQ-5D-5L	—	✓	—	—	—	✓	—
Pain with VAS[f]	—	✓	—	—	—	—	✓
aCOPM[g]	—	✓	—	—	—	—	✓
CSSRI-EU[h]	—	✓	—	—	—	—	—
Personal interview	—	—	—	—	—	—	✓

[a]Not applicable.

[b]Age; sex; time after disability; height of the lesion; completeness of disability; brand and model of powered wheelchair; spasticity and tonus of upper extremities; mobility of upper extremities; home setting.

[c]PSSUQ: Post-Study System Usability Questionnaire.

[d]NASA-TLX: National Aeronautics and Space Administration Task Load Index.

[e]QoL: quality of life.

[f]Pain with VAS: pain with Visual Analogue Scale.

[g]aCOPM: Questionnaire on independence in everyday life, adapted from Canadian Occupational Performance Measure.

[h]CSSRI-EU: Client Sociodemographic and Service Receipt Inventory.

Data Management

Quantitative study data from the forthcoming evaluation of the robotic arm and the economic evaluation will be managed using REDCap electronic data capture tools. REDCap is a secure, web-based software platform designed to support data capture for research studies. Password-protected accounts will be created only for authorized study team members, and the level of database access granted to each member depends on their respective roles within the study [36]. Due to the stacked structure of the clinical phase, it will be possible for the research team members to identify the study participants based on the time of data collection; however, further processing of the data takes place anonymously. Daily log files capturing how, when, and where the robotic arm is used, along with data on GUI interactions during ADLs, will be automatically stored in a secure Microsoft Teams folder accessible only to the research team. They will be used to document usage patterns, monitor system performance, and support both technical diagnostics and research evaluation of the WMRA. The log files will contribute to further development and optimization of the system based on the real-world user interaction.

Qualitative data will be saved in a protected file on Microsoft Teams with access only to members of the research group. Data will be pseudonymized with a number, but study participant identification can take place with the time point of data collection. All data, including both electronic and paper-based records, will be retained for an indefinite period. The project leader is responsible for securely storing the research data for

XSL•FO
RenderX

at least 10 years following the completion or early termination of the study.

Data Analysis

Quantitative data will be analyzed using the R statistical computing environment (version 2025.05.0+496; R Core Team) [37], MATLAB (version R2024a; MathWorks) [38] and/or Python (version 3.13; Python Software Foundation) [39]. Descriptive statistics will be used to describe the characteristics of each participant, depending on the nature and distribution of the data, including mean (SD) or median (IQR) for age, disability, and questionnaire data, while categorical variables such as sex will be summarized using frequencies and percentages. The use of standardized instruments such as the EQ-5D-5L will allow for both between- and within-group effects about pre- and postintervention differences. For the assessment of changes in perceived independence in daily activities, a split-plot design will be applied to aCOPM data, enabling both within-subject and between-group comparisons. Repeated measures ANOVA will be used to assess main effects and interaction effects over time. PSSUQ data will be analyzed using ordinal regression to model Likert-scale responses, supplemented by descriptive statistics and mean item scores for usability subscales. NASA-TLX scores will be evaluated using repeated measures ANOVA or nonparametric equivalents to assess cognitive and physical workload changes over time. Task performance (pick-up, drink, and push button) will be analyzed using completion rates and time-to-complete measures, with baseline (without robotic arm) and pre and post (with robotic arm) comparisons performed using paired tests to assess feasibility.

For the health evaluation, quality-adjusted life years (QALYs) will be calculated using the EQ-5D-5L instrument, following the methodology outlined by Sassi [40]. The QALY estimates will be combined with data from the Care-Related Support Index (CRSI) to calculate the Incremental Cost-Effectiveness Ratio (ICER), providing a measure of cost-effectiveness for the intervention. The robustness of ICER results will be examined through both deterministic and probabilistic sensitivity analyses. The latter will use Monte Carlo simulations to account for parameter uncertainty. Given the absence of a binding or explicitly defined ICER threshold in Switzerland, cost-effectiveness acceptability curves (CEACs) will be generated to illustrate the probability that the intervention is cost-effective across a range of willingness-to-pay thresholds. Subgroup analyses will explore heterogeneity in cost-effectiveness.

Qualitative data from informal postquestionnaire responses and semistructured interviews will be analyzed using qualitative content analysis, following an inductive approach to allow themes to emerge from the data. Interviews, transcribed according to predefined rules (Kuckartz and Rädiker [41]), will be analyzed using the AI-based software noScribe [42]. The rapid and rigorous data reduction technique will be used for iterative data reduction, enabling synthesis and thematic comparison with quantitative outcomes [43].

In the final stage of the data interpretation, the quantitative and qualitative results will be integrated in line with the convergent parallel mixed methods design. Findings will be compared across methodological strands to generate a comprehensive understanding of the feasibility, usability, and personal impact of the WMRA. With this, an explanation for possible divergences between performance outcomes (3 tasks, PSSUQ, and NASA-TLX) and user satisfaction, as well as contextual factors influencing variability in quantitative outcomes, can be identified and given. The refinement of the WMRA as a prototype can be established further incorporating relevant user-centered adjustments. In order to minimize the bias of disproportionate emphasis on either type of data, which come along with the different nature of the data types and represent a challenge of mixed methods research [44], regular multidisciplinary analysis meetings will be held. The meetings involve experts in physiotherapy, engineering, health economics, and will ensure a balance of both data types. Furthermore, findings will be cross-validated through triangulation where quantitative and qualitative results will be examined in relation to each other [45].

Missing Data

To minimize missing data, a stepwise protocol will be implemented to ensure timely and complete data collection at all assessment points. Participants will receive detailed instructions and reminders via REDCap, and research team members will be responsible for verifying questionnaire completion during the study. Despite preventive measures, incomplete data sets will be included in the analysis. If missing values occur, they will be addressed using a multiple imputation approach suitable to the type and extent of the missing data. Recruitment will continue until the predefined sample size is achieved, ensuring adequate statistical power for both usability and health economic analyses.

Ethical Considerations

A formal clarification of jurisdiction was submitted to the competent cantonal ethics committees to determine whether the study falls under the scope of the Swiss Human Research Act. The committees reviewed the project and declared themselves not responsible, as the study does not fall under the Human Research Act (Art. 2, Abs. 1) [46]. The project does not involve a medical product, biological material, or any modification of standard care. This decision is documented under BASEC-Nr: REq-2022 - 01492. Although formal ethics approval was not required, the study will be conducted in accordance with recognized ethical standards, including respect for the dignity, privacy, and autonomy of all participants. All participants provided informed consent prior to participation in the study. Participant data will be pseudonymized and stored securely, with access restricted to authorized study personnel. Participants may withdraw from the study at any time without providing a reason and without any negative consequences. Appropriate safety measures will be implemented throughout the study, including continuous monitoring for adverse events and predefined protocols for incident response, in line with the Declaration of Helsinki [47].

XSL•FO
RenderX

Results

Recruitment for this study began in April 2025. As outlined in this protocol, up to 15 participants meeting the inclusion criteria will be enrolled in the phase of the robotic arm intervention, with an additional group of up to 15 participants contributing data only for the health economic evaluation. At the time of the manuscript submission, the study is actively recruiting participants, and enrollment and data collection are proceeding according to the predefined study timeline, with no major deviations from the protocol.

As this manuscript reports a study protocol, no outcome data are available at this stage. Upon study completion, results related to feasibility, usability, and the health economic evaluation will be reported.

Discussion

Anticipated Findings

This study aims to evaluate the feasibility, usability, and potential health-economic impact of a WMRA, focusing on whether the system is user-friendly and applicable in the daily lives of individuals with tetraplegia in posthospitalization settings.

Based on existing evidence, a reduction in care time for assistance in basic ADLs through the use of the robotic arm may be expected. For instance, the Jaco robotic arm has been shown to save approximately 1.31 hours per day for 31 out of 34 users across 16 standard tasks such as grasping a bottle or pressing a button [12]. Similarly, the iARM system demonstrated an average time saving of 1.25 hours per day [48]. These findings suggest that our robotic system has the potential to significantly reduce the time users depend on external support for basic and instrumental ADLs. From an economic standpoint, the implications of such time savings are considerable. In Switzerland, persons with SCI receive on average 27 hours of care per week [49]. If this care were fully provided for, it would amount to approximately CHF 60,000 (US $78,154.14) per year. Therefore, even partial automation of care could alleviate substantial costs, providing a compelling case for broader implementation and reimbursement of assistive robotic technologies. Importantly, these potential cost savings must be considered in conjunction with the anticipated improvements in QoL. By linking reduced care time to measurable gains in QALYs, the intervention's overall value can be assessed through cost-effectiveness analysis. Furthermore, our study builds on previous user needs assessments, including the work of Hutmacher et al [28], which emphasized the psychological and physical burden placed on informal caregivers, often over extended periods [6]. By offering relief in routine tasks, the robotic arm may help reduce caregiver strain and promote greater autonomy and self-efficacy among individuals with tetraplegia [7]. This is particularly relevant, as increased autonomy is one of the most frequently reported unmet needs among individuals with high cervical injuries. Through independent task performance and reduced reliance on caregivers, robotic arms may contribute positively to self-perceived independence and mental well-being. Our study is designed to assess these domains using validated instruments such as the EQ-5D-5L and aCOPM, providing quantitative insight into the real-life impact of robotic AT.

Challenges and Limitations

Despite its promise, the implementation of assistive robotic arms is not without technical and logistical challenges. As the robotic system under investigation is a prototype, hardware malfunctions or software issues could require intervention by technical staff or, in worst-case scenarios, require repair. Such disruptions could significantly hinder the daily use and long-term sustainability of the system in real-world environments. Technical issues might also influence not only feasibility but also secondary outcomes including perceived usability, task performance, cognitive workload, and even potentially responses to the QoL. To account for these technical incidents, they will be logged and considered in the interpretation of feasibility, usability, and economic endpoints.

Participant recruitment presents another challenge, a known barrier in comparable AT studies. Previous projects have struggled to recruit sufficient participants, largely due to the complex logistics of home-based interventions. These often disrupt fragile support systems essential to daily survival, placing additional stress on participants and caregivers. By mixing different inpatient settings forms (inpatient rehabilitation facilities, home settings, nursing homes, or long-term care facilities), the difficulty of participant recruitment is addressed. This approach also allows for the identification of possible integration challenges that are essential for scalable deployment of the WMRA and could be further explored in future studies.

Recruitment is further constrained by the small size of the target population. We address this risk by using a snowball recruitment strategy and diversifying outreach through social media, rehabilitation centers, and caregiver networks. Nonetheless, we remain aware of the limitations this poses for the generalizability of our findings.

The expected heterogeneity within the study sample arising from differences in motor capacity, control interface requirements, and available caregiver support poses another relevant challenge in this study. However, this is a variability inherent to the target user population of the WRMA, and we therefore consider this to ensure ecological validity. Neurological level and ASIA grade will be recorded and included as covariates in the analysis to account for these differences and where subgroup analyses will be conducted where appropriate. Additionally, qualitative data from the user will be used to contextualize the quantitative outcomes, ensuring that differences in independence and support availability are properly interpreted.

Finally, the short intervention period of 6 days limits the ability to capture long-term effects on QoL or sustained behavioral adaptation, which are essential for robust QALY estimation. We are aware that within this period, only immediate psychological impacts of first use can be captured and that it will be used exploratorily in this feasibility study. To measure sustained changes in mental well-being, a longer study period needs to be addressed in future works. Since the WMRA tested

XSL•FO
RenderX

is still in the prototype stage and there are currently very few comparable publications evaluating feasibility of a WMRA [11,13,50,51], the test duration of 6 consecutive days was chosen for logistical reasons and to give priority to safety monitoring and minimize user fatigue. Considering future studies such as a pilot clinical trial, the use period could be extended to 2 or 3 weeks to improve the quantitative and qualitative insights gained and measured.

Conclusion and Impact

This protocol describes a real-world convergent parallel mixed methods feasibility study evaluating a wheelchair-mounted robotic arm designed to support activities of daily living for individuals with tetraplegia across posthospitalization settings in Switzerland. By combining standardized usability and workload measures with task-performance testing, usage logs, qualitative interviews, and a health-economic evaluation, the study is designed to generate complementary evidence on functional feasibility, user experience, and potential cost-effectiveness in everyday contexts.

Pending positive findings, this study may contribute to larger-scale evaluation and long-term clinical trials that assess real-world effectiveness, cost-efficiency, and user benefit of assistive robotic arm systems for persons with tetraplegia. The co-development with end users ensures future iteration and that diverse user needs and preferences of the robotic arm are met, making future versions more user-centered and effective. Ongoing collaboration with engineers and health care professionals could help overcome cost barriers and facilitate broader adoption of the robotic arm.

Patient and Public Involvement

The development process and design of the robotic arm system was iterative and informed by a previous research cross-sectional study that examined AT needs for individuals with tetraplegia [28]. Importantly, individuals with tetraplegia were actively involved from the outset, contributing their lived experiences and feedback throughout the design and laboratory-based testing phases. A first version of the prototype was showcased at the CYBATHLON 2024 in the Assistance Robot Race, successfully completing 10 daily living tasks, including retrieving a parcel from a mailbox or navigating a touchscreen [52].

Acknowledgments

The authors extend their gratitude to all participants for sharing their expertise and dedicating time and effort to this research. They also acknowledge the project team at the Bern University of Applied Sciences for their contributions. According to the International Association of Scientific, Technical, and Medical Publishers Recommendations for a classification of artificial intelligence (AI) use in academic manuscript preparation [53], generative AI was used in refinement, correction, editing, and formatting of this manuscript to improve the clarity of language. Furthermore, generative AI was used in drafting and refinement of some manuscript content and in the translation of specific words of the manuscript.

Funding

The overall project was financially supported by the Paraplegiker-Stiftung, which provided a grant in 2023. The funder is not invovled in role in study design, data collection, analysis, interpretation, or publication decisions.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the nature of the participant population and the potential for reidentification even after pseudonymization but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: BCW, IDB, AMR, and AC
Data curation: BCW, IDB, and AC
Formal analysis: BCW, IDB, and AC
Funding acquisition: AMR and GG
Investigation: BCW, IDB, RR, and AC
Methodology: BCW, IDB, RR, and AC
Project administration: AMR and GG
Resources: BCW, IDB, and AC
Software: BCW, IDB, RR, JFR, and AC
Supervision: AMR
Validation: AMR
Visualization: FV, BCW, IDB, and AC
Writing original draft: FV
Data curation: BCW, IDB, and AC

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Assessments – 3 Tasks.
[DOCX File, 274 KB - resprot_v15i1e78339_app1.docx]

Multimedia Appendix 2
Semistructured interview guide.
[DOCX File, 18 KB - resprot_v15i1e78339_app2.docx]

Checklist 1
GRAMMS checklist.
[PDF File, 84 KB - resprot_v15i1e78339_app3.pdf]

References

1. Rupp R. Chapter 6 - spinal cord lesions. In: Ramsey NF, Millán J del R, editors. Handbook of Clinical Neurology: Elsevier; 2020:51-65. [doi: 10.1016/B978-0-444-63934-9.00006-8]
2. Spinal cord injury. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/spinal-cord-injury [accessed 2026-02-03]
3. Leone S, Giunta L, Rino V, et al. Design of a wheelchair-mounted robotic arm for feeding assistance of upper-limb impaired patients. Robotics 2024 Mar;13(3):38. [doi: 10.3390/robotics13030038]
4. Readioff R, Siddiqui ZK, Stewart C, Fulbrook L, O'Connor RJ, Chadwick EK. Use and evaluation of assistive technologies for upper limb function in tetraplegia. J Spinal Cord Med 2022 Nov;45(6):809-820. [doi: 10.1080/10790268.2021.1878342] [Medline: 33606599]
5. Pacheco Barzallo D, Oña A, Brach M, Gemperli A. Healthcare service utilization and related costs. A comparison of the situation of persons with spinal cord injury to the general population. Am J Phys Med Rehabil 2024 Nov 1;103(11S Suppl 3):S251-S259. [doi: 10.1097/PHM.0000000000002539] [Medline: 39453854]
6. Keihanian F, Kouchakinejad-Eramsadati L, Yousefzadeh-Chabok S, Homaie Rad E. Burden in caregivers of spinal cord injury patients: a systematic review and meta-analysis. Acta Neurol Belg 2022 Jun;122(3):587-596. [doi: 10.1007/s13760-022-01888-2] [Medline: 35157242]
7. Beaudoin M, Lettre J, Routhier F, Archambault PS, Lemay M, Gélinas I. Impacts of robotic arm use on individuals with upper extremity disabilities: a scoping review. Can J Occup Ther 2018 Dec;85(5):397-407. [doi: 10.1177/0008417418820878] [Medline: 30866682]
8. Zahlen und fakten [Web page in German]. Schweizer Paraplegiker-Gruppe. 2022. URL: https://www.paraplegie.ch/de/ueber-uns/organisation/zahlen-und-fakten/ [accessed 2026-02-03]
9. Gallenberger D, Bhattacharjee T, Kim Y, Srinivasa SS. Transfer depends on acquisition: analyzing manipulation strategies for robotic feeding. Presented at: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI); Mar 11-14, 2019; Daegu, Korea (South) p. 267-276. [doi: 10.1109/HRI.2019.8673309]
10. Bekey G, Ambrose R, Kumar V, Sanderson A, Wilcox B, Zheng Y. WTEC panel report on international assessment of research and development in robotics. World Technology Evaluation Center, Inc. 2006 URL: https://de.scribd.com/document/39996852/Screen-Robotics-Final-Report [accessed 2026-02-12]
11. Graser A, Heyer T, Fotoohi L, et al. A supportive FRIEND at work: robotic workplace assistance for the disabled. IEEE Robot Automat Mag 2013 Dec;20(4):148-159. [doi: 10.1109/MRA.2013.2275695]
12. Maheu V, Frappier J, Archambault PS, Routhier F. Evaluation of the JACO robotic arm: clinico-economic study for powered wheelchair users with upper-extremity disabilities. Presented at: 2011 IEEE International Conference on Rehabilitation Robotics; Jun 29 to Jul 1, 2011. [doi: 10.1109/ICORR.2011.5975397]
13. Kyrarini M, Lygerakis F, Rajavenkatanarayanan A, et al. A survey of robots in healthcare. Technologies (Basel) 2021 Jan 18;9(1):8. [doi: 10.3390/technologies9010008]
14. Mekki M, Delgado AD, Fry A, Putrino D, Huang V. Robotic rehabilitation and spinal cord injury: a narrative review. Neurotherapeutics 2018 Jul;15(3):604-617. [doi: 10.1007/s13311-018-0642-3] [Medline: 29987763]
15. EXXOMOVE® GmbH [Webpage in German]. Exxomove. URL: https://www.exxomove.de/ [accessed 2026-02-03]
16. Reporting results in mixed methods research. ATLAS.ti. URL: https://atlasti.com/guides/the-guide-to-mixed-methods-research/reporting-results-in-mixed-methods-research [accessed 2026-02-03]
17. Kollaborative roboter (cobots) [Web page in German]. Phoenix Mecano. 2025 May 22. URL: https://www.phoenix-mecano.ch/de/produkte/produktion/kollaborative-roboter/ [accessed 2026-02-03]
18. Adaptive grippers. Robotiq. 2025 May 22. URL: https://robotiq.com/products/adaptive-grippers [accessed 2026-02-03]

XSL•FO
RenderX

19. SpeechRecognizer class (Android.Speech). Microsoft Learn. 2025 May 22. URL: https://learn.microsoft.com/en-us/dotnet/api/android.speech.speechrecognizer?view=net-android-35.0 [accessed 2026-02-03]
20. Emary PC, Stuber KJ, Mbuagbaw L, et al. Quality of reporting using good reporting of a mixed methods study criteria in chiropractic mixed methods research: a methodological review. J Manipulative Physiol Ther 2023;46(3):152-161. [doi: 10.1016/j.jmpt.2023.11.004] [Medline: 38142381]
21. Rupp R, Biering-Sørensen F, Burns SP, et al. International standards for neurological classification of spinal cord injury: revised 2019. Top Spinal Cord Inj Rehabil 2021;27(2):1-22. [doi: 10.46292/sci2702-1] [Medline: 34108832]
22. System usability scale - an overview. ScienceDirect. URL: https://www.sciencedirect.com/topics/computer-science/system-usability-scale [accessed 2026-02-03]
23. Lewis JR. Psychometric evaluation of the Post-Study System Usability Questionnaire: the PSSUQ. Proc Hum Factors Soc Annu Meet 1992 Oct;36(16):1259-1260. [doi: 10.1177/154193129203601617]
24. Hart SG, Staveland LE. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. Human Mental Workload: North-Holland; 1988:139-183. [doi: 10.1016/S0166-4115(08)62386-9]
25. EQ-5D-5L. EuroQol. 2025. URL: https://euroqol.org/information-and-support/euroqol-instruments/eq-5d-5l/ [accessed 2026-02-03]
26. Heller GZ, Manuguerra M, Chow R. How to analyze the Visual Analogue Scale: myths, truths and clinical relevance. Scand J Pain 2016 Oct;13:67-75. [doi: 10.1016/j.sjpain.2016.06.012] [Medline: 28850536]
27. Chatfield J, Beckett D. The Canadian Occupational Performance Measure: use in an independent living centre. Int J Ther Rehabil 2007 Jun;14(6):280-283. [doi: 10.12968/ijtr.2007.14.6.23900]
28. Hutmacher N, Bellwald A, Rätz R, et al. Identification of needs for an assistive robotic arm in individuals with tetraplegia: a mixed-methods approach. J NeuroEngineering Rehabil 2025;22(1):113. [doi: 10.1186/s12984-025-01642-8] [Medline: 40394679]
29. Chisholm D, Knapp MRJ, Knudsen HC, et al. Client Socio-Demographic and Service Receipt Inventory – European Version: development of an instrument for international research. Br J Psychiatry 2000 Jul;177(S39):s28-s33. [doi: 10.1192/bjp.177.39.s28] [Medline: 10945075]
30. Folan A, Barclay L, Cooper C, Robinson M. Exploring the experience of clients with tetraplegia utilizing assistive technology for computer access. Disabil Rehabil Assist Technol 2015 Jan;10(1):46-52. [doi: 10.3109/17483107.2013.836686] [Medline: 24050283]
31. Zimmermann G, Bolter LM, Sluka R, et al. Sample sizes and statistical methods in interventional studies on individuals with spinal cord injury: a systematic review. J Evid Based Med 2019 Aug;12(3):200-208. [doi: 10.1111/jebm.12356] [Medline: 31231977]
32. Javaremi MN, Argall BD. Characterization of assistive robot arm teleoperation: a preliminary study to inform shared control. arXiv. Preprint posted online on Jul 31, 2020. [doi: 10.48550/arXiv.2008.00109]
33. Teresi JA, Yu X, Stewart AL, Hays RD. Guidelines for designing and evaluating feasibility pilot studies. Med Care 2022 Jan 1;60(1):95-103. [doi: 10.1097/MLR.0000000000001664] [Medline: 34812790]
34. Allin S, Eckel E, Markham H, Brewer BR. Recent trends in the development and evaluation of assistive robotic manipulation devices. Phys Med Rehabil Clin N Am 2010 Feb;21(1):59-77. [doi: 10.1016/j.pmr.2009.09.001] [Medline: 19951778]
35. Asadi-Lari M, Packham C, Gray D. Unmet health needs in patients with coronary heart disease: implications and potential for improvement in caring services. Health Qual Life Outcomes 2003 Jul 23;1(1):26. [doi: 10.1186/1477-7525-1-26] [Medline: 12917018]
36. Software. REDCap. 2025. URL: https://projectredcap.org/software/ [accessed 2026-02-03]
37. RStudio builds: version 2025.05.0+496. RStudio. 2025. URL: https://dailies.rstudio.com/version/2025.05.0+496/ [accessed 2026-02-03]
38. MATLAB. MathWorks. 2025. URL: https://www.mathworks.com/products/matlab.html [accessed 2026-02-03]
39. What's new in python 3.13. Python documentation. 2025 May 28. URL: https://docs.python.org/3/whatsnew/3.13.html [accessed 2026-02-03]
40. Sassi F. Calculating QALYs, comparing QALY and DALY calculations. Health Policy Plan 2006 Sep;21(5):402-408. [doi: 10.1093/heapol/czl018] [Medline: 16877455]
41. Kuckartz U, Rädiker S. Qualitative inhaltsanalyse: methoden, praxis, computerunterstützung. 5., überarbeitete auflage [Report in German]. : Weinheim Basel: Beltz Juventa; 2022 URL: https://www.beltz.de/fileadmin/beltz/leseproben/978-3-7799-3344-1.pdf [accessed 2026-02-12]
42. Kaixxx/noscribe. GitHub. 2025. URL: https://github.com/kaixxx/noScribe [accessed 2026-02-03]
43. Watkins DC. Rapid and rigorous qualitative data analysis: The "RADaR" technique for applied research.. Int J Qual Methods 2017 Dec 1;16:1609406917712131. [doi: 10.1177/1609406917712131]
44. A discussion of some controversies in mixed methods research for emerging researchers. Sage. 2025. URL: https://journals.sagepub.com/doi/epub/10.1177/20597991221123398 [accessed 2026-02-03]
45. Triangulation in mixed methods research. ATLAS.ti. 2025 Nov 4. URL: https://atlasti.com/guides/the-guide-to-mixed-methods-research/triangulation-in-mixed-methods-research [accessed 2026-02-03]

XSL•FO
RenderX

46. SR 81030 - federal act of 30 september 2011 on research involving human beings (human research act, HRA). Fedlex. 2025. URL: https://www.fedlex.admin.ch/eli/cc/2013/617/en [accessed 2026-02-03]
47. WMA declaration of Helsinki – ethical principles for medical research involving human participants. World Medical Association. 2025. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki/ [accessed 2026-02-03]
48. Romer G, Stuyt HJA, Peters A. Cost-savings and economic benefits due to the assistive robotic manipulator (ARM). Presented at: 9th International Conference on Rehabilitation Robotics, 2005 ICORR 2005; Jun 28 to Jul 1, 2005; Chicago, IL, USA p. 201-204. [doi: 10.1109/ICORR.2005.1501084]
49. Pacheco Barzallo D, Hernandez R, Brach M, Gemperli A. The economic value of long-term family caregiving. The situation of caregivers of persons with spinal cord injury in Switzerland. Health Soc Care Community 2022 Sep;30(5):e2297-e2307. [doi: 10.1111/hsc.13668] [Medline: 34854509]
50. Routhier F, Archambault P, Maheu V, Cyr MC, Lemay M, Gélinas I. Benefits of jaco robotic arm on independent living and social participation: an exploratory study. Presented at: RESNA Annual Conference; Jun 11-15, 2014 URL: https://www.resna.org/sites/default/files/conference/2014/Robotics/Routhier.html [accessed 2026-02-03]
51. Gbètoho Atigossou OL, Faieta J, Corcuff M, et al. Social impact of the JACO wheelchair-mounted robotic arm on users and their caregivers. medRxiv. Preprint posted online on September 19, 2024. [doi: 10.1101/2024.09.18.24313906]
52. Assistenzroboter-Rennen [Web page in German]. CYBATHLON ETH Zürich. URL: https://cybathlon.ethz.ch/de/event/disziplinen/rob [accessed 2026-02-03]
53. Recommendations for a classification of AI use in academic manuscript preparation. : STM Association; 2025 URL: https://s3.eu-west-2.amazonaws.com/stm.offloadmedia/wp-content/uploads/2025/04/23020709/STM_AI_Classification_Recs_19_Sept2025-1.pdf [accessed 2026-02-03]

Abbreviations

aCOPM: Questionnaire on independence in everyday life, adapted from Canadian Occupational Performance Measure
ADL: activity of daily living
AIS: ASIA Impairment Scale
ASIA: American Spinal Injury Association
AT: assistive technology
CEACs: cost-effectiveness acceptability curves
CRSI: Care-Related Support Index
CSSRI-EU: Client Socio-Demographic and Service Receipt Inventory-European Version
FRIEND: User Friendly Interface for Disabled People
GRAMMS: Good Reporting of a Mixed Methods Study
GUI: graphical user interface
ICER: Incremental Cost-Effectiveness Ratio
NASA-TLX: National Aeronautics and Space Administration-Task Load Index
PSSUQ: Post-Study System Usability Questionnaire
QALY: quality-adjusted life year
QoL: quality of life
REDCap: Research Electronic Data Capture
SCI: spinal cord injury
VAS: Visual Analogue Scale
WMRA: wheelchair-mounted robotic arm

Edited by A Schwartz; submitted 05.Jun.2025; peer-reviewed by M Adeleke, R Gassert; accepted 30.Nov.2025; published 03.Mar.2026.

Please cite as:
Fosbrooke V, Christen A, Wortmann BC, de Boer ITM, Rätz R, Rösch JF, Gruener G, Raab AM
Assistive Robotic Arm to Support Activities of Daily Living in Individuals With Tetraplegia: Protocol for a Real-World Convergent Parallel Mixed Methods Feasibility Study
JMIR Res Protoc 2026;15:e78339
URL: https://www.researchprotocols.org/2026/1/e78339
doi:10.2196/78339

© Vera Fosbrooke, Aline Christen, Barbara Catherine Wortmann, Iris Theodora Maria de Boer, Raphael Rätz, Julian Frederik Rösch, Gabriel Gruener, Anja M Raab. Originally published in JMIR Research Protocols (https://www.researchprotocols.org),

XSL•FO
RenderX

3.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Use of an AI-Based Tool (Human Experiences and Reflections Data Connector) to Improve Discovery and Reuse of Archived Qualitative Data: Protocol for an Algorithm Development and Validation Study

Tammy Leonard[1,2], PhD; Jim P Stimpson[1,2,3], PhD; Miguel Ángel Cano[4], PhD; Wenqi Shi[5], PhD; Ada Abaragu[1], BS; Song Zhang[5], PhD

[1]Department of Health Economics, Systems, and Policy, Peter O'Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX, United States

[2]Harold C. Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, TX, United States

[3]Department of Internal Medicine, School of Medicine, University of Texas Southwestern Medical Center, Dallas, TX, United States

[4]Department of Social and Behavioral Sciences, Peter O'Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, Dallas, TX, United States

[5]Department of Health Data Science and Biostatistics, Peter O'Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, Dallas, TX, United States

Corresponding Author:
Tammy Leonard, PhD
Department of Health Economics, Systems, and Policy, Peter O'Donnell Jr. School of Public Health, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX, United States

Abstract

Background: Despite the growing emphasis on open science and equity in research, qualitative data capturing diverse human experiences and perspectives are rarely reused beyond the original study. Increasingly, data repositories are used to make these data publicly available, but it is unclear whether these data can be effectively identified by researchers interested in secondary data analysis.

Objective: We describe a protocol for identifying and characterizing archived qualitative datasets in leading public repositories, developing an artificial intelligence–based tool to enhance qualitative data reuse, and validating that tool using existing data.

Methods: We will search 4 leading repositories to assess the scope and identifiability of existing publicly available qualitative datasets. We will subsequently build the Human Experiences and Reflections (HEARs) Archive, a directory of deidentified study data that is only accessible indirectly through the use of the HEARs Portal. The HEARs Portal will be supported by large language model–based tools using the retrieval-augmented generation framework. The artificial intelligence tools' performance will be assessed across 3 domains: relevance of identified studies, validity as evaluated by comparison with human qualitative data analysis, and robustness against the addition of irrelevant information.

Results: A preliminary review of existing data repositories has begun. The anticipated study completion date is December 31, 2026.

Conclusions: The proposed project will provide evidence regarding the existing capacity for identifying and accessing qualitative data through leading repositories. It will also provide evidence on the validity of the HEARs Data Connector for identifying and describing qualitative datasets in ways that can assist researchers interested in secondary analysis. Establishing the validity of the HEARs Data Connector and developing an evidence-based ongoing improvement and monitoring strategy will be essential for establishing trust within the qualitative research community.

International Registered Report Identifier (IRRID): PRR1-10.2196/82725

(JMIR Res Protoc 2026;15:e82725) doi:10.2196/82725

KEYWORDS

LLM; large language models; qualitative research; secondary data analysis; information science; transcripts

XSL•FO
RenderX

Introduction

Background

Most scientific research conducted in the health, social, economic, and behavioral sciences, and related fields is aimed at improving human welfare. As public demand for "evidence-based" solutions continues to grow, there is often an implicit emphasis on quantitative evidence. However, qualitative data—data derived from transcribed conversations, open-ended survey questions, or original text from archival materials [1]—are essential for keeping research grounded in real-world experiences and responsive to social and contextual challenges [2]. Qualitative methods can lead to the discovery of determinants that influence behavior, help enhance the scientific rigor of research, facilitate the translation of observational data into interventions, and improve the likelihood that interventions implemented are sustainable and acceptable in the target population [3-7]. Much of the evidence base for programs to improve human welfare is derived from survey or administrative data that are coded numerically and then analyzed using statistical methods. However, data only exist to measure things that some researcher or administrator chose to collect [8]. Qualitative data are the basis of information from which those choices can be best informed. Importantly, qualitative data often return unexpected responses to a question because respondents are not told a priori what response possibilities the researcher has in mind. Qualitative data are necessary to inform accurate interpretation of quantitative results, develop hypotheses and research questions regarding the populations of interest, and improve the relevance of evidence-based solutions—especially those created to serve populations whose perspectives differ from those of the mainstream.

Importance of Qualitative Data Sharing and Reuse

Qualitative data are costly to collect, and economies of scale for qualitative data collection are less attainable compared to quantitative data. As a result, there are unique economic barriers to increasing the use and impact of qualitative data. Each new instance of interview or focus group data collection requires the same additional researcher time to administer, record, and transcribe the data. In contrast, once quantitative data collection instruments are developed, they can generally be scaled to collect large sample sizes with little additional cost beyond participant recruitment. Despite numerous calls for greater data sharing and reuse and the existence of many large data repositories, secondary analysis of qualitative data is extremely rare. This is regrettably inefficient considering the potential benefits that can come from the reuse of qualitative data and the high cost of primary qualitative data collection.

Increased capacity for sharing and combining qualitative data from multiple studies has been identified as an important goal for improving research reliability and relevance. The Joint Declaration of Data Citation Principles highlights this prioritization within the scientific community by making greater access to data for the purpose of secondary analysis a critical goal [9]. As with quantitative data, data sharing is essential for ensuring that analysis is reproducible. However, this concern is heightened in the case of qualitative data, where analysis can be more heavily influenced by the skill set and perspective of the researcher [10]. Could a different qualitative researcher interpret the same dataset differently? Secondary analysis of qualitative data is also important for facilitating faster dissemination links, whereby new research questions emerge from existing data. Could a researcher from a different disciplinary background analyze the same dataset but extract themes that inform important ongoing work different from the original study? Meta-analysis of qualitative datasets and meta-synthesis of qualitative data findings have shown promise for identifying new research evidence from existing data, but they have largely failed to reach their potential because of the difficulty in identifying and accessing related qualitative studies [11,12]. Finally, qualitative data, when shared, can allow future researchers to identify the early developments of new social challenges that were not even being considered by researchers at the time of data collection [8].

Challenges to Qualitative Data Sharing and Reuse

Data repositories are increasingly being used to deposit qualitative data for reuse, and some are being created for the sole purpose of facilitating the sharing of qualitative data. However, there remain gaps between the vision (easy sharing and robust reuse of qualitative data) and the reality (sharing is possible, but reuse is infrequent). First, qualitative data are not easy to identify within repositories that also contain quantitative data. Repositories do not generally provide filters that allow searchers to distinguish between quantitative and qualitative data types. Second, across all repository types, qualitative data are often stored behind firewalls to protect the confidentiality of human participants, and institutional review board (IRB) approval is often required just to inspect the data and assess its suitability for a planned research project. Third, guidelines for sharing qualitative data through repositories are varied and often lacking altogether. In a review of 32 English-language multidisciplinary or social science data repositories, only 12 had qualitative data sharing guidelines, and most guidelines remained vague regarding the removal of indirect identifiers [1]. Fourth, when qualitative data can be obtained, it requires additional time inputs to review and assess whether there is sufficient content related to desired themes to warrant use in secondary analysis.

Opportunities related to the better use of qualitative data are being explored. ChatGPT has been shown to perform "reasonably well" in identifying themes within qualitative datasets, but has some noted shortcomings when themes are subtle or require nuanced interpretations [13-15]. However, artificial intelligence (AI) has not been examined as a tool specifically adapted to engage in meta-analysis of qualitative datasets while incorporating core methodological insights of the qualitative field—especially through the development of appropriate metadata to prompt the AI regarding the unique study context and nuance. Researchers are also engaging in greater sharing of qualitative data by building large open-access qualitative datasets to improve monitoring of the evolution of social processes and outcomes. For example, the American Voices Project is building a large, repeated cross-sectional repository of immersive interviews from across the United States.

XSL•FO
RenderX

The Human Experiences and Reflections Data Connector

The Human Experiences and Reflections (HEARs) Data Connector will fill existing gaps to improve qualitative data sharing and reuse capacity. Importantly, the HEARs Data Connector is intended to complement existing repositories and collaborate with the services they provide. For the purposes of this protocol, we focus on qualitative data collected within research studies, which most often includes transcripts from interviews, focus groups, or ethnographic observations. We also limit our work to English-language text, although we hope to remove this restriction in future studies. The HEARs Data Connector consists of 2 parts: the HEARs Archive and the HEARs Portal.

The HEARs Archive will house deidentified qualitative data, serving as a foundational knowledge base for the HEARs Portal. Unlike traditional data repositories, it will not provide users with direct access to the underlying data. Instead, it will support the Portal's functionality by enabling data-informed tools and resources. The HEARs Data Connector intends to provide a "middle ground" solution to address the complex challenges related to qualitative data sharing. Deidentified qualitative data can inform summary information accessed by researchers through the HEARs Portal, but direct access to the data itself must be obtained through collaboration with the original investigator or their designee, or by accessing data in an existing public repository. Thus, it allows for data discovery and facilitates reuse, while allowing for heterogeneous, context-specific approaches to ethical data sharing.

The HEARs Portal will be powered by AI and will allow users to perform advanced queries of the HEARs Archive. These queries will generate summaries of key themes within the archive and identify the studies that contributed to each thematic result. To support transparency and reproducibility, the HEARs Portal will display links to the specific studies that contributed to each thematic summary, enabling users to review the source material and assess the analytic pathway. The Portal will provide sufficient detail to support hypothesis generation, guide the selection of qualitative datasets for secondary analysis, and provide information on how to request access to those datasets, such as through links to existing repositories.

The purpose of this protocol is to describe the initial development of the HEARs Data Connector. We will describe the planned foundational research to establish the availability of existing data that can be used within HEARs, as well as the development of the first HEARs demonstration project and its associated performance evaluation.

Methods

Study Design

The study protocol is organized into 2 primary phases. Phase 1 encompasses the development and curation of the HEARs Archive, a specialized knowledge base of qualitative research data. This phase involves identifying and acquiring relevant datasets from established repositories, securing investigator permissions, and applying a rigorous preprocessing and deidentification workflow. Phase 2 details the design, implementation, and validation of the HEARs Portal, a web-based tool powered by large language models (LLMs) for querying the HEARs Archive. Specifically, the HEARs Portal allows researchers to query the HEARs Archive, identify related studies from a large sample of qualitative datasets, conduct preliminary descriptive meta-analyses of the datasets in response to queries, and obtain descriptions of the available datasets related to the query (eg, instructions for accessing the data from the investigator or their designee, publications associated with the dataset, study setting, and context). The HEARs Portal will be powered by LLMs with a hybrid retrieval-augmented generation (RAG) framework, whose performance is optimized through synergistic metadata and prompt design. We will assess the performance of the HEARs Portal across 3 domains: relevance of identified studies in the HEARs Archive, validity by comparison with human qualitative data analysis, and robustness against the addition of irrelevant information within the HEARs Archive.

Patient and Public Involvement

We do not plan for direct patient or public involvement in the designing, conducting, or reporting of this project because its main objective is to develop the HEARs Data Connector, which is primarily a technical task. However, patient and public involvement is anticipated in subsequent projects using and refining the tool for applied use.

Phase 1: HEARs Archive

Sample Selection

We have identified 4 leading US-based data repositories that are likely to contain qualitative datasets: Open Science Framework, Inter-university Consortium for Political and Social Research, Qualitative Data Repository, and Harvard Dataverse. These repositories are maintained by nonprofit organizations (The Center for Open Science) or academic institutions (University of Michigan, Syracuse University, and Harvard University, respectively) and have been built to promote data sharing among academic researchers. We will first meet with representatives from each repository to understand how best to identify qualitative datasets and use the repositories' query tools. Next, we will attempt to identify all qualitative datasets available within each repository. From the available studies, we will attempt to identify clusters of topically related studies and select one or more clusters from which to develop and test the HEARs Portal. The goal of identifying clusters is simply to drive efficiency in our initial development of the HEARs Data Connector; thus, we will not use a strict methodology in identifying clusters. We simply seek to find topical areas where there are multiple related studies, and we will prioritize adding studies from these topical areas to the HEARs Archive first to begin development and testing of the portal regarding how it uses information from related studies. We will populate the HEARs Archive with deidentified qualitative datasets, including studies from the identified clusters. We will prioritize populating the HEARs Archive with studies from any repository that are topically related to the cluster that was identified. Before including any study in the HEARs Archive—even studies that are publicly available for download—we will seek permission

XSL•FO
RenderX

from the study's principal investigator and obtain evidence that the data were collected under the supervision of an IRB.

Preprocessing and Deidentification

To prepare the raw data for the HEARs Portal, each transcript will undergo a systematic preprocessing and curation workflow to transform it into an indexed, queryable record within the HEARs Archive. First, to ensure participant confidentiality, all transcripts will be deidentified using a Python-based anonymization pipeline and named entity recognition models to automatically detect and redact direct identifiers (eg, names, locations, and so on) and quasi-identifiers. We will test the success of anonymization by manually reviewing a 25% sample of the anonymized data.

Repository Structure

Each separate file within the repository will be labeled with a 2-part identifier: study ID and record ID. The 7-digit study ID will be unique to each study. The 4-digit record ID will be unique within each study ID. Metadata will then be created for each 2-part identifier. This will allow metadata to identify both characteristics that apply to an entire study (eg, study settings and study authors) as well as characteristics unique to a particular study record (eg, race or ethnicity of the participant). The repository data will be stored in a secure, encrypted, and password-protected drive in compliance with IRB requirements.

Metadata fields are described in Table 1. Study-level metadata fields include those describing the overall study, the methods used to collect the data, the type and size of study data, and publications related to the study data. Record-level metadata fields describe the record and the characteristics of the participants who provided information for the record. Study-level metadata will be made available so that users of the HEARs Portal can know the full list of studies that are being used. We will implement Schema.org/Dataset via JSON-LD to ensure indexing of metadata can be completed by general search engines like Google Dataset Search, and we will also align our metadata schema with the Data Documentation Initiative Codebook or Lifecycle standard for deep research interoperability. Additionally, we will map our metadata fields (Table 1) by implementing a hybrid ontology mapping that uses the European Language Social Science Thesaurus and MeSH (Medical Subject Headings). Furthermore, the named entity recognition models used in our preprocessing pipeline will be tuned to recognize and align entities with these ontologies during the indexing phase.

XSL•FO
RenderX

Table . Human Experiences and Reflections (HEARs) Archive metadata fields.

Field category and name	Field level	Example[a]
Study description		
Study ID	Study	7801004
Study principal investigator	Study	John Smith
Study title	Study	Caregiver impact on adherence to discharge instructions
Study purpose or summary	Study	Following an acute illness, patients may be overwhelmed with discharge instructions. This study explores the ability of patients to understand and adhere to instructions upon hospital discharge. A qualitative design of interviews with 63 patients and caregivers was used. Data were collected within 3 weeks of discharge from hospital.
Subject	Study	Medicine, health, and life sciences
Institutional review board of record	Study	University of Texas Southwestern
Study methods		
Study design	Study	Randomized controlled trial
Mode of data collection	Study	Face-to-face interview
Recruitment method	Study	Random sample
Time period	Study	2016-2017
Geographic coverage	Study	Alabama, United States
Study data		
Number of participants in study	Study	126
Number of records	Study	63
Data types	Study	Coded qualitative data
Study publications		
Related publication(s)	Study	Smith, John. 2019. Protocol for Caregiver Impact on Adherence to Discharge Instructions. *Protocol Journal*. 54(2), p111-123
Repository archive	Study	Qualitative data repository
Repository archive identifier	Study	https://doi.org/12.7066/F9MKT6G
Contact for more information	Study	Please refer to instructions at repository.
Record characteristics		
Record identifier	Record	5460
Number of participants in record	Record	2
Data collection method	Record	Interview
Year of data collection	Record	2017
Participant race or ethnic composition		
Non-Hispanic Black	Record	50%
Non-Hispanic White	Record	50%
Asian	Record	0%
Hispanic	Record	0%
American Indian or Alaskan Native	Record	0%

XSL•FO
RenderX

Field category and name	Field level	Example[a]
Native Hawaiian or Other Pacific Islander	Record	0%
Participant gender		
Man	Record	50%
Woman	Record	50%
Nonbinary	Record	0%
Participant age (y)		
0 - 18	Record	0
19 - 65	Record	100
>65	Record	0

[a]The example provided is not a real study.

Phase 2: HEARs Portal Design

Overview

We will develop the HEARs Portal using a hybrid RAG framework to improve the accuracy and reliability of LLMs for analyzing the HEARs repository [16]. We will leverage the application programming interface–based GPT-series LLMs from the Health Insurance Portability and Accountability Act (HIPAA)-compliant Microsoft Azure OpenAI service as backbone LLMs. The user workflow is explicitly designed as two linked stages: (1) a hybrid retrieval process to identify relevant information and (2) a contextualized generation process to synthesize that information into a coherent response.

The HEARs Portal will identify relevant studies within the repository, summarize key concepts and themes related to the research question contained within these studies, and provide a list of related questions that may be further investigated using the available data in the repository. Users will then have the ability to engage in translation by repeating this process to query competing interpretations and further related questions. With each response, the HEARs Portal will provide a complete list of studies from which the results were obtained, along with summary metadata about the study elements contributing to the response. Importantly, the HEARs Portal will provide a summary of the range of responses to particular questions or thematic analyses alongside the frequency of elements providing support for each response.

Stage 1: Hybrid Information Retrieval

We will first create a web interface for querying the AI tool and provide instructions and examples to guide effective use [17]. The instructions will guide users in iteratively using the AI tool in alignment with established methods for meta-analysis of qualitative data steps [11]. When a user submits a research query, the Portal initiates a sophisticated retrieval process designed to locate the most relevant information from the HEARs Archive. To maximize both precision and comprehensiveness, we will use a hybrid retrieval strategy that combines 2 distinct modes of search. First, in explicit search (metadata filtering), users can leverage the rich, structured metadata to perform targeted database-style queries. This allows researchers to narrow the search space to studies or records with specific characteristics (eg, “all interviews with female participants over the age of 65 years from randomized controlled trials”). This acts as a powerful preliminary filter. Second, in implicit search (content-based retrieval)**,** to search the textual content of the transcripts (especially when metadata is incomplete), the Portal will use a combination of state-of-the-art retrieval models to ensure that the results are relevant both lexically and semantically. A sparse retriever, such as a term-frequency model like BM25 [18], will identify documents based on keyword overlap, excelling at matching specific terms. Concurrently, a dense retriever built from text-embedding-3-large embeddings (deployed in our secure Azure environment), with candidate lists fused using reciprocal rank fusion and then reranked using the bge-reranker-large cross-encoder model (Beijing Academy of Artificial Intelligence), will use a semantic similarity model like DRAGON [19] to generate vector embeddings to retrieve passages that are contextually related to the query, even without direct keyword matches. The outputs from both retrievers will be fused and reranked to create an optimized context for the generation stage.

Stage 2: Contextualized Generation

In the generation stage, the retrieved text chunks and their associated metadata are fed to an LLM, which synthesizes the information into a coherent, evidence-based narrative. To manage hallucination and ensure traceable attribution, generation will be strictly retrieval-gated with versioned prompt templates that instruct the model to use only the provided evidence and to abstain when evidence is insufficient. Pydantic models will be used to enforce a strict JSON schema on responses, including (1) a theme or summary section, (2) a complete list of contributing study IDs, and (3) chunk-level evidence pointers for each major claim. We will implement automated “no phantom citation” checks that verify every cited study or chunk is present in the retrieved context, and we will treat retrieved transcript text as untrusted input (explicitly disabling instruction-following from the corpus) to reduce prompt-injection risk. The Portal will surface citations as study-level identifiers with repository links and an auditable mapping from each synthesized theme to the supporting studies or records, enabling users to verify provenance and supporting evidence.

XSL•FO
RenderX

While the procedures described are the current most plausible baseline design choices for each component, we recognize that these are inherently empirical decisions; accordingly, we will systematically test reasonable alternatives and select the configuration that performs best on our expert-labeled validation benchmarks.

Data Analysis

We will validate the HEARs Portal's performance across several areas: identification of relevant studies, agreement with human analysis of qualitative data, and stability in response to the addition of irrelevant information. First, we will validate that the Portal is able to accurately identify relevant studies and return accurate summaries of the metadata from those studies. To do this, we will assess the accuracy of the retrieval system by comparing its output against a gold-standard evaluation set created by human domain experts. For a diverse set of test queries, we will use standard information retrieval metrics, including precision@k, recall@k, and mean reciprocal rank, to quantify the system's ability to retrieve relevant documents within the top-ranked results. We will also query the HEARs Portal on subject matter not included in the Archive and test whether the HEARs Portal is able to accurately assess when there is no relevant information available for a given query. Next, we will test for agreement with thematic analysis of qualitative data. We will use peer-reviewed and published studies recording primary analysis of datasets in the HEARs Archive and assess how well the Portal, when examining only a single study, is able to reproduce the results. We will also conduct a meta-analysis of related studies published in the peer-reviewed literature and compare that analysis to results returned from the HEARs Portal when examining data from all studies represented in the meta-analysis. Finally, we will examine the stability of the HEARs Portal by exposing it incrementally to additional studies contained within the full HEARs Archive and repeating the analysis above. We will assess the extent to which responses to queries change when seemingly unrelated, irrelevant studies are added to the corpus of materials examined by the Portal.

Reproducibility and Sharing

We will develop and maintain Docker containers [20] that capture the complete runtime environment, including all dependencies (Python packages, embedding models, and retrieval frameworks), library versions, and configurations. We will make Docker images available through a public registry (eg, Docker Hub) alongside version-controlled source code. This will enable users and reviewers to verify and reproduce our validation experiments.

Each query of the HEARs portal will generate a persistent ID. The persistent ID will capture the exact query parameters, the specific version of the HEARs Archive queried, and the AI model version used. This will ensure that researchers can cite the specific synthesis provided by the HEARs Portal and thus allow for verification of the analytic pathway.

Ethical Considerations

This study has been approved by the IRB at the University of Texas Southwestern Medical Center (STU-2024-1045). We have received a waiver of informed consent for our analysis of secondary data. All data will be deidentified and anonymized prior to inclusion in the HEARs archive. We are not enrolling any participants for primary data collection; hence, there is no participant compensation in this study.

Dissemination Plan

Study results will be disseminated via peer-reviewed studies and presentations at academic conferences. We anticipate at least 2 studies from this work: the first reporting on the availability and ease of accessing qualitative data from existing repositories, and the second reporting on the validation of the HEARs AI tool.

The HEARs Data Connector will become publicly available through a web portal. Users will be required to register prior to using HEARs. We will also work to disseminate our efforts in collaboration with the leadership of other large repositories. Since the HEARs repository will not allow users to access the raw data, we aim to form collaborative relationships with other established repositories so that HEARs can provide incremental benefits by allowing the discovery of qualitative data without redundantly taking on the task of archiving the data for others to access.

Following the initial development and evaluation of the HEARs portal, we will consult with a broad range of stakeholders (eg, repository representatives, researchers, librarians, and funding agency representatives) to discuss our findings and identify next steps for improving HEARs in future projects. We anticipate some of these areas might include the development of strategies to improve the quantity and quality of data in the HEARs archive, investments to improve the sustainability of HEARs, and improvements to metadata and AI tools. Currently, HEARs will be hosted and maintained locally at the University of Texas Southwestern Medical Center; findings from this protocol will inform recommendations for maintaining or adjusting this plan.

Results

As of February 6, 2026, we have completed work to establish protocols for the identification and processing of studies for inclusion in HEARs. Specifically, we have developed and tested procedures for searching for qualitative studies in existing archives (Open Science Framework, Inter-university Consortium for Political and Social Research, Qualitative Data Repository, and Harvard Dataverse), and we have met with representatives from each of the archives to obtain input on our procedures. We have also developed protocols for receiving studies into the HEARs portal. Additionally, we have developed tools in REDCap (Research Electronic Data Capture, Vanderbilt University) to collect metadata about each study, developed and tested an AI tool to deidentify qualitative data, and piloted all of these processes across 2 existing studies conducted by collaborators of our core research team. We have also begun work to identify qualitative studies available in existing repositories. We anticipate publishing the results of that formative investigation in Fall 2026.

XSL•FO
RenderX

Discussion

The HEARs Data Connector offers the potential to improve the scientific knowledge creation system by making it easier for existing qualitative data to be reused to inform new study development, in secondary analysis to obtain additional insights, or to provide context in the interpretation of quantitative results when qualitative data were not collected. Because qualitative data are the foundation upon which decisions regarding the scope of quantitative data collection (eg, what survey questions to ask) are made and the interpretation of quantitative data analysis is undertaken, the benefits of being able to reuse existing qualitative data more easily are likely significant. Anticipated findings of the project described in this protocol include the creation of a pilot version of the HEARs Data Connector that can be used to demonstrate its effectiveness in improving capacity for qualitative data reuse.

Obtaining data to populate the HEARs Archive is essential for its long-term success. Our work will illuminate potential gaps in the existing practice of qualitative data sharing as we explore the capacity for including studies from existing repositories into the HEARs Archive. The work is likely to inform new collaborative approaches between researchers and repository managers that can improve the efficiency of secondary analysis of qualitative data. It might also inform ongoing questions regarding the refinement of methods for the secondary analysis of qualitative data [21].

Our project, if successful, will demonstrate an innovative use of AI in qualitative research. Numerous ongoing studies are attempting to use AI to speed up the primary analysis of qualitative data and have identified challenges that the field is working to overcome [22-24]. Our use of AI is more limited in scope and perhaps more suited to the strengths of current-generation AI tools [25], in that we are not attempting to analyze qualitative data with the HEARs Portal, but rather to identify and describe qualitative data resources that can be used in human- or machine-based analysis.

Validation of the initial implementation of the HEARs Data Connector will be essential for establishing efficacy and informing future improvements to the Data Connector. The HEARs Data Connector's long-term success will largely be determined by the willingness of researchers and repository leaders to adopt the tool [26]. While guidelines for evaluating the ethical and scientifically grounded use of AI for the novel application we propose are not available, our results can be compared to established principles related to AI use in other contexts [27,28]. Establishing the validity of the tool's output and developing an evidence-based, ongoing improvement and monitoring strategy will be essential for establishing trust within the qualitative research community.

Acknowledgments

No generative artificial intelligence tools were used in the production of this manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' Contributions

TL, MÁC, JPS, WS, and SZ contributed to the conceptual design of the study, as well as the writing and reviewing of the manuscript. AA contributed to reviewing the manuscript and designing the sample selection protocol.

Conflicts of Interest

None declared.

References

1. Antes AL, Walsh HA, Strait M, Hudson-Vitale CR, DuBois JM. Examining data repository guidelines for qualitative data sharing. J Empir Res Hum Res Ethics 2018 Feb;13(1):61-73. [doi: 10.1177/1556264617744121] [Medline: 29226747]
2. Shelton RC, Philbin MM, Ramanadhan S. Qualitative research methods in chronic disease: introduction and opportunities to promote health equity. Annu Rev Public Health 2022 Apr 5;43(37-57):37-57. [doi: 10.1146/annurev-publhealth-012420-105104] [Medline: 34936827]
3. Qualitative methods in implementation science. : National Cancer Institute; 2018 URL: https://cancercontrol.cancer.gov/IS/docs/NCI-DCCPS-ImplementationScience-WhitePaper.pdf [accessed 2025-07-21]
4. Castro FG, Kellison JG, Boyd SJ, Kopak A. A methodology for conducting integrative mixed methods research and data analyses. J Mix Methods Res 2010 Sep 20;4(4):342-360. [doi: 10.1177/1558689810382916] [Medline: 22167325]
5. Busetto L, Wick W, Gumbinger C. How to use and assess qualitative research methods. Neurol Res Pract 2020;2(1):14. [doi: 10.1186/s42466-020-00059-z] [Medline: 33324920]
6. Yardley L, Bradbury K, Morrison L. Using qualitative research for intervention development and evaluation. In: Qualitative Research in Psychology: Expanding Perspectives in Methodology and Design: American Psychological Association; 2021:263-282. [doi: 10.1037/0000252-013]

XSL•FO
RenderX

7. Duggleby W, Williams A. Methodological and epistemological considerations in utilizing qualitative inquiry to develop interventions. Qual Health Res 2016 Jan;26(2):147-153. [doi: 10.1177/1049732315590403] [Medline: 26063607]
8. Edin KJ, Fields CD, Grusky DB, et al. Listening to the voices of America. RSF 2024 Sep;10(4):1-31. [doi: 10.7758/RSF.2024.10.4.01]
9. Fenner M, Crosas M, Grethe JS, et al. A data citation roadmap for scholarly data repositories. Sci Data 2019 Apr 10;6(1):28. [doi: 10.1038/s41597-019-0031-8] [Medline: 30971690]
10. Nuzzo R. How scientists fool themselves—and how they can stop. Nature 2015 Oct 8;526(7572):182-185. [doi: 10.1038/526182a] [Medline: 26450039]
11. McCormick J, Rodney P, Varcoe C. Reinterpretations across studies: an approach to meta-analysis. Qual Health Res 2003 Sep;13(7):933-944. [doi: 10.1177/1049732303253480] [Medline: 14502959]
12. Vogel JA, Rising KL, Jones J, Bowden ML, Ginde AA, Havranek EP. Reasons patients choose the emergency department over primary care: a qualitative metasynthesis. J Gen Intern Med 2019 Nov;34(11):2610-2619. [doi: 10.1007/s11606-019-05128-x] [Medline: 31428988]
13. Wachinger J, Bärnighausen K, Schäfer LN, Scott K, McMahon SA. Prompts, pearls, imperfections: comparing ChatGPT and a human researcher in qualitative data analysis. Qual Health Res 2025 Aug;35(9):951-966. [doi: 10.1177/10497323241244669] [Medline: 38775392]
14. Morgan DL. Exploring the use of artificial intelligence for qualitative data analysis: the case of ChatGPT. Int J Qual Methods 2023 Oct;22:16094069231211248. [doi: 10.1177/16094069231211248]
15. De Paoli S. Performing an inductive thematic analysis of semi-structured interviews with a large language model: an exploration and provocation on the limits of the approach. Soc Sci Comput Rev 2024 Aug;42(4):997-1019. [doi: 10.1177/08944393231220483]
16. Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. Presented at: NIPS'20: Proceedings of the 34th International Conference on Neural Information Processing Systems; Dec 6-12, 2020 URL: https://dl.acm.org/doi/abs/10.5555/3495724.3496517 [accessed 2026-03-20]
17. Zhang H, Wu C, Xie J, Lyu Y, Cai J, Carroll JM. Harnessing the power of AI in qualitative research: exploring, using and redesigning ChatGPT. Comput Hum Behav 2025 May;4:100144. [doi: 10.1016/j.chbah.2025.100144]
18. Robertson S, Zaragoza H. The probabilistic relevance framework: BM25 and beyond. Found Trends Inf Retr 2009;3(4):333-389. [doi: 10.1561/1500000019]
19. Lin SC, Asai A, Li M, et al. How to train your DRAGON: diverse augmentation towards generalizable dense retrieval. Presented at: Findings of the Association for Computational Linguistics: EMNLP 2023; Dec 6-10, 2023. [doi: 10.18653/v1/2023.findings-emnlp.423]
20. Docker. URL: https://www.docker.com/ [accessed 2026-04-07]
21. Sharp EA, Munly K. Reopening a can of words: qualitative secondary data analysis. J Fam Theory Rev 2022 Mar;14(1):44-58. [doi: 10.1111/jftr.12443]
22. Cook DA, Ginsburg S, Sawatsky AP, Kuper A, D'Angelo JD. Artificial intelligence to support qualitative data analysis: promises, approaches, pitfalls. Acad Med 2025 Oct 1;100(10):1134-1149. [doi: 10.1097/ACM.0000000000006134] [Medline: 40560241]
23. Christou PA. Reliability and validity in qualitative research revisited and the role of AI. Qual Rep 2025;30(3):3306-3314. [doi: 10.46743/2160-3715/2025.7523]
24. Kabir SMA, Ali F, Ahmed RL, Sulaiman-Hill R. Exploring the use of AI in qualitative data analysis: comparing manual processing with Avidnote for theme generation. Int J Qual Methods 2025 Apr;24:16094069251336810. [doi: 10.1177/16094069251336810]
25. Amer E, Elboghdadly T. The end of the search engine era and the rise of generative AI: a paradigm shift in information retrieval. Presented at: 2024 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC); Nov 13-14, 2024 URL: https://ieeexplore.ieee.org/abstract/document/10783559 [accessed 2026-03-20] [doi: 10.1109/MIUCC62295.2024.10783559]
26. Mohammadi E, Thelwall M, Cai Y, Collier T, Tahamtan I, Eftekhar A. Is generative AI reshaping academic practices worldwide? A survey of adoption, benefits, and concerns. Inf Process Manag 2026 Jan;63(1):104350. [doi: 10.1016/j.ipm.2025.104350]
27. Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. Nature New Biol 2023 Jan 26;613(7945):612. [doi: 10.1038/d41586-023-00191-1] [Medline: 36694020]
28. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell 2019;1(9):389-399. [doi: 10.1038/s42256-019-0088-2]

Abbreviations

AI: artificial intelligence
HEARs: Human Experiences and Reflections
HIPAA: Health Insurance Portability and Accountability Act

XSL•FO
RenderX

IRB: institutional review board
LLM: large language model
MeSH: Medical Subject Headings
RAG: retrieval-augmented generation
REDCap: Research Electronic Data Capture

Edited by J Sarvestan; submitted 20.Aug.2025; peer-reviewed by D Verran, S Lachmuth; revised version received 24.Feb.2026; accepted 25.Feb.2026; published 21.Apr.2026.

Please cite as:
Leonard T, Stimpson JP, Cano MÁ, Shi W, Abaragu A, Zhang S
Use of an AI-Based Tool (Human Experiences and Reflections Data Connector) to Improve Discovery and Reuse of Archived Qualitative Data: Protocol for an Algorithm Development and Validation Study
JMIR Res Protoc 2026;15:e82725
URL: https://www.researchprotocols.org/2026/1/e82725
doi:10.2196/82725

© Tammy Leonard, Jim P Stimpson, Miguel Ángel Cano, Wenqi Shi, Ada Abaragu, Song Zhang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Life-Course Framework for Institute-Based Surveillance of Overweight, Obesity, and Metabolic Risk Factors in India: Protocol for Surveillance System Development

Sai Ram Challa[1,2*], MBBS, MD; Biswanath Dash[1*], PhD

[1]Department of Humanities and Social Sciences, Birla Institute of Technology and Science - Hyderabad Campus, Hyderabad, India

[2]Maternal and Child Health and Nutrition, ICMR–National Institute of Nutrition (NIN), Jamai Osmania PO, Hyderabad, India

[*]all authors contributed equally

Corresponding Author:
Sai Ram Challa, MBBS, MD
Department of Humanities and Social Sciences, Birla Institute of Technology and Science - Hyderabad Campus, Hyderabad, India

Abstract

Background: India faces a complex spectrum of malnutrition, with undernutrition, micronutrient deficiencies, and rising overweight and obesity coexisting across populations and life stages. National surveys, such as the National Family Health Survey, Comprehensive National Nutrition Survey, and WHO STEPwise approach to Surveillance, provide essential population-level estimates but are episodic, limiting their utility for timely, institution-level prevention and early detection of nutrition and metabolic risks. Continuous, life course–oriented surveillance is required to support responsive public health action and double-duty approaches addressing multiple forms of malnutrition.

Objective: This study aims to describe a life course framework and study protocol for an institute-based, standardized surveillance system to monitor overweight, obesity, and metabolic risk factors in India.

Methods: This study protocol was developed through a structured synthesis of international surveillance models and public health frameworks, combined with operational insights from ongoing implementation initiatives in India. The proposed surveillance system is institution-anchored and spans key life stages from preconception to young adulthood. Core domains include anthropometry, diet, selected biochemical markers, and behavioral risk factors. Data collection is designed using a hybrid model incorporating optical mark recognition–based paper tools, digital platforms, and connected measurement devices, guided by principles of primary health care, including equity, community participation, intersectoral coordination, and appropriate technology.

Results: The surveillance protocol has transitioned from conceptual development to implementation through 2 intramural research platforms: NULRISC (sanctioned July 2024) and COPRIME (sanctioned May 2025). Formative phase activities are ongoing in Telangana across approximately 15 high schools and colleges, as well as primary health care–linked platforms, including 3 primary health centers, 2 kindergartens, and 5 primary schools. At the time of manuscript revision, 10 educational institutions (5 schools and 5 colleges) were onboarded, with expansion to 30 institutions planned. The first surveillance round is scheduled for June 2026 in Hyderabad and Medchal–Malkajgiri districts, and initial analytical outputs are anticipated in 2027.

Conclusions: An institute-based, standardized surveillance system has the potential to shift nutrition and metabolic risk monitoring in India from episodic assessment to continuous prevention. By embedding surveillance within routine institutional platforms and aligning with primary health care principles, the proposed protocol offers a feasible and sustainable approach to strengthening evidence-informed policy and practice across the life course.

International Registered Report Identifier (IRRID): DERR1-10.2196/86084

(JMIR Res Protoc 2026;15:e86084) doi:10.2196/86084

KEYWORDS

nutrition surveillance; institute-based monitoring; overweight; obesity; metabolic risk factors; malnutrition; life course; adolescents; first 1000 days; primary health care; double-duty actions; India; protocol; systems

Introduction

Overweight, obesity, and other forms of malnutrition now coexist as part of a continuum of nutritional imbalance that increasingly shapes global health profiles across the life course. Worldwide, the prevalence of overweight and obesity has nearly tripled since 1975, with more than 1 billion individuals affected in 2022, including an estimated 159 million children and

XSL•FO
RenderX

adolescents [1]. This growing burden contributes substantially to noncommunicable diseases (NCDs) and interacts with persistent undernutrition and micronutrient deficiencies, particularly in low- and middle-income countries.

India exemplifies this complex epidemiological reality through what has been described as a "triple burden of malnutrition," characterized by the simultaneous presence of undernutrition, micronutrient deficiencies, and rising overweight and obesity within the same populations, communities, and even households [2,3]. This coexistence poses distinct public health challenges, as early-life undernutrition and later-life excess energy intake often intersect to amplify cardiometabolic risk, perpetuating intergenerational cycles of poor health and inequity.

The concept of nutrition transition, first articulated by Popkin [4], provides a useful framework for understanding these shifts. Rapid economic growth, urbanization, and globalization have altered dietary patterns and physical activity environments in India, leading to increased consumption of refined grains, added sugars, edible oils, and ultra-processed foods, alongside declining intake of fruits, vegetables, and pulses [5]. These changes have been accompanied by reduced physical activity and increased sedentary behavior, contributing to a rising prevalence of obesity and diet-related NCDs, such as hypertension, type 2 diabetes, cardiovascular disease, and certain cancers. At the same time, undernutrition continues to predispose vulnerable groups to infections, low birth weight, stunting, and impaired cognitive development [6].

Evidence from life course epidemiology underscores that nutritional and behavioral exposures during early and transitional stages of life have long-term consequences for metabolic health. The period spanning preconception, pregnancy, and the first 1000 days is particularly critical for metabolic programming and future disease risk [7]. Subsequent stages—early childhood, school age, adolescence, and young adulthood—represent additional windows during which dietary patterns, physical activity, sleep, and other behaviors consolidate. The surveillance of nutrition and metabolic risk factors across these stages is therefore essential for identifying early deviations in energy balance, body composition, and lifestyle before they become entrenched and translate into adult NCDs [8].

India currently relies on several large-scale nutrition and health surveys, including the National Family Health Survey (NFHS), the Comprehensive National Nutrition Survey (CNNS), ICMR-INDIAB, and WHO STEPS, to generate population-level estimates of nutritional status and risk factors. While these surveys provide invaluable national benchmarks, they are largely cross-sectional and episodic, typically conducted at intervals of 5 years or longer. As a result, they lack the frequency, institutional linkage, and life-stage continuity required for timely local action, program monitoring, or rapid course correction. Moreover, substantial delays between data collection and dissemination—often exceeding 2 to 3 years—further limit their utility for preventive decision-making at institutional or district levels [9,10]. In contrast, infectious disease surveillance systems, such as India's Integrated Disease Surveillance Programme, demonstrate that continuous, institution-anchored data flows are operationally feasible and can support timely public health responses [11].

Addressing obesity and other nutrition-related NCDs in India, therefore, requires a complementary surveillance approach that moves beyond episodic surveys toward routine, institute-based monitoring across the life course. Institutions such as kindergartens, Anganwadi centers, schools, colleges, and primary health centers offer stable platforms with defined populations, governance structures, and recurring contact points that are well suited for periodic surveillance. Embedding standardized surveillance within these settings can leverage existing national programs, such as Rashtriya Bal Swasthya Karyakram (RBSK), Rashtriya Kishor Swasthya Karyakram (RKSK), POSHAN 2.0, PM POSHAN, and the Ayushman Bharat School Health Programme, while also enabling the participation of private educational institutions to improve representativeness and equity.

In this context, this study describes a protocol for a life course, institute-based standardized surveillance system for overweight, obesity, and metabolic risk factors in India. The proposed protocol emphasizes appropriate and context-sensitive technology, including optical mark recognition (OMR)–based paper tools for school settings and digital or Internet of Things (IoT)-enabled platforms where feasible, to ensure scalability, data quality, and ethical oversight. Importantly, the protocol is informed by practical operational experience from ongoing intramural implementation initiatives at ICMR-National Institute of Nutrition, which provide early insights into feasibility, indicator selection, and system integration.

Despite the availability of multiple data sources, India currently lacks a coherent and sustainable framework that systematically links institutional screening, surveillance, and preventive action across life stages. Establishing an institute-based surveillance ecosystem offers the opportunity to transform fragmented data into continuous, actionable intelligence. Such a system can support the early detection of risk, enable double-duty actions addressing multiple forms of malnutrition, and strengthen evidence-informed policy and practice across the life course.

Methods

Study Design and Purpose

This study presents a study protocol for a life course, institute-based standardized surveillance system to monitor overweight, obesity, and metabolic risk factors in India. The proposed system is designed as a continuous, periodic, and institution-anchored surveillance mechanism rather than a one-time survey or intervention trial. Its primary purpose is to generate timely, actionable data across key life stages to support early risk identification, preventive counseling, program monitoring, and policy feedback.

The protocol emphasizes routine surveillance embedded within existing institutional platforms, including kindergartens, Anganwadi centers, schools, colleges, primary health centers, and selected workplaces, to complement large-scale national surveys by providing higher-frequency, institute-level intelligence. The design is informed by principles of primary

XSL•FO
RenderX

health care (PHC), including equity, community participation, intersectoral coordination, and appropriate use of technology.

Development of the Conceptual Framework

The conceptual framework underpinning this surveillance protocol was developed through a structured synthesis of international surveillance models, relevant public health theories, and operational experience from ongoing implementation projects in India. Given the complex and coexisting forms of malnutrition observed across the Indian population, a multitheoretical and systems-oriented approach was adopted.

Review of Existing Surveillance Models

A targeted review of established international surveillance and cohort systems was undertaken to identify design features relevant to a life course, institute-based approach. The models reviewed included the Youth Risk Behavior Surveillance System (YRBS), the National Health and Nutrition Examination Survey (NHANES), the Environmental Influences on Child Health Outcomes (ECHO) program, the WHO European Childhood Obesity Surveillance Initiative, and the Global School Health Policies and Practices Survey [12-16].

These models informed key elements of the proposed protocol, including periodic data collection; standardized indicators; integration of anthropometric, behavioral, and biological measures; and the use of institutional platforms as units of observation and action. However, none of these systems individually addressed the dual requirements of life course coverage and routine institute-level surveillance within the Indian context, necessitating contextual adaptation and integration.

Theoretical Foundations

Five complementary theoretical perspectives were integrated to guide the design and operational logic of the surveillance system.

Nutrition transition theory provides a lens to understand the simultaneous rise in overweight and obesity alongside persistent undernutrition in India, driven by rapid dietary shifts, urbanization, and changing physical activity patterns [4,17]. Surveillance systems must therefore be capable of capturing both energy excess and nutrient deficiency within the same populations.

The life course perspective emphasizes that nutritional and behavioral exposures from preconception through young adulthood influence long-term metabolic risk. Surveillance across sequential life stages, including preconception, pregnancy, the first 1000 days, early childhood, school age, adolescence, and young adulthood, enables the early detection of risk trajectories and supports timely preventive action.

The social ecological model recognizes that health behaviors are shaped by interactions across individual, interpersonal, institutional, community, and policy levels [18]. Institute-based surveillance operates primarily at the institutional level, where schools, colleges, health facilities, and workplaces provide structured environments for measurement, feedback, and linkage to program action.

Behavior change communication models, including the health belief model and the social cognitive theory, highlight the role of awareness, perceived risk, and self-efficacy in behavior modification [19,20]. Surveillance systems that provide individualized and institution-level feedback, such as growth charts, dietary quality indicators, or physical activity summaries, can therefore function not only as monitoring tools but also as catalysts for behavioral change.

Finally, the surveillance-as-action paradigm reframes surveillance as an active public health function, in which data collection is explicitly linked to response, accountability, and program refinement [11]. Lessons from infectious disease surveillance systems, such as India's Integrated Disease Surveillance Programme, demonstrate the feasibility of decentralized, real-time reporting mechanisms that inform timely action. Extending this paradigm to nutrition and NCD risk surveillance requires designing systems that are responsive, continuous, and locally actionable.

Integrative Framework

Integrating these theoretical perspectives yields a multidimensional framework for institute-based surveillance (Figure 1). The nutrition transition theory defines the evolving drivers of risk, the life course perspective identifies critical windows for monitoring, the social ecological model situates institutions as operational nodes, behavior change communication models ensure engagement through feedback, and the surveillance-as-action paradigm embeds responsiveness and accountability. Together, these elements provide a coherent foundation for designing a technology-enabled, life course surveillance system aligned with PHC principles.

XSL•FO
RenderX

Figure 1. Conceptual framework for the institute-based standardized surveillance system of risk factors for overweight, obesity, and other forms of malnutrition in India. PHCs: primary health centers.

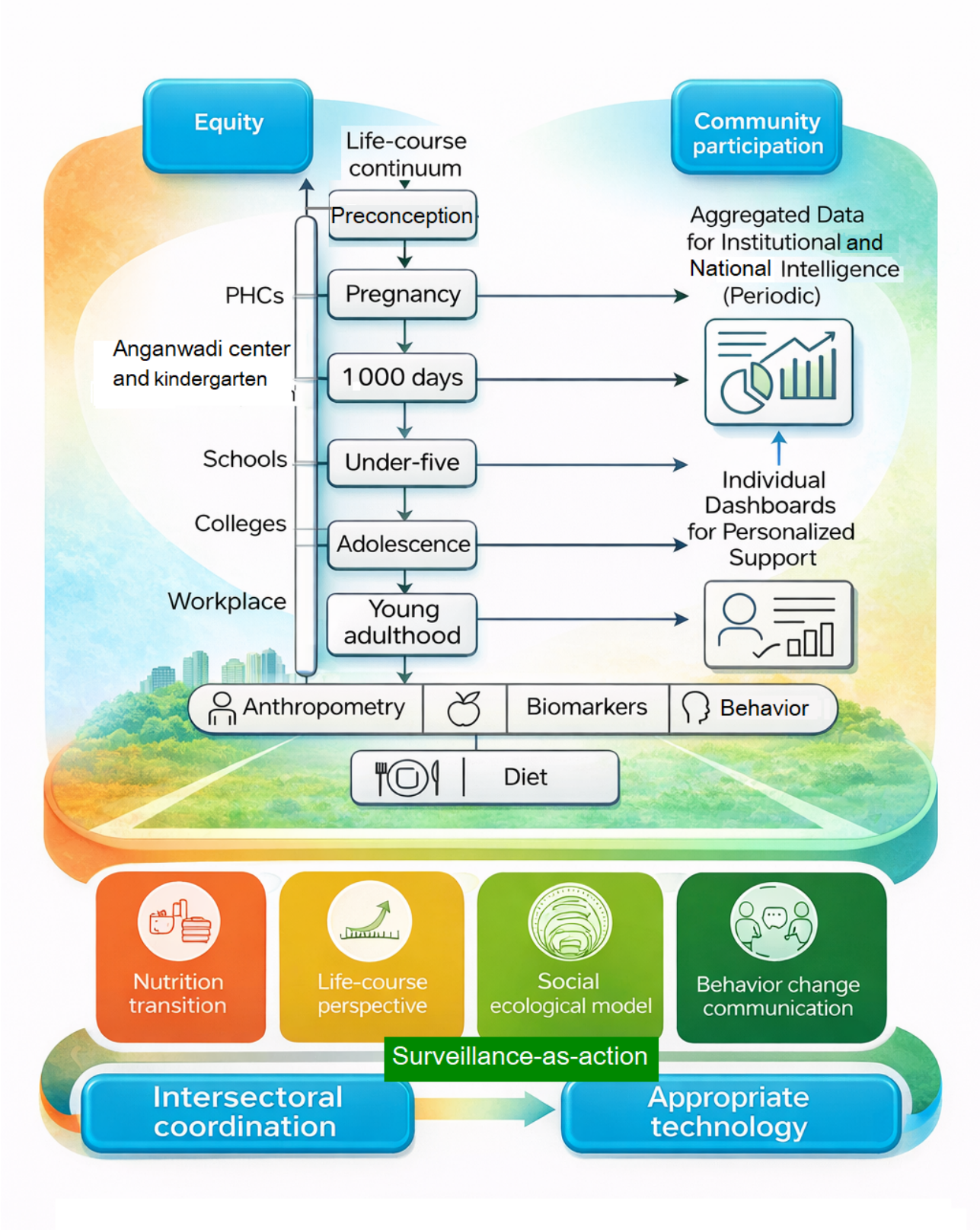

The framework integrates 5 theoretical lenses (nutrition transition theory, life course perspective, social ecological model, behavior change communication models, and surveillance-as-action paradigm) within the 4 principles of PHC (equity, community participation, intersectoral coordination, and appropriate technology). Institutions, such as Anganwadi centers, schools, colleges, and primary health care facilities, serve as sentinel nodes for the periodic collection of anthropometric, dietary, biochemical, and behavioral indicators. The data are processed through hybrid OMR, digital, and IoT technologies, feeding upward into district and national

XSL•FO
RenderX

dashboards, while feedback loops ensure local action and double-duty nutrition interventions across life stages.

Operational Refinement and Contextualization

The protocol design was further refined through practical insights from initial implementation phases of intramural projects funded by the Indian Council of Medical Research. These experiences informed decisions related to indicator selection, technology mix, data flow architecture, quality assurance mechanisms, and ethical safeguards.

No new human participant data were collected for the preparation of this study. Ethical approvals for the referenced implementation projects were obtained in accordance with the Indian Council of Medical Research guidelines.

Study Setting and Surveillance Units

India has generated substantial evidence on population nutrition and health status through large-scale surveys, such as the NNFHS, the CNNS, the National NCD Monitoring Survey, and WHO STEPS. These surveys have produced nationally representative estimates of anthropometry, dietary patterns, and behavioral risk factors and remain indispensable for benchmarking and policy planning [2,6,21,22]. However, their episodic nature—typically conducted at intervals of 5 years or longer—creates substantial temporal gaps, limiting their usefulness for continuous monitoring, institutional feedback, and timely preventive action.

In addition, existing program platforms, such as the RBSK and the RKSK, provide important opportunities for screening and counseling among children and adolescents. Nevertheless, the data generated through these platforms are largely clinical, fragmented across sectors, and seldom integrated into longitudinal nutrition or metabolic risk surveillance systems [23]. As a result, despite a strong foundation for population-level assessment, India currently lacks a routine, institute-anchored surveillance mechanism that connects educational and health institutions as recurring observation points within communities.

Limitations of Survey-Based Monitoring

While national surveys, such as the NFHS and the CNNS, are invaluable for estimating prevalence and trends at state and national levels, they are not designed to support routine institutional feedback, early detection of emerging risk patterns, or rapid program adjustment. Large sampling frames and complex field operations constrain their applicability for local decision-making, and delays between data collection, analysis, and dissemination frequently exceed 2 to 3 years [9]. Furthermore, many surveys prioritize women of reproductive age and children under 5 years, resulting in the relative underrepresentation of adolescents (10 - 19 y) and young adults (20 - 24 y)—life stages that are critical for the consolidation of dietary, physical activity, and other behavioral risk factors influencing long-term NCD risk [24].

By contrast, infectious disease surveillance systems, such as India's Integrated Disease Surveillance Programme, demonstrate that continuous, institution-based data flows are operationally feasible and can support early detection and timely response [11]. Applying similar surveillance logic to nutrition and metabolic risk factors offers an opportunity to address a major systems gap in NCD prevention.

Rationale for an Institute-Based Surveillance Approach

Institutions—including kindergartens, Anganwadi centers, schools, colleges, primary health centers, and selected workplaces—represent stable environments with defined populations, established governance structures, and recurring service or academic calendars. Anchoring surveillance within these institutions offers several operational advantages.

First, *sustainability* is enhanced by aligning data collection with routine institutional cycles, reducing dependence on externally funded survey rounds. Second, *continuity* is facilitated through repeated measurements within the same institutions, enabling longitudinal assessment of trends and early identification of deviations in anthropometry, dietary quality, or metabolic risk markers. Third, *inclusivity* is strengthened by integrating both public and private institutions, improving representation across socioeconomic groups. Finally, *actionability* is increased, as institution-level data can directly inform local health promotion activities, school meal planning, counseling services, and community interventions.

In educational settings, surveillance may use OMR-based paper tools—similar in principle to systems, such as the Youth Risk Behavior Surveillance System—allowing standardized data collection in contexts with limited digital infrastructure [25]. In health care institutions and workplaces, digital or IoT-enabled platforms can be used where feasible to support real-time data capture and integration.

Life Course Relevance of Institutional Platforms

An institute-based surveillance system is inherently life course oriented. During adolescence and young adulthood, schools and colleges provide access to large populations for monitoring body mass index, dietary patterns, physical activity, and other behavioral risk factors. In the preconception and pregnancy periods, primary health centers and Anganwadi centers serve as key surveillance nodes for indicators, such as maternal nutritional status, gestational weight gain, anemia, and gestational diabetes. For infants and young children, growth monitoring and infant and young child feeding indicators can be integrated within existing Integrated Child Development Services (ICDS) and POSHAN 2.0 platforms [26].

This institutional continuum enables the systematic documentation of nutritional exposures across successive life stages and facilitates linkage between early-life risk factors and later adolescent or adult outcomes. By embedding surveillance within institutions that individuals routinely access, the system supports repeated observation without imposing additional survey burden.

Standardization and Interoperability

The proposed protocol emphasizes standardized anthropometric and behavioral indicators, uniform data formats, and interoperable digital dashboards across institutions. The data are envisioned to flow upward from classrooms and clinics to district and state repositories, creating a distributed yet integrated surveillance architecture. Periodic, nonstigmatizing feedback

XSL•FO
RenderX

reports at institutional and district levels are intended to transform surveillance from a descriptive exercise into an actionable public health function, supporting double-duty actions that address both undernutrition and overweight.

Target Population and Life Course Stratification

The surveillance protocol adopts a life course approach, structured around institutions that serve as fixed and recurring community anchors, including Anganwadi centers, schools, colleges, workplaces, and primary health care facilities (Table 1). These institutions function as sentinel surveillance units, periodically capturing nutrition-related and NCD-related risk indicators and generating data for both local action and higher-level aggregation. The system design aligns with World Health Organization recommendations for public health surveillance systems to be continuous, simple, flexible, acceptable, and action-oriented [16,22].

Table . Life course surveillance framework for overweight, obesity, and metabolic risk factors in India.

Life stage	Typical institution	Priority indicators	Surveillance purpose
Preconception	Community platforms, Eligible Couple Register, preconception counseling platforms	BMI, hemoglobin, dietary adequacy, physical activity, menstrual health	Early detection of undernutrition and overnutrition prior to conception
Pregnancy and lactation	PHC[a], antenatal clinics	Gestational weight gain, anemia, GDM[b], HDP[c], dietary intake	Early identification of maternal risk and counseling
First 1000 days (0 - 2 y)	PHC, Anganwadi centers	Breastfeeding, complementary feeding, infant growth	Strengthening IYCF[d] practices and growth monitoring
Under-five (0 - 5 y)	ICDS[e], preschools (kindergarten)	Weight-for-height, MUAC[f], anemia, dietary diversity	Monitoring wasting and stunting trajectories
School age (6 - 9 y)	Primary schools	BMI-for-age, diet, physical activity, screen time	Identification of emerging overweight patterns
Adolescents (10 - 19 y)	Schools, colleges	BMI-for-age, WHtR[g], diet, physical activity, body image, substance use	Informing life-skills education and prevention
Young adults (20 - 24 y)	Colleges, workplaces	BMI, waist circumference, blood pressure, stress, diet, sleep	Early detection of metabolic risk

[a]PHC: primary health center.

[b]GDM: gestational diabetes mellitus.

[c]HDP: hypertensive disorders of pregnancy.

[d]IYCF: infant and young child feeding.

[e]ICDS: Integrated Child Development Services.

[f]MUAC: mid-upper arm circumference.

[g]WHtR: waist-to-height ratio.

Surveillance windows are defined across the continuum from preconception to young adulthood, consistent with the first 1000 days framework and the adolescent-youth continuum [22,27]. Stratification by life stage enables the identification of critical periods for early risk detection, counseling, and preventive intervention.

This institution-anchored stratification supports the continuity of surveillance across life stages while leveraging existing human resources, such as teachers, auxiliary nurse midwives, and accredited social health activists, thereby minimizing additional operational burden.

Indicators and Measurement Domains

To ensure comparability across institutions and over time, the protocol proposes a standardized core indicator set, supplemented by context-specific expanded indicators where resources permit. Indicators are organized into four primary domains.

Anthropometric Indicators

Core anthropometric measures include height, weight, BMI (or BMI-for-age *z* scores), waist circumference, and waist-to-height ratio. Mid-upper arm circumference and triceps skinfold thickness may be included as supplemental measures in selected settings [28].

Dietary Indicators

Dietary assessment includes simplified 24-hour dietary recall or food-frequency modules, diet diversity scores, and frequency of consumption of ultra-processed foods. In school and household contexts, basic food environment mapping may be conducted to contextualize dietary patterns [29].

Biochemical Indicators

Biochemical assessment focuses on hemoglobin as a core indicator, with periodic or point-of-care measurement of fasting glucose, lipid profile, hemoglobin A_{1c}, ferritin, and other markers, such as urinary iodine or thyroid function tests, included where feasible and ethically appropriate [30].

XSL•FO
RenderX

Behavioral and Environmental Indicators

Behavioral indicators include physical activity (using adapted Global Physical Activity Questionnaire [GPAQ] or Youth Risk Behavior Surveillance System [YRBS] modules), sedentary behavior and screen time, sleep duration, tobacco and alcohol use, perceived stress, and body image perception [15,25].

Core indicators include height, weight, BMI (or BMI-for-age), waist circumference, hemoglobin, key dietary quality markers, and physical activity. Expanded indicators, such as biochemical markers, skinfold thickness, or detailed dietary modules, are included in selected institutions based on feasibility and ethical approval.

Together, these domains capture the balance between energy intake and expenditure, behavioral determinants, and biological manifestations that define malnutrition in all its forms.

Data Capture Tools and Technology Architecture

The surveillance system uses a hybrid technology model designed to accommodate India's heterogeneous digital infrastructure.

OMR-Based Paper Tools

In school and college settings, OMR-based paper tools enable standardized, low-cost data capture at scale, without reliance on continuous internet connectivity. Completed forms can be processed using locally available scanners, allowing rapid aggregation while preserving privacy, analogous to systems such as the YRBS [25].

Digital Platforms and Web Portals

In urban institutions and primary health care facilities, app-based or web-based platforms may be used to facilitate real-time data entry, automated dashboards, and the generation of individual or group-level feedback reports. Where feasible, interoperability with Ayushman Bharat Digital Health Mission identifiers may be explored to support longitudinal linkage.

IoT-Enabled Measurement Devices

Digital weighing scales, stadiometers, blood pressure monitors, glucometers, and selected point-of-care devices can be integrated to directly upload measurements to secure databases, reducing transcription errors and improving data quality [31].

Data Flow and Feedback Loops

The data collected at the institutional level are transmitted to district and state repositories through secure, encrypted app programming interfaces, with provisions for offline storage and periodic synchronization in low-connectivity settings. After validation and aggregation, automated, nonstigmatizing feedback reports are generated for institutions, PHC teams, and caregivers to enable local interpretation and timely action. These reports support targeted counseling, referral linkage, and planning of school-level and PHC-level health promotion activities, thereby operationalizing the surveillance-as-action paradigm within routine institutional workflows. Model one-page feedback report formats for educational institutions and for PHC-Anganwadi platforms are provided as Multimedia Appendices 1 and 2.

Ethical Considerations

Ethical safeguards are integral to the acceptability and sustainability of the surveillance system. Key provisions include informed consent and assent procedures, particularly for minors; confidentiality through anonymization and restricted access to identifiable data; and delivery of feedback using nonstigmatizing language to minimize the risk of labeling or harm [32]. Data governance will be overseen by institutional ethics committees and aligned with the Digital Personal Data Protection Act, 2023 [33]. An explicit equity lens will guide implementation to ensure the inclusion of low-resource institutions and marginalized populations.

Integration With Existing National Programs and Systems

Rationale for Program Integration

The proposed institute–based surveillance protocol is designed to complement and strengthen existing national nutrition and health programs rather than duplicate or replace them. By providing a standardized, institution-anchored surveillance layer, the protocol enables periodic data generated across programs to be consolidated into a coherent life course monitoring system. This integration is intended to reduce fragmentation, improve data use for prevention, and support double-duty actions addressing both undernutrition and overweight and obesity [34].

Rashtriya Bal Swasthya Karyakram

The RBSK provides a nationwide platform for screening children from birth to 18 years through mobile health teams and district early intervention centers. Within this protocol, RBSK platforms can serve as surveillance nodes for anthropometric and selected metabolic risk indicators among school-age children and adolescents. The standardization of nutrition-related measures, such as BMI-for-age, waist circumference, and selected dietary and physical activity indicators, would enable a shift from primarily curative screening toward preventive, trend-based surveillance while remaining aligned with the RBSK's existing operational workflows [23].

Rashtriya Kishor Swasthya Karyakram

The RKSK targets adolescents aged 10 to 19 years through school-based and community-based platforms, including adolescent-friendly health clinics and peer educator networks. The integration of the surveillance protocol within the RKSK would enable periodic, gender-disaggregated monitoring of anthropometry, anemia, dietary patterns, physical activity, and other lifestyle risk factors relevant to adolescent health [35]. This linkage supports continuity between school-based surveillance and health-facility–based counseling, addressing a critical life stage during which behavioral and metabolic risk trajectories are established.

POSHAN 2.0 and the ICDS Platform

Mission POSHAN 2.0 consolidates ICDS, Anganwadi services, and the Poshan Tracker platform for growth and nutrition monitoring during pregnancy, early childhood, and the first 1000 days. The institute-based surveillance protocol is designed to interface with POSHAN 2.0 data streams by aligning indicator definitions, periodicity, and reporting formats. This

XSL•FO
RenderX

interoperability enables under-five and maternal nutrition data, such as stunting, wasting, anemia, and dietary diversity, to be analytically linked with school-level and adolescent-level surveillance, facilitating life course continuity across the malnutrition spectrum [36].

PM POSHAN (School Meal Program)

The PM POSHAN Scheme provides daily meals to a large population of primary and upper-primary school children. Embedding surveillance within PM POSHAN–covered schools allows linkage between nutritional inputs (meals) and outcomes (anthropometry, diet quality, and physical activity indicators). Periodic feedback from school-based surveillance can inform menu planning, fortification strategies, and hygiene practices. The inclusion of private and unaided schools—currently outside PM POSHAN—within the surveillance protocol is proposed to enhance representativeness and enable comparative analysis across socioeconomic groups [36].

Ayushman Bharat and Digital Health Integration

Ayushman Bharat, through Health and Wellness Centres and the School Health Programme, provides a platform for preventive and promotive care across the life course. Within this protocol, Health and Wellness Centres function as data aggregation, validation, and counseling hubs, while school health teams support periodic surveillance rounds using standardized tools. Where feasible, linkage with Ayushman Bharat Digital Health Mission identifiers enables interoperability and longitudinal tracking, subject to ethical approvals and data governance safeguards [37].

Private-Sector and Institutional Partnerships

To improve inclusivity and sustainability, the protocol incorporates the voluntary participation of private schools, colleges, and workplaces, which collectively serve a substantial proportion of adolescents and young adults in India [38]. Partnerships with private hospitals, academic institutions, and laboratories are envisioned for capacity building, biomarker validation, and quality assurance. Such public-private collaboration strengthens surveillance coverage while maintaining standardized protocols and ethical oversight.

Enabling Double-Duty Actions and Policy Convergence

By integrating data across maternal, child, adolescent, and young adult platforms, the surveillance system enables double-duty actions that address multiple forms of malnutrition simultaneously. Consolidated dashboards support the identification of convergent risk patterns, such as early-life undernutrition followed by adolescent overweight, informing coordinated responses across food systems, physical activity promotion, and health services. This approach aligns with global and national commitments to double-duty actions and the Sustainable Development Goals (SDGs) related to nutrition and NCD prevention [34,39].

Summary

Integration with existing national programs positions the institute-based surveillance protocol as a unifying operational layer across health, nutrition, education, and welfare systems. By standardizing indicators, aligning data flows, and incorporating both public and private institutions, the protocol supports a scalable, inclusive, and life course–oriented surveillance architecture capable of generating actionable nutrition intelligence in real time.

Alignment With Primary Health Care Principles

Overview

The institute-based surveillance protocol is explicitly grounded in the principles of primary health care as articulated in the Alma-Ata Declaration and reaffirmed in the Astana Declaration [40,41]. These principles—equity, community participation, intersectoral coordination, and appropriate technology—are operationalized within the design, implementation, and governance of the surveillance system to ensure that data generation and use remain people-centered, ethically grounded, and action-oriented.

Equity

Equity is a foundational consideration in the surveillance design, recognizing the pronounced social gradients in malnutrition and metabolic risk across populations. The protocol embeds surveillance within both public and private institutions, including Anganwadi centers, government schools, private schools, colleges, and workplaces, to ensure representation across socioeconomic strata. By standardizing indicators and data formats across institutions, the system enables the identification of high-risk clusters while supporting equitable resource allocation and program targeting.

To minimize exclusion related to digital access, the protocol uses context-appropriate technology. OMR-based paper tools are used in resource-constrained settings, while digital platforms and connected devices are used where infrastructure permits. This tiered approach is intended to prevent the widening of the digital divide and to position surveillance data as a public good accessible for population-level planning and prevention [10,29].

Community Participation

Consistent with PHC principles, the protocol treats communities not merely as sources of data but as active participants in surveillance and response. Implementation involves the engagement of students, parents, teachers, accredited social health activists, and Anganwadi workers in data collection, interpretation, and feedback processes.

In educational settings, age-appropriate self-report modules adapted from established surveillance instruments are used to capture dietary and lifestyle behaviors [25]. Automated, nonstigmatizing summary reports are generated for parents, institutions, and local health teams to support counseling and informed decision-making. The periodic dissemination of aggregated findings through community platforms, such as nutrition campaigns and health and wellness activities, closes the feedback loop between surveillance and behavior change. This participatory design supports transparency, trust, and sustained compliance.

Intersectoral Coordination

Malnutrition and metabolic risk are shaped by determinants spanning health, education, nutrition, and broader social systems.

XSL•FO
RenderX

The surveillance protocol incorporates mechanisms for intersectoral coordination by aligning institutional roles across sectors. Primary health centers and Health and Wellness Centres support the validation of anthropometric and biochemical data, schools and colleges operationalize periodic data collection, nutrition programs provide contextual linkage to food and feeding environments, and digital health platforms facilitate secure data integration and reporting.

Such coordination enables surveillance findings to inform actions beyond the health sector, including school meal planning, physical activity promotion, and community-level nutrition interventions. By structuring data flows and responsibilities across sectors, the protocol operationalizes PHC's emphasis on collaborative action on the social determinants of health [41].

Appropriate Technology

The protocol applies the principle of appropriate technology by combining scientific validity with social acceptability and economic feasibility [42]. A hybrid technology model is adopted to suit diverse institutional contexts. OMR-based paper tools enable low-cost, scalable data capture in settings with limited digital infrastructure. IoT-enabled measurement devices, such as digital weighing scales, blood pressure monitors, glucometers, and portable hemoglobin meters, support accurate data capture in health care facilities. Digital dashboards and mobile apps facilitate real-time visualization and decision support for district and institutional stakeholders.

This flexible approach ensures that technology enhances, rather than replaces, human judgment and institutional capacity, allowing the surveillance system to scale while remaining responsive to local constraints [31].

Summary

By embedding equity, community participation, intersectoral coordination, and appropriate technology into its operational design, the institute-based surveillance protocol translates PHC principles into measurable practice. This alignment ensures that surveillance data are ethically generated, locally relevant, and actionable, supporting prevention-oriented responses to malnutrition and metabolic risk across the life course.

Operational Workflow for Institutional Surveillance

The surveillance workflow is designed as a station-based model aligned with the academic or service delivery schedule of the institution. On a typical school surveillance day, students are sequentially routed through 4 stations: registration and unique identifier assignment; anthropometric measurements; self-administered or interviewer-assisted behavioral and dietary assessment; and, where applicable, biochemical assessment.

Each station is managed by trained personnel, including teachers or institutional staff, project technical staff, and health functionaries. Anthropometric measurements are conducted using standardized protocols, and the average time per student is estimated at 5 to 7 minutes. A team of 4 to 6 trained personnel can complete measurements for approximately 120 to 150 students per day, depending on the institutional schedule and the availability of space.

Surveillance periodicity varies by life stage and institutional platform, with annual or biennial rounds in schools and colleges, and more frequent measurements integrated within routine PHC contacts for maternal and early childhood platforms.

Data Flow and Data Governance

Data captured through OMR-based tools or digital devices undergo field-level validation, including completeness checks and basic plausibility screening. Digitized data are transmitted through secure, encrypted channels to a central server. In settings with limited internet connectivity, offline data storage with periodic batch synchronization is used.

Each participant is assigned a unique surveillance identifier to enable longitudinal linkage while maintaining confidentiality. Role-based access controls define permissions for data entry, validation, analysis, and dashboard visualization.

Automated dashboards generate institution-level summaries, which are shared with authorized stakeholders. Individual-level reports are accessible only to designated health personnel for counseling and referral.

Data retention follows institutional and national data governance policies, with deidentified datasets used for analysis and reporting.

Quality Assurance and Standardization

All measurement devices are calibrated at predefined intervals. A minimum of 5% of anthropometric measurements are repeated for intraobserver and interobserver reliability assessment. Automated data checks flag biologically implausible values and missing fields at the point of digitization.

Quality indicators for each surveillance round include coverage, completeness of core variables, the proportion of implausible values, and turnaround time from data collection to feedback generation.

Phased Implementation and Pilot Evaluation

The surveillance system will be implemented in a phased manner, beginning with sentinel institutions in Hyderabad and the Medchal–Malkajgiri districts of Telangana. A pilot evaluation will assess feasibility and system performance using predefined indicators, including institutional coverage, proportion of complete records, turnaround time for feedback reports, stakeholder acceptability, and the proportion of identified high-risk individuals linked to counseling or referral services.

The findings from the pilot phase will be used to refine workflows, training modules, and technology architecture prior to scale-up.

Results

The institute-based surveillance protocol has transitioned from conceptual development to implementation through 2 intramural research platforms at the ICMR-National Institute of Nutrition.

The NULRISC project (sanctioned in July 2024) represents the school-based and college-based implementation arm of the surveillance framework. Formative phase activities, including

XSL•FO
RenderX

stakeholder engagement, tool refinement, workflow standardization, and feasibility assessments, are currently underway in approximately 15 high schools and colleges in Telangana.

The COPRIME project (sanctioned in May 2025) operationalizes the early life course and primary care components of the surveillance model. Formative implementation is ongoing in 3 primary health centers, 2 kindergartens, and 5 primary schools to standardize measurement protocols, data flow architecture, and referral linkages.

At the time of study revision, 5 schools and 5 colleges have been formally onboarded as sentinel surveillance institutions, with a target of 30 institutions in the first implementation phase across the Hyderabad and Medchal–Malkajgiri districts of Telangana.

The first periodic surveillance round is scheduled for June 2026. The initial phase will generate institution-level coverage estimates, data completeness indicators, turnaround time metrics, and feasibility and acceptability assessments. Analytical outputs from the first surveillance cycle are expected in 2027.

Discussion

Principal Findings

The increasing prevalence of overweight, obesity, and other forms of malnutrition among Indian children, adolescents, and young adults underscores the need for surveillance systems that are continuous, standardized, and embedded within routine institutional settings. While national surveys remain essential for population-level benchmarking, their episodic design limits their ability to support early detection, local decision-making, and preventive action. The institute-based surveillance protocol presented in this paper addresses this gap by proposing a life course–oriented, institution-anchored system that links data generation directly to action across education, health, and nutrition platforms.

Value of an Institute-Based Surveillance Model

Anchoring surveillance within institutions, such as Anganwadi centers, schools, colleges, primary health centers, and workplaces, offers several advantages over traditional survey-based approaches. First, repeated measurements aligned with institutional calendars enable timely monitoring of trends, allowing emerging risks to be identified before they consolidate into established disease patterns. Second, the protocol spans critical life stages, from preconception and the first 1000 days through adolescence and young adulthood, supporting a life course perspective that is essential for interrupting intergenerational cycles of malnutrition and metabolic risk [24,27].

Third, institution-level data enhance local relevance and usability, enabling context-specific responses such as school meal modifications, targeted nutrition counseling, or physical activity promotion. Fourth, integration across platforms, such as RBSK, RKSK, POSHAN 2.0, and PM POSHAN, enables double-duty actions that simultaneously address undernutrition and overweight, aligning with global recommendations for integrated malnutrition prevention [34]. Finally, the hybrid use of OMR-based paper tools and digital or IoT-enabled technologies supports cost-efficient scalability while remaining sensitive to infrastructural diversity and ethical considerations [31].

Positioning Within Global Surveillance Practice

International experience demonstrates the feasibility and utility of institution-based and school-linked surveillance systems for child and adolescent health. Systems, such as the Youth Risk Behavior Surveillance System in the United States and the WHO European Childhood Obesity Surveillance Initiative, illustrate how standardized, low-burden data collection in schools can generate actionable insights at scale [25,43]. The NHANES and the ECHO further demonstrate the value of integrating behavioral, anthropometric, and biological measures within continuous or longitudinal frameworks [13,14].

Compared with these systems, the proposed institute-based model is distinguished by its explicit life course coverage, its integration of both public and private institutions, and its alignment with existing national programs rather than stand-alone surveillance infrastructure (Table 1). By combining behavioral, anthropometric, and selected biochemical indicators within a unified framework, the protocol adapts global best practices to India's epidemiological and institutional context while retaining international comparability.

Anticipated Implementation Challenges

Despite its conceptual and operational strengths, the implementation of the proposed surveillance system may encounter several challenges. Capacity constraints related to the training and supervision of teachers, frontline health workers, and institutional staff may affect the consistency of data collection. Ensuring data quality and standardization across diverse settings will require robust calibration protocols, periodic validation, and automated data checks. Ethical concerns, including the protection of minors' data, informed consent, and the prevention of stigma, necessitate strong governance mechanisms and careful design of feedback processes [32].

Interoperability across existing program databases and the inclusion of private institutions will require sustained intersectoral coordination and alignment of data standards. Long-term sustainability will depend on policy commitment and financing mechanisms that transition surveillance from time-limited projects to routine institutional practice. These challenges are not unique to this protocol and can be mitigated through phased implementation, pilot testing within sentinel sites, and integration with digital public infrastructure and existing program workflows.

Implications for Policy and Practice

By shifting nutrition and metabolic risk monitoring from episodic surveys to continuous, institution-anchored surveillance, the proposed protocol offers a pathway to strengthen preventive public health practice in India. The system is designed to generate timely, disaggregated evidence that can inform program refinement, resource allocation, and local action while maintaining alignment with national and global nutrition

XSL•FO
RenderX

and NCD goals. If implemented at scale, this approach has the potential to transform how malnutrition in all its forms is monitored and addressed across the life course. The use of structured institutional feedback reports ensures that surveillance outputs are directly translated into decentralized preventive action within schools, Anganwadi centers, and PHC systems.

Conclusion

India is at a critical juncture in its nutrition transition, with undernutrition, overweight, obesity, and diet-related NCDs coexisting within the same populations and life stages. Although national surveys continue to provide essential population-level estimates, their episodic design limits their capacity to support timely, local, and life course–specific preventive action. The institute-based standardized surveillance protocol presented in this paper offers a structured approach to addressing this gap.

By embedding surveillance within routine institutional settings—Anganwadi centers, schools, colleges, primary health centers, and workplaces—and aligning it with existing national programs, the proposed system reorients nutrition and metabolic risk monitoring from episodic assessment to continuous prevention. The use of a hybrid technology model combining OMR-based paper tools, digital platforms, and IoT-enabled devices enables scalability while remaining sensitive to infrastructural diversity and ethical considerations. Importantly, the protocol operationalizes the core principles of PHC—equity, community participation, intersectoral coordination, and appropriate technology—ensuring that surveillance functions as a public good rather than a purely technical exercise.

International experience from systems, such as YRBS, NHANES, ECHO, Global School Health Policies and Practices Survey, and the WHO European Childhood Obesity Surveillance Initiative, demonstrates that institutionalized surveillance is both feasible and impactful when integrated into education and health systems. Adapted to India's context, the proposed life course surveillance model has the potential to generate timely, disaggregated evidence that informs double-duty actions, strengthens program convergence, and supports national commitments to the SDGs, particularly SDG 2 (Zero Hunger) and SDG 3 (Good Health and Well-being).

Reframing surveillance as a participatory, institute-anchored process transforms data collection into a continuous cycle of learning, accountability, and prevention. If implemented and sustained, this approach can contribute meaningfully to strengthening India's public health response to malnutrition in all its forms across the life course.

Acknowledgments

The authors gratefully acknowledge the guidance and institutional support of Dr Bharati Kulkarni, Director of the ICMR–National Institute of Nutrition (NIN), for enabling the development and implementation of platforms that informed this protocol. This work draws on operational insights generated through the intramural research initiatives NULRISC (Developing a Model Nutrition and Lifestyle Risk Surveillance System for Schools and Colleges in India; IIRP-2023-0000178) and COPRIME (Combating Childhood Obesity in Urban India: An Early Life-Course Approach through Primary Care; INTR-IG-2024-01-0017). The authors acknowledge the contributions of the project research teams, participating educational institutions, primary health care facilities, and field and technical staff involved in the formative phases of these initiatives. The authors also acknowledge the academic support provided by the Department of Humanities and Social Sciences, BITS Pilani, Hyderabad Campus, toward the doctoral research of the corresponding author (SRC). The authors acknowledge the broader public health, education, and primary health care stakeholders whose engagement is essential for the development of institute-based life course surveillance in India. The authors thank the JMIR Research Protocols editorial and peer-review team for their constructive feedback, which helped strengthen the conceptual clarity and reporting of this protocol.

The authors declare that they had insufficient funding to fully support the open access publication of this manuscript. JMIR Publications provided partial article processing fee (APF) support for this article.

Funding

The authors acknowledge APF support from the ICMR intramural research initiative NULRISC (Developing a Model Nutrition and Lifestyle Risk Surveillance System for Schools and Colleges in India; IIRP-2023-0000178).

Disclaimer

The authors used generative artificial intelligence tools for limited language editing and formatting support. All scientific content, study design, analysis, interpretation, and final manuscript preparation were performed and verified by the authors, who take full responsibility for the accuracy and integrity of the work.

Village names mentioned in the model institutional feedback reports are dummy villages and do not represent a real village.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Model institutional feedback report—institute-based nutrition and metabolic risk surveillance.
[PDF File, 883 KB - resprot_v15i1e86084_app1.pdf]

Multimedia Appendix 2
Model institutional feedback report—primary health center or Anganwadi nutrition and early life surveillance.
[PDF File, 212 KB - resprot_v15i1e86084_app2.pdf]

References

1. Obesity and overweight. World Health Organization. 2024. URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2026-03-23]
2. National family health survey (NFHS-5), 2019–21: India report. : International Institute for Population Sciences (IIPS), Ministry of Health and Family Welfare (India), ICF International; 2021 URL: https://dhsprogram.com/pubs/pdf/FR375/FR375.pdf [accessed 2026-03-23]
3. Global nutrition report 2022: stronger commitments for nutrition in a new era. : Development Initiatives Poverty Research Ltd; 2022 URL: https://media.globalnutritionreport.org/documents/2022_Global_Nutrition_Report_updated.pdf [accessed 2026-03-23]
4. Popkin BM. Nutritional patterns and transitions. Popul Dev Rev 1993 Mar;19(1):138. [doi: 10.2307/2938388]
5. Dietary guidelines for Indians: a manual. : National Institute of Nutrition (NIN), Indian Council of Medical Research (ICMR); 2011 URL: https://www.nin.res.in/downloads/DietaryGuidelinesforNINwebsite.pdf [accessed 2026-03-23]
6. National Non-Communicable Disease Monitoring Survey (NNMS) 2017–18. : Indian Council of Medical Research (ICMR), National Centre for Disease Informatics and Research (NCDIR); 2020 URL: https://www.ncdirindia.org/nnms/resources/NNMS%202017-18%20-%20Report.pdf [accessed 2026-03-23]
7. Victora CG, de Onis M, Hallal PC, Blössner M, Shrimpton R. Worldwide timing of growth faltering: revisiting implications for interventions. Pediatrics 2010 Mar;125(3):e473-e480. [doi: 10.1542/peds.2009-1519] [Medline: 20156903]
8. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence. Lancet Child Adolesc Health 2018 Mar;2(3):223-228. [doi: 10.1016/S2352-4642(18)30022-1] [Medline: 30169257]
9. SDG India Index 2023–24. : NITI Aayog; 2024 URL: https://www.niti.gov.in/sites/default/files/2024-07/SDG_India_Index_2023-24.pdf [accessed 2026-03-23]
10. Operational guidelines on school health programme under Ayushman Bharat. : NITI Aayog, Ministry of Health and Family Welfare; 2021 URL: https://nhm.gov.in/New_Updates_2018/NHM_Components/RMNCHA/AH/guidelines/Operational_guidelines_on_School_Health_Programme_under_Ayushman_Bharat.pdf [accessed 2026-03-23]
11. Thacker SB, Berkelman RL. Public health surveillance in the United States. Epidemiol Rev 1988;10:164-190. [doi: 10.1093/oxfordjournals.epirev.a036021] [Medline: 3066626]
12. Youth risk behavior survey data summary & trends report for dietary, physical activity, and sleep behaviors: 2013–2023. : Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services (HHS); 2025 URL: https://www.cdc.gov/yrbs/dstr/pdf/DSTR-Dietary-PhysicalActivity-SleepBehavior2013-2023-508.pdf [accessed 2025-10-12]
13. National health and nutrition examination survey (NHANES). National Center for Health Statistics (NCHS). 2023. URL: https://www.cdc.gov/nchs/nhanes [accessed 2026-03-23]
14. Gillman MW, Blaisdell CJ. Environmental influences on child health outcomes, a research program of the National Institutes of Health. Curr Opin Pediatr 2018 Apr;30(2):260-262. [doi: 10.1097/MOP.0000000000000600] [Medline: 29356702]
15. Breda J, McColl K, Buoncristiano M, et al. Methodology and implementation of the WHO European Childhood Obesity Surveillance Initiative (COSI). Obes Rev 2021 Nov;22 Suppl 6(Suppl 6):e13215. [doi: 10.1111/obr.13215] [Medline: 34738283]
16. Global school health policies and practices survey (GSHPPS). World Health Organization. 2021. URL: https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/global-school-health-policies-and-practices-survey [accessed 2026-03-23]
17. Popkin BM. Nutrition in transition: the changing global nutrition challenge. Asia Pac J Clin Nutr 2001;10 Suppl:S13-S18. [Medline: 11708577]
18. Bronfenbrenner U. The Ecology of Human Development: Experiments by Nature and Design: Harvard University Press; 1979. URL: https://www.jstor.org/stable/j.ctv26071r6 [accessed 2026-03-23]
19. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. Health Educ Q 1988;15(2):175-183. [doi: 10.1177/109019818801500203] [Medline: 3378902]
20. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory: Prentice-Hall; 1986. URL: https://books.google.co.in/books?id=HJhqAAAAMAAJ [accessed 2026-03-23]
21. Comprehensive National Nutrition Survey (CNNS) 2016–2018: National report. : UNICEF India and Ministry of Health and Family Welfare; 2019 URL: https://nhm.gov.in/WriteReadData/l892s/1405796031571201348.pdf [accessed 2026-03-23]
22. 2017-2018 STEPS fact sheet India. : World Health Organization; 2017 URL: https://www.who.int/publications/m/item/2017-2018-steps-fact-sheet-india [accessed 2026-03-23]

XSL•FO
RenderX

23. Operational guidelines for Rashtriya Bal Swasthya Karyakram (RBSK). : Ministry of Health and Family Welfare; 2013 URL: https://nhm.gov.in/images/pdf/programmes/RBSK/Operational_Guidelines/Operational%20Guidelines_RBSK.pdf [accessed 2026-03-23]
24. Patton GC, Sawyer SM, Santelli JS, et al. Our future: a Lancet commission on adolescent health and wellbeing. Lancet 2016 Jun 11;387(10036):2423-2478. [doi: 10.1016/S0140-6736(16)00579-1] [Medline: 27174304]
25. Centers for Disease Control and Prevention (CDC), Brener ND, Kann L, et al. Methodology of the Youth Risk Behavior Surveillance System—2013. MMWR Recomm Rep 2013 Mar 1;62(RR-1):1-20. [Medline: 23446553]
26. Guidelines of Pradhan Mantri POSHAN Shakti Nirman (PM POSHAN) Scheme. : Ministry of Education; 2024 URL: https://dsel.education.gov.in/sites/default/files/schemes_guidelines/Guidelines_PM%20POSHAN_SCHEME.pdf [accessed 2026-03-23]
27. Bhutta ZA, Das JK, Rizvi A, et al. Evidence-based interventions for improvement of maternal and child nutrition: what can be done and at what cost? Lancet 2013 Aug 3;382(9890):452-477. [doi: 10.1016/S0140-6736(13)60996-4] [Medline: 23746776]
28. WHO child growth standards: methods and development. : World Health Organization; 2007 URL: https://iris.who.int/server/api/core/bitstreams/0c57f82c-feec-4211-9cf5-7467b94a8e20/content [accessed 2026-03-23]
29. Popkin BM, Corvalan C, Grummer-Strawn LM. Dynamics of the double burden of malnutrition and the changing nutrition reality. Lancet 2020 Jan 4;395(10217):65-74. [doi: 10.1016/S0140-6736(19)32497-3] [Medline: 31852602]
30. Mukherjee A, Gupta R, Mukhopadhyay D, Mukherjee E. The growth of ultra-processed foods in India: an analysis of trends, issues and policy recommendations. 2023 URL: https://iris.who.int/bitstream/handle/10665/372414/9789290210672-eng.pdf [accessed 2026-03-23]
31. Boulos MNK, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014;5(3):229. [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]
32. Puhl RM, Latner JD. Stigma, obesity, and the health of the nation's children. Psychol Bull 2007 Jul;133(4):557-580. [doi: 10.1037/0033-2909.133.4.557] [Medline: 17592956]
33. The Digital Personal Data Protection Bill, 2023. PRS legislative research. URL: https://prsindia.org/billtrack/digital-personal-data-protection-bill-2023 [accessed 2025-10-18]
34. Hawkes C, Ruel MT, Salm L, Sinclair B, Branca F. Double-duty actions: seizing programme and policy opportunities to address malnutrition in all its forms. Lancet 2020 Jan 11;395(10218):142-155. [doi: 10.1016/S0140-6736(19)32506-1] [Medline: 31852603]
35. Rashtriya Kishor Swasthya Karyakram (RKSK). URL: https://rksk.in/ [accessed 2025-10-18]
36. Mission Saksham Anganwadi and POSHAN 2.0 Scheme guidelines. : Ministry of Women and Child Development; 2022 URL: https://cdnbbsr.s3waas.gov.in/s34c144c47ecba6f8318128703ca9e2601/uploads/2024/12/2024121116634619 18.pdf [accessed 2026-03-23]
37. Mishra US, Yadav S, Joe W. The Ayushman Bharat Digital Mission of India: an assessment. Health Syst Reform 2024 Dec 17;10(2):2392290. [doi: 10.1080/23288604.2024.2392290] [Medline: 39437234]
38. Unified District Information System for Education Plus (UDISE+) 2021–22. : Ministry of Education; 2022 URL: https://www.education.gov.in/sites/upload_files/mhrd/files/statistics-new/udise_21_22.pdf [accessed 2026-03-23]
39. The double burden of malnutrition: policy brief. : World Health Organization; 2021 URL: https://iris.who.int/server/api/core/bitstreams/03b3b543-2f0a-4978-af9b-3a559b235692/content [accessed 2026-03-23]
40. Rifkin SB. Lessons from community participation in health programmes: a review of the post Alma-Ata experience. Int Health 2009 Sep;1(1):31-36. [doi: 10.1016/j.inhe.2009.02.001] [Medline: 24036293]
41. Declaration of Astana. : World Health Organization (WHO); 2018 URL: https://iris.who.int/server/api/core/bitstreams/2211376b-d04d-4da8-bd99-8c5ffa2f426c/content [accessed 2026-03-23]
42. Bhattacharyya O, Khor S, McGahan A, Dunne D, Daar AS, Singer PA. Innovative health service delivery models in low and middle income countries - what can we learn from the private sector? Health Res Policy Syst 2010 Jul 15;8:24. [doi: 10.1186/1478-4505-8-24] [Medline: 20630108]
43. Wijnhoven TMA, van Raaij JMA, Spinelli A, et al. WHO European Childhood Obesity Surveillance Initiative 2008: weight, height and body mass index in 6-9-year-old children. Pediatr Obes 2013 Apr;8(2):79-97. [doi: 10.1111/j.2047-6310.2012.00090.x] [Medline: 23001989]

Abbreviations

CNNS: Comprehensive National Nutrition Survey
ECHO: Environmental Influences on Child Health Outcomes
GPAQ: Global Physical Activity Questionnaire
ICDS: Integrated Child Development Services
IoT: Internet of Things
NCD: noncommunicable disease

XSL•FO
RenderX

NFHS: National Family Health Survey
NHANES: National Health and Nutrition Examination Survey
OMR: optical mark recognition
PHC: primary health care
RBSK: Rashtriya Bal Swasthya Karyakram
RKSK: Rashtriya Kishor Swasthya Karyakram
SDG: Sustainable Development Goal
WHO: World Health Organization
yrbs: YRBS is appearing twice
YRBS: Youth Risk Behavior Surveillance System
YRBS: Youth Risk Behavior Surveillance System

Edited by J Sarvestan; submitted 18.Oct.2025; peer-reviewed by H Yusuf, R Ruparel; revised version received 27.Feb.2026; accepted 27.Feb.2026; published 13.Apr.2026.

Please cite as:
Challa SR, Dash B
Life-Course Framework for Institute-Based Surveillance of Overweight, Obesity, and Metabolic Risk Factors in India: Protocol for Surveillance System Development
JMIR Res Protoc 2026;15:e86084
URL: https://www.researchprotocols.org/2026/1/e86084
doi:10.2196/86084

© Sai Ram Challa, Biswanath Dash. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Standardized Methods for Evaluating Physical and Eating Behaviors: The WEALTH Cross-Sectional Study Protocol

Grainne Hayes[1], BSc, PhD; Christoph Buck[2], PhD; Greet Cardon[3], PhD; Richard Cimler[4], PhD; Steriani Elavsky[5], PhD; Leopold Fezeu K[6], PhD; Janas M Harrington[7], PhD; Jitka Kühnová[4], PhD; Jean-Michel Oppert[6,8], PhD; Luis Sigcha[1], PhD; Pepijn Van de Ven[9], PhD; Tomas Vetrovsky[4], MD, PhD; Catherine B Woods[1], PhD; Antje Hebestreit[2*], PhD; Alan E Donnelly[1*], PhD

[1]Department of Physical Education and Sport Sciences, Health Research Institute, Physical Activity for Health Research Centre, University of Limerick, Limerick, Ireland

[2]Leibniz Institute for Prevention Research and Epidemiology - BIPS, Bremen, Germany

[3]Department of Movement and Sports Sciences, Ghent University, Ghent, Belgium

[4]Faculty of Science, University of Hradec Králové, Hradec Králové, Czech Republic

[5]Department of Human Movement Studies, University of Ostrava, Ostrava, Czech Republic

[6]Université Sorbonne Paris Nord et Université Paris Cité, INSERM U1153, INRAE U1125, CNAM, Centre de Recherche en Epidémiologie et Statistiques (CRESS), Équipe de Recherche en Épidémiologie Nutritionnelle (EREN), Bobigny, France

[7]Centre for Health and Diet Research, School of Public Health, University College Cork, Cork, Ireland

[8]Department of Nutrition, Pitie-Salpetriere Hospital (AP-HP), Sorbonne University, Paris, France

[9]Health Research Institute, University of Limerick, Limerick, Ireland

[*]these authors contributed equally

Corresponding Author:
Grainne Hayes, BSc, PhD
Department of Physical Education and Sport Sciences, Health Research Institute, Physical Activity for Health Research Centre, University of Limerick, Limerick, Ireland

Abstract

Background: The accurate measurement of physical behaviors (PBs) and eating behaviors (EBs) is critical for designing, monitoring, and implementing public health guidelines and intervention strategies. The objective of the Wearable Sensor Assessment of Physical and Eating Behaviours (WEALTH) project was to develop standardized methods to identify daily PBs and EBs from wearable research- and consumer-grade sensors and evaluate the interaction and contexts of these behaviors.

Objective: The aim of this paper is to describe the study design and methods and report on the descriptive characteristics of the participants.

Methods: Within the framework of the WEALTH project, a cross-sectional study (spring 2023 to spring 2024) was completed in 5 European research centers in the Czech Republic, France, Germany, and Ireland. In each center, participants attended a research lab, completed an online questionnaire, and provided measures of anthropometry and handgrip strength. The participants were then fitted with 2 research-grade and 2 consumer-grade devices and participated in a standardized semistructured lab-based activity protocol. The latter was specifically designed to collect labeled data that simulated common PBs and EBs typical for a daily routine. Participants were then followed during a 9-day free-living data collection period, which combined the assessment of PB and EB via wearable devices and time-based, event-based, and self-initiated ecological momentary assessments (EMAs). The EMA surveys were complemented by three 24-hour dietary recalls, using validated web-based programs. Upon the completion of the survey protocol, participants completed a questionnaire that assessed the feasibility of the procedures.

Results: The final sample includes 627 participants, of whom 44% (n=275) were male. The mean age was 32.7 (SD 13.3) years, and the mean body mass index was 24.5 (SD 4.0) kg/m². The WEALTH study data will be used to develop machine learning (ML) models for classifying daily activities from wrist and hip-worn accelerometer data, evaluate EMA methods for studying interactions between PB and EB, and evaluate the feasibility and compliance of the methods. Data processing and ML model development are currently underway, with primary results expected to be published in 2026.

Conclusions: The output of the WEALTH project will be provided via a repository and a comprised toolbox of publicly available labeled data, ML models for behavior classification from accelerometer data, and a methodology to simultaneously capture EB and PB, thereby producing an integrated data collection system to support future research.

International Registered Report Identifier (IRRID): DERR1-10.2196/70186

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e70186) doi:10.2196/70186

KEYWORDS

physical behaviors; eating behaviors; ecological momentary assessment; machine learning; wearables

Introduction

Engaging in healthy physical behaviors (PBs), including regular physical activity (PA), low sedentary behavior, and optimal sleep routines, as well as in healthy eating behaviors (EBs), is essential for reducing the risk of noncommunicable diseases [1,2]. The accurate measurement of PBs is critical for designing, monitoring, and implementing public health guidelines and interventions. However, most existing assessment methodologies rely on self-report measures, which are susceptible to recall bias and social desirability bias and often demonstrate weak validity and reliability [3]. These limitations restrict the ability to examine how PBs interact with eating patterns and behaviors, collectively referred to here as EBs [4].

Device-based measures have improved the objectivity of PB assessment in epidemiological research [5]. Yet, comparisons across studies remain challenging due to differences in research-grade devices, nonstandardized intensity cut-points, and inconsistent validation analysis techniques [6]. Consumer-grade wearables are increasingly accessible, but their validity varies widely across device models and manufacturers [7-9], and many do not meet the standards required for academic research [9,10]. While conventional approaches to processing accelerometer data have relied on heuristic methodologies, machine learning (ML) techniques now enable more precise estimations [11,12]. These advancements allow for the improved classification of PB and energy expenditure from raw accelerometer data [13], compared with cut-point–based approaches [14]. However, standardized protocols for implementing ML in this context (PB) are still lacking.

Similar methodological inconsistencies affect EB measurement. Emerging digital and sensing technologies enable the objective and passive measurement of EBs; for example, wearable devices, such as smart watches, are equipped with accelerometers and gyroscopes that can track hand-to-mouth movements and hand gestures, providing insights into eating frequency, speed, and duration, as well as eating-related activities [15-17]. These technologies still encounter limitations when used in free-living environments [15,18]; for example, the diversity of available devices and sensors contributes to inconsistencies in data collection and analysis; thus, the absence of standardized measurement methodologies hinders cross-study comparisons and data interpretation. As a result, integrating PB and EB data in a meaningful way remains challenging.

Standardized measurement approaches (ie, standardized accelerometers, instruments, protocols, processing procedures) are therefore essential for producing comparable, high-quality PB and EB data and for enabling the examination of how these behaviors co-occur in daily life. Such standardization ensures that PB and EB data can be meaningfully integrated and compared. Despite the advantages of accelerometers for capturing movement-related aspects of PBs, these devices cannot provide information on the contextual factors that shape either EB or PB. Ecological momentary assessment (EMA) can address this limitation by capturing real-time self-reported data on behaviors and their surrounding contexts. By leveraging mobile technologies, EMA offers insight into the environmental, social, and psychological circumstances in which behaviors occur, thereby complementing accelerometry-based measures and strengthening multimethod approaches to PB and EB assessment [19]. At the same time, ML techniques have the potential to enable the use of sensor data from wrist-worn devices to detect and classify EB, such as meal onset or eating speed [18], but harmonized and validated models applicable across devices are still lacking.

Together, EMA and ML can complement conventional measures and contribute to more precise and integrated assessments of PBs and EBs. A web-based data processing framework incorporating ML models trained and validated for the processing and analysis of wearable device data that works seamlessly both for existing research-grade and commercial devices is feasible and warranted. To address the identified gaps, the aim of the "Wearable Sensor Assessment of Physical and Eating Behaviours" (WEALTH) project was to develop standardized methods for classifying daily PBs and EBs from accelerometry or sensor devices using EMA and ML methods and to investigate the interrelationship of both behaviors. To achieve this, the objectives of WEALTH were to (1) establish the feasibility of integrating accelerometer-derived PB data with EMA-derived EB data in free-living environments; (2) generate high-quality, standardized datasets using harmonized protocols; and (3) develop and validate ML models capable of classifying PBs. Secondary objectives included the creation of a web-based data processing infrastructure to support ML implementation, cross-device compatibility, and open-source data workflows and additional feasibility outcome measures. This paper describes the design and methods of the WEALTH project, providing insights into the available data and the quality measures to derive this unique dataset.

Methods

Study Design

Initiated in 2022, the WEALTH project was funded by the EU Joint Programming Initiative Healthy Diet Healthy Life [20] under the call on STAMIFY (standardised measurement, monitoring and/or biomarkers to study food intake, physical activity and health).

In the framework of the WEALTH project, a multicountry mixed methods cross-sectional survey was conducted, which focused on enrolling and monitoring PBs and EBs of adult participants from 5 research centers across Europe, in the Czech Republic, France, Germany, and Ireland. Participants were invited to participate between March 2023 and March 2024 via stratified and convenience sampling methods through letters, emails,

XSL•FO
RenderX

telephone calls, posters, flyers, and social media. France was the exception, recruiting their participants from the NutriNet-Santé study. With a target sample size of 600 participants (150 participants: 50% (n=75) female, aged 18 - 64 years per country), enrollment was distributed across the centers meeting the project's eligibility criteria (see below). A small cash reward incentive of €30-€40 (US $35-$47) per participant was given in the Czech Republic, Germany, and Ireland to increase response rates, participant compliance, and ensure the return of all devices. Participants who dropped out during the study or provided insufficient data were eliminated. Participants who dropped out before the center total was reached were replaced. A total of 627 adults completed the study (Table 1).

XSL•FO
RenderX

Table . Study population baseline characteristics.

	Total			Male			Female		
	Partici-pants, n	Mean (SD)	IQR	Partici-pants, n (%)	Mean (SD)	IQR	Partici-pants, n (%)	Mean (SD)	IQR
Total									
Height (m)	627	1.7 (0.1)	1.6-1.8	275 (43.9)	1.8 (0.1)	1.7-1.8	352 (56.1)	1.7 (0.1)	1.6-1.7
Weight (kg)	627	72.1 (14.3)	61.2-80.8	275 (43.9)	80.8 (12.7)	72.9-88.0	352 (56.1)	65.3 (11.6)	56.9-70.8
BMI (kg/m^2)	627	24.5 (4.0)	21.6-26.6	275 (43.9)	25.1 (3.7)	22.8-26.9	352 (56.1)	24.0 (4.1)	21.1-26.1
Age (y)	627	32.7 (13.3)	24.0-40.0	275 (43.9)	33.2 (13.0)	24.0-41.0	352 (56.1)	32.3 (13.5)	24.0-40.0
Waist circumference (cm)	626	83.7 (11.9)	74.2-91.0	274 (43.8)	89.5 (10.9)	81.3-95.0	352 (56.1)	79.1 (10.6)	71.2-86.0
Max handgrip strength (kg)	627	35.1 (10.7)	26.9-43.2	275 (43.9)	44.5 (7.9)	39.1-50.0	352 (56.1)	27.7 (5.6)	23.9-31.5
Czech Republic									
Height (m)	162	1.7 (0.1)	1.6-1.8	60 (37.0)	1.8 (0.1)	1.8-1.9	102 (63.0)	1.7 (0.1)	1.6-1.7
Weight (kg)	162	74.1 (15.6)	61.2-83.9	60 (37.0)	83.4 (14.3)	74.6-88.9	102 (63.0)	68.5 (13.7)	58.3-76.4
BMI (kg/m^2)	162	24.9 (4.4)	21.7-26.9	60 (37.0)	25.3 (3.7)	23.3-26.6	102 (63.0)	24.7 (4.7)	21.3-27.4
Age (y)	162	37.4 (11.4)	29.0-46.0	60 (37.0)	36.7 (9.9)	29.0-44.0	102 (63.0)	37.8 (12.2)	29.0-47.0
Waist circumference (cm)	161	84.7 (11.3)	74.5-92.0	59 (36.6)	89.6 (10.4)	81.0-95.5	102 (63.0)	81.9 (10.9)	72.5-89.5
Max handgrip strength (kg)	162	34.7 (10.9)	27.7-42.0	60 (37.0)	45.2 (8.8)	38.6-50.5	102 (63.0)	28.5 (6.4)	25.2-32.3
France									
Height (m)	159	1.7 (0.1)	1.6-1.8	68 (42.8)	1.8 (0.1)	1.7-1.8	91 (57.2)	1.6 (0.1)	1.6-1.7
Weight (kg)	159	72.5 (14.6)	61.3-82.4	68 (42.8)	81.5 (13.1)	73.8-90.4	91 (57.2)	65.8 (11.8)	56.6-71.5
BMI (kg/m^2)	159	25.2 (4.4)	22.0-27.5	68 (42.8)	25.9 (4.3)	22.4-28.5	91 (57.2)	24.6 (4.4)	21.2-26.8
Age (y)	159	36.7 (17.6)	20.0-57.0	68 (42.8)	39.3 (17.7)	21.0-57.5	91 (57.2)	34.7 (17.3)	20.0-52.0
Waist circumference (cm)	159	85.7 (14.2)	74.5-94.0	68 (42.8)	93.4 (13.0)	84.5-104.4	91 (57.2)	80.0 (12.2)	71.4-86.0
Max handgrip strength (kg)	159	33.8 (9.9)	25.4-42.2	68 (42.8)	43.0 (7.5)	39.5-47.5	91 (57.2)	27.0 (4.6)	23.5-30.8
Germany									
Height (m)	156	1.7 (0.1)	1.6-1.8	72 (46.2)	1.8 (0.1)	1.7-1.8	84 (53.8)	1.6 (0.1)	1.6-1.7
Weight (kg)	156	68.7 (12.5)	60.0-76.1	72 (46.2)	76.1 (11.9)	68.5-84.0	84 (53.8)	62.3 (9.1)	55.3-67.3
BMI (kg/m^2)	156	23.7 (3.3)	21.3-25.6	72 (46.2)	24.1 (3.3)	21.6-26.0	84 (53.8)	23.4 (3.2)	21.2-25.4
Age (y)	156	24.3 (2.0)	24.0-25.0	72 (46.2)	24.5 (0.5)	24.0-25.0	84 (53.8)	24.1 (2.7)	24.0-25.0
Waist circumference (cm)	156	82.6 (10.6)	74.4-89.2	72 (46.2)	87.9 (10.4)	80.7-93.5	84 (53.8)	78.1 (8.4)	71.2-85.2
Max handgrip strength (kg)	156	32.6 (9.9)	24.5-39.7	72 (46.2)	41.4 (7.0)	37.4-46.5	84 (53.8)	25.1 (4.1)	22.1-27.8
Ireland									
Height (m)	150	1.7 (0.1)	1.7-1.8	75 (50)	1.8 (0.1)	1.8-1.9	75 (50)	1.7 (0.1)	1.6-1.7
Weight (kg)	150	73.1 (13.9)	62.0-81.5	75 (50)	82.6 (10.4)	76.5-88.1	75 (50)	63.7 (10.0)	57.0-67.7
Body Mass Index (kg/m^2)	150	24.2 (3.5)	21.5-26.0	75 (50)	25.4 (3.2)	23.2-27.1	75 (50)	23.0 (3.5)	20.3-25.0
Age (y)	150	32.2 (12.3)	22.0-40.0	75 (50)	33.1 (12.1)	23.0-41.0	75 (50)	31.3 (12.6)	22.0-39.0

XSL•FO
RenderX

	Total			Male			Female		
	Partici-pants, n	Mean (SD)	IQR	Partici-pants, n (%)	Mean (SD)	IQR	Partici-pants, n (%)	Mean (SD)	IQR
Waist circumference (cm)	150	81.4 (10.8)	72.5-88.0	75 (50)	87.5 (8.7)	81.0-92.0	75 (50)	75.4 (9.3)	67.5-82.0
Max handgrip strength (kg)	150	39.5 (10.9)	31.2-48.7	75 (50)	48.4 (6.7)	44.0-52.5	75 (50)	30.5 (5.6)	27.5-33.5

Eligibility Criteria

To be deemed eligible for inclusion, participants were required to be aged between 18 and 64 years. To maintain protocol homogeneity and ensure the relevance of the data processing methods applied, individuals with overt chronic diseases, physical impairments, and those engaged in shift work were excluded from the study. These exclusion criteria were implemented to uphold the integrity of the prescribed semistructured activity protocol and to prevent potential disturbances in the free-living period arising from participants' unique circumstances and, consequently, enhance the fidelity of the ensuing analytical outcomes.

Study Procedures and Data Collection: Overview of Study Protocol

The WEALTH data collection encompassed 2 integral components (Figure 1). The first component involved a standardized semistructured lab–based activity protocol, specifically designed to collect labeled data that simulated common PBs and EBs. The second component extended the data collection beyond the laboratory setting. It incorporated a 9-day EMA data collection combined with worn accelerometer and wearable devices, thereby enabling the continuous data collection of individuals' behaviors during free-living. To record the participants' experience of the study, an end-of-study feasibility instrument was completed when the participants returned their devices. To uphold methodological rigor and ensure conformity across all stages of the study, a WEALTH standard operating procedures manual was developed to define the WEALTH research group's study protocol consisting of different modules (wearables, anthropometric measurements, dietary assessment, and general questionnaire), field access with interview guidelines, order of measurements, measurement procedures, equipment requirements, prerequisites, and data entry. The manual served as a comprehensive reference tool, offering researchers clear guidelines to enhance the precision and effectiveness of the data collection processes across the various study locations.

XSL•FO
RenderX

Figure 1. WEALTH (Wearable Sensor Assessment of Physical and Eating Behaviours) study protocol. EMA: ecological momentary assessment.

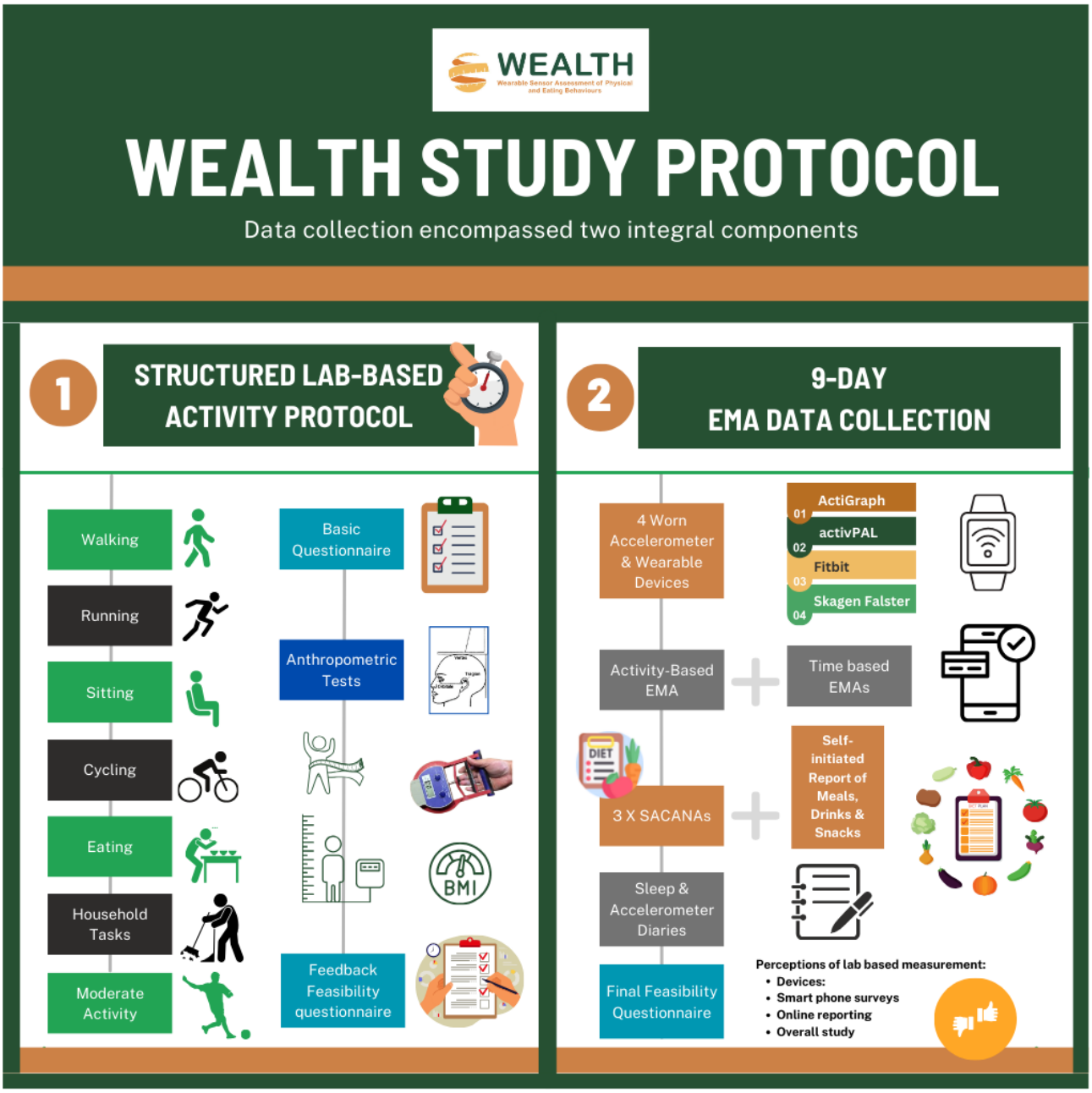

Measurement of Baseline Information and Anthropometric Characteristics

Measurement of Baseline Information

Upon participant recruitment and prior to participating in the laboratory testing session, participants were required to provide online written consent, download the Fitbit app and a custom-built mobile app HealthReact for delivering the EMA surveys, and complete a Physical Activity Readiness Questionnaire [21].

Participants attended the center wearing light gym clothing and running shoes. Upon arrival, they were provided with a verbal description of the semistructured laboratory and free-living components of the study. In Ireland, participants then completed an online baseline WEALTH questionnaire, while researchers verified the installation of the HealthReact and Fitbit apps. In the Czech Republic and France, participants filled in the online baseline WEALTH questionnaire through a secured link prior to participating in the laboratory session. In Germany, participants were provided with a link to the online baseline WEALTH questionnaire and were asked to fill it out in the upcoming week. The general WEALTH questionnaire comprised several instruments assessing important co-variates, such as participant sociodemographics and lifestyle behaviors, such as smoking, alcohol drinking, sleep duration, PA, and food consumption. It also assessed the general health status, including the Short Form 36 Health Survey Questionnaire instrument [22], to assess health-related quality of life. All questionnaires were provided in the language of the respective country (translated and back translated to control for translation errors or using established translated versions of the tools).

XSL•FO
RenderX

Measurement of Anthropometric Characteristics

Anthropometric measurements of the participants were measured without shoes and in light clothing. Height was measured to the nearest 0.25 cm, using a portable stadiometer (Seca Ltd), and body mass was measured to the nearest 0.1 kg using a portable electronic scale (Seca Ltd). BMI was recorded using the standard formula (kg/m^2). Waist circumference was recorded to the nearest 0.1 cm with an adjustable anthropometric inelastic tape (Seca Ltd). The measurement was taken at the level of the narrowest point of the abdomen at the end of normal exhalation as observed from the anterior aspect. Alternatively, if no one point was evident, waist circumference was assessed at the midpoint of the lowest rib and iliac crest at the end of normal exhalation. The measurement of the waist circumference was completed twice.

Handgrip strength was measured with the dominant and nondominant hand using a digital hand dynamometer with adjustable grip (Takei Scientific Instruments Co). Participants were instructed to self-adjust the dynamometer so that it fitted comfortably to their hand size to obtain their best performance [23]. Prior to the measurement, participants were taken through a familiarization trial for each hand so that they could get acquainted with the instrument and procedures and choose the best adjustment [24]. Participants were then instructed to grip the dynamometer with maximum strength in response to a voice command [25]. For the measurements, participants stood holding the dynamometer by their side with both arms fully extended, avoiding contact with any other body part [26]. Two trials were performed on each side, alternately, with a rest period of at least 1 minute between trials of the same hand [27].

Measurement of PBs and EBs

Participants were fitted with 2 research-grade (ActiGraph wGT3X-BT, activPAL 3 micro) and 2 consumer-grade (Fitbit Charge 5, LifeQ enabled smartwatch) devices [28]. The Fitbit (Fitbit) and a LifeQ-enabled smartwatch (Skagen Designs Generation 6) were worn on the participants' nondominant and dominant wrist, respectively. The smartwatch was supported by LifeQ to enable the storage and download of tri-axial accelerometer data and heart rate from green photoplethysmography. The Fitbit device was synchronized with the Health React app on either the participant's or the study smartphone. The number of steps and heart rate data from Fitbit were collected on the Health React server. The ActiGraph wGT3X-BT was attached to the right iliac crest using an elastic band, and the activPAL was attached to their right thigh as per manufacturer guidelines [28]. The characteristics and specifications of each activity monitor are described in Table 2. The activPAL device has demonstrated excellent reliability and validity in measuring sedentary behavior and PA patterns in free-living conditions [28]. Similarly, the ActiGraph device has shown robust reliability and validity in assessing PA levels across various populations and settings [29,30]. Favorable reliability and validity metrics have been reported for the Fitbit Charge 5, particularly in tracking steps, heart rate, and sleep patterns [31]. While the LifeQ-enabled smartwatch shows promising reliability and validity in monitoring PA and providing personalized feedback, further research is needed to establish its performance fully [32]. To address the potential risk of timestamp misalignment across the 4 wearable devices, all devices were synchronized to a single study laptop or to a server time prior to deployment, ensuring consistent timing across devices. Continuous data collection protocols will incorporate periodic checks and automatic synchronization routines (where supported by the device manufacturer) to maintain temporal alignment throughout the monitoring period. Any residual discrepancies in timestamps will be handled during the data curation process prior to the use in ML pipelines.

Table . Device specifications.

Specification	activPAL 3 micro	ActiGraph wGT3X-BT	LifeQ-enabled smartwatch	Fitbit Charge 5
Size (mm)	23.5×43×5	33×46×15	32.5 active area diameter	36.8×23×11
Mass (g)	10	19	100	29
Placement	Midpoint of anterior right thigh	Right iliac crest	Dominant wrist	Nondominant wrist
Application	Waterproofed with nitrile sleeve and tegaderm	Elastic Belt	Wrist strap	Wrist strap
Internal sensors	Tri-axial accelerometer	Tri-axial Accelerometer	Uni-axial accelerometer, heart rate monitor	Tri-axial accelerometer, heart rate monitor
Sample rate (Hz)	20 - 80	30 - 100	25	100
Accelerometer amplitude range (±g)	6	2 - 8	Not reported	Not reported
Connectivity (Wifi or Bluetooth or USB)	USB	USB and Bluetooth	Wifi, USB, and Bluetooth	USB and Bluetooth
Battery life	7 - 10 days	7 - 14 days	24 - 36 hours	7 days

In the German study center, a subsample of participants was fitted with a portable metabolic unit (MetaMax 3b, CORTEX Biophysik, Leipzig) for the duration of the standardized activity protocol. Oxygen uptake (VO_2 mL/kg/min), carbon dioxide production (VCO_2, mL/kg/min), and heart rate were measured

XSL•FO
RenderX

breath-by-breath. The data were internally stored during the standardized activity protocol and uploaded to the computer running the dedicated software (MetaSoft, Cortex Biophysik, Leipzig) afterward.

For collecting labeled data, the participants then followed a 2-hour standardized semistructured activity protocol to record PBs and EBs, designed to emulate activities of daily living. The participants performed a maximum of 11 common activities (Table 3) across 4 intensity categories (sedentary, light, moderate, and vigorous). Sedentary and light-intensity activities were performed for a duration of 3 - 5 minutes, while moderate- and vigorous-intensity activities were performed for a duration of 4 - 10 minutes. The eating a meal task was performed over a duration of 10 minutes.

XSL•FO
RenderX

Table . Standardized semistructured activity protocol.

Time (min)	Activity intensity	Description
10	Sedentary	Participants sit and fill in cross-word puzzles (5 min) and watch the news (5 min).
6	Standing and snacking	Passive standing: relaxed (3 min). Stand and snack while watching the TV[a] (3 min).
5	Light activity	Slow walking: individual speed, not controlled. Include 2 random stops for about 20 seconds (5 min).
5	Moderate activity	Brisk walking: individual speed, not controlled (5 min).
4	Vigorous activity	Jogging or running: individual speed, not controlled (4 min).
3	Transition	Short rest: provide water especially to participants with spirometry masks. Allow participants to sit, check phones, and so forth. Explain games and the next task.
8	Moderate activity	Let participants choose active games and switch in between (eg, 1 vs 1 for each game). Participants do not need to attend all possible games but rather games that they are able or experienced to play (8 min). • Football playing soccer back and forth • Badminton • Table tennis • Basketball (no competition rather playing to the basket)
4	Transition	Get the bikes and give directions (4 min)
5	Moderate activity	Cycling: cycle round outside of the building on a real bike or stationary bike. Ensure that participants have previously cycled and are comfortable in doing so. All participants are required to wear a helmet throughout this task. Include at least 2 stops for about 20 seconds (5 min).
5	Rest	Short-seated: rest particularly for the participant wearing spirometry—ask participants if they want to keep on the masks—provide water. Explain household tasks for next tasks (5 min).
10	Light activity	Household tasks: let participants switch between tasks, for example, participant A starts hoovering, participant B is filling the dishwasher, then participant A is emptying the dishwasher and B is hoovering. Participants do not need to execute all provided tasks (10 min). • Hoovering • Sweeping or mopping (without water) • Emptying or filling the dishwasher • Chopping vegetables • Spread butter on toast, prepare sandwiches for task 9 (could be seated at the table).

XSL•FO
RenderX

Time (min)	Activity intensity	Description
10	Eating a meal (with drinks)	Eat and drink while sitting provided food and drinks. Sandwiches, vegetables or fruits, yogurt (provide different options: vegan, with milk, plain, or with flavor). Proposed sequence: • Minute 1: drink out of a cup (cold beverage) • Minute 2 - 3: Use knife or fork • Minute 4 - 5: Use spoon (cereal or yogurt) • Minute 6 - 7: Eat with hand From minute 7: let participants eat as preferred Activity session ends after eating a meal task.

[a]TV: television.

9-Day Free-Living Data Collection

Time-Based, Event-Based, and Self-Initiated Ecological Momentary Assessments

Following the standardized activity protocol, participants were required to continue to wear the 4 devices during a 9-day free-living data collection. During this time, participants received time-based and event-based EMA prompts to answer short questionnaire surveys that appeared on their smartphones. In addition, they were instructed to self-initiate surveys on specific occasions when they were eating, snacking, or drinking (except water).

The EMA data were collected using the HealthReact platform developed at the University of Hradec Kralove, Czechia (a license to use HealthReact is held by a company owned by Richard Cimler, a member of the WEALTH consortium). HealthReact is a software suite comprising a server-side component and a mobile app for iOS and Android smartphones [33]. The HealthReact server can collect data from a diverse range of sensors, including wearable devices, such as Fitbit, and it can evaluate these data in real time to automatically trigger event-based EMA surveys based on predefined rules. These rules can be customized by researchers through a user-friendly web interface. The EMA surveys are then pushed to and displayed on the HealthReact app on participants' smartphones. Additionally, HealthReact supports traditional time–based surveys, which are triggered at random times within a specified time frame. Finally, HealthReact also enables the possibility of self-initiated surveys. For the WEALTH study, the Fitbit Charge 5 was chosen as a suitable wearable sensor for triggering the event-based surveys within HealthReact. While HealthReact is not the only platform that enables accelerometry-based triggers, most commonly used EMA platforms do not support this feature, which may limit comparability with other studies.

For the 9-day EMA protocol, the data collection was standardized, with the day when participants attended the lab (day -1) and the first full day (day 0) considered as "lead-in" days, followed by 7 full days of the monitoring period (ie, day 1 -to day 7). Participants without a compatible smartphone or without a mobile data plan were provided with a project smartphone, which included a continuous mobile internet connection. The EMA protocol consisted of a combination of time-based, event-based, and self-initiated surveys. There were 7 time-based surveys (08:00-09:45 morning survey, 10:15-11:45, 12:15-13:45, 14:15-15:45, 16:15-17:45, 18:15-19:45 daily surveys, 20:15-21:45 evening survey) triggered randomly in the predefined time windows. These windows were designed with a time gap to ensure the surveys were 30 minutes apart. In addition to time-based surveys, the event-based surveys were triggered by data collected by the Fitbit Charge 5 that was worn on the participants' wrists. The Fitbit device provided accurate minute-level number of steps [34] and several heart rate recordings per minute. These data were collected continuously and required continuous internet connection for regular synchronization with the Fitbit server, which occurs automatically approximately every 15 minutes or when the participant manually syncs the app. A sedentary survey was triggered after 20 consecutive minutes of 0 steps with concurrent detection of heart rate to distinguish between sedentary behavior and nonwear. A maximum of 2 sedentary surveys were administered within the 8:00-14:00 time window and the 14:00-20:00 time window (ie, a maximum of 4 surveys per day), and there was a minimum gap of 90 minutes between 2 sedentary surveys. A walking survey was triggered after 5 consecutive minutes with ≥60<140 steps/minute, with a maximum of 3 walking surveys per day. A running survey was triggered after 5 consecutive minutes with ≥140 steps/minute, with a maximum of 3 running surveys per day. It should be noted that because the triggers were step-based, the event-based EMA surveys did not capture nonambulatory forms of PA (eg, cycling, resistance training, yoga). These activities could still be reported as "exercise" via time-based or self-initiated surveys but were not part of the automated event-based triggering. For both time-based and event-based surveys, participants were instructed to complete the survey within 15 minutes of the prompt. They were reminded 3 times during the 15-minute validity period of the questionnaire: after 5 minutes, after 10 minutes, and 3 minutes before the questionnaire expired. The survey was sent to the server when it was completed or after the time for completion had expired (even when not completed).

Consumed meals, snacks, and drinks (except for water) were monitored through self-initiated surveys completed at the time of consumption or shortly thereafter. Reminders for the completion of the self-initiated surveys were incorporated within the time-based surveys. Depending on the number of triggered event-based surveys, the maximum number of surveys a participant received per day was 17.

XSL•FO
RenderX

24-Hour Dietary Recall and Food Frequency Questionnaire

The self-initiated meal, snack, and drink surveys were complemented by 3 nonconsecutive 24-hour dietary recalls (24HDR) completed either on the SACANA (Self-Administered Children, Adolescent and Adults Nutrition Assessment 24-hour Dietary Recall in Czech Republic, Germany and Ireland) or the NutriNet-Santé [35] online platforms. Both instruments are validated web-based programs for the assessment of 24HDR [36-38]. Participants were asked to complete at least three 24HDR on 2 nonconsecutive weekdays and 1 weekend day during the 9 days of data collection. They were encouraged to complete the first 24HDR after the devices had been worn for at least 2 days. The online platforms provided a timestamp for each eating, snacking, or drinking occasion, place of consumption, company, and simultaneous activities. Thus, the 24HDR tools also provided data for cross-linking the context of snacking or eating behavior with the EMA data. Participants were further asked to report the frequency of intake of food items during the previous 4 weeks using a food frequency questionnaire. The food frequency questions were grouped into 15 food groups, namely vegetables, fresh fruits, drinks, breakfast cereals, milk, yogurt, fish, meat and meat products, eggs and mayonnaise, meat replacement products and plant-based products, cheese, spreadable products, olive oil, cereal products, and snacks [39].

Since measurement error is inevitable, correction methods are strongly advised to reduce bias in effect estimates when analyzing 24HDR data [40]. Therefore, usual dietary intake will be calculated according to the National Cancer Institute method [41,42] using the 24HDR and food frequency questionnaire data from the general questionnaire. Upon the completion of the study, participants completed a feasibility and acceptability questionnaire designed to assess study procedures used within the laboratory and during the 9-day protocol period.

Data Storage and Management

All data comprising questionnaires and device data were compiled, processed, and stored in a quality-controlled common database in accordance with European Union data protection regulations. The accuracy of the EMA reports was evaluated using the activPAL's CREA classification, which served as a gold standard for scarcely labeled free-living data [43]. The activPAL device was selected due to its unobtrusive and strategic placement on the body, enabling reliable continuous monitoring over 24-hour periods [44]. Data completeness was verified on a daily basis for each sensor to ensure integrity, and potential nonwear time was identified through 2 complementary approaches: participant self-reports collected via questionnaires and automated predictions derived from either the activPAL's CREA algorithm or the GGIR package. This approach provided a systematic framework to assess data loss, ensure data quality, and mitigate limitations associated with the accuracy of self-reported labels. Fitbit device data and EMA prompted via the HealthReact software are stored on a European Union server with project-specific access. Fitbit data were initially transmitted and stored on a US server before retrieval by the HealthReact app. In accordance with the WEALTH standard operating procedures specifications and to ensure data quality management measures were upheld, daily monitoring processes were implemented to evaluate participant adherence or nonadherence across all centers. All study sites received email notifications regarding data inconsistencies, and the subsequent resolution of queries was conducted to increase data completeness and data quality. Identified discrepancies included inconsistent data, absent data (relevant EMAs, context-based questionnaire), range validations, and deviations from the protocol. In addition, weekly rollouts for monitoring the compliance with 24HDR entries and the general questionnaire were provided to all centers.

Data Analysis Plan

The data collected in the WEALTH project will be used initially for 2 purposes. For hip- and wrist-worn devices, ML models will be developed to classify PB and EB from raw tri-axial accelerometer data. Model development will be based on strict labeled data from the semistructured protocol as well as labeled data from EMA time stamps with valid behavior reports of participants during the free-living survey. Classifiers for PBs and EB will be modeled using multiple ML methods ranging from random forests to convolutional neural networks or more complex deep learning applications. In the German subsample, the data on energy expenditure from the semistructured protocol will be used to develop ML models estimating metabolic equivalent of task units to each behavior.

Predictive models will be developed using the lab-based protocol and subsequently externally validated and calibrated using free-living sensor data labeled via EMA responses. To ensure robust evaluation, a unified validation strategy will be applied, combining subject-independent holdout validation, k-fold cross-validation, and performance testing on EMA-labeled free-living data. Model performance will be assessed using established metrics, including balanced accuracy, precision, recall, and *F*-score, computed only on data not used for training.

To ensure comparability across devices, accelerometer data will be harmonized by resampling to 20 Hz, converting to gravitational units (g), and synchronizing all sensors using the start time recorded during the laboratory session. EMA, accelerometer, and dietary recall data will be aligned using time stamps, enabling consistent temporal integration of self-reported and sensor-derived behaviors.

EMA-labeled data used to train ML models will be further validated by comparing them with external references, including the activPAL CREA algorithm and established ActiGraph classification methods. Low-agreement segments will be excluded using threshold-based criteria, ensuring that only high-confidence labels contribute to model development as done in Sigcha et al [43]. Model success will also be assessed through consistency across devices and quantitative agreement between EMA reports, accelerometer-derived predictions, and dietary-recall events. Agreement will be evaluated using temporal overlap metrics, Bland-Altman comparisons, intraclass correlation coefficients, and concordance indices, such as Cohen's or weighted kappa for categorical outcomes.

XSL•FO
RenderX

The second purpose of the WEALTH study is to evaluate the potential of the use of EMA methods to study the interrelation between PB and EB. Interrelationships will be examined between PB and EB as predicted from ML models and data on context as assessed by EMA time- and context-sensitive results on incident behaviors. Multibehavioral patterns on the data from these different sources will be identified to investigate the association on body composition, well-being, and health status (self-rated health and health-related quality of life). Based on the feedback questionnaires, feasibility will be evaluated for the included measurement methods, and factors (eg, country, age, gender, activity level) associated with the degree of acceptance will be defined.

Power Analysis and Sample Size

Sample size in each test center was guided by an estimate for a sample needed to develop the ML models. The sample size required to train a ML model successfully depends on the complexity of the model, the use of prior knowledge within the model, and the acceptable difference between performance observed during training and performance that can be expected on yet unseen data. It was estimated that a sample of at least 100 participants would allow for a substantial range of ML models to be trained successfully. This estimate is further supported by evidence from large-scale free-living accelerometer datasets, where empirical learning-curve analyses demonstrate that model performance improves markedly with increasing sample size up to approximately 60 - 100 participants, after which performance gains plateau, with only marginal improvements (<1% - 2% in macro-F1) for both classical ML algorithms (eg, random forest, gradient boost), and deep learning algorithms (eg, convolutional or recurrent neural networks) [45]. These findings suggest that a cohort of 100 participants provides sufficient data to train robust, population-level models while maintaining generalizability across diverse activity contexts. To ensure that the amount of available data would not be a limiting factor in the choice of ML models, and to safeguard against data loss or participant attrition, 150 participants were recruited in each center.

Ethical Considerations

This multicenter study is being conducted in accordance with globally accepted standards of good research practice in agreement with the Declaration of Helsinki, and with local institutional research policies, procedures, and regulations. The study design was developed by academic and clinical expertise underpinned by scientific evidence that emphasizes the requirement for standardized measurement of PBs and EBs. Patients or the public were not involved in the development of the study protocol.

Ethics committee approval was granted in each of the 4 study centers prior to study commencement. In the University of Limerick, ethics committee approval was granted by the Education and Health Sciences Faculty Research Ethics Committee (approval no 22_09_10_EHS_); in Bremen, approval was granted by the Ethics Committee of the University of Bremen (approval no 2022 - 25); in the Czech Republic, approval was granted by the Committee for Research Ethics at the University of Hradec Kralove (approval no 11/2022); and in France, approval was by Comité de Protection des Personnes CPP Ile-de-France VI (approval no 2022-A02208-35).

All participants received detailed information about the purpose, procedures, risks, and benefits of the study prior to participation. Written informed consent was obtained from all participants before data collection commenced. Participants were informed of their right to withdraw at any time without penalty. Participants' privacy and confidentiality were strictly protected. Identifying details (including names, initials, , or other personal identifiers) have been omitted.

JPI-WEALTH partners abided by the guidelines on consent [46] and transparency [47] as adopted by the European Data Protection Board and will remain in line with any new guidance that may be issued by the European Data Protection Board. The JPI-WEALTH project partners conform to the relevant convention and protocols of the Council of Europe, including the Convention on Human Rights and Biomedicine [48], and protocols on Human Rights and Biomedicine [49]. In addition, data collection was in accordance with the International Epidemiological Association's Guidelines for Proper Conduct of Epidemiological Research [50] and the Council for International Organizations of Medical Science's International Ethical Guidelines for Health-related Research Involving Humans [51].

Results

The WEALTH Project was funded by the Joint Programming Initiative a Healthy Diet for a Healthy Life, a research and innovation initiative of EU member states and associated countries under grant agreement number 727565, under STAMIFY. The data collection in 2023 to 2024 resulted in rich datasets including 627 participants providing data from 2 research-grade and 2 consumer-grade devices, lab-based activity measurement, a 9-day free-living data collection period which combined the assessment of PBs and EBs via wearable devices and time-based, event-based, and self-initiated EMA, complemented by three 24HDRs. The final sample includes 627 participants, of which 44% (n=275) were male. The mean age was 32.7 (SD 13.3) years, and the mean body mass index was 24.5 (SD 4.0) kg/m². Data processing and machine-learning model development are currently underway, with primary results expected to be published in 2026.

Discussion

Principal Findings

The WEALTH project is expected to generate one of the most comprehensive multicountry datasets integrating accelerometer-derived PBs, wearable-derived EBs, and EMA data collected in both controlled and free-living settings. Based on the study design, we anticipate producing (1) harmonized, labeled datasets for PB and EB classification, (2) ML models capable of classifying behavioral states from raw accelerometry, and (3) an evidence-based methodology to support the simultaneous measurement of PBs and EBs in real-world contexts. These data will support the development of harmonized, labeled datasets and ML models capable of

XSL•FO
RenderX

classifying PBs and EBs from raw accelerometry. This is essential given longstanding concerns about the lack of standardized device protocols, heterogeneity in cut points, and poor comparability across studies and countries [5,6,45,46]. The outputs from WEALTH will be accessible through a dedicated website outlining the project's aims, methods, and products [52]. Critically, the ML models developed will be publicly available, enabling researchers to apply them to similar data. Providing openly accessible models and processing frameworks will facilitate detailed monitoring and assessment of PBs, EBs, and their determinants, thereby supporting future research beyond the project's duration and contributing to data sharing and protocol harmonization.

The wearable dataset collected in this study will be made publicly available following the completion of the study and the end of an embargo period, in line with data protection and ethical approvals. The data will be shared following the Findable, Accessible, Interoperable, and Reusable (FAIR)principles, starting with a comprehensive data catalog and detailed data description, to facilitate transparency, reproducibility, and further research in this domain. During the embargo period, these data are available upon reasonable request in agreement with the research team.

Prior research has demonstrated the potential of ML to improve the classification of PB beyond traditional cut-point methods [11,12,45]. However, most existing work has been constrained to single-device datasets, laboratory environments, or narrow demographic groups. Similarly, while emerging studies have shown that wrist-worn sensors can help detect eating episodes and components of meal behaviors [15-18], free-living detection remains challenging due to variability in movement patterns and context.

Existing EMA studies have demonstrated strong value in capturing contextual, emotional, and environmental correlates of behavior [4,19] but have rarely integrated automated event-triggered surveys based on wearable sensor data. Previous work has also been limited by participants needing to self-initiate logs, contributing to recall bias and incomplete data [4]. The WEALTH design advances this field by combining 4 wearable devices, automated EMA triggers, multicountry standardization, 24HDR data, and cross-device timestamp harmonization. To our knowledge, this integrated approach has not previously been attempted at this scale in free-living behavioral research. This level of methodological integration directly addresses persistent methodological challenges documented in the literature, including cross-device comparability and harmonization of accelerometry outputs [5,6], the limited contextual fidelity of self-report and EMA-based behavioral monitoring [4,19], and the substantial complexities associated with accurately detecting and classifying EBs using wearable sensing technologies in free-living settings [15-18].

The aspiration is for WEALTH methodologies to be implemented in national and international nutritional and PA surveillance efforts. The ultimate goal is to establish a more robust system for assessing real-world data on PBs, EBs, and their determinants. This enhanced assessment capability will aid in designing public health interventions that promote health-enhancing PA and healthy diets, particularly in at-risk populations. These methodological advancements are expected to build a stronger evidence base on the health impacts of PBs and EBs, informing the development of improved national and European guidelines and recommendations aimed at enhancing citizen health. Furthermore, this accessibility is intended to support national and international efforts to achieve the United Nations Sustainable Development Goals by informing public health interventions and policies targeting PA and healthy diets in at-risk populations.

The project will deliver several key products and methodologies:

- The WEALTH taxonomy: a comprehensive classification of daily PB and EB.
- The WEALTH protocols: standardized measurement methods and integration protocols for PBs and EBs, adaptable for use in at least 4 different countries and languages.
- The WEALTH web-based infrastructure: tools for the instant processing, aggregation, and analysis of accelerometer data to classify PBs and EBs.
- The WEALTH ML models: open-access models for data processing, available via the WEALTH toolbox on the WEALTH website.
- The WEALTH detailed reports: standardized and interpretable data translation from multiple devices, facilitated by the WEALTH toolbox.
- Recommendations: guidelines for using WEALTH methods in large-scale epidemiological studies, mass surveillance, and the evaluation of intervention studies.
- Training and capacity building resources: resources for using the WEALTH toolbox, including a workshop at the final symposium and support through online teaching methodologies.

These outputs are designed to maximize the use of new knowledge and methods developed within WEALTH by project stakeholders and European Union citizens, ultimately contributing to public health improvement and policymaking. By ensuring both ML models and datasets are publicly available, WEALTH actively promotes reproducibility, transparency, and wider scientific collaboration.

Strengths and Limitations

We used a standardized protocol across multiple countries to capture both lab-based and free-living PBs and EBs. The combination of research-grade and consumer-grade wearable devices enabled the comprehensive monitoring of participant activities and physiological responses. The 9-day EMA and 24HDR provided rich, real-time data for examining behavioral patterns in real-life contexts. The backend engineering protocols of HealthReact are proprietary due to commercial licensing, which may limit direct replicability. The intensive data collection and device reliance may have introduced participant burden and technological challenges, affecting data quality. Self-reported EMA and dietary data could be subject to recall and social desirability bias, despite correction efforts.

XSL•FO
RenderX

Conclusion

This protocol outlines the design and methodology of the WEALTH project, a multicountry initiative to develop standardized, integrated approaches for assessing PBs and EBs using wearable sensors, ecological momentary assessment, and ML techniques. Although the findings are forthcoming, the anticipated outputs include harmonized datasets, openly accessible ML models, and an integrated data processing framework. These outputs have the potential to strengthen behavioral surveillance, support epidemiological and interventional research, and inform public health guidance. By combining rigorous standardization with advanced sensing and analytical methods, WEALTH is positioned to make a substantive contribution to the harmonization of PB and EB measurement across Europe and to future digital health innovation.

Acknowledgments

We would like to express our sincere appreciation to the volunteers and field staff in the Czech Republic, France, Germany, and Ireland for their invaluable participation. We are deeply grateful to the WEALTH (Wearable Sensor Assessment of Physical and Eating Behaviours) Advisory Panel for their essential guidance and unwavering support throughout this study. A special thanks goes to Ms Ella Tyrrell for her outstanding project management, steadfast support, and significant contributions to the success of this research.

Funding

The WEALTH (Wearable Sensor Assessment of Physical and Eating Behaviours) project is funded by the Joint Programming Initiative a Healthy Diet for a Healthy Life, a research and innovation initiative of European Union member states and associated countries under grant agreement number 727565, under STAMIFY (standardised measurement, monitoring and/or biomarkers to study food intake, physical activity, and health). The funding agencies supporting this work are (in alphabetical order of participating countries) as follows: Belgium: Research Foundation – Flanders; Czechia: Ministry of Education, Youth and Sports; France: French National Research Agency; Germany: Federal Ministry of Education and Research; and Ireland: Health Research Board.

Data Availability

The data are available from the corresponding author on reasonable request.

Authors' Contributions

Each of the authors have participated at all stages of this article and conconcurs with the content in the final manuscript. GH (lead author) has undertaken and completed all aspects of the research and writing process and thus takes full responsibility for the paper. AH and AD are joint Principle Investigators, and shared last authorship.

Conflicts of Interest

RC is the owner of a company “TallWell s.r.o.” that holds a license to use HealthReact, a software developed by the University of Hradec Králové. All remaining authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Dumuid D, Stanford TE, Martin-Fernández JA, et al. Compositional data analysis for physical activity, sedentary time and sleep research. Stat Methods Med Res 2018 Dec;27(12):3726-3738. [doi: 10.1177/0962280217710835] [Medline: 28555522]
2. Sun J, Li Y, Zhao M, et al. Association of the American Heart Association’s new “Life’s Essential 8” with all-cause and cardiovascular disease-specific mortality: prospective cohort study. BMC Med 2023 Mar 29;21(1):116. [doi: 10.1186/s12916-023-02824-8] [Medline: 36978123]
3. Dowd KP, Szeklicki R, Minetto MA, et al. A systematic literature review of reviews on techniques for physical activity measurement in adults: a DEDIPAC study. Int J Behav Nutr Phys Act 2018 Feb 8;15(1):15. [doi: 10.1186/s12966-017-0636-2] [Medline: 29422051]
4. Bruening M, van Woerden I, Todd M, Brennhofer S, Laska MN, Dunton G. A mobile ecological momentary assessment tool (devilSPARC) for nutrition and physical activity behaviors in college students: a validation study. J Med Internet Res 2016 Jul 27;18(7):e209. [doi: 10.2196/jmir.5969] [Medline: 27465701]
5. Loyen A, Verloigne M, Van Hecke L, et al. Variation in population levels of sedentary time in European adults according to cross-European studies: a systematic literature review within DEDIPAC. Int J Behav Nutr Phys Act 2016 Jun 28;13(1):71. [doi: 10.1186/s12966-016-0397-3] [Medline: 27350251]
6. Hayes G, Dowd K, MacDonncha C, Donnely A. Simultaneous validation of count-to-activity thresholds for five commonly used activity monitors in adolescent research: a step toward data harmonization. J Meas Phys Behav 2021;4(4):333-342. [doi: 10.1123/jmpb.2021-0023]

XSL•FO
RenderX

7. Van Laerhoven K, Wenzel M, Geelen A, et al. Experiences from a wearable-mobile acquisition system for ambulatory assessment of diet and activity. Presented at: iWOAR '17; Sep 21-22, 2017. [doi: 10.1145/3134230.3134239]
8. Martinato M, Lorenzoni G, Zanchi T, et al. Usability and accuracy of a smartwatch for the assessment of physical activity in the elderly population: observational study. JMIR Mhealth Uhealth 2021 May 5;9(5):e20966. [doi: 10.2196/20966] [Medline: 33949953]
9. O'Driscoll R, Turicchi J, Hopkins M, et al. The validity of two widely used commercial and research-grade activity monitors, during resting, household and activity behaviours. Health Technol 2020 May;10(3):637-648. [doi: 10.1007/s12553-019-00392-7]
10. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: systematic review. JMIR Mhealth Uhealth 2020 Sep 8;8(9):e18694. [doi: 10.2196/18694] [Medline: 32897239]
11. Farrahi V, Niemelä M, Kangas M, Korpelainen R, Jämsä T. Calibration and validation of accelerometer-based activity monitors: a systematic review of machine-learning approaches. Gait Posture 2019 Feb;68:285-299. [doi: 10.1016/j.gaitpost.2018.12.003] [Medline: 30579037]
12. Narayanan A, Desai F, Stewart T, Duncan S, Mackay L. Application of raw accelerometer data and machine-learning techniques to characterize human movement behavior: a systematic scoping review. J Phys Act Health 2020 Mar 1;17(3):360-383. [doi: 10.1123/jpah.2019-0088] [Medline: 32035416]
13. Mackintosh KA, Montoye AHK, Pfeiffer KA, McNarry MA. Investigating optimal accelerometer placement for energy expenditure prediction in children using a machine learning approach. Physiol Meas 2016 Oct;37(10):1728-1740. [doi: 10.1088/0967-3334/37/10/1728] [Medline: 27653339]
14. Bassett DR. Device-based monitoring in physical activity and public health research. Physiol Meas 2012 Nov;33(11):1769-1783. [doi: 10.1088/0967-3334/33/11/1769] [Medline: 23110847]
15. Heydarian H, Adam M, Burrows T, Collins C, Rollo ME. Assessing eating behaviour using upper limb mounted motion sensors: a systematic review. Nutrients 2019 May 24;11(5):1168. [doi: 10.3390/nu11051168] [Medline: 31137677]
16. Tang Z, Patyk A, Jolly J, Goldstein SP, Thomas JG, Hoover A. Detecting eating episodes from wrist motion using daily pattern analysis. IEEE J Biomed Health Inform 2024;28(2):1054-1065. [doi: 10.1109/JBHI.2023.3341077]
17. Wang C, Kumar TS, De Raedt W, Camps G, Hallez H, Vanrumste B. Eating speed measurement using wrist-worn imu sensors towards free-living environments. IEEE J Biomed Health Inform 2024 Oct;28(10):5816-5828. [doi: 10.1109/JBHI.2024.3422875] [Medline: 38959146]
18. Skinner A, Toumpakari Z, Stone C, Johnson L. Future directions for integrative objective assessment of eating using wearable sensing technology. Front Nutr 2020;7:80. [doi: 10.3389/fnut.2020.00080] [Medline: 32714939]
19. Dunton GF. Ecological momentary assessment in physical activity research. Exerc Sport Sci Rev 2017 Jan;45(1):48-54. [doi: 10.1249/JES.0000000000000092] [Medline: 27741022]
20. Healthy Diet Healthy Life. URL: https://www.healthydietforhealthylife.eu/ [accessed 2026-02-28]
21. Warburton DE, Jamnik V, Bredin SSD, Shephard RJ, Gledhill N. The 2018 Physical Activity Readiness Questionnaire for Everyone (PAR-Q+) and electronic Physical Activity Readiness Medical Examination (ePARmed-X+). Health Fit J Can 2018;11(1):31-34. [doi: 10.14288/hfjc.v11i1.260]
22. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [doi: 10.1097/00005650-199206000-00002]
23. Boadella JM, Kuijer PP, Sluiter JK, Frings-Dresen MH. Effect of self-selected handgrip position on maximal handgrip strength. Arch Phys Med Rehabil 2005 Feb;86(2):328-331. [doi: 10.1016/j.apmr.2004.05.003] [Medline: 15706562]
24. Marion R, Niebuhr BR. Effect of warm-up prior to maximal grip contractions. J Hand Ther 1992 Jul;5(3):143-146. [doi: 10.1016/S0894-1130(12)80349-8]
25. Johansson CA, Kent BE, Shepard KF. Relationship between verbal command volume and magnitude of muscle contraction. Phys Ther 1983 Aug;63(8):1260-1265. [doi: 10.1093/ptj/63.8.1260] [Medline: 6878436]
26. Balogun JA, Akomolafe CT, Amusa LO. Grip strength: effects of testing posture and elbow position. Arch Phys Med Rehabil 1991 Apr;72(5):280-283. [Medline: 2009042]
27. Watanabe T, Owashi K, Kanauchi Y, Mura N, Takahara M, Ogino T. The short-term reliability of grip strength measurement and the effects of posture and grip span. J Hand Surg Am 2005 May;30(3):603-609. [doi: 10.1016/j.jhsa.2004.12.007] [Medline: 15925174]
28. Edwardson CL, Winkler EAH, Bodicoat DH, et al. Considerations when using the activPAL monitor in field-based research with adult populations. J Sport Health Sci 2017 Jun;6(2):162-178. [doi: 10.1016/j.jshs.2016.02.002] [Medline: 30356601]
29. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S531-S543. [doi: 10.1249/01.mss.0000185657.86065.98] [Medline: 16294116]
30. Aadland E, Ylvisåker E. Reliability of the Actigraph GT3X+ accelerometer in adults under free-living conditions. PLoS ONE 2015;10(8):e0134606. [doi: 10.1371/journal.pone.0134606] [Medline: 26274586]
31. Moreno-Pino F, Porras-Segovia A, López-Esteban P, Artés A, Baca-García E. Validation of Fitbit Charge 2 and Fitbit Alta HR against polysomnography for assessing sleep in adults with obstructive sleep apnea. J Clin Sleep Med 2019 Nov 15;15(11):1645-1653. [doi: 10.5664/jcsm.8032] [Medline: 31739855]

XSL•FO
RenderX

32. Muellmann S, Bragina I, Voelcker-Rehage C, et al. Development and evaluation of two web-based interventions for the promotion of physical activity in older adults: study protocol for a community-based controlled intervention trial. BMC Public Health 2017 May 25;17(1):512. [doi: 10.1186/s12889-017-4446-x] [Medline: 28545506]
33. Janek M, Kuhnova J, Cardon G, et al. Ecological momentary assessment of physical and eating behaviours: the WEALTH feasibility and optimisation study with recommendations for large-scale data collection. PLoS ONE 2025;20(2):e0318772. [doi: 10.1371/journal.pone.0318772] [Medline: 39932907]
34. Delobelle J, Lebuf E, Dyck DV, et al. Fitbit's accuracy to measure short bouts of stepping and sedentary behaviour: validation, sensitivity and specificity study. Digit Health 2024;10:20552076241262710. [doi: 10.1177/20552076241262710] [Medline: 38894943]
35. Etude NutriNet-Santé [Website in French]. URL: https://etude-nutrinet-sante.fr/ [accessed 2026-02-28]
36. Hebestreit A, Wolters M, Jilani H, Eiben G, Pala V. Web-based 24-h dietary recall: the SACANA program. In: Instruments for Health Surveys in Children and Adolescents: Springer; 2019:77-102. [doi: 10.1007/978-3-319-98857-3_5]
37. Andreeva VA, Deschamps V, Salanave B, et al. Comparison of dietary intakes between a large online cohort study (Etude NutriNet-Santé) and a nationally representative cross-sectional study (Etude Nationale Nutrition Santé) in France: addressing the issue of generalizability in e-epidemiology. Am J Epidemiol 2016 Nov 1;184(9):660-669. [doi: 10.1093/aje/kww016] [Medline: 27744386]
38. Lassale C, Castetbon K, Laporte F, et al. Validation of a Web-based, self-administered, non-consecutive-day dietary record tool against urinary biomarkers. Br J Nutr 2015 Mar 28;113(6):953-962. [doi: 10.1017/S0007114515000057] [Medline: 25772032]
39. Pala V, Reisch LA, Lissner L. Dietary behaviour in children, adolescents and families: the eating habits questionnaire (EHQ). In: Instruments for Health Surveys in Children and Adolescents: Springer; 2019:103-133. [doi: 10.1007/978-3-319-98857-3_6]
40. Shaw PA, Deffner V, Keogh RH, et al. Epidemiologic analyses with error-prone exposures: review of current practice and recommendations. Ann Epidemiol 2018 Nov;28(11):821-828. [doi: 10.1016/j.annepidem.2018.09.001] [Medline: 30316629]
41. Kipnis V, Midthune D, Buckman DW, et al. Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. Biometrics 2009 Dec;65(4):1003-1010. [doi: 10.1111/j.1541-0420.2009.01223.x] [Medline: 19302405]
42. Tooze JA, Midthune D, Dodd KW, et al. A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. J Am Diet Assoc 2006 Oct;106(10):1575-1587. [doi: 10.1016/j.jada.2006.07.003] [Medline: 17000190]
43. Sigcha L, Swenne A, Kuhnova J, et al. Data labelling for free-living physical activity recognition using thigh-worn wearables and event-based ecological momentary assessment. Research Square. Preprint posted online on Aug 5, 2025. [doi: 10.21203/rs.3.rs-6835979/v1]
44. Stevens ML, Gupta N, Inan Eroglu E, et al. Thigh-worn accelerometry for measuring movement and posture across the 24-hour cycle: a scoping review and expert statement. BMJ Open Sport Exerc Med 2020;6(1):e000874. [doi: 10.1136/bmjsem-2020-000874] [Medline: 33408875]
45. Chan S, Hang Y, Tong C, et al. CAPTURE-24: a large dataset of wrist-worn activity tracker data collected in the wild for human activity recognition. Sci Data 2024 Oct 16;11(1):1135. [doi: 10.1038/s41597-024-03960-3] [Medline: 39414802]
46. Guidelines 05/2020 on consent under regulation 2016/679. : European Data Protection Board; 2020 URL: https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf [accessed 2026-02-25]
47. Guidelines on transparency under regulation 2016/679. : European Data Protection Board; 2017 URL: https://www.edpb.europa.eu/system/files/2023-09/wp260rev01_en.pdf [accessed 2026-02-25]
48. Council of Europe. Convention for protection of human rights and dignity of the human being with regard to the application of biology and biomedicine: convention on human rights and biomedicine. Kennedy Inst Ethics J 1997 Sep;7(3):277-290. [doi: 10.1353/ken.1997.0021]
49. Council of Europe. Additional protocol to the convention on human rights and biomedicine, concerning biomedical research. Jahrbuch für Wissenschaft und Ethik 2005 Oct 17;10(1):391-402. [doi: 10.1515/9783110182521.391]
50. IEA guidelines for proper conduct of epidemiological research. Yumpu. 2007. URL: https://www.yumpu.com/en/document/view/19409945/iea-guidelines-for-proper-conduct-of-epidemiological-research-hrsa [accessed 2026-02-25]
51. International ethical guidelines for health-related research involving humans. : Council for International Organizations of Medical Sciences; 2016 URL: https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf [accessed 2026-02-25]
52. WEALTH. URL: https://wealth-stamify.com/ [accessed 2026-02-28]

Abbreviations

24HDR: 24-hour dietary recall
EB: eating behavior
EMA: ecological momentary assessment

XSL•FO
RenderX

ML: machine learning
PA: physical activity
PB: physical behavior
SACANA: Self-Administered Children, Adolescent and Adults Nutrition Assessment
STAMIFY: standardised measurement, monitoring and/or biomarkers to study food intake, physical activity and health
WEALTH: Wearable Sensor Assessment of Physical and Eating Behaviours

Edited by J Sarvestan; submitted 17.Dec.2024; peer-reviewed by C Wang, J Hernandez, M Liu; revised version received 09.Dec.2025; accepted 10.Dec.2025; published 06.Mar.2026.

Please cite as:
Hayes G, Buck C, Cardon G, Cimler R, Elavsky S, Fezeu K L, Harrington JM, Kühnová J, Oppert JM, Sigcha L, Van de Ven P, Vetrovsky T, Woods CB, Hebestreit A, Donnelly AE
Standardized Methods for Evaluating Physical and Eating Behaviors: The WEALTH Cross-Sectional Study Protocol
JMIR Res Protoc 2026;15:e70186
URL: https://www.researchprotocols.org/2026/1/e70186
doi:10.2196/70186

© Grainne Hayes, Christoph Buck, Greet Cardon, Richard Cimler, Steriani Elavsky, Leopold Fezeu K, Janas M Harrington, Jitka Kühnová, Jean-Michel Oppert, Luis Sigcha, Pepijn Van de Ven, Tomas Vetrovsky, Catherine B Woods, Antje Hebestreit, Alan E Donnelly. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 6.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

International Clinical Practice Guidelines on Traditional Chinese Medicine Therapy for Dysmenorrhea: Development Protocol

Li Zhang[1], PhD; Xiao Liang[2], PhD; Junlu Chen[3], MMed; Yuhan Du[3], MMed; Huilan Du[2], PhD; Xiaohui Wang[4], PhD

[1]First Clinical School (Affiliated Hospital), Shaanxi University of Traditional Chinese Medicine, Xianyang, Shaanxi Province, China

[2]Collaborative Innovation Center of Integrated Chinese and Western Medicine on Reproductive Disease, Hebei Key Laboratory of Integrative Medicine on Liver-kidney Patterns, College of Integrated Traditional Chinese and Western Medicine, Hebei University of Chinese Medicine, No 326, Xinshi South Road, Qiaoxi District, Shijiazhuang, Hebei, China

[3]Graduate School, Hebei University of Chinese Medicine, Shijiazhuang, China

[4]School of Public Health, Lanzhou University, Lanzhou, China

Corresponding Author:
Huilan Du, PhD
Collaborative Innovation Center of Integrated Chinese and Western Medicine on Reproductive Disease, Hebei Key Laboratory of Integrative Medicine on Liver-kidney Patterns, College of Integrated Traditional Chinese and Western Medicine, Hebei University of Chinese Medicine, No 326, Xinshi South Road, Qiaoxi District, Shijiazhuang, Hebei, China

Abstract

Background: Traditional Chinese medicine (TCM) functions according to the concepts of “holism” and treatment based on syndrome differentiation, and it has achieved good clinical results in treating patients with dysmenorrhea, which is a common gynecological disorder. However, there are currently no international clinical practice guidelines involving TCM therapies for dysmenorrhea. This study aims to establish a protocol for the development of such guidelines.

Objective: This protocol will provide a road map for the development of the first international clinical practice guidelines on TCM therapy for dysmenorrhea.

Methods: The guidelines will be developed with reference to the *General Rules of Preparation of Diagnosis and Treatment Guideline in Traditional Chinese Medicine* and the *Western Medicine Guideline and Traditional Chinese Medicine Guideline: Improving Together in Mutual Learning*. They will be developed in accordance with the Appraisal of Guidelines for Research and Evaluation II and the World Health Organization guideline handbook and will make recommendations based on systematic reviews. We have established a guideline working group and will formulate clinical questions using the participants, interventions, comparisons, outcomes, and study design framework. The recommendations will be developed through evidence retrieval, synthesis, and the Delphi method. We will consider the preferences and values of patients, as well as the costs and the pros and cons of interventions.

Results: This work was supported by the National Key R&D Program of China (grant 2019YFC1712000) in 2019. The writing group for guidelines was formed in January 2021. The literature search and screening process began in May 2022. So far, the literature search and 2 rounds of the Delphi method have been completed. The protocol will provide a road map for the development of the first international clinical practice guidelines on TCM therapy for dysmenorrhea.

Conclusions: The guideline will be developed in accordance with this protocol, which may provide support and evidence for TCM treatment in patients with dysmenorrhea. There is currently a need for clinical practice guidelines in TCM for the treatment of dysmenorrhea.

International Registered Report Identifier (IRRID): DERR1-10.2196/77423

(JMIR Res Protoc 2026;15:e77423) doi:10.2196/77423

KEYWORDS

endometriosis; adenomyosis; dysmenorrhea; traditional Chinese medicine; international clinical practice guidelines; protocol

Introduction

Description of Dysmenorrhea

Dysmenorrhea is a common disorder characterized by painful cramps that occur with menstruation and lower abdominal pain, and it may be accompanied by other symptoms such as nausea, diarrhea, headache, and lightheadedness. The prevalence of dysmenorrhea is highest in adolescent girls, with estimates ranging from 45% to 97% in women of different ages and nationalities [1]. However, only a small percentage of women with this disorder consult a physician about their condition, and many choose to self-medicate [2]. Up to 15% of women with dysmenorrhea have symptoms that are severe enough to cause

XSL•FO
RenderX

absences from school or work, as well as restrictions on social, academic, and sports activities [3]. Dysmenorrhea is often underreported and undertreated, and proper dysmenorrhea management will improve the quality of life and alleviate the academic and financial burdens of many women [4].

Dysmenorrhea is classified as primary or secondary dysmenorrhea according to the identification of the underlying cause [5]. Primary dysmenorrhea is defined as pain without any organic cause and is most common in women aged <25 years, with pain beginning within 3 years after menarche. The duration of the pain is commonly 8 to 72 hours and is typically related to the onset of menstrual cramps. In contrast, secondary dysmenorrhea is more likely to occur at any time after menarche but may appear as a new symptom during a woman's third and fourth decades, after the onset of an underlying causative condition [6]. The pain is related to conditions such as leiomyomas, endometriosis, and pelvic inflammatory disease [3]. The physiological causes of dysmenorrhea include increased levels of prostaglandin in the menstrual fluid [7] and overproduction of vasopressin, a hormone that stimulates uterine muscle contraction [8]. Secondary dysmenorrhea is caused by specific pelvic pathology, and its treatment focuses on the causative pelvic pathology or medical diseases [3]. Risk factors for dysmenorrhea include heavy menstrual flow, nulliparity, depression, and smoking [9].

Description of Interventions

There are multiple treatment options for dysmenorrhea, many of which have reasonable efficacy and safety profiles. Therefore, there is no one-size-fits-all algorithm for the treatment of dysmenorrhea. Specific factors, including use of contraception, contraindications, ease of use, patient preferences, potential side effects, and cost, can help guide the choice of initial treatment [3]. Nonsteroidal anti-inflammatory drugs and hormonal contraceptives are considered first-line treatments for dysmenorrhea, and they provide effective pain relief for most women [10]. However, the long-term use of nonsteroidal anti-inflammatory drugs and hormonal contraception may have adverse effects, and some patients have contraindications to the use of these medications, including a history of thrombosis or digestive problems. In addition, approximately 10% to 18% of women affected by dysmenorrhea do not respond to these medications. In all these cases, women seek alternatives, including vasodilators, antispasmodics, oxytocin, and vasopressin receptor antagonists, as well as complementary or alternative therapy [11]. The use of complementary or alternative therapy has been welcomed by both consumers and mainstream medical practitioners [6]. These interventions include heat, Chinese herbal medicine, acupuncture, transcutaneous electrical nerve stimulation, yoga, and exercise [12].

Traditional Chinese medicine (TCM) is considered to be a viable option for dysmenorrhea treatment. It has been used in China and many countries in Southeast Asia for a long time. In recent years, it has been used worldwide due to its lack of side effects, safety, and efficacy. TCM uses herbal medicine combined with acupuncture, massage, and acupoint stimulation to treat disease, and these approaches have demonstrated therapeutic benefits. However, there is a worldwide lack of clinical research and clinical practice guidelines for TCM treatment of dysmenorrhea.

Principles of TCM

Dysmenorrhea is a common gynecological disorder, with qi-blood circulation running not fluently, deficiency of both qi and blood, or asthenia of liver and kidney as the main pathogenetic factors, and the disease locations are the uterus and the thoroughfare-conception meridian. Dysmenorrhea occurs along with the menstrual cycle and is related to the specific physiological state of the menstrual period itself, as well as before and after the menstrual period, that is, during the period of nonmenstruation. The pathogenic factors and the insufficient qi-blood harmonization lead to qi stagnation and blood stasis or insufficiency. Before and after menstruation, blood changes from fullness to overflow, during which time the body becomes more vulnerable to pathogenic influences. This may result in impaired circulation of qi and blood or underlying qi-blood deficiency. The treatment is based on the principle of coordinating the thoroughfare-conception meridian and harmonizing the qi-blood. During the menstruation period, the emphasis is on promoting blood circulation for removing blood stasis and relieving pain to achieve symptom relief. The medicine is usually taken 3 to 5 days before the pain is expected and is continued until the pain is relieved. The emphasis is on treatment upon syndrome differentiation to achieve resolution of the symptoms in chronic cases. This requires continuous treatment for more than 3 menstrual cycles.

Syndrome Differentiation of Dysmenorrhea in TCM

Treatment upon syndrome differentiation is the TCM principle used for the recognition and treatment of diseases. Syndrome differentiation should be based on the time of occurrence of dysmenorrhea and the nature and degree of the pain, combined with menstrual conditions and systemic syndromes to distinguish patterns of cold vs heat and deficiency vs excess. TCM usually classifies dysmenorrhea into the following syndromes: (1) qi stagnation and blood stasis syndrome, where delayed qi movement causes disorders of blood circulation, with the pathological changes coexisting with qi stagnation and blood stasis; (2) deficiency of both qi and blood syndrome, characterized by pathological changes in parts of the body that are deprived of nourishment and thus have reduced function; (3) cold congealing and blood stasis syndrome, where cold causes congealing, affecting qi and causing the blood to stagnate, and where the meridians are blocked; (4) cold damp stagnation syndrome, where, due to the intrusion of cold and moisture, the moisture cannot be excreted smoothly, resulting in blood stasis symptoms; (5) liver-kidney depletion syndrome, where there is deficiency in the liver and kidney, resulting in deficiency of the thoroughfare-conception meridian, and where the sea of blood is empty, and the uterus and thoroughfare-conception meridian lose nourishment; (6) damp-heat stasis syndrome, where accumulated heat and damp rush into the uterus, the circulation of qi and blood is blocked, the sea of blood before menstruation is full of qi and blood, the heat and damp are intertwined with blood, and stagnation cannot be overcome; and (7) yang deficiency and internal cold syndrome, where there is weakness

XSL•FO
RenderX

of the kidney yang, endogenous cold pathogen, loss of warmth in the uterus, and impaired circulation of qi and blood.

Objective

This protocol aims to provide a road map for the international guideline for the diagnosis and treatment of dysmenorrhea in TCM. There is a worldwide lack of clinical research and clinical practice guidelines for the TCM treatment of dysmenorrhea. These guidelines would represent the first international clinical practice guidelines on TCM therapy for dysmenorrhea and will provide substantial evidence and standardized guidance for treating dysmenorrhea using TCM. As this guideline aims to develop a TCM symptom guideline for dysmenorrhea, it is divided into primary and secondary dysmenorrhea (endometriosis and adenomyosis) for research.

Methods

Principles

The study will reference the *General Rules of Preparation of Diagnosis and Treatment Guideline in Traditional Chinese Medicine* [13-15] and the *Western Medicine Guideline and Traditional Chinese Medicine Guideline: Improving Together in Mutual Learning* [16,17]. It will follow the Chinese Clinical Guidelines Evaluation System (AGREE-China) [18-20] and the World Health Organization guideline handbook [21] and will make recommendations based on systematic reviews. We have established a guideline working group and will formulate clinical questions using the participants, interventions, comparisons, outcomes, and study design (PICOS) framework. The recommendations will be developed through evidence retrieval, synthesis, and the Delphi method. We will also consider the preferences and values of the patients, as well as the cost and the pros and cons of the interventions. The guideline has been registered on the International Practice Guidelines Registry Platform (registration number IPGRP-2021CN138).

Guideline Development Institutions, Target Users, and Population

The guidelines will be developed by Hebei University of Chinese Medicine. The Evidence-Based Medicine Center of Lanzhou University will provide methodological support. The target users of the guidelines are Chinese and Western medicine obstetricians and gynecologists, TCM physicians, physicians of integrated Chinese and Western medicine, nursing staff, and Chinese pharmacists. The target population is patients who are clinically diagnosed with primary dysmenorrhea or endometriosis or adenomyosis.

Guideline Working Group

The guideline working group was established in June 2021 and consists of 6 subgroups: the guideline steering committee, conflict of interest management committee, consensus expert group, secretary group, evidence evaluation group, and external review group.

The steering committee consists of 7 members, including 2 masters of TCM, 4 gynecologists practicing TCM, and 1 expert in evidence-based medicine. The guideline steering committee will be mainly responsible for (1) determining the topic and scope of the guideline; (2) establishing other working groups for the guideline; (3) managing conflicts of interest; (4) approving the guideline plan, recommendations, and full text; (5) supervising the development process of the guideline; (6) documenting the guideline development process; (7) writing and submitting the first draft of the guideline; and (8) monitoring and evaluating the published guideline.

The conflict of interest management committee is composed of 3 members who will manage, evaluate, and approve the declaration of interest at each stage of the guideline formulation. The main principles are as follows: (1) as far as possible, members with conflicts of interest should not be included; (2) all participants in the guideline development process will be required to declare their own and their family members' conflicts of interest and continue to update them before participating in any link; and (3) the conflict of interest statement should report all possible conflicts of interest in the previous 3 years.

Taking into account the balance of disciplines, majors, gender, and geographical distribution, we established a consensus expert group, which is composed of 30 experts in related fields, including 2 methodological experts, 16 Chinese medicine experts, 10 Western medicine experts, and 2 pharmaceutical experts. This group is mainly responsible for the determination of clinical issues, conducting consensus and votes on recommendations, and the finalization of the full text of the guidelines.

The secretary group consists of 5 members. These 5 members are mainly responsible for (1) coordinating the work of other working groups, (2) drafting the guideline plan, (3) conducting research into clinical problems, (4) organizing recommendation consensus meetings, (5) recording the entire process of guideline formulation in detail, and (6) writing the first draft of the guideline and submitting the guideline.

The evidence evaluation group consists of 7 members. The 7 members are mainly responsible for (1) retrieving, evaluating, synthesizing, and grading evidence; (2) the production of a systematic review; (3) the production of evidence summary tables; and (4) the production of recommendation and opinion decision tables.

The external review group is composed of 5 members who are mainly responsible for reviewing the final version of the guideline; ensuring its scientific worth, clarity, and fairness; and giving feedback and suggestions on the major risks or problems of the guideline and the content of specific recommendations.

Several foreign guideline development agencies, including the World Health Organization, National Institute for Health and Care Excellence, and Scottish Intercollegiate Guidelines Network [22-24], require the participation of patient representatives. In the guideline formulation process, the participation of patients in the formulation of the scope of the guideline and clinical problems can assist in ensuring that the problems to be solved by the guideline are relevant to the patients [25]. Therefore, according to these requirements of guideline development, the guideline working group recruited 2 patients. The 2 patients participated in raising questions,

XSL•FO
RenderX

evaluating evidence, and forming recommendations for the guideline. During the question-raising stage, the guideline development team understood patients' perspectives on the research topics, helping to ensure that the chosen research questions truly align with clinical practice and patients' concerns. In the evidence collection and evaluation stage, patients' opinions were sought to understand their real experiences and needs during actual treatment. In the recommendation formation stage, the guideline development team could directly learn patients' views and feedback on the recommendations, further clarifying the key factors that should be considered in treatment decisions. In addition, the guideline development team could survey different patient groups to obtain information on variations in treatment preferences, ensuring that the recommendations can meet the needs of different patients. Draft recommendations are published in clear and understandable language, inviting patients to provide feedback to ensure that the recommendations not only align with the latest evidence but also address patients' real concerns.

Declaration of Interests

All members of the guideline steering committee, consensus expert group, secretarial group, external review group, and other experts or consultants participating in the guideline development process will be required to fill out the conflict of interest declaration form before participating in the guideline development work to ensure that potential conflicts of interest are declared and disclosed.

Formulating Questions and Choosing Outcomes

According to the topic of the guideline, the guideline steering committee and secretary group will collect clinical questions from different levels of clinical medical workers in different levels of hospitals and will compile a preliminary list of clinical questions based on the full texts of existing literature and international guidelines. They will categorize clinical issues, deconstruct PICOS issues, conduct 2 to 3 rounds of expert consensus, and will finally determine the clinical issues to be included. The format of the clinical questions will follow the PICOS principles [25,26]. For example: Is Chinese medicine an effective way to treat dysmenorrhea?

The *P* in PICOS stands for population or participants; here, it refers to all patients with primary dysmenorrhea, endometriosis, or adenomyosis. *I* stands for intervention; here, it refers to patients who use Chinese medicine. *C* stands for comparison; here, it refers to patients who do not use Chinese medicine. *O* stands for outcome; here, it refers to degree of pain and adverse reaction.

On the basis of a preliminary analysis of the published literature and clinical practice, the program will select outcome indicators, including the degree of dysmenorrhea; duration of dysmenorrhea; scores from the visual analogue scale, Cox Menstrual Symptom Scale, Dysmenorrhea Symptom Score, Short-Form McGill Pain Questionnaire, pain numerical rating scale, and Cox Retrospective Symptom Scale; serum prostaglandin F2α/prostaglandin E2 levels; uterine hemodynamic indicators; lesion size; menstrual volume; carbohydrate antigen 125 level; patient fertility; recurrence rate; and adverse reactions. These outcome indicators represent the efficacy of dysmenorrhea treatment. Each outcome indicator will be scored on a 9-point scale, with 1 to 3 points indicating that the outcome indicator is not important, 4 to 6 points indicating that the outcome indicator is important, and 7 to 9 points indicating that the outcome indicator plays a vital role in decision-making [25].

Evidence Retrieval and Synthesis

Databases

We will systematically review the literature from MEDLINE, Web of Science, Embase, Cochrane Library, and 4 Chinese databases (CNKI, Weipu, Wanfang, and China Biomedical Literature Database). The date of publication will be restricted from January 1, 2001, to June 30, 2022. The publications will be limited to those published in Chinese and English.

Search Terms

The following Medical Subject Headings (MeSH) items and free words will be used to balance the search sensitivity and specificity: dysmenorrhea, primary dysmenorrhea, secondary dysmenorrhea, endometriosis, adenomyosis, herbal, Chinese medicine, traditional Chinese medicine, complementary and alternative medicine, acupuncture, and moxibustion, among others. We will first search for MeSH terms in the MEDLINE database to determine subject terms and free words. On this basis, evidence-based medicine experts from Lanzhou University will provide guidance and develop detailed retrieval strategies. Using PubMed to retrieve "primary dysmenorrhea" as an example, the search process is detailed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Search terms.

1. "dysmenorrhea"[MeSH] OR "primary dysmenorrhea"[MeSH]
2. "dysmenorrhea" OR "primary dysmenorrhea" OR "painful menstruations" OR "dysmenorrheas" OR "menstrual pain" OR "painful menstruation" OR "painful period" OR "menstrual pains" OR "menstrual cramp"
3. #3 #1 OR #2
4. "Chinese medicine"[MeSH] OR "traditional Chinese medicine" [MeSH] OR "acupuncture" [MeSH] OR "complementary and alternative medicine" [MeSH] OR "moxibustion" [MeSH] OR "(herbal medicine" [MeSH] OR "TCM" [MeSH]
5. "Chinese medicine" OR "traditional Chinese medicine" OR "acupuncture" OR "complementary and alternative medicine" OR "moxibustion" OR "herbal medicine" OR "TCM" OR "powders" OR "Granules" OR "Decoction" OR "san" OR "Decoction"
6. #4 OR #5
7. "randomized controlled trial" [Publication Type] OR "Meta-Analysis" [Publication Type] OR "Systematic review" [Publication Type] OR "randomized controlled trial" [MeSH] OR "Meta-Analysis" [MeSH] OR "Systematic review"[MeSH]
8. "random" OR "randomly" OR "allocation" OR "blind" OR "concealment" OR "placebo" OR "clinical" OR "RCT" OR "Systematic review" OR "Meta" OR "control" OR "study" OR "trial"
9. #7 OR #8
10. #3 AND #6 AND #9

Literature Selection

Randomized controlled trials, systematic reviews, and other types of clinical studies related to this study will be retrieved. After eliminating duplicate and irrelevant studies, the guide secretary group will screen the included literature based on the title and abstract and will then rescreen it by reading the full text to finally determine the selected literature.

Evidence Syntheses

High-quality systematic reviews published within the previous 2 years will be used directly. In the absence of relevant systematic reviews, if the quality of existing systematic reviews is not high, or the results of the systematic reviews are of low applicability to the issues addressed by the guidelines, we will consider conducting or updating a systematic review [27].

Evidence Assessment

The Grade of Recommendations Assessment Development and Evaluation (GRADE) [28,29] instrument is used for the classification of evidence, including evidence quality classification and recommendation classification. The quality of evidence is divided into very low, low, medium, and high, and the strength of recommendation is divided into 2 levels, weak and strong. Randomized controlled trials are considered high-quality evidence, while observational studies are considered low-quality evidence. Methodologists and clinical experts will jointly adjust the level of evidence based on the evaluation criteria of 5 downgrading factors and 3 upgrading factors to form the final version.

We will evaluate the evidence for the entire study on the basis of different results. The methodologists will be responsible for evaluating the quality of evidence, drafting evidence summaries, and presenting these summaries at the guideline development group meeting. The grading instrument will be used to assess the level of evidence for making recommendations. The evidence evaluation group will develop an evidence summary table, and the guideline development group will divide the relevant evidence for each recommendation into 4 levels based on the scoring instrument: high, medium, low, and very low.

Patients' Preferences and Values

The secretary group will be responsible for investigating the preferences and values of patients with dysmenorrhea in using TCM treatment. The opinions of patients will be fed back to the steering committee and consensus expert group for consideration when formulating recommendations.

Developing Recommendations

After completing the evidence classification according to the GRADE system, the preliminary recommendations will be drafted with reference to the paper titled "Criterion and Detailed Judgments for the Transformation From Evidence to Recommendations in the Clinical Practice Guidelines of Chinese Medicine" by Yang et al [30]. Expert consensus will be reached through 2 to 4 rounds of the Delphi method to formulate draft recommendations, which will then be submitted to the steering committee for final approval [31,32]. The GRADE grid structured approach will be used to reach consensus [33]. Each item in the questionnaire will have 5 levels of choice: strongly or weakly recommending for it, making no clear recommendation, and strongly or weakly recommending against it. For each item, if more than 50% of the experts vote for any option other than "no clear recommendation," or if more than 70% of the experts vote for 1 of the 2 options on the same side, consensus will be deemed to have been reached. If this does not occur, the item will be deemed to have not reached a consensus, and the recommendations will enter the next round of the Delphi process.

Drafting the Guidelines

The guidelines will be drafted based on the level of evidence and strength of each recommendation. The reporting specifications of this guideline will refer to the Reporting Items for Practice Guidelines in Healthcare (RIGHT) for TCM [34] during its development. The main text of the guideline will be

XSL•FO
RenderX

written at the same time as the corresponding preparation instructions.

Peer Review of Guidelines

The draft guidelines will be submitted to the external review group for review and confirmation. The guideline working group will record the review process and collect opinions from the review experts. After fully discussing these opinions, the draft guidelines will be revised appropriately.

Publication and Promotion of Guidelines

The guidelines are expected to be published within 1 or 2 years and will be translated into Chinese and English and published in relevant journals. After the guideline is released, it will be promoted in the following ways: (1) the guidelines will be printed as a manual and disseminated in the academic community; (2) the guidelines will be used in lectures at academic conferences; and (3) the Chinese, English, or other language versions of the guidelines will be promoted on popular medical or official health websites or apps.

Implementation, Evaluation, and Updating of the Guidelines

The guideline working group will regularly monitor the implementation of the guidelines. This will include process evaluation, implementation effect, performance evaluation, and evaluation of the guideline itself. When there is new relevant evidence or recommendations, or if the strength of recommendations changes after the release of the guideline, the guideline will be updated in a timely manner. The update cycle of the general guidelines will be between 2 and 5 years.

Ethical Considerations

This study protocol pertains solely to the development of a clinical practice guideline based on a literature review, expert consensus, and analysis of existing, fully anonymized, publicly available, or aggregated statistical data. There is no direct involvement of human subjects, no primary data collection, and no experimental intervention. Expert consultation within this process is for professional opinion formation and does not constitute human subjects research as defined by our institutional policies. Given the nature of the study design, the type of data used, and the minimal risk involved, the study was reviewed and classified as exempt from full ethics review by the Hebei University of Chinese Medicine Science and Technology Ethics Committee.

Results

This work was supported by the National Key R&D Program of China (grant 2019YFC1712000) in 2019. The writing group for guidelines was formed in January 2021. The literature search and screening process began in May 2022. An initial search conducted in May 2022 across 8 databases (including 4 Chinese databases and 4 English databases) identified 54,029 studies that met the search criteria. Of the 54,029 articles, 51,066 (94.5%) duplicates were removed, leaving 2963 (5.5%) articles for screening. Before screening, an additional search will be conducted to update the identified articles and ensure they are the most current. Two independent reviewers finished the literature search and screening process. After completing the evidence classification according to the GRADE system, the preliminary recommendations were drafted with reference to the paper by Yang et al [30]. Expert consensus was reached through 2 to 4 rounds of the Delphi method to formulate draft recommendations. After 2 rounds of Delphi method expert questionnaire surveys, a consensus among experts was reached, and 41 and 46 feedback responses were collected in the 2 rounds.

Discussion

Anticipated Findings

The establishment of the first international clinical practice guidelines on TCM therapy for dysmenorrhea will provide a basis for such therapy. The program aims to adopt an appropriate approach to the development of the guideline. The guidelines will be developed in accordance with this protocol, which may provide support and evidence for TCM treatment for patients with dysmenorrhea. Chinese herbal medicine, acupuncture, massage, acupoint stimulation, and other TCM methods used solely or in combination for the treatment of dysmenorrhea have demonstrated considerable therapeutic benefits. TCM gives good clinical results for the treatment of dysmenorrhea based on the basic principles of holism and syndrome differentiation.

Dysmenorrhea is a common gynecological condition that is further divided into 2 categories, namely, primary and secondary dysmenorrhea. The guideline will clarify the definition, diagnosis, and TCM syndrome classification for dysmenorrhea and recommend different TCM treatments based on the level of evidence, including Chinese herbal medicine, Chinese patent medicine, acupuncture, and other auxiliary therapies. This guideline will provide substantive evidence and standardized guidance for the treatment of dysmenorrhea using TCM, improve the quality and safety of medical services, and standardize diagnosis and treatment plans for dysmenorrhea.

Strengths and Limitations

This guideline was developed with the involvement and guidance of methodology experts to ensure methodological rigor and transparency. Data screening and extraction were carried out by 2 independent reviewers, with a third reviewer resolving any discrepancies, thereby minimizing selection and extraction bias. However, some limitations can be anticipated. The exclusion of studies in languages other than Chinese and English may reduce the comprehensiveness of the evidence base due to translation limitations. In addition, including nonrandomized studies may introduce bias. Despite these limitations, this guideline will provide substantial evidence and standardized guidance for the treatment of dysmenorrhea with TCM, improve the quality and safety of medical services, and standardize the diagnosis and treatment of dysmenorrhea.

Future Directions

In the future, we will continue to pay attention to research progress on the treatment of dysmenorrhea with TCM, continuously update the guidelines, and provide guidance for clinical TCM treatment of dysmenorrhea.

XSL•FO
RenderX

Conclusions

This plan proposes the specific steps for developing guidelines, aiming to establish international clinical practice guidelines for the treatment of dysmenorrhea with TCM and providing guidance for clinical practice. The findings will be disseminated through professional clinical and academic networks, ensuring that research findings are translated into practice and guiding future research directions.

Acknowledgments

The authors would like to thank all the reviewers who participated in the review, as well as MJEditor for providing English editing services during the preparation of this manuscript.

Funding

This work was supported by the National Key R&D Program of China (grant 2019YFC1712000).

Conflicts of Interest

None declared.

References

1. Sharghi M, Mansurkhani SM, Larky DA, et al. An update and systematic review on the treatment of primary dysmenorrhea. JBRA Assist Reprod 2019 Jan 31;23(1):51-57. [doi: 10.5935/1518-0557.20180083] [Medline: 30521155]
2. Pellow J, Nienhuis C. Medicinal plants for primary dysmenorrhoea: a systematic review. Complement Ther Med 2018 Apr;37:13-26. [doi: 10.1016/j.ctim.2018.01.001] [Medline: 29609924]
3. Ferries-Rowe E, Corey E, Archer JS. Primary dysmenorrhea: diagnosis and therapy. Obstet Gynecol 2020 Nov;136(5):1047-1058. [doi: 10.1097/AOG.0000000000004096] [Medline: 33030880]
4. Iacovides S, Avidon I, Baker FC. What we know about primary dysmenorrhea today: a critical review. Hum Reprod Update 2015;21(6):762-778. [doi: 10.1093/humupd/dmv039] [Medline: 26346058]
5. Latthe PM, Champaneria R. Dysmenorrhoea. BMJ Clin Evid 2014 Oct 21;2014:183. [Medline: 25338194]
6. Proctor ML, Hing W, Johnson TC, Murphy PA. Spinal manipulation for primary and secondary dysmenorrhoea. Cochrane Database Syst Rev 2006 Jul 19;2006(3):CD002119. [doi: 10.1002/14651858.CD002119.pub3] [Medline: 16855988]
7. Dawood MY. Primary dysmenorrhea: advances in pathogenesis and management. Obstet Gynecol 2006 Aug;108(2):428-441. [doi: 10.1097/01.AOG.0000230214.26638.0c] [Medline: 16880317]
8. Bettendorf B, Shay S, Tu F. Dysmenorrhea: contemporary perspectives. Obstet Gynecol Surv 2008 Sep;63(9):597-603. [doi: 10.1097/OGX.0b013e31817f15ff] [Medline: 18713479]
9. Morrow C, Naumburg EH. Dysmenorrhea. Prim Care 2009 Mar;36(1):19-32. [doi: 10.1016/j.pop.2008.10.004] [Medline: 19231600]
10. Ju H, Jones M, Mishra G. The prevalence and risk factors of dysmenorrhea. Epidemiol Rev 2014;36(1):104-113. [doi: 10.1093/epirev/mxt009] [Medline: 24284871]
11. Oladosu FA, Tu FF, Hellman KM. Nonsteroidal antiinflammatory drug resistance in dysmenorrhea: epidemiology, causes, and treatment. Am J Obstet Gynecol 2018 Apr;218(4):390-400. [doi: 10.1016/j.ajog.2017.08.108] [Medline: 28888592]
12. Sachedina A, Todd N. Dysmenorrhea, endometriosis and chronic pelvic pain in adolescents. J Clin Res Pediatr Endocrinol 2020 Feb 6;12(Suppl 1):7-17. [doi: 10.4274/jcrpe.galenos.2019.2019.S0217] [Medline: 32041388]
13. Office of Standardization of Traditional Chinese Medicine of the State Administration of Traditional Chinese Medicine. General Rules of Preparation of Diagnosis and Treatment Guideline in Traditional Chinese Medicine: China Press of Traditional Chinese Medicine; 2015.
14. Lu CJ, Yang KH. Manual for the Development of Integrated Traditional Chinese and Western Medicine Diagnosis and Treatment Guidelines: People's Medical Publishing House; 2016.
15. Graham R, Mancher M, Miller Wolman D, Greenfield S, Steinberg E. Clinical Practice Guidelines We Can Trust: National Academies Press (US); 2011. [Medline: 24983061]
16. Rosenfeld RM, Shiffman RN, Robertson P, Department of Otolaryngology State University of New York Downstate. Clinical Practice Guideline Development Manual, Third Edition: a quality-driven approach for translating evidence into action. Otolaryngol Head Neck Surg 2013 Jan;148(1 Suppl):S1-55. [doi: 10.1177/0194599812467004] [Medline: 23243141]
17. Chen YL, Liu X, Zhang YQ, et al. Western medicine guideline and traditional Chinese medicine guideline: improving together in mutual learning. Med J PUMCH 2020;11(5):615-620. [doi: 10.3969/j.issn.1674-9081.2020.05.019]
18. AGREE Collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 2003 Feb;12(1):18-23. [doi: 10.1136/qhc.12.1.18] [Medline: 12571340]
19. Xie LM, Wang WY. A brief introduction to appraisal of guidelines for research and evaluation II. Zhong Xi Yi Jie He Xue Bao 2012 Feb;10(2):160-165. [doi: 10.3736/jcim20120206] [Medline: 22313883]

XSL•FO
RenderX

20. Wang SC, Zhao X, Yu S, et al. Quality evaluation of evidence-based TCM clinical diagnosis and treatment guidelines--AGREE II tool and its application. Chin J Tradit Chin Med Pharm 2016;31(8):2963-2967 [FREE Full text]
21. WHO Handbook for Guideline Development: World Health Organization; 2014. URL: https://www.who.int/publications/i/item/9789241548960 [accessed 2026-03-02]
22. Scottish Intercollegiate Guidelines Network (SIGN). SIGN 50: A Guideline Developer's Handbook, 3rd edition: SIGN Publications. URL: https://www.sign.ac.uk/guidelines/sign-50-a-guideline-developers-handbook/ [accessed 2026-03-02]
23. Developing NICE Guidelines: The Manual: National Institute for Health and Care Excellence; 2015. URL: https://www.ncbi.nlm.nih.gov/books/NBK310375/pdf/Bookshelf_NBK310375.pdf [accessed 2026-03-02]
24. SIGN 100: A Handbook for Patient and Carer Representatives: Scottish Intercollegiate Guidelines Network (SIGN); 2019. URL: https://guidelines.ebmportal.com/handbook-patients-and-carer-representatives-sign-cpg-100 [accessed 2025-05-25]
25. Chen W, Guo BY, Fang SN, et al. Determination of clinical problems in clinical practice guidelines of traditional Chinese medicine. Chin J Integr Tradit Wes Med 2020;40(2):220-223.
26. Li H, Chen YL, Wei D, et al. Development and revision methods of clinical practice guideline (CGP) on traditional Chinese medicine (integrated medicine): Clinical issues construction. Chin J Tradit Chin Med Pharm 2016;31(6):2202-2205 [FREE Full text]
27. Liu Q, Chen T, Zhao G, Liu J. Clinical practice guidelines for traditional Chinese medicine treatment of influenza (2021). J Tradit Chin Med 2022;63(1):85-98. [doi: 10.13288/j.11-2166/r.2022.01.016]
28. Atkins D, Best D, Briss PA, et al. Grading quality of evidence and strength of recommendations. BMJ 2004 Jun 19;328(7454):1490. [doi: 10.1136/bmj.328.7454.1490] [Medline: 15205295]
29. Chen YL. The Application of GRADE in Systematic Review and Practice Guidelines: China Peking Union Medical University Press; 2021. URL: https://book.douban.com/subject/35445154/ [accessed 2025-05-25]
30. Yang SH, Chen KJ, Liu JP, Chen W. Criterion and detailed judgments for the transformation from evidence to recommendations in the clinical practice guidelines of Chinese medicine. Chin J Integr Tradit West Med 2021;41(4):494-498 [FREE Full text]
31. Li JC. Guidelines for chinese medicine rehabilitation of chronic obstructive pulmonary disease. World J Tradit Chin Med 2020;6(3):295-306. [doi: 10.4103/wjtcm.wjtcm_44_20B2]
32. Liao X, Hu J, Xie YM, et al. A study on exploring consensus developing standards in clinical practice guideline for traditional Chinese medicine. Chin J Chin Materi Med 2017;42(8):1518-1524.
33. Jaeschke R, Guyatt GH, Dellinger P, et al. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. BMJ 2008 Jul 31;337:a744. [doi: 10.1136/bmj.a744] [Medline: 18669566]
34. Xie R, Xia Y, Chen Y, et al. The RIGHT extension statement for traditional Chinese medicine: development, recommendations, and explanation. Pharmacol Res 2020 Oct;160:105178. [doi: 10.1016/j.phrs.2020.105178] [Medline: 32889127]

Abbreviations

GRADE: Grade of Recommendations Assessment Development and Evaluation
MeSH: Medical Subject Headings
PICOS: participants, interventions, comparisons, outcomes, and study design
RIGHT: Reporting Items for Practice Guidelines in Healthcare
TCM: traditional Chinese medicine

Edited by J Sarvestan; submitted 13.May.2025; peer-reviewed by Z Lin; revised version received 01.Feb.2026; accepted 02.Feb.2026; published 11.Mar.2026.

Please cite as:
Zhang L, Liang X, Chen J, Du Y, Du H, Wang X
International Clinical Practice Guidelines on Traditional Chinese Medicine Therapy for Dysmenorrhea: Development Protocol
JMIR Res Protoc 2026;15:e77423
URL: https://www.researchprotocols.org/2026/1/e77423
doi:10.2196/77423

© Li Zhang, Xiao Liang, Junlu Chen, Yuhan Du, Huilan Du, Xiaohui Wang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 11.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Noninvasive Measurement of Cerebrospinal Fluid Flow in Shunted Hydrocephalus: Protocol for Scanner Calibration and Multisite Data Collection

Jason M Toliao[1]; Matthew T Borzage[2,3,4], PhD; Pradip P Chaudhari[5,6], MS, MD; Joseph H Ha[7,8], BA; Isabella Friedli[1]; Madison E Gutierrez[7], BS; Meghan Drastal[9], RN, MSN, MBA; Peter Calabrese[10], PhD; Eamon K Doyle[11], PhD; Isabel Torres[7,11], BS; Jacob Al-Husseini[6], BS; Clio González-Zacarías[7,11], PhD; John C Wood[3,6,11,12], MD, PhD; Jason Hauptman[13], MD, PhD; J Gordon McComb[6,7], MD; Stefan Blüml[11], PhD; Peter A Chiarelli[6,7,13], MD, DPhil

[1]University of Southern California, Los Angeles, CA, United States

[10]Department of Quantitative and Computational Biology, University of Southern California, Los Angeles, CA, United States

[11]Department of Radiology, Children's Hospital of Los Angeles, Los Angeles, CA, United States

[12]Division of Cardiology, Children's Hospital of Los Angeles, Los Angeles, CA, United States

[13]Division of Neurosurgery, Phoenix Children’s Hospital, 1919 E Thomas Rd, Phoenix, AZ, United States

[2]Fetal and Neonatal Institute, Division of Neonatology, Children's Hospital of Los Angeles, Los Angeles, CA, United States

[3]Alfred E Mann Department of Biomedical Engineering, Viterbi School of Engineering, University of Southern California, Los Angeles, CA, United States

[4]Department of Regulatory and Quality Sciences, Alfred E. Mann School of Pharmacy and Pharmaceutical Sciences, University of Southern California, Los Angeles, CA, United States

[5]Division of Emergency and Transport Medicine, Children's Hospital Los Angeles, Los Angeles, CA, United States

[6]Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

[7]Division of Neurosurgery, Department of Surgery, Children's Hospital of Los Angeles, Los Angeles, CA, United States

[8]Stanford School of Medicine, Palo Alto, CA, United States

[9]Department of Nursing Administration, Children's Hospital Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Peter A Chiarelli, MD, DPhil
Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

Abstract

Background: Ventricular shunts divert cerebrospinal fluid (CSF) in patients with hydrocephalus, which can be lifesaving. Untreated shunt failure may lead to increased intracranial pressure and neurological injury. The process of diagnosing shunt malfunction can be complex, and historically, there has not been a simple method for noninvasive and quantitative measurement of CSF flow through shunts. The demonstration of successful clinical application of phase-contrast magnetic resonance imaging (PC-MRI) to noninvasively quantify shunt flow is relatively new and will benefit from standardization across varying types of magnetic resonance (MR) hardware to facilitate implementation at multiple medical sites. PC-MRI CSF flow measurement through ventricular shunts has not yet been compared across different types of MR hardware (ie, differing field strengths, radiofrequency coils, slew rates, gradient strengths, models, and manufacturers).

Objective: This study describes a protocol aimed to standardize and optimize PC-MRI shunt flow measurements for widespread use on any MR scanner.

Methods: To study shunt flow in a manner that would translate effectively to the patient care setting, a phantom model incorporating a human shunt catheter can be constructed to obtain data over a typical physiological range of CSF flow rates. Calibration curves are used to model data comparing known flow rates to flow measured via PC-MRI. The accuracy of each MR scanner is assessed using linear regression. This protocol will be repeated on 8 MR scanners before multisite data collection.

Results: The shunt flow phantom was constructed in November 2024. Single-site phantom calibration began in April 2025 and concluded in December 2025. Registration of clinical data collection sites will take place between November 2025 and July 2026. Approval from the Children’s Hospital of Los Angeles Institutional Review Board was obtained in January 2023 (before scanner calibration), and single-site data collection is ongoing. Multi-institutional clinical data collection will begin in July 2026 and continue until January 2027. Results and statistical analyses are expected by April 2027.

Conclusions: This study protocol provides a methodology to test and implement PC-MRI on any MR scanner using a phantom model that (1) represents real flow and catheter conditions, (2) provides a repeatable means for data collection, (3) is easy to

XSL•FO
RenderX

assemble, (4) minimizes artifacts, and (5) is transportable. We also describe analytic methods for interscanner calibration across a range of hardware parameters and provide a framework for multisite data collection.

International Registered Report Identifier (IRRID): DERR1-10.2196/85918

(JMIR Res Protoc 2026;15:e85918) doi:10.2196/85918

KEYWORDS

cerebrospinal fluid shunt; phase contrast; magnetic resonance imaging; hydrocephalus; shunt malfunction; study protocol; method; cerebrospinal fluid

Introduction

Over 380,000 patients worldwide are diagnosed annually with hydrocephalus [1], a pathological imbalance between cerebrospinal fluid (CSF) production and absorption [2], often causing markedly increased intracranial pressure [3,4]. If left untreated, elevated intracranial pressure can lead to severe neurological damage, disability, or death [3-5]. Placement of a ventricular shunt to treat hydrocephalus is a frequently performed neurosurgical procedure. Ventricular shunts redirect CSF to a different site in the body for absorption [3,4,6]. Mechanical obstruction or migration of shunt hardware leads to failure in nearly 30% of patients within 6 months of shunt placement [7], and more than 38,000 hospitalizations occur annually related to CSF shunt systems [8].

Evaluation of shunt function in clinical practice can include the following approaches applied independently or in combination: review of presenting signs and symptoms [3,5,9], structural imaging of ventricular size [3,10-12], x-ray imaging of hardware integrity [3,11,13,14], percutaneous shunt tap [15], radionuclide shunt flow study [16-18], or continued inpatient monitoring for evolution or resolution of symptoms [5,8]. The ability to reliably detect shunt malfunction remains imperfect. Approximately a quarter of patients with shunt failure exhibit no change in ventricular size [16], rendering accurate detection of malfunctioning shunts a composite assessment and highlighting the importance of clinical experience [5,19,20]. There is potential to improve modern clinical practice and provide additional rigor to the determination of shunt function or malfunction by measuring shunt flow noninvasively through magnetic resonance imaging (MRI) [21].

Phase-contrast magnetic resonance imaging (PC-MRI) was recently demonstrated [21-23] as a fast and reliable scan sequence to measure bulk movement of CSF in ventricular shunts. The PC-MRI technique for shunt flow measurement has been validated for multiple patients at a single institution [21], with active plans to expand use of the technology through multi-institutional implementation and data assimilation. Adoption of the scan sequence at different hospitals and on different magnetic resonance (MR) scanners requires thorough confirmation that shunt flow measurements are of comparable accuracy and reliability at these sites. Results are influenced by software (scanning features and parameters), hardware (field strength, radiofrequency coil design, gradient strength, gradient slew rate, manufacturer, and model), or both. A systematic framework to investigate these factors will assist in guiding expectations for the success of flow measurement on different scanners and will provide a potentially necessary means to account for variable bias in flow measurement if this effect were to occur on a given scanner or hardware combination. Scanner calibration and quantification of expected error will enhance the reliability of data acquired using different types of MR hardware during multi-institutional data collection and standardize PC-MRI protocols for implementation on any MR scanner.

This work presents (1) a consolidated description of the hardware variables that must be accounted for on new scanners; (2) a method for any institution to implement a simple phantom design to measure shunt flow; (3) a protocol for flow rate testing via PC-MRI on different MR scanners; (4) an appropriate suite of basic and more stringent analytic methods to calculate flow; (5) a statistical approach to determine potential hardware-induced bias, as well as to correct for bias potentially introduced on different scanners; and (6) plans to validate PC-MRI for multisite clinical data collection. By using a homogeneous phantom medium with the same hardware used in human shunts, we ensure consistent measurements that can effectively model in vivo shunt flow. The combination of a linear single-catheter phantom with a screw-driven pump mechanism allows for precise and reliable delivery of preset flows across a specific spectrum of physiological flow rates. PC-MRI flow measurements can be compared with predetermined flow rates to obtain a calibration curve for individual scanners and their particular hardware. After a thorough investigation of scanner-related variability is performed, we will collect clinical data across multiple institutions to evaluate the feasibility of implementing standardized PC-MRI data collection platforms.

Methods

Overview

We constructed a phantom for data acquisition and developed a standardized method to calibrate PC-MRI for accurate shunt flow assessment on any MR scanner. Our protocol applies phantom calibration to evaluate variation in PC-MRI measurements introduced by different MR hardware and validates multisite PC-MRI data collection. Rigorous methods for statistical comparison between types of scanner hardware are also provided.

Apparatus for Data Collection

We constructed and evaluated a versatile suite of phantoms (Multimedia Appendix 1). A single-tube, linear phantom was used for data collection based on its simple design and reproducible flow production. The single-tube phantom was

XSL•FO
RenderX

constructed using an antibiotic-impregnated catheter with a 1.3-mm inner diameter (Ares model; Medtronic), a common surgically placed distal shunt catheter in patients at our institution. The catheter was secured longitudinally through a rectangular 500-mL biologic media flask filled with water to eliminate air pockets and optimize the magnetic homogeneity of the imaged region.

An Alaris screw-driven syringe pump (Becton, Dickinson and Company) provided modifiable flow rates to drive water through the phantom in a continuous fashion without peristaltic or pulsatile action. The pump was connected to one end of the phantom using small-bore Luer lock tubing (ICU Medical). Materials and a step-by-step procedure for phantom construction are provided in Multimedia Appendix 1.

We calibrated the combination of flow phantom and pump before data collection to ensure accurate and precise volumetric fluid delivery during subsequent testing (Figure 1). An analytical balance (AG204 DeltaRange; Mettler Toledo) was used to collect fluid output from the shunt system at 10 predetermined rates spanning the expected physiological range of CSF flow (0-22 mL per hour). The catheter was oriented flat on the benchtop to allow the liquid to drip freely, avoiding potential siphoning effects from gravitational pressure differences. Each pump setting was tested 3 times against the corresponding volumetric fluid output to evaluate the reproducibility of the pump's flow.

Figure 1. A schematic of the setup for calibration of the pump mechanism and single-tube phantom on an analytical balance.

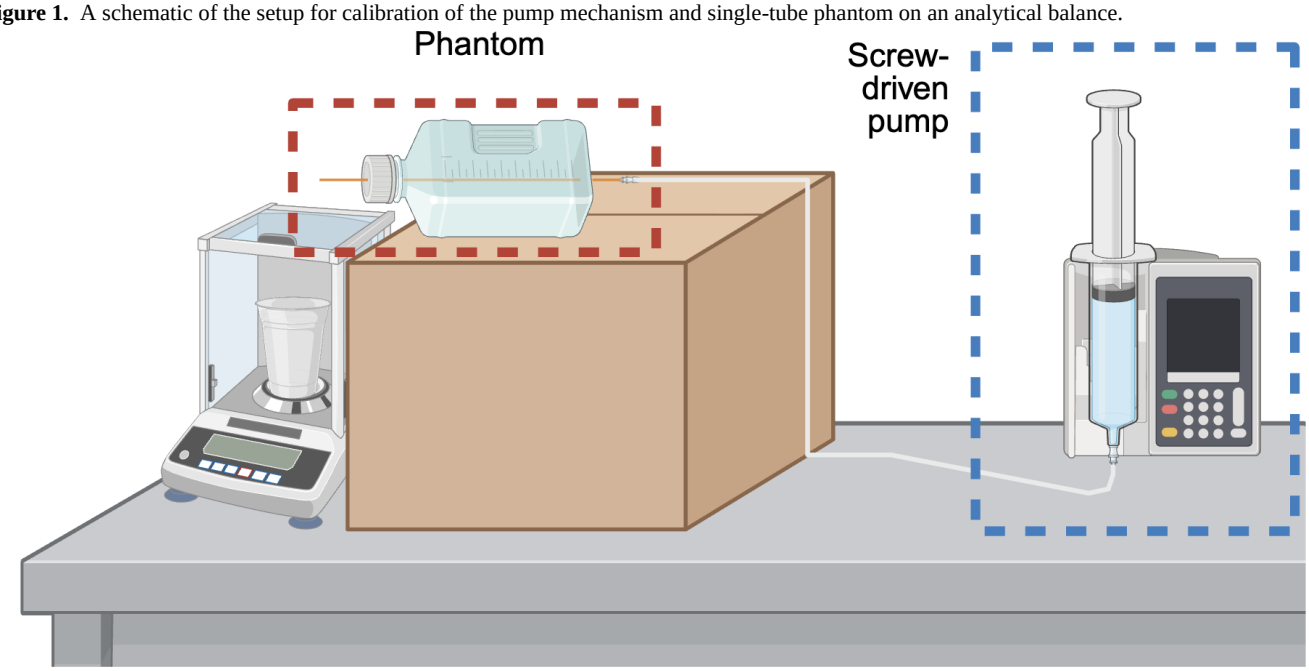

Protocol for PC-MRI Image Acquisition

PC-MRI flow acquisition is suggested with the following base sequence parameters: field of view of 30 × 30 mm, slice thickness of 10 mm, 2 signal averages, matrix size of 92 × 91, acquisition voxel dimensions of 0.20 × 0.20 mm^2, reconstructed voxel dimensions of 0.10 × 0.10 mm^2, compressed sensitivity encoding factor of 2, and velocity encoding of 1 cm per second [21].

During image acquisition, the phantom is advised to be positioned parallel to the length of scanner bore and level with the MR bed. The screw-driven pump, located outside the MR room, is attached to the phantom, and connection tubing is carefully positioned to remain free of kinks or air (Figure 2). Triplanar (coronal, axial, and sagittal) T2-weighted images are used for PC-MRI slice localization, and PC-MRI image planes are placed orthogonally (transverse) to the direction of catheter flow such that the catheter lumen appears circular on phase-contrast images. The screw-driven pump is used to establish flow through the phantom at 10 benchtop-tested rates (0, 1, 3, 5, 8, 11, 14, 17, 19, and 22 mL/h) chosen for their overlap with the typical physiological range of CSF production. Four repeat PC-MRI images are obtained at each flow rate. Images are saved as Digital Imaging and Communications in Medicine files and exported to an external processing platform for flow measurement in a typical fashion.

To facilitate multi-institutional implementation, this acquisition and processing protocol is designed to be reproducible across centers with varying MR hardware configurations. Home institution plans for multiscanner calibration include repeating this procedure on 8 MR scanners in which field strength, gradient slew rate, gradient strength, radiofrequency coil type, model, and manufacturer vary (Table 1). Sites participating in multicenter data collection may apply this protocol to calibrate scanners for use in PC-MRI data acquisition.

XSL•FO
RenderX

Figure 2. Scanner setup for phantom measurement: (A) schematic and magnitude image of the catheter lumen obtained on the single-tube phantom and (B) image plane through the phantom. PC-MRI: phase-contrast magnetic resonance imaging.

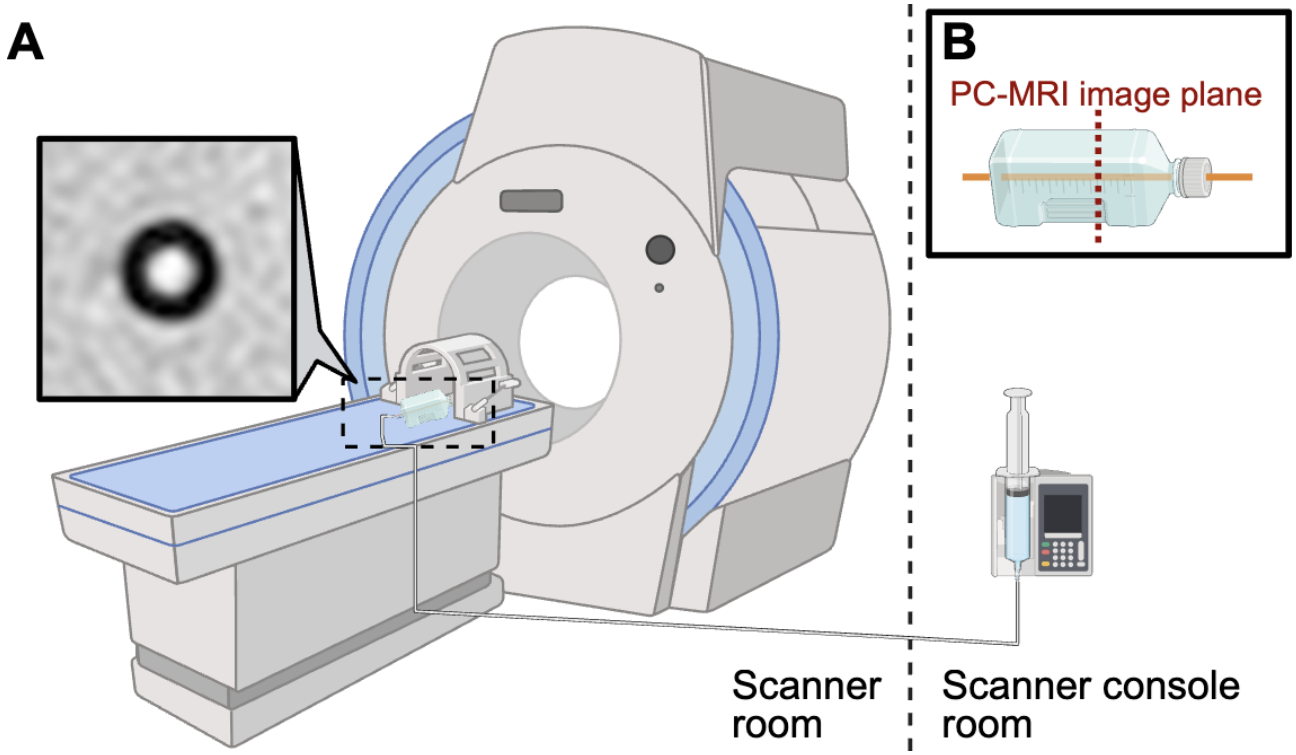

XSL•FO
RenderX

Table . Variables that may affect phase-contrast magnetic resonance imaging (PC-MRI) measured flow rate.

Variable	Example parameters
Biological variation	
Flow	• 0-22 mL/h
Catheter type	• Barium impregnated (Codman model [Integra LifeSciences] and Ares model [Medtronic]) • Antibiotic-impregnated (Codman model and Ares model)
Image quality[a]	
Vendor or model	• Siemens and Phillips
Field strength	• 0.55-7 T
RF[b] coil type	• 8, 16, and 32 channels
Gradient slew rate	• Covaries with field strength
Gradient strength	• 20-200 mT/m
TR[c] and TE[d]	• TR: 38-135 ms • TE: 25-50 ms
Echo train length	• 0-3
Echo numbers	• 1
Slice thickness	• 5-10 mm
Number of averages	• 2-8
Acquisition matrix	• Optimized for each manufacturer and system
Acquisition voxel dimensions	• Optimized for each manufacturer and system
Bandwidth	• 24-230 Hz
Flip angle	• 5-45°
Acceleration factor	• CSENSE[e] or GRAPPA[f]
Scan duration	• 48-180 s
Acquisition[a]	
Angle	• 0-25°
VENC[g]	• –1 to 1 cm/s
Shim location	• Default shim • One-quarter FOV[h] shim offset • Three-quarters FOV shim offset
Image processing	
ROI[i] size	• <1.3-mm inner diameter • 1.3-mm inner diameter • >1.3-mm inner diameter
ROI location	• 6 individuals will independently prescribe ROIs

XSL•FO
RenderX

Variable	Example parameters
Analytical method	• Mean value • Maximum value • Parabolic model

[a]Image quality and acquisition parameters that can affect contrast-to-noise ratio on the PC-MRI image.
[b]RF: radiofrequency.
[c]TR: repetition time.
[d]TE: echo time.
[e]CSENSE: compressed sensitivity encoding.
[f]GRAPPA: generalized autocalibrating partially parallel acquisitions.
[g]VENC: velocity encoding gradient.
[h]FOV: field of view.
[i]ROI: region of interest.

Analytic Methods of Flow Calculation

We established three analytical approaches to measure volumetric flow from each PC-MRI image: (1) mean intensity within a circular region of interest prescribed over the catheter lumen; (2) maximum intensity within a small circular region of interest at the center of the catheter lumen, thereafter converted via a formula to total flow; and (3) full parabolic model of a laminar flow profile within the lumen (Figure 3). A description of each method of analysis and equations for flow calculation are provided in Multimedia Appendix 2.

Figure 3. Methods for flow estimation: (A) mean value within a circular region of interest, (B) maximum value at the center of a parabolic flow profile, and (C) full parabolic fit of a laminar flow profile within the lumen.

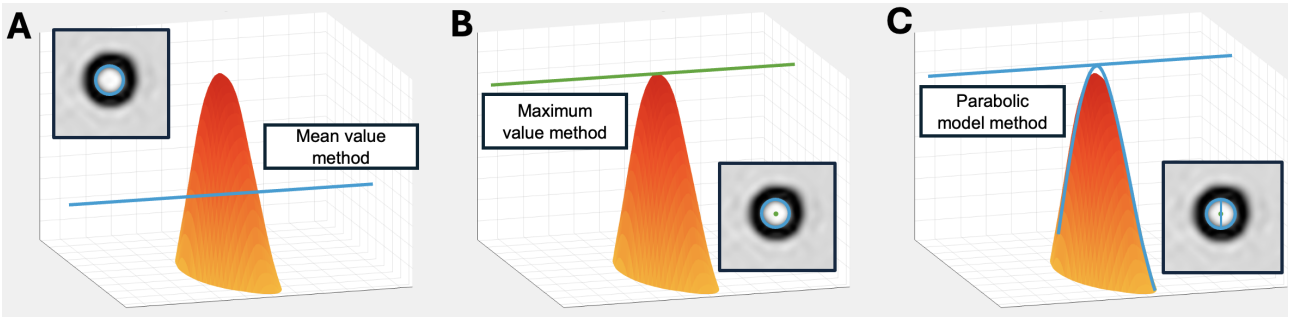

Regression Modeling and Statistical Comparison

For each scanner and hardware combination, linear regression will be used to evaluate the relationship between PC-MRI–derived flow rates and true flow outputs from the pump. An ANOVA will compare mean error between datasets, with a significance level set at α=.05. These statistical approaches will be repeated to evaluate the precision and accuracy of each analytical technique for flow measurement. Descriptive statistics will also be used to examine the effect of variables such as gradient strength, field strength, and slew rate on calibration measurements as these variables often covary with MR scanner.

Multisite Data Collection: Design and Setting

We aim to validate the feasibility of implementing multisite PC-MRI data collection platforms by evaluating PC-MRI image quality and logistical factors such as scheduling feasibility, personnel training requirements, and processes for data transfer. Image quality will be reported as a proportion of analyzable scans within the dataset. Logistical factors surrounding barriers to personnel training or PC-MRI integration into scan protocols will be described individually by each site.

Patients who undergo a PC-MRI flow study will be included in the cohort regardless of age, etiology upon presentation, clinical setting (emergency department, inpatient, or outpatient), or MR hardware used during image acquisition. Patients with distal slit valves, valveless shunts, cystoperitoneal shunts, subdural shunts, and shunts with non–MR-resistant valves will be excluded. To maintain independence among data points, only the first PC-MRI flow measurement for each patient will be included in the analysis of shunt flow over a population of patients. Individuals with repeat PC-MRI measurements during unrelated clinical encounters will be analyzed separately to evaluate longitudinal trends in CSF flow.

Data Collection and Management Plan

Clinical PC-MRI images will be collected across sites following standardized acquisition and processing protocols. A single PC-MRI sequence will be incorporated into routine and emergent full-length structural brain MRI exams for shunted patients at participating institutions. At each site, 2 independent data abstractors will manually review patient charts and clinical records according to predetermined standard operating procedures. Demographic, clinical, and imaging data (Multimedia Appendix 3) will be extracted from the electronic health records and collected using the REDCap (Research Electronic Data Capture; Vanderbilt University) tool [24].

XSL•FO
RenderX

PC-MRI images in which flow data cannot be measured (eg, low resolution, improper shim, and nonorthogonal image plane prescription) will be reported as "missing" and accompanied by documentation of image quality issues. Each institution will use local REDCap collection forms, and periodic queries for data validation will take place at 4-month intervals. For analysis, data will be merged between institutions via secure file transfer with Health Insurance Portability and Accountability Act (HIPAA) compliance. Datasets will be stored for 10 years after the study.

Ethical Considerations

Home institution plans for PC-MRI imaging and data collection have been approved by the Children's Hospital Los Angeles Institutional Review Board (IRB; CHLA-20-00041), and external sites will rely on local IRB approval along with data sharing agreements. As data for this study will be acquired retrospectively, a waiver of consent will be obtained; patient recruitment and informed consent before imaging will not be necessary. Identifying information will be correlated with unique study identification numbers stored on a local encrypted server.

Safety Considerations

MRI is generally considered a safe imaging technique, with primary concerns related to implantable devices and projectile risks from ferromagnetic objects [25,26]. All patients will undergo standard MRI safety screening before imaging.

Sample Size Estimates

To evaluate the feasibility of implementing a multisite PC-MRI data collection platform, we plan to enroll 100 participants across 3 sites. This target is based on our preliminary work and is expected to be feasible based on site selection and the number of eligible patients over the study period. A formal sample size calculation was not performed as the planned multisite study is focused on feasibility rather than hypothesis testing or estimation of effect sizes. The sample size is intended to ensure adequate experience with clinical data collection and will allow us to identify challenges related to PC-MRI acquisition, image quality, and flow measurement.

Results

The shunt flow phantom was constructed in November 2024. Single-site phantom calibration began in April 2025 and concluded in December 2025. Registration of clinical data collection sites will take place between November 2025 and July 2026. Home institution IRB approval was obtained in January 2023 (before scanner calibration), and single-site data collection is ongoing. Multi-institutional clinical data collection will begin in July 2026 and continue until January 2027. Results and statistical analyses are expected by April 2027.

Figure 4 shows the results of hardware calibration of the screw-driven syringe pump across appropriate flow values. Confirmation of pump accuracy is important before use in the MR scanner environment to establish correct baseline flow values and account for potential temporal drift in real pump output. The coefficient of determination (R^2) between true and expected flow output from the pump in our case was 0.9997. Because the slope of the regression was close to 1 and the intercept was near 0, no calibration correction was applied.

Figure 4. Syringe pump output setting (mL per hour; x-axis) vs true flow rate (mL per hour) calibrated on the analytical balance (y-axis).

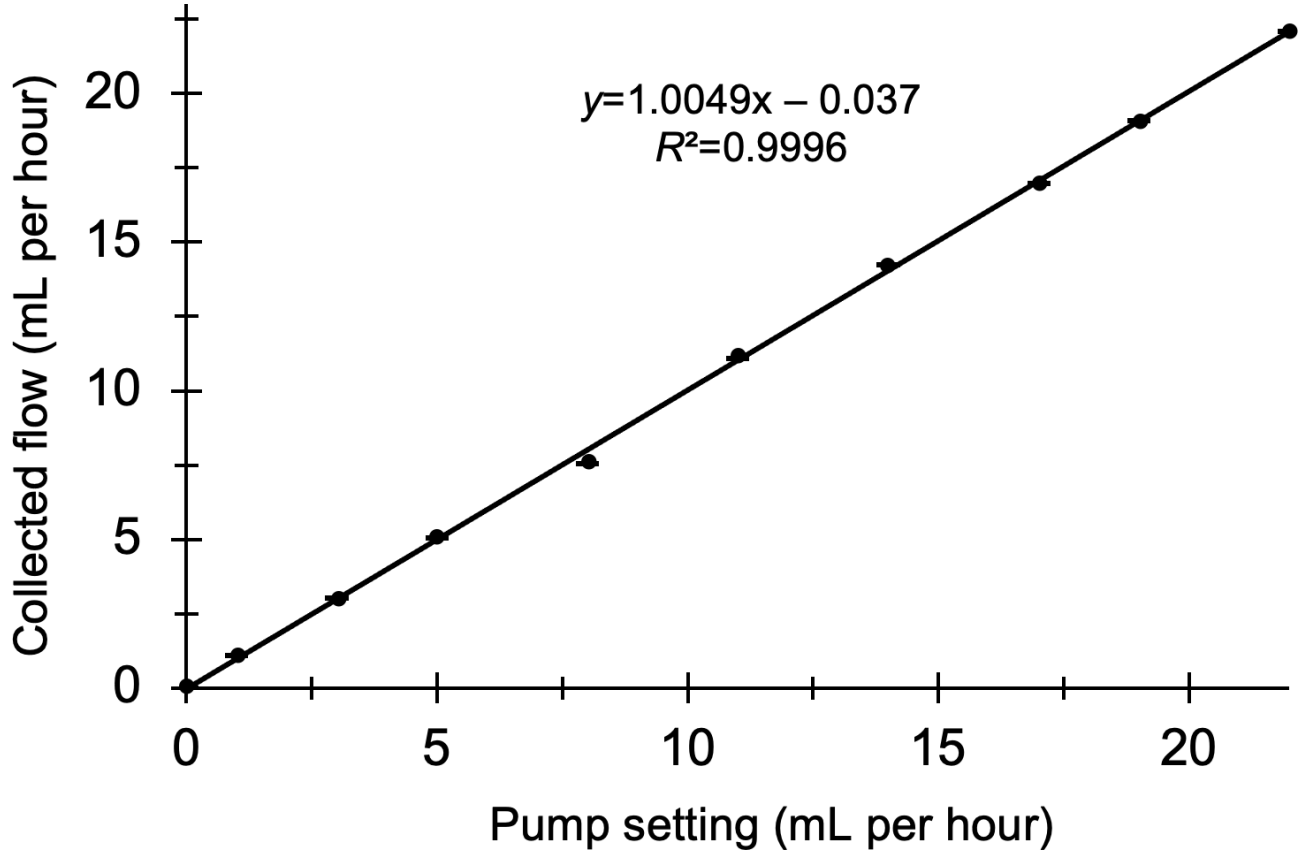

XSL•FO
RenderX

Discussion

Standardizing PC-MRI With Reproducible Phantom Measurements

PC-MRI can provide additional utility in the diagnosis of shunt malfunction by offering a rapid and noninvasive approach to accurately quantify shunt flow. Our protocol presents a framework to calibrate individual MR scanners for accurate PC-MRI flow measurement in multi-institutional clinical data collection. Standardizing PC-MRI measurements and accounting for error introduced by variable MR hardware will allow individual institutions to perform appropriate calibration corrections. The ability to acquire reliable flow measurements between imaging centers can improve the validity of comparison between patient population flow values even when imaging is performed at different sites. Characterizing shunt flow across a large sample during multicenter data collection may improve the clinical assessment of shunt failure and our understanding of shunted hydrocephalus.

We designed a transportable and easy-to-use phantom that ensures magnetic homogeneity for reliable data acquisition. The single-catheter phantom and screw-driven syringe pump combination enables simple, consistent data collection across a physiological range of flows. To represent human shunt flow conditions, our phantom included a common catheter model used in patients at our institution, and water acted as a surrogate for CSF because both fluids have similar magnetic susceptibility and viscosity [27-29]. While CSF flow is known to vary with cardiac and respiratory cycles [30,31], PC-MRI measurements in this study will not be physiologically gated [21]. The potential influence of cardiac [32] or respiratory gating [33] on shunt flow measurement has not yet been studied and remains a potential area of future investigation.

Understanding the reproducibility of PC-MRI flow measurement across MR scanners is useful both to understand the influence of hardware parameters on measured shunt flow and to implement an appropriate calibration correction as needed. MR scanners can vary in field strength, slew rate, gradient strength, number of radiofrequency channel coils, model, and manufacturer at different institutions. This method for phantom model calibration will allow institutions to confirm the reliability of their MR scanners for accurate PC-MRI flow measurement. The protocol outlined in this work aims to streamline the implementation of PC-MRI by preemptively calibrating the scan sequence on a range of scanners and identifying challenges to multi-institutional data collection.

Potential Limitations

As MR scanner time is primarily reserved for clinical use at most institutions, obtaining access for scanner calibration via phantom data collection remains a potential barrier. As PC-MRI images must be exported and externally processed by trained study personnel, the study may be limited by staffing constraints and processing delays. Outside of error introduced by MR hardware, it is important to note that human physiology (eg, susceptibility of the surrounding tissue [34] or altered pulsatility [30,35,36]) and movement within the scanner may provide additional potential variability. While this protocol aims to isolate the hardware-introduced variability in flow measurement, future research should investigate real-world sources of variation in flow measurement.

Conclusions

This work details a protocol for multi-institutional clinical PC-MRI data collection. We also describe the construction of an easy-to-use phantom model of shunt flow and an interinstitutional procedure for scanner calibration to enable standardized collection of PC-MRI flow measurements. The tested phantom and pump combination exhibits accurate and precise outputs at both low and high flow rates while minimizing magnetic field inhomogeneities. An appropriate next step in progress at our institution and forthcoming at others is to generate scanner-specific calibration curves by acquiring PC-MRI data using the single-catheter phantom design. Widespread scanner calibration and validation of clinical feasibility will enable institutions to reliably implement PC-MRI, providing a quantitative method to evaluate shunt function in a broader patient population.

Funding

The authors thank the Los Angeles Kings Care Foundation (awarded to PAC), Do A Little (awarded to PAC), the Rudi Schulte Research Institute (awarded to IF, IT, JMT, PAC, and SB), the Hydrocephalus Association (awarded to PAC), the National Institutes of Health National Institute of Neurological Disorders and Stroke K08 Award (K08NS125175; awarded to PAC), the Kuk Schwartz Family (awarded to PAC), and the University of Southern California Provost Undergraduate Research Fellowship (awarded to JMT) for indirect funding for this study. The funders did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability

The data that support the findings of this study will be available from the corresponding author on reasonable request after study completion.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Phantom fabrication and pump evaluation.
[DOCX File, 875 KB - resprot_v15i1e85918_app1.docx]

Multimedia Appendix 2
Description of methods for flow calculation.
[DOCX File, 26 KB - resprot_v15i1e85918_app2.docx]

Multimedia Appendix 3
List of variables for extraction.
[DOCX File, 30 KB - resprot_v15i1e85918_app3.docx]

References

1. Dewan MC, Rattani A, Mekary R, et al. Global hydrocephalus epidemiology and incidence: systematic review and meta-analysis. J Neurosurg 2018 Apr 1;130(4):1065-1079. [doi: 10.3171/2017.10.JNS17439] [Medline: 29701543]
2. Rekate HL. The definition and classification of hydrocephalus: a personal recommendation to stimulate debate. Cerebrospinal Fluid Res 2008 Jan 22;5:2. [doi: 10.1186/1743-8454-5-2] [Medline: 18211712]
3. Koleva M, De Jesus O. Hydrocephalus. In: StatPearls: StatPearls Publishing; 2023. [Medline: 32809710]
4. Orešković D, Klarica M. Development of hydrocephalus and classical hypothesis of cerebrospinal fluid hydrodynamics: facts and illusions. Prog Neurobiol 2011 Aug;94(3):238-258. [doi: 10.1016/j.pneurobio.2011.05.005] [Medline: 21641963]
5. Turgut M, Jallo G, Di Rocco C, Martínez-Lage JF. Complications of CSF Shunting in Hydrocephalus: Prevention, Identification, and Management: Springer; 2014.
6. Fowler JB, De Jesus O, Mesfin FB. Ventriculoperitoneal shunt. In: StatPearls: StatPearls Publishing; 2023. [Medline: 29083724]
7. Rossi NB, Khan NR, Jones TL, Lepard J, McAbee JH, Klimo P Jr. Predicting shunt failure in children: should the global shunt revision rate be a quality measure? J Neurosurg Pediatr 2016 Mar;17(3):249-259. [doi: 10.3171/2015.5.PEDS15118] [Medline: 26544083]
8. Simon TD, Riva-Cambrin J, Srivastava R, et al. Hospital care for children with hydrocephalus in the United States: utilization, charges, comorbidities, and deaths. J Neurosurg Pediatr 2008 Feb;1(2):131-137. [doi: 10.3171/PED/2008/1/2/131] [Medline: 18352782]
9. Lehner KR, Alawneh RJ, Lee RP, et al. Clinical and socioeconomic predictors of shunt malfunction in the pediatric emergency department. J Neurosurg Pediatr 2024 Apr 5;34(1):75-83. [doi: 10.3171/2024.1.PEDS23489] [Medline: 38579347]
10. Pitetti R. Emergency department evaluation of ventricular shunt malfunction: is the shunt series really necessary? Pediatr Emerg Care 2007 Mar;23(3):137-141. [doi: 10.1097/PEC.0b013e3180328c77] [Medline: 17413426]
11. Jayanth A, Benabbas R, Chao JH, Sinert RH. Diagnostic modalities to determine ventriculoperitoneal shunt malfunction: a systematic review and meta-analysis. Am J Emerg Med 2021 Jan;39:180-189. [doi: 10.1016/j.ajem.2020.09.024] [Medline: 33067062]
12. Zhang L, Hussain Z, Ren Z. Recent advances in rational diagnosis and treatment of normal pressure hydrocephalus: a critical appraisal on novel diagnostic, therapy monitoring and treatment modalities. Curr Drug Targets 2019;20(10):1041-1057. [doi: 10.2174/1389450120666190214121342] [Medline: 30767741]
13. Desai KR, Babb JS, Amodio JB. The utility of the plain radiograph “shunt series” in the evaluation of suspected ventriculoperitoneal shunt failure in pediatric patients. Pediatr Radiol 2007 May;37(5):452-456. [doi: 10.1007/s00247-007-0431-3] [Medline: 17380325]
14. Docter TA, Patel BH, Brennan JJ, Castillo EM, Lee RR, Vilke GM. Utility of shunt series in the evaluation of ventriculoperitoneal shunt dysfunction in adults. J Emerg Med 2020 Mar;58(3):391-397. [doi: 10.1016/j.jemermed.2019.10.001] [Medline: 31806435]
15. Rocque BG, Lapsiwala S, Iskandar BJ. Ventricular shunt tap as a predictor of proximal shunt malfunction in children: a prospective study. J Neurosurg Pediatr 2008 Jun;1(6):439-443. [doi: 10.3171/PED/2008/1/6/439] [Medline: 18518693]
16. Quezada JJ, McComb JG. Reliability of the radiopharmaceutical shunt flow study for the detection of a CSF shunt malfunction in the presence of stable ventricular size. J Neurosurg Pediatr 2020 Jul 17;26(4):364-370. [doi: 10.3171/2020.4.PEDS2020] [Medline: 32679559]
17. Tsai SY, Wang SY, Shiau YC, Yang LH, Wu YW. Clinical value of radionuclide shuntography by qualitative methods in hydrocephalic adult patients with suspected ventriculoperitoneal shunt malfunction. Medicine (Baltimore) 2017 Apr;96(17):e6767. [doi: 10.1097/MD.0000000000006767] [Medline: 28445308]
18. Graham MM. Practical aspects of nuclear medicine ventriculoperitoneal shunt evaluation. J Nucl Med Technol 2023 Dec 5;51(4):271-276. [doi: 10.2967/jnmt.123.266203] [Medline: 37963777]
19. Reddy GK. Ventriculoperitoneal shunt surgery and the incidence of shunt revision in adult patients with hemorrhage-related hydrocephalus. Clin Neurol Neurosurg 2012 Nov;114(9):1211-1216. [doi: 10.1016/j.clineuro.2012.02.050] [Medline: 22472352]

XSL•FO
RenderX

20. Di Rocco C, Massimi L, Tamburrini G. Shunts vs endoscopic third ventriculostomy in infants: are there different types and/or rates of complications? A review. Childs Nerv Syst 2006 Dec;22(12):1573-1589. [doi: 10.1007/s00381-006-0194-4] [Medline: 17053941]
21. Ha JH, Borzage MT, Vanstrum EB, et al. Quantitative noninvasive measurement of cerebrospinal fluid flow in shunted hydrocephalus. J Neurosurg 2023 Oct 6;140(4):1117-1128. [doi: 10.3171/2023.7.JNS231326] [Medline: 38564811]
22. König RE, Stucht D, Baecke S, et al. Phase-contrast MRI detection of ventricular shunt CSF flow: proof of principle. J Neuroimaging 2020 Nov;30(6):746-753. [doi: 10.1111/jon.12794] [Medline: 33146931]
23. Kurwale NS, Agrawal D. Phase-contrast magnetic resonance imaging of intracranial shunt tube: a valuable adjunct in the diagnosis of ventriculoperitoneal shunt malfunction. Clin Neurosurg 2011;58:138-142. [doi: 10.1227/neu.0b013e31822784cf] [Medline: 21916139]
24. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
25. Shellock FG. Reference Manual for Magnetic Resonance Safety, Implants, and Devices: Biomedical Research Publishing Group; 2014.
26. Shah A, Aran S. A review of magnetic resonance (MR) safety: the essentials to patient safety. Cureus 2023 Oct 19;15(10):e47345. [doi: 10.7759/cureus.47345] [Medline: 38021512]
27. Bloomfield IG, Johnston IH, Bilston LE. Effects of proteins, blood cells and glucose on the viscosity of cerebrospinal fluid. Pediatr Neurosurg 1998 May;28(5):246-251. [doi: 10.1159/000028659] [Medline: 9732257]
28. Straub S, Schneider TM, Emmerich J, et al. Suitable reference tissues for quantitative susceptibility mapping of the brain. Magn Reson Med 2017 Jul;78(1):204-214. [doi: 10.1002/mrm.26369] [Medline: 27529579]
29. Dimov AV, Nguyen TD, Spincemaille P, et al. Global cerebrospinal fluid as a zero-reference regularization for brain quantitative susceptibility mapping. J Neuroimaging 2022 Jan;32(1):141-147. [doi: 10.1111/jon.12923] [Medline: 34480496]
30. Mestre H, Tithof J, Du T, et al. Flow of cerebrospinal fluid is driven by arterial pulsations and is reduced in hypertension. Nat Commun 2018 Nov 19;9(1):4878. [doi: 10.1038/s41467-018-07318-3] [Medline: 30451853]
31. Gutiérrez-Montes C, Coenen W, Vidorreta M, et al. Effect of normal breathing on the movement of CSF in the spinal subarachnoid space. AJNR Am J Neuroradiol 2022 Sep;43(9):1369-1374. [doi: 10.3174/ajnr.A7603] [Medline: 35981761]
32. Karki P, Murphy MC, Ganji S, et al. Real-time 2D phase-contrast MRI to assess cardiac- and respiratory-driven CSF movement in normal pressure hydrocephalus. J Neuroimaging 2025;35(1):e70000. [doi: 10.1111/jon.70000] [Medline: 39737803]
33. Akçakaya M, Gulaka P, Basha TA, Ngo LH, Manning WJ, Nezafat R. Free-breathing phase contrast MRI with near 100% respiratory navigator efficiency using k-space-dependent respiratory gating. Magn Reson Med 2014 Jun;71(6):2172-2179. [doi: 10.1002/mrm.24874] [Medline: 23900942]
34. Duyn JH, Schenck J. Contributions to magnetic susceptibility of brain tissue. NMR Biomed 2017 Apr;30(4). [doi: 10.1002/nbm.3546] [Medline: 27240118]
35. Causemann M, Vinje V, Rognes ME. Human intracranial pulsatility during the cardiac cycle: a computational modelling framework. Fluids Barriers CNS 2022 Nov 1;19(1):84. [doi: 10.1186/s12987-022-00376-2] [Medline: 36320038]
36. Linninger AA, Tsakiris C, Zhu DC, et al. Pulsatile cerebrospinal fluid dynamics in the human brain. IEEE Trans Biomed Eng 2005 Apr;52(4):557-565. [doi: 10.1109/TBME.2005.844021] [Medline: 15825857]

Abbreviations:

CSF: cerebrospinal fluid
HIPAA: Health Insurance Portability and Accountability Act
IRB: institutional review board
MR: magnetic resonance
MRI: magnetic resonance imaging
PC-MRI: phase-contrast magnetic resonance imaging
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

Edited by J Sarvestan; submitted 15.Oct.2025; peer-reviewed by D Garic, SA Mohammad; revised version received 16.Dec.2025; accepted 27.Dec.2025; published 20.Feb.2026.

Please cite as:
Toliao JM, Borzage MT, Chaudhari PP, Ha JH, Friedli I, Gutierrez ME, Drastal M, Calabrese P, Doyle EK, Torres I, Al-Husseini J, González-Zacarías C, Wood JC, Hauptman J, McComb JG, Blüml S, Chiarelli PA
Noninvasive Measurement of Cerebrospinal Fluid Flow in Shunted Hydrocephalus: Protocol for Scanner Calibration and Multisite Data Collection
JMIR Res Protoc 2026;15:e85918
URL: https://www.researchprotocols.org/2026/1/e85918
doi:10.2196/85918

© Jason M Toliao, Matthew T Borzage, Pradip P Chaudhari, Joseph H Ha, Isabella Friedli, Madison E Gutierrez, Meghan Drastal, Peter Calabrese, Eamon K Doyle, Isabel Torres, Jacob Al-Husseini, Clio González-Zacarías, John C Wood, Jason Hauptman, J Gordon McComb, Stefan Blüml, Peter A Chiarelli. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Exploring Risk Factors and Neurophysiological Mechanisms Underlying the Development of Chronic Postsurgical Pain After Thoracic Surgery: Protocol for an Observational Feasibility Study

Mustaali Hussain[1], BSc; Stevie Dante Foglia[2], MSc, PhD; Ameer Hamoodi[1], BSc; Jinhui Ma[3], PhD; John Agzarian[4], MD; Christian John Finley[4], MD; Yaron Shargall[4], MD; Aimee Jennifer Nelson[1,2], PhD; Harsha Shanthanna[5], MD, PhD

[1]Department of Kinesiology, Faculty of Science, McMaster University, Hamilton, ON, Canada

[2]School of Biomedical Engineering, McMaster University, Hamilton, ON, Canada

[3]Department of Health Research Methods, Evidence, and Impact, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

[4]Division of Thoracic Surgery, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

[5]Department of Anesthesia, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Harsha Shanthanna, MD, PhD
Department of Anesthesia
Faculty of Health Sciences
McMaster University
1280 Main St W
Hamilton, ON, L8S4K1
Canada
Phone: 1 9055259140
Email: shanthh@mcmaster.ca

Abstract

Background: Chronic postsurgical pain (CPSP) is a debilitating chronic pain condition that particularly impacts patients undergoing thoracic surgery, with incidence rates of up to 50%. The current understanding of risk factors is limited, and preoperative neurophysiological risk factors that may predict the development of CPSP have not yet been explored. Additionally, the specific neural mechanisms underlying the transition to CPSP are not well characterized. As a novel approach, we propose the use of transcranial magnetic stimulation and electroencephalography, along with other patient and surgical factors, to understand the neurophysiological mechanisms underlying the onset of CPSP after thoracic surgery.

Objective: The primary objective of this study is to evaluate the feasibility of our study design to inform a larger observational cohort study. Secondary objectives include exploring preoperative neurophysiological markers along with clinical characteristics associated with a higher risk of developing CPSP, as well as exploring postoperative differences in cortical function between patients who undergo thoracic surgery and develop CPSP compared with those who do not develop CPSP.

Methods: A total of 30 participants undergoing video-assisted thoracic surgery or a robotic-assisted thoracic lobectomy, wedge resection, segmental section, or minimally invasive esophagectomy, will be recruited to take part in 2 assessment sessions. The first assessment will take place 2 to 3 weeks before surgery, and the second assessment will take place 3 months after surgery, during which the CPSP diagnosis of each participant will be assessed by the experimenter using a validated definition. Feasibility outcomes include recruitment and retention rates of study participants. The secondary objectives include exploring factors associated with the development of CPSP, as well as examining postoperative differences in neurophysiological measures between patients with and without CPSP. We will consider the following neurophysiological measures for these objectives: transcranial magnetic stimulation measures of short-latency intracortical inhibition, cortical silent period, and motor evoked potentials; electroencephalography measures of resting band activity, event-related desynchronization, and corticomuscular coherence; and quantitative sensory testing of mechanical detection threshold and pressure pain threshold.

Results: This is an ethics-approved, ongoing study. Initial funding for this study was provided in March 2023. Recruitment for the study began in January 2025. A total of 22 participants have been recruited for the study. We anticipate completing data collection for this study by April 2026, with data analysis to follow.

XSL•FO
RenderX

Conclusions: This protocol details our study design for a feasibility study exploring the neurophysiological markers and patient characteristics associated with the development of CPSP. Demonstration of feasibility is expected to lead to a larger study. Improved understanding of the risk factors and mechanisms underlying CPSP may inform the delivery of targeted therapies and preventive measures to reduce the incidence of CPSP after thoracic surgery.

International Registered Report Identifier (IRRID): DERR1-10.2196/81042

(JMIR Res Protoc 2026;15:e81042) doi:10.2196/81042

KEYWORDS

chronic postsurgical pain; thoracic surgery; transcranial magnetic stimulation; TMS; electroencephalography; EEG; risk factors; neurophysiological mechanisms

Introduction

Chronic postsurgical pain (CPSP) is defined by the International Association for the Study of Pain as pain localized to the operated area that persists for at least 3 months after surgery and cannot be explained by other potential causes [1]. CPSP is a debilitating condition that negatively impacts patients' quality of life [2,3], while also placing a significant burden on health care providers and hospitals [2]. While the incidence of CPSP can vary across surgical procedures, thoracic surgery shows a markedly high rate, with rates between 30% and 50% [4-7].

Understanding the preoperative profile that predicts the development of CPSP is important to identify patients at higher risk and implement early interventions. Exploring neurophysiological risk factors using transcranial magnetic stimulation (TMS) and electroencephalography (EEG) may improve our understanding beyond the currently identified risk factors, such as female sex [6,8], depression and anxiety [5,8], younger age, and preoperative pain [8-10], which alone show only moderate prognostic value [11]. In previous work, measures of cortical function via TMS or EEG have successfully predicted the development of chronic neuropathic pain [12] and other neurological conditions [13].

These neurophysiological tools may also provide insight into the maladaptive mechanisms underlying the persistence of CPSP after thoracic surgery. While central sensitization of pain processing centers to nociceptive input has been the prevailing theory behind the persistence of CPSP [14], the specific dysregulations involved remain incompletely understood. In other chronic pain conditions, TMS assessments have revealed reductions in cortical excitability and intracortical inhibition [15-19], suggesting impairments in glutamatergic and γ-aminobutyric acid-mediated neural transmission. These markers may also explain the associations between depression and the risk of CPSP [20] but are unlikely to account for other risk factors, such as younger age [21]. EEG assessments among individuals with various chronic pain conditions have demonstrated elevated band power across various frequency bands at rest [22-27] and elevated cortical desynchronization during movements across various cortical regions [25,28], which indicate impairments in sensory processing. Such EEG markers may also provide insight into maladaptive cognitive states, such as pain catastrophizing, another risk factor for CPSP [8]. Collectively, these findings indicate that TMS and EEG measurements may provide valuable insights into the neurophysiological risk factors and mechanisms underlying the development and presence of CPSP after thoracic surgery.

Beyond these mechanistic aims, an important consideration is assessing the feasibility of longitudinally studying patients undergoing thoracic surgery before and after surgery. Given that these individuals have been shown to exhibit higher levels of fear of surgery preoperatively compared with those undergoing other surgeries [29], the ability to recruit them for a neurophysiological assessment study may present challenges, due to their elevated concerns regarding the surgery itself. In addition, the fact that thoracic surgery leads to postoperative complications in up to one-third of patients [30] alongside hospital readmission rates of 10% to 20% up to 90 days after surgery [30] suggests that performing follow-up neurophysiological testing on these patients after surgery may also be challenging and lead to high attrition rates.

Therefore, the primary objective of this study is to assess the feasibility of recruitment and retention of patients undergoing thoracic surgery before and after surgery, which will help inform the planning of a larger cohort study. Secondary objectives for this pilot study include exploring preoperative neurophysiological markers (in conjunction with clinical profile characteristics) associated with a higher risk of CPSP, as well as identifying postoperative differences in cortical function between patients who develop CPSP (CPSP+) and those who do not develop CPSP (CPSP–).

Methods

Recruitment

This study will recruit 30 participants aged between 18 and 80 years who are scheduled to undergo video-assisted thoracoscopic surgery or a robotic-assisted thoracic lobectomy, wedge resection, segmental resection, or a minimally invasive esophagectomy (MIE) at St Joseph's Healthcare (SJH; Hamilton, Ontario). Although open thoracotomies and esophagectomies are associated with a higher risk of CPSP, they are not commonly performed. Participants must have comprehension of the English language or have an interpreter present to participate. Participants will be excluded if (1) they have any contraindications to undergoing TMS, including (but not limited to) the presence of a pacemaker, metal, electrical, or magnetic implants near the head or neck (not including titanium); a known history of untreated or uncontrolled psychological disorders, pregnancy, history of seizure, or a diagnosis of epilepsy, or if they are taking prescription medications that increase the risk

XSL•FO
RenderX

of seizure; (2) they have a known significant psychological impairment that may affect comprehension of the study's instructions or their ability to participate; or (3) they are participating in another research study intervention that may impact the development of CPSP. The presence of other comorbidities or pre-existing pain conditions will not be a reason for exclusion from our study.

The research team will obtain a rolling list of patients consented for surgery from the thoracic surgery team at SJH, and these patients will subsequently be contacted to be given a detailed explanation of the study protocol and to assess eligibility through completing the TMS screening questionnaire. If eligible and willing to participate, participants will be invited to complete a written informed consent form. Participants will be informed that they can withdraw at any time during the study period. Further, they have the right to withdraw their data from the study for up to 2 weeks following their withdrawal or completion of the study. All assessment sessions will be conducted at the McMaster University.

Ethical Considerations

This study will be conducted in compliance with the protocol and the principles laid down in the Declaration of Helsinki. Approval to conduct this study has been granted by the Hamilton Integrated Research Ethics Board (17806). Participants will provide written informed consent prior to participation, and they will be made aware that they may withdraw at any time during the study period.

Study Design

This is an observational feasibility study in individuals undergoing video-assisted thoracic surgery or robotic-assisted thoracic lobectomy, wedge resection, segmental resection, or MIE. Participants will attend 2 assessment sessions: the first will occur 2 to 3 weeks before their scheduled procedure, and the second will take place 3 months after surgery (Figure 1). Each session will span about 3 hours and will be scheduled flexibly, with ample breaks provided as needed to minimize participant burden.

Figure 1. Experimental timeline of the study. All participants will take part in 2 assessments. Assessment 1 (T0) will occur 2-3 weeks before the participants' scheduled procedure. Participants will then complete assessment 2 (T1) at 3 months after surgery. Both assessments will involve collecting clinical profile information using questionnaires, neurophysiological measures acquired using transcranial magnetic stimulation (TMS) and electroencephalography (EEG), and measures of sensation with quantitative sensory testing (QST). See table 1 for the specific measurements acquired. Diagnosis of chronic postsurgical pain (CPSP) will be evaluated in the second assessment. Each assessment will take approximately 3 hours to complete.

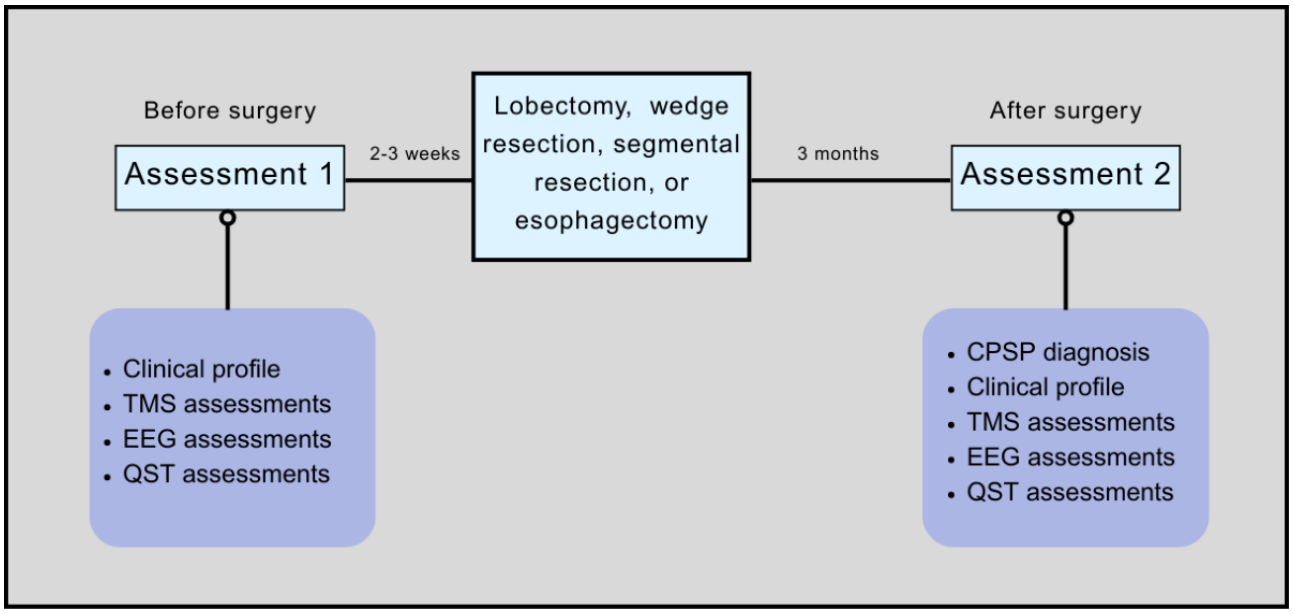

Study Outcomes

Feasibility

The primary objective is to assess the feasibility of the study design. Feasibility outcomes will include (1) participant recruitment and (2) participant retention for the postsurgical assessment. On the basis of past records of the thoracic surgery division at SJH, we aim to recruit 3 to 4 patients per month over a 10-month study period. We will consider monthly recruitment of 3 to 4 patients and >90% retention to be feasible, monthly recruitment of 1 to 2 patients and 85% to 90% retention as feasible with appropriate modifications, and no patients recruited in 5 out of 10 months and <85% retention as not feasible. These cutoffs will be used to evaluate the need for protocol revisions before conducting a large-scale study.

Demographic and Clinical Information

In the presurgical assessment, information regarding each participant's age, sex, history of prior surgery, history of preexisting pain, reason for upcoming surgery, other comorbidities, adjuvant therapies, and pain medication use will be documented. In the postsurgical assessment, details of each patient's surgery type, number of ports, adjuvant therapies, and use of pain medications will be documented by the experimenter. In addition, the anesthesia protocols used for each participant's surgery will be collected through electronic chart data, including the anesthetic technique (volatile or total intravenous), regional anesthesia type (thoracic epidural, paravertebral, erector spinae

XSL•FO
RenderX

plane, or intercostal block), delivery mode (single-shot or continuous catheter), and duration of postoperative infusion.

This information will be reported descriptively and considered during data analysis so that its potential impact can be considered in the interpretation of the results.

Questionnaires

The questionnaires that will be completed in a self-administered manner during each assessment session are presented in Textbox 1.

Textbox 1. Questionnaires.

Patient-Reported Outcomes Measurement Information System-29 (PROMIS-29; v2.0)

The PROMIS-29 questionnaire will be used to assess 7 health domains, including physical function, anxiety, depression, fatigue, sleep disturbances, ability to participate in social roles and activities, and pain interference [31]. The 0 to 10 numerical rating scale included in the PROMIS will be used to assess the presence and severity of pain. Scores from this questionnaire will be expressed as *t* scores, which are standardized relative to a reference healthy population (mean 50, SD 10). Higher scores correspond to increased levels of the measured domain [32]. This questionnaire was chosen as it can efficiently capture a broad range of health domains.

Patient Health Questionnaire-4 (PHQ-4)

The PHQ-4 [33] is a 4-item self-report screening questionnaire for depression and anxiety symptoms experienced in the past 2 weeks, consisting of the Generalized Anxiety Disorder-2 (GAD-2; 2 core criteria for generalized anxiety disorder) [34] and the PHQ-2 (2 core criteria for depression) [35]. This will be used to assess symptoms of depression and anxiety. A score of ≥3 for GAD-2 suggests the presence of anxiety, and a score of ≥3 for PHQ-2 suggests the presence of depression. Given that depression and anxiety are cited as potential risk factors for developing chronic postsurgical pain [5,8], depression and anxiety levels assessed from these questionnaires will also be incorporated as features within our analysis. This questionnaire was selected for its validity [36] and its ultrabrief design to minimize participant burden.

TMS Assessments

Overview

TMS will be performed using a 50-mm figure-of-8 coil attached to a Magstim 200^2 stimulator (Magstim). This study will focus on stimulating the motor hot spot of the right first dorsal interosseus (FDI) muscle for all TMS measures. Surface electrodes (9 mm Ag-AgCl) will be used to record electromyography (EMG) activity from the FDI muscle of the right hand.

Resting motor threshold, defined as the lowest intensity required to evoke a motor evoked potential (MEP) ≥50 µV in 50% of the time, will be determined for each participant's FDI muscle [37]. Active motor threshold, defined as the lowest intensity required to evoke an MEP≥200-µV in 5 out of 10 consecutive trials [37] while participants maintain a sustained isometric contraction of the right FDI muscle to 10% of their maximum voluntary contraction (MVC), will also be recorded. This MVC value will be determined beforehand, via participants performing a brief 5-second isometric voluntary maximum contraction of the right FDI muscle, with maximal EMG activity recorded.

MEP Analysis

MEPs will be acquired by stimulating the motor hot spot for the right FDI muscle at 120% of the resting motor threshold for a total of 30 pulses at rest. MEPs will also be collected during voluntary contraction of the target FDI muscle to 50% MVC, with 20 pulses delivered at an intensity that evokes a resting MEP with approximately 1 mV peak-to-peak amplitude (determined beforehand). Average MEP amplitude will be calculated for each condition, and this measure will be used to assess the levels of corticospinal excitability.

Short-Latency Intracortical Inhibition

Short-latency intracortical inhibition (SICI), a measure of cortical inhibition, is measured using paired-pulse TMS with an interstimulus interval of 2 ms between the conditioning stimulus (CS) and test stimulus (TS). The CS will be set to 90% of the active motor threshold, and the TS will be set to the intensity evoking an MEP with a peak-peak amplitude of approximately 1 mV in the right FDI. SICI will be calculated as the ratio of the average conditioned MEP (CS-TS) to the amplitude of the average unconditioned MEP (MEP-TS only), with smaller values indicating greater levels of intracortical inhibition, and vice versa. A total of 25 TS and 25 CS+TS pulses will be delivered.

Cortical Silent Period

Cortical silent period (CSP), a measure of spinal and cortical inhibition, is an interruption in the ongoing voluntary EMG signal that occurs following an MEP evoked by a suprathreshold TMS pulse. The CSP will be probed by delivering single-pulse TMS during voluntary contraction at approximately 50% MVC of the FDI. A total of 20 pulses will be delivered at an intensity evoking a resting MEP with a peak-peak amplitude of approximately 1 mV (determined beforehand). CSP duration will be measured for each trial and averaged. CSP onset will be defined as the beginning of the TMS-evoked MEP, and CSP offset will be defined as when the voluntary EMG signal reappears. Once CSP onset and offset are determined, CSP duration will be calculated.

Quantitative Sensory Testing

Quantitative sensory testing (QST) will be used in this study to assess somatosensory function before and after surgery. The series of tests to be used includes the mechanical detection threshold (MDT) and the pressure pain threshold (PPT) tests. For patients undergoing lobectomy, wedge resection, or segmental resection, QST will be performed unilaterally at 3 sites on the sides of the trunk (approximately the fourth, seventh, and ninth intercostal spaces) [38,39]. For patients undergoing MIE, QST will be performed bilaterally at 3 sites on the sides of the trunk (approximately the fourth, seventh, and ninth intercostal spaces) and 1 site at the abdomen (around the navel)

XSL•FO
RenderX

[40]. QST results will be analyzed in their original measurement units (lb for PPT and mN for MDT).

EEG Assessments

EEG will be recorded using an active EEG electrode system (Quik-Cap Neo Net; Compumedics Neuroscan) connected to a Grael 4K EEG amplifier (Compumedics Neuroscan) with 32 channels of recording electrodes positioned across the cortical hemisphere using the international 10 to 20 system. EEG data will be used to derive the following 3 outcomes.

Band Activity

Spontaneous EEG will be recorded under 2 conditions: eyes open (EO) and eyes closed (EC). Each condition will last 5 minutes. During the EO condition, a white fixation cross will be presented at the center of the screen for patients to focus on. During the EC condition, patients will sit relaxed with their eyes closed. Experimenters will ensure that patients remain awake during the EC condition through continuous monitoring.

Following findings from Vuckovic et al [41], features will be based on the power spectral density calculated for each frequency band (alpha [8-12 Hz] and beta [13-30 Hz]) for the EO and EC tasks. Additionally, the ratio between the EC and EO states for each frequency band will be calculated [41]. Features will be computed for each of the 32 channels.

Event-Related Desynchronization

Event-related desynchronization (ERD) will be measured during periodic submaximal isometric contractions of the abductor pollicis brevis muscle.

ERD measures the reduction in the power of synchronized neural oscillations across various frequency bands in response to movement and serves as a marker of increased cortical excitability [28]. ERD will be assessed using an EEG recorded from electrodes positioned over the primary motor cortex (C3 and C4). ERD will be calculated by comparing alpha and beta band power during the active movement phase (0.5-1.5 seconds) to a premovement baseline period (−3.5 to −0.5 s), using standard event-related spectral perturbation analysis [42].

Corticomuscular Coherence

Corticomuscular coherence (CMC) will be measured during periodic submaximal isometric contractions of the abductor pollicis brevis muscle.

CMC is a measure of the synchrony between oscillations in EEG signals and EMG signals at the muscle, suggested to reflect reciprocal communication between the motor cortex and active muscle during movement [43]. Deficits in CMC have been shown to reflect impairments in sensorimotor integration in chronic pain [44]. In this study, CMC will be calculated from EEG recordings over C3 and C4 and EMG signals from the right abductor pollicis brevis muscle using established correlation-based methods [45], focusing on the mu (8-12 Hz) and beta (15-30 Hz) frequency bands.

Measurements Acquired in Each Assessment

The measurements acquired during each assessment are summarized in Table 1. Each of these TMS and EEG measures used has shown good to excellent within-session or between-session reliability [46-51], and therefore provides a reliable basis for evaluating cortical activity, excitability, and inhibition.

XSL•FO
RenderX

Table 1. Measurements acquired in each assessment.

Assessment	Outcomes measured
T0 (before surgery)	• Demographic and clinical information (age, sex, surgery history, comorbidities, adjuvant therapies, and pain medications) • Questionnaires (PROMIS-29[a] v2.1 and PHQ-4[b]) • TMS[c] (MEPs[d], SICI[e], and CSP[f]) • EEG[g] (band activity, ERD[h], and CMC[i]) • QST[j] (MDT[k] and PPT[l])
T1 (3 months after surgery)	• Demographic and clinical information (age, sex, surgery history, surgery type, number of ports, comorbidities, adjuvant therapies, and pain medications) • Questionnaires (PROMIS-29 v2.1 and PHQ-4) • TMS (MEPs, SICI, and CSP) • EEG (band activity, ERD, and CMC) • QST (MDT and PPT)

[a]PROMIS-29: Patient-Reported Outcomes Measurement Information System–29.

[b]PHQ-4: Patient Health Questionnaire-4.

[c]TMS: transcranial magnetic stimulation.

[d]MEP: motor evoked potential.

[e]SICI: short-latency intracortical inhibition.

[f]CSP: cortical silent period.

[g]EEG: electroencephalography.

[h]ERD: event-related desynchronization.

[i]CMC: corticomuscular coherence.

[j]QST: quantitative sensory testing.

[k]MDT: mechanical detection threshold.

[l]PPT: pressure pain threshold.

Diagnosis of CPSP

At T1, participants will be categorized as CPSP+ or CPSP–. CPSP– will be defined as a participant experiencing no persistent pain related to the surgery 3 months after the operation. CPSP+ will be defined as the presence of new persistent pain that is localized to the operated area (of any severity) or pain of increased severity (if there was pre-existing pain in that site) 3 months after the operation, which cannot be attributable to any other identifiable cause. This diagnosis will be evaluated by asking the following question: Do you have persistent pain (that started or increased with or after surgery) at or near the surgical area? For patients who answer yes, the severity of pain localized to the surgical area will be assessed using the numeric rating scale [52].

Statistical Analysis

Power

The sample size for this study is based on evaluating feasibility for a larger study, and thus is not sufficiently powered to detect clinical differences. We based our sample size on pilot trial sample size calculations using a CI approach suggested by Thabane et al [53]. In addition, our sample size was determined in consultation with the thoracic surgery team at SJH, who provided data on the number of surgeries performed within the department.

Identifying Prospective Markers Associated With CPSP Development

Univariable logistic regression will be conducted to explore the associations between preoperative and perioperative characteristics and CPSP status. These analyses will be considered hypothesis-generating to inform the planned larger cohort study. We will test the individual association of the following independent variables with the diagnosis of CPSP: demographic and clinical information (age, sex, baseline pain intensity, depression and anxiety levels, surgery type, and regional anesthesia information), TMS measures (SICI, CSP, and MEPs), EEG measures (band activity, ERD, and CMC), and QST measures (MDT and PPT). Variables that show a statistically significant relationship ($P<.05$) will be considered independent predictors of CPSP development in this pilot study. Odds ratios with 95% CIs will also be reported.

Identifying Neurophysiological Mechanisms Underlying the Presence of CPSP

Exploratory comparisons of postsurgical TMS and EEG measures between patients who are CPSP+ and CPSP– will be conducted using 2-sample 2-tailed *t* tests. Significance will be set to $\alpha=.05$, and effect sizes (Cohen *d*) will be reported with 95% CIs. Before this analysis, normality will be assessed with the Shapiro-Wilk test, and outliers will be removed with Grubbs test.

These secondary exploratory aims will be analyzed using an available-case approach, whereby participants with missing data

XSL•FO
RenderX

for a given variable will be excluded only from analyses involving that variable and retained in all others.

Results

Initial funding for this study was provided in March 2023. Recruitment for the study began in January 2025. A total of 22 participants have been enrolled in the study and have completed the presurgical assessment. Of these, 10 participants have completed the postsurgical assessment. We anticipate completing data collection for this study by April 2026, with data analysis to follow.

Discussion

Understanding and mitigating the high incidence of CPSP after thoracic surgery remains an unmet medical need. This pilot study takes a novel approach by investigating the potential neurophysiological risk factors and mechanisms underlying its development. This pilot study will allow for an understanding of the feasibility of conducting such an investigation on a large scale, through assessing participant recruitment and retention rates, which may allow us to identify potential recruitment or study design challenges that can lead to attrition and burden. Prior use of TMS [15,18,19,54-57] and EEG [22-25] to successfully identify impairments in cortical function in other chronic pain conditions supports the potential of our work to improve the understanding of factors that increase the risk of developing CPSP after thoracic surgery, as well as to advance knowledge on neural mechanisms underlying its presence. Such an investigation may serve to support the understanding of CPSP to improve preoperative identification and prevention, as well as inform the development of effective treatment strategies.

This study has several limitations. As a single-center feasibility study with a sample size of 30 participants, the analyses will not be powered for any definitive conclusions. In addition, pain trajectories during the immediate postoperative period and longer-term follow-up beyond 3 months will not be assessed due to logistical constraints. In addition, the assessment of CPSP status did not include further information on pain categorization. These are important considerations for future studies to strengthen their clinical relevance. The neurophysiological assessments used here are time-intensive and largely confined to research settings. However, they are essential for establishing a mechanistic understanding, which may inform more simplified, bedside-compatible assessments in the future, such as bedside EEG. Despite these limitations, this study will provide critical feasibility data and preliminary mechanistic insight to guide the design of a larger cohort study examining the neurophysiological mechanisms of CPSP.

Funding

This work was supported by the Canadian Research Chairs Program to AJN, the Canadian Institutes of Health Research Scholarship Program, and the Michael G DeGroote Institute for Pain Research and Care.

Authors' Contributions

Conceptualization: SDF (lead), HS (equal), AJN (equal), MH (supporting)
Data curation: MH (lead), SDF (equal), HS (equal), AJN (equal), AH (supporting)
Funding acquisition: AJN (lead), HS (equal), SDF (equal), MH (equal)
Investigation: MH (lead), SDF (supporting), AH (supporting), AJN (supporting), HS (supporting)
Methodology: SDF (lead), AJN (equal), HS (equal), MH (supporting)
Project administration: HS (lead), AJN (equal)
Resources: HS (lead), AJN (equal), JA (supporting), CJF (supporting), YS (supporting)
Supervision: HS (lead), AJN (equal)
Writing—original draft: MH (lead), SDF (supporting), AH (supporting), JM (supporting)
Writing—review and editing: MH (lead), SDF (supporting), AH (supporting), JM (supporting), JA (supporting), CJF (supporting), YS (supporting), AJN (supporting), HS (supporting)

Conflicts of Interest

None declared.

References

1. Schug SA, Lavand'homme P, Barke A, Korwisi B, Rief W, Treede RD. The IASP classification of chronic pain for ICD-11: chronic postsurgical or posttraumatic pain. Pain 2019 Jan;160(1):45-52. [doi: 10.1097/j.pain.0000000000001413] [Medline: 30586070]
2. Parsons B, Schaefer C, Mann R, Sadosky A, Daniel S, Nalamachu S, et al. Economic and humanistic burden of post-trauma and post-surgical neuropathic pain among adults in the United States. J Pain Res 2013 Jun 17;6:459-469 [FREE Full text] [doi: 10.2147/JPR.S44939] [Medline: 23825931]
3. Peng Z, Li H, Zhang C, Qian X, Feng Z, Zhu S. A retrospective study of chronic post-surgical pain following thoracic surgery: prevalence, risk factors, incidence of neuropathic component, and impact on qualify of life. PLoS One 2014 Feb 28;9(2):e90014 [FREE Full text] [doi: 10.1371/journal.pone.0090014] [Medline: 24587187]

XSL•FO
RenderX

4. Bayman EO, Brennan TJ. Incidence and severity of chronic pain at 3 and 6 months after thoracotomy: meta-analysis. J Pain 2014 Sep;15(9):887-897 [FREE Full text] [doi: 10.1016/j.jpain.2014.06.005] [Medline: 24968967]
5. Khan JS, Dana E, Xiao MZ, Rao V, Djaiani G, Seltzer Z, et al. Prevalence and risk factors for chronic postsurgical pain after thoracic surgery: a prospective cohort study. J Cardiothorac Vasc Anesth 2024 Feb;38(2):490-498. [doi: 10.1053/j.jvca.2023.09.042] [Medline: 39093584]
6. Chen WC, Bai YY, Zhang LH, Liu CY, Liang JW, He HF. Prevalence and predictors of chronic postsurgical pain after video-assisted thoracoscopic surgery: a systematic review and meta-analysis. Pain Ther 2023 Feb;12(1):117-139 [FREE Full text] [doi: 10.1007/s40122-022-00439-0] [Medline: 36227420]
7. Zhang Y, Zhou R, Hou B, Tang S, Hao J, Gu X, et al. Incidence and risk factors for chronic postsurgical pain following video-assisted thoracoscopic surgery: a retrospective study. BMC Surg 2022 Mar 02;22(1):76 [FREE Full text] [doi: 10.1186/s12893-022-01522-1] [Medline: 35236334]
8. Sydora BC, Whelan LJ, Abelseth B, Brar G, Idris S, Zhao R, et al. Identification of presurgical risk factors for the development of chronic postsurgical pain in adults: a comprehensive umbrella review. J Pain Res 2024 Jul 30;17:2511-2530 [FREE Full text] [doi: 10.2147/JPR.S466731] [Medline: 39100136]
9. Clephas PR, Hoeks SE, Singh PM, Guay CS, Trivella M, Klimek M, et al. Prognostic factors for chronic post-surgical pain after lung and pleural surgery: a systematic review with meta-analysis, meta-regression and trial sequential analysis. Anaesthesia 2023 Aug;78(8):1005-1019 [FREE Full text] [doi: 10.1111/anae.16009] [Medline: 37094792]
10. Lim J, Chen D, McNicol E, Sharma L, Varaday G, Sharma A, et al. Risk factors for persistent pain after breast and thoracic surgeries: a systematic literature review and meta-analysis. Pain 2022 Jan 01;163(1):3-20. [doi: 10.1097/j.pain.0000000000002301] [Medline: 34001769]
11. Montes A, Roca G, Cantillo J, Sabate S, GENDOLCAT Study Group. Presurgical risk model for chronic postsurgical pain based on 6 clinical predictors: a prospective external validation. Pain 2020 Nov;161(11):2611-2618. [doi: 10.1097/j.pain.0000000000001945] [Medline: 32541391]
12. Vuckovic A, Gallardo VJ, Jarjees M, Fraser M, Purcell M. Prediction of central neuropathic pain in spinal cord injury based on EEG classifier. Clin Neurophysiol 2018 Aug;129(8):1605-1617 [FREE Full text] [doi: 10.1016/j.clinph.2018.04.750] [Medline: 29886266]
13. Benussi A, Dell'Era V, Cantoni V, Cotelli MS, Cosseddu M, Spallazzi M, et al. TMS for staging and predicting functional decline in frontotemporal dementia. Brain Stimul 2020;13(2):386-392 [FREE Full text] [doi: 10.1016/j.brs.2019.11.009] [Medline: 31787557]
14. Richebé P, Capdevila X, Rivat C. Persistent postsurgical pain: pathophysiology and preventative pharmacologic considerations. Anesthesiology 2018 Sep;129(3):590-607. [doi: 10.1097/ALN.0000000000002238] [Medline: 29738328]
15. Pacheco-Barrios K, Lima D, Pimenta D, Slawka E, Navarro-Flores A, Parente J, et al. Motor cortex inhibition as a fibromyalgia biomarker: a meta-analysis of transcranial magnetic stimulation studies. Brain Netw Modul 2022;1(2):88-101 [FREE Full text] [doi: 10.4103/2773-2398.348254] [Medline: 35845034]
16. Caumo W, Deitos A, Carvalho S, Leite J, Carvalho F, Dussán-Sarria JA, et al. Motor cortex excitability and BDNF levels in chronic musculoskeletal pain according to structural pathology. Front Hum Neurosci 2016 Jul 15;10:357 [FREE Full text] [doi: 10.3389/fnhum.2016.00357] [Medline: 27471458]
17. Chiang MC, Hsueh HW, Yeh TY, Cheng YY, Kao YH, Chang KC, et al. Maladaptive motor cortical excitability and connectivity in polyneuropathy with neuropathic pain. Eur J Neurol 2022 May;29(5):1465-1476. [doi: 10.1111/ene.15247] [Medline: 35020255]
18. Lefaucheur JP, Drouot X, Ménard-Lefaucheur I, Keravel Y, Nguyen JP. Motor cortex rTMS restores defective intracortical inhibition in chronic neuropathic pain. Neurology 2006 Nov 14;67(9):1568-1574. [doi: 10.1212/01.wnl.0000242731.10074.3c] [Medline: 17101886]
19. Parker RS, Lewis GN, Rice DA, McNair PJ. Is motor cortical excitability altered in people with chronic pain? A systematic review and meta-analysis. Brain Stimul 2016;9(4):488-500. [doi: 10.1016/j.brs.2016.03.020] [Medline: 27133804]
20. Bunse T, Wobrock T, Strube W, Padberg F, Palm U, Falkai P, et al. Motor cortical excitability assessed by transcranial magnetic stimulation in psychiatric disorders: a systematic review. Brain Stimul 2014;7(2):158-169. [doi: 10.1016/j.brs.2013.08.009] [Medline: 24472621]
21. Bhandari A, Radhu N, Farzan F, Mulsant BH, Rajji TK, Daskalakis ZJ, et al. A meta-analysis of the effects of aging on motor cortex neurophysiology assessed by transcranial magnetic stimulation. Clin Neurophysiol 2016 Aug;127(8):2834-2845 [FREE Full text] [doi: 10.1016/j.clinph.2016.05.363] [Medline: 27417060]
22. Pinheiro ES, de Queirós FC, Montoya P, Santos CL, do Nascimento MA, Ito CH, et al. Electroencephalographic patterns in chronic pain: a systematic review of the literature. PLoS One 2016 Feb 25;11(2):e0149085 [FREE Full text] [doi: 10.1371/journal.pone.0149085] [Medline: 26914356]
23. Stern J, Jeanmonod D, Sarnthein J. Persistent EEG overactivation in the cortical pain matrix of neurogenic pain patients. Neuroimage 2006 Jun;31(2):721-731. [doi: 10.1016/j.neuroimage.2005.12.042] [Medline: 16527493]
24. Sarnthein J, Stern J, Aufenberg C, Rousson V, Jeanmonod D. Increased EEG power and slowed dominant frequency in patients with neurogenic pain. Brain 2006 Jan;129(Pt 1):55-64 [FREE Full text] [doi: 10.1093/brain/awh631] [Medline: 16183660]

XSL•FO
RenderX

25. Vuckovic A, Hasan MA, Fraser M, Conway BA, Nasseroleslami B, Allan DB. Dynamic oscillatory signatures of central neuropathic pain in spinal cord injury. J Pain 2014 Jun;15(6):645-655 [FREE Full text] [doi: 10.1016/j.jpain.2014.02.005] [Medline: 24589821]
26. Fallon N, Chiu Y, Nurmikko T, Stancak A. Altered theta oscillations in resting EEG of fibromyalgia syndrome patients. Eur J Pain 2018 Jan;22(1):49-57 [FREE Full text] [doi: 10.1002/ejp.1076] [Medline: 28758313]
27. Zebhauser PT, Hohn VD, Ploner M. Resting-state electroencephalography and magnetoencephalography as biomarkers of chronic pain: a systematic review. Pain 2023 Jun 01;164(6):1200-1221 [FREE Full text] [doi: 10.1097/j.pain.0000000000002825] [Medline: 36409624]
28. Marques LM, Castellani A, Barbosa SP, Imamura M, Battistella LR, Simis M, et al. Neuroplasticity changes in knee osteoarthritis (KOA) indexed by event-related desynchronization/synchronization during a motor inhibition task. Somatosens Mot Res 2024 Sep;41(3):149-158. [doi: 10.1080/08990220.2023.2188926] [Medline: 36921090]
29. Akutay S, Ceyhan Ö. The relationship between fear of surgery and affecting factors in surgical patients. Perioper Med (Lond) 2023 Jun 09;12(1):22 [FREE Full text] [doi: 10.1186/s13741-023-00316-0] [Medline: 37296486]
30. Ahmadi N, Mbuagbaw L, Finley C, Agzarian J, Hanna WC, Shargall Y. Impact of the integrated comprehensive care program post-thoracic surgery: a propensity score-matched study. J Thorac Cardiovasc Surg 2021 Jul;162(1):321-30.e1 [FREE Full text] [doi: 10.1016/j.jtcvs.2020.05.095] [Medline: 32713635]
31. Ader DN. Developing the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care 2007 May;45(5):S1-S2 [FREE Full text] [doi: 10.1097/01.mlr.0000260537.45076.74]
32. PROMIS. HealthMeasures. URL: https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis [accessed 2025-11-28]
33. Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
34. Kroenke K, Spitzer RL, Williams JB, Monahan PO, Löwe B. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. Ann Intern Med 2007 Mar 06;146(5):317-325. [doi: 10.7326/0003-4819-146-5-200703060-00004] [Medline: 17339617]
35. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
36. Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. J Affect Disord 2010 Apr;122(1-2):86-95. [doi: 10.1016/j.jad.2009.06.019] [Medline: 19616305]
37. Rossini PM, Burke D, Chen R, Cohen LG, Daskalakis Z, Di Iorio R, et al. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. Committee. Clin Neurophysiol 2015 Jun;126(6):1071-1107 [FREE Full text] [doi: 10.1016/j.clinph.2015.02.001] [Medline: 25797650]
38. Wildgaard K, Ringsted TK, Hansen HJ, Petersen RH, Werner MU, Kehlet H. Quantitative sensory testing of persistent pain after video-assisted thoracic surgery lobectomy. Br J Anaesth 2012 Jan;108(1):126-133 [FREE Full text] [doi: 10.1093/bja/aer325] [Medline: 21980121]
39. Wang G, He M, Ji XL, Wang XL, Feng Y. Identifying patients at high risk of chronic pain after video-assisted thoracoscopic surgery using thermal quantitative sensory testing. J Cardiothorac Vasc Anesth 2022 Aug;36(8):2406-2411. [doi: 10.1053/j.jvca.2022.03.015] [Medline: 35568656]
40. Harrington C, Molena D. Minimally invasive Ivor Lewis esophagectomy in 10 steps. JTCVS Tech 2021 Aug 08;10:489-494 [FREE Full text] [doi: 10.1016/j.xjtc.2021.04.038] [Medline: 34977791]
41. Vuckovic A, Jajrees M, Purcell M, Berry H, Fraser M. Electroencephalographic predictors of neuropathic pain in subacute spinal cord injury. J Pain 2018 Nov;19(11):1256.e1-125.e17 [FREE Full text] [doi: 10.1016/j.jpain.2018.04.011] [Medline: 29751110]
42. Meirovitch Y, Harris H, Dayan E, Arieli A, Flash T. Alpha and beta band event-related desynchronization reflects kinematic regularities. J Neurosci 2015 Jan 28;35(4):1627-1637 [FREE Full text] [doi: 10.1523/JNEUROSCI.5371-13.2015] [Medline: 25632138]
43. Witham CL, Riddle CN, Baker MR, Baker SN. Contributions of descending and ascending pathways to corticomuscular coherence in humans. J Physiol 2011 Aug 01;589(15):3789-3800 [FREE Full text] [doi: 10.1113/jphysiol.2011.211045] [Medline: 21624970]
44. Chen BJ, Liu TY, Wu HC, Tsai MW, Wei SH, Chou LW. Effects of sling exercises on pain, function, and corticomuscular functional connectivity in individuals with chronic low back pain-preliminary study. PLoS One 2023 Nov 30;18(11):e0288405 [FREE Full text] [doi: 10.1371/journal.pone.0288405] [Medline: 38032998]
45. Chowdhury A, Raza H, Meena YK, Dutta A, Prasad G. An EEG-EMG correlation-based brain-computer interface for hand orthosis supported neuro-rehabilitation. J Neurosci Methods 2019 Jan 15;312:1-11. [doi: 10.1016/j.jneumeth.2018.11.010] [Medline: 30452976]
46. Sfreddo HJ, Wecht JR, Alsalman OA, Wu YK, Harel NY. Duration and reliability of the silent period in individuals with spinal cord injury. Spinal Cord 2021 Aug;59(8):885-893. [doi: 10.1038/s41393-021-00649-x] [Medline: 34099882]

XSL•FO
RenderX

47. Kamen G. Reliability of motor-evoked potentials during resting and active contraction conditions. Med Sci Sports Exerc 2004 Sep;36(9):1574-1579. [doi: 10.1249/01.mss.0000139804.02576.6a] [Medline: 15354040]
48. Maeda F, Gangitano M, Thall M, Pascual-Leone A. Inter- and intra-individual variability of paired-pulse curves with transcranial magnetic stimulation (TMS). Clin Neurophysiol 2002 Mar;113(3):376-382. [doi: 10.1016/s1388-2457(02)00008-1] [Medline: 11897538]
49. Popov T, Tröndle M, Baranczuk-Turska Z, Pfeiffer C, Haufe S, Langer N. Test-retest reliability of resting-state EEG in young and older adults. Psychophysiology 2023 Jul;60(7):e14268. [doi: 10.1111/psyp.14268] [Medline: 36894751]
50. Pohja M, Salenius S, Hari R. Reproducibility of cortex-muscle coherence. Neuroimage 2005 Jul 01;26(3):764-770. [doi: 10.1016/j.neuroimage.2005.02.031] [Medline: 15955485]
51. Espenhahn S, de Berker AO, van Wijk BC, Rossiter HE, Ward NS. Movement-related beta oscillations show high intra-individual reliability. Neuroimage 2017 Feb 15;147:175-185 [FREE Full text] [doi: 10.1016/j.neuroimage.2016.12.025] [Medline: 27965146]
52. Shanthanna H, Wang L, Paul J, Lovrics P, Devereaux PJ, Bhandari M, et al. A prospective cohort study of chronic postsurgical pain after ambulatory surgeries. Curr Med Res Opin 2024 Jul;40(7):1187-1193. [doi: 10.1080/03007995.2024.2360128] [Medline: 38809229]
53. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 06;10:1 [FREE Full text] [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
54. Barr MS, Farzan F, Davis KD, Fitzgerald PB, Daskalakis ZJ. Measuring GABAergic inhibitory activity with TMS-EEG and its potential clinical application for chronic pain. J Neuroimmune Pharmacol 2013 Jun;8(3):535-546. [doi: 10.1007/s11481-012-9383-y] [Medline: 22744222]
55. Krause P, Foerderreuther S, Straube A. Effects of conditioning peripheral repetitive magnetic stimulation in patients with complex regional pain syndrome. Neurol Res 2005 Jun;27(4):412-417. [doi: 10.1179/016164105X17224] [Medline: 15949240]
56. Cardinal TM, Antunes LC, Brietzke AP, Parizotti CS, Carvalho F, De Souza A, et al. Differential neuroplastic changes in fibromyalgia and depression indexed by up-regulation of motor cortex inhibition and disinhibition of the descending pain system: an exploratory study. Front Hum Neurosci 2019;13:138 [FREE Full text] [doi: 10.3389/fnhum.2019.00138] [Medline: 31105542]
57. Barbosa LM, Valerio F, da Silva VA, Rodrigues AL, Galhardoni R, Yeng LT, et al. Corticomotor excitability is altered in central neuropathic pain compared with non-neuropathic pain or pain-free patients. Neurophysiol Clin 2023 Jun;53(3):102845 [FREE Full text] [doi: 10.1016/j.neucli.2023.102845] [Medline: 36822032]

Abbreviations

CMC: corticomuscular coherence
CPSP: chronic postsurgical pain
CS: conditioning stimulus
CSP: cortical silent period
EC: eyes closed
EEG: electroencephalography
EMG: electromyography
EO: eyes open
ERD: event-related desynchronization
FDI: first dorsal interosseous
MDT: mechanical detection threshold
MEP: motor evoked potential
MIE: minimally invasive esophagectomy
MVC: maximum voluntary contraction
PPT: pressure pain threshold
QST: quantitative sensory testing
SICI: short-latency intracortical inhibition
SJH: St Joseph's Healthcare
TMS: transcranial magnetic stimulation
TS: test stimulus

XSL•FO
RenderX

Edited by J Sarvestan; submitted 21.Jul.2025; peer-reviewed by P O'Scanaill; comments to author 29.Sep.2025; accepted 19.Nov.2025; published 12.Jan.2026.

Please cite as:
Hussain M, Foglia SD, Hamoodi A, Ma J, Agzarian J, Finley CJ, Shargall Y, Nelson AJ, Shanthanna H
Exploring Risk Factors and Neurophysiological Mechanisms Underlying the Development of Chronic Postsurgical Pain After Thoracic Surgery: Protocol for an Observational Feasibility Study
JMIR Res Protoc 2026;15:e81042
URL: https://www.researchprotocols.org/2026/1/e81042
doi:10.2196/81042
PMID:

©Mustaali Hussain, Stevie Dante Foglia, Ameer Hamoodi, Jinhui Ma, John Agzarian, Christian John Finley, Yaron Shargall, Aimee Jennifer Nelson, Harsha Shanthanna. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 12.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Retrieving Transgender and Gender Diverse Literature: Protocol for the Development and Validation of 2 Search Hedges

Chi Dinh[1], BHSc; Zack Marshall[1,2,3], MSW, PhD; Scott Marsalis[4], BA, MLiS; Ryn Gagen[5], BA, MLiS; Preet Kang[6], BA; Avery Everhart[7], BA, MA, PhD

[1]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[2]School of Social Work, McGill University, Montreal, QC, Canada

[3]O'Brien Institute for Public Health, University of Calgary, Calgary, AB, Canada

[4]Sciences, Agriculture & Engineering, University of Minnesota Libraries, University of Minnesota, Saint Paul, MN, United States

[5]Health Sciences Library, University of Minnesota, Saint Paul, MN, United States

[6]Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada

[7]Department of Geography, Faculty of Arts, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Zack Marshall, MSW, PhD
Department of Community Health Sciences
Cumming School of Medicine
University of Calgary
3280 Hospital Dr NW
Calgary, AB, T2N 4Z6
Canada
Phone: 1 4032206940
Email: zack.marshall@ucalgary.ca

Abstract

Background: Searching for transgender and gender diverse (TGD) references within large academic databases can be a challenging process, partly due to the dynamic and diverse definitions of words and terminologies used by multiple interest holders. Search hedges are preestablished search strings that aid in the efficacy of identifying and screening relevant articles. Validated search hedges focused on TGD people and topics will aid in identifying relevant literature.

Objective: This study aims to develop and validate the sensitivity and precision of 2 interdisciplinary and cross-cultural TGD search hedges designed for retrieving references from MEDLINE and APA PsycInfo, both on the Ovid platform.

Methods: Searches were conducted using the finalized search hedges via Ovid on June 7, 2024, yielding 31,055 references from MEDLINE and 22,924 references from APA PsycInfo. A random sample of 2330 records from MEDLINE and 2293 records from APA PsycInfo will be independently screened by at least 2 team members. At the title and abstract screening stage, references will be excluded if they (1) use solely binary terminology to describe gender, (2) focus on psychometric measurement of gender, or (3) focus on intersex or differences of sex development (DSD) topics. References will be included if they (1) report on transgender or gender diverse people, or both, in their sample; or (2) specifically discuss TGD communities or TGD topics. References without an abstract will be categorized as *No_Abstract.* References in which the TGD population is unclear will be categorized as *LGB_Maybe_T* or *Mixed_Topics.* Only references in the *No_Abstract*, *LGB_Maybe_T*, or *Mixed_Topics* categories will proceed to the full-text screening phase. In the full-text screening phase, references will be categorized as included if they (1) clearly distinguish between sexual identity and gender identity, (2) mention or discuss TGD topics or experiences in the Methods or Results sections, (3) communicate consideration for participants' gender self-identification and experiences, or (4) consider TGD populations as a distinct subpopulation. The results of the screening process will be used to calculate precision and sensitivity, with a targeted sensitivity of 100% and a targeted precision of 76% for each search hedge.

Results: Validation and data analysis are projected to be finished by December 2025, with results expected to be published in 2026.

Conclusions: Rigorous and transparent knowledge synthesis processes, starting with a high-quality search hedge, can help inform and equip community members, clinicians, policymakers, and other key decision-makers with scientifically sound evidence.

International Registered Report Identifier (IRRID): DERR1-10.2196/76055

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e76055) doi:10.2196/76055

KEYWORDS

evidence synthesis; database searching; search hedge validation; search filter validation; Two-Spirit; transgender; nonbinary; gender minority; information justice

Introduction

Background

Knowledge synthesis, also called evidence synthesis, is the process of identifying, evaluating, and critically integrating all available evidence on a topic [1]. This process is a foundational practice for the development of evidence-based decision-making [1]. Effective and comprehensive search strategies are one of the core components of knowledge synthesis [2]. To support robust search strategies, search hedges are increasingly used. A search hedge consists of pre-established search strings designed to comprehensively retrieve references in a topic area from one or multiple bibliographic databases [3]. A search hedge is typically developed by librarians working with subject matter experts and should be validated for its performance [4]. Validated search hedges can help ensure all relevant literature is retrieved, reduce the number of irrelevant articles screened, and increase the efficiency of resource use in the knowledge synthesis process.

In addition to hedges with a specific subject matter focus, more recent publications have emphasized other search attributes, such as geographic, methodological, or population-specific searches [5-7]. While search hedges focused on combined sexual and gender minority groups are available, there are currently no validated and published search hedges that focus on transgender and gender diverse (TGD) communities and topics [8,9]. In this protocol, we use the term *TGD* to describe people who do not identify with their assigned gender at birth. This definition has the potential to also describe intersex people and people with differences of sex development (DSDs), and there are indeed certain overlapping experiences between TGD and intersex people. However, we follow the recommendations of organizations such as interACT: Advocates for Intersex Youth who make plain that being intersex and being TGD are distinct categories of identity and experience. Thus, we will not include articles studying DSDs or intersex communities unless it is clear within the article that the study sample includes people who are both intersex and transgender to respect these identities and experiences as distinct from those of TGD communities.

Additionally, we recognize that *TGD* can be understood as an anglophone export or a category imposed onto autochthonous sex and gender systems that have been devalued, erased, or even criminalized through imperialism and settler colonialism [10,11]. Thus, our search hedges incorporate terminology beyond that centered within the medical literature of the anglophone world and Global North. This decision to incorporate culturally specific terms, while intended to be inclusive, also runs the risk of reproducing the centrality of not only English-language terms but also medicalized models of understanding gender nonconformity and gender diversity.

TGD people face systematic marginalization and oppression unique from those of the cisgender heterosexual population and their sexual minority peers [12,13]. Misconceptions and disinformation regarding the experiences, quality of life, and health of TGD people further fuel the discrimination against this population [14-16]. Rigorous and transparent knowledge synthesis processes, starting with high-quality search hedges, can help inform and equip community members, clinicians, policymakers, and other key decision-makers with scientifically sound evidence.

Validation Structure

Overview

Developing structured search strategies focused on TGD populations is complex for several reasons, and our nonstandard screening process was developed to address these complexities. In addition to the *Include* and *Exclude* categories in a standard literature screening process, 2 additional categories were created for the title and abstract screening phase: *Mixed_Topics* and *LGB_Maybe_T*. The purpose of the *Mixed_Topics* category is to differentiate articles related to TGD people and topics from those related to cisgender people and topics. The purpose of the *LGB_Maybe_T* category is to differentiate articles on gender identity topics relevant to TGD people from those on Two-Spirit, lesbian, gay, bisexual, transgender, queer, intersex, asexual, or additional sexual and gender diverse (2SLGBTQIA+) topics.

Mixed_Topics Category

One challenge in identifying and retrieving literature relevant to TGD populations is the variations in context and usage of polysemic terms such as *gender diversity* or *gender inclusivity* (for a complete list of polysemic terms relevant to these hedges, see Multimedia Appendix 1). For example, the term *gender inclusivity* can be used to describe the inclusion of TGD people in historically underrepresented fields, but it can also be used to refer to the inclusion of cisgender women in fields and occupations historically dominated by men. Due to this usage in both cisgender and TGD contexts, inclusion of these terms in a search hedge has the potential to increase the number of nonrelevant references being retrieved, imposing additional burden on the subsequent literature screening process. However, exclusion of these terms risks compromising the comprehensiveness of the search hedges. Thus, the category *Mixed_Topics* was created to evaluate the accuracy and usefulness of including these polysemic terms.

LGB_Maybe_T Category

An additional challenge with searching for TGD literature is the conflation of gender identity and sexual orientation: authors often report gender identity inconsistently, if at all [17-19]. This lack of specificity means that gender minority populations are often homogenized within the broader 2SLGBTQIA+ acronym without differentiating the experiences of gender minority

XSL•FO
RenderX

groups as distinct from those of sexual minority groups. While sexual and gender minority groups share close social and political struggles, and TGD people can also have a sexual minority identity (eg, a nonbinary bisexual individual), people with gender minority identities often experience unique forms of discrimination stemming from cisnormativity [20]. Therefore, articles that are relevant to 2SLGBTQIA+ populations are not guaranteed to be relevant to TGD populations, and articles that only use the term 2SLGBTQIA+ (or its variations) in their titles and abstracts do not provide sufficient information to determine their relevancy to TGD people and topics. The category *LGB_Maybe_T* was created to (1) confirm the relevance of articles using the acronym 2SLGBTQIA+ to the TGD population and (2) analyze patterns of usage of TGD terms in 2SLGBTQIA+ research.

Precision and Sensitivity

Precision and sensitivity are 2 measurements commonly used in the validation process of search hedges and together offer quantifiable reference points in the performance of a hedge. Precision is the ability of a search hedge to minimize the retrieval of nonrelevant articles [21,22]. Precision is calculated as the proportion of relevant over total retrieved articles [22]. Sensitivity, also called relative recall, is the ability of a search hedge to retrieve all relevant articles [21,22]. Sensitivity is calculated as the proportion of articles from a gold-standard set that are retrieved by a search hedge [22]. The most common method of establishing a gold-standard set for the validation of a search hedge is through manually searching a journal, a set of journals, or a database [22]. However, the main disadvantage of this method is its high time consumption and low article yield [22]. Another approach to finding a gold standard is using a previously assembled collection [22]. In this study, the gold-standard set consists of 144 TGD articles from Knowsy (search year: 2019; for a full list of articles in the gold-standard set, see Multimedia Appendix 2). Knowsy is an online dataset of scoping reviews, systematic reviews, and structured reviews focused on 2SLGBTQIA+ topics [23,24]. Articles in Knowsy were identified by systematically searching 4 academic databases (MEDLINE, APA PsycInfo, Embase, and CINAHL) and selected in accordance with the standard 2-step screening process performed by 2 independent, trained reviewers [23,24]. Each reference in Knowsy has been labeled with standardized category tags for topics, populations, and type of knowledge synthesis by a trained reviewer, with independent validation by a second reviewer [23,24].

Special consideration also needs to be paid to ensure ethical and responsible research when working with a marginalized community, such as gender minority communities. Accountability and transparency in the search and research process are 2 key components to ensure that researchers do not further alienate already marginalized people. One of the best practices to accomplish these goals is the publication of a protocol at the start of the research process [25]. A protocol offers detailed background information on the search and screening process and increases the reproducibility of the results. It also provides an accountability mechanism as deviations from the protocol need to be identified when reporting study findings. As validations of search hedges are still uncommon, protocols for the validation of these hedges are even more limited [26-28]. The scant literature regarding search hedge validation methods could further discourage this type of research from being conducted as lack of established methods can increase the amount of time and resources dedicated to this process. Thus, publishing a search hedge validation protocol that includes details about the methods used aids in knowledge and resource sharing, as well as facilitating the peer review process and increasing research rigor. In this paper, we present a protocol for the validation of 2 TGD search hedges using a novel validation structure.

Objectives

This protocol describes the methods for a search hedge validation. The objectives are to (1) validate 2 interdisciplinary and cross-cultural TGD search hedges designed for retrieving references in MEDLINE and APA PsycInfo, (2) determine the usage pattern of TGD polysemic terms in the literature, and (3) assess the usage pattern of TGD terms in 2SLGBTQIA+ literature.

Methods

Ethical Considerations

Ethics approval will not be sought for this study because it will only include analysis of published articles. This study will not involve human participants, medical records, patient information, observations of public behaviors, or secondary data analyses.

Search Hedge Design

To develop and validate the 2 search hedges, we followed a 2-stage process: a preliminary search and a finalized search.

Stage 1: Preliminary Search Hedges

Two preliminary search hedges were developed in 2023 by 2 librarians with the goal of obtaining new insights into the effectiveness of the searches and inform sample size calculations for the full validation study. The MEDLINE search produced 49,796 results, and the APA PsycInfo search produced 38,344 references. References were imported into EPPI-Reviewer (EPPI-Centre) and randomly allocated into screening groups [29]. Two team members then screened 0.5% (239/49,796) of the references from MEDLINE and 0.5% (192/38,344) of the references from APA PsycInfo.

In terms of precision and sensitivity, the preliminary search hedges identified 37.7% (90/239) of articles from MEDLINE and 37.5% (72/192) of articles from APA PsycInfo that met the eligibility criteria, achieving a precision of 39% for both databases. The hedges retrieved all articles from the gold-standard set, achieving a 100% sensitivity rate. Additionally, 5.7% (11/192) of the records from APA PsycInfo were not available in full text.

Stage 2: Finalization of the Search Hedges

On the basis of the precision and sensitivity of the preliminary search hedges, the team refined the search strategies to reduce the number of nonrelevant results. On June 7, 2024, we conducted 2 searches using the finalized search hedges, which

XSL•FO
RenderX

produced 31,055 references from MEDLINE and 22,924 references from APA PsycInfo (Multimedia Appendix 3).

Sample Size Calculations

Informed by the preliminary results of our pilot hedges, we assume that a similar proportion of sampled records will meet the inclusion criteria (approximately 39%) and expect approximately 10% of records to be unretrievable (although our pilot search in APA PsycInfo only showed 6% of records to be unretrievable). For the MEDLINE hedge, a sample size of 2330 will achieve a targeted sensitivity of 100% (99% CI 99.4%-100%). For the APA PsycInfo hedge, a sample size of 2293 will achieve a targeted sensitivity of 100% (99% CI 99.3%-100%).

The calculation of precision for each search hedge requires the specificity rate. Specificity is defined as the proportion of nonrelevant articles correctly not retrieved by a search hedge [21]. In the context of this project, to calculate specificity, all articles indexed in MEDLINE and APA PsycInfo would need to be screened. This is not a feasible option. Thus, specificity will not be calculated. However, we opted for a targeted specificity of 80% (99% CI 77.2%-83%) for MEDLINE and a targeted specificity of 80% (99% CI 77.1%-83%) for APA PsycInfo. These targeted sensitivity and specificity levels correspond to a precision of 76% for both hedges.

Search Hedge Validation

Title and Abstract Eligibility Criteria

The goal of the title and abstract screening process is to identify references with TGD-relevant content. When screening, team members will select 1 of 5 categories for each reference: *Exclude*, *No_Abstract*, *Mixed_Topics*, *LGB_Maybe_T*, and *Include*.

Articles categorized as *Include* or *Exclude* in the screening based on title and abstract phase will not undergo full-text review because each category already demonstrates sufficient data on the relevance or nonrelevance to TGD topic areas, respectively. Abstracts categorized as *No_Abstract*, *Mixed_Topics*, or *LGB_Maybe_T* have insufficient information to determine their relevance and, therefore, will proceed to the full-text screening phase (Figure 1). We will incorporate rates at which references are labeled as *Mixed_Topics* and *LGB_Maybe_T* at the title and abstract screening phase and the rates at which those references with these specific labels are ultimately included after full-text review. Tracking these rates will potentially enable us to derive further semantic clarity on search terms within the hedges, as well as provide recommendations on more targeted use of terminology in future peer-reviewed literature.

Figure 1. Flow chart depicting the screening process for references relevant to trans and gender diverse (TGD) subject topics from MEDLINE. A similar process was used for results retrieved from APA PsycInfo. Only references categorized as *No_Abstract*, *Mixed_Topics*, or *LGB_Maybe_T* will proceed to the full-text screening phase to validate their relevance to TGD subject topics.

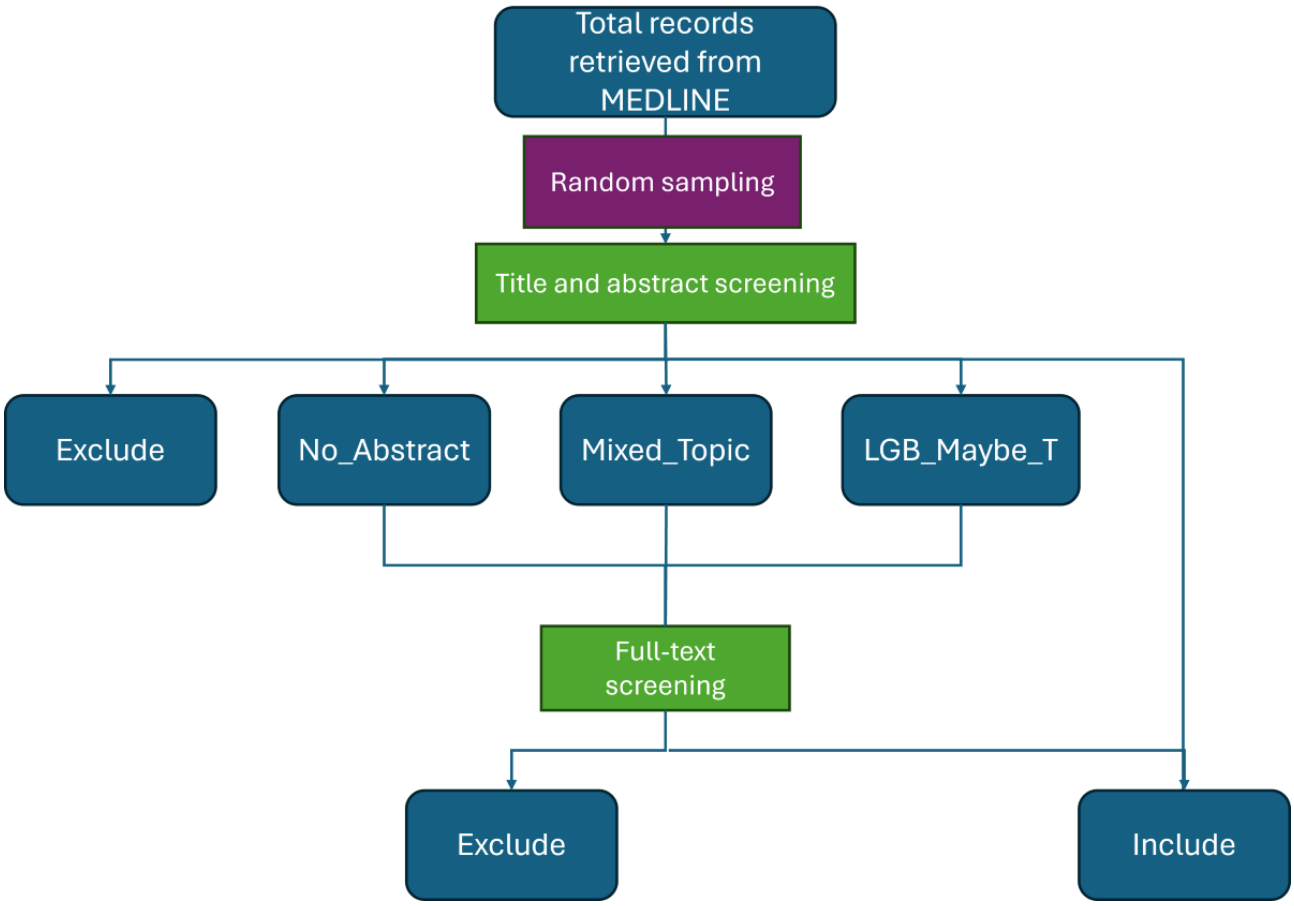

XSL•FO
RenderX

Exclude Category

References will be excluded if they (1) use solely binary terminology to describe gender (eg, boys and girls, men and women, *both genders*, or *opposite gender*) or (2) focus on psychometric measurement of gender without reporting participants' self-identification or narrative description (eg, Bem Sex-Role Inventory, including sex typing using feminine, masculine, and androgynous categories) [30]. Abstracts focusing on intersex or DSD topics will also be excluded.

Include Category

References will be included if they (1) report on TGD people, or both, in their sample; or (2) specifically discuss TGD communities or TGD topics, such as gender-affirming health care, gender expression, or drag performance. References containing historical terminology referring to Two-Spirit or TGD people, as well as Indigenous and non-Eurocentric gender-expansive terminology, will be included.

No_Abstract Category

References without abstracts will be categorized as *No_Abstract* unless their titles clearly indicate relevancy to TGD populations (eg, using the term *transgender* in the title), in which case they will be categorized as *Include*.

LGB_Maybe_T or Mixed_Topics Category

For abstracts with unclear subject matter relevant to TGD topic areas, 2 coding options will be available. Abstracts with an unclear TGD or 2SLGBTQIA+ population or TGD or 2SLGBTQIA+–related study topic will be categorized as *Mixed_Topics* (eg, using the term *gender inclusivity* but with no further clarification of whether this means inclusion of TGD people or only cisgender women). If abstracts use the term *transgender* or *trans* solely associated with or as part of elaborating on sexual and gender identity acronyms (eg, *LGBTQ+* or *2SLGBTQIA+*), these references will be categorized as *LGB_Maybe_T.* Abstracts in the *Mixed_Topics* category will be presumed to relate to the cisgender population with unclear relevance to TGD people and topics until further confirmation via the full text. Abstracts in the *LGB_Maybe_T* category will be presumed to relate to the 2SLGBTQIA+ population with unclear relevance to TGD people and topics until further confirmation via the full text.

Full-Text Eligibility Criteria

During full-text screening, articles will be categorized as *Include* or *Exclude*. Similar to the screening based on title and abstract phase, the objective is to accurately identify whether the article is relevant to TGD subject matter.

Articles that are included will (1) clearly distinguish among sex, sexual identity, and gender identity; (2) communicate consideration for participants' gender self-identification and experiences (eg, self-declaration options in a demographic survey); or (3) consider the TGD population as a distinct subpopulation within the 2SLGBTQIA+ umbrella (eg, report on transgender or gender diverse participants separately in a demographic section or as part of the data analysis strategy). For primary research papers, these considerations should be reflected in the Methods or Results sections.

Information Sources

No restrictions will be placed on the type of reference, publication date, or language. All efforts will be made to translate articles that are not in English, French, Spanish, or Mandarin using Google Translate or DeepL Translator (DeepL SE), an online translation software [31]. Additional professional translation services will also be sought for articles in languages that are not stated above.

Study Records

Data Management

Results from the MEDLINE and APA PsycInfo searches will be imported to the EPPI-Reviewer online systematic review software [29]. Deduplication will not be conducted as validation and data analysis will be performed separately for each database. Random samples will be allocated using the *Create reference groups* feature in EPPI-Reviewer. Full-text references will be manually or automatically retrieved via Zotero (Corporation for Digital Scholarship), a citation manager, and uploaded to EPPI-Reviewer. EPPI-Reviewer will also be used to track and manage full text retrieval.

Selection Process

The review team includes a total of 6 members with lived experience in the TGD community. To ensure intercoder agreement, team member training will consist of a prescreening walkthrough of the eligibility criteria followed by trial coding until the disagreement rate between all coders is equal to or below the recommended 20% [32]. An existing internally circulated team notebook will be created at the start of the screening process to record notable discussions and decisions throughout the process. This notebook will be available to all team members to ensure consistency between coders. Two team members will independently screen each randomly sampled title, abstract, and full text based on the predefined eligibility criteria. Any disagreements will be addressed through discussion facilitated by a third team member, with at least one of the original screening coders present to ensure clarity and context. If differences cannot be resolved through discussion, the project team leads will make the final decision.

Data Analysis

Intercoder Agreement Calculation

The intercoder agreement in each screening phase will be calculated using the following formula [32]:



Data Items and Planned Statistical Tests

Precision and sensitivity will be calculated to evaluate the validity of each search hedge. Precision quantifies the ability of a search hedge to identify relevant TGD articles. This measurement will be calculated as the number of relevant articles retrieved divided by the total number of screened articles [21]:



XSL•FO
RenderX

Sensitivity assesses the ability of a search hedge to comprehensively retrieve all relevant articles [21]. Sensitivity will be calculated as the proportion of articles successfully retrieved by the search hedge compared to the gold-standard set as predefined from Knowsy:



Search Term Insights

Polysemic terms identified from the finalized search with unclear relevancy to TGD topics will be assessed. A verbatim search of each term will be conducted in EPPI-Reviewer, and accuracy will be calculated as a proportion of the articles in each category (*Mixed_Topics*, *LGB_Maybe_T*, *Exclude*, and *Include* from both the title and abstract and full-text screening) over the total number of articles containing said term. A higher proportion from the *Mixed_Topics* or *LGB_Maybe_T* category suggests that a term is not being used in a TGD-specific context. A higher proportion from the *Exclude* category suggests that a term does not have high relevance to TGD topics:



Results

As of November 2025, we have finished screening all 2330 articles from MEDLINE and all 2293 articles from APA PsycInfo. Data analysis is projected to be finished by December 2025. The results are expected to be published in 2026.

Discussion

Expected Findings

While a number of validated and unvalidated search hedges focused on 2SLGBTQIA+ populations exist, to the best of our knowledge, this protocol is the first to detail the process of evaluating 2 search hedges that exclusively focus on the TGD population for MEDLINE and APA PsycInfo [8,9,33,34]. In this paper, we developed a novel screening structure for terminologies with dynamic meanings, with full-text screening only for references where further confirmation was required. We anticipate that both the MEDLINE and APA PsycInfo search hedges will achieve a similar sensitivity of 100% to that of their pilot hedges. We also anticipate that the hedges will achieve higher precision than their pilot counterparts due to the fine-tuning conducted in between. The precision, sensitivity, and search term insight results will provide guidance for librarians and researchers to customize the hedges to best fit their research questions.

While various published and validated search filters exist, most of these hedges focus on retrieving articles based on geographic locations, type of study, and treatment [5,26,27,35]. Validated search hedges focusing on populations, specifically marginalized communities, are particularly scant. Schilperoort et al [34] documented the development and validation process of search filters for PubMed focusing on queer women. Wafford et al [36] documented the development and validation process of a search filter for PubMed focusing on immigrant populations. Both of these population filters relied on the relative recall method to create their respective gold-standard sets. Relative recall is the process of retrieving structured reviews, such as systematic reviews, scoping reviews, and evidence gap maps, related to the specified population and searching through the citations of each review to further identify relevant articles. Schilperoort et al [34] used an internal validation approach in which the gold-standard set was used to generate the search terms as well as the standard for testing the sensitivity of the filters. On the other hand, Wafford et al [36] used an external validation approach in which the team followed the same process of relative recall but partitioned the retrieved articles into 2 separate sets: a development set from which to derive search terms and a validation set (occupying the same role of a gold standard) to test the filter. External validation offers more rigorous results as it mimics a testing environment closer to how a filter will be deployed in a real-world setting [35]. In this protocol, we are able to use an external validation approach due to the existence of the TGD subset in Knowsy. The Knowsy dataset is entirely independent in development and screening to the 2 current search hedges, allowing for a more stringent validation process.

Our validation method further diverged from the methods used by Schilperoort et al [34] and Wafford et al [36]. Both groups validated the precision of their filters against known sets established through the relative recall process, whereas the existence of the Knowsy dataset as our gold standard allows us to test and calculate the precisions and sensitivities of our hedges against unknown sets, offering a closer resemblance to real-world settings [34,36]. While Wafford et al [36] also used external validation, the team only accessed articles at the title and abstract level. For a large portion of our retrieved articles, the titles and abstracts do not provide sufficient information regarding our population of interest, a similar problem that Schilperoort et al [34] also encountered with a search filter focused on queer women. Thus, without the need to generate a gold-standard or validation set, our screening process can explore the semantics and context of terms through our *Mixed_Topics* and *LGB_Maybe_T* categories while minimizing the number of articles needing to be screened in full text.

Knowledge Mobilization Plan

Once validated, the search hedges will be submitted to the InterTASC Information Specialists' Sub-Group Search Filter Resource, the McMaster Health Knowledge Refinery, and the University of Alberta Library's health sciences search filter database [37-39]. We plan to periodically update the 2 hedges to keep abreast with indexing changes in their respective databases. Additionally, future plans will include development and adaptation of the hedges for other bibliographic databases, such as Embase, PubMed, CINAHL, and Scopus. Other knowledge mobilization plans include presentations at relevant research and knowledge synthesis conferences such as Moving Trans History Forward and the Global Evidence Summit [40,41].

Limitations and Considerations

One primary limitation of the hedges is the limited semantic search. This is related to the standard search mechanisms in MEDLINE, APA PsycInfo, and other databases. The default search mechanism for these databases is the keyword search,

XSL•FO
RenderX

which relies on exact match or known variants of the search strings [42]. For example, if a search string specifies the term *transgender*, the keyword search will retrieve only items with the exact match. Truncation operators can increase recall by looking for variants based on the word root (eg, plurals) but not semantic meaning. However, a keyword search does not take into account the context in which the term occurred, thus retrieving results with false semantic matches. In this example, articles with phrases such as *not transgender* would still be retrieved by the hedges. We anticipate that this search mechanism will decrease the precision rates of the hedges.

Another consideration when using the search hedges is the trade-off between sensitivity and precision. The primary user audience for the search hedges are information professionals supporting evidence syntheses, and thus, the hedges were designed with sensitivity as the top priority. The trade-off of this prioritization is precision, and certain types of reviews with time and resource constraints, such as rapid reviews, should keep this in mind. By examining specific search terms that are anticipated to retrieve more nonrelevant articles, we hope to provide librarians and other readers with additional information to make guided adaptations of the hedges that best fit individual research questions and purposes.

Conclusions

This protocol provides a blueprint for future studies to respectfully and comprehensively search literature related to self-described identities used by TGD communities. This protocol, along with other published articles of its kind, will contribute to the openness of science and the practice of sharing resources within the research community.

Acknowledgments

The authors would like to extend their heartfelt gratitude to Arthur Dave Miller (Sexual Health and Gender (SHaG) lab, Dalhousie University), Meng Wang (University of Calgary), and Atlas Davis (University of Calgary) for providing their valuable insights in developing the search hedges. This project received no direct grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: AE, CD, RG, SM, ZM
Data curation: RG, SM
Formal analysis: AE, CD
Investigation: AE, CD, RG, SM, ZM
Methodology: AE, CD, PK, RG, SM, ZM
Project administration: CD
Supervision: AE, CD, ZM
Validation: AE, CD
Visualization: CD
Writing—original draft: CD, ZM
Writing—review and editing: AE, CD, PK, RG, SM, ZM

Conflicts of Interest

The gold-standard dataset comes from Knowsy, a free online portal of 2SLGBTQIA+ focused systematic reviews, scoping studies, and evidence maps. The studies in the dataset are from the 2019 iteration of Knowsy and were identified from initial search strategies used to retrieve sexual and gender minority research. ZM cofounded Knowsy and now coleads this initiative with CD.

Multimedia Appendix 1
Lists of polysemic terms identified during pilot review.
[DOCX File , 1915 KB - resprot_v15i1e76055_app1.docx]

Multimedia Appendix 2
List of trans and gender diverse gold standard references from Knowsy (searched in 2019).
[DOCX File , 48 KB - resprot_v15i1e76055_app2.docx]

Multimedia Appendix 3
Trans search hedges in Ovid MEDLINE and APA PsycInfo.
[DOCX File , 2041 KB - resprot_v15i1e76055_app3.docx]

XSL•FO
RenderX

References

1. Grimshaw J. A knowledge synthesis chapter. Canadian Institutes of Health Research. URL: https://www.cihr.gc.ca/e/documents/knowledge_synthesis_chapter_e.pdf [accessed 2025-03-04]
2. Rethlefsen ML, Farrell AM, Osterhaus Trzasko LC, Brigham TJ. Librarian co-authors correlated with higher quality reported search strategies in general internal medicine systematic reviews. J Clin Epidemiol 2015 Jun;68(6):617-626. [doi: 10.1016/j.jclinepi.2014.11.025] [Medline: 25766056]
3. Campbell S. What is the difference between a filter and a hedge? J Eur Assoc Health Inf Libr 2016 Apr 01;12(1):4-5 [FREE Full text] [doi: 10.7939/R35T3GF4C]
4. Damarell RA, May N, Hammond S, Sladek RM, Tieman JJ. Topic search filters: a systematic scoping review. Health Info Libr J 2018 Dec 21;36(1):4-40. [doi: 10.1111/hir.12244] [Medline: 30578606]
5. Ayiku L, Levay P, Hudson T, Craven J, Finnegan A, Adams R, et al. The Embase UK filter: validation of a geographic search filter to retrieve research about the UK from OVID Embase. Health Info Libr J 2019 Jun;36(2):121-133 [FREE Full text] [doi: 10.1111/hir.12252] [Medline: 30912233]
6. Damarell RA, Lewis S, Trenerry C, Tieman JJ. Integrated Care Search: development and validation of a PubMed search filter for retrieving the integrated care research evidence. BMC Med Res Methodol 2020 Jan 21;20(1):12 [FREE Full text] [doi: 10.1186/s12874-020-0901-y] [Medline: 31964347]
7. Ioerger M, Flanders RM, Goss KD, Turk MA. Developing a systematic search strategy related to people with disability: a brief report testing the utility of proposed disability search terms in a search about opioid use. Disabil Health J 2019 Apr;12(2):318-322. [doi: 10.1016/j.dhjo.2018.11.009] [Medline: 30470478]
8. Kennedy M. Filter to retrieve studies related to LGBTQIA2S+ from the Ovid APA PsycInfo database. University of Alberta. 2025 Apr 4. URL: https://docs.google.com/document/d/1ZNt2GkByI5EpgDnxH0HCqBts-XStVr90VTsAk6zpX8M/edit?usp=sharing&usp=embed_facebook [accessed 2025-09-04]
9. Kennedy M. Filter to retrieve studies related to LGBTQIA2S+ from the Ovid MEDLINE database. University of Alberta. 2025 Apr 4. URL: https://docs.google.com/document/d/1ZNt2GkByI5EpgDnxH0HCqBts-XStVr90VTsAk6zpX8M/edit?tab=t.0#heading=h.qi55eeyvgzy9 [accessed 2025-11-28]
10. Boellstorff T, Cabral M, Cárdenas M, Cotten T, Stanley EA, Young K, et al. Decolonizing transgender: a roundtable discussion. Transgender Stud Q 2014 Aug 01;1(3):419-439. [doi: 10.1215/23289252-2685669]
11. Everhart AR. Bones without flesh and (trans)gender without bodies: querying desires for trans historicity. Hypatia 2023 Feb 23;37(4):601-618 [FREE Full text] [doi: 10.1017/hyp.2022.50]
12. Deal C, Doshi RD, Gonzales G. Gender minority youth experiencing homelessness and corresponding health disparities. J Adolesc Health 2023 May;72(5):763-769. [doi: 10.1016/j.jadohealth.2022.11.229] [Medline: 36646565]
13. Delozier AM, Kamody RC, Rodgers S, Chen D. Health disparities in transgender and gender expansive adolescents: a topical review from a minority stress framework. J Pediatr Psychol 2020 Sep 01;45(8):842-847. [doi: 10.1093/jpepsy/jsaa040] [Medline: 32626901]
14. Billard TJ. The politics of transgender health misinformation. Political Commun 2024 Jan 08;41(2):344-352. [doi: 10.1080/10584609.2024.2303148]
15. Coldwater KD. Decoding the Misinformation-Legislation Pipeline: an analysis of Florida Medicaid and the current state of transgender healthcare. J Med Libr Assoc 2023 Oct 02;111(4):750-761 [FREE Full text] [doi: 10.5195/jmla.2023.1724] [Medline: 37928129]
16. McNamara M, McLamore Q, Meade N, Olgun M, Robinson H, Alstott A. A thematic analysis of disinformation in gender-affirming healthcare bans in the United States. Soc Sci Med 2024 Jun;351:116943. [doi: 10.1016/j.socscimed.2024.116943] [Medline: 38759383]
17. Connolly MD, Zervos MJ, Barone CJ, Johnson CC, Joseph CL. The mental health of transgender youth: advances in understanding. J Adolesc Health 2016 Nov;59(5):489-495. [doi: 10.1016/j.jadohealth.2016.06.012] [Medline: 27544457]
18. Edmiston EK, Donald CA, Sattler AR, Peebles JK, Ehrenfeld JM, Eckstrand KL. Opportunities and gaps in primary care preventative health services for transgender patients: a systemic review. Transgend Health 2016 Nov 01;1(1):216-230 [FREE Full text] [doi: 10.1089/trgh.2016.0019] [Medline: 28861536]
19. Suen LW, Lunn MR, Katuzny K, Finn S, Duncan L, Sevelius J, et al. What sexual and gender minority people want researchers to know about sexual orientation and gender identity questions: a qualitative study. Arch Sex Behav 2020 Oct;49(7):2301-2318 [FREE Full text] [doi: 10.1007/s10508-020-01810-y] [Medline: 32875381]
20. Berger I, Ansara YG. Cisnormativity. In: Goldberg AE, Beemyn G, editors. The SAGE Encyclopedia of Trans Studies. Thousand Oaks, CA: SAGE Publications; 2021.
21. Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: simple definition and calculation of accuracy, sensitivity and specificity. Emerg (Tehran) 2015;3(2):48-49 [FREE Full text] [Medline: 26495380]
22. Jenkins M. Evaluation of methodological search filters--a review. Health Info Libr J 2004 Sep;21(3):148-163 [FREE Full text] [doi: 10.1111/j.1471-1842.2004.00511.x] [Medline: 15318913]
23. Farley N, Marshall Z. Creating a database for LGBTQ2SI+ knowledge synthesis (Knowsy): a mapping review of evidence and gaps to inform future research. 2020 Presented at: What Works Global Summit; October 30, 2020; Online URL: https://web.archive.org/web/20210515083935/https://www.wwgs2020.org/30-october.html

XSL•FO
RenderX

24. Home. Knowsy. URL: https://knowsy.ca/ [accessed 2025-03-17]
25. Ohtake PJ, Childs JD. Why publish study protocols? Phys Ther 2014 Sep 01;94(9):1208-1209 [FREE Full text] [doi: 10.2522/ptj.2014.94.9.1208]
26. Muente C, Pachanov A, Hirt J, Ayiku L, Pieper D. Validated geographic search filters for bibliographic databases: a scoping review protocol. JBI Evid Synth 2024 Mar 01;22(3):441-446. [doi: 10.11124/JBIES-23-00445] [Medline: 38344846]
27. Wilczynski NL, Morgan D, Haynes RB, Hedges Team. An overview of the design and methods for retrieving high-quality studies for clinical care. BMC Med Inform Decis Mak 2005 Jun 21;5:20 [FREE Full text] [doi: 10.1186/1472-6947-5-20] [Medline: 15969765]
28. Lunny CA, Salzwedel DM, Liu T, Ramasubbu C, Gerrish S, Puil L, et al. Protocol for the validation of four search strategies for retrieval of clinical practice guidelines in MEDLINE, Embase and PubMed. PeerJ Preprints Preprint posted online on August 28, 2018. [doi: 10.7287/peerj.preprints.27149v1]
29. Thomas J, Graziosi S, Brunton J, Ghouze Z, O'Driscoll P, Bond M, et al. EPPI-Reviewer: advanced software for systematic reviews, maps and evidence synthesis. EPPI Centre. 2023. URL: https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=2967 [accessed 2025-11-28]
30. Bem SL. Bem Sex Role Inventory. American Psychological Association. 2011 Sep 12. URL: https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft00748-000 [accessed 2025-11-28]
31. DeepL translate: the worlds most accurate translator. DeepL. URL: https://www.deepl.com/translator [accessed 2025-03-04]
32. Gisev N, Bell JS, Chen TF. Interrater agreement and interrater reliability: key concepts, approaches, and applications. Res Social Adm Pharm 2013;9(3):330-338. [doi: 10.1016/j.sapharm.2012.04.004] [Medline: 22695215]
33. Parker R, Foster M, Wanner A, Lackey M. Design and validation of a search filter for LGBTQ+ populations. J Can Health Libr Assoc 2018 Jul 24;39(2):94 [FREE Full text]
34. Schilperoort HM, Hickner A, Morgan-Daniel J, Parker RM. Development and validation of LGBTQIA+ search filters: report on process and pilot filter for queer women. J Med Libr Assoc 2025 Apr 18;113(2):123-132. [doi: 10.5195/jmla.2025.2002] [Medline: 40342307]
35. Lefebvre C, Glanville J, Beale S, Boachie C, Duffy S, Fraser C, et al. Assessing the performance of methodological search filters to improve the efficiency of evidence information retrieval: five literature reviews and a qualitative study. Health Technol Assess 2017 Nov;21(69):1-148 [FREE Full text] [doi: 10.3310/hta21690] [Medline: 29188764]
36. Wafford QE, Miller CH, Wescott A, Kubilius RK. Meeting a need: development and validation of PubMed search filters for immigrant populations. J Med Libr Assoc 2024 Jan 16;112(1):22-32 [FREE Full text] [doi: 10.5195/jmla.2024.1716] [Medline: 38911528]
37. Health sciences search filters. University of Alberta. 2025. URL: https://guides.library.ualberta.ca/health-sciences-search-filters [accessed 2025-11-28]
38. Hedges project. McMaster University. URL: https://hiruweb.mcmaster.ca/hkr/hedges/ [accessed 2025-09-18]
39. The ISSG search filters resource. The InterTASC Information Specialists' Sub-Group. URL: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home [accessed 2025-09-18]
40. Moving trans history forward. University of Victoria. 2025. URL: https://www.uvic.ca/mthf2025/index.php [accessed 2025-09-18]
41. Leaders unite at Global Evidence Summit. Joanna Briggs Institute. URL: https://jbi.global/news/article/global-evidence-summit [accessed 2025-09-18]
42. Malve A, Chawan PM. A comparative study of keyword and semantic based search engine. Int J Innov Res Sci Eng Technol 2015 Nov;4(11):11156-11161 [FREE Full text] [doi: 10.15680/IJIRSET.2015.0411039]

Abbreviations

DSD: difference of sex development
2SLGBTQIA+: Two-Spirit, lesbian, gay, bisexual, transgender, queer, intersex, asexual, and additional sexual and gender diverse terminologies
TGD: transgender and gender diverse

Edited by J Sarvestan; submitted 20.Apr.2025; peer-reviewed by N Smallheiser; comments to author 30.Jul.2025; accepted 10.Nov.2025; published 23.Jan.2026.

Please cite as:
Dinh C, Marshall Z, Marsalis S, Gagen R, Kang P, Everhart A
Retrieving Transgender and Gender Diverse Literature: Protocol for the Development and Validation of 2 Search Hedges
JMIR Res Protoc 2026;15:e76055
URL: https://www.researchprotocols.org/2026/1/e76055
doi:10.2196/76055
PMID:41575780

XSL•FO
RenderX

©Chi Dinh, Zack Marshall, Scott Marsalis, Ryn Gagen, Preet Kang, Avery Everhart. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Anxiety-Free Public Dentistry for Adults With Disabilities by Using Head-Mounted Virtual Reality Technology: Protocol for a Feasibility Mixed Methods Study

Dung T Bui[1,2], PhD; Fiona McAlinden[1], MBM; Alice Urban[3], BASc; Charmine E J Hartel[4,5], PhD; Kristian Rotaru[4,5,6], PhD; Kadek Ananta Satriadi[7], PhD; Tanner Person[4,5], MA; Wei Wang[4,5], PhD; Libby Callaway[2,8,9], PhD

[1]School of Primary and Allied Health Care, Faculty of Medicine, Nursing and Health Sciences, Monash University, Frankston, VIC, Australia

[2]National Centre for Healthy Ageing, Faculty of Medicine, Nursing and Health Sciences, Monash University, Frankston, VIC, Australia

[3]Integrated Care and Dental, Peninsula Health, Frankston, VIC, Australia

[4]Faculty of Business and Economics, Monash University, Caulfield, VIC, Australia

[5]Opportunity Tech Lab, Monash University, Clayton, VIC, Australia

[6]School of Psychological Sciences, Monash University, Clayton, VIC, Australia

[7]Faculty of Information Technology, Monash University, Clayton, VIC, Australia

[8]Rehabilitation, Ageing and Independent Living Research Centre, Faculty of Medicine, Nursing and Health Sciences, Monash University, Frankston, VIC, Australia

[9]Occupational Therapy Department, Faculty of Medicine, Nursing and Health Sciences, Monash University, Frankston, VIC, Australia

Corresponding Author:
Dung T Bui, PhD
School of Primary and Allied Health Care
Faculty of Medicine, Nursing and Health Sciences
Monash University
47-49 Moorooduc Hwy
Frankston, VIC, 3199
Australia
Phone: 61 0399044553
Email: trungdung.bui@monash.edu

Abstract

Background: Oral disease remains a global public health concern, disproportionately affecting socioeconomically disadvantaged populations. Adults with disabilities or health conditions face additional barriers to dental care, including physical accessibility, communication challenges, and heightened anxiety. These factors contribute to care avoidance and poorer oral health outcomes. While virtual reality (VR) has shown promise in reducing procedural anxiety in pediatric and private dental settings, its application in adult public dentistry, particularly for people with disabilities, remains underexplored.

Objective: This study aims to evaluate the feasibility, usability, and acceptability of Smileyscope, a Therapeutic Goods Administration–approved head-mounted VR headset, in reducing dental anxiety and enhancing care experiences for adults with disabilities in public dental clinics.

Methods: A mixed methods convergent design will be implemented across community dental sites in Victoria, Australia. In total, 50 adult patients and up to 10 dental staff members will be recruited. Primary feasibility indicators include recruitment rate (≥60% consent), completion rate (≥80% System Usability Scale [SUS] completion), and usability threshold (mean SUS ≥68). The primary analysis will be descriptive, with 95% CIs reported. Quantitative data will be collected using the visual analog scale for willingness, the SUS, and the technology acceptance model questionnaire. Qualitative data from semistructured interviews will be thematically analyzed. The study is co-designed with a disability advocate and integrates lived experience throughout all phases, including recruitment, evaluation, and dissemination.

Results: The project was funded in 2025, and ethics approval was granted by Peninsula Health Human Research Ethics Committee (project ID 117565). Data collection is scheduled for October 2025 to March 2026 at the participating community dental sites. Primary feasibility estimates and qualitative themes are expected to be submitted for publication in mid-2026.

XSL•FO
RenderX

Conclusions: This protocol outlines a feasibility study that will inform scalable models for VR integration into public dental services. The findings will contribute to improved oral health equity and patient-centered care, advancing the evidence base for inclusive digital health innovation in dentistry.

International Registered Report Identifier (IRRID): PRR1-10.2196/85916

(JMIR Res Protoc 2026;15:e85916) doi:10.2196/85916

KEYWORDS

virtual reality; dentistry; disability; inclusive health care; dental anxiety reduction; feasibility; oral health; innovation

Introduction

Background

In 2021, the World Health Assembly approved a resolution on oral care, recognizing it as a key preventive primary health care measure and setting targets for a global strategy and action plan for oral health [1]. Oral disease remains a significant global issue and disproportionately affects poor and socially disadvantaged populations. Unlike private or pediatric settings, public dental clinics face throughput constraints, variable sensory needs, and fewer standardized accommodations. These limitations can amplify anxiety and care avoidance among adults with disabilities.

Adults with disabilities, including those with different health conditions or varying levels of impairment or activity limitations [2,3], can face persistent and multifaceted challenges in accessing public dental care, including physical accessibility barriers, communication difficulties, and heightened sensory sensitivities. These challenges are compounded by systemic issues such as clinician reluctance, inadequate training in disability-inclusive care, and limited infrastructure to support tailored interventions [4,5]. Public dental services often operate under resource constraints, which can exacerbate anxiety and reduce treatment adherence among this population [6]. Compared to pediatric dental settings, in which virtual reality (VR) has shown promise in reducing procedural anxiety, adult public dentistry presents a more complex context. This includes greater variability in cognitive and physical abilities, fewer accommodations for sensory needs, and a lack of standardized protocols for VR integration [7]. Moreover, adults with disabilities or health conditions can be more likely to delay or avoid dental care due to fear, discomfort, or previous negative experiences, leading to poorer oral health outcomes [7-10]. These factors underscore a critical evidence gap: although VR has demonstrated efficacy in reducing procedural anxiety in pediatric and private dental settings, its feasibility and acceptability among adults with disabilities in resource-constrained public clinics remain largely unexplored.

As technology advances, the use of head-mounted displays for VR application–based purposes—including in the delivery of primary health care interventions such as dental care—is increasing and can offer novel approaches to patient engagement and treatment uptake [11]. The existing literature on VR technology within dental care highlights several critical points from both clinician and patient perspectives. Clinicians perceive VR as valuable for treatment planning and enhancing patient communication. Research also indicates that VR technology can facilitate the simulation of treatment outcomes and can assist in procedures such as dental implants [12,13]. Furthermore, clinicians report that VR technology can significantly enhance the patient experience by alleviating anxiety, particularly among pediatric populations and populations with disabilities. Studies indicate that VR positively affects patient anxiety levels and procedural efficiency, thereby rendering care procedures less daunting [14-17]. Clinicians regard VR as an essential tool for improving dental education and training as it can also offer immersive learning opportunities and allows for the practice of complex procedures within a controlled environment [13,18-21].

Research has demonstrated that patients similarly appreciate the contribution of VR in reducing dental anxiety and pain. This technology has previously been found to offer a distraction during procedures, thereby enhancing overall comfort [14,15,17]. Research also demonstrates notable reductions in heart rate and stress levels when VR is used as a distraction during dental interventions. Moreover, patients find value in VR's educational offering, which aids them in comprehending treatment plans and procedures more effectively [13,22].

Despite the favorable perceptions held by many clinicians and patients regarding VR, issues have also been found. A systematic review of factors associated with VR motion sickness in head-mounted displays found that sickness profiles were influenced by both the type and duration of content used (with visual stimulation or locomotion content and increased exposure times linked to motion sickness) and the age and ability of the user [11]. From a dental clinician perspective, research trials have indicated that headset bulkiness may impact the treatments that can be administered [23].

Several significant gaps also remain in the research regarding both the perspectives of dental staff and patients on the potential of this technology in dental care. Most studies that do exist primarily focus on immediate outcomes, such as anxiety and pain alleviation during a single dental visit. There is a lack of research examining the efficacy of VR across various dental procedures and its long-term benefits, particularly regarding its impact on patient outcomes and clinician performance over extended periods [14,24]. In addition, current literature often fails to account for diversity in patient demographics, including age, socioeconomic status, and cultural background [17]. Many studies have concentrated predominantly on pediatric patients, with limited research explicitly addressing the experiences and outcomes of adult patients. This narrow focus constrains the generalizability of findings and the capacity to tailor VR interventions to diverse patient groups.

XSL•FO
RenderX

There is a pressing need for comprehensive studies addressing the level of training that clinicians receive in using VR technology and their overall receptiveness to its integration into routine dental practice [25]. Moreover, the establishment of standardized protocols and guidelines for the integration of VR in dental practices has to ensure consistency and efficacy [26]. Addressing these gaps has the potential to offer a more comprehensive understanding of how VR technology can be effectively integrated into dental care, benefitting both clinicians and adult patients.

Smileyscope (Smileyscope Pty Ltd) is a Therapeutic Goods Administration–approved VR headset (Australian Register of Therapeutic Goods ID297181) [27] designed for health care settings with compatibility with standard infection control protocols and a library of immersive VR experiences tailored to health care settings. The headset is a low-cost, easy-to-deploy head-mounted VR display that has demonstrated efficacy in reducing procedural pain and anxiety in pediatric hospital environments [28-30]. It is widely used across Australia, the United States, and the United Kingdom, with multiple independent studies validating its effectiveness in health care delivery [31,32]. While Smileyscope has been trialed in various medical contexts, including palliative care and heart failure management [33-35], its application in dental care—particularly for adults with disabilities—remains unexplored.

Research Aim

This study aims to address the research gaps by evaluating the feasibility, acceptability, and usability of Smileyscope in public dental settings for adults, including people with disabilities (hereafter referred to as "patients"). By integrating lived experience and inclusive design principles, this study seeks to understand both patient and clinician perspectives on VR use during routine dental procedures. The findings will inform scalable models for VR integration into public oral health services, contributing to improved equity and patient-centered care, and prespecify feasibility benchmarks to inform a subsequent pilot or implementation study.

Research Questions

This study will address the following 6 research questions (RQs):

- RQ 1—what proportion of patients are willing to use head-mounted VR technology during routine dental care, and what factors influence willingness?
- RQ 2—what are the patients' subjective experiences of using the head-mounted VR technology as measured through usability (System Usability Scale; SUS) and qualitative accounts?
- RQ 3—what are patients' perceptions of the potential for head-mounted VR technology to enhance their experience during dental care?
- RQ 4—how willing are dental staff to trial and support VR technology use in their practice?
- RQ 5—what is the dental staff's subjective experience of using the head-mounted VR technology during care provision?
- RQ 6—what are the dental staff's perceptions of the potential for head-mounted VR technology to improve the patient experience during dental care?

Methods

This feasibility study will use a mixed methods convergent design guided by the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 checklist for interventional trials (Multimedia Appendix 1) [36].

Ethical Considerations

This research was approved by the Peninsula Health Human Research Ethics Committee (project ID117565; Multimedia Appendix 2). All participants will be required to provide informed consent before enrollment. Participation is voluntary, and individuals can withdraw at any time without impact on their care or employment.

All data will be deidentified before being entered into the data analysis software at the end of the data collection period. No identifiable information will be published or reported.

Electronic data will be stored in a password-protected cloud storage system provided by Monash University. Hard-copy files will be kept in a locked filing cabinet. Access to the data will be restricted to researchers only. The storage of collected data will comply with Peninsula Health policies and Monash University regulations, and the data will be accessible only to study investigators. All data will be retained for at least 5 years following the study publication date. After this period, all electronic files will be permanently deleted, and any paper files will be shredded and placed in a secure document bin for disposal at Monash University.

Adverse events will be monitored, including VR sickness, dizziness, distress, or device fit discomfort. Stop rules include immediate headset removal upon any adverse reaction, with events logged and patients offered standard care without the device.

Lived Experience Advisor Involvement

People with lived experience of disability and health conditions and community-based primary health care organizations in the health and disability sectors are actively involved in this project. This includes a formal collaboration with a lived experience advisor, who is an adult with cerebral palsy, and access to community advisory groups available through the collaborating health care network [37]. The project lived experience advisor will contribute to all components of the project, including the co-design of accessible recruitment materials, participant engagement strategies, and inclusive evaluation methods, providing specific paid deliverables, including co-design workshops and easy-read brief development. His lived experience of significant disability and assistive technology ensures that authentic disability perspectives guide all project phases from protocol development to dissemination. Lived experience involvement follows GRIPP2-SF (Guidance for Reporting Involvement of Patients and the Public 2–Short Form) reporting: the advocate will codesign recruitment materials and evaluation methods, with input directly shaping content selection

XSL•FO
RenderX

and accessibility protocols through documented workshop outcomes.

Settings

The study will be conducted at a community dental clinic, which is part of a public health network located at a community-based health and well-being hub in Victoria, Australia. This community-facing facility has been designed to deliver integrated health and well-being services. It hosts over a dozen organizations focused on intergenerational health, disability support, and inclusive health and well-being care. Within the hub, the community dental service provides general dental treatment to people of all ages in an accessible environment via prebooked appointments.

As part of this protocol development, 4 dental staff members from the hub were invited to an introductory session led by the study's chief investigator. They were informed of the proposed study design and had the opportunity to trial Smileyscope. They also consulted on the preparation of 1 of the 2 dental chairs that will be used for the VR trials following the clinical sanitation guidelines of the dental service and also following the clinical sanitation guidelines of Smileyscope. It was explained at this time that the feasibility pilot of the VR-enabled treatment would be used during routine dental checkups but would be excluded from complex or emergency dental surgeries. Infection control protocols include disposable face shields or covers, antimicrobial wipe-down between uses, and designated downtime for sterilization compliance.

Participants

Two distinct groups will be recruited.

- Group 1—adult patients aged ≥18 years who can provide informed consent for dental treatment and research participation. Exclusion criteria include emergency dental needs, contraindications for VR use, or inability to complete study instruments. The term "adult patients" is used operationally to reflect the real-world intake model of public dental services, in which adults may present with a diverse range of abilities or experience health conditions or disabilities. These may include physical, sensory, cognitive, psychosocial, and chronic health conditions or disabilities, which may be permanent, episodic, diagnosed, or self-identified [38-40]. All eligible patients visiting the Peninsula Health dental services will be invited to participate in the study. Both patients coming in for their first visit or subsequent dental appointments will be included. Given the current volume of local patients, this research aims to recruit 50 participants for group 1 from October 2025 to March 2026.
- Group 2—dental staff, comprising dentists, dental nurses, and other staff working at participating Peninsula Health sites during the VR trial period. All Peninsula Health dental staff who have been present at a participating Peninsula Health site will be invited to participate in the study. Given the staffing allocation at dental sites, we aim to recruit up to 10 dental staff members for group 2.

Recruitment

Recruitment flyers will be placed on trays in the waiting room of the dental service at both the Healthy Futures Hub and the adjacent Belvedere Community Centre. Dental reception staff will also invite patients to consider using VR during their dental care. The VR intervention will be offered across multiple weekdays between October 2025 and December 2025 using a schedule established in consultation with the local dental team to avoid congestion and ensure adequate time for patients and staff to engage with the technology. Patients can choose to trial the VR headset during their current visit or defer participation to a future appointment.

An experienced dental staff member was appointed by the Peninsula Health integrated and dental care manager to implement the VR trials in dental care as her rostered working day was the scheduled VR trial day (Fridays).

Study Intervention

On designated VR trial days, a research team member will approach patients and provide further information about the study, including a verbal explanation and printed versions of the project explanatory statement, for them to consider consenting to participate in the study (Figure 1). They can then choose to trial the VR headset in the waiting room and during dental treatment.

To initiate the experience in the waiting room, a researcher will first clean the headset using an antimicrobial wipe and then power it on by holding the red button on the front display. Using the touch screen interface, the researcher will select a suitable program based on the patient's age group, personal interest topic, and desired duration (rapid, short, medium, or long). Volume and language settings can also be adjusted, and for certain programs, the procedure site must be confirmed before proceeding (Figure 2).

Once the program is selected, the researcher will introduce the headset to the patient and ensure that the patient is in a safe and comfortable position. The headset will then be placed over the patient's eyes, with straps adjusted for a snug but comfortable fit. Patients who wear glasses can typically keep them on. After connecting the faceplate, the researcher will triple-click the green play button to start the VR experience. A chime will confirm activation, and verbal cues will guide the patient through the session. Upon completion, the headset will be carefully removed, cleaned, and recharged for future use.

The patients will then be asked whether they would like to continue wearing the VR headset during their dental procedure. Those who decline will be invited to complete a visual analog scale (VAS) and semistructured interviews to capture their perspectives on VR use in dental care. Patients who agree to use the headset will proceed with the VR experience during their dental treatment and subsequently complete the VAS, the SUS, the technology acceptance model (TAM) questionnaire, and the semistructured questionnaire.

XSL•FO
RenderX

Figure 1. Patient recruitment and intervention. SUS: System Usability Scale; TAM: technology acceptance model; VR: virtual reality.

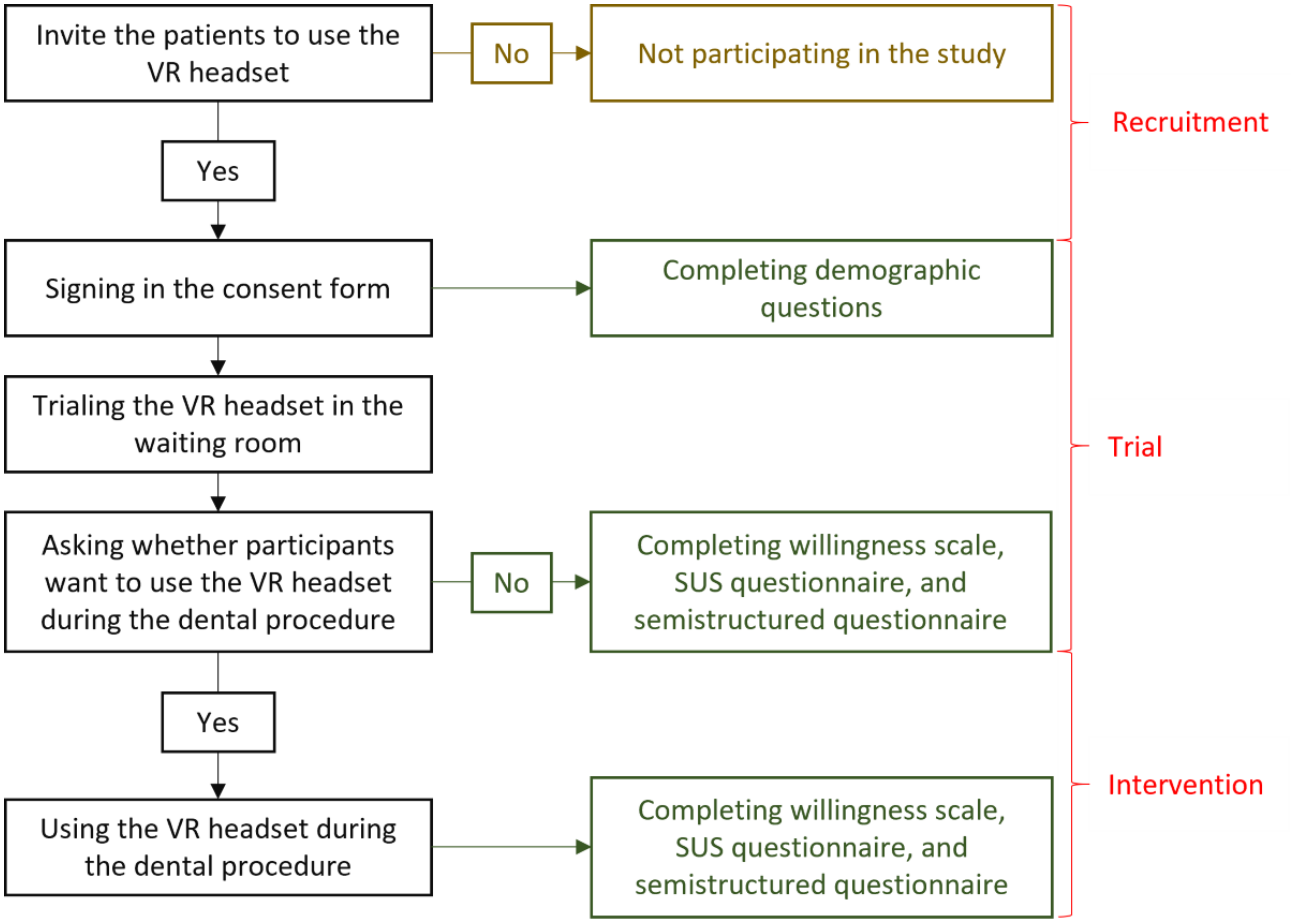

Figure 2. Map for selecting the experience on Smileyscope. VR: virtual reality.

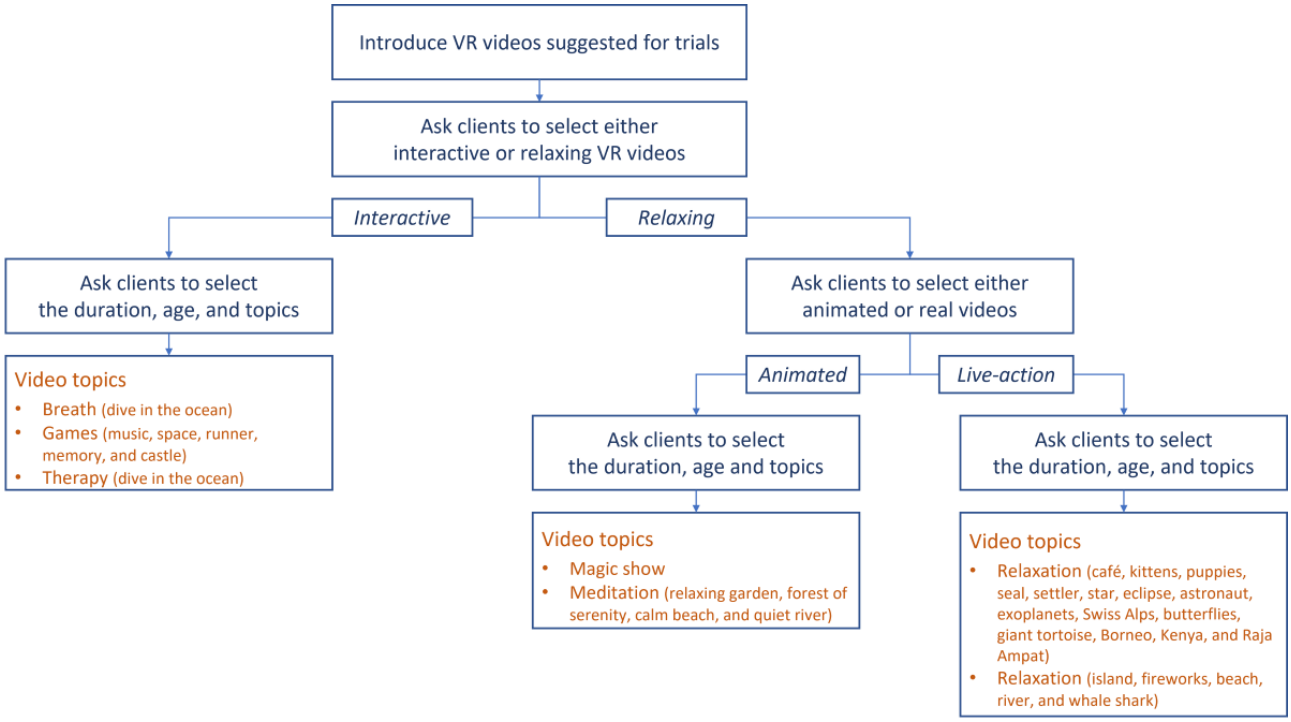

XSL•FO
RenderX

Outcome Measurements

Primary outcomes include the feasibility, acceptability, and usability of the VR intervention during routine dental care from both the patient and dental staff perspectives.

The SUS is a widely used tool for assessing the usability of digital systems [41]. It consists of 10 items that measure users' subjective perceptions of system usability, including aspects such as complexity, ease of use, and confidence in use. Each item is rated on a 5-point Likert scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). The SUS provides a single score from 0 to 100, which reflects overall usability. It is particularly suitable for health care settings due to its brevity, reliability, and ease of administration, typically requiring less than 5 minutes to complete.

The TAM is a theoretical framework used to evaluate user acceptance of technology [42]. The adapted TAM questionnaire for this study includes 3 core constructs: perceived usefulness, attitude toward use, and behavioral intention to use. Each construct is measured using 3 to 4 items rated on a 5-point Likert scale. Perceived ease of use, another TAM construct, is assessed through overlapping items in the SUS. This combined approach enables the study to assess both the usability and acceptance of the Smileyscope VR headset among dental staff.

The SUS-only version (Multimedia Appendix 3) is intended for the patients as it focuses on usability, which is central to understanding how patients experience the Smileyscope VR headset. The combined SUS and TAM version (Multimedia Appendix 4) is designed for the staff as it assesses technology acceptance, including intent to use, perceived usefulness, and attitudes, which are more relevant for staff, who decide whether to adopt the technology in practice.

In addition, a separate 7-point VAS (Multimedia Appendix 5) will be developed to assess the willingness of both patients and dental staff to use VR headsets during dental treatments. This scale will range from −3 (indicating no willingness at all) to +3 (indicating a high level of willingness).

A brief demographic survey and semistructured interviews will also be used with both groups. Indicative questions to be used in the short interview are provided in Multimedia Appendix 6.

Disability or health condition status will be documented primarily via patient self-report in the demographic survey or in the interview [43]. Formal diagnostic verification is not required, consistent with disability-inclusive research principles and public health care practice, where functional experience and access needs are prioritized over diagnostic categorization.

Complementary performance indicators will also be collected, including the type and length of procedures and the duration of VR headset use for each participant.

Data Collection

Table 1 summarizes the alignment between each RQ, the target participant group, the instruments used for data collection, and the timing of measurement throughout the study. This mapping ensures that feasibility, usability, willingness, and experiential outcomes related to VR use in dental care are captured systematically from both the patient and dental staff perspectives at appropriate stages of the trial.

Table 1. Data collection.

RQ[a]	Target participants	Measurement tool	Collection period
RQ 1: what proportion of patients are willing to use head-mounted VR[b] technology during routine dental care, and what factors influence willingness?	Patients	VAS[c]	After the trial and after the intervention
RQ 2: what are the patients' subjective experiences of using the head-mounted VR technology as measured through usability (SUS[d]) and qualitative accounts?	Patients	SUS	After the intervention
RQ 3: what are patients' perceptions of the potential for head-mounted VR technology to enhance their experience during dental care?	Patients	Semistructured questionnaire	After the intervention
RQ 4: how willing are dental staff to trial and support VR technology use in their practice?	Dental staff	VAS	Before the trial
RQ 5: what is the dental staff's subjective experience of using the head-mounted VR technology during care provision?	Dental staff	SUS and TAM[e] questionnaire	After the intervention
RQ 6: what are the dental staff's perceptions of the potential for head-mounted VR technology to improve the patient experience during dental care?	Dental staff	Semistructured questionnaire	After the intervention

[a]RQ: research question.

[b]VR: virtual reality.

[c]VAS: visual analog scale.

[d]SUS: System Usability Scale.

[e]TAM: technology acceptance model.

XSL•FO
RenderX

Data Management

The demographic survey, VAS, SUS, and TAM questionnaire, and semistructured questionnaire will contain reidentifiable information, with a participant code applied to each data collection sheet.

Participants can complete the survey in hard copy and return it to the researchers in person, send it via mail, or scan and email it to the research group. Alternatively, the survey can be completed electronically and emailed to the research group.

Statistical Considerations and Planned Analysis

The thematic analysis framework by Braun and Clarke [44] will be used to analyze the qualitative data gathered from the semistructured questionnaire. One researcher will enter the data gathered from the semistructured interviews according to the themes identified in each question. A second researcher will verify the accuracy of the data, whereas a third researcher will resolve any conflicts in data interpretation.

The thematic analysis will involve the following steps:

- Data familiarization—reading the responses multiple times to grasp the overall content
- Code generation—assigning labels or "codes" to key ideas or patterns in the data
- Review and refinement of themes—revisiting the data to ensure that the themes accurately represent the content and capture essential aspects of the feedback
- Presentation of results—description of the themes and how they relate to the RQs or objectives

Demographic data will be analyzed using descriptive statistics.

The SUS will be analyzed using the standard scoring method described by Brooke [45]. Each of the 10 items will be scored on a 5-point Likert scale. For positively worded items (odd numbered), the score contribution is the scale position minus 1; for negatively worded items (even numbered), the contribution is 5 minus the scale position. The total score is then multiplied by 2.5 to yield a usability score ranging from 0 to 100. Higher scores indicate better perceived usability. Descriptive statistics, including means and SDs, will be calculated to assess overall usability and variability in staff responses.

The TAM questionnaire will be analyzed by calculating the mean and SD for each of the 3 core constructs: perceived usefulness, attitude toward use, and behavioral intention to use. Perceived ease of use will be inferred from overlapping SUS items. A staff member will be considered to have accepted the technology if their mean score across behavioral intention to use items is of ≥4 (on a 5-point scale). The proportion of staff members meeting this threshold will be calculated. The TAM data will also be used to explore correlations between constructs and identify factors influencing acceptance.

Primary feasibility indicators include recruitment rate (≥60% consent), completion rate (≥80% SUS completion), and usability threshold (mean SUS ≥68). The primary analysis will be descriptive, with 95% CIs reported.

Given the feasibility nature of the study and the anticipated sample size, formal stratification by disability subtype is not planned [38,46]. However, to support interpretability and transparency, participants' self-reported disability or health condition characteristics will be captured descriptively. Descriptive categories may include physical, sensory, cognitive, psychosocial, and chronic health conditions, recognizing that participants may identify across multiple domains. These data will be reported using descriptive summaries and narrative interpretation rather than inferential subgroup analyses [40,46]. Where sample size permits, exploratory noncomparative observations may be presented to inform the design of future pilot or implementation studies.

Expected Outcomes

This study will generate foundational evidence in Australia on the feasibility, usability, and acceptability of integrating VR technology into adult public dental care offered within a community dental service. By evaluating the Smileyscope VR headset, the research will provide critical insights into patient and clinician experiences, willingness to use the technology, and its integration into routine dental workflows.

Key outcomes will include the development of an implementation toolkit containing costed solutions to common VR integration barriers, such as staff training protocols and accessibility modifications. Additionally, the study will produce evidence-informed recommendations for scaling sustainable VR integration in public dental services supported by specific cost-benefit analyses. These findings will directly inform service design improvements and contribute to the growing evidence base on digital health innovation in dentistry.

Results

The study received institutional funding support in 2025 and Peninsula Health Human Research Ethics Committee approval before participant enrollment (project ID117565). Data collection is planned for October 2025 to March 2026 across participating public dental clinics. Initial feasibility and usability results are targeted for submission in mid-2026. Recruitment data have been cleaned and analyzed at this time.

Discussion

Expected Findings

This paper outlines an a priori research protocol for a feasibility study designed to evaluate the integration of head-mounted VR technology into routine public dental care for adults with disabilities or health conditions. The intervention will use the Smileyscope VR headset to deliver immersive, calming experiences aimed at reducing anxiety and enhancing patient engagement. By assessing usability, acceptability, and willingness to use VR from both the patient and clinician perspectives, the study aims to generate foundational evidence to inform future implementation and scale-up efforts.

Previous research has demonstrated the effectiveness of VR in pediatric dental and hospital settings, particularly in reducing procedural anxiety and improving patient cooperation

XSL•FO
RenderX

[14-17,28-32]. However, there is limited evidence on the use of VR in adult dental care, especially among populations with disabilities or chronic health conditions. This study addresses that gap by focusing on adult patients in community-based public dental services using a co-designed approach that incorporates lived experience and inclusive evaluation methods.

The protocol also builds on broader health care literature that supports VR as a tool for distraction, pain management, and procedural efficiency [20-22]. Unlike many existing studies that focus on single-session outcomes or pediatric populations, this study is designed to explore the feasibility of integrating VR into routine workflows and understand the perspectives of both patients and clinicians.

Benefits and Risks

The anticipated benefits of VR integration include reduced dental anxiety, improved patient cooperation, and enhanced clinician focus during procedures. These outcomes may contribute to more efficient care delivery and improved patient satisfaction. For dental staff, the use of VR may foster a more supportive work environment by reducing procedural stress and improving patient engagement.

Potential risks include discomfort or disorientation for some patients, particularly those with vestibular or other sensory sensitivities or cognitive communication changes. The headset may not be suitable for all procedures, and its use requires appropriate screening and staff training. Technical issues, hygiene protocols, and workflow disruptions are also considerations that will be addressed through the implementation toolkit.

Limitations of the Study and Future Directions

As a feasibility study, this protocol is limited by its short duration (October 2025 to December 2025) and relatively small sample size. The intervention is restricted to routine dental checkups and excludes emergency care and complex procedures. Only patients capable of providing informed consent will be included, which may limit generalizability to all populations with cognitive communication changes.

Future research should explore the long-term impact of VR integration on oral health outcomes, treatment adherence, and clinician performance. Studies involving broader and more diverse populations, including children, older adults, and individuals with cognitive communication impairments, are needed to assess the scalability and inclusivity of VR interventions. Additionally, further investigation into cost-effectiveness and integration into other areas of primary and allied health care will be essential for sustainable adoption.

Conclusions

This protocol presents a feasibility study that will evaluate the use of head-mounted VR technology to reduce dental anxiety and improve care experiences for adults with disabilities. By assessing usability, acceptability, and willingness to use Smileyscope in public dental settings, the study will generate foundational evidence to inform scalable, inclusive models for VR integration into oral health care. The findings will support service design improvements and contribute to the digital health innovation literature in dentistry.

Acknowledgments

The authors wish to thank Jonathan Bredin for the disability lived experience he contributed to this protocol development and the staff of the Healthy Futures Hub and Peninsula Health Community Dental Service for their participation and support throughout study protocol development and preparation.

Funding

This paper was produced with the financial support of a grant from the Monash Assistive Technology and Society Centre. The sponsor had no role in the design; implementation; or data collection, analysis, and interpretation of the study or in the decision to submit the manuscript for publication.

Authors' Contributions

DTB led the evaluation conceptualization, protocol development, technology testing, and manuscript drafting. LC led project conceptualization and contributed to methodology design, partnership engagement, supervision, and manuscript review. FM supported study coordination and contributed to manuscript editing. AU provided clinical expertise and facilitated site implementation. CEJH, KR, and TP contributed to the design of evaluation tools and data analysis frameworks. KAS supported the integration of virtual reality technology and the technical setup. WW contributed to statistical planning and data management strategies. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
SPIRIT checklist.
[DOC File , 125 KB - resprot_v15i1e85916_app1.doc]

XSL•FO
RenderX

Multimedia Appendix 2
Ethics approval letter.
[PDF File (Adobe PDF File), 595 KB - resprot_v15i1e85916_app2.pdf]

Multimedia Appendix 3
System Usability Scale for patients.
[DOCX File , 27 KB - resprot_v15i1e85916_app3.docx]

Multimedia Appendix 4
System Usability Scale and technology acceptance model questionnaire for staff.
[DOCX File , 31 KB - resprot_v15i1e85916_app4.docx]

Multimedia Appendix 5
Visual analog scale for willingness.
[DOCX File , 35 KB - resprot_v15i1e85916_app5.docx]

Multimedia Appendix 6
Semistructured questions.
[DOCX File , 21 KB - resprot_v15i1e85916_app6.docx]

References

1. Oral health. World Health Organization. 2021 Jan 21. URL: https://apps.who.int/gb/ebwha/pdf_files/EB148/B148_R1-en.pdf [accessed 2025-09-17]
2. People with disability in Australia. Australian Institute of Health and Welfare. 2024. URL: https://www.aihw.gov.au/reports/disability/people-with-disability-in-australia/contents/summary [accessed 2025-10-10]
3. Noncommunicable diseases. World Health Organization. 2025 Sep 25. URL: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2025-10-10]
4. Asiri FY, Tennant M, Kruger E. Barriers to oral health care for persons with disabilities: an overview of systematic reviews. Community Dent Health 2024 Nov 29;41(4):256-264. [doi: 10.1922/CDH_00173Asiri09] [Medline: 39535395]
5. Park BY, Cho HA, Shin H. Disparity in access for people with disabilities to outpatient dental care services: a retrospective cohort study. BMC Oral Health 2023 Apr 14;23(1):213 [FREE Full text] [doi: 10.1186/s12903-023-02948-6] [Medline: 37060034]
6. Bhadauria US, Purohit B, Priya H. Access to dental care in individuals with disability: a systematic review. Evid Based Dent 2024 Jun;25(2):111. [doi: 10.1038/s41432-024-00970-3] [Medline: 38225370]
7. Hoffmann B, Erwood K, Ncomanzi S, Fischer V, O'Brien D, Lee A. Management strategies for adult patients with dental anxiety in the dental clinic: a systematic review. Aust Dent J 2022 Mar;67 Suppl 1(Suppl 1):S3-13 [FREE Full text] [doi: 10.1111/adj.12926] [Medline: 35735746]
8. Borg-Bartolo R, Roccuzzo A, Gambetta-Tessini K, Schimmel M, Molinero-Mourelle P, Sabatini G, et al. Association between oral health-related quality of life and structural determinants of health among elderly populations. A systematic review and meta-analysis. Int J Equity Health 2025 Jul 01;24(1):189 [FREE Full text] [doi: 10.1186/s12939-025-02557-5] [Medline: 40597152]
9. World Health Assembly Resolution paves the way for better oral health care. World Health Organization. 2021 May 27. URL: https://www.who.int/news/item/27-05-2021-world-health-assembly-resolution-paves-the-way-for-better-oral-health-care [accessed 2025-10-15]
10. Taqi M, Zaidi SJ, Javaid J, Alam Z, Saleem A, Khan SA. Patient perceptions and experiences of dental fear of different dental specialties: a mixed-method study. BMC Oral Health 2023 Nov 19;23(1):884 [FREE Full text] [doi: 10.1186/s12903-023-03626-3] [Medline: 37981667]
11. Saredakis D, Szpak A, Birckhead B, Keage HA, Rizzo A, Loetscher T. Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. Front Hum Neurosci 2020 Mar 31;14:96 [FREE Full text] [doi: 10.3389/fnhum.2020.00096] [Medline: 32300295]
12. Monaghesh E, Negahdari R, Samad-Soltani T. Application of virtual reality in dental implants: a systematic review. BMC Oral Health 2023 Aug 28;23(1):603 [FREE Full text] [doi: 10.1186/s12903-023-03290-7] [Medline: 37641060]
13. Monterubbianesi R, Tosco V, Vitiello F, Orilisi G, Fraccastoro F, Putignano A, et al. Augmented, virtual and mixed reality in dentistry: a narrative review on the existing platforms and future challenges. Appl Sci 2022 Jan 15;12(2):877 [FREE Full text] [doi: 10.3390/app12020877]
14. Cunningham A, McPolin O, Fallis R, Coyle C, Best P, McKenna G. A systematic review of the use of virtual reality or dental smartphone applications as interventions for management of paediatric dental anxiety. BMC Oral Health 2021 May 07;21(1):244 [FREE Full text] [doi: 10.1186/s12903-021-01602-3] [Medline: 33962624]

XSL•FO
RenderX

15. La Paglia F, Belluardo L, Lo Baido R, La Cascia C, Wiederhold BK, La Barbera D. Feasibility of virtual reality environments use in reduction of dental anxiety during treatment. Ann Rev Cyberther Telemed 2023;21:228-232 [FREE Full text]
16. Pawar M, Pandya P, Rimple CM, Agrawal K, Dasarraju RK, Jain S. Virtual reality distraction: a novel behaviour management technique. J Pharm Bioallied Sci 2024 Feb;16(Suppl 1):S53-S55 [FREE Full text] [doi: 10.4103/jpbs.jpbs_965_23] [Medline: 38595521]
17. Tesfaye EM, Garito IP, Lewis-Fung S, Calabia D, Schlosser R, Appel L. Using virtual reality to manage pain and anxiety during dental treatments in patients with stroke: a case series. SAGE Open Med Case Rep 2024 Apr 30;12:2050313X241249910 [FREE Full text] [doi: 10.1177/2050313X241249910] [Medline: 38694905]
18. Genaro LE, Capote TS. Use of virtual reality in dentistry: literature review. Metaverse 2021;21(1) [FREE Full text] [doi: 10.54517/m.v2i1.1799]
19. Moussa R, Alghazaly A, Althagafi N, Eshky R, Borzangy S. Effectiveness of virtual reality and interactive simulators on dental education outcomes: systematic review. Eur J Dent 2021 Aug 24;16(1):14-31 [FREE Full text] [doi: 10.1055/s-0041-1731837] [Medline: 34428851]
20. Özdemi˙r G, Albayrak B, Yüzbaşioğlu E, Ölçer Us Y. Virtual articulators, virtual occlusal records and virtual patients in dentistry. J Exp Clin Med 2021 May;38(SI-2):129-135. [doi: 10.52142/omujecm.38.si.dent.9]
21. Gandedkar NH, Wong M, Prasad S, Ali Darendeliler M. Augmented/virtual reality and artificial intelligence in dental education and research. In: Geroimenko V, editor. Augmented Reality and Artificial Intelligence. Cham, Switzerland: Springer; 2023:147-169.
22. Hu J, Luo E, Song E, Xu X, Tan H, Zhao Y, et al. Patients' attitudes towards online dental information and a web-based virtual reality program for clinical dentistry: a pilot investigation in China. Int J Med Inform 2009 Mar;78(3):208-215. [doi: 10.1016/j.ijmedinf.2008.07.003] [Medline: 18752987]
23. Ghobadi A, Moradpoor H, Sharini H, Khazaie H, Moradpoor P. The effect of virtual reality on reducing patients' anxiety and pain during dental implant surgery. BMC Oral Health 2024 Feb 05;24(1):186 [FREE Full text] [doi: 10.1186/s12903-024-03904-8] [Medline: 38317209]
24. Al-Khaled I, Al-Khaled A, Abutayyem H. Augmented reality in dentistry: uses and applications in the digital era. Edelweiss Appl Sci Technol 2021;5:25-32 [FREE Full text] [doi: 10.33805/2576-8484.191]
25. Masood Z, Qabool H, Fida M, Sukhia RH, Fida M. Exploring the knowledge and awareness on applications of virtual reality and augmented reality technology among dental healthcare professionals - a crosssectional survey. J Pak Med Assoc 2024 Apr;74(4 (Supple-4)):S10-S16. [doi: 10.47391/JPMA.AKU-9S-03] [Medline: 38712404]
26. Fahim S, Maqsood A, Das G, Ahmed N, Saquib S, Lal A, et al. Augmented reality and virtual reality in dentistry: highlights from the current research. Appl Sci 2022 Apr 07;12(8):3719 [FREE Full text] [doi: 10.3390/app12083719]
27. Smileyscope Pty Ltd - VR apparatus - diversionary therapy device (297181). Department of Health, Disability and Ageing Australia. 2017 Dec 5. URL: https://www.tga.gov.au/resources/artg/297181 [accessed 2025-09-15]
28. Chan E, Hovenden M, Ramage E, Ling N, Pham JH, Rahim A, et al. Virtual reality for pediatric needle procedural pain: two randomized clinical trials. J Pediatr 2019 Jun;209:160-7.e4. [doi: 10.1016/j.jpeds.2019.02.034] [Medline: 31047650]
29. Tas FQ, van Eijk CA, Staals LM, Legerstee JS, Dierckx B. Virtual reality in pediatrics, effects on pain and anxiety: a systematic review and meta-analysis update. Paediatr Anaesth 2022 Dec;32(12):1292-1304 [FREE Full text] [doi: 10.1111/pan.14546] [Medline: 35993398]
30. Storey K, Dimanopoulos TA, Plummer K, Kimble R, Xiang H, Griffin B. Acceptability and usability of smileyscope virtual reality for paediatric pain management during burn procedures: perspectives of patients, carers and clinicians. J Adv Nurs 2025 Mar;81(3):1568-1582. [doi: 10.1111/jan.16417] [Medline: 39227177]
31. Gordon SF, Lam J, Vasquez JT, Cercone R, Tenneti N, Hart J, et al. A tailored COVID-19 vaccination pathway for children 5-11 years in Victoria, Australia. Vaccine 2023 May 22;41(22):3436-3445 [FREE Full text] [doi: 10.1016/j.vaccine.2023.04.032] [Medline: 37120401]
32. Samnakay S, Bell E, Evans D, Sommerfield D, Sommerfield A, Hauser N, et al. Assessing the use and acceptability of virtual reality to assist coping in children undergoing clinical procedures. J Spec Pediatr Nurs 2025 Jan;30(1):e70002. [doi: 10.1111/jspn.70002] [Medline: 39722555]
33. Codrington J, Qureshi FH, Rahman F, Saltzman RG, Deebel NA, Cartaya S, et al. Virtual reality for improved vasectomy experience: insights from a single-center clinical trial. J Sex Med 2024;21(Supplement_1):qdae001.250. [doi: 10.1093/jsxmed/qdae001.250]
34. Groninger H, Stewart D, Fisher JM, Tefera E, Cowgill J, Mete M. Virtual reality for pain management in advanced heart failure: a randomized controlled study. Palliat Med 2021 Dec 25;35(10):2008-2016. [doi: 10.1177/02692163211041273] [Medline: 34431398]
35. Vayssiere P, Constanthin PE, Baticam NS, Herbelin B, Degremont C, Blanke O, et al. Use of virtual reality to improve the quality of the hospital stay for patients in neurosurgery. Front Virtual Real 2021 Oct;2:736122. [doi: 10.3389/frvir.2021.736122]
36. Chan A, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med 2013 Feb 05;158(3):200-207 [FREE Full text] [doi: 10.7326/0003-4819-158-3-201302050-00583] [Medline: 23295957]

XSL•FO
RenderX

37. Consumer advisory groups. Bayside Health Peninsula. URL: https://www.peninsulahealth.org.au/support-us/consumer-advisory-groups/ [accessed 2025-10-17]
38. Herman C, Eisenberg Y, Vanderbom K, Tempio D, Gardner J, Rimmer J. Feasibility of implementing disability inclusive evidence-based health promotion. J Public Health Manag Pract 2023;29(1):82-92 [FREE Full text] [doi: 10.1097/PHH.0000000000001671] [Medline: 36448761]
39. Kuper H, Shakespeare T. Disability inclusion in research and health care. Nat Rev Gastroenterol Hepatol 2024 Sep 15;21(9):601-602. [doi: 10.1038/s41575-024-00958-5] [Medline: 39009891]
40. Bailie J, Fortune N, Plunkett K, Gordon J, Llewellyn G. A call to action for more disability-inclusive health policy and systems research. BMJ Glob Health 2023 Mar 23;8(3):e011561 [FREE Full text] [doi: 10.1136/bmjgh-2022-011561] [Medline: 36958749]
41. Lewis JR. The system usability scale: past, present, and future. Int J Hum Comput Interact 2018 Mar 30;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
42. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003 [FREE Full text] [doi: 10.1287/mnsc.35.8.982]
43. Simon L, Gölz F, Schenk O, Bührmann T, Kauff M, Kraus de Camargo O, et al. The international classification of functioning, disability and health in clinical practice, research findings and their impact on training and education. Front Rehabil Sci 2024 Sep 12;5:1420498. [doi: 10.3389/fresc.2024.1420498] [Medline: 39329060]
44. Braun V, Clarke V. Thematic Analysis: A Practical Guide. Thousand Oaks, CA: SAGE Publications; 2021.
45. Brooke J. SUS: a 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation in Industry. London, UK: CRC Press; 1996.
46. Evans E, Zengul A, Knight A, Willig A, Cherrington A, Mehta T, et al. Stakeholders' perspectives, needs, and barriers to self-management for people with physical disabilities experiencing chronic conditions: focus group study. JMIR Rehabil Assist Technol 2023 Dec 18;10:e43309. [doi: 10.2196/43309] [Medline: 38109170]

Abbreviations

GRIPP2-SF: Guidance for Reporting Involvement of Patients and the Public 2–Short Form
RQ: research question
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SUS: System Usability Scale
TAM: technology acceptance model
VAS: visual analog scale
VR: virtual reality

Edited by GRM La Rosa; submitted 17.Oct.2025; peer-reviewed by M Pang, D Shah; comments to author 06.Jan.2026; revised version received 12.Jan.2026; accepted 14.Jan.2026; published 13.Feb.2026.

Please cite as:
Bui DT, McAlinden F, Urban A, Hartel CEJ, Rotaru K, Satriadi KA, Person T, Wang W, Callaway L
Anxiety-Free Public Dentistry for Adults With Disabilities by Using Head-Mounted Virtual Reality Technology: Protocol for a Feasibility Mixed Methods Study
JMIR Res Protoc 2026;15:e85916
URL: https://www.researchprotocols.org/2026/1/e85916
doi:10.2196/85916
PMID:

©Dung T Bui, Fiona McAlinden, Alice Urban, Charmine E J Hartel, Kristian Rotaru, Kadek Ananta Satriadi, Tanner Person, Wei Wang, Libby Callaway. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Cumulative Stress Burden and Association With DNA Methylation in Ethiopian American Immigrants: Protocol for a Community-Engaged, Biopsychosocial Study

Lara Stone[1], MA; Lili Ding[2,3], PhD; Meron Hirpa Kassa[4], MD; Beakal Amsalu[1], MPH; Bruktawit Getnet[1], BA; Hermella Eshete[1], MPH; Tesfaye B Mersha[2,5,6], PhD; Lisa M Vaughn[1,2], MA, MEd, PhD

[1]Division of Emergency Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[2]Department of Pediatrics, College of Medicine, University of Cincinnati, Cincinnati, OH, United States

[3]Division of Biostatistics/Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[4]The Christ Hospital Health Network, Cincinnati, OH, United States

[5]Division of Asthma, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[6]Precision Pulmonary Medicine Program, Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
Lisa M Vaughn, MA, MEd, PhD
Department of Pediatrics
College of Medicine
University of Cincinnati
3333 Burnet Ave
ML 2008
Cincinnati, OH, 45229
United States
Phone: 1 5136369424
Fax: 1 5136367967
Email: lisa.vaughn@cchmc.org

Related Article:

Correction of: https://www.researchprotocols.org/2026/1/e85971

(JMIR Res Protoc 2026;15:e95656) doi:10.2196/95656

In "Cumulative Stress Burden and Association With DNA Methylation in Ethiopian American Immigrants: Protocol for a Community-Engaged, Biopsychosocial Study" [1] the authors noted two errors.

The seventh author (TBM) has added the middle initial "B" to their name.

In addition, the following affiliation has been added as affiliation 6 and linked to TBM:

Precision Pulmonary Medicine Program, Department of Medicine, Indiana University School of Medicine, Indianapolis, IN

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Stone L, Ding L, Hirpa Kassa M, Amsalu B, Getnet B, Eshete H, et al. Cumulative stress burden and association with DNA methylation in Ethiopian American immigrants: protocol for a community-engaged, biopsychosocial study. JMIR Res Protoc 2026 Mar 10;15:e85971 [FREE Full text] [doi: 10.2196/85971] [Medline: 41667122]

XSL•FO
RenderX

Submitted 19.Mar.2026; this is a non–peer-reviewed article;accepted 19.Mar.2026; published 31.Mar.2026.

Please cite as:
Stone L, Ding L, Kassa MH, Amsalu B, Getnet B, Eshete H, Mersha TB, Vaughn LM
Correction: Cumulative Stress Burden and Association With DNA Methylation in Ethiopian American Immigrants: Protocol for a Community-Engaged, Biopsychosocial Study
JMIR Res Protoc 2026;15:e95656
URL: https://www.researchprotocols.org/2026/1/e95656
doi:10.2196/95656
PMID:

©Lara Stone, Lili Ding, Meron Hirpa Kassa, Beakal Amsalu, Bruktawit Getnet, Hermella Eshete, Tesfaye B Mersha, Lisa M Vaughn. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Cultural Adaptation of a Digital Mobile App for Bipolar Disorder (PolarUs): Protocol for a Qualitative Co-Design Study

Leena W Chau[1], PhD; Jill K Murphy[2], PhD; Emma Morton[3], PhD; Martin D Provencher[4], PhD; Delphine Raucher-Chene[5,6], MD, PhD; Steven J Barnes[7], PhD; Erin E Michalak[1], PhD

[1]Department of Psychiatry, Faculty of Medicine, University of British Columbia, 2255 Wesbrook Mall, Vancouver, BC, Canada

[2]Interdisciplinary Health Program, St. Francis Xavier University, Antigonish, NS, Canada

[3]Faculty of Medicine, Nursing and Health Sciences, School of Psychological Sciences, Monash University, Clayton, Victoria, Australia

[4]Department of Psychology, Université Laval, Quebec City, QC, Canada

[5]Department of Psychiatry, McGill University, Montreal, QC, Canada

[6]Douglas Research Centre, Montreal, QC, Canada

[7]Department of Psychology, Faculty of Arts, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Leena W Chau, PhD
Department of Psychiatry, Faculty of Medicine, University of British Columbia, 2255 Wesbrook Mall, Vancouver, BC, Canada

Abstract

Background: Bipolar disorder (BD) affects approximately 40 million people worldwide and is a chronic, potentially disabling mood disorder. Although effective treatments exist, access to evidence-informed psychosocial care remains limited, particularly for culturally and linguistically diverse populations, contributing to persistent global treatment gaps. Digital mental health interventions (DMHIs), such as smartphone apps, offer a promising means to improve access to self-management support and quality of life (QoL), an outcome prioritized by people with BD and in clinical guidelines. However, most apps for BD lack quality and are not culturally adapted or co-designed with people with BD, limiting relevance and engagement. PolarUs (mobile app) is an evidence-informed DMHI developed using co-design with people with BD. The app is structured on the core 14 domains from the Quality of Life in BD scale, the only BD-tailored scale, combined with psychoeducation on self-management strategies and QoL. A recent pilot study demonstrated promising QoL, clinical, and feasibility outcomes.

Objective: This study aims to culturally and linguistically adapt the PolarUs app into French, Chinese, and Spanish for the North American context using qualitative and co-design methods.

Methods: Guided by community-based participatory research principles, whereby end users are engaged throughout the research process, and the Ecological Validity Framework of Bernal et al, we will engage advisory groups of people with lived experience from each linguistic community throughout the cultural adaptation process. Semimonthly virtual meetings will support systematic cultural adaptation of the self-management strategies, affirmations, and resources while maintaining fidelity to core evidence-based components. This will include cultural tailoring of app content and the identification of culturally appropriate resources. Advisory groups will also contribute to the cointerpretation of findings and the co-design of culturally appropriate recruitment and implementation strategies of PolarUs for a future clinical trial. Meetings will be recorded and coanalyzed as research data with advisory groups using qualitative reflexive thematic analysis to capture advisory group perspectives and experiences.

Results: This study was funded in October 2024. As of January 31, 2026, we enrolled 7 participants, and the results are expected to be published in the fall of 2026.

Conclusions: The findings will support the development of a culturally appropriate DMHI for BD for additional linguistic communities, advance cultural adaptation methodologies, and inform preparation for a future clinical trial. This study will produce the first culturally adapted, BD-specific DMHI developed through co-design using a community-based participatory research approach with multilingual end users from traditionally underserved communities, advancing equitable access, engagement, and scalability of DMHIs for BD and digital health care more broadly.

International Registered Report Identifier (IRRID): PRR1-10.2196/92600

(JMIR Res Protoc 2026;15:e92600) doi:10.2196/92600

KEYWORDS

digital mental health; bipolar disorder; cultural adaptation; co-design; quality of life; mood disorders

XSL•FO
RenderX

Introduction

Background

Approximately 40 million people worldwide live with bipolar disorder (BD [1]), a chronic and potentially disabling mood disorder, with lifetime prevalence estimated to be 2.4% [2] and onset typically in late adolescence or early adulthood [3]. Prevalence is consistent across cultures and ethnic groups [2]. BD is characterized by recurrent mood episodes ranging from manic or hypomanic states to major depressive episodes [4], with individuals frequently experiencing residual symptoms and impairment between mood episodes [5]. Lifelong treatment with ongoing symptom monitoring is thus often essential to mitigating risk of relapse, supporting recovery, and enhancing quality of life (QoL) [6]. QoL outcomes supported by psychoeducation on self-management are prioritized by people with BD [7,8] and emphasized in treatment guidelines [9]; optimal interventions therefore require attention to QoL with self-management psychoeducation, in addition to symptom management to support long-term recovery and resilience in BD [5].

Effective treatments for BD are available, and people with the condition can experience good QoL and overall health. Despite this, access to evidence-informed treatments is limited, especially in low-income countries [10] and among culturally diverse communities in high-income countries [11], who also tend to experience higher misdiagnosis rates [12], reflecting the inadequate understanding and application of culture-bound aspects of BD presentation, diagnosis, and treatment [13]. As a result, this mental health treatment gap can lead to people with BD experiencing substantial alterations in mood, energy, cognition, and functional capacity, resulting in debilitating personal and social impact [14]. Up to 10% of people with BD die by suicide [15]. The economic costs attributable to BD are also substantial [16]; in the United States, this was estimated at US $202 billion in 2015 [17].

Digital mental health interventions (DMHIs), particularly smartphone apps, offer one promising means to address the treatment gap by improving access to self-management information and real-time support. App-based interventions offer an innovative, highly accessible way of providing accessible, user-friendly, and self-directed support for people living with chronic conditions such as BD [18], even for those living in rural or remote areas [19]. Numerous apps exist for BD; however, most lack quality [20], BD-specific tailoring [21], QoL focus [22], co-design [23], and adaptation for cultural appropriateness supporting equitable access [24]. Many also lack an evidence base for their efficacy and in worst-case scenarios may expose individuals to harmful or misleading advice or fail to meet basic privacy standards [25]. Evidence from the broader digital mental health literature has also shown that the desires and goals of end users are not adequately taken into account during app development, and as a result, user engagement with mental health apps, particularly long term, remains low [23,26]. Using advanced co-design methods, the Collaborative Research Team to Study Psychosocial Issues in BD (CREST.BD), led by principal investigator (PI) and senior author EEM, developed the comprehensive PolarUs DMHI to help address this gap.

PolarUs Intervention

The PolarUs mobile app is a DMHI for people with BD designed to support users to self-monitor and self-manage key aspects of QoL in BD. A full description of features, content, and the co-design process of the PolarUs Alpha app is outside the scope of this paper; interested readers are referred to the protocol for more details [22]. Briefly, the app is organized according to life areas assessed by the QoL in BD scale (QoL.BD). The QoL.BD scale, developed by CREST.BD, is the first and only QoL scale tailored for BD and has been translated or validated in multiple languages [27]. The full QoL.BD assesses 14 domains (12 core: Physical, Sleep, Mood, Cognition, Leisure, Social, Spirituality, Finance, Household, Self-Esteem, Independence, Identity, and 2 optional: Work, Study) across 56 items, while the Brief QoL.BD scale assesses the 12 core domains with 12 corresponding items. The app also contains evidence-based content adapted from CREST.BD's Bipolar Wellness Centre, along with additional psychoeducation content and self-management strategies, all co-developed by the team of researchers and clinicians, along with multiple "peer researchers" and 2 highly engaged advisory groups with lived experience over a 2-year period. Self-management resources supporting the 99 strategies were both curated from evidence-informed sources to link out to, as well as developing new resources in-house where existing resources were not available or appropriate. After completing the QoL.BD assessment at baseline, users select QoL areas and self-management strategies to focus on. To support engagement, app users receive encouragement to self-monitor their daily QoL, mood, and sleep through reminder notifications and tailored daily affirmations written by people with BD (see Figure 1 for screenshots of the PolarUs interface).

XSL•FO
RenderX

Figure 1. Illustrations of PolarUs app content.

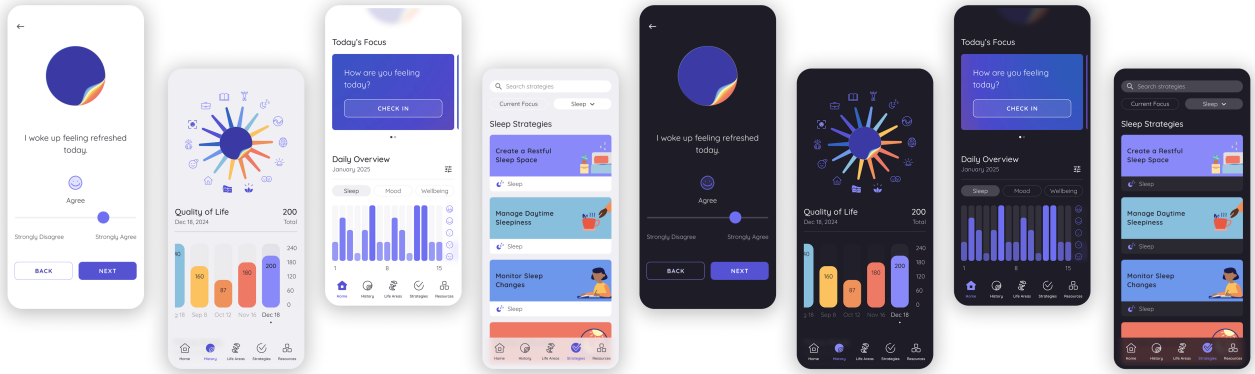

PolarUs underwent rigorous user testing before being launched. The app is available on both iOS (1.0.20; released on March 5, 2025) and Android (1.0.0; released on October 23, 2025) platforms and will remain free for downloading and usage by anyone. A mixed methods feasibility study and pilot evaluation [22] was conducted across North America with 170 participants (of whom 70.0% (n=119) identified as women, 4.7% (n=8) as nonbinary, 70.6% (n=120) as White, and 92.9% (n=158) diagnosed with bipolar type I disorder). The quantitative results showed significant improvements for the primary outcome measure (Brief QoL.BD) and several of the secondary outcome measures (depression, mania, generic QoL, and self-compassion) [28]. Qualitative interviews conducted with a purposive subset of the participants (n=25) based on patterns of app usage supported feasibility and effectiveness, with participants indicating the app supported increasing awareness of fluctuations in QoL and mood and for selecting appropriate corresponding self-management strategies [29,30]. Modifications were made based on pilot feedback, such as enhanced methods for daily mood, sleep, and QoL tracking. Most recently, the app was selected competitively for a rigorous assessment by the Mental Health Commission of Canada. One of the core 5 domains is cultural safety (cultural safety, social responsibility, and equity standards), which we strive to attain with this rigorous adaptation and evaluation process. The assessment is currently underway. Upon the successful completion of the assessment, PolarUs will be published in a National Repository.

Cultural Adaptation

Cultural adaptation is "the process of systematically modifying an evidence-based intervention to be congruent with the language, values, beliefs, and context corresponding to clients' cultural backgrounds" [31]. This is particularly important for QoL-focused interventions, as QoL is inherently subjective and understood from an individual's own perspective, which is shaped by cultural context and value systems [32]. Cultural adaptation seeks to identify components that are cultural to be tailored for target populations while maintaining fidelity, or the degree to which intervention is designed and delivered as intended to evidence-based components [33]. The cultural adaptation of existing interventions has been shown to be a cost-effective and scalable way of delivering interventions and more likely to generate positive impacts [34]. There has, however, been limited attention on adaptation methods to ensure cultural relevance, appropriateness, and ultimately their effectiveness [35,36]. Many apps for BD lack appropriate adaptation, with representatives from ethnocultural and linguistic minorities often not included in development and testing [24], limiting equitable access. It is critical to ensure that all those who are attracted to DMHIs, especially traditionally underserved groups such as those from culturally diverse communities, have access to culturally appropriate resources [37]. Strategies to enhance access by diverse communities include engaging these end users in the co-design of not just the DMHIs but also randomized controlled trial (RCT) methodologies [38].

Objectives

This study will culturally and linguistically adapt the PolarUs app to French, Chinese, and Spanish, the 3 most spoken languages in North America (Canada and the United States) after English [39,40], using qualitative advanced co-design methods and user-centered principles engaging with knowledge holders from the linguistic groups to ensure acceptability, fidelity, and responsiveness of the adapted app. The culturally adapted apps will be tested in a future RCT in North America comparing the effects of the culturally adapted app to enhanced waitlisted usual care. This work will advance comprehensive cultural adaptation and co-design methods in a DMHI more broadly, for both BD specifically and mental health in general, to leverage their responsiveness and scalability to decrease health inequities globally. This work will also produce the first culturally adapted DMHI tailored for BD using advanced co-design with end users in multiple languages.

Methods

Frameworks

Community-Based Participatory Research

This study is guided by community-based participatory research (CBPR [41]), a collaborative research approach that enables community members living with a health condition to actively and equitably participate in the full spectrum of a research study, from research design to analysis and interpretation, and communication of results [42]. In CBPR studies, community members and researchers partner equitably to create and share knowledge, making CBPR an ideal form of research to use with disadvantaged or stigmatized populations, including cultural minority communities. CBPR is also, importantly, the guiding approach for our CREST.BD research network [42,43].

XSL•FO
RenderX

Cultural Adaptation Framework

The Ecological Validity Framework of Bernal et al [44] will be used to guide the cultural adaptation process. The Ecological Validity Framework, developed originally for psychotherapy and mental health interventions, has been used often in evidence-based treatment adaptation of DMHIs [44,45]. It provides a framework for a systematic and comprehensive approach, integrating sociocultural and psychological relevance, supporting a thorough cultural adaptation process that delves into the values, concepts, and context. Importantly, it emphasizes the components of an intervention that must be modified for cultural relevance, while keeping intact the critical therapeutic components to maintain fidelity. The Ecological Validity Framework includes 8 domains to examine when adapting a psychological intervention, detailed in Table 1.

Table . Ecological validity framework [44].

Domain	Definition
Language	Translation of materials and use of culturally appropriate terms. Ensures readability, comprehension, and resonance with users.
Persons	Adaptation of who delivered the intervention and how they relate to participants. Includes therapists' or facilitators' ethnicity, credibility, and cultural sensitivity.
Metaphors	Incorporation of culturally meaningful symbols, sayings, and stories.
Content	Adaptation of topics, examples, scenarios, case studies, and materials to align with cultural experiences and priorities.
Concepts	Aligning intervention constructs with culturally understood meanings.
Goals	Ensuring intervention goals match the values and expectations of the cultural group.
Methods	Adapting how intervention components are delivered (eg, format, structure).
Context	Consideration of the broader familial, social, and environmental factors that affect engagement.

Study Design and Objectives

This study will culturally and linguistically adapt the PolarUs app to French, Chinese, and Spanish for the North American context using advanced qualitative co-design methods and user-centered design engaging with knowledge holders from the different linguistic groups. Specific objectives are to (1) culturally and linguistically adapt and advance the current beta version of the PolarUs app in French, Chinese (Mandarin), and Spanish to increase accessibility and equity and (2) build knowledge on and capacity for effective cultural adaptation of DMHIs for BD in these linguistic groups.

Advisory groups consisting of individuals with the lived experience of BD from relevant linguistic groups (French, Chinese, Spanish) will guide multiple aspects of the research process, such as providing input and establishing consensus on the cultural and linguistic adaptation process, supporting the cultural and linguistic translation process of app content and accompanying materials, reviewing and providing feedback on language-specific resources and affirmations, developing additional culturally specific resources, and testing and revising features of the culturally adapted apps. Recognizing there is diversity among people who share a language, the cultural adaptation process will emphasize conceptual equivalence and inclusivity, with translations and adaptations emphasizing meaning rather than literal wording. This is well supported by the digital format, which allows for flexible content design, including culturally varied scenarios. Adaptation will be iterative, incorporating ongoing feedback from the advisory groups and will be further supported by the involvement of advisory group members in the data analysis and development of appropriate recruitment and implementation strategies for the future RCT. Activities will support research development and advance the PolarUs app to be culturally acceptable, appropriate, and responsive to end-user needs, while maintaining fidelity to core evidence-based intervention components.

Ethical Considerations

All procedures were approved by the Behavioural Research Ethics Board at the University of British Columbia (UBC) in Canada (H25-02835). Written informed consent will be obtained from all advisory group member participants. Participants will receive CAD $55 (US $40) per meeting attended, following a standard CREST.BD peer-researcher remuneration rate of CAD $27.5 (US $20) per hour. This remuneration is inclusive of approximately 75 minutes of meeting time and 45 minutes of meeting preparation. Remuneration will be sent as an e-gift card (Amazon) or cheque within 6 weeks of each meeting. Additional remuneration may be allocated for supporting specific tasks undertaken by individual advisory group members between meetings.

Study Setting and Participants

This study will take place in North America (Canada and the United States). Research activities will be coordinated by CREST.BD, located within the Department of Psychiatry at UBC in Vancouver, BC, Canada. We seek to recruit approximately 5 participants per advisory group (one for each linguistic group) for a total of 15 participants.

XSL•FO
RenderX

The inclusion criteria are adults age 18 years or older; have a self-reported diagnosis of BD (self-report enhances equity and inclusion for those individuals who face diagnostic barriers and often aligns well with clinical records [46]); have regular access to a smartphone (a mobile phone capable of running apps; smartphone operating system requirement: iOS 13/Android 10 or later); have sufficient understanding of written and spoken English and one of French, Mandarin, or Spanish; and are residents of Canada or the United States, or have lived in Canada or the United States previously. Individuals will be excluded from participation if the inclusion criteria are not met. Individuals will also be excluded if their current mental or mental health status (eg, symptoms of hypomania or mania, psychosis, acute distress) interferes with participation in study-related activities, in the opinion of PI and senior author EEM, who is experienced in clinical studies with vulnerable populations. The research team will follow a risk management and distress protocol if participants verbalize significant risk or suicidal ideation during a meeting (conducted virtually over Zoom). When this occurs, research staff will end Zoom meetings immediately. A research staff member will guide the participant to a breakout room, while another research staff member will debrief the remaining group and offer support if needed. Next, protocol steps include acknowledging the person's distress and offering information on risk management procedures, providing a list of mental health resources, and informing the PI. If a participant abruptly ends the meeting, they will be contacted immediately, according to details (eg, their location, emergency contacts) provided in a safety questionnaire provided before the session.

Recruitment Process

The recruitment of advisory group members will occur through several avenues, including through CREST.BD's network affiliations. For example, the study will be promoted through a network blog post (Multimedia Appendix 1) and newsletter to people who have signed up for the CREST.BD e-newsletter (which has over 4000 subscribers) and will include a link for participants to sign up. In addition, we will reach out to various CREST.BD network members who are active in communities with large populations of speakers of French, Spanish, or Chinese and request their support for distributing informational (eg, study information sheets; Multimedia Appendix 2) and promotional material through their networks. Recruitment will also involve targeted discussions with individuals who have been involved previously with CREST.BD projects for their assistance in distributing within their networks. We will also use CREST.BD's social media platforms, including LinkedIn, Facebook (~2500 followers), Instagram (~1600 followers), and the WhatsApp channel for advertising using study posters. These social media posts will also contain links to a study landing page and contact information for study lead and first author LWC. All recruitment material will be in English and the 3 languages.

We will also enlist the help of our partner organizations (eg, the Canadian Network for Mood and Anxiety Treatment and Hope + Me) on this project. Finally, we will use innovative recruitment strategies to recruit racially diverse participants developed in collaboration with language-tailored community network organization partners, such as Relief, Inter-Cultural Online Health Network, and International Bipolar Foundation (French-speaking, Mandarin-speaking, and Spanish-speaking communities, respectively). Community outreach with network organizations is an effective strategy to recruit participants from specific populations, especially those who may be underrepresented or hard to reach. These strategies emphasize trust-building, cultural relevance, and collaboration. Prospective participants can express interest by completing a UBC Qualtrics [47] form or contacting LWC or other members of the team directly. Qualtrics is UBC's top-tier secure, BC Freedom of Information and Protection of Privacy Act–compliant tool for administering surveys and analyzing survey data.

Informed Consent

Participants who self-identify as being eligible will be contacted to schedule a meeting to discuss their potential participation and to confirm eligibility based on the opinion of PI EEM, who is experienced in clinical studies with vulnerable populations. Participants will be given a minimum of 7 days to decide whether they choose to participate in the project, as well as any additional time needed or requested on an individual basis. Confirmed eligible participants will be emailed a copy of the terms of reference and consent form, both available in English, French, Chinese, and Spanish. In the consent form, we will remind participants of their right not to take part in any aspect of the study and to withdraw their participation at any time, without providing a reason. Participants will be encouraged to contact us if they have questions or concerns about the study.

Study Procedure

Preparatory Work

Guided by a cultural adaptation working group comprising BD researchers and clinicians fluent in at least one of the 3 languages, PolarUs content (resources and affirmations) will be synthesized and culturally contextualized in advance for input from the advisory groups. This involves converting all self-management strategy titles and descriptions into plain language to facilitate translation, eliminating redundancies and updating resources, and identifying new opportunities for in-house content creation, working in collaboration with key community organizations such as the International Society of Bipolar Disorders for capturing diverse expertise. The English-language modifications will be incorporated into the English version.

In addition, existing resources available in the target languages will be marked for review by the advisory groups. For resources not available in the new target languages, the working group will identify core English language resources deemed essential for the implementation of the strategies for in-house creation. This process will involve the creation of the resource in English first, followed by translation using a large language model to produce first drafts in the new target languages. These drafts will then be provided to the advisory groups to verify linguistic accuracy and cultural appropriateness, capturing any necessary adaptations.

XSL•FO
RenderX

Advisory Group Work

Advisory groups will function both as a consultative body for the project and as a source of qualitative data through advisory group meetings. Meetings with each advisory group will occur semimonthly and will be analyzed as research data to capture advisory group perspectives and experiences. Meetings will be structured to generate input on the cultural and linguistic adaptation processes, with a balance between presenting information for discussion and feedback, and subsequent meetings will be structured based on advisory group feedback from the previous meeting or meetings depending on aspects of the adaptation process identified as the most important. Accordingly, guides will be developed for each meeting based on the discussion from the previous meeting. The initial meetings will focus on the cultural adaptation process. A subsequent extended meeting will center on the coproduction of the findings, with the aim of collaboratively interpreting the results through participant feedback to ensure accuracy, shared understanding, and contextual relevance. A final extended meeting will gather participants' advice on culturally appropriate marketing, recruitment, and outreach strategies for the future RCT, providing guidance to support PolarUs implementation and scale-up in the community. Using an iterative design, advancements in PolarUs app content based on qualitative findings will be made to the app to be utilized in the RCT (see Figure 2 for an overview of the advisory group work).

Figure 2. Advisory group work phases. QoL: quality of life; RCT: randomized controlled trial.

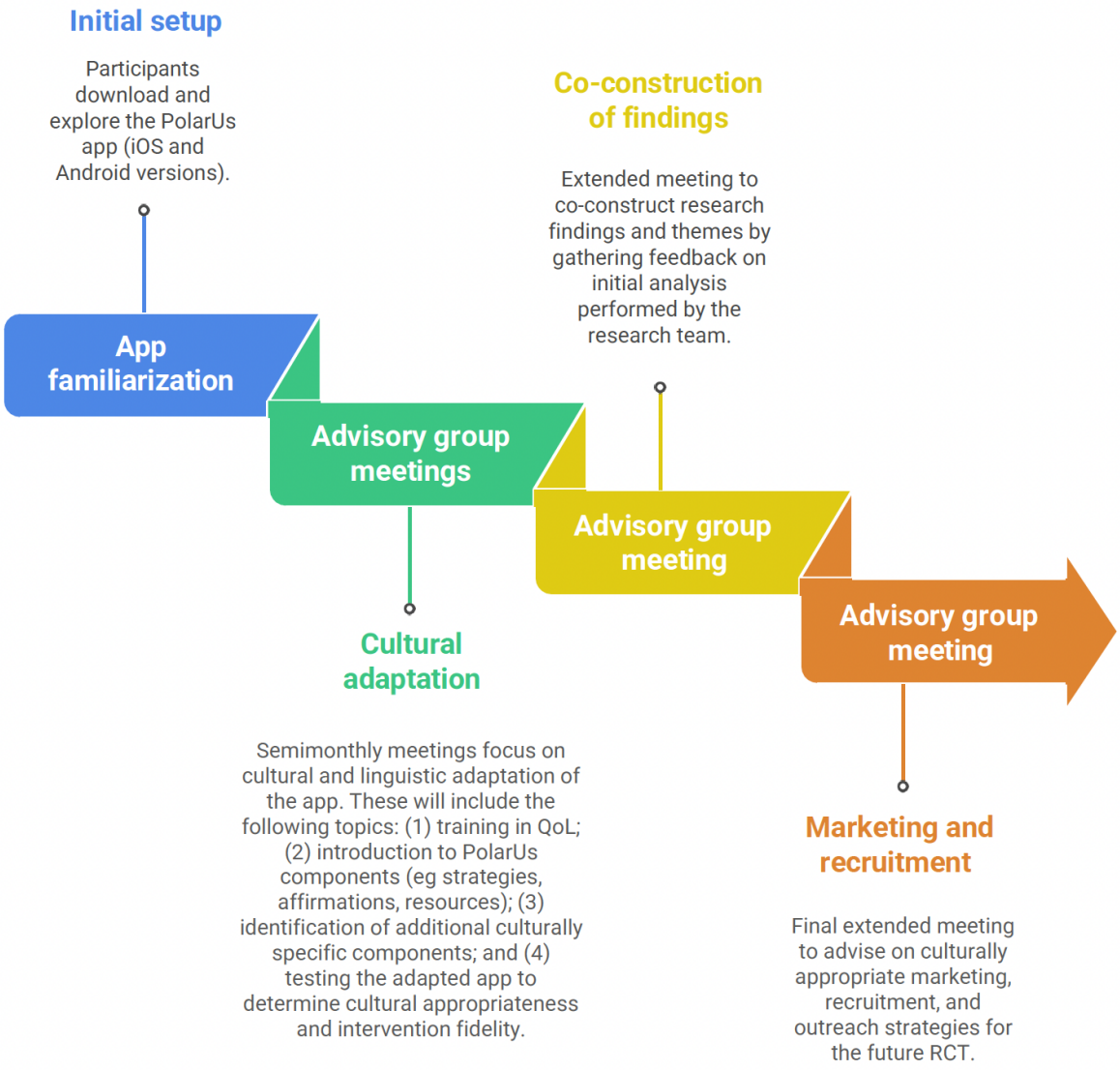

During the meetings, which will be held virtually by Zoom, participants will be reminded that UBC Zoom may collect data (usernames, email addresses), which are stored on a secure UBC server. Participants will be told they have the option of turning their camera on or off and using a nickname instead of their real name. Meetings will be recorded, transcribed, and translated. Transcription and translation will be performed in-house by multilingual research assistants trained on project objectives and key terminology.

An integrated knowledge translation approach guided by CBPR principles will be used to ensure the inclusion of knowledge holders throughout the research process, from cultural adaptation to knowledge dissemination. New knowledge outputs will be created and disseminated using various platforms, including

XSL•FO
RenderX

open-access publications and digital tools (eg, the CREST.BD TalkBD podcast), in consultation with the advisory groups. End-user involvement in this study will be supported through a synergistic combination of community-based research and best practices with respect to user-centered design.

Data Analysis

Reflexive thematic analysis will be used for the qualitative data to identify and analyze common themes across transcripts to capture important patterns and meaning in the data [48]. Reflexive thematic analysis emphasizes the researcher's role in constructing meaning from the data, and as such, the involvement of multiple coders for intercoder agreement is discouraged [49]. Transcripts will first be familiarized by LWC through repeated reading, followed by initial open coding to identify meaningful units of text. An initial coding framework will be developed by LWC both deductively, informed by the Ecological Validity Framework, and inductively, through analysis of the qualitative research data. This initial coding framework will be shared with the advisory groups for their input and feedback, supporting the integration of diverse perspectives and deeper interpretation of the data. Following the integration of advisory group feedback, codes will be reviewed and refined by LWC, with related codes grouped into themes. Themes will be examined for coherence and distinctiveness and defined. The final report will be shared with advisory group members for their input. This process will be conducted separately in each of the 3 languages, recognizing that there may be different culturally influenced nuances to data interpretation.

Reflexivity will involve an explicit consideration and discussion of how assumptions are shaped in the analysis through rigorous and transparent engagement with the data [48]. Reflexivity, aided by detailed meeting notes taken by LWC and any other members of the research team and during the coding process, will be used to help ensure the trustworthiness and rigor of the analysis and research findings [48,50]. These strategies contribute an audit trail of the analysis process to support reproduction by other researchers following the same processes [51]. The audit trail will also include detailed records of the study methods and procedures. The NVivo software [52] (version 15; QSR International) will be used for all qualitative analysis.

Data Management, Confidentiality, and Security

The participant sign-up form will be administered via Qualtrics using our university's Qualtrics account. Qualtrics, designed to gather sensitive questionnaire data online, incorporates several procedures for minimizing the risk of a breach of confidentiality. Qualtrics servers are protected by high-end firewall systems, and vulnerability scans are performed regularly. Complete penetration tests are performed yearly. All services have quick failover points and redundant hardware, and complete backups are performed nightly. Qualtrics survey data will be downloaded locally to the research coordinator's password–protected and encrypted computer to be transferred to a secure UBC server. The data will only be accessible to authorized team members. Meeting recordings may contain audio and video information. Audio recordings and notes taken during the meetings will be stored on a secure UBC server, and all data will be deidentified and password protected. Recordings may be downloaded locally to the research coordinator's password–protected and encrypted computer to be transferred to a secure UBC server. They will only be accessible to authorized team members.

We have minimized the collection of information that would link a participant's identity with their data (eg, names, IP addresses, and addresses will not be part of the data file). Both the consent form and other identifiable participant data will be stored in a password-protected file separately from study data, which will be linked only to an assigned participant identification number. The data from this study will be coded by the project name and assigned a participant identification number. A master list will be created linking codes to names. This list will be password protected, encrypted, and stored separately from the rest of the data. Only authorized team members will have access to this list.

While participants will be using the app to support cultural adaptation processes, their app data will not be stored or used for any analysis purposes. However, for reference, the PolarUs app uses end-to-end encryption for all data transferred between our server (located at UBC) and the app. Firebase is utilized for real-time database management, authentication, and cloud storage, supporting seamless user experience and data security. All app data will be deidentified, encrypted, password protected, and stored using end-to-end encryption on a secure Canadian server. All paper study files will be retained for a minimum of 5 years in a locked file cabinet in the Department of Psychiatry at UBC. Digital files will be stored securely on UBC OneDrive for at least 5 years after publication in accordance with UBC policy. Any data containing identifiers will be destroyed 5 years after study findings are published; digital recordings will be erased and paper copies shredded.

Results

As of January 2026, we are nearing completion of the preparatory work, including converting all self-management titles and descriptions into plain language to facilitate translation, updating resources, creating new in-house content, and identifying core English language resources for translation. We have also enrolled 7 participants in the advisory groups, with advisory group activities expected to commence in February 2026. The results are expected to be published in fall 2026.

Discussion

Strengths

Worldwide, there is a persistent and growing mental health treatment gap [10,53], with disparities between and within countries. Addressing this inequality is a global health priority identified by the World Health Organization in its 2022 World Mental Health Report [54]. Contributing factors include a shortage and maldistribution of health human resources and stigma that impacts help-seeking [55]. This is despite there being effective treatment for many mental disorders, including for severe conditions such as BD. Furthermore, cultural minorities tend to have a higher prevalence of common mental

XSL•FO
RenderX

disorders, yet they are less likely to seek professional help due to language barriers, a lack of trust in traditional health care systems, and stigma [55,56]. Cultural minorities are also more likely to experience disparities in treatment access when they do seek care and are more likely to drop out of care [55]. French-speaking, Chinese-speaking, and Spanish-speaking populations are cultural and linguistic minorities in North America and face similar challenges [57-59].

DMHIs present as a promising way to help bridge the treatment gap, particularly for cultural minorities, considering their accessibility, affordability, and ability to support anonymity. Research consistently supports the efficacy of culturally tailored in-person interventions compared to nontailored interventions, and this is expected to also be critical for DMHIs [55]. A recent systematic review and meta-analysis showed that DMHIs culturally adapted for racial and ethnic minorities were effective compared to control conditions (waitlist and treatment-as-usual) [60]. However, evidence remains limited for culturally adapted DMHIs for these populations [61,62], along with evidence-based guidance for how to adapt them [26]. The findings from this study will support the first BD-tailored DMHI to be adapted for different cultural contexts closely involving people with lived experience from the respective cultural communities to support equitable mental health care access. This will help ensure that DMHIs are codeveloped not only by users with lived experience but by those who experience BD according to the uniqueness of their cultural backgrounds.

Poor user engagement and adherence to DMHIs have been widely attributed to insufficient consultation with end users, particularly individuals from cultural minority groups, undermining their therapeutic potential. Recent evidence of 158,930 individuals who downloaded the MoodTools app showed that nearly 50% did not log in a second time and that a third of the active sessions lasted only between 0 and 10 seconds [63]. Promisingly, a recent qualitative study involving interviews with people with BD (N=25) who were given 3-month access to PolarUs generated illuminating findings on feature-related engagement factors (motivations, salience, perceived effort) and contextual influences (smartphone ecosystem, daily life, mood symptoms, involvement in a research study) [29]. This study highlighted the importance of capturing lived experience to better understand engagement with app-based DMHIs and the importance of context. Consistent with this, user involvement has been cited as the gold standard for culturally adapting DMHIs. Consistent with this, user involvement has been cited as the gold standard approach to the cultural adaptation of DMHIs [64].

While the PolarUs app was developed over 2 years in close collaboration with 2 highly engaged advisory groups of people with lived experience, clinicians, and researchers, the needs and preferences of individuals from diverse cultural backgrounds need specific consideration. For cultural minorities, the challenge of engagement with DMHIs is further complicated by language and cultural barriers. DMHIs, therefore, need to be tailored for inclusivity, including adaptation for ethnic minorities and the inclusion of customizable content that aligns with users' culture and values. This study will provide critical evidence using advanced co-design and CBPR principles in the cultural adaptation of a comprehensive DMHI for BD that has been shown to be effective in a nonrandomized pilot study. Lived experience engagement in this cultural adaptation study will ensure PolarUs is responsive and acceptable to the needs of diverse people living with BD in different contexts.

The findings and developments from this study will support the upcoming RCT, the first study to test the effectiveness of a new evidence-based BD-tailored app focused on QoL outcomes, and the first DMHI tailored for BD to be culturally adapted using advanced co-design with end users in multiple languages. Should results demonstrate PolarUs' effectiveness, the linguistically adapted versions of the app can be scaled up to other French-speaking, Spanish-speaking, and Chinese-speaking communities worldwide, with appropriate additional language and cultural adaptations. Importantly, the research processes will provide not only critical cultural adaptation evidence for the PolarUs app for BD itself but will advance comprehensive cultural adaptation and co-design methods in DMHI research more broadly, supporting widespread scale-up and application to help decrease mental health inequities globally.

Limitations and Directions for Future Research

This study has several challenges that may limit its broad applicability. The first is that the language adaptations are focused on the North American context (ie, Latin American Spanish, rather than Castilian Spanish; Canadian French vs Metropolitan French) and that Chinese is focused on Mandarin, which is very different from Cantonese. This may limit the adapted app's generalizability to other French-speaking, Spanish-speaking, and Chinese-speaking communities given linguistic diversity and regional variants. However, our study is focused on the North American population, with efforts currently targeted at producing a feasible, culturally appropriate, and ultimately clinically effective DMHI for French-speaking, Spanish-speaking, and Chinese-speaking communities in North America.

We also acknowledge that this study will engage bilingual advisory group members (ie, fluent in one of the three languages in addition to English), yet the perspectives of bilingual users may not fully reflect the needs of monolingual or non-English speaking users, whose levels of acculturation, maintenance of cultural traditions, and development or adoption of new cultural norms shaped by migration experiences may differ. However, there is a pressing need for improved access to care for BD, and our app seeks to first address that treatment gap. We anticipate that with the robust cultural and linguistic adaptation processes in place, this study represents an important first step. Further adaptations in these languages, both in other countries and with monolingual or non-English speaking users, could be carried out with relatively limited resource requirements using the same protocol. In addition, there may be opportunities for monolingual app users in these language groups to be involved in usability testing of the adapted app. Looking even further ahead, widespread scale-up and implementation to advance equity, accessibility, and cultural responsiveness of BD treatment and care globally is our ultimate aim. We have already submitted applications for scale-up to other contexts, including in Africa.

XSL•FO
RenderX

An additional consideration is that the cultural and linguistic adaptation methods described in this protocol differ from more resource-intensive translation approaches commonly recommended for patient-reported outcome measures, such as the Professional Society for Health Economics and Outcomes Research (ISPOR) guidelines [65], which emphasize multiple professional translators and forward-back translation processes. While these approaches are well suited for the translation of discrete instruments, their application to complex, large-scale DMHIs, characterized by dynamic content and interactivity, may pose feasibility challenges. Accordingly, this protocol involves a more pragmatic and participatory approach that prioritizes cultural relevance, contextual appropriateness, and lived experience input across multiple languages.

Conclusions

Globally, substantial gaps persist in both the availability and uptake of mental health care. DMHIs offer an important opportunity to help address this gap and improve population mental health. However, their widespread and equitable impact is often limited by the lack of co-design and cultural tailoring. This study will add to the growing evidence base in digital mental health by demonstrating how cultural adaptation using CBPR and co-design methods of a DMHI for BD involving people with lived experience from different cultural linguistic groups can support widespread and equitable application and engagement, particularly from traditionally underserved communities. If the adapted PolarUs app is found to be effective in the future clinical trial, research findings can contribute further evidence that a low-barrier intervention for BD can help improve mood symptoms and QoL in people living with BD across different cultural contexts.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the GAIDeT (Generative AI Delegation Taxonomy; 2025), the following task was delegated to GAI tools under full human supervision: visualization. The GAI tool used was Napkin AI (v1.0.0). Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. This declaration was submitted by LWC. We used Napkin AI (v1.0.0) to generate Figure 2.

Funding

This study is funded by the Canadian Institutes of Health Research Project Grant (RN538779). DR-C is supported by a salary award through the Fonds de Recherche du Québec Santé (#367157).

Data Availability

Data from the cultural adaptation study will be available upon request.

Authors' Contributions

Conceptualization: EEM (equal), LWC (lead)
Funding acquisition: EEM (lead), EM, JKM, MDP, SJB
Methodology: DRC, EEM, EM, JKM, LWC (lead)
Project administration: EEM, LWC (lead)
Resources: EEM
Supervision: EEM
Writing – original draft: LWC
Writing – review & editing: DR-C, EM, JKM, LWC (lead), MDP, SJB

Conflicts of Interest

DR-C reports personal fees from Otsuka Canada outside the submitted work. EM has received an honorarium for advising on the development of unrelated educational materials for Neurotorium, an online nonprofit educational platform supported by the Lundbeck Foundation. The other authors declare no conflicts of interest.

Multimedia Appendix 1
Recruitment blog posts (English, French, Chinese, and Spanish).
[PDF File, 16080 KB - resprot_v15i1e92600_app1.pdf]

Multimedia Appendix 2
Study information sheets (English, French, Chinese, and Spanish).
[PDF File, 1057 KB - resprot_v15i1e92600_app2.pdf]

XSL•FO
RenderX

Peer Review Report 1
Peer review report by the Randomized Controlled Trials 2 Committee, Canadian Institutes of Health Research (CIHR).
[PDF File, 310 KB - resprot_v15i1e92600_app3.pdf]

References

1. Bipolar disorder. World Health Organization. 2025 Sep 8. URL: https://www.who.int/news-room/fact-sheets/detail/bipolar-disorder [accessed 2026-03-20]
2. Merikangas KR, Jin R, He JP, et al. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Arch Gen Psychiatry 2011 Mar;68(3):241-251. [doi: 10.1001/archgenpsychiatry.2011.12] [Medline: 21383262]
3. Bolton S, Warner J, Harriss E, Geddes J, Saunders KEA. Bipolar disorder: trimodal age-at-onset distribution. Bipolar Disord 2021 Jun;23(4):341-356. [doi: 10.1111/bdi.13016] [Medline: 33030292]
4. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR): American Psychiatric Association; 2022. URL: https://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425787 [accessed 2014-09-12]
5. Morton E, Michalak EE, Murray G. What does quality of life refer to in bipolar disorders research? A systematic review of the construct's definition, usage and measurement. J Affect Disord 2017 Apr 1;212:128-137. [doi: 10.1016/j.jad.2017.01.026] [Medline: 28160685]
6. Oliva V, Fico G, De Prisco M, Gonda X, Rosa AR, Vieta E. Bipolar disorders: an update on critical aspects. Lancet Reg Health Eur 2025 Jan;48:101135. [doi: 10.1016/j.lanepe.2024.101135] [Medline: 39811787]
7. Eiring Ø, Nylenna M, Nytrøen K. Patient-important outcomes in the long-term treatment of bipolar disorder: a mixed-methods approach investigating relative preferences and a proposed taxonomy. Patient 2016 Apr;9(2):91-102. [doi: 10.1007/s40271-015-0128-x] [Medline: 25990222]
8. Michalak EE, Suto MJ, Barnes SJ, et al. Effective self-management strategies for bipolar disorder: a community-engaged Delphi Consensus Consultation study. J Affect Disord 2016 Dec;206:77-86. [doi: 10.1016/j.jad.2016.06.057] [Medline: 27466745]
9. Yatham LN, Kennedy SH, Parikh SV, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. Bipolar Disord 2018 Mar;20(2):97-170. [doi: 10.1111/bdi.12609] [Medline: 29536616]
10. Chakrabarti S. Digital psychiatry in low-and-middle-income countries: new developments and the way forward. World J Psychiatry 2024 Mar 19;14(3):350-361. [doi: 10.5498/wjp.v14.i3.350] [Medline: 38617977]
11. Fontanella CA, Nguyen KM, Llamocca EN, et al. Racial and ethnic disparities in treatment of youth with bipolar disorders. J Am Acad Child Adolesc Psychiatry 2026 Feb;65(2):279-290. [doi: 10.1016/j.jaac.2025.02.018] [Medline: 40188867]
12. Akinhanmi MO, Biernacka JM, Strakowski SM, et al. Racial disparities in bipolar disorder treatment and research: a call to action. Bipolar Disord 2018 Sep;20(6):506-514. [doi: 10.1111/bdi.12638] [Medline: 29527766]
13. Oedegaard CH, Berk L, Berk M, et al. An ISBD perspective on the sociocultural challenges of managing bipolar disorder: a content analysis. Aust N Z J Psychiatry 2016 Nov;50(11):1096-1103. [doi: 10.1177/0004867416668790] [Medline: 27742913]
14. Grande I, Berk M, Birmaher B, Vieta E. Bipolar disorder. Lancet 2016 Apr 9;387(10027):1561-1572. [doi: 10.1016/S0140-6736(15)00241-X] [Medline: 26388529]
15. Nordentoft M, Mortensen PB, Pedersen CB. Absolute risk of suicide after first hospital contact in mental disorder. Arch Gen Psychiatry 2011 Oct;68(10):1058-1064. [doi: 10.1001/archgenpsychiatry.2011.113] [Medline: 21969462]
16. Jin H, McCrone P. Cost-of-illness studies for bipolar disorder: systematic review of international studies. Pharmacoeconomics 2015 Apr;33(4):341-353. [doi: 10.1007/s40273-014-0250-y] [Medline: 25576148]
17. Cloutier M, Greene M, Guerin A, Touya M, Wu E. The economic burden of bipolar I disorder in the United States in 2015. J Affect Disord 2018 Jan 15;226:45-51. [doi: 10.1016/j.jad.2017.09.011] [Medline: 28961441]
18. Fletcher K, Murray G. Beyond pharmacotherapy for bipolar disorder: the role of adjunctive psychological treatments. Psychiatr Ann 2019 Sep;49(9):399-404. [doi: 10.3928/00485713-20190808-01]
19. Naslund JA, Aschbrenner KA, Araya R, et al. Digital technology for treating and preventing mental disorders in low-income and middle-income countries: a narrative review of the literature. Lancet Psychiatry 2017 Jun;4(6):486-500. [doi: 10.1016/S2215-0366(17)30096-2] [Medline: 28433615]
20. Nicholas J, Larsen ME, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. J Med Internet Res 2015 Aug 17;17(8):e198. [doi: 10.2196/jmir.4581] [Medline: 26283290]
21. Morton E, Nicholas J, Yang L, et al. Evaluating the quality, safety, and functionality of commonly used smartphone apps for bipolar disorder mood and sleep self-management. Int J Bipolar Disord 2022 Apr 4;10(1):10. [doi: 10.1186/s40345-022-00256-6] [Medline: 35368207]
22. Michalak EE, Barnes SJ, Morton E, et al. Supporting self-management and quality of life in bipolar disorder with the PolarUs App (Alpha): protocol for a mixed methods study. JMIR Res Protoc 2022 Aug 4;11(8):e36213. [doi: 10.2196/36213] [Medline: 35925666]

XSL•FO
RenderX

23. Majid S, Reeves S, Figueredo G, et al. The extent of user involvement in the design of self-tracking technology for bipolar disorder: literature review. JMIR Ment Health 2021 Dec 20;8(12):e27991. [doi: 10.2196/27991] [Medline: 34931992]
24. Whitehead L, Talevski J, Fatehi F, Beauchamp A. Barriers to and facilitators of digital health among culturally and linguistically diverse populations: qualitative systematic review. J Med Internet Res 2023 Feb 28;25:e42719. [doi: 10.2196/42719] [Medline: 36853742]
25. Lagan S, Ramakrishnan A, Lamont E, Ramakrishnan A, Frye M, Torous J. Digital health developments and drawbacks: a review and analysis of top-returned apps for bipolar disorder. Int J Bipolar Disord 2020 Dec 1;8(1):39. [doi: 10.1186/s40345-020-00202-4] [Medline: 33259047]
26. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
27. Michalak EE, Murray G, Collaborative RESearch Team to Study Psychosocial Issues in Bipolar Disorder (CREST.BD). Development of the QoL.BD: a disorder-specific scale to assess quality of life in bipolar disorder. Bipolar Disord 2010 Nov;12(7):727-740. [doi: 10.1111/j.1399-5618.2010.00865.x] [Medline: 21040290]
28. Michalak EE, Morton EE, Meyer D, Murray G, O Brien HL, Barnes SJ. Empowering quality of life monitoring and self-management in bipolar disorder: pilot evaluation of the PolarUs app. Bipolar Disord 2026 May;28(3):e70096. [doi: 10.1111/bdi.70096] [Medline: 41784296]
29. Morton E, Hole R, O'Brien H, Li L, Barnes SJ, Michalak EE. What influences engagement with a bipolar disorder self-management app? A qualitative investigation of use of the PolarUs app. In: Ayatollahi H, editor. PLOS Digit Health 2025 Oct;4(10):e0001017. [doi: 10.1371/journal.pdig.0001017] [Medline: 41071762]
30. Morton E, Hole R, O'Brien H, Li LC, Barnes SJ, Michalak EE. The experience of self-monitoring using the PolarUs bipolar disorder self-management app: a qualitative report of impacts and unmet needs. J Affect Disord 2025 Aug 15;383:374-384. [doi: 10.1016/j.jad.2025.04.107] [Medline: 40286926]
31. Bernal G, Jiménez-Chafey MI, Domenech Rodríguez MM. Cultural adaptation of treatments: a resource for considering culture in evidence-based practice. Prof Psychol Res Pr 2009 Aug;40(4):361-368. [doi: 10.1037/a0016401]
32. WHOQOL: measuring quality of life. World Health Organization. URL: https://www.who.int/toolkits/whoqol [accessed 2026-01-20]
33. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. Implement Sci 2007 Nov 30;2:40. [doi: 10.1186/1748-5908-2-40] [Medline: 18053122]
34. Todd NJ, Jones SH, Lobban FA. What do service users with bipolar disorder want from a web-based self-management intervention? A qualitative focus group study. Clin Psychol Psychother 2013;20(6):531-543. [doi: 10.1002/cpp.1804] [Medline: 22715161]
35. Eustis EH, LoPresti J, Aguilera A, Schueller SM. Cultural responsivity in technology-enabled services: integrating culture into technology and service components. J Med Internet Res 2023 Oct 3;25(1):e45409. [doi: 10.2196/45409] [Medline: 37788050]
36. Lal S, Gleeson J, Malla A, et al. Cultural and contextual adaptation of an eHealth intervention for youth receiving services for first-episode psychosis: adaptation framework and protocol for Horyzons-Canada Phase 1. JMIR Res Protoc 2018 Apr 23;7(4):e100. [doi: 10.2196/resprot.8810] [Medline: 29685867]
37. Torous J, Jän Myrick K, Rauseo-Ricupero N, Firth J. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. JMIR Ment Health 2020 Mar 26;7(3):e18848. [doi: 10.2196/18848] [Medline: 32213476]
38. Friis-Healy EA, Nagy GA, Kollins SH. It Is Time to REACT: opportunities for digital mental health apps to reduce mental health disparities in racially and ethnically minoritized groups. JMIR Ment Health 2021 Jan 26;8(1):e25456. [doi: 10.2196/25456] [Medline: 33406050]
39. What languages do we speak in the United States? United States Census Bureau. 2022. URL: https://www.census.gov/library/stories/2022/12/languages-we-speak-in-united-states.html?utm_source=chatgpt.com [accessed 2026-01-20]
40. The great Canadian language mash-up. Statistics Canada. 2025. URL: https://www.statcan.gc.ca/o1/en/plus/7869-great-canadian-language-mash?utm_source=chatgpt.com [accessed 2026-01-20]
41. Collins SE, Clifasefi SL, Stanton J, et al. Community-based participatory research (CBPR): towards equitable involvement of community in psychology research. Am Psychol 2018 Oct;73(7):884-898. [doi: 10.1037/amp0000167] [Medline: 29355352]
42. Michalak EE, Jones S, Lobban F, et al. Harnessing the potential of community-based participatory research approaches in bipolar disorder. Int J Bipolar Disord 2016 Dec;4(1):4. [doi: 10.1186/s40345-016-0045-5] [Medline: 26856996]
43. Michalak EE, Morton E, Barnes SJ, Hole R, Murray G, CREST.BD. Supporting self-management in bipolar disorder: mixed-methods knowledge translation study. JMIR Ment Health 2019 Apr 15;6(4):e13493. [doi: 10.2196/13493] [Medline: 30985287]
44. Bernal G, Bonilla J, Bellido C. Ecological validity and cultural sensitivity for outcome research: issues for the cultural adaptation and development of psychosocial treatments with Hispanics. J Abnorm Child Psychol 1995 Feb;23(1):67-82. [doi: 10.1007/BF01447045] [Medline: 7759675]

XSL•FO
RenderX

45. Perera C, Salamanca-Sanabria A, Caballero-Bernal J, et al. No implementation without cultural adaptation: a process for culturally adapting low-intensity psychological interventions in humanitarian settings. Confl Health 2020;14:46. [doi: 10.1186/s13031-020-00290-0] [Medline: 32684948]
46. Rutter LA, Howard J, Lakhan P, Valdez D, Bollen J, Lorenzo-Luaces L. "I haven't been diagnosed, but I should be"—Insight into self-diagnoses of common mental health disorders: cross-sectional study. JMIR Form Res 2023 Jan 13;7:e39206. [doi: 10.2196/39206] [Medline: 36637885]
47. UBC survey tool. The University of British Columbia. URL: https://it.ubc.ca/services/teaching-learning-tools/survey-tool [accessed 2026-03-20]
48. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2023;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
49. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. Qual Quant 2022 Jun;56(3):1391-1412. [doi: 10.1007/s11135-021-01182-y]
50. Krefting L. Rigor in qualitative research: the assessment of trustworthiness. Am J Occup Ther 1991 Mar;45(3):214-222. [doi: 10.5014/ajot.45.3.214] [Medline: 2031523]
51. Lincoln YS, Guba EG. Naturalistic Inquiry: Sage Publications; 1985. URL: https://books.google.co.in/books?id=2oA9aWlNeooC [accessed 2026-03-20]
52. NVivo. Lumivero. URL: https://lumivero.com/products/nvivo/ [accessed 2025-11-07]
53. Carter H, Araya R, Anjur K, Deng D, Naslund JA. The emergence of digital mental health in low-income and middle-income countries: a review of recent advances and implications for the treatment and prevention of mental disorders. J Psychiatr Res 2021 Jan;133:223-246. [doi: 10.1016/j.jpsychires.2020.12.016] [Medline: 33360867]
54. World mental health report: transforming mental health for all. : World Health Organization; 2022 URL: https://iris.who.int/server/api/core/bitstreams/40e5a13a-fe50-4efa-b56d-6e8cf00d5bfa/content [accessed 2025-11-11]
55. Torous J, Linardon J, Goldberg SB, et al. The evolving field of digital mental health: current evidence and implementation issues for smartphone apps, generative artificial intelligence, and virtual reality. World Psychiatry 2025 Jun;24(2):156-174. [doi: 10.1002/wps.21299] [Medline: 40371757]
56. Clement S, Schauman O, Graham T, et al. What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. Psychol Med 2015 Jan;45(1):11-27. [doi: 10.1017/S0033291714000129] [Medline: 24569086]
57. Forcén FE, Vélez Flórez MC, Bido Medina R, et al. Deconstructing cultural aspects of mental health care in Hispanic/Latinx people. Psychiatr Ann 2023 Mar;53(3):127-132. [doi: 10.3928/00485713-20230215-02] [Medline: 37781171]
58. Park J, Lee S, Park G, Woodward S. Mental health help-seeking behaviours of East Asian immigrants: a scoping review. Eur J Psychotraumatol 2025 Dec;16(1):2514327. [doi: 10.1080/20008066.2025.2514327] [Medline: 40631378]
59. de Moissac D, Bowen S. Impact of language barriers on quality of care and patient safety for official language minority Francophones in Canada. J Patient Exp 2019 Mar;6(1):24-32. [doi: 10.1177/2374373518769008] [Medline: 31236448]
60. Ellis DM, Draheim AA, Anderson PL. Culturally adapted digital mental health interventions for ethnic/racial minorities: a systematic review and meta-analysis. J Consult Clin Psychol 2022 Oct;90(10):717-733. [doi: 10.1037/ccp0000759] [Medline: 36227330]
61. Jiwani Z, Tatar R, Dahl C, et al. Examining equity in access and utilization of a freely available meditation app. NPJ Ment Health Res 2023;2:5. [doi: 10.1038/s44184-023-00025-y] [Medline: 37159797]
62. Danese A, Martsenkovskyi D, Remberk B, et al. Scoping review: digital mental health interventions for children and adolescents affected by war. J Am Acad Child Adolesc Psychiatry 2025 Feb;64(2):226-248. [doi: 10.1016/j.jaac.2024.02.017] [Medline: 38735431]
63. Su L, Anderson PL. User behavior of a publicly available, free-to-use, self-guided mHealth app for depression: observational study in a global sample. JMIR Form Res 2022 Oct 25;6(10):e35538. [doi: 10.2196/35538] [Medline: 36282559]
64. Nittas V, Chavez SJ, Daniore P. Current practice and expert perspectives on cultural adaptations of digital health interventions: qualitative study. JMIR mHealth uHealth 2025 Jul 18;13:e59965. [doi: 10.2196/59965] [Medline: 40680174]
65. Wild D, Grove A, Martin M, et al. Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value Health 2005;8(2):94-104. [doi: 10.1111/j.1524-4733.2005.04054.x] [Medline: 15804318]

Abbreviations

BD: bipolar disorder
CBPR: community-based participatory research
CREST.BD: Collaborative Research Team to Study Psychosocial Issues in BD
DMHI: digital mental health intervention
ISPOR: The Professional Society for Health Economics and Outcomes Research
QoL: quality of life
QoL.BD: Quality of Life in BD

XSL•FO
RenderX

RCT: randomized controlled trial
UBC: University of British Columbia

Edited by J Sarvestan; submitted 03.Feb.2026; this is a non–peer-reviewed article;accepted 06.Mar.2026; published 08.Apr.2026.

Please cite as:
Chau LW, Murphy JK, Morton E, Provencher MD, Raucher-Chene D, Barnes SJ, Michalak EE
Cultural Adaptation of a Digital Mobile App for Bipolar Disorder (PolarUs): Protocol for a Qualitative Co-Design Study
JMIR Res Protoc 2026;15:e92600
URL: https://www.researchprotocols.org/2026/1/e92600
doi:10.2196/92600

© Leena W Chau, Jill K Murphy, Emma Morton, Martin D Provencher, Delphine Raucher-Chene, Steven J Barnes, Erin E Michalak. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

All-Hazards Vulnerability and Adaptation Assessment of Canadian Kidney Care Systems: Protocol for a Qualitative Study

Saly El Wazze, MA; Shinjini Mondal, PhD; Shaifali Sandal, MD

Division of Nephrology, Department of Medicine, McGill University Health Centre, 1001 Boul Decarie D05-7160, Montreal, QC, Canada

Corresponding Author:
Shaifali Sandal, MD
Division of Nephrology, Department of Medicine, McGill University Health Centre, 1001 Boul Decarie D05-7160, Montreal, QC, Canada

Abstract

Background: Health and health care systems are progressively challenged by more frequent natural and human-caused hazards. A hazard becomes a disaster when vulnerability and exposure interact in complex ways with system capacity to trigger major disruptions in the functioning of a community or society. Patients with kidney diseases, especially those on dialysis, are particularly vulnerable.

Objective: We will conduct a vulnerability and adaptation assessment (VAA) of Canadian provincial health systems and individual kidney care networks (KCNs) with respect to natural and human-caused disasters and emergencies. Our research aims to map existing KCNs across Canada, contextualize the current and future vulnerability of health systems and patients to disasters and emergencies, and identify and prioritize adaptation measures in partnership with key interest holders.

Methods: We are conducting a bottom-up VAA following a qualitative descriptive approach and using the framework previously developed as an operational guide. We started by conducting an environmental scan and document review and identified 75 KCNs across Canada. After collecting directives and policies on disaster management from these KCNs, we will conduct a content analysis guided by our framework in disaster management. We are now conducting semistructured interviews with kidney health care professionals across Canada recruited using purposive and snowballing techniques to better understand their perspectives, expertise, and lived experiences. We will conduct thematic analysis using an inductive-deductive approach guided by the framework. Findings from these multiple data sources will then be triangulated to generate a robust VAA.

Results: We want to identify the existing operational and human vulnerabilities and the risks associated, as well as explore lessons learned, to develop needed adaptations. Emerging data reflect a variety of experiences across programs and provinces. The findings will explore analysis at the program, provincial, and health system levels. The environmental scan was initiated in November 2024, and the semistructured interviews started in February 2025. We have recruited 71 participants from 55 KCNs. Interviews are ongoing, and coding has been initiated simultaneously. The first VAA focusing on in-center hemodialysis services is expected to conclude in April 2026. These findings are projected to be written up by May 2026, and manuscripts are expected to be submitted for publication to peer-reviewed journals. Thereafter, similar analyses will be conducted focusing on other kidney replacement therapies and pediatric programs.

Conclusions: This is the first VAA of KCNs across Canada. Findings will provide a critical foundation for understanding vulnerability, supporting strategic planning, and guiding adaptation measures that can strengthen health system resilience and mitigate associated risks of disasters and emergencies to patients with kidney diseases.

International Registered Report Identifier (IRRID): DERR1-10.2196/90059

(JMIR Res Protoc 2026;15:e90059) doi:10.2196/90059

KEYWORDS

vulnerability and adaptation assessment; disasters; emergencies; nephrology; dialysis; Canada; patients

Introduction

Background

More frequent hazards are progressively challenging health and health care systems [1,2]. Hazards can be natural (earthquakes, tsunamis, cyclones, extreme temperatures, and floods), biological (disease outbreaks), technological (explosions and infrastructure failures), or societal (conflict, wars, and humanitarian emergencies) [3]. Their impact reflects the interplay between the exposure and the vulnerability of the affected population [4-7]. A hazard becomes a disaster when these 2 factors interact in complex ways with system capacity to trigger a major disruption in the functioning of a community or society [8]. In contrast, emergencies involve hazardous events that do not precipitate large-scale societal disruption [8]. Several populations are more vulnerable to the impacts of these disasters

XSL•FO
RenderX

and emergencies, and patients with kidney diseases in particular experience a disproportionately higher risk with hazard exposure [9].

The survival of those with end-stage kidney disease who receive dialysis depends on the safe provision of this therapy multiple times a week [10-12]. Receiving dialysis is dependent on intact and functioning infrastructure, the availability of medical personnel, and the supply of medical equipment and medications, all of which can be disrupted during a disaster or emergency setting [11,13-25]. Surveys of health care providers such as physicians, nurses, and technicians have reported challenges they experienced in getting to a dialysis center and providing services during or following disasters [20,22-24]. Several studies have reported disruptions in dialysis care delivered to patients in such settings [10,20,22,24,26-28]. For example, in a survey of patients impacted by Hurricane Sandy in 2012, a total of 26.3% of participants reported missing dialysis sessions during the storm, and 8% reported missing 3 or more sessions [26,29]. Similarly, postdisaster analyses have reported an increased risk of emergency department visits, hospitalizations, and short-term mortality among patients receiving hemodialysis exposed to a disaster [27,30,31]. All-cause mortality with hurricane exposure among 187,388 patients undergoing dialysis in the United States from 1997 to 2017 increased by 13% compared to those who were not exposed [31]. Furthermore, vulnerability is compounded by both medical and socioeconomic factors, particularly given that kidney disease disproportionately affects racial and ethnic minority groups and those experiencing structural disadvantages across several other social determinants of health [32-39].

The Canadian Disaster Database reported 424 climate-, conflict-, and technology-related disasters from 2000 to 2020 that have resulted in 1741 deaths and 12,792 injuries, displaced 603,549 individuals, and cost more than CAD $19 billion (US $14 billion) [40]. Anecdotal experiences and reports in the lay press suggest that the kidney care community in Canada has faced numerous disasters and emergencies over the past 2 decades [41-45]. For example, in 2023, many remotely located Indigenous people who received dialysis in the satellite facilities had to be evacuated due to the wildfires. In 2021, torrential rain and winds impacted the care of 72 patients, who needed to be evacuated and repatriated to other dialysis units [45]. While, in the United States, dialysis facilities are mandated by federal regulations to have a disaster plan [16], in Canada, there is no such national initiative in disaster preparedness, response, and recovery in kidney care, although some regional isolated efforts exist. We have highlighted the need to champion disaster and emergency risk reduction and management (DERRM) in kidney care in Canada and underscored the importance of a unified national approach to strengthening system resilience [9]. However, advancing such an approach first requires a clearer understanding of existing systems.

Framework

Overview

To better contextualize this, we first developed a framework in DERRM following a scoping review and a content analysis of contemporary literature addressing disaster management in kidney care across the domains of disaster preparedness, response, and recovery [46]. This work included guidelines, recommendations, commentaries, and review articles. Of the 3381 titles and abstracts screened, 85 (2.5%) articles were selected for full-text review, and ultimately, 45 (1.3%) articles were included in our analysis. Notably, none emerged from Canada. The details of the road map have been published, and brief details are mentioned below.

The ABC4S of Disaster Preparedness

This entails a chronological set of processes as follows: assess needs, risks, and vulnerabilities (regional risks and patients at risk); build a task force network; capacity building (tangible resources, intangible resources, monetary considerations, and transportation); communication (network and protocol, patients' medical and dialysis information, contact information of all stakeholders, inclusive approach, and reliable medium); coaching (patients, caregivers, health care personnel, and reinforce and repeat coaching); contingency planning (surge capacity, rationing care, and resource distribution); and strategic partnerships (Figure 1) [46].

Figure 1. The ABC4S approach to disaster preparedness that we previously developed, which serves as a framework for this vulnerability and adaptation assessment.

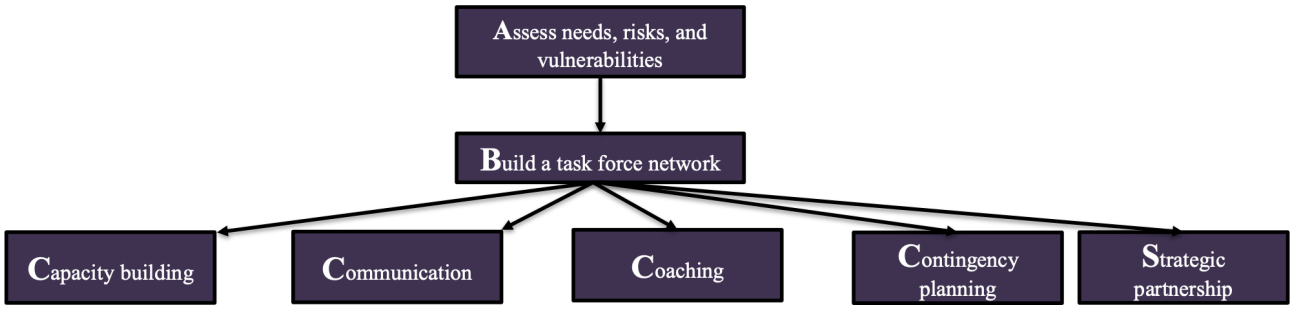

The DIAL Disaster Response

On the basis of the existing evidence, we proposed the following circular series of steps: damage and scope assessment; initiate action plan (choose the plan, apply preparedness tenets, and implications for receiving facilities); appraise the action plan regularly (reassess, maintain ethical standards, and address psychosocial needs); and liaise, engage, and update [46].

The ARC to Recovery

Although we were guided by limited evidence, we proposed the following steps to create a learning health system: assess

XSL•FO
RenderX

damage; return to the (new) norm; and collect data to evaluate, improve, and share [46].

Research Goals

In the Canadian setting, a better understanding of the experiences, resiliency, and vulnerabilities of existing systems is needed to strengthen preparedness efforts and other risk reduction strategies and build "resilient health systems capable of anticipating, responding to, coping with, recovering from, and adapting to climate-related shocks and stress" [7]. Moreover, a better understanding of system weaknesses, practice gaps, and response capacities can help develop adaptation priority actions that can mitigate risks to patients with kidney diseases. The goal of this research is to conduct a vulnerability and adaptation assessment (VAA) of Canadian provincial health systems and individual kidney care networks (KCNs) with respect to natural and human-caused disasters and emergencies. Our research aims are to map existing KCNs across Canada, contextualize the current and future vulnerability of health systems and patients to disasters and emergencies, and identify and prioritize adaptation measures in partnership with key interest holders. Our overarching goal is to strengthen system preparedness and resilience.

Methods

We are conducting a bottom-up VAA that follows a qualitative descriptive approach and uses the DERRM framework that we have previously developed (Figure 1). We are using the COREQ (Consolidated Criteria for Reporting Qualitative Research) to guide the conduct and reporting of this work [47].

Approach

A VAA is a key tool in identifying critical health risks, vulnerable populations, capacities, and weaknesses in the involved systems and the interventions needed to mitigate, manage, and respond to arising risks [7,48,49]. The World Health Organization suggests that VAAs are a foundational step in preparing health systems for hazards. Such an exercise is essential to assess the potential impacts of climate- and human-caused hazards on health systems, profiling the impacted populations, whether patients or health care personnel, and the capacity of health institutions and systems to address arising situations [50]. A VAA explores the components of a health system to assess potential areas of vulnerability to disasters, where capacities stand and/or fail if future risks arise, and what adaptive measures can be taken to achieve a more resilient health system [48]. A bottom-up approach to VAA focuses on the context instead of the outcome of a hazard. It starts with the stakeholders impacted and examines the elements of the involved health systems [51]. Rather than a top-down approach focusing on a particular hazard and estimating its future impact, we used the bottom-up approach to explore the underlying contexts and elements of a system that potentially develop vulnerabilities and limit DERRM [51,52]. Moreover, instead of using model-generated climate data, this approach relies on participatory input from different stakeholder sources to understand disaster management directly from the perspective of those who have experienced or may experience it [52].

Approach

Often used in nursing and health research, a qualitative descriptive approach is important for gaining insights into underexplored research areas [53]. It is the method of choice when straight descriptions of phenomena are desired and for understanding user experiences of processes within health care systems [54]. Moreover, a qualitative descriptive approach recognizes the subjective nature of research questions in health research and explores the answers through the different perspectives of participants' experiences [53]. In practice, this translates into understanding important clinical issues within their unique contexts. This approach also benefits from a flexible inductive process to pursue deeper descriptive knowledge, contributing to change and quality improvement in the practice setting without the need to transform data into a theoretical or conceptual reflection [55,56]. Overall, we are adopting a qualitative descriptive design with the intention of obtaining straightforward descriptions of experiences and perceptions in an area with little prior investigation [53,57].

Data Collection

Overview

An approach to our data collection is presented in Figure 2 and summarized below.

XSL•FO
RenderX

Figure 2. Approach to data collection and analysis in this vulnerability and adaptation assessment research plan of Canadian kidney care networks.

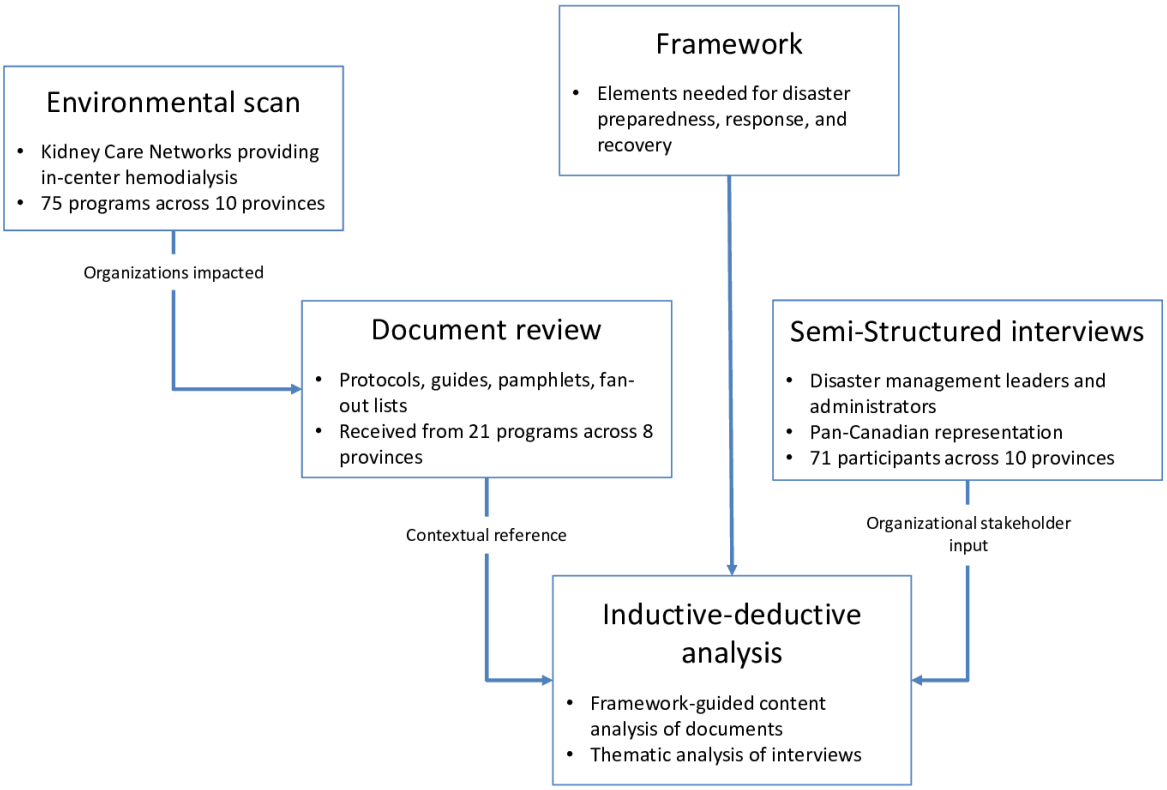

Environmental Scan

In Canada, health care is federally mandated but delivered at the provincial and territorial level across 10 provinces and 3 territories. To establish a clear picture of the kidney care landscape, we began by using data from the Canadian Organ Replacement Register to generate a preliminary map of existing KCNs and dialysis facilities (adult, pediatric, those that are located in cities, and satellite units), home dialysis programs, and transplant programs embedded within each network [58]. Publicly available data from the Canadian Organ Replacement Register provided a foundational national overview; however, as our data collection advanced, additional information from provincial programs, our team, and participants allowed us to refine and update this network structure. The final list of 75 KCNs reflecting these iterative updates and focusing only on adult in-center hemodialysis programs is presented in Table 1. Similar lists will be generated for pediatric, home dialysis, and transplant programs.

XSL•FO
RenderX

Table . A list of Canadian kidney care networks.

Province, provincial network, and health region	Dialysis network and jurisdiction
British Columbia	
BC Renal	
Fraser Health	• Abbotsford Regional Hospital and Cancer Centre • Royal Columbian Hospital • Surrey Memorial Hospital
Interior Health	• Royal Inland Hospital • Kelowna General Hospital • Penticton Regional Hospital • Kootenay Boundary Regional Hospital
Island Health	• Nanaimo Regional General Hospital • Royal Jubilee Hospital
Northern Health	• University Hospital of Northern British Columbia
Vancouver Coastal Health and Providence Health Care	• Vancouver General Hospital • St. Paul's Hospital
Alberta	
Alberta Kidney Care – North	• University of Alberta Hospital
Alberta Kidney Care – South	• University of Calgary
Saskatchewan	
Regina	• Regina General Hospital
Saskatoon	• St. Paul's Hospital
Manitoba	
Prairie Mountain Health	• Brandon Regional Health Centre
Winnipeg Regional Health Authority	• Seven Oaks General Hospital
Shared Health, Northern Health, Interlake-Eastern Regional Health Authority, Prairie Mountain Health, and Southern Health	• Health Sciences Centre
Ontario	
Ontario Renal Network	
Dialysis Management Clinics Inc	—[a]
Erie St. Clair	• Windsor Regional Hospital
South West	• London Health Sciences Centre
Waterloo-Wellington	• Grand River Hospital
Hamilton Niagara Haldimand Brant	• Niagara Health • St. Joseph's Healthcare Hamilton
Central West	• William Osler Health System
Mississauga Halton	• Halton Healthcare • Trillium Health Partners
Toronto Central	• St. Joseph's Health Centre • St. Michael's Hospital • Sunnybrook Health Sciences Centre • University Health Network

XSL•FO
RenderX

Province, provincial network, and health region	Dialysis network and jurisdiction
Central	• Humber River Hospital • Mackenzie Health
Central East	• Lakeridge Health • Peterborough Regional Health Centre • Scarborough Health Network
South East	• Kingston Health Sciences Centre
Champlain	• Renfrew Victoria Hospital • The Ottawa Hospital
North Simcoe Muskoka	• Orillia Soldiers' Memorial Hospital • Royal Victoria Regional Health Centre
North East	• Health Sciences North and Horizon Santé-Nord • North Bay Regional Health Centre • Sault Area Hospital • Timmins and District Hospital
North West	• Thunder Bay Regional Health Sciences Centre
Quebec	

XSL•FO
RenderX

Province, provincial network, and health region	Dialysis network and jurisdiction
CISSS[b] du Bas-Saint-Laurent and CISSS de la Gaspésie	• Hôpital régional de Rimouski
CIUSSS[c] Saguenay–Lac-St-Jean	• Hôpital de Chicoutimi
CHU[d] de Québec–Université Laval, CIUSSS de la Capitale-Nationale, CISSS de la Côte-Nord, and CISSS des Îles	• Pavillon L'Hôtel-Dieu de Québec
CIUSSS Mauricie-et-du-Centre-du-Québec	• CHAUR[e] Trois-Rivières
CIUSSS de l'Estrie – CHUS[f]	• CHU de Sherbrooke – Hôpital Fleurimont
MUHC[g] and Conseil cri de la santé et des services sociaux de la Baie James	• MUHC
CIUSSS de l'Ouest-de-l'Île-de-Montréal	• Centre hospitalier de St. Mary
CIUSSS du Centre-Ouest-de-l'Île-de-Montréal	• Hôpital général juif Sir Mortimer B. Davis
CIUSSS de l'Est-de-l'Île-de-Montréal	• Hôpital Maisonneuve-Rosemont
CIUSSS du Nord-de-l'Île-de-Montréal Clinique Bois-de-Boulogne	• Hôpital du Sacré-Cœur de Montréal
CIUSSS du Centre-Sud-de-l'Île-de-Montréal	• Hôpital de Verdun
CHUM[h]	• CHUM
CISSS de l'Outaouais	• Hôpital de Hull
CISSS de Chaudière-Appalaches	• Hôtel-Dieu de Lévis
CISSS de Laval	• Hôpital de la Cité-de-la-Santé
CISSS de Lanaudière	• Centre hospitalier régional de Lanaudière – Hôpital de Joliette
CISSS des Laurentides and CISSS de l'Abitibi-Témiscamingue	• Hôpital régional de Saint-Jérôme
CISSS de la Montérégie-Centre	• Hôpital Charles-LeMoyne
CISSS de la Montérégie-Est	• Hôpital Honoré-Mercier
CISSS de la Montérégie-Ouest	• Hôpital Anna-Laberge
Newfoundland	
Newfoundland and Labrador Health Services	
Eastern, Western, Central, and Labrador-Grenfell Health	• Health Sciences Center Hemodialysis Unit • Corner Brook Hemodialysis Unit • Grand Falls-Windsor Hemodialysis Unit
New Brunswick	
Horizon Health Network	• Saint John Regional Hospital
Vitalité Health Network	• Dr. Georges-L.-Dumont University Hospital Centre
Prince Edward Island	
PEI Renal Program	• Queen Elizabeth Hospital
Nova Scotia	

XSL•FO
RenderX

Province, provincial network, and health region	Dialysis network and jurisdiction
Eastern, Western, Northern, and Central Health	• Queen Elizabeth II Health Sciences Centre • Cape Breton Regional Hospital

[a]Dialysis Management Clinics Inc includes multiple out-of-hospital dialysis units across the province of Ontario and does not fall under a particular dialysis network or jurisdiction.

[b]CISSS: Les centres intégrés de santé et de services sociaux.

[c]CIUSSS: Les centres intégrés universitaires de santé et de services sociaux.

[d]CHU: Centre Hospitalier Universitaire.

[e]CHAUR: Centre hospitalier affilié universitaire régional.

[f]CHUS: Centre hospitalier universitaire de Sherbrooke.

[g]MUHC: McGill University Health Centre.

[h]CHUM: Centre hospitalier de l'Université de Montreal.

Document Review

Once a preliminary list of dialysis networks was generated, we solicited documents to better understand existing practices and policies related to disaster and emergency management. This allows for an examination of different standing policies, directives, and resources on disaster management in each program and will serve as a complementary resource during data analysis [59]. Documents for review were obtained prior to, after, or in conjunction with interviews and included any policy, guideline, directive, resource, or presentation pertaining to disaster management within each dialysis network. Documents nonspecific to these networks, such as hospital-wide emergency protocols, were ineligible. Of the 55 KCNs that have responded thus far, we have procured documents from 21 (38.2%) networks, with 39 (70.9%) confirming they have no formal kidney-specific protocols in place. A list of these documents is provided in Table 2.

XSL•FO
RenderX

Table . List of documents solicited that will be analyzed during the vulnerability and adaptation assessment.

Province and document title	Health region or dialysis network
British Columbia	
Emergency Management Plan	BC Renal
Renal Department Disaster Response Initiation Chain of Command and Events	BC Renal
Are You Ready: Emergency Preparedness Information for Dialysis Patients	BC Renal
Emergency Preparedness Wallet Card	BC Renal
Emergency Preparedness Month: Planning for disaster - Patient flyer	BC Renal
Emergency Kit Essentials	BC Renal
Alberta	
Pandemic Response Plan – Alberta Kidney Care	Alberta Kidney Care – North
Emergency Preparedness and Response – Hemodialysis	Alberta Kidney Care – South
Emergency Meal Plan for People on Hemodialysis	Alberta Kidney Care – South
Bad Weather Patient Information Package[a]	Alberta Kidney Care – South
Saskatchewan	
Ready or not: Emergency Preparedness Information Booklet for Saskatchewan Dialysis Patients	Regina Kidney Health
Manitoba	
Manitoba Renal Program - Disaster Management Manual	Winnipeg Regional
Ontario	
Emergency Preparedness Plan	Dialysis management clinics
Planning Resource Checklist	Dialysis management clinics
Clinical Services Listing	Dialysis management clinics
Priority Services Checklist	Dialysis management clinics
CKD Surge Demand and Capabilities Checklist	Dialysis management clinics
Critical Infrastructure Checklist	Dialysis management clinics
Emergency Renal Product Supply Checklist	Dialysis management clinics
Dialysis Capital Equipment Inventory	Dialysis management clinics
CKD Hazard Identification and Risk Assessment Workbook	Dialysis management clinics
CKD Hazard Identification and Risk Assessment Instructions for Use	Dialysis management clinics
Emergency Contact Information List	Dialysis management clinics
Staff Planning: Roles and Responsibility	Dialysis management clinics
Staff Planning: Operational and Emergency Full-Time Equivalent Requirements	Dialysis management clinics
Annual Training and Education Checklist	Dialysis management clinics
The Final Checklist	Dialysis management clinics
Emergency Management Planning Guide 2015	Ontario Renal Network
Nephrology Emergency Response Plan	Health Sciences North
Renal Program Leadership Emergency Management Guide	London Health Sciences Centre
Emergency Preparedness Information for Peterborough Regional Renal Program Hemodialysis Patients	Peterborough Regional Health Centre
Emergency Preparedness Information for Peterborough Regional Renal Program Multi-Care Kidney Clinic Patients	Peterborough Regional Health Centre

XSL•FO
RenderX

Province and document title	Health region or dialysis network
Emergency Preparedness Information for Peterborough Regional Renal Program Peritoneal Dialysis Patients	Peterborough Regional Health Centre
Peterborough Regional Renal Program Hazardous Identification and Risk Assessment - Emergency Management Plan 2020	Peterborough Regional Health Centre
Regional Renal Program - Emergency Management Plan	Peterborough Regional Health Centre
Algoma Regional Renal Program - Emergency Management Plan	Sault Area Hospital
Dialysis Patients: Being Prepared for an Emergency	Sault Area Hospital
Kidney Care Contingency Plan	St. Joseph's Health Centre
Emergency Management Plan - Regional Kidney Care Program Simcoe Muskoka	Orillia Soldiers' Memorial Hospital
Quebec	
Plan d'urgence - Plan particulier d'intervention (PPI)	Centre hospitalier universitaire de Québec
Procédure interne de communication entre les coordonnateurs d'activités de soins et les AIC du Centre de services ambulatoires de dialyse de Gaspé et des sites résiduels de Notre-Dame et Saint-Luc.	Centre hospitalier de l'Université de Montréal
Plan des mesures d'urgence de la dialyse - Eau Contaminée	Centre hospitalier de l'Université de Montréal
Plan des mesures d'urgence de la dialyse - Panne de Courant	Centre hospitalier de l'Université de Montréal
Plan des mesures d'urgence de la dialyse - Panne d'alimentation en eau ou du traitement d'eau	Centre hospitalier de l'Université de Montréal
Plan des mesures d'urgence de la dialyse - Évacuation	Centre hospitalier de l'Université de Montréal
Mesure d'urgence pour patients - Gaspé	Centre hospitalier de l'Université de Montréal
Mesure d'urgence Mesure d'urgence – Fournitures essentielles d'accès vasculaire[a]	Centre hospitalier de l'Université de Montréal
Plan d'urgence - Suppléance Rénale	CIUSSS[b] de l'Est-de-l'Île-de-Montréal
Liste des machines et du nombre de patients dialysés	CIUSSS du Nord-de-l'Île-de-Montréal
Dialysis Patients: Being Prepared for an Emergency	MUHC[c]
Procedure for Evacuation of Dialysis Patients During Hemodialysis	MUHC
Plan De Contingence Hémodialyse	CIUSSS de l'Estrie – CHUS[d] Sherbrooke
Plan Contingence Plan Répartition Quotas 4 Patients/Inf.	CIUSSS de l'Estrie – CHUS Sherbrooke
Prince Edward Island	
Renal Program Storm Planning Quick Reference Checklist	PEI Renal Network
Provincial Renal Program Storm Plan	PEI Renal Network
Storm Planning Notification	PEI Renal Network
Queen Elizabeth Hospital Hemodialysis Storm Planning	PEI Renal Network
Nova Scotia	
Emergency Backup Hemodialysis Infrastructure Table	Nova Scotia Health Authority
Emergency Patient Status/Transfer Requirements Summary Form	Nova Scotia Health Authority
Renal Program Business Continuity Status Checklist	Nova Scotia Health Authority
NSHA Renal Program Disaster Evacuation Response	Nova Scotia Health Authority
NSHA Renal Program Emergency Management Action Plan - allocation table	Nova Scotia Health Authority
NSHA Renal Program Essential Services	Nova Scotia Health Authority
NSHA Renal Program Extended Power Outage	Nova Scotia Health Authority
NSHA Renal Program Severe Weather Procedure	Nova Scotia Health Authority
NSHA Renal Program Short Term Power Outage Procedure	Nova Scotia Health Authority

XSL•FO
RenderX

Province and document title	Health region or dialysis network
NSHA Renal Program Short Term Water Disruption Procedure	Nova Scotia Health Authority

[a]Title created based on document content.

[b]CIUSSS: Les centres intégrés universitaires de santé et de services sociaux.

[c]MUHC: McGill University Health Centre.

[d]CHUS: Centre hospitalier universitaire de Sherbrooke.

Semistructured Interviews

Finally, we are conducting semistructured interviews with health care professionals (clinicians, administrators, technicians, nurses, and other allied health professionals) exploring the vulnerabilities and service gaps of existing systems, identifying strengths, and discussing potential adaptations to increase the resilience of kidney care in the face of disasters. The guide explores participants' experiences with disasters and disaster management within their networks or their specific dialysis unit. We investigate participants' roles, and depending on the extent of their experience with a disaster, participants are prompted to describe lived experiences and their perspectives on the state of disaster preparedness and response at their units or networks. Questions explore different components of the health system and disaster vulnerability, current capacities, associated challenges, and potential adaptations. For example, topics include the comprehensiveness and maintenance of current protocols at the unit; leadership and coordination; aspects of patient preparedness and vulnerability; personnel training and capacities; and different infrastructure systems such as physical, financial, and communication.

The second part of the interview solicits feedback on the developed DERRM framework as a potential adaptation at their unit. All 3 components of the framework are shown to participants, and feedback is collected on its different elements, comprehensiveness, feasibility, and applicability at their units. A complete interview guide is provided in Multimedia Appendix 1. Interviews follow an iterative approach whereby issues or ideas identified by participants are discussed with subsequent participants to enable further definition and refinement of themes [60]. Interviews are conducted remotely through Zoom (Zoom Video Communications) in English or French by trained qualitative researchers. All interviews are being digitally recorded and transcribed.

Eligibility and Recruitment

Our eligibility criteria are participants who have a key leadership, administrative, or operational role in a KCN. Relying on the comprehensive list of 75 identified KCNs following the environmental scan, recruitment will focus on the main units within each network as managerial and administrative functions are set by them and followed accordingly by other units within the network. We use purposive criterion sampling to identify these individuals through Google searches, phone calls to kidney units, and the assistance of our collaborators. We focus on identifying a minimum of 1 participant per KCN. Considering varying geographical contexts and different levels of disaster exposure, and to allow for more nuanced experiences, some participants in nonmanagerial roles or from satellite KCNs are also being invited using a snowballing method [61,62]. We will pursue recruitment at each of the 75 KCNs and plan to invite at least 2 contacts within each network a maximum of 3 times. This will ensure accuracy, comprehensiveness, and integrity as all KCNs will be offered an opportunity to participate [63]. If a main unit is nonresponsive, we will move on to other units within that network prior to labeling it as a nonresponding KCN. Recruitment will continue beyond the point of data saturation to pursue pan-Canadian representation.

Data Analysis

Overview

We will use an inductive-deductive approach whereby the DERRM framework guides our data interpretation while also allowing space for emerging themes. Although inductive analysis is more common in qualitative research [64], a deductive framework-guided approach also provides a conceptual map to contextually incorporate data and interpret them in connection with existing literature and evidence [64-67]. This combined analysis approach creates space for different nuances—such as interprovincial differences, rural vs urban settings, and different stakeholder inputs—while maintaining conceptual integrity. As a result, a comprehensive assessment of Canadian KCN vulnerabilities and the needed adaptations can integrate current risk considerations into existing decision-making and health system components [51]. We will use the following stepwise approach.

Document Review

We will conduct a deductive framework-guided content analysis. Using the DERRM framework categories, we will develop a coding scheme matrix of the different elements and subelements of disaster management. Document passages will be organized into predefined matrix categories and coded accordingly, with attention paid to data that fit more than one category. The documents' contents will be coded line by line with the DERRM framework elements as a guide and cross-referenced to populate the matrix with the relevant data. Once the matrix is filled, we will synthesize which categories were more strongly, weakly, or not represented across the documents. This systematic approach will help understand the existing components of disaster management at Canadian KCNs, keeping in mind that the presence of fragmented directives or policies or the absence of any directives is a similarly significant finding. We will also note substantial contradictions and remain open to any categories and subcategories that might emerge [68]. Supporting data, such as rural vs urban location of the unit, size of the unit, size of the network, and province, will also be considered during the analysis.

XSL•FO
RenderX

Semistructured Interviews

The interviews will be analyzed using inductive and deductive thematic analysis. Thematic analysis entails identifying and analyzing patterns of meaning and mapping regularities and variations across different accounts [69-72]. An inductive and deductive approach emerges by using the existing DERRM framework to fit our data into an a priori concept [73,74] while also allowing for any new themes and codes to emerge through a close reading of the data [72,75,76]. Interview transcripts will be analyzed independently by 2 senior qualitative researchers. NVivo (version 15; Lumivero) will be used to support data management and analysis.

Transcripts will be read and highlighted line by line to preliminarily derive data anchored in the DERRM framework elements and tease out emerging ideas and themes. These initial findings will be organized into codes and subcodes separately by each researcher to form 2 parallel initial coding schemes. The team will meet to discuss and establish whether intercoder reliability is achieved and decide on a first draft of the codebook. We will ensure rigor throughout the research process, drawing on concepts of transparency, credibility, transferability, and reflexivity as indicators of research quality in qualitative research [77,78]. Transparency will be ensured through the maintenance of a research journal documenting all decisions related to sampling and analysis. Reflexivity will be supported through ongoing discussions of emergent codes and themes alongside critical reflection on analytic decisions and the researchers' potential influence on the process. The final themes will be presented to the collaborators prior to manuscript preparation.

Ethical Considerations

Ethics approval for this study was obtained from the Research Ethics Board at McGill University Health Centre, Montreal, Quebec (2025-11213). This study is being conducted in accordance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2 (2022) [79].

Written or electronic informed consent will be obtained from interview participants via email prior to commencing the interviews. It will briefly explain the study and its purpose, the voluntary nature of participation, the duration of the interview, and that any personally identifiable information will not be collected or will be omitted from the transcripts. Participants will be compensated with a CAD $50 (US $38) gift card following their participation in the interviews.

Results

Funding for this research was obtained in April 2025. The environmental scan was initiated in November 2024, and the semistructured interviews were initiated in February 2025. Interviews are ongoing, and the analysis has been initiated simultaneously. We have recruited 71 participants from 55 KCNs from March 2025 to December 2025 and are attempting to recruit at least one participant from the remaining 20 KCNs. The first VAA focusing on in-center hemodialysis services is expected to conclude in April 2026. These findings are projected to be written up by May 2026, and manuscripts are expected to be submitted for publication to peer-reviewed journals in the summer of 2026 (tentatively June 2026). Thereafter, similar VAAs will be conducted focusing on other kidney replacement therapies and pediatric programs. Separate analysis and reporting will be considered depending on emerging codes and resulting themes. Key contacts at KCNs across the different provinces who have supported participant identification will be involved in data interpretation and manuscript preparation. Key insights are emerging with respect to differences in the provision of kidney care; a separate analysis reporting these findings will be considered with the input of our collaborators.

Discussion

Expected Findings

This study aims to operationalize a DERRM framework to conduct a VAA of KCNs in Canada. This pan-Canadian initiative will identify and address the vulnerability of the kidney health community (ie, health care providers, patients with kidney disease, kidney care units, and provincial renal health authorities), identify approaches to adaptation, facilitate DERRM, and strengthen health system resilience to disasters. We want to identify the existing operational and human vulnerabilities and the associated risks, as well as explore lessons learned from previous experiences, to develop needed adaptations. Emerging data suggest that this varies widely between programs and provinces. As mentioned above, only 38.2% (21/55) of the responding programs have written documents specifying policies on disaster management. Input gathered has covered managerial and operational aspects, ensuring relevance to applicable policy and lived practices and experiences [7]. The snowball recruitment approach enabled us to reach individuals who are often difficult to engage, including those supporting rural communities and working in direct, frontline roles with patients [46]. These participants contributed valuable experiences in managing disasters through working with different dialysis modalities, holding tenure in their institutions, working with remote communities, and being exposed to a multitude of hazards. All participants interviewed thus far have offered differing perspectives on disaster management in different Canadian contexts. Document review will serve as complementary data collection and as a means of triangulation with interview data [59]. By examining both existing directives and lived experiences and the expertise of health care professionals (clinicians, administrators, technicians, nurses, and other allied health professionals), this approach provides a comprehensive assessment of DERRM at the individual, provincial, and health system levels while directly advancing our research aims.

This will be the first VAA of kidney care facilities across Canada. In addition to investigating vulnerabilities and adaptations at the KCN level, analysis will highlight provincial and regional variations based on prior exposure to disasters and perceived current and future risks. Future research will explore patient and caregiver perspectives and adaptation recommendations to generate actionable goals to guide relevant interest holders in prioritizing and planning responses to current and future health threats. Furthermore, as per the World Health

XSL•FO
RenderX

Organization, VAA can serve as a baseline for monitoring changes in risk and evaluating the effectiveness of adaptation measures over time.

Limitations

The following limitations may be present in our analysis, and we are undertaking steps to address some of them. First, despite extensive consultations with collaborators and ongoing verification during interviews, we may not have fully captured all existing KCNs in Canada. Although we attempted to engage every KCN to achieve pan-Canadian representation, several regions could not be reached to validate their inclusion, and we are continuing efforts to engage with them. During our ongoing data collection, significant differences are emerging regarding the vulnerabilities of patients on various kidney replacement therapies. We are attempting to recruit participants who will represent the full scale of therapies across adult and pediatric populations.

Regarding our methods, our recruitment approach may be susceptible to nonresponse bias as only those with a personal interest in the research topic may have agreed to participate. Additionally, those who perceived their KCN as insufficiently prepared may have opted not to participate, potentially limiting the range of perspectives captured. This may potentially skew insights into system vulnerabilities and the applicability of potential adaptations. However, we are undertaking several measures to increase our outreach efforts to ensure that all perspectives are represented. Finally, although the document review serves as a contextual reference for existing policies on disaster management at KCNs complementary to interview data, we acknowledge that it may not be exhaustive as documents are shared through participants and dependent on their knowledge of their own unit's directives or lack thereof.

The target population focused on the main dialysis center within each KCN, and we only recruited participants from satellite units through snowballing when recommended. We acknowledge that other units within a network can also offer valuable, nuanced input; however, considering that administrative functions and managerial functions are localized to main units, we believe that this level of representation is sufficient. In parallel, the study examines organizational-level vulnerabilities and adaptations without addressing individual-level risk factors, including patient experience with different kidney replacement therapy modalities. Patient perspectives will be pursued in the following stages of the study.

Conclusions

Patients with kidney diseases are uniquely vulnerable to the impacts of a hazard. Delay in the provision of dialysis or medications can be a life-threatening or organ-threatening event for these patients. This VAA, guided by a framework we previously developed within the DERRM model and informed by an extensive literature review and experiential knowledge, represents the first initiative of its kind in kidney care. The findings will provide a critical foundation for understanding vulnerability, supporting strategic planning, and guiding adaptation measures that can strengthen health system resilience. Ultimately, this work will help protect patients with kidney diseases and the broader kidney community from the escalating risks of evolving hazards within Canada and may serve as a model for similar efforts elsewhere.

Acknowledgments

The authors would like to thank the Sustainable Nephrology Action Planning committee of the Canadian Society of Nephrology, BC Renal, and the Can-SOLVE CKD Network (Canada's largest-ever kidney research initiative) for their support of this study and assistance with participant recruitment. SS is supported by the Chercheuses-boursières cliniciennes – Junior 1 from the Fonds de recherche du Québec – Santé (an education grant to allow for protected research time for clinician scientists).

Funding

This project is funded by Canadian Institutes of Health Research Project Grant - Priority Announcement: Pandemic Preparedness and Health Emergencies Research (multiyear funding; grant PNM – 198199).

Data Availability

The datasets generated and analyzed during this study are not publicly available because informed consent to share transcribed data for secondary use beyond this research was not obtained from participants.

Authors' Contributions

SS conceived, designed, and conceptualized the study. SM and SEW are conducting data collection. SEW wrote the first draft of the paper. SS and SM contributed to the development of the paper and provided comments for improvement. All authors have read and approved the final manuscript.

Conflicts of Interest

SS has received a one-time speaking honorarium from AstraZeneca. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Semistructured interview guide for health care professionals.

XSL•FO
RenderX

[DOCX File, 23 KB - resprot_v15i1e90059_app1.docx]

Peer Review Report 1
Peer review report by the Clinical Investigation C: Digestive, Endocrine and Excretory Systems (CIC) Review Committee, Canadian Institutes of Health Research (CIHR).
[PDF File, 199 KB - resprot_v15i1e90059_app2.pdf]

References

1. Climate change and health: vulnerability and adaptation assessment. World Health Organization. 2021. URL: https://www.who.int/publications/i/item/9789240036383 [accessed 2026-04-06]
2. Critical infrastructure including schools, health facilities and roads threatened by climate crisis. United Nations Office for Disaster Risk Reduction (UNDRR). 2019. URL: https://www.undrr.org/news/critical-infrastructure-including-schools-health-facilities-and-roads-threatened-climate [accessed 2026-04-06]
3. WHO guidance on research methods for health emergency and disaster risk management, revised 2022. World Health Organization. 2022. URL: https://iris.who.int/handle/10665/363502 [accessed 2026-04-06]
4. Tselios V, Tompkins EL. What causes nations to recover from disasters? An inquiry into the role of wealth, income inequality, and social welfare provisioning. Int J Disaster Risk Reduct 2019 Feb;33:162-180. [doi: 10.1016/j.ijdrr.2018.10.003]
5. Lindersson S, Raffetti E, Rusca M, Brandimarte L, Mård J, Di Baldassarre G. The wider the gap between rich and poor the higher the flood mortality. Nat Sustain 2023;6:995-1005. [doi: 10.1038/s41893-023-01107-7]
6. Choong JJ, Wagenaar D, Rabonza ML, Hamel P, Switzer AD, Lallemant D. Shared hazards, unequal outcomes: income-driven inequities in disaster risk. npj Nat Hazards 2025;2:33. [doi: 10.1038/s44304-025-00087-2]
7. Operational framework for building climate resilient and low carbon health systems. World Health Organization. 2023. URL: https://www.who.int/publications/i/item/9789240081888 [accessed 2026-04-06]
8. Disaster risk reduction terminology. United Nations Office for Disaster Risk Reduction. URL: https://www.undrr.org/drr-glossary/terminology [accessed 2026-04-06]
9. Sandal S, El Wazze S, Stigant CE, et al. From risk and vulnerability to preparedness and resiliency: an opinion piece on championing disaster and emergency risk reduction and management in kidney care. Can J Kidney Health Dis 2025;12:20543581251391277. [doi: 10.1177/20543581251391277] [Medline: 41281374]
10. Pawłowicz-Szlarska E, Vanholder R, Sever MS, et al. Distribution, preparedness and management of Ukrainian adult refugees on dialysis-an international survey by the Renal Disaster Relief Task Force of the European Renal Association. Nephrol Dial Transplant 2023 Sep 29;38(10):2407-2415. [doi: 10.1093/ndt/gfad073] [Medline: 37326036]
11. Sever MS, Vanholder R, Luyckx V, et al. Armed conflicts and kidney patients: a consensus statement from the Renal Disaster Relief Task Force of the ERA. Nephrol Dial Transplant 2023 Jan 23;38(1):56-65. [doi: 10.1093/ndt/gfac247] [Medline: 35998320]
12. Gopolan T, Ornelas-Brauer CM, Barbar T, Mithani Z, Silberzweig J. Conflict nephrology: war and natural disasters. Kidney360 2023 Mar 1;4(3):405-408. [doi: 10.34067/KID.0000000000000071] [Medline: 36763799]
13. Pape L, Ahlenstiel-Grunow T, Birtel J, Krohne TU, Hoppe B. Oxalobacter formigenes treatment combined with intensive dialysis lowers plasma oxalate and halts disease progression in a patient with severe infantile oxalosis. Pediatr Nephrol 2020 Jun;35(6):1121-1124. [doi: 10.1007/s00467-019-04463-9] [Medline: 32107618]
14. Vanholder R, Gallego D, Sever MS. Wars and kidney patients: a statement by the European Kidney Health Alliance related to the Russian-Ukrainian conflict. J Nephrol 2022 Mar;35(2):377-380. [doi: 10.1007/s40620-022-01301-4] [Medline: 35246798]
15. Vanholder R, Sever MS, Lameire N. The role of the Renal Disaster Relief Task Force. Nat Clin Pract Nephrol 2008 Jul;4(7):347. [doi: 10.1038/ncpneph0862] [Medline: 18580958]
16. Kopp JB, Lempert K, Finne K. Disaster preparedness for patients with kidney disease. Nat Rev Nephrol 2023 Mar;19:147-148. [doi: 10.1038/s41581-023-00678-0] [Medline: 36747083]
17. Sapkota A, Kotanko P. Climate change-fuelled natural disasters and chronic kidney disease: a call for action. Nat Rev Nephrol 2023 Mar;19(3):141-142. [doi: 10.1038/s41581-023-00682-4] [Medline: 36670273]
18. Sever MS, Luyckx V, Tonelli M, et al. Disasters and kidney care: pitfalls and solutions. Nat Rev Nephrol 2023 Oct;19(10):672-686. [doi: 10.1038/s41581-023-00743-8] [Medline: 37479903]
19. Bonomini M, Stuard S, Dal Canton A. Dialysis practice and patient outcome in the aftermath of the earthquake at L'Aquila, Italy, April 2009. Nephrol Dial Transplant 2011 Aug;26(8):2595-2603. [doi: 10.1093/ndt/gfq783] [Medline: 21248293]
20. Sever MS, Erek E, Vanholder R, et al. Features of chronic hemodialysis practice after the Marmara earthquake. J Am Soc Nephrol 2004 Apr;15(4):1071-1076. [doi: 10.1097/01.asn.0000119145.40232.67] [Medline: 15034111]
21. Abir M, Jan S, Jubelt L, Merchant RM, Lurie N. The impact of a large-scale power outage on hemodialysis center operations. Prehosp Disaster Med 2013 Dec;28(6):543-546. [doi: 10.1017/S1049023X13008844] [Medline: 23965738]

XSL•FO
RenderX

22. Lin CJ, Pierce LC, Roblin PM, Arquilla B. Impact of Hurricane Sandy on hospital emergency and dialysis services: a retrospective survey. Prehosp Disaster Med 2014 Aug;29(4):374-379. [doi: 10.1017/S1049023X14000715] [Medline: 25068276]
23. Nishimura H, Kagara I, Inokuchi S, et al. Local dialysis disaster relief during two torrential downpours on Amami-Ohshima Island. J Disaster Res 2014;9(1):86-91. [doi: 10.20965/jdr.2014.p0086]
24. Sekkarie MA, Zanabli AR, Rifai AO, Murad LB, Al-Makki AA. The Syrian conflict: assessment of the ESRD system and response to hemodialysis needs during a humanitarian and medical crisis. Kidney Int 2015 Feb;87(2):262-265. [doi: 10.1038/ki.2014.336] [Medline: 25635715]
25. Alasfar S, Alashawi HS, Murad L, Burnham GM. Assessment of dialysis operations in conflict-affected areas: insights from survey results from Northwest Syria. J Am Soc Nephrol 2023;34(11S):736-737. [doi: 10.1681/ASN.20233411S1736b]
26. Murakami N, Siktel HB, Lucido D, Winchester JF, Harbord NB. Disaster preparedness and awareness of patients on hemodialysis after Hurricane Sandy. Clin J Am Soc Nephrol 2015 Aug 7;10(8):1389-1396. [doi: 10.2215/CJN.10181014] [Medline: 26220814]
27. Anderson AH, Cohen AJ, Kutner NG, Kopp JB, Kimmel PL, Muntner P. Missed dialysis sessions and hospitalization in hemodialysis patients after Hurricane Katrina. Kidney Int 2009 Jun;75(11):1202-1208. [doi: 10.1038/ki.2009.5] [Medline: 19212421]
28. Hyre AD, Cohen AJ, Kutner N, Alper AB, Muntner P. Prevalence and predictors of posttraumatic stress disorder among hemodialysis patients following Hurricane Katrina. Am J Kidney Dis 2007 Oct;50(4):585-593. [doi: 10.1053/j.ajkd.2007.07.013] [Medline: 17900458]
29. Damasiewicz MJ, Magliano DJ, Daly RM, et al. Serum 25-hydroxyvitamin D deficiency and the 5-year incidence of CKD. Am J Kidney Dis 2013 Jul;62(1):58-66. [doi: 10.1053/j.ajkd.2013.03.010] [Medline: 23623574]
30. Kelman J, Finne K, Bogdanov A, et al. Dialysis care and death following Hurricane Sandy. Am J Kidney Dis 2015 Jan;65(1):109-115. [doi: 10.1053/j.ajkd.2014.07.005] [Medline: 25156306]
31. Blum MF, Feng Y, Anderson GB, Segev DL, McAdams-DeMarco M, Grams ME. Hurricanes and mortality among patients receiving dialysis. J Am Soc Nephrol 2022 Sep;33(9):1757-1766. [doi: 10.1681/ASN.2021111520] [Medline: 35835459]
32. Thomas DA, Huang A, McCarron MC, et al. A retrospective study of chronic kidney disease burden in Saskatchewan's First Nations people. Can J Kidney Health Dis 2018;5:2054358118799689. [doi: 10.1177/2054358118799689] [Medline: 30245841]
33. United States Renal Data System. USRDS Annual Data Report: Epidemiology of Kidney Disease in the United States. : National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases; 2023 URL: https://usrds-adr.niddk.nih.gov/2023 [accessed 2026-04-10]
34. Yeates K. Health disparities in renal disease in Canada. Semin Nephrol 2010 Jan;30(1):12-18. [doi: 10.1016/j.semnephrol.2009.10.014] [Medline: 20116643]
35. Garcia-Garcia G, Jha V. Chronic kidney disease in disadvantaged populations. Indian J Nephrol 2015;25(2):65-69. [doi: 10.4103/0971-4065.150078] [Medline: 25838641]
36. Nash DM, Dirk JS, McArthur E, et al. Kidney disease and care among First Nations people with diabetes in Ontario: a population-based cohort study. CMAJ Open 2019;7(4):E706-E712. [doi: 10.9778/cmajo.20190164] [Medline: 31822501]
37. Harasemiw O, Komenda P, Tangri N. Addressing inequities in kidney care for indigenous people in Canada. J Am Soc Nephrol 2022 Aug;33(8):1474-1476. [doi: 10.1681/ASN.2022020215] [Medline: 35667707]
38. Canadian Institute for Health Information. Organ replacement in Canada: CORR annual statistics. 2023. URL: https://web.archive.org/web/https://www.cihi.ca/en/organ-replacement-in-canada-corr-annual-statistics [accessed 2026-04-10]
39. Vart P, Gansevoort RT, Joosten MM, Bültmann U, Reijneveld SA. Socioeconomic disparities in chronic kidney disease: a systematic review and meta-analysis. Am J Prev Med 2015 May;48(5):580-592. [doi: 10.1016/j.amepre.2014.11.004] [Medline: 25891058]
40. Natural hazards. Government of Canada. URL: https://natural-resources.canada.ca/science-data/science-research/natural-hazards [accessed 2026-04-06]
41. Public Safety Canada. National Risk Profile: Strengthening Canada's All-Hazards Approach to Emergency Management. 2024. URL: https://www.publicsafety.gc.ca/cnt/mrgnc-mngmnt/ntnl-rsk-prfl/index-en.aspx [accessed 2026-04-10]
42. Kane P. Yellowknife's wildfire evacuation was tailored for the privileged. The Walrus. 2024 May 16. URL: https://thewalrus.ca/yellowknifes-wildfire-evacuation-was-tailored-for-the-privileged [accessed 2026-04-06]
43. Bongiorno J. As tourists flock to Montreal for F1 race, Cree dialysis patients who fled fires are on the move again. CBC News. 2023. URL: https://www.cbc.ca/news/canada/montreal/dialysis-montreal-cree-wildfires-hotel-1.6879932 [accessed 2026-04-06]
44. Not one run missed: how Williams Lake maintained full care levels during largest fire in BC history. BC Renal. 2018. URL: http://www.bcrenal.ca/about/news-stories/news/not-one-run-missed-how-williams-lake-maintained-full-care-levels-during-largest-fire-in-bc-history [accessed 2026-04-06]
45. BC hemodialysis patient and nurse who adapted to extreme flood conditions share their experience. BC Renal. 2022. URL: http://www.bcrenal.ca/about/news-stories/people-in-kidney-care/bc-hemodialysis-patient-and-nurse [accessed 2026-04-06]

XSL•FO
RenderX

46. Sandal S, El Wazze S, Nijjar D, et al. A roadmap for disaster risk reduction and management in kidney care: a scoping review and content analysis. J Am Soc Nephrol 2025 Feb 6;36(7):1363-1379. [doi: 10.1681/ASN.0000000635] [Medline: 39913200]
47. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
48. Climate change. In: Compendium of WHO and Other UN Guidance in Health and Environment, 2024 Update: World Health Organization; 2024:112 URL: https://scopeclimatechange.scopeonline.in/wp-content/uploads/2025/01/PDF-5.pdf [accessed 2026-04-06]
49. Climate change and health vulnerability and adaptation assessments: a knowledge to action resource guide. : Government of Canada; 2021 URL: https://www.canada.ca/en/health-canada/services/publications/healthy-living/climate-health-adapt-vulnerability-adaptation-assessments-resource-guide.html [accessed 2026-04-06]
50. Methods of assessing human health vulnerability and public health adaptation to climate change. : World Health Organization; 2003 URL: https://www.who.int/publications/i/item/methods-of-assessing-human-health-vulnerability-and-public-health-adaptation-to-climate-change [accessed 2026-04-06]
51. A framework for climate change vulnerability assessments. : Deutsche Gesellschaft für Internationale Zusammenarbeit; 2014 URL: https://www.adaptationcommunity.net/download/va/vulnerability-guides-manuals-reports/Framework_for_Climate_Change_Vulnerability_Assessments_-_GIZ_2014.pdf [accessed 2026-04-06]
52. Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ). A Framework for Climate Change Vulnerability Assessments. Ministry of Environment, Forests and Climate Change, Government of India. 2014 URL: https://weadapt.org/wp-content/uploads/2023/08/5476022698f9agiz2014-1733en-framework-climate-change.pdf [accessed 2026-04-11]
53. Doyle L, McCabe C, Keogh B, Brady A, McCann M. An overview of the qualitative descriptive design within nursing research. J Res Nurs 2020 Aug;25(5):443-455. [doi: 10.1177/1744987119880234] [Medline: 34394658]
54. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
55. Bradshaw C, Atkinson S, Doody O. Employing a qualitative description approach in health care research. Glob Qual Nurs Res 2017;4:2333393617742282. [doi: 10.1177/2333393617742282] [Medline: 29204457]
56. Chafe R. The value of qualitative description in health services and policy research. Healthc Policy 2017 Feb;12(3):12-18. [Medline: 28277201]
57. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
58. Canadian Institute for Health Information. Canadian Organ Replacement Register Directory. 2024. URL: https://web.archive.org/web/https://www.cihi.ca/en/canadian-organ-replacement-register-directory [accessed 2026-04-10]
59. Bowen GA. Document analysis as a qualitative research method. Qual Res J 2009 Aug 3;9(2):27-40. [doi: 10.3316/QRJ0902027]
60. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health 2019;7(2):e000057. [doi: 10.1136/fmch-2018-000057] [Medline: 32148704]
61. Reinhold S. Local Conflict and Ideological Struggle: "Positive Images" and Section 28: University of Sussex; 1994. URL: https://books.google.co.in/books/about/Local_Conflict_and_Ideological_Struggle.html?id=KTvFoAEACAAJ&redir_esc=y [accessed 2026-04-06]
62. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches: SAGE Publications; 2016. URL: https://books.google.co.in/books/about/Qualitative_Inquiry_and_Research_Design.html?id=DLbBDQAAQBAJ&redir_esc=y [accessed 2026-04-06]
63. Ethical considerations associated with qualitative research methods. UK Statistics Authority. 2022. URL: https://uksa.statisticsauthority.gov.uk/publication/ethical-considerations-associated-with-qualitative-research-methods/pages/5 [accessed 2026-04-06]
64. Garvey CM, Jones R. Is there a place for theoretical frameworks in qualitative research? Int J Qual Methods 2021;20(5):1609406920987959. [doi: 10.1177/1609406920987959]
65. Luft JA, Jeong S, Idsardi R, Gardner G. Literature reviews, theoretical frameworks, and conceptual frameworks: an introduction for new biology education researchers. CBE Life Sci Educ 2022 Sep;21(3):rm33. [doi: 10.1187/cbe.21-05-0134] [Medline: 35759629]
66. Walter TG, Bricknell LK, Preston RG, Crawford EGC. Climate change adaptation methods for public health prevention in Australia: an integrative review. Curr Environ Health Rep 2024 Mar;11(1):71-87. [doi: 10.1007/s40572-023-00422-7] [Medline: 38221599]
67. George AZ. Research frameworks: critical components for reporting qualitative health care research. J Patient Cent Res Rev 2024;11(1):4-7. [doi: 10.17294/2330-0698.2068] [Medline: 38596348]
68. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]

XSL•FO
RenderX

69. Nowell LS, Norris JM, White DE, Moules NJ. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017;16(1). [doi: 10.1177/1609406917733847]
70. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
71. Vaismoradi M, Jones J, Turunen H, Snelgrove S. Theme development in qualitative content analysis and thematic analysis. J Nurs Educ Pract 2016;6(5):100-110. [doi: 10.5430/jnep.v6n5p100]
72. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide no. 131. Med Teach 2020 Aug;42(8):846-854. [doi: 10.1080/0142159X.2020.1755030] [Medline: 32356468]
73. Fife ST, Gossner JD. Deductive qualitative analysis: evaluating, expanding, and refining theory. Int J Qual Methods 2024;23(6). [doi: 10.1177/16094069241244856]
74. Naeem M, Ozuem W, Howell K, Ranfagni S. A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. Int J Qual Methods 2023;22. [doi: 10.1177/16094069231205789]
75. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
76. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006 Jun;27(2):237-246. [doi: 10.1177/1098214005283748]
77. Noble H, Smith J. Issues of validity and reliability in qualitative research. Evid Based Nurs 2015 Apr;18(2):34-35. [doi: 10.1136/eb-2015-102054] [Medline: 25653237]
78. Given LM. The SAGE Encyclopedia of Qualitative Research Methods: SAGE Publications; 2008. URL: https://books.google.co.in/books/about/The_Sage_Encyclopedia_of_Qualitative_Res.html?id=55JYAQAACAAJ&redir_esc=y [accessed 2026-04-06]
79. Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2 (2022). Government of Canada. URL: https://ethics.gc.ca/eng/policy-politique_tcps2-eptc2_2022.html [accessed 2026-04-15]

Abbreviations:

COREQ: Consolidated Criteria for Reporting Qualitative Research
DERRM: disaster and emergency risk reduction and management
KCN: kidney care network
VAA: vulnerability and adaptation assessment

Edited by A Schwartz; submitted 22.Dec.2025; this is a non–peer-reviewed article;accepted 24.Mar.2026; published 23.Apr.2026.

Please cite as:
El Wazze S, Mondal S, Sandal S
All-Hazards Vulnerability and Adaptation Assessment of Canadian Kidney Care Systems: Protocol for a Qualitative Study
JMIR Res Protoc 2026;15:e90059
URL: https://www.researchprotocols.org/2026/1/e90059
doi:10.2196/90059

© Saly El Wazze, Shinjini Mondal, Shaifali Sandal. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Supporting Unpaid Caregivers of Persons Living With Dementia: Protocol for a Pilot Feasibility Study to Explore Caregiver Outcomes and Impact of a Co-Designed Simulation–Based Psychoeducation Program in Virtual Reality

Mary Chiu[1,2], PhD; Sarah C Pistritto[2], MHSc; Kristina M Kokorelias[3,4,5], PhD; Krystina M Clarke[2], MHSc; Andrei Torres[6,7], PhD; Cole G Craven[6], MDM; Adam Dubrowski[2,7], PhD; Bill Kapralos[6,7,8], PhD; Erica O'Hare[1], MSc; Joel Sadavoy[9,10], MD; Adriana Shnall[11,12], PhD; Michael S D Smith[13], MSc; Nusrat Choudhury[14], MEng; Amer M Burhan[1,10], MSc, MBChB

[1]Research & Academics, Ontario Shores Centre for Mental Health Sciences, 700 Gordon St, Whitby, ON, Canada

[10]Department of Psychiatry, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[11]The Koschitzky Centre for Innovations in Caregiving, Baycrest Academy for Research and Education, Toronto, ON, Canada

[12]Factor Inwentash Faculty of Social Work, University of Toronto, Toronto, ON, Canada

[13]Medical Devices, National Research Council Canada, Winnipeg, ON, Canada

[14]Medical Devices, National Research Council Canada, Boucherville, ON, Canada

[2]Faculty of Health Sciences, Ontario Tech University, Oshawa, ON, Canada

[3]Section of Geriatrics, Sinai Health and University Health Network, Toronto, ON, Canada

[4]Rehabilitation Sciences Institute and Department of Occupational Sciences and Occupational Therapy, Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada

[5]KITE-University Health Network, Toronto, ON, Canada

[6]Faculty of Computer Science, Ontario Tech University, Oshawa, ON, Canada

[7]maxSIMhealth Group, Ontario Tech University, Oshawa, ON, Canada

[8]Game Development and Interactive Media, Ontario Tech University, Oshawa, ON, Canada

[9]Department of Psychiatry, Mount Sinai Hospital, Toronto, ON, Canada

Corresponding Author:
Mary Chiu, PhD
Research & Academics, Ontario Shores Centre for Mental Health Sciences, 700 Gordon St, Whitby, ON, Canada

Abstract

Background: Dementia is a global public health concern, with prevalence projected to reach 78 million individuals by 2030 and 139 million by 2050. Most persons living with dementia reside in community settings and are supported by family caregivers. As caregiving demands grow, caregivers experience significant psychosocial, emotional, and financial burden, including high rates of stress, social isolation, and depressive symptoms. Access to effective support services remains limited, highlighting the urgent need for innovative and accessible caregiver interventions.

Objective: This pilot study first aims to assess the feasibility, acceptability, and tolerability of VR-SIM Carers, a virtual reality (VR)–based psychoeducational training program for family caregivers or care partners of people living with dementia. Second, it will aim to provide a preliminary evaluation of potential impact on caregiver outcomes, including caregiver competence, stress, resilience, empathy, and quality of life. The study is not designed to support causal inference regarding the effectiveness of VR-SIM Carers.

Methods: A mixed methods design will be used with a sample of 30 family caregivers of people living with dementia. Participants will complete 3 immersive, self-paced VR caregiving scenarios—Managing Apathy, Crisis Response, and Refusal of Care—receiving real-time feedback from simulated characters, including clinician and person living with dementia avatars. Primary outcomes (feasibility and educational impact) include recruitment, retention, adherence, usability, acceptability, and tolerability and caregiver competence (Pearlin Caregiving Competence Scale), perceived stress (Cohen Perceived Stress Scale), resilience (Connor-Davidson Resilience Scale), and empathy (Empathy Assessment Scale). Secondary outcomes (preliminary efficacy) include caregiver quality of life (Adult Carer Quality of Life Questionnaire), caregiver burden (Burden Questionnaire), and behavioral symptoms (Neuropsychiatric Inventory, Center for Epidemiologic Studies Depression Scale Short Form) assessed at baseline, postintervention, and 1-month follow-up. Feasibility and user engagement will be evaluated via the 18-item Gaming Use Engagement and Severity Scale, Igroup Presence Questionnaire, qualitative interviews, reflection notes, and open-ended

XSL•FO
RenderX

feedback. Quantitative data will be analyzed using repeated-measures ANOVA and paired 2-tailed *t* tests, while qualitative data will be analyzed using an inductive thematic coding framework. Data analyses are descriptive and exploratory only, and no causal claims regarding intervention effectiveness will be made. Consistent with CONSORT (Consolidated Standards of Reporting Trials) guidance for pilot and feasibility studies, caregiver outcomes (eg, competence, stress, resilience, empathy) are treated as exploratory.

Results: The findings from this study will inform the feasibility, acceptability, and educational value of immersive VR for caregiver training, while providing preliminary evidence regarding the efficacy of VR-SIM Carers as a training tool to improve psychoeducation outcomes for family caregivers of people living with dementia and reduce caregiving burden. Data collection commenced in March 2024 with a projected end date of March 2026. As of the submission of the manuscript (December 2025), 21 participants have been enrolled. Data analysis will be completed in April 2026, and the results are expected to be published in fall 2026.

Conclusions: VR-SIM Carers represents an innovative, scalable intervention designed to enhance caregiver preparedness, psychosocial outcomes, and sustainable community–based dementia care. This pilot study will provide critical evidence to guide further development and implementation of VR-based caregiver support programs.

International Registered Report Identifier (IRRID): PRR1-10.2196/87107

(JMIR Res Protoc 2026;15:e87107) doi:10.2196/87107

KEYWORDS

protocol; dementia; family caregivers; resilience; empathy; virtual reality; simulation-based learning; caregiver stress; burden; experiential learning

Introduction

Background

Dementia is a growing global public health challenge with profound personal, societal, and economic implications. As the world's population ages, the prevalence of dementia is projected to increase dramatically: in 2019, an estimated 55.2 million people were living with dementia worldwide—a figure expected to rise to 78 million by 2030 and 139 million by 2050 [1]. This demographic shift is mirrored in Canada, where the incidence of dementia is rising alongside the aging population. By 2030, nearly 1 million Canadians are expected to be living with dementia, and this number is projected to double by 2050 [2]. In Canada, approximately 61% of people with dementia reside at home, instead of in long-term care facilities [3]. This trend reflects both the wishes of older adults to age in place and the increasing reliance on unpaid caregivers—primarily family members, friends, and neighbors—to provide daily support [3]. In low- and middle-income countries, where formal care services are often limited, family caregiving is nearly universal; however, even in high-income countries, such as Canada, the majority of people living with dementia continue to rely on unpaid care arrangements [4].

Unpaid caregivers are the backbone of dementia care, making a critical yet often under-recognized contribution to the health system and society. In Canada alone, family caregivers or care partners provide an estimated 330 million hours of unpaid care annually, a figure projected to nearly double to 640 million hours by 2040 [5]. The economic value of this contribution is staggering; in 2016, the combined health system and out-of-pocket costs for people with dementia reached $10.4 billion, with unpaid caregiving accounting for a substantial proportion of this total [5]. By 2050, the annual hours of unpaid care provided by friends and family are expected to reach 1.4 billion—the equivalent of over 690,000 full-time jobs [2].

Chronic Stress in Family Caregivers Associated with Higher Risk of Dementia

Caring for a person living with dementia is associated with substantial emotional, psychosocial, and financial burdens. Numerous studies have documented elevated rates of stress, anxiety, depression, and social isolation among family caregivers or care partners of people living with dementia [6,7]. A recent Canadian study found that 86% of those caring for both aging parents and their own children—sandwich generation caregivers—reported negative impacts on their health and well-being, including high levels of stress, exhaustion, and depressive symptoms [3]. Women constitute the majority of caregivers, and those who fall under the category of "sandwich generation" are particularly vulnerable to emotional distress and burnout, likely due to both cultural expectations and the hands-on nature of their caregiving roles [3,6]. Similarly, caregivers who are concurrently employed experience disproportionately high levels of stress combined with a low sense of mastery when faced with caring for mental and behavioral demands of older adults with dementia [8].

In particular, the chronic stress associated with dementia caregiving is a risk factor for increased risk of cognitive decline and dementia in caregivers themselves [7,9]. Prolonged exposure to caregiving stress is associated with elevated cortisol levels, immune dysregulation, and a greater incidence of cardiovascular disease and metabolic syndrome [9]. Furthermore, caregivers with high perceived stress are less able to engage in self-care, leading to a vicious cycle of declining health and reduced caregiving capacity [7]. Given these risks, interventions may need to consider both supporting caregivers in their roles but also addressing their own health and well-being. This is particularly crucial for family caregivers or care partners of people living with dementia who are burdened, stressed, and lacking the skills and knowledge to provide effective care.

XSL•FO
RenderX

Evidence-Based Interventions for Family Caregivers or Care Partners of People Living with Dementia

A broad array of interventions has been developed to support family caregivers or care partners of people living with dementia, including support groups, counseling, psychoeducational programs, and respite care [10,11]. Among these, psychoeducational and multimodal structured psychotherapy have shown promise in reducing caregiver burden, stress, and depressive symptoms while enhancing caregiver competence and emotional resilience [7,12]. Structured programs that combine disease-specific knowledge, practical caregiving skills, psychotherapy, and emotional support have been shown to improve caregiver competence, reduce anxiety, and foster effective care strategies [7,8,13]. Multimodal interventions that integrate education, communication skills training, and problem-solving techniques offer a holistic approach to managing the complex demands of dementia care [12]. Support groups, whether in-person or online, foster a sense of community and shared understanding, reducing social isolation and promoting psychological well-being [11,14].

Despite their effectiveness, many traditional interventions are limited by in-person delivery models that can be difficult for caregivers to access regularly, especially those balancing work and/or living in rural and underserved communities. Barriers, such as time constraints, transportation challenges, and stigma, further impede engagement with support services. There is a need for innovative, accessible, and evidence-based interventions that can be delivered flexibly and at scale, particularly for caregivers who are most burdened, stressed, and at risk of declining health.

Opportunity: Tech-Enabled Caregiver Support and Intervention

Simulation-based learning has long been a gold standard in professional health care education, offering a structured and psychologically safe environment to practice clinical skills, improve communication, and develop emotional intelligence [15]. Technologies, such as virtual reality (VR), can support the delivery of immersive simulation–based psychoeducation in an accessible, scalable, and engaging format. Recent studies have shown that VR-based training for both professional and family caregivers of people living with dementia can improve knowledge, compassion, and behavioral competence, in addition to supporting meaningful changes in caregiving practice [16]. Importantly, VR-based interventions can be designed to be self-paced and are potentially accessible from home or community settings, in urban or rural settings, offering flexibility and autonomy that are particularly valuable for caregivers juggling multiple responsibilities [16]. Interventions that include options for caregiver reflection, expert guidance, and resource linkage further enhance user experience and implementation potential [17-19]. Although evidence on effectiveness is still emerging, the preliminary findings are promising; immersive VR simulations can support empathy-building, skill acquisition, and confidence in care, all of which are central to caregiver preparedness, confidence, and well-being [19,20].

VR simulations not only enhance skill development but may also promote empathy by allowing caregivers to hear feedback from the character representing people living with dementia. When coupled with opportunities for guided reflection and clinical debrief, this form of intervention holds strong potential to shift caregiving behavior and improve caregiver outcomes. Moreover, as immersive technologies become more affordable and user-friendly, they represent a promising avenue for caregivers to access simulation-based, experiential learning and practice that is skills-based and emotionally engaging. However, few initiatives to date have rigorously explored the feasibility and acceptability of delivering simulation-based psychoeducation for family caregivers or care partners of people living with dementia through VR or evaluated its potential impact on caregivers' psychosocial and mental health outcomes.

VR-SIM Carers: Program at a Glance

The VR-SIM Carers Minimally Viable Prototype (MVP) is a self-paced, VR-based training tool designed to help family caregivers or care partners of people living with dementia develop practical caregiving strategies through immersive simulation–based learning. Developed using a co-design approach with input from older adults, caregivers, clinicians, simulation educators, and computer scientists, the MVP is accessible via commercially available headsets and aims to replicate real-world caregiving challenges, allowing users to learn and practice in a safe and controlled environment. The current prototype includes 3 interactive VR scenarios focused on challenging situations commonly encountered by family caregivers or care partners of people living with dementia. These scenarios are "high-fidelity," meaning they closely replicate real-world caregiver situations with a high degree of realism. The content was inspired by real-life caregiver experiences discussed in a clinical setting during the Reitman Centre CARERS Program [8,21]. The 3 scenarios included in the MVP—Managing Apathy, Crisis Response, and Refusal of Care—were chosen through a priority setting survey completed by 16 knowledge users, family caregivers or care partners of people living with dementia and health or social providers and clinicians. Each scenario takes approximately 10 minutes to complete and may be repeated as needed. The further functionality of the MVP is detailed in the *Methods* section.

Psychoeducation is a therapeutic approach focused on educating individuals about their mental health conditions, treatments, and coping strategies, empowering them with essential knowledge for effective management. The development of our VR-based psychoeducational program adheres to the VR-CORE framework described by Birckhead et al [22], a set of guidelines and recommendations for the development and testing of VR treatments through 3 phases (VR1, VR2, and VR3), akin to the U.S. Food and Drug Administration's Phase I-III framework for pharmaceuticals but adapted for VR. The objectives of this pilot study (see below) fit within the VR2 phase, which emphasizes early testing to assess the feasibility, acceptability, and tolerability of the VR intervention, along with its potential clinical efficacy. End user feedback will be obtained in order to explore the program's readiness and identify implementation challenges. Our program targets outcomes, such as enhancing caregiver relationships with people living with dementia, boosting caregiver competency and resilience, and reducing caregiver stress.

XSL•FO
RenderX

Research Objectives

The primary objective of the pilot study is to evaluate the feasibility, acceptability, and tolerability of immersive VR technology as a tool to deliver psychoeducation content to family caregivers or care partners of people living with dementia, as well as its readiness for implementation.

The secondary objective is to examine the initial efficacy of VR-based training in improving psychosocial outcomes of family caregivers or care partners of people living with dementia.

Methods

Methodological Approaches and Timeline

The pilot study will use a mixed methods approach to assess feasibility, acceptability, and tolerability of the VR-SIM Carers MVP and collect preliminary data on caregiver outcomes. Quantitative outcome measures and assessment tools will be administered at 3 time points: baseline, postintervention, and 1-month postintervention, as illustrated in Table 1. For both objectives, qualitative data will be collected through semistructured interviews with participants (Supplementary file B in Multimedia Appendix 1). In addition, multiple sources of qualitative data—1) field notes taken at the week 2 or 3 check-in phone call, 2) participants' written reflection notes, and 3) open text responses to survey questions at 1-month follow-up—will be analyzed to assess feasibility, acceptability, and tolerability of VR-SIM Carers from the perspectives of the caregivers. Data collection began in March 2025, and we anticipate all data analyses to be completed by April 2026. Reporting preliminary efficacy will follow the CONSORT (Consolidated Standards of Reporting Trials) extension for pilot and feasibility trials; a CONSORT flow diagram will be included to report numbers screened, eligible, enrolled, allocated, assessed at each time point, and retained, with the reasons for exclusions and losses to follow-up.

Table . Self-completed assessment scales and respective completion time points.

Scale or survey	Baseline (week 1)	Post intervention (week 4)	1-month follow-up
Exploratory outcomes and descriptors			
Caregiver competence	✓	✓	✓
Connor-Davidson Resilience Scale	✓	✓	✓
Empathy Assessment Scale	✓	✓	✓
Perceived Stress Scale	✓	✓	✓
Neuropsychiatric Inventory	✓	✓	✓
Quality of life scale	✓	✓	✓
Baseline descriptors only			
Center for Epidemiological Studies Depression Scale	✓		
The Burden Questionnaire	✓		
Demographic survey		✓	
Feasibility and tolerability			
18-item Gaming Use Engagement and Severity Scale			✓
Igroup Presence Questionnaire			✓
Qualitative responses (4 questions)			✓

Attrition, missing surveys, and incomplete scale data will be treated as feasibility outcomes and reported transparently at each assessment time point. Quantitative analyses will primarily use available case summaries. Where appropriate, simple descriptive sensitivity analyses (eg, last observation carried forward or comparison of complete vs incomplete cases) may be conducted to assess the robustness of the observed trends. No formal imputation or model-based inference is planned, consistent with CONSORT recommendations for pilot and feasibility studies. Standardized assessment tools for depressive symptoms (Center for Epidemiological Studies Depression Scale) and caregiver burden (the Burden Questionnaire) are completed at baseline only and are intended as descriptive sample characteristics and contextual variables rather than outcomes or indicators of intervention effectiveness.

Intervention Description

As aforementioned, the VR-SIM Carers MVP is a self-paced, VR-based training tool designed to support family caregivers and care partners of people living with dementia in developing practical caregiving strategies through immersive simulation–based learning. The MVP consists of three key components:

XSL•FO
RenderX

1. Functional system menu: Upon entering the VR-SIM Carers platform, users are greeted with a brief narration describing how to select and start a scenario. Users may select from 1 of 3 scenarios.
2. Three high-fidelity VR scenarios: They allow for simulation-based learning and practice. Users choose from responses presented to them to move forward in the scenarios, while receiving real-time guidance and feedback from 2 nonplayer characters in the VR scenario: a virtual clinician (Figure S1 in Multimedia Appendix 1) and a simulated person living with dementia (a virtual character playing the role of a person living with dementia in different settings). The 3 scenarios included are as follows:
    - Managing Apathy (Figure S2 in Multimedia Appendix 1): The user will learn strategies to interact with a "mother" character displaying symptoms of apathy.
    - Crisis Response (Figure S3 in Multimedia Appendix 1): The user will respond to and problem-solve a possible fire-hazard situation and interact with a "father" character who may be unaware of the smoking pan in the kitchen.
    - Refusal of Care (Figure S4 in Multimedia Appendix 1): The user will interact with a "father" character who is displaying resistance to care, specifically refusal to get dressed.
3. Feedback and reflection: To reinforce users' learning after scenario completion, custom feedback is provided by a holographic clinician avatar and simulated person living with dementia within the same scenario environment, based on the particular decisions selected by the user. Immediate feedback and guided reflections from these characters will help caregivers understand the emotional and practical implications of their decisions.

The intervention is intentionally designed to be self-paced and may be repeated to reflect real-world use by caregivers. No minimum exposure dose is specified beyond prespecified adherence thresholds, which are used solely to inform feasibility. Exposure will be operationalized descriptively using engagement metrics including the number of scenarios completed, total time spent in scenarios, and whether virtual clinician feedback was reviewed. These metrics will be summarized to characterize the patterns of use and feasibility. No post hoc exposure-based subgroup analyses or inferential dose-response analyses are planned, since such analyses would risk bias and overinterpretation in a small pilot sample.

Participants

Participants will be recruited from various partner sites and complete an expression of interest form to self-refer to the study. Partner sites are responsible for sharing our research ethics board (REB)-approved recruitment material for this study. To be eligible to participate, the following inclusion criteria should be met:

- 18 years of age or older.
- Able to provide informed consent.
- Current unpaid family caregivers or care partners of people living with dementia.
- Caring for an individual with either suspected or diagnosed dementia. A formal diagnosis is not required.
- No self-report of active or unstable major psychiatric conditions (eg, major depressive disorders, bipolar disorder, schizophrenia). The study co-principal investigator or joint senior author and geriatric psychiatrist AMB will make the final decisions if necessary.
- Able to read and speak English and able to (1) participate in prototype testing, (2) participate in a postintervention feedback survey, and (3) follow instructions from research staff and complete self-administered surveys or assessment tools.

Study Enrollment and Timeline

Upon enrollment in the study, participants will complete three key activities (Textbox 1):

1. In-person orientation session: The study team led by primary author (MC) will guide participants one-on-one on how to safely use and take care of a VR-SIM Carers VR headset. A commercially available VR headset (Meta Quest 3), with the VR-SIM Carers software installed, will be loaned to participants for 4 weeks; a user manual and other supplementary materials will also be provided for review.
2. Complete high-fidelity simulation scenarios: Using the VR headset, participants will engage in a series of self-paced immersive scenarios—Managing Apathy, Crisis Response, and Refusal of Care—that represent caregiving situations commonly encountered by family caregivers or care partners of people living with dementia. Participants will review the user manual to guide the completion of the 3 caregiving scenarios (8-10 minutes each) on the VR-SIM Carers headset at their own pace. After completing the scenarios, users will write reflections in the user manual following predefined prompts.
3. Self-assessment tools: Participants will complete a set of 9 brief, self-assessment tools (20-30 minutes) at 3 time points (before, after, and 1 month after completing the VR scenarios). These will be emailed to participants and delivered through SimpleSurvey [23].

XSL•FO
RenderX

Textbox 1. Timeline or workflow of research activities to be completed by study participants.

Week 1: In-person orientation event (90 min)

- In-person visit at the recruitment site
- Complete REB-approved consent form and baseline assessment scales (via SimpleSurvey)
- Sign out study VR headset; complete orientation RE: headset use led by the study team. Complete Simulator Sickness Questionnaire

Weeks 2 and 3: Self-paced learning on VR-SIM Carers

- Self-completion of VR simulation scenarios
- Review virtual clinician's feedback; program resources
- Complete check-in phone call from the study team

Week 4: In-person visit by study team (1 h)

- Complete time 2 assessment scales (via SimpleSurvey)
- Semistructured interview to capture participants' experience with VR-SIM Carers
- Return study virtual reality headset; submit written reflection notes

One month after

- Complete time 3 assessment scales (via SimpleSurvey)
- Provide written feedback on overall study experience (via SimpleSurvey)

Sample Size

Thirty caregivers will be recruited for the study. A cohort of 5 caregiver participants will be recruited from each recruitment site for equal representation, and enrollment will continue until the targeted number is reached. Recruitment sites are Ontario Shores Centre for Mental Health Sciences in Oshawa, Ontario; Baycrest Centre in Toronto, Ontario; Carefirst Community & Senior Services Association, in Toronto, Ontario; London Health Sciences Centre in London, Ontario; and St. Mary's Hospital in Montreal, Quebec. Recruitment will take place once local ethics clearance has been received from each site's research ethics board.

Power Calculation

Based on previous literature, a minimum sample size of 20 participants has been identified as appropriate for a pilot study to provide preliminary estimates of variability and to evaluate key feasibility parameters. These estimates are informed by existing caregiver pilot study interventions [24,25]. To account for expected attrition, a total sample size of 30 participants will be recruited. This target provides a buffer to maintain the minimum sample needed to evaluate feasibility outcomes and to estimate variance for future sample size calculations, while accounting for potential data loss due to loss to follow-up.

Progression Criteria

We prespecified quantitative and qualitative progression criteria to guide decisions about whether to proceed to a future definitive trial. Key thresholds are recruitment success 80% or higher (≥24/30 participants) of the target, amend-and-reassess if 50% to 79% (15-23/30) are enrolled, and stop or redesign if less than 50% (<15 participants) are enrolled. Retention success is 80% or higher completion of both postintervention and 1-month follow-up assessments, amend if 60% to 79% (18-23/30) complete, and stop if less than 60% (<18 participants). Intervention adherence success is 75% or more of the participants (≥23/30 participants) completing at least 2 of 3 scenarios and reviewing virtual clinician feedback during the loan period, amend if 50% to 74% (15-22/30), and stop if less than 50% (<15 participants). Data completeness success is 90% or higher (≥27/30) at baseline and 80% or higher (≥24/30) at post-intervention and 1-month follow-up for primary caregiver measures; amend for baseline 75% to 89% (≥23/30) or post-intervention or 1-month 60% to 79% (18-23/30). Safety or tolerability success is 15% or lower (<4 participants) reporting moderate or severe simulator sickness or adverse events attributable to device or software and no serious adverse events, amend if 16% to ‐30% (5-9/30) report moderate symptoms that are remediable, or stop if more than 30% (>9 out of 30 participants) report moderate or severe symptoms or any serious adverse event attributable to the intervention. Acceptability success is 70% or higher (≥21/30 participants) reporting willingness to use the MVP again or recommend it, amend if 50% to 69% (15-21/30), or stop if less than 50% (<15/30). Qualitative signals from interviews, reflection notes, and field notes will be used alongside these thresholds to interpret feasibility. Decision rules are to proceed to a definitive trial without substantive changes if most key criteria meet success thresholds, proceed with specified protocol modifications if criteria fall in the amend range but no major safety signals exist, and stop and redesign if multiple key criteria fall in the stop range or qualitative data reveal fundamental, unresolvable concerns. All progression decisions will be reported and justified in the Results section, and the CONSORT pilot flow diagram will present numbers screened, enrolled, completed at each time point, and the reasons for nonparticipation or dropout.

XSL•FO
RenderX

Outcome Measures

The primary focus of this pilot is feasibility. Feasibility outcomes that will be reported include recruitment and enrollment rates, consent rate, retention and follow-up rates, adherence to the intervention (eg, time spent in scenarios and number of scenarios completed), data completeness, and adverse events or dropouts related to tolerability. Acceptability and usability outcomes will be assessed through participant responses and standardized instruments. Caregiver outcomes are collected as exploratory measures to describe change over time and to estimate variability for future studies. These caregiver outcome measures, completed at the 3 time points where indicated (Table 1), include:

- Caregiver competence: Caregiver Competence Scale [26], a 4-item scale assessing perceptions of ability to manage caregiving tasks effectively.
- Resilience: Connor-Davidson Resilience Scale [27], a 10-item self-report questionnaire assessing ability to cope with adversity.
- Empathy: Empathy Assessment Scale [28], a 13-item scale evaluating cognitive and affective components relevant to caregiving.
- Stress: Perceived Stress Scale [29], a 10-item scale measuring perceived stress.
- Secondary exploratory measures include behavioral disturbances in people living with dementia (Neuropsychiatric Inventory) [30], depressive symptoms (Center for Epidemiologic Studies Depression Scale) [31], caregiver burden (the Burden Questionnaire) [32], and quality of life (Adult Carer Quality of Life Questionnaire) [33].

VR-SIM Carers is also being evaluated for its feasibility, acceptability, and tolerability on caregiver outcomes, and these will be assessed at multiple points. For technical feasibility, participants will be asked about their ability to carry out the study workflow and the amount of time they engaged with the VR scenarios. Recruitment and retention will be tracked to inform feasibility. For acceptability, participants will be asked about willingness to use the MVP on the expression of interest form and consent form. For tolerability, participants will be asked about any side effects related to hardware or software at the week 1 phone check-in, and the number of dropouts due to adverse events will be tracked.

Participants will also complete evidence-based scales that inform acceptability and tolerability. The 18-item Gaming Use Engagement and Severity Scale [34] assesses user satisfaction, usability, and system quality, while the Igroup Presence Questionnaire [35] measures spatial presence, involvement, and realism in VR environments. The Simulator Sickness Questionnaire documents the type and severity of side effects due to hardware and/or software use.

Data Analysis

Qualitative Data

An inductive thematic analysis will be conducted on qualitative data collected from the week 4 audio-recorded interviews with study participants. Interviews will take 30 to 45 minutes and will be facilitated by SCP, KMC, and EO following a semistructured interview guide (Supplementary file B in Multimedia Appendix 1). Secondary sources of qualitative data will include participant reflection notes, the week 2 or 3 check-in phone call, and the 1-month qualitative written feedback survey questions. This multimodal approach is intended to capture participants' experiences and perspectives over time. The audio recordings of the semistructured interviews will be transcribed verbatim by research staff and will serve as the primary qualitative data source. Transcripts, reflection notes, and field notes will be imported to Dedoose (9.0.107) for analysis [36].

To create the codebook in Dedoose, 3 (MC, SCP, and KMC) members of the study team trained in coding will independently review 3 interview transcripts to generate an initial set of codes and to collaboratively develop a codebook through consensus [23]. To reduce potential bias, the member who transcribed an audio-recorded interview will not be assigned to code that same transcript. A tracking sheet will be created to manage coding assignments and ensure consistent review.

Quantitative data

Descriptive statistics will summarize participants' demographic characteristics and feasibility metrics. Statistical analysis will be completed using SPSS (version 29; IBM Corp) [37]. Paired 2-tailed *t* tests may be used to assess within-subject changes between specific time points as exploratory analyses. Repeated measures statistics will be used to describe change across the 3 time points. Consistent with CONSORT guidance for pilot and feasibility studies, caregiver outcomes (eg, competence, stress, resilience, empathy) are treated as exploratory, and no causal claims regarding intervention effectiveness will be made. Observed changes over time may reflect nonspecific effects such as time, expectancy, repeated measurement, or general study contact (eg, orientation and check-in calls). Repeated measures are collected to describe the direction and magnitude of change over time and assess the feasibility and sensitivity of the selected measures in this population and delivery format. All quantitative analyses are intended to estimate variability and inform sample size calculations and protocol decisions for a future definitive trial. A 2-sided P value of $<.05$ may be reported for descriptive purposes, but hypothesis testing is not the primary aim of this pilot.

Study Monitoring

Upon enrollment, each participant will be assigned a unique study number. Any research-related information collected (eg, surveys, scales, and questionnaires) will be labeled and stored using this study number to deidentify personal information and uphold confidentiality. A master list linking participant names to their unique study number will be stored separately on the Ontario Shores secured network, which will be password-protected and accessible only to authorized study personnel. Written documents (eg, consent forms) will be stored in a locked filing cabinet at the Ontario Shores Research and Academics Office, with access restricted to study-approved staff.

Survey and deidentified interview data will be password-protected and accessible only to authorized study

XSL•FO
RenderX

personnel. Audio recording files will be deleted once transcribed, and identifying information will be removed. Check-in call field notes will be deidentified as well. All study data will be stored on the Ontario Shores Secured Network Drive and will be aggregated for reporting. All assessment tools and scales will be provided on SimpleSurvey [38], a secure survey platform that will allow each participant easy access to completing surveys at their own pace. The data will remain in SimpleSurvey for 60 days following termination of the account in November 2025 [39]. The data collected through this research study will be stored securely for a 7-year period and will be monitored and reviewed by study investigators to ensure confidentiality protocols are upheld, following REB guidelines and protocol procedures.

Patient and Public Involvement

As per the guidelines of the VR-CORE framework (VR1 phase), the MVP was shaped through a collaborative and iterative process involving demonstration outreach and knowledge summit activities [40,41]. Initial exposure of the pre-MVP prototype to attendees from these events provided the opportunity for diverse backgrounds (clinicians, researchers, and potential end users) to provide feedback on their interaction. Although this feedback did not contribute to the overall study design, attendees' feedback provided input regarding usability, technological requirements, clinical relevance, and practical considerations that informed the iterative development of the MVP.

As part of these outreach efforts, a detailed manual was presented to all participants at the knowledge summit to guide them through the MVP's features and activities and support more structured reflection and feedback. This helped to ensure users had a basic understanding of the prototype's intended functions and use. Additionally, this project benefited from an advisory board that included knowledge users (Ron Beleno and Carla Velastegui), who provided further review and guidance.

To assess the feasibility of the MVP developed, protocol, and user experience, 3 clinician externs from Ontario Shores Centre for Mental Health Sciences engaged in a mock structured run-through of the protocol. First, an orientation session was conducted; then, externs took the VR headset home for 1 week to experiment with the scenarios and review the manual. Their involvement allowed for the real-time identification of the duration of our set activities and addressed any accessibility or technological concerns about the MVP and orientation flow. Feedback led to improvements in system navigation, onboarding materials, and improved preparedness of the study team for subsequent caregiver cohort orientation sessions.

Ethical Considerations

This study has received ethical approval from the Ontario Shores Centre for Mental Health Sciences Joint Research Ethics Board (JREB 22 - 031-B) and secondary approval from the NRC Research Ethics Board (NRC-REB 2023 - 23). All participants will provide written informed consent at the time of the in-person orientation session, prior to taking part in any study-related activities. Participation is voluntary and compensated as approved by the research ethics boards, and all data will be deidentified and stored securely in accordance with institutional data protection guidelines. Any identifying information will be kept separate from research data and accessible only to trained research personnel. Participants will be informed of their right to withdraw from the study at any time without consequence. Study findings will be broadly disseminated through peer-reviewed publications, conference presentations, and reports to stakeholders and funders. Knowledge users' feedback gathered throughout the pilot study will inform future iterations of the VR-SIM Carers MVP and implementation. In addition, a publicly accessible project website will feature blog posts and community updates to extend our outreach effort. By sharing our findings with both academic and community audiences, we aim to contribute to capacity building, as well as establishing the evidence base for scalable, technology-enabled psychoeducation tools for family caregivers or care partners of people living with dementia.

As aforementioned, local ethics clearance has also been received from each recruitment site's research ethics board.

Results

The forthcoming results of this pilot study will provide preliminary evidence on the feasibility, acceptability, tolerability, and preliminary efficacy of the VR-SIM Carers MVP as a psychoeducational tool for family caregivers of people living with dementia. Quantitative results will report on changes in caregiver competence, resilience, empathy, and stress across 3 time points (baseline, postintervention, and 1-month follow-up), as well as secondary outcomes, such as caregiver burden, quality of life, depressive symptoms, and perceptions of behavioral disturbances in people living with dementia. These findings will indicate whether the VR-SIM Carers intervention holds promise in improving psychosocial outcomes for caregivers, while also highlighting areas for refinement prior to larger-scale trials.

Qualitative results, derived from interviews, reflection notes, and written survey responses, will complement the quantitative findings by exploring participants' lived experiences with the intervention, including perceived benefits, usability, and barriers to engagement. Additional feasibility outcomes, including recruitment and retention rates, adverse events, and system usability scores, will further inform whether the MVP is suitable for broader testing.

Funding for the overall project commenced in September 2022, with funds for the study as described in this protocol paper coming in April 2024. Data collection began in March 2025 with a projected end date of March 2026. As of the submission of the manuscript (December 2025), 21 participants have been enrolled. Analysis of both quantitative and qualitative data is expected to be completed by April 2026. The results will then be published in fall 2026 through academic publications, conference presentations, and community-facing outputs, such as reports and project website updates. These findings will directly inform the next stage of intervention development and scaling.

XSL•FO
RenderX

Discussion

Principal Results

The current protocol will be the first to evaluate the feasibility of a simulation-based psychoeducation prototype delivered in an immersive VR environment called "VR-SIM Carers." The study will first assess feasibility, usability, and educational outcomes of using immersive VR to support family caregivers or care partners of people living with dementia. In addition, this study will explore the preliminary efficacy of VR-SIM Carers in improving psychosocial outcomes such as resilience, empathy, and caregiving competence, measured using standardized assessment tools. The qualitative component of the study will explore users' experiences with the VR technology and the applicability of psychoeducational content in day-to-day, real-life caregiving encounters.

A key strength of the study is the innovative use of commercially available VR technology to deliver high-fidelity caregiving simulations co-designed with knowledge users, such as family caregivers or care partners of people living with dementia and older adults. The participatory approach used in developing the prototype ensured it was meaningfully co-designed with, and for, caregivers. Users' feedback was integrated into the design process to ensure content relevance and usability. Importantly, content and technology development without co-design would be unlikely to achieve the same level of relevance, acceptability, or impact, as interventions developed in isolation from knowledge users often fail to address real-world needs and contexts [40,42]. Orientation and technical support may reduce barriers to VR adoption and encourage participation among users who may not be comfortable with technology use. This delivery approach and the self-paced nature of the program enhance accessibility for caregivers, allowing them to learn at a time and pace that suits their individual needs, which improves scalability. Additionally, the MVP for this study was developed according to the VR-CORE framework's best practices for content development and user-centered design (VR1 phase), actively involving patients and health care providers.

Another strength is the incorporation of both quantitative and qualitative data collection across 3 time points. Quantitative assessments via validated scales (eg, Caregiver Competence Scale, Connor-Davidson Resilience Scale, Empathy Assessment Scale, Perceived Stress Scale) alongside qualitative data obtained from reflection notes, semistructured interviews, and written reflections are expected to yield nuanced insights into outcomes and users' learning experiences and acceptance of the VR-based intervention.

The findings from this pilot study will provide foundational evidence for the acceptability and feasibility of immersive, simulation–based psychoeducation delivered via VR for family caregivers or care partners of people living with dementia. Should the VR-SIM Carers intervention demonstrate positive outcomes in reducing caregiver stress and improving resilience, empathy, and caregiving competence, future research will focus on large-scale, multisite randomized controlled trials to confirm efficacy and generalizability across diverse caregiver populations and care settings [16]. Further investigation is warranted to identify which caregiver subgroups benefit most, optimal frequency and duration of VR-based interventions, and ways to integrate such technology into existing community and health system supports.

To enhance clinical relevance and impact, future studies of VR-SIM Carers may incorporate the assessment of dementia staging to better target and tailor the intervention for family caregivers or care partners most likely to benefit [43,44]. Longitudinal research will be critical to assess sustained impacts on caregiver well-being, care recipient outcomes, and potential reductions in health system utilization or delayed institutionalization [45]. Additionally, qualitative studies exploring user experience, barriers to adoption, and co-design with caregivers from marginalized or rural communities will be essential to ensure broad accessibility and cultural relevance [16]. Clinically, if VR-SIM Carers proves feasible and acceptable, it could be incorporated as a scalable adjunct to traditional caregiver education, offering flexible, experiential training that complements in-person or online resources. Orientation materials may be adopted into an online Learning Management System to further improve accessibility. Integrating immersive VR simulation into caregiver support programs may ultimately contribute to more sustainable, person-centered dementia care by empowering caregivers with practical skills and emotional coping strategies [12,16].

Limitations

Several limitations should be acknowledged. The small sample size, although appropriate for a pilot study, limits the statistical power to detect moderate effects and reduces generalizability. For this initial pilot, only English-speaking participants are included, and participants will not be experiencing significant psychological distress, which may exclude some caregivers most in need of support. The pilot's self-referral recruitment process through the expression of interest form may introduce selection bias, favoring individuals more open to using novel technology. Additionally, variability in participants' familiarity with VR technology could affect the consistency of the intervention's delivery. The in-person orientation session, although centrally located, may be a barrier to participation. Finally, individuals in the early stages of dementia typically do not exhibit the behavioral and psychological symptoms of dementia that the VR-SIM Carers MVP is designed to address; as such, the intervention may have limited relevance for caregivers supporting individuals in these early stages [43,44]. These limitations will be carefully considered in interpreting pilot findings and designing future larger-scale studies in diverse caregiver populations (eg, cultural groups) and settings (eg, community vs long-term care).

Comparison With Prior Work Published in the Literature

Traditional caregiver support programs (in-person or online) have been shown to significantly reduce caregiver strain [46]. Previous VR interventions for caregivers (eg, dementia-simulation programs) demonstrated modest gains in competence [47]. VR-SIM Carers is novel in its integration of evidence-based psychoeducational content into a high-fidelity, self-paced VR simulation. This experiential format goes beyond

XSL•FO
RenderX

the didactic approaches of earlier studies, potentially amplifying learning through realistic immersion. The VR-SIM Carers prototype was developed iteratively with input from family caregivers and clinicians, following the VR-CORE framework's recommendation to engage end users in content design (VR1 phase). Involving caregivers as co-designers enhances relevance and acceptability; indeed, research shows that active patient and caregiver involvement in intervention design improves outcomes and satisfaction. Our participatory approach helps ensure that scenarios and lessons reflect genuine day-to-day challenges faced by caregivers.

Earlier VR-for-caregiver studies often focused on a narrow set of outcomes. For example, a recent meta-analysis found that VR training modestly improved caregiver competence [47]. In comparison, our protocol measures a broad suite of psychosocial outcomes (caregiver stress, resilience, empathy, competence) using validated scales at multiple time points. Coupling these quantitative measures with qualitative interviews and reflections provides richer insights into both the efficacy of the intervention and users' learning processes. This mixed methods evaluation is more comprehensive than many prior trials, which typically lacked longitudinal follow-up or in-depth user feedback.

Conclusions

This protocol describes a pilot study to assess a pioneering VR-based intervention that leverages self-paced, immersive simulation to provide psychoeducational content for family caregivers or care partners of people living with dementia. The study is poised to offer valuable insights into the feasibility, acceptability, and preliminary efficacy of such an approach, despite inherent limitations related to sample size and variability in technology acceptance. As immersive technologies become more integrated into caregiver support and skills-building, it is essential to evaluate their real-world effectiveness, usability, and ethical implications. This protocol lays the groundwork for advancing VR-based psychoeducation as a tool to build caregiver competence, resilience, and well-being—an urgent priority in the context of caregiver stress and growing dementia care burden. The outcomes of this pilot study could inform the design and delivery of future VR caregiver interventions and could facilitate building partnerships across the dementia care, technology, and community support sectors. The careful, evidence-based advancement of VR-SIM Carers and similar interventions holds meaningful potential to complement existing resources and address the evolving needs of family caregivers or care partners of people living with dementia in Canada and beyond.

Acknowledgments

This study and the "VR-SIM Carers initiative" are supported by the National Research Council Canada, NRC Aging in Place Program, Canadian Institute of Health Research, and the institutional host Ontario Shores Centre for Mental Health Sciences (Ontario Shores). The authors would like to thank Carla Velastegui and Ron Beleno for reviewing and providing comments on the manuscript. The authors would also like to thank EO for coordinating the initiative and this study and Dhakshi Krishnalingam for her editing and formatting support.

Generative artificial intelligence tools were not used in the writing, editing, or preparation of this manuscript. All content was written, reviewed, and revised by the listed authors.

Funding

This work was supported by National Research Council Canada and the Canadian Institutes of Health Research (CIHR 202203AIP-483968).

Data Availability

There are no data to report.

Authors' Contributions

MC is responsible for the overall content, originally planned the protocol with NC and AMB, wrote the initial draft, and is the guarantor of this work. MC, SCP, KMC, and KMK contributed text to the Introduction and Methods sections. JS, MSDS, NC, and AMB critically reviewed and edited an initial draft. All other authors (AT, CGC, BK, and AD) provided feedback and revisions and approved the final version for submission.
NC and AMB are joint senior authors.

Conflicts of Interest

AMB has received research grant support from NIA/NIH, CIHR, Brain Canada Foundation, Weston Foundation, CABHI, CCNA, Western University, LHRI, Ontario Shores Foundation and Research Chair support from Ontario Shores Hospital; speaker/consulting fees from Eisai, Avenir, Janssen, Roche, Otsuka/Lundbeck, Boehringer Ingelheim, Amica Seniors Lifestyles, Health Tech Connex, atheneum partners, and contracting investigator with Headland Research (Toronto Memory Program). All other authors declare no conflicts of interest.

XSL•FO
RenderX

Multimedia Appendix 1
Supplementary materials: Appendix A: Screenshots of non-player characters and virtual environments in the 3 scenarios; Appendix B: Semi-Structured Interview Guide for post-intervention interview with study participants.
[DOCX File, 675 KB - resprot_v15i1e87107_app1.docx]

Peer Review Report 1
Peer review report by the NRC-CIHR Aging in Place Call 2021 Review Committee, National Research Council/Canadian Institutes of Health Research (Canada).
[PDF File, 165 KB - resprot_v15i1e87107_app2.pdf]

References

1. World failing to address dementia challenge. World Health Organization. URL: https://www.who.int/news/item/02-09-2021-world-failing-to-address-dementia-challenge [accessed 2026-02-05]
2. Navigating the path forward for dementia in Canada. : Alzheimer Society of Canada; 2022 URL: https://alzheimer.ca/sites/default/files/documents/Landmark-Study-Report-1-Path_Alzheimer-Society-Canada.pdf [accessed 2026-02-05]
3. "Sandwiched" between unpaid care for children and care-dependent adults: a gender-based study. : Statistics Canada, Government of Canada; 2024 URL: https://www150.statcan.gc.ca/n1/en/pub/89-652-x/89-652-x2024002-eng.pdf?st=GjdLCyrx [accessed 2026-02-05]
4. Hogan U, Bingley A, Morbey H, Walshe C. The experience of informal caregivers in providing patient care in hospitals in low- and middle-income countries: a qualitative meta-synthesis. J Health Serv Res Policy 2022 Oct;27(4):321-329. [doi: 10.1177/13558196221101968] [Medline: 35592881]
5. Prevalence and monetary costs of dementia in Canada. : Alzheimer Society of Canada; 2016 URL: https://alzheimer.ca/sites/default/files/documents/prevalenceandcostsofdementia_en_0.pdf [accessed 2026-02-05]
6. Brodaty H, Donkin M. Family caregivers of people with dementia. Dialogues Clin Neurosci 2009;11(2):217-228. [doi: 10.31887/DCNS.2009.11.2/hbrodaty] [Medline: 19585957]
7. Liu J, Lou Y, Cheung ESL, Wu B. Use of tangible, educational and psychological support services among Chinese American dementia caregivers. Dementia (London) 2022 Aug;21(6):1914-1932. [doi: 10.1177/14713012221099781] [Medline: 35511805]
8. Sadavoy J, Sajedinejad S, Duxbury L, Chiu M. The impact on employees of providing informal caregiving for someone with dementia. Aging Ment Health 2022 May;26(5):1035-1043. [doi: 10.1080/13607863.2021.1871878] [Medline: 33459037]
9. Vitaliano PP, Murphy M, Young HM, Echeverria D, Borson S. Does caring for a spouse with dementia promote cognitive decline? A hypothesis and proposed mechanisms. J Am Geriatr Soc 2011 May;59(5):900-908. [doi: 10.1111/j.1532-5415.2011.03368.x] [Medline: 21568959]
10. Chien LY, Chu H, Guo JL, et al. Caregiver support groups in patients with dementia: a meta-analysis. Int J Geriatr Psychiatry 2011 Oct;26(10):1089-1098. [doi: 10.1002/gps.2660] [Medline: 21308785]
11. Carter G, Monaghan C, Santin O. What is known from the existing literature about peer support interventions for carers of individuals living with dementia: a scoping review. Health Soc Care Community 2020 Jul;28(4):1134-1151. [doi: 10.1111/hsc.12944] [Medline: 31919974]
12. Amin Abdelhalim DS, Ahmed MM, Hussein HA, Sarhan MD, Khalaf OO. A skill-based multimodal intervention for dementia caregivers: impact on burden and anxiety. Aging Clin Exp Res 2025 Mar 17;37(1):95. [doi: 10.1007/s40520-025-02985-x] [Medline: 40095192]
13. Frias CE, Garcia-Pascual M, Montoro M, Ribas N, Risco E, Zabalegui A. Effectiveness of a psychoeducational intervention for caregivers of people with dementia with regard to burden, anxiety and depression: a systematic review. J Adv Nurs 2020 Mar;76(3):787-802. [doi: 10.1111/jan.14286] [Medline: 31808211]
14. Gallagher-Thompson D, Coon DW. Evidence-based psychological treatments for distress in family caregivers of older adults. Psychol Aging 2007 Mar;22(1):37-51. [doi: 10.1037/0882-7974.22.1.37] [Medline: 17385981]
15. Lateef F. Simulation-based learning: just like the real thing. J Emerg Trauma Shock 2010 Oct;3(4):348-352. [doi: 10.4103/0974-2700.70743] [Medline: 21063557]
16. Scerbe A, O'Connell ME, Astell A, et al. Digital tools for delivery of dementia education for caregivers of persons with dementia: a systematic review and meta-analysis of impact on caregiver distress and depressive symptoms. PLoS One 2023;18(5):e0283600. [doi: 10.1371/journal.pone.0283600] [Medline: 37196022]
17. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implementation Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0]
18. Carroll J, Hopper L, Farrelly AM, Lombard-Vance R, Bamidis PD, Konstantinidis EI. A scoping review of augmented/virtual reality health and wellbeing interventions for older adults: redefining immersive virtual reality. Front Virtual Real 2021;2:655338. [doi: 10.3389/frvir.2021.655338]
19. Dickinson R, Kimball J, Fahed M, Chang T, Sekhon H, Vahia IV. Augmented reality (AR) in dementia care: understanding its scope and defining its potential. Am J Geriatr Psychiatry 2023 Mar;31(3):S132-S133. [doi: 10.1016/j.jagp.2022.12.183]

XSL•FO
RenderX

20. Jones C, Jones D, Moro C. Use of virtual and augmented reality-based interventions in health education to improve dementia knowledge and attitudes: an integrative review. BMJ Open 2021 Nov 1;11(11):e053616. [doi: 10.1136/bmjopen-2021-053616] [Medline: 34725081]
21. Carey JM, Rossler K. The how when why of high fidelity simulation. In: StatPearls [Internet]: StatPearls Publishing; 2025. URL: http://www.ncbi.nlm.nih.gov/books/NBK559313 [accessed 2025-07-10]
22. Birckhead B, Khalil C, Liu X, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973. [doi: 10.2196/11973] [Medline: 30702436]
23. Giovannetti O, Tomalty D, Greco S, et al. Patient and provider perspectives on LEEP/LLETZ treatment and outcomes: an interview study. J Sex Med 2023 Jun 28;20(7):977-990. [doi: 10.1093/jsxmed/qdad026] [Medline: 36918387]
24. Meyer K, Glassner A, Norman R, et al. Caregiver self-efficacy improves following complex care training: results from the learning skills together pilot study. Geriatr Nurs (Lond) 2022 May;45:147-152. [doi: 10.1016/j.gerinurse.2022.03.013]
25. Reynolds L, Buchanan B, Crabtree S, Bordanave E. A program to improve the quality of life of caregivers of older adults with chronic conditions: a pilot study. Open J Occup Ther 2023;11(2):1-11. [doi: 10.15453/2168-6408.2073]
26. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
27. Connor KM, Davidson JRT. Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). Depress Anxiety 2003;18(2):76-82. [doi: 10.1002/da.10113] [Medline: 12964174]
28. Malakcioglu C. Empathy Assessment Scale. North Clin Istanb 2022;9(4):358-366. [doi: 10.14744/nci.2022.55649] [Medline: 36276566]
29. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
30. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The neuropsychiatric inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994 Dec;44(12):2308-2308. [doi: 10.1212/WNL.44.12.2308]
31. Center for Epidemiological Studies-Depression. American Psychological Association. URL: https://www.apa.org/pi/about/publications/caregivers/practice-settings/assessment/tools/depression-scale [accessed 2026-02-05]
32. Pendergrass A, Malnis C, Graf U, Engel S, Graessel E. Screening for caregivers at risk: extended validation of the short version of the Burden Scale for Family Caregivers (BSFC-s) with a valid classification system for caregivers caring for an older person at home. BMC Health Serv Res 2018 Apr 2;18(1):229. [doi: 10.1186/s12913-018-3047-4] [Medline: 29609600]
33. Elwick H, Joseph S, Becker S, Becker F. Manual for the Adult Carer Quality of Life Questionnaire (AC-QoL). : The Princess Royal Trust for Carers; 2010 URL: https://studylib.net/doc/18125085/adult-carer-quality-of-life-questionnaire--ac-qol-? [accessed 2026-02-05]
34. Phan MH, Keebler JR, Chaparro BS. The development and validation of the Game User Experience Satisfaction Scale (GUESS). Hum Factors 2016 Dec;58(8):1217-1247. [doi: 10.1177/0018720816669646] [Medline: 27647156]
35. Berkman M, Çatak G. I-group presence questionnaire: psychometrically revised English version. Mugla J Sci Technol 2021 Oct 16;7:1-10. [doi: 10.22531/muglajsci.882271]
36. Meet Dedoose: great research made easy. Dedoose. 2023. URL: https://www.dedoose.com/ [accessed 2026-02-05]
37. IBM SPSS software. IBM. URL: https://www.ibm.com/products/spss [accessed 2025-07-22]
38. Canada's online survey software. Simple. Secure. Trusted. SimpleSurvey. URL: https://simplesurvey.com/ [accessed 2026-02-05]
39. Simple survey master subscription agreement. SimpleSurvey. URL: https://form.simplesurvey.com/oss/external/msa.aspx?lang=EN+-+Section+15%29 [accessed 2026-02-05]
40. Greenhalgh T, Jackson C, Shaw S, Janamian T. Achieving research impact through co-creation in community-based health services: literature review and case study. Milbank Q 2016 Jun;94(2):392-429. [doi: 10.1111/1468-0009.12197] [Medline: 27265562]
41. Masterson D, Areskoug Josefsson K, Robert G, Nylander E, Kjellström S. Mapping definitions of co-production and co-design in health and social care: a systematic scoping review providing lessons for the future. Health Expect 2022 Jun;25(3):902-913. [doi: 10.1111/hex.13470] [Medline: 35322510]
42. Slattery P, Saeri AK, Bragge P. Research co-design in health: a rapid overview of reviews. Health Res Policy Syst 2020 Feb 11;18(1):17. [doi: 10.1186/s12961-020-0528-9] [Medline: 32046728]
43. Cerejeira J, Lagarto L, Mukaetova-Ladinska EB. Behavioral and psychological symptoms of dementia. Front Neurol 2012;3:73. [doi: 10.3389/fneur.2012.00073] [Medline: 22586419]
44. Fazio S, Pace D, Maslow K, Zimmerman S, Kallmyer B. Alzheimer's association dementia care practice recommendations. Gerontologist 2018 Jan 18;58(suppl_1):S1-S9. [doi: 10.1093/geront/gnx182] [Medline: 29361074]
45. Boots LMM, de Vugt ME, van Knippenberg RJM, Kempen G, Verhey FRJ. A systematic review of Internet-based supportive interventions for caregivers of patients with dementia. Int J Geriatr Psychiatry 2014 Apr;29(4):331-344. [doi: 10.1002/gps.4016] [Medline: 23963684]

XSL•FO
RenderX

46. Cintoli S, Tommasini LL, Del Prete E, Cerri M, Ceravolo R, Tognoni G. The psychoeducational interventions: a valuable communication tool to support the caregiver of people with dementia. BMC Geriatr 2024 Dec 19;24(1):1004. [doi: 10.1186/s12877-024-05562-5] [Medline: 39702001]
47. Wang J, Li Q, Cui J, et al. Effectiveness of virtual reality on the caregiving competence and empathy of caregivers for elderly with chronic diseases: a systematic review and meta-analysis. J Nurs Manag 2023;2023:5449955. [doi: 10.1155/2023/5449955] [Medline: 40225657]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
MVP: Minimally Viable Prototype
VR: virtual reality

Edited by J Sarvestan; submitted 04.Nov.2025; this is a non–peer-reviewed article;accepted 06.Jan.2026; published 05.Mar.2026.

Please cite as:
Chiu M, Pistritto SC, Kokorelias KM, Clarke KM, Torres A, Craven CG, Dubrowski A, Kapralos B, O'Hare E, Sadavoy J, Shnall A, Smith MSD, Choudhury N, Burhan AM
Supporting Unpaid Caregivers of Persons Living With Dementia: Protocol for a Pilot Feasibility Study to Explore Caregiver Outcomes and Impact of a Co-Designed Simulation–Based Psychoeducation Program in Virtual Reality
JMIR Res Protoc 2026;15:e87107
URL: https://www.researchprotocols.org/2026/1/e87107
doi:10.2196/87107

© Mary Chiu, Sarah C Pistritto, Kristina M Kokorelias, Krystina M Clarke, Andrei Torres, Cole G Craven, Adam Dubrowski, Bill Kapralos, Erica O'Hare, Joel Sadavoy, Adriana Shnall, Michael S D Smith, Nusrat Choudhury, Amer M Burhan. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Acceptance and Commitment to Empowerment Intervention to Reduce HIV Stigma and Promote Community Resilience: Protocol for an Implementation Study

Josephine Pui-Hing Wong[1], PhD; Alan Tai-Wai Li[2], MD; Carla T Hilario[3], PhD; Mandana Vahabi[4], PhD; Egbe Etowa[1], PhD; Isaac Luginaah[5], PhD; Aniela M dela Cruz[6], PhD; Miya Narushima[7], PhD; Kenneth Po-Lun Fung[8], MD, FRCP

[1]Daphne Rockwell School of Nursing, Toronto Metropolitan University, 350 Victoria St, Toronto, ON, Canada

[2]Primary Care, Maple Leaf Medical Clinic, Toronto, ON, Canada

[3]School of Nursing, University of British Columbia Okanagan, Kelowna, BC, Canada

[4]Lawrence Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada

[5]Department of Geography and Environment, Western University, London, ON, Canada

[6]Faculty of Nursing, University of Calgary, Calgary, AB, Canada

[7]Faculty of Applied Health Sciences, Brock University, St. Catharines, ON, Canada

[8]Department of Psychiatry, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Josephine Pui-Hing Wong, PhD
Daphne Rockwell School of Nursing, Toronto Metropolitan University, 350 Victoria St, Toronto, ON, Canada

Abstract

Background: Racialized immigrants are disproportionately impacted by HIV in Canada. In 2022, 69.5% of first-time HIV cases, where race and/or ethnicity was reported, were among racialized individuals: Indigenous (22.6%), Black (18.0%), Asian and Arab (16.3%), and Latinx (10.1%). HIV vulnerability in racialized communities is reinforced by HIV stigma intersecting with prejudice and discrimination associated with racism, gendered oppression, homophobia, and economic marginalization. Stigma and discrimination impede public health efforts in HIV prevention, testing, treatment, and care.

Objective: The goal of the study is to reduce HIV-related stigma in racialized immigrant communities in 6 Canadian cities in Alberta (Calgary, Edmonton) and Ontario (London, Niagara, Ottawa, Toronto). We will implement an evidence-based 6-session online intervention, Acceptance and Commitment to Empowerment (ACE), to promote skills for psychological flexibility (mindfulness, acceptance, cognitive defusion, values clarification, and committed action) and collective empowerment (critical reflection, dialogue, interconnectedness, and collective action). The objectives are to (1) assess community contexts and needs to inform the refinement of ACE; (2) build community capacity in stigma reduction; (3) evaluate the facilitators and barriers of implementing stigma reduction programs; and (4) disseminate knowledge to advance research, policy, and practice.

Methods: Guided by the principles of social justice and health equity, we will apply the context-based PRISM (Practical Robust Implementation and Sustainability Model) and outcome-based RE-AIM (reach, effectiveness, adoption, implementation, maintenance) frameworks to inform the contextual adaptation, implementation, and evaluation of the ACE intervention. In phase 1 (year 1), we aim to survey 12 local organizations on their existing stigma reduction programs; engage 30 service providers and 60 community members to explore the local contexts and needs of stigma reduction. In phase 2 (year 2), phase 1 results will be used to refine the ACE intervention and apply a train-the-trainer (TTT) approach to engage 48 service providers or community leaders in the ACE training. ACE TTT graduates will be mentored to become ACE Community Facilitators. In phase 3 (years 3 - 4), project staff and ACE Community Facilitators will work in small 3-member teams to deliver the online ACE intervention to qualified participants (n=288) in 6 cities. Mixed methods will be used to examine the contextual factors on implementation processes and intervention effectiveness. Data analyses will include reflexive thematic analysis of qualitative data, inferential statistics, and analyses of variance to determine the effectiveness of the intervention over time.

Results: The implementation of project ACE was delayed by 2 years due to the COVID-19 pandemic. The requirement of obtaining research ethics approvals from all affiliated universities (n=8) further delayed the launch of research activities involving participants until September 2023. Between October 2023 and August 2024, the project team recruited 8 community leaders from the 6 project sites to form the Project Advisory Committee (PAC), worked with PAC members to identify outreach and recruitment strategies for phase 1, and recruited and secured 19 community organizations as project collaborators for phase 2. As of March 2025, phase 1 data collection has been completed, and data analysis and knowledge translation are underway.

XSL•FO
RenderX

Conclusions: Based on results from previous research, we anticipate generating critical knowledge on the effectiveness of ACE in reducing internalized and enacted stigma through collective empowerment, and the facilitators/barriers in implementing ACE among racialized immigrants in real-world settings.

International Registered Report Identifier (IRRID): DERR1-10.2196/80669

(JMIR Res Protoc 2026;15:e80669) doi:10.2196/80669

KEYWORDS

acceptance commitment to empowerment; HIV stigma reduction; racialized communities; train-the-trainer approach; capacity building

Introduction

Racialized immigrants and refugees are disproportionately impacted by HIV in Canada. Although racialized people constituted 26.6% of the Canadian population as of the 2021 census [1], they continue to be disproportionately impacted by HIV. In 2022, 69.5% (n=519) of first-time HIV cases, where race and/or ethnicity was reported (n=766)**,** were racialized individuals: Indigenous (n=175, 22.6%); Black (n=140, 18%), Asian and Arab (n=126, 16.3%), and Latinx (n=78, 10.1%) [2]. Among Black people in Canada, the proportion of first-time HIV diagnoses across populations increased from 15.6% (n=288) in 2021 to 18.0% (n=140) in 2022, even though they made up 4.3% of the Canadian population [3]. HIV statistics by geographical region show that Ontario maintains the highest number of cases of new HIV diagnoses in Canada. In 2022, for example, Ontario reported 49.1% (n=723) of first-time HIV diagnoses in Canada [3]. Following an increase in the nationwide diagnosis rate in 2023, there was a dramatic increase in new diagnoses in Alberta from 15.6% (n=190) in 2022 to 20.8% (n=255) in 2023, making it the largest increase since HIV was first reported in Alberta in 1998 [3].

Elevated HIV vulnerability for racialized minority groups, including individuals with sexual and gender minority identities, immigrants and refugees, and people without immigrant status, is associated with complex factors, including social isolation [4]; migration/settlement stress [5,6]; systemic barriers to accessing resources [7-9]; and intersecting marginalization related to transphobia/homophobia [10,11], racism [7,12,13], gender-based inequities [4,14], and poverty [15,16]. In addition, HIV stigma and discrimination reinforce HIV vulnerability by creating unsafe environments that deter people from testing and disclosure, resulting in isolation, depression, delayed diagnosis and linkage to treatment and care, and poor health outcomes [17,18]. Stigma also leads to the invisibility of people living with HIV, reinforces community denial, undermines HIV prevention efforts, and impedes provision of accessible community care [19]. However, knowledge gaps in effective and sustainable stigma reduction strategies and programs continue to exist [20]. Further, implementation research that examines effective HIV stigma reduction interventions in real-world settings is also scarce.

The Acceptance and Commitment to Empowerment (ACE) intervention is built on existing knowledge and evidence generated from recent research completed by our team [21,22]. HIV stigma is a priority issue that negatively affects people living with HIV or vulnerable to HIV, especially in terms of accessing HIV and mental health care [19,22,23]. Between 2011 and 2015, our team worked closely with Black, Latinx, and Asian communities in Toronto to design and pilot the Community Champions HIV/AIDS Advocates Mobilization Project (CHAMP), funded by the Canadian Institutes of Health Research. The CHAMP study engaged community leaders both living and not living with HIV from faith-based, media, and social justice sectors to address HIV stigma and promote social justice. We evaluated the effectiveness of 2 evidence-based trainings in reducing stigma at the individual and the collective level: (1) acceptance and commitment therapy (ACT) [24,25], an intervention that promotes psychological flexibility through mindfulness-based exercises and experiential activities that are underpinned by 6 core processes: diffusion (observing thoughts as thoughts), acceptance (of experiences of emotions and feelings), contact with the present moment (mindfulness), self-as-context (awareness and self-perspective), values (being clear about what matters), and committed action (based on values) [26]; and (2) social justice capacity building (SJCB) [27,28], an empowerment education that promotes a critical understanding of stigma in the context of power relations in society, recognition of interdependence and interconnectedness as important determinants of health equity, and the importance of tapping into lived experiences and community strengths to take action to reduce stigma. Given their effectiveness, our team integrated ACT and SJCB into a comprehensive model and intervention—ACE, as shown in Figure 1. ACE has been scaled up with service providers and university students in a Global Alliance for Chronic Diseases partnership project in Jinan, China, to reduce mental illness stigma and promote mental health literacy [29,30]. Existing research shows that while online interventions are more costly to develop, well-designed and deployed online interventions can be cost-effective to implement and sustain [31,32].

evieweThe overall goal of this study is to reduce HIV-related stigma in racialized immigrant and refugee communities in 6 Canadian cities in Alberta and Ontario. The objectives are to assess community contexts and needs to inform the refinement of the ACE intervention; build community capacity in stigma reduction using a TTT approach; evaluate the facilitators and barriers of implementing stigma reduction programs; and disseminate knowledge to advance research, policy, and practice. Hence, the overarching research questions are as follows: (1) What are the contextual facilitators and barriers in implementing the evidence-based online ACE intervention in real-world settings? (2) How effective is the online ACE intervention in reducing HIV-related stigma and promoting collective

XSL•FO
RenderX

empowerment? Specifically, we will (1) examine the effectiveness of the online training in reducing HIV stigma and promoting collective action; (2) identify the contextual factors contributing to the effective implementation of the online training; (3) determine participants' characteristics contributing to acceptability and feasibility of the online training; and (4) assess organizational characteristics determining the acceptability and feasibility of the online training.

Figure 1. Acceptance and Commitment to Empowerment intervention presented as an integrated model of the acceptance and commitment therapy [24] and the social justice capacity building empowerment model [28].

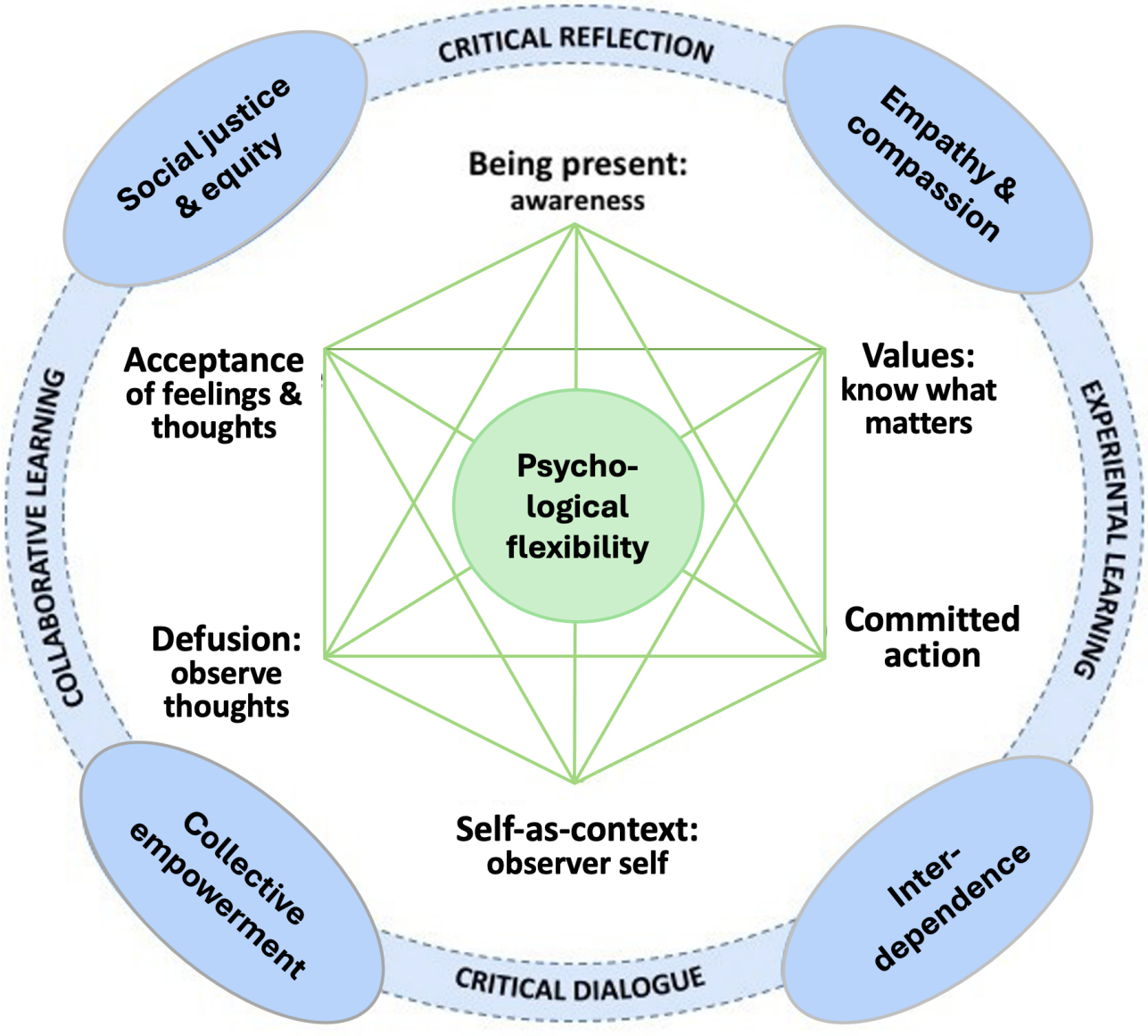

Methods

Conceptual Frameworks

The proposed study is underpinned by the principles of equity, access, and social justice. It is guided by a decolonized critical health promotion framework [33], which centers voices of affected communities and applies strategies such as meaningful engagement [34], shared leadership and responsibilities [35], and capacity building [36] to promote collective empowerment throughout all phases of the research. In addition, we will draw on the PRISM (Practical Robust Implementation and Sustainability Model) and RE-AIM (reach, effectiveness, adoption, implementation, maintenance), well-tested integrated frameworks, to guide the implementation and evaluation of this project. We will apply the RE-AIM framework [37,38] to operationalize the proposed study. The RE-AIM categories will enable us to use consistent methods of evaluation, data collection, analysis, and reporting of results. In addition, the PRISM [39] will enable us to identify and document specific contextual factors that affect implementation at the macro level (eg, public policies, resource distributions, intersecting marginalization, media discourses), meso level (eg, community networks, organizational capacity, practice culture, service user engagement), and micro level (eg, personal values and beliefs, lived experiences). Insights and knowledge of the contextual factors will enable us to identify infrastructures needed for effective implementation and sustainability (eg, policy change, professional development, practice competence,

XSL•FO
RenderX

audit-and-feedback loops, and integrated practice, and resources distribution).

Project ACE will be implemented in 6 project sites: Calgary and Edmonton in Alberta, and London, Ottawa, Niagara, and the Greater Toronto Area in Ontario. Guided by PRISM and RE-AIM, research activities of community engagement, implementation, evaluation, and knowledge translation and exchange (KTE) will be carried out in 3 phases, as presented in Figure 2.

Figure 2. Implementation phases of the Acceptance and Commitment to Empowerment (ACE) intervention.

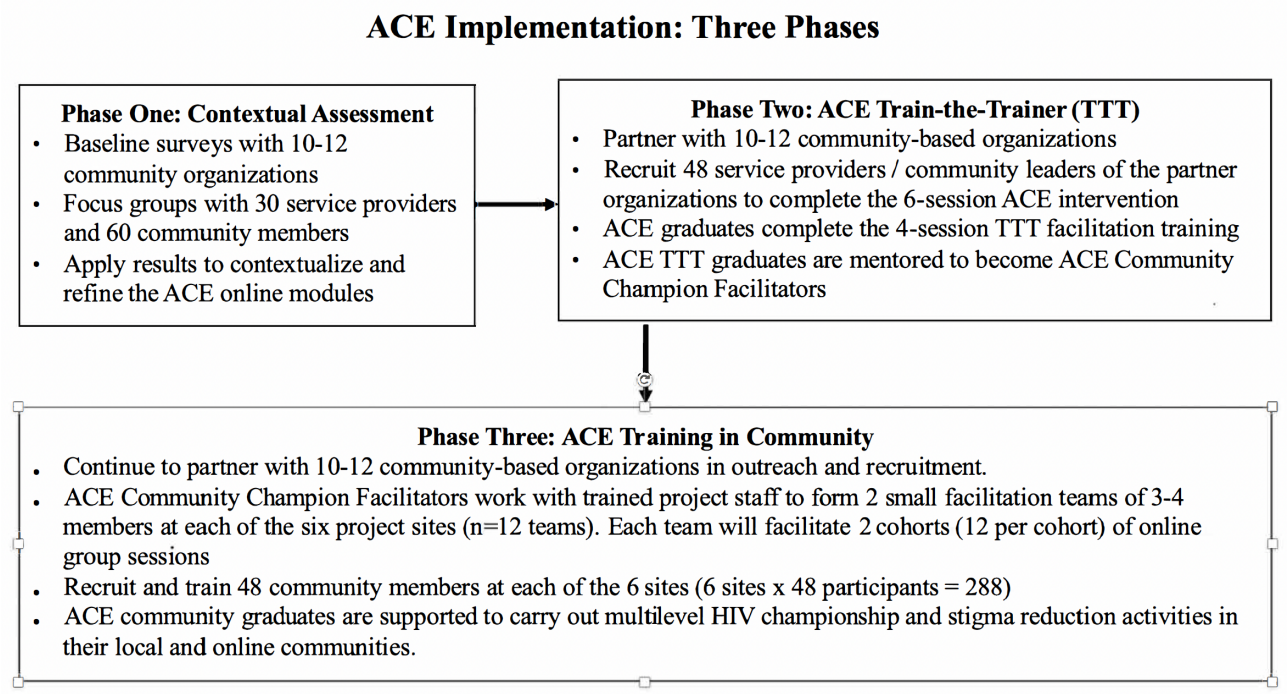

Ethical Considerations

The project team has received research ethics approvals from all institutions affiliated with this study. Specifically, the study protocols of the current implementation project have been approved by the research ethics boards of all participating institutions (University of Alberta Pro00127653; University of British Columbia H23-00246; Brock University 22‐122, University of Calgary Pro00127653/pSite-23‐0024; University of Ottawa S-01-23-8683; University of Toronto 43739; Toronto Metropolitan University 2022‐267; Western University 122170; York University e2023-050).

All participants will provide written informed consent before participating in the study. Informed consent will be obtained from all participants before data collection. At each of the research activities, participants will be informed and reminded of the voluntary nature of the research (ie, they can withdraw from the study at any time), and their withdrawals will not affect their relationships with any institutions or service organizations affiliated with this study. The informed consent includes consent for the recording and publication of anonymized information shared during individual participations. Participants will recieve a $30 honorarium as a token of appreciation for their participation in the evaluation survey (i.e., each time they start filling out the pre-survey, post-survey, or 3-month follow-up survey), They will also receive $20 for filling out each of the six online self-directed learning modules, and $10 for each of the three activity logs after the training. Participants will also receive $40 if they are randomly selected and agree to take part in the 3-month post-training focus group.

All research data will be anonymized to protect the privacy or confidentiality of participants. Survey results will be reported in aggregated form without any individual-level identifiers. For participants taking part in the focus groups or group discussions, they may be concerned about their privacy. There may be a risk that their sharing within the focus groups may be shared outside of the group, thus breaching their privacy. The risk will be minimized by (1) informing participants that in a group context, absolute confidentiality cannot be guaranteed and that participants are encouraged to decide what they wish to share and what they do not wish to share; (2) we have included a statement of confidentiality agreement in the participant consent form; and (3) we have also established a set of guiding principles that will be reviewed before all focus groups or group discussions.

Three Phases of Project ACE

The following section describes the 3 phases of Project ACE. We use a sequential approach in which each phase is informed by the results of the preceding phase.

Phase 1 (Year 1): Contextual Assessment

Formation of Teams and Project Advisory Committees

In the first 6 months, the team will establish a Project Steering Committee consisting of principal researchers who developed the evidence-based ACE intervention; Regional Research Teams consisting of the coprincipal researchers and coresearchers; and a Project Advisory Committee made up of members including people living with HIV, service providers, community leaders, and collaborators. Project staff will include both people living

XSL•FO
RenderX

and not living with HIV drawn from Alberta and Ontario sites. The Project Steering Committee will work with Regional Research Teams to obtain research ethics approval from affiliated universities and organizations.

Local Contextual Assessment Specific to HIV Stigma Reduction

Group 1: Participating and Collaborating Organizations

We will conduct a 3-month baseline survey with participating organizations to explore the types of services or programs they have offered to address HIV-related stigma and the quantities of services used by their clients and community members (see Multimedia Appendix 1). These data will be used as the baseline to measure reach and effectiveness. Also, we will engage 10 to 12 decision-makers and knowledge users drawn from collaborating organizations every 6 to 9 months in a session of Collaborator Dialogue to share project updates and their perspectives on adopting the ACE intervention to reduce HIV-related stigma, including perceived level of complexity about ACE; organizational values and norms towards new programs or practices; and their perception of potential benefits and challenges in adopting ACE into their organizational practice. Our research team includes 8 community-based organizations as formal collaborators, who will support recruitment of other relevant organizations.

Group 2: Service Providers/Community Leaders

We will recruit and engage 5 participants in each of the 6 project sites (n=30) in focus groups to explore their perspectives on the needs of HIV stigma reduction in their organizations and communities, challenges and facilitators of stigma reduction, available resources, and critical aspects of stigma reduction interventions (see Multimedia Appendix 2). Participation criteria will include being aged 18 and older, self-identified as a service provider or community leader working with racially minoritized communities affected by HIV-related stigma in one of the 6 project sites. Prior to the focus groups, participants will be asked to complete a short sociodemographic survey (see Multimedia Appendix 3) to identify the project site they work in, age, sex, gender, ethnicity, first language, religion, and level of community engagement in HIV stigma reduction.

Group 3: People Living With HIV And Community Members Affected by HIV-Related Stigma

We will recruit and engage 10 participants in each of the 6 project sites (n=60) in 6 gender-specific and mixed-gender focus groups to explore their experiences and perspectives on HIV stigma, their resistance and resilience strategies, and access to relevant support (see Multimedia Appendix 2). Participation criteria will include being aged 18 years and older, self-identifying as belonging to a racially minoritized community, living with and/or affected by HIV-related stigma, and living in one of the 6 project sites. We will pay special attention to recruiting members of diverse sexual and gender identities. Prior to the focus groups, participants will also be asked to complete a short sociodemographic survey (see Multimedia Appendix 3) to identify: the project site they reside in, age, sex, gender, ethnicity, first language, religion, and level of community engagement in HIV stigma reduction.

Groups 2 and 3 Recruitment

Participants will be recruited with support from our community collaborating agencies, using common methods of outreach (ie, poster, e-flyers, e-distributions, and social media). Collaborators will be given an electronic copy of the recruitment flyer, which can be printed and posted in their organizations' public bulletin boards. The e-flyer can be distributed through Twitter, Facebook, collaborators' e-newsletter, and the project website. This method of distribution has been chosen to ensure that collaborators and partners help to distribute recruitment materials, but they will not do direct recruitment of participants to eliminate any potential conflicts of interest. Direct recruitment will be carried out by project staff. The use of social media will strictly be limited to distribution of our e-flyers; there will be no function for external users to post comments or engage in public communication or interactions. Potential participants will be directed to contact the project coordinator directly, as indicated in the e-flyer.

Refinement of ACE Intervention

Data on the local contextual assessment will be analyzed, and the results will be used to refine the ACE intervention to best suit the contexts and needs of the participant populations. We will share results from the context assessments with the Local Advisory Committees and our collaborators and engage them in the refinement and adaptation of ACE.

Phase 2 (Year 2): ACE TTT

Overview

We will engage and build capacity among existing service providers and community leaders. We will recruit 8 service providers and community leaders at each of the 6 project sites (n=48) in the HIV and related sectors (eg, mental health, addiction, sexual health, settlement services). We will apply a TTT capacity-building approach to implement and evaluate the effectiveness of the ACE intervention in reducing stigma and motivating the ACE graduates to become cofacilitators to implement ACE in the community in phase 3. Phase 2 participation criteria include being aged 18 years and older, self-identifying as a service provider or community leader working with racially minoritized communities affected by HIV-related stigma in one of the 6 project sites, being available and committed to complete the ACE intervention and become an ACE Champion to co-facilitate the ACE intervention in their local communities, and having fluency in written and spoken English. Phase 2 learning activities include the following: pretraining self-directed learning ("Stigma Reduction 101"), the ACE intervention, TTT practice workshops, and posttraining support.

Pretraining Self-Directed Learning ("Stigma Reduction 101")

Participants will complete an online preintervention survey and one 30-minute online module called HIV Stigma 101, hosted on a Learning Management System. The HIV Stigma 101 will be developed by the team, drawing on the results of our Contextual Analysis in phase 1, as well as culturally safe and up-to-date HIV information on the HIV care cascade.

XSL•FO
RenderX

The ACE Intervention

A 6-week online experiential learning activities to (1) promote participants' psychological flexibility; (2) reduce HIV-related stigma; (3) promote resilience/effective coping; and (4) increase individual and collective empowerment that enhance community championship. The online training consists of 6 weekly 60-minute self-directed interactive online learning modules, which will be reinforced by a 1.5-hour weekly videoconferencing session. Each module will cover one of the 6 ACT processes along with their corresponding SJCB principles and processes. Our proposed online ACE intervention is innovative in that, unlike many online psychological interventions that offer individual self-directed learning [40,41], it includes group collaborative learning through videoconferencing to promote engagement and collective empowerment.

TTT Practice Workshops

The ACE TTT graduates will receive the ACE Intervention TTT Manual and access to training videos. They will also take part in four online 2.5-hour TTT workshops to learn about: outreach and recruitment; data collection; the ACE concepts in detail; online facilitation and debriefing of the 6 online sessions; and practice facilitation with peer and mentor feedback.

Posttraining Support

We will develop a community of practice utilizing the Learning Management System to engage the ACE TTT graduates and support them in taking on the role as ACE Community Facilitators. They will be engaged in asynchronous and scheduled online structured peer-sharing and periodic mentorship sessions during the 6-month practicum period to enact phase 3 (ie, ACE facilitation in the community). They will be encouraged to engage in regular reflections post-training on the impact of ACE on their personal, professional, and community lives and submit an online activity log of their championship activities.

Phase 3 (Years 2-4): ACE Community Champions in Action

The ACE Community Facilitators will be "mobilized" to implement HIV-related stigma reduction activities, including the following: recruitment, delivery of the ACE intervention, and stigma reduction championship.

Recruitment

The ACE Community Facilitators at each site (n=8) will be divided into 2 teams of 3 to 4 members (accounting for job change and attrition). With the support of their organizations, each team will work with Project ACE staff to recruit 24 service users and/or community members as participants (6 sites × 2 teams × 24 persons; n=288). Since Project ACE is an implementation study that focuses on identifying the facilitators and barriers to implementing evidence-based practice in real-world settings, we will not use any control groups. We will pay special attention to recruiting members of diverse sexual and gender identities.

Phase 3 participation criteria for the community ACE intervention includes: aged 18 years and older, living in one of the 6 project sites, self-identifying as belonging to a racially minoritized community and a priority group affected by HIV stigma and/or identified by the partner organizations (eg, LGBTQ+ [lesbian, gay, bisexual, transgender, queer, and other gender and sexual minorities], newcomers, youth, women), is available and committed to attending all data collection and training sessions and interested in becoming an HIV community champion to reduce HIV-related stigma (eg, substance use, racism, transphobia) in the community.

Delivery of the ACE Intervention

Each ACE Community Facilitator team will deliver 2 cohorts of online ACE facilitation. The aim of the ACE intervention is to reduce HIV-related stigma and promote resilience among the participants. Graduates of the community ACE intervention training will become ACE community champions. In phase 3, we will use the same online ACE intervention used in phase 2, consisting of 6 weekly modules of 60-minute self-directed interactive online learning activities reinforced by a 1.5-hour weekly videoconferencing session. Our proposed online ACE intervention is innovative in that, unlike many online psychological interventions that offer individual self-directed learning only [40,41], it includes group collaborative learning through videoconferencing to promote peer support, community engagement, and collective empowerment.

Stigma Reduction Championship

The ACE Intervention Community Facilitators will continue to engage in the community of practice to consolidate their ACE knowledge and skills and contribute to community-wide stigma reduction by connecting the newly trained ACE graduates (n=288) to local organizations, groups, or initiatives, and providing them with support in carrying out HIV-related stigma reduction activities (eg, peer workshops, media interviews, newsletters). ACE graduates will be invited to submit 3 monthly community champion activity logs that capture these activities.

Evaluation of ACE Intervention (Years 2-4)

We will use mixed methods to collect data, including surveys with sociodemographic questions (see Multimedia Appendix 3) and validated scales (listed below), focus groups, dialogical sessions, document analysis, field observation, and tracking of activities of interest. We will use PRISM [39] to contextualize and RE-AIM [37,38] outcomes in evaluating the implementation of the ACE intervention.

Reach

We will measure the absolute number, proportion, and representativeness of the intended population willing to, or taking part in the ACE intervention. The reach of the ACE intervention will be the proportion of service providers, community leaders, service users, and community members recruited and successfully trained compared to the estimated corresponding eligible populations. As we have set a predetermined recruitment target (8 TTT participants per site; 48 community participants per site) and composition (eg, service providers/community leaders, people affected by HIV stigma), we will keep track of those who expressed interest to be trained beyond the target as an estimate of "maximal reach." The total number of eligible participants will be obtained based on

XSL•FO
RenderX

information from the local advisory committees. We will keep track of noncompleters and examine for any significant differences in demographics and other baseline variables between completers and noncompleters, as well as between completers and known data of the populations in each site.

Effectiveness

This measures the impact of the ACE intervention on our intended outcomes (ie, stigma reduction and collective empowerment). We will evaluate the effectiveness of the ACE intervention with corresponding validated measures at preintervention, immediate postintervention, and 3-month postintervention. In addition to a short sociodemographic survey on the project site participants work in or reside in, age, sex, gender, ethnicity, first language, religion, and level of community engagement in HIV stigma reduction, relevant validated scales to be used include (1) the 12-item HIV Stigma Scale [42], which include four subscales on personalized stigma, disclosure concerns, concerns with public attitudes and negative self-image, (2) the 4-item Perceived Discrimination from Service Providers Scale [43] with dichotomized score (yes/no) to measure perceived HIV-related discrimination from health care providers, (3) the 12-item Cognitive and Affective Mindfulness Scale (CAMS-R) [44,45] that assesses mindfulness using language that is not connected to any specific style of meditation practices, (4) the 7-item Acceptance and Action Questionnaire Version II (AAQ-II) [46] that measures experiential psychological inflexibility and cognitive fusion, (5) the 28-item Brief Coping Orientation to Problems Experienced (Brief COPE) [47] that measures 14 different coping strategies (self-distraction, active coping, denial, substance use, use of emotional support, use of instrumental support, behavioral disengagement, venting, positive reframing, planning, humor, acceptance, religion, and self-blame); and (6) the 12-item Self-Compassion Scale Short (SCS-S) [48] that measures self-kindness, self-judgment, common humanity, isolation, mindfulness, and over-identification. Focus groups (see Multimedia Appendix 4) will also be conducted at 3-month post-training to explore (1) participants' experiences of taking part in the ACE intervention; (2) how participants have applied what they learned in ACE in their everyday life to address stigma; and (3) how ACE can be improved.

Adoption

This component measures the absolute number, proportion, and representativeness of organizations willing to participate in this project. The adoption of the ACE intervention, or the support of organizations/settings to implement ACE, will be assessed in terms of perceived feasibility, acceptability, and fit based on results from the local Contextual Assessment (phase 1). We will also track expressed interest and diffusion attributes by stakeholders in the semiannual dialogical sessions, contextual factors that facilitate or impede adoption/adaptation, and the potential and actual uptake of the ACE intervention by organizations within or outside of the HIV/sexual health sector in each site, and in other cities in Alberta and Ontario.

Implementation

We will assess the extent to which the intervention is delivered as intended. To achieve this, fidelity checklists will be used to evaluate the implementation of the ACE intervention by the staff facilitators and the trained ACE Intervention Community Facilitators. Participants will complete a postsession feedback form at each training session to assess their self-report of satisfaction, knowledge gained, and confidence in specific skills. ACE facilitators will complete a debriefing form after each training session to document implementation issues and group dynamics. Finally, we will conduct a cost analysis on community-engaged implementation [49] of the ACE intervention.

Maintenance

We will measure maintenance and sustainability of the ACE intervention at both the individual and organizational levels over the period of follow-up, especially with a focus on activities that can be sustained beyond the project, and diffusion networks (ie, potential uptake by organizations and interest holders not directly involved in the study). This includes activities by ACE Champions and community participants captured in monthly activity logs, the online responses in the Learning Management System, and community requests for ACE intervention. In addition, an online qualitative survey will be conducted 3 months postintervention with participants and interest holders to evaluate: the impact of ACE at the personal, professional, organizational, and community levels, acceptability of ACE, desire for adoption, appropriateness, and feasibility.

Data Analysis (Years 1-4)

Overview

In addition to using the RE-AIM Framework, we will use a multiple-case study design to examine the contextual factors that support or hinder effective implementation of ACE in the six project sites. The case study approach is suitable for implementation research because it allows for an in-depth investigation of real-life phenomena within contexts (eg, an individual, a group, an organization, a community, an event, or a problem) [50,51]. In case study research, the contextual conditions are not controlled but considered part of the investigation. The multiple case design allows us to perform within-case (eg, online ACE intervention, Black women, or the Calgary site) as well as across-case analyses (eg, comparison of effectiveness of site-specific intervention groups, implementation facilitators, and barriers across all six sites, or across different ethno-cultural communities). We will apply triangulation as a data collection and analytic strategy to achieve detailed case descriptions and to gain a critical understanding of "how" and "why" things happen [52,53]. We will also apply cross-case analysis to identify similarities and differences in the factors, conditions, and contexts that support or hinder effective implementation of the ACE intervention in the 6 sites.

Quantitative Data Analysis

Descriptive and group-based analyses will be conducted to gain insights into the sampled population. Inferential statistics, including analyses of variance, will be conducted to determine the effectiveness of the interventions over time in terms of their

XSL•FO
RenderX

contribution to decreasing stigma (stigma scales); increasing psychological flexibility (AAQ-II); increasing mindfulness (CAMS-R); increasing self-compassion (SCS-S); and decreasing negative coping (Brief Cope). Regression, path analyses, and hierarchical structural equation modeling will be conducted to examine the potential mediating and moderating impact of other variables, such as background demographics, level of participation, roles in community (service provider, leader, service user), and involvement in HIV-AIDS championship activities, on project and intervention outcomes.

Qualitative Data Analysis

Focus group interviews (see Multimedia Appendix 4) will be transcribed verbatim. A computer software program (N-Vivo) will be used to aid in data management. For data analysis, we will use both inductive and deductive approaches to manually look for broad categories that were *voice-centered* articulated by the participants (inductive) and *sensitizing* drawn from pre-existing theories and the research questions (deductive) in the transcripts [54]. Guided by a decolonized critical health promotion framework [33] and using an iterative and systematic approach [55-57], we will (1) familiarize ourselves with the data; (2) review our focus group discussion guide and the PRISM/RE-AIM framework; (3) generate codes as the smallest units of analysis responsive to the research questions [58]; (4) develop a coding tree with subcodes to facilitate coding for analysis and interpretation; (5) make sense of and use codes as building blocks to generate themes as central organizing concepts [57,59]; (6) engage team members in reflexive analysis and interpretation; and (7) produce reports and relevant knowledge products based on the results. These iterative team processes will ensure that collective reflexive insights are integrated into the data analysis and consensus is reached in understanding the data.

Results

Although Project ACE was funded in October 2020, project implementation was delayed by almost 2 years due to the impact of the COVID-19 pandemic. In June 2022, research preparation activities were launched; they included the establishment of a Project Steering Committee, hiring and training of project staff, and the development of the online training platform. However, the requirement of receiving research ethics approvals from all affiliated universities (n=8) further delayed the launch of research activities involving participants until September 2023. Between October 2023 and August 2024, the project team recruited 8 community leaders from the 6 project sites to form the Project Advisory Committee (PAC), worked with PAC members to identify outreach and recruitment strategies for phase 1, recruited and secured 19 community organizations as project collaborators for phase 2. As of March 2025, 18 organizations have completed the baseline survey, and 32 service providers and 53 community members have been successfully recruited in the phase 1 contextual assessment. Data analysis is underway. A community infographic report on phase 1 is expected to be published in February 2026, and 3 peer-reviewed manuscripts are expected to be published by August 2026.

Discussion

Drawing on the results of our previous research, we anticipate that Project ACE will contribute critical new knowledge to inform stigma reduction and effective implementation. Although HIV stigma occurs in multiple levels (intrapersonal, interpersonal, organizational, community, and societal), most stigma reduction interventions only address stigma at the personal or interpersonal levels [60]. Building on a decade of stigma reduction research [19,22], our team has developed the ACE intervention, which addresses stigma through skills building to increase psychological flexibility and readiness for value-guided committed action at multiple levels [61].

In this project, we are drawing on a decolonized critical health promotion framework [33] and the well-established PRISM/RE-AIM implementation evaluation framework [37-39]. The integrated use of these frameworks places us in a strong position to identify, at multiple levels, factors that perpetuate HIV stigma and the contextual facilitators and barriers in implementing antistigma intervention in real-world settings. For example, in our previous research, the application of a critical health promotion framework enabled us to challenge the popular perception of stigma as individual behaviors. We documented how structural and system determinants, such as homophobia, racism, gender-based discrimination, and unequal power relations, interacted to produce and perpetuate stigma and misconceptions that deterred community leaders from engaging in collective action to reduce stigma [19].

Our innovative pre-post-follow-up mixed methods design also positions our team to generate important knowledge. Data collection will include sociodemographic surveys, validated scales, focus groups, and meeting notes. In addition, all ACE participants will be invited to complete regular activity logs for 3 months postintervention. The use of activity logs will enable us to measure not only changes in participants' attitude, emotional regulation, degree of coping, and sense of readiness to engage in collective action on HIV stigma, but we will be able to document their actual behavioral changes at personal, interpersonal, organizational, and community levels. For example, we used activity logs for data collection in our CHAMP study, in which we recruited and engaged 104 participants, and 62 completed a similar intervention. These 62 graduates carried out 1090 multilevel stigma reduction and equity-based activities over 9 months postintervention. The champion activities included providing peer support, volunteering at an HIV/AIDS organization, leading health literacy workshops, establishing new social groups, and confronting homophobia at a faith-based organization [62].

Similarly, the Linking Hearts project engaged 146 multidisciplinary professionals at 6 universities in China to promote student mental health and reduce stigma. A total of 124 participants completed the in-person ACE-LYNX-PRO (Acceptance and Commitment to Empowerment–Linking Youth and "Xin" (hearts) with Professionals) intervention, online pre-post-follow-up surveys with validated scales, 3-month postintervention focus groups, and weekly online activity logs for 3 months postintervention, yielding 1083 entries. Survey

XSL•FO
RenderX

results showed statistically significant psychological empowerment measured through reduction in stress and in stigmatizing attitude. Analysis of the activity logs and focus groups showed committed action at multiple levels: (1) individual-level activities included counseling individuals, self and family care, building personal skills; (2) group-level activities included structured ACE-based group counseling; facilitating peer dialogue on mental health; engaging in collaborative learning; (3) organizational-level activities included initiating psychoeducation and mental health advocacy; integrating ACE into mental health education curricula; coordinating mental health awareness campaigns; establishing book clubs; and (4) community-level activities included ACE-based experiential workshops and mental health literacy promotion among neighborhood residents [61].

Based on previous research results, it is anticipated that participants in Project ACE will experience psychological empowerment and demonstrate committed action in HIV championship and stigma reduction at multiple levels. At the same time, we also anticipate some potential limitations. Since this is an implementation study that demands a lengthy commitment from participants, the study withdrawal rates may be high. We are addressing this potential challenge using intentional over-recruitment. In addition, the extent of successful implementation of the ACE intervention is dependent on many social, political, and economic contextual factors at the organizational, community, and societal levels that are beyond the control of our team. However, as an implementation study, our meticulous systematic documentation and rigorous analysis will enable us to generate knowledge that is informative and pragmatic.

The research team is strongly committed to creating knowledge and putting knowledge into action. An integrated knowledge translation approach will be applied to engage community members, service providers, and decision makers throughout all phases of the study. This will include (1) regular updates and Knowledge Translation involving posting and circulation of three newsletters per year to disseminate research updates and to promote community engagement. Also, semi-annual dialogical sessions will be held every 6 to 9 months to promote evidence uptake and ACE integration into their programs and services; (2) participant-driven KTE through collaboration with our ACE champions to deliver a variety of KTE activities to reduce stigma within their communities (eg, media interviews, presentations at faith-based events, digital stories, graphic novels, educational forums); and (3) end-of-grant KTE through collaboration with people living with HIV, ACE champions, collaborators, service providers, community leaders, health policy/decision-makers, and researchers. The collaborative team will produce policy briefs, media reports, infographic factsheets, community reports, digital stories, graphic novels, peer-reviewed conference presentations, and peer-reviewed publications. Finally, a KTE summit will be held to bring together different interest-holders to dialogue and advance policy, practice, and action.

Acknowledgments

The authors attest that there was no use of generative artificial intelligence in the generation of texts, figures, or any other informational contents of this manuscript.

Funding

This implementation research is funded by the Canadian Institutes of Health Research (CIHR, PJT - 173362). The funding source had no role in the design of this study and will not have any role during the implementation, data analysis and interpretation, and decision regarding the dissemination of findings.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to consent and agreement with participants but are available from the corresponding author on reasonable request.

Authors' Contributions

JPHW, ATWL, and KPLF created the ACE Model and Intervention. JPHW (nominated principal investigator) and ATWL, AMdC, CTH, KPLF, IL, MN, MV (co-principal investigators) work closely to implement the Acceptance and Commitment to Empowerment (ACE) protocol across the 6 project sites. EE is a postdoctoral fellow. JPHW led the drafting of this protocol manuscript with support from EE; all authors have reviewed, edited, and approved the final submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Phase 1 contextual assessment: collaborator baseline survey.
[PDF File, 222 KB - resprot_v15i1e80669_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
Phase 1 Acceptance and Commitment to Empowerment (ACE) contextual assessment: focus group discussion guides.
[PDF File, 208 KB - resprot_v15i1e80669_app2.pdf]

Multimedia Appendix 3
Participant sociodemographic survey (all phases).
[PDF File, 178 KB - resprot_v15i1e80669_app3.pdf]

Multimedia Appendix 4
Phases 2 and 3 postintervention focus group discussion guide.
[PDF File, 262 KB - resprot_v15i1e80669_app4.pdf]

Peer Review Report 1
Peer review report by Psychosocial, Sociocultural & Behavioral Determinants of Health Review Committee, Canadian Institutes for Health Research (CIHR).
[PDF File, 40 KB - resprot_v15i1e80669_app5.pdf]

References

1. Hou F, Schimmele C, Stick M. Changing demographics of racialized people in Canada. : Statistics Canada; 2023 URL: https://www150.statcan.gc.ca/n1/en/pub/36-28-0001/2023008/article/00001-eng.pdf?st=EaBX_Lk7 [accessed 2026-01-06]
2. HIV in Canada, surveillance report. : Public Health Agency of Canada; 2022 URL: https://www.canada.ca/content/dam/phac-aspc/documents/services/publications/diseases-conditions/hiv-canada-surveillance-report-december-31-2022/hiv-in-canada-surveillance-report-to-december-31-2022-en.pdf [accessed 2026-01-06]
3. HIV in Canada, surveillance report. : Public Health Agency of Canada; 2023 URL: https://publications.gc.ca/collections/collection_2025/aspc-phac/HP37-51-2023-eng.pdf [accessed 2026-01-06]
4. Logie CH, Lacombe-Duncan A, Lee-Foon N, Ryan S, Ramsay H. "It's for us -newcomers, LGBTQ persons, and HIV-positive persons. You feel free to be": a qualitative study exploring social support group participation among African and Caribbean lesbian, gay, bisexual and transgender newcomers and refugees in Toronto, Canada. BMC Int Health Hum Rights 2016 Jul 2;16(1):18. [doi: 10.1186/s12914-016-0092-0] [Medline: 27369374]
5. Wanigaratne S, Rashid M, Gagnon A, et al. Refugee mothers, migration pathways and HIV: a population-based cohort study. AIDS Care 2020 Jan;32(1):30-36. [doi: 10.1080/09540121.2019.1612009] [Medline: 31060379]
6. Baidoobonso S, Bauer GR, Speechley KN, Lawson E, BLACCH Study Team. Social and proximate determinants of the frequency of condom use among African, Caribbean, and other Black people in a Canadian city: results from the BLACCH study. J Immigr Minor Health 2016 Feb;18(1):67-85. [doi: 10.1007/s10903-014-9984-z] [Medline: 24488693]
7. Husbands W, Oakes W, Mbulaheni T, Ongoïba F, Pierre-Pierre V, Luyombya H. Resourceful masculinities: exploring heterosexual Black men's vulnerability to HIV in Ontario, Canada. Ethn Health 2020 Jan;25(1):17-33. [doi: 10.1080/13557858.2017.1395817] [Medline: 29082777]
8. Zhou YR, Majumdar B, Vattikonda N. Culture, but more than culture: an exploratory study of the HIV vulnerability of Indian immigrants in Canada. Cult Health Sex 2016 Sep;18(9):1067-1080. [doi: 10.1080/13691058.2016.1162328] [Medline: 27070278]
9. Blondell SJ, Kitter B, Griffin MP, Durham J. Barriers and facilitators to HIV testing in migrants in high-income countries: a systematic review. AIDS Behav 2015 Nov;19(11):2012-2024. [doi: 10.1007/s10461-015-1095-x] [Medline: 26025193]
10. Batchelder AW, Ehlinger PP, Boroughs MS, et al. Psychological and behavioral moderators of the relationship between trauma severity and HIV transmission risk behavior among MSM with a history of childhood sexual abuse. J Behav Med 2017 Oct;40(5):794-802. [doi: 10.1007/s10865-017-9848-9] [Medline: 28396969]
11. Nelson LE, Wilton L, Zhang N, et al. Childhood exposure to religions with high prevalence of members who discourage homosexuality is associated with adult HIV risk behaviors and HIV infection in Black men who have sex with men. Am J Mens Health 2017 Sep;11(5):1309-1321. [doi: 10.1177/1557988315626264] [Medline: 26758708]
12. Husbands W, Miller D, McCready LT, et al. Sexuality and sexual agency among heterosexual Black men in Toronto: tradition, contradiction, and emergent possibilities in the context of HIV and health. Can J Soc 2019 Jan 1;44(4):399-424. [doi: 10.29173/cjs29506]
13. Martin-Storey A, Benner A. Externalizing behaviors exacerbate the link between discrimination and adolescent health risk behaviors. J Youth Adolesc 2019 Sep;48(9):1724-1735. [doi: 10.1007/s10964-019-01020-3] [Medline: 31175511]
14. Kteily-Hawa RN, Islam S, Loutfy M. Immigration as a crisis tendency for HIV vulnerability among racialised women living with HIV in Ontario, Canada: an anti-oppressive lens. Cult Health Sex 2019 Feb;21(2):121-133. [doi: 10.1080/13691058.2018.1453087] [Medline: 29658837]

XSL•FO
RenderX

15. Shokoohi M, Bauer GR, Kaida A, et al. Social determinants of health and self-rated health status: a comparison between women with HIV and women without HIV from the general population in Canada. PLoS One 2019;14(3):e0213901. [doi: 10.1371/journal.pone.0213901] [Medline: 30897144]
16. Goldenberg SM, Krüsi A, Zhang E, Chettiar J, Shannon K. Structural determinants of health among im/migrants in the indoor sex industry: experiences of workers and managers/owners in metropolitan Vancouver. PLoS One 2017;12(1):e0170642. [doi: 10.1371/journal.pone.0170642] [Medline: 28141835]
17. Chambers LA, Rueda S, Baker DN, et al. Stigma, HIV and health: a qualitative synthesis. BMC Public Health 2015 Sep 3;15:848. [doi: 10.1186/s12889-015-2197-0] [Medline: 26334626]
18. Gagnon M. Re-thinking HIV-related stigma in health care settings: a qualitative study. J Assoc Nurses AIDS Care 2015;26(6):703-719. [doi: 10.1016/j.jana.2015.07.005] [Medline: 26300466]
19. Li ATW, Wong JPH, Cain R, Fung KPL. Engaging African-Caribbean, Asian, and Latino community leaders to address HIV stigma in Toronto. Int J Migr Health Soc Care 2016 Dec 12;12(4):288-300. [doi: 10.1108/IJMHSC-07-2014-0029]
20. Stangl AL, Lloyd JK, Brady LM, Holland CE, Baral S. A systematic review of interventions to reduce HIV - related stigma and discrimination from 2002 to 2013: how far have we come? J Int AIDS Soc 2013 Nov;16(3 Suppl 2):18734. [doi: 10.7448/IAS.16.3.18734] [Medline: 24242268]
21. Chen YYB, Li AT, Fung KP, Wong JP. Improving access to mental health services for racialized immigrants, refugees, and non-status people living with HIV/AIDS. J Health Care Poor Underserved 2015 May;26(2):505-518. [doi: 10.1353/hpu.2015.0049] [Medline: 25913347]
22. Wang F, Gao J, Hao S, et al. Empowering Chinese university health service providers to become mental health champions: insights from the ACE-LYNX intervention. Front Psychiatry 2024;15:1349476. [doi: 10.3389/fpsyt.2024.1349476] [Medline: 38585479]
23. Fung KPL, Liu JJW, Wong JPH. Exploring mechanisms of mental illness stigma reduction in Asian Canadian men. Can J Psychiatry 2022 Jun;67(6):490-498. [doi: 10.1177/07067437211018674] [Medline: 34027706]
24. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006 Jan;44(1):1-25. [doi: 10.1016/j.brat.2005.06.006] [Medline: 16300724]
25. Faezipour M, Ghanbaripanah A, Seyedalinaghi S, Hajiabdolbaghi M, Voltarelli F. Effectiveness of acceptance and commitment therapy on reducing depression among people living with HIV/AIDS. J Int Transl Med 2018;6:125-129. [doi: 10.11910/2227-6394.2018.06.03.04]
26. Hayes SC, Strosahl KD, Wilson KG. Acceptance and Commitment Therapy: The Process and Practice of Mindful Change, 2nd edition: The Guilford Press; 2012. URL: https://www.guilford.com/books/Acceptance-and-Commitment-Therapy/Hayes-Strosahl-Wilson/9781462528943 [accessed 2026-01-06]
27. Wong JP, Li AT. The exclusionary circle game: a tool to promote critical dialogue about HIV stigma and social justice. Prog Community Health Partnersh 2015;9(3):431-438. [doi: 10.1353/cpr.2015.0063] [Medline: 26548795]
28. Li A, Wong JP. CHAMP: mobilizing people living with HIV and allies to champion HIV prevention and care in ethno-racial communities. TheBodyPro. 2016. URL: https://www.thebodypro.com/article/champ-mobilizing-people-living-with-hiv-and-allies [accessed 2026-01-06]
29. Wong JPH, Jia CX, Vahabi M, et al. Promoting the mental health of university students in China: protocol for contextual assessment to inform intervention design and adaptation. JMIR Res Protoc 2021 May 11;10(5):e25009. [doi: 10.2196/25009] [Medline: 33973869]
30. Fung K, Cheng SL, Ning X, et al. Mental health promotion and stigma reduction among university students using the Reach, Efficacy, Adoption, Implementation, and Maintenance (RE-AIM) framework: protocol for a mixed methods study. JMIR Res Protoc 2021 Aug 26;10(8):e25592. [doi: 10.2196/25592] [Medline: 34435956]
31. Jacobs RJ, Caballero J, Ownby RL, Kane MN. Development of a culturally appropriate computer-delivered tailored Internet-based health literacy intervention for Spanish-dominant Hispanics living with HIV. BMC Med Inform Decis Mak 2014 Nov 30;14:103. [doi: 10.1186/s12911-014-0103-9] [Medline: 25433489]
32. Ownby RL, Waldrop-Valverde D, Jacobs RJ, Acevedo A, Caballero J. Cost effectiveness of a computer-delivered intervention to improve HIV medication adherence. BMC Med Inform Decis Mak 2013 Feb 28;13:29. [doi: 10.1186/1472-6947-13-29] [Medline: 23446180]
33. Chandanabhumma PP, Narasimhan S. Towards health equity and social justice: an applied framework of decolonization in health promotion. Health Promot Int 2020 Aug 1;35(4):831-840. [doi: 10.1093/heapro/daz053] [Medline: 31236575]
34. Suarez-Balcazar Y. Meaningful engagement in research: community residents as co - creators of knowledge. Am J Community Psychol 2020 Jun;65(3-4):261-271. [doi: 10.1002/ajcp.12414] [Medline: 31907950]
35. Derrien M, Brinkley W, Blahna D, et al. Joining collective impact and community science: a framework for core collaborative community science. Ecol Soc 2024;29(1). [doi: 10.5751/ES-14867-290128]
36. Israel BA, Coombe CM, Cheezum RR, et al. Community-based participatory research: a capacity-building approach for policy advocacy aimed at eliminating health disparities. Am J Public Health 2010 Nov;100(11):2094-2102. [doi: 10.2105/AJPH.2009.170506] [Medline: 20864728]
37. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [doi: 10.2105/ajph.89.9.1322] [Medline: 10474547]

XSL•FO
RenderX

38. Glasgow RE, Harden SM, Gaglio B, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64. [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
39. McCreight MS, Rabin BA, Glasgow RE, et al. Using the Practical, Robust Implementation and Sustainability Model (PRISM) to qualitatively assess multilevel contextual factors to help plan, implement, evaluate, and disseminate health services programs. Transl Behav Med 2019 Nov 25;9(6):1002-1011. [doi: 10.1093/tbm/ibz085] [Medline: 31170296]
40. Simister HD, Tkachuk GA, Shay BL, Vincent N, Pear JJ, Skrabek RQ. Randomized controlled trial of online acceptance and commitment therapy for fibromyalgia. J Pain 2018 Jul;19(7):741-753. [doi: 10.1016/j.jpain.2018.02.004] [Medline: 29481976]
41. Trompetter HR, Bohlmeijer ET, Lamers SMA, Schreurs KMG. Positive psychological wellbeing is required for online self-help acceptance and commitment therapy for chronic pain to be effective. Front Psychol 2016;7:353. [doi: 10.3389/fpsyg.2016.00353] [Medline: 27014159]
42. Reinius M, Wettergren L, Wiklander M, Svedhem V, Ekström AM, Eriksson LE. Development of a 12-item short version of the HIV stigma scale. Health Qual Life Outcomes 2017 May 30;15(1):115. [doi: 10.1186/s12955-017-0691-z] [Medline: 28558805]
43. Kinsler JJ, Wong MD, Sayles JN, Davis C, Cunningham WE. The effect of perceived stigma from a health care provider on access to care among a low-income HIV-positive population. AIDS Patient Care STDS 2007 Aug;21(8):584-592. [doi: 10.1089/apc.2006.0202] [Medline: 17711383]
44. Feldman G, Hayes A, Kumar S, Greeson J, Laurenceau JP. Cognitive and Affective Mindfulness Scale--Revised (CAMS-R). APA PsycTests 2007. [doi: 10.1037/t49653-000]
45. Feldman G, Hayes A, Kumar S, Greeson J, Laurenceau JP. Mindfulness and emotion regulation: the development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R). J Psychopathol Behav Assess 2007 Jul 13;29(3):177-190. [doi: 10.1007/s10862-006-9035-8]
46. Bond FW, Hayes SC, Baer RA, et al. Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: a revised measure of psychological inflexibility and experiential avoidance. Behav Ther 2011 Dec;42(4):676-688. [doi: 10.1016/j.beth.2011.03.007] [Medline: 22035996]
47. Carver CS. You want to measure coping but your protocol's too long: consider the brief COPE. Int J Behav Med 1997;4(1):92-100. [doi: 10.1207/s15327558ijbm0401_6] [Medline: 16250744]
48. Neff KD, Tóth-Király I, Yarnell LM, et al. Examining the factor structure of the Self-Compassion Scale in 20 diverse samples: support for use of a total score and six subscale scores. Psychol Assess 2019 Jan;31(1):27-45. [doi: 10.1037/pas0000629] [Medline: 30124303]
49. Bowser DM, Henry BF, McCollister KE. Cost analysis in implementation studies of evidence-based practices for mental health and substance use disorders: a systematic review. Implement Sci 2021 Mar 12;16(1):26. [doi: 10.1186/s13012-021-01094-3] [Medline: 33706780]
50. Yin RK. Case Study Research: Design and Methods, 5th edition: Sage Publications; 2014. [doi: 10.3138/CJPE.30.1.108]
51. Stake RE. Qualitative case studies. In: Lincoln YS, Denzin NK, editors. The SAGE Handbook of Qualitative Research, 3rd edition: Sage Publications; 2005:443-466 URL: https://emu.tind.io/record/47727 [accessed 2026-01-06]
52. Dooley LM. Case study research and theory building. Adv Dev Hum Resour 2002 Aug 1;4(3):335-354. [doi: 10.1177/15222302004003007]
53. Eisenhardt KM. Building theories from case study research. Acad Manage Rev 1989 Oct;14(4):532. [doi: 10.2307/258557]
54. Merriam SB, Tisdell EJ. Qualitative Research: A Guide to Design and Implementation: Jossey-Bass; 2009. URL: https://download.e-bookshelf.de/download/0003/7195/84/L-G-0003719584-0007575839.pdf [accessed 2026-01-06]
55. Sandelowski M. Qualitative analysis: what it is and how to begin. Res Nurs Health 1995 Aug;18(4):371-375. [doi: 10.1002/nur.4770180411] [Medline: 7624531]
56. Byrne D. A worked example of Braun and Clarke's approach to reflexive thematic analysis. Qual Quant 2022 Jun;56(3):1391-1412. [doi: 10.1007/s11135-021-01182-y]
57. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health 2019 Aug 8;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
58. Braun V, Clarke V, Hayfield N. 'A starting point for your journey, not a map': Nikki Hayfield in conversation with Virginia Braun and Victoria Clarke about thematic analysis. Qual Res Psychol 2022 Apr 3;19(2):424-445. [doi: 10.1080/14780887.2019.1670765]
59. Clarke V, Braun V. Thematic analysis. J Posit Psychol 2017 May 4;12(3):297-298. [doi: 10.1080/17439760.2016.1262613]
60. Rao D, Elshafei A, Nguyen M, Hatzenbuehler ML, Frey S, Go VF. A systematic review of multi-level stigma interventions: state of the science and future directions. BMC Med 2019 Feb 15;17(1):41. [doi: 10.1186/s12916-018-1244-y] [Medline: 30770756]
61. Wang F, Gao J, Yao ZY, et al. Reducing stigma and promoting empowerment: a pre-post evaluation of ACE-LYNX intervention on the mental health literacy of university providers. Int J Ment Health Promot 2025;27(10):1497-1514. [doi: 10.32604/ijmhp.2025.069458]

XSL•FO
RenderX

62. Li ATW, Fung KPL, Maticka-Tyndale E, Wong JPH. Effects of HIV stigma reduction interventions in diasporic communities: insights from the CHAMP study. AIDS Care 2018 Jun;30(6):739-745. [doi: 10.1080/09540121.2017.1391982] [Medline: 29063782]

Abbreviations

AAQ-II: Acceptance and Action Questionnaire Version II
ACE: Acceptance and Commitment to Empowerment
ACE-LYNX-PRO: Acceptance and Commitment to Empowerment–Linking Youth and 'Xin' (hearts) with Professionals
ACT: acceptance and commitment therapy
Brief COPE: Brief Coping Orientation to Problems Experienced
CAMS-R: Cognitive and Affective Mindfulness Scale
CHAMP: Community Champions HIV/AIDS Advocates Mobilization Project
KTE: knowledge translation and exchange
LGBTQ+: lesbian, gay, bisexual, transgender, queer, and other gender and sexual minorities
PRISM: Practical Robust Implementation and Sustainability Model
RE-AIM: reach, effectiveness, adoption, implementation, maintenance
SCS-S: Self-Compassion Scale Short
SJCB: social justice capacity building
TTT: train-the-trainer

Edited by A Schwartz, T Leung; submitted 15.Jul.2025; this is a non–peer-reviewed article;accepted 30.Nov.2025; published 26.Jan.2026.

Please cite as:
Wong JPH, Li ATW, Hilario CT, Vahabi M, Etowa E, Luginaah I, dela Cruz AM, Narushima M, Fung KPL
Acceptance and Commitment to Empowerment Intervention to Reduce HIV Stigma and Promote Community Resilience: Protocol for an Implementation Study
JMIR Res Protoc 2026;15:e80669
URL: https://www.researchprotocols.org/2026/1/e80669
doi:10.2196/80669

© Josephine Pui-Hing Wong, Alan Tai-Wai Li, Carla T Hilario, Mandana Vahabi, Egbe Etowa, Isaac Luginaah, Aniela M dela Cruz, Miya Narushima, Kenneth Po-Lun Fung. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Scaling Up a Diabetes Prevention Program in Geographically and Ethnoculturally Diverse Urban Regions of Canada: Protocol for a Hybrid Type 2 Implementation-Effectiveness Study

Katie Weatherson[1], MSc; Jessica E Bourne[1], PhD; Kaela Cranston[1], PhD; Kyra Braaten[1], MSc; Natalie Grieve[1], MSc; Joseph Kelly[1], MSc; Mary E Jung[1], PhD

Department of Health and Exercise Sciences, University of British Columbia Okanagan, Kelowna, BC, Canada

Corresponding Author:
Mary E Jung, PhD
Department of Health and Exercise Sciences
University of British Columbia Okanagan
1238 Discovery Avenue
Kelowna, BC, V1V 1V9
Canada
Phone: 1 2508079670
Email: mary.jung@ubc.ca

Abstract

Background: It is estimated that type 2 diabetes (T2D) impacts an estimated 5.3 million Canadians, despite the condition being largely preventable. Laboratory-based diabetes prevention programs (DPPs) have limited effectiveness when translated into community settings due to their low-quality delivery and inability to reach people in the community most in need. To date, no community-based DPPs have been implemented nationwide across Canada. Small Steps for Big Changes (SSBC) is a diet and physical activity counseling intervention that significantly reduces the risk of developing T2D and has been designed for feasible delivery by community-dwelling peers. To ensure maximum public health impact, SSBC must be optimally implemented, demonstrate effectiveness for diverse groups, and be sustainable over time.

Objective: This project aims to adapt SSBC and evaluate the implementation, effectiveness, and sustainability of SSBC in diverse urban communities across Canada.

Methods: A hybrid type 2 implementation-effectiveness study design using multiple and mixed methods will be used to evaluate the implementation and effectiveness of SSBC over 6 years in partnership with 11 regional Young Men's Christian Associations across 8 provinces in Canada. Beginning in 2024, we will (1) adapt and implement SSBC in diverse urban cities across Canada; (2) examine the implementation (including implementation strategies), effectiveness, and cost-effectiveness of SSBC (2024-2028); and (3) determine the sustainability of SSBC at each delivery location (2028-2029). Data will be collected from SSBC clients, coaches, site leads, and senior leadership municipality partners. The project will be overseen by an advisory group and 3 committees focused on sex, gender, and inclusivity; program evaluation; and diabetes prevention engagement. This study has received ethical approval from the University of British Columbia Clinical Research Ethics Board (H23-01930).

Results: Funding for this project began in October 2022, and Institutional Review Board approval was obtained in October 2023. Program implementation within each region is occurring in a phased approach, with partners beginning program delivery in one site (2024-2025) before expanding to any additional locations (2025-2027). Program enrollment occurs continuously during the implementation phase across all sites. From 2024-2025, a total of 13 delivery sites began program delivery, 722 participants have enrolled in the program, and 406 have begun the program (153/342, 44.7% non-Western or Eastern European; 80/345, 23.2% men or 82/347, 23.6% male). An additional 13 sites have confirmed they will launch the program in 2026.

Conclusions: This study will demonstrate that SSBC can be scaled up nationwide to effectively and equitably reduce the Canadian population-level risk of T2D. This work will determine best practice implementation determinants, outcomes, and strategies critical for sustaining DPP implementation across Canada and beyond. Project findings will be shared with municipality partners and will be copresented with partners and SSBC clients to community organizations, local interested parties, and academics.

Trial Registration: ClinicalTrials.gov NCT06440395; https://clinicaltrials.gov/study/NCT06440395

International Registered Report Identifier (IRRID): DERR1-10.2196/80276

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e80276) doi:10.2196/80276

KEYWORDS

diabetes prevention program; implementation science; implementation effectiveness study; sustainability

Introduction

The incidence of type 2 diabetes (T2D) is continuing to rise worldwide [1,2] despite overwhelming evidence that this chronic disease is largely preventable [3-5]. As the fastest-growing chronic condition, T2D impacts an estimated 5.3 million Canadians [6], with estimated annual treatment costs of CAD $30 billion (a currency exchange rate of CAD $1=US $0.71 was applicable) [7]. An additional 6 million Canadians [6] live with prediabetes, a clinical condition recognized by the World Health Organization that precedes T2D. Further, 70% of individuals with prediabetes are estimated to progress to T2D in the next 10 years [8]. Rates of T2D are higher in areas with lower household income and education level, higher levels of food insecurity, unemployment, immigrant status, and among individuals of African, South Asian, and Indigenous origins [7-10]. Racial, sex, and gender-based health inequities in accessing care exacerbate these barriers [11-16]. Preventing the progression of prediabetes to T2D and addressing health inequities is a public health priority in Canada [17].

T2D prevention programs (diabetes prevention program [DPPs]) have been shown to reduce the progression of prediabetes to T2D in diverse populations and geographic regions of the world. Three landmark trials involving more than 4000 individuals with prediabetes demonstrated that increasing physical activity (PA) and inducing modest weight loss through dietary intervention can reduce the risk of developing T2D by between 50% to 60% [18-20]. Notably, diet and PA approaches were twice as efficacious as pharmaceutical interventions [18]. However, translating these highly efficacious DPPs into community settings for public health impact often results in reduced effectiveness [21,22]. This drop in efficacy is often attributed to inequitable access [23,24], poor fidelity [25-29], lack of cultural tailoring [11,30-34], and lack of sustainability plans [30,35-37].

To date, no community-based nationwide DPP has been implemented across Canada. To make a measurable impact on T2D prevention at a population level, potent evidence-based interventions must be implemented effectively at scale. Nationwide uptake of DPPs is hindered by the absence of evidence-informed implementation plans tailored to local contexts [38] and the examination of individual- and between-delivery site differences to understand the impact of context on implementation. Site contextual differences are common, as each community has unique challenges and opportunities that impact implementation, and remain an understudied paradox of implementation efforts [39]. Additionally, equity-based contextual factors, including access and cultural relevance for populations traditionally underserved and racialized against, are pivotal considerations for implementation scale-up [40,41]. It is imperative to disaggregate multisite data and understand how community-, organization-, provider-, and client-level factors contribute to implementation outcomes when organizations deliver the same program in different contexts [41,42]. Implementation studies with a focus on context, and in particular health equity, are needed to compare within and between-site implementation outcomes and determinants [41]. Understanding these contextual factors is also key for identifying methods (implementation strategies) required to enhance the implementation and scale-up of DPPs [43].

Further limiting DPPs from achieving population-level health impact is a failure to prioritize and support the sustainability of these evidence-based programs after they have been implemented. Sustainability, defined as the continued use of program components at sufficient intensity for the sustained achievement of desirable program goals and population outcomes [44], is now considered “one of the most significant translational research problems of our time” [45]. Sustaining evidence-based interventions is imperative to assess long-term health outcomes, inform scale-up of implemented programs, avoid debilitating policy and ethical implications of abandoning funded effective interventions, and maintain community support and trust in research [44]. To maximize the public health impact of evidence-based interventions, it is critical to investigate the factors and processes that influence both implementation and sustainability of these interventions, and the health behaviors and outcomes they produce [44,46]. Adaptability, partnership with organizations delivering the intervention, and commitment from the community have been suggested as key factors in sustaining disease prevention programs [47,48]. The call for sustainable solutions to T2D prevention has yet to be answered [49], with no study to date examining how to sustain evidence-based DPPs in “real-world,” community contexts. Implementing and sustaining evidence-based DPPs has implications for substantially reducing the incidence of T2D in Canada, and beyond, and simultaneously fulfilling an important knowledge gap in implementation science.

Small Steps for Big Changes (SSBC) is an evidence-based DPP, consisting of behavior change techniques for diet and PA modification delivered in six 1-hour sessions. Since development in 2013, SSBC has been systematically scaled up from preliminary efficacy testing in a laboratory setting to real-world effectiveness trials in 1 urban and 5 rural community settings in British Columbia (BC) [37,50-55]. Our research to date has demonstrated that SSBC can be adapted for implementation in community settings, delivered in-person and virtually by “laypersons” with no specialized education or experience [56,57], and results in short- [58], medium- [52,59], and long-term behavior changes [60] (J Bourne et al, unpublished data, 2025) leading to clinically relevant improvements. Most notably, 41% of clients who participate in SSBC no longer have prediabetes (glycated hemoglobin [HbA_{1c}] <5.7%; American Diabetes Association criteria) 1 year post program, with a further 16% on the cusp of no longer being prediabetic with HbA_{1c} levels of 5.7% (J Bourne et al, unpublished data, 2025).

XSL•FO
RenderX

Nationwide scale-up of SSBC presents a promising public health solution to equitably reach and improve the health of the growing number of Canadians who are at risk of developing T2D. As such, this project aims to evaluate the implementation, effectiveness, and sustainability of SSBC as it is scaled up nationwide across diverse urban communities in Canada. To address this aim, the project has three objectives (Figure 1): (1) To adapt and implement SSBC in diverse urban cities across Canada; (2) to examine the implementation (determinants, outcomes, and strategies), effectiveness, and cost-effectiveness of SSBC; and (3) determine the sustainability (determinants and outcomes) of delivering SSBC in these contextually varied urban locations.

Figure 1. Study objectives. CFIR: Consolidated Framework for Implementation Research; CFIR-ERIC: Consolidated Framework for Implementation Research–Expert Recommendations for Implementing Change; FRAME: Framework for Reporting Adaptations and Modifications-Expanded; FRAME-IS: Framework for Reporting Adaptations and Modifications for Implementation Strategies; SSBC: Small Steps for Big Changes.

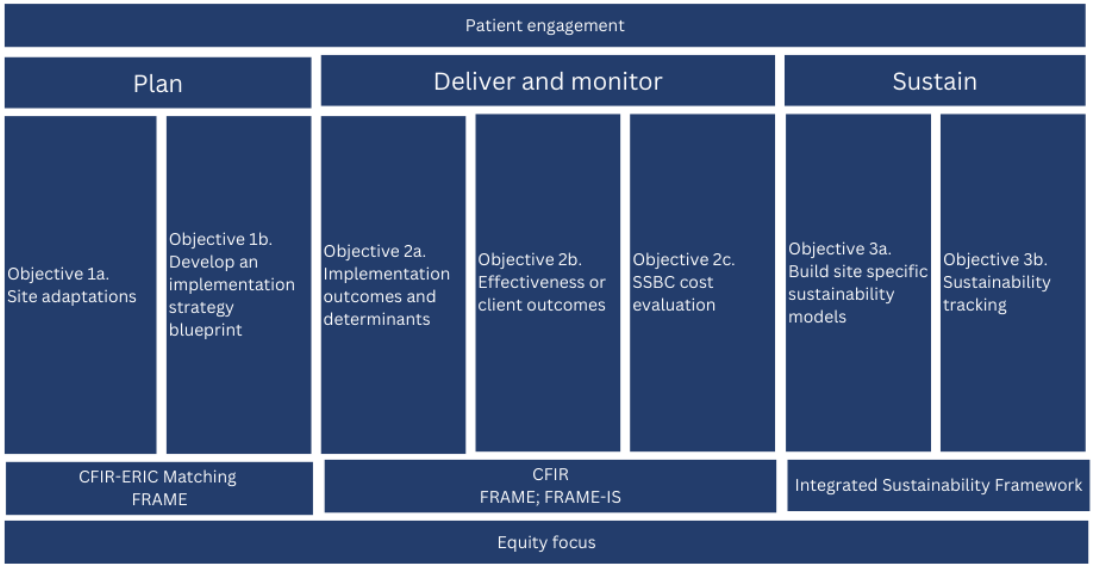

Methods

Study Design and Procedures

A hybrid type 2 implementation-effectiveness study design [61] using multiple and mixed methods will be used to evaluate the implementation and effectiveness of SSBC over 6 years (2024-2029). Years 1 and 2 will serve as a planning period (preimplementation phase). Specifically, SSBC will be adapted to the site and local contexts, recruitment and referral plans will be developed, and coaches will be trained to deliver SSBC. Years 2-4 will involve the delivery and monitoring of SSBC in diverse urban Young Men's Christian Association (YMCA) health and fitness facilities (implementation phase). Individuals who enroll in SSBC will be followed for 24 months after completion of the program (until the end of year 6). Data collected during the implementation phase will be used to guide the cocreation of sustainability models with partner organizations for continued delivery of SSBC in each municipality (sustainability phase in years 5 and 6; see Figure 2 for project timeline).

Figure 2. Small Steps for Big Changes delivery sites in Canada. YMCA: Young Men's Christian Association.

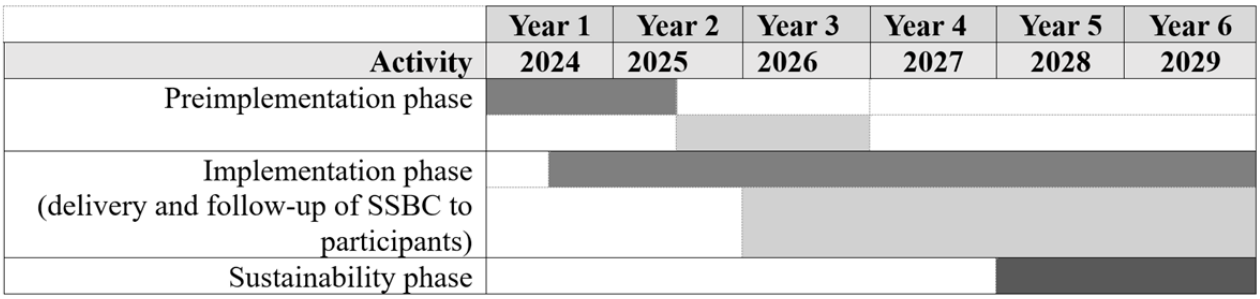

Intervention

SSBC aims to empower individuals with prediabetes to make long-lasting diet and PA changes. Trained SSBC coaches meet with clients for 6 one-on-one sessions focused on PA, reducing intake of added sugar, improving carbohydrate choices, and preventing relapse, with each session followed by supervised aerobic exercise. Coaches use a motivational interviewing-informed approach to bolster client autonomy. The core components of the SSBC program can be found in Textbox 1, and specific behavior change techniques are published elsewhere [62].

XSL•FO
RenderX

Textbox 1. The core components of Small Steps for Big Changes.

Structural elements

- Six sessions incorporating guidance on diet and physical activity behavior (≈40 minutes) and structured exercise (≈20 minutes).
- Delivered one-on-one.
- Choice-based high-intensity interval training or moderate-intensity continuous training aerobic exercise.
- Coach-supported exercise.
- Counseling style informed by motivational interviewing.

Context

- Delivered in the community.
- Deliverable by community lay persons who have received brief virtual coach training.

Intervention practices (diet and exercise)

- Brief action planning.
- Relapse prevention planning.
- Self-monitoring.
- Offering education regarding diet (specifically carbohydrate and added sugar content) and physical activity choices.
- Independent practice of goal-oriented diet or physical activity.
- Supporting the client to reflect on behaviors.

SSBC Coach Training

SSBC is delivered by community-dwelling peers (eg, YMCA staff or volunteers and past clients) without postsecondary education or counseling experience requirements. All individuals who deliver SSBC complete coach training on an online platform, consisting of 7 asynchronous modules, a posttraining knowledge check, and a mock scenario. The mock scenario consists of a coach meeting with an SSBC staff member acting as a standardized client, where the coach must demonstrate correct delivery of intervention content and exercise protocols using a client-centered level of motivational interviewing. The current training was developed with feedback from SSBC coaches [63] and is effective and acceptable [64], with coaches retaining the learnt counseling skills over time (K Cranston et al, unpublished data, 2025).

Organization Implementation Sites

Eleven distinct regional YMCAs across 8 provinces in Canada will participate in this study (see Figure 3 for specific locations). In year 2, SSBC will be delivered in 11 initial implementation sites, one in each distinct YMCA region. Sites were purposively selected to maximize ethnocultural, population density, and socioeconomic diversity, as these contextual factors may influence implementation, effectiveness, and sustainability. Specifically, each site represents a region with high prediabetes prevalence, access to priority populations (operationalized in this project as health equity-owed groups who may face systemic barriers to access, opportunities, and resources to DPPs due to factors such as race, ethnicity, gender, and socioeconomic status; [11-16]), and staffing capacity to deliver SSBC within the organization. Each regional YMCA oversees several health and fitness facility buildings (often in distinct cities) into which SSBC delivery will scale up in years 3 and 4, for a total of 44 Canadian implementation sites.

XSL•FO
RenderX

Figure 3. Small Steps for Big Changes Delivery Sites in Canada.



Each implementation site will select a staff member to act as the site lead for SSBC implementation. The role of each site lead will be to oversee all SSBC coaches and the day-to-day delivery of the program at their site. Senior leadership includes managerial staff (eg, chief executive officer) who are responsible for implementation, site selection, and will help identify site leads, but will not oversee the day-to-day operations of SSBC.

Data Collection

This study consists of 3 interconnected objectives, highlighted above, that align with the 3 phases of work. Data collection and analysis plans are outlined for each objective below.

Preimplementation Phase

Objective 1a: Adapt SSBC to Each Implementation Site

SSBC has been successfully implemented in 1 urban location in BC [36,50-52]. However, the implementation and sustainability of SSBC in other urban municipalities requires adaptation of these materials and processes. Site-level adaptations will ensure program content is relevant to the local context and delivery is compatible with each site, without altering the core, evidence-based principles of SSBC (Textbox 1). To maximize shared decision-making, an adaptation meeting will be held with each site in the preimplementation phase. This meeting will involve 1 site lead, a minimum of 1 YMCA coach, a regional research coordinator (member of the research team), and the research manager. In addition, regular meetings will be held between the regional research coordinator and the site lead during which program adaptations or the need to make program adaptations will be discussed. All adaptations will be captured using standard templates created by the research team. Site-specific adaptations will be reported and recorded by the research team on an ongoing basis during the implementation phase. Content analysis [65] using the FRAME (Framework for Reporting Adaptations and Modifications-Expanded) [66] will be used to classify adaptations and modifications to SSBC.

Objective 1b: Develop an Implementation Strategy Plan and Tracking System

This work will involve developing an implementation strategy plan, or "blueprint," along with an associated toolkit that formally outlines the strategies for scaling up SSBC across Canada. Implementation strategies are the "methods or techniques used to enhance the adoption, implementation, and sustainability of a program" [43]. The development of this plan will be guided by the identification of barriers and facilitators to SSBC implementation from three sources: (1) findings from the delivery of SSBC at three YMCA facilities in an urban location in BC [50], (2) findings from the delivery of SSBC in five rural locations in BC (K Braaten et al, unpublished data, 2025), and (3) meetings with the research team and YMCA senior leadership in planning for the urban expansion of SSBC. Barriers and facilitators to implementation identified from the 3 sources will be matched to the Consolidated Framework for Implementation Research (CFIR) [67]. The CFIR–Expert Recommendations for Implementing Change matching tool [68] will be used to identify potential implementation strategies that can be used to address each identified factor. Each implementation strategy will be operationalized for the current

XSL•FO
RenderX

context, and information on the actor (ie, who delivers the strategy), temporality (ie, when or at what time), and the targeted implementation outcome will be provided such that this information can be easily used by the delivery sites. This process will provide a rigorous and transparent method for linking implementation barriers and facilitators identified across diverse delivery settings to implementation strategies for use in the current project. Use of implementation strategies will be captured through regular meetings between the research team and each delivery site. Content analysis of regular meeting minutes will be conducted using the FRAME-IS (Framework for Reporting Adaptations and Modifications for Implementation Strategies) [69]. In addition, reasons for implementation strategy adaptation, use, or exclusion will be captured. The integrated findings from objectives 1a and 1b will inform the development of each site's sustainability planning in Objective 3.

Implementation Phase

Objective 2a: Examine the Implementation of SSBC and Associated Implementation Strategies

Implementation will be measured via 4 implementation outcomes at both the level of the intervention and implementation strategies within and between sites. These outcomes include (1) adoption, (2) reach, (3) dose delivered, and (4) fidelity. Table 1 details the implementation outcomes and determinants, how they are operationalized, the source, and the method of data collection. Table 2 provides details of data collection time points for each participant (ie, client, coach, site lead, etc) involved in this study.

XSL•FO
RenderX

Table 1. Overview of implementation outcomes and determinants, their operationalization, data source, and measurement tool.

Intervention implementation			Implementation strategies		
Indicators or definition	Data source	Measurement or tool	Indicators or definition	Data source	Measurement or tool
Implementation outcomes					
Adoption					
Number of coaches, proportion of staff who are coaches, and representativeness of coaches	C[a] and SL[b]	Annual reporting and qualitative interviews	Representativeness of the delivery support team (SL, managers) that use IS[c]	C, SL, and Y-SL[d]	Demographics survey and monthly meetings (SL and RRC[e])
Demographics of coaches and site leads	C and SL	Coach survey	Representativeness of the research team (RRC and DPRG[f]) that uses IS	RRC and DPRG	Demographic survey with the research team
Reach					
Number of individuals interested in SSBC[g] but do not sign up, number of clients enrolled in SSBC, and number that withdraw after enrollment	Client	REDCap[h] audit of eligibility survey	Number of delivery staff interested, trained, and retained as SSBC coaches	C and SL	Annual report and monthly meetings (SL and RRC); virtual coach training platform tracking
Demographic characteristics of clients	Client	Client survey	Recruitment strategies reaching participant	Client	Client eligibility survey
Dose delivered					
Number of sessions attended and length of session	C	Session checklists	Number of times IS is used at the delivery site and by the research team	SL, RRC, and DPRG	Monthly meetings (SL and RRC) and the research team IS tracking
Fidelity					
Frequency with which participants receive each core component (checklists)	C	Session checklists and audio audits	The extent to which IS was implemented as prescribed in the implementation blueprint	SL	Monthly meetings (SL and RRC)
Quality of counseling skills	C	Audio audits of sessions and A-MICA[i] coded	The extent to which IS was implemented as prescribed in the implementation blueprint	SL	Monthly meetings (SL and RRC)
Implementation determinants					
Program cost					
Cost per site (trainer costs and delivery time)	SL	Monthly invoices and annual report	Costs incurred from conducting IS	SL, RRC, and DPRG	Annual report and monthly meetings (SL and RRC)
Membership referral rates	Client and SL	Client survey and annual report	Costs incurred from conducting IS	SL, RRC, and DPRG	Annual report and monthly meetings (SL and RRC)
Retention and conversion (number of YMCA[j] memberships purchased, SSBC completion)	Client and SL	Client survey and annual report	Costs incurred from conducting IS	SL, RRC, and DPRG	Annual report and monthly meetings (SL and RRC)
Adaptations					
Number and type of adaptations made to the SSBC program before delivery	C, SL, and Y-SL	Adaptation workshop minutes and preimplementation tracking sheets	Extent to which IS was adapted, tailored, refined, or reinvented to meet the needs of organizations at scale-up	C, SL, Y-SL, RRC, and DPRG	Monthly meetings (SL and RRC) and interviews with C, SL, Y-SL, RRC, and DPRG

XSL•FO
RenderX

Intervention implementation			Implementation strategies		
Indicators or definition	Data source	Measurement or tool	Indicators or definition	Data source	Measurement or tool
Number and type of adaptations made to the SSBC program during delivery	C, SL, and RRC	Monthly meetings (SL and RRC), and C and SL interviews	Extent to which IS was adapted, tailored, refined, or reinvented to meet the needs of organizations at scale-up	C, SL, Y-SL, RRC, and DPRG	Monthly meetings (SL and RRC), and interviews with C, SL, Y-SL, RRC, and DPRG
Number and type of adaptations made to sustain SSBC in years 5 and 6	C, SL, Y-SL, and PP[k]	Focus groups and monthly meetings (SL and RRC)	Extent to which IS was adapted, tailored, refined, or reinvented to meet the needs of organizations at scale-up	C, SL, Y-SL, RRC, and DPRG	Monthly meetings (SL and RRC), and interviews with C, SL, Y-SL, RRC, and DPRG
Acceptability					
Client satisfaction and acceptability	Client	Client survey (net promoter score, program satisfaction, theoretical framework of acceptability), and interviews	Perceptions among the site lead, coaching staff, or support system that IS are agreeable, palatable, or satisfactory (including SSBC training, program manuals, etc)	C, SL, Y-SL, RRC, and DPRG	Interviews with C, SL, Y-SL, RRC, and DPRG
Coach and site lead satisfaction and acceptability	C and SL	C and SL survey (BISS[l], AIM[m], IAM[n], and FIM[o]), and interviews (after 10 clients)	Perceptions among the site lead, coaching staff, or support system that IS are agreeable, palatable, or satisfactory (including SSBC training, program manuals, etc)	C, SL, Y-SL, RRC, and DPRG	Interviews with C, SL, Y-SL, RRC, and DPRG
Context					
Aspects of the larger social, political, and economic environment that may influence SSBC implementation	C, SL, Y-SL, and RRC	Monthly meetings (SL and RRC), and interviews with C, SL, and Y-SL. Interview guide based on CFIR[p] and ISF[q]	Aspects of the larger social, political, and economic environment that may influence the delivery of the implementation strategies	C, SL, Y-SL, RRC, and DPRG	Interviews with C, SL, Y-SL, RRC, and DPRG
Feasibility					
Perceptions among the delivery team that SSBC can be successfully used and carried out within their organization	C, SL, Y-SL, and RRC	Monthly meetings (SL and RRC), and interviews with C, SL, and Y-SL. Interview guide based on CFIR and ISF	Perceptions among the delivery team and research team that the implementation strategies can be successfully used or carried out at scale within different organizations or settings	C, SL, Y-SL, RRC, and DPRG	Interviews with C, SL, Y-SL, RRC, and DPRG
Compatibility					
Extent to which SSBC fits with the mission, priorities, and values of the delivery site	C, SL, Y-SL, and RRC	Monthly meetings (SL and RRC), and interviews with C, SL, and Y-SL. Interview guide based on CFIR and ISF	Extent to which implementation strategies fit with the mission, priorities, and values of organizations at scale (including SSBC training)	C, SL, Y-SL, and RRC	Interviews: C, SL, Y-SL, and RRC
Culture					
The organizations' norms, values, and basic assumptions around SSBC	C, SL, Y-SL, and RRC	Monthly meetings (SL and RRC), and interviews with C, SL, and Y-SL. Interview guide based on CFIR and ISF	Organizations' norms, values, and basic assumptions around selected implementation strategies	C, SL, Y-SL, and RRC	Interviews: C, SL, Y-SL, and RRC
Dose (satisfaction)					

XSL•FO
RenderX

Intervention implementation			Implementation strategies		
Indicators or definition	Data source	Measurement or tool	Indicators or definition	Data source	Measurement or tool
Delivery team's satisfaction with SSBC and interactions with the research team	C, SL, Y-SL, and RRC	Monthly meetings (SL and RRC), and interviews with C, SL, and Y-SL. Interview guide based on CFIR and ISF	Delivery team and support systems' satisfaction with the implementation strategies	C, SL, Y-SL, and RRC	Interviews: C, SL, Y-SL, and RRC
Complexity					
Perceptions of the ease or difficulty in implementing SSBC among the delivery team	C, SL, Y-SL, and RRC	Monthly meetings (SL and RRC), and interviews with C, SL, and Y-SL. Interview guide based on CFIR and ISF	Perceptions among the delivery team and support system that implementation strategies are relatively difficult to understand and use, number of different strategies. Related to implementation setting (eg, coach community of practice, SSBC training, and program manual etc)	C, SL, Y-SL, and RRC	Interviews: C, SL, Y-SL, and RRC
Self-efficacy					
Coaches' belief in their ability to execute the course of action to achieve the implementation goals	C, SL, Y-SL, and RRC	Monthly meetings (SL and RRC); interviews with C, SL, and Y-SL. Interview guide based on CFIR and ISF	Support systems' belief in their ability to execute courses of action to achieve implementation goals	C, SL, Y-SL, and RRC	Interviews: C, SL, Y-SL, and RRC
Organizational readiness					
Organizational readiness	C and SL	Survey at baseline using organizational readiness for change [70]	Not applicable	Not applicable	Not applicable

[a]C: coach.

[b]SL: site lead.

[c]IS: implementation strategies.

[d]Y-SL: Young Men's Christian Association Senior Leadership.

[e]RRC: Regional Research Coordinator.

[f]DPRG: Diabetes Prevention Research Group.

[g]SSBC: Small Steps for Big Changes.

[h]REDCap: Research Electronic Data Capture (Vanderbilt University).

[i]A-MICA: Abbreviated Motivational Interviewing Competency Assessment.

[j]YMCA: Young Men's Christian Association.

[k]PP: patient partner.

[l]BISS: Behavioral Interventionist Satisfaction Survey.

[m]AIM: Acceptablity of Intervention Measure.

[n]IAM: Intervention Appropriateness Measure.

[o]FIM: Feasibility of Intervention Measure.

[p]CFIR: Consolidated Framework for Intervention Research.

[q]ISF: Integrated Sustainability Framework.

XSL•FO
RenderX

Table 2. Data collection time points.

Who and what	Time point							
	Coach training	Adaptation workshop	Baseline	During delivery	Postprogram completion	12 months into delivery	24 months (end of delivery)	Sustainability (years 3-4)
SSBC[a] client								
Consent			✓					
Intake survey			✓					
Questionnaires and in-person measures			✓		✓	✓	✓	
Interviews (2-3 clients per site)					✓[b]			
Site focus group[c]							✓	✓ (x2)
SSBC coach								
Coach training and demographics	✓							
Consent			✓					
Adaptation workshop		✓						
Survey			✓ (ORQ[d])	✓ (BISS[e], AIM[f], IAM[g], FIM[h]; at 5, 10, and 20 clients if applicable)				
Delivery checklists or audio audit				✓				
Interviews (2 coaches per site)				✓[i] (10 clients)				
Site focus groups							✓	✓ (x2)
Delivery site lead								
Consent			✓					
Coach training and demographics	✓							
Adaptation workshop		✓						
Survey			✓ (ORQ)			✓ (AIM, IAM, and FIM)		
Monthly meetings			✓	✓	✓	✓	✓	
Interviews						✓	✓	
Site focus groups							✓	✓ (x2)
Delivery site senior leadership								
Consent and demographics						✓		
Meeting minutes								
Interviews						✓	✓	
Site focus groups							✓	✓ (x2)

XSL•FO
RenderX

Who and what	Time point							
	Coach training	Adaptation workshop	Baseline	During delivery	Postprogram completion	12 months into delivery	24 months (end of delivery)	Sustainability (years 3-4)
Municipality partners[j]: focus group							✓	
Regional research coordinator: interview							✓	
Diabetes prevention research group project manager: interview							✓	

[a]SSBC: Small Steps for Big Changes.

[b]These interviews will occur 2-3 months after the client has completed Small Steps for Big Changes.

[c]Site focus groups focused on creating sustainability models if deemed appropriate and identifying implementation strategies used.

[d]ORQ: Organizational Readiness Questionnaire.

[e]BISS: Behavioral Interventionist Satisfaction Survey.

[f]AIM: Acceptability of Intervention Measure.

[g]IAM: Intervention Appropriateness Measure.

[h]FIM: Feasibility of Intervention Measure.

[i]Coaches must have had 10 clients complete the Small Steps for Big Changes before being interviewed. Note that quantitative measures at 10 clients are to be completed ahead of the interview.

[j]Municipality partners include patient partners, Small Steps for Big Changes alumni, and persons with lived experience.

In addition, the multilevel factors posited to influence intervention implementation and use of implementation strategies (ie, the implementation determinants) will be examined across multiple settings and contexts. This information will be captured qualitatively through monthly meetings between the research team and delivery sites and through one-on-one interviews with 2 coaches at each site, site leads (after 1 year of delivery), and senior leadership at each site (after 2 years of delivery). Interview guides will be guided by the CFIR [67] to understand the contextual factors, organizational factors, processes factors, implementer characteristics, and intervention characteristics that impact implementation.

Quantitatively, organizational readiness will be assessed at baseline [70] through surveys with site leads and coaches. Implementer satisfaction with the program will be captured using the Behavioral Interventionist Satisfaction Survey [71] and the Intervention Acceptability, Appropriateness, and Feasibility [72] measures during the delivery process. Client receptivity and satisfaction will be assessed quantitatively using the 8-item Theoretical Framework for Acceptability [73], a 7-item overall program satisfaction measure, and the 1-item net promoter score [74]. In addition, client interviews will be conducted with a purposive subsample (2-3 clients per site) post-SSBC completion to assess client satisfaction and acceptability. Program costs will be captured through monthly invoices and annual reports from sites.

Quantitative data from adoption, reach, dose, fidelity, program cost, and client and coach receptivity measures will be analyzed descriptively. Client demographics will be used to assess representativeness relative to people with prediabetes across Canada. SSBC program sessions will be audio-recorded using tablets with data saved on the University of British Columbia's secure data research network in real-time; 2 program sessions per year from each coach will be randomly sampled and coded using an abbreviated version of the validated Motivational Interviewing Competency Assessment [75]. Motivational Interviewing Competency Assessment scores will be used to assess coach counseling fidelity, as done previously [57]. All implementation strategies used will be identified and classified according to definitions from the Expert Recommendations for Implementing Change project [76].

Qualitative data from interviews will be audio-recorded, transcribed verbatim, and analyzed using coding reliability thematic analysis [77]. This method aligns with the pragmatic research objectives for these interviews to assess satisfaction and acceptability (clients) and barriers and facilitators of implementation (delivery team), thereby supporting considerations of meaningful and methodological coherence [78]. Interviews will be deductively coded by 2 independent coders to increase the credibility and transferability of the findings with the aim of arriving at one objective truth [78]. Additionally, selecting an aim that will improve the practice of implementation efforts and offer significant contributions to the field further considers rigor and quality in qualitative inquiry [78].

Objective 2b: Examine the Effectiveness of SSBC

All clients will complete in-person measurements (weight and waist circumference) at baseline, 12-, and 24-month post program. At locations with an appropriate space (ie, walking track), cardiorespiratory fitness will also be assessed via a 6-minute walk test [79]. In addition, all clients will complete surveys at baseline, session 6, 12, and 24 months post program via REDCap (Research Electronic Data Capture). Surveys will assess self-reported T2D status (12- and 24-months only

XSL•FO
RenderX

[80,81]), moderate-to-vigorous PA, dietary intake, and health-related quality of life (EQ-5D-5L) [82]. Table 3 outlines the measures that will be used in the evaluation of SSBC effectiveness. Using an equity-focused lens, a prospective cohort subsample of clients (n=1285) will be followed across 24 months to evaluate the differential impact of SSBC on T2D status, assessed by change in HbA_{1c}, in addition to self-reported T2D status. All effectiveness outcomes will be analyzed using descriptive statistics. Hierarchical multivariable regression models will be used to assess the impact of SSBC on effectiveness outcomes, adjusting for potential confounders and accounting for within-subject and within-site clustering. Covariates (ie, sex, gender, and ethnicity) will be included to adjust for within-subject and within-site correlations. A sensitivity analysis using missing data imputation will be conducted to evaluate the potential impact of nonadherence on study outcomes. Data will be disaggregated to examine differences in effectiveness by ethnicity, socioeconomic status, municipality, sex, and gender within and between sites. For the subsample, a multilevel binary logistic regression will be used to examine differences in sociodemographic factors between those who do and do not experience a clinically significant reduction in HbA_{1c} at 24 months.

XSL•FO
RenderX

Table 3. Overview of effectiveness measures and associated time points.

	Eligibility	Baseline measurement	Postprogram measurement	Follow-up measurements	
Measure	Online intake survey	Baseline on session 1	After session 6	12-month follow-up	24-month follow-up
Age	✓				
Current T2D[a] diagnosis	✓				
Prediabetes diagnosis or ADA[b] diabetes risk score	✓				
T2D status					
1-item self-report T2D status (in questionnaire)		✓		✓	✓
Laboratory-based HbA_{1c}[c] (subsample)		✓		✓	✓
Client questionnaire					
Demographic characteristics		✓			
GLTEQ[d] [83]		✓	✓	✓	✓
PAV[e] [84]		✓	✓	✓	✓
4-item 24-hour movement measure (adapted for SSBC[f] urban scale up in line with Canadian 24-hour movement guidelines [85])		✓	✓	✓	✓
Brief Food Frequency Measure [86]		✓	✓	✓	✓
Self-compassion [87]		✓			
EuroQol-5D-5 level HRQOL[g], Weighted Health State Utility Value [82]		✓	✓	✓	✓
HRU[h] [88]		✓		✓	✓
General program feedback (7 items created for SSBC: scored on a scale of 1=extremely dissatisfied to 7=extremely satisfied)			✓		
Program acceptability (8-item theoretical framework of acceptability [73])			✓		
Net promoter score (1-item)			✓		
Membership (1-item)			✓		
Referral (3-items)			✓		
In-person measurements					
Height (m)		✓			
Weight (lbs)		✓	✓	✓	✓
Waist circumference (cm)		✓	✓	✓	✓
Resting heart rate (BPM)		✓	✓	✓	✓
CRF[i] (6MWT[j]) [89]		✓	✓	✓	✓

[a]T2D: type 2 diabetes.

[b]ADA: American Diabetes Association.

[c]HbA_{1c}: glycated hemoglobin.

[d]GLTEQ: Godin Leisure Time Exercise Questionnaire.

[e]PAV: Physical Activity Vital Sign.

[f]SSBC: Small Steps for Big Changes.

[g]HRQOL: Health Related Quality of Life.

[h]HRU: Health Resource Utilization Questionnaire.

[i]CRF: cardiorespiratory fitness.

[j]6MWT: 6-minute walk test.

XSL•FO
RenderX

Objective 2c: Examine the Cost-Effectiveness of SSBC

The cost-effectiveness of SSBC will be determined using organizational data collected in objective 2a and client-level costs. Health-related client-level costs will be detailed using the Health Resource Utilization Questionnaire [88] delivered in a survey. The Health Resource Utilization Questionnaire asks for specific details regarding health professional visits, admissions to the hospital, laboratory tests or investigations, and use of medications. On a per-client basis, costs will be assigned to health care resource use using a fully allocated hospital cost model (for in-patient costs) and provincial guides to medical fees (for outpatient costs) every 6 months to account for potential fluctuations in medical costs. The 5 dimensions of EQ-5D-5L will be converted into a weighted health state utility value that will provide weightings for quality-adjusted life years (QALYs) and Canadian conversion tariffs for transforming health state profiles into utility scores [90]. QALYs are calculated based on the quality of life of a person (measured using health state utility values estimated from the EQ-5D-5L) in a given health state and the time spent in that health state. Incremental cost per client and effects generated by SSBC will be assessed using a Canadian health care system perspective. The outcome of cost-effectiveness analysis is the incremental cost-effectiveness ratio, where incremental cost-effectiveness ratio=Δ Cost/Δ HbA_{1c} and is the incremental cost of SSBC delivery per client. The outcome of our cost-utility analysis is the incremental cost-utility ratio, where incremental cost-utility ratio=Δ Cost/Δ QALY [91].

Sustainability Phase

Objective 3a: Cocreate Sustainability Models

In collaboration with the partnering YMCAs and affiliated municipality partners, the research team will engage in a process of codeveloping a feasible and reciprocally beneficial model of sustaining SSBC within each organization. Data from objectives 1 and 2 will be used to generate reports for each YMCA to directly inform this phase. Next, focus groups will be held at each site. Guided by the Integrated Sustainability Framework (ISF) [44], municipality-level models will establish which program components of SSBC will be continued (if at all), and at what intensity. Focus groups will also elucidate what factors result in the sustainment of evidence-based interventions in community settings. The outcome of this objective will be the development of a logic model of sustainment satisfactory to all parties.

Objective 3b: Examine SSBC Sustainability

Sustainability of SSBC will be examined at the individual client level (the number of clients who enroll and complete SSBC in the sustainability phase), individual coach level (the number of coaches that continue to deliver SSBC in the sustainability phase), and the site level (the number of sites that continue SSBC delivery in the sustainability phase; see Table 4). Potential contextual factors that influence the sustainability of SSBC will be explored through interviews with site leads and senior leadership 1 year into the sustainability phase of the project (2028) and at the end of the sustainability phase (2029). Interview guides will be developed using the ISF [44]. In addition, regional research coordinators will document any contextual factors impacting sustainability during regular meetings with site leads. All sustainability determinants will be deductively coded to the categories represented in the ISF (outer context, organizational context, processes, intervention characteristics, and implementer characteristics) by 2 independent coders. Implementation strategies used during this sustainability phase will be documented and analyzed as per the same processes used for implementation.

XSL•FO
RenderX

Table 4. Overview of sustainability outcomes and determinants, their operationalization, data sources, and measurement tools. The outcome is sustainability.

Indicators or definition	Data source	Measurement tool and time point
Site level: number of sites continuing delivery, intention to deliver the program, actual delivery of the program, and attempts to seek funds to run the program if appropriate	SL[a], Y-SL[b], and RRC[c]	Interviews with SL and Y-SL, focus groups, annual monitoring, and monthly meetings with RRC
Staff level: number of coaches continuing to deliver the program	SL and RRC	Site monthly meetings during the sustainability phase and annual reporting
Individual level: number of clients who participated in the presustainability phase and number who participated in the sustainability phase, if applicable	SL and RRC	Site monthly meetings with RRC and REDCap[d]
Sustainability determinants: contextual factors as assessed through the ISF at 24 months into implementation, after 1 year of sustainability, and end of sustainability	C[e], SL, Y-SL, MP[f], PP[g], and RRC	Interviews, focus groups, and site monthly meetings with RRC
Implementation strategies: strategies continued, discontinued, and reasons	C, SL, Y-SL, MP, PP, and RRC	Interviews, focus groups, and site monthly meetings with RRC

[a]SL: site lead.

[b]Y-SL: Young Men's Christian Association senior leadership.

[c]RRC: Regional Research Coordinator.

[d]REDCap: Research Electronic Data Capture.

[e]C: coach.

[f]MP: municipality partners.

[g]PP: patient partner.

Recruitment and Consent

SSBC Clients

Recruitment strategies will vary across delivery sites, based on the local connections previously or newly established by the research team and delivery team. A series of generic recruitment materials (eg, posters, pamphlets, or social media posts) and potential recruitment pathways (eg, health assessment days and physician referrals) have been generated by the research team and will be shared with all delivery sites. These recruitment strategies will be included in the implementation strategies' blueprint and toolkit. Each site will tailor their recruitment plan to the local contextual needs with support from the research team.

Interested or referred individuals will complete an eligibility survey to determine their eligibility for SSBC. Individuals are eligible for SSBC if they are 18 years of age or older, able to read and speak English, and meet one of the following criteria: (1) have been diagnosed with prediabetes by a physician, (2) have an HbA_{1c} of 5.7%-6.4%, (3) have an American Diabetes Association risk assessment score indicating high risk (≥5), or (4) have achieved T2D remission (have an HbA_{1c} of <6.5%) without any diabetes-related medications for a minimum of 3 months. Eligible individuals will be directed to an online consent form. Following completion, individuals will be scheduled for their first SSBC session at their local YMCA implementation site.

Individuals who express an interest in the substudy will be given a blood requisition form for an HbA_{1c} measurement from a partnering local blood requisition laboratory. The substudy will aim to oversample individuals identifying as visible minority to ensure that priority populations are represented in line with the most recent Canadian Government statistics (based and derived directly from the concept of visible minority in the census [31.3% of the Canadian population identify as being from a racialized group, with 25.2% as visible minority and 6.1% from Indigenous ancestry]). The substudy will also oversample individuals identifying as male, who are at greater risk for having prediabetes and developing T2D [6], yet are disproportionately represented in DPPs [92]. Demographic information of clients enrolling in the substudy will be assessed every 6 months to determine the proportion of individuals self-identifying as visible minorities and males. Purposive sampling may be used during the substudy if the demographic data of subsample clients demonstrates over- or underrepresentation.

SSBC Delivery Team

YMCA staff from the delivery sites will be invited to participate in this study to provide feedback on adaptation, implementation processes, and factors influencing implementation and sustainability of SSBC at their site. Site leads and coaches will be identified by senior leadership at each YMCA facility. After completing coach training, site leads and coaches will receive a consent form to participate in surveys, interviews, and focus groups. Senior leadership will receive a consent form 12 months into the program delivery, inviting them to participate in interviews and focus groups.

Project Governance

A holacratic approach to organizational operations was codeveloped with delivery sites to enhance conflict resolution and achieve project objectives. Specifically, each team member is considered integral and acknowledges their interconnected role in multiple project processes and outcomes (Figure 4).

XSL•FO
RenderX

Municipality partners (made up of patient partners, SSBC alumni, and persons with lived experience) are central to advising the project. The delivery team—the community-dwelling staff who deliver SSBC and liaise directly with researchers, senior YMCA leadership, and municipality partners—will provide ongoing feedback on referral and recruitment strategies, SSBC implementation and adaptations, and factors influencing organizational and staff readiness. YMCA senior management at each site oversees the delivery team and reports to the research team. Three committees will provide feedback and guidance on methods used, analyses performed, outputs reported, and the sustainability plan based on their expertise and experiences for sex, gender, and inclusivity; program evaluation; and the diabetes prevention engagement committee. Finally, an advisory committee will act as an external committee distinct from the project team. This committee will ensure that safe and ethical principles are applied throughout the project, project milestones are met promptly, and all decisions are made with clients and delivery organizations in mind.

Figure 4. Project governance model. SSBC: Small Steps for Big Changes; YMCA: Young Men's Christian Association.

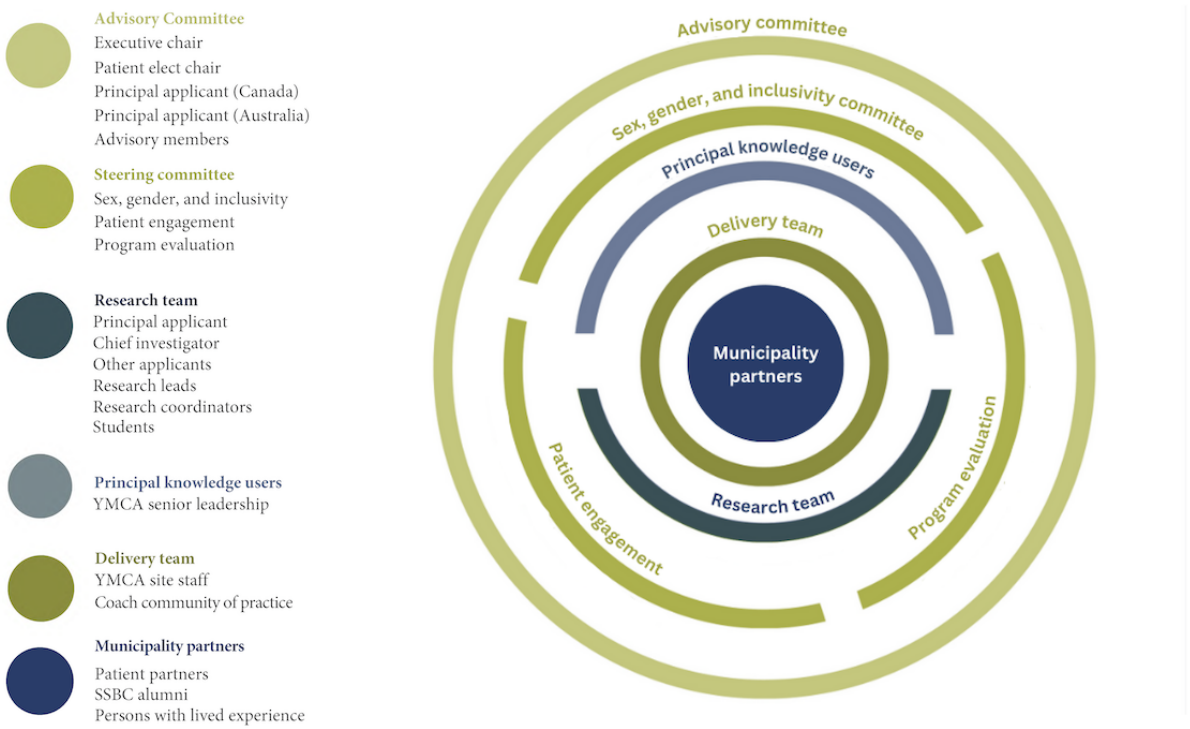

Ethical Considerations

This study has been approved by the University of British Columbia Clinical Research Ethics Board (H23-01930) and other university and provincial ethics boards (Brock University Health Science Research Ethics Board [REB-23-157] and Newfoundland and Labrador Health Research Ethics Board [HREB 20250577]). All activities will comply with the Tri-Council Policy Statement: Ethical Conduct of Research Involving Humans. Written informed consent will be obtained from all participants. Privacy and confidentiality of participant data or identity will be maintained. Consent forms will be stored separately from deidentified data. Pseudonyms will be used in interview transcripts and any reporting of the results. Participants will be compensated for their participation. Individuals who participate in the program will receive a 1-month free membership to the facility they attend. All client participants will receive a CAD $20 gift card for each of the follow-up appointments that they attend (CAD $40 total), as well as a CAD $30 gift card for each time they get their bloodwork done if they are part of the subsample (CAD $90 total). All participants who participate in an interview or focus group will receive a CAD $30 gift card. This study has been registered at ClinicalTrials.gov. The Standards for Reporting Implementation Studies checklist [93] was used to guide the reporting of this protocol paper and can be found in Multimedia Appendix 1.

Results

Program implementation within each region is occurring in a phased approach, with partners beginning program delivery in 1 site (2024-2025) before expanding to any additional locations (2025-2027). Program enrollment occurs continuously during the implementation phase across all sites. From May 2024 to November 2025, a total of 13 delivery sites began program delivery (1 regional YMCA decided to launch in 3 sites at once), 722 participants have enrolled in the program, and 406 have begun the program (153/342, 44.7% non-Western or Eastern European, 17/342, 5% of whom report identifying as Indigenous; 80/345, 23.2% men or 82/347, 23.6% male). Further, 334 participants have expressed interest in the substudy, of which 89 have had their first blood draw (19/89, 21.8% identify as a visible minority; 4/89, 5.3% indigenous; 16/82, 19.5% men or

XSL•FO
RenderX

16/81, 19.8% male). An additional 13 sites have confirmed they will launch the program in 2026.

Discussion

Principal Findings

This study will provide evidence that SSBC can be scaled up nationwide across community settings to equitably and effectively reduce population-level risk of T2D. Nationwide uptake of DPPs has been previously hindered by the absence of evidence-informed implementation plans tailored to local contexts [38]. This project will identify which implementation strategies work best for whom, where, and how, deepening our understanding of how to translate SSBC to multiple contexts. We move beyond a one-size-fits-all strategy to apply evidence-based implementation principles that will result in equitable program implementation for diverse and priority populations across Canada. This protocol paper documents our key objectives and associated methods for other implementation scientists outside of the realm of T2D prevention to consider when planning to implement evidence-based interventions at scale.

Strengths and Limitations

Sustainable DPP scale-up is a necessary condition for improved population health outcomes. This is the first study to examine how to sustain an evidence-based DPP in real-world community contexts. Unlike other DPPs that have been delivered at a national scale (eg, US National DPP [94] and English National DPP [95]), SSBC was designed to include characteristics that are key for achieving sustainability—for example, being brief, adaptable, community-based, and delivered by laypersons with no specialized training. By exploring the underlying factors underpinning the dynamic processes of implementation and sustainability when delivered at a broad scale, and engaging early with municipality partners to develop and prioritize sustainability planning to overcome identified barriers, this research will discern best practice implementation determinants, outcomes, and strategies critical for sustaining DPP implementation across Canada and beyond.

Implementation science is critical in improving health equity by ensuring health programs are accessible and effective to those most in need [40,41]. Of the participants that have enrolled thus far, we are on track for reaching an ethnically representative sample of the Canadian population, which will allow us to understand if SSBC is effective for visible minorities; however, individuals self-identifying as men and male continue to be underrepresented in SSBC participation. Establishing buy-in from health care professionals in clinical settings may help us to reach this demographic, but developing partnerships with this sector has proven challenging. Extending diabetes prevention to community settings is key to advancing health equity; however, staff shortages and turnover are ongoing challenges faced when working with community partners.

Dissemination Plan

End-of-project and integrated knowledge translation activities are planned and designed around fundamental knowledge translation principles [96-98]. To ensure broad reach, we will share our knowledge with multiple audiences using multiple formats (see Table 5 for a selected list of knowledge mobilization activities).

XSL•FO
RenderX

Table 5. Select knowledge mobilization activities for the national scale-up of Small Steps for Big Changes.

Key audiences and sectors for KT[a]	Selected sample tactics or knowledge products (outputs to audiences)	Sample tactics for knowledge input from audiences to the partnership
Delivery Team (YMCA[b] staff, senior leadership), their organizational stakeholders (donors, board of directors, and government) and the health and recreation sector	• Post resources and peer-delivered information on the project's website and social media accounts • Papers in YMCA and stakeholder magazines, newsletters, etc • Mass media press releases • Flow research and knowledge products through YMCA + stakeholders' networks (eg, present at AGMsc, in annual reports, etc) • Publish papers in open-access journals • Publish papers in trade journals • Develop and disseminate evidence-based guides, toolkits, videos, and decision aids	• Advisory committee and committee meetings • Coach community of practice (ongoing) • Site lead community of practice • Knowledge-sharing meetings (2/year) • Train the trainer website (ongoing) • Board of directors meetings
Organizations providing services to diverse communities at T2D[d] risk (eg, immigration services, indigenous friendship centers, or food banks)	• Disseminate 1-page project briefs and summary reports • Host webinars and post resources and information on the project and partners' websites and social media accounts • Present at organizations' AGMs and other events • Publication in open-access journals • Mass media press releases	• Advisory committee and committee discussions with organizations • Municipality partners committed to supporting SSBCe
Policy makers (municipality stakeholders, First Nation, and other health authorities)	• 1-page policy briefs that will be brought forward at our scheduled meetings with Government and Health Authority Partners • Reports for government ministries	• Meetings with municipality and health partners: when and how to influence policy decisions re: T2D prevention and Bill C-237
Academics and Diabetes Care Practitioners	• Peer-reviewed journals • Present at academic conferences	• Engagement with National Diabetes Framework plan and funding calls • Identify relevant and impactful journals and conferences

[a]KT: knowledge translation.

[b]YMCA: Young Men's Christian Association.

[c]AGM: Annual General Meeting.

[d]T2D: type 2 diabetes.

[e]SSBC: Small Steps for Big Changes.

Acknowledgments

The authors would like to acknowledge all Canadian and Australian Small Steps for Big Changes Urban Scale Up Collaborators. This work was conducted on the lands of the Syilx Okanagan Nation. Small Steps for Big Changes was developed on the unceded, traditional territory of the Syilx Okanagan Nation, and our team respectfully acknowledges that the program is delivered on the lands of many Indigenous Peoples across Turtle Island.

Funding

This study was funded by research grants from the Canadian Institutes of Health Research (grant #HG4 185015) and the Public Health Agency of Canada (grant #2324-HQ-000001). The funders have no role in the design, collection, analysis, and interpretation of data; in the writing of this paper; and in the decision to submit this paper for publication.

Data Availability

The future datasets generated or analyzed during this study will be made available from the corresponding author on reasonable request.

Authors' Contributions

MEJ and KW led the conceptualization of this study and its design. JEB, KC, KB, NG, and JK contributed to the conceptualization of this study's aims, procedures, and data analysis plan. MEJ, KW, JEB, and KC drafted this paper. All authors revised this paper critically for important intellectual content, read and approved the final paper.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from Healthy Cities Implementation Science (HCIS) Team Grants Review Committee, Canadian Institutes for Health Research (CIHR).
[PDF File (Adobe PDF File), 574 KB - resprot_v15i1e80276_app1.pdf]

References

1. GBD 2021 Diabetes Collaborators. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease study 2021. Lancet 2023;402(10397):203-234 [FREE Full text] [doi: 10.1016/S0140-6736(23)01301-6] [Medline: 37356446]
2. Magliano DJ, Boyko EJ. IDF Diabetes Atlas 10th edition scientific committee. In: IDF Diabetes Atlas. Brussels: International Diabetes Federation; 2021.
3. Haw JS, Galaviz KI, Straus AN, Kowalski AJ, Magee MJ, Weber MB, et al. Long-term sustainability of diabetes prevention approaches: a systematic review and meta-analysis of randomized clinical trials. JAMA Intern Med 2017;177(12):1808-1817 [FREE Full text] [doi: 10.1001/jamainternmed.2017.6040] [Medline: 29114778]
4. Schellenberg ES, Dryden DM, Vandermeer B, Ha C, Korownyk C. Lifestyle interventions for patients with and at risk for type 2 diabetes: a systematic review and meta-analysis. Ann Intern Med 2013;159(8):543-551 [FREE Full text] [doi: 10.7326/0003-4819-159-8-201310150-00007] [Medline: 24126648]
5. Uusitupa M, Khan TA, Viguiliouk E, Kahleova H, Rivellese AA, Hermansen K, et al. Prevention of type 2 diabetes by lifestyle changes: a systematic review and meta-analysis. Nutrients 2019;11(11):2611 [FREE Full text] [doi: 10.3390/nu11112611] [Medline: 31683759]
6. Diabetes in Canada: estimated prevalence and cost of diabetes. Diabetes Canada. 2024. URL: https://www.diabetes.ca/research-(1)/advocacy-reports/national-and-provincial-backgrounders/diabetes-in-canada [accessed 2025-11-21]
7. Diabetes 360: a framework for a diabetes strategy for Canada. Diabetes Canada. 2018. URL: https://www.diabetes.ca/DiabetesCanadaWebsite/media/Advocacy-and-Policy/Diabetes-360-Recommendations.pdf [accessed 2025-12-05]
8. Tabák AG, Herder C, Rathmann W, Brunner EJ, Kivimäki M. Prediabetes: a high-risk state for diabetes development. Lancet 2012;379(9833):2279-2290. [doi: 10.1016/s0140-6736(12)60283-9]
9. Key health inequalities in Canada: a national portrait - executive summary. Public Health Agency of Canada. 2018. URL: https://www.canada.ca/en/public-health/services/publications/science-research-data/key-health-inequalities-canada-national-portrait-executive-summary.html [accessed 2025-11-21]
10. Fazli GS, Moineddin R, Bierman AS, Booth GL. Ethnic differences in prediabetes incidence among immigrants to Canada: a population-based cohort study. BMC Med 2019;17(1):100 [FREE Full text] [doi: 10.1186/s12916-019-1337-2] [Medline: 31122233]
11. D'Silva C, Hafleen N, Mansfield E, Martel S, Fierheller D, Banerjee A, et al. Service provider perspectives on exploring social determinants of health impacting type 2 diabetes management for South Asian adults in peel region, Canada. Can J Diabetes 2022;46(6):611-619.e1 [FREE Full text] [doi: 10.1016/j.jcjd.2022.05.011] [Medline: 35843835]
12. Scheim AI, Coleman T, Lachowsky N, Bauer GR. Health care access among transgender and nonbinary people in Canada, 2019: a cross-sectional survey. CMAJ Open 2021;9(4):E1213-E1222 [FREE Full text] [doi: 10.9778/cmajo.20210061] [Medline: 34933879]
13. Tami A, Ferguson T, Bauer GR, Scheim AI. Avoidance of primary healthcare among transgender and non-binary people in Canada during the COVID-19 pandemic. Prev Med Rep 2022;27:101789 [FREE Full text] [doi: 10.1016/j.pmedr.2022.101789] [Medline: 35402149]
14. Fante-Coleman T, Wilson CL, Cameron R, Coleman T, Travers R. 'Getting shut down and shut out': exploring ACB patient perceptions on healthcare access at the physician-patient level in Canada. Int J Qual Stud Health Well-Being 2022;17(1):2075531 [FREE Full text] [doi: 10.1080/17482631.2022.2075531] [Medline: 35585792]
15. Socías ME, Koehoorn M, Shoveller J. Gender inequalities in access to health care among adults living in British Columbia, Canada. Womens Health Issues 2016;26(1):74-79. [doi: 10.1016/j.whi.2015.08.001] [Medline: 26384547]
16. In plain sight: addressing indigenous-specific racisim and discrimination in B.C. health care. BC Ministry of Health. 2020. URL: https://engage.gov.bc.ca/app/uploads/sites/613/2020/11/In-Plain-Sight-Summary-Report.pdf [accessed 2025-11-21]
17. Framework for diabetes in Canada. Canada. 2022. URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/framework-diabetes-canada.html [accessed 2025-11-21]
18. Knowler WC, Barrett-Connor E, Fowler SE. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002;346(6):393-403. [doi: 10.1056/nejmoa012512]
19. Pan XR, Li GW, Hu YH, Wang JX, Yang W, An Z, et al. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance: the Da Qing IGT and diabetes study. Diabetes Care 1997;20(4):537-544. [doi: 10.2337/diacare.20.4.537] [Medline: 9096977]

XSL•FO
RenderX

20. Lindström J, Louheranta A, Mannelin M, Rastas M, Salminen V, Eriksson J, Finnish Diabetes Prevention Study Group. The Finnish Diabetes Prevention Study (DPS): lifestyle intervention and 3-year results on diet and physical activity. Diabetes Care 2003;26(12):3230-3236. [doi: 10.2337/diacare.26.12.3230] [Medline: 14633807]
21. Chambers DA, Glasgow RE, Stange KC. The dynamic sustainability framework: addressing the paradox of sustainment amid ongoing change. Implement Sci 2013;8(1):117 [FREE Full text] [doi: 10.1186/1748-5908-8-117] [Medline: 24088228]
22. Lane C, McCrabb S, Nathan N, Naylor PJ, Bauman A, Milat A, et al. How effective are physical activity interventions when they are scaled-up: a systematic review. Int J Behav Nutr Phys Act 2021;18(1):16 [FREE Full text] [doi: 10.1186/s12966-021-01080-4] [Medline: 33482837]
23. Ariel-Donges AH, Gordon EL, Dixon BN, Eastman AJ, Bauman V, Ross KM, et al. Rural/urban disparities in access to the National Diabetes Prevention Program. Transl Behav Med 2020;10(6):1554-1558 [FREE Full text] [doi: 10.1093/tbm/ibz098] [Medline: 31228199]
24. Venkataramani M, Pollack CE, Yeh HC, Maruthur NM. Prevalence and correlates of diabetes prevention program referral and participation. Am J Prev Med 2019;56(3):452-457. [doi: 10.1016/j.amepre.2018.10.005] [Medline: 30661888]
25. French DP, Hawkes RE, Bower P, Cameron E. Is the NHS diabetes prevention programme intervention delivered as planned? An observational study of fidelity of intervention delivery. Ann Behav Med 2021;55(11):1104-1115 [FREE Full text] [doi: 10.1093/abm/kaaa108] [Medline: 33580647]
26. Hawkes RE, Cameron E, Miles LM, French DP. The fidelity of training in behaviour change techniques to intervention design in a National Diabetes Prevention Programme. Int J Behav Med 2021;28(6):671-682 [FREE Full text] [doi: 10.1007/s12529-021-09961-5] [Medline: 33559009]
27. Hawkes RE, Miles LM, Bower P, Cotterill S, French DP. Assessing and ensuring fidelity of the nationally implemented English NHS diabetes prevention programme: lessons learned for the implementation of large-scale behaviour change programmes. Health Psychol Behav Med 2022;10(1):498-513 [FREE Full text] [doi: 10.1080/21642850.2022.2077205] [Medline: 35646476]
28. Azar KMJ, Nasrallah C, Szwerinski NK, Petersen JJ, Halley MC, Greenwood D, et al. Implementation of a group-based diabetes prevention program within a healthcare delivery system. BMC Health Serv Res 2019;19(1):694 [FREE Full text] [doi: 10.1186/s12913-019-4569-0] [Medline: 31615525]
29. Brunisholz KD, Kim J, Savitz LA, Hashibe M, Gren LH, Hamilton S, et al. A formative evaluation of a diabetes prevention program using the RE-AIM framework in a learning health care system, Utah, 2013-2015. Prev Chronic Dis 2017;14:160556 [FREE Full text] [doi: 10.5888/pcd14.160556] [Medline: 28727546]
30. Whittemore R. A systematic review of the translational research on the diabetes prevention program. Transl Behav Med 2011;1(3):480-491 [FREE Full text] [doi: 10.1007/s13142-011-0062-y] [Medline: 24073067]
31. Venditti EM, Wylie-Rosett J, Delahanty LM, Mele L, Hoskin MA, Edelstein SL, Diabetes Prevention Program Research Group. Short and long-term lifestyle coaching approaches used to address diverse participant barriers to weight loss and physical activity adherence. Int J Behav Nutr Phys Act 2014;11(1):16 [FREE Full text] [doi: 10.1186/1479-5868-11-16] [Medline: 24521153]
32. McCurley JL, Gutierrez AP, Gallo LC. Diabetes prevention in U.S. Hispanic adults: a systematic review of culturally tailored interventions. Am J Prev Med 2017;52(4):519-529 [FREE Full text] [doi: 10.1016/j.amepre.2016.10.028] [Medline: 27989451]
33. Ritchie ND, Christoe-Frazier L, McFann KK, Havranek EP, Pereira RI. Effect of the National Diabetes Prevention Program on weight loss for English- and Spanish-speaking latinos. Am J Health Promot 2018;32(3):812-815. [doi: 10.1177/0890117117698623] [Medline: 28320212]
34. Ritchie ND, Phimphasone-Brady P, Sauder KA, Amura CR. Perceived barriers and potential solutions to engagement in the National Diabetes Prevention Program. ADCES Pract 2021;9(1):16-20 [FREE Full text] [doi: 10.1177/2633559x20966275]
35. Garst J, L'Heveder R, Siminerio LM, Motala AA, Gabbay RA, Chaney D, et al. Sustaining diabetes prevention and care interventions: a multiple case study of translational research projects. Diabetes Res Clin Pract 2017;130:67-76 [FREE Full text] [doi: 10.1016/j.diabres.2017.04.025] [Medline: 28575728]
36. Bean C, Sewell K, Jung ME. A winning combination: collaborating with stakeholders throughout the process of planning and implementing a type 2 diabetes prevention programme in the community. Health Soc Care Community 2020;28(2):681-689. [doi: 10.1111/hsc.12902] [Medline: 31795016]
37. Wurz A, Bean C, Shaikh M, Culos-Reed SN, Jung ME. From laboratory to community: three examples of moving evidence-based physical activity into practice in Canada. Health Soc Care Community 2022;30(5):e1690-e1700. [doi: 10.1111/hsc.13596] [Medline: 34623004]
38. Gruss SM, Nhim K, Gregg E, Bell M, Luman E, Albright A. Public health approaches to type 2 diabetes prevention: the US national diabetes prevention program and beyond. Curr Diab Rep 2019;19(9):78 [FREE Full text] [doi: 10.1007/s11892-019-1200-z] [Medline: 31385061]
39. Aziz Z, Mathews E, Absetz P, Sathish T, Oldroyd J, Balachandran S, et al. A group-based lifestyle intervention for diabetes prevention in low- and middle-income country: implementation evaluation of the Kerala Diabetes Prevention Program. Implement Sci 2018;13(1):97 [FREE Full text] [doi: 10.1186/s13012-018-0791-0] [Medline: 30021592]

XSL•FO
RenderX

40. Brownson RC, Kumanyika SK, Kreuter MW, Haire-Joshu D. Implementation science should give higher priority to health equity. Implement Sci 2021;16(1):28 [FREE Full text] [doi: 10.1186/s13012-021-01097-0] [Medline: 33740999]
41. Baumann AA, Cabassa LJ. Reframing implementation science to address inequities in healthcare delivery. BMC Health Serv Res 2020;20(1):190 [FREE Full text] [doi: 10.1186/s12913-020-4975-3] [Medline: 32164706]
42. Durlak JA, DuPre EP. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. Am J Community Psychol 2008;41(3-4):327-350. [doi: 10.1007/s10464-008-9165-0] [Medline: 18322790]
43. Proctor EN, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. Implement Sci 2013;8:139 [FREE Full text] [doi: 10.1186/1748-5908-8-139] [Medline: 24289295]
44. Shelton RC, Cooper BR, Stirman SW. The sustainability of evidence-based interventions and practices in public health and health care. Annu Rev Public Health 2018;39:55-76 [FREE Full text] [doi: 10.1146/annurev-publhealth-040617-014731] [Medline: 29328872]
45. Proctor E, Luke D, Calhoun A, McMillen C, Brownson R, McCrary S, et al. Sustainability of evidence-based healthcare: research agenda, methodological advances, and infrastructure support. Implement Sci 2015;10(1):88 [FREE Full text] [doi: 10.1186/s13012-015-0274-5] [Medline: 26062907]
46. Glasgow RE, Chambers D. Developing robust, sustainable, implementation systems using rigorous, rapid and relevant science. Clin Transl Sci 2012;5(1):48-55 [FREE Full text] [doi: 10.1111/j.1752-8062.2011.00383.x] [Medline: 22376257]
47. Scheirer MA, Dearing JW. An agenda for research on the sustainability of public health programs. Am J Public Health 2011;101(11):2059-2067. [doi: 10.2105/ajph.2011.300193]
48. Shediac-Rizkallah MC, Bone LR. Planning for the sustainability of community-based health programs: conceptual frameworks and future directions for research, practice and policy. Health Educ Res 1998;13(1):87-108. [doi: 10.1093/her/13.1.87] [Medline: 10178339]
49. Wareham NJ, Herman WH. The clinical and public health challenges of diabetes prevention: a search for sustainable solutions. PLoS Med 2016;13(7):e1002097 [FREE Full text] [doi: 10.1371/journal.pmed.1002097] [Medline: 27458730]
50. Dineen TE, Bean C, Jung ME. Implementation of a diabetes prevention program within two community sites: a qualitative assessment. Implement Sci Commun 2022;3(1):11 [FREE Full text] [doi: 10.1186/s43058-022-00258-6] [Medline: 35123582]
51. Dineen TE, Banser T, Bean C, Jung ME. Fitness facility staff demonstrate high fidelity when implementing an evidence-based diabetes prevention program. Transl Behav Med 2021;11(10):1814-1822. [doi: 10.1093/tbm/ibab039] [Medline: 33963868]
52. Bean C, Dineen T, Locke SR, Bouvier B, Jung ME. An evaluation of the reach and effectiveness of a diabetes prevention behaviour change program situated in a community site. Can J Diabetes 2021;45(4):360-368. [doi: 10.1016/j.jcjd.2020.10.006] [Medline: 33323314]
53. Dineen TE, Bean C, Jung ME. Successes and challenges from a motivational interviewing-informed diabetes prevention program situated in the community. Health Promot Pract 2024;25(2):274-284 [FREE Full text] [doi: 10.1177/15248399221115066] [Medline: 36047459]
54. Cranston KD, MacPherson MM, Sim JA, Jung ME. Small steps towards an inclusive diabetes prevention program: how small steps for big changes is improving program equity and inclusion. Community Health Equity Res Policy 2023;45(3):227-236. [doi: 10.1177/2752535x231189932]
55. MacPherson M, Merry K, Locke S, Jung M. Developing mobile health interventions with implementation in mind: application of the multiphase optimization strategy (MOST) preparation phase to diabetes prevention programming. JMIR Form Res 2022;6(4):e36143 [FREE Full text] [doi: 10.2196/36143] [Medline: 35471473]
56. McGavock J, Chauhan BF, Rabbani R, Dias S, Klaprat N, Boissoneault S, et al. Layperson-led vs professional-led behavioral interventions for weight loss in pediatric obesity: a systematic review and meta-analysis. JAMA Netw Open 2020;3(7):e2010364 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.10364] [Medline: 32658289]
57. Cranston K, Ivanova E, Davis C, Jung M. Fidelity and maintenance of motivational interviewing skills in diabetes prevention program coaches: a pilot study Internet. Research Square Preprint posted online on November 2, 2021 [FREE Full text] [doi: 10.21203/rs.3.rs-1025262/v1]
58. Jung ME, Bourne JE, Beauchamp MR, Robinson E, Little JP. High-intensity interval training as an efficacious alternative to moderate-intensity continuous training for adults with prediabetes. J Diabetes Res 2015;2015:191595 [FREE Full text] [doi: 10.1155/2015/191595] [Medline: 25918728]
59. Locke SR, Bourne JE, Beauchamp MR, Little JP, Barry J, Singer J, et al. High-intensity interval or continuous moderate exercise: a 24-week pilot trial. Med Sci Sports Exercise 2018;50(10):2067-2075. [doi: 10.1249/MSS.0000000000001668] [Medline: 29762252]
60. Jung ME, Locke SR, Bourne JE, Beauchamp MR, Lee T, Singer J, et al. Cardiorespiratory fitness and accelerometer-determined physical activity following one year of free-living high-intensity interval training and moderate-intensity continuous training: a randomized trial. Int J Behav Nutr Phys Act 2020;17(1):25 [FREE Full text] [doi: 10.1186/s12966-020-00933-8] [Medline: 32102667]
61. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]

XSL•FO
RenderX

62. MacPherson MM, Dineen TE, Cranston KD, Jung ME. Identifying behaviour change techniques and motivational interviewing techniques in small steps for big changes: a community-based program for adults at risk for type 2 diabetes. Can J Diabetes 2020;44(8):719-726. [doi: 10.1016/j.jcjd.2020.09.011] [Medline: 33279096]
63. Cranston KD, Grieve NJ, Dineen TE, Jung ME. Designing and developing online training for diabetes prevention program coaches using an integrated knowledge translation approach: development and usability study. JMIR Form Res 2024;8(1):e50942 [FREE Full text] [doi: 10.2196/50942] [Medline: 38277214]
64. Grieve NJ, Cranston KD, Jung ME. Examining the effectiveness of an e-learning training course for coaches of a type 2 diabetes prevention program. J Technol Behav Sci 2024;9(2):180-190 [FREE Full text] [doi: 10.1007/s41347-023-00316-3] [Medline: 37362065]
65. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
66. Wiltsey Stirman S, Baumann AA, Miller CJ. The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. Implement Sci 2019;14(1):58 [FREE Full text] [doi: 10.1186/s13012-019-0898-y] [Medline: 31171014]
67. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
68. Waltz TJ, Powell BJ, Fernández ME, Abadie B, Damschroder LJ. Choosing implementation strategies to address contextual barriers: diversity in recommendations and future directions. Implement Sci 2019;14(1):42 [FREE Full text] [doi: 10.1186/s13012-019-0892-4] [Medline: 31036028]
69. Miller CJ, Barnett ML, Baumann AA, Gutner CA, Wiltsey-Stirman S. The FRAME-IS: a framework for documenting modifications to implementation strategies in healthcare. Implement Sci 2021;16(1):36 [FREE Full text] [doi: 10.1186/s13012-021-01105-3] [Medline: 33827716]
70. Shea CM, Jacobs SR, Esserman DA, Bruce K, Weiner BJ. Organizational readiness for implementing change: a psychometric assessment of a new measure. Implement Sci 2014;9(1):7 [FREE Full text] [doi: 10.1186/1748-5908-9-7] [Medline: 24410955]
71. McLean KA. Healthcare provider acceptability of a behavioral intervention to promote adherence - University of Miami. University of Miami. 2013. URL: https://scholarship.miami.edu/esploro/outputs/graduate/Healthcare-Provider-Acceptability-of-a-Behavioral/991031447949202976 [accessed 2025-11-21]
72. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
73. Sekhon M, Cartwright M, Francis JJ. Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. BMC Health Serv Res 2022;22(1):279 [FREE Full text] [doi: 10.1186/s12913-022-07577-3] [Medline: 35232455]
74. Net Promoter Score. Project Management Institute. 2021. URL: https://dabrowser.pmi.org/?_gl=1*ns9hxa*_gcl_au*MTE0NDY4OTMxNC4xNzY0NzIwODk4#item:-2004084846 [accessed 2025-11-21]
75. Vossen J, Burduli E, Barbosa-Leiker C. MICA Manual. 2017. URL: https://micacoding.com/validity-testing [accessed 2025-11-22]
76. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the expert recommendations for implementing change (ERIC) project. Implement Sci 2015;10(1):21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
77. Braun V, Clarke V, Hayfield N, Terry G. Handbook of research methods in health social sciences. In: Thematic Analysis. Singapore: Springer; 2019:843-860.
78. Tracy SJ. Qualitative quality: eight “big-tent” criteria for excellent qualitative research. Qual Inq 2010;16(10):837-851. [doi: 10.1177/1077800410383121]
79. Rikli RE, Jones CJ. The reliability and validity of a 6-minute walk test as a measure of physical endurance in older adults. J Aging Phys Act 1998;6(4):363-375 [FREE Full text] [doi: 10.1123/japa.6.4.363]
80. Comino EJ, Tran DT, Haas M, Flack J, Jalaludin B, Jorm L, et al. Validating self-report of diabetes use by participants in the 45 and up study: a record linkage study. BMC Health Serv Res 2013;13(1):481 [FREE Full text] [doi: 10.1186/1472-6963-13-481] [Medline: 24245780]
81. Margolis KL, Lihong Q, Brzyski R, Bonds DE, Howard BV, Kempainen S, Women Health Initiative Investigators. Validity of diabetes self-reports in the Women's Health Initiative: comparison with medication inventories and fasting glucose measurements. Clin Trials 2008;5(3):240-247 [FREE Full text] [doi: 10.1177/1740774508091749] [Medline: 18559413]
82. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
83. Godin G, Shephard RJ. Leisure-time exercise questionnaire. APA PsycTests. 1985. URL: https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft31334-000 [accessed 2025-11-21]

XSL•FO
RenderX

84. Physical activity vital sign. American College of Sports Medicine. 2021. URL: https://exerciseismedicine.org/wp-content/uploads/2021/04/EIM-Physical-Activity-Vital-Sign.pdf [accessed 2025-11-21]
85. Ross R, Chaput JP, Giangregorio LM, Janssen I, Saunders TJ, Kho ME, et al. Canadian 24-hour movement guidelines for adults aged 18-64 years and adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and sleep. Appl Physiol Nutr Metab 2020;45(Suppl 2):S57-S102 [FREE Full text] [doi: 10.1139/apnm-2020-0467] [Medline: 33054332]
86. Block G, Azar KMJ, Romanelli RJ, Block TJ, Palaniappan LP, Dolginsky M, et al. Improving diet, activity and wellness in adults at risk of diabetes: randomized controlled trial. Nutr Diabetes 2016;6(9):e231-e231 [FREE Full text] [doi: 10.1038/nutd.2016.42] [Medline: 27643726]
87. Raes F, Pommier E, Neff KD, Van Gucht D. Construction and factorial validation of a short form of the Self-Compassion Scale. Clin Psychol Psychother 2011;18(3):250-255. [doi: 10.1002/cpp.702] [Medline: 21584907]
88. Maetzel A, Li LC, Pencharz J, Tomlinson G, Bombardier C. The economic burden associated with osteoarthritis, rheumatoid arthritis, and hypertension: a comparative study. Ann Rheum Dis 2004;63(4):395-401 [FREE Full text] [doi: 10.1136/ard.2003.006031] [Medline: 15020333]
89. Guyatt GH, Sullivan MJ, Thompson PJ, Fallen EL, Pugsley SO, Taylor DW, et al. The 6-minute walk: a new measure of exercise capacity in patients with chronic heart failure. Can Med Assoc J 1985;132(8):919-923 [FREE Full text] [Medline: 3978515]
90. Bansback N, Tsuchiya A, Brazier J, Anis A. Canadian valuation of EQ-5D health states: preliminary value set and considerations for future valuation studies. PLoS One 2012;7(2):e31115 [FREE Full text] [doi: 10.1371/journal.pone.0031115] [Medline: 22328929]
91. Drummond M, Manca A, Sculpher M. Increasing the generalizability of economic evaluations: Recommendations for the design, analysis, and reporting of studies. Int J Technol Assess Health Care 2005;21(2):165-171. [doi: 10.1017/s0266462305050221]
92. Jackson MC, Dai S, Skeete RA, Owens-Gary M, Cannon MJ, Smith BD, et al. An examination of gender differences in the National Diabetes Prevention Program's lifestyle change program. Diabetes Educ 2020;46(6):580-586 [FREE Full text] [doi: 10.1177/0145721720964585] [Medline: 33063641]
93. Pinnock H, Barwick M, Carpenter CR, Eldridge S, Grandes G, Griffiths CJ, et al. Standards for Reporting Implementation Studies (StaRI) statement. BMJ 2017;356:i6795 [FREE Full text] [doi: 10.1136/bmj.i6795] [Medline: 28264797]
94. Ely EK, Gruss SM, Luman ET, Gregg EW, Ali MK, Nhim K, et al. A national effort to prevent type 2 diabetes: participant-level evaluation of CDC's national diabetes prevention program. Diabetes Care 2017;40(10):1331-1341 [FREE Full text] [doi: 10.2337/dc16-2099] [Medline: 28500215]
95. Valabhji J, Barron E, Bradley D, Bakhai C, Fagg J, O'Neill S, et al. Early outcomes from the English National Health Service Diabetes Prevention Programme. Diabetes Care 2020;43(1):152-160 [FREE Full text] [doi: 10.2337/dc19-1425] [Medline: 31719054]
96. Lavis JN, Robertson D, Woodside JM, McLeod CB, Abelson J, Knowledge Transfer Study Group. How can research organizations more effectively transfer research knowledge to decision makers? Milbank Q 2003;81(2):221-248 [FREE Full text] [doi: 10.1111/1468-0009.t01-1-00052] [Medline: 12841049]
97. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
98. Grimshaw JM, Eccles MP, Lavis JN, Hill SJ, Squires JE. Knowledge translation of research findings. Implement Sci 2012;7(1):50 [FREE Full text] [doi: 10.1186/1748-5908-7-50] [Medline: 22651257]

Abbreviations

BC: British Columbia
CFIR: Consolidated Framework for Implementation Research
DPP: diabetes prevention program
FRAME: Framework for Reporting Adaptations and Modifications-Expanded
FRAME-IS: Framework for Reporting Adaptations and Modifications for Implementation Strategies
HbA1c: glycated hemoglobin
ISF: Integrated Sustainability Framework
PA: physical activity
QALY: quality-adjusted life year
REDCap: Research Electronic Data Capture
SSBC: Small Steps for Big Changes
T2D: type 2 diabetes
YMCA: Young Men's Christian Association

XSL•FO
RenderX

Edited by A Schwartz; submitted 07.Jul.2025; this is a non–peer-reviewed article;accepted 14.Nov.2025; published 19.Jan.2026.

Please cite as:
Weatherson K, Bourne JE, Cranston K, Braaten K, Grieve N, Kelly J, Jung ME
Scaling Up a Diabetes Prevention Program in Geographically and Ethnoculturally Diverse Urban Regions of Canada: Protocol for a Hybrid Type 2 Implementation-Effectiveness Study
JMIR Res Protoc 2026;15:e80276
URL: https://www.researchprotocols.org/2026/1/e80276
doi:10.2196/80276
PMID:

©Katie Weatherson, Jessica E Bourne, Kaela Cranston, Kyra Braaten, Natalie Grieve, Joseph Kelly, Mary E Jung. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Context-Adapted Diabetes Prevention Program (Small Steps for Big Changes) in Australia: Protocol for a Hybrid Implementation-Effectiveness Study

Genevieve N Healy[1], PhD; Sjaan R Gomersall[1,2], PhD; David W Dunstan[3,4], PhD; Elisabeth A H Winkler[1], PhD; Katherine A Heseltine[1], BCExP (Hons); Grace Q Sim[5], MHSM; Fiona Mason[6], BPT; Denis Giguere[6], BPhty; Michael Tsiamis[7], BA; Mary E Jung[8], PhD; Ana D Goode[1], PhD

[1]Health and Wellbeing Centre for Research Innovation, School of Human Movement and Nutrition Sciences, The University of Queensland, Brisbane, Queensland, Australia

[2]School of Health and Rehabilitation Sciences, The University of Queensland, Brisbane, Australia

[3]Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Burwood, BC, Australia

[4]Physical Activity Laboratory, Baker Heart and Diabetes Institute, Melbourne, Victoria, Australia

[5]School of Human Movement and Nutrition Sciences, The University of Queensland, Brisbane, Australia

[6]Logan Healthy Living, UQ Health Care, Logan, Australia

[7]The Y Queensland, Brisbane, Australia

[8]School of Health and Exercise Sciences, Faculty of Health and Social Development, University of British Columbia, Kelowna, BC, Canada

Corresponding Author:
Genevieve N Healy, PhD
Health and Wellbeing Centre for Research Innovation
School of Human Movement and Nutrition Sciences
The University of Queensland
Building 26B
Brisbane, Queensland, 4072
Australia
Phone: 61 33656240
Email: g.healy@uq.edu.au

Abstract

Background: With type 2 diabetes rates escalating worldwide, including in Australia, effective, acceptable, sustainable, and scalable diabetes prevention programs are needed. Small Steps for Big Changes (SSBC) is a Canadian-developed, community-delivered diet and exercise counseling intervention for individuals at risk of type 2 diabetes. The 3- to 6-week intervention can be delivered by trained non–health professionals, with all SSBC coaches receiving training in the delivery of the program, including motivational interviewing. However, the suitability and feasibility of the program in the Australian context are unknown. To address this gap, funding has been secured to adapt, implement, and evaluate SSBC in Australia (SSBC Australia), providing evidence on its effectiveness, acceptability, and implementation in this context.

Objective: The aim of the study is to describe the protocol for the type 2 cluster nonrandomized single-arm hybrid effectiveness-implementation trial of SSBC Australia.

Methods: SSBC Australia will be delivered and evaluated in 5 community-based sites across 2 organizations in South-East Queensland, Australia. One organization (1 site) will trial training students on clinical and project placements as coaches. The evaluation period is 4 years. For the first 2 years, sites receive funding for program delivery, after which, ongoing delivery will be self-funded. The recruitment target is 500 participants completing the 6-session intervention across the 5 sites within 2 years, with approximately 50 coaches trained. Data will be collected from the organization, site, coach, and client using a variety of methods (surveys, objective assessments, interviews, site audits, website analytics, meeting minutes, and project tracking). The integrated Practical, Robust, Implementation Sustainability Model and Reach, Effectiveness, Adoption, Implementation, Maintenance framework and the Affordability, Practicality, Effectiveness and Cost-Effectiveness, Acceptability, Side-Effects/Safety, and Equity criteria will guide implementation and evaluation and inform iterative adaptations as required. Data will be collected on the context for delivery; adoption and reach (number of coaches and clients and their characteristics); effectiveness of the coach training and the intervention (client pre- and post changes in measured clinical indicators [body composition, cardiorespiratory

XSL•FO
RenderX

fitness, strength, and balance], self-reported health behaviors [movement behaviors, healthy eating, and program behaviors], psychosocial indicators [self-efficacy and social supports], quality of life, diabetes status, and health care use); implementation of the coach training and program delivery (fidelity and acceptability); and maintenance of program delivery (sites) and client outcomes at 3, 6, and 12 months post end-of-program.

Results: Ethics approval and trial registration were completed. All 5 sites have been recruited and undergone preimplementation evaluation, with SSBC Australia coach training underway. Client recruitment started in September 2025.

Conclusions: This study will provide evidence whether the contextually adapted SSBC diabetes prevention program can be successfully implemented and is effective within an Australian context. Findings will inform potential expansion to additional community sites and health service contexts.

Trial Registration: Australian and New Zealand Clinical Trials Registry ACTRN12624001194550; https://tinyurl.com/588jkmva

International Registered Report Identifier (IRRID): DERR1-10.2196/81195

(JMIR Res Protoc 2026;15:e81195) doi:10.2196/81195

KEYWORDS

diabetes prevention; community; exercise; diet; motivational interviewing; cultural adaptation; implementation; intervention

Introduction

Type 2 diabetes is a global epidemic: in 2021, an estimated 529 million adults had diabetes with projections that this will reach 1.31 billion by 2050 [1]. Given the significant health, social, and financial costs of diabetes (eg, diabetes costs an estimated US $2.25 billion to the Australian health care system in 2020-2021 [2]), there has been significant investment in understanding how to prevent and reduce the risk of developing this chronic condition. The pivotal US Diabetes Prevention Program demonstrated that a lifestyle modification intervention involving physical activity and dietary changes was effective at reducing progression to type 2 diabetes in those at elevated risk [3], with these effects sustained long-term [4]. Diabetes prevention programs have now been adapted for community-based delivery across multiple communities, with demonstrated cost-effectiveness [5], and effectiveness robust to various cultural adaptations and translational strategies [6]. However, the length and intensity of these programs, which typically require at least a 12-month commitment, impact on their reach, retention, and maintenance [7,8]. Lower intensity programs (ie, shorter duration and fewer sessions) may be more attractive to potential participants and align with calls to action to reconsider how diabetes prevention is approached [9]; however, the evidence for these programs is currently relatively limited [10]. One such lower-intensity diabetes prevention program is Small Steps for Big Changes (SSBC).

SSBC is an evidence- and community-based diet and exercise counseling intervention for individuals at risk of type 2 diabetes [11-15]. The program, developed by Jung et al [15] in Canada, was designed for feasible and sustainable translation into communities nationally [11,16] and enhanced to assist in providing equitable access and inclusive care [14], with the development, efficacy testing, and scale-up of the SSBC program within the Canadian context described in detail elsewhere [11,12,14,16,17]. The intervention was based on social cognitive theory [18] and behavior change techniques specifically known for enhancing adherence to diet and exercise modification [13]. Mapping to the behavior change taxonomy identified 43 behavior change techniques used as part of the SSBC program [12]. The implementation strategies of SSBC have been iteratively tested as part of the research-to-practice and scale-up process.

A key implementation strategy tested was the use of lay health coaches to deliver the program. Here, coaches attended a 3-day in-person workshop where they were trained in motivational interviewing (MI) as well as the diet, exercise, and diabetes content of the program. This strategy was successfully tested with a community-based partner (YMCA) [16,19-21]. To further enhance scalability, this training was adapted to be delivered online using an integrated knowledge translation approach [22] and shown to be successful and acceptable to coaches [11,21]. The coaches deliver the 6-session program within community fitness facilities (eg, YMCAs). Each session lasts approximately 60-75 minutes and is comprised of approximately 40 minutes of MI-informed one-on-one counseling on diet and exercise behavior and 20-30 minutes of structured exercise at either a moderate or high intensity. The ability for community members and previous participants to take on the role as coach distinguishes SSBC from other diabetes prevention programs that typically require delivery by health professionals. Findings to date have demonstrated that this delivery approach is feasible, acceptable, and effective [11,16,23]. Based on the success of these findings, funding was granted to Jung et al [24] to expand the delivery of the SSBC program in YMCA sites across Canada, including in areas with low socioeconomic status and ethnoculturally diverse populations. Details on the implementation strategies used in the promotion, implementation, and evaluation of the Canadian delivery of the SSBC program, and the associated measurement of these strategies, will be reported in detail elsewhere [25].

Funding support was also granted to Healy et al [26] to understand the suitability, implementation, and effectiveness of the program within an Australian context. Like other countries, diabetes is a public health concern in Australia, with around 125 people a day diagnosed with type 2 diabetes and almost 1.2 million people living with the condition [2]. These rates vary by socioeconomic status, with age-standardized rates of type 2 diabetes in Australia almost twice as high in those living in the lowest socioeconomic areas (5.1%) compared to

XSL•FO
RenderX

the highest socioeconomic areas (2.6%) [2]. Importantly, 1 in 6 Australian people have prediabetes: a condition in which blood glucose levels are higher than normal but not high enough to be classified as diabetes [27]. Notably, despite the position statement highlighting the need for screening of prediabetes [28], Australia does not yet have a national screening program for diabetes, so rates—particularly of prediabetes—may be an underestimate.

Australia offers a range of free diabetes prevention programs [29], with most provided at a state level and funded through government and Diabetes Australia or state-based partners. Various delivery modes are used, including telephone, online, and in-person, with most programs delivered by expert health professionals. In the state of Queensland, where SSBC in Australia (SSBC Australia) will be delivered, the freely available programs are My Health for Life [30] and Beat It [31]. My Health for Life is a government-funded 18-week program delivered over 6 sessions by a qualified health professional [30], while Beat It is a free 8-week program delivered by exercise professionals [31]. The Canadian-developed SSBC may provide an important addition to these offerings by providing an in-person program that does not require a health professional for delivery, extending the potential reach of the program into communities where such services are not readily available or where traditional diabetes prevention programs may not be widely accepted. The ability for SSBC to be delivered by a non–health professional may also increase the sustainability of the program.

Similar to Canada, the Australian implementation of SSBC (SSBC Australia) will have a delivery arm through the Queensland YMCA community sites (rebranded to The Y Queensland), with sites identified based on need and demand for such services [32], as well as providing a mix of regional and urban settings. Understanding the potential for delivery of SSBC Australia in regional settings is particularly important, given the geographical size of Queensland (1.727 million square kilometers, larger than all but 16 countries) and the population spread (more than half of the population lives outside the capital city area) [33] can make access to health care professionals to deliver diabetes prevention programs difficult. A community health care service, Logan Healthy Living by UQ Health Care, will also be included as a delivery site. UQ Health Care is a not-for-profit controlled entity of the University of Queensland. Logan Healthy Living is an interprofessional clinic with a range of allied health services that specializes in providing evidence-based interventions to manage and improve outcomes for individuals with various chronic diseases including those with and at risk of type 2 diabetes [34]. It uses a student-infused work-integrated learning placement model, with students from across multiple allied health professions and from both undergraduate and postgraduate programs working together to provide an interdisciplinary model of care [35]. One novel element of the SSBC Australia delivery will be trialing these placement students as SSBC coaches. In addition to providing a potentially sustainable delivery model, this also provides a mechanism to build the knowledge, skills, and confidence of these preprofessional students and enhance workforce capacity in lifestyle management interventions.

The Australian delivery of the SSBC program has been informed by a program of exploratory preimplementation work. This included a needs assessment within the community, preliminary discussions with the sites, and exploration of the appropriateness of the Canadian-developed SSBC coach training for the Australian context. The needs assessment was conducted via a cross-sectional survey with 575 community members of The Y (formerly known as The YMCA) across 20 Y Queensland sites [32]. Findings showed that there was both a need and a desire for a diabetes prevention program, with 46% of participants showing elevated risk of developing type 2 diabetes in the next 5 years according to the Australian Type 2 Diabetes Risk Assessment Tool (AUSDRISK) score [36], and 68% of participants expressing interest in taking part in such a program [32].

The preliminary site-level discussions involved interviews with site leads and other key staff members at each of the sites, with the Consolidated Framework for Implementation Research (CFIR) [37] used to guide the semistructured interview questions. An initial interview was conducted online, followed up with interviews on-site at each of the 5 sites. Findings highlighted that there was variation across sites in terms of the equipment available (both for assessment and exercise), services offered, staffing availability, and experience in implementing research projects and working with university partners. All sites catered to an aging population (appropriate for the SSBC intervention), were focused on serving the communities in which they worked, and had established communication channels available for program promotion. Findings confirmed that the sites identified for delivery were suitable, but site-specific protocols for implementation would be required.

The appropriateness of the online SSBC coach training for the Australian context was also explored, with findings reported in detail elsewhere [38]. In brief, focus groups and interviews, along with “think-alouds” of the online training experience, were conducted with a range of potential future coaches of SSBC Australia. These included current site staff, tertiary students, and allied health professionals and students who identify as Aboriginal or Torres Strait Islander. Overall, the SSBC program and the online coach training were well received and were seen as appropriate and feasible by all stakeholders. Key content modification suggestions were to reference the Australian physical activity and dietary guidelines (rather than the Canadian guidelines) and enhance the cultural safety components to be suitable for the Australian context. Key delivery modification suggestions were to allow strength training (which is now available as an augmented addition in the Canadian delivery) and provide opportunities to practice in a face-to-face setting prior to the mock session. These modifications were developed and tested with potential end users [39] and then incorporated into the coach training process and material for Australian delivery. The aim of this paper is to describe the protocol for the implementation and evaluation of SSBC Australia.

XSL•FO
RenderX

Methods

Study Design

In line with the SSBC Canada evaluation, SSBC Australia will be assessed via a type 2 hybrid effectiveness-implementation cluster nonrandomized single-arm trial, where the primary effectiveness outcome will be end-of-program change in self-reported moderate- to vigorous-intensity physical activity (MVPA) behavior, and the primary implementation outcome will be fidelity of program delivery within the provider settings. The study design with respect to establishing effectiveness is a cluster single-arm pre- and posttrial. The evaluation period is 4 years; however, research funding for site-level implementation is provided for the first 2 years only. The evaluation is facilitated by the research team across the 4 years. The timeline for adaption, delivery, and evaluation is shown in Table 1.

Table 1. Timeline of the project.

	Year 1	Year 2	Year 3	Year 4	Year 5
Site-level adaptations	✓				
Research-funded delivery		✓	✓		
Site-funded delivery				✓	✓
Research evaluation		✓	✓	✓	✓

Frameworks for Evaluation and Reporting

The integrated Practical, Robust, Implementation Sustainability Model (PRISM) and Reach, Effectiveness, Adoption, Implementation, Maintenance (RE-AIM) framework [40] and the Affordability, Practicality, Effectiveness and Cost-Effectiveness, Acceptability, Side-Effects/Safety, and Equity (APEASE) criteria [41] will be used to guide the implementation and evaluation, assess progress over time, and make modifications as required. The PRISM and RE-AIM frameworks incorporate evaluation of contextual factors, implementation strategies, and RE-AIM outcomes. It has been widely used to understand the implementation of evidence-based programs into practice, including diabetes prevention programs [42]. Learnings from the Canadian implementation [25,43] as well as unique barriers and enablers within the Australian context will inform the SSBC Australia implementation strategies. In parallel and in collaboration with the Canadian SSBC research project [25], the SSBC Australia project will use the Expert Recommendations for Implementing Change taxonomy [44] and Proctor's recommendations for specifying and reporting [45] to map, name, and define the implementation strategies for SSBC Australia, with this information to be reported in detail elsewhere. The Framework for Reporting Adaptations and Modifications to Evidence-Based Implementation Strategies [46] will be used to track adaptations to the implementation strategies. The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) guidelines for reporting outcomes in trial protocols [47] and the StaRI (Standards for Reporting Implementation Studies) checklist [48] are used to guide reporting (Multimedia Appendix 1).

Ethical Considerations

The study has received ethics approval from The University of Queensland Human Research Committee for the adaptation to the Australian context (#HE001455, #HE000500, and #HE001547), understanding of context and barriers and enablers to implementation from a site and provider perspective (#HE001455), and program delivery and evaluation (#HE001123 and #HE001280). All participants are provided with an electronic participant information and consent sheet prior to participation, where it is emphasized that participation in the research is voluntary. Participants can withdraw at any time without consequence; however, data collected up to that point may be used in analyses. Data are identifiable when collected and deidentified for analyses. Client participants are provided with up to 6 weeks of membership at their SSBC Australia site (with this paid for by the delivery site) and provided the program for free. Sites are provided with funding from the grant to deliver the program for the first 2 years.

Aims of the SSBC Australia Trial

The overall aim of the trial is to evaluate the contextually adapted version of the Canadian-developed SSBC program (SSBC Australia). Specifically, we aim to evaluate the uptake (adoption) by coaches; reach of the program; program implementation, including program acceptability (at the provider site, coach, and client level) and contextual factors influencing implementation; program effectiveness; and sustainability of the program, including delivery costs and ongoing delivery.

SSBC Australia Components

Implementation of the Canadian-developed SSBC program within the Australian context will involve four key elements: (1) understanding site-level requirements (site adaptations), (2) SSBC Australia coach training, (3) delivery of the program by site (delivery funded by research), and (4) delivery of the program by site (delivery funded by site).

Site Adaptations

In recognition that resource availability, workflows, and processes may differ between sites, understanding site adaptation requirements is built into the preimplementation phase process, with these processes designed and developed by SSBC Canada. A core part of this process is a preimplementation workshop. This workshop is to understand what site-level adaptations are needed as well as to clarify roles and responsibilities. This workshop is intended to build and extend on preliminary conversations with each site. Prior to the workshop, site leads will be sent a preworkshop checklist on elements to reflect on or ascertain prior to the workshop to help facilitate decision-making. The workshop will be recorded, and field

XSL•FO
RenderX

notes taken and summarized. This summary will be shared with the site team and modified as required to appropriately reflect the discussions and decisions, both during the workshop and through other correspondence. Together, this information will be used to develop a site-level standard operating protocol and inform site-level implementation strategies.

SSBC Australia Coach Training

Overview

The SSBC Australia coach training consists of 4 main elements: the Canadian training modules, the Australian-specific modules and resources, the measurement assessment and exercise supervision training, and the competency evaluation. The Canadian and Australian modules are housed on an online training platform from the 3C Institute (Durham, North Carolina). A separate user journey is available for the Australian coaches so that they only access resources appropriate for SSBC Australia. A REDCap (Research Electronic Data Capture) coach training database was also created specifically for SSBC Australia (described in the SSBC Australia REDCap Coach Training Database section). Discussions around the training process will be held as part of the site preimplementation workshop, with any adaptions or additions to the training protocol (eg, the use of peers to practice skills prior to competency testing) for coaches recorded. In addition to any site-level variations, variations at an individual coach level will also be recorded via REDCap and asked via interviews.

Online Canadian Modules

Coaches will complete the Canadian modules as the first part of their SSBC training process. The online training consists of 7 modules, including a welcome and overview of the program, an overview of MI, and how to use the program resources, with each module including embedded quizzes [22]. The Canadian online training modules take approximately 3 to 4 hours to complete.

Australian-Specific Modules

The SSBC Australia coaches will then be asked to review 4 Australian-specific modules as part of their training process as well as the SSBC Australia client workbook (Table 2). These modules were designed to address the changes required to ensure that the program was suitable for delivery within the Australian context as well as provide training for the additional evaluation components that will be used in the Australian trial. An optional module is also provided, which provides further information on MI to complement the resources provided in the Canadian online training. Collectively, the Australian modules are expected to take approximately 2 hours to complete. Coaches can access these modules at any point via the online resource hub. This mechanism also provides flexibility to add additional training content if identified as part of the implementation process. Any additions will be tracked by the project team. Each module includes an Acknowledgement of Country at the start of the module in recognition that the research is being conducted on Aboriginal and Torres Strait Islander lands.

XSL•FO
RenderX

Table 2. Australian-specific modules for the Small Steps for Big Changes (SSBC) Australia implementation.

Module	Core content
1. Introduction to SSBC Australia and diabetes prevention	• Introduction to SSBC Australia program • Preventing type 2 diabetes (including Australian statistics) • Physical activity (including reference to the importance of considering the 24-hour day) • Healthy eating (in line with Australian dietary guidelines) • Supporting behavior change
2. Fostering a safe and inclusive environment	• Cultural safety in practice (Australian statistics) • Guiding principles for culturally safe practice (adapted from Australian materials) • Essentials for communicating clearly during appointments • Additional cultural awareness resources (Australian-specific)
3. SSBC Australia coach guide	• Overview of SSBC Australia (including summary of differences between SSBC Australia and SSBC Canada training) • SSBC Australia data collection process • Physical measurements protocol for SSBC Australia including demonstration videos • SSBC Australia session guides (each session has a summary infographic, MI[a] guide, and key resources) • FAQ[b] for SSBC Australia • Additional resources (Australian specific)
4. Using REDCap[c] in SSBC Australia	• Overview of REDCap (SSBC Australia–specific) • Scheduling guide for SSBC Australia • Session guide for SSBC Australia • Troubleshooting guide and FAQ
5. SSBC Australia client workbook	• A copy of the SSBC Australia workbook clients receive
6. MI (optional)	• Summary of core MI components (SPIRIT[d], OARS[e], and Ask-Tell-Ask) • Processes that guide MI conversations with examples • Change talk with examples

[a]MI: motivational interviewing.

[b]FAQ: frequently asked question.

[c]REDCap: Research Electronic Data Capture.

[d]SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials.

[e]OARS: Open-Ended Questions, Affirmations, Reflections, Summaries.

Assessment Training and Exercise Supervision Training

Site leads or other trained staff will provide training in the processes required for the in-person measurements, completing the fidelity checklists, and supervising the exercise sessions.

Competency Assessment

Coaches are required to undertake a competency assessment prior to delivery of the SSBC Australia program. The protocol and standards for coach competency assessment were designed and tested by SSBC Canada. In line with the Canadian model, assessment of competency includes 4 elements: knowledge of the program content, ability to conduct the in-person measurement protocol, ability to appropriately supervise the exercise sessions (or provide guidance on appropriate exercise sessions to attend if a site chooses to use group-based exercise), and ability to use appropriate MI skills. Knowledge will be assessed via a 20-item survey, which has been adapted from the Canadian knowledge check to align with the Australian-specific content. A pass mark of at least 70% (14/20) is required to achieve competency, with coaches able to repeat the survey as required until they meet this pass mark. Scores and the number of passes on the first attempt will be recorded. The site-level trainer or the research team will assess competency in the in-person measurement and exercise supervision protocol. For assessment of competency in MI, coaches will undertake a mock session of the first health coaching session, with assessment conducted by the research team using the Abbreviated Motivational Interviewing Competency Assessment tool [49] and a checklist against delivery of session content. In line with the Canadian SSBC coach training [11,22], the coach will be provided with feedback on their mock session by the research team (who have been trained in the use of the Abbreviated Motivational Interviewing Competency Assessment tool), with a minimum score of 0.5 of 2.0 for each of the 5 criteria required to meet competency. Potential coaches can repeat any component or part thereof to meet competence.

SSBC Australia REDCap Coach Training Database

An Australian-specific implementation strategy was the development of a REDCap database to allow staff to easily guide and track the coach training process and evaluate the training. The database also enabled the collection of data from coaches regarding their perspectives on the SSBC Australia program at various intervals after completing the training. For

XSL•FO
RenderX

student coaches, data were collected at the end of their placement and at 1 month and 1 year later; for all other coaches, data were collected at 2 months, 1 year, 2 years, 3 years, and 4 years after the training. Research staff and site staff performing the training use this database to add records for new trainee coaches, send out the required training steps in order, receive alerts indicating when each step is done, fill in competency checklists for the mock session assessments (MI, measurement, and exercise supervision), and send iterative and final feedback to the trainees, as well as to notify relevant staff when competent coaches will require access to the REDCap client database. The steps trainees perform are completing consent forms, completing pretraining surveys, following instructions to join the Online 3C Training platform (and undertake that training), successfully complete a knowledge check or a supplemental knowledge check (if the first check is failed), and use links they are sent of the practice coaching forms for the mock session assessments. After they have completed their mock session assessments, they complete posttraining surveys, receive final feedback on their mock assessments, and complete ongoing surveys tracking training outcomes and their perspectives on the training at various time points after the training.

Delivery of Program by Site (Delivery Funded by Research)

For the first 2 years, the delivery of the SSBC Australia intervention to clients will be primarily funded by the research grant, with in-kind support given by the delivery provider in relation to free gymnasium membership for the duration of the 6-session delivery (4 to 6 weeks).

Aligned with the Canadian model, the SSBC Australia intervention involves 2 key elements: the 6 individual coaching and exercise sessions with the trained SSBC coach and the client workbook.

Individual Coaching Sessions

The 6 individual sessions are delivered using an MI approach across a 3- to 4-week period (noting additional time up to 6 weeks is allowable if required by the site or client). Each session involves a coaching component for approximately 30-40 minutes and an exercise component for 20-30 minutes. Sessions 1 and 6 also have an assessment component (see Clinical Indicators (Objective Physical Assessments) section). Each coaching session includes a check-in with the client, provision of information around a session topic, planning and goal setting, and wrap-up.

Session topics covered in SSBC Australia include an introduction to the SSBC Australia program; healthy eating and identifying hidden sugars; physical sensations of cardiovascular exercise and strength training; carbohydrates, glycemic index, and a healthy plate; identifying barriers and enablers to active living; and managing setbacks and planning for success. Key changes from the SSBC Canada program were the alignment of recommendations to the Australian Physical Activity [50] and Healthy Eating [51] guidelines. This included the addition of material on hidden sugars, strength training (in line with evidence of the benefits of strength training for diabetes prevention [52]), and the importance of considering the 24-hour day [53], including the identification and impact of high sedentary time [54]. Selected by the participant, each exercise session involves either moderate-intensity continuous training or high-intensity interval training, with the duration of the exercise component increasing across the 6 sessions from 12 to 20 minutes in the first session, up to 30 minutes in each of the last 2 sessions. Delivery sites can choose to substitute or augment this core program with tailored elements such as virtual delivery of the coaching and use of group-based exercise classes. Provision and take-up of any substitutions or augmentations will be tracked. Scripts embedded in REDCap will be used to guide SSBC Australia coaches through the key points to cover in each session, in conjunction with data collection to monitor implementation fidelity.

Client Workbook

Clients will be provided with an SSBC Australia workbook, which will be used to guide the session activities. In addition to the educational content and session-based activities, there is also space for clients to monitor and track their progress toward their goals and reflect on their progress if they choose. The SSBC workbook was adapted to be suitable for an Australian delivery context (eg, the use of Australian rather than Canadian guidelines and examples of Australian rather than Canadian foods).

Community of Practice

The SSBC Canada team developed 2 communities of practice: one for SSBC coaches and one for site leads implementing the program. The Australian coaches and site leads will be invited to these virtual communities of practice to enhance and encourage learnings across all implementation sites. Engagement with the Canadian community-of-practice site will be monitored through the Canadian website [55] and a Microsoft Teams site. In addition, an online sharing site (Microsoft Teams) will be established to facilitate an Australian-based community-of-practice among SSBC Australia coaches and site leads at participating Australian sites, with engagement and meeting minutes recorded and reviewed.

Delivery of Program by Site (Delivery Funded by Site)

After 2 years of primarily research-funded delivery, sites will have the opportunity to decide whether they want to continue to deliver the program (with sites fully funding the delivery) or to not continue delivery. The research grant will continue to fund the evaluation of the program until the end of 4 years of implementation regardless of whether sites choose to continue the program or not after 2 years. Site and organization leads have committed to this evaluation for 4 years. A sustainability adaptation workshop will be conducted with each site that decides to continue delivery at both the end of year 2 (end of research-funded delivery) and the end of year 4 (end of research evaluation). SSBC Australia will be considered affordable on the APEASE criteria if the site decides to continue delivery of the program after 2 years.

XSL•FO
RenderX

SSBC Australia Participants: Delivery Providers

Overview

SSBC Australia will be delivered across 2 providers and 5 sites in South-East Queensland: The Y Queensland (4 sites) and Logan Healthy Living (1 site). These sites were identified (1) to align with the delivery partner in Canada (Y Queensland) and (2) to address an identified need for a diabetes prevention program, which can be delivered by non–health professionals (including students), from a community-based primary care partner (Logan Healthy Living). Identification and recruitment of sites occurred as part of the grant funding submission process.

The Y Queensland

The Y Queensland is a not-for-profit organization with services including childcare, fitness, camping, education, recreation, and community programs [56]. The organization is aligned with the YMCA's in Canada, which are the delivery sites for the SSBC Canada program [16]. There are 9 Y sites across South-East Queensland. In partnership with The Y leadership team, 4 sites were chosen as delivery sites for SSBC Australia: 1 inner city, 1 outer city, and 2 regional. It was expected that these sites would provide an appropriate mix of participants and resource availability to be able to inform the sustainability and potential scale-up of the program.

Logan Healthy Living

Logan Healthy Living by UQ Health Care is a community-based allied health service designed to support adults living in the Logan area with, or at risk of, type 2 diabetes [34]. Logan is a culturally diverse and relatively low-income region in South-East Queensland. The region experiences rates of diabetes higher than the Queensland average (5.2% vs 4.5% [57]), with diabetes the most common cause of potentially preventable hospitalizations [58]. Logan Healthy Living was designed to address this need, with details of the clinic development and service provision reported previously [34]. In brief, this multisectorial health hub model, which was established in 2021, is delivered by UQ Health Care in partnership with Health and Wellbeing Queensland, along with key stakeholders: The University of Queensland, Griffith University, Metro South Hospital and Health Service, and Brisbane South Primary Health Network. The clinic provides a range of group and individual allied health services, including physiotherapy, exercise physiology, dietetics, social work, and health psychology. Students from across these disciplines, and others including public health and pharmacy, are integrated into the service delivery, supervised by interprofessional clinical educators.

SSBC Australia Coaches

Overview

SSBC Australia coaches will be recruited at each of the sites using recruitment strategies identified in partnership with the delivery site. For The Y, it is expected that most coaches will be existing staff identified by the site leads as suitable for program delivery. For Logan Healthy Living, it is expected that most coaches will be drawn from students undertaking their placement at the site during the study period. Importantly, SSBC Australia coaches are not required to have any professional qualifications and can include lay volunteers and students. All coaches are required to be 18 years or older, meet the site-level requirements for volunteers at their respective sites, and pass the competency assessment for SSBC Australia intervention delivery. All coaches are also required to sign a nondisclosure agreement prior to undertaking the SSBC training as part of the Canadian-based requirements.

For Logan Healthy Living, all students, from any degree, undergoing at least a 5-week work integrated learning placement during the study period (and beyond depending on findings from the study) will be required to undertake the SSBC Australia training as part of their placement experience. Students can then opt in to provide data collected as part of this standard process to the research project and to deliver the SSBC Australia program to clients. Only those who provide consent for the research will be eligible to deliver the program to clients, providing they meet other eligibility requirements.

Target Sample Size—Coaches

There is no target sample size for coaches. Each Y site is required to have at least 2 trained coaches, while the Logan Healthy Living site is expected to have a small number of trained clinicians and a larger number of trained student coaches (anticipated total n≈50, mostly from Logan Healthy Living).

Program Participants (Clients)

Overview

Program participants (clients) will be recruited at each of the sites using recruitment strategies identified in partnership with the delivery site. Recruitment strategies are expected to include but are not limited to listserv mailouts, on-site promotion, establishment of primary care and public health care referral pathways, promotion through local community groups, and health information showcases. Referral sources and referral strategies will be tracked by the project team.

The recruitment materials will direct potentially interested clients to complete an online expression of interest (EOI) and screening survey (REDCap). This survey, which was based on the Canadian SSBC eligibility survey, is anonymous until participants reach the consent stage, with the survey stopped if a participant does not meet the inclusion criteria at each of the staged steps. Ethics for this EOI and screening survey was obtained from The University of Queensland Low and Negligible Risk Human Research Ethics Committee (#2024/HE001280).

The EOI and screening survey will include the AUSDRISK as well as questions on glycosylated hemoglobin and fasting blood glucose level (if known). Permission to use the AUSDRISK, which is a simple yet valid and reliable risk assessment tool to predict incident type 2 diabetes within the next 5 years [36], was granted from the Australian Government Department of Health and Aged Care. It comprises a series of questions about diabetes risk factors including age, sex, personal and family history of high blood glucose, First Nations background (Aboriginal, Torres Strait Islander, Pacific Islander, and Māori), other ethnic minority background (Asian, Middle Eastern, North African, and Southern European background), if taking

XSL•FO
RenderX

medication for hypertension, smoking status, diet, exercise habits, and waist measurement. A diabetes risk score between 0 and 38 will then be calculated, with scores being assigned according to risk for each factor, where a higher score indicates a higher risk of developing type 2 diabetes in the next 5 years.

Eligibility and Consent

Adults who are not pregnant (18 years or older) with a AUSDRISK score of 12 or over (considered high risk) or between 6 and 11 (considered intermediate risk) with at least 1 modifiable risk factor (smoking, low fruit and vegetables, low physical activity, and high waist circumference), or self-reported prediabetes or a glycosylated hemoglobin of 5.7% to 6.4% or fasting blood glucose of 5.6 to 6.9 mmol/L (if provided), able to provide informed consent, and able to attend at least 1 of the delivery sites, will be considered eligible and invited to complete the consent form. Information on comorbid conditions will also be captured at this point.

Participants who provide consent will then be automatically sent an information pack (via REDCap), tailored to their preferred site, which provides a guide to the SSBC Australia program and instructions for next steps. The site that the client indicates as their preferred site will also be notified of this potential client, along with information about their risk level. Participants will be encouraged to contact the site directly to support the enrollment process, depending on site preferences. Eligibility will be reconfirmed as part of this enrollment process. Once the client has been confirmed eligible by the site, they will be onboarded into the REDCap client database.

Target Sample Size—Participants

The target sample size is 500 participants enrolled (400 from the Y Queensland sites and 100 from the Logan Healthy Living site) for the 2-year researcher-led recruitment window, with no maximum sample size. This target sample size was determined based on the anticipated feasibility of recruitment. This sample size of 500 across 5 sites is sufficient to provide 80%-90% power with 5% 2-tailed significance to detect a 60 minute per week change in self-reported MVPA, which requires sample size of 394-528 based on assumptions of r=0.4, SD 330, 20% attrition, and intraclass correlation=0.001 (design effect=1+0.001(100–1)=1.099) informed by unpublished data from similar programs delivered in Logan and The Y sites.

Data Collection

Overview

A variety of data collection methods will be used, including surveys, objective assessments, checklists, qualitative interviews, site audits, website analytics, meeting minutes, and project tracking. Data collection will take place at multiple levels including the organization, provider site, coach, and client. Both online and in-person measurements will take place. Survey data specific to the trial will be collected via Qualtrics for the data from the organization and site leads and via REDCap (version 15.0.7) for the SSBC coach information and SSBC coach training, client EOI and screening, and client data. Broader site-level data (eg, memberships) will be extracted from the organizational databases. Interviews will be audio-recorded and transcribed. Meeting minutes will be taken and reviewed; website analytics (from the 3C training platform and from the SSBC website coach forums) will be downloaded and reviewed. Further details are provided below.

Organizational- and Site-Level Data Collection

The organizational- and site-level leads will provide or have provided informed consent for the action-research process of data collection throughout the project (ethics approval from The University of Queensland HREC #HE001455). Data collected at both the organizational and site level will include information on contextual factors, organizational readiness, acceptability, appropriateness, feasibility, and sustainability. Data will be collected via surveys, interviews, meeting minutes, site visits, and document reviews of operating protocols to understand how the program is working within the site workflow (Table 3).

XSL•FO
RenderX

Table 3. Data collection time points for organization leads and site-level leads.

Data collection	Preimplementation	Postimplementation				
		≈3 months	1 year	2 years	3 years	4 years
Survey						
Demographics and work characteristics[a]	✓					
Organizational readiness for change	✓			✓		✓
Barriers and facilitators to implementation	✓			✓		✓
Program perceptions	✓	✓	✓	✓	✓	✓
Sustainability assessment			✓	✓	✓	✓
Preimplementation adaptation workshop	✓					
Interviews	✓	✓	✓	✓	✓	✓
Document review of site operating protocol	✓	✓	✓	✓	✓	✓
Site visit and audit	✓	✓	✓	✓	✓	✓
Sustainability adaptation workshop				✓		✓
Meeting minutes	✓	✓	✓	✓	✓	✓
Membership conversions		✓	✓	✓	✓	✓
Access and input to community of practice		✓	✓	✓	✓	✓

[a]If organization or site-level leads change, then the new leads will complete demographics and work characteristics questions.

SSBC Australia Coach-Level Data Collection

Data collected from coaches will include information on their demographic and work characteristics, their experiences with the training, their perceptions of the program, and their fidelity to program delivery, with data collected via surveys, website analytics, and interviews. Data collection for general coaches will continue up to 4 years (Table 4); data collection for the student coaches will continue up to 12 months (Table 5). Data will be collected on the coach community of practice and the session checklists throughout program delivery.

Table 4. Data collection time points for general health coaches.

	Profile or pretraining	End training	≈2-month post	12-month post	2 years	3 years	4 years
Surveys							
Demographics	✓						
Referral pathway							
Training and experience	✓						
Work and study	✓			✓	✓	✓	✓
Training motivation							
Confidence and knowledge	✓	✓	✓	✓			
Knowledge check		✓					
Feedback on the training		✓					
Perceptions of program		✓	✓	✓	✓	✓	✓
Competency assessment		✓					
3C website analytics		✓					
Interviews			✓	✓	✓	✓	✓

XSL•FO
RenderX

Table 5. Data collection time points for student health coaches.

	Profile or pretraining	End training	End placement[a]	1-month post placement[a]	12 months
Surveys					
Demographics	✓				
Work and study	✓				
Student study profile	✓				
Confidence and knowledge	✓	✓	✓	✓	✓
Knowledge check		✓			
Feedback on the training		✓			
Student placement impact			✓		
Student training impact			✓	✓	✓
Perceptions of program		✓	✓	✓	✓
Competency assessment		✓			
3C website analytics		✓			
Interview			✓		

[a]Placement duration varies from 5 to 20+ weeks depending on the program requirements.

Client-Level Data Collection

Data from clients will be collected via online surveys as part of the eligibility, screening, and consent process and as part of their profile, preprogram, end-of-program, and at 3, 6, and 12 months post program following consent (Table 6). In-person measurements will occur at the preprogram, end-of-program, and 12-month time points. A subsample of clients will also be invited to take part in interviews at the end-of-program and at 12 months post program.

XSL•FO
RenderX

Table 6. Data collection time points for clients.

Client-level measure	Profile	Preprogram	End-of-program	3 months post program	6 months post program	12 months post program
	Enrollment	≈Session 1	≈Session 6	≈Time after session 6		
Surveys						
EOI[a] and screening	✓					
Demographics	✓					
Work and study	✓					
Emergency contact	✓					
Mailing list	✓					
Connections and access		✓				✓
Adult Pre-Exercise Screening System		✓				
Movement behaviors		✓	✓	✓	✓	✓
Dietary behaviors		✓	✓	✓	✓	✓
Program behaviors		✓	✓	✓	✓	✓
Confidence and supports		✓	✓	✓	✓	✓
Health-related QoL[b]		✓	✓		✓	✓
Loneliness		✓	✓		✓	✓
Health care use		✓				✓
Program feedback			✓			
Program acceptability			✓			
Prediabetes and diabetes status			✓	✓	✓	✓
Involvement in fitness and community		✓	✓	✓	✓	✓
Adverse events			✓	✓	✓	✓
Referral		✓	✓	✓	✓	✓
In-person clinical measures		✓	✓			✓
Interviews[c]			✓			✓

[a]EOI: expression of interest.

[b]QoL: quality of life.

[c]Interviews in the subset only. Only those who are eligible and consent go on to be scheduled to receive the program.

Measures and Outcomes

The measures and outcomes are mapped to the PRISM and RE-AIM framework [40], with information ordered by context, adoption and reach, implementation, effectiveness, and maintenance.

Context

Context will be captured under the 4 PRISM dimensions: characteristics of the implementation settings or delivery providers, perspectives on the intervention, the external environment, and implementation and sustainability infrastructure.

Characteristics of Implementation Settings and Delivery Providers

Characteristics captured will include the location of the delivery site, which will then be mapped to the Modified Monash Model [59] and Index of Relative Socio-Economic Advantage and Disadvantage [60] to understand remoteness and socioeconomic status of the areas in which the program is being provided; the number and type (eg, staff qualifications) of staff; the services offered (eg, group classes and types of classes); and the client base (number of members and types of memberships). Data will be collected via the site audit and reviewed annually.

Other characteristics will be collected via a CFIR-informed audit and interview. This will include, for example, understanding of other programs delivered by the site (especially those catering to a similar client base), alignment with the

XSL•FO
RenderX

priorities of the site and organization, resource availability (including facilities), and existing communication channels to their members and the community.

Perspectives on the Intervention

Readiness for change will be assessed via the Organizational Readiness for Implementing Change tool [61] administered to both organizational- and site-level leads at preimplementation, 3 months post implementation, and then yearly. The measure consists of 12 items, each measured on a 5-point Likert scale (1=disagree and 5=agree). The Organizational Readiness for Implementing Change has demonstrated good content adequacy and reliability [61].

The pragmatic context assessment tool will be used to understand contextual barriers to change from the perspective of site leads [62]. This questionnaire, which maps to the CFIR [37], is designed to help facilitate identification of barriers and facilitators prior to implementing change [62]. It consists of 14 statements with participants asked to indicate whether they disagree (barrier), are neutral, or agree (facilitator) with each statement, and then indicate the likely effect (weak or no effect or strong effect) of this barrier or facilitator on their ability to implement the improvement. This information can then be used to map to recommended implementation strategies [62]. The pragmatic context assessment tool will be measured at preimplementation and then at 2 and 4 years to understand contextual barriers or facilitators to sustainability. Interviews, site visits, document reviews, and meeting minutes will be used to complement the survey data.

External Environment

Information on a range of external factors will be collected including links with community health partners, links to other support services in the area, and relevant policies and market forces. Data will be collected via interviews and audits with site leads at preimplementation and through the adaptation workshop, then annually.

Implementation and Sustainability Infrastructure

Information on the support provided such as ongoing training and feedback by the dedicated implementation team via communities of practice will be collected via meeting minutes (Australia) as well as analytics data from the coach forum section of the Canadian SSBC website (Canada). Each site will have its own program operating protocol (including recruitment, enrollment, evaluation, reporting, communication, and retention protocols) with this site operating protocol developed in partnership with the site, then reviewed at approximately 2 months after implementation, then annually with protocol modifications tracked. Information regarding sustainability planning will be collected via interviews with site and organizational leads at preimplementation, then annually, as well as through the sustainability adaptation workshops at 2 and 4 years.

Adoption and Reach

Adoption and reach measures are reported in Table 7, with data collected via a mix of surveys, interviews, and project tracking.

XSL•FO
RenderX

Table 7. Adoption and reach outcome measures.

Outcome	Measure	Instrument
Adoption (organization, site, and coach)		
Reasons for taking up program	• Organizational reasons • Site reasons • Coach reasons[a]	• Preimplementation interview • Preimplementation interview • Coach profile
Uptake of health coaches	• Number of coaches interested (n) • Number of coaches eligible or ineligible (n) • Number of coaches SSBC[b] Australia certified (n)	• Coach consent • Project tracking • Competency check
Referral pathways coach[a]	• How they heard about the program	• Coach profile
Characteristics of coaches	• Demographic characteristics • Work and experience characteristics • Study profile[c]	• Coach demographics • Coach work and experience • Student coach study profile
Withdrawals or reasons for withdrawals	• Number of withdrawals (n: site and coach) • Reasons for withdrawal site • Reason for withdrawal coach	• Project tracking • Withdrawal survey or interview • Withdrawal survey or interview
Adverse events	• External factors impacting on adoption (site renewal, natural disasters, and person-made disasters)	• Project tracking
Reach (participants)		
Uptake of participants	• Number interested (n) • Number eligible or ineligible (n) • Number consented (n) • Number enrolled (n)	• EOI[d] and screening • EOI and screening • Participant consent • Number onboarded to REDCap[e] client database
Referral pathways	• How heard about the program	• EOI and screening
Characteristics of enrolled participants	• AUSDRISK[f] score, site, comorbid health conditions • Demographic, work, and study	• EOI and screening • Demographics, suitability for exercise, work, and study
Characteristics of ineligible participants	• Demographics, AUSDRISK score (if available)	• EOI and screening
Withdrawals	• Number of withdrawals (n) • Reasons for withdrawals	• Participant tracking • Withdrawal survey

[a]Not asked for the student coaches at Logan Healthy Living, as undertaking the training is a requirement of their placement.

[b]SSBC: Small Steps for Big Changes.

[c]Asked of students only.

[d]EOI: expression of interest.

[e]REDCap: Research Electronic Data Capture.

[f]AUSDRISK: Australian Type 2 Diabetes Risk Assessment Tool.

Demographic Characteristics (All Participants)

Characteristics to be measured on all participants (organizational lead, site leads, health coaches, and clients) include age (year of birth), sex assigned at birth, gender, location (provider site), ethnic origins, country of birth (Australia or other), main language spoken at home (English or other), disability status, highest education level, home postcode, and paid employment and study status.

Work History and Experience (Health Coaches)

All health coaches will additionally be asked about their qualifications, time worked in the health or health promotion field, completion of any courses in behavior or behavior change, amount of training in health or health promotion, cultural safety, MI (5-point scale from 1=none to 5=extensive), reasons for doing this training, and experience in using MI and working with clients with type 2 diabetes or prediabetes (5-point scale from 1=none to 5=extensive).

XSL•FO
RenderX

Student Study Profile (Student Health Coaches)

Student health coaches will be asked additional questions to the general health coaches that are specific to their university training, namely, student status (domestic or international), university, program of study, degree level (undergraduate or postgraduate), year of study, and number of full-time placements completed (0 to 5+).

Suitability for Exercise (Clients)

Clients will be asked to complete the Adult Pre-Exercise Screening System [63] before the intervention. None of the survey responses are automatic disqualifications for participation. Responses will be used by coaches to discuss any relevant safety concerns about participation with participants at their first session, including escalating to a review by an appropriate allied health professional or medical practitioner as required, based on the site-level protocols. This questionnaire also provides open text for the client to record anything else they would like to talk about with their coach or like their coach to be aware of.

Contextual Measures (Clients)

Clients will be asked about their social connections and access to healthy food, transport, recreation spaces, assistance with legal matters, housing or accommodation, and health services using the 6-item access subscale and the single item on connections outside of family (on a 9-point scale from 0=not at all connected or poor access to 8=very connected or easy access) from the Steps to Better Health Questionnaire [64]. Clients will also be asked about their involvement in activities outside the home (social, exercise groups, art and craft, and performing arts), their membership with any fitness facilities (the study site and other), and their referral of others to the SSBC Australia program and to the site.

Other

All coaches are asked to indicate their level of motivation to do the required training to become a SSBC Australia coach (0=not at all to 10=extremely). Further, health coaches who are not students will be asked their reasons for becoming a coach (open text) and how they heard about the program. This will not be asked of the student coaches, as they will all do the training as part of their placement requirements. Referral pathways will also be captured as part of the screening and consent process for clients.

Implementation

The implementation outcomes will consider both the SSBC Australia coach training and program delivery. SSBC Australia will be considered practicable on the APEASE criteria if it is able to be appropriately implemented (high fidelity) for both these elements.

Implementation—Health Coach Training

Fidelity of the training process will be considered in terms of the number of core training components completed: the 7 online Canadian modules (assessed via website analytics), the 4 core Australian-specific modules and review of the client workbook (assessed via self-reported completion), the knowledge check quiz, and the competency assessments. All components are expected to be completed for training fidelity.

Acceptability, appropriateness, and feasibility of the training will be measured following completion of the competency check process via a bespoke survey. In total, 5 items ask about the acceptability, usefulness, appropriateness, feasibility, and value (to current and future career) with responses on a 5-point Likert scale (1=not at all to 5=extremely), and 2 items will be asked on satisfaction regarding the length and content of the training in general (1=extremely dissatisfied to 5=extremely satisfied). For each of the training components the coach indicated they did, they will be asked how useful the component was (1=not at all useful to 5=extremely useful). One question will be open text for any other feedback. The self-reported time spent on each of the training components will be captured, along with details about training for their competency checks (since the sites have flexibility in how these are conducted).

Implementation—Program

Provider Level (Coach, Site Lead, and Organization Lead)

Acceptability, appropriateness, and feasibility of the program from the perspective of coaches, site leads, and organization leads will be measured by the Acceptability of Intervention Measure, the Intervention Appropriateness Measure, and the Feasibility of Intervention Measure [65]. These measures, which can be used separately or together, have demonstrated high test-retest reliability (R=0.73 to 0.88) and good structural validity [65]. All measures use the same 5-point Likert scale (1=completely disagree and 5=completely agree), and each measure has 4 items. Scores for the items within each domain are averaged, with higher scores indicating higher acceptability or appropriateness or feasibility. Data will be collected at preimplementation, approximately 2-3 months after implementation (to capture short-term impressions), and then yearly until year 4. Student coaches will complete the measures at preimplementation, end-of-placement, and 1 month and 12 months after implementation. These factors will also be explored via interview and open-text questions approximately 2-3 months after training, then yearly until year 4 for the general coaches and site and organizational leads and at the end-of-placement for the student coaches.

Fidelity of program delivery (site level) will be considered in terms of adherence to the site operating protocols, assessed during quarterly project site meetings.

Fidelity of coaching delivery (coaches) will be considered in terms of adherence to the SSBC Australia coaching session checklist. This checklist, which is embedded into each of the 6 coaching sessions, will capture whether the core education components for each session were covered.

Quality of the coaching (coaches) will be self-rated by the health coaches at the end of each session via a checklist. Items include client engagement with the session (single item, 1=none to 6=excellent) and overall quality of the sessions as defined by the fidelity checklist (single item, 1=none to 6=excellent). A random sample of sessions will be recorded (with permission

XSL•FO
RenderX

from the client), and coaches will be asked to rate their use of MI skills (5 items, all 1=strongly disagree to 5=strongly agree). These sessions will be reviewed by the research team against these same criteria (engagement, quality, and the use of MI skills). These recordings will also be used to confirm fidelity indicators.

The number and type of adaptations to the coach training and to program delivery made prior to and then during delivery will be tracked via the research team. The number and type of adaptations made in order to sustain SSBC in the Australian context (should sites decide to continue delivery) will be recorded as part of the sustainability adaptation workshop.

Sustainability of program delivery (site level) will be determined by the decision of the delivery site to continue delivery of the program at 2 years (end of research-funded support) and 4 years (end of evaluation support). Reasons for the decision will be captured as part of the yearly interviews with site and organization leads.

Costs will be considered in terms of time for coach training, program delivery and evaluation, site lead time, administration time, and provision of the in-kind membership. Coach training time will be asked of coaches as part of their posttraining survey suite. Each of the 6 client sessions will be time-stamped (start and finish) to track program delivery and evaluation time. Site lead time will be determined from their scheduled time allocated to the project (minus any training or program delivery time). Estimation of administration time will be asked at the regular site meetings and recorded by the research team.

Client Level

Program acceptability from a client perspective will be asked via survey at the end of the program using the Acceptability of Intervention Measure and Intervention Appropriateness Measure [65]. Clients will also be asked 4 open-ended questions on barriers experienced as part of the program, how they think the program can improve, what they liked about the program, and any other feedback relating to their experience. A subset of participants, randomly selected, will be asked to take part in an interview to gain further insights on their experience, with interviews conducted at the end of the program and at 12 months (anticipated n=10-15 at each time point).

Program satisfaction will be measured at the end of the program via a survey developed specifically for this evaluation. In total, 9 questions will ask about satisfaction in terms of their overall experience, their coach, the nutrition content, the physical activity content, the program workbook, the location of sessions, the length of sessions, the length of program, and their experience with administration (scheduling), with response options on a 7-point Likert scale (1=extremely dissatisfied and 7=extremely satisfied). Clients will also be asked to indicate on a scale of 0=not likely at all to 10=definitely their likelihood of promoting the SSBC Australia program to family, friends, or colleagues.

Fidelity of program receipt will be considered in terms of the extent to which the program was received by the client as intended. It will be measured by the number of coaching and exercise sessions successfully completed by the client (a maximum 12: 6 coaching and 6 exercise) and will be tracked by the health coach as part of their session checklist.

Effectiveness

Training Effectiveness (Coaches)

Effectiveness of the training for SSBC Australia coaches will be assessed via attainment of competency and their change from pre- to posttraining in self-rated knowledge and confidence scores (1-7). These scores incorporate knowledge and confidence in cultural safety, prevention and management of type 2 diabetes and prediabetes, MI, physical activity, and nutrition (all collected via 7-point Likert scales from 1=strongly disagree to 7=strongly agree).

Program Effectiveness (Clients)

Overview

Client-level effectiveness outcomes are listed below. The primary effectiveness outcome is self-reported moderate to vigorous physical activity time. Effectiveness outcomes are collected either via survey or by in-person measurement at preprogram, end-of-program (primary end point), and 12 months after the program ends. The self-report outcomes are also collected by surveys at 3 and 6 months after the program ends, except that health-related quality of life and loneliness are skipped at 3 months (Table 6).

Clinical Indicators (Objective Physical Assessments)

In-person measures of body composition (weight [kg] and waist circumference [cm]), cardiorespiratory fitness (resting blood pressure and 2-minute step test), and strength and balance (hand grip strength and balance test) will be conducted at baseline, end-of-program, and 12 months with the primary time point change before program to end-of-program.

Height (cm) will be measured at baseline only (Seca 217 Stadiometer) to enable calculation of body mass index. Weight will be measured by digital scales (ADE Digital Scales with a capacity of 250 kg) with outer clothing (ie, jackets) and shoes removed. Waist circumference will be measured by a tape measure at the abdomen at its narrowest point between the lower costal (10th rib) border and the top of the iliac crest, perpendicular to the long axis of the trunk.

The Queens College protocol will be used for the 2-minute step test [66]. The 2-minute step test has been shown to have high intraclass reliability (R=0.9) and criterion validity (r=0.74) [67]. Blood pressure will be measured by an automatic blood pressure monitor (model HEM-907; Omron) after the client has been seated for at least 5 minutes. The average of 2 readings will be used.

Hand grip strength will be measured with a dynamometer (Camry model EH101), with the average of 3 trials on each hand will be used. A systematic review reported this test to be a valid and reliable measure [68]. The 4-stage balance test with ceiling [69] will be used for assessment of balance, with participants asked to hold each position for at least 30 seconds if they can. This test has been shown to have acceptable reliability (test-retest r=0.66), validity, and discriminant ability [70].

XSL•FO
RenderX

Self-Report Measures

MVPA will be captured with the Physical Activity Vital Sign (PAVS) questionnaire [71]. The PAVS is recommended by the American College of Sports Medicine as part of their Exercise is Medicine resources and has demonstrated good concurrent validity [72]. An additional question will be added to the PAVS to identify how much of that time is spent specifically in vigorous (strenuous) activity.

Strength training will be captured through the question from the PAVS on strength training [71].

Movement behaviors across the 24-hour day will be measured using a bespoke single-item (4-part) question that has shown good validity against an activPAL criterion in in-house testing and good acceptability from participants. Here, participants will be asked their time (in hours and minutes) spent in the following activities for a typical day, with the total needing to add up to 24 hours: sleeping, moderate or strenuous (vigorous) activity (noting this is prefilled from their responses to the PAVS), other lighter moving or standing, and sitting or lying down (sedentary).

The percentage of prolonged sedentary time (sitting or lying for 30 minutes or more continuously) will be captured using a single-item question shown to have adequate measurement properties in terms of both test-retest and criterion validity [73].

Sleep quality will be measured by the single-item 11-point sleep quality scale (0=terrible and 10=excellent), which has demonstrated good criterion validity (r=–0.76 to –0.92) and test-retest reliability (intraclass correlation=0.62) [74].

Dietary behaviors will be measured by scores from the 9-item Mini-Eating Assessment Tool, which has shown good correlation with the Healthy Eating Index assessed by a validated food frequency questionnaire (r=0.71) [75]. Data are collected on the frequency of consumption of fruits, vegetables, legumes, nuts and seeds, fish or seafood, whole grains, refined grains, low-fat dairy, high-fat dairy and saturated fats, and sweets and sweet foods. Permission for use was granted by the Mini-Eating Assessment Tool team (March 11, 2025). An additional item on sugar-sweetened beverages was also included in the items to enable comparison to population health surveys [76].

Participants will be asked how often they do the behaviors targeted in the program, with each of the 6 questions using a 5-point Likert scale (1=never or very rarely to 5=very often or always). The questions, developed for this study, ask about tracking and setting goals for physical activity, reading food labels, adding sugar (item reversed scored), choosing low glycemic index options, and choosing a well-balanced plate.

Psychosocial measures include those related to social support, self-efficacy, and loneliness. Social support for physical activity and for diet will be measured by 2 items adapted from the Social Support and Exercise Survey [77], with respondents asked to respond to the prompt “During the past month, how often have your friends, family or members of your household a) encouraged me to do physical activity; and, b) supported me to eat healthy foods.” Response options will be on a 5-point scale (1=none of the time to 5=very often or always).

Self-efficacy for healthy eating will be measured by the 7-item healthy eating scale of the Healthy Eating and Weight Self-Efficacy scale [78]. Response options are on a 5-point Likert scale (1=strongly disagree to 5=strongly agree). This subscale has good internal consistency (α=.81) and adequate test-retest reliability (r=0.72) [78]. An additional question, specific to the SSBC Australia program, was added “based on my knowledge of glycaemic index (GI), I am able to choose foods that are low GI.”

Self-efficacy for physical activity will be measured by an adapted version of the exercise self-efficacy scale, which has shown good internal consistency (α=.92) and construct and criterion validity [79]. Participants will be asked “How confident are you right now that you could achieve your physical activity goals if ...” for 4 of the scenarios (felt pain when exercising, did not enjoy it, too busy with other activities, and felt tired) and 1 scenario (were in a bad mood) adapted from the feeling stressed or feeling depressed items. Response options are on an 11-point scale from 0=not confident to 10=very confident.

Loneliness will be measured using a single-item measure, which asks “How often do you feel lonely?” 1=often or always to 5=never [80].

Health-Related Measures

Health-related quality of life will be measured by the EQ-5D-5L [81], with registration for the use of this project obtained on July 2, 2024. This is collected as an effectiveness outcome and for its potential use for economic evaluation. The EuroQol-5D-5L asks participants to rate their health today on a visual analogue scale, where 0 is the worst health you can imagine, and 100 is the best health you imagine. It also asks participants to rate their level of difficulty (1=no problems to 5=extreme problems) with 5 dimensions (mobility, self-care, usual activities, pain or discomfort, and anxiety or depression). Their health states reported on these items will be converted to utility scores (where 0 represents dead and 1 represents full health), using value sets [82] consistent with reported norms for the Australian population [83]. The utility scores can be used to derive quality-adjusted life years.

Diabetes and prediabetes status are not effectiveness outcomes of such a short intervention but will be assessed at each time point by a single multiresponse question, along with the date of any diagnosis that occurred.

Health care use is collected in the preprogram and 12-month surveys for its potential in estimating economic benefit. Participants are asked to report over the last 12 months their number of emergency department presentations, urgent care clinic presentations, day-only hospital admissions, overnight hospital admissions (including the number of nights spent in hospital), as well as the number of times they visited each of the following: general practitioner or family doctor, allied health professional, and specialist. The questions were designed to be broadly consistent with most health care use measures [84], adding in urgent care clinics, which are a relatively new health care service in Queensland, Australia, that are designed to reduce hospital admissions and emergency department presentations.

XSL•FO
RenderX

Adverse events will be collected for trial monitoring and to inform the APEASE criteria of side effects or safety. The coach will record any adverse events occurring during sessions. Clients will be shown anything their coach recorded and asked if they have experienced any new problems with their health since they started the program (at the end of the program) or since they were last asked (at 3, 6, and 12 months post program). They will be prompted to include any new diagnoses of medical conditions or illnesses, any periods of hospitalization or surgery, any muscle injuries, bone or joint problems, or any new symptoms or worsening of preexisting conditions and told they do not need to mention anything their coach already recorded. For up to 5 problems per time point, they will be asked to describe the problem, rate the likelihood it was related to program participation (highly unlikely or unlikely or about 50:50 chance or likely or highly likely or I do not know), and the severity of the problem (mild or moderate or severe).

Maintenance

Sustainability of the program will be measured yearly by the Short Program Sustainability Assessment Tool (version 2.0) [85], with the measure completed by site and organizational leads. This tool measures sustainability across 8 domains: environmental support, funding stability, partnerships, organizational capacity, program evaluation, program adaptation, communications, and strategic planning. Each domain has 3 statements (24 in total) with a 7-point response scale for each statement (1=to little or no extent to 7=to a very great extent) as well as a "not able to answer" option. The average score in each domain and the average total score will be considered.

Data on membership conversions and retention rates will be collected at quarterly intervals from the site-level client-management system and considered at both a site and organizational level.

Maintenance of client-level outcomes will be assessed in 2 ways. Primarily, effects will be considered maintained if there is still an improvement over preprogram levels that is present following a period of noncontact after completion of the intervention. Secondarily, the degree of change from the end of the program will be evaluated against a margin of error (δ) to determine whether outcomes are worsened (by an amount more than δ), further improved (by an amount more than δ), or maintained (unchanged to within ±δ). The primary end point for maintenance is 12 months, with other time points (3 and 6 months) collected to inform how long effects might be maintained.

Data Management

REDCap (version 15.0.7) will be used to collect all survey data for the coaches and clients, the client clinical assessment measures, and the implementation checklists. REDCap is a mature, secure web application for building and managing online surveys and databases. It provides automated export procedures for seamless data downloads to Microsoft Excel and common statistical packages and reporting features to facilitate project management. Qualtrics will be used for the organization- and site-level survey data collection. Any additional data collection (eg, interview recordings, meeting minutes, and participant tracking) will be stored on The University of Queensland Research Data Manager with access only provided to those named on the ethics. This research data manager will also be used to store any combined deidentified datasets. These deidentified datasets will be used for data pooling with the Canadian implementation of the SSBC program as required to address additional research questions beyond just the Australian evaluation. All staff and students involved in this study will be provided with sufficient training in the collection of data and entry into all applications. To comply with the Australian Code for the Responsible Conduct of Research [86], research data will be retained for a minimum period of 15 years from the date of publication. A deidentified datasets will be uploaded to a data repository following publication of the primary papers.

Data Analysis

Context, adoption, reach, implementation, and sustainability outcomes will be described. To evaluate client-level effectiveness and maintenance, changes from pre- to postintervention, as well as over the relevant maintenance time points, will be assessed using linear mixed models, correcting for repeated measures and cluster (site) using random intercepts, and adjusting for baseline values of the outcome. Analyses will be of evaluable cases, adjusting for predictors of missing data, with sensitivity analysis involving multiple imputation to test the robustness of conclusions to missing data handling choices. Effectiveness of the training will be assessed by describing competency attainment and testing changes from pre- to posttraining in confidence and knowledge scores, using 2-tailed paired *t* tests or nonparametric paired tests as appropriate. These tests ignore clustering due to the expected very small sample within each of the Y sites as well as the fact the training is provided centrally (unlike the intervention). Instead, exploratory analyses will test effectiveness in relevant subgroups (eg, Y coaches and student coaches from Logan Healthy Living).

Results

The project is funded for 5 years. Ethics has been obtained, the trial has been registered, and the 5 sites have been recruited, with SSBC Australia coach training commenced at all 5 sites. Client recruitment started in September 2025 with a staggered start across sites. Results will be reported in conference abstracts and publications (with authorship determined according to scientific authorship guidelines) and annual stakeholder reports. Findings will be shared with stakeholders using a range of channels including newsletters, the community-of-practice, the Canadian website, and social media.

Discussion

Principal Findings

This trial will provide key data to inform whether the Canadian-developed SSBC diabetes prevention program can be implemented and whether it is effective within an Australian context. It is expected that the SSBC Australia intervention will demonstrate improvements in clients' health behaviors, psychosocial indicators, and selected clinical outcomes while

XSL•FO
RenderX

being feasible to implement in a range of community-based settings by trained non–health professionals. Findings will directly inform the decision-making of participating sites regarding ongoing delivery beyond the trial period and provide insights into potential scale-up nationally.

Comparison to Prior Work and Contribution to the Evidence Base

Current diabetes prevention programs in Australia typically require health professionals to deliver the program [30,31], which may limit scalability due to workforce and cost constraints. If this trial of SSBC Australia is successful, it will demonstrate that an evidence-based, community-delivered program led by trained non–health professionals can be implemented in a new health care and cultural context. This aligns with calls to action to reconsider how we approach diabetes prevention [9] and extends international evidence on the type and importance of contextual adaptations to enable successful and sustained implementation [87]. Furthermore, while undergraduate and graduate students have been involved as coaches in the Canadian SSBC delivery [11,88], there has not yet been detailed exploration regarding the role of the coach training in a student's learning experience from multiple perspectives (eg, student, clinical educator, and university), the impact of the training for their future careers, and any indirect impacts of the training and delivery on student experience (eg, on interprofessional development). Student health professionals are typically required to complete placements as part of their training, with supporting client behavior change, one of the key skills that can be developed within this setting [89]. Development of the SSBC Australia student-led model involved working with clinical educators, university placement coordinators, and students to adapt the Canadian SSBC coach training materials to be suitable for students as part of their placement experience. The coach training also builds on several of the core competencies identified that all health professionals should know for movement behavior change [90]. Importantly, students receive the benefit of learning, and potentially applying, MI skills within their clinical or project placement. This person-centered and goal-oriented communication style is used, and has shown effectiveness, across multiple fields [91]; however, it is perceived that students do not receive sufficient training in this skill [92]. This trial will help evaluate how useful students find this MI training in their future placement and roles.

Strengths and Limitations

There are several components designed to enhance the sustainability of SSBC Australia. This includes investment in the support infrastructure within each site through upskilling staff in both the program delivery and research methods. Further, the multiphase process to adapt, integrate, evaluate, and potentially sustain the SSBC Australia program is a collaborative and shared decision-making process between the research team and the organization and site leads. This pragmatic and user-centered approach helps to ensure program suitability within existing workflows; that is, SSBC Australia is designed to not just be an "add-on" research project. The evaluation is guided by the PRISM and RE-AIM implementation framework [40], with data collection embedded into delivery and usual practice where possible to facilitate ongoing evaluation beyond the research project. The APEASE criteria [41] will be used to inform the business case for further investment and expansion to other sites and organizations. Limitations include the self-report of the primary effectiveness outcomes and the self-report of diabetes status. Further funding will be sought to collect data from device-based measures in a subsample. The nonrandomized, single-arm design limits causal inference regarding effectiveness; however, this design is appropriate for the study's primary focus on implementation feasibility and contextual adaptation.

Future Directions and Dissemination

Findings from SSBC Australia will inform potential expansion to additional community sites and health service contexts across Australia. Data will also be used as part of a larger pooled dataset with Canadian partners to examine cross-country variations in program implementation and implementation strategies. Dissemination will include peer-reviewed publications, conference presentations, community presentations, and plain-language summaries. Findings will also be communicated via the SSBC Canada and organizational websites, newsletters, and social media channels. If the trial demonstrates positive outcomes, it could provide a scalable model for diabetes prevention that complements existing services and leverages the emerging health workforce.

Acknowledgments

The authors thank the Australian SSBC Urban Scale-Up collaborators: Jacqui Hain, Karen Peters, Karla Flores Schain, Kym McIntosh, Narelle Van Der Hoeven, Suz Reilly, and Tracey Smith for their contribution. The authors acknowledge the students, staff, and end users who contributed to the adaptation work and thank them for their input. This work was conducted on the lands of Turrbal, Gubbi Gubbi, Yagara and Yugambeh, Quandamooka, and Gidhabal people.

Funding

This work was supported through funding awarded from the 2021 National Health and Medical Research Council-Canadian Institutes of Health Research Healthy Cities Implementation Science Team Grant Scheme (#2018644). GNH was funded by an Australian Medical Research Future Fund Emerging Leadership Fellowship (award 1193815). SRG is partly funded by the Health and Wellbeing Centre for Research Innovation, which is co-funded by The University of Queensland and Health and Wellbeing Queensland.

XSL•FO
RenderX

Data Availability

Data sharing is not applicable to this paper, as no datasets were generated or analyzed during this study.

Authors' Contributions

MEJ developed the original Small Steps for Big Changes (SSBC) program. GNH, MEJ, SRG, ADG, DWD, DG, and MT obtained funding. GNH, SRG, ADG, KAH, and FM conducted the work involved in the exploratory phase. EAHW developed the analysis and data management plan. GNH, SRG, ADG, and KAH wrote the initial draft. All authors read and approved the final draft. All SSBC content, standard operating procedures for training, evaluating coaches, implementation and adaptation workshop materials, REDCap (Research Electronic Data Capture) surveys, implementation strategies, and all recruitment materials were developed by the Canadian research team. All SSBC Australia materials have been coadapted from the original Canadian materials and codeveloped by the Australian team in collaboration with end users.

Conflicts of Interest

MEJ is the developer of the Small Steps for Big Changes program.

Multimedia Appendix 1
SPIRIT outcomes and StaRi checklist.
[PDF File (Adobe PDF File), 381 KB - resprot_v15i1e81195_app1.pdf]

Multimedia Appendix 2
Peer review comments from Healthy Cities Implementation Science (HCIS) Team Grants Review Committee, Canadian Institutes of Health Research (CIHR).
[PDF File (Adobe PDF File), 132 KB - resprot_v15i1e81195_app2.pdf]

References

1. GBD 2021 Diabetes Collaborators. Global, regional, and national burden of diabetes from 1990 to 2021, with projections of prevalence to 2050: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2023;402(10397):203-234 [FREE Full text] [doi: 10.1016/S0140-6736(23)01301-6] [Medline: 37356446]
2. Diabetes: Australian facts. Australian Institute of Health and Welfare. 2023. URL: https://www.aihw.gov.au/reports/diabetes/diabetes/contents/how-common-is-diabetes/type-2-diabetes [accessed 2024-01-30]
3. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]
4. Knowler WC, Doherty L, Edelstein SL, Bennett PH, Dabelea D, Hoskin M, et al. Long-term effects and effect heterogeneity of lifestyle and metformin interventions on type 2 diabetes incidence over 21 years in the US Diabetes Prevention Program randomised clinical trial. Lancet Diabetes Endocrinol 2025;13(6):469-481. [doi: 10.1016/S2213-8587(25)00022-1] [Medline: 40311647]
5. Kuo S, Ye W, Wang D, McEwen L, Villatoro Santos C, Herman WH. Cost-effectiveness of the National Diabetes Prevention Program: a real-world, 2-year prospective study. Diabetes Care 2025;48(7):1180-1188. [doi: 10.2337/dc24-1110] [Medline: 39565893]
6. Neamah HH, Sebert Kuhlmann AK, Tabak RG. Effectiveness of program modification strategies of the diabetes prevention program: a systematic review. Diabetes Educ 2016;42(2):153-165 [FREE Full text] [doi: 10.1177/0145721716630386] [Medline: 26879459]
7. Venkataramani M, Pollack CE, Yeh H, Maruthur NM. Prevalence and correlates of diabetes prevention program referral and participation. Am J Prev Med 2019;56(3):452-457. [doi: 10.1016/j.amepre.2018.10.005] [Medline: 30661888]
8. Cannon MJ, Masalovich S, Ng BP, Soler RE, Jabrah R, Ely EK, et al. Retention among participants in the National Diabetes Prevention Program lifestyle change program, 2012-2017. Diabetes Care 2020;43(9):2042-2049 [FREE Full text] [doi: 10.2337/dc19-2366] [Medline: 32616617]
9. Golovaty I, Ritchie ND, Tuomilehto J, Mohan V, Ali MK, Gregg EW, et al. Two decades of diabetes prevention efforts: a call to innovate and revitalize our approach to lifestyle change. Diabetes Res Clin Pract 2023;198:110195 [FREE Full text] [doi: 10.1016/j.diabres.2022.110195] [Medline: 36470316]
10. Tseng E, Lam KY, Meza KA, O'Brien MJ, Maruthur NM. Lower-intensity interventions for prediabetes: a systematic review. Am J Prev Med 2023;65(5):906-915. [doi: 10.1016/j.amepre.2023.05.012] [Medline: 37217038]
11. Grieve NJ, Cranston KD, Jung ME. Examining the effectiveness of an e-learning training course for coaches of a type 2 diabetes prevention program. J Technol Behav Sci 2023:1-11 [FREE Full text] [doi: 10.1007/s41347-023-00316-3] [Medline: 37362065]

XSL•FO
RenderX

12. MacPherson MM, Dineen TE, Cranston KD, Jung ME. Identifying behaviour change techniques and motivational interviewing techniques in Small Steps for Big Changes: a community-based program for adults at risk for type 2 diabetes. Can J Diabetes 2020;44(8):719-726. [doi: 10.1016/j.jcjd.2020.09.011] [Medline: 33279096]
13. Bourne JE, Ivanova E, Gainforth HL, Jung ME. Mapping behavior change techniques to characterize a social cognitive theory informed physical activity intervention for adults at risk of type 2 diabetes mellitus. Transl Behav Med 2020;10(3):705-715. [doi: 10.1093/tbm/ibz008] [Medline: 30715534]
14. Cranston KD, MacPherson MM, Sim JA, Jung ME. Small steps towards an inclusive diabetes prevention program: How small steps for big changes is improving program equity and inclusion. Commun Health Equity Res Policy 2025;45(3):227-236. [doi: 10.1177/2752535X231189932] [Medline: 37478867]
15. Jung ME, Locke SR, Bourne JE, Beauchamp MR, Lee T, Singer J, et al. Cardiorespiratory fitness and accelerometer-determined physical activity following one year of free-living high-intensity interval training and moderate-intensity continuous training: a randomized trial. Int J Behav Nutr Phys Act 2020;17(1):25 [FREE Full text] [doi: 10.1186/s12966-020-00933-8] [Medline: 32102667]
16. Bean C, Sewell K, Jung ME. A winning combination: collaborating with stakeholders throughout the process of planning and implementing a type 2 diabetes prevention programme in the community. Health Soc Care Commun 2020;28(2):681-689. [doi: 10.1111/hsc.12902] [Medline: 31795016]
17. Bean C, Dineen T, Locke SR, Bouvier B, Jung ME. An evaluation of the reach and effectiveness of a diabetes prevention behaviour change program situated in a community site. Can J Diabetes 2021;45(4):360-368. [doi: 10.1016/j.jcjd.2020.10.006] [Medline: 33323314]
18. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
19. Dineen TE, Banser T, Bean C, Jung ME. Fitness facility staff demonstrate high fidelity when implementing an evidence-based diabetes prevention program. Transl Behav Med 2021;11(10):1814-1822 [FREE Full text] [doi: 10.1093/tbm/ibab039] [Medline: 33963868]
20. Dineen TE, Bean C, Jung ME. Implementation of a diabetes prevention program within two community sites: a qualitative assessment. Implement Sci Commun 2022;3(1):11 [FREE Full text] [doi: 10.1186/s43058-022-00258-6] [Medline: 35123582]
21. Dineen TE, Bean C, Cranston KD, MacPherson MM, Jung ME. Fitness facility staff can be trained to deliver a motivational interviewing-informed diabetes prevention program. Front Public Health 2021;9:728612 [FREE Full text] [doi: 10.3389/fpubh.2021.728612] [Medline: 34950623]
22. Cranston KD, Grieve NJ, Dineen TE, Jung ME. Designing and developing online training for diabetes prevention program coaches using an integrated knowledge translation approach: development and usability study. JMIR Form Res 2024;8:e50942 [FREE Full text] [doi: 10.2196/50942] [Medline: 38277214]
23. Dineen TE, Bean C, Jung ME. Successes and challenges from a motivational interviewing-informed diabetes prevention program situated in the community. Health Promot Pract 2024;25(2):274-284 [FREE Full text] [doi: 10.1177/15248399221115066] [Medline: 36047459]
24. Jung ME, Alexander A, Casmey S, Coes C, Compton S, Cullen L, et al. Small Steps for Big Changes: implementing an evidence-based diabetes prevention program into diverse urban communities. Canadian Institutes of Health Research: Team Grant: Healthy Cities Implementation Science (HCIS) Team Grants. 2023. URL: https://webapps.cihr-irsc.gc.ca/decisions/p/project_details.html?applId=471208&lang=en [accessed 2025-12-06]
25. Grieve NJ, Bourne JE, Braaten K, Goode AD, Jung ME. Setting up for success: the process of creating an implementation strategy plan for intervention scale-up. Submitted 2025 Oct.
26. Healy GN, Jung ME, Gomersall SR, Goode A, Dunstan D. Small Steps for Big Changes: implementing an evidence-based diabetes prevention program into diverse urban communities. NHMRC-CIHR Healthy Cities Implementation Science Team Grant Scheme. 2023. URL: https://hmns.uq.edu.au/project/small-steps-big-changes-implementing-evidence-based-diabetes-prevention-program-diverse-urban-communities [accessed 2025-12-06]
27. Shaw J, Tanamas S. Diabetes: The Silent Pandemic and Its Impact on Australia. Melbourne: Baker IDI Heart and Diabetes Institute; 2012.
28. Bell K, Shaw JE, Maple-Brown L, Ferris W, Gray S, Murfet G, et al. A position statement on screening and management of prediabetes in adults in primary care in Australia. Diabetes Res Clin Pract 2020;164:108188 [FREE Full text] [doi: 10.1016/j.diabres.2020.108188] [Medline: 32360708]
29. Health and wellbeing programs. Diabetes Australia. 2024. URL: https://www.diabetesaustralia.com.au/prevention-programs/ [accessed 2025-12-04]
30. Seib C, Moriarty S, McDonald N, Anderson D, Parkinson J. Changes in health behaviours in adults at-risk of chronic disease: primary outcomes from the My health for life program. BMC Public Health 2022;22(1):1648 [FREE Full text] [doi: 10.1186/s12889-022-14056-1] [Medline: 36042442]
31. Kirwan M, Chiu CL, Hay M, Laing T. Community-based exercise and lifestyle program improves health outcomes in older adults with type 2 diabetes. Int J Environ Res Public Health 2021;18(11):6147 [FREE Full text] [doi: 10.3390/ijerph18116147] [Medline: 34200198]

XSL•FO
RenderX

32. Campbell LE, Gomersall SR, Tsiamis M, Goode AD, Healy GN. Understanding diabetes risk in the Y Community of Greater Brisbane: findings from a cross-sectional survey. Health Promot J Austr 2025;36(1):e889. [doi: 10.1002/hpja.889] [Medline: 38866388]
33. Interesting facts about Queensland. Queensland Government. 2017. URL: https://www.qld.gov.au/about/about-queensland/statistics-facts/facts [accessed 2025-10-15]
34. Gomersall SR, Giguere DY, Cotugno J, Munro J, Westbrook WJ, Littlewood R, et al. Development and impact of a community-delivered, multisectoral lifestyle management service for people living with type 2 diabetes (Logan Healthy Living): protocol for a pragmatic, single-arm intervention study. JMIR Res Protoc 2025;14:e69477 [FREE Full text] [doi: 10.2196/69477] [Medline: 40063948]
35. Lehman E, Healy GN, Forbes R, Laroiya M, Gomersall SR. Impact of a clinical placement on pre-professional health students perceived knowledge, skills, and confidence in the delivery of behaviour change and physical activity support. Aust J Clin Educ 2024;13(1):133-169 [FREE Full text] [doi: 10.53300/001c.123458]
36. Australian Government Department of Health and Aged Care. The Australian Type 2 Diabetes Risk Assessment Tool (AUSDRISK). Commonwealth of Australia. 2022. URL: https://www.health.gov.au/resources/apps-and-tools/the-australian-type-2-diabetes-risk-assessment-tool-ausdrisk?language=en [accessed 2025-11-26]
37. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
38. Goode AD, Gomersall SR, Heseltine KA, Jung M, Healy GN. Understanding cultural adaptations for training Australian health coaches in the Small Steps for Big Changes diabetes prevention program. 2025 Jun Presented at: International Society of Behavioural Nutrition and Physical Activity; June 11-14, 2025; Auckland, New Zealand.
39. Heseltine KA, Goode AD, Gomersall SR, Giguere D, Mason F, Jung ME, et al. Co-design and adaptations for a student-led model of the Small Steps for Big Changes diabetes prevention program. 2025 Presented at: International Society of Behavioural Nutrition and Physical Activity; June 11-14, 2025; Auckland, New Zealand.
40. Holtrop JS, Estabrooks PA, Gaglio B, Harden SM, Kessler RS, King DK, et al. Understanding and applying the RE-AIM framework: clarifications and resources. J Clin Transl Sci 2021;5(1):e126 [FREE Full text] [doi: 10.1017/cts.2021.789] [Medline: 34367671]
41. Michie S, Atkins L, West R. The Behaviour Change Wheel: A Guide to Designing Interventions. Sutton, England: Silverback; 2014:1003-1010.
42. Loeb TB, Gholami M, Ramm K, Shedd K, Soetenga S, Jackson NJ, et al. Multilevel perceptions of the virtual delivery of the University of California Diabetes Prevention Program on RE-AIM domains due to COVID-19 mandates. Front Public Health 2024;12:1327429 [FREE Full text] [doi: 10.3389/fpubh.2024.1327429] [Medline: 38525342]
43. Weatherson K, Bourne JE, Cranston KD, Braaten K, Grieve NJ, Kelly J, et al. Scaling up a diabetes prevention program in geographically and ethnoculturally diverse urban regions of Canada: a study protocol. JMIR Res Protoc 2025. [doi: 10.2196/80276]
44. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015;10:21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
45. Proctor EK, Powell BJ, McMillen JC. Implementation strategies: recommendations for specifying and reporting. Implement Sci 2013;8:139 [FREE Full text] [doi: 10.1186/1748-5908-8-139] [Medline: 24289295]
46. Miller CJ, Barnett ML, Baumann AA, Gutner CA, Wiltsey-Stirman S. The FRAME-IS: a framework for documenting modifications to implementation strategies in healthcare. Implement Sci 2021;16(1):36 [FREE Full text] [doi: 10.1186/s13012-021-01105-3] [Medline: 33827716]
47. Butcher NJ, Monsour A, Mew EJ, Chan A, Moher D, Mayo-Wilson E, et al. Guidelines for reporting outcomes in trial protocols: the SPIRIT-outcomes 2022 extension. JAMA 2022;328(23):2345-2356. [doi: 10.1001/jama.2022.21243] [Medline: 36512367]
48. Pinnock H, Barwick M, Carpenter CR, Eldridge S, Grandes G, Griffiths CJ, et al. Standards for Reporting Implementation Studies (StaRI) statement. BMJ 2017;356:i6795 [FREE Full text] [doi: 10.1136/bmj.i6795] [Medline: 28264797]
49. Hall A. Abbreviated Motivational Interviewing Competency Assessment rating form. CCMI - Centre for Collaboration, Motivation & Innovation. URL: https://centrecmi.ca/a-mica-mini-course/ [accessed 2025-04-29]
50. Physical activity and exercise guidelines for all Australians. Australian Government Department of Health and Aged Care. 2021. URL: https://www.health.gov.au/topics/physical-activity-and-exercise/physical-activity-and-exercise-guidelines-for-all-australians [accessed 2025-11-26]
51. Australian guide to healthy eating. Australian Government Department of Health and Ageing. URL: https://www.eatforhealth.gov.au/guidelines/australian-guide-healthy-eating [accessed 2025-04-29]
52. Qadir R, Sculthorpe NF, Todd T, Brown EC. Effectiveness of resistance training and associated program characteristics in patients at risk for type 2 diabetes: a systematic review and meta-analysis. Sports Med Open 2021;7(1):38 [FREE Full text] [doi: 10.1186/s40798-021-00321-x] [Medline: 34050828]
53. Ross R, Chaput J, Giangregorio LM, Janssen I, Saunders TJ, Kho ME, et al. Canadian 24-hour movement guidelines for adults aged 18-64 years and adults aged 65 years or older: an integration of physical activity, sedentary behaviour, and

XSL•FO
RenderX

sleep. Appl Physiol Nutr Metab 2020;45(10 (Suppl 2)):S57-S102 [FREE Full text] [doi: 10.1139/apnm-2020-0467] [Medline: 33054332]
54. Saunders TJ, McIsaac T, Douillette K, Gaulton N, Hunter S, Rhodes RE, et al. Sedentary behaviour and health in adults: an overview of systematic reviews. Appl Physiol Nutr Metab 2020;45(10 (Suppl. 2)):S197-S217 [FREE Full text] [doi: 10.1139/apnm-2020-0272] [Medline: 33054341]
55. Small Steps for Big Changes. 2023. URL: https://www.smallstepsforbigchanges.com/ [accessed 2025-03-29]
56. The Y Queensland. 2025. URL: https://ymcaqueensland.org.au/ [accessed 2025-03-29]
57. Logan: 2021 Census All persons QuickStats. Australian Bureau of Statistics. 2021. URL: https://www.abs.gov.au/census/find-census-data/quickstats/2021/LGA34590 [accessed 2024-09-30]
58. Brisbane South Primary Health Network needs assessment 2019-2022. Brisbane South Primary Health Network. 2018. URL: https://www.scribd.com/document/491682115/Brisbane-South-PHN-2019-22-Needs-Assessment-Final [accessed 2025-11-26]
59. Modified Monash Model. Commonwealth of Australia: Australian Government Department of Health and Aged Care. 2023. URL: https://www.health.gov.au/topics/rural-health-workforce/classifications/mmm [accessed 2024-06-26]
60. Socio-Economic Indexes for Areas (SEIFA), Australia 2021. Australian Bureau of Statistics. 2023. URL: https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release [accessed 2024-06-26]
61. Shea CM, Jacobs SR, Esserman DA, Bruce K, Weiner BJ. Organizational readiness for implementing change: a psychometric assessment of a new measure. Implement Sci 2014;9:7 [FREE Full text] [doi: 10.1186/1748-5908-9-7] [Medline: 24410955]
62. Robinson CH, Damschroder LJ. A pragmatic context assessment tool (pCAT): using a think aloud method to develop an assessment of contextual barriers to change. Implement Sci Commun 2023;4(1):3 [FREE Full text] [doi: 10.1186/s43058-022-00380-5] [Medline: 36631914]
63. Adult Pre-Exercise Screening System. Exercise & Sports Science Australia (ESSA), Sports Medicine Australia (SMA), Fitness Australia. URL: https://www.essa.org.au/Web/Web/Resources/Tools-and-templates/Adult-Pre-Exercise-Screening-System-User-Guide.aspx [accessed 2025-11-26]
64. Oster C, Gransbury B, Anderson D, Martin V, Skuza P, Leibbrandt R. Development and validation of a self-report social determinants of health questionnaire in Australia. Health Promot Int 2023;38(3):daac029. [doi: 10.1093/heapro/daac029] [Medline: 35325112]
65. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
66. Wood R. Queens College Step Test. Topend Sports. 2008. URL: https://www.topendsports.com/testing/tests/step-queens.htm [accessed 2024-10-02]
67. Rikli R, Jones C. Development and validation of a functional fitness test for community-residing older adults. J Aging Phys Act 1999;7(2):129-161. [doi: 10.1123/japa.7.2.129]
68. Bobos P, Nazari G, Lu Z, MacDermid JC. Measurement properties of the hand grip strength assessment: a systematic review with meta-analysis. Arch Phys Med Rehabil 2020;101(3):553-565. [doi: 10.1016/j.apmr.2019.10.183] [Medline: 31730754]
69. The 4-stage balance test. Centres for Disease Control and Prevention. 2017. URL: https://www.cdc.gov/steadi/media/pdfs/STEADI-Assessment-4Stage-508.pdf [accessed 2024-08-13]
70. Rossiter-Fornoff JE, Wolf SL, Wolfson LI, Buchner DM. A cross-sectional validation study of the FICSIT common data base static balance measures. Frailty and injuries: cooperative studies of intervention techniques. J Gerontol A Biol Sci Med Sci 1995;50(6):M291-M297. [doi: 10.1093/gerona/50a.6.m291] [Medline: 7583799]
71. Greenwood JLJ, Joy EA, Stanford JB. The physical activity vital sign: a primary care tool to guide counseling for obesity. J Phys Act Health 2010;7(5):571-576. [doi: 10.1123/jpah.7.5.571] [Medline: 20864751]
72. Ball TJ, Joy EA, Gren LH, Shaw JM. Concurrent validity of a self-reported physical activity "Vital Sign" questionnaire with adult primary care patients. Prev Chronic Dis 2016;13:E16 [FREE Full text] [doi: 10.5888/pcd13.150228] [Medline: 26851335]
73. Clark BK, Stephens SK, Goode AD, Healy GN, Winkler EAH. Alternatives for measuring sitting accumulation in workplace surveys. J Occup Environ Med 2021;63(12):e853-e860 [FREE Full text] [doi: 10.1097/JOM.0000000000002387] [Medline: 34538838]
74. Snyder E, Cai B, DeMuro C, Morrison MF, Ball W. A new single-item sleep quality scale: results of psychometric evaluation in patients with chronic primary insomnia and depression. J Clin Sleep Med 2018;14(11):1849-1857 [FREE Full text] [doi: 10.5664/jcsm.7478] [Medline: 30373688]
75. Lara-Breitinger KM, Medina Inojosa JR, Li Z, Kunzova S, Lerman A, Kopecky SL, et al. Validation of a brief dietary questionnaire for use in clinical practice: Mini-EAT (Eating Assessment Tool). J Am Heart Assoc 2023;12(1):e025064 [FREE Full text] [doi: 10.1161/JAHA.121.025064] [Medline: 36583423]
76. NSW Population Health Survey. NSW Health. NSW Government. URL: https://www.health.nsw.gov.au/surveys/methods-results/Documents/questionnaire-2023.pdf [accessed 2025-12-04]

XSL•FO
RenderX

77. Sallis JF, Grossman RM, Pinski RB, Patterson TL, Nader PR. The development of scales to measure social support for diet and exercise behaviors. Prev Med 1987;16(6):825-836. [doi: 10.1016/0091-7435(87)90022-3] [Medline: 3432232]
78. Wilson-Barlow L, Hollins TR, Clopton JR. Construction and validation of the healthy eating and weight self-efficacy (HEWSE) scale. Eat Behav 2014;15(3):490-492. [doi: 10.1016/j.eatbeh.2014.06.004] [Medline: 25064304]
79. Resnick B, Jenkins LS. Testing the reliability and validity of the self-efficacy for exercise scale. Nurs Res 2000;49(3):154-159. [doi: 10.1097/00006199-200005000-00007] [Medline: 10882320]
80. Measuring loneliness: guidance for use of the national indicators on surveys. Office for National Statistics. 2018. URL: https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/measuringlonelinessguidanceforuseofthenationalindicatorsonsurveys [accessed 2025-11-26]
81. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
82. Norman R, Mulhern B, Lancsar E, Lorgelly P, Ratcliffe J, Street D, et al. The use of a discrete choice experiment including both duration and dead for the development of an EQ-5D-5L value set for Australia. Pharmacoeconomics 2023;41(4):427-438 [FREE Full text] [doi: 10.1007/s40273-023-01243-0] [Medline: 36720793]
83. Redwood L, Currow D, Kochovska S, Thomas SJ. Australian population norms for health-related quality of life measured using the EQ-5D-5L, and relationships with sociodemographic characteristics. Qual Life Res 2024;33(3):721-733 [FREE Full text] [doi: 10.1007/s11136-023-03558-z] [Medline: 38085452]
84. Leggett LE, Khadaroo RG, Holroyd-Leduc J, Lorenzetti DL, Hanson H, Wagg A, et al. Measuring resource utilization: a systematic review of validated self-reported questionnaires. Medicine (Baltimore) 2016;95(10):e2759 [FREE Full text] [doi: 10.1097/MD.0000000000002759] [Medline: 26962773]
85. Short Program Sustainability Assessment Tool. Centre for Public Health Systems Science. 2023. URL: https://www.sustaintool.org/wp-content/uploads/2024/06/Short-PSATv2-with-Scoring.pdf [accessed 2024-10-03]
86. Australian Code for the Responsible Conduct of Research 2018. National Health and Medical Research Council, Australian Research Council and Universities Australia. URL: https://www.nhmrc.gov.au/about-us/publications/australian-code-responsible-conduct-research-2018 [accessed 2025-12-06]
87. Tabak RG, Sinclair KA, Baumann AA, Racette SB, Sebert Kuhlmann A, Johnson-Jennings MD, et al. A review of diabetes prevention program translations: use of cultural adaptation and implementation research. Transl Behav Med 2015;5(4):401-414 [FREE Full text] [doi: 10.1007/s13142-015-0341-0] [Medline: 26622913]
88. Cranston KD, Ivanova E, Davis C, Jung ME. Fidelity and maintenance of motivational interviewing skills in diabetes prevention program coaches: a pilot study. Int J Health Promot Educ 2022. [doi: 10.1080/14635240.2022.2136976]
89. Gates AB, Swainson MG, Moffatt F, Kerry R, Metsios GS, Ritchie I. Undergraduate examination and assessment of knowledge and skills is crucial in capacity planning for the future healthcare workforce in physical activity interventions. Br J Sports Med 2020;54(17):1015-1016. [doi: 10.1136/bjsports-2019-101646] [Medline: 31937575]
90. Alsop T, Lehman E, Brauer S, Forbes R, Hanson CL, Healy GN, et al. What should all health professionals know about movement behaviour change? An international Delphi-based consensus statement. Br J Sports Med 2023;57(22):1419-1427. [doi: 10.1136/bjsports-2023-106870] [Medline: 37793699]
91. Bischof G, Bischof A, Rumpf H. Motivational interviewing: an evidence-based approach for use in medical practice. Dtsch Arztebl Int 2021;118(7):109-115 [FREE Full text] [doi: 10.3238/arztebl.m2021.0014] [Medline: 33835006]
92. North R, Pincus R. The state of motivational interviewing training in school counseling graduate programs. J Humanistic Counseling 2023;63(1):10-23 [FREE Full text] [doi: 10.1002/johc.12209]

Abbreviations

APEASE: Affordability, Practicality, Effectiveness and Cost-Effectiveness, Acceptability, Side-Effects/Safety, and Equity
AUSDRISK: Australian Type 2 Diabetes Risk Assessment Tool
CFIR: Consolidated Framework for Implementation Research
EOI: expression of interest
MI: motivational interviewing
MVPA: moderate- to vigorous-intensity physical activity
PAVS: Physical Activity Vital Sign
PRISM: Practical, Robust, Implementation Sustainability Model
RE-AIM: Reach, Effectiveness, Adoption, Implementation, Maintenance
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SSBC: Small Steps for Big Changes
SSBC Australia: SSBC in Australia
StaRI: Standards for Reporting Implementation Studies

XSL•FO
RenderX

Edited by A Schwartz; submitted 24.Jul.2025; this is a non–peer-reviewed article;accepted 21.Nov.2025; published 19.Jan.2026.

Please cite as:
Healy GN, Gomersall SR, Dunstan DW, Winkler EAH, Heseltine KA, Sim GQ, Mason F, Giguere D, Tsiamis M, Jung ME, Goode AD A Context-Adapted Diabetes Prevention Program (Small Steps for Big Changes) in Australia: Protocol for a Hybrid Implementation-Effectiveness Study
JMIR Res Protoc 2026;15:e81195
URL: https://www.researchprotocols.org/2026/1/e81195
doi:10.2196/81195
PMID:

©Genevieve N Healy, Sjaan R Gomersall, David W Dunstan, Elisabeth A H Winkler, Katherine A Heseltine, Grace Q Sim, Fiona Mason, Denis Giguere, Michael Tsiamis, Mary E Jung, Ana D Goode. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Determining the Impact of a Physiotherapist-Led Primary Care Model for Low Back Pain: Protocol and Analysis Plan for a Cluster Randomized Controlled Trial and Embedded Process Evaluation

Jordan Miller[1,2], PT, PhD; Catherine Donnelly[1,2], OT, PhD; Chad McClintock[1,2], PT, MSc; Kevin Varette[1], PT, MSc; Yeimi Camargo[1,2], MD; Jacquelyn Marsh[3], PhD; Monica Taljaard[4,5], PhD; Mir Sanim Al Mamun[1], MPH; Geneviève Bacchus[1], BHSc; David Barber[6], MD; Lynn Cooper[7], MA; Simon French[8], PhD; Jonathan Hill[9], PhD; Michael Green[2,6,10], MD, MPH; Joy MacDermid[3], PhD; Kathleen Norman[1], PhD; Julie Richardson[11], PhD; Joan Tranmer[12], PhD; Timothy Wideman[13], PhD

[1]School of Rehabilitation Therapy, Faculty of Health Sciences, Queen's University, Kingston, ON, Canada

[2]Health Services and Policy Research Institute, Faculty of Health Sciences, Queen's University, Kingston, ON, Canada

[3]School of Physical Therapy, Western University, London, ON, Canada

[4]Ottawa Hospital Research Institute, Ottawa, ON, Canada

[5]Epidemiology and Public Health, University of Ottawa, Ottawa, ON, Canada

[6]Department of Family Medicine, Faculty of Health Sciences, Queen's University, Kingston, ON, Canada

[7]Canadian Injured Worker's Alliance, Thunder Bay, ON, Canada

[8]Spinal Pain Research Centre, Department of Chiropractic, Macquarie University, Sydney, New South Wales, Australia

[9]School of Allied Health Professionals, Keele University, Newcastle-under-Lyme, England, United Kingdom

[10]NOSM University, Thunder Bay, ON, Canada

[11]School of Rehabilitation Science, McMaster University, Hamilton, ON, Canada

[12]School of Nursing, Faculty of Health Sciences, Queen's University, Kingston, ON, Canada

[13]School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada

Corresponding Author:
Jordan Miller, PT, PhD
School of Rehabilitation Therapy
Faculty of Health Sciences
Queen's University
31 George Street
Kingston, ON, K7L3N6
Canada
Phone: 1 (613) 533 6103
Email: jordan.miller@queensu.ca

Abstract

Background: Low back pain (LBP) is a common and disabling condition that is costly for health systems and society. Interprofessional primary care models may improve care quality and reduce this burden.

Objective: This protocol and analysis plan communicates the methods for a cluster randomized trial with the following objectives: (1) evaluate the effectiveness of a physiotherapist-led (PT-led) primary care model for LBP at improving disability (primary outcome), pain intensity, quality of life, global rating of change, patient satisfaction, and adverse events compared with usual physician-led primary care; and (2) determine the impact of the PT-led primary care model for LBP on the health care system and society (health care access, health care use, missed work, cost-effectiveness). Both objectives are evaluated over a 1-year period. A multimethod process evaluation is embedded to assess model implementation, mechanisms, perspectives of patients and providers, and contextual influences.

Methods: This study is a cluster randomized controlled trial with 20 primary care practices (clusters) in Canada, randomized 1:1 to a PT-led or usual physician-led primary care model for LBP. Adults seeking care from their primary care team for LBP are recruited over 1 year. Data collection occurs at baseline, 6 weeks, and 3, 6, 9, and 12 months. Effectiveness will be analyzed using linear mixed regression. The process evaluation analysis will include: descriptive and comparative analyses to assess

XSL•FO
RenderX

implementation; descriptive and mediation analyses to assess potential mechanisms; qualitative interpretive description to understand experiences and perspectives of patients, PTs, and other health professionals; and mixed methods to determine contextual influences on implementation.

Results: Recruitment of primary care sites (clusters) was completed in June 2023, following delays related to the COVID-19 pandemic. Cluster randomization occurred in July 2023. Recruitment of patient participants began in October 2023 and concluded in November 2024 (n=739). The final self-reported patient data was collected on November 25, 2025. Extraction of electronic health record data is scheduled for completion on December 19, 2025. Data analysis will be conducted in accordance with the study protocol and analysis plan and will begin once all data collection activities are complete. No interim analyses have been performed.

Conclusions: The results of this trial will provide evidence for knowledge users to determine whether a PT-led primary care model for LBP is effective and should be adopted more widely. Knowledge users have identified the impact of the new model of care on disability, quality of life, and cost-effectiveness as key evidence needed to inform key decision-making. The multimethod process evaluation will provide critical evidence to interpret trial results and inform future scale and spread of this model of care if effective.

Trial Registration: ClinicalTrials.gov NCT04287413; https://clinicaltrials.gov/study/NCT04287413

International Registered Report Identifier (IRRID): DERR1-10.2196/89004

(JMIR Res Protoc 2026;15:e89004) doi:10.2196/89004

KEYWORDS

primary health care; clinical trial; health services; cluster randomized controlled trial; physiotherapy

Introduction

Background

Low back pain (LBP) is the world's largest contributor to years lived with disability [1], costs the Canadian health care system between CAD $6 and CAD $12 billion (US $4.4 to US $8.8 billion) annually [2], and is a leading contributor to lost work productivity [3,4]. The burden of LBP on the health care system is evidenced by frequent health care use, including unnecessary specialist consultations, diagnostic procedures [4,5], opioid prescriptions [6], and emergency department (ED) visits [7].

LBP is among the most common reasons for physician visits, with primary care physicians being the most frequent first point of contact within the health care system for people with LBP [8-10]. A *Lancet* [11-13] series on LBP highlighted international expert consensus on the need to evaluate new primary care models that better support physicians, who receive limited training in the management of musculoskeletal conditions and report low confidence addressing LBP [14-16]. One of these models involves integrating physiotherapists (PTs) within primary care teams at the first point of contact, which has the potential to provide patients with a more focused LBP consultation, assist with evidence-based treatment delivery, reduce low-value care (such as inappropriate imaging, specialist physician referrals, and opioid prescriptions), and reduce primary care physician visits. If integrating PTs can reduce primary care physician visits, it could contribute to increasing the capacity of primary care teams to help address the growing challenge of providing universal access to primary care [17] in the context of more complex patient encounters, increasing prevalence of multiple chronic conditions [18], and an aging population [19].

Evidence suggests PTs can provide collaborative care and implement recommendations from LBP primary care guidelines [20-23], including: screening for serious pathology, the need for diagnostics [24] and identifying risk factors for poor recovery [25]; providing reassurance, encouragement for early return to work, and exercise and physical activity recommendations [26]; and delivering targeted, psychologically informed interventions for those at risk of prolonged recovery [27]. Evidence from outside of Canada suggests that early guideline-adherent PT care for LBP improves function and disability [28,29] while reducing diagnostic imaging, opioid prescriptions, and physician specialist referrals [29-31], and reducing per-person health care costs [32,33].

Evidence from observational research suggests that involving PTs in providing first-point-of-contact care (ie, PT-led care) for those with work-related injuries in the US military resulted in workers being more satisfied with their care, receiving faster access to treatment, having fewer sickness absences, and using PTs and specialist physicians more appropriately [34-39]. Observational studies from the UK National Health Service on first-contact PT models of care indicate similar health outcomes, high levels of patient and physician satisfaction, shorter physician wait times, fewer work absences and diagnostic images, lower prescription medication use, and reduced costs [40-43].

The absence of high-quality randomized trials of PT-led primary care models leaves important unanswered questions about the process and impact of integrating PTs within primary care teams for people with LBP [44]. Specifically, there is a need for higher quality evidence on the impact of PT-led primary care on patient-oriented outcomes (eg, function, pain, and quality of life), health system outcomes (eg, health care access, physician workload, ED visits, specialist physician referrals, medication use, diagnostic imaging), and societal outcomes (eg, missed work, cost-effectiveness). Furthermore, it is unclear how PTs will navigate primary care challenges, such as providing care for people presenting with multiple health concerns or fulfilling requests for medications, diagnostic imaging, or notes for work

XSL•FO
RenderX

absences. This cluster-randomized trial will address these gaps by assessing the impact of a PT-led primary care model for LBP. The results will inform primary care transformations across multiple health systems and potentially improve outcomes for patients with LBP.

The purpose of publishing this protocol and analysis plan is to transparently report our design and methods, and to transparently communicate our analytic plan in advance of completing data collection and carrying out our planned analysis to reduce the risk of analytic or reporting bias.

Research Questions

The research questions involved in this study are as follows:

1. Is a PT-led primary care model for LBP effective at improving disability (primary outcome), pain intensity, quality of life, global rating of change (GROC), patient satisfaction, and adverse events compared with usual physician-led primary care, when evaluated over a 1-year period?
2. What is the impact of a PT-led primary care model for LBP on the health system and society (health care access, primary care physician workload, health care use, missed work, cost-effectiveness), evaluated over a 1-year period?

Methods

Trial Design

The trial is a parallel arm cluster randomized controlled trial at 20 primary care sites randomized 1:1 to a PT-led or usual physician-led primary care model for LBP. Randomization of practices, rather than patient participants, allows evaluation of a model where PTs are able to fully integrate within the primary care team and reduces potential contamination between study arms [45]. This protocol and analysis plan includes all of the items included in the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trial) reporting guidelines [46,47]. Figure 1 provides a summary of the schedule of enrolment, interventions, and assessments.

XSL•FO
RenderX

Figure 1. Schedule of enrolment, interventions, and assessments. EHR: electronic health record; EQ-5D-5L: EuroQOL-5D; GROC: global rating of change; LBP: low back pain; PCS: Pain Catastrophizing Scale; PHQ-2: 2-item Patient Health Questionnaire; PSEQ: Pain Self-Efficacy Questionnaire; PT: physiotherapist; RMDQ: Roland-Morris Disability Questionnaire; TSK-11: Tampa Scale of Kinesiophobia 11.

	STUDY PERIOD							
	Enrolment of clusters	Allocation of clusters	Enrolment of patient participants	Post-allocation				Close-out
TIMEPOINT	Pre-intervention	Pre-intervention	0 (baseline)	6 wk	12 wk	6 mo.	9 mo.	12 mo.
ENROLMENT OF PRIMARY CARE SITES:								
Eligibility screen	✓							
Informed consent	✓							
Allocation to study group		✓						
ENROLMENT OF PATIENT PARTICIPANTS:								
Eligibility screen			✓					
Informed consent			✓					
INTERVENTIONS:								
PT-led primary care			←	—	—	—	—	→
Usual care			←	—	—	—	—	→
ASSESSMENTS:								
Participant characteristics and demographic information			✓					
Participant health outcomes (RMDQ, pain intensity, EQ-5D-5L, PSEQ, PCS, TSK-11, PHQ-2)			✓	✓	✓	✓	✓	✓
Participant health and experience outcomes (Satisfaction, GROC, adverse events)				✓	✓	✓	✓	✓
Health system and societal outcomes (provider encounters, EHR utilization, healthcare utilization, missed occupational activities, assistance required)				✓	✓	✓	✓	✓
Implementation measures (timely access, PT as first point of contact, access to PT, LBP management)			✓	✓	✓	✓	✓	✓
Implementation measures (patient adherence)				✓				
QUALITATIVE INVERVIEWS:								
Patient participants					✓			
Primary care providers and leads, medical office assistants*							✓	✓

The trial incorporates a mixed methods process evaluation informed by United Kingdom Medical Research Council guidance for developing and evaluating complex interventions [48,49], the Consolidated Framework for Implementation Research (CFIR) [50,51], and findings from the process evaluation from our pilot trial [52]. Process evaluations assess how interventions were implemented and under what conditions. They are important to help interpret trial results (eg, explaining why an intervention fails or has unexpected consequences, or why it works and how it can be refined). Process evaluations can be particularly valuable for informing future implementation and sustainability of complex health care interventions [49]. Our process evaluation is informed by and intended to refine our program theory [48] for the PT-led primary care model for LBP (Figure 2). The program theory describes how the PT-led primary care model for LBP is intended to lead to improvements in trial outcomes.

XSL•FO
RenderX

Figure 2. Physiotherapy-led primary care model program theory. LBP: low back pain; PT: physiotherapist.

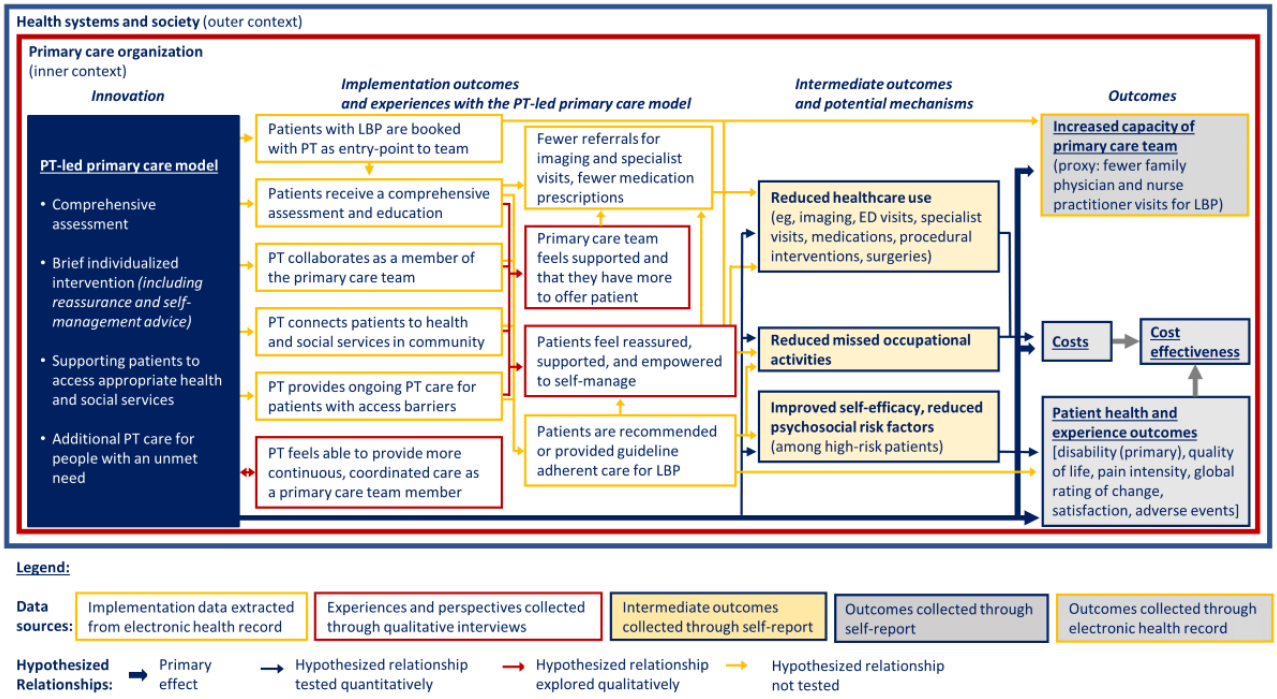

Allocation of Participating Primary Care Teams to Trial Groups

Restricted randomization is recommended for cluster randomized trials to ensure the arms are balanced at baseline. We used covariate-constrained randomization [53] with a 1:1 ratio to the intervention and comparison arms to retain the merits of random allocation while ensuring baseline balance across the arms. We stratified by location (Southeastern Ontario or Interior British Columbia) and included the following covariates: number of active patients and rural versus urban setting. Our maximum tolerable difference in the rurality indicator was 1, and the strata balancing criteria for the number of active patients was set to 10%. An independent statistician implemented the procedure using a SAS macro [54]. Concealment was maintained by ensuring each practice had an anonymized code and performing randomization after all sites were recruited.

Methods for Protecting Against Sources of Bias

Due to the trial design and interventions being compared, the PTs, patient participants, other members of the primary care team, and research assistants have not been blinded. While limitations due to the inability to blind are unavoidable, we have taken additional steps to minimize the risk of bias suggested for cluster trials [55]. A common challenge in cluster trials is identification and recruitment bias when patient recruitment must take place after cluster allocation [56]. To minimize the risks of these biases, we invited consecutive patients to participate, provided the same information about the trial to both groups prior to participants consenting (inviting participation in a study that involved either seeing their family physician as usual or seeing a PT added to their team in order not to reveal the cluster allocation until after consent and baseline data collection), employed research assistants unfamiliar with the patients to recruit and consent participants, and rigorously applied inclusion/exclusion criteria at all sites [57].

Inclusion/Exclusion Criteria

For primary care sites, inclusion criteria were ≥2 family physicians or nurse practitioners and 1500 or more active patients. Exclusion criteria were: already having a PT on the team or no space available for a PT to practice. Patient participants were screened prior to consent by a trained research assistant at both intervention and comparison sites. Inclusion criteria for patient participants were: adults (≥19 years of age) with LBP of any duration, who were able to read, write, and speak English. Exclusion criteria were: known cancer that could possibly contribute to LBP, or inability to complete the scheduled follow-ups over 1 year.

Recruitment

We recruited 20 sites (14 in Ontario and 6 in British Columbia, Canada). Site recruitment focused primarily on the engagement of contacted sites to reduce the risk of site withdrawal [58]. For patient participant recruitment, staff at each primary care site screened patients for their willingness to participate when they called to book an appointment. If patients booked via a web-based booking system, their reason for visit was screened by the clinic staff, and staff reached out to potential participants to ask about their willingness to be contacted by a research assistant. Trained research assistants screened potential participants using inclusion/exclusion criteria and enrolled consenting patients. Patients who raised LBP with any provider during a clinical encounter were also identified as potential participants. In these cases, they were directed to the administrative team, who then connected them with a research assistant for screening and enrollment. Balance in recruitment across arms and across clusters was monitored throughout the study, and strategies were implemented to maintain consistent implementation of recruitment processes across sites.

XSL•FO
RenderX

Sample Size and Power Calculations

We used the methodology of Teerenstra et al [59] to calculate the required number of clusters based on an analysis of covariance for the primary outcome (Roland-Morris Disability Questionnaire [RMDQ]) at 12 months, adjusting for baseline. Our target sample was 20 clusters (10 per arm), allowing for 1560 patient participants (78 per practice before attrition). Conservatively accounting for attrition of 2 clusters (resulting in 18 clusters) and 20% of patients (resulting in 63 patients/cluster), this target would achieve 90% power to detect a minimally important mean difference of 2.5 points [60] (Cohen d=0.4) using a 2-sided α of 0.05 and assuming a SD of 5.7 points based on pilot study data, a conservative intracluster correlation coefficient of 0.1 [61], a cluster autocorrelation coefficient (correlation between cluster means at baseline and follow-up) of 0.5, and an individual autocorrelation coefficient (correlation between participant scores at baseline and follow-up) of 0.6 informed by our pilot study [62]. An average rate of 1.5 patient participants per site each week over 1 year would achieve the recruitment target of 1560 participants, which we anticipated based on our pilot study, where we recruited and retained 4 sites and achieved a recruitment rate of >1.7 participants/site/week. Recruitment of 400 patient participants across 18 clusters (22 participants/cluster) would achieve 80% power.

Clinical Partner Sites

This study is being coordinated from a primary Research Coordinating Centre at Queen's University with 20 participating primary care sites (clusters) between Southeastern Ontario (n=14) and Interior British Columbia (n=6), Canada. Sites include representation from urban and rural settings in both regions to facilitate generalizability.

Patient and Public Involvement

A senior advisor for the Ontario Ministry of Health contributed to the plan for this study. An individual with lived experience is a coauthor on this protocol (LC) who has been engaged in the design of this study, including the identification and decision on the trial outcome measures that are important to patients. Their involvement in carrying out the study includes pretesting of data collection methods, interpretation of process evaluation results, and participation in the development of knowledge mobilization products (ie, designing summaries appropriate for patient organizations).

Trial Interventions

Index Intervention: PT-Led Primary Care Model for LBP

Overview

The index intervention involves integrating a PT within the primary care team and making them available at the first point of contact for people with LBP. Patients with LBP are encouraged to book with the PT except when the primary reason for their visit is medication renewal or when additional health concerns require another provider's attention. The intervention has four components: (1) assessment and screening; (2) brief individualized intervention; (3) supporting patients to access appropriate health services based on assessment findings; and (4) providing additional PT care to people with an unmet need. Ten registered PTs, who have completed 2 days of training on this care model, are participating in the delivery of the model across the 10 practices randomized to the PT-led primary care arm.

Assessment and Screening

The assessment and screening includes: taking a history; screening for pathology (eg, cauda equina syndrome, traumatic fracture, cancer); physical and neurological examination; application of evidence-based tools to identify comorbid health conditions (eg, 2-item Patient Health Questionnaire [PHQ-2] [63] for depression) requiring additional care; and using a validated tool (Keele STarT Back Tool [27,64]) to identify risk factors associated with persistent LBP and disability.

Brief Individualized Intervention

The PTs provide a brief individualized intervention at the initial visit. This intervention is intended to be based on primary care guidelines for LBP [23] and consists of effective communication to validate the patient's experiences [65], cognitive reassurance [66], individually tailored exercises [67,68], and advice/strategies to stay active [69].

Supporting Patients to Access Appropriate Health Services Based on Assessment Findings

It is intended that the PT collaborates with the patient to identify appropriate health services based on assessment findings, collaborates with relevant primary care team members to provide the needed care, and integrates health service providers from outside of the team, considering the patient's needs and access to health services. First, the PT identifies potential pathology needing urgent referral (eg, cauda equina, fracture, and infection). Next, they identify comorbid conditions that require collaboration with other primary care team members (eg, physicians, nurse practitioners, nurses, social workers, occupational therapists, pharmacists, and dieticians). For example, people who screen positive for depression are referred to their physician, nurse practitioner, or a member of the mental health team. Finally, primary care PTs refer to community PT as informed by the patient's score on the Keele STarT Back tool [27,64] (if appropriate). The STarT Back tool categorizes patients with LBP into low, medium, or high risk of persistent pain and disability based on physical and psychosocial risk factors [64]. The recommended matched treatment for low-risk patients is to provide self-management advice and to avoid referral and investigations where possible. This intervention is brief, and it is expected that it can take place at the first visit. The recommended matched treatment for medium-risk patients is referral for standard community PT, and the matched treatment for high-risk patients is PT care from the primary care PT who received specific training aimed at reducing physical and psychosocial risk factors for chronic pain and disability as part of the 2-day training. This stratified approach to care has demonstrated improved function, quality of life, and cost-effectiveness in comparison to usual care in the United Kingdom [27]. When a need for PT care is identified, the primary care PT helps these patients navigate the available PT resources (ie, private PT clinics when the patient has private

XSL•FO
RenderX

health insurance; government-funded PT for those who meet the criteria; or PT in primary care for those without access to a PT elsewhere).

Providing Additional PT Care to People With an Unmet Need

The PT provides additional care at the primary care site for patients who are appropriate for PT (based on STarT Back score) and who have barriers to accessing PT care in their community. For example, people who may benefit from physiotherapy but do not have access to services—due to a lack of private or government insurance coverage or because of geographic or transportation barriers—are offered additional physiotherapy care. Care includes evidence-based, guideline-consistent management, such as individualized education, exercise, and cognitive behavioral approaches. To avoid duplication of available PT services, patient participants with private- or government-funded health insurance for PT are referred to external PT services.

Comparison: Usual Physician-Led Care

The usual physician- or nurse practitioner-led primary care intervention has been unstandardized to reflect usual primary care clinical practice in Canada. Our pilot study suggested the most common management approaches in the usual care model included diagnostic imaging (12% of patients), medication renewal (21%), new medication prescription (14%), notes to employers (12%), and referral (PT 23%, chronic pain clinic 5%, massage therapy 5%, physician specialists 5%, and dietician 2%). Interventions provided or recommended by the physician have been and will continue to be recorded and monitored throughout the full trial.

There are no required or prohibited concomitant treatment options for either treatment arm. Should a participant choose to explore other or additional treatment options within their primary care team or from other health service providers, they are free to do so. This was made clear in the letter of information and consent form provided to the participants upon entry to the study.

Duration of the Treatment Period

The PT-led primary care model is being implemented over a 1-year period from the time of consent. All participants were offered an initial assessment with the PT and have access to the PT for follow-up needs for a 1-year period after their initial assessment. The majority of participants classified as low-risk using the STarT Back tool, and those with private or government-funded health insurance for PT are intended to refine our program theory for the PT-led primary care model for LBP, but will have access to the PT as a member of their primary care team throughout the 1-year follow-up period if they need or want a follow-up visit. Participants identified as medium- or high-risk without access to PT elsewhere are offered additional PT care from the PT in primary care. The frequency and duration of the treatment plan is determined by the PT and the patient participant. In our pilot study, participants who took part in ongoing care received an average of 4 visits over 8 weeks.

Intervention Modifications

We do not anticipate encountering a situation that would require the withdrawal of a participant for safety concerns related to the PT-led primary care or usual physician-led primary care. As per normal primary care and PT practice for LBP, the intervention will be modified by the primary care team to maintain participant safety as part of the model of care (eg, modification of exercise in response to increases in pain that may be experienced with an exercise intervention, discontinuation of medications if adverse effects are experienced). Due to the low-risk nature of the study and the fact that patient participants maintained access to their usual primary care providers, a Data Monitoring Committee was not established. Ongoing oversight and patient participant safety monitoring were provided by the study investigators, including monitoring all adverse events reported at all time points.

Data Collection and Management

All patient participant-reported outcome measures are being collected from participants using REDCap (Research Electronic Data Capture; Vanderbilt University), a secure online survey and data capture tool hosted at Queen's University [70,71]. A distinct link for the surveys is being sent to each participant at each assessment time point. We are providing the option of completing the questionnaires in person or by phone if participants have barriers to completing them online. Data are also being extracted from participants' electronic health records (EHRs) related to care provided for LBP. Data are being extracted using prepiloted EHR extraction forms.

At study completion, responses to the surveys will be exported directly from REDCap to encrypted- and password-protected datasets and stored securely in Microsoft OneDrive at Queen's University. All data collected from the EHR, along with a master linking log that links study identification numbers with participants, will be stored in password-protected and encrypted files in OneDrive. The linking log will be permanently destroyed at the end of the data analysis period. Qualitative interview recordings will be transcribed, deidentified, and stored securely in OneDrive.

We implemented multiple strategies to promote participant retention across all time points. Research assistants sent reminders every 2-3 days through personalized emails, phone calls, and text messages to encourage timely survey completion. When requested, surveys were completed in-person or by phone to strengthen engagement and minimize loss to follow-up.

Frequency and Duration of Follow-Up

Patient characteristics and demographic information were collected at baseline. All patient-reported outcome measures are being collected at baseline, 6 weeks, and 3, 6, 9, and 12 months from the initial visit, with the primary comparison at 12 months for all outcomes. Patient satisfaction, GROC, adverse effects, and health care use are being collected at follow-up time points only. Patient adherence to PT recommendations is being collected at the 6-week follow-up.

XSL•FO
RenderX

Participant Characteristics

Sociodemographic and Clinical Characteristics

The following characteristics and demographic information were collected from patients at baseline: age, first 3 digits of their postal code (to determine rural/urban status), biological sex, gender, identification as a member of a racialized group, duration of the current episode of LBP, previous history of LBP, number of other pain locations, highest level of education achieved, household income, and work status. In addition, the following questionnaires were administered at baseline as potential covariates and to inform subgroup analyses as part of our process evaluation.

Functional Comorbidity Index

An 18-item list of comorbidities that are associated with physical functioning. Each comorbidity is assigned a score of 1, and the total score is the sum of the comorbidity elements [72,73].

Keele STarT Back Tool

The Keele STarT Back Tool categorizes patients with LBP into low, medium, or high risk of persistent pain and disability based on physical and psychosocial risk factors [27,64,74].

Patient Health and Experience Outcomes

Self-Reported Disability (Primary Outcome)

Self-reported disability (primary outcome) is measured using the RMDQ, which demonstrates reliability, validity, and responsiveness in people with acute and chronic LBP [75,76].

Pain Intensity

Pain intensity is measured using a Numeric Pain Rating Scale (0-10) [77] for pain at rest, pain when walking, and pain when lifting a bag of groceries from the floor. Each will be reported on a scale of 0 (no pain at all) to 10 (worst imaginable pain).

Health-Related Quality of Life

Health-related quality of life is assessed using the EuroQOL-5D (EQ-5D-5L) [78], which demonstrates good reliability and validity, and is suitable for economic evaluation in LBP [79]. The EQ-5D-5L Visual Analog Scale score (0-100) will be reported as a patient health outcome. The EQ-5D responses will be converted to an EQ-5D index value using the value set calculated for the Canadian context [80]. The index value will be reported as a patient-level health outcome and will also be used to calculate quality-adjusted life years (QALYs) for our outcome in the economic evaluation.

Global Rating of Change

Global Rating of Change is measured using an 11-point GROC scale that asks participants, “with respect to your functional abilities, how would you describe yourself now compared to how you were at your initial appointment for this episode of back pain?” Participants are asked to respond on a scale from a great deal better (+5) to a great deal worse (–5), as has been recommended for self-reported rating of change [81].

Participant Satisfaction

Participant satisfaction is measured using an 11-point scale for satisfaction with the care they received for their LBP, extremely dissatisfied (–5) to extremely satisfied (+5).

Adverse Events

Adverse events are assessed using a questionnaire consistent with reporting guidelines [82,83] that asks: (1) if the patient participant has experienced any adverse events as a result of any of the treatments received, (2) what adverse events were experienced, (3) how long the events lasted, and (4) how severe each adverse event was. For analysis, adverse events will be identified as serious or nonserious. An adverse event will be identified as serious if any of the following criteria are met: the participant requires inpatient hospitalization or an ED visit due to the adverse event, the adverse event results in significant and persistent disability (beyond 72 hours), or the adverse event is life-threatening or results in death. These responses are being monitored as they are completed in order to provide ongoing oversight of patient participant safety.

The following measures will be assessed as secondary outcomes, reported in the trial results, and included in the process evaluation as potential mechanisms through which the PT-led primary care model influences LBP-related disability.

Self-Efficacy

Self-efficacy is assessed as confidence in the ability to participate in usual activities using the Pain Self-Efficacy Questionnaire (PSEQ) [84,85]

Psychosocial Risk Factors for Persistent Pain and Disability

The Pain Catastrophizing Scale (PCS) [86,87], Tampa Scale of Kinesiophobia 11 (TSK-11) [88,89], and PHQ-2 [63,90] will measure psychosocial factors associated with pain-related disability.

Health System and Societal Outcomes

Primary Care Physician or Nurse Practitioner Encounters

The number of new and repeat primary care physician or nurse practitioner visits for LBP per patient. This measure will be considered a proxy for a potential increase in primary care team capacity achieved if the LBP-related workload of primary care physicians or nurse practitioners is reduced, thus increasing their availability to provide care to other patients.

Health Care Usage Within the Primary Care Team

Health care usage within the primary care team is assessed using data collected from the EHR, including consultations with other primary care team members (eg, physicians, nurse practitioners, nurses, social workers, and occupational therapists) and group programming accessed within the primary care organization.

Health Care Use Outside of the Primary Care Team

Health care use outside of the primary care team is assessed using self-report data from follow-up surveys, cross-checked with EHR reports when possible, including medications used;

XSL•FO
RenderX

walk-in clinic visits; ED visits; inpatient hospital stays; surgeries, injections, and other interventional procedures; visits to specialist physicians; diagnostic imaging; and visits to other health professionals outside the primary care team (eg, chiropractors, massage therapists, occupational therapists, PTs, chronic pain clinics).

Missed Occupational Activities

Missed occupational activities are assessed using self-report data from follow-up surveys, including time (days) lost from paid employment, volunteer, homemaking, or educational activities related to LBP.

Assistance Required

Assistance required is assessed using self-report data from follow-up surveys: paid and unpaid assistance required. For example, self-care (eg, taking medications, dressing/undressing, going to the bathroom, bathing/showering, grooming), shopping/groceries, meal preparation, housework, managing finances, or transportation (eg, to a medical appointment).

Costs

Total costs per person will be calculated by summing direct health care costs and indirect costs using a human capital approach for missed occupational activities.

Sources of Direct Health Care Cost Data

Intervention costs will include the PT salary and training needed to carry out the intervention. Costs for publicly funded health care services will be obtained from the Ontario Ministry of Health Schedule of Benefits [91]. Medication costs will be obtained from the Ontario Drug Benefit formulary. Expenses related to health services funded by private insurance or out of pocket will be collected through self-report at all follow-up assessments. Other costs incurred by the participant related to their LBP are also being collected by self-report, including support or assistance for self-care, housework, shopping, or transportation (eg, to health care appointments). The total direct costs will be determined by multiplying the quantity of resource use by the corresponding unit cost, summing the total cost over each follow-up interval, and then calculating the mean cost at each follow-up time point, as well as an overall mean cost for the entire study period.

Indirect Costs

Non-health care costs will be limited to loss of productivity using a human capital approach [92]. The mean provincial wage reported by Statistics Canada will be used to assign a monetary value to time lost from paid employment. The minimum wage value in Ontario and British Columbia will be used to place a value on time lost by those who were retired and time lost from volunteer, homemaking, caregiving, or educational activities.

Cost-Effectiveness

We will conduct a cost-utility analysis from both societal (primary) and health payer (secondary) perspectives to meet the needs of all knowledge users. For both societal and health payer perspectives, we will estimate the incremental cost-per-QALY gained [93-95].

Implementation Measures

The following measures will be used to assess how the PT-led primary care model and the usual physician- or nurse practitioner-led primary care model were implemented.

Timely Access to LBP Care

Timely access to LBP care is assessed using the percentage of patient participants with LBP who are assessed within 48 hours of calling for an appointment. Participants who learned about the study and were invited to participate during an appointment for their LBP will not be included in this analysis.

PT as the First Point of Contact

Using the percentage of patient participants with LBP in the PT-led primary care arm who visited a PT as their first point of contact for the current episode of LBP.

Access to PT Services

PT as the first point of contact is assessed using the percentage of patients who are classified as medium- or high-risk on the STarT Back screening tool who access PT.

LBP Management

LBP management is described using data collected from the EHR, including education; exercise; psychological approaches; referrals to other primary care team members (eg, primary care physicians, nurse practitioners, nurses, social workers, occupational therapists, and group programming); referrals made to health professionals outside of the primary care team; medications prescribed, deprescribed, and suggested; diagnostic imaging ordered; lab work ordered and received; notes to employers or insurers; interprofessional communications with the primary care team; and other interventions provided.

Patient Adherence to Recommendations

Adherence to recommendations from the primary care PT is being collected at the 6-week follow-up. For patient participants identified as medium- or high-risk using the STarT Back classification, whether or not they accessed recommended PT (either through a referral to a community PT or through the PT in primary care) is being collected as part of our health use survey questions.

Qualitative Interviews

Semistructured qualitative interviews are being conducted with patients, PTs, other health professionals, primary care team members, medical office assistants, and primary care organizational leaders. Interviews with patient participants, PTs, and other health professionals have two goals: (1) to explore the experiences and perspectives with the PT-led primary care model for LBP and (2) to understand how the model of care was implemented, how the intervention interacted with its context, and barriers/facilitators to implementation. Interview guides for interviews with patient participants, PTs, and other health professionals start by exploring experiences with the PT-led primary care model for LBP using open-ended questions and probing. The interview guides then focus on asking participants about how the model of care was implemented, how the intervention interacted with its context, and barriers/facilitators to implementation using questions

XSL•FO
RenderX

constructed to align with the CFIR domains. The interview guides for medical office assistants and primary care organizational leaders will focus on how the model was implemented and contextual factors influencing implementation using the CFIR. Interview guides were prepiloted with persons with lived experience as patients and primary care team members prior to conducting the interviews.

We are using purposive sampling to recruit 8 to 12 patient participants with diversity in terms of age, gender, race, household income, LBP-related disability, LBP duration, STarT Back risk categories, and primary care site. We are inviting all PTs who are involved in implementing the PT-led primary care model for LBP to participate in an interview. We are purposively sampling 10 to 15 primary care health professionals who work with a PT in the PT-led primary care model for LBP, ensuring variation in terms of professional background, gender, and primary care site. We are using purposive sampling to identify 4 to 8 medical office assistants and 4 to 8 primary care organizational leaders who have experienced implementation of the PT-led primary care model. The concept of information power [96] related to our study objectives is being used to determine when to stop interviewing within each informant group based.

Patient participants were asked during their initial consent process for the main trial whether they were willing to be contacted for a follow-up qualitative interview about their experiences with the PT-led model of care. Willing patient participants were contacted within 12 weeks of enrollment. Based on our purposive sampling criteria, research assistants contacted potential participants who had agreed to be contacted to provide additional details about the interview purpose, review the consent process, and arrange a convenient time for the interview. Participants who agreed to take part were then sent a letter of information and a consent form in advance. At the time of the interview, the research assistant confirms that the participant has reviewed the consent form, responds to any questions, and obtains verbal consent before proceeding. Recruitment of health care providers, medical office assistants, and organizational leaders is being carried out by a study coordinator and research assistant who are familiar with the participating clinical teams. Potential nonpatient participants are being invited via email or in-person. Those who indicate interest are scheduled for an interview. As with patient participants, they will receive the letter of information and consent form ahead of the interview, and verbal consent is obtained at the start of the session after confirming their understanding and answering any questions.

Protocol and Analysis Plan Amendments

Changes to the protocol and analysis plan will be communicated by amending the trial registry at ClinicalTrials.gov and reported in the full trial publication. Investigators and participants will be communicated with as appropriate based on the changes.

Analysis

Effectiveness Analysis

All quantitative analyses will be conducted as per the Intention-To-Treat principle. Descriptive statistics for baseline characteristics and primary and secondary outcomes will be reported by arm using means (SD) or medians (IQR) for continuous variables and counts (percent) for categorical variables. All analyses will be performed in SAS (version 9.4; SAS Institute Inc). Differences between arms will be compared, accounting for site clustering using linear mixed models and generalized estimating equations (GEEs), and significance will be reported with *P* values.

Patient Health and Experience Outcomes

For our patient health outcomes, we will use linear mixed models to estimate individual patient participant outcomes adjusting for clustering by primary care center. Our primary outcome (RMDQ) with repeated measures at baseline and the 6-week, 3-, 6-, 9-, and 12-month follow-up time points will be analyzed using linear mixed regression (using PROC MIXED in SAS), estimated using restricted maximum likelihood estimation and a Kenward–Roger degrees of freedom correction [97] to account for a small number of clusters. The model will include fixed effects for time, intervention group by time interaction, (omitting the group main effect to ensure baseline differences are constrained to zero [98]), factors used in the covariate-constrained allocation procedure [99] (rurality of the cluster, number of active patients) and other prespecified covariates associated with LBP-related disability (patient participant age [100], sex [101], income [102], highest level of education achieved, duration of current episode of LBP [103], Functional Comorbidity Index score [100]). The correlation in repeated measures on the same participant will be modeled using a suitable covariance structure, identified using information criteria (Akaike and/or Bayesian information criteria). To account for clustering within practices, the site will be modeled as a random effect. The intervention effect will be obtained as the adjusted least square mean difference between arms at 12 months, with 95% CIs. Secondary comparisons will be obtained using least square mean differences at intermediate time points.

The use of restricted maximum likelihood estimation under an assumption of missing at random (MAR) allows the use of all available data without the need for multiple imputation. To examine the risk of bias due to missing data, we will compare the characteristics of those remaining and those lost to follow-up to identify factors associated with attrition. We will perform a sensitivity analysis for a missing not at random departure from our MAR assumption using a delta-adjusted imputation pattern mixture model [104,105] approach to investigate the robustness of our trial outcomes with regard to the missing values of the RMDQ. Within this sensitivity analysis approach, we will start with the posterior distributions suggested by an imputation model using multiple imputation by chained equations. Our imputation model will incorporate all variables in our primary analysis model, along with the last-observed-before-time covariates, and additional covariates needed for the pattern mixture. The sensitivity parameter (delta) will be introduced to explore how a departure from MAR affects results by specifying a maximum delta for each pattern of missing. We will set the maximum delta to be twice the residual sample SD from the observed data fit of the primary linear mixed model. Our sensitivity analysis will use multiple imputation 9 times to

XSL•FO
RenderX

generate estimated treatment effects for a range of sensitivity [106].

Pain intensity, health-related quality of life (EQ-5D-5L), PSEQ, PCS, TSK-11, and PHQ-2 outcomes will be analyzed as described for the primary outcome, adjusted for the same covariates. Patient satisfaction and GROC outcomes have no baseline measures, and will be analyzed using ordinal logistic regression with clusters as random effects, adjusting for the same covariates as described above. When individual items are missing from within any of the questionnaires, we will use simple mean imputation as suggested by Chavance [106]. Serious adverse events will be presented descriptively by arm due to low expected counts. Any adverse events (yes/no) will be presented as incidence rates with CIs and compared by calculating relative risks with CIs from robust Poisson regression using GEE-type robust variance estimators (using PROC GLIMMIX with EMPIRICAL option in SAS) to account for clustering [107] and using an exchangeable working correlation matrix. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard [108] correction to account for the small number of clusters in all models comparing incidence rates. In the case of nonconvergence or unstable estimates due to the small number of clusters, we would attempt to fit the model using an independent working correlation matrix. Other information related to nonserious adverse events (ie, severity and duration) will be reported descriptively.

We have planned a secondary analysis, a responder analysis [109], to compare the proportion of participants who experience a meaningful improvement in disability (RMDQ) in the PT-led primary care model arm versus the usual care arm. We will define a meaningful improvement as an improvement of greater than or equal to 30% improvement on the RMDQ, corresponding to an established minimally important difference among people with LBP [110,111]. We will calculate the proportion of participants who experience a meaningful improvement in each arm and compare between groups using robust Poisson regression, with GEE-type robust variance estimators to account for clustering [112]. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard [108] correction to account for small number of clusters.

Health Care Use and Missed Occupational Activity Outcomes

For health care use and missed occupational activities outcomes, we will estimate the average effect at the patient level across the population (marginal models), accounting for clustering using GEE-type robust variance estimators and robust Poisson regression with an exchangeable working correlation matrix. All of the models generated for our health care use and missed occupational activities outcomes will include the same covariates as our analysis for patient health outcomes: age, sex, highest level of education achieved, income, duration of current episode of LBP, and Functional Comorbidity Index score (individual level), and primary care site rurality and number of active patients (cluster level). In the case of nonconvergence or model instability, a possibility for any binary outcomes with very low or high event rates given our small number of clusters [113], we will attempt to model outcomes using an independence working correlation matrix. If we continue to experience issues with nonconvergence or instability with alternate covariance structures, we will remove covariates, starting with duration of pain and income, based on theoretical grounds and existing evidence on the strength of relationships between our covariates and our outcomes.

Primary care physician or nurse practitioner visits will be presented as rates and compared using rate ratios with adjusted Poisson or negative binomial regression. Other health care use within the primary care team (whether there were any other consultations with interprofessional team members and whether or not there was group programming accessed), and health care use outside of the primary care team (medications, diagnostic imaging, walk-in clinic visits, ED visits, specialist physician visits, ED visits, hospital admissions, interventional procedures, surgeries, other health provider visits) will be presented as incidence rates and compared by calculating relative risks with CIs using robust Poisson regression, accounting for clustering [107]. These models will use time as an offset to account for variable follow-up times and will incorporate empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for a small number of clusters [108,114,115]. Time (days) lost from occupational activities (paid employment, volunteer, homemaking, or educational activities) and assistance required (hours of paid assistance, hours of unpaid assistance) due to LBP will be presented as rates and compared using rate ratios with negative binomial regression.

Economic Evaluation

We will estimate the cost-effectiveness of the PT-led care model from both a societal (primary) and health system payer (secondary) perspective to meet the needs of all knowledge users. The total costs will be determined by multiplying the quantity of resource use (or lost days) by the corresponding unit cost (or hourly wage), summing the total cost over each follow-up interval, and then calculating the mean cost at each follow-up time point, as well as an overall mean cost for the entire 1-year study period. Results will be presented as aggregated and disaggregated costs. Utility data will be generated using EQ-5D-5L index values (ie, utility scores) from all follow-up assessment time points. We will estimate QALYs for every participant, using the area under the curve approach, assuming linear interpolation between the measurements. To accommodate the hierarchical structure of the data, we will use bivariate multilevel modeling to estimate the incremental cost-effectiveness ratio using a calculation of cost-per-QALY gained for PT-led primary care versus usual care [93-95] and to estimate the incremental net benefit at various willingness-to-pay values. To account for clustering of study sites, as well as heterogeneity in costs and treatment effect across jurisdictions, we will model treatment group as a fixed effect and the study site as a random effect, adjusting for the same covariates as the primary analyses as fixed effects in our models. We will conduct a probabilistic sensitivity analysis using 10,000 Monte Carlo simulations to present the uncertainty in our cost-effectiveness estimates. Simulation results will be plotted

XSL•FO
RenderX

on cost-effectiveness planes and we will generate cost-effectiveness acceptability curves to display the probability that the PT-led care model is cost-effective across a range of willingness to pay thresholds.

Subgroup Analyses

In alignment with Sex and Gender Equity in Research (SAGER) guidelines [116], we plan to conduct exploratory analyses for each of our effectiveness outcomes for males and females to explore potential sex differences in each of these outcomes. We will include sex and its interaction with time and group by time in the models. We will report the interaction *P* value along with forest plots to visualize the subgroup treatment effects, along with 95% CIs.

Process Evaluation Analysis

The multimethod process evaluation analysis will assess how the PT-led primary care model for LBP was implemented, the potential mechanisms of the model, the experiences and perspectives of patients and primary care team members toward the model, and how the context influenced implementation and outcomes. The analysis is guided by and intended to inform refinements to our program theory [48,117] for the PT-led primary care model for LBP (Figure 2).

We will assess how the PT-led primary care model and usual care model were implemented and how the model influenced health care for people with LBP by the following:

1. Describing and comparing the proportion of patient participants who received timely access to LBP care between trial arms. The intended implementation outcome is that patient participants seeking primary care for LBP receive timely access (within 48 hours) to an LBP assessment from a PT in the PT-led primary care model for LBP. Access within 48 hours will be presented as incidence rates with CIs and compared by calculating relative risks with CIs using GEE-type variance estimators to account for clustering [107], using an exchangeable working correlation matrix. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for a small number of clusters. We will incorporate the same covariates as with our effectiveness analysis.
2. Describing the proportion of patient participants in the PT-led primary care arm who saw the PT as the first point of contact for the current episode of LBP.
3. Describing and comparing the proportion of patient participants who are categorized as medium or high risk on the STarT Back tool who access PT services between trial arms. The intended implementation outcome is that patient participants with LBP at medium or high risk of ongoing pain and disability access PT services, either through a referral (for participants who have access to PT services) or through additional PT care from the primary care PT (for participants who would not otherwise have access to PT services). We will report the proportion of participants at medium or high risk receiving PT care as incidence rates with CIs and compared by calculating relative risks with CIs using GEE-type variance estimators, accounting for clustering [107], using an exchangeable working correlation matrix. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for small number of clusters [118-120]. We will incorporate the same covariates as with our effectiveness analysis.
4. Describing and comparing the LBP management provided between trial arms and comparing the care provided to practice guidelines using recent World Health Organization (WHO) LBP guidelines [121] to assess the alignment of the care provided with practice guideline recommendations. The 10 recommended interventions in the WHO guidelines are: structured education, exercise, needling therapies, spinal manipulation, massage, operant therapy, cognitive behavioral therapy, nonsteroidal anti-inflammatory drugs, topical cayenne pepper, and multicomponent biopsychosocial care. Interventions that are recommended against include: traction, ultrasound, transcutaneous electrical nerve stimulation, and lumbar supports. Additionally, referrals for diagnostic imaging or physician specialist visits for spinal injections or surgery consultations are rarely needed for LBP and thus will be compared between arms for a potential reduction in low-value care [122,123]. The intended implementation outcome is that patient participants receive care recommended in practice guidelines and do not receive care recommended against in practice guidelines, and that a low proportion of people receive referrals or prescriptions for diagnostic imaging, physician specialists, or medications outside of nonsteroidal anti-inflammatory drugs. Receipt of each intervention recommended by the guidelines, recommended against by the guidelines, and referrals (for imaging, spinal injections, or surgical consults) will be presented as incidence rates with CIs and compared by calculating relative risks with CIs using GEE-type robust variance estimates, accounting for clustering [107], using an exchangeable working correlation matrix. We will use empirical covariance ("sandwich") bias-adjusted (residual-based) estimators and the Fay and Graubard correction to account for a small number of clusters [108]. We will incorporate the same covariates as with our effectiveness analysis.
5. Describing adherence to PT primary care recommendations. The intended implementation outcome is that patient participants report actioning the initial recommendations for physical activity and exercise, and that those who were recommended to access PT services report doing so.

To assess potential mechanisms of the PT-led primary care model, if there is a significant treatment effect, we will conduct a series of mediation analyses [124] to assess whether changes in LBP-related disability are explained by changes in self-efficacy (PSEQ) or changes in psychosocial risk factors (PHQ-2, PCS, TSK-11) for persistent LBP-related disability. We will carry out this mediation analysis for each of the self-efficacy and psychosocial risk factor variables in our entire sample and in the subgroup of patient participants identified as high-risk using the STarT Back classification. We hypothesize that change in self-efficacy will mediate the relationship between the intervention arm and RMDQ score in the full patient participant sample and that a change in psychosocial risk factors

XSL•FO
RenderX

will mediate the relationship between the treatment arm and RMDQ score in the high-risk subgroup. We will use a stepwise approach to exploring the temporal trends and dynamics of the treatment effect across repeated measures as proposed by Beril et al [125], along with our theoretical insights regarding the mediation effect, when considering the appropriate mediation model. The indirect effect, or the intervention effect that can be explained by the mediator, will be determined as the difference between the total effect of the intervention and the direct effect of the intervention [126,127]. The significance of this effect can be used to statistically evaluate the possibility of mediation [128,129]. In this context, the total effect is the intervention effect obtained in the full model described in the trial. To be consistent with our primary outcome measure in the trial, we will be concerned with the time-specific mediation on the RMDQ outcome at the 12-month follow-up [130,131]. These effect measures will allow us to present the proportion of the total effect that is mediated through the respective measures of interest [132]. Development of the causal/associated conceptual model has allowed us to consider, and control for where needed, mediation analysis assumptions; that is, there is no intervention-outcome, mediator-outcome, or intervention-mediator confounding or mediator-outcome confounding that is influenced by the intervention itself [112,127,133]. Additionally, as part of exploring potential mechanisms, we will describe the proportion of any differences in total costs that are due to differences in health care use costs (and if so, what health care services) and what proportion of any differences in total costs are due to differences in costs due to missed occupational activities.

We will further explore potential mechanisms for intervention effects by conducting a planned subgroup analysis for all process outcomes, RMDQ, and costs based on STarT Back risk classification (low, medium, or high), recognizing that this analysis is exploratory and likely to be underpowered. We hypothesize that the participants in the PT-led group classified as low risk will be less likely to receive requisitions for diagnostic imaging, referrals to specialist physicians, and prescriptions for medications; demonstrate reduced costs; and show similar disability outcomes in comparison to patient participants in the usual care group. We hypothesize that participants in the PT-led group classified as medium risk will: be more likely to access PT services; receive LBP management closer aligned with guidelines; receive fewer requisitions for diagnostic imaging, prescriptions for medications, and referrals to specialist physicians; demonstrate reduced costs; and have reduced LBP-related disability. Finally, we hypothesize that participants in the PT-led group classified as high risk will: be more likely to access PT services; receive LBP management closer aligned with guidelines and more targeted care for psychosocial risk factors of persistent LBP-related disability; show greater reductions in the psychosocial risk factors; demonstrate reduced costs; and have reduced LBP-related disability.

To explore the experiences and perspectives of patients, PTs, and other primary care health professionals who have participated in the PT-led primary care model for LBP, qualitative interviews with patient participants, the PTs, and other health professional primary care team members will be recorded, transcribed, and coded independently by 2 investigators. An inductive qualitative analysis will be completed in an interpretive description tradition [118,119]. Interpretive description was chosen because of its emphasis on adding an interpretive lens during the analysis process in order identify meaningful themes from the data that can be applied in practice. This qualitative approach is therefore well aligned with our process evaluation goals to understand experiences and perspectives with the PT-led primary care model, potentially leading to refinements in the model of care, the program theory, or plans for scaling the intervention if effective. To promote rigor, we will use: 2 independent coders for the first 2 to 3 transcripts for each group (patient participants, PTs, other health professionals) and meet to reach agreement on the initial coding structure; reflexivity journaling and reflexive dialogue amongst team members throughout the analytic process; field notes and written memos; prolonged engagement within the data; and an audit trail of the research process and analytic decisions [120,134].

To explore how the context influenced the implementation of the PT-led primary care model for LBP, a concurrent mixed methods analysis will be conducted using the quantitative data collected to describe how the intervention is being implemented (analysis described above) and qualitative interview data from patient participants, PTs, other primary care health professionals, medical office assistants, and primary care organizational leaders. Qualitative analysis will begin by having multiple research team members immerse themselves in the data. Data will be coded by 2 independent research team members using an in-depth deductive (codes derived from CFIR constructs) and inductive (codes derived from the data) coding process in alignment with the CFIR User Guide [135] and previous research. Further, we will code relationships between constructs to capture how constructs interact, and how implementation determinants relate to how the model was implemented. Integration of qualitative and quantitative will be achieved through the design (convergent mixed methods), methods (merging), and interpretation (narrative and joint displays) [136]. Merging of qualitative and quantitative data will be achieved by linking participant characteristics, implementation data, and qualitative interview data. By bringing together the quantitative and qualitative data for analysis and linking this data at the level of the patient participant and cluster, we will be able to carry out an in-depth analysis of how context influenced implementation. Integration at the interpretation phase will be achieved using a weaving approach to interpret and report qualitative and quantitative findings together [136]. Joint displays will be used to bring data together in a visual display if appropriate.

Ethical Considerations

Ethics approval for this study has been obtained from the Queen's University Health Science and Affiliated Teaching Hospitals Research Ethics Board (HSREB #6027847). Written consent was obtained from all participants willing to participate. Research participants' data was collected and stored securely to maintain privacy and confidentiality. Participants were

XSL•FO
RenderX

provided with a CAD $10 (US $7.33) gift card for each completed assessment to thank them for their time.

Knowledge Dissemination

We plan to mobilize the knowledge generated through this cluster randomized trial and process evaluation through a series of peer-reviewed manuscripts, with the following foci: (1) effectiveness of the PT-led primary care model for people with LBP (including one manuscript on patient health outcomes, and one on health system and societal outcomes), (2) how the PT-led primary care model for LBP was implemented, (3) potential mechanisms through which the PT-led primary care model for LBP influences patient health and health system outcomes (if the intervention is effective), (4) experiences and perspectives of patients with LBP who participated in the PT-led primary care model, (5) experiences and perspectives of PTs who participated in implementing the PT-led primary care model for people with LBP, (6) experiences and perspectives of other primary care team members who participated in implementing the PT-led primary care model for people with LBP, and (7) how the context influenced implementation of the PT-led primary care model for people with LBP. We plan to present these results at national and international conferences on primary care, health services, physiotherapy, and back pain. We will create tailored summary reports for each manuscript for each of the following knowledge user groups: patients, health professionals, primary care team leaders, and health system decision makers.

Results

This trial was registered prospectively in January 2020 at ClinicalTrials.gov (NCT04287413). Recruitment of primary care sites (clusters) was completed in June 2023 (after being delayed due to the COVID-19 Pandemic). Cluster randomization was performed in July 2023. Patient participant recruitment was initiated in October 2023 and completed in November 2024. Last self-report patient data collected on November 25, 2025. EHR data extraction will be completed on December 19, 2025. Data analysis will be carried out in accordance with this protocol and analysis plan and will commence following the completion of data collection. No interim analyses have been conducted. Results are expected to be published in 2026.

Discussion

The anticipated primary results of this trial are a robust estimate of the effectiveness of a PT-led primary care model on disability for people seeking primary care for LBP, as well as important estimates of the impact of the model of care on cost-effectiveness, health care use patterns, primary care team visits, and a number of secondary individual outcomes.

This cluster randomized trial will build on existing evidence by providing the most robust evidence to date on the effectiveness of a PT-led primary care model for people with LBP. A recent systematic review indicated that evidence on the effectiveness of PT-led primary care models on individual health outcomes and costs was limited to low or very low quality of evidence [137].

The strengths of this protocol are the pragmatic design and the detail in which we describe the methods and prespecified analysis plan to provide transparency and reproducibility of the methods and statistical considerations. Limitations of the study design include the fact that participant recruitment occurs after randomization of clinical sites, which prevents blinding of providers and may increase the risk of revealing allocation status to eligible patient participants, introducing a potential source of selection bias. Additionally, while the study is adequately powered to detect an effect for the primary outcome, the relatively small number of clusters may limit statistical power for secondary outcomes.

The results of this trial will inform new models of primary care in Canada and will be applicable to health systems around the world. Our detailed analysis plan has integrated the perspectives of knowledge users, including people living with LBP, primary care providers, PTs, researchers, and health system decision makers. These knowledge users have helped define study outcomes (ie, disability, quality of life, and cost-effectiveness) that will meet the needs of key knowledge users and decision makers. This protocol and analysis plan builds on our trial registration by articulating our analytic decisions in advance of analyzing our data to reduce the risk of analytic or reporting bias.

Our knowledge user team has also contributed to our detailed process evaluation, where we aim to explore how the PT-led primary care model for LBP was implemented, potential mechanisms of the model of care, experiences of people involved in implementation, and how context influences implementation. The updated Medical Research Council guidance for the development and evaluation of complex interventions emphasizes the importance of moving beyond only questions of effectiveness to also explore how complex interventions, like this model of care, will be accepted, adopted, implemented, scaled, and transferred across contexts [48]. Our analysis plan for an in-depth process evaluation aims to support health system decision makers by providing clear and transparent analytic processes to answer the key questions needed to inform the scale and spread of this model of care if it is effective.

Funding

This study is funded by a Canadian Institutes of Health Research (CIHR) project grant (reference number 166177). CIHR played no part in the writing of this protocol and analysis plan. J Miller is also supported by the Arthritis Society of Canada and the CIHR STARS award, which provides financial support for him to focus more of his time on research.

XSL•FO
RenderX

Data Availability

The datasets generated and analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review documents from the Randomized Controlled Trials 2 Review Committee, Canadian Institutes of Health Research (CIHR).
[PDF File (Adobe PDF File), 1041 KB - resprot_v15i1e89004_app1.pdf]

References

1. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the global burden of disease study 2010. Lancet 2012;380(9859):2163-2196 [FREE Full text] [doi: 10.1016/S0140-6736(12)61729-2] [Medline: 23245607]
2. Low back pain. Bone and Joint Canada. 2022. URL: http://boneandjointcanada.com/low-back-pain/ [accessed 2022-08-26]
3. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J 2008;8(1):8-20. [doi: 10.1016/j.spinee.2007.10.005] [Medline: 18164449]
4. Katz JN. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. J Bone Joint Surg Am 2006;88 Suppl 2:21-24. [doi: 10.2106/JBJS.E.01273] [Medline: 16595438]
5. Gandjour A, Telzerow A, Lauterbach KW, INTERCARE International Investigators. European comparison of costs and quality in the treatment of acute back pain. Spine (Phila Pa 1976) 2005;30(8):969-975. [doi: 10.1097/01.brs.0000158944.54033.60] [Medline: 15834342]
6. Reid MC, Engles-Horton LL, Weber MB, Kerns RD, Rogers EL, O'Connor PG. Use of opioid medications for chronic noncancer pain syndromes in primary care. J Gen Intern Med 2002;17(3):173-179 [FREE Full text] [doi: 10.1046/j.1525-1497.2002.10435.x] [Medline: 11929502]
7. Singer J. A snapshot of health care in Canada as demonstrated by top 10 lists, 2011. Government of Canada Publications.: Canadian Institute of Health Information; 2011. URL: https://publications.gc.ca/site/archivee-archived.html?url=https%3A%2F%2Fpublications.gc.ca%2Fcollections%2Fcollection_2012%2Ficis-cihi%2FH115-54-2011-eng.pdf [accessed 2026-01-24]
8. Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. Spine (Phila Pa 1976) 2006;31(23):2724-2727. [doi: 10.1097/01.brs.0000244618.06877.cd] [Medline: 17077742]
9. Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. BMC Musculoskelet Disord 2010;11:144 [FREE Full text] [doi: 10.1186/1471-2474-11-144] [Medline: 20598124]
10. Côté P, Cassidy JD, Carroll L. The treatment of neck and low back pain: who seeks care? who goes where? Med Care 2001;39(9):956-967. [doi: 10.1097/00005650-200109000-00006] [Medline: 11502953]
11. Buchbinder R, van Tulder M, Öberg B, Costa LM, Woolf A, Schoene M, Lancet Low Back Pain Series Working Group. Low back pain: a call for action. Lancet 2018;391(10137):2384-2388 [FREE Full text] [doi: 10.1016/S0140-6736(18)30488-4] [Medline: 29573871]
12. Clark S, Horton R. Low back pain: a major global challenge. Lancet 2018;391(10137):2302. [doi: 10.1016/S0140-6736(18)30725-6] [Medline: 29573869]
13. Foster NE, Anema JR, Cherkin D, Chou R, Cohen SP, Gross DP, Lancet Low Back Pain Series Working Group. Prevention and treatment of low back pain: evidence, challenges, and promising directions. Lancet 2018;391(10137):2368-2383. [doi: 10.1016/S0140-6736(18)30489-6] [Medline: 29573872]
14. Freedman KB, Bernstein J. Educational deficiencies in musculoskeletal medicine. J Bone Joint Surg Am 2002;84(4):604-608. [doi: 10.2106/00004623-200204000-00015] [Medline: 11940622]
15. Day CS, Yeh AC, Franko O, Ramirez M, Krupat E. Musculoskeletal medicine: an assessment of the attitudes and knowledge of medical students at Harvard Medical School. Acad Med 2007;82(5):452-457. [doi: 10.1097/ACM.0b013e31803ea860] [Medline: 17457065]
16. Darlow B, Dean S, Perry M, Mathieson F, Baxter GD, Dowell A. Acute low back pain management in general practice: uncertainty and conflicting certainties. Fam Pract 2014;31(6):723-732 [FREE Full text] [doi: 10.1093/fampra/cmu051] [Medline: 25192904]
17. Petterson SM, Liaw WR, Phillips RL, Rabin DL, Meyers DS, Bazemore AW. Projecting US primary care physician workforce needs: 2010-2025. Ann Fam Med 2012;10(6):503-509 [FREE Full text] [doi: 10.1370/afm.1431] [Medline: 23149526]

XSL•FO
RenderX

18. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012;380(9836):37-43 [FREE Full text] [doi: 10.1016/S0140-6736(12)60240-2] [Medline: 22579043]
19. Public Health Agency of Canada. The chief public health officer's report on the state of public health in Canada 2010. In: Growing Older: Adding Life to Years. Canada: Public Health Agency of Canada; 2010.
20. Delitto A, George SZ, Van Dillen L, Whitman JM, Sowa G, Shekelle P, Orthopaedic Section of the American Physical Therapy Association. Low back pain. J Orthop Sports Phys Ther 2012;42(4):A1-57 [FREE Full text] [doi: 10.2519/jospt.2012.42.4.A1] [Medline: 22466247]
21. Koes BW, van Tulder M, Lin CW, Macedo LG, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. Eur Spine J 2010;19(12):2075-2094 [FREE Full text] [doi: 10.1007/s00586-010-1502-y] [Medline: 20602122]
22. Pillastrini P, Gardenghi I, Bonetti F, Capra F, Guccione A, Mugnai R, et al. An updated overview of clinical guidelines for chronic low back pain management in primary care. Joint Bone Spine 2012;79(2):176-185 [FREE Full text] [doi: 10.1016/j.jbspin.2011.03.019] [Medline: 21565540]
23. Low Back Pain Working Group. Evidence-informed primary care management of low back pain - clinical practice guideline, 3rd edition. Institute of Health Economics. Alberta, Canada: Toward Optimized Practice; 2017. URL: https://www.ihe.ca/public/uploaded/lbp-guideline%20(1).pdf [accessed 2025-12-01]
24. Comans T, Raymer M, O'Leary S, Smith D, Scuffham P. Cost-effectiveness of a physiotherapist-led service for orthopaedic outpatients. J Health Serv Res Policy 2014;19(4):216-223. [doi: 10.1177/1355819614533675] [Medline: 24819380]
25. Childs JD, Whitman JM, Sizer PS, Pugia ML, Flynn TW, Delitto A. A description of physical therapists' knowledge in managing musculoskeletal conditions. BMC Musculoskelet Disord 2005;6:32 [FREE Full text] [doi: 10.1186/1471-2474-6-32] [Medline: 15963232]
26. Freburger JK, Carey TS, Holmes GM, Wallace AS, Castel LD, Darter JD, et al. Exercise prescription for chronic back or neck pain: who prescribes it? who gets it? what is prescribed? Arthritis Rheum 2009;61(2):192-200 [FREE Full text] [doi: 10.1002/art.24234] [Medline: 19177524]
27. Hill JC, Whitehurst DGT, Lewis M, Bryan S, Dunn KM, Foster NE, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. Lancet 2011;378(9802):1560-1571 [FREE Full text] [doi: 10.1016/S0140-6736(11)60937-9] [Medline: 21963002]
28. Fritz JM, Cleland JA, Brennan GP. Does adherence to the guideline recommendation for active treatments improve the quality of care for patients with acute low back pain delivered by physical therapists? Med Care 2007;45(10):973-980. [doi: 10.1097/MLR.0b013e318070c6cd] [Medline: 17890995]
29. Rutten GM, Degen S, Hendriks EJ, Braspenning JC, Harting J, Oostendorp RA. Adherence to clinical practice guidelines for low back pain in physical therapy: do patients benefit? Phys Ther 2010;90(8):1111-1122. [doi: 10.2522/ptj.20090173] [Medline: 20488978]
30. Riis A, Jensen CE, Bro F, Maindal HT, Petersen KD, Bendtsen MD, et al. A multifaceted implementation strategy versus passive implementation of low back pain guidelines in general practice: a cluster randomised controlled trial. Implement Sci 2016;11(1):143 [FREE Full text] [doi: 10.1186/s13012-016-0509-0] [Medline: 27769263]
31. Tahvonen P, Oikarinen H, Niinimäki J, Liukkonen E, Mattila S, Tervonen O. Justification and active guideline implementation for spine radiography referrals in primary care. Acta Radiol 2017;58(5):586-592. [doi: 10.1177/0284185116661879] [Medline: 27609905]
32. Fritz JM, Childs JD, Wainner RS, Flynn TW. Primary care referral of patients with low back pain to physical therapy: impact on future health care utilization and costs. Spine (Phila Pa 1976) 2012;37(25):2114-2121. [doi: 10.1097/BRS.0b013e31825d32f5] [Medline: 22614792]
33. Childs JD, Fritz JM, Wu SS, Flynn TW, Wainner RS, Robertson EK, et al. Implications of early and guideline adherent physical therapy for low back pain on utilization and costs. BMC Health Serv Res 2015;15:150 [FREE Full text] [doi: 10.1186/s12913-015-0830-3] [Medline: 25880898]
34. Donato EB, DuVall RE, Godges JJ, Zimmerman GJ, Greathouse DG. Practice analysis: defining the clinical practice of primary contact physical therapy. J Orthop Sports Phys Ther 2004;34(6):284-304. [doi: 10.2519/jospt.2004.34.6.284] [Medline: 15233391]
35. Greathouse DG, Schreck RC, Benson CJ. The United States army physical therapy experience: evaluation and treatment of patients with neuromusculoskeletal disorders. J Orthop Sports Phys Ther 1994;19(5):261-266. [doi: 10.2519/jospt.1994.19.5.261] [Medline: 8199619]
36. James JJ, Stuart RB. Expanded role for the physical therapist. screening musculoskeletal disorders. Phys Ther 1975;55(2):121-131. [doi: 10.1093/ptj/55.2.121] [Medline: 124877]
37. McGill T. Effectiveness of physical therapists serving as primary care musculoskeletal providers as compared to family practice providers in a deployed combat location: a retrospective medical chart review. Mil Med 2013;178(10):1115-1120. [doi: 10.7205/MILMED-D-13-00066] [Medline: 24083925]
38. Murphy BP, Greathouse D, Matsui I. Primary care physical therapy practice models. J Orthop Sports Phys Ther 2005;35(11):699-707. [doi: 10.2519/jospt.2005.35.11.699] [Medline: 16355912]

XSL•FO
RenderX

39. Ross MD, Childs JD, Middel C, Kujawa J, Brown D, Corrigan M, et al. Physical therapist vs. family practitioner knowledge of simple low back pain management in the U.S. Air Force. Mil Med 2014;179(2):162-168. [doi: 10.7205/MILMED-D-13-00099] [Medline: 24491612]
40. Walsh NE, Halls S, Thomas R, Berry A, Liddiard C, Cupples ME, et al. First contact physiotherapy: an evaluation of clinical effectiveness and costs. Br J Gen Pract 2024;74(747):e717-e726 [FREE Full text] [doi: 10.3399/BJGP.2023.0560] [Medline: 38429110]
41. Thompson J, Macintosh F, Beaumont N, Bedford L, Powley A, Bailey S. The experiences and perceptions of first contact practitioners in primary care-a qualitative systematic review. Musculoskeletal Care 2024;22(4):e1955. [doi: 10.1002/msc.1955] [Medline: 39467815]
42. Lamb K, Comer C, Walsh N, McHugh G. Patient access to first contact practitioner physiotherapists in the UK: a national survey. Musculoskeletal Care 2023;21(4):1554-1562. [doi: 10.1002/msc.1834] [Medline: 37837408]
43. Stynes S, Jordan KP, Hill JC, Wynne-Jones G, Cottrell E, Foster NE, et al. Evaluation of the first contact physiotherapy (FCP) model of primary care: patient characteristics and outcomes. Physiotherapy 2021;113:199-208 [FREE Full text] [doi: 10.1016/j.physio.2021.08.002] [Medline: 34656297]
44. Docking S, Sridhar S, Haas R, Mao K, Ramsay H, Buchbinder R, et al. Models of care for managing non-specific low back pain. Cochrane Database Syst Rev 2025;3(3):CD015083. [doi: 10.1002/14651858.CD015083.pub2] [Medline: 40052535]
45. Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. Qual Saf Health Care 2003;12(1):47-52 [FREE Full text] [doi: 10.1136/qhc.12.1.47] [Medline: 12571345]
46. Chan A, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ 2013;346:e7586 [FREE Full text] [doi: 10.1136/bmj.e7586] [Medline: 23303884]
47. Chan AW, Boutron I, Hopewell S, Moher D, Schulz KF, Collins GS, et al. SPIRIT 2025 statement: updated guideline for protocols of randomized trials. Nat Med 2025;31(6):1784-1792. [doi: 10.1038/s41591-025-03668-w] [Medline: 40295741]
48. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, et al. A new framework for developing and evaluating complex interventions: update of medical research council guidance. BMJ 2021;374:n2061 [FREE Full text] [doi: 10.1136/bmj.n2061] [Medline: 34593508]
49. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: medical research council guidance. BMJ 2015;350:h1258 [FREE Full text] [doi: 10.1136/bmj.h1258] [Medline: 25791983]
50. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated consolidated framework for implementation research based on user feedback. Implement Sci 2022;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
51. Damschroder LJ, Reardon CM, Opra Widerquist MA, Lowery J. Conceptualizing outcomes for use with the consolidated framework for implementation research (CFIR): the CFIR outcomes addendum. Implement Sci 2022;17(1):7 [FREE Full text] [doi: 10.1186/s13012-021-01181-5] [Medline: 35065675]
52. Vader K, Donnelly C, French SD, Grady C, Hill JC, Tripp DA, et al. Implementing a new physiotherapist-led primary care model for low back pain: a qualitative study of patient and primary care team perspectives. BMC Prim Care 2022;23(1):201 [FREE Full text] [doi: 10.1186/s12875-022-01817-5] [Medline: 35948876]
53. Moulton LH. Covariate-based constrained randomization of group-randomized trials. Clin Trials 2004;1(3):297-305. [doi: 10.1191/1740774504cn024oa] [Medline: 16279255]
54. Chaudhary MA, Moulton LH. A SAS macro for constrained randomization of group-randomized designs. Comput Methods Programs Biomed 2006;83(3):205-210. [doi: 10.1016/j.cmpb.2006.04.011] [Medline: 16870302]
55. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. BMJ 2003;327(7418):785-789 [FREE Full text] [doi: 10.1136/bmj.327.7418.785] [Medline: 14525877]
56. Campbell MJ. Challenges of cluster randomized trials. J Comp Eff Res 2014;3(3):271-281 [FREE Full text] [doi: 10.2217/cer.14.21] [Medline: 24969154]
57. Eldridge S, Kerry S, Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? BMJ 2009;339:b4006 [FREE Full text] [doi: 10.1136/bmj.b4006] [Medline: 19819928]
58. Flynn TN, Whitley E, Peters TJ. Recruitment strategies in a cluster randomized trial--cost implications. Stat Med 2002;21(3):397-405. [doi: 10.1002/sim.1025] [Medline: 11813226]
59. Teerenstra S, Eldridge S, Graff M, de Hoop E, Borm GF. A simple sample size formula for analysis of covariance in cluster randomized trials. Stat Med 2012;31(20):2169-2178. [doi: 10.1002/sim.5352] [Medline: 22495809]
60. Bombardier C, Hayden J, Beaton DE. Minimal clinically important difference. low back pain: outcome measures. J Rheumatol 2001;28(2):431-438. [Medline: 11246692]
61. Campbell MK, Fayers PM, Grimshaw JM. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. Clin Trials 2005;2(2):99-107. [doi: 10.1191/1740774505cn071oa] [Medline: 16279131]

XSL•FO
RenderX

62. Miller J, Barber D, Donnelly C, French S, Green M, Hill J, et al. Determining the impact of a new physiotherapist-led primary care model for back pain: protocol for a pilot cluster randomized controlled trial. Trials 2017;18(1):526 [FREE Full text] [doi: 10.1186/s13063-017-2279-7] [Medline: 29121989]
63. Thombs BD, Benedetti A, Kloda LA, Levis B, Nicolau I, Cuijpers P, et al. The diagnostic accuracy of the patient health questionnaire-2 (PHQ-2), patient health questionnaire-8 (PHQ-8), and patient health questionnaire-9 (PHQ-9) for detecting major depression: protocol for a systematic review and individual patient data meta-analyses. Syst Rev 2014;3:124 [FREE Full text] [doi: 10.1186/2046-4053-3-124] [Medline: 25348422]
64. Hill JC, Dunn KM, Main CJ, Hay EM. Subgrouping low back pain: a comparison of the STarT back tool with the orebro musculoskeletal pain screening questionnaire. Eur J Pain 2010;14(1):83-89 [FREE Full text] [doi: 10.1016/j.ejpain.2009.01.003] [Medline: 19223271]
65. Linton SJ. Intricacies of good communication in the context of pain: does validation reinforce disclosure? Pain 2015;156(2):199-200. [doi: 10.1097/01.j.pain.0000460297.25831.67] [Medline: 25599438]
66. Pincus T, Holt N, Vogel S, Underwood M, Savage R, Walsh DA, et al. Cognitive and affective reassurance and patient outcomes in primary care: a systematic review. Pain 2013;154(11):2407-2416. [doi: 10.1016/j.pain.2013.07.019] [Medline: 23872104]
67. Gordon R, Bloxham S. A systematic review of the effects of exercise and physical activity on non-specific chronic low back pain. Healthcare (Basel) 2016;4(2):25 [FREE Full text] [doi: 10.3390/healthcare4020022] [Medline: 27417610]
68. Searle A, Spink M, Ho A, Chuter V. Exercise interventions for the treatment of chronic low back pain: a systematic review and meta-analysis of randomised controlled trials. Clin Rehabil 2015;29(12):1155-1167. [doi: 10.1177/0269215515570379] [Medline: 25681408]
69. Dahm KT, Brurberg KG, Jamtvedt G, Hagen KB. Advice to rest in bed versus advice to stay active for acute low-back pain and sciatica. Cochrane Database Syst Rev 2010;2010(6):CD007612. [doi: 10.1002/14651858.CD007612.pub2] [Medline: 20556780]
70. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
71. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
72. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. J Clin Epidemiol 2005;58(6):595-602. [doi: 10.1016/j.jclinepi.2004.10.018] [Medline: 15878473]
73. Rundell SD, Resnik L, Heagerty PJ, Kumar A, Jarvik JG. Performance of the functional comorbidity index (FCI) in prognostic models for risk adjustment in patients with back pain. PM R 2020;12(9):891-898. [doi: 10.1002/pmrj.12315] [Medline: 31901004]
74. Main CJ, Sowden G, Hill JC, Watson PJ, Hay EM. Integrating physical and psychological approaches to treatment in low back pain: the development and content of the STarT back trial's 'high-risk' intervention (StarT back; ISRCTN 37113406). Physiotherapy 2012;98(2):110-116. [doi: 10.1016/j.physio.2011.03.003] [Medline: 22507360]
75. Chiarotto A, Maxwell LJ, Terwee CB, Wells GA, Tugwell P, Ostelo RW. Roland-morris disability questionnaire and oswestry disability index: which has better measurement properties for measuring physical functioning in nonspecific low back pain? systematic review and meta-analysis. Phys Ther 2016;96(10):1620-1637. [doi: 10.2522/ptj.20150420] [Medline: 27081203]
76. Roland M, Fairbank J. The Roland-Morris disability questionnaire and the oswestry disability questionnaire. Spine (Phila Pa 1976) 2000;25(24):3115-3124. [doi: 10.1097/00007632-200012150-00006] [Medline: 11124727]
77. Haefeli M, Elfering A. Pain assessment. Eur Spine J 2006;15 Suppl 1(Suppl 1):S17-S24 [FREE Full text] [doi: 10.1007/s00586-005-1044-x] [Medline: 16320034]
78. van RM, Janssen B, Stolk E, Boye KS, Herdman M, Kennedy-Martin M. EQ-5D-5L user guide. EuroQol.: EuroQol Research Foundation; 2019. URL: https://euroqol.org/wp-content/uploads/2023/11/EQ-5D-5LUserguide-23-07.pdf [accessed 2026-01-24]
79. Finch AP, Dritsaki M, Jommi C. Generic preference-based measures for low back pain: which of them should be used? Spine (Phila Pa 1976) 2016;41(6):E364-E374 [FREE Full text] [doi: 10.1097/BRS.0000000000001247] [Medline: 26583478]
80. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, Canadian EQ-5D-5L Valuation Study Group. A time trade-off-derived value set of the EQ-5D-5L for Canada. Med Care 2016;54(1):98-105 [FREE Full text] [doi: 10.1097/MLR.0000000000000447] [Medline: 26492214]
81. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther 2009;17(3):163-170 [FREE Full text] [doi: 10.1179/jmt.2009.17.3.163] [Medline: 20046623]
82. Eldridge SM, Chan CL, Campbell MJ, Bond CM, Hopewell S, Thabane L, PAFS consensus group. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. Pilot Feasibility Stud 2016;2:64 [FREE Full text] [doi: 10.1186/s40814-016-0105-8] [Medline: 27965879]

XSL•FO
RenderX

83. Ioannidis JPA, Evans SJW, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004;141(10):781-788 [FREE Full text] [doi: 10.7326/0003-4819-141-10-200411160-00009] [Medline: 15545678]
84. Chiarotto A, Vanti C, Cedraschi C, Ferrari S, de Lima E Sà Resende F, Ostelo RW, et al. Responsiveness and minimal important change of the pain self-efficacy questionnaire and short forms in patients with chronic low back pain. J Pain 2016;17(6):707-718 [FREE Full text] [doi: 10.1016/j.jpain.2016.02.012] [Medline: 26975193]
85. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. Eur J Pain 2007;11(2):153-163. [doi: 10.1016/j.ejpain.2005.12.008] [Medline: 16446108]
86. Osman A, Barrios FX, Kopper BA, Hauptmann W, Jones J, O'Neill E. Factor structure, reliability, and validity of the pain catastrophizing scale. J Behav Med 1997;20(6):589-605. [doi: 10.1023/a:1025570508954] [Medline: 9429990]
87. Sullivan MJ, Thorn B, Haythornthwaite JA, Keefe F, Martin M, Bradley LA, et al. Theoretical perspectives on the relation between catastrophizing and pain. Clin J Pain 2001;17(1):52-64. [doi: 10.1097/00002508-200103000-00008] [Medline: 11289089]
88. Roelofs J, Goubert L, Peters ML, Vlaeyen JWS, Crombez G. The Tampa scale for kinesiophobia: further examination of psychometric properties in patients with chronic low back pain and fibromyalgia. Eur J Pain 2004;8(5):495-502. [doi: 10.1016/j.ejpain.2003.11.016] [Medline: 15324781]
89. Swinkels-Meewisse EJCM, Swinkels RAHM, Verbeek ALM, Vlaeyen JWS, Oostendorp RAB. Psychometric properties of the Tampa scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. Man Ther 2003;8(1):29-36. [doi: 10.1054/math.2002.0484] [Medline: 12586559]
90. Manea L, Gilbody S, McMillan D. A diagnostic meta-analysis of the patient health questionnaire-9 (PHQ-9) algorithm scoring method as a screen for depression. Gen Hosp Psychiatry 2015;37(1):67-75. [doi: 10.1016/j.genhosppsych.2014.09.009] [Medline: 25439733]
91. NA. OHIP schedule of benefits and fees. Government of Ontario. URL: https://www.ontario.ca/page/ohip-schedule-benefits-and-fees [accessed 2024-11-13]
92. Drummond M, Scupher M, Claxton K, Stoddard G, Torrance G. Methods for the Economic Evaluation of Health Care Programs. Oxford, UK: Oxford University Press; 2005.
93. Gomes M, Grieve R, Nixon R, Edmunds WJ. Statistical methods for cost-effectiveness analyses that use data from cluster randomized trials: a systematic review and checklist for critical appraisal. Med Decis Making 2012;32(1):209-220. [doi: 10.1177/0272989X11407341] [Medline: 21610256]
94. Gomes M, Ng ES, Grieve R, Nixon R, Carpenter J, Thompson SG. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. Med Decis Making 2012;32(2):350-361 [FREE Full text] [doi: 10.1177/0272989X11418372] [Medline: 22016450]
95. Grieve R, Nixon R, Thompson SG. Bayesian hierarchical models for cost-effectiveness analyses that use data from cluster randomized trials. Med Decis Making 2010;30(2):163-175. [doi: 10.1177/0272989X09341752] [Medline: 19675321]
96. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
97. Kenward MG, Roger JH. An improved approximation to the precision of fixed effects from restricted maximum likelihood. Comput Stat Data Anal 2009;53(7):2583-2595 [FREE Full text] [doi: 10.1016/j.csda.2008.12.013]
98. Hooper R, Forbes A, Hemming K, Takeda A, Beresford L. Analysis of cluster randomised trials with an assessment of outcome at baseline. BMJ 2018;360:k1121 [FREE Full text] [doi: 10.1136/bmj.k1121] [Medline: 29559436]
99. Hemming K, Taljaard M. Key considerations for designing, conducting and analysing a cluster randomized trial. Int J Epidemiol 2023;52(5):1648-1658 [FREE Full text] [doi: 10.1093/ije/dyad064] [Medline: 37203433]
100. Shmagel A, Foley R, Ibrahim H. Epidemiology of chronic low back pain in US adults: data from the 2009-2010 national health and nutrition examination survey. Arthritis Care Res (Hoboken) 2016;68(11):1688-1694 [FREE Full text] [doi: 10.1002/acr.22890] [Medline: 26991822]
101. George SZ, Fritz JM, Childs JD, Brennan GP. Sex differences in predictors of outcome in selected physical therapy interventions for acute low back pain. J Orthop Sports Phys Ther 2006;36(6):354-363. [doi: 10.2519/jospt.2006.2270] [Medline: 16776485]
102. Mathieu J, Roy K, Robert ME, Akeblersane M, Descarreaux M, Marchand AA. Sociodemographic determinants of health inequities in low back pain: a narrative review. Front Public Health 2024;12:1392074 [FREE Full text] [doi: 10.3389/fpubh.2024.1392074] [Medline: 39324158]
103. Bekkering GE, Hendriks HJM, van Tulder MW, Knol DL, Simmonds MJ, Oostendorp RAB, et al. Prognostic factors for low back pain in patients referred for physiotherapy: comparing outcomes and varying modeling techniques. Spine (Phila Pa 1976) 2005;30(16):1881-1886. [doi: 10.1097/01.brs.0000173901.64181.db] [Medline: 16103860]
104. Moreno-Betancur M, Chavance M. Sensitivity analysis of incomplete longitudinal data departing from the missing at random assumption: methodology and application in a clinical trial with drop-outs. Stat Methods Med Res 2016;25(4):1471-1489. [doi: 10.1177/0962280213490014] [Medline: 23698867]
105. Troendle JF, Sur A, Leifer ES, Powell-Wiley T. Sensitivity analyses for missing in repeatedly measured outcome data. Stat Med 2025;44(23-24):e70282. [doi: 10.1002/sim.70282] [Medline: 41057147]

XSL•FO
RenderX

106. Chavance M. Handling missing items in quality of life studies. Commun Stat Theory Methods 2004;33(6):1371-1383. [doi: 10.1081/sta-120030155]
107. Yelland LN, Salter AB, Ryan P. Performance of the modified poisson regression approach for estimating relative risks from clustered prospective data. Am J Epidemiol 2011;174(8):984-992. [doi: 10.1093/aje/kwr183] [Medline: 21841157]
108. Fay M, Graubard B. Small-sample adjustments for wald-type tests using sandwich estimators. Biometrics 2001;57(4):1198-1206. [doi: 10.1111/j.0006-341x.2001.01198.x] [Medline: 11764261]
109. Henschke N, van Enst A, Froud R, Ostelo RWG. Responder analyses in randomised controlled trials for chronic low back pain: an overview of currently used methods. Eur Spine J 2014;23(4):772-778 [FREE Full text] [doi: 10.1007/s00586-013-3155-0] [Medline: 24419902]
110. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, Cleeland CS, Farrar JT, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008;9(2):105-121 [FREE Full text] [doi: 10.1016/j.jpain.2007.09.005] [Medline: 18055266]
111. Ostelo RWJG, Deyo RA, Stratford P, Waddell G, Croft P, Von Korff M, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine (Phila Pa 1976) 2008;33(1):90-94. [doi: 10.1097/BRS.0b013e31815e3a10] [Medline: 18165753]
112. VanderWeele TJ. Bias formulas for sensitivity analysis for direct and indirect effects. Epidemiology 2010;21(4):540-551 [FREE Full text] [doi: 10.1097/EDE.0b013e3181df191c] [Medline: 20479643]
113. Qiu H, Cook AJ, Bobb JF. Evaluating tests for cluster-randomized trials with few clusters under generalized linear mixed models with covariate adjustment: a simulation study. Stat Med 2024;43(2):201-215. [doi: 10.1002/sim.9950] [Medline: 37933766]
114. Li P, Redden DT. Small sample performance of bias-corrected sandwich estimators for cluster-randomized trials with binary outcomes. Stat Med 2015;34(2):281-296 [FREE Full text] [doi: 10.1002/sim.6344] [Medline: 25345738]
115. Thompson JA, Leyrat C, Fielding KL, Hayes RJ. Cluster randomised trials with a binary outcome and a small number of clusters: comparison of individual and cluster level analysis method. BMC Med Res Methodol 2022;22(1):222 [FREE Full text] [doi: 10.1186/s12874-022-01699-2] [Medline: 35962318]
116. Heidari S, Babor TF, De Castro P, Tort S, Curno M. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. Res Integr Peer Rev 2016;1:2 [FREE Full text] [doi: 10.1186/s41073-016-0007-6] [Medline: 29451543]
117. Funnell SC, Rogers PJ. Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. New York: John Wiley & Sons; 2011.
118. Hunt MR. Strengths and challenges in the use of interpretive description: reflections arising from a study of the moral experience of health professionals in humanitarian work. Qual Health Res 2009;19(9):1284-1292. [doi: 10.1177/1049732309344612] [Medline: 19690208]
119. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative alternative for developing nursing knowledge. Res Nurs Health 1997;20(2):169-177. [doi: 10.1002/(sici)1098-240x(199704)20:2<169::aid-nur9>3.0.co;2-i] [Medline: 9100747]
120. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. Int J Qual Methods 2002;1(2):13-22. [doi: 10.1177/160940690200100202]
121. O'Sullivan K. Appraisal of clinical practice guideline: World Health Organization guideline for non-surgical management of chronic primary low back pain in adults in primary and community care settings. J Physiother 2024;70(2):156 [FREE Full text] [doi: 10.1016/j.jphys.2024.02.008] [Medline: 38472050]
122. Buchbinder R, Underwood M, Hartvigsen J, Maher CG. The lancet series call to action to reduce low value care for low back pain: an update. Pain 2020;161 Suppl 1(1):S57-S64 [FREE Full text] [doi: 10.1097/j.pain.0000000000001869] [Medline: 33090740]
123. Hartvigsen J, Kamper SJ, French SD. Low-value care in musculoskeletal health care: is there a way forward? Pain Pract 2022;22 Suppl 2(Suppl 2):65-70 [FREE Full text] [doi: 10.1111/papr.13142] [Medline: 36109843]
124. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression Based Approach. New York: Guilford Press; 2017.
125. Berli C, Inauen J, Stadler G, Scholz U, Shrout PE. Understanding between-person interventions with time-intensive longitudinal outcome data: longitudinal mediation analyses. Ann Behav Med 2021;55(5):476-488 [FREE Full text] [doi: 10.1093/abm/kaaa066] [Medline: 32890399]
126. VanderWeele TJ. A three-way decomposition of a total effect into direct, indirect, and interactive effects. Epidemiology 2013;24(2):224-232 [FREE Full text] [doi: 10.1097/EDE.0b013e318281a64e] [Medline: 23354283]
127. VanderWeele TJ. Mediation analysis: a practitioner's guide. Annu Rev Public Health 2016;37:17-32. [doi: 10.1146/annurev-publhealth-032315-021402] [Medline: 26653405]
128. Zhao X, Lynch JG, Chen Q. Reconsidering Baron and Kenny: myths and truths about mediation analysis. J Consum Res 2010;37(2):197-206. [doi: 10.1086/651257]
129. Preacher KJ. Advances in mediation analysis: a survey and synthesis of new developments. Annu Rev Psychol 2015;66(1):825-852. [doi: 10.1146/annurev-psych-010814-015258]

XSL•FO
RenderX

130. Selig JP, Preacher KJ. Mediation models for longitudinal data in developmental research. Res Hum Dev 2009;6(2-3):144-164. [doi: 10.1080/15427600902911247]
131. Goldsmith KA, MacKinnon DP, Chalder T, White PD, Sharpe M, Pickles A. Tutorial: the practical application of longitudinal structural equation mediation models in clinical trials. Psychol Methods 2018;23(2):191-207 [FREE Full text] [doi: 10.1037/met0000154] [Medline: 29283590]
132. Ananth CV. Proportion mediated in a causal mediation analysis: how useful is this measure? BJOG 2019;126(8):983. [doi: 10.1111/1471-0528.15691] [Medline: 30893492]
133. Richiardi L, Bellocco R, Zugna D. Mediation analysis in epidemiology: methods, interpretation and bias. Int J Epidemiol 2013;42(5):1511-1519 [FREE Full text] [doi: 10.1093/ije/dyt127] [Medline: 24019424]
134. Holloway I, Galvin K. Qualitative Research in Nursing and Healthcare. New Jersey: John Wiley & Sons; 2016.
135. Reardon CM, Damschroder LJ, Ashcraft LE, Kerins C, Bachrach RL, Nevedal AL, et al. The consolidated framework for implementation research (CFIR) user guide: a five-step guide for conducting implementation research using the framework. Implement Sci 2025;20(1):39 [FREE Full text] [doi: 10.1186/s13012-025-01450-7] [Medline: 40819067]
136. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
137. Abuhl B, Ehrmantraut D, Wolden M. First-contact physical therapy compared to usual primary care for musculoskeletal disorders: a systematic review and meta-analysis of randomized controlled trials. Phys Ther 2025;105(8). [doi: 10.1093/ptj/pzaf080] [Medline: 40511847]

Abbreviations

CFIR: Consolidated Framework for Implementation Research
ED: emergency department
EHR: electronic health record
GEE: generalized estimating equation
GROC: global rating of change
LBP: low back pain
MAR: missing at random
PCS: Pain Catastrophizing Scale
PHQ-2: 2-item Patient Health Questionnaire
PSEQ: Pain Self-Efficacy Questionnaire
PT: physiotherapist
QALY: quality-adjusted life year
REDCap: Research Electronic Data Capture
RMDQ: Roland-Morris Disability Questionnaire
SAGER: Sex and Gender Equity in Research
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trial
TSK-11: Tampa Scale of Kinesiophobia 11
WHO: World Health Organization

Edited by J Sarvestan; submitted 04.Dec.2025; this is a non–peer-reviewed article;accepted 19.Jan.2026; published 26.Feb.2026.

Please cite as:
Miller J, Donnelly C, McClintock C, Varette K, Camargo Y, Marsh J, Taljaard M, Mamun MSA, Bacchus G, Barber D, Cooper L, French S, Hill J, Green M, MacDermid J, Norman K, Richardson J, Tranmer J, Wideman T
Determining the Impact of a Physiotherapist-Led Primary Care Model for Low Back Pain: Protocol and Analysis Plan for a Cluster Randomized Controlled Trial and Embedded Process Evaluation
JMIR Res Protoc 2026;15:e89004
URL: https://www.researchprotocols.org/2026/1/e89004
doi:10.2196/89004
PMID:

©Jordan Miller, Catherine Donnelly, Chad McClintock, Kevin Varette, Yeimi Camargo, Jacquelyn Marsh, Monica Taljaard, Mir Sanim Al Mamun, Geneviève Bacchus, David Barber, Lynn Cooper, Simon French, Jonathan Hill, Michael Green, Joy MacDermid, Kathleen Norman, Julie Richardson, Joan Tranmer, Timothy Wideman. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 26.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and

XSL•FO
RenderX

reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Updating the Good Reporting of a Mixed Methods Study (GRAMMS) Reporting Guidelines: Protocol for a Methodological Review and Modified Delphi Process

Sarah Munce[1], PhD; Sergi Fàbregues[2], PhD; Quan Nha Hong[3,4], PhD; Alicia O'Cathain[5], PhD; Timothy Guetterman[6], PhD; John Creswell[6], PhD; Cheryl Poth[7], PhD; Katie Dainty[8], PhD; Andrea C Tricco[9,10], PhD; Mandy M Archibald[11,12], PhD; Clementine Jarrett[13]; Jennifer Boyle[1], PhD; Dorothy Luong[1], MSc; Frank Kiwanuka[14], PhD; Ahtisham Younas[15,16], PhD

[1]Holland Bloorview Kids Rehabilitation Hospital, Bloorview Research Institute, Toronto, ON, Canada
[2]Department of Psychology and Education, Universitat Oberta de Catalunya, Barcelona, Spain
[3]School of Rehabilitation, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada
[4]Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal, Montréal, QC, Canada
[5]School of Medicine and Population Health, University of Sheffield, South Yorkshire, United Kingdom
[6]Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States
[7]Faculty of Education, University of Alberta, Edmonton, AB, Canada
[8]Research and Innovation, North York General Hospital, Toronto, ON, Canada
[9]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada
[10]Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada
[11]College of Nursing, University of Manitoba, Winnipeg, MB, Canada
[12]Children's Hospital Research Institute of Manitoba, University of Manitoba, Winnipeg, MB, Canada
[13]Carleton University, Ottawa, ON, Canada
[14]Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland
[15]Memorial University of Newfoundland, St. John's, NL, Canada
[16]Biruni University, Istanbul, Turkey

Corresponding Author:
Sarah Munce, PhD
Holland Bloorview Kids Rehabilitation Hospital
Bloorview Research Institute
150 Kilgour Road
Toronto, ON, M4G 1R8
Canada
Phone: 1 (416) 425 6220
Email: smunce@hollandbloorview.ca

Abstract

Background: In mixed methods research (MMR), researchers combine elements of qualitative and quantitative methodologies, methods of data collection and analysis, viewpoints, and integration procedures to gain a deeper understanding of what is being studied, design culturally specific tools, and explore the conditions under which health care interventions succeed or fail. Integration is considered the hallmark of MMR and can occur at various levels, such as sampling, data collection, and analysis. MMR is particularly useful for investigating complex, multilevel programs and interventions and is well-suited to address research problems involving knowledge translation, program evaluations, or comparisons of therapeutic interventions. Although there are many potential benefits of mixed methods in health research, the extent to which mixed methods studies implement integration remains limited, with this specific gap persisting for almost 20 years. The Good Reporting of a Mixed Methods Study (GRAMMS) reporting guidelines were developed in 2008 to help improve the quality of reporting in mixed methods reports and articles. Since then, the field of mixed methods has evolved rapidly, and the guideline no longer reflects current practices and innovations.

Objective: The objective of this study is to develop an updated GRAMMS 2.0 guideline. This project aims to develop the GRAMMS 2.0 guideline and checklist to improve consistency, transparency, and quality of reporting on studies that use mixed methods in health services research. The specific research objectives of this protocol are to (1) examine the extent of mixed

XSL•FO
RenderX

methods methodological literature and identify relevant reporting quality criteria for inclusion in GRAMMS 2.0 (ie, a methodological review), and (2) using the results of the first objective, prioritize the components of the updated GRAMMS 2.0 (ie, modified Delphi approach and consensus meeting).

Methods: This study will follow established methodological frameworks for reporting guideline development and will include a methodological review followed by a modified Delphi process.

Results: The project has received funding from the Canadian Institutes of Health Research in April 2025.

Conclusions: The GRAMMS 2.0 guideline will improve the consistency, transparency, and quality of reporting of mixed methods studies by researchers and multiple knowledge users in policy and practice. Peer reviewers and editors may also use GRAMMS 2.0 to improve the review of manuscripts involving MMR. Ultimately, the updated guideline will increase the clarity of mixed methods research findings, thereby improving their potential transferability to practice and facilitating efficient use of new results in mixed methods in health research, bringing better returns on research investments.

International Registered Report Identifier (IRRID): PRR1-10.2196/82364

(JMIR Res Protoc 2026;15:e82364) doi:10.2196/82364

KEYWORDS

mixed-methods research; methodological review; modified delphi process; guideline development; evidence-informed guidelines

Introduction

Health research often involves investigating complex processes and systems using quantitative (ie, numerical) and qualitative (eg, participant experiences) research approaches [1,2]. The specific research question or objective drives the selection of research methods [3,4]. Quantitative methodologies are often used to address questions related to causality, efficacy, effectiveness, generalizability, size of effects, and prediction. Qualitative methodologies are typically used to understand individual experiences or perspectives, explain why or how a phenomenon occurs, or develop hypotheses or theories [5]. Mixed methods research (MMR) is defined as "the type of research in which researchers combine elements of qualitative and quantitative methodologies, methods of data collection and analysis, viewpoints, and integration procedures to gather viewpoints, breadth and depth of understanding, and corroboration of a particular phenomenon" [6]. Mixed methods research leverages the strengths of both quantitative and qualitative approaches and is an essential methodology for studying the "wicked problems" defined as "those that involve multiple interacting systems, are replete with social and institutional uncertainties, and for which only imperfect knowledge about their nature and solutions exist" [7]. Such problems are inherent in health research due to the intricacy of health care systems and service delivery. The explicit mixing or linking of qualitative and quantitative components within a mixed methods study allows researchers to answer questions with greater depth and understanding than possible with one methodology alone [3,4]. This explicit integration is increasingly necessary as health system challenges are becoming more complex, requiring systems-based solutions inclusive of various data types [8]. Mixed methods research is particularly useful for investigating complex, multilevel programs and interventions [3,8,9] and is well-suited to address research problems involving program evaluations, comparisons of therapeutic interventions, or implementation [10,11].

Integration refers to "purposeful interdependence between different sources, methods, approaches" [12]. Integration lies at the heart of MMR [9,13-15], and can occur at multiple levels or dimensions through theory, conceptual models, design, methods, results, analysis, interpretation, visualization, presentation, publication, and teams [15]. Meaningful integration allows researchers to achieve the goal of mixed methods to "produce a whole through integration that is greater than the sum of the individual qualitative and quantitative parts" [16]. For example, qualitative inquiry can be used to provide greater depth to the quantitative findings by integrating qualitative and quantitative results (ie, merging) to examine complementarity and expansion. Quantitative results can also be used to generate the qualitative sample (ie, connecting). Qualitative inquiry can lead to developing or refining quantitative instruments (ie, building) or interventions or generating hypotheses in the qualitative component for testing in the quantitative component [3,5,17,18]. Joint displays, which provide a visual means for implementing and reporting on mixed methods integration, including generating metainferences, are increasingly seen as an area of innovation for advancing integration. However, joint displays in health research remain underused [19-21]. This, along with the emergence of new integrative modalities in MMR (eg, timelines) [22], highlights the imperative for their inclusion in MMR reporting.

Researchers conducting mixed methods studies also need to decide on the timing of the qualitative and quantitative components (ie, convergent or sequential data collection and analysis) and each method's intent [3,4,23]. For example, Stuber and colleagues [24] used a sequential explanatory design, where the quantitative component was first conducted and was followed by interviews and focus groups (qualitative) to explain initial survey results (quantitative) in a study of patient perceptions of patient-driven care in chiropractic practice. In addition to the mixing or linking of 2 unique research approaches—qualitative and quantitative—mixed methods studies may also involve data transformation (ie, converting qualitative data into quantitative data ("quantitizing") or vice versa ("qualitizing") [9]. As such, mixed methods studies can become complex investigations that require additional time and resources and a team of researchers with expertise in

XSL•FO
RenderX

quantitative, qualitative, and mixed methodologies, specifically [25].

In 2015, Fetters and Freshwater [16] put forward the 1+1=3 challenge, encouraging health services researchers and others interested in MMR to consider the question, "What synergy was gained by the additional work of using both qualitative and quantitative data methods?" The authors urged researchers to plan their studies judiciously with purposeful choices that leverage integration. Although there are many potential benefits of MMR in health services research, the extent to which mixed methods studies provide or report evidence of integration remains limited, with this specific gap in the field persisting for almost 20 years [3,26-28]. This gap remains despite substantial progress in specific areas of integration including integration strategies, assessment of fit of integration, metainferences, and joint displays, as well as other critical aspects, such as methodological quality, reporting of cross-cultural mixed methods studies, intersectional mixed methods, sex/gender-based analysis, considerations of equity and diversity reporting in MMR, and arts-based MMR [3,12,17,18,20,25,29-36]. Moreover, quality guidance and reporting criteria do not reflect the current developments in integration.

Recent reviews of empirical studies in specific fields (eg, primary care, palliative and end-of-life care, and chiropractic services) have examined the mixed methods literature and highlighted areas for improvement in the quality of reporting (ie, whether integration procedures are described and whether evidence of integration is provided) [8,27,33,37-42]. For instance, a review of health services research found that authors of mixed methods studies typically did not describe or justify the need for a mixed methods design or integrate data and findings from the individual quantitative and qualitative components [5,41]. This lack of integration prevents readers from knowing what new insights have been generated by combining quantitative and qualitative components (ie, beyond the results obtained from the two separate components), thereby limiting the methodological potential of this research strategy [5].

The review of health services research formed the basis of the Good Reporting of a Mixed Methods Study (GRAMMS) in 2008 [5]. GRAMMS includes the following components: (1) describe the justification for using a mixed methods approach to the research question; (2) describe the design in terms of purpose, priority, and sequence of methods; (3) describe each method in terms of sampling, data collection, and analysis; (4) describe where integration has occurred, how it has occurred, and who has participated in it; (5) describe any limitation of one method associated with the presence of the other method; and (6) describe any insights gained from mixing or integrating methods. Despite the foundational contribution of the GRAMMS, its application in health research is underused. GRAMMS is more than 15 years old and needs to be updated to better reflect current practices and innovations in the field of MMR. Indeed, the field has advanced substantially [37,43] in specific areas of integration, including integration strategies [3,12,17,18,27,30,31], assessment of fit of integration and meta-inferences [3,25,30,32], joint displays [20,33,34], and other critical aspects, such as methodological quality and reporting [35,37], generalization [36,44], as well as cross-cultural MMR [36]. More critically, some of the GRAMMS criteria require additional detail and attention, especially the criterion related to integration (criterion 4). This criterion does not consider recent advances in this area, and since integration is an essential aspect of an MMR study, more than one criterion on this topic is necessary. Indeed, recently Hirose and Creswell [45] put forward a shortlist of quality criteria in MMR that highlighted the need for much more detail related to integration, reflecting the advances in the field, including mixed methods diagrams to depict points of integration, integration in the form of a joint display, and meta-inferences. They also highlighted that other important criteria include the skills of researchers, the use of philosophy, and the use of literature. There is also increasing recognition that evaluation occurs within a complex set of cultural dimensions, which mixed methods are ideally suited to address and could be featured in the future iteration of the GRAMMS [46]. Mixed methods research is increasingly collaborative, and efforts to advance collaboration in MMR have also had associated reporting criteria associated with it [47]. Furthermore, the criterion regarding limitations of one method related to the presence of the other method (criterion 5) is hardly present in most studies, as indicated by several methodological reviews, which may indicate that researchers do not know how to think about MMR limitations and tend to think about generic limitations of qualitative and quantitative research [8,35].

It is essential to close these gaps to leverage the true benefits of MMR in health research. Developing an updated GRAMMS guideline and checklist based on the Enhancing the Quality and Transparency of Health Research (EQUATOR) methodology (ie, involving a methodological review of quality reporting criteria in MMR as well as a modified Delphi process) is one way to close this gap and is essential to promote quality reporting of mixed methods studies in health research and interdisciplinary fields so that researchers, practitioners, decision makers, and other relevant knowledge users have access to high quality MMR study findings to inform practice and policy making. This proposal outlines the aim to develop the GRAMMS 2.0 guideline and checklist to improve the consistency, transparency, and quality of reporting of studies that use mixed methods in health research.

Methods

Study Design

The development of GRAMMS 2.0 will follow the established methodological framework for reporting guideline development, as outlined by the EQUATOR Network [48] and the Guidance for Developers of Health Research Reporting Guidelines [49]. The EQUATOR Network is an international initiative that aims to improve the quality of health research publications by promoting transparent and accurate reporting [48]. The Network's website contains a comprehensive and up-to-date database of reporting guidelines relevant to health research. As part of this process, we have registered our GRAMMS 2.0 proposal on the EQUATOR Network website to ensure

XSL•FO
RenderX

alignment with international standards and enhance transparency.

We will also adopt an integrated knowledge translation (iKT) approach, engaging a diverse panel that includes patient partners with experience in MMR, as well as researchers, journal editors, and funders. This panel will be guided by the Ontario Brain Institute's "Ways Community Members Can Participate in the Stages of Research" framework [50] to ensure meaningful engagement throughout the project. To further support inclusive and respectful collaboration, we will use the reflective exercise titled "Strategy for Patient-Oriented Research Evidence Alliance Reflective EDI Exercise" to encourage open dialogue and understanding around equity, diversity, and inclusion topics [51]. These discussions will be conducted in a way that values diverse perspectives and promotes trust, respect, and inclusion. The iKT panel will play an active role in various stages of the project, including shaping the literature search strategy, reviewing preliminary findings, and planning how the final GRAMMS 2.0 guideline and checklist will be shared and applied. The Guidance for Reporting Involvement of Patients and Public (GRIPP2) will guide the reporting of our patient engagement efforts for the iKT panel as well as participation in study phases [52].

Study Phases

Phase 1: Methodological Review of Quality Reporting Criteria in Mixed Methods Research

We will conduct a methodological review based on the guidelines for these types of reviews proposed in the literature [53-55] to identify and summarize existing literature on the quality of reporting of mixed methods research. This approach is well-suited for summarizing a broad range of methodological sources, including MMR systematic methodological reviews as well as discussion papers focused on quality reporting in MMR across various disciplines. The goal is to systematically identify and analyze key themes and proposed reporting criteria, providing a comprehensive overview of the current evidence in the field. Our review will specifically address the question: "What quality reporting criteria are proposed in the methodological literature on mixed methods?" This protocol paper will serve as the published protocol for the scoping review component of the study. To support transparency and accessibility, all review data will be made publicly available on the OSF. The completed PRISMA-ScR checklist is shown in Multimedia Appendix 1.

Selection Criteria and Search Strategy

We will apply the following inclusion criteria: (1) mixed methods methodological literature published in 2008 or later when the GRAMMS was first published; and (2) systematic, scoping, prevalence, or methodological reviews of empirical mixed methods studies examining the quality of reporting in those studies or discussion papers about quality reporting in MMR. We will also consider any existing MMR reporting tools. We will include any review and discussion paper regardless of the discipline or field that has examined the methodological reporting quality of mixed methods research studies, and the publication language. For reviews published in languages other than English, French, Spanish, and Finnish, DeepL [56] will be used to determine whether they meet the inclusion criteria. The use of DeepL for the initial screening of non-English publications may introduce minor screening inaccuracies due to translation limitations; however, we will mitigate this risk by having a fluent speaker of the respective language review the translation of any of these articles to ensure accuracy. We will exclude methodological literature that is not peer-reviewed, such as editorials, commentaries, mixed methods guides, books, and book chapters, to ensure that the included evidence has undergone scholarly peer review and presents sufficient methodological transparency for critical appraisal. However, to mitigate the risk of omitting foundational MMR guidance that may appear in seminal books or chapters, we will scan the reference lists of included reviews to identify any such works. If these sources are recurrently cited as influential, we will describe them narratively to contextualize the peer-reviewed evidence base. We will search the following electronic databases from 2008: Medline, CINAHL, Scopus, Embase, Cochrane, PsycINFO, ERIC, Web of Science, JSTOR, Sociological Abstracts, and ECONLit.

A librarian with expertise in knowledge synthesis will design a structured search strategy using both subject headings and relevant text words related to MMR, quality, reporting, and reviews. The search strategy will be peer-reviewed by another librarian using the Peer Review of Electronic Search Strategies (PRESS) checklist [57] as well as reviewed by our iKT panel. In addition, to capture relevant literature not indexed in these databases, we will spot-check key discipline-specific journals, including the *Journal of Mixed Methods Research* and the *International Journal of Multiple Research Approaches*. This approach ensures that foundational and highly relevant methodological articles are not missed.

Screening Process and Data Extraction

All screening (ie, levels 1 and 2) will occur in duplicate and independently. For level 1 screening, reviewers will screen the titles and abstracts for inclusion using the screening form. A pilot test of the level 1 screening form on a random sample of approximately 100 articles will be conducted to increase reliability. The descriptions of the eligibility criteria will be revised if necessary by the team or if a low agreement (ie, < 70%) is observed to improve the consistent application of the selection criteria. For level 2, the full text of potentially relevant articles will be collected and screened to determine final inclusion. A pilot test of the level 2 screening form will also be performed on approximately 1% of the articles, similar to the process for level 1 screening. Another reviewer knowledgeable in the research area will be available to resolve conflicts when necessary.

Extracted data will include the first author's name, year of publication, number of studies included, types of mixed methods designs, time frame (in years) of the studies included in the article, first author's institution's country affiliation, type of review, geographical focus, population focus, the components of the GRAMMS addressed, components of the quality reporting not addressed by the GRAMMS, primary findings related to reporting quality, limitations of each review, and

XSL•FO
RenderX

recommendations for reporting of mixed methods studies. Additional categories will likely be identified through the completion of the search and discussions with the research team, including the iKT Panel. The data extraction form will be pilot-tested and refined with at least two research team members. Like the level 1 and 2 screening processes, data extraction will occur in duplicate and independently. Discussion or the involvement of another reviewer will resolve disagreements. Covidence will be used to manage the records and data throughout the review.

Quality Assessment

Two reviewers will use the Critical Appraisal Tool for Health Promotion and Prevention Reviews (CAT HPPR) to assess all systematic, rapid, and scoping reviews. This tool is well-established to evaluate the quality of reviews and meta-analyses and evaluates the reviews in terms of review question, search strategy, risk of bias assessment, data extraction and synthesis, and reporting quality [58]. Although the tool is relatively new and was not specifically designed for methodological or scoping reviews, it was selected because its domains align with core quality criteria relevant to assessing the transparency, rigor, and reproducibility of review processes. For scoping reviews, which are not typically subject to formal critical appraisal, the CAT-HPPR will be applied in a descriptive rather than evaluative manner to identify methodological strengths and gaps. Any adaptations or interpretive decisions will be documented and reported to ensure transparency.

The methodological review is expected to begin in Fall 2025 and is anticipated to conclude within one year. The results of this methodological review will be used to inform the surveys implemented in Phase 2, as well as the draft GRAMMS 2.0 guideline. The review's findings will help identify which items need to remain, be modified, or be added to the GRAMMS 2.0. Findings will be published in a peer-reviewed publication, as well as disseminated via a newsletter and webinar to promote engagement for Phase 2.

Phase 2: Modified Delphi Process to Prioritize Items for the GRAMMS 2.0 Guideline

To prioritize the items for inclusion in the GRAMMS 2.0 guideline, we will use a modified Delphi process, a well-established method for building consensus in the development of reporting guidelines [49]. This process involves a series of surveys administered over multiple rounds to gather expert opinions and refine the list of items based on collective feedback. Participants are typically asked to rate or rank the importance of various items, with each round helping to narrow the list and move toward consensus. The modified Delphi differs from the conventional Delphi process in that it begins the process with preselected items drawn from earlier work, rather than using the experts to brainstorm on a particular subject [59]. In line with the approach used to develop other reporting guidelines [60,61], we anticipate conducting 3 rounds of the Delphi process. Each round of the Delphi process will identify items with strong agreement and reduce the number of items based on participant responses and expert consensus. A virtual consensus meeting will then be held following the Delphi. The ACcurate COnsensus Reporting Document (ACCORD) will be used to report the findings from the modified Delphi approach and consensus meeting [62].

Identification of Participants for the Modified Delphi Process and Virtual Consensus Meeting

An international, interdisciplinary, and diverse group of patients, funders, policymakers, researchers, and journal editors with experience or expertise in MMR will be invited to participate in the modified Delphi process. Participants will be identified through various sources, including authors of published MMR studies, authors of mixed methods books, instructors of mixed methods courses, and leadership and members of the Mixed Methods International Research Association. We will also use social media outlets, including X, Facebook, Threads, Blue Sky, and LinkedIn. Recognizing and addressing the appropriate use of sex and gender terms in reporting guidelines has been noted as an important gap [63]. As such, we are committed to ensuring that GRAMMS 2.0 reflects best practices in this area by including perspectives from a broad range of groups and ensuring proper representation throughout the Delphi process [63].

We will use a purposive and snowball sampling strategy to achieve the recommended minimum sample size of 100-150, consistent with other reporting guideline initiatives [60]. Potential expert participants will be invited to participate by email, including an explanatory statement about the initiative. Participants who may have challenges accessing technology will be provided with the means to participate, for example, by conducting surveys over the phone, through Zoom, or by mailing the surveys.

Survey and Administration for Rounds 1-3

For Round 1, the survey will be based on the results of the proposed methodological review, as described above. The initial list of GRAMMS 2.0 guideline items and the survey will be reviewed with the research team; they will be pilot-tested for content and clarity and revised before dissemination. Only those participants who complete the survey in Round 1 will be invited to participate in the survey in Round 2. Similarly, participants must complete the Round 2 survey to be invited for the Round 3 survey. Demographic and descriptive information about the participants will be collected (eg, sex, gender identity, race, country of residence, expert participant type, career stage, etc) to ensure a diverse sample. However, these items will be optional as some people may feel uncomfortable reporting them.

Survey administration will last 4 weeks per round, with reminder emails sent every 7 days after the initial invitation to optimize participation. For all rounds, participants will be asked to rate their agreement with the inclusion of each of the proposed guideline items using a 5-point Likert scale (from 1="strongly disagree" to 5="strongly agree"). Surveys will be administered in English. Each survey item will include an optional text box for participants to provide comments, including suggestions for any additions, deletions, aggregation, and refinement of items (eg, there will be opportunities to retain items in a modified form).

XSL•FO
RenderX

Survey Analysis

Demographic and descriptive information of the participants will be analyzed with proportions expressed as percentages (eg, sex, gender identity, country of residence, race, knowledge user type, career stage, etc). For all 3 rounds, a threshold of 70% agreement will be applied for each guideline item to indicate a consensus among the expert participants. That is, at least 70% of participants must mostly or entirely agree (ie, values of 4 or 5 on the Likert scale) with the inclusion or exclusion of the guideline item as part of GRAMMS 2.0. This threshold is consistent with developing other reporting guidelines, such as the PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) [60]. If less than 70% of participants mostly or entirely agree with the item's inclusion or exclusion, it will be considered discrepant. Persistent discrepant items will be brought forward for structured discussion during the consensus meeting. Where possible, we will investigate whether item agreement differs by sex, gender, other aspects of diversity, and knowledge user type (ie, chi-square analysis).

When applicable, content analysis [64] will be conducted on the comments. These results will be summarized and used to inform subsequent surveys. Furthermore, for the Round 2 and 3 surveys, participants will be given their results and the overall group distribution, median, and IQR from the previous survey rounds (ie, Round 1 and 2 surveys). These data will promote transparency and encourage careful reflection on items where consensus has not been reached. In addition, as part of this exercise, in the Round 2 and 3 surveys, we will ask participants to respond to the question, "After reviewing your survey results for this item (ie, each discrepant item), please comment on why you rated this item the way you did." Content analysis [64] will be conducted on these comments.

Consensus Meeting

Two half-day virtual consensus meetings will be held after completing the 3 rounds of surveys. Approximately 25-30 expert participants who completed the survey in Rounds 1-3 will be invited to participate in the consensus meeting. Consideration will be given to include 25-30 expert participants with a diverse range of characteristics, for example, sex, gender, country, expertise in MMR, that is, mixed methods trainees, as well as senior-career mixed methods researchers, and discipline/subdiscipline. A skilled facilitator will lead the consensus meeting, generate meeting minutes, and compile the discussion to produce a summary of results. The facilitator will be an independent moderator in order to ensure balanced participation. A summary of the results from Round 3 and any accompanying recommendations will be presented. These data will be shared to generate and inform the discussion on any critical discrepant items (ie, key issues and considerations relating to both items and wording). Participants will be asked to elaborate on why they rated a discrepant item the way they did. To mitigate potential dominance of academic or research voices, the iKT panel will co-develop a plan to share with the facilitator to promote equitable engagement during the consensus meeting. This plan will specify how discussions, feedback, and decision-making will occur and will outline concrete facilitation strategies to ensure balanced participation and prevent overrepresentation of any single expert group.

A draft version of the GRAMMS 2.0 guideline will be created by the research team with the iKT panel and then presented for discussion at the consensus meeting. The expert participants will discuss these proposed items and reach a consensus on which items should be included in the final version of GRAMMS 2.0 and their specific wording. As part of this exercise, we will also discuss the potential recommendation of a core set of items of the guideline and checklist (including a discussion of how each guideline item will be linked to a checklist item) and the possible weighting of items, if appropriate/feasible. As such, we will consider developing a short and long version of GRAMMS 2.0.

Furthermore, we will discuss how the GRAMMS 2.0 guideline and checklist can promote diverse perspectives and the reporting of these diverse perspectives (eg, use of culturally specific paradigms). We will also use the consensus meeting to discuss the overall dissemination of our results. We will conduct a comprehensive recording (ie, digital recording as well as note-taking) of all the discussions related to the GRAMMS 2.0 guideline and checklist development, including providing a record of the decisions made during the meeting. Content analysis, performed in duplicate, will be used to analyze the digital recording of the meeting [64]. Notes taken at the meeting will be used to supplement the recordings and aid the analysis. After the consensus meeting, the GRAMMS 2.0 guideline and checklist will be distributed to the expert participants to ensure they reflect the decisions made.

Finalization of GRAMMS 2.0

The final GRAMMS 2.0 guideline and checklist will be tested with approximately 15 researchers/scientists, decision-makers, and students by applying it to at least one of their MMR for relevance and feasibility (ie, persons not involved in the Delphi). Following the testing, the research team with iKT panel members will meet in person to discuss and finalize GRAMMS 2.0.

The final GRAMMS 2.0 guideline and accompanying checklist will also be posted on the EQUATOR Network's website. To promote the uptake of GRAMMS 2.0, we will create a 2-minute video outlining how to operationalize each item, webinars for key organizations that conduct MMR, and a 1-page infographic. We may publish the GRAMMS 2.0 guideline and checklist in multiple journals simultaneously to promote uptake and dissemination. The first author's (SM's) website will provide updates on the project and promote dissemination. We will endeavor to locate and email authors who have registered a mixed methods study to disseminate GRAMMS 2.0. We will also explore automated tools to promote the uptake of GRAMMS 2.0, including an email system in which authors would be sent the GRAMMS 2.0 upon registering or publishing their mixed methods protocol. Penelope.ai is another online tool that could be leveraged, which verifies manuscripts, for example, provides comprehensive reporting, and provides authors with feedback. It should be noted that these specific strategies have been adopted with the PRISMA-ScR [60] guideline and have contributed to it being cited >23,000 times since 2018.

XSL•FO
RenderX

Additionally, we will share the GRAMMS 2.0 guideline and checklist widely within our networks.

Ethical Considerations

Research Ethics Board approval for this component of the study will be sought at Holland Bloorview Kids Rehabilitation Hospital (ie, the lead investigator's primary institution).

Results

This work is supported by a Project Grant from the Canadian Institutes of Health Research (CIHR) received in April 2025 (202409PJT-527857-HS1-CEAS-178733; Multimedia Appendix 2). An iKT panel, including patient partners with experience in mixed methods research, as well as researchers, journal editors, and funders, will be established to contribute to various stages of the project. We anticipate that Phase 1 of the study will begin in Fall 2025 and be completed within one year. Ethics approval for Phase 1 is not required as it will not involve collecting or using data from participants. Ethics approval for Phase 2 of the study will be sought in Fall 2026. Phase 2, including the modified Delphi process and consensus meeting, is anticipated to be completed by 2027.

Discussion

Principal Findings

The collective results of this project will lead to the development of a ready-to-implement GRAMMS 2.0 guideline that outlines a minimum set of items to include in health research using mixed methods.

Challenges and Mitigating Strategies

Recruiting diverse participants for our modified Delphi process may be difficult. To address this challenge, we will leverage our association with and support of the Mixed Methods International Research Association and review our recruitment material across a range of demographic and geographical diversity biases. We will also invite mixed methods researchers who have worked with or are known to the research team. To ensure that members of the modified Delphi process and consensus panel are not limited in their participation due to a lack of access to devices and technology (eg, individuals from rural and remote communities), we will mail paper copies of the surveys on request. Finally, we also acknowledge that holding half-day consensus meetings may be challenging for some participants. Organizing meetings across time zones will also be challenging. We will ask our participants about their preferences regarding scheduling and timing, and, if preferred, hold the consensus meeting for a few hours, but over a couple of days/weeks. We will also make accommodations for those with vision and hearing impairments (eg, closed captioning).

Conclusions

The GRAMMS 2.0 guideline will improve the consistency, transparency, and quality of reporting of mixed methods studies by researchers and multiple knowledge users in policy and practice and across different disciplines. A key strength of this work is leveraging the existing GRAMMS reporting guideline as well as the expertise of its developer (AOC). Peer reviewers and editors will also use the GRAMMS 2.0 to improve the review of manuscripts involving MMR. Ultimately, the GRAMMS 2.0 guideline will increase the clarity and credibility of MMR findings, increasing their potential for practical application and contributing to the continued advancement of mixed methods methodology. In turn, this will facilitate more effective use of new results in mixed methods in health research and maximize the return on investment.

Funding

This work is supported by a Project Grant from the Canadian Institutes of Health Research (CIHR) received in April 2025 (202409PJT-527857-HS1-CEAS-178733).

Data Availability

The data collected during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: SM (lead), AY (equal)
Methodology: SM (lead), AY (equal), DL (supporting), CJ (supporting), JB (supporting)
Project administration: SM (lead), DL (supporting), CJ (supporting), JB (supporting)
Writing – original draft: SM (lead), AY (equal)
Writing – review & editing: All authors
Supervision: SM (lead)

Conflicts of Interest

SM is the Editor-in-Chief of *JMIR Rehabilitation and Assistive Technologies*. None declared by the other authors.

Multimedia Appendix 1
PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist.

XSL•FO
RenderX

[PDF File (Adobe PDF File), 178 KB - resprot_v15i1e82364_app1.pdf]

Multimedia Appendix 2 [PDF File (Adobe PDF File), 72 KB - resprot_v15i1e82364_app2.pdf]

References

1. Creswell JW, Fetters MD, Ivankova NV. Designing a mixed methods study in primary care. Ann Fam Med 2004;2(1):7-12 [FREE Full text] [doi: 10.1370/afm.104] [Medline: 15053277]
2. Curry LA, Krumholz HM, O'Cathain A, Plano Clark VL, Cherlin E, Bradley EH. Mixed methods in biomedical and health services research. Circ Cardiovasc Qual Outcomes 2013;6(1):119-123 [FREE Full text] [doi: 10.1161/CIRCOUTCOMES.112.967885] [Medline: 23322807]
3. Fetters MD, Molina-Azorin JF. New requirements to include the methodological contribution in articles published in the Journal of Mixed Methods Research. J Mix Methods Res 2019;13(2):138-142. [doi: 10.1177/1558689819834753]
4. Creswell JW, Clark VLP. Designing and Conducting Mixed Methods Research. United States: SAGE Publications Inc; 2017.
5. O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. J Health Serv Res Policy 2008;13(2):92-98. [doi: 10.1258/jhsrp.2007.007074] [Medline: 18416914]
6. Johnson RB, Onwuegbuzie AJ, Turner LA. Toward a definition of mixed methods research. J Mix Methods Res 2007;1(2):112-133. [doi: 10.1177/1558689806298224]
7. Mertens DM. Mixed methods and wicked problems. J Mix Methods Res 2014;9(1):3-6 [FREE Full text] [doi: 10.1177/1558689814562944]
8. Emary PC, Stuber KJ, Mbuagbaw L, Oremus M, Nolet PS, Nash JV, et al. Quality of reporting using good reporting of a mixed methods study criteria in chiropractic mixed methods research: a methodological review. J Manipulative Physiol Ther 2023;46(3):152-161 [FREE Full text] [doi: 10.1016/j.jmpt.2023.11.004] [Medline: 38142381]
9. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
10. Palinkas LA, Aarons GA, Horwitz S, Chamberlain P, Hurlburt M, Landsverk J. Mixed method designs in implementation research. Adm Policy Ment Health 2011;38(1):44-53 [FREE Full text] [doi: 10.1007/s10488-010-0314-z] [Medline: 20967495]
11. Palinkas L, Cooper B. Dissemination and Implementation Research in Health: Translating Science to Practice. Oxford, England: Oxford University Press; 2017.
12. Bazeley P. Integrating Analyses in Mixed Methods Research. Thousand Oaks: SAGE Publications Ltd; 2017.
13. Fielding NG. Triangulation and mixed methods designs: data integration with new research technologies. J Mix Methods Res 2012;6(2):124-136. [doi: 10.1177/1558689812437101]
14. Guetterman TC, Manojlovich M. Grand rounds in methodology: designing for integration in mixed methods research. BMJ Qual Saf 2024;33(7):470-478. [doi: 10.1136/bmjqs-2023-016112] [Medline: 38575310]
15. Fetters MD, Molina-Azorin JF. The starts a new decade: the mixed methods research integration trilogy and its dimensions. J Mix Methods Res 2017;11(3):291-307. [doi: 10.1177/1558689817714066]
16. Fetters MD, Freshwater D. The 1 + 1 = 3 integration challenge. J Mix Methods Res 2015 Apr 09;9(2):115-117. [doi: 10.1177/1558689815581222]
17. Pluye P, Bengoechea EG, Granikov V, Kaur N, Tang DL. A world of possibilities in mixed methods: review of the combinations of strategies used to integrate qualitative and quantitative phases, results and data. IJMRA 2018;10(1):41-56. [doi: 10.29034/ijmra.v10n1a3]
18. Tunarosa A, Glynn MA. Strategies of integration in mixed methods research: insights using relational algorithms. Organ Res Methods 2016 Jun 22;20(2):224-242. [doi: 10.1177/1094428116637197]
19. Guetterman TC, Fetters MD, Creswell JW. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann Fam Med 2015;13(6):554-561 [FREE Full text] [doi: 10.1370/afm.1865] [Medline: 26553895]
20. Guetterman TC, Fàbregues S, Sakakibara R. Visuals in joint displays to represent integration in mixed methods research: a methodological review. Methods Psychol 2021 Dec;5:100080. [doi: 10.1016/j.metip.2021.100080]
21. Younas A, Pedersen M, Durante A. Characteristics of joint displays illustrating data integration in mixed-methods nursing studies. J Adv Nurs 2020;76(2):676-686. [doi: 10.1111/jan.14264] [Medline: 31713252]
22. Archibald MM, Mazzon SJ, Bourque MA. Examining the integrative potential of visual timelines in mixed methods research. J Mix Methods Res 2024;19(3):287-303. [doi: 10.1177/15586898241303319]
23. Moseholm E, Fetters MD. Conceptual models to guide integration during analysis in convergent mixed methods studies. Methodol Innov 2017;10(2):205979911770311. [doi: 10.1177/2059799117703118]
24. Stuber KJ, Langweiler M, Mior S, McCarthy PW. A pilot study assessing patient-centred care in patients with chronic health conditions attending chiropractic practice. Complement Ther Med 2018;39:1-7. [doi: 10.1016/j.ctim.2018.05.006] [Medline: 30012379]

XSL•FO
RenderX

25. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Kolner Z Soz Sozpsychol 2017;69(Suppl 2):107-131 [FREE Full text] [doi: 10.1007/s11577-017-0454-1] [Medline: 28989188]
26. Bryman A. Integrating quantitative and qualitative research: how is it done? Qual Res 2006;6(1):97-113. [doi: 10.1177/1468794106058877]
27. Lee SD, Iott B, Banaszak-Holl J, Shih S, Raj M, Johnson KE, et al. Application of mixed methods in health services management research: a systematic review. Med Care Res Rev 2022;79(3):331-344. [doi: 10.1177/10775587211030393] [Medline: 34253078]
28. Lewin S, Glenton C, Oxman AD. Use of qualitative methods alongside randomised controlled trials of complex healthcare interventions: methodological study. BMJ 2009;339(sep10 1):b3496-b3496 [FREE Full text] [doi: 10.1136/bmj.b3496] [Medline: 19744976]
29. Åkerblad L, Seppänen-Järvelä R, Haapakoski K. Integrative strategies in mixed methods research. J Mix Methods Res 2020;15(2):152-170. [doi: 10.1177/1558689820957125]
30. Hitchcock J, Onwuegbuzie A. The Routledge Handbook for Advancing Integration in Mixed Methods Research. Milton Park: Routledge; 2022.
31. Younas A, Durante A. Decision tree for identifying pertinent integration procedures and joint displays in mixed methods research. J Adv Nurs 2023;79(7):2754-2769. [doi: 10.1111/jan.15536] [Medline: 36524303]
32. Younas A, Fàbregues S, Creswell JW. Generating metainferences in mixed methods research: a worked example in convergent mixed methods designs. Methodological Innovations 2023;16(3):276-291. [doi: 10.1177/20597991231188121]
33. Haynes-Brown TK, Fetters MD. Using joint display as an analytic process: an illustration using bar graphs joint displays from a mixed methods study of how beliefs shape secondary school teachers' use of technology. Int J Qual Methods 2021;20:160940692199328. [doi: 10.1177/1609406921993286]
34. McCrudden MT, Marchand G, Schutz PA. Joint displays for mixed methods research in psychology. Methods Psychol 2021;5:100067. [doi: 10.1016/j.metip.2021.100067]
35. Fàbregues S, Hong QN, Escalante-Barrios EL, Guetterman TC, Meneses J, Fetters MD. A methodological review of mixed methods research in palliative and end-of-Life care (2014–2019). IJERPH 2020;17(11):3853. [doi: 10.3390/ijerph17113853]
36. Younas A, Durante A. The logics of and strategies to enhance generalization of mixed methods research findings. Methodology 2023 Jun 30;19(2):170-191. [doi: 10.5964/meth.10863]
37. Guetterman TC, Molina-Azorin JF, Fàbregues S. The need to rigorously develop common quality guidelines for reporting mixed methods research. J Mix Methods Res 2022;17(1):6-11. [doi: 10.1177/15586898221143561]
38. Kaur N, Vedel I, El Sherif R, Pluye P. Practical mixed methods strategies used to integrate qualitative and quantitative methods in community-based primary health care research. Fam Pract 2019 Oct 08;36(5):666-671 [FREE Full text] [doi: 10.1093/fampra/cmz010] [Medline: 30888015]
39. Irvine FE, Clark MT, Efstathiou N, Herber OR, Howroyd F, Gratrix L, et al. The state of mixed methods research in nursing: a focused mapping review and synthesis. J Adv Nurs 2020;76(11):2798-2809. [doi: 10.1111/jan.14479] [Medline: 32896959]
40. Younas A, Pedersen M, Tayaben J. Review of mixed-methods research in nursing. Nurs Res 2019;68(6):464-472. [doi: 10.1097/NNR.0000000000000372] [Medline: 31693552]
41. Wisdom JP, Cavaleri MA, Onwuegbuzie AJ, Green CA. Methodological reporting in qualitative, quantitative, and mixed methods health services research articles. Health Serv Res 2012;47(2):721-745 [FREE Full text] [doi: 10.1111/j.1475-6773.2011.01344.x] [Medline: 22092040]
42. de Haan M, van Eijk-Hustings Y, Vrijhoef HJ. Using mixed methods in health services research: a review of the literature and case study. J Health Serv Res Policy 2021 Apr;26(2):141-147. [doi: 10.1177/1355819620955223] [Medline: 32957813]
43. Younas A, Durante A, Fàbregues S. Understanding the nature of and identifying and formulating "research problems" in mixed methods research. J Mix Methods Res 2023 Jul 25;18(4):483-502. [doi: 10.1177/15586898231191441]
44. Hong QN, Fàbregues S. A critical reflection of generalization in mixed methods research. Eval Rev 2025 [FREE Full text] [doi: 10.1177/0193841x251331723]
45. Hirose M, Creswell JW. Applying core quality criteria of mixed methods research to an empirical study. J Mix Methods Res 2022 May 02;17(1):12-28. [doi: 10.1177/15586898221086346]
46. Poth CN, editor. The Sage Handbook of Mixed Methods Research Design. California: Sage Publications; 2023.
47. Archibald MM. Investigator triangulation: a collaborative strategy with potential for mixed methods research. J Mix Methods Res 2015;10(3):228-250. [doi: 10.1177/1558689815570092]
48. Enhancing the quality and transparency of health research. EQUATOR Network. 2025. URL: https://www.equator-network.org/ [accessed 2025-06-04]
49. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. PLoS Med 2010;7(2):e1000217 [FREE Full text] [doi: 10.1371/journal.pmed.1000217] [Medline: 20169112]
50. Ways community members can participate in the stages of research. Ontario Brain Institute. 2025. URL: https://braininstitute.ca/img/Patient-engagement-in-the-stages-of-research.pdf [accessed 2025-06-04]
51. Reflective exercise. SPOR Evidence Alliance. URL: https://sporevidencealliance.ca/wp-content/uploads/2023/02/4.-SPOREA_Reflective-EDI-Exercise-UPDATED_2021.pdf [accessed 2025-06-04]

XSL•FO
RenderX

52. Staniszewska S, Brett J, Simera I, Seers K, Mockford C, Goodlad S, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. BMJ 2017 Aug 02;358:j3453 [FREE Full text] [doi: 10.1136/bmj.j3453] [Medline: 28768629]
53. Fàbregues S, Younas A, Escalante-Barrios EL, Molina-Azorin JF, Vázquez-Miraz P. Toward a framework for appraising the quality of integration in mixed methods research. J Mix Methods Res 2024;18(3):270-280. [doi: 10.1177/15586898241257555]
54. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol 2018;18(1):5. [doi: 10.1186/s12874-017-0468-4] [Medline: 29316881]
55. Aguinis H, Ramani RS, Alabduljader N. Best-practice recommendations for producers, evaluators, and users of methodological literature reviews. Organ Res Methods 2020;26(1):46-76. [doi: 10.1177/1094428120943281]
56. DeepL translate: the world's most accurate translator. URL: https://www.deepl.com/en/translator [accessed 2025-07-14]
57. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016;75:40-46 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
58. Heise TL, Seidler A, Girbig M, Freiberg A, Alayli A, Fischer M, et al. CAT HPPR: a critical appraisal tool to assess the quality of systematic, rapid, and scoping reviews investigating interventions in health promotion and prevention. BMC Med Res Methodol 2022;22(1):334 [FREE Full text] [doi: 10.1186/s12874-022-01821-4] [Medline: 36567381]
59. Custer RL, Scarcella JA, Stewart BR. The modified delphi technique - a rotational modification. J Career Tech Educ 1999;15(2):00. [doi: 10.21061/jcte.v15i2.702]
60. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
61. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71 [FREE Full text] [doi: 10.1136/bmj.n71] [Medline: 33782057]
62. Patient and public partner appreciation policy and protocol. SPOR Evidence Alliance. 2022. URL: https://sporevidencealliance.ca/wp-content/uploads/2022/01/SPOREA_Patient-and-Public-Appreciation-Policy_2021.01.14-1.pdf [accessed 2025-12-12]
63. Gogovor A, Zomahoun HTV, Ekanmian G, Adisso EL, Deom Tardif A, Khadhraoui L, et al. Sex and gender considerations in reporting guidelines for health research: a systematic review. Biol Sex Differ 2021;12(1):62 [FREE Full text] [doi: 10.1186/s13293-021-00404-0] [Medline: 34801060]
64. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]

Abbreviations

ACCORD: ACcurate COnsensus Reporting Document
CAT HPPR: Critical Appraisal Tool for Health Promotion and Prevention Reviews
EQUATOR: Enhancing the Quality and Transparency of Health Research
GRAMMS: Good Reporting of a Mixed Methods Study
GRIPP2: Guidance for Reporting Involvement of Patients and Public
iKT: integrated knowledge translation
MMR: mixed methods research
PRESS: Peer Review of Electronic Search Strategies
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews

Edited by J Sarvestan; submitted 14.Aug.2025; this is a non–peer-reviewed article;accepted 24.Nov.2025; published 17.Mar.2026.

Please cite as:
Munce S, Fàbregues S, Hong QN, O'Cathain A, Guetterman T, Creswell J, Poth C, Dainty K, Tricco AC, Archibald MM, Jarrett C, Boyle J, Luong D, Kiwanuka F, Younas A
Updating the Good Reporting of a Mixed Methods Study (GRAMMS) Reporting Guidelines: Protocol for a Methodological Review and Modified Delphi Process
JMIR Res Protoc 2026;15:e82364
URL: https://www.researchprotocols.org/2026/1/e82364
doi:10.2196/82364
PMID:

XSL•FO
RenderX

©Sarah Munce, Sergi Fàbregues, Quan Nha Hong, Alicia O'Cathain, Timothy Guetterman, John Creswell, Cheryl Poth, Katie Dainty, Andrea C Tricco, Mandy M Archibald, Clementine Jarrett, Jennifer Boyle, Dorothy Luong, Frank Kiwanuka, Ahtisham Younas. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Objective Assessment of Financial Decision-Making With a Simulated Online Money Management Task in Older Adults: Protocol for a Prospective Observational Study

Preeti Sunderaraman[1,2], PhD; Madison Bouchard-Liporto[3], BA; Zachary Kunicki[1], MS, MPH, PhD; Edward D Huey[1], MD; Stephanie Cosentino[4,5,6], PhD

[1]Department of Psychiatry and Human Behavior, Alpert Medical School, Brown University, 222 Richmond Street, Providence, RI, United States

[2]Memory and Aging Program, Butler Hospital, 345 Blackstone Boulevard, Providence, RI, United States

[3]Department of Neurology, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA, United States

[4]Taub Institute for Research on Alzheimer's Disease and the Aging Brain, Columbia University Irving Medical Center, New York, NY, United States

[5]The Gertrude H. Sergievsky Center, College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY, United States

[6]Department of Neurology, College of Physicians and Surgeons, Columbia University and the New York Presbyterian Hospital, New York, NY, United States

Corresponding Author:
Preeti Sunderaraman, PhD
Department of Psychiatry and Human Behavior, Alpert Medical School, Brown University, 222 Richmond Street, Providence, RI, United States

Abstract

Background: Technology-enabled tasks to conduct financial transactions are ubiquitous around the world. In a recent survey, about 75% of the respondents endorsed the use of technology to perform financial activities such as reviewing bank statements and keeping track of money spent. However, assessment of financial decision-making (FDM) is limited by tasks that use traditional paper-and-pencil methods or by relying on self or informant reports. Furthermore, such tools have weak psychometric properties, are prone to biases, and are at times limited in scope. Thus, there is an urgent need to develop modern, technology-based tools that have strong psychometric properties and that can assess FDM comprehensively and accurately.

Objective: This study aimed to develop and establish the psychometric properties of a novel, simulated Online Money Management (OMM) credit card task. Based on existing gaps identified in the literature, this task relied on objective measurement, assessed multiple dimensions within a single task, and mimicked a real-world task to bridge the gap between a controlled, clinical setting and real-world functioning.

Methods: This was a prospective cohort study that enrolled cognitively healthy older adults. This study was funded by the National Institutes of Health. Various recruitment sites were involved, which allowed for the recruitment of older adults across the United States. The OMM task was developed in collaboration with an interdisciplinary team of computer scientists, economists, psychologists, and geriatricians. The tasks consist of both online and offline components, with subcomponents examining the ability to navigate, basic and complex credit card literacy, and statement monitoring. Data about participants' perception of their financial abilities and a self-report survey on financial exploitation were collected. The test battery consisted of an array of cognitive, financial, and psychosocial tasks. Participants provided written informed consent, and all procedures received institutional review board approval.

Results: Data collection began in September 2019, and enrollment stopped in July 2025. A total of 272 participants completed the baseline visit, while 147 completed the longitudinal follow-up visit. Data analysis is underway as of August 2025, and results are expected to be published in 2026.

Conclusions: Rigorous standards have been deployed for developing this novel OMM credit card task. If the measurement properties of the task are found adequate, the OMM task can be used to assess FDM in clinical evaluations for early detection and prevention or mitigation of financial mismanagement.

International Registered Report Identifier (IRRID): DERR1-10.2196/82488

(JMIR Res Protoc 2026;15:e82488) doi:10.2196/82488

KEYWORDS

financial management; decision-making; neuropsychology; technology; cognition; functional status; older adults

XSL•FO
RenderX

Introduction

Older adults have faced economic hardship in part due to the COVID-19 pandemic and the rising rates of inflation [1]. The level of debt incurred by this group has tripled [2]. According to the National Council on Aging, 15 million households for older adults aged 60 years and older have no assets with a median income of US $18,000 in 2020 [3]. Additionally, age-related cognitive decline, including that associated with Alzheimer disease (AD), increases susceptibility to poor decision-making [4,5], with older adults exhibiting compromised financial decision-making (FDM) in areas such as investment choices [6,7], bankruptcy claims [8], and credit card behaviors [4]. Additionally, older adults are exploited by family members, friends, and strangers more frequently than younger adults [9-13]. Compromised self-awareness can further impact FDM [14,15]. It is therefore critical to be able to identify deteriorating FDM abilities objectively (without relying on self or informant report) in older adults to prevent potentially devastating financial losses [12,16-18]. Early identification can also empower older adults and their care partners to make choices with the goal of preserving wealth.

Use of technology to conduct financial activities has become ubiquitous, even among older adults [19]. In a recent study specifically designed to assess financial habits, the majority of respondents (≈75%) endorsed the use of technology (eg, online services accessed via a computer) to perform financial activities such as reviewing bank statements and keeping track of money spent [19]. The youngest baby boomers, born in 1964 and currently aged 59 years, are expected to be tech-savvy in similar ways to younger cohorts [20,21]. Questions pertaining to online financial management abilities will become increasingly relevant as the number of baby boomers continues to increase precipitously. Simultaneously, there will be a corresponding increase in age-related disorders such as dementia [8,20,22,23]. Timely and modern ways of assessing FDM abilities in vulnerable individuals are critical to preventing and mitigating catastrophic financial loss.

However, there is a lack of instrumentation to evaluate one's ability to manage finances using modern approaches [24]. Existing instruments are limited because they rely on subjective reports (prone to biases) [25,26] or use traditional, paper-and-pencil methods (eg, checkbook management), which are less relevant to modern, electronic financial management [24,27,28]. Using the COSMIN (Consensus-Based Standards for the Selection of Health Measurement Instruments), a recent review concluded that available instruments had weak psychometric properties [27]. The US Social Security Administration found that current approaches to FDM assessment (eg, clinical interview) were grossly inadequate in predicting real-world FDM and underscored the need to develop sound tools [24,27]. Therefore, the proposed study aimed to develop a simulated Online Money Management (OMM) credit card task to assess FDM and financial awareness in typical cognitive aging.

The process leading to the development of the OMM task and the K99/R00 project is depicted in Figure 1. Based on survey data collected during the principal investigator's (PI)'s National Institutes of Health (NIH)–funded F32 project, 61% of baby boomers reported using OMM methods, and 67% reported that using technology was a change from how they previously performed financial transactions[19]. Analysis for different types of financial habits revealed that 65% of older adults were using technology to pay bills, and 80% to check their credit card and bank statements[19]. Conceptual models of FDM have not yet taken OMM methods into account—this is understandable, as the financial landscape has undergone dramatic changes only over the past decade, and some of the influential FDM models were developed before the digital technology proliferation. The 3 leading conceptual models that served as the basis for the OMM task are detailed in the following paragraphs.

The person-centered model [25] focuses on a single impending financial transaction using a person-centered approach that views the individual as a whole with unique strengths, weaknesses, and experiences. It seeks to merge a person's decisional abilities (choice, understanding, appreciation, and reasoning) with financial awareness of the situation to detect risk of exploitation or fraud. The Lichtenberg Financial Decision-Making Rating Scale [25] and Lichtenberg Financial Decision-Making Rating Screening Scale [28] have been derived from this model, and these instruments are still undergoing psychometric validation [13,28-30]. Though comprehensive in many respects, a limitation of these instruments is that they are based on self-reports, and information gleaned is difficult to objectively verify and/or may be invalid in those with impaired self-awareness or memory problems.

XSL•FO
RenderX

Figure 1. Flowchart depicting the process leading up to the development of the OMM credit card task. OMM: Online Money Management.

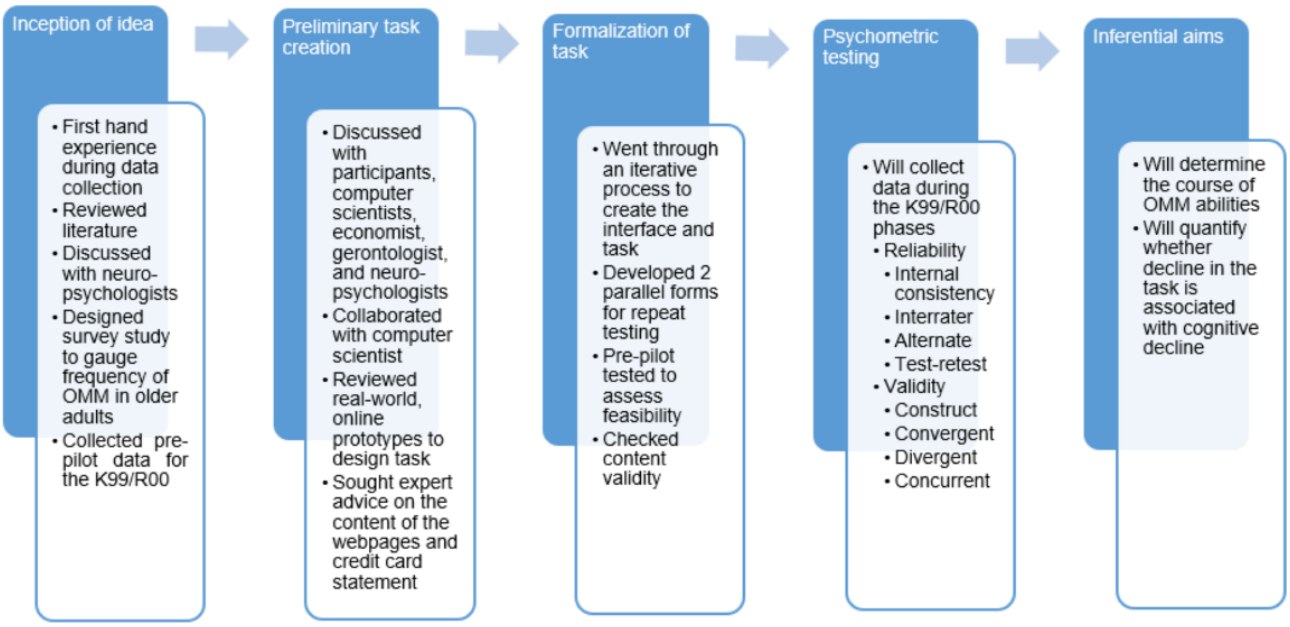

From the clinical-cognitive model [31-34], primarily based on a clinical perspective and borrowing concepts from cognitive psychology, this model conceptualizes financial skills to range from basic (identify and count coins or currency) to complex (make investment decisions) that are necessary for independent functioning. Per the model, financial capability comprises performance, procedural, knowledge or declarative, and judgment-based dimensions [35]. Several instruments, such as the Financial Capacity Instrument – Short Form (FCI-SF), have been derived from this model [35,36]. This model acknowledges that the 3 FDM dimensions may interact, for example, creating a will or investing in a bond requires a combination of all 3 abilities [31]. Thus, financial assessment should ideally include an examination of all 3 dimensions (declarative, procedural, and judgment-based) within a single task. A recent exploratory factor analysis identified 4 factors in the FCI-SF: basic monetary knowledge and calculation skills (eg, calculate a tip), financial judgment (eg, identify mail fraud), financial conceptual knowledge (eg, define debt), and financial procedural knowledge (eg, write a check) [37]. While this model highlights the role of performance-based financial assessment, a limitation of the instrument is that it favors breadth over depth in that it does not fully explore any single dimension of FDM, such as conceptual knowledge and its manipulation.

From the Institute of Medicine Model [24], it and the National Academy of Sciences convened an expert panel for the US Social Security Administration to conduct an exhaustive review of literature and develop a conceptual model of financial capability. Building on the existing models described above, the newly proposed model distinguished between financial competence (typically assessed in a controlled, clinical setting) vs financial performance (real-world functioning). They further divided financial competence into knowledge and judgment and emphasized that to understand real-world functioning, it is important to consider environmental and personal factors (together referred to as “contextual” factors). Personal factors refer to concepts intrinsic to the person (eg, religion, substance addiction, and mental health issues), and environmental factors involve exogenous facets (eg, support system, access to autopay, and access to bank account and financial products). These factors can either impede or facilitate real-world financial performance. The committee concluded that a gap exists between financial competence and financial performance. The OMM task attempts to bridge this gap by embedding experimental, contextual manipulations of specific FDM components (eg, financial knowledge and financial awareness) within a simulated real-world task.

Currently, only a handful of online tools exist to examine FDM. These include the functional skills and assessment training [38], the banking app [39], and the web-based evaluation of banking skills [40]. However, none of these tools comprehensively assesses credit card management. The OMM credit card task will begin to fill in the lacunae in existing FDM models and instruments by (1) measuring abilities directly and objectively, (2) assessing multiple clinical or cognitive dimensions within a single FDM task, and (3) mimicking a real-world task to bridge the gap between a controlled, clinical setting and real-world functioning. The OMM task is developed to make FDM assessment more accurate. It can be used by any trained and qualified professional, including social workers, occupational therapists, physicians, and psychologists. Ultimately, this task will aid in the detection and prevention or mitigation of financial mismanagement and loss experienced by older adults.

The K99/R00 study had four specific aims: (1) to evaluate the reliability, factor structure, and validity of the OMM credit card task in cognitively healthy older adults; (2) to determine the natural course of overall OMM and specific OMM components over a span of 2 years; (3) to examine differences in OMM across cognitively healthy older adults and those with clinical AD; and (4) to examine the nature and strength of the association between financial exploitation and the OMM task in pathological aging.

XSL•FO
RenderX

Methods

Study Design

This was a prospective cohort study enrolling a convenience sample of older adults. This study was funded by the NIH, and the grant review summary statement is included in Peer Review Report 1.

Study Population and Setting

Cognitively healthy older adults and 50 participants with a clinical diagnosis of AD were expected to complete the OMM task. They were to be recruited from the Alzheimer's Disease Research Center (ADRC) at Columbia University Medical Center (K99 phase) and Boston University School of Medicine (R00 phase). Additional sources of recruitment included the Aging and Dementia Clinic in the Department of Neurology at Columbia University Irving Medical Center (K99 phase) and the Saint Elizabeth's Hospital in Brighton, MA; Boston Medical Center (BMC) in Boston, MA; and from ResearchMatch, which allowed for recruitment of older adults across the United States (R00 phase).

Ethical Considerations

This study was approved by the Columbia University Medical Center Institutional Review Board (IRB; K99 phase; protocol number 19 - 00364) and by the Boston University IRB (R00 phase; protocol number 66322). Across these institutes, the IRBs determined that this research involves no greater than minimal risk. Study participants provided written informed consent. Participant data are stored in a secure manner using encrypted and password-protected devices and software. Participants were paid US $50 via gift card upon completion of every study time interval they participated in (baseline, 3-month follow, and longitudinal follow-up).

Eligibility Criteria

Study-specific inclusion and exclusion criteria are detailed in Textbox 1. Aside from the clause about the absence of a clinical diagnosis and independent living, all other criteria apply to those recruited with clinical AD.

Textbox 1. Inclusion and exclusion criteria for Aims 1 and 2.

Inclusion criteria

- Living independently in the community.
- Able to read and understand the consent form.
- Native English speaker.
- Reading level above fifth grade.
- Comfortable navigating a computer.
- Manages bank statements or credit card accounts online.

Exclusion criteria

- Montreal Cognitive Assessment score ≤26.
- Diagnosis of neurological disorder.
- Diagnosis of past or current psychotic disorder.
- Current or recent (past 5 years) history of major depressive disorder, bipolar disorder, or anxiety disorder.
- Recent use (past month) of recreational drugs.
- Diagnosed with a learning disability, dyslexia, or attention deficit hyperactivity disorder.

Due to the COVID-19 pandemic, this study procedure had to be initially ramped down and then adapted to use remote testing procedures. The following is a list of study procedures that were changed:

1. The K99 aims could not be fully examined in the proposed timeline of two years due to the circumstances posed by the pandemic. Therefore, baseline data collection on cognitively healthy older adults continued in the R00 phase.
2. During the K99 phase, the PI proposed collecting data on two alternate forms of the OMM credit card statement. Although an alternate version of the OMM credit card task was developed, data collection did not begin due to difficulty recruiting participants online for this study's sessions during the COVID-19 pandemic.
3. Referral rates for cognitively healthy older adults were low. Consequently, individuals were recruited from the community using flyers posted at public locations (eg, cafes) and through flyers distributed on online websites such as Research Match.
4. Despite recruiting through multiple sources, referrals of patients with AD were infrequent. Therefore, aim 3 of this study comparing the performance of cognitively healthy participants with those having clinical AD was not conducted.

Study Assessments and Outcome Measures

The OMM Credit Card Task was the primary outcome variable. This task was developed iteratively, in collaboration with a computer scientist and other experts (Figure 1). As shown in Figure 2, the task consists of 2 broad areas (navigation and

XSL•FO
RenderX

content) and 4 specific content-related aspects (simple and complex literacy, monitoring, and awareness). Screenshots of the webpages for this task are depicted in Figure 3. For the task, the participant first navigates 3 webpages (login page, accounts activity page, and view statements page; Figure 4) to access the credit card statement. The participants have to navigate these pages sequentially. They are able to progress to the next page (eg, from login to accounts activity page) only after they successfully enter the required information or click on the correct icon. Second, they answer 7 questions one at a time about the statement by clicking on the simulated pages (credit card basic literacy) and next identify erroneous transactions (statement monitoring). Third, they answer 9 questions about credit card management (complex credit card literacy) and, finally, rate their level of confidence about their performance before and after each task (self-awareness component). After completing the task, participant feedback about their perceived difficulty and comfort while performing this task is recorded. Performance metrics for this task include the number of clicks, accuracy, and total time to completion. Additional metrics are being developed and refined. Before each component of the task, participants are provided with practice to click on the responses or shown a demonstration as per the task requirements.

Figure 2. Composition of the OMM task**.** There are 2 broad components of the task: Content and Navigation, which are theoretically independent. The Content component consists of 4 theoretically related subcomponents. OMM: Online Money Management.

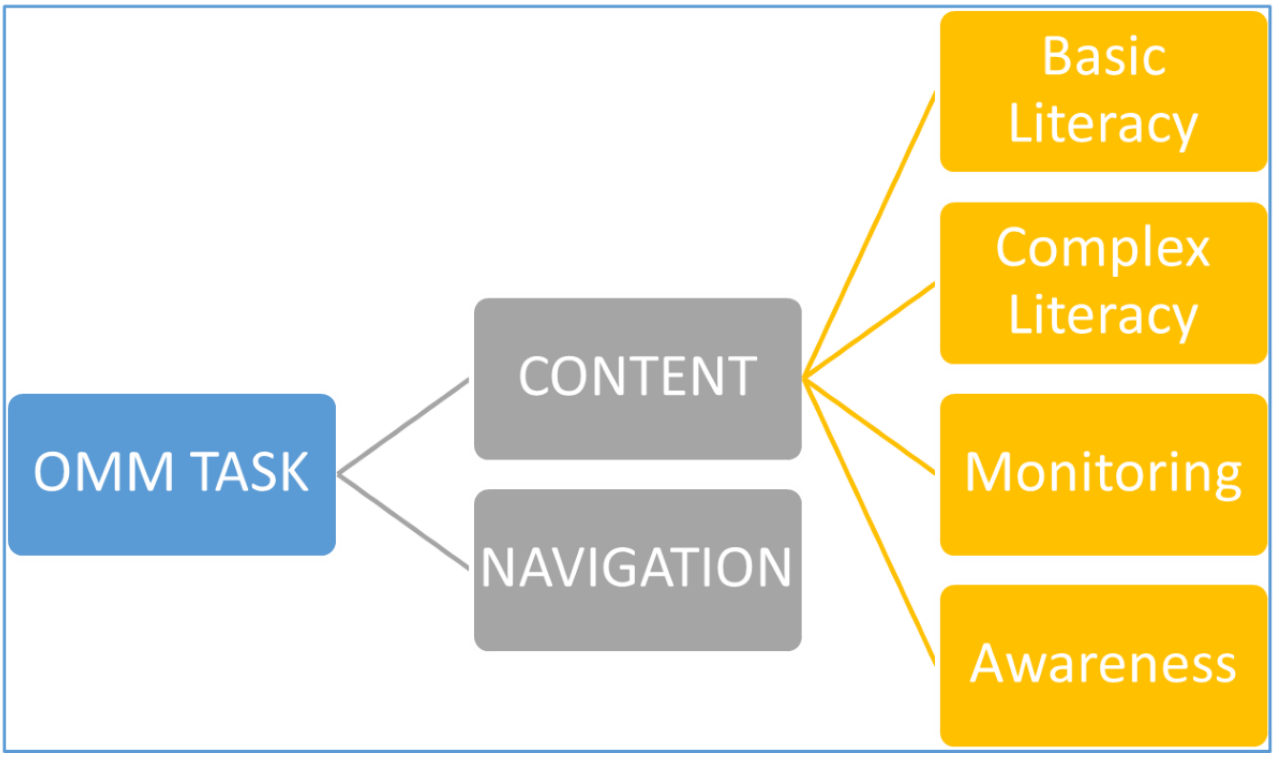

XSL•FO
RenderX

Figure 3. Screenshot of the simulated accounts activity web page.

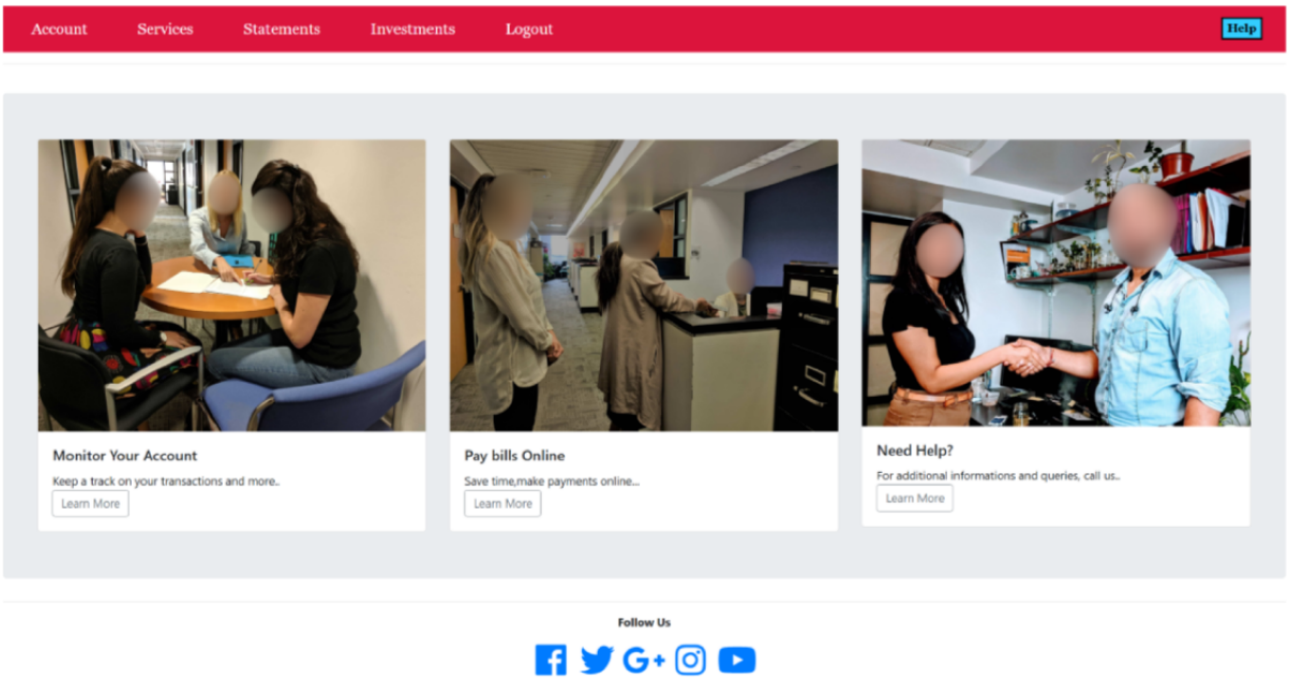

Figure 4. Screenshot of the view statements simulated web page.

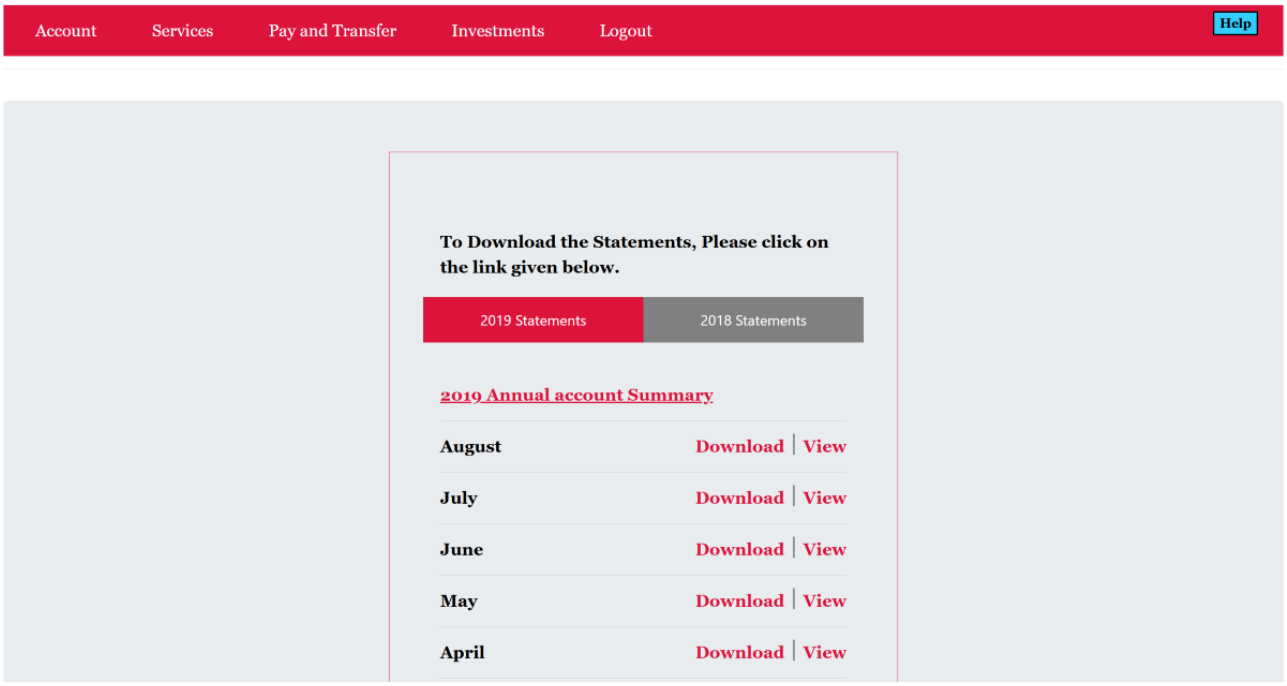

For participants enrolled via the ADRCs, cognitive data collected from those centers were going to be used to examine part of aim 1, which seeks to examine the validity of the task. In the ADRC, participants complete a standard battery of cognitive tests from the National Alzheimer's Coordinating Center Uniform Data Set Battery. Additional encompassing sociocognitive-emotional domains are outlined in Table 1. Moreover, a set of financial tasks was included in this study. The FCI-SF was adapted for online administration during the K99 phase, and 2 additional novel tasks and 2 online tasks, including the online ATM (automated teller machine) and banking tasks from the iFunction battery [41-43], were included in the R00 stage. Questions about experience with financial management and use of technology were completed via self-report, in which participants rated their confidence across various aspects of financial management, such as their everyday handling of finances, mathematical abilities, and use of an ATM.

XSL•FO
RenderX

Table . Select neuropsychological battery and key FDM[a] measures.

Domain	Test
General cognitive ability	Montreal Cognitive Assessment–Alternate Form
Memory	NIHT[b] Auditory Verbal Learning Test
Language or crystallized intelligence	NIHT Picture Vocabulary
Episodic memory	NIHT Picture Sequence Memory Test
Working memory	NIHT List Sorting Working Memory Test
Language or estimated intelligence	NIHT Oral Reading Recognition Test
Numeracy	Berlin Numeracy Scale
Practical judgment	Test of Practical Judgment
Financial capability[c]	Online ATM[d] and banking task (iFunction)

[a]FDM: financial decision-making.

[b]NIHT: National Institutes of Health ToolBox Measure.

[c]Study-specific tests added to the battery.

[d]ATM: automatic teller machine.

A measure of financial exploitation was also incorporated during the R00 study phase. Financial exploitation is operationalized as financial loss due to identity theft, scam, or fraud that could have happened in person, or via phone, electronic, or online transaction. Participants completed a questionnaire about whether financial loss had ever occurred, the number of times it had occurred, and how it occurred (eg, theft, scam, fraud, repetitive spending, unnecessary purchases, etc). The survey was designed following the methodological recommendations of a recent meta-analysis, which found that multiple close-ended questions about specific fraud-scam events identified victims at a higher rate than studies using a single, general self-report question [44-46]. Preliminary data in people diagnosed with dementia indicates that 3 of 5 caregivers endorsed that financial loss had occurred in patients with instances ranging from 1 to 5 times. Of these 3 instances, 2 patients were endorsed by the caregivers as having “no concerns” about their ability to handle financial matters. For this project, a continuous, composite score ranging from 0 (never occurred) to n (number of times it occurred) will be used to quantify exploitation or loss for aim 4’s analysis.

Recruitment and Screening Strategies

ADRC

The ADRC Data Core first ran a search query in its registry that fit our screening criteria. The ADRC recruitment coordinators or research staff conducted an initial screening. They referred participants to our study team once the ADRC staff made contact with an interested candidate and received consent to share their information with our study team. Additionally, during weekly ADRC consensus meetings, ADRC staff identified participants who fit our eligibility criteria and reached out to them for consent to share their information with us. This study’s staff contacted participants via phone or email to give study information and conduct further screening if the participant was interested. Consent and study procedures were conducted on the same day or at a later time point, depending on the participants' preferences. Individuals who expressed interest in this study were told the purpose of this study, the details of participation, and the potential risks and benefits before scheduling a study visit.

Saint Elizabeth’s Medical Center- Brighton/Boston, MA

Participants with and without cognitive impairment were to be recruited from Saint Elizabeth’s hospital in Brighton, MA. The PI provided Saint Elizabeth’s staff with Boston University’s IRB-approved and stamped flyers for them to give to patients. Interested individuals who contacted our study team were then screened, consented, and enrolled in this study. Participants recruited from Saint Elizabeth’s Medical Center were asked to sign a HIPAA (Health Insurance Portability and Accountability Act) release of medical records form so that the PI could access their medical records.

BMC-Boston, MA

Participants with and without cognitive impairment were to be recruited from BMC. BMC would identify and send out an initial opt-out message to individuals who may be eligible to participate in this study. The opt-out message was signed by the providers known to the patients recruited from BMC. Individuals were identified by their providers (during clinic visits) or by conducting a database search among the patients who received clinical care from neurologists or nurse practitioners using the criteria listed in the HIPAA section of our IRB. This study’s team then conducted screening and consented participants for this study. Participants recruited from BMC were asked to consent to link their data from BMC with the data collected in this study.

Community and Social Media

Healthy controls were recruited via social media platforms or from ResearchMatch with Boston University–stamped and approved posters. Interested individuals were contacted by this study’s team via phone or email. Individuals who expressed interest in this study were told the purpose of this proposed study, the details of participation, and the potential risks and benefits before obtaining their informed consent. We informed

XSL•FO
RenderX

participants at the time of the prescreening phone call that we would do additional prescreening before consenting them to this study and that they may or may not be eligible at that time. The cognitive prescreener was administered to determine if a participant met the eligibility criteria.

Data Collection and Monitoring

Participants consisted of older adults aged 60 to 75 years who are cognitively normal (Clinical Dementia Rating=0) as per the ADRC consensus conference. For participants recruited from the community, a Montreal Cognitive Assessment (MOCA) score of 26 or above was required for inclusion in this study. For those with a lower level of education, an education correction was applied to the total MOCA score. We focus on this age band to ensure that we are maximizing the probability of capturing older adults who use online services to manage their money. Participants complete this study with an examiner via a Zoom (Zoom Communications, Inc) call. The same set and order of tests are administered to all participants.

Statistical Analysis

Overview

Primary covariates will include sex and years of education. For aims 1 and 2, we will conduct sensitivity analysis and then include variables such as socioeconomic status and financial literacy as secondary covariates to explore their influence on FDM. Secondary analysis will involve stratifying based on race and ethnicity and using these as covariates or as silent regressors.

Aim 1

For aim 1, we hypothesized that the OMM task would have high internal consistency, interrater, and test-retest reliability. Cronbach ω will be used to measure internal consistency reliability, Cohen κ for interrater reliability, and the Pearson product-moment correlation coefficient of stability for test-retest. Standard descriptive statistics of frequency, central tendency, and dispersion will be used to describe the sample. The NIH criteria of rigor will be incorporated when reporting the confirmatory factor analysis results, with emphasis on presenting information on the proposed models, modifications (if any) that were made, the specific measures that identify the latent variables, the intercorrelations among measures, the correlations between latent variables, and whether any constraints were used. When reporting the model fit statistics, the chi-square test, the root mean square error of approximation, the comparative fit index, and the standardized root mean square residual will be reported. A subgroup of participants has been retested on the same OMM credit card statement task after 3 months to evaluate test-retest reliability. The 2 raters will score the responses for the Complex Financial Literacy subcomponent following objectively defined criteria.

Similar to a previous study's sample size [29], power analysis was conservatively conducted using α=.001. With n=200 (older adults; sex equally stratified), the minimum detectable effect size of r=0.20 at 80% power. A previous study, examining different FDM domains in a mixed group of older adults and those diagnosed with AD, reported strong correlations when assessing internal consistency (r ranged between 0.85 to 0.93; n=73), test-retest reliability (above 0.85, interval period=23 days, n=17), and high interrater reliability (n=11) [32]. Another study found money management to be associated with Dementia Rating Score (r=.80, P<.05; n=83) [32,47]. Thus, we have enough power to detect strong correlations.

Aim 2

For aim 2, we hypothesized that the rate of decline in statement monitoring and complex literacy would be associated with the rate of decline in executive functioning. Mixed effect models will be conducted with OMM task measures as dependent variables, time (2 levels: baseline and follow-up), domain (4 levels: basic financial literacy, complexity literacy, statement monitoring, and awareness), executive function scores, a 3-way interaction (time × domain × executive function) and their nested 2-way interactions as fixed effects controlling for age, gender, financial literacy, socioeconomic status, race, and education. The within-subject correlation is accounted for as a random intercept, and the correlation between domains will be specified as unstructured. We will first test the 3-way interaction term to test whether the change in executive function is differently associated with changes in OMM. If the 3-way interaction term is significant, we will run a contrast analysis to identify which domain changes show a significant association with changes in executive function. If the 3-way interaction term is not significant, we will rerun the mixed effect model without the 3-way interaction term and test the time × executive function. Power analyses were conducted with the assumption of moderate within-subject correlation (r=0.3) and conservatively estimating 20% attrition. With the sample size of 200, we expect to have at least 160 subjects at the end of this study. If the 2 domains show more than 0.38 SD decline (small effect size), we would be able to detect time × domain interaction with 80% power at the 5% significance level. If the association between change in executive function and change in complex literacy is greater than r=0.22, which is smaller than the medium effect size, we would have more than 80% power to detect such an association at the 5% significance level.

Aims 3 and 4

As the data from the AD sample could not be collected, aim 3 was not conducted, and aim 4 was modified. Originally, we had hypothesized that for aim 4, in those with clinical AD, compared to FDM as measured by OMM metrics (basic and complex literacy, and monitoring) and demographics (age, sex, and education), worse performance on OMM financial awareness will be strongly associated with a greater average number of times being financially exploited. Currently, we intend to examine the prevalence of financial loss in cognitively healthy older adults. Bivariate correlations will be conducted between demographics, FDM metrics (literacy and monitoring), financial awareness, and financial exploitation. Depending on the type of predictors, the three most significant correlations will be entered into a multiple linear regression analysis as predictors and the number of times financial exploitation occurred as the outcome variable. The mean difference in number of times financial exploitation occurred between groups for each predictor will be reported as β, along with 95% CIs. If the data are skewed or the range for the number of financial exploitation

XSL•FO
RenderX

events is restricted, then multivariable logistic regression analysis will be conducted, with odds ratios and 95% CIs reported, to test whether each predictor is associated with financial exploitation as a dichotomous (yes or no) outcome. Separate models for each predictor will be constructed, along with a combined model that includes all 3 predictors included in order to assess whether each predictor is associated with financial exploitation independent of the other predictors. Models will be adjusted for age, sex, and education. Statistical significance will be based on a *P* value less than .05.

Power Analysis

For the primary hypothesis that lower financial awareness is associated with a higher likelihood of being financially exploited, power analysis was conducted under the assumption that the average proportion of having low financial awareness among those who reported being financially exploited is 1.54(p1), based on my preliminary data. With a sample size of 50 participants and assuming a 0.5 SD difference (ie, medium effect size) in p1 compared to p2, which represents the average proportion of having high financial awareness among those who report being financially exploited, we will have 83% power to detect such difference at the 5% significance level. This result implies that at our given sample size, we would have even greater power if we observe a larger difference (ie, greater than 0.5 SD difference). However, if the observed difference is smaller than 0.5 SDs, then we may be underpowered (ie, <80%) to detect such difference under our given sample size and assumptions.

Results

Funding for the K99 and the R00 phases of the project was received on August 15, 2019, and January 1, 2021, respectively. This study was approved by the Columbia University Medical Center IRB on January 17, 2019, and by the Boston University IRB on March 30, 2022. Participant recruitment for the K99 phase began at Columbia University Medical Center in September 2019, while recruitment for the R00 phase began at Boston University in November 2022.

Enrollment concluded in July 2025. A total of 230 participants from the R00 group and 42 participants from the K99 group consented to this study, combining to a total of 272. Further, 147 participants returned to complete a longitudinal study visit. A subcohort of R00 participants (n=53) also completed a 3-month follow-up in which only the OMM task was completed. Preliminary results have been presented at various national and international conferences, including at the Gerontological Society of America's meeting in 2020, and the International Neuropsychological Society's meetings from 2023 to 2025 [48-50] and at the Gerontological Society of America's meeting in 2020 [51]. Data analysis is underway as of August 2025, and results are expected to be published in 2026.

Discussion

Principal Findings and Implications

This proposed study was designed to use rigorous standards for developing the OMM credit card task. Study procedures were adapted during the K99 phase of this study. During the R00 phase, in sync with the K99 procedures, remote testing procedures were continued to maximize data and maintain standardization.

This proposed project has substantial implications for researchers, clinicians, and older adults. Digital currency is rapidly becoming the most preferred form of financial transaction. In fact, as per several studies such as the Pew Research Center's data, the majority of older adults are internet savvy, with 78% of them being online and 77% owning a smartphone [52-54]. Development of the OMM task is highly relevant for the baby boomer cohort as well as the following generations and will reflect FDM abilities in a more appropriate and task-congruent manner. By objectively collecting FDM data in concert with comprehensive information about cognition, the proposed project will help researchers to expand their conceptualization of FDM and provide clinicians with a psychometrically sound, easy-to-administer, up-to-date assessment tool, ultimately helping older adults and their caregivers make informed determinations about managing their finances. Moreover, the simulation technology-based OMM task will provide new information that can be used to inform public health policies and perhaps banking and judicial policies, thereby preventing financial losses. Indeed, in a survey study of financial professionals, 74% compliance officers and 61% financial advisors reported encountering individuals with diminished capacity and feeling unprepared to handle the situation [55]. The availability of OMM will allow banks and judicial organizations to confidently develop roadmaps for handling cases related to questionable FDM and thereby create a balance between autonomy and safety.

Potential Problems and Alternative Strategies

Challenges associated with confounding factors, low reliability, low validity, and attrition will be explored. We will attempt to report the rate of attrition along with the possible reasons (death, refusal to return, lost to follow-up, etc). A mixed-model statistical approach will be adopted to mitigate the effects of attrition. If we find specific patterns in attrition, the Inverse Probability of Treatment Weighting approach will be used to adjust for differential attrition rates [56]. There is a possibility that no change will be detected in older adults at longitudinal follow-up. If changes are detected, then the OMM task can serve as a sensitive test for detecting functional impairments in preclinical MCI. If, however, no changes are detected, then the data will serve as a useful reference for comparing other clinical groups and for deriving normative data.

Dissemination Plan

Once the task is validated, a training manual with administration instructions, scoring, and interpretation procedures, along with a demonstration video of the OMM task, will be developed to enable standardization of the OMM application. Online training will be provided to researchers and clinicians on an annual basis as per institutional guidelines. We anticipate peer-reviewed publication of our findings after 2025. Findings will be shared at national and international conferences via oral talks and poster presentations. Deidentified data will be provided upon request

XSL•FO
RenderX

by the researcher and as per journal and institutional policies to encourage reproducibility across diverse settings.

Future Directions

The findings from this study will be used to build several closely related projects. First, for the next iteration of the OMM task, the plan is to automate the task such that participants can complete the task remotely without the involvement of research staff. The results will be automatically generated, and if the researcher or clinician needs to view the task performance, they will be able to review the performance via automatically recorded videos. Next, the OMM task needs to be validated in clinical and demographically diverse populations. Similarly, the links between this task and individuals with varying types and levels of biomarker statuses need to be examined. Third, the neural basis of the OMM task should be evaluated as this might have implications for diagnosis, monitoring, and rehabilitation.

Acknowledgments

No generative artificial intelligence tools were used in the creation of this paper or the presented protocol.

Funding

This project is funded by the National Institute of Health (K99/R00AG062783). The funder had no involvement in this study's design, data collection, analysis, interpretation, or the writing of this paper.

Data Availability

Data generated from this study will be available upon reasonable request. The principal investigator will approve of such requests and will provide the data at no cost to individual investigators. This study's team used a REDCap (Research Electronic Data Capture; Vanderbilt University) system and SPSS (IBM Corp) software to share the data. Research results will be disseminated publicly through publication in scientific journals and through oral presentations or posters presented at scientific meetings. Such data publication will occur during the term of this project or after its conclusion, as appropriate and in keeping with normal scientific practices for reporting results.

Authors' Contributions

Conceptualization: PS, SC
Data curation: PS, SC, MB-L
Formal analysis: PS, ZK
Funding acquisition: PS
Investigation: PS, MB-L
Methodology: PS, SC, MB-L, EDH
Project administration: PS, MB-L
Resources: PS, SC
Software: PS, ZK
Supervision: PS, SC
Validation: PS, MB-L
Visualization: PS, MB-L
Writing – original draft: PS, SC
Writing – review & editing: all authors

Conflicts of Interest

None declared.

Peer Review Report 1
Peer review report by Pathway to Independence Award Committee (National Institutes of Health, United States).
[PDF File, 101 KB - resprot_v15i1e82488_app1.pdf]

References

1. Miller DB, Brown KM, Adams KB. Economic fragility of older adults: factors contributing to a continuing crisis. J Aging Soc Policy 2025;37(6):1147-1163. [doi: 10.1080/08959420.2024.2384190] [Medline: 39190857]
2. Zhe L. Household debt among older Americans, 1989-2016. : Library of Congress; 2019 URL: https://www.congress.gov/crs-product/R45911 [accessed 2026-03-28]
3. Popham L, Silberman S, Berke L, Tavares J, Cohen M. The 80% the true scope of financial insecurity in retirement. : National Council on Aging and LeadingAge LTSS Center at UMass Boston; 2020 URL: https://assets.ncoa.org/

XSL•FO
RenderX

ffacfe7d-10b6-0083-2632-604077fd4eca/302be397-f6e1-482b-ac71-ff2957dbe5bb/2021-FWP-DG02_The%2080_Issue%20Brief_FINAL.pdf [accessed 2026-03-28]
4. Agarwal S, Driscoll JC, Gabaix X, Laibson D. The age of reason: financial decisions over the life cycle and implications for regulation. ECA 2009 Sep;2009(2):51-117. [doi: 10.1353/eca.0.0067]
5. Tymula A, Rosenberg Belmaker LA, Ruderman L, Glimcher PW, Levy I. Like cognitive function, decision making across the life span shows profound age-related changes. Proc Natl Acad Sci U S A 2013 Oct 15;110(42):17143-17148. [doi: 10.1073/pnas.1309909110] [Medline: 24082105]
6. Korniotis GM, Kumar A. Do older investors make better investment decisions? Rev Econ Stat 2011 Feb;93(1):244-265. [doi: 10.1162/REST_a_00053]
7. Shivapour SK, Nguyen CM, Cole CA, Denburg NL. Effects of age, sex, and neuropsychological performance on financial decision-making. Front Neurosci 2012;6(82):82. [doi: 10.3389/fnins.2012.00082] [Medline: 22715322]
8. Pottow JAE. The rise in elder bankruptcy filings and the failure of U.S. bankruptcy law. Elder Law J 2011;19:119-158 [FREE Full text]
9. Report and recommendations: fighting elder financial exploitation through community networks. : Consumer Financial Protection Bureau: Office of Financial Protection for Older Americans; 2016 URL: https://www.consumerfinance.gov/data-research/research-reports/report-and-recommendations-fighting-elder-financial-exploitation-through-community-networks/ [accessed 2026-04-11]
10. Acierno R, Hernandez MA, Amstadter AB, et al. Prevalence and correlates of emotional, physical, sexual, and financial abuse and potential neglect in the United States: the National Elder Mistreatment Study. Am J Public Health 2010 Feb;100(2):292-297. [doi: 10.2105/AJPH.2009.163089] [Medline: 20019303]
11. Conrad KJ, Iris M, Ridings JW, Langley K, Wilber KH. Self-report measure of financial exploitation of older adults. Gerontologist 2010 Dec;50(6):758-773. [doi: 10.1093/geront/gnq054] [Medline: 20667945]
12. Spreng RNP, Karlawish JM, Marson DCM. Cognitive, social, and neural determinants of diminished decision-making and financial exploitation risk in aging and dementia: a review and new model. J Elder Abuse Negl 2016;28(4-5):320-344. [doi: 10.1080/08946566.2016.1237918] [Medline: 27644698]
13. Lichtenberg PA, Stickney L, Paulson D. Is psychological vulnerability related to the experience of fraud in older adults? Clin Gerontol 2013 Mar 1;36(2):132-146. [doi: 10.1080/07317115.2012.749323] [Medline: 23997404]
14. Lichtenberg PA, Campbell R, Hall L, Gross EZ. Context matters: financial, psychological, and relationship insecurity around personal finance is associated with financial exploitation. Gerontologist 2020 Aug 14;60(6):1040-1049. [doi: 10.1093/geront/gnaa020] [Medline: 32211847]
15. Shao J, Zhang Q, Ren Y, Li X, Lin T. Why are older adults victims of fraud? Current knowledge and prospects regarding older adults' vulnerability to fraud. J Elder Abuse Negl 2019 Mar;31(3):225-243. [doi: 10.1080/08946566.2019.1625842] [Medline: 31159679]
16. Lusardi A, Mitchell OS. Baby Boomer retirement security: the roles of planning, financial literacy, and housing wealth. J Monet Econ 2007 Jan;54(1):205-224. [doi: 10.1016/j.jmoneco.2006.12.001]
17. Han SD, Boyle PA, James BD, Yu L, Bennett DA. Mild cognitive impairment is associated with poorer decision-making in community-based older persons. J Am Geriatr Soc 2015 Apr;63(4):676-683. [doi: 10.1111/jgs.13346] [Medline: 25850350]
18. Boyle PA, Yu L, Wilson RS, Gamble K, Buchman AS, Bennett DA. Poor decision making is a consequence of cognitive decline among older persons without Alzheimer's disease or mild cognitive impairment. PLoS ONE 2012;7(8):e43647. [doi: 10.1371/journal.pone.0043647] [Medline: 22916287]
19. Sunderaraman P, Ho S, Chapman S, et al. Technology use in everyday financial activities: evidence from online and offline survey data. Arch Clin Neuropsychol 2020 May 22;35(4):385-400. [doi: 10.1093/arclin/acz042] [Medline: 31696205]
20. Pollard K, Scommegna P. Just how many baby boomers are there? Population Reference Bureau. 2014 Apr 16. URL: https://www.prb.org/resources/just-how-many-baby-boomers-are-there [accessed 2026-03-28]
21. Lee CC, Czaja SJ, Moxley JH, et al. Attitudes toward computers across adulthood from 1994 to 2013. Gerontologist 2019 Jan 9;59(1):22-33. [doi: 10.1093/geront/gny081] [Medline: 29982458]
22. Rolfe A. Bank of America trends in consumer mobility report. : Payments Industry Intelligence; 2017 URL: https://www.paymentscardsandmobile.com/bank-america-consumer-mobility-report [accessed 2026-03-28]
23. Buchman AS, Boyle PA, Bennett DA. Expanding the toolkit for studies of aging. J Prev Alzheimers Dis 2017;4(2):69-70. [doi: 10.14283/jpad.2017.14] [Medline: 29177136]
24. Appelbaum PS, Spicer CS, Valliere FR. Informing Social Security's Process for Financial Capability Determination: National Academies Press; 2016. [doi: 10.17226/21922]
25. Lichtenberg PA, Stoltman J, Ficker LJ, Iris M, Mast B. A person-centered approach to financial capacity assessment: preliminary development of a new rating scale. Clin Gerontol 2015 Jan 1;38(1):49-67. [doi: 10.1080/07317115.2014.970318] [Medline: 25866438]
26. Sunderaraman P, Cosentino S, Lindgren K, James A, Schultheis M. Informant report of financial capacity for individuals with chronic acquired brain injury: an assessment of informant accuracy. J Head Trauma Rehabil 2018;33(6):E85-E94. [doi: 10.1097/HTR.0000000000000382] [Medline: 29601341]

XSL•FO
RenderX

27. Engel L, Chui A, Beaton DE, Green RE, Dawson DR. Systematic review of measurement property evidence for 8 financial management instruments in populations with acquired cognitive impairment. Arch Phys Med Rehabil 2018 Sep;99(9):1848-1875. [doi: 10.1016/j.apmr.2018.02.004] [Medline: 29524397]
28. Lichtenberg PA PhD Abpp, Ficker L PhD, Rahman-Filipiak A Ma, et al. The Lichtenberg Financial Decision Screening Scale (LFDSS): a new tool for assessing financial decision making and preventing financial exploitation. J Elder Abuse Negl 2016;28(3):134-151. [doi: 10.1080/08946566.2016.1168333] [Medline: 27010780]
29. Lichtenberg PA, Ocepek-Welikson K, Ficker LJ, Gross E, Rahman-Filipiak A, Teresi JA. Conceptual and empirical approaches to financial decision-making by older adults: results from a financial decision-making rating scale. Clin Gerontol 2018;41(1):42-65. [doi: 10.1080/07317115.2017.1367748] [Medline: 29077531]
30. Teresi JA, Ocepek-Welikson K, Lichtenberg PA. Item response theory analysis of the Lichtenberg Financial Decision Screening Scale. J Elder Abuse Negl 2017;29(4):213-228. [doi: 10.1080/08946566.2017.1338170] [Medline: 28590882]
31. Marson D. Conceptual models and guidelines for clinical assessment of financial capacity. Arch Clin Neuropsychol 2016 Sep;31(6):541-553. [doi: 10.1093/arclin/acw052] [Medline: 27506235]
32. Marson DC, Sawrie SM, Snyder S, et al. Assessing financial capacity in patients with Alzheimer disease: a conceptual model and prototype instrument. Arch Neurol 2000 Jun;57(6):877-884. [doi: 10.1001/archneur.57.6.877] [Medline: 10867786]
33. Griffith HR, Belue K, Sicola A, et al. Impaired financial abilities in mild cognitive impairment: a direct assessment approach. Neurology (ECronicon) 2003 Feb 11;60(3):449-457. [doi: 10.1212/wnl.60.3.449] [Medline: 12578926]
34. Dubois B, Feldman HH, Jacova C, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. Lancet Neurol 2007 Aug;6(8):734-746. [doi: 10.1016/S1474-4422(07)70178-3] [Medline: 17616482]
35. Marson DC, Triebel K, Knight A. Civil Capacities in Clinical Neuropsychology: Research Findings and Practical Applications: Oxford University Press; 2012:39-68.
36. Marson D, Hethcox A, Shawaryn M, Novack T. Assessing financial capacity in patients with traumatic brain injury. Arch Clin Neuropsychol 2000 Nov 1;15(8):828-829. [doi: 10.1093/arclin/15.8.828]
37. Gerstenecker A, Triebel K, Eakin A, Martin R, Marson D. Exploring the factor structure of financial capacity in cognitively normal and impaired older adults. Clin Gerontol 2018;41(1):33-41. [doi: 10.1080/07317115.2017.1387211] [Medline: 29182443]
38. Rahman S, Kordovski VM, Tierney SM, Woods SP. Internet navigation skills for financial management: associations with age, neurocognitive performance, and functional capacity. Neuropsychology 2021 Sep;35(6):630-642. [doi: 10.1037/neu0000759] [Medline: 34292027]
39. Woods SP, Iudicello JE, Morgan EE, et al. Household everyday functioning in the internet age: online shopping and banking skills are affected in HIV-associated neurocognitive disorders. J Int Neuropsychol Soc 2017 Aug;23(7):605-615. [doi: 10.1017/S1355617717000431] [Medline: 28625210]
40. Penheiro R, Kiselica AM, Webber TA, Kamar JLT, Woods SP. Banking on variability: a conceptual replication of the association between cognitive dispersion and financial management. Clin Neuropsychol 2025 Jun 23;23:1-20. [doi: 10.1080/13854046.2025.2522249] [Medline: 40549434]
41. Czaja SJ, Kallestrup P, Harvey PD. Evaluation of a novel technology-based program designed to assess and train everyday skills in older adults. Innov Aging 2020;4(6):igaa052. [doi: 10.1093/geroni/igaa052] [Medline: 33324761]
42. Czaja SJ, Loewenstein DA, Sabbag SA, Curiel RE, Crocco E, Harvey PD. A novel method for direct assessment of everyday competence among older adults. J Alzheimers Dis 2017;57(4):1229-1238. [doi: 10.3233/JAD-161183] [Medline: 28304300]
43. Harvey PD, Tibiriçá L, Kallestrup P, Czaja SJ. A computerized functional skills assessment and training program targeting technology based everyday functional skills. J Vis Exp 2020 Feb 13(156):e60330. [doi: 10.3791/60330] [Medline: 32116309]
44. Nguyen AL, Mosqueda L, Windisch N, Weissberger G, Axelrod J, Han SD. Perceived types, causes, and consequences of financial exploitation: narratives from older adults. J Gerontol B Psychol Sci Soc Sci 2021 Apr 23;76(5):996-1004. [doi: 10.1093/geronb/gbab010] [Medline: 33423064]
45. Jackson SL. A systematic review of financial exploitation measures in prevalence studies. J Appl Gerontol 2018 Sep;37(9):1150-1188. [doi: 10.1177/0733464816650801] [Medline: 27231244]
46. Greene AJ. Elder financial abuse and electronic financial instruments: present and future considerations for financial capacity assessments. Am J Geriatr Psychiatry 2022 Jan;30(1):90-106. [doi: 10.1016/j.jagp.2021.02.045] [Medline: 33781661]
47. Marson DC. Loss of financial competency in dementia: conceptual and empirical approaches. Aging, Neuropsychol, Cognit 2001 Sep;8(3):164-181. [doi: 10.1076/anec.8.3.164.827]
48. Sunderaraman P, Chapman S, Hartstone W, et al. Ushering in modern and objective ways of assessing financial decision making in clinical settings: the development and validation of an online money management credit card task. J Int Neuropsychol Soc 2023 Dec;29(s1):208-209. [doi: 10.1017/S1355617723003144]
49. Bouchard-Liporto M, Goyal K, Chapman S, Cosentino S, Sundararaman P. Cognitive and numeracy correlates of financial decision-making tasks in older adults [Abstract]. Presented at: 52nd Annual North American Meeting of the International Neuropsychological Society; Feb 14-17, 2024.

XSL•FO
RenderX

50. Bouchard-Liporto M, Goyal K, Thomas H, et al. Congruencies and discrepancies in objective performance and self-reported perceptions for two simulated online financial decision-making tasks [Abstract]. Presented at: 53rd Annual North American Meeting of the International Neuropsychological Society; Feb 12-15, 2025.
51. Sunderaraman P, Bouchard-Liporto M, Chapman S, Rizer S, Yajima M, Cosentino S. Cognitive correlates of a novel simulated online money management credit card task. Innovation Aging 2025 Dec 1;9(Supplement_2):igaf122.106. [doi: 10.1093/geroni/igaf122.106]
52. Damodaran L, Olphert CW, Sandhu J. Falling off the bandwagon? Exploring the challenges to sustained digital engagement by older people. Gerontology 2014;60(2):163-173. [doi: 10.1159/000357431] [Medline: 24356488]
53. Kuerbis A, Mulliken A, Muench F, A. Moore A, Gardner D. Older adults and mobile technology: factors that enhance and inhibit utilization in the context of behavioral health. Ment Health Addict Res 2017;2(2):1-11. [doi: 10.15761/MHAR.1000136]
54. Anderson M, Perrin A. Technology use among seniors. Pew Research Center. 2017 May 17. URL: https://www.pewresearch.org/internet/2017/05/17/technology-use-among-seniors/ [accessed 2026-03-28]
55. Karp N. Protecting older investors: the challenge of diminished capability. Gener: J Am Soc Aging 2012;36(2):33-38 [FREE Full text]
56. Rosenbaum PR. Model-based direct adjustment. J Am Stat Assoc 1987 Jun;82(398):387-394. [doi: 10.1080/01621459.1987.10478441]

Abbreviations

AD: Alzheimer disease
ADRC: Alzheimer's Disease Research Center
ATM : automated teller machine
BMC: Boston Medical Center
COSMIN: Consensus-Based Standards for the Selection of Health Measurement Instruments
FCI-SF: Financial Capacity Instrument – Short Form
FDM: financial decision-making
HIPAA : Health Insurance Portability and Accountability Act
IRB: institutional review board
MOCA: Montreal Cognitive Assessment
NIH: National Institutes of Health
OMM: Online Money Management
PI: principal investigator

Edited by A Schwartz; submitted 18.Aug.2025; this is a non–peer-reviewed article;accepted 05.Dec.2025; published 20.Apr.2026.

Please cite as:
Sunderaraman P, Bouchard-Liporto M, Kunicki Z, Huey ED, Cosentino S
Objective Assessment of Financial Decision-Making With a Simulated Online Money Management Task in Older Adults: Protocol for a Prospective Observational Study
JMIR Res Protoc 2026;15:e82488
URL: https://www.researchprotocols.org/2026/1/e82488
doi:10.2196/82488

© Preeti Sunderaraman, Madison Bouchard-Liporto, Zachary Kunicki, Edward D. Huey, Stephanie Cosentino. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Strengthening Actions for Menstrual Health and Hygiene Interventions for Promotion of Women's Health in Nepal (SAMIP): Protocol for a Participatory Intervention Development Study Using Realist Synthesis, Human-Centered Design, Intervention Mapping, and Arts-Based Methods

Sara Elizabeth Baumann[1], PhD, MPH; Hannah Schwarz[1], MPS; Annika Agarwal[2], BA; Bhimsen Devkota[1,3,4], PhD; Elizabeth Miller[2], PhD, MD; Ada Youk[5], PhD; Sara Parker[6], PhD, MSc; Sanjana Murthy[7], BS, BA; Nikita Sharma[4], MPH; Archana Shrestha[8], PhD, MPH; Madhusudan Subedi[9], MPhil; Nirajan Khadka[10], MA, MHSc; Mary Hawk[1], DrPH, LSW

[1]Department of Behavioral and Community Health Sciences, School of Public Health, University of Pittsburgh, 130 Desoto St, Pittsburgh, PA, United States

[10]Aama Surakshya Nepal, Kathmandu, Nepal

[2]School of Medicine, University of Pittsburgh, Pittsburgh, PA, United States

[3]Department of Health Education, Tribhuvan University, Kirtipur, Bagmati Province, Nepal

[4]Bikash Shrot Kendra, Kathmandu, Bagmati Province, Nepal

[5]Department of Biostatistics, School of Public Health, University of Pittsburgh, Pittsburgh, PA, United States

[6]Department of Humanities and Social Science, Liverpool John Moores University, Liverpool, England, United Kingdom

[7]Department of Anthropology, Dietrich School of Arts and Sciences, University of Pittsburgh, Pittsburgh, PA, United States

[8]Department of Public Health, School of Medical Sciences, Kathmandu University, Dhulikel, Bagmati Province, Nepal

[9]Department of Community Health Sciences, Patan Academy of Health Sciences, Patan, Bagmati Province, Nepal

Corresponding Author:
Sara Elizabeth Baumann, PhD, MPH
Department of Behavioral and Community Health Sciences, School of Public Health, University of Pittsburgh, 130 Desoto St, Pittsburgh, PA, United States

Abstract

Background: Menstrual restrictions remain widespread in Nepal, where approximately 90% of women and girls follow at least one restriction. One of the most harmful practices, chhaupadi (severe menstrual seclusion), requires women and girls to isolate in huts during menstruation and is associated with substantial physical and psychological risks, including injury, assault, smoke inhalation, and social exclusion. Although interventions have attempted to address chhaupadi, effectiveness has varied, and no comprehensive assessment has identified which components work, for whom, and under what conditions. Few prior efforts have meaningfully engaged communities in intervention design to ensure contextual relevance and sustainability.

Objective: This protocol describes a three-aim study designed to (1) elucidate mechanisms underlying the success or failure of chhaupadi interventions; (2) co-design a culturally grounded, theory- and evidence-informed intervention in partnership with communities; and (3) pilot and evaluate the intervention using a controlled trial design.

Methods: This multiphase study integrates realist synthesis, intervention mapping, human-centered design, and arts-based research. In aim 1, we will conduct a realist synthesis of published and gray literature, complemented by knowledge-sharing workshops and expert interviews to identify underlying mechanisms explaining chhaupadi intervention success or failure. In aim 2, synthesis findings will inform participatory intervention development. Using the intervention mapping framework, we will convene a co-design workshop incorporating human-centered design and arts-based research methods (eg, asset mapping, impact ladder, concept posters, and communal storytelling) with women, adolescents, and community collaborators to develop intervention components and implementation strategies. In aim 3, the co-designed intervention will be piloted in a controlled trial comparing intervention communities (full community-designed intervention) with control communities (menstrual health education only). Quantitative measures will assess knowledge, attitudes, and practices related to chhaupadi, self-efficacy, mental health, and reproductive health outcomes. Implementation indicators (eg, enrollment, acceptability, appropriateness, and feasibility) will also be assessed. Qualitative interviews and discussions will examine implementation processes, barriers, facilitators, and contextual influences.

XSL•FO
RenderX

Results: The study was funded in September 2023. Aim 1 activities began in April 2024. To date, 69 materials met the inclusion criteria, and a workshop with 8 organizations and 44 interviews has been conducted. Analysis was completed in February 2026, with results expected in the summer of 2026. Aim 2 co-design workshops began in March 2025; eleven women participated in the workshop, and an additional 123 community members have been consulted. This has resulted in a draft intervention called "Our Voice: Transforming Lives through Ending Chhaupadi." Aim 2 results are expected in Fall 2026. Aim 3 data collection will begin in Winter 2027.

Conclusions: By systematically identifying effective intervention mechanisms and embedding community leadership throughout design and implementation, this study aims to develop a sustainable strategy to reduce harms associated with chhaupadi and improve women's and girls health. This integrated framework may also inform interventions addressing other socially embedded health practices in low-resource settings.

International Registered Report Identifier (IRRID): DERR1-10.2196/89117

(JMIR Res Protoc 2026;15:e89117) doi:10.2196/89117

KEYWORDS

Nepal; menstrual health; human centered design; intervention mapping; community engaged research; arts-based research; realist synthesis; menstruation; intervention design

Introduction

Menstruation Studies and the Nepal Context

Despite being a natural and physiological process [1,2], menstruation is often associated with health and safety challenges, stigma, and human rights concerns, especially in low-resource settings [3]. Guided by the field of critical menstrual studies, addressing health challenges associated with menstruation requires critical engagement with systems of power, socially constructed realities, and lived life course experiences, extending past anatomical explanations to understand how menstruation is interpreted and experienced [4].

In many cultures, menstruation is shrouded in stigma and taboos, leading women to adopt harmful practices [5-12]. In Nepal specifically, menstrual restrictions are widespread; 90% of women and girls follow at least one menstrual restriction [13]. One of these traditions, chhaupadi (severe menstrual seclusion), is a social-religious tradition in which women and girls isolate in menstrual huts or sheds during monthly menstruation (referred to as minor chhau) [14-17]. Many women also isolate to these sheds during the postnatal period (referred to as major chhau or chhau period) [18]. The practice is prevalent in the Karnali and Sudurpashchim provinces of Nepal, where overall development and gender equality are the lowest nationally [16,19]. Chhaupadi was banned by the Supreme Court in 2005 [20], declared a form of gender-based violence in 2010, and criminalized in 2017; however, it prevails in parts of the country [21].

Chhaupadi poses significant health and safety risks for women and families [13,16,17,22]. Physical health risks include snake bites and animal attacks, hypothermia [23], pneumonia [24], suffocation from lighting fires in the sheds to keep warm [14], rape [14], and, in some cases, death [16,17,19,25]. Poor hygiene while practicing chhaupadi puts women at increased risk of infectious disease, as well as infections that can lead to adverse pregnancy outcomes, including preterm delivery [26] and acquisition of sexually transmitted infections [27]. Chhaupadi also impacts the well-being of children who sleep with their mothers or female family members in the sheds; cough, fever, diarrhea, and pneumonia are commonly reported [28]. Women and girls are also psychologically affected during their stays in sheds. Following the practice, which often involves limited security, can leave women in perpetual states of fear [14,29], along with feelings of loneliness [23], stress [30], humiliation [31], isolation [30], and low self-esteem [16]. One study found that 31.4% experienced sadness or depression, and 20% experienced fear of being abused while sleeping in the sheds [31].

Girls who follow chhaupadi have expressed that they follow the tradition due to several deeply held reasons, such as religious and spiritual beliefs, family tradition, fear of negative consequences (eg, potential illness or harm to family if they do not uphold the tradition), and social pressure to maintain the practice [22]. These factors are particularly important to consider when designing interventions.

Chhaupadi sits at the intersection of 2 well-documented bodies of literature: menstrual stigma and harmful traditional practices. Cultural taboos perpetuate menstrual stigma, restrict menstrual hygiene, and hinder efforts to improve knowledge and female well-being; as such, millions of women worldwide face challenges in managing menstruation that negatively affect their health, education, labor force participation, and productivity [32]. Scholars have noted that policy responses alone are often insufficient to dismantle these norms: while public policies can recognize menstrual stigma and set out to break the silence surrounding menstruation, alone, they do not necessarily contribute to dismantling it [33]. Chhaupadi exemplifies this policy-practice gap—criminalized in Nepal since 2017, yet still widely practiced [21]—and thus offers a critical case study in the limits of legislative approaches to gender-based harmful practices.

In the scholarly literature, chhaupadi has been classified as a harmful traditional practice—a term used to describe culturally embedded customs that violate human rights and cause physical, psychological, or social harm, particularly to women and girls [34]. Scholarship on harmful practices in Nepal has documented the wide-ranging health and development consequences of

XSL•FO
RenderX

culturally and socially embedded customs targeting women and girls—among them early marriage, dowry, and menstrual seclusion practices such as chhaupadi [12]. While chhaupadi is specific to Nepal, menstrual restrictions and taboos are documented across South Asia and beyond. Menstruation poses particular challenges to young women in India and other South Asian countries due to high levels of menstrual stigma, where menstruating girls and women are often considered dirty or impure and face taboos and restrictions on their mobility, daily activities, and hygiene practices [35]. Chhaupadi thus represents a particularly severe manifestation of a regionally and globally documented phenomenon.

Efforts to Address Chhaupadi Harms

A range of strategies have been used to address chhaupadi, including awareness raising, education, the physical removal of menstrual sheds, legal sanctions, and the conditional withdrawal of state services [36-38]. Demolition of menstrual sheds has been one structural intervention used by government and NGO actors to eliminate chhaupadi, yet its effectiveness has been limited. In some communities, the removal of sheds has been linked to increased danger for women and girls, who face pressure to practice seclusion regardless of hut destruction, and may resort to sleeping in caves or open spaces without shelter. Local officials have acknowledged that women have rebuilt sheds out of fear, suggesting that physical demolition without attitude change does not constitute a meaningful victory [23,39].

Other intervention approaches that have seen some success are working in partnership with dhamis (traditional healers) and health workers. Interventions that engage traditional healers as agents of change have shown some promise in addressing chhaupadi, reflecting their significant community authority [40]. Other recent studies also suggest that harnessing the expertise of female community health volunteers [41,42] and community health nurses [43] can offer a way to shift attitudes around chhaupadi and associated harms.

Despite these efforts, among others, progress has been modest and often short-lived [38]. The literature points to deeply held beliefs linking the abandonment of menstrual restrictions to illness and misfortune as a significant impediment to change [38]. Effective responses must also account for variation across caste, ethnicity, and religion, as menstrual practices are embedded in complex socio-cultural and religious frameworks that resist one-size-fits-all solutions [13,22,32,41].

Approach

No known studies to date have comprehensively assessed which components of chhaupadi interventions hold promise. Additionally, no known prior interventions have fully engaged communities in the intervention design process to ensure suitability, sustainability, and centering of community values and voices. Pinpointing salient mechanisms and partnering with communities in crafting solutions is needed for sustainable change, and thus is a central objective of this 5-year study.

We intend to first rigorously review intervention evidence using realist synthesis (RS), a method for reviewing and synthesizing evidence that allows researchers to move beyond whether an intervention is effective, and rather understand the complexities of what works, for whom, and under which circumstances [44,45]. An RS is well-suited to study chhaupadi interventions because it allows for an in-depth examination of how, why, and under what contextual conditions these interventions succeed or fail, capturing the complex social, cultural, and structural factors that shape their effectiveness.

These findings will be used to inform the co-design of a chhaupadi intervention with community members. To facilitate this process, we will harness Intervention Mapping (IM), a systematic and evidence-based framework for planning health promotion interventions [46]. By engaging community members throughout intervention design and gathering feedback via an iterative process, we expect to codevelop a culturally appropriate intervention to pilot test in a clinical trial, ultimately aiming to improve women's health in Nepal.

This paper describes a protocol for conducting participatory and creative menstrual health intervention research in Nepal that aims to fill gaps in understanding how to sustainably address harms associated with the behavioral practice of chhaupadi (severe menstrual seclusion). This study is innovative and timely, as it is the first study to develop a comprehensive assessment of existing chhaupadi interventions and the first to use a co-designed intervention approach to address the practice. This integrative 3-pillar approach—drawing on IM, human-centered design (HCD), and arts-based research (ABR)—seeks to generate not only a cocreated, theory- and evidence-based intervention, but also an innovative methodological framework adaptable to addressing multifaceted health issues in other low-resource contexts.

Specific Aims

Three specific aims guide our inquiry:

- Aim 1: Elucidate the mechanisms of success or failure of chhaupadi interventions to date. Via review of scientific and gray literature and key-informant interviews (n=35), we will conduct an RS [45] of extant studies/interventions to identify opportunities and develop a theoretical model to inform chhaupadi intervention design.
- Aim 2: Co-design a chhaupadi intervention with community members. Based on aim 1 findings, we will conduct a 2-week co-design workshop with women and girls who practice chhaupadi (n=10), guided by principles of IM, and using HCD and ABR methods to ultimately develop a culturally relevant chhaupadi intervention. We will elicit feedback on the intervention via community dialogue sessions (n=2, 10 people each), community-level key informant interviews (KIIs) (n=8), and menstrual health expert interviews at the national level (n=8). Additional insights will be gathered via rapid situation analysis in 4 additional chhaupadi practicing communities to understand how the intervention may need to be tailored to specific contexts in future scale-up work (n=75 interviews).
- Aim 3: Pilot the chhaupadi intervention with community members to assess feasibility and acceptability. Two matched communities will be identified and randomly assigned to intervention and control (menstrual education only) at the community level (n=25 intervention, n=25

XSL•FO
RenderX

control). Surveys and focus groups with participants will be conducted to assess feasibility, acceptability, and changes in health attitudes and behaviors.

Methods

Study Design

In this 3-phase study called Strengthening Actions for Menstrual Health and Hygiene Interventions for Promotion of Women's Health in Nepal (SAMIP), we will apply a suite of innovative and tested research methods to address our study objectives (Figure 1). We will conduct an RS (aim 1) to elucidate mechanisms of success or failure of chhaupadi interventions to date, which will inform community-led intervention development using principles of IM, as well as HCD and ABR tools (aim 2). Next, we will pilot the community-designed chhaupadi intervention in a clinical trial, using a control group (aim 3). Guided by community-engaged approaches at all stages of the study, we will collect qualitative, quantitative, and visual data (eg, photographs, community maps, drawings) across multiple districts in Nepal where chhaupadi is practiced.

Figure 1. Schematic overview of the SAMIP (Strengthening Actions for Menstrual Health and Hygiene Interventions for Promotion of Women's Health in Nepal) study workflow. The study is conducted in 3 linked phases. Aim 1: realist synthesis (RS) to identify mechanisms of success and failure of prior chhaupadi interventions. Aim 2: co-design of a community-informed, theory- and evidence-based intervention using intervention mapping (IM), human-centered design (HCD), and arts-based research (ABR) methods. Aim 3: pilot test and evaluation of the intervention in a controlled trial. Iterative feedback loops reflect adaptive refinement of program theories and intervention components throughout the study.

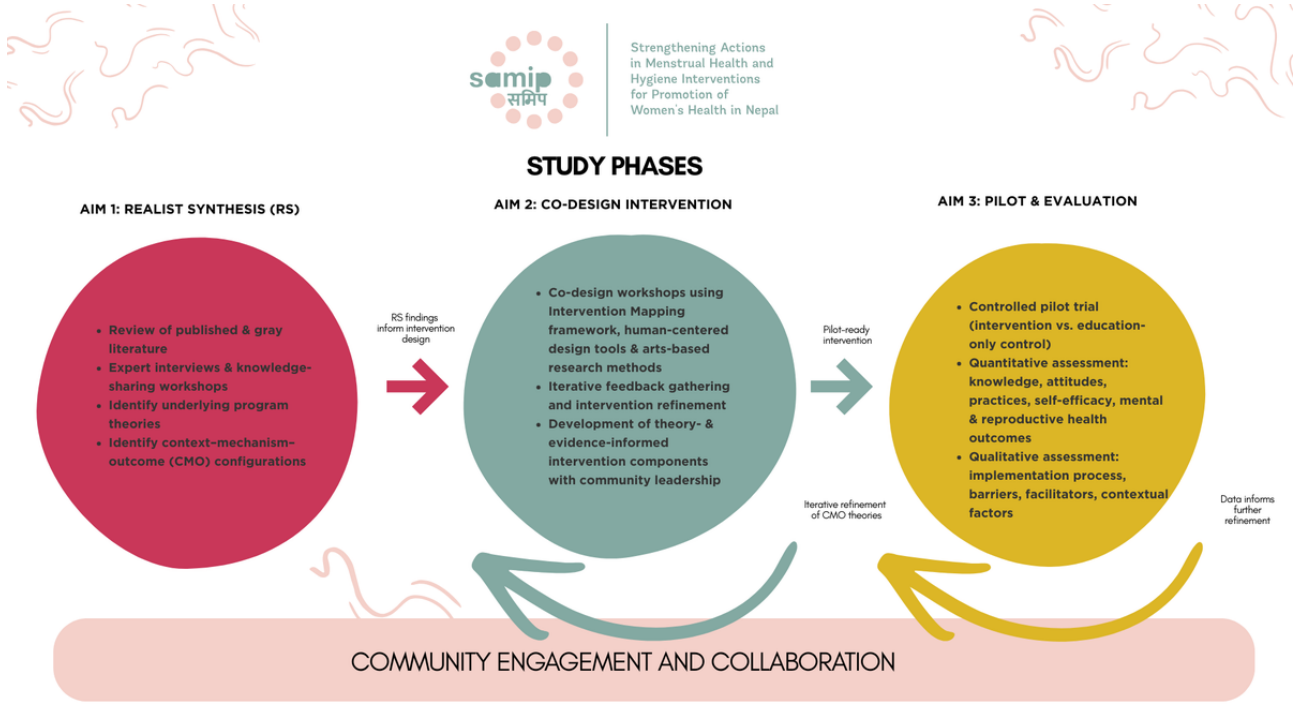

Overview of Study Phases and Data Collection Sites

The study will be conducted in several sites across Nepal (Table 1). In aim 1, the RS will be initiated in the capital Kathmandu, where the research team will consult with organizations that have implemented chhaupadi interventions and will collect reports, evaluations, and manuscripts. We will hold a Knowledge Sharing Workshop to which all organizations that have or are currently implementing chhaupadi interventions will be invited to share lessons learned. These organizations will be identified via the Menstrual Health and Hygiene Partner's Alliance (MHMPA), which tracks the actions of ongoing menstrual health efforts in Nepal [47]. Based on initial findings from an analysis of documentation identified via gray literature and organizations, as well as a review of the literature with the support of a health sciences librarian, follow-up KIIs will be conducted to triangulate findings in several locations across the country. KIIs will be conducted with (1) intervention experts in Kathmandu, (2) staff at the district level where chhaupadi interventions have been implemented, and (3) staff and community members impacted by the interventions in communities where chhaupadi is prevalent. Districts for field-level data collection in aim 1 will be considered if they have a chhaupadi intervention that had/has promising aspects, based on the RS initial analysis.

XSL•FO
RenderX

Table . Overview of study activities and site types by study aim.

Study activities	Study site types/locations	Inclusion/exclusion criteria
Aim 1: realist synthesis (RS)		
• Knowledge sharing workshop (national level) • KIIs[a] with intervention experts (national level) • KIIs with implementation staff (district level) • KIIs/FGDs[b] with communities impacted by promising interventions	• Kathmandu • Districts/communities with promising chhaupadi interventions	Inclusion criteria: • Staff of organizations that have implemented chhaupadi interventions • Individuals who have received a chhaupadi intervention that had promising aspects, based on RS initial analysis
Aim 2: co-design chhaupadi intervention		
• Intervention co-design workshop using IM[c], HCD[d], and ABR[e] methods • KIIs - Community dialogue sessions	• One chhaupadi practicing district with high prevalence • 4 additional medium/high prevalence districts	Inclusion criteria: • Chhaupadi-practicing districts with medium or high prevalence of the practice • Accessible by jeep within 2 hours from district headquarters Exclusion criteria: • Communities with current chhaupadi intervention or within the past 5 years
Aim 3: pilot test		
• Clinical trial of community-designed intervention	• 1 chhaupadi-practicing district • 2 matched communities, random allocation to intervention versus control	Inclusion criteria: • Chhaupadi-practicing district (medium or high prevalence) • Two communities that are sufficiently distant from each other to prevent contamination Exclusion criteria: • Communities from previous aims

[a]KII: key informant interview.

[b]FGD: focus group discussion.

[c]IM: intervention mapping.

[d]HCD: human-centered design.

[e]ABR: arts-based research.

In aim 2, we will use IM, HCD, and ABR methods to generate unique, contextual, viable solutions with communities to address chhaupadi, based on existing evidence as consolidated through the RS (aim 1) [46,48-53]. In aim 2, a co-design workshop will be held in one “action community” that will be selected based on the latest evidence regarding chhaupadi prevalence. District selection will be informed by the following inclusion criteria: (1) has a high prevalence of chhaupadi, (2) does not currently have an ongoing chhaupadi intervention, and (3) is accessible by jeep (within 2 hours from district headquarters). The exclusion criterion is as follows: (1) communities that have had a chhaupadi intervention in the past 5 years. The SAMIP team will work closely with 10 women in the community, referred hereafter as community design team (CDT) members, to create a chhaupadi intervention by using a series of creative and participatory activities and iterative design tools. Activities for the co-design workshop will be guided by the six steps of IM and will include a series of participatory activities harnessing HCD and ABR tools (Table 2).

XSL•FO
RenderX

Table . Detailed data collection activities for human-centered design (HCD) co-design workshop (aim 2).

Activity	Description and timing
Intervention mapping (IM) step 1: develop a logic model of the problem[a]	
Develop the logic model of the problem (LMOP) visual (before fieldwork)	Based on a rapid literature review conducted by the SAMIP[b] team and an RS[c] study to investigate interventions to date, we will develop a logic model of the health problem. We will also bring in theories that guide the logic model development. This will be translated into a visual figure in Nepali for ease of communication with community members.
Share logic model of the problem visual with community design team (CDT) and gather insights (in the field)	A key visual and clear language in Nepali will be used to communicate key aspects of the LMOP with the CDT. The CDT will be invited to provide feedback on what changes they wish to make by using the HCD tool Rose, Thorn, Bud. Next, the SAMIP team will incorporate the changes into (1) the visual model, and (2) the table model based on CDT feedback. (2 h)
Photo journaling	Photo journaling will be used within an asset mapping exercise to identify key assets in the community to guide the overall study. The CDT will be invited to use iPads in pairs to walk around their community and reflect upon the following prompt: “What are things that you love in or about your community?” The CDT will participate in a group discussion to view and discuss key aspects captured in each photo. (3 h)
Stakeholder mapping	The CDT will brainstorm gatekeepers/key stakeholders of chhaupadi (eg, those who practice, those who perpetuate, and those who have the power to change it). Mindsets for each stakeholder will be discussed. The CDT will vote on 5 key groups of people they deem most important for promoting change. (2 h)
Experience diagramming	The CDT will map their chhaupadi journeys, specifically key aspects of the practice and determine which aspects are liked/beneficial, and which aspects are not liked/could be improved. Participants can use a range of art tools and prepopulated icons. Ideas will be discussed as a group and combined into one combined Experience Diagram based on a group discussion. If needed, the CDT will be provided voting tokens to narrow the scope of the program goals based on (1) major influence on chhaupadi, (2) urgency, and (3) feasibility. (2 h)
Walk-a-mile immersion	The research team will visit CDT households (n=10) to explore how chhaupadi is practiced. This will result in empathy and trust building through first-hand experience and deeper understanding of the range of chhaupadi traditions. (8 h)
Impact ladder	The CDT will explore the lasting social change they wish to see in their community related to chhaupadi. They will also identify immediate outcomes that are expected to contribute to this change. They will specify a target group and timeframe for each goal. (2 h)
Community dialogue session 1 and key informant interviews (KIIs)	We will hold a community dialogue session (n=10) (2 h) and conduct KIIs with key leaders (n=8) (45 min) to share results, gather alternative perspectives, and communicate findings to community members. Findings will also be incorporated into the logic model of the problem to ensure a holistic understanding of the intervention context, including the setting and community and assets. Discussions will be audio-recorded.
IM step 2: specify program outcomes (ie, target behaviors), performance objectives, and logic model of change[d]	
Review impact ladder and share community dialogue session insights with CDT	The research team will summarize the insights gathered from the community dialogue session. The CDT will have an opportunity to revisit their impact ladder, refine, and edit based on community input. (1 h)
Change objectives, performance objectives, and logic model of change (internal preparation with CDT feedback)	Based on the insights gathered using the impact ladder review, the research team will compile the findings into a matrix of change objectives tool used in IM. The research team will share the draft matrix with the CDT members using icons and visuals. After incorporating CDT feedback, the research team will compile the feedback and draft a Logic Model of Change using the templates from the IM process. (3 h)
IM step 3: program design[e]	

XSL•FO
RenderX

Activity	Description and timing
Importance difficulty (ID) matrix	As a group, the CDT will decide where change objectives fall on an ID Matrix. On a large piece of paper, there will be an option of 4 quadrants (low difficulty, low importance; low difficulty, high importance; high difficulty, high importance; and high difficulty, low importance). Using icons for each program objective, CDT members will place them in the appropriate quadrant depending on how difficult it would be to implement, and the importance for bringing about change related to chhaupadi. (2 h)
Buy a feature	Each quadrant of the ID will be assigned a price. Participants will be given currency notes, and they will be asked to "purchase" program components within a limited budget. This will allow the CDT to explore the importance of specific activity combinations and provide an opportunity for the CDT to consider how different program components might fit together in an intervention. (2 h)
Methods for changing behavior presentation	The SAMIP team will provide an overview of evidence-based approaches to changing behavior, which participants can draw from for future exercises. (30 min)
Concept posters	The CDT will break into teams of 2 people to develop Concept Posters for the program components that they "purchased" in the previous "Buy A Feature" exercise. Each team of 2 people will be provided a poster with key areas to consider for their intervention development and will be invited to use creative supplies (eg, paints, icons, drawing and writing utencils, etc). Posters will include the following prompts: intervention goals, desired outcomes, target groups, key implementers, main components, locations, duration, and inputs required. Upon completion, they will pitch their ideas to the wider CDT. (4 h)
Rose, thorn, bud	After pitching concept posters, the CDT will use the rose, thorn, bud technique to identify aspects of the intervention that are expected to be successful ("roses"), aspects that may fail ("thorns"), and aspects that have potential but could be improved ("buds"). (2 h)
Visualize the vote	The CDT will vote for the 5 intervention components they like the most, particularly those that they believe will bring about the greatest change. (2 h)
"Coolers" color palette brainstorm	The CDT will be invited to brainstorm what colors come to mind based on their ideal chhaupadi program that they have designed in the previous steps. Everyone picks a color and explains why they selected it. (1 h)
Dream and design	The CDT will be asked to go around the community and gather materials (ie, natural elements like leaves, sticks, rocks, dirt, spices, makeup, cloth, etc) that could be used to communicate or represent their chhaupadi program. They will come back together to create a rangoli-like collage, present, and discuss. Throughout the discussion, the research team will seek to identify common themes, symbols, words, etc, around their ideal chhaupadi program. The CDT will also finalize the title of the intervention. These created materials will be used in the intervention materials to ensure it is grounded in the vision, designs, and creative ideas of participants. (4 h)
Model making	The CDT will break into teams to create "sample" activities and outputs related to their intervention. For example, one group may create a sample video if they wish to use social media videos and campaigns in their intervention. The CDT will share their outputs and creations with the full CDT. (3 h)
Visual voices	As a closing visioning exercise, the CDT will participate in visual voices. They will complete a blind drawing, and painting and writing exercises to ultimately create a group collage/mural reflecting on the following prompt: If your intervention is successful, what will your community look and feel like? (4 h)
Community validation meeting 2	The same community members from the community validation meeting 1 will be invited back to review the CDT concept posters, provide feedback, or offer alternative insights. They will vote on the intervention ideas that they believe will have the greatest impact. (2 h)

XSL•FO
RenderX

Activity	Description and timing
Focus group discussion (FGD)	The CDT will participate in an FGD, with the goal of understanding experiences using HCD to develop interventions, explore how the activities impacted them, what worked well, what could be improved, and ideas for the future. (1.5 h)

[a]Goal: Finalize a logic model of the problem (LMOP), which incorporates findings from realist synthesis (RS; aim 1), with input from the community.

[b]SAMIP: Strengthening Actions for Menstrual Health and Hygiene Interventions for Promotion of Women's Health in Nepal.

[c]RS: realist synthesis.

[d]Goal: Specify what changes related to chhaupadi and the surrounding environment need to occur to improve health and quality of life.

[e]Goal: Design a chhaupadi intervention that is evidence-based and incorporates creative ideas generated from the CDT using HCD and ABR tools, with input from community members.

After initial intervention ideas are developed in partnership with the action community, we will conduct a "rapid situation analysis," where local data enumerators will elicit feedback on the initial intervention ideas in an additional 4 chhaupadi-practicing districts with medium or high prevalence of the practice, known as "collaboration communities." Enumerators will summarize key findings from the HCD workshop, share the intervention concept with a variety of community members in the collaboration communities, and collect feedback on perceived strengths, areas for refinement, and contextual barriers to implementation. Insights gathered from collaboration communities will be critical to generating insights to guide contextual adaptation of the intervention across districts and to equip the team for effective scale-up of the approach. Finally, the team will hold an "insight workshop" at the community level to share insights gathered from the rapid situation analysis and share a final intervention concept via a one-day workshop.

In aim 3, we will conduct a clinical trial to pilot test the intervention approach in one chhaupadi-practicing community and include a control group. One chhaupadi-practicing district will be selected, but it will exclude communities visited during the previousaims. The control and pilot communities will be matched and will be far enough away to prevent spillover effects.

Study Population and Recruitment

For all aims, individuals will be included across the lifespan, age 15 years and older; there is no maximum age range for the study overall. Given the focus of this study on practices (ie, chhaupadi) that specifically impact women and girls who currently experience monthly menstruation, or have experienced menstruation, the minimum age for enrollment is 15 years. The study will be open to men, women, and gender-expansive individuals, as all these groups are expected to play key roles in the menstrual movement in Nepal and, therefore, will be eligible to participate in various aspects of the research study if they meet the inclusion criteria for each aim.

For aim 1, key informant participants at the national level must be current or previous members of a team/organization that has implemented a chhaupadi intervention in Nepal, and have expertise in chhaupadi interventions (eg, a representative from the Family Welfare Division). For the district and community level KIIs with program staff, participants must currently work or have worked at the time of the intervention for an organization implementing a chhaupadi intervention that is included in the RS. Community-level interviews will also be conducted with those who were direct recipients of the chhaupadi intervention, or who are leaders in the community where the chhaupadi intervention took place, and can speak to the effectiveness and response to the intervention. Those under the age of 15 years will be excluded. We aim to enroll a minimum of 35 participants in aim 1.

For aim 2, we will establish one CDT of 10 participants to work with the study team to co-design a chhaupadi intervention. Sample size is guided by standards in the literature for the ideal focus group size. The CDT will be comprised of women and girls who practice chhaupadi (menstruating age and older), from diverse backgrounds including occupation, age, caste/ethnicity, religion, literacy, and marital status. CDT participants will be recruited via snowball sampling and through community-level collaborators who have strong ties with local NGOs and government leaders. The CDT and all participants will be compensated for their time. Community dialogue sessions (2, n=10) will include youth, elders, religious leaders, teachers, health care workers, and other community members who can speak to the potential for a chhaupadi intervention in the community. We will also conduct KIIs (n=8) with key community leaders such as local police, government leaders, and activists with chhaupadi expertise to inform our overall understanding of the context.

The "rapid situation analysis," will focus on gathering quick snapshots from a variety of chhaupadi-practicing districts, from a range of community members, including key informants (eg, police, teachers, health post in charge, members of government), to those with lived experience practicing chhaupadi. The participants will be selected in partnership with the local municipality office of each study district.

Finally, the "insight workshop," which will be held at the district level in one of the chhaupadi-practicing districts to ensure accessibility for community members and local leaders, will bring together the original CDT members from aim 2.

In aim 3, two communities (1 intervention and 1 control) will be selected in one chhaupadi-practicing district of Nepal with high prevalence of chhaupadi based on the latest evidence at the time of data collection. We seek to identify communities that have not been exposed to a chhaupadi intervention in the past 5 years. Control and intervention communities will be matched based on exposure to chhaupadi interventions, population, rurality, and caste/ethnic makeup.

XSL•FO
RenderX

Eligibility criteria for the clinical trial will be refined based on the target groups of the final intervention developed in aim 2, but are expected to include women and girls who currently practice chhaupadi and have been residents of the communities for at least one year. We will use a sample size that allows for preliminary testing of feasibility and acceptability. Before enrolling, the research team will screen participants based on eligibility criteria, obtain informed consent, and document reasons for nonparticipation, if necessary. Communities will be randomized into the intervention or control group to eliminate bias; the control group will receive health education materials only.

Intervention Material Development

After aim 2 activities are complete, we will hire a Nepal-based educational consultant to review plans and materials, and to support with drafting the intervention materials that meet the vision of the CDT members. The consultant will also partner with the SAMIP team to develop an implementation plan, which will consider timing, locations, key players, and other implementation logistics. Additionally, we will partner with a graphic design company to ensure the materials are created, designed, and executed according to the vision of the CDT members. The graphic design team will be responsible for the design and finalizing the layout, look, and feel of the intervention materials, which will include written content (in English, Nepali, and local dialect, if necessary), illustrative graphics, an implementation guide, intervention sequence, training instructions, and other materials as needed. Overall, both the educational consultant and the graphic design company will help finalize intervention materials in preparation for the clinical trial in aim 3.

Data Collection Activities

Aim 1: RS

We seek to extend the literature by assessing existing chhaupadi interventions by exploring what works, for whom, and under what circumstances (ie, mechanisms of success). To do so, we will build upon extant studies by (1) conducting an RS to classify complex, heterogeneous intervention data; (2) triangulating intervention data with KIIs at community, district, and national levels; and (3) assessing underlying theories of interventions.

In Step 1, we will define and clarify the scope of the review, identify initial intervention theories, and conduct a preliminary search. This process will include 3 brainstorming sessions with menstrual health program experts from the MHMPA to (1) refine the research question; (2) define key terms relevant to the search (eg, How is chhaupadi defined? What are the parameters of a chhaupadi intervention?); (3) finalize search strategies; (4) develop a draft theoretical framework; (5) set study rigor threshold; and (6) finalize data extraction fields.

In Step 2, we will search and appraise the evidence. The research team will develop a search strategy and timeline, then 2 reviewers will screen titles and abstracts against inclusion/exclusion criteria using bespoke extraction forms in DistillerSR; the form will be pretested with a subset of data until disagreements are resolved. Studies/evaluations will be included if they (1) address chhaupadi via an intervention in Nepal (eg, behavioral, educational, policy, media, health systems, etc), or (2) address chhaupadi as a component of a broader intervention (eg, chhaupadi educational material as a part of a broader sexuality education program). Studies/evaluations will be excluded if they (1) are anecdotal, (2) are not relevant to chhaupadi, (3) are not based in Nepal, or (4) are published in a language other than English or Nepali.

In Step 3, the team will extract and synthesize evidence. The extraction tool will include questions related to the theory of change and the central review question and will include direct quotes from data sources. Given the iterative nature of RS, we will add additional information that is deemed important for informing our theoretical framework throughout the process, as is standard practice [54]. We will assess each study/intervention for rigor and quality by using established ranking criteria [55], and credibility and trustworthiness criteria [44]. Only studies that meet the agreed-upon threshold will be included; notably, even lesser quality evidence can be relevant for theory development in RS [56].

To reduce biases and increase the validity of the data collected via secondary sources, we will conduct interviews at the national (n=10), district (n=10), and community levels (n=15) to verify findings with experts who have implemented chhaupadi interventions, similar to Kantilal et al [55], approach for capturing real-world feedback. Sample size was determined based on recommendations from empirical studies on qualitative saturation, in which saturation is often reached within 9 interviews [57]; however, additional interviews will be added if saturation is not reached (Multimedia Appendix 1 provides details on topic guides).

In step 4, presentation of our findings will be organized according to the theoretical framework developed and refined throughout the RS, using constructs of context, mechanism, and outcome. Recommendations will be established via a final brainstorming session with experts from Step 1. Results will be reported according to the Realist and Metanarrative Evidence Synthesis Standards for quality and publication (RAMESES; Multimedia Appendix 2 provides details on the checklist) [44].

Aim 2: Intervention Cocreation

Guided by findings from aim 1, in aim 2, we aim to co-design a chhaupadi intervention that meets the unique contextual needs of communities while also incorporating current evidence and theory. This study will significantly build upon the intervention development literature by harnessing the systematic and theoretical approaches offered by IM, combined with the creative and community-centered benefits that HCD and ABR methods offer.

Drawing inspiration from Leung et al's HCD model [58] and refined through pilot testing [41,48], we will use the six steps of IM [46] combined with several creative and participatory activities to develop a contextual understanding of gatekeepers of chhaupadi, their mindsets, experiences of chhaupadi, and drivers and facilitators of the practice. We will work with an action community to co-design a chhaupadi intervention, along with community dialogue sessions and conduct KIIs to share

XSL•FO
RenderX

results, gather alternative perspectives, and communicate findings to community members (Table 2).

The design and implementation of the co-design workshop will be overseen by a PhD-trained researcher with expertise in participatory and community-engaged methods, who identifies as a woman, a mother, and a menstruator, which is expected to support in establishing rapport and trust with participants. The PI has intermediate Nepali language skills and will be supported by additional research assistants who are fluent in English and Nepali, including the local dialects, to ensure all activities are understandable for all participants and team members.

Next, local data enumerators will conduct a rapid situation analysis (n=75), in which they will summarize key findings from the HCD co-design workshop and share with community members and leaders in several chhaupadi-practicing districts. Through KIIs, they will gather feedback and considerations for contextualizing the intervention and future scale-up. Finally, the CDT will be invited to give feedback during an insight workshop, in which we will conduct several creative and participatory activities to gather insights to finalize the intervention design and implementation plan (additional data collection activities outlined in Table 2).

Aim 3: Pilot Trial

We will pilot test the chhaupadi intervention in one community and include a matched control community located sufficiently far away to minimize spillover effects. The trial will focus on assessing feasibility and acceptability, as well as collecting preliminary data on knowledge, attitudes, behaviors, and health outcomes expected to be impacted by the intervention. The intervention will be delivered by trained community organizers with expertise and experience conducting health interventions in rural Nepal. As a research organization, our partner Bikash Shrot Kendra will provide support with intervention evaluation and ensure implementation is consistent across pilot communities. We estimate that the intervention will include 8 - 10 two-hour sessions in the community, over the course of 4 months, estimated based on other intervention approaches for shifting cultural norms.

Quantitative outcome measures will be used to assess feasibility and acceptability, and results will inform the design of a subsequent trial. We will conduct focus group discussions (FGDs) with participants and the intervention team to discuss challenges, barriers, and opportunities to adjust the intervention material and protocol (mid-point and post). Last, we will follow up with participants six months after completion to further assess outcomes.

Analysis

Overview

We will use StataSE for quantitative analyses. For qualitative analyses, all audio-recorded data will be transcribed and translated from Nepali to English by trained translators who are fluent in both languages. Photographs, visuals, and diagrams collected as part of the HCD activities will also be transcribed/translated using standardized spreadsheets in Microsoft Excel. Qualitative analyses will be conducted in Dedoose. For all aims, we will collect demographic information about each study participant, which will be used to describe the study sample based on key demographic variables. Intervention reporting will follow the Template for Intervention Description and Replication (TIDieR) checklist (Checklist 1 provides more information).

Specific Analysis Plans for Aim 1

In the analysis of studies/intervention reports for the RS, we will scrutinize the evidence, specifically each study/intervention by context, mechanisms, and outcomes, aligned with RS standards [45]. Context will explore the conditions under which the intervention impacts outcomes (eg, geographic location and cultural norms). Mechanisms, or processes operating within an intervention, will explore how participants use the resources offered through the intervention. Outcomes will include the consequences of the study/intervention, intended or unintended, resulting from the interaction between context and mechanism (eg, improved mental health outcomes and reduced number of women practicing chhaupadi) [56]. We will use these components to describe the extent and circumstances under which an intervention works or does not, and to test and refine our theoretical framework drafted in Step 1. Screened data will be exported from Qualtrics forms and discussed with the review team to create data evidence tables [45] organized by key theme (eg, context, mechanism, and outcome). Next, we will assess the data for a chain of inference to explore connections across studies/interventions, as is standard practice in RS [45,56].

KIIs (conducted in Nepali or English based on participant preference) will be audio recorded, transcribed, and translated into English as necessary. Two coders will analyze data thematically in NVivo (Lumivero). Presentation of our findings will be organized according to the theoretical framework developed and refined throughout the RS, using constructs of context, mechanism, and outcome. Recommendations will be established via a final brainstorming session with experts from Step 1. Results will be reported according to the RAMESES for quality and publication [44].

Specific Analysis Plans for Aim 2

The resulting data from the HCD workshops will include the following. (1) Field notes: reflections on what worked effectively and what could be improved. (2) HCD artifacts: audio recordings and photographs of the outputs for all HCD activities (eg, stakeholder map, walk a mile immersion, photo journaling, experience diagram, impact ladder, etc) transcribed/translated into standardized Excel sheets for further analysis. This will include direct quotes from the activity discussion from audio recordings, which will be transcribed and translated into English. (3) FGDs and KIIs: audio recordings of FGDs and KIIs will be transcribed and translated from Nepali to English. All data will be qualitatively analyzed thematically in Dedoose and guided by the Health Stigma and Discrimination Framework [59].

Based on findings from the co-analysis workshops, the PI will lead the development of the intervention materials, implementation protocol, and training manual. The educational consultant will be responsible for creating presentations

XSL•FO
RenderX

providing overviews of each intervention module, including material lists, goals, and activity instructions. Intervention materials will include written content (in English, Nepali, and local dialect, if necessary), illustrative graphics, an implementation guide, and activity instructions. The graphic design company will ensure that all intervention materials, such as training instructions or an implementation guide, are in accordance with the vision of the CDT members. Based on piloting, we posit that the core intervention could include (1) activities at multiple levels of the social ecological model, considering the influence of individual knowledge, family beliefs, social pressure, and policies on menstrual behaviors; (2) multiple target groups for behavior change, including but not limited to mothers/women, youth, female community health volunteers, and religious leaders, each with their own behavior change messaging and approaches (eg, community group sessions vs radio/newspaper vs one-on-one education); and (3) sessions over the course of several months, since social norms and behaviors require considerable time to shift.

We will elicit feedback on the intervention content and protocol using cognitive interviewing, an evidence-based method for determining if the proposed tool fulfills its planned purpose [60], with 10 experts from MHMPA and BSK before the pilot test in aim 3.

Specific Analysis Plans for Aim 3

This pilot is not intended to power between-arm comparisons; however, findings will allow for an initial assessment of feasibility and acceptability in preparation for a future randomized controlled trial. Aligned with pilot study best practices, we will describe the intervention sample using participant characteristics (n and % for categorical variables and mean and SD for continuous variables). We will calculate the approached-to-enroll and completion rates compared with prespecified benchmarks (Table 3) and report the percent of sessions completed as planned, participant completion rates, as well as outcomes for each measure (frequency and %, mean and SD, or median and range). Since this is an initial pilot test, no formal efficacy tests will be conducted by the comparison group; however, we will run exploratory analyses on outcomes by caste/ethnicity, age, education, and religion for descriptive purposes only. We will also test for differences in potential confounders between the intervention and control groups and adjust for them in regression models. All quantitative analyses will be conducted in Stata SE (StataCorp LLC).

XSL•FO
RenderX

Table . Measures and tools for aim 3 data collection.

Measures/tools	Pre	Mid	Post	6 month
Demographics (eg, sex, caste/ethnicity, religion, education, occupation, and age)	✓			
Intervention Pilot Testing Measures				
Enrollment: Enrollment pace, and retention (Benchmark: ≥80% of approached individuals enrolled; ≥80% retention)			✓	
Acceptability: For participants and implementers, acceptability of intervention measure [61] a valid and reliable measure of acceptability with 4 items on a 5-point Likert (Benchmark: ≥3.5) and open-ended questions on acceptability of intervention components			✓	
Appropriateness: For participants, the intervention appropriateness measure [61] a valid and reliable measure of appropriateness with 4 items on a 5-point Likert scale (Benchmark: ≥3.5)			✓	
Feasibility: Extent to which the intervention is successfully executed. Feasibility of Intervention Measure [61] a valid and reliable measure of feasibility with 4 items on a 5-point Likert scale (Benchmark: ≥3.5)			✓	
Challenges, barriers, opportunities: FGDs[a] with implementers (n=1) and participants (n=2)		✓	✓	
Exploratory intervention outcome measures				
Mental Health: Beck Anxiety Inventory Validated for Nepal [62] 21 items, 4-point Likert	✓		✓	✓
Reproductive Health Questionnaire [17]	✓		✓	✓
Chhaupadi knowledge, attitudes, and behaviors	✓		✓	✓
Self-efficacy: Generalized Self-Efficacy Scale, 10 items, 4-point Likert [63]	✓		✓	✓

[a]FGD: focus group discussion.

Qualitative analyses of open-ended and FGD data will be coded thematically in NVivo. We will professionally transcribe and translate the data, and 2 independent members of the research team will analyze data using template analysis to enhance rigor, allowing for inductive and deductive analysis [64].

XSL•FO
RenderX

Ethical Considerations

Given the stigmatized nature of menstruation in Nepal, as well as the legal measures banning and criminalizing the practice, it is necessary to take every precaution to prevent and mitigate risk or harm to participants. To that end, this study received approval from both the University of Pittsburgh Institutional Review Board (study number: MOD23080043-001) as well as the Nepal Health Research Council Ethical Review Board (Protocol Number: 640_2023). Every effort will be made to ensure there are no changes in the risk/benefit ratio during the study and that confidentiality of research data is maintained. All records related to this study will be stored in secure databases only accessible by research staff or in a locked file cabinet. Additionally, all research staff have completed training on the ethical principles of research, highlighting confidentiality to minimize the risk of breaches of confidentiality.

The research team will obtain informed consent from every adult participant in a manner that is suitable to the context. In aim one, the research team will collect consent via a Qualtrics survey via mobile device or pen and paper, given the participant's preference and technological literacy. In aim 2, informed consent for participation in the co-design workshops, community validation meetings, KIIs, and cognitive interviews will be obtained in person via Qualtrics on a mobile device prior to collecting data. Obtaining informed consent for pilot testing the intervention in aim 3 will follow similar procedures to aim 2.

Participants younger than 18 years will be asked to provide assent, and the research team will ask a parent or guardian to sign the consent form. All consent forms will outline the following details: goals of the study; foreseeable risks or potential discomfort; how confidentiality will be maintained; potential benefits; contact information for the PI and research assistant; information about the voluntary nature of the research study, that they can withdraw at any time, and can skip any questions that they do not wish to answer. Consent materials will be available in English and Nepali, and participants can choose which language they prefer.

To minimize coercion and undue influence, participants will be informed about the study in advance of the study team coming to the village, so they will have ample time to reflect on whether they would like to participate. We will also hold an informational meeting with potential participants to answer any questions that they may have before initiating the consent process.

The research team developed additional safety plans for the pilot study in aim 3 to address unlikely but possible inadvertent effects of the intervention on stigmatizing the practice. This includes (1) continuously monitoring community impacts of the intervention design and associated discussions (eg, through community validation meetings and KIIs); (2) stigma will not be used to promote global health in any aspect of the chhaupadi intervention [65], which will be discussed with the CDT during aim 2; (3) applying an intersectional lens throughout the study to ensure that intersecting forms of one's identity are considered and protected throughout the study (eg, caste/ethnicity, religious, gender, rurality, marital status, education); (4) establishing a community safety advisory board (CSAB) in the intervention community to focus on risk management. During the intervention implementation, the PI will meet with the CSAB monthly to monitor data collection, management processes, examine the integrity of data analysis, and review feedback. Any concerns regarding potential risk to the participants and broader community will be seriously considered, and the PI will develop rapid risk mitigation plans as needed; (5) implementing measures in pre, mid, and postintervention surveys to assess shifts in stigma in real-time within the intervention community. This will allow the team to capture potentially negative impacts early and address them in partnership with the CSAB; and (6) asking directly about stigma experiences in FGDs at the mid and end points of the intervention to better understand the nature of stigma experiences and how they may be impacted by the intervention.

Results

The study was funded in September 2023. Aim 1 activities began in April 2024. During the review, the study team identified 2033 resources for potential inclusion in the RS. After reviewing, 69 materials met the inclusion criteria and were analyzed in detail, which included peer-reviewed manuscripts, evaluation documents, program reports, videos, and more. To supplement findings from secondary sources, a workshop with 8 organizations was held in the capital to generate initial chhaupadi program theories and identify initial drafts of causal mechanisms for further analysis. These initial program theories and insights from the secondary sources were discussed at the national and district levels via consultations with 44 individuals. Analysis was completed in February 2026. Evidence suggests that sustainable change requires the creation of enabling environments, incremental behavior change, and the active involvement of families and local leaders, alongside attention to structural determinants such as economic insecurity and women's financial empowerment. Full results are expected in Summer 2026.

Aim 2 data collection began in March 2025. A 10-day co-design workshop using HCD tools and ABR methods was implemented in Dailekh with 11 women and CDT members, and they created a draft intervention titled, : (Our Voice: Transforming Lives through Ending Chhaupadi). An additional 123 community members have been consulted, both in the design community as well as across 4 additional districts (Achham, Baitadi, Doti, and Surkhet). The refinement of the intervention is ongoing based on community feedback. Results are anticipated in Fall 2026.

Aim 3 data collection is projected to begin in Winter 2027.

Discussion

Principal Findings

This study protocol offers several methodological innovations for global health intervention synthesis and development. It is the first known study to harness IM, partnered with HCD and ABR methods for intervention development in Nepal. The

XSL•FO
RenderX

approach outlined in this paper builds upon pilot work our team conducted in Dailekh, Nepal, in 2022 [41,48,50].

This study protocol offers several methodological innovations to the field of global women's health. We are the first to apply RS to menstrual health in Nepal. Unlike previous approaches, before developing an intervention, we will first assess interventions, enabling us to understand what elements work, in what contexts, and how [45].

This is also the first known example of partnering with communities to co-design an intervention on this topic and with this population. Although several interventions have targeted chhaupadi, such as elimination campaigns, community sensitization, and radio shows, the practice persists in parts of the country [38]. Other initiatives have dismantled menstrual sheds or withheld state support services, but have often resulted in increased vulnerabilities, forcing women to live in harsher conditions [37,66]. Shortcomings of these interventions (eg, top-down approaches) have prevented advancement of the field, likely because they lack the integration of community voices in all aspects of the intervention design, implementation, and assessment. Our preliminary pilot evidence suggests a community-engaged approach will result in a locally relevant response to shifting deep-rooted behaviors. This protocol offers HCD and ABR tools as relevant ways in which communities with diverse abilities, expertise, and backgrounds can be incorporated at all stages of the intervention design and planning process.

Finally, this will be a novel use of HCD and ABR methods for intervention development in Nepal. While HCD has been successfully applied to co-design interventions in several contexts in which culture and social norms play key roles [58,67], including for reproductive health [68,69], this will be the first application of HCD to develop a full menstrual health intervention in Nepal. By developing and piloting the intervention, guided by IM principles, we will contribute significant knowledge of culturally appropriate intervention attributes that can be scaled and rigorously tested in future studies.

Limitations

While our study is built on the strengths of community engagement and theory- and evidence-based intervention development, it is not without limitations. A key limitation of the RS in aim 1 is that its findings depend on the quality and detail of existing literature and stakeholder input, which may vary and result in incomplete or context-specific explanations that are not universally generalizable. While we will seek to recruit a diverse sample throughout all aims, we are limited by time, budget, and transportation/road access, and therefore, our sample is not representative of all communities and individuals who partake in the chhaupadi tradition. Additionally, while community members will design intervention concepts, provide multiple rounds of feedback, and we will triangulate the findings in multiple districts, we anticipate that additional intervention development design work will be required by the research team to turn the community ideas into replicable intervention activities. Additional contributions from the research team, designers, content area experts, trainers, and program developers will be harnessed to build upon the foundation established by community members. We will incorporate the voices of the community advisory board as we develop the protocols, training manuals, and tools, but this part of the process will primarily be led by the research team and is a limitation due to time and budget. Finally, limitations in our timeline and budget may influence the complexity of the intervention that we are able to fully develop and test.

Conclusions

Despite numerous efforts, a comprehensive understanding of the effects of chhaupadi interventions across Nepal remains absent. This study will use RS to identify what works, for whom, and under which circumstances. Insights from the RS will directly inform co-designed interventions with community members, leveraging IM, HCD, and ABR methods. By embedding diverse community perspectives at every stage of the study, this approach ensures interventions are contextually grounded and culturally relevant. Beyond addressing a culturally situated menstrual health practice—chhaupadi—this community-centered model has the potential to inform strategies for reducing the harms associated with other harmful practices in diverse global settings.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. In accordance with the GAIDeT taxonomy (2025), GAI tools were used under full human supervision for literature search and systematization, evaluation of research novelty and identification of knowledge gaps, visualization, proofreading and editing, summarizing text, and reformatting. The GAI tool used was ChatGPT-5 Mini. Responsibility for the content and integrity of the final manuscript rests entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

Research reported in this publication was supported by Fogarty International Center of the National Institutes of Health under award number K01TW012424 (PI: SEB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

XSL•FO
RenderX

Data Availability

Due to the sensitive nature of the data and the illegal status of chhaupadi in Nepal, individual- and community-level data cannot be publicly shared to protect participants. Deidentified data supporting the findings may be made available from the corresponding author upon reasonable request, subject to a data use agreement to ensure confidentiality and ethical compliance.

Authors' Contributions

SEB conceptualized the study's aims, overall design, and research questions guiding the multiphase project. SEB, MH, BD, SP, EM, AY, HS, and AA developed the methodological frameworks, integrating realist synthesis, intervention mapping, human-centered design, and arts-based research. SEB, MH, BD, EM, AY, HS, AA, and SM designed the analytical approaches for quantitative, qualitative, and mixed methods data, including synthesis frameworks and pilot outcome analysis plans. SEB, NS, HS, and AA conducted the investigation, including literature searches, key informant interviews at national and field levels, and co-design workshop activities, as well as coordinating participant recruitment and data collection. Data curation was carried out by SEB, BD, NS, HS, SM, and AA, who managed and prepared datasets—including qualitative transcripts, survey data, and visual artifacts—for analysis, including transcription and translation processes. SEB, HS, and AA drafted the initial manuscript. BD, EM, AY, SP, SM, NS, AS, MS, NK, and MH reviewed and provided substantive revisions, including refining the conceptual framework, interpreting study progress, and contextualizing findings. Project administration was coordinated by NS, BD, and SEB, who managed overall study implementation, including multisite logistics, community engagement, and collaboration with local partners. Supervision was provided by SEB, MH, BD, SP, EM, AY, AS, and MS, who oversaw research activities and provided methodological guidance and mentorship to research assistants and junior team members. SEB secured funding for the project and managed grant reporting requirements. BD, NS, NK, MS, AS, and SP provided resources by facilitating access to community partners, organizations implementing chhaupadi interventions, and logistical support for field activities.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview and focus group topic guides.
[DOCX File, 19 KB - resprot_v15i1e89117_app1.docx]

Multimedia Appendix 2
RAMESES publication standards for realist synthesis.
[DOCX File, 17 KB - resprot_v15i1e89117_app2.docx]

Checklist 1
The TIDieR checklist.
[DOCX File, 29 KB - resprot_v15i1e89117_app3.docx]

Peer Review Report 1
Peer review report by the ZRG1 ICP-P (55) - Center for Scientific Review Special Emphasis Panel, PAR Panel: International and Cooperative Projects (National Institutes of Health, USA).
[PDF File, 3792 KB - resprot_v15i1e89117_app4.pdf]

References

1. Joshi L. Chhaupadi Pratha: socio-cultural violence against women in the far-western region of Nepal. JNEPHA 2015;6(6):22-32.
2. Sharma S, Deuja S, Saha CG. Menstrual pattern among adolescent girls of Pokhara Valley: a cross sectional study. BMC Womens Health 2016 Dec 9;16(1):74. [doi: 10.1186/s12905-016-0354-y] [Medline: 27938370]
3. Guidance on menstrual health and hygiene. UNICEF. 2019. URL: https://www.unicef.org/documents/guidance-menstrual-health-and-hygiene [accessed 2026-04-07]
4. Bobel C. Introduction: menstruation as lens - menstruation as opportunity. In: Bobel C, Winkler I, Fahs B, Hasson KA, Kissling EA, Roberts TA, editors. The Palgrave Handbook of Critical Menstruation Studies: Palgrave Macmillan; 2020.
5. Kuhlmann AS, Henry K, Wall LL. Menstrual hygiene management in resource-poor countries. Obstet Gynecol Surv 2017 Jun;72(6):356-376. [doi: 10.1097/OGX.0000000000000443] [Medline: 28661550]
6. Mason L, Nyothach E, Alexander K, et al. We keep it secret so no one should know'--a qualitative study to explore young schoolgirls attitudes and experiences with menstruation in rural western Kenya. PLoS ONE 2013;8(11):e79132. [doi: 10.1371/journal.pone.0079132] [Medline: 24244435]

XSL•FO
RenderX

7. Ali TS, Rizvi SN. Menstrual knowledge and practices of female adolescents in urban Karachi, Pakistan. J Adolesc 2010 Aug;33(4):531-541. [doi: 10.1016/j.adolescence.2009.05.013] [Medline: 19589587]
8. Garg S, Sharma N, Sahay R. Socio-cultural aspects of menstruation in an urban slum in Delhi, India. Reprod Health Matters 2001 May;9(17):16-25. [doi: 10.1016/s0968-8080(01)90004-7] [Medline: 11468832]
9. Hennegan J, Montgomery P. Do menstrual hygiene management interventions improve education and psychosocial outcomes for women and girls in low and middle income countries? a systematic review. PLoS ONE 2016;11(2):e0146985. [doi: 10.1371/journal.pone.0146985] [Medline: 26862750]
10. Sommer M, Ackatia-Armah N, Connolly S, Smiles D. A comparison of the menstruation and education experiences of girls in Tanzania, Ghana, Cambodia and Ethiopia. Compare 2015 Jul 4;45(4):589-609. [doi: 10.1080/03057925.2013.871399]
11. Menstrual hygiene a human rights issue: a simple guide to ending discrimination, abuse. Human Rights Watch. 2017 Aug 27.
12. Literature review of harmful practices in Nepal. : United Nations Population Fund (UNFPA); 2019 URL: https://nepal.unfpa.org/sites/default/files/pub-pdf/Literature%20Review%20on%20Harmful%20Practices%20in%20Nepal.pdf [accessed 2026-04-17]
13. Baumann SE, Lhaki P, Burke JG. Assessing the role of caste/ethnicity in predicting menstrual knowledge, attitudes, and practices in Nepal. Glob Public Health 2019 Sep;14(9):1288-1301. [doi: 10.1080/17441692.2019.1583267] [Medline: 30785854]
14. Budha A. Violence in a form of social practice against women in nepal. In: Asian Center for Women's Studies: Ewha Womans University; 2018, Vol. 1:58-65.
15. Khadka N. Chhaupadi Pratha: Women's Condition and Suffering: Tribhuvan University; 2014.
16. Kadariya S, R. Aro A. Chhaupadi practice in Nepal – analysis of ethical aspects. MB 2015:53. [doi: 10.2147/MB.S83825]
17. Ranabhat C, Kim CB, Choi EH, Aryal A, Park MB, Doh YA. Chhaupadi culture and reproductive health of women in Nepal. Asia Pac J Public Health 2015 Oct;27(7):785-795. [doi: 10.1177/1010539515602743] [Medline: 26316503]
18. Joshi PR. Role of social norms and education in utilization of postnatal services in bajhang district, nepal. In: Dissertation for the Degree of Doctor of Philosophy in Health Education: Tribhuvan University; 2023.
19. Karki KB, Poudel PC, Rothchild J, et al. Scoping review and preliminary mapping: menstrual health and hygiene management in nepal. : Population Services International Nepal; 2017.
20. Abolishing chhaupadi, breaking the stigma of menstruation in rural Nepal. UN Women. 2017. URL: https://www.unwomen.org/en/news/stories/2017/4/feature-abolishing-chhaupadi-breaking-the-stigma-of-menstruation-in-rural-nepal [accessed 2026-04-07]
21. Baumann SE, Merante MM, Wiltrout CT, et al. Is criminalization the answer? perspectives of community members and police on menstrual seclusion policy in far-west Nepal. Health Policy Plan 2021 Aug 12;36(7):1003-1012. [doi: 10.1093/heapol/czab032] [Medline: 34051090]
22. Baumann SE, Lhaki P, Terry MA, et al. Beyond the menstrual shed: exploring caste/ethnic and religious complexities of menstrual practices in far-west Nepal. Womens Reprod Health (Phila) 2021 Jan 2;8(1):1-28. [doi: 10.1080/23293691.2020.1861413]
23. Amatya P, Ghimire S, Callahan KE, Baral BK, Poudel KC. Practice and lived experience of menstrual exiles (Chhaupadi) among adolescent girls in far-western Nepal. PLoS One 2018;13(12):e0208260. [doi: 10.1371/journal.pone.0208260] [Medline: 30532183]
24. Cousins S. In Nepal, tradition is killing women. Foreign Policy. 2019 Jan 6.
25. United Nations Resident and Humanitarian Coordinator's Office. Chhaupadi in the far-west. United Nations. 2011. URL: https://un.info.np/Net/NeoDocs/View/2401 [accessed 2026-04-07]
26. Svare JA, Schmidt H, Hansen BB, Lose G. Bacterial vaginosis in a cohort of Danish pregnant women: prevalence and relationship with preterm delivery, low birthweight and perinatal infections. BJOG 2006 Dec;113(12):1419-1425. [doi: 10.1111/j.1471-0528.2006.01087.x] [Medline: 17010117]
27. Das P, Baker KK, Dutta A, et al. Menstrual hygiene practices, WASH access and the risk of urogenital infection in women from Odisha, India. PLoS One 2015;10(6):e0130777. [doi: 10.1371/journal.pone.0130777] [Medline: 26125184]
28. Upadhyay P. Menstruation pollution taboos and gender based violence in western Nepal. NEHU J 2017;15(2):101-111.
29. Cardoso LF, Clark CJ, Rivers K, Ferguson G, Shrestha B, Gupta J. Menstrual restriction prevalence and association with intimate partner violence among Nepali women. BMJ Sex Reprod Health 2018 Sep 28;45(1):38-43. [doi: 10.1136/bmjsrh-2017-101908] [Medline: 30266716]
30. Lama D, Kamaraj R. Maternal and Child Health Care in Chhaupadi Pratha, Social Seclusion of Mother and Child after Delivery in Achham: SRM University Public Health Research Series; 2015:22-33.
31. Abolition of chhaupadi in the far and mid-western region of Nepal: baseline report. : Restless Development Nepal; 2015.
32. Morgan R, Kalbarczyk A, de la Paz M, Boy-Mena E, Coates A. Addressing harmful gender norms in sexual health services. Bull World Health Organ 2024 Dec 1;102(12):910-912. [doi: 10.2471/BLT.23.291155] [Medline: 39611189]
33. Olson MM, Alhelou N, Kavattur PS, Rountree L, Winkler IT. The persistent power of stigma: a critical review of policy initiatives to break the menstrual silence and advance menstrual literacy. PLOS Glob Public Health 2022;2(7):e0000070. [doi: 10.1371/journal.pgph.0000070] [Medline: 36962272]

XSL•FO
RenderX

34. Fact sheet no 23, harmful traditional practices affecting the health of women and children. OHCHR. URL: https://www.ohchr.org/sites/default/files/Documents/Publications/FactSheet23en.pdf [accessed 2026-02-28]
35. Garg S, Anand T. Menstruation related myths in India: strategies for combating it. J Family Med Prim Care 2015;4(2):184-186. [doi: 10.4103/2249-4863.154627] [Medline: 25949964]
36. Parker S, Standing K. Nepal's menstrual huts: what can be done about this practice of confining women to cow sheds? Article 2019 Jan 23. [doi: 10.64628/AB.x93gtcpfp]
37. Thakuri DS, Thapa RK, Singh S, Khanal GN, Khatri RB. A harmful religio-cultural practice (chhaupadi) during menstruation among adolescent girls in Nepal: prevalence and policies for eradication. PLoS ONE 2021;16(9):e0256968. [doi: 10.1371/journal.pone.0256968] [Medline: 34469491]
38. Thapa S, Aro AR. "Menstruation means impurity": multilevel interventions are needed to break the menstrual taboo in Nepal. BMC Womens Health 2021 Feb 28;21(1):84. [doi: 10.1186/s12905-021-01231-6] [Medline: 33639917]
39. Nepal Fertility Care Center (NFCC). Assessment study on chhaupadi in Nepal: towards a harm reduction strategy. NHSP. 2015. URL: http://nhsp.org.np/wp-content/uploads/formidable/7/Chhaupadi-FINAL.pdf [accessed 2026-04-07]
40. Dignity for women project (DWP) 2021-2024. Green Tara. 2025. URL: http://greentara.org.np/project/15/details#:~:text=Vision%20Mission%20Goal,3 [accessed 2026-02-28]
41. Baumann SE, Rabin MA, Hawk M, et al. From stigma to solutions: harnessing local wisdom to tackle harms associated with menstrual seclusion (chhaupadi) in Nepal. Cult Health Sex 2025 Mar 4;27(3):299-320. [doi: 10.1080/13691058.2024.2373793] [Medline: 38970806]
42. Betron M, Thapa A, Amatya R, et al. Should female community health volunteers (FCHVs) facilitate a response to gender-based violence (GBV)? a mixed methods exploratory study in Mangalsen, Nepal. Glob Public Health 2021 Oct;16(10):1604-1617. [doi: 10.1080/17441692.2020.1839929] [Medline: 33186506]
43. Ghimire A, Neupane MS. Educating for equity, navigating reality: community health nurses implementing interventions to eradicate Chhaupadi in Karnali, Nepal. BMC Nurs 2025;24(1). [doi: 10.1186/s12912-025-04012-w]
44. Wong G, Greenhalgh T, Westhorp G, Buckingham J, Pawson R. RAMESES publication standards: realist syntheses. BMC Med 2013 Jan 29;11(1):21. [doi: 10.1186/1741-7015-11-21] [Medline: 23360677]
45. Rycroft-Malone J, McCormack B, Hutchinson AM, et al. Realist synthesis: illustrating the method for implementation research. Implement Sci 2012 Apr 19;7(1):33. [doi: 10.1186/1748-5908-7-33] [Medline: 22515663]
46. Bartholomew EL, Markham C, Ruiter R, Fernández M, Kok G, Parcel G. Planning Health Promotion Programs: An Intervention Mapping Approach, 4th edition: Jossey-Bass; 2016.
47. Menstrual Health and Hygiene Management Programme Alliance (MHMPA). URL: https://mhmpa.org/ [accessed 2026-04-07]
48. Baumann SE, Rabin MA, Devkota B, et al. Centering communities in global health: using human-centered design to facilitate collaboration and intervention development. Community Health Equity Res Policy 2025 Jan;45(2):167-185. [doi: 10.1177/2752535X241264331] [Medline: 39044545]
49. Holeman I, Kane D. Human-centered design for global health equity. Inf Technol Dev 2020 Jul 2;26(3):477-505. [doi: 10.1080/02681102.2019.1667289]
50. Baumann S, Hawk M, Devkota B, Burke J. Harnessing human-centered design (HCD) for intervention development in global public health: a community-engaged pilot study in nepal. Presented at: Consortium of Universities for Global Health Annual Conference; Apr 14-16, 2023.
51. Leavy P. Method Meets Art: Arts-Based Research Practice, 2nd edition: The Guilford Press; 2015.
52. Fraser KD, al Sayah F. Arts-based methods in health research: a systematic review of the literature. Arts Health 2011 Sep;3(2):110-145. [doi: 10.1080/17533015.2011.561357]
53. Boydell KM, Gladstone BM, Volpe T, Allemang B, Stasiulis E. The production and dissemination of knowledge: a scoping review of arts-based health research. Forum Qual Soc Res 2012. [doi: 10.17169/fqs-13.1.1711]
54. Wong G, Greenhalgh T, Pawson R. Internet-based medical education: a realist review of what works, for whom and in what circumstances. BMC Med Educ 2010 Feb 2;10(1):12. [doi: 10.1186/1472-6920-10-12] [Medline: 20122253]
55. Kantilal K, Hardeman W, Whiteside H, Karapanagiotou E, Small M, Bhattacharya D. Realist review protocol for understanding the real-world barriers and enablers to practitioners implementing self-management support to people living with and beyond cancer. BMJ Open 2020 Sep 3;10(9):e037636. [doi: 10.1136/bmjopen-2020-037636] [Medline: 32883731]
56. Khalil H, Peters M, Godfrey CM, McInerney P, Soares CB, Parker D. An evidence-based approach to scoping reviews. Worldviews Evid Based Nurs 2016 Apr;13(2):118-123. [doi: 10.1111/wvn.12144] [Medline: 26821833]
57. Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation. Qual Health Res 2017 Mar;27(4):591-608. [doi: 10.1177/1049732316665344] [Medline: 27670770]
58. Leung CL, Naert M, Andama B, et al. Human-centered design as a guide to intervention planning for non-communicable diseases: the BIGPIC study from Western Kenya. BMC Health Serv Res 2020 May 12;20(1):415. [doi: 10.1186/s12913-020-05199-1] [Medline: 32398131]
59. Stangl AL, Earnshaw VA, Logie CH, et al. The health stigma and discrimination framework: a global, crosscutting framework to inform research, intervention development, and policy on health-related stigmas. BMC Med 2019 Feb 15;17(1):31. [doi: 10.1186/s12916-019-1271-3] [Medline: 30764826]

XSL•FO
RenderX

60. Abi Ramia J, Harper Shehadeh M, Kheir W, et al. Community cognitive interviewing to inform local adaptations of an e-mental health intervention in Lebanon. Glob Ment Health (Camb) 2018;5:e39. [doi: 10.1017/gmh.2018.29] [Medline: 30637112]
61. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
62. Kohrt BA, Koirala NR, Sharma VD, Nepal MK. Validation of the Nepali version of beck anxiety inventory. J Inst Med Nepal 2024;25(3 & amp; amp; 4):5-9. [doi: 10.59779/jiomnepal.166]
63. Luszczynska A, Scholz U, Schwarzer R. The general self-efficacy scale: multicultural validation studies. J Psychol 2005 Sep;139(5):439-457. [doi: 10.3200/JRLP.139.5.439-457] [Medline: 16285214]
64. Brooks J, King N. Doing Template Analysis: Evaluating an End-of-Life Care Service: SAGE Publications, Ltd; 2014. [doi: 10.4135/9781446273050135127555]
65. Brewis A, Wutich A. Why we should never do it: stigma as a behaviour change tool in global health. BMJ Glob Health 2019;4(5):e001911. [doi: 10.1136/bmjgh-2019-001911] [Medline: 31750003]
66. Adhikari R. Bringing an end to deadly “menstrual huts” is proving difficult in Nepal. BMJ 2020 Feb 14;368:m536. [doi: 10.1136/bmj.m536] [Medline: 32060070]
67. Tijani B, Jaiyeola T, Oladejo B, Kassam Z. Improving data integrity in public health: a case study of an outbreak management system in Nigeria. Glob Health Sci Pract 2021 Nov 29;9(Suppl 2):S226-S233. [doi: 10.9745/GHSP-D-21-00240] [Medline: 34845046]
68. Ippoliti N, Sekamana M, Baringer L, Hope R. Using human-centered design to develop, launch, and evaluate a national digital health platform to improve reproductive health for Rwandan youth. Glob Health Sci Pract 2021 Nov 29;9(Suppl 2):S244-S260. [doi: 10.9745/GHSP-D-21-00220] [Medline: 34845048]
69. Blynn E, Harris E, Wendland M, et al. Integrating human-centered design to advance global health: lessons from 3 programs. Glob Health Sci Pract 2021 Nov 29;9(Suppl 2):S261-S273. [doi: 10.9745/GHSP-D-21-00279] [Medline: 34845049]

Abbreviations

ABR : Arts-Based Research
CDT: Community Design Team
CSAB: Community Safety Advisory Board
FGD : Focus Group Discussion
HCD: Human-Centered Design
IM : Intervention Mapping
KII: Key Informant Interview
MHMPA: Menstrual Health and Hygiene Partner’s Alliance
RAMSES: Realist and Metanarrative Evidence Synthesis Standards for quality and publication
RS: Realist Synthesis
SAMIP: Strengthening Actions for Menstrual Health and Hygiene Interventions for Promotion of Women’s Health in Nepal
TIDieR: Template for Intervention Description and Replication

Edited by A Schwartz; submitted 07.Dec.2025; this is a non–peer-reviewed article;accepted 11.Mar.2026; published 23.Apr.2026.

Please cite as:
Baumann SE, Schwarz H, Agarwal A, Devkota B, Miller E, Youk A, Parker S, Murthy S, Sharma N, Shrestha A, Subedi M, Khadka N, Hawk M
Strengthening Actions for Menstrual Health and Hygiene Interventions for Promotion of Women’s Health in Nepal (SAMIP): Protocol for a Participatory Intervention Development Study Using Realist Synthesis, Human-Centered Design, Intervention Mapping, and Arts-Based Methods
JMIR Res Protoc 2026;15:e89117
URL: https://www.researchprotocols.org/2026/1/e89117
doi:10.2196/89117

© Sara Elizabeth Baumann, Hannah Schwarz, Annika Agarwal, Bhimsen Devkota, Elizabeth Miller, Ada Youk, Sara Parker, Sanjana Murthy, Nikita Sharma, Archana Shrestha, Madhusudan Subedi, Nirajan Khadka, Mary Hawk. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is

XSL•FO
RenderX

properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Assessing Neonatal Transport, Health Outcomes, and Psychosocial Inequities Among Black, Very Preterm Infants and Their Families to Advance Health Equity: Protocol for a Mixed Methods Study

Karen F Warren[1], RN, PhD; Tisha Felder[1], MSW, PhD; Nansi Boghossian[2], MPH, PhD; Robin M Dawson[1], APRN, PhD; Robin B Dail[1], RN, PhD

[1]Department of Biobehavioral Health and Nursing Science, College of Nursing, University of South Carolina, 1601 Greene Street, Columbia, SC, United States

[2]Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC, United States

Corresponding Author:
Karen F Warren, RN, PhD
Department of Biobehavioral Health and Nursing Science, College of Nursing, University of South Carolina, 1601 Greene Street, Columbia, SC, United States

Abstract

Background: Despite medical advancements, preterm birth rates in the United States remain high and contribute substantially to infant mortality and long-term morbidity, with Black families disproportionately affected. Very preterm (VPT) infants—born between 24 and 30 weeks' gestation—are frequently delivered at community hospitals that lack advanced neonatal care and therefore require postnatal transport to tertiary neonatal intensive care units (NICUs). While neonatal transport is essential for optimizing infant outcomes, it can create additional challenges for families, including geographic separation from the NICU, disruptions to early parent-infant bonding, and increased parental psychosocial stress. These challenges may be further intensified by social vulnerability and experiences of perceived discrimination in health care settings.

Objective: This study examines associations between neonatal transport, infant health outcomes, and the psychosocial well-being of Black parents with VPT infants. It also examines how structural inequities and social vulnerability shape experiences from delivery through NICU hospitalization.

Methods: This protocol describes a longitudinal mixed methods multiple case study at a single perinatal center in South Carolina. Up to 10 cases of VPT Black infants transported from community hospitals between May 2023 and January 2024 are included. Each case includes the infant and mother, with partners invited. Quantitative data include infant outcomes from electronic medical records, neighborhood disadvantage using the Area Deprivation Index, and parent surveys of stress, anxiety, depression, bonding, perceived discrimination, and access to care. Qualitative data come from semistructured interviews guided by an adaptation of the National Institute on Minority Health and Health Disparities Research Framework and analyzed using reflexive thematic analysis. Data are integrated within each case to create structured summaries, followed by cross-case comparisons to identify patterns and variation.

Results: The study was funded by the National Institute of Nursing Research in July 2023. Participant enrollment occurred between May 2023 and January 2024, with 6 cases enrolled. Data collection and electronic medical record abstraction were completed in May 2024, with analysis completed by July 2024. This manuscript reports the protocol and analytic plan. Detailed results will be reported in 2 companion manuscripts currently under review; publication of one is anticipated in fall 2026.

Conclusions: This protocol outlines a novel mixed methods case study of neonatal transport and infant and parent outcomes among Black families. The study is exploratory and intended to generate hypotheses and inform future larger scale research and intervention development to support parental mental health, reduce disparities, and improve care experiences following preterm birth and neonatal transport.

International Registered Report Identifier (IRRID): RR1-10.2196/84626

(JMIR Res Protoc 2026;15:e84626) doi:10.2196/84626

KEYWORDS

neonatal transport; maternal mental health; very preterm infant; health disparities; neonatal intensive care unit

XSL•FO
RenderX

Introduction

Background

Despite medical advancements, global preterm birth rates have remained largely unchanged over the past decade [1]. As survival rates of preterm infants have improved [2], neonatal intensive care unit (NICU) admissions and length of stay have increased [3]. Preterm birth—defined as delivery before 37 weeks of completed gestation—remains a leading cause of neonatal mortality and lifelong morbidity worldwide [4]. Prematurity and low birth weight are the second-leading causes of infant mortality in the United States and major contributors to both short- and long-term morbidities, including respiratory, neurodevelopmental, and behavioral complications [4]. Annually, more than 70,000 infants are transported from community hospitals to NICUs to manage complications related to preterm birth [5].

While neonatal transport is essential for optimizing infant outcomes, it often creates additional challenges for families, especially those living far from the NICU. Parents of very preterm (VPT) infants—particularly those from low socioeconomic backgrounds—experience disproportionately high levels of psychological distress, including elevated levels of stress, anxiety, depression, and posttraumatic stress symptoms [6-9]. Recent evidence suggests that 30% to 50% of parents of preterm infants experience clinically significant mental health symptoms, which may persist for months to years after discharge [10,11].

Families with fewer socioeconomic resources or those living far from tertiary NICUs face additional barriers, including transportation costs, missed employment, childcare challenges for siblings, and limited access to their infant [12,13]. These stressors are associated with higher rates of depression and psychological distress among parents, particularly those from historically marginalized communities [7,14].

Elevated parental stress and depression have been linked to reduced participation in NICU care, altered parent-infant bonding, and adverse neurodevelopmental and behavioral outcomes that extend into early childhood [15]. Racial inequities in parental stress and mental health outcomes persist, with Black parents of preterm infants reporting higher levels of stress, depression, and perceived discrimination in health care settings, even after accounting for illness severity in their infant and socioeconomic status [16]. Structural inequities in access to high-quality perinatal care and mental health care may further compound these challenges. Despite this growing body of literature, no studies have examined how neonatal transport specifically contributes to parental psychosocial outcomes among Black families.

To contextualize these disparities, this study is informed by the National Institute on Minority Health and Health Disparities (NIMHD) Research Framework, which conceptualizes health outcomes as the result of interacting influences across multiple domains. These domains (biological, behavioral, sociocultural, and health care system factors) operate at the individual, interpersonal, community, and societal levels. Applying this framework allows for the examination of how multilevel determinants contribute to inequalities experienced during birth, transport, and NICU hospitalization.

Study Aims

Aim 1: Examining the Impact of Neonatal Transport on Infant Health Outcomes

Pathways resulting in poor infant outcomes are often complex and multifactorial. Community hospitals and staff vary greatly in their resources to stabilize and resuscitate VPT infants. Birth in community hospitals has been associated with increased morbidity such as hypothermia, severe intraventricular hemorrhage, chronic lung disease, and other poor outcomes resulting from physical stress and lifesaving procedures [17-19]. We aim to examine cases for conditions that may have an association with the stabilization trajectory at the birth hospital or during transport, which may negatively affect an infant's functioning or quality of life and pose challenges and costs to families. Studies have shown that VPT infants born in a community hospital and requiring transport have significantly higher odds of morbidities, cognitive impairment, and mortality risk [20-27]. Morbidities of interest known to be higher in the VPT-transported infant population were selected for data collection during NICU hospitalization. Any additional described morbidities were also noted to contextualize each case. Long-term health outcomes are less than optimal for this population, as multiple studies reveal that infants requiring transport have a higher chance of developmental impairment, cerebral palsy, and overall higher odds of death than those born in facilities with optimal capabilities [20-22,28].

Aim 2: Examining Parent Psychosocial Health, Social Vulnerability, and Socioeconomic or Racial Inequities

In the United States, Black women are more likely to deliver preterm infants than White women [29]. Mothers of VPT infants are more likely to live in areas with higher social vulnerability [30]. Further exacerbating these resource issues, VPT infants have considerably higher costs and use more health care resources across all health care service categories than their full-term peers during their first year of life [31]. Psychological morbidities in parents are common following childbirth, with depressive symptoms estimated to be 4.5% to 20% in mothers and 3% to 10% in fathers [32-34]. Research has shown that one partner's depressive symptoms can negatively affect the other partner's ability to bond with their child and that stress experienced by one parent often contributes to increased stress in the other [33,35]. Distance from the parents' home to the hospital NICU is another significant stress indicator [12], also affecting parent-infant bonding and further compounding stress and anxiety or leading to depression. Because structural racism in the health care system may also play an important role in racial disparities [36], we aim to investigate parents' perceptions of bias and racism in relation to neonatal transport and NICU hospitalization.

The purpose of this paper is to report a study protocol designed to accomplish the stated aims by examining the association between neonatal transport, the health outcomes of Black VPT infants, and the psychosocial well-being of their parents, with

XSL•FO
RenderX

particular attention to social vulnerability, and socioeconomic and racial inequities experienced during birth, transport, and NICU hospitalization. A deeper understanding of these relationships is crucial for improving access to individualized, high-quality care and for advancing health equity.

Methods

Study Setting

The study was conducted in South Carolina, where the preterm birth rate for Black women is 55% higher than that for other racial and ethnic groups, and the infant mortality rate exceeds the national average [1,37]. Recruitment took place at a tertiary level 3 NICU at a local health system in a medium-sized Southeastern US city. This health system includes a children's hospital and serves as the region's perinatal center, which covers 11 community hospitals and 16 counties, all of which are ranked moderate to high on the Centers for Disease Control and Prevention Social Vulnerability Index [38]. The most distant site in this region is 1.2 hours away by ambulance. Each case starts at the community hospital where the preterm infant is born. The perinatal center has a specialized neonatal and pediatric critical care transport team stationed onsite, which is responsible for all infant and pediatric patients returning to this center, resulting in more than 300 infant transports annually. Transport may occur by specialized critical care ambulance or by helicopter, depending on the distance and the illness severity of the infant, making the transport vehicle the next setting of the case. The final case setting is the only tertiary level III NICU in the region, with 70 beds and more than 800 infants delivered annually. The NICU manager and the neonatal medical director both provided letters of support for this study.

Study Design

This descriptive, exploratory study used a multiple-case study design and a mixed methods approach to describe the trajectory of Black preterm infants and their families from birth at a community hospital through transport to a referral center and subsequent NICU hospitalization. A case study design allows for a longitudinal understanding of complex and dynamic processes from multiple perspectives and contexts over time [37] and enables the description, explanation, or exploration of everyday phenomena in their natural settings [39]. Case study research is appropriate for descriptive and exploratory studies that investigate research questions that have not been studied in depth previously [40]. Multiple data collection methods provide a detailed investigation and analysis of the context and processes involved in the phenomenon under study. An adapted version of the NIMHD Research Framework [41] informed the research design and methods [42].

Ethical Considerations

This study was approved by the Prisma Health Midlands Institutional Review Board (1999826-1). All collected materials, including interviews, medical record information, and surveys, were deidentified and entered into a password-protected database stored on a secure server to protect participant privacy. Informed consent was obtained from all participants, who were informed that they could withdraw from the study at any time without consequence. Phone numbers and email addresses were collected at consent. Participants received financial compensation of US $25 for each set of surveys or interviews completed to acknowledge their time and reduce attrition.

Sample

The study was conducted between May 2023 and June 2024. Each case consisted of the anchor infant: a Black VPT infant born between 24 and 30 weeks' gestational age (GA), transferred to the NICU, and the infant's mother, with an invitation to the other parent to participate. Our aim was to enroll 10 infant cases using purposive convenience sampling with deliberate stratified sampling to ensure a range of GAs across infant cases. The sample size was set at 10 to achieve the purpose of the study and to allow for an in-depth examination. In case study methodology, sample size is not determined by statistical power but by the depth, richness, and purpose of the cases, with an emphasis on understanding complex phenomena within a real-world context [43]. Multiple-case study designs with 6 to 10 cases are recommended as sufficient to enable cross-case comparison while maintaining feasibility [44,45]. The intensive data collection also necessitates a limited number of cases to maintain rigor. Importantly, the purpose of this study is not to generalize findings but to generate grounded insights that can inform hypothesis development and intervention design to guide future larger-scale studies.

Study Procedures

This project was supported by the National Institutes of Health–National Institute of Nursing Research award (1F31NR020731-01A), awarded in July 2023. Institutional review board approval was obtained in advance from the hospital where the research was conducted. The NICU charge nurse notified the research team of any new admissions that met the inclusion criteria and provided contact information. Mothers provided informed consent at the infant's bedside, if available, or via phone, both for themselves and on behalf of their infants. Inclusion criteria consisted of Black mothers and their infants born between 24 and 30 weeks' gestation. The other parent was also invited to participate. "Parent" was defined as the other person who would be actively participating in the infant's care upon discharge. Eligible infants had to be delivered at an outside hospital and transported to the NICU, where recruitment took place. Infants with significant congenital anomalies and mothers who could not speak English were excluded. English-speaking participants were approached because the research team conducting the qualitative interviews was English speaking.

To achieve the objectives of aim 1, we used publicly available information to gather characteristics of the birth hospital. The first author (KFW) performed an electronic medical record (EMR) data review to provide a complete picture of the infant's transport trajectory. The transport team provided records from the birth hospital and an admission summary that informed the case on stabilization procedures that were given immediately after birth. A thorough review of EMR data from birth through 36 weeks' corrected GA was conducted for any morbid health outcomes that may have a critical impact on the health and well-being of each infant and their family.

XSL•FO
RenderX

To accomplish the goals of aim 2, parents completed a series of surveys (Table 1). An email was sent with a direct link to a set of surveys on Research Electronic Data Capture (REDCap), which is a Health Insurance Portability and Accountability Act (HIPAA)–compliant web-based application created to capture data for clinical research and to create databases and projects [46]. These surveys were previously published and validated instruments, apart from the infant access questionnaire, which was developed by the study team. Sets of surveys were administered 2 to 3 times (depending on the infant's length of stay) to measure how responses changed over the course of NICU admission and to understand experiences across multiple phases (refer to Figure 1 for example). The average time to complete an entire set of surveys was 23 minutes. Infant data collected from the EMR by the first author included infant GA, gender, and home address. This address was used to obtain the Area Deprivation Index (ADI) on the Neighborhood Atlas website (a scientifically validated measure that shows socioeconomic disadvantage by neighborhood) [47,48], which was added to REDCap. The ADI is composed of 17 measures related to education, employment, housing quality, and poverty, originally drawn from long-form census data and updated to incorporate more recent American Community Survey data [47,48]. Scores within a state are ranked from lowest to highest and then divided into deciles (1 to 10), with 10 representing the most disadvantaged neighborhood groups. Research shows that socioeconomic status can be a fundamental cause of health disparities [49].

XSL•FO
RenderX

Figure 1. Data collection timeline example. EMR: electronic medical record.

XSL•FO
RenderX

Table . Study outcomes and corresponding measures.

Expected outcomes	Measure	Description
Parental bonding	Parent-to-Infant Attachment Questionnaire [50]	It is a 25-item questionnaire using a 6-point Likert scale to assess the occurrence of disturbances in postpartum bond formation between mother and child.
Barriers to infant access	Team created a survey	It is an 8-item survey assessing descriptive factors (distance and cost of travel, transportation and accommodation, and visitation barriers or hardships).
Parental perceptions of bias and discrimination—health care experience	Discrimination in Medical Settings [51]	It is a 7-item instrument using a 5-point Likert scale to assess previous experiences of mistreatment in health care attributed to race, ancestry, or national origin.
Parental perceptions of bias and discrimination—comparative experience	Single question from Behavioral Risk Factor Surveillance System [52]	It assesses whether, in the past 12 months, health care experiences were worse than, the same as, or better than those of individuals of other races.
Parental perception of free choice to make informed health care choices	Adapted Medical Outcomes Study [53]	It is a measure that asks: (1) If there were a choice between treatments, how often would your child's doctor ask you to help make the decision? (2) How often does your child's doctor give you some control over your treatment? (3) How often does your child's doctor ask you to take some of the responsibility for your treatment?
Parental psychosocial morbidities—stress	Parental Stress Scale: Neonatal Intensive Care Unit [54]	It is a 26-item questionnaire assessed with a 5-point Likert scale to quantify parental stress in the particular situation of preterm birth.
Parental psychosocial morbidities—anxiety	Beck Anxiety Inventory [55]	It is a 21-item self-report questionnaire using 4-point scale to assess the intensity of physical and cognitive anxiety symptoms during the past week.
Parental psychosocial morbidities—depression	Edinburgh Postnatal Depression Scale [56]	It is a 10-item questionnaire using a 4-point Likert scale to assess depressive symptoms and coping related to pregnancy and childbirth
Socioeconomic status and disadvantage	Survey and Area Deprivation Index [57]	It is a multiple-choice 8-item survey assessing descriptive characteristics (age, race, gender, marital status, education, paid leave, occupation, annual household income, and insurance status) and Area Deprivation Index rank for social disadvantage.

To gain deeper insight into the psychosocial health of mothers—including their experiences during pregnancy, of having a preterm infant, and of navigating neonatal transport—the study team developed semistructured interview guides informed by key constructs from an integrated theoretical framework. This framework is an adapted version of the National Institutes of Health's NIMHD Research Framework, which is described in detail in a separate publication [42]. Probing questions were used to elicit deeper reflections on emotional support, staff interactions, perceived needs, and related experiences. Sample questions from the interview guide are presented in Table 2. To maximize participation, these interviews were conducted by phone or via Zoom teleconferencing.

Table . Sample interview questions.

Interview question	Additional probing example
Tell me about your pregnancy experience over the past __ months and what really stands out to you about this time in your life.	Who is the person or people you feel gave you emotional support during your pregnancy?
Tell me about your delivery experience with your baby at __ hospital.	Are there any parts of your labor and delivery experience at the hospital that stand out to you, either good or bad?
The transport team brought your baby to this hospital. Did someone talk with you before your baby was taken away in the ambulance?	Tell me any ways the process of taking your baby to another hospital might have been better for you.

XSL•FO
RenderX

Data Analysis

Overview

The case study approach allows for an in-depth, multifaceted examination of an underresearched phenomenon in a real-life setting. Case study methodology integrates well with mixed methods, which seeks a more complete understanding through the integration of qualitative and quantitative research [37]. We chose this approach to address both aims across the trajectory of the case. This multiple-case, longitudinal, exploratory study, with within-case and across-case analyses, used a personalized, concurrent mixed methods approach with repeated data collection points, as parental experiences are shaped by the unique trajectory of their infant's NICU experience and their own experiences within the NICU.

Quantitative Analysis

Three types of quantitative data were collected: (1) surveys administered to the parents at sequential time points for variables of interest, (2) the ADI to measure neighborhood disadvantage, and (3) infant morbidity outcomes from the EMR. All data, including survey data, were uploaded into REDCap as research capture tools. Because this study is exploratory and not powered for inference due to the small sample size, we used descriptive statistics (including means, SDs, and ranges) to analyze and summarize all data for each case. These findings were used to characterize each case, summarize patterns, and support cross-case comparisons. The primary purpose of this study was to generate hypotheses rather than test hypotheses. By describing infant clinical trajectories and parental psychosocial experiences, this study aims to identify relationships that may be examined in future studies with larger samples.

Qualitative Analysis

Research questions used for semistructured interviews were guided by an adapted model of the NIMHD Research Framework [58]. Parent interviews were recorded, typically lasting 20 to 40 minutes. Transcription was performed by the first author and was reviewed and compared to the recordings to ensure that they represent participants' utterances as accurately as possible. Reflexive thematic analysis, as outlined by Braun and Clarke [59], was used to explore the experiences of Black mothers navigating the emotional journey of giving birth to a VPT infant who required both neonatal transport and NICU hospitalization. This robust 6-phase framework includes (1) familiarization with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) writing the report [60]. The inductive approach was used, with no attempt to fit the data into a researcher's preconceptions [61]. Themes and subthemes were identified by the first author, RMD, and RBD.

After quantitative and qualitative data analyses, data were integrated to comprehensively summarize each enrolled case and reviewed for deviations from the interpretive summary by the first author and RBD. Each structured case summary included the infant's clinical course (with data spanning from birth at a community hospital until NICU discharge), longitudinal parental psychosocial measures, and themes derived from parent interviews. Quantitative data were summarized descriptively. Data were then applied to a visual timeline of the trajectory, comparing the infant's clinical status to the parent's psychosocial health to enhance interpretation. Cross-case analyses explored commonalities and divergences in participants' aggregate experiences. After case summarization, between-case comparisons were discussed by the entire research team. These results will be used to generate hypotheses for future studies.

Reflexivity

The first author was a novice qualitative researcher. She was a doctoral student studying in the field of nursing science, with an interest in preterm infant outcomes and maternal experiences. Her clinical background, spanning almost 2 decades, was in neonatal critical care, both in the NICU and as a NICU or pediatric ICU flight nurse. The familiar context provided benefits such as access and a working knowledge of both NICU and neonatal transport policies and procedures. Awareness of positionality was vital, as she identifies as a White researcher interviewing Black women. Particular attention was given to how racial and social position may have shaped the dynamics and data interpretation. Efforts were made to build trust through transparency about the study's purpose, active listening, and validation of participants' experiences, while recognizing that trust is not assumed but earned. Notes were taken throughout the analysis to document assumptions and emotional responses and to examine how the first author's positionality may have influenced theme interpretation and development. Discussions with coauthors were used to further interpret the data and enhance reflexive rigor.

To support methodological rigor, training in qualitative data collection and analysis was provided by the methodological expert (RMD) to the first author, as well as to members of the research team prior to study initiation and reinforced at intervals throughout the study. The intention was for author KFW to develop the knowledge and skills to be an independent researcher. Author RMD is a faculty member at the study university with expertise in qualitative methods and research focused on communication processes in vulnerable and underserved populations and their contributions to health disparities. Authors KFW and RMD each conducted independent reflexive thematic analysis of the interview data. These interpretations and themes were used to enhance depth rather than achieve consensus.

Results

This study examined the relationships between neonatal transport and both infant health outcomes and parental psychosocial well-being among Black families with VPT infants. Participant enrollment occurred between May 2023 and January 2024, during which time 6 cases were enrolled. During the enrollment period, neonatal transport admissions were reduced following the implementation of a high-risk maternal transport program, and all eligible cases that met the inclusion criteria were recruited. All enrolled cases were included in the analysis to inform the interpretation of findings.

XSL•FO
RenderX

Study activities, including EMR review and abstraction across NICU hospitalization, survey administration, and interviews, were completed in May 2024, with analysis completed by July 2024. Data collection was comprehensive and encompassed detailed clinical information, parental psychosocial measures, and qualitative data necessary to address the study aims. Detailed findings from this study will be reported separately in 2 companion manuscripts: one presenting overall study results with an illustrative case exemplar and another focusing on the qualitative results of parental psychosocial outcomes. Portions of this work have also been presented at regional, national, and international conferences.

Discussion

Important Implications for Practice

Health systems and social service networks have a significant opportunity to mitigate the psychosocial challenges experienced by NICU mothers in the postpartum period by providing avenues of support to help mothers cope with the transition process. Guidelines have been proposed to support the integration of psychologists into NICU settings, and a growing network of NICU psychologists has emerged [62]. There is growing evidence that interventions can improve parental mental well-being in the NICU; however, further research is needed to evaluate emerging technology-based and telehealth approaches [63,64].

Previous literature has highlighted the role of structural factors (such as limited access to postpartum mental health services, communication challenges within health care systems, and perceived discrimination) in shaping maternal experiences and contributing to inequitable outcomes [10,65]. Provider education addressing bias and increasing awareness about the barriers faced by Black mothers can reduce issues perpetuated within systems. Legislative and policy solutions to improve maternal access and quality of care and to sustain funding for community-based maternal health programs could improve future outcomes for Black women and infants.

Limitations

Eligibility for this study was limited to a small population of Black mothers who delivered a preterm infant between 24 and 30 weeks' GA at a community hospital and required transport to a NICU. Recruitment of mothers for this study was within a single perinatal center in the southern United States; therefore, the experiences of mothers in other regions may be different.

Conclusions

This multiple-case study design and mixed methods approach provides exploratory insight into the real-life experiences of Black mothers following preterm birth and neonatal transport. To our knowledge, this is the first study to examine the experiences of Black women who delivered a preterm infant in a community hospital that required neonatal transport, using a mixed methods case study lens. Through intensive, case-level data collection, the protocol outlines an approach for capturing maternal perspectives, which can highlight the challenges and systemic barriers present within our health care system across all phases, from pregnancy to several months postpartum. This protocol outlines a small foundational study designed to generate hypotheses and inform future larger-scale research. The findings generated from this study will be used to guide the development of subsequent research aimed at supporting mothers during this challenging and stressful period, tailoring interventions to reduce disparities, and improving maternal mental health outcomes.

Funding

This research project was funded by the National Institutes of Health–National Institute of Nursing Research (award number 1F31NR020731-01A) from July 2023 to August 2024.

Data Availability

The datasets generated or analyzed during this study cannot be publicly shared due to ethical or legal reasons but are available from the corresponding author on reasonable request.

Authors' Contributions

KFW and RBD contributed to the conceptualization, methodology, and writing (original draft). All the authors contributed to the validation, funding acquisition, supervision, and writing (review and editing) of the manuscript.

Conflicts of Interest

None declared.

Peer Review Report 1
National Institutes of Health–National Institute of Nursing Research F31 Summary Sheets
[PDF File, 148 KB - resprot_v15i1e84626_app1.pdf]

References

1. 2023 March of Dimes report card: the state of maternal and infant health for American families. March of Dimes. 2023. URL: https://www.marchofdimes.org/peristats/assets/s3/reports/documents/March-of-Dimes-2023-Full-Report-Card.pdf [accessed 2026-03-31]

XSL•FO
RenderX

2. Bell EF, Hintz SR, Hansen NI, et al. Mortality, in-hospital morbidity, care practices, and 2-year outcomes for extremely preterm infants in the US, 2013-2018. JAMA 2022 Jan 18;327(3):248-263. [doi: 10.1001/jama.2021.23580] [Medline: 35040888]
3. Harrison W, Goodman D. Epidemiologic trends in neonatal intensive care, 2007-2012. JAMA Pediatr 2015 Sep;169(9):855-862. [doi: 10.1001/jamapediatrics.2015.1305] [Medline: 26214387]
4. Ohuma EO, Moller AB, Bradley E, et al. National, regional, and global estimates of preterm birth in 2020, with trends from 2010: a systematic analysis. Lancet 2023 Oct;402(10409):1261-1271. [doi: 10.1016/S0140-6736(23)00878-4] [Medline: 37805217]
5. Gisondo C, Stanley K. Learning objectives for a neonatal-perinatal medicine transport curriculum based on a national needs assessment. Pediatrics 2020 Jul 1;146(1_MeetingAbstract):496-497 [FREE Full text] [doi: 10.1542/peds.146.1MA6.496]
6. Ganguly R, Patnaik L, Sahoo J, Pattanaik S, Sahu T. Assessment of stress among parents of neonates admitted in the neonatal intensive care unit of a tertiary care hospital in Eastern India. J Educ Health Promot 2020;9:288. [doi: 10.4103/jehp.jehp_169_20] [Medline: 33282993]
7. Sandnes R, Le Floch M, Riquin E, Nocus I, Müller JB, Bacro F. Parental stress and mental health outcomes following very preterm birth: a systematic review of recent findings. J Affect Disord 2024 Jun 15;355:513-525. [doi: 10.1016/j.jad.2024.03.154] [Medline: 38556094]
8. Hall EM, Shahidullah JD, Lassen SR. Development of postpartum depression interventions for mothers of premature infants: a call to target low-SES NICU families. J Perinatol 2020 Jan;40(1):1-9. [doi: 10.1038/s41372-019-0473-z] [Medline: 31439918]
9. Enlow E, Faherty LJ, Wallace-Keeshen S, Martin AE, Shea JA, Lorch SA. Perspectives of low socioeconomic status mothers of premature infants. Pediatrics 2017 Mar;139(3):e20162310. [doi: 10.1542/peds.2016-2310] [Medline: 28223372]
10. Hansen K, Gladdis T, Kukora S. Parental mental health & well-being in the NICU: addressing the Surgeon General's Advisory. Semin Perinatol 2025 Oct;49(6):152111. [doi: 10.1016/j.semperi.2025.152111] [Medline: 40603153]
11. Shetty AP, Halemani K, Issac A, et al. Prevalence of anxiety, depression, and stress among parents of neonates admitted to neonatal intensive care unit: a systematic review and meta-analysis. Clin Exp Pediatr 2024 Feb;67(2):104-115. [doi: 10.3345/cep.2023.00486] [Medline: 37986569]
12. Bourque SL, Williams VN, Scott J, Hwang SS. The role of distance from home to hospital on parental experience in the NICU: a qualitative study. Children (Basel) 2023 Sep 20;10(9):1576. [doi: 10.3390/children10091576] [Medline: 37761537]
13. Gourevitch RA, Ellicott E, Kim C, et al. Non-health care costs associated with neonatal intensive care unit visitation. Health Aff Sch 2025 Feb;3(3):qxaf043. [doi: 10.1093/haschl/qxaf043] [Medline: 40092401]
14. Adappa R, Barr S. Social determinants of health and the neonate in the neonatal intensive care. Paediatr Child Health 2023 Jun;33(6):154-157. [doi: 10.1016/j.paed.2023.03.002]
15. Cîmpian DM, Strete GE, Cîmpian CI, et al. The impact of preterm birth on parents' mental health and the role of family-centred interventions: a narrative review. Children (Basel) 2025 Sep 29;12(10):1311. [doi: 10.3390/children12101311] [Medline: 41153493]
16. Karvonen KL, Baer RJ, Blebu B, et al. Racial disparities in emergency mental healthcare utilization among birthing people with preterm infants. Am J Obstet Gynecol MFM 2022 Mar;4(2):100546. [doi: 10.1016/j.ajogmf.2021.100546] [Medline: 34871781]
17. Craig AK, Cutler A, Kerecman J, et al. Association of low hospital birth volume and adverse short-term outcomes for neonates treated with therapeutic hypothermia in rural states. J Perinatol 2025 Dec;45(12):1729-1734. [doi: 10.1038/s41372-025-02352-8] [Medline: 40617961]
18. Helenius K, Longford N, Lehtonen L, Modi N, Gale C, Neonatal Data Analysis Unit and the United Kingdom Neonatal Collaborative. Association of early postnatal transfer and birth outside a tertiary hospital with mortality and severe brain injury in extremely preterm infants: observational cohort study with propensity score matching. BMJ 2019 Oct 16;367:l5678. [doi: 10.1136/bmj.l5678] [Medline: 31619384]
19. Hogeveen M, Hooft L, Onland W. Hypothermia and adverse outcomes in very preterm infants: a systematic review. Pediatrics 2025 May 1;155(5):e2024069668. [doi: 10.1542/peds.2024-069668] [Medline: 40262762]
20. Sasaki Y, Ishikawa K, Yokoi A, et al. Short- and long-term outcomes of extremely preterm infants in Japan according to outborn/inborn birth status. Pediatr Crit Care Med 2019 Oct;20(10):963-969. [doi: 10.1097/PCC.0000000000002037] [Medline: 31232855]
21. Amer R, Moddemann D, Seshia M, et al. Neurodevelopmental outcomes of infants born at <29 weeks of gestation admitted to Canadian neonatal intensive care units based on location of birth. J Pediatr 2018 May;196:31-37.e1. [doi: 10.1016/j.jpeds.2017.11.038] [Medline: 29305231]
22. Fang JL, Mara KC, Weaver AL, Clark RH, Carey WA. Outcomes of outborn extremely preterm neonates admitted to a NICU with respiratory distress. Arch Dis Child Fetal Neonatal Ed 2020 Jan;105(1):33-40. [doi: 10.1136/archdischild-2018-316244] [Medline: 31079068]
23. Hohlagschwandtner M, Husslein P, Klebermass K, Weninger M, Nardi A, Langer M. Perinatal mortality and morbidity. Comparison between maternal transport, neonatal transport and inpatient antenatal treatment. Arch Gynecol Obstet 2001 Aug;265(3):113-118. [doi: 10.1007/s004040100197] [Medline: 11561737]

XSL•FO
RenderX

24. Chan K, Ohlsson A, Synnes A, et al. Survival, morbidity, and resource use of infants of 25 weeks' gestational age or less. Am J Obstet Gynecol 2001 Jul;185(1):220-226. [doi: 10.1067/mob.2001.115280] [Medline: 11483932]
25. Thomas K, Shah PS, Canning R, Harrison A, Lee SK, Dow KE. Retinopathy of prematurity: risk factors and variability in Canadian neonatal intensive care units. J Neonatal Perinatal Med 2015;8(3):207-214. [doi: 10.3233/NPM-15814128] [Medline: 26485554]
26. Good WV, Hardy RJ, Dobson V, et al. The incidence and course of retinopathy of prematurity: findings from the early treatment for retinopathy of prematurity study. Pediatrics 2005 Jul;116(1):15-23. [doi: 10.1542/peds.2004-1413] [Medline: 15995025]
27. Shipley L, Gyorkos T, Dorling J, Tata LJ, Szatkowski L, Sharkey D. Risk of severe intraventricular hemorrhage in the first week of life in preterm infants transported before 72 hours of age. Pediatr Crit Care Med 2019 Jul;20(7):638-644. [doi: 10.1097/PCC.0000000000001937] [Medline: 31013263]
28. Cho IY, Lee HM, Kim SY, Kim ES. Impact of outborn/inborn birth status of infants born at <29 weeks of gestation on neurodevelopmental impairment: a nationwide cohort study in Korea. Int J Environ Res Public Health 2022 Sep 16;19(18):11718. [doi: 10.3390/ijerph191811718] [Medline: 36141991]
29. Brown J, Chang X, Matson A, et al. Health disparities in preterm births. Front Public Health 2023;11:1275776. [doi: 10.3389/fpubh.2023.1275776] [Medline: 38162611]
30. Givens M, Teal EN, Patel V, Manuck TA. Preterm birth among pregnant women living in areas with high social vulnerability. Am J Obstet Gynecol MFM 2021 Sep;3(5):100414. [doi: 10.1016/j.ajogmf.2021.100414] [Medline: 34082172]
31. Sarda SP, Abogunrin S, Zhang Y, Sarri G. Economic burden of very preterm birth- a systematic literature review. Value Health 2017;20(5) [FREE Full text]
32. Massoudi P, Hwang CP, Wickberg B. How well does the Edinburgh Postnatal Depression Scale identify depression and anxiety in fathers? A validation study in a population based Swedish sample. J Affect Disord 2013 Jul;149(1-3):67-74. [doi: 10.1016/j.jad.2013.01.005] [Medline: 23499163]
33. Kerstis B, Aarts C, Tillman C, et al. Association between parental depressive symptoms and impaired bonding with the infant. Arch Womens Ment Health 2016 Feb;19(1):87-94. [doi: 10.1007/s00737-015-0522-3] [Medline: 25854998]
34. Kerstis B, Nohlert E, Öhrvik J, Widarsson M. Association between depressive symptoms and parental stress among mothers and fathers in early parenthood: a Swedish cohort study. Ups J Med Sci 2016;121(1):60-64. [doi: 10.3109/03009734.2016.1143540] [Medline: 26947219]
35. Faisal-Cury A, Tabb KM, Ziebold C, Matijasevich A. The impact of postpartum depression and bonding impairment on child development at 12 to 15 months after delivery. J Affect Disord Rep 2021 Apr;4:100125. [doi: 10.1016/j.jadr.2021.100125]
36. Churchwell K, Elkind MS, Benjamin RM, et al. Call to action: structural racism as a fundamental driver of health disparities: a presidential advisory from the American Heart Association. Circulation 2020 Dec 15;142(24):e454-e468. [doi: 10.1161/CIR.0000000000000936] [Medline: 33170755]
37. Guetterman TC, Fetters MD. Two methodological approaches to the integration of mixed methods and case study designs: a systematic review. Am Behav Sci 2018 Jun;62(7):900-918. [doi: 10.1177/0002764218772641]
38. Social Vulnerability Index. Agency for Toxic Substances and Disease Registry. 2024. URL: https://www.atsdr.cdc.gov/place-health/php/svi/index.html [accessed 2026-03-31]
39. Tight M. Understanding Case Study Research: Small-Scale Research with Meaning, 1st edition: SAGE Publications; 2017. [doi: 10.4135/9781473920118]
40. Stake RE. The Art of Case Study Research: SAGE Publications; 1995. URL: https://books.google.co.in/books/about/The_Art_of_Case_Study_Research.html?hl=es&id=ApGdBx76b9kC&redir_esc=y [accessed 2026-03-31]
41. Alvidrez J, Castille D, Laude-Sharp M, Rosario A, Tabor D. The National Institute on Minority Health and Health Disparities research framework. Am J Public Health 2019 Jan;109(S1):S16-S20. [doi: 10.2105/AJPH.2018.304883] [Medline: 30699025]
42. Warren KF, Dail RB, Dawson RM, Boghossian NS, Felder TM. When a baby is born, so is a parent: understanding the effects of preterm birth on Black parents through the lens of the NIMHD framework. Nurs Outlook 2024;72(5):102246. [doi: 10.1016/j.outlook.2024.102246] [Medline: 39116649]
43. Sandelowski M. Sample size in qualitative research. Res Nurs Health 1995 Apr;18(2):179-183. [doi: 10.1002/nur.4770180211] [Medline: 7899572]
44. Hennink M, Kaiser BN. Sample sizes for saturation in qualitative research: a systematic review of empirical tests. Soc Sci Med 2022 Jan;292:114523. [doi: 10.1016/j.socscimed.2021.114523] [Medline: 34785096]
45. Constantinou CS, Georgiou M, Perdikogianni M. A comparative method for themes saturation (CoMeTS) in qualitative interviews. Qual Res 2017;17(5):571-588. [doi: 10.1177/1468794116686650]
46. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
47. Kind AJ, Jencks S, Brock J, et al. Neighborhood socioeconomic disadvantage and 30-day rehospitalization: a retrospective cohort study. Ann Intern Med 2014 Dec 2;161(11):765-774. [doi: 10.7326/M13-2946] [Medline: 25437404]
48. Kind AJ, Buckingham WR. Making neighborhood-disadvantage metrics accessible - the neighborhood atlas. N Engl J Med 2018 Jun 28;378(26):2456-2458. [doi: 10.1056/NEJMp1802313] [Medline: 29949490]

XSL•FO
RenderX

49. Stormacq C, Van den Broucke S, Wosinski J. Does health literacy mediate the relationship between socioeconomic status and health disparities? Integrative review. Health Promot Int 2019 Oct 1;34(5):e1-e17. [doi: 10.1093/heapro/day062] [Medline: 30107564]
50. Brockington IF, Fraser C, Wilson D. The Postpartum Bonding Questionnaire: a validation. Arch Womens Ment Health 2006 Sep;9(5):233-242. [doi: 10.1007/s00737-006-0132-1] [Medline: 16673041]
51. Peek ME, Nunez-Smith M, Drum M, Lewis TT. Adapting the everyday discrimination scale to medical settings: reliability and validity testing in a sample of African American patients. Ethn Dis 2011;21(4):502-509. [Medline: 22428358]
52. Mokdad AH. The Behavioral Risk Factors Surveillance System: past, present, and future. Annu Rev Public Health 2009;30:43-54. [doi: 10.1146/annurev.publhealth.031308.100226] [Medline: 19705555]
53. Jalil YF, Villarroel GS, Silva AA, et al. Reliability and validity of the revised impact on family scale (RIOFS) in the hospital context. J Patient Rep Outcomes 2019 May 14;3(1):28. [doi: 10.1186/s41687-019-0118-1] [Medline: 31089824]
54. Miles MS, Funk SG, Carlson J. Parental Stressor Scale: neonatal intensive care unit. Nurs Res 1993;42(3):148-152. [Medline: 8506163]
55. Beck AT, Steer RA, Carbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin Psychol Rev 1988;8:77-100. [doi: 10.1016/0272-7358(88)90050-5]
56. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 1987 Jun;150:782-786. [doi: 10.1192/bjp.150.6.782] [Medline: 3651732]
57. Area Deprivation Index v20. University of Wisconsin School of Medicine and Public Health. 2015. URL: https://www.neighborhoodatlas.medicine.wisc.edu [accessed 2026-04-01]
58. NIMHD research framework. National Institute on Minority Health and Health Disparities. URL: https://www.nimhd.nih.gov/resources/nimhd-research-framework [accessed 2026-04-01]
59. Braun V, Clarke V. One size fits all? What counts as quality practice in (reflexive) thematic analysis? Qual Res Psychol 2021;18:328-352. [doi: 10.1080/14780887.2020.1769238]
60. Braun V, Clarke V. Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: a review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). Palliat Med 2024 Jun;38(6):608-616. [doi: 10.1177/02692163241234800] [Medline: 38469804]
61. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
62. Saxton SN, Dempsey AG, Willis T, et al. Essential knowledge and competencies for psychologists working in neonatal intensive care units. J Clin Psychol Med Settings 2020 Dec;27(4):830-841. [doi: 10.1007/s10880-019-09682-8] [Medline: 31776758]
63. Wagenaar J, Mah C, Bodell F, et al. Opportunities for telemedicine to improve parents' well-being during the neonatal care journey: scoping review. JMIR Pediatr Parent 2024 Dec 2;7:e60610. [doi: 10.2196/60610] [Medline: 39622079]
64. Feng YY, Korale-Liyanage S, Jarde A, McDonald SD. Psychological or educational eHealth interventions on depression, anxiety or stress following preterm birth: a systematic review. J Reprod Infant Psychol 2021 Apr;39(2):140-152. [doi: 10.1080/02646838.2020.1750576] [Medline: 32336122]
65. Lund C, Brooke-Sumner C, Baingana F, et al. Social determinants of mental disorders and the Sustainable Development Goals: a systematic review of reviews. Lancet Psychiatry 2018 Apr;5(4):357-369. [doi: 10.1016/S2215-0366(18)30060-9] [Medline: 29580610]

Abbreviations

ADI: Area Deprivation Index
EMR: electronic medical record
GA: gestational age
HIPAA: Health Insurance Portability and Accountability Act
NICU: neonatal intensive care unit
NIMHD: National Institute on Minority Health and Health Disparities
REDCap: Research Electronic Data Capture
VPT: very preterm

XSL•FO
RenderX

Edited by A Schwartz; submitted 25.Sep.2025; this is a non–peer-reviewed article;accepted 09.Mar.2026; published 10.Apr.2026.

Please cite as:
Warren KF, Felder T, Boghossian N, Dawson RM, Dail RB
Assessing Neonatal Transport, Health Outcomes, and Psychosocial Inequities Among Black, Very Preterm Infants and Their Families to Advance Health Equity: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e84626
URL: https://www.researchprotocols.org/2026/1/e84626
doi:10.2196/84626

© Karen F Warren, Tisha Felder, Nansi Boghossian, Robin M Dawson, Robin B Dail. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating the Combined Effects of an Adverse Childhood Experiences–Focused Family Advocate Model and the Strengthening Families Program: Study Protocol for a Hybrid Type 1 Effectiveness-Implementation Trial in 36 New Jersey Communities

Dallas Elgin[1*], PhD; Phillip W Graham[1*], DrPH, MPH; Kiersten Johnson[1], PhD; Sazid Khan[1], PhD; Michael Mills[1], MA; Feker Wondimagegnehu[1], BS; Brent Gibbons[2], PhD; Jason Williams[1], PhD; Xinyi Jiang[3], PhD; Andrea Zapcic[4], MPA; Brittany Simon[5], MA

[1]RTI International, 3040 East Cornwallis Road, Durham, NC, United States

[2]Independent Scholar, Rockville, MD, United States

[3]National Center for Injury Prevention and Control, Centers for Disease Control and Prevention, Atlanta, GA, United States

[4]New Jersey Prevention Network, Tinton Falls, NJ, United States

[5]RWJBarnabas Health Institute for Prevention and Recovery, Eatontown, NJ, United States

[*]these authors contributed equally

Corresponding Author:
Dallas Elgin, PhD
RTI International, 3040 East Cornwallis Road, Durham, NC, United States

Abstract

Background: Early exposure to adverse childhood experiences (ACEs), such as parental substance use, elevates the risk of future substance use and drug overdose, and in the absence of intervention, could perpetuate a cycle of substance-related ACEs across generations. Although research suggests that effectively decreasing the prevalence and impact of ACEs and substance use can benefit from addressing both family- and community-level factors in tandem, there is a critical gap in the evidence base pertaining to interventions that effectively integrate these 2 factors to prevent substance use and ACEs.

Objective: This study aims to conduct a rigorous evaluation of a novel intervention that integrates the established, evidence-based Strengthening Families Program with clinically trained, trauma-informed family advocates who will assist families in accessing resources related to ACEs and social determinants of health.

Methods: This study uses a hybrid type 1 effectiveness-implementation trial design, which is used to test a clinical intervention while gathering information on its delivery during the effectiveness trial and on the potential for implementation in real-world settings. Our hybrid type 1 design consists of 3 components. A cluster-randomized controlled trial will be used to test the effectiveness of the intervention on substance use, overdose, and ACEs in 36 New Jersey communities. A robust process evaluation informed by the Consolidated Framework for Implementation Research will be used to explore implementation barriers and facilitators. A cost evaluation will be conducted to accurately estimate the costs required to implement the intervention components and assess the cost-effectiveness of the intervention.

Results: Funding was awarded in 2022, and institutional review board approval was obtained in November 2023. Intervention and evaluation development lasted approximately 1 year, with study partners engaged in the co-design process to ensure alignment with the principles of community-placed behavioral health research. Study recruitment began in November 2023 and is anticipated to continue until September 2026. Final data collection is anticipated by March 2027, followed by data analysis.

Conclusions: This study evaluates a novel intervention that integrates New Jersey's established, evidence-based Strengthening Families Program with clinically trained, trauma-informed family advocates who will assist families in accessing community resources. By evaluating both family- and community-level outcomes through a hybrid type 1 design, the study bridges a key gap in the evidence base and provides a framework for combining multilevel interventions to reduce the long-term impacts of ACEs and substance use on youth and families. Future findings will provide critical insight related to intervention effectiveness, implementation barriers and facilitators, and the intervention's associated costs and cost-effectiveness.

Trial Registration: ClinicalTrials.gov NCT06236100; https://clinicaltrials.gov/study/NCT06236100

International Registered Report Identifier (IRRID): DERR1-10.2196/85151

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e85151) doi:10.2196/85151

KEYWORDS

adverse childhood experiences; substance use; social determinants of health; clinical trial; implementation science; cost-effectiveness analysis

Introduction

Adverse childhood experiences (ACEs) are preventable, potentially traumatic events that occur during childhood (0 - 17 y), such as experiencing neglect, experiencing or witnessing violence, or having a family member attempt or die by suicide [1,2]. ACEs are common among US adults—in a survey conducted from 2011 to 2020, approximately two-thirds of respondents reported having had at least 1 ACE, and approximately 1 in 6 reported 4 or more ACEs [3]. Toxic stress resulting from ACEs can negatively affect healthy brain development, leading to increased risk for adverse mental and physical health outcomes, violence victimization and perpetration, and reduced economic and life opportunities [4-7]. Furthermore, ACEs can play a role in substance use initiation among youth [8,9]. Existing research indicates that ACEs are associated with an earlier age of initiating opioid use and injection drug use, developing substance use disorders [10,11], and experiencing lifetime overdose, all of which have played a significant role in fueling the opioid epidemic [8,12]. Given the acceleration of overdose deaths in the United States since at least 1999 [13], efforts to prevent risk factors for substance use initiation, such as ACEs, are important.

To improve public health across the lifespan, the US Centers for Disease Control and Prevention (CDC) invested resources in ACE prevention, spanning the focus areas of surveillance, research, and programmatic efforts [14]. In addition, the CDC's National Center for Injury Prevention and Control (NCIPC) has synthesized the best available evidence on ACEs prevention into 6 comprehensive strategies to prevent ACEs and mitigate their long-term consequences if they do occur. One of these strategies is intervening to mitigate both immediate and long-term harms [2].

The Strengthening Families Program (SFP), a highly structured, evidence-based family skills training preventive intervention, has been recognized for its effectiveness in reducing known risk factors for ACEs and decreasing youth substance use, including opioid misuse [2,15-19]. The SFP7–17 curriculum is a new 11-session class curriculum from the developers of the original SFP [20]. Parents, youth, and children in SFP attend weekly 2-hour family group sessions for 11 weeks. A recent scoping review underscores the positive effects of SFP on substance use prevention and protective parenting factors in the United States [21]. Nevertheless, this scoping review highlighted a lack of available evidence regarding SFP adoption and maintenance at the setting level by the implementing organizations [21]. Indeed, SFP facilitators ("family coaches") often lack clinical training to make effective referrals, limiting their ability to identify and address substance use–related issues or make service linkages in the community. Meanwhile, family-level interventions—such as patient navigators, family advocates (FAs), and family navigation programs—have been shown to increase protective factors against ACEs and reduce disparities while improving reach via service linkages in the community [22-25]. Research suggests that effectively decreasing the prevalence and impact of ACEs and substance use can benefit from addressing both family- and community-level factors in tandem [26,27]. However, little is known about the evidence base pertaining to interventions that effectively integrate family- and community-level factors to prevent or mitigate substance use and ACEs.

To address this critical gap in the evidence base, Research Triangle Institute (RTI) International and its partners, the New Jersey Prevention Network and the RWJBarnabas Health Institute for Prevention and Recovery, will evaluate a novel intervention that combines family- and community-level factors in 36 New Jersey communities experiencing a disproportionate burden of substance use and ACEs. This intervention will integrate clinically trained, trauma-informed FAs into New Jersey's established evidence-based SFP. FAs will help families connect with community resources dedicated to the prevention and treatment of substance use and ACEs. This study will test whether integrating a combined family- and community-level intervention increases referrals and linkages to community resources, increases positive family functioning, and reduces youth and parent substance use and ACEs compared with families participating in SFP alone. The study will also test whether integrating a combined family- and community-level intervention produces system-level changes by aligning services and resources and subsequently decreasing the prevalence of substance use and ACEs at the community level. The purpose of this protocol paper is to provide an overview of an ongoing study by describing the study rationale, evaluation design, analysis methods, associated ethical requirements, and dissemination plan. Study findings will be presented in subsequent publications.

Methods

Research Objective and Aims

Our objective is to conduct a rigorous evaluation of a novel intervention approach that incorporates family- and community-level factors to mitigate the harms of exposure to ACEs and prevent future ACEs while aiming to prevent substance use and overdose. The study has three aims, that are, (1) use a cluster-randomized controlled trial (RCT) to test effectiveness on substance use, overdose, and ACEs of the SFP+FA intervention (treatment group) and SFP-only (control group), in 18 communities each; (2) conduct a robust process evaluation informed by the Consolidated Framework for Implementation Research (CFIR) [28] to explore implementation barriers and facilitators; and (3) conduct a cost evaluation to accurately estimate the costs required to start up and implement the SFP and FA intervention components and to assess the cost-effectiveness of the intervention.

XSL•FO
RenderX

Intervention Condition

Families residing in communities randomized to the intervention group will be connected to clinically trained, trauma-informed FAs who will assess families and refer them to community services. The intervention will provide wraparound supports to prevent ACEs and substance use and, critically, enable providers and community-based partners to align their services in a way that addresses the social determinants of health (SDOH) and other community-level factors that affect substance use and the relationship between social connection and ACEs. The FA component of the intervention will run concurrently with the SFP7–17 sessions, with the FAs interacting weekly with families over the 10- to 14-week intervention period. Each week, FAs will conduct 1-hour, post-session check-ins with each family. This 1-hour period will consist of a 20-minute telephone call with families to discuss their needs, with the remaining 40 minutes used to debrief, make service referrals, and complete documentation.

Intervention Development

The intervention development process was guided by the Problem, Objective, Design, (end-)Users, Co-creators, Evaluation, Scalability (PRODUCES) cocreation framework [29], operationalized in conjunction with the established principles of community-engaged research. PRODUCES offers a systems-oriented schema for structuring population health intervention research, emphasizing iterative coproduction of knowledge and reproducibility across all phases of inquiry—planning, implementation, evaluation, and dissemination. This framework delineates seven interdependent domains to be systematically addressed, such as (1) articulation of the public health problem or behavioral determinant targeted by the intervention; (2) specification of the overarching objectives and proximal aims of the participatory design process; (3) selection of participatory and design methodologies for cocreation; (4) identification of the intended end users and beneficiaries of the intervention; (5) characterization of the stakeholder groups engaged as codesigners; (6) determination of evaluation criteria, including fidelity, feasibility, and effectiveness metrics; and (7) specification of strategies for scalability and population-level translation.

Within this structure, we employed a co-design methodology to integrate the FA component into the evidence-based SFP. Given the existence of multiple SFP curricula with variable dosages, content emphases, and delivery modalities, we conducted final curriculum selection through an iterative, evidence-informed process that triangulated input from local implementers, assessments of organizational capacity, and feasibility analyses of delivery platforms (eg, virtual or in-person implementation). This selection process was anchored in both implementation science principles (eg, alignment with CFIR constructs) and pragmatic trial readiness (ie, feasibility of embedding the intervention in diverse real-world service systems).

The FA component itself was refined through a series of structured, iterative design sprints and stakeholder consultations with agency personnel, community representatives, and potential end users. These engagements used consensus-building techniques to ensure that the resulting intervention achieved cultural consonance with community norms and practices while maintaining fidelity to core evidence-based elements of SFP. The process also incorporated explicit attention to implementation outcomes (eg, acceptability, appropriateness, and sustainability) to strengthen downstream evaluability and translational potential. Taken together, this approach produced an intervention that is methodologically rigorous, contextually tailored, and positioned for scalable implementation, consistent with an emphasis on reproducibility, stakeholder relevance, and impact at the population level.

Control Condition

Families residing in communities randomized to the control group will participate in the SFP7–17 group class curriculum. Parents and children participate in SFP7–17, both separately and together, as the curriculum has lessons for parents, teens, and children, plus a joint family practice class. SFP7–17 meetings are 2 hours long and are typically held in person (but families can participate remotely during extenuating circumstances). Families complete 11 sessions over a 10- to 14-week period.

Evaluation Design

Overview

The study team will evaluate the intervention's effectiveness using a modified version of a hybrid type 1 design, which is used to test a clinical intervention while gathering information on its delivery during the effectiveness trial and on the potential for implementation in real-world settings [30]. Hybrid type 1 designs are ideally situated for rigorous evaluations where there is a strong evidence base for an intervention (such as SFP) and supporting evidence that the intervention would support applicability to new delivery methods (eg, FAs) and new settings and populations (eg, communities with high rates of overdose and ACEs). Our modified hybrid type 1 design integrates process and outcome evaluation components with a cost evaluation component to evaluate the intervention's impacts on family outcomes (eg, substance use prevalence, risk or harm perceptions, and ACEs) over a 3-year period.

Participants in each of the 36 communities (18 treatment and 18 control communities) will complete the study-related tasks for the effectiveness trial, process, and cost evaluation components over approximately 9 - 11 months. Figure 1 illustrates how each evaluation component fits within the study and the associated time frames. The process evaluation's pre- and posttrial study activities will occur 1 - 3 weeks before and after the effectiveness trial, respectively. Completion of the SFP7–17 curriculum over 10 - 14 weeks will be the primary end point for the effectiveness trial, and completion of 6-month postintervention surveys will be the secondary end point. The cost evaluation study activities will be conducted throughout the pretrial, effectiveness trial, and posttrial phases and reported during the posttrial component. The effectiveness trial, process, and cost evaluation components are discussed in greater detail in subsequent sections, followed by the study's data analysis plan.

XSL•FO
RenderX

Figure 1. Hybrid type 1 effectiveness-implementation design. ACE: adverse childhood experience; FA: family advocate; SFP: Strengthening Families Program; SFP+FA: Strengthening Families Program and family advocate (ie, treatment condition).

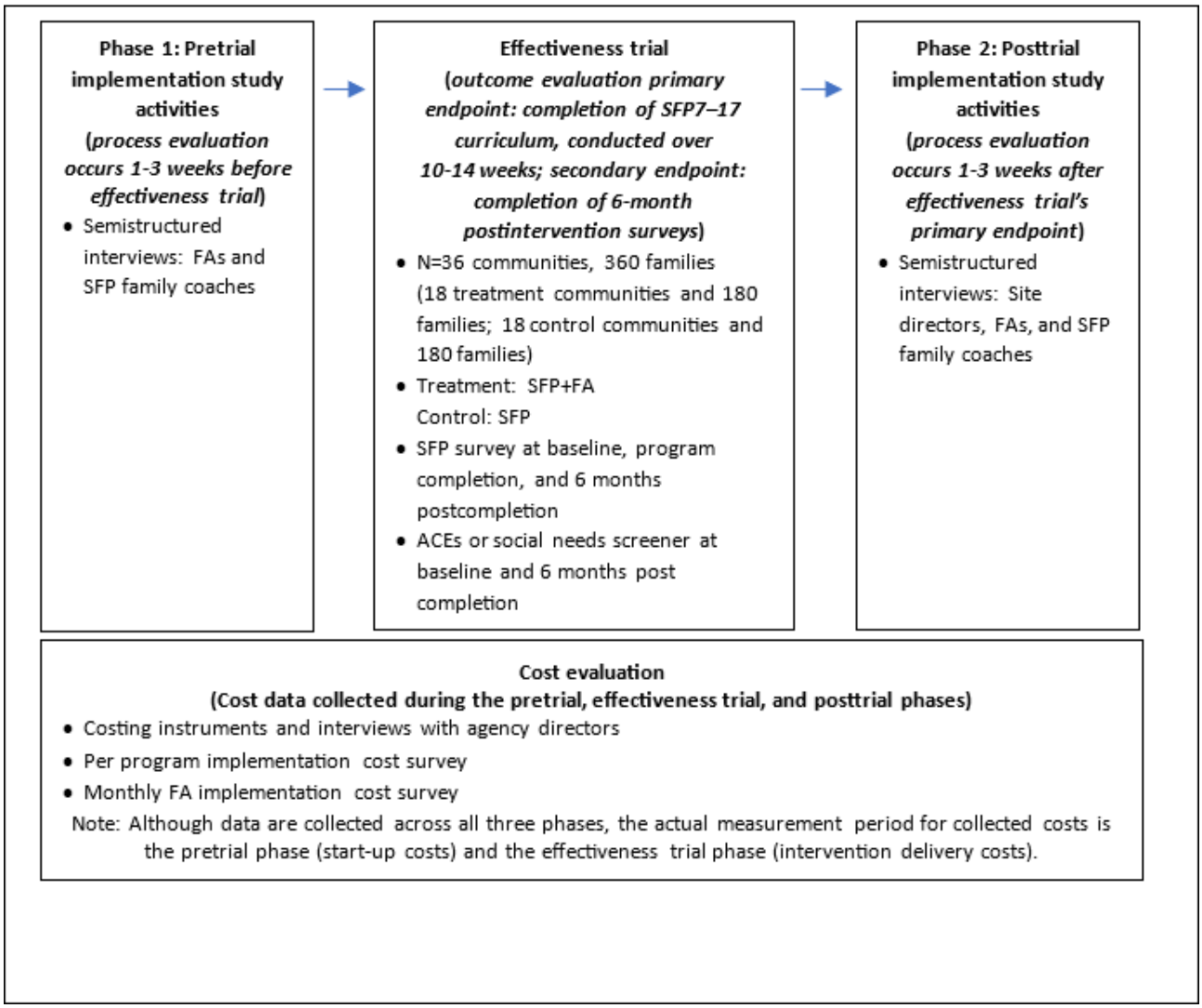

Effectiveness Trial

The effectiveness trial uses a cluster RCT design to randomize 36 New Jersey communities experiencing a disproportionate burden of substance use and ACEs [31] to treatment and control conditions. Cluster RCTs offer a rigorous design for evaluating community-level interventions and provide key benefits in the form of mitigating spillover effects, reducing attrition, and increasing compliance from program staff and participants [32].

Community Selection

All families participating in the study will come from preidentified communities across the state of New Jersey. A "community" is akin to a classified municipality in the state. New Jersey has 564 municipalities, comprising 5 location types—252 boroughs, 52 cities, 15 towns, 241 townships, and 4 villages. Communities with population sizes less than 10,000 per US Census estimates were removed from eligibility on the front end, leaving approximately 200 communities for the selection process. Eligible communities were then ranked using data comprising ACEs (collectively, these secondary data sources provided data on the prevalence of 8 of the 10 original ACEs [24], that is, physical abuse, emotional abuse, sexual abuse, physical neglect, emotional neglect, substance misuse, mental illness, and incarcerated relatives) and opioid-related vulnerability metrics, such as overdose-focused SDOH, opioid overdose–involved hospitalizations, and child maltreatment case rates [33-35].

These metrics were derived from the following secondary data sources:

- The RTI Rarity project uses supervised machine learning, including random forests and other state-of-the-art predictive methods, to create local social inequity scores at the census-tract level that draw on SDOH and other contextual, neighborhood-level measures. This study used RTI Rarity's Local Index in drug overdose model, which provides percentile-ranked scores for each US Census tract, with higher scores indicating higher predicted drug overdose mortality rates. The model comprises 40 measures and explains 99% of the variance in drug overdose mortality rates across census tracts [36].
- Healthcare Cost and Utilization Project (HCUP) databases bring together the data collection efforts of state data

XSL•FO
RenderX

organizations, hospital associations, private data organizations, and the federal government (HCUP partners) to create a national information resource of encounter-level health care data. HCUP includes the largest collection of longitudinal hospital care data in the United States, with all-payer, encounter-level information beginning in 1988.
- The National Child Abuse and Neglect Data System is a federally sponsored effort that annually collects and analyzes data on child abuse and neglect known to child protective services agencies in the United States.

Communities were subsequently assigned into quadrants based on average rankings across the 3 metrics to prioritize disproportionately affected communities. County providers with eligible communities were approached to ask about their interest in and ability to participate in the first year of the study. Vulnerable communities in counties that responded positively were then added to the list of communities to be randomized. The team also noted counties that would be able to participate in subsequent years.

Randomization

For the purpose of random assignment, during the 3 years of the effectiveness trial, the study team plans to enroll 12 communities from 6 counties from the initial list of counties. Each county will contribute 1 community implementing SFP-only (ie, control condition) and 1 community implementing SFP+FA (ie, treatment condition).

Counties with multiple communities will be preferred to facilitate randomization at the county level, thus helping to balance the geographical and background characteristics of the participating communities. As a contingency plan, if a county has only 1 community (or more than 1 community but with limited capacity), the study team will attempt to find matching communities in other counties. Simultaneous implementation of SFP and SFP+FA within selected counties will be preferred but is not required for successful implementation; asynchronous starts may be used in counties, if necessary.

In total, 3 general randomization approaches were identified. Feasibility and preference for a method varied by implementation and community readiness factors, as well as methodological concerns, such as control of potential confounds. The three potential approaches were pairwise assignment, pool randomization, and rolling assignment. First, under pairwise assignment, (1) communities within each county are ranked sequentially according to their ACEs, substance use disorder, and SDOH risk profile or rank; (2) sequential pairs are then randomized to SFP and SFP+FA conditions; (3) assignment is repeated for each pair; and (4) recruitment is done at the pair level. Second, under pool randomization, (1) within each county, each community is randomly assigned to 1 of 2 or more blocks or groups of communities. Block size will depend on the total number of communities available in the county. Multiple patterns of assignment of communities within each block are generated. Assignment is determined by a randomly selected pattern for each block. (2) Recruitment is then done in each county for a single SFP-only and single SFP+FA community for entry into the study. Finally, under rolling assignment, (1) assignments are based on enrollment flow and not a priori assignment; (2) recruitment takes place before assignment; and (3) the first community enrolled is randomly assigned to either SFP or SFP+FA, and the next enrolled group is automatically placed in the group not assigned to the first community.

On the basis of conversations with the study team's New Jersey partners and information gathered from local partners, randomization was conducted in a process that combined the pairwise and pool randomization approaches. First, 2 communities were chosen at random from each county. Next, multiple blocks were generated for each county. Then, 10 blocks were generated for each county; in each block, 1 community was randomly assigned to SFP and the other was assigned to SFP+FA. The block containing the community assignment to be used was then chosen at random. The study team will repeat this process in each successive year of the effectiveness trial, for 3 rounds of randomization involving 36 communities.

Sample

The study team conducted power analyses to determine the minimum number of families needed to detect statistically significant effects in the effectiveness trial. Calculations were based on the use of a 3-time-point latent growth model (LGM) with linear slope conditional on treatment group (ie, a difference-in-differences model). Power analyses used a Type I error rate of 0.05 and an intraclass correlation of 0.01 to 0.05. The estimated minimal detectable effect size (based on power=0.80) for 36 communities with an assumed 10 families per site (for a total of 360 families) is 0.40 or below.

Eligibility Criteria

To be eligible to participate in this study, a family must meet all of the following criteria:

- Either reside or attend SFP meetings in 1 of the 36 New Jersey communities assigned to treatment or control conditions
- Consist of 1 or more adult caregivers and 1 or more children aged 7 - 17 years
- Provide informed consent (caregivers) or assent (youth) to participate in the study
- Express a willingness to adhere to the regimens of the SFP and FA interventions
- Have access to necessary resources (ie, computer, smartphone, and internet access) for completing data collection instruments.

Objectives and End Points

Table 1 describes the study objectives and end points for the effectiveness trial. The trial will use a combination of assessments at baseline, completion, and 6-month follow-up to examine youth and parent or caregiver substance use and ACEs among families in the randomized communities. The ACEs or social needs screener will be used for both screening and data collection purposes. To mitigate potential ethical and privacy concerns, the screener does not gather separate information on types of ACEs experienced; rather, respondents are asked to indicate the number of relevant ACEs in each of 3 categories—abuse, neglect, and household challenges. Participating families will complete a baseline screener during

XSL•FO
RenderX

SFP session 3, and families in the treatment group may elect to share their data with the FAs, who will assess the data and refer families to relevant services in the community. Families in the control group will be subject to a treatment-as-usual condition. The ACEs or social needs screener will be readministered 6 months post intervention to measure the intervention's impact on the prevalence of ACEs and social needs among participating families. The study team will use weekly family referral summaries to collect outcome data to measure the intervention's impact on clinical and nonclinical referrals.

Table . Effectiveness trial: study objectives, end points, and justifications.

Objectives	End points	Justification for end points
To use a cluster-randomized controlled trial to test the comparative effectiveness of the intervention and control condition on substance use, overdose, and ACEs[a] in 18 communities each	Primary end points • SFP[b] posttest survey (completed by families at SFP graduation) • Weekly family referral summaries (completed by SFP Family Coaches and FAs[c] throughout intervention until SFP[b] graduation) • Community service provider follow-up survey (completed by SFP family coaches, FAs, and community service providers after SFP graduation) Secondary end points • ACEs or social needs screener (completed by families at session 3 of SFP and 6 months after SFP graduation) • SFP 6-month follow-up survey (completed by families 6 months after SFP graduation)	Primary end points • The SFP posttest survey will collect key outcome data for measuring the intervention's impact on substance use, perceptions of harm, and risk • The weekly family referral summaries will collect key outcome data for measuring the intervention's impact on referrals to substance use disorder services and nonclinical community service providers • The community service provider follow-up survey will collect key outcome data for measuring the intervention's long-term impact on community-level change via change in system linkages Secondary end points • The ACEs or social needs screener will collect key outcome data for measuring the intervention's impact on the prevalence of ACEs and social needs at baseline and 6 months postintervention among participating families • The SFP 6-month follow-up survey will collect key outcome data for measuring the intervention's long-term impact on substance use, perceptions of harm, and risk

[a]ACE: adverse childhood experience.

[b]SFP: Strengthening Families Program.

[c]FA: family advocate.

To assess community-level impact, the trial will use a survey of community service providers to collect data on system-level changes in the referral networks among FAs, SFP facilitators, and community service providers in the 36 participating communities. The survey will ask respondents about the characteristics of their respective organizations (eg, number of clients served and staff size) and services offered to individuals. Via a series of network questions, respondents will be asked to list up to 10 of the main organizations that their organization refers to or collaborates with. Finally, respondents will be asked whether their organization refers to or collaborates with specific community service providers. Responses to these surveys will provide a matrix of interactions among FAs, SFP facilitators, and community service providers that will help construct egocentric clinical and nonclinical referral networks within each of the 36 communities for the periods before and after intervention implementation. The analysis methods for the effectiveness trial and other evaluation components are discussed in the Data Analysis section of this paper.

Process Evaluation

The process evaluation will consist of pretrial implementation data collection, fidelity monitoring during intervention implementation, and posttrial implementation data collection. Pre- and posttrial process evaluation data collection activities were informed by the CFIR to allow the study team to explore implementation barriers and facilitators by identifying strategies that facilitate the integration of the FA role into SFP. The process evaluation aims to help clarify (1) the barriers and facilitators that families in participating communities may encounter in accessing services; (2) the challenges that participating agencies face in implementing SFP and integrating FA into SFP, as well as any solutions; and (3) the ways in which the implementation of SFP+FA affects family access to services and cross-system coordination and communication among agencies. Table 2 describes the study objectives and end points for the process evaluation.

XSL•FO
RenderX

Table . Process evaluation: study objectives, end points, and justifications.

Objectives	Primary end points	Justification for primary end points
To conduct a robust process evaluation informed by the Consolidated Framework for Implementation Research in order to explore implementation barriers and facilitators	• Posttrial interviews (completed with SFP[a] and FA[b] delivery staff after SFP[a] graduation) • Posttrial program lead interviews (completed with prevention agency program leads after SFP graduation)	Both types of posttrial interviews will collect key process evaluation data related to • observed challenges and associated solutions that participating agencies encountered when integrating the FA component into SFP • observed barriers and facilitators that families encountered when accessing services • ways in which the implementation of the intervention affected family access to services, cross-system coordination, and communication

[a]SFP: Strengthening Families Program.

[b]FA: family advocate.

During the pretrial implementation phase, the study team will conduct 2 series of qualitative interviews. The first series will interview designated SFP facilitators from each participating community. These semistructured interviews will gather qualitative data about SFP facilitators' past experiences implementing the SFP, their understanding of the SFP facilitator's role in the program, anticipated challenges of implementing the program, and anticipated barriers to and facilitators of accessing services that may be encountered by families in their communities. The second series of interviews will be conducted with FAs, with each FA participating in 1 pretrial interview. These semistructured interviews will gather qualitative data regarding FAs' understanding of their role and any anticipated challenges and facilitators pertaining to their role.

During intervention implementation, the study team will monitor implementation fidelity through weekly fidelity reporting. Each site will have 1 designated SFP facilitator and parents or caregivers who will respond to brief fidelity questionnaires after each SFP session. SFP facilitators will report session attendance, adherence to the curricula, and participant engagement, whereas parents and caregivers will report on their personal satisfaction with the session and their perceptions of SFP facilitators' knowledge, empathy, and preparedness.

During the posttrial process evaluation, the study team will conduct 3 series of qualitative interviews. The first series of semistructured interviews will gather qualitative data about facilitators' experiences implementing SFP and about barriers to and facilitators of accessing services encountered by families in their communities. These interviews will also ask SFP facilitators from intervention communities about their experiences with the addition of the FAs. The second series will interview FAs and will gather qualitative data regarding FAs' experiences in their role, including any challenges and facilitators they encountered. The third and final series will interview prevention agency program leads who oversee SFP groups to gather qualitative data regarding each agency's experience implementing SFP and regarding barriers to and facilitators of accessing services encountered by families in their communities. The study team will also ask program leads from intervention communities about their experiences with the integration of the FA role.

Cost Evaluation

As illustrated in Table 3, the cost evaluation will use data collected from prevention agency program leads, SFP facilitators, and FAs to identify the activities and costs associated with starting up and implementing the intervention in treatment and control communities. The data collected from prevention agency program leads will include cost data on staff salaries, overhead costs, training materials and fees, and other SFP implementation resources to inform accurate cost estimates for prespecified activities (eg, management); these data will be supplemented by survey data collected from SFP facilitators and FAs. Given that SFP facilitators and FAs are the main staff delivering the intervention, their survey data will help to understand how their time is spent in various activities, including travel time and distance for SFP sessions, time spent in various activities related to implementing SFP, and time spent in specific FA activities (eg, engaging with families). The study team will use the activity-based data collected from SFP facilitators and FAs, combined with the cost information from the prevention agency program leads, to calculate the costs of the treatment and control conditions and to examine potential sources of variation in costs, such as activity cost differences or staffing mix differences. Start-up costs for the intervention will also be estimated from preimplementation activity data, primarily from the cost instrument completed by prevention agency program leads.

XSL•FO
RenderX

Table . Cost evaluation: study objectives, end points, and justifications.

Objectives	Primary end points	Justification for primary end points
• To conduct a cost evaluation to accurately estimate the costs required to implement SFP[a]+FA[b] intervention components and assess the comparative cost-effectiveness of the treatment intervention and control • To estimate the start-up costs associated with the intervention	• Costing instrument and interview (completed by prevention agency program leads after SFP graduation) • Per-program implementation cost survey (completed by SFP[a] Family Coaches after SFP graduation) • Monthly implementation cost survey (completed by FAs monthly)	• The costing instrument and interview, per program implementation cost survey, and monthly implementation cost survey will collect key cost evaluation data to inform start-up cost estimations, implementation cost estimations, and cost-effectiveness analyses

[a]SFP: Strengthening Families Program.

[b]FA: family advocate.

Data Analysis

Figure 2 details the mixed methods analysis approach that will be used to analyze data from the evaluation's process, effectiveness trial, and cost components. In the sections below, we provide an overview of the methods used for each component, followed by the approach used to systematically integrate and analyze the findings across components.

Figure 2. Overview of the mixed methods analysis for the study's hybrid type 1 design. ACE: adverse childhood experience; FA: family advocate; SFP: Strengthening Families Program.

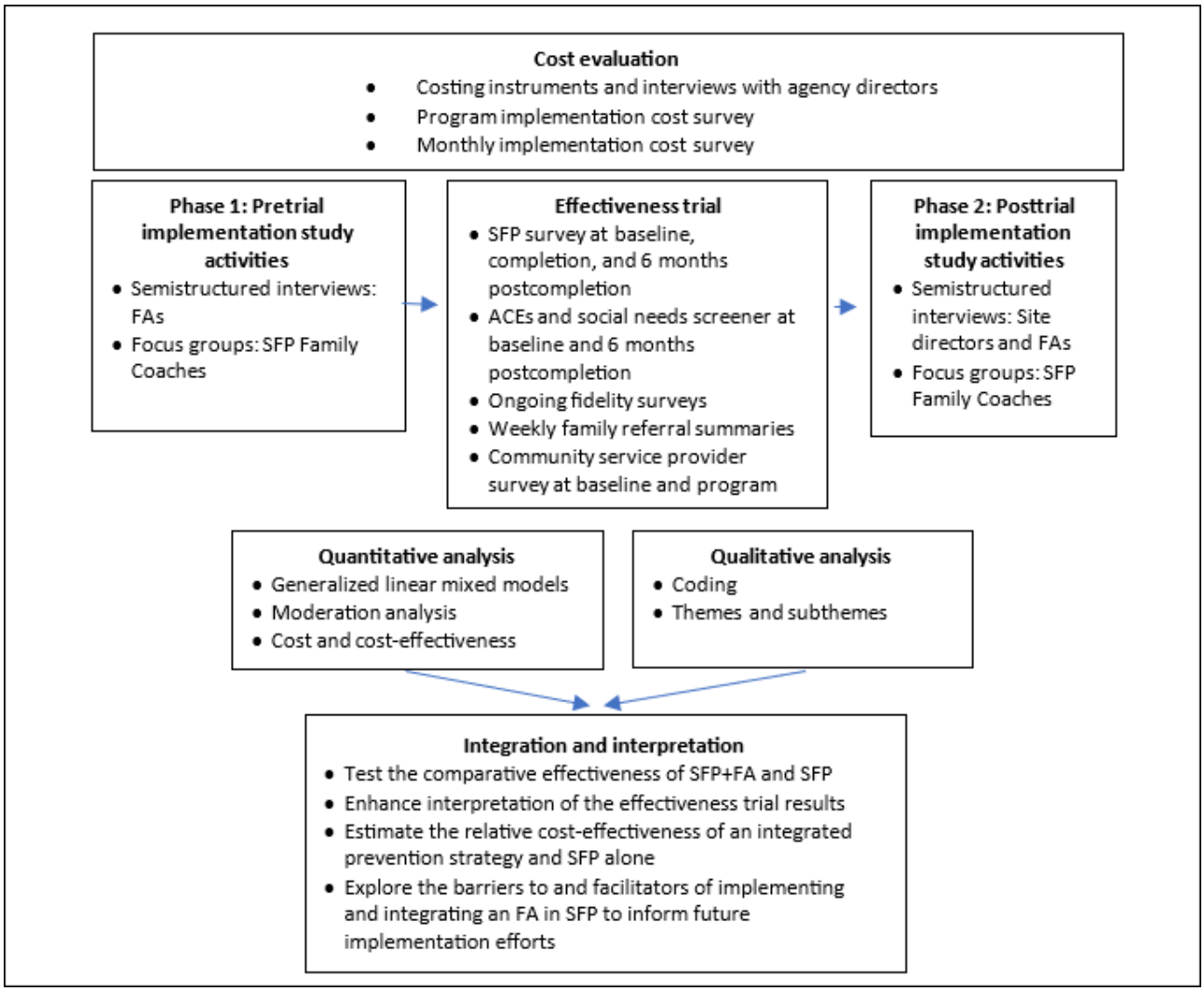

Quantitative Analysis of Effectiveness Trial Data

Overview

All inferential statistics will use a Type I error rate of 0.05 and 2-tailed tests. The study team will apply a multiple comparison adjustment to mitigate the likelihood of Type 1 errors across the analyses of effectiveness trial data. The Benjamini-Hochberg procedure (Benjamini and Hochberg [37]) will be applied separately for each set of outcomes within the primary and

XSL•FO
RenderX

secondary end points specified in Table 1 to adjust *P* values and account for multiple hypothesis testing.

Descriptive Statistics

Preliminary analyses will be used to characterize and explore the data. Means, frequencies, and correlations will describe the distributions of demographics, key covariates, and the outcome variables. Simple bivariate regression models will be used to examine the groups for baseline differences. Variables will be assessed for outliers and other distributional violations.

Program Effects

The difference in family outcomes between exposure to SFP+FA and SFP-only will be examined with generalized linear mixed models (GLMMs)—that is, hierarchical linear models or multilevel regression models. These models can accommodate the repeated assessments at baseline, immediate post treatment, and 6-month follow-up. GLMM can simultaneously account for nesting of families in communities to account for any nonheterogeneity attributable to location. The evaluation model will consist of a 3-time-point LGM estimated within the GLMM framework. This model examines how family outcomes (eg, substance use prevalence, risk or harm perceptions, and ACEs) change over time and incorporates random effects (variance components) for each family's initial status (intercept) and rate of change or slope [38]. The LGM model will test how the SFP and SFP+FA groups change from baseline through the 6-month follow-up and the difference in those 2 rates of change. The difference in rates of change (sometimes called a difference-in-differences model) [39,40] will be the primary evaluation outcome.

Figure 3 illustrates a difference-in-difference model in which panel A shows no intervention effect (the difference A2−A1 is the same as the difference B2−B1), and panel B demonstrates an intervention effect (A2−A1 is not equal to B2−B1).

Figure 3. Illustration of difference-in-differences evaluation model. (A) No association between exposure and measured outcome, and (B) association of exposure and measured outcome.

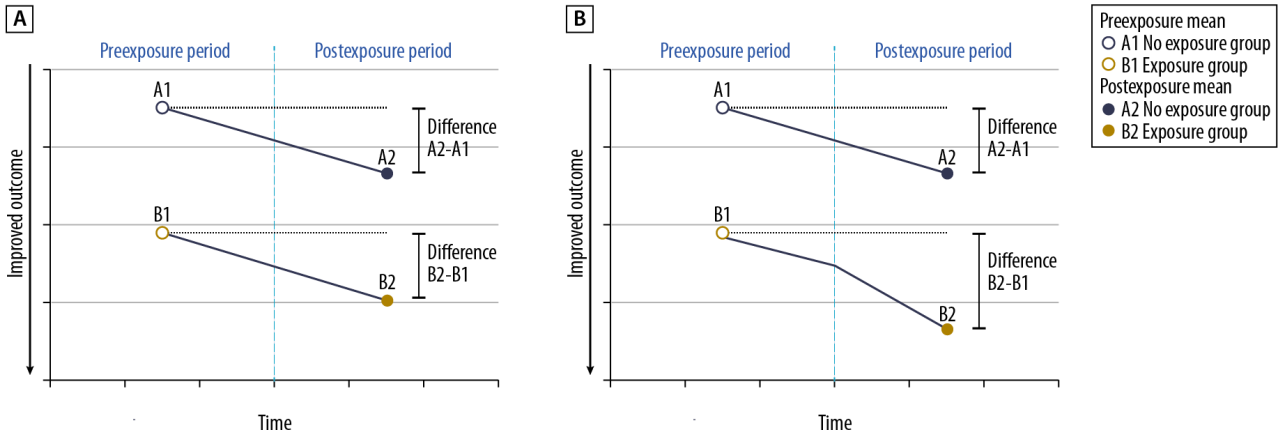

The LGM approach provides more options than ANOVA or other traditional mean comparison techniques by being able to incorporate covariates into an analysis. Covariates may be at any level included in the GLMM, such as the individual, family, or community levels. All outcomes assessed will be examined with these models and will begin with random effects at the family and community levels for both the intercept and slope, and the covariance between intercepts and slopes at each level will be included. Models may be iteratively adjusted to remove random effects and covariances as necessary, because it is common for not all outcomes of interest to require all sources of variability.

Missing Data Strategies

The impact of missing data will be minimized through the use of likelihood-based estimators, which are available for multilevel models of the proposed difference-in-differences model for the evaluation. These likelihood methods yield unbiased estimates and maximally efficient standard errors without sacrificing cases (thus maximizing statistical power). The proposed method of accommodating missing data is fully appropriate when data are missing truly at random (missing completely at random [MCAR]) or predicted by other variables (such as baseline scores and covariates) in a given model but independent of the potential values of the outcome itself (ie, missing at random [MAR]). The MAR and MCAR mechanisms are termed ignorable missing data because they can be rendered unproblematic for analysis with the proper estimation technique and model (eg, including the variable that predicts missingness as a predictor of the outcome). Although there is no general test of the MAR assumption, MAR-appropriate estimators are typically maximally efficient and unbiased even when the data are not truly MAR or MCAR (ie, the missing data are nonignorable). However, if the pattern of missing data is suspected to be nonignorable, then the research team will incorporate auxiliary variables in the analysis model or use pattern-mixture models that are adept at obtaining estimates for longitudinal studies with attrition caused by levels of outcomes.

Network Analysis of Effectiveness Trial Data

The findings from the community service provider survey will be used to examine the intervention's effectiveness in achieving community-level change via change in system linkages within the referral networks between FAs, SFP family coaches, and community service providers in each of the 36 communities participating in the effectiveness trial. Surveys will be administered pre- and postimplementation of the intervention in each of the 36 communities, with individuals in the 3 respondent categories asked a series of questions about the organizations they refer to or collaborate with for any ACEs,

XSL•FO
RenderX

substance use, or mental health services. Respondents (referred to as egos in the community service provider surveys) will be asked to list up to 10 of the main organizations that they refer or collaborate with and then to respond to a second set of questions asking about referrals and collaborations with selected community service providers. Answers to these questions will provide a matrix of interactions among FAs, SFP family coaches, and community service providers that will facilitate the construction of egocentric networks within each of the 36 communities for the periods before and after implementation of the intervention.

This study will use a multistep process to analyze the networks for the treatment and control communities. In the first step, network diagrams will be constructed using the R software package (R Core Team). Descriptive network analyses will be conducted on each of the networks to examine key network measures for the communities in the treatment and control conditions. These analyses will move from the simple to more complex network measures:

- Network size: The number of nodes (ie, FAs, SFP family coaches, and community service providers) in a given network.
- Edges: Connections (directed and undirected) among nodes within a network.
- Density: How close a network is to being fully connected, ranging from no connections at all to all possible connections among nodes being made.
- Centrality: The degree to which a network is organized around a central node by comparing the proportion of edges to a central node to all other edges in a network.
- Reciprocity: The fraction of all possible connections in which nodes are mutually connected.
- Triadic closure: The addition of an edge that closes a 2-path (ie, 3 nodes connected by 2 edges) to form a triangle (or among 3 nodes, A, B, and C, connections A–B and A–C exist, and a new connection B–C is formed).
- Assortativity: The extent to which nodes in a network are associated with other nodes in the network, being of similar sort or being of opposing sort.

In the next step of the analysis, mixed effects models will be estimated to examine the intervention's effectiveness in achieving community-level change via overall and configurational changes in the referrals and linkages between FAs, SFP family coaches, and community service providers in treatment and control communities. The random intercept and random coefficient mixed effects models will incorporate a 2-level structure with repeated measurements (ie, pre- and postimplementation) nested within communities [41]. The model will incorporate random slopes for communities and random intercepts for time. The following 2-level model will estimate the impacts of the intervention on referrals and linkages between FAs, SFP family coaches, and community service providers:

$$yij = \beta 0 + \mu 0j + \beta 1Treat1\ ij + \beta 2X2\ ij + \ldots + \beta pXp\ ij + \mu 1X1ij + \varepsilon ij$$

where yij refers to the outcome for time point i in community j, β0 is the overall mean of the outcome across all groups, μ0jrepresents the random intercept, β1Treat1ij is an indicator equal to 1 for communities assigned to the treatment group and zero for those assigned to the control group, β2X2 ij+…+βpXp ij represents a vector of covariates, μ1X1ij represents the random coefficient, and εij is an individual-level error term.

The final step of the analysis will consist of formulating exponential random graph models [42] (ERGMs), a class of statistical models for social networks, to examine the structure of networks in the treatment and control communities post implementation. ERGMs account for the presence and absence of network ties (ie, edges) and provide an approach to modeling overall network structure by modeling small, local, tie-based structures, such as reciprocity and triangles. These models permit inferences about whether a given network has significantly more or less of a given feature (such as reciprocity or triadic closures) than one would expect and provide a deeper understanding of how and why network ties arise. This study will estimate a series of ERGMs to model network features in the treatment and control group communities [42]:

$$P\theta\ (G) = ce\theta 1z1\ (G) + \theta 2z2\ (G) + \ldots + \theta pzp\ (G)$$

where the probability of a given network *G* is given by a sum of network statistics (z in the equation above, which represents counts of the number of network configurations in the given network *G*) weighted by parameters (θ in the equation above) inside an exponential (and where c is a normalizing constant).

Qualitative Analysis of Process Evaluation Data

The study team will analyze qualitative data using a hybrid deductive-inductive analytic approach that emphasizes addressing key constructs of the CFIR, such as intervention characteristics, outer and inner settings, characteristics of individuals, and process, while allowing for emerging patterns. Interview transcripts will be uploaded to the qualitative data analysis software ATLAS.ti (Lumivero). Individual team members will code interview transcripts using a codebook that is guided by CFIR domains and the constructs within them, and that allows team members to identify emerging themes in the data. Furthermore, 2 coders (KJ and FW) will code a subset of transcripts and compare the level of agreement to assess reliability. The study team will calculate the Cohen κ statistic to check for coder consistency in applying the codebook. If a high level of consensus is not reached, team members will work to resolve conflicts until reaching a level of high intercoder agreement (0.80).

Through an iterative process, the qualitative study team will continually create, refine, and discard codes. The qualitative study team will hold twice-weekly data meetings during the analysis phase to discuss coding. As transcripts are discussed and coded, the study team will also write analytical memos to record observations, refine preliminary themes, and capture emerging themes. When all transcripts have been coded and discussed, the study team will engage in an iterative process of searching data using ATLAS.ti, comparing data segments from searches, and writing analytic memos to answer the process evaluation's research questions, thereby developing a full understanding of the barriers to and facilitators of implementing FAs within SFP.

XSL•FO
RenderX

Quantitative Analysis of Cost Evaluation Data

The study team will use the collected cost data to calculate average program start-up costs and ongoing program implementation costs, primarily from the funder perspective but also from the societal perspective, reflecting the costs from the perspectives of all funders, plus the value of any in-kind donations of time or other supplies or equipment. In cost analyses, costs will be calculated for the following resource categories: labor, materials or supplies, equipment, and other resource categories. Mean costs will be calculated by intervention assignment (SFP+FA or SFP-only), and cost variability will be examined by intervention, staff mix, or communities. The sensitivity of costs will also be examined against assumptions about unit costs or other input values. Cost estimates, particularly the costs of SFP, will be compared with other estimates in the research literature, along with the primary drivers of differences or similarities in costs across implementations of SFP. These costs will be reported as costs per site per year and costs per family per year.

The total program costs for SFP+FA and SFP will be estimated and subsequently used to calculate the incremental cost of SFP+FA relative to SFP-only. These analyses will compare differences in start-up costs and ongoing program costs to implement the programs during the study period. Fixed costs, which do not vary based on the number of families served, will be compared with variable costs, which increase for each family or site added. Understanding the breakdown of costs between start-up and implementation is necessary for accurately estimating the costs of program scale-up or new program implementation in other settings.

The study team will also explore the feasibility of conducting a budget impact analysis. A budget impact analysis would first identify the organizations and agencies that incur costs and experience cost savings resulting from SFP+FA or SFP-only. These interventions may affect service usage in multiple sectors, including the health care system, the criminal justice system, and the child welfare or protective services system. A budget impact analysis will consider a time horizon beyond the trial period, such as 5 or 10 years, to analyze whether and how the intervention affects longer-term outcomes in each of the various systems that affect families. One aspect of this feasibility study is to determine whether data are available to estimate the budget impact from the state, federal, or other funder perspectives. A potential barrier is limited access to appropriate data from all relevant systems to conduct the analysis. Although it is possible to assess impacts on emergency department and inpatient health care usage using local data from Medicaid claims or hospital discharges, it may be more challenging to obtain data from nonhealth social services systems. An important part of this study will be to assess the feasibility of obtaining relevant data from nonhealth social services systems.

Mixed Methods Analysis of Process, Effectiveness, and Cost Data

In the final phase of the analysis, the evaluation data will be analyzed using a mixed methods analysis approach in which qualitative and quantitative data collected from the process, outcomes, and cost components of the hybrid type 1 design are systematically integrated and analyzed. Mixed methods analysis helps answer questions that could not be answered by the qualitative or quantitative approaches alone, thereby enriching results, strengthening the overall reliability of a study's findings, and providing the opportunity to uncover unique insights and findings that would have been otherwise neglected using a single method. The study's mixed methods analysis provides a systematic approach to integrating and interpreting data from the outcome evaluation's effectiveness trial, the process evaluation's focus on implementation barriers and facilitators, and the cost evaluation's cost and cost-effectiveness estimates.

The findings from the qualitative and quantitative analyses of process, outcome, and cost data will be systematically integrated using joint display tables that provide a visual framework for intentionally integrating findings with a clear rationale, thereby illuminating insights beyond separate quantitative and qualitative analyses. This mixed methods analysis provides a critical benefit to the study by providing a robust understanding of the intervention's effectiveness, implementation barriers and facilitators, and cost-effectiveness, and thereby enhancing the likelihood that the intervention strategies shown to be effective will be sustainable when implemented in future communities.

Ethical Considerations

The study protocol, data collection instruments, and consent forms were approved by the Advarra Institutional Review Board (#00071669). The study is conducted in accordance with the ethical principles of the Belmont Report, with the study's effectiveness trial conducted in accordance with International Council on Harmonisation Good Clinical Practice, applicable United States Code of Federal Regulations, and the NCIPC Terms and Conditions of Award. Important protocol modifications will be shared with the institutional review board, trial registries, journals, and trial participants.

Informed consent will be obtained from all participating parents and caregivers, and parental permission and assent will be obtained for all participating youth. Consent and assent procedures will be conducted electronically and designed to ensure participants' understanding of study procedures and the associated risks and benefits of participation. Participants will be compensated for their time using electronic gift cards for engaging with intervention activities and completing surveys or interviews. All consent forms, surveys, and interview data are stored within a Health Insurance Portability and Accountability Act (HIPAA)–compliant network that is accessible only to authorized project staff using password-protected devices. No identifiable information will appear in reports or publications.

Results

Funding was awarded in 2022, and institutional review board approval was obtained in November 2023. Intervention and evaluation development lasted approximately 1 year, with study partners engaged in the co-design process to ensure alignment with the principles of community-placed behavioral health research. Study recruitment began in November 2023 and is anticipated to continue until September 2026. Final data

XSL•FO
RenderX

collection is anticipated by March 2027, followed by data analysis.

Discussion

Strengths and Limitations

One of the key strengths of this study is the use of a hybrid type 1 effectiveness-implementation design, which allows for the simultaneous evaluation of both the clinical outcomes and the implementation processes in real-world settings. This comprehensive approach provides valuable insights into not only whether the intervention works, but also how it can be implemented effectively in community settings. The cluster RCT design involving 36 New Jersey communities further strengthens the study by reducing the likelihood of selection bias and enabling the comparison of intervention and control groups at the community level. The study's integration of both family- and community-level interventions addresses multiple SDOH and enhances the potential for sustainable, long-term impact on reducing ACEs and substance use.

However, the study has certain limitations that could affect the interpretation of its results. First, measuring compliance with family referrals and tracking the use of community resources may be challenging, potentially leading to incomplete data. Furthermore, the aggregation of family-level outcomes to assess community-level impact could introduce variability that might obscure the intervention's true effect. The reliance on self-reported data, particularly regarding sensitive topics such as ACEs and substance use, could also introduce response bias. Finally, the study's focus on communities in New Jersey limits the generalizability of its findings to other regions with different demographic and socioeconomic characteristics.

Implications for Practice

The findings of this study have important implications for both policy and practice, particularly in addressing the dual challenges of substance use and ACEs through community-based interventions. If successful, the integration of trauma-informed FAs with the SFP could serve as a scalable model for addressing ACEs and substance use prevention in other high-risk communities. By aligning services and resources across family and community levels, this approach could inform the development of more comprehensive, multilevel interventions that are responsive to the specific needs of families and communities. Additionally, the study's process evaluation will provide valuable insights into the barriers to and facilitators of implementing such interventions, helping practitioners and policymakers identify strategies to enhance service delivery and improve family outcomes.

Conclusions

This study addresses a key gap in the evidence base by rigorously evaluating a novel intervention that effectively integrates family- and community-level factors to prevent substance use, overdose, and ACEs. The study's research design will build the evidence base for community-level interventions by using a hybrid type 1 design that integrates outcome, process, and cost evaluations to optimize the likelihood that the intervention strategies shown to be effective will be sustainable when implemented in future communities. The findings from this study will provide a roadmap for how disproportionately affected communities can implement the intervention to prevent substance use, overdose, and ACEs.

Acknowledgments

The authors thank Drs. Losby, DePadilla, and Anderson from the

Centers for Disease Control and Prevention (CDC) for the helpful feedback that they provided on an early version of the manuscript.

The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the CDC.

Funding

This study underwent independent peer review by the CDC (Cooperative Agreement # 1 U01CE003482-01-00).

Data Availability

The datasets from this study will be deidentified and stored at the CDC NCIPC data repository (in accordance with the NCIPC Data Management Plan) for use by other researchers.

Authors' Contributions

PG and DE are coprincipal investigators and led the conception and design of the study. KJ, SK, BS, and AZ are coinvestigators and contributed to the design and implementation of the study along with FW, MM, BG, and JW. XJ is the CDC NCIPC Division scientific collaborator for the study.

Conflicts of Interest

None declared.

XSL•FO
RenderX

References

1. Merrick MT, Ford DC, Ports KA, et al. Vital signs: estimated proportion of adult health problems attributable to adverse childhood experiences and implications for prevention - 25 states, 2015-2017. MMWR Morb Mortal Wkly Rep 2019 Nov 8;68(44):999-1005. [doi: 10.15585/mmwr.mm6844e1] [Medline: 31697656]
2. Adverse childhood experiences prevention strategy. : National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2021 URL: https://restoredcdc.org/www.cdc.gov/aces/media/pdfs/ACEs-Strategic-Plan_Final_508.pdf [accessed 2026-03-20]
3. Swedo EA, Aslam MV, Dahlberg LL, et al. Prevalence of adverse childhood experiences among U.S. adults - behavioral risk factor surveillance system, 2011-2020. MMWR Morb Mortal Wkly Rep 2023 Jun 30;72(26):707-715. [doi: 10.15585/mmwr.mm7226a2] [Medline: 37384554]
4. Shonkoff JP, Garner AS, Committee on Psychosocial Aspects of Child and Family Health, Committee on Early Childhood, Adoption, and Dependent Care, Section on Developmental and Behavioral Pediatrics. The lifelong effects of early childhood adversity and toxic stress. Pediatrics 2012 Jan;129(1):e232-e246. [doi: 10.1542/peds.2011-2663] [Medline: 22201156]
5. Peterson C, Aslam MV, Niolon PH, et al. Economic burden of health conditions associated with adverse childhood experiences among US adults. JAMA Netw Open 2023 Dec 1;6(12):e2346323. [doi: 10.1001/jamanetworkopen.2023.46323] [Medline: 38055277]
6. Metzler M, Merrick MT, Klevens J, Ports KA, Ford DC. Adverse childhood experiences and life opportunities: shifting the narrative. Child Youth Serv Rev 2017 Jan 14;72(141-149):141-149. [doi: 10.1016/j.childyouth.2016.10.021] [Medline: 37961044]
7. Matjasko JL, Herbst JH, Estefan LF. Preventing adverse childhood experiences: the role of etiological, evaluation, and implementation research. Am J Prev Med 2022 Jun;62(6 Suppl 1):S6-S15. [doi: 10.1016/j.amepre.2021.10.024] [Medline: 35597583]
8. Swedo EA, Sumner SA, de Fijter S, et al. Adolescent opioid misuse attributable to adverse childhood experiences. J Pediatr 2020 Sep;224:102-109. [doi: 10.1016/j.jpeds.2020.05.001] [Medline: 32437756]
9. Rothman EF, Edwards EM, Heeren T, Hingson RW. Adverse childhood experiences predict earlier age of drinking onset: results from a representative US sample of current or former drinkers. Pediatrics 2008 Aug;122(2):e298-e304. [doi: 10.1542/peds.2007-3412] [Medline: 18676515]
10. Choi NG, DiNitto DM, Marti CN, Choi BY. Association of adverse childhood experiences with lifetime mental and substance use disorders among men and women aged 50+ years. Int Psychogeriatr 2017 Mar;29(3):359-372. [doi: 10.1017/S1041610216001800] [Medline: 27780491]
11. Douglas KR, Chan G, Gelernter J, et al. Adverse childhood events as risk factors for substance dependence: partial mediation by mood and anxiety disorders. Addict Behav 2010 Jan;35(1):7-13. [doi: 10.1016/j.addbeh.2009.07.004] [Medline: 19720467]
12. Stein MD, Conti MT, Kenney S, et al. Adverse childhood experience effects on opioid use initiation, injection drug use, and overdose among persons with opioid use disorder. Drug Alcohol Depend 2017 Oct 1;179:325-329. [doi: 10.1016/j.drugalcdep.2017.07.007] [Medline: 28841495]
13. Understanding the opioid overdose epidemic.: Centers for Disease Control and Prevention; 2023. URL: https://www.cdc.gov/overdose-prevention/about/understanding-the-opioid-overdose-epidemic.html [accessed 2024-04-30]
14. Gervin DW, Holland KM, Ottley PG, Holmes GM, Niolon PH, Mercy JA. Centers for Disease Control and Prevention investments in adverse childhood experience prevention efforts. Am J Prev Med 2022 Jun;62(6 Suppl 1):S1-S5. [doi: 10.1016/j.amepre.2021.11.014] [Medline: 35597578]
15. Spoth R, Lopez Reyes M, Redmond C, Shin C. Assessing a public health approach to delay onset and progression of adolescent substance use: latent transition and log-linear analyses of longitudinal family preventive intervention outcomes. J Consult Clin Psychol 1999 Oct;67(5):619-630. [doi: 10.1037//0022-006x.67.5.619] [Medline: 10535229]
16. Kumpfer KL, Magalhães C. Strengthening Families Program: an evidence-based family intervention for parents of high-risk children and adolescents. J Child Adolesc Subst Abuse 2018 May 4;27(3):174-179. [doi: 10.1080/1067828X.2018.1443048]
17. Kumpfer KL, Magalhães C, Ahearn Greene J. Chapter 18: strengthening families program. In: Evidence-Based Parenting Education: Routledge; 2015:303-318. [doi: 10.4324/9781315766676-28]
18. Spoth RL, Redmond C, Shin C. Randomized trial of brief family interventions for general populations: adolescent substance use outcomes 4 years following baseline. J Consult Clin Psychol 2001 Aug;69(4):627-642. [doi: 10.1037//0022-006x.69.4.627] [Medline: 11550729]
19. Spoth R, Trudeau L, Shin C, et al. Longitudinal effects of universal preventive intervention on prescription drug misuse: three randomized controlled trials with late adolescents and young adults. Am J Public Health 2013 Apr;103(4):665-672. [doi: 10.2105/AJPH.2012.301209] [Medline: 23409883]
20. What is the strengthening families program? Strengthening Families Foundation. URL: https://strengtheningfamiliesfoundation.org/program-overview/#:~:text=The%20low%2Dcost%20SFP%207,downloadable%20handouts%20and%20tracking%20sheets [accessed 2024-04-30]

XSL•FO
RenderX

21. Pinheiro-Carozzo NP, Murta SG, Vinha LDA, da Silva IM, Fontaine A. Beyond effectiveness of the Strengthening Families Program (10-14): a scoping RE-AIM-based review. Psicol Reflex Crit 2021 Jun 15;34(1):16. [doi: 10.1186/s41155-021-00182-z] [Medline: 34131838]
22. Kirby S, Edwards K, Yu S, et al. Improving outcomes for marginalised rural families through a care navigator program. Health Promot J Austr 2021 Apr;32(2):285-294. [doi: 10.1002/hpja.348] [Medline: 32323411]
23. Summers NA, Colasanti JA, Feaster DJ, et al. Predictors for poor linkage to care among hospitalized persons living with HIV and co-occurring substance use disorder. AIDS Res Hum Retroviruses 2020 May;36(5):406-414. [doi: 10.1089/AID.2019.0153] [Medline: 31914790]
24. Stitzer M, Matheson T, Cunningham C, et al. Enhancing patient navigation to improve intervention session attendance and viral load suppression of persons with HIV and substance use: a secondary post hoc analysis of the Project HOPE study. Addict Sci Clin Pract 2017 Jun 27;12(1):16. [doi: 10.1186/s13722-017-0081-1] [Medline: 28651612]
25. Radigan M, Wang R, Chen Y, Xiang J. Youth and caregiver access to peer advocates and satisfaction with mental health services. Community Ment Health J 2014 Nov;50(8):915-921. [doi: 10.1007/s10597-014-9709-8] [Medline: 24522553]
26. Anda R. The role of adverse childhood experiences in substance misuse and related behavioral health problems. : SAMHSA's Center for the Application of Prevention Technologies; 2018 URL: https://naswwv.socialworkers.org/LinkClick.aspx?fileticket=vt4_UPPKMz4%3D&portalid=13 [accessed 2026-03-20]
27. Adverse Childhood Experiences (ACEs) Prevention Resource for Action: A Compilation of the Best Available Evidence. : National Center for Injury Prevention and Control, Centers for Disease Control and Prevention; 2019 URL: https://www.cdc.gov/violenceprevention/pdf/aces-prevention-resource_508.pdf [accessed 2026-03-20]
28. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
29. Leask CF, Sandlund M, Skelton DA, et al. Framework, principles and recommendations for utilising participatory methodologies in the co-creation and evaluation of public health interventions. Res Involv Engagem 2019;5(1):2. [doi: 10.1186/s40900-018-0136-9] [Medline: 30652027]
30. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226. [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
31. Felitti VJ, Anda RF, Nordenberg D, et al. Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. The Adverse Childhood Experiences (ACE) Study. Am J Prev Med 1998 May;14(4):245-258. [doi: 10.1016/s0749-3797(98)00017-8] [Medline: 9635069]
32. Glennerster R, Takavarasha K. Running Randomized Evaluations: A Practical Guide: Princeton University Press; 2013. [doi: 10.2307/j.ctt4cgd52]
33. Lines LM, Long MC, Humphrey JL, et al. Artificially intelligent social risk adjustment: development and pilot testing in Ohio [internet]. : Research Triangle Park (NC): RTI Press; 2022. [Medline: 37011120]
34. HCaUP. HCUP State Inpatient databases (SID).: Agency for Healthcare Research and Quality; 2005. URL: www.hcup-us.ahrq.gov/sidoverview.jsp
35. Healthcare Cost and Utilization Project (HCUP). HCUP State Emergency Department databases (SEDD).: Agency for Healthcare Research and Quality; 2009. URL: www.hcup-us.ahrq.gov/seddoverview.jsp
36. Cesare N, Lines LM, Chandler R, et al. Development and validation of a community-level social determinants of health index for drug overdose deaths in the HEALing Communities Study. J Subst Use Addict Treat 2024 Feb;157:209186. [doi: 10.1016/j.josat.2023.209186] [Medline: 37866438]
37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 1995 Jan 1;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
38. Kwok OM, Underhill AT, Berry JW, Luo W, Elliott TR, Yoon M. Analyzing longitudinal data with multilevel models: an example with individuals living with lower extremity intra-articular fractures. Rehabil Psychol 2008 Aug;53(3):370-386. [doi: 10.1037/a0012765] [Medline: 19649151]
39. Dimick JB, Ryan AM. Methods for evaluating changes in health care policy: the difference-in-differences approach. JAMA 2014 Dec 10;312(22):2401-2402. [doi: 10.1001/jama.2014.16153] [Medline: 25490331]
40. Wing C, Simon K, Bello-Gomez RA. Designing difference in difference studies: best practices for public health policy research. Annu Rev Public Health 2018 Apr 1;39:453-469. [doi: 10.1146/annurev-publhealth-040617-013507] [Medline: 29328877]
41. Jordan DG, Winer ES, Salem T. The current status of temporal network analysis for clinical science: considerations as the paradigm shifts? J Clin Psychol 2020 Sep;76(9):1591-1612. [doi: 10.1002/jclp.22957] [Medline: 32386334]
42. Lusher D, Koskinen J, Robins G, editors. Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications: Cambridge University Press; 2013. [doi: 10.1017/CBO9780511894701]

XSL•FO
RenderX

Abbreviations

ACE: adverse childhood experience
CDC: Centers for Disease Control and Prevention
CFIR: Consolidated Framework for Implementation Research
ERGM: exponential random graph models
FA: family advocate
GLMM: generalized linear mixed model
HCUP: Healthcare Cost and Utilization Project
HIPAA: Health and Insurance Portability and Accountability Act
LGM: latent growth model
MAR: missing at random
MCAR: missing completely at random
NCIPC: National Center for Injury Prevention and Control
PRODUCES: Problem, Objective, Design, (end-)Users, Co-creators, Evaluation, Scalability
RCT: randomized controlled trial
RTI: Research Triangle Institute
SDOH: social determinants of health
SFP: Strengthening Families Program

Edited by J Sarvestan; submitted 01.Oct.2025; this is a non–peer-reviewed article;accepted 23.Jan.2026; published 09.Apr.2026.

Please cite as:
Elgin D, Graham PW, Johnson K, Khan S, Mills M, Wondimagegnehu F, Gibbons B, Williams J, Jiang X, Zapcic A, Simon B
Evaluating the Combined Effects of an Adverse Childhood Experiences–Focused Family Advocate Model and the Strengthening Families Program: Study Protocol for a Hybrid Type 1 Effectiveness-Implementation Trial in 36 New Jersey Communities
JMIR Res Protoc 2026;15:e85151
URL: https://www.researchprotocols.org/2026/1/e85151
doi:10.2196/85151

© Dallas Elgin, Phillip W Graham, Kiersten Johnson, Sazid Khan, Michael Mills, Feker Wondimagegnehu, Brent Gibbons, Jason Williams, Xinyi Jiang, Andrea Zapcic, Brittany Simon. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Feasibility of Internet-Based Mind-Body Training for Adults With Subjective Cognitive Decline: Protocol for a Randomized Controlled Trial

Ruchika S Prakash[1], PhD; Madhura Phansikar[1], PhD; Megan E Fisher[1], MS; Sarah Dove[2], PhD; Rebecca Andridge[3], PhD; Doug Scharre[4], MD; Kathy D Wright[5], PhD; Elizabeth Head[6], PhD; Claire O'Callaghan[7], PhD; Quinn K Meyerson[1], BA; Elizabeth J Duraney[1], PhD; Matthew Schroeder[1], MS; Niyathi Chakrapani[1], BA; Elise K Grant[1], BS; Tyreek Jones[1], BA; James Teng[1], MS; Michael Esterman[8,9], PhD; Joshua M Smyth[1], PhD; Thomas Gregoire[10], PhD; Rond Malhas[6], BS

[1]Department of Psychology, College of Arts & Sciences, The Ohio State University, 1835 Neil Avenue, Columbus, OH, United States

[10]College of Social Work, The Ohio State University, Columbus, OH, United States

[2]Office of Curriculum and Scholarship, College of Medicine, The Ohio State University, Columbus, OH, United States

[3]Division of Biostatistics, College of Public Health, The Ohio State University, Columbus, OH, United States

[4]Department of Neurology, College of Medicine, The Ohio State University, Columbus, OH, United States

[5]College of Nursing, Wexner Medical Center, The Ohio State University, Columbus, OH, United States

[6]Department of Pathology and Laboratory Medicine, University of California, Irvine, Irvine, CA, United States

[7]School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[8]National Center for PTSD, VA Boston Healthcare System, Boston, MA, United States

[9]Department of Psychiatry, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, United States

Corresponding Author:
Ruchika S Prakash, PhD
Department of Psychology, College of Arts & Sciences, The Ohio State University, 1835 Neil Avenue, Columbus, OH, United States

Abstract

Background: In the United States, the prevalence of Alzheimer disease (AD) is projected to double over the next 30 years, with associated familial and societal costs estimated at US $1 trillion annually if current trends continue. Although pharmacological treatments of AD are showing promise, the adoption of healthy lifestyle behaviors, particularly during the preclinical phase of AD, may reduce dementia rates by up to 45%. Subjective cognitive decline (SCD), defined as persistent self-perceived declines in cognitive functioning compared with previously normal cognitive abilities, has been identified as a potential preclinical stage of AD.

Objective: This randomized controlled trial aims to evaluate the feasibility and acceptability of an internet-based, asynchronous mindfulness-based stress reduction program compared with an active control group (an internet-based lifestyle education program). Secondary objectives include examining preliminary effects of each intervention on SCD, plasma-based biomarkers of amyloid and tau pathology, and everyday mind-wandering.

Methods: Sixty adults aged 50 years and older will be screened for SCD in the absence of objective cognitive impairment, based on the Uniform Data Set Neuropsychological Battery (version 3.0) from the National Alzheimer's Coordinating Center. Eligible and consenting participants will complete behavioral and imaging-based tasks of sustained attention and mind-wandering, as well as blood draws at baseline and after the 8-week intervention. After baseline assessments, participants will be randomized to either an internet-based, asynchronous mindfulness-based stress reduction program or the internet-based lifestyle education program. Both programs have been adapted from our manualized in-person programs and refined through focus group interviews with the target population.

Results: The study was funded in April 2024. Phase 1 focused on iterative development of the 2 programs based on focus group feedback. Recruitment for the randomized controlled trial (internet-based mind-body training trial) began in June 2025 and is ongoing. Recruitment is expected to conclude in September 2026, with data collection ending in December 2026.

Conclusions: Behavioral, lifestyle-based interventions that emphasize experiential practices show promise as preventative strategies to prevent decline in cognitive and brain health. Yet, there remain significant barriers to engaging with in-person programs, including limited accessibility, time and schedule constraints, and travel logistics. The internet-based mind-body training trial will evaluate the feasibility and acceptability of 2 fully online, mind-body training programs for adults at risk for

XSL•FO
RenderX

AD. Future Stage II and Stage III studies will be necessary to establish the efficacy of these programs for improving AD biomarkers and cognitive outcomes and their broader dissemination to adults noticing subtle changes in cognitive functioning.

Trial Registration: Clinicaltrials.gov NCT07019402; https://clinicaltrials.gov/study/NCT07019402

International Registered Report Identifier (IRRID): DERR1-10.2196/86276

(JMIR Res Protoc 2026;15:e86276) doi:10.2196/86276

KEYWORDS

subjective cognitive decline; amyloid and tau pathology; online mind-body interventions; mindfulness; pilot randomized controlled trial

Introduction

Alzheimer disease (AD) is the leading cause of dementia worldwide, affecting 6.7 million Americans aged 65 years and older. According to the *2023 Alzheimer's Disease Facts and Figures* [1], this number is projected to double by 2050, with individuals diagnosed with AD expected to comprise 22% of the US population. With limited treatment options, there has been a renewed focus on targeting neurodegenerative and pathophysiological processes through behavioral and lifestyle-based interventions during the prodromal phase of AD. These preclinical stages represent a window of opportunity for preventative interventions aimed at slowing the rate of accumulation of AD-related proteinopathies or delaying the functional impact of neurodegeneration on cognition [2].

Subjective cognitive decline (SCD), defined as perceived, persistent declines in cognitive functioning relative to one's previous abilities, was first introduced in 2014 as a potential preclinical stage of AD [3]. Individuals with SCD are identified through self-reported concerns about cognitive functioning, typically assessed through administering simple paper-and-pencil questionnaires, such as the Everyday Cognition (ECog) Questionnaire [4], the Subjective Cognitive Decline Questionnaire [5], the Cognitive Change Index [6], or the Cognitive Function Index [7]. The classification of SCD additionally requires normal performance on standardized cognitive measures, a criterion that allows individuals with SCD to be distinguished from those with mild cognitive impairment (MCI) [3]. SCD has been associated later in life with the classic neural signatures of AD, including cortical thinning in key regions of the temporal lobe [8,9]; progressive whole-brain atrophy [10]; loss of gray matter volume, especially in the hippocampus [11]; and white matter hyperintensities [12]. It has also been associated with β-amyloid plaques [13,14] and other cerebrospinal fluid (CSF)–based AD biomarkers [15]. Moreover, it is estimated that 25% to 40% of adults seeking medical advice at Memory Disorders Clinics are reported to have SCD [16,17], which would allow for preventative interventions to be more easily implemented, potentially with higher rates of adherence compared to a community sample [2].

Mindfulness meditation, traditionally offered through manualized 8-week programs such as the mindfulness-based stress reduction program or the mindfulness-based cognitive therapy program, has a growing evidence base for its potential to benefit cognitive and affective health [18-20]. Mindfulness, as taught in these programs, involves the intentional cultivation of present moment awareness to fully observe and experience the unfolding of each moment [21]. Eight weeks of training in mindfulness practices typically include experiential exercises ranging from focused attention practices, which anchor attention to a single phenomenon (eg, the breath), to open monitoring practices, which involve flexibly switching attention to whichever phenomenon becomes most salient.

Accordingly, attention and its various components, including orientation, alertness, and executive control, as well as lapses in attention, often termed mind-wandering, have been investigated as primary outcome variables in many clinical trials of mindfulness meditation (see [22] for a scoping review). Results from randomized controlled trials (RCTs) in our laboratory have also demonstrated that mindfulness training reduces off-task thinking or mind-wandering [23,24], although in our Stage II efficacy study [24], these effects were not maintained at long-term follow-up. These reductions in mind-wandering propensity, replicated by other clinical trials (see Feruglio et al [25] for a systematic review), provide preliminary support that training in mindfulness meditation may strengthen goal maintenance systems, thereby enhancing one's ability to prioritize task-relevant representations and filter out irrelevant, internal thoughts or external stimuli in the environment.

Additionally, across cross-sectional and intervention studies, mindfulness training has been shown to reduce activity within the default mode network (DMN) during tasks of sustained attention, strengthen intranetwork connectivity of the DMN, and increase connectivity between the DMN and the frontoparietal networks. In a cross-sectional study of healthy older adults, our laboratory provided the first evidence of trait mindfulness to be associated with greater intranetwork DMN connectivity [1,2,26]. Expert meditators with more than 10,000 hours of meditative practices consistently show reduced activity in the posterior cingulate cortex and the medial prefrontal cortex (core DMN hubs) both during cognitively demanding tasks [27] and while meditating in the magnetic resonance imaging (MRI) scanner [28]. Short-term RCTs in meditation novices have also shown reduced low-frequency amplitude within DMN regions [29] and increased negative connectivity between DMN and frontoparietal networks following the training [30,31].

There is also preliminary support for mindfulness training to strengthen the intranetwork connectivity of regions of the DMN, including the posterior cingulate cortex, the medial prefrontal cortex, and the hippocampus in individuals with MCI [32]. In the same study, 8 weeks of mindfulness training also resulted

XSL•FO
RenderX

in the maintenance of hippocampal volume, a validated biomarker for clinical trials, including those conducted during predementia stages, compared to the volumetric decline observed in the control group. More recently, findings from the PREVENT-AD (Pre-symptomatic Evaluation of Experimental or Novel Treatments for AD) study established direct links between mindfulness and AD pathology in adults aged older than 60 years and at risk for dementia (ie, having either a parent or at least two siblings with AD-like dementia) [33]. Trait mindfulness, assessed using the Five Facet Mindfulness Questionnaire, was associated with a lower β-amyloid burden in the temporoparietal and frontal regions, including the precuneus, the posterior cingulate cortex, and the orbitofrontal cortex [34]. Facets of mindfulness were also associated with lower tau burden in the bilateral temporal regions, such as the entorhinal cortices; amygdala; and parahippocampal, inferior temporal, and fusiform gyri. Finally, trait mindfulness was also linked to reduced cognitive decline in global cognition, sustained attention, immediate memory, and delayed memory. This study thus provided the first direct evidence of a multivariate relationship among mindfulness, AD pathophysiology, and cognitive decline.

Building on these cross-sectional studies, our RCT, the internet-based mind-body training (iMBT) program, will evaluate the feasibility and acceptability of 2 fully asynchronous mind-body interventions: an internet-based mindfulness-based stress reduction (MBSR) program and an internet-based lifestyle education (iLifeEd) program. The 2 programs were co-developed with input from focus groups, including members of our target population, namely, adults aged 50 years and older with SCD. We will additionally examine preliminary effects on plasma $A\beta_{42}/A\beta_{40}$ and phosphorylated tau (p-tau217), biomarkers with high sensitivity and specificity in detecting AD amyloid and tau pathology, with validation studies showing a strong correspondence between plasma-based biomarkers and positron emission tomography (PET), CSF, and autopsy measures [35-37]. Other secondary outcomes include ecological momentary assessment (EMA) of mind-wandering and cognitive variability, DMN integrity during sustained attention, and behavioral measures of mind-wandering.

Methods

Study Design

The primary goal of our iMBT trial is to examine the feasibility and acceptability of the internet-based, asynchronous mindfulness-based stress reduction (iMBSR) program using a Stage I study design proposed in the National Institutes of Health (NIH) Stage Model [38]. Pilot studies are foundational for large-scale efficacy studies and help assess recruitment potential, delivery of the intervention, adherence and acceptability, barriers to compliance, and potential safety issues before the launch of large clinical trials [39-41]. These studies can also be used to generate estimates of a treatment's effects and its variance and covariance to provide necessary data for large-scale trials [42]. For this RCT, we are recruiting a sample of adults with SCD, and participants are being randomized into 1 of 2 groups: iMBSR or iLifeEd. We are collecting data on the feasibility and acceptability of the program using a priori defined criteria. Additionally, blood plasma will also be collected before randomization and after 8 weeks of training to determine the effect of mindfulness training on $A\beta_{42}/A\beta_{40}$ ratio and p-tau217. To examine the preliminary effects of the iMBSR program on brain health, neuroimaging sessions will also be completed at both time points with the strength of the DMN serving as the measure of neuronal health. Participants will additionally complete computerized measures of sustained attention, EMA of mind-wandering and cognitive functioning, and several self-report questionnaires to assess their psychosocial profile. Figure 1 illustrates the timeline of assessment and intervention sessions for the ongoing clinical trial.

Figure 1. Timeline of the internet-based mind-body training randomized controlled trial, including screening sessions, baseline assessments, and intervention sessions. Following a combination of web or phone screening and an in-person screening session, participants complete behavioral and magnetic resonance imaging (MRI) sessions, which are repeated immediately after the 8-week training program. Blood and ecological momentary assessment (EMA) data are also collected before and after the intervention. LifeEd: lifestyle education; MBSR: mindfulness-based stress reduction.

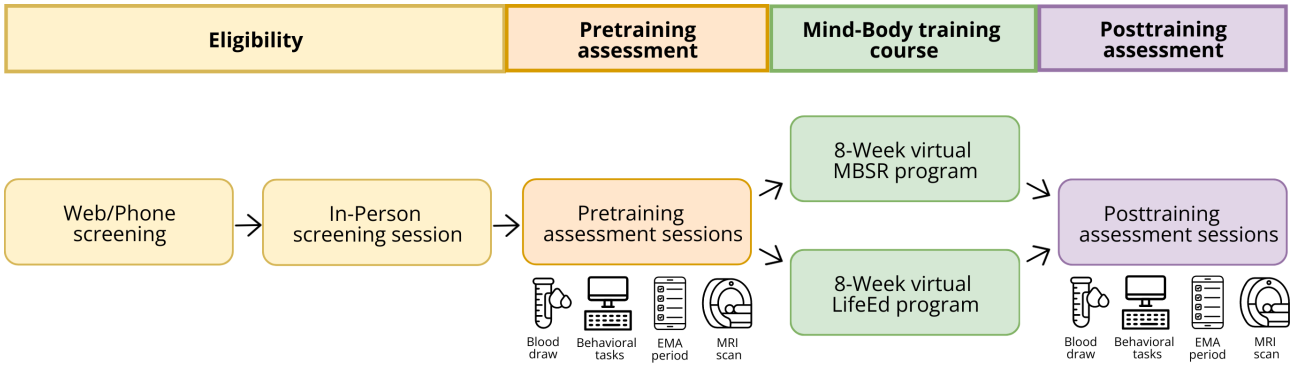

Setting

All behavioral data are being collected in the Clinical Neuroscience Lab, located within the Department of Psychology at The Ohio State University. The Clinical Neuroscience Lab, directed by the study's principal investigator (PI; RSP), contains 3 dedicated testing rooms equipped for neuropsychological testing. Blood samples are also obtained at the laboratory by trained phlebotomists from the Clinical Research Center at The Ohio State University. Neuroimaging data are being acquired at the Center for Cognitive and Behavioral Brain Imaging, an interdisciplinary research facility housed within the Department of Psychology at The Ohio State University.

XSL•FO
RenderX

Participant Eligibility

The target population for this study is adults aged older than 50 years with SCD. Following the criteria identified in a study by Jessen et al [43] to recruit participants with SCD, we will administer the ECog Questionnaire to quantify self-reported deficits in cognitive functioning. The ECog Questionnaire is a self-reported and an informant-reported 39-item questionnaire assessing cognitively based everyday activities [41]. This measure was recently found to be reliable and valid in an ethnoracially diverse sample of older adults, with scores on this measure correlating reliably with concurrent and longitudinal changes in cognition and brain volumes [44]. In the proposed study, we chose to use the self-reported ECog Questionnaire to reduce participant burden. Emerging evidence suggests that inclusion criteria requiring input from informants, such as family members, disproportionately exclude traditionally underrepresented groups, including individuals with low socioeconomic status and those from racial and ethnic minority backgrounds [45,46]. Furthermore, for this study, we will use the previously validated cutoff score of 1.31 to recruit participants with SCD. This cutoff has been independently validated in 2 separate studies [47,48], with individuals scoring greater than 1.31 on the self-reported ECog twice as likely to be diagnosed with MCI [47]. Only participants with elevated scores (>1.31) will be invited for an assessment of objective cognitive performance.

To screen out potential cognitive impairment and keep participant burden to a minimum, participants who meet our cutoff for the ECog Questionnaire will first complete the Telephone Montreal Cognitive Assessment (T-MOCA) [49] and the Instrumental Activities of Daily Living (IADL) [50] to screen out individuals with potential AD and related dementias. The T-MOCA is a widely used screener that can be administered over the phone to assess potential cognitive impairment. Only participants who score greater than or equal to 19 on the T-MOCA (out of 22) and score 8 (out of 8) on the IADL are then invited for an in-person detailed neuropsychological assessment session where they are administered the Uniform Data Set Neuropsychological Battery (version 3.0; UDS3-NB) from the National Alzheimer's Coordinating Center [51]. The UDS3-NB includes 10 widely used and well-standardized neuropsychological measures. Confirmatory factor analyses of the cognitive tests in the UDS3-NB in an ethnoracially diverse sample of more than 20,000 older adults from the National Alzheimer's Coordinating Center identified a 4-factor latent structure, including the domains of attention, episodic memory, language, and processing speed/executive functioning [52]. Table 1 lists the specific measures and metrics that quantify performance in the 4 cognitive domains. Using the Jak/Bondi diagnostic criteria [53,54], participants will be classified as having MCI or dementia and excluded from the study if they score (1) higher than 1.5 SD below the demographically adjusted normative mean on all 3 measures of 1 domain or (2) higher than 1.5 SD below the demographically adjusted normative mean on at least one measure of all 4 domains.

Table . Cognitive domains, tasks, and specific metrics from the Uniform Data Set Neuropsychological Battery version 3.0 (National Alzheimer's Coordinating Center) used to quantify objective cognitive performance.

Domain	Measure	Metric
Memory	• Craft Story 21 Recall (immediate) • Craft Story 21 Recall (delayed) • Benson complex figure copy (delayed)	• Verbatim score for the immediate recall • Verbatim score for the delayed recall • Total score for the delayed recall
Attention	• Number span test: forward • Number span test: backward • Benson complex figure copy (immediate)	• Number of correct trials • Number of correct trials • Total score for the Immediate Recall
Processing speed or executive functioning	• Trail making test: part A • Trail making test: part B • Verbal fluency: phonemic test	• Adjusted score: number of correct lines or times • Adjusted score: number of correct lines or times • Total number of correct "F" words and "L" words
Language	• Multilingual naming test • Category fluency • Category fluency	• Total score • Total number of correct "animal" words • Total number of correct "vegetable" words

Sample Size and Power

Our feasibility study will involve 60 adults with SCD randomized to one of 2 groups: iMBSR or iLifeEd. Our primary outcome variables for this study are the feasibility and acceptability of the protocols. With 30 participants per group, we will be able to estimate within-group feasibility percentages (eg, the percentage of participants who find the program prospectively acceptable and the percentage of people who indicate adequate levels of satisfaction) with a margin of error of no larger than +18% or −18%. For intervention adherence, we will have 80% power for detecting a proportion of greater than 50% adherence if the true proportion of participants who complete all modules is 75%.

Our key secondary outcome variables for this study are the plasma $A\beta_{42}/A\beta_{40}$ ratio and plasma p-tau217, quantifying reduction in amyloid and tau pathology following training. We

XSL•FO
RenderX

conducted sensitivity analyses to determine the effect size that we would need to observe to detect significant effects. Specifically, with 30 participants per group, we would need to observe a large effect size (Cohen *d*=0.74) to detect significant effects. Although there are, to our knowledge, no mindfulness-based interventions examining effects on plasma-based measures of amyloid and tau, prior work by one of the co-authors (EH) has demonstrated that 4 weeks of slow-paced breathing significantly decreases plasma $A\beta_{42}$, with a large effect size (Cohen *d*=0.81). As such, we are powered to detect the effect of our mindfulness intervention on plasma markers of amyloid. Other secondary variables of interest include behavioral and self-report measures of mind-wandering, including both laboratory- and EMA-based assessments of mind-wandering as well as the network strength of the DMN during the Gradual-Onset Continuous Performance Task. Given the pilot nature of the study, emphasis will be placed on the estimation of effect sizes and CIs for these effect sizes.

Study Procedures

Participants who meet the demographic criteria, along with criteria for the ECog Questionnaire, the T-MOCA, the IADL questionnaire, and MRI safety, are invited to the Clinical Neuroscience Laboratory at The Ohio State University for an initial in-person screening session. At this visit, visual acuity is assessed using the Snellen chart, and a blinded research assistant administers the UDS3-NB.

Participants meeting criteria for intact objective cognitive performance are then invited to a behavioral testing session at the Clinical Neuroscience Lab and a neuroimaging session at the Center for Cognitive and Behavioral Brain Imaging. During the behavioral session, a trained nurse from the Clinical Research Center at The Ohio State University collects the blood plasma sample, after which participants complete 2 behavioral measures of sustained attention and mind-wandering. As psychoaffective factors can influence perceived ratings of SCD, participants also complete a battery of questionnaires assessing stress (the Perceived Stress Scale [55]), depression (the Center for Epidemiologic Studies Depression Scale [56]), anxiety (the Beck Anxiety Inventory [57]), worry (the Penn State Worry Questionnaire [58]), quality of life (the World Health Organization Quality of Life–Brief [59]), and perceived racial discrimination (the Everyday Discrimination Scale [60]). We also administer the Motivation to Change Lifestyle and Health Behaviors for Dementia Risk Reduction [61] to quantify psychological and cognitive factors that influence engaging in health behaviors.

At the same session, participants receive training on EMA procedures and are provided with study-specific smartphones (Samsung Galaxy phones) to collect EMA data on mind-wandering, cognitive functioning, perceived stress, affect, and arousal for 5 days between sessions. Finally, structural and functional MRI data are collected at the neuroimaging session to quantify changes in the integrity of the DMN.

After completion of all assessment sessions, participants are randomized to either the iMBSR or the iLifeEd group with a 1:1 allocation ratio. All assessments will be readministered immediately following the 8-week intervention.

Randomization and Blinding

The study biostatistician (RA) generated the randomization sequence using a block size of 4 with stratification by sex. The project coordinator (EKG) received concealed envelopes for the 2 strata and is responsible for assigning eligible participants to one of the 2 groups. The project coordinator, the study facilitators, and the study biostatistician are the only team members unblinded to group assignment. All personnel conducting assessments, behavioral or neuroimaging, will remain blind to group assignment.

Internet-Based Mind-Body Training Programs

For this study, we developed 2 asynchronous, 8-week online mind-body training programs: the iMBSR and the iLifeEd training programs. Both programs were provisionally adapted from our standardized MBSR and lifestyle education manuals, which have previously been used to conduct Stage II efficacy studies of mindfulness meditation [24,62]. Our team's expertise in contemplative sciences, RCTs, and psychosocial interventions was complemented by instructional design expertise from the Office of Distance Education at The Ohio State University. In designing these platforms, we ensured inclusion of the critical components of the MBSR and lifestyle education program, including didactics, experiential learning, homework practices, social support and community building, and facilitator interactions. We additionally conducted focus groups with our target population, adults aged older than 50 years with SCD, to better understand their perspectives on healthy living, brain health, and engagement with digital interventions for improving brain health. Themes from the focus groups (N=23) were incorporated throughout the platforms, resulting in the final versions of the iMBSR and the iLifeEd programs.

The iMBSR and iLifeEd programs are delivered through The Ohio State University's Canvas for Continuing Education platform. Each program consists of 8 modules, with each module requiring approximately 2 hours per week to complete. Participants are expected to complete 1 module per week over 8 weeks. In addition, they are invited to optional weekly drop-in hours with the study PI, a licensed psychologist with more than 17 years of experience facilitating mind-body interventions.

The 2 programs are matched on the following metrics: (1) format of delivery,—both programs are online and quasi–self-paced; (2) timing—each of the 8 modules for the 2 programs is expected to take 2 hours per week, plus optional 30-minute drop-in hours; (3) content—both programs have been designed to integrate didactics, experiential practices, and homework practices; and (4) peer support—participants in both groups have access to a community board for discussions following experiential practices. The community boards are expected to foster a sense of community throughout the program. Table 2 outlines the themes of each module for both programs and provides an overview of the content covered in each module.

XSL•FO
RenderX

Table . Weekly themes, practices, and homework assignments for each of the 8 weeks of the internet-based mindfulness-based stress reduction (iMBSR) and internet-based lifestyle education (iLifeEd) programs developed for the internet-based mind-body training study.

	iMBSR	iLifeEd
Module 1	The potential of mindfulness meditation • Pt[a] 1. Introduction to study aims (40 min) • Pt 2. Defining mindfulness (45 min) • Pt 3. Putting it into practice! (30 min) • Hw[b] 1. Mindful eating (15 min) • Hw 2. Body scan practice (15 min)	Introduction to the study • Pt 1. Introduction to study aims (45 min) • Pt 2. Shift from the biomedical model to the biopsychosocial model (40 min) • Pt 3. Watch, reflect, and apply! (40 min) • Hw 1. Blue zones (20 min) • Hw 2. Stretching and toning (15 min) • Hw 3. On health literacy (5 min)
Module 2	Breath as our ever-present anchor • Pt 1. Cultivating awareness by tuning into our bodies (35 min) • Pt 2. Cultivating mindful awareness with the breath (50 min) • Pt 3. Putting it into practice! (30 min) • Hw 1. Body scan practice (15 min) • Hw 2. Breath awareness practice (15 min)	All about sedentary behavior • Pt 1. Introduction to sedentary behavior (40 min) • Pt 2. Benefits of sitting less (30 min) • Pt 3. Watch, reflect, and apply! (30 min) • Hw 1. Posture for back pain (5 min) • Hw 2. Stretching and toning (15 min) • Hw 3. Move more, sit less (10 min)
Module 3	Mindful awareness of sensations • Pt 1. Invitation to taking your seats (30 min) • Pt 2. Cultivating awareness through mindful yoga (45 min) • Pt 3. Putting it into practice! (35 min) • Hw 1. Breath awareness practice (20 min) • Hw 2. Mindful movement (15 min)	All about physical activity • Pt 1. Introduction to physical activity (30 min) • Pt 2. Benefits of physical activity and practical tips (35 min) • Pt 3. Watch, reflect, and apply! (35 min) • Hw 1. Exercising for brain health (5 min) • Hw 2. Stretching and toning (15 min) • Hw 3. Motivation for exercise (15 min)
Module 4	Mindful of emotions: impact on stress • Pt 1. Understanding and experiencing emotions (35 min) • Pt 2. Introduction to stress (45 min) • Pt 3. Putting it into practice! (30 min) • Hw 1. Mindful movement (15 min) • Hw 2. Sitting meditation practice (15 min)	On brain plasticity • Pt 1. Brain plasticity and cognitive reserve (40 min) • Pt 2. Benefits of cognitively stimulating activities (30 min) • Pt 3. Watch, reflect, and apply! (35 min) • Hw 1. More on neuroplasticity (15 min) • Hw 2. Stretching and toning (15 min) • Hw 3. Interview with Yaakov Stern: founder of cognitive reserve (5 min)
Module 5	Finding the space to make choices • Pt 1. Introduction to choiceless awareness (55 min) • Pt 2. How do we contribute to our own stress? (25 min) • Pt 3. Putting it into practice! (30 min) • Hw 1. Choiceless awareness practice (15 min) • Hw 2. Body scan practice (15 min)	All about sleep • Pt 1. Introduction to sleep (40 min) • Pt 2. Risks associated with sleeping less (35 min) • Pt 3. Watch, reflect, and apply! (30 min) • Hw 1. Busting sleep myths (5 min) • Hw 2. Stretching and toning (15 min) • Hw 3. One more reason to get good sleep (10 min)
Module 6	Firming and stabilizing your practice • Pt 1. Deepening your meditation practices (45 min) • Pt 2. More on breath awareness (40 min) • Pt 3. Putting it into practice! (30 min) • Hw 1. Standing movements (15 min) • Hw 2. Choiceless awareness practice (15 min)	All about stress • Pt 1. Introduction to stress (45 min) • Pt 2. Coping with stress (25 min) • Pt 3. Watch, reflect, and apply! (35 min) • Hw 1. Why you need stress in your life (15 min) • Hw 2. Stretching and toning (15 min) • Hw 3. Mindfulness to reduce stress (5 min)

XSL•FO
RenderX

	iMBSR	iLifeEd
Module 7	Kindness toward self and others • Pt 1. Starting with foundational practices (45 min) • Pt 2. Practicing loving-kindness meditation (40 min) • Pt 3. Putting it into practice! (30 min) • Hw 1. Mindful yoga (15 min) • Hw 2. Any 15-minute practice (15 min)	All about diet • Pt 1. Introduction to diet – Focus on nutrition (45 min) • Pt 2. Introduction to diet – Focus on hydration (35 min) • Pt 3. Watch, reflect, and apply! (35 min) • Hw 1. Why we can't stop eating unhealthy foods (15 min) • Hw 2. Stretching and toning (15 min) • Hw 3. Dehydration with aging (5 min)
Module 8	Mindfulness: a lifelong practice! • Pt 1. Returning to the basics (30 min) • Pt 2. Final closing (50 min) • Pt 3. Building your personalized mindfulness practice (30 min) • Hw 1. Any 15-minute practice (15 min) • Hw 2. Any 15-minute practice (15 min)	All about social support • Pt 1. Introduction to social engagement and social support (30 min) • Pt 2. Social engagement – Risks and practical strategies (20 min) • Pt 3. Watch, reflect, and apply! (45 min) • Hw 1. What makes a good life? (15 min) • Hw 2. Stretching and toning (15 min) • Hw 3. The value of human connection (15 min)

[a]Pt: part.

[b]Hw: homework.

Primary Outcomes

Feasibility will be determined by evaluating the recruitment and retention of study participants, with high feasibility defined as the successful recruitment and enrollment of 60 participants and an attrition rate of no greater than 25%. We will also examine changes in SCD, with reductions on the ECog Questionnaire considered as evidence for feasibility of the intervention. Acceptability will be measured using the Acceptability of Intervention Measure [63], with high acceptability defined as an average score greater than 3 (on a 1‐5 scale, ranging from "completely agree" to "completely disagree"). Program satisfaction will be measured by our custom Postintervention Acceptability Questionnaire, which focuses on program enjoyment, satisfaction, and potential barriers. High program satisfaction will be defined as an average score greater than 5 (on a 0‐10 scale, ranging from "not at all" to "extremely" scale) for items focused on satisfaction. Participant attendance will also factor into acceptability, with high participation defined as at least 50% of participants successfully completing 6 or more modules on the class platform Scarlet Canvas.

Secondary Outcomes: Plasma-Based Biomarkers of Amyloid and Tau

Blood draws are scheduled during the behavioral session. At each time point, we will collect 20 mL of blood (15 mL for plasma and 5 mL for serum). Following standard operating procedures, blood will be drawn via venipuncture from each participant into one 10 mL and one 5 mL lavender-top ethylenediaminetetraacetic acid tubes and one 5 mL red-top serum tube. Immediately after collecting blood, each ethylenediaminetetraacetic acid tube will be gently mixed by inverting 8 to 10 times to ensure proper mixing of blood and anticoagulant and kept at 4 °C. The red-top tube will also be gently mixed by inverting it 5 times and placed upright at room temperature for 30 minutes to allow for complete clot formation. Furthermore, within 24 hours of acquisition, all samples will be centrifuged for 10 minutes at 2000×*g* at 4 °C, and one-third of the aliquots will be shipped to our collaborators at the University of California, Irvine, for analysis. In this study, the secondary outcome variables are the plasma $A\beta_{42}/A\beta_{40}$ ratio, used to quantify amyloid pathology, and plasma p-tau217, used to determine tau pathology. Plasma samples will be analyzed using the automated Single Molecule Array (Simoa) HD-X analyzer with commercially available kits from Quanterix. A total of 100 µL of plasma will be used for the tau assays (p-tau217), and an additional 100 µL will be used for the 4-Plex panel.

Other Outcomes

Behavioral and Imaging Measures of Sustained Attention and Mind-Wandering

Overview

Laboratory-based assessments of sustained attention and mind-wandering will include 2 behavioral tasks—the Shapes Expectation task [64,65], designed to place minimal demands on the participants, and the Go/No-Go task, a standardized task of sustained attention with embedded thought probes assessing mind-wandering. Participants will additionally complete the Gradual-Onset Continuous Performance Task in the MRI scanner.

Shapes Expectation Task

In this low-effort task, participants will simply be asked to focus on brightly colored geometric shapes on a computer screen. They will not be asked to remember the stimuli or respond to them. Following the presentation of the stimulus, a thought

XSL•FO
RenderX

probe will appear on the screen instructing participants to report on the content of the thoughts they experienced while viewing the shape. This "thinking-aloud" technique in the context of a low-effort perceptual task has been shown to elicit a wide range of task-unrelated thoughts as well as to engage the activity of the 2 canonical networks critically implicated in mind-wandering—the default mode and the frontoparietal networks [64,65]. Participants will be provided with training examples before the commencement of the task. There will be a total of 9 trials with trials being of a short duration (≤20 s), a medium duration (30 - 60 s), and a long duration (≥90 s). There will be no time limit imposed for the thought probes, and all responses will be recorded via a dictaphone.

Go/No-Go Task

In this task, participants will be presented with one of 2 visual stimuli ("X" and "M" or "Z" and "/") for 750 ms, followed by an intertrial interval of 750 ms. They will be asked to press the corresponding key on frequent Go trials and to withhold responses on the less frequent (occurring 10% of the time) No-Go trials, which will be signaled by an auditory tone. Participants will complete 6 blocks of this task, with each block containing 54 Go trials, 6 No-Go trials, and 3 mind-wandering probes, resulting in a total of 18 mind-wandering probes. The quasi-randomly presented mind-wandering probes will ask participants to respond to categorize their immediately preceding thought as (1) "on task"; (2) "thinking about performance on the task"; or (3) "thinking about personal worries, daydreaming, fantasizing, or just lost in thought."

Gradual-Onset Continuous Performance Task

In the MRI scanner, participants view a sequence of gradually appearing mountains and city scenes. They are asked to respond to frequently occurring city scenes (~90%) but to withhold their responses to mountain scenes [66,67]. Scenes transition gradually over 800 ms, and each of the 2 runs includes 9 blocks of varying durations (short: 44 s, medium: 52 s, and long: 60 s). At the end of each block, following the last trial, the scene transitions into a scrambled image, after which 2 thought probes are presented: (1) "To what degree was your focus just on the task or on something else?" and (2) "To what degree were you aware of your focus?" For the first probe, participants respond on a graded scale ranging from only else to only task, and for the second probe, from unaware to aware.

Ambulatory Assessment of Mind-Wandering and Cognitive Variability

In alignment with recent recommendations for EMA-based studies [68], this study will collect EMA data 5 times per day over a 5-day period using study-provided smartphones. These assessments will be conducted before and after training. Participants will be provided with study-specific Android smartphones equipped with the MetricWire platform for EMA data collection, and they will undergo individualized training on EMA procedures. This training will help ensure participants can readily use the study application, understand the different measurements fully, and respond accurately to the EMA probes. The EMA platform will be programmed to signal participants with an audible beep and a pop-up notification at semirandom intervals throughout the day, structured into 5 time bins across waking hours. Mobile-based cognitive assessments of processing speed, working memory, and inhibitory control will include 3 tasks taken from the Mobile Monitoring of Cognitive Change (M2C2) battery [69]. The cognitive tasks will be completed immediately after each EMA report of mind-wandering.

Mind-Wandering Prompts

Overview

Consistent with current best practice in EMA mind-wandering research [68,70,71], each signal will prompt participants to provide information about mind-wandering and contextual details from the moment they were signaled. At each signal, participants will be presented with the following probes: (1) In the moments just before the notification, to what degree had your mind wandered to something other than what you were doing? (2) In the moments just before the notification, to what degree were you actively directing what was on your mind? (3) In the moments just before the notification, to what degree were you aware of what was on your mind? and (4) In the moments just before the notification, to what degree were your thoughts moving freely? For each of these prompts, responses will be collected on a 5-point Likert-type scale, ranging from not at all (0) to very much so (4), with intermediate options of a little, somewhat, and quite a bit. Our primary variable will be the overall rate of mind-wandering in daily life.

Symbol Search Task

This is a speeded performance task, measuring processing speed and attention. Participants will be presented with a set of 3 symbol pairs located on the upper half of the screen and 2 symbol pairs on the lower half of the screen. They are instructed to select which of the bottom symbol pairs matches 1 of the 3 symbol pairs above, as quickly and accurately as possible. A total of 20 trials with an intertrial interval of 200 ms will be administered, and each set of stimuli will be presented until the participant selects a response. The SD of response times will be computed, with higher scores reflecting higher processing speed variability.

Grid Memory Task

This task is a delayed short-term recall task and measures visual working memory abilities. First, participants complete a 3-second encoding phase in which they are presented with a 5×5 grid containing 3 red dots. Participants are instructed to remember the locations of the 3 red dots. Second, participants complete an 8-second distraction phase where they are presented with a 5×6 grid of "F" letters among an array of "E" letters and are directed to select all the "F" letters as quickly as possible. Finally, they complete a retrieval phase where they are presented with an empty 5×5 grid and must indicate the location of the 3 red dots; this grid will be displayed until the participant submits their selection. A total of 4 trials will be administered. The SD of the Euclidean distance of the location of the incorrect dot to the correct dot location will be computed, with higher scores indicating more variable placement.

Letter Go/No-Go Task

This task is a continuous performance task measuring attention and inhibitory control. Participants will be presented with a

XSL•FO
RenderX

series of 30 letters, one at a time, and are instructed to tap a button as quickly as possible when they are presented with each letter. However, when the participant is presented with the letter "X," they are instructed to withhold their response. Each letter will be presented for 1000 ms with an intertrial interval of 500 ms, and the ratio of Go/No-Go trials is 5:1.

Data Management and Analysis

Our laboratory is dedicated to conducting rigorous and reproducible randomized clinical trials that evaluate the feasibility and efficacy of lifestyle-based psychosocial interventions. We have established detailed procedures for data collection, reduction, storage, and analysis, and these protocols will be followed in this study.

All data collected as part of this study, including behavioral, self-report, and neuroimaging data, will be stored in the PI's laboratory, accessible only to authorized laboratory personnel. Blood samples will be stored at the Clinical Research Center, and deidentified EMA data will be housed on MetricWire's secure cloud server. For screening and behavioral assessments, 2 independent raters will code the measures (either manually or via custom scripts), with a third rater confirming agreement. Questionnaire data will be coded using custom scripts that have been independently reviewed by 2 laboratory members and validated against manual coding. EMA data will be processed using tools provided by MetricWire and subsequently analyzed with custom scripts.

Structural MRI and functional MRI (fMRI) preprocessing will be performed using the standardized fMRIPrep (version 1.4.1) pipeline [72]. T1-weighted (T1w) images will be corrected for intensity nonuniformity [73], skull-stripped (with brain surfaces reconstructed using FreeSurfer, version 6.0.1) [74], and spatially normalized to the Montreal Neurological Institute (MNI) space through nonlinear registration [75]. Functional data will be corrected for susceptibility distortion [76] and motion [77] and co-registered to the T1w using boundary-based registration [78]. Motion-correcting transformations, BOLD-to-T1w transformations, and T1w-to-template (MNI) warps will be concatenated and applied in a single step using Lanczos interpolation (ANTs, version 2.1.0). Physiological regressors [79], framewise displacement [80], and global signal will be extracted from the fMRI data. Preprocessed fMRI data will then be denoised using the signal clean function within Nilearn [81]. Potentially confounding sources of signal variance will be removed: mean CSF, white matter, global signal, a 24-parameter motion model, and the linear trend. To further control for motion-related confounds, we will add a spike regressor for each frame with a framewise displacement of ≥0.5 mm [80]. A high-pass temporal filtering at 0.01 Hz will also be applied.

Preprocessed fMRI data will then be parcellated using the Glasser functional atlas with 360 cortical regions and 19 subcortical regions [82,83]. The functional atlas will be transformed from MNI space to each participant's native space. We will then extract the BOLD signal time courses from each of the 379 regions by averaging the signal across all voxels within each region. To compute the statistical dependence between BOLD signal across 2 nodes, we will first z-score each time series, compute their element-wise product, and then average this product to obtain the Pearson correlation coefficient between each pair of nodes. Calculating the Pearson correlation for all pairs of nodes will result in a 379×379 functional connectivity matrix for each participant, which will be used to compute the DMN connectivity.

Following the preprocessing of self-report, behavioral, and neuroimaging data, analyses will be conducted to determine intervention feasibility and acceptability. For preliminary effects on plasma biomarkers of AD pathology, behavioral and neural metrics of mind-wandering, and other secondary outcomes, we will fit linear mixed models to account for within-subject correlations arising from measuring each subject at 2 time points. In these models, emphasis will be placed on estimating effect sizes and CIs, rather than on hypothesis testing. As we are including adults with SCD based on self-reported concerns, we will also conduct sensitivity analyses using the total score on the Penn State Worry Questionnaire [58] to examine whether feasibility, acceptability, and preliminary effects differ as a function of worry scores within the current sample.

Adverse Event Reporting

Adverse events (AEs) during the clinical trial will be reported in accordance with a detailed protocol developed in alignment with the NIH Adverse Event Reporting Guidelines. Any event brought to the study team's attention will be immediately reported to the study PI and the clinical trial safety monitor. The safety monitor, a cognitive neurologist with extensive expertise in neurocognitive disorders, will evaluate the event and determine its relationship to the study. Each event will be classified as unrelated, possibly related, probably related, or definitely related using standard criteria for clinical trials.

For all events deemed possibly, probably, or definitely related to the study, they will then be categorized by severity as (1) mild AE—does not require treatment; (2) moderate AE—resolved with treatment; (3) severe AE—inability to carry on normal activities and requires professional medical attention; (4) life-threatening or disabling AE; and (5) fatal AE. For mild AEs, we will document the event, and the study team will determine whether changes to the study procedures are warranted. For moderate or more severe events that are related to the study, a detailed report will be prepared and submitted to the institutional review board (IRB) and the NIH Program Officer assigned to the study. If 2 participants experience the same AE, recruitment will be stopped, a report will be submitted to the IRB and NIH, and any revisions to the study protocol suggested by either the IRB or the NIH will be implemented. In addition, annual reports to the IRB and progress reports to the NIH will be submitted, including information on enrollments, withdrawals, completions, and the occurrence of AE.

Ethical Considerations

The Ohio State Institutional Review Board has granted approval to conduct this study (IRB 2024H0401). All participants will provide written informed consent that details the aims of the study, outlines the elements of the study protocol, and describes the risks and benefits associated with participation. Participants are informed both verbally and through the written consent

XSL•FO
RenderX

document that participation is voluntary in nature and that they may withdraw from the study at any time without penalty. All study personnel are trained in the ethical conduct of research and are explicitly prohibited from using any kind of coercion during study participation. Participants receive compensation at a rate of US $15 per hour for the in-person screening session, behavioral sessions, and MRI sessions (both before and after assessment). For EMA, participants receive US $10 per day for completing at least 4 of the 5 daily prompts (up to US $50 before and after assessment). Thus, for completing all pre- and post-assessments, participants can receive up to US $310. Enrolled participants can also receive additional compensation for referring others to participate in the study (US $5 per referral, up to US $35). Intervention sessions are provided to participants at no cost.

Although identifying information is collected for the study, where possible, data are collected using coded IDs. All data are stored on a secure cloud server housed at The Ohio State University. Upon study completion, all data will be fully deidentified using the Safe Harbor approach, and only coded data will be retained for study analyses.

Results

This study was funded by the National Institute on Aging in April 2024. Phase I focused on iterative development of the 2 programs based on focus group feedback. Recruitment for the Phase II RCT (iMBT) began on June 13, 2025, and is ongoing. As of December 1, 2025, we have 31 participants randomized into the study: 15 participants have been randomized to the iMBSR group, and 16 participants have been randomized to the iLifeEd group. We anticipate that recruitment will end in September 2026, with all 60 participants randomized by that date. Although data analyses will begin after all participants have completed the intervention and assessment sessions, all analysis scripts have been finalized, and data reduction and processing are ongoing. We expect to complete the processing and analysis of the primary outcome variables by March 2027, with a targeted manuscript submission date of June 2027.

Discussion

AD is a growing public health concern, particularly as targeted pharmacological interventions have not demonstrated their expected effects. Preventative interventions, targeting at-risk individuals when pathological processes are only beginning to manifest, have the potential to stave off the functional impairments associated with AD neurodegeneration. In this Stage I RCT, iMBT, we aim to systematically test the feasibility, acceptability, and preliminary effects of an internet-based mindfulness program. Sixty adults with SCD, a well-established late preclinical phase of AD, will be recruited. Participants will asynchronously complete 8 modules of either the iMBSR or the iLifeEd program, with behavioral and neuroimaging assessments conducted before and immediately after the intervention. Our main hypothesis is that iMBSR, carefully curated to include the active components of the MBSR program and designed to be self-paced and online, will be feasible and acceptable to adults at higher risk for developing AD and related dementias.

MBSR is a manualized 8-week program that provides systematic and incremental training in the principles, attitudinal foundations, and practices of mindfulness meditation [84,85]. These programs are traditionally taught in group-based settings with weekly 2.5-hour sessions. Despite its demonstrated clinical utility, the in-person format poses significant accessibility challenges, resulting in a growing need for more widely available adaptations [86]. Internet-based mindfulness programs offer a promising and cost-effective solution and have been rated as a preferred delivery format among the general population [86]. Several internet-based interventions, ideographically tailored for the target population, including acquired brain injury [87], fibromyalgia [88], and heart disease [89], show promising support. However, key methodological limitations necessitate the systematic testing of the iMBSR and iLifeEd protocols developed for this trial.

First, much of the extant literature on web-based mindfulness interventions lacks active control conditions, thus failing to account for nonspecific factors and expectancy effects [22]. For example, in one of the largest internet-delivered mindfulness trials, 150 individuals with cancer reported reductions in self-reported anxiety and depression [90]. However, because comparisons were made against a wait-list control group, it is unclear whether improvements reflected mindfulness practices, demand characteristics, or even placebo effects. Second, the majority of existing internet-based mindfulness training programs are based on mindfulness-based cognitive therapy [91-93] and acceptance and commitment therapy [94-96], or they simply teach participants how to meditate [97,98]. In contrast, surprisingly few studies have adapted MBSR for internet delivery.

Of the existing internet MBSR adaptations, most are synchronous formats [87,99,100], whereas only one study, to our knowledge, has developed an asynchronous internet-based MBSR program [101]. In that study, participants completed 30 minutes of class material per week compared to the 2.5 hours common among standard in-person programs, thus questioning the treatment fidelity of this program [102]. Although synchronous formats more closely resemble traditional in-person mindfulness-based programs in terms of course content and class time, they fail to offer participants the flexibility to complete components on their personal schedule. Asynchronous formats may be particularly advantageous for individuals unable to commit to predetermined group meeting times, as is common among midlife adults (a sample included in this RCT). Finally, only approximately half of internet-based interventions report on participant adherence [103], limiting conclusions about their efficacy in targeting key outcomes. In fact, in a systematic review of all psychosocial wellness and stress management applications, only 2.08% (21 of 1009) mobile apps had published, peer-reviewed evidence of feasibility and/or efficacy [104]. There is thus a critical need to develop and rigorously test an internet-based mindfulness program that is derived from the standardized 8-week MBSR protocol, designed for the target population (here, adults aged ≥50 years with SCD), including an active control comparison matched for nonspecific factors,

XSL•FO
RenderX

and has in-built elements for increasing and maintaining adherence. Leveraging our team's expertise in delivering in-person mindfulness programs and designing mobile and web-based apps for the delivery of health-based interventions [24,105,106], the iMBT trial incorporates state-of-the-art research elements into an internet-delivered, guided, self-paced mindfulness program for adults at risk for AD. We have also designed a comparable control program, created by the same team and focus group participants, against which we will compare the iMBSR program in future studies to ensure internal validity before making causal claims about efficacy.

In addition to examining feasibility and acceptability, this study will examine the preliminary effects of mindfulness practice on amyloid and tau pathology. The deposition of these proteins in the densely connected nodes of the DMN is a characteristic hallmark of AD [107,108]. Early aggregation of amyloid has been observed in anterior midcortical structures, such as the medial orbitofrontal gyrus and the anterior cingulate cortex, as well as in posterior midcortical structures, including the posterior cingulate cortex and the precuneus. Tau pathology is known to spread in a prion-like manner, beginning in the transentorhinal cortex and progressing to the perirhinal cortex, hippocampus, limbic areas, and, eventually, the association cortices [109-112]. The spatial affinities of Aβ and tau-based tangles for DMN regions, which also appear to support the prophylactic benefits of mindfulness meditation [26-31], provide the scientific rationale for this RCT. We hypothesize that mindfulness training will facilitate present moment awareness and reduce mind-wandering in daily life, thereby strengthening DMN connectivity and exerting downstream cascading effects on amyloid and tau pathology.

AD pathology in this study will be quantified using the recently validated plasma-based biomarkers of Aβ and tau [113-117]. Emerging evidence demonstrates that low plasma Aβ ($A\beta_{42}/A\beta_{40}$) levels are associated with higher brain amyloid deposition as measured using PET imaging [118-120]. Although high levels of p-tau217 show a positive association with amyloid PET data, growing evidence also suggests that plasma p-tau217 is even more closely associated with brain tau deposition, particularly in later disease stages [115,121,122]. Importantly, these plasma-based measures were recently validated in a multiethnic community study of 300 participants that included 46% non-Hispanic White participants and 54% participants from other ethnoracial groups [123]. The iMBT clinical trial will provide the first test of whether online mindfulness training can reduce amyloid and tau pathology in adults at risk for AD.

In conclusion, the iMBT trial is designed to test the feasibility and acceptability of 2 asynchronous, online mind-body training programs. Both interventions were systematically adapted from in-person formats with demonstrated high adherence rates [62] and further refined for online delivery using instructional design principles and feedback from the target population. Findings from this trial will provide critical groundwork for disseminating these accessible interventions to adults reporting everyday cognitive concerns. Given the limited availability of evidence-based preventative programs, these interventions have the potential to expand access to scalable, mind-body interventions for adults at risk for cognitive impairment.

Acknowledgments

We acknowledge the contributions of videographer Mr Joshua Howey for recording all intervention videos curated for the 2 study groups. We also thank yoga instructors Ms Anabelle McDonald and Ms Valerie Williams for thoughtfully developing the low-intensity mindful yoga videos used in the interventions. We further acknowledge the Center for Cognitive and Behavioral Brain Imaging at the College of Arts & Sciences for providing essential support for the neuroimaging component of the study and the Clinical Research Center at the College of Medicine for their assistance in the collection, processing, and storage of blood samples. We acknowledge the use of ChatGPT (OpenAI) for proofreading and minimal language copyediting. Microsoft Office's built-in editing tools (spelling and grammar checks) were also used to proofread the manuscript. Generative artificial intelligence was not used in this manuscript for scientific content generation or conceptualization, literature review and/or synthesis, power analysis, or text generation.

Funding

This work was supported by the National Institute on Aging of the National Institutes of Health (R61AG081982-01A1 awarded to RSP). The study protocol has undergone full external peer review by the funding body. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

Conflicts of Interest

None declared.

References

1. Alzheimer's Association. 2023 Alzheimer's disease facts and figures. Alzheimer's & Dementia 2023 Apr;19(4):1598-1695. [doi: 10.1002/alz.13016]

XSL•FO
RenderX

2. Buckley RF, Villemagne VL, Masters CL, et al. A conceptualization of the utility of subjective cognitive decline in clinical trials of preclinical Alzheimer's disease. J Mol Neurosci 2016 Nov;60(3):354-361. [doi: 10.1007/s12031-016-0810-z] [Medline: 27514526]
3. Jessen F, Amariglio RE, van Boxtel M, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. Alzheimers Dement 2014 Nov;10(6):844-852. [doi: 10.1016/j.jalz.2014.01.001] [Medline: 24798886]
4. Tomaszewski Farias S, Mungas D, Harvey DJ, Simmons A, Reed BR, Decarli C. The measurement of everyday cognition: development and validation of a short form of the Everyday Cognition scales. Alzheimers Dement 2011 Nov;7(6):593-601. [doi: 10.1016/j.jalz.2011.02.007] [Medline: 22055976]
5. Rami L, Mollica MA, García-Sanchez C, et al. The Subjective Cognitive Decline Questionnaire (SCD-Q): a validation study. J Alzheimers Dis 2014;41(2):453-466. [doi: 10.3233/JAD-132027] [Medline: 24625794]
6. Rattanabannakit C, Risacher SL, Gao S, et al. The Cognitive Change Index as a measure of self and informant perception of cognitive decline: relation to neuropsychological tests. Journal of Alzheimer's Disease 2016 Apr 12;51(4):1145-1155. [doi: 10.3233/JAD-150729]
7. Amariglio RE, Mormino EC, Pietras AC, et al. Subjective cognitive concerns, amyloid-β, and neurodegeneration in clinically normal elderly. Neurology (ECronicon) 2015 Jul 7;85(1):56-62. [doi: 10.1212/WNL.0000000000001712]
8. Lin Y, Shan PY, Jiang WJ, Sheng C, Ma L. Subjective cognitive decline: preclinical manifestation of Alzheimer's disease. Neurol Sci 2019 Jan;40(1):41-49. [doi: 10.1007/s10072-018-3620-y] [Medline: 30397816]
9. Verfaillie SCJ, Slot RE, Tijms BM, et al. Thinner cortex in patients with subjective cognitive decline is associated with steeper decline of memory. Neurobiol Aging 2018 Jan;61:238-244. [doi: 10.1016/j.neurobiolaging.2017.09.009] [Medline: 29029762]
10. Saykin AJ, Wishart HA, Rabin LA, et al. Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. Neurology (ECronicon) 2006 Sep 12;67(5):834-842. [doi: 10.1212/01.wnl.0000234032.77541.a2]
11. Cherbuin N, Sargent-Cox K, Easteal S, Sachdev P, Anstey KJ. Hippocampal atrophy is associated with subjective memory decline: the PATH Through Life study. Am J Geriatr Psychiatry 2015 May;23(5):446-455. [doi: 10.1016/j.jagp.2014.07.009] [Medline: 25204687]
12. Dhana A, DeCarli C, Dhana K, et al. Association of subjective memory complaints with white matter hyperintensities and cognitive decline among older adults in Chicago, Illinois. JAMA Netw Open 2022 Apr 1;5(4):e227512. [doi: 10.1001/jamanetworkopen.2022.7512] [Medline: 35426922]
13. Verfaillie SCJ, Timmers T, Slot RER, et al. Amyloid-β load is related to worries, but not to severity of cognitive complaints in individuals with subjective cognitive decline: the SCIENCe project. Front Aging Neurosci 2019;11:7. [doi: 10.3389/fnagi.2019.00007] [Medline: 30760996]
14. Perrotin A, La Joie R, de La Sayette V, et al. Subjective cognitive decline in cognitively normal elders from the community or from a memory clinic: differential affective and imaging correlates. Alzheimer's & Dementia 2017 May;13(5):550-560. [doi: 10.1016/j.jalz.2016.08.011]
15. Wolfsgruber S, Jessen F, Koppara A, et al. Subjective cognitive decline is related to CSF biomarkers of AD in patients with MCI. Neurology (ECronicon) 2015 Mar 24;84(12):1261-1268. [doi: 10.1212/WNL.0000000000001399]
16. van der Flier WM, Scheltens P. Amsterdam dementia cohort: performing research to optimize care. Journal of Alzheimer's Disease 2018 Mar 13;62(3):1091-1111. [doi: 10.3233/JAD-170850]
17. Garcia-Ptacek S, Cavallin L, Kåreholt I, et al. Subjective cognitive impairment subjects in our clinical practice. Dement Geriatr Cogn Disord Extra 2014 Nov 11;4(3):419-430. [doi: 10.1159/000366270]
18. Kraemer KM, Luberto CM, Hall DL, Ngo LH, Yeh GY. A systematic review and meta-analysis of mindfulness- and acceptance-based interventions for affect intolerance/sensitivity. Behav Res Ther 2020 Dec;135:103746. [doi: 10.1016/j.brat.2020.103746]
19. Prakash RS. Mindfulness meditation: impact on attentional control and emotion dysregulation. Arch Clin Neuropsychol 2021 Oct 13;36(7):1283-1290. [doi: 10.1093/arclin/acab053] [Medline: 34651648]
20. Zainal NH, Newman MG. Mindfulness enhances cognitive functioning: a meta-analysis of 111 randomized controlled trials. Health Psychol Rev 2024 Jun;18(2):369-395. [doi: 10.1080/17437199.2023.2248222] [Medline: 37578065]
21. Kabat-Zinn J. Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life: Hyperion; 1994.
22. Prakash RS, Fountain-Zaragoza S, Kramer AF, Samimy S, Wegman J. Mindfulness and attention: current state-of-affairs and future considerations. J Cogn Enhanc 2020 Sep;4(3):340-367. [doi: 10.1007/s41465-019-00144-5] [Medline: 33817547]
23. Whitmoyer P, Fountain-Zaragoza S, Andridge R, et al. Mindfulness training and attentional control in older adults: a randomized controlled trial. Mindfulness (N Y) 2020 Jan;11(1):203-218. [doi: 10.1007/s12671-019-01218-3]
24. Prakash RS, Fountain-Zaragoza S, Fisher M, et al. Protocol for a randomized controlled trial of mindfulness-based stress reduction to improve attentional control in older adults (HealthyAgers trial). BMC Geriatr 2022 Aug 13;22(1):666. [doi: 10.1186/s12877-022-03334-7] [Medline: 35964000]
25. Feruglio S, Matiz A, Pagnoni G, Fabbro F, Crescentini C. The impact of mindfulness meditation on the wandering mind: a systematic review. Neurosci Biobehav Rev 2021 Dec;131:313-330. [doi: 10.1016/j.neubiorev.2021.09.032] [Medline: 34560133]

XSL•FO
RenderX

26. Shaurya Prakash R, De Leon AA, Klatt M, Malarkey W, Patterson B. Mindfulness disposition and default-mode network connectivity in older adults. Soc Cogn Affect Neurosci 2013 Jan;8(1):112-117. [doi: 10.1093/scan/nss115] [Medline: 23051900]
27. Hasenkamp W, Wilson-Mendenhall CD, Duncan E, Barsalou LW. Mind wandering and attention during focused meditation: a fine-grained temporal analysis of fluctuating cognitive states. Neuroimage 2012 Jan 2;59(1):750-760. [doi: 10.1016/j.neuroimage.2011.07.008] [Medline: 21782031]
28. Brewer JA, Worhunsky PD, Gray JR, Tang YY, Weber J, Kober H. Meditation experience is associated with differences in default mode network activity and connectivity. Proc Natl Acad Sci USA 2011 Dec 13;108(50):20254-20259. [doi: 10.1073/pnas.1112029108]
29. Yang CC, Barrós-Loscertales A, Li M, et al. Alterations in brain structure and amplitude of low-frequency after 8 weeks of mindfulness meditation training in meditation-naïve subjects. Sci Rep 2019 Jul 29;9(1):10977. [doi: 10.1038/s41598-019-47470-4] [Medline: 31358842]
30. Bauer CCC, Rozenkrantz L, Caballero C, et al. Mindfulness training preserves sustained attention and resting state anticorrelation between default-mode network and dorsolateral prefrontal cortex: a randomized controlled trial. Hum Brain Mapp 2020 Dec 15;41(18):5356-5369. [doi: 10.1002/hbm.25197] [Medline: 32969562]
31. Kral TRA, Imhoff-Smith T, Dean DC III, et al. Mindfulness-based stress reduction-related changes in posterior cingulate resting brain connectivity. Soc Cogn Affect Neurosci 2019 Jul 31;14(7):777-787. [doi: 10.1093/scan/nsz050]
32. Wells RE, Yeh GY, Kerr CE, et al. Meditation's impact on default mode network and hippocampus in mild cognitive impairment: a pilot study. Neurosci Lett 2013 Nov 27;556:15-19. [doi: 10.1016/j.neulet.2013.10.001] [Medline: 24120430]
33. Strikwerda-Brown C, Ozlen H, Pichet Binette A, et al. Trait mindfulness is associated with less amyloid, tau, and cognitive decline in individuals at risk for Alzheimer's disease. Biol Psychiatry Glob Open Sci 2023 Jan;3(1):130-138. [doi: 10.1016/j.bpsgos.2022.01.001] [Medline: 36712573]
34. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
35. Karikari TK, Pascoal TA, Ashton NJ, et al. Blood phosphorylated tau 181 as a biomarker for Alzheimer's disease: a diagnostic performance and prediction modelling study using data from four prospective cohorts. Lancet Neurol 2020 May;19(5):422-433. [doi: 10.1016/S1474-4422(20)30071-5] [Medline: 32333900]
36. West T, Kirmess KM, Meyer MR, et al. A blood-based diagnostic test incorporating plasma Aβ42/40 ratio, ApoE proteotype, and age accurately identifies brain amyloid status: findings from a multi cohort validity analysis. Mol Neurodegener 2021 May 1;16(1):30. [doi: 10.1186/s13024-021-00451-6] [Medline: 33933117]
37. Smirnov DS, Ashton NJ, Blennow K, et al. Plasma biomarkers for Alzheimer's disease in relation to neuropathology and cognitive change. Acta Neuropathol 2022 Apr;143(4):487-503. [doi: 10.1007/s00401-022-02408-5] [Medline: 35195758]
38. Onken LS, Carroll KM, Shoham V, Cuthbert BN, Riddle M. Reenvisioning clinical science: unifying the discipline to improve the public health. Clin Psychol Sci 2014 Jan 1;2(1):22-34. [doi: 10.1177/2167702613497932] [Medline: 25821658]
39. Campbell M, Fitzpatrick R, Haines A, et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 Sep 16;321(7262):694-696. [doi: 10.1136/bmj.321.7262.694] [Medline: 10987780]
40. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. Int J Nurs Stud 2013 May;50(5):587-592. [doi: 10.1016/j.ijnurstu.2012.09.010] [Medline: 23159157]
41. Tickle-Degnen L. Nuts and bolts of conducting feasibility studies. Am J Occup Ther 2013;67(2):171-176. [doi: 10.5014/ajot.2013.006270] [Medline: 23433271]
42. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 6;10(1):1. [doi: 10.1186/1471-2288-10-1]
43. Jessen F, Amariglio RE, Buckley RF, et al. The characterisation of subjective cognitive decline. Lancet Neurol 2020 Mar;19(3):271-278. [doi: 10.1016/S1474-4422(19)30368-0]
44. Filshtein T, Chan M, Mungas D, et al. Differential item functioning of the Everyday Cognition (ECog) scales in relation to racial/ethnic groups. J Int Neuropsychol Soc 2020 May;26(5):515-526. [doi: 10.1017/S1355617719001437]
45. Nuño MM, Gillen DL, Grill JD, Alzheimer's Disease Cooperative Study. Study partner types and prediction of cognitive performance: implications to preclinical Alzheimer's trials. Alzheimers Res Ther 2019 Nov 27;11(1):92. [doi: 10.1186/s13195-019-0544-6] [Medline: 31775871]
46. Stites SD, Largent EA, Gill J, Gurian A, Harkins K, Karlawish J. Predictors of who serves as an Alzheimer's disease research participant's study partner and the impact of their relationship on study partners' reports on participants. Res Aging 2022;44(9-10):734-746. [doi: 10.1177/01640275221075739] [Medline: 35324354]
47. van Harten AC, Mielke MM, Swenson-Dravis DM, et al. Subjective cognitive decline and risk of MCI: the Mayo Clinic Study of Aging. Neurology (ECronicon) 2018 Jul 24;91(4):e300-e312. [doi: 10.1212/WNL.0000000000005863] [Medline: 29959257]
48. Kang JM, Manjavong M, Jin C, et al. Subjective cognitive decline predicts longitudinal neuropsychological test performance in an unsupervised online setting in the Brain Health Registry. Alzheimers Res Ther 2025 Jan 7;17(1):10. [doi: 10.1186/s13195-024-01641-2] [Medline: 39773247]

XSL•FO
RenderX

49. Katz MJ, Wang C, Nester CO, et al. T - MoCA: a valid phone screen for cognitive impairment in diverse community samples. Alz & Dem Diag Ass & Dis Mo 2021 Jan;13(1):e12144. [doi: 10.1002/dad2.12144]
50. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [doi: 10.1093/geront/9.3_Part_1.179] [Medline: 5349366]
51. Weintraub S, Besser L, Dodge HH, et al. Version 3 of the Alzheimer Disease Centers' Neuropsychological Test Battery in the Uniform Data Set (UDS). Alzheimer Dis Assoc Disord 2018;32(1):10-17. [doi: 10.1097/WAD.0000000000000223] [Medline: 29240561]
52. Matusz EF, Fiala J, Kiselica AM, et al. Cognitive factor structure of the NACC UDS-3 neuropsychological battery across ethno-racial, linguistic, and cognitive status groups. Clin Neuropsychol 2025 Oct 24:1-23. [doi: 10.1080/13854046.2025.2576154] [Medline: 41131969]
53. Bondi MW, Edmonds EC, Jak AJ, et al. Neuropsychological criteria for mild cognitive impairment improves diagnostic precision, biomarker associations, and progression rates. J Alzheimers Dis 2014;42(1):275-289. [doi: 10.3233/JAD-140276] [Medline: 24844687]
54. Jak AJ, Bondi MW, Delano-Wood L, et al. Quantification of five neuropsychological approaches to defining mild cognitive impairment. Am J Geriatr Psychiatry 2009 May;17(5):368-375. [doi: 10.1097/JGP.0b013e31819431d5]
55. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
56. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977:385-401 [FREE Full text] [doi: 10.1177/014662167700100306]
57. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. J Consult Clin Psychol 1988;56(6):893-897. [doi: 10.1037/0022-006X.56.6.893]
58. Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the Penn State Worry Questionnaire. Behav Res Ther 1990;28(6):487-495. [doi: 10.1016/0005-7967(90)90135-6] [Medline: 2076086]
59. The Whoqol Group. Development of the World Health Organization WHOQOL-BREF Quality of Life Assessment. Psychol Med 1998 May;28(3):551-558. [doi: 10.1017/S0033291798006667]
60. Williams DR, Jackson JS, Anderson NB. Racial differences in physical and mental health: socio-economic status, stress and discrimination. J Health Psychol 1997 Jul;2(3):335-351. [doi: 10.1177/135910539700200305] [Medline: 22013026]
61. Kim S, Sargent-Cox K, Cherbuin N, Anstey KJ. Development of the motivation to change lifestyle and health behaviours for dementia risk reduction scale. Dement Geriatr Cogn Disord Extra 2014 May 22;4(2):172-183. [doi: 10.1159/000362228]
62. Fisher M, Phansikar M, Teng J, et al. Effects of mindfulness training on sustained attention and mind-wandering in older adults: a randomized controlled trial. Open Science Framework. Preprint posted online on Apr 30, 2025. [doi: 10.31219/osf.io/5zpyt_v1]
63. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implementation Sci 2017 Dec;12(1):108. [doi: 10.1186/s13012-017-0635-3]
64. O'Callaghan C, Shine JM, Lewis SJG, Andrews-Hanna JR, Irish M. Shaped by our thoughts – a new task to assess spontaneous cognition and its associated neural correlates in the default network. Brain Cogn 2015 Feb;93:1-10. [doi: 10.1016/j.bandc.2014.11.001]
65. O'Callaghan C, Shine JM, Hodges JR, Andrews-Hanna JR, Irish M. Hippocampal atrophy and intrinsic brain network dysfunction relate to alterations in mind wandering in neurodegeneration. Proc Natl Acad Sci USA 2019 Feb 19;116(8):3316-3321. [doi: 10.1073/pnas.1818523116]
66. Esterman M, Noonan SK, Rosenberg M, Degutis J. In the zone or zoning out? Tracking behavioral and neural fluctuations during sustained attention. Cereb Cortex 2013 Nov;23(11):2712-2723. [doi: 10.1093/cercor/bhs261] [Medline: 22941724]
67. Kucyi A, Esterman M, Riley CS, Valera EM. Spontaneous default network activity reflects behavioral variability independent of mind-wandering. Proc Natl Acad Sci U S A 2016 Nov 29;113(48):13899-13904. [doi: 10.1073/pnas.1611743113] [Medline: 27856733]
68. Kawashima I, Hinuma T, Tanaka SC. Ecological momentary assessment of mind-wandering: meta-analysis and systematic review. Sci Rep 2023 Feb 18;13(1):2873. [doi: 10.1038/s41598-023-29854-9] [Medline: 36801931]
69. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and validity of ambulatory cognitive assessments. Assessment 2018 Jan;25(1):14-30. [doi: 10.1177/1073191116643164]
70. Kane MJ, Smeekens BA, Meier ME, Welhaf MS, Phillips NE. Testing the construct validity of competing measurement approaches to probed mind-wandering reports. Behav Res Methods 2021 Dec;53(6):2372-2411. [doi: 10.3758/s13428-021-01557-x] [Medline: 33835393]
71. Christoff K, Irving ZC, Fox KCR, Spreng RN, Andrews-Hanna JR. Mind-wandering as spontaneous thought: a dynamic framework. Nat Rev Neurosci 2016 Nov;17(11):718-731. [doi: 10.1038/nrn.2016.113] [Medline: 27654862]
72. Esteban O, Markiewicz CJ, Blair RW, et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat Methods 2019 Jan;16(1):111-116. [doi: 10.1038/s41592-018-0235-4] [Medline: 30532080]
73. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 2010 Jun;29(6):1310-1320. [doi: 10.1109/TMI.2010.2046908] [Medline: 20378467]

XSL•FO
RenderX

74. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. Neuroimage 1999 Feb;9(2):179-194. [doi: 10.1006/nimg.1998.0395]
75. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 2008 Feb;12(1):26-41. [doi: 10.1016/j.media.2007.06.004]
76. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Computers and Biomedical Research 1996 Jun;29(3):162-173. [doi: 10.1006/cbmr.1996.0014]
77. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 2002 Oct;17(2):825-841. [doi: 10.1006/nimg.2002.1132]
78. Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. Neuroimage 2009 Oct;48(1):63-72. [doi: 10.1016/j.neuroimage.2009.06.060]
79. Behzadi Y, Restom K, Liau J, Liu TT. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 2007 Aug 1;37(1):90-101. [doi: 10.1016/j.neuroimage.2007.04.042] [Medline: 17560126]
80. Power JD, Mitra A, Laumann TO, Snyder AZ, Schlaggar BL, Petersen SE. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 2014 Jan 1;84:320-341. [doi: 10.1016/j.neuroimage.2013.08.048] [Medline: 23994314]
81. Abraham A, Pedregosa F, Eickenberg M, et al. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 2014;8:14. [doi: 10.3389/fninf.2014.00014] [Medline: 24600388]
82. Glasser MF, Coalson TS, Robinson EC, et al. A multi-modal parcellation of human cerebral cortex. Nature New Biol 2016 Aug;536(7615):171-178. [doi: 10.1038/nature18933]
83. Ji JL, Spronk M, Kulkarni K, Repovš G, Anticevic A, Cole MW. Mapping the human brain's cortical-subcortical functional network organization. Neuroimage 2019 Jan 15;185:35-57. [doi: 10.1016/j.neuroimage.2018.10.006] [Medline: 30291974]
84. Kabat-Zinn J. Full Catastrophe Living: The Program of the Stress Reduction Clinic at the University of Massachusetts Medical Center: Delta; 1990.
85. Santorelli SF, Meleo-Meyer F, Koerbel L, Kabat-Zinn J. Mindfulness-based stress reduction (MBSR) authorized curriculum guide © 2017. : Center for Mindfulness in Medicine, Health Care, and Society, University of Massachusetts Medical School; 2017 URL: https://mbsr.website/sites/default/files/docs/mbsr-curriculum-guide-2017.pdf [accessed 2026-03-23]
86. Wahbeh H, Svalina MN, Oken BS. Group, one-on-one, or internet? Preferences for mindfulness meditation delivery format and their predictors. Open Med J 2014;1:66-74. [doi: 10.2174/1874220301401010066] [Medline: 27057260]
87. Johansson B, Bjuhr H, Karlsson M, Karlsson JO, Rönnbäck L. Mindfulness-based stress reduction (MBSR) delivered live on the internet to individuals suffering from mental fatigue after an acquired brain injury. Mindfulness (N Y) 2015 Dec;6(6):1356-1365. [doi: 10.1007/s12671-015-0406-7]
88. Davis MC, Zautra AJ. An online mindfulness intervention targeting socioemotional regulation in fibromyalgia: results of a randomized controlled trial. Ann Behav Med 2013 Dec;46(3):273-284. [doi: 10.1007/s12160-013-9513-7] [Medline: 23670111]
89. Younge JO, Wery MF, Gotink RA, et al. Web-based mindfulness intervention in heart disease: a randomized controlled trial. PLOS One 2015;10(12):e0143843. [doi: 10.1371/journal.pone.0143843] [Medline: 26641099]
90. Nissen ER, O'Connor M, Kaldo V, et al. Internet-delivered mindfulness-based cognitive therapy for anxiety and depression in cancer survivors: a randomized controlled trial. Psychooncology 2020 Jan;29(1):68-75. [doi: 10.1002/pon.5237] [Medline: 31600414]
91. Segal ZV, Williams JMG, Teasdale JD. Mindfulness Based Cognitive Therapy for Depression: A New Approach to Preventing Relapse: Guilford Press; 2002.
92. Dowd H, Hogan MJ, McGuire BE, et al. Comparison of an online mindfulness-based cognitive therapy intervention with online pain management psychoeducation: a randomized controlled study. Clin J Pain 2015 Jun;31(6):517-527. [doi: 10.1097/AJP.0000000000000201] [Medline: 25565584]
93. Thompson NJ, Patel AH, Selwa LM, et al. Expanding the efficacy of Project UPLIFT: distance delivery of mindfulness-based depression prevention to people with epilepsy. J Consult Clin Psychol 2015 Apr;83(2):304-313. [doi: 10.1037/a0038404] [Medline: 25495361]
94. Hayes SC, Pistorello J, Levin ME. Acceptance and commitment therapy as a unified model of behavior change. Couns Psychol 2012 Oct;40(7):976-1002. [doi: 10.1177/0011000012460836]
95. Trompetter HR, Bohlmeijer ET, Veehof MM, Schreurs KMG. Internet-based guided self-help intervention for chronic pain based on Acceptance and Commitment Therapy: a randomized controlled trial. J Behav Med 2015 Feb;38(1):66-80. [doi: 10.1007/s10865-014-9579-0] [Medline: 24923259]
96. Buhrman M, Skoglund A, Husell J, et al. Guided internet-delivered acceptance and commitment therapy for chronic pain patients: a randomized controlled trial. Behav Res Ther 2013 Jun;51(6):307-315. [doi: 10.1016/j.brat.2013.02.010] [Medline: 23548250]
97. Cavanagh K, Churchard A, O'Hanlon P, et al. A randomised controlled trial of a brief online mindfulness-based intervention in a non-clinical population: replication and extension. Mindfulness (N Y) 2018;9(4):1191-1205. [doi: 10.1007/s12671-017-0856-1] [Medline: 30100934]

XSL•FO
RenderX

98. Bhayee S, Tomaszewski P, Lee DH, et al. Attentional and affective consequences of technology supported mindfulness training: a randomised, active control, efficacy trial. BMC Psychol 2016 Nov 29;4(1):60. [doi: 10.1186/s40359-016-0168-6] [Medline: 27894358]
99. Aikens KA, Astin J, Pelletier KR, et al. Mindfulness goes to work: impact of an online workplace intervention. J Occup Environ Med 2014 Jul;56(7):721-731. [doi: 10.1097/JOM.0000000000000209] [Medline: 24988100]
100. Wolever RQ, Bobinet KJ, McCabe K, et al. Effective and viable mind-body stress reduction in the workplace: a randomized controlled trial. J Occup Health Psychol 2012;17(2):246-258. [doi: 10.1037/a0027278]
101. Mak WW, Chan AT, Cheung EY, Lin CL, Ngai KC. Enhancing web-based mindfulness training for mental health promotion with the health action process approach: randomized controlled trial. J Med Internet Res 2015 Jan 19;17(1):e8. [doi: 10.2196/jmir.3746]
102. Crane RS, Brewer J, Feldman C, et al. What defines mindfulness-based programs? The warp and the weft. Psychol Med 2017 Apr;47(6):990-999. [doi: 10.1017/S0033291716003317]
103. Sevilla-Llewellyn-Jones J, Santesteban-Echarri O, Pryor I, McGorry P, Alvarez-Jimenez M. Web-based mindfulness interventions for mental health treatment: systematic review and meta-analysis. JMIR Ment Health 2018 Sep 25;5(3):e10278. [doi: 10.2196/10278] [Medline: 30274963]
104. Lau N, O'Daffer A, Colt S, et al. Android and iPhone mobile apps for psychosocial wellness and stress management: systematic search in app stores and literature review. JMIR Mhealth Uhealth 2020 May 22;8(5):e17798. [doi: 10.2196/17798] [Medline: 32357125]
105. Prakash RS, Manglani HR, Duraney EJ, et al. TRACking health behaviors in people with Multiple Sclerosis (TRAC-MS): study protocol and description of the study sample. Contemp Clin Trials Commun 2022 Dec;30:101006. [doi: 10.1016/j.conctc.2022.101006] [Medline: 36203849]
106. Phansikar M, Duraney EJ, Manglani HR, et al. TRACking health behavior in people with multiple sclerosis: effects of a randomized trial on physical activity and working memory. Rehabil Psychol 2025 Aug;70(3):227-239. [doi: 10.1037/rep0000578] [Medline: 39207422]
107. Buckner RL, Snyder AZ, Shannon BJ, et al. Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. J Neurosci 2005 Aug 24;25(34):7709-7717. [doi: 10.1523/JNEUROSCI.2177-05.2005]
108. Elman JA, Madison CM, Baker SL, et al. Effects of beta-amyloid on resting state functional connectivity within and between networks reflect known patterns of regional vulnerability. Cereb Cortex :bhu259. [doi: 10.1093/cercor/bhu259]
109. Kaufman SK, Del Tredici K, Thomas TL, Braak H, Diamond MI. Tau seeding activity begins in the transentorhinal/entorhinal regions and anticipates phospho-tau pathology in Alzheimer's disease and PART. Acta Neuropathol 2018 Jul;136(1):57-67. [doi: 10.1007/s00401-018-1855-6]
110. Adams JN, Lockhart SN, Li L, Jagust WJ. Relationships between tau and glucose metabolism reflect Alzheimer's disease pathology in cognitively normal older adults. Cereb Cortex 2019 May 1;29(5):1997-2009. [doi: 10.1093/cercor/bhy078] [Medline: 29912295]
111. Franzmeier N, Rubinski A, Neitzel J, et al. Functional connectivity associated with tau levels in ageing, Alzheimer's, and small vessel disease. Brain (Bacau) 2019 Apr 1;142(4):1093-1107. [doi: 10.1093/brain/awz026]
112. Cope TE, Rittman T, Borchert RJ, et al. Tau burden and the functional connectome in Alzheimer's disease and progressive supranuclear palsy. Brain (Bacau) 2018 Feb 1;141(2):550-567. [doi: 10.1093/brain/awx347]
113. Palmqvist S, Janelidze S, Quiroz YT, et al. Discriminative accuracy of plasma phospho-tau217 for Alzheimer disease vs other neurodegenerative disorders. JAMA 2020 Aug 25;324(8):772-781. [doi: 10.1001/jama.2020.12134] [Medline: 32722745]
114. Leuzy A, Cullen NC, Mattsson-Carlgren N, Hansson O. Current advances in plasma and cerebrospinal fluid biomarkers in Alzheimer's disease. Curr Opin Neurol 2021 Apr 1;34(2):266-274. [doi: 10.1097/WCO.0000000000000904] [Medline: 33470669]
115. VandeVrede L, Cho H, Sanderson-Cimino M, et al. Detection of Alzheimer neuropathology in Alzheimer and non-Alzheimer clinical syndromes with blood-based biomarkers. JAMA Neurol 2025 Feb 10;82(4):344-354. [doi: 10.1001/jamaneurol.2024.5017] [Medline: 39928343]
116. Palmqvist S, Insel PS, Stomrud E, et al. Cerebrospinal fluid and plasma biomarker trajectories with increasing amyloid deposition in Alzheimer's disease. EMBO Mol Med 2019 Dec;11(12):e11170. [doi: 10.15252/emmm.201911170] [Medline: 31709776]
117. Thijssen EH, La Joie R, Strom A, et al. Plasma phosphorylated tau 217 and phosphorylated tau 181 as biomarkers in Alzheimer's disease and frontotemporal lobar degeneration: a retrospective diagnostic performance study. Lancet Neurol 2021 Sep;20(9):739-752. [doi: 10.1016/S1474-4422(21)00214-3] [Medline: 34418401]
118. Nakamura A, Kaneko N, Villemagne VL, et al. High performance plasma amyloid-β biomarkers for Alzheimer's disease. Nature New Biol 2018 Feb;554(7691):249-254. [doi: 10.1038/nature25456]
119. Palmqvist S, Janelidze S, Stomrud E, et al. Performance of fully automated plasma assays as screening tests for Alzheimer disease-related β-amyloid status. JAMA Neurol 2019 Sep 1;76(9):1060-1069. [doi: 10.1001/jamaneurol.2019.1632] [Medline: 31233127]

XSL•FO
RenderX

120. Schindler SE, Bollinger JG, Ovod V, et al. High-precision plasma β-amyloid 42/40 predicts current and future brain amyloidosis. Neurology (ECronicon) 2019 Oct 22;93(17):e1647-e1659. [doi: 10.1212/WNL.0000000000008081] [Medline: 31371569]
121. Mattsson-Carlgren N, Janelidze S, Bateman RJ, et al. Soluble P-tau217 reflects amyloid and tau pathology and mediates the association of amyloid with tau. EMBO Mol Med 2021 Jun 7;13(6):e14022. [doi: 10.15252/emmm.202114022] [Medline: 33949133]
122. Mattsson-Carlgren N, Janelidze S, Palmqvist S, et al. Longitudinal plasma p-tau217 is increased in early stages of Alzheimer's disease. Brain (Bacau) 2020 Dec 5;143(11):3234-3241. [doi: 10.1093/brain/awaa286] [Medline: 33068398]
123. Brickman AM, Manly JJ, Honig LS, et al. Plasma p‐tau181, p‐tau217, and other blood‐based Alzheimer's disease biomarkers in a multi‐ethnic, community study. Alzheimer's & Dementia 2021 Aug;17(8):1353-1364. [doi: 10.1002/alz.12301]

Abbreviations

AD: Alzheimer disease
AE: adverse event
CSF: cerebrospinal fluid
DMN: default mode network
ECog: Everyday Cognition
EMA: ecological momentary assessment
IADL: Instrumental Activities of Daily Living
IRB: institutional review board
MBSR: mindfulness-based stress reduction
MCI: mild cognitive impairment
MNI: Montreal Neurological Institute
MRI: magnetic resonance imaging
NIH: National Institutes of Health
PET: positron emission tomography
PI: principal investigator
RCT: randomized controlled trial
SCD: subjective cognitive decline
T-MOCA: Telephone Montreal Cognitive Assessment
T1w: T1-weighted
UDS-NB: Uniform Data Set Neuropsychological Battery

Edited by J Sarvestan; submitted 22.Oct.2025; this is a non–peer-reviewed article;accepted 03.Dec.2025; published 13.Apr.2026.

Please cite as:
Prakash RS, Phansikar M, Fisher ME, Dove S, Andridge R, Scharre D, Wright KD, Head E, O'Callaghan C, Meyerson QK, Duraney EJ, Schroeder M, Chakrapani N, Grant EK, Jones T, Teng J, Esterman M, Smyth JM, Gregoire T, Malhas R
Feasibility of Internet-Based Mind-Body Training for Adults With Subjective Cognitive Decline: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e86276
URL: https://www.researchprotocols.org/2026/1/e86276
doi:10.2196/86276

© Ruchika S Prakash, Madhura Phansikar, Megan E Fisher, Sarah Dove, Rebecca Andridge, Doug Scharre, Kathy D Wright, Elizabeth Head, Claire O'Callaghan, Quinn K Meyerson, Elizabeth J Duraney, Matthew Schroeder, Niyathi Chakrapani, Elise K Grant, Tyreek Jones, James Teng, Michael Esterman, Joshua M Smyth, Thomas Gregoire, Rond Malhas. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

The National Dental Practice-Based Research Network Dental Implant Restoration Registry to Evaluate Dental Implant Outcomes in Community Practice Settings: Protocol for a Prospective Observational Study

Joana Cunha-Cruz[1], DDS, PhD; Danyelle Barton[2]; David Cochran[3], DDS, PhD; Phillip Crawford[2], MS; Farahnaz Fahimipour[4], DMD, MS, PhD; Ellen Funkhouser[1], DrPH; Nicolaas C Geurs[4], DDS, MS; Gregg H Gilbert[1], DDS, MBA; James D Johnson[5]; Ning Smith[2], PhD; Kaur Maninder[4], BDS, MPH, MS; Sarah Jane Startley[4], DMD; Kimberly Stewart[2], MPH; Lisa Waiwaiole[2], MS; National Dental PBRN Collaborative Group[6,7]

[1]Department of Clinical and Community Sciences, School of Dentistry, University of Alabama at Birmingham, 1717 11th Avenue S, Suite 402E, Birmingham, AL, United States

[2]Kaiser Permanente, Portland, OR, United States

[3]Department of Periodontology, University of Texas Health San Antonio, San Antonio, TX, United States

[4]Department of Periodontology, School of Dentistry, University of Alabama at Birmingham, Birmingham, AL, United States

[5]College of Dentistry, University of Florida, Gainesville, FL, United States

[6]See Acknowledgments

[7]

Corresponding Author:
Joana Cunha-Cruz, DDS, PhD
Department of Clinical and Community Sciences, School of Dentistry, University of Alabama at Birmingham, 1717 11th Avenue S, Suite 402E, Birmingham, AL, United States

Abstract

Background: Dental implants are a widely used therapeutic option for tooth replacement; however, biological and prosthetic complications may compromise implant success. While prior research has largely focused on academic or specialty settings, data on implant outcomes in community dental practices remain limited.

Objective: This study aims to establish a national registry within the National Dental Practice-Based Research Network to evaluate the incidence of and factors associated with biological and prosthetic complications following implant therapy in community practice settings.

Methods: This prospective observational cohort study plans to include approximately 2000 implant restorations from 1550 patients receiving implant restorations from 150 practitioners across 6 regional network nodes. Clinical, implant, and prosthetic characteristics will be recorded at baseline, with annual follow-up visits collecting data on complications, implant failures, prosthetic issues, and patient-reported outcomes. Digitized radiographs will be centrally reviewed for peri-implant bone changes, emergence angle, and prosthetic fit. The primary outcome is the incidence of biological and prosthetic complications. Secondary analyses will evaluate patient-centered outcomes and identify risk factors for complications. Multilevel mixed-effects regression and time-to-event analyses will be used to account for clustering and censored data.

Results: Of 430 practitioners who expressed initial interest, 187 (43.5%) completed training and began patient enrollment, with 143 (76.5%) of these practitioners enrolling at least one participant. A total of 1453 patients were registered between August 2022 and October 2024. Practitioner activity and enrollment peaked in late 2023. Overall, 1178 (81.1%) baseline patient surveys and 1453 (100%) baseline clinical forms were completed, covering 2087 dental implant restorations.

Conclusions: The registry will generate robust, real-world evidence on implant-related complications and their predictors in community practice. The findings will enhance clinical decision-making; support personalized risk assessment; and, ultimately, improve patient outcomes in implant dentistry.

International Registered Report Identifier (IRRID): DERR1-10.2196/82795

(JMIR Res Protoc 2026;15:e82795) doi:10.2196/82795

KEYWORDS

dental implants; dental prosthesis; dentures, implant supported; oral health; prospective studies; practice-based research

XSL•FO
RenderX

Introduction

Dental implants have become a standard treatment modality for rehabilitating partially and fully edentulous patients worldwide. The global dental implant market, valued at approximately US $9 billion in 2025 with an estimated 5 million implants placed annually, is projected to experience steady growth due to rising patient demand and advances in implant technologies [1]. While dental implants have demonstrated long-term success in many cases, they are not free from complications associated with various factors, including treatment planning, patient-related issues, surgical and prosthetic execution, material failure, and lack of maintenance [2-5].

The longitudinal survival rates of osseointegrated dental implants range between 90% and 96.2% after 5 years [2-5]. However, these figures may not fully capture the rates of peri-implant disease and health. Biological and prosthetic complications can interfere with the health of the peri-implant tissues and the function and aesthetics of implant restorations [6]. Peri-implant diseases are classified into peri-implant mucositis (inflammation restricted to the peri-implant mucosa) and peri-implantitis (characterized by peri-implant bone loss). The prevalence of peri-implant mucositis and peri-implantitis varies widely, with reported ranges of 19% to 65% and 1% to 47%, respectively, and meta-analyses estimating weighted mean prevalences of 43% to 63% for mucositis and 22% to 25% for peri-implantitis [7-12]. Prosthetic complications, such as screw loosening, fractures, chipping, and decementation, pose significant clinical challenges in implant dentistry. Their prevalence ranges from 10% to 16% during midterm follow-up influenced by prosthetic design and materials [13-15].

The current body of literature on biological and prosthetic complications is limited by small studies primarily conducted in European academic and specialty settings [4,16,17]. Many of these studies have been retrospective and may not accurately represent community practice settings. The National Dental Practice-Based Research Network (PBRN) Dental Implant Restoration Registry aims to address these knowledge gaps by creating an implant registry to record implant and prosthetic procedures and characteristics and any complications in community practice settings. The primary objective of this study is to quantify the incidence of biological and prosthetic complications in dental implants over a 3-year period. Secondary objectives are to assess risk factors associated with implant complications, develop a risk assessment tool for implant complications, and evaluate patient-reported outcomes related to implant therapy.

Methods

Study Design and Setting

This is a prospective observational cohort study designed to evaluate outcomes of dental implant restorations over a 3-year follow-up period in community practice settings. It will be conducted across approximately 150 dental practices in the United States. Eligible practitioners who are members of the National Dental PBRN from all 6 regional nodes and provide dental implant restorative therapy will be invited to participate.

The study protocol and data collection instruments have been reviewed by the network's Practitioner Executive Committee. The funding agency and the institutional review board (IRB) also reviewed and approved the protocol (version 6; November 11, 2024). This study protocol is reported following the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) [18] reporting guidelines for protocols and the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guidelines for observational studies [19].

The National Dental PBRN is a nationwide initiative funded by the National Institute of Dental and Craniofacial Research (NIDCR) that engages dental practitioners from a broad range of community settings across the United States in conducting clinical research relevant to routine dental care [20,21]. The network is structured into 6 regional nodes, each housed at distinct academic or research institutions. Each node is led by node directors and supported by node research coordinators, who serve as the primary liaisons between the network and participating dental practitioners. Node coordinators facilitate communication, recruitment, and coordination of study activities, fostering close collaboration with dental providers to ensure study feasibility and relevance.

An administrative and resource center (ARC) housed at the University of Alabama at Birmingham oversees centralized functions, including ethical review through the central IRB, administrative agreements, and management of payments to study participants and dental practitioners acting as individual investigators. The ARC leads national engagement efforts and communication and dissemination activities. The Practitioner Executive Committee, composed of dental practitioner members from the network, provides critical input on study concepts, protocol development, and implementation strategies. The National Coordinating Center based at Kaiser Permanente Center for Health Research manages data systems, quality control, and monitoring activities. Funding for the network is coordinated through distinct requests for applications issued by the National Institutes of Health (NIH) and NIDCR; separate collaborative agreements fund the ARC, the National Coordinating Center, and individual research studies. Program officers from the NIH and NIDCR work closely with the network leadership across all components to ensure alignment and successful execution. This organizational structure promotes close collaboration among dental practitioners, researchers, research coordinators, data managers, biostatisticians, and principal investigator teams to facilitate the successful completion of studies.

Study Population

Participant inclusion criteria are patients with planned receipt of dental implant restorations from a participating practice who are older than 18 or 19 years (depending on the regional node) and are available for the duration of the study. Exclusion criteria include an inability to comply with study procedures or provide informed consent. Each enrolled patient can contribute more than one implant if multiple implants will be restored at the baseline visit.

XSL•FO
RenderX

Recruitment and Retention

Practitioners will recruit and enroll their patients during the prosthetic phase of implant therapy after they complete study readiness and training activities. The network's regional coordinators will assist practitioners in implementing the study protocol and monitoring recruitment progress. Strategies to enhance recruitment include regular communication with participating practitioners, provision of recruitment materials, offering continuing education credits for study participation, and periodic recruitment progress reports to motivate practitioners. Patient retention will be promoted through reminder emails, SMS text messages, or telephone calls for follow-up visits and surveys, and flexible scheduling of follow-up visits.

Study Schedule of Assessments

The study timeline includes a baseline visit (prosthesis insertion) for enrollment, consent, and baseline data collection, followed by annual follow-up visits at 1, 2, and 3 years (±30 days and +150 days from the target date).

Data Collection Procedures and Variables

Overview

All study data will be collected electronically using secure, web-based forms accessible via tablets, smartphones, or computers. Practitioners will enter clinical examination and procedural data through electronic case report forms (CRFs) at baseline and during each annual follow-up visit [22,23]. Radiographs taken as part of routine clinical care at baseline (prosthesis placement) and at each follow-up visit will be uploaded by practitioners to a secure centralized electronic repository. Separately, enrolled participants will complete electronic surveys following the baseline visit and at each annual follow-up interval.

Patient-Completed Surveys

Participants will provide self-reported demographic data, including age, sex, race, ethnicity, insurance type, educational level, household income, and residential environment. They will also report on general health status and behaviors, including medical conditions, allergies, medication use, and tobacco product use (cigarettes, cigars, smokeless tobacco, and electronic nicotine delivery systems). Oral health–related data will include measures of oral health–related quality of life (OHRQoL) and symptoms of dry mouth.

OHRQoL will be assessed using 10 items with responses on a 4-point Likert scale ("never," "sometimes," "always," and "not sure"). Items will query functional limitations (eg, restrictions on eating and difficulty with chewing or speaking), pain or discomfort during eating or oral hygiene, medication use for dental pain, social impacts, and satisfaction with dental appearance and implant-supported prostheses drawing on domains used in established OHRQoL instruments [24,25]. Dry mouth symptoms will be assessed using 4 questions covering frequency of dry mouth diagnosis, difficulty swallowing, oral dryness during meals, and burning tongue sensations.

At each annual follow-up, participants will update their health status and medication and tobacco use and recomplete the OHRQoL questions.

Practitioner-Completed Clinical Research Forms

Practitioners will complete standardized electronic CRFs at baseline and each follow-up visit to capture comprehensive patient-, implant-, and prosthesis-level data:

- Patient-level clinical data—at baseline, practitioners will assess and record the participants' overall oral health, detailed periodontal status (health, gingivitis, and periodontitis stage and progression) around natural teeth, the presence and extent of peri-implant disease in existing implants, routine dental maintenance history (frequency of maintenance visits in the prior 12 months), and parafunctional habits (eg, bruxism and occlusal guard use).
- Implant-level characteristics—for each study implant, practitioners will document the reason for tooth loss at the implant site; implant brand, length, and diameter; timing of placement in relation to extraction; results of bone and soft tissue grafting; and any early healing complications.
- Peri-implant mucosal characteristics—practitioners will document their assessments of peri-implant soft tissue (bleeding on probing, presence of purulence, deepest probing depth and location, mucosal recession and its extent, and width of keratinized mucosa), patient-reported pain at the implant site (with pain characteristics), and radiographic evaluation of bone levels.
- Prosthetic characteristics—data will be collected on prosthetic connection type, temporary prosthesis use and timing, final prosthesis type (fixed or removable), retention method, material, occlusal scheme, classification of opposing dentition, number of units and pontics, cantilever presence, abutment and cement types, seating and fit, and the presence and location of open contacts.

At each follow-up, CRFs will be used to record changes in periodontal and peri-implant status, occurrences of peri-implant disease or other complications, updates on implant survival (removal and reason), patient-reported pain, and radiographic bone loss. Practitioners will also document any biological or prosthetic complications (eg, screw loosening or fracture, prosthesis or abutment misfit, loss of retention, chipping or fracture of restorative materials, cement remnants, and adverse events), as well as management procedures and outcomes.

Radiographic Measurement for the Assessment of Bone Level and Prosthetic Fit

Radiographs obtained at baseline (prosthetic placement) and at annual follow-up visits (1, 2, and 3 years) will be used to quantify peri-implant bone levels and evaluate the fit of implant-supported prostheses. Calibrated evaluators will measure bone levels and prosthetic fit using the NIH ImageJ software [26] following a standardized protocol, with annotated images and measurements securely saved and uploaded to the study database.

Alveolar bone levels will be measured at the mesial and distal surfaces of each study implant as the linear distance from the implant platform to the most coronal bone-implant contact

XSL•FO
RenderX

visible on the radiograph. As images will originate from different clinical systems with varying resolutions, measurements will be recorded in pixels initially and converted to millimeters using implant-specific geometry parameters (implant length or diameter) to standardize measurements across images.

Radiographs will also be examined for gaps at the implant-abutment and abutment-prosthesis interfaces, with fit discrepancies recorded as a dichotomous variable. The emergence angle will be measured mesially and distally using the implant platform as the vertex and lines drawn along the implant's long axis and the gingival contour of the abutment or prosthesis to assess prosthetic design and potential biomechanical risk factors.

Data Management and Monitoring

All data will be entered directly into the electronic data capture system, which includes built-in quality control checks. All radiographs will be anonymized and reviewed centrally for quality assurance. Two independent and calibrated evaluators blinded to clinical data and each other's measurements will perform radiographic measurements. Discrepancies exceeding predefined thresholds will prompt involvement of a third adjudicator, who will conduct an independent assessment. All measurement data and marked images will be stored in a centralized repository.

Data will be encrypted for transmission to the coordinating center and stored on password-protected computers. Unique study codes will be used for participants to maintain confidentiality. The data management team will perform regular data quality checks and generate reports on missing data, inconsistencies, and protocol deviations. Research coordinators will support practitioners in resolving data inconsistencies and minimizing missing data whenever possible.

The study team will implement a clinical site monitoring plan to ensure the quality and integrity of the data collected. The monitoring strategy will include remote and/or on-site monitoring, regular quality control assessments, and ongoing monitoring of protocol adherence. These procedures will maintain high standards of data quality and ensure that study conduct is compliant with all applicable human subject research and regulatory requirements.

Although the study is observational, participant safety will remain a priority. Safety monitoring will focus on unanticipated problems involving risks to participants, including unanticipated problems that meet the definition of a serious adverse event. All reported events will be followed up on by the study team until resolution or stabilization. The principal investigator will be responsible for overall study oversight, including monitoring and assessment of adverse events and unanticipated problems and ensuring data integrity and adherence to the approved study protocol.

Dissemination and Data Sharing

Results from this study will be disseminated through multiple channels, including publication in peer-reviewed scientific journals, presentation at national and international dental conferences, dissemination to network practitioners through newsletters and webinars, and public release of findings through press releases and the network website. Authorship will be determined according to the International Committee of Medical Journal Editors guidelines. The network will acknowledge the efforts of all participating practitioners by listing them as a group author in primary publications. Deidentified participant-level data and statistical code will be made available after the publication of primary study results in the PBRN's website consistent with NIH policy [27].

Statistical Considerations

Sample Size

Approximately 1550 patients will be enrolled to achieve a sample of 2000 implants. This sample size was determined based on anticipated complication rates of 10% to 15% over 3 years, the desire to detect differences in complication rates as small as 5% between groups, accounting for clustering of implant restorations within patients and practitioners, and allowing for up to 20% loss to follow-up. The sample size provides 80% power to detect clinically meaningful differences in complication rates between subgroups, with a 2-sided α of .05.

Statistical Analysis Plan

The study's primary outcomes will be the incidence of biological and prosthetic complications occurring over the 3-year follow-up period. *Biological complications* will be defined according to the following criteria:

- Peri-implant mucositis—presence of bleeding on probing and/or purulent exudate around the implant without concomitant radiographic bone loss
- Peri-implantitis—radiographic peri-implant bone loss will be calculated by comparing the bone level at each annual follow-up to the baseline measurement, with bone loss defined as a negative change in millimeters
- Implant failure—documented implant loss or planned removal as reported by practitioners during follow-up

Prosthetic complications will include any event compromising the aesthetics or function of the implant restoration. These complications encompass screw loosening or fracture, structural fractures, resin or porcelain fractures, loss of retention, abutment or implant misfit, and other adverse prosthetic events captured via practitioner-completed data forms.

Patient-centered outcomes are critical components in evaluating the overall success of dental treatments. Secondary outcomes will include patient-reported measures of satisfaction with implant aesthetics and function as well as OHRQoL.

Descriptive statistics will be presented to summarize baseline characteristics and outcome variables. A study flow diagram will illustrate the number of participants at each stage of the study, including reasons for exclusion, loss to follow-up, and other participant dispositions following reporting guidelines [18,19]. Incidence and prevalence rates of biological and prosthetic complications will be estimated per implant and per year of follow-up. The primary analysis will use multilevel mixed-effects models to account for clustering of implant

XSL•FO
RenderX

restorations within patients and practitioners. Time-to-event analyses will be conducted for complication-free survival. Multivariable regression models incorporating patient characteristics, implant-specific factors, prosthetic details, and behavioral variables will be estimated to identify predictors of biological and prosthetic complications. Subgroup analyses will explore effect modification by implant and prosthesis type and patient demographics. Missing data will be handled using multiple imputation techniques. Sensitivity analyses will be conducted to assess the impact of missing data and loss to follow-up on the study conclusions.

Ethical Considerations

The study will be conducted in accordance with the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use guidelines, the Code of Federal Regulations on the Protection of Human Subjects (title 45, part 46), and the NIDCR Clinical Terms of Award. IRB approval will be obtained from the National Dental PBRN Central IRB at the University of Alabama at Birmingham, with local context review by the local IRBs from each regional node. Any modifications to the protocol that may impact the conduct of the study or patient safety will be submitted for IRB approval before implementation. Protocol amendments will be communicated to all participating sites and investigators.

Informed consent will be obtained verbally at the prosthetic placement visit by trained personnel who have completed study protocol and human subject research training. The consent process will include a thorough explanation of the study procedures, risks, and benefits and will be documented in the electronic data capture system. Patients will have the opportunity to ask questions before providing consent. Participant confidentiality will be maintained through the use of unique study codes, data encryption, and password-protected storage. Access to identifiable information will be restricted to authorized study personnel on a need-to-know basis. The principal investigator and study statisticians will have access to the final dataset. Other investigators may request access to deidentified data after the completion of the primary analyses.

A check in the amount of USD $25 will be provided to participants for each survey answered for up to 4 surveys (US $100) per participant.

All investigators will disclose any potential conflicts of interest related to the study. These will be reviewed by the institutional conflict of interest committee and managed appropriately.

Results

Of the 430 practitioners who initially expressed interest in the study, 194 (45.1%) completed human subjects research training and fulfilled other research and administrative requirements. Subsequently, 187 (96.4%) of these practitioners were trained in the study protocol and data management procedures and were cleared to begin patient enrollment. Of those 187 practitioners, 143 (76.5%) enrolled at least one patient.

Of these 143, practitioners were distributed across the 6 regions: Southwest (n=38, 26.6%), South Atlantic (n=31, 21.7%), Midwest (n=28, 19.6%), South Central (n=18, 12.6%), Northeast (n=18, 12.6%), and Western (n=10, 7%). Practice settings varied, with 63 (44.1%) in solo private practices, 40 (28%) in small group practices (2 - 4 dentists), 5 (3.5%) in large group practices (5 or more dentists), 18 (12.6%) in academic settings, and 17 (11.9%) distributed among managed care or corporate practices and community health centers. More than half (n=85, 59.4%) of practitioners had advanced or specialty training, including 51 (35.7%) who completed advanced general dentistry programs (such as an advanced education in general dentistry program or general practice residency), 32 (22.4%) holding fellowships or masterships with the Academy of General Dentistry, 13 (9.1%) in prosthodontics, and 35 (24.5%) in other specialties (eg, dental implantology, orthodontics, and dental public health). Most practitioners (n=131, 91.6%) identified themselves as general dentists providing primary care, whereas 12 (8.4%) identified themselves as specialists.

Practitioner activation and monthly enrollment activity increased rapidly after study launch in June 2022, with the number of practitioners actively enrolling participants increasing from 7 in August 2022 to more than 60 by mid-2023 and peaking at just under 90 in late 2023 (Figure 1). Monthly activity, assessed via the number of practitioners enrolling at least one participant, peaked between months 15 and 16, with 90 practitioners actively enrolling, and then declined to 31 by the final study month. The cumulative activation curve crossed the 25% and 50% thresholds by approximately December 2022 and June 2023, respectively, and surpassed 75% by the third quarter of 2023 as additional practices from all regions began enrolling; after the peak, the number of active enrolling practitioners declined gradually until October 2024 even as the cumulative number of active practitioners continued to increase to 100%.

XSL•FO
RenderX

Figure 1. Percentage of active practitioners and monthly number of practitioners actively enrolling participants stratified by National Dental Practice-Based Research Network regional nodes: Dental Implant Restoration Registry from August 2022 to October 2024.

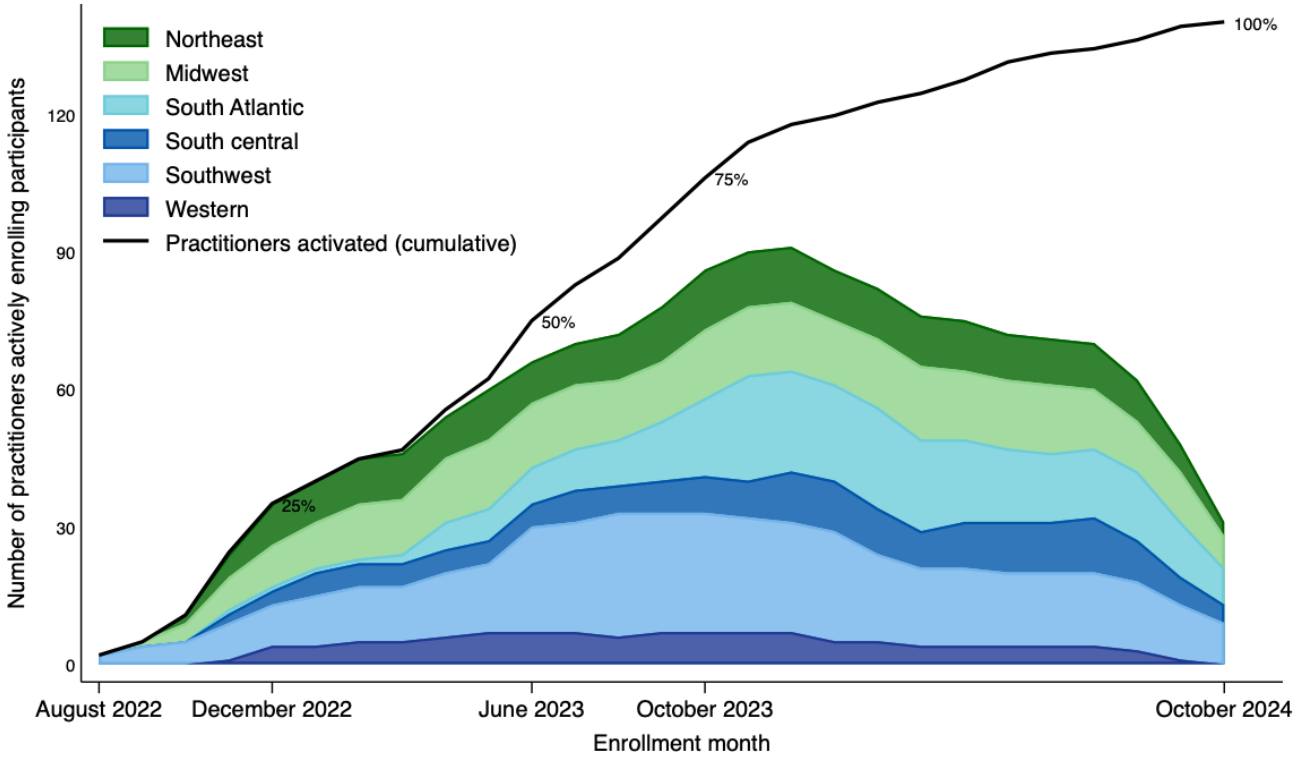

Cumulative patient enrollment began gradually and increased steadily across all 6 regional nodes. The southwest and midwest nodes contributed the largest shares, producing a near-linear rise in total accrual over time. The cumulative enrollment curve crossed 25% of the final sample by June 2023 (month 11), 50% by November 2023 (month 15), and 75% by April 2024 (month 21), reaching the target sample by October 2024 (Figure 2).

Figure 2. Number of participants enrolled in the Dental Implant Restoration Registry stratified by National Dental Practice-Based Research Network regional nodes.

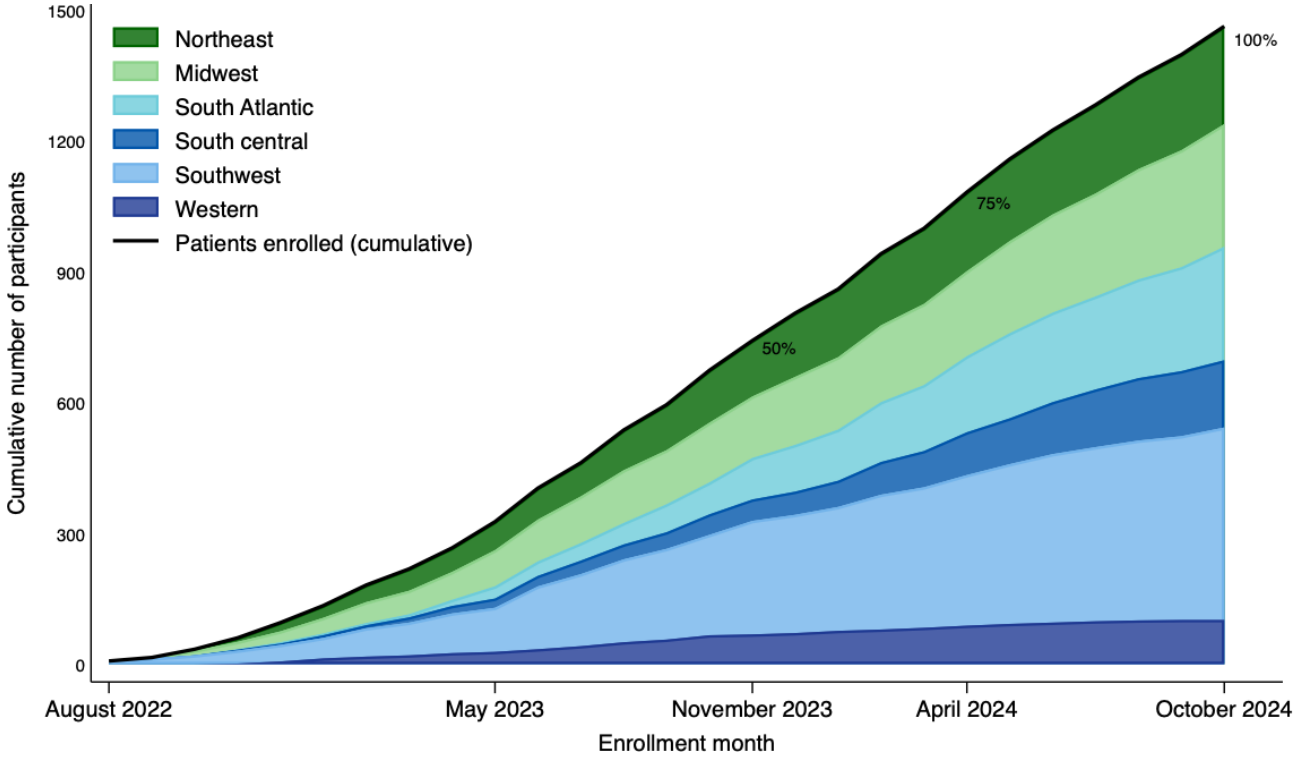

On average, each practitioner enrolled 10 (SD 8.5) patient participants. The median number per practitioner was 8 (IQR 4-13), with a range from 1 to 39 participants. On average, each practitioner enrolled 14.6 (SD 12) implant restorations, with a median of 11 (IQR 5-21). The maximum number of implant restorations enrolled by a practitioner was 54.

XSL•FO
RenderX

The total number of participants enrolled in the registry was 1453. Practitioners completed all baseline CRFs for 2087 dental implant restorations, and 1178 (81.1%) participants completed their baseline surveys.

Discussion

This paper details the successful establishment and operational metrics of a large-scale, geographically diverse, prospective dental implant registry in the United States. Over a 26-month period (August 2022 to October 2024), the registry enrolled 1453 participants and more than 2000 dental implant restorations. This achievement was driven predominantly by general dentists operating in solo and small group private practices, reflecting the real-world clinical context in the United States. Practitioner activation and patient enrollment peaked later in the study, with the 75% patient milestone reached in month 21. Methodological rigor was maintained, with completion of all baseline CRFs by practitioners and an 81.1% (1178/1453) data completion rate for patient surveys.

A key methodological weakness in prior studies of dental implant outcomes is their reliance on data from university and specialized implant clinics, which limits generalizability to general practice settings [2,7]. This registry addresses this limitation by engaging practitioners from diverse community settings across 6 US regions, positioning the cohort as representative of real-world effectiveness in implant therapy [8,9]. The final sample size is considerably larger than those of many previous epidemiological studies in implantology, which have typically included between 100 and 662 patients [7,8]. This scale overcomes the issue of limited sample size often cited in the literature [7] and provides a robust foundation for assessing outcomes and validating risk factors for peri-implant diseases [9].

The recruitment dynamics observed—a phased approach over 26 months with a late peak in activity—mirror the operational challenges typical of pragmatic trials [28]. The need to enroll and sustain the engagement of numerous practicing clinicians often leads to initial enthusiasm waning, resulting in premature discontinuation due to poor recruitment [29]. This registry overcame these challenges through staggered training, ramp-up periods, and ongoing activation of new practitioners throughout the enrollment period. This approach aligns with the operational model of the PBRN [21]. The rapid increase in activity after the initial launch phase demonstrates the effectiveness of the training and support infrastructure provided to participating sites.

The enrollment window, with key accrual milestones reached only after a prolonged ramp-up, is consistent with reports that large, multisite pragmatic studies in community practice settings often require extended periods to achieve target sample sizes [21,28]. Seasonal patterns in health care use and regional variation in practice organization contribute to fluctuating recruitment rates, which must be considered in planning and interpreting accrual trajectories [30,31]. Regional heterogeneity in patient demographics, insurance coverage, and practice models influences both participation rates and case mix, potentially affecting complication rates and treatment patterns [7,10]. Documentation of these temporal and regional enrollment patterns will enable sensitivity analysis in the future when estimating peri-implant disease burden and implant survival [7,11,12].

Limitations of this study include a few methodological constraints. The use of convenience sampling based on the willingness of practitioners and patients to participate limits the ability to generalize findings to the broader population [7,10]. Heterogeneity across multiple practitioners and multiple practice environments introduces interoperator variability in diagnostic and treatment approaches, as well as in outcome assessment procedures [14]. While this variability enhances external validity by reflecting real-world practice patterns, it may complicate causal inference and necessitates sensitivity analyses to explore the impact of practice-level effects [14]. Systematic reviews stress that assessing technical and biological events requires a mean observation period of at least 5 years [2,15], whereas monitoring implant loss in high-risk patients often demands follow-up of 10 years or more [4]. Additionally, loss to follow-up is common in prospective registries and may disproportionately affect patients with poorer adherence or complex needs, potentially biasing outcome estimates [1,13] Within a 5- to 7-year funding cycle, the combined demands of start-up, recruitment, and long-term follow-up may not allow for sufficient time to fully characterize late implant failures, particularly among high-risk groups.

Despite these limitations, the registry provides a substantially larger, more diverse, and methodologically rigorous dataset than most previous studies on dental implant outcomes. A key strength of the registry is the combination of high data completeness [22] with broad representation of general practice settings, providing a unique platform for refining contemporary estimates of peri-implant mucositis and peri-implantitis prevalence in the United States [8,14,21]. These data can be directly compared with recent global estimates reporting high patient-level prevalence of peri-implant inflammation and disease, thereby contextualizing US practice-based outcomes within the international literature [7,11,12]. The detailed capture of patient-, implant-, and procedure-level variables enables rigorous evaluation of established and emerging risk indicators, including history of periodontitis, smoking, diabetes, and other systemic or behavioral factors identified as influential in recent meta-analyses [4,11]. Such analyses will be particularly informative for clarifying potential interactions with types of implant fixtures and restoration techniques, which remain incompletely characterized in existing cohort studies [4,10]. Recent evaluations of dental implant prediction models point to limited methodological transparency and poor external validation [14]. By supporting both model development and true external validation in an independent, heterogeneous, practice-based cohort, the Dental Implant Restoration Registry study can contribute to the evolution of more robust clinical prediction tools that support precision dentistry and individualized risk-based prevention strategies [14].

The results of this study have the potential to significantly impact clinical practice by informing evidence-based recommendations for implant therapy, risk assessment, and complication prevention. The establishment of a national implant

XSL•FO
RenderX

registry creates opportunities for future targeted studies on specific complications identified through this research. This protocol outlines a comprehensive approach to addressing current knowledge gaps in dental implant and prosthetic outcomes in dental practice settings, contributing to improved patient care and clinical decision-making in implant dentistry.

Acknowledgments

The opinions and assertions contained herein are those of the authors and are not to be construed as necessarily representing the views of the respective organizations or the National Institutes of Health. There is an internet site devoted to the network [30]. The authors used the generative artificial intelligence tool ChatGPT developed by OpenAI (GPT-4 architecture) to assist in drafting manuscript content. Artificial intelligence–generated text was subsequently reviewed, edited, and approved by the authors to ensure accuracy and appropriateness. The authors thank the National Dental Practice-Based Research Network (PBRN) dentists, clinic staff, and patients who participated in this study and the practitioners of the National Dental PBRN Practitioner Executive Committee, who provided important suggestions during protocol development and early study implementation. In addition to those who are already acknowledged by serving as coauthors on this manuscript, the authors are also grateful to the investigators and staff of the National Dental PBRN administrative and resource center at the Birmingham, Alabama, site (Muna Anabtawi, national program manager; Brittni Ball, national program coordinator; Aleena Potluri, node coordinator; and Trishathi Malagar Nandakumar, data coordinator); the Gainesville, Florida, site (Valeria Gordan, national director of the practitioner training and practitioner recruitment and engagement components and director of the south Atlantic region, and Brenda Thacker, regional program manager); the Rochester, New York, site (Dorota Kopycka-Kedzierawski, director of the northeast region; Cyril Meyerowitz, assistant director of the northeast region; Kathy Bohn, regional program manager; Rita Cacciato, node coordinator; Pat Regusa, node coordinator; and Victoria Thomas, node coordinator); the San Antonio, Texas, site (Rahma Mungia, assistant director of the southwest node; Ashley Spencer, node coordinator; and Caitlin Sangdahl, node coordinator); the Minneapolis, Minnesota, site (Brad Rindal, director of the midwest node; Jose Maldonado-Ortiz, assistant director; Sarah Basile, program manager; Chris Enstad, node coordinator; Amanda Gillesby, node coordinator; Kimberly Johnson, node coordinator; and Heather Weidner, node coordinator); and the Portland, Oregon, site (Jeffrey Fellows, director of the western region; Chalinya Ingphakorn, node coordinator; and Christine Catlin, node coordinator) and the investigators and staff of the National Dental PBRN National Coordinating Center (Mary Ann McBurnie, principal investigator and director; Kim Funkhouser, technical director; Suzanne Gillespie, administrative codirector; Tamara Lischka, study manager; and Celeste Machen).

Funding

This work was supported by National Institutes of Health grants UH3-DE030090, U19-DE028717, and U01-DE028727.

Data Availability

The datasets generated or analyzed during this study will be publicly available at National Dental Practice-Based Research Network website [23] when the results for the primary outcomes are published but are available now from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JC-C, FF, NCG, KM
Data curation: DB, PC, EF, NS, KS
Formal analysis: JC-C, EF, NS
Funding acquisition: JC-C, NCG, GHG
Investigation: DB, DC, JC-C, FF, NCG, GHG, JDJ, KM, SJS, LW
Methodology: DB, DC, PC, JC-C, FF, EF, NCG, GHG, JDJ, KM, NS, SJS, LW
Project administration: DB, JC-C, NCG
Supervision: DC, JC-C, NCG
Validation: DB, EF, NS
Visualization: JC-C
Writing—original draft: JC-C
Writing—review and editing: DB, DC, JC-C, FF, EF, NCG, GHG, JDJ, KM, NS, SJS, LW

Conflicts of Interest

None declared.

References

1. Dental implants market size & share: industry report, 2030. Grand View Research. 2024. URL: https://www.grandviewresearch.com/industry-analysis/dental-implants-market [accessed 2025-05-25]

XSL•FO
RenderX

2. Pjetursson BE, Asgeirsson AG, Zwahlen M, Sailer I. Improvements in implant dentistry over the last decade: comparison of survival and complication rates in older and newer publications. Int J Oral Maxillofac Implants 2014;29 Suppl:308-324. [doi: 10.11607/jomi.2014suppl.g5.2] [Medline: 24660206]
3. Moraschini V, Mourão C, Montemezzi P, Mourão C, et al. Clinical comparation of extra-short (4 mm) and long (>8 mm) dental implants placed in mandibular bone: a systematic review and metanalysis. Healthcare (Basel) 2021 Mar 12;9(3):315. [doi: 10.3390/healthcare9030315] [Medline: 33809203]
4. Annunziata M, Cecoro G, Guida A, et al. Effectiveness of implant therapy in patients with and without a history of periodontitis: a systematic review with meta-analysis of prospective cohort studies. J Periodontal Res 2025 Jun;60(6):524-543. [doi: 10.1111/jre.13351] [Medline: 39466662]
5. Peña-Cardelles JF, Akhondi S, Markovic J, Pala K, Lanis A, Gallucci GO. Survival rate and marginal bone loss in titanium vs titanium-zirconia single-unit narrow diameter implants. a systematic review and meta-analysis of prospective studies. Med Oral Patol Oral Cir Bucal 2025 Jul 1;30(4):e499-e511. [doi: 10.4317/medoral.27042] [Medline: 40192111]
6. Berglundh T, Armitage G, Araujo MG, et al. Peri-implant diseases and conditions: consensus report of workgroup 4 of the 2017 world workshop on the classification of periodontal and peri-implant diseases and conditions. J Periodontol 2018 Jun;89 Suppl 1:S313-S318. [doi: 10.1002/JPER.17-0739] [Medline: 29926955]
7. Derks J, Tomasi C. Peri-implant health and disease. A systematic review of current epidemiology. J Clin Periodontol 2015 Apr;42 Suppl 16:S158-S171. [doi: 10.1111/jcpe.12334] [Medline: 25495683]
8. Zitzmann NU, Berglundh T. Definition and prevalence of peri-implant diseases. J Clin Periodontol 2008 Sep;35(8 Suppl):286-291. [doi: 10.1111/j.1600-051X.2008.01274.x] [Medline: 18724856]
9. Lindhe J, Meyle J, Group D of European Workshop on Periodontology. Peri-implant diseases: consensus report of the sixth European workshop on periodontology. J Clin Periodontol 2008 Sep;35(8 Suppl):282-285. [doi: 10.1111/j.1600-051X.2008.01283.x] [Medline: 18724855]
10. Reis I, Huamán-Mendoza AA, Ramadan D, et al. The prevalence of peri-implant mucositis and peri-implantitis based on the world workshop criteria: a systematic review and meta-analysis. J Dent 2025 Sep;160:105914. [doi: 10.1016/j.jdent.2025.105914] [Medline: 40523497]
11. Galarraga-Vinueza ME, Pagni S, Finkelman M, Schoenbaum T, Chambrone L. Prevalence, incidence, systemic, behavioral, and patient-related risk factors and indicators for peri-implant diseases: an AO/AAP systematic review and meta-analysis. J Periodontol 2025 Jun;96(6):587-633. [doi: 10.1002/JPER.24-0154] [Medline: 40489307]
12. Lee CT, Huang YW, Zhu L, Weltman R. Prevalences of peri-implantitis and peri-implant mucositis: systematic review and meta-analysis. J Dent 2017 Jul;62:1-12. [doi: 10.1016/j.jdent.2017.04.011] [Medline: 28478213]
13. Molinero-Mourelle P, Fonseca M, Grossen ME, Zuercher AN, Schimmel M, Katsoulis J. Clinical and radiographic evaluation of zirconia bars for mandibular implant overdentures: a prospective cohort study with up to 3 years of follow-up. Clin Implant Dent Relat Res 2025 Aug;27(4):e70066. [doi: 10.1111/cid.70066] [Medline: 40709726]
14. Zhu Y, Du M, Li P, Lu H, Li A, Xu S. Prediction models for the complication incidence and survival rate of dental implants-a systematic review and critical appraisal. Int J Implant Dent 2025 Jan 23;11(1):5. [doi: 10.1186/s40729-025-00590-1] [Medline: 39847174]
15. Pjetursson BE, Thoma D, Jung R, Zwahlen M, Zembic A. A systematic review of the survival and complication rates of implant-supported fixed dental prostheses (FDPs) after a mean observation period of at least 5 years. Clin Oral Implants Res 2012 Oct;23 Suppl 6(22-38):22-38. [doi: 10.1111/j.1600-0501.2012.02546.x] [Medline: 23062125]
16. Srinivasan M, Meyer S, Mombelli A, Müller F. Dental implants in the elderly population: a systematic review and meta-analysis. Clin Oral Implants Res 2017 Aug;28(8):920-930. [doi: 10.1111/clr.12898] [Medline: 27273468]
17. Kupka JR, König J, Al-Nawas B, Sagheb K, Schiegnitz E. How far can we go? A 20-year meta-analysis of dental implant survival rates. Clin Oral Investig 2024 Sep 21;28(10):541. [doi: 10.1007/s00784-024-05929-3] [Medline: 39305362]
18. Chan AW, Boutron I, Hopewell S, et al. SPIRIT 2025 statement: updated guideline for protocols of randomised trials. BMJ 2025 Apr 28;389:e081477. [doi: 10.1136/bmj-2024-081477] [Medline: 40294953]
19. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Epidemiology 2007 Nov;18(6):800-804. [doi: 10.1097/EDE.0b013e3181577654] [Medline: 18049194]
20. Cunha-Cruz J, Gilbert GH, Allareddy V, et al. Characteristics of dentists in the National Dental Practice-Based Research Network. J Dent 2023 Oct;137:104653. [doi: 10.1016/j.jdent.2023.104653] [Medline: 37572986]
21. Gilbert GH, Fellows JL, Allareddy V, et al. Structure, function, and productivity from the National Dental Practice-Based Research Network. J Clin Transl Sci 2022;6(1):e87. [doi: 10.1017/cts.2022.421] [Medline: 35989860]
22. Welcome to the National Dental PBRN. National Dental Practice-Based Research Network. URL: https://www.nationaldentalpbrn.org [accessed 2025-12-05]
23. National dental practice-based research network. recruiting, ongoing, upcoming, and completed studies. National Dental Practice-Based Research Network. URL: https://www.nationaldentalpbrn.org/recruiting-ongoing-upcoming-completed [accessed 2025-12-05]

XSL•FO
RenderX

24. Locker D, Matear D, Stephens M, Lawrence H, Payne B. Comparison of the GOHAI and OHIP-14 as measures of the oral health-related quality of life of the elderly. Community Dent Oral Epidemiol 2001 Oct;29(5):373-381. [doi: 10.1034/j.1600-0528.2001.290507.x] [Medline: 11553110]
25. Manfredini M, Pellegrini M, Rigoni M, et al. Oral health-related quality of life in implant-supported rehabilitations: a prospective single-center observational cohort study. BMC Oral Health 2024 May 4;24(1):531. [doi: 10.1186/s12903-024-04265-y] [Medline: 38704566]
26. ImageJ. National Institutes of Health. URL: https://imagej.nih.gov/ij [accessed 2025-05-25]
27. Final NIH policy for data management and sharing: NIH notice NOT-OD-21-013. National Institutes of Health. 2020. URL: https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html [accessed 2025-12-05]
28. Worsley SD, Oude Rengerink K, Irving E, et al. Series: pragmatic trials and real world evidence: paper 2. Setting, sites, and investigator selection. J Clin Epidemiol 2017 Aug;88:14-20. [doi: 10.1016/j.jclinepi.2017.05.003] [Medline: 28502811]
29. Kasenda B, Liu J, Jiang Y, et al. Prediction of recruitment in randomized clinical trials (RECRUIT-IT)-rationale and design for an international collaborative study. Trials 2020 Aug 21;21(1):731. [doi: 10.1186/s13063-020-04666-8] [Medline: 32825846]
30. Olaisen RH, Manski RJ. Dental Utilization and Expenditures, US Civilian Noninstitutionalized Population Aged 2 and Older, 2019–2021. Statistical Brief (Medical Expenditure Panel Survey (US).: Agency for Healthcare Research and Quality (US); 2024 URL: https://www.ncbi.nlm.nih.gov/books/NBK603185/ [accessed 2026-02-24]
31. Pace WD, Dickinson LM, Staton EW. Seasonal variation in diagnoses and visits to family physicians. Ann Fam Med 2004;2(5):411-417. [doi: 10.1370/afm.73] [Medline: 15506572]

Abbreviations

ARC: administrative and resource center
CRF: case report form
IRB: institutional review board
NIDCR: National Institute of Dental and Craniofacial Research
NIH: National Institutes of Health
OHRQoL: oral health–related quality of life
PBRN: Practice-Based Research Network
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology

Edited by A Schwartz; submitted 22.Aug.2025; this is a non–peer-reviewed article;accepted 08.Jan.2026; published 27.Feb.2026.

Please cite as:
Cunha-Cruz J, Barton D, Cochran D, Crawford P, Fahimipour F, Funkhouser E, Geurs NC, Gilbert GH, Johnson JD, Smith N, Maninder K, Startley SJ, Stewart K, Waiwaiole L, National Dental PBRN Collaborative Group
The National Dental Practice-Based Research Network Dental Implant Restoration Registry to Evaluate Dental Implant Outcomes in Community Practice Settings: Protocol for a Prospective Observational Study
JMIR Res Protoc 2026;15:e82795
URL: https://www.researchprotocols.org/2026/1/e82795
doi:10.2196/82795

© Joana Cunha-Cruz, Danyelle Barton, David Cochran, Phillip Crawford, Farahnaz Fahimipour, Ellen Funkhouser, Nicolaas C Geurs, Gregg H Gilbert, James D Johnson, Ning Smith, Kaur Maninder, Sarah Jane Startley, Kimberly Stewart, Lisa Waiwaiole, National Dental PBRN Collaborative Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

World Health Organization–Recommended Periodic Presumptive Treatment Versus Doxycycline Post-Exposure Prophylaxis for Sexually Transmitted Infection Control Among Men Who Have Sex With Men in Kenya: Protocol for a Randomized Controlled Trial

Susan M Graham[1,2], MPH, MD, PhD; Fredrick O Otieno[3], MPH, PhD; Joshua Kimani[4], MPH, MBChB; Alé Barrientos[5], MS; Olusegun O Soge[5], MSc, PhD; Monisha Sharma[5], MS, PhD; Deven T Hamilton[6], MPH, PhD; Connie Celum[1,2], MPH, MD; R. Scott McClelland[1,2,7], MPH, MD; Eduard J Sanders[8], MPH, MD, PhD

[1]Department of Global Health, School of Medicine, University of Washington, Seattle, WA, United States

[2]Department of Medicine, School of Medicine, University of Washington, Seattle, WA, United States

[3]Nyanza Reproductive Health Society, Kisumu, Kenya

[4]Partners for Health and Development in Africa, Nairobi, Kenya

[5]Department of Global Health, School of Public Health, University of Washington, Seattle, WA, United States

[6]Center for Studies in Demography & Ecology, University of Washington, Seattle, WA, United States

[7]Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, United States

[8]Aurum Institute, Johannesburg, South Africa

Corresponding Author:
Susan M Graham, MPH, MD, PhD
Department of Global Health
School of Medicine
University of Washington
Hans Rosling Center
3980 15th Ave NE
Seattle, WA, 98105
United States
Phone: 1 206 351 0414
Fax: 1 206 744 6564
Email: grahamsm@uw.edu

Abstract

Background: Men who have sex with men (MSM) are at high risk for bacterial sexually transmitted infections (STIs), including gonorrhea, chlamydia, and syphilis, in Kenya. Because nucleic acid amplification testing (NAAT) is not widely accessible, most gonorrhea and chlamydia infections go undiagnosed and are treated only if symptomatic. World Health Organization (WHO)–recommended periodic presumptive treatment (PPT) and doxycycline post-exposure prophylaxis (doxyPEP) are both potential interventions to reduce the burden of bacterial STIs in this population. Neither has been rigorously tested among MSM in Africa.

Objective: This study aims to evaluate the effectiveness of WHO-recommended PPT versus doxyPEP, compared with standard syndromic treatment, in reducing the prevalence of bacterial STIs, including gonorrhea, chlamydia, and syphilis, among MSM in Kenya.

Methods: We are conducting an open-label randomized controlled trial with 2900 participants assigned in a 2:2:1 ratio to WHO-recommended PPT given every 3 months, doxyPEP taken 24-72 hours after condomless sex, or standard treatment. Sociodemographic, psychosocial, and behavioral data are collected by audio computer-assisted self-interview. Syphilis testing and treatment are provided as part of standard care. Throat and rectal swabs and urine are collected, pooled, and batch tested for gonorrhea and chlamydia by NAAT; these results are not used to guide treatment. The primary trial outcome is the combined prevalence of laboratory-diagnosed gonorrhea, chlamydia, and early syphilis after baseline; secondary outcomes include the prevalence of each pathogen individually and antimicrobial resistance in *Neisseria gonorrhoeae*. Primary and secondary outcomes

XSL•FO
RenderX

will be compared between each intervention and the common control arm by estimating relative risks over follow-up (months 3-18) using a modified Poisson model fitted with generalized estimating equations. We will also assess implementation outcomes, including acceptability, feasibility, and safety of each intervention compared to standard care among providers and patients using a mixed methods approach. Finally, we will evaluate the potential health and economic impact of scaling up WHO-recommended PPT and doxyPEP compared to standard of care on STI control among MSM and their partners in Kenya using a stochastic, network-based model and cost-effectiveness analysis on trial data.

Results: Enrollment commenced on October 29, 2025. As of November 25, 2025, a total of 357 participants (12.3% of target) have been enrolled, including 133 in Kisumu, 122 in Nairobi, and 102 in Mombasa. Full enrollment is expected to take 6 months, with follow-up occurring over 18 months per participant. Results will be published in 2028.

Conclusions: Results of this trial will provide critical data needed to inform guidelines to improve STI control among MSM in sub-Saharan Africa and other resource-limited settings where NAAT is not routinely available. Modeled estimates of the health and economic impact of scaling up these two interventions on STI control among MSM and their partners in Kenya will provide critical information to guide policymakers considering adoption of either intervention.

Trial Registration: ClinicalTrials.gov NCT06468462; https://clinicaltrials.gov/ct2/show/NCT06468462

International Registered Report Identifier (IRRID): PRR1-10.2196/81113

(JMIR Res Protoc 2026;15:e81113) doi:10.2196/81113

KEYWORDS

chlamydia; gonorrhea; men who have sex with men; MSM; sexually transmitted infections; STI; syphilis

Introduction

Gay and bisexual men and other men who have sex with men (MSM) face disproportionate risk for sexually transmitted infections (STIs), especially in resource-limited settings. The burden of STIs is greatest in low- and middle-income countries, and overlapping, synergistic epidemics of HIV and bacterial STIs in Africa have been recognized since the earliest days of the HIV epidemic [1-3]. Among MSM, STIs caused by *Neisseria gonorrhoeae* (NG) and *Chlamydia trachomatis* (CT) are significant causes of morbidity, ranging from urethritis, epididymitis, and proctitis to disseminated gonococcal infection [4]. Bacterial STIs can also lead to reproductive health problems among female partners of MSM, including pelvic inflammatory disease, chronic pelvic pain, tubal infertility, pregnancy complications, and fetal and neonatal death [1,5-10]. In sub-Saharan Africa, published prevalence estimates among MSM have ranged from 3%-11.5% for NG, 10%-14.5% for CT, and 1.1%-5.7% for early syphilis [11-17]. NG and CT increase the risk of HIV-1 transmission [18], with estimates from one study in the United States suggesting that 10.2% of all HIV infections among MSM are attributable to NG and CT [19]. Effective interventions to reduce the disproportionate burden of STIs among MSM in resource-limited settings are urgently needed.

Syndromic treatment is the standard of care for STI management in most of sub-Saharan Africa; it is less costly than etiologic testing [3], has relatively good sensitivity for male urethritis (87%-99%), and high cure rates when administered [20]. Unfortunately, syndromic treatment misses over 90% of STIs among MSM because they are asymptomatic, especially rectal and pharyngeal infections [12,21]. While laboratory-based or point-of-care nucleic acid amplification testing (NAAT) for NG and CT is widely used in well-resourced settings [22,23], NAAT for NG and CT is prohibitively expensive in African settings, at approximately US $19 per test conducted and US $97 per infection diagnosed, even if urethral, rectal, and pharyngeal samples are pooled [11]. In contrast, the estimated health expenditure per capita in Kenya in 2018 was US $88.39 [24]. The challenges with funding STI diagnostic tests are explicitly acknowledged in the 2022 World Health Organization (WHO) *Implementation Tool for PrEP for HIV Infection*, which recommends a stepwise approach to integration of STI services with HIV pre-exposure prophylaxis (PrEP) services, starting with syndromic management and integrating rapid diagnostic tests and eventually molecular tests depending on resource availability [25]. While the development of low-cost point-of-care diagnostics is an urgent priority [21,26], low-cost and feasible interventions to control STIs among MSM are needed now, as we await technology advances.

One intervention that holds promise to reduce STIs among MSM globally is doxycycline post-exposure prophylaxis (doxyPEP). In 2018, the IPERGAY substudy, a randomized open-label clinical trial of doxycycline 200 mg within 24-72 hours after condomless sex (doxyPEP) among MSM in France, reported a good safety profile, high levels of adherence, and a 47% relative reduction in new bacterial STIs among HIV PrEP users who also took 200 mg of doxycycline following every sexual encounter [27]. This reduction was driven by decreases in incident CT (70% reduction) and syphilis infections (73% reduction); there was no significant impact on NG [27]. Since that time, the DoxyPEP study, an open-label trial among MSM and transgender women living with HIV or taking PrEP in Seattle and San Francisco, found a 66% significant reduction in bacterial STIs in the PrEP cohort and a 62% reduction in the HIV cohort [28]. Reductions for specific infections in these cohorts were lower for NG than for CT and syphilis, at 55% and 57% for NG versus 88% and 74% for CT and 87% and 77% for syphilis, respectively [28]. More recently, the ANRS 174 DOXYVAC study in France, an open-label randomized trial of doxyPEP and meningococcal group B vaccine with a 2 × 2 factorial design, reported a significant reduction in all 3 STIs (ie, NG, CT, and syphilis) in the doxyPEP group with early

XSL•FO
RenderX

termination of the control arm. The adjusted hazard ratio in the doxyPEP arm was 0.17 (95% CI 0.12-0.26; *P*<.0001) for CT or syphilis (or both) but only 0.78 (95% CI 0.60-1.01; *P*=.06) for NG; there was no evidence that the vaccine was effective [29]. DoxyPEP as a patient-controlled intervention to reduce STIs could be effective in resource-limited settings, although the lack of quarterly NAAT to identify and treat "breakthrough" NG infections could limit effectiveness of this strategy. Because NG is the only STI that has been associated with increased HIV acquisition among MSM in Kenya [30], its effective control should be a priority. Research to evaluate the overall risk-benefit ratio of doxyPEP, in terms of efficacy, long-term safety, impact on antimicrobial resistance (AMR), and cost-effectiveness among MSM in resource-limited settings is needed.

One additional STI control strategy has been proposed. In 2011, the WHO recommended periodic presumptive treatment (PPT) to eradicate asymptomatic NG and CT infections for MSM and transgender women who report condomless receptive anal intercourse and either multiple sex partners or a sex partner with an STI in the past 6 months [4]. No specific frequency of this intervention was recommended, and readers were referred to syndromic treatment guidelines for regimen details [4]. Although targeting criteria for this recommendation have been evaluated by our research in Kenya [31,32], evidence of the efficacy of PPT among MSM is lacking, as is an evaluation of its potential downsides. Several studies on single-dose or monthly antibiotics among female sex workers in Asia and Africa demonstrated reduced disease burden using various PPT strategies [33-38]. Moreover, a meta-analysis of PPT intervention effects showed a statistically significant impact on the incidence of curable STIs in 13 of 14 studies included [39]. The premise that PPT could help reduce STI burden among MSM in resource-limited settings is supported by this evidence, yet an investigation of the potential impact, risks, and benefits of PPT for MSM in a randomized controlled trial has not been conducted. As a provider-delivered strategy that can be directly observed, PPT using cefixime and azithromycin, the currently recommended syndromic regimen for CT and NG [40], could prove more effective than doxyPEP if doxyPEP adherence is suboptimal or if AMR limits the impact of doxyPEP on NG control.

The potential emergence of AMR is an important concern for both doxyPEP and PPT. Following early studies of empiric single-dose antibiotics to reduce NG burden among female sex workers in the Philippines, some women (12%) reported self-prescribing antibiotics for empiric treatment, and high rates of penicillin-resistant NG were associated with self-prescription [41]. Today, multidrug resistant NG is a growing international public health issue [42-47]. In sub-Saharan Africa, reports of cephalosporin resistance in NG have been relatively rare (<1%) to date; however, high levels of tetracycline and quinolone resistance have been reported, with detection of plasmid-mediated tetracycline-resistant NG estimated at 73%-97% [42,43,48-50]. In Kenya specifically, no confirmed cephalosporin resistance has been reported [42,51,52], although one study reported increasing minimum inhibitory concentrations for ceftriaxone in isolates collected between 2002 and 2009 [48]. Of note, doxycycline resistance in NG is estimated at 56% in France [53,54], where doxyPEP had unclear efficacy against NG in the IPERGAY study, compared to 20%-30% in the United States, where the doxyPEP study demonstrated a significant decrease in NG incidence [28], but has led to an increase in tetracycline resistance in NG in the community [55]. Rigorous studies are needed to quantify the scope of potential resistance that might accompany doxyPEP or PPT in resource-limited settings with high levels of tetracycline resistance, to mitigate this critical problem.

In this study, we will use a hybrid type 1 implementation-effectiveness trial to address three objectives: (1) to evaluate the effectiveness of two interventions: WHO-recommended PPT given every 3 months and doxyPEP taken 24-72 hours after condomless anal sex, compared to standard syndromic treatment, for reducing the combined prevalence of laboratory-diagnosed gonorrhea, chlamydia, and early syphilis (primary trial outcome) among Kenyan MSM; (2) to assess implementation outcomes including the acceptability, feasibility, and safety of each intervention compared to standard care among providers and patients, using a mixed methods approach; and (3) to evaluate the potential health and economic impact of scaling up WHO-recommended PPT and doxyPEP compared to standard of care on STI control among MSM and their partners in Kenya using a stochastic, network-based model and cost-effectiveness analysis based on trial data.

Methods

Study Design

We will conduct an open-label randomized controlled trial with a hybrid type 1 implementation-effectiveness component, enrolling 2900 participants for 18 months of follow-up. The trial will evaluate 2 interventions (ie, PPT vs doxyPEP) versus the standard of care assigned in a 2:2:1 ratio.

Setting

The trial will be conducted at 3 research clinics in Kenya. In Nairobi, participants will be enrolled and followed at the Sex Workers Outreach Programme TRANSFORM Clinic, which is managed by Partners for Health and Development in Africa. In Kisumu, participants will be enrolled and followed at the Anza Mapema Clinic, which is managed by the Nyanza Reproductive Health Society. In Mombasa, participants will be enrolled and followed at Hapa-Kenya, a community-based organization serving the LGBTQ+ community, in partnership with the Pwani Research Center, University of Washington, University of Nairobi, and the Aurum Institute.

Population

To participate, individuals must be aged 18-29 years, be assigned male sex at birth, identify as male, report condomless anal intercourse and either multiple male sex partners or a male sex partner with a syndromic (urethritis, proctitis, or genital ulcer disease) or laboratory-diagnosed STI (NG, CT, or syphilis) in the past 6 months, and be willing and able to give written informed consent. Exclusion criteria include: allergy to cephalosporin, macrolide, or tetracycline class antibiotics; recent use of prolonged antibiotics (≥14-day course) in the month before enrollment; use of medications that impact cefixime,

XSL•FO
RenderX

azithromycin, or doxycycline metabolism; ongoing participation in another clinical trial; or active, clinically significant medical or psychiatric conditions that would interfere with study participation, at the discretion of the site investigator.

Recruitment and Retention

At each site, recruitment will be conducted by fliers, posters, information sessions, word of mouth, and social media apps (Facebook, WhatsApp) to increase awareness of the study at each research clinic and at local partner organizations in each site's catchment area. In addition, we will recruit participants through experienced peer outreach workers at previously mapped MSM hot spots, including bars, discos, restaurants, and social halls. Retention efforts will be supported by ensuring that the trial sites offer high-quality MSM-friendly services, offer flexible hours, monitor clinic flow and wait times, and provide HIV and STI treatment free of charge. In addition, we will use text or WhatsApp reminders 3 days and 1 day before a scheduled visit to promote timeline attendance. Tracing of participants who miss visits will be the responsibility of the outreach teams at each research site.

Procedures

Interested individuals will receive information on the study interventions and randomization and randomization process, then screened for eligibility. Sociodemographic data will be collected for all individuals screened, whether or not they are determined eligible or choose to enroll. Those who are eligible and choose to enroll will provide written informed consent. Participants will then be randomized in variable-sized blocks stratified by site, using REDCap (Research Electronic Data Capture; Vanderbilt University), a secure web application for building and managing databases for research studies [56]. Participants will attend quarterly visits, which are standard of care for HIV prevention and care services, over 18 months of follow-up. At each visit, we will collect participant data via REDCap using an audio computer-assisted self-interview (ACASI), with audio recordings available in English, Swahili, and Dholuo, depending on participant preference. Clinical data on symptoms, exam findings, specimen collection, and intervention delivery will be entered in REDCap data forms. Figure 1 presents a CONSORT (Consolidated Standards of Reporting Trials) flow diagram of the study, from enrollment and allocation through quarterly follow-up. Multimedia Appendix 1 provides the REDCap codebook for the different forms used, including the ACASI survey and clinical data forms; of note, many fields in the ACASI relate to the audio recordings in English, Swahili, and Dholuo, which follow each question and are made active using skip rules.

XSL•FO
RenderX

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flow diagram. ACASI: audio computer-assisted self-interview; CT: *Chlamydia trachomatis*; doxyPEP: doxycycline post-exposure prophylaxis; ITT: intention-to-treat; NG: *Neisseria gonorrhoeae*; PPT: periodic presumptive treatment; STI: sexually transmitted infection; WHO: World Health Organization.

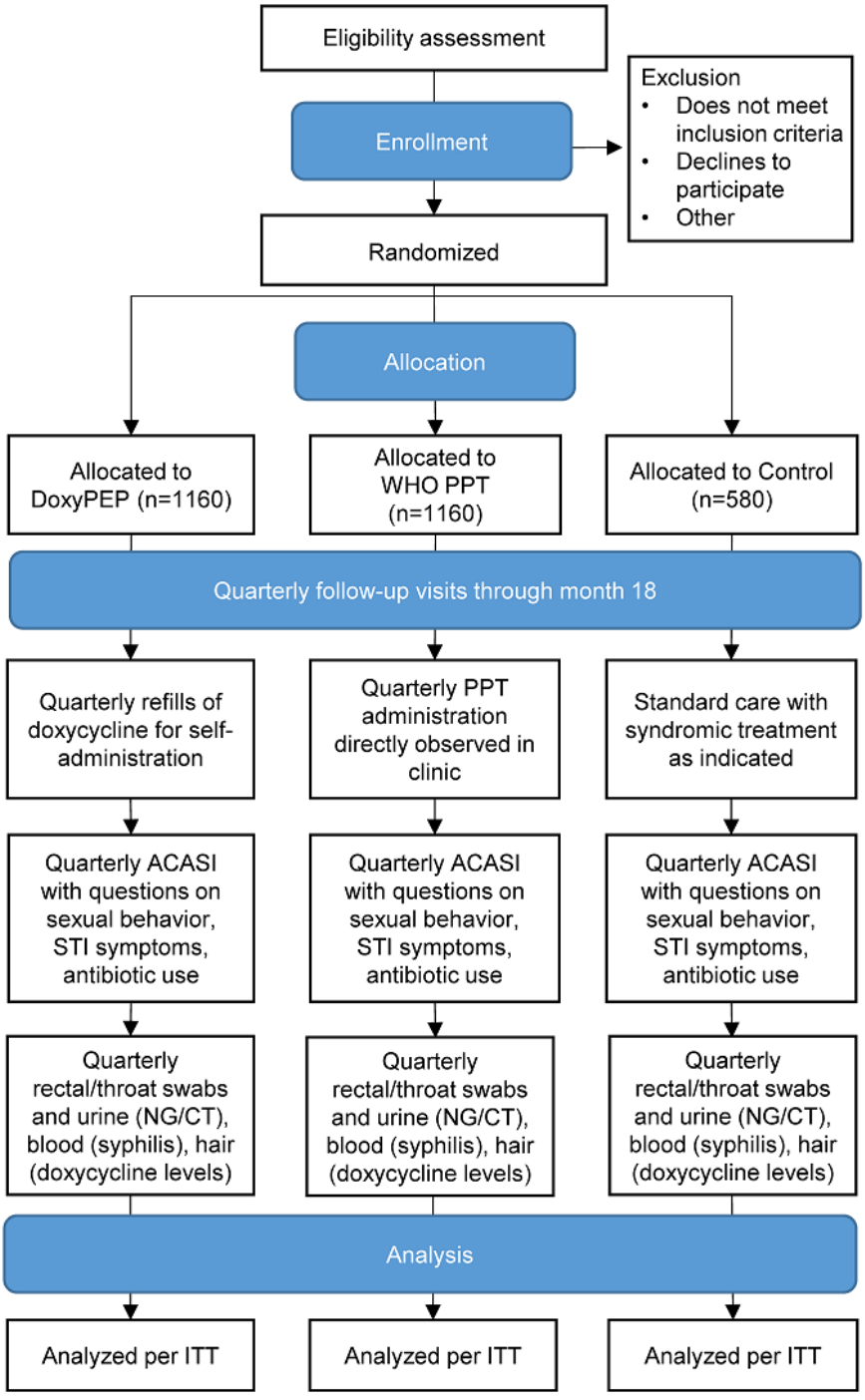

Study Arms

The study comprises three arms: PPT intervention, doxyPEP intervention, and standard care.

Periodic Presumptive Treatment Intervention

Participants assigned to the PPT arm will be evaluated at baseline and every 3 months thereafter for PPT eligibility based on having had condomless anal sex and either multiple sex partners or a sex partner with an STI in the past 6 months (at baseline) [32] or in the past 3 months (at follow-up visits). If eligible at a given visit, they will be offered 400 mg by mouth (po) cefixime plus 1 gram azithromycin by mouth under direct observation, using the same regimen as for syndromic treatment per the latest WHO and Kenyan recommendations [57].

XSL•FO
RenderX

Doxycycline Post-Exposure Prophylaxis Intervention

Participants assigned to the doxyPEP arm will be provided with a 30-dose supply (ie, 60 capsules) of doxycycline hyclate at each quarterly visit and given 1:1 counseling on the self-administration of 200 mg of doxycycline by mouth within 24-72 hours after condomless anal or vaginal sex as frequently as daily if indicated but not more than once daily, in accordance with the DoxyPEP trial in the United States [28]. Counseling to promote doxyPEP adherence will be repeated at each follow-up visit and any interim visits, and refills will be provided by the study team upon request by telephone, WhatsApp, or in person by participants in this arm. Of note, participants in the DoxyPEP trial used a median of 15 doses per quarter (IQR 4-30), with a similar number of doses used (17 doses per quarter, IQR 7-32) in the open-label extension [58].

Standard Care

Participants assigned to the standard care arm will receive no STI control intervention (ie, neither PPT nor doxyPEP). Standard care provided at every scheduled visit to all participants (regardless of study arm) includes syphilis testing, with treatment as indicated in accordance with Kenyan guidelines [59], and screening for symptoms of urethritis or proctitis. Participants with symptomatic urethritis or proctitis will receive syndromic treatment with 400 mg of cefixime and 1 gram of azithromycin by mouth under direct observation, in accordance with Kenyan guidelines [59]. This regimen will be updated if Kenyan treatment recommendations change.

Operational Components

Textbox 1 summarizes the operational components of the randomized controlled trial. Participants may return to the study clinic for interim visits at any time STI symptoms are present, regardless of study arm. All treatment will be documented, whether provided at a scheduled quarterly or an interim visit. Fidelity of delivery of each intervention will be closely monitored and documented, using a detailed standard operating procedure with a checklist of steps completed.

Textbox 1. Operational components of the randomized controlled trial.

Inclusion criteria

- Aged 18-29 years; assigned male sex at birth; identifies as male; reports condomless anal intercourse and either multiple male sex partners or a male sex partner with a syndromic (urethritis, proctitis, or genital ulcer disease) or laboratory-diagnosed sexually transmitted infection (STI, such as *Neisseria gonorrhoeae* [NG], *Chlamydia trachomatis* [CT], or syphilis) in the past 6 months; willing and able to give written informed consent.

Exclusion criteria

- Allergy to cephalosporin, macrolide, or tetracycline class antibiotics; recent use of prolonged antibiotics (≥14-day course) in the month before enrollment; use of medications that impact cefixime, azithromycin, or doxycycline metabolism; ongoing participation in another clinical trial; or active, clinically significant medical or psychiatric conditions that would interfere with study participation at the discretion of the site investigator.

Sample size

- A total of 2900 men who have sex with men (MSM); randomized 2:2:1 to directly observed periodic presumptive treatment (PPT) or doxycycline post-exposure prophylaxis (doxyPEP) versus standard care (syndromic STI treatment only).

Visit schedule

- Screening and enrollment, then quarterly for 18 months; reminders and tracing in case of missed visits to ensure high retention.

Standard of care package

- All participants: syphilis testing and treatment, syndromic STI treatment, partner notification per current guidelines as indicated, risk reduction counseling, condoms, lubricants, and standard HIV prevention or care services according to HIV status.

All study sites include programmatic HIV prevention and care services. During the informed consent process, participants will be asked permission to link their programmatic data to the trial records. Within these programs, HIV testing and counseling will be conducted for previously HIV-negative participants at each quarterly visit, per Kenyan guidelines for key populations. Antiretroviral therapy (ART) will be provided to participants with HIV, and PrEP will be provided to those who test HIV negative and are willing to take daily or intermittent tenofovir disoproxil fumarate-based oral PrEP.

Specimen Collection, Testing, and Storage

At each quarterly follow-up visit, participants will undergo hair collection for testing of doxycycline levels, blood collection for syphilis serology, and 3-site screening for NG and CT, with provider-collected or participant-collected (if provider collection is refused) pharyngeal and rectal swabs (2 of each) and a clean-catch urine. Participants who return to the study clinic for interim STI symptoms will undergo collection of the same pharyngeal, rectal, and urine samples, but no blood sample collection unless acute syphilis is suspected. We will not diagnose and treat infections other than syphilis in real time. Syphilis serology including rapid plasma reagin (Becton Dickinson) and, if reactive, *Treponema pallidum* hemagglutination assay (Biotec) will be tested at quarterly visits, with treatment provided to participants with a new syphilis diagnosis as soon as results are available. Participants who have symptomatic urethritis at any study visit (estimated at 5% of visits or ≈1015 visits) will undergo swab collection of urethral discharge for NG culture using modified Thayer-Martin Agar, a selective medium supplemented with nystatin, colistin, vancomycin, and trimethoprim that favors growth of NG while

XSL•FO
RenderX

suppressing genital and extragenital commensal microbial flora growth. Antimicrobial susceptibility testing will be performed using disk diffusion for penicillin and tetracycline and E-test for cefixime, azithromycin, and doxycycline, using WHO control strains for quality assurance of antimicrobial susceptibility testing at each of the site's microbiology laboratories [60].

All samples for NAAT will be transferred to and stored in the University of Washington–University of Nairobi East Africa STI Laboratory in Mombasa, which has operated with external quality assurance for >10 years. Aptima Combo 2 testing for NG and CT will be conducted in batches on a Hologic Panther platform by experienced laboratory staff blinded to randomization assignment. To reduce testing costs, we will pool the 3 samples from each participant for initial testing following published protocols [61,62]. All NG-positive pools will be broken down and individually tested to identify the infected site or sites [61]. All NG, CT, and early syphilis diagnoses will be reviewed by an Endpoint Adjudication Committee consisting of 2 STI experts on the University of Washington faculty who will be blinded to treatment arm. A random sample of 10% of all hair samples collected (≈812 samples) from the month 6 and month 18 visits in the doxyPEP arm will be shipped to the Hair Analytical Laboratory at the University of California at San Francisco for testing for doxycycline levels [63].

NG Antimicrobial Resistance Testing

The potential for selection of AMR in NG by PPT or doxyPEP is one of the greatest risks to the success of these strategies; however, the magnitude of this risk cannot be known without empiric data. Cephalosporin resistance in NG is happily still rare in Kenya. Azithromycin resistance in NG is uncommon in Kenya (0%-8.3% in recent studies) [42,51], and PPT may not increase it substantially given co-coverage with cefixime. Resistance to azithromycin in CT has not been seen, although it is known to be less effective at eradicating rectal chlamydia than doxycycline [64]. Plasma-mediated doxycycline resistance in NG is common and may impact the efficacy of doxyPEP [50].

Given the importance of AMR in NG as a public health problem, we will focus on this organism. We will perform NG culture and susceptibility testing of urethral swabs from symptomatic urethritis cases, as described above, due to the high yield of these specimens on culture [65], but will supplement this testing with rigorous molecular methods using residual NAAT samples and noncultured swab or urine specimens to identify well-characterized genetic AMR determinants conferring cefixime, azithromycin, or tetracycline resistance in NG. Of note, we have validated and successfully used these methods for strain typing and characterization of resistance determinants in NG [50,66,67].

We will select a random sample of 25% of all NG-positive specimens (≈761 samples) from each intervention arm (50% at baseline and 50% at follow-up) for shipment to the University of Washington for molecular testing. DNA will be purified and extracted from residual NAAT samples, swabs, or urine in the Soge Laboratory using the Total Nucleic Acid Isolation Kit on the Roche MagNA Pure 24 System (Roche Diagnostics), an automated nucleic acid purification system that has been validated to yield high-quality DNA from NAAT and clinical specimens [68]. Molecular characterization of NG will include strain typing using the internationally used *N. gonorrhoeae* multiantigen sequencing typing system [69]; detection of plasmid-mediated *tetM*, which confers high-level tetracycline resistance, and the chromosomally mediated mutation rpsJ V57M, which confers low-level tetracycline resistance [70]; detection of single-nucleotide polymorphisms associated with reduced cephalosporin (cefixime or ceftriaxone) susceptibility and resistance (mosaic penA alleles, *penA* A311V, *penA* N513Y, ponA L421P, mtrR −35delA, and porB G120K/A121D/N) [71-74]; and detection of azithromycin resistance (23S rRNA A2059G, C2611T) [72,75] by polymerase chain reaction and sequencing. Well-characterized positive control strains and negative controls will be included in all polymerase chain reaction assays [76].

Primary and Secondary Trial Outcomes

The primary trial outcome will be the combined prevalence of NG, CT, and early syphilis ascertained by laboratory diagnosis (eg, positive NAAT for NG or CT or new syphilis diagnosis, based on first positive rapid plasma reagin or a fourfold increase in nontreponemal titers). Secondary trial outcomes will be the prevalence of each pathogen individually and the prevalence of AMR to the different antibiotics assessed. Outcomes will be assessed at baseline and every 3 months for 18 months. We have selected prevalence instead of incidence as our outcome of interest because a major drawback of syndromic treatment is its failure to identify and treat asymptomatic infections that can therefore persist as a reservoir for transmission, despite spontaneous clearance of some infections [77].

Power and Sample Size

The trial will follow a standard randomized superiority design, with comparison of each intervention separately to the standard of care arm. To detect at 33% reduction in STI prevalence (from 16% to 10.8%) for each intervention compared to standard care, we will need 2465 participants total: 986 participants in each intervention group and 493 in the control group (α=0.05, β=0.20). Accounting for 15% attrition after randomization, we will need 2900 participants total: 1160 participants in each intervention group and 580 in the control group. We will target enrollment of ≈967 participants at each of our three research sites to attain the required sample size. For analysis of rates of AMR in NG, our power will be limited for very rare outcomes (eg, cefixime resistance), but we will have 80% power to detect an increase in tetracycline resistance from 80% to 86.7% or from 85% to 90.8% in each intervention arm compared to the control arm.

Statistical Analysis

The main trial analyses will follow intention-to-treat principles, including participants whether or not they accept or continue PPT or doxyPEP. Of note, in a small pilot program of ≈80 MSM cohort participants at a research clinic on the Kenyan coast in 2018-2019, acceptability of PPT at quarterly visits was over 90%. Bacterial STI diagnoses (NG, CT, and early syphilis) will be reported at baseline and each follow-up visit by arm and by

XSL•FO
RenderX

study site. The combined STI outcome will be compared between each intervention and the common control arm by estimating the relative risks of any STI over follow-up (months 3-18) using a modified Poisson model fitted with generalized estimating equations (GEE) methods to account for repeated observations for each participant, assuming an independent covariance structure, with site and study arm as the only covariates [78]. The test for significance will use a 2-sided α=.05, and 95% CIs will be calculated using robust standard errors. The same approach will be used for each STI individually.

Appropriate missing data methods (eg, imputation) or causal estimation of direct effects will be conducted to assess the potential influence of differential missingness by arm. Syndromic infections diagnosed and treated at interim visits after 28 days will be defined as infections for the current quarter and handled in the analysis as if they were present at the next scheduled visit. Per-protocol analyses will also be conducted, as well as analyses controlling for potential confounders such as HIV status if they are unbalanced across arms. Resistance to each antibiotic assessed (ie, penicillin, tetracycline, cefixime, azithromycin, and doxycycline) among NG cases will be tabulated by study arm, research site, site of infection, and AMR assessment method. The odds of resistance among NG cases with resistance results will be compared between each intervention and the common control arm using logistic regression, with repeated measures GEE if repeat infections occur. We will also analyze nonparametric trends over time in the proportion of NG isolates demonstrating resistance to each antibiotic in the study overall and by study arm. Additional analyses with adjustment for baseline STI, sexual risk behavior, and HIV status will also be performed.

Implementation Outcomes

Considering the unique needs of Kenyan MSM with respect to sexual health will be essential to developing effective strategies to address their STI burden. In addition, any STI control intervention must be acceptable, feasible, and safe for the patient. We will use a mixed methods approach to measure acceptability, feasibility, and safety of these interventions, using a conceptual model based on Proctor's Implementation Science Framework [79,80]. Figure 2 presents the conceptual model for this work, which includes assessment of implementation outcomes (acceptability, feasibility, uptake, adherence, and cost), service outcomes (effectiveness, AMR rates, and safety), and client outcomes (symptoms, sexual behavior, and satisfaction). While the main trial analysis will focus on effectiveness and AMR rates, we will also collect acceptability, feasibility, and safety data from providers and from participants in all arms. Secondary outcomes related to the PPT and doxyPEP interventions that will be assessed include uptake and adherence (related to acceptability and sustainability of the interventions) and client outcomes including satisfaction, symptoms, and sexual behavior. The quantitative and qualitative measures we will use to assess these outcomes are presented in Table 1 [79]. Participants will be asked questions about acceptability (the 4-item Acceptability of Intervention Measure), feasibility (the 4-item Feasibility of Intervention Measure), and doxyPEP uptake and adherence in the ACASI instrument (Multimedia Appendix 1). In addition, their eligibility for and receipt of PPT (PPT arm participants) and acceptance of doxyPEP refills (doxyPEP arm participants), as well as symptoms reflecting potential side effects, will be captured in a clinical data form. To capture staff views and any challenges that arise, minutes will be taken at monthly staff meetings at each site. At the end of the study, acceptability and feasibility for providers will also be assessed using the Acceptability of Intervention Measure and the Feasibility of Intervention Measure in a questionnaire administered to research staff [81].

Figure 2. Conceptual model based on Proctor's Implementation Science Framework. The conceptual model for assessment of each intervention tested includes assessment of implementation outcomes (acceptability, feasibility, uptake, adherence, and cost), service outcomes (effectiveness, antimicrobial resistance [AMR] rates, and safety), and client outcomes (symptoms, sexual behavior, and satisfaction). doxyPEP: doxycycline post-exposure prophylaxis; PPT: periodic presumptive treatment; STI: sexually transmitted infection.

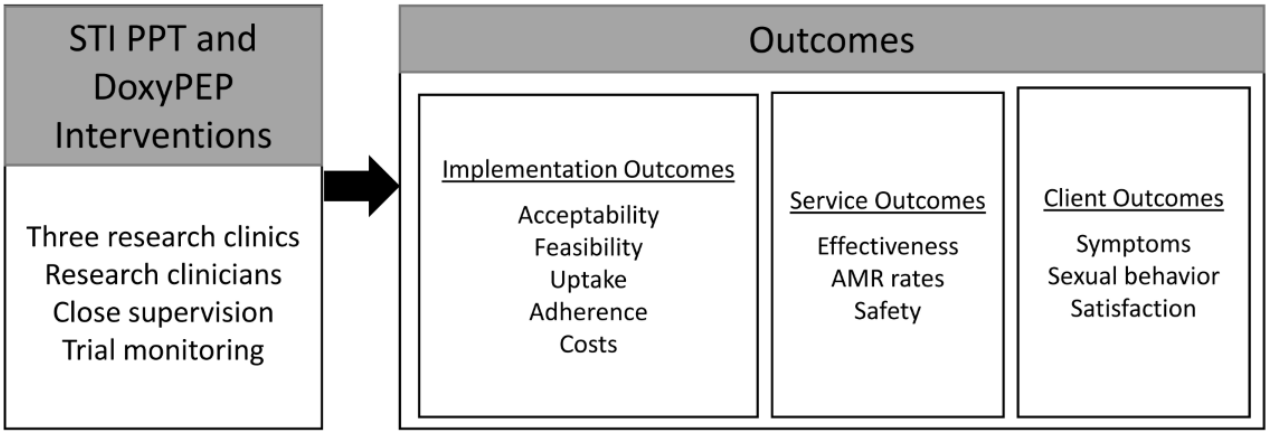

XSL•FO
RenderX

Table 1. Conceptual model with implementation, service, and client outcome measures.

Outcome type and specific outcome	Approach	Result
Implementation		
Acceptability	Brief questionnaire for eligible individuals who decline to enroll; ACASI[a] questions to participants; brief exit interviews; staff meeting minutes and questionnaire	Understanding how well PPT[b] and doxyPEP[c] are received by and are perceived to meet the needs of the target population
Feasibility	ACASI questions to participants; staff meeting minutes and questionnaire	Understanding how easily PPT and doxyPEP can be carried out
Uptake	Acceptance of directly observed PPT or of doxycycline tablets at study visits	Estimated PPT and doxyPEP acceptance at baseline documented in clinic records
Adherence	Administration records for PPT participants; ACASI questions on doxyPEP use and adherence; doxycycline levels in hair for doxyPEP participants	Estimated PPT and doxyPEP acceptance over time documented in clinic records
Costs	Microcosting; time-motion surveys (described below)	Cost to system, cost to patient
Service		
Effectiveness	Impact on STI[d] prevalence and on AMR[e] (trial results)	Reduction in STI prevalence relative to increased AMR
Safety	Rigorous monitoring of safety, including social harms	Adverse events attributed to PPT or doxyPEP; discontinuation due to adverse events
Client		
Symptoms	Data on symptoms of potential side effects	Understanding how symptoms differ by arm
Sexual behavior	ACASI questions on sexual behavior and antiretroviral use; brief exit interviews	Understanding of any behavioral disinhibition with PPT or doxyPEP
Satisfaction	Brief exit interviews	Understanding how well PPT and doxyPEP met the needs of the target population

[a]ACASI: audio computer-assisted self-interview.

[b]PPT: periodic presumptive treatment.

[c]doxyPEP: doxycycline post-exposure prophylaxis.

[d]STI: sexually transmitted infection.

[e]AMR: antimicrobial resistance.

Quantitative Data Collection

Due to its enhanced privacy, ACASI will be the main method of data collection, in English, Dholuo, or Kiswahili. We have successfully used ACASI to capture data in multiple projects at our research sites [30,82-85]. Because participants occasionally have difficulties using ACASI, a research assistant will be available to assist participants if needed. Measures that will be collected are based on our conceptual model, extensive experience conducting research among Kenyan MSM, and published literature on correlates of sexual behavior, HIV and STI incidence, and medication adherence. Table 2 presents an overview of data elements, content, tools for assessment, frequency of assessment, and applicable study arm.

XSL•FO
RenderX

Table 2. Quantitative data elements collected from participants: content, collection method, frequency, and arm.

Category and data element	Method and content	Frequency and arm
Primary trial outcome variable: NG[a], CT[b], or early infection	Aptima NAAT[c]: pooled specimen (throat, rectal, and urine) positive for NG or CT; Syphilis testing: first positive rapid plasma reagin or a fourfold increase in nontreponemal titers) for early syphilis	All arms: quarterly and at interim visits for STI[d] symptoms
Secondary trial outcome variables		
Each infection individually	Same as above	All arms: quarterly and at interim visits for STI symptoms
AMR[e]	Resistance to penicillin, tetracycline, cefixime, azithromycin, or doxycycline by culture of urethral discharge (symptomatic urethritis cases) or molecular testing (25% of all NG-positive samples)	All arms: quarterly and at interim visits for STI symptoms
Implementation outcome variables		
Acceptability	ACASI[f]: Acceptability of Intervention Measure [81]	All arms: quarterly
Feasibility	ACASI: Feasibility of Intervention Measure [81]	All arms: quarterly
Uptake	Direct observation: administration of cefixime and azithromycin in PPT arm; acceptance of doxycycline in doxyPEP[g] arm	Intervention arms: quarterly
Adherence	ACASI: self-report questions on adherence to doxyPEP; direct observation of PPT as above; hair doxycycline levels.	Intervention arms: quarterly
Safety including social harms	Study monitoring: adverse events and social harms reported by participants or study staff	All arms: quarterly
Symptoms	ACASI: urethral or rectal symptoms since past visit; whether STI treatment was sought outside the study; side effects of medications	All arms: quarterly
Sexual risk behavior	ACASI: anal intercourse, vaginal intercourse, condom use, partner numbers, partner HIV status, concurrency, sexual networks, transactional sex, intimate partner violence	All arms: quarterly
PrEP or ART use and adherence	ACASI: self-reported use of and adherence to PrEP (if HIV-negative) or ART[h] (if HIV-positive); program records: chart data to confirm use if consents to linkage	All arms: quarterly
Study retention	Study records: scheduled visits completed within 1-month window	All arms: quarterly
Potential correlates of study outcomes		
Demographics	ACASI: age, education, religion, marital status, sexual orientation, employment	All arms: baseline
Clinic access	ACASI: transportation costs, wait times	All arms: quarterly
STI risk perception	ACASI: “What do you think is your risk of getting a sexually transmitted infection?” (4 item Likert scale)	All arms: quarterly
HIV status	ACASI: self-reported HIV status; Program Records: programmatic test results if consents to linkage	All arms: quarterly
Antibiotic use	ACASI: antibiotic use outside the study in the past 3 months	All arms: quarterly
Substance use	ACASI: Alcohol Use Disorders Identification Test [86,87], Drug Abuse Screening Test [88]	All arms: quarterly

XSL•FO
RenderX

Category and data element	Method and content	Frequency and arm
Childhood abuse	ACASI: Childhood Experience of Care and Abuse [89]	All arms: baseline
Sexual stigma	ACASI: sexual stigma, as in our previous work [90,91]	All arms: every 6 months
Depressive symptoms	ACASI: Patient Health Questionnaire-9 [92]	All arms: quarterly

[a]NG: *Neisseria gonorrhoeae.*

[b]CT: *Chlamydia trachomatis.*

[c]NAAT: nucleic acid amplification testing.

[d]STI: sexually transmitted infection.

[e]AMR: antimicrobial resistance.

[f]ACASI: audio computer-assisted self-interview.

[g]doxyPEP: doxycycline post-exposure prophylaxis.

[h]ART: antiretroviral therapy.

Qualitative Data Collection

During study screening, a brief questionnaire with 2 open-ended questions will be used to capture reasons for declining to enroll, including poor individual-level acceptability of PPT, doxyPEP, or randomization. During the trial, qualitative data collection will include brief exit interviews with a subset of 30 participants in each intervention arm (10 at each site), to understand participant-level acceptability and satisfaction with the interventions. Interviews will be conducted using participants' preferred language (English, Dholuo, or Kiswahili). Participants will be sampled for these interviews to include a range of ages and reported risk behaviors (eg, transactional sex) using a stratified-purposive sampling approach. At any visit on which a participant withdraws, a similar exit interview will be conducted to obtain feedback on the intervention received. Interviews will be recorded if permission is granted; detailed notes will also be taken to document participant feedback. The topic guide for these interviews is included in Multimedia Appendix 2.

Quantitative Analysis of Implementation Outcomes

GEE with appropriate links and robust standard errors will be used to test retention across all study arms and to evaluate associations between study arm and the quantitative implementation outcomes collected for both interventions (eg, acceptability and feasibility measures, uptake, adverse events, symptoms, and sexual risk behavior). Factors associated with adherence to doxyPEP will be analyzed in the population for whom hair samples are tested, based on semiquantitative assessment of high versus low doxycycline concentrations [63]. For all quantitative analyses, patterns of missing data will be evaluated, and multiple imputation will be used where appropriate.

Qualitative Analysis of Implementation Outcomes

Exit interview notes and transcripts and staff meeting minutes will be translated into English if necessary and imported into Dedoose (SocioCultural Research Consultants) qualitative analysis software. These documents will be reviewed separately by 2 investigators for completeness and initial theme generation. Coding and analysis will focus on participant-level acceptability and satisfaction and on staff-level acceptability and feasibility, starting with a deductive approach informed by the theoretical framework of acceptability, a recently developed approach to acceptability research of medical interventions [93]. The components of the theoretical framework of acceptability are affective attitude, burden, ethicality, intervention coherence, opportunity costs, perceived effectiveness, and self-efficacy. Additional codes emerging from an inductive process will be added as they arise. Codes will be reviewed for consistency and intercoder agreement with consensus code generation by themes. Descriptive thematic analyses will be done separately for individuals who refuse participation, trial participants (including those who withdraw), and staff. Within these groupings, results will be triangulated with relevant quantitative data using joint displays.

Cost Analysis

During the field work, we will conduct microcosting, semistructured interviews or focus group discussions with staff regarding time spent on different work activities, and time and motion observation of the intervention and standard of care treatments within the randomized controlled trial. Costs will be divided into direct costs of diagnostic evaluation or testing, direct costs of treatment, and indirect costs to patients. Research time (eg, informed consent) will be removed from programmatic costs. We will estimate the incremental cost per participant as the total annual program costs divided by the number of participants. The topic guide for staff is included in Multimedia Appendix 3.

Modeling Impact

To assess impact, we will develop a stochastic, network-based model of NG, CT, and HIV transmission in Kenya. Because syphilis is relatively uncommon among Kenyan MSM and its transmission has not yet been extensively modeled, our model will not include syphilis. The general modeling approach will follow that developed by Jenness et al [94] for modeling NG, CT, and HIV transmission among MSM in the United States. The model will simulate a population of 100,000 individuals distinguished by the following attributes: age; partner preference; sexual role preference (insertive, receptive, or versatile); relationship status; NG, CT and HIV status; diagnosis status; time since infection; treatment status; and HIV viral load,

XSL•FO
RenderX

among possible relevant others revealed by our fieldwork. Individuals can undergo any or all of the following transitions: enter the sexually active population; age; form or dissolve relationships of different types; become infected with NG, CT, or HIV; experience STI or HIV symptoms; test; become diagnosed; undergo treatment; discontinue ART treatment; reinitiate ART treatment; age out of the population; or die due to background age-specific mortality or HIV-related mortality. Within each partnership, individuals will engage in anal intercourse with coital frequency based on a Poisson model fit to study participant data, use condoms based on a binomial model also fit to study participant data, choose a sexual role (insertive or receptive) for each act of anal intercourse based on individual preferences, and disclose STI or HIV status. Network dynamics will be modeled using the *statnet* package in R to estimate exponential-family random graph models and their dynamic extension temporal exponential-family random graph models, which provide a foundation for statistically principled simulation of local and global network structure given a set of target statistics from empirical data [95]. Data sources and parameters will be based on the latest available literature at the time of model building. Sensitivity analyses will be conducted, especially around key parameters of the individual-level intervention efficacy. Comparison will also be made to NAAT-guided treatment as done in well-resourced settings. Model outcomes will include life-years lived with untreated NG or CT, incident NG and CT cases averted, and disability-adjusted life years averted.

Cost-Effectiveness Analysis

After the trial is completed, mathematical models will be used to simulate long-term health and cost outcomes based on study data and to consider clinical outcomes beyond the scope and time horizon of the trial. Using cost and effectiveness outcomes estimated by the model, we will calculate the cost-effectiveness of PPT and of doxyPEP as interventions. The cost-effectiveness analysis will follow WHO guidelines and Cost of Implementing New Strategies recommendations [96,97], and report summary estimates that allow comparison to standard-care syndromic management. In addition, a comparison to NAAT-guided treatment (as done in well-resourced settings) will be made. Results of the cost-effectiveness analysis will provide important information to decision makers charged with defining priorities and allocating resources.

We will combine effectiveness estimates from the updated stochastic model with the cost estimates collected from each intervention to calculate cost per NG or CT case identified and treated and the incremental cost-effectiveness ratio measured as cost per incident NG or CT case averted and cost per disability-adjusted life year averted. Disability weights for calculation of disability-adjusted life years will be obtained from the latest global burden of disease study [98]. We will calculate the incremental cost-effectiveness ratio associated with PPT and with doxyPEP as the ratio of difference in costs divided by difference in disability-adjusted life years compared to the next most costly strategy over a 10-year horizon. Following updated guidelines recommending use of a supply-side threshold to reflect the opportunity cost of additional health investment, PPT and doxyPEP will be considered cost-effective if the incremental cost-effectiveness ratio is ≤US $500 per disability-adjusted life year averted [99]. We will use both the governmental and societal perspectives. Univariate and probabilistic sensitivity analyses will be performed to determine model robustness and the impact of varying different parameters through their plausible ranges on the estimate of cost-effectiveness.

Data Safety and Monitoring

Adverse events will be recorded in study documentation, and serious adverse events will be reviewed for relatedness to study products or procedures by the study leadership team (the principal investigator of record and a lead investigator at each study site). Social impacts and any harms of study participation will also be assessed. An independent data safety and monitoring board composed of an experienced trial biostatistician and 2 internationally regarded clinical researchers, one Kenyan and the other Canadian, has been established to review study progress, including participant accrual, visit attendance, retention, and collection of primary and secondary endpoint data. This board will also review safety and outcome data in a closed report on a biannual basis to ensure safety and equipoise. The data safety and monitoring board will make recommendations based on their review of safety and efficacy data at biannual meetings. The trial will be terminated early if the incidence of serious adverse events related to either intervention exceeds 2% and is statistically significantly higher than the rate observed in the standard care arm (P<.01) in 2 consecutive interim (ie, every 6 months) safety analyses.

Ethical Considerations

The protocol, consent forms, participant education and recruitment materials, and any subsequent modifications were reviewed and approved by the Kenyan Medical Research Institute's Ethical Review Committee (Protocol 4962) and the Human Subjects Division of the University of Washington (STUDY00018588). The trial protocol (currently version 1.1, dated June 10, 2024) and any approved modifications will be posted at ClinicalTrials.gov. Trained research staff obtain written informed consent from all trial participants, for all exit interviews, and from staff who participate in feedback on acceptability and feasibility of the study interventions, following standardized protocols. Confidentiality of personal information is strictly maintained. Participant identifiers are not recorded in the study database but are included in a link log at each site. The link log is used to record participant names, contact information, and study ID numbers and will be maintained until participant enrollment and follow-up are complete for the purposes of avoiding duplicate enrollments, assisting with retention, and tracking participants with positive syphilis results. Participants are reimbursed 1000 KSh (approximately US $7.72) per visit. Clinical trial insurance and protocols to address social harms are in place to mitigate any adverse events resulting from trial participation.

Results

During community engagement activities in preparation for the trial, the study was named the "*Mambo Matatu*" (Swahili for "three ways") study by the gay, bisexual, and other MSM

XSL•FO
RenderX

communities served by our research clinics. Regulatory approval was obtained from the Kenya Pharmacy and Poisons Board in December 2024, and a research permit was obtained from the Kenyan National Commission for Science, Technology, and Innovation in March 2025. After a delay due to late receipt of year 2 funding from the US National Institutes of Health, enrollment commenced on October 29, 2025. As of November 25, 2025, a total of 357 participants (12.3% of target) have been enrolled, including 133 in Kisumu, 122 in Nairobi, and 102 in Mombasa. Full enrollment is expected to take 6 months, with follow-up occurring over 18 months per participant. Results will be published in 2028. Table 3 presents the participant timeline as a schedule of enrollment, interventions, and assessments.

Table 3. Participant timeline: schedule of enrollment, interventions, and assessments.

Category	Trial period timepoints[a]							
	Enrollment		Postrandomization					Close-out
	$-t_i$ to 0[b]	Month 0	Month 3	Month 6	Month 9	Month 12	Month 15	Month 18
Enrollment								
Eligibility screen	✓							
Informed consent	✓							
Locator information	✓							
Randomization		✓						
Intervention or comparator								
doxyPEP[c] provision[d]		✓	✓	✓	✓	✓	✓	✓
PPT[e]		✓	✓	✓	✓	✓	✓	✓
Standard care[f]		✓	✓	✓	✓	✓	✓	✓
Assessments								
ACASI Survey		✓	✓	✓	✓	✓	✓	✓
Symptom review and clinical exam		✓	✓	✓	✓	✓	✓	✓
Specimen collection		✓	✓	✓	✓	✓	✓	✓
Syndromic treatment if indicated		✓	✓	✓	✓	✓	✓	✓
Syphilis treatment if indicated		✓	✓	✓	✓	✓	✓	✓
Adverse event monitoring		✓	✓	✓	✓	✓	✓	✓
Exit interview								✓

[a]Acceptable visit window ±30 days.

[b]Pre-enrollment.

[c]doxyPEP: doxycycline post-exposure prophylaxis.

[d]Supply of 60 capsules of doxycycline are provided at each scheduled visit, with refills provided upon request between visits.

[e]Periodic presumptive treatment; provided when participant meets eligibility criteria (ie, having had condomless anal sex and either multiple sex partners or a sex partner with a sexually transmitted infection (STI) in the past 6 months at baseline [32] or in the past 3 months at follow-up visits).

[f]Limited to syndromic treatment when symptoms of an STI are present. Note that syndromic treatment is provided to participants in each study arm per Kenyan STI treatment guidelines when indicated.

Discussion

Overview

This study of two interventions to control bacterial STIs including CT, NG, and syphilis among MSM in African settings will yield important insights into how to improve STI control in resource-limited settings while balancing benefits, risks, and costs. The potential benefits of implementing either doxyPEP or PPT for the control of CT, NG, and syphilis among MSM in African settings are substantial; at the same time, these strategies are untested, and potential risks—including adverse events, increased AMR, risk compensation impacting sexual behavior, and increased costs—could counteract these benefits. Only a well-designed randomized trial comparing these two interventions to a standard care arm can help address these questions and determine whether the benefits outweigh risks of adopting either strategy. While we will be able to compare the two interventions to each other in addition to the primary analysis comparing each intervention to standard care, it is the comparison to standard care that is of greatest interest for policymakers, who will need to consider not only each intervention's effectiveness, but also its acceptability, feasibility, and disadvantages in terms of safety, AMR, and costs relative to the current standard of care in determining the overall value added if either intervention is adopted.

XSL•FO
RenderX

We have carefully considered our study design and research approach to optimize feasibility while maintaining high rigor and ensuring that results will be reproducible. Kenya is an ideal setting for evaluating PPT and doxyPEP among MSM, as the Government of Kenya's visionary 20-year plan for evidence-based HIV and STI prevention, the *Kenya HIV Prevention Revolution Road Map: Count Down to 2030*, aims to drive new HIV infections toward zero between 2013 and 2030 using effective HIV and STI prevention initiatives [100]. As HIV transmission decreases, risk has become more concentrated in key populations such as MSM, who also experience high rates of bacterial STIs. The Kenyan Ministry of Health endorses STI control in key populations, has worked to address health disparities among MSM, and supports this proposal, ensuring that our trial results will be linked to policy. In addition, each research site has collaborative relationships with MSM-led community-based organizations with large membership rosters, laying the groundwork for community feedback on the proposed work and its implementation and for successful outreach, recruitment, and retention.

We are using an oral regimen of cefixime and azithromycin instead of the Centers for Disease Control and Prevention's currently recommended STI treatment regimen consisting of ceftriaxone by intramuscular injection and doxycycline by mouth [64]. We decided against the use of this regimen due to concerns about low acceptability of regular injections for presumptive treatment and because most low- and middle-income countries still using cefixime and azithromycin have not seen the same levels of AMR as in high-income countries, likely due the limitation of treatment to syndromic infections [101,102]. Azithromycin provides both CT coverage and dual coverage of most NG. Although azithromycin resistance can emerge in NG due to the long half-life of this drug [103,104], data from the NG surveillance programs at the WHO and Centers for Disease Control and Prevention suggest that cefixime resistance decreased after the addition of azithromycin to the NG treatment regimen [105,106]. In contrast, a recent study of our isolates in Kenya suggested that doxycycline can select for plasmid-mediated penicillin resistance only or both plasmid-mediated penicillin resistance and high-level plasma-mediated tetracycline resistance and therefore could potentially increase the risk of cephalosporin resistance [42]. Therefore, it is critical to evaluate NG AMR to all 3 antibiotics in all study arms.

Of note, the WHO did not specify a frequency for PPT assessment and delivery in their recommendations [4]. We debated offering this intervention on a 6-month basis, since a 6-month recall period is used for eligibility, but felt this might weaken the intervention's efficacy. We will therefore provide PPT at each quarterly visit on which participants in the PPT arm are eligible and assess eligibility for PPT based on sexual behavior in the past 3 months. We anticipate being able to explore the potential efficacy of different frequencies of PPT administration in our modeling work. A consideration for both PPT and doxyPEP is the treatment of uninfected individuals. In our prior work, we have estimated a number needed to treat one infection with PPT ranging from 4-12 [31,32]. In the DoxyPEP study in US settings with similar STI incidence as in Kenya, the number needed to treat was estimated at 5 [28]. Treatment of uninfected individuals occurs in the setting of epidemiologic treatment of partners; for example, a recent study estimated that 65% of individuals receiving expedited partner treatment (often with cefixime by mouth) are asymptomatic and NAAT negative [107]. An important aspect of the proposed research is delivery of PPT or doxyPEP in a setting in which the vast majority of bacterial STIs are currently undiagnosed and therefore left untreated.

This study is particularly timely given recent evidence related to doxyPEP in particular and to STI control in general that has emerged. Recent publications have documented that presumptive treatment is common even in settings with NAAT and that both overtreatment and undertreatment occur at varying rates across studies and care settings [108,109]. In settings where NAAT is possible and follow-up is high, evidence supports testing first to avoid unnecessary ceftriaxone use for STI symptoms in MSM [110] and to avoid unnecessary antibiotic treatment in asymptomatic partners exposed to CT or NG [111]. Rising concerns about frequent NAAT and high levels of antibiotic consumption among MSM led to a randomized trial in Belgium that compared quarterly 3-site screening to not screening of asymptomatic MSM for NG and CT at all [112]. This "Gonoscreen" study demonstrated that frequent screening lowered CT rates but led to higher antibiotic consumption and had no effect on NG incidence [112]. Available evidence suggests that mass treatment of CT, NG, and syphilis in high prevalence populations such as MSM has only resulted in temporary reductions in the prevalence of these infections while likely increasing AMR [113]. DoxyPEP has, not unexpectedly, been associated with increases in antimicrobial resistance in NG as well as in other organisms present in the microbiome [55,114-116], leading to concern about the long-term benefits and harms of widespread use [117-119]. For example, the estimated percentage of NG isolates carrying the plasmid-borne gene *tetM*, which confers high-level tetracycline resistance, increased from less than 10% in 2020 to more than 30% in the first quarter of 2024 [120]. While modeling has shown that doxyPEP could avert a substantial portion of all STI diagnoses among at-risk MSM [121], a study that modeled the impact of doxyPEP on NG found that after initial decreases in NG prevalence and incidence, resistance would spread quickly, resulting in the total loss of doxyPEP's clinical efficacy against NG within 10-20 years [122]. While continued impact of doxyPEP on CT and syphilis [123] will likely provide benefit for this strategy, evaluating other possible interventions, such as PPT, that are less costly and would lead to less overall antibiotic consumption is an important consideration.

While there are many important strengths of this study, it is not without limitations. First, selection bias is likely given that we will enroll MSM who accept research screening and participation. To mitigate this bias, we will collect data on characteristics of those who decline trial participation. Second, our results may not be generalizable to all geographic areas of Kenya, although data from earlier studies among MSM in Kenya suggest that the behavior risk profiles, sexual networks, and social interactions of MSM across the 3 sites are similar. Third, contamination can be a problem, so we will collect data on

XSL•FO
RenderX

receipt of antibiotics in the past 30 days to monitor for this, especially in the standard care arm. Fourth, providing PPT or doxyPEP to intervention arm participants may reduce STI rates in the control arms by reducing NG and CT prevalence within sexual networks. We have data on STI prevalence at these sites prior to the study, which we can use to identify temporal decreases in STI rates [124]. Fifth, data we collect on stigmatized behaviors, such as condomless receptive anal intercourse and substance use, may be underreported due to social desirability bias. Our experience with ACASI in prior studies suggests that this method of data collection may yield more frank responses to questions on sensitive topics [82]. Sixth, there may be differential retention given the open-label design. We will try to address attrition bias by minimizing the burden of study visits and of each intervention on participants, keeping procedures within a reasonable timeframe, and monitoring clinic flow. Finally, we have not included testing for resistance in other organisms such as *Staphylococcus aureus* or *Streptococcus pneumoniae* due to a ceiling on our allowed budget. We will seek additional funding for additional testing of stored swabs and hair samples to explore the relationship between PPT and doxyPEP and AMR in off-target bacteria.

Conclusions

The prevalence and incidence of bacterial STIs, including syphilis, NG, and CT infections, among MSM remain high in resource-limited settings such as Kenya. Results of this trial will provide critical data needed to inform guidelines to improve STI control among MSM in sub-Saharan Africa and other resource-limited settings where NAAT is not routinely available. Trial data will be used to model the health and economic impact of scaling up WHO-recommended PPT and doxyPEP compared to standard care on STI control among MSM and their partners in Kenya using a stochastic, agent-based model and standard cost-effectiveness evaluation. Modeled estimates of the health and economic impact of scaling up these two interventions on STI control among MSM and their partners in Kenya will provide critical information to guide policymakers.

Acknowledgments

We are grateful for the support of community-based organizations serving the gay, bisexual, and other men who have sex with men (MSM) communities in Kisumu, Nairobi, and Mombasa and for County Health Departments.

Authors' Contributions

SMG contributed to the conceptualization, writing of the original draft, funding acquisition, methodology, and supervision. FOO contributed to writing, reviewing and editing, as well as methodology. JK contributed to writing, reviewing and editing, and methodology. AB contributed to writing, reviewing and editing. OOS contributed to writing, reviewing and editing, and methodology. MS contributed to writing, reviewing and editing, and methodology. DTH contributed to writing, reviewing and editing, and methodology. CC contributed to writing, reviewing and editing, and methodology. RSM contributed to writing, reviewing and editing, and methodology. EJS contributed to writing, reviewing and editing, funding acquisition, and methodology.

Conflicts of Interest

RSM receives research funding, paid to the University of Washington, from Hologic Corporation. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1
REDCap project with all data forms.
[PDF File (Adobe PDF File), 3891 KB - resprot_v15i1e81113_app1.pdf]

Multimedia Appendix 2
Exit interview topic guide.
[DOCX File , 31 KB - resprot_v15i1e81113_app2.docx]

Multimedia Appendix 3
Staff feedback guide.
[DOCX File , 30 KB - resprot_v15i1e81113_app3.docx]

Multimedia Appendix 4
Peer-review report from HSQE Health Services: Quality and Effectiveness Study Section, Healthcare Delivery and Methodologies Integrated Review Group (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 161 KB - resprot_v15i1e81113_app4.pdf]

References

XSL•FO
RenderX

1. Steen R, Wi TE, Kamali A, Ndowa F. Control of sexually transmitted infections and prevention of HIV transmission: mending a fractured paradigm. Bull World Health Organ 2009;87(11):858-865 [FREE Full text] [doi: 10.2471/blt.08.059212] [Medline: 20072772]
2. Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al. Chlamydia, gonorrhoea, trichomoniasis and syphilis: global prevalence and incidence estimates, 2016. Bull World Health Organ 2019;97(8):548-562P [FREE Full text] [doi: 10.2471/BLT.18.228486] [Medline: 31384073]
3. Terris-Prestholt F, Vyas S, Kumaranayake L, Mayaud P, Watts C. The costs of treating curable sexually transmitted infections in low- and middle-income countries: a systematic review. Sex Transm Dis 2006;33(10 Suppl):S153-S166. [doi: 10.1097/01.olq.0000235177.30718.84] [Medline: 17003680]
4. Guidelines: prevention and treatment of HIV and other sexually transmitted infections among men who have sex with men and transgender people: recommendations for a public health approach. In: World Health Organization. Geneva, Switzerland: WHO Press; 2011.
5. Centers for Disease Control and Prevention. Sexually transmitted infections in developing countries: Current concepts and strategies on imporving STI prevention, treatment, and control. Center for Disease Control. 2008. URL: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/867421468313772326 [accessed 2025-11-15]
6. Paavonen J, Eggert-Kruse W. Chlamydia trachomatis: impact on human reproduction. Hum Reprod Update 1999;5(5):433-447. [doi: 10.1093/humupd/5.5.433] [Medline: 10582782]
7. Ville Y, Leruez M, Glowaczower E, Robertson JN, Ward ME. The role of Chlamydia trachomatis and Neisseria gonorrhoeae in the aetiology of ectopic pregnancy in Gabon. Br J Obstet Gynaecol 1991;98(12):1260-1266. [doi: 10.1111/j.1471-0528.1991.tb15399.x] [Medline: 1777459]
8. Stephens AJ, Aubuchon M, Schust DJ. Antichlamydial antibodies, human fertility, and pregnancy wastage. Infect Dis Obstet Gynecol 2011;2011:525182 [FREE Full text] [doi: 10.1155/2011/525182] [Medline: 21949601]
9. Amornkul PN, Vandenhoudt H, Nasokho P, Odhiambo F, Mwaengo D, Hightower A, et al. HIV prevalence and associated risk factors among individuals aged 13-34 years in Rural Western Kenya. PLoS One 2009;4(7):e6470 [FREE Full text] [doi: 10.1371/journal.pone.0006470] [Medline: 19649242]
10. Masese L, Baeten JM, Richardson BA, Bukusi E, John-Stewart G, Graham SM, et al. Changes in the contribution of genital tract infections to HIV acquisition among Kenyan high-risk women from 1993 to 2012. AIDS 2015;29(9):1077-1085 [FREE Full text] [doi: 10.1097/QAD.0000000000000646] [Medline: 26125141]
11. De Baetselier I, Vuylsteke B, Yaya I, Dagnra A, Diandé S, Mensah E, CohMSM-PrEP pooling study group. To pool or not to pool samples for sexually transmitted infections detection in men who have sex with men? An evaluation of a new pooling method using the GeneXpert instrument in West Africa. Sex Transm Dis 2020;47(8):556-561. [doi: 10.1097/OLQ.0000000000001191] [Medline: 32355106]
12. Jones J, Sanchez TH, Dominguez K, Bekker L, Phaswana-Mafuya N, Baral SD, et al. Sexually transmitted infection screening, prevalence and incidence among South African men and transgender women who have sex with men enrolled in a combination HIV prevention cohort study: the sibanye methods for prevention packages programme (MP3) project. J Int AIDS Soc 2020;23(Suppl 6):e25594 [FREE Full text] [doi: 10.1002/jia2.25594] [Medline: 33000886]
13. Otieno F, Ng'ety G, Okall D, Aketch C, Obondi E, Graham SM, et al. Incident gonorrhoea and chlamydia among a prospective cohort of men who have sex with men in Kisumu, Kenya. Sex Transm Infect 2020;96(7):521-527. [doi: 10.1136/sextrans-2019-054166] [Medline: 31974213]
14. Sanders EJ, Thiong'o AN, Okuku HS, Mwambi J, Priddy F, Shafi J, et al. High prevalence of Chlamydia trachomatis and Neisseria gonorrhoeae infections among HIV-1 negative men who have sex with men in coastal Kenya. Sex Transm Infect 2010;86(6):440-441. [doi: 10.1136/sti.2010.043224] [Medline: 20656722]
15. Grijsen ML, Graham SM, Mwangome M, Githua P, Mutimba S, Wamuyu L, et al. Screening for genital and anorectal sexually transmitted infections in HIV prevention trials in Africa. Sex Transm Infect 2008;84(5):364-370 [FREE Full text] [doi: 10.1136/sti.2007.028852] [Medline: 18375645]
16. Ouedraogo HG, Zida S, Compaore TR, Lanou BH, Rao A, Sagna T, et al. Seroepidemiology of syphilis among men who have sex with men in Burkina Faso, West Africa. Eur J Clin Microbiol Infect Dis 2019;38(10):1803-1809. [doi: 10.1007/s10096-019-03610-2] [Medline: 31201641]
17. Twahirwa Rwema JO, Herbst S, Hamill MM, Liestman B, Nyombayire J, Lyons CE, et al. Cross-sectional assessment of determinants of STIs among men who have sex with men and transgender women in Kigali, Rwanda. Sex Transm Infect 2022;98(3):178-187 [FREE Full text] [doi: 10.1136/sextrans-2020-054753] [Medline: 33958492]
18. Leu RH. Complications of coexisting chlamydial and gonococcal infections. Postgrad Med 1991;89(7):56-60. [doi: 10.1080/00325481.1991.11700939] [Medline: 2031031]
19. Jones J, Weiss K, Mermin J, Dietz P, Rosenberg ES, Gift TL, et al. Proportion of incident human immunodeficiency virus cases among men who have sex with men attributable to gonorrhea and chlamydia: A modeling analysis. Sex Transm Dis 2019;46(6):357-363 [FREE Full text] [doi: 10.1097/OLQ.0000000000000980] [Medline: 31095100]
20. Pettifor A, Walsh J, Wilkins V, Raghunathan P. How effective is syndromic management of STDs?: A review of current studies. Sex Transm Dis 2000;27(7):371-385. [doi: 10.1097/00007435-200008000-00002] [Medline: 10949428]

XSL•FO
RenderX

21. Prevention and control of sexually transmitted infections (STIs) in the era of oral pre-exposure prophylaxis (PrEP) for HIV. In: World Health Organization. Geneva, Switzerland: WHO Press; 2019.
22. Agutu CA, Ngetsa CJ, Price MA, Rinke de Wit TF, Omosa-Manyonyi G, Sanders EJ, et al. Systematic review of the performance and clinical utility of point of care HIV-1 RNA testing for diagnosis and care. PLoS One 2019;14(6):e0218369 [FREE Full text] [doi: 10.1371/journal.pone.0218369] [Medline: 31246963]
23. Voirin N, Allam C, Charre C, Fernandez C, Godinot M, Oria F, et al. Optimizing strategies for Chlamydia trachomatis and Neisseria gonorrhoeae screening in men who have sex with men: A modeling study. Clin Infect Dis 2020;70(9):1966-1972. [doi: 10.1093/cid/ciz510] [Medline: 31198933]
24. Current health expenditure per capita (current US$) - Kenya. World Bank. 2008. URL: https://data.worldbank.org/indicator/SH.XPD.CHEX.PC.CD?locations=KE [accessed 2025-11-15]
25. WHO implementation tool for pre-exposure prophylaxis (PrEP) of HIV infection. Module 13. Integrating STI services. In: World Health Organization. Geneva, Switzerland: WHO Press; 2022.
26. Wi TE, Ndowa FJ, Ferreyra C, Kelly-Cirino C, Taylor MM, Toskin I, et al. Diagnosing sexually transmitted infections in resource-constrained settings: challenges and ways forward. J Int AIDS Soc 2019;22(Suppl 6):e25343 [FREE Full text] [doi: 10.1002/jia2.25343] [Medline: 31468679]
27. Molina JM, Charreau I, Chidiac C, Pialoux G, Cua E, Delaugerre C, ANRS IPERGAY Study Group. Post-exposure prophylaxis with doxycycline to prevent sexually transmitted infections in men who have sex with men: an open-label randomised substudy of the ANRS IPERGAY trial. Lancet Infect Dis 2018;18(3):308-317. [doi: 10.1016/S1473-3099(17)30725-9] [Medline: 29229440]
28. Luetkemeyer AF, Donnell D, Dombrowski JC, Cohen S, Grabow C, Brown CE, DoxyPEP Study Team. Postexposure doxycycline to prevent bacterial sexually transmitted infections. N Engl J Med 2023;388(14):1296-1306 [FREE Full text] [doi: 10.1056/NEJMoa2211934] [Medline: 37018493]
29. Molina JM, Bercot B, Assoumou L, Rubenstein E, Algarte-Genin M, Pialoux G, ANRS 174 DOXYVAC Study Group. Doxycycline prophylaxis and meningococcal group B vaccine to prevent bacterial sexually transmitted infections in France (ANRS 174 DOXYVAC): a multicentre, open-label, randomised trial with a 2 × 2 factorial design. Lancet Infect Dis 2024;24(10):1093-1104. [doi: 10.1016/S1473-3099(24)00236-6] [Medline: 38797183]
30. Sanders EJ, Okuku HS, Smith AD, Mwangome M, Wahome E, Fegan G, et al. High HIV-1 incidence, correlates of HIV-1 acquisition, and high viral loads following seroconversion among MSM. AIDS 2013;27(3):437-446 [FREE Full text] [doi: 10.1097/QAD.0b013e32835b0f81] [Medline: 23079811]
31. Sanders EJ, Wahome E, Okuku HS, Thiong'o AN, Smith AD, Duncan S, et al. Evaluation of WHO screening algorithm for the presumptive treatment of asymptomatic rectal gonorrhoea and chlamydia infections in at-risk MSM in Kenya. Sex Transm Infect 2014;90(2):94-99 [FREE Full text] [doi: 10.1136/sextrans-2013-051078] [Medline: 24327758]
32. Quilter LAS, Obondi E, Kunzweiler C, Okall D, Bailey RC, Djomand G, et al. Prevalence and correlates of and a risk score to identify asymptomatic anorectal gonorrhoea and chlamydia infection among men who have sex with men in Kisumu, Kenya. Sex Transm Infect 2019;95(3):201-211 [FREE Full text] [doi: 10.1136/sextrans-2018-053613] [Medline: 30242143]
33. Steen R, Dallabetta G. The use of epidemiologic mass treatment and syndrome management for sexually transmitted disease control. Sex Transm Dis 1999;26(4 Suppl):S12-20; discussion S21. [doi: 10.1097/00007435-199904001-00004] [Medline: 10227695]
34. Steen R, Dallabetta G. Sexually transmitted infection control with sex workers: regular screening and presumptive treatment augment efforts to reduce risk and vulnerability. Reprod Health Matters 2003;11(22):74-90 [FREE Full text] [doi: 10.1016/s0968-8080(03)02295-x] [Medline: 14708399]
35. Fonck K, Kaul R, Kimani J, Keli F, MacDonald KS, Ronald AR, et al. A randomized, placebo-controlled trial of monthly azithromycin prophylaxis to prevent sexually transmitted infections and HIV-1 in Kenyan sex workers: study design and baseline findings. Int J STD AIDS 2000;11(12):804-811. [doi: 10.1258/0956462001915327] [Medline: 11138916]
36. Kaul R, Kimani J, Nagelkerke NJ, Fonck K, Ngugi EN, Keli F, Kibera HIV Study Group. Monthly antibiotic chemoprophylaxis and incidence of sexually transmitted infections and HIV-1 infection in Kenyan sex workers: a randomized controlled trial. JAMA 2004;291(21):2555-2562. [doi: 10.1001/jama.291.21.2555] [Medline: 15173146]
37. Wi T, Ramos ER, Steen R, Esguerra TA, Roces MCR, Lim-Quizon MC, et al. STI declines among sex workers and clients following outreach, one time presumptive treatment, and regular screening of sex workers in the Philippines. Sex Transm Infect 2006;82(5):386-391 [FREE Full text] [doi: 10.1136/sti.2005.018283] [Medline: 17012514]
38. Steen R, Dallabetta G, Neilsen G. Antibiotic chemoprophylaxis and HIV infection in Kenyan sex workers. JAMA 2004;292(8):921-922. [doi: 10.1001/jama.292.8.921-a] [Medline: 15328320]
39. Steen R, Chersich M, Gerbase A, Neilsen G, Wendland A, Ndowa F, et al. Periodic presumptive treatment of curable sexually transmitted infections among sex workers: a systematic review. AIDS 2012;26(4):437-445. [doi: 10.1097/QAD.0b013e32834ed991] [Medline: 22095197]
40. Guidelines for the management of symptomatic sexually transmitted infections. In: World Health Organization. Geneva, Switzerland: WHO Press; 2021.

XSL•FO
RenderX

41. Klausner JD, Aplasca MR, Mesola VP, Bolan G, Whittington WL, Holmes KK. Correlates of gonococcal infection and of antimicrobial-resistant Neisseria gonorrhoeae among female sex workers, Republic of the Philippines, 1996-1997. J Infect Dis 1999;179(3):729-733. [doi: 10.1086/314625] [Medline: 9952388]
42. Cehovin A, Harrison OB, Lewis SB, Ward PN, Ngetsa C, Graham SM, et al. Identification of novel Neisseria gonorrhoeae lineages harboring resistance plasmids in coastal Kenya. J Infect Dis 2018;218(5):801-808 [FREE Full text] [doi: 10.1093/infdis/jiy240] [Medline: 29701830]
43. Kularatne R, Maseko V, Gumede L, Kufa T. Trends in antimicrobial resistance over a ten-year surveillance period, johannesburg, South Africa, 2008 2017. Antibiotics (Basel) 2018;7(3):58 [FREE Full text] [doi: 10.3390/antibiotics7030058] [Medline: 30002329]
44. Wi T, Lahra MM, Ndowa F, Bala M, Dillon JR, Ramon-Pardo P, et al. Antimicrobial resistance in Neisseria gonorrhoeae: Global surveillance and a call for international collaborative action. PLoS Med 2017;14(7):e1002344 [FREE Full text] [doi: 10.1371/journal.pmed.1002344] [Medline: 28686231]
45. Soge OO, Salipante SJ, No D, Duffy E, Roberts MC. In vitro activity of delafloxacin against clinical Neisseria gonorrhoeae Isolates and selection of gonococcal delafloxacin resistance. Antimicrob Agents Chemother 2016;60(5):3106-3111 [FREE Full text] [doi: 10.1128/AAC.02798-15] [Medline: 26976873]
46. Lagace-Wiens PRS, Duncan S, Kimani J, Thiong'o A, Shafi J, McClelland S, et al. Emergence of fluoroquinolone resistance in Neisseria gonorrhoeae isolates from four clinics in three regions of Kenya. Sex Transm Dis 2012;39(5):332-334 [FREE Full text] [doi: 10.1097/OLQ.0b013e318248a85f] [Medline: 22504591]
47. Lahra MM, Ryder N, Whiley DM. A new multidrug-resistant strain of Neisseria gonorrhoeae in Australia. N Engl J Med 2014;371(19):1850-1851. [doi: 10.1056/NEJMc1408109] [Medline: 25372111]
48. Mehta SD, Maclean I, Ndinya-Achola JO, Moses S, Martin I, Ronald A, et al. Emergence of quinolone resistance and cephalosporin MIC creep in Neisseria gonorrhoeae isolates from a cohort of young men in Kisumu, Kenya, 2002 to 2009. Antimicrob Agents Chemother 2011;55(8):3882-3888 [FREE Full text] [doi: 10.1128/AAC.00155-11] [Medline: 21606224]
49. Fayemiwo SA, Müller EE, Gumede L, Lewis DA. Plasmid-mediated penicillin and tetracycline resistance among Neisseria gonorrhoeae isolates in South Africa: prevalence, detection and typing using a novel molecular assay. Sex Transm Dis 2011;38(4):329-333. [doi: 10.1097/OLQ.0b013e3181fc695a] [Medline: 21042234]
50. Soge OO, Issema R, Bukusi E, Baeten JM, Mujugira A, Celum C, Partners PrEP Study Team. Predominance of high-level tetracycline-resistant Neisseria gonorrhoeae in Kenya: Implications for global implementation of doxycycline postexposure prophylaxis for prevention of sexually transmitted infections. Sex Transm Dis 2023;50(5):317-319 [FREE Full text] [doi: 10.1097/OLQ.0000000000001769] [Medline: 36728331]
51. Kivata MW, Mbuchi M, Eyase F, Bulimo WD, Kyanya CK, Oundo V, et al. Plasmid mediated penicillin and tetracycline resistance among Neisseria gonorrhoeae isolates from Kenya. BMC Infect Dis 2020;20(1):703 [FREE Full text] [doi: 10.1186/s12879-020-05398-5] [Medline: 32977759]
52. Duncan S, Thiong'o AN, Macharia M, Wamuyu L, Mwarumba S, Mvera B, et al. High prevalence of quinolone resistance in Neisseria gonorrhoeae in coastal Kenya. Sex Transm Infect 2011;87(3):231. [doi: 10.1136/sti.2010.048777] [Medline: 21307154]
53. Molina JM, Charreau I, Chidiac C, Pialoux G, Cua E, Delaugerre C, ANRS IPERGAY Study Group. Post-exposure prophylaxis with doxycycline to prevent sexually transmitted infections in men who have sex with men: an open-label randomised substudy of the ANRS IPERGAY trial. Lancet Infect Dis 2018;18(3):308-317. [doi: 10.1016/S1473-3099(17)30725-9] [Medline: 29229440]
54. La Ruche G, Goubard A, Bercot B, Cambau E, Semaille C, Sednaoui P. Gonococcal infections and emergence of gonococcal decreased susceptibility to cephalosporins in France, 2001 to 2012. Euro Surveill 2014;19(34):20885 [FREE Full text] [doi: 10.2807/1560-7917.es2014.19.34.20885] [Medline: 25188611]
55. Soge OO, Thibault CS, Cannon CA, McLaughlin SE, Menza TW, Dombrowski JC, et al. Potential impact of doxycycline post-exposure prophylaxis on tetracycline resistance in Neisseria gonorrhoeae and colonization with tetracycline-resistant staphylococcus aureus and group a streptococcus. Clin Infect Dis 2025;80(6):1188-1196. [doi: 10.1093/cid/ciaf089] [Medline: 40036749]
56. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
57. WHO guidelines for the treatment of Neisseria gonorrhoeae. In: World Health Organization. Geneva, Switzerland: WHO Press; 2016.
58. Luetkemeyer AF, Donnell D, Cohen SE, Dombrowski JC, Grabow C, Haser G, et al. Doxycycline to prevent bacterial sexually transmitted infections in the USA: final results from the DoxyPEP multicentre, open-label, randomised controlled trial and open-label extension. Lancet Infect Dis 2025;25(8):873-883. [doi: 10.1016/S1473-3099(25)00085-4] [Medline: 40147465]
59. Kenya national guidelines for prevention, management, and control of sexually transmitted infections. In: Kenya Ministry of Health. Nairobi, Kenya: Government of Kenya; 2018.

XSL•FO
RenderX

60. Unemo M, Golparian D, Sánchez-Busó L, Grad Y, Jacobsson S, Ohnishi M, et al. The novel 2016 WHO Neisseria gonorrhoeae reference strains for global quality assurance of laboratory investigations: phenotypic, genetic and reference genome characterization. J Antimicrob Chemother 2016;71(11):3096-3108 [FREE Full text] [doi: 10.1093/jac/dkw288] [Medline: 27432602]
61. Durukan D, Read TRH, Bradshaw CS, Fairley CK, Williamson DA, De Petra V, et al. Pooling pharyngeal, anorectal, and urogenital samples for screening asymptomatic men who have sex with men for Chlamydia trachomatis and Neisseria gonorrhoeae. J Clin Microbiol 2020;58(5):e01969 [FREE Full text] [doi: 10.1128/JCM.01969-19] [Medline: 32132192]
62. Wilson JD, Wallace HE, Loftus-Keeling M, Ward H, Davies B, Vargas-Palacios A, et al. Swab-yourself trial with economic monitoring and testing for infections collectively (SYSTEMATIC): Part 2. A diagnostic accuracy and Cost-effectiveness study comparing rectal, pharyngeal, and urogenital samples analyzed individually, versus as a pooled specimen, for the diagnosis of gonorrhea and chlamydia. Clin Infect Dis 2021;73(9):e3183-e3193. [doi: 10.1093/cid/ciaa1546] [Medline: 33044490]
63. Angelakis E, Armstrong N, Nappez C, Richez M, Chabriere E, Raoult D. Doxycycline assay hair samples for testing long-term compliance treatment. J Infect 2015;71(5):511-517. [doi: 10.1016/j.jinf.2015.08.003] [Medline: 26299894]
64. Workowski KA, Bachmann LH, Chan PA, Johnston CM, Muzny CA, Park I, et al. Sexually transmitted infections treatment guidelines, 2021. MMWR Recomm Rep 2021;70(4):1-187 [FREE Full text] [doi: 10.15585/mmwr.rr7004a1] [Medline: 34292926]
65. Bignell C, Ison CA, Jungmann E. Gonorrhoea. Sex Transm Infect 2006;82(Suppl 4):iv6-iv9 [FREE Full text] [doi: 10.1136/sti.2006.023036] [Medline: 17151055]
66. Pabbaraju K, Wong S, Song JJ, Singh AE, Read R, Drews SJ. Utility of specimens positive for Neisseria gonorrhoeae by the Aptima Combo 2 assay for assessment of strain diversity and antibiotic resistance. J Clin Microbiol 2013;51(12):4156-4160 [FREE Full text] [doi: 10.1128/JCM.01694-13] [Medline: 24108603]
67. Pond MJ, Hall CL, Miari VF, Cole M, Laing KG, Jagatia H, et al. Accurate detection of Neisseria gonorrhoeae ciprofloxacin susceptibility directly from genital and extragenital clinical samples: towards genotype-guided antimicrobial therapy. J Antimicrob Chemother 2016;71(4):897-902 [FREE Full text] [doi: 10.1093/jac/dkv432] [Medline: 26817487]
68. Mayes J, Liyanarachchy N, Balm MND. Nucleic acid extraction from residual aptima clinical samples: Evaluation of automated platform to enhance workflow. J Clin Microbiol 2017;55(6):1968-1969 [FREE Full text] [doi: 10.1128/JCM.00342-17] [Medline: 28404675]
69. Martin IMC, Ison CA, Aanensen DM, Fenton KA, Spratt BG. Rapid sequence-based identification of gonococcal transmission clusters in a large metropolitan area. J Infect Dis 2004;189(8):1497-1505. [doi: 10.1086/383047] [Medline: 15073688]
70. Mortimer TD, Zhang JJ, Ma KC, Grad YH. Loci for prediction of penicillin and tetracycline susceptibility in Neisseria gonorrhoeae: a genome-wide association study. Lancet Microbe 2022;3(5):e376-e381 [FREE Full text] [doi: 10.1016/S2666-5247(22)00034-9] [Medline: 35544098]
71. Mortimer TD, Grad YH. Applications of genomics to slow the spread of multidrug-resistant Neisseria gonorrhoeae. Ann N Y Acad Sci 2019;1435(1):93-109 [FREE Full text] [doi: 10.1111/nyas.13871] [Medline: 29876934]
72. Peterson SW, Martin I, Demczuk W, Barairo N, Naidu P, Lefebvre B, et al. Multiplex real-time PCR assays for the prediction of cephalosporin, ciprofloxacin and azithromycin antimicrobial susceptibility of positive Neisseria gonorrhoeae nucleic acid amplification test samples. J Antimicrob Chemother 2020;75(12):3485-3490. [doi: 10.1093/jac/dkaa360] [Medline: 32830242]
73. Peterson SW, Martin I, Demczuk W, Bharat A, Hoang L, Wylie J, et al. Molecular assay for detection of genetic markers associated with decreased susceptibility to cephalosporins in neisseria gonorrhoeae. J Clin Microbiol 2015;53(7):2042-2048 [FREE Full text] [doi: 10.1128/JCM.00493-15] [Medline: 25878350]
74. Wong LK, Hemarajata P, Soge OO, Humphries RM, Klausner JD. Real-time PCR targeting the mosaic XXXIV Type for prediction of extended-spectrum-cephalosporin susceptibility in clinical neisseria gonorrhoeae isolates. Antimicrob Agents Chemother 2017;61(11):e01339 [FREE Full text] [doi: 10.1128/AAC.01339-17] [Medline: 28848021]
75. Trembizki E, Buckley C, Donovan B, Chen M, Guy R, Kaldor J, et al. Direct real-time PCR-based detection of Neisseria gonorrhoeae 23S rRNA mutations associated with azithromycin resistance. J Antimicrob Chemother 2015;70(12):3244-3249. [doi: 10.1093/jac/dkv274] [Medline: 26338048]
76. Chopra I, Roberts M. Tetracycline antibiotics: mode of action, applications, molecular biology, and epidemiology of bacterial resistance. Microbiol Mol Biol Rev 2001;65(2):232-260 [FREE Full text] [doi: 10.1128/MMBR.65.2.232-260.2001] [Medline: 11381101]
77. Teker B, de Vries H, Heijman T, van Dam A, Schim van der Loeff M, Jongen VW. Spontaneous clearance of asymptomatic anogenital and pharyngeal : a secondary analysis from the NABOGO trial. Sex Transm Infect 2023;99(4):219-225 [FREE Full text] [doi: 10.1136/sextrans-2022-055488] [Medline: 35820778]
78. Yelland LN, Salter AB, Ryan P. Performance of the modified Poisson regression approach for estimating relative risks from clustered prospective data. Am J Epidemiol 2011;174(8):984-992. [doi: 10.1093/aje/kwr183] [Medline: 21841157]
79. Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C, Mittman B. Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. Adm Policy Ment Health 2009;36(1):24-34 [FREE Full text] [doi: 10.1007/s10488-008-0197-4] [Medline: 19104929]

XSL•FO
RenderX

80. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
81. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
82. van der Elst EM, Okuku HS, Nakamya P, Muhaari A, Davies A, McClelland RS, et al. Is audio computer-assisted self-interview (ACASI) useful in risk behaviour assessment of female and male sex workers, Mombasa, Kenya? PLoS One 2009;4(5):e5340 [FREE Full text] [doi: 10.1371/journal.pone.0005340] [Medline: 19412535]
83. Wahome E, Thiong'o AN, Mwashigadi G, Chirro O, Mohamed K, Gichuru E, et al. An empiric risk score to guide PrEP targeting among MSM in coastal Kenya. AIDS Behav 2018;22(Suppl 1):35-44 [FREE Full text] [doi: 10.1007/s10461-018-2141-2] [Medline: 29767324]
84. Kunzweiler CP, Bailey RC, Okall DO, Graham SM, Mehta SD, Otieno FO. Factors associated with prevalent HIV infection among Kenyan MSM: The anza mapema study. J Acquir Immune Defic Syndr 2017;76(3):241-249. [doi: 10.1097/QAI.0000000000001512] [Medline: 28746167]
85. Mehta SD, Okall D, Graham SM, N'gety G, Bailey RC, Otieno F. Behavior change and sexually transmitted incidence in relation to PREP use among men who have sex with men in Kenya. AIDS Behav 2021;25(7):2219-2229. [doi: 10.1007/s10461-020-03150-3] [Medline: 33403516]
86. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993;88(6):791-804. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
87. Peltzer K, Davids A, Njuho P. Alcohol use and problem drinking in South Africa: findings from a national population-based survey. Afr J Psychiatry (Johannesbg) 2011;14(1):30-37. [doi: 10.4314/ajpsy.v14i1.65466] [Medline: 21509408]
88. Skinner HA. The drug abuse screening test. Addict Behav 1982;7(4):363-371. [doi: 10.1016/0306-4603(82)90005-3] [Medline: 7183189]
89. Smith N, Lam D, Bifulco A, Checkley S. Childhood experience of care and abuse questionnaire (CECA.Q). Validation of a screening instrument for childhood adversity in clinical populations. Soc Psychiatry Psychiatr Epidemiol 2002;37(12):572-579. [doi: 10.1007/s00127-002-0589-9] [Medline: 12545234]
90. Secor AM, Wahome E, Micheni M, Rao D, Simoni JM, Sanders EJ, et al. Depression, substance abuse and stigma among men who have sex with men in coastal Kenya. AIDS 2015;29(3):S251-S259 [FREE Full text] [doi: 10.1097/QAD.0000000000000846] [Medline: 26562814]
91. Korhonen CJ, Flaherty BP, Wahome E, Macharia P, Musyoki H, Battacharjee P, et al. Validity and reliability of the Neilands Sexual Stigma Scale among Kenyan gay, bisexual, and other men who have sex with men. BMC Public Health 2022;22(1):754 [FREE Full text] [doi: 10.1186/s12889-022-13066-3] [Medline: 35421967]
92. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
93. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
94. Jenness SM, Weiss KM, Goodreau SM, Gift T, Chesson H, Hoover KW, et al. Incidence of gonorrhea and chlamydia following human immunodeficiency virus preexposure prophylaxis among men who have sex with men: A modeling study. Clin Infect Dis 2017;65(5):712-718 [FREE Full text] [doi: 10.1093/cid/cix439] [Medline: 28505240]
95. Krivitsky P, Handcock M, Hunter D. Statnet. 2015. URL: https://cran.r-project.org/web/packages/tergm/index.html [accessed 2025-11-15]
96. Saldana L, Chamberlain P, Bradford WD, Campbell M, Landsverk J. The cost of implementing new strategies (COINS): A method for mapping implementation resources using the stages of implementation completion. Child Youth Serv Rev 2014;39:177-182 [FREE Full text] [doi: 10.1016/j.childyouth.2013.10.006] [Medline: 24729650]
97. Saldana L, Ritzwoller DP, Campbell M, Block EP. Using economic evaluations in implementation science to increase transparency in costs and outcomes for organizational decision-makers. Implement Sci Commun 2022;3(1):40 [FREE Full text] [doi: 10.1186/s43058-022-00295-1] [Medline: 35410434]
98. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the global burden of disease study 2010. Lancet 2012;380(9859):2129-2143 [FREE Full text] [doi: 10.1016/S0140-6736(12)61680-8] [Medline: 23245605]
99. Woods B, Revill P, Sculpher M, Claxton K. Country-level cost-effectiveness thresholds: Initial estimates and the need for further research. Value Health 2016;19(8):929-935 [FREE Full text] [doi: 10.1016/j.jval.2016.02.017] [Medline: 27987642]
100. Mburu G, Ram M, Siu G, Bitira D, Skovdal M, Holland P. Intersectionality of HIV stigma and masculinity in eastern Uganda: implications for involving men in HIV programmes. BMC Public Health 2014;14:1061 [FREE Full text] [doi: 10.1186/1471-2458-14-1061] [Medline: 25304035]

XSL•FO
RenderX

101. Van Dijck C, Laumen J, Zlotorzynska M, Manoharan-Basil SS, Kenyon C. Association between STI screening intensity in men who have sex with men and gonococcal susceptibility in 21 States in the USA: an ecological study. Sex Transm Infect 2020;96(7):537-540. [doi: 10.1136/sextrans-2019-054313] [Medline: 32066589]
102. Kenyon C, Laumen J, Van Dijck C. Could intensive screening for gonorrhea/chlamydia in preexposure prophylaxis cohorts select for resistance? Historical lessons from a mass Treatment campaign in greenland. Sex Transm Dis 2020;47(1):24-27. [doi: 10.1097/OLQ.0000000000001092] [Medline: 31856072]
103. Chisholm SA, Dave J, Ison CA. High-level azithromycin resistance occurs in Neisseria gonorrhoeae as a result of a single point mutation in the 23S rRNA genes. Antimicrob Agents Chemother 2010;54(9):3812-3816 [FREE Full text] [doi: 10.1128/AAC.00309-10] [Medline: 20585125]
104. Wind CM, de Vries E, Schim van der Loeff MF, van Rooijen MS, van Dam AP, Demczuk WHB, et al. Decreased azithromycin susceptibility of neisseria gonorrhoeae isolates in patients recently treated with azithromycin. Clin Infect Dis 2017;65(1):37-45. [doi: 10.1093/cid/cix249] [Medline: 28510723]
105. Gonococcal antimicrobial surveillance programme (GASP). World Health Organization. URL: https://www.who.int/initiatives/gonococcal-antimicrobial-surveillance-programme [accessed 2025-11-15]
106. Weston EJ, Kirkcaldy RD, Stenger M, Llata E, Hoots B, Torrone EA. Narrative review: Assessment of neisseria gonorrhoeae infections among men who have sex with men in national and sentinel surveillance systems in the United States. Sex Transm Dis 2018;45(4):243-249 [FREE Full text] [doi: 10.1097/OLQ.0000000000000740] [Medline: 29465702]
107. Rowlinson E, Golden MR, Berzkalns A, Thibault C, Barbee LA. Epidemiologic treatment for contacts to neisseria gonorrhoeae and chlamydia trachomatis infection in sexually transmitted disease clinic patients in Seattle, WA; 1994 to 2018. Sex Transm Dis 2020;47(10):665-671 [FREE Full text] [doi: 10.1097/OLQ.0000000000001234] [Medline: 32649579]
108. Allen KS, Hinrichs R, Heumann CL, Titus MK, Duszynski TJ, Valvi NR, et al. Findings from a scoping review: Presumptive treatment for chlamydiatrachomatis and neisseria gonorrhoeae in the United States, 2006-2021. Sex Transm Dis 2023;50(4):209-214 [FREE Full text] [doi: 10.1097/OLQ.0000000000001762] [Medline: 36584164]
109. Llata E, Braxton J, Asbel L, Huspeni D, Tourdot L, Kerani RP, et al. Presumptive and follow-up treatment associated with gonorrhea and chlamydia testing episodes in sexually transmitted disease clinics: Impact of changing treatment guidelines for gonorrhea, sexually transmitted disease surveillance network, 2015-2018. Sex Transm Dis 2023;50(1):5-10 [FREE Full text] [doi: 10.1097/OLQ.0000000000001714] [Medline: 36194764]
110. Wong A, Applegate T, Boettiger DC, Varma R, Guy R, Medland N. Unnecessary antibiotic use in men who have sex with men (MSM) with anogenital symptoms attending a sexual health clinic: a retrospective analysis. Sex Transm Infect 2024;100(7):435-441. [doi: 10.1136/sextrans-2024-056120] [Medline: 38871452]
111. Said LB, Conrad A, Souza SD, Alfaiate D, Ader F, Batalla A, et al. Partner treatment strategies for chlamydia and gonorrhea: Time for a reappraisal. Infect Dis Now 2025;55(3):105038 [FREE Full text] [doi: 10.1016/j.idnow.2025.105038] [Medline: 39947479]
112. Vanbaelen T, Tsoumanis A, Florence E, Van Dijck C, Huis In 't Veld D, Sauvage A, et al. Effect of screening for neisseria gonorrhoeae and Chlamydia trachomatis on incidence of these infections in men who have sex with men and transgender women taking HIV pre-exposure prophylaxis (the Gonoscreen study): results from a randomised, multicentre, controlled trial. Lancet HIV 2024;11(4):e233-e244 [FREE Full text] [doi: 10.1016/S2352-3018(23)00299-0] [Medline: 38423024]
113. Vanbaelen T, Manoharan-Basil SS, Kenyon C. Effect of mass treatment on the long-term prevalence of gonorrhoea, chlamydia and syphilis-a systematic review. Int J STD AIDS 2024;35(7):550-564. [doi: 10.1177/09564624241239994] [Medline: 38506648]
114. Kenyon C, Gestels Z, Vanbaelen T, Abdellati S, Van Den Bossche D, De Baetselier I, et al. Doxycycline PEP can induce doxycycline resistance in in a model of PEP. Front Microbiol 2023;14:1208014 [FREE Full text] [doi: 10.3389/fmicb.2023.1208014] [Medline: 37711686]
115. Chu VT, Glascock A, Donnell D, Grabow C, Brown CE, Ward R, et al. Impact of doxycycline post-exposure prophylaxis for sexually transmitted infections on the gut microbiome and antimicrobial resistome. Nat Med 2025;31(1):207-217 [FREE Full text] [doi: 10.1038/s41591-024-03274-2] [Medline: 39363100]
116. Chu VT, Glascock A, Donnell D, Grabow C, Brown CE, Ward R, et al. Impact of doxycycline post-exposure prophylaxis for sexually transmitted infections on the gut microbiome and antimicrobial resistome. Nat Med 2025;31(1):207-217 [FREE Full text] [doi: 10.1038/s41591-024-03274-2] [Medline: 39363100]
117. Cornelisse VJ, Ong JJ, Ryder N, Ooi C, Wong A, Kenchington P, et al. Interim position statement on doxycycline post-exposure prophylaxis (Doxy-PEP) for the prevention of bacterial sexually transmissible infections in Australia and Aotearoa New Zealand - the Australasian Society for HIV, Viral Hepatitis and Sexual Health Medicine (ASHM). Sex Health 2023;20(2):99-104. [doi: 10.1071/SH23011] [Medline: 36927481]
118. Werner RN, Schmidt AJ, Potthoff A, Spornraft-Ragaller P, Brockmeyer NH, German STI Society (DSTIG). Position statement of the German STI Society on the prophylactic use of doxycycline to prevent STIs (Doxy-PEP, Doxy-PrEP). J Dtsch Dermatol Ges 2024;22(3):466-478. [doi: 10.1111/ddg.15282] [Medline: 38123738]
119. Sugarman J, Taylor HA, Bachmann LH, Barbee LA, Cahill S, Celum C, et al. Ethical considerations in implementing doxycycline postexposure prophylaxis for the prevention of bacterial sexually transmitted infections. Sex Transm Dis 2024;51(3):135-138. [doi: 10.1097/OLQ.0000000000001907] [Medline: 38079243]

XSL•FO
RenderX

120. Helekal D, Mortimer TD, Grad YH. Expansion of -Carrying in the United States, 2018-2024. N Engl J Med 2025;393(2):198-200. [doi: 10.1056/NEJMc2504010] [Medline: 40632956]
121. Traeger MW, Mayer KH, Krakower DS, Gitin S, Jenness SM, Marcus JL. Potential impact of doxycycline post-exposure prophylaxis prescribing strategies on incidence of bacterial sexually transmitted infections. Clin Infect Dis 2023:ciad488. [doi: 10.1093/cid/ciad488] [Medline: 37595139]
122. Reichert E, Grad YH. Effects of doxycycline post-exposure prophylaxis for prevention of sexually transmitted infections on gonorrhoea prevalence and antimicrobial resistance among men who have sex with men in the USA: a modelling study. Lancet Microbe 2024;5(11):100926 [FREE Full text] [doi: 10.1016/S2666-5247(24)00168-X] [Medline: 39374606]
123. Tantalo LC, Luetkemeyer AF, Lieberman NAP, Nunley BE, Avendaño C, Greninger AL, et al. In vitro exposure of treponema pallidum to subbactericidal doxycycline did not induce Resistance: Implications for doxycycline postexposure prophylaxis. J Infect Dis 2025;231(3):729-733. [doi: 10.1093/infdis/jiae381] [Medline: 39067061]
124. Sanders EJ, Wahome E, Otieno F, Kimani J, Bentzer A, Okall D, et al. Chlamydia and gonorrhoea infections in young Kenyan HIV-negative cisgender men who have sex with men and transgender women: a multicentre cohort study. BMJ Open 2025;15(7):e098916 [FREE Full text] [doi: 10.1136/bmjopen-2025-098916] [Medline: 40750281]

Abbreviations

ACASI: audio computer-assisted self-interview
AMR: antimicrobial resistance
ART: antiretroviral therapy
CONSORT: Consolidated Standards of Reporting Trials
CT: Chlamydia trachomatis
doxyPEP: doxycycline post-exposure prophylaxis
GEE: generalized estimating equations
MSM: men who have sex with men
NAAT: nucleic acid amplification testing
NG: Neisseria gonorrhoeae
PPT: periodic presumptive treatment
PrEP: pre-exposure prophylaxis
REDCap: Research Electronic Data Capture
STI: sexually transmitted infection
WHO: World Health Organization

Edited by T Leung, A Schwartz; submitted 24.Jul.2025; this is a non–peer-reviewed article;accepted 30.Nov.2025; published 06.Jan.2026.

Please cite as:
Graham SM, Otieno FO, Kimani J, Barrientos A, Soge OO, Sharma M, Hamilton DT, Celum C, McClelland RS, Sanders EJ
World Health Organization–Recommended Periodic Presumptive Treatment Versus Doxycycline Post-Exposure Prophylaxis for Sexually Transmitted Infection Control Among Men Who Have Sex With Men in Kenya: Protocol for a Randomized Controlled Trial
JMIR Res Protoc 2026;15:e81113
URL: https://www.researchprotocols.org/2026/1/e81113
doi:10.2196/81113
PMID:

©Susan M Graham, Fredrick O Otieno, Joshua Kimani, Alé Barrientos, Olusegun O Soge, Monisha Sharma, Deven T Hamilton, Connie Celum, R. Scott McClelland, Eduard J Sanders. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 06.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Virtual Mentored Implementation to Improve Care Transitions in Chronic Obstructive Pulmonary Disease: Protocol for a Pragmatic Implementation Study

Cathryn T Lee[1], MD, MS; Leah Traeger[1], MHA; Mahima Akula[1], BA; April E Fegley[2], BA; Jenna Goldstein[2], MA; Kim Erwin[3], MDes; Laura J Damschroder[4], MPH; Jean Rommes[5], PhD; Hannah Pick[6], MA; Andrew Auerbach[7], MD; Peter Lindenauer[8], MD; Wen Wan[1], PhD; Andrea Jackson Sagredo[1], MS; Valerie G Press[1], MD, MPH

[1]Section of Pulmonary and Critical Care Medicine, Department of Medicine, University of Chicago, Chicago, IL, United States

[2]Society of Hospital Medicine, Philadelphia, PA, United States

[3]Institute of Design, Illinois Institute of Technology, Chicago, IL, United States

[4]Implementation Pathways LLC, Chelsea, MI, United States

[5]Chronic Obstructive Pulmonary Disease Foundation, Des Moines, IA, United States

[6]Onda Collective LLC, Chicago, IL, United States

[7]Division of Hospital Medicine, University of California, San Francisco, San Francisco, CA, United States

[8]Division of Healthcare Delivery and Population Sciences, UMass Chan Medical School- Baystate, Springfield, MA, United States

Corresponding Author:
Valerie G Press, MD, MPH
Section of Pulmonary and Critical Care Medicine, Department of Medicine
University of Chicago
5841 S Maryland Ave
Chicago, IL, 60637
United States
Phone: 1 773 702 5170
Email: vpress@bsd.uchicago.edu

Abstract

Background: Chronic obstructive pulmonary disease (COPD) is a leading cause of mortality and morbidity among US adults, including recurrent emergency department (ED) visits and unplanned hospital admissions. Despite this, the transition of care (TOC) from the inpatient to outpatient setting remains under-studied.

Objective: The objectives of the Reduce Respiratory Emergent Visits using Implementation Science Interventions Tailored to Setting (Reduce REVISITS) study are to conduct contextual assessments to inform implementation plans for COPD TOC interventions, conduct a cluster randomized trial evaluating implementation over 1 year of COPD TOC bundles, and monitor sustainment of implementation over a 2-year period across 20 sites.

Methods: This pragmatic, multisite study uses a hybrid type II effectiveness-implementation design to evaluate clinical and implementation outcomes of COPD TOC programs across 20 sites. Sites are cluster-randomized to 1 of 4 intervention groups, varying by program delivery method (in-person vs virtual) and implementation strategy (mentored implementation with or without co-design). Sites select evidence-based interventions they wish to incorporate into their COPD TOC program and are randomized to in-person or virtual delivery. During the 1-year active implementation period of the study, assigned mentors will meet monthly with sites (for a total of 12 sessions) to enable on-the-ground troubleshooting of site-specific difficulties with TOC interventions. The primary effectiveness outcome for this study will be COPD-specific acute health care use, defined as a composite of all ED visits and hospitalizations within 30 days of index hospitalization for a COPD exacerbation. The primary implementation outcome will be reach, defined as the proportion of patients receiving their assigned TOC interventions (the whole bundle).

Results: As of August 2025, 21 sites completed the contextual assessments and developed site-specific implementation plans. Publication of the qualitative data from this pre-implementation phase is anticipated in December 2025. Site randomization is complete; sites randomized to co-design have completed 3 sessions. Baseline data collection on use is complete. Implementation-year data collection on use is nearly complete. Year 1 and 2 post-implementation-phase data collection on use is ongoing. Quantitative data analyses of the baseline and implementation-phase reports are nearly complete. Manuscript submission for the primary implementation-phase manuscript is anticipated for December 2025. Manuscript submission for the implementation-sustainment

XSL•FO
RenderX

analyses are anticipated for September 2026. Qualitative data collection for year 1 of the post-implementation phase is complete, and analysis is under way. Qualitative data collection for year 2 began in August 2025.

Conclusions: The Reduce REVISITS study will use novel integrated implementation science and human-centered design methodology to investigate bundles of effective COPD TOC interventions with the goal of reducing COPD hospital revisits. The study will evaluate evidence-based programs for effectiveness and implementation across a wide variety of health care sites to ultimately improve outcomes in this high-risk patient population.

Trial Registration: ClinicalTrials.gov NCT05568043; https://clinicaltrials.gov/study/NCT05568043

International Registered Report Identifier (IRRID): DERR1-10.2196/82043

(JMIR Res Protoc 2026;15:e82043) doi:10.2196/82043

KEYWORDS

COPD; chronic obstructive pulmonary disease; hospital readmissions; transitions of care; COPD exacerbation; hybrid effectiveness-implementation design

Introduction

Chronic obstructive pulmonary disease (COPD) is one of the leading causes of morbidity and mortality for US adults [1], but the effectiveness of implementing COPD chronic disease management programs has been under-studied. This dearth of data became salient in 2014 when the Centers for Medicare and Medicaid Services added COPD to their Hospital Readmission Reduction Program (HRRP), which financially penalizes hospitals receiving federal funding for excessive 30-day readmissions [2]. Since COPD was the third leading cause of readmission among Medicare beneficiaries [3], including COPD among penalized conditions was logical. However, an expert panel convened in 2016 by the American Thoracic Society (ATS) to identify the best evidence and practices to reduce readmissions for acute exacerbations of COPD found that at the time of HRRP implementation there was little published evidence on effective hospital-based programs to reduce readmissions [4].

In the decade since the HRRP implementation, a handful of interventions have shown promise at reducing readmissions and facilitating transitions of care (TOC) from the inpatient to outpatient setting. These include early pulmonary rehabilitation, inhaler education, medication reconciliation and education via pharmacists, smoking cessation counseling, and early outpatient follow-up. However, a single intervention is unlikely to reduce readmissions enough to avoid financial penalties [5]. Hence, bundles of interventions through transition of care programs hold more promise. One study evaluated a COPD Chronic Care Management Collaborative comprised of education through subject matter experts and implementation support through peer coaching [6]. This collaborative aimed to support sites' implementation of interventions through quality improvement initiatives to reduce emergency department (ED) and hospital revisits across 47 US hospitals. With only about half of sites reporting, the collaborative found that they were able to successfully support the majority of sites in reducing ED and hospital revisits. In addition to this collaborative, a handful of institutions have published their attempts to address preventable acute care revisits after COPD hospitalization with mixed results [7]. For instance, one single-site randomized controlled trial (RCT) evaluating a 3-month transitional care and long-term self-management program for patients hospitalized with COPD found that patients enrolled in the intervention unexpectedly experienced harm in the form of more COPD acute care visits without improvement in quality of life [8]. Single-center studies of COPD bundled programs responding to the voluntary Bundled Payments for Care Innovation value-based care initiative of the Centers for Medicare and Medicaid Services have demonstrated mixed results regarding the effectiveness of readmission rates as well as the impact on overall health care costs [9,10].

One of the key concerns from the ATS report was a dearth of rigorously obtained data using RCTs along with limited sample size and consequently statistical power in single-center studies. Thus, we sought to study multiple COPD TOC bundle of care programs simultaneously in the National Institutes of Health–funded virtual mentored implementation to reduce Respiratory Emergent Visits using Implementation Science Interventions Tailored to Setting (Reduce REVISITS) study (R01HL146644). The purpose of this study is to evaluate tailored COPD TOC program delivery (in-person versus virtual) with implementation support (mentoring with or without co-design support) to identify ideal, scalable solutions for reducing excessive preventable COPD acute care revisits across diverse health care institutions. We used implementation science and human-centered design methods to develop the site-specific implementation plans [11]. We also will use these innovative methods to conduct the trial, specifically a hybrid type II effectiveness-implementation cluster RCT as a novel approach for rigorously and efficiently improving standards of care for TOC in high-risk patients with COPD. This hybrid approach will allow us to simultaneously evaluate the method of delivering the evidence-based interventions (in person versus virtual) along with measuring the level of implementation support (mentoring with or without co-design). The objectives of the Reduce REVISITS study are to (1) conduct contextual assessments using mixed methods to inform site-specific implementation plans for 2-3 evidence-based COPD TOC interventions (aim 1, methods published separately) [11]; (2) conduct a cluster randomized trial evaluating implementation over one year of the COPD TOC bundles (aim 2); and (3) monitor sustainment of implementation over a 2-year period (aim 3).

XSL•FO
RenderX

Methods

Study Design

This pragmatic, multisite study uses a hybrid type II effectiveness-implementation design to evaluate both clinical and implementation outcomes of COPD care transition programs across 20 diverse US hospital sites. After recruitment and enrollment, sites underwent a contextual assessment to develop site-specific implementation plans (published elsewhere) in aim 1 [11]. Participating hospitals will then be cluster-randomized to implement their intervention bundles using their site-specific implementation plans. Sites will be randomized to one of 4 intervention implementation groups, varying by program delivery method (in-person vs virtual) and implementation strategy (mentored implementation with or without co-design). Each site implements a tailored bundle of 2-3 evidence-based care transition interventions guided by the Society of Hospital Medicine (SHM)'s mentored implementation model (MIM) over a one-year period. Half of the sites also receive co-design support to actively engage end-users in shaping implementation approaches. The intervention groups are assessed through both quantitative outcomes, including acute care use, such as 30-day COPD-related revisits (primary effectiveness outcome) and bundle reach (primary implementation outcome). Deidentified patient-level data will be provided via Research Electronic Data Capture (REDCap; Vanderbilt University) and analyzed as below [12]. After one year of implementation, the project will enter the sustainment phase (aim 3), where ongoing effectiveness data and implementation data will be collected for an additional 2 years. Further, sustainment, fidelity, and cost will be evaluated through surveys and interviews. These post-implementation outcomes are measured over an additional 2-year follow-up period to assess the long-term sustainment and impact of implementation strategies and to inform future scale-up and dissemination (Figure 1). This study has been registered on ClinicalTrials.gov (NCT05568043).

Figure 1. Overall study schematic. This protocol paper focuses on the aim 2 methods. COPD: chronic obstructive pulmonary disease; ED: emergency department.

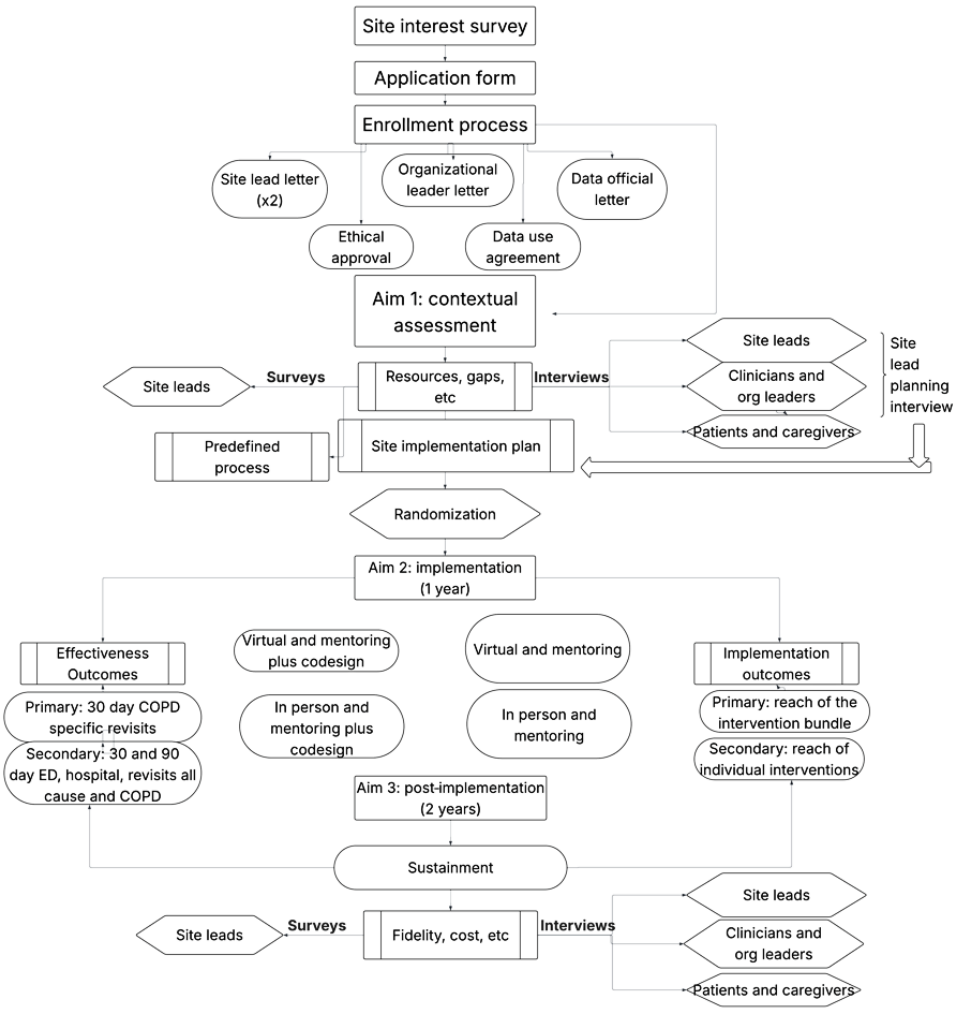

Population and Setting

Hospitals interested in improving their COPD TOC programs for patients hospitalized with COPD will be recruited into the initial phase of the study. The sites will be primarily identified from the Hospital Medicine Reengineering Network (HOMERuN), a collaborative of hospitals and care teams aimed at creating and improving best practices via a learning network [13]. The HOMERuN sites are geographically diverse and include both academic and community hospitals. Additional sites will be recruited through the SHM [14]. The study aims to enroll at least 20 hospitals in the first phase (aim 1). To allow

XSL•FO
RenderX

for sufficient power of the co-primary outcomes (see power calculation below) with up to 20% attrition in between this contextual assessment (aim 1) and randomization (aim 2), a minimum of 16 of the 20+ sites recruited into aim 1 sites are needed to participate in the trial (aims 2 and 3).

Site Recruitment and Onboarding

Sites are selected into the first phase of the study via a 3-step process designed to ensure diversity, feasibility, and readiness [15]. First, a site interest survey is disseminated broadly across the HOMERuN and SHM networks and through word-of-mouth referrals to gauge willingness and capacity to participate. Second, interested sites complete a detailed application capturing information about their hospital type (academic vs community), location (urban, suburban, and rural), baseline COPD care transition practices, population demographics, institutional priorities, and motivation for participating. The application also included questions to identify potential site leads and assess internal resources needed to support COPD program implementation.

Third, the formal enrollment process requires each site to (1) identify and commit at least 2 site leads to serve as project leads and points of contact; (2) submit letters of support from senior hospital leadership demonstrating institutional commitment; (3) secure institutional review board or ethical oversight approval; (4) confirm ability to collect, deidentify, and securely transmit data on a monthly basis via REDCap; and (5) engage in pre-implementation planning activities, including site lead interviews, providing names and contact information for stakeholder interviews, and discussions to develop a tailored implementation plan. All survey responses, applications, and enrollment forms are submitted via REDCap and tracked by the study team.

Once onboarding is complete, each enrolled site will participate in aim 1 activities, including contextual assessments and development of their site-specific implementation plans for 2-3 evidence-based COPD TOC interventions. Sites become eligible for randomization into aim 2 only after completing all aim 1 activities.

Mentor Recruitment, Onboarding, and Training

Mentors will be recruited using a snowball approach from membership in the SHM and ATS [16]. Mentors will be expected to have quality improvement expertise. An onboarding process will consist of the SHM team and Reduce REVISITS team meeting virtually with mentors. Content will include processes related to baseline data collection and data reporting.

Evidence-Based Interventions for TOC

During phase 1, individual sites select evidence-based interventions they wish to incorporate into their COPD TOC program (Table 1) [11]. Sites will then be randomized to deliver these COPD TOC programs either in person or virtually. All selected interventions are either part of standard clinical practice or supported by strong evidence for effectiveness. These interventions include, first, COPD action plans; these are structured tools that patients and health care providers review together to support patient self-management. Both the American Lung Association and the COPD Foundation have published free action plans that were available for use by sites [17,18]. Second, the interventions include gold standard inhaler education delivered using the "teach-to-goal" method [1,19,20]. Teach-to-goal adheres to the guideline-recommended "teach-back" technique [21] and is an iterative, patient-centered approach that ensures correct inhaler technique through cycles of teaching and testing ("teach-back"). It adapts to diverse learning needs and delivery methods and has demonstrated lasting improvements in both inhaler technique and health outcomes, making it a best-practice method in COPD inhaler education [19,22]. Third, medication reconciliation enhancements were also available for sites to adopt or expand upon, aiming to improve adherence and patient safety [23]. Fourth, post-discharge follow-up visits were used, whether in-person or virtual, as they have been associated with lower 90-day readmission rates; sites could opt to initiate new visit types or strengthen existing workflows and show rates [24,25]. Similarly, sites could choose to focus on patient navigation programs using community health workers, which draws on successful asthma models to support care continuity and address social determinants of health [26]. Fifth, smoking cessation assistance is provided, as it has proven associations with reduced mortality, slowed disease progression, and decreased acute care use in COPD patients [27,28]. Sixth, pulmonary rehabilitation is included in the interventions, as it is a well-established intervention that lowers readmission risks [29]. Sites had the option to launch a new program or work to enhance referrals and attendance in existing programs. Seventh, improved diagnostic testing through increased spirometry testing was also offered as an intervention option, addressing frequent underdiagnosis or misdiagnosis of COPD [30]. Eighth, sites had the option to update electronic health record order sets for COPD patients to streamline inpatient and outpatient COPD management by standardizing best practices for diagnosis and treatment. Consults for COPD care, general COPD education, and standardized note templates incorporating the above best practices were also interventions that could be included in site plans.

XSL•FO
RenderX

Table 1. Data supporting implementation of selected interventions aimed at reducing readmissions.

Intervention	In-person outcomes or care standards	Virtual outcomes or care standards	References
Post-discharge follow-up visit	• Reduction in 90-day readmissions	• Noninferior to traditional home-based post-discharge program in time to first exacerbation	• Saxena et al [31] • Mínguez Clemente et al [32]
Inhaler teach-to-goal	• Reduction in 30-day acute care visits	• Noninferior in rates of inhaler misuse	• Press et al [33] • Press et al [34]
COPD action plan	• Standard practice	• Faster exacerbation recovery time compared to no action plan	• Farias et al [35]
Smoking cessation education	• Hospital-based education program increased prevalence of cessation at 12 months	• High rates of quit attempts and satisfaction with services received	• Fung et al [36] • Chase et al [37]
Medication reconciliation	• Toolkits available to disseminate best practices	• Pilot study was feasible and resulted in a high number of discrepancies recognized	• Mueller et al [38] • Heyworth et al [39]
Pulmonary rehabilitation	• Reduction in 30-day readmissions, increase in one-year readmissions	• Reduction in 30-day readmissions	• Myers et al [40] • Bhatt et al [41]
Remote monitoring	• No “in-person” option	• Some reduction in COPD-related readmissions • Low quality of evidence, no improvement in quality of life or mortality	• Janjua et al [42] • Nagase et al [43]

Implementation Support via MIM With or Without Co-Design

The MIM is a well-validated strategy to enhance the implementation, effectiveness, and sustainability of hospital-based quality initiatives (Table 2). Mentors will be randomly assigned to either one or 2 sites, based on their indicated time commitment to the study. For mentors with 2 assigned sites, both sites will be in the same randomized group (MIM only or MIM plus co-design). Mentor training will be in the form of a 2-day virtual “mentor university,” during which time they will receive best practices training in mentoring, delineate their roles and responsibilities, and discuss “case studies” of anticipated implementation issues they may encounter.

During the active implementation period of the study, mentors will meet monthly with their assigned sites (over the course of one year, for a total of 12 sessions) to enable on-the-ground troubleshooting of site-specific difficulties with TOC interventions. Monthly mentor calls will be highly structured and consist of initial updates of current events that could impact hospitalization rates, review of data, overview of implementation progress, review of plan-do-study-act cycles, brief review of further action items, and finally sustainability planning.

In addition to monthly mentoring sessions, mentors will also perform a virtual site visit (within the first several months of TOC implementation), allowing them to observe and respond to challenges occurring at the site level. These visits will be completed during the active implementation period of the study and consist of several meetings with multiple stakeholder groups (including the core intervention team, executive leadership, and implementing care providers). Templates of agendas and presentations will be provided to the mentors for their site-specific revisions, and site visit debriefs will be developed by mentors at the conclusion of the site visit to document the team’s ongoing efforts, direct their focus toward opportunities for greater adherence, and potentially garner additional institutional support.

Half of the sites will additionally be randomized to receive co-design support with our partner, Onda Collective. Co-design is human-centered design approach that invites end-users into the design process as partners to improve acceptability, usability, and feasibility of intervention implementation. This approach will give site leads, clinicians, patients, and caregivers the opportunity to think through detailed aspects of how a specific intervention will be implemented together (Table 2). Sites will work with Onda’s trained designers to co-design one of their 2 to 3 interventions in their COPD TOC bundle.

XSL•FO
RenderX

Table 2. Features of mentored implementation and co-design approaches.

Feature	Mentored Implementation [44-50]	Co-design [51-55]
Goals	• Culture change • Practice change • Capacity building • Leadership development	• Shared power • Team alignment • Integration of diverse expertise • Meeting end-user needs and objectives
Approach	• Consultative guidance • Ongoing implementation support; expertise	• Collaborative design • Patients, staff, and clinicians as participants, not just "informants" • Front-end process facilitation
Outputs	• Monthly agendas, summary notes • Virtual site visit records and reports • Site implementation planning workbooks	• Customer journey maps • Design criteria for implementation and alignment • Intervention implementation blueprints
Key activities	• Mentorship meetings and virtual site visit	• Participatory workshops
Team breakdown	• Physicians with quality improvement expertise • Project management support	• Academically trained designers

Randomization and Assignment of Interventions

Sites will be randomized to receive virtual or in-person delivery of COPD TOC programs along with either virtual or in-person MIM of these programs. Four study arms will thus be created, and sites will be randomized in a 1:1:1:1 fashion. Mentors will also be randomized to sites. Covariate-based constrained randomization will be used to achieve balance between groups. Mentor randomization will be performed via stratified randomization given uneven efforts of mentors; that is, some mentors worked at 2 sites while others did not.

Outcomes

The primary effectiveness outcome for this study will be COPD-specific acute health care use, defined as a composite of all ED visits and hospitalizations within 30 days of index hospitalization for a COPD exacerbation. The primary implementation outcome will be reach, defined as the proportion of patients receiving their assigned TOC interventions (the whole bundle). Secondary effective outcomes will include additional acute care metrics related to COPD-specific and all-cause 30-, 60-, and 90-day revisits, ED visits, and readmissions. Secondary implementation outcomes include the reach of individual interventions within sites' COPD TOC programs (ie, component-level reach).

Data Collection

Quantitative data (aims 2 and 3): deidentified patient-level data related to program effectiveness and intervention delivery will be collected monthly on all patients discharged with an index hospitalization for COPD exacerbation and all subsequent encounters (Textbox 1). These data include patient demographics (eg, age, race, ethnicity, and gender), zip code (first 3 digits), smoking history, insurance, acute care use using specific *International Classification of Diseases, Tenth Revision* (*ICD-10*) codes, and all subsequent hospitalizations, ED visits, and relevant ambulatory visits (eg, primary care, pulmonology, and pulmonary rehabilitation). [56] The *ICD-10* codes used to identify COPD will be J40, J41.0, J41.1, J41.8, J42, J43.0, J43.1, J43.2, J43.8, J43.9, J44.0, J44.1, and J44.9. The *ICD-10* codes used to identify COPD-related hospitalizations will be J96.00, J96.01, J96.02, J96.10, J96.20, J96.21, J96.22, J96.90, and R09.2 WITH a secondary code for COPD. The interventions received (ie, patient education and post-discharge encounters) will also be captured.

Qualitative data (study aims 1 and 3): For both the contextual assessments (aim 1) and the sustainment assessments (aim 3), site leads will be first surveyed and then interviewed, followed by interviews with clinicians, organizational leaders, patients, and caregivers. For the contextual assessment, a final interview session with site leads was conducted to develop their site-specific implementation plans (manuscript in press). For the sustainment evaluations, we will collect data on sustainment, fidelity to the site implementation plan, and cost analysis (Tables 3 and 4).

XSL•FO
RenderX

Textbox 1. Deidentified patient-level data.

Identification

- Fake medical record number (dummy number to track patients over time without being able to identify them)
- Fake encounter number (dummy number to track encounter-related data without being able to identify specific encounter)

Encounter information

- Encounter type
- Days since index discharge (we could not use real dates, sites provided us with the number of days since discharge from the index hospitalization for chronic obstructive pulmonary disease (COPD) whereby admission for the index was a negative number [0-day prior to discharge] and all subsequent encounters including ED visits, rehospitalizations, and outpatient visits were a positive number [0 + days until that encounter])
- Length of stay
- Discharge location

Diagnosis codes (for the index COPD hospitalization, specific *International Classification of Diseases, Tenth Revision (ICD-10)* codes were used; for all other encounters we captured all *ICD-10* codes)

- Primary *ICD-10* code
- Secondary *ICD-10* codes
- Diagnostic-related groups (optional)

Characteristics

- Primary insurance
- Secondary insurance
- Age at encounter
- Gender
- Race
- Ethnicity
- First 3 digits of zip code

COPD data (optional)

- Smoking status
- Disease duration
- Disease severity
- Pulmonary rehabilitation referral
- Smoking cessation referral
- Outpatient oxygen order
- Home health ordered

Intervention data (tailored to each site whereby primary metric was reach (yes or no) with other metrics decided at the site level)

- Interventions 1, 2, and 3 completed
- Other metrics for interventions 1, 2, and 3

Table 3. Sustainment, fidelity, and cost data collection by end-user group.

End-user group	Sustainment	Fidelity	Cost analysis
Site leads	✓	✓	✓
Organizational leaders	✓		
Clinicians	✓	✓	
Patients and caregivers	✓		

XSL•FO
RenderX

Table 4. Sustainment, fidelity, and cost domains.

Domain	Focus	End-user group	Data collection methods
Sustainment	Extent to which transitions of care (TOC) interventions are maintained post-implementation phase	Site leads, clinicians, organizational leaders, patients, caregivers	Survey, qualitative interviews
Fidelity	Degree to which sites adhere to their site-specific implementation plans	Site leads, clinicians	Survey, qualitative interviews
Cost analysis	Resources required to maintain the intervention and implementation supports	Site leads	Survey, qualitative interviews

Statistical Analysis and Power Calculation

This primary analysis will be an intention-to-treat analysis with both unadjusted and adjusted analyses using generalized estimating equations. The primary intention-to-treat analysis will be applied to all enrolled hospitals and their outcomes, regardless of duration of intervention or implementation and actual adherence to the randomization at each site. Two-factor generalized estimating equations with a Poisson or negative distribution will be used to assess the effect of delivery mode (virtual vs in-person) and implementation support strategy (mentoring vs mentoring plus co-design), along with their interaction. Randomization "group 1" (deliver interventions virtually with mentoring plus co-design implementation support) will be compared to the remainder of the groups (Table 1). Covariates will include patient age, gender, and race, among others; hospital region; hospital type (academic vs community); setting (urban vs suburban vs rural); and hospital size. Models will also adjust for the baseline level of the primary outcomes at each individual hospital. Hospitals will be considered as a random intercept in a generalized linear mixed model, while within-hospital associations will be adjusted as clusters in a generalized estimating equation. We will also conduct pairwise comparisons among the four groups. For the primary effectiveness (30-day COPD-specific revisits) and the primary implementation outcome (reach of bundle), we will use 2-factor GEE with binomial distribution. A 2-sided type I error of 1% will be used for primary outcomes to account for multiple testing. Secondary outcomes, including other COPD-specific and all-cause 30-, 60-, and 90-day ED visits, readmissions, and revisits, will use the same methods as described above. These analyses will be exploratory and will not spend type I error of 5%. Time-to-event outcomes will be assessed via survival analysis with Cox proportional hazard regression models and Kaplan-Meier curves with the log-rank test. Tests of interaction to assess for any relationship between the intervention delivery (virtual versus in-person) and implementation support (mentoring only versus mentoring plus co-design) will also be performed. Given the testing of multiple hypotheses, a 2-sided *P* value of 1% will be considered significant. For analyses of secondary outcomes, a *P* value <5% will be considered significant regardless of type I error spending. For binary or count outcomes such as intervention provided (yes or no), the same statistical methods used in the primary analyses will be applied.

The sample size required for the trial is a minimum of 16 sites. For the primary effectiveness outcome of 30-day COPD-specific acute health care use, assuming a 30-day acute care revisit difference of 0.24 between virtual and in-person groups and a cluster size of 602, as was seen in pilot data [19,22,57], along with an intraclass correlation of 0.03, a total of 16 clusters and 4628 patients is needed to have 91% power to detect this difference. For the implementation outcomes of bundle reach, using a modified *z* test and assuming a 20% difference between receipt of interventions between arms and 90% power to detect a difference, 15 clusters with 4560 patients are needed. Both of these power calculations are at a 2-sided significance level of 1%.

Sensitivity Analyses

We will assess the robustness of our findings with per-protocol and as-treated sensitivity analyses. Per protocol analysis is defined as comparing sites that adhered to their randomization for delivering interventions (all in person or all virtual), and as treated is defined as comparing sites regardless of randomization that delivered interventions all in person versus any virtual (hybrid virtual and in person). These will be repeated in each of the 4 arms, including testing for interaction between intervention strategy and implementation approach and among patient subgroups (eg, gender, age, and socioeconomic status) and hospital subgroups (region and hospital size). We will use 2-way generalized linear mixed models with hospitals considered as random intercepts.

Results

As of August 2025, 21 sites completed the contextual assessments and developed site-specific implementation plans. Publication of the qualitative data from this pre-implementation phase is anticipated in December 2025. Randomization of sites has been performed, and the sites that were randomized to co-design have completed 3 sessions. Baseline use data collection is complete. Use data collection for the implementation year is nearly complete. Use data collection for years 1 and 2 of the post-implementation phase is ongoing. Quantitative data analyses of the baseline and implementation phase reports are nearly complete, with anticipation of submission of the primary implementation phase manuscript in December 2025, pending submission of final implementation-phase data reports. Anticipate manuscript submission for the post-implementation sustainment analyses is anticipated for September 2026. Qualitative data collection for year 1 of the post-implementation phase is complete, and analysis is under way. Qualitative data collection for year 2 is being initiated as of August 2025.

XSL•FO
RenderX

Discussion

Principal Findings

The Reduce REVISITS study will use novel integrated implementation science and human-centered design methodology to investigate bundles of effective COPD TOC interventions with the goal of reducing COPD hospital revisits. Using a virtual adaptation of the MIM strategy with or without co-design support, the interventions under study, delivered in person or virtually, will be tested for feasibility and endurance in addition to effectiveness. By specifically studying virtual delivery of these interventions and their implementation, our study evaluates whether this approach improves reach by integrating into the rapidly advancing telehealth landscape and may be optimally translated into practice even when resources are finite. Our study outcomes include both effectiveness and implementation measures to enhance dissemination and broader-scale implementation of COPD TOC programs.

Prior literature has investigated the effectiveness of multilevel interventions in a variety of disease states. Studies targeting the transition from the adolescent to young adult health systems, for instance, are ongoing and are using multimodal approaches to target information gaps among patients, health care providers, and caregivers [58]. Congestive heart failure and COPD readmission care bundles and post-discharge follow-up programs have been evaluated as a group for overall effectiveness with varying results, particularly for the outcomes of readmission and mortality [24,59,60]. However, which elements of these programs are the most effective and can be easily delivered to multiple health systems are current gaps in knowledge that the Reduce REVISITS study will fill.

The post-pandemic era has ushered in a plethora of virtual health care measures aimed at improving quality and decreasing cost [61]. Even before the pandemic, a randomized controlled trial found that a telehealth-based transitions of care program involving patients with multiple chronic diseases increased medication reconciliation and adherence compared to an in-person–based program, although telehealth did not reduce health care use in the form of ED visits or readmissions [62]. Specific to COPD, virtual pulmonary rehabilitation has been found to be comparably effective to in-person therapy, but routine clinical monitoring provided virtually has not been demonstrated to reduce health care use or cost compared to standard of care in-person visits [63,64]. The use of emerging wearable technologies for remote patient monitoring is also being actively investigated in relation to clinical outcomes in patients with congestive heart failure or COPD [65]. Ensuring the effectiveness of these virtual interventions through studies such as the Reduce REVISITS study and others will undoubtedly improve the care of chronically ill hospitalized patients.

While this study aims to find the most effective and scalable approaches to facilitate COPD TOC, it is unlikely that a one-size-fits-all approach exists for this complex, diverse, and comorbid patient population. Thus, we have specifically designed our study across a wide variety of hospital settings to capture as much of this variation as possible. Another potential limitation could be that patients could seek care outside of their previously used health care system, limiting data collection. By randomizing sites, we hope to mitigate this bias. We additionally will evaluate sites via value-based purchasing programs that have complete revisit data when possible.

In summary, TOC programs serve a vital role in the necessary goal of reducing admissions for COPD and other chronic diseases. The Reduce REVISITS study, with its novel methodology, wide scope of tested interventions, and introduction of co-design support, is poised to transform the landscape of multilevel systems-based interventions for this vulnerable patient population.

Conclusions

The Reduce REVISITS study uses a pragmatic hybrid type II effectiveness-implementation design to evaluate COPD care transition programs delivered across hospital settings. By integrating mentored implementation with human-centered co-design, this work seeks to identify practical and scalable strategies for delivering evidence-based COPD TOC interventions in both virtual and in-person formats. The study will generate needed evidence on how implementation supports influence reach, effectiveness, and long-term sustainment. Findings will help guide hospitals in optimizing COPD care transitions and reducing preventable acute care use in this high-risk patient population.

Funding

This study was supported by the National Heart, Lung, and Blood Institute (R01HL146644).

Data Availability

No data have been generated or analyzed yet, as this manuscript reports the study protocol. In accordance with the study's registration on ClinicalTrials.gov (NCT05568043), individual participant data will not be shared. Upon study completion, aggregated results and related study materials (eg, survey instruments, codebooks, and analysis plans) will be made available in subsequent publications and, where appropriate, in National Institutes of Health–approved repositories in accordance with National Heart, Lung, and Blood Institute data-sharing policies and upon reasonable request.

XSL•FO
RenderX

Authors' Contributions

CTL contributed to writing (original draft, review, and editing). LT and MA contributed to data collection, data analysis, and writing (editing). AEF contributed to data collection and writing (editing). JG, KE, LD, JR, and HP contributed to data collection. AA and PL contributed to data collection and writing (editing). WW andAJS contributed to data collection, data analysis, and writing (editing). VGP contributed to conceptualization, methodology, and writing (review and editing).

Conflicts of Interest

VGP also receives support from the National Institutes of Health (K24HL163408) and the Agency for Healthcare Research and Quality (R01HS027804) and reports consulting fees from Humana. AA receives royalties from UpToDate, for editing role. Grants from NHLBI, PCORI, AHRQ, and NCATS (Paid to institution). All other authors declare no conflicts of interest. Portions of this work were presented as poster presentations at the International American Thoracic Society meeting in May 2025.

Multimedia Appendix 1
Peer review report from DIRH - Dissemination and Implementation Research in Health Study Section, Healthcare Delivery and Methodologies Integrated Review Group (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 173 KB - resprot_v15i1e82043_app1.pdf]

References

1. Press VG, Arora VM, Shah LM, Lewis SL, Ivy K, Charbeneau J, et al. Misuse of respiratory inhalers in hospitalized patients with asthma or COPD. J Gen Intern Med 2011;26(6):635-642 [FREE Full text] [doi: 10.1007/s11606-010-1624-2] [Medline: 21249463]
2. Hospital readmissions reduction program (HRRP). Centers for Medicare and Medicaid Services. URL: https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program [accessed 2022-01-06]
3. Jencks SF, Williams MV, Coleman EA. Rehospitalizations among patients in the medicare fee-for-service program. N Engl J Med 2009;360(14):1418-1428 [FREE Full text] [doi: 10.1056/NEJMsa0803563] [Medline: 19339721]
4. Press VG, Au DH, Bourbeau J, Dransfield MT, Gershon AS, Krishnan JA, et al. Reducing chronic obstructive pulmonary disease hospital readmissions. an official American thoracic society workshop report. Ann Am Thorac Soc 2019;16(2):161-170 [FREE Full text] [doi: 10.1513/AnnalsATS.201811-755WS] [Medline: 30707066]
5. Press VG, Myers LC, Feemster LC. Preventing COPD readmissions under the hospital readmissions reduction program: how far have we come. Chest 2021;159(3):996-1006 [FREE Full text] [doi: 10.1016/j.chest.2020.10.008] [Medline: 33065106]
6. Press VG, Randall K, Hanser A. Evaluation of COPD chronic care management collaborative to reduce emergency department and hospital revisits across U.S. hospitals? Chronic Obstr Pulm Dis 2022;9(2):209-225 [FREE Full text] [doi: 10.15326/jcopdf.2021.0273] [Medline: 35322625]
7. Prieto-Centurion V, Markos MA, Ramey NI, Gussin HA, Nyenhuis SM, Joo MJ, et al. Interventions to reduce rehospitalizations after chronic obstructive pulmonary disease exacerbations. A systematic review. Ann Am Thorac Soc 2014;11(3):417-424 [FREE Full text] [doi: 10.1513/AnnalsATS.201308-254OC] [Medline: 24423379]
8. Aboumatar H, Naqibuddin M, Chung S, Chaudhry H, Kim SW, Saunders J, et al. Effect of a hospital-initiated program combining transitional care and long-term self-management support on outcomes of patients hospitalized with chronic obstructive pulmonary disease: a randomized clinical trial. JAMA 2019;322(14):1371-1380 [FREE Full text] [doi: 10.1001/jama.2019.11982] [Medline: 31593271]
9. Bhatt SP, Wells JM, Iyer AS, Kirkpatrick DP, Parekh TM, Leach LT, et al. Results of a medicare bundled payments for care improvement initiative for chronic obstructive pulmonary disease readmissions. Ann Am Thorac Soc 2017;14(5):643-648 [FREE Full text] [doi: 10.1513/AnnalsATS.201610-775BC] [Medline: 28005410]
10. Waltman A, Konetzka RT, Chia S, Ghani A, Wan W, White SR, et al. Effectiveness of a bundled payments for care improvement program for chronic obstructive pulmonary disease. J Gen Intern Med 2023;38(12):2662-2670 [FREE Full text] [doi: 10.1007/s11606-023-08249-6] [Medline: 37340256]
11. Akula M, Traeger L, Damschroder L, Rommes JM, Erwin K, Press VG. The revisits study: rapid qualitative analysis of multi-group participant data for COPD transition of care bundle implementation planning across 21 U.S. hospitals. Am J Respir Crit Care Med 2025 May;211(Abstracts):A4952-A4952. [doi: 10.1164/ajrccm.2025.211.Abstracts.A4952]
12. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
13. Auerbach AD, Patel MS, Metlay JP, Schnipper JL, Williams MV, Robinson EJ, et al. The hospital medicine reengineering network (HOMERuN): a learning organization focused on improving hospital care. Acad Med 2014;89(3):415-420 [FREE Full text] [doi: 10.1097/ACM.0000000000000139] [Medline: 24448050]
14. Respiratory emergency visits using implementation science interventions tailored to settings. Society of Hospital Medicine. URL: https://www.hospitalmedicine.org/clinical-topics/revisits-study/ [accessed 2025-08-05]

XSL•FO
RenderX

15. Press VG, Goldstein J, Traeger L, Fegley A, Rommes JM, Auerbach A. Challenges and opportunities with recruiting and enrolling diverse U.S. hospitals for a cluster-randomized trial of COPD transition-of-care programs. Am J Respir Crit Care Med 2023;207(1):A3147. [doi: 10.1164/ajrccm-conference.2023.207.1_MeetingAbstracts.A3147]
16. Press V. Adapting A mentored implementation model for a multi-site copd transitions of care quality improvement study. Journal of Hospital Medicine. URL: https://shmabstracts.org/abstract/adapting-a-mentored-implementation-model-for-a-multi-site-copd-transitions-of-care-quality-improvement-study/ [accessed 2025-07-30]
17. My COPD Action and Management Plan. American Lung Association. URL: https://www.lung.org/lung-health-diseases/lung-disease-lookup/copd/living-with-copd/copd-management-tools [accessed 2024-11-21]
18. My COPD action plan. COPD Foundation. URL: https://www.copdfoundation.org/Praxis/Resource-Repository/Search/Article/404/My-COPD-Action-Plan.aspx [accessed 2024-11-21]
19. Press VG, Arora VM, Shah LM, Lewis SL, Charbeneau J, Naureckas ET, et al. Teaching the use of respiratory inhalers to hospitalized patients with asthma or COPD: a randomized trial. J Gen Intern Med 2012;27(10):1317-1325 [FREE Full text] [doi: 10.1007/s11606-012-2090-9] [Medline: 22592354]
20. Paasche-Orlow MK, Riekert KA, Bilderback A, Chanmugam A, Hill P, Rand CS, et al. Tailored education may reduce health literacy disparities in asthma self-management. Am J Respir Crit Care Med 2005;172(8):980-986 [FREE Full text] [doi: 10.1164/rccm.200409-1291OC] [Medline: 16081544]
21. Agustí A, Celli BR, Criner GJ, Halpin D, Anzueto A, Barnes P, et al. Global initiative for chronic obstructive lung disease 2023 report: gold executive summary. Am J Respir Crit Care Med 2023;207(7):819-837 [FREE Full text] [doi: 10.1164/rccm.202301-0106PP] [Medline: 36856433]
22. Press VG, Arora VM, Trela KC, Adhikari R, Zadravecz FJ, Liao C, et al. Effectiveness of interventions to teach metered-dose and diskus inhaler techniques. a randomized trial. Ann Am Thorac Soc 2016;13(6):816-824 [FREE Full text] [doi: 10.1513/AnnalsATS.201509-603OC] [Medline: 26998961]
23. Mueller SK, Kripalani S, Stein J, Kaboli P, Wetterneck TB, Salanitro AH, et al. A toolkit to disseminate best practices in inpatient medication reconciliation: multi-center medication reconciliation quality improvement study (MARQUIS). Jt Comm J Qual Patient Saf 2013;39(8):371-382 [FREE Full text] [doi: 10.1016/s1553-7250(13)39051-5] [Medline: 23991510]
24. Saxena FE, Bierman AS, Glazier RH, Wang X, Guan J, Lee DS, et al. Association of early physician follow-up with readmission among patients hospitalized for acute myocardial infarction, congestive heart failure, or chronic obstructive pulmonary disease. JAMA Netw Open 2022;5(7):e2222056 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.22056] [Medline: 35819782]
25. Mínguez Clemente P, Pascual-Carrasco M, Mata Hernández C, Malo de Molina R, Arvelo LA, Cadavid B, et al. Follow-up with telemedicine in early discharge for COPD exacerbations: randomized clinical trial (TELEMEDCOPD-Trial). COPD 2021;18(1):62-69 [FREE Full text] [doi: 10.1080/15412555.2020.1857717] [Medline: 33307857]
26. Parekh TM, Copeland CR, Dransfield MT, Cherrington A. Application of the community health worker model in adult asthma and COPD in the U.S.: a systematic review. BMC Pulm Med 2019;19(1):116 [FREE Full text] [doi: 10.1186/s12890-019-0878-7] [Medline: 31242944]
27. Tashkin DP. Smoking cessation in chronic obstructive pulmonary disease. Semin Respir Crit Care Med 2015;36(4):491-507. [doi: 10.1055/s-0035-1555610] [Medline: 26238637]
28. Godtfredsen NS, Vestbo J, Osler M, Prescott E. Risk of hospital admission for COPD following smoking cessation and reduction: a Danish population study. Thorax 2002;57(11):967-972 [FREE Full text] [doi: 10.1136/thorax.57.11.967] [Medline: 12403880]
29. Myers LC, Faridi MK, Hasegawa K, Camargo CA. Pulmonary rehabilitation and readmission rates for medicare beneficiaries with acute exacerbation of chronic obstructive pulmonary disease. Chronic Obstr Pulm Dis 2021;8(4):427-440 [FREE Full text] [doi: 10.15326/jcopdf.2020.0193] [Medline: 34329550]
30. Diab N, Gershon AS, Sin DD, Tan WC, Bourbeau J, Boulet L, et al. Underdiagnosis and overdiagnosis of chronic obstructive pulmonary disease. Am J Respir Crit Care Med 2018;198(9):1130-1139. [doi: 10.1164/rccm.201804-0621CI] [Medline: 29979608]
31. Saxena FE, Bierman AS, Glazier RH, Wang X, Guan J, Lee DS, et al. Association of early physician follow-up with readmission among patients hospitalized for acute myocardial infarction, congestive heart failure, or chronic obstructive pulmonary disease. JAMA Netw Open 2022 Jul 01;5(7):e2222056 [FREE Full text] [doi: 10.1001/jamanetworkopen.2022.22056] [Medline: 35819782]
32. Mínguez Clemente P, Pascual-Carrasco M, Mata Hernández C, Malo de Molina R, Arvelo LA, Cadavid B, et al. Follow-up with telemedicine in early discharge for COPD exacerbations: randomized clinical trial (TELEMEDCOPD-Trial). COPD 2021 Feb;18(1):62-69 [FREE Full text] [doi: 10.1080/15412555.2020.1857717] [Medline: 33307857]
33. Press VG, Arora VM, Trela KC, Adhikari R, Zadravecz FJ, Liao C, et al. Effectiveness of interventions to teach metered-dose and diskus inhaler techniques. A randomized trial. Ann Am Thorac Soc 2016 Jun;13(6):816-824 [FREE Full text] [doi: 10.1513/AnnalsATS.201509-603OC] [Medline: 26998961]

XSL•FO
RenderX

34. Press VG, Arora VM, Kelly CA, Carey KA, White SR, Wan W. Effectiveness of virtual vs in-person inhaler education for hospitalized patients with obstructive lung disease: a randomized clinical trial. JAMA Netw Open 2020 Jan 03;3(1):e1918205 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.18205] [Medline: 31899529]
35. Farias R, Sedeno M, Beaucage D, Drouin I, Ouellet I, Joubert A, et al. Innovating the treatment of COPD exacerbations: a phone interactive telesystem to increase COPD Action Plan adherence. BMJ Open Respir Res 2019;6(1):e000379 [FREE Full text] [doi: 10.1136/bmjresp-2018-000379] [Medline: 31178998]
36. Fung PR, Snape-Jenkinson SL, Godfrey MT, Love KW, Zimmerman PV, Yang IA, et al. Effectiveness of hospital-based smoking cessation. Chest 2005 Jul;128(1):216-223. [doi: 10.1378/chest.128.1.216] [Medline: 16002938]
37. Chase W, Zurmehly J, Amaya M, Browning KK. Implementation of a smoking cessation e-Learning education program for oncology clinic healthcare providers: evaluation with implications for evidence-based practice. Worldviews Evid Based Nurs 2020 Dec;17(6):476-482. [doi: 10.1111/wvn.12476] [Medline: 33215842]
38. Mueller SK, Kripalani S, Stein J, Kaboli P, Wetterneck TB, Salanitro AH, et al. A toolkit to disseminate best practices in inpatient medication reconciliation: multi-center medication reconciliation quality improvement study (MARQUIS). The Joint Commission Journal on Quality and Patient Safety 2013 Aug;39(8):371-AP3. [doi: 10.1016/S1553-7250(13)39051-5]
39. Heyworth L, Paquin AM, Clark J, Kamenker V, Stewart M, Martin T, et al. Engaging patients in medication reconciliation via a patient portal following hospital discharge. J Am Med Inform Assoc 2014 Feb;21(e1):e157-e162 [FREE Full text] [doi: 10.1136/amiajnl-2013-001995] [Medline: 24036155]
40. Myers LC, Faridi MK, Hasegawa K, Camargo CA. Pulmonary rehabilitation and readmission rates for medicare beneficiaries with acute exacerbation of chronic obstructive pulmonary disease. Chronic Obstr Pulm Dis 2021 Oct 28;8(4):427-440 [FREE Full text] [doi: 10.15326/jcopdf.2020.0193] [Medline: 34329550]
41. Bhatt SP, Patel SB, Anderson EM, Baugh D, Givens T, Schumann C, et al. Video telehealth pulmonary rehabilitation intervention in chronic obstructive pulmonary disease reduces 30-day readmissions. Am J Respir Crit Care Med 2019 Aug 15;200(4):511-513 [FREE Full text] [doi: 10.1164/rccm.201902-0314LE] [Medline: 30978302]
42. Janjua S, Carter D, Threapleton CJD, Prigmore S, Disler RT. Telehealth interventions: remote monitoring and consultations for people with chronic obstructive pulmonary disease (COPD). Cochrane Database Syst Rev 2021 Jul 20;7(7):CD013196 [FREE Full text] [doi: 10.1002/14651858.CD013196.pub2] [Medline: 34693988]
43. Nagase FI, Stafinski T, Avdagovska M, Stickland MK, Etruw EM, Menon D. Effectiveness of remote home monitoring for patients with chronic obstructive pulmonary disease (COPD): systematic review. BMC Health Serv Res 2022 May 14;22(1):646 [FREE Full text] [doi: 10.1186/s12913-022-07938-y] [Medline: 35568904]
44. Schnipper JL, Reyes Nieva H, Mallouk M, Mixon A, Rennke S, Chu E, MARQUIS2 Site Leaders, MARQUIS2 Study Group. Effects of a refined evidence-based toolkit and mentored implementation on medication reconciliation at 18 hospitals: results of the MARQUIS2 study. BMJ Qual Saf 2022;31(4):278-286 [FREE Full text] [doi: 10.1136/bmjqs-2020-012709] [Medline: 33927025]
45. Li J, Hinami K, Hansen LO, Maynard G, Budnitz T, Williams MV. The physician mentored implementation model: a promising quality improvement framework for health care change. Acad Med 2015;90(3):303-310. [doi: 10.1097/ACM.0000000000000547] [Medline: 25354069]
46. Maynard G, Budnitz T, Nickel W, Greenwald J, Kerr K, Miller J, et al. 2011 John M. Eisenberg patient safety and quality awards. Mentored implementation: building leaders and achieving results through a collaborative improvement model. Innovation in patient safety and quality at the national level. Jt Comm J Qual Patient Saf 2012;38(7):301-310. [doi: 10.1016/s1553-7250(12)38040-9] [Medline: 22852190]
47. Mixon AS, Smith GR, Mallouk M, Nieva HR, Kripalani S, Rennke S, MARQUIS2 Site Leaders, MARQUIS2 Study Group. Design of MARQUIS2: study protocol for a mentored implementation study of an evidence-based toolkit to improve patient safety through medication reconciliation. BMC Health Serv Res 2019;19(1):659 [FREE Full text] [doi: 10.1186/s12913-019-4491-5] [Medline: 31511070]
48. O'Leary KJ, Johnson JK, Manojlovich M, Goldstein JD, Lee J, Williams MV. Redesigning systems to improve teamwork and quality for hospitalized patients (RESET): study protocol evaluating the effect of mentored implementation to redesign clinical microsystems. BMC Health Serv Res 2019;19(1):293 [FREE Full text] [doi: 10.1186/s12913-019-4116-z] [Medline: 31068161]
49. Starmer AJ, Spector ND, O'Toole JK, Bismilla Z, Calaman S, Campos M, I-PASS SHM Mentored Implementation Study Group. Implementation of the I-PASS handoff program in diverse clinical environments: a multicenter prospective effectiveness implementation study. J Hosp Med 2023 Jan;18(1):5-14 [FREE Full text] [doi: 10.1002/jhm.12979] [Medline: 36326255]
50. O'Leary KJ, Johnson JK, Williams MV, Estrella R, Hanrahan K, Leykum LK, et al. Effect of complementary interventions to redesign care on teamwork and quality for hospitalized medical patients : a pragmatic controlled trial. Ann Intern Med 2023;176(11):1456-1464. [doi: 10.7326/M23-0953] [Medline: 37903367]
51. Trischler J, Pervan SJ, Kelly SJ, Scott DR. The value of codesign: the effect of customer involvement in service design teams. J Serv Res 2018;21(1):75-100. [doi: 10.1177/1094670517714060]
52. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008;4(1):5-18. [doi: 10.1080/15710880701875068]

XSL•FO
RenderX

53. Donetto S, Pierri P, Tsianakas V, Robert G. Experience-based Co-design and healthcare improvement: realizing participatory design in the public sector. Des J 2015;18(2):227-248. [doi: 10.2752/175630615X14212498964312]
54. Kynoch K, Ramis M. Experience based co-design in acute healthcare services: a scoping review protocol. JBI Database System Rev Implement Rep 2019;17(1):3-9. [doi: 10.11124/JBISRIR-2017-003655] [Medline: 30024434]
55. Bird M, McGillion M, Chambers EM, Dix J, Fajardo CJ, Gilmour M, et al. A generative co-design framework for healthcare innovation: development and application of an end-user engagement framework. Res Involv Engagem 2021;7(1):12 [FREE Full text] [doi: 10.1186/s40900-021-00252-7] [Medline: 33648588]
56. Stein BD, Bautista A, Schumock GT, Lee TA, Charbeneau JT, Lauderdale DS, et al. The validity of international classification of diseases, ninth revision, clinical modification diagnosis codes for identifying patients hospitalized for COPD exacerbations. Chest 2012;141(1):87-93 [FREE Full text] [doi: 10.1378/chest.11-0024] [Medline: 21757568]
57. Press VG, Arora VM, Kelly CA, Carey KA, White SR, Wan W. Effectiveness of virtual vs in-person inhaler education for hospitalized patients with obstructive lung disease: a randomized clinical trial. JAMA Netw Open 2020;3(1):e1918205 [FREE Full text] [doi: 10.1001/jamanetworkopen.2019.18205] [Medline: 31899529]
58. Garland BH, Majumder M, Wiemann CM, Sanchez-Fournier B, Babla J, Hergenroeder AC. Development of a multi-level/multi-modal intervention for health care transition preparation. Health Care Transit 2024;2:100063 [FREE Full text] [doi: 10.1016/j.hctj.2024.100063] [Medline: 39712581]
59. Reese RL, Clement SA, Syeda S, Hawley CE, Gosian JS, Cai S, et al. Coordinated-transitional care for veterans with heart failure and chronic lung disease. J Am Geriatr Soc 2019;67(7):1502-1507 [FREE Full text] [doi: 10.1111/jgs.15978] [Medline: 31081946]
60. Hsiao YL, Bass EB, Wu AW, Kelly D, Sylvester C, Berkowitz SA, Johns Hopkins Community Health Partnership (J-CHiP) Team. Preventing avoidable rehospitalizations through standardizing management of chronic conditions in skilled nursing facilities. J Am Med Dir Assoc 2023;24(12):1910-1917.e3. [doi: 10.1016/j.jamda.2023.08.010] [Medline: 37690461]
61. Herzer KR, Pronovost PJ. Ensuring quality in the era of virtual care. JAMA 2021;325(5):429-430. [doi: 10.1001/jama.2020.24955] [Medline: 33528544]
62. Noel K, Messina C, Hou W, Schoenfeld E, Kelly G. Tele-transitions of care (TTOC): a 12-month, randomized controlled trial evaluating the use of telehealth to achieve triple aim objectives. BMC Fam Pract 2020;21(1):27 [FREE Full text] [doi: 10.1186/s12875-020-1094-5] [Medline: 32033535]
63. Michaelchuk W, Oliveira A, Marzolini S, Nonoyama M, Maybank A, Goldstein R, et al. Design and delivery of home-based telehealth pulmonary rehabilitation programs in COPD: a systematic review and meta-analysis. Int J Med Inform 2022;162:104754. [doi: 10.1016/j.ijmedinf.2022.104754] [Medline: 35395474]
64. Soriano JB, García-Río F, Vázquez-Espinosa E, Conforto JI, Hernando-Sanz A, López-Yepes L, et al. A multicentre, randomized controlled trial of telehealth for the management of COPD. Respir Med 2018;144:74-81 [FREE Full text] [doi: 10.1016/j.rmed.2018.10.008] [Medline: 30366588]
65. Harris S, Paynter K, Guinn M, Fox J, Moore N, Maddox TM, et al. Post-hospitalization remote monitoring for patients with heart failure or chronic obstructive pulmonary disease in an accountable care organization. BMC Health Serv Res 2024;24(1):69 [FREE Full text] [doi: 10.1186/s12913-023-10496-6] [Medline: 38218820]

Abbreviations

ATS: American Thoracic Society
COPD: chronic obstructive pulmonary disease
ED: emergency department
HOMERuN: Hospital Medicine Reengineering Network
HRRP: Hospital Readmission Reduction Program
ICD-10: International Classification of Diseases, Tenth Revision
MIM: mentored implementation model
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
REVISITS: Respiratory Emergent Visits using Implementation Science Interventions Tailored to Setting
SHM: Society of Hospital Medicine
TOC: transition of care

XSL•FO
RenderX

Edited by J Sarvestan; submitted 08.Aug.2025; this is a non–peer-reviewed article;accepted 06.Nov.2025; published 13.Jan.2026.

Please cite as:
Lee CT, Traeger L, Akula M, Fegley AE, Goldstein J, Erwin K, Damschroder LJ, Rommes J, Pick H, Auerbach A, Lindenauer P, Wan W, Jackson Sagredo A, Press VG
Virtual Mentored Implementation to Improve Care Transitions in Chronic Obstructive Pulmonary Disease: Protocol for a Pragmatic Implementation Study
JMIR Res Protoc 2026;15:e82043
URL: https://www.researchprotocols.org/2026/1/e82043
doi:10.2196/82043
PMID:

©Cathryn T Lee, Leah Traeger, Mahima Akula, April E Fegley, Jenna Goldstein, Kim Erwin, Laura J Damschroder, Jean Rommes, Hannah Pick, Andrew Auerbach, Peter Lindenauer, Wen Wan, Andrea Jackson Sagredo, Valerie G Press. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Factors Associated With Alcohol Use After Metabolic and Bariatric Surgery: Protocol for an Ecological Momentary Assessment

Lisa R Miller-Matero[1,2,3], PhD; Daniel Saleh[2], BA; Brittany Christopher[2], BA; Maha Albujuq[2], BS; Alyssa Vanderziel[2,3], PhD; Erin N Haley[2,3], PhD; Jordan M Braciszewski[2,3], PhD; Roland S Moore[4], PhD; Arthur M Carlin[5], MD; Kristina M Jackson[6], PhD

[1]Behavioral Health, Henry Ford Health System, Detroit, MI, United States

[2]Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, United States

[3]Michigan State University, East Lansing, MI, United States

[4]Pacific Institute For Research and Evaluation, Berkeley, CA, United States

[5]Department of Surgery, Henry Ford Health System, Detroit, MI, United States

[6]Department of Psychiatry, Robert Wood Johnson Medical School, Rutgers, The State University of New Jersey, New Brunswick, NJ, United States

Corresponding Author:
Lisa R Miller-Matero, PhD
Behavioral Health
Henry Ford Health System
1 Ford Pl, 5E
Detroit, MI, 48083-5070
United States
Phone: 1 3132148240
Email: lmatero1@hfhs.org

Abstract

Background: Individuals who undergo metabolic and bariatric surgery (MBS) are at increased risk for postoperative alcohol use disorder. Reducing postoperative alcohol use could prevent the development of alcohol use disorder; however, the factors leading to episodic alcohol use are not known.

Objective: The purpose of this paper is to describe the protocol for a study that will examine distal and proximal factors associated with episodic alcohol use and hazardous alcohol use among individuals who undergo MBS.

Methods: We will enroll 100 participants who undergo MBS at a single health care system. Participants will complete measures of substance use, psychiatric symptoms, and disordered eating behaviors at baseline and at 6- and 12-week follow-ups. Participants will also complete a 3-week ecological momentary assessment protocol in which they will complete brief surveys each morning and evening, reporting on their mood, disordered eating, and substance use.

Results: This study received funding from the National Institute on Alcohol Abuse and Alcoholism (R21 AA029423) in May 2023. This part of the grant was approved by the institutional review board in March 2024, and data collection occurred between November 2024 and December 2025. We anticipate that our study protocol will be feasible and that we will observe at least 80% participant retention at the follow-up assessments and their response to at least 75% of ecological momentary assessment signals. We hypothesize that depressive symptoms (distal factor) and negative affect (proximal factor) will be associated with increased alcohol use, and alcohol use will occur in lieu of disordered eating behaviors.

Conclusions: Findings will help us understand distal and proximal factors leading to episodic alcohol use after undergoing MBS. This knowledge will allow us to construct better monitoring strategies for postoperative alcohol use within MBS programs and identify targets for intervention to reduce alcohol use after undergoing MBS.

International Registered Report Identifier (IRRID): DERR1-10.2196/87209

(JMIR Res Protoc 2026;15:e87209) doi:10.2196/87209

KEYWORDS

metabolic and bariatric surgery; alcohol use; affect; disordered eating; ecological momentary assessment

XSL•FO
RenderX

Introduction

Metabolic and bariatric surgery (MBS) is the most effective method for treating class 3 obesity, even compared to glucagon-like peptide-1 receptor agonist medications [1,2]. Despite the ability of MBS to result in significant weight loss and improvement in medical comorbidities, some patients experience adverse outcomes. One potential adverse outcome following MBS is the increased risk for alcohol use disorder (AUD). Up to 1 in 5 individuals may have an AUD 5 years after MBS, and most of these are considered new onset [3,4]. The increased risk of AUD appears to begin between 1 and 2 years after surgery and continues to rise as time elapses [3,5].

There are several possible explanations for why patients develop new problems with alcohol after surgery. The pharmacokinetics of alcohol change after MBS. For patients who have undergone sleeve gastrectomy (SG) or Roux-en-Y gastric bypass (RYGB; the most common procedure types), blood alcohol concentrations increase faster and peak higher [6-8]. Patients report increased sensitivity to the effects of alcohol after MBS [9-11]. MBS alters the ghrelin system, which may increase the rewarding effects of alcohol [12,13]. There is also a hypothesis that patients may be "transferring" their addiction [14,15]. Certain types of food act similarly in the brain as other addictive substances [16-23]. In functional magnetic resonance imaging studies, drug-related cues for those with a drug addiction and food-related cues for those with a "food addiction" show activation in the same areas of the brain (ie, the amygdala, orbitofrontal cortex, striatum, and insula) [16-23]. Patients identified as having preoperative problems with high-sugar or low-fat or high-glycemic-index foods (ie, foods thought to be addictive) were more likely to have a new onset of AUD after MBS [24]. However, other studies have not supported the addiction transfer hypothesis [25]. Previous work has also identified factors that are associated with increased risk of AUD following MBS. Characteristics associated with postoperative alcohol use and AUD include being male, younger age, preoperative alcohol use, other substance use, depression, and disordered eating behaviors [25,26].

Despite some useful scholarly work informing mechanisms of AUD, existing studies have several limitations. First, much of the work that has been done has examined preoperative factors while neglecting postoperative factors associated with alcohol use. Symptoms such as depression and disordered eating behaviors can change following MBS and may be worthwhile targets of intervention in the postoperative period [27,28]. Second, the preoperative factors that have been identified are broad and apply to a wide range of patients, making it challenging to identify high-risk individuals. Third, existing research has only evaluated average alcohol use with retrospective recall. Finally, no studies have examined proximal factors that lead to episodic alcohol use after MBS. Identifying postoperative factors associated with episodic alcohol use after MBS could lead to new targets for interventions to reduce alcohol use after MBS, thus preventing the development of postoperative AUD.

To identify proximal factors associated with episodic alcohol use (eg, mood and eating), we propose to use ecological momentary assessment (EMA). In EMA, participants respond to brief questions on their mobile device (eg, smartphone) in "real time" while involved in usual daily activities. EMA methodology improves the accuracy of reporting because it provides in-the-moment, real-world assessments, which capture nuanced patterns of alcohol use and the dynamics between use and proximal factors of episodic use that are unreliable with retrospective evaluation [29,30]. Measurement of proximal factors with distal factors (eg, psychiatric symptoms and drinking motives) could inform goals for various types of interventions to reduce alcohol use.

In this paper, we describe the study protocol of a project that seeks to characterize episodic alcohol use after MBS. The aim of this investigation is to determine the feasibility of the study protocol and understand the distal and proximal factors associated with postoperative episodic alcohol use.

Methods

Participants and Procedures

The surgery types performed at this health system are RYGB and SG. We plan to enroll 100 participants who are between 6 months and 5 years after MBS. This will allow us to capture a range of drinking behavior, from reinitiation (to minimize recall bias) to continued use, by enrolling those who may have begun to escalate their frequency and amount of alcohol use [4,31]. Patients who underwent MBS at our health system will be randomly selected in batches and contacted via email with a brief description of the study and a link to a brief online eligibility survey through REDCap (Research Electronic Data Capture; Vanderbilt University) software, a secure and Health Insurance Portability and Accountability Act (HIPAA)–compliant web-based program. Patients will be asked if they have consumed alcohol since surgery. If so, they will report the frequency of use over the previous month (ie, not at all, once per month, 2-3 times per month, once per week, 2-3 times per week, 4-6 times per week, or every day). Patients will be eligible if they have consumed alcohol at least 2 to 3 times in the previous month, ensuring enrollment of those with current alcohol use. This criterion was also selected, given our goal to capture at least 2 drinking episodes during the EMA period. Once eligibility is determined, a research assistant will contact eligible patients via email or phone to schedule a time to complete informed consent and the baseline assessment. After obtaining participant consent to participate in the study, the research assistant will email a link to the baseline assessment, which will be completed through REDCap. Participants will also complete the assessment battery through REDCap at 6- and 12-week follow-ups (Figure 1). Participants will receive up to 5 reminders via email or phone if they have not completed each assessment. In addition, after completing the baseline assessment, each participant will complete 3 weeks of EMA. The EMA will include brief assessments (eg, 2-3 min) delivered twice per day to each participant's smartphone. We propose to collect data each morning (eg, 8 AM) and in the evening before bedtime (eg, 8 PM); administering 2 surveys will reduce recall

XSL•FO
RenderX

bias while also being cognizant of participant burden. Morning and evening assessments are frequently used in other EMA studies that have used a time-based approach [30]. An automatic push notification from the EMA app will be sent to the participant to signal when the items can be completed. Participants will be informed that the survey will be available on their dashboard 2 hours before the push notification is scheduled to allow for flexibility in the event that the participant wishes to complete it before receiving the push notification. Automatic reminders will be sent at 1 and 2 hours after the push notification if the survey has not yet been completed. Each signal will expire 4 hours after the push notification is sent. During the consent appointment, the research assistant will instruct participants on how to download and use the EMA app. If a participant misses multiple EMA signals in a row, the research assistant will attempt to contact the participant via email or phone.

Figure 1. Timeline of participation. EMA: ecological momentary assessment.

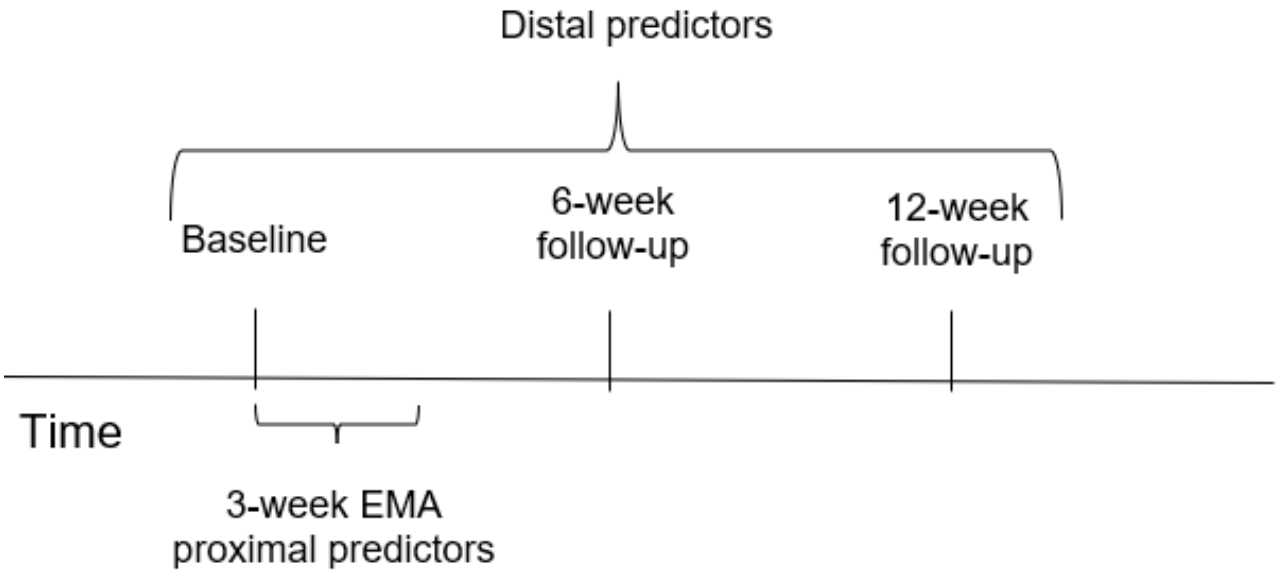

Measures: Distal Factors

Measures for the distal factors will be administered at baseline and at 6 and 12 weeks unless otherwise noted (Table 1).

Table 1. Measures at the baseline and 6- and 12-week assessments.

	Baseline	6 weeks	12 weeks
Demographics and background information	✓		
Preoperative substance use	✓		
Timing of postoperative substance use	✓		
Alcohol use (past 6 weeks)	✓	✓	✓
Hazardous alcohol use	✓		
Motives for alcohol use	✓	✓	✓
Cannabis use (past 6 weeks)	✓	✓	✓
Hazardous cannabis use	✓		
Depressive symptoms	✓	✓	✓
Anxiety symptoms	✓	✓	✓
Emotional eating	✓	✓	✓
Disordered eating behaviors	✓	✓	✓
Food addiction	✓		
Weight bias	✓		

XSL•FO
RenderX

Demographics and Background Information (Baseline Only)

Information collected at baseline will include age, sex and gender, race and ethnicity, height, weight, education level, current medications, lifetime history of psychiatric diagnoses, history of psychiatric treatment, and history of substance use treatment. We will also collect data from electronic health records, including surgery type (RYGB or SG) and preoperative height, weight, and BMI.

Preoperative Substance Use (Baseline Only)

Participants will report the frequency of alcohol use and types of alcohol consumed the year before MBS. They will report the frequency of cannabis use and the forms of cannabis used in the year before MBS. We will also collect data from the preoperative psychosocial evaluation regarding preoperative substance use (eg, type, frequency, and amount of substance use).

Timing of Postoperative Substance Use (Baseline Only)

Participants will report the number of months after MBS when they first consumed alcohol. Those reporting cannabis use after MBS will report the number of months since MBS when they first used cannabis.

Alcohol Use

Participants will respond to questions regarding their alcohol use over the previous 6 weeks. They will report types of alcohol consumed, frequency of alcohol use, typical number of drinks consumed per occasion, number of times they felt drunk, and number of binge drinking episodes (ie, 5 or more drinks for male individuals or 4 or more drinks for female individuals in 1 occasion).

Hazardous Alcohol Use (Baseline Only)

Participants will complete the Alcohol Use Disorders Identification Test [32]. The Alcohol Use Disorders Identification Test consists of 10 items that query the frequency and amount of alcohol use over the previous 6 months as well as the frequency of consequences from alcohol use. A score of 8 or greater suggests the presence of hazardous alcohol use.

Motives for Alcohol Use

The Drinking Motives Questionnaire-Revised [33] assesses respondents' reasons for drinking across 4 categories over the past 6 weeks: conformity, coping, enhancement, and social. Each subscale consists of 5 items with responses ranging from 1 (strongly disagree) to 4 (strongly agree). Items from each subscale are added together for a total subscale score.

Cannabis Use

Although our primary outcome is alcohol use, we will also explore factors associated with cannabis use. Participants will respond to questions regarding their cannabis use over the previous 6 weeks. They will report frequency of cannabis use, forms of cannabis used, methods of ingestion, and average number of hours high or under the influence.

Hazardous Cannabis Use (Baseline Only)

The Cannabis Use Disorders Identification Test measures frequency of cannabis use over the past 6 months as well as consequences from cannabis use. A score of 8 or greater suggests the presence of hazardous cannabis use.

Depressive Symptoms

We will measure current (eg, past 2 weeks) depressive symptoms with the Patient Health Questionnaire-8 depression scale [34]. Participants will report the frequency of depressive symptoms from "not at all" to "every day." Scores range from 0 to 24, with greater scores indicating greater severity. A score of 10 or greater indicates a high likelihood of current depression.

Anxiety Symptoms

We will also assess current (ie, past 2 weeks) anxiety symptoms with the Generalized Anxiety Disorder-7 [35]. Participants will report the frequency of anxiety symptoms from "not at all" to "every day." Scores range from 0 to 21, with greater scores indicating greater severity. A score of 10 or greater indicates a high likelihood of current anxiety.

Emotional Eating

The Emotional Eating Scale will assess the urge to eat in response to 25 emotions that fall into 3 factors (ie, anger or frustration, anxiety, and depression) and can also be combined for a total overall score [36]. Greater scores indicate higher levels of emotional eating. This measure was validated among individuals with obesity [36] and has been previously used with individuals planning to undergo MBS [26,37].

Disordered Eating

Items from the Eating Disorder Examination-Questionnaire will be used to assess the frequency of disordered eating behaviors over the past 6 weeks. Participants will report whether they had a disordered eating behavior in the past 6 weeks (eg, loss of control, large portions, binge eating, grazing, and night eating) and, if so, the number of episodes each behavior occurred.

Food Addiction (Baseline Only)

Participants will complete the Yale Food Addiction Scale 2.0, which is a 35-item measure that assesses symptoms of food addiction that map onto the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* criteria for substance use disorders [38]. This measure has been validated for use with patients planning to undergo MBS [39].

Weight Bias

Weight bias will be assessed to collect preliminary data on postoperative weight bias to support future work. The Modified Weight Bias Internalization Scale is a 10-item measure in which participants respond to the degree to which they agree with statements regarding weight bias [40]. Higher scores indicate greater internalized weight bias.

Measures: Proximal Factors

Proximal factors are measured during the EMA period. Table 2 presents the measures to be used in the daily EMA surveys. Completion of the complete set of items is expected to take less than 3 minutes.

XSL•FO
RenderX

Table 2. Measures for the ecological momentary assessment.

Construct	Morning measure	Evening measure
Current mood		
Positive and Negative Affect Schedule–Short Form	10 items (Likert scale)	10 items (Likert scale)
Disordered eating		
Binge eating	Yes or no[a]	Yes or no[a]
Large portions	Yes or no[a]	Yes or no[a]
Loss of control	Yes or no[a]	Yes or no[a]
Grazing	Yes or no[a]	Yes or no[a]
Poor food choices	Yes or no[a]	Yes or no[a]
Emotional eating	Yes or no[a]	Yes or no[a]
Alcohol use	Intention to use (yes, no, or maybe)	Actual use (yes or no)[b]
Alcohol use	Willingness to use (Likert scale)	—[c]
Cannabis use	Use the previous day (yes or no)[d]	—[c]

[a]If yes, what time did this begin?

[b]If yes, additional questions will query the type used, the amount, the time of use, and perceived intoxication.

[c]These items were not used on the evening survey.

[d]If yes, additional questions will query the form used, the time of use, and perceived high.

Mood

Participants will be asked to rate their current mood using the Positive and Negative Affect Scale–Short Form [41]. This measure has 5 items on each subscale (ie, positive and negative affect), allowing a range of affect to be assessed while remaining brief, which is important in EMA studies.

Disordered Eating

Participants will be asked if they engaged (yes or no) in 6 types of disordered eating behaviors (ie, binge eating, large portions, loss of control, grazing, poor food choices, and emotional eating) [42-44]. In the morning, the participant will report eating behaviors from the previous day. At the evening assessment, the participant will report on eating behaviors that occurred on that day. For each eating behavior endorsed, the participant will report the time at which the eating behavior began.

Alcohol Use

Participants will be presented with an image and description of a standard drink on every survey. At the morning assessment, participants will indicate whether they intend to use alcohol that day (yes, no, or maybe). They will also be asked about their willingness to drink (eg, “If somebody offered you an alcoholic beverage later today, would you drink it?”), to which they will respond using a scale from “definitely not” to “definitely yes.” They will also report the amount of alcohol use on the previous day (morning survey) or on that day (evening survey). If alcohol consumption is reported, they will report the time they started their first drink and the time they finished their last drink (with the option to select that they are still drinking). Participants will also report the types of alcohol consumed and their level of intoxication (“not at all” to “very intoxicated”). These items have been used in previous EMA studies on mood and substance use [45,46].

Cannabis Use

On the morning survey, participants will respond to whether they used cannabis the previous day (yes or no). If yes, they will report on the form of cannabis used, the time of use, and how high they felt (on a scale from “not at all” to “very high”).

Power Analysis

Power estimates assume an α of .05, with 2-tailed tests. For the regression models, G*Power calculations indicate that with 100 participants, power to detect a moderate effect (Cohen f^2=0.15) is 0.94; power to detect a small-to-moderate effect (Cohen f^2=0.10) is 0.80 [47]. For the multilevel modeling, a rule of thumb for estimating models with random effects is approximately 50 higher-level units (participants) with 20 observations per unit (time points) [48]. Our sample size and sampling design exceed this recommendation. For survival analysis, we computed power to detect a difference in hazard for 2 groups (eg, time to first drink as a function of men and women), assuming exponentially distributed survival times and all individuals followed for the same duration [49]. As we cover the full interval during which individuals initiate alcohol use, we can assume a follow-up period of twice the average median lifetime (2.0). For a large effect size (R=2.00, the ratio of the median interval to event in group 1 vs group 2), 88 participants are required to attain power of 0.80; for a moderate effect size (R=1.75), 134 participants are needed. Notably, the primary goal of this project is to establish the feasibility and acceptability of the study methodology; therefore, for all analyses, we will focus on the magnitude and direction of effects.

XSL•FO
RenderX

Ethical Considerations

This study is approved by the Henry Ford Health Institutional Review Board, where recruitment will take place (17061). Institutional review board approval was obtained in March 2024. To protect privacy, participants will be asked whether they consent to share their data in a repository. Participants who do not consent to this will not have their data shared. No identifying information of any participant will be shared or published. Participants will be compensated for their participation. They will receive US $50 for completion of each of the assessments at the baseline, 6-, and 12-week follow-ups. During the EMA, they will receive US $2 for each signal completed (up to US $4 per day) for a total of up to US $84 for the 21 days. Participants who complete at least 80% of the EMA signals will also receive a US $50 bonus.

Results

Overview

This study received funding from the National Institute on Alcohol Abuse and Alcoholism (R21 AA029423) in May 2023. The primary goal of this study is to evaluate the feasibility of the study design. We will calculate the enrollment rate (the number of those who complete the baseline assessment of those eligible) and retention rates at the 6- and 12-week follow-ups. We anticipate that at least 50% of the eligible individuals will enroll and that at least 80% will complete the 6- and 12-week follow-ups. We will also calculate adherence to the morning and evening assessments during the EMA period. We anticipate that 75% of the participants or more will have at least a 75% response rate to the EMA assessments across the full 3-week EMA period.

The second goal of this study is to determine distal and proximal predictors of alcohol use. We plan to measure alcohol use in several ways. Our primary outcome is alcohol use on a given day during the EMA period (yes or no), and the secondary outcomes include the number of drinks and hazardous alcohol use. We will also explore whether there are differences in predictors for planned versus unplanned drinking. Data collection began in November 2024 and will conclude in December 2025.

Distal Predictors

We will conduct multiple regression analyses predicting alcohol outcomes at 6 and 12 weeks. Predictors will include baseline measures of depression, anxiety, emotional eating scores, number of food addiction symptoms, frequency of disordered eating behaviors (binge eating, loss of control eating, large portions, grazing, and unhealthy food choices), and drinking motives; we will control for sex and examine sex differences by testing interactions with sex. We will control for time since drinking reinitiation, given that there will be variability in timing. We hypothesize that negative emotionality will be associated with increased drinking in lieu of disordered eating. We will also produce path models that include data from all 3 time points (baseline and 6 and 12 weeks) in a single model within Mplus (Muthén & Muthén).

We will use continuous time survival analysis [49-52] to investigate initiation of alcohol use following MBS (time 0). We can predict time to first drink from time-invariant variables, such as sex and history of preoperative alcohol use.

Proximal Predictors

To examine proximal predictors of drinking, we will use multilevel modeling [53,54], which accounts for clustering (clustering for days within each person) and permits varying numbers of observations to account for missing reports. Level 1 (L1) effects correspond to within-person effects (eg, proximal behaviors at the level of the day), and level 2 effects indicate between-person effects (eg, distal predictors such as sex or L1 variables such as eating behaviors or depression aggregated over time). Consistent with recommendations, level 2 effects will be grand mean centered, and L1 effects will be person mean centered, such that the L1 effect reflects a deviation from one's own typical behavior [55,56]. Outcomes will include any drinking and amount of drinking as well as our measure of unplanned drinking; predictors will include same-day mood and eating behaviors. Moreover, we can use the survey data to code whether drinking precedes mood, follows mood, or occurs contemporaneously to examine directionality within a day; these categorical variables can be used in subsequent analyses. In addition, daily-level data can be lagged to examine whether there are mood effects on next-day alcohol use and vice versa. We will also examine daily associations among the mood and eating variables. We hypothesize that alcohol use will occur on days characterized by greater negative mood is present and in lieu of disordered eating, given that we expect maladaptive coping strategies (ie, alcohol use or disordered eating in response to negative mood) to be used at different times, and we can code directionality among mood and drinking and eating behaviors for use in analyses.

Discussion

The primary goals of this study are to determine the feasibility of the study design and identify distal and proximal factors associated with episodic alcohol use after MBS. We believe our study design will be feasible and that we will meet our benchmarks for enrollment and retention in the waves as well as adherence to the EMA. Our benchmarks were established based on previous EMA work [57]. These findings will support future use of this EMA design among those who underwent MBS.

Our study will examine distal and proximal factors associated with alcohol use. We hypothesize that depressive symptoms (distal factor) and negative affect (proximal factor) will be associated with higher levels of drinking. As many as 50% of the individuals planning to undergo MBS report a history of depression [58-60]. Our team has found that although depressive symptoms improve in the first year after MBS, depressive symptoms return 1 to 2 years following MBS [27]. This is the same time frame in which we found that postoperative alcohol use is commonly initiated [31]. We will also explore whether alcohol consumption occurs in lieu of disordered eating behaviors. Previous qualitative work has suggested that patients who underwent MBS do not consume alcohol and food together

XSL•FO
RenderX

[61]. This study will extend this through a more thorough quantitative investigation. We will also determine whether patients consume alcohol in response to depression and negative affect in lieu of disordered eating behaviors. In addition, we plan to examine whether those who consume alcohol in lieu of disordered eating are more likely to engage in hazardous drinking. If affirmative, this would provide support for the addiction transfer hypothesis [25].

There are several strengths of this study, including that it will be the first to provide a granular examination of episodic alcohol use among individuals who underwent MBS. We are capturing real-world data at the time events are occurring, which improves the rigor of previous work that has relied on retrospective recall. The EMA design also allows for the temporal examination of mood, eating, and alcohol use, which can test the addiction transfer hypothesis in a way that has not previously been done. In addition, it is novel to look at intention to drink (ie, planned vs unplanned drinking and willingness to drink). However, it is important to note several limitations. First, because we are capturing real-world behavior, it is possible that there may be a low number of drinking episodes that occur during the EMA period. To prevent this, our recruitment will be targeted to those who endorse more frequent drinking behaviors (eg, at least 2-3 times per month). In addition, affect can fluctuate throughout the day. Because of the study design, we may not capture affect immediately preceding drinking. However, this work will support future investigation of these variables in a study with greater resources, allowing a more rigorous design among a larger sample.

This line of research will help us understand distal and proximal factors leading to episodic alcohol use after MBS. The results from this project will inform a future proposal with longer-term follow-up that will examine the progression from reinitiation of alcohol use to heavy use to the development of an AUD. This knowledge will allow us to construct better monitoring strategies for postoperative alcohol use within MBS programs. Findings will also assist in the development of interventions to reduce negative affect and alcohol use after MBS. By reducing postoperative alcohol use, we can prevent AUDs in this vulnerable, high-risk population.

Funding

This project was funded by the National Institute on Alcohol Abuse and Alcoholism (R21 AA029423).

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from AA-2 - Epidemiology, Prevention and Behavior Research Study Section, National Institute on Alcohol Abuse and Alcoholism Initial Review Group (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 150 KB - resprot_v15i1e87209_app1.pdf]

References

1. Arterburn DE, Telem DA, Kushner RF, Courcoulas AP. Benefits and risks of bariatric surgery in adults: a review. JAMA 2020 Sep 01;324(9):879-887. [doi: 10.1001/jama.2020.12567] [Medline: 32870301]
2. Sarma S, Palcu P. Weight loss between glucagon-like peptide-1 receptor agonists and bariatric surgery in adults with obesity: a systematic review and meta-analysis. Obesity (Silver Spring) 2022 Nov;30(11):2111-2121. [doi: 10.1002/oby.23563] [Medline: 36321278]
3. King WC, Chen JY, Courcoulas AP, Dakin GF, Engel SG, Flum DR, et al. Alcohol and other substance use after bariatric surgery: prospective evidence from a U.S. multicenter cohort study. Surg Obes Relat Dis 2017 Aug;13(8):1392-1402 [FREE Full text] [doi: 10.1016/j.soard.2017.03.021] [Medline: 28528115]
4. Miller-Matero LR, Orlovskaia J, Hecht LM, Braciszewski JM, Martens KM, Hamann A, et al. Hazardous alcohol use in the four years following bariatric surgery. Psychol Health Med 2022 Oct 06;27(9):1884-1890. [doi: 10.1080/13548506.2021.1930075] [Medline: 34096405]
5. Svensson PA, Anveden Å, Romeo S, Peltonen M, Ahlin S, Burza MA, et al. Alcohol consumption and alcohol problems after bariatric surgery in the Swedish obese subjects study. Obesity (Silver Spring) 2013 Dec 31;21(12):2444-2451 [FREE Full text] [doi: 10.1002/oby.20397] [Medline: 23520203]
6. Acevedo MB, Eagon JC, Bartholow BD, Klein S, Bucholz KK, Pepino MY. Sleeve gastrectomy surgery: when 2 alcoholic drinks are converted to 4. Surg Obes Relat Dis 2018 Mar;14(3):277-283 [FREE Full text] [doi: 10.1016/j.soard.2017.11.010] [Medline: 29305304]

XSL•FO
RenderX

7. Pepino MY, Okunade AL, Eagon JC, Bartholow BD, Bucholz K, Klein S. Effect of Roux-en-Y gastric bypass surgery: converting 2 alcoholic drinks to 4. JAMA Surg 2015 Nov;150(11):1096-1098 [FREE Full text] [doi: 10.1001/jamasurg.2015.1884] [Medline: 26244751]
8. Steffen KJ, Engel SG, Pollert GA, Li C, Mitchell JE. Blood alcohol concentrations rise rapidly and dramatically after Roux-en-Y gastric bypass. Surg Obes Relat Dis 2013 May;9(3):470-473 [FREE Full text] [doi: 10.1016/j.soard.2013.02.002] [Medline: 23507629]
9. Smith KE, Engel SG, Steffen KJ, Garcia L, Grothe K, Koball A, et al. Problematic alcohol use and associated characteristics following bariatric surgery. Obes Surg 2018 May 6;28(5):1248-1254 [FREE Full text] [doi: 10.1007/s11695-017-3008-8] [Medline: 29110243]
10. Buffington CK. Alcohol use and health risks: survey results. Bariatric Times 2007;4(2):1-21 [FREE Full text]
11. Miller-Matero LR, Hecht LM, Barnett NP, Moore RS, Jackson KM, McCarthy D, et al. Patient perceptions regarding alcohol use after bariatric surgery. Surg Endosc 2023 May 13;37(5):3669-3675 [FREE Full text] [doi: 10.1007/s00464-023-09868-0] [Medline: 36639579]
12. Polston JE, Pritchett CE, Tomasko JM, Rogers AM, Leggio L, Thanos PK, et al. Roux-en-Y gastric bypass increases intravenous ethanol self-administration in dietary obese rats. PLoS One 2013 Dec 31;8(12):e83741 [FREE Full text] [doi: 10.1371/journal.pone.0083741] [Medline: 24391816]
13. Hajnal A, Zharikov A, Polston JE, Fields MR, Tomasko J, Rogers AM, et al. Alcohol reward is increased after Roux-en-Y gastric bypass in dietary obese rats with differential effects following ghrelin antagonism. PLoS One 2012 Nov 7;7(11):e49121 [FREE Full text] [doi: 10.1371/journal.pone.0049121] [Medline: 23145091]
14. Yoder R, MacNeela P, Conway R, Heary C. How do individuals develop alcohol use disorder after bariatric surgery? A grounded theory exploration. Obes Surg 2018 Mar 14;28(3):717-724. [doi: 10.1007/s11695-017-2936-7] [Medline: 29032488]
15. Steffen KJ, Engel SG, Wonderlich JA, Pollert GA, Sondag C. Alcohol and other addictive disorders following bariatric surgery: prevalence, risk factors and possible etiologies. Eur Eat Disord Rev 2015 Nov 08;23(6):442-450. [doi: 10.1002/erv.2399] [Medline: 26449524]
16. Blumenthal DM, Gold MS. Neurobiology of food addiction. Curr Opin Clin Nutr Metab Care 2010;13(4):359-365. [doi: 10.1097/mco.0b013e32833ad4d4]
17. Gearhardt AN, Yokum S, Orr PT, Stice E, Corbin WR, Brownell KD. Neural correlates of food addiction. Arch Gen Psychiatry 2011 Aug 01;68(8):808-816 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2011.32] [Medline: 21464344]
18. Corsica JA, Pelchat ML. Food addiction: true or false? Curr Opin Gastroenterol 2010 Mar;26(2):165-169. [doi: 10.1097/MOG.0b013e328336528d] [Medline: 20042860]
19. Pelchat ML. Food addiction in humans. J Nutr 2009 Mar;139(3):620-622 [FREE Full text] [doi: 10.3945/jn.108.097816] [Medline: 19176747]
20. Schulte EM, Yokum S, Jahn A, Gearhardt AN. Food cue reactivity in food addiction: a functional magnetic resonance imaging study. Physiol Behav 2019 Sep 01;208:112574 [FREE Full text] [doi: 10.1016/j.physbeh.2019.112574] [Medline: 31181233]
21. Hsu JS, Wang PW, Ko CH, Hsieh TJ, Chen CY, Yen JY. Altered brain correlates of response inhibition and error processing in females with obesity and sweet food addiction: a functional magnetic imaging study. Obes Res Clin Pract 2017 Nov;11(6):677-686. [doi: 10.1016/j.orcp.2017.04.011] [Medline: 28552670]
22. Domingo-Rodriguez L, Ruiz de Azua I, Dominguez E, Senabre E, Serra I, Kummer S, et al. A specific prelimbic-nucleus accumbens pathway controls resilience versus vulnerability to food addiction. Nat Commun 2020 Feb 07;11(1):782 [FREE Full text] [doi: 10.1038/s41467-020-14458-y] [Medline: 32034128]
23. Pursey KM, Contreras-Rodriguez O, Collins CE, Stanwell P, Burrows TL. Food addiction symptoms and amygdala response in fasted and fed states. Nutrients 2019 Jun 06;11(6):1285 [FREE Full text] [doi: 10.3390/nu11061285] [Medline: 31174338]
24. Fowler L, Ivezaj V, Saules KK. Problematic intake of high-sugar/low-fat and high glycemic index foods by bariatric patients is associated with development of post-surgical new onset substance use disorders. Eat Behav 2014 Aug;15(3):505-508. [doi: 10.1016/j.eatbeh.2014.06.009] [Medline: 25064307]
25. Ivezaj V, Benoit SC, Davis J, Engel S, Lloret-Linares C, Mitchell JE, et al. Changes in alcohol use after metabolic and bariatric surgery: predictors and mechanisms. Curr Psychiatry Rep 2019 Aug 13;21(9):85 [FREE Full text] [doi: 10.1007/s11920-019-1070-8] [Medline: 31410716]
26. Miller-Matero LR, Hamann A, LaLonde L, Martens KM, Son J, Clark-Sienkiewicz S, et al. Predictors of alcohol use after bariatric surgery. J Clin Psychol Med Settings 2021 Sep 17;28(3):596-602. [doi: 10.1007/s10880-020-09751-3] [Medline: 33205321]
27. Martens K, Hamann A, Miller-Matero LR, Miller C, Bonham AJ, Ghaferi AA, et al. Relationship between depression, weight, and patient satisfaction 2 years after bariatric surgery. Surg Obes Relat Dis 2021 Feb;17(2):366-371. [doi: 10.1016/j.soard.2020.09.024] [Medline: 33127323]
28. Smith KE, Orcutt M, Steffen KJ, Crosby RD, Cao L, Garcia L, et al. Loss of control eating and binge eating in the 7 years following bariatric surgery. Obes Surg 2019 Jun 28;29(6):1773-1780 [FREE Full text] [doi: 10.1007/s11695-019-03791-x] [Medline: 30820886]

XSL•FO
RenderX

29. Shiffman S. Ecological momentary assessment (EMA) in studies of substance use. Psychol Assess 2009 Dec;21(4):486-497 [FREE Full text] [doi: 10.1037/a0017074] [Medline: 19947783]
30. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
31. Miller-Matero LR, Coleman JP, LaLonde L, Martens KM, Hamann A, Carlin AM. Patient recall of education about the risks of alcohol use following bariatric surgery. Obes Surg 2019 Aug 16;29(8):2707-2710. [doi: 10.1007/s11695-019-03940-2] [Medline: 31098786]
32. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun 24;88(6):791-804. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
33. Cooper ML. Motivations for alcohol use among adolescents: development and validation of a four-factor model. Psychol Assess 1994;6(2):117-128. [doi: 10.1037//1040-3590.6.2.117]
34. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
35. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
36. Arnow B, Kenardy J, Agras WS. The Emotional Eating Scale: the development of a measure to assess coping with negative affect by eating. Int J Eat Disord 1995 Jul;18(1):79-90. [doi: 10.1002/1098-108x(199507)18:1<79::aid-eat2260180109>3.0.co;2-v] [Medline: 7670446]
37. Miller-Matero LR, Bryce K, Saulino CK, Dykhuis KE, Genaw J, Carlin AM. Problematic eating behaviors predict outcomes after bariatric surgery. Obes Surg 2018 Jul 7;28(7):1910-1915. [doi: 10.1007/s11695-018-3124-0] [Medline: 29417489]
38. Gearhardt AN, Corbin WR, Brownell KD. Development of the Yale Food Addiction Scale version 2.0. Psychol Addict Behav 2016 Feb;30(1):113-121. [doi: 10.1037/adb0000136] [Medline: 26866783]
39. Clark SM, Martens K, Smith-Mason CE, Hamann A, Miller-Matero LR. Validation of the Yale Food Addiction Scale 2.0 among a bariatric surgery population. Obes Surg 2019 Sep 22;29(9):2923-2928. [doi: 10.1007/s11695-019-03927-z] [Medline: 31119701]
40. Pearl RL, Puhl RM. Measuring internalized weight attitudes across body weight categories: validation of the modified weight bias internalization scale. Body Image 2014 Jan;11(1):89-92. [doi: 10.1016/j.bodyim.2013.09.005] [Medline: 24100004]
41. Thompson ER. Development and validation of an internationally reliable short-form of the positive and negative affect schedule (PANAS). J Cross Cult Psychol 2007 Mar 01;38(2):227-242. [doi: 10.1177/0022022106297301]
42. Mitchell JE, Christian NJ, Flum DR, Pomp A, Pories WJ, Wolfe BM, et al. Postoperative behavioral variables and weight change 3 years after bariatric surgery. JAMA Surg 2016 Aug 01;151(8):752-757 [FREE Full text] [doi: 10.1001/jamasurg.2016.0395] [Medline: 27096225]
43. Mitchell JE, King WC, Courcoulas A, Dakin G, Elder K, Engel S, et al. Eating behavior and eating disorders in adults before bariatric surgery. Int J Eat Disord 2015 Mar 09;48(2):215-222 [FREE Full text] [doi: 10.1002/eat.22275] [Medline: 24719222]
44. Goldschmidt AB, Crosby RD, Cao L, Engel SG, Durkin N, Beach H, et al. Ecological momentary assessment of eating episodes in obese adults. Psychosom Med 2014;76(9):747-752 [FREE Full text] [doi: 10.1097/PSY.0000000000000108] [Medline: 25373891]
45. Fatseas M, Serre F, Swendsen J, Auriacombe M. Effects of anxiety and mood disorders on craving and substance use among patients with substance use disorder: an ecological momentary assessment study. Drug Alcohol Depend 2018 Jun 01;187:242-248. [doi: 10.1016/j.drugalcdep.2018.03.008] [Medline: 29684892]
46. Serre F, Fatseas M, Debrabant R, Alexandre J, Auriacombe M, Swendsen J. Ecological momentary assessment in alcohol, tobacco, cannabis and opiate dependence: a comparison of feasibility and validity. Drug Alcohol Depend 2012 Nov 01;126(1-2):118-123. [doi: 10.1016/j.drugalcdep.2012.04.025] [Medline: 22647899]
47. Faul F, Erdfelder E, Buchner A, Lang A. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/brm.41.4.1149]
48. Hox JJ, Moerbeek M, Van de Schoot R. Multilevel Analysis: Techniques and Applications. 2nd edition. New York, NY: Routledge; 2010.
49. Singer JD, Willett JB. Modeling the days of our lives: using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. Psychol Bull 1991;110(2):268-290. [doi: 10.1037//0033-2909.110.2.268]
50. Singer JD, Willett JB. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford, UK: Oxford University Press; 2003.
51. Singer JD, Willett JB. Designing and analyzing studies of onset, cessation, and relapse: using survival analysis in drug abuse prevention research. NIDA Res Monogr 1994;142:196-263. [Medline: 9243538]
52. Jackson KM, Janssen T. Developmental considerations in survival models as applied to substance use research. Addict Behav 2019 Jul;94:36-41 [FREE Full text] [doi: 10.1016/j.addbeh.2018.11.028] [Medline: 30538054]

XSL•FO
RenderX

53. Hayes AF. A primer on multilevel modeling. Human Comm Res 2006 Oct;32(4):385-410. [doi: 10.1111/j.1468-2958.2006.00281.x]
54. Raudenbush SW, Bryk AS. Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: Sage Publications; 1999.
55. Curran PJ, Bauer DJ. The disaggregation of within-person and between-person effects in longitudinal models of change. Annu Rev Psychol 2011 Jan 10;62(1):583-619 [FREE Full text] [doi: 10.1146/annurev.psych.093008.100356] [Medline: 19575624]
56. Kreft IG, de Leeuw J, Aiken LS. The effect of different forms of centering in hierarchical linear models. Multivariate Behav Res 1995 Jan 01;30(1):1-21 [FREE Full text] [doi: 10.1207/s15327906mbr3001_1] [Medline: 26828342]
57. Jones A, Remmerswaal D, Verveer I, Robinson E, Franken IH, Wen CK, et al. Compliance with ecological momentary assessment protocols in substance users: a meta-analysis. Addiction 2019 Apr 21;114(4):609-619 [FREE Full text] [doi: 10.1111/add.14503] [Medline: 30461120]
58. Miller-Matero LR, Tobin ET, Clark S, Eshelman A, Genaw J. Pursuing bariatric surgery in an urban area: gender and racial disparities and risk for psychiatric symptoms. Obes Res Clin Pract 2016 Jan;10(1):56-62. [doi: 10.1016/j.orcp.2015.03.006] [Medline: 25858423]
59. Sarwer DB, Cohn NI, Gibbons LM, Magee L, Crerand CE, Raper SE, et al. Psychiatric diagnoses and psychiatric treatment among bariatric surgery candidates. Obes Surg 2004 Oct 01;14(9):1148-1156. [doi: 10.1381/0960892042386922] [Medline: 15527626]
60. Kalarchian MA, Marcus MD, Levine MD, Courcoulas AP, Pilkonis PA, Ringham RM, et al. Psychiatric disorders among bariatric surgery candidates: relationship to obesity and functional health status. Am J Psychiatry 2007 Feb;164(2):328-374. [doi: 10.1176/ajp.2007.164.2.328] [Medline: 17267797]
61. Miller-Matero LR, Vanderziel A, Haley EN, Jackson KM, Moore RS, Hamann A, et al. Alcohol use after metabolic and bariatric surgery: a qualitative investigation of the relation with mood and food. Health Psychol Behav Med 2025 Mar 13;13(1):2478029 [FREE Full text] [doi: 10.1080/21642850.2025.2478029] [Medline: 40098646]

Abbreviations

AUD: alcohol use disorder
EMA: ecological momentary assessment
HIPAA: Health Insurance Portability and Accountability Act
L1: level 1
MBS: metabolic and bariatric surgery
REDCap: Research Electronic Data Capture
RYGB: Roux-en-Y gastric bypass
SG: sleeve gastrectomy

Edited by J Sarvestan; submitted 06.Nov.2025; this is a non–peer-reviewed article;accepted 11.Dec.2025; published 14.Jan.2026.

Please cite as:
Miller-Matero LR, Saleh D, Christopher B, Albujuq M, Vanderziel A, Haley EN, Braciszewski JM, Moore RS, Carlin AM, Jackson KM
Factors Associated With Alcohol Use After Metabolic and Bariatric Surgery: Protocol for an Ecological Momentary Assessment
JMIR Res Protoc 2026;15:e87209
URL: https://www.researchprotocols.org/2026/1/e87209
doi:10.2196/87209
PMID:

©Lisa R Miller-Matero, Daniel Saleh, Brittany Christopher, Maha Albujuq, Alyssa Vanderziel, Erin N Haley, Jordan M Braciszewski, Roland S Moore, Arthur M Carlin, Kristina M Jackson. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 14.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Central Nervous System Mechanisms and Treatment Response in Chronic Ocular Surface Pain: Protocol for a Cross-Sectional Observational Phenotyping Study

Lindsey B De Lott[1], MS, MD; Steven E Harte[2], PhD; Chelsea Kaplan[2], PhD; David A Williams[2], PhD; Roni Shtein[1], MD; Maria A Woodward[1], MD, MSc; Alexander Tsodikov[3], PhD; Tatiana Deveney[1], MD; Anat Galor[4,5], MD, MPH; Anne Shea[6], MD; Charles Schultz[6], MD; Clare McKolay[1], COT; Kathy A Scott[2], BSN, RN; Daniel J Clauw[2], MD

[1]Department of Ophthalmology and Visual Sciences, University of Michigan, Ann Arbor, MI, United States

[2]Chronic Pain and Fatigue Research Center, Department of Anesthesiology, University of Michigan, Ann Arbor, MI, United States

[3]Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, United States

[4]Bascom Palmer Eye Institute, University of Miami, Miami, FL, United States

[5]Surgical services, Miami Veterans Administration Medical Center, Miami, FL, United States

[6]Department of Neurology, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Lindsey B De Lott, MS, MD
Department of Ophthalmology and Visual Sciences
University of Michigan
1000 Wall Street
Ann Arbor, MI, 48104
United States
Phone: 1 734 232 8241
Fax: 1 734 232 2332
Email: ldelott@med.umich.edu

Abstract

Background: Chronic ocular surface pain (COSP), occurring either in isolation or as part of numerous ocular conditions, such as dry eye syndrome, is a leading cause of eye care visits in the United States. Conventional treatments directed at the ocular surface—the perceived pain source—are often inadequate for pain relief. We hypothesize that some individuals with COSP are experiencing symptoms driven by central nervous system (CNS) dysfunction, similar to chronic overlapping pain conditions, rather than solely pathological problems in the eye. Some individuals with chronic overlapping pain conditions (eg, fibromyalgia) show evidence of nociplastic pain mechanisms, where the pain results from amplified or dysregulated CNS signaling and sensory processing. Although data exist suggesting the presence of nociplastic pain features in COSP, there is a need for comprehensive studies.

Objective: Our aim is to rigorously define the role of nociplastic pain in COSP with a large, representative cohort of 200 participants with COSP using established clinical, neurobiological, and treatment response features. We propose that as sign and symptom discordance increases, features indicative of nociplastic pain will also increase.

Methods: In aim 1, we will clinically phenotype participants with COSP, using validated patient-reported outcome measures and standard ocular exams. In aim 2, we will compare the neurobiological features of nociplastic pain across the discordance spectrum among a subset of aim 1 participants using multimodal evoked sensory testing (pressure, thermal, and visual testing) at sites both local to and remote from the eye. Participants will also complete structural and functional brain magnetic resonance imaging to assess regions important for pain perception and modulation. In aim 3, we will examine and validate predictors of COSP pain responses before and after application of a topical anesthetic to the ocular surface, which should block peripherally induced discomfort to allow for clarification of pain origination.

Results: We received funding for this project in August 2024. Recruitment and enrollment began in January 2025 after protocol development and piloting were completed. This study is ongoing with 59 participants enrolled and 51 participants completing the study visits as of January 2026.

Conclusions: Findings from the study have the potential to fundamentally change the way ocular pain syndromes are conceptualized, diagnosed, and treated. This work will not only help identify new CNS-directed pain treatments for a subset of

XSL•FO
RenderX

patients with COSP but also help us better understand why many peripherally directed therapies are destined to be ineffective in a subset of individuals experiencing COSP.

International Registered Report Identifier (IRRID): DERR1-10.2196/84240

(JMIR Res Protoc 2026;15:e84240) doi:10.2196/84240

KEYWORDS

ocular pain; chronic pain; nociplastic; dry eye; protocol; observational study

Introduction

Chronic ocular surface pain (COSP) encompasses a number of interrelated ocular pain symptoms, such as burning pain or eye irritation. Although COSP can occur as a symptom of several ocular conditions, including dry eye disease, it is well recognized that it can also occur in isolation. COSP is a leading cause of eye care visits in the United States, affecting more than 5% of people over their lifetime [1-3]. Because the symptoms are relentless and can be lifelong, patients experience decreased quality of life [4,5], similar to those with an immobilizing hip fracture or moderate to severe angina [6], decreased visual functioning [7,8], and higher odds of suicidal ideation [9].

A complex interaction of loss of tear film homeostasis and ocular surface inflammation (leading to pain [ie, nociceptive pain]) or pain arising from damaged corneal nerves (ie, neuropathic pain) drives pain symptoms in many individuals with COSP [10]. However, treatments aimed at restoring tear film homeostasis, reducing ocular inflammation, or even promoting tissue and nerve repair often do not relieve the pain for many people with COSP. Such is the case with the 10 currently recognized chronic overlapping pain conditions (COPCs)—a set of chronic pain conditions, including fibromyalgia, irritable bowel syndrome, temporomandibular disorders, chronic migraine, chronic tension type headache, chronic low back pain, chronic fatigue syndrome, vulvodynia, endometriosis, and interstitial cystitis or bladder pain syndrome [11]. In regional COPCs, such as temporomandibular disorder, interstitial cystitis, and bladder pain syndrome, a subset of individuals experiences peripherally mediated pain, as we know occurs in COSP. However, a prominent subset experiences centrally mediated nociplastic pain, where amplified or dysregulated neural signaling and sensory processing within the central nervous system (CNS) is now thought to be the primary or maintaining pathogenic pain mechanism [12-15]. In fact, these same nociplastic pain conditions are now referred to as “primary pain” conditions because the pain is the problem and not an underlying disease-causing damage or inflammation leading to pain. In nociplastic pain, therapeutic strategies that target the CNS are key to delivering pain relief.

We hypothesize that a subset of individuals with COSP experience symptoms driven primarily by CNS dysfunction rather than pathological problems exclusively with the eye. This premise is supported by observed nociplastic pain features among subsets of individuals with COSP [16], including (1) pain with minimal ocular surface signs (ie, sign and symptom discordance) [8,17], (2) multisite pain and CNS-mediated somatic symptoms (eg, fatigue) [18-20], (3) augmented CNS pain processing on evoked sensory testing and neuroimaging [21-23], and (4) persistent pain after topical anesthetic application [21]. However, no comprehensive studies have investigated clinical, neurobiological, and treatment response features of nociplastic pain in COSP. To address our hypothesis, we will conduct a cross-sectional, observational study to rigorously define the role of nociplastic pain in COSP (inclusive of dry eye disease but without structural, mechanical, and infectious causes), with a large, representative cohort of participants using established clinical phenotypic, neurobiological, and treatment response features.

Methods

Overview and Study Aims

This prospective, observational cohort study will recruit 200 individuals with COSP, with or without ocular surface signs (eg, dry eye syndrome but no other structural, mechanical, or infectious causes) from the University of Michigan (U-M). Our study design and selection of clinical and neurobiological measures are adapted from consensus methods developed as part of the National Institutes of Health (NIH) Helping to End Addiction Long-term (HEAL) initiative and several pain research networks [24-27].

The first aim of our project is to perform a clinical pain phenotyping study of participants with COSP along the sign and symptom discordance continuum. Participants will complete an ocular surface examination and standardized patient-reported outcome measures, including pain intensity assessment before and following application of topical anesthetic (proparacaine) to the ocular surface. On the basis of the ocular surface symptoms and signs, participants will be placed on a sign and symptom discordance continuum as described by Vehof et al [28]. This method has been used in observational studies and clinical trials [8,29].

The second aim of our project is to determine the neurobiological features of participants with COSP across the sign and symptom discordance continuum. A subset of 80 participants with COSP (discordant: n=40, 50%; concordant: n=40, 50%) and 40 controls (healthy controls: n=20, 50%; “positive” controls with a diagnosis of fibromyalgia: n=20, 50%) will also undergo multimodal quantitative sensory testing (QST), such as pressure, temperature, and visual modalities, and structural and functional magnetic resonance imaging (fMRI). We will also capture corneal subbasal nerve fiber density using in vivo corneal microscopy for exploratory analyses. Fibromyalgia was selected as a “positive” control as it is the prototypical nociplastic pain condition [12].

XSL•FO
RenderX

The final aim of our project is to integrate clinical and neurobiological data to determine which features predict response to a peripherally acting topical anesthetic. Our model of clinical predictors will be validated in 50 newly recruited participants with COSP. The clinical, neurobiological, and treatment response data will be integrated into our analyses to provide a robust, contextualized, and actionable interpretation of our results.

The SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist is provided in Multimedia Appendix 1.

Study Design and Site

Our study adopts an observational cross-sectional cohort design. This study will occur at a single center (U-M), leveraging the resources at the U-M Kellogg Eye Center (Department of Ophthalmology and Visual Sciences) and U-M Chronic Pain and Fatigue Research Center (CPFRC), Department of Anesthesiology.

Ethical Considerations

This study has been approved by institutional review boards of the University of Michigan Medical School (HUM00256221). Participants will receive incentive payments upon completion of the study visits, including US $50 for the clinical assessment and US $150 for the neurobiological assessment. Informed consent is obtained either in person or over the phone after the details of the study and possible risks of the study are explained. All study procedures are in accordance with the ethical standards of the institutional review boards and with the Declaration of Helsinki. Data will be stored as discussed in the section on Data Management and Sharing.

Study Participants

Individuals with COSP aged between 18 and 70 years will be invited to participate. For the first 2 aims of our study, recruitment of participants with COSP will end when 200 participants are enrolled. We will also recruit an additional 40 control participants, 20 (50%) healthy controls and 20 (50%) “positive” controls with a diagnosis of fibromyalgia. For the third aim of our study, we will recruit an additional 50 participants with COSP following initial analyses to validate our clinical phenotyping results.

Our study will use 2 sets of inclusion and exclusion criteria. The first set of criteria is for enrollment into the clinical phenotyping project for all participants (Textbox 1). The second set of criteria is necessary for the safe and valid conduct of the neurobiological testing protocol (subset of participants with COSP: n=80; control participants: n=40; Textbox 2).

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria for clinical phenotyping assessment.

Inclusion criteria

- Provide a signed and dated informed consent form
- Willing to comply with all study procedures and be available for the duration of the study
- Adults aged between 18 and 70 years
- Ocular surface disease index score of 13 or greater
- Patients reporting ocular surface pain lasting 3 months or more, with pain present for more than half of those days
- Ability to read and speak English to allow for written informed consent and patient-reported outcomes measures

Exclusion criteria

- Adults older than 70 years (excluded due to potential vascular disease leading to abnormal functional imaging findings in older patients)
- Intraocular pressure lower than 5 mm Hg and higher than 22 mm Hg in each eye
- Current diagnosis of any eye infection or intraocular inflammation (uveitis)
- History of ocular herpetic keratitis
- Ocular surgery within 6 months (eg, cataract surgery)
- Previous corneal surgery (eg, laser-assisted in situ keratomileusis)
- Eyelid abnormalities that affect lid function (eg, ectropion, entropion, and lagophthalmos)
- Extensive ocular surface scarring or a condition that may compromise ocular surface integrity (eg, Stevens-Johnson syndrome)
- Current opioid use
- Diagnosed cognitive impairment, serious mental illness (eg, schizophrenia; anxiety and depression are not exclusions), or other condition that may impair decision-making capacity
- Pregnancy
- Persons in prisons
- Anything that would place the individual at increased risk or preclude the individual's full compliance with or completion of the study
- Individuals receiving disability benefits or compensation within the past year or involved in litigation
- Any other diseases or conditions that would make a patient unsuitable for study participation, as determined by the principal investigator (including, but not limited to, severe psychiatric disorders, active suicidal ideations or history of suicide attempts, uncontrolled drug or alcohol addiction, autoimmune conditions on active biologics and immunosuppressive therapies, and visual or hearing difficulties precluding participation).

XSL•FO
RenderX

Textbox 2. Additional inclusion and exclusion criteria for the subset of participants undergoing neurobiological testing.

Additional inclusion criteria

- Normal visual acuity or correctable (with corrective lenses, glasses, or contacts) to at least 20/40 for reading instructions in the magnetic resonance imaging (MRI) and visual sensitivity testing
- No contraindications to MRI (eg, metal implants)
- Willingness to refrain from pain medications, such as nonsteroidal anti-inflammatory drugs and acetaminophen, for 12 hours before neuroimaging and quantitative sensory testing
- Willingness to refrain from alcohol and nicotine on the day of quantitative sensory testing and neuroimaging (alcohol and nicotine consumption allowed after testing is completed)
- Willingness to refrain from physical activity or exercise that would cause muscle or joint soreness for 48 hours before testing (routine exercise or activity not leading to soreness is acceptable)
- Able to lie still on their back for 1 hour during MRI
- Completed all clinical testing, including a complete ocular exam and patient-reported outcome measures

Additional exclusion criteria

- Severe claustrophobia precluding MRI and evoked pain testing during scanning
- Unable to lie comfortably during MRI
- Current, recent (within the last 6 months), or habitual use of artificial nails or nail enhancements (artificial nails can influence pressure pain sensitivity at the thumbnail)
- Diagnosed or self-reported epilepsy because the visual testing strategy uses flashing lights
- Diagnosed or self-reported visual-evoked migraines because the visual testing strategy uses flashing lights
- Tattoos above the shoulders (interfere with MRI)

Recruitment and Screening

The U-M Kellogg Eye Center clinics will be the primary site for participant identification and recruitment. Our U-M online active research studies website [30], established patient research registry (HUM00254693), and Michigan Medicine community advertisements will also be used for recruitment. These multiple strategies allow for recruitment beyond Kellogg Eye Center clinics and Michigan Medicine to reach a broader and more representative population of individuals with COSP. For identifying control participants, individuals participating in ongoing CPFRC studies who provided consent for future correspondence about additional research opportunities will be contacted by letter with study details and an opt-out form if they do not wish to be further contacted. Contact with all interested participants will be initiated by phone or in the clinic by a dedicated research coordinator. Potential participants will be screened using study forms. Study visits will be scheduled. Both the clinical and neurobiological assessments will take place in person and on separate days.

Our priority is to facilitate and encourage participation. When possible, we will schedule appointments on the same day as standard care appointments. Research appointments are scheduled through our electronic medical record system, and participants will receive both email and phone reminders. Questionnaires can be completed 24 hours before the appointment to decrease overall appointment times.

Clinical Assessment

Participants will first complete a battery of questionnaires via the web-based Qualtrics electronic data capture system. These can be completed up to 24 hours before their scheduled visit. A structured medical and ophthalmic history will be obtained and an ophthalmologic examination will be conducted. Participants who provided consent but were later found to be ineligible will be considered screen failures and withdrawn.

Validated questionnaires will assess eye pain, pain in other parts of the body, nonpain CNS-mediated symptoms (eg, mood, cognitive functioning, and sleep), other measures of psychological functioning, and sociodemographic information (Table 1). Questionnaires were selected based on recommendations from the NIH HEAL Initiative Common Data Elements for adult chronic pain. A structured medical, ocular, and surgical history will be obtained using case report forms. Data on the duration of COSP, previous diagnoses potentially underlying COSP, systemic and ocular interventions (medications and procedures) used to treat COSP, and response to these interventions will be collected. Information about current medications used and family members with chronic pain conditions will also be collected.

XSL•FO
RenderX

Table 1. Key questionnaires for all participants.

Category and concepts	Measures
Pain and sensory symptoms	
Multisite pain	• Michigan Body Map (widespread pain index)
Symptom severity	• Symptom Severity Score from the American College of Rheumatology 2016 fibromyalgia criteria
Neuropathic pain	• painDETECT • Neuropathic Pain Symptom Inventory-Eye
Ocular surface symptoms	• Ocular Surface Disease Index
Other chronic pain conditions with a nociplastic basis	• Chronic overlapping pain condition screener
Chronic sensory symptoms	• General Sensory Sensitivity Questionnaire-8
Photophobia	• Visual Light Sensitivity Questionnaire-8
Pain duration and frequency	• Two items adapted for the eye from the National Institutes of Health Research Task Force Minimum Dataset
Pain self-efficacy	• Pain Self-Efficacy Questionnaire
Pain acceptance	• Chronic Pain Acceptance Questionnaire
Pain catastrophizing	• Pain Catastrophizing Scale
Pain interference, enjoyment, and general activity	• Pain, Enjoyment of Life, and General Activity Scale-3
Pain interference-specific activity	• PROMIS SF[a] version 1.0–Pain Interference 4a
Pain intensity	• PROMIS SF version 2.0 0–Pain Intensity 3a
Pain severity	• Pain Numeric Rating Scale
Nonpain central nervous system–mediated symptoms	
Physical functioning or QoL[b]	• PROMIS SF version 2.0–Physical Function 4a
Sleep	• PROMIS SF version 1.0–Sleep Disturbance 4a
Depression	• PROMIS SF version 1.0–Depression 4a
Anxiety	• PROMIS SF version 1.0–Anxiety 4a
Fatigue	• PROMIS SF version 1.0–Fatigue 4a
Social roles and activities	• PROMIS SF version 2.0–Ability to Participate 4a
Memory and cognition	• PROMIS SF version 2.0–Cognitive Function 4a
Sleep duration	• Single item of sleep duration adapted from the Pittsburgh Sleep Quality Index
Dysautonomia	• Composite Autonomic Symptom Score-31
QoL	
Visual QoL	• National Eye Institute Visual Functioning Questionnaire-9 and 2 pain subscale items
Overall QoL	• World Health Organization Quality of Life-brief version

XSL•FO
RenderX

Category and concepts	Measures
Social and psychological	
Substance use	• Tobacco, Alcohol, Prescription Medication, and other Substance Use tool
Optimism	• Life Orientation Test-Revised
Trauma	• Brief Trauma Questionnaire
Change measure	
Impression of change	• Patient Global Impression of Change scale

[a]PROMIS SF: Patient-Reported Outcomes Measurement Information System Short Form.

[b]QoL: quality of life.

A structured ophthalmic examination will be conducted at the U-M Kellogg Eye Center. Briefly, we will collect the best corrected visual acuity, intraocular pressures, and pupillary assessment. The optic disc will also be assessed to ensure there is no evidence of optic disc edema to suggest elevated intracranial pressure, which can cause head pain, or an ongoing ophthalmic condition as a cause for COSP. Ocular surface measures were selected based on their use in clinical trials (eg, the Dry Eye Evaluation and Management study) [31] and routine eye examinations to maximize translation to clinical care. All participants will receive proparacaine hydrochloride 0.5%, a rapid-acting topical anesthetic instilled in the inferior fornix of each eye. Ocular pain intensity will be assessed before and 30 seconds after instillation using the Numeric Rating Scale (0-10; 0=no pain and 10=worst pain imaginable). Following the proparacaine challenge, participants will undergo Schirmer testing with an anesthetic.

Neurobiological Assessment

A subset of participants (participants with COSP: n=80; control participants: n=40) will undergo a neurobiological assessment during a separate appointment at the CPFRC. Participants with COSP from across the sign and symptom discordance spectrum will be selected to participate. We will use a standard discordance scoring system generated from the clinical assessment data as described by Vehof et al [28] and used in previous studies [29]. During this visit, participants will undergo a battery of static and dynamic QST, in addition to a structural and fMRI of the brain. Briefly, our QST battery includes assessments of (1) corneal sensation detection and pain threshold testing using a commercially available gas esthesiometer (Brill Engines); (2) local and remote pressure pain thresholds measured at the forehead above the eye and the trapezius, respectively, by algometry (FPX 25; Wagner Instruments) and suprathreshold pressure pain sensitivity measured at the thumbnail bed (Multimodal Automated Sensory Testing System; Arbor Medical Innovations); (3) local and remote heat pain threshold testing at the forehead above the eye and forearm (QST.lab); (4) conditioned pain modulation using trapezius pressure pain threshold as the test stimulus and hand immersion in 10 °C water as the conditioning stimulus; (5) mechanical temporal summation of pain at the volar forearm and above the eye (PinPrick stimulators; MRC Systems GmbH); and (6) visual sensitivity to a flashing annular checkboard presented at multiple levels of brightness (Michigan Visual Aversion Stress Test; U-M).

The neuroimaging assessment includes 4 different magnetic resonance imaging (MRI) procedures, including resting state functional connectivity MRI, evoked pain fMRI using thumbnail pressure, evoked visual stimulation fMRI using a flashing annular checkboard, and voxel-based morphometry. Imaging data will be obtained on a 3.0 T General Electric MRI scanner located at the fMRI laboratory at U-M.

Finally, participants will undergo noninvasive, in vivo corneal confocal microscopy to obtain images of the subbasal corneal nerve fiber layer. The Heidelberg Retinal Tomograph II Rostock Cornea Module, a laser scanning confocal biomicroscope, provides 400 × 400 µm coronal images of the cornea. After placement of topical anesthetic eye drops and a moisturizing gel, contact examination will be performed to image the central cornea. Five representative still images of the central subbasal nerve plexus will be evaluated using image analysis software to quantify nerve density and morphology (eg, microneuromas). Images can be preprocessed using software incorporated in the instrument or NeuronJ, a plug-in of the NIH ImageJ analysis software. No gold standard exists for corneal image preprocessing. We will use whichever of these methods or any newly developed method with the best empirical support at the time the analyses are performed. Textbox 3 depicts key elements of the neurobiological testing protocol.

XSL•FO
RenderX

Textbox 3. Key elements of the deep phenotyping assessment for a subset of participants (participants with chronic ocular surface pain: n=80; controls: n=40).

Quantitative sensory testing

- Pressure pain sensitivity
- Heat pain threshold
- Corneal sensation
- Visual sensitivity
- Conditioned pain modulation
- Temporal summation

Neuroimaging

- Structural magnetic resonance imaging (T1 and T2)
- Resting state functional magnetic resonance imaging (before and after evoked pain)
- Evoked pain functional magnetic resonance imaging

Study Oversight

Study oversight will be under the direction of an independent study monitor (TD) and the principal investigator (LBDL). A study monitor was selected as an additional protection against risks for human participants. The study monitor will be available in real time to review and recommend appropriate actions regarding adverse events and other safety issues to the principal investigator and the study team.

Statistical Design and Analysis

Our central hypothesis is that as the degree of symptoms out of proportion to signs (ie, discordance) increases, established clinical, neurobiological, and treatment response features indicative of nociplastic pain, such as multisite pain, multisensory hypersensitivity, and nonpain CNS-mediated symptoms, will also increase. We will use descriptive statistics to explore the data across aims, including cross-sectional correlation analyses of discordance with clinical and neurobiological data, as well as changes in pain intensity scores pre- and posttopical anesthetic application.

Likewise, linear regression models will be fit to assess the effects on the response variables, including discordance measures regressed on clinical nociplastic pain features, both before and after adding neurobiological features. Analysis of covariance–style linear regressions of change in pain intensity will be performed with nociplastic pain features, discordance, and pain intensity at baseline as covariates. For all models, the list of covariates will include systemic medications that potentially affect pain and ocular surface dryness treatment responses and other clinical and demographic variables. Backward variable selection procedures will be used along with Bayesian information criteria for best model selection. In high-dimensional settings, an elastic net regularization will be done to stabilize the regressions. Preference will be given to the least absolute shrinkage and selection operator penalties that assign 0 weights to nonsignificant predictors. Cross validation (10 fold) with 200 participants will be done to provide an approximate, unbiased assessment of model performance. This will be followed by second-order validation using an external validation set of 50 participants with COSP. Residuals will be examined for possible departures from normality. Generalized linear model and feature transformations will be used, if needed, to deal with violations of linear model assumptions. Simultaneous model-based CIs will be reported that characterize the joint character of hypothesis testing rather than a multiple comparisons adjustment.

Logistic models will be used with dichotomous outcomes for discordance (yes or no) in aims 1 and 2 and 50% relative improvement in pain intensity (yes or no) in aim 3 (including a mixed logistic generalized linear mixed model in the longitudinal analytic setup). Linear mixed models (or generalized linear mixed models) will be used as an alternative in aim 3. Gaussian random intercept shared within the participant will be used to model the participant-specific effect, with measurements before and after treatment serving as the response. The model will be used to characterize the variance components. Estimates will be obtained by maximum likelihood, followed by best model selection and hypothesis testing using Bayesian information criteria and likelihood ratio model–based tests.

We will perform 2 exploratory analyses. First, a latent class or latent variable model will be used to explore an unsupervised measurement of the CNS involvement status (ie, nociplastic pain). This will allow us to suggest an alternative to defining discordance as an empirical measure. Second, in a separate exploratory exercise, we will reconcile the results of this project with historical CPFRC studies of other chronic pain cohorts (eg, chronic low back pain and knee osteoarthritis) that explored CNS pain amplification and persistent pain after surgery. Intuitively, we anticipate general common features of CNS involvement across chronic pain conditions and will look for a possible joint model of multiple disease cohorts to assess disease-specific versus general features that characterize nociplastic pain.

Reasons for missingness will be analyzed using multinomial logistic regression. The results of formal missing data imputation by predictive-matching algorithms will be compared to the results of missing data exclusion in a sensitivity analysis. We

XSL•FO
RenderX

do not expect the fraction of missing data to exceed 10%; therefore, 1 round of imputation will likely be sufficient. An alternative approach is multiple imputation [32].

Power Calculations

We will be able to reliably detect a correlation (r) of 0.2 or higher with 81% power when tested at the 5% significance level with 200 participants. Higher correlations (r=0.40) were observed in our previous knee osteoarthritis studies between pain and the widespread pain index (included in the Michigan Body Map), measuring CNS amplification [33]. Published analyses of discordance in patients with dry eye disease based on an objective tear production measure (Schirmer tests) confirm effects at least as strong as in this study [8]. In the QST studies, the observed correlation was 0.2 between the forehead and forearm hot and cold pain intensity measurements and discordance measures [17]. Correlations of 0.25 were observed in the literature between burning ocular pain and the presence of hypersensitivity to heat in the forearm [34]. For the treatment response study (aim 3), 92% power can be obtained with as few as 40 participants with nociplastic pain, which is easily attainable with the sample size of 200 in our study. This assessment is based on the 2-tailed t test comparing pain intensity among participants with and without persistent pain in a study of 224 veterans (mean 6.7, SD 2.6 vs mean 4.5, SD 2.9) [21].

Data Management and Sharing

Data will be stored on Health Insurance Portability and Accountability Act (HIPAA)–compliant systems (eg, [REDCap; Research Electronic Data Capture; Vanderbilt University]) to protect the personal health information of participants. Participant demographics will be collected using NIH-endorsed common data elements, for example, age, sex, race, and ethnicity. Participant clinical data will be collected using forms from the NIH Common Data Elements repository when possible [35]. Some participant-reported outcome data (eg, Patient Reported Outcomes Measurement Information System 29; version 2) have standard collection forms available in the NIH Common Data Elements repository. Others that do not (eg, Ocular Surface Disease Index) will be collected using standard survey methodology procedures.

Sufficient data from this project will be preserved to enable validation and replication of the research findings described in the aims. Deidentified data, including participant clinical data, sensory testing data, and survey responses, will be shared. To protect participant privacy and confidentiality, shared data will be deidentified using the “safe harbor” method, in compliance with the HIPAA privacy rule. We will archive our corneal image data in the Deep Blue Data (DBD) repository. DBD is a generalist repository developed at the U-M in support of sharing and preserving research data. DBD provides free public access and downloading of data. All other data, including neuroimaging data, will be deposited in the National Institute of Mental Health Data Archive (NDA). NDA makes human participants’ data collected from hundreds of research projects across many scientific domains available. NDA provides infrastructure for sharing research data, tools, methods, and analyses, enabling collaborative science and discovery. Data deposited in DBD will be accessible for at least 10 years, after which they will be appraised for continued preservation by the university library personnel. Data deposited in the NDA will be archived indefinitely.

Results

Our study is supported by a grant from the National Eye Institute (NIH R01EY036357). Recruitment and enrollment commenced in January 2025 following finalization and piloting of the study protocol. As of January 2026, we have enrolled 59 participants and completed 51 clinical phenotyping study visits. Of those 51 participants, 15 have completed neurobiological phenotyping visits. The remaining participants are enrolled but have not completed either study visit.

Discussion

Expected Outcomes and Future Directions

The transfer of experience and knowledge from other COPCs with a nociplastic component (eg, fibromyalgia, interstitial cystitis or bladder pain syndrome, and irritable bowel syndrome) to studies of the eye has strong potential to change how we treat COSP and will advance our understanding of ocular pain mechanisms more broadly. This is the first-of-its-kind comprehensive study of clinical and neurobiological nociplastic pain features in COSP using validated, leading-edge pain assessments. The proposed work will open a new view of COSP as part of a broader systemic chronic pain state driven preferentially by CNS dysfunction in many patients. Our extensive clinical and neurobiological data can be leveraged to (1) support larger studies that validate our findings in broader populations, (2) design clinical criteria for detecting subsets of patients with COSP with nociplastic pain, (3) identify those who may respond preferentially to CNS-directed therapies and monitoring both clinical and neurobiological responses to treatment, and (4) identify and modify antecedents of nociplastic pain development in patients with COSP. Ultimately, successful future treatment strategies for COSP depend not only on the success of addressing relevant nociceptive and neuropathic pain but also on the degree to which nociplastic pain can be identified and managed.

Conclusions

COSP is disabling for many patients, and our current treatments aimed at the ocular surface—the presumed origin of pain—are insufficient. Advancing our understanding of pain mechanisms in COSP is critically needed to target treatments effectively. This research will rigorously define the role of nociplastic pain in COSP with a large representative cohort of patients using established clinical phenotypic, neurobiological, and treatment response features. This represents a crucial step forward in reaching the ultimate goal of realizing targeted therapeutic strategies tailored to individual patients experiencing COSP.

XSL•FO
RenderX

Acknowledgments

Generative artificial intelligence was not used in the research or writing of this manuscript. Responsibility for the final manuscript lies entirely with the authors.

Funding

This research protocol was funded by the National Institutes of Health (grant R01EY036357).

Data Availability

The datasets generated during this study will be available in the Deep Blue Data and National Institute of Mental Health Data Archive repositories.

Authors' Contributions

SEH, CK, DAW, RS, AT, AG, MAW, and DJC conceived the study and contributed to the study design, study protocol, and manuscript preparation and revision. LBDL conceived the study, contributed to the design of the study, prepared the study protocol and manuscript, and secured funding. KAS, CM, TD, AS, and CS revised the study protocol and the manuscript.

Conflicts of Interest

SEH receives grants and equity from Arbor Medical Innovations. DJC has performed consulting for Tonix, Axsome, Virios, and Merck. All other authors have no conflicts to declare.

Multimedia Appendix 1
SPIRIT checklist.
[DOCX File , 34 KB - resprot_v15i1e84240_app1.docx]

References

1. Schaumberg DA, Dana R, Buring JE, Sullivan DA. Prevalence of dry eye disease among US men: estimates from the Physicians' Health Studies. Arch Ophthalmol 2009 Jun;127(6):763-768 [FREE Full text] [doi: 10.1001/archophthalmol.2009.103] [Medline: 19506195]
2. Schaumberg DA, Sullivan DA, Buring JE, Dana MR. Prevalence of dry eye syndrome among US women. Am J Ophthalmol 2003 Aug;136(2):318-326. [doi: 10.1016/s0002-9394(03)00218-6] [Medline: 12888056]
3. Stapleton F, Alves M, Bunya VY, Jalbert I, Lekhanont K, Malet F, et al. TFOS DEWS II epidemiology report. Ocul Surf 2017 Jul;15(3):334-365. [doi: 10.1016/j.jtos.2017.05.003] [Medline: 28736337]
4. Morthen MK, Magno MS, Utheim TP, Snieder H, Hammond CJ, Vehof J. The physical and mental burden of dry eye disease: a large population-based study investigating the relationship with health-related quality of life and its determinants. Ocul Surf 2021 Jul;21:107-117. [doi: 10.1016/j.jtos.2021.05.006] [Medline: 34044135]
5. Sayegh RR, Yu Y, Farrar JT, Kuklinski EJ, Shtein RM, Asbell PA, Dry Eye Assessment and Management (DREAM) Study Research Group. Ocular discomfort and quality of life among patients in the dry eye assessment and management study. Cornea 2021 Jul 01;40(7):869-876 [FREE Full text] [doi: 10.1097/ICO.0000000000002580] [Medline: 33290317]
6. Mertzanis P, Abetz L, Rajagopalan K, Espindle D, Chalmers R, Snyder C, et al. The relative burden of dry eye in patients' lives: comparisons to a U.S. normative sample. Invest Ophthalmol Vis Sci 2005 Jan 01;46(1):46-50. [doi: 10.1167/iovs.03-0915] [Medline: 15623753]
7. Le Q, Zhou X, Ge L, Wu L, Hong J, Xu J. Impact of dry eye syndrome on vision-related quality of life in a non-clinic-based general population. BMC Ophthalmol 2012 Jul 16;12(1):22 [FREE Full text] [doi: 10.1186/1471-2415-12-22] [Medline: 22799274]
8. Shtein RM, Harper DE, Pallazola V, Harte SE, Hussain M, Sugar A, et al. Discordant dry eye disease (an American Ophthalmological Society thesis). Trans Am Ophthalmol Soc 2016 Aug;114:T4 [FREE Full text] [Medline: 28050051]
9. Um SB, Yeom H, Kim NH, Kim HC, Lee HK, Suh I. Association between dry eye symptoms and suicidal ideation in a Korean adult population. PLoS One 2018 Jun 20;13(6):e0199131 [FREE Full text] [doi: 10.1371/journal.pone.0199131] [Medline: 29924835]
10. Bron AJ, de Paiva CS, Chauhan SK, Bonini S, Gabison EE, Jain S, et al. TFOS DEWS II pathophysiology report. Ocul Surf 2017 Jul;15(3):438-510. [doi: 10.1016/j.jtos.2017.05.011] [Medline: 28736340]
11. Maixner W, Fillingim RB, Williams DA, Smith SB, Slade GD. Overlapping chronic pain conditions: implications for diagnosis and classification. J Pain 2016 Sep;17(9 Suppl):T93-107 [FREE Full text] [doi: 10.1016/j.jpain.2016.06.002] [Medline: 27586833]

XSL•FO
RenderX

12. Fitzcharles MA, Cohen SP, Clauw DJ, Littlejohn G, Usui C, Häuser W. Nociplastic pain: towards an understanding of prevalent pain conditions. Lancet 2021 May 29;397(10289):2098-2110. [doi: 10.1016/S0140-6736(21)00392-5] [Medline: 34062144]
13. Kutch JJ, Ichesco E, Hampson JP, Labus JS, Farmer MA, Martucci KT, MAPP Research Network. Brain signature and functional impact of centralized pain: a multidisciplinary approach to the study of chronic pelvic pain (MAPP) network study. Pain 2017 Oct 29;158(10):1979-1991 [FREE Full text] [doi: 10.1097/j.pain.0000000000001001] [Medline: 28692006]
14. Schrepf A, Williams DA, Gallop R, Naliboff BD, Basu N, Kaplan C, MAPP Research Network. Sensory sensitivity and symptom severity represent unique dimensions of chronic pain: a MAPP research network study. Pain 2018 Oct 28;159(10):2002-2011 [FREE Full text] [doi: 10.1097/j.pain.0000000000001299] [Medline: 29863527]
15. Harper DE, Schrepf A, Clauw DJ. Pain mechanisms and centralized pain in temporomandibular disorders. J Dent Res 2016 Sep;95(10):1102-1108 [FREE Full text] [doi: 10.1177/0022034516657070] [Medline: 27422858]
16. De Lott LB, Kaplan C, Harte S, Clauw DJ, Galor A, Vehof J, et al. Nociplastic pain among individuals with chronic ocular surface pain: one cause for "pain without stain"? Surv Ophthalmol 2025 May;70(3):536-543. [doi: 10.1016/j.survophthal.2025.01.004] [Medline: 39814104]
17. Ong ES, Felix ER, Levitt RC, Feuer WJ, Sarantopoulos CD, Galor A. Epidemiology of discordance between symptoms and signs of dry eye. Br J Ophthalmol 2018 May 18;102(5):674-679. [doi: 10.1136/bjophthalmol-2017-310633] [Medline: 28821553]
18. Magno MS, Utheim TP, Snieder H, Hammond CJ, Vehof J. The relationship between dry eye and sleep quality. Ocul Surf 2021 Apr;20:13-19 [FREE Full text] [doi: 10.1016/j.jtos.2020.12.009] [Medline: 33421635]
19. Galor A, Covington D, Levitt AE, McManus KT, Seiden B, Felix ER, et al. Neuropathic ocular pain due to dry eye is associated with multiple comorbid chronic pain syndromes. J Pain 2016 Mar;17(3):310-318 [FREE Full text] [doi: 10.1016/j.jpain.2015.10.019] [Medline: 26606863]
20. Galor A, Seiden BE, Park JJ, Feuer WJ, McClellan AL, Felix ER, et al. The association of dry eye symptom severity and comorbid insomnia in US veterans. Eye Contact Lens 2018 Sep;44(Suppl 1):S118-S124 [FREE Full text] [doi: 10.1097/ICL.0000000000000349] [Medline: 28181961]
21. Crane AM, Feuer W, Felix ER, Levitt RC, McClellan AL, Sarantopoulos KD, et al. Evidence of central sensitisation in those with dry eye symptoms and neuropathic-like ocular pain complaints: incomplete response to topical anaesthesia and generalised heightened sensitivity to evoked pain. Br J Ophthalmol 2017 Sep 18;101(9):1238-1243. [doi: 10.1136/bjophthalmol-2016-309658] [Medline: 28100479]
22. Vehof J, Kozareva D, Hysi PG, Harris J, Nessa A, Williams FK, et al. Relationship between dry eye symptoms and pain sensitivity. JAMA Ophthalmol 2013 Oct 01;131(10):1304-1308. [doi: 10.1001/jamaophthalmol.2013.4399] [Medline: 23907167]
23. Choudhury A, Reyes N, Galor A, Mehra D, Felix E, Moulton EA. Clinical neuroimaging of photophobia in individuals with chronic ocular surface pain. Am J Ophthalmol 2023 Feb;246:20-30 [FREE Full text] [doi: 10.1016/j.ajo.2022.09.020] [Medline: 36223850]
24. Clemens JQ, Mullins C, Kusek JW, Kirkali Z, Mayer EA, Rodríguez LV, MAPP Research Network Study Group. The MAPP research network: a novel study of urologic chronic pelvic pain syndromes. BMC Urol 2014 Aug 01;14(1):57 [FREE Full text] [doi: 10.1186/1471-2490-14-57] [Medline: 25085007]
25. Hassett AL, Williams DA, Harris RE, Harte SE, Kaplan CM, Schrepf A, et al. An interventional response phenotyping study in chronic low back pain: protocol for a mechanistic randomized controlled trial. Pain Med 2023 Aug 04;24(Suppl 1):S126-S138 [FREE Full text] [doi: 10.1093/pm/pnad005] [Medline: 36708026]
26. Greco CM, Wasan AD, Schneider MJ, Mehling W, Williams DA, Darwin J, et al. Biobehavioral assessments in BACPAC: recommendations, rationale, and methods. Pain Med 2023 Aug 04;24(Suppl 1):S61-S70 [FREE Full text] [doi: 10.1093/pm/pnac175] [Medline: 36370080]
27. Mauck MC, Lotz J, Psioda MA, Carey TS, Clauw DJ, Majumdar S, et al. The Back Pain Consortium (BACPAC) Research Program: structure, research priorities, and methods. Pain Med 2023 Aug 04;24(Suppl 1):S3-12 [FREE Full text] [doi: 10.1093/pm/pnac202] [Medline: 36622041]
28. Vehof J, Sillevis Smitt-Kamminga N, Nibourg SA, Hammond CJ. Predictors of discordance between symptoms and signs in dry eye disease. Ophthalmology 2017 Mar;124(3):280-286. [doi: 10.1016/j.ophtha.2016.11.008] [Medline: 28024826]
29. Yu K, Asbell PA, Shtein RM, Ying GS, for Dry Eye Assessment and Management Study Research Group. Dry eye subtypes in the Dry Eye Assessment and Management (DREAM) study: a latent profile analysis. Transl Vis Sci Technol 2022 Nov 01;11(11):13 [FREE Full text] [doi: 10.1167/tvst.11.11.13] [Medline: 36383391]
30. Health research. University of Michigan. URL: https://umhealthresearch.org/ [accessed 2025-12-30]
31. Asbell PA, Maguire MG, Peskin E, Bunya VY, Kuklinski EJ, Dry Eye Assessment and Management (DREAM©) Study Research Group. Dry Eye Assessment and Management (DREAM©) study: study design and baseline characteristics. Contemp Clin Trials 2018 Aug;71:70-79 [FREE Full text] [doi: 10.1016/j.cct.2018.06.002] [Medline: 29883769]
32. Rubin DB. Multiple imputation after 18+ years. J Am Stat Assoc 1996 Jun 21;91(434):473-489. [doi: 10.1080/01621459.1996.10476908]

XSL•FO
RenderX

33. Neville SJ, Clauw AD, Moser SE, Urquhart AG, Clauw DJ, Brummett CM, et al. Association between the 2011 fibromyalgia survey criteria and multisite pain sensitivity in knee osteoarthritis. Clin J Pain 2018 Oct;34(10):909-917 [FREE Full text] [doi: 10.1097/AJP.0000000000000619] [Medline: 29642237]
34. Galor A, Levitt RC, McManus KT, Kalangara JP, Seiden BE, Park JJ, et al. Assessment of somatosensory function in patients with idiopathic dry eye symptoms. JAMA Ophthalmol 2016 Nov 01;134(11):1290-1298. [doi: 10.1001/jamaophthalmol.2016.3642] [Medline: 27685701]
35. Use of CDEs supports the NIH Data Management and Sharing Policy. National Institutes of Health Common Data Element Repository. URL: https://cde.nlm.nih.gov/home [accessed 2025-12-30]

Abbreviations

CNS: central nervous system
COPC: chronic overlapping pain condition
COSP: chronic ocular surface pain
CPFRC: Chronic Pain and Fatigue Research Center
DBD: Deep Blue Data
fMRI: functional magnetic resonance imaging
HEAL: Helping to End Addiction Long-term
HIPAA: Health Insurance Portability and Accountability Act
MRI: magnetic resonance imaging
NDA: National Institute of Mental Health Data Archive
NIH: National Institutes of Health
QST: quantitative sensory testing
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
U-M: University of Michigan

Edited by J Sarvestan; submitted 17.Sep.2025; this is a non–peer-reviewed article;accepted 19.Dec.2025; published 15.Jan.2026.

Please cite as:
De Lott LB, Harte SE, Kaplan C, Williams DA, Shtein R, Woodward MA, Tsodikov A, Deveney T, Galor A, Shea A, Schultz C, McKolay C, Scott KA, Clauw DJ
Central Nervous System Mechanisms and Treatment Response in Chronic Ocular Surface Pain: Protocol for a Cross-Sectional Observational Phenotyping Study
JMIR Res Protoc 2026;15:e84240
URL: https://www.researchprotocols.org/2026/1/e84240
doi:10.2196/84240
PMID:

©Lindsey B De Lott, Steven E Harte, Chelsea Kaplan, David A Williams, Roni Shtein, Maria A Woodward, Alexander Tsodikov, Tatiana Deveney, Anat Galor, Anne Shea, Charles Schultz, Clare McKolay, Kathy A Scott, Daniel J Clauw. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 15.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Very Low–Carbohydrate Breakfast Intervention for Adults with Type 2 Diabetes and Persistent Hyperglycemia: Protocol for a Digital, Nonrandomized Pre-Post Study

Annika Hansen[1], MA; Kaitlyn Raymond[2], BS; Alexander Simon[2], MS; Sarah Kim[3], MD; Hovig Bayandorian, MA; Deanna Marriott[2], PhD; Laura R. Saslow[2], PhD

[1]School of Medicine, University of Utah, Salt Lake City, UT, United States

[2]School of Nursing, University of Michigan, Ann Arbor, MI, United States

[3]School of Medicine, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Laura R. Saslow, PhD
School of Nursing
University of Michigan
2178 400 North Ingalls Building
Ann Arbor, MI, 94709
United States
Phone: 1 7347647836
Email: saslowl@umich.edu

Abstract

Background: More than 15% of US adults with type 2 diabetes have persistent hyperglycemia. Adults with persistent hyperglycemia and type 2 diabetes have an elevated health risk of a variety of outcomes, including amputation and mortality from cardiovascular disease and from all causes. Nutrition-focused interventions can be effective for improving glycemic control, reducing antihyperglycemic medications, and reducing body weight, all of which are critical outcomes for adults with type 2 diabetes. Carbohydrate intake impacts postprandial glycemia more than any other dietary factor. The American Diabetes Association now recommends a very low–carbohydrate diet, because of its ability to improve glycemic control, for the treatment of type 2 diabetes. However, typical nutrition-focused interventions can be burdensome, as the interventions often have complex instructions and require changing one's diet completely. Additionally, adults with type 2 diabetes and persistent hyperglycemia may be more likely to have low health literacy levels, which can be a barrier to adherence to complex interventions.

Objective: This study aimed to evaluate the effectiveness of a digital, small-steps intervention that focuses on implementing a very low–carbohydrate dietary pattern specifically at breakfast for adults with type 2 diabetes and persistent hyperglycemia. The goal is to determine whether this targeted dietary modification can lead to reductions in hemoglobin A_{1c} levels and decreased use of antihyperglycemic medications, without requiring participants to change their entire diet.

Methods: The Breakfast Study will enroll adults with a hemoglobin A_{1c} (HbA_{1c}) of 7.0% or higher to our online, 4-month intervention, which will teach participants to change their breakfasts to be very low in carbohydrates. We will measure acceptability and feasibility, plus critical efficacy outcomes, such as changes in HbA_{1c}, antihyperglycemic medications, glycemic variability, body weight, blood pressure, and lipids. We will also test whether factors such as sex and baseline insulin resistance significantly moderate the impact of the intervention on change in HbA_{1c} and antihyperglycemic medications. If the results are promising, we will conduct a follow-up, powered, longer randomized controlled trial of this approach. As the prevalence of type 2 diabetes and the understanding of personalized interventions continue to increase, there is a critical need to provide additional effective options for population-level type 2 diabetes treatment strategies, especially for adults with type 2 diabetes and persistent hyperglycemia.

Results: As of October 2025, we have enrolled 119 participants. The results will be published separately.

Conclusions: The Breakfast Study is a nonrandomized, pre-post trial to assess the acceptability, feasibility, and preliminary effectiveness of an accessible, very low–carbohydrate breakfast for adults with type 2 diabetes and persistent hyperglycemia. This study could provide support for continued research investigating how to lower barriers to dietary interventions for type 2 diabetes.

Trial Registration: ClinicalTrials.gov NCT05986097; https://clinicaltrials.gov/study/NCT05986097

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/81041

(JMIR Res Protoc 2026;15:e81041) doi:10.2196/81041

KEYWORDS

type 2 diabetes; glycemic control; very low-carbohydrate breakfast; nutrition; glycemic variability; lifestyle intervention

Introduction

Background and Rationale

Type 2 diabetes is one of the largest and most prevalent contemporary public health problems in the United States. If the current trajectory of prevalence continues, 1 in 3 adults in the United States will have type 2 diabetes by 2050 [1]. It is the most expensive health condition in terms of total health care spending nationwide [2], and these costs have continued to grow, from roughly US $175 billion in 2007 to US $327 billion in 2017 [3,4].

More than 15% of US adults with type 2 diabetes have persistent hyperglycemia [5]. Patients with type 2 diabetes have an elevated risk of a variety of health outcomes, including stroke, blindness, renal failure, premature death, and mortality from cardiovascular disease and all causes [6]. More than half of all amputations in the United States are due to complications from type 2 diabetes [7].

Nutrition-focused interventions may be effective for improving critical outcomes for adults with type 2 diabetes, such as improving glycemic control, reducing antihyperglycemic medications, and reducing body weight. Carbohydrate intake has the strongest impact on postprandial glycemia of any dietary factor. A very low–carbohydrate diet, because of its ability to improve glycemic control and reduce antihyperglycemic medications, is now recommended by the American Diabetes Association for the treatment of type 2 diabetes [8]. However, typical nutrition-focused interventions can be burdensome, as they often use complex instructions and require a complete revision of one's diet. Additionally, adults with type 2 diabetes and persistent hyperglycemia are more likely to have low health literacy levels, which can be a barrier to adherence to complex self-management interventions [9,10].

The benefits of a very low-carbohydrate diet may be available to individuals who change only their breakfasts to be very low in carbohydrates, rather than modifying their entire diet. For example, in a 3-month randomized trial of 59 adults with overweight or obesity with type 2 diabetes, there was a greater reduction in hemoglobin A_{1c} (HbA_{1c}) and antihyperglycemic medications in those assigned to eat a higher-calorie, reduced-carbohydrate breakfast compared to those assigned to eat a smaller, higher-carbohydrate breakfast. However, the participants already had well-controlled type 2 diabetes (average baseline HbA_{1c} level was 6.9%) and the reduced carbohydrate breakfast still had a moderate amount of carbohydrates (39% of calories from carbohydrates) and therefore would not be considered a very low–carbohydrate meal [11].

In a feeding trial of 23 adults with type 2 diabetes, participants consumed isocaloric diets that differed only in whether breakfast was very low in carbohydrates or higher in carbohydrates. On days that the participants consumed the very low–carbohydrate diet breakfast, their morning postprandial hyperglycemia was reduced, and their before-dinner perception of hunger was lower [12].

A different large trial found that postprandial glucose variability in response to a standardized breakfast predicts subsequent hunger and, thus, breakfast may be an important target for changing the day-long eating behavior. Researchers tracked the continuous glucose responses of more than 1000 people and found a correlation between postprandial glucose levels and subsequent hunger and calorie consumption. Researchers found that larger blood glucose dips after eating led to greater later-day calorie consumption. This suggests that, by eating meals that reduce postprandial glycemic variability, subsequent hunger and, therefore, subsequent eating might be reduced [13]. This trial was conducted in people without type 2 diabetes. Previous research shows that people with type 2 diabetes have, on average, twice as large a blood glucose spike in response to eating carbohydrates as compared to people without type 2 diabetes [14]. This suggests that reducing carbohydrate consumption for glycemic control in people with type 2 diabetes is of even greater importance.

A 3-month randomized trial investigated the effects of a low-carbohydrate breakfast versus a low-fat breakfast on glycemic control in individuals with type 2 diabetes. The low-carbohydrate diet showed a reduction in HbA_{1c} levels. Continuous glucose monitoring revealed improved metrics for the low-carbohydrate group, including lower mean and maximum glucose levels, reduced glycemic variability, and increased time within the target glucose range. However, this study only included participants who had an HbA_{1c} level of 6.5% to 8.5% and not those with persistent hyperglycemia [15].

The above research suggests the benefit of a very low–carbohydrate diet breakfast for improving glycemic control, reducing the need for antihyperglycemic medications, reducing postprandial glucose variability, and reducing both subsequent hunger and calorie consumption.

Objectives

To conduct an acceptability, feasibility, and preliminary effectiveness trial in 119 adults with type 2 diabetes and persistent hyperglycemia. We will assign adults with an HbA_{1c} of 7.0% or higher to our 4-month, digital, small-steps, very low-carbohydrate breakfast–focused program (The Breakfast Study).

We will examine the acceptability and feasibility of the very low-carbohydrate breakfast intervention, including participants' satisfaction with the intervention and their diabetes treatment satisfaction, health-related quality of life, and dietary adherence.

XSL•FO
RenderX

For feasibility, we will record the mean number of completed sessions (out of 16 possible). We have denoted a 50% attendance of assigned sessions as our a priori benchmark for feasibility. Using an SD of 30%, this trial has >85% power to detect this level of completion (vs 50% completion) using a 2-tailed *t* test at an level of 5%.

We will assess health and psychological outcomes, including HbA_{1c}, antihyperglycemic medications, glycemic variability, body weight, lipids, insulin resistance, blood pressure, self-efficacy for dietary adherence, and satiety.

We will test whether factors such as sex, emotional eating, and baseline insulin resistance significantly moderate the impact of the intervention on change in HbA_{1c}.

An effective and simple intervention for adults with diabetes and persistent hyperglycemia is needed to reduce both HbA_{1c} levels and antihyperglycemic medications. We hypothesize that a digital, small-steps, very low-carbohydrate breakfast–focused program for adults with type 2 diabetes and persistent hyperglycemia could be effective at achieving these aims. Our primary aim is to assess the acceptability and feasibility of such a study.

Methods

Trial Design

This online, remote trial enrolls adults with an HbA_{1c} of 7.0% or higher and assigns them to a 4-month intervention focused on teaching participants to adopt very low–carbohydrate breakfasts.

Study Setting

The intervention will be delivered remotely and is technology supported. Participants can live anywhere in the United States.

Eligibility Criteria

Key inclusion criteria include (1) a current HbA_{1c} level of >7.0%, (2) the ability to read and speak English, (3) age 18-80 years, and (4) the willingness to regularly check blood glucose levels.

Exclusion criteria include (1) unable to provide informed consent; (2) pregnant or planning to become pregnant in the next 12 months, or breastfeeding or less than 6 months post partum; (3) vegan; (4) a C-peptide test result that suggests type 1 diabetes, a previous diagnosis of type 1 diabetes, or latent autoimmune diabetes of adults (if C-peptide is <0.75 ng/mL, participants are ineligible; if C-peptide is between 0.75 ng/mL and 1.5 ng/mL, these participants will need to get an additional test, Glutamic Acid Decarboxylase 65, to rule out latent autoimmune diabetes of adults; if the Glutamic Acid Decarboxylase 65 test result is ≥5 U/mL, participants will be ineligible); (5) cancer; (6) heart failure; (7) kidney failure; (8) liver failure; (9) an untreated mental health condition; (10) currently following a very low–carbohydrate diet or breakfasts; (11) abnormal thyroid levels, if baseline thyroid stimulating hormone is out of range, a free T3 or free T4 test may be ordered; (12) triglyceride >999 mg/dL; (13) a change in diabetes medication within the last 3 months; (14) alcohol use disorder; (15) previous bariatric surgery; (16) difficulty chewing or swallowing; (17) dependence on others for food preparation; and (18) currently enrolled in another investigative study that might conflict with this research.

Participants with a triglyceride level reported on the initial blood draw of 300-499 mg/dL will be enrolled with the requirement of a follow-up lipids test 2 weeks into the intervention. For triglyceride levels 500-999 mg/dL, we will have the participant first discuss the result with their primary care physician so that any medication or dietary changes can occur before starting the study. If any significant changes were made by the participant's physician (eg, addition or removal of a medication), we would measure a new baseline at the appropriate time. If triglycerides are greater than 999 mg/dL, the participant will not be enrolled. Such tests may also be ordered after an initial meeting with the study physician, based on clinical judgment.

Who Will Take Informed Consent?

An authorized team member, who has undergone the necessary ethical training and received clearance from the institutional review board, will discuss the details of the study and consent forms, as well as respond to any questions. It will be emphasized that participation in the study is completely optional.

Intervention

Intervention Description

We propose to enroll 84 adults, after attrition, with type 2 diabetes and persistent hyperglycemia, in our 4-month online intervention. Assessments will take place at baseline and 4 months later, at the end of the intervention. Participants will complete an online course for a very low–carbohydrate breakfast diet program. This online program allows us to conduct the intervention remotely. Participants will receive weekly classes online, a food guide and recipe booklet through the mail, and a diet coach to support them. For details, see the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) flowchart in Figure 1 for a schedule of the intervention timeline (checklist in Multimedia Appendix 1).

XSL•FO
RenderX

Figure 1. Timeline of planned contact for enrollment, study intervention, and assessments for pre-post program for adults with type 2 diabetes. CGM: continuous glucose monitor.

	STUDY PERIOD																		
	SCREENING/ ENROLLMENT				INTERVENTION														
TIMEPOINT (weeks)	0	1	2	3	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
SCREENING / ENROLLMENT:																				
Screening Survey	X																			
Screening/Baseline Blood Draw	X																			
Video Consent Visit	X																			
CGM Placement (remote or in person)		X																X		
Med consult (if necessary)				X																
INTERVENTION:																				
Lesson: Video/cumulative handout of materials					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Survey: Including blood glucose levels (reviewed by staff)					X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Coach check-in including optional monthly group session]					X		X		X		X		X		X		X		X	
ASSESSMENTS:																				
CGM 14-day wear			X	X															X	X
One-to-two 24-hour dietary recalls			X	X															X	X
Blood pressure 8 times over 2 weeks (home, self-report)			X	X															X	X
Weight on scale (reports to team automatically)				X																X
Baseline / 4-Month Survey			X	X															X	X

Very Low–Carbohydrate Breakfasts

Study materials will encourage eating a very low–carbohydrate breakfast each day. The very low–carbohydrate diet intervention will focus on food choices, without calorie counting or calorie restriction. We will recommend that breakfast meals should have no more than 10 nonfiber grams of carbohydrates or no more than 20% of total calories from carbohydrates. To avoid confusion about how to interpret this, we will provide participants with copious food options that will vary in their complexity and cost. For examples, see Textbox 1. Through coaching, we will help participants to find options that they enjoy and find accessible.

XSL•FO
RenderX

Textbox 1. Examples of very low–carbohydrate diet breakfast options as part of a pre-post 4-month program for adults with type 2 diabetes.

Breakfast options

- Classic eggs, bacon
- Sausage breakfast casserole
- Low-carbohydrate pancakes and waffles
- Cottage cheese with fruit and nuts
- Almond and coconut crunch cereal
- Vanilla pecan bars
- Cheesy breakfast pizza
- Low-carbohydrate English muffin in a mug

Criteria for Discontinuing or Modifying Allocated Interventions

In the event of any serious adverse events resulting from the intervention, participants will immediately stop the diet and will be included in the intention-to-treat analysis. If there are any changes to the participants' medical treatment plan that conflict with the intervention, participants will also stop the intervention. Furthermore, participants will be notified that they have the right to retract their consent to participate at any time and may refuse to answer any questions from the study team.

Strategies to Improve Adherence to Interventions

To improve adherence to the intervention, there will be responsive coaching (via phone, video, and email) so that participants' questions and comments are replied to in a timely and supportive manner. Reminders about targeted behaviors are tied to greater adherence, so participants will be sent text messages to provide feedback, as well as to motivate, educate, and remind them about targeted behaviors and skills about 5 times per week. Monthly group video sessions will also be offered as an optional way for participants to receive support.

Relevant Concomitant Care Permitted or Prohibited During the Trial

Relevant concomitant care is permitted during the trial.

Provisions for Posttrial Care

We will not be providing posttrial care. We do not anticipate harm due to trial participation.

Outcomes

Aim 1: Assess Acceptability and Feasibility—Acceptability: Intervention Satisfaction

At month 4, we will ask participants, "How would you rate your overall satisfaction with the program?" on a 0-6 Likert scale (0=very dissatisfied; 6=very satisfied).

Additionally, each week, participants will receive a check-in survey. The survey will inquire (1) how often they ate a very low–carbohydrate breakfast on a scale from 1 to 7 and how much they enjoyed eating a very low–carbohydrate breakfast on a scale from 1 to 7 (1=not at all to 7=very much so), (2) what their highest and lowest blood sugar readings were over the past week and what were the circumstances of the readings (only asked of those taking insulin, sulfonylureas, or meglitinides), and (3) any medication changes they may have made over the past 7 days.

Other Exploratory Outcomes for Acceptability

Diabetes Treatment Satisfaction

The Diabetes Treatment Satisfaction Questionnaire, an 8-item scale, measures 2 factors: treatment satisfaction (along dimensions such as flexibility and convenience) and burden from hyperglycemia and hypoglycemia. It is a short, easy-to-answer scale with low participant burden. We will use the Diabetes Treatment Satisfaction Questionnaire at baseline and 4 months later [16].

Health-Related Quality of Life

Health-related quality of life is a central aspect of well-being. We will use the Patient-Reported Outcomes Measurement Information System 29 at baseline and 4 months later [17]. Patient-Reported Outcomes Measurement Information System 29 is a questionnaire [17] that assesses physical functioning, anxiety, depression, fatigue, sleep disturbance, social functioning, and pain.

Physical Symptoms

At baseline and 4 months later, we will also measure physical symptoms using a short, face-valid measure of physical symptoms [18]. Participants report how often over the past month they have had symptoms such as dizziness, diarrhea, constipation, headaches, blurred vision, nausea, pain in their hands or feet, heartburn, and acne.

Exploratory Open-Ended Survey Questions and Qualitative Interviews

In the 4-month survey, all participants will be asked open-ended questions about the aspects of the program participants enjoyed, which aspects could be improved, what made it easier to adhere to the program, and which factors made adherence more difficult.

To better inform the next phase of our study, we will conduct qualitative semistructured interviews at the end of the trial. This will allow us to explore the experience of the participants in greater depth, including their perceptions of the intervention, as well as barriers to and facilitators of their ability to make long-lasting dietary changes. We will model the interview guide

XSL•FO
RenderX

and data analysis on prior research, using a semistructured interview guide that includes open-ended questions to encourage participants' perspectives, thoughts, and beliefs about the intervention and its components. Participants will be interviewed by nonintervention staff from our research team. These interviews will last approximately 20-30 minutes. We will create a summary of the qualitative data, with evidence from participant quotes, to explore participant experiences.

Other Exploratory Outcomes for Feasibility

Participant Recruitment and Progress Through the Trial

As recommended by the CONSORT (Consolidated Standards of Reporting Trials) guidelines, we will report on the percentage enrolled from each recruitment method, the number excluded and why, the number who chose not to enter the trial and why, and the number and reasons for withdrawal [19].

Dietary Adherence

We will assess dietary adherence with 1-2 unannounced 24-hour dietary recalls at baseline and at 4 months, conducted by a dietician. At the 4-month outcome mark, this will serve as a check on our intervention fidelity in terms of whether we have achieved our targeted macronutrient levels for breakfasts. We will explore what percentage of calories in participants' breakfasts were derived from carbohydrates, as well as what percentage of participants ate breakfasts that had 20% or less of their calories from carbohydrates. We will also examine the macronutrient content and quality of other eating occasions throughout the day to assess whether breakfast quality impacted other eating patterns.

Aim 2: Assess Changes in Health and Psychological Outcomes

As appropriate, for the following measures, we will explore how each of these outcomes changed from the baseline to the end of the study. All blood tests will be done at Labcorp lab locations, with the rare exception that if participants are unable to go to a Labcorp location, we will mail them a DTI Laboratories, Inc home HbA_{1c} kit for the 4-month outcome.

HbA_{1c}

HbA_{1c} is a standard measure of overall glycemic control in the clinical care of type 2 diabetes. A higher HbA_{1c} is associated with increased risk of microvascular complications from diabetes [20].

Antihyperglycemic Medications

The study team will ask participants about their current diabetes-related medication regimen using questions that have successfully been used in other studies [21]. We will use a medication-effect score, which combines dosage and strength of medications to assess the overall intensity of an antihyperglycemic medication regimen [22].

Diabetes Control

We will explore metrics of diabetes control and report on the percentage of participants who (1) lower their HbA_{1c} levels to below 7.0% without increasing their medication effect score, (2) lower their HbA_{1c} levels to below 6.5% without increasing their medication effect score, and (3) lower their HbA_{1c} levels to below 6.5%, without increasing their medication effect score, and take no antihyperglycemic medications other than metformin.

Glycemic Variability

Data will be analyzed from a continuous glucose monitor at baseline and 4 months. The research team will download sensor data at the end of the measurement period. We will assess the glucose variability and the proportion of time spent in the euglycemic (3-7.8 mmol/L) and hyperglycemic (≥ 11.1 mmol/L) states, following previous standards for interstitial glucose concentrations [23-25].

Body Weight

We will give participants a body weight scale and ask them to weigh themselves weekly in order to help tailor coaching support. However, our outcome of interest will be weight change from baseline to 4 months. We will examine average weight change and the percentage of participants who lose 5% of their body weight, a clinically meaningful amount of weight loss.

Lipids

At baseline and at 4 months, we will measure triglycerides and fractionated cholesterol using Labcorp's Nuclear Magnetic Resonance LipoProfile [26]. This advanced lipid assay provides measurements that are better associated with elevated cardiovascular risk than conventional lipid assays [27].

Insulin Resistance

At baseline and 4 months, fasting insulin and glucose will be used to estimate insulin resistance by calculating Homeostatic Model Assessment-Insulin Resistance. Homeostatic Model Assessment-Insulin Resistance uses a single fasting blood draw to estimate insulin resistance [28].

Inflammation

At baseline and 4 months, high-sensitivity C-reactive protein, which is an acute phase reactant, will be assessed via blood draw.

Blood Pressure

We will mail participants a blood pressure cuff with instructions from the American Heart Association to correctly measure blood pressure. To control for the white-coat effect, we will use a home-based measurement [29]. We will ask participants to measure their blood pressure level 8 times during the 2 weeks before the study begins and the last 2 weeks of the study. We will use the average over those 2-week periods as our assessment of interest, with the change from baseline to 4 months being our outcome of interest.

Self-Efficacy for Dietary Adherence

At baseline and 4 months later, we will measure this with the 8-item Weight Efficacy Lifestyle Questionnaire Short-Form, which assesses self-efficacy for dietary adherence when facing a variety of challenging situations, including poor mood and social situations [30].

XSL•FO
RenderX

Satiety and Food Responsiveness

At baseline and 4 months later, we will assess whether the intervention improves these 2 factors using two 4-item subscales of the Adult Eating Behavior Questionnaire, one for each factor [31].

Participant Timeline

Potential participants will make initial contact with the study by visiting the study website or reaching out to the study team through a dedicated study phone number or email address. Participants will be directed to a study website where they will take an initial screening survey. If participants appear initially eligible, the participant is invited to watch an online video describing the study and answer questions about their goals for participation and the pros and cons of participation (which previous research suggests can improve trial retention) [32]. Following the online video, they fill out the blood draw consent form. Once consented, participants will go for a screening and baseline fasting blood draw to assess parameters required for inclusion.

If potential participants remain eligible after the blood draw, our study staff will consent them on the phone or via a video meeting and review the study schedules, procedures, and assessments. We will send a fax to the primary care provider explaining the study and that the participant is eligible unless the physician objects. If we are not contacted by the physician in 2 weeks, we assume the physician is supportive. Participants will then (1) complete a baseline survey online, (2) participate in 1-2 unannounced 24-hour dietary recalls over the phone, and (3) wear a continuous glucose monitor mailed to them, or attached in person if they live in the greater Ann Arbor area, and they will be taught how to place the continuous glucose monitor via a video call with a trained staff member. If the participant already has a continuous glucose monitor as part of their normal care, we will ask to be connected to their data. After wearing the continuous glucose monitor for 14 days, participants will mail the device back to the program staff, (4) report their blood pressure measured on the monitor provided by the study 8 times over 2 weeks, (5) measure their body weight on the scale provided twice within 5 minutes to provide an accurate average, and (6) meet online with a study physician to discuss medications and glucose testing, if they are using insulin or a secretagogue to manage their diabetes. Once all baseline steps have been completed, the program will begin asynchronously for each participant. For details, see the SPIRIT flowchart in Figure 1 for the screening and intervention timeline.

Sample Size

To estimate the number of subjects needed for the outcome of HbA_{1c}, we note that the Diabetes Prevention Program showed a 0.1% decrease in HbA_{1c} with an SD of the difference approximately equal to 0.1% [33]. This decrease in HbA_{1c} was deemed clinically significant because it led to a clear reduction in type 2 diabetes incidence compared to the control. The DPP's SD estimate of 0.1% is consistent with 2 studies of very low–carbohydrate diets in adults with prediabetes [34,35]. Thus, we used 0.1% as the SD in our sample size calculation. For a 2-sided, paired *t* test to detect a clinically meaningful difference of 0.1% and a SD of 0.1%, we have more than 90% power with 84 retained participants.

Progression

We have denoted the progression criteria that would inform a decision to proceed to a full-scale trial. We would proceed if the 95% CIs for our measures of acceptability, feasibility, as well as for our health outcome of $HbA_{1c,}$ all meet the following benchmarks, respectively: a mean rating of at least 3.0 on the 0-6 satisfaction scale, at least 40% attendance of assigned sessions, and a decrease of at least 0.1% SD.

Recruitment

To reach our target sample size, we will advertise on social media to reach participants across the United States. We will hang fliers in community centers in various locations in the United States. We will also search the medical records of Michigan Medicine for potentially eligible participants. Each year of the study, we will initially contact the pool of potentially eligible participants by sending them a letter and flyer using a Health Insurance Portability and Accountability Act (HIPAA)–compliant company describing the study, which will refer them to a study webpage and an initial online screening questionnaire. All recruitment materials will be written at no higher than a fifth-grade reading level. We may follow up with a phone call.

Data Collection and Management

Plans for Assessment and Collection of Outcomes

Assessment data will be reviewed for accuracy and completion by study staff. Loss to follow-up and incomplete data will be recorded.

Plans to Promote Participant Retention and Complete Follow-Up

Participants will be paid US $100 for completing the 4-month assessments and will be able to keep any of the intervention-related items we send them, including a body weight scale, a blood pressure cuff, and a glucometer. We also allow them to choose a few kitchen items, such as a skillet, a baking sheet, or ingredients, to send to them at the beginning of the study.

Data Management

Trial data will be collected through online questionnaires via Qualtrics. Labcorp will send laboratory data via electronic files to be added to the study database. There will be a regular review of cases during the enrollment process to ensure proper eligibility. All study data will be stored in REDCap (Research Electronic Data Capture; Vanderbilt University), a cloud-based HIPAA-compliant database. REDCap allows for specified ranges and automatic calculations to reduce errors. Data will be cleaned by the research team upon completion of data collection.

XSL•FO
RenderX

Plans for Collection, Laboratory Evaluation, and Storage of Biological Specimens for Genetic or Molecular Analysis in This Trial or Future Use

Participants will have blood samples drawn at baseline and 4 months later. All blood samples will be drawn, analyzed, and then destroyed by Labcorp, excluding the rare exception where participants will be mailed a DTI Laboratories kit for their 4-month HbA_{1c} measurement.

Statistical Analysis

A biostatistician will perform all outcome statistical analyses. An intention-to-treat analysis will be performed on all participants. A per-protocol analysis will also be used on participants who had high compliance with the very low–carbohydrate breakfast intervention. We defined high compliance as breakfasts that have 20% or less of their calories from net carbohydrates.

Descriptive statistics will be computed for all outcome measures. Continuous variables will be reported using means, medians, IQRs, and SDs based on the distribution. Categorical variables will be presented using frequencies and percentages.

Aim 1: Assess Acceptability and Feasibility

Overview

For this aim, we will explore how measures of acceptability and feasibility change from baseline to the end of the study using a 2-tailed paired *t* test for continuous measures and a test of proportions for dichotomous measures. The magnitude and direction of the effect will be reported. Graphical summaries and Q-Q plots will assess the appropriateness of the normality assumption for the *t* tests. If normality is not indicated, we will take remedial measures such as transformations or nonparametric analyses.

Power Aim 1

For the primary outcome measure of acceptability, we have denoted a mean rating of 4.0 on the 0-6 satisfaction scale as our *a priori* benchmark for acceptability, above the midpoint of the scale, with a meaningful difference of 5.0, one point above. We based our estimate of variance on the satisfaction rating from our most recent 2 trials of adults using our 16-week program (SD=1.3 and SD=0.7). Our proposed trial has >90% power to detect this level of satisfaction, using the conservative values of 1.3 for the SD, using a 2-tailed *t* test at an level of 5%.

Thus, we have at least 85% power for both primary hypotheses of Aim 1.

Aim 2: Assess Changes in Health and Psychological Outcomes

For continuous health and psychological measures, we will explore the change from baseline to the end of the study for each outcome, using a 2-tailed paired *t* test, with a significance level of 0.05 and 95% CIs. As in Aim 1, we will test the assumption of normality and remediate as necessary. For the dichotomous variable measuring achievement of glycemic control, we will report descriptive statistics in comparison to other interventions, such as the Diabetes Prevention Program [36].

Interim Analyses

There are no planned interim analyses or stopping rules for this trial. All intervention-related serious adverse events will be reviewed by the Data and Safety Monitoring Board to determine if the study should be stopped.

Methods for Additional Analyses (eg, Subgroup Analysis)

We will analyze whether baseline characteristics modify the benefits of the low-carbohydrate breakfast diet on HbA_{1c}. We will explore whether there seem to be differences in the magnitude of HbA_{1c} changes from baseline to 4 months later across 3 predefined subgroups: levels of obesity, levels of insulin resistance, and women versus men. Linear mixed models similar to those described above will be used, with the addition of candidate moderators (each in separate models), and interactions between the moderator, intervention arm, and time. These models will be used to estimate the change in the outcome within subgroups and differences in change between subgroups. The focus of these exploratory analyses will be the magnitude and direction of change within and between subgroups. We will also report the statistical test on the interaction term, which represents an overall test of the moderation effect.

Methods in Analysis to Handle Protocol Nonadherence and Any Statistical Methods to Handle Missing Data

To assess the impact of missing data on our outcomes, baseline characteristics of participants will be compared with and without missing data. Primary outcome analyses will use random-intercept-random-slope mixed effects models to approximate changes from the start to the end of the study, 4 months later. Mixed effects models using maximum likelihood estimation allow an appropriate assessment of repeated measures despite the effects of missing data.

Plans to Give Access to the Full Protocol and Participant-Level Data

After the trial's prespecified outcomes have been published, a deidentified dataset will be available to other investigators by request. For access to the dataset, the request must align with institutional review board protocols, and the trial steering committee must agree to the request.

Oversight and Monitoring

Composition of the Coordinating Center and Trial Steering Committee

The primary decision-making body of this study is the investigative team comprising the principal investigator (PI; LRS) and the coinvestigators. The PI is accountable for the overall management of the study. She directs the operations of the study, reviews questions or problems that arise from the study between team meetings, and presents issues to the research team for decision. The PI acts as the intermediary with the funding organization, handling tasks such as submitting annual reports, overseeing fiscal and administrative operations, and managing the coordination and implementation of the study.

XSL•FO
RenderX

The trial steering committee will consist of the PI (LRS), trial coordinator (KR), and study physician (SK). Weekly meetings with the study team will be conducted to review study implementation and discuss adverse events. Additionally, monthly meetings will be held with other study investigators and staff to address any overarching study issues and assess trial progress. Ongoing communication through group email will facilitate discussions related to enrollment inquiries and address issues like recommending medication changes for participants.

The day-to-day operations of the study, encompassing recruitment, data collection, and intervention processes, are overseen by project coordinators and research assistants. Their responsibilities extend to coordinating institutional review board revisions, managing data, monitoring reports, and documenting the completion of necessary trainings. Staff members are tasked with participant recruitment and screening, securing informed consent, and scheduling and conducting follow-up assessments. The lead project manager supervises the development of the study's data tracking system and surveys.

Composition of the Data Monitoring Committee, its Role, and Reporting Structure

A data safety monitoring board, composed of 3 individuals, has been established. Biannual board meetings will be conducted through video conference to assess recruitment, retention, and safety reports, as well as preliminary results.

Adverse Event Reporting and Harms

In the event of a serious adverse event with a likely connection to study participation, a special meeting with the data and safety monitoring board will be convened to assess the necessity for any required modifications or potential early termination of the trial. An adverse event is broadly defined as any occurrence that causes or elevates the risk of harm to the participant or others. Serious adverse events encompass instances leading to death, inpatient hospitalization, or extension of existing hospital stay, persistent or significant disability or incapacity, or the emergence of a congenital anomaly or birth defect. The study team meticulously evaluates all potential adverse events reported by participants, considering their relation to the study intervention, expectedness, and severity.

Following a reduced-carbohydrate diet may result in reduced glucose levels. For individuals taking medications like insulin or a secretagogue, this could elevate the risk of hypoglycemia if appropriate medication adjustments are not implemented. Medical adjustments, particularly reductions (or discontinuations at low doses), will be supervised by study physicians. The order of prioritizing reductions or discontinuations typically follows the sequence outlined: (1) insulin, (2) secretagogues, (3) meglitinides, (4) sodium-glucose transport-2 inhibitors, (5) glucagon-like peptide -1 agonists, glucagon-like peptide 1/glucose-dependent insulinotropic polypeptide dual agonists, and dipeptidyl peptidase-4 inhibitors, (5) alpha glucosidase inhibitors, and (6) thiazolidinediones.

Upon enrollment in the study, we will provide participants' primary care physicians with study information.

The study team will consult the study endocrinologist for potential medication changes if the participants report a low blood glucose level below 90 mg/dL on their weekly survey or if they report their blood glucose as dropping below 110 mg/dL on 2 separate occasions. If the endocrinologist advises medication changes, this will be communicated to the participant. We will fax any changes to the primary care physician.

If participants decide not to accept the medication changes, the study team will also fax the primary care physician our recommendation and the participants' rejection of the suggestion.

Frequency and Plans for Auditing Trial Conduct

The trial will be closely monitored by study investigators and the data and safety monitoring board, convening twice annually. The study team will submit annual progress reports to both the institutional review board and the National Institute of Diabetes and Digestive Kidney Diseases, the study sponsor.

Dissemination Plans

The trial results will be disseminated through conference presentations, uploaded onto ClinicalTrials.gov (NCT05986097, initially registered on August 2, 2023), and published in peer-reviewed journals. All final peer-reviewed manuscripts will be submitted to the digital archive PubMed Central. Additionally, relevant data will be deposited in suitable public repositories where applicable.

Ethical Considerations

This trial uses a single institutional review board, with the primary application approved by the University of Michigan Institutional Review Board (HUM00225646; Office for Human Research Protections Institutional Review Board registration number IRB00000244). When communicating important protocol modifications, the study team will request approval from the University of Michigan Institutional Review Board via eResearch and communicate approved changes to the relevant parties via email, phone, or online meetings. We plan to confirm that all groups understand and agree with the proposed changes. All trial participants provide informed, written consent using forms approved by the institutional review board. We will deidentify and code all surveys and forms with a unique participant number. To encourage participant retention, participants will be paid US $100 for completing the 4-month assessments.

Results

As of October 2025, we have enrolled 119 participants. Recruitment was initiated on August 2, 2023, and the date for completion was October 17, 2025. The results will be published separately.

Discussion

Primary Findings

This paper describes the protocol for the Breakfast Study, an acceptability, feasibility, and preliminary effectiveness trial in

XSL•FO
RenderX

115 adults with type 2 diabetes and persistent hyperglycemia. We will assign adults with an HbA_{1c} of 7.0% or higher to our 4-month, digital, small-steps, very low–carbohydrate breakfast-focused program (The Breakfast Study).

This study aims to establish the feasibility and effectiveness of a dietary intervention that only requires changing breakfast to a very low–carbohydrate one. A previously conducted trial reported improvements in glucose control (reduction in antihyperglycemic medications and HbA_{1c}), with a reduced carbohydrate (moderate carbohydrate), higher calorie breakfast, when compared to a higher carbohydrate breakfast of a smaller caloric portion [11]. Another trial that compared a very low- to a high-carbohydrate breakfast for people with type 2 diabetes showed a reduction in hyperglycemia, postprandially, and a lower perception of hunger predinner [12]. This trial is aimed at addressing important knowledge gaps that have been noted by the American Diabetes Association and other experts in making clinical guidelines for the nutritional management of type 2 diabetes.

Limitations

There are several limitations in this trial. First, because it aims to establish acceptability, feasibility, and preliminary effectiveness, it is not a randomized trial; all participants will be in the intervention group. There are limitations inherent in the single-arm design of this study, as it may not fully address key feasibility aspects required to inform a future randomized controlled intervention. Specifically, we will need to address issues such as participant satisfaction with randomization and adherence to the future control arm of the study. Additionally, the absence of a control group may result in inflated effect estimates. However, the single-arm design was chosen as a pragmatic first step to explore initial feasibility and acceptability, refine the intervention, and identify potential barriers to implementation of the intervention. If this trial is successful, the intent would be to conduct a larger randomized controlled intervention of a very low–carbohydrate breakfast diet intervention.

The trial is also limited in the method of measuring dietary adherence, continually asking participants throughout the trial about their diet, but there is no way of validating the information that is received. The design of the trial is completely online and remote (with rare exceptions based on participant need), thus participants must be technologically literate to participate. The study team is trying to mitigate this limitation by mailing physical copies of the curriculum and recipe book, as well as participant lab forms for their blood work, if requested. This trial also relies on the participants having a physical address so that materials can be mailed to them throughout the study. This is not a meal replacement study, so participants must have the financial means and the mental capacity to adhere to our dietary recommendations.

Conclusion

To our knowledge, the Breakfast Study is the first of its kind assessing the feasibility and efficacy of changing to a very low–carbohydrate breakfast in patients with type 2 diabetes and persistent hyperglycemia. If such an intervention is feasible and efficacious at controlling diabetes, it would be an important intervention to study more, as it is much easier to implement than a completely ketogenic diet.

Acknowledgments

We extend our gratitude to the Data Safety and Monitoring Board: Dr. David Cavan, Dr. Michael Tamber, and Sanjib Basu. We would also like to thank our dietary coach, Brandalyn Riedel, PhD.

Funding Statement

This study was funded by the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health (DK132099). The funding body had no role in the design of the study and will not be involved in the collection, analysis, and interpretation of data, writing reports, or the decision to submit to a journal.

Data Availability

Deidentified participant data will be available upon reasonable request to the corresponding author and pending institutional review board approval.

Authors' Contributions

LRS, SK, and HB were responsible for the conception and design of the research. LRS and AH drafted the manuscript, which was reviewed, edited, and approved by all coauthors.

Conflicts of Interest

LRS’s partner, HB, is an inventor of software used in this study, which purchased a software services agreement for its use.

Multimedia Appendix 1
SPIRIT checklist.
[PDF File (Adobe PDF File), 424 KB - resprot_v15i1e81041_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
Peer-review report from: ZRG1 RPHB-C (90) - Center for Scientific Review Special Emphasis Panel, Risk Prevention and Lifestyle Change, Behavioral Medicine, Psychosocial Development, and Behavioral Health Outcomes (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 107 KB - resprot_v15i1e81041_app2.pdf]

References

1. Boyle JP, Thompson TJ, Gregg EW, Barker LE, Williamson DF. Projection of the year 2050 burden of diabetes in the US adult population: dynamic modeling of incidence, mortality, and prediabetes prevalence. Popul Health Metr 2010;8:29 [FREE Full text] [doi: 10.1186/1478-7954-8-29] [Medline: 20969750]
2. Dieleman JL, Baral R, Birger M, Bui AL, Bulchis A, Chapin A, et al. US spending on personal health care and public health, 1996-2013. JAMA 2016;316(24):2627-2646 [FREE Full text] [doi: 10.1001/jama.2016.16885] [Medline: 28027366]
3. American Diabetes Association. Economic costs of diabetes in the U.S. in 2017. Diabetes Care 2018;41(5):917-928 [FREE Full text] [doi: 10.2337/dci18-0007] [Medline: 29567642]
4. American Diabetes Association. Economic costs of diabetes in the U.S. in 2012. Diabetes Care 2013;36(4):1033-1046 [FREE Full text] [doi: 10.2337/dc12-2625] [Medline: 23468086]
5. Carls G, Huynh J, Tuttle E, Yee J, Edelman SV. Achievement of glycated hemoglobin goals in the US remains unchanged through 2014. Diabetes Ther 2017;8(4):863-873 [FREE Full text] [doi: 10.1007/s13300-017-0280-5] [Medline: 28646411]
6. Cavero-Redondo I, Peleteiro B, Álvarez-Bueno C, Rodriguez-Artalejo F, Martínez-Vizcaíno V. Glycated haemoglobin A1c as a risk factor of cardiovascular outcomes and all-cause mortality in diabetic and non-diabetic populations: a systematic review and meta-analysis. BMJ Open 2017;7(7):e015949 [FREE Full text] [doi: 10.1136/bmjopen-2017-015949] [Medline: 28760792]
7. National diabetes statistics report: estimates of diabetes and its burden in the United States. Centers for Disease Control and Prevention. 2014. URL: https://stacks.cdc.gov/view/cdc/148231 [accessed 2025-11-07]
8. Evert AB, Dennison M, Gardner CD, Garvey WT, Lau KHK, MacLeod J, et al. Nutrition therapy for adults with diabetes or prediabetes: a consensus report. Diabetes Care 2019;42(5):731-754 [FREE Full text] [doi: 10.2337/dci19-0014] [Medline: 31000505]
9. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Viera A, Crotty K, et al. Health literacy interventions and outcomes: an updated systematic review. Evid Rep Technol Assess (Full Rep) 2011(199):1-941. [Medline: 23126607]
10. Schillinger D, Grumbach K, Piette J, Wang F, Osmond D, Daher C, et al. Association of health literacy with diabetes outcomes. JAMA 2002;288(4):475-482. [doi: 10.1001/jama.288.4.475] [Medline: 12132978]
11. Rabinovitz HR, Boaz M, Ganz T, Jakubowicz D, Matas Z, Madar Z, et al. Big breakfast rich in protein and fat improves glycemic control in type 2 diabetics. Obesity (Silver Spring) 2014;22(5):E46-E54 [FREE Full text] [doi: 10.1002/oby.20654] [Medline: 24311451]
12. Chang CR, Francois ME, Little JP. Restricting carbohydrates at breakfast is sufficient to reduce 24-hour exposure to postprandial hyperglycemia and improve glycemic variability. Am J Clin Nutr 2019;109(5):1302-1309 [FREE Full text] [doi: 10.1093/ajcn/nqy261] [Medline: 30968140]
13. Wyatt P, Berry SE, Finlayson G, O'Driscoll R, Hadjigeorgiou G, Drew DA, et al. Author correction: postprandial glycaemic dips predict appetite and energy intake in healthy individuals. Nat Metab 2021;3(7):1032. [doi: 10.1038/s42255-021-00436-1] [Medline: 34253908]
14. Bantle JP, Laine DC, Castle GW, Thomas JW, Hoogwerf BJ, Goetz FC. Postprandial glucose and insulin responses to meals containing different carbohydrates in normal and diabetic subjects. N Engl J Med 1983;309(1):7-12. [doi: 10.1056/NEJM198307073090102] [Medline: 6343873]
15. Oliveira BF, Chang CR, Oetsch K, Falkenhain K, Crampton K, Stork M, et al. Impact of a low-carbohydrate compared with low-fat breakfast on blood glucose control in type 2 diabetes: a randomized trial. Am J Clin Nutr 2023;118(1):209-217 [FREE Full text] [doi: 10.1016/j.ajcnut.2023.04.032] [Medline: 37257563]
16. Bradley C. Diabetes treatment satisfaction questionnaire. Change version for use alongside status version provides appropriate solution where ceiling effects occur. Diabetes Care 1999;22(3):530-532. [doi: 10.2337/diacare.22.3.530] [Medline: 10097946]
17. Hinchcliff M, Beaumont JL, Thavarajah K, Varga J, Chung A, Podlusky S, et al. Validity of two new patient-reported outcome measures in systemic sclerosis: patient-reported outcomes measurement Information System 29-item health profile and functional assessment of chronic illness therapy-dyspnea short form. Arthritis Care Res (Hoboken) 2011;63(11):1620-1628 [FREE Full text] [doi: 10.1002/acr.20591] [Medline: 22034123]
18. McColl E, Steen IN, Meadows KA, Hutchinson A, Eccles MP, Hewison J, et al. Developing outcome measures for ambulatory care--an application to asthma and diabetes. Soc Sci Med 1995;41(10):1339-1348. [doi: 10.1016/0277-9536(95)00120-v] [Medline: 8560301]
19. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 2010;8:18 [FREE Full text] [doi: 10.1186/1741-7015-8-18] [Medline: 20334633]

XSL•FO
RenderX

20. Stratton I, Adler A, Neil H, Matthews D, Manley S, Cull C, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ 2000;321(7258):405-412 [FREE Full text] [doi: 10.1136/bmj.321.7258.405] [Medline: 10938048]
21. Saslow LR, Kim S, Daubenmier JJ, Moskowitz JT, Phinney SD, Goldman V, et al. A randomized pilot trial of a moderate carbohydrate diet compared to a very low carbohydrate diet in overweight or obese individuals with type 2 diabetes mellitus or prediabetes. PLoS One 2014;9(4):e91027 [FREE Full text] [doi: 10.1371/journal.pone.0091027] [Medline: 24717684]
22. Alexopoulos A, Yancy WS, Edelman D, Coffman CJ, Jeffreys AS, Maciejewski ML, et al. Clinical associations of an updated medication effect score for measuring diabetes treatment intensity. Chronic Illn 2021;17(4):451-462 [FREE Full text] [doi: 10.1177/1742395319884096] [Medline: 31653175]
23. Longato E, Acciaroli G, Facchinetti A, Maran A, Sparacino G. Simple linear support vector machine classifier can distinguish impaired glucose tolerance versus type 2 diabetes using a reduced set of CGM-based glycemic variability indices. J Diabetes Sci Technol 2020;14(2):297-302 [FREE Full text] [doi: 10.1177/1932296819838856] [Medline: 30931604]
24. Chakarova N, Dimova R, Grozeva G, Tankova T. Assessment of glucose variability in subjects with prediabetes. Diabetes Res Clin Pract 2019;151:56-64. [doi: 10.1016/j.diabres.2019.03.038] [Medline: 30935927]
25. Dimova R, Chakarova N, Grozeva G, Kirilov G, Tankova T. The relationship between glucose variability and insulin sensitivity and oxidative stress in subjects with prediabetes. Diabetes Res Clin Pract 2019;158:107911. [doi: 10.1016/j.diabres.2019.107911] [Medline: 31707004]
26. Jeyarajah EJ, Cromwell WC, Otvos JD. Lipoprotein particle analysis by nuclear magnetic resonance spectroscopy. Clin Lab Med 2006;26(4):847-870. [doi: 10.1016/j.cll.2006.07.006] [Medline: 17110242]
27. Krauss R. Lipoprotein subfractions and cardiovascular disease risk. Curr Opin Lipidol 2010;21(4):305-311. [doi: 10.1097/MOL.0b013e32833b7756] [Medline: 20531184]
28. Wallace TM, Levy JC, Matthews DR. Use and abuse of HOMA modeling. Diabetes Care 2004;27(6):1487-1495. [doi: 10.2337/diacare.27.6.1487] [Medline: 15161807]
29. Pioli MR, Ritter AMV, de Faria AP, Modolo R. White coat syndrome and its variations: differences and clinical impact. Integr Blood Press Control 2018;11:73-79 [FREE Full text] [doi: 10.2147/IBPC.S152761] [Medline: 30519088]
30. Ames G, Heckman M, Diehl N, Grothe K, Clark M. Further statistical and clinical validity for the weight efficacy lifestyle questionnaire-short form. Eat Behav 2015;18:115-119. [doi: 10.1016/j.eatbeh.2015.05.003] [Medline: 26042918]
31. Hunot C, Fildes A, Croker H, Llewellyn CH, Wardle J, Beeken RJ. Appetitive traits and relationships with BMI in adults: development of the adult eating behaviour questionnaire. Appetite 2016;105:356-363 [FREE Full text] [doi: 10.1016/j.appet.2016.05.024] [Medline: 27215837]
32. Mayhew M, Leo MC, Vollmer WM, DeBar LL, Kiernan M. Interactive group-based orientation sessions: A method to improve adherence and retention in pragmatic clinical trials. Contemp Clin Trials Commun 2020;17:100527 [FREE Full text] [doi: 10.1016/j.conctc.2020.100527] [Medline: 32083219]
33. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]
34. Yancy WS, Westman EC, McDuffie JR, Grambow SC, Jeffreys AS, Bolton J, et al. A randomized trial of a low-carbohydrate diet vs orlistat plus a low-fat diet for weight loss. Arch Intern Med 2010;170(2):136-145. [doi: 10.1001/archinternmed.2009.492] [Medline: 20101008]
35. Sepah SC, Jiang L, Ellis RJ, McDermott K, Peters AL. Engagement and outcomes in a digital diabetes prevention program: 3-year update. BMJ Open Diabetes Res Care 2017;5(1):e000422 [FREE Full text] [doi: 10.1136/bmjdrc-2017-000422] [Medline: 28948027]
36. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the diabetes prevention program outcomes study. Lancet 2009;374(9702):1677-1686 [FREE Full text] [doi: 10.1016/S0140-6736(09)61457-4] [Medline: 19878986]

Abbreviations

CONSORT: Consolidated Standards of Reporting Trials
HbA_{1c}: glycosylated hemoglobin
HIPAA: Health Insurance Portability and Accountability Act
PI: principal investigator
REDCap: Research Electronic Data Capture
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

XSL•FO
RenderX

Edited by A Mavragani; submitted 23.Jul.2025; this is a non–peer-reviewed article;accepted 04.Nov.2025; published 28.Jan.2026.

Please cite as:
Hansen A, Raymond K, Simon A, Kim S, Bayandorian H, Marriott D, Saslow LR
Very Low–Carbohydrate Breakfast Intervention for Adults with Type 2 Diabetes and Persistent Hyperglycemia: Protocol for a Digital, Nonrandomized Pre-Post Study
JMIR Res Protoc 2026;15:e81041
URL: https://www.researchprotocols.org/2026/1/e81041
doi:10.2196/81041
PMID:

©Annika Hansen, Kaitlyn Raymond, Alexander Simon, Sarah Kim, Hovig Bayandorian, Deanna Marriott, Laura R. Saslow. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 28.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Real-Time Exposure to Intersectional Minority Stressors and Alcohol Use: Protocol for an Ecological Momentary Assessment Study With Latinx and Non-Latinx Sexual Minority Youth

Robert Rosales[1*], MSW, PhD; Suzanne M Colby[2*], PhD; Kristina M Jackson[3*], PhD; Christina S Lee[4*], PhD; Jacob John van den Berg[5*], PhD; Madelyn Clancy[6*], BA; Ethan H Mereish[7*], PhD; Robert Miranda Jr[2*], ABPP, PhD

[1]Center for Alcohol and Addiction Studies; Center for Health Promotion and Health Equity, Department of Behavioral and Social Sciences, Brown University, Providence, RI, United States

[2]Center for Alcohol and Addiction Studies, Department of Behavioral and Social Sciences; Department of Psychiatry and Human Behavior, Brown University, Providence, RI, United States

[3]Rutgers Robert Wood Johnson Medical School, Rutgers Addiction Research Center, Department of Psychiatry, Rutgers, The State University of New Jersey, Piscataway, NJ, United States

[4]School of Social Work, Boston University, Boston, MA, United States

[5]Department of Psychology and Education, College of Arts and Sciences, Norwich University, Northfield, VT, United States

[6]School of Public Health, Center for Alcohol and Addiction Studies, Brown University, Providence, RI, United States

[7]Department of Psychology, University of Maryland, College Park, College Park, MD, United States

[*]all authors contributed equally

Corresponding Author:
Robert Rosales, MSW, PhD
Center for Alcohol and Addiction Studies; Center for Health Promotion and Health Equity
Department of Behavioral and Social Sciences
Brown University
121 South Main Street
Providence, RI, 02912
United States
Phone: 1 4018636737
Email: robert_rosales@brown.edu

Abstract

Background: Sexual minority youth (SMY) are significantly more likely to use alcohol compared with their heterosexual peers. Recent national data also suggest a turning point in alcohol use disparities: Latinx youth now report higher alcohol use than non-Latinx youth. Despite this, little is known about the social context and reasons why Latinx SMY may engage in alcohol use.

Objective: This manuscript describes the protocol for a study designed to assess real-time exposure to minority stressors and protective factors, and their relationship to alcohol use among Latinx and non-Latinx White SMY.

Methods: The project is being conducted in 3 phases with a combined sample of approximately 140 participants. Phase 1 (completed) involved cognitive interviews with 23 SMY participants, which refined and adapted survey measures to ensure cultural and developmental appropriateness for the next study phases. Phase 2 was a pilot ecological momentary assessment (EMA) survey with 20 participants to evaluate feasibility, acceptability, and compliance (completed). Phase 3 will recruit approximately 100 SMY aged 15-19 years, with equal representation of Latinx and non-Latinx White SMY. Participants will complete a baseline survey and repeated EMA surveys to capture daily experiences of stressors, protective factors, and alcohol use.

Results: Results for the study sample, recruitment (between April 2022 and November 2023), and challenges confronted are presented for Phase 1. Findings showed that the study sample included 23 Latinx and non-Latinx SMY, split almost in half by ethnicity. Participants were mostly female and affluent. Recruitment efforts showed that certain flyers and locations (eg, Facebook/Instagram) performed better at recruiting this sample. We present issues faced with screening out ineligible participants and bots, recruiting participants assigned male at birth, recruiting 15- to 17-year-old participants, and building overall trust with this population. Results from the rest of the data in this study will be analyzed and disseminated through peer-reviewed scientific journals.

XSL•FO
RenderX

Conclusions: This study will provide novel data on the real-time contexts of alcohol use among SMY with particular attention to Latinx youth, an understudied and marginalized population. By identifying stress and protective mechanisms linked to alcohol use, findings can inform tailored prevention and intervention strategies. Furthermore, the protocol offers a replicable framework for future EMA research on intersectionality, minority stress, and alcohol use among diverse SMY populations.

International Registered Report Identifier (IRRID): DERR1-10.2196/87201

(JMIR Res Protoc 2026;15:e87201) doi:10.2196/87201

KEYWORDS

Latinx; sexual minority youth; alcohol; minority stress; intersectionality; ecological momentary assessment

Introduction

Overview

Underage alcohol use is a significant public health concern in the United States, with particularly heightened risks among sexual minority youth (SMY). Data from the 2021 Youth Risk Behavior Surveillance Survey, a nationally representative survey of middle and high school youth, show that nearly half of adolescents have consumed alcohol in their lifetime, and approximately 30% drank in the past 30 days [1]. SMY are at higher risk of alcohol use and related problems compared with heterosexual peers [2-4]. SMY also tend to initiate drinking at earlier ages, and more frequently progress to alcohol use disorders [5]. These disparities may be partly attributable to the stress associated with heterosexism and other forms of discrimination [6].

Data from the Monitoring the Future study suggest that Latinx youth may now be more likely than non-Latinx White youth to consume alcohol and engage in high-risk drinking behaviors [7,8]. Longitudinal national data further suggest that discrimination may contribute to this increase in alcohol use among Latinx youth [9]. Despite these concerning trends, a recent review emphasized substantial gaps in the literature, noting that little is currently known about the risk and protective factors predicting substance use among Latinx SMY [10].

Minority stress theory (MST) posits that sexual minority individuals may use alcohol as a means of coping with negative affect arising from experiences of oppression-based stress [11-16]. MST further emphasizes the importance of social context, noting that the setting in which minority stress occurs—such as at school or at home—can be a critical predictor of substance use [17]. However, evidence on alcohol, tobacco, and other drugs (ATOD) use among Latinx SMY, compared with White SMY, remains mixed. Some studies report higher rates of ATOD use among Latinx SMY [18,19], whereas others indicate lower rates relative to White SMY [20]. To explain these inconsistencies, MST offers 2 competing hypotheses. The resilience hypothesis posits that exposure to racism may foster resilience to heterosexism, potentially buffering Latinx SMY from substance use risk compared with White SMY [6]. Support for this hypothesis comes from research on Latinx youth more broadly, which highlights unique sociocultural resilience factors. For example, ethnic identity may serve as a protective resource, mitigating the additive stress associated with discrimination based on cultural values and language [21]. Aligned with Intersectional Theory, the double jeopardy hypothesis suggests that Latinx SMY may face an elevated risk of ATOD misuse due to the compounded stress of occupying both sexual and ethno-racial minority statuses [6]. However, research suggests that the effects of minority stress on ATOD use among this population can also be explained by momentary changes in affect.

Affect and Alcohol Use

Affect refers to subjective emotional states, typically categorized as positive or negative [22,23]. Although findings are mixed, prior research has demonstrated that negative affect is correlated with greater substance use, whereas positive affect tends to be associated with lower levels of use [24]. One proposed mechanism is that individuals may use substances to reduce or regulate negative emotions [25]. Negative emotional experiences have been identified as risk factors for substance use among minority populations, including sexual minority and Latinx youth [4,26]. Such negative emotions are often linked to minority-related stressors, including the challenges of coming out and experiences of discrimination [4,27]. Importantly, recent ecological momentary assessment (EMA) research has shown that negative affect can mediate the relationship between minority stress and substance use (specifically tobacco use) among SMY [28]. However, considerably less is known about the role of ethnicity in these processes, including whether there are sociocultural protective factors that may buffer the impact of minority stress on ATOD use.

Study Objective

This paper outlines the protocol for a study designed to examine the relationships among intersectional minority stress, protective factors, affect (both positive and negative), and alcohol use in the natural environment between Latinx and non-Latinx White SMY (ages 15-19 years). Although White SMY experience similar experiences with minority stress for their sexual orientation, Latinx SMY may face additional stress because of having to deal with cultural differences, language discrimination, and xenophobia in the United States. High school students aged 15 to 19 years are at a transition age where they are developing their identity and forming an independent relationship with their parents. This independence also puts them at greater risk of being influenced by their peers, who may discriminate against them, before they have developed all their coping mechanisms [29].

The overarching goal is to address gaps in the literature on alcohol use risk within this population and to identify protective

XSL•FO
RenderX

mechanisms that may be leveraged in intervention [10]. This study protocol consists of three phases: (1) a qualitative phase to explore minority stress and alcohol use and to refine/adapt measures for the EMA phase, (2) a pilot study with follow-up interviews to assess the feasibility and acceptability of the survey, and (3) a 21-day EMA protocol. Participants in both the pilot and full EMA phases (2 and 3) also complete an extensive baseline survey to assess control and moderating variables.

EMA is a data collection approach that captures experiences as they occur, ideally reducing recall bias and enhancing ecological validity [30]. In this study, EMA is used to address the overarching research aim of examining the additional stress burden experienced by Latinx SMY due to daily experiences of minority stress, when compared with White SMY. EMA methods make it possible to assess attitudes and behaviors related to ATOD use in real-time, including momentary affect, alcohol use episodes, and contextual risk and protective factors as youth encounter minority stress throughout the day [31,32]. The repeated measures design of EMA allows us to evaluate whether minority stress and affect are contemporaneously associated with alcohol use (eg, within the same day or week). In addition, the contextual risk and protective factors at the individual, family, and environmental levels are assessed in real time. Participants report when, where, and with whom minority stress occurred, along with their perceptions of whether these experiences are related to alcohol use.

This protocol paper presents (1) the study methods for all 3 phases, including design, procedures, compensation, measures, and analytic approach; (2) sample description and lessons learned from the first phase of the study related to recruitment strategies; and (3) the strategies used to address challenges during implementation of Phase 1. The purpose of providing this information is to offer a roadmap that may help future researchers navigate the complexities of conducting research with this population more efficiently and with fewer obstacles.

Methods

Design

The study protocol was organized into 3 phases. The targeted combined sample of 140 Latinx and non-Latinx White SMY ensures equal representation of Latinx and non-Latinx White participants in each phase. The target sample of 140 participants is not powered to find between-person differences in alcohol use. Thus, this study will provide the feasibility and acceptability of this protocol for future fully powered EMA studies on this topic. It will also serve as preliminary findings showing potential signals for between-person and within-person differences in alcohol use. Across all phases, participants are recruited according to the same study inclusion and exclusion criteria. Eligibility criteria are (1) self-identify as a sexual minority with one or more sexual orientation dimensions (same-sex behavior, attraction, or identity) [33]; (2) identify as either Hispanic/Latinx (any race) or non-Hispanic/Latinx White; (3) between the ages of 15 and 19 years; (4) report at least one minority stressor (ethnic/racial or sexual, such as victimization or discrimination) within the past 30 days; (5) report alcohol use within the past 30 days; (6) speak English or Spanish; (7) currently reside in the United States; and (8) have access to a phone or other device capable of completing daily surveys and interviews.

Ethical Considerations

All study procedures were reviewed and approved by the Brown University Institutional Review Board (IRB; protocol numbers 556 (Phase 1), 383 (Phases 2 and 3).

Across all study phases, informed consent (or assent for minors) was obtained either by having participants read the consent form themselves or by having a research team member read it aloud. The consent process was designed to ensure comprehension and voluntariness. The purpose and procedures of the study were presented in clear, accessible language, available in both English and Spanish. Potential participants were given adequate time to consider their decision, had the opportunity to ask questions, and were informed that participation was voluntary. Because all participants were required to have functional proficiency in English or Spanish, translation services were not necessary. The principal investigator and most research team members were fluent in both languages and able to conduct the process bilingually. All participants were provided with contact information for study staff and encouraged to reach out with any questions. It was emphasized that all survey information would remain confidential.

Prior research has documented the potential for psychological or physical harm to youth who disclose their sexual orientation to parents [34,35]. In addition, studies indicate that many SMY decline to participate in research requiring parental consent, thereby increasing self-selection bias [36-38]. Such bias threatens the validity of findings and may limit the applicability of recommendations for this population. The Institute of Medicine has similarly noted that “certain studies that are important to adolescent health and well-being will not be feasible without such a waiver [of parental consent]” [39]. Given these considerations—and the minimal risks associated with participation—the IRB determined that parental consent was not appropriate for this study. The waiver of parental consent was granted in accordance with Brown University’s IRB requirements, which stipulate that (1) the research involves no more than minimal risk to participants, (2) the waiver does not adversely affect participants’ rights or welfare, and (3) the research could not practicably be conducted without the waiver. The IRB approved a waiver of documentation of consent, allowing participants aged 18 years and older to provide verbal consent and participants aged 15-17 years to provide verbal assent without parental permission.

Since this is not a clinical trial, the IRB also approved that we do not access EMA data in real time. The data are only reviewed at the end of the study or when all data are collected. Given that the focus is on underage alcohol use in a marginalized population, we also provide participants with a behavioral health resource document that provides them with contacts for national and local behavioral health organizations that can be contacted through phone call or text during an emergency. All data was deidentified before the data analytic process.

XSL•FO
RenderX

Recruitment

The study used the same recruitment strategies across all phases (target sample size=140). Recruitment efforts were multi-pronged. First, printed materials (eg, posters and brochures) were distributed, with organizational approval, to community-based settings nationwide (eg, Los Angeles, California; Boston, Massachusetts; Providence, Rhode Island; and Phoenix, Arizona), including high schools, colleges and universities, health centers, libraries, Latinx-serving organizations, and organizations serving SMY youth. Second, study staff continued outreach at community events such as Pride celebrations and Latinx cultural gatherings in the northeastern US, as well as in high-traffic areas (eg, bus stops, convenience stores, markets, and laundromats). Third, targeted digital recruitment was implemented through targeted ads (by age and location) on social media platforms (eg, Facebook and Instagram) to enhance reach to SMY populations throughout all study phases. All recruitment materials were available in both Spanish and English.

All recruitment flyers contained standardized information. Each flyer included (1) a call to study action (eg, “Help your community!”), (2) eligibility criteria regarding age (eg, “Are you a teenager 15-19 years old?”), (3) a general study overview (eg, “We invite you to join a voluntary research study on stress and health”), (4) a description of study procedures (eg, “The study is fully remote and will involve two interviews and completion of phone-based surveys over three weeks”), (5) information about compensation (eg, “You can receive up to $200 for participating”), (6) instructions for next steps (eg, “To check eligibility, scan the QR code or complete a brief survey”), and (7) study team identifiers (eg, Brown University logo, study phone number, email address, website, and IRB protocol number).

Multiple flyer versions were developed to reach distinct participant groups. Variations were based on recruitment modality. Flyers distributed in person were rectangular and included a QR code to direct participants to the eligibility screening survey. In contrast, digital flyers were square to meet Facebook and Instagram advertisement specifications; these did not include a QR code or link, as clicking the advertisement automatically redirected users to the screening survey. Flyers also varied according to the intended target population. For instance, materials aimed at Latinx SMY featured images of individuals visually identifiable as people of color and included the term “Latino/Hispanic/Latinx”. Additional variants included or excluded the term “LGBTQ+” (lesbian, gay, bisexual, transgender, and queer or questioning) to reach both youth who explicitly identify as LGBTQ+ and those who do not adopt such labels but report same-sex attractions or behaviors. Examples of flyers used in Phase 1 are presented in Figure 1, with additional examples from Phases 2 and 3 shown in Figure 2.

XSL•FO
RenderX

Figure 1. Examples of phase 1 study flyers. LGBTQ+: lesbian, gay, bisexual, transgender, and queer or questioning.



English flyer with LGBTQ+ mentioned



English flyer with Latinx mentioned



English flyer without Latinx or LGBTQ+ mentioned



Spanish flyer without Latinx or LGBTQ+ mentioned

Procedures

Screening Process

Interested individuals were redirected to an initial screening survey administered via Qualtrics. The screener assessed age, English/Spanish fluency, sexual identity, sexual behavior, sexual attraction, recent experiences with minority stressors, current alcohol use, ethnicity (Latinx vs non-Latinx White), race, and access to a phone or device with internet capability. The age item was revised from “Are you between 15 and 19 years of age?” (response categories: Yes/No) to “What is your age?” (open-ended response) after it was observed that some respondents older than 19 selected “Yes” to the original question to gain study access. The open-ended format was intended to reduce this misrepresentation by making eligibility criteria less transparent.

Participants were also asked to report their sex assigned at birth to facilitate a balance between those assigned male and those assigned female at birth. Beginning in Phases 2 and 3, an additional item was added to confirm access to an electronic device capable of supporting surveys and interviews. Finally, the screener included an item asking participants to identify where they saw or heard about the study, which provided information to refine and improve recruitment strategies. Screening questions and corresponding inclusion criteria are presented in Table 1.

XSL•FO
RenderX

Table 1. Qualtrics screening questions.

Inclusion criteria	Questions	Response categories	Inclusion response
Age			
Original	Are you between 15 and 19 years of age?	(1) Yes or (2) No	Yes
Revised	What is your age?	Open-ended	15-19 years old
English or Spanish fluency	Do you speak fluent English or Spanish?	(1) Yes or (2) No	Yes
Recently experienced discrimination	Have you been discriminated against in the last 30 days because of your sexual or racial identity?	(1) Yes or (2) No	Yes
Current alcohol use	Have you had an alcoholic beverage in the last 30 days?	(1) Yes or (2) No	Yes
Race and ethnicity			Either (1) Latinx and any racial category or (2) not Latinx and White
Hispanic or Latinx	Do you identify as Hispanic/Latino/Latinx or were you or your parents born in Latin America?	(1) Yes or (2) No	
Race	What is your Race?	(1) White, (2) Black/African American, (3) Asian, (4) Native Hawaiian or Other Pacific Islander, (5) American Indian or Alaska Native, and (6) Other	
Dimensions of sexual orientation			Only participants who respond in the following way are excluded (1) Straight/heterosexual (sexual minority identity), (2) Person of the same sex (sexual minority behavior), and (3) Person of the same sex (sexual minority attraction)
Sexual minority identity	Do you identify as:	(1) Straight/heterosexual, (2) Gay, lesbian, bisexual, queer, (3) Other, (4) Not sure, and (5) Don't know	
Sexual minority behavior	During your life, with whom have you had sexual contact?	(1) I have never had sexual contact, (2) Person of the same sex, (3) Person of the opposite sex, and (4) Both sexes	
Sexual minority attraction	Who are you sexually attracted to?	(1) I am not sexually attracted to anyone, (2) Person of the same sex, (3) Person of the opposite sex, and (4) Both sexes	
Ability to complete electronic surveys	Do you have a phone, computer, or device that you can use to respond to daily surveys and do video interviews?	(1) Yes or (2) No	Yes
Recruitment location	Where did you find out about this study?	(1) Facebook, (2) TikTok, (3) Instagram, (4) Flyer at a clinic, (5) Flyer at a bus station, (6) Flyer at school, (7) A friend, (8) A teacher, (9) My clinician (example: social worker, psychologist), (10) Someone else (Please tell us who):, and (11) I saw a flyer somewhere else (Please tell us where):	N/A[a]
Sex assigned at birth	What sex were you assigned at birth (on your original birth certificate)?	(1) Female, (2) Intersex, (3) Male, and (4) Prefer not to answer	Respondents assigned a certain sex at birth will be excluded or brought into the study when we are able to get a better balance of participants' sex assigned at birth

[a]Not applicable.

XSL•FO
RenderX

Challenges: Screening Out Bots

To further ensure that ineligible individuals did not enter the study, we implemented an additional screening interview. Initially, eligibility was determined solely through the Qualtrics screener. However, the first author observed indicators of ineligible participation. For example, some respondents appeared to be located outside of the United States, as suggested by discrepancies such as dark video backgrounds during morning hours in the United States and non-US time zones displayed in the Calendly scheduling system. In addition, as noted above, some participants self-reported ages exceeding the upper cutoff of 19 years.

Given such recruitment challenges, a structured screening interview was added to confirm eligibility. During this process, participants were asked to provide their zip code, state, birthdate, and age, and responses were cross-verified for consistency (eg, confirming that a reported zip code matched the reported state). The interviewer was authorized to determine eligibility by integrating all available data, including participant responses, time-of-day consistency, and indicators of potential fabrication (eg, evidence of online searching during questioning). We believe this procedure successfully excluded nearly all ineligible participants. Because a substantial number of screening interviews continued to involve ineligible respondents, we retained this additional verification step throughout the study.

Prior research indicates that open-ended questions, such as those asking about age, can be useful in detecting discrepancies that help identify bots and fraudulent participants [40,41]. Accordingly, we incorporated 4 additional items into the screening process to enhance data integrity: “What is your age?” “What is your birthday?” “Please share anything that will help prepare for our meeting” and “How did you hear about this study?” Responses to these items were reviewed and compared with the digital screener to verify consistency before scheduling the screening interview.

For further quality control, participant IP addresses were collected in Qualtrics. Cases were excluded when responses were inconsistent (eg, evidence of bot activity or mass-produced responses), when multiple entries were submitted by the same users, or when individuals appeared to be intentionally misrepresenting themselves (eg, attempting to qualify for an ineligible survey, providing contradictory information). In such instances, participants were not enrolled in the study, and compensation was withheld.

Phase 1: Qualitative Phase to Refine and Adapt Interview

Phase 1 consisted of cognitive interviews with approximately 20 SMY (half Latinx; half non-Latinx White) to refine and adapt survey items, reduce participant burden in the subsequent EMA survey, and ensure cultural and contextual appropriateness for SMY. Each interview lasted approximately 2 hours and was conducted via Zoom. The interview had 2 components. First, participants engaged in a think-aloud exercise, a method designed to capture participants’ real-time cognitive processes while completing survey items [42-44]. To familiarize participants with the approach, the session began with a practice question unrelated to the study (eg, “How many windows are there in your house, apartment, or dorm?”) [45]. Participants were then asked to verbalize their thoughts while completing the draft survey planned for Phases 2 and 3 [46]. These think-aloud responses were used iteratively to identify comprehension issues, detect errors, and refine survey items, consistent with best-practice recommendations [42]. Following the think-aloud portion, participants responded to a series of open-ended questions regarding (1) general feedback on survey items, (2) impressions of recruitment flyers, (3) alcohol use, (4) sociocultural risk and protective factors, (5) alcohol-related consequences, and (6) access to treatment.

Phase 2: Pilot Study With Follow-Up Interviews to Assess Survey Feasibility and Acceptability

During Phase 2, participants completed all study procedures remotely via Zoom. At the baseline session, participants were introduced to the study app on their personal mobile device and trained in the use of the EMA survey system. The baseline survey assessed demographics, covariates, and variables relevant to moderation analyses. At the beginning of the session, a trained research team member administered the Timeline Followback interview to capture alcohol and cannabis use over the previous 28 days [47]. Participants were then provided with a link to complete the remainder of the baseline survey. Surveys were completed while the research team member remained present to answer any questions or clarify item wording. Pilot testing indicated that the baseline interview required approximately 50 minutes to complete (range = 45 minutes to 1 hour). The Catalyst mobile app (MetricWire Inc), was used to administer EMA surveys in this study. Participants received brief training on how to download and install the app on their personal devices. Once installed, a study team member provided structured training on EMA procedures, including the types of prompts, survey frequency, procedures for problem reports, and options for participant-initiated reporting. Training also covered clarification of study terminology (eg, definitions of identity within the EMA survey, the meaning of a standard drink), using a live demonstration within the app.

Participants were asked to complete a final interview designed to inform refinements to the Phase 3 EMA survey methodology. During this interview, the Timeline Followback was re-administered to assess alcohol and cannabis use, allowing verification of whether any episodes had been missed in the EMA reporting. Participants also completed a mixed methods Qualtrics survey in which they provided feedback on the EMA survey, including any difficulties encountered or suggestions for improvement.

Phase 3: 21-Day EMA Protocol

Due to the pilot showing high feasibility and acceptability (manuscript expected in March 2026), the protocol for Phase 3 looks similar, with only changes to the final interview. Rather than reporting on the feasibility of the study, participants are asked to report on treatment adaptation needs, help-seeking, readiness to change, mobile health intervention preferences, and barriers to treatment. The aim of these questions is to provide preliminary findings on how to adapt an intervention for this population. In addition, because the feasibility and

XSL•FO
RenderX

acceptability are high for this EMA protocol, all Phase 2 EMA responses have been combined with Phase 3 responses.

EMA Survey Structure

Participants engage in up to 21 consecutive days of EMA. Surveys assess alcohol and other substance use, sexual activity, mental health, and exposures to stressors (including discrimination). Consistent with prior EMA protocols assessing substance use [28,48], participants complete up to 6 compensated surveys per day: one morning survey, one bedtime survey, and 4 randomly prompted surveys (referred to as Teen Electronic Diary [TED] reports). In addition, participants have the option to initiate a problem report (to indicate technical or content-related issues) and a begin drink report (to mark the onset of alcohol consumption). These optional reports are not required and do not carry compensation.

The EMA protocol includes 3 types of scheduled reports—morning, bedtime, and random—as well as optional self-initiated drinking reports. Figure 3 provides a description of the scheduling, prompts, and reminders for each type of these reports. Morning reports (ie, surveys completed by participants in the morning) assess sleep quality, unreported substance use from the previous day, and anticipated substance use. Bedtime reports assess social support and belongingness experienced during the day. Random reports (also called TED reports) assess real-time affect, substance use, and minority stress. Our random reports will be scheduled daily within 4 windows, and participants are instructed to complete each survey within 59 minutes of the first prompt. If a survey is still incomplete, SMS reminders will be sent (eg, “Hi, [participant First Name]. Please complete your TED Report ASAP. It will expire soon and your answers are important to us.”). Additional information about the timing of each report can be found in Figure 3. To minimize burden, participants are advised not to complete surveys in unsafe or inappropriate contexts (eg, while driving or during class). The final type of report, self-initiated drinking reports, allows participants to document the onset of alcohol use and associated experiences of minority stress in real time.

Figure 3. Description of morning, random, and bedtime report timing.

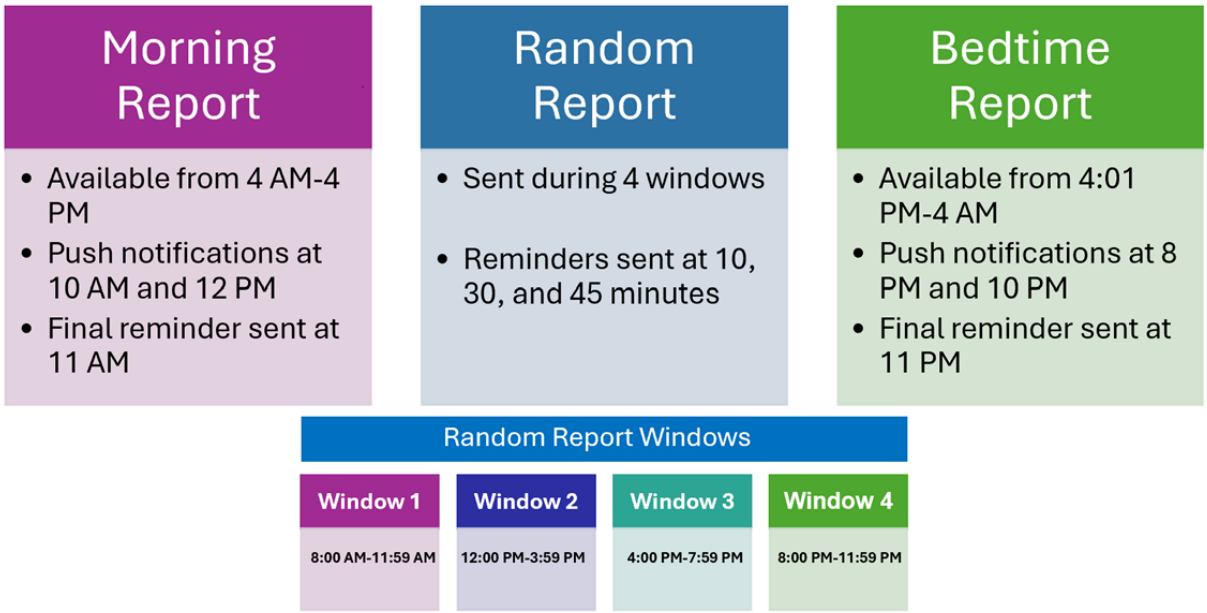

Compensation

Participant compensation will vary by study phase but will be consistently delivered via the participants’ preferred e-gift card option. A detailed breakdown is presented in Table 2. In Phase 1, participants received US $50 for completing the 2-hour cognitive interview. In Phase 2, participants earned up to US $200 for full study participation. This included US $40 for the 1.5-hour baseline interview and US $25 for the 30-minute final interview. To encourage EMA compliance, participants were compensated per survey rather than through unannounced bonuses, a strategy shown to enhance adherence in EMA studies [49]. Specifically, participants earned US $1 per completed survey and an additional US $3 incentive for completing at least 75% of surveys within a 7-day period, for a potential EMA total of US $135. EMA compensation was disbursed weekly, with lump-sum payments issued each Monday after participants’ first 7 days in the study. Phase 3 compensation follows the same structure as Phase 2.

XSL•FO
RenderX

Table 2. Compensation breakdown.

Study activity	Compensation (US $)	Total possible compensation (US $)
Phase 1		50
2-hour cognitive interview	50	
Phase 2		200
First interview	40	
21-day surveys on your phone	Up to 126 (1/survey)	
Incentive for completing 75% of a week	Up to 9 (3/week)	
Final interview	25	
Phase 3		175
First interview	40	
21-day surveys on your phone	Up to 126 (1/survey)	
Incentive for completing 75% of a week	Up to 9 (3/week)	

Measures

Phase 1: Qualitative Phase to Refine and Adapt Interview

During the item refinement portion of the qualitative interviews, participants responded to a battery of questions intended for use in the baseline and EMA portions of Phases 2 and 3. To align with the originally planned 30-day EMA period, all items were asked with a 30-day recall frame. First, participants completed the Everyday Identity Stress Scale (9 items), developed to assess daily experiences of identity-related stress among SMY [50] and subsequently adapted for momentary assessment in EMA studies [28]. Items included questions about distal discrimination (eg, “I was targeted or harassed because of my identity”) and attribution of experiences to specific identities (eg, “My race”). Second, participants completed the LGBTQ+ People of Color Microaggressions Scale (16 items), which measures 3 domains: LGBT racism (eg, “I felt misunderstood by White LGBTQ+ people”), People of Color heterosexism (eg, “I felt misunderstood by other Latino/Latinx/Hispanic people/person”), and LGBTQ+ relationship racism (eg, “I was rejected by potential dating or sexual partners because I am Latino/Hispanic”) [51]. Finally, participants were asked to respond to additional items identified as potential moderators or mediators, including victimization, social support, coping strategies, alcohol-related consequences, and drinking context. A complete list of items and response categories is provided in Table 3.

XSL•FO
RenderX

Table 3. Constructs measured in Phase 1 qualitative interviews.

Construct and questions	Response categories
Victimization	
1) Were you bullied or harassed by someone in the last 30 days?	1) No 2) Yes
2) If yes, how did it happen: (Check all that apply)	1) in-person contact 2) by telephone 3) by text message 4) via online interaction 5) through some other way
3) Who bullied or harassed you? (Check all that apply)	1) Friends I drink with 2) Friends I don't drink with 3) Classmates who are not friends 4) Someone online 5) Coworkers 6) Others
Social support	
1) After one of these negative experiences, someone expressed care/concern for me.	1) No 2) Yes
2) If yes, who? (Check all that apply)	1) Friends I drink with 2) Friends I don't drink with 3) Classmates who are not friends 4) Boy/girlfriend/intimate partner 5) Child/children 6) Mother 7) Father 8) Siblings, brothers, or sisters 9) Other relatives 10) Friend's parents 11) Coworkers 12) Others _____
Coping	
1) Did you use one or more of the following coping strategies to manage your emotions after experiencing discrimination in the last 30 days? (Check all that apply)	1) keep busy 2) socialize 3) think positively 4) do something good for self 5) calm self 6) find perspective 7) sit with feelings until they pass 8) use substances 9) get angry 10) ruminate 11) Other _____
Alcohol consequences	
1) Did you have any of the following problems in the last 30 days because of your drinking?	1) Legal problems (for example: related to citizenship) 2) School problems (for example: suspended or expelled from school) 3) Work problems (for example: being fired or decreased hours) 4) Relationship problems
Drinking context	

XSL•FO
RenderX

Construct and questions	Response categories
Prompt: Think about the last time you drank alcohol. Answer the following questions as if you were in that situation.	
1) Where were you? (Check all that apply)	1) Home (a) 2) Dorm (b) 3) Friend's home (c) 4) School (d) 5) Work (e) 6) At a party (f) 7) Restaurant (g) 8) Club or bar (h) 9) Outside in public place (i) 10) Inside in a public place (j) 11) Car, bus, other transportation (k) 12) Study visit (Brown) (l) 13) Elsewhere _____ (m)
2) Were you in a place where you usually drink?	1) No 2) Yes
3) Who was with you? (Check all that apply)	1) No one 2) Friends I drink with 3) Friends I don't drink with 4) Family member 5) Friend's family member 6) Coworkers 7) Others
4) How many people are with you?	Open ended
5) Were you with one of the following: (Check all that apply)	1) A Latino/Latinx/Hispanic person/people 2) An LGBTQ+[a] person/people 3) An LGBTQ+ Latino/Latinx/Hispanic person/people

[a]LGBTQ+: lesbian, gay, bisexual, transgender, and queer or questioning.

Phase 2 and 3 EMA Survey

The EMA protocol incorporated morning, random, and bedtime reports to assess minority stress, protective factors, and ATOD use in real time. The morning and random report survey items were adapted from an EMA battery previously validated for alcohol use [52]. Random reports assessed momentary affect, substance use, cravings, situational context (eg, location, companions, drinking environment), and experiences of minority stress. The bedtime report questionnaire was adapted from prior research examining minority stress, affect, protective factors, and tobacco use among SMY [53]. This measure also incorporated items assessing ethnicity as a protective factor (eg, I was involved in an activity, meeting, event, or overall community for my ethnic group today). Bedtime reports captured participants' daily experiences of minority stress and their responses to these experiences, including the use of coping strategies and the receipt of social support. Example questions for each report can be seen in Figure 4.

XSL•FO
RenderX

Figure 4. Examples of morning, random, and bedtime report questions. LGBTQ+: lesbian, gay, bisexual, transgender, and queer or questioning.



Analytic Approach

Phase 1: Qualitative Phase to Refine and Adapt Interview Analytic Plan

For Phase 1, descriptive statistics were used to evaluate whether survey measures require refinement (see items in Table 3; manuscript expected in March 2026). In addition, thematic analysis was conducted using NVivo (QSR International) software to identify themes emerging from the qualitative interviews' open-ended questions and to assess whether measures should be adapted using the think-aloud responses. Thematic analysis is a qualitative analysis technique for coding and deriving meaning from interview data [54]. Identified themes were connected to existing literature on minority stress and protective factors related to alcohol use. Based on these findings, the research team refined measures by incorporating items that reflect participants' lived experiences (manuscript expected in March 2026). Participants were also asked to provide feedback on their experiences completing the survey. Follow-up questions included the following: (1) "How was your experience with filling out the questionnaire?" (2) "How much time do you think it took for you to complete the questionnaire? How do you feel about that?" (3) "How would you feel about taking it 3-5 times a day? For 30 days?" (4) "How much would you want to get paid if you had to do it this many times?" (5) "Were there any questions that you did not understand?" (6) "Were there any questions that were not related to your experience?" and (7) "Was there a question that you were expecting that was not there?"

Phase 2: Pilot Study With Follow-Up Interviews to Assess Survey Feasibility and Acceptability Analytic Plan

In Phase 2, qualitative data from the final interview following the 21-day EMA pilot will be thematically analyzed to identify themes related to the feasibility and acceptability of the EMA protocol (analyses planned for January 2026, and manuscript expected in March 2026). These themes will be integrated with EMA response rate data to evaluate potential modifications aimed at enhancing protocol adherence and improving the design of the finalized EMA survey in Phase 3.

Phase 3: 21-Day EMA Protocol Analytic Plan

Multilevel modeling (MLM) is the standard analytic approach for EMA data (manuscript expected March 2026 [52,55,56]. In Phase 3, MLM will be conducted in Mplus to examine the relationship between minority stress and alcohol use, minority stress and sociocultural risk and protective factors, sociocultural risk and protective factors and alcohol use, and whether these sociocultural risk and protective factors influence the association between minority stress and alcohol use. Four hypotheses will guide the analyses: (1) minority stress experienced in the moment will predict increases in SMYs' alcohol use, (2) negative affect will mediate the association between minority stress and alcohol use, (3) ethnic identity (Latinx) will exacerbate the relationship between minority stress and alcohol use, and (4) protective factors (eg, social support, ethnic identity salience, and outness) will mitigate the relationship between momentary minority stress and alcohol use. The primary outcome of interest for phase 3 will be within-person alcohol use assessed in three separate ways: (1) in the moment, (2) the same day, and (3) on the weekends. Since we are not powered to assess between-person relationships, we also aimed to assess differences in between-person differences in these alcohol use rates.

All analyses will adjust for covariates that may function as potential confounders (ie, sex, age, and household income). A 2-level MLM will be estimated, in which survey responses (Level 1: L1) are nested within participants (Level 2: L2) [57]. L1 represents within-person effects (eg, momentary minority stress, affect, alcohol use), encompassing both event-level and random assessments and thus capturing within-subject variability. L2 represents between-person effects (eg, sex, ethnicity, or L1 variables aggregated across time), reflecting between-subject variability. To facilitate interpretation, L1 variables will be person-centered and L2 variables grand mean-centered. Compositional effects will be examined by aggregating the L1 variables to partition between-subject and within-subject variance. Minority stress responses from EMA surveys will be modeled as L1 predictors of the primary dependent variable, alcohol use. This approach will allow us to test whether slopes for alcohol use differ significantly between

XSL•FO
RenderX

participants who experience minority stress and those who do not.

Moderators and mediators will be incorporated into the base multilevel model to test additional hypotheses. To evaluate hypothesis 2—that negative affect mediates the relationship between minority stress and alcohol use—we will include EMA reports of affective state (L1) as a mediator. This analysis will proceed in 2 stages (planned for February 2026). First, we will examine contemporaneous associations, testing whether minority stress (independent variable) is associated with alcohol use (dependent variable) and negative affect (mediator) at the same survey time point. This will assess whether stress and alcohol use co-occur within an event, as affect may be immediately influenced by stress experiences (or vice versa). Second, we will examine lagged associations, testing whether minority stress or negative affect at time *t* predicts alcohol use at Time *t*+1 (ie, the subsequent EMA prompt, several hours later). This temporal ordering provides a stronger test of mediation by assessing whether earlier stress or affect predicts later alcohol use.

We will also estimate a lagged model to examine whether minority stress predicts alcohol use on the following day or weekend. Formal tests of multilevel mediation will be conducted using the procedures outlined by Zhang, Zyphur, and Preacher [58]. Given the relatively modest sample size, emphasis will be placed on effect sizes rather than significance tests when interpreting mediated effects. Consistent with hypothesis 3, ethnicity (White vs Latinx) will be included as a moderator in the base model to assess support for either the double jeopardy or resilience hypothesis. Specifically, we hypothesize that Latinx SMY who experience discrimination (double jeopardy) will report higher alcohol use relative to White SMY. Finally, additional moderators will be examined to evaluate protective influences on the minority stress-alcohol use relationship. These will include baseline characteristics (eg, coping strategies, ethnic identity, level of outness, and gender) and daily-level variables (eg, social support and belongingness).

Results

Overview

In this section, we will present descriptive findings on sample characteristics, recruitment sources, feedback from the Youth Advisory Board (YAB), and data from our digital recruitment platform (Meta Business). Together, these results provide insights into the barriers and facilitators of conducting research with this population and may inform the design of future studies.

Sample Characteristics

Phase 1 was completed with 23 participants, recruited between April 2022 and November 2023. Descriptive statistics for the sample are presented in Table 4. Thirteen (57%) participants identified as Latinx and 10 as non-Latinx White, which met our original goal of dividing participant ethnicity in half. The average age was 18.1 years, and the mean reported household income was US $112,449.50. A majority of participants were enrolled in college (17/23, 74%) rather than high school. Although most participants were assigned female sex at birth (18/23, 78%), the sample reflected considerable diversity in gender identity. Specifically, participants self-identified as gender fluid (n=1), gender queer (n=5), men (n=4), non-binary (n=5), transmen (n=5), and women (n=10).

We assessed both sexual identity and sexual attraction as dimensions of participants' sexual orientation. Regarding sexual identity, the largest proportion of participants (9/23, 39%) identified as bisexual. With respect to sexual attraction, the largest group (10/23, 44%) reported attraction to people of any or all genders.

Participants were also asked to report their national background. The majority (19/23, 82%) were born in the United States. Latinx participants were additionally asked to specify their national heritage. Among this subgroup, the largest proportions identified as Dominican (3/13, 23%) and Colombian (3/13, 23%).

XSL•FO
RenderX

Table 4. Phase 1 participant demographics.

Construct	Values
Ethnicity, n (%)	
Latinx	13 (57)
Non-Latinx White	10 (43)
Current education, n (%)	
In high school	6 (26)
In college	17 (74)
Sex assigned at birth, n (%)	
Female	18 (78)
Male	5 (22)
Gender identity[a], n (%)	
Gender fluid	1 (4)
Gender queer	5 (22)
Man	4 (17)
Non-binary	5 (22)
Transman	2 (9)
Woman	10 (44)
Sexual orientation[a], n (%)	
Asexual	1 (4)
Bisexual	9 (39)
Gay	4 (17)
Lesbian	5 (22)
Pansexual	3 (13)
Queer	1 (4)
Not listed	1 (4)
Sexual attraction[a], n (%)	
Both men and women	5 (22)
Men	4 (17)
People of any/all gender	10 (44)
Women	5 (22)
Non-binary people	1 (4)
Nativity, n (%)	
US born	19 (82)
Dominican Republic	2 (9)
Colombian	2 (9)
National heritage[a,b], n (%)	
Mexican	1 (8)
Dominican	3 (23)
Central American	1 (8)
Colombian	3 (23)
Cuban	1 (8)
Brazilian	1 (8)

XSL•FO
RenderX

Construct	Values
Other	1 (8)
Not reported	2 (16)
Age (years), mean (SD)	18.13 (1.29)
Household income, mean (SD)	112,449.50 (88,860.91)

[a]Numbers may exceed 23 because participants were allowed to report more than one response.

[b]Only asked to the Latinx participants.

Recruitment

Phase 1: Participant Location

During Phase 1, recruitment efforts reached (ie, number of participants who completed the digital screener) 896 potential participants. The majority were referred through Instagram (448/896, 50%), followed by Facebook (143/896, 16%), school-based flyers (81/896, 9%), friends (72/896, 8%), and other sources such as family, college email announcements, or community organizations (63/896, 7%). Due to confidentiality constraints, it was not possible to directly link screening questionnaires with the participants who ultimately enrolled in the study.

Youth Advisory Board Feedback

We also consulted Dr. Jackson's YAB on 2 occasions to gather feedback on recruitment strategies and flyer design. The YAB consisted of racially diverse, primarily female high school students aged 15-18 years. Members emphasized the importance of visually appealing recruitment materials, noting a preference for flyers with vibrant colors and professional design. Trust emerged as a central theme in their feedback. Specifically, they advised that the inclusion of the Brown University logo and a study email with a Brown University domain increased their confidence that the study was legitimate. Conversely, the use of a tinyurl link to the study screening survey reduced trust, as it did not reflect a university-affiliated domain. Although the tinyurl was ultimately used because it was shorter and easier to type than the original Qualtrics link, future studies may consider adopting university-branded short links to strengthen perceptions of credibility and trust among potential participants.

Meta Business Suite Flyer Data

An iterative process was used to evaluate and refine flyer effectiveness using analytics from Meta's Business Suite. This platform provides metrics, including results (link clicks), reach (number of unique individuals who viewed the advertisement), impressions (total number of times the advertisement was displayed), cost per click, and total spending. For example, between October 3 and October 10, one campaign generated 60 link clicks, reached 5922 individuals, and accumulated 10,407 impressions. The average cost per link click was US $0.73, with a total expenditure of US $43.52. To assess performance, we categorized advertisements with a cost per click below US $0.30 as high-performing, from US $0.30-US $0.40 as average-performing, and above US $0.40 as poor-performing. Advertisements exceeding US $0.40 per click were automatically replaced with alternative flyers. Initially, flyer performance was evaluated over a 7-day review period. However, this timeframe was extended to 30 days, as one week was insufficient to accurately assess campaign performance. Only a small number of flyers achieved high-performance status (ie, <US $0.30 per click). The strongest performer to date was the English language flyer for 18- to 19-year-olds (Figure 2), which consistently achieved a cost of approximately US $0.09 per click. We speculate that the flyer's imagery, which featured a more sensual depiction of male bodies, may have contributed to its effectiveness in attracting participants.

XSL•FO
RenderX

Figure 2. Examples of phase 2 and 3 study flyers. LGBTQ+: lesbian, gay, bisexual, transgender, and queer or questioning.



Findings from the initial 7-day performance reviews also informed decisions about which flyer to use for in-person recruitment. Flyers that demonstrated the greatest efficiency in attracting participants online were prioritized for physical distribution. For example, Flyer 1 (Figure 1) performed strongly in digital recruitment and received high ratings from the YAB. As a result, this flyer was frequently used in in-person recruitment, with minor modifications to tailor messaging to specific communities. For instance, flyers disseminated in Latinx neighborhoods included explicit references to Latinx/Hispanic identity to better align with the target audience.

Phase 2 and 3: Recruitment Status

Phase 2 recruitment started in May 2024 and ended in November 2024. Phase 3 began immediately after Phase 2 and has a projected end date of February 2026. As of the initial submission of this manuscript (November 6, 2025), 10,197 people have completed the screening survey, and 355 of them have scheduled a screening interview. From this, 260 participants were excluded (ie, not eligible [n=23], no-shows [n=20], suspected to be fraudulent [n=72], and other reasons [n=145]), and 95 participants were deemed eligible. Of the 95 eligible participants, 79 participants completed the baseline session and EMA protocol (15 participants were lost to contact, and one participant was no longer interested). Fifty-three participants have completed the final interview (with 4 participants still actively completing EMA). Data has been cleaned for 54 participants.

Challenges Confronted

The primary challenges encountered in this study centered on recruitment (Figure 5). Specifically, we faced difficulties with ineligible participants attempting to enter the study, underrepresentation of participants assigned male at birth, limited recruitment of younger adolescents, and the need to build trust within the community. Similar challenges have been documented in other research with racial/ethnic and sexual minority populations. In particular, the shift toward digital research during and after the COVID-19 pandemic has been associated with increased instances of participants misrepresenting their demographic information, as well as automated “bots” attempting to complete screening surveys, issues that disproportionately affect studies involving underrepresented groups [40,59]. Our experience demonstrated that a web-based screener alone was insufficient to exclude ineligible respondents. To address this, we implemented additional safeguards, including digital screening interviews, open-ended questions, the collection of general location information, and cross-verification of responses to detect

XSL•FO
RenderX

discrepancies. These procedures substantially reduced the likelihood of ineligible participants entering the study.

Figure 5. Challenges encountered during the study.



Second, consistent with other research [60], we encountered challenges in recruiting participants assigned male at birth. Despite initially implementing strategies intended to enhance male participation (eg, waiving parental consent) [61], recruitment of male participants—particularly during Phase 1—remained limited. We observed some improvement later by tailoring recruitment materials specifically for this group, including flyers and campaigns that featured male-presenting individuals and people of color in the visuals, particularly to engage Latinx SMY boys and young men [62]. Although this approach increased male participation, enrollment levels did not reach parity with female participants.

Thirdly, we faced challenges in recruiting adolescents aged 15-17 years during Phase 1. To increase enrollment among this age group, we implemented several strategies: (1) incorporating age-appropriate imagery into flyers, (2) expanding outreach and advertising efforts within high schools, and (3) launching a Facebook advertisement campaign, specifically targeting youth aged 15-17 years. These strategies yielded modest improvements, though recruitment in this age group remained more difficult than among older participants.

Finally, we confronted challenges related to building trust among this population, which may have limited participation among lower-income youth and 17-19-year-old youth not enrolled in college. The Phase 1 sample was more affluent and more highly educated than the general US population. One possible explanation is that the use of the Brown University logo enhanced credibility among college-attending and affluent youth, while offering little name recognition in lower-income communities. In some cases, institutional branding may even deter participation if potential participants fear being exploited by researchers. Partnering with nationally or locally recognized community organizations (eg, The Trevor Project at the national level or Youth Pride Inc. of Rhode Island locally) may therefore represent a more effective strategy for fostering trust and engagement among underrepresented subgroups in future research.

Discussion

Principal Findings

This paper describes the protocol and lessons learned from a 3-phase mixed methods study examining minority stress, risk and protective factors, and ATOD use in real time among Latinx and non-Latinx White SMY. More specifically, we predicted the following: in-the-moment minority stress will predict increases in SMY alcohol use; negative affect will mediate the association between minority stress and alcohol use; Latinx identity will exacerbate the relationship between minority stress and alcohol use; and protective factors will mitigate this relationship. The roadmap developed through this work has enabled the recruitment and enrollment of 101 participants as of November 2025 and provides a practical framework for future researchers seeking to recruit, verify eligibility, administer surveys and qualitative interviews, or implement EMA protocols with SMY populations. Elements of this roadmap may be adapted or adopted in whole or in part to enhance the efficiency of similar studies. We highlight examples of effective recruitment flyers and strategies, including benchmarks used to assess campaign performance. We also present insights from our YAB, whose feedback improved campaign effectiveness by shaping flyer appropriateness, design, and perceived trustworthiness. Importantly, this roadmap reflects nearly 5 years of iterative research and troubleshooting during which numerous challenges were encountered and addressed.

Building Off Prior Work

This study protocol builds on prior seminal work using EMA to assess the effects of minority stress on behavioral health. Our EMA protocol and survey were modeled after the study by Mereish et al [28] assessing the effects of minority stressors on nicotine use among SMY, which used a similar method (eg, inclusion criteria, EMA battery, and incentive structure). This study found that (1) minority stressors were associated with greater momentary nicotine craving and negative affect, and (2) nicotine craving and positive affect mediated the relationship between minority stressors and nicotine use [28]. We also built off the EMA study by Livingston et al [14] with sexual minority adults, which showed that minority stressors in the moment were associated with greater odds of momentary nicotine and substance use. These studies have all been informed by the Psychological Mediation Framework, which states that emotion regulation (eg, negative affect) mediates the relationship between distal stigma-related stressors (eg, SMY discrimination) and psychopathology (eg, substance use) [63]. Our study goes beyond this prior research to assess whether ethnicity and the social oppressive structures that ethnic minority people experience differentiate the experiences between Hispanic and non-Hispanic White SMY.

XSL•FO
RenderX

Strengths and Limitations

This study is not without limitations. First, our study is not fully powered to assess between-person differences in the effects of predictors on alcohol use. However, it may provide signals suggesting statistically significant relationships for a future fully powered EMA survey. Second, our samples of interest may be heterogeneous due to differences by country of origin, sexual orientation, and other factors related to ethnicity and sexual minority status. Findings for this study will be presented as general to the youth who identify as Hispanic and sexual minority or sexual minority. The findings will inform general experiences of these populations, and we will advocate for a larger study that can disentangle the heterogeneous experiences and studies on specific groups who may be at greatest risk for alcohol use (eg, Puerto Rican SMY). Third, our study has primarily enrolled participants who are in college, upper-middle-income, and assigned female at birth. Future studies should use other recruitment techniques to recruit greater numbers of lower-income and male participants. One such technique, which was not approved by the Brown University IRB, could involve using respondent-driven sampling to compensate low-income, non-college-educated, and male participants to recruit similar participants in their community. Respondent-driven sampling has been used to recruit large samples of participants from marginalized communities into similar research studies within weeks [64,65].

Future Directions

The lessons learned from this study provide a foundation for several ongoing and planned projects within our research group. These include manuscripts examining (1) differences in minority stress and alcohol use between Latinx SMY and non-Latinx SMY using Phase 1 qualitative interviews, (2) feasibility and acceptability of the EMA protocol from Phase 2, and (3) risk and protective factors associated with momentary alcohol and cannabis use among SMY. Beyond informing our own work, we hope this paper will serve as a roadmap for other researchers working with this population. More broadly, our aspiration is to contribute to the development of best practices for conducting research with intersectionally marginalized youth, including those who are both sexual minority and racial/ethnic minority.

Dissemination Plan

In addition to submitting the above-mentioned manuscripts to peer-reviewed academic journals, we plan to disseminate the findings of these studies to lay audiences. We have already used some of the lessons learned about behavioral health in this population to write an op-ed in the Miami Herald calling for the repeal of the colloquially termed “Don’t Say Gay” and “Stop Woke” acts [66]. We plan to submit similar op-eds using the findings from this study to support policy change for Hispanic and non-Hispanic SMY. In addition, we plan to write policy briefs that can be presented to policymakers at policy summits, in reports, and to support anti-oppressive legislation. Finally, we aim to have these findings readily available to parents of SMY. Thus, we plan to share our findings with community organizations and parenting groups with the aim of providing caregivers with skills on how to best support youth in this community. We also plan to submit our findings to media targeting this caregiver population, such as Spanish news media (eg, La Opinión) and major news media with articles focused on supporting this population (eg, ABC News’ Family section).

Acknowledgments

Generative artificial intelligence was not used in any part of the manuscript generation.

Funding

This research protocol was funded by the National Institute on Minority Health and Health Disparities (number 5K08MD015289-05; principal investigator RR). No external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this article. National Institutes of Health had no role in the study design, collection, analysis, or interpretation of the data; writing the manuscript; or the decision to submit the paper for publication.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary Statement from NIH Reviewers.
[PDF File (Adobe PDF File), 142 KB - resprot_v15i1e87201_app1.pdf]

References

1. Hoots BE, Li J, Hertz MF, Esser MB, Rico A, Zavala EY, et al. Alcohol and other substance use before and during the COVID-19 pandemic among high school students - youth risk behavior survey, United States, 2021. MMWR Suppl 2023;72(1):84-92 [FREE Full text] [doi: 10.15585/mmwr.su7201a10] [Medline: 37104552]

XSL•FO
RenderX

2. Marshal MP, Friedman MS, Stall R, King KM, Miles J, Gold MA, et al. Sexual orientation and adolescent substance use: a meta-analysis and methodological review. Addiction 2008;103(4):546-556 [FREE Full text] [doi: 10.1111/j.1360-0443.2008.02149.x] [Medline: 18339100]
3. Russell ST, Driscoll AK, Truong N. Adolescent same-sex romantic attractions and relationships: implications for substance use and abuse. Am J Public Health 2002;92(2):198-202. [doi: 10.2105/ajph.92.2.198] [Medline: 11818291]
4. Goldbach JT, Tanner-Smith EE, Bagwell M, Dunlap S. Minority stress and substance use in sexual minority adolescents: a meta-analysis. Prev Sci 2014;15(3):350-363. [doi: 10.1007/s11121-013-0393-7] [Medline: 23605479]
5. Corliss HL, Rosario M, Wypij D, Fisher LB, Austin SB. Sexual orientation disparities in longitudinal alcohol use patterns among adolescents: findings from the growing up today study. Arch Pediatr Adolesc Med 2008;162(11):1071-1078 [FREE Full text] [doi: 10.1001/archpedi.162.11.1071] [Medline: 18981356]
6. Meyer IH. Identity, stress, and resilience in lesbians, gay men, and bisexuals of color. Couns Psychol 2010;38(3):442-454 [FREE Full text] [doi: 10.1177/0011000009351601] [Medline: 24347674]
7. Miech RA, Johnston LD, Patrick ME, O'Malley PM, Bachman JG. Monitoring the Future National Survey Results on Drug Use, 1975-2023: Overview and Detailed Results for Secondary School Students. Michigan: Institute for Social Research; 2024.
8. Miech RA, Johnston LD, Patrick ME, O'Malley PM, Bachman JG, Schulenberg JE. Monitoring the Future National Survey Results on Drug Use, 1975-2022: Secondary School Students. Michigan: Institute for Social Research; 2023.
9. Rosales R, Veliz PT, Jardine J, Weigard AS, McCabe SE. Ethnic discrimination's role on increased substance susceptibility and use among U.S. youth. Am J Prev Med 2025;69(4):107956 [FREE Full text] [doi: 10.1016/j.amepre.2025.107956] [Medline: 40578581]
10. Rosales R, Zelaya DG, Moreno O, Figuereo V, Chavez SJ, Ordoñez S, et al. Latinx sexual minority adolescent substance use: state of the science and call for intersectional minority stressors and protective factors. Curr Addict Rep 2023;10(3):396-411 [FREE Full text] [doi: 10.1007/s40429-023-00503-5] [Medline: 38774111]
11. Brooks VR. Minority Stress and Lesbian Women. Lexington, Mass: Lexington Books; 1981.
12. Meyer IH. Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. Psychol Bull 2003;129(5):674-697 [FREE Full text] [doi: 10.1037/0033-2909.129.5.674] [Medline: 12956539]
13. Livingston NA, Christianson N, Cochran BN. Minority stress, psychological distress, and alcohol misuse among sexual minority young adults: a resiliency-based conditional process analysis. Addict Behav 2016;63:125-131. [doi: 10.1016/j.addbeh.2016.07.011] [Medline: 27471197]
14. Livingston NA, Flentje A, Heck NC, Szalda-Petree A, Cochran BN. Ecological momentary assessment of daily discrimination experiences and nicotine, alcohol, and drug use among sexual and gender minority individuals. J Consult Clin Psychol 2017;85(12):1131-1143 [FREE Full text] [doi: 10.1037/ccp0000252] [Medline: 29189029]
15. Mereish EH. Oppression-based stress and alcohol inequities among sexual and gender minority people: an intersectional multilevel framework. Alcohol Res 2024;44(1):05 [FREE Full text] [doi: 10.35946/arcr.v44.1.05] [Medline: 39246430]
16. Mereish EH, Abramson JR, Lee H, Watson RJ. Intersectional oppression-based stress, drinking to cope motives, and alcohol use and hazardous drinking among sexual and gender minority adolescents who are black, indigenous, and people of color. LGBT Health 2025;12(2):125-133 [FREE Full text] [doi: 10.1089/lgbt.2024.0023] [Medline: 38973422]
17. Goldbach JT, Gibbs JJ. A developmentally informed adaptation of minority stress for sexual minority adolescents. J Adolesc 2017;55(1):36-50 [FREE Full text] [doi: 10.1016/j.adolescence.2016.12.007] [Medline: 28033502]
18. Pollitt AM, Mallory AB, Fish JN. Homophobic bullying and sexual minority youth alcohol use: do sex and race/ethnicity matter? LGBT Health 2018;5(7):412-420 [FREE Full text] [doi: 10.1089/lgbt.2018.0031] [Medline: 30334685]
19. Bishop MD, Moran SA, Turpin RE, Aparicio EM, Mereish EH, Russell ST, et al. Binge drinking disparities by grade, race and ethnicity, sexual orientation, and gender identity. Pediatrics 2024;154(3):e2023063978 [FREE Full text] [doi: 10.1542/peds.2023-063978] [Medline: 39104357]
20. Kann L, Olsen EO, Kinchen S, Morris E, Wolitski RJ. HIV-related risk behaviors among male high school students who had sexual contact with males - 17 large urban school districts, United States, 2009-2013. MMWR Morb Mortal Wkly Rep 2016;65(5):106-109 [FREE Full text] [doi: 10.15585/mmwr.mm6505a2] [Medline: 26867146]
21. Torres L, Ong AD. A daily diary investigation of Latino ethnic identity, discrimination, and depression. Cultur Divers Ethnic Minor Psychol 2010;16(4):561-568 [FREE Full text] [doi: 10.1037/a0020652] [Medline: 21058819]
22. Watson D, Tellegen A. Toward a consensual structure of mood. Psychol Bull 1985;98(2):219-235 [FREE Full text] [doi: 10.1037//0033-2909.98.2.219] [Medline: 3901060]
23. Watson D, Clark LA. Emotions, moods, traits, and temperaments: conceptual distinctions and empirical findings. In: Ekman P, Davidson RJ, editors. The Nature of Emotion: Fundamental Questions. Oxford, United Kingdom: Oxford University Press, Incorporated; 1994:89-93.
24. Wills TA, Sandy JM, Shinar O, Yaeger A. Contributions of positive and negative affect to adolescent substance use: test of a bidimensional model in a longitudinal study. Psychol Addict Behav 1999;13(4):327-338 [FREE Full text] [doi: 10.1037//0893-164x.13.4.327]

XSL•FO
RenderX

25. Colder CR, Chassin L, Lee MR, Villalta IK. Developmental perspectives: affect and adolescent substance use. Subst Abuse Emotion 2010:109-135. [doi: 10.1037/12067-005]
26. Unger JB. Cultural influences on substance use among hispanic adolescents and young adults: findings from project RED. Child Dev Perspect 2014;8(1):48-53 [FREE Full text] [doi: 10.1111/cdep.12060] [Medline: 24729791]
27. Pilkington NW, D'Augelli AR. Victimization of lesbian, gay, and bisexual youth in community settings. J Community Psychol 1995;23(1):34-56. [doi: 10.1002/1520-6629(199501)23:1<34::aid-jcop2290230105>3.0.co;2-n]
28. Mereish EH, Treloar Padovano H, Parlette B, Miranda R. Momentary associations among minority stress, craving, affect, and nicotine use among sexual minority youth. J Clin Child Adolesc Psychol 2022;51(6):877-891 [FREE Full text] [doi: 10.1080/15374416.2022.2093208] [Medline: 35895311]
29. Graham S, Echols L. Race and ethnicity in peer relations research. In: Handbook of Peer Interactions, Relationships, and Groups. New York, NY: The Guilford Press; 2018:590-614.
30. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4(1):1-32 [FREE Full text] [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
31. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006;31(1):13-20 [FREE Full text] [Medline: 16496031]
32. Stone A, Shiffman S, Atienza A, Nebeling L. The Science of Real-Time Data Capture: Self-Reports in Health Research. England: Oxford University Press; 2007.
33. Laumann EO, Gagnon JH, Michael RT, Michaels S. The Social Organization of Sexuality: Sexual Practices in the United States. Chicago, IL: University of Chicago press; 2000.
34. Weir K. Studying adolescents without parents' consent: a new APA resolution supports mature minors' participation in research without parental permission. Am Psychol Assoc 2019;50(2):20 [FREE Full text]
35. Smith AU, Schwartz SJ. Waivers of parental consent for sexual minority youth. Account Res 2019;26(6):379-390. [doi: 10.1080/08989621.2019.1632200] [Medline: 31195827]
36. Flores D, McKinney R, Arscott J, Barroso J. Obtaining waivers of parental consent: a strategy endorsed by gay, bisexual, and queer adolescent males for health prevention research. Nurs Outlook 2018;66(2):138-148 [FREE Full text] [doi: 10.1016/j.outlook.2017.09.001] [Medline: 28993074]
37. Sims JP, Nolen C. "I Wouldn't Trust the Parents To 'Do No Harm' To a Queer Kid": rethinking parental permission requirements for youth participation in social science research. J Empir Res Hum Res Ethics 2021;16(1-2):35-45. [doi: 10.1177/1556264620983134] [Medline: 33356824]
38. Mustanski B. Ethical and regulatory issues with conducting sexuality research with LGBT adolescents: a call to action for a scientifically informed approach. Arch Sex Behav 2011;40(4):673-686. [doi: 10.1007/s10508-011-9745-1] [Medline: 21528402]
39. Field MJ, Behrman RE. Children IoMCoCRI. The necessity and challenges of clinical research involving children. In: Ethical conduct of clinical research involving children. Washington, DC: National Academies Press (US); 2004.
40. Bybee S, Cloyes K, Ellington L, Baucom B, Supiano K, Mooney K. Bots and nots: safeguarding online survey research with underrepresented and diverse populations. Psychol Sex 2022;13(4):901-911 [FREE Full text] [doi: 10.1080/19419899.2021.1936617] [Medline: 36439051]
41. Xu Y, Pace S, Kim J, Iachini A, King L, Harrison T. Threats to online surveys: recognizing, detecting, and preventing survey bots. Soc Work Res 2022;46(4):343-350. [doi: 10.1093/swr/svac023]
42. Honey M, Aspden T, Brackley K, Amende H, French A, Chan A. Using a" think-aloud" process to aid co-design in development of a medicine information needs assessment tool. Online J Nurs Inform 2022;26(1):7-11.
43. Mes MA, Chan AHY, Wileman V, Katzer CB, Goodbourn M, Towndrow S, et al. Patient involvement in questionnaire design: tackling response error and burden. J Pharm Policy Pract 2019;12(1):17 [FREE Full text] [doi: 10.1186/s40545-019-0175-0] [Medline: 31236279]
44. Wolcott MD, Lobczowski NG. Using cognitive interviews and think-aloud protocols to understand thought processes. Curr Pharm Teach Learn 2021;13(2):181-188. [doi: 10.1016/j.cptl.2020.09.005] [Medline: 33454077]
45. Mulhern B, Bansback N, Brazier J, Buckingham K, Cairns J, Devlin N, et al. Preparatory study for the revaluation of the EQ-5D tariff: methodology report. Health Technol Assess 2014;18(12):vii-xxvi [FREE Full text] [doi: 10.3310/hta18120] [Medline: 24568945]
46. Ryan M, Watson V, Entwistle V. Rationalising the 'irrational': a think aloud study of discrete choice experiment responses. Health Econ 2009;18(3):321-336. [doi: 10.1002/hec.1369] [Medline: 18651601]
47. Robinson SM, Sobell LC, Sobell MB, Leo GI. Reliability of the timeline followback for cocaine, cannabis, and cigarette use. Psychol Addict Behav 2014;28(1):154-162 [FREE Full text] [doi: 10.1037/a0030992] [Medline: 23276315]
48. Jackson K, Meisel M, Sokolovsky A, Chen K, Barnett N. Detecting and understanding social influence during drinking situations: protocol for a bluetooth-based sensor feasibility and acceptability study. JMIR Res Protoc 2024;13:e50650 [FREE Full text] [doi: 10.2196/50650] [Medline: 38842927]
49. Howard AL, Lamb M. Compliance trends in a 14-week ecological momentary assessment study of undergraduate alcohol drinkers. Assessment 2024;31(2):277-290 [FREE Full text] [doi: 10.1177/10731911231159937] [Medline: 36914966]

XSL•FO
RenderX

50. Mereish EH, Miranda R, Liu Y, Hawthorne DJ. A daily diary study of minority stress and negative and positive affect among racially diverse sexual minority adolescents. J Couns Psychol 2021;68(6):670-681 [FREE Full text] [doi: 10.1037/cou0000556] [Medline: 34166050]
51. Balsam KF, Molina Y, Beadnell B, Simoni J, Walters K. Measuring multiple minority stress: the LGBT people of color microaggressions scale. Cultur Divers Ethnic Minor Psychol 2011;17(2):163-174 [FREE Full text] [doi: 10.1037/a0023244] [Medline: 21604840]
52. Miranda R, MacKillop J, Treloar H, Blanchard A, Tidey JW, Swift RM, et al. Biobehavioral mechanisms of topiramate's effects on alcohol use: an investigation pairing laboratory and ecological momentary assessments. Addict Biol 2016;21(1):171-182 [FREE Full text] [doi: 10.1111/adb.12192] [Medline: 25353306]
53. Parnes JE, Le TP, Mereish EH, Miranda R. Daily associations between resilience factors, substance use, and affect among sexual minority youth. Psychol Addict Behav 2024;38(7):805-817 [FREE Full text] [doi: 10.1037/adb0000998] [Medline: 38358656]
54. Braun V, Clarke V, Terry G, Hayfield N. Handbook of research methods in health and social sciences. In: Thematic Analysis. Germany: Springer; 2019:843-860.
55. Treloar Padovano H, Miranda R. Using ecological momentary assessment to identify mechanisms of change: an application from a pharmacotherapy trial with adolescent cannabis users. J Stud Alcohol Drugs 2018;79(2):190-198 [FREE Full text] [doi: 10.15288/jsad.2018.79.190] [Medline: 29553345]
56. Bogers RP, Bolte JFB, Houtveen JH, Lebret E, van Strien RT, Schipper CMA, et al. Design of an ecological momentary assessment study of exposure to radiofrequency electromagnetic fields and non-specific physical symptoms. BMJ Open 2013;3(8):e002933 [FREE Full text] [doi: 10.1136/bmjopen-2013-002933] [Medline: 23988360]
57. Terhorst L, Beck KB, McKeon AB, Graham KM, Ye F, Shiffman S. Hierarchical linear modeling for analysis of ecological momentary assessment data in physical medicine and rehabilitation research. Am J Phys Med Rehabil 2017;96(8):596-599. [doi: 10.1097/PHM.0000000000000690] [Medline: 28092277]
58. Zhang Z, Zyphur MJ, Preacher KJ. Testing multilevel mediation using hierarchical linear models. Organ Res Methods 2008;12(4):695-719. [doi: 10.1177/1094428108327450]
59. Shaw TJ, Cascalheira CJ, Helminen EC, Brisbin CD, Jackson SD, Simone M, et al. Yes stormtrooper, these are the droids you are looking for: identifying and preliminarily evaluating bot and fraud detection strategies in online psychological research. Psychol Methods 2025;30(6):1294-1310 [FREE Full text] [doi: 10.1037/met0000724] [Medline: 40029298]
60. Borg DJ, Haritopoulou-Sinanidou M, Gabrovska P, Tseng HW, Honeyman D, Schweitzer D, et al. Barriers and facilitators for recruiting and retaining male participants into longitudinal health research: a systematic review. BMC Med Res Methodol 2024;24(1):46 [FREE Full text] [doi: 10.1186/s12874-024-02163-z] [Medline: 38389065]
61. Ybarra ML, Prescott TL, Phillips GL, Parsons JT, Bull SS, Mustanski B. Ethical considerations in recruiting online and implementing a text messaging-based HIV prevention program with gay, bisexual, and queer adolescent males. J Adolesc Health 2016;59(1):44-49 [FREE Full text] [doi: 10.1016/j.jadohealth.2016.03.020] [Medline: 27185621]
62. Sullivan PS, Khosropour CM, Luisi N, Amsden M, Coggia T, Wingood GM, et al. Bias in online recruitment and retention of racial and ethnic minority men who have sex with men. J Med Internet Res 2011;13(2):e38 [FREE Full text] [doi: 10.2196/jmir.1797] [Medline: 21571632]
63. Hatzenbuehler ML. How does sexual minority stigma "get under the skin"? a psychological mediation framework. Psychol Bull 2009;135(5):707-730 [FREE Full text] [doi: 10.1037/a0016441] [Medline: 19702379]
64. Garcini LM, Galvan T, Peña JM, Chen N, Klonoff E. Effectiveness of respondent-driven sampling for conducting health studies among undocumented immigrants at a time of heightened immigration enforcement. J Immigr Minor Health 2022;24(1):102-110 [FREE Full text] [doi: 10.1007/s10903-020-01112-4] [Medline: 33136244]
65. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. Soc Probl 1997;44(2):174-199 [FREE Full text] [doi: 10.1525/sp.1997.44.2.03x0221m]
66. Diarra B, Ordonez S, Cristino M, Rosales R, Color of Change. Don’t Ignore Effects of ‘Don’t say gay’ and ‘Stop WOKE’ on LGBTQ+ students of color. In: Miami Herald. Florida: Miami Herald; 2022.

Abbreviations

ATOD: alcohol, tobacco, and other drugs
EMA: ecological momentary assessment
IRB: Institutional Review Board
L1: Level 1
L2: Level 2
LGBTQ+: lesbian, gay, bisexual, transgender, and queer or questioning
MLM: multilevel modeling
MST: minority stress theory
SMY: sexual minority youth
TED: Teen Electronic Diary

XSL•FO
RenderX

YAB: Youth Advisory Board

Edited by J Sarvestan; submitted 06.Nov.2025; this is a non–peer-reviewed article;accepted 09.Dec.2025; published 30.Jan.2026.

Please cite as:
Rosales R, Colby SM, Jackson KM, Lee CS, van den Berg JJ, Clancy M, Mereish EH, Miranda Jr R
Real-Time Exposure to Intersectional Minority Stressors and Alcohol Use: Protocol for an Ecological Momentary Assessment Study With Latinx and Non-Latinx Sexual Minority Youth
JMIR Res Protoc 2026;15:e87201
URL: https://www.researchprotocols.org/2026/1/e87201
doi:10.2196/87201
PMID:

©Robert Rosales, Suzanne M Colby, Kristina M Jackson, Christina S Lee, Jacob John van den Berg, Madelyn Clancy, Ethan H Mereish, Robert Miranda Jr. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Cumulative Stress Burden and Association With DNA Methylation in Ethiopian American Immigrants: Protocol for a Community-Engaged, Biopsychosocial Study

Lara Stone[1], MA; Lili Ding[2,3], PhD; Meron Hirpa Kassa[4], MD; Beakal Amsalu[1], MPH; Bruktawit Getnet[1], BA; Hermella Eshete[1], MPH; Tesfaye B Mersha[2,5,6], PhD; Lisa M Vaughn[1,2], MA, MEd, PhD

[1]Division of Emergency Medicine, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[2]Department of Pediatrics, College of Medicine, University of Cincinnati, Cincinnati, OH, United States

[3]Division of Biostatistics/Epidemiology, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[4]The Christ Hospital Health Network, Cincinnati, OH, United States

[5]Division of Asthma, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States

[6]Precision Pulmonary Medicine Program, Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, United States

Corresponding Author:
Lisa M Vaughn, MA, MEd, PhD
Department of Pediatrics
College of Medicine
University of Cincinnati
3333 Burnet Ave
ML 2008
Cincinnati, OH, 45229
United States
Phone: 1 5136369424
Fax: 1 5136367967
Email: lisa.vaughn@cchmc.org

Related Article:

This is a corrected version. See correction statement: https://www.researchprotocols.org/2026/1/e95656

Abstract

Background: Immigrants experience significant health disparities, which are exacerbated by a heavy stress burden, which in turn affects the epigenetic profiles of the immune system, leading to chronic diseases. Cumulative stress burden for immigrants ranges from immigration-related stressors to unique psychosocial, environmental, and everyday challenges, all of which contribute to negative psychological and biological impacts on their health over the life-course. Although social and environmental conditions have been established as key factors driving disparities in health outcomes, the effects of stress and epigenetic change among immigrants remain poorly understood, impeding the development of novel and robust intervention approaches aimed at reducing health disparities. Epigenetic changes can act as surrogate markers for the stress effect. However, very few studies have examined epigenetic marks associated with stress among African immigrants. Ethiopians form one of the largest groups of African immigrants in the United States; yet, this is the first study of this kind among Ethiopian American immigrants.

Objective: This protocol aims to quantify cumulative stress burden and determine DNA methylation (DNAm) associated with stress among Ethiopian American immigrants.

Methods: Working with community partners, a community advisory board, and 3 community research coordinators, we will conduct a community-engaged research study of Ethiopian American immigrants. Data collection occurs at public events, church gatherings, and festivals. We use a multistation model composed of five stations through which participants rotate: (1) eligibility screening and consent, (2) stress-related questionnaires, (3) clinical measures, (4) health status and demographic questionnaires, and (5) saliva collection and incentive receipt. The measures used are: Stress of Immigration Survey, Boen's stress exposure measure, 10-item Perceived Stress Scale, Patient-Reported Outcomes Measurement Information System Global Health Scale (version 1.2), 8-item Patient Health Questionnaire, and 7-item Generalized Anxiety Disorder. Latent profile analysis, chi-square

XSL•FO
RenderX

tests, and logistic regression will be used. Saliva samples will be tested using genome-wide DNAm. With a subset of the sample, we will also conduct and thematically analyze qualitative interviews to understand additional experiences of stress and buffers to stress among participants.

Results: This study received National Institutes of Health R21 grant funding. Data collection began in October 2024 and will continue until November 2025 (currently at 89% completion). In November 2025, we will start data cleaning and analysis of questionnaires, clinical measures, and DNAm. We plan to complete data analyses and prepare scientific and community outputs by the Spring of 2026.

Conclusions: The study will provide new insights to address health disparities among growing immigrant populations in the United States, informing novel and robust intervention approaches to reduce chronic illness and associated sequelae for vulnerable populations. Findings from this study may highlight the need for policy change, such as providing more support in the form of infrastructure and social service agencies for immigrants.

International Registered Report Identifier (IRRID): DERR1-10.2196/85971

(JMIR Res Protoc 2026;15:e85971) doi:10.2196/85971

KEYWORDS

stress; immigrant health; DNA methylation; cumulative stress burden; chronic illness; Ethiopia

Introduction

Stress is ubiquitous for immigrants and ranges from daily stressors to unique psychosocial, environmental, and acculturative challenges contributing to adverse psychological and biological impacts on their health over the life-course [1-4]. Immigration-related stress includes acculturative stress compounded by language, financial, and isolation barriers [5] and the pressure of sending remittances to their native country [6,7]. Some immigrants have undergone challenging and traumatic journeys to the United States or, in their native country, experienced traumatic events (premigration stress) [6-9]. Additionally, an anti-immigrant sociopolitical environment in the new host country can exacerbate stress for immigrants [10-14]. Immigrants can also experience other forms of psychosocial stress and adversities from "everyday stressors" of family, work, acute stressful events, as well as social determinants that contribute to ill health (eg, poverty and community violence) [7,15-17]. In nontraditional migration areas like the Midwest, there is less infrastructure and social service agencies to support immigrants, less social acceptance of immigrants in general [18], and fewer ethnic enclaves, which often provide social support to immigrants [19], all of which can contribute to further exacerbation of stress. Taken together, immigrants may experience multiple types and sources of stress that accumulate to impact health risk and disease [20-24] and contribute to the decline of the initial immigrant health advantage [25,26]. It is well established that stress experiences and adversities have significant health consequences in a dose-response relationship such that cumulative exposure to stress, regardless of type or source, is associated with decrements and disparities in physical and mental health for vulnerable and minority groups such as immigrants [22-24,27,28]. To better understand immigrant-related stress processes, studies that assess immigrant stress burden multidimensionally.

Epigenetic changes are heritable; thus, they represent a key biological mechanism through which the effects of stress can be transmitted across generations [29,30]. This intergenerational transmission is further shaped by environmental factors that influence stress and, consequently, immune-related epigenetic changes [31,32]. Mounting evidence suggests that changes in the environment alter stress levels, which in turn affect the host epigenome profiles of the immune system, leading to health-related diseases [33]. There is growing evidence suggesting that immigration-related stress can leave measurable epigenetic signatures, particularly in DNA methylation (DNAm) patterns of stress-response genes [31,34]. For instance, psychosocial stressors in Latino immigrants have been associated with altered methylation profiles related to cardiometabolic risk [35]. Additionally, epigenetic modifications in DNAm have been recently documented with Syrian refugees experiencing war-related violence [31]. Hence, epigenetics is the missing link between migration-related experience and diseases outcome [36]. Several studies have supported epigenetics as a mechanism linking immigrant stress and cardiometabolic diseases (obesity, diabetes, hypertension, and stroke) and mental disorders (depression and anxiety disorders) [35-44]. Higher levels of cumulative family adversity, characterized by co-occurring stressors in a family context, were associated with DNAm at 7 sites, primarily in stress- and immune-related genes [45]. While stress has been suggested to induce epigenetic modifications resulting in racial differences in stress [46], studies are lacking in immigrant populations.

Within the United States, Black African immigrants experience disproportionate economic, health, and social disadvantage compared to other immigrant groups [47]. African immigrants have been referred to as "invisible immigrants" in the United States because they are rarely included in immigration policy discussions or research studies on immigrants [48,49]. Although not monolithic, Black African United States immigrants face the likelihood of compounded stress resulting from additional challenges associated with socioeconomic disadvantage, racism, and acculturation [50]. Thus, research is lacking on health care needs, priorities, and challenges experienced by this significantly growing segment of the US immigrant population, particularly in nontraditional migration areas. Currently, the second largest African immigrant population in the United States [48,51], Ethiopian American immigrants in the United States have steadily increased over the years. The United States is the most common destination for Ethiopian immigrants and a top source

XSL•FO
RenderX

for remittances to Ethiopia [52]. Despite the large number of Ethiopian immigrants living in the United States, limited health research has been conducted with this population [6].

In collaboration with our established Ethiopian American immigrant Community Advisory Board (CAB) and community partners, our team recently conducted a community health needs assessment of Ethiopian immigrants in Cincinnati and Columbus, OH, funded by the Cincinnati Academic Health Clinical Translational Science Awards Center [53]. Understanding the health beliefs, behaviors, needs, and priorities of Ethiopian American immigrants informs and supports this study. Importantly, we established an Ethiopian American immigrant CAB that met monthly during the needs assessment. Building on our previous work in immigrant health and this recent health needs assessment of local Ethiopian American immigrants, our goal is to better understand how psychological, social, environmental, and acculturative stressors affect the risk for chronic diseases through epigenomic mechanisms. Figure 1 outlines the conceptual model and purpose of this study, which is to (1) measure and describe cumulative stress burden and (2) determine the association between DNAm and stress among Ethiopian American immigrants. In secondary analyses, we will also examine the association of stress and specific chronic diseases—obesity, diabetes, hypertension, depression, and anxiety. This study directly addresses gaps in the literature investigating the multiple pathways through which cumulative stress becomes embodied and "gets under the skin" of Ethiopian American immigrants. This study will set the stage for action to address health disparities, including future interventions that are contextually grounded in the everyday realities of Ethiopian American immigrants' lives.

Figure 1. Conceptual model of the proposed research. EAI: Ethiopian American immigrant.

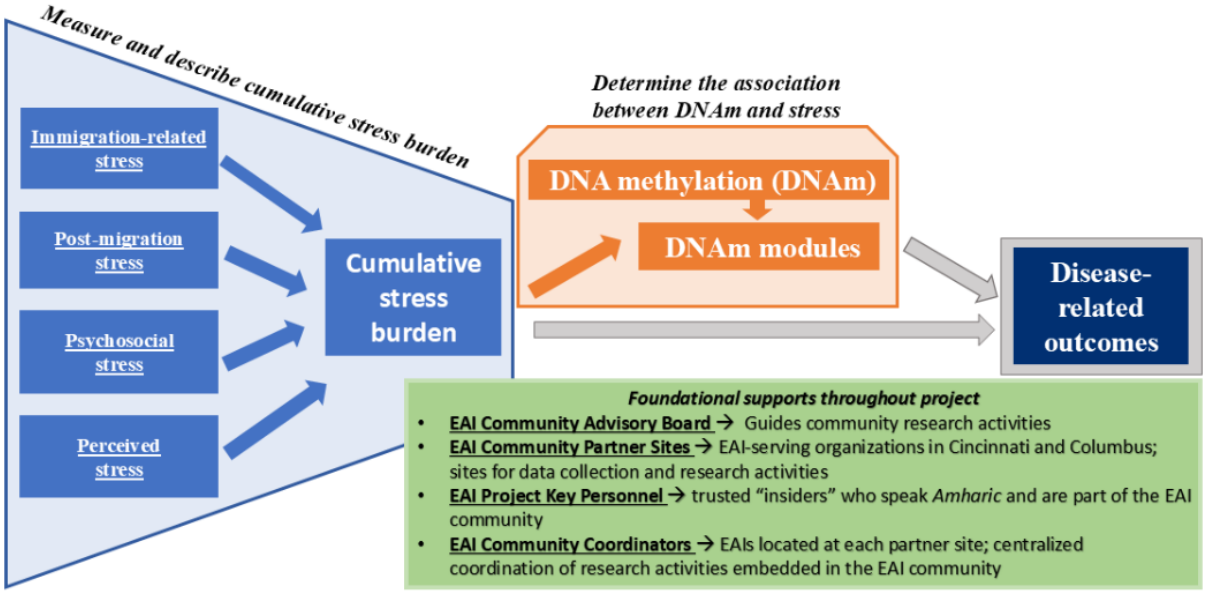

Methods

Overview and Study Design

In this biopsychosocial cross-sectional and qualitative study, we use a community-engaged research approach [54] working with Ethiopian community partners at each step of the research process. Data collection occurs in the community at public events (eg, health fairs, festivals, and meetings) using a multistation model comprised of five stations through which participants rotate: (1) eligibility screening and consent, (2) stress-related questionnaires to assess cumulative stress burden, (3) clinical measures (hemoglobin A_{1c} [HbA_{1c}], height, weight, and blood pressure [BP]), (4) health status and demographic questionnaires, and (5) saliva collection and incentive receipt. See supplementary files for Observational Study Protocol and Qualitative Study Protocol Checklists [55].

Study Setting

To capture a range of Ethiopian American immigrant experiences, data collection will occur with Ethiopian American immigrant community samples in 2 geographic areas in the Midwest: Greater Columbus and Greater Cincinnati, Ohio. These Midwest cities have not traditionally been destinations for immigrants, and culturally and linguistically appropriate social services and social networks are less established.

Community Partners

To ensure that our study is relevant and feasible for Ethiopian American immigrants and for a wider translation to the community, we are partnering with 2 Ethiopian American immigrant–serving community sites- an Ethiopian social service agency in Columbus and an Ethiopian Orthodox church in Cincinnati. We are working with an established CAB of Ethiopian leaders and 3 Ethiopian community research coordinators (BA, BG, and HE) hired for this study.

A CAB already established for a previous study will continue to meet for this study. The CAB has 6 members with representation from each of the 2 community partner sites. The CAB will help translate measures into Amharic and ensure the project is culturally and contextually relevant, sensitive, and feasible for Ethiopian American immigrants. The CAB will meet with the research team at least quarterly to provide input

XSL•FO
RenderX

on study progression to ensure the project is relevant and meaningfully engaging to the Ethiopian American immigrant communities, review study materials, and interpret and disseminate findings. CAB members will be compensated with US $50 for attendance at each CAB meeting.

The Ethiopian community research coordinators have relevant knowledge of health and community services and have received formal research training from the principal investigators to coordinate data collection. Community research coordinators have been trained in the appropriate collection of biometric screening tests for hypertension, diabetes, and obesity. The training sessions integrated and balanced the cultural and contextual expertise of the community research coordinators and the research expertise of the academic members. The community research coordinators hired for this study completed CIRTification, which is a training program in human research protections that is tailored to the unique roles of community research partners [56].

Study Sample

Over 12 months, we will recruit and enroll 150 Ethiopian American immigrant (approximately 75 per site) participants from 2 Ethiopian American immigrant–serving community partner sites in Columbus and Cincinnati, Ohio. Recruitment, consent, and conduct of the research will occur in collaboration with each site at their sites and associated public events (eg, health fairs and festivals). At each public data collection event, we will use maximal variation sampling [57] to identify potential participants of diverse ages, genders, and varied immigration length and status. After a potential participant is identified, we will ask screening questions to determine eligibility. Eligible participants must be (1) older than >18 years, (2) Ethiopian American immigrants born in Ethiopia, and (3) able to speak in Amharic or English.

Data Collection

Overview

At each public data collection event, eligible participants, in their preferred language (ie, English or Amharic) will receive a full verbal explanation of the study (ie, information about the study, eligibility criteria, compensation, benefits and risks to study participation, steps to ensure data confidentiality, and right to withdraw from study at any time) and written informed consent will be obtained prior to participation. Participants can ask questions about the study to a study team member, and if they agree to participate, sign the informed consent document (Station 1). Proceeding through the remaining stations, participants will complete verbally administered questionnaires (stress measures, health status, and demographics), provide biological samples (ie, finger prick point of care HbA_{1c} and saliva for DNA), and complete screening tests for obesity, hypertension, diabetes, depression, and anxiety (Table 1).

XSL•FO
RenderX

Table 1. Study measures.

Categories	Variables	Measures
Cumulative stress burden		
To capture both the unique and joint contributions of stress burden for EAIs[a], we will measure cumulative stress burden in a robust, multidimensional way, which includes validated stress scales assessing immigration-related stress post migration stress, and psychosocial perceived stress (recent and chronic stress).	• Immigration-related stress • Post migration and psychosocial stress • Perceived stress	• 21-item SOIS[b] • 36-item Boen's stress exposure measure • PSS-10[c]
Health conditions and health status		
Estimates of perceived and actual health status will be obtained using self-reported measures of health status and screening tests for obesity, diabetes, hypertension (high BP[d]), depression, and anxiety.	• Perceived health status • Overweight, obesity • Hypertension • Diabetes • Depression • Anxiety	• 10-item PROMIS[e] Global Health Scale (v1.2) • Weight, height, and BMI • Systolic and diastolic BPs • Glycosylated HbA1c[f] • PHQ-8[g] • GAD-7[h]
DNA methylation		
We will develop the WGCNA[i] module associated with stress burden variables and identify methylation modules and functional pathways.	• DNA	• Saliva samples or oral-rinse specimens will be obtained by a trained study team member
Demographics		
10 items regarding family and background	—	—

[a]EAI: Ethiopian American immigrant.

[d]SOIS: 21-item Stress of Immigration Survey.

[e]PSS-10: 10-item Perceived Stress Scale.

[b]BP: blood pressure.

[f]PROMIS: Patient-Reported Outcomes Measurement Information System.

[g]HbA_{1c}: hemoglobin A_{1c}.

[h]PHQ-8: 8-item Patient Health Questionnaire.

[i]GAD-7: 7-item Generalized Anxiety Disorder.

[c]WGCNA: weighted gene co-expression network analysis.

Data are anonymous, and each participant will receive a unique ID number, which is used on all measures. All study data and tracking will be entered into the REDCap (Research Electronic Data Capture; Vanderbilt University) database, a secure research electronic data management system that provides an interface for validated data entry, audit trails for tracking data manipulation, and export procedures. REDCap is compliant with the HIPAA (Health Insurance Portability and Accountability Act) and satisfies all local, state, and federal regulations for the capture and storage of private health information for research purposes.

Participants who express any distress during verbal administration of questionnaires, those who indicate high levels of stress, and those with abnormal or elevated clinical measures will be referred, and the on-call physician will determine the need for emergency room evaluation. Participants with an abnormal HbA_{1c} result (≥5.7), elevated BMI (≥25), and/or elevated BP (≥120/80 mm Hg) will be counseled on healthy lifestyle modifications and the need for medical evaluation with a primary care provider. Participants will then be provided with contact information for local primary care clinics. If a participant has a screening test that requires urgent or emergent evaluation (HbA_{1c} >9, random blood glucose of ≥ 400, and/or BP ≥180/120 mm Hg), the community coordinator will contact the on-call physician for consultation. The on-call physician will speak with the participant and determine the need for emergency room evaluation. Participants who complete the data collection event will receive a US $75 gift card.

Cumulative Stress Burden

To capture both the unique and joint contributions of stress burden for Ethiopian American immigrants, we will measure cumulative stress burden in a robust, multidimensional way [58] which includes validated stress scales assessing immigration-related stress, post-migration stress, and psychosocial perceived stress (recent and chronic stress) in combination with a saliva specimen (biological measure of stress) and qualitative semistructured interviews to capture additional missed experiences of stress and buffers to stress. The scales and interviews will be verbally administered by the trained community research coordinators. We will work with the CAB and a professional translator to translate scales and the interview guide into Amharic. We will work with the CAB

XSL•FO
RenderX

to ensure that all measures are culturally appropriate for Ethiopian American immigrant.

Immigration-related stress will be measured using the 21-item Stress of Immigration Survey (SOIS), which assesses 5 domains of stress: language, immigrant status, work issues, yearning for family and home country, and cultural dissonance among immigrants living in the United States [59]. The SOIS has robust psychometric properties with immigrant samples. Post migration, psychosocial, and perceived stress will be assessed using Boen's stress exposure measure [20] which considers stress and adversity in 6 domains of social life: lifetime traumas; recent stressful life events (last 5 y); financial strain; ongoing chronic strains (eg, health, work, relationship, and housing); everyday discrimination; and major lifetime discrimination. The individual measures have been previously used to assess the association between psychosocial stress and a variety of mental and physical health outcomes (eg, depressive symptoms, chronic illnesses, and physical limitations) [21]. Following the same procedure used by Boen and Hummer [20] and Cuevas et al [27], we will create 2 measures of cumulative exposure across these stress domains: (1) high risk stress burden (numbers of stressors for which the respondent is in the highest quartile and (2) composite total stress exposure across domains of stress (no cutoffs). The 10-item Perceived Stress Scale (PSS-10) is a 10-item scale that is used to assess the degree of perceived stress a person experiences [60]. The items relate to how unpredictable, uncontrollable, and overloaded respondents find their lives in general and do not focus on a specific event. The PSS-10 has been used in multiple languages and has well-established psychometric properties, including in Ethiopian university students [61-64]. A biological measure of stress will be collected with saliva samples or oral-rinse specimens obtained by trained study team members. Nonfasting saliva will be collected from each participant following oral rinsing with tap water. Participants will expectorate into an Oragene-RNA RE-100 Self-Collection Kit (DNA Genotek), and samples will be stored at room temperature for further analysis [65]. Buccal cells for DNA extraction will be collected using DNA Genotek saliva collection kits [66]. Saliva has been used as a noninvasive and stress-free diagnostic alternative to blood to examine associations of stress and resilience factors with DNAm by our groups and others [37,67,68].

Health Status and Health Conditions

Estimates of perceived and actual health status will be obtained using self-reported measures of health status and screening tests for obesity, diabetes, hypertension (high BP), depression, and anxiety. Perceived health status will be assessed with the Patient-Reported Outcomes Measurement Information System (PROMIS) Global v1.2 scale [69-71], a self-reported measure of global, physical, mental, and social health. Participants' BMI will be measured and used to categorize participants' weight status. Screening for hypertension will be conducted using measurements of systolic and diastolic BP. BP measurements will be used to categorize participants into categories (normal, elevated, stage 1 hypertension, and stage 2 hypertension). Screening for diabetes will be conducted using a point-of-care capillary blood glucose test (HbA_{1c}), which will be used to categorize participants (normal blood sugar, prediabetes, or diabetes). Depression and anxiety symptoms will be measured using the 8-item Patient Health Questionnaire (PHQ-8) and the 7-item Generalized Anxiety Disorder (GAD-7) assessments, both widely used and well-validated in different settings and population groups [72-75]. PHQ-8 and GAD-7 have been validated in Ethiopia using different cutoff points (5 or above on the PHQ-8; 10 or above on the GAD-7) [76-78].

Qualitative Interviews

To explore the nuances of Ethiopian American immigrant stress burden, trained members of the team will conduct qualitative semistructured interviews [79] with a subset of Aim 1 Ethiopian American immigrants who vary in levels of reported stress on the above measures. These interviews will be conducted to ensure we understand Ethiopian American immigrant experiences with stress that may be missed from standardized measures and to explore potential buffers to the adverse effects of stress (Multimedia Appendix 1 for interview guide reviewed and pilot tested with members of the CAB). We will contact participants by phone, text, or WhatsApp (Meta; depending on preference indicated during the public data collection events) and schedule interested individuals for a ~60-minute web-conference (eg, Zoom and Microsoft Teams) interview. We will use maximal variation sampling [57] to identify potential interviewees of diverse ages, genders, and varied immigration length and status. Data collection will continue until saturation, or informational redundancy [80-82] of major categories and themes is achieved. Specifically, we will operationalize saturation, for the purpose of this study, as the point at which no new substantive information is gained from 3 consecutive interviews. We will conduct a minimum of 10 interviews in each geographic area (Cincinnati and Columbus) and expect that we will reach data saturation within 20-25 total interviews based on our prior experience in similar qualitative projects with immigrant populations.

Data Analysis

We will use latent profile analysis (LPA) [83,84], combining all the stress subscales to identify, within the study sample of Ethiopian American immigrants, subgroups having similar within-group, but different between-group profiles of cumulative stress. When combined with epigenetic profiling and network-based methylation module analysis, a multidimensional understanding of stress is achieved. Instead of using a single-variable analysis, this method yields a systems-level perspective of how stress may shape immune-related gene activity and how it contributes to immigrants' vulnerability to disease. This type of analysis examines whether DNAm modules mediate the relationship between stress and chronic health conditions, including obesity, diabetes, hypertension, depression, and anxiety.

Prior to LPA, descriptives for the stress subscales will be generated, and normality, skewness, kurtosis, or outliers will be examined. Z-standardized mean subscale scores of all measures will be generated and used in LPA. As all measures are continuous variables, the maximum likelihood approach or maximum likelihood with robust SEs will be used for LPA. The number of profiles will rely on statistical fit, as well as theoretical and content-related considerations. Models with

XSL•FO
RenderX

different numbers of latent profiles will be fitted using a stepwise approach, starting with 2 profiles and successively adding profiles. In each step, we will examine the fit information criteria (Bayesian information criterion, [85] sample-size adjusted Bayesian information criterion, [86] and Akaike information criterion [87,88]). To compare models with one fewer or more profiles, we will apply the bootstrapped Likelihood Ratio Test [89,90] and the adjusted Lo-Mendell-Rubin test [91]. A profile that includes <1% of the total sample size or fewer than 25 cases will not be considered to avoid lower power, precision, and less parsimony. Profiles will be inspected for theoretical and content-related considerations and assigned labels. An alternative approach for LPA will be to use one measure of cumulative stress (eg, Stress of Immigration Survey). Missing data is assumed missing at random. Missing data will be examined for patterns and handled using either the full information maximum likelihood estimation method or multiple imputation. Based on past research and rules of thumb, the simulation study of Nylund et al [92] concluded that a smaller sample size may be adequate with simpler models and well-separated profiles.

Secondary Analyses A

Association between the composite stress burden and chronic physical and mental health conditions. In a secondary analysis using chi-square tests and logistic regression with adjustment for covariates, the identified latent stress burden profiles will be tested for associations with physical and mental health conditions for which we screened (obesity, diabetes, hypertension, depression, and anxiety). This also serves as validation of the profiles. We will also examine how the identified profiles are related to demographics and other covariates, and whether and how latent stress burden profiles differ by cities.

Genome-wide DNAm

We will use a genome-wide approach (Illumina EPIC arrays) to identify Cytosine-phosphate-Guanine (CpG) sites associated with stress and chronic illnesses (hypertension, obesity, cardiovascular disease, and diabetes). One-microgram DNA from saliva samples will be treated with bisulfite and submitted to the DNA Sequencing and Genotyping Core for DNAm quantification. The genome-wide DNAm analysis will be performed using the Infinium Human Methylation 850 BeadChip "860K array" (Illumina), [93] which covers 99% RefSeq genes, 333,265 CpG sites located in enhancer regions identified by the ENCODE (Encyclopedia of DNA Elements), 5' untranslated region (5'UTR), exon, gene, and 3' untranslated region (3'UTR) regions [94]. The 850K array has recently become an attractive choice for DNAm studies due to its effective cost and high coverage of promoter regions [95]. Samples will be randomly assigned to avoid batch effects due to chips and plates. The MethylationEPIC BeadChip images will be captured using the Illumina iScan. Samples with high-quality bisulfite conversion efficiency (intensity signal >4000) that also pass all GenomeStudio quality control steps based on built-in control probes for staining, hybridization, extension, and specificity will be filtered. Individual probes will be filtered away based on mean detection P>.01. The raw DNAm data will be exported from GenomeStudio and analyzed using Bioconductor and the lumi package.

Single CpG Site Analysis

Briefly, signal intensities of samples that pass quality control will be background-adjusted using out-of-band probes (noob) and normalized using subset-quantile within array normalization ("swan"). Beta values calculated as $signal_{methylation}/(signal_{methylation}+signal_{unmethylation})$ will be used as the measure for DNAm levels, while m values will be used in the analysis. Since DNAm is collected from saliva, and it may be affected by cell composition [96] and cell dilution, we will correct for cell composition using Houseman reference-based deconvolution algorithms (R package "minfi"). We will subsequently control for possible batch effects, as well as hidden structure in DNAm data, using surrogate variable analysis. To test association of DNAm with stress, we will use linear regression models (R limma package), [97] where DNAm level of a CpG locus will be the dependent variable, stress factors in aim 1 are the independent variables and the following covariates will be adjusted: significant surrogate variables from surrogate variable analysis, estimated proportions of saliva cell types, immigration duration, age, and sex, as well as a random effect for recruitment site. Differentially methylated CpG sites with methylation differences of at least 5% and P<.0001 will be identified.

Construction of Methylation Networks and Identification of Methylation Modules

Because exposure to cumulative stress most likely affects methylation of a variety of genes, single CpG site analysis may fail to fully capture the epigenetic pathways impacted by stress. We plan to use weighted gene coexpression network analysis (WGCNA) to identify methylation modules (ie, groups of genes with highly correlated methylation values and that may represent biological pathways) and then relate these modules to stress profiles. We apply the coexpression approach because differentially expressed genes may fail to fully capture the molecular pathways, as genes do not function in isolation [98]. Briefly, WGCNA takes the methylation correlation matrix and transforms it into a weighted undirected network by raising the absolute value of each pair-wise correlation to a power β. The higher the value of β, the more weight is given to high correlations at the expense of low correlations [99]. This weighting scheme preserves the continuous nature of the methylation data and produces highly robust results. Modules are then identified using average linkage hierarchical clustering. In addition, the use of network approaches reduces the burden of multiple testing, as we will test the association of a few hundred modules, representing biological pathways, with the stressors of interest, rather than test hundreds of thousands of methylation sites. TM and LD have conducted WGCNA-related experiments in the past [98,100].

Cumulative Stress - DNAm Modules Association Analysis

We will calculate for each module the eigengene E that is defined as the first principal component of the methylation values of a given module. The epigenome can be considered as

XSL•FO
RenderX

representative of the methylation profiles in a module. The association between the identified modules and cumulative stress can then be evaluated using linear models. Analyses will be adjusted for principal components of methylation variation (calculated from the epigenome-wide methylation data), estimated proportions of saliva cell types, other covariates (eg, immigration duration, age, and sex), and random recruitment site effect. We will also examine what genes and/or CpG sites have a greater contribution to the stressor-module association by looking into their weights on the module eigengene.

Secondary Analyses B

Determine the mediation effects of DNAm modules between stress profiles and specific outcome Studies have shown an association between stress and health outcomes. However, none of these studies has examined whether health outcomes in immigrants are potentially mediated by DNAm modules. A mediation analysis will be performed to assess whether or not the observed association between stress and health outcome is mediated through DNAm modules. Mediating variables are selected on sound theories that provide evidence for pathways of relationship [101-103].

Testing for DNAm Modules Mediation Effects

Our mediation analysis aims to explore the relationship between the independent variable (stress, X), the mediator (DNAm, M), and the dependent variable (disease-related outcome, Y) [102-104]. The test for mediation will be accomplished with a 3-step model. We will include covariates (age, sex, and immigration duration) in all of the models. However, for simplicity, we will omit them in the following equations: Step 1: logit(Pr(Y=1, x)) = α_1 + C'X, testing for path C' (total effect) when X predicting Y; Step 2: M = α_2 + AX + ε_2, testing for path A when X predicting M; and Step 3: logit(Pr(Y=1|x, m))= α_3 + BM + CX, multiple regression analysis with X and M predicting Y. The direct effect (path C) is the effect of stress on health outcome, adjusting for the mediator (methylation modules). The mediation effect is the indirect effect (path A*B) [105]. We will estimate the average causal mediation effects [106].

Power Analysis

Power analysis is based on the association between DNAm module eigengenes and cumulative stress burden. We expect to analyze about 100 normalized module eigengenes from 100 methylation modules using genome-wide methylation data. With a sample size of 150, for a multiple linear regression model which already includes ~6 covariates with a squared multiple correlation ρ^2 of 0.1, a sample size of 150 will have 80% power to detect at a significance level of 3.125×10^{-5} (=0.05/100 methylation modules × 16 stress subscales) an increase in ρ^2 of 0.136 due to including 1 additional stress subscale.

Qualitative Analysis

We will use an inductive process of thematic analysis [107,108] to code the interview data to the point of saturation (described above). Guided by an expert in qualitative research (LV), we will use Dedoose, a qualitative software for data coding and management [109]. At least two study team members with prior experience in qualitative research and study-specific training will double-code each interview, modifying the coding structure as needed. Using established standards for qualitative research [110,111], the coding will be compared, and disagreements resolved through discussion. As coding progresses, individual concepts will be grouped into categories and eventually themes. Once we have preliminary themes, we will use a "reflexive participant collaboration" approach [112] and conduct validation (peer review) interviews [113] with members of the CAB and additional participants.

Ethical Considerations

The institutional review board at Cincinnati Children's Hospital Medical Center approved this study (#2024-0176). Informed consent in the participant's preferred language (English or Amharic) was obtained for all participants. The study data are deidentified and entered directly into a password-protected REDCap database. Only study staff have access to this database. Participants are compensated via a reloadable debit card. They receive US $75 for initial data collection and US $40 if selected for a qualitative interview. CAB members receive US $50 for their participation in meetings.

Results

The study will span 2 years (Figure 2). After receipt of funding, the first 3 months were devoted to institutional board review approval, hiring and training community research coordinators, reconvening the CAB from a previous project, ordering data collection supplies (eg, clinical tests and tools and saliva collection tools), and initial planning conversations with community partners regarding recruitment, challenges, and data collection strategies. October 2024 to November 2025 will be focused on data collection (currently at 89% completion). In November 2025, we will start data cleaning and analysis of questionnaires, clinical measures, and DNAm. We have begun qualitative interviews with a subset of the participants to understand additional experiences of stress and buffers to stress among participants. Although data analysis has not started, preliminary insights from the study include participants' discomfort and unwillingness to use the word stress in stress-related questionnaires and interviews, and instead emphasizing the words "worry" and "concerns," and perceived stigma of naming and/or identifying with mental health disorders such as depression and anxiety on the PHQ-8 and GAD-7. The importance of religion, prayer, turning to God, and reliance on priests has consistently been emphasized by participants during data collection and interviews. Social connections within and across Ethiopians living in Cincinnati and Columbus versus non-Ethiopians seem to be prioritized based on participants' responses to the questionnaires and in the interviews thus far. We plan to complete all data analyses and prepare scientific and community outputs by Spring 2026.

XSL•FO
RenderX

Figure 2. Study timeline.

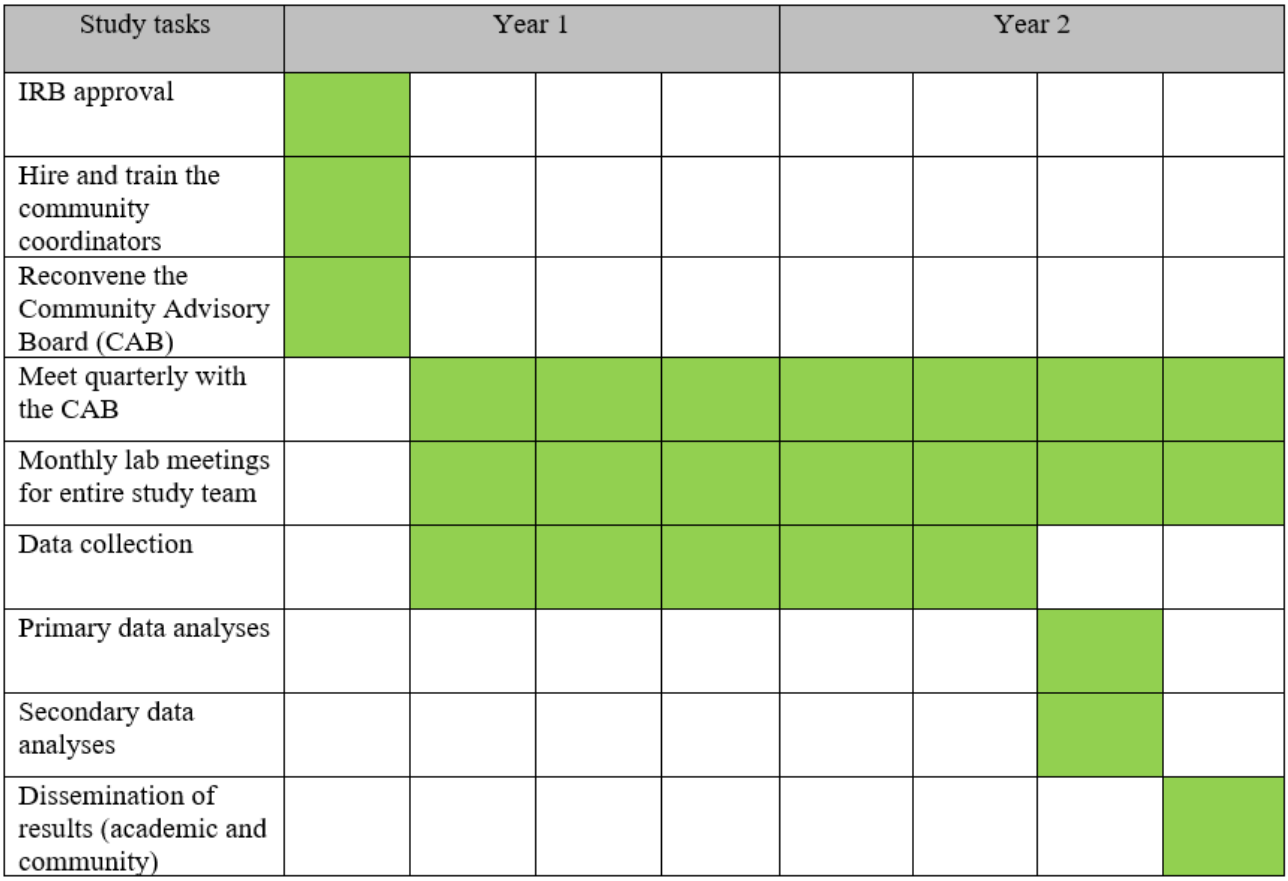

Discussion

Anticipated Findings

This study protocol investigates how cumulative stress burden affects chronic disease risk among Ethiopian American immigrants. It examines how cumulative stress uses both psychosocial and biological pathways. To this aim, we used a combination of validated stress measures, clinical health indicators, and genome-wide DNAm analysis. The framework for this study acknowledges that the stress experienced by immigrants is complex and can include premigration trauma, postmigration challenges, and sociopolitical adversity. Compounding the mental health risks of stress among immigrants is a significant disparity in adequate treatment for depression, mental health concerns, and substance abuse [114-116]. Barriers to access and quality mental health care for immigrants are multifactorial and include complex drivers across societal and individual levels [114,117-120]. Barriers at the societal and structural level include lack of health insurance, low socioeconomic status, limited English proficiency, low educational attainment, documentation status, logistics, scheduling, and culturally and linguistically congruent health care providers [114,119,121-125]. At the individual level, cultural mistrust of the mental health system, perceived discrimination from health care providers, focus on physical versus psychiatric symptoms, and cultural stigma can lead immigrants to attempt to manage stress and mental health issues on their own [11,114,119,126-131]. Because of the increased risk of stress burden, immigrants are in dire need of secondary prevention via mental health support that is accessible, asset-based, and culturally relevant [50]. Without such support, stress burden and barriers to mental health care experienced by many immigrants increase the risk of suicide, addiction, chronic disease, and early death [124,128,132-141]. It is important to note that the study's community-engaged approach fosters cultural sensitivity, builds trust within the Ethiopian American immigrant community, and strengthens recruitment and retention efforts among a population that is typically difficult to study through conventional research strategies [53,142].

Our initial qualitative findings indicate that the way stress is expressed and understood, such as the culturally preferred terms like "worry" or "concern" instead of stress, along with the influence of religious beliefs and intraethnic social ties, are essential for capturing how Ethiopian American immigrants experience and interpret stress. This knowledge will assist in the interpretation of quantitative results and will inform future interventions. US Black African immigrants have been underrepresented in research about health. The results of the study will make an important contribution to the literature on immigrant health disparities by focusing on Black African immigrants from Ethiopia. The use of epigenetic data gives us the opportunity to study how stress may affect us biologically, and whether it may be transmitted across generations. Results of this study will inform the development of culturally relevant interventions that can positively impact the health and well-being of US immigrants.

XSL•FO
RenderX

Limitations

While the study offers substantial strengths, several limitations should be acknowledged. The participant pool is restricted to Ethiopian American immigrants living in Columbus and Cincinnati, Ohio—areas considered nontraditional migration destinations with distinct social environments. As a result, the findings may not be broadly applicable to Ethiopian American immigrants in other regions with differing sociopolitical contexts, community structures, or support systems. Additionally, the study design limits the ability to draw causal conclusions. Although mediation analyses are planned, longitudinal data will be essential to establish temporal relationships between stress exposure, DNAm changes, and health outcomes.

Many of the study's measures rely on self-reported data, which may be affected by language barriers and cultural stigma, especially regarding mental health. While saliva collection is a convenient and noninvasive method suitable for community-based research, it may not fully reflect methylation patterns in specific tissues involved in chronic diseases. Furthermore, although the study is adequately powered to examine stress profiles and their associations with health indicators, the sample size may be relatively small for genome-wide DNAm analyses—particularly when detecting small effect sizes or conducting subgroup analyses.

Comparison With Prior Work

This study builds on our previous work in immigrant health and a recent health needs assessment of local Ethiopian American immigrants [53].

Conclusions

This study is innovative in its exploration of existing health disparities among immigrant populations in the United States. Results may offer valuable insights that support the development of new and effective intervention strategies to reduce chronic illness in these vulnerable communities. Specifically, findings from this study may highlight the need for policy change, such as providing more support in the form of infrastructure and social service agencies to support immigrants.

Acknowledgments

The authors are grateful for their participants' willingness to participate in this study.

Funding

This work was supported by the National Institutes of Health National Institute on Minority Health and Health Disparities (1R21MD019192-01A1).

Data Availability

The datasets generated or analyzed during this study will be available from the principal investigators (TM and LV) on reasonable request once the study is completed.

Authors' Contributions

Conceptualization: LV and TM
Data curation: LS and LV
Formal analysis: LS, LD, MHK, BA, BG, HE, TM, and LMV
Funding acquisition: LV and TM
Investigation: LS, LD, MHK, BA, BG, HE, TM, and LMV
Methodology: LV and TM
Project administration: LV, TM, and LS
Resources: LV, TM, LS, and MHK
Software: LD and LS
Supervision: LV, TM, LS, LD, and MHK
Validation: LV, TM, and LD
Visualization: LV
Writing – LS, LD, MHK, BA, BG, HE, TM, and LMV
Writing – review and editing: LS, LD, MHK, BA, BG, HE, TM, and LMV

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from MESH - Biobehavioral Mechanisms of Emotion, Stress and Health Study Section, Biobehavioral and Behavioral Processes Integrated Review Group (National Institutes of Health, USA).

XSL•FO
RenderX

[PDF File (Adobe PDF File), 147 KB - resprot_v15i1e85971_app1.pdf]

References

1. Garcini LM, Domenech Rodríguez MM, Mercado A, Silva M, Cadenas G, Galvan T, et al. Anti-immigration policy and mental health: risk of distress and trauma among deferred action for childhood arrivals recipients in the United States. Psychol Trauma 2023;15(7):1067-1075 [FREE Full text] [doi: 10.1037/tra0001228] [Medline: 35482682]
2. Garcini LM, Nguyen K, Lucas-Marinelli A, Moreno O, Cruz PL. "No one left behind": a social determinant of health lens to the wellbeing of undocumented immigrants. Curr Opin Psychol 2022;47:101455 [FREE Full text] [doi: 10.1016/j.copsyc.2022.101455] [Medline: 36055083]
3. Ferguson GM, Causadias JM, Simenec TS. Acculturation and psychopathology. Annu Rev Clin Psychol 2023;19:381-411 [FREE Full text] [doi: 10.1146/annurev-clinpsy-080921-080622] [Medline: 36854286]
4. Mercado A, Venta A, Morales F, Palomin A, Garcini L, Silva M, et al. Trauma in the American asylum process: experiences of immigrant families under the migrant protection protocols. Psychol Trauma 2024;16(Suppl 2):S379-S388 [FREE Full text] [doi: 10.1037/tra0001368] [Medline: 36222662]
5. Salas-Wright CP, Kagotho N, Vaughn MG. Mood, anxiety, and personality disorders among first and second-generation immigrants to the United States. Psychiatry Res 2014;220(3):1028-1036 [FREE Full text] [doi: 10.1016/j.psychres.2014.08.045] [Medline: 25223256]
6. Ayele SG, Thompson-Robinson M, Andrews J, Dodge Francis C. Factors associated with mental health service utilization among ethiopian immigrants and refugees. J Immigr Minor Health 2020;22(5):965-972. [doi: 10.1007/s10903-020-00984-w] [Medline: 32130570]
7. Covington-Ward Y, Agbemenu K, Matambanadzo A. "We feel like it was better back home:" stress, coping, and health in a U.S. dwelling African immigrant community. J Health Care Poor Underserved 2018;29(1):253-265. [doi: 10.1353/hpu.2018.0018] [Medline: 29503299]
8. Sangalang CC, Becerra D, Mitchell FM, Lechuga-Peña S, Lopez K, Kim I. Trauma, post-migration stress, and mental health: a comparative analysis of refugees and immigrants in the United States. J Immigr Minor Health 2019;21(5):909-919 [FREE Full text] [doi: 10.1007/s10903-018-0826-2] [Medline: 30244330]
9. Finklestein M, Solomon Z. Cumulative trauma, PTSD and dissociation among Ethiopian refugees in Israel. J Trauma Dissociation 2009;10(1):38-56. [doi: 10.1080/15299730802485151] [Medline: 19197711]
10. Morey BN. Mechanisms by which anti-immigrant stigma exacerbates racial/ethnic health disparities. Am J Public Health 2018;108(4):460-463 [FREE Full text] [doi: 10.2105/AJPH.2017.304266] [Medline: 29470116]
11. Szaflarski M, Bauldry S. The effects of perceived discrimination on immigrant and refugee physical and mental health. Adv Med Sociol 2019;19:173-204 [FREE Full text] [doi: 10.1108/S1057-629020190000019009] [Medline: 31178652]
12. Salas-Wright CP, Vaughn MG, Goings TC, Oh S, Delva J, Cohen M, et al. Trends and mental health correlates of discrimination among Latin American and Asian immigrants in the United States. Soc Psychiatry Psychiatr Epidemiol 2020;55(4):477-486. [doi: 10.1007/s00127-019-01811-w] [Medline: 31811317]
13. Viruell-Fuentes EA. Beyond acculturation: immigration, discrimination, and health research among Mexicans in the United States. Soc Sci Med 2007;65(7):1524-1535. [doi: 10.1016/j.socscimed.2007.05.010] [Medline: 17602812]
14. Beiser M, Taa B, Fenta-Wube H, Baheretibeb Y, Pain C, Araya M. A comparison of levels and predictors of emotional problems among preadolescent Ethiopians in Addis Ababa, Ethiopia, and Toronto, Canada. Transcult Psychiatry 2012;49(5):651-677. [doi: 10.1177/1363461512457155] [Medline: 23015641]
15. Bigby J, Ko LK, Johnson N, David MMA, Ferrer B, REACH Boston 2010 BreastCervical Cancer Coalition. A community approach to addressing excess breast and cervical cancer mortality among women of African descent in Boston. Public Health Rep 2003;118(4):338-347 [FREE Full text] [doi: 10.1093/phr/118.4.338] [Medline: 12815081]
16. Akinsulure-Smith AM. Resilience in the face of adversity: African immigrants' mental health needs and the American transition. J Immigr Refug Stud 2016;15(4):428-448. [doi: 10.1080/15562948.2016.1238989]
17. Linton JM, Choi R, Mendoza F. Caring for children in immigrant families: vulnerabilities, resilience, and opportunities. Pediatr Clin North Am 2016;63(1):115-130. [doi: 10.1016/j.pcl.2015.08.006] [Medline: 26613692]
18. Jacquez F, Vaughn LM, Pelley T, Topmiller M. Healthcare experiences of Latinos in a nontraditional destination area. J Community Pract 2015;23(1):76-101. [doi: 10.1080/10705422.2014.983629]
19. Shandy DJ, Fennelly K. A comparison of the integration experiences of two African immigrant populations in a rural community. J Relig Spiritual Soc Work 2006;25(1):23-45. [doi: 10.1300/j377v25n01_03]
20. Boen CE, Hummer RA. Longer-but harder-lives?: The hispanic health paradox and the social determinants of racial, ethnic, and immigrant-native health disparities from midlife through late life. J Health Soc Behav 2019;60(4):434-452 [FREE Full text] [doi: 10.1177/0022146519884538] [Medline: 31771347]
21. Sternthal MJ, Slopen N, Williams DR. Racial disparities in health: the impact of multi-dimensional measures of race/ethnicity on the self-reported health status of Latinos. Du Bois Rev 2011;8(1):95-113 [FREE Full text] [doi: 10.1017/S1742058X11000087] [Medline: 29887911]
22. Williams DR. Stress and the mental health of populations of color: advancing our understanding of race-related stressors. J Health Soc Behav 2018;59(4):466-485 [FREE Full text] [doi: 10.1177/0022146518814251] [Medline: 30484715]

XSL•FO
RenderX

23. Nurius PS, Fleming CM, Brindle E. Life course pathways from adverse childhood experiences to adult physical health: a structural equation model. J Aging Health 2019;31(2):211-230. [doi: 10.1177/0898264317726448] [Medline: 28845729]
24. Nurius PS, Green S, Logan-Greene P, Longhi D, Song C. Stress pathways to health inequalities: embedding ACEs within social and behavioral contexts. Int Public Health J 2016;8(2):241-256 [FREE Full text] [Medline: 27274786]
25. Kaestner R, Pearson JA, Keene D, Geronimus AT. Stress, allostatic load and health of Mexican immigrants. Soc Sci Q 2009;90(5):1089-1111 [FREE Full text] [doi: 10.1111/j.1540-6237.2009.00648.x] [Medline: 21165158]
26. Cohen S, Janicki-Deverts D. Who's stressed? Distributions of psychological stress in the United States in probability samples from 1983, 2006, and 2009. J Applied Social Pyschol 2012;42(6):1320-1334. [doi: 10.1111/j.1559-1816.2012.00900.x]
27. Cuevas AG, Chen R, Thurber KA, Slopen N, Williams DR. Psychosocial stress and overweight and obesity: findings from the Chicago community adult health study. Ann Behav Med 2019;53(11):NP [FREE Full text] [doi: 10.1093/abm/kaz008] [Medline: 30917198]
28. Ferraro KF, Shippee TP. Aging and cumulative inequality: how does inequality get under the skin? Gerontologist 2009;49(3):333-343 [FREE Full text] [doi: 10.1093/geront/gnp034] [Medline: 19377044]
29. Tobi EW, Goeman JJ, Monajemi R, Gu H, Putter H, Zhang Y, et al. DNA methylation signatures link prenatal famine exposure to growth and metabolism. Nat Commun 2014;5(1):5592 [FREE Full text] [doi: 10.1038/ncomms6592] [Medline: 25424739]
30. Yehuda R, Lehrner A. Intergenerational transmission of trauma effects: putative role of epigenetic mechanisms. World Psychiatry 2018;17(3):243-257 [FREE Full text] [doi: 10.1002/wps.20568] [Medline: 30192087]
31. Mulligan CJ, Quinn EB, Hamadmad D, Dutton CL, Nevell L, Binder AM. Epigenetic signatures of intergenerational exposure to violence in three generations of Syrian refugees. Sci Rep 2025;15(1):5945 [FREE Full text] [doi: 10.1038/s41598-025-89818-z] [Medline: 40016245]
32. Banushi B, Collova J, Milroy H. Epigenetic echoes: bridging nature, nurture, and healing across generations. Int J Mol Sci 2025;26(7):3075 [FREE Full text] [doi: 10.3390/ijms26073075] [Medline: 40243774]
33. Pilkay S, Riffer A, Carroll A. Trauma context exerts intergenerational effects on child mental health via DNA methylation. Epigenetics 2024;19(1):2333654 [FREE Full text] [doi: 10.1080/15592294.2024.2333654] [Medline: 38577817]
34. Zhen-Duan J, Canenguez KM, Wilson AE, Gu Y, Valluri HG, Chavez AD. Religion, spirituality, and DNA methylation in HPA-axis genes among Hispanic/Latino adults. Epigenomics 2025;17(3):155-166 [FREE Full text] [doi: 10.1080/17501911.2024.2442293] [Medline: 39707707]
35. Clausing ES, Non AL. Epigenetics as a mechanism of developmental embodiment of stress, resilience, and cardiometabolic risk across generations of Latinx immigrant families. Front Psychiatry 2021;12:696827 [FREE Full text] [doi: 10.3389/fpsyt.2021.696827] [Medline: 34354616]
36. Chilunga FP, Henneman P, Venema A, Meeks KA, Gonzalez JR, Ruiz-Arenas C, et al. DNA methylation as the link between migration and the major noncommunicable diseases: the RODAM study. Epigenomics 2021;13(9):653-666 [FREE Full text] [doi: 10.2217/epi-2020-0329] [Medline: 33890479]
37. Boateng D, Agyemang C, Beune E, Meeks K, Smeeth L, Schulze M, et al. Migration and cardiovascular disease risk among Ghanaian populations in Europe: the RODAM study (research on obesity and diabetes among African migrants). Circ Cardiovasc Qual Outcomes 2017;10(11):e004013. [doi: 10.1161/CIRCOUTCOMES.117.004013] [Medline: 29150534]
38. Agyemang C, Addo J, Bhopal R, Aikins ADG, Stronks K. Cardiovascular disease, diabetes and established risk factors among populations of sub-Saharan African descent in Europe: a literature review. Global Health 2009;5:7 [FREE Full text] [doi: 10.1186/1744-8603-5-7] [Medline: 19671137]
39. van der Laan LC, Meeks KAC, Chilunga FP, Agyemang C, Venema A, Mannens MMAM, et al. Epigenome-wide association study for perceived discrimination among sub-Saharan African migrants in Europe - the RODAM study. Sci Rep 2020;10(1):4919 [FREE Full text] [doi: 10.1038/s41598-020-61649-0] [Medline: 32188935]
40. Agyemang C, Meeks K, Beune E, Owusu-Dabo E, Mockenhaupt FP, Addo J, et al. Obesity and type 2 diabetes in sub-Saharan Africans - is the burden in today's Africa similar to African migrants in Europe? The RODAM study. BMC Med 2016;14(1):166 [FREE Full text] [doi: 10.1186/s12916-016-0709-0] [Medline: 27769239]
41. Gudsnuk K, Champagne FA. Epigenetic influence of stress and the social environment. ILAR J 2012;53(3-4):279-288 [FREE Full text] [doi: 10.1093/ilar.53.3-4.279] [Medline: 23744967]
42. Krieger N. Embodiment: a conceptual glossary for epidemiology. J Epidemiol Community Health 2005;59(5):350-355 [FREE Full text] [doi: 10.1136/jech.2004.024562] [Medline: 15831681]
43. Lussier AA, Zhu Y, Smith BJ, Simpkin AJ, Smith ADAC, Suderman MJ, et al. Sensitive periods for the effect of childhood adversity on DNA methylation: updated results from a prospective, longitudinal study. Biol Psychiatry Glob Open Sci 2023;3(3):567-571 [FREE Full text] [doi: 10.1016/j.bpsgos.2022.04.002] [Medline: 37519470]
44. Clausing ES, Binder AM, Non AL. Epigenetic age associates with psychosocial stress and resilience in children of Latinx immigrants. Epigenomics 2021;13(21):1677-1699. [doi: 10.2217/epi-2019-0343] [Medline: 33749330]
45. Hiu-ming CM, Merrill S, Fatima F, MacIsa J, Obradović J, Boyce W, et al. Cross-tissue specificity of pediatric DNA methylation associated with cumulative family adversity. bioRxiv 2023 [FREE Full text] [doi: 10.1101/2023.10.04.559423]
46. Aroke EN, Joseph PV, Roy A, Overstreet DS, Tollefsbol TO, Vance DE, et al. Could epigenetics help explain racial disparities in chronic pain? JPR 2019;12:701-710. [doi: 10.2147/jpr.s191848] [Medline: 30863142]

XSL•FO
RenderX

47. Tefera GM. Understanding access and utilization of healthcare services among African immigrant women in the United States: the application of health belief model. Appl Res Qual Life 2024;19(3):1097-1115. [doi: 10.1007/s11482-024-10283-3]
48. Omenka OI, Watson DP, Hendrie HC. Understanding the healthcare experiences and needs of African immigrants in the United States: a scoping review. BMC Public Health 2020;20(1):27 [FREE Full text] [doi: 10.1186/s12889-019-8127-9] [Medline: 31914960]
49. Donnella L. Black immigrants are growing in numbers, but in the US many often feel invisible. NPR. 2023. URL: https://www.npr.org/2023/06/04/1165058041/black-immigrants-nashville-tennessee [accessed 2026-01-09]
50. Ibe-Lamberts K, Ajibewa T, Onyeise D, McNeil-Young V, Nmezi N, Williams J. "I'm Black and I'm Stressed": an exploratory literature review on stress and coping mechanisms among African Americans, afro-Caribbean, and African immigrants in the USA. J Racial Ethn Health Disparities 2025:1-10 [FREE Full text] [doi: 10.1007/s40615-025-02394-w] [Medline: 40146467]
51. Anderson M. African immigrant population in US steadily climbs. Pew Research Center. 2017. URL: https://www.pewresearch.org/fact-tank/2017/02/14/african-immigrant-population-in-u-s-steadily-climbs/ [accessed 2022-05-23]
52. RAD diaspora profile-the Ethiopian diaspora in the United States. Migration Policy Institute. 2014. URL: https://www.migrationpolicy.org/sites/default/files/publications/RAD-Ethiopia.pdf [accessed 2021-07-15]
53. Kassa MH, Vaughn LM, Amsalu B, Schnadower Z, Jorga I, Habte A, et al. Community health needs assessment of Ethiopian American immigrants in two nontraditional migration cities. J Health Care Poor Underserved 2025;36(3):830-854. [doi: 10.1353/hpu.2025.a967335] [Medline: 40820778]
54. Haapanen KA, Christens BD. Community-engaged research approaches: multiple pathways to health equity. Am J Community Psychol 2021;67(3-4):331-337. [doi: 10.1002/ajcp.12529] [Medline: 34312882]
55. Low GK, Subedi S, Omosumwen OF, Jiee SF, Devkota S, Shanmuganathan S, et al. Development and validation of observational and qualitative study protocol reporting checklists for novice researchers (ObsQual checklist). Eval Program Plann 2024;106:102468 [FREE Full text] [doi: 10.1016/j.evalprogplan.2024.102468] [Medline: 39029287]
56. Center for Clinical and Translational Science. CIRTification: A human research protections training program tailored to the unique roles of community research partners. University of Illinois Chicago. 2025. URL: https://ccts.uic.edu/resources/cirtification/ [accessed 2024-10-02]
57. Patton M. Purposeful sampling. In: Patton M, editor. Qualitative Evaluation and Research Methods. Washington DC: Sage; 1990.
58. Epel ES, Crosswell AD, Mayer SE, Prather AA, Slavich GM, Puterman E, et al. More than a feeling: a unified view of stress measurement for population science. Front Neuroendocrinol 2018;49:146-169 [FREE Full text] [doi: 10.1016/j.yfrne.2018.03.001] [Medline: 29551356]
59. Sternberg RM, Nápoles AM, Gregorich S, Paul S, Lee KA, Stewart AL. Development of the stress of immigration survey: a field test among Mexican immigrant women. Fam Community Health 2016;39(1):40-52 [FREE Full text] [doi: 10.1097/FCH.0000000000000088] [Medline: 26605954]
60. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983 Dec;24(4):385-396. [doi: 10.2307/2136404] [Medline: 6668417]
61. Lee EH. Review of the psychometric evidence of the perceived stress scale. Asian Nurs Res (Korean Soc Nurs Sci) 2012;6(4):121-127 [FREE Full text] [doi: 10.1016/j.anr.2012.08.004] [Medline: 25031113]
62. Taylor JM. Psychometric analysis of the ten-item perceived stress scale. Psychol Assess 2015;27(1):90-101. [doi: 10.1037/a0038100] [Medline: 25346996]
63. Ramírez MTG, Hernández RL. Factor structure of the perceived stress scale (PSS) in a sample from Mexico. Span J Psychol 2007;10(1):199-206 [FREE Full text] [doi: 10.1017/s1138741600006466] [Medline: 17549893]
64. Manzar MD, Salahuddin M, Peter S, Alghadir A, Anwer S, Bahammam AS, et al. Psychometric properties of the perceived stress scale in Ethiopian university students. BMC Public Health 2019;19(1):41 [FREE Full text] [doi: 10.1186/s12889-018-6310-z] [Medline: 30626438]
65. Iwasiow R, Desbois A, Birnboim H. Long-term stability of DNA from saliva samples stored in the Oragene self-collection kit. Oragene 2011(3):1-2 [FREE Full text]
66. Folger AT, Ding L, Ji H, Yolton K, Ammerman RT, Van Ginkel JB, et al. Neonatal NR3C1 methylation and social-emotional development at 6 and 18 months of age. Front Behav Neurosci 2019;13:14 [FREE Full text] [doi: 10.3389/fnbeh.2019.00014] [Medline: 30804765]
67. Desai GS, Mathews ST. Saliva as a non-invasive diagnostic tool for inflammation and insulin-resistance. World J Diabetes 2014;5(6):730-738 [FREE Full text] [doi: 10.4239/wjd.v5.i6.730] [Medline: 25512775]
68. Parikh MN, Brokamp C, Rasnick E, Ding L, Mersha TB, Bowers K, et al. Epigenome-wide association of neonatal methylation and trimester-specific prenatal PM exposure. Environ Epidemiol 2022;6(5):e227 [FREE Full text] [doi: 10.1097/EE9.0000000000000227] [Medline: 36249271]
69. Barile JP, Reeve BB, Smith AW, Zack MM, Mitchell SA, Kobau R, et al. Monitoring population health for healthy people 2020: evaluation of the NIH PROMIS® Global Health, CDC healthy days, and satisfaction with life instruments. Qual Life Res 2013;22(6):1201-1211 [FREE Full text] [doi: 10.1007/s11136-012-0246-z] [Medline: 23404737]

XSL•FO
RenderX

70. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009;18(7):873-880 [FREE Full text] [doi: 10.1007/s11136-009-9496-9] [Medline: 19543809]
71. Revicki DA, Kawata AK, Harnam N, Chen WH, Hays RD, Cella D. Predicting EuroQol (EQ-5D) scores from the patient-reported outcomes measurement information system (PROMIS) global items and domain item banks in a United States sample. Qual Life Res 2009;18(6):783-791 [FREE Full text] [doi: 10.1007/s11136-009-9489-8] [Medline: 19472072]
72. Kroenke K, Spitzer RL, Williams JBW, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010;32(4):345-359. [doi: 10.1016/j.genhosppsych.2010.03.006] [Medline: 20633738]
73. Gelaye B, Tadesse MG, Williams MA, Fann JR, Vander Stoep A, Andrew Zhou X. Assessing validity of a depression screening instrument in the absence of a gold standard. Ann Epidemiol 2014;24(7):527-531 [FREE Full text] [doi: 10.1016/j.annepidem.2014.04.009] [Medline: 24935465]
74. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
75. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
76. Yeakey MP, Chipeta E, Rijken Y, Taulo F, Tsui AO. Experiences with fistula repair surgery among women and families in Malawi. Glob Public Health 2011;6(2):153-167 [FREE Full text] [doi: 10.1080/17441692.2010.491833] [Medline: 20614361]
77. Belete A, Andaregie G, Tareke M, Birhan T, Azale T. Prevalence of anxiety and associated factors among people living HIV/AIDS at debretabor general hospital anti retro viral clinic Debretabor, Amhara, Ethiopia, 2014. American Journal of Psychiatry Neuroscience 2014;2(6):109-114. [doi: 10.11648/j.ajpn.20140206.15]
78. Hanlon C, Medhin G, Selamu M, Breuer E, Worku B, Hailemariam M, et al. Validity of brief screening questionnaires to detect depression in primary care in Ethiopia. J Affect Disord 2015;186:32-39. [doi: 10.1016/j.jad.2015.07.015] [Medline: 26226431]
79. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Community Health 2019;7(2):e000057 [FREE Full text] [doi: 10.1136/fmch-2018-000057] [Medline: 32148704]
80. Guest G, Namey E, Chen M. A simple method to assess and report thematic saturation in qualitative research. PLoS One 2020;15(5):e0232076 [FREE Full text] [doi: 10.1371/journal.pone.0232076] [Medline: 32369511]
81. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
82. Morse JM. "Data were saturated . . . ". Qual Health Res 2015;25(5):587-588. [doi: 10.1177/1049732315576699] [Medline: 25829508]
83. Gibson WA. Three multivariate models: factor analysis, latent structure analysis, and latent profile analysis. Psychometrika 2025;24(3):229-252. [doi: 10.1007/bf02289845]
84. Oberski D. Mixture models: latent profile and latent class analysis. In: Modern Statistical Methods for HCI. USA: Springer; 2016:275-287.
85. Schwarz G. Estimating the dimension of a model. Ann Statist 1978;6(2):461-464. [doi: 10.1214/aos/1176344136]
86. Sclove SL. Application of model-selection criteria to some problems in multivariate analysis. Psychometrika 2025;52(3):333-343. [doi: 10.1007/bf02294360]
87. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected Papers of Hirotugu Akaike. New York: Springer; 1998:199-213.
88. Akaike H. Factor analysis and AIC. Psychometrika 2025;52(3):317-332. [doi: 10.1007/bf02294359]
89. McCutcheon AL. Latent Class Analysis. USA: Sage; 1987.
90. McLachlan GJ, Peel D. Finite Mixture Models. USA: John Wiley & Sons; 2004.
91. Lo Y, Mendell NR, Rubin DB. Testing the number of components in a normal mixture. Biometrika 2001;88(3):767-778. [doi: 10.1093/biomet/88.3.767]
92. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Structural Equation Modeling: A Multidisciplinary Journal 2007;14(4):535-569. [doi: 10.1080/10705510701575396]
93. Touleimat N, Tost J. Complete pipeline for infinium(®) human methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics 2012;4(3):325-341. [doi: 10.2217/epi.12.21] [Medline: 22690668]
94. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics 2016;8(3):389-399 [FREE Full text] [doi: 10.2217/epi.15.114] [Medline: 26673039]
95. Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F. Evaluation of the infinium methylation 450K technology. Epigenomics 2011;3(6):771-784. [doi: 10.2217/epi.11.105] [Medline: 22126295]

XSL•FO
RenderX

96. Groen K, Lea RA, Maltby VE, Scott RJ, Lechner-Scott J. Letter to the editor: blood processing and sample storage have negligible effects on methylation. Clin Epigenetics 2018;10:22 [FREE Full text] [doi: 10.1186/s13148-018-0455-6] [Medline: 29456766]
97. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47 [FREE Full text] [doi: 10.1093/nar/gkv007] [Medline: 25605792]
98. Dessie EY, Gautam Y, Ding L, Altaye M, Beyene J, Mersha TB. Development and validation of asthma risk prediction models using co-expression gene modules and machine learning methods. Sci Rep 2023;13(1):11279 [FREE Full text] [doi: 10.1038/s41598-023-35866-2] [Medline: 37438356]
99. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559 [FREE Full text] [doi: 10.1186/1471-2105-9-559] [Medline: 19114008]
100. Phelan KJ, Dill-McFarland KA, Kothari A, Segnitz RM, Burkle J, Grashel B, et al. Airway transcriptome networks identify susceptibility to frequent asthma exacerbations in children. J Allergy Clin Immunol 2023;152(1):73-83 [FREE Full text] [doi: 10.1016/j.jaci.2023.02.031] [Medline: 36918038]
101. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology 2013;38(1):23-38 [FREE Full text] [doi: 10.1038/npp.2012.112] [Medline: 22781841]
102. MacKinnon DP, Luecken LJ. Statistical analysis for identifying mediating variables in public health dentistry interventions. J Public Health Dent 2011;71(Suppl 1):S37-S46 [FREE Full text] [doi: 10.1111/j.1752-7325.2011.00252.x] [Medline: 21656950]
103. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986;51(6):1173-1182. [doi: 10.1037//0022-3514.51.6.1173] [Medline: 3806354]
104. Jasti S, Dudley WN, Goldwater E. SAS macros for testing statistical mediation in data with binary mediators or outcomes. Nurs Res 2008;57(2):118-122. [doi: 10.1097/01.NNR.0000313479.55002.74] [Medline: 18347484]
105. Preacher KJ, Hayes AF. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behav Res Methods Instrum Comput 2004;36(4):717-731. [doi: 10.3758/bf03206553] [Medline: 15641418]
106. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods 2010;15(4):309-334. [doi: 10.1037/a0020761] [Medline: 20954780]
107. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
108. Braun V, Clarke V. Thematic Analysis: A Practical Guide. USA: Sage; 2022.
109. Dedoose Version 9.0.107, Cloud Application for Managing, Analyzing, and Presenting Qualitative and Mixed Method Research Data. Los Angeles, CA: SocioCultural Research Consultants, LLC; 2023.
110. Giacomini MK, Cook DJ. Users' guides to the medical literature: XXIII. Qualitative research in health care A. Are the results of the study valid? Evidence-based medicine working group. JAMA 2000;284(3):357-362 [FREE Full text] [doi: 10.1001/jama.284.3.357] [Medline: 10891968]
111. Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999;34(5 Pt 2):1189-1208 [FREE Full text] [Medline: 10591279]
112. Motulsky SL. Is member checking the gold standard of quality in qualitative research? Qualitative Psychology 2021;8(3):389-406. [doi: 10.1037/qup0000215]
113. Buchbinder E. Beyond checking: experiences of the validation interview. Qualitative Social Work 2010;10(1):106-122. [doi: 10.1177/1473325010370189]
114. Appel HB, Nguyen PD. Eliminating racial and ethnic disparities in behavioral healthcare in the United States. J Health Soc Sci 2021;7(1):52-56. [doi: 10.5114/jhi.2021.107008]
115. Dedania R, Gonzales G. Disparities in access to health care among US-born and foreign-born US adults by mental health status, 2013–2016. Am J Public Health 2019;109(S3):S221-S227 [FREE Full text] [doi: 10.2105/ajph.2019.305149] [Medline: 31242018]
116. Elshahat S, Moffat T, Newbold K. Understanding the healthy immigrant effect in the context of mental health challenges: a systematic critical review. J Immigr Minor Health 2022;24(6):1564-1579 [FREE Full text] [doi: 10.1007/s10903-021-01313-5] [Medline: 34807354]
117. Bucay-Harari L, Page KR, Krawczyk N, Robles YP, Castillo-Salgado C. Mental health needs of an emerging Latino community. J Behav Health Serv Res 2020;47(3):388-398 [FREE Full text] [doi: 10.1007/s11414-020-09688-3] [Medline: 32002728]
118. Giacco D, Matanov A, Priebe S. Providing mental healthcare to immigrants: current challenges and new strategies. Curr Opin Psychiatry 2014;27(4):282-288. [doi: 10.1097/YCO.0000000000000065] [Medline: 24840159]
119. Alegría M, Alvarez K, Ishikawa RZ, DiMarzio K, McPeck S. Removing obstacles to eliminating racial and ethnic disparities in behavioral health care. Health Aff (Millwood) 2016;35(6):991-999 [FREE Full text] [doi: 10.1377/hlthaff.2016.0029] [Medline: 27269014]

XSL•FO
RenderX

120. Alegría M, Zhen-Duan J, O'Malley IS, DiMarzio K. A new agenda for optimizing investments in community mental health and reducing disparities. Am J Psychiatry 2022;179(6):402-416 [FREE Full text] [doi: 10.1176/appi.ajp.21100970] [Medline: 35599537]
121. Van der Boor CF, White R. Barriers to accessing and negotiating mental health services in asylum seeking and refugee populations: the application of the candidacy framework. J Immigr Minor Health 2020;22(1):156-174 [FREE Full text] [doi: 10.1007/s10903-019-00929-y] [Medline: 31444614]
122. Ohtani A, Suzuki T, Takeuchi H, Uchida H. Language barriers and access to psychiatric care: a systematic review. Psychiatr Serv 2015;66(8):798-805 [FREE Full text] [doi: 10.1176/appi.ps.201400351] [Medline: 25930043]
123. Cabral J, Cuevas AG. Health inequities among Latinos/Hispanics: documentation status as a determinant of health. J Racial Ethn Health Disparities 2020;7(5):874-879 [FREE Full text] [doi: 10.1007/s40615-020-00710-0] [Medline: 32026286]
124. Rousseau C, Frounfelker RL. Mental health needs and services for migrants: an overview for primary care providers. J Travel Med 2019;26(2):tay150 [FREE Full text] [doi: 10.1093/jtm/tay150] [Medline: 30561687]
125. Khullar D, Chokshi DA. Challenges for immigrant health in the USA-the road to crisis. The Lancet 2019;393(10186):2168-2174 [FREE Full text] [doi: 10.1016/s0140-6736(19)30035-2] [Medline: 30981536]
126. Falgas I, Ramos Z, Herrera L, Qureshi A, Chavez L, Bonal C, et al. Barriers to and correlates of retention in behavioral health treatment among Latinos in 2 different host countries: the United States and Spain. J Public Health Manag Pract 2017;23(1):e20-e27 [FREE Full text] [doi: 10.1097/PHH.0000000000000391] [Medline: 26910867]
127. Fripp JA, Carlson RG. Exploring the influence of attitude and stigma on participation of African American and Latino populations in mental health services. J Multicult Couns & Deve 2017;45(2):80-94. [doi: 10.1002/jmcd.12066]
128. Mays VM, Jones A, Delany-Brumsey A, Coles C, Cochran SD. Perceived discrimination in health care and mental health/substance abuse treatment among blacks, Latinos, and whites. Med Care 2017;55(2):173-181 [FREE Full text] [doi: 10.1097/MLR.0000000000000638] [Medline: 27753743]
129. Mental health disparities: Hispanics and Latinos. American Psychiatric Association. Washington, DC: American Psychiatric Association; 2017. URL: https://www.psychiatry.org [accessed 2021-05-19]
130. Byrow Y, Pajak R, Specker P, Nickerson A. Perceptions of mental health and perceived barriers to mental health help-seeking amongst refugees: a systematic review. Clin Psychol Rev 2020;75:101812 [FREE Full text] [doi: 10.1016/j.cpr.2019.101812] [Medline: 31901882]
131. Chang CD. Social determinants of health and health disparities among immigrants and their children. Curr Probl Pediatr Adolesc Health Care 2019;49(1):23-30 [FREE Full text] [doi: 10.1016/j.cppeds.2018.11.009] [Medline: 30595524]
132. Satcher D. Mental Health: Culture, Race, and Ethnicity-A Supplement to Mental Health: A Report of the Surgeon General. Rockville (MD): US Department of Health and Human Services; 2001.
133. Grecco GG, Chambers RA. The Penrose effect and its acceleration by the war on drugs: a crisis of untranslated neuroscience and untreated addiction and mental illness. Transl Psychiatry 2019;9(1):320 [FREE Full text] [doi: 10.1038/s41398-019-0661-9] [Medline: 31780638]
134. Kessler RC, Berglund PA, Bruce ML, Koch JR, Laska EM, Leaf PJ, et al. The prevalence and correlates of untreated serious mental illness. Health Serv Res 2001;36(6 Pt 1):987-1007 [FREE Full text] [Medline: 11775672]
135. Bijl RV, de Graaf R, Hiripi E, Kessler RC, Kohn R, Offord DR, et al. The prevalence of treated and untreated mental disorders in five countries. Health Aff (Millwood) 2003;22(3):122-133 [FREE Full text] [doi: 10.1377/hlthaff.22.3.122] [Medline: 12757277]
136. Garcia ME, Hinton L, Gregorich SE, Livaudais-Toman J, Kaplan C, Karliner L. Unmet mental health need among Chinese and Latino primary care patients: intersection of ethnicity, gender, and English proficiency. J Gen Intern Med 2020;35(4):1245-1251 [FREE Full text] [doi: 10.1007/s11606-019-05483-9] [Medline: 31667737]
137. Lee M, Bhimla A, Lu W, Ma GX. Correlates of mental health treatment receipt among Asian Americans with perceived mental health problems. J Behav Health Serv Res 2021;48(2):199-212 [FREE Full text] [doi: 10.1007/s11414-020-09704-6] [Medline: 32347427]
138. Behrman P, Fitzgibbon M, Buscemi J, Sanchez-Johnsen L, Laroche H, Yanez B, et al. Society of behavioral medicine position statement: congress should protect immigrants seeking health care. Transl Behav Med 2019;9(6):1244-1247 [FREE Full text] [doi: 10.1093/tbm/ibz119] [Medline: 31367753]
139. Lee M, Lu W, Mann-Barnes T, Nam J, Nelson J, Ma G. Mental health screening needs and preference in treatment types and providers in African American and Asian American older adults. Brain Sci 2021;11(5):597 [FREE Full text] [doi: 10.3390/brainsci11050597] [Medline: 34063018]
140. Brenes F. Hispanics, mental health, and the interpersonal-psychological theory of suicide: brief report. Hisp Health Care Int 2021;19(2):77-81 [FREE Full text] [doi: 10.1177/1540415320951489] [Medline: 32844688]
141. Alegría M, Nakash O, NeMoyer A. Increasing equity in access to mental health care: a critical first step in improving service quality. World Psychiatry 2018;17(1):43-44 [FREE Full text] [doi: 10.1002/wps.20486] [Medline: 29352534]
142. Vaughn LM, Whetstone C, Boards A, Busch MD, Magnusson M, Määttä S. Partnering with insiders: a review of peer models across community-engaged research, education and social care. Health Soc Care Community 2018;26(6):769-786 [FREE Full text] [doi: 10.1111/hsc.12562] [Medline: 29512217]

XSL•FO
RenderX

Abbreviations

BP: blood pressure
CAB: Community Advisory Board
CpG: cytosine-phosphate-guanine
DNAm: DNA methylation
ENCODE: Encyclopedia of DNA Elements
GAD-7: 7-item Generalized Anxiety Disorder
HbA1c: hemoglobin A1c
HIPAA: Health Insurance Portability and Accountability Act
LPA: latent profile analysis
PHQ-8: 8-item Patient Health Questionnaire
PROMIS: Patient-Reported Outcomes Measurement Information System
PSS-10: 10-item Perceived Stress Scale
REDCap: Research Electronic Data Capture
SOIS: Stress of Immigration Survey
WGCNA: weighted gene co-expression network analysis

Edited by A Schwartz; submitted 15.Oct.2025; this is a non–peer-reviewed article;accepted 19.Jan.2026; published 10.Feb.2026.

Please cite as:
Stone L, Ding L, Hirpa Kassa M, Amsalu B, Getnet B, Eshete H, Mersha TB, Vaughn LM
Cumulative Stress Burden and Association With DNA Methylation in Ethiopian American Immigrants: Protocol for a Community-Engaged, Biopsychosocial Study
JMIR Res Protoc 2026;15:e85971
URL: https://www.researchprotocols.org/2026/1/e85971
doi:10.2196/85971
PMID:

©Lara Stone, Lili Ding, Meron Hirpa Kassa, Beakal Amsalu, Bruktawit Getnet, Hermella Eshete, Tesfaye B Mersha, Lisa M Vaughn. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Migration Effects on Cognition: Protocol for the Aging in Kerala Americans Research Study

Kelly Cotton[1], PhD; Dristi Adhikari[1], MA; Anne Felicia Ambrose[1], MBBS; Emmeline Ayers[1], MPH; Helena M Blumen[1], PhD; Mirnova E Ceïde[2,3], MD; VG Pradeep Kumar[4], DM; Mairim Melecio-Vazquez[2], MA; Sanish Sathyan[1,5], PhD; Alben Sigamani[6], MBBS; Kavita Sivaramakrishnan[7], PhD; Marnina Stimmel[2], PhD; Erica F Weiss[2], PhD; Jie Yang[8], PhD; Joe Verghese[1], MBBS

[1]Department of Neurology, Renaissance School of Medicine, Stony Brook University, Stony Brook, NY, United States

[2]Department of Neurology, Albert Einstein College of Medicine, Bronx, NY, United States

[3]Department of Psychiatry and Behavioral Sciences, Albert Einstein College of Medicine, Bronx, NY, United States

[4]Institute of Neurosciences, Baby Memorial Hospital, Kozhikode, India

[5]Department of Molecular Biology, Kannur University, Kannur, India

[6]Carmel Research Consultancy Pvt Ltd, Bengaluru, India

[7]Department of Sociomedical Sciences, Mailman School of Public Health, Columbia University, New York, NY, United States

[8]Department of Family, Population & Preventive Medicine, Renaissance School of Medicine, Stony Brook University, Stony Brook, NY, United States

Corresponding Author:
Kelly Cotton, PhD
Department of Neurology
Renaissance School of Medicine
Stony Brook University
100 Nicolls Road
Stony Brook, NY, 11794
United States
Phone: 1 631 689 8333
Email: kelly.cotton@stonybrookmedicine.edu

Abstract

Background: In the United States, Asian American people represent the fastest growing population group, and are highly diverse linguistically, culturally, and demographically. Yet, in most national studies, Asian American groups are aggregated, masking potential health disparities. Racial and ethnic minorities, especially first-generation immigrants, are also at a particularly elevated risk of cognitive impairment.

Objective: The Aging in Kerala Americans Research (AKKARE) study aims to examine both positive and negative migration effects on health in the first-generation Kerala American population, focusing on cognition and dementia. We will assess the effect of immigrant and cultural factors and social relations on cognitive aging from epidemiological, biological, and vascular perspectives. This protocol describes the study design and procedures for the AKKARE study.

Methods: The AKKARE study proposes to enroll 400 older first-generation Kerala American individuals from the tristate area. A smaller subset of these participants will complete blood tests (n=360) and neuroimaging studies (n=160). We will assess the role of immigration and cultural effects on cognitive function, mood, and quality of life, as well as biological and vascular aging. We will conduct follow-up assessments at 12-month intervals for up to 5 years.

Results: The AKKARE study (grant #1R01AG084567-01) was funded by the US National Institutes of Health in 2024 and received approval from the Stony Brook University Institutional Review Board to start the study in 2025. Enrollment began in September 2025.

Conclusions: As there is presently a lack of fundamental data on the epidemiology in diseases of aging in Indian American immigrants, the AKKARE study will provide new insights into factors of risk and resilience associated with cognitive impairment in this group and in the broader older adult population.

International Registered Report Identifier (IRRID): DERR1-10.2196/85493

(JMIR Res Protoc 2026;15:e85493) doi:10.2196/85493

XSL•FO
RenderX

KEYWORDS

immigration; India; cognitive impairment; neuroimaging; aging; dementia; acculturation; biomarkers

Introduction

Background

In the United States, Asian American people represent the fastest-growing population group, more than doubling in size from 10.5 million in 2000 to more than 23 million in 2020. Asian American people trace their ancestry to at least 19 countries in East and Southeast Asia and the Indian subcontinent. The Asian American population (including Indian American people) is highly diverse linguistically, culturally, and demographically. In most national surveys and studies, however, these groups are aggregated, masking health disparities in specific groups.

The Indian American population is the second-largest immigrant group in the United States. The enactment of the 1965 Immigration and Naturalization law led to an uptick in immigration to the United States from India, and the Indian American population in the United States has increased dramatically, from 3000 in the 1960s to 4.6 million people in 2019 [1-3]. Immigrants from the state of Kerala in South India account for approximately 7% of the current Indian American population in the United States [1], particularly in the tristate area (New York, New Jersey, and Connecticut) [3]. Kerala has low per capita income (US $2775/year) and high unemployment (6%), but is on par with many developed countries in education (literacy rate 94%) and health [4,5], a paradox termed the "Kerala phenomenon" [4,6]. Our interviews and experience with Kerala American individuals reveal 2 major immigrant groups: one group who migrated to the United States in middle adulthood for education and job opportunities and have aged in place, and another group who migrated at older ages, following retirement in India, to join their children in the United States. Unlike the first group of migrants, the older second group has had less in-country time to establish social relations and assimilate the local culture. Hence, these 2 immigrant groups from Kerala likely vary in migrant experiences and social determinants of health.

The incidence of mild cognitive impairment and Alzheimer disease is high among racial or ethnic minorities, especially first-generation immigrants [7-10], and immigrant populations spend a greater proportion of their later years with cognitive impairment and dementia than native-born US populations [8,11]. Psychological stress, depression, and other migration-related factors (eg, age, acculturation, and social relations) may accelerate cognitive decline in migrants [7,12], but the situation is complex. Both worse health and better health outcomes ("healthy migrant effect") are seen in migrants; however, benefits appear to wane with longer US residency [10,12-14]. Still, a 2021 trans–National Institutes of Health workshop on the Asian American population noted that there is a paucity of fundamental data on the epidemiology as well as risk and resilience factors in diseases of aging in this large US population segment.

Here, we propose the Aging in Kerala Americans Research (AKKARE) study. In Malayalam, the language spoken by most people in Kerala, the word *akkare* translates to "on the other bank (shore)," and is also used to refer to overseas. In the AKKARE study, we aim to examine the effects, both positive and negative, of migrating to "the other bank" on health in first-generation Kerala American individuals, focusing on cognition and dementia. We will use a social-cultural conceptual model (Figure 1) to address migration effects on cognition that incorporates Zimmerman's stages of migration (predeparture, travel, and destination phases; Table 1) [15] as well as social determinants of health. This conceptual visual model is meant to organize our study approach and to identify relationships between risk and protective factors for cognitive health in Kerala American individuals. We will focus on the effect of immigrant and cultural factors and social relations on cognitive aging from epidemiological, biological, and vascular perspectives.

XSL•FO
RenderX

Figure 1. Conceptual model.

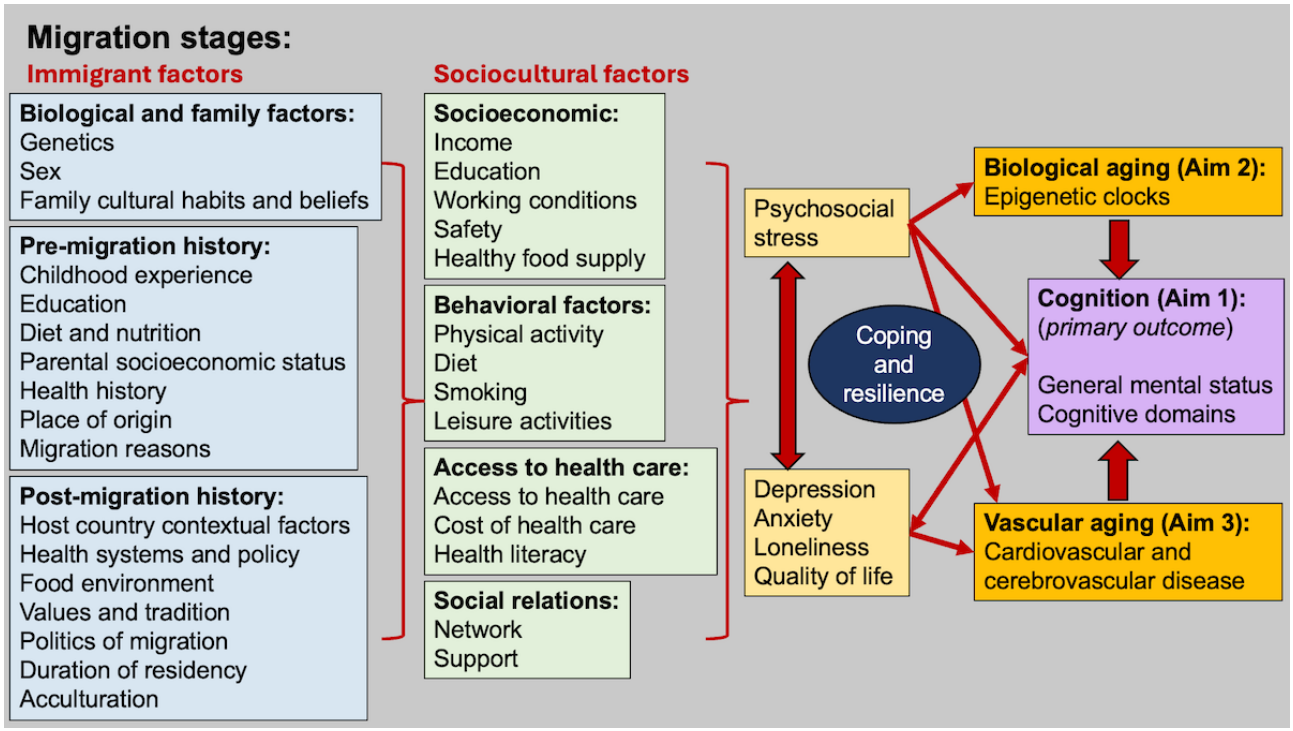

XSL•FO
RenderX

Table 1. Measures and analyses by migration phase for each aim.

Migration phase	Predeparture	Destination	Predeparture	Destination	Predeparture	Destination
Study aim	Aim 1: identify immigrant and cultural factors and social relationships that contribute to dementia risk in older Kerala American people	Aim 1: identify immigrant and cultural factors and social relationships that contribute to dementia risk in older Kerala American people	Aim 2: determine the contribution of immigrant and cultural factors and social relationships to biological aging in older Kerala American people	Aim 2: determine the contribution of immigrant and cultural factors and social relationships to biological aging in older Kerala American people	Aim 3: determine the impact of immigrant and cultural factors and social relationships on vascular aging in older Kerala American people	Aim 3: determine the impact of immigrant and cultural factors and social relationships on vascular aging in older Kerala American people
Outcome	Global cognition, individual cognitive domains, mood, and quality of life	Global cognition, individual cognitive domains, mood, and quality of life	AgeAccel (biological aging)	AgeAccel (biological aging)	Blood pressure, ECG[a], and small vessel disease	Blood pressure, ECG, and small vessel disease
Exposures	Age at migration, migration reasons, and marital status	Residency period in the United States, percent of life spent in the United States, occupation, acculturation, discrimination, stress, social network, perceived social support, and loneliness	Age at migration, migration reasons, and marital status	Residency period in the United States, percent of life spent in the United States, occupation, acculturation, discrimination, stress, social network, perceived social support, and loneliness	Age at migration, migration reasons, and marital status	Residency period in the United States, percent of life spent in the United States, occupation, acculturation, discrimination, stress, social network, perceived social support, and loneliness
Planned contrasts	Middle vs late-life migrants	Middle vs late-life migrants (United States vs Kerala cohorts)	Middle vs late-life migrants	Middle vs late-life migrants	Middle vs late-life migrants	Middle vs late-life migrants (United States vs Kerala cohorts)
Moderators	Diet, physical activity, leisure activity, smoking or alcohol use, and functional status	Diet, physical activity, leisure activity, smoking or alcohol use, and functional status	—[b]	—	—	—
Mediators	Biological aging and vascular markers	Biological aging and vascular markers	—	—	—	—
Covariates	—[c]	—[c]	—[c]	—[c]	—[c]	—[c]

[a]ECG: electrocardiogram.

[b]Not applicable.

[c]Shared across aims: age, sex, education, and chronic medical illness.

Study Aims

There are 3 main aims of the AKKARE study. First, the study aims to identify immigrant and cultural factors and social relationships (such as social networks and social support) that contribute to the risk of dementia in older Kerala American individuals. We will assess the effects of immigrant and cultural factors (such as age at migration, residency duration, migration reasons, sex, education, and acculturation) on global cognition, as well as individual cognitive domains, mood, and quality of life. Furthermore, to address the "healthy migrant" effect [8], we will examine coping and resilience mechanisms that may help preserve cognition. The second aim is to determine the contribution of immigrant and cultural factors and social relationships to biological aging (quantified using biological clocks) in older Kerala American individuals. Third, the study aims to determine the impact of immigrant and cultural factors and social relationships on vascular aging in older Kerala American individuals. We will compare vascular aging (eg, cardiovascular and cerebrovascular indices) between the US and Kerala cohorts.

Methods

Study Design and Setting

We propose a research study of 400 older first-generation Kerala American individuals from the tristate area (New York, New Jersey, and Connecticut). A subset of participants will complete blood tests (n=360) and neuroimaging studies (n=160). We will also conduct follow-up assessments at 12-month intervals for up to 5 years.

XSL•FO
RenderX

Eligibility Criteria

The general inclusion criteria include (1) age 60 years or older, (2) first-generation immigrant with at least 1 year US residency and currently living in the United States most (at least 6 months) of the year, (3) speaks Malayalam or English at a level sufficient to complete assessments, and (4) 1 participant per household to avoid bias on social measures.

The general exclusion criteria include (1) presence of dementia (Montreal Cognitive Assessment Blind/Telephone <10) [16]; inability to complete activities of daily living [17], or previous dementia diagnosis); (2) terminal illness with life expectancy <12 months, (3) progressive, degenerative, neurologic disease (eg, Parkinson disease and amyotrophic lateral sclerosis); and (4) major psychiatric symptoms, such as delusions or psychosis that will prevent assessments. Additionally, for participants who agree to participate in neuroimaging studies, the standard neuroimaging exclusion criteria are (1) presence of dementia; (2) presence of any surgically implanted metallic devices, such as aneurysm clips or pacemakers, and large amounts of dental or surgical hardware in the head and neck; and (3) a known history of claustrophobia.

Recruitment

We will implement recruitment strategies with proven success in our center: outreach to Kerala American cultural and religious organizations, health fairs for the Kerala community, referrals from local clinicians, and targeted advertisements in Malayalam language print and online media. We will appoint a bilingual community outreach member to do outreach, recruitment, and engagement. The bilingual outreach staff member will engage with the community (eg, conducting needs assessment with participants, facilitating community engagement studios, and reaching out to community organizations), publicize the study, and promote enrollment.

Assessment Procedures

Potential participants will be sent a letter describing the study, followed by a telephone call from bilingual research staff to assess interest and conduct a preliminary eligibility screen. The telephone interview includes verbal consent, confirmation of Kerala origin and US residency, and brief dementia screeners. We will arrange a home or clinic visit by our staff for potential recruits to obtain written informed consent and complete in-person assessments. We chose to offer the option of home-based or clinic-based testing for the convenience of participants, to build rapport, and to remove logistical barriers, such as travel, that may influence participation and retention. We will offer flexibility in the completion of self-report measures and questionnaires, and participants will complete these surveys on their own or remotely using technology (phone or internet), based on their preference. We will partner with a mobile phlebotomy and clinical exam company for participants to complete physical assessments, blood draw, and an electrocardiogram (ECG) on a separate day from the cognitive assessments. Transportation will be provided for a third visit to our research center for magnetic resonance imaging (MRI) studies. A subset of 15-20 participants, as well as community leaders, will also be invited to complete an additional day where participants will undergo interviews focused on ethnography and life course history. Participants will be compensated for their time.

Measures

Sociodemographic

We will collect basic sociodemographic information (age, sex, education, income, and occupation) using a questionnaire. We will assess language use with the Language Experience and Proficiency Questionnaire [18], a valid and reliable tool for assessing language profiles of multilingual adults.

Cognition

We will assess cognition with validated tests in Malayalam and English, depending on participant preference. Cognitive testing will be administered by in-person bilingual research assistants. These tests overlap with the Kerala Einstein Study (KES), a cohort study of older adults residing in Kerala [19], facilitating cross-national comparisons.

Global cognition will be assessed with Addenbrooke's Cognitive Examination (ACE)-III [20], a cognitive screening tool that incorporates the Mini-Mental State Examination as well as additional memory, language, visuospatial, and verbal fluency components, without specialized equipment, an advantage for home assessments. Scores on the ACE range from 0 to 100. The internal consistency, test-retest reliability, and interrater reliability of ACE were reported to be excellent worldwide [21,22]. The ACE predicts incident Alzheimer disease [23] and discriminates cognitively normal from mild cognitive impairment [20] and dementia [23,24]. The Malayalam version of the ACE has been previously validated [21,25].

We will also complete more robust assessments of individual cognitive domains, including memory, processing speed, executive function and attention, visuospatial reasoning, and verbal fluency (Table 2). Although participants will be able to choose their preference for language of interaction and assessment, we will evaluate and, if necessary, adjust for language proficiency in bilingual individuals by completing verbal fluency measures in both languages.

XSL•FO
RenderX

Table 2. Cognitive domains and associated tests.

Domain	Test
Global cognitive screener	• Addenbrooke's Cognitive Examination-III [20,21]
Memory	• Verbal Learning Test[a] [26,27] • Picture Memory Impairment Screen [28,29] • Modified Taylor Complex Figure, Recall Trials[a] [26,30]
Processing speed	• Symbol Match [31] • Digit Symbol Substitution Test [32] • Trail-Making (Black and White), Trial 1[a] [26,33]
Executive function and attention	• Digit Span [25,32] • Trail-Making (Black and White), Trial 2[a] [26,33]
Visuospatial reasoning	• Modified Taylor Complex Figure copy[a] [26,30]
Verbal fluency	• Letter Fluency[a] [26,34] • Category Fluency[a] [26]

[a]Test from the Indian Council of Medical Research-Neurocognitive Toolbox.

Mood

We will use the Geriatric Depression Scale (GDS)-15 [35] and Generalized Anxiety Disorder-7 item (GAD-7) [36] to assess symptoms of depression and anxiety, respectively. The Malayalam versions of the GDS and GAD-7 have been previously validated [37,38].

Quality of Life

Quality of life will be measured with the 12-Item Short-Form Health Survey, a well-validated and reliable scale [39]. The scale results in 2 summary scores—a mental component score and a physical component score.

Immigrant and Cultural Factors

Premigrant and postmigration factors will be collected. Premigration factors include sex, education, marital status, and migration reason (education, economic, or to join family). Postmigration factors include residency period in the United States, percent of life spent in the United States (years of residence in the United States divided by age), occupation, acculturation, discrimination, and stress. We will assess acculturation with the Mediators of Atherosclerosis in South Asians Living in America acculturation questionnaire, which was developed in South Asian immigrants [40]. The questionnaire asks participants, "How much would you wish these traditions from Kerala would be practiced in America?" with 7 items, including religious ceremonies, serving sweets at ceremonies, fasting on specific occasions, living in a joint family, arranged marriage, eating a traditional diet, and using spices for health and healing. Higher scores represent weaker traditional Indian beliefs. High Cronbach α (0.83) has been reported with similar reliability in men and women. Based on this scale, acculturation is coded into one of three categories: (1) assimilation (preference for US culture), (2) separation (preference for Kerala culture), and (3) integration (similar preference for both cultures).

We will measure perceived discrimination using the Everyday Discrimination Scale, a highly reliable scale [41]. The scale has 9 items that ask about mistreatment a participant may have experienced, such as "people act as if you are dishonest" and "you are threatened or harassed." Participants respond to each item using a 6-point scale with options ranging from "never" (0) to "almost every day" (6).

Life stress will be measured using the Presumptive Stressful Life Events scale [42]. This scale includes 51 items about the occurrence of a variety of events in the past year and in their lifetime, such as having financial difficulties, family problems, job problems, or problems with people outside the family; suffering serious injury or the death of a spouse or friend; being a victim of a crime; and moving to a new home. Higher scores indicate greater life stress.

Social Relations

The Social Network Index (SNI) [43] will be used to quantify the number of high-contact social roles (SNI-1) a person interacts with at least biweekly and the total number of network members (SNI-2). A person can have a maximum of 12 high-contact social roles, including spouse, parent, child, child-in-law, close relative, close friend, religious group member, student, employee, neighbor, volunteer, and group member. The total number of network members is obtained by summing up the number of people a person interacts with at least biweekly across these 12 roles. Few high-contact social roles are defined as 4 or fewer social relationships that a person communicates with biweekly or more. A small network is defined as 18 or fewer social relationships.

We will assess perceived social support using the Medical Outcomes Study Social Support Survey [44], a validated and reliable social support scale that distinguishes between emotional or informational support, tangible support, affectionate support, and positive social interactions. The scale includes 19 items, such as "someone to share your most private worries and fears

XSL•FO
RenderX

with" (emotional or informational), "someone to help you if you were confined to bed" (tangible), "someone who hugs you" (affectionate), and "someone to do something enjoyable with" (positive social interaction). For each item, participants are asked to indicate how often each type of support is available if needed, with options ranging from "None of the time" (1) to "All of the time" (5). Higher scores indicate higher levels of social support.

Finally, we will measure loneliness using the 3-item UCLA Loneliness scale [45,46]. This scale includes 3 items that assess 3 dimensions of loneliness: relational connectedness, social connectedness, and self-perceived isolation. For each item, participants indicate how often that item applies to them, (1) "Hardly ever," (2) "Some of the time," or (3) "Often." Higher scores indicate greater loneliness.

Health and Lifestyle

We will collect information on medical history, medication use, and health care access. We will collect physical measurements (height, weight, and waist or hip circumference). We will assess health literacy using the Short Assessment of Health Literacy [47].

Traumatic brain injury history will be assessed using the Kerala Brain Injury Questionnaire, which is a reliable and valid assessment of chronic postconcussive symptoms in older Indian adults [48]. The Kerala Brain Injury Questionnaire can be used by nonclinicians, making it suitable for widespread case detection in community settings and for prescreening individuals before clinic visits, helping clinicians identify potential undiagnosed health issues.

We will assess function and disability with the Kerala Instrumental Activities of Daily Living scale [49,50], which has been shown to have excellent internal consistency and discriminative validity in Kerala [49,50].

We will assess cognitive and physical leisure activities using a previously validated scale [51,52]. We will collect information on current, past, and lifetime tobacco, marijuana, and alcohol use. We will assess dietary habits by asking participants to recall everything they ate or drank in the previous 24 hours, broken down by meal.

Physical Assessments

We will measure gait speed using a timed walk at a normal pace over a fixed distance (2.5 m). Use of any walking aid will be noted.

We will assess grip strength using a Jamar dynamometer. After identifying the dominant hand, the maximum voluntary contraction in the dominant hand will be measured.

Additional Questionnaires

We will assess subjective cognitive change using the Cognitive Change Index, which is a valid and consistent measure of self-perceived cognitive decline [53]. We will assess subjective change in memory with the 12 Cognitive Change Index items that assess memory performance (eg, "recalling information when I really try"). Participants are asked to compare their present ability to that of 5 years ago, rating each item from 1 ("no change/normal ability") to 5 ("much worse/severe problem"). Higher scores indicate greater subjective cognitive decline. Participants will also complete a subjective cognitive and motor complaint questionnaire, which asks participants about mobility (eg, "How far can you walk in an hour?") and cognitive ability (eg, "Daily problems with thinking and/or memory").

Coping and Resilience Mechanisms

We will conduct semistructured topic-guide interviews that focus on life course histories and address topics such as acculturative stress, language history, and coping [54-56]. A bilingual interviewer will conduct hour-long home-based interviews with 15-20 Kerala American people, including participants, families, and community representatives or leaders, on days and timings of their convenience. Sampling with maximum diversity will be done, and people with different views will be selected. These interviews will be done on a separate day from the study visits. The interviews will be digitally recorded and transcribed verbatim.

Biological Aging

We will quantify biological aging using a DNA methylation (DNAm) clock. The clock is built from epigenetic DNAm markers that can accurately quantify an age-related phenotype or outcome, or both [57-61]. We will mainly use second-generation DNAm aging algorithms (GrimAge and PhenoAge), but we will adapt for all available epigenetic clocks, including first-generation clocks (Horvath [353CpGs] and Hannum Clocks [71CpGs] as well as any new clocks developed in the future with methylation data. DNAm clocks are built with a supervised machine learning method, such as a penalized regression trained against chronological age or on health-related outcomes to identify an informative and sparse predictive set of CpGs [61]. GrimAge was developed based on the 7 DNAm surrogates of plasma proteins (DNAm adrenomedullin levels, DNAm beta-2 microglobulin, DNAm cystatin C, DNAm growth differentiation factor 15, DNAm leptin, DNAm plasminogen activation inhibitor 1, and DNAm tissue inhibitor metalloproteinase 1) and smoking pack years in a 2-stage procedure. DNAm GrimAge is based on the 1030 CpGs methylation level. The GrimAge version 2 leverages 2 additional DNAm-based estimators of plasma proteins: high sensitivity log C-reactive protein and log hemoglobin A1C, and will also be tested in our participants. DNAm phenotypic age is another successful second-generation clock predicting aging outcomes, including all-cause mortality, health span, and Alzheimer disease. The epigenetic age calculation is carried out using the DNAm age website [62] or by using the R package (R Core Team), methylclock [63]. Epigenome-wide DNAm data are proposed to be generated using the Infinium Human MethylationEPIC v2.0 BeadChip (Illumina Inc). However, with ongoing advancements in methylation microarray technology, we plan to use the most updated array available at the time of profiling to ensure comprehensive and high-resolution data capture.

AgeAccel is defined as the residual resulting from regressing predicted age on chronological age, and is independent of chronological age [61,64]. A positive AgeAccel value indicates

XSL•FO
RenderX

faster aging, and a negative value suggests slower aging. This measure is derived from clocks described above and will be used for our analyses.

Vascular Aging

We will assess vascular aging with 2 indices: cardiovascular (blood pressure and ECG) and cerebrovascular (small vessel disease). We will measure blood pressure as per the Systolic Blood Pressure Intervention Trial (SPRINT) protocol [65]. The participant will be positioned in a seated position with back support and arm at heart level. The trained mobile phlebotomists will measure blood pressure 3 times with an appropriately sized cuff after 3-5 minutes of quiet rest, and readings will be averaged. Blood pressure will be analyzed as a continuous value and using established high and low cut-scores.

We will use a 12-lead ECG at home to maximize data collection. This will be conducted on the same day as the blood draw. The 12-lead ECG systems are practical for home assessments as they are digital, compact, and portable. ECG traces will undergo technical review and be codified with Minnesota code classification as reported in KES [66]. ECG outcomes include left ventricular hypertrophy, abnormality of ST-segment and T-wave, and abnormal rhythms. All ECGs will be read and assessed for normality/abnormality by a trained physician. Participants will be informed of abnormal results.

The cerebrovascular index will focus on small vessel disease and includes lacunar infarctions, white matter hyperintensities, and microbleeds. Details of this measure are discussed further in this study.

Neuroimaging: MRI Protocol and Assessment

We will assess cerebral small vessel disease with a focus on lacunar infarctions, white matter hyperintensities, and microbleeds. We will additionally measure cortical volume and thickness and white matter integrity to generate new hypotheses.

Image Acquisition

All MRI images will be acquired on a 3 Siemens PRISMA 3 Tesla Scanner (Siemens Healthineers). T1-weighted 3D-MP-RAGE images will be acquired over a 256 mm field of view (FOV) with echo time (TE)=2.49 milliseconds, repetition time (TR)=2260.0 milliseconds, 2.0 SENSE factor, and 1.0 mm isotropic resolution. Diffusion-weighted images will be acquired in 64 independent noncollinear directions over a 256 mm FOV with TE=69.0 milliseconds, TR=4200 milliseconds, and 2 mm slice thickness. 3D-FLAIR (fluid attenuating inverse recovery) will also be acquired over a 240 mm FOV with TE=326 milliseconds, TR=4300 milliseconds, and 1 mm slice thickness. 3D susceptibility-weighted images will be acquired with a TE=20.20 milliseconds, TR=27 milliseconds, 64 slices, and 0.86 voxel size. Finally, a resting-state functional MRI (or blood oxygenation level dependent) sequence will be acquired over a 224 mm FOV with TE=28.00 milliseconds, TR=2000 milliseconds, 2.0 mm slice thickness, and a multiband acceleration factor of 2.

Lacunar Infarctions

Lacunar infarctions are focal hyperintensities on T2-weighted or FLAIR images ≥3 mm in diameter. The whole brain will be visually searched for lacunar infarcts. Infarcts located in white matter must be hypointense on T1-weighted and FLAIR images to distinguish from white matter hyperintensities (discussed further in this study).

White Matter Hyperintensities

White matter hyperintensity (WMH) refers to the appearance on MRI of damage in white matter regions of the brain, and will be automatically quantified from 3D-FLAIR images using the lesion segmentation toolbox, which will be implemented with SPM12 (Wellcome Department of Imaging Neuroscience at University College London) or MATLAB (MathWorks) [67]. The lesion segmentation toolbox pipeline provides overall WMH lesion count and lesion volume, as well as a voxel-based WMH probability map that can be entered into subsequent analyses.

Cortical Microbleeds

Cortical microbleeds are small foci of chronic blood products in normal (or near normal) brain tissue. Microbleed Anatomical Rating Scale [68] will be applied to susceptibility-weighted images to rate definite and possible microbleeds in lobar, deep, and infratentorial regions [69].

Structural MRI

T1-weighted images will be parcellated into anatomical brain regions, which is important for examining cortical thickness and volume and a necessary step in processing data for examining structural connectivity (refer to "Diffusion-Weighted Imaging" section). Each T1-weighted image will be reconstructed using FreeSurfer (version 7.2; Martinos Center for Biomedical Imaging). FreeSurfer's subcortical segmentation and cortical parcellation are comparable with manual labeling [70]. Each participant's white and gray matter boundaries, as well as gray matter and cerebral spinal fluid boundaries, will be visually inspected slice by slice. The subcortical structure borders will be plotted with FreeView visualization tools (part of the FreeSurfer package), compared with actual brain regions, and manually corrected if necessary. Gray matter segmentation and longitudinal alignment will be performed with the FreeSurfer longitudinal pipeline, which is robust to initialization points and avoids biasing toward any 1 time point by generating a median template for each participant based on the T1-weighted images at both time points using cubic spline interpolation. Several processing steps, such as skull stripping, Talairach transforms, atlas registration, as well as spherical surface maps and parcellations, will then be initialized with common information from the subject-specific template, significantly increasing reliability and statistical power. Gray matter volumes and cortical thickness of 68 brain regions will be extracted from this FreeSurfer processing pipeline at each time point and then entered into subsequent statistical models. FreeSurfer 7.2 also provides additional measures that can be explored beyond the specific aims, including volumes of hippocampal and amygdalar subfields, subcortical and brain stem structures, and hypothalamic nuclei.

Diffusion-Weighted Imaging

Preprocessing will consist of field distortion corrections using top-up distributed as part of FSL (Functional Magnetic Resonance Imaging of the Brain Software Library) and eddy

XSL•FO
RenderX

current and movement correction using the eddy tool in FSL. To evaluate any white matter changes, we will use TRACULA (Tracts Constrained by Underlying Anatomy) distributed as part of the FreeSurfer v.7.2 library [71,72], which uses probabilistic tractography to extract 42 major white matter tracts. The software performs informed automatic tractography by incorporating anatomical information from a training dataset, provided by the software, with the anatomical segmentation of the T1 image of the current dataset, thus increasing the accuracy of the white matter tract placement for each participant by incorporating each participant's anatomical data into the tractography algorithm. Parcellation results from the FreeSurfer longitudinal stream will be applied to TRACULA to increase sensitivity to longitudinal changes in white matter tracts. The software outputs white matter integrity measures for each voxel inside the 42 tracts, with a mean of about 500 voxels per tract. Thus, for each participant, voxel-wise white matter integrity measures for 42 tracts at baseline and follow-up will be used in subsequent analyses.

Blood Collection

The phlebotomy service will be provided with the names, addresses, and phone numbers of consenting participants and will be responsible for contacting them to schedule blood draw appointments. Collection materials will be mailed to participants' homes in advance of the scheduled visit. Every effort will be made to conduct the blood draw within the same week or the following week of the interview. Fasting will be recommended and preferred, but not required. Phlebotomists will record the fasting status at the time of collection. The phlebotomy service will maintain regular communication with the study coordinator, reporting participants who decline to schedule appointments or miss scheduled visits. Follow-up efforts will be made to identify potential issues and to reschedule appointments when possible. Blood samples will be collected using both ethylenediaminetetraacetic acid (EDTA) tubes and serum separation tubes. EDTA tubes will be used for plasma isolation and for extracting DNA, which will be used for performing a methylation array for epigenetic aging clocks and other DNA-based analyses in the future. Phlebotomists will ship aliquots of plasma, serum, and EDTA whole blood overnight to the Stony Brook University *Freezer Farm*, a core facility that will ensure safe and reliable storage of samples at –80 °C for future research use. Serum and plasma will be stored for a wide range of biomarker analyses, including proteomic, metabolomic, and biochemical assays in the future. DNA will be extracted from EDTA whole blood using QIAamp DNA Blood Mini Kits (Qiagen), followed by quality assessment and quantification using the Agilent 2200 TapeStation or the Invitrogen Qubit fluorometer. Extracted DNA will also be stored at –80 °C until further analysis.

Data Processing and Analysis

Quantitative Analyses

All analyses will be preceded by extensive data checking and verification to resolve any missing, inconsistent, or out-of-range values, followed by standard descriptive analyses to summarize variables. Transformation of continuous biomarkers will be considered if severe skewness in distributions is detected.

For aim 1, the primary analyses will evaluate cross-sectional associations of immigrant and cultural and social factors with global cognition (ACE-III score) and tests of episodic memory, executive function, and attention using a series of linear regression models and reported as adjusted beta estimates with 95% CIs. The predictors will be included as continuous scores or categorical variables. All models will be adjusted for age, sex, education, chronic medical illness, and other confounders. We will first examine predictors separately, as these relationships have not been established in the Kerala American population.

As a hypothesis-generating analysis for future directions, we will evaluate whether loneliness, diet, physical activity, and leisure activity moderate the association between immigration related predictors and cognition, by adding these variables and their interactions with the predictors in the linear regression models described above. We will perform mediation analysis using the product of coefficients method to test if any significant associations of immigrant, cultural, and social factors with cognition can be explained by a mediating effect of biological aging or vascular markers. We will also address sex in our analyses, both adjusting for it as a covariate and conducting stratified analysis by sex. Finally, we will conduct exploratory analyses to examine the association of predictors at baseline with longitudinal change on cognitive measures using linear mixed effects models adjusted for confounders. Given the sample size and limited follow-up, these analyses are exploratory and intended to generate new hypotheses.

To address the potential problems arising from multiple hypotheses being tested, we will examine correlations among the outcome variables and among the predictors, and use principal component analysis to reduce the dimensionality if the predictors are highly correlated. Then, we will prioritize the models being tested based on the literature and our data that is independent from this cohort. The relationship between variables will be visually examined, transformations on the variables will be performed, and extra terms will be added to the regression model if necessary to model nonlinear and nonmonotonic relationships.

Age acceleration is defined as the residual resulting from regressing predicted age on chronological age. For aim 2, the primary analysis will use linear regression models (adjusted for the same covariates described above) to examine the association of immigrant, cultural, or social measures with biological aging (age acceleration value). In addition, we will compare these predictors in those defined as fast agers and slow agers. A positive age acceleration value indicates faster aging, and a negative value suggests slower aging.

For aim 3, the primary analyses will use linear regression models to examine the association of immigrant and cultural factors and social relations with cardiovascular disease (blood pressure and ECG variables). Linear mixed effects models will be applied to cerebral small vessel disease (vascular and structural brain abnormality measures) to compare abnormalities between participants across strata of immigrant and cultural factors, as well as social relations, adjusting for confounders described above. As a hypothesis-generating analysis, we will perform a

XSL•FO
RenderX

mediation analysis (product of coefficients method) to evaluate whether the vascular measures mediate any significant relationships seen in the primary analyses.

For all aims, we will contrast 2 Kerala American migrant groups (those who migrated in middle life and those who migrated in later life). For aims 1 and 3, we will also conduct a cross-national comparison of cognition and risk factors between the Kerala American and KES cohorts. We are unable to assess biological aging due to a lack of appropriate data in the KES cohort.

Qualitative Analyses

For the ethnographic life course history interviews, we will independently analyze textual data deductively and inductively and reach a consensus to minimize personal bias and interpretation errors, thereby ensuring validity and reliability. The study's trustworthiness (ie, credibility, transferability, dependability, and confirmability) will be ensured by tracking the internal analysis processes. Data will be analyzed by the qualitative content analysis method with MaxQda software (2020; VERBI GmbH).

Power

Power estimation was performed for the primary hypothesis in each aim, which will all use baseline cross-sectional data and hence will not be affected by longitudinal attrition. For power calculation using multiple linear regressions, the detectable effect size was reported as the increase in R^2, which is a standard and popularly used measure for effect size [73-75]. We did not translate such a change in R^2 to the change or differences in any outcome because such numerical translation would require multiple, largely speculative assumptions about the unexplained variance and predictor distributions, and hence it could be misleading rather than clarifying, because it would suggest a level of precision about "clinically meaningful" changes that is not supported by available data. Power calculation considers confounders by assuming a R^2 of 25%, the percentage of variance in the predictor of interest explained by confounders. With a sample size of 400 and estimated prevalence of predictors ranging from 15% to 40% based on KES and other Indian American cohort studies, we have greater than 80% power to detect an increase of R^2=0.033 at a conservative significance level 0.005, chosen under the consideration of multiple testing. The optimal size for epigenetic studies (DNAm) to create an accurate calibration model for age estimation is estimated to be 134 individuals or more. Given the estimated sample of 360 Kerala American individuals with biological aging measures, we have >80% power to detect an increase of R^2=0.032 at a conservative significance level of 0.01, with the consideration of multiple testing. We anticipate that all 400 participants will complete cardiovascular disease assessment (blood pressure and/or ECG). Given sample of 400, and at a conservative significance level of 0.005, we have >80% power to detect an increase of R^2=0.033 for the linear model; and under the assumption of 20% prevalence of cardiovascular disease (either blood pressure or ECG abnormalities), 0.8 power to detect an odds ratio (OR) 1.66 with an increase of 1 SD of the predictor. In cognitively normal older adults in KES, the prevalence of hypertension was 31%, and ECG abnormalities varied from 4% to 16%, and higher prevalence was seen in dementia cases. Given the sample of 160 for MRIs, and under the same assumption, we have 0.8 power to detect an OR of 2.33 at a significance level of 0.005. We will also combine MRIs from KES and Kerala American cohorts to examine cerebrovascular disease. This combined 460 MRI samples (KES=300 and Kerala American=160), under the same assumptions, will provide >0.8 power to detect an OR 1.5 at a significance level of 0.005.

Ethical Considerations

This study has been approved by the Stony Brook University Institutional Review Board (Stony Brook IRB 2025-00139). All participants will provide informed consent. Data will be deidentified, and protective measures will be taken to ensure participant privacy. Data will be maintained by assigning each new participant a unique study ID. That ID will be used in each database file to identify and link subject data. ID and name associations will be password-protected in an encrypted master file to which only the research staff have access. Participant data, including computer data disks, will be kept in a locked room. Identifying information about a subject will not be used during the discussion, presentation, or publication of any research data. Full backups of the database will be performed daily. For added security, copies of the database will be kept in 2 separate physical locations in locked, fire-resistant containers. Participants will be compensated for their time (US $25 per study visit, US $50 for blood collection visit, and US $50 for MRI visit).

Results

The AKKARE study was funded by the US National Institutes of Health in 2024 and received approval from the Stony Brook University Institutional Review Board to start the study in 2025. We began recruitment in September 2025. As of January 2026, we enrolled 31 participants.

Discussion

Principal Findings

The proposed study aims to investigate the effects of migration on health in the Kerala American population, focusing on dementia risk and cognition. Studying Kerala American individuals can provide insights into risk and protection of cognition as well as Alzheimer and related dementia in other immigrant groups. Our interdisciplinary clinical and biological research will inform researchers and clinicians to better identify at-risk older US immigrants for targeted assessments and interventions, understand ethnic aging, introduce state-of-the-art biological investigations to immigrant health research, and create a cohort for future aging studies in the United States.

Various migration- and social-related factors can lead to physical as well as psychosocial stresses over months to years, and along with depression, apathy, anxiety, and loneliness, can result in adverse health and cognitive consequences, such as Alzheimer and related dementias. The term "resilience" generally refers to the capacity of individuals to survive in the face of stress and shocks [76], but it has not received much research attention in

XSL•FO
RenderX

the context of cognition and dementia in migrants. In general, personal traits (self-esteem, motivation, optimism, intellect, and coping skills) and collective resources (community pride, ethnic networks, cultural practices, and faith-based networks) are recognized as protective factors that strengthen migrants' capacity to overcome challenges [77-83]. These variables may buffer the negative impacts of stressful situations and life events on brain structure and networks [84,85]. Our study will use a mixed methods approach by implementing qualitative focus group interviews to provide context to our quantitative findings and to spur new research questions.

Aging is the biggest risk factor for dementia and is described in chronological or biological terms [86]. Chronological aging (number of years a person has been alive) increases at the same rate for everyone, whereas biological aging (how old a person seems) does not. However, chronological age by itself fails to explain the heterogeneity in health outcomes in individuals. For instance, individuals with the same chronological age can be either functionally independent or frail with multiple chronic diseases. Epigenetic aging is a novel measure of biological age that reflects exposures and disease risks independent of chronological age [57-59], and epigenetic clocks may be useful for early detection of adverse health outcomes in aging immigrants. Epigenetic modifications are theorized to help explain the persistence of racial and social health disparities, such as the increased burden of cardiovascular disease among Black American individuals relative to White American individuals [87]. Few studies have linked immigration status to age acceleration measures, as we propose to do. Epigenetic clocks are fairly new, and their relationship with cognitive aging and dementia is only beginning to be elucidated [64,88-91].

Cardiovascular disease (CVD) burden is increasing faster in developing countries, where it also occurs earlier in life [92-94]. Compared with migrants from other countries, Indian American people had the highest prevalence of overweight and were less physically active [95]. Further, chronic stress heightens risk for CVD [96]. Because racial or ethnic minorities face unique stressors (eg, acculturative stress and racial discrimination), the proposed study aims to test whether these experiences are linked to the presence of MRI markers of vascular pathology, and in the future, extend these findings to Alzheimer-related pathology. Disparities in Alzheimer disease and dementia are hypothesized to persist because immigrants have depleted resources to adapt to risks and brain pathology. Interventions aimed at vascular risks may prevent Alzheimer disease and dementia, especially in migrant and native US populations with high rates of CVD or unfavorable metabolic profiles.

The proposed study has several strengths. Most notably, our study is a first-of-its-kind study of cognition in first-generation Kerala American immigrants, and our previous KES study cohort will allow us to make cross-national comparisons. Furthermore, we will examine a breadth of factors related to immigration and culture, coping and resilience mechanisms, biological aging, and vascular aging to better understand the impact of migration on health. We also plan to create a biorepository for future biological studies. Finally, while our aims in this study are cross-sectional, we also propose longitudinal follow-up as a foundation for future studies of causality of Alzheimer disease and dementias, and this study provides the foundation for longitudinal epidemiological and biological studies of immigrant Asian American health in the United States.

Conclusion

The AKKARE study will provide an in-depth understanding of migration effects on cognitive health in the Kerala American population, and shed light on underlying biological pathways and mechanisms, and interrelationships of biological and migrant related sociocultural factors. This research will have a major impact on our understanding of migrant health and Alzheimer risk among the fast-growing US population segment of Indian American individuals and lay the groundwork for future interventions to maintain cognitive health in other US immigrant populations.

Funding

This project is funded by the National Institute on Aging (1R01AG084567-01).

Authors' Contributions

All authors contributed to the study conception and design. The first draft of the manuscript was written by KC, and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report by AIMR - Aging, Injury, Musculoskeletal, and Rheumatologic Disorders Study Section, Population Sciences and Epidemiology Integrated Review Group (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 144 KB - resprot_v15i1e85493_app1.pdf]

References

XSL•FO
RenderX

1. Badrinathan S, Kapur D, Vaishnav M, Kay J. Social realities of Indian Americans: Results from the 2020 Indian American attitudes survey [Internet]. Social realities of Indian Americans: Carnegie Endowment for International Peace Washington, DC; 2024.
2. Chakravorty S, Kapur D, Singh N. The other one percent. Walton Street, Oxford, England: Indians in America. Oxford University Press; 2016.
3. Budiman A, Ruiz NG. Key facts about Asian origin groups in the US Pew Research Center;. 2021. URL: https://www.pewresearch.org/short-reads/2021/04/29/key-facts-about-asian-origin-groups-in-the-u-s/ [accessed 2026-01-22]
4. Rajan SI. Aging in Kerala: one more population problem? Asia-Pac Popul J 1989;4:19-48. [doi: 10.18356/ead11af5-en]
5. Nag M. The Kerala formula. World Health Forum 1988;9(2):258-262. [Medline: 3254226]
6. Liebig PS, Rajan SI. An aging India: Perspectives, prospects, and policies. New York: Routledge; 2013.
7. Garcia MA, Reyes AM, Downer B, Saenz JL, Samper-Ternent RA, Raji M. Age of Migration and the Incidence of Cognitive Impairment: A Cohort Study of Elder Mexican-Americans. Innov Aging 2017 Nov;1(3):igx037 [FREE Full text] [doi: 10.1093/geroni/igx037] [Medline: 30480123]
8. Markides K, Rote S. The Healthy Immigrant Effect and Aging in the United States and Other Western Countries. Gerontologist 2019 Mar 14;59(2):205-214. [doi: 10.1093/geront/gny136] [Medline: 30383212]
9. Rizzi L, Rosset I, Roriz-Cruz M. Global epidemiology of dementia: Alzheimer's and vascular types. Biomed Res Int 2014;2014:908915 [FREE Full text] [doi: 10.1155/2014/908915] [Medline: 25089278]
10. Moon H, Badana AN, Hwang SY, Sears JS, Haley WE. Dementia Prevalence in Older Adults: Variation by Race/Ethnicity and Immigrant Status. Am J Geriatr Psychiatry 2019 Mar;27(3):241-250. [doi: 10.1016/j.jagp.2018.11.003] [Medline: 30573327]
11. Garcia MA, Saenz J, Downer B, Wong R. The role of education in the association between race/ethnicity/nativity, cognitive impairment, and dementia among older adults in the United States. DemRes 2018 Jan 11;38:155-168. [doi: 10.4054/demres.2018.38.6]
12. Alegría M, Canino G, Shrout PE, Woo M, Duan N, Vila D, et al. Prevalence of mental illness in immigrant and non-immigrant U.S. Latino groups. Am J Psychiatry 2008 Mar;165(3):359-369 [FREE Full text] [doi: 10.1176/appi.ajp.2007.07040704] [Medline: 18245178]
13. Jimenez DE, Alegría M, Chen C, Chan D, Laderman M. Prevalence of psychiatric illnesses in older ethnic minority adults. J Am Geriatr Soc 2010 Mar;58(2):256-264 [FREE Full text] [doi: 10.1111/j.1532-5415.2009.02685.x] [Medline: 20374401]
14. Wallace M, Khlat M, Guillot M. Mortality advantage among migrants according to duration of stay in France, 2004-2014. BMC Public Health 2019 Mar 21;19(1):327 [FREE Full text] [doi: 10.1186/s12889-019-6652-1] [Medline: 30898125]
15. Zimmerman C, Kiss L, Hossain M. Migration and health: a framework for 21st century policy-making. PLoS Med 2011 May;8(5):e1001034 [FREE Full text] [doi: 10.1371/journal.pmed.1001034] [Medline: 21629681]
16. Lipton RB, Katz MJ, Kuslansky G, Sliwinski MJ, Stewart WF, Verghese J, et al. Screening for dementia by telephone using the memory impairment screen. J Am Geriatr Soc 2003 Oct;51(10):1382-1390. [doi: 10.1046/j.1532-5415.2003.51455.x] [Medline: 14511157]
17. Shelkey M, Wallace M. Katz Index of Independence in Activities of Daily Living. J Gerontol Nurs 1999 Mar;25(3):8-9. [doi: 10.3928/0098-9134-19990301-05] [Medline: 10362969]
18. Marian V, Blumenfeld H, Kaushanskaya M. The Language Experience and Proficiency Questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. J Speech Lang Hear Res 2007 Aug;50(4):940-967. [doi: 10.1044/1092-4388(2007/067)] [Medline: 17675598]
19. Sathyan S, Ayers E, Blumen H, Weiss E, Adhikari D, Stimmel M, et al. Epidemiology of Motoric Cognitive Risk Syndrome in the Kerala Einstein Study: Protocol for a Prospective Cohort Study. JMIR Res Protoc 2023 Aug 17;12:e49933 [FREE Full text] [doi: 10.2196/49933] [Medline: 37590054]
20. Mioshi E, Dawson K, Mitchell J, Arnold R, Hodges JR. The Addenbrooke's Cognitive Examination Revised (ACE-R): a brief cognitive test battery for dementia screening. Int J Geriatr Psychiatry 2006 Nov;21(11):1078-1085. [doi: 10.1002/gps.1610] [Medline: 16977673]
21. Mathuranath PS, Hodges JR, Mathew R, Cherian PJ, George A, Bak TH. Adaptation of the ACE for a Malayalam speaking population in southern India. Int J Geriatr Psychiatry 2004 Dec;19(12):1188-1194. [doi: 10.1002/gps.1239] [Medline: 15526301]
22. Takenoshita S, Terada S, Yoshida H, Yamaguchi M, Yabe M, Imai N, et al. Validation of Addenbrooke's cognitive examination III for detecting mild cognitive impairment and dementia in Japan. BMC Geriatr 2019 Apr 29;19(1):123 [FREE Full text] [doi: 10.1186/s12877-019-1120-4] [Medline: 31035933]
23. Dudas RB, Berrios GE, Hodges JR. The addenbrooke's cognitive examination (ACE) in the differential diagnosis of early dementias versus affective disorder. The American Journal of Geriatric Psychiatry 2005;13(3):218-226. [doi: 10.1097/00019442-200503000-00007]
24. Davies RR, Dawson K, Mioshi E, Erzinçlioğlu S, Hodges JR. Differentiation of semantic dementia and Alzheimer's disease using the Addenbrooke's Cognitive Examination (ACE). Int J Geriatr Psychiatry 2008 Apr;23(4):370-375. [doi: 10.1002/gps.1887] [Medline: 17768772]

XSL•FO
RenderX

25. Mathuranath PS, Cherian JP, Mathew R, George A, Alexander A, Sarma SP. Mini mental state examination and the Addenbrooke's cognitive examination: effect of education and norms for a multicultural population. Neurol India 2007;55(2):106-110 [FREE Full text] [doi: 10.4103/0028-3886.32779] [Medline: 17558112]
26. Menon RN, Varghese F, Paplikar A, Mekala S, Alladi S, Sharma M, et al. Validation of Indian Council of Medical Research Neurocognitive Tool Box in Diagnosis of Mild Cognitive Impairment in India: Lessons from a Harmonization Process in a Linguistically Diverse Society. Dement Geriatr Cogn Disord 2020;49(4):355-364 [FREE Full text] [doi: 10.1159/000512393] [Medline: 33412549]
27. Schmidt M. Rey Auditory Verbal Learning Test. Torrance, CA: Western Psychological Services Los Angeles; 1996.
28. Verghese J, Noone ML, Johnson B, Ambrose AF, Wang C, Buschke H, et al. Picture-based memory impairment screen for dementia. J Am Geriatr Soc 2012 Nov;60(11):2116-2120 [FREE Full text] [doi: 10.1111/j.1532-5415.2012.04191.x] [Medline: 23039180]
29. Malik R, Weiss EF, Gottesman R, Zwerling J, Verghese J. Picture-Based Memory Impairment Screen: Effective Cognitive Screen in Ethnically Diverse Populations. J Am Geriatr Soc 2018 Aug;66(8):1598-1602 [FREE Full text] [doi: 10.1111/jgs.15422] [Medline: 29808583]
30. Casarotti A, Papagno C, Zarino B. Modified Taylor Complex Figure: normative data from 290 adults. J Neuropsychol 2014 Sep;8(2):186-198. [doi: 10.1111/jnp.12019] [Medline: 23647550]
31. Carmona G, Nsubayi C, Stimmel M, Ayers E, Verghese J, Weiss EF. A - 173 symbol match validity at the Montefiore Memory Disorder Centers. Archives of Clinical Neuropsychology 2023;38(7):1345 [FREE Full text] [doi: 10.1093/arclin/acad067.190]
32. Wechsler D. WMS-R: Wechsler memory scale-revised: manual. San Antonio, TX: Psychological Corporation; 1984.
33. Kim HJ, Baek MJ, Kim S. PLoS One 2014;9(2):e89078 [FREE Full text] [doi: 10.1371/journal.pone.0089078] [Medline: 24551221]
34. Mathuranath PS, George A, Cherian PJ, Alexander A, Sarma SG, Sarma PS. Effects of age, education and gender on verbal fluency. J Clin Exp Neuropsychol 2003 Dec;25(8):1057-1064. [doi: 10.1076/jcen.25.8.1057.16736] [Medline: 14566579]
35. Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982;17(1):37-49. [doi: 10.1016/0022-3956(82)90033-4] [Medline: 7183759]
36. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
37. Thomas AM, Cherian V, Antony A. Translation, validation and cross-cultural adaptation of the geriatric depression scale (GDS-30) for utilization amongst speakers of Malayalam; the regional language of the South Indian State of Kerala. J Family Med Prim Care 2021 May;10(5):1863-1867 [FREE Full text] [doi: 10.4103/jfmpc.jfmpc_1813_20] [Medline: 34195117]
38. De Man J, Absetz P, Sathish T, Desloge A, Haregu T, Oldenburg B, et al. Are the PHQ-9 and GAD-7 Suitable for Use in India? A Psychometric Analysis. Front Psychol 2021;12:676398 [FREE Full text] [doi: 10.3389/fpsyg.2021.676398] [Medline: 34054677]
39. Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [doi: 10.1097/00005650-199603000-00003] [Medline: 8628042]
40. Needham BL, Mukherjee B, Bagchi P, Kim C, Mukherjea A, Kandula NR, et al. Acculturation Strategies and Symptoms of Depression: The Mediators of Atherosclerosis in South Asians Living in America (MASALA) Study. J Immigr Minor Health 2018 Aug;20(4):792-798 [FREE Full text] [doi: 10.1007/s10903-017-0635-z] [Medline: 28748299]
41. Williams DR, Yan Yu, Jackson JS, Anderson NB. Racial Differences in Physical and Mental Health: Socio-economic Status, Stress and Discrimination. J Health Psychol 1997 Jul;2(3):335-351 [FREE Full text] [doi: 10.1177/135910539700200305] [Medline: 22013026]
42. Singh G, Kaur D, Kaur H. Presumptive stressful life events scale (psles) - a new stressful life events scale for use in India. Indian J Psychiatry 1984 Apr;26(2):107-114 [FREE Full text] [Medline: 21965968]
43. Cohen S, Doyle WJ, Turner R, Alper CM, Skoner DP. Sociability and susceptibility to the common cold. Psychol Sci 2003 Sep;14(5):389-395. [doi: 10.1111/1467-9280.01452] [Medline: 12930466]
44. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32(6):705-714. [doi: 10.1016/0277-9536(91)90150-b] [Medline: 2035047]
45. Domènech-Abella J, Lara E, Rubio-Valera M, Olaya B, Moneta M, Rico-Uribe L, et al. Loneliness and depression in the elderly: the role of social network. Soc Psychiatry Psychiatr Epidemiol 2017 Apr;52(4):381-390 [FREE Full text] [doi: 10.1007/s00127-017-1339-3] [Medline: 28154893]
46. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Mar;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
47. Lee SYD, Stucky BD, Lee JY, Rozier RG, Bender DE. Short Assessment of Health Literacy-Spanish and English: a comparable test of health literacy for Spanish and English speakers. Health Serv Res 2010 Aug;45(4):1105-1120 [FREE Full text] [doi: 10.1111/j.1475-6773.2010.01119.x] [Medline: 20500222]

XSL•FO
RenderX

48. Ambrose AF, Adhikari D, Kumar VGP, Sigamani A, Weiss E, Jin Y, et al. Kerala Brain Injury Questionnaire: Feasibility, Reliability, and Validation Study. Neurol India 2025 Nov 01;73(6):1193-1198 [FREE Full text] [doi: 10.4103/neurol-india.Neurol-India-D-24-00864] [Medline: 41215511]
49. Mathuranath PS, George A, Cherian PJ, Mathew R, Sarma PS. Instrumental activities of daily living scale for dementia screening in elderly people. Int Psychogeriatr 2005 Sep;17(3):461-474 [FREE Full text] [doi: 10.1017/s1041610205001547] [Medline: 16252378]
50. Mathuranath PS, Cherian PJ, Mathew R, Kumar S, George A, Alexander A, et al. Dementia in Kerala, South India: prevalence and influence of age, education and gender. Int J Geriatr Psychiatry 2010 Mar;25(3):290-297 [FREE Full text] [doi: 10.1002/gps.2338] [Medline: 19621355]
51. Verghese J, LeValley A, Derby C, Kuslansky G, Katz M, Hall C, et al. Leisure activities and the risk of amnestic mild cognitive impairment in the elderly. Neurology 2006 Mar 28;66(6):821-827 [FREE Full text] [doi: 10.1212/01.wnl.0000202520.68987.48] [Medline: 16467493]
52. Verghese J, Lipton RB, Katz MJ, Hall CB, Derby CA, Kuslansky G, et al. Leisure Activities and the Risk of Dementia in the Elderly. N Engl J Med 2003 Jun 19;348(25):2508-2516 [FREE Full text] [doi: 10.1056/nejmoa022252]
53. Rattanabannakit C, Risacher SL, Gao S, Lane KA, Brown SA, McDonald BC, et al. The Cognitive Change Index as a Measure of Self and Informant Perception of Cognitive Decline: Relation to Neuropsychological Tests. Journal of Alzheimer's Disease 2016 Feb 25;51(4):1145-1155 [FREE Full text] [doi: 10.3233/jad-150729]
54. Hicken MT, Lee H, Hing AK. The weight of racism: Vigilance and racial inequalities in weight-related measures. Soc Sci Med 2018 Mar;199:157-166 [FREE Full text] [doi: 10.1016/j.socscimed.2017.03.058] [Medline: 28372829]
55. Lamb S. Permanent personhood or meaningful decline? Toward a critical anthropology of successful aging. J Aging Stud 2014 Apr;29:41-52. [doi: 10.1016/j.jaging.2013.12.006] [Medline: 24655672]
56. Lamb S. On Being (Not) Old: Agency, Self-care, and Life-course Aspirations in the United States. Med Anthropol Q 2019 Jun;33(2):263-281. [doi: 10.1111/maq.12498] [Medline: 30575125]
57. Bocklandt S, Lin W, Sehl ME, Sánchez FJ, Sinsheimer JS, Horvath S, et al. Epigenetic predictor of age. PLoS One 2011;6(6):e14821 [FREE Full text] [doi: 10.1371/journal.pone.0014821] [Medline: 21731603]
58. Higgins-Chen AT, Thrush KL, Levine ME. Aging biomarkers and the brain. Semin Cell Dev Biol 2021 Aug;116:180-193 [FREE Full text] [doi: 10.1016/j.semcdb.2021.01.003] [Medline: 33509689]
59. Wu JW, Yaqub A, Ma Y, Koudstaal W, Hofman A, Ikram MA, et al. Biological age in healthy elderly predicts aging-related diseases including dementia. Sci Rep 2021 Aug 05;11(1):15929 [FREE Full text] [doi: 10.1038/s41598-021-95425-5] [Medline: 34354164]
60. Xiao FH, Wang HT, Kong QP. Dynamic DNA Methylation During Aging: A "Prophet" of Age-Related Outcomes. Front Genet 2019;10:107 [FREE Full text] [doi: 10.3389/fgene.2019.00107] [Medline: 30833961]
61. Bell CG, Lowe R, Adams PD, Baccarelli AA, Beck S, Bell JT, et al. DNA methylation aging clocks: challenges and recommendations. Genome Biol 2019 Nov 25;20(1):249 [FREE Full text] [doi: 10.1186/s13059-019-1824-y] [Medline: 31767039]
62. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol 2013;14(10):R115 [FREE Full text] [doi: 10.1186/gb-2013-14-10-r115] [Medline: 24138928]
63. Pelegí-Sisó D, de Prado P, Ronkainen J, Bustamante M, González JR. methylclock: a Bioconductor package to estimate DNA methylation age. Bioinformatics 2021 Jul 19;37(12):1759-1760. [doi: 10.1093/bioinformatics/btaa825] [Medline: 32960939]
64. Gao X, Chen Q, Yao H, Tan J, Liu Z, Zhou Y, et al. Epigenetics in Alzheimer's Disease. Front Aging Neurosci 2022;14:911635 [FREE Full text] [doi: 10.3389/fnagi.2022.911635] [Medline: 35813941]
65. Drawz PE, Ix JH. BP Measurement in Clinical Practice: Time to SPRINT to Guideline-Recommended Protocols. JASN 2017 Oct 19;29(2):383-388. [doi: 10.1681/asn.2017070753]
66. Buss S, Noone ML, Tsai R, Johnson B, Pradeep VG, Salam KA, et al. Objective cardiac markers in dementia: results from the Kerala-Einstein study. Int J Cardiol 2013 Jul 31;167(2):595-596 [FREE Full text] [doi: 10.1016/j.ijcard.2012.09.220] [Medline: 23103139]
67. Schmidt P, Gaser C, Arsic M, Buck D, Förschler A, Berthele A, et al. An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. Neuroimage 2012 Mar 15;59(4):3774-3783. [doi: 10.1016/j.neuroimage.2011.11.032] [Medline: 22119648]
68. Gregoire SM, Chaudhary UJ, Brown MM, Yousry TA, Kallis C, Ja¨ger HR, et al. The Microbleed Anatomical Rating Scale (MARS). Neurology 2009 Nov 24;73(21):1759-1766 [FREE Full text] [doi: 10.1212/wnl.0b013e3181c34a7d]
69. Greenberg SM, Vernooij MW, Cordonnier C, Viswanathan A, Al-Shahi Salman R, Warach S, et al. Cerebral microbleeds: a guide to detection and interpretation. The Lancet Neurology 2009 Feb;8(2):165-174. [doi: 10.1016/s1474-4422(09)70013-4]
70. Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, et al. Automatically parcellating the human cerebral cortex. Cereb Cortex 2004 Jan;14(1):11-22. [doi: 10.1093/cercor/bhg087] [Medline: 14654453]
71. Yendiki A, Panneck P, Srinivasan P, Stevens A, Zöllei L, Augustinack J, et al. Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. Front Neuroinform 2011;5:23 [FREE Full text] [doi: 10.3389/fninf.2011.00023] [Medline: 22016733]

XSL•FO
RenderX

72. Maffei C, Lee C, Planich M, Ramprasad M, Ravi N, Trainor D, et al. Using diffusion MRI data acquired with ultra-high gradient strength to improve tractography in routine-quality data. Neuroimage 2021 Dec 15;245:118706 [FREE Full text] [doi: 10.1016/j.neuroimage.2021.118706] [Medline: 34780916]
73. Maxwell SE, Kelley K, Rausch JR. Sample size planning for statistical power and accuracy in parameter estimation. Annu Rev Psychol 2008;59:537-563. [doi: 10.1146/annurev.psych.59.103006.093735] [Medline: 17937603]
74. Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput Sci 2021;7:e623 [FREE Full text] [doi: 10.7717/peerj-cs.623] [Medline: 34307865]
75. Cohen J. Statistical Power Analysis for the Behavioral Sciences (Second edition). United Kingdom: Routledge Academic; 1988.
76. Siriwardhana C, Ali SS, Roberts B, Stewart R. A systematic review of resilience and mental health outcomes of conflict-driven adult forced migrants. Confl Health 2014;8:13 [FREE Full text] [doi: 10.1186/1752-1505-8-13] [Medline: 25177360]
77. Nieto M, Visier ME, Silvestre IN, Navarro B, Serrano JP, Martínez-Vizcaíno V. Relation between resilience and personality traits: The role of hopelessness and age. Scand J Psychol 2023 Mar;64(1):53-59 [FREE Full text] [doi: 10.1111/sjop.12866] [Medline: 36057793]
78. McSorley VE, Howard C, Shah RC, James BD, Boyle PA, Barnes LL. The Relationship of John Henryism With Cognitive Function and Decline in Older Black Adults. Psychosom Med 2022 Jul 28;84(7):766-772. [doi: 10.1097/psy.0000000000001113]
79. Aschwanden D, Sutin AR, Ledermann T, Luchetti M, Stephan Y, Sesker AA, et al. Subjective Cognitive Decline: Is a Resilient Personality Protective Against Progression to Objective Cognitive Impairment? Findings from Two Community-Based Cohort Studies. Journal of Alzheimer's Disease 2022 Jul 15;89(1):87-105. [doi: 10.3233/jad-220319]
80. Améndola L, Weary D, Zobel G. Effects of personality on assessments of anxiety and cognition. Neurosci Biobehav Rev 2022 Oct;141:104827 [FREE Full text] [doi: 10.1016/j.neubiorev.2022.104827] [Medline: 35970418]
81. Walters MJ, Sterling J, Quinn C, Ganzer C, Osorio RS, Andrews RD, et al. Associations of lifestyle and vascular risk factors with Alzheimer's brain biomarker changes during middle age: a 3-year longitudinal study in the broader New York City area. BMJ Open 2018 Nov 25;8(11):e023664 [FREE Full text] [doi: 10.1136/bmjopen-2018-023664] [Medline: 30478117]
82. Perneczky R, Kempermann G, Korczyn AD, Matthews FE, Ikram MA, Scarmeas N, et al. Translational research on reserve against neurodegenerative disease: consensus report of the International Conference on Cognitive Reserve in the Dementias and the Alzheimer's Association Reserve, Resilience and Protective Factors Professional Interest Area working groups. BMC Med 2019 Mar 27;17(1):47 [FREE Full text] [doi: 10.1186/s12916-019-1283-z] [Medline: 30808345]
83. Joshi MS, Galvin JE. Cognitive Resilience in Brain Health and Dementia Research. JAD 2022 Nov 08;90(2):461-473. [doi: 10.3233/jad-220755]
84. Lindsay-Smith G, O'Sullivan G, Eime R, Harvey J, van Uffelen JGZ. A mixed methods case study exploring the impact of membership of a multi-activity, multicentre community group on social wellbeing of older adults. BMC Geriatr 2018 Sep 24;18(1):226 [FREE Full text] [doi: 10.1186/s12877-018-0913-1] [Medline: 30249200]
85. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychological Bulletin 1985 Sep;98(2):310-357. [doi: 10.1037/0033-2909.98.2.310]
86. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell 2013 Jun 06;153(6):1194-1217 [FREE Full text] [doi: 10.1016/j.cell.2013.05.039] [Medline: 23746838]
87. Kuzawa CW, Sweet E. Epigenetics and the embodiment of race: developmental origins of US racial disparities in cardiovascular health. Am J Hum Biol 2009;21(1):2-15. [doi: 10.1002/ajhb.20822] [Medline: 18925573]
88. Gonzales M, Garbarino V, Pollet E, Palavicini J, Kellogg D, Kraig E. Biological aging processes underlying cognitive decline and neurodegenerative disease. J Clin Invest 2022:e158453 [FREE Full text] [doi: 10.1172/jci158453]
89. Othman Z, Abdul Halim AS, Azman KF, Ahmad AH, Zakaria R, Sirajudeen KNS, et al. Profiling the Research Landscape on Cognitive Aging: A Bibliometric Analysis and Network Visualization. Front Aging Neurosci 2022;14:876159 [FREE Full text] [doi: 10.3389/fnagi.2022.876159] [Medline: 35572132]
90. Sherazi SAM, Abbasi A, Jamil A, Uzair M, Ikram A, Qamar S, et al. Molecular hallmarks of long non-coding RNAs in aging and its significant effect on aging-associated diseases. Neural Regen Res 2023 May;18(5):959-968 [FREE Full text] [doi: 10.4103/1673-5374.355751] [Medline: 36254975]
91. Zhou A, Wu Z, Zaw Phyo AZ, Torres D, Vishwanath S, Ryan J. Epigenetic aging as a biomarker of dementia and related outcomes: a systematic review. Epigenomics 2022 Sep;14(18):1125-1138 [FREE Full text] [doi: 10.2217/epi-2022-0209] [Medline: 36154448]
92. Gaziano TA, Galea G, Reddy KS. Scaling up interventions for chronic disease prevention: the evidence. The Lancet 2007 Dec;370(9603):1939-1946. [doi: 10.1016/s0140-6736(07)61697-3]
93. Gupta R, Joshi P, Mohan V, Reddy KS, Yusuf S. Epidemiology and causation of coronary heart disease and stroke in India. Heart 2008 Jan;94(1):16-26. [doi: 10.1136/hrt.2007.132951] [Medline: 18083949]
94. Reddy S. Regional case studies? India. In: Kalhan SC, Prentice AM, Yajnik CS, Editors. Emerg Soc - coexistence child malnutrition obes 63rd nestlé nutr inst workshop pediatr program New Delhi March-April. S.Karger AG 2008:15-24 [FREE Full text] [doi: 10.1159/isbn.978-3-8055-9010-5]

XSL•FO
RenderX

95. Koirala B, Turkson-Ocran RA, Baptiste D, Koirala B, Francis L, Davidson P, et al. Heterogeneity of Cardiovascular Disease Risk Factors Among Asian Immigrants: Insights From the 2010 to 2018 National Health Interview Survey. J Am Heart Assoc 2021 Jul 06;10(13):e020408 [FREE Full text] [doi: 10.1161/JAHA.120.020408] [Medline: 34182790]
96. Rosengren A, Hawken S, Ôunpuu S, Sliwa K, Zubaid M, Almahmeed W, et al. Association of psychosocial risk factors with risk of acute myocardial infarction in 11 119 cases and 13 648 controls from 52 countries (the INTERHEART study): case-control study. The Lancet 2004 Sep;364(9438):953-962. [doi: 10.1016/s0140-6736(04)17019-0]

Abbreviations

ACE: Addenbrooke Cognitive Examination
AKKARE: Aging in Kerala Americans Research Study
CVD: cardiovascular disease
DNAm: DNA methylation
ECG: electrocardiogram
FLAIR: fluid-attenuated inversion recovery
FOV: field of view
FSL: Functional Magnetic Resonance Imaging of the Brain Software Library
GAD-7: Generalized Anxiety Disorder
GDS: Geriatric Depression Scale
KES: Kerala Einstein Study
MRI: magnetic resonance imaging
OR: odds ratio
SNI: Social Network Index
SPRINT: Systolic Blood Pressure Intervention Trial
TE: echo time
TR: repetition time
TRACULA: Tracts Constrained by Underlying Anatomy
WMH: white matter hyperintensity

Edited by J Sarvestan; submitted 27.Oct.2025; this is a non–peer-reviewed article;accepted 16.Jan.2026; published 13.Feb.2026.

Please cite as:
Cotton K, Adhikari D, Ambrose AF, Ayers E, Blumen HM, Ceïde ME, Kumar VGP, Melecio-Vazquez M, Sathyan S, Sigamani A, Sivaramakrishnan K, Stimmel M, Weiss EF, Yang J, Verghese J
Migration Effects on Cognition: Protocol for the Aging in Kerala Americans Research Study
JMIR Res Protoc 2026;15:e85493
URL: https://www.researchprotocols.org/2026/1/e85493
doi:10.2196/85493
PMID:

©Kelly Cotton, Dristi Adhikari, Anne Felicia Ambrose, Emmeline Ayers, Helena M Blumen, Mirnova E Ceïde, VG Pradeep Kumar, Mairim Melecio-Vazquez, Sanish Sathyan, Alben Sigamani, Kavita Sivaramakrishnan, Marnina Stimmel, Erica F Weiss, Jie Yang, Joe Verghese. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Evaluation of the Positive Peers Mobile App for Supporting the Viral Suppression of Young People With HIV: Protocol for a Concurrent Mixed Methods Evaluation With Randomized Controlled Clinical Trial and Observational Cohort

Mary M Step[1], MACT, PhD; L Anthony Catania[2], BA; Jennifer McMillen Smith[2], LPSW; Steven A Lewis[3], MPH; Yanis Bitar[2], MPH; Vinay K Cheruvu[1], PhD; Kristen A Berg[3], PhD; Jeffrey S Hallam[1], PhD; Ann K Avery[4], MD

[1]College of Public Health, Kent State University, Kent, OH, United States

[2]Division of Infectious Disease, MetroHealth Medical Center, Cleveland, OH, United States

[3]Case Western Reserve University School of Medicine at MetroHealth, Cleveland, OH, United States

[4]Department of Infectious Disease, Case Western Reserve University School of Medicine at MetroHealth, Cleveland, OH, United States

Corresponding Author:
Mary M Step, MACT, PhD
College of Public Health
Kent State University
750 Hilltop Drive
Kent, OH, 44242
United States
Phone: 1 330 672 2630
Email: mstep@kent.edu

Abstract

Background: People at greatest risk for poor HIV outcomes include young (13-34) people of color who have sex with men. Individuals in this population are least likely to be aware of their HIV status and are at the highest risk for disengaging from medical care and antiretroviral therapy. The Positive Peers mobile app (PPA) was designed to engage this population with real-time social support, HIV and healthy lifestyle information, and medical management tools. We expect that greater PPA engagement will predict key HIV care outcomes. Study predictions are grounded in a user-centric model of digital media use and the perceived affordances of the PPA.

Objective: This study aims to determine the optimal deployment of the PPA in clinical settings (at enrollment vs delayed start) and to compare intervention outcomes with a no-intervention, observation-only condition. The PPA is designed specifically for key HIV disparity populations, including younger sexual, gender, racial, and ethnic minorities.

Methods: The Positive Peers Intervention Trial (PoPIT) is a multisite, randomized clinical trial designed to evaluate the effectiveness of the PPA as a tool for use in clinical settings. Trial arms compare the immediate deployment of the PPA with usual care and observation only. This protocol outlines a mixed methods design consisting of concurrent prospective self-report questionnaires, in-depth interviews with PPA users, and medical record review. PoPIT questionnaires include measures of social determinants of health, HIV-related stigma, perceptions of digital media use, self-efficacy, substance use, and social support. Multiple aspects of PPA intervention engagement are measured natively within the app. Outcomes include HIV National Quality Forum indicators and perceived HIV-related stigma.

Results: The research protocol (1R01MD019185-01) was funded for US $903,363.00 in direct costs by the National Institute for Minority Health and Health Disparities on September 24, 2023, for 5 years. The study began recruiting patients on June 3, 2024, and will continue accrual until November 27, 2026.

Conclusions: Findings will provide evidence of the usefulness of the PPA as a support tool in HIV clinical care. Primary results will inform optimization of PPA deployment and evaluate a theoretical model of user engagement with a mobile health management app. Qualitative data will provide a phenomenological description of intervention engagement and perceived user efficacy.

Trial Registration: ClinicalTrials.gov NCT06388109; https://clinicaltrials.gov/ct2/show/NCT06388109

International Registered Report Identifier (IRRID): DERR1-10.2196/87601

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e87601) doi:10.2196/87601

KEYWORDS

HIV; young people with HIV; mHealth; HIV support; social support; retention in HIV care; viral suppression; peer support; social media affordances; HIV stigma

Introduction

Background

Younger people under the age of 35 continue to carry a disproportionate burden of HIV risk [1-4]. Within this group, social risk factors such as poverty, drug use, job insecurity, food scarcity, and housing instability present significant barriers to HIV testing, retention in care, and viral suppression [5,6]. These factors contribute to younger people being less likely to be retained in HIV care and to achieve sustained viral suppression, especially those in marginalized communities [1,7]. There is a critical need to establish effective and sustainable tools to address these barriers and improve HIV outcomes for disproportionately impacted young people [4,7].

Several electronic tools have been designed to engage young people with HIV (YPWH) in clinical care and to support medication adherence [8-10]. However, fewer tools link newly diagnosed and out-of-care clinical populations to a live peer cohort that shares the unique challenge of living with HIV [11]. Mobile health (mHealth) apps are gaining credibility as reliable and accessible communication modalities that provide a direct platform for addressing social barriers to adherence, retention, and viral suppression [10,12,13]. Implementation of mHealth interventions is particularly well suited to high-priority jurisdictions where PWH may be socially and geographically isolated from trustworthy information and reliable support [14-16]. The Positive Peers mobile app (PPA) was created in response to these needs as a clinical mHealth tool to help newly diagnosed young people, or those returning to care, connect with each other and learn about living with HIV [17].

The Positive Peers Mobile App

PPA is a free, multifunction mobile app that features interactive functions organized around a Mind/Body/Spirit theme. The Mind theme is characterized by information provision, including more than 400 vetted articles concerning issues faced by YPWH, as well as first-person stories from current app users living with HIV. The Body theme offers tangible support in the form of interactive medication and laboratory reminders and trackers, and curated state-by-state community resources addressing a variety of living needs (clothing, housing, food). The Spirit theme fosters human connection through a community forum that allows users to ask questions and share experiences and encouragement, as well as a chat function for 1-on-1 conversations. The app is freely available to eligible young people (ages 13-34) who provide evidence of their HIV-positive status. App activity is monitored and supported by a group of trained app administrators at our home medical institution in Cleveland, Ohio, as well as at key partner locations across the United States.

Following development and successful feasibility testing with a pilot group [17], a cohort comparison medical record review of PPA users and clinically eligible but study-naïve nonusers showed that PPA users were significantly more likely to be regularly engaged in HIV care and to achieve higher rates of viral suppression than nonusers, particularly among young adults aged 13-24 [18]. This success, paired with positive reactions identified in qualitative work [19], provided the basis for planning a randomized controlled trial to evaluate the PPA as a clinical tool for supporting YPWH who are newly diagnosed or seeking care in 6 federally designated Ending the Epidemic priority jurisdictions in the United States [14].

The primary objective of the Positive Peers Intervention Trial (PoPIT) is to evaluate the impact of the mHealth app on engagement in care and viral suppression outcomes among newly diagnosed and reengaged young patients. Study research questions and hypotheses are informed by, and evaluated based on, social networking, media uses and effects, and social support theories.

Social Networking Theory and Social Support

Social Networking Theory (SNT) is an approach to understanding how ideas and behaviors disseminate through relational ties [20]. A key line of research in SNT examines whether and how social networks influence health behavior [21,22]. SNT has guided thinking about information flow and innovation uptake for decades, with more contemporary approaches positioning social network membership as central to identity, community, and health [23-25]. Notably, although several mechanisms of influence have been postulated (eg, modeling, presentation of group norms), the source of behavior change in SNT approaches stems from social ties rather than individual cognition [26-28]. A key mechanism underlying social network effects is social support.

Although social support was initially conceptualized as a dyadic process (see [29]), health-related research has produced ample evidence of the efficacy and benefits of networked social support on various social media platforms, including reduced depression and mitigation of stigma [30-32]. As social media enables a more flexible, collaborative communication process, it offers greater opportunities for engagement with similar others [33-35]. Participants in online support groups have reported strong feelings of empowerment associated with exchanging information, receiving emotional support, experiencing validation and understanding, helping others, and sharing amusement [36,37]. These empowerment processes, in turn, are associated with feeling more informed, experiencing greater social well-being, demonstrating greater efficacy in interactions with physicians and in managing treatment, and greater acceptance of illness [38].

Among PWH, stronger social support has been associated with safer sex [39], better medication adherence [40], fewer disclosure-related concerns [41], lower depression [42], and better overall quality of life [31,43]. Not surprisingly, younger

XSL•FO
RenderX

people, and those recently diagnosed with HIV, are more likely to use online rather than face-to-face support, a trend that reverses as they gain experience and begin to trust others [44,45]. Support messaging in this context most frequently takes the form of informational and emotional support [43,45].

Social networking among PWH particularly benefits younger, isolated, or marginalized people who report less in-person support [46]. Stress rooted in chronic illness is known to be associated with a greater prevalence of adverse mental and behavioral health outcomes [47,48]. Ongoing stressors, such as living with HIV, homelessness, or identifying with a population minority, can also motivate social media use and online support seeking [49]. Positive outcomes associated with stress-induced social media use include active coping, the creation of positive group norms, increased self-efficacy, and reduced perceived stigma [37,50-52]. Importantly, unique sources of stress can be mitigated through connection with similarly stigmatized others, particularly within a context of social support and expressed self-compassion [53-56]. The resulting sense of belonging, or social connectedness, generated in these situations can catalyze resilience, increased esteem, and potentially improved self-care [55,57]. These effects are known to be robust even within weak-tie social networks, such as those involving acquaintances [58]. The PPA leverages these processes and effects by providing a secure space where social connections with similar others can aid the management of ongoing chronic illness stressors. The PPA builds on these ideas by using HIV status as a criterion for belonging rather than exclusion, providing secure opportunities for individual or group communication, and disseminating information targeted to network characteristics.

Digital Media Uses, Effects, and Affordances

The PPA app is informed by 2 user-centric theories of media use: media uses and effects theory and media affordances theory [59-63]. The media uses and effects approach is a model of media use that posits that needs experienced by users are expressed as motives, that is, intentions for media use [64]. Active user motives (eg, information- or support-seeking), rather than passive motives (eg, passing time), are associated with more direct information and affective processing, which further determine the likelihood of cognitive, affective, or behavioral change [65]. This paradigm has been useful for understanding internet use in general and social media effects in particular [61,66]. For example, Phua et al [24] established that social networking site design and interface features were associated with different motives for use and subsequent gratifications received from those sites.

Media affordances theory developed from the media uses and effects approach by accommodating the rapid evolution of technology as a mediator of human communication. Within media affordances theory, specific features, functions, or constraints of social media technology provide opportunities to act relative to a user's needs or purposes within a given context [67,68]. Media affordances are defined as variable user perceptions of technology features that offer pathways for the gratification of perceived needs [69]. For example, interacting with a peer in the PPA community forum may address a user's need for social inclusion in a more acceptable way than through a public social media platform, creating a perceived affordance associated with the app. Not all users perceive the same affordances of the same technology. Instead, users' perceptions of app or device affordances emerge through interaction with that technology. For example, one person may use an mHealth app because it affords a confidential space to connect with similar others, whereas another may be drawn to a comprehensive medication tracking system. These theoretical frameworks suggest that user needs and perceptions of app affordances create a context for seeking gratification from that app that may or may not be reinforced. Experimentation and adaptation shape users' perceptions of how an app can work for them. These related user-centric approaches to understanding media influence depict a communication process that accommodates variation in user-technology interactions [67,68]. Variability can be found in (1) user characteristics and needs, (2) technology design features, and (3) user-technology interactions and outcomes [68].

Study Rationale and Objectives

We expect the PPA to support engagement in HIV care by normalizing and rewarding care visits, promoting health literacy, reducing stigma, and enhancing inclusion, thereby allowing users to help others with similar challenges. Our goal is to increase HIV viral suppression rates by at least 15% by offering practical and peer support through a smartphone app. Specific aims addressed by the protocol described in this manuscript are (1) comparing the effectiveness of HIV care supported by the PPA with usual care for improving retention in HIV care and viral suppression among high-priority adolescents and young adults with HIV; (2) identifying factors that predict user engagement with primary PPA components and associated effects on retention in care, viral suppression, and HIV-related perceived stigma; and (3) evaluating PoPIT implementation across sites and over time. Figure 1 illustrates the conceptual model of the PoPIT study's influence and outcomes. This paper describes the protocol used to achieve these aims in a randomized clinical trial with a concurrent qualitative evaluation focused on user experience.

XSL•FO
RenderX

Figure 1. Positive Peers Intervention Trial (PoPIT) study conceptual model. PPA: Positive Peers mobile app.

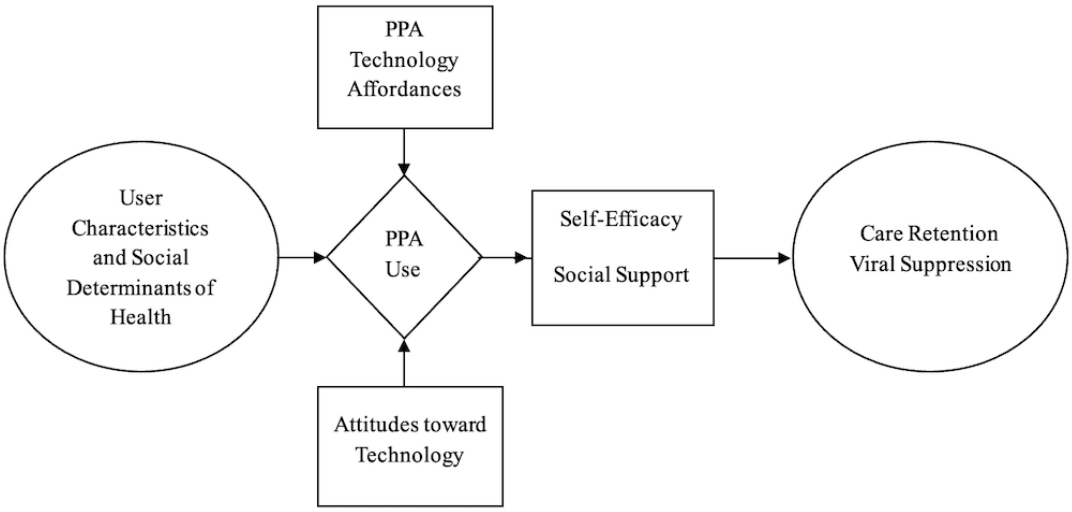

Methods

Study Design

Specific aims and corresponding hypotheses will be evaluated in a randomized controlled trial that compares HIV outcomes in patient groups randomized to either immediate or delayed access to the PPA. The delayed-start design is intended to ensure that both randomized study conditions receive identical study procedures, follow-up, reminders, incentives, and assessments during the first 6 months of participation. The only difference is the timing of app access. Any improvement observed in the delayed-start group during this period can therefore be attributed primarily to nonspecific effects of trial participation. The primary randomized comparison at 6 months thus estimates the incremental benefit of the app beyond these common trial effects. A greater improvement in outcomes in the immediate-start group at this point would provide evidence of a specific impact of the app.

Prospective questionnaires will be administered at baseline and at 3, 6, 9, and 12 months to provide evidence for model-specific hypotheses and the mediating effects of covariates and demographics. We have included an opportunity for patients who decline full study participation to grant permission to the PoPIT team to abstract their HIV outcomes from medical records on the same schedule as full study participants. This observation-only arm includes only the baseline questionnaire and medical record review of study outcomes. The inclusion of an observational cohort in the randomized controlled trial design strengthens external validity by enabling comparison of trial participants with the broader clinic population. By comparing outcomes from both randomized controlled trial arms with those of the observational cohort, we will gain additional insight into how trial participation may influence outcomes and how the intervention might be further adapted for a more diverse clinical population.

Allocation (1:1) to study conditions is determined by the REDCap system's randomization module following signed consent and completion of the baseline questionnaire, and before PPA download. Participants are informed of their randomization assignment after completing the baseline questionnaire. If randomized to the immediate intervention condition, site staff assist the participant in downloading the PPA to their personal smartphone and orient them to the app's functions and study guidelines. Participants randomized to the delayed-start arm receive usual clinic care before returning to the clinic 6 months later to download and begin using the app. This delayed-start cohort serves as a control group, allowing comparison of the effects of intervention exposure with usual care.

The Practical Robust Implementation and Sustainability Model (PRISM) was used to inform intervention implementation throughout the study [70,71] (Table 1). The PRISM model was selected because it accommodates multilevel and dynamic interactions among the intervention, stakeholders' perspectives, and recipient characteristics. The PRISM model is an expansion of the well-used RE-AIM (ie, Reach, Effectiveness, Adoption, Implementation, and Maintenance) approach, with additional emphasis on sustainability factors both within and external to the study setting [70].

XSL•FO
RenderX

Table 1. PRISM[a]-informed implementation plan for the PoPIT[b] study.

RE-AIM[c] domain	Before implementation	During implementation	After implementation
Intervention delivery	• Estimate population parameters with site staff and set accrual benchmarks. • Develop recruitment strategies with site staff.	• Log and monitor recruitment and enrollment. • Deploy boosting efforts if needed.	• Calculate reach and enrollment trajectories.
Effectiveness	• Preintervention site visits to determine setting and population characteristics, and resources.	• Collect prospective questionnaire data. • Medical record review. • Conduct user qualitative interviews.	• Primary and secondary outcome analysis and qualitative analysis of user interview data.
Adoption	• Tailor recruitment materials for sites. • Provide clear user support materials and training curriculum for site staff. • Create instructional animations for PPA[d] functions.	• Track all admin requests for app assistance, addressing potential barriers as needed. • Analyze native app engagement tracking. • Review monthly performance trajectories.	• Report adoption facilitators and barriers, construct app engagement variables, and track use and attrition trajectories over the project life.
Implementation	• Deliver group and individual site trainings to promote fidelity. • Design a REDcap[e] data management system for onboarding and data collection.	• Log site administrator app activity and user compliance with data collection. • Conduct monthly all-site check-in meetings.	• Postimplementation staff interviews.
Maintenance	• Develop oversight strategies, meeting agendas, staff interview approach, and questions.	• Monthly site staff meetings. • Qualitative staff interviews.	• Evaluate cost, review systems, identify potential adaptations, and determine dissemination prospects.

[a]PRISM: Practical Robust Implementation and Sustainability Model.

[b]PoPIT: Positive Peers Intervention Trial.

[c]RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance.

[d]PPA: Positive Peers mobile app.

[e]REDcap: Research Electronic Data Capture.

Implementation planning with PRISM identified a set of evaluation activities before, during, and after the study that provide evidence of intervention exposure, effectiveness, staff adoption of study procedures, and intervention sustainability. A detailed description of the implementation plan and associated outcomes will be reported in a future manuscript.

Sample Population

The study is designed to evaluate the effects of the PPA intervention among the age group most likely to be newly diagnosed with HIV in the United States: adolescents and young adults. This cohort is challenged by a propensity for greater risk behavior, lower health literacy and management skills, and typically less favorable social determinants of health. Potential participants are identified as patients at 1 of 5 HIV care organizations serving as study sites, located across 6 priority jurisdictions where HIV transmission occurs most frequently. Patients are screened to determine whether they meet all necessary inclusion criteria for study participation (Textbox 1).

XSL•FO
RenderX

Textbox 1. Positive Peers Intervention Trial eligibility criteria.

1. Inclusion criteria

- HIV-positive status, either newly diagnosed within the last 12 months, out of care (without an office visit in the last 12 months), or not virally suppressed (viral load >200 copies/mL within the last 12 months).
- Adolescents and young adults aged 13-34 receiving care at one of the study sites.
- Identification with a sexual, gender, racial, or ethnic minority.
- Participant's capacity to provide informed consent (aged >16 years) or assent to participate (aged <16 years).
- Ownership of a smartphone/tablet (either iOS or Android) capable of supporting the Positive Peers mobile app.
- Ability to read English.

1. Exclusion criteria

- Previous use of the Positive Peers mobile app.
- Adults older than age 34 at the time of enrollment.
- Not a patient at one of the participating clinical sites.

Study Setting

Ending the Epidemic aims to decrease new infections by identifying PWH, increasing access to biomedical and HIV prevention interventions, and improving engagement in care for those already diagnosed [72]. The PoPIT study is being implemented at 6 study sites within high-priority jurisdictions in the United States. Three of these sites are within infectious disease clinics at academic medical centers in Texas, New Jersey, and Washington, while the remaining 3 are located at federally qualified health centers in California and Ohio.

Recruitment

Eligible persons seeking care or registered with a participating clinic are invited to participate in the study through their preferred contact channel on record with the clinic. They are asked by trained clinic staff at participating sites whether they would be interested in helping to evaluate a new HIV support app in a research study. Persons who agree to participate review the informed consent document with the study staff before any study procedures. Participants can earn up to US $250 in gift cards by completing study questionnaires and interviews. Participants receive US $50 for completing the baseline questionnaire and a qualitative interview, and US $25 for completing each follow-up questionnaire. There is also a monthly raffle for ecological momentary assessment (EMA) completion, in which participants have a chance to win an additional US $25 for completing an EMA questionnaire during that month. If a person declines participation in the randomized trial, they are asked whether they would participate in the more passive observational monitoring arm, which includes a baseline questionnaire and prospective medical record data collection. Participants in the observational cohort are compensated with a US $50 gift card at the time of study consent and completion of the baseline questionnaire.

A reliance review institutional review board (IRB) system is in place, with agreements from all sites to accept the lead study site IRB for regulatory oversight. The general consent form is written in simple, plain language and is devoid of logos and colors. One study site is required to use its own institution's regulatory language; however, study information is consistent across all sites. The primary investigative team is clearly identified in the consent form, along with corresponding contact information.

Enrollment, Randomization, and Data Collection Procedures

All participants are consented, enrolled, and randomized at the partner institution where they receive care, in partnership with the PoPIT staff, who are needed to set up the user's PPA account. Specifically, following in-person consent at the partner site, all participants complete a preloaded baseline questionnaire on a study-dedicated tablet in the presence of study staff. While they are completing the questionnaire, the trial participant is randomly assigned in REDCap to either the intervention (immediate PPA start) or control (6-month delayed PPA start) arm at a 1:1 ratio. Participants randomized to the delayed-start arm receive usual clinic care before beginning the clinical intervention. Given our pilot findings indicating a stronger PPA effect on HIV viral suppression among participants in the youngest age group (ie, <24 years), we employ a block randomization strategy in which study arm assignment occurs separately for participants within the age groups 13-24 and 25-34 years.

Following randomization, site study staff assist participants randomized to the immediate-start arm in downloading the PPA from the public app store using the participant's personal phone. Site study staff guide participants through in-app tutorials demonstrating how to use app features. Participants randomized to the delayed-start arm are provided with a folder containing general HIV self-management information and contact numbers for the clinic and the site coordinator. Data are collected via online forms (enrollment and medical record review), questionnaires, and SMS text messaging. Participants are consented and enrolled by study staff in person, and the baseline questionnaire is completed independently at the same visit. Remaining follow-up questionnaires at 3, 6, 9, and 12 months are prompted by an SMS text message or email and completed

XSL•FO
RenderX

online. Participants and study staff are aware of study arm assignment. Delayed-start participants are given app access after completing the 6-month questionnaire. Site study staff follow the same procedure to assist with downloading the app and reviewing tutorials for participants in both the immediate- and delayed-start arms.

Intervention

As described above and elsewhere, PPA is a password-protected, multifunction mobile app for supporting retention in HIV care and viral suppression [17]. Before clinical trial use, the color scheme and fonts were refreshed, the digital architecture of the app was evaluated and debugged to ensure reliability, and accessibility was upgraded to comply with Web Content Accessibility Guidelines (WCAG) 2.1 Level AA and Section 508 standards. All original app functions—that is, the article feed, first-person narratives, customizable reminders, local resources, community forum, and private chat function—remain organized as they were in the pilot study. A Positive Peers brand guide was created in the initial design phase by a professional marketing team to preserve PPA's visual and content design. Informative articles are written by agency copywriters and study staff and are then vetted by clinicians. Participants are not directed or required to use the app in any specific way.

PPA is continuously monitored by a dedicated digital design team and PoPIT study staff. Conferences are held regularly to address issues identified by app administrators, reported by users, or noted by other study staff. The study coordinator relies on an electronic reporting system to alert the technical team, who respond promptly and effectively to app issues or bugs. PPA provides instructions for use, community standards, a privacy policy, and accessibility standards in an "About" tab within the app. All participants are made aware of data collection practices during informed consent. PPA does not sell or display advertising or rely on commercial sponsorships at this stage of development.

Data Security and Accessibility

Existing infrastructure and security features of the PPA technology allow comprehensive and increasingly secure storage of PPA activity data. A centralized REDCap data management system is used to manage all study records and data. Unique modules are designed to securely store and track site recruitment, enrollment, informed consent and randomization, study questionnaires, qualitative interview participation, medical record review, EMAs, and study event reminders. All PPA usage data are stored on a secure, cloud-based, Health Insurance Portability and Accountability Act (HIPAA)– and System and Organization Controls 2 (SOC 2)–compliant server without any intrinsic ties to users' demographic identities. Each study participant is assigned an ID upon enrollment, which is connected to their app and project activity. Questionnaire data are entered directly into and maintained in a secure REDCap database. PPA is compliant with WCAG for people with disabilities [73].

Study Staff Training

The PoPIT study team delivers study training to staff at each collaborating site. Study staff at the trial sites reflect a range of expertise, with some highly trained in research support and others new to research work. The PoPIT training curriculum is organized into 3 parts: (1) orientation to the study, with emphasis on study staff roles and study flow (ie, prospective data collection); (2) specific instruction in the enrollment protocol, including review of study documents, consent, and randomization procedures; and (3) steps for downloading the PPA app and engaging with app functions. Staff new to research complete online basic research training through the host institution. PoPIT study instruction includes the formulation of a site-specific recruitment and promotion plan, as well as mock enrollment with PoPIT staff. The PoPIT study team conducts these trainings both in person and virtually, with an in-person session conducted at the site initiation visit. New site staff receive standardized onboarding in line with the training curriculum before actively enrolling participants in the study. Initial study training is supported by monthly site-specific videoconferencing meetings to reinforce the protocol and troubleshoot technical or other study-relevant problems. Tip sheets and notes supporting study procedures are maintained in a secure shared cloud storage system available to all study staff.

Site principal investigators and study staff meet monthly to discuss recruitment, best practices for patient enrollment, share successes and challenges, and troubleshoot issues. Concurrent with staff onboarding and site initiation, sites provide information to the PoPIT study team for integration into the PPA's app infrastructure, including local contacts, logos, general information about the organization, and community resources that clinic staff wish to make readily available to their patient population. These resources include relevant medical clinics and contact information for local social service organizations. Many sites choose to include contacts for housing assistance, substance use and addiction support, mental health providers, food pantries, and support groups. These resources are transcribed into the app's geographically organized resource database. All resources added to the database are independently verified by the PoPIT study team before being manually added to the app's infrastructure.

PPA Administrators

Each study site has a clinical study leader and a dedicated PPA administrator to carry out study activities. Some sites engage additional staff to support study-related activities such as recruitment, enrollment reporting, and medical record review. Study site PPA administrators receive separate training to maintain an online presence in the app that is welcoming but secondary to study participant interaction. For example, new participants are sent a welcome message by their site PPA administrator after entering the app. PPA administrators are asked to refrain from creating new posts in the community forum to avoid an administrator-dominated conversation, while commenting on posts when applicable. To date, as study enrollment has increased, PPA study app administrators have not initiated forum posts or discussions. Local PPA administrators are asked to maintain an unobtrusive presence in the app by introducing new users, providing occasional local or regional announcements, or engaging in private chats with participants to offer troubleshooting or technical support.

XSL•FO
RenderX

Generally, the PPA intervention is entirely user-directed and does not require additional human involvement. However, if the app is adopted as a clinical tool, a staff member at the clinic should be assigned to manage recruitment, registration, in-app support, and navigation services.

Reminders and Notifications

The PoPIT study uses an automated alert and reminder system for follow-up questionnaire invitations. Reminders are keyed to each participant's study enrollment date and study arm through the REDCap database system. Based on participant preferences, questionnaire invitations and reminders are sent via SMS text message or email. If contact information changes during the study, this is either reported directly by a participant to site staff or becomes apparent when a routine questionnaire invitation or reminder message is returned as undeliverable. Regular verification of contact information is emphasized to site staff as a critical component of project implementation and is continuously monitored. First, this routine outreach is an integral component of the PPA, functioning as an extension tool for clinical staff to promote engagement in care and support strategies outside the clinical encounter. Outside the trial environment, users receive notifications about updates and new features. Second, PPA includes a customizable notification system for the medication tracker, community forum, chat function, and event scheduling. Users can enable or disable these reminders and schedule the dates and times at which they are sent. Users may also opt in or out of notifications related to forum and chat activity. A third type of notification used in PPA is EMAs.

Ecological Momentary Assessments

EMA is a method widely used to study behavior and mood in the settings in which they naturally occur [74,75]. Because of this immediacy, EMAs can support ecological validity by mitigating the memory limitations inherent in retrospective self-reports. Four general characteristics define EMAs: (1) data collection in real-world environments, (2) a focus on the individual's current state of mind, (3) strategic selection of assessment moments, and (4) repeated assessments over time. Although the PPA records and maintains a log of each user action within the app, it remains unclear how users perceive the app while actively using it.

The PPA is premised on a user-centered model of digital media effects [59,61,63]. This approach suggests that users' needs and the characteristics of the technology (ie, its affordances) interact to shape user engagement. Greater engagement with the mobile app is expected to influence self-efficacy in HIV management and increase the likelihood of favorable HIV clinical outcomes. As EMAs occur repeatedly in real time, they provide an opportunity to assess user engagement with the PPA more immediately than retrospective self-report. The goal of the EMA methodology is to link tracked physical interactions with the app (ie, natively recorded usage data) to users' immediate perceptions of engagement in real time.

User engagement with interactive media comprises 4 validated attributes: (1) physical interaction, (2) interface assessment, (3) absorption, and (4) digital outreach [76]. Analyses indicate that greater physical interaction with the app and more positive assessments of the interface predict cognitive absorption with the content, which in turn is associated with stronger behavioral intentions to manage and socially share the content [76]. Our conceptual model has focused primarily on frequencies of physical interaction. EMAs can provide insight into user perceptions of the app interface and the extent to which users feel absorbed or attentive during use. We may also include a single-item measure of digital outreach in the form of the likelihood of recommending the app to others. Ultimately, we anticipate that increased user engagement will be associated with higher medication adherence, lower viral load, greater engagement in care, and reduced stigma. In this study, EMAs are delivered as SMS text messages sent to participants. The first EMA is deployed to all users 1 month after they begin using the app and focuses on physical interaction with the PPA (see Table 2). The second EMA is sent 2 months after app initiation and includes brief questions assessing current participant engagement with the PPA interface, as well as 4 items adapted from the System Usability Scale [77]. These micro-questionnaires serve as evaluative tools for assessing intervention implementation and effectiveness.

XSL•FO
RenderX

Table 2. EMA[a] items at time 1 and time 2.

EMA items	Response option
Time 1 EMA items: interaction with the PPA[b]	
Is the Positive Peers mobile app downloaded to your smartphone right now?	• Yes/no
Did you use the Positive Peers mobile app today?	• Yes/no
How many days in the last week did you use the Positive Peers mobile app?	• Daily • 4-6 days • 2 or 3 days • Once • None
Did you take your medication yesterday?	• Yes/no
Time 2 EMA items: interface assessment, absorption, and digital outreach	
When was the last time you used the Positive Peers app?	• I haven't used the Positive Peers app. • Yesterday or today • Within the last week • Within the last month
In general, the app is interesting to me.	• Strongly agree to strongly disagree
This app holds my interest.	• Strongly agree to strongly disagree
While I was using the app, I was able to focus on the task at hand.	• Strongly agree to strongly disagree
I would recommend this app to other people like me.	• Yes/no

[a]EMA: ecological momentary assessment.

[b]PPA: Positive Peers mobile app.

Quantitative Data Collection and Measures

Data collection is planned for a total of 18 months for each participant. Three modes of data collection are used in the study. First, site staff collect medical record data, including date of diagnosis, HIV viral load, T-cell counts, and office visit dates, for the 12 months before enrollment and for 18 months after enrollment.

Second, participants receive an email or SMS text message notification to complete follow-up electronic questionnaires at 3, 6, 9, and 12 months after enrollment. Automated reminders are sent on the scheduled completion date; if the questionnaire is not completed, additional reminders are sent 1 week and 2 weeks later. If the questionnaire coincides with an office visit or cannot be completed remotely for any other reason, study staff can administer the pending questionnaire on site using a laptop or iPad (Apple Inc). Delayed-start participants are onboarded to the app, either in person or remotely by site staff, 6 months after enrollment. This group completes baseline, 3-month, and 6-month questionnaires before downloading and initiating app use.

Third, during the intervention period, EMAs are used to collect in-the-moment information about app use (eg, "Is the Positive Peers mobile app currently downloaded to your phone?") and related perceptions. PPA provides a natural technological interface through which such perceptions or behaviors, and their correlates within the app or the broader social environment, can be assessed in a nuanced, dynamic, and longitudinal framework.

Measures

Selection of Standardized and Validated Study Measures

Selected study measures make use of the PhenX toolkit [78], Patient-Reported Outcomes Measurement Information System (PROMIS) Toolbox [79,80], and Health Resources and Services Administration (HRSA) HIV/AIDS Bureau reporting standards [81]. Measures were selected based on our preliminary work, theoretical model, and relevance to the population (Table 3). The PhenX Toolkit is a searchable online database of recommended measurement protocols for human studies [82]. Database protocols are well established, broadly validated, reliable, and low burden for study participants. Measures for this study, derived from the PhenX toolkit, include a set of demographic characteristics and social determinants of health, as well as the Medical Outcomes Study Social Support Survey.

XSL•FO
RenderX

Table 3. PoPIT[a] constructs, measures, and prospective data collection.

Construct	Measure	Collection
Participant characteristics [82]	• PhenX: age, ethnicity and race, biological sex, sexual orientation, sexual identity, marital status, and educational attainment	Baseline[b]
Social determinants of health [83]	• AAFP[c] Social Needs Screening Tool: housing, food and transportation insecurity, utilities, employment, finances, and personal safety • PhenX: internet access, access to health technology, and general social media use	Baseline and 12 months
Attitudes toward technology [84]	• Social Media Usage & Attitudes Scale (MTUAS[d])	Baseline, 6, and 12 months
Technology affordances [85]	• Mobile Application Affordances Measure	3 and 6 months after app initiation
PPA[e] engagement	• Number of logins, duration (minutes), number of sessions, and engagement counts for each PPA function (articles, med tracker, resources, forum, and chat)	Concurrent with participant app use
Self-efficacy for managing chronic illness [86]	• PROMIS[f] subscales: managing emotions, social interactions, medication adherence, and general self-efficacy	Baseline, 3, 6, 9, and 12 months
Social support [87]	• Medical Outcomes Study • Subscales: emotional/informational support, tangible/instrumental support, and positive social interaction	Baseline, 3, 6, 9, and 12 months
Perceived HIV stigma [88]	• HIV Stigma Scale subscales: personalized, disclosure, negative self-image, and public attitudes	Baseline, 3, 6, 9, and 12 months
Engagement in HIV care [89]	• Pre- and postclinic visit dates and count	12 months before enrollment and 18 months after study completion
Viral suppression [89]	• Preenrollment CD4 count and date • Pre- and post-HIV viral load values	12 months before enrollment and 18 months after study completion

[a]PoPIT: Positive Peers Intervention Trial.

[b]Baseline data collection occurs at consent.

[c]AAFP: American Academy of Family Physicians

[d]MTUAS: Media and Technology Usage and Attitudes Scale.

[e]PPA: Positive Peers mobile app.

[f]PROMIS: Patient-Reported Outcomes Measurement Information System.

Participant Characteristics

Measured sociodemographic characteristics collected at baseline include age, ethnicity and race, biological sex, sexual orientation, sexual identity, marital status, and educational attainment. Measures were derived from the PhenX toolkit [82]. The toolkit measures are concise and widely used across health behavior studies involving marginalized groups.

Social Determinants of Health

The American Academy of Family Physicians developed an index of brief measures to serve as a screening tool for social determinants of health [83]. Measured social determinants of health include employment, housing, food and transportation insecurity, finances, and personal safety. The PhenX Toolkit provided measures of internet access, access to health technology, and general social media use.

Attitudes Toward Technology

The Media and Technology Usage and Attitudes Scale (MTUAS) [84] is a comprehensive 60-item measure designed to assess behaviors and attitudes across a wide variety of technologies and platforms. This study uses a version of the MTUAS attitudes subscale adapted to the PPA. This scale includes 12 five-point response items focused on measuring positive and negative attitudes toward technology use in general, as well as a third subscale that measures technology dependence. The subscales are reliable, and scores are significantly correlated with other measures of daily media use and technology-related anxiety.

Technology Affordances

Affordances are defined as features of user-technology interactions that offer gratification of perceived needs [85,89]. These features are independent of content and reflect how a user experiences a particular type of technology. Although more or less specific typologies of media affordances have been

XSL•FO
RenderX

presented [62,63,90], this study relies on a set of 10 affordances (39 items) typical of social media use and tailored to mHealth app use. These include accessibility, bandwidth, social presence, privacy, network association, personalization, persistence, editability, conversation control, and anonymity. Items were adapted to PPA use, for example, "The Positive Peers app allows me to carefully craft my message before sending it" and "The Positive Peers app makes it seem like the other person is present," and were set on a 7-point Likert response scale.

PPA Engagement

Following pilot study methods, intervention exposure is measured in terms of frequency of app use (number of logins and time spent) and engagement with its functions (eg, number of articles read, number of entries into a function, number of posts or replies) [17]. Participant intervention exposure is tracked natively in the PPA in real time and associated with the user's study ID. Every screen tap for each app function is counted and aggregated in monthly reports. Community forum posts and replies are tracked verbatim, with time and date stamps. User data are separated from administrator and study staff data for evaluation using unique study IDs.

Self-Efficacy for Managing Chronic Conditions

Self-efficacy is measured in this study using the PROMIS instruments on self-efficacy for managing chronic conditions [91] and includes self-efficacy for managing (1) medications and treatments (4 items), (2) emotions (8 items), and (3) social interactions (8 items). Published validation efforts showed good concurrent validity and internal consistency [91]. Subscale items have 5 response options ranging from 1=I am not at all confident to 5=I am very confident.

Social Support

The PPA offers social support from others through the community forum and chat functions. Perceptions of social support will be tracked prospectively from baseline using the Medical Outcomes Study Social Support measure [87]. This measure is highly reliable and predictive of psychosocial outcomes across multiple chronic diseases [92,93].

This version includes 5-point Likert items that measure support across 3 dimensions: (1) emotional/informational support, (2) tangible/instrumental support, and (3) positive social interaction (ie, companionship). A fourth subscale, affectionate support, was beyond the scope of this study and was not included in the questionnaires.

Outcomes

Perceived HIV Stigma

The 12-item (short version) Perceived HIV Stigma Scale is used to measure overall perceived stigma and distinct dimensions (3 items each), including (1) personalized stigma, (2) disclosure concerns, (3) concerns about public attitudes, and (4) negative self-image. A 4-point Likert response format assesses each domain. Scores can be summed or averaged for each 3-item subscale and across all 12 items to yield a total stigma score, with higher scores indicating greater perceived stigma. Psychometric evaluation in a national sample of adults with HIV supported a 4-factor structure with good fit and acceptable internal consistency across subscales.

Clinical Outcomes

Retention in HIV care and viral suppression are measured using standardized performance measures developed by the HRSA/HIV/AIDS Bureau [89]. Measures selected for this study are retention in HIV medical care (2 or more visits separated by at least 90 days for HIV care within the measurement year) and HIV load suppression (the most recent viral load <200 copies/mL during the measurement year). Values for outcome variables (engagement in care and viral suppression) are derived from a retrospective medical record review conducted by site study staff using a study-specific form. Engagement in care is determined by at least one office visit with an HIV clinician per 6-month interval for 12 months before and 6 months after study enrollment. Viral suppression is determined by assessing the participant's viral load, as noted in laboratory values, typically required at least every 6 months per the standard of care. A value of <200 copies/mL is considered suppressed. Values for each variable are established for the year before study enrollment, at baseline enrollment, and then at 6-month intervals for 18 months.

Power Analysis

Group sample sizes of 106 in the Positive Peer group and 106 in the control group achieve 90% power to detect a difference between group proportions of 0.20. This is the minimum sample size required in each group to detect the desired difference in proportions. The proportion in group 1 (the treatment group) is assumed to be 0.60 under the null hypothesis and 0.80 under the alternative hypothesis. The proportion in group 2 (the control group) is 0.60. The test statistic used is the 2-sided Z test with pooled variance. The significance level of the test was set at .05. Several assumptions regarding baseline viral suppression were tested based on the literature. The current sample size is based on 63% viral suppression at baseline and a planned 80% in the intervention group. If baseline viral suppression is lower, a smaller sample size would be needed to demonstrate a 20% improvement in viral suppression. Based on this estimation, we aim to recruit and randomize 250 eligible individuals to the study from 6 clinic sites over 24 months, allowing for potential loss to follow-up. It is important to note that patients without documentation of a clinic visit during the assessment period are designated as out of care and will be assumed to be HIV virally unsuppressed. Available data on other outcomes and covariates, however, will be handled using mixed-effects models, which employ maximum likelihood estimation and assume missing-at-random patterns to support valid inferences.

Planned Statistical Analysis

Aims 1 and 2 entail data collection and analysis to provide evidence of PPA effectiveness. Outcomes include retention in HIV care (proportion of enrolled patients with laboratory results in each 6-month period of the study), HIV load suppression (proportion of enrolled patients whose viral load measurement at 6 months in the study is <200 copies/mL), as well as self-reported psychosocial variables, specifically perceived technology affordances, attitudes toward technology, and

XSL•FO
RenderX

self-efficacy for managing chronic conditions. We expect that key demographics, technology covariates, and PPA use will predict self-efficacy, which will be directly associated with (1) retention in care, (2) viral suppression, and (3) HIV-related perceived stigma. Additionally, we plan to test potential covariate mediation and examine theoretical linkages in our model of PPA effects.

Analyses of repeated measures will be performed using generalized linear mixed-effects models with a logit link function for binary outcomes (ie, viral suppression and engagement in care) and linear mixed-effects models for continuous outcomes (eg, perceived stigma, resilience). Time will be treated as a fixed, repeated factor within patients, and patient-specific random effects will be included to account for within-patient correlation. To appropriately model the hierarchical structure of observations nested within patients and patients nested within sites, the site will be included as a random effect, providing an estimate of between-site variability. Multilevel mediation models will be used to evaluate potential mediating effects of covariates. *P* values will be adjusted using Bonferroni or Benjamini-Hochberg correction procedures to control the family-wise type I error rate due to multiple comparisons across distinct psychological outcomes.

Qualitative Data Collection

Longitudinal Semistructured Interviews and Interpretive Thematic Analysis

In-depth semistructured interviews with participants explore experiences with PPA initiation and use, including participants' reflections on joining a mobile app community focused on HIV care. Interviews are conducted longitudinally at approximately 3, 7, and 11 months after initiating use within both immediate- and delayed-start cohorts, with largely independent samples of 8-10 participants from each cohort at each time point. Interviews explore how participants learn, navigate, and maintain PPA use, focusing on perceived benefits, costs, and motivations for ongoing engagement. Anchoring the convergence of mixed methods, interviews also elicit reflections on YPWH's experiences with the study's measured quantitative constructs (eg, experienced social needs that thwart or facilitate engagement in HIV care, perspectives and attitudes toward social media use, perceived self-efficacy for health management).

For sampling and analysis, participants are selected purposively to balance participation across key demographic strata of interest (eg, racial and ethnic identity, gender identity). All interviews are professionally transcribed verbatim and analyzed using NVivo (Lumivero) software with an interpretive, phenomenologically informed approach [94] and a constant comparative approach [95] to thematic analysis. Our approach to analysis is phenomenologically informed in that study participants share a core set of experiences related to their identities as YPWH. To enhance rigor and validity, a subgroup of 5 YPWH participants will serve as thematic auditors and critically respond to qualitative findings and interpretations.

Qualitative Analysis

In support of aims 2 and 3, interviews with a subset of the delayed- and immediate-start cohorts will focus on YPWH's experiences with initiating and continuing PPA use, as well as any thoughts or concerns participants may have had about joining a mobile app community focused on HIV care. Anchoring the convergence of methods, interviews will also elicit YPWH's reflections on their experiences with the study's measured quantitative constructs (eg, experienced social needs that thwart or facilitate engagement in HIV care, perspectives and attitudes toward social media use, perceived self-efficacy for health management). Interviews will be conducted 2-4, 6-8, and 10-12 months after initiating use. These interviews will explore YPWH's experiences with learning, navigating, and maintaining PPA intervention use; for example, interviews will focus on eliciting how YPWH make sense of perceived benefits or costs of intervention use and what personal motivations facilitate ongoing engagement.

Participants from both cohorts will be asked to reflect on and compare their thoughts and preferences defined before engaging with the PPA intervention. Approximately 40 participants across the 6 clinic sites will be selected, stratified by immediate- (n=20) and delayed-start (n=20) intervention engagement. Based on prior research in similar phenomenological investigations of clinical interventions, thematic saturation is expected to be reached with interviews involving between 8 and 20 participants. Enrollment will continue until thematic saturation is achieved (ie, until no new thematic insights are identified). Qualitative participants will be selected purposively to ensure balanced participation across demographic strata of interest (eg, racial and ethnic identity, socioeconomic position). All interviews will be professionally transcribed verbatim.

Interviews will be coded using NVivo qualitative data analysis software. Our approach to analysis is phenomenologically informed, as study participants share a core set of experiences related to their identities as YPWH. Given the goal of convergent interpretation with quantitative study results, a thematic constant comparative approach will be used to analyze transcripts. Analysis will consist of thematic organization and concept clustering, conducted to illuminate participants' health maintenance preferences and their holistic understanding and experience of the measured outcomes.

A mixed methods convergence matrix will further clarify the synthesis, overlap, and divergence of key study findings and will provide a validity check. The interview guide, informed by the perspectives of patients in the demonstration pilot, will be refined before study initiation and following the initial rounds of interviews. Interviews will be conducted via secure videoconferencing. Two site personnel will serve as thematic auditors to critically respond to qualitative findings and interpretations. Reporting of qualitative and mixed methods findings will adhere to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines [96].

Ethical Considerations

This study is being conducted in accordance with guidelines outlined in the Declaration of Helsinki and the Common Rule

XSL•FO
RenderX

[97]. The PoPIT protocol was approved on October 26, 2023, by the MetroHealth System IRB (Protocol: CR00001226), and a reliance agreement was established with each of the academic sites. The 2 federally qualified health centers do not have their own IRB and deferred to the MetroHealth IRB. Informed consent is completed in person at the enrollment visit.

Results

The research protocol (1R01MD019185-01) was funded for US $903,363.00 in direct costs by the National Institute for Minority Health and Health Disparities on September 24, 2023, for 5 years. The study began recruiting patients on June 3, 2024, and will continue accrual until November 27, 2026.

Discussion

Anticipated Findings

Mobile apps to support chronic disease management are appealing, and evidence for their efficacy is building [35,98,99]. Social networking technologies offer quick and constant access to social ties and create diverse virtual social networks, despite potentially discriminating social factors such as age or race [100,101]. These virtual interpersonal networks allow users the opportunity to observe and create norms for positive self-care behaviors, including illness self-management and social support [50,58,100].

The goal of this project is to evaluate the PPA as a clinic-based tool for helping patients access peer support, remain in HIV care, and sustain viral suppression. Although traditional models of patient navigation can be effective for PWH in care [102,103], navigation augmented with tracking tools and a constant source of diverse, relevant information paired with peer experience can (1) rapidly build trust, (2) demystify HIV care complexity, and (3) build self-efficacy toward independent self-care. HIV clinic or nonprofit staff can use the PPA to offer navigation services ranging from medication and viral load testing reminders to peer support and assistance with accessing local living resources. Study results will determine whether this tool helps organizations and clinics reach the Ending the Epidemic goals. Furthermore, data will inform optimal deployment of this mHealth tool for newly diagnosed and reengaging PWH.

Concerns about mHealth scale-up suggest a gap between what researchers consider meaningful and appropriate and what target users find relevant [104,105]. Some social media–based interventions have relied on generic commercial platforms adapted to a disease with little input from end users [106,107]. The PPA was conceived and continues to be shaped by YPWH. Formative app development included user input and decision-making for most app features, content, and imagery, including the look of the app (eg, avatars, backgrounds, badges), article topics, resource needs, and the arrangement of features and functions [17]. The PPA offers a decentralized pathway to care engagement grounded in peer solidarity rather than structured, top-down case management approaches [108]. Clear evidence of effectiveness will allow us to expand PPA availability to new regions and infuse vitality and diversity into the existing PPA social network.

Conclusions

The PPA offers a mobile social network that provides instrumental, emotional, and tangible support in a space where the stigma of HIV status is neutralized. Validation of the PPA as a clinical support tool may provide underresourced clinics with a relevant resource for augmenting outreach and psychosocial support. Completion of these aims will contribute to a national partnership among geographically distinct providers and their patients. Broad dissemination, award-winning creative content, and secure, stable, and responsive tracking tools create an unprecedented opportunity to improve HIV outcomes for high-risk and disparity populations of YPWH.

Acknowledgments

We extend our thanks to the young people who visualized this project first, and to those who followed with enthusiasm, altruism, and hope. We declare that no generative artificial intelligence was used by any author of this manuscript.

Funding

This research protocol was fully funded by the National Institute for Minority Health and Health Disparities (grant id 1R01MD019185-01). The study has been funded for a 5-year period starting in September 2023. The study was funded for US$ 903,363.00 in direct costs. The funder has no involvement in the study design, data collection, analysis, interpretation, or the writing of the manuscript.

Authors' Contributions

MMS conceptualized the study, acquired funding, supervised research activities, prepared and wrote the original draft, and managed critical review and revision. LAC contributed to methodology development and project administration and participated in writing and critically reviewing the original draft. JMS conceptualized key study logistics, supervised research activity planning and project administration, and contributed to writing and critical review of the original draft. VKC led the formal analysis, conceptualized and drafted the primary quantitative analysis plan and power analysis for the original draft, and provided critical review. SAL contributed to methodological assessment, formal analysis, and data curation, and participated in writing and critically reviewing the original draft. YB led data curation and contributed to the original draft and critical review. KAB contributed to and drafted the qualitative analytic plan and provided a critical review of the original draft. JSH contributed to study implementation

XSL•FO
RenderX

planning and provided review and revision of the original draft. AKA conceptualized the study, acquired funding, supervised research activities, and contributed to the original draft, critical review, and revision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report from the ZMD1 IN (A1) - National Institute on Minority Health and Health Disparities Special Emphasis Panel, Promoting Viral Suppression and HIV Prevention Interventions Amongst Health Disparity Populations (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 131 KB - resprot_v15i1e87601_app1.pdf]

References

1. HIV diagnoses, deaths, and prevalence: 2025 update. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/hiv-data/nhss/hiv-diagnoses-deaths-and-prevalence-2025.html [accessed 2026-03-09]
2. GBD 2021 HIV Collaborators. Global, regional, and national burden of HIV/AIDS, 1990-2021, and forecasts to 2050, for 204 countries and territories: the Global Burden of Disease Study 2021. Lancet HIV 2024 Dec;11(12):e807-e822 [FREE Full text] [doi: 10.1016/S2352-3018(24)00212-1] [Medline: 39608393]
3. Sullivan PS, Satcher Johnson A, Pembleton ES, Stephenson R, Justice AC, Althoff KN, et al. Epidemiology of HIV in the USA: epidemic burden, inequities, contexts, and responses. Lancet 2021 Mar 20;397(10279):1095-1106. [doi: 10.1016/S0140-6736(21)00395-0] [Medline: 33617774]
4. Allan-Blitz L, Mena LA, Mayer KH. The ongoing HIV epidemic in American youth: challenges and opportunities. Mhealth 2021;7:33 [FREE Full text] [doi: 10.21037/mhealth-20-42] [Medline: 33898602]
5. Babel RA, Wang P, Alessi EJ, Raymond HF, Wei C. Stigma, HIV risk, and access to HIV prevention and treatment services among men who have sex with men (MSM) in the United States: a scoping review. AIDS Behav 2021 Nov;25(11):3574-3604 [FREE Full text] [doi: 10.1007/s10461-021-03262-4] [Medline: 33866444]
6. Dang T, Anderson J, Powell A. Factors associated with retention in HIV care within 12 months postpartum: a retrospective cohort study. J Aquir Immune Defic Syndr 2025;98(5):478-483. [doi: 10.1097/qai.0000000000003593]
7. Crepaz N, Dong X, Wang X, Hernandez AL, Hall HI. Racial and ethnic disparities in sustained viral suppression and transmission risk potential among persons receiving HIV care - United States, 2014. MMWR Morb Mortal Wkly Rep 2018 Feb 02;67(4):113-118 [FREE Full text] [doi: 10.15585/mmwr.mm6704a2] [Medline: 29389918]
8. Casale M, Carlqvist A, Cluver L. Recent interventions to improve retention in HIV care and adherence to antiretroviral treatment among adolescents and youth: a systematic review. AIDS Patient Care STDS 2019 Jun;33(6):237-252 [FREE Full text] [doi: 10.1089/apc.2018.0320] [Medline: 31166783]
9. Sohail MH, Hussen SA, Dougherty Sheff S, Mugavero M, Schneider J, Hightow-Weidman L, et al. An mHealth app-based social capital intervention (PrEP US NoW) to improve sexual health and uptake of pre-exposure prophylaxis among young, Black, sexual minority men: protocol for intervention development and a pilot randomized controlled trial. JMIR Res Protoc 2025 Sep 18;14:e66326 [FREE Full text] [doi: 10.2196/66326] [Medline: 40966683]
10. Agbo CE, Agu MC, Ogieuhi IJ, Chima UE, Okoye CO, Akinmeji O, et al. The role of digital tools in enhancing antiretroviral therapy adherence among adolescents and young adults with HIV: a systematic review of randomized controlled trials. BMC Infect Dis 2025 Mar 13;25(1):353 [FREE Full text] [doi: 10.1186/s12879-025-10724-w] [Medline: 40082764]
11. Marcolino M, Oliveira J, D'Agostino M, Ribeiro A, Alkmim M, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23. [doi: 10.2196/mhealth.8873] [Medline: 29343463]
12. Mustanski B, Garofalo R, Monahan C, Gratzer B, Andrews R. Feasibility, acceptability, and preliminary efficacy of an online HIV prevention program for diverse young men who have sex with men: the keep it up! intervention. AIDS Behav 2013 Nov;17(9):2999-3012 [FREE Full text] [doi: 10.1007/s10461-013-0507-z] [Medline: 23673793]
13. Muessig K, Nekkanti M, Bauermeister J, Bull S, Hightow-Weidman L. A systematic review of recent smartphone, internet and Web 2.0 interventions to address the HIV continuum of care. Curr HIV/AIDS Rep 2015 Mar;12(1):173-190 [FREE Full text] [doi: 10.1007/s11904-014-0239-3] [Medline: 25626718]
14. Ending the HIV epidemic in the US. U.S. Department of Health and Human Services. 2022. URL: http:/www.hiv.gov/federal-response/ending-the-hiv-epidemic/overview [accessed 2026-02-14]
15. Henkhaus M, Hussen S, Brown D, Del Rio C, Fletcher MR, Jones MD, et al. Barriers and facilitators to use of a mobile HIV care model to re-engage and retain out-of-care people living with HIV in Atlanta, Georgia. PLoS One 2021;16(3):e0247328 [FREE Full text] [doi: 10.1371/journal.pone.0247328] [Medline: 33705421]
16. Sullivan P, Hightow-Weidman L. Mobile apps for HIV prevention: how do they contribute to our epidemic response for adolescents and young adults? Mhealth 2021;7:36 [FREE Full text] [doi: 10.21037/mhealth-20-71] [Medline: 33898605]

XSL•FO
RenderX

17. Step M, McMillen Smith J, Kratz J, Briggs J, Avery A. "Positive peers": function and content development of a mobile app for engaging and retaining young adults in HIV care. JMIR Form Res 2020 Jan 30;4(1):e13495 [FREE Full text] [doi: 10.2196/13495] [Medline: 32012035]
18. Step M, McMillen Smith J, Lewis S, Avery A. Using the positive peers mobile app to improve clinical outcomes for young people with HIV: prospective observational cohort comparison. JMIR Mhealth Uhealth 2022 Sep 28;10(9):e37868. [doi: 10.2196/37868] [Medline: 36170001]
19. Step M, Knight K, McMillen Smith J, Lewis S, Russell T, Avery A. Positive peers mobile application reduces stigma perception among young people living with HIV. Health Promot Pract 2020 Sep;21(5):744-754. [doi: 10.1177/1524839920936244] [Medline: 32757838]
20. Rogers E. Diffusion of Innovations. 3rd edition. New York, NY: Free Press; 1983.
21. Centola D. The spread of behavior in an online social network experiment. Science 2010 Sep 03;329(5996):1194-1197. [doi: 10.1126/science.1185231] [Medline: 20813952]
22. Valente T. In: Glanz K K, Rimer B B, Viswanath K K, editors. Health Behavior: Theory, Research, and Practice. 5th edition. New York, NY: Wiley; 2015:205-222.
23. Papacharissi Z. In: Papacharissi Z, editor. A Networked Self: Identity, Community, and Culture on Social Network Sites. New York, NY: Routledge; 2011:304-318.
24. Phua J, Jin S, Kim J. Uses and gratifications of social networking sites for bridging and bonding social capital: a comparison of Facebook, Twitter, Instagram, and Snapchat. Computers in Human Behavior 2017 Jul;72:115-122 [FREE Full text] [doi: 10.1016/j.chb.2017.02.041]
25. Gokcekus S, Firth J, Regan C, Sheldon B. Recognising the key role of individual recognition in social networks. Trends Ecol Evol 2021 Nov;36(11):1024-1035. [doi: 10.1016/j.tree.2021.06.009] [Medline: 34256987]
26. Valente T, Pitts S. An appraisal of social network theory and analysis as applied to public health: challenges and opportunities. Annu Rev Public Health 2017 Mar 20;38:103-118 [FREE Full text] [doi: 10.1146/annurev-publhealth-031816-044528] [Medline: 27992729]
27. Berkman L, Glass T. Social integration, social networks, social support, and health. Social Epidemiology 2000;1(6):137-173 [FREE Full text] [doi: 10.1093/oso/9780195083316.003.0007]
28. Smith KN, Christakis NA. Social networks and health. Annu Rev Sociol 2008 Aug 01;34(1):405-429 [FREE Full text] [doi: 10.1146/annurev.soc.34.040507.134601]
29. Burleson B, MacGeorge E. Supportive communication. In: Knapp M, Daly J, editors. Handbook of Interpersonal Communication. 3rd edition. Beverly Hills, CA: Sage; 2002:374-424.
30. Abramowitz S, Koenig LJ, Chandwani S, Orban L, Stein R, Lagrange R, et al. Characterizing social support: global and specific social support experiences of HIV-infected youth. AIDS Patient Care STDS 2009 May;23(5):323-330. [doi: 10.1089/apc.2008.0194] [Medline: 19320599]
31. Bekele T, Rourke S, Tucker R, Greene S, Sobota M, Koornstra J, Positive Spaces Healthy Places Team. Direct and indirect effects of perceived social support on health-related quality of life in persons living with HIV/AIDS. AIDS Care 2013;25(3):337-346. [doi: 10.1080/09540121.2012.701716] [Medline: 22774876]
32. Cole DA, Nick EA, Zelkowitz RL, Roeder KM, Spinelli T. Online social support for young people: does it recapitulate in-person social support; can it help? Comput Human Behav 2017 Mar;68:456-464 [FREE Full text] [doi: 10.1016/j.chb.2016.11.058] [Medline: 28993715]
33. Walther JB, Jang J. Communication processes in participatory websites. J Comput-Mediat Comm 2012 Oct 10;18(1):2-15. [doi: 10.1111/j.1083-6101.2012.01592.x]
34. Merolli M, Gray K, Martin-Sanchez F. Health outcomes and related effects of using social media in chronic disease management: a literature review and analysis of affordances. J Biomed Inform 2013 Dec;46(6):957-969 [FREE Full text] [doi: 10.1016/j.jbi.2013.04.010] [Medline: 23702104]
35. Taggart T, Grewe M, Conserve D, Gliwa C, Roman Isler M. Social media and HIV: a systematic review of uses of social media in HIV communication. J Med Internet Res 2015 Nov 02;17(11):e248 [FREE Full text] [doi: 10.2196/jmir.4387] [Medline: 26525289]
36. van Uden-Kraan C, Drossaert C, Taal E, Seydel E, van de Laar M. Participation in online patient support groups endorses patients' empowerment. Patient Educ Couns 2009 Jan;74(1):61-69. [doi: 10.1016/j.pec.2008.07.044] [Medline: 18778909]
37. Bartlett Y, Coulson N. An investigation into the empowerment effects of using online support groups and how this affects health professional/patient communication. Patient Educ Couns 2011 Apr;83(1):113-119. [doi: 10.1016/j.pec.2010.05.029] [Medline: 20599338]
38. van Uden-Kraan CF, Drossaert C, Taal E, Shaw B, Seydel E, van de Laar MAFJ. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. Qual Health Res 2008 Mar;18(3):405-417. [doi: 10.1177/1049732307313429] [Medline: 18235163]
39. Arnold E, Sterrett-Hong E, Jonas A, Pollack L. Social networks and social support among ball-attending African American men who have sex with men and transgender women are associated with HIV-related outcomes. Glob Public Health 2018 Feb;13(2):144-158 [FREE Full text] [doi: 10.1080/17441692.2016.1180702] [Medline: 27169632]

XSL•FO
RenderX

40. Wood S, Gross R, Shea J, Bauermeister JA, Franklin J, Petsis D, et al. Barriers and facilitators of PrEP adherence for young men and transgender women of color. AIDS Behav 2019 Oct;23(10):2719-2729 [FREE Full text] [doi: 10.1007/s10461-019-02502-y] [Medline: 30993479]
41. Laschober T, Serovich J, Brown-Smith M, Kimberly J, Lescano C. Decision readiness to disclose human immunodeficiency virus-positive status to family members. Sex Transm Dis 2019 May;46(5):342-346 [FREE Full text] [doi: 10.1097/OLQ.0000000000000966] [Medline: 30985635]
42. Tabler J, Mykyta L, Schmitz R, Kamimura A, Martinez DA, Martinez RD, et al. Getting by with a little help from our friends: the role of social support in addressing HIV-related mental health disparities among sexual minorities in the lower Rio Grande Valley. J Homosex 2021 Jan 28;68(2):269-289. [doi: 10.1080/00918369.2019.1651112] [Medline: 31411539]
43. Mo P, Coulson N. Developing a model for online support group use, empowering processes and psychosocial outcomes for individuals living with HIV/AIDS. Psychol Health 2012;27(4):445-459. [doi: 10.1080/08870446.2011.592981] [Medline: 21854088]
44. Kalichman S, Sikkema K, Somlai A. People living with HIV infection who attend and do not attend support groups: a pilot study of needs, characteristics and experiences. AIDS Care 1996 Oct;8(5):589-599. [doi: 10.1080/09540129650125542] [Medline: 8893909]
45. Flickinger T, DeBolt C, Waldman A, Reynolds G, Cohn WF, Beach MC, et al. Social support in a virtual community: analysis of a clinic-affiliated online support group for persons living with HIV/AIDS. AIDS Behav 2017 Nov;21(11):3087-3099. [doi: 10.1007/s10461-016-1587-3] [Medline: 27766448]
46. Marcus R, Dasgupta S, Taussig J, Tie Y, Nair P, Prejean J. Social determinants of health and unmet needs for services among young adults with HIV: Medical Monitoring Project, 2018-2021. J Acquir Immune Defic Syndr 2025 May 01;99(1):9-19. [doi: 10.1097/QAI.0000000000003605] [Medline: 39780305]
47. Weeks S, Renshaw T, Vinal S. Minority stress as a multidimensional predictor of LGB+ adolescents' mental health outcomes. J Homosex 2023 Apr 16;70(5):938-962. [doi: 10.1080/00918369.2021.2006000] [Medline: 34806974]
48. Hoy-Ellis C. Minority stress and mental health: a review of the literature. J Homosex 2023 Apr 16;70(5):806-830. [doi: 10.1080/00918369.2021.2004794] [Medline: 34812698]
49. Comulada W, Step MM, Fletcher J, Tanner AE, Dowshen NL, Arayasirikul S, Special Projects Of National Significance Social Media Initiative Study Group. Predictors of internet health information-seeking behaviors among young adults living with HIV across the United States: longitudinal observational study. J Med Internet Res 2020 Nov 02;22(11):e18309 [FREE Full text] [doi: 10.2196/18309] [Medline: 33136057]
50. Geber S, Baumann E, Klimmt C. Where do norms come from? Peer communication as a factor in normative social influences on risk behavior. Communication Research 2017 Jul 12;46(5):708-730 [FREE Full text] [doi: 10.1177/0093650217718656]
51. Mulawa MI, Rosengren AL, Amico KR, Hightow-Weidman LB, Muessig KE. mHealth to reduce HIV-related stigma among youth in the United States: a scoping review. Mhealth 2021;7:35 [FREE Full text] [doi: 10.21037/mhealth-20-68] [Medline: 33898604]
52. Sun C, Shato T, Steinbaugh A, Pradeep S, Rivet Amico K, Horvath K. Virtual voices: examining social support exchanged through participant-generated and unmoderated content in a mobile intervention to improve HIV antiretroviral therapy adherence among GBMSM. AIDS Care 2023 Jan;35(1):7-15 [FREE Full text] [doi: 10.1080/09540121.2022.2038364] [Medline: 35156485]
53. Kelleher C. Minority stress and health: implications for lesbian, gay, bisexual, transgender, and questioning (LGBTQ) young people. Counselling Psychology Quarterly 2009 Dec;22(4):373-379 [FREE Full text] [doi: 10.1080/09515070903334995]
54. Toomey RB, Ryan C, Diaz RM, Russell ST. Coping with sexual orientation-related minority stress. J Homosex 2018;65(4):484-500 [FREE Full text] [doi: 10.1080/00918369.2017.1321888] [Medline: 28441107]
55. Williams S, Mann A, Fredrick E. Proximal minority stress, psychosocial resources, and health in sexual minorities. Journal of Social Issues 2017 Sep 21;73(3):529-544 [FREE Full text] [doi: 10.1111/josi.12230]
56. Frost D, Meyer I, Lin A, Wilson BDM, Lightfoot M, Russell ST, et al. Social change and the health of sexual minority individuals: do the effects of minority stress and community connectedness vary by age cohort? Arch Sex Behav 2022 May;51(4):2299-2316 [FREE Full text] [doi: 10.1007/s10508-022-02288-6] [Medline: 35411489]
57. Kohut H. The Restoration of the Self. Chicago, IL: The University of Chicago Press; 2009.
58. Sandstrom GM, Dunn EW. Social interactions and well-being: the surprising power of weak ties. Pers Soc Psychol Bull 2014 Jul;40(7):910-922. [doi: 10.1177/0146167214529799] [Medline: 24769739]
59. Rubin A. Audience activity and media use. Commun Monogr 1993;60(1):98-105 [FREE Full text] [doi: 10.1080/03637759309376300]
60. Kim J, Rubin A. The variable influence of audience activity on media effects. Communication Research 1997 Apr 01;24(2):107-135 [FREE Full text] [doi: 10.1177/009365097024002001]
61. Sundar S, Limperos A. Uses and Grats 2.0: new gratifications for new media. Journal of Broadcasting & Electronic Media 2013 Dec 02;57(4):504-525 [FREE Full text] [doi: 10.1080/08838151.2013.845827]
62. Ronzhyn A, Cardenal A, Rubio A. Defining affordances in social media research: a literature review. A New Media Soc 2023;25(11):3165-3188 [FREE Full text] [doi: 10.1177/14614448221135187]

XSL•FO
RenderX

63. Schrock AR. Communicative affordances of mobile media: Portability, availability, locatability, and multimediality. Int J Comm 2015;15(9):1229-1246 [FREE Full text]
64. Cappella JN, Kim HS, Albarracín D. Selection and transmission processes for information in the emerging media environment: psychological motives and message characteristics. Media Psychol 2015 Jul 01;18(3):396-424 [FREE Full text] [doi: 10.1080/15213269.2014.941112] [Medline: 26190944]
65. Levy M, Windahl S. Audience activity and gratification: a conceptual clarification and exploration. Comm Res 1984;11(1):51-78.
66. Bryant J, Zillmann D. A retrospective prospective look at media effects. In: Nabi RL, Oliver MB, editors. The Sage Handbook of Media Processes and Effects. Thousand Oaks, CA: Sage; 2009:9-17.
67. Rathnayake C, Winter J. Carrying forward the uses and Grats 2.0 agenda: an affordance-driven measure of social media uses and gratifications. Journal of Broadcasting & Electronic Media 2018 Jul 12;62(3):371-389 [FREE Full text] [doi: 10.1080/08838151.2018.1451861]
68. Bucher T, Anne H. The affordances of social media platforms. In: Burgess J, Poell T, Marwick A, editors. The SAGE Handbook of Social Media. Thousand Oaks, CA: Sage; 2017:233-254.
69. Rice R, Evans S, Pearce K, Sivunen A, Vitak J, Treem J. Organizational media affordances: operationalization and associations with media use. J Commun 2017 Jan 17;67(1):106-130 [FREE Full text] [doi: 10.1111/jcom.12273]
70. Feldstein A, Glasgow R. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. Jt Comm J Qual Patient Saf 2008 Apr;34(4):228-243 [FREE Full text] [doi: 10.1016/s1553-7250(08)34030-6] [Medline: 18468362]
71. Rabin B, Cakici J, Golden C, Estabrooks P, Glasgow R, Gaglio B. A citation analysis and scoping systematic review of the operationalization of the Practical, Robust Implementation and Sustainability Model (PRISM). Implement Sci 2022 Sep 24;17(1):62 [FREE Full text] [doi: 10.1186/s13012-022-01234-3] [Medline: 36153628]
72. Fauci A, Redfield R, Sigounas G, Weahkee M, Giroir B. Ending the HIV epidemic: a plan for the United States. JAMA 2019 Mar 05;321(9):844-845. [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
73. WCAG 2 overview. Web Accesibility Initiative. URL: https://www.w3.org/WAI/standards-guidelines/wcag/ [accessed 2025-05-06]
74. Courvoisier D, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: the effect of time and personality characteristics. Psychol Assess 2012 Sep;24(3):713-720 [FREE Full text] [doi: 10.1037/a0026733] [Medline: 22250597]
75. Shiffman S, Stone A, Hufford M. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
76. Oh J, Bellur S, Sundar SS. Clicking, assessing, immersing, and sharing: an empirical model of user engagement with interactive media. Communication Research 2015 Sep 21;45(5):737-763. [doi: 10.1177/0093650215600493]
77. Brooke J. SUS: a retrospective. J Usability Stud 2013;8(2):29-40 [FREE Full text]
78. Krzyzanowski M, Ives C, Jones N, Entwisle B, Fernandez A, Cullen TA, et al. The PhenX toolkit: measurement protocols for assessment of social determinants of health. Am J Prev Med 2023 Sep;65(3):534-542 [FREE Full text] [doi: 10.1016/j.amepre.2023.03.003] [Medline: 36935055]
79. Salsman JM, Schalet BD, Merluzzi TV, Park CL, Hahn EA, Snyder MA, et al. Calibration and initial validation of a general self-efficacy item bank and short form for the NIH PROMIS. Qual Life Res 2019 Sep;28(9):2513-2523 [FREE Full text] [doi: 10.1007/s11136-019-02198-6] [Medline: 31140041]
80. Schalet B, Pilkonis P, Yu L, Dodds N, Johnston KL, Yount S, et al. Clinical validity of PROMIS depression, anxiety, and anger across diverse clinical samples. J Clin Epidemiol 2016 May;73:119-127. [doi: 10.1016/j.jclinepi.2015.08.036] [Medline: 26931289]
81. Valdiserri RO, Forsyth AD, Yakovchenko V, Koh HK. Measuring what matters: development of standard HIV core indicators across the U.S. Department of Health and Human Services. Public Health Rep 2013;128(5):354-359 [FREE Full text] [doi: 10.1177/003335491312800504] [Medline: 23997280]
82. Hamilton C, Strader L, Pratt J, Maiese D, Hendershot T, Kwok RK, et al. The PhenX toolkit: get the most from your measures. Am J Epidemiol 2011 Aug 01;174(3):253-260 [FREE Full text] [doi: 10.1093/aje/kwr193] [Medline: 21749974]
83. O'Gurek DT, Henke C. A practical approach to screening for social determinants of health. Fam Pract Manag 2018;25(3):7-12 [FREE Full text] [Medline: 29989777]
84. Rosen L, Whaling K, Carrier L, Cheever NJ, Rokkum J. The media and technology usage and attitudes scale: an empirical investigation. Comput Human Behav 2013 Nov 01;29(6):2501-2511 [FREE Full text] [doi: 10.1016/j.chb.2013.06.006] [Medline: 25722534]
85. Fox J, McEwan B. Distinguishing technologies for social interaction: the perceived social affordances of communication channels scale. Communication Monographs 2017 Jun 28;84(3):298-318 [FREE Full text] [doi: 10.1080/03637751.2017.1332418]
86. Gruber-Baldini A, Velozo C, Romero S, Shulman LM. Validation of the PROMIS measures of self-efficacy for managing chronic conditions. Qual Life Res 2017 Jul;26(7):1915-1924 [FREE Full text] [doi: 10.1007/s11136-017-1527-3] [Medline: 28239781]

XSL•FO
RenderX

87. Sherbourne C, Stewart A. The MOS social support survey. Soc Sci Med 1991;32(6):705-714. [doi: 10.1016/0277-9536(91)90150-b] [Medline: 2035047]
88. Reinius M, Wettergren L, Wiklander M, Svedhem V, Ekström AM, Eriksson LE. Development of a 12-item short version of the HIV Stigma Scale. Health Qual Life Outcomes 2017 May 30;15(1):115 [FREE Full text] [doi: 10.1186/s12955-017-0691-z] [Medline: 28558805]
89. HIV/AIDS Bureau core performance measures. Health Resources and Services Administration. 2019. URL: https://ryanwhite.hrsa.gov/sites/default/files/ryanwhite/grants/core-measures.pdf [accessed 2026-03-10]
90. Wu M, Wu T, Pei Y. What drives health information exchange on social media? Social media affordances and social support perspectives. Health Commun 2024 Nov;39(13):3365-3379. [doi: 10.1080/10410236.2024.2321408] [Medline: 38477307]
91. Hicks S, Jiang W, Kibugi J, Badia J, Richardson BA, Beima-Sofie K, et al. Sociobehavioral determinants of nonretention among adolescents and young adults with HIV. AIDS 2025 Jun 01;39(7):886-898. [doi: 10.1097/QAD.0000000000004141] [Medline: 39912749]
92. Smith C, Fernengel K, Holcroft C, Gerald K, Marien L. Meta-analysis of the associations between social support and health outcomes. Ann Behav Med 1994;16(4):352-362 [FREE Full text] [doi: 10.1093/abm/16.4.352]
93. Robitaille A, Orpana H, McIntosh C. Psychometric properties, factorial structure, and measurement invariance of the English and French versions of the Medical Outcomes Study social support scale. Health Rep 2011 Jun;22(2):33-40 [FREE Full text] [Medline: 21848131]
94. Moustakas C. Phenomenological Research Methods. Thousand Oaks, CA: Sage; 1994.
95. Lincoln YS, Guba EG, Pilotta JJ. Naturalistic Inquiry. Thousand Oaks, CA: Sage; Apr 1, 1985.
96. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
97. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
98. Dillingham R, Ingersoll K, Flickinger TE, Waldman AL, Grabowski M, Laurence C, et al. PositiveLinks: a mobile health intervention for retention in HIV care and clinical outcomes with 12-month follow-up. AIDS Patient Care STDS 2018 Jun;32(6):241-250 [FREE Full text] [doi: 10.1089/apc.2017.0303] [Medline: 29851504]
99. McCoy S, Buzdugan R, Grimball R, Natoli L, Mejia CM, Klausner JD, et al. Stick to it: pilot study results of an intervention using gamification to increase HIV screening among young men who have sex with men in California. Mhealth 2018;4:40 [FREE Full text] [doi: 10.21037/mhealth.2018.09.04] [Medline: 30363751]
100. Kim H. Enacted social support on social media and subjective well-being. International Journal of Communication 2014;8(2014):2340 [FREE Full text]
101. Rains S, Tsetsi E. Social support and digital inequality: does internet use magnify or mitigate traditional inequities in support availability? Communication Monographs 2016 Sep 09;84(1):54-74 [FREE Full text] [doi: 10.1080/03637751.2016.1228252]
102. Arayasirikul S, Turner C, Trujillo D, Le V, Wilson E. Efficacy and Impact of Digital HIV Care Navigation in Young People Living With HIV in San Francisco, California: Prospective Study. JMIR Mhealth Uhealth 2020 May 08;8(5):e18597 [FREE Full text] [doi: 10.2196/18597] [Medline: 32383680]
103. Mizuno Y, Higa D, Leighton C, Roland K, Deluca J, Koenig L. Is HIV patient navigation associated with HIV care continuum outcomes? AIDS 2018 Nov 13;32(17):2557-2571 [FREE Full text] [doi: 10.1097/QAD.0000000000001987] [Medline: 30102661]
104. Giovenco D, Muessig K, Horvitz C, Biello KB, Liu AY, Horvath KJ, et al. Adapting technology-based HIV prevention and care interventions for youth: lessons learned across five U.S. Adolescent Trials Network studies. Mhealth 2021;7:21 [FREE Full text] [doi: 10.21037/mhealth-20-43] [Medline: 33898590]
105. Kankanhalli A, Xia Q, Zhao X. Understanding personalization for health behavior change applications: a review and future directions. THCI 2021 Sep 30;13:316-349. [doi: 10.17705/1thci.00152]
106. Subramaniam A, Hensley E, Stojancic R, Vaughn J, Shah N. Careful considerations for mHealth app development: lessons learned from QuestExplore. Mhealth 2022;8:24 [FREE Full text] [doi: 10.21037/mhealth-21-51] [Medline: 35928515]
107. Lau N, O'Daffer A, Yi-Frazier J, Rosenberg A. Goldilocks and the three bears: a just-right hybrid model to synthesize the growing landscape of publicly available health-related mobile apps. J Med Internet Res 2021 Jun 07;23(6):e27105 [FREE Full text] [doi: 10.2196/27105] [Medline: 34096868]
108. Jones M, Dyer K, Nedell E, Fletcher MR, Grimsley Ackerley C, Hussen SA, et al. One size does not fit all: Preferences for HIV care delivery among out-of-care people living with HIV in the Southeastern United States. PLoS One 2023;18(1):e0276852 [FREE Full text] [doi: 10.1371/journal.pone.0276852] [Medline: 36649350]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
EMA: ecological momentary assessment
HIPAA: Health Insurance Portability and Accountability Act

XSL•FO
RenderX

HRSA: Health Resources and Services Administration
IRB: institutional review board
mHealth: mobile health
MTUAS: Media and Technology Usage and Attitudes Scale
PoPIT: Positive Peers Intervention Trial
PPA: Positive Peers mobile app
PRISM: Practical Robust Implementation and Sustainability Model
PROMIS: Patient-Reported Outcomes Measurement Information System
PWH: people with HIV
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
SNT: Social Networking Theory
SOC 2: System and Organization Controls 2
WCAG: Web Content Accessibility Guidelines
YPWH: young people with HIV

Edited by J Sarvestan; submitted 11.Nov.2025; this is a non–peer-reviewed article;accepted 24.Dec.2025; published 17.Mar.2026.

Please cite as:
Step MM, Catania LA, McMillen Smith J, Lewis SA, Bitar Y, Cheruvu VK, Berg KA, Hallam JS, Avery AK
Evaluation of the Positive Peers Mobile App for Supporting the Viral Suppression of Young People With HIV: Protocol for a Concurrent Mixed Methods Evaluation With Randomized Controlled Clinical Trial and Observational Cohort
JMIR Res Protoc 2026;15:e87601
URL: https://www.researchprotocols.org/2026/1/e87601
doi:10.2196/87601
PMID:

©Mary M Step, L Anthony Catania, Jennifer McMillen Smith, Steven A Lewis, Yanis Bitar, Vinay K Cheruvu, Kristen A Berg, Jeffrey S Hallam, Ann K Avery. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Characterizing AIDS Drug Assistance Program Practices and Policies for Sustained Viral Suppression Using the Consolidated Framework for Implementation Research: Protocol for a Qualitative Study

Kathleen A McManus[1], MD, MSCR; Erin Q Rogers[1], MPH, PhD; Amber Steen[2], MA; Amy Killelea[1], JD; Tim Horn[3], MPH; Reanna Panagides[1], BSN, MSDS; Steven Laymon[1], PhD; Jessica Keim-Malpass[2], RN, CPNP-AC, PhD

[1]Division of Infectious Diseases & International Health, School of Medicine, University of Virginia, Charlottesville, VA, United States

[2]School of Nursing, University of Virginia, Charlottesville, VA, United States

[3]National Alliance of State and Territorial AIDS Directors, Washington, DC, United States

Corresponding Author:
Kathleen A McManus, MD, MSCR
Division of Infectious Diseases & International Health
School of Medicine
University of Virginia
1300 Jefferson Park Ave
Charlottesville, VA, 22903
United States
Phone: 1 434 982 1700
Email: km8jr@uvahealth.org

Abstract

Background: AIDS Drug Assistance Programs (ADAPs), a key part of the federally funded Ryan White HIV/AIDS Program, provide antiretroviral therapy (ART) to low-income and uninsured or underinsured people with HIV. As long as they meet federal regulations, jurisdictions maintain flexibility in implementing ADAPs, allowing for a range of operational and programmatic options, including eligibility criteria, formulary design for ART (and non-ART) drugs, and the nature and character of insurance support. These programmatic and policy decisions ultimately impact ADAP client outcomes, including engagement in care, differences in the populations benefiting from ADAP, and viral suppression rates. Despite ADAPs achieving an 85% national average viral suppression rate, this falls short of the 90% goal needed to end the HIV epidemic, and no formal evaluation has examined how specific operational or policy factors drive success across states.

Objective: This qualitative study aims to examine how operational, organizational, and contextual factors shape the ability of ADAPs to improve viral suppression rates among people with HIV enrolled in ADAPs, specifically by identifying the mechanisms that facilitate high performance and viral suppression for all people with HIV. Results will inform how viral suppression outcomes could be improved through policy, program, operational, and organizational changes in ADAP design and implementation.

Methods: Using a qualitative descriptive approach guided by the Consolidated Framework for Implementation Research, we conducted semistructured interviews with state-level ADAP leaders, operational and programmatic decision-makers, and staff. Interview questions covered 5 domains: leadership and staffing; operations, integration, and collaboration; formulary design and treatment; eligibility and enrollment; and innovation and implementation. Qualitative data from the interviews will be analyzed using deductive and inductive approaches along with situational mapping.

Results: This study was funded in April 2023. In total, 33 individuals from 16 states have participated in this qualitative study, resulting in a participation rate of 31.4% (16/51). States from all 4 US census regions were represented. Interviews were conducted between February 12, 2025, and March 20, 2025. As of March 2026, data analysis is underway.

Conclusions: We anticipate that the study will highlight features, organizational practices, and operational norms that contribute to the successes of ADAPs with high rates of viral suppression. Successful completion of this analysis will provide evidence to inform state and federal regulations, program development, resource allocation, and prioritization to advance the goal of helping people achieve viral suppression and interrupt HIV transmission, which is a vital public health objective. Using other data alongside our qualitative findings, we can provide context for the environment in which these programs or policies impact ADAP clients

XSL•FO
RenderX

and the broader public health objectives of the Ryan White HIV/AIDS Program. Ultimately, we expect that disseminating these results will support improvements in viral suppression, enhancing both individual health outcomes and public health.

International Registered Report Identifier (IRRID): DERR1-10.2196/90008

(JMIR Res Protoc 2026;15:e90008) doi:10.2196/90008

KEYWORDS

AIDS Drug Assistance Program; ADAP; viral suppression; Consolidated Framework for Implementation Research; CFIR; Ryan White HIV/AIDS Program; antiretroviral therapy

Introduction

Despite the availability of safe and effective antiretroviral therapy (ART), the most recent estimate of viral suppression among people with HIV in the United States was 67% in 2023 [1]. To accomplish the goals of the federal Ending the HIV Epidemic initiative, all people with HIV, including those with low incomes, must have uninterrupted, affordable access to ART [2] to achieve sustained viral suppression, which is critical to individual health (reduced morbidity, comorbidities, and mortality) [3,4] and public health (preventing transmission of HIV and thereby reducing incidence) [5,6]. Additionally, viral suppression could reduce health care costs because each HIV infection averted saves up to US $1.2 million [7-9].

As a key part of the federally funded US Ryan White HIV/AIDS Program (RWHAP) health care delivery safety net, AIDS Drug Assistance Programs (ADAPs) provide free ART to people with HIV with low incomes, either through direct provision or ADAP-subsidized insurance plans [10,11]. People with HIV are eligible for ADAP support if they are uninsured or underinsured and have an income less than 200% to 500% of the federal poverty level, depending on their state's eligibility criteria [11,12]. In 2023, ADAPs provided ART for 20% of people with HIV in the United States, with an annual national budget of US $2.5 billion [13]. Although people with HIV enrolled in ADAP achieved a higher average viral suppression rate (85%) [13] than the US average (67%) [1], this rate still falls short of the 90% goal needed to end the HIV epidemic [14]. Viral suppression rates also vary across states and among ADAPs, with a mix of external policy factors and internal programmatic decisions driving these differences. Nevertheless, the 85% viral suppression rate, as a national average across all ADAPs, demonstrates the value of intentionally focusing on delivering attentive care to low-income uninsured or underinsured people with HIV [13].

Since their inception, ADAPs have evolved to meet clients' changing needs and have often expanded their formularies beyond the federal requirement of including at least 1 drug from each class of HIV antiretroviral medications [15]. This improves treatment effectiveness and adherence and addresses individual patient needs. As more ARTs have been approved by the US Food and Drug Administration, ADAPs have expanded their formularies to include new medications, such as single-tablet and long-acting injectable regimens [16]. Many have also incorporated non-ART medications to address comorbidities, such as hepatitis C and substance use disorders, and have developed premium and cost-sharing support to help clients maintain insurance and prescription drug coverage [17].

ADAPs have also adapted to a changing insurance landscape, where far more clients have public and private insurance coverage, both because of Affordable Care Act (ACA), Medicaid, and private insurance expansion and because of an aging population newly qualifying for Medicare [18]. ADAPs have expanded their programs to be able to not only provide free ART to low-income people with HIV without another coverage source but also assist insured people with HIV with their premiums and medication cost sharing [19]. Overall, ADAPs have demonstrated implementation agility by adapting to changes in their client base, the needs of people with HIV, innovations in treatment and care (including evolving formularies), and shifts in policy and funding [13]. With the expiration of pandemic-era continuous Medicaid coverage and significant changes to state Medicaid programs, state ADAPs will need to find new ways to preserve support for people with HIV and accomplish the broader aims of improving viral suppression rates for their clients [20].

Despite their demonstrated adaptability, differences in ADAP use [17] and outcomes [13] have emerged, and the drivers of these differences are poorly understood. Evidence indicates persistent differences in ADAP use, suggesting that smaller proportions of Black people with HIV and women are served by ADAP than White people with HIV and men, respectively [17]. ADAP performance for subgroups of people with HIV is shaped not only by client characteristics but also by the policy and operational contexts in which they function. A key challenge for ADAPs is how they collect, analyze, and act upon client data to assess program performance; identify differences in outcomes; and guide responsive interventions. The ways in which ADAPs approach this challenge will vary by jurisdiction, reflecting differences in context, policy, and infrastructure.

Although all state ADAPs operate under a common federal legislative and regulatory framework, implementation varies considerably [15]. Each ADAP has considerable latitude in how it structures its medication procurement and delivery systems, as well as how it formulates its day-to-day client service, case management, and communications practices [15]. These differences are present in several ways, including the extent to which ADAP assists insured people with HIV with premiums and cost sharing, program eligibility criteria (including income thresholds and recertification frequency), formulary composition (both ART and non-ART drugs), pharmacy network structure, and availability of mail-order drugs. States also vary in client mix and demographics, the percentage of clients supported through health insurance assistance vs full-pay ADAP, and the organizational placement of the ADAP within the broader state health department or social service agencies. Further variation

XSL•FO
RenderX

exists in ADAPs' relationships with the larger public health apparatus, including HIV surveillance and prevention efforts and state Medicaid offices, and in the degree of coordination with community-based and faith-based organizations that promote ADAP awareness, access, and enrollment. Programmatic decisions made by ADAPs also occur against a backdrop of budget constraints [21], with relatively flat federal appropriations for RWHAP over the past decade.

To date, no formal evaluation has examined how these factors and characteristics influence ADAP implementation or their capacity to improve viral suppression rates among people with HIV. This represents a key gap in the literature and a primary motivation for this study. The inner workings of ADAPs within states—how context, operational architecture, economics, and organizational practices translate into outcomes—remain to be described. The most recent qualitative study of ADAPs was performed in 2011, focusing on how ADAPs would adapt to the ACA [22]. Since the ACA's implementation in 2014, both HIV care delivery and ADAP operations have evolved [13,23]. Additionally, the population served by ADAPs has changed, with new diagnoses increasingly occurring among people who are Black or African American and/or Hispanic or Latino. There are also increasing comorbidities among older people with HIV. Understanding how ADAPs adapt to these changing demographic and clinical needs is essential to supporting their mission of improving viral suppression and advancing toward the 90% goal of the Ending the HIV Epidemic initiative [2].

The objective of this research is to understand how operational, organizational, economic, and contextual factors shape ADAPs' ability to improve viral suppression rates among low-income people with HIV. The theoretical framework guiding this qualitative research study is the Consolidated Framework for Implementation Research (CFIR) [24]. The CFIR is a flexible and useful framework for understanding policy implementation and the impact of innovation across a range of operational settings. Health care researchers have used the CFIR to understand organizational practices and policies, operational norms, and the challenges and contexts accompanying the implementation of innovation and the introduction of new systems. The value of the CFIR lies in its reach, enabling an understanding of how a variety of factors, including participants' needs and characteristics, organizational culture and operational routines, relevant governmental or regulatory guidelines, and the character and quality of organizational leadership, shape or determine outcomes. The CFIR allows researchers to understand how a variety of context-specific variables impact how well a system or set of practices serves the clients, patients, and participants it is designed to serve. We seek to better understand how ADAP structures and processes integrate within heterogeneous state contexts.

One feature of the CFIR as a conceptual instrument is the continued engagement of the instrument's designers in ongoing efforts to refine and supplement the CFIR as a research and evaluative tool [24]. The codevelopers behind the CFIR have assembled data that provide a glimpse into how health care and public health researchers have used the CFIR to guide data analysis and to evaluate and design or redesign implementation strategies [24]. This results in a broad community of practice offering crowdsourced solutions to research, application, and evaluation questions. Figure 1 shows our CFIR model of the ADAP operational landscape.

Figure 1. Application of the Consolidated Framework for Implementation Research to inform the design and analysis of qualitative interviews.

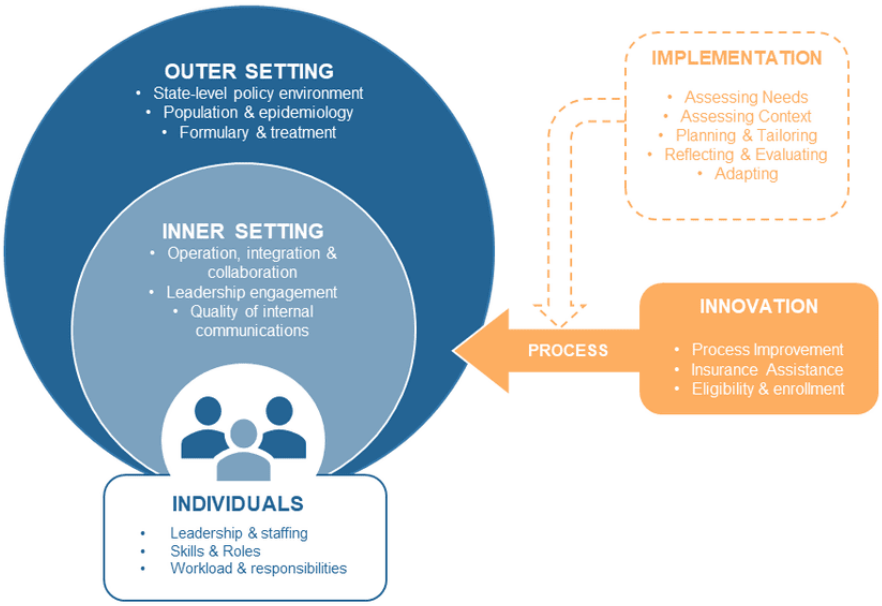

XSL•FO
RenderX

Methods

Overview

This study takes a qualitative descriptive approach [25,26] to identify ADAPs' programs and policies that could help ADAP clients achieve sustained viral suppression for individual and public health benefits. Enrollment was at the state level.

Data Collection

Data collection involved individual semistructured interviews and was conducted via Health Insurance Portability and Accountability Act (HIPAA)–compliant videoconferencing to maintain the confidentiality of responses. Interviews were recorded and transcribed verbatim. Participants also responded to a brief survey focused on the characteristics of ADAP structures.

While the interview questions focused on ADAP structures and processes, some respondents may have offered insight into other aspects of the RWHAP and the general health infrastructure within the state, and we will incorporate these findings. Each interview consisted of open-ended questions organized into five parts covering five domains: (1) Opening questions focused on individuals and their roles and experiences in the operation and oversight of their state ADAP; this included questions about ADAP successes and challenges. (2) Operation, integration, and collaboration were discussed, which included understanding connections to other state agencies and health department personnel. We also assessed input from community stakeholders; connections to other social service providers; and interconnectedness with ADAPs in other states, the National Alliance for State and Territorial AIDS Directors (NASTAD), and other external partners, as well as how these are operationalized. (3) ADAP medication procurement and delivery policies were explored, including insurance assistance programs and processes shaping ADAP policy decisions, including questions about program eligibility, formulary design, and pharmacy network design. (4) Questions about eligibility criteria and enrollment policy, practices, and operational machinery included questions about state ADAP internal assessments of viral suppression rates and goals; differences in viral suppression among subpopulations who are disproportionately affected and/or historically underserved; and suggestions for ideal resources, policies, and programs that they think would improve viral suppression and reduce differences in outcomes for ADAP clients. (5) Concluding questions focused on innovation and implementation, which included program assessment, stakeholders involved in evaluating outcomes, data and data literacy challenges, and perceptions of the value of ADAP, as well as needed state or federal support, resources, or regulatory changes to implement best practices and improve health outcomes.

The interviews concluded with open-ended questions that invited participants to share additional thoughts. They will also be asked to identify any key stakeholders whom we should invite to participate in the study.

The study used maximum variance sampling to obtain qualitative interviews across a broad range of state ADAPs to maximize contextual insight. States will be selected based on the following key variables: (1) viral suppression rates [27] and (2) contextual variables, including region or geographic diversity [28], HIV prevalence [29], number of ADAP clients [12], state financial contributions to ADAP [12], and state Medicaid expansion status [30], as has been done by us [31] and others in previous work [22,32]. This sampling strategy will enable us to gain insight into a wide range of state ADAPs. These variables are critical as they represent external policy and resource factors (outer setting in CFIR) that constrain or enable ADAP operations.

From the 50 US states and the District of Columbia (n=51), state health department HIV leaders, including AIDS directors and ADAP directors and staff, were approached to participate in this study. The list of eligible participants (with their email addresses) is publicly available and was verified by NASTAD [33]. We used modified snowball sampling to recruit additional participants based on enrolled participants' suggestions for key stakeholders in HIV care whom we should interview (eg, politically appointed health department personnel, community opinion leaders, and implementation team members), aligning with CFIR roles [27]. When more than 1 participant was interested in a state, we offered a joint interview or individual interviews.

The interview questions were drafted to uncover rich information about ADAP operations, policies, and practices, including eligibility criteria and re-enrollment requirements, organizational architecture and leadership, and participant and client profiles. For example, we anticipate learning about individual roles within ADAPs, staffing, and staff workloads; patterns of integration and cooperation with external partners and the wider community of actors conceptualizing and guiding ADAPs' policies and practices; how state agencies and political and operational leadership cooperate and support ADAPs' work; variations in program eligibility criteria and how criteria are adapted to respond to shifts in the level and character of need; and the ways in which innovations and new practices are implemented and adopted. Additionally, resources, including ADAP budget and revenue generation opportunities, are a key part of what drives operational and programmatic decisions, and these topics are covered as well. Table 1 shows the domains of qualitative interview questions and their alignment with the CFIR. The full interview guide is available in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Domains of qualitative interview questions and alignment with the Consolidated Framework for Implementation Research (CFIR).

Domains	CFIR settings	Questions asked in the survey and key analytical concerns
Leadership and staffing	Individuals	• How many people work full time in your ADAP[a]? • How many work part time? • Describe the workload and responsibilities. • Is work outsourced or assigned to external contractors? • Do you have an ADAP advisory committee? • What is the advisory committee's composition? Does it meet regularly?
Operation, integration, and collaboration	Inner setting	• What is the relationship with the state health department? • Is it part of the same agency responsible for Ryan White HIV/AIDS Program Part B? • Is it part of the same agency responsible for HIV prevention and surveillance? • What is the role of politically appointed health department personnel? • What medical and nonmedical casework is completed and what is the relationship between the two? • What input do you receive from community stakeholders? • What connections does ADAP have with other social service agencies or providers? • What connections exist between ADAP and programs in other states, the National Alliance of State and Territorial AIDS Directors, and other external partners?
Operations	Outer setting	• What ideal resources, policies, and programs would improve viral suppression for ADAP clients?
Formulary design and treatment	Outer setting	• What is the formulary design for antiretroviral therapy? • What is the formulary design for non–antiretroviral therapy medications? • How are these decisions made? • What is the allowed duration of medication supply? • What is the pharmacy network design? • Do you offer or require mailing of medications? • Does the full-pay formulary design differ from the insurance assistance formulary design? • How does ADAP obtain client-level data from clinics and laboratories? • Do you have a process for clients to self-report laboratory results? • How are data harmonized across various laboratories and clinics?
Eligibility and enrollment	Outer setting and innovation	• What is the income threshold and the mechanisms used for determining it? • What is the process by which eligibility is determined? • How are clients notified about or screened for re-enrollment? • How do clients lose ADAP coverage? • How can they regain eligibility? What is this process? • How can individuals regain eligibility, and what does this process involve? • What is the process for enrolling in full-pay ADAP or ADAP-assisted insurance coverage?
Innovation and implementation	Innovation	• Does your ADAP perform internal assessments of disparities in viral suppression related to race or other factors? • Does your ADAP perform internal assessments of engagement in the ADAP? • What are the data challenges, data system challenges, and data literacy issues for your ADAP? • Who is involved in policy and implementation decision-making? • Are there any recent or planned changes related to eligibility? How are these changes made?

[a]ADAP: AIDS Drug Assistance Program.

Outcomes and Data Analysis

Analyses will include descriptive statistics, conventional qualitative content analysis [34], and situational mapping [31], in which themes are inductively derived from data [34]. Data analysis will combine inductive and deductive methods of qualitative content analysis, with deductive components guided by CFIR [34]. Specifically, we will develop an initial coding dictionary using relevant CFIR domains (eg, inner setting, outer setting, innovation, and characteristics of individuals). Segments of interview data will be coded to these domains to systematically identify implementation determinants, such as leadership engagement, resource availability, interagency coordination, data infrastructure, and policy constraints. This approach ensures that analysis remains theoretically anchored and allows examination of how multilevel contextual factors influence ADAP operations.

The remaining analyses will be divided between research team members, who have in-depth experience with both forms of qualitative content analysis used in the study. Analyses will be

XSL•FO
RenderX

discussed at full team meetings. These will be conducted using Dedoose (version 10.0.59; SocioCultural Research Consultants LLC). Situational mapping will be used to visualize and analyze the complex interplay between ADAP (the "what" or innovation), organizational context (inner setting), external environment (outer setting), and resultant client outcomes (viral suppression) across different states, as guided by CFIR.

Research team members will maintain reflective journal entries throughout the research process. Methods to ensure rigor and trustworthiness [35] of the data analysis will include aspects of demonstrating (1) credibility (direct observation of online communication, iterative questioning, and frequent debriefing), (2) transferability (contextual review), (3) dependability (maintaining an audit trail), and (4) confirmability (bracketing and investigator triangulation).

We will use the Consolidated Criteria for Reporting Qualitative Research reporting criteria when we publish the primary analysis [36].

Ethical Considerations

This study has received ethics approval from the University of Virginia Institutional Review Board for the Social and Behavioral Sciences (Protocol #5707; modification approved on December 1, 2025). NASTAD and JSI Research and Training Institute Inc are collaborating partners supporting implementation, analysis, and dissemination activities. University of Virginia Institutional Review Board–Social and Behavioral Sciences serves as the institutional review board of record. JSI obtained a supplemental exempt determination (institutional review board #24-56E) for the participation of its staff. All participants were adults. Prior to participation, the study purpose, procedures, confidentiality protections, and potential risks and benefits were explained to each prospective participant. Verbal informed consent was obtained, and participants were informed that their participation is voluntary and that they could decline to answer questions or withdraw from the study at any point without consequences. No financial compensation was provided for participation. In the dissemination process, we will not identify individual people or states.

Results

This study was funded in April 2023. A small adjustment to the study aims was made in 2025, as requested by the National Institutes of Health. The original and revised aims can be accessed in the National Institutes of Health RePORTER database [37,38]. As of March 2026, 16 states have participated, including 33 individual participants in our qualitative interviews, often with more than 1 interview per state. The participation rate was 31.4% (16/51). Characteristics of participating interviewees (N=33) and states (N=16) are included in Table 2. More than half of the participants were ADAP-related administrators (16/33, 48.5%), followed by HIV-related state health department administrators (15/33, 45.5%) and other state health department administrators (2/33, 6.1%). Participating states were geographically diverse, with the largest share in the West (6/16, 37.5%), equal representation from the South and Midwest (4/16, 25% each), and the fewest from the Northeast (2/16, 12.5%). Regarding financial eligibility requirements, half of the participating ADAPs set their federal poverty level threshold at ≥500% (8/16, 50%), while the remaining programs were evenly split between 400% and 500% (4/16, 25%) and 200% to 400% (4/16, 25%). Interviews were conducted between February 12, 2025, and March 20, 2025. As of March 2026, analyses are currently ongoing and expected results to be published in 2026.

XSL•FO
RenderX

Table 2. Sample characteristics of study participants (individual and states).

Characteristics	Participants, n (%)
Role of individual participants (N=33)[a]	
HIV-related state health department administrator	15 (45.5)
ADAP[b]-related administrator	16 (48.5)
Other state health department administrator	2 (6.1)
Geographic region of participating state (N=16)	
West	6 (37.5)
South	4 (25)
Midwest	4 (25)
Northeast	2 (12.5)
ADAP financial eligibility threshold (FPL[c]), 2025, of participating State (N=16)[d]	
>500% of FPL	8 (50)
400%-500% of FPL	4 (25)
200%-400% of FPL	4 (25)

[a]Summarizes the roles of participants who completed interviews. Across categories, roles spanned varying levels of responsibility, including program coordinators, managers, and directors.

[b]ADAP: AIDS Drug Assistance Program.

[c]FPL: federal poverty line.

[d]For all states included, ADAP's financial eligibility requirements as of January 1, 2025, were concordant across both full-pay medication programs and ADAP-funded insurance programs.

Discussion

Context for Anticipated Findings

The overall viral suppression rate in the United States is 67% [1], which contrasts with the higher rate of 85% among people with HIV who receive services from ADAP [13]. This difference highlights an important opportunity to leverage ADAPs to improve HIV outcomes for low-income people with HIV [39]. Additionally, as ADAPs are helping people with HIV achieve higher-than-average outcomes, part of the reasoning behind focusing on ADAPs is that getting more people with HIV into ADAP could help improve viral suppression rates in the United States. It could be a key component of ending the HIV epidemic in the United States.

Our study focuses on this opportunity by studying ADAPs—a critical safety-net system that provides HIV medications for low-income people with HIV, either through direct provision of medication or assistance with premiums and medication cost sharing. Our approach is both theoretically and conceptually innovative. Specifically, we situate our analysis within the broader field of implementation science using the CFIR to guide our interview guide and subsequent analysis. This framework enables us to identify the key factors and contextual conditions that facilitate or hinder the adoption of innovations designed to improve viral suppression among people with HIV served by ADAPs.

A variety of mechanisms, organizational structures, operational practices, eligibility criteria, and related factors contribute to the success of particular ADAPs in elevating their clients' viral suppression rates, providing continuity of care, and accomplishing broader public health aims that guide their operations. This study, given its reach across a broad range of state ADAPs representing the variety of policies and practices permitted under the RWHAP, can help identify policies, practices, approaches, and organizational structures that promote better outcomes for individual ADAP clients and public health objectives of wider health care and civic effort.

Our qualitative methodology, when used within a broader understanding of the health care needs and objectives related to people with HIV, is designed to assemble rich operational-level details about the mechanisms, policy structures, eligibility criteria, and decision-making processes that shape how effectively and responsively ADAPs serve an evolving client base in a changing policy and health care environment (Multimedia Appendix 2). The dynamic nature of the landscape in which viral suppression efforts unfold requires a better understanding of how decisions are implemented, innovations are adopted, and health care professionals and organizational and political leaders collaborate to implement best practices and ultimately improve health outcomes.

Much of HIV research overlooks the real-world barriers to achieving viral suppression among subpopulations who are disproportionately affected, historically underserved groups, and low-income populations. Despite the availability of safe and effective ART, more than one-third of people with HIV in the United States remain with detectable viral loads [1]. More work is needed to ensure that the tools available to safely and effectively treat HIV are reaching those who need them [40]. Our team's recent quantitative work demonstrates that while state ADAPs serve less than a quarter of people with HIV,

XSL•FO
RenderX

almost a third of the viral suppression in the United States can be attributed to people with HIV accessing medications through ADAPs [41]. ADAPs are essential for ending the HIV epidemic in the United States, and recent experts have declared ADAPs a model for chronic disease treatment support [42].

We anticipate that ADAPs with high rates of viral suppression will share information about programs and policies that could be disseminated to improve viral suppression and reduce differences in outcomes for subpopulations who are disproportionately affected and/or historically underserved. Because we will also have data on aspects of the states (such as Medicaid expansion status) as well as viral suppression rates, we can provide context for the environments in which these programs or policies affect ADAP clients. The Health Resources and Services Administration and state ADAPs could use these results to inform future regulatory and funding priorities. Following completion of the interviews, there were additional cost-containment measures imposed [43,44]. Therefore, these interviews will serve as a critical baseline assessment prior to the implementation and evaluation of the impacts of these cost-containment measures. In close collaboration with NASTAD, findings will be disseminated to traditional academic audiences and translated into actionable policy recommendations. Dissemination to stakeholders will include the development of brief, accessible policy reports and presentations at key national- and state-level RWHAP-ADAP convenings to facilitate shared best practices and immediate translation into operational changes.

Despite the study strengths, several challenges remain. Perspectives are drawn primarily from state health department HIV leaders, ADAP leaders and staff, and key stakeholders, which may underrepresent frontline staff and client experiences and may be influenced by institutional framing or social desirability. Although we use maximum variance sampling, participation is voluntary, and some states may differ systematically from those that did not participate. Viral suppression is shaped by many factors beyond ADAP design, including the broader HIV care infrastructure and health system navigation, Medicaid policies, and social and structural determinants. Findings also reflect a dynamic policy period (eg, the end of the COVID-19 public health emergency and the related Medicaid continuous coverage unwinding, as well as delayed funding and defunding of state public health systems), which may affect generalizability over time. Finally, we will mitigate concerns related to generalizability using team-based coding, audit trails, reflexivity, and triangulation.

Dissemination Plan

We will disseminate the results of this study through traditional academic channels, including peer-reviewed manuscripts and presentations at national HIV conferences. In addition, study findings will inform the development of policy reports tailored to federal and state partners. In collaboration with NASTAD, we will share results directly with ADAP leadership and program staff to support translation of evidence into practice. We will align dissemination products with the Translational Science Benefit Model for optimal translation to impact policy and health outcomes [45].

Conclusions

This qualitative study will address a critical gap by characterizing the operational, organizational, and contextual factors that enable ADAPs to achieve sustained viral suppression. By identifying key mechanisms that drive success across diverse state contexts, this work is well suited to inform policy and program development aimed at improving the overall health of people living with HIV. Findings will contribute essential evidence to guide ADAPs, policymakers, and public health leaders in strengthening the nation's HIV care safety net.

Acknowledgments

The authors thank the team members from 2 state AIDS Drug Assistance Programs (ADAPs) who generously met with us to advise on the study design. The authors also thank state health department leaders and ADAP leaders and teams for their time and effort in participating in the study.

Funding

This work is supported by the National Institute of Allergy and Infectious Diseases at the National Institutes of Health (grant R01AI170093 to KAM). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to the sensitive nature of the interview content and the terms of participant consent but deidentified quotations are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: AK, AS, EQR, JK-M, KAM, RP, SL, TH
Data curation: AS, EQR, JK-M, KAM
Funding acquisition: AK, AS, JK-M, KAM, TH
Methodology: AS, EQR, JK-M, KAM

XSL•FO
RenderX

Project administration: AS, EQR, JK-M, KAM
Supervision: AS, JK-M, KAM
Visualization: AS, EQR, JK-M, KAM, RP, SL
Writing—original draft: AK, AS, EQR, JK-M, KAM, RP, SL, TH
Writing—review and editing: AK, AS, EQR, JK-M, KAM, RP, SL, TH

Conflicts of Interest

AS owns stocks from Merck and Co. JK-M owns equity in ArteraAI, a cancer biomarker company that is unrelated to this current work.

Multimedia Appendix 1
Interview guide.
[DOCX File , 29 KB - resprot_v15i1e90008_app1.docx]

Multimedia Appendix 2
Peer review report from PPAH - Population and Public Health Approaches to HIV/AIDS Study Section, Healthcare Delivery and Methodologies Integrated Review Group (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 147 KB - resprot_v15i1e90008_app2.pdf]

References

1. Data table: monitoring national HIV prevention goals. Center for Disease Control and Prevention. 2025 Feb 07. URL: https://www.cdc.gov/hiv-data/initiatives/prevention-goals-table.html [accessed 2026-03-23]
2. Fauci AS, Redfield RR, Sigounas G, Weahkee MD, Giroir BP. Ending the HIV epidemic: a plan for the United States. JAMA 2019 Mar 05;321(9):844-845. [doi: 10.1001/jama.2019.1343] [Medline: 30730529]
3. INSIGHT START Study Group, Lundgren JD, Babiker AG, Gordin F, Emery S, Grund B, et al. Initiation of antiretroviral therapy in early asymptomatic HIV infection. N Engl J Med 2015 Aug 27;373(9):795-807 [FREE Full text] [doi: 10.1056/NEJMoa1506816] [Medline: 26192873]
4. Marcus JL, Leyden WA, Alexeeff SE, Anderson AN, Hechter RC, Hu H, et al. Comparison of overall and comorbidity-free life expectancy between insured adults with and without HIV infection, 2000-2016. JAMA Netw Open 2020 Jun 01;3(6):e207954 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.7954] [Medline: 32539152]
5. Rodger AJ, Cambiano V, Bruun T, Vernazza P, Collins S, van Lunzen J, et al. Sexual activity without condoms and risk of HIV transmission in serodifferent couples when the HIV-positive partner is using suppressive antiretroviral therapy. JAMA 2016 Jul 12;316(2):171-181 [FREE Full text] [doi: 10.1001/jama.2016.5148] [Medline: 27404185]
6. Rodger AJ, Cambiano V, Bruun T, Vernazza P, Collins S, Degen O, et al. Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive partner taking suppressive antiretroviral therapy (PARTNER): final results of a multicentre, prospective, observational study. Lancet 2019 Jun 15;393(10189):2428-2438. [doi: 10.1016/S0140-6736(19)30418-0] [Medline: 31056293]
7. Farnham PG, Gopalappa C, Sansom SL, Hutchinson AB, Brooks JT, Weidle PJ, et al. Updates of lifetime costs of care and quality-of-life estimates for HIV-infected persons in the United States: late versus early diagnosis and entry into care. J Acquir Immune Defic Syndr 2013 Oct 01;64(2):183-189. [doi: 10.1097/QAI.0b013e3182973966] [Medline: 23615000]
8. Cohen JP, Beaubrun A, Ding Y, Wade RL, Hines DM. Estimation of the incremental cumulative cost of HIV compared with a non-HIV population. Pharmacoecon Open 2020 Dec;4(4):687-696 [FREE Full text] [doi: 10.1007/s41669-020-00209-8] [Medline: 32219732]
9. Cohen JP, Anupindi VR, Doshi R, Yeaw J, Zhou X, Christoph MJ, et al. Estimation of lifetime costs among insured persons with HIV in the United States. Pharmacoecon Open 2025 Nov;9(6):869-886 [FREE Full text] [doi: 10.1007/s41669-025-00584-0] [Medline: 40490607]
10. Part B: AIDS Drug Assistance Program (ADAP). Health Resources & Services Administration. URL: https://hab.hrsa.gov/about-ryan-white-hivaids-program/part-b-aids-drug-assistance-program [accessed 2025-03-22]
11. AIDS Drug Assistance Programs (ADAPs). Kaiser Family Foundation. URL: https://files.kff.org/attachment/Fact-Sheet-AIDS-Drug-Assistance-Programs? [accessed 2025-06-25]
12. 2021-2022 National RWHAP Part B and ADAP monitoring project report. NASTAD. URL: https://nastad.org/partb-adap-2021-2022-report [accessed 2025-03-25]
13. 2025 Annual report: National Ryan White HIV/AIDS Program Part B AIDS Drug Assistance Program monitoring project. NASTAD. URL: https://nastad.org/2025-rwhap-part-b-adap-monitoring-report [accessed 2025-04-21]
14. Eisinger RW, Fauci AS. Ending the HIV/AIDS pandemic. Emerg Infect Dis 2018 Mar;24(3):413-416 [FREE Full text] [doi: 10.3201/eid2403.171797] [Medline: 29460740]
15. AIDS Drug Assistance Program (ADAP) manual. Ryan White HIV/AIDS Program. URL: https://ryanwhite.hrsa.gov/sites/default/files/ryanwhite/resources/dshap-adap-manual-june-2025-final.pdf [accessed 2026-03-23]

XSL•FO
RenderX

16. National ADAP formulary database. NASTAD. URL: https://nastad.org/adap-formulary-database [accessed 2025-04-22]
17. McManus KA, Killelea A, Horn T, Steen A, Sun B, Keim-Malpass J, et al. Trends in AIDS drug assistance program support for people with HIV, by age, sex, race, and US Region, 2008 2021. Am J Public Health 2025 Aug;115(8):1288-1298. [doi: 10.2105/AJPH.2025.308101] [Medline: 40472297]
18. Dawson L, Kates J, Roberts T. The health insurance and financing landscape for people with and at risk for HIV. Kaiser Family Foundation. 2023 May 25. URL: https://www.kff.org/hiv-aids/the-health-insurance-and-financing-landscape-for-people-with-and-at-risk-for-hiv/ [accessed 2026-03-23]
19. McManus KA, Engelhard CL, Dillingham R. Current challenges to the United States' AIDS Drug Assistance Program and possible implications of the Affordable Care Act. AIDS Res Treat 2013;2013:350169. [doi: 10.1155/2013/350169] [Medline: 23573418]
20. Federal policy changes and their impact on AIDS Drug Assistance Programs. NASTAD. URL: https://nastad.org/sites/default/files/2025-10/nastad_adap_issuebrief_103025.pdf [accessed 2026-03-23]
21. The Ryan White HIV/AIDS Program: the basics. Kaiser Family Foundation. 2025 May 29. URL: https://www.kff.org/hivaids/fact-sheet/the-ryan-white-hivaids-program-the-basics/ [accessed 2025-08-19]
22. Martin EG, Meehan T, Schackman BR. AIDS Drug Assistance Programs: managers confront uncertainty and need to adapt as the Affordable Care Act kicks in. Health Aff (Millwood) 2013 Jun;32(6):1063-1071 [FREE Full text] [doi: 10.1377/hlthaff.2012.0123] [Medline: 23733980]
23. McManus KA, Rhodes A, Bailey S, Yerkes L, Engelhard CL, Ingersoll KS, et al. Affordable Care Act qualified health plan coverage: association with improved HIV viral suppression for AIDS Drug Assistance Program clients in a Medicaid nonexpansion state. Clin Infect Dis 2016 Aug 01;63(3):396-403 [FREE Full text] [doi: 10.1093/cid/ciw277] [Medline: 27143661]
24. Damschroder LJ, Reardon CM, Widerquist MA, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75 [FREE Full text] [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
25. Sandelowski M. What's in a name? Qualitative description revisited. Res Nurs Health 2010 Feb;33(1):77-84. [doi: 10.1002/nur.20362] [Medline: 20014004]
26. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
27. Patton MQ. Qualitative Research & Evaluation Methods: Integrating Theory and Practice. Thousand Oaks, CA: SAGE Publications; 2014.
28. Geographic areas reference manual. United States Census Bureau. 1994. URL: https://www2.census.gov/geo/pdfs/reference/GARM/GARMcont.pdf [accessed 2026-03-23]
29. AtlasPlus Glossary. Centers for Disease Control and Prevention. URL: https://gis.cdc.gov/grasp/nchhstpatlas/Content/docs/AtlasPlusGlossary.pdf [accessed 2025-01-24]
30. Status of state action on the medicaid expansion decision. Kaiser Family Foundation. 2025 Sep. URL: http://kff.org/health-reform/state-indicator/state-activity-around-expanding-medicaid-under-the-affordable-care-act/ [accessed 2025-03-20]
31. McManus KA, Killelea A, Honeycutt E, An Z, Keim-Malpass J. Assisters succeed in insurance navigation for people living with HIV and people at increased risk of HIV in a complex coverage landscape. AIDS Res Hum Retroviruses 2020 Oct;36(10):842-851 [FREE Full text] [doi: 10.1089/AID.2020.0013] [Medline: 32631076]
32. Kapadia SN, Johnston CD, Marks KM, Schackman BR, Martin EG. Strategies for improving hepatitis C treatment access in the United States: state officials address high drug prices, stigma, and building treatment capacity. J Public Health Manag Pract 2019;25(3):245-252 [FREE Full text] [doi: 10.1097/PHH.0000000000000829] [Medline: 29927900]
33. Member directory. NASTAD. URL: https://nastad.org/member-directory [accessed 2025-04-04]
34. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
35. Lincoln YS, Guba EG. Naturalistic Inquiry. Thousand Oaks, CA: SAGE Publications; 1985.
36. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
37. Viral suppression for people with HIV with low incomes: study of disparities, health equity, and best practices. NIH RePORT. URL: https://reporter.nih.gov/search/2ATeK54dE0eAUju0DDSL8Q/project-details/10619064 [accessed 2026-03-12]
38. Best practices to improve viral suppression for people with HIV. NIH RePORT. URL: https://reporter.nih.gov/search/2ATeK54dE0eAUju0DDSL8Q/project-details/11101276 [accessed 2026-03-12]
39. Ghosn J, Taiwo B, Seedat S, Autran B, Katlama C. HIV. Lancet 2018 Aug 25;392(10148):685-697. [doi: 10.1016/S0140-6736(18)31311-4] [Medline: 30049419]
40. Fauci AS, Lane HC. Four decades of HIV/AIDS - much accomplished, much to do. N Engl J Med 2020 Jul 02;383(1):1-4. [doi: 10.1056/NEJMp1916753] [Medline: 32609976]

XSL•FO
RenderX

41. McManus KA, Killelea A, Rogers EQ, Liu F, Horn T, Steen A, et al. State AIDS Drug Assistance Programs' contribution to the United States' viral suppression, 2015-2022. medRxiv Preprint posted online on April 6, 2025 [FREE Full text] [doi: 10.1101/2025.04.04.25325288] [Medline: 40236406]
42. Dombrowski JC, Thompson A, Feinberg J. The AIDS Drug Assistance Program is a model for chronic disease treatment support. Am J Public Health 2025 Aug;115(8):1187-1189. [doi: 10.2105/AJPH.2025.308189] [Medline: 40633000]
43. AIDS Drug Assistance Program (ADAP) cost containment considerations. NASTAD. URL: https://nastad.org/resources/aids-drug-assistance-program-adap-cost-containment-considerations [accessed 2026-02-11]
44. NASTAD ADAP Watch - February, 2026. URL: https://nastad.org/resources/nastad-adap-watch-february-2026 [accessed 2026-03-23]
45. Luke DA, Sarli CC, Suiter AM, Carothers BJ, Combs TB, Allen JL, et al. The Translational Science Benefits Model: a new framework for assessing the health and societal benefits of clinical and translational sciences. Clin Transl Sci 2018 Jan;11(1):77-84 [FREE Full text] [doi: 10.1111/cts.12495] [Medline: 28887873]

Abbreviations

ACA: Affordable Care Act
ADAP: AIDS Drug Assistance Program
ART: antiretroviral therapy
CFIR: Consolidated Framework for Implementation Research
HIPAA: Health Insurance Portability and Accountability Act
NASTAD: National Alliance for State and Territorial AIDS Directors
RWHAP: Ryan White HIV/AIDS Program

Edited by A Schwartz; submitted 23.Dec.2025; this is a non–peer-reviewed article;accepted 17.Mar.2026; published 16.Apr.2026.

Please cite as:
McManus KA, Rogers EQ, Steen A, Killelea A, Horn T, Panagides R, Laymon S, Keim-Malpass J
Characterizing AIDS Drug Assistance Program Practices and Policies for Sustained Viral Suppression Using the Consolidated Framework for Implementation Research: Protocol for a Qualitative Study
JMIR Res Protoc 2026;15:e90008
URL: https://www.researchprotocols.org/2026/1/e90008
doi:10.2196/90008
PMID:41858071

©Kathleen A McManus, Erin Q Rogers, Amber Steen, Amy Killelea, Tim Horn, Reanna Panagides, Steven Laymon, Jessica Keim-Malpass. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Combining Digital Cognitive Behavioral Therapy With Mindfulness Training for Binge Eating Disorder: Protocol for a Feasibility Trial

Margaret Sala[1], PhD; Corey Roos[2], PhD; Dante Ascarrunz[1,3], BA; Casey M Stern[1], BA; Jonathan Bricker[4,5], PhD; Jonathan M Feldman[1,3], PhD; Cheri Levinson[6], PhD; Mihaela Aslan[7,8], PhD; Jennifer L Hay[9], PhD; Hedy Kober[10], PhD

[1]Ferkauf Graduate School of Psychology, Yeshiva University, Bronx, NY, United States
[2]Department of Psychiatry, School of Medicine, Yale University, New Haven, CT, United States
[3]Department of Pediatrics, Albert Einstein College of Medicine, Bronx, NY, United States
[4]Division of Public Health Sciences, Fred Hutch Cancer Center, Seattle, WA, United States
[5]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States
[6]Department of Psychological and Brain Sciences, University of Louisville, Louisville, KY, United States
[7]Department of Internal Medicine, School of Medicine, Yale University, New Haven, CT, United States
[8]VA Connecticut Healthcare System, West Haven, CT, United States
[9]Department of Psychiatry and Behavioral Sciences, Memorial Sloan Kettering Cancer Center, New York, NY, United States
[10]Department of Psychology, University of California, Berkeley, Berkeley, CA, United States

Corresponding Author:
Margaret Sala, PhD
Ferkauf Graduate School of Psychology
Yeshiva University
Louis E. and Dora Rousso Building
1165 Morris Park Ave
Bronx, NY, 10461
United States
Phone: 1 9723694267
Email: margarita.sala@yu.edu

Abstract

Background: Cognitive behavioral therapy (CBT) is the most effective treatment for binge eating disorder (BED) but is limited by modest efficacy and low reach of in-person delivery. Combining mindfulness training with CBT could enhance efficacy by targeting emotion dysregulation, a key factor in BED that CBT does not adequately target. Digital delivery can enhance the reach of treatment. Accordingly, we developed *CBT-based Mindful Courage*, a 16-session digital intervention combining CBT and mindfulness training for BED.

Objective: The aim of this study is to conduct a pilot randomized controlled trial to evaluate the feasibility (intervention and assessment completion) and acceptability (usability, overall satisfaction, engagement, visual appeal of content, understandability of program content, desire to continue the program, perceived skill acquisition, perceived confidence in implementing skills, and perceived helpfulness) of *CBT-based Mindful Courage*. We will also evaluate changes in outcomes (binge eating frequency, eating disorder symptoms, clinical impairment, and depression symptoms) and mechanisms (emotion dysregulation, dietary overrestriction, and trait mindfulness).

Methods: This is a parallel-group, single-blinded, randomized controlled trial with a 1:1 allocation ratio. It will be conducted from November 2025 to November 2026, with study assessments at baseline, midtreatment, end-of-treatment, and 2-month follow-up. We will recruit a volunteer sample of adults with BED (N=40) aged 18 to 75 years via online advertisements. Exclusion criteria include a BMI <18.5, current anorexia or bulimia nervosa, and current treatment for BED or weight loss. Participants will be randomly allocated to an active control that includes coached self-monitoring (self-monitoring of food intake in *Recovery Record* and coaching calls by a research assistant) or to *CBT-based Mindful Courage*+coached self-monitoring. Both conditions will have a duration of 18 weeks. All study sessions will take place over Zoom.

Results: The study was funded in April 2023. Data collection began in November 2025, and we anticipate recruitment to be completed by November 2026. Results are expected to be analyzed by May 2027 and published by May 2028.

XSL•FO
RenderX

Conclusions: This will be an important first step in creating a scalable and efficacious treatment for BED.

Trial Registration: ClinicalTrials.gov NCT07212673; https://clinicaltrials.gov/study/NCT07212673

International Registered Report Identifier (IRRID): DERR1-10.2196/91761

(JMIR Res Protoc 2026;15:e91761) doi:10.2196/91761

KEYWORDS

cognitive behavioral therapy; mindfulness-based therapy; mindfulness-based intervention; eating disorder; binge eating disorder; digital

Introduction

Binge eating disorder (BED) is the most prevalent eating disorder (ED), affecting 17.3 million individuals worldwide [1]. Individuals with BED are at increased risk of several psychosocial and medical consequences [2-4] and elevated mortality [5]. Furthermore, BED is linked with high comorbidity [6] and reduced quality of life [7].

Conceptualizations of BED and other EDs have consistently focused on the notion that overrestriction (ie, excessive rigid restrictions in food intake) is a primary contributor to BED [8,9]. Consequently, cognitive behavioral therapy (CBT), the gold standard treatment for BED [10], has several behavioral components that directly target overrestriction [9].

However, CBT for BED has modest efficacy. Notably, approximately 50% of individuals who receive CBT for BED continue to report difficulties with binge eating after treatment [11]. The modest efficacy of CBT may be due to an inadequate focus on emotion dysregulation. Emotion dysregulation has been identified as a major contributing factor in the onset and maintenance of BED [12-14] but receives minimal attention in CBT treatments for BED. CBT interventions lead to only slight improvements in emotion dysregulation [8,15].

Mindfulness-based interventions (MBIs) are particularly effective for ameliorating emotion dysregulation [16,17] because they improve the ability to process and respond effectively to emotions. MBIs directly teach individuals to respond to emotions with awareness rather than automatically engaging in maladaptive emotion regulation strategies [18]. As such, they may help individuals with BED to accept emotions rather than react to them habitually (eg, by binge eating). Therefore, mindfulness training, the core component of MBIs, could be a promising addition to CBT for BED.

A few studies to date have investigated the use of MBIs in BED and produced preliminary evidence supporting the use of MBIs in the treatment of BED [19,20]. In addition, a recent review suggests that several MBIs for EDs result in significant reductions in emotion dysregulation [21]. Furthermore, our preliminary work indicates that MBIs result in large and significant decreases in emotion dysregulation in other EDs [22,23].

Integrating CBT and mindfulness training could lead to a potent intervention that targets two core mechanisms in BED: (1) excessive dietary restriction and (2) emotion dysregulation. Furthermore, mindfulness training may be synergistic with CBT. For example, mindfulness training may increase adherence to CBT skills by increasing the ability to tolerate emotional discomfort related to deploying CBT skills [24]. Similarly, CBT may enhance the ability to engage with the present moment and increase awareness by reducing focus on eating, weight, and shape.

Despite the promise of CBT and MBIs in the treatment of BED, a recent systematic review showed that only 20% of individuals in the community with a diagnosable ED can access ED-specific treatment [25] due to barriers such as a lack of experienced therapists, high cost of treatment, lack of health insurance, and shame and stigma associated with EDs and mental health treatment [26]. One way to overcome these barriers to treatment is to offer a digital treatment, that is, a multimedia application that can be delivered via smartphone or web. Digital MBIs and CBTs were already shown to be efficacious for several other psychiatric disorders [27-30]. While digital delivery of CBT has been shown to be efficacious for EDs [31], there are no digital MBIs to date for EDs outside of the ones developed in our preliminary work [22,23,32]. A recent meta-analysis of digital interventions for EDs included 8 treatments (4 for BED), and all interventions were based on CBT, with no interventions including mindfulness training [31]. Furthermore, the interventions included in this meta-analysis had various limitations, including high dropout rates, lack of engagement of the end user in the development stage, and modest effect sizes [31].

We developed a digital treatment combining CBT and mindfulness training for BED, called *CBT-based Mindful Courage*. We will conduct a small-scale randomized controlled trial (RCT) to assess the feasibility and acceptability of *CBT-based Mindful Courage*. The primary goals are to (1) evaluate feasibility, including intervention and assessment completion; and (2) evaluate acceptability, including dimensions of usability, overall satisfaction, engagement, visual appeal of content, understandability of program content, desire to continue the program, perceived skill acquisition, perceived confidence in implementing skills, and perceived helpfulness. Notably, the study is funded by the National Center for Complementary and Integrative Health (NCCIH) and guided by the NCCIH research framework [33] and the National Institutes of Health stage model for intervention development [34]. Accordingly, we will not conduct between-group analyses at this stage of development. Instead, we will evaluate within-treatment-condition mean changes in outcomes (binge eating frequency, ED symptoms, clinical impairment, and depression) and mechanisms (emotion dysregulation, dietary overrestriction, and trait mindfulness) as exploratory objectives.

XSL•FO
RenderX

Methods

Participant Recruitment and Eligibility Criteria

Participants will be recruited throughout the United States using advertisements posted on social media. The target sample size is 40 (n=20 in each treatment arm). This sample size was based on balancing various considerations, including ethical considerations (eg, participants' time), resource use, prior feasibility trials [35], and availability of funding. Participant eligibility criteria are shown in Textbox 1.

Textbox 1. Participant eligibility criteria.

Inclusion criteria

- Aged 18 to 75 years
- Ability to speak English fluently
- Meeting *Diagnostic and Statistical Manual of Mental Disorders, 5th edition* criteria for current binge eating disorder (BED)
- Willing and able to commit to the entire study protocol

Exclusion criteria

- BMI <18.5
- Requiring immediate treatment for medical complications
- Having current anorexia or bulimia nervosa or purging behaviors within the past year
- Being pregnant or breastfeeding
- Experiencing other severe psychopathology or medical illness that would limit the participants' ability to comply with the demands of this study
- Currently receiving BED or weight loss treatment (treatment for other conditions will be allowed, as long as the treatment is not mindfulness based)
- Currently taking medications for weight loss or beginning medications that affect eating or weight within the last 6 months

Study Design

This study is a 2-arm pilot feasibility RCT with parallel assignment and a 1:1 allocation ratio. All procedures reported here are derived from protocol version 1.0. Participants will be randomized to one of the following groups: (1) coached self-monitoring or (2) *CBT-based Mindful Courage*+coached self-monitoring. Participants will be blinded as to whether they received the active intervention or control condition. Outcome assessors and data analysts will be blinded to group allocation. All recruitment and study activities will be conducted online; the principal investigator and study staff will be located at Yeshiva University. Participants will complete interview-based and self-report assessments assessed at baseline, midtreatment, end-of-treatment, and 2-month follow-up on Health Insurance Portability and Accountability Act (HIPAA)–secured Zoom.

Ethical Considerations

Ethics and Participant Safety

This study has received approval from the institutional review board at Yeshiva University (43318060.0). To our knowledge, the intervention does not pose any significant risks to participants, although some participants may become distressed while engaging in the treatment because we are addressing sensitive content (eg, difficult emotions, ED urges, and eating). All questionnaires have been validated in numerous prior studies. However, some questionnaires involve questions on ED symptoms and diagnoses and other psychological difficulties, which may elicit distress in some participants. All participants will provide signed consent after being informed of study procedures and objectives, possible risks, and the right to withdraw from participation at any time.

Compensation

Participants will receive up to US $180 for completing study assessments. This will include US $20 for the baseline visit, US $30 for the midtreatment assessments, US $40 for the end-of-study assessments, US $60 for the 2-month follow-up assessments, and US $30 for completing all lessons (for participants in the *Mindful Courage*+coached self-monitoring condition only). We decided to provide incentivization for app adherence because we are in an early stage of treatment development, and we wanted to ensure that participants were fully exposed to the intervention to provide acceptability ratings.

Data Privacy and Security

All data will be collected and stored via Qualtrics. Qualtrics is HIPAA secured and uses industry-standard encryption, supports access controls, and has audit logs. Deidentified data will be stored on secure cloud servers and will be retained for a minimum of 5 years. Full access to data will be given only to the principal investigator and approved research staff. Zoom sessions will not be recorded.

Overview of CBT-Based Mindful Courage

Core Lessons

The digital intervention *CBT-based Mindful Courage* was developed through an iterative process considering patient feedback and consists of 16 sequential core lessons, which are completed across 18 weeks (Multimedia Appendix 1). Each lesson takes approximately 40 to 50 minutes to complete and

XSL•FO
RenderX

includes several key features to promote engagement [36-39], including multiple animated videos, audio-guided meditation practices, interactive activities and questions to help participants understand the intervention content in a personally relevant manner, and supportive feedback. After completing each core lesson, participants are instructed to complete 2 audio-guided mindfulness meditation sessions that last 10 to 15 minutes before the next core lesson. The program includes 2 narrators, both clinical psychologists, who appear as avatars in the animated videos to guide participants through the lessons. The program also features several fictional patients, voiced by professional actors hired through Fiverr, who share their experiences through client testimonials and discussions with the 2 psychologists.

The CBT content is derived from enhanced CBT for EDs [9] and a condensed CBT for EDs [40]. The mindfulness content is adapted from a digital MBI for substance use disorder (NCT05852015) developed by Roos et al [35], a prior digital MBI for anorexia nervosa developed by Sala et al [22,23], and several standardized MBI and acceptance and commitment therapy protocols [41-45], particularly mindfulness-based relapse prevention [41].

The intervention also combines CBT and mindfulness techniques. For example, (1) we combine self-monitoring and mindful awareness to increase awareness of eating patterns that may trigger binge eating, (2) we encourage mindful awareness of the consequences of irregular eating, (3) we instruct participants to use mindfulness strategies to notice in their daily life how using CBT strategies (such as reducing dietary restraint and eating satiating meals) affects hunger or fullness and binge eating urges, (4) we combine reduction of valuation of weight and shape with values work, (5) we encourage awareness of ED behaviors and of the relationships between ED behaviors and ED symptoms, and (6) we encourage awareness of factors that may contribute to lapses and relapses.

Animated Content

The animated videos were built with Powtoon, an online animation creation program; the digital lessons were created in Qualtrics; and videos were hosted in Vimeo. Audio content was recorded and edited with Audacity. See Figure 1 for examples of *CBT-based Mindful Courage* visuals.

Figure 1. Example animations and video clips of cognitive behavioral therapy (CBT)–based Mindful Courage. BOAT: Breathe, Observe, Accept, Take a Moment; ED: eating disorder; SOAK: Stop, Observe, Appreciate, Keep Curious.



XSL•FO
RenderX

Overview of Coached Self-Monitoring

Recovery Record

Participants in both conditions will be asked to self-monitor their food intake using *Recovery Record* [46], a frequently used digital intervention among individuals with BED [47]. Participants will not be asked to engage with any other features of *Recovery Record*. We chose to ask participants from both conditions to track their food intake as regular eating is the foundation of ED recovery [48].

Phone Coaching

Participants in both conditions will receive phone coaching by psychology students during the 18-week treatment period but not during the follow-up period. Phone coaching for both conditions will include weekly coaching calls (10-15 minutes, except for week 1, which will be 30 minutes) and extra technical support as needed via phone and SMS text message. In addition, for participants in *CBT-based Mindful Courage*, there will be 1 to 2 weekly SMS text message reminders to use *CBT-based Mindful Courage* and additional phone or SMS text message contact (up to 1 weekly call and 2 weekly SMS text messages) for participants not engaging in *CBT-based Mindful Courage*. The aim of the phone coaching is to provide general motivational and technical support rather than provide any specific therapeutic intervention.

Treatment Conditions

Coached Self-Monitoring

Participants in this condition will not have access to *CBT-based Mindful Courage* but will be asked to self-monitor their food intake in *Recovery Record* and will receive phone coaching.

CBT-Based Mindful Courage+Coached Self-Monitoring

Participants in this condition will have access to *CBT-based Mindful Courage* for 18 weeks. Participants will also be encouraged to engage with *Recovery Record* self-monitoring of food intake and will also receive phone coaching.

Measures

Multimedia Appendix 2 [49-61] provides a summary of assessments and measures.

Screening Measures

We will establish BED diagnosis using the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders, 5th edition* [49]. Other inclusion criteria will be assessed with 1-item questions in an interview-based assessment.

Feasibility Benchmarks

Benchmarks for feasibility are based on previous digital interventions for EDs [31,62-64], and include (1) feasibility of recruitment (40 participants randomized); (2) feasibility of randomization (75% of eligible screens randomized); (3) intervention completion (34/40, 85% of randomized individuals completing at least 1 module; average completion rates of 60% for the 16 lessons) assessed objectively via Qualtrics; (4) completion rates for midtreatment, end-of-treatment, and 2-month follow-up visits (32/40, 80% of participants completing the midtreatment visit, 30/40, 75% of participants completing the end-of-treatment visit, and 26/40, 65% of participants completing the follow-up visit).

Acceptability Benchmarks

Acceptability will include dimensions of usability from the System Usability Scale [65]. We also created additional single-item questions, drawing upon the Technology Acceptance Model [66], evaluating overall satisfaction, engagement, visual appeal of content, understandability of program content, desire to continue the program, perceived skill acquisition, perceived confidence in implementing skills, and perceived helpfulness of *CBT-based Mindful Courage*. Participants in the *CBT-based Mindful Courage*+coached self-monitoring condition will rate their responses to each item from 1 (*strongly disagree*) to 5 (*strongly agree*). Given that this is the first pilot feasibility trial of *CBT-based Mindful Courage*, and our primary intention is to evaluate whether this intervention has preliminary promise for acceptability, benchmarks for these items will be conservatively set to an average score of >3.5 (*agree)*.

ED-Related Outcomes

BED-related outcomes, examined on an exploratory basis to estimate effect sizes for a full-scale RCT, will include (1) frequency of binge eating episodes in the last 28 days, which will be assessed with the Eating Disorder Examination Questionnaire (EDE-Q) [50]; and (2) ED symptoms, assessed with the global EDE-Q score. Additional outcomes will include (3) changes in clinical impairment measured via the Clinical Impairment Assessment [51] and (4) depressive symptoms measured via the 9-item Patient Health Questionnaire-9 [52]. All these measures have demonstrated adequate psychometric properties [51,67-72].

Mechanisms

We will measure mechanisms of change due to the intervention on an exploratory basis. This will include (1) emotion dysregulation, measured via the Difficulties in Emotion Regulation Scale-Short Form [53]; (2) dietary overrestriction, measured via the restraint subscale of the EDE-Q [50]; and (3) trait mindfulness, measured via the Cognitive and Affective Mindfulness Scale-Revised [54]. All these measures have demonstrated adequate psychometric properties [54,67,73].

Additional Measures

We will collect additional measures to be used in exploratory analyses. At baseline, we will collect data on duration of illness, past treatment experiences, past meditation experiences, and eating self-monitoring app use. We will also collect data on adverse events, positive and negative affect measured via the Positive and Negative Affect Schedule [55], savoring measured via the Savoring Beliefs Inventory [56], food craving measured via the Food Cravings Questionnaire-Trait-reduced [57], anxiety symptoms measured via the Generalized Anxiety Disorder-7 item assessment [58], body satisfaction measured via the Body Shape Questionnaire [59], binge eating symptoms measured via the Binge Eating Scale [60], and BMI derived from self-reported height and weight.

XSL•FO
RenderX

Procedure

Screening

See Figure 2 for a graphic of the study timeline. Online advertisements will include a link for an online prescreening (via HIPAA-secured Qualtrics surveys). Interested participants can then use this link to be assessed for preliminary eligibility; this will include questions on age, ability to engage in the research study, and the Eating Disorder Diagnostic Scale [61]. If individuals meet preliminary eligibility based on the online prescreener, they will be invited to a complete a HIPAA-secured Zoom-based screening interview with a trained psychology student. The Zoom-based screening interview will include informed consent, administration of the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders, 5th edition* [49] to confirm BED diagnosis, and a semistructured interview-based assessment to evaluate inclusion and exclusion criteria (refer to the Measures section).

Figure 2. Cognitive behavioral therapy (CBT)–based Mindful Courage timeline. Note: at baseline, participants are randomized to either (1) CBT-based Mindful Courage+coached self-monitoring or (2) coached self-monitoring.

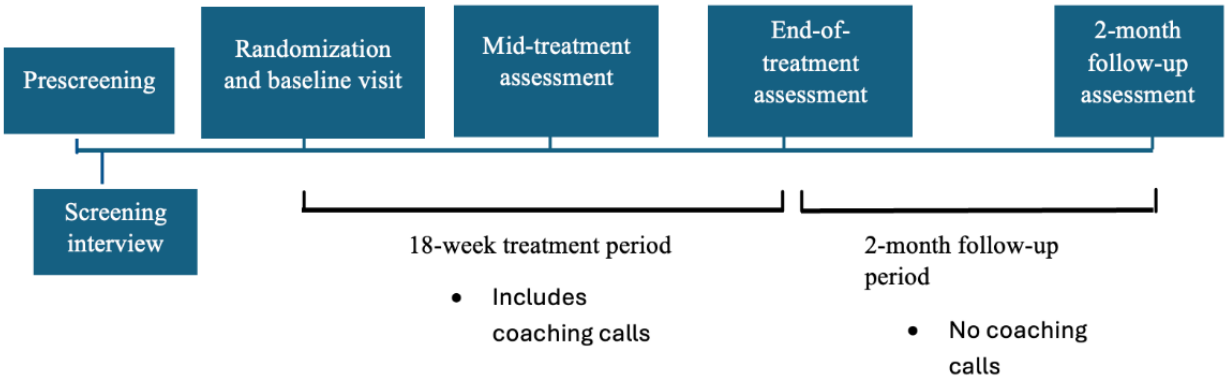

Randomization and Baseline Session

If participants are eligible and willing to continue, they will be randomly assigned to either coached self-monitoring or *CBT-based Mindful Courage*+coached self-monitoring by the psychology student who completed the screening interview. Following our prior work [35], randomization will be achieved using a web-based computerized urn randomization program that ensures balance across treatment groups based on sex, age, binge eating severity, and race or ethnicity. Allocation concealment will be ensured because the program reveals assignments only after participant enrollment. After randomization, participants will proceed to the baseline study session (also conducted over Zoom by the psychology student; Multimedia Appendix 2) and will be instructed on how to track their food intake in *Recovery Record.* They will also receive the orientation for phone coaching. Participants randomized to the *CBT-based Mindful Courage*+coached self-monitoring condition will be instructed on how to access *CBT-based Mindful Courage* (either on their computers or via Android or iPhone) and will complete session 1 of the intervention as part of the visit. After completing session 1, participants will be informed that they will have 18 weeks of access to the app to complete 16 lessons and the 2 accompanying mindfulness practices, with weekly support from a student coach.

Midtreatment, End-of-Treatment, and 2-Month Follow-Up Assessments

After the baseline visit, participants will complete 3 follow-up assessments: (1) midtreatment assessment (weeks 8-9), (2) end-of-treatment assessment (to be administered within 2 weeks of completing the final lesson or reaching the end of the study period; weeks 16-20), and (3) a 2-month follow-up (2 months after the end of treatment; weeks 24-28). The exact assessment timeline may vary for each participant according to scheduling availability and whether the 16 lessons are completed before the end of the study period.

Statistical Analysis

Feasibility and Acceptability

We will use descriptive statistics to examine feasibility and acceptability of *CBT-based Mindful Courage*. Both groups will be included in analyses of recruitment feasibility, randomization, and assessment completion, whereas only participants in the *CBT-based Mindful Courage* group will be included for evaluating acceptability and feasibility of intervention completion. Specifically, we will compute average scores and frequencies for all feasibility and acceptability benchmarks.

Outcomes and Mechanisms

On an exploratory basis, we will evaluate changes in outcomes and mechanisms from baseline to each subsequent assessment point using linear mixed models, separately for participants in each of the conditions (n=20 in each condition). Mixed models enable the inclusion of all participants regardless of missing data (ie, an intent-to-treat analysis). We will also calculate Cohen *d* effect sizes for changes within condition.

Results

The study was funded in April 2023. Data collection began in November 2025. We anticipate that the recruitment will be completed by November 2026. At the time of initial manuscript submission, no participants had been recruited. We expect results to be analyzed by May 2027 and published by May 2028.

XSL•FO
RenderX

Discussion

Anticipated Findings

This protocol describes the design of a pilot RCT with the primary objectives of evaluating the feasibility and acceptability of a digital program combining CBT and mindfulness training for BED—*CBT-based Mindful Courage*. As exploratory objectives, we will evaluate changes in binge eating frequency, global ED symptoms, clinical impairment, and depressive symptoms (as outcomes), as well as emotion dysregulation, dietary overrestriction, and trait mindfulness (as potential mechanisms).

On the basis of our prior research [22,23,32], we anticipate that *CBT-based Mindful Courage* will be acceptable and feasible, setting the stage for an efficacy trial. Our recent study evaluating an 8-week version of *Mindful Courage* for anorexia nervosa and bulimia nervosa concluded that intervention feasibility was excellent and that the intervention was acceptable to participants [23]. Furthermore, our pilot trial evaluating a single-session version of *Mindful Courage* for BED demonstrated excellent acceptability and 100% intervention completion [32]. Preliminary efficacy was demonstrated in both cases, with significant improvements in ED symptoms as well as secondary outcomes (eg, clinical impairment).

Strengths and Limitations

This study will advance our understanding of the feasibility and acceptability of integrating CBT with mindfulness training for EDs. This trial is also innovative as the first digital intervention for BED to place a major emphasis on mindfulness training. Overall, this study advances our understanding of the viability of digital mindfulness–based treatment as a modality for treating EDs in general, and BED in particular.

This study has several potential limitations. First, a common problem with digital treatments and online research studies is high dropout rates, from both the treatment and study assessments. To overcome this problem, we will implement various retention strategies such as graduated incentives and the inclusion of coaches. Second, we have a relatively low sample size. Nevertheless, the sample size is sufficient to evaluate our primary feasibility and acceptability aims, and this type of test follows the NCCIH model for treatment development [33]. Third, the sample may not be representative of the population.

Future Directions

Nevertheless, the study has the potential to be an important step in developing a digital treatment that will provide large-scale delivery of a treatment for BED. This project can serve as a foundation for a full-scale efficacy RCT. A digital treatment can offer unprecedented possibilities for dissemination in the real world. Finally, although this study focuses on BED, results from this study could inform the development of a similar intervention for other EDs.

Acknowledgments

Generative artificial intelligence was not used in any portion of the manuscript generation.

Funding

Research reported in this publication was supported by the National Center for Complementary and Integrative Health of the National Institutes of Health (K23AT012126). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data Availability

The datasets generated or analyzed during this study will be available from the corresponding author on reasonable request

Authors' Contributions

Conceptualization: MS, HK
Funding acquisition: MS, HK
Methodology: MS, HK
Project administration: MS
Software: MS, CR, CS, HK
Supervision: MS
Visualization: MS, DA
Writing – original draft: MS
Writing – review and editing: MS, CR, DA, CS, JB, JMF, CL, MA, JLH, HK

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
Summary of the 16 core lessons in Cognitive behavioral therapy–based Mindful Courage.
[PDF File (Adobe PDF File), 101 KB - resprot_v15i1e91761_app1.pdf]

Multimedia Appendix 2
Summary of study measures.
[PDF File (Adobe PDF File), 177 KB - resprot_v15i1e91761_app2.pdf]

Multimedia Appendix 3
Peer review report from ZAT1 JM (13) - National Center for Complementary and Integrative Health Special Emphasis Panel, NCCIH Training and Education Review Panel (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 209 KB - resprot_v15i1e91761_app3.pdf]

References

1. Santomauro DF, Melen S, Mitchison D, Vos T, Whiteford H, Ferrari AJ. The hidden burden of eating disorders: an extension of estimates from the Global Burden of Disease Study 2019. Lancet Psychiatry 2021 Apr;8(4):320-328 [FREE Full text] [doi: 10.1016/S2215-0366(21)00040-7] [Medline: 33675688]
2. Guerdjikova AI, Mori N, Casuto LS, McElroy SL. Binge eating disorder. Psychiatr Clin North Am 2017 Jun;40(2):255-266. [doi: 10.1016/j.psc.2017.01.003] [Medline: 28477651]
3. Hudson JI, Hiripi E, Pope HG, Kessler RC. The prevalence and correlates of eating disorders in the National Comorbidity Survey Replication. Biol Psychiatry 2007 Feb 01;61(3):348-358 [FREE Full text] [doi: 10.1016/j.biopsych.2006.03.040] [Medline: 16815322]
4. Mitchell JE. Medical comorbidity and medical complications associated with binge-eating disorder. Int J Eat Disord 2016 Mar;49(3):319-323. [doi: 10.1002/eat.22452] [Medline: 26311499]
5. Crow SJ, Swanson SA, Peterson CB, Crosby RD, Wonderlich SA, Mitchell JE. Latent class analysis of eating disorders: relationship to mortality. J Abnorm Psychol 2012 Feb;121(1):225-231 [FREE Full text] [doi: 10.1037/a0024455] [Medline: 21707126]
6. Kowalewska E, Bzowska M, Engel J, Lew-Starowicz M. Comorbidity of binge eating disorder and other psychiatric disorders: a systematic review. BMC Psychiatry 2024 Aug 13;24(1):556 [FREE Full text] [doi: 10.1186/s12888-024-05943-5] [Medline: 39138440]
7. Ágh T, Kovács G, Pawaskar M, Supina D, Inotai A, Vokó Z. Epidemiology, health-related quality of life and economic burden of binge eating disorder: a systematic literature review. Eat Weight Disord 2015 Mar;20(1):1-12 [FREE Full text] [doi: 10.1007/s40519-014-0173-9] [Medline: 25571885]
8. Fairburn CG, Cooper Z, Shafran R. Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. Behav Res Ther 2003 May;41(5):509-528. [doi: 10.1016/s0005-7967(02)00088-8] [Medline: 12711261]
9. Fairburn CG. Cognitive Behavior Therapy and Eating Disorders. New York City, NY: Guilford Press; 2008.
10. National Guideline Alliance (UK). Eating Disorders: Recognition and Treatment. London, UK: National Institute for Health and Care Excellence (NICE); 2017.
11. Fairburn CG, Bailey-Straebler S, Basden S, Doll HA, Jones R, Murphy R, et al. A transdiagnostic comparison of enhanced cognitive behaviour therapy (CBT-E) and interpersonal psychotherapy in the treatment of eating disorders. Behav Res Ther 2015 Jul;70:64-71 [FREE Full text] [doi: 10.1016/j.brat.2015.04.010] [Medline: 26000757]
12. Haedt-Matt AA, Keel PK. Revisiting the affect regulation model of binge eating: a meta-analysis of studies using ecological momentary assessment. Psychol Bull 2011 Jul;137(4):660-681 [FREE Full text] [doi: 10.1037/a0023660] [Medline: 21574678]
13. Kober H, Boswell RG. Potential psychological and neural mechanisms in binge eating disorder: implications for treatment. Clin Psychol Rev 2018 Mar;60:32-44. [doi: 10.1016/j.cpr.2017.12.004] [Medline: 29329692]
14. Leehr EJ, Krohmer K, Schag K, Dresler T, Zipfel S, Giel KE. Emotion regulation model in binge eating disorder and obesity--a systematic review. Neurosci Biobehav Rev 2015 Feb;49:125-134. [doi: 10.1016/j.neubiorev.2014.12.008] [Medline: 25530255]
15. Wonderlich SA, Peterson CB, Crosby RD, Smith TL, Klein MH, Mitchell JE, et al. A randomized controlled comparison of integrative cognitive-affective therapy (ICAT) and enhanced cognitive-behavioral therapy (CBT-E) for bulimia nervosa. Psychol Med 2014 Feb;44(3):543-553 [FREE Full text] [doi: 10.1017/S0033291713001098] [Medline: 23701891]
16. Crane RS, Brewer J, Feldman C, Kabat-Zinn J, Santorelli S, Williams JM, et al. What defines mindfulness-based programs? The warp and the weft. Psychol Med 2017 Apr;47(6):990-999. [doi: 10.1017/S0033291716003317] [Medline: 28031068]
17. Lindsay EK, Creswell JD. Mechanisms of mindfulness training: Monitor and Acceptance Theory (MAT). Clin Psychol Rev 2017 Feb;51:48-59 [FREE Full text] [doi: 10.1016/j.cpr.2016.10.011] [Medline: 27835764]

XSL•FO
RenderX

18. Hölzel BK, Lazar SW, Gard T, Schuman-Olivier Z, Vago DR, Ott U. How does mindfulness meditation work? Proposing mechanisms of action from a conceptual and neural perspective. Perspect Psychol Sci 2011 Nov;6(6):537-559. [doi: 10.1177/1745691611419671] [Medline: 26168376]
19. Godfrey KM, Gallo LC, Afari N. Mindfulness-based interventions for binge eating: a systematic review and meta-analysis. J Behav Med 2015 Apr;38(2):348-362. [doi: 10.1007/s10865-014-9610-5] [Medline: 25417199]
20. Turgon R, Ruffault A, Juneau C, Blatier C, Shankland R. Eating disorder treatment: a systematic review and meta-analysis of the efficacy of mindfulness-based programs. Mindfulness 2019 Aug 20;10:2225-2244. [doi: 10.1007/s12671-019-01216-5]
21. Barney JL, Murray HB, Manasse SM, Dochat C, Juarascio AS. Mechanisms and moderators in mindfulness- and acceptance-based treatments for binge eating spectrum disorders: a systematic review. Eur Eat Disord Rev 2019 Jul;27(4):352-380 [FREE Full text] [doi: 10.1002/erv.2673] [Medline: 30887695]
22. Sala M, Levinson CA, Kober H, Roos CR. A pilot open trial of a digital mindfulness-based intervention for anorexia nervosa. Behav Ther 2023 Jul;54(4):637-651. [doi: 10.1016/j.beth.2023.01.003] [Medline: 37330254]
23. Sala M, Kober H, Levinson C, Ascarrunz D, Stern C, Roos C. Acceptability and feasibility of a digital mindfulness-based intervention for anorexia nervosa and bulimia nervosa. Int J Eat Disord 2026 Jan 22 (forthcoming). [doi: 10.1002/eat.70037] [Medline: 41572612]
24. Roos CR, Sala M, Kober H, Vanzhula IA, Levinson CA. Mindfulness-based interventions for eating disorders: the potential to mobilize multiple associative-learning change mechanisms. Int J Eat Disord 2021 Sep;54(9):1601-1607. [doi: 10.1002/eat.23564] [Medline: 34061387]
25. Hart LM, Granillo MT, Jorm AF, Paxton SJ. Unmet need for treatment in the eating disorders: a systematic review of eating disorder specific treatment seeking among community cases. Clin Psychol Rev 2011 Jul;31(5):727-735. [doi: 10.1016/j.cpr.2011.03.004] [Medline: 21501580]
26. Neyland MK, Bardone-Cone AM. Treatment experiences of Latinas with current or past binge eating disorder and/or bulimia nervosa. Eat Disord 2019;27(2):253-265. [doi: 10.1080/10640266.2019.1591827] [Medline: 31084427]
27. Pauley D, Cuijpers P, Papola D, Miguel C, Karyotaki E. Two decades of digital interventions for anxiety disorders: a systematic review and meta-analysis of treatment effectiveness. Psychol Med 2023 Jan;53(2):567-579 [FREE Full text] [doi: 10.1017/S0033291721001999] [Medline: 34047264]
28. Soh HL, Ho RC, Ho CS, Tam WW. Efficacy of digital cognitive behavioural therapy for insomnia: a meta-analysis of randomised controlled trials. Sleep Med 2020 Nov;75:315-325. [doi: 10.1016/j.sleep.2020.08.020] [Medline: 32950013]
29. Spijkerman MP, Pots WT, Bohlmeijer ET. Effectiveness of online mindfulness-based interventions in improving mental health: a review and meta-analysis of randomised controlled trials. Clin Psychol Rev 2016 Apr;45:102-114 [FREE Full text] [doi: 10.1016/j.cpr.2016.03.009] [Medline: 27111302]
30. Sztein DM, Koransky CE, Fegan L, Himelhoch S. Efficacy of cognitive behavioural therapy delivered over the internet for depressive symptoms: a systematic review and meta-analysis. J Telemed Telecare 2018 Sep;24(8):527-539. [doi: 10.1177/1357633X17717402] [Medline: 28696153]
31. Linardon J, Shatte A, Messer M, Firth J, Fuller-Tyszkiewicz M. E-mental health interventions for the treatment and prevention of eating disorders: an updated systematic review and meta-analysis. J Consult Clin Psychol 2020 Nov;88(11):994-1007. [doi: 10.1037/ccp0000575] [Medline: 32852971]
32. Sala M, Roos CR, Kober H, Bricker JB, Stern CM, Plutchik J, et al. Combining cognitive-behaviour therapy with mindfulness training in a digital intervention for binge eating disorder: a single-session pilot trial. Eur Eat Disord Rev 2025 Sep;33(5):1051-1060. [doi: 10.1002/erv.3204] [Medline: 40350587]
33. NCCIH clinical trials funding opportunities: research framework. National Center for Complementary and Integrative Health. URL: https://www.nccih.nih.gov/grants/nccih-clinical-trials-funding-opportunities/research-framework [accessed 2025-02-24]
34. Onken LS, Carroll KM, Shoham V, Cuthbert BN, Riddle M. Reenvisioning clinical science: unifying the discipline to improve the public health. Clin Psychol Sci 2014 Jan 01;2(1):22-34 [FREE Full text] [doi: 10.1177/2167702613497932] [Medline: 25821658]
35. Roos CR, Bricker J, Kiluk B, Trull TJ, Bowen S, Witkiewitz K, et al. A smartphone app-based mindfulness intervention to enhance recovery from substance use disorders: protocol for a pilot feasibility randomized controlled trial. Contemp Clin Trials Commun 2024 Jul 26;41:101338 [FREE Full text] [doi: 10.1016/j.conctc.2024.101338] [Medline: 39233850]
36. Alkhaldi G, Hamilton FL, Lau R, Webster R, Michie S, Murray E. The effectiveness of prompts to promote engagement with digital interventions: a systematic review. J Med Internet Res 2016 Jan 08;18(1):e6 [FREE Full text] [doi: 10.2196/jmir.4790] [Medline: 26747176]
37. Andersson G. Internet-delivered psychological treatments. Annu Rev Clin Psychol 2016;12:157-179. [doi: 10.1146/annurev-clinpsy-021815-093006] [Medline: 26652054]
38. Barak A, Klein B, Proudfoot JG. Defining internet-supported therapeutic interventions. Ann Behav Med 2009 Aug;38(1):4-17. [doi: 10.1007/s12160-009-9130-7] [Medline: 19787305]
39. Proudfoot J, Klein B, Barak A, Carlbring P, Cuijpers P, Lange A, et al. Establishing guidelines for executing and reporting internet intervention research. Cogn Behav Ther 2011;40(2):82-97. [doi: 10.1080/16506073.2011.573807] [Medline: 25155812]

XSL•FO
RenderX

40. Juarascio AS, Manasse SM, Schumacher L, Espel H, Forman EM. Developing an acceptance-based behavioral treatment for binge eating disorder: rationale and challenges. Cogn Behav Pract 2017 Feb;24(1):1-13 [FREE Full text] [doi: 10.1016/j.cbpra.2015.12.005] [Medline: 29881247]
41. Bowen S, Chawla N, Witkiewitz K. Mindfulness-based relapse prevention for addictive behaviors. In: Baer RA, editor. Mindfulness-Based Treatment Approaches (2nd Edition). Cambridge, MA: Academic Press; 2014.
42. Juarascio AS, Parker MN, Hunt R, Murray HB, Presseller EK, Manasse SM. Mindfulness and acceptance-based behavioral treatment for bulimia-spectrum disorders: a pilot feasibility randomized trial. Int J Eat Disord 2021 Jul;54(7):1270-1277 [FREE Full text] [doi: 10.1002/eat.23512] [Medline: 33851734]
43. Kabat-Zinn J. Mindfulness-based stress reduction (MBSR). Constr Hum Sci 2003;8(2):73-107.
44. Kristeller J, Wolever RQ, Sheets V. Mindfulness-based eating awareness training (MB-EAT) for binge eating: a randomized clinical trial. Mindfulness 2013 Feb 1;5:282-297. [doi: 10.1007/s12671-012-0179-1]
45. Segal ZV, Williams JM, Teasdale JD. Mindfulness-Based Cognitive Therapy for Depression: A New Approach to Preventing Relapse. New York City, NY: Guilford Publications; 2002.
46. Tregarthen JP, Lock J, Darcy AM. Development of a smartphone application for eating disorder self-monitoring. Int J Eat Disord 2015 Nov 27;48(7):972-982. [doi: 10.1002/eat.22386] [Medline: 26213130]
47. Reilly EE, Gorrell S, Johnson C, Duffy A, Blalock DV, Mehler P, et al. Characterising use of recovery record among a large, transdiagnostic sample of adults with eating disorders across higher levels of care. Eur Eat Disord Rev 2024 May;32(3):404-416. [doi: 10.1002/erv.3053] [Medline: 37997259]
48. Kim JP, Sadeh-Sharvit S, Welch HA, Neri E, Tregarthen J, Lock J. Eating disorders early app use mediates treatment effect on clinical improvement. Int J Eat Disord 2022 Mar;55(3):382-387. [doi: 10.1002/eat.23652] [Medline: 34904745]
49. First MB, Williams JB, Karg RS, Spitzer RL. The Structured Clinical Interview for DSM-5 Disorders: Research Version (SCID-5-RV). Washington, DC: American Psychiatric Association; 2015.
50. Fairburn CG, Beglin S. Assessment of eating disorders: interview or self-report questionnaire? Int J Eat Disord 1994 Dec;16(4):363-370. [Medline: 7866415]
51. Bohn K, Doll HA, Cooper Z, O'Connor M, Palmer RL, Fairburn CG. The measurement of impairment due to eating disorder psychopathology. Behav Res Ther 2008 Oct;46(10):1105-1110 [FREE Full text] [doi: 10.1016/j.brat.2008.06.012] [Medline: 18710699]
52. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
53. Kaufman EA, Xia M, Fosco G, Yaptangco M, Skidmore CR, Crowell SE. The Difficulties in Emotion Regulation Scale Short Form (DERS-SF): validation and replication in adolescent and adult samples. J Psychopathol Behav Assess 2016 Sep;38(3):443-455. [doi: 10.1007/s10862-015-9529-3] [Medline: 41522882]
54. Feldman G, Hayes A, Kumar S, Greeson J, Laurenceau JP. Mindfulness and emotion regulation: the development and initial validation of the Cognitive and Affective Mindfulness Scale-Revised (CAMS-R). J Psychopathol Behav Assess 2006 Nov 7;29:177-190. [doi: 10.1007/s10862-006-9035-8]
55. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037/0022-3514.54.6.1063]
56. Bryant F. Savoring Beliefs Inventory (SBI): a scale for measuring beliefs about savouring. J Ment Health 2009 Jul 06;12(2):175-196. [doi: 10.1080/09638230310001034891]
57. Cepeda-Benito A, Gleaves DH, Williams TL, Erath SA. The development and validation of the state and trait food-cravings questionnaires. Behav Ther 2000;31(1):151-173. [doi: 10.1016/S0005-7894(00)80009-X]
58. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
59. Cooper PJ, Taylor MJ, Cooper Z, Fairbum CG. The development and validation of the body shape questionnaire. Int J Eat Disord 1987 Jul;6(4):485-494. [doi: 10.1002/1098-108X(198707)6:4<485::AID-EAT2260060405>3.0.CO;2-O]
60. Gormally J, Black S, Daston S, Rardin D. The assessment of binge eating severity among obese persons. Addict Behav 1982;7(1):47-55. [doi: 10.1016/0306-4603(82)90024-7] [Medline: 7080884]
61. Stice E, Telch CF, Rizvi SL. Development and validation of the eating disorder diagnostic scale: a brief self-report measure of anorexia, bulimia, and binge-eating disorder. Psychol Assess 2000;12(2):123-131. [doi: 10.1037/1040-3590.12.2.123]
62. Fitzsimmons-Craft EE, Taylor CB, Graham AK, Sadeh-Sharvit S, Balantekin KN, Eichen DM, et al. Effectiveness of a digital cognitive behavior therapy-guided self-help intervention for eating disorders in college women: a cluster randomized clinical trial. JAMA Netw Open 2020 Aug 03;3(8):e2015633 [FREE Full text] [doi: 10.1001/jamanetworkopen.2020.15633] [Medline: 32865576]
63. Melioli T, Bauer S, Franko DL, Moessner M, Ozer F, Chabrol H, et al. Reducing eating disorder symptoms and risk factors using the internet: a meta-analytic review. Int J Eat Disord 2016 Jan;49(1):19-31. [doi: 10.1002/eat.22477] [Medline: 26607683]
64. Pretorius N, Arcelus J, Beecham J, Dawson H, Doherty F, Eisler I, et al. Cognitive-behavioural therapy for adolescents with bulimic symptomatology: the acceptability and effectiveness of internet-based delivery. Behav Res Ther 2009 Sep;47(9):729-736. [doi: 10.1016/j.brat.2009.05.006] [Medline: 19515360]

XSL•FO
RenderX

65. Brooke J. SUS: a 'quick and dirty' usability scale. In: Usability Evaluation in Industry. Boca Raton, FL: CRC Press; 1996.
66. Marangunić N, Granić A. Technology acceptance model: a literature review from 1986 to 2013. Univ Access Inf Soc 2014 Feb 16;14:81-95. [doi: 10.1007/s10209-014-0348-1]
67. Berg KC, Peterson CB, Frazier P, Crow SJ. Psychometric evaluation of the eating disorder examination and eating disorder examination-questionnaire: a systematic review of the literature. Int J Eat Disord 2012 Apr;45(3):428-438 [FREE Full text] [doi: 10.1002/eat.20931] [Medline: 21744375]
68. Pook M, Tuschen-Caffier B. Sensitivity to change of scales assessing symptoms of bulimia nervosa. Psychiatry Res 2004 Aug 30;128(1):71-78. [doi: 10.1016/j.psychres.2003.03.001] [Medline: 15450916]
69. Pook M, Tuschen-Caffier B, Brähler E. Evaluation and comparison of different versions of the Body Shape Questionnaire. Psychiatry Res 2008 Feb 28;158(1):67-73. [doi: 10.1016/j.psychres.2006.08.002] [Medline: 18037499]
70. Ford J, Thomas F, Byng R, McCabe R. Use of the patient health questionnaire (PHQ-9) in practice: interactions between patients and physicians. Qual Health Res 2020 Nov;30(13):2146-2159 [FREE Full text] [doi: 10.1177/1049732320924625] [Medline: 32564676]
71. Becker AE, Thomas JJ, Bainivualiku A, Richards L, Navara K, Roberts AL, et al. Adaptation and evaluation of the Clinical Impairment Assessment to assess disordered eating related distress in an adolescent female ethnic Fijian population. Int J Eat Disord 2010 Mar;43(2):179-186 [FREE Full text] [doi: 10.1002/eat.20665] [Medline: 19308992]
72. Reas DL, Rø O, Kapstad H, Lask B. Psychometric properties of the clinical impairment assessment: norms for young adult women. Int J Eat Disord 2010 Jan;43(1):72-76. [doi: 10.1002/eat.20653] [Medline: 19260038]
73. Skutch JM, Wang SB, Buqo T, Haynos AF, Papa A. Which brief is best? Clarifying the use of three brief versions of the difficulties in emotion regulation scale. J Psychopathol Behav Assess 2019 Sep 15;41(3):485-494 [FREE Full text] [doi: 10.1007/s10862-019-09736-z] [Medline: 34446987]

Abbreviations

BED: binge eating disorder
CBT: cognitive behavioral therapy
ED: eating disorder
EDE-Q: Eating Disorder Examination Questionnaire
HIPAA: Health Insurance Portability and Accountability Act
MBI: mindfulness-based intervention
NCCIH: National Center for Complementary and Integrative Health
RCT: randomized controlled trial

Edited by J Sarvestan; submitted 19.Jan.2026; this is a non–peer-reviewed article;accepted 11.Mar.2026; published 17.Apr.2026.

Please cite as:
Sala M, Roos C, Ascarrunz D, Stern CM, Bricker J, Feldman JM, Levinson C, Aslan M, Hay JL, Kober H
Combining Digital Cognitive Behavioral Therapy With Mindfulness Training for Binge Eating Disorder: Protocol for a Feasibility Trial
JMIR Res Protoc 2026;15:e91761
URL: https://www.researchprotocols.org/2026/1/e91761
doi:10.2196/91761
PMID:40729081

©Margaret Sala, Corey Roos, Dante Ascarrunz, Casey M Stern, Jonathan Bricker, Jonathan M Feldman, Cheri Levinson, Mihaela Aslan, Jennifer L Hay, Hedy Kober. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Engaging Hospital Staff to Identify Levers for Adoption of Clinical Decision Support: Protocol for a Single-Site Case Study Using System Dynamics Group Model Building

Nina Rachel Sperber[1,2], PhD; Lindsey Eileen Zimmerman[3,4], PhD; Lauren Caton[5], PhD; Tom Rust[6], PhD; Armando Bedoya[7], MD; Sarah Haas[1], BS; Lori Orlando[8], MD; Ben Goldstein[1,7], PhD

[1]Department of Population Health Sciences, Duke University School of Medicine, 215 Morris St., Suite 210, Durham, NC, United States

[2]Durham VA Health Care System, Durham, NC, United States

[3]VA Palo Alto Health Care System, Dissemination and Training Division, National Center for PTSD, Menlo Park, CA, United States

[4]Division of Public Mental Health and Population Sciences, Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Palo Alto, CA, United States

[5]University of North Carolina, Chapel Hill, NC, United States

[6]United States Department of Veterans Affairs, Washington, DC, United States

[7]Department of Biostatistics and Bioinformatics, Duke University School of Medicine, Durham, NC, United States

[8]Department of General Internal Medicine, Wake Forest University School of Medicine, Winston Salem, NC, United States

Corresponding Author:
Nina Rachel Sperber, PhD
Department of Population Health Sciences, Duke University School of Medicine, 215 Morris St., Suite 210, Durham, NC, United States

Abstract

Background: Clinical decision support (CDS) tools that provide patient-specific and evidence-based information to clinicians and care managers regarding patient risk for adverse outcomes have been a part of health care for decades. However, modern CDS, which consists of automated predictions based on complex machine learning models and hundreds of complex input variables, faces obstacles to adoption related to health care providers' perceptions of lack of transparency and utility. Often, the expertise of data scientists and clinical end users is not well integrated, creating implementation gaps from CDS development to adoption and ongoing implementation.

Objective: This protocol describes the use of group model building (GMB) from the field of system dynamics to engage health system staff in identifying dynamic facilitators and barriers to implementing 1 class of CDS—early warning scores (EWSs)—in general medical-surgical wards. We aim to produce a causal model that reflects the insights and feedback shared during these sessions. We will also evaluate the GMB process as a potential strategy for CDS implementation and adoption more generally.

Methods: The protocol consists of 3 sequential GMB sessions designed to elicit key variables for inclusion in the model, understand how changes in variable behavior over time affect adoption, and develop a causal loop diagram. Pre- and postsession questionnaires assess changes in perceived acceptability, appropriateness, and feasibility of the EWS and collect feedback on the GMB process. A stock-and-flow simulation model will be developed from the causal loop diagram to quantify how feedback loops influence variables over time and test assumptions.

Results: The project was funded in 2022-2025, and 3 GMB sessions and qualitative causal loop diagrams were completed during that time. Data analysis is ongoing. This analysis consists of translating the causal loop diagram from the GMB into a stock-and-flow simulation model to quantify how feedback loops influence variables over time. Results will include a causal loop diagram accompanied by a detailed narrative that together tell a story about system behavior surrounding EWS adoption that is supported by session transcripts, the simulation model and test results, and the data on GMB participants' views about the EWS itself and the modeling process.

Conclusions: These findings will have broader applicability beyond just EWSs. Future work will build on the EWS system dynamics model by incorporating multiple clinical use cases to fully capture multilevel factors that determine real-world adoption and sustainability of machine learning CDS.

International Registered Report Identifier (IRRID): DERR1-10.2196/8084

(JMIR Res Protoc 2026;15:e80848) doi:10.2196/80848

XSL•FO
RenderX

KEYWORDS

clinical decision support; machine learning; early warning score; system dynamics; group model building; implementation science

Introduction

Problem

Clinical decision support (CDS) tools provide patient-specific and evidence-based information to health care providers and care managers regarding patient risk for adverse outcomes. CDS has been a part of health care for decades and is a key component of the learning health care system [1]. Overall, CDS has led to improvements in patient and health care provider outcomes [1-3]. With the advent of integrated electronic health record (EHR) systems over the past decade, there has been a proliferation of predictive models, that is, algorithms based on clinical, typically EHR, data that generate estimates of the probability of outcomes (eg, sepsis, readmission, and mortality) for any given patient, and in turn can alert health care providers to the need for specific interventions [3,4]. Predictive models can use a variety of approaches to compute algorithms ranging from simple logic-based rules to interpretable statistical models to, more recently, machine learning (ML) algorithms that learn from examples and include the use of both structured and unstructured data [5].

While predictive ML models of any type can improve the identification of high-risk patients, there are obstacles to implementation and adoption. Clinicians have responded to a perceived overabundance of CDS alerts by developing "alert fatigue" and ignoring the CDS [6,7]. In fact, prior research indicates that between 49% and 96% of alerts are ignored or overridden [8-10]. Nonresponse to alerts often occurs when clinical users do not understand or trust the CDS [11,12]. ML-based CDS faces particular obstacles to adoption by users who perceive a lack of transparency in how predictions are made, affecting their views about its validity or usefulness in clinical care [9,13]. Participatory approaches that include input from health system staff and others throughout the CDS lifecycle, from algorithm development to implementation and maintenance, could improve the adoption of ML-based CDS [14]. Typically, only a few select clinical users contribute to the developmental process, and data scientists are often not well integrated within the clinical environment, even though their design decisions impact downstream use of the CDS tools [15]. This lack of integrated participation creates a barrier in communication and knowledge transfer between the technical and clinical realms, thereby limiting the utility and application of CDS in routine clinical practice.

Objective

The objective of this project is to bring health system staff together to identify facilitators and barriers to ML CDS implementation in their contexts. To do this, we use group model building (GMB), a methodology from the field of system dynamics that facilitates collaborative modeling of complex problems [16,17]. In system dynamics, complex processes are modeled by specifying variables, feedback loops, and time delays. In GMB, accuracy, impact, and usability of system dynamics models are enhanced by directly involving individuals with lived experiences of the problem being addressed to develop the models [18]. Used most commonly in organizational change management as a strategic planning technique to identify solutions from the bottom-up, GMB has more recently emerged as a recommended strategy to address problems with adoption of evidence-based practices for health care [19,20]. In this case, GMB is used to foster a shared understanding of why the implementation problem persists over time and to identify potential solutions as compared to the status quo [16]. Our research uses GMB to engage a multidisciplinary group of health system staff involved in CDS implementation.

Methods

Study Design

This protocol uses a single-site case study design within a single health care system. GMB is used within the case study to elicit and represent stakeholders' perspectives on adoption dynamics. Together, this design and methodology enable an in-depth exploration of conditions influencing adoption of ML-enabled CDS in a real-world implementation setting.

Use Case

This case study focuses on 1 class of CDS, early warning scores (EWSs), which have been adopted well in some settings but not others. EWSs are widely used by both large academic and small community hospitals to identify patients at risk for decompensation, leading to intensive care or mortality [21]. Using predominantly vital signs, EWSs typically predict the need for an intensive care unit transfer and/or the risk of death within a 12- to 48-hour time horizon, typically referred to as unexpected decompensation. The first widely used EWS was the National Early Warning Score (NEWS), a point-based score of 7 vital signs measurements [22]. Since the introduction of the NEWS, over 20 EWSs have been introduced in the literature, ranging from simple modifications of the NEWS (eg, extra laboratory measurement) to more complex scores that integrate laboratory data, demographics, and comorbidities [23]. Over the last several years, EWS scores have evolved from point-based scores to logistic regression–based scores (with variable weights) to complicated ML-based scores. Benefits of EWS adoption have been observed. For example, a large analysis across Kaiser Health indicated a 16% (10% - 22%) reduction in the adjusted relative risk of mortality in hospitals that used their EWS [24]. As a result of findings such as these, investments are continually made to develop and improve EWS CDS, and there are many choices involved in EWS implementation. These range from the target user (doctor or nurse), type of model (ML or points based), and temporality of the alert (real time or interval based). Our prior research showed that due to lack of adoption, the technical integration of NEWS into the EHR of study sites did not change rates of mortality or unanticipated intensive care unit transfer, even when the accuracy of the ML score indicated it was performing well [12]. Frontline nurses in our health care system ignored the alert 86% of the time [12]. In discussion groups, nurses expressed neither

XSL•FO
RenderX

understanding of how the score operates nor how to use it clinically and described experiences with unhelpful CDS behaviors, including reports of alerts that were not accepted continuing to fire on the same patient. The EWS developers subsequently amended the score to present a more qualitative—and easily interpreted—assessment of risk and to fire only around shift changes (8 AM and 8 PM) as opposed to continually throughout the day [25].

Setting

This study leverages Duke University Health System's (DUHSs) broad experience with EWS, often talked about monolithically despite existing within a variety of contexts. Within the DUHS, different EWS implementation contexts create a natural variation. This variation allows us to assess how different conditions affect EWS adoption.

We selected 2 hospitals within the DUHS, a large quaternary care hospital, Duke University Hospital (DUH), and a smaller community hospital, Durham Regional Hospital (DRH), that use 2 different EWS types. While the hospitals have used a shared DUHS Epic-based EHR system since 2013, they also have separate governance structures as to how and which CDS tools are implemented. In 2015, DUHS implemented the NEWS score across 3 hospitals. After the impact evaluation in 2019, the score was turned off at DUH and replaced with a homegrown EWS trained on DUH data [12]. This resulted in the continued use of the NEWS at DRH and the use of Duke EWS at DUH, providing an opportunity to compare conditions that affect adoption [25].

The NEWS and Duke EWS differ in 5 specific ways. First, the underlying scores are fundamentally different with regard to the input variables and the information that the score will display. The NEWS is a points-based score consisting of 7 input variables. Users can see both a patient's overall score as well as the points associated with each of the input variables. Conversely, the Duke EWS is an ML-based score consisting of over 40 input variables. As such, it is not possible to discern how any 1 variable impacts a patient's overall score. Second, while both scores are directly integrated into the EHR, they are operationalized differently in terms of how they are integrated into the workflow with timing of alerts. The NEWS is operationalized as a continuously firing alert that provides real-time feedback. The Duke EWS, based on user input, is designed to be used at shift change only (ie, discrete time points), to inform which patients are most critical. Third, each score has different target users. The NEWS broadly targets staff nurses caring for admitted patients, while the Duke EWS targets primarily charge nurses, responsible for overseeing staff and operations of the staff nurses. Fourth, the scores differ in the information displayed. The NEWS is presented to the end user as an overall score. The Duke EWS presents as a color-coded 3-level risk category (red, yellow, and green). Fifth, overall model interpretability differs. The overall score that the NEWS displays has no specific probabilistic interpretation. The alert fires when the score crosses a prespecified threshold. However, Duke EWS risk categories indicate a patient's probabilistic risk level. The risk category thresholds were chosen based on desired operational performance with respect to sensitivity and positive predictive value.

Participant Recruitment

We identified developers (data scientists), implementers, and users who had interacted with 1 of the 2 EWS versions in use at DUHS through snowball sampling, asking an initial contact identified by a member of our study team to refer us to other interested stakeholders. They were invited to participate with an emailed letter signed by the principal investigators. We continued in this manner until we recruited at least 1 developer and implementer for each or both of the 2 EWS scores and at least 2 users of each score. Everyone was asked to participate in 3 sessions. To minimize attrition, we scheduled the 3 sessions as a series at 1 time and according to the participants' availability and kept each session to a maximum of 2 hours.

Ethical Considerations

This study was reviewed and approved by the DUHS Institutional Review Board (Pro00111061). All participants provided informed consent prior to participation. Participants who consented and attended 1 or more GMB sessions were included in the study, with attendance across all 3 sessions encouraged but not required. Participant privacy is protected through use of unique study ID numbers, secure storage of reidentification logs in restricted-access Duke Box and network drives, deidentification of all materials, aggregate reporting, and destruction of audio recordings after publication; no identifiable names are required during group sessions or recordings. Participants received no compensation.

GMB Overview

Overview

The 3 sessions were designed to engage hospital staff in developing a model of EWS adoption using the established GMB activities of variable elicitation (to identify key variables for inclusion in the model), graphing behavior over time (to understand how change in variable behavior affects adoption), and causal loop diagramming (to visualize how feedback loops affect adoption) [26-28]. Each session was built on the outputs of the previous one. These sessions were delivered using the Zoom application (Zoom Communications, Inc) with a password entry, audio recording, and transcriptions generated. Miro electronic whiteboard (Miro, Inc), a collaborative platform, was used as a group workspace [29,30]. The core modeling team planned and participated in all sessions, with each person having a distinct perspective and role: the team included a coordinator who organized and managed the flow of the sessions; a trained GMB facilitator who conducted the sessions; a qualitative researcher who served as a cofacilitator; a biostatistician or data scientist with experience developing CDS ML algorithms who observed as a subject matter expert; and a clinician data scientist who provided a community-based perspective. A detailed facilitation guide specified timing, roles, and scripts (Multimedia Appendix 1).

Session 1: Variable Elicitation

The first session used variable elicitation to identify key variables affecting EWS adoption in general medicine settings

XSL•FO
RenderX

for inclusion in the system dynamics model [26]. This session included 10 staff who worked with NEWS or Duke EWS. They were brought together in 1 session with the end goal of converging on a set of key variables. It began with the facilitator presenting the problem of failure to use the EWS CDS, showing a depiction of the CDS development-to-use gap in the form of a conceptually stylized graph with a line representing the hoped-for trajectory of a high, steady level of commitment compared to the observed trajectory of declining commitment. This type of image is used in GMB to define the problem with the group and create a shared reference [31]. The facilitator asked the group whether the image captured their perspectives and if changes were needed. The facilitator then asked participants to describe components that influence the current trajectory with questions about perceived acceptability (What are key factors that inhibited or facilitated satisfaction with EWS?), perceived appropriateness (What are key factors that inhibited or facilitated compatibility with practice settings or users?), and perceived feasibility (What are factors that inhibited or facilitated successful use of the EWS in their setting?). The cofacilitator posted responses on the whiteboard in thematic clusters. The facilitator probed for detail, such as whether the clusters resonated with the group, whether other themes should be included, or if themes should be removed, and asked participants to prioritize factors in order of importance.

Session 2: Behavior Over Time

The second session aimed to understand how the variables identified in the first session changed or could change over time [27]. In this session, the facilitator asked participants to graph trajectories of the variables on the whiteboard, with the x-axis indicating progression of time and the y-axis indicating changing values of the concept. The graphs included hoped-for, expected ("business as usual"), or feared trajectories. The facilitator probed participants to discuss their hypotheses about what causes and is caused by these trajectories. For example, participants identified quality of data input as a factor affecting use of the score, and the facilitator asked them to describe more about how they had observed the quality of inputs changing during a patient's stay in the unit. The DUH staff, DRH staff, and developers met separately from each other to facilitate identification of themes specific to their contexts.

Session 3: Causal Loop Diagrams

The third session engaged all staff to converge on a hypothesis about structural drivers of EWS adoption, that is, the underlying factors and relationships that shape how the EWS is used in practice, by contributing to the development of a causal loop diagram, which visually depicts these relationships through feedback loops [28]. Prior to the session, the study team developed a simple causal loop diagram to provide an initial scaffold of causal connections for participants to review, critique, and elaborate by reviewing notes from the prior sessions, "walking through" causal relationships and loops, and considering how the diagram fit with what they knew about the behavior of hospital units. The facilitator initiated the GMB session by reminding the group of the problem and the variables and graphs discussed in the earlier exercises. The facilitator then presented the causal loop diagram and solicited input by asking participants whether the diagram reflected their experiences, if they had any changes, and whether the directionality of relationships between variables depicted (eg, increase in one variable causes an increase in another vs an increase in one variable causes a decrease in another) was accurate. After this session, the study team created a new version of the causal loop diagram according to content from the session and their own observations of the health system. The team met with health system staff to validate this synthesis through an individual interview with a charge nurse who had participated in the sessions and a group interview with 2 staff from in-hospital and nursing leadership who worked with EWSs and had not participated in the sessions.

Measures

A questionnaire administered at 4 time points, prior to the first session and following each session, measured changes in participants' views of the EWS over the course of their involvement with GMB. Validated implementation science measures of perceived acceptability of an intervention (whether they like it), perceived appropriateness of an intervention (whether it fits with their practice), and perceived feasibility of an intervention (whether they can successfully use it) were used [32]. These measures were selected because collectively the constructs can act as a proxy for assessing commitment to adopting a new tool such as the EWS [32]. After each session, participants also provided feedback on the GMB process (Multimedia Appendix 2).

Analysis

System Dynamics Modeling

This analysis consists of integrating quantitative and qualitative information. The GMB causal loop diagram and transcripts will serve as the primary qualitative data sources. We are translating the causal loop diagram from GMB into a stock-and-flow simulation model to quantify how feedback loops influence variables over time. In a stock-and-flow model, stocks represent quantities that accumulate or decrease, such as patient counts, while flows indicate the rates at which these stocks change. The stock-and-flow simulation model will reflect the insights and feedback shared during the GMB sessions, with adjustable parameters to explore what would happen to system behavior if 1 or more values change. We will define each stock and flow relationship with mathematical equations and capture delays, nonlinearities, and feedback effects using Vensim (Ventana Systems) simulation software's built-in functions. We anticipate generating time-series graphs, which show how key variables, such as patient counts or adoption of CDS, change over time. We will also conduct scenario comparisons, running the model under different assumptions to see how outcomes vary. The research team will thus use the simulation model to test the internal logic of the causal loop diagram and refine its underlying assumptions.

Statistical Analysis

The statistical analysis will consist of descriptive summaries of presession and postsession questionnaire data to compare responses by developers and the DRH and DUH settings. We will calculate the mean, SD, and range of the Acceptability of

XSL•FO
RenderX

Intervention Measure, Intervention Appropriateness Measure, and Feasibility of Intervention Measure and the feedback items for each group [32]. We will construct a dot plot using RStudio for the Acceptability of Intervention Measure, Intervention Appropriateness Measure, and Feasibility of Intervention Measure that shows the mean score for each of the 3 groups at the 4 time points when the questionnaire was administered (prior to the first session and after each of the 3 sessions). We will also construct a dot plot that shows the mean score for each of the feedback items after each of the 3 sessions along with the open-ended responses.

Results

The project was funded in 2022-2025, and three GMB sessions were completed during that time. Data analysis is ongoing. Results will include a causal loop diagram accompanied by a detailed narrative that together tell a story about system behavior surrounding EWS adoption. The causal loop diagram will be supported by session transcripts, the simulation model and test results, and the data on GMB participants' views about the EWS itself and the modeling process. These results will illustrate processes and mechanisms that affect EWS adoption under complex, dynamic real-world conditions.

Discussion

The results will have broader applicability beyond just EWSs. When developing and implementing ML-enabled CDS tools in general, many decisions need to be made. These include the nature of the underlying score, who the user base is, and how and when it is presented to those users. Often, these decisions stem from siloed and artificial test environments that limit the ability to evaluate how they function in real-world settings. Early in the model building process, developers often make design decisions that impact the overall implementation of the CDS, typically with usability studies of single tools conducted with individual user groups in research settings [33,34]. Although these studies demonstrate technical capability of the tools, a major limitation is that they do not consider applied function within the context of typical clinical care workflows. Once the CDS is embedded in the clinical setting, there is often a disconnect between what the developer intended and users' interaction. Such an orientation is likely to cause low adoption by end users. This approach to engage developers, implementers, and users to create a shared visual model during implementation planning or execution will help to improve their understanding of the problem that the CDS is meant to address and, in turn, improve buy-in and the potential for adoption [17,35].

Future work will build on the EWS system dynamics model that is the focus of this protocol by incorporating multiple clinical use cases to fully capture multilevel factors that determine real-world adoption and sustainability of artificial intelligence (AI)-CDS. The resulting interactive simulation model will provide an environment to test the implementation of AI-enabled CDS more generally across health systems and identify intervenable levers to improve efficiency, reducing the reliance on costly real-world research usually needed to identify implementation barriers, facilitators, and strategies. The simulation model will provide an interactive learning infrastructure for responsible AI deployment to understand the behavior of AI tools as they move from development to implementation by allowing users to visualize effects over time.

Interactively modeling system dynamics through GMB and a simulation model promotes a learning health care system environment by allowing those who develop, implement, and use CDS tools to reflect on what did and did not work in that process [36]. This approach could offer value during the exploratory phase, when decision-making occurs, to better understand implementation issues and help define implementation strategies that consider not only CDS features but also system complexity, and thus improve sustainability of ML-enabled CDS over time [37]. Having a way to support the ongoing planning and monitoring of any ML-enabled CDS in health care is important to ensure its safe and effective use given an evolving regulatory framework, and given that most ML-enabled tools depend on nuances of local data [33]. This approach to understanding facilitators and barriers to EWS adoption will be useful for researchers and practitioners interested in proposing an informatics implementation project.

Acknowledgments

The manuscript was initially drafted entirely by the authors without AI assistance. Generative AI (OpenAI's ChatGPT, Microsoft's Copilot, GPT-5.4) was subsequently used solely to clarify text, improve flow, and ensure consistency throughout the manuscript. AI-generated suggestions were limited to wording, transitions, and formatting; all scientific content, study design, and interpretation were determined solely by the authors. All AI-generated text was reviewed, verified, and edited by the authors for accuracy, completeness, and alignment with the intended meaning. To support transparency, we have included a table of prompts used during AI-assisted text refinement, along with the corresponding sections of the manuscript where AI input was applied (Multimedia Appendix 3).

Funding

This study was funded by the National Institutes of Health (5R21-LM013649-02). The funder had no role in the design of the study, data collection, analysis, or interpretation of the results. This study was partially supported by a National Research Service Award Pre-Doctoral Traineeship from the Agency for Healthcare Research and Quality (AHRQ) sponsored by The Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill (grant T32-HS000032).

XSL•FO
RenderX

Authors' Contributions

NRS and BG contributed to the conceptualization of the study and acquired funding. LEZ contributed to the conceptualization of the study, drafted the manuscript, and is contributing to ongoing data analysis. LC facilitated the sessions and contributed to manuscript preparation. TR is assisting with data analysis. AB and LO will contribute to the data analysis and interpretation. SH is assisting with literature review. All authors reviewed the manuscript prior to submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Facilitation guide.
[DOCX File, 219 KB - resprot_v15i1e80848_app1.docx]

Multimedia Appendix 2
Questionnaire for feedback.
[DOC File, 44 KB - resprot_v15i1e80848_app2.doc]

Multimedia Appendix 3
Artificial intelligence prompts.
[DOCX File, 13 KB - resprot_v15i1e80848_app3.docx]

References

1. Middleton B, Sittig DF, Wright A. Clinical decision support: a 25 year retrospective and a 25 year vision. Yearb Med Inform 2016 Aug 2;Suppl 1(Suppl 1):S103-S116. [doi: 10.15265/IYS-2016-s034] [Medline: 27488402]
2. Grechuta K, Shokouh P, Alhussein A, et al. Benefits of clinical decision support systems for the management of noncommunicable chronic diseases: targeted literature review. Interact J Med Res 2024 Nov 27;13:e58036. [doi: 10.2196/58036] [Medline: 39602213]
3. Elhaddad M, Hamam S. AI-driven clinical decision support systems: an ongoing pursuit of potential. Cureus 2024 Apr;16(4):e57728. [doi: 10.7759/cureus.57728] [Medline: 38711724]
4. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc 2017 Jan;24(1):198-208. [doi: 10.1093/jamia/ocw042] [Medline: 27189013]
5. Rajkomar A, Dean J, Kohane I. Machine learning in medicine. N Engl J Med 2019 Apr 4;380(14):1347-1358. [doi: 10.1056/NEJMra1814259] [Medline: 30943338]
6. Ancker JS, Edwards A, Nosal S, et al. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. BMC Med Inform Decis Mak 2017 Apr 10;17(1):36. [doi: 10.1186/s12911-017-0430-8] [Medline: 28395667]
7. Johnson KR, Hagadorn JI, Sink DW. Alarm safety and alarm fatigue. Clin Perinatol 2017 Sep;44(3):713-728. [doi: 10.1016/j.clp.2017.05.005] [Medline: 28802348]
8. van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006;13(2):138-147. [doi: 10.1197/jamia.M1809] [Medline: 16357358]
9. Abbas Q, Jeong W, Lee SW. Explainable AI in clinical decision support systems: a meta-analysis of methods, applications, and usability challenges. Healthcare 2025 Aug 29;13(17):2154. [doi: 10.3390/healthcare13172154]
10. Olakotan OO, Mohd Yusof M. The appropriateness of clinical decision support systems alerts in supporting clinical workflows: a systematic review. Health Informatics J 2021;27(2):14604582211007536. [doi: 10.1177/14604582211007536] [Medline: 33853395]
11. Schwartz JM, George M, Rossetti SC, et al. Factors influencing clinician trust in predictive clinical decision support systems for in-hospital deterioration: qualitative descriptive study. JMIR Hum Factors 2022 May 12;9(2):e33960. [doi: 10.2196/33960] [Medline: 35550304]
12. Bedoya AD, Clement ME, Phelan M, Steorts RC, O'Brien C, Goldstein BA. Minimal impact of implemented early warning score and best practice alert for patient deterioration. Crit Care Med 2019 Jan;47(1):49-55. [doi: 10.1097/CCM.0000000000003439] [Medline: 30247239]
13. Matheny ME, Whicher D, Thadaney Israni S. Artificial intelligence in health care: a report from the National Academy of Medicine. JAMA 2020 Feb 11;323(6):509-510. [doi: 10.1001/jama.2019.21579] [Medline: 31845963]
14. Sittig DF, Boxwala A, Wright A, et al. A lifecycle framework illustrates eight stages necessary for realizing the benefits of patient-centered clinical decision support. J Am Med Inform Assoc 2023 Aug 18;30(9):1583-1589. [doi: 10.1093/jamia/ocad122] [Medline: 37414544]

XSL•FO
RenderX

15. Bastian G, Baker GH, Limon A. Bridging the divide between data scientists and clinicians. Intell Based Med 2022;6:100066. [doi: 10.1016/j.ibmed.2022.100066]
16. Hovmand PS. Group model building and community-based system dynamics process. In: Community Based System Dynamics: Springer; 2014:17-30. [doi: 10.1007/978-1-4614-8763-0_2]
17. Vennix J. Group Model Building: Facilitating Team Learning Using System Dynamics: John Wiley & Sons; 1996.
18. Király G, Miskolczi P. Dynamics of participation: system dynamics and participation—an empirical review. Syst Res Behav Sci 2019;36(2):199-210. [doi: 10.1002/sres.2580]
19. Holmes BJ, Finegood DT, Riley BL, Best A. Systems thinking in dissemination and implementation research. In: Brownson RC, Colditz GA, Proctor EK, editors. Dissemination and Implementation Research in Health: Translating Science to Practice: Oxford University Press; 2012.
20. Brownson RC, Colditz GA, Proctor EK. Dissemination and Implementation Research in Health: Translating Science to Practice: Oxford University Press; 2017.
21. McGaughey J, Alderdice F, Fowler R, Kapila A, Mayhew A, Moutray M. Outreach and early warning systems (EWS) for the prevention of intensive care admission and death of critically ill adult patients on general hospital wards. Cochrane Database Syst Rev 2007 Jul 18(3):CD005529. [doi: 10.1002/14651858.CD005529.pub2] [Medline: 17636805]
22. Smith GB, Prytherch DR, Meredith P, Schmidt PE, Featherstone PI. The ability of the National Early Warning Score (NEWS) to discriminate patients at risk of early cardiac arrest, unanticipated intensive care unit admission, and death. Resuscitation 2013 Apr;84(4):465-470. [doi: 10.1016/j.resuscitation.2012.12.016] [Medline: 23295778]
23. Gerry S, Bonnici T, Birks J, et al. Early warning scores for detecting deterioration in adult hospital patients: systematic review and critical appraisal of methodology. BMJ 2020 May 20;369:m1501. [doi: 10.1136/bmj.m1501] [Medline: 32434791]
24. Escobar GJ, Liu VX, Schuler A, Lawson B, Greene JD, Kipnis P. Automated identification of adults at risk for in-hospital clinical deterioration. N Engl J Med 2020 Nov 12;383(20):1951-1960. [doi: 10.1056/NEJMsa2001090] [Medline: 33176085]
25. O'Brien C, Goldstein BA, Shen Y, et al. Development, implementation, and evaluation of an in-hospital optimized early warning score for patient deterioration. MDM Policy Pract 2020 Jan 10;5(1):2381468319899663. [doi: 10.1177/2381468319899663] [Medline: 31976373]
26. Scriptapedia/variable elicitation. Wikibooks. URL: https://en.wikibooks.org/wiki/Scriptapedia/Variable_Elicitation [accessed 2026-03-26]
27. Scriptapedia/graphs over time. Wikibooks. URL: https://en.wikibooks.org/wiki/Scriptapedia/Graphs_over_Time [accessed 2026-03-26]
28. Scriptapedia/initiating and elaborating a causal loop diagram. Wikibooks. URL: https://en.wikibooks.org/wiki/Scriptapedia/Initiating_and_Elaborating_a_Causal_Loop_Diagram [accessed 2026-03-26]
29. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using Zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. Int J Qual Methods 2019;18:160940691987459. [doi: 10.1177/1609406919874596]
30. Wilkerson B, Aguiar A, Gkini C, Czermainski de Oliveira I, Lunde Trellevik LK, Kopainsky B. Reflections on adapting group model building scripts into online workshops. Syst Dyn Rev 2020;36(3):358-372. [doi: 10.1002/sdr.1662]
31. Rouwette E, Vennix JAM, Mullekom TV, Mullekom T. Group model building effectiveness: a review of assessment studies. Syst Dyn Rev 2002 Mar;18(1):5-45. [doi: 10.1002/sdr.229]
32. Weiner BJ, Lewis CC, Stanick C, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Aug 29;12(1):108. [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
33. Pencina MJ, Silcox C, Economou-Zavlanos N, McClellan M. Bridging the gap between developers and implementers in health AI. JAMA Health Forum 2025 Jun 7;6(6):e251692. [doi: 10.1001/jamahealthforum.2025.1692] [Medline: 40540282]
34. Wiens J, Saria S, Sendak M, et al. Do no harm: a roadmap for responsible machine learning for health care. Nat Med 2019 Sep;25(9):1337-1340. [doi: 10.1038/s41591-019-0548-6] [Medline: 31427808]
35. Damschroder LJ, Reardon CM, Widerquist MA, Lowery J. The updated Consolidated Framework for Implementation Research based on user feedback. Implement Sci 2022 Oct 29;17(1):75. [doi: 10.1186/s13012-022-01245-0] [Medline: 36309746]
36. Steel PAD, Wardi G, Harrington RA, Longhurst CA. Learning health system strategies in the AI era. npj Health Syst 2025;2(1):21. [doi: 10.1038/s44401-025-00029-0]
37. Aarons GA, Hurlburt M, Horwitz SM. Advancing a conceptual model of evidence-based practice implementation in public service sectors. Adm Policy Ment Health 2011 Jan;38(1):4-23. [doi: 10.1007/s10488-010-0327-7] [Medline: 21197565]

Abbreviations

AI: artificial intelligence
CDS: clinical decision support
DRH: Duke Regional Hospital
DUH: Duke University Hospital
DUHS: Duke University Health System

XSL•FO
RenderX

EHR: electronic health record
EWS: early warning score
GMB: group model building
ML: machine learning
NEWS: National Early Warning Score

Edited by J Sarvestan; submitted 21.Jul.2025; peer-reviewed by K Trinkley, MK Lemke; revised version received 20.Feb.2026; accepted 20.Feb.2026; published 21.Apr.2026.

Please cite as:
Sperber NR, Zimmerman LE, Caton L, Rust T, Bedoya A, Haas S, Orlando L, Goldstein B
Engaging Hospital Staff to Identify Levers for Adoption of Clinical Decision Support: Protocol for a Single-Site Case Study Using System Dynamics Group Model Building
JMIR Res Protoc 2026;15:e80848
URL: https://www.researchprotocols.org/2026/1/e80848
doi:10.2196/80848

© Nina Rachel Sperber, Lindsey Eileen Zimmerman, Lauren Caton, Tom Rust, Armando Bedoya, Sarah Haas, Lori Orlando, Ben Goldstein. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Synthesis and Standardization of Outcomes in Severe Malaria Treatment Trials: Protocol for the Development of a Core Outcome Set (the COSSMaT Study)

Gideon Darko Asamoah[1*], BSc, MSc; Sharon B Love[1*], BSc, MSc, PhD; Diana M Gibb[1*], MD, MSc; Kathryn Maitland[2,3*], MD, PhD; Elizabeth C George[1*], BSc, MSc, PhD

[1]Institute of Clinical Trials and Methodology, MRC Clinical Trials Unit at UCL, 90 High Holborn, London, United Kingdom

[2]Institute of Global Health Innovation, Department of Surgery and Cancer, Faculty of Medicine, Imperial College London, London, United Kingdom

[3]KEMRI Wellcome Trust Programme, Kilifi, Kenya

[*]all authors contributed equally

Corresponding Author:
Gideon Darko Asamoah, BSc, MSc
Institute of Clinical Trials and Methodology, MRC Clinical Trials Unit at UCL, 90 High Holborn, London, United Kingdom

Abstract

Background: Severe malaria remains a major global health challenge, particularly in endemic regions, where it causes substantial morbidity and mortality, especially among children and pregnant women. Although numerous clinical trials have evaluated treatments for severe malaria, heterogeneity in outcome selection, definition, and measurement limits comparability across studies, hampers evidence synthesis, and contributes to research waste. There is currently no established core outcome set (COS) for trials specifically focused on severe malaria treatment. Developing a COS is therefore essential to improve outcome standardization, strengthen evidence synthesis, and support evidence-based clinical practice.

Objective: This study aims to develop a COS for trials evaluating the treatment of severe malaria.

Methods: This protocol follows guidance from the Core Outcome Measures in Effectiveness Trials (COMET) Initiative and the Core Outcome Set–Standardised Protocol Items (COS-STAP) checklist. An updated systematic review of outcomes reported in randomized controlled trials of severe malaria treatment has been conducted. Searches covered the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE via Ovid, and Literatura Latino-Americana em Ciências da Saúde (LILACS), as well as clinical trial registries, including the International Standard Randomised Controlled Trial Number (ISRCTN) registry, ClinicalTrials.gov, and the Pan-African Clinical Trial Registry (PACTR). Study selection followed PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidance. Outcomes important to patients and families are being identified through qualitative interviews with parents and caregivers of patients treated for severe malaria, with data analyzed using thematic synthesis. Outcomes identified across these sources will be prioritized through a 2-round, multistakeholder Delphi survey, with rounds separated by 4 weeks and a target sample of 60 to 80 participants. An online consensus meeting will then be held to agree on the final COS. Subsequently, outcome measurement instruments used in severe malaria trials will be identified and assessed, and a dissemination and implementation strategy will be developed to support uptake of the final COS.

Results: Preliminary results from the updated systematic review identified 326 records through database and registry searches. Following screening and full-text review, 5 eligible randomized controlled trials conducted between 2020 and 2024 in Africa and Asia were included. Qualitative interviews with parents and caregivers have been conducted with 26 participants recruited and interviewed. The Delphi survey, consensus meeting, and outcome measurement instrument selection are ongoing and expected to be completed by September 2026.

Conclusions: This protocol outlines a rigorous and inclusive approach to developing the first COS for severe malaria treatment trials. The resulting COS aims to improve outcome standardization, reduce reporting bias, and enhance the comparability and impact of future severe malaria trials.

Trial Registration: COMET 2908; https://www.comet-initiative.org/Studies/Details/2908

International Registered Report Identifier (IRRID): DERR1-10.2196/78616

(JMIR Res Protoc 2026;15:e78616) doi:10.2196/78616

KEYWORDS

severe malaria; core outcome sets; Delphi survey; clinical trials; treatment

XSL•FO
RenderX

Introduction

Background

Malaria persists as a formidable global health challenge, exacting a heavy toll on both morbidity and mortality rates worldwide [1-3]. Africa, in particular, bears a disproportionate burden, with 94% of malaria cases (246 million) and 95% (n=569,000) of malaria deaths recorded by the World Health Organization (WHO) in 2023 [2]. Children younger than 5 years constituted approximately 76% of malaria-related fatalities, making malaria a leading cause of childhood illness and death in the region [2,4]. Beyond its impact on children, malaria significantly influences maternal outcomes directly and neonatal outcomes indirectly because it heightens the risks of abortion, low birth weight, preterm delivery, stillbirth, and neonatal death [5-7].

The pursuit of improved malaria treatment remains an ongoing endeavor, particularly for severe disease [8,9]. However, a significant impediment arises from the lack of consensus on appropriate outcomes for use in clinical trials, thereby posing a challenge to measuring and evaluating the efficacy of antimalarial and/or adjunctive treatments. A systematic review has highlighted the considerable heterogeneity, with 101 diverse outcome measures reported in 27 trials spanning from 2010 to 2020 [10]. This disparity has led to the measurement of diverse outcomes and the use of various instruments for assessing the same outcome, resulting in inconsistencies in reported results and complicating the comparison of outcomes in systematic reviews and meta-analyses [10]. Creating a core outcome set (COS) for severe malaria treatment in trials could help mitigate the current deficiencies in addressing reporting challenges.

A COS is defined by the Core Outcome Measures in Effectiveness Trials (COMET) as "an agreed standardized set of outcomes that should be measured and reported, as a minimum, in all clinical trials in specific areas of health or health care" [11]. While a previous study has developed COS for malaria in the context of vaccine trials [12], a notable gap exists concerning the treatment of severe malaria, especially with the inclusion of adjunctive therapies [10]. This void is a critical concern for global health, given the severe and widespread impact of severe malaria. Urgent attention is warranted to address this gap.

The scope of the COS specifies the particular health care domain or condition it addresses, detailing its target population, interventions, and settings. This clarity ensures its appropriate and consistent use, thereby avoiding any ambiguity [13-15]. For this COS, the scope encompasses severe malaria, characterized by complications such as impaired consciousness, acidosis, and severe anemia. The target population includes all groups affected by severe malaria, including children under five, adults, pregnant women, and travelers who reside in endemic and nonendemic regions. The interventions involve antimalarial treatments and adjunctive therapies to manage complications. The setting is within clinical trials to standardize outcome selection and reporting, ensuring relevance and consistency across studies. The aim of the study is to develop a COS for trials in the treatment of severe malaria.

Study Summary

We conducted a search for an existing COS for severe malaria treatment trials, adhering to the COMET guidelines' recommendation to prevent duplication of ongoing or preexisting studies. No such study was found.

We will develop the COS through the following phases:

- Phase 1: develop a "long list" of outcomes to guide the creation of a COS for severe malaria treatment trials, we will use the following information sources:
  i. Systematic review update: identify additional outcome sets reported for trials in the treatment of severe malaria
  ii. Consultation with clinical experts in severe malaria to incorporate recent and unpublished research insights
  iii. Qualitative study: identify severe malaria outcomes reported by patients, parents, and guardians or caregivers for measurement in treatment trials
- Phase 2: develop a preliminary COS (informed by the "long list" of outcomes from phase 1) with key stakeholders through a multistage (2- or 3-round, web-based) Delphi survey
- Phase 3: hold a consensus meeting with stakeholders to discuss and agree on the final COS for severe malaria treatment
- Phase 4: identify outcome measurement instruments for the outcome sets reported in severe malaria treatment trials
- Phase 5: develop a dissemination and implementation strategy for the final COS

Methods

Overview

This study protocol will adhere to the guidelines outlined in the Core Outcome Set–Standardised Protocol Items (COS-STAP) statement checklist [16] (Checklist 1). The study has been registered on the COMET Initiative website [17]. The study will follow the COMET Initiative guidelines for developing a COS to produce a reliable, valid, and responsive minimum set of outcomes to be measured and reported in all clinical trials considering the treatment of severe malaria [13]. The schematic overall flow and link between the different phases of the study are presented in Figure 1.

XSL•FO
RenderX

Figure 1. Schematic figure showing the overall workflow and link between the different phases. COS: core outcome set.

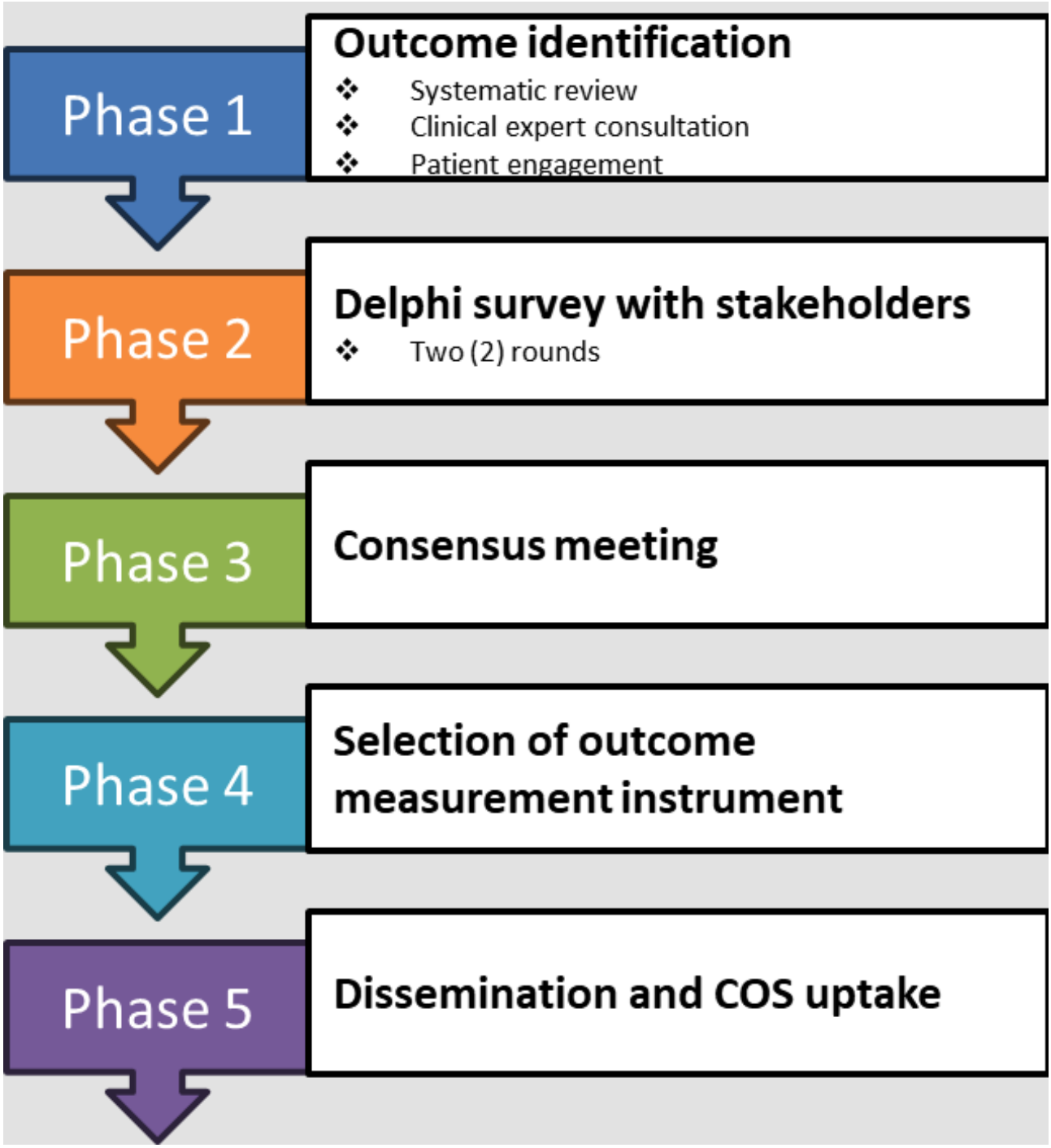

Phase 1: "Long List" of Outcomes to Guide the Creation of a COS for Severe Malaria Treatment Trials: Systematic Review Update

The systematic review of the study sought to build upon an earlier systematic review that examined outcomes reported in trials for severe malaria treatment conducted between January 1, 2010, and June 30, 2020 [10]. The objective of the update was to capture and integrate any newly published or registered trials focusing on severe malaria treatment conducted from July 1, 2020, to July 26, 2024. The updated systematic review followed the guidelines outlined in the "Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions" checklist (Checklist 2) [18].

Eligibility Criteria

The eligibility criteria for this systematic review were established to ensure the inclusion of studies directly addressing the research objectives.

Inclusion Criteria

- Study design: Only randomized controlled trials (RCTs) conducted between 2020 and 2024 were included, representing the COS's scope of use.
- Population: The review included studies focusing exclusively on hospitalized adults or children diagnosed

XSL•FO
RenderX

with severe malaria, including cerebral malaria, based on the WHO criteria [19]. This population was chosen to provide insights into treatment efficacy and outcomes within the context of severe disease.
- Intervention and comparison: Eligible studies assessed the use of antimalarial therapies and adjunctive treatments for severe malaria. This criterion enabled a focused synthesis of evidence on therapeutic strategies for managing the condition.

Exclusion Criteria

- Studies involving animal models or nonhuman subjects
- Observational studies, literature reviews, and secondary analyses of RCTs
- Nonrandomized trials and research focusing on malaria prevention or vaccination
- Trials that included both severe and uncomplicated malaria without distinguishing between the two
- Conference abstracts, posters, or other non–peer-reviewed publications

Search Strategy

The search strategy for this study was carefully designed to capture all relevant studies about severe malaria treatment within the defined scope. The search terms focused on 4 critical concepts: severe malaria, treatment, hospital, and RCTs. The search was restricted to English-language publications to ensure clarity and accessibility of the content. The time frame was deliberately chosen to provide an updated and comprehensive overview of outcomes from RCTs published between July 2020 and the search date. This time frame bridges the gap left by a prior systematic review [10], whose last search concluded in June 2020. The final search was conducted on July 26, 2024, covering publications up to that date. A complete list of the search terms and the detailed construction of the search strategy for MEDLINE, Literatura Latino-Americana em Ciências da Saúde (LILACS), and ClinicalTrials.gov are provided in Multimedia Appendix 1.

Information Sources

Searches were conducted across key databases and clinical trial registries, including the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE via Ovid, and LILACS, to ensure comprehensive geographical coverage. The search also extended to registries such as the International Standard Randomized Controlled Trial Number (ISRCTN), ClinicalTrials.gov, and the Pan-African Clinical Trial Registry (PACTR) to identify ongoing and unpublished trials, addressing potential publication bias. To enhance robustness, citation lists of retrieved trials were reviewed for additional records, ensuring no relevant studies were overlooked [20,21].

Study Selection Process

The search results were organized using EndNote X9 (Clarivate) reference management software. Duplicate entries were removed, followed by a double-screening of records. Titles and abstracts were initially reviewed to identify relevant trials, with eligible trials undergoing a detailed full-text review to ensure they met the inclusion criteria. For trials from registries without published results, entries were cross-referenced using key identifiers. We discussed and resolved any eligibility concerns by consensus. Trials not meeting the criteria were excluded, with explanations documented. The screening and selection process is detailed in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 2).

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) diagram of included studies.

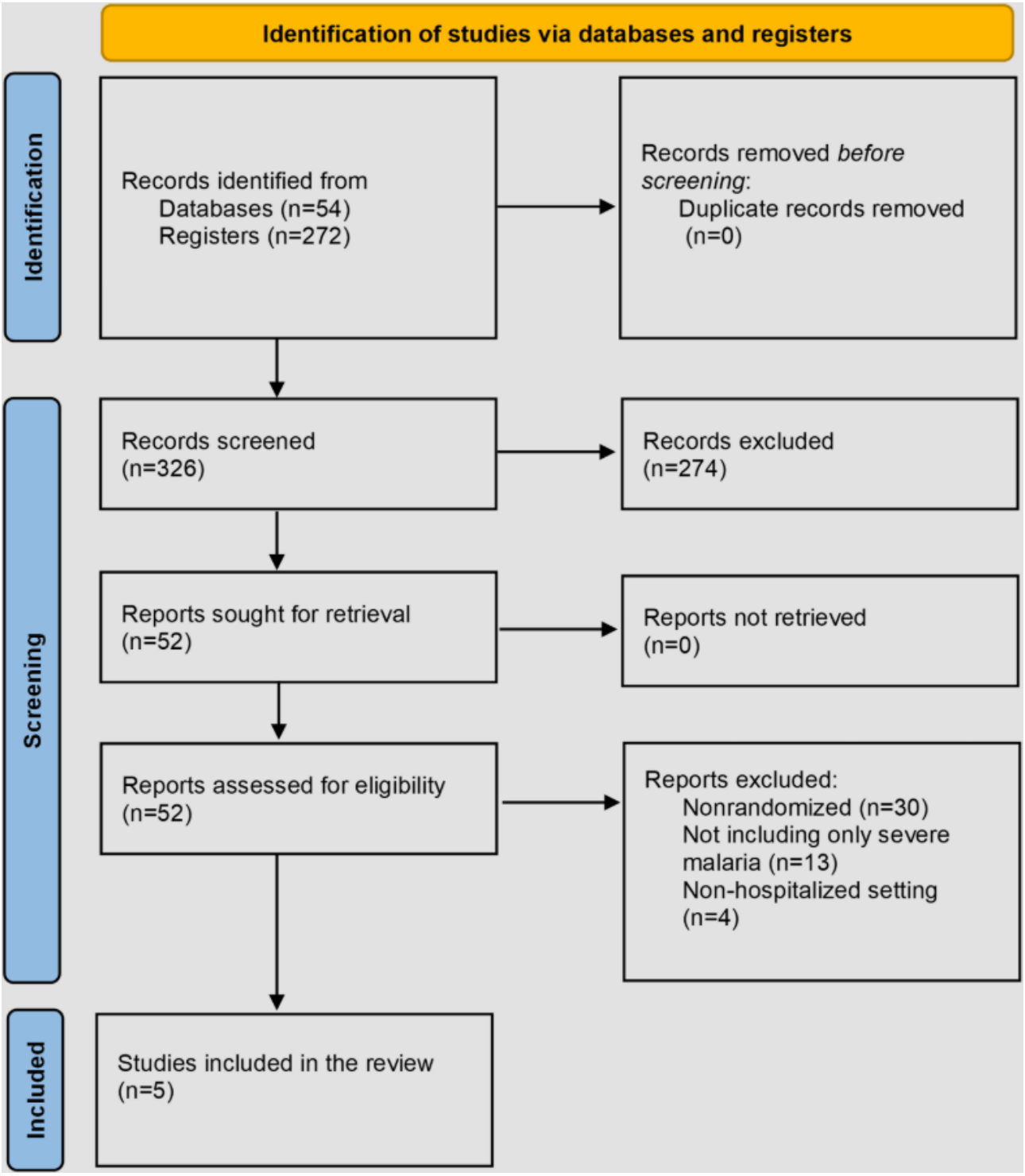

Data Charting and Extraction

To systematically organize and analyze information, a data charting form was developed in Microsoft Excel (Multimedia Appendix 2). This form outlined specific variables to be collected, aligning with the study's aim. The variables captured essential trial details, ensuring thorough data collection. The extracted data included trial characteristics such as study title, publication year, participant demographics, study design and phase, intervention type, and outcome measures. Outcome measures were described verbatim from study protocols or trial registry entries to maintain accuracy. Details about the instruments or methods used to measure these outcomes were also included, providing context for how the outcomes were quantified.

Consultation With Clinical Experts

Clinical experts were engaged to provide essential insights into practical and clinically relevant outcomes. Input was sought from clinical experts in the field of severe malaria, whose insights and familiarity with recent and unpublished research

XSL•FO
RenderX

added an extra layer of comprehensiveness to the search strategy.

Qualitative Study

The COMET guidelines emphasize the importance of actively involving patients and the public throughout the development process to ensure that the selected outcomes are relevant and meaningful to those directly impacted [13]. By engaging with parents and caregivers, this qualitative research directly addresses the COMET Initiative guidelines on patient and public involvement [13]. Such engagement ensures that "all potentially relevant outcomes" are considered, thereby enhancing the comprehensiveness and relevance of the outcome list [22,23].

The study was conducted following the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines (Checklist 3) [24] and the reporting recommendations for qualitative research methods in COS development outlined by Jones et al [25], as used in previous COS studies [22].

Aim

The study aims to use qualitative research methods to explore and identify the outcomes of severe malaria that parents, guardians, and carers consider important to measure in treatment trials.

Design

A descriptive qualitative research design was used in this study to identify outcomes reported by parents, guardians, and caregivers of children affected by severe malaria. Descriptive qualitative research aims to provide a detailed and comprehensive account of a phenomenon by focusing on the "who, what, where, and when" aspects, without delving into theoretical interpretations or complex explanations [26-28].

Without patient involvement in the COS development process, there is a risk that important outcomes will be overlooked, which could hinder the ability to assess treatment effectiveness. Several studies have used qualitative research in COS development to identify relevant outcomes, capture stakeholders' perspectives, and ensure that the outcomes included in clinical trials reflect the priorities of patients, caregivers, and health care professionals [22,29-31]. Incorporating qualitative methods into the outcome identification process offers several benefits, particularly in providing a deeper understanding of why certain outcomes are important to patients, parents, and caregivers [22,32]. This helps ensure that the outcomes presented in the COS reflect the true priorities and needs of those most affected by severe malaria.

Setting

Participants for this study were recruited from Ghana, a country in sub-Saharan Africa, which is one of the regions that bears a disproportionately high burden of malaria [33-35]. Targeting recruitment in these settings allowed for the inclusion of populations facing significant challenges with severe malaria, enriching our understanding of its impact. This study was conducted in Kumasi, the capital of Ghana's Ashanti Region, a malaria hotspot with an increase in cases in recent years [36,37], with clinical malaria prevalence rates ranging from 3.4% to 16.9% [38]. The study was conducted at the Komfo Anokye Teaching Hospital (KATH) in Kumasi, within its Child Health Directorate. KATH was chosen for its role as the principal referral center for the region, providing specialized care for severe and complex conditions, including severe malaria, and offering access to a diverse patient population. The Child Health Directorate's expertise in pediatric care enabled engagement with a broad range of stakeholders and grounded the research in a context reflecting the challenges and priorities of populations at the epicenter of the malaria burden.

Recruitment Strategy

Participants were recruited from the Severe Malaria Africa: A Research and Trials Consortium (SMAART) observational study, which aims to describe the treatment for children with severe malaria from hospital presentation to 6 months after discharge. The study enrolled 2 cohorts: hospitalized children with severe malaria and time-matched controls with nonsevere malaria, with follow-up over 6 months to assess outcomes. Children aged 3 months to 15 years were included to generate insights into severe malaria management.

In total, 3 supervisors involved in this study are members of the SMAART Consortium. They facilitated contact with the local principal investigator at KATH, enabling the identification of key gatekeepers. These gatekeepers reviewed patient records to identify eligible participants, whose parents, guardians, or caregivers were then informed about the study and invited to participate.

Sampling and Sample Characteristics

We initially planned to recruit parents, guardians, or caregivers of children currently receiving treatment for severe malaria at KATH. However, we revised our protocol to include those whose children had been treated for severe malaria within the past year. This change was made after observing that the severity of the children's conditions often left their caregivers too distracted to engage fully in interviews. The 1-year time frame aimed to reduce recall bias [39,40]. Participants were purposefully sampled based on the following criteria: (1) aged ≥18 years; (2) parent and guardian or caregiver caring for an infant, diagnosed with and treated for severe malaria within a year at the time of recruitment; and (3) those admitted to and treated at the KATH during the time of recruitment.

Data Collection Method

Experienced qualitative researchers, GDA and 2 SMAART study team members, conducted face-to-face semistructured interviews with parents, guardians, or caregivers. Semistructured interviews were chosen for their flexibility and ability to explore experiences, attitudes, and perceptions in depth [41,42]. This method allows researchers to guide discussions while adapting to participant responses, making it effective for collecting open-ended data and delving into personal and sensitive issues. The interviews were guided by a topic guide based on predefined themes from previous research, ensuring a thorough exploration of participant responses (Multimedia Appendix 3) [22,43,44].

Interviews were conducted in the local language to ensure participants felt comfortable. Consent was obtained before recording, respecting participants' autonomy and privacy. The

XSL•FO
RenderX

interviews took place in a private office within the Child Health Directorate, providing a confidential environment. During the interviews, demographic information of participants was collected, and a "concept-elicitation" approach encouraged open dialogue. Topics included understanding of severe malaria, important outcomes, hospital-based treatment experiences, and the broader impact of severe malaria on their lives. This approach ensured participants' voices were central to the study's findings.

Management of Risk

The research team provided detailed explanations of the study procedures to potential participants, ensuring they had sufficient time to review the participant information sheet (PIS) and informed consent form (ICF) thoroughly. Participants were encouraged to ask questions and seek clarifications to ensure they fully understood the study's purpose, requirements, and role. The research staff, all trained in ethical research practices, carefully assessed participants' comprehension of the study procedures. Participants were informed that their participation was voluntary, with the right to withdraw at any time without affecting their medical care. A distress protocol adopted from a previous study (Multimedia Appendix 4) [22] was in place to support participants during potentially distressing interviews. If participants appeared uncomfortable, interviews were paused or stopped, and support options, such as referrals to health care professionals, were offered.

Data Analysis

Overview

All interview recordings will be transcribed verbatim and translated into English. To ensure accuracy and preserve meaning, translations will be independently verified by a second researcher fluent in both languages. This verification process is essential for maintaining the integrity of participants' accounts, particularly in relation to culturally specific meanings and idiomatic expressions [45-47].

NVivo software (version 14; Lumivero) [48] will be used to manage and organize the qualitative data. Each transcript will be treated as a distinct case, representing an individual participant [49]. Data analysis will be conducted using thematic synthesis. An inductive coding approach will be applied, comprising both semantic coding to identify outcomes explicitly reported by participants and interpretive coding to capture inferred concerns and underlying meanings [22,50].

To enhance rigor and consistency, 2 experienced qualitative researchers will independently code an initial subset of 10 transcripts. The coding framework will then be discussed and refined through iterative meetings before being applied to the remaining transcripts. This process will support analytical consistency and transparency across the dataset.

Generation of the "Long List" of Outcomes

In line with COMET guidance, all identified outcomes will be collated into a comprehensive long list and reviewed for duplication and conceptual overlap. Outcomes that are similar or closely related will be consolidated, and the refined list will be organized into broader outcome domains using the COMET-recommended outcome taxonomy for medical research [51], supported by frameworks applied in previous COS studies [10,52]. Domain allocation will be undertaken iteratively and refined through consultation with clinical experts to ensure clinical relevance, conceptual coherence, and consistency with the objectives of the COSSMaT (core outcome sets for severe malaria trials) study.

This structured and transparent approach to outcome identification and domain development will provide a robust foundation for subsequent Delphi surveys and consensus meetings, as recommended by COMET, and will support the development of a COS that is methodologically rigorous, stakeholder informed, and relevant to the evaluation of treatments for severe malaria.

Phase 2: Delphi Survey With Stakeholders

Overview

The purpose of a Delphi survey is to collect opinions from a panel of participants, including experts in severe malaria and clinicians who regularly manage patients with severe malaria, on a specific topic to achieve group consensus [53]. Key stakeholders will be presented with plain language explanations of the "long list" of outcomes and will then rate the importance of each item for inclusion in the COS through a multistage Delphi survey.

Stakeholder Groups and Recruitment

Key stakeholders, including patient or support group representatives, health care workers, pharmaceutical industry representatives, policymakers, and researchers, will be involved in the COS development process. Stakeholders will be identified and recruited through expert networks associated with public health agencies, regulatory authorities, research institutions, pharmaceutical industries, and civil society organizations. Notable examples of such networks include the WHO, the Malaria Consortium, the SMAART Consortium, the COMET Initiative, Novartis, Cipla, support groups representing individuals with lived experiences, and first authors of related published articles.

We aim to promote the study and recruit stakeholders by developing a project webpage, featuring in network newsletters, presenting at seminars and conferences, using social media, and distributing leaflets. PIS and ICF will be provided electronically.

Multistage Delphi Survey

For the Delphi survey, a 2-round process will be conducted using the long list of outcomes derived from the identified information sources [13], with rounds separated by a 4-week interval. We will aim to recruit 60 to 80 participants overall, consistent with methodological evidence indicating that this sample size provides high replicability (≥80%) in multistakeholder Delphi surveys [54]. To ensure balanced representation, we will recruit at least 15 participants per stakeholder group. This distribution is intended to preserve meaningful comparison across groups and to maintain robust consensus thresholds in the event of attrition. Reminder emails and flexible completion windows will be used to support retention between rounds. Participants will have 4 weeks to

XSL•FO
RenderX

complete each survey, with the option to save and resume their responses later. A third round may be considered depending on the analysis of the first 2 rounds. Descriptive statistics will be used to analyze overall scores for each stakeholder group, ultimately determining inclusion or exclusion of outcomes.

In round 1, participants will rate the importance of outcomes for inclusion in the COS using a 9-point scale, categorized as 9 to 7 (critically important), 6 to 4 (important but not critical), and 3 to 1 (not important). They can also briefly justify their choices, suggest changes, and propose additional items not covered in the survey. Suggested outcomes from the first round will be reviewed by the study committee for inclusion in the second round. Only responses from participants who rated at least 50% of items will be included.

In round 2, a summary of overall scores and scores from stakeholder groups on each outcome will be provided to participants along with their own scoring. Participants will be aware of others' views but will not know individual scores to maintain anonymity, allowing unbiased expression of opinions. Participants can maintain or adjust their ratings in the second round if desired. Each round will have a 4-week completion time frame.

Consensus Definition

Outcomes that receive scores of 7 to 9 from 70% or more participants and scores of 1 to 3 from less than 15% of participants across all stakeholder groups will be considered for inclusion in the COS during the consensus meetings. Outcomes will be excluded if they are scored between 1 and 3 by at least 70% of participants across all stakeholder groups and between 7 and 9 by less than 15% of participants. Outcomes that do not fall into either category will be classified as having no consensus and will be discussed at the consensus meeting, as applied in other COS studies [55].

Phase 3: Consensus Meetings

All participants who complete the Delphi process will be invited to participate in an online consensus meeting conducted via Zoom. To ensure that discussions reflect a range of perspectives while remaining manageable, each stakeholder group will be represented by a minimum of 2 participants. This approach is intended to balance inclusivity with feasibility and to ensure that all stakeholder voices are adequately represented during the outcome agreement. If an invited participant cannot attend, efforts will be made to find a suitable replacement from the same stakeholder group. To facilitate broad international participation, meetings will be scheduled to accommodate different time zones and participant availability.

Before the meetings, all necessary materials (outcomes rated through the Delphi process and voting patterns from different stakeholder groups) will be circulated to ensure participants are well prepared. To enhance engagement and decision-making, participants will be provided with a secure weblink to cast their votes remotely during the meeting.

The primary objective of the consensus meeting is to finalize the COS through structured discussions involving a diverse international panel. Outcomes identified through the Delphi process will be reviewed alongside voting patterns from different stakeholder groups. Discussions will be chaired by experienced, nonvoting facilitators to ensure neutrality. Following these discussions, participants will engage in anonymous, computerized voting. An outcome will be included in the final COS if it receives at least 70% approval from participants, with representation from all stakeholder groups.

Phase 4: Outcome Measurement Instruments

Following agreement on the final COS, outcome measurement instruments for each outcome will be identified and evaluated in accordance with COSMIN guidelines[56]. This process will ensure that recommended instruments are valid, reliable, and appropriate for use in severe malaria treatment trials.

Systematic literature searches will be conducted to identify existing instruments used to measure the outcomes included in the COS. Sources will include published clinical trials and methodological studies relevant to severe malaria and related conditions. Identified instruments will then be assessed for methodological quality using the COSMIN Risk of Bias checklist, with the evaluation of key measurement properties such as validity, reliability, and responsiveness. Where evidence is insufficient or no appropriate instrument exists, gaps in outcome measurement will be highlighted, with recommendations for future research or instrument development.

Phase 5: Dissemination Strategy

The results of this research will be written as part of a PhD dissertation and publications and will be disseminated at conferences, on the project webpage, featured in network newsletters, presented at seminars, as well as distribution of leaflets to all expert networks associated with the treatment of severe malaria in trials. To enhance accessibility, a plain English summary will be provided in infographics and on social media to facilitate broader engagement. The study protocol and results will be publicly available on the Open Science Framework to facilitate free and widespread sharing of the research. Participants will receive the findings of this research if they consent to it.

Ethical Considerations

At the interview venue, participants were provided with a written ICF (Multimedia Appendix 5) after being guided through the PIS (Multimedia Appendix 6) to ensure they fully understood the study's purpose and procedures. For participants unable to read or write, the PIS and ICFs were read aloud and translated into the local language to facilitate comprehension. Local gatekeepers played a crucial role in assisting participants by printing and signing the forms when necessary.

Participants unable to sign the consent form had the option to provide a thumbprint, ensuring inclusivity and respect for diverse literacy levels. All signed consent forms were securely stored in a locked filing cabinet at the local co-investigator's office, safeguarding participants' confidentiality and adhering to ethical research practices. All data were managed according to the University College London (UCL) Research Ethics Committee and the Komfo Anokye Teaching Hospital Institutional Review Board (KATH-IRB) guidelines.

XSL•FO
RenderX

Ethical approval for the qualitative interview phase of this study was obtained from the UCL Research Ethics Committee (project ID 27289/001) and the KATH-IRB, Ghana (reference KATH IRB/AP/087/24). All interview participants received written study information and provided written informed consent before participation. Data were treated with strict confidentiality, with participants assigned system-generated study identification numbers, and all personally identifiable information anonymized. No names or identifiable details appear in the thesis or resulting publications. Audio recordings and transcripts were securely stored on UCL OneDrive and accessible only to authorized members of the research team, with anonymized data retained for thesis write-up and repository deposit. Interview participants received GH¢50 (US $4.5) airtime to compensate for time, inconvenience, and transport costs.

The subsequent Delphi surveys and consensus meeting phases will be conducted under the same ethical approvals. Participants will be provided with written study information and will give informed consent before participation. Survey participants will be able to pause, save, and return to the survey, and participation will be voluntary with the option to withdraw at any time. Where applicable, names will be collected solely for acknowledgment and email addresses for survey administration and dissemination of study results; participants will be able to opt out of both. All data will be anonymized using study IDs, stored securely, and accessible only to authorized researchers. Email addresses will be retained until completion of the PhD project (anticipated September 30, 2026) and then deleted. No financial compensation is planned for participation in the survey or consensus meeting phases.

Results

Systematic Review Update

Study Selection

Following the update of the systematic review, a total of 326 studies were identified through database and trial registry searches aimed at capturing additional studies. After the initial screening of titles and abstracts, 272 studies were excluded for not meeting the inclusion criteria.

Of the 54 studies selected for full-text review, 5 were deemed eligible and met the predefined inclusion criteria. The inclusion process and exclusion rationale are comprehensively illustrated in Figure 2. This review builds upon a prior systematic review, which included 27 trials from an initial pool of 282 screened studies [10].

Study Characteristics

The included studies, conducted between 2021 and 2024, were geographically distributed across malaria-endemic regions in Africa (Mozambique, Malawi, Zambia, Nigeria, and Uganda) [57-60] and Asia (India) [61]. Participants were predominantly children, with ages ranging from 6 months to 17 years in 4 studies and 1 study in India focusing on adults aged 36 to 45 years. Sample sizes varied from 30 participants in a single-blind trial in India to 668 in a large double-blind trial in Malawi and Zambia [59,61].

All studies were RCTs with varying levels of blinding (single, double, and triple) [58-61] and research phases ranging from phases I or II [57,58] to phase III [60]. The interventions assessed were primarily adjunctive treatments to standard antimalarial therapies [57-59,61], with 1 study focusing exclusively on antimalarial treatment [60]. Primary outcomes varied, including mortality [60,61], pediatric cerebral malaria [59], acute kidney injury [57], and biomarkers such as angiopoietin-2 [58] (Table 1).

XSL•FO
RenderX

Table . Characteristics and results of sources of evidence.

Author	Year	Country	Participants	Sample size	Study design	Intervention type	Primary outcome
Varo et al [58]	2023	Mozambique	Children (aged 1 - 12 y)	180	Phase IIb parallel arm, equally randomized, placebo-controlled, double-blind RCT[a]	Adjunctive treatment	Rate of decline in angiopoietin 2 from admission levels in children with SM[b]
Chilombe et al [59]	2022	Malawi, Zambia	Children (aged 2 - 11 y)	668	Double-blinded, placebo-controlled, 2-armed RCT	Adjunctive treatment	Reduce T_{max} in pediatric CNS[c] malaria and evaluate the "proof-of-concept" for potential neuroprotective efficacy
Alao et al [60]	2021	Nigeria	Children (aged 6 mo to 17 y)	235	Phase III triple-blinded, parallel RCT	Antimalarial	Number of deaths in 12, 24, and 48 h after randomization
Tiwari et al [61]	2023	India	Adults (aged 36-45 y)	30	Single-blind clinical trial	Adjunctive treatment	Death or discharge
Paasi et al [57]	2022	Uganda	Children (aged >6 mo to <12 y)	40	Phase I or II unblinded RCT	Adjunctive treatment	Correction of AKI[d] at 48 h

[a]RCT: randomized controlled trial.

[b]SM: severe malaria.

[c]CNS: central nervous system.

[d]AKI: acute kidney injury.

Qualitative Interview

The study involved 26 participants, the majority of whom were female (n=20, 76.9%), aged 21 to 70 years. The largest age group was 40 to 49 years (n=9, 34.6%), followed by 30 to 39 years (n=8, 30.8%).

Educational backgrounds varied: 50% (n=13) completed junior high school, and 7.7% (n=2) had no formal education.

Most participants were married (n=15, 57.7%) and mothers (n=16, 61.5%). Occupations included trading (n=16, 61.5%), skilled work (n=5, 19.2%), farming (n=3, 11.5%), professionals, and retired (n=1, 3.8% each). The patients were mostly male (n=16, 61.5%), aged 2 to 14 years, and older than 5 years (n=17, 65.4%), with the majority in primary school (n=14, 53.8%).

Discussion

This study is expected to address a critical gap in malaria research by developing the first COS specifically for trials evaluating treatments for severe malaria. At present, there is no established COS for severe malaria treatment trials, and outcomes reported across studies vary widely in definition, measurement, and prioritization. The anticipated outcome of this work is a stakeholder-informed COS that provides a structured framework for the consistent selection and reporting of outcomes in future trials.

By standardizing outcomes and promoting the use of clearly defined and consistently measured end points, the proposed COS has the potential to improve comparability across severe malaria trials and strengthen the quality and transparency of evidence generation in this field. This standardization is expected to reduce outcome reporting bias and research waste while facilitating meaningful synthesis of findings across studies, including meta-analyses [62-64]. In the longer term, a COS for severe malaria treatment may support the selection of a single valid and reliable measurement instrument for each core outcome, thereby enhancing the interpretability and clinical usefulness of trial results within the broader malaria research landscape.

The methodological approach underpinning this protocol reflects current best practice in COS development by the COMET guidelines and in previous studies [11,16,65,66].

By integrating evidence from systematic reviews, qualitative research with parents and caregivers, and consensus methods involving multiple stakeholder groups, the study prioritizes inclusivity and relevance [10,63,67-70]. In particular, the explicit incorporation of parents and caregiver perspectives is expected to ensure that the resulting COS captures outcomes that reflect lived experience, functional impact, and quality of life, alongside traditional clinical endpoints [23,30,71,72]. This approach aligns with contemporary efforts to make clinical research more patient centered and responsive to stakeholder priorities.

XSL•FO
RenderX

Several limitations are acknowledged at the protocol stage. The qualitative component is being conducted within a single malaria-endemic country, which may limit the diversity of perspectives captured and may not fully reflect the experiences and priorities of patients and caregivers in other endemic regions. Furthermore, due to resource constraints, the Delphi survey and consensus processes will be conducted in English only, which may restrict participation from some stakeholders and reduce inclusivity across non–English-speaking settings. The findings will depend on the successful recruitment and sustained engagement of diverse stakeholder groups throughout the Delphi and consensus stages. As with all consensus-based methodologies, the final COS will represent areas of agreement among participants and may not capture all context-specific priorities relevant to every setting.

Future work following completion of the COS will focus on dissemination and implementation. Planned dissemination includes publication in peer-reviewed journals, presentation at relevant conferences, and sharing findings with study participants and the wider malaria research community. Uptake of COS in future trials will be critical to realizing its intended benefits, and engagement with trialists, funders, and guideline developers will be important to support adoption.

Acknowledgments

We would like to express our sincere gratitude to Professor Daniel Ansong and the SMAART Consortium Team in Ghana for their invaluable assistance with the Komfo Anokye Teaching Hospital Institutional Review Board ethics application process.

Funding

The study is funded by the United Kingdom Medical Research Council as part of GDA's research degree (under the core support to the Medical Research Council Clinical Trials Unit at University College London; MC_UU_00004/05). Funding for ECG, SBL, and DMG was provided under the same funding to Medical Research Council Clinical Trials Unit at University College London. Additional support was received from Wellcome Collaborative Grant in Science (Severe Malaria Africa: A Research and Trials Consortium): grant 209265/Z/17/Z.

Authors' Contributions

GDA serves as the chief investigator, having led the proposal and protocol development. SBL contributed methodological expertise to the protocol's development. ECG, the primary supervisor, conceived the study, contributed to its design, and developed and drafted the study proposal and protocol. All authors were involved in further conceptualizing and designing the study and have approved the submitted version of the manuscript. All authors agree to be personally accountable for their contributions and to ensure that any questions regarding the accuracy or integrity of any part of the work, even those they were not directly involved in, are appropriately investigated, resolved, and documented in the literature. GDA, SBL, DMG, KM, and ECG contributed to conceptualization. GDA, SBL, and ECG contributed to methodology. GDA, SBL, DMG, KM, and ECG contributed to resources. GDA, SBL, and ECG contributed to original draft preparation. GDA, SBL, DMG, KM, and ECG contributed to reviewing and editing the article. SBL, DMG, KM, and ECG contributed to supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy for MEDLINE and LILACS.
[DOCX File, 19 KB - resprot_v15i1e78616_app1.docx]

Multimedia Appendix 2
Data chart.
[DOCX File, 15 KB - resprot_v15i1e78616_app2.docx]

Multimedia Appendix 3
Interview (topic) guide.
[DOCX File, 18 KB - resprot_v15i1e78616_app3.docx]

Multimedia Appendix 4
Distress protocol for interviews.
[DOCX File, 15 KB - resprot_v15i1e78616_app4.docx]

Multimedia Appendix 5

XSL•FO
RenderX

Informed consent form.
[DOCX File, 19 KB - resprot_v15i1e78616_app5.docx]

Multimedia Appendix 6
Participant information leaflet.
[DOCX File, 18 KB - resprot_v15i1e78616_app6.docx]

Checklist 1
PRISMA checklist.
[PDF File, 1450 KB - resprot_v15i1e78616_app7.pdf]

Checklist 2
The COS-STAP statement checklist.
[DOCX File, 19 KB - resprot_v15i1e78616_app8.docx]

Checklist 3
COREQ checklist.
[PDF File, 1147 KB - resprot_v15i1e78616_app9.pdf]

References

1. Sarfo JO, Amoadu M, Kordorwu PY, et al. Malaria amongst children under five in sub-Saharan Africa: a scoping review of prevalence, risk factors and preventive interventions. Eur J Med Res 2023 Feb 17;28(1):80. [doi: 10.1186/s40001-023-01046-1] [Medline: 36800986]
2. Fact sheet about malaria. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/malaria [accessed 2023-12-06]
3. Malaria's impact worldwide. CDC. 2021. URL: https://www.cdc.gov/malaria/php/impact/?CDC_AAref_Val=https://www.cdc.gov/malaria/malaria_worldwide/impact.html [accessed 2023-12-06]
4. Malaria in Africa. UNICEF DATA. URL: https://data.unicef.org/topic/child-health/malaria/ [accessed 2023-12-06]
5. Cardona-Arias JA, Carmona-Fonseca J. Frequency of gestational malaria and maternal-neonatal outcomes, in Northwestern Colombia 2009-2020. Sci Rep 2022 Jun 29;12(1):10944. [doi: 10.1038/s41598-022-15011-1] [Medline: 35768619]
6. Hartman TK, Rogerson SJ, Fischer PR. The impact of maternal malaria on newborns. Ann Trop Paediatr 2010;30(4):271-282. [doi: 10.1179/146532810X12858955921032] [Medline: 21118620]
7. Eisele TP, Larsen DA, Anglewicz PA, et al. Malaria prevention in pregnancy, birthweight, and neonatal mortality: a meta-analysis of 32 national cross-sectional datasets in Africa. Lancet Infect Dis 2012 Dec;12(12):942-949. [doi: 10.1016/S1473-3099(12)70222-0] [Medline: 22995852]
8. WHO guidelines for malaria. World Health Organization. 2023. URL: https://app.magicapp.org/#/guideline/7661 [accessed 2023-12-06]
9. FDA approves only drug in US to treat severe malaria. FDA. URL: https://www.fda.gov/news-events/press-announcements/fda-approves-only-drug-us-treat-severe-malaria [accessed 2023-12-06]
10. Lampro L, George EC. Outcomes reported in trials of treatments for severe malaria: the need for a core outcome set. Trop Med Int Health 2022 Sep;27(9):767-775. [doi: 10.1111/tmi.13803] [Medline: 35916146]
11. Core outcome measures in effectiveness trials. COMET Initiative. URL: https://www.comet-initiative.org/ [accessed 2024-02-05]
12. Moorthy V, Reed Z, Smith PG, WHO Study Group on Measures of Malaria Vaccine Efficacy. Measurement of malaria vaccine efficacy in phase III trials: report of a WHO consultation. Vaccine (Auckl) 2007 Jul 9;25(28):5115-5123. [doi: 10.1016/j.vaccine.2007.01.085] [Medline: 17577487]
13. Williamson PR, Altman DG, Bagley H, et al. The COMET handbook: version 1.0. Trials 2017 Jun;18(S3):280. [doi: 10.1186/s13063-017-1978-4]
14. Kirkham JJ, Davis K, Altman DG, et al. Core Outcome Set-STAndards for Development: the COS-STAD recommendations. PLoS Med 2017 Nov;14(11):e1002447. [doi: 10.1371/journal.pmed.1002447] [Medline: 29145404]
15. Millward CP, Armstrong TS, Barrington H, et al. Opportunities and challenges for the development of "core outcome sets" in neuro-oncology. Neuro Oncol 2022 Jul 1;24(7):1048-1055. [doi: 10.1093/neuonc/noac062] [Medline: 35287168]
16. Kirkham JJ, Gorst S, Altman DG, et al. Core Outcome Set-STAndardised Protocol Items: the COS-STAP Statement. Trials 2019 Feb 11;20(1):116. [doi: 10.1186/s13063-019-3230-x] [Medline: 30744706]
17. Developing a core outcome set for trials in the treatment of severe malaria. COMET INITIATIVE. URL: https://www.comet-initiative.org/Studies/Details/2908 [accessed 2026-02-03]

XSL•FO
RenderX

18. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 2009 Jul 21;339:b2700. [doi: 10.1136/bmj.b2700] [Medline: 19622552]
19. Severe malaria. Tropical Med Int Health 2014 Sep;19(s1):7-131. [doi: 10.1111/tmi.12313_2]
20. Choong MK, Tsafnat G. Role of citation tracking in updating of systematic reviews. AMIA Jt Summits Transl Sci Proc 2014;2014:18. [Medline: 25954571]
21. Tucker C. Research guides: CMGT 540: uses of communication research: citation tracking. USC Libraries. URL: https://libguides.usc.edu/c.php?g=235146&p=1561960 [accessed 2024-11-29]
22. Quirke F, Ariff S, Battin M, et al. Core outcomes in neonatal encephalopathy: a qualitative study with parents. BMJ Paediatr Open 2022 Jul;6(1):e001550. [doi: 10.1136/bmjpo-2022-001550] [Medline: 36053648]
23. Keeley T, Williamson P, Callery P, et al. The use of qualitative methods to inform Delphi surveys in core outcome set development. Trials 2016 May 4;17(1):230. [doi: 10.1186/s13063-016-1356-7] [Medline: 27142835]
24. Tong A, Sainsbury P, Craig J. Consolidated Criteria for Reporting Qualitative Research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
25. Jones JE, Jones LL, Keeley TJH, Calvert MJ, Mathers J. A review of patient and carer participation and the use of qualitative research in the development of core outcome sets. PLOS ONE 2017 Mar 16;12(3):e0172937. [doi: 10.1371/journal.pone.0172937]
26. Ayton D. Chapter 5: qualitative descriptive research. Open Educational Resources Collective. 2023. URL: https://oercollective.caul.edu.au/qualitative-research/chapter/__unknown__-5/ [accessed 2024-11-20]
27. Doyle L, McCabe C, Keogh B, Brady A, McCann M. An overview of the qualitative descriptive design within nursing research. J Res Nurs 2020 Aug;25(5):443-455. [doi: 10.1177/1744987119880234]
28. Kim H, Sefcik JS, Bradway C. Characteristics of qualitative descriptive studies: a systematic review. Res Nurs Health 2017 Feb;40(1):23-42. [doi: 10.1002/nur.21768]
29. Remus A, Smith V, Wuytack F. Methodology in core outcome set (COS) development: the impact of patient interviews and using a 5-point versus a 9-point Delphi rating scale on core outcome selection in a COS development study. BMC Med Res Methodol 2021 Jan 7;21(1):10. [doi: 10.1186/s12874-020-01197-3] [Medline: 33413129]
30. Hamzeh H, Kelly C, Spencer S. Outcomes of physiotherapy for people living with bronchiectasis: qualitative study to inform development of a core outcome set. Physiotherapy 2023 Dec;121:37-45. [doi: 10.1016/j.physio.2023.07.001] [Medline: 37812851]
31. Shepherd V, Wood F, Robling M, Randell E, Hood K. Development of a core outcome set for the evaluation of interventions to enhance trial participation decisions on behalf of adults who lack capacity to consent: a mixed methods study (COnSiDER Study). Trials 2021 Dec 19;22(1):935. [doi: 10.1186/s13063-021-05883-5] [Medline: 34924004]
32. Cleland JA. The qualitative orientation in medical education research. Korean J Med Educ 2017 Jun;29(2):61-71. [doi: 10.3946/kjme.2017.53] [Medline: 28597869]
33. Snow RW, Omumbo JA. In: Jamison DT, Feachem RG, Makgoba MW, Bos ER, Baingana FK, Hofman KJ, editors. Disease and Mortality in Sub-Saharan Africa, 2nd edition: The International Bank for Reconstruction and Development/The World Bank; 2006.
34. Sempungu JK, Choi M, Lee EH, Lee YH. The burden of malaria in sub-Saharan Africa and its association with development assistance for health and governance: an exploration of the GBD 2019 study. Glob Public Health 2023 Jan;18(1):2229892. [doi: 10.1080/17441692.2023.2229892] [Medline: 37438859]
35. Papaioannou I, Utzinger J, Vounatsou P. Malaria-anemia comorbidity prevalence as a measure of malaria-related deaths in sub-Saharan Africa. Sci Rep 2019 Aug 5;9(1):11323. [doi: 10.1038/s41598-019-47614-6]
36. Malaria prevalence in children. Ghana Statistical Services. URL: https://web.archive.org/web/20241002132206/https://statsghana.gov.gh/gssmain/fileUpload/pressrelease/World%20Malaria%20Day%20Press%20Release%20from%20GSS.pdf [accessed 2024-04-18]
37. Ghana. Severe Malaria Observatory. URL: https://www.severemalaria.org/countries/ghana [accessed 2024-04-19]
38. Ronald LA, Kenny SL, Klinkenberg E, et al. Malaria and anaemia among children in two communities of Kumasi, Ghana: a cross-sectional survey. Malar J 2006 Nov 9;5(1):105. [doi: 10.1186/1475-2875-5-105] [Medline: 17094806]
39. Recall bias. Catalog of Bias. 2017. URL: https://catalogofbias.org/biases/recall-bias/ [accessed 2024-11-20]
40. Definition of recall bias. National Cancer Institute. URL: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/recall-bias [accessed 2024-11-20]
41. Focus group and semi-structured interviews. Queen's University. URL: https://www.queensu.ca/vpr/ethics/guidelines-policies/focus-groups-semi-structured-interviews-guidelines [accessed 2024-11-22]
42. DeJonckheere M, Vaughn LM. Semistructured interviewing in primary care research: a balance of relationship and rigour. Fam Med Com Health 2019 Mar;7(2):e000057. [doi: 10.1136/fmch-2018-000057]
43. Sanatinia R, Crawford MJ, Quirk A, et al. Identifying features associated with higher-quality hospital care and shorter length of admission for people with dementia: a mixed-methods study. Health Serv Deliv Res 2020;8(22):1-92. [doi: 10.3310/hsdr08220]

XSL•FO
RenderX

44. Busetto L, Wick W, Gumbinger C. How to use and assess qualitative research methods. Neurol Res Pract 2020 Dec;2(1):14. [doi: 10.1186/s42466-020-00059-z]
45. Dolan HR, Alvarez AA, Freylersythe SG, Crane TE. Methodology for analyzing qualitative data in multiple languages. Nurs Res 2023;72(5):398-403. [doi: 10.1097/NNR.0000000000000673] [Medline: 37625183]
46. Yunus NA, olde Hartman T, Lucassen P, et al. Reporting of the translation process in qualitative health research: a neglected importance. Int J Qual Methods 2022 Apr;21:16094069221145282. [doi: 10.1177/16094069221145282]
47. Clark L, Birkhead AS, Fernandez C, Egger MJ. A transcription and translation protocol for sensitive cross-cultural team research. Qual Health Res 2017 Oct;27(12):1751-1764. [doi: 10.1177/1049732317726761] [Medline: 28936930]
48. Ask more from your data with the depth and power of NVivo. Lumivero. URL: https://lumivero.com/products/nvivo/ [accessed 2024-11-22]
49. Create case classifications. NVIVO. URL: https://help-nv.qsrinternational.com/12/win/v12.1.115-d3ea61/Content/classifications/create-case-classifications.htm [accessed 2024-11-22]
50. Kiger ME, Varpio L. Thematic analysis of qualitative data: AMEE Guide No. 131. Med Teach 2020 Aug;42(8):846-854. [doi: 10.1080/0142159X.2020.1755030] [Medline: 32356468]
51. Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. J Clin Epidemiol 2018 Apr;96:84-92. [doi: 10.1016/j.jclinepi.2017.12.020] [Medline: 29288712]
52. White NJ. Severe malaria. Malar J 2022 Oct 6;21(1):284. [doi: 10.1186/s12936-022-04301-8] [Medline: 36203155]
53. Nasa P, Jain R, Juneja D. Delphi methodology in healthcare research: how to decide its appropriateness. World J Methodol 2021 Jul 20;11(4):116-129. [doi: 10.5662/wjm.v11.i4.116] [Medline: 34322364]
54. Manyara AM, Purvis A, Ciani O, Collins GS, Taylor RS. Sample size in multistakeholder Delphi surveys: at what minimum sample size do replicability of results stabilize? J Clin Epidemiol 2024 Oct;174:111485. [doi: 10.1016/j.jclinepi.2024.111485] [Medline: 39069013]
55. Ryan JM, Devane D, Simiceva A, et al. Surgical Handover CORE Outcome Measures (SH-CORE): a protocol for the development of a CORE outcome set for trials in surgical handover. Trials 2024 Jun 10;25(1):373. [doi: 10.1186/s13063-024-08201-x] [Medline: 38858749]
56. Guideline for selecting instruments for a core outcome set. COSMIN. URL: https://www.cosmin.nl/tools/guideline-selecting-proms-cos/ [accessed 2026-03-31]
57. Paasi G, Okalebo CB, Ongodia P, et al. PARIST study protocol: a phase I/II randomised, controlled clinical trial to assess the feasibility, safety and effectiveness of paracetamol in resolving acute kidney injury in children with severe malaria. BMJ Open 2023 Jul 31;13(7):e068260. [doi: 10.1136/bmjopen-2022-068260] [Medline: 37524553]
58. Varo R, Crowley VM, Mucasse H, et al. Adjunctive rosiglitazone treatment for severe pediatric malaria: a randomized placebo-controlled trial in Mozambican children. Int J Infect Dis 2024 Feb;139:34-40. [doi: 10.1016/j.ijid.2023.11.031] [Medline: 38013152]
59. Chilombe MB, McDermott MP, Seydel KB, Mathews M, Mwenechanya M, Birbeck GL. Aggressive antipyretics in central nervous system malaria: study protocol of a randomized-controlled trial assessing antipyretic efficacy and parasite clearance effects (Malaria FEVER study). PLoS ONE 2022;17(10):e0268414. [doi: 10.1371/journal.pone.0268414] [Medline: 36206262]
60. Alao MA, Orimadegun AE, Ibrahim OR, et al. Efficacy and safety of DUAL intravenous artesunate plus quinine compared to intravenous artesunate for cerebral malaria in a triple blinded parallel multisite randomized controlled TRIAL in Nigerian children: DUAL PAQ TRIAL Protocol. Trials 2021 Oct 20;22(1):721. [doi: 10.1186/s13063-021-05634-6] [Medline: 34670598]
61. Tiwari S, Kursange S, Goyal A, Safi D. Efficacy of pulse methylprednisolone in treatment of acute respiratory distress syndrome due to malaria: a randomized controlled clinical trial. J Assoc Physicians India 2023 Nov;71(11):36-39. [doi: 10.59556/japi.71.0379] [Medline: 38720494]
62. Williamson PR, Altman DG, Blazeby JM, et al. Developing core outcome sets for clinical trials: issues to consider. Trials 2012 Dec;13(1):132. [doi: 10.1186/1745-6215-13-132]
63. Young AE, Brookes ST, Avery KNL, Davies A, Metcalfe C, Blazeby JM. A systematic review of core outcome set development studies demonstrates difficulties in defining unique outcomes. J Clin Epidemiol 2019 Nov;115:14-24. [doi: 10.1016/j.jclinepi.2019.06.016]
64. Kirkham JJ, Williamson P. Core outcome sets in medical research. bmjmed 2022 Oct;1(1):e000284. [doi: 10.1136/bmjmed-2022-000284]
65. Taneri PE, Kirkham JJ, Molloy EJ, et al. Protocol for the development of a core outcome set for neonatal sepsis (NESCOS). PLOS ONE 2023 Dec 5;18(12):e0295325. [doi: 10.1371/journal.pone.0295325]
66. Quirke FA, Healy P, Bhraonáin EN, et al. COHESION: core outcomes in neonatal encephalopathy (protocol). Trials 2021 Feb 8;22(1):125. [doi: 10.1186/s13063-021-05030-0] [Medline: 33557892]
67. Rodrigues M, Syed Z, Dufort A, et al. Heterogeneity across outcomes reported in clinical trials for older adults with depression: a systematic survey. J Clin Epidemiol 2023 May;157:59-73. [doi: 10.1016/j.jclinepi.2023.03.002]

XSL•FO
RenderX

68. Manera KE, Johnson DW, Cho Y, et al. Scope and heterogeneity of outcomes reported in randomized trials in patients receiving peritoneal dialysis. Clin Kidney J 2021 Jul;14(7):1817-1825. [doi: 10.1093/ckj/sfaa224] [Medline: 34221389]
69. Clarke M, Williamson PR. Core outcome sets and systematic reviews. Syst Rev 2016 Dec;5(1):11. [doi: 10.1186/s13643-016-0188-6]
70. Saldanha IJ, Dodd S, Gorst SL, Williamson PR. More than half of systematic reviews have relevant core outcome sets. J Clin Epidemiol 2021 Aug;136:168-179. [doi: 10.1016/j.jclinepi.2021.04.019]
71. Connolly B, Milton-Cole R, Blackwood B, Pattison N. Using patient and care partner experiences to confirm outcomes of relevance for inclusion in a core outcome set for trials of physical rehabilitation in critical illness: a qualitative interview study. Aust Crit Care 2024 Nov;37(6):912-923. [doi: 10.1016/j.aucc.2024.05.005] [Medline: 38845285]
72. Cao H, Chen Y, Yang Z, et al. The methodological quality of systematic reviews regarding the Core Outcome Set (COS) development. BMC Med Res Methodol 2024 Mar 11;24(1):65. [doi: 10.1186/s12874-024-02182-w] [Medline: 38468223]

Abbreviations

CENTRAL: Cochrane Central Register of Controlled Trials
COMET: Core Outcome Measures in Effectiveness Trials
COREQ: Consolidated Criteria for Reporting Qualitative Research
COS: core outcome set
COS-STAP: Core Outcome SetStandardized Protocol Items
COSSMaT: core outcome sets for severe malaria trials
ICF: informed consent form
ISRCTN: International Standard Randomized Controlled Trial Number
KATH-IRB: Komfo Anokye Teaching Hospital Institutional Review Board
LILACS: Literatura Latino-Americana em Ciências da Saúde
PACTR: Pan-African Clinical Trial Registry
PIS: participant information leaflet
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SMAART: Severe Malaria Africa - A Consortium for Research and Trials
UCL: University College London
WHO: World Health Organization

Edited by A Schwartz; submitted 05.Jun.2025; peer-reviewed by H Anandan, M Munikumar; revised version received 30.Jan.2026; accepted 17.Feb.2026; published 13.Apr.2026.

Please cite as:
Asamoah GD, Love SB, Gibb DM, Maitland K, George EC
Synthesis and Standardization of Outcomes in Severe Malaria Treatment Trials: Protocol for the Development of a Core Outcome Set (the COSSMaT Study)
JMIR Res Protoc 2026;15:e78616
URL: https://www.researchprotocols.org/2026/1/e78616
doi:10.2196/78616

© Gideon Darko Asamoah, Sharon B Love, Diana M Gibb, Kathryn Maitland, Elizabeth C George. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Integrating Gait Analysis and AI Into Knee Osteoarthritis Care: Protocol for a 3-Phase Participatory Qualitative Study

Owen Ryan Lindsay, MSc; Janie L Astephen Wilson, MSc, PhD

School of Biomedical Engineering, Department of Surgery, Faculties of Engineering and Medicine, Dalhousie University, Dentistry Bldg, 5th Fl., 5981 University Avenue, Halifax, NS, Canada

Corresponding Author:
Owen Ryan Lindsay, MSc
School of Biomedical Engineering, Department of Surgery, Faculties of Engineering and Medicine, Dalhousie University, Dentistry Bldg, 5th Fl., 5981 University Avenue, Halifax, NS, Canada

Abstract

Background: Knee osteoarthritis (OA) leads to pain, disability, and reduced quality of life. For advanced stages, knee arthroplasty surgery is the standard treatment, yet dissatisfaction and persistent mobility deficits remain common. Current surgical decision-making processes seldom incorporate objective predictors of outcomes, such as biomechanical data. Advances in computer vision, wearable sensors, and artificial intelligence now enable efficient capture and interpretation of clinically prognostic gait features in real-world settings. However, the clinical adoption of such innovations remains limited, hindered by usability challenges, misalignment with stakeholder needs, and system-level barriers.

Objective: This study aims to inform the development of a knee OA clinical decision support (CDS) tool that integrates gait analysis and artificial intelligence with clinical decision-making. More broadly, it seeks to advance a framework for participatory digital health CDS tool implementation by identifying and addressing suboptimal workflows, stakeholder priorities, and organizational challenges.

Methods: This study uses a 3-phase participatory design process, with in-depth interviews conducted in each phase to generate insights that inform subsequent stages. Phase 1 investigates perceived barriers to and facilitators of the implementation of a digital CDS tool and examines current workflow processes in knee OA management. Phase 2 focuses on identifying and defining user requirements to guide solution ideation for translational digital decision support. Phase 3 assesses user acceptance of prototyped solutions. Field observations in phase 1 and usability testing in phase 3 will supplement interview findings with real-world evidence.

Results: The study was funded in November 2024. Ethics approval was granted by the Nova Scotia Health Research Ethics Board on July 28, 2025. Data collection began in September 2025. As of February 2026, a total of 6 clinicians have been interviewed, with all 6 interviews transcribed and 4 coded as part of phase 1 analysis. Study completion is anticipated by August 2026. Findings will be published by January 2027.

Conclusions: Our stakeholder-driven approach prioritizes both technological rigor and practical usability, addressing key human factors for successful implementation. While initially focused on Nova Scotia's orthopedic setting, this research provides a scalable framework for integrating digital decision support tools into broader clinical environments, advancing innovation in knee OA care and beyond.

International Registered Report Identifier (IRRID): PRR1-10.2196/82860

(JMIR Res Protoc 2026;15:e82860) doi:10.2196/82860

KEYWORDS

knee osteoarthritis; arthroplasty; clinical decision support; artificial intelligence; AI; qualitative research; gait analysis; patient participation; interview; user-centered design; implementation

Introduction

Background

Knee osteoarthritis (OA) is a progressive and debilitating disease characterized by pain, disability, and poor health-related quality of life, with its incidence on the rise. For advanced stages, knee arthroplasty surgery is the standard treatment, intended to relieve symptoms and restore function and joint mobility. Despite the common use of arthroplasty, the spectrum of relevant and multifaceted [1] knee arthroplasty surgery outcomes is not universally positive [1-5]. Approximately 1 in 5 patients is not satisfied after surgery [5], and upward of 70% experience continued walking deficits years after surgery [2]. These mixed results reflect the fact that treatment decisions are often based on limited indicators, such as pain severity or radiographic

XSL•FO
RenderX

changes, rather than a multidimensional understanding of the patient. Evidence shows that combinations of factors, including pain, gait function, joint structures, and patient expectations, better predict recovery [6,7], yet these predictors are rarely integrated systematically into clinical workflows.

Objective measures of function, particularly gait analysis outcomes, provide critical insights into OA progression and surgical response [8,9]. Deficits in knee joint mechanics are strongly linked to patient-reported pain, functional outcomes [7], and even implant migration and early failure [10-12] following knee arthroplasty surgery. However, traditional gait analysis has required costly, complex laboratory systems, limiting uptake in clinical practice. Recent advances in computer vision, wearable sensors, and machine learning now make it feasible to capture prognostic gait features more readily in real-world clinical environments [13-16]. Broad investment in advanced computing and artificial intelligence (AI) augments the potential of such systems, enabling the integration of biomechanical biomarkers with patient-reported outcome measures, activity levels, joint morphometrics, and demographics for clinically meaningful, individualized risk predictions and decision support [17].

A major gap in translating these technological advances into practice lies in implementation. Despite their promise, clinical decision support (CDS) tools of this nature seldom progress beyond the research stage due to poor usability, lack of alignment with clinician and patient needs, and organizational constraints that hinder adoption [18,19]. These challenges highlight the importance of early, systematic engagement with stakeholders to ensure that CDS tools are designed with contextual, organizational, and human factors in mind.

This protocol describes a participatory qualitative study to inform the development of a multidimensional knee OA decision support tool. Guided by implementation science [20] and design thinking principles, the study will iteratively investigate clinical workflows (phase 1), specify stakeholder values and user requirements (phase 2), and test early prototypes for usability and acceptance (phase 3). By situating technical innovation within a stakeholder-driven design process, this study aims to inform how to bridge the persistent gap between digital health potential and real-world clinical adoption for knee OA management.

Objectives

Our long-term goal is to develop an objective, evidence-driven decision support tool that promotes consistent, enhanced patient outcomes both during the surgical wait period and after knee arthroplasty surgery. This protocol outlines the qualitative methods guiding the participatory design and development of a CDS tool prototype, which will form the foundation for future research into clinical implementation.

Phase 1: Contextual Inquiry

The aim of phase 1 is to understand and map workflows comprising surgical evaluations, follow-up visits, and decision-making processes for knee arthroplasty surgery candidates; assess how clinicians communicate risks to patients and determine appropriateness for surgery; elucidate potential challenges or barriers to integrating gait analysis and AI-enabled CDS tools into routine clinical practice; and identify relevant stakeholders, including interdisciplinary clinical team members, patients, and decision-makers, who may further inform the development and implementation of the tool.

Phase 2: Value Specification

The aim of phase 2 is to identify and characterize touchpoints within clinical workflows and across the perioperative period where a digital support tool would provide optimal value to users; refine features using stakeholder insights, ensuring that multidimensional patient data (eg, gait biomarkers, patient-reported outcome measures, activity levels, morphometrics, and demographics) are translated into actionable, transparent risk predictions; and assess the role of CDS and AI on patient-clinician communication and decision-making, including users' willingness to engage with a tool and its specific components.

Phase 3: Design

The aim of phase 3 is to design a functioning CDS system prototype; assess the tool's intuitive usability by evaluating interface navigation, data visualization, and integration into clinical workflows; and identify user-perceived benefits, barriers, and areas for improvement, thereby informing translational and iterative refinements to tool development.

Methods

Study Design

Drawing from literature on technology adoption and sustainability in health care [21,22], this protocol deliberately moves beyond the typical solution-driven approach, whereby system intelligence and technical capacity are the primary focus of design. Instead, it grounds these components in human experiences, offering a more empathetic approach that balances system capabilities with diverse stakeholder perspectives. In the context of implementation, where stakeholder perspectives must span many levels of health care, it is necessary that methodologies be designed to capture, organize, and integrate inputs from multidisciplinary team members and end users with varying roles and levels of expertise.

To accomplish this, our protocol follows a holistic framework guided by the Centre for eHealth and Wellbeing Research (CeHRes) Roadmap [20] and enriched by principles of design thinking [23]. Our methodologies expand on the CeHRes Roadmap by incorporating complementary theoretical constructs known to influence implementation success, scalability, and sustainability, including organizational readiness, health literacy, and contextual resilience [21]. This protocol is grounded in three core principles derived from these frameworks: (1) active engagement and oversight by a multidisciplinary team of key opinion leaders, clinicians, designers, and engineers; (2) iterative evaluation cycles with user-driven feedback loops to ensure that tool functionality reflects the lived experiences and expertise of unique stakeholder groups; and (3) ongoing consideration of adoption and scalability across all stages of research to support future uptake.

XSL•FO
RenderX

Setting and Participants

In phase 1, we will recruit orthopedic surgeons responsible for evaluating patient appropriateness for knee arthroplasty surgery. These clinicians will participate in individual interviews conducted either at their clinic site or remotely for convenience. They may also optionally participate in field observations of clinical consultations, which will be conducted at their clinic site. A purposive sampling strategy will be used to recruit participants across a range of experience levels, from surgical fellows to senior clinicians. Clinicians and fellows who are not directly involved in perioperative care or decision-making for knee OA will be excluded.

Building on phase 1 outcomes, phase 2 will use respondent-driven sampling to re-engage previous participants and expand recruitment to additional stakeholders whose perspectives are relevant to tool development. Eligible participants will include allied health professionals, health care administrators, and policy-level decision-makers involved in clinical workflow design or technology adoption, whereas industry representatives or individuals not directly connected to clinical implementation will be excluded. Additionally, patients with clinical diagnoses of knee OA who are being evaluated or scheduled for knee arthroplasty surgery will be recruited during OA clinic visits, and those with prior knee arthroplasty surgery may also be included if they are awaiting revision surgery or undergoing surgery on the contralateral limb. Participants from phase 2 will be invited to continue into phase 3, with supplementary recruitment as necessary. To minimize travel burdens, interviews and usability testing in these phases will be offered both in person at participants' workplaces and virtually using videoconferencing software when appropriate.

Ethical Considerations

This study was reviewed and approved by the Nova Scotia Health Research Ethics Board (file 1031703; approval date: July 28, 2025). Eligible patients will first be asked for verbal consent to be contacted by a co–principal investigator during a clinical encounter and, if interested, will complete written informed consent electronically via REDCap (Research Electronic Data Capture; Vanderbilt University) prior to participation. Clinical staff and decision-makers will be recruited through email invitations and will similarly provide written informed consent before participation. All identifying information will be used solely for scheduling and coordination and will be stored securely in accordance with institutional privacy and confidentiality protocols. Study data will be deidentified prior to analysis. No financial compensation or other incentives will be provided to participants for their involvement in this study.

Sample Size

Estimating sample sizes in qualitative research a priori is inherently challenging due to its evolving and emergent nature. For health care research, Hamilton and Finley [22] advise recruitment of 5 to 10 participants in key roles within a single health care setting. However, saturation often extends beyond this range [24-28]. Acknowledging the role of methodologies, researcher expertise, study objectives, and other factors in reaching saturation, Mthuli et al [29] developed the "define, explain, justify, apply" tool, which has been adopted as our guiding framework:

- Define—a nonprobability purposive sampling strategy will be used to capture context-rich insights. Initially, we will recruit orthopedic surgeons responsible for knee arthroplasty surgery decisions. In subsequent phases, respondent-driven sampling will broaden our participant pool to include patients, other care team members, and decision-makers.
- Explain—this sampling strategy aligns with our aim of deeply understanding real-world clinical workflows and stakeholder experiences, ensuring that early phases focus on clinician insights whereas later phases incorporate diverse perspectives.
- Justify—prioritizing depth over breadth is critical in exploratory qualitative research. Focusing first on clinicians secures alignment with clinical practices, and expanding the sample later enhances the tool's relevance across various health care contexts.
- Apply—in phase 1, we aim for 5 to 10 participants to gather in-depth, context-rich insights from key clinical stakeholders. In phase 2, this pool will be expanded to 15 to 30 participants across multiple user groups (ie, clinicians, allied health professionals, administrators, organizational decision-makers, and patients), incorporating a broader range of perspectives. Finally, phase 3 will involve iterative evaluations with 5 to 10 participants per primary user group, who will provide ongoing feedback in response to prototypes of increasing fidelity (estimate of 3 iterations). These sample sizes are consistent with previous research addressing contextual inquiry [25,30], value specification [31-33], and design objectives [34-36].

Research Phases

Overview

Our research uses in-depth qualitative inquiry to investigate the tool's intended environment and user groups (Figure 1). Specifically, we aim to understand perceived barriers and facilitators, current workflow processes (phase 1), user requirement identification and definition (phase 2), and prototype functionality (phase 3). The methods are designed to systematically build on each phase: phase 1 identifies workflow barriers, phase 2 expands on these barriers and reframes them as design opportunities, and phase 3 tests solutions to these barriers through prototypes. Methods include one-on-one interviews, field observations, and think-aloud usability testing, all of which are well suited to achieve these objectives based on the research team's prior experience.

XSL•FO
RenderX

Figure 1. Schematic representation of the 3 phases comprising preliminary clinical decision support tool development.

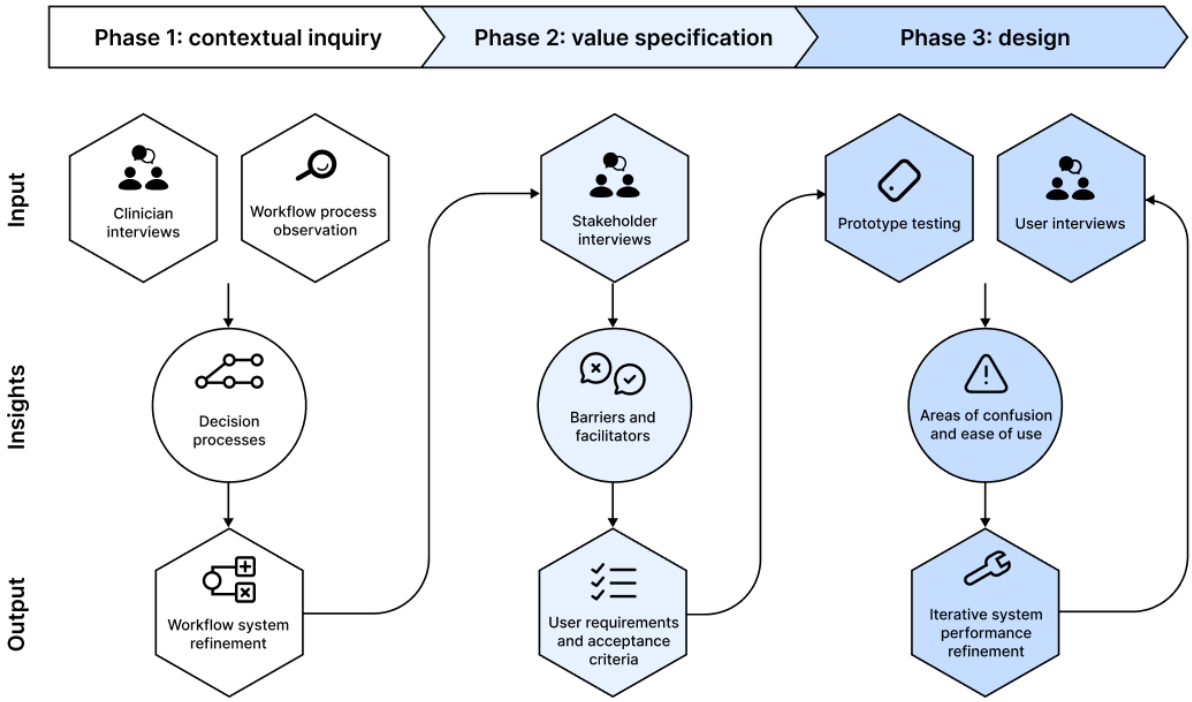

Phase 1: Contextual Inquiry

The first phase of this study involves formative research to understand process workflows throughout the perioperative patient journey and during clinic visits. The CeHRes Roadmap toolkit acknowledges that the very development of health care technologies creates new processes within health care delivery [19]. Understanding these processes is critical as previous studies have shown that technology abandonment may result from disrupted workflows, changes in staff roles, and a low organizational capacity to support technology innovation [18,21]. Semistructured interviews comprising open-ended questions will explore the process for determining appropriateness for knee arthroplasty surgery, including which decision-making resources and information are used to determine appropriateness and timing for surgery (Multimedia Appendix 1). Clinicians will also be given the opportunity to provide insights into perceived anticipated challenges or barriers to the digital CDS concept, including organizational readiness, workflow fit, and capacity to support implementation and sustained use.

To corroborate interview data [37] and ensure that tool development is grounded in a real-world clinical context, participating clinicians may optionally consent to have a research member conduct observations of patient consultations. Prior to each observed consultation, clinicians will inform patients that a researcher will be present to observe clinic workflow and will request verbal consent for the researcher to remain. Patients may decline without impact on their care. Field notes will document only workflow elements (eg, sequence of tasks, decision points, information sources, and physical environment) [38] and will exclude all patient-identifiable information and clinically sensitive details. Both clinicians and patients may request the researcher's departure at any time, and interactions with patients will be limited during observations to avoid disrupting clinical workflows.

Direct observations will allow us to capture unspoken processes, contextual factors, and decision-making dynamics that may not emerge in verbal accounts. Additionally, field notes will serve as a reference to map workflow processes by capturing inputs, outputs, and steps of decision-making as they unfold in real time [38,39]. Where discrepancies between interview-reported and researcher-observed workflows emerge, these will be explored through follow-up member checking to understand underlying reasons (eg, situational variation, informal workarounds, and idealized vs actual practice) as such gaps may reveal implementation barriers and design requirements for workflow variability.

Phase 1 will conclude when contextual understanding of current workflows, decision-making processes, and implementation barriers has reached thematic saturation, yielding validated workflow maps and a barrier framework to inform further inquiry into user values and requirements.

Phase 2: Value Specification

Building on the findings from the contextual inquiry, this phase will further explore user requirements to drive solution ideation. Using an inductive approach, we will analyze themes from phase 1 to identify the high-level barriers and challenges anticipated by stakeholders. These insights will guide a deeper investigation into specific opportunities for development, helping us prioritize user requirements based on their values.

XSL•FO
RenderX

Specifically, we will explore how technologies such as motion capture, AI, system integrations, and supporting devices (eg, video cameras, wearable sensors, desktop computers, smartphones, and tablets) relate to identified barriers. Understanding how users ideally wish to interact with these technologies, as well as their preferences for how the system should generate, translate, and present knowledge for decision-making, is essential for defining system inputs and establishing benchmark usability requirements. Additionally, interviews will explore deployment material needs, including training resources, workflow integration guides, and technical support requirements necessary for successful adoption. Insights from these interviews will be used to align tool development with previously identified barriers, informing both user needs definition and feature acceptance criteria [21,40]. Phase 2 will conclude when consensus on user-defined priority features and deployment needs has been reached, providing the specifications necessary to guide prototype development in phase 3. Because the value specification process builds on previous findings, the interview guide will be developed after completing phase 1 analysis to ensure alignment with emerging themes.

Phase 3: Design

The design cycle focuses on translating identified user needs and operational requirements into a refined prototype and deployment materials. Prototypes will be interactive to support evaluation of user experiences. For usability testing, participants will be provided with training and deployment materials developed based on earlier-phase evidence and asked to complete hypothetical, user-specific scenarios using the CDS system prototype while verbalizing their thought processes. Following prototype interaction, participants will be asked brief follow-up questions to capture overall impressions and suggestions for improvement.

Interactive prototypes will display synthetic, representative patient data to generate simulated patient-specific risk profiles. These outputs will be visualized in case report format in a digital user interface, presenting users with realistic (but synthetic) patient scenarios. It is important to note that this study does not serve to test model accuracy, validate predictive performance, or make effectiveness claims regarding clinical outcomes. Rather, it explores users' behavioral responses to model outputs and visualizations in the context of clinical decision-making, identifying design requirements and implementation barriers prior to formal validation studies.

Usability evidence generated through phase 3 will inform refinement of the CDS tool prototype and deployment protocols, positioning these outputs for subsequent validation studies and implementation planning. Future research will assess the tool's impact on surgical decision-making, patient outcomes, and health care resource use.

Analysis

Interviews and Think-Aloud Testing

Interview sessions and think-aloud testing will be audio recorded and transcribed verbatim, with transcripts verified against recordings by a second research team member for accuracy. In each phase, the 2 researchers (ORL and JLAW) will review and analyze transcripts via NVivo (version 14.24.3; Lumivero) using an inductive thematic approach without prior theoretical influences [41].

The researchers will independently code early transcripts to generate initial inductive codes, bringing their unique disciplinary backgrounds to the interpretation. For each coding decision, the researchers will document their rationale in reflexive memos, creating an audit trail that captures the reasoning behind interpretive choices. Codes and associated memos will be shared at regular meetings to discuss coding decisions, examine how the researchers' perspectives on AI and clinical decision-making may influence interpretation, and explore whether coding discrepancies reflect genuine ambiguity in the data or researcher assumptions. Discrepancies will be treated as opportunities to surface alternative perspectives, refine understanding, and mitigate individual bias.

Through this iterative process, codes will be expanded, collapsed, or modified as new transcripts are analyzed, with ongoing documentation of analytical decisions and their relationship to emerging vs preexisting beliefs about CDS tool implementation. Previously coded transcripts will be revisited as needed when the coding framework evolves. Data collection and analysis will proceed concurrently to allow for the full exploration of emerging themes. This iterative and reflexive approach to thematic analysis aligns with the methodology described by Braun and Clarke [41]. Sampling will continue until inductive thematic saturation is reached, meaning that no new themes or codes are expected to emerge from further data [42].

Analysis of usability data will explore user preferences regarding data interpretability, information hierarchy, and communication approaches. The role of this tool in clinical decision-making will be assessed by how clinicians interpret and integrate prognostic information into clinical reasoning and the perceived utility of this information for triage and treatment decisions. Usability findings will be interpreted in conjunction with contextual evidence gathered in phases 1 and 2, such as participants' roles, organizational constraints (workflow constraints and technical infrastructure), and individual characteristics (health literacy and technical proficiency) to enrich our interpretation of user interactions.

No formal quantitative usability analysis or hypothesis testing will be conducted. Task completion times will be noted during usability testing sessions as contextual information to support the interpretation of qualitative findings.

Process Mapping

Realistic accounts of field observations during clinical consultations will be recorded as they occur (ie, in situ) [38]. A deductive framework will guide the initial classification of workflow components (eg, inputs, decisions, and outputs), whereas inductive coding will capture emergent factors that extend beyond process mapping but are still relevant to implementation. To construct a process map, sequences of events comprising clinical workflows will be charted chronologically, identifying points where a CDS tool could integrate into and support existing workflows.

XSL•FO
RenderX

Quality and Rigor

Ensuring rigor in qualitative research requires strategies aligned with the study's methodological approach and objectives [43]. Accordingly, this study integrates measures to enhance validity, reliability, and rigor across all stages of data collection and analysis.

To enhance data richness, interview guides will first be pilot-tested with participant representatives, including clinicians, patients, and organizational decision-makers, to refine question clarity, relevance, and reliability prior to data collection [22]. Following pilot-testing, the guides will remain flexible and may be iteratively adapted during data collection in response to emerging themes. Insights generated during the contextual inquiry phase will also inform subsequent sampling strategies, ensuring that the study remains responsive to participants and to data not initially anticipated within the tool's scope [44]. This adaptive approach will be combined with recruitment strategies designed to promote continuous participant engagement across compounding phases, enabling continuous checking that the data are contextually grounded.

To enhance trustworthiness, real-time member checking will be used during interviews, wherein participant responses will be summarized or restated to confirm accuracy and encourage clarification or elaboration [43]. Given the lack of standardized guidelines for modifying data following transcription or analyses, full transcripts or completed analyses will not be returned to participants for validation [44].

Triangulation will occur at multiple stages to verify data accuracy. Inductive thematic analysis will include ongoing comparison, with independently generated codes shared between the 2 researchers to identify discrepancies, examine interpretive assumptions, and iteratively refine the analytical framework. Additionally, triangulating interview data with direct observations of clinician workflows will help corroborate the findings and identify discrepancies between reported and observed practices (eg, clinician-reported vs researcher-observed workflows) [45].

Results

This study received funding in November 2024. Recruitment and pilot-testing of study procedures commenced in September 2025. As of February 2026, a total of 6 clinicians have been interviewed, with all 6 interviews transcribed and 4 coded as part of phase 1 analysis. Data collection is expected to continue until August 2026, with findings from each phase feeding into iterative design cycles. The results from this study will be synthesized to inform a framework for integrated CDS technologies in orthopedics and are anticipated to be disseminated in peer-reviewed publications and conference presentations beginning in January 2027.

Discussion

Expected Findings

This study protocol describes a participatory, multiphase qualitative approach to developing a digital CDS tool incorporating gait technologies to support clinical decisions for advanced knee OA guided by the CeHRes Roadmap and design thinking principles. By centering the research process on implementation, our methodologies actively challenge designer bias and prior assumptions, incorporating real-world insights into clinical workflows and stakeholder experiences to validate anticipated barriers and uncover unforeseen challenges to adoption and sustainability. In doing so, our approach extends beyond technical development and directly addresses why CDS tools often fail to achieve adoption or sustainability.

Implementing decision support technologies is an inherently delicate process as they introduce workflows and ideologies that may potentially disrupt established health care ecosystems [46,47]. If executed poorly, this can challenge the longevity of CDS tools, possibly resulting in nonadoption, abandonment, or limited scalability [21]. Long-term viability hinges on tools that effectively recognize and balance diverse personal and operational factors [48], usually through design that acknowledges both typical and edge use cases [49,50]. Our approach to this design challenge is to engage clinicians, patients, health care leaders, and decision-makers from the outset, continuously checking that prototype development meets the distinct goals of these user groups at critical decision points in the development process [51].

Research aimed at advancing health care technologies operates in a challenging space where rapid innovation must be balanced with rigorous, evidence-based methodologies. From a methodological perspective, this study contributes to the broader digital health field by demonstrating how participatory design can be operationalized in the context of CDS tool development. Traditional linear models of health technology design, where design, development, testing, and deployment occur in distinct and sequential stages, provide a logical structure but are poorly suited to the rapid iteration required for digital health innovation [52]. Our 3-phase design introduces structured feedback loops that allow evidence from interviews, observations, and usability testing to directly inform functional requirements and software prototypes. This agile framework combines the rigor of qualitative research with the flexibility required for real-world translation, offering a replicable model for testing other CDS technologies where implementation has proved challenging.

In the context of knee OA, this research also carries important clinical implications. A digital CDS tool that incorporates gait biomarkers has the potential to substantially improve both patient- and system-level outcomes. By providing clinicians and patients with objective, individualized prognostic information, the tool could help optimize surgical timing, reduce inappropriate or premature procedures, and support monitoring in the postoperative period. At the health system level, gait-informed CDS may streamline triage, reduce unnecessary clinic visits, and improve resource allocation while also generating large-scale datasets to advance research on OA progression. Future validation studies and implementation planning will build on the outcomes of this study to assess the tool's impact on surgical decision-making, patient outcomes, and health care resource use.

XSL•FO
RenderX

Limitations

While these methods were chosen to enhance the overall quality of this study, we acknowledge the limitation that this research is focused specifically on an orthopedic setting in Nova Scotia in Canada's health care system, which is a specific ecosystem within a publicly funded, resource-constrained environment. This targeted approach enables a deep, context-specific analysis and a tailored implementation strategy but may limit the generalizability of our findings to other health care environments. However, we view this as a necessary first step in establishing a robust, small-scale implementation model that can later be adapted and scaled to more diverse settings.

Conclusions

This protocol sets out an approach to developing a knee OA decision support tool that is tethered to implementation, combining stakeholder perspectives, participatory design, and iterative refinement. Rather than treating technology development and adoption as separate challenges, our methodologies integrate them into a single process, ensuring that translational tools evolve in step with the needs and realities of their intended users. Although initially situated within an orthopedic context in Nova Scotia, this work represents an important step toward leveraging digital innovation to improve decision-making and patient outcomes in OA care while also offering transferable insights into how complex health technologies can progress from promising concepts to sustainable tools.

Acknowledgments

The authors are grateful to Dr Michael Dunbar for his clinical expertise and role in shaping the phase 1 interviews and for his support in facilitating future recruitment through his clinic.

Funding

This research was funded by the Nova Scotia Health Research Fund in November 2024. The funding organization had no role in the study design, data collection, data analysis, interpretation of results, or the decision to submit this manuscript for publication.

Authors' Contributions

Conceptualization: ORL (lead), JLAW (supporting)
Funding acquisition: JLAW
Methodology: ORL (lead), JLAW (supporting)
Resources: JLAW
Supervision: JLAW
Visualization: ORL
Writing—original draft: ORL
Writing—review and editing: ORL, JLAW (equal)

Conflicts of Interest

None declared.

Multimedia Appendix 1
Phase 1 interview script—clinical integration of gait analysis for decision support.
[DOCX File, 20 KB - resprot_v15i1e82860_app1.docx]

References

1. Singh JA, Dohm M, Sprowson AP, et al. Outcome domains and measures in total joint replacement clinical trials: can we harmonize them? An OMERACT collaborative initiative. J Rheumatol 2015 Dec;42(12):2496-2502. [doi: 10.3899/jrheum.141201] [Medline: 25834208]
2. Astephen Wilson JL, Lamontagne M, Wilson DR, Beaulé PE, Mwale F, Yee A. Patient-specific functional analysis: the key to the next revolution towards the treatment of hip and knee osteoarthritis. J Orthop Res 2019 Aug;37(8):1754-1759. [doi: 10.1002/jor.24317] [Medline: 31042316]
3. Noble PC, Conditt MA, Cook KF, Mathis KB. The John Insall Award: patient expectations affect satisfaction with total knee arthroplasty. Clin Orthop Relat Res 2006 Nov;452:35-43. [doi: 10.1097/01.blo.0000238825.63648.1e] [Medline: 16967035]
4. Shan L, Shan B, Suzuki A, Nouh F, Saxena A. Intermediate and long-term quality of life after total knee replacement: a systematic review and meta-analysis. J Bone Joint Surg Am 2015 Jan 21;97(2):156-168. [doi: 10.2106/JBJS.M.00372] [Medline: 25609443]

XSL•FO
RenderX

5. Bourne RB, Chesworth BM, Davis AM, Mahomed NN, Charron KDJ. Patient satisfaction after total knee arthroplasty: who is satisfied and who is not? Clin Orthop Relat Res 2010 Jan;468(1):57-63. [doi: 10.1007/s11999-009-1119-9] [Medline: 19844772]
6. van der Esch M, Knoop J, van der Leeden M, et al. Clinical phenotypes in patients with knee osteoarthritis: a study in the Amsterdam osteoarthritis cohort. Osteoarthritis Cartilage 2015 Apr;23(4):544-549. [doi: 10.1016/j.joca.2015.01.006] [Medline: 25596322]
7. Young-Shand KL, Dunbar MJ, Astephen Wilson JL. Individual gait features are associated with clinical improvement after total knee arthroplasty. JB JS Open Access 2020;5(2):e0038. [doi: 10.2106/JBJS.OA.19.00038] [Medline: 33123659]
8. Astephen JL, Deluzio KJ, Caldwell GE, Dunbar MJ. Biomechanical changes at the hip, knee, and ankle joints during gait are associated with knee osteoarthritis severity. J Orthop Res 2008 Mar;26(3):332-341. [doi: 10.1002/jor.20496] [Medline: 17960658]
9. Astephen Wilson JL, Deluzio KJ, Dunbar MJ, Caldwell GE, Hubley-Kozey CL. The association between knee joint biomechanics and neuromuscular control and moderate knee osteoarthritis radiographic and pain severity. Osteoarthritis Cartilage 2011 Feb;19(2):186-193. [doi: 10.1016/j.joca.2010.10.020] [Medline: 21074628]
10. Wilson DA, Hubley-Kozey CL, Wilson JL, Dunbar MJ. Pre-operative muscle activation patterns during walking are associated with TKA tibial implant migration. Clin Biomech (Bristol) 2012 Nov;27(9):936-942. [doi: 10.1016/j.clinbiomech.2012.06.012] [Medline: 22858133]
11. Astephen Wilson JL, Wilson DA, Dunbar MJ, Deluzio KJ. Preoperative gait patterns and BMI are associated with tibial component migration. Acta Orthop 2010 Aug;81(4):478-486. [doi: 10.3109/17453674.2010.501741] [Medline: 20809746]
12. Hilding MB, Lanshammar H, Ryd L. Knee joint loading and tibial component loosening. RSA and gait analysis in 45 osteoarthritic patients before and after TKA. J Bone Joint Surg Br 1996 Jan;78(1):66-73. [Medline: 8898130]
13. Kanko RM, Laende E, Selbie WS, Deluzio KJ. Inter-session repeatability of markerless motion capture gait kinematics. J Biomech 2021 May 24;121:110422. [doi: 10.1016/j.jbiomech.2021.110422]
14. Kanko RM, Laende EK, Davis EM, Selbie WS, Deluzio KJ. Concurrent assessment of gait kinematics using marker-based and markerless motion capture. J Biomech 2021 Oct 11;127:110665. [doi: 10.1016/j.jbiomech.2021.110665] [Medline: 34380101]
15. Kanko RM, Laende EK, Strutzenberger G, et al. Assessment of spatiotemporal gait parameters using a deep learning algorithm-based markerless motion capture system. J Biomech 2021 Jun 9;122:110414. [doi: 10.1016/j.jbiomech.2021.110414] [Medline: 33915475]
16. Cotton RJ. KinTwin: imitation learning with torque and muscle driven biomechanical models enables precise replication of able-bodied and impaired movement from markerless motion capture. arXiv. Preprint posted online on May 19, 2025. [doi: 10.48550/arXiv.2505.13436]
17. Hak F, Guimarães T, Santos M. Towards effective clinical decision support systems: a systematic review. PLoS One 2022 Aug 15;17(8):e0272846. [doi: 10.1371/journal.pone.0272846] [Medline: 35969526]
18. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. NPJ Digit Med 2020 Feb 6;3:17. [doi: 10.1038/s41746-020-0221-y] [Medline: 32047862]
19. van Gemert-Pijnen JE, Nijland N, van Limburg M, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011 Dec 5;13(4):e111. [doi: 10.2196/jmir.1672] [Medline: 22155738]
20. Kip H, Beerlage-de Jong N, van Gemert-Pijnen LJ, Kelders SM. The CeHRes Roadmap 2.0: update of a holistic framework for development, implementation, and evaluation of eHealth technologies. J Med Internet Res 2025 Jan 13;27:e59601. [doi: 10.2196/59601] [Medline: 39805104]
21. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 1;19(11):e367. [doi: 10.2196/jmir.8775] [Medline: 29092808]
22. Hamilton AB, Finley EP. Qualitative methods in implementation research: an introduction. Psychiatry Res 2019 Oct;280:112516. [doi: 10.1016/j.psychres.2019.112516] [Medline: 31437661]
23. Krolikowski KA, Bi M, Baggott CM, Khorzad R, Holl JL, Kruser JM. Design thinking to improve healthcare delivery in the intensive care unit: promise, pitfalls, and lessons learned. J Crit Care 2022 Jun;69:153999. [doi: 10.1016/j.jcrc.2022.153999] [Medline: 35104692]
24. Lin D, Papi E, McGregor AH. Exploring the clinical context of adopting an instrumented insole: a qualitative study of clinicians' preferences in England. BMJ Open 2019 Apr 20;9(4):e023656. [doi: 10.1136/bmjopen-2018-023656] [Medline: 31005908]
25. Zheng L, Ohde JW, Overgaard SM, et al. Clinical needs assessment of a machine learning-based asthma management tool: user-centered design approach. JMIR Form Res 2024 Jan 15;8:e45391. [doi: 10.2196/45391] [Medline: 38224482]
26. Gould DJ, Dowsey MM, Glanville-Hearst M, et al. Patients' views on AI for risk prediction in shared decision-making for knee replacement surgery: qualitative interview study. J Med Internet Res 2023 Sep 18;25:e43632. [doi: 10.2196/43632] [Medline: 37721797]

XSL•FO
RenderX

27. Frankel L, Sanmartin C, Hawker G, et al. Perspectives of orthopaedic surgeons on patients' appropriateness for total joint arthroplasty: a qualitative study. J Eval Clin Pract 2016 Apr;22(2):164-170. [doi: 10.1111/jep.12449] [Medline: 26347053]
28. Bunzli S, Nelson E, Scott A, French S, Choong P, Dowsey M. Barriers and facilitators to orthopaedic surgeons' uptake of decision aids for total knee arthroplasty: a qualitative study. BMJ Open 2017 Nov 12;7(11):e018614. [doi: 10.1136/bmjopen-2017-018614] [Medline: 29133333]
29. Mthuli SA, Ruffin F, Singh N. 'Define, explain, justify, apply' (DEJA): An analytic tool for guiding qualitative research sample size. Int J Soc Res Methodol 2022 Nov 2;25(6):809-821. [doi: 10.1080/13645579.2021.1941646]
30. Thomas J, Li J, Scowen C, Eigenstetter A, Georgiou A. Emergency department pathology order support tool (ED POST): a protocol using qualitative inquiry to inform design and development of a prototype to reduce low value care. Health Serv Outcomes Res Method 2024 Sep;24:268-280. [doi: 10.1007/s10742-023-00314-1]
31. Papi E, Murtagh GM, McGregor AH. Wearable technologies in osteoarthritis: a qualitative study of clinicians' preferences. BMJ Open 2016 Jan;6(1):e009544. [doi: 10.1136/bmjopen-2015-009544] [Medline: 26810998]
32. Apell P, Eriksson H. Artificial intelligence (AI) healthcare technology innovations: the current state and challenges from a life science industry perspective. Technol Anal Strateg Manag 2023 Feb 1;35(2):179-193. [doi: 10.1080/09537325.2021.1971188]
33. Pacheco-Brousseau L, Poitras S, Charette M, Amor SB, Desmeules F, Stacey D. Exploring appropriateness criteria for informing the total knee arthroplasty decision-making process: an interpretive descriptive study. J Eval Clin Pract 2023 Sep;29(6):942-954. [doi: 10.1111/jep.13893] [Medline: 37410789]
34. Zheng H, Tulu B, Choi W, Franklin P. Using mHealth app to support treatment decision-making for knee arthritis: patient perspective. EGEMS (Wash DC) 2017 Apr 20;5(2):7. [doi: 10.13063/2327-9214.1284] [Medline: 29930969]
35. Grant S, Tonkin E, Craddock I, et al. Toward enhanced clinical decision support for patients undergoing a hip or knee replacement: focus group and interview study with surgeons. JMIR Perioper Med 2023 Apr 24;6:e36172. [doi: 10.2196/36172] [Medline: 37093626]
36. Stevenson RD, Chowdhury EA, Lobo J, Western MJ, Bilzon JL. Influence of the intelligent knee osteoarthritis lifestyle app (iKOALA) on knee joint pain. BMC Musculoskelet Disord 2024 Jan 24;25(1):90. [doi: 10.1186/s12891-024-07198-3] [Medline: 38263125]
37. Karsh BT. Clinical practice improvement and redesign: how change in workflow can be supported by clinical decision support. Agency for Healthcare Research and Quality. 2009. URL: https://digital.ahrq.gov/key-topics/clinical-decision-support/clinical-practice-improvement-and-redesign-how-change-workflow-can-be-supported-clinical-decision [accessed 2026-03-26]
38. Mulhall A. In the field: notes on observation in qualitative research. J Adv Nurs 2003 Feb;41(3):306-313. [doi: 10.1046/j.1365-2648.2003.02514.x] [Medline: 12581118]
39. Lu AD, Kaul B, Reichert J, Kilbourne AM, Sarmiento KF, Whooley MA. Implementation strategies for frontline healthcare professionals: people, process mapping, and problem solving. J Gen Intern Med 2021 Feb;36(2):506-510. [doi: 10.1007/s11606-020-06169-3] [Medline: 32918200]
40. Park Y, Jackson GP, Foreman MA, Gruen D, Hu J, Das AK. Evaluating artificial intelligence in medicine: phases of clinical research. JAMIA Open 2020 Oct;3(3):326-331. [doi: 10.1093/jamiaopen/ooaa033] [Medline: 33215066]
41. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
42. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907. [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
43. Morse JM. Critical analysis of strategies for determining rigor in qualitative inquiry. Qual Health Res 2015 Sep;25(9):1212-1222. [doi: 10.1177/1049732315588501] [Medline: 26184336]
44. Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. Int J Qual Methods 2002 Jun;1(2):13-22. [doi: 10.1177/160940690200100202]
45. Bunce AE, Gold R, Davis JV, et al. Ethnographic process evaluation in primary care: explaining the complexity of implementation. BMC Health Serv Res 2014 Dec 5;14:607. [doi: 10.1186/s12913-014-0607-0] [Medline: 25475025]
46. Carlisle N, Watson HA, Carter J, et al. Clinicians' experiences of using and implementing a medical mobile phone app (QUiPP V2) designed to predict the risk of preterm birth and aid clinical decision making. BMC Med Inform Decis Mak 2021 Nov 18;21(1):320. [doi: 10.1186/s12911-021-01681-w] [Medline: 34794405]
47. Jacob C, Sanchez-Vazquez A, Ivory C. Clinicians' role in the adoption of an oncology decision support app in Europe and its implications for organizational practices: qualitative case study. JMIR Mhealth Uhealth 2019 May 3;7(5):e13555. [doi: 10.2196/13555] [Medline: 31066710]
48. Wosny M, Strasser LM, Hastings J. Experience of health care professionals using digital tools in the hospital: qualitative systematic review. JMIR Hum Factors 2023 Oct 17;10:e50357. [doi: 10.2196/50357] [Medline: 37847535]
49. Choudhury A, Asan O, Medow JE. Clinicians' perceptions of an artificial intelligence-based blood utilization calculator: qualitative exploratory study. JMIR Hum Factors 2022 Oct 31;9(4):e38411. [doi: 10.2196/38411] [Medline: 36315238]
50. Cronin S, Kane B, Doherty G. A qualitative analysis of the needs and experiences of hospital-based clinicians when accessing medical imaging. J Digit Imaging 2021 Apr;34:385-396. [doi: 10.1007/s10278-021-00446-1]

XSL•FO
RenderX

51. Hogg HD, Al-Zubaidy M, Technology Enhanced Macular Services Study Reference Group, et al. Stakeholder perspectives of clinical artificial intelligence implementation: systematic review of qualitative evidence. J Med Internet Res 2023 Jan 10;25:e39742. [doi: 10.2196/39742] [Medline: 36626192]
52. Ray JM, Ratwani RM, Sinsky CA, et al. Six habits of highly successful health information technology: powerful strategies for design and implementation. J Am Med Inform Assoc 2019 Oct 1;26(10):1109-1114. [doi: 10.1093/jamia/ocz098] [Medline: 31265064]

Abbreviations

AI: artificial intelligence
CDS: clinical decision support
CeHRes: Centre for eHealth and Wellbeing Research
OA: osteoarthritis
REDCap: Research Electronic Data Capture

Edited by J Sarvestan; submitted 09.Sep.2025; peer-reviewed by GD Giebel, N Shah; revised version received 18.Feb.2026; accepted 18.Feb.2026; published 13.Apr.2026.

Please cite as:
Lindsay OR, Astephen Wilson JL
Integrating Gait Analysis and AI Into Knee Osteoarthritis Care: Protocol for a 3-Phase Participatory Qualitative Study
JMIR Res Protoc 2026;15:e82860
URL: https://www.researchprotocols.org/2026/1/e82860
doi:10.2196/82860

© Owen Ryan Lindsay, Janie L Astephen Wilson. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Application of the ADAPT Framework to Contextualize a Participatory Learning and Action Community Intervention for the Prevention and Control of Type 2 Diabetes Mellitus in Urban and Rural Settings in Afghanistan and Pakistan: Protocol for Intervention Adaptation

Amber Tahir[1*], MSc; Noor Sanauddin[2*], PhD; Sara Imtiaz[1], MSBE; Nadia Khaleeq[3], MSc; Mariam AbdeAli[1], MSc; Asima Khan[1], MPH; Farrukh Ahmed[1], MSc; Hannah Maria Jennings[4], PhD; Rubia Zafar[1], MPhil, PGD; Khalid Rehman[3], PhD; Abdul Rahman Shahab[5], MPH; Saima Afaq[6], PhD

[1]Health Promotion Foundation, Karachi, Pakistan

[2]Department of Sociology, University of Peshawar, Peshawar, Pakistan

[3]Institute of Public Health and Social Sciences, Khyber Medical University, Peshawar, Pakistan

[4]Department of Health Sciences, Hull York Medical School, University of York, Heslington, United Kingdom

[5]HealthNet TPO, Afghanistan Country Program, Kabul, Afghanistan

[6]Department of Health Sciences, University of York, Seebohm Rowntree Building, Heslington, York, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Saima Afaq, PhD
Department of Health Sciences, University of York, Seebohm Rowntree Building, Heslington, York, United Kingdom

Abstract

Background: Type 2 diabetes mellitus (T2DM) is a critical global health issue, especially in low- and middle-income countries like Afghanistan and Pakistan, where many cases remain undiagnosed and specialized care is costly. Community Health Participatory Research, which actively involves community members in addressing health issues, is increasingly recognized as an effective approach to deliver sustainable and culturally relevant health solutions. Participatory Learning and Action (PLA), a method of Community Health Participatory Research, was implemented in the D-Magic (Diabetes Mellitus: Action Through Community Groups or mHealth Information for Better Control) trial in Bangladesh that empowered communities to identify challenges and develop locally relevant strategies to prevent and control T2DM.

Objective: This study aims to adapt the PLA-based D-Magic intervention for the prevention and control of T2DM in rural and urban communities of Afghanistan and Pakistan.

Methods: This trial, titled "Engagement of Community Through Participatory Learning and Action for Control and Prevention of Type 2 Diabetes and Its Risk Factors" (EMPOWER-D), will systematically adapt the D-Magic intervention using the ADAPT guidance framework. The study will be conducted across 3 sites: rural Afghanistan (Kabul), rural Pakistan (Peshawar and Swabi), and urban Pakistan (Karachi). Stakeholders and community members will be involved to ensure cultural relevance and authenticity. The process includes identifying barriers and facilitators to the implementation of community-based interventions through a qualitative study and a scoping review, co-designing intervention materials, and piloting for feasibility and effectiveness.

Results: EMPOWER-D was funded and launched in November 2022 by the United Kingdom's National Institute for Health and Care Research (NIHR203248). Following a year-long development phase, the adaptation protocol was finalized, and a mock run of the interviews was conducted. Recruitment and qualitative data collection, including in-depth interviews and focus group discussions, began in July 2024 and have been completed in Peshawar, Swabi, and Karachi. As of January 2025, data collection has commenced in Afghanistan. Results are anticipated to be published in mid-2026 following the completion of data analysis.

Conclusions: The EMPOWER-D adaptation will align the PLA-based intervention with the local sociocultural context, ensuring health equity and participation. Building on the D-Magic trial's success, this adaptation will ensure the intervention's effectiveness and cost-effectiveness in preventing and managing T2DM, not only in Afghanistan and Pakistan but also in other low- and middle-income countries.

Trial Registration: ClinicalTrials.gov NCT06561126: https://clinicaltrials.gov/study/NCT06561126; NCT06570057: https://clinicaltrials.gov/study/NCT06570057; NCT07350694: https://clinicaltrials.gov/study/NCT07350694

XSL•FO
RenderX

International Registered Report Identifier (IRRID): DERR1-10.2196/71602

(JMIR Res Protoc 2026;15:e71602) doi:10.2196/71602

KEYWORDS

implementation; co-design; behavior change; noncommunicable disease; contextual, public health; mixed method; low- and middle-income countries; participatory research; formative study; Southeast Asia

Introduction

Background

Type 2 diabetes mellitus (T2DM) is a major global health issue, with 463 million adults affected globally, 80% of whom are living in low- and middle-income countries (LMICs) [1]. This health crisis is further exacerbated by conflicts, political instability, and natural disasters that disproportionately affect LMICs [2,3] such as Afghanistan and Pakistan. Afghanistan faces significant challenges, with an estimated 11.07% of adults having T2DM and 10.32% having impaired hyperglycemia [4]. Similarly, in Pakistan, the world's fifth most populous country, 26.7% of adults (approximately 33 million people) have T2DM [5]. Additionally, many cases of T2DM in both countries are undiagnosed, worsening the health and economic burden of the condition.

Environmental changes, epidemiological transition, genetic predisposition, and behavioral factors such as diet and physical inactivity contribute to the rising incidence of T2DM in LMICs, in countries such as Afghanistan and Pakistan [6,7]. The disease impacts physical health, imposing significant emotional, social, and economic strains on individuals and communities [8]. T2DM resulted in 75.3 million (95% uncertainty interval 63.5 - 90.2) global disability-adjusted life years in 2021 and was responsible for 2.6% (95% uncertainty interval 2.3% - 2.9%) of total global disability-adjusted life years [9].

Recent advances in health care have suggested addressing T2DM through specialized care, including a multidisciplinary approach to address diabetes-related complications [10]. However, this approach places a continuous and significant burden on individuals due to high out-of-pocket costs [11,12] for frequent medical visits, expensive medications, and diagnostic tests. Furthermore, the focus is usually placed on the management rather than the prevention of the disease, which further adds strain on already stretched health care systems, making diabetes management less accessible for many. In contrast, community-based participatory approaches are becoming increasingly popular because of their effectiveness in improving diabetes outcomes [13]. Moreover, community-based participatory approaches to health address the underlying causes of the disease and help in the prevention as well as the management of the health issue [14].

Community Health Participatory Research (CHPR) is an inclusive approach that actively involves community members in identifying areas for intervention, defining the problem, and generating potential solutions [15]. These approaches prioritize collaboration, equity, and the cocreation of knowledge to produce more relevant and effective interventions [16]. One CHPR approach is Participatory Learning and Action (PLA). PLA was inspired by Paulo Freire's philosophy, where community members actively participate in learning about, defining, and addressing their health issues through facilitated sessions and activities [17]. PLA empowers marginalized communities to identify and solve their problems through a cycle of problem identification, solution planning, implementation, and reflection rather than relying on top-down strategies [18]. PLA has been systematized into a group-based intervention and has proven effective in improving neonatal and child health [19,20]. More recently, this approach was adapted to address diabetes in rural Bangladesh through the D-Magic (Diabetes Mellitus: Action Through Community Groups or mHealth Information for Better Control) trial [21]. The PLA intervention that was tested led to a 48% reduction in the prevalence of T2DM compared to the control group (305/3757, 8% vs 493/3821, 13%, respectively), with an adjusted odds ratio of 0.52 (95% CI 0.38 - 0.71; $P<.001$). Additionally, the intervention increased awareness of T2DM prevention and motivated healthy dietary changes [21]. The D-Magic trial successfully used a participatory approach, organizing communities into small groups that held monthly meetings led by a lay facilitator. These groups followed a 4-phase PLA cycle to prevent and control T2DM, identifying barriers to healthy behavior and developing and implementing context-specific strategies, such as raising awareness, group exercise, income generation, and kitchen gardening [21].

The EMPOWER-D (Engagement of Community Through Participatory Learning and Action for Control and Prevention of Type 2 Diabetes and Its Risk Factors) trial is grounded in the conviction that PLA can empower communities to take ownership of diabetes prevention and control efforts. By organizing communities for collective action, PLA aims to increase awareness and motivation for healthy behaviors, thereby enabling individuals to make informed choices about their health. As a result, individuals are expected to gain confidence and motivation to seek health care services and adopt healthier habits. Ultimately, this approach is anticipated to lead to the development of a self-sustainable model for the prevention and control of diabetes, which can also have a positive impact on the prevention and control of other noncommunicable diseases (NCDs) in the region.

The theory of change logic model for our PLA intervention in EMPOWER-D (Multimedia Appendix 1) illustrates the hypothesized pathways through which participatory group activities are expected to lead to improved diabetes prevention and control in fragile LMIC contexts. It was developed by the research team, based on the findings from the D-Magic process evaluation and the context of Pakistan and Afghanistan. This model underpins our adaptation approach and will guide both implementation and evaluation.

XSL•FO
RenderX

Participatory interventions like PLA are highly context-sensitive, requiring adaptation to achieve sustainable outcomes when implemented in new settings or with different populations [22]. Despite its effectiveness and greater acceptability among the people, evidence on how PLA interventions like D-Magic can be adapted for fragile, crisis-affected, and socially diverse LMICs remains scarce. This is particularly relevant in countries with high T2DM burdens, such as Afghanistan and Pakistan, where direct replication could overlook important contextual distinctions [23,24]. These include variations in health care system capacity, cultural differences, gender-segregated communities, urban-rural disparities, and the added burden of conflict and political instability [25]. In Afghanistan, decades of conflict have overshadowed public health efforts, including addressing NCDs like T2DM, likely resulting in underreported prevalence [26]. In Pakistan, although the situation is relatively stable, the country has faced repercussions due to its proximity to unstable regions and a rapidly changing political environment, which adds to the issues of rapid urbanization and health care disparities, particularly between urban and rural areas, and for women in conservative rural regions [27]. These challenges contribute to rising T2DM rates and a fragmented health care system, with private urban centers offering better care than underfunded public clinics [28].

These differences shape how communities perceive health, engage with interventions, and mobilize collective actions. Thus, this study does not aim for simple transferability of D-Magic but for systematic adaptation that will generate new evidence on how participatory approaches can be tailored for fragile, gender-segregated, and highly diverse LMIC contexts.

Aims and Objectives

This work aims to provide the following:

1. Empirical insights into barriers and facilitators of community-based diabetes prevention in crisis-affected and urbanizing LMICs
2. Methodological advances by combining scoping reviews, qualitative inquiry, and co-design
3. Practical frameworks for incorporating participatory diabetes interventions in health systems where trust, equity, and access are major challenges.

Methods

Ethical Considerations

Ethical approval for this study has been granted by the ethics review committees of the Aga Khan University, Karachi, Pakistan (2024-9340-28927); Khyber Medical University, Peshawar, Pakistan (KMU/IPHSS/Ethics/2023/EO/0136); HealthNetTPO, Kabul, Afghanistan (A-12-24-464); and the National Bioethics Committee (NBC-R-1070). Researchers will approach identified participants and inform them about the research verbally and through a participant information sheet in their local language (Pashto, Dari, and Urdu). Participants will be encouraged to ask about the research before deciding whether they would like to participate. For those who consent to participate and provide a written consent, a time and date for an in-depth interview (IDI) or focus group discussion (FGD) will be arranged. Interviews will be conducted in a comfortable and confidential environment to ensure participants' privacy. Participants will be assigned unique identification codes, and no personal identifiers will be included in transcripts or analysis files. Audio recordings will be stored in password-protected, encrypted folders and deleted after transcription. Deidentified transcripts will be used for analysis. All electronic data will be stored on secure, password-protected devices, and hard copies (eg, consent forms) will be kept in locked cabinets. Only authorized research team members will have access to the data, and findings will be reported in aggregate to protect participant identity. Participants will receive a monetary incentive to cover their travel expenses and compensate them for their time spent participating in the interview.

The full version of the EMPOWER-D trial is registered with ClinicalTrials.gov (NCT07350694 for rural Afghanistan, registered January 20, 2026, NCT06561126 for rural Pakistan, registered August 23, 2024, and NCT06570057 for urban Pakistan, registered August 26, 2024).

Study Overview

This study is part of the research program EMPOWER-D. The EMPOWER-D trial, a project of Centre for Improving Mental and Physical Health Together (C4IMPACT), was launched to specifically address the prevention and control of T2DM in rural and urban communities of Afghanistan and Pakistan.

Conceptual Framework

We will utilize the ADAPT guidance [22] to adapt the intervention systematically. The ADAPT guidance framework emphasizes the importance of involving stakeholders and community members in the adaptation process to ensure that the adapted evidence-based intervention (EBI) is culturally appropriate and contextually relevant.

The framework comprises 5 steps to guide the adaptation process (Figure 1). This study will adapt the D-Magic intervention by carrying out three main activities: (1) identifying the barriers and facilitators in the implementation of community-based interventions (CBIs) and their contextual appropriateness through a qualitative study (using the socioecological model (SEM) [29]) with local stakeholders and a scoping review of the published literature, (2) co-designing the intervention materials, and (3) piloting the adapted materials for feasibility and effectiveness.

The adaptation of intervention materials across the 3 sites (rural Afghanistan, rural Pakistan, and urban Pakistan) will vary according to each context, but the process of adaptation will be the same. The process and findings will be documented using the FRAME (Framework for Reporting Adaptations and Modifications in Evidence-based Interventions) framework [30] (Figure 1), detailing the changes made, the reasons for these changes, and the implementation procedures. The outcomes and implications of the study will be recorded and reported, with recommendations for future research and practice.

XSL•FO
RenderX

Figure 1. Conceptual framework of the process of adaptation in EMPOWER-D (Engagement of Community Through Participatory Learning and Action for Control and Prevention of Type 2 Diabetes and Its Risk Factors). FRAME: Framework for Reporting Adaptations and Modifications in Evidence-Based Interventions; LMICs: low- and middle-income countries; NCDs: noncommunicable diseases; PLA: Participatory Learning and Action; T2DM: type 2 diabetes mellitus; PICO: Population, Intervention, Comparison, Outcome.

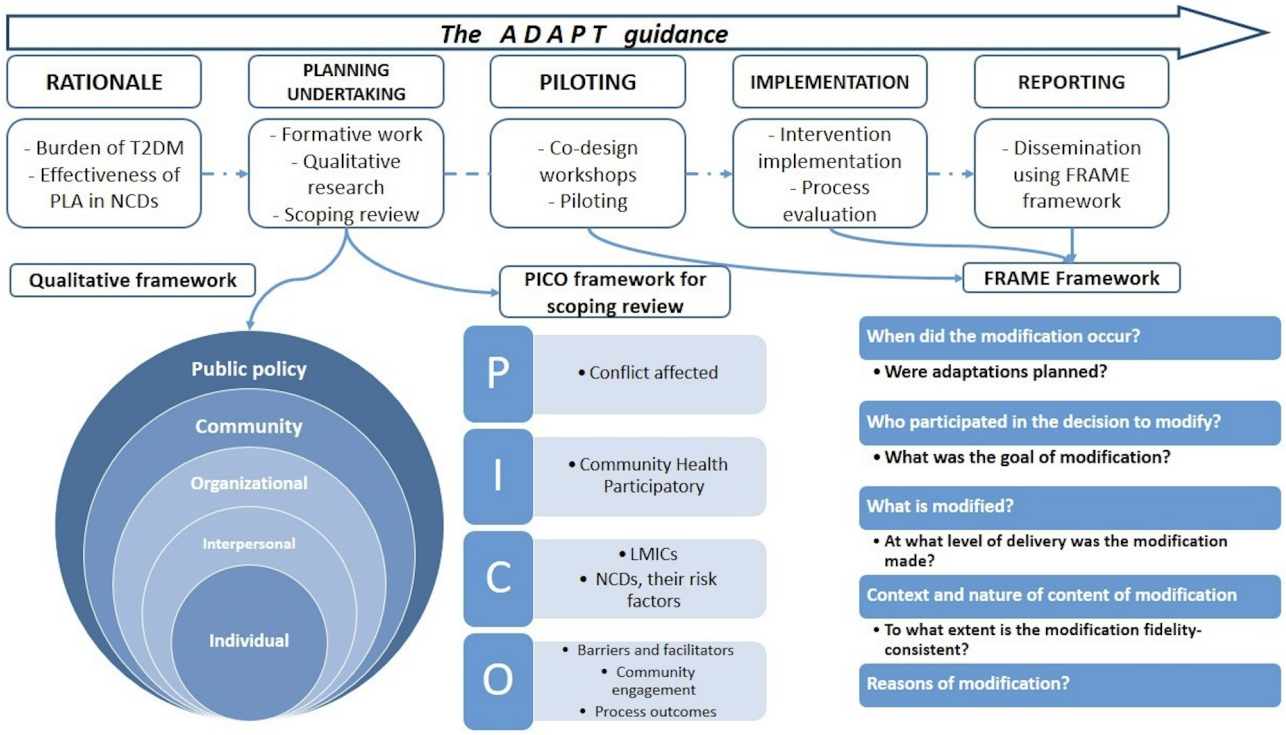

Phase 1: Identifying Barriers and Facilitators to the Implementation of CBIs

Step 1A: Qualitative Study

Overview

We will conduct a qualitative study to explore the understanding of T2DM in rural and urban communities of Pakistan and Afghanistan, the barriers and facilitators to the uptake and maintenance of healthy behaviors related to T2DM, and their willingness to engage in CBIs. We selected a qualitative approach because health-related behaviors are deeply shaped by cultural beliefs, gender roles, and community structures that cannot be adequately captured through quantitative methods. Qualitative inquiry will also allow us to capture the lived experiences and perspectives of diverse community stakeholders, contributing to context-rich insights that are essential for tailoring participatory interventions to fragile and heterogeneous settings. The SEM emphasizes the interaction of individual, interpersonal, organizational, and community-level factors in shaping health outcomes [28]. The SEM was preferred as it offers a comprehensive framework to explore diabetes-related barriers and facilitators across multiple levels, which is essential for designing and adapting CBIs.

Study Settings

In Afghanistan, the feasibility trial will be conducted in Farza and Dehsabz, Kabul district. The trial will mainly focus on the rural population in the Kabul district, which is populated mainly by people from lower socioeconomic backgrounds. Protracted conflicts and political instability have greatly affected the health care system in the district.

In Khyber Pakhtunkhwa, the trial will be conducted in the districts of Peshawar and Swabi. Peshawar, the densely populated capital, has 4.7 million [31] residents and offers comprehensive health care through public and private sectors. Swabi, with 1.9 million people [31], provides primary and secondary care through government facilities supplemented by private health care. Both predominantly rural districts, inhabited by Pashtuns, faced challenges of poverty, limited infrastructure, under-resourced education, and conflicts, resulting in reduced economic growth and development for the past decade.

For the urban feasibility trial in Pakistan, the District Central of Karachi, Sindh, has been selected. With a population of approximately 3.8 million [32], it is one of the most densely populated areas in Karachi, which itself houses 17 to 20 million people. The district's diverse population includes various ethnic groups, and livelihoods range from white-collar jobs to informal labor. Education access is uneven, with significant disparities between elite private schools and underfunded public institutions. Karachi's extensive health care system includes public, private, and nonprofit sectors, but it faces overcrowding and unequal access, particularly in public hospitals.

Participants and Recruitment

Participants will be recruited through purposive sampling from the 3 study sites. The inclusion criteria are adult (over 18 y of age) male and female participants residing within the geographical boundaries of the selected sites. Participants must be able to communicate effectively in either Pashto, Urdu, or Dari (or any other local language) for the interviews and be willing to share their experiences and opinions openly. Participants who have already participated in similar interviews or focus groups for the same study will be excluded to avoid

XSL•FO
RenderX

redundancy and bias. Maximum variation will be ensured by recruiting participants with different demographic characteristics (gender, socioeconomic status, geographical spread, age, ethnicity, etc). A team of male and female researchers will go to the field to recruit participants through their professional contacts and snowballing. In addition, community advisory panels (CAPs) were formed at all study sites, comprising members of the community, such as religious leaders, community elders, health care professionals, individuals living with T2DM, and caregivers. These CAP members will also be requested to help identify potential participants.

Sampling

Five categories of stakeholders, guided by the SEM, will be recruited to explore the views and perspectives of people living with T2DM and people with unknown T2DM status, carers, influential people, and health care providers (Table 1).

Specifically, we will recruit the following:

1. Carers of people with T2DM, adults who are the primary caregivers for someone living with T2DM, are to provide insights into household dynamics and support systems.
2. People with unknown T2DM status, adults who have never been diagnosed with T2DM, and who do not self-identify as having the condition. This group may include individuals with undiagnosed T2DM, as they are unaware of their glycemic status at the time of recruitment, to capture perceptions, beliefs, and practices around prevention, lifestyle, and risk awareness.
3. Influential people, Imams (religious leaders), community elders, teachers, elected representatives, etc, are engaged, given their influence on collective norms and decision-making.
4. A range of health care providers, including doctors, nurses, pharmacists, and community health workers, who treat and care for people with T2DM will be interviewed to reflect organizational and health system perspectives.
5. People with T2DM, adults living with T2DM (who have had a medical diagnosis), to capture both lived experiences of disease management and perceptions of risk and prevention.

As shown in Table 1, a total of 60 IDIs and 16 FGDs with the aforementioned 5 categories of participants will be conducted at each site. Efforts will be made to recruit an equal number of male and female participants across the study sites. In line with qualitative research principles, our sampling strategy is guided by the principle of thematic saturation. The planned sample size is sufficiently large to capture a diverse range of perspectives across major stakeholder groups while allowing flexibility to adjust numbers until saturation is reached.

Table . Participants' categories and sample size for in-depth interviews (IDIs) and focus group discussions (FGDs).

Participants' category and interview type	Sample size			
	Rural Pakistan (Peshawar and Swabi)	Urban Pakistan (Karachi)	Rural Afghanistan (Kabul)	Total sample
People living with T2DM[a]				
IDIs	4	4	4	12
FGDs	4	2	2	8
People with unknown T2DM status				
IDIs	4	4	4	12
FGDs	4	2	2	8
Carers of people with T2DM				
IDIs	4	4	4	12
Health care providers				
IDIs	4	4	4	12
Community influentials				
IDIs	4	4	4	12

[a]T2DM: type 2 diabetes mellitus.

Data saturation will be evaluated through an iterative process of concurrent data collection and analysis. After each round of interviews and FGDs, the research team will review the transcripts to identify emerging codes and themes. Saturation will be considered reached when successive interviews no longer yield new codes or meaningfully expand existing categories. To enhance rigor, at least 2 researchers will independently assess saturation and maintain a saturation grid to document the point at which no new information emerges. This approach ensures a clear, evidence-based justification for the adequacy of the sample and the credibility of the findings.

Data Collection

The IDIs and FGDs will be conducted by a team of male and female research assistants trained in qualitative research methods. An experienced research team member will supervise the fieldwork. The research team will develop topic guides based on the research objectives and the conceptual framework. FGDs

XSL•FO
RenderX

and IDIs are valuable research techniques, and by triangulating these methods and participants, we aim to obtain rich information from various sources. The data collection techniques are described below:

1. In-depth interviews: Interviews will be conducted by research assistants with support from a notetaker and are expected to last 45 to 60 minutes. IDIs will be conducted with all 5 categories of participants in a peaceful place convenient to them. With participants' consent, interviews will be recorded using a digital audio recorder (Multimedia Appendices 2-6).
2. Focus group discussions: FGDs will involve people living with T2DM and people with unknown T2DM status to foster interactive discussions and gain insights into community beliefs and experiences. Separate FGDs will be held for males and females and those with T2DM and with unknown T2DM status to capture diverse perspectives. Discussions will occur at convenient locations, such as local community centers or public spaces, and will be facilitated by a moderator and a notetaker and are expected to last 60 to 90 minutes. With participants' consent, FGDs will be audio-recorded (Multimedia Appendices 7 and 8).

Data Analysis

The recordings of IDIs and FGDs will be transcribed and translated into English. Transcripts and field notes will be uploaded onto a secure server, which is accessible to the research team only. Sections of transcriptions and translations will be checked for accuracy and quality by a senior research team member who is not directly involved in the data collection and transcription and who is bilingual.

The transcripts will be carefully read to reflect participants' articulated meanings and experiences. Initially, open and inductive coding of five transcripts will be conducted to create a list of recurring codes within the data. Based on these initial codes, a coding framework will be prepared in a Microsoft Excel worksheet. After that, the transcripts will be deductively coded and charted into the coding framework, which will be flexible enough to accommodate more codes during the charting process [33]. Next, codes will be merged into categories and themes [34]. The themes identified deductively will be inductively organized using SEM [29] to provide a broader context for the findings. Finally, summaries of all the themes will be prepared to address the research questions.

To ensure the quality of the analysis, data sources (IDIs and FGDs) will be triangulated to confirm the validity and consistency of the findings. Throughout the analysis process, researchers will be encouraged to reflect on their positionalities and place in the study and discuss coding schemes within the team (reflexivity). Peer debriefing will involve discussions with colleagues or experts to review and critique the coding and thematic development process, enhancing the credibility of the analysis. Oversight by experienced researchers will be conducted to review the coding and thematic development for consistency and reliability.

The data will be analyzed separately for Pakistan's urban and rural areas and Afghanistan's rural areas, resulting in 3 distinct reports. A combined summary of these reports will then be prepared to inform the adaptation of the PLA-based intervention.

Step 1B: Scoping Review

Overview

We will conduct a scoping review to map evidence on CHPR interventions for NCDs in crisis-affected LMICs. Because the evidence on PLA for the prevention of NCDs and control in crisis-affected LMICs is still limited and poorly characterized, a scoping review is more appropriate than a full systematic review. This approach allows us to maintain methodological rigor while synthesizing a wide range of study designs and contexts that might otherwise remain fragmented. It will also highlight critical insights into how participatory approaches have been implemented and sustained or perhaps failed in crisis-affected settings. The findings will complement our qualitative inquiry by highlighting transferable elements and best practices, while also identifying context-specific gaps that require tailored adaptation in Afghanistan and Pakistan.

The objective of this review is to identify, describe, and characterize community-based participation interventions and the process outcomes of CHPR interventions; explore the barriers and facilitators to community involvement in such interventions; and provide insights into how participatory approaches have been implemented and sustained in fragile settings.

Guidelines and Framework

The review will follow the Joanna Briggs Institute guidelines and comply with the Arksey and O'Malley Framework [35], and the findings will be reported as per the PRISMA-ScR (Preferred Reporting Items for Systematic Review and Meta-Analysis Extension for Scoping Reviews) checklist [36]. The scoping review has been prospectively registered at Open Science Framework [37], providing detailed information on screening, extraction, and outcomes.

Eligibility Criteria

Inclusion criteria will comprise all original papers published in English on CHPR interventions and NCDs in crisis-affected LMICs, focusing on the adult population (age 18 y or more). In the case of mixed-aged populations (those below and above 18 y), pertinent information specific to the eligible population will be extracted, where possible. Studies with crisis-affected populations, including refugee and internally displaced populations, as specific groups or subsets within the target population, will also be considered and included for data extraction.

Studies will be excluded if they have a pharmacological or genetic focus, target infectious, communicable diseases, focus on domestic, sexual, or gender-based violence, or are conducted in a noncommunity or clinical setting.

Databases and Screening

A total of 4 databases will be searched: PubMed, Web of Science, Scopus, and Google Scholar. The search strategy will focus on 4 key areas, such as community participation, NCDs, LMIC, and crisis-affected settings, and will be adapted for each

XSL•FO
RenderX

database. The reference list of the included studies will be assessed for additional sources.

All retrieved records will be imported into Covidence for de-duplications and screening. Two reviewers will independently review screen titles, abstracts, and full text against eligibility criteria as presented in Figure 2.

Figure 2. Flowchart of the screening strategy.

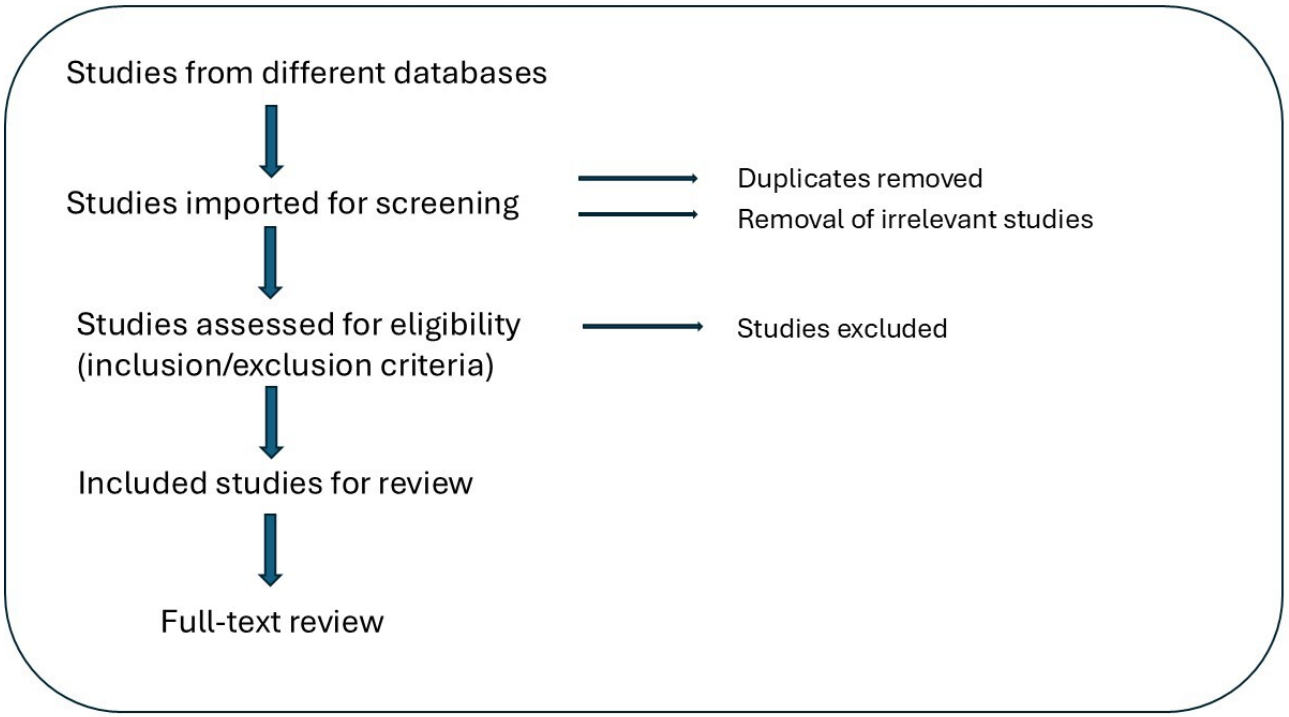

Data Extraction and Charting

A standardized data extraction template will be developed, piloted, and refined before use. Two independent reviewers will independently extract data on predefined characteristics as per study objectives. Disagreements at any stage of review will be resolved through discussion or by consulting a third reviewer.

Data will be synthesized narratively to highlight patterns across studies, identify transferable elements and best practices, and outline evidence gaps. As consistent with the scoping review, no formal risk of bias assessment will be undertaken for included studies. Although this is appropriate for the review type, it represents a methodological limitation that will be acknowledged in interpreting the findings.

Linking Formative Findings

The findings of the qualitative study and the scoping review will be discussed with the CAP members, including community representatives, to validate and increase the relevance of the findings to the local contexts. Furthermore, these findings will not remain at a descriptive level but will be systematically fed into the adaptation process. Subsequently, a draft of major components and materials of the PLA intervention will be developed and shared with stakeholders in co-design workshops for their input and consensus.

Phase 2: Co-Designing the Intervention Materials

Co-design methods will be employed to culturally, relevantly, and context-specifically adapt PLA interventions for the prevention and control of T2DM, using the findings from qualitative and scoping reviews [38]. Relevant stakeholders, including the research team, representatives of the local communities, health care providers, individuals with T2DM or those at risk, and graphic designers (for the preparation of materials), will be engaged in the co-design process to foster a collaborative approach to creating and refining strategies that are effective, sustainable, and acceptable to those who will implement and benefit from them. Co-design methods will ensure that the intervention is grounded in the sociocultural context and the lived experiences of the community, leading to more comprehensive and long-term solutions.

Process of Integration Through Co-Design

Through a series of co-design workshops (2 - 4 per site), with the stakeholders mentioned above, the PLA intervention will be adapted for each site. The details of the workshops will be decided closer to the time; we will also allow for flexibility. However, each workshop will combine findings from formative qualitative research with insights drawn from the scoping review, highlighting transferable lessons and best practices as well as gaps where evidence is weak. This deliberate integration of global evidence with local perspectives will ensure that the adaptation process is both evidence-informed and context-responsive, allowing the formative work (scoping review and qualitative study) to inform the intervention design iteratively and in real time.

To account for contextual diversity, the data collected from the 3 study sites (urban Pakistan, rural Pakistan, and rural Afghanistan) will undergo separate analysis, and distinct

XSL•FO
RenderX

co-design workshops will be conducted in each site. This approach ensures that although the overall adaptation framework is consistent, the outputs will be site-specific PLA interventions, tailored to the unique needs and realities of each context, instead of a one-size-fits-all approach. Scoping review findings will be systematically mapped onto key decision points in the workshops (eg, content gaps, delivery mechanisms, and best practices), enabling participants to cross-validate global evidence with local insights. This structured process of triangulation and documentation will help prevent fragmentation and instead generate cohesive, contextually adaptive intervention packages.

The workshops are likely to look as follows:

- Workshop 1: In the first workshop, participants will be introduced to the intervention and materials, with an overview of the EMPOWER-D project, PLA, the target population, key components, and the goals of the intervention. Formative research findings will be shared, highlighting community needs, preferences, and challenges related to T2DM prevention and control. Group activities and discussions will then identify areas of the intervention materials that need review or adaptation.
- Workshop 2: In the second workshop, specific sections of the materials will be reviewed in detail—this is likely to be done through group work. Groups will assess content and context and propose changes or new aspects. The groups will present their findings, and the discussion will focus on reaching a consensus on the necessary changes. These changes will be prioritized based on feasibility, impact, and relevance to the local context.
- Workshop 3: The third workshop will review the changes made to the intervention materials. Revised materials will be presented, and participants will review and provide feedback on the effectiveness and clarity of the revisions. Any final adjustments will be discussed and agreed upon.
- Workshop 4: The final workshop will focus on practical aspects of the intervention, such as meeting places, session frequencies, and group sizes. The discussion will consider logistical factors, such as accessibility, timing, cultural appropriateness, and translation of materials in the local language. Participants will collaborate to finalize the format of the intervention, document decisions, and review the implementation plan and timeline.

These workshops will include approximately 10 to 12 participants representing different stakeholders and take place 1 week apart, leaving enough time for the research team to prepare the agenda, required documents, and activity plan for the next workshop.

Managing Contextual Challenges and Trade-Offs

Given the fragile and diverse contexts in which this study will be conducted, we anticipate potential challenges related to political instability, resource constraints, and logistical barriers. To safeguard feasibility, we have built in strategies to prioritize essential components, adapt sequencing, and manage trade-offs without compromising the study's core objectives.

1. Prioritization of core components*:* The qualitative study and co-design workshops are the nonnegotiable elements essential for this adaptation. The scoping review and pilot are designed to be complementary, but timelines can be adjusted if external constraints arise.
2. Phased adaptation: Activities will not run in parallel across all 3 sites. Instead, learnings from one setting (ie, Pakistan) may inform and streamline work in subsequent sites (ie, Afghanistan).
3. Monitoring*:* Oversight committees, comprising senior team members and local and international partners, will monitor feasibility in real time and make decisions about scaling back or sequencing activities if political or logistical barriers emerge.
4. Contextual safeguards: Given the anticipated instability in Afghanistan, for example, we have planned remote support and contingency arrangements, such as virtual supervision, flexible timelines, and coordination through local leaders, to safeguard study continuity. Engagement with community leaders and CAP members will guide trust-building and help mitigate political or cultural resistance to community mobilization. Regular reviews of risk assessments, clear incident reporting and supervision mechanisms, and contingency plans for alternative data collection will ensure researcher safety and continuity of fieldwork in case of restricted access to certain districts.

Expected Outcome of the Adaptation Process

The adaptation process will result in the development of the following:

1. A finalized description of the intervention (including meeting content, the composition of meetings, community meetings, and areas for discussion). The detailed intervention may vary across different sites.
2. A manual for the recruitment and training of staff for PLA implementation in the field.
3. A manual for community mobilizers—a guide to help them deliver the PLA intervention in the field.
4. Completed materials for group meetings, such as flipcharts and other visual aids.

Phase 3: Piloting of Adapted PLA Materials

We will conduct a focused pilot of 1 - 2 meetings at each site to obtain feedback on the co-designed PLA materials and assess initial feasibility and acceptability before moving to implementation. Given resource and time constraints, this will not be a full-scale pilot of the intervention but rather an opportunity to examine key aspects of how the adapted materials and group processes function in real settings.

In rural Pakistan, the adapted PLA materials will be tested in 2 rural clusters (one in Peshawar and one in Swabi), with two community groups (male and female) per cluster. In rural Afghanistan (Kabul) and urban Pakistan (Karachi), the pilot will be conducted in 1 cluster each, with 2 community groups (male and female). Groups will be facilitated by trained male and female community mobilizers under the supervision of senior staff.

Two pilot meetings will be conducted at each site, and the evaluation will focus on 2 core domains:

XSL•FO
RenderX

1. Group dynamics (feasibility of group process): Attendance records (proportion of invitees attending), levels of participation and engagement (observed interaction, interest, and willingness to contribute), and mobilizer reflections on ease of facilitation.
2. Acceptability of PLA materials: Clarity of content, cultural appropriateness, and relevance of flipcharts and participatory activities. These will be assessed through structured field notes, observation checklists, and short exit interviews with a purposive sample of group members.

Data sources will therefore include the following:

1. Attendance record (to capture the feasibility of participation)
2. Observation checklists/field notes (to assess dynamics, participation, and fidelity to facilitation guides)
3. Exit interviews (at least 2 per site) with participants (to capture perceptions of clarity, appropriateness, and suggestions for improvement)
4. Debriefs with community mobilizers and supervisors (to reflect on what was feasible and acceptable to deliver)

Success criteria will be pragmatic: for example, reasonable attendance (≥50% of invited participants attending at least one of two meetings), visible engagement in discussions, and no major concerns about the cultural acceptability of the materials.

Findings will be summarized in a short pilot report structured around these domains. Any revisions to the PLA materials will be informed by participants, mobilizers, and supervisors' feedback. All modifications will be documented using the FRAME framework, including the rationale for the change [30].

Patient and Public Involvement

This study has an active program of community engagement for which we have constituted CAPs at each of the project sites. The CAP members represent both men and women from the community who have knowledge and understanding of the local contexts and cultures. The members of the CAPs include individuals living with diabetes, individuals without diabetes, caregivers of diabetic individuals, youth representatives, social workers, community health workers, persons from the seminary or mosque, school, and local women's representation. CAP members have been involved in designing the topic guides and their translation into the local languages. They have also been involved in identifying participants for the study and will help in disseminating the study findings. Their involvement will improve research quality and help foster a sense of ownership of outputs by the community.

Results

Under the C4IMPACT, EMPOWER-D was funded and launched in November 2022 by NIHR-United Kingdom (NIHR203248). Following that, the main protocol of the project was developed, which took almost a year. This was followed by detailed planning for adaptation work during which the adaptation protocol was conceived. In early 2024, the team started working simultaneously on scoping reviews and the qualitative study. Topic guides for IDIs and FGDs were prepared and piloted in May 2024. Recruitment and invitations for stakeholders or participants were extended, and IDIs and FGDs began in July 2024 and have been completed in Peshawar, Swabi, and Karachi. As of August 2025, data collection is currently ongoing in Afghanistan. The scoping review has also been completed. Results are anticipated to be published in mid-2026 following the completion of data analysis.

Discussion

Framework-Guided Contextual Adaptation

This study hypothesizes that adapting the D-Magic intervention through systematic use of the ADAPT guidance and the FRAME framework will result in contextually relevant, acceptable, and feasible community-based participatory interventions tailored for the prevention and control of T2DM in Afghanistan and Pakistan. By incorporating sociocultural preferences, local resources, and community participation, the adapted intervention is expected to enhance both its effectiveness and sustainability in these fragmented settings.

Principal Findings

We anticipate that the adaptation process will generate interventions that are culturally appropriate, acceptable to participants, and feasible to implement across 3 diverse sites, building a pathway for future implications in other LMICs. The process is designed such that it maintains the internal logic and core of the original PLA intervention, while allowing for necessary contextual modifications. This balance between fidelity and flexibility is expected to strengthen community ownership and improve the potential for sustained impact on diabetes prevention and control.

Comparison to Prior Work

Our approach builds upon evidence from the D-Magic trial in Bangladesh, which demonstrated the effectiveness and cost-effectiveness of PLA in reducing diabetes risk factors in a collectivist, low-resource setting. Afghanistan and Pakistan share some social structures and health system challenges, making PLA a potentially transferable and scalable strategy. At the same time, this study contributes to the growing literature on how systematic adaptation frameworks can be applied to EBIs to ensure relevance across diversified populations.

Strengths and Limitations

A major strength of this study lies in its systematic and participatory approach to adaptation, which focuses on community engagement, ensures transparency in documenting changes, and provides insights into maintaining fidelity. However, we understand that given fragile health systems and potential political instability or security issues, feasibility and fidelity may be affected in ways that are difficult to fully anticipate. As this is a protocol of an adaptation study, the conclusions are based on anticipated rather than observed outcomes.

Dissemination

The adapted PLA toolkit will be disseminated within the project team for its implementation. Key adaptation findings will also be shared with the CAP members. Our qualitative research

XSL•FO
RenderX

findings and the scoping review will be disseminated through high-impact, open-access peer-reviewed journals and at local and international conferences. Reports will also be prepared for policymakers and stakeholders in the countries involved.

Future Directions

The findings of this study will emphasize the importance of systematic adaptation processes for shaping intervention delivery, effectiveness, and sustainability in implementation research, particularly in resource-constrained and crisis-affected settings. This study will also provide a replicable framework for other EBIs targeting NCDs and related risk factors in LMICs.

Conclusion

The adaptation of D-Magic for EMPOWER-D is expected to synthesize a culturally grounded intervention that both adheres to its psychosocial origin and responds to the contextual realities of Afghanistan and Pakistan. By applying structured frameworks to guide and document this process, the results are expected to provide insights into how fidelity and contextual relevance can be balanced in complex and diverse, resource-limited settings. The hypothesized findings will inform the design of a definitive trial and contribute to the wider evidence base on the adaptation and implementation of participatory CBIs for NCD prevention and control in LMICs.

Acknowledgments

The authors thank Gul Ghutai, Sana Rahman, Alishba Zahid, and Amna Mansoor for their valuable support and the D-Magic team for sharing their study materials. The authors declare that no generative AI tools were used in the generation of this manuscript.

Funding

This study is funded by “NIHR Global Health Research Centre for Improving Mental and Physical Health Together (C4IMPACT), NIHR203248” (National Institute for Health and Care Research-United Kingdom). The C4IMPACT is a virtual noncommunicable disease (NCD) research center that collaborates with stakeholders across Afghanistan, Pakistan, and the United Kingdom to address the burden of NCDs in Afghanistan and Pakistan. The views expressed in this publication are those of the author(s), and the funders do not have any role in study design, data collection and analysis, manuscript preparation, and the publication processes.

Data Availability

This article describes the study protocol for the intervention adaptation process and does not report primary data. All relevant supporting documents are included within the article and in Multimedia Appendices 2-8.

Authors' Contributions

AT conceptualized and designed the study, drafted the manuscript, tables, and figures, reviewed and edited the manuscript, was responsible for the final version of the manuscript, and will lead the qualitative study in urban Pakistan.
NS conceptualized and designed the study, drafted the manuscript, tables, and figures, was responsible for the final version of the manuscript, and will lead the qualitative research in rural Pakistan.
SI conceptualized and designed the study, prepared the manuscript, designed the study data management systems, and will lead the scoping review.
NK prepared the manuscript and will colead the scoping review.
MA prepared the manuscript and will contribute to the qualitative study and scoping review.
AK conceptualized and designed the study and reviewed the manuscript.
FA prepared the manuscript, will provide qualitative research methods expertise, and contribute to translations of all study tools in local languages.
HMJ conceptualized and designed the study, reviewed the manuscript, and will oversee qualitative research for all sites, along with the scoping review expertise and overall adaptation processes.
RZ prepared the manuscript, designed the study data management systems, and will contribute to trial assessments and process evaluation, managing the trial for the urban Pakistan site.
KR conceptualized and designed the study and prepared and reviewed the manuscript, leading the trial for rural Pakistan.
ARS conceptualized and designed the study and reviewed the manuscript, leading the trial in Afghanistan.
SA conceptualized and designed the study, leading the adaptation process for the project, and prepared and reviewed the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The theory of change logic model for PLA intervention in EMPOWER-D.
[PNG File, 346 KB - resprot_v15i1e71602_app1.png]

XSL•FO
RenderX

Multimedia Appendix 2
Topic guide for in-depth interviews with participants from community with unknown diabetes status.
[DOCX File, 19 KB - resprot_v15i1e71602_app2.docx]

Multimedia Appendix 3
Topic guide for in-depth interviews with participants with diabetes.
[DOCX File, 20 KB - resprot_v15i1e71602_app3.docx]

Multimedia Appendix 4
Topic guide for in-depth interviews with carers of people with diabetes.
[DOCX File, 20 KB - resprot_v15i1e71602_app4.docx]

Multimedia Appendix 5
Topic guide for in-depth interviews with health care providers.
[DOCX File, 19 KB - resprot_v15i1e71602_app5.docx]

Multimedia Appendix 6
Topic guide for in-depth interviews with representatives of the community.
[DOCX File, 19 KB - resprot_v15i1e71602_app6.docx]

Multimedia Appendix 7
Focus group discussion topic guide for participants with unknown diabetes status.
[DOCX File, 17 KB - resprot_v15i1e71602_app7.docx]

Multimedia Appendix 8
Focus group discussion topic guide for participants living with diabetes.
[DOCX File, 17 KB - resprot_v15i1e71602_app8.docx]

References

1. Flood D, Seiglie JA, Dunn M, et al. The state of diabetes treatment coverage in 55 low-income and middle-income countries: a cross-sectional study of nationally representative, individual-level data in 680 102 adults. Lancet Healthy Longev 2021 Jun;2(6):e340-e351. [doi: 10.1016/s2666-7568(21)00089-1] [Medline: 35211689]
2. Mueller H, Techasunthornwat C. Poverty and shard prosperity 2020: conflict and poverty. : World Bank Group; 2020 URL: https://documents1.worldbank.org/curated/en/519741603804458786/pdf/Conflict-and-Poverty.pdf [accessed 2026-02-26]
3. Truppa C, Yaacoub S, Valente M, Celentano G, Ragazzoni L, Saulnier D. Health systems resilience in fragile and conflict-affected settings: a systematic scoping review. Confl Health 2024 Jan 3;18(1):2. [doi: 10.1186/s13031-023-00560-7] [Medline: 38172918]
4. Dadras O, Nyaboke Ongosi A, Wang CW. Prevalence and correlates of diabetes and impaired fasting glucose among adults in Afghanistan: insights from a national survey. SAGE Open Med 2024;12:20503121241238147. [doi: 10.1177/20503121241238147] [Medline: 38516644]
5. Magliano DJ, Boyko EJ. Chapter 3, global picture. In: IDF Diabetes Atlas, 10th Edition: International Diabetes Federation; 2021. URL: https://diabetesatlas.org/resources/previous-editions/ [accessed 2026-03-05]
6. Hills AP, Arena R, Khunti K, et al. Epidemiology and determinants of type 2 diabetes in South Asia. Lancet Diabetes Endocrinol 2018 Dec;6(12):966-978. [doi: 10.1016/S2213-8587(18)30204-3] [Medline: 30287102]
7. Kengne AP, Ramachandran A. Feasibility of prevention of type 2 diabetes in low- and middle-income countries. Diabetologia 2024 May;67(5):763-772. [doi: 10.1007/s00125-023-06085-1] [Medline: 38355989]
8. Arage MW, Kumsa H, Asfaw MS, et al. Exploring the health consequences of armed conflict: the perspective of Northeast Ethiopia, 2022: a qualitative study. BMC Public Health 2023 Oct 24;23(1):2078. [doi: 10.1186/s12889-023-16983-z] [Medline: 37875885]
9. GBD 2021 Diseases and Injuries Collaborators. Global incidence, prevalence, years lived with disability (YLDs), disability-adjusted life-years (DALYs), and healthy life expectancy (HALE) for 371 diseases and injuries in 204 countries and territories and 811 subnational locations, 1990–2021: a systematic analysis for the Global Burden of Disease Study 2021. Lancet 2024 May 18;403(10440):2133-2161. [doi: 10.1016/S0140-6736(24)00757-8] [Medline: 38642570]
10. Bain SC, Cummings MH, McKay GA. Multidisciplinary approach to management and care of patients with type 2 diabetes mellitus. EMJ Diabet 2019;7:73-81. [doi: 10.33590/emjdiabet/10314374]

XSL•FO
RenderX

11. Sun H, Saeedi P, Karuranga S, et al. IDF Diabetes Atlas: global, regional and country-level diabetes prevalence estimates for 2021 and projections for 2045. Diabetes Res Clin Pract 2022 Jan;183:109119. [doi: 10.1016/j.diabres.2021.109119] [Medline: 34879977]
12. Butt MD, Ong SC, Wahab MU, et al. Cost of illness analysis of type 2 diabetes mellitus: the findings from a lower-middle income country. Int J Environ Res Public Health 2022 Oct 2;19(19):12611. [doi: 10.3390/ijerph191912611] [Medline: 36231911]
13. Campbell JA, Yan A, Egede LE. Community-based participatory research interventions to improve diabetes outcomes: a systematic review. Diabetes Educ 2020 Dec;46(6):527-539. [doi: 10.1177/0145721720962969] [Medline: 33353510]
14. Nyame S, Boateng D, Heeres P, et al. Community-based strategies to improve health-related outcomes in people living with hypertension in low- and middle-income countries: a systematic review and meta-analysis. Glob Heart 2024;19(1):51. [doi: 10.5334/gh.1329] [Medline: 38883258]
15. Wallerstein NB, Duran B. Using community-based participatory research to address health disparities. Health Promot Pract 2006 Jul;7(3):312-323. [doi: 10.1177/1524839906289376] [Medline: 16760238]
16. Wallerstein N, Oetzel JG, Sanchez-Youngman S, et al. Engage for equity: a long-term study of community-based participatory research and community-engaged research practices and outcomes. Health Educ Behav 2020 Jun;47(3):380-390. [doi: 10.1177/1090198119897075] [Medline: 32437293]
17. Orlowski P. Freirean conceptions of participatory action research and teaching for social justice - same struggle, different fronts. CJAR 2019;20(1):30-51. [doi: 10.33524/cjar.v20i1.445]
18. Seal AJ, Mohamed HA, Stokes-Walter R, et al. Use of an adapted participatory learning and action cycle to increase knowledge and uptake of child vaccination in internally displaced persons camps (IVACS): a cluster-randomised controlled trial. Vaccine 2023 May 5;41(19):3038-3046. [doi: 10.1016/j.vaccine.2023.02.016] [Medline: 36906409]
19. Prost A, Colbourn T, Seward N, et al. Women's groups practising participatory learning and action to improve maternal and newborn health in low-resource settings: a systematic review and meta-analysis. Lancet 2013 May 18;381(9879):1736-1746. [doi: 10.1016/S0140-6736(13)60685-6] [Medline: 23683640]
20. Morrison J, Osrin D, Alcock G, et al. Exploring the equity impact of a maternal and newborn health intervention: a qualitative study of participatory women's groups in rural South Asia and Africa. Int J Equity Health 2019 Apr 11;18(1):55. [doi: 10.1186/s12939-019-0957-7] [Medline: 30971254]
21. Fottrell E, Ahmed N, Morrison J, et al. Community groups or mobile phone messaging to prevent and control type 2 diabetes and intermediate hyperglycaemia in Bangladesh (DMagic): a cluster-randomised controlled trial. Lancet Diabetes Endocrinol 2019 Mar;7(3):200-212. [doi: 10.1016/S2213-8587(19)30001-4] [Medline: 30733182]
22. Moore G, Campbell M, Copeland L, et al. Adapting interventions to new contexts-the ADAPT guidance. BMJ 2021 Aug 3;374:n1679. [doi: 10.1136/bmj.n1679] [Medline: 34344699]
23. Shams S. Nation, religion, and language ideology: the case of postcolonial Bangladesh. J-BILD 2022;6(1):97-110 [FREE Full text]
24. Gartenstein-Ross D, Vassefi T. The forgotten history of Afghanistan-Pakistan relations. Yale J Int'l Aff 2012;7:38 [FREE Full text]
25. Perera C, Salamanca-Sanabria A, Caballero-Bernal J, et al. No implementation without cultural adaptation: a process for culturally adapting low-intensity psychological interventions in humanitarian settings. Confl Health 2020;14(1):46. [doi: 10.1186/s13031-020-00290-0] [Medline: 32684948]
26. Stanikzai MH, Wafa MH, Rahimi BA, Sayam H. Conducting health research in the current afghan society: challenges, opportunities, and recommendations. Risk Manag Healthc Policy 2023;16:2479-2483. [doi: 10.2147/RMHP.S441105] [Medline: 38024503]
27. Habib SS, Jamal WZ, Zaidi SMA, et al. Barriers to access of healthcare services for rural women-applying gender lens on TB in a rural district of Sindh, Pakistan. Int J Environ Res Public Health 2021 Sep 26;18(19):10102. [doi: 10.3390/ijerph181910102] [Medline: 34639403]
28. Muhammad Q, Eiman H, Fazal F, Ibrahim M, Gondal MF. Healthcare in Pakistan: navigating challenges and building a brighter future. Cureus 2023 Jun;15(6):e40218. [doi: 10.7759/cureus.40218] [Medline: 37435262]
29. Caperon L, Saville F, Ahern S. Developing a socio-ecological model for community engagement in a health programme in an underserved urban area. PLoS ONE 2022;17(9):e0275092. [doi: 10.1371/journal.pone.0275092] [Medline: 36155664]
30. Wiltsey Stirman S, Baumann AA, Miller CJ. The FRAME: an expanded framework for reporting adaptations and modifications to evidence-based interventions. Implement Sci 2019 Dec;14(1):58. [doi: 10.1186/s13012-019-0898-y]
31. Population by sex, sex ratio, average household size, and growth rate – Khyber Pakhtunkhwa. Pakistan Bureau of Statistics. 2023. URL: https://www.pbs.gov.pk/census/ [accessed 2026-03-02]
32. Population by sex, sex ratio, average household size, and growth rate – Sindh. Pakistan Bureau of Statistics. 2023. URL: https://www.pbs.gov.pk/census/ [accessed 2026-03-02]
33. Henderson DX, Baffour TD. Applying a socio-ecological framework to thematic analysis using a statewide assessment of disproportionate minority contact in the United States. Qual Rep 2015;20(12):1960-1973. [doi: 10.46743/2160-3715/2015.2405]

XSL•FO
RenderX

34. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
35. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
36. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
37. Community health participatory (CHP) interventions in the prevention and control of non-communicable diseases in conflict-affected lmics- a scoping review protocol. Open Science Framework. URL: https://osf.io/eu95b/overview [accessed 2026-02-26]
38. Ssanyu JN, Kananura RM, Birabwa C, et al. How a co-design process led to more contextually relevant family planning interventions in emerging urban settings in Eastern Uganda. PLOS Glob Public Health 2023;3(9):e0002421. [doi: 10.1371/journal.pgph.0002421] [Medline: 37773920]

Abbreviations

C4IMPACT: Centre for Improving Mental and Physical Health Together Declarations
CAP: community advisory panel
CBIs: community-based interventions
CHPR: Community Health Participatory Research
D-Magic Trial: Diabetes Mellitus: Action Through Community Groups or mHealth Information for Better Control
EBIs: evidence-based intervention
EMPOWER-D: Engagement of Community Through Participatory Learning and Action for Control and Prevention of Type 2 Diabetes and Its Risk Factors
FGD: focus group discussion
FRAME: The Framework for Reporting Adaptations and Modifications in Evidence-based Interventions
IDI: in-depth interview
LMICs: low- and middle-income countries
NCDs: noncommunicable diseases
PLA: Participatory Learning and Action
PRISMA-ScR: Preferred Reporting Items for Systematic Review and Meta-Analysis Extension for Scoping Reviews
SEM: socioecological model
T2DM: type 2 diabetes mellitus

Edited by J Sarvestan; submitted 23.Jan.2025; peer-reviewed by AN Ali; revised version received 16.Oct.2025; accepted 16.Oct.2025; published 27.Mar.2026.

Please cite as:
Tahir A, Sanauddin N, Imtiaz S, Khaleeq N, AbdeAli M, Khan A, Ahmed F, Jennings HM, Zafar R, Rehman K, Shahab AR, Afaq S Application of the ADAPT Framework to Contextualize a Participatory Learning and Action Community Intervention for the Prevention and Control of Type 2 Diabetes Mellitus in Urban and Rural Settings in Afghanistan and Pakistan: Protocol for Intervention Adaptation JMIR Res Protoc 2026;15:e71602
URL: https://www.researchprotocols.org/2026/1/e71602
doi:10.2196/71602

© Amber Tahir, Noor Sanauddin, Sara Imtiaz, Nadia Khaleeq, Mariam AbdeAli, Asima Khan, Farrukh Ahmed, Hannah Maria Jennings, Rubia Zafar, Khalid Rehman, Abdul Rahman Shahab, Saima Afaq. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Chat-Based Decision Support System for the Maternal Health Journey in Assam, India: Protocol for a Mixed Methods Multiphase Implementation Study

Ashita Munjral[1], MSc; Swapnil Rawat[2], PhD; Charlotte E Warren[3], PhD; Jitender Nagpal[4], MD; Ben Bellows[5], PhD; Aditi Aditi[6], PhD; Pompy Sridhar[7], MBA; Sowmya Ramesh[1], MS, MPH

[1]Maternal Health, Population Council Consultancy, New Delhi, Delhi, India

[2]Research Department, Sitaram Bhartia Institute of Science and Research, B-16 Qutab Institutional Area, New Delhi, Delhi, India

[3]Social Behavioral Science Research, Population Council Inc, New York, NY, United States

[4]Department of Pediatrics, Sitaram Bhartia Institute of Science and Research, New Delhi, Delhi, India

[5]Digital Health, Nivi Inc., Sudbury, MA, United States

[6]Research Department, Population Council Institute, New Delhi, Delhi, India

[7]MSD For Mothers, MSD (India), Mumbai, Maharashtra, India

Corresponding Author:
Swapnil Rawat, PhD
Research Department, Sitaram Bhartia Institute of Science and Research, B-16 Qutab Institutional Area, New Delhi, Delhi, India

Abstract

Background: Assam, India, has the country's highest maternal mortality ratio (195 per 100,000 live births), mainly due to poor access to and quality of maternal health (MH) care. Many women receive inadequate antenatal and postnatal services, made worse by isolation, socioeconomic barriers, and weak health care infrastructure. Digital tools like mobile messaging and chatbots have improved antenatal care (ANC) and facility-based deliveries in similar settings. The e-SAATHI (Strengthening ANC/PNC via AskNivi Tailored Health Information, Referrals, and Follow-Up) project aims to provide personalized, stage-specific maternal health support through a chat-based system in Assam.

Objective: This study assesses the acceptability, feasibility, and effectiveness of the e-SAATHI chatbot in increasing women's access to MH information and improving ANC and PNC service uptake across public and private facilities. Objectives include increasing ANC/PNC use (eg, ≥4 ANC visits and timely PNC), promoting respectful care, and gathering insights for scaling digital health in high-burden regions.

Methods: Phase 1 (0 - 3 mo) involves co-designing and pilot testing aligned with World Health Organization and national guidelines. Phase 2 (4 - 24 mo) involves enrolling pregnant and postpartum women via health facilities and social media. The chatbot sends 2-3 messages weekly from 10-week pregnancy to 15 weeks postpartum. About 300 health care providers will be trained and engaged for onboarding and feedback. Phase 3 (25 - 36 mo) involves scaling up across districts, reaching 225,000 women. Data collection includes interviews, surveys, facility assessments, and chatbot analytics. Qualitative analysis will explore experiences; quantitative data (ANC completion, facility delivery, PNC follow-up, and satisfaction) will compare pre- and post-interventions. Ethical approvals, informed consent, and data confidentiality are observed.

Results: The study was funded in September 2022. As of August 2025, 210 facilities have been onboarded, and 201,813 women were enrolled. Chatbot-based data collection began in April 2023 and will continue through the study period. Qualitative and quantitative evaluation data collection started in November 2023 and is expected to complete in June 2027. Interim analyses will be conducted after midline data collection in 2026; final analyses will be performed after endline data collection in 2027. The primary outcome will be the change in the composite quality score of maternal and newborn care. Secondary outcomes will include service uptake indicators, user-reported knowledge and self-care practices, and satisfaction with care. Operational feasibility—including provider integration and barriers such as digital literacy and connectivity—will also be assessed. Ongoing collaborative learning and adapting cycles are expected to capture intervention adaptations and inform optimal strategies for scale-up.

Conclusions: e-SAATHI offers a scalable digital approach to improve MH across a variety of sociodemographic, linguistic, and risk settings. By delivering timely, personalized support, the chatbot may enhance health-seeking behavior and outcomes in Assam and in similar low-resource areas globally.

International Registered Report Identifier (IRRID): DERR1-10.2196/81873

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e81873) doi:10.2196/81873

KEYWORDS

maternal health; digital solution; protocol; Assam; India

Introduction

Despite decades of investments and initiatives to address intergenerational and deep-rooted social and cultural norms around maternal health (MH), progress has been slow in selected geographies of India, such as Assam. The state records the highest maternal mortality ratio in the country, nearly double the national average of 195 per 100,000 live births (2018 - 2020) compared to 103 per 100,000 live births in (2017 - 2019) [1-3]. Key indicators reflect significant gaps in care quality and utilization—only half of pregnant women receive the recommended 4 or more antenatal care (ANC) visits—approximately 48% received postnatal care (PNC) from trained health personnel, and, among Janani Suraksha Yojana beneficiaries, only 45% of newborns initiated breastfeeding within the critical first hour [4,5].

The persistently high maternal mortality rate in Assam stems from multiple interconnected factors [6,7]. Geographic barriers limit health care access, while socioeconomic challenges compound these difficulties [6,8]. The public health system faces an overwhelming delivery burden insufficiently coordinated with private MH providers [9]. Gaps include limited access to personalized pregnancy and childbirth information, inadequate self-care guidance, low insurance coverage, and high out-of-pocket expenditures [8,10,11]. Women's restricted autonomy further constrains their access to, and use of, quality reproductive and MH services—collectively contributing to poor health outcomes for mothers and newborns [8,12]. While India has achieved a notable decline in maternal and newborn mortality over the past 2 to 3 decades, partly attributable to various maternal and newborn health policy reforms, significant challenges remain, particularly in high-burden states like Assam [6,7].

Direct-to-beneficiary mobile health (mHealth) programs that provide stage-based health information to new and expectant mothers have proliferated rapidly over the last decade [13]. In low-income and middle-income countries (LMICs), mHealth messaging programs represent some of the most successful examples of digital health programs, demonstrating measurable improvements in MH services [13,14]. Several digital MH services have achieved significant scale in Bangladesh, India, South Africa, and Tanzania, with each program reaching over 1 million subscribers and demonstrating potential for population-level impact [13,15].

The Government of India has prioritized digital health initiatives, leveraging the fact that more than 3-quarters of men and two-thirds of women own a mobile phone [16]. Smartphone ownership reaches approximately half of men (49%) and just over a quarter of women (26%) with notable variation between urban and rural settings.

Digital platforms and mHealth apps offer significant potential to improve the health information delivery and service access, even in contexts with low literacy levels [17-19]. Most digital health interventions use telephone-based systems that can engage individuals in their everyday lives, help them navigate into and across the health system, and capture client feedback following provider interactions.

Evidence demonstrates that messaging systems using digital apps or chatbots positively impact key MH indicators, including improved attendance at 4 or more ANC visits and increased facility-based deliveries [20-22].

This protocol paper describes the e-SAATHI (*Strengthening Antenatal Care/Postnatal Care via AskNivi Tailored Health Information, Referrals, and Follow-Up*) project—an iterative, collaborative learning and adapting (CLA) approach to document and monitor the introduction and scale-up of a chat-based digital support system (askNivi) for pregnant and postnatal women residing in the state of Assam, India, throughout their MH journey. Using WhatsApp, health information messages are sent to pregnant and postnatal women weekly based on the World Health Organization (WHO) guidelines during pregnancy and up to 15 weeks postnatally. The e-SAATHI project also identifies public and private MH providers to engage with the chatbot—to provide quality services to pregnant and postnatal women.

Methods

Aim and Objectives

The main aim of e-SAATHI is to evaluate the effectiveness of designing, implementing, and evaluating a digital health approach (chatbot) for improving women's access to MH information and quality services in public and private health facilities in the state of Assam.

Key objectives are as follows: (1) to assess the acceptability and feasibility of an MH chatbot approach in increasing access/uptake of respectful MH services in public and private facilities; (2) to design, implement, monitor, and evaluate the effect of MH chatbot to improve the uptake of services; and (3) to document and assess the dynamics of implementing interventions to introduce an MH chatbot and generate lessons for replication/expansion at scale.

Population

Our geographic target area is in Northeast India in the demographically and topographically diverse state of Assam where progress in improving key MH indicators has stalled and interventions are required to accelerate toward sustainable development goal targets. Given the challenges that Assamese pregnant women face regarding equitable access to timely, respectful, and quality MH care across geography and socioeconomic strata, a weak and overburdened public health system, and an underutilized private health sector, the e-SAATHI project intends to adapt, test, and scale a digital marketplace approach (chat-based digital support system:

XSL•FO
RenderX

askNivi). Respectful maternity care is defined by WHO as care organized for and provided to all women in a manner that maintains their dignity, privacy, and confidentiality, ensures freedom from harm and mistreatment, and enables informed choice and continuous support during labor and childbirth [23]. This project will be implemented in up to 10 districts in the state of Assam, with a total of 300 health care facilities. Initially, 5 facilities will be brought on board, followed by additional facilities in 3 districts and finally facilities across 10 districts by the end of 3 years. For monitoring and evaluation purposes, care providers and users will both be participants for the qualitative work, while the chatbot users will be part of the quantitative dipstick surveys.

Implementation

Understanding the urgent need to enhance MH outcomes in Assam, the Population Council Consulting Pvt. Ltd. (PCC), in collaboration with partners like Nivi Inc., Sitaram Bhartia Institute of Science and Research (SBISR), the Federation of Obstetric and Gynecological Societies of India, and the Centre for Northeast Studies and Policy Research, has launched the e-SAATHI initiative. This program aims to pilot and expand a chat-based digital support system for MH, designed to increase access to vital health information and strengthen maternal and newborn health across 10 districts in Assam through a comprehensive, multifaceted strategy.

e-SAATHI forms partnerships around a successful digital platform that empowers women to make informed choices about their reproductive health, promotes self-care, and connects users to reliable MH information and services throughout pregnancy and after childbirth. Given the obstacles pregnant women encounter in Assam, our approach is designed to both enhance and work alongside current public health programs related to MH. Most important, this solution helps reduce delays in seeking care by supporting women along their entire MH journey [24]. By maintaining engagement with expectant mothers from pregnancy through the postnatal period, the chat-based platform can help build greater awareness around healthy pregnancies (through timely and gestationally appropriate messages), encourage self-care practices, reinforce the value of ANC, facility-based births, and postnatal follow-up. It also facilitates timely connections with quality MH service providers, ensuring women get the safe and respectful care they need during this important period.

Chatbot

A key component of the e-SAATHI project is the chatbot messaging on MH information. It consists of a set of weekly tips and periodic reminders/check-ins that we send to subscribed users from prenatal week 10 through postnatal week 15. Women can subscribe to the free program in 1 of 2 ways: (1) clicking on a social media advertisement; or (2) scanning a QR code while visiting a health facility (public or private) that has engaged with the project. Both actions launch an automated conversation in WhatsApp that prompts the woman to onboard to the askNivi chatbot (accept terms and conditions and provide basic demographic information) and then subscribe to the ANC/PNC program (initially provide expected due date and preferred provider). Women who scan a QR code at a facility have the option of getting 1-on-1 support from a provider—typically a nurse—trained in user onboarding. The messages will be developed in alignment with the National Health Mission and WHO, ANC, and PNC guidelines (Figure 1).

XSL•FO
RenderX

Figure 1. Proposed digital chat-based pathway for maternal health. ANC: antenatal care; EDD: expected date of delivery; LMP: last menstrual period; PNC: postnatal care; WHO: World Health Organization.

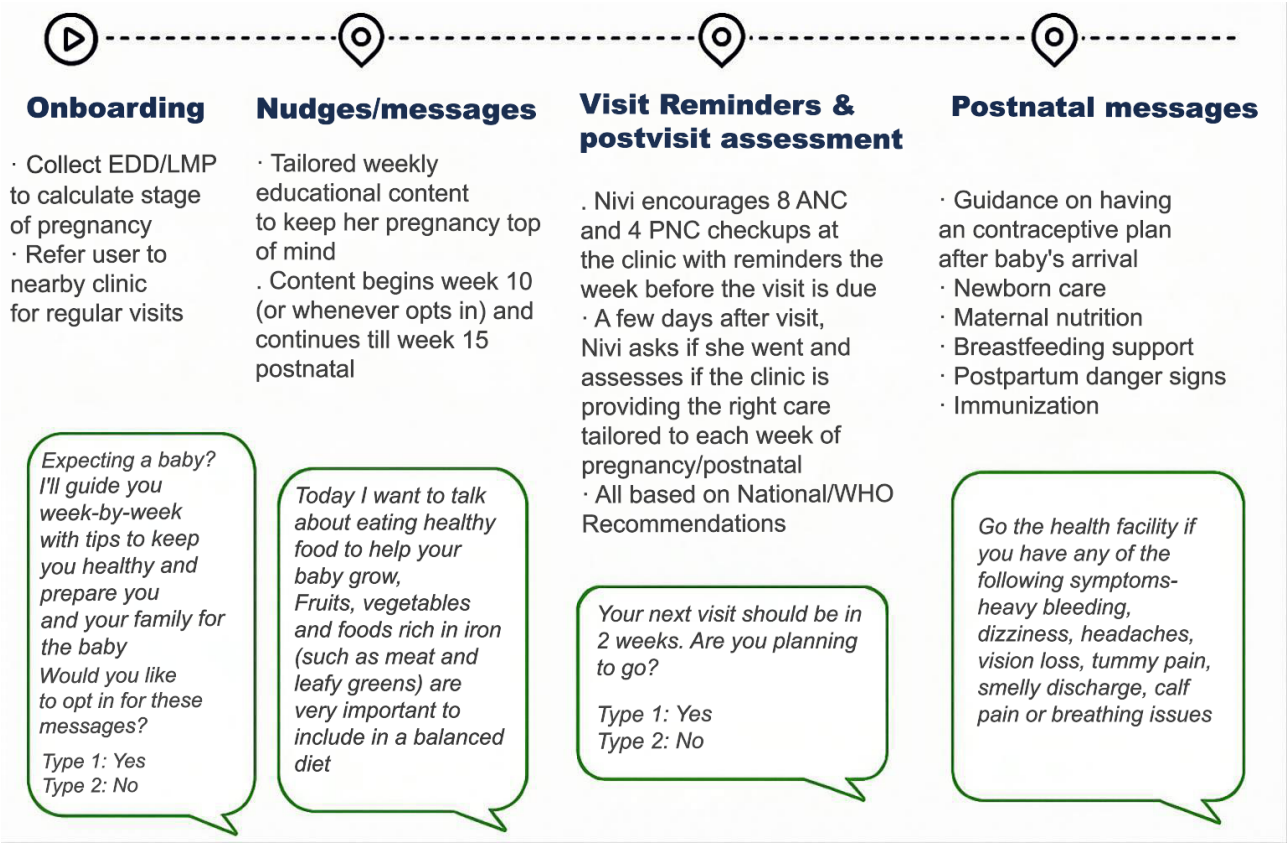

The intervention will be implemented over 3 distinct phases: (1) establish and operationalize MH chatbot (0 - 3 mo); (2) engagement, monitoring, and iterative improvement (4 - 24 mo); and (3) scale-up (25 - 36 mo).

Phase 1: Establish and Operationalize MH Chatbot (January 2023-March 2023)

The MH chatbot will be developed in partnership with gynecologists, adhering to ANC and PNC guidelines established by the WHO and the National Health Mission. To ensure that all content aligns with current evidence and clinical standards, a diverse panel of health professionals will be convened. Additionally, a Women's Ambassador Forum, comprising local community members—including pregnant and recently delivered women—will be formed to review the chatbot's content for clarity, contextual relevance, and completeness. Upon completion of development, pilot testing will be conducted with a select group of approximately 15 - 20 pregnant and postpartum women, including members of the Women's Ambassador Forum, to assess message comprehension, cultural appropriateness, and usability. Feedback from users and subject matter experts will continually inform ongoing refinements to message wording, sequencing, and delivery as content is periodically updated.

We will collaborate with health care providers in both private and public hospitals to facilitate the enrollment of pregnant and postnatal women onto the digital system via the WhatsApp chatbot. Once enrolled, participants will receive 2 to 3 tailored messages per week, providing guidance aligned with their stage of gestation or weeks postnatal—including reminders for visits to designated MH providers—and offering opportunities to request additional support or submit feedback regarding their care. Additionally, this solution enhances MH provider workflows by using the MH digital support system. It streamlines onboarding processes, delivers automated reminders and follow-up prompts for both users and providers, and implements structured feedback mechanisms for ongoing improvement. These features collectively improve operational efficiency, enable prompt engagement, and ensure user feedback is systematically integrated into service delivery.

Phase 2: Engagement, Monitoring, and Iterative Improvement (April 2023-September 2025)

Engagement

Onboarded women will be regularly engaged by health care providers or the research team whenever visiting the health facility to identify any challenges they face while using the chatbot. Additionally, feedback will be collected from users regarding the quality and usefulness of the services provided by the e-SAATHI chatbot.

Monitoring and Iterative Improvement

The monitoring and evaluation process will be conducted on users and health care providers. This study uses a mixed method, pre-post implementation design to evaluate the introduction and scale-up of the e-SAATHI MH chatbot in Assam. While this design enables comprehensive assessment of both quantitative outcomes and qualitative experiences, we acknowledge that the absence of randomization limits the strength of causal inference. To address potential sources of bias, we will implement several

XSL•FO
RenderX

strategies, including adjustment for district-level clustering, accounting for secular trends over time, and controlling for measured confounders in the analysis. These approaches aim to strengthen the validity of our findings and provide a nuanced understanding of the intervention's impact in real-world settings.

All the developed data collection instruments and procedures will be pretested in settings like those in which they will be administered. Guidance will be sought from postnatal women who have recently given birth when finalizing the instruments to ensure that questions and the data collection procedures are conducted with appropriate sensitivity and are not perceived as stigmatizing. The study tools will be pretested among a small group of women with characteristics similar to those of the study population to identify potential negative consequences and will be modified accordingly. Study tools for postnatal women will be translated and pretested in Assamese and Hindi.

The study comprises 5 main data collection tools that are described in Table 1: (1) in-depth qualitative interviews with MH providers; (2) in-depth qualitative interviews with postnatal women; (3) telephone interviews with postnatal women and MH chatbot users; (4) health facility inventory; and (5) MH digital platform program data.

Table . Data collection activities.[a]

	1. IDIs[b] with health care providers	2. IDIs with postnatal (PN[c]) women	3. Survey with PN women	4. Facility inventory and routine data	5. MH[d] data from chatbot platform
Study population	Purposively selected providers from various care levels (public and private)	Purposively sampled PN women using chatbot <6 mo since childbirth	PN women recently given birth in last 2 - 8 wk	Facility providing MH services	Pregnant and PN women onboarded to chatbot (deidentified)
Sample size	10 from range of facility levels periodically	2 - 4 women per facility—periodically	200 PN women	Each facility taking part in the project	All women who use chatbot
Location of activity	Private room at health facility or mutually agreed site	Private room at health facility or mutually agreed site	Recruitment at PNC[e] ward then short interview (2 - 8 wk later)	Facility level data (deidentified)	Online program data
Timing	Baseline and periodically 6 - 9 mo	Baseline, midline, and endline	Baseline, then periodically (9 mo after chatbot introduction)	Quarterly	From initiation of chatbot to end of project
Method	In-depth face-to-face interview by researcher	In-depth face-to-face interview by researcher	Short in-person/telephone survey by researcher	Facilities visited by researcher	Analysis of chatbot program data

[a]Analysis of baseline, midline, and endline data will occur after the completion of each respective phase; interim analyses will be performed after midline (2026) and final analyses after endline (2027).

[b]IDI: in-depth interview.

[c]PN: postnatal.

[d]MH: maternal health.

[e]PNC: postnatal care.

First, this includes any MH provider (at least 10) over 18 years, who consent to be interviewed and work in one of the public or private facilities in the study areas, nurses, midwives, auxiliary nurse midwives, Accredited Social Health Activists (ASHAs), and doctors at health facilities engaged in the digital MH support system. The purpose of the interviews with MH providers is to identify facility and provider-specific issues that may affect the delivery of care and uptake of women onboarding onto the chatbot. Researchers will use a short in-depth interview (IDI) guide to ask respondents about their expectations regarding the MH chatbot, its strengths and weaknesses, develop an in-depth understanding of the process of introducing the chatbot, and how it might be improved and subsequently scaled up. Data will be analyzed iteratively after each collection wave to inform ongoing implementation and comprehensively the completion of phase 2.

Second, this will be conducted with a subsample of women (n=10) who used the MH chatbot and accessed its services in a facility. These interviews will be used to gain a deeper understanding of the motivations, perceptions, and priorities of women using the chatbot and facility services. Researchers will use a short IDI guide to ask respondents on their experience of using the MH chatbot, its strengths and weaknesses, and how it might be improved. Women will be at least 18 years old, have consented to the interview, and have given birth in a facility within the preceding 6 months. Data will be analyzed iteratively after each collection wave to inform ongoing implementation and the comprehensive completion of phase 2.

Third, postnatal women who have given birth and have experienced MH journey and/or used *askNivi* chatbot (>18 y age) will be approached in a postnatal ward at a facility. Women who consent to participate will be contacted telephonically between 2 to 8 weeks after the date of delivery for a telephonic or in-person survey. The study team will note their identifying information for future contact separately from the data. During the recruitment or enrollment at the facility, the research assistant will ask permission to contact the woman. A member of the study team will call and agree on a time that suits them. That same member of the team will then call to interview at the conveniently decided time. The short dipstick survey will cover

XSL•FO
RenderX

the quality of care, experiences during pregnancy (number of ANC contacts, etc), childbirth and postnatally, and experience of using askNivi chatbot. Women who had loss or severe complications during childbirth were excluded. We will use the dipstick survey parameters field feedback to modify/adapt the composite quality score of maternal and newborn care as a composite measure designed to assess the quality of maternal and newborn health services received [24]. Each indicator will be scored based on whether the recommended standard will be met, typically using a binary (yes/no) scale. The scores for all indicators will then be summed to produce a total score, which will often be standardized to a percentage for ease of interpretation. This approach will allow for comparison of service quality across facilities and overtime and will help identify specific areas for improvement [24]. Quantitative analyses will be conducted after baseline, midline, and endline data collection points. Data cleaning and quality checks will occur fortnightly during collection with a comprehensive pass before midline and endline analyses.

Fourth, a short checklist including facility staffing levels; maternity unit capacity, medicines, supplies, and infrastructure; educational materials, job aids and protocols; health management information system and information and communication technology; facilitation and management of emergency referrals and transport; and other MH interventions available in the facility. This checklist will be administered to the in-charge or delegate of the facility. Statistics related to routine program data on utilization or uptake of maternal and newborn health services (ANC, PNC, births, and deaths) will be collected. Monthly trends in the numbers of new and continuing ANC and PNC clients will be obtained from facility records.

Fifth, this includes responses from women who have joined the MH chatbot (e-SAATHI), including demographic information and use of services. The data used is from all women who interact with the MH platform. Women who are on board with the chatbot permit their data to be used (deidentified).

The study uses a mixed methods approach to comprehensively document and monitor the rollout of the mental health chatbot for pregnant and postnatal women in Assam. Implementation is guided by continuous field feedback, data on onboarding and user engagement with the chatbot, the effectiveness of online campaigns, and qualitative and quantitative impact assessments. By adopting a participatory CLA methodology, the program integrates women's perspectives and feedback from the earliest proof of concept, beginning in 5 facilities and then expanding across 3 districts to ultimately reach 10 districts within 3 years. These iterative CLA cycles, paired with regular evaluations of operational and financial sustainability, foster a deeper understanding of implementation challenges, including acceptability and feasibility. Upon completion of phase 2, a detailed impact evaluation will assess the effectiveness of the intervention, ensuring that crucial insights and real-world evidence are retained throughout the study's progression. User engagement with the chatbot will be tracked as an automatically generated metric, supporting ongoing learning and adaptation to enhance the system. For the qualitative interviews and dipstick surveys, which are single, one-time interactions, with no follow-up assessments. However, we will systematically record and report the number of individuals who refuse participation or cannot be contacted at recruitment, along with reasons for refusal when volunteered. For the chatbot, we will also report the number of users who disengage before completing the interaction.

Phase 3: Scale-Up and Sustainability (October 2025-June 2027)

In the final phase, the intervention will be scaled up to cover health care facilities across all 10 districts. This gradual expansion is planned over a period of 3 years, allowing for iterative refinement of the approach, capacity building of providers, and sustained engagement with stakeholders at each stage of implementation. The study population includes 300 MH providers and managers (nurses, midwives, auxiliary nurse midwives, doctors, and community health workers [ASHAs]) in selected private and public facilities across 10 districts in Assam with diverse geographic settings, as well as state-level managers or policymakers. We also anticipate reaching 225,000 pregnant/postnatal women through the chatbot.

Data Management and Analysis

Overview

The data collection with postnatal women, providers, facility assessments, and service use statistics will be completed by trained research assistants using tablet computers. Upon completion of any electronic data collection form, all data will be immediately encrypted. Encrypted data cannot be decrypted without the unique private encryption key, to which only study investigators will have access. Encrypted survey data will be stored on research tablets until it can be transmitted to password-protected, study-specific servers using the local cellular data network. If cellular data service is unavailable at a given facility, the data collection team leader's laptop computer will function as a central synchronization and backup point until the team reaches an area where mobile data service is available.

Qualitative Analysis

Qualitative data from the IDIs will be transcribed verbatim, translated into English where appropriate, and typed in Word. The data will be analyzed using Nvivo15 (Version 15; Lumivero) [25]. The analysis will entail open coding and progressive categorization of issues based on inductive (examining relevant issues that emerge from the data) and deductive approaches (examining key issues identified at the design stage). Themes will be further modified as more issues emerge from the data analysis. Categories derived from the data will be further refined through analysis charts. Hard copies of study documents will be stored under lock and key in secure drawers/rooms at PCC in New Delhi. We will upload digital recordings and store them on password-protected computers in folders on a secure network. We will store all data for at least 3 years before being destroyed by PCC.

Quantitative Analysis

Raw quantitative data collected from the postnatal women's survey (dipstick survey) will be cleaned and checked for errors

XSL•FO
RenderX

and outliers. Data cleaning and quality checks will occur fortnightly during collection with a comprehensive pass before midline and endline analyses. Descriptive statistics (means, SDs, medians, and frequencies) will summarize baseline characteristics and outcomes. The primary outcome, the composite quality score of maternal and newborn care [24] and secondary outcomes (service uptake indicators, user-reported knowledge and self-care practices, and satisfaction with care) will be compared between baseline and follow-up.

Normality will be tested (Shapiro-Wilk). Parametric (*t* tests, ANOVA) or nonparametric (Wilcoxon and Mann-Whitney) tests will be used as appropriate. Categorical variables will be analyzed using chi-square or Fisher exact test. Multivariable analyses will use generalized linear/logistic mixed models to adjust for clustering (facility/district) and confounders (age, parity, socioeconomic status, dose of exposure-engagement frequency, and enrollment channel). For repeated measures, mixed-effects models will assess changes over time. Multiple comparisons will be adjusted using Bonferroni or FDR. All analyses will be performed in IBM SPSS Statistics (Version 29.0; IBM Corp) and R Core Team. Results will be visualized with tables and graphs. Statistical significance is set at $P<.05$, with effect sizes and 95% CIs reported.

Interim analyses will be conducted after midline data (Phase 2) collection completion in 2026; final analyses will be performed after endline data collection in 2027.

Data Confidentiality

Any personal, organizational, or identifier information will not be attached to the data, nor will it be linked to the reports in a way that is traceable to an individual. Specifically, no names will be included in the transcripts, facility, and village names (that may characterize a community) will be described vaguely as a "rural" or "peri-urban site." Pseudonyms or numbers will be assigned in naming stored data files to protect respondents from being identified. During dissemination of any type, all data will be presented in aggregate form so that names and other identifying information of a respondent will not be exposed—for example, an IDI with a postnatal woman will be classified in these ambiguous terms. Similarly, an IDI with a health provider will be described without explicitly naming a facility or area in which the data was collected. All data will be collected under the regulation of the data protection and management guidelines provided by the Government of India.

Ethical Considerations

The research protocol has been reviewed by key stakeholders and ethical clearance has been granted by the Population Council's Institutional Review Board (517), and the SBISR Ethics Committee (F.2/SBISR-EC/PR 01 - 23). All researchers and research assistants will be trained on the conduct of ethical procedures and will be monitored during field work by the coalition led by PPC/SBISR. Informed consent will be obtained separately for each study participant and each component. All participants will be given detailed information about the study, including aims and methods of study; institutional affiliations of the research; anticipated benefits, risks, and discomfort it may cause (expected to be minimal); follow-up of the study; expected duration of an interview; a participant's choice not to answer any questions; that they have the right to abstain from participating in the study or to withdraw from it at any time, without reprisal; measures taken to ensure confidentiality and anonymity of information provided; the conduct of interviews in places chosen by a participant that may maximize audio privacy; and contact details of the study coordinator for any questions or concerns. All data will be stored in password-protected computer files. Hard copies of questionnaires and anonymized transcriptions will be stored securely in a locked cabinet, in accordance with the PCC and SBISR policy. No compensation will be provided to participants.

Results

The e-SAATHI study was formally funded in September 2022, marking the commencement of project activities. Since the initiation of the program, significant progress has been achieved in both the onboarding of health care facilities and participant recruitment. As of December 2025, a total of 210 health facilities across Assam have been successfully integrated into the e-SAATHI network. A total of 201,813 women have been enrolled to receive MH support through the chatbot platform.

Chatbot-based data collection was launched in April 2023 and is scheduled to continue through June 2027, ensuring comprehensive coverage of the intervention period. The collection of qualitative and quantitative evaluation data began in November 2023, providing a robust foundation for assessing both user experiences and measurable outcomes.

Interim analyses are planned following the completion of midline data collection in 2026, which will offer early insights into the effectiveness and operational feasibility of the intervention. Final analyses are scheduled to take place after endline data collection in 2027, at which point the full impact of the e-SAATHI program will be evaluated and published. The primary outcome of interest is the change in the composite quality score of maternal and newborn care, serving as a key indicator of the intervention's effectiveness. Secondary outcomes include a range of service uptake indicators, user-reported knowledge and self-care practices, and satisfaction with care received. The study will also assess operational feasibility, including the integration of providers into the digital system and the identification of barriers such as digital literacy and connectivity challenges.

As of the writing of this paper (December 2025), we are completing phase 2 and initiating in-depth impact evaluation.

Discussion

e-SAATHI's Implications

The introduction of the e-SAATHI intervention represents a significant advance in the realm of MH care, particularly within low-resource settings such as Assam, India. Assam has the highest maternal mortality ratio in India at 195 per 100,000 live births, nearly double the national average [1]. Given this context, there exists an urgent need for innovative solutions to address systemic gaps and improve MH outcomes. mHealth interventions, including chatbots and digital communication

XSL•FO
RenderX

platforms, have demonstrated promise in improving MH care use in LMICs [26]. The implementation of e-SAATHI aligns with these global priorities, offering tailored health information, decision-making support, and continuous engagement to pregnant and postpartum women through digital channels. The observed changes in the composite quality score of maternal and newborn care [24] and domains will provide robust evidence of improvements in the quality and comprehensiveness of MH services. Secondary outcomes, including increased service uptake and improved knowledge/self-care, will further support the effectiveness and scalability of the intervention.

The e-SAATHI protocol leverages global evidence on the efficacy of mHealth tools in promoting maternal care use. Meta-analyses have consistently highlighted that digital health interventions such as chatbots and mobile messaging platforms significantly enhance key MH metrics, including ANC visits and facility-based delivery rates [26]. As a chatbot-driven initiative, e-SAATHI addresses longstanding gaps in MH education and accessibility, particularly in underserved regions. By providing stage-specific health information tailored to users' gestation periods, the intervention reinforces global efforts aimed at improving maternal knowledge, autonomy, and continuity of care [27].

Global Context and Precedent for Digital MH Interventions

In LMIC contexts, the adoption of digital health interventions has proven effective in mitigating barriers to health care engagement [4]. South Asia, in particular, has witnessed the large-scale implementation of mobile messaging programs for MH. For instance, India's Kilkari program delivered weekly audio messages related to maternal and child health to over 10 million subscribers, improving pregnancy-related knowledge and self-care practices among participants [28]. While challenges such as subscriber retention and equitable reach were reported, the program underscored the feasibility of deploying gestation-tailored messaging as a complement to routine antenatal education. Similarly, Bangladesh's "Aponjon" initiative provided stage-specific information to pregnant women via mobile devices, contributing to improved care-seeking behaviors, though coverage gaps persisted in hard-to-reach areas [29]. These findings emphasize the critical role of interactive, stage-specific messaging in augmenting MH services.

Sub-Saharan African countries provide additional examples of successful mHealth interventions. Chatbot and mobile messaging services have been widely implemented to support maternal and neonatal health, demonstrating high user demand and acceptability. African Researchers Magazine, 2023 [30] highlights the transformative impact of health chatbots in improving care-seeking behaviors among underserved populations. Such global precedents reinforce the rationale for implementing e-SAATHI, which is grounded in established evidence and adapted to local needs.

Feasibility and Public Health Significance

The feasibility of e-SAATHI is supported by the widespread penetration of mobile phones and growing accessibility to internet services in LMICs. In India, over 70% of women own mobile phones, with an increasing proportion gaining access to smartphones and basic internet services [26]. High mobile penetration rates in Africa and Latin America have similarly facilitated the scalability of digital MH interventions. The public health significance of such initiatives is substantial, as they address critical gaps in health systems constrained by limited human resources. Women in low-resource settings often experience brief, infrequent clinic visits and overburdened health care providers, leading to missed opportunities for counseling and education. Chatbot programs such as e-SAATHI partially bridge this gap by delivering timely and personalized health messages, encouraging engagement with ANC services, immunizations, and institutional deliveries [30]. By supporting informed decision-making and early care-seeking behaviors, e-SAATHI contributes to achieving national and global targets for maternal and neonatal mortality reduction.

Qualitative evidence from prior interventions suggests that digital messaging services provide psychological support and empowerment to users. Women perceive these tools as continuous sources of information, which improves their confidence in navigating pregnancy-related challenges [26]. e-SAATHI's interactive features, including quizzes and 2-way prompts, are designed to enhance user engagement and memory retention, particularly among individuals with limited literacy.

Addressing Digital Inclusion and Equity

Despite the promise of digital health solutions, challenges related to equity and inclusion must be addressed. Vulnerable subgroups, such as women with low literacy levels or those residing in remote areas without smartphones, may face barriers in accessing e-SAATHI's services. These concerns are explicitly addressed in the intervention's design. The chatbot platform is delivered via widely available messaging apps such as WhatsApp, ensuring accessibility for diverse user groups. Content is provided in local languages, including Assamese and Hindi, to maximize comprehension and cultural relevance. Collaborations with ASHAs and nurses further enhance trust and engagement among users. This blended approach mirrors the successful integration of digital and human-mediated services observed in South Africa's MomConnect initiative [31,32].

Infrastructure obstacles, such as poor connectivity in rural areas, are mitigated through asynchronous messaging capabilities, allowing users to access content even in regions with intermittent internet access. By prioritizing inclusive design features, e-SAATHI ensures its accessibility to underserved populations, thereby enhancing its scalability and impact.

Study Strengths and Limitations

Strengths

The e-SAATHI implementation stands out for its innovative use of a scalable, chat-based digital platform to deliver personalized MH information, a particularly valuable approach in resource-constrained regions like Assam where it addresses persistent gaps in MH education and access. The intervention was co-designed in partnership with local community members, health care providers, and national experts, ensuring that it is contextually relevant, culturally appropriate, and responsive to

XSL•FO
RenderX

user needs. Using a mixed methods evaluation, the study integrates qualitative and quantitative methodologies to capture both measurable outcomes and nuanced experiences of users and providers, thereby enhancing the overall assessment of feasibility, acceptability, and effectiveness. Throughout the study, robust ethical and data security protocols—including thorough informed consent processes, stringent data confidentiality procedures, and comprehensive ethical oversight—safeguard participant privacy and uphold the integrity of all collected data.

Limitations

The implementation of the e-SAATHI intervention faces several limitations. First, digital access barriers remain a concern, as users in areas with limited or no internet connectivity may be unable to access the chatbot. Some pregnant and postnatal women may not have smartphones or internet access, potentially excluding those who most need information. Although mobile phone density is generally adequate in the study districts, alternative delivery mechanisms—such as dissemination through ASHAs or other media—are being explored to address this gap. Additionally, systemic and operational challenges could arise, as the alignment of partner facilities and district health officials with the intervention's goals may vary, potentially causing delays. Microplanning differences in implementation could also result in varied onboarding and user experiences across sites. The study design presents another constraint; while the pre-post monitoring and evaluation framework is practical for real-world application, it limits the strength of causal inferences. Without randomization, observed changes may be influenced by external factors or secular trends, despite efforts to adjust for confounders and clustering. There is also the potential for selection bias, as women who are able to access and use the chatbot may differ systematically from those who cannot, which could affect the generalizability of the findings.

Conclusion and Future Implications

The e-SAATHI intervention exemplifies a scalable and inclusive approach to improve MH outcomes in Assam and beyond. Its design reflects lessons from global experiences, emphasizing digital inclusion, cultural adaptation, and user engagement. While its effectiveness remains to be evaluated, the intervention demonstrates potential to complement existing MH initiatives, such as cash incentive programs and community health worker outreach. The implementation protocol includes mixed methods evaluations to capture quantitative metrics, user experiences, and integration within health systems. These insights will inform the scalability and sustainability of chatbot-based interventions for MH care.

By addressing persistent gaps in MH services, e-SAATHI contributes to the broader objective of equitable health care access. Its participatory design and evidence-based approach position it as a promising addition to the MH toolkit in India and similar LMIC contexts. Continued investigation and adaptation of the intervention will enhance its relevance and impact, advancing global efforts to improve maternal and neonatal health outcomes.

Acknowledgments

The Strengthening Antenatal Care/Postnatal Care via AskNivi Tailored Health Information, Referrals, and Follow-Up is a coalition of Population Council Consulting, in partnership with Nivi, Inc., and Sitaram Bhartia Institute of Science and Research, in collaboration with the Federation of Obstetric and Gynecological Societies of India, Guwahati ObGyn Society, and the Centre for Northeast Studies and Policy Research. We acknowledge Jacquelyn Caglia and Iyadunni Olubode for reviewer-level inputs and strategic guidance.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The authors declare that this study received funding from Merck Sharp & Dohme (MSD) through its MSD for Mothers initiative (grant MFM-23-173706). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication. MSD for Mothers is an initiative of Merck & Co., Inc., Rahway, NJ, USA

Data Availability

The data sets generated during this study are not publicly available due to privacy and confidentiality concerns but are available from the author on reasonable request.

Authors' Contributions

Study conception and design: SR, JN, CEW
Writing—Original draft: AM
Review and editing: AA, SwR, CEW, BB
Drafting or significant revision of the manuscript: all authors
Approval of the manuscript submission: all authors.

XSL•FO
RenderX

Conflicts of Interest

BB is a cofounder and Chief Business Officer of Nivi, Inc., which developed and operates the askNivi platform evaluated in this study. BB holds an ownership stake in Nivi, Inc. All other authors declare no conflicts of interest.

References

1. Meh C, Sharma A, Ram U, et al. Trends in maternal mortality in India over two decades in nationally representative surveys. BJOG 2022 Mar;129(4):550-561. [doi: 10.1111/1471-0528.16888] [Medline: 34455679]
2. Faridi SB, Selvam I. Maternal health situation in Uttar Pradesh, India. Eras J Med Res 2022 Dec;9(2):239-242. [doi: 10.24041/ejmr2022.37]
3. Purohit T, Purohit P, Satender C, Batwal P. A correlation analysis between maternal health service indicators aggregate scoring and maternal mortality ratio: a quantitative study of states in India. SN Compr Clin Med 2024;7(1):15. [doi: 10.1007/s42399-024-01758-3]
4. Kayaroganam R, Saya GK, Kar SS. Utilization of maternal health services among Janani Suraksha Yojana beneficiaries in Puducherry, India. Int J Adv Med Health Res 2016;3(2):73. [doi: 10.4103/2349-4220.195935]
5. Sen S, Chatterjee S, Khan PK, Mohanty SK. Unintended effects of Janani Suraksha Yojana on maternal care in India. SSM Popul Health 2020 Aug;11:100619. [doi: 10.1016/j.ssmph.2020.100619] [Medline: 32642548]
6. Tripathi P, Chakrabarty M, Singh A, Let S. Geographic disparities and determinants of full utilization of the continuum of maternal and newborn healthcare services in rural India. BMC Public Health 2024 Dec 5;24(1):3378. [doi: 10.1186/s12889-024-20714-3] [Medline: 39639301]
7. Bhushan H, Ram U, Scott K, et al. Making the health system work for over 25 million births annually: drivers of the notable decline in maternal and newborn mortality in India. BMJ Glob Health 2024 May 6;9(Suppl 2):e011411. [doi: 10.1136/bmjgh-2022-011411] [Medline: 38770806]
8. Hamal M, Dieleman M, De Brouwere V, de Cock Buning T. Social determinants of maternal health: a scoping review of factors influencing maternal mortality and maternal health service use in India. Public Health Rev 2020;41(1):13. [doi: 10.1186/s40985-020-00125-6] [Medline: 32514389]
9. Purbey A, Kumar A, Mozumdar A, Mishra P, Acharya R, Saggurti N. Trends in the utilisation of maternal and child healthcare services from the public and private health sectors in India, 2005-2021: an analysis of cross-sectional survey data. BMJ Open 2025 Feb 7;15(2):e082241. [doi: 10.1136/bmjopen-2023-082241] [Medline: 39920061]
10. Gebremedhin TA, Mohanty I, Niyonsenga T. Public health insurance and maternal health care utilization in India: evidence from the 2005-2012 mothers' cohort data. BMC Pregnancy Childbirth 2022 Feb 25;22(1):155. [doi: 10.1186/s12884-022-04441-4] [Medline: 35216564]
11. Lusome R, Sivan A, Kumar MA. Out of pocket expenditure on institutional deliveries in India. Matern Child Health J 2025 Mar;29(3):386-395. [doi: 10.1007/s10995-025-04060-3] [Medline: 39847255]
12. Mondal D, Karmakar S, Banerjee A. Women's autonomy and utilization of maternal healthcare in India: evidence from a recent national survey. PLoS ONE 2020;15(12):e0243553. [doi: 10.1371/journal.pone.0243553] [Medline: 33296428]
13. Scott K, Ummer O, Shinde A, et al. Another voice in the crowd: the challenge of changing family planning and child feeding practices through mHealth messaging in rural central India. BMJ Glob Health 2021 Jul;6(Suppl 5):e005868. [doi: 10.1136/bmjgh-2021-005868] [Medline: 34312156]
14. Murthy N, Chandrasekharan S, Prakash MP, et al. The impact of an mHealth voice message service (mMitra) on infant care knowledge, and practices among low-income women in India: findings from a pseudo-randomized controlled trial. Matern Child Health J 2019 Dec;23(12):1658-1669. [doi: 10.1007/s10995-019-02805-5] [Medline: 31584144]
15. Kabongo EM, Mukumbang FC, Delobelle P, Nicol E. Understanding the influence of the MomConnect programme on antenatal and postnatal care service utilisation in two South African provinces: a realist evaluation protocol. BMJ Open 2019 Jul 1;9(7):e029745. [doi: 10.1136/bmjopen-2019-029745] [Medline: 31266842]
16. Ranganathan R. Towards a holistic digital health ecosystem in india. : Observer Research Foundation (ORF); 2020 URL: https://www.orfonline.org/public/uploads/posts/pdf/20230524152812.pdf [accessed 2026-01-24]
17. Muthappan S, Elumalai R, Shanmugasundaram N, et al. AYUSH digital initiatives: harnessing the power of digital technology for India's traditional medical systems. J Ayurveda Integr Med 2022;13(2):100498. [doi: 10.1016/j.jaim.2021.07.014] [Medline: 34969591]
18. Patel K, Nayak B, Behera TR, et al. Factors influencing, SWOT analysis and emerged strategies for improving the use of digital data platforms in immunisation program by frontline health care providers: a study from rural eastern India. Int J Community Med Public Health 2022;9(11):4091. [doi: 10.18203/2394-6040.ijcmph20222901]
19. Rasekaba TM, Pereira P, Rani G V, Johnson R, McKechnie R, Blackberry I. Exploring telehealth readiness in a resource limited setting: digital and health literacy among older people in rural India (DAHLIA). Geriatrics (Basel) 2022 Mar 1;7(2):28. [doi: 10.3390/geriatrics7020028] [Medline: 35314600]
20. Bahja M, Abuhwaila N, Bahja J. An antenatal care awareness prototype chatbot application using a user-centric design approach. 2020 Presented at: HCI International 2020-Late Breaking Papers: Multimodality and Intelligence: 22nd HCI International Conference, HCII 2020; Jul 19-24, 2020; Copenhagen, Denmark p. 20-31. [doi: 10.1007/978-3-030-60117-1_2]

XSL•FO
RenderX

21. Grenier L, Suhowatsky S, Kabue MM, et al. Impact of group antenatal care (G-ANC) versus individual antenatal care (ANC) on quality of care, ANC attendance and facility-based delivery: a pragmatic cluster-randomized controlled trial in Kenya and Nigeria. PLoS ONE 2019;14(10):e0222177. [doi: 10.1371/journal.pone.0222177] [Medline: 31577797]
22. Rahman MO, Yamaji N, Nagamatsu Y, Ota E. Effects of mHealth interventions on improving antenatal care visits and skilled delivery care in low- and middle-income countries: systematic review and meta-analysis. J Med Internet Res 2022 Apr 22;24(4):e34061. [doi: 10.2196/34061] [Medline: 35451987]
23. WHO recommendations on intrapartum care for a positive childbirth experience. : World Health Organization; 2018 URL: https://iris.who.int/server/api/core/bitstreams/ba043cf7-cba4-484d-bf7e-ec79c4102d54/content [accessed 2026-01-24]
24. Lee HY, Rana MJ, Kim R, Subramanian SV. Small area variation in the quality of maternal and newborn care in India. JAMA Netw Open 2022 Nov 1;5(11):e2242666. [doi: 10.1001/jamanetworkopen.2022.42666] [Medline: 36441555]
25. NVivo (version 15). Lumivero. 2025. URL: https://lumivero.com/products/nvivo/ [accessed 2026-02-03]
26. Ramesh S, Warren CE, Bellows B, et al. Leveraging health financing, digital health and self-care approaches to strengthen maternal health journeys in India: perspectives from Assam. Front Glob Womens Health 2025;6:1469328. [doi: 10.3389/fgwh.2025.1469328] [Medline: 40270690]
27. Muthelo L, Mbombi MO, Bopape MA, et al. Reflections on digital maternal and child health support for mothers and community health workers in rural areas of Limpopo Province, South Africa. Int J Environ Res Public Health 2023 Jan 19;20(3):1842. [doi: 10.3390/ijerph20031842] [Medline: 36767230]
28. Mohan D, Bashingwa JJH, Scott K, et al. Optimising the reach of mobile health messaging programmes: an analysis of system generated data for the Kilkari programme across 13 states in India. BMJ Glob Health 2022 Aug;6(Suppl 5):e009395. [doi: 10.1136/bmjgh-2022-009395] [Medline: 35940611]
29. Alam M, Banwell C, Olsen A, Lokuge K. Patients' and doctors' perceptions of a mobile phone-based consultation service for maternal, neonatal, and infant health care in Bangladesh: a mixed-methods study. JMIR mHealth uHealth 2019 Apr 22;7(4):e11842. [doi: 10.2196/11842] [Medline: 31008716]
30. Phiri M, Munoriyarwa A. Health chatbots in Africa: scoping review. J Med Internet Res 2023 Jun 14;25:e35573. [doi: 10.2196/35573] [Medline: 35584083]
31. Ochieng H, Awosiku O. An overview of healthcare chatbots in Africa. Digital Health Africa. URL: https://digitalhealth-africa.org/an-overview-of-healthcare-chatbots-in-africa/ [accessed 2026-01-24]
32. Verma T. Using AI to improve maternal health chatbots in South Africa. IDinsight. 2022. URL: https://www.idinsight.org/article/using-ai-to-improve-maternal-health-chatbots-in-south-africa/ [accessed 2026-01-24]

Abbreviations

ANC: antenatal care
ASHAs: Accredited Social Health Activists
CLA: collaborative learning and adapting
e-SAATHI: Strengthening ANC/PNC via AskNivi Tailored Health Information, Referrals, and Follow-Up
IDI: in-depth interview
LMICs: low-income and middle-income countries
MH: maternal health
mHealth: mobile health
PCC: Population Council Consulting Pvt. Ltd.
PNC: postnatal care
SBISR: Sitaram Bhartia Institute of Science and Research
WHO: World Health Organization

Edited by J Sarvestan; submitted 05.Aug.2025; peer-reviewed by A Dozier; revised version received 02.Jan.2026; accepted 05.Jan.2026; published 13.Feb.2026.

Please cite as:
Munjral A, Rawat S, Warren CE, Nagpal J, Bellows B, Aditi A, Sridhar P, Ramesh S
Chat-Based Decision Support System for the Maternal Health Journey in Assam, India: Protocol for a Mixed Methods Multiphase Implementation Study
JMIR Res Protoc 2026;15:e81873
URL: https://www.researchprotocols.org/2026/1/e81873
doi:10.2196/81873

XSL•FO
RenderX

© Ashita Munjral, Swapnil Rawat, Charlotte E Warren, Jitender Nagpal, Ben Bellows, Aditi Aditi, Pompy Sridhar, Sowmya Ramesh. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Developing, Implementing, and Evaluating a Care Management Model for Patients With Neuromuscular Diseases: Protocol for a Participatory Mixed Methods Study

Veronika Waldboth[1], PhD; Christina Schuler[1], MSc; Raffaella Willmann[2], PhD; Heidi Petry[3], PhD; Markus Weber[4], MD; Barbara Grädel Messerli[5], MSc; Hannele Hediger[1], MSc; Iris Kramer[1], MSc; Gabriela Stettler-Nemecek[1], MSc; Martin Knoblauch[6], lic.oec; Georg Martin Stettner[7], MD, PD

[1]Department of Health Sciences, Institute of Nursing, ZHAW Zurich University of Applied Sciences, Winterthur, Switzerland

[2]Swiss Foundation for Research on Muscle Diseases, Colombier, Switzerland

[3]Centre for Clinical Nursing Science, University Hospital Zurich, Zurich, Switzerland

[4]Neuromuscular Diseases Unit/ALS Clinic, Health Ostschweiz (HOCH), Cantonal Hospital St. Gallen, St. Gallen, Switzerland

[5]Medical division for children and adolescents, Bern University Hospital, Bern, Switzerland

[6]Swiss ALS Foundation, Winterthur, Switzerland

[7]Neuromuscular Center Zurich and Department of Pediatric Neurology, University Children's Hospital Zurich, Zurich, Switzerland

Corresponding Author:
Veronika Waldboth, PhD
Department of Health Sciences
Institute of Nursing
ZHAW Zurich University of Applied Sciences
Katharina Sulzer Platz 9
Winterthur, 8401
Switzerland
Phone: 41 58 934 6499
Email: veronika.waldboth@zhaw.ch

Abstract

Background: Neuromuscular diseases (NMDs) are often progressive conditions that can reduce quality of life and life expectancy and require long-term caregiving. As patients age, caregivers face increasing demands, often resulting in significant caregiver burden. Treatment and care are complex, and a coordinated, family-centered approach is indicated. The current care situation for individuals with NMDs and their families in Switzerland is poorly understood, and there is no standardized NMD care management available.

Objective: The proposed study aims to assess the current care practices for individuals with NMDs in Switzerland, identifying unmet needs and challenges faced by patients, families, and health care providers. On the basis of these findings, a care management model will be developed, implemented, and evaluated to enhance support structures, strengthen specialized neuromuscular centers, and improve overall health care outcomes for affected individuals and their families.

Methods: We planned a 3-phase participatory mixed methods study (the Care-NMD-CH study). First, qualitative descriptive data and quantitative survey data were used to assess the current state of care practices. Then, a care management model and training program were codeveloped based on these insights. Finally, the care management service was implemented in pilot specialized neuromuscular centers and evaluated using quantitative, qualitative, and integrated methods. In addition to interview data and patient records, we measured patient-reported outcomes (quality of life and self-efficacy), family-reported outcomes (family functioning and caregiver burden), and stakeholder-reported outcomes (quality of care and interprofessional collaboration). The study population comprises individuals with NMDs and their families living in Switzerland.

Results: Ethics approval was obtained in 2020. Baseline data (T0) were collected and analyzed in 2021 (phase A). The NMD care management model (phase B) was developed based on these data and finalized in 2022. Following the start of implementation in 8 centers in 2023, data collection for process and outcome evaluation (T1-T4) commenced and was completed in May 2025. Data analysis is currently ongoing, and results will be published separately. This paper reports only the study protocol.

XSL•FO
RenderX

Conclusions: This study represents an important step toward implementing and evaluating an evidence-based, family-centered care management service for individuals with NMDs and their families in specialized neuromuscular centers. Building on a 3-phase approach, the project identified gaps in current care; developed a structured care management model, including training; and evaluated its impact on patients, family members, and interprofessional care teams. Expected outcomes include possible improvements in the areas of care coordination, quality of life, and self-efficacy and family functioning. In addition, potential alleviation of disease and caregiver burden, as well as reduction in health care costs for those affected and the health system, will be investigated.

International Registered Report Identifier (IRRID): DERR1-10.2196/82833

(JMIR Res Protoc 2026;15:e82833) doi:10.2196/82833

KEYWORDS

neuromuscular diseases; rare diseases; care management; caregivers; quality of life; caregiver burden; mixed methods; study protocol; Switzerland

Introduction

Many neuromuscular diseases (NMDs) follow a degenerative course. They can be associated with a high burden of disease, reduced quality of life, and shortened life expectancy [1,2]. However, owing to new disease-modifying therapies, improvements in symptomatic treatment, and advanced health technologies, NMDs now often have a better prognosis than ever before [3,4]. Despite these advancements, individuals with NMDs still require regular therapies as well as assistance in activities of daily living. In many families with affected children or adults, family members take over caregiving responsibilities [5-7]. As individuals with NMDs grow older, their caregivers must continue providing care for longer periods, which poses additional practical and psychosocial challenges [8,9]. The prolonged and intensive caregiving often leads to an increased caregiver burden [6,10,11]. Individuals with NMDs, their families, and the health system are often faced with considerable financial burdens [12-14].

Treatment and care of patients with NMDs is inherently complex and requires a coordinated, multidisciplinary, and interprofessional approach [15-17]. On the basis of disease stage, symptom burden, and personal priorities, the needs of individuals with NMDs and their family members can change constantly [17]. Therefore, families benefit from targeted support, assistance with care, care coordination, and family-strengthening interventions [5,7,18]. This is particularly relevant during health crises and important stages, such as the loss of ambulation, the affected young person's transition into adulthood, and the end-of-life phase [6,19]. It is strongly advised to use a systemic approach when caring for the well-being of an individual with an NMD and their entire family.

Strengthening the support system for individuals living with these rare but often serious diseases is essential to ensure a sustainable care model for the future. Internationally recognized care guidelines exist for specific NMDs, emphasizing the necessity of multidisciplinary models that incorporate medical specialists and varied health care professional groups [20,21]. Individuals with NMDs benefit from a multidisciplinary approach through more comprehensive support, including clinical follow-up and resource use, improving their quality of life [22,23]. Care management services provided by health centers can support care coordination, improve the quality of life of frequent health care users, and enhance health care provider satisfaction, all while potentially reducing costs to the health care system [24,25]. Care management prioritizes interdisciplinary collaboration and effective communication among all the professionals involved in patient care to plan, implement, coordinate, evaluate, and prioritize services based on patient needs [26,27]. Through care management, individuals living with NMDs are empowered, and their self-efficacy and ability to self-manage their condition can improve [25,28,29]. The establishment and translation of evidence-based care management into practice would greatly benefit individuals with NMDs, as well as their families, by addressing their needs throughout all phases of care [30]. There is a need for an evidence-based, coordinated, multidisciplinary, and family-centered approach to care for the NMD population [20,30,31] to strengthen the support system for patients living with these rare and often severe conditions and create a sustainable care situation despite limited financial resources in the future.

However, in Switzerland, there is no standardized NMD care management available, and the current care situation and organization for individuals with NMDs is poorly understood. Given these circumstances, a comprehensive analysis and description of the current care practices for individuals with NMDs in Switzerland are crucial to identify effective strategies and areas for improvement.

The objectives of this study are to (1) describe the current state of care practices for individuals living with NMDs and identify unmet needs and challenges related to patients, their families, and the care team; (2) develop a care management model and training program for NMD care managers capable of addressing identified gaps in the support chain and enhancing the support offered by specialized neuromuscular centers; and (3) implement the standardized care management model and evaluate its impact on patients with NMDs and their families, on the interprofessional care team, and at the systems level.

Methods

Overview

This study uses a participatory mixed methods design that includes qualitative and quantitative data collection at different stages [32]. This design is suitable for studies that address health

XSL•FO
RenderX

disparities through empowerment of marginalized or underrepresented populations, involve the target population in the research [32], and examine complex phenomena such as the evaluation of services in which various perspectives need to be included [33,34]. This study uses a pragmatic philosophical stance and a family systems theoretical framework [35]. Data integration occurs through methods (building and merging) and during interpretation (joint display and narrative) [32].

Setting and Participants

Initial steps involved identifying the study cohort; establishing an expert panel (physicians, advanced practice nurses, and researchers) and a sounding board (affected individuals, family members, professionals, and other stakeholders) for feedback, guidance, and quality control; and creating a participant recruitment plan. The study cohort for the Care-NMD-CH study comprised patients with NMDs and their families living in Switzerland. In Switzerland, 7 neuromuscular centers provide specialized outpatient and inpatient care for individuals with NMDs. These centers are hospital-based tertiary national reference centers accredited by the National Coordination of Rare Diseases. Four groups of study participants were identified: (1) patients with NMDs, (2) family members (parents, siblings, partners, and other relatives), (3) professionals who are involved in the care, and (4) other stakeholders (eg, nurses, physiotherapists, and members of patient organizations). The study included participants from all linguistic regions of Switzerland. Recruitment included the identification of collaborating institutions and gatekeepers, who informed potential participants about study participation.

Study Phases

Overview

According to the study aims, there were 3 phases (A-C) over a duration of 5 years (2020-2025; Figure 1). In phase A, qualitative and quantitative data were collected using instruments tailored to the needs of the individuals being studied to assess the current state of care practice [36]. In phase B, a new care management model and training program for NMD care managers was developed involving relevant stakeholders, including patients with NMDs, their families, and health care professionals. In phase C, the new care management model was implemented in pilot specialized neuromuscular health centers, and processes and outcomes were evaluated using a pretest and posttest design. The following criteria were used to reflect on the quality of the study: rigor, credibility, and transferability [37-39].

Figure 1. Overview of the study phases.

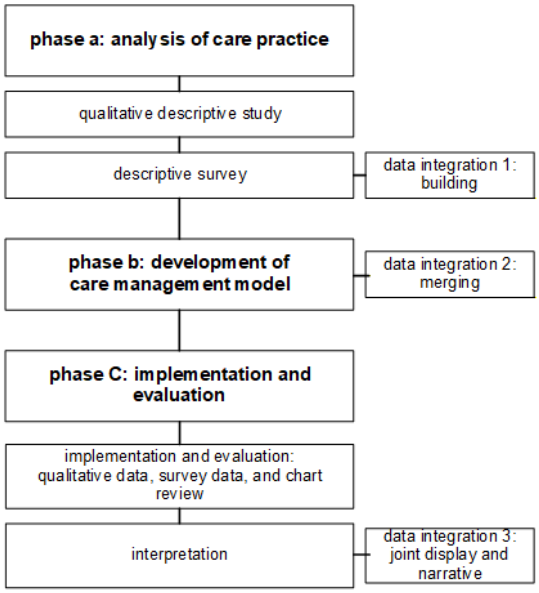

Study Phase A: Analysis of the Current State of Care Practice

Qualitative Descriptive Study

At first, a qualitative descriptive research design was applied to explore and describe the current state of care practice [40,41]. This approach facilitates close engagement with the data and enables a subjective exploration of people's everyday experiences and needs [42,43]. A purposive sampling strategy was used to recruit participants who met the inclusion criteria and were able to provide in-depth information addressing the research question [44]. Data were collected through one-on-one and focus group interviews with patients with NMDs, family members, professionals who were involved in their care, and other stakeholders (planned: n=45). For ethical reasons, patients and family members participating in interviews were aged ≥8 years.

Separate semistructured interview guides were developed for each participant group containing age-appropriate open-ended and probing questions. The questions assessed the general care situation, benefits of care, gaps in the support chain, unmet needs, experiences within the care network, barriers to sustainability, communication, and collaboration. Audio data were transcribed verbatim according to the guidelines set by Dresing and Pehl [45]. The latest version of MAXQDA (VERBI GmbH) was used for data analysis [46]. The transcripts were analyzed using the qualitative data analysis principles outlined by Saldaña [47]. The inductive analysis process involved four

XSL•FO
RenderX

steps: (1) open coding without using prior concepts or predefined categories, (2) code allocation into main areas, (3) bundling into subcategories, and (4) comparison and revision of subcategories and generation of categories. The sociodemographic data of participants were collected and analyzed descriptively.

Descriptive Survey

The quantitative part of phase A consisted of an online descriptive survey to collect further information on the current practices of care from a larger group of participants. The results from the qualitative study were used to develop the survey components. This was the first data integration point in our mixed methods study, where results from one data collection phase (qualitative) informed the next data collection phase (quantitative), an integration approach referred to as "building" [32,36].

Quantitative Data Collection

The online survey included self-developed questions based on the obtained qualitative results and standardized instruments (Table 1). Four questionnaires were developed, one for each group of participants (patients, family members, professionals, and other stakeholders; planned sample: n=500). Patients and family members who were expected to complete the online survey independently were aged ≥14 years. The questionnaires consisted of Likert-type questions, visual analog scales, and closed- and open-ended questions. Topics common to all participant groups included sociodemographics, information about the care situation, satisfaction with care, gaps in the support network or services, unmet needs, experiences within the care network, barriers to sustainability, and collaboration. Group-specific measures included the primary end point (quality of life for patients) and the following secondary end points: self-efficacy for patients, family functioning and caregiver burden for family members, and quality of care and interprofessional collaboration for professionals. The survey was conducted and managed using the web-based REDCap (Research Electronic Data Capture; Vanderbilt University) tool [48]. The survey was made available online, and participants were recruited using a multitiered approach to gather a range of viewpoints. To ensure anonymization of data and to be able to link data over different time points, each participant marked the questionnaire with a personal code developed according to a defined scheme.

Table 1. Overview of survey components.

	Patients with NMDs[a]	Family members	Professionals	Other stakeholders
Component 1	Self-developed questions based on the qualitative study	Self-developed questions based on the qualitative study	Self-developed questions based on the qualitative study	Self-developed questions based on the qualitative study
Component 2	Health-related quality of life (WHOQOL-BREF[b] [49])	Family functioning (FAD[c] [50])	Quality of care (single-item measure [51])	Acute events or hospitalization, morbidity, and mortality (documents and statistics)
Component 3	Self-efficacy (GSE[d] [52])	Caregiver burden (BSFC[e]–Short Version [53])	Interprofessional collaboration (IPC[f] [54])	Costs, financing, and reimbursement (documents and statistics)

[a]NMD: neuromuscular disease.

[b]WHOQOL-BREF: World Health Organization Quality of Life Scale–Brief.

[c]FAD: Family Assessment Device.

[d]GSE: General Self-Efficacy Scale.

[e]BSFC: Burden Scale for Family Caregivers.

[f]IPC: interprofessional collaboration.

Quantitative Data Analysis

The data analysis will be carried out using the latest version of SPSS Statistics (IBM Corp) [55]. The dataset will be examined for missing and out-of-range values using frequency tables, and imputation procedures will be applied according to missing value mechanisms. Statistical assumptions will be assessed before analysis, and the choice of tests will be adapted as appropriate based on the data characteristics and the actual sample. We will describe the sample using frequencies, percentages, means, SDs, and ranges. The instrument's internal consistency, measured via the Cronbach α, must be above 0.80. Minimum values of 0.60 can be accepted for small sample sizes and scales with few items. The significance level will be set at <.05. Quantitative findings will be discussed with the expert panel and sounding board to check for clinical relevance and credibility.

Study Phase B: Development of a Care Management Model

The care management model was developed following the structure of a logic model [56]. Therefore, the qualitative and quantitative results from phase A were integrated to delineate unmet needs and challenges for the participant groups, which can be addressed by NMD care managers. This was the second data integration point in our mixed methods study, where integration of data occurred for comparison through merging quantitative and qualitative findings from phase A [32,36]. The developed model outlined the roles and responsibilities of NMD care managers and contained a comprehensive requirement and competency profile that covered both pediatric and adult care. Experts from the expert panel and sounding board were involved in the development of the model following a creative cyclical

XSL•FO
RenderX

process of repeated phases of reflection, planning, action, observation, reflection, and replanning [57].

On the basis of the developed NMD care management model, a training program for NMD care managers was developed, which consisted of classroom and online learning sessions and covered a broad spectrum of essential topics related to NMD-specific patient care. Key areas included pathophysiology, diagnostics, supportive interventions, and nursing and treatment options; psychosocial aspects, palliative care, and the relevant transitions (eg, from pediatric to adult care); and systemic counseling, interprofessional collaboration, and leadership. The training program incorporated existing standards of care for individual NMDs as well as care priorities and counseling topics that are shared across diagnoses. NMD care managers were expected to network, complete an internship, and prepare a case report. The 2-year part-time training program started with the implementation of the care management model in the pilot specialized health centers.

Study Phase C: Implementation and Evaluation

Implementation

Phase C comprised the recruitment of pilot specialized neuromuscular health centers for the 2-year implementation and a context analysis ahead of the implementation. The context analysis was necessary to better understand the specific institutional situation, overcome implementation challenges, and identify factors that may influence intervention uptake and study outcomes [58,59]. The context analysis was based on the Integrated Promoting Action on Research Implementation in Health Services framework [58,60]. The analysis was performed through focus group interviews (planned: n=8) using a semistructured interview guide based on the Integrated Promoting Action on Research Implementation in Health Services implementation framework, which focuses on recipients, evidence, and inner (local) and outer (organization) context [60]. The aim was to identify barriers to and facilitators of adapting the NMD care management model to individual institutional needs. Interviews were audio recorded and transcribed verbatim and analyzed as outlined in phase A.

Evaluation

The process and outcome evaluation was performed using formative and summative evaluation strategies. Formative evaluation assesses the local implementation strategy, monitors the process, and allows for real-time adjustments, improving the chances of success [61,62]. This was achieved through the training program, regular contact between the clinical team and research team, and documentation of clinical practice. After each contact with patients and their families or any other activity related to their role, this information was documented by the NMD care manager and collected by the research team for further analysis (care manager record using a case report form). The record included anonymized data, such as the type of contact, counseling topics, assessment, and quality of life. Summative evaluation of the NMD care management model was conducted by collecting qualitative and quantitative data through interviews, surveys, and the care manager record on 4 levels and at different time points (T1-T4; Table 2).

Table 2. Overview of data collection across time points.

Participants	Phase A (2020-2021)-T0 (baseline)	Phase B (2022)	Phase C (2023-2025)			
		Context analysis	T1 (6 months)	T2 (12 months)	T3 (18 months)	T4 (24 months)
Patients with NMDs[a]	Interview and survey	—[b]	—	Survey and care manager record	Interview	Survey and care manager record
Family members	Interview and survey	—	—	Survey	Interview	Survey
Professionals	Interview and survey	Interview	Interview	Survey	—	Survey
Stakeholders	Interview and survey	—	—	Survey	—	Survey

[a]NMD: neuromuscular disease.

[b]Not applicable.

Qualitative Descriptive Study, Data Collection, and Analysis

Six months after the start of implementation (T1), focus group interviews were conducted with professionals involved in the care at the pilot specialized neuromuscular health centers to evaluate the implementation and impact of the NMD care management model (planned: n=8). At 18 months after implementation (T3), semistructured interviews were undertaken with patients and their families to assess the potential benefits of the care management services (planned: n=16). The semistructured interview guides, data transcription, and analysis corresponded to those applied for the qualitative study in phase A.

Quantitative Data Collection and Analysis

The survey described in phase A was repeated in phase C under the same conditions: the same quantitative data were collected at 12 (T2) and 24 months (T4) after the start of implementation.

The T2 and T4 samples will be described using descriptive statistics (frequencies, percentages, means, SDs, and ranges). Changes over time (T0, T1, and T2) will be examined using inferential statistics and linear mixed models. Statistical assumptions will be assessed before analysis, and the choice of tests will be adapted as appropriate based on the data characteristics and the actual sample. A minimum sample size of 24 patients was estimated using a simulation-based approach to detect a medium effect, accounting for attrition in repeated-measure analyses of the primary end point. Quantitative

XSL•FO
RenderX

findings will be triangulated with qualitative results to explain observed patterns, and not significant changes will be discussed with the expert panel and sounding board to determine whether they are still clinically meaningful and inform refinement of the care management model and its implementation strategy.

Data Integration During Interpretation

Finally, findings will be integrated through joint displays and narrative at a third data integration point [32]. Therefore, both qualitative and quantitative findings will be reported together on a theme-by-theme basis for selected themes (weaving approach). Quantitative data will present the general trends and patterns, whereas qualitative data will be used to illustrate and deepen understanding of those trends through participant narratives. This integrated approach will be used in reporting and presentations to ensure a coherent and holistic account of the study's outcomes. Additionally, data will be brought together through a joint display using a table or matrix to draw out new insights beyond the information gained separately [32]. Therefore, a triangulation protocol will be applied.

Ethical Considerations

This study adheres to Swiss national laws and the ethical principles of the Declaration of Helsinki [55]. Ethics approval was received from the lead ethics committee in Ethikkommission Zürich (Business Administration System for Ethics Committees ID 2020-01882) and the ethics committees of the relevant study sites (Kantonale Ethikkommission Bern, Ethikkommission Nordwest- und Zentralschweiz EKNZ, Commission cantonale d'éthique de la recherche sur l'être humain CER-VD, Ethikkommission Ostschweiz EKOS, Ethikkommission Tessin). This study was classified as human research with the exception of clinical trials (Human Research Ordinance; chapter 2), risk category A (low risk), and registered in the Registry of All Projects in Switzerland (2020-01882). Participants provided informed consent before inclusion in the study. All participants were informed that their participation was voluntary and that they could withdraw from the study at any time. For data security, consent forms and other documents containing identifying information will be stored in a restricted archive for 10 years after the study ends. No compensation was provided to participants.

Results

This study received ethics approval in 2020. It was funded in 3 tranches in 2020, 2022, and 2024. Baseline data on the current state of care practices (T0) were collected in 2021 (phase A). A total of 52 interviews with individuals with NMDs or caregivers and health care professionals were conducted. In total, 360 people took part in the online survey. In 2022, based on the data from phase A, a care management model was codeveloped (phase B), which will be reported elsewhere. In phase C, the care management model was implemented in 8 pilot neuromuscular centers in 2023 with a concurrent training program of 150 training hours. Subsequently, data collection continued until May 2025, where data were collected at different time points (T1-T4) to evaluate the possible effects of implementing the care model. In phase C, a total of 24 interviews were conducted, and 409 people took part in the online survey. Data analysis and integration in this phase are currently ongoing, and findings will be published separately. This publication reports on the study protocol.

Discussion

Expected Findings

The proposed study investigates the current care situation for individuals with NMDs and their families in Switzerland and aims to develop, implement, and evaluate a standardized NMD care management model targeted at this population. The study's objective is to improve clinical outcomes, patient and family well-being, and the efficiency of health care services. Expected outcomes include enhanced patient and caregiver satisfaction through more coordinated, family-centered care. Codeveloping an NMD care management model with patients, family members, and professionals is expected to result in a comprehensive model and targeted implementation strategies. This model will meet real-life needs, with the potential to significantly improve health-related quality of life, self-efficacy, and family functioning while simultaneously reducing caregiver burden and the frequency of acute health care events such as hospitalizations. Strengthening collaboration among health care professionals within specialized neuromuscular centers and with external institutions through effective information exchange and optimized resource use can reduce overall health care costs and support a more sustainable model of care for patients with NMD. Additionally, this study aims to build and strengthen a network of specialized NMD care managers within and across health care settings, facilitate education and the exchange of NMD expertise, establish a standardized NMD care management model, and develop a training program for NMD care managers. This participatory mixed methods protocol developed for the Care-NMD-CH study can be adapted and applied to other settings and patient populations with complex health care needs, providing a scalable and generalizable framework for improving care management practices. This project represents the first time in Switzerland that the care of patients with NMDs, in terms of both the quality and quantity of services, is being examined.

Strengths and Limitations

This study targets an underserved population and introduces a framework for the development and implementation of care management services. The needs analysis and cocreation conducted in phases A and B with relevant stakeholders can enhance the credibility of the findings and may positively influence the acceptability and long-term sustainability of the care management model. Methodological choices and a mixed methods design in process and outcome evaluation (phase C), drawing on multiple data sources, enable triangulation across patient, family, professional, and system levels and support methodological rigor.

There are some potential limitations that should be noted at this stage. We recruited from different study sites in Switzerland, all of which have their own subcultures, focused specializations and interests, and differences in internal and external processes. A comparable level of diversity among participants may lead to heterogeneous results, and common themes in our data may be difficult to identify. However, it is important to recruit study

XSL•FO
RenderX

sites and participants from across Switzerland to better understand the current practices with regard to this group of rare disorders and the impact on the multifaceted quality of patient care in the country. Another potential challenge is that only a specific group of participants, such as those with a higher educational level, greater familiarity with digital media, and a lower disease burden, participated in the study. In addition, study sites involved in participant recruitment may have preferentially contacted individuals for interviews who were perceived as more suitable, which could introduce selection bias. Language may pose a challenge during the training course for new care managers, as many Swiss residents do not speak all three main official languages, and the English proficiency of potential participants may be inconsistent. The quantitative evaluation relies on pretest-posttest comparisons without a control group, limiting causal inference. The range of assessed outcomes can diminish the interpretability of the findings, and analysis is pending, so conclusions about broader impact remain preliminary.

Conclusions

The introduction of an evidence-based and family-centered care management model in pilot specialized neuromuscular centers will have a multifaceted impact on the care and treatment of this vulnerable patient group. It is expected that a standardized NMD care management model will strengthen the support system for patients living with these rare but severe conditions and create a sustainable care situation despite limited financial resources in the future. The Care-NMD-CH project represents an important step toward a comprehensive family-centered care management model for individuals with NMDs, which can be transferred to other chronic and rare disorders using this study framework. This study will advance care for individuals with NMDs and their families by identifying gaps in current practices; developing a targeted NMD care management model with training; and evaluating its impact on patients, families, and care teams. The findings will guide coordinated, family-centered, and sustainable care, offering a structure to improve clinical practice, patient and family experience, team collaboration, and health system support.

Acknowledgments

The authors thank the patient organizations and funding bodies for their support in developing this study protocol. The authors used ChatGPT (GPT-4-turbo; Open AI) and ProWritingAid (ProWritingAid Ltd) for structuring, language editing, and proofreading during the final phase of this manuscript. They confirm that ChatGPT or any other artificial intelligence–based tools were not used to develop this research project or the content of this manuscript. The authors take full responsibility for the content and interpretation of the manuscript.

Funding

Funding was provided by the Swiss Association for Muscle Diseases (Schweizerische Muskelgesellschaft), the Swiss ALS Foundation (Schweizerische ALS Stiftung), and the Swiss Foundation for Research on Muscle Diseases. The Swiss Association for Muscle Diseases received grants from the following pharmaceutical companies specifically to support this research project: Biogen, Novartis, and Roche. Open access funding was provided by ZHAW Zurich University of Applied Sciences. This study is being conducted independently. Pharmaceutical companies had no role in the study design; data collection, analysis, and interpretation; or reporting. The authors are committed to transparent reporting and will publish results regardless of the outcome.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical restrictions. Instruments and templates are available from the corresponding author upon reasonable request.

Authors' Contributions

VW, HP, GMS, MW, BGM, RW, and MK contributed to the study concept. CS, VW, and GMS drafted the manuscript. All authors critically revised the content of the manuscript and read and approved the final version.

Conflicts of Interest

GMS served on advisory boards for Novartis Gene Therapies, Biogen, Novartis, Pfizer, Roche, and Italfarmaco; received speaker's honoraria from Novartis; and is a member of the steering board of the Swiss registry for neuromuscular disorders. All other authors declare no other conflicts of interest.

References

1. Cruz KL, Santos IC, de Jesus Alves de Baptista CR, Mattiello-Sverzut A. Quality of life assessment instruments in children and adolescents with neuromuscular diseases: a systematic scoping review. Health Qual Life Outcomes 2024 Feb 16;22(1):18 [FREE Full text] [doi: 10.1186/s12955-024-02232-3] [Medline: 38360791]

XSL•FO
RenderX

2. García I, Rodríguez AA, Angelini C, García-Sanchoyerto M, Espinosa-Blanco P, Martínez O. Effects of coping strategies on health-related quality of life of people with neuromuscular diseases. Muscles 2024 Apr 03;3(2):110-120 [FREE Full text] [doi: 10.3390/muscles3020011]
3. Birnkrant DJ, Bushby K, Bann CM, Apkon SD, Blackwell A, Brumbaugh D, DMD Care Considerations Working Group. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and neuromuscular, rehabilitation, endocrine, and gastrointestinal and nutritional management. Lancet Neurol 2018 Mar;17(3):251-267 [FREE Full text] [doi: 10.1016/S1474-4422(18)30024-3] [Medline: 29395989]
4. Wu YF, Chen JA, Jong YJ. Treating neuromuscular diseases: unveiling gene therapy breakthroughs and pioneering future applications. J Biomed Sci 2025 Feb 21;32(1):30 [FREE Full text] [doi: 10.1186/s12929-025-01123-z] [Medline: 39985020]
5. Rodríguez AA, García M, Martínez O, López-Paz JF, García I, Pérez-Nuñez P, et al. Predictors of overload in parents of children with neuromuscular diseases. Front Neurol 2024;15:1349501-1342024 [FREE Full text] [doi: 10.3389/fneur.2024.1349501] [Medline: 38585358]
6. Waldboth V, Patch C, Mahrer-Imhof R, Metcalfe A. The family transition experience when living with childhood neuromuscular disease: a grounded theory study. J Adv Nurs 2021 Apr;77(4):1921-1933 [FREE Full text] [doi: 10.1111/jan.14754] [Medline: 33433037]
7. Waldboth V, Hediger H, Nemecek G, Fröhli R, Mezger M, Schubert M. Eine Untersuchung der Versorgungssituation von Menschen mit neuromuskulären Erkrankungen in der Schweiz (Projekt Care-NMD-CH) Kurzbericht Phase A: Analyse der aktuellen Versorgungssituation. Zurich University of Applied Sciences (ZHAW). URL: https://www.zhaw.ch/storage/gesundheit/forschung/fe-pflege/20221121-care-nmd-ch-kurzbericht-phase1-zhaw-gesundheit.pdf [accessed 2025-05-29]
8. Landfeldt E, Edström J, Buccella F, Kirschner J, Lochmüller H. Duchenne muscular dystrophy and caregiver burden: a systematic review. Dev Med Child Neurol 2018 Oct;60(10):987-996 [FREE Full text] [doi: 10.1111/dmcn.13934] [Medline: 29904912]
9. Mekhuri S, Dale C, Buchanan F, Hammash N, Ambreen M, Saha S, et al. Family caregivers of individuals with neuromuscular disease participating in a randomized controlled trial of a digital peer support program: nested qualitative study. J Med Internet Res 2025 Jul 28;27:e72141 [FREE Full text] [doi: 10.2196/72141] [Medline: 40720845]
10. Mitchell LA, Hirdes J, Poss JW, Slegers-Boyd C, Caldarelli H, Martin L. Informal caregivers of clients with neurological conditions: profiles, patterns and risk factors for distress from a home care prevalence study. BMC Health Serv Res 2015 Aug 28;15:350 [FREE Full text] [doi: 10.1186/s12913-015-1010-1] [Medline: 26315104]
11. Travlos V, Patman S, Downs J, Hince D, Wilson A. Parent carer quality of life and night-time attendance in non-ambulant youth with neuromuscular disorders. Dev Neurorehabil 2021 Oct;24(7):456-465 [FREE Full text] [doi: 10.1080/17518423.2021.1908440] [Medline: 33944675]
12. Paracha N, Hudson P, Mitchell S, Sutherland CS. Systematic literature review to assess the cost and resource use associated with spinal muscular atrophy management. Pharmacoeconomics 2022 Apr 11;40(Suppl 1):11-38 [FREE Full text] [doi: 10.1007/s40273-021-01105-7] [Medline: 34761360]
13. Elshahawy R, Elezbawy B, Ashmawy R, Elshahawy R, Mahmoud YS, Korra N, et al. Global economic burden of spinal muscular atrophy: a systematic literature review. Cureus 2025 Mar;17(3):e81023. [doi: 10.7759/cureus.81023] [Medline: 40264596]
14. Handberg C, Rudolfsen JH, Andersen H, Vissing J, Rossau CD, Dreyer P, et al. Societal costs, healthcare utilisation and labour market affiliation of persons with adult-onset myotonic dystrophy type 1 (DM1)-a register-based study II. Eur J Neurol 2025 Jun;32(6):e70211 [FREE Full text] [doi: 10.1111/ene.70211] [Medline: 40522020]
15. Mercuri E, Finkel RS, Muntoni F, Wirth B, Montes J, Main M, et al. Diagnosis and management of spinal muscular atrophy: part 1: recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. Neuromuscul Disord 2018 Feb;28(2):103-115 [FREE Full text] [doi: 10.1016/j.nmd.2017.11.005] [Medline: 29290580]
16. Birnkrant DJ, Bushby K, Bann CM, Alman BA, Apkon SD, Blackwell A, DMD Care Considerations Working Group. Diagnosis and management of Duchenne muscular dystrophy, part 2: respiratory, cardiac, bone health, and orthopaedic management. Lancet Neurol 2018 Apr;17(4):347-361 [FREE Full text] [doi: 10.1016/S1474-4422(18)30025-5] [Medline: 29395990]
17. Senyel D, Senn K, Boyd J, Nagels K. A systematic review of telemedicine for neuromuscular diseases: components and determinants of practice. BMC Digit Health 2024 May 09;2(1):17. [doi: 10.1186/S44247-024-00078-9]
18. Young HM, Kilaberia TR, Whitney R, Link BM, Bell JF, Tonkikh O, et al. Needs of persons living with ALS at home and their family caregivers: a scoping review. Muscle Nerve 2023 Sep;68(3):240-249 [FREE Full text] [doi: 10.1002/mus.27849] [Medline: 37248728]
19. Matthews EJ, Gelech J, Graumans R, Desjardins M, Gélinas I. Mediating a fragmented system: partnership experiences of parents of children with neurodevelopmental and neuromuscular disabilities. J Dev Phys Disabil 2020 May 23;33(2):311-330 [FREE Full text] [doi: 10.1007/S10882-020-09750-0]
20. Ørstavik K, Rosenberger AD, Fossmo HL, Nordstrøm M, de Visser M. Multidisciplinary management and care in rare neuromuscular disorders: a call for action. Eur J Neurol 2024 Jun 29;31(6):e16265. [doi: 10.1111/ene.16265] [Medline: 38425128]

XSL•FO
RenderX

21. Bushby K, Finkel R, Birnkrant DJ, Case LE, Clemens PR, Cripe L, DMD Care Considerations Working Group. Diagnosis and management of Duchenne muscular dystrophy, part 2: implementation of multidisciplinary care. Lancet Neurol 2010 Feb;9(2):177-189. [doi: 10.1016/S1474-4422(09)70272-8] [Medline: 19945914]
22. Paganoni S, Nicholson K, Leigh F, Swoboda K, Chad D, Drake K, et al. Developing multidisciplinary clinics for neuromuscular care and research. Muscle Nerve 2017 Nov 29;56(5):848-858 [FREE Full text] [doi: 10.1002/mus.25725] [Medline: 28632945]
23. Hess SM, Adu-Amankwah D, Elaiho CR, Butler LR, Ranade SC, Shah BJ, et al. Qualitative feedback from caregivers in a multidisciplinary pediatric neuromuscular clinic. J Pediatr Rehabil Med 2024;17(2):237-246 [FREE Full text] [doi: 10.3233/PRM-230011] [Medline: 38427510]
24. Hudon C, Chouinard MC, Brousselle A, Bisson M, Danish A. Evaluating complex interventions in real context: logic analysis of a case management program for frequent users of healthcare services. Eval Program Plann 2020 Apr;79:101753 [FREE Full text] [doi: 10.1016/j.evalprogplan.2019.101753] [Medline: 31835149]
25. Hudon C, Chouinard MC, Lambert M, Dufour I, Krieg C. Effectiveness of case management interventions for frequent users of healthcare services: a scoping review. BMJ Open 2016 Sep 29;6(9):e012353 [FREE Full text] [doi: 10.1136/bmjopen-2016-012353] [Medline: 27687900]
26. Aiken LH, Clarke SP, Sloane DM, Sochalski JA, Busse R, Clarke H, et al. Nurses' reports on hospital care in five countries. Health Aff (Millwood) 2001;20(3):43-53 [FREE Full text] [doi: 10.1377/hlthaff.20.3.43] [Medline: 11585181]
27. Campagna V, Lee-Riley L. Interprofessional collaboration as a best practice across the care continuum. Prof Case Manag 2025;30(3):100-106. [doi: 10.1097/NCM.0000000000000793] [Medline: 39838513]
28. Chouinard MC, Hudon C, Dubois MF, Roberge P, Loignon C, Tchouaket E, et al. Case management and self-management support for frequent users with chronic disease in primary care: a pragmatic randomized controlled trial. BMC Health Serv Res 2013 Feb 07;13:49 [FREE Full text] [doi: 10.1186/1472-6963-13-49] [Medline: 23391214]
29. Izumi S, Mood L, Velk L, Szper C, Fifield A, Ayala D, et al. Key components of chronic care management programs: qualitative descriptions of care managers' practice. BMC Health Serv Res 2025 Nov 22;25(1):1542 [FREE Full text] [doi: 10.1186/s12913-025-13788-1] [Medline: 41272783]
30. Gruebner O, van Haasteren A, Hug A, Elayan S, Sykora M, Albanese E, et al. Mental health challenges and digital platform opportunities in patients and families affected by pediatric neuromuscular diseases - experiences from Switzerland. Digit Health 2023;9:20552076231213700 [FREE Full text] [doi: 10.1177/20552076231213700] [Medline: 38025108]
31. Waldboth V, Patch C, Mahrer-Imhof R, Metcalfe A. Living a normal life in an extraordinary way: a systematic review investigating experiences of families of young people's transition into adulthood when affected by a genetic and chronic childhood condition. Int J Nurs Stud 2016 Oct;62:44-59. [doi: 10.1016/j.ijnurstu.2016.07.007] [Medline: 27450665]
32. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
33. Grove SK, Burns N, Gray J. The Practice of Nursing Research: Appraisal, Synthesis, and Generation of Evidence. 6th edition. New York, NY: Elsevier Inc; 2009.
34. Westhues A, Ochocka J, Jacobson N, Simich L, Maiter S, Janzen R, et al. Developing theory from complexity: reflections on a collaborative mixed method participatory action research study. Qual Health Res 2008 May;18(5):701-717 [FREE Full text] [doi: 10.1177/1049732308316531] [Medline: 18420539]
35. McGoldrick Carter GP. The Expanded Family Life Cycle: Individual, Family, and Social Perspectives. 4th edition. New York, NY: Pearson; 2014.
36. Creswell JW, Creswell JD. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. 5th edition. Thousand Oaks, CA: Sage Publications; 2018.
37. Creswell JW, Clark VL. Designing and Conducting Mixed Methods Research. Thousand Oaks, CA: Sage Publications Ltd; 2017.
38. Lincoln YS, Guba EG, Pilotta JJ. Naturalistic inquiry. In: Given LM, editor. The SAGE Encyclopedia of Qualitative Research Methods. Thousand Oaks, CA: Sage Publication; 1985:438-439.
39. Tashakkori A, Teddlie C. SAGE Handbook of Mixed Methods in Social & Behavioral Research. Thousand Oaks, CA: Sage Publications; 2010.
40. Brandenburg H, Panfil E, Mayer H, Schrems B. Lehrund Arbeitsbuch zur Einführung in die Methoden der Pflegeforschung. 3 ed. Germany: Hogrefe; 2018.
41. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
42. Miles MB, Huberman AM, Saldaña J. Qualitative Data Analysis: A Methods Sourcebook. 3rd edition. Thousand Oaks, CA: Sage Publications; 2014.
43. Creswell JW, Poth CN. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: Sage Publications; 2016.
44. Kelly SE, Bourgeault I, Dingwall R. Qualitative interviewing techniques and styles. In: Bourgeault I, Dingwall R, De Vries R, editors. The SAGE Handbook of Qualitative Methods in Health Research. Thousand Oaks, CA: Sage Publications; 2010:307-326.

XSL•FO
RenderX

45. Dresing T, Pehl T. Manual on Transcription Techniques: A Guide for Qualitative Research. Cham, Switzerland: Springer; 2015.
46. Advanced qualitative & mixed methods research for Windows and Mac. MAXQDA. URL: https://www.maxqda.com/products/maxqda [accessed 2025-05-29]
47. Saldaña J. The Coding Manual for Qualitative Researchers. 2nd edition. Thousand Oaks, CA: Sage Publications; Oct 18, 2016.
48. REDCap. URL: https://www.project-redcap.org/ [accessed 2024-03-10]
49. Angermeyer M, Kilian R, Matschinger H. Handbuch für die deutschsprachigen Version der WHO Instrumente zur Erfassung von Lebensqualität. Göttingen: Hogrefe; 2000.
50. Angermeyer RK, Matschinger H. Diagnostiscke Verfahren für die Beratung: Materialen zur Diagnostik und Therapie in Ehe-, Familien-und Lebensberatung. 2003. URL: https://www.hogrefe.com/ch/shop/diagnostische-verfahren-fuer-die-beratung-65116.html [accessed 2026-02-15]
51. Zúñiga F, Ausserhofer D, Hamers JP, Engberg S, Simon M, Schwendimann R. Are staffing, work environment, work stressors, and rationing of care related to care workers' perception of quality of care? a cross-sectional study. J Am Med Dir Assoc 2015 Oct 01;16(10):860-866 [FREE Full text] [doi: 10.1016/j.jamda.2015.04.012] [Medline: 26027721]
52. Generalized self-efficacy scale. 1995. URL: https://userpage.fu-berlin.de/~health/selfscal.htm [accessed 2026-02-15]
53. Graessel E, Berth H, Lichte T, Grau H. Subjective caregiver burden: validity of the 10-item short version of the Burden Scale for Family Caregivers BSFC-s. In: BMC Geriatr. Häusliche Pflege-Skala HPS: zur Erfassung der Belastung bei betreuuenden oder pflegenden Angehörigen. 2 ed. VLESS Verlag; Feb 20, 2014:23.
54. Kenaszchuk C, Reeves S, Nicholas D, Zwarenstein M. Validity and reliability of a multiple-group measurement scale for interprofessional collaboration. BMC Health Serv Res 2010 Mar 30;10:83 [FREE Full text] [doi: 10.1186/1472-6963-10-83] [Medline: 20353577]
55. IBM SPSS Statistics (version 31.0.0.0). IBM Corp. URL: https://www.ibm.com/products/spss-statistics [accessed 2025-05-25]
56. Silverman B, Cara M, Boulet S, O'Leary L. Logic models for planning and evaluation. Centers for Disease Control and Prevention. URL: https://stacks.cdc.gov/view/cdc/11569/cdc_11569_DS1.pdf? [accessed 2025-05-29]
57. Glasson JB, Chang EM, Bidewell JW. The value of participatory action research in clinical nursing practice. Int J of Nursing Practice 2008 Jan 07;14(1):34-39 [FREE Full text] [doi: 10.1111/j.1440-172x.2007.00665.x]
58. McCormack B, Kitson A, Harvey G, Rycroft-Malone J, Titchen A, Seers K. Getting evidence into practice: the meaning of 'context'. J Adv Nurs 2002 Apr;38(1):94-104 [FREE Full text] [doi: 10.1046/j.1365-2648.2002.02150.x] [Medline: 11895535]
59. Nilsen P, Bernhardsson S. Context matters in implementation science: a scoping review of determinant frameworks that describe contextual determinants for implementation outcomes. BMC Health Serv Res 2019 Mar 25;19(1):189 [FREE Full text] [doi: 10.1186/s12913-019-4015-3] [Medline: 30909897]
60. Harvey G, Kitson A. PARIHS revisited: from heuristic to integrated framework for the successful implementation of knowledge into practice. Implement Sci 2016 Mar 10;11:33 [FREE Full text] [doi: 10.1186/s13012-016-0398-2] [Medline: 27013464]
61. Kowalski CP, Veeser M, Heisler M. Formative evaluation and adaptation of pre-and early implementation of diabetes shared medical appointments to maximize sustainability and adoption. BMC Fam Pract 2018 Jul 07;19(1):109-123 [FREE Full text] [doi: 10.1186/s12875-018-0797-3] [Medline: 29981568]
62. Stetler CB, Legro MW, Wallace CM, Bowman C, Guihan M, Hagedorn H, et al. The role of formative evaluation in implementation research and the QUERI experience. J Gen Intern Med 2006 Feb;21(S2):S1-S8 [FREE Full text] [doi: 10.1007/S11606-006-0267-9]

Abbreviations

NMD: neuromuscular disease
REDCap: Research Electronic Data Capture

XSL•FO
RenderX

Edited by J Sarvestan; submitted 22.Aug.2025; peer-reviewed by M Tomita; comments to author 11.Nov.2025; revised version received 05.Feb.2026; accepted 05.Feb.2026; published 24.Feb.2026.

Please cite as:
Waldboth V, Schuler C, Willmann R, Petry H, Weber M, Grädel Messerli B, Hediger H, Kramer I, Stettler-Nemecek G, Knoblauch M, Stettner GM
Developing, Implementing, and Evaluating a Care Management Model for Patients With Neuromuscular Diseases: Protocol for a Participatory Mixed Methods Study
JMIR Res Protoc 2026;15:e82833
URL: https://www.researchprotocols.org/2026/1/e82833
doi:10.2196/82833
PMID:

©Veronika Waldboth, Christina Schuler, Raffaella Willmann, Heidi Petry, Markus Weber, Barbara Grädel Messerli, Hannele Hediger, Iris Kramer, Gabriela Stettler-Nemecek, Martin Knoblauch, Georg Martin Stettner. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Implementation of a Personalized Medicine Approach in Patients With Type 2 Diabetes Mellitus Receiving Multiple Daily Insulin Injections (POMA Project): Protocol for a Before-and-After Intervention Study

Rosa Giné-Balcells[1], RN, MSc; Bogdan Vlacho[2,3], PharmD, MSc, PhD; Olga Carmen Vidal-Pérez[4], RN; Albert Lecube[2,5,6], MD, PhD; Josep Franch-Nadal[2,3,7], MD, PhD; Dídac Mauricio[2,3,8,9,10], MD, PhD; Àngels Molló-Iniesta[1,3*], MD, PhD; Marta Hernández[4,11*], MD, PhD; POMA Study Group[4,12]

[1]Primary Health Care Center Tàrrega, Gerència d'Atenció Primària Lleida, Institut Català de la Salut, Lleida, Spain

[10]Faculty of Medicine, University of Vic - Central University of Catalonia (UVIC/UCC), Vic, Spain

[11]Institut de Recerca Biomèdica de Lleida Fundació Dr. Pifarré, University of Lleida, Lleida, Spain

[12]Primary Health Care Center, Gerència d'Atenció Primària Lleida, Institut Català de la Salut, Lleida, Spain

[2]CIBER de Diabetes y Enfermedades Metabólicas Asociadas, Instituto de Salud Carlos III, Barcelona, Spain

[3]DAP-Cat group, Unitat de Suport a la Recerca Barcelona, Fundació Institut Universitari d'Investigació en Atenció Primària Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain

[4]Department of Endocrinology and Nutrition, University Hospital Arnau de Vilanova, Avda. Alcalde Rovira Roure, 80, Lleida, Spain

[5]Endocrinology Department, Vall d'Hebron University Hospital, Autonomous University Barcelona, Barcelona, Spain

[6]Diabetes and Metabolism Research Group, Vall d'Hebron Research Institute (VHIR), Barcelona, Spain

[7]Primary Health Care Center, Gerència d'Àmbit d'Atenció Primària Barcelona Ciutat, Institut Català de la Salut, Barcelona, Spain

[8]Research Group in Endocrinology, Diabetes and Nutrition, Institut de Recerca Sant Pau, Barcelona, Spain

[9]Department of Endocrinology and Nutrition, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain

[*]these authors contributed equally

Corresponding Author:
Marta Hernández, MD, PhD
Department of Endocrinology and Nutrition, University Hospital Arnau de Vilanova, Avda. Alcalde Rovira Roure, 80, Lleida, Spain

Related Article:

Correction of: https://www.researchprotocols.org/2026/1/e85375

(JMIR Res Protoc 2026;15:e95267) doi:10.2196/95267

In "Implementation of a Personalized Medicine Approach in Patients With Type 2 Diabetes Mellitus Receiving Multiple Daily Insulin Injections (POMA Project): Protocol for a Before-and-After Intervention Study" [1], the authors made three corrections.

At the beginning of the Acknowledgements section, the following group author contributor information was added:

> *The POMA Study Group members are as follows: Magda Mateu Amorós, Anaïs Arqué Badia, Núria Balsells Bailón, Marta Bueno Díez, Marta Carrasco Marín-Blázquez, Alba Dalmau Vila, Carolina López Cano, Raquel Ruano Esteban, Ana Gloria Soler Beunza, Javier Suárez Balaguer, Gabriela Monroy Rodríguez, Josep León Mengíbar, Ferran Rius Riu, Maribel Martí Sotus, Mª José Rosauro Gómez, Laia Naté López, Josefa Tomàs Anastasi, Rocío Quintillà Pueo, Sara Areny Rodrigo, Fernando Chilangua Sepúlveda, Mercè Oliver García, Ruben Perales Sánchez, Anna Serra Peruchet, Jaume Martí Montesa, Irene Gómez Companys, Janette Sandrà Medina, Silvia Guiu Vendrell, Roxana Lenuta Mocanu, David José De la Rica Escuin, Mª Teresa Guasch Clapés, Eva Ribalta Calvet, Rosa M Aran Padulles, Encarna López Esteve, Meritxell Cunillera Batlle, Laura Camats Pascual, Esther Call Camps, Laura García Esteve, Teresa Ibañez Plana, Laia Llubes Arria, Irene Muñoz del Campo, Agustí Jové Subirada, Silvia Tierz Puyuelo, Teresa Vidal Bertran, Maria Mesalles Rodríguez, Natalia Reñé Puigfel, Asunción Garrido Díaz, Sandra Maria Millera, Mª Antònia Lafarga Giribets, Marta Ferminán Miró, Mònica Solanes Cabús, Mª Abel Cucurull Llobet, Laura Roure Chiquero, Salvador Salanova Montané, Mª Fe Diaz Fernandez, Arantxa Morales Valle, Cristina Cebrián Aiguadé, Anna Pagès Cónsul, Neus Miró Vallvé, Raquel Rodríguez Aguilar, Maite Puig Solé, Montserrat Torra Solé, Úrsula Belén Flix Pau, Sara Porta Acosta, and Marta Tolosa Fortuny.*

XSL•FO
RenderX

At the end of the Acknowledgments section, the following sentences were added:

> *The authors also acknowledge that the Diabetes Treatment Satisfaction Questionnaire (DTSQ) is a copyrighted instrument developed by Professor Clare Bradley. Permission to use the DTSQ was obtained from Health Psychology Research Ltd (HPR), UK.*

In addition, reference 23 was originally published as:

> *Gomis R, Herrera-Pombo JL, Calderón A, Rubio-Terrés C, Sarasa P. RETRACTED ARTICLE: Validación del cuestionario "Diabetes Treatment Satisfaction Questionnaire" (DTSQ) en la población española. Pharmacoecon Span Res Artic. 2006 Jan;3(1):7–18. doi:10.1007/BF03320906.*

This has been replaced with the following:

> *Bradley C. Diabetes Treatment Satisfaction Questionnaire: (DTSQ). In: Bradley C, editor. Handbook of Psychology and Diabetes: A Guide to Psychological Measurement in Diabetes Research and Practice. Routledge; 1994. ISBN: 9783718655625*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Giné-Balcells R, Vlacho B, Vidal-Pérez OC, et al. Implementation of a personalized medicine approach in patients with type 2 diabetes mellitus receiving multiple daily insulin injections (POMA Project): protocol for a before-and-after intervention study. JMIR Res Protoc 2026 Feb 24;15:e85375. [doi: 10.2196/85375] [Medline: 41734338]

Submitted 13.Mar.2026; this is a non–peer-reviewed article;accepted 13.Mar.2026; published 13.Apr.2026.

Please cite as:
Giné-Balcells R, Vlacho B, Vidal-Pérez OC, Lecube A, Franch-Nadal J, Mauricio D, Molló-Iniesta À, Hernández M, POMA Study Group
Correction: Implementation of a Personalized Medicine Approach in Patients With Type 2 Diabetes Mellitus Receiving Multiple Daily Insulin Injections (POMA Project): Protocol for a Before-and-After Intervention Study
JMIR Res Protoc 2026;15:e95267
URL: https://www.researchprotocols.org/2026/1/e95267
doi:10.2196/95267

© Rosa Giné-Balcells, Bogdan Vlacho, Olga Carmen Vidal-Pérez, Albert Lecube, Josep Franch-Nadal, Dídac Mauricio, Àngels Molló-Iniesta, Marta Hernández, POMA Study Group. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Evaluation of the Comparative Efficacy of an Ayurvedic Formulation (Nimba-Amalakyadi Powder) vs Metformin in the Management of Type 2 Diabetes Mellitus: Protocol for a Pilot Study

Punam Sawarkar[1*], BAMS, MD, PhD; Gaurav Rajendra Sawarkar[2*], BAMS, MD, PhD

[1]Department of Panchakarma, Mahatma Gandhi Ayurved College, Hospital and Research Centre, Datta Meghe Institute of Higher Education and Research, Salod(H), Wardha, Maharashtra, India

[2]Department of Rachana Sharir, Mahatma Gandhi Ayurved College, Hospital and Research Centre, Datta Meghe Institute of Higher Education and Research, Salod(H), Wardha, Maharashtra, India

[*]all authors contributed equally

Corresponding Author:
Punam Sawarkar, BAMS, MD, PhD
Department of Panchakarma, Mahatma Gandhi Ayurved College, Hospital and Research Centre, Datta Meghe Institute of Higher Education and Research, Salod(H), Wardha, Maharashtra, India

Related Article:

Correction of: https://www.researchprotocols.org/2026/1/e63574

(JMIR Res Protoc 2026;15:e96344) doi:10.2196/96344

In "Evaluation of the Comparative Efficacy of an Ayurvedic Formulation (*Nimba-Amalakyadi* Powder) vs Metformin in the Management of Type 2 Diabetes Mellitus: Protocol for a Pilot Study" [1], the authors made one addition.

The following institution name was added to both affiliations:

> *Datta Meghe Institute of Higher Education and Research*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Sawarkar P, Sawarkar GR. Evaluation of the comparative efficacy of an Ayurvedic formulation (Nimba-Amalakyadi powder) vs metformin in the management of type 2 diabetes mellitus: protocol for a pilot study. JMIR Res Protoc 2026 Mar 13;15:e63574. [doi: 10.2196/63574] [Medline: 41824950]

Submitted 30.Mar.2026; this is a non–peer-reviewed article;accepted 30.Mar.2026; published 09.Apr.2026.

Please cite as:
Sawarkar P, Sawarkar GR
Correction: Evaluation of the Comparative Efficacy of an Ayurvedic Formulation (Nimba-Amalakyadi Powder) vs Metformin in the Management of Type 2 Diabetes Mellitus: Protocol for a Pilot Study
JMIR Res Protoc 2026;15:e96344
URL: https://www.researchprotocols.org/2026/1/e96344
doi:10.2196/96344

© Punam Sawarkar, Gaurav Rajendra Sawarkar. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The

XSL•FO
RenderX

complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Correction: Smartphone App–Based Eating Behavior Monitoring and Feedback Intervention for Glucocorticoid-Induced Appetite Increase in Patients With Systemic Lupus Erythematosus: Protocol for a Pilot Randomized Controlled Trial

Takashi Yamaguchi[1*], BPharm, PhD; Nobuyuki Takahashi[1*], BPharm, PhD; Ryohei Inanaga[2,3*], MD; Ryuhei So[4], MD, MPH, PhD; Hiroe Kikuchi[5], MD, PhD; Hisashi Noma[6], MPH, PhD; Hiroyuki Sasai[7], PhD; Kazuro Kamada[8], MD; Tomohiro Sugimoto[9], MD, PhD; Hiroshi Tsushima[10], MD, PhD; Takanori Ichikawa[11], MD, PhD; Hirofumi Miyake[12], MD; Shunichi Fujita[13], MD, PhD; Keisuke Ono[14], MD, PhD; Yusuke Miwa[15], MD, PhD; Anna Hasegawa[16], MD; Naoki Suzuki[17], MD; Akira Onishi[18], MPH, MD, PhD; Toshihiro Matsui[19], MD, PhD; Ryu Watanabe[20], MD, PhD; Yasuhiro Hasegawa[21], MD, PhD; Rei Ono[22], BPharm, PhD; Takeo Isozaki[22], MD, PhD; Yuichi Ishikawa[23], MD, MPH; Nobuyuki Yajima[3,15], MPH, MD, PhD; Noriaki Kurita[3,15,24], MD, PhD

[1]Department of Hospital Pharmaceutics, School of Pharmacy, Showa Medical University, Shinagawa-ku, Japan

[10]Department of Internal Medicine and Rheumatology, Juntendo University Shizuoka Hospital, Izunokuni, Japan

[11]Department of Medicine (Neurology and Rheumatology), School of Medicine, Shinshu University, Matsumoto, Japan

[12]Department of General Internal Medicine, Tenri Hospital, Tenri, Japan

[13]Department of Rheumatology, Kawasaki Medical School, Kurashiki, Japan

[14]Department of Nephrology and Rheumatology, School of Medicine, Kyorin University, Mitaka, Japan

[15]Division of Rheumatology, Department of Medicine, Showa Medical University, Shinagawa-ku, Japan

[16]Department of Rheumatology, Dokkyo Medical University, Mibu, Japan

[17]Department of Stem Cell and Immune Regulation, Graduate School of Medicine, Yokohama City University, Yokohama, Japan

[18]Department of Advanced Medicine for Rheumatic Diseases, Kyoto University Graduate School of Medicine, Kyoto, Japan

[19]Department of Rheumatology Research, Clinical Research Center for Allergy and Rheumatology, National Hospital Organization Sagamihara National Hospital, Sagamihara, Japan

[2]Department of Nephrology, Shin-yurigaoka General Hospital, Kawasaki, Japan

[20]Department of Clinical Immunology, Graduate School of Medicine, Osaka Metropolitan University, Osaka, Japan

[21]Department of Rheumatology, Kitasato University, Sagamihara, Japan

[22]Department of Pathogenesis and Translational Medicine, Graduate School of Pharmacy, Showa Medical University, Shinagawa-ku, Japan

[23]The First Department of Internal Medicine, University of Occupational and Environmental Health Japan, Kitakyushu, Japan

[24]Department of Innovative Research and Education for Clinicians and Trainees (DiRECT), Fukushima Medical University Hospital, Fukushima, Japan

[3]Department of Clinical Epidemiology, Graduate School of Medicine, Fukushima Medical University, 1 Hikarigaoka, Fukushima, Japan

[4]Department of Psychiatry, Okayama Psychiatric Medical Center, Okayama, Japan

[5]Department of Psychosomatic Medicine, National Center for Global Health and Medicine, Japan Institute for Health Security, Shinjuku-ku, Japan

[6]Department of Interdisciplinary Statistical Mathematics, The Institute of Statistical Mathematics, Tachikawa, Japan

[7]Research Team for Promoting Independence and Mental Health, Tokyo Metropolitan Institute for Geriatrics and Gerontology, Itabashi-ku, Japan

[8]Department of Rheumatology, Endocrinology and Nephrology, Faculty of Medicine and Graduate School of Medicine, Hokkaido University, Sapporo, Japan

[9]Department of Clinical Immunology and Rheumatology, Hiroshima University Hospital, Hiroshima, Japan

[*]these authors contributed equally

Corresponding Author:
Noriaki Kurita, MD, PhD
Department of Clinical Epidemiology, Graduate School of Medicine, Fukushima Medical University, 1 Hikarigaoka, Fukushima, Japan

Related Article:

Correction of: https://www.researchprotocols.org/2025/1/e78612

(JMIR Res Protoc 2026;15:e90010) doi:10.2196/90010

XSL•FO
RenderX

In “Smartphone App–Based Eating Behavior Monitoring and Feedback Intervention for Glucocorticoid-Induced Appetite Increase in Patients With Systemic Lupus Erythematosus: Protocol for a Pilot Randomized Controlled Trial” [1], the authors noted an error.

The eighth author’s name has been changed from the following:

Kazuhiro Kamada

The name now appears as:

Kazuro Kamada

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Yamaguchi T, Takahashi N, Inanaga R, et al. Smartphone app-based eating behavior monitoring and feedback intervention for glucocorticoid-induced appetite increase in patients with systemic lupus erythematosus: protocol for a pilot randomized controlled trial. JMIR Res Protoc 2025 Nov 26;14:e78612. [doi: 10.2196/78612] [Medline: 41305946]

Submitted 19.Dec.2025; this is a non–peer-reviewed article;accepted 23.Dec.2025; published 21.Jan.2026.

Please cite as:
Yamaguchi T, Takahashi N, Inanaga R, So R, Kikuchi H, Noma H, Sasai H, Kamada K, Sugimoto T, Tsushima H, Ichikawa T, Miyake H, Fujita S, Ono K, Miwa Y, Hasegawa A, Suzuki N, Onishi A, Matsui T, Watanabe R, Hasegawa Y, Ono R, Isozaki T, Ishikawa Y, Yajima N, Kurita N
Correction: Smartphone App–Based Eating Behavior Monitoring and Feedback Intervention for Glucocorticoid-Induced Appetite Increase in Patients With Systemic Lupus Erythematosus: Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2026;15:e90010
URL: https://www.researchprotocols.org/2026/1/e90010
doi:10.2196/90010

© Takashi Yamaguchi, Nobuyuki Takahashi, Ryohei Inanaga, Ryuhei So, Hiroe Kikuchi, Hisashi Noma, Hiroyuki Sasai, Kazuro Kamada, Tomohiro Sugimoto, Hiroshi Tsushima, Takanori Ichikawa, Hirofumi Miyake, Shunichi Fujita, Keisuke Ono, Yusuke Miwa, Anna Hasegawa, Naoki Suzuki, Akira Onishi, Toshihiro Matsui, Ryu Watanabe, Yasuhiro Hasegawa, Rei Ono, Takeo Isozaki, Yuichi Ishikawa, Nobuyuki Yajima, Noriaki Kurita. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Deep Learning–Based Pattern Recognition for Detecting Penile Abnormalities: Protocol for Developing a Mobile App for Circumcision Eligibility

Irfan Wahyudi[1], MD, PhD, Prof Dr; Chandra Prasetyo Utomo[2], BSc, MS; Samsuridjal Djauzi[3], MD, PhD, Prof Dr; Muhamad Fathurahman[2], BSc, MSc; Gerhard Reinaldi Situmorang[1], MD, PhD; Arry Rodjani[1], MD; Putu Angga Risky Raharja[1], MD; Kevin Yonathan[1], MD; Budi Santoso[1], MD; Dwidian Khresna[1], MD; Marco Raditya[1], MD

[1]Department of Urology, Faculty of Medicine, Universitas Indonesia - Cipto Mangunkusumo Hospital, Jakarta, Indonesia

[2]YARSI E-Health Research Center, Faculty of Information Technology, YARSI University, Jakarta, Indonesia

[3]Department of Internal Medicine, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia

Corresponding Author:
Irfan Wahyudi, MD, PhD, Prof Dr
Department of Urology
Faculty of Medicine
Universitas Indonesia - Cipto Mangunkusumo Hospital
Jl. Pangeran Diponegoro No.71
Jakarta, 10430
Indonesia
Phone: 62 213160493
Email: irf.wahyudi2011@gmail.com

Related Article:

Correction of: https://www.researchprotocols.org/2025/1/e65811

(JMIR Res Protoc 2026;15:e89351) doi:10.2196/89351

In "Deep Learning–Based Pattern Recognition for Detecting Penile Abnormalities: Protocol for Developing a Mobile App for Circumcision Eligibility" [1], the authors noted an error in the funding information in the Acknowledgments section of the published version. The correction has been made to ensure accurate attribution of grant sources and contract numbers.

In the originally published article, the Acknowledgments section read:

> *We would like to acknowledge the support of the Badan Riset dan Inovasi Nasional for their financial support to fund this research (grant NKB-912/UN2.RST/HKP.05.00/2022). We would also like to extend our sincere appreciation to the entire medical team for their vital contributions in executing this protocol and gathering the research data. Huge appreciation is also given to the staff members of the E-Health Research Center for their help in providing support throughout the AI development process.*

The corrected Acknowledgments section now reads:

This study was funded by the RIIM LPDP Grant Program and Badan Riset dan Inovasi Nasional (BRIN; grants 97/IV/KS/11/2022 and 530./PKS/WRIII-DISTP/UI/2022), awarded through Universitas Indonesia. We would also like to extend our sincere appreciation to the entire medical team for their vital contributions in executing this protocol and gathering the research data. Huge appreciation is also given to the staff members of the YARSI E-Health Research Center for their help in providing support throughout the AI development process.

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Wahyudi I, Utomo CP, Djauzi S, Fathurahman M, Situmorang GR, Rodjani A, et al. Deep Learning-Based Pattern Recognition for Detecting Penile Abnormalities: Protocol for Developing a Mobile App for Circumcision Eligibility. JMIR Res Protoc 2025 Sep 10;14:e65811 [FREE Full text] [doi: 10.2196/65811] [Medline: 40929720]

XSL•FO
RenderX

Submitted 11.Dec.2025; this is a non–peer-reviewed article;accepted 12.Dec.2025; published 02.Mar.2026.

Please cite as:
Wahyudi I, Utomo CP, Djauzi S, Fathurahman M, Situmorang GR, Rodjani A, Raharja PAR, Yonathan K, Santoso B, Khresna D, Raditya M
Correction: Deep Learning–Based Pattern Recognition for Detecting Penile Abnormalities: Protocol for Developing a Mobile App for Circumcision Eligibility
JMIR Res Protoc 2026;15:e89351
URL: https://www.researchprotocols.org/2026/1/e89351
doi:10.2196/89351
PMID:

©Irfan Wahyudi, Chandra Prasetyo Utomo, Samsuridjal Djauzi, Muhamad Fathurahman, Gerhard Reinaldi Situmorang, Arry Rodjani, Putu Angga Risky Raharja, Kevin Yonathan, Budi Santoso, Dwidian Khresna, Marco Raditya. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 02.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Vaping, Acculturation, and Social Media Use Among Mexican American College Students: Protocol for a Mixed Methods Web-Based Cohort Study

Bara S Bataineh[1], MD, PhD; C Nathan Marti[2], PhD; Dhiraj Murthy[2], PhD; David Badillo[1], MPH; Sherman Chow[2], MPH; Alexandra Loukas[2], PhD; Anna V Wilkinson[1], PhD

[1]University of Texas Health Science Center at Houston, Austin, TX, United States

[2]University of Texas at Austin, Austin, TX, United States

Corresponding Author:
Bara S Bataineh, MD, PhD
University of Texas Health Science Center at Houston
1836 San Jacinto
Austin, TX, 78701
United States
Phone: 1 9723527755
Email: Bara.bataineh@uth.tmc.edu

Related Article:

Correction of: https://www.researchprotocols.org/2025/1/e63584/

(JMIR Res Protoc 2026;15:e95651) doi:10.2196/95651

In "Vaping, Acculturation, and Social Media Use Among Mexican American College Students: Protocol for a Mixed Methods Web-Based Cohort Study" [1], an error in the institutional review board number was identified in the Methods section under the subheading Human Participant Ethics Review Approvals.

The following statement has been corrected:

> *The IRB at UTHealth Houston (HSC-SPH-19-0796) approved all aspects of the study protocol.*

The statement now reads:

> *The IRB at UTHealth Houston (HSC-SPH-21-0976) approved all aspects of the study protocol.*

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Bataineh BS, Marti CN, Murthy D, Badillo D, Chow S, Loukas A, et al. Vaping, Acculturation, and Social Media Use Among Mexican American College Students: Protocol for a Mixed Methods Web-Based Cohort Study. JMIR Res Protoc 2025 Mar 24;14:e63584 [FREE Full text] [doi: 10.2196/63584] [Medline: 40127433]

Submitted 18.Mar.2026; this is a non–peer-reviewed article;accepted 26.Mar.2026; published 02.Apr.2026.

Please cite as:
Bataineh BS, Marti CN, Murthy D, Badillo D, Chow S, Loukas A, Wilkinson AV
Correction: Vaping, Acculturation, and Social Media Use Among Mexican American College Students: Protocol for a Mixed Methods Web-Based Cohort Study
JMIR Res Protoc 2026;15:e95651
URL: https://www.researchprotocols.org/2026/1/e95651
doi:10.2196/95651
PMID:

XSL•FO
RenderX

©Bara S Bataineh, C Nathan Marti, Dhiraj Murthy, David Badillo, Sherman Chow, Alexandra Loukas, Anna V Wilkinson. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 02.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: Real-World Effectiveness and Safety of Liuwei Dihuang Pill for Menopausal Syndrome: Protocol for a Prospective, Observational, Multicenter Cohort Study

Shuoshuo Wei[1], MM; Hongyan Zhang[1], MMSc; Mengmeng Wang[1], MMSc; Yuhui Sun[2], MD; Tao Song[2], MD; Weifang Zheng[3], MBBS; Yunmei Ye[3], MBBS; Liran Ma[4], MM; Yunshuang Yang[5], PhD; Tong Wei[6], MM; Lan Cheng[7], MM; Chunquan Sun[8], MD; Xingting Ding[9], MM; Shichu Zhao[7*], PhD; Qingqiao Song[10*], MD; Lianxin Wang[1*], PhD

[1]Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China
[2]First Affiliated Hospital of Harbin Medical University, Heilongjiang, China
[3]Lanxi Hospital of Traditional Chinese Medicine, Zhejiang, China
[4]Beijing Gulou Hospital of Traditional Chinese Medicine, Beijing, China
[5]Beijing Longfu Hospital, Beijing, China
[6]Hepingli Community Health Service Center, Dongcheng District, Beijing, China
[7]Beijing Puren Hospital, Beijing, China
[8]Tsinghua University Yuquan Hospital (Tsinghua University Hospital of Integrative Traditional Chinese and Western Medicine), Beijing, China
[9]Chaoyangmen Community Health Service Center, Dongcheng District, Beijing, China
[10]Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China
[*]these authors contributed equally

Corresponding Author:
Lianxin Wang, PhD
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences
No. 16, Dongzhimen South Street, Dongcheng District
Beijing, 100700
China
Phone: 86 13521781839
Email: wlxing@126.com

Related Article:

Correction of: https://www.researchprotocols.org/2026/1/e84803

(JMIR Res Protoc 2026;15:e84867) doi:10.2196/84867

In "Real-World Effectiveness and Safety of Liuwei Dihuang Pill for Menopausal Syndrome: Protocol for a Prospective, Observational, Multicenter Cohort Study" [1], the authors made several corrections.

The following changes were made to the author affiliations:

First, the affiliation for author Y Ye has been updated. Y Ye is affiliated with institution 3. Institution 4 in the original manuscript has been removed.

Second, affiliation 6 has been corrected from the following:

> *Hepingli Community Health Service Center, Beijing, China*

Affiliation 6 now reads:

> *Hepingli Community Health Service Center, Dongcheng District, Beijing, China*

Third, affiliation 7 has been corrected from the following:

> *Beijing Purun Hospital, Beijing, Beijing, China*

Affiliation 7 now reads:

> *Beijing Puren Hospital, Beijing, China*

Fourth, affiliation 8 has been corrected from the following:

> *Tsinghua University Yuquan Hospital (Tsinghua University Hospital of Integrative Traditional Chinese and Western Medicine), Beijing, Beijing, China*

Affiliation 8 now reads:

> *Tsinghua University Yuquan Hospital (Tsinghua University Hospital of Integrative Traditional Chinese and Western Medicine), Beijing, China*

Fifth, affiliation 9 has been corrected from the following:

XSL•FO
RenderX

Chaoyangmen Community Health Service Center, Beijing, China

Affiliation 9 now reads:

Chaoyangmen Community Health Service Center, Dongcheng District, Beijing, China

Finally, affiliation 10 has been corrected from the following:

Guang'anmen Hospital Jinan Hospital, China Academy of Chinese Medical Sciences, Beijing, China

Affiliation 10 now reads:

Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China

The corrected author affiliations will appear as follows:

Shuoshuo Wei[1], MM; Hongyan Zhang[1], MMSc; Mengmeng Wang[1], MMSc; Yuhui Sun[2], MD; Tao Song[2], MD; Weifang Zheng[3], MBBS; Yunmei Ye[3], MBBS; Liran Ma[4], MM; Yunshuang Yang[5], PhD; Tong Wei[6], MM; Lan Cheng[7], MM; Chunquan Sun[8], MD; Xingting Ding[9], MM; Shichu Zhao[7], PhD; Qingqiao Song[10*], MD; Lianxin Wang[1*], PhD*

[1]Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, Beijing, China

[2]First Affiliated Hospital of Harbin Medical University, Heilongjiang, China

[3]Lanxi Hospital of Traditional Chinese Medicine, Zhejiang, China

[4]Beijing Gulou Hospital of Traditional Chinese Medicine, Beijing, China

[5]Beijing Longfu Hospital, Beijing, China

[6]Hepingli Community Health Service Center, Dongcheng District, Beijing, China

[7] Beijing Puren Hospital, Beijing, China

[8]Tsinghua University Yuquan Hospital (Tsinghua University Hospital of Integrative Traditional Chinese and Western Medicine), Beijing, China

[9]Chaoyangmen Community Health Service Center, Dongcheng District, Beijing, China

[10]Guang'anmen Hospital, China Academy of Chinese Medical Sciences, Beijing, China

[*]these authors contributed equally

In the Study Participants and Acknowledgments section, the hospital name was corrected from the following:

Beijing Purun Hospital

The hospital name has been corrected to the following:

Beijing Puren Hospital

These corrections do not affect the scientific conclusions of the article.

The correction will appear in the online version of the paper on the JMIR Publications website, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Reference

1. Wei S, Zhang H, Wang M, Sun Y, Song T, Zheng W, et al. Real-World Effectiveness and Safety of Liuwei Dihuang Pill for Menopausal Syndrome: Protocol for a Prospective, Observational, Multicenter Cohort Study. JMIR Res Protoc 2026 Mar 25;15:e84803 [FREE Full text] [doi: 10.2196/84803] [Medline: 41879810]

Submitted 02.Apr.2026; this is a non–peer-reviewed article;accepted 09.Apr.2026; published 21.Apr.2026.

Please cite as:
Wei S, Zhang H, Wang M, Sun Y, Song T, Zheng W, Ye Y, Ma L, Yang Y, Wei T, Cheng L, Sun C, Ding X, Zhao S, Song Q, Wang L
Correction: Real-World Effectiveness and Safety of Liuwei Dihuang Pill for Menopausal Syndrome: Protocol for a Prospective, Observational, Multicenter Cohort Study
JMIR Res Protoc 2026;15:e84867
URL: https://www.researchprotocols.org/2026/1/e84867
doi:10.2196/84867
PMID:23725242

©Shuoshuo Wei, Hongyan Zhang, Mengmeng Wang, Yuhui Sun, Tao Song, Weifang Zheng, Yunmei Ye, Liran Ma, Yunshuang Yang, Tong Wei, Lan Cheng, Chunquan Sun, Xingting Ding, Shichu Zhao, Qingqiao Song, Lianxin Wang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is

XSL•FO
RenderX

properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluating Experiences With the Newly Enacted Law on Assisted Suicide in Austria: Protocol for an Interdisciplinary Mixed Methods Study

Klara Doppler[1*], Mag; Tamina-Laetitia Vielgrader[1*], MSc; Julia Fischer[1], PhD; Helene Mayer[1], MA; Jana Marica Hluch[1], BA; Marlene Walser[2], Mag; Sabine Parrag[1], Mag; Stefan Dinges[1], Dr Theol; Karl Stöger[1,2], Prof Dr; Thomas Wochele-Thoma[3], Dr med; Maria Kletecka-Pulker[1,3], Priv Doz Dr

[1]Department for Ethics and Law in Medicine, University of Vienna/Medical University of Vienna, Spitalgasse 2-4, Hof 2.8, Vienna, Austria

[2]Department of Constitutional and Administrative Law, Faculty of Law, University of Vienna, Vienna, Austria

[3]Ludwig Boltzmann Gesellschaft, Ludwig Boltzmann Institute Digital Health and Patient Safety, Vienna, Austria

[*]these authors contributed equally

Corresponding Author:
Klara Doppler, Mag
Department for Ethics and Law in Medicine, University of Vienna/Medical University of Vienna, Spitalgasse 2-4, Hof 2.8, Vienna, Austria

Abstract

Background: Austria has only recently established a legal framework for assisted dying. Individuals seeking assistance in suicide must navigate a multistage process that has been criticized for its complexity both for the individuals seeking assistance and for the health care and legal professionals involved in the procedure.

Objective: To address the critical gap in empirical research that sheds light on the law's implementation and its impact on the lived experiences of those involved, as well as on the broader health care sector, we have designed an evaluation that examines the various perspectives of individuals seeking assisted suicide; their family members and other relatives; and all professionals specified in the legal framework—notably physicians, notaries and other relevant legal professionals, pharmacists, and psychologists.

Methods: Our study uses an interdisciplinary approach that integrates theories, concepts, and methodologies from legal science, ethics, and social science. The legal and ethical components examine aspects of the Austrian legal framework, such as the legal foundations, moral implications, and ethical challenges involved in assisted dying. The main part of this study lies in its comprehensive social science section, which consists of 2 online cross-sectional surveys. The surveys assess sociodemographic characteristics, knowledge and understanding of the framework, encounters and experiences with assisted suicide, and reported need for support and information on the legalization of assisted suicide. In parallel, we are in the process of conducting semistructured in-depth interviews with (1) individuals seeking assistance in dying, (2) family members and other relatives of the individuals seeking assistance in dying, and (3) professionals mentioned in the legal framework (eg, physicians). Thematic analysis will be used to interpret the interview data.

Results: The initial research proposal received approval from the University of Vienna Ethics Committee in November 2023. The legal and ethical analyses are ongoing. Recruitment for interview participants began in January 2024. In total, 49 participants have been interviewed as of February 2026. Given the novelty of the legislation and the limited practical experience to date, further interviews are scheduled until April 2026. Data collection for both surveys took place from June to October 2024. The final samples consisted of 239 medical and nursing directors and 304 physicians. Data analysis was completed in May 2025. Completion and dissemination of all study components are anticipated by the end of 2026.

Conclusions: Our multiperspective evaluation aims to assess the framework on assisted suicide in Austria. By evaluating the perspectives of relevant key stakeholders, we aim to provide a nuanced understanding of the law's societal, legal, and ethical implications.

International Registered Report Identifier (IRRID): DERR1-10.2196/86740

(JMIR Res Protoc 2026;15:e86740) doi:10.2196/86740

KEYWORDS

assisted suicide; assisted dying; end of life; interdisciplinary studies; empirical research; legal aspects; medical ethics; social sciences; stakeholder participation

XSL•FO
RenderX

Introduction

Background

In January 2022, Austria established a legal framework for assisted suicide, making it one of the most recent countries to regulate assisted dying amid intensifying global debate and a growing number of nations either adopting or considering relevant legal frameworks [1-7].

The political debate on self-determined dying and assisted suicide in Austria had been underway for many years, increasing as more European countries introduced legal frameworks. However, apart from a 2015 statement by the Austrian Bioethics Commission advocating reform [8], no significant political steps were taken until the Austrian Constitutional Court's 2020 ruling, which declared the outright prohibition of assisted suicide to be unconstitutional [9], ultimately leading to the introduction of the Sterbeverfügungsgesetz (Dying Decree Act) on January 1, 2022. The law supplements the Austrian Criminal Code and introduces the concept of the "dying decree" (*Sterbeverfügung*). A dying decree is a formal written declaration in which an individual who has decided to pursue assisted suicide records their voluntary and autonomous decision to do so. Health care and legal professionals are involved in preparing a dying decree, which permits the substance pentobarbital to be dispensed by a public pharmacy.

The eligibility criteria for obtaining assistance in suicide are shaped by a restrictive legal framework, ultimately limiting access to this option to a small group of individuals. To qualify, individuals must be of legal age and fully capable of making informed decisions. Additionally, the option is only available to individuals who are either Austrian citizens or who have permanent residence in Austria.

Moreover, the law enforces strict conditions with respect to the individual's state of health: a dying decree may only be issued to people with either an incurable terminal illness or a severe, permanent illness with persistent symptoms that permanently impair their overall quality of life. In both cases, the illness must result in a subjectively unbearable level of suffering that cannot otherwise be alleviated.

Beyond these strict personal requirements, individuals who wish to terminate their lives must navigate significant hurdles during the subsequent procedural steps of preparing a dying decree. The formal process begins with a medical assessment by 2 physicians, one of whom must be certified in the field of palliative medicine. If there is an indication that the individual has a psychological condition that could prompt the desire to end their life, this must additionally be assessed by a psychiatric specialist or clinical psychologist.

After a 12-week waiting period (shortened to 2 weeks if the terminal phase has already been reached), the next step is for a notary or legally qualified employee of a public patient advocacy institution to formally establish or issue the decree. The process is then finalized by providing the individual who wishes to terminate their life via dying decree with a lethal preparation of pentobarbital that can be dispensed by a public pharmacy (Textbox 1).

Textbox 1. Dying Decree Act, Austria (2022).

Eligibility criteria

- Permanent residence in Austria or Austrian citizenship
- At least 18 years of age
- Full capability of making an informed decision
- An incurable terminal illness or a severe, permanent illness with persistent symptoms, the effects of which permanently impair the person's entire lifestyle; the illness must result in a level of suffering that cannot otherwise be alleviated

Procedural requirements

- Two independent eligibility assessments by physicians, one of whom must be certified in the field of palliative medicine
- The physicians must inform the individual seeking assisted suicide about (1) alternative courses of action; (2) the availability of psychotherapeutic and other counseling services; and (3) the dosage, accompanying medication, and method of administration of the lethal preparation
- Optional third assessment by a psychiatric specialist or clinical psychologist in cases in which there is an indication that the individual seeking assisted suicide has a psychological condition
- Waiting period of 12 weeks between the first medical consultation and the effective issue of a dying decree; this period can be shortened to 2 weeks if the illness has already progressed to the terminal phase
- The decree is formally issued by a notary or a legally qualified employee of a public patient advocacy institution
- Proof of the valid dying decree must be presented when the preparation is dispensed by the pharmacy

In terms of the final steps of assisted suicide, the law provides no further procedural guidelines. In Austria, assisted suicide is not institutionalized, nor are physicians or other health care providers obligated to accompany individuals through the last step. The legislator has intentionally left the final steps of assisted suicide largely open-ended. In practice, this lack of guidance has resulted in hurdles for all parties involved.

While the Austrian legal framework is broadly in line with international practices [1-7], the regulatory approach toward the final step of the assisted suicide process is comparatively

XSL•FO
RenderX

atypical. In most other jurisdictions—with the exception of Germany (which lacks a specific statutory framework) and Switzerland (where the law does not formally mandate such involvement, although, in practice, a highly structured process ensures professional participation)—medical professionals hold clearly defined responsibilities regarding the application of the lethal preparation [1]. In contrast, the Austrian model leaves the application of the barbiturate unsupervised, lacking a formal requirement for health care professionals to accompany the individual at the time of the assisted suicide.

The unique challenges posed by the Dying Decree Act in Austria have been the subject of theoretical discourse, particularly in the legal [10-13] and ethical [14-16] realms. The law was highly controversial and widely debated, which eventually brought it to the Austrian Constitutional Court for examination. In December 2024, the Constitutional Court confirmed that most provisions of the Dying Decree Act conformed to the Federal Constitution. Two provisions (the need for full renewal of a dying decree after 1 year and the strict ban on information about relevant services) were declared unconstitutional.

However, 3 years after the law's implementation, by early 2025, there remains a critical lack of empirical research examining the law's implications for those directly affected. Our study aims to address this gap by analyzing the lived experiences of individuals navigating the Austrian assisted suicide framework and evaluating its legal implementation. This approach is designed to contribute to the existing scholarship in evaluating assisted dying laws [17-21]. In addition to our scholarly objective, we aim to provide insights that may inform policy development and professional practice. To this end, we will disseminate our findings both within academia and beyond.

Aims and Objectives

Our multiperspective evaluation aims to assess the newly enacted Dying Decree Act. By comparing and contrasting the perspectives of relevant key stakeholders, we aim to provide a nuanced understanding of the law's societal, legal, and ethical implications while also exploring its practical consequences for the health care system and health care professionals. We define key stakeholders as individuals seeking assistance in dying; their family members and other close relatives; and all the professionals specified in the legal framework, notably physicians, notaries, other relevant legal professionals, pharmacists, and psychologists. Thus, our evaluation is grounded in a perspectivist approach, which maintains that any assessment of the law must be built upon the perspectives of those who are centrally affected by it, meaning the key stakeholders, as listed above [22,23]. Our study uses an interdisciplinary approach that incorporates theories, concepts, and methodologies from legal science, ethics, and social science.

Research Questions

The research questions are as follows:

- What ethical and philosophical challenges arise from the implementation of the Dying Decree Act?
- How do key stakeholders experience and evaluate the law?
- What legal and practical issues arise from the implementation of the Dying Decree Act?

Methods

Legal Study Section

The study sets out to examine the Austrian legal framework for assisted suicide applying generally recognized methods of legal interpretation [24,25]. This analysis will be complemented by a comprehensive legal literature review of academic law journals, databases, and library resources. The review will focus on legal texts and their scholarly interpretation and dissemination, excluding empirical studies. Our aim lies in generating a structured understanding of the legal framework governing assisted suicide in Austria. Extracted material will be organized thematically according to key legal topics (eg, eligibility, procedural safeguards, and institutional responsibilities) to provide a structured narrative of the legal landscape. Searches will be conducted using relevant keywords related to the applicable statutory provisions.

Beyond the Austrian context, the study will also explore alternative regulatory approaches. This includes an analysis of other national and European Union (EU) provisions, guidelines, rulings, and proposed solutions. Particular attention will be paid to the case law of the European Court of Human Rights and further relevant EU documents as a means of illustrating the European framework in which the discourse on assisted suicide is embedded. In addition to EU regulations, the study will include a comparative legal analysis of how other nations regulate assisted suicide. By comparing these approaches with the Austrian framework, the study aims to provide both an overview of international best practice models and an insight into the Austrian legislators' legal reasoning. The objective is twofold: (1) to understand the reasoning behind the Austrian legislators' choices and (2) to apply a comparative perspective that provides the scientific community with new insights.

The legal analysis will be conducted by the authors in collaboration with legal experts from the Forum on End-of-Life Autonomy (Forum Autonomie am Lebensende). This forum is an expert panel that was established in 2022 and is headed by the Department for Ethics and Law in Medicine of the University of Vienna (Institut für Ethik und Recht in der Medizin). The forum members meet on a regular basis to discuss the practical implementation of the legal framework surrounding assisted suicide. It comprises experts from various fields with practical insights into the law and its application. By involving the forum, the study ensures that legal questions will be addressed and answered from multiple perspectives and areas of expertise.

Ethics Study Section

The primary objective of this section of the study is to evaluate the moral implications and ethical challenges of the Dying Decree Act from a philosophical and applied ethical perspective. This is achieved by analyzing the line of argumentation used in the process of legalizing assisted suicide and its manifestation in the legal framework in Austria. We will then consider whether the Dying Decree Act can adequately address ethical concerns at the level of the individual, the health care system, and society as a whole.

XSL•FO
RenderX

Thus, we base our ethical reflections on a multiperspective, systemic approach. At the micro level, we focus on patient–health care professional interactions. Our applied ethics are guided by the 4 bioethical principles [26] as well as interprofessional health care ethics, referencing nursing and medical ethics as professional ethics. At the meso level, we focus on interactions between decision-makers and teams in health care settings, applying organizational ethics theories [27] from a systemic perspective. These highlight the urgent need for team and organizational development to implement the Dying Decree Act. At the macro level, it is important to clarify the various moral arguments influencing social discourse and reflect on managing tensions and conflicts in the Austrian health care system. This necessitates the active involvement of health care professionals, policymakers, and patients, ensuring that diverse perspectives are represented. It also demands careful attention to and respect for pluralism within a solidarity-based health care system. Therefore, the ethical analysis will consider practical consequences at the level of patients, health care professionals, and the health care system as a whole.

Building on this framework, we will select philosophical and ethical approaches equipped to identify and reflect on potential moral concerns associated with assisted suicide. Our aim is to show how these concerns and objections are more or less meaningful according to whether they target individuals, organizations, or society. Second, we will highlight those arguments and objections that arise during the legalization process and are specifically considered in the Dying Decree Act to discuss their consequences and implications for individuals, organizations, and society. We will focus primarily on the Austrian Constitutional Court texts as well as the legal text of the Dying Decree Act itself. Third, our ethical reflection is integrated into both the quantitative and qualitative research components.

Social Science Study Section

Surveys

To date, the perspectives of physicians and those in leadership roles, such as medical and nursing directors, on this subject remain largely unexplored [28]. To address this gap, 2 online questionnaires were developed. One survey targets physicians of all specialties, whereas the other is targeted at directors of hospitals and long-term care facilities. By engaging both groups, the questionnaires provide valuable insights into their experiences with assisted suicide cases in Austria. Additionally, the surveys aim to determine whether physicians and directors of hospitals and nursing homes feel adequately supported in navigating the complex legal framework or whether additional resources and guidance are needed.

Findings from cross-sectional studies on assisted suicide conducted in various countries informed the development of the questionnaires [29-32]. In addition, experts in medical law, ethics, social science, public health, psychology, and medicine contributed to refining the questions and provided further input on the questionnaires' structure and content. Before distribution, the questionnaires were pretested by the researchers regarding technical functionality.

We obtained a list of all hospitals and nursing facilities in Austria from the Ministry of Health and sent the questionnaires to each institution via email. For the physicians' questionnaire, we asked the Austrian Medical Chamber to distribute the questionnaire to all practicing physicians in Austria.

Both surveys consisted of a mix of closed- and open-ended questions (24 items in the medical and nursing directors' survey and 33 items in the physicians' survey). The surveys commenced with sociodemographic questions regarding age, gender, and length of employment at their respective institutions. The main part of the questionnaire can thematically be grouped into (1) knowledge and understanding of the Dying Decree Act, (2) encounters and experiences with assisted suicide, (3) institutional guidelines, and (4) participants' reported need for more information and support regarding assisted suicide. Considering the sensitive nature of the study, participants were allowed to skip questions.

Both questionnaires were programmed using the online tool SoSci Survey (version 3.5.02; SoSci Survey GmbH) [33]. The Cochran formula was used to calculate the required sample size by incorporating the estimated size of the target population as well as the desired α error and confidence coefficient. An α error of 5% was selected for both questionnaires, corresponding to a confidence coefficient of 95%. As there were 50,631 registered physicians in Austria in 2023 [34], the calculation indicated that a total of 385 participants would be required to obtain a representative sample to gain a first insight into Austrian physicians' experiences and perceived challenges regarding assisted suicide. For the questionnaire targeted at medical and nursing directors, our population of interest consisted of 1251 Austrian health facilities, and we calculated that a minimum of 295 completed surveys would be needed.

We used purposive sampling by distributing the questionnaire targeted at physicians directly via the Austrian federal and state medical boards to reach our target population of physicians residing in Austria. Publicly available catalogs maintained by governmental institutions listing all nursing and medical facilities in Austria were used to directly contact the medical and nursing directors' respective facilities.

For the data analysis, we started by conducting a descriptive statistical analysis of our sample using bar charts and contingency tables. This step focused on examining the distribution of responses to each item in the questionnaires to gain an initial overview of participants' answers to our survey. On the basis of the identified patterns in the descriptive results, we subsequently conducted exploratory inferential analyses to further investigate potential associated variables within our data. As the data derived from the questionnaires primarily consisted of dichotomous (yes or no) variables and categorical variables, which correspond to a nominal scale level, participants' answers were mainly analyzed using nonparametric methods, such as chi-square tests. ANOVAs and 2-tailed *t* tests were used to investigate differences in relation to the few metrically scaled items contained in the questionnaire (eg, Likert scales).

An English translation of the complete surveys can be found in the designated Open Science Framework (OSF) repository [35]. The CHERRIES (Checklist for Reporting Results of Internet

XSL•FO
RenderX

E-Surveys) checklist was used to guide the reporting of the survey [36] as provided in Checklist 1.

Semistructured Interviews

We are conducting semistructured in-depth interviews with (1) individuals seeking assistance in dying; (2) family members and other relatives of individuals seeking assistance in dying; and (3) professionals mentioned in the legal framework, notably physicians, notaries and other relevant legal professionals, pharmacists, and psychologists. Interviews are conducted either online via a videoconferencing platform or in person at a location chosen by the participants (eg, their workplace or home). All interviews are audio recorded using the recording function of the corresponding videoconferencing platform or a digital voice recorder and subsequently transcribed verbatim by a professional transcriber. In addition, field notes are taken to document contextual details and supplement the content of each interview.

We recruit participants nationwide across all stakeholder groups to identify potential regional differences (Table 1). Calculations about the number of interviews and when to stop data collection cannot be determined in advance of analysis as such calculations are not consistent with the interpretative epistemological foundation of our interview study. In an iterative process, we recruit participants in parallel with transcribing and analyzing interviews to best determine when saturation is reached. Saturation is evaluated separately for each interview sample group and defined as the point at which no new information, codes, or themes emerge [37]. This process, in turn, ensures a comprehensive and thorough understanding of the law's effects.

Table . Recruitment procedures for the interview study.

	Professionals mentioned in the legal framework	Individuals (or their family members and other relatives) seeking assistance in dying
Inclusion criteria	Professionals working in Austria with relevant experience according to their legally defined role in the Dying Decree Act	Adults (aged ≥18 years) residing in Austria (or their adult relatives) currently seeking or planning on seeking assistance in dying
Recruitment strategy	Nationwide recruitment via official professional associations and the Forum on End-of-Life Autonomy, supplemented by snowball sampling	A call for participation will be disseminated through relevant organizations and professionals connected to the practice of assisted suicide in Austria; if interested, potential interview participants will be able to contact the research group via the contact information provided in the call for study participation

All interviews follow an interview guide developed by the research team. An outline of the interview guide is provided in Multimedia Appendix 1, whereas the full version can be found in the designated OSF repository [35]. Different versions of the interview guide were prepared and adapted for each key stakeholder, tailoring them to the particularities of the stakeholders' specific roles. However, the different versions of the interview guide are similar in terms of the central questions on the interviewees' personal attitudes toward assisted suicide, their experiences with assisted suicide in Austria and potentially other countries, and their opinion on the Austrian law. Throughout the course of the study, the interview guides are continuously revised, expanded, and adapted. Interviews are prepared to accommodate unplanned questions that arise during the course of the dialogue. All interviewers have training and experience in leading interviews prior to the study and are familiar with the sensitive nature of the research topic. Furthermore, most interviews are conducted in pairs—with one lead interviewer and a notetaker also tasked with ensuring that all aspects of the interview guide are addressed.

We are conducting a thematic analysis to interpret and make sense of the interview data following the approach outlined by Braun and Clarke [38,39] and using the MAXQDA software (VERBI GmbH) [40]. This method facilitates comprehensive understanding of the dataset by identifying and analyzing latent patterns that encapsulate the overarching narrative conveyed by the data. Thematic analysis provides a framework for uncovering meaningful insights by systematically exploring how various aspects of the data interconnect. The process of analysis begins with an in-depth familiarization with the data, achieved through the repeated and immersive reading of transcribed interviews. To develop themes, we systematically search the interview transcripts for recurring passages and shared ideas, identifying clusters of meaning that reflect common messages or concepts. These passages are then compared and collated, enabling preliminary themes to be articulated. The next phase is to review and refine these themes by consistently relating them back to the original interview data, ensuring that they remain grounded in the participants' perspectives and experiences. To strengthen interpretative validity and reduce potential researcher bias, the analysis is conducted collaboratively by multiple members of the research team. Furthermore, the researchers debrief on a regular basis in structured team discussions as a reflexivity practice [41].

Final reporting of the qualitative component of the study will be in accordance with the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines as provided in Checklist 2, which ensures transparent and comprehensive reporting of the study design, data collection, and analysis [42].

Ethical Considerations

This study is designed to comply with the Declaration of Helsinki. The University of Vienna Ethics Committee approved the initial research proposal on November 27, 2023 (01071). An addendum was submitted in September 2024, to which ethical consent was granted on November 4, 2024.

XSL•FO
RenderX

All participants were asked to provide informed consent prior to taking part. For the online survey, participants provided informed consent by clicking a confirmation button prior to proceeding to the questionnaire. Before being asked to provide their consent, participants were presented with detailed information about the study, including its purpose, the voluntary nature of participation, and data protection measures. The survey was conducted anonymously. Providing contact details for potential participation in a follow-up interview was optional, and they were collected separately from the survey responses.

For the interview study, all participants are asked to provide written informed consent using the official consent form issued by the Ethics Committee of the University of Vienna. Participation is voluntary, and participants may withdraw from the study at any time without providing a reason. Due to the sensitive nature of the topic, interview participants who are either individuals seeking assistance in dying or family members or relatives will also be informed that professional counseling services are available and may be accessed through the study team if desired.

No financial compensation or incentives were offered for participation.

Furthermore, all applicable data protection regulations, including the EU's General Data Protection Regulation, have been strictly adhered to, and concise rules on data security have been formulated (eg, all data will be deleted when the research is concluded and results are published). All interviews will be transcribed by a professional transcriber who is subject to strict confidentiality requirements. All identifying details (eg, names and places) will be removed after transcription. Transcripts will be coded using unique identification numbers in accordance with pseudonymization procedures, and identifying information will be stored separately in a password-protected file within the secure university data system until completion of the project.

Results

The study was funded in November 2023. The literature review and data collection for the legal and ethical analyses commenced in January 2024 and will continue until the study's completion. Recruitment and data collection for the interview study began in January 2024 and are expected to be completed by April 2026. In total, 49 participants have been interviewed as of February 2026, including 5 (10.2%) individuals seeking assistance in dying, 10 (20.4%) family members and other relatives of individuals seeking assistance in dying, and 34 (69.4%) professionals mentioned in the legal framework (with physicians making up the largest demographic group with n=16, 47.1% interviewees). One online survey was launched in June 2024, the other in August 2024. The final samples consisted of 239 medical and nursing directors and 304 physicians. Data collection ended in October 2024, whereas data analysis was completed in May 2025. Completion and dissemination of all study components are anticipated by the end of 2026. Given the novelty of the legislation and the limited practical experience to date, further interviews are scheduled until April 2026.

Discussion

Expected Contribution

Our study aims to provide a comprehensive understanding of the lived experiences and practical challenges encountered by all stakeholders involved in assisted suicide under Austria's newly established legal framework. By examining these perspectives, the study seeks to address critical gaps in empirical research and inform future policy and practice. Previous scholarship has highlighted that end-of-life regulation cannot be adequately understood through isolated legal or ethical analysis alone but requires a holistic approach [43]. In line with this identified need, our study uses an interdisciplinary approach to evaluate Austria's new framework on assisted suicide, bringing stakeholders' perspectives and lived experiences to the forefront. The study components, which incorporate multiple stakeholder perspectives, will aim to advance academic discourse by examining the experiences of (1) individuals seeking assistance in dying, (2) their family members and other relatives, and (3) involved professionals. Thus, this study generates empirical insights into how the legal framework shapes clinical roles, responsibilities, and accessibility. Ultimately, the study aims to bridge the gap between theoretical discussions on assisted suicide and the practical realities faced by those involved by offering evidence-based recommendations that are relevant not only to Austria but also to international debates on assisted dying legislation.

Potential Findings and Comparison With Existing Literature

With our project, we aim to contribute to the international landscape of research on assisted dying regulation. It is essential to recognize that cultural and legal differences necessitate caution when transferring findings across countries. It is crucial to consider the lived experiences of all involved parties within their specific national context. An example illustrating this need is the variation in physicians' attitudes and practices regarding assisted dying across countries [44-46].

In other countries, studies have already adopted multiperspective and interdisciplinary approaches to gain a comprehensive understanding of national practices and their impact on all involved stakeholders [18,43]. In Austria, research on assisted suicide has largely been limited to specific stakeholder groups [47-50], leaving a gap in understanding the interconnected nature of these groups' various experiences. Our study is the first holistic approach evaluating the Austrian legal framework.

We anticipate that the framework's lack of guidance regarding the final steps—where health care providers are neither required nor formally instructed to accompany individuals—creates practical uncertainties across multiple stakeholder groups. This may shape the experiences of individuals seeking assisted dying, their relatives, and health care professionals by influencing access, redistributing responsibilities, and creating uncertainty regarding professional roles. Health care institutions, including hospitals and nursing homes, may also encounter procedural ambiguities due to the informal handling of assisted suicide.

XSL•FO
RenderX

Additional anticipated findings include physicians' reasoning regarding participation in the assessment process and eligibility determination, the (emotional) impact of assisted dying on relatives, and practical or financial barriers influencing equitable access to assisted suicide.

Strengths and Limitations

The study sets out to provide a robust and practical study design capable of providing a thorough evaluation of the newly enacted law on assisted suicide from a multiperspective standpoint. Its large-scale approach is enabled by an interdisciplinary research team and a clear emphasis on including all stakeholder groups affected by the legal framework. This inclusive perspective—combined with a mixed methods approach suitable for collecting both qualitative and quantitative data—positions the study as the most extensive research design on this topic in Austria. As such, it is able to make a significant contribution to the research in this highly relevant field.

We are aware that our evaluation will face several challenges that may turn into limitations of our study. First, the recent enactment of the law may lead to a lack of routine in its practical implementation, which could cause the relevance of some identified issues to shift over time. Second, given the highly sensitive and personal nature of assisted dying, recruiting study participants—particularly individuals seeking assisted suicide and their family members—may prove challenging. Third, we need to take care of avoiding response bias. Study participants may be hesitant to fully disclose their experiences or opinions due to fear of judgment or perceived legal and professional repercussions. We aim to mitigate this by using state-of-the-art interviewing techniques and ensuring strict confidentiality throughout the research process.

Dissemination

The results of this study will be disseminated to a broad audience through a variety of publications and international and national conferences. To maximize accessibility and facilitate unrestricted knowledge exchange, the results will be submitted to journals that support open-access publication. Additionally, the study has been preregistered through the OSF repository [35]. For the general public, the findings will be presented in plain, accessible language on the designated website. Finally, study results will be presented at roundtable discussions and expert panels organized by the Department of Ethics and Law in Medicine to ensure public engagement and knowledge exchange until the final research findings can be presented at the institute's closing conference.

Acknowledgments

The authors thank their study assistants, Sarah Franziska Ulmer and Noah Fida, for their invaluable support in editing the manuscript. Furthermore, they extend special thanks to all members of the Forum on End-of-Life Autonomy (Forum Autonomie am Lebensende). They are also grateful to the study participants, whose time, openness, and effort made this research possible. The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading and editing. The GenAI tool used was GPT-4.5 (OpenAI). Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes. The final draft of the manuscript was proofread by a professional language editor.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Open access funding was provided by the University of Vienna.

Data Availability

The datasets generated and analyzed for the survey component are available from the corresponding author on reasonable request. Due to privacy and ethical considerations, the qualitative interview datasets are not publicly available. The interview guides and surveys are provided in Multimedia Appendix 1. The study protocol and all generated results will be registered on the Open Science Framework to enhance transparency and accessibility.

Authors' Contributions

Conceptualization: KD, TLV, JF, SP, MK-P
Data curation: KD, TLV, JF, JMH
Formal analysis: KD, TLV, JF, JMH
Investigation: KD, TLV, JF, JMH
Methodology: KD, TLV, JF, JMH, SD, SP, MK-P
Project administration: KD, TLV, MK-P
Software: JMH, TLV, SP
Supervision: JF, KS, SF, MK-P
Validation: JMH, TLV
Visualization: HM, JMH

XSL•FO
RenderX

Writing—original draft: KD, TLV, JF, HM, JMH, SP, KS, MK-P
Writing—review and editing: KD, TLV, JF, JMH

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide outline.
[PDF File, 207 KB - resprot_v15i1e86740_app1.pdf]

Checklist 1
CHERRIES checklist.
[PDF File, 296 KB - resprot_v15i1e86740_app2.pdf]

Checklist 2
COREQ checklist.
[PDF File, 182 KB - resprot_v15i1e86740_app3.pdf]

References

1. White BP. Research Handbook on Voluntary Assisted Dying Law, Regulation and Practice: Edward Elgar Publishing; 2025.
2. Dyer O, White C, García Rada A. Assisted dying: law and practice around the world. BMJ 2015;351:h4481. [doi: 10.1136/bmj.h4481] [Medline: 26290517]
3. Mroz S, Dierickx S, Deliens L, Cohen J, Chambaere K. Assisted dying around the world: a status quaestionis. Ann Palliat Med 2021 Mar;10(3):3540-3553. [doi: 10.21037/apm-20-637] [Medline: 32921084]
4. Richardson S. An international expansion in voluntary euthanasia/assisted dying: the implications for nursing. Int Nurs Rev 2023 Mar;70(1):117-126. [doi: 10.1111/inr.12807] [Medline: 36271824]
5. Krüger JM. Comparing the decriminalisation of assisted dying in Europe. Eur J Health Law 2022 Oct 26;30(3):243-271. [doi: 10.1163/15718093-bja10101] [Medline: 37582531]
6. Grosse C, Grosse A. Assisted suicide: models of legal regulation in selected European countries and the case law of the European Court of Human Rights. Med Sci Law 2015 Oct;55(4):246-258. [doi: 10.1177/0025802414540636] [Medline: 25143341]
7. Looi MK. Assisted dying laws around the world. BMJ 2024;387:q2385. [doi: 10.1136/bmj.q2385] [Medline: 39477363]
8. Sterben in würde: empfehlungen zur begleitung und betreuung von menschen am lebensende und damit verbundene fragestellungen. Bundeskanzleramt. 2015. URL: https://www.bundeskanzleramt.gv.at/dam/jcr:6d5d655b-e11d-4e65-b6a6-9fd531a033fb/sterben_in_wuerde.pdf [accessed 2026-02-26]
9. Erkenntnis g 139/2019-71. Verfassungsgerichtshof. 2020. URL: https://www.vfgh.gv.at/downloads/VfGH-Erkenntnis_G_139_2019_vom_11.12.2020.pdf [accessed 2026-04-09]
10. Stöger K. Bis Zur Bahre – Die Notare: Manz Verlag; 2021.
11. Birklbauer A. Möglichkeiten und Grenzen untersagter Suizidassistenz in Pflege- und Betreuungseinrichtungen. J Med Gesundheitsrecht 2022;2022(1):10-16. [doi: 10.37942/jmg202201001001]
12. Ganner M. Neues Sterbeverfügungsgesetz. ÖZPR 2021;107:180-185 [FREE Full text]
13. Kopetzki C. Ein Jahr Sterbeverfügung: Manz Verlag; 2022.
14. Niedermayr H. Sterbehilfe In Österreich: Ein Liberaler Gesetzesentwurf und Dessen Ethische Erörterung Anhand Ausgewählter Fragestellungen: Jan Sramek; 2022.
15. Feichtner A, Körtner U, Likar R, Watzke H, Weixler D. Assistierter Suizid: Hintergründe, Spannungsfelder und Entwicklungen: Springer; 2022.
16. Halmich M, Klein A. Sterbehilfe/Suizidbeihilfe in Österreich: Educa Verlag Halmich; 2021.
17. Young J, Dehkhoda A, Ahuriri-Driscoll A, et al. Exploring the early experiences of assisted dying in Aotearoa New Zealand: a qualitative study protocol. BMJ Open 2024 Oct 9;14(10):e090118. [doi: 10.1136/bmjopen-2024-090118] [Medline: 39384236]
18. Schwabe S, Herbst FA, Stiel S, Schneider N. Suicide assistance in Germany: a protocol for a multi-perspective qualitative study to explore the current practice. PLoS One 2024;19(12):e0311880. [doi: 10.1371/journal.pone.0311880] [Medline: 39671457]
19. Shapiro GK, Tong E, Nissim R, et al. Exploring key stakeholders' attitudes and opinions on medical assistance in dying and palliative care in Canada: a qualitative study protocol. BMJ Open 2021 Dec 3;11(12):e055789. [doi: 10.1136/bmjopen-2021-055789] [Medline: 34862301]

XSL•FO
RenderX

20. Vallès-Poch M, Parra Jounou I, Ortega-Lozano R, et al. End-of-life narratives of patients who request medical assistance in dying: a qualitative study protocol. Int J Qual Methods 2023;22. [doi: 10.1177/16094069231202196]
21. Verdaguer M, Vallès-Peris N, Busquet-Duran X, et al. Implementation of assisted dying in Catalonia: impact on professionals and development of good practices. Protocol for a qualitative study. Int J Qual Methods 2023;22. [doi: 10.1177/16094069231186133]
22. Berghofer P. Scientific perspectivism in the phenomenological tradition. Eur J Philos Sci 2020;10(3):30. [doi: 10.1007/s13194-020-00294-w] [Medline: 32566043]
23. Reginster BH. The paradox of perspectivism. Philos Phenomenol Res 2007;62(1):217-233. [doi: 10.1111/j.1933-1592.2001.tb00053.x]
24. Bydlinski P. Grundzüge der Juristischen Methodenlehre, 4th edition: Facultas Verlags- und Buchhandels AG; 2023.
25. Bydlinski F. Juristische Methodenlehre und Rechtsbegriff, 2nd edition: Springer; 1982.
26. Beauchamp TL, Childress JF. Principles of Biomedical Ethics, 8th edition: Oxford University Press; 2019.
27. Woellert K. Klinische Ethik systemisch betrachtet – vom Einfluss systemischer Grundannahmen und Methoden auf die Gestaltung einer effektiven Ethikberatung. Ethik Med 2022 Dec;34:529-548. [doi: 10.1007/s00481-022-00710-0]
28. Lenhart A, Weixler D, Schaden E, Fruhwald SM, Friesenecker B. Assistierter suizid ist noch keine„area of interest“ für ärzt:innen. Anästhesie Nachr 2024;6:1-14. [doi: 10.1007/s44179-024-00203-z]
29. Hamer MK, Baugh CM, Bolcic-Jankovic D, et al. Conscience-based barriers to medical aid in dying: a survey of Colorado physicians. J Gen Intern Med 2024 Dec;39(16):3138-3145. [doi: 10.1007/s11606-024-08782-y] [Medline: 38710866]
30. Kraak-Steenken FW, Renckens SC, Pasman HR, Bosma F, van der Heide A, Onwuteaka-Philipsen BD. Euthanasia and physician-assisted suicide in people with an accumulation of health problems related to old age: a cross-sectional questionnaire study among physicians in the Netherlands. Int J Public Health 2024;69:1606962. [doi: 10.3389/ijph.2024.1606962] [Medline: 38698912]
31. Pottash M, Saikaly K, Stevenson M, Krohmal B. A survey of clinicians who provide aid in dying. Am J Hosp Palliat Care 2024 Sep;41(9):1045-1050. [doi: 10.1177/10499091231205841] [Medline: 37776055]
32. Smets T, Cohen J, Bilsen J, Van Wesemael Y, Rurup ML, Deliens L. Attitudes and experiences of Belgian physicians regarding euthanasia practice and the euthanasia law. J Pain Symptom Manage 2011 Mar;41(3):580-593. [doi: 10.1016/j.jpainsymman.2010.05.015] [Medline: 21145197]
33. Leiner D, Leiner S. SoSci Survey: GESIS; 2021. URL: https://search.gesis.org/publication/zis-LeinerLeiner2021SoSci [accessed 2026-04-09]
34. Ärztestatistik für das Jahr 2023. Österreichische Ärztekammer. 2023. URL: https://www.aerztekammer.at/statistik-2023 [accessed 2026-02-26]
35. Vielgrader TL, Doppler K, Fischer J. Multiperspective evaluation of the Dying Decree Law in Austria. OSF Registries. Preprint posted online on Apr 23, 2025. [doi: 10.17605/OSF.IO/BGPSA]
36. Eysenbach G. Improving the quality of web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34. [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
37. Braun V, Clarke V. To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. Qualit Res Sport Exerc Health 2021 Mar 1;13(2):201-216. [doi: 10.1080/2159676X.2019.1704846]
38. Braun V, Clarke V. What can “thematic analysis” offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014;9:26152. [doi: 10.3402/qhw.v9.26152] [Medline: 25326092]
39. Braun V, Clarke V. Toward good practice in thematic analysis: avoiding common problems and be(com)ing a knowing researcher. Int J Transgend Health 2022;24(1):1-6. [doi: 10.1080/26895269.2022.2129597] [Medline: 36713144]
40. MAXQDA. URL: https://www.maxqda.com [accessed 2026-02-26]
41. Olmos-Vega FM, Stalmeijer RE, Varpio L, Kahlke R. A practical guide to reflexivity in qualitative research: AMEE Guide no. 149. Med Teach 2022:1-11. [doi: 10.1080/0142159X.2022.2057287] [Medline: 35389310]
42. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
43. White BP, Willmott L, Close E. Better regulation of end-of-life care: a call for a holistic approach. J Bioeth Inq 2022 Dec;19(4):683-693. [doi: 10.1007/s11673-022-10213-8] [Medline: 36251135]
44. Voorhees JR, Rietjens JA, van der Heide A, Drickamer MA. Discussing physician-assisted dying: physicians’ experiences in the United States and the Netherlands. Gerontologist 2014 Oct;54(5):808-817. [doi: 10.1093/geront/gnt087] [Medline: 24000266]
45. McCormack R, Clifford M, Conroy M. Attitudes of UK doctors towards euthanasia and physician-assisted suicide: a systematic literature review. Palliat Med 2012 Jan;26(1):23-33. [doi: 10.1177/0269216310397688] [Medline: 22190615]
46. Willmott L, White BP, Sellars M, Yates PM. Participating doctors’ perspectives on the regulation of voluntary assisted dying in Victoria: a qualitative study. Med J Aust 2021 Aug 2;215(3):125-129. [doi: 10.5694/mja2.51123] [Medline: 34109641]
47. Zeilinger EL, Petersen A, Brunevskaya N, et al. Experiences and attitudes of nurses with the legislation on assisted suicide in Austria. Palliat Support Care 2024 Oct 11:1-7. [doi: 10.1017/S147895152400107X] [Medline: 39390951]

XSL•FO
RenderX

48. Stolz E, Mayerl H, Gasser-Steiner P, Freidl W. Attitudes towards assisted suicide and euthanasia among care-dependent older adults (50+) in Austria: the role of socio-demographics, religiosity, physical illness, psychological distress, and social isolation. BMC Med Ethics 2017 Dec 7;18(1):71. [doi: 10.1186/s12910-017-0233-6] [Medline: 29212490]
49. Kitta A, Ecker F, Zeilinger EL, Kum L, Adamidis F, Masel EK. Statements of Austrian hospices and palliative care units after the implementation of the law on assisted suicide: a qualitative study of web-based publications. Wien Klin Wochenschr 2024 Jul;136(13-14):382-389. [doi: 10.1007/s00508-023-02157-9] [Medline: 36894787]
50. Fischer L, Le NS, Kirchner S, et al. Evaluating physicians' knowledge of assisted suicide and palliative sedation therapy in Austria: an exploratory survey study. Wien Klin Wochenschr 2025 Jun;137(11-12):357-367. [doi: 10.1007/s00508-025-02509-7] [Medline: 40035840]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
COREQ: Consolidated Criteria for Reporting Qualitative Research
EU: European Union
OSF: Open Science Framework

Edited by A Schwartz; submitted 29.Oct.2025; peer-reviewed by AVD Heide; revised version received 26.Feb.2026; accepted 13.Mar.2026; published 16.Apr.2026.

Please cite as:
Doppler K, Vielgrader TL, Fischer J, Mayer H, Hluch JM, Walser M, Parrag S, Dinges S, Stöger K, Wochele-Thoma T, Kletecka-Pulker M
Evaluating Experiences With the Newly Enacted Law on Assisted Suicide in Austria: Protocol for an Interdisciplinary Mixed Methods Study
JMIR Res Protoc 2026;15:e86740
URL: https://www.researchprotocols.org/2026/1/e86740
doi:10.2196/86740

© Klara Doppler, Tamina-Laetitia Vielgrader, Julia Fischer, Helene Mayer, Jana Marica Hluch, Marlene Walser, Sabine Parrag, Stefan Dinges, Karl Stöger, Thomas Wochele-Thoma, Maria Kletecka-Pulker. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Directly Observing and Characterizing Adolescents' Self-Generated Social Media Posts: Protocol for Creation and Implementation of a Cyberethnography Informed Codebook

Kylie Boyd[1], MPH; Lydia Bliss[1], BA; Tingting Fan[1], PhD; Kayla Kern[1], MPH; Caitlin M Carlson[1], MS; Megan A Moreno[1], MPH, MD, MSED; Christopher N Cascio[2], PhD; Ellen Selkie[1], MD, MPH

[1]Department of Pediatrics, School of Medicine and Public Health, University of Wisconsin-Madison, 800 University Bay Drive, Suite 310, Madison, WI, United States

[2]School of Journalism and Mass Communications, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Kylie Boyd, MPH
Department of Pediatrics, School of Medicine and Public Health, University of Wisconsin-Madison, 800 University Bay Drive, Suite 310, Madison, WI, United States

Abstract

Background: Adolescent social media research has primarily focused on frequency of platform use and self-report measures. There has been limited focus on the self-generated content posted by adolescents and how this might relate to their well-being.

Objective: This study describes a researcher-observed codebook for characterizing adolescents' self-generated content in a longitudinal sample.

Methods: Participants in the study provided informed assent (and parental informed consent) for researchers to follow them and passively observe their self-generated content on Instagram (Meta Platforms), TikTok (ByteDance Ltd), Facebook (Meta Platforms), and X (formerly known as Twitter; X Corp). Guided by Bronfenbrenner's social ecological biopsychosocial model, the research team created a codebook incorporating prior cyberethnographic observation of self-generated social media content. After codebook refinement, coders (research staff and student research assistants) were trained through multiple rounds of test coding, and the codebook was applied to participant data with periodic quality control measures to ensure interrater reliability.

Results: This study was funded in early 2023 and began data collection in March 2023 and will conclude in 2027. So far, the interrater reliability agreement scores (AC1) between coders have shown strong interrater reliability. For Year 1, scores were Facebook 0.89, Instagram 0.89, TikTok 0.88, X 0.87, and combined 0.88; for Year 2, Facebook 0.95, Instagram 0.96, TikTok 0.96, X 0.96, and combined 0.96. This project provides replicable guidance to categorize social media data from adolescent participants using human coders who can contextualize content through longitudinal observation. The method that our team chose and followed paved the way for many strengths to be recognized as well as lessons learned by our team that allowed for adaptation and growth to occur while this study has been ongoing.

Conclusions: Cyberethnography, the chosen method for this research protocol, has allowed this research project to collect self-generated content for adolescent social media in a comprehensive manner. Thus, allowing our team to be able to cross-analyze this data with the well-being data that are being collected under the grander project for patterns. Sharing our protocol through this paper will also allow other researchers to draw from our methodology for future projects to aid in social media and adolescent understanding.

International Registered Report Identifier (IRRID): DERR1-10.2196/84461

(JMIR Res Protoc 2026;15:e84461) doi:10.2196/84461

KEYWORDS

methodology; social media; adolescents; qualitative coding; codebook creation; cyberethnography; longitudinal observational study

Introduction

Background

Social media has become a cornerstone of adolescent daily life, with 96% of US teens aged 13 - 17 years reporting daily internet use and 46% engaging with platforms "almost constantly" [1]. These platforms (eg, TikTok [ByteDance Ltd], Instagram [Meta Platforms], Facebook [Meta Platforms], and X [X Corp]) offer adolescents unprecedented opportunities to interact with peers [2], explore their identities [3], and access global culture [4].

XSL•FO
RenderX

Although this study does cite from recent years (2023), we do recognize that social media data and information are constantly being updated, which has brought about a need for a study such as this to examine the type of content that adolescents post online.

Uncovering what specific content adolescents choose to post online is crucial, as it reflects lived experiences, developmental priorities, and cultural contexts [5,6]. These methods capture displayed content, which includes content analysis and ethnography. Within content analysis, there is objective coding but with limits to what meets our criteria [7]. Ethnographic studies of social media to date primarily focus on a narrow topic [8]. To address this gap, this study focuses broadly on cyberethnography of adolescents' social media postings. Cyberethnography, by definition, is a type of research methodology that involves viewing media online and observing interactions in a digital space [9]. In this study, we conceptualize social media posting as self-generated content (SGC), defined as material that is created and shared by adolescents themselves on social media platforms. Though there is similar terminology, such as user-generated content (UGC), SGC differs in the sense that it is specific to what the individual participant posts versus just any form of content that is posted (retweets, reposts, etc). Our study also builds on previous social media coding tools such as the Application Programming Interface (API) or Python (Python Software Foundation) to use manual human coders instead of data scraping for data collection. Our research group proposes a cyberethnographic codebook designed to systematically categorize adolescent self-generated social media content across platforms, grounded in the lived experiences of teens' posts [10].

Adolescence is a pivotal developmental period marked by milestones such as identity exploration [11,12], intimate friendship-building [13,14], and cultural socialization [15], all of which are commonly mediated through social media. Our codebook creates greater insight into these developmental periods through creation of a self-generated content categorization (coding) scheme informed by Bronfenbrenner (1979)'s ecological model [16]. Ranging from individual experiences to broader cultural systems, we unpack layers of contexts that may emerge from teens' self-generated posts in the following section.

Adolescent Social Media Posting and Development

Engagement with cultural trends, politics, global issues, and social norms situates teens within broader digital and societal contexts. Adolescents prepare for adulthood by understanding social norms and gaining insight into the world around them [17,18]. By posting sociocultural content, adolescents explore civic engagement and political views and test their values and beliefs within the context of society. Engaging with these topics online and participating in larger communities allows adolescents to express and reinforce their values, helping them feel understood and validated within their social and societal contexts [19]. This sense of validation fosters a greater sense of belonging and safety [20], which in turn enhances their self-esteem and emotional well-being [21].

Moving on to interactions with family, peers, and romantic partners reflects immediate social ecosystems. One of the core developmental tasks during adolescence is the formation and maintenance of meaningful interpersonal and social relationships. Social media plays a significant role in this process, serving as a platform where adolescents can share aspects of their personal lives and sustain social connections [22,23]. These interactions often begin within the family unit, where adolescents engage in intimate relationships and shared activities. It is common for adolescents to post content about family events or milestones, as these interactions form the foundation of their social networks, with 44% of US teens reporting that they post content about their family on social media [24]. As adolescents spend more time outside the home, particularly in school, they learn to navigate and build friendships with classmates and peers in environments separate from their families [25].

Then we will review institutional activities that support adolescent development by providing structure, fostering social connections, and creating opportunities for personal growth [26-30]. Participation in these activities also contributes to identity development by allowing youth to explore and express aspects of themselves within institutional frameworks; joining a basketball team, for example, offers the opportunity to learn the sport while also conferring the identity of a basketball player. Furthermore, these activities offer social opportunities for adolescents to gain in-group affiliation (eg, being a member of the junior varsity basketball team). Social media may extend this process by providing a space to publicly display and reinforce these identities [2,31].

The next topic being adolescent behaviors, specifically displayed on social media, including self-presentation, identity exploration, and sharing of personal achievements, reflects adolescents' developmental imperative to construct a coherent self-concept and assert autonomy [11]. Guided by the psychosocial task of identity formation—a hallmark of adolescence [32]—teens use platforms to experiment with roles, express authenticity, and document milestones. For instance, self-presentation through curated selfies and photos, or candid mental health disclosures, allows adolescents to negotiate how they are perceived by others, while identity-related posts enable them to explore and affirm marginalized aspects of their identity [33,34]. Furthermore, these activities scaffold identity integration, bolster self-esteem, and enhance emotional regulation, processes that underpin socioemotional well-being [22].

Finally, in the digital age, social media provides a key avenue for maintaining peer relationships. Adolescents use social media not only to interact with friends but also to update their social status, share group activities, and express romantic interests [23]. For example, posting pictures or updates about time spent with friends or significant others allows adolescents to communicate and reinforce social bonds. Posting relational content online is crucial for adolescents' socioemotional well-being [23]. Adolescence is a time for learning social norms and building identity, and sharing experiences with peers, family, or romantic partners on social media reinforces a sense of belonging. By publicly validating their social identities, adolescents strengthen their connections to important social

XSL•FO
RenderX

groups [35]. These interactions also provide feedback and validation, which are essential for boosting self-esteem and feelings of belonging [36]. Moreover, posting relational content helps adolescents navigate social norms and satisfy their social needs. These fulfilled needs promote overall well-being by fostering secure attachments, emotional resilience, and a positive sense of self.

Building on the developmental foundations with the existing research on adolescents' social media use, we introduce an adolescent-centered, cyberethnographically informed codebook that objectively captures adolescents' SGC across 4 major platforms. This study was created to observe adolescent social media and their related well-being, with our ultimate goal being to create reliable, replicable categories to be used in future analyses about how posting specific content on social media influences adolescent well-being. This work will, in turn, allow parents or guardians and potentially policymakers to make more informed decisions about youth in relation to social media and then ultimately how these decisions will affect the well-being of the adolescents.

Methods

Creating the Codebook

To categorize the themes in adolescents' online posts, the research team created a detailed codebook aligned with Bronfenbrenner's ecological systems model (Figure 1) [16]. Our codebook was designed to be applicable for posts made across 4 platforms, including Instagram, Facebook, X, and TikTok.

Figure 1. Social ecological model.

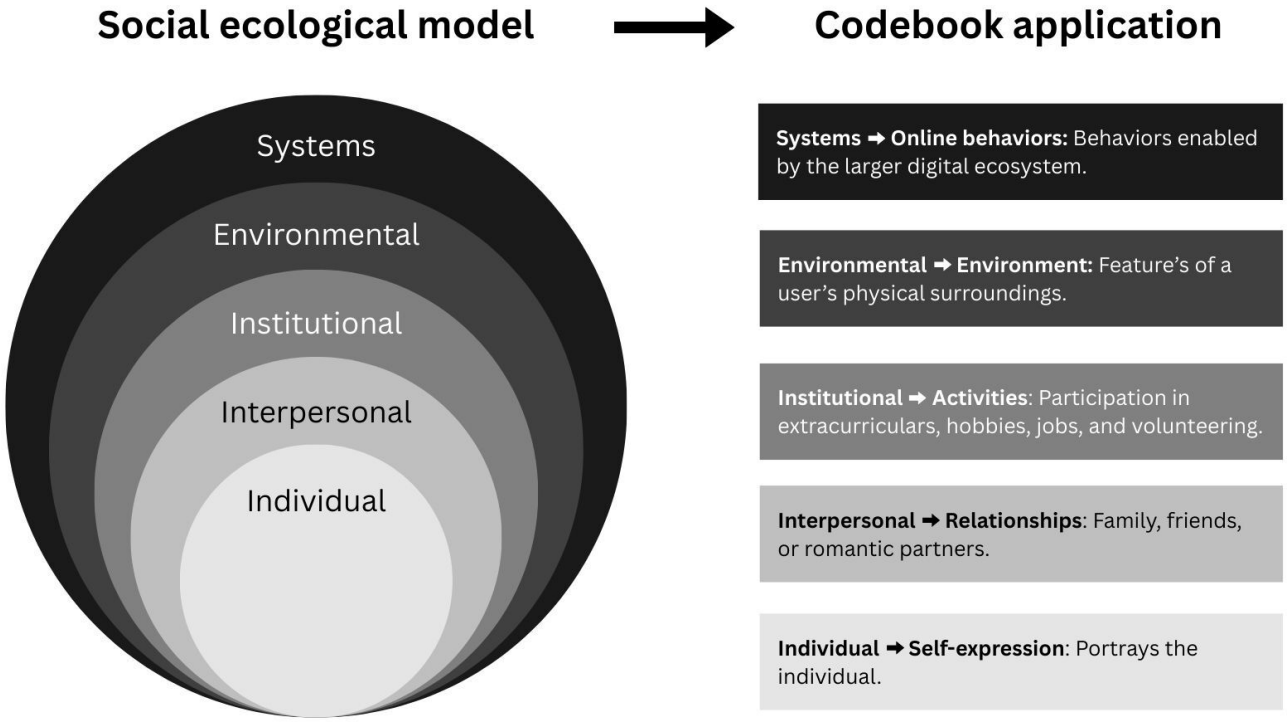

The initial version of this codebook was informed by the lab's prior research on adolescent posting behaviors. Between 2019 and 2022, the lab conducted a study that directly observed 145 youth over 2 years through social media monitoring and dense cyberethnographic field notes [37]. These field notes were used to identify common patterns and themes across participants, informing the development of the codebook for the present study. The team identified 5 overarching themes in adolescents' social media posts that reflect developmental processes unfolding across each ecological level (Figure 1), including self-expression, relationships, activities, environment, and online behavior. Through these 5 categories, the team drew from template analysis and framework analysis to create the concepts for the codes using the data that were previously collected by this study team [38]. Table 1 presents the complete codebook of themes, including subcodes and examples. Multimedia Appendix 1 includes sample posts and the applicable codes for each post.

XSL•FO
RenderX

Table . Codebook table.

Category and code	Definitions	Examples
Self-expression		
Selfie	Photograph or video of the participant, taken by the participant or someone else in the photo, while holding their phone.	• Participant takes a front-camera selfie. • Participant takes a mirror selfie with the back camera. • Friend takes a selfie with the participant. • Participant lip-syncs in a front-camera Tik-Tok.
Portrait	Posed photo of the participant by themselves, taken by someone else, not engaged in an activity. Does not need to be formal or professional.	• Participant's formal senior photo. • Participant poses in a solo birthday photo. • Participant stands in front of a scenic view.
Self-promotion	Promoting personal work or ventures (eg, business, organization, event, social media) with an implied call to action (eg, to buy something, donate, visit a website, be recruited).	• "Come to our school's fundraiser concert Friday night." • "Donate to our Key Club fundraiser!" • Participant's sports highlight reel.
Positive mental health	Positively discussing mental well-being (eg, happiness, emotional growth, gaining skills).	• Photo of participant's artwork with caption: "So happy with my progress." • "Protecting my peace."
Mental health challenges	Discussing mental health difficulties (eg, sadness, suicidal thoughts, self-harm, worsening depression).	• "My parents hate me. They don't understand." • "I feel so alone." • "Life sucks"
Politics	Engaging in political activism or discourse, social justice, and political critiques.	• "Make sure you go vote!" • Meme making fun of the president • Infographic about a current political issue
Religion	Following the beliefs and rituals of a certain faith (eg, Attending church, participating in religious ceremonies, expressing devotion to God).	• Photo of participant attending a religious service. • "Be strong and courageous, Joshua 1:9"
LGBTQ+[a]	Poster directly mentions something about their gender identity or sexuality or experiences with LGBTQ+ topics.	• "Non-binary and proud" • Photos of participant at a Pride Parade with caption: "Happy Pride to me!"
Relationships		
Family	Parents, siblings, and extended family.	• Photo of participant with siblings on vacation.
Peers	Peers, friends, classmates (ie, anybody that is not the participant or a family member).	• Selfie of participant and friends getting food together.
Romantic relationships	Content referencing any stage or aspect of a romantic relationship, including the beginning (eg, starting to date), challenges or conflict (eg, breakups or relationship difficulties), positive experiences, expressions of being single, or general references to romantic involvement.	• Photo of participant holding hands with someone, caption: "6 months with you <3." • Selfie of participant alone with caption: "Being single >>>." • TikTok of participant using trending audio about relationship red flags.
Activities		
Organized group	Participant engaged in an activity that is structured and pre-planned by a school or organization (eg, sports, theater, youth group), done in a group setting (eg, team, club), and is ostensibly supervised or directed by an adult mentor (eg, coach, instructor, teacher).	• Photo of participant playing soccer with their club team. • Group photo of participant with theater cast. • Selfie of participant with school marching band. • "Our debate team won first place at the state competition!"

XSL•FO
RenderX

Category and code	Definitions	Examples
Nonorganized group	Participant engaged in an individual activity performed within a group setting without formal organization, not including general socialization (hanging out with friends, getting dinner, etc).	• Photo of participant fishing with a friend. • Photo of participant's artwork with caption: "I'm so proud of my progress!" • Video of participant and friends lifting weights.
Self-directed	Participant engaged in leisure activity or interest pursued independently. Includes expressing and engaging with personal interests (eg, online fandoms, discourse about sports, music, movies, etc).	• Photo of participant fishing alone. • Photo of participant crocheting on a picnic blanket. • Screenshot of Taylor Swift album with caption: "I'm the biggest Swiftie."
Event	Attending an event (eg, school dance, concerts, sporting events). For events that have performances or competitions, this code should only be considered if the participant is a spectator (not an active participant).	• Photo of participant with date at school homecoming dance. • Photo of concert with caption: "Front row at the Taylor Swift concert in Chicago!" • Photo of participant and family at Brewers baseball game.
Job and/or volunteering	Participants talk about or show themselves being employed or volunteering (eg, animal shelter and restaurant). Does NOT include household chores (laundry, family farm work etc).	• Photo of participant in fast food uniform with caption: "Closing shift." • Photo of participant walking dogs at animal shelter.
Holiday	Observing or celebrating a particular holiday.	• "Happy Diwali to all my Desi girls." • Photos from a Christmas parade.
Birthday	Birthday celebrations or happy birthday wishes.	• Photos from participants's 16th birthday party. • Photo of birthday cake with candles lit.
Vacation and/or travel	Traveling, going on vacation, taking a trip, or visiting one's cabin. Clear indication that the participant left their town of residence and visited a separate location (eg, captions may use the words "vacation," "trip," "break," "travel," etc).	• Photo of participant on a beach with caption: "Spring Break 2025." • Video montage of cabin trip with outdoor scenery and campfire.
Environment		
School	Content explicitly created at school or contains topics related to academics or schooling.	• Photo of participant and friends in school hallway. • Photo of open laptop with caption: "Late night grind."
Nature	Outdoor landscapes and plants. No human faces in the photo. Nature is the prominent or sole feature in the post.	• Photo of mountain range taken during hike. • Photo of sunset over a lake.
Animals	Posts that feature real animals, including pets, wild animals, zoo animals, or animals encountered outdoors.	• Photo of participant's dog lying on bed. • Selfie with a goat at a petting zoo.
Food and/or beverage	Food or beverage is visible in the photo (eg, plates of food, cups, water bottles, etc).	• Selfie of participant holding a Starbucks cup. • Photo of Chick-fil-A bag and drink in car seat.
Online behavior		

XSL•FO
RenderX

Category and code	Definitions	Examples
Seeking online engagement	Seeking other users to engage with a post or a profile (eg, comments, likes, DMs) in a way that is directly or indirectly related to the participant.	• "Comment 3 things you know about me." • "Repost if you think animal cruelty is wrong." • "What songs should I add to my playlist?"
Media	Content from mass media (TV, movies, video games, podcasts, music, professional sports, etc) and internet content (eg, memes, clips, etc). Content must be directly taken from source (not videos of a screen) but can be manipulated (eg, fancams, edits). Mere references to pop culture are not coded.	• Screen-recorded TikTok reposting BTS fancam. • Meme using SpongeBob still with humorous caption. • Spotify screenshot of currently playing song.
Trend	Participant is engaging in a viral trend (eg, dances, lip-syncing, pranks, using CapCut templates, etc).	• The participant performs a trending dance to popular audio. • Video of participant doing the "choose your character" trend.

[a]LGBTQ+" lesbian, gay, bisexual, transgender, and queer and or questioning.

Self-Expression

This theme includes content focused on portraying the individual adolescent user, such as photos of themselves, posts promoting personal work or ventures, discussions of mental well-being, and expressions of personal identity. This theme aligns with the individual level of Bronfenbrenner's nested ecological model, as it reflects the adolescent user's personal characteristics, experiences, and self-concept.

Relationships

This theme includes posts featuring or discussing family, friends, or romantic partners, such as selfies with a friend, group photos, or family portraits. This theme reflects the "interpersonal" level of the nested ecological model, representing immediate social networks and reciprocal relationships that directly shape an individual's development.

Activities

This theme includes content about extracurriculars, jobs, volunteering, and hobbies. This theme aligns largely with the "institutional" level of the ecological model, as these activities occur in structured settings, such as schools, workplaces, and community organizations. These institutional settings provide opportunities, resources, and social contexts that shape individual development.

Environment

This theme captures aspects of a person's physical surroundings that may reflect or influence their physical, mental, or socioemotional well-being. This includes posts about individual items, such as pets or food, as well as broader settings, like schools or outdoor spaces, that may be linked to well-being. This theme corresponds to the "environmental" level of the ecological model, encompassing immediate, community, and cultural contexts that shape an individual's experiences and development.

Online Behaviors

This theme captures actions performed specifically on the internet, including sharing mass media content and engaging with online trends. This category also includes online behaviors intended to increase engagement with a post or profile. In our adaptation of the ecological systems model, we conceptualize posting behavior as occurring within an overarching digital ecosystem; therefore, behaviors enabled by digital infrastructure and specific to online contexts align with the largest "systems" level of this model.

Codebook Refinement and Back-Coding

These 5 major themes formed the preliminary codebook, which was first implemented in April 2023. Over the first year of coding, the codebook was applied to all collected data and continuously refined to address emerging patterns and discrepancies, ensuring that codes both accurately reflected the data and had strong interrater reliability. This refinement process was done through frequent team meetings to discuss and compare coding between posts reviewed by multiple coders, all of which active coders were required to attend, and if they could not make it, they would be required to watch the recording during their next shift. Adding a new code to the codebook consisted of having a new theme brought up during individual coding that didn't fit in with any existing ones. Then the team would keep a list of all emergent themes, and if it was brought up more than one time by more than one person, the team would then need to reach a consensus to agree to add in that new code to the codebook. All updates were carefully documented in a shared document between the team to maintain a timeline of revisions for future reference. The codebook was finalized once no new emergent codes were identified by any coder during the coding process, indicating theoretical saturation. After multiple iterations, the codebook was finalized in June 2024. Since the codebook has been finalized, no additional codes have been created; however, if a subject does come up that appears to not fit into a category, the team will meet to discuss what the best codes are that do fit that. All subsequently collected data were coded using this final version.

To ensure consistency across the dataset, all data coded before the codebook was finalized (April 2023-June 2024) was systematically back-coded according to the final codebook. Using the documentation of all codebook changes, the study

XSL•FO
RenderX

team created a detailed guide for coders that included a comprehensive list of changes and instructions on how to address these changes while back-coding. The back-coding process, which ran from July to December 2024, took place alongside ongoing data collection. See Figure 2 for a timeline of codebook development and application.

Figure 2. Codebook timeline.



Data Collection Tools

Research Electronic Data Capture (REDCap; Vanderbilt University) is our primary data collection tool, chosen for being accessible, flexible, and secure. REDCap is a free, web-based, open-source platform that is Health Insurance Portability and Accountability Act (HIPAA)–compliant and hosted on our institution's servers. REDCap allows research teams to create multiple customizable forms for data collection and coding tracking and provides storage for the repository of coding data. Although we do recognize that there are other tools that could be used for this project, REDCap has a special team at the University of Wisconsin–Madison that is ready and able to help with issues or questions on programming. This has been incredibly helpful for project creation and execution. This team had previously used Qualtrics (Qualtrics International Inc) for a different project and found that the data collection was fairly easy; however, when it came to data exportation and analysis, this platform was not as user-friendly as REDCap.

To view participants' social media accounts, we created multiple study-specific profiles on each platform (Facebook, Instagram, TikTok, and X). Study accounts use fictitious names and photos designed to resemble typical user profiles and are periodically updated with posts. This approach reduces the likelihood of being flagged as bots or inadvertently revealing participants' involvement in the study. Multiple profiles were also created in advance to serve as backups in case of accessibility issues. With parental consent and minor assent, we request to follow participants' accounts on each platform using the study profiles during enrollment and conduct six-month follow-ups to verify that the information remains current. All research team members use the same accounts for data collection.

Unit of Analysis and Contextual Coding Procedures

To ensure consistency in coding, our team first defined the unit of analysis—the portion of a social media post that would be coded. In this study, the unit of analysis is the entire self-generated post, including its media (eg, photos and videos), audio, and captions. Because our focus is on content created by participants, comments are not thematically coded. However, on platforms where engagement is visible, coders record the number of comments the participant writes, reacts to, or responds to, providing additional insight into how participants interact with their own posts. Each platform varied in the information that is to be collected. From Instagram having regular posts to reels to highlights to stories, we created a variety of forms to encapsulate this information in a meaningful way for this study. Instructions for the coding procedures can be found in Multimedia Appendix 2.

Our team uses a holistic approach, using participants' shared social media accounts and other publicly available data to properly code each post. When needed, coders are instructed to gather contextual information from any of the participant's social media accounts followed by the study team. Coders may also reference other public information, such as usernames of tagged individuals, to support coding determinations. For example, if a post features the participant alongside another similarly aged individual, it might be unclear whether this person is a friend or a family member. By reviewing the participant's social media history, a coder may observe that this individual only appears in family-related posts. Additionally, a publicly visible username may reveal a shared last name, suggesting a familial relationship and allowing for more precise coding.

Coders are encouraged to explore social media platforms and the broader internet to contextualize posts to identify online trends, interpret online vernacular, or understand niche or unknown references. For instance, if a participant mentioned going to an "LARP event," a quick Google search could reveal that "LARP" stands for "live-action role play." Coders are also instructed to use search tools and platform-specific features—for example, TikTok's shared public audio function—to identify trends or clarify content. However, coders are strictly prohibited from using personally identifiable information in any online searches and are instructed not to collect or record data from accounts that have not explicitly consented to participate in the study.

XSL•FO
RenderX

Study Design

This project and team are a part of a larger longitudinal observational study that is conducted over the course of 24 months, using a mixed methods study design consisting of both qualitative and quantitative data collection [39]. There are 3 principal investigators (PIs) who man individual teams varying in size from 5 to 20 research staff at any time, which are made up of undergraduate students, graduate students, postdoctoral students, and professionals.

Well-Being Data

The larger longitudinal observation study is collecting well-being data at different time points throughout the study. Some of the well-being metrics being collected include Media and Technology Usage and Attitudes Scale (MTUA), General Anxiety Disorder 7-item (GAD-7), Emotional Responses to Social Media Experiences Measure (ERSMEM), 8-item Patient Health Questionnaire (PHQ-8), and more [40-43]. These metrics are collected at 3 time points over the 24 months and will be used to connect to the participants' social media data at different time points throughout the study.

Participants

The participants in this study are adolescents aged 13 - 15 years from the state of Wisconsin who have at least one active social media profile on Facebook, X, Instagram, and/or TikTok. These participants were sourced from local events in Wisconsin such as the state fair, Wisconsin Science Festival, and more, where team members from the study talked to potential participants and their parent or guardian about the study and enrollment. Our team used the census data from 2020 for the state of Wisconsin for our targeted enrollment demographics. Each participant, when joining the study, shares their social media platforms and is followed by the study-specific profiles mentioned above. This team then goes into each participant's profile every month to code and collect the data.

Coding Team

Our team began coding study data in April 2023, with the initial coding conducted by research staff with backgrounds in adolescent medicine, public health, and psychology. Beginning in September 2023, undergraduate student research assistants were hired to assist with data collection. The size of the coding team has fluctuated over the course of the study based on workload and coder availability. As of the time of writing, 28 coders have contributed to data collection, each participating for at least 4 months. The standing study team for this project is built of PI (ES), 5 research staff (KB, LB, TF, KK, and CMC), and alternating undergraduate student research assistants.

Coder Training

First, our study team developed a comprehensive coding guide to support coders in both learning the coding process and referencing updated procedures throughout the project. The guide was designed to be highly visual and easy to navigate and included screenshots of social media and data collection platforms, step-by-step coding instructions, and troubleshooting strategies. Any updates to the coding process are immediately reflected in the guide, which is used as the definitive reference to ensure consistency across coders and throughout the data collection process.

Coders begin training by meeting one-on-one with the lead study team member who reviews the coding process and explains each code in detail. Coders then practice applying this codebook to a sample social media post with the assistance of their supervisor to assess their understanding. Using the coding guide and codebook, coders then complete a training exercise by coding 24 preselected social media posts (6 from each platform). Upon completion, a REDCap report identifies agreements and discrepancies between their coding and the answer key, created by their supervisor. Each coder then meets with their supervisor to review results post-by-post, discussing agreements, discrepancies, and refining their understanding of the codebook.

Coders who achieve at least 80% agreement with the master coder begin official data collection under close supervision to address any lingering questions. While most coders reach this threshold after the first round, those who do not repeat the process with a new set of posts. As of writing, no coders have required more than two rounds, though a third set of posts was prepared as a contingency.

Coding

Coding is conducted monthly for each participant, with each 30-day interval following enrollment representing an observation period. Participants are followed for 24 observation periods as part of a larger longitudinal, multimethod study [39]. Content is coded from participants' Facebook, Instagram, TikTok, and X accounts (Figure 3). Something to note is that each post can be assigned more than one thematic code, as many posts are complex and include a variety of subjects.

XSL•FO
RenderX

Figure 3. Coding flowchart by platform.

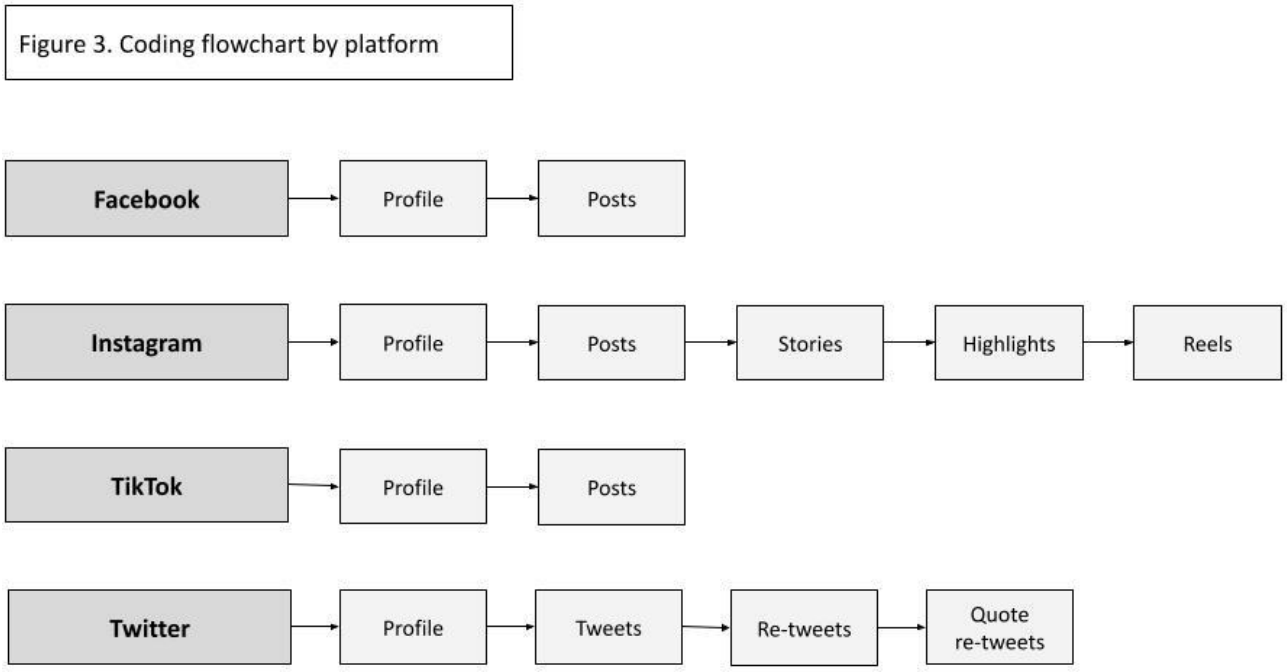

Assigning Coding

We assign coding using REDCap's Data Resolution Workflow, a feature typically intended for formally flagging data issues (eg, missing, incorrect, or invalid values). This workflow enables users to open a query attached to a specific data field, assign it to another user, and track its resolution. Although not designed for this purpose, our team adapted the feature to assign coding tasks to individual team members.

Coding is organized by "participant-months," where each unit represents one month of coding for a single participant. At enrollment, each participant's full coding schedule was programmed into REDCap for the duration of their study participation. When a participant-month became due, it appeared on REDCap's calendar.

To make assignments, the data manager reviews the calendar daily to identify participant-months scheduled for coding. For each assignment, the manager opened a query in a form on the participant's profile, specifying the participant-month to be completed. Coders accessed their assignments through the data resolution dashboard by filtering queries by their username.

Quality Control

Given the volume of data, number of coders, and duration of the study, it was essential to establish rigorous procedures to ensure data integrity. To support accuracy and consistency, our team implemented annual interrater reliability checks, double-coding procedures, and weekly alignment meetings.

Annual Interrater Reliability Checks

Each year, we randomly selected 6 participant posts from each social media platform (Facebook, Instagram, TikTok, and X), yielding a total of 24 posts. All active coders independently coded each post, and the results were compiled and assessed using the Gwet First-Order agreement coefficient (AC1), which is well-suited for evaluating interrater reliability with binary outcomes and multiple raters [44]. A passing threshold was set at >80%. If this threshold was not met, the team completed a second round of interrater coding and recalculated the AC1 score. These checks were typically conducted each spring and repeated annually for the duration of the study.

Weekly Team Meetings

Our team holds weekly meetings to keep coders aligned on the codebook and informed about any protocol or administrative updates. Each week, one coder shares a social media post with the team, typically one that is challenging to code or raises questions about applying specific codes. The coder presents the post along with their coding decisions, interpretation, and any questions about applying the codebook to unique features of the content. Team members then share their own coding determinations, and the group discusses discrepancies until reaching consensus. This practice helps the coding team refine their understanding of how codes are applied and highlights areas where additional clarification may be needed in the coding guide.

Concerning Content Protocol

Because 2 of our PIs are mandated reporters, our study team has both a legal and ethical obligation to comply with mandated reporting requirements [45]. To support this, we developed a standard operating procedure for how to identify and report content on participants' social media profiles that might meet reporting criteria. Our standard operating procedure ensures that any reportable content is identified and escalated to the investigators in a timely and responsible manner.

For the purposes of this study, concerning content is defined as posts that referenced suicidality or self-harm, sexual or physical abuse, or threats of harm or violence toward others. Such content can appear in the participant's own words or in reposted material that suggests the participant may be at personal risk of harm. A made-up example of this would be a participant posting "I'm

XSL•FO
RenderX

going to school with a gun tomorrow," which would require prompt action from our team. Notably, concerning content does not include all illegal or risky behaviors, which do not always meet the threshold for mandated reporting [46].

Ethical Considerations

Institutional Review Board approval was obtained from Application Review for Research Oversight at Wisconsin (ARROW) under Institutional Review Board number 2022‐1280. Informed consent is given through a survey through REDCap that each participant's parent or guardian must sign before being enrolled in the study. All participant and related data are stored in the secure University of Wisconsin-Madison REDCap and are deidentified when exported for analysis. Participants can receive a total of up to US $300 for completing all study-related activities.

Results

Overview

This study was funded in 2022 and began data collection in early 2023, with the codebook still in development, with the study concluding in 2027; however, there is a possibility of continuation. Figure 2 outlines the timeline from beginning to the expected end of the study. There are 369 participants at the time of this writing, with 336 actively enrolled and 33 withdrawn.

Coding Progress

Figure 4 illustrates the trend in participant-months requiring coding. The workload peaks around the midpoint of the study (mid-2025) and then declines as participants reach the end of the 2-year enrollment period. Month-to-month variation reflects the staggered and fluctuating nature of participant enrollment.

Figure 4. Frequency of participant-months per month.

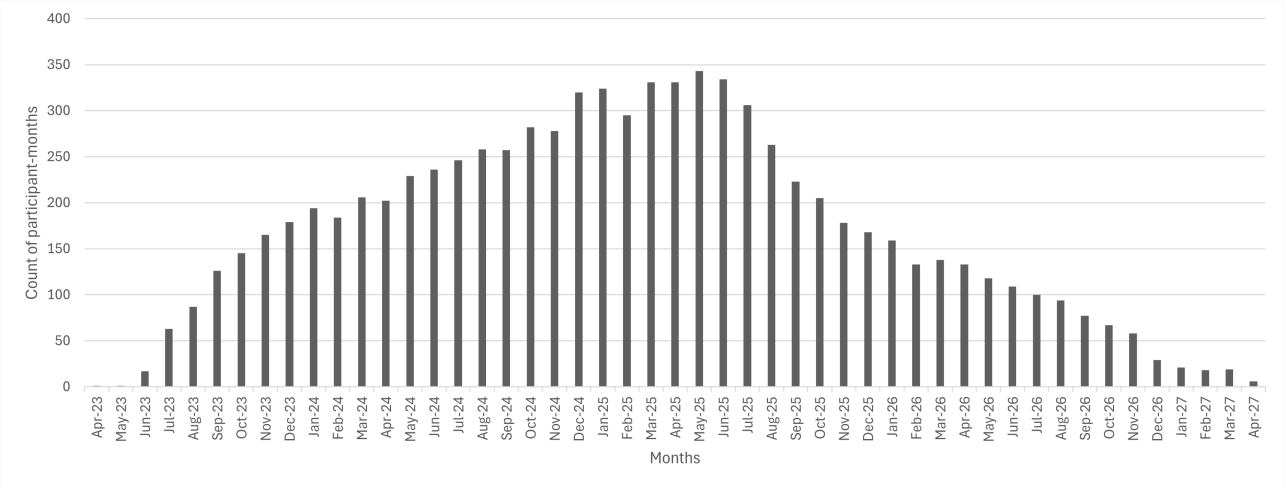

Quality Control Measures

Annual interrater reliability assessments produce 5 AC1 scores, one for each platform (Facebook, Instagram, TikTok, and X) and one overall score. Year 1 results, shown in Table 2, indicate that Instagram had the highest agreement (0.89) and X the lowest (0.87), with all scores exceeding the 0.80 acceptability threshold. In Year 2 to date, X shows the highest agreement (0.96), while Facebook has the lowest (0.95).

Table . Interrater reliability agreement scores.

Years	Interrater reliability agreement scores				
	Facebook	Instagram	TikTok	Twitter	Overall
Year 1	0.89	0.89	0.88	0.87	0.88
Year 2	0.95	0.96	0.96	0.96	0.96

Discussion

Strengths and Unique Contributions

Using Human Coders

Using research staff and hiring student research assistants for social media data collection has been a cornerstone of our methodology. Human coders add essential cultural and contextual understanding to content coding, particularly when interpreting the highly complex and dynamic nature of adolescent social media content. Human coders can be better equipped than automated tools (such as artificial intelligence [AI] systems or data scraping algorithms) in understanding context, recognizing tone, subtext, peer norms, and internet colloquialisms. These nuances are central to online communication and often carry significant meaning; even the most advanced automated tools can often remain unreliable in detecting them accurately. However, there is a time and a place where AI may have a value that is considered to be more useful than a human coder, such as for large-scale data collection that would not be feasible for a human coder to complete.

XSL•FO
RenderX

Human coders, our lab's student research assistants, carry out much of the coding and play a central role in data collection. Aged 18-22 years, they are close in age to our adolescent study participants, aged 13-17 years. Both groups belong to the same generation of "digital natives," young people who have grown up during the age of digital technology and are intimately familiar with social media. This generational proximity enhances the coders' ability to recognize platform-specific norms, contextual nuances, and forms of expression that might otherwise be overlooked, supporting more accurate and culturally informed data collection and interpretation.

Human coders are also highly adaptive. Given the nearly limitless variety of social media posts adolescents create, the tools used to collect and interpret this data must be equally flexible. Human coders can respond to unexpected or ambiguous content in ways that preprogrammed systems cannot, especially in a research context where interpretive accuracy is critical. This approach also enables real-time, context-sensitive interpretation. Unlike retrospective analysis, which may fail to account for fleeting internet trends or evolving slang, human coders can engage with content as it emerges, an essential feature when working in rapidly changing digital environments.

Unlike automated data scraping (ie, using machine learning and other big data approaches), which is often limited by technical barriers and platform-specific API restrictions, human coders can access and analyze content across platforms without relying on unstable or restricted infrastructure. This is particularly advantageous in longitudinal studies where interruptions to data access could compromise the dataset. Human coders are also able to synthesize multimodal content, including video, audio, text, and comments, which often carry interdependent and context-specific meaning. While automated tools may struggle to parse and integrate these multiple layers of information, human coders can evaluate them holistically, adding coherence and nuance to coding decisions.

Human coding also ensures greater transparency and traceability. Unlike the opaque decision-making processes of AI-driven models, human coding decisions can be documented, discussed, revisited, and replicated. Weekly team meetings allow coders to refine their approach, address inconsistencies, and ensure the integrity of the dataset through shared understanding and collaborative troubleshooting. Ethically, human coders provide a more secure and accountable approach to data handling. Our method does not rely on third-party tools, reducing the risk of privacy breaches. Additionally, coders can manually flag content of concern in a thoughtful and timely manner, an important safeguard in studies that involve mandated reporters.

Cyberethnography

From a methodological standpoint, the foundations of our codebook in cyberethnography bring unique advantages to social media research. Foundationally, ethnography in general is meant to provide a picture of daily life, and for adolescents, daily life occurs online. Given the concordance of social media displays and adolescent offline behavior [47,48], the online space represents an important natural environment for observation of adolescent social behavior. Traditional measures of social media experiences, such as survey measures, ecological momentary assessment, and interviews, rely on self-report. The aforementioned methods have the strength of the inclusion of youth voices and their perspectives on social media, particularly in the presence of rigorous data collection instruments. The method described in this manuscript, cyberethnography, allows the research team to objectively observe adolescents' social media activity to further analyze what each participant is trying to portray through their SGC. Whereas, if the team were to rely upon just interviews, the information could fall prey to bias from individuals reporting on their own social media activity rather than benefiting from objective observation through cyberethnography.

Large Number of Posts and Longitudinal Approach

With a considerable sample size of social media posts such as ours, we are able to add additional dimensions to traditional cyberethnographic analysis. While sample size is an area of much debate in qualitative research [49], arguably with a group of hundreds of participants and thousands of posts, theoretical saturation is easily reached. We are then able to scale up our codebook to create quantitative variables that can be used as longitudinal predictors of outcomes of interest, for example, socioemotional well-being indicators. With participants enrolled for 2 years with multiple data points, our research is responsive to calls for longitudinal research to determine bidirectional relationships between social media experiences and socioemotional well-being [50].

Informed Data Collection

Another unique advantage of this methodology is its emphasis on ethical data collection that is completed internally within the research team. We very carefully collect informed consent and assent for all participants. This in itself is unique for the collection of social media data at the level of detail we use and has been met with both hesitation and trust. We note that social media companies collect much more data than this project, with very little transparency. With long, complex terms of service documentation that users must accept to use social media platforms, the "consent" that users provide for this data collection is questionably "informed." We believe that an informed consent model for social media data collection ultimately fosters trust in researchers through transparency about exactly what is collected and why [51]. That being said, the use of human researchers to observe social media data may cause trepidation in adolescents and their parents. Future research could explore further with adolescents and their parents such concerns and possible mitigation strategies.

In addition, human review raises ethical concerns as to sensitive content; if a coder observes content depicting illegal behavior, self-harm, or abuse, what is the duty of care to address this? In our study, we are guided by the principles of mandated reporting, the same reporting structure that guides the physician members of our investigative team in clinical settings. This approach is particularly useful in a cohort study in which we have regular interactions with participants and prioritize safety and confidentiality [52].

XSL•FO
RenderX

REDCap

REDCap was chosen as the data repository for this project because it met HIPAA compliance and supported our diverse data collection needs. REDCap allowed for all team members to be able to be added to the specific projects and work simultaneously on data collection. The use of REDCap also enabled the team to import the participants from a spreadsheet. The functionality of the calendar and scheduling then became useful for creating the 24 months of observation periods. The use of the reports and exports also allowed the team to perform quality control checks and data analysis for project tracking and manuscript writing. All of these processes and actions could then be followed through the use of the dashboards via REDCap, which allowed for the data and tracking to be observed. REDCap is also fully supported by our academic institution, which has increased efficiency in troubleshooting any issues that have arisen. With an on-campus data manager, our research team could rely upon timely responses from the REDCap support team at our institution.

Lessons Learned

Applying this method included important considerations, including the team, participant social media use, and policies impacting social media use.

Research Team Training and Fluctuations

The research team changed over the course of the codebook's development, training, and application. While team turnover may lead to more generalizable codes, the changes also created rationale and interpretation gaps. The team maintained interpretable codes by writing clear definitions, using a single document to house definitions and decisions, hosting regular staff meetings to re-emphasize code applications, and adjusting definitions or adding codes as the team collected more data. The high interrater reliability scores highlight the strength of the strategy. Future applications of this method should ensure written definitions of codes in one place, flexibility in code definitions, and regular communication with the coding team.

Researcher Reflexivity

Any study with a qualitative component carries concerns for "blind spots" in interpretation based on the lived experiences and positionality of the individuals conducting the research and designing the codebook. Our team has attempted to mitigate potential bias by active inclusion of individuals of diverse age, race and/or ethnicity, sexuality, gender identity, and discipline. Furthermore, frequent meetings during both refinement and implementation of the codebook allow for the contribution of new voices as the study goes on.

Account Access Challenges

One of the challenges that was fairly constant over the course of the study was with access to participant social media accounts. Participants adjusted accounts throughout the study, such as changing usernames, unfollowing our research accounts, adding new accounts, deleting current accounts, and having accounts suspended for various reasons. Troubleshooting changes to account access was a continual issue with the risk of nonaccess resulting in missing data for that time point. To address this, a point person was assigned to handle all issues with username access as they arose, which allowed research staff to continue to code available profiles and maintain coding progress. All accounts with access issues were assigned and coded once the issue was resolved.

Another issue the research team faced was the terms and conditions policies related to each social media platform. Having a shared research team profile for each social media platform is an atypical use case but is necessary for maximal access to participant profiles in the absence of collaboration with platforms using their APIs. Having multiple researchers accessing the same social media account and scrolling back through posts for each participant every month carried risk for research group accounts to be flagged as spam or bot accounts. After being flagged, our research team is not able to follow or add any new participants to that account. Once the research account is flagged as spam, there is little possibility for restoration, a common experience for users of these platforms. Over the course of the study, 3 of our research accounts were flagged as spam, and we addressed this by creating new profiles to follow new research participants. Flexibility and quick troubleshooting were needed to make sure all participant accounts were accessible throughout the study. To maintain security and accessibility of multiple coders accessing accounts at the same time, using the University VPN and deleting saved cache files or saved passwords on computers, as well as using different browser platforms to access different profile accounts. By adopting these policies, the team was able to maintain the four profiles without other flagging issues occurring.

Federal Platform Regulations

Accessing certain social media platforms among evolving federal regulations—particularly regarding TikTok—was another unforeseen issue the research team had to troubleshoot. As a public institution, accessing special permissions for research was sought for our study to have access to TikTok on university-owned computers. Further, the passing of HR 7521, the "Protecting Americans from Foreign Adversary Controlled Applications Act" (2024), would delay or terminate our data collection on this platform. The bill prohibits the distribution, maintenance, and updating of certain foreign-owned applications. As a foreign-owned company, TikTok is at risk of being removed from United States app stores and web hosting services. Enforcement of the bill would significantly harm our understanding of self-generated content in adolescents, especially since an increasing percentage of enrolled participants are users with multiple accounts on TikTok. Their active use demonstrates the popularity and importance of this platform to our adolescent participants. Ultimately, the ban only went into effect for less than one day before being temporarily delayed by the Trump administration, interrupting data collection for only a short period [53]. As of writing, the ban deadline has been extended 3 times via executive order [54]. The future of TikTok remains uncertain, and enforcement of the ban threatens to disrupt future data collection efforts.

Platform Variability and Change

The team needed to adjust for the distinct functionalities within social media platforms. Adolescents share self-generated content

XSL•FO
RenderX

differently between social media platforms. For example, Facebook and X tend to be associated with text-based content, whereas Instagram and TikTok are more visual (ie, photos and short-form videos). Further, users can share content on Instagram in more spaces (eg, Stories, Highlights, the "grid") compared to other social media platforms. Previous research explores how different functionalities may hint at different user motivations [55,56]. Therefore, this codebook applied to a wide range of SGCs. To navigate platform differences, the team used tailored REDCap forms by platform to ensure all possible self-generated content was captured.

We also needed to adapt to changes within platforms. For example, in February 2023, TikTok added "Reposts," a functionality that allows content created by one user to be reshared by a second user. Unlike Facebook and Instagram, TikTok reposts only show the likes and comments from the original post and do not include timestamps of when the content was reposted. Additionally, participants who reposted content on TikTok were likely to repost a heavy volume within the observation period, challenging the team's ability to effectively code these posts separately. Therefore, the team needed to make 2 separate determinations: first, to what extent are TikTok reposts self-generated content? Second, how should TikTok reposts be incorporated?

To navigate this change, we relied on multiple key strengths, including routine, open communication with team members, the flexibility of the method, and the versatility of REDCap. We first discussed the feasibility and importance of coding reposts with the coding team. Through this discussion, we decided that all reposts are important to capture, but that theming each post individually is not feasible due to the volume of reposts. Therefore, we adapted our method by counting the number of reposts within an observation period and summarizing the content generally. We also added a section in REDCap to mark the last repost shared in the previous observation period to more accurately count the number of reposts in the next observation period.

Other less notable examples of platform changes include Facebook adding Stories, Instagram increasing their cap on the number of photos in a post, and platforms hiding the number of likes and views on participant posts. In all cases, the coding guide was changed, and the coding team was notified. Social media platforms are distinct, and their evolution is inevitable. Researchers who apply this method should prepare to navigate platform differences throughout the course of their project and lean on the expertise and communication of their team to navigate challenges.

Other Limitations

More notable restrictions that researchers may face include API restrictions, which we have encountered in regard to a different project, forcing the team to find alternative ways to collect the needed data. This would also vary across different countries in the world as well, since each country has differing laws on data access and collection.

Importance

The implications of our methodology are many. First, the provision of granular, researcher-observed data regarding self-generated social media content not only densely phenotypes adolescent online behavior but also allows for identification of more precise online experiences that may have positive or negative impacts on a multitude of physical and mental outcomes. The impact of social media on the health and well-being of adolescents is of concern to adolescents, parents, policymakers, and the public [53,54]. Objective identification of risky and healthy online behaviors can aid in future development of interventions, from parental, adolescent, and clinician-based education to detection and amplification at the design level. The latter, however, would require close partnership and transparency between researchers and social media companies, which may need to be enforced at a policy level.

Across other research studies, there have been a variety of different approaches taken to code adolescent social media. One study focused on coding race- and ethnicity-related content (text and images) in Facebook public posts through a hybrid inductive-deductive approach [57]. Which is similar to our study in that we used a hybrid approach but focused mostly on psychosocial and/or well-being–related content across 4 platforms, and we also included private posts as well. Another study focused on health-risk–related content, which also followed a theoretical theory (deductive approach) first followed by an inductive approach, so it was very similar to our approach that we used [58]. As for content coding of profiles versus just posts, one other paper coded users' profiles and then further developed a codebook for sexualized behaviors on Facebook, which draws to our study as well because we not only collected posting data but also coded profiles for each account per platform [59]. Finally, some of our team's previous social media research work includes creating a codebook that was designed for YouTube video content, which ultimately helped us to have an idea of how to navigate creating a codebook for social media data collection [60].

This project offers both methodological and theoretical advances to adolescent social media research. First, it adds to the field's previous evidence using self-reported screen time or broad "active–passive" classifications by objectively analyzing the actual posts adolescents create. Each participant's content (video, text, photo, etc) is captured across multiple platforms (TikTok, Instagram, Facebook, and X) over a 2-year span, providing a granular, longitudinal view of their digital lives that is unattainable with single-platform or cross-sectional designs. Second, the study introduces an adolescent-centered cyberethnographic coding framework. The codebook was applied by rigorously trained coders, ensuring high interrater reliability while eliminating recall bias. This participatory approach yields a faithful record of what adolescents choose to share online. Third, the coding scheme is grounded in Bronfenbrenner's ecological model, capturing posting themes at 5 nested levels, including individual, interpersonal, institutional, environmental, and systems. By aligning observable social media behavior with a well-established developmental theory, the study links online self-generated

XSL•FO
RenderX

content to the broader contextual layers that shape adolescent well-being. Together, these innovations furnish a holistic portrait of adolescents' online expression and lay the groundwork for future research examining how specific posting patterns relate to socioemotional outcomes.

Planned Analysis

Our study team is already in the process of running initial analysis on the descriptives of these data that we are collecting. Beyond a descriptive analysis, we have been working on cross-analyzing the engagement of the participants' peers on their social media posts in relation to the participants' overall well-being from different time points throughout the study. These are just a few examples of how this data will be used; however, overall, we are planning to run many different types of analyses on the well-being data that are being collected and the SGC that is also being coded by this team. We will be able to show how one impacts the other and vice versa, which is information that will hopefully help parents or guardians and policymakers.

Conclusion

This formative research approach that was taken with our lab has allowed for us to have insight and collect data that will give a new perspective on social media and adolescents. This, in turn, will help our research team to further show the connection between social media and socioemotional well-being. Specifically, this team plans to write a descriptive paper on the data that is being collected. We then plan to work with the other teams on the bigger project to connect participants' well-being with the social media data that we have collected. This team works with undergraduate and graduate students, and postdoc and research professionals across the University of Wisconsin–Madison, who will also use the data that we have collected to answer differing research questions. This approach will aid future studies in project creation and execution on studies similar in nature. The foundation that has been laid out by this team should be a guidance to further the research on the effects of social media on adolescents and within other age groups as well. This research strategy could also be extended to conducting research in different ethnically diverse populations across the globe. This research is critical to understanding the well-being of individuals who participate in social media use, especially with the ever-changing nature of social media. This team will continue to publish research outcomes from the data collected from this methodological approach and will carry on studying the well-being of adolescents and their interactions with social media.

Acknowledgments

The authors would like to thank the undergraduate research assistants who assist this team with coding. Additionally, the authors would like to thank Victoria Adkins and the participants who have shared their social media data with our team.

Funding

This protocol is funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (P01HD109850-01) and the University of Wisconsin-Madison Vice Chancellor for Research and Graduate Education Matching Grant.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

KB: conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, visualization, writing-original draft, writing - reviewing and editing.
LB: conceptualization, data curation, formal analysis, investigation, methodology, project administration, supervision, visualization, writing-original draft, writing - reviewing and editing.
TF: conceptualization, writing - original draft, writing - review and editing.
KK: writing - original draft, writing - review and editing.
C Carlson: writing - original draft, writing - review and editing.
MM: writing - review and editing.
C Cascio: writing - review and editing.
ES: conceptualization, funding acquisition, methodology, resources, supervision, writing - original draft, writing - review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample social media posts with applicable codes.
[PNG File, 454 KB - resprot_v15i1e84461_app1.png]

XSL•FO
RenderX

Multimedia Appendix 2
Coding guide.
[PDF File, 14059 KB - resprot_v15i1e84461_app2.pdf]

References

1. Faverio M, Teens SO. Social media and technology. Pew Research. 2024. URL: https://www.pewresearch.org/internet/2024/12/12/teens-social-media-and-technology-2024/ [accessed 2026-03-22]
2. Nesi J, Choukas-Bradley S, Prinstein MJ. Transformation of adolescent peer relations in the social media context: part 1-A theoretical framework and application to dyadic peer relationships. Clin Child Fam Psychol Rev 2018 Sep;21(3):267-294. [doi: 10.1007/s10567-018-0261-x] [Medline: 29627907]
3. Yang Q, van den Bos K, Zhang X, Adams S, Ybarra O. Identity lost and found: self-concept clarity in social network site contexts. Self Identity 2022 May 19;21(4):406-429. [doi: 10.1080/15298868.2021.1940270]
4. Reid Chassiakos YL, Radesky J, Christakis D, Moreno MA, Cross C, COUNCIL ON COMMUNICATIONS AND MEDIA. Children and Adolescents and Digital Media. Pediatrics 2016 Nov;138(5):e20162593. [doi: 10.1542/peds.2016-2593] [Medline: 27940795]
5. Bates A, Hobman T, Bell BT. "Let Me Do What I Please With It... Don't Decide My Identity For Me": LGBTQ+ youth experiences of social media in narrative identity development. J Adolesc Res 2020 Jan;35(1):51-83. [doi: 10.1177/0743558419884700]
6. Sanchez G, Jenkins JH. Social media & subjectivity: adolescent lived experiences with social media in a Southern California middle school. Soc Sci Med 2024 May;348:116839. [doi: 10.1016/j.socscimed.2024.116839] [Medline: 38581816]
7. Harwood TG, Garry T. An overview of content analysis. Mark Rev 2003 Dec 1;3(4):479-498. [doi: 10.1362/146934703771910080]
8. Hammersley M. Ethnography: problems and prospects. Ethnogr Educ 2006 Mar;1(1):3-14. [doi: 10.1080/17457820500512697]
9. Keeley-Browne E. Cyber-ethnography: the emerging research approach for 21st century research investigation. In: Kurubacak G, Yuzer TV, editors. Handbook of Research on Transformative Online Education and Liberation: Models for Social Equality: IGI Global Scientific Publishing; 2011:330-338. [doi: 10.4018/978-1-60960-046-4.ch017]
10. Kozinets RV. The International Encyclopedia of Digital Communication and Society: John Wiley and Sons; 2015:1-8. [doi: 10.1002/9781118767771]
11. Erikson EH. Identity: Youth and Crisis: W W Norton and Company; 1968, Vol. 68. URL: https://psycnet.apa.org/record/1968-35041-000 [accessed 2026-03-22]
12. McLean KC, Syed M. Personal, master, and alternative narratives: an integrative framework for understanding identity development in context. Hum Dev 2016 Jul 14;58(6):318-349. [doi: 10.1159/000445817]
13. Berndt TJ. Friendship quality and social development. Curr Dir Psychol Sci 2002 Feb;11(1):7-10. [doi: 10.1111/1467-8721.00157]
14. Sullivan HS, editor. The Interpersonal Theory of Psychiatry: Routledge; 2013:416. [doi: 10.4324/9781315014029]
15. Del Toro J, Wang MT. School cultural socialization and academic performance: examining ethnic-racial identity development as a mediator among African American adolescents. Child Dev 2021 Jul;92(4):1458-1475. [doi: 10.1111/cdev.13467] [Medline: 33205402]
16. Bronfenbrenner U. Ecological models of human development. In: Readings on the Development of Children, 2nd edition 1993, Vol. 3:37-43.
17. Arnett JJ. Emerging adulthood. A theory of development from the late teens through the twenties. Am Psychol 2000 May;55(5):469-480. [doi: 10.1037/0003-066X.55.5.469] [Medline: 10842426]
18. Smetana JG. Adolescents, Families, and Social Development: How Teens Construct Their Worlds: Wiley Blackwell; 2011:327. [doi: 10.1002/9781444390896]
19. Popat A, Tarrant C. Exploring adolescents' perspectives on social media and mental health and well-being - a qualitative literature review. Clin Child Psychol Psychiatry 2023 Jan;28(1):323-337. [doi: 10.1177/13591045221092884] [Medline: 35670473]
20. O'Reilly M, Dogra N, Hughes J, Reilly P, George R, Whiteman N. Potential of social media in promoting mental health in adolescents. Health Promot Int 2019 Oct 1;34(5):981-991. [doi: 10.1093/heapro/day056] [Medline: 30060043]
21. Valkenburg PM, Koutamanis M, Vossen HGM. The concurrent and longitudinal relationships between adolescents' use of social network sites and their social self-esteem. Comput Human Behav 2017 Nov;76:35-41. [doi: 10.1016/j.chb.2017.07.008] [Medline: 29104364]
22. Toma CL, Chen YA. The international encyclopedia of media psychology. In: The International Encyclopedia of Media Psychology, 1st edition 2020:1-10 URL: https://onlinelibrary.wiley.com/doi/book/10.1002/9781119011071 [doi: 10.1002/9781119011071]
23. Valkenburg PM, Peter J. Online communication among adolescents: an integrated model of its attraction, opportunities, and risks. J Adolesc Health 2011 Feb;48(2):121-127. [doi: 10.1016/j.jadohealth.2010.08.020] [Medline: 21257109]

XSL•FO
RenderX

24. Jiang MA. 1. Teens and their experiences on social media. Pew Research Center. 2018 Nov 28. URL: https://www.pewresearch.org/internet/2018/11/28/teens-and-their-experiences-on-social-media/ [accessed 2018-11-28]
25. Brown BB, Larson J. Peer relationships in adolescence. In: Handbook of Adolescent Psychology: Contextual Influences on Adolescent Development, 3rd edition 2009, Vol. 2:74-103. [doi: 10.1002/9780470479193]
26. Berger C, Deutsch N, Cuadros O, et al. Adolescent peer processes in extracurricular activities: identifying developmental opportunities. Child Youth Serv Rev 2020 Nov;118:105457. [doi: 10.1016/j.childyouth.2020.105457]
27. Dworkin JB, Larson R, Hansen D. Adolescents' accounts of growth experiences in youth activities. J Youth Adolesc 2003 Feb;32(1):17-26. [doi: 10.1023/A:1021076222321]
28. Eccles JS, Barber BL, Stone M, Hunt J. Extracurricular activities and adolescent development. J Soc Issues 2003 Dec;59(4):865-889. [doi: 10.1046/j.0022-4537.2003.00095.x]
29. Feldman AF, Matjasko JL. The role of school-based extracurricular activities in adolescent development: a comprehensive review and future directions. Rev Educ Res 2005 Jun;75(2):159-210. [doi: 10.3102/00346543075002159]
30. Schaefer DR, Simpkins SD, Vest AE, Price CD. The contribution of extracurricular activities to adolescent friendships: new insights through social network analysis. Dev Psychol 2011 Jul;47(4):1141-1152. [doi: 10.1037/a0024091] [Medline: 21639618]
31. Pérez-Torres V. Social media: a digital social mirror for identity development during adolescence. Curr Psychol 2024 Jul;43(26):22170-22180. [doi: 10.1007/s12144-024-05980-z]
32. Steinberg L, Morris AS. Adolescent development. Annu Rev Psychol 2001;52(1):83-110. [doi: 10.1146/annurev.psych.52.1.83] [Medline: 11148300]
33. Hjetland GJ, Finserås TR, Sivertsen B, Colman I, Hella RT, Skogen JC. Focus on self-presentation on social media across sociodemographic variables, lifestyles, and personalities: a cross-sectional study. Int J Environ Res Public Health 2022 Sep 5;19(17):11133. [doi: 10.3390/ijerph191711133] [Medline: 36078843]
34. Selkie E, Adkins V, Masters E, Bajpai A, Shumer D. Transgender adolescents' uses of social media for social support. J Adolesc Health 2020 Mar;66(3):275-280. [doi: 10.1016/j.jadohealth.2019.08.011] [Medline: 31690534]
35. Kaakinen M, Sirola A, Savolainen I, Oksanen A. Shared identity and shared information in social media: development and validation of the identity bubble reinforcement scale. Media Psychol 2020 Jan 2;23(1):25-51. [doi: 10.1080/15213269.2018.1544910]
36. Davis K. Young people's digital lives: the impact of interpersonal relationships and digital media use on adolescents' sense of identity. Comput Human Behav 2013 Nov;29(6):2281-2293. [doi: 10.1016/j.chb.2013.05.022]
37. Calvin AJ. Online self-presentation in adolescence. In: The International Encyclopedia of Media Psychology: John Wiley and Sons; 2020:1-10. [doi: 10.1002/9781119011071]
38. Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychology research. Qual Res Psychol 2015 Apr 3;12(2):202-222. [doi: 10.1080/14780887.2014.955224] [Medline: 27499705]
39. Cascio CN, Selkie E, Moreno MA. Effect of technology and digital media use on adolescent health and development: protocol for a multimethod longitudinal study. JMIR Res Protoc 2023 Sep 13;12(1):e50984. [doi: 10.2196/50984] [Medline: 37703071]
40. Rosen LD, Whaling K, Carrier LM, Cheever NA, Rokkum J. The media and technology usage and attitudes scale: an empirical investigation. Comput Human Behav 2013 Nov 1;29(6):2501-2511. [doi: 10.1016/j.chb.2013.06.006] [Medline: 25722534]
41. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
42. Nesi J, Rothenberg WA, Bettis AH, et al. Emotional responses to social media experiences among adolescents: longitudinal associations with depressive symptoms. J Clin Child Adolesc Psychol 2022 Nov 2;51(6):907-922. [doi: 10.1080/15374416.2021.1955370]
43. Kroenke K, Strine TW, Spitzer RL, Williams JBW, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026]
44. Gwet K. Handbook of Inter-Rater Reliability: How to Estimate the Level of Agreement Between Two or Multiple Raters: STATAXIS Publishing Company; 2001.
45. Moreno MA, Jelenchick LA, Egan KG, et al. Feeling bad on Facebook: depression disclosures by college students on a social networking site. Depress Anxiety 2011 Jun;28(6):447-455. [doi: 10.1002/da.20805] [Medline: 21400639]
46. Reporting requirements office of compliance. Office of Compliance. URL: https://compliance.wisc.edu/titleix/mandatory-reporting/ [accessed 2025-08-22]
47. Ilakkuvan V, Johnson A, Villanti AC, Evans WD, Turner M. Patterns of social media use and their relationship to health risks among young adults. J Adolesc Health 2019 Feb;64(2):158-164. [doi: 10.1016/j.jadohealth.2018.06.025] [Medline: 30269907]
48. Moreno MA, Cox ED, Young HN, Haaland W. Underage college students' alcohol displays on Facebook and real-time alcohol behaviors. J Adolesc Health 2015 Jun;56(6):646-651. [doi: 10.1016/j.jadohealth.2015.02.020] [Medline: 26003580]
49. Ahmed SK. Sample size for saturation in qualitative research: debates, definitions, and strategies. J Med Surg Public Health 2025 Apr;5:100171. [doi: 10.1016/j.glmedi.2024.100171]

XSL•FO
RenderX

50. Hamilton JL, Nesi J, Choukas-Bradley S. Reexamining social media and socioemotional well-being among adolescents through the lens of the COVID-19 pandemic: a theoretical review and directions for future research. Perspect Psychol Sci 2022 May;17(3):662-679. [doi: 10.1177/17456916211014189] [Medline: 34756118]
51. Buck AM, Ralston DF. I didn't sign up for your research study: the ethics of using "public" data. Comput Compos 2021 Sep;61:102655. [doi: 10.1016/j.compcom.2021.102655]
52. Cristello JV, Strowger M, Moreno MA, Trucco EM. Navigating the modern landscape of social media: ethical considerations for research with adolescents and young adults. Transl Issues Psychol Sci 2024 Jun;10(2):123-134. [doi: 10.1037/tps0000408] [Medline: 39898233]
53. H.R.7521 - Protecting Americans from foreign adversary controlled applications act. CONGRESSGOV. 2024 Mar 14. URL: https://www.congress.gov/bill/118th-congress/house-bill/7521/text [accessed 2024-03-14]
54. Extending the TikTok enforcement delay- executive orders. The White House. 2025 Apr 4. URL: https://www.whitehouse.gov/presidential-actions/2025/04/extending-the-tiktok-enforcement-delay/ [accessed 2025-08-22]
55. Alhabash S, Smischney TM, Suneja A, Nimmagadda A, White LR. So similar, yet so different: how motivations to use Facebook, Instagram, Twitter, and TikTok predict problematic use and use continuance intentions. Sage Open 2024 Apr;14(2). [doi: 10.1177/21582440241255426]
56. Masciantonio A, Bourguignon D, Bouchat P, Balty M, Rimé B. Don't put all social network sites in one basket: Facebook, Instagram, Twitter, TikTok, and their relations with well-being during the COVID-19 pandemic. PLOS ONE 2021;16(3):e0248384. [doi: 10.1371/journal.pone.0248384] [Medline: 33705462]
57. Criss S, Gonzales SM, Mane H, et al. Race and ethnicity in Facebook images and text: thematic analysis. JMIR Form Res 2025 Sep 4;9:e62713. [doi: 10.2196/62713] [Medline: 40906523]
58. Moreno MA, Egan KG, Brockman L. Development of a researcher codebook for use in evaluating social networking site profiles. J Adolesc Health 2011 Jul;49(1):29-35. [doi: 10.1016/j.jadohealth.2011.04.015] [Medline: 21700153]
59. Sarabia I, Estévez A. Sexualized behaviors on Facebook. Comput Human Behav 2016 Aug;61:219-226. [doi: 10.1016/j.chb.2016.03.037]
60. Bliss L, Zhao Q, Chao I, Haimson OL, Selkie E. Medical information provided by transgender and gender-diverse content creators on YouTube: descriptive content analysis. JMIR Form Res 2025 Aug 29;9:e75787. [doi: 10.2196/75787] [Medline: 40882233]

Abbreviations

AC1: Gwet First-Order agreement coefficient
AI: artificial intelligence
API: Application Programming Interface
ARROW: Application Review for Research Oversight at Wisconsin
ERSMEM: Emotional Responses to Social Media Experiences Measure
GAD-7: General Anxiety Disorder 7-item
HIPAA: Health Insurance Portability and Accountability Act
MTUA: Media and Technology Usage and Attitudes scale
PHQ-8: 8-item Patient Health Questionnaire
PI: principal investigator
REDCap: Research Electronic Data Capture
SGC: self-generated content

Edited by A Schwartz; submitted 19.Sep.2025; peer-reviewed by J Wang, R Nooripour; accepted 17.Feb.2026; published 31.Mar.2026.

Please cite as:
Boyd K, Bliss L, Fan T, Kern K, Carlson CM, Moreno MA, Cascio CN, Selkie E
Directly Observing and Characterizing Adolescents' Self-Generated Social Media Posts: Protocol for Creation and Implementation of a Cyberethnography Informed Codebook
JMIR Res Protoc 2026;15:e84461
URL: https://www.researchprotocols.org/2026/1/e84461
doi:10.2196/84461

© Kylie Boyd, Lydia Bliss, Tingting Fan, Kayla Kern, Caitlin M Carlson, Megan A Moreno, Christopher N Cascio, Ellen Selkie. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR

XSL•FO
RenderX

Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Undergraduate Interprofessional Education: Protocol for a Mixed Methods Study

Hannah Snead[*], BS; Dhwaniben Hada[*]; Lori Ann Eldridge[*], PhD, CTRS

East Carolina University, 3204 Carol G. Belk Building, Greenville, NC, United States

[*]all authors contributed equally

Corresponding Author:
Lori Ann Eldridge, PhD, CTRS
East Carolina University, 3204 Carol G. Belk Building, Greenville, NC, United States

Abstract

Background: Interprofessional education (IPE) is essential for fostering collaboration and communication among health care professionals to improve patient outcomes. Early IPE training for prehealth undergraduate students has been shown to enhance teamwork, dismantle stereotypes, and build self-efficacy in health care settings. Self-efficacy is the belief that one can achieve their goals and succeed at tasks. These outcomes contribute to creating a cohesive and equitable health care environment focused on patient-centered care. Despite evidence supporting the benefits of IPE training, there is a lack of consistent implementation at the undergraduate level across institutions.

Objective: Programs that have adopted IPE have demonstrated improved student preparedness, better interprofessional collaboration, and enhanced patient care outcomes. Our study addresses this gap in undergraduate IPE among institutions by examining the impact of a structured IPE training program, Pathway to Health Professions, designed for prehealth East Carolina University undergraduate students who are members of Buff in Scrubs, an East Carolina University organization.

Methods: Our study uses a mixed methods approach, including retrospective surveys and open-ended questions, to assess changes in participants' attitudes, skills, and competencies before and after training. Data will be collected from January 2025 to April 2025.

Results: We hypothesize that early and structured exposure to IPE training will significantly enhance prehealth undergraduate students' confidence, understanding of interprofessional roles, and ability to collaborate effectively. By focusing on self-efficacy–based skills, we hope to equip prehealth students with the confidence to apply to their desired health professions program.

Conclusions: This paper outlines the protocol for Pathway to Health Professions, a prehealth preparation IPE program. Future testing of the use of this program is outlined. We hypothesize that the results from the future testing of this program will show an improvement in undergraduate prehealth students' self-efficacy related to graduate school readiness.

International Registered Report Identifier (IRRID): DERR1-10.2196/74394

(JMIR Res Protoc 2026;15:e74394) doi:10.2196/74394

KEYWORDS

interprofessional education; undergraduate; health care; self-efficacy; protocol

Introduction

Background

Health care is a diverse field that encompasses many occupations collaborating to provide patient care. It is important for future health care workers to be aware and informed of how to interact with one another to have a productive health care team. Our program, Pathway to Health Professions, is based on current research regarding the best interprofessional education (IPE) content and practices. This IPE program is intended to equip prehealth students with different skills to help prepare them for their future careers. Some of these skills include but are not limited to teamwork, leadership, collaboration between health care disciplines, networking, and decision-making. This program hopes to inspire prehealth undergraduates to invest in their future health care occupations and learn about IPE within the health care field. IPE involves developing and structuring programs to educate and facilitate collaboration across various health care professions [1]. It focuses on collaboration and communication between professionals for better-working teams [2]. Multiple studies have demonstrated that the early implementation of IPE training for undergraduate students can have a positive impact on teamwork, collaboration, and confidence [3].

Studies have found that introducing IPE before starting graduate training programs enhances the understanding of various health professions [4]. IPE training helps dismantle stereotypes that

XSL•FO
RenderX

many prehealth students and professionals hold about health care teams [5,6]. Eliminating these stereotypes and receiving information on the role of each profession creates a more cohesive health care environment for all health care professionals to work together [7]. Early focus on teamwork and collaboration can create more responsive students who can better adapt to group environments, which affects patient care and outcomes [8,9]. IPE training can focus on caring for vulnerable populations and ensure equitable care [10]. Not only is equitable care an outcome of IPE training, but it has also been shown that the use of IPE training creates a more comprehensive team that further promotes better health outcomes and shorter hospital stays for patients [11]. Furthermore, it has been shown that IPE training offers students a safe environment in which to practice caring for patients. This overall improves predicted patient outcomes and encourages confidence in students [12]. Major universities and medical programs have recognized the gaps present and the outcomes of IPE training and have begun implementing IPE training into their programs to bridge these gaps and prevent lapses in care [13]. Subsequently, IPE training has been successfully implemented in specific medical programs, yielding positive outcomes [14]. Those programs that have implemented IPE training have observed better graduate outcomes in students who completed it, as well as an overall increase in positive health care environments [15]. However, despite the success that several medical programs have experienced, many others have not incorporated IPE training [16].

Studies suggest that implementing IPE training sessions at the undergraduate level can address the improvement in teamwork, mutual respect, and patient-centered care approaches required by upper-level programs [17]. This has been supported by multiple studies that show the outcomes of implementing IPE programs [18]. In one study, multiple prehealth students focused on different professions were brought together. They participated in a short IPE training in which collaboration and education were a focus. After the study, students felt more competent and confident in their leadership skills and felt that they could better understand their interprofessional coworkers [3]. Another cross-sectional study of students who completed surveys while in their undergraduate and postgraduate programs found that identity and teamwork were affected by involvement in IPE events [15]. These short IPE studies show the need for longer programs with multiple sessions focusing on early intervention for prehealth undergraduate students. Furthermore, the most effective IPE training is multidisciplinary and uses various teaching methods [3]. With the increased interest in the continuation and implementation in universities and higher education programs, IPE training programs with the mission to educate prehealth students will become more common in the future [19].

IPE Training

IPE refers to the education and overall design of programs that focus on the education and facilitation of learning among different health care professions. This is especially important in the health care field, where people from different occupations collaborate to treat patients. Given the variations among patients, it is best for the health care team to operate smoothly for fast and reliable care [2].

A meta-study by Spaulding et al [1] examined 19 articles to determine the benefits of exposing health care students to other health care fields. Attitudes toward other disciplines and the value placed on a team-based approach were found to improve statistically in 17 of the 19 studies. Due to the importance of IPE in the health care field, many universities now require IPE-specific health accreditation standards for disciplines such as dentistry, medicine, nursing, occupational therapy, pharmacy, physical therapy, physician assistant studies, and public health [20]. Spaulding et al [1] identified 7 studies that demonstrated statistically significant improvements in collaborative behavior through IPE. Another study examined whether the introduction of IPE influenced students' perceptions of certain occupations. The study found that students who participated in IPE activities exhibited significant changes in behavior and attitudes toward other occupations [21]. While exposure to collaborative environments in health care education fosters improved interdisciplinary attitudes and behaviors, another critical factor influencing student success is self-efficacy.

Significant research has been dedicated to improving undergraduate students' confidence, which has been linked to an overall enhancement in their professional experiences. A study specifically focused on preparing medical students for the health field found that having a strong professional identity and effective collaboration skills was important for future success [15]. Another study examined the relationship between health care and IPE centered on participant impact [4]. During students' clinical rotations, they received IPE training and were later surveyed about their experiences. The findings revealed that introducing IPE before students started their programs enhanced their understanding of various health professions [4]. This literature indicates that self-efficacy and IPE may enhance one another and offer better support to prehealth students.

Self-Efficacy: Comparison With Prior Work

Prior interventions to prepare prehealth students or early-stage trainees often focus on a single component, such as mentoring, shadowing, standardized patient encounters, or graduate program application workshops, and are typically delivered as stand-alone workshops or brief series. In contrast, the Pathway to Health Professions program integrates these elements into a coherent, longitudinal, IPE-focused program (Multimedia Appendix 1).

Session 1 ("Mentor Match Up") extends existing mentoring models by pairing undergraduates with graduate-level health profession students, providing near-peer guidance that may be particularly salient for first-generation students and underrepresented students, such as those from rural communities; women; Black, Latino, and Native American students; and students with a low socioeconomic status. Session 2 ("Rounds Revealed") adapts clinical reflection and ward-round models to an undergraduate context to demystify clinical environments and normalize early professional socialization. Session 3 ("Medical Mock Trials") builds on simulation and standardized patient approaches to foster communication skills, role clarity, and critical thinking in interprofessional teams. Session 4 ("Equity in Action") explicitly centers health equity

XSL•FO
RenderX

and social determinants of health, which are often underemphasized in prehealth curricula despite their relevance to graduate competencies. Sessions 5 and 6 ("Pathway Prep" and "Health Roles Unveiled") respond to literature documenting confusion and anxiety about competitive graduate program application processes and limited understanding of diverse health careers by offering structured, concrete guidance on applications, interviews, and interprofessional roles.

By linking these evidence-informed components into a sequenced curriculum, this program seeks to address multiple domains of self-efficacy simultaneously, academic and professional preparedness, networking, communication, equity-focused practice, and role clarity rather than targeting any single skill in isolation.

Study Purpose and Objectives

An IPE-based training program called Pathway to Health Professions was developed to further support prehealth undergraduate students and prepare them for higher-level graduate programs. The Pathway to Health Professions program (n=6 sessions) will be tested to determine whether it was effective in improving self-efficacy regarding graduate school for undergraduate students in prehealth majors. We hypothesize that prehealth students who engage in all 6 Pathway to Health Professions IPE sessions will have improved self-efficacy regarding graduate training programs. This program differs from those in previous studies in that it spans a longer period and has multiple sessions addressing various aspects of health care IPE.

Methods

Study Design: Theory

In 1984, Bandura [22] introduced the theory of self-efficacy, which described the relationship between confidence and motivation. It refers to the fact that people are more likely to complete a task or move forward with a goal if they believe they have the capability to do so. Self-efficacy theory has 4 main components: experience, vicarious experience, social persuasion, and physiological feedback [22] (Figure 1).

Figure 1. The 4 elements of self-efficacy.

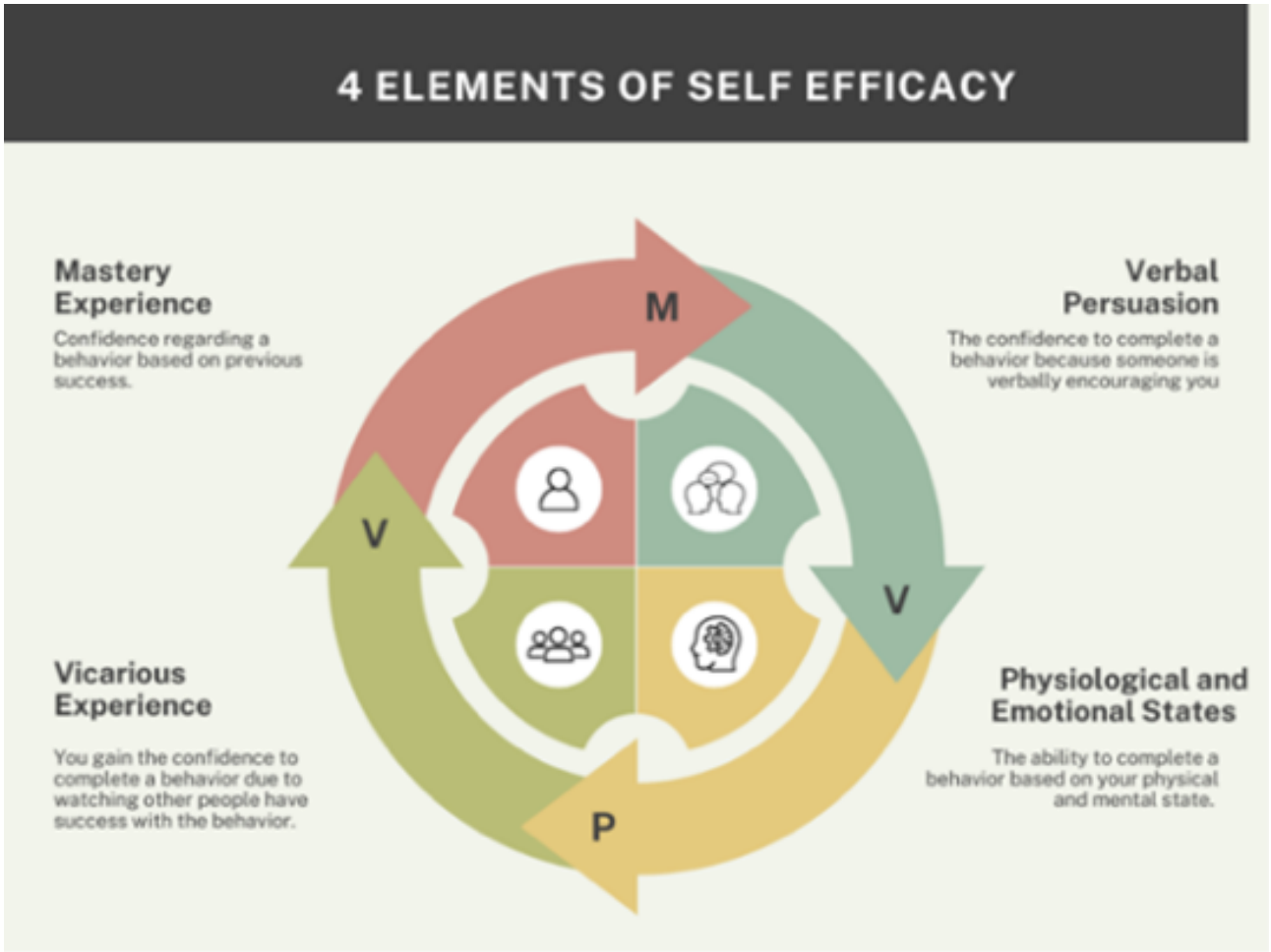

Participation and Recruitment

Inclusion

To participate in this study, one must be an undergraduate student at East Carolina University majoring in a prehealth field. Prehealth majors include but are not limited to biology, chemistry, psychology, public health, physics, kinesiology, and recreation therapy.

Exclusion

Undergraduate students who are not in a prehealth major (eg, business, computer science, English, or geology) and graduate students will be excluded.

XSL•FO
RenderX

Recruitment

Participants for this study will be recruited from undergraduate prehealth programs. The university offers multiple programs, including but not limited to nursing, physician assistant, physical therapy, occupational therapy, speech therapy, recreational therapy, and audiology programs. The university is known for its large population of prehealth students.

The primary recruitment form will be distributed via GroupMe and social media messaging. We will verbally announce the same message to Buff in Scrubs organization members at meetings. Emails will be sent to students via the university honors program and the preprofessional advising center's weekly newsletter. Flyers will be posted on campus in the following locations: dormitories, the public health building, and the library.

Pathway to Health Professions Intervention

Pathway to Health Professions consists of 6 sessions; each session was developed based on literature focusing on IPE to improve undergraduate student self-efficacy (Table 1). At the conclusion of each IPE session, the participants will be asked to complete a retrospective quantitative survey.

XSL•FO
RenderX

Table . Pathway to Health Professions sessions.

Sessions	Session description	Self-efficacy skills	References
Session 1: Mentor Match Up	Included participants will be paired up with students currently in graduate health programs. During this matchup, undergraduate students will ask graduate students questions and advice about their specific programs.	• Self-esteem • Being informed about tasks to reach goals • Confidence in ability to network • Confidence in written and verbal communication • Confidence in ability to perform tasks and activities required to achieve career goals	[23,24]
Session 2: Rounds Revealed	During this session, students will be provided with literature and videos that outline what it would look like to shadow a physician. Students will also have an opportunity to talk about their own experiences while shadowing and what they learned.	• Confidence in ability to network • Confidence in cold communication with health care providers • Ability to ask questions and learn from health care providers	[2,25]
Session 3: Medical Mock Trials	During this session, students will pair up and portray mock patient-health care provider relationships to gain an understanding of the need for patient education. Students will portray diverse types of providers and work together on a case study, mirroring real-life patient care teams.	• Consideration of whether the chosen health career is suitable for the student	[26,27]
Session 4: Equity in Action: Health Challenges and Solutions	During this session, a guest speaker will convey the importance of health equity and how providers can work together to address issues patients face. The speaker will address how different populations face different issues and how, as future providers, the students' job is to reduce those differences.	• Confidence in ability to work effectively with medically underserved communities • Understanding social, economic, historical, political, psychosocial, and cultural factors that influence the health of underserved populations • Familiarity with community resources available to assist members of health care teams that treat underserved populations • Confidence in ability to communicate with and assist people with different needs	[13,28]
Session 5: Pathway Prep: Application and Shadowing Tips/Graduate Applications Walkthrough	During this session, similarities and differences among the most popular graduate program application processes (MD, PT[a], PA[b], dentistry, and nursing) for current Buff in Scrubs students will be discussed and reviewed. A slideshow presentation detailing application requirements and recommendations will be shown. For shadowing, students will be provided with guidance on how to contact physicians and a sample email template. Students can begin the application process during the session and ask questions. The 3 most popular processes will be determined based on the members' interests.	• Preparedness to present oneself during an interview • Confidence in ability to perform all the tasks and activities required to achieve health science career goals	[29,30]

XSL•FO
RenderX

Sessions	Session description	Self-efficacy skills	References
Session 6: Health Roles Unveiled: Occupation Walkthroughs or Profession Profiles: Interprofessional Skits and Talks	During this session, students will learn about different professions and what a day in the life looks like. This will encourage students to learn more about different occupations. Additionally, it will allow participants to be aware of what their counterparts in other professions do when treating patients.	• Being informed about what needs to be done to reach the goal of becoming a health science professional • Ability to see oneself as capable of achieving and practicing their chosen health science career	[30,31]

[a]PT: physical therapist.

[b]PA: physician assistant.

The 6 sessions (Table 1) will be completed over 3 months, with two 90-minute or 1-hour sessions per month. Sessions will address common concerns or issues that prehealth students experience (Multimedia Appendix 1). Each session has been developed through a literature review and is characterized by various objectives that have been shown to benefit prehealth students in multiple studies.

Implementation of the Intervention

A detailed description of each session is provided in Multimedia Appendix 1. This description includes the purpose, name of the session, location of the session, session description, materials needed, the goals and objectives, the facilitation process, variations to the session, and references. At the conclusion of each session, a survey will be administered. Participants will scan a QR code and complete the survey on a personal device. The same instructions will be given to participants by each session facilitator to ensure consistent delivery of the survey. Participants will be told that survey participation is optional and anonymous to reduce expectation and confirmation bias.

Outcomes

Measures and Outcomes

Six surveys (Multimedia Appendix 2) were developed comprising a Likert scale using questions from academically sourced journals. This will allow the survey to be reputable and valid. Surveys and questionnaires that assess self-efficacy were used as sources. Some questions' wording was altered to fit the Pathway to Health Professions program language. These alterations were small and would not affect the validity of the questions.

At a student's first session in the program, they complete a brief demographic questionnaire that includes current year of study, age, gender, race, ethnicity, and aspired profession. Participants also complete a brief survey at the end of each session that includes demographics, a repeated primary outcome measure, session-specific items, and open-ended feedback.

The primary outcome is self-efficacy, assessed using the same 6 items at each session (refer to Multimedia Appendix 1 for the exact wording). Four items use Likert-type agreement response options ("strongly disagree" to "strongly agree"), and 2 items use a confidence response option ("not confident" to "completely confident"). Each session also includes topic-specific questions unique to that session, all assessed using a Likert-type agreement scale ("strongly disagree" to "strongly agree"). The number of session-specific items varies by session: session 1 (1 item), session 2 (2 items), session 3 (1 item), session 4 (4 items), session 5 (2 items), and session 6 (1 item). In addition, participants respond to a global program satisfaction item ("program recommendation") rated from "not at all recommend" to "highly recommend" and an open-ended qualitative item ("Please share your feedback"). Open-ended responses are used to capture qualitative suggestions for program improvement. Surveys are administered using Qualtrics (Qualtrics International Inc) at the end of each session. Participants will be informed that completing the surveys is optional.

Survey Matching Procedures

We will follow a previously established method that will use facts about a person and generate a numbered code to match them to their responses [32]. For our surveys, we will use the number of older siblings of the participants, the name of their first pet, the color of their backpack, and the last 2 digits of their phone number. This will allow us to match the survey data anonymously. Furthermore, matching will allow us to assess change over time in the participant.

Analytic Methods

A Wilcoxon signed-rank test will be used to assess changes in responses over time for Likert-scale data. The Likert-scale items, ranging from, for example, "strongly disagree" (1) to "strongly agree" (5), will be treated as ordinal data, and changes will be examined retrospectively during each session.

Given the ordinal nature of Likert-scale data, nonparametric tests will be used over parametric methods that assume normality. Specifically, the Friedman test will be used to assess the overall differences in median scores across multiple time points. Post hoc pairwise comparisons with Bonferroni correction will be performed if the Friedman test indicates a statistically significant result. All analyses will be conducted using SPSS (version 28; IBM Corp).

Ethical Considerations

This study has been approved by the East Carolina University and Medical Center Institutional Review Board (UMCIRB 24-001921).

Informed consent was obtained from all participants prior to data collection. Participants were provided with detailed information about the study purpose, procedures, risks, and benefits and were informed that their participation was voluntary

XSL•FO
RenderX

and that they could withdraw at any time without penalty. Participants were not paid.

To support anonymous tracking of responses across multiple sessions, participants were asked to generate responses to a set of anonymous identification questions (see Survey Matching Procedures section). These identifiers were designed to protect participant confidentiality while allowing for linkage of responses over time. No personally identifiable information was collected, and all data were stored securely in accordance with institutional review board guidelines.

Results

Data were collected from January 2025 to April 2025. As of June 2025, 9 individuals were recruited. Data analysis is currently underway. Results are expected to be published in January 2026.

Discussion

Expected Findings

Pathway to Health Professions is designed to strengthen prehealth undergraduates' self-efficacy for pursuing graduate health profession training. We anticipate that students who participate in all 6 IPE-based sessions over the 3-month period will report meaningful improvements in perceived preparedness for graduate applications and interviews, confidence in networking and communicating with health professionals, clarity regarding career fit, and confidence in working with medically underserved communities. We further hypothesize that repeated exposure to interprofessional content and role models across sessions will reinforce students' sense of belonging in health professions and their belief that they can successfully navigate the pathway to graduate training.

This protocol offers several contributions. First, it provides a replicable, session-by-session structure for an IPE-based prehealth preparation program that can be adopted or adapted by institutions seeking to strengthen their health profession pipeline. Second, the curriculum is explicitly mapped to specific self-efficacy domains (eg, confidence in networking, communicating with health care providers, navigating graduate program applications, and working with underserved communities), allowing future evaluation to examine which components are most strongly associated with gains in different domains. Third, the program brings together multiple types of experiential learning, mentoring, reflective discussion, simulation, and career exploration within an interprofessional framework, which may better mirror the collaborative nature of contemporary health care.

A further strength is the program's emphasis on equity. The inclusion of a dedicated health equity session and the repeated attention to medically underserved populations across activities are intended to reinforce students' sense of responsibility for reducing health disparities and align their career aspirations with the needs of marginalized communities. Finally, the program is designed to be feasible within a typical academic semester (six 90-minute sessions over 3 months), which may facilitate adoption at institutions with limited curricular flexibility.

Future Directions and Dissemination

Following implementation and evaluation, we plan to refine the program based on participant feedback, outcome data, and feasibility considerations. Future work could include longer-term follow-up to assess maintenance of self-efficacy gains and downstream outcomes such as graduate program applications, acceptances, and persistence in health profession training. Adaptations for different institution types (eg, community colleges and minority-serving institutions) and delivery formats (hybrid or fully online) will also be important to enhance scalability. In addition, subgroup analyses (eg, by first-generation status or underrepresented racial or ethnic groups) may help determine whether the program differentially benefits students who have historically faced barriers to entering health professions.

Dissemination of the Pathway to Health Professions program will occur through multiple channels, including open access publication of the curriculum, sharing of session materials with campus partners, and presentations. If future evaluations demonstrate positive effects on student self-efficacy and graduate school readiness, this protocol could serve as a model for institutions seeking structured, equity-oriented IPE experiences to strengthen the health profession workforce pipeline.

Limitations

This protocol has several limitations, the first of which is the surveys. The surveys themselves may be difficult for students to complete. For example, we may have students who attend a certain session but not others, so we will not have a full dataset including all 6 sessions for each student. The Likert scale itself has limitations. Likert scales present an inherent bias that can affect how respondents answer questions. Furthermore, the scale does not offer many options and limits respondents to answers they may disagree with [33]. The retrospective nature of the survey introduces potential recall bias and may affect the accuracy of the results. We will attempt to mitigate recall bias by conducting the surveys immediately following each session.

Conclusions

This protocol paper outlines Pathway to Health Professions, a prehealth preparation IPE program. Future testing of the use of this program is outlined. We hypothesize that the results from the future testing of this program will show an improvement in undergraduate prehealth students' self-efficacy related to graduate school readiness.

XSL•FO
RenderX

Acknowledgments

Artificial intelligence use was limited to coming up with topic names for the sessions. This was the only time artificial intelligence was used in this study.

Funding

This project was completed without funding.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Session descriptions.
[PDF File, 723 KB - resprot_v15i1e74394_app1.pdf]

Multimedia Appendix 2
Session surveys.
[PDF File, 1995 KB - resprot_v15i1e74394_app2.pdf]

References

1. Spaulding EM, Marvel FA, Jacob E, et al. Interprofessional education and collaboration among healthcare students and professionals: a systematic review and call for action. J Interprof Care 2021;35(4):612-621. [doi: 10.1080/13561820.2019.1697214] [Medline: 31865823]
2. Walton V, Hogden A, Long JC, Johnson JK, Greenfield D. How do interprofessional healthcare teams perceive the benefits and challenges of interdisciplinary ward rounds. J Multidiscip Healthc 2019;12:1023-1032. [doi: 10.2147/JMDH.S226330] [Medline: 31849478]
3. Jung H, Park KH, Min YH, Ji E. The effectiveness of interprofessional education programs for medical, nursing, and pharmacy students. Korean J Med Educ 2020 Jun;32(2):131-142. [doi: 10.3946/kjme.2020.161] [Medline: 32486622]
4. Gunaldo TP, Augustus-Wallace A, Schilling D, Hicks M, Goumas A, Levitzky E. Gaining an understanding of interprofessional roles and responsibilities among pre-professional health students. Health Interprof Pract Educ ;4(1):1191. [doi: 10.7710/2641-1148.1191]
5. Mohammed CA, Anand R, Saleena Ummer V. Interprofessional education (IPE): a framework for introducing teamwork and collaboration in health professions curriculum. Med J Armed Forces India 2021 Feb;77(Suppl 1):S16-S21. [doi: 10.1016/j.mjafi.2021.01.012] [Medline: 33612927]
6. Bosch B, Mansell H. Interprofessional collaboration in health care: lessons to be learned from competitive sports. Can Pharm J (Ott) 2015 Jul;148(4):176-179. [doi: 10.1177/1715163515588106] [Medline: 26448769]
7. Schot E, Tummers L, Noordegraaf M. Working on working together. A systematic review on how healthcare professionals contribute to interprofessional collaboration. J Interprof Care 2020;34(3):332-342. [doi: 10.1080/13561820.2019.1636007] [Medline: 31329469]
8. Murray M. The impact of interprofessional simulation on readiness for interprofessional learning in health professions students. Teach Learn Nurs 2021 Jul;16(3):199-204. [doi: 10.1016/j.teln.2021.03.004]
9. O'Leary N, Salmon N, Clifford AM. "It benefits patient care": the value of practice-based IPE in healthcare curriculums. BMC Med Educ 2020 Nov 12;20(1):424. [doi: 10.1186/s12909-020-02356-2] [Medline: 33183276]
10. Carey MJ, Taylor M. The impact of interprofessional practice models on health service inequity: an integrative systematic review. J Health Organ Manag 2021 Oct 8;35(6):682-700. [doi: 10.1108/JHOM-04-2020-0165] [Medline: 33829715]
11. Recto P, Lesser J, Farokhi MR, et al. Addressing health inequities through co-curricular interprofessional education: a secondary analysis scoping review. J Interprof Educ Pract 2022 Dec;29:100549. [doi: 10.1016/j.xjep.2022.100549]
12. Data S, Dubé MM, Bajunirwe F, et al. Feasibility of an interprofessional, simulation-based curriculum to improve teamwork skills, clinical skills, and knowledge of undergraduate medical and nursing students in Uganda. Sim Healthcare 2021;16(6):e100-e108. [doi: 10.1097/SIH.0000000000000531]
13. Maldonado ME, Fried ED, DuBose TD, Nelson C, Breida M. The role that graduate medical education must play in ensuring health equity and eliminating health care disparities. Ann Am Thorac Soc 2014 May;11(4):603-607. [doi: 10.1513/AnnalsATS.201402-068PS] [Medline: 24708150]
14. Berger-Estilita J, Fuchs A, Hahn M, Chiang H, Greif R. Attitudes towards interprofessional education in the medical curriculum: a systematic review of the literature. BMC Med Educ 2020 Aug 6;20(1):254. [doi: 10.1186/s12909-020-02176-4] [Medline: 32762740]

XSL•FO
RenderX

15. Chaou CH, Yu SR, Chang YC, et al. The evolution of medical students' preparedness for clinical practice during the transition of graduation: a longitudinal study from the undergraduate to postgraduate periods. BMC Med Educ 2021 May 6;21(1):260. [doi: 10.1186/s12909-021-02679-8] [Medline: 33957907]
16. Sezgin MG, Bektas H. Effectiveness of interprofessional simulation-based education programs to improve teamwork and communication for students in the healthcare profession: a systematic review and meta-analysis of randomized controlled trials. Nurse Educ Today 2023 Jan;120:105619. [doi: 10.1016/j.nedt.2022.105619] [Medline: 36343420]
17. Samarasekera DD, Nyoni CN, Amaral E, Grant J. Challenges and opportunities in interprofessional education and practice. The Lancet 2022 Oct;400(10362):1495-1497. [doi: 10.1016/S0140-6736(22)02086-4]
18. Marion-Martins AD, Pinho DLM. Interprofessional simulation effects for healthcare students: a systematic review and meta-analysis. Nurse Educ Today 2020 Nov;94:104568. [doi: 10.1016/j.nedt.2020.104568] [Medline: 32932058]
19. Zaher S, Otaki F, Zary N, Al Marzouqi A, Radhakrishnan R. Effect of introducing interprofessional education concepts on students of various healthcare disciplines: a pre-post study in the United Arab Emirates. BMC Med Educ 2022 Jul 2;22(1):517. [doi: 10.1186/s12909-022-03571-9] [Medline: 35780117]
20. Greer AG, Clay M, Blue A, Evans CH, Garr D. The status of interprofessional education and interprofessional prevention education in academic health centers: a national baseline study. Acad Med 2014 May;89(5):799-805. [doi: 10.1097/ACM.0000000000000232] [Medline: 24667499]
21. Hoffman SJ, Harnish D. The merit of mandatory interprofessional education for pre-health professional students. Med Teach 2007 Oct;29(8):e235-e242. [doi: 10.1080/01421590701551672] [Medline: 18236267]
22. Bandura A. Self-efficacy mechanism in human agency. Am Psychol 1982;37(2):122-147. [doi: 10.1037/0003-066X.37.2.122]
23. Brown BK, Hanson SH. Development of a student mentoring program. Am J Pharm Educ 2003;67(4):121. [doi: 10.1016/S0002-9459(24)00650-8]
24. Nimmons D, Giny S, Rosenthal J. Medical student mentoring programs: current insights. Adv Med Educ Pract 2019;10:113-123 [FREE Full text] [doi: 10.2147/AMEP.S154974] [Medline: 30881173]
25. Hamilton D, Taylor C, Maben J. How does a group reflection intervention (Schwartz rounds) work within healthcare undergraduate settings? A realist review. Perspect Med Educ 2023;12(1):550-564. [doi: 10.5334/pme.930] [Medline: 38144671]
26. Flanagan OL, Cummings KM. Standardized patients in medical education: a review of the literature. Cureus 2023 Jul;15(7):e42027. [doi: 10.7759/cureus.42027] [Medline: 37593270]
27. Song C, Jang A. Mock trial as a simulation strategy allowing undergraduate nursing students to experience evidence-based practice: a scoping-review. PLOS ONE 2023;18(8):e0289789. [doi: 10.1371/journal.pone.0289789] [Medline: 37561724]
28. Cooper LA, Purnell TS, Showell NN, et al. Progress on major public health challenges: the importance of equity. Public Health Rep 2018;133(1_suppl):15S-19S. [doi: 10.1177/0033354918795164] [Medline: 30426874]
29. Castillo E. Graduate school application accessibility for international students [thesis]. The University of Texas at El Paso. 2020 URL: https://www.proquest.com/dissertations-theses/graduate-school-application-accessibility/docview/2423836940/se-2 [accessed 2026-03-27]
30. Estien C, Chapman M, Schell CJ, Lowy N, Gerson JR. Demystifying the graduate school application process. Life Sciences Preprint posted online on 2022. [doi: 10.32942/OSF.IO/E5UMR]
31. Beggs RL, Sidwell DK, Rattray ME, Neville GE, Zimmerman PAP. Undergraduate nursing students' perceptions of the graduate application process: a qualitative exploration. Collegian 2022 Apr;29(2):213-219. [doi: 10.1016/j.colegn.2021.08.006]
32. Agley J, Tidd D, Jun M, et al. Developing and validating a novel anonymous method for matching longitudinal school-based data. Educ Psychol Meas 2021 Feb;81(1):90-109. [doi: 10.1177/0013164420938457] [Medline: 33456063]
33. Tanujaya B, Prahmana RCI, Mumu J. Likert scale in social sciences research: problems and difficulties. FWU Journal 2022 Dec 27;16(4):89-101. [doi: 10.51709/19951272/Winter2022/7]

Abbreviations

IPE: interprofessional education

Edited by J Sarvestan; submitted 23.Mar.2025; peer-reviewed by E Levine; revised version received 08.Feb.2026; accepted 16.Feb.2026; published 08.Apr.2026.

Please cite as:
Snead H, Hada D, Eldridge LA
Undergraduate Interprofessional Education: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e74394
URL: https://www.researchprotocols.org/2026/1/e74394
doi:10.2196/74394

XSL•FO
RenderX

© Hannah Snead, Dhwani Hada, Lori Ann Eldridge. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 8.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Validity and Reliability of the Psychodynamic Organizational Diagnostic Instrument SyMOA: Protocol for a Mixed Methods Study

Anne-Maria Müller, MA; Sophia Sachs, MSc; Yannik Rieder, MSc; Claas Lahmann, Prof Dr Med

Department of Psychosomatic Medicine and Psychotherapy, Faculty of Medicine, University Medical Center Freiburg, Hauptstrasse 8, Freiburg, Germany

Corresponding Author:
Anne-Maria Müller, MA
Department of Psychosomatic Medicine and Psychotherapy, Faculty of Medicine, University Medical Center Freiburg, Hauptstrasse 8, Freiburg, Germany

Abstract

Background: A comprehensive understanding of organizations is fundamental for implementing successful change measures. However, to date, there is no empirically testable, operationalized systems-psychodynamic organizational diagnostic method that can capture the deeper, more complex dynamics that are crucial for sustainable transformation. To address this gap, we developed the Systematic Multidimensional Organizational Assessment (SyMOA), a qualitative instrument based on an evidence-based clinical diagnostic framework, the Operationalized Psychodynamic Diagnostics III. SyMOA integrates clinical, organizational, and systemic psychodynamic theory and analyzes an organization's challenges based on invisible and unconscious aspects, that is, those lurking beneath the surface. It hypothesizes 3 organizational dimensions: (1) current challenges based on the sociotechnical integration and organizational internal functioning level, (2) internal relationship dynamics, and (3) unconscious organizational conflicts. The SyMOA dimensions are operationalized into a semistructured interview guide and coding protocols for the analysis of the content. By capturing the underlying dynamics, SyMOA aims to provide a deeper understanding of an organization's challenges and establish a solid foundation for targeted interventions.

Objective: This study aims to evaluate the validity and intercoder reliability of the SyMOA instrument.

Methods: For this purpose, semistructured interviews will be conducted with employees of at least 3 different companies in Germany. The evaluation will be carried out by calculating Krippendorff α to determine intercoder reliability. In addition, construct validity, content validity, and external validity will also be analyzed.

Results: Recruitment and training commenced in May 2025. Data collection is planned for the second half of 2025, with analysis to follow thereafter. As this is a study protocol, no results are available yet. At the time of submission, 46 participants have been recruited.

Conclusions: This study will give methodological insights into the validity, reliability, feasibility, and acceptability of the SyMOA instrument. The findings are expected to help further instrument refinement and inform the application of systems-psychodynamic approaches in organizational diagnostics.

Trial Registration: German Clinical Trials Register DRKS00039411; https://drks.de/search/en/trial/DRKS00039411

International Registered Report Identifier (IRRID): DERR1-10.2196/77554

(JMIR Res Protoc 2026;15:e77554) doi:10.2196/77554

KEYWORDS

organizational assessment; systems psychodynamics; psychodynamic theory; organizational health; validation

Introduction

Background

Organizations are complex social systems that are shaped by the interaction of their members. These individuals bring their own history, personality, and emotions, which significantly influence the dynamics within an organization [1]. At the same time, organizations are in a constant state of change in order to adapt to external changes and secure their long-term existence [2]. Nevertheless, planned change processes are often slow, ineffective, and meet with resistance [3,4]. In addition, organizational employees experience change processes as a significant source of stress [4]. Resistance toward change and adaptation processes is common and signals the activation of—at least initially mainly unconscious—anxiety [5].

XSL•FO
RenderX

According to many researchers, the basis for meaningful and sustainable change management is a thorough and focused organizational diagnosis [3,4,6,7]. It is evident, though, that organizational structures and processes go far beyond the visible level, and models of organizational behavior based on behaviorist concepts have not been sufficient to look beneath the surface and identify the deep structure of underlying problems [8]. Systems psychodynamic research has contributed to understanding these hidden layers by addressing the interplay between collective structures, norms and practices, as well as individual cognitions, motivations, and emotions [1]. Systems psychodynamic theory combines the theory of open systems, which describes organizations as self-regulating systems that interact with their environment [1], with psychodynamic theory that understands organizational conflicts as unavoidable and structure-shaping elements. Psychodynamic theory postulates that the organization becomes a symbolic replica of the family in which members try to fulfill basic psychological needs. In this sense, organizational culture can be understood as an institutionalized defense mechanism that provides members of the organization relief from the anxiety caused by unavoidable change [5]. These social defense mechanisms are embedded in the organizational structure and thus influence the thinking, feeling, and behavior of its members [9].

Literature Gap and Problem Statement

Theoretical and practice-oriented systems-psychodynamic approaches to organizational analysis are existent and well established [1,8,10]. Nevertheless, to date, there is no empirically testable, standardized, and operationalized diagnostic instrument capable of systematically assessing psychodynamic content within organizations while providing a structured procedure and a structured coding and interpretation protocol that translates systems-psychodynamic organizational theory into a transparent and reproducible diagnostic instrument.

The lack of an empirically testable and operationalized systems-psychodynamic diagnostic instrument is a problem for research and practice alike. Without this tool, unconscious conflict patterns, relational dynamics, organizational defenses, and the inner-organizational level of functioning cannot be analyzed in a structured, transparent, or reproducible manner. On the side of organizational practice, this limitation restricts organizations' ability to understand the deeper causes of resistance to change and failed change initiatives. On the side of research, the lack of an operationalized instrument hinders the systematic empirical testing and comparison of findings in systems-psychodynamic organizational research.

Prior Work

Overview

To close this gap, the authors of this article developed the Systematic Multidimensional Organizational Assessment (SyMOA, publication in process). SyMOA is a qualitative, systems-psychodynamic diagnostic instrument that aims to assess organizations based on the 3 dimensions of (1) current challenges analyzed on the basis of sociotechnical integration and internal level of functioning, (2) underlying relationship dynamics, and (3) unconscious organizational conflicts. SyMOA was developed by the authors using a theory-to-research approach based on clinical, psychodynamic, and organizational theories and was refined through an iterative process of theory development, empirical application in small- and medium-sized enterprises (SMEs), and theoretical refinement. The model from clinical practice that served as the basis for SyMOA's in-depth organizational diagnostics is the Operationalized Psychodynamic Diagnostics III [11]. It enables a comprehensive analysis of psychodynamic processes and is successfully used for therapy planning [11,12]. Adapted and applied to organizations, such diagnostics can help to specifically support a company in understanding the causes of resistance and enable its ability to learn and adapt [13]. There is precedence for transferring psychological instruments, developed with the individual in mind, into the organizational context. Psychological concepts like emotional intelligence, motivation theories, and psychodynamic coaching have already been successfully adapted to organizations. Furthermore, both individuals and organizations exhibit complex psychological dynamics such as defense mechanisms, conflicts, and relational patterns [8]. SyMOA in its current version is a comprehensive and in-depth qualitative instrument that aims to capture systemic and psychodynamic aspects within an organization that no existing instrument has been able to map so far. It provides a semistructured interview procedure and a structured, operationalized rating system designed to enable psychodynamic assessment. This developmental work has been done separately and prior to the present study.

SyMOA Dimension I: Current Challenges Based in the Sociotechnical Integration and the Organization's Internal Level of Functioning

The first dimension describes the organization's understanding of current and past challenges, including its resources and barriers to change. Additionally, this dimension delves deeper beneath the surface to examine the organization's internal level of functioning. This includes how the organization, as a system, perceives itself and its environment, how it regulates itself both internally and externally in everyday situations as well as in critical moments, how internal and external communication functions, and how well it establishes bonds—both internally among employees and externally with stakeholders such as customers, the market, and supply chains. Furthermore, this dimension captures the defense mechanisms typically used by the organization.

SyMOA Dimension II: Internal Relationship Dynamics of the Organization

The second dimension assesses the relationship dynamics between different groups within the organization. This includes relationship patterns in typical situations between individual employees, within teams, between teams, between management and stakeholders, and between management and employees. Relationship dynamics are increasingly seen as the key to overcoming nonroutine challenges. Assessment of the relational dynamics in the organization focuses on observable relational behavior and conscious emotional experience in specific situations. Of high importance for the assessment is the perspective of experience and how the specific constellations

XSL•FO
RenderX

of perspectives influence (dys)functional patterns of relationships. It can shed light on recurring interaction patterns in organizations, where party X experiences itself and others (party Y) in a given situation, and how the other party (Y) experiences party X, as well as how party Y experiences itself in interaction with party X.

SyMOA Dimension III: Unconscious Organizational Conflicts

Based on psychodynamic theories of motivation, emotions, and unconscious conflicts, this dimension conceptualizes the internal conflicts within organizations. Typical conflicts include dependence versus autonomy, control versus subordination, intrinsic versus extrinsic organization-based self-worth conflicts, provision versus self-sufficiency, organizational exploitation versus prosocial engagement, market dominance, and identity conflicts. Conflicts are initially external, rooted in the early history of the company, but over time they become internalized and thus develop into unconscious internal conflicts. Conflict represents the clash of opposing positions, the collision of motives, desires, needs, values, and ideas. Conflicts between competing needs are a normal and everyday occurrence, such as the need for self-determination in the workplace on one hand and the desire for leadership and guidance on the other. If the underlying conflict is adequately resolved, individuals within the organization can shift flexibly between these needs depending on the situation. However, stress or dysfunction arises when the organization fails to develop an adequate strategy to resolve the fundamental conflict and persistently tries to rigidly implement one pole of these needs. These conflicts are repeatedly triggered by situational factors throughout the organization's development.

Research Objectives

SyMOA has been developed from a theory-driven perspective and has been refined through preliminary applications in 2 SMEs. To date, SyMOA has not been applied outside of its developmental testing. Furthermore, the psychometric properties and methodological robustness have not yet been empirically evaluated. There is still a lack of evidence on the instrument's reliability, validity, feasibility, and acceptability. For SyMOA to become established as a scientifically sound diagnostic instrument and to be used confidently in organizational practice, this gap needs to be closed. Specifically, the study examines (1) the intercoder reliability of the SyMOA rating system; (2) the construct, content, and external validity of the instrument; and (3) the feasibility and acceptability of its interview and coding procedures in organizational practice. The present study therefore constitutes the first systematic evaluation of SyMOA's reliability and validity.

Methods

Ethical Considerations

The study was approved by Ethikkommission Universitätsklinikum Freiburg on March 3, 2025, under the trial ID 24-1561-S2. The study is being conducted in accordance with the local legislation and institutional requirements. The participants will provide their written informed consent to participate in this study. When presenting data from the interviews, particular care will be taken to ensure that published statements—especially those concerning conflict situations—do not negatively impact individuals or the company. Consent for publication is included in the informed consent process. The trial was registered with the Freiburger Clinical Trials Register (Freiburger Register Klinischer Studien; trial ID FRKS005508) on December 17, 2024.

Participants' privacy and confidentiality will be fully respected. No identifying personal information will be collected. Interviews will focus exclusively on participants' perspectives on organizational dynamics and working relationships. While participants may refer to personal experiences, all data will be analyzed and reported in an anonymized form.

Audio recordings will be stored without personal identifiers using encrypted storage systems and will be accessible only to the research team. The recordings will not be shared with third parties and will not be publicly available. No images or visual materials of participants will be collected or included. The study does not involve patients or clinical data.

No financial or nonfinancial compensation will be provided to participants for taking part in the study.

Objectives, Design, and Setting

The aim of the study is to determine intercoder reliability (Krippendorff α) and validity (construct, content, and external validity) of the newly designed SyMOA instrument. The study follows an exploratory qualitative multicase research design aiming to recruit at least 3 SMEs.

Study Population

The study focuses on SMEs in the industrial sector in Germany. Organizations with fewer than 30 employees are excluded. Participants within the selected organization must meet the following criteria: (1) aged above 18 years, (2) with the company for more than 1 month, and (3) proficient in German. Employees who do not meet these criteria will be excluded. Factors such as age, gender, and social status will not be a priority in the selection, as all employees in the respective critical department of the company will be approached. Sociodemographic data will be documented.

Study Procedure

The study flow is illustrated in Figure 1.

The first phase will be dedicated to training of the researchers and coders in the interview guide and rating criteria. The second phase will focus on recruitment of interested organizations and obtaining informed consent from the organization's management and participating employees. This will be followed by data collection, including semistructured interviews supplemented by structured on-site observations documented in a research diary (logbook with templates). If necessary, follow-up data collection may be arranged with the organizations. Three researchers will independently code the interviews using the audio recordings and the transcripts using MAXQDA24 (VERBI Software) [14]. The final phase will be dedicated to data analysis and reporting, culminating in oral and written feedback to the

XSL•FO
RenderX

participating companies, including recommendations for potential interventions.

Figure 1. Overview of the study procedure. SyMOA: Systematic Multidimensional Organizational Assessment.

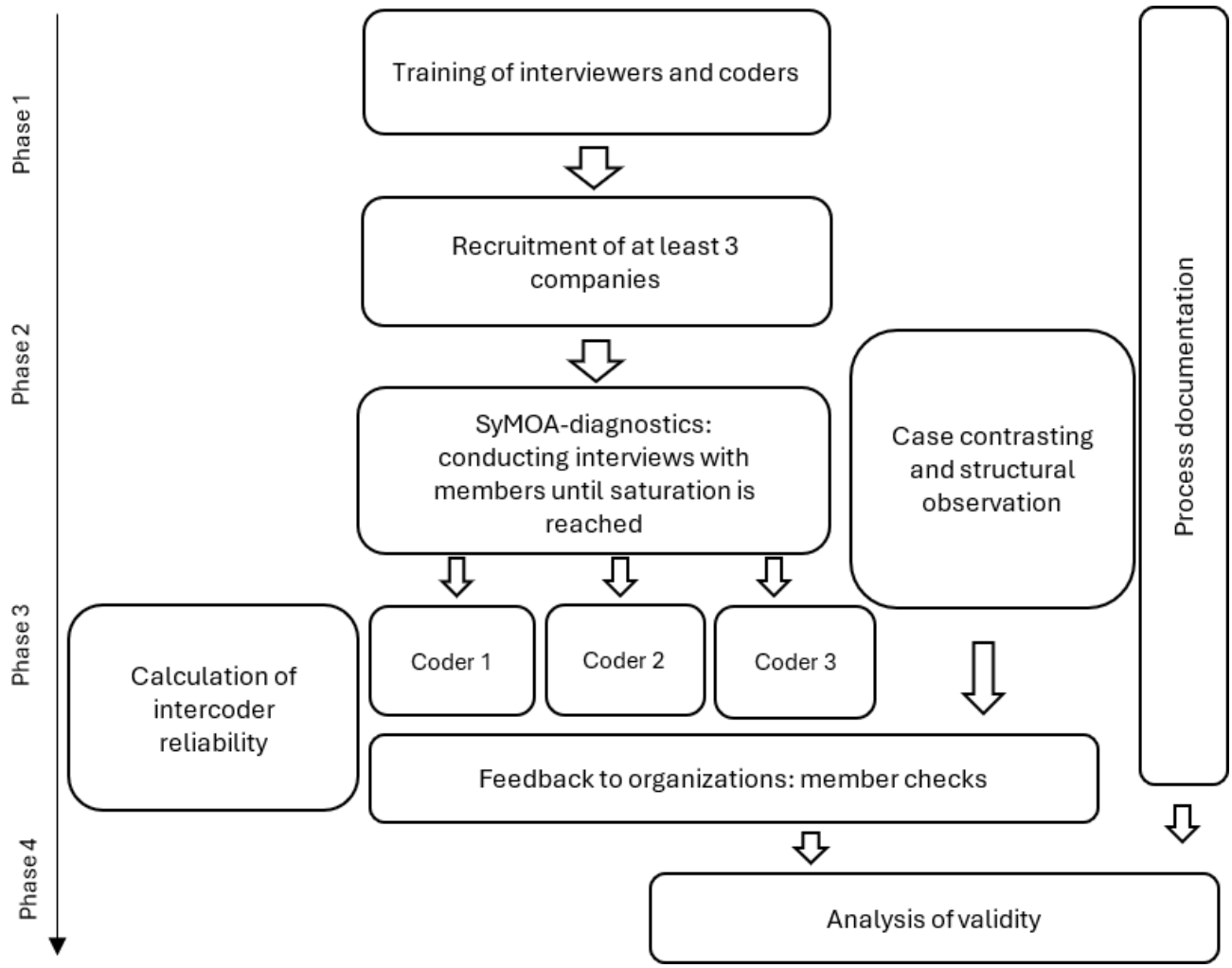

Training of Raters

All interviewers and raters are qualified psychologists. The training was carried out by 1 of the writers of the SyMOA framework. The training contains 2 sessions of 2.5 hours each and interim exercises. The first session focuses on imparting the theoretical buildup of the interview guide, the questioning process, operationalization of the constructs, and their rating. The basis is a logical understanding of the content, followed by an exploration of psychological meaning through the examination of contradictions and experiential context. A scenic understanding to assess interactional patterns and organizational culture, culminating in the interpretation of unconscious elements will be trained. While this interpretative process remains open to emergent themes, it is guided by the predefined dimensions of SyMOA and incorporates interpretative phenomenological analysis [15] to systematically capture the organizational dynamics. The raters will practice with each other in this session. After the first training, the raters will receive audio tapes of prerecorded interviews, which they will independently rate and bring to the second session. During the second session, the ratings will be evaluated, and further questions and uncertainties will be discussed. Furthermore, the interviewers will conduct a training interview, which will be audio-recorded and rated by all involved researchers of the team for feedback.

Recruitment

Three companies will be recruited through direct contact and researcher networks. At least 10 employees per company, representing different hierarchical levels and workspaces, will be asked to participate. The recruitment process begins with the chief executive officers or human resources management, followed by direct outreach to employees through personal contact, email, and public notice on boards. Management will facilitate initial contact, but participant information will remain confidential. Recruitment materials include the project's homepage, a white paper, and an information sheet. Participants will receive detailed information about the study, including potential risks and their right to withdraw at any time without consequences. Additionally, participants are informed about data security and their participant rights. Study personnel will then obtain written informed consent.

Data Collection

The interview will be conducted (and analyzed) in German with each of the three interviewers taking turns. The other interviewers will not be present during the interview but will

XSL•FO
RenderX

rate it independently afterward. The interviews will be audio-recorded only. This decision was made in consideration of data protection requests of participating companies and concerns regarding identifiability of employees. Nonverbal and interactional cues will nevertheless be systematically captured by the interviewer through structured observer notes in the research logbook. In order to assess the influence of the lack of some nonverbal cues (gestures, facial expressions, body language, etc) on the rating of the interviews, the coding of the interviewer will be compared for significant discrepancies with the coding of the two independent nonpresent raters.

Data collection will continue until at least 10 interviews have been conducted, and afterward until information saturation is reached, defined as the point where no new themes emerge and the evaluation process between interviews becomes predictable. Theme saturation is considered achieved when the same themes are repeated in at least 2 consecutive interviews without the emergence of any new relevant themes.

Materials

Interview Guide

The interview is based on a semistructured interview guide previously constructed and tested for feasibility. The interview guide has been designed to facilitate a flowing, natural conversation, which lasts about 1 hour. The guide is structured into 5 sections. The first section focuses on the current challenges in the organization, their development and impact as well as resources and barriers for overcoming the challenges. The second section focuses on relational patterns and dynamics in the organization, particularly on how organizational challenges are influenced by these patterns. The third section focuses on the historical development of the organization including insights into the impact of important figures and events in its history. The final section focuses on the level of inner functioning. This encompasses, among other things, the organization's ability to reflect on itself, deal with critical situations, feedback culture, communicational capabilities, the ability to form healthy and steady working relationships, and their flexible use of adaptive and functional defense mechanisms. The interview guide not only includes the specific questions, but also provides contextual information and what to observe during the interview. The interviewees are encouraged to describe their perspectives and experiences in the organization through examples of interactions and situations.

Scoring Sheets

The coding sheets contain 124 coding units that need to be coded for each interview. The coding units are sectioned in respect to the main dimensions of the SyMOA construct. Most coding involves ratings on 3- or 4-point scales. If the raters feel that they do not have sufficient information for the rating of the coding unit, they have the option to leave the coding blank, and it will be treated as missing data in the analysis.

Coding Book

All coding units are detailed in a coding book with detailed information on dimension category, definition of coding unit, inclusion and exclusion criteria, detailed information on rating criteria, and anchor examples. The coding book will be prepared in advance. The coding of the first interviews will be used to go through all ratings of the three raters, discuss discrepancies in coding, and clarify and refine the coding criteria. After this initial process, the coding book will remain fixed and is the basis for the coding of all subsequent interviews in the study.

Data Analysis and Measures

Each dimension has a scoring sheet for evaluation and documentation (see the Materials section). For the quantifiable constructs, the scoring sheets list the operationalized items connected to the questions and have options for ratings on a scale from 0 to 3. Furthermore, the scoring sheets have free text spaces for more descriptive evaluations. Each of the three raters will code and analyze the interview independently, and results will not be discussed with each other until after the data collection is completed. Each interview is analyzed consecutively. After completion of data collection, the scoring sheets for each interview will be analyzed for intercoder reliability and the final analysis data will be used to evaluate SyMOA's validity. Furthermore, the interviewers as well as the interviewees will be asked to give feedback on the feasibility and acceptability of the interview through a short feedback survey.

Intercoder reliability will be tested by coding agreement of the three raters using Krippendorff α in MAXQDA24. This statistical measure will assess coding consistency across evaluators, ensuring the robustness of the SyMOA framework in capturing organizational dynamics. In order to assess the influence of the presence during the interview and being part of the interview interactional dynamic, the coding of the interviewer will be compared with the coding of the two independent raters. This comparison is an exploratory concordance analysis and is not a part of the primary interrater reliability estimate. The analysis levels will be the following:

- Primary intercoder reliability:
  - Alpha (Rater1, Rater2) → Intercoder Reliability of two independent coders as the primary index ($\alpha \geq .80$, good; .67-.79, acceptable)
- Secondary bias analysis and consistency exploration:
  - Alpha (R1, Interviewer) → Bias-Analysis 1
  - Alpha (R2, Interviewer) → Bias-Analysis 2
  - Alpha (R1, R2, Interviewer) → Consistency (explorative)

To ensure methodological rigor, the study evaluates various forms of validity [16].

So far, there is no existing gold standard for an operationalized organizational psychodynamic instrument in Germany; therefore, *construct validity* will be evaluated through theory-based and expert-informed approaches rather than through criterion validity. It is examined by evaluating whether the dimensions operationalized in SyMOA are representative and encompassing of systems-psychodynamic organizational theory and diagnostics. This is done through theoretical coherence validation, construct representation, and known-groups validity. *Theoretical validity* is established by

XSL•FO
RenderX

assessing whether SyMOA's findings align with or contribute to existing theoretical frameworks in organizational diagnostics. This will be done by a critical discussion of SyMOA's results on the background of the current literature on systems-psychodynamics and how it is in line with this literature or may even further develop and contribute to the literature. Furthermore, there will be a critical discussion of SyMOA in relation to other established organizational diagnostics, how it contributes to organizational diagnostic theory, and the merits and the limits of SyMOA in comparison to other organizational diagnostics. The first step is a top-down theoretical mapping, comparing the operationalized SyMOA dimensions with core systems-psychodynamic concepts (eg, defense mechanisms). A mapping matrix will be used to document conceptual matches, partial matches, and conceptual gaps.

The second step is a bottom-up *construct representation* analysis following a combined framework from Messick [17,18] unified validity theory and qualitative validity approach by Whittemore et al [19] and Maxwell [20]. Emergent patterns identified in the coded interviews will be grouped into themes and compared with established theoretical constructs of SyMOA to determine whether the material aligns with or diverges from the framework.

The third step of theoretical validity assessment will be a comparative analysis of frameworks [21,22]. Instruments of well-established classical organizational assessment, systems-psychodynamic organizational assessment, and qualitative interview methods will be chosen for comparison. A matrix including key concepts and aspects representing systems-psychodynamics (eg, interactional patterns, conflicts, defense mechanisms, etc), as well as standardization, objectivity, and quantification, will be used to rate each instrument on whether it captures, partially captures, or does not capture these aspects.

Furthermore, *known-groups validity* as an additional form of construct validity will be applied. Based on theoretical assumptions, a valid instrument should be able to differentiate between groups that are expected to exhibit distinct psychodynamic profiles (eg, inner level of functioning or teams with differing conflict levels). Therefore, the SyMOA score profiles for the 3 organizations will be compared according to their ability to capture the theoretically predictable differences across these groups. Evidence of differences consistent with theoretical assumptions may be interpreted as discriminant support for construct validity.

Expert validity as a source of evidence for *content validity* will be assessed through feedback from at least five specialists in systems-psychodynamic organizational research, who will evaluate whether SyMOA adequately captures relevant organizational dynamics. The limited number of experts reflects the scarcity of specialists in this field in Germany. Experts will be given the conceptual background of SyMOA, the interview guide as well as the rating sheets. They will be asked to fill in a feedback form, asking among other things about the appropriateness of the interview guide questions, the understandability, the completeness with which the questions represent the construct, the order of the questions, and if any important questions are missing. The feedback from the experts will be used to revise the interview guide.

External validity is tested by assessing how well SyMOA's diagnostic results correspond to measurable organizational factors such as employee turnover, sick days, and conflict rates (reported complaints or conflict cases to human resources, at the leadership level, or to an ombudsman). For analysis of correspondence to these data, the interview data will be grouped into thematic categories given by SyMOA's construct structure (eg, "organizational regulational capscities," "communicative capacities," "organizational relationship pattern," etc). Frequencies or intensity values will be assigned according to how often a topic was mentioned by each interviewee and how strongly it was rated, thus creating ordinal scaled data. Sick days, conflict rates, and fluctuation will be asserted on the same level of analysis (department and team). A Spearman correlation will be calculated with the 2 data sets. A moderate correlation and argument for external validity would be a Spearman ρ between 0.3 and 0.5.

Process validation is ensured through detailed documentation of data collection, analysis, and interpretation procedures to ensure reliability and transparency of the procedure. The documentation system consists of a logbook with a template for the coding process, a template for reflection processes, and a template for observational data.

Communicative validation, or member checking, is incorporated by inviting interview participants to review and confirm interpretations of their interview data to ensure that the conclusions drawn are valid. This is done after the SyMOA analysis is finalized.

Data triangulation is applied by combining multiple data sources, including interviews and structured on-site observations, to strengthen construct validity. Observational data will include direct observations of employees in their natural work environments and detailed field notes documenting organizational behaviors and interactions. The documentation of the observational data is also based on the 3 dimensions and constructs of SyMOA. However, direct observation focuses more on the nonverbal, unconscious, interpersonal, and situational aspects that can be less well-expressed during interviews. These observations will serve as supplementary contextual information for the analysis and are documented in the logbook.

Assessment of Feasibility and Acceptability

Feasibility will be assessed by keeping a recruiting and implementation log including a recruitment feasibility protocol (number of contacted organizations and participants who accepted or declined, with reasons if possible), an interview feasibility protocol (no-show rate to interviews, average interview duration, disturbances, comprehension problems, and completeness of interview implementation) and a coding feasibility protocol (data completeness, technical feasibility [incidents that prevented implementation of interview or rating by all raters], reachability of saturation, problems with coding instructions, and ambiguity in coding).

XSL•FO
RenderX

Acceptability will be assessed on the level of the organization, the participants, and the interviewers. The interview participants will receive a short survey asking about the comprehensibility of the questions, how they experienced the atmosphere during the interview, the overall load of the interview, and the appropriateness of the questions for understanding the organization and its challenges. Acceptability on the level of the organization will be assessed by member check and feedback of results to the organization. Finally, the acceptability on the level of the interviewers will be assessed by a rating of the interviewers after each interview and a rating of the feasibility of conducting the interview in the given timeframe, whether the questions were sufficient to allow a deeper understanding of the interviewee's experience, if normal conversational flow was possible, and if they could ask all questions.

Data Management and Data Protection

In this study, personal data will be collected in the form of audio recordings of interviews, signed informed consent forms, and interview transcripts. These do not include health-related information, nor will any sensitive internal company data be requested. The audio recordings will be stored password-protected, with names and company identifiers removed. However, complete anonymization is not possible, as voices and interview content may still allow identification. The transcriptions will be pseudonymized so that they can only be attributed to the participants via an externally created ID list. All data will be stored on a university hospital server that complies with General Data Protection Regulation. Participation is voluntary, and consent for data processing can be withdrawn at any time without providing a reason or facing any consequences. Upon withdrawal, no further data will be collected.

Dissemination Plans

The results of the study will be made public using channels such as peer-reviewed publication and presentation in national and international conferences. The participating organizations will receive oral and written feedback on the results of the interviews with particular care that the feedback is not traceable to the individual interview partner, that is, employees of the organization. The study results will be published in a scientific journal in a timely manner.

Results

Recruitment and interviewer or coder training began in May 2025 and are expected to be completed by July 2025. Data collection is scheduled to take place between September 2025 to February 2026, followed by data analysis and reporting from February 2026 to April 2026. As this is a study protocol, no study results are available at this stage.

Discussion

Principal Findings

This study protocol describes a qualitative validation study of the SyMOA, a systems-psychodynamic diagnostic tool designed to identify organizational challenges, inner functioning, relational dynamics, and unconscious conflict patterns. The planned study is expected to enhance our knowledge of the feasibility of administering SyMOA in organizational settings, the validity of the semistructured interview, and the consistency with which different coders can apply the coding framework. Although no results are available yet, the study is designed to examine whether the operationalized dimensions of SyMOA are meaningful and can be reliably identified by different raters in different organizational contexts.

Comparison to Prior Work

Most existing assessments of the organizational context focus on directly observable concepts, such as structures, climate, or leadership, but rarely integrate deeper levels of analysis derived from an understanding of unconscious dynamics, relational patterns, and inner-systemic functioning. Systems-psychodynamic research emphasizes the relevance of these aspects, yet there is still a lack of validated diagnostic instruments. SyMOA contributes to this gap by picking up concepts from a clinical instrument (Operationalized Psychodynamic Diagnostics III) and further developing them into an operationalized system that enables a psychodynamic analysis of organizational systems using a semistructured interview guide and a structured coding and rating protocol grounded in systems-psychodynamic theory. This study aims to advance psychodynamic organizational assessment by systematically examining how well such themes and content can be coaxed through a semistructured interview and reliably coded.

Strengths and Limitations

Strengths of the study include the application of an innovative diagnostic approach based on established psychological diagnostic methods, ensuring both reliability and practical relevance. A further strength of this study protocol is the implementation of multiple validity assessments—including content, construct and external validity, process validation, communicative validation, and researcher triangulation. The validity assessments are complemented with intercoder reliability testing. Observer notes and on-site observations enhance methodological rigor. The use of structured rater training, rating protocols, and a detailed coding book further enhances transparency and reproducibility.

However, certain limitations must be acknowledged. While the study aims to provide evidence, the reliance on qualitative interview data introduces the possibility of subjective interpretation, which data and researcher triangulation are designed to limit. Another limitation is the extent to which interpretative qualitative data can be contextualized with external quantitative data to assess the external validity of the construct. Furthermore, the determination of external validity will be limited as the assessment relies on correlations with staff turnover, sick days, and conflict data. However, these indicators may be influenced by various external factors (eg, economic conditions and industry trends) that are not accounted for in the analysis. In addition, a potential selection bias in the recruitment process for the interviews could occur if certain characteristics of the employees (eg, attitudes, greater self-reflection, or time constraints) influence their willingness to participate. With

XSL•FO
RenderX

regard to the interview data, a potential resistance or social desirability bias can be expected due to fear of repercussions or a desire to present the organization in a positive light. The use of audio recordings only as the basis for the coding process limits access to nonverbal cues that may be important for accurate interpretation of the interviews. These limitations will be addressed through triangulation, structured observer documentation, and consistency protocols, and a bias analysis between the interviewer coding and the two independent raters. This analysis is not included in the primary intercoder reliability analysis, but the comparison can reveal systematic differences. Such differences possibly reflect richer contextual understanding when present in the situation. Nevertheless, the limitations remain important considerations for interpreting the findings. Finally, the study will collect data from 3 organizations which may restrict generalizability.

Future Directions

The results of this study will guide refinement of the SyMOA interview guide and the coding framework. Future research may include application in diverse sectors, quantitative validation components, and integration with organizational outcome measures.

Acknowledgments

We used deepl.com and ChatGPT-5 for refinement, correction, and editing of the manuscript to improve the quality and clarity of language.

Funding

The study is funded by the University Medical Center Freiburg Department of Mental Health and Department of Psychosomatic Medicine and Psychotherapy. There is no external funding.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

CL: Contributed to conceptualization, methodology, and writing – review & editing
AMM: Contributed to conceptualization, background, methodology, writing – original draft, writing – review & editing, and project administration
YR: Contributed to writing – review & editing and methodology
SS: Contributed substantially to the preparation of ethics approval, writing – original draft, writing – review & editing, and methodology
Each author has made a substantial contribution to the work. All authors have approved the final manuscript. All authors agree both to be personally accountable for their own contributions and to ensure that questions related to the accuracy or integrity of any parts of the work, even ones in which the author was not personally involved, are appropriately investigated and resolved and the resolution documented in the literature.

Conflicts of Interest

None declared.

References

1. Petriglieri G, Petriglieri JL. The return of the oppressed: a systems psychodynamic approach to organization studies. Acad Manag Ann 2020 Jan;14(1):411-449. [doi: 10.5465/annals.2017.0007]
2. Wolfe N. The Living Organization: Transforming Business to Create Extraordinary Results: Quantum Leaders; 2011.
3. McFillen JM, O'Neil DA, Balzer WK, Varney GH. Organizational diagnosis: an evidence-based approach. J Chang Manag 2013 Jun;13(2):223-246. [doi: 10.1080/14697017.2012.679290]
4. Stouten J, Rousseau DM, De Cremer D. Successful organizational change: integrating the management practice and scholarly literatures. Acad Manag Ann 2018 Jun;12(2):752-788. [doi: 10.5465/annals.2016.0095]
5. Möller H, Giernalczyk T, Hinn D, et al. Individual and collective defense mechanisms. In: Greif S, Möller H, Scholl W, editors. International Handbook of Evidence-Based Coaching: Springer International Publishing; 2022:459-470. [doi: 10.1007/978-3-030-81938-5_37]
6. Beer M. Organization Change and Development: A Systems View: Scott Foresman; 1980.
7. Kanter RM, Stein B, Jick T, editors. The Challenge of Organizational Change: How Companies Experience It and Leaders Guide It: Simon and Schuster; 1992.
8. Kets de Vries M. Organizations on the couch: a clinical perspective on organizational dynamics. Eur Manag J 2004 Apr;22(2):183-200. [doi: 10.1016/j.emj.2004.01.008]

XSL•FO
RenderX

9. Mikkelsen EN, Gray B, Petersen A. Unconscious processes of organizing: intergroup conflict in mental health care. J Manage Stud 2020 Nov;57(7):1355-1383. [doi: 10.1111/joms.12611]
10. Obholzer A, Roberts VZ, editors. The Unconscious at Work: A Tavistock Approach to Making Sense of Organizational Life, 2nd edition: Routledge; 2019. [doi: 10.4324/9781351104166]
11. Arbeitskreis OPD, editor. OPD-3 - Operationalisierte Psychodynamische Diagnostik: Das Manual Für Diagnostik Und Therapieplanung [Book in German], 1st edition: Hogrefe AG; 2023. [doi: 10.1024/86263-000]
12. Both LM, Bastos AG, Freitas LHM. Operationalized psychodynamic diagnosis: a systematic review of the literature. Trends Psychiatry Psychother 2019 Jul 10;41(2):201-210. [doi: 10.1590/2237-6089-2018-0020] [Medline: 31291412]
13. Kontoghiorghes C, Awbre SM, Feurig PL. Examining the relationship between learning organization characteristics and change adaptation, innovation, and organizational performance. Hum Resour Dev Q 2005;16(2):185-212. [doi: 10.1002/hrdq.1133]
14. Qualitative inhaltsanalyse software [Website in German]. MAXQDA. URL: https://www.maxqda.com/de/software-inhaltsanalyse [accessed 2025-04-02]
15. Biggerstaff D, Thompson AR. Interpretative phenomenological analysis (IPA): a qualitative methodology of choice in healthcare research. Qual Res Psychol 2008 Sep 25;5(3):214-224. [doi: 10.1080/14780880802314304]
16. Mayring P. Einführung in Die Qualitative Sozialforschung [Book in German]: Beltz; 2016.
17. Messick S. Validity. In: Linn RL, editor. Educational Measurement, 3rd edition: American Council on Education and Macmillan Publishing Co, Inc; 1989:13-104.
18. Messick S. Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. Am Psychol 1995;50(9):741-749. [doi: 10.1037/0003-066X.50.9.741]
19. Whittemore R, Chase SK, Mandle CL. Validity in qualitative research. Qual Health Res 2001 Jul;11(4):522-537. [doi: 10.1177/104973201129119299] [Medline: 11521609]
20. Maxwell JA. A Realist Approach for Qualitative Research: SAGE; 2012.
21. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, editors. Analyzing Qualitative Data: Routledge; 1994:173-194. [doi: 10.4324/9780203413081_chapter_9]
22. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117. [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]

Abbreviations

SME: small- and medium-sized enterprise
SyMOA: Systematic Multidimensional Organizational Assessment

Edited by J Sarvestan; submitted 15.May.2025; peer-reviewed by RCS Kay; revised version received 03.Dec.2025; accepted 04.Dec.2025; published 05.Mar.2026.

Please cite as:
Müller AM, Sachs S, Rieder Y, Lahmann C
Validity and Reliability of the Psychodynamic Organizational Diagnostic Instrument SyMOA: Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e77554
URL: https://www.researchprotocols.org/2026/1/e77554
doi:10.2196/77554

© Anne-Maria Müller, Sophia Sachs, Yannik Rieder, Claas Lahmann. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

International Case Studies to Identify Success Factors and Contextual Conditions in the Digital Transformation of Health Care Systems and Derive Lessons for Germany: Study Protocol for a Mixed Methods Study

Lena Kraft[1*], MA; Anna-Lena Brecher[1*], MSc; Sophia Sgraja[1], MA; Reinhard Busse[2], Prof Dr; Volker Eric Amelung[1], Prof Dr

[1]Hannover Medical School, Institute for Epidemiology, Social Medicine, and Health Systems Research, Carl-Neuberg-Straße 1, Hannover, Germany

[2]Department of Health Care Management, Technische Universität Berlin, Berlin, Germany

[*]these authors contributed equally

Corresponding Author:
Lena Kraft, MA
Hannover Medical School, Institute for Epidemiology, Social Medicine, and Health Systems Research, Carl-Neuberg-Straße 1, Hannover, Germany

Abstract

Background: Germany's health care system continues to face significant challenges in its digital transformation due to outdated structures, interoperability issues, strict data protection regulations, and low user acceptance, despite numerous legislative initiatives, such as the Digital Care Act in 2019, which was intended to promote practical use and innovation. In contrast, several international health care systems have successfully advanced their digital transformation, offering valuable insights and potential lessons for the German health care system.

Objective: This study, as part of the research project "NADI: Benefits and Acceptance of Digital Health," analyzes international health care systems to identify key success factors and develop pragmatic recommendations for German policymakers to enhance the country's digital health implementation.

Methods: This study uses a mixed methods triangulation approach, combining case study selection, qualitative expert interviews, and a quantitative online survey to develop actionable policy recommendations for the digital transformation of health care in Germany. The study applies the conceptual framework of tipping points and success factors to identify critical factors in the digital transformation of health care systems, where certain actions or conditions fundamentally influence adoption and success. A total of more than 100 interviews were conducted with experts representing 8 stakeholder groups from 9 different health care systems. The qualitative data are evaluated using qualitative content analysis according to Kuckartz and Rädiker. In an online survey, a minimum of 305 participants from the German health care system will be surveyed regarding the relevance and feasibility of the key success factors identified in the international case studies. The dataset will be analyzed statistically using SPSS, both descriptively and inferentially (eg, subgroup analyses).

Results: Between November 2024 and September 2025, interviews with international health care experts were conducted. As of October 2025, the qualitative content analysis is still ongoing. The recruitment phase for the online survey is planned from October 15 to December 15, 2025. Initial results are expected to be available in 2026. The study protocol was submitted during the qualitative data collection phase before the commencement of the quantitative survey. Analysis had not yet begun at the time of submission.

Conclusions: The use of a case study methodology has been demonstrated to facilitate the acquisition of invaluable insights into international best practices, while concurrently offering the opportunity to identify specific success and failure factors. The integration of qualitative expert interviews serves to contextualize international findings on tipping points and success factors in the implementation and use of digital health tools. The transfer of the international results to the German context represents a central component of the research project, which aims to investigate practical implementation. The combination of these approaches forms a comprehensive basis for deriving specific recommendations for action for the German health care system.

International Registered Report Identifier (IRRID): DERR1-10.2196/80301

(JMIR Res Protoc 2026;15:e80301) doi:10.2196/80301

XSL•FO
RenderX

KEYWORDS

digital health; international health care systems; case study research; success factors; tipping points

Introduction

Germany is encountering challenges in its digital transformation, particularly in the health care sector. For instance, the electronic patient file was legally integrated into the German Social Code (Sozialgesetzbuch) as early as 2004 [1]. This was followed by a period of 20 years characterized by numerous legislative initiatives, such as the resolutions on the E-Health Act in 2015 to expand the telematics infrastructure or the Digital Care Act in 2019, which was intended to promote the practical use and innovation of digital health technologies [1,2]. It was not until March 2024 that the Act to Accelerate the Healthcare System came into force as part of the Federal Ministry of Health's digitalization strategy, in which the establishment of the electronic health record as a central component is now being driven forward [2]. Thus, there was clearly no shortage of political initiatives [3]. Rather, the reasons for the weak implementation of digital health in Germany seem to be attributable to a number of factors, including the heterogeneity of information technology (IT) systems and the lack of interoperability [4], which are exacerbated by the fragmented health care system in Germany. In addition, significant public concerns regarding unauthorized data access and data protection issues are holding back the digital transformation [5]. Moreover, a dearth of acceptance among users, such as patients or physicians, is frequently attributable to a paucity of skills [3].

Other European countries, including Denmark and Estonia, as well as countries in more distant regions, such as Israel, have demonstrated a high level of development in digital structures in the health care sector compared to Germany and face fewer challenges in the areas described above and beyond [6]. For example, Estonia and Denmark are focusing on a central platform for health data, ease of use, and public trust in digital health tools [7]. Israel, on the other hand, is strongly committed to start-ups through the provision of grants, tax incentives, and favorable regulations [8]. Nevertheless, the digital transformation within the German health care system is already underway. Electronic health cards, electronic prescriptions, and reimbursable digital health applications, among other digital components, have been implemented and are being used to varying degrees [9]. To exploit the full potential of this transformation, the right framework conditions are needed, such as education and incentive mechanisms. It is important to identify success factors in this context to improve conditions and accelerate the implementation and use of digital tools.

This study constitutes a component of the research project *NADI: Benefits and Acceptance of Digital Health*, which is funded by the German Innovation Fund. The project involves the analysis of international experiences, options for action, and patient preferences. This part of the study aims to identify success and failure factors, as well as tipping points, in the digital transformation of international health care systems. By examining a selection of international case studies and best practices, this study will uncover insights into what contributes to successful digitalization in health care, for example, success factors with regard to regulation or financing. The core findings will then be evaluated for their transferability to the German health care system. Considering the findings from all project components of the other project partners, a series of pragmatic recommendations is set to be developed. These recommendations are designed to assist German policymakers in effectively overseeing the ongoing transformation toward digital health. Furthermore, these recommendations are intended to ensure that the requisite measures are implemented to ensure a successful and sustainable digitalization process.

The aim of this study is therefore to identify success and failure factors in the process of digitalization of health care systems and to analyze these in the context of the individual conditions and structures of the respective health care system. Finally, recommendations for action to promote digital transformation in the German health care system will be derived. The focus is not on specific technological innovations or applications but rather on the political and sociocultural processes that have set the course for successful digitalization in health care, adopting a rather broad perspective on digital health. The subsequent research questions are as follows.

- What are the relevant success and failure factors in the digital transformation of health care systems?
- What are the relevant tipping points in the digital transformation of health care systems?
- Which success factors from other countries in the digital transformation of health care systems can be transferred into recommendations for action for the German health care system?

Methods

Study Design

The study follows a mixed methods triangulation approach. It includes case studies, comprising extensive literature research and semistructured expert interviews, as well as a quantitative online survey. In a concluding stage, the results of the analysis of the interviews and the quantitative data will inform the development of practical recommendations and measures for health policy regarding the digital transformation of the German health care system (Figure 1). The study was developed with the support of the scientific advisory board. Patients were not involved in the design, development, or dissemination plans of the research.

XSL•FO
RenderX

Figure 1. Study design. This figure shows the flowchart of the project process in 2024-2026.

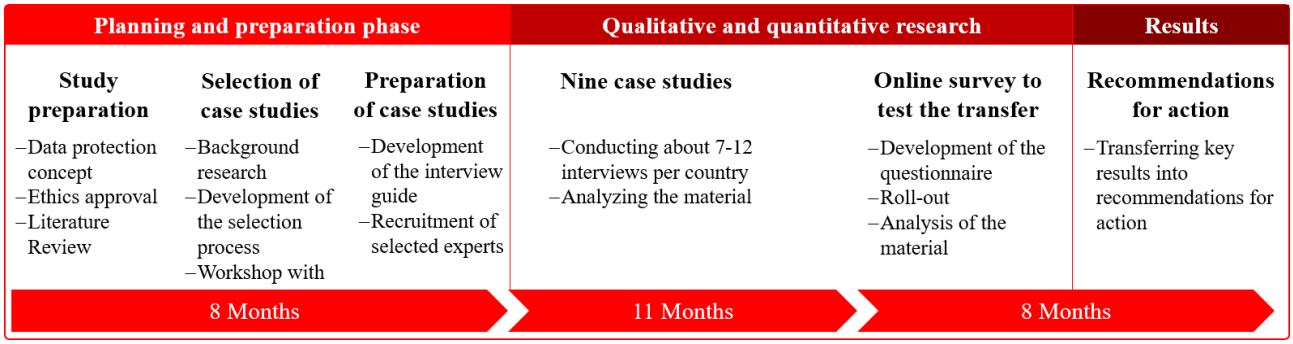

Success Factor Research

The research of success factors constitutes a subdiscipline of business administration. Its purpose is to identify the factors that have contributed positively or negatively to the success of a company. According to Steiner [10], "Strategic factors refer to an action, element, or condition which for a business may be of critical importance in its success or failure." In this study, we will apply success factor research to a different research discipline, in which the focus is not on individual companies but on entire health care systems [10]. Success in this case is the successful implementation of key digital health tools, characterized by a high application rate and user acceptance, as well as a positive impact of the tools on care and communication processes in the health care system.

Theory of Tipping Points

Another theory that is being addressed as part of this research project is the concept of tipping points. This theory was popularized by the book *The Tipping Point* by Malcolm Gladwell, first published in 2000 [11]. Since then, it has been increasingly used in many different research disciplines [12,13]. The tipping point is defined by Gladwell [11] as "that magic moment when an idea, trend, or social behavior crosses a threshold, tips, and spreads like wildfire." In this study, the concept of tipping points is used to identify critical phases in which the digital transformation in the health care systems examined has been fundamentally influenced. Triggers could include certain actions, such as political strategies, the implementation of particular technologies, legal changes, adjustments to provider remuneration, or public campaigns, as well as overarching conditions, such as changes in governance or the occurrence of a pandemic. Recognizing these pivotal moments is crucial to identifying the catalysts and analyzing them as potential success factors.

Case Studies

To identify success factors and tipping points of digital health transformation, it is essential to consider the specific context, for example, the conditions of the respective health care system. According to Yin [14], a case can be defined as "a contemporary phenomenon within its real-life context, especially when the boundaries between a phenomenon and context are not clear and the researcher has little control over the phenomenon and the context". Therefore, the case study approach is well suited to systematically gathering international experiences regarding the digitalization of health care systems.

This study uses an exploratory case study approach, which is 1 of 3 possible approaches. However, the boundaries between descriptive, explanatory, and exploratory research are flexible and overlap to some extent [15]. In addition to Yin, Merriam and Stake [15,16] have particularly shaped and further developed the case study research approach. The approach applied in this study is primarily based on the recommendations of Merriam, which are associated with the constructivist case study paradigm. In accordance with Merriam's recommendations, the case study design and research questions are based on a literature review. The qualitative data collection for this study is based on interviews and the analysis of documents. As Merriam describes it, the analysis of the data is "the process of making meaning" [11] and is approached through qualitative content analysis [9]. Merriam further recommends various strategies for validating the results of case study research in terms of internal and external validity, such as triangulation and exchange with experts, as well as a very comprehensive description of the cases, as is also implemented in the context of this study [9,10].

The following section delineates the steps involved in conducting the case studies and the planned evaluation.

Case Study Selection

The selection of countries for case studies is often perceived as arbitrary, and there is currently insufficient research on this approach. Nevertheless, a reasoned and transparent approach is necessary to reduce bias in case selection and to maximize knowledge gain [17]. In this study, 9 international case studies were identified through a multistage, criteria-driven process (Figure 2). The study uses the method of diverse case study selection, which is particularly recommended for exploratory research approaches. An inherent limitation of the case study approach is its lack of representativeness. Nevertheless, the selection of cases that demonstrate a wide range of characteristics with respect to relevant variables can enhance representativeness by reflecting the diversity of the overall pool [12].

XSL•FO
RenderX

Figure 2. Case study selection. This figure shows the 3-stage process of case study selection.

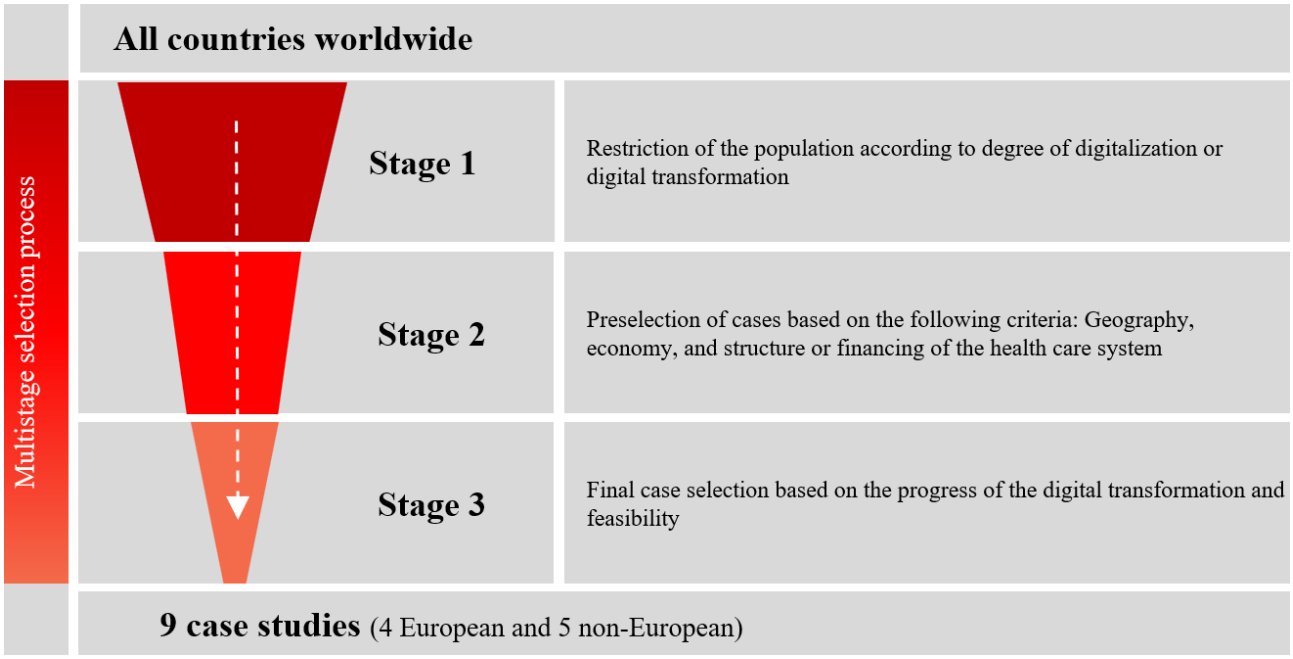

The selection process for the case study countries was conducted in 3 stages, following an iterative approach. The first stage involves narrowing down the initial pool, encompassing all health care systems worldwide, based on their degree of digitalization, thereby including both countries with well-established digital health systems and those that have undergone significant digital transformation in recent years. To assess the level of digitalization, several international indices were consulted, including the World Digital Competitiveness Ranking 2023, the Global Digital Health Monitor (latest update in 2023), the Digital Economy and Society Index 2022, the Digital Riser Reports 2020 and 2021, and the Bertelsmann Foundation's 2018 study on international health care system comparisons with a focus on digitalization [6,18-22]. Relevant indicators include strategy and policy, workforce, standards and interoperability, mindset, and infrastructure. Additionally, expert opinions from the project's scientific advisory board were incorporated to account for recent developments in the digitalization of various health care systems.

In the second stage, a set of relevant criteria was selected to characterize the countries in the preliminary selection. These criteria were grouped into three overarching categories: (1) geography, (2) economy, and (3) health care system (Table 1).

This structured overview aimed to facilitate the selection of a diverse set of case studies. The goal was to ensure a broad spectrum of insights while maximizing the relevance and transferability of findings to the German health care system. The characterization of a short list of countries included in stage 2 can be found in Multimedia Appendix 1.

The final selection of countries in stage 3 was based on an in-depth assessment of a short list of countries identified as suitable in the previous stages. This assessment particularly considered the actual state of health care digitalization and recent developments in the implementation of key digital health tools. In addition, practical considerations, including access to scientific literature, availability of experts, political stability, and the logistical and financial feasibility of conducting the interviews on site, were taken into account. The selection process was further refined through expert discussions and insights gained from semistructured expert interviews. The 9 selected case studies are as follows (Figure 3): Estonia, Denmark, Poland, and Portugal (European countries); Israel, Japan, Saudi Arabia, South Africa, and the United States (specifically the United States Department of Veterans Affairs; non-European countries).

laGiven the complexity of the health care system in the United States, this case study will focus on the Veterans Affairs system, the largest integrated health care system in the country, providing universal health care to approximately 9.1 million eligible veterans [23,24].

XSL•FO
RenderX

Table . Case study characterization criteria. This table shows the criteria that were considered in stage 2 of the selection of the case studies.

Category and criteria	Indicators
Geography	
Geographic region	WHO[a] regions
Population	Number of inhabitants
Area	Area in km^2
Economy	
Economic development	High-income/low-income countries
Economic performance	GDP[b] in US $ per capita
Health care system	
Health care system typology	National health service/social insurance/private insurance
Centralization in health care governance	Centralized/decentralized
Health care expenditure	In US $ per capita

[a]WHO: World Health Organization

[b]GDP: Gross Domestic Product.

Figure 3. Map of selected case studies. This figure illustrates the selected case studies on a world map. Source: The authors' illustration created using Python code and the Folium library (Copyright 2013–2025 Rob Story) [25]. Geospatial data were obtained from OpenStreetMap (Copyright OpenStreetMap contributors; data licensed under the Open Database License version 1.0) [26].



Qualitative Survey

Background Research

Following the selection of case studies, a comprehensive background research phase was conducted. This phase focuses on characterizing the health care system of each selected country, identifying key institutions involved in health care delivery, and assessing the development and current state of digitalization in the health care sector. The research also identified relevant stakeholders, such as government agencies, health care providers, technology companies, and patient organizations, to understand their roles in the digital transformation process. This background information provides essential context for the case studies, ensuring a thorough understanding of the health care systems and their digitalization trajectories before conducting the interviews.

XSL•FO
RenderX

Conception of the Expert Interview

The conceptualization and planned reporting of the qualitative data collection are grounded in the COREQ ("Consolidated Criteria for Reporting Qualitative Research") guidelines (Checklist 1), thereby ensuring a systematic approach [19]. As the analysis is ongoing, not all the criteria in the checklist have been met yet. To generate specific information on the digital transformation in the selected countries, it was decided to conduct expert interviews. People involved in the context of digital health and who have relevant specialist knowledge are designated as experts. If the individuals have specific role knowledge or can be considered to have special expertise in the field of digital health, these experts are consulted. Using Helfferich's methodology, a structured framework of a semistructured interview guide was devised for the interviews with these experts, focusing on factual questions to elicit their expertise [20]. This approach ensures the collection of pertinent role-specific knowledge. In accordance with this methodology, interview sections are thematically constrained to enable experts to efficiently reproduce their knowledge, given the limited time available. Experts are permitted to respond in detail to specific domains of knowledge or to provide concise statements. The interviewer is assisted by the semistructured interview guide to present themselves professionally and competently, thereby reducing the power imbalance between interviewer and interviewee [20].

Development of the Semistructured Interview Guide

The structure of the interview guide was developed in accordance with the requirements and principles of guide development as outlined by Helfferich and Kruse [27,28]. The design of the semistructured interview guide was informed by these principles, with the objective of addressing the research interest in an appropriate manner, while maintaining a high degree of openness and structure. The 4 steps of Helfferich's procedure were used as a framework, encompassing collection, verification, organization, and integration [27].

In the initial step, a range of questions was compiled, which was then reviewed in the subsequent step with respect to their expressive potential and relevance in addressing the research queries. In the third step, pertinent questions were categorized according to their content. In addition, overarching modules and subcategories were formulated in this step. In the final step, narrative-generating impulses were assigned to the developed modules. The content of the semistructured interview guide is based on findings from previous research and workshops with the project's advisory board. The following categories have been developed and will guide the interviewer and interviewee through the interview: (1) tipping points, (2) governance, (3) incentives and sanctions, (4) technical regulations and structures of interoperability, (5) digital health tools, and (6) recommended actions.

In the inaugural module of tipping points, the expert is queried on the events that have precipitated the digital health transformation in the respective nation. This is followed by inquiries into the influence of the health system's structure on the transformation, incentive mechanisms, and potential sanctions. Additionally, questions regarding technical infrastructure and interoperability are posed. Ultimately, the implementation of digital health tools, including diagnosis-, system-, or prevention-oriented tools and tools for collaboration with other health care providers, is subjected to scrutiny. Finally, the conversation is concluded with a question regarding the key lessons for the digital transformation of a health care system. A series of documents was created for each country, with adaptations made to align with the stakeholder groups and the 4 distinct tools. This approach was adopted to ensure the collection of more precise and detailed information. However, to maintain the comparability of the interviews, the structure of the interview guide remained unaltered, with only minor adjustments made to the content. The organization of pertinent questions, the development of modules, and the allocation of narrative-generating prompts were subjects of repeated deliberations with other researchers, leading to successive revisions [27,28]. The basic interview guide for the interviews was subjected to rigorous testing in accordance with the prevailing standards of scientific research [29,30]. Initially, the semistructured interview guide was submitted for independent evaluation to 2 experienced methodologists specializing in qualitative research. Concurrently, 4 pretest remote interviews were conducted with experts from Portugal and Poland to assess the efficacy of the semistructured guide in a field setting. The comments of the 2 experts in qualitative methods and the experiences from the 4 pretest interviews were discussed and reflected upon by the project team, and the interview guide was revised and finalized accordingly. The semistructured interview guide can be found in Multimedia Appendix 2.

Recruitment of the Interviewees

Participants for the case studies were identified through background research for each country and were sought from the following areas: health care payers, service provision (outpatient and inpatient care), IT and industry sector, patient representatives, science, and other categories (such as nonprofit organizations).

The basis for the inclusion of suitable participants was predicated on digital health expertise. Further interviews were conducted until theoretical saturation was achieved, as described by Glaser and Strauss [31]. A total of approximately 8 to 12 interviews per country were sufficient to capture both the breadth and depth of relevant insights.

The process of recruiting participants for interviews was initiated on August 1, 2024, and has been concluded by September 30, 2025. The initial contact was made via email. The email contained a brief overview of the project and an offer to meet online to discuss any queries. Responses were followed up, and where no response was received, a follow-up email was sent, and in some cases, telephone contact was made. Attempts were made to arrange further interview appointments through existing contacts in the country. The recruitment period for each country commenced approximately 2 to 3 months before the onsite data collection. Before the interview, the participants receive information on the general conditions of the interviews. This includes the aim of the project, the type and duration of the interview, and the further use of the data. Consent was also

XSL•FO
RenderX

obtained from the interviewees to make audio recordings and to ensure later transcription and analysis of the data.

Collecting and Analyzing the Interview Data

A team of 4 experienced researchers (LK, ALB, SS, and VEA) conducted individual interviews with experts in 8 of the 9 countries on site from November 2024 to April 2025. Further interviews were carried out remotely. The data collection for all case studies has been completed by the end of September 2025. A total of more than 100 expert interviews were conducted. The schedule of the interviews can be seen in Figure 4. In Israel, the interviews were conducted remotely over the entire survey period until the end of September 2025. If an onsite appointment at the workplace of the expert or a public place was not feasible, an online interview was arranged. The approach to the interviews was based on the specially created semistructured interview guide. The interviews are scheduled for 30 to 45 minutes. Both the onsite interviews and the digital conversations were recorded by an appropriate recording device. The audio recordings were transcribed with the support of artificial intelligence within the software MAXQDA Transcription (VERBI Software GmbH) and the integrated transcription function in Microsoft Word. The transcriptions are revised in terms of content and formatting. The transcription is based on the content-semantic transcription according to Dresing and Pehl [27]. Consequently, citations are not made according to spoken language, but verbally. Furthermore, time stamps are set, and the paragraphs are divided by interviewer and expert. The evaluation of the data within the finished transcripts is carried out manually within the MAXQDA software using the qualitative content analysis according to Kuckartz and Rädiker [28]. In accordance with the procedure of qualitative content analysis, inductive and deductive coding is carried out by researchers within the framework of a category system. Deductive codes are developed based on the semistructured interview guide, which covers topics such as governance, incentive mechanisms, and interoperability. Inductively, researchers primarily add subcategories that are assigned to the main codes. This qualitative content analysis will enable them to evaluate interviews based on relevant criteria, such as success factors or tipping points. Consequently, the analysis will yield hypotheses about possible success factors, the transfer of which will be evaluated in a subsequent quantitative survey.

Figure 4. Schedule of the interviews. This figure shows the schedule for the expert interviews, which are conducted onsite from November 2024 to April 2025.

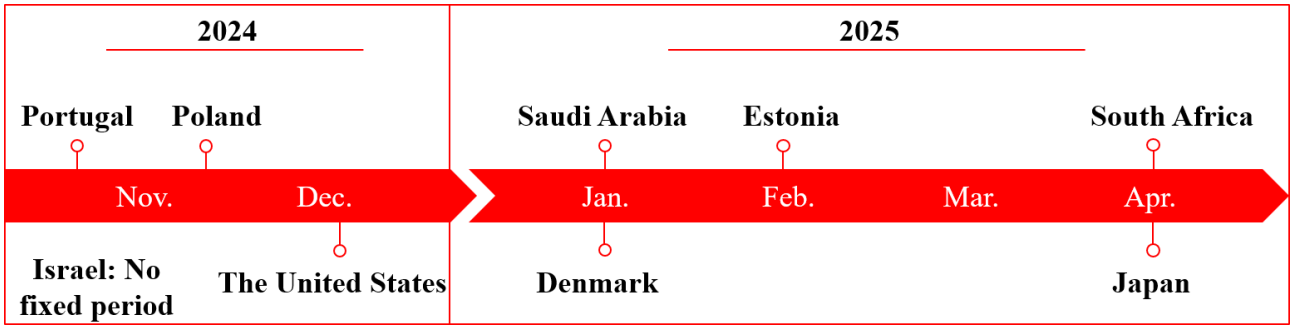

Quantitative Survey

In an iterative process, the quantitative survey will be carried out based on the results of the qualitative interviews. The aim of this cross-sectional data collection is to assess the transferability of key findings from the case studies to the context and status quo of digital health in Germany. Furthermore, it will be analyzed which successful strategies and measures, observed in other health care systems, could be implemented in the German system. To this end, a customized online questionnaire will be developed based on the qualitative data obtained. A set of closed and semi-open questions will be designed. The developed questionnaire will be pretested in a 2-step procedure. In the first step, the wording of the questions and the response categories will be checked in an internal pretest with 2 to 3 methodological experts. If required, the questionnaire will then be optimized as necessary. In the second step, the adapted questionnaire will be field-tested in the targeted population, with a sample of at least 10 participants, with each subgroup being represented by 2 people [29]. The period for recruitment is scheduled to take place from October 15 to December 15, 2025. The survey instrument will be disseminated in various iterations to relevant stakeholder groups within the German health care system, following a convenience sampling approach, including patient representatives, health insurance funds, associations of statutory health insurance, physicians, professional organizations, and other stakeholders in health care policy. The online questionnaire will be distributed via email, social media, and professional networks. Before completing the questionnaire, participants are informed about data collection, data protection, and data confidentiality, and consent is obtained.

The required minimum sample size to achieve adequate statistical power was calculated a priori using the software program G*Power [31]. To identify differences and correlations between the 5 subgroups of a medium effect size with 95% power, a sample size of a minimum of 305 participants is required, equating to 60 participants per stakeholder group. The objective is to obtain a response rate of minimum of a minimum of 305 to 400 participants.

The collected questionnaires are subjected to a process of review, during which the answers will be checked for completeness and plausibility. If necessary, data cleansing will be performed according to predefined criteria. The final quantitative data will be analyzed using the statistical analysis software IBM SPSS (version 30.0.0). Therefore, both descriptive and inferential tests will be used. For the descriptive analysis, measures of absolute and relative frequencies, central tendency

XSL•FO
RenderX

(mode, median), and dispersion (range, IQR) will be reported. Correlations and subgroup differences will primarily be analyzed using the Spearman rank correlation and the Kruskal-Wallis test. Finally, the key results of the quantitative evaluation will inform the development of recommendations for action.

Ethical Considerations

Before the initiation of the study, a data protection concept was formulated, encompassing the methodology for data collection and the protocol for data protection, as delineated by the researchers. A consent form was developed as part of the data protection concept and presented to each participant in an expert interview. By signing this consent form, the interviewer and interviewee agreed to the processing of the data. The audio files and transcripts are accessible exclusively to the project team at Hannover Medical School. The results of the qualitative content analysis will be disseminated without identification of the interview partners. Participants in the quantitative survey were also informed about data collection and privacy, and consent was obtained before participation in the survey as a mandatory part of participation. The collection and analysis of data are conducted anonymously. Subsequent to the formulation of this concept, the ethics vote of the ethics committee of Hannover Medical School, Ethikkommission der Medizinischen Hochschule Hannover, was obtained. The study approach was approved by the members of the ethics committee (approval number: 11440_BO_S_2024) and was deemed to be "without any concerns." The study procedure is consistently aligned with the Declaration of Helsinki and the principles of good clinical practice [32]. The interviewees and participants in the survey did not receive any compensation.

Results

The data collection process for our qualitative research method was initiated in November 2024 and completed by the end of September 2025. A total of 109 expert interviews were conducted. The analysis and interpretation of the data will be completed in December 2025. The initial results of the qualitative survey are targeted for publication in spring 2026. In October 2025, the online questionnaire was disseminated to the relevant stakeholders. The analysis of the quantitative data and its subsequent submission for publication are scheduled for 2026.

Discussion

Anticipated Findings

In light of the challenges facing the digital transformation of the German health care system, as outlined in the introduction, the focus of this research project is particularly timely and relevant. Despite recent progress in digital health, Germany continues to face significant barriers that hinder the full implementation and realization of digital health's potential benefits. By systematically analyzing the success factors behind the digitalization efforts of other health care systems, the results of this research project will contribute to the development of evidence-based policy recommendations for the digital transformation of health care systems in Germany. By integrating insights from qualitative expert interviews and quantitative survey data, the study provides a comprehensive understanding of the political, social, and technological factors driving digitalization. This approach will help identify key success factors and address existing research gaps in the national context. Ultimately, the findings will support tailored recommendations that not only advance Germany's health care policies but also enhance its competitiveness at an international level.

In the following, the strengths and limitations of the methodological approach, the selection of the case studies, as well as the qualitative and quantitative survey are discussed.

Scope

This study adopts a broad perspective on the digital transformation of health care systems, encompassing various health care systems as case studies and incorporating insights from diverse stakeholders and experts across different countries. Unlike research that focuses on specific technological innovations or applications, this study examines the political and social processes that shape digital transformation in health care. The advantage of this approach is its ability to generate comprehensive insights and provide a holistic understanding of a highly complex topic. However, a significant challenge is ensuring that all relevant perspectives and aspects are adequately captured, analyzed, and interpreted, given the multitude of factors influencing digitalization processes in different health care systems.

Methodological Approach

The study uses a mixed methods approach, which strengthens the validity and robustness of the findings. Combining qualitative and quantitative methods allows the research to benefit from the in-depth understanding generated through semistructured expert interviews and to obtain more generalizable results through the quantitative online survey. The qualitative component of this approach is based on case study research. While case studies are sometimes criticized for lacking systematic structure and methodological flexibility, they are well suited to addressing the research questions posed in this study. The flexibility intrinsic to case study research enables an in-depth exploration of contextual factors and dynamics within each health care system, providing rich insights that would be difficult to obtain through more rigid methodological frameworks. To enhance the reliability and objectivity of the findings, this study follows the recommendations of Merriam [11], ensuring a structured approach to case study research and mitigating potential biases. The quantitative survey, in turn, provides a broader empirical foundation to validate these insights, particularly in the context of the German health care system.

The selection of case study countries posed a considerable challenge. In other studies, the selection of cases is often presented in a nontransparent or overly concise manner. In contrast, this study is based on a targeted selection process grounded in extensive research and discussions with experts in the field. The primary difficulty lies in accounting for the numerous criteria that influence the success factors of digital

XSL•FO
RenderX

transformation within a health care system and evaluating the potential for the transfer of these factors to the German health care system. In addition, the selection process must account for limited human and financial resources available. The diverse case selection approach used in this study is well suited for exploratory research [12]. By selecting countries that exhibit a broad range of characteristics in terms of geographical region and size, economic power, health care system typology, and governance structure, this approach ensures a wide spectrum of insights into digital transformation processes. While this method does not prioritize representativeness, it allows for a more nuanced understanding of the diversity of digital health strategies and facilitates the identification of relevant success factors that may inform digital health strategies in Germany. Despite the inherent challenges, the selection of case studies is appropriate for addressing the research questions and maximizing knowledge gain.

Qualitative Survey

The data collection within the case studies was carried out using a qualitative methodology based on expert interviews. The choice of this qualitative method proved to be suitable as it enabled the researchers to gain in-depth and comprehensive insights into the experiences and perspectives of the experts in the field of digital health.

The selection of experts was made on the basis of their demonstrated proficiency in the domain of digital health, with the objective of acquiring particular insights and experiences during the course of the interviews. Furthermore, given the heterogeneity of the occupational distribution of the interviewees, it was anticipated that a broad spectrum of perspectives would be obtained. Accordingly, the interviewees work in the areas of health care policy, payers, service provision (outpatient and inpatient), IT and the industrial sector, patient representation, science, and beyond. The comprehensive perspectives incorporated in this study are a notable strength. The number of interviewees per health care system was set at a minimum of 8 to 12, with the intention of attaining a state of saturation. However, due to the unavailability of some interviewees during the specified period, alternative dates were arranged for the remaining interviews to be conducted remotely. The researchers also encountered challenges in recruiting participants due to cultural differences in scheduling appointments and infrequent feedback from the experts.

The use of a structured interview guide was instrumental in facilitating the systematic conduct of the interviews, thereby ensuring consistency in the formulation of pertinent inquiries by the researchers. This systematic approach was meticulously designed to guarantee the exploration of specific key topics. The use of a semistructured design was deliberate, with the objective being to capture the diverse perspectives of the experts.

The majority of the interviews were planned as face-to-face conversations with the experts on site. This allows the researchers to get a first impression of the digital conditions in the country and in different institutions. The advantage of conducting interviews in person is that nonverbal signals, such as body language, tone of voice, and gestures, can be captured. However, some of the interviews were carried out in a location with pronounced background noise, which made transcription difficult, although not impossible largely due to the use of AI transcription tools. The addition of remote interviews was also extremely useful to interview experts who were difficult to reach. However, it should be noted that remote interviews are susceptible to technical challenges, including unstable internet connections and poor audio quality. Nonetheless, with adequate preparation, these issues can be mitigated.

Concerning the evaluation of results, it is important to acknowledge the inherent subjectivity of researchers [27]. The management of data, its categorization, and the allocation of locations may consequently yield a distorted representation of the data. Ensuring the analysis is conducted by 2 researchers and that there is a regular exchange of information on the status of the evaluation to date can ensure intersubjective traceability and achieve an equal understanding of the analysis process. In addition, the findings will be presented transparently [27].

Quantitative Survey

As part of the planned quantitative survey, within which the findings from the case studies are to be tested for the German health care system, it is essential to create a customized questionnaire based on the qualitative preliminary work. This will enable the analysis of relevant success factors. However, the challenge here is to precisely operationalize complex concepts, such as digital acceptance and implementation barriers, so that they remain measurable and interpretable.

Another challenge could lie in recruiting a sufficiently large and diversified sample to enable reliable statements for different stakeholder groups. The aim is to achieve a sample size of 305 to 400 completed questionnaires, and it is acknowledged that self-selection could lead to bias, as people with a strong interest in the topic could participate. This should be counteracted by a broad distribution of the online questionnaire. It should also be mentioned that possible limitations in data analysis due to sample selection bias or nonresponse bias must be taken into account. The statistical analysis, which is to be conducted using the statistical analysis software SPSS, will enable a systematic evaluation.

Following the analysis of the quantitative survey, the core results that can be transferred to the German health care system will be determined, as well as the extent and form in which they can be transferred. This issue will be tackled with a clearly defined plan that forms the basis of the research project and is regularly reviewed through regular dialogue within the research team and with the project's cooperation partners.

Conclusions

In summary, the integration of qualitative case studies and quantitative transferability assessment within a mixed methods approach is expected to generate valuable insights into the key factors and tipping points that influence the successful implementation of digital health care solutions. In addition, all relevant stakeholders in the health care sector can be considered comprehensively. The insights gained from this study will not only shed light on the success factors that have fundamentally contributed to digital transformation in international health care systems. They will also inform effective recommendations for

XSL•FO
RenderX

action to accelerate the digital transformation in the German health care system.

Acknowledgments

The authors would like to acknowledge the former and present members of the scientific advisory board of the "NADI" project from the European Observatory on Health Systems and Politics, the German Digital Healthcare Association, the Leiden University Medical Centre, the Association of Statutory Health Insurance Physicians of Lower Saxony, the Company Health Insurance Fund (BKK) Umbrella Organization (German: BKK Dachverband), and the Federal Association of Managed Care (BMC) for their valuable insights and guidance throughout the development of this study. The authors would also extend their gratitude to their project partners from the University of Applied Sciences Neubrandenburg and the Technical University of Berlin for their collaboration and support. Their contributions have been essential in shaping the study design and methodology.

Artificial intelligence tools, such as DeepL Write (DeepL SE), ChatGPT (OpenAI), MAXQDA Transcription, and the integrated transcription function in Microsoft Word, were used to assist the researchers in their work. The English language and writing style of this manuscript were improved by using the assistance of the DeepL Write. The transcription of the audio recordings was assisted by the MAXQDA Transcription software and the integrated transcription function within Microsoft Word. Figure 3 is the authors' own illustration, created using Python code and the Folium library. Geospatial data were sourced from OpenStreetMap. ChatGPT was used to assist with drafting and refining the Python code. All suggestions were reviewed and incorporated by the authors.

Funding

This work was funded by the German Innovation Fund (grant 01VSF23031). The German Innovation Fund did not contribute to the design of the study or the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to data protection reasons but are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: LK, ALB, VEA
Data curation: LK, ALB
Funding acquisition: VEA, RB
Investigation: LK, ALB, SS, VEA
Methodology: LK, ALB, SS, VEA
Project administration: LK, ALB
Validation: LK, ALB
Visualization: LK, ALB
Writing – original draft: LK, ALB (both authors contributed equally to this work)
Writing – review & editing: LK, ALB, SS, VEA, RB

Conflicts of Interest

None declared.

Multimedia Appendix 1
Shortlist of countries considered for case study selection and their key characteristics.
[DOCX File, 40 KB - resprot_v15i1e80301_app1.docx]

Multimedia Appendix 2
Semistructured interview guide.
[PDF File, 258 KB - resprot_v15i1e80301_app2.pdf]

Checklist 1
COREQ checklist.
[PDF File, 424 KB - resprot_v15i1e80301_app3.pdf]

References

XSL•FO
RenderX

1. Law for Secure Digital Communication and Applications in Healthcare: E-Health Act [Report in German]. : German Federal Parliament; 2015 URL: http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl115s2408.pdf [accessed 2025-12-30]
2. Law for Better Care Through Digitalization and Innovation: Digital Care Act – DVG [Report in German]. : German Federal Parliament; 2019 URL: http://www.bgbl.de/xaver/bgbl/start.xav?startbk=Bundesanzeiger_BGBl&jumpTo=bgbl119s2562.pdf [accessed 2025-12-30]
3. Bratan T, Schneider D, Heyen NB, et al. E-health in Germany: development prospects and international comparison: studies on the German innovation system [Report in German]. : Expert Commission on Research and Innovation; 2022 URL: https://www.econstor.eu/bitstream/10419/251366/1/1795368888.pdf [accessed 2025-12-30]
4. Bertram N, Püschner F, Gonçalves ASO, et al. Introduction of an electronic patient record in Germany against the backdrop of international experience. In: Klauber J, Geraedts M, Friedrich J, editors. Hospital Report 2019: The Digital Hospital: Springer Berlin Heidelberg; 2019:3-16. [doi: 10.1007/978-3-662-58225-1_1]
5. eGovernment MONITOR 2017 – Use and acceptance of digital administrative services – Germany, Austria and Switzerland in comparison [Report in German]. : Initiative D21; 2017 URL: https://initiatived21.de/app/uploads/2017/10/egovernmentmonitor2017_20171129.pdf [accessed 2025-12-11]
6. Thiel R, Deimel L, Schmidtmann D, et al. #SmartHealthSystems: an international comparison of digitalisation strategies [Report in German]. : Bertelsmann Foundation; 2018 URL: https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/smarthealthsystems/ [accessed 2025-12-30]
7. Pariso P, Picariello M, Marino A. Driving technological efficiency in European healthcare: the role of digital innovation across EU. IJEEP 2025;15(4):490-498. [doi: 10.32479/ijeep.17898]
8. Israel country commercial guide - digital economy. International Trade Administration. 2024. URL: https://www.trade.gov/country-commercial-guides/israel-digital-economy [accessed 2025-12-11]
9. Volpert A, Hery D, Kraushaar J, et al. Digital healthcare 2033: trends, scenarios and theories [Report in German]. : Fraunhofer Society; 2023 URL: https://publica-rest.fraunhofer.de/server/api/core/bitstreams/a621ba91-9a73-4b7d-b9d8-496b07a1a057/content [accessed 2025-12-30]
10. Steiner GA. Strategic Factors in Business Success: Financial Executives Research Foundation; 1969.
11. Gladwell M. The Tipping Point: How Little Things Can Make a Big Difference: Little Brown & Company; 2000. URL: https://www.hachettebookgroup.com/titles/malcolm-gladwell/the-tipping-point/9780316316965/?lens=little-brown [accessed 2025-12-12]
12. van Nes EH, Arani BMS, Staal A, et al. What do you mean, “Tipping Point”? Trends Ecol Evol 2016 Dec;31(12):902-904. [doi: 10.1016/j.tree.2016.09.011] [Medline: 27793466]
13. Brohmann B, David M. Transformation strategies and models of change for sustainable social change: tipping point concepts in the context of sustainable social change: report of AP3 [Report in German]. : Federal Environment Agency; 2015 URL: https://www.umweltbundesamt.de/sites/default/files/medien/378/publikationen/texte_67_2015_tipping_point_konzeptionen_im_kontext_eines_nachhaltigen_gesellschaftlichen_wandels_1.pdf [accessed 2025-12-30]
14. Yin RK. Case Study Research and Applications: Design and Methods, 6th edition: SAGE Publications; 2018. URL: https://quicksearch.lib.iastate.edu/discovery/fulldisplay?vid=01IASU_INST:01IASU&tab=Everything&docid=alma9921205216802756&lang=en&context=L [accessed 2025-12-12]
15. Sibbald SL, Paciocco S, Fournie M, Van Asseldonk R, Scurr T. Continuing to enhance the quality of case study methodology in health services research. Healthc Manage Forum 2021 Sep;34(5):291-296. [doi: 10.1177/08404704211028857] [Medline: 34227408]
16. Yazan B. Three approaches to case study methods in education: Yin, Merriam, and Stake. TQR 2015. [doi: 10.46743/2160-3715/2015.2102]
17. Seawright J, Gerring J. Case selection techniques in case study research. Polit Res Q 2008 Jun;61(2):294-308. [doi: 10.1177/1065912907313077]
18. World digital competitiveness ranking. IMD. URL: https://www.imd.org/centers/wcc/world-competitiveness-center/rankings/world-digital-competitiveness-ranking/ [accessed 2025-12-30]
19. Digital Economy and Society Index (DESI) 2022. : European Commission; 2022 URL: https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2022 [accessed 2025-12-12]
20. Digital riser report 2021. : European Center for Digital Competitiveness; 2021 URL: https://digital-competitiveness.eu/digitalriser/ [accessed 2025-03-07]
21. Digital riser report 2020. : European Center for Digital Competitiveness; 2020 URL: https://digital-competitiveness.eu/wp-content/uploads/ESCP_Digital-Riser-Report_2020-1.pdf [accessed 2025-03-11]
22. World Map 2023. Global Digital Health Monitor. URL: https://monitor.digitalhealthmonitor.org/map [accessed 2025-10-09]
23. Palm K, Brantnell A, Peolsson M, Özbek N, Hedström G. National eHealth strategies: a comparative study of nine OECD health systems. BMC Health Serv Res 2025 Feb 18;25(1):269. [doi: 10.1186/s12913-025-12411-7] [Medline: 39966936]
24. Veterans Health Administration. Veterans Affairs. URL: https://www.va.gov/health/ [accessed 2025-03-13]

XSL•FO
RenderX

25. Folium. URL: https://python-visualization.github.io/folium/latest/ [accessed 2026-01-20]
26. OpenStreetMap. 2025. URL: https://www.openstreetmap.org/copyright [accessed 2025-02-13]
27. Helfferich C, Baur N, Blasius J. Guided and expert interviews. In: Handbook of Empirical Social Research Methods: Springer Fachmedien Wiesbaden; 2022:875-892. [doi: 10.1007/978-3-658-37985-8_55]
28. Kruse J. Qualitative Interviewforschung [Book in German]: Beltz; 2015. URL: https://www.beltz.de/fachmedien/soziologie/produkte/details/15051-qualitative-interviewforschung.html [accessed 2025-12-12]
29. Garcia D. Qualitative methodologies for questionnaire assessment: development of a survey on gender-based violence. : ISTAT; 2017.
30. Bhalla S, Bahar N, Kanapathy K. Pre-testing semi-structured interview questions using expert review and cognitive interview methods. IJBM 2023;7(5):11-19. [doi: 10.26666/rmp.ijbm.2023.5.2]
31. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007 May;39(2):175-191. [doi: 10.3758/bf03193146] [Medline: 17695343]
32. German Research Foundation. Guidelines for safeguarding good research practice. code of conduct. Zenodo. 2025. URL: https://zenodo.org/records/14281892 [accessed 2025-01-12]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research

Edited by A Schwartz; submitted 08.Jul.2025; peer-reviewed by F Fonda, L Maaß; revised version received 14.Oct.2025; accepted 31.Oct.2025; published 21.Jan.2026.

Please cite as:
Kraft L, Brecher AL, Sgraja S, Busse R, Amelung VE
International Case Studies to Identify Success Factors and Contextual Conditions in the Digital Transformation of Health Care Systems and Derive Lessons for Germany: Study Protocol for a Mixed Methods Study
JMIR Res Protoc 2026;15:e80301
URL: https://www.researchprotocols.org/2026/1/e80301
doi:10.2196/80301

© Lena Kraft, Anna-Lena Brecher, Sophia Sgraja, Reinhard Busse, Volker Eric Amelung. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Perinatal Health Care Among Climate Migrant Women: Protocol for a Scoping Review

Giti Khalaj[1,2,3], MSc; James Plaisimond[1,2,3], MA; Thierno Diallo[1,2,3], PhD; Lisa Merry[4,5], PhD; Steven Ouellet[1], PhD; Mohamed Ali Ag Ahmed[5,6,7], PhD; Lara Maillet[8], PhD; Samira Amil[2,3,9], MD; Marielle M'bangha[10], MA; Erika Adriana Corona[1], MA; Marie-Pierre Gagnon[1,2,3], PhD

[1]Faculty of Nursing, Université Laval, 1050 Av. de la Médecine, Québec, QC, Canada

[10]Service de référence en périnatalité pour les femmes immigrantes de Québec, Québec, QC, Canada

[2]CHU de Québec – Université Laval Research Center, Québec, QC, Canada

[3]VITAM – Centre for Research in Sustainable Health, Québec, QC, Canada

[4]Faculty of Nursing, Université de Montréal, Montréal, QC, Canada

[5]SHERPA University Institute, Montréal, QC, Canada

[6]Administration publique et gestion des services de santé, École des hautes études publiques (HEP), Université de Moncton, Moncton, NB, Canada

[7]Department of Health Management, Evaluation and Policy, Université de Montréal, Montréal, QC, Canada

[8]École nationale d'administration publique, Montréal, QC, Canada

[9]School of Nutrition, Université Laval, Québec, QC, Canada

Corresponding Author:
Marie-Pierre Gagnon, PhD
Faculty of Nursing, Université Laval, 1050 Av. de la Médecine, Québec, QC, Canada

Abstract

Background: Climate change–induced international migration has the potential to negatively impact the health and well-being of displaced populations. Pregnancy often serves as a point of entry into the health care system for migrant women; however, these women often face reduced access to maternal health care services compared to nonmigrants. In the context of climate-related international migration, these disparities may be further exacerbated, increasing the risk of maternal morbidity and adverse perinatal outcomes. While the intersections between climate change, migration, and health are increasingly acknowledged, literature specifically focused on international climate-related migrant women—particularly during the perinatal period—remains limited and dispersed. Thus, there is a growing need for research and synthesized data on climate change, population movements, and the perinatal health care needs of childbearing women.

Objective: The aim of this study is to examine and describe the scope and nature of available evidence on maternal health and care for international climate-related migrant women, from pregnancy through the postpartum period.

Methods: We will conduct a scoping review following the Joanna Briggs Institute methodology. A tailored search strategy using key terms related to climate change, migration, women, and perinatal health care will be applied to four databases—Embase, CINAHL, PsycINFO, and Ovid MEDLINE—without restriction on publication date. Relevant gray literature sources will also be searched and considered for inclusion. Only literature published in English, French, Portuguese, or Spanish will be included. Two reviewers will independently screen full-text records based on predefined inclusion criteria and extract the relevant data.

Results: A total of 741 studies were screened from 1113 records. Results summarizing perinatal health outcomes and needs, care experiences, barriers faced by international climate-related migrant women, and recommendations will be reported using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram. We anticipate finalizing the manuscript for this work in 2026.

Conclusions: Considering vulnerability factors related to migration status is essential to improving access to integrated perinatal health care and reducing health inequities among immigrant women. This review will provide valuable insights to tailor interventions to the social and cultural needs of climate-affected migrant women during the perinatal period.

Trial Registration: OSF Registries osf.io/cs4zm; https://osf.io/cs4zm/overview

International Registered Report Identifier (IRRID): PRR1-10.2196/84176

(JMIR Res Protoc 2026;15:e84176) doi:10.2196/84176

KEYWORDS

climate change; migration; women; maternal health; perinatal health care; prenatal period

XSL•FO
RenderX

Introduction

Background

Climate change, as defined by the Intergovernmental Panel on Climate Change [1], refers to a change in the state of the climate that can be identified (eg, using statistical tests) by changes in the mean and/or the variability of its properties and that persists for an extended period, typically decades or longer. It refers to any change in climate over time, whether due to natural variability or as a result of human activity. Slow-onset climate processes refer to gradual environmental changes caused by climate change, such as sea-level rise, desertification, salinization, land degradation, biodiversity loss, glacial retreat, and changing rainfall patterns [2]. Unlike sudden-onset events, these processes unfold over longer time frames and can progressively undermine the viability of livelihoods, leading to delayed or protracted migration decisions.

Human migration in response to ecological changes has occurred since the origin of our species [3]. However, the current pressures that climate change exerts on human mobility are relatively recent and continue to intensify [4]. Climate change is increasingly understood as a "threat multiplier" because of its interaction with other economic, political, and social factors—exacerbating existing socioeconomic and political vulnerabilities and drivers of migration [5,6]. According to several projections, climate change is expected to substantially increase the scale of population displacement in the coming decades [7-11].

Climate migration is defined by the International Organization of Migration (IOM) as the movement of a person or groups of persons who, predominantly for reasons of sudden or progressive change in the environment due to climate change, are obliged to leave their habitual place of residence or choose to do so, either temporarily or permanently, within a state or across an international border [12]. These environmental changes can be sudden-onset events, such as floods or hurricanes, or slow-onset processes, such as prolonged droughts or sea-level rise, which gradually render areas uninhabitable [2]. For the purposes of this review, we will adopt this IOM definition of climate migration to guide study selection and screening.

Despite the growing importance of this phenomenon, global data remain fragmented and underresearched, particularly concerning migrations related to climate changes. Currently, there is no internationally recognized legal framework or formal immigration status for individuals displaced by climate change, which makes it nearly impossible to accurately estimate the prevalence of climate-related migration. This limitation is further complicated by the fact that climate-related factors often intersect with economic, political, and social drivers of migration, making causal attribution complex [13,14]. However, the increasing number of internal displacements triggered by sudden-onset climate events in recent years, alongside sustained international migration from countries severely affected by slow-onset climate changes, suggests that climate-related migration is steadily rising [1,15].

While migration is increasingly recognized for its adaptive potential in the context of climate change, and research on the health impacts of climate change is expanding, the specific health implications of climate-related migration remain underexplored. Climate-induced migration can have serious consequences for the health of migrated individuals, particularly in contexts of forced migration. The 2015 Paris Agreement acknowledges that women are disproportionately affected by climate change [16], suggesting that women's health, especially during the perinatal period, is also disproportionately impacted on a global scale. Emerging evidence indicates that climate change can affect childbearing health through increased exposure to heat stress, air pollution, food and water insecurity, and disruptions to health care systems. These conditions are associated with heightened risks of adverse maternal and neonatal outcomes, such as preterm birth, low birth weight, stillbirth, and complications during labor [17,18].

Pregnancy often serves as a critical entry point into the health care system for migrant women [19]. Yet, migrant women are generally less likely than nonmigrant women to access adequate maternal health care services [19-22]. A systematic review by Heaman et al [23] reported that migrant women in Western industrialized countries are more likely to receive inadequate prenatal care, begin care later in pregnancy, and have fewer than the recommended number of prenatal visits. Childbearing migrant women—including those who are internally displaced, are living in refugee camps, or have crossed international borders—face increased health risks throughout the migration process. These groups often encounter multiple vulnerabilities, such as unstable housing, food insecurity, trauma, and limited access to culturally appropriate health care [24-27]. They are at greater risk of maternal morbidity and mortality and are less likely to access health care, particularly in humanitarian or low-resource contexts [28,29]. Given these intersecting vulnerabilities and the additional strain climate change places on health systems and migration patterns, women experiencing climate-related migration may face compounded risks and heightened barriers to perinatal care. Most studies have focused on women remaining within disaster-affected regions. As such, the perinatal risks associated with climate migration can only be postulated by extrapolating from existing evidence on the effects of migration and internal displacement on pregnancy. Although most migration driven by climate change tends to occur within national borders, some people are also compelled to cross international boundaries. However, comprehensive global data—like movements linked to slow-onset climate processes such as sea-level rise or land degradation—remain scarce [30].

While some reviews have addressed aspects of climate-related migration and perinatal or sexual and reproductive health [31,32], there remains a need for a focused scoping review that specifically maps evidence on the perinatal health and care experiences of international climate-related migrant women. Due to the limited evidence available on migration specifically linked to climate processes, a scoping review is warranted to map the existing evidence on the perinatal health and care experiences of different groups of international migrant women affected by climate change, including those affected by both

XSL•FO
RenderX

sudden- and slow-onset climate processes, in order to inform practice and policy, identify knowledge gaps, and guide future research priorities aimed at improving maternal and newborn health outcomes in the context of a changing climate. Specifically, it aims to examine and describe the scope and nature of available evidence on maternal health and care for international climate-related migrant women, from pregnancy through the postpartum period.

Review Question

What is known about the perinatal health and care of migrant women who have experienced climate-related migration?

Subquestions

There are two subquestions this review aims to answer:

1. How do climate-related migrant women perceive perinatal health care services in host countries?
2. What barriers and challenges do they face in accessing perinatal health care services?

Methods

Overview

This scoping review will be conducted in accordance with the Joanna Briggs Institute (JBI) methodology for scoping reviews [33]. This framework involves a systematic process that includes the identification of the research questions, the comprehensive search for relevant literature, the selection of eligible literature, the extraction and charting of data, synthesis of the findings, and, where applicable, stakeholder consultation. To ensure methodological transparency and reproducibility, the review will adhere to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist and explanation [34]. This protocol will also be registered on the Open Science Framework to enhance transparency and facilitate access.

Inclusion Criteria

The inclusion will align with the JBI framework including population, concept, and context (PCC) [35]. Articles addressing the perinatal health care received by climate-related migrant women during prepartum and postpartum will be included in our review.

We will include all types of evidence matching the PCC (population, concept, context) criteria.

Population

We will include literature focusing on international climate-related migrant women during the perinatal period. During the screening process, studies will be included if participants are explicitly identified as climate-related migrants, defined as women who, in the perinatal period, predominantly for reasons of sudden or progressive environmental changes due to climate change (eg, floods, droughts, extreme heat, temperature anomalies, cyclones), are obliged to leave their habitual place of residence or choose to do so, either temporarily or permanently, within a state or across an international border.

Concept

The concept under investigation is perinatal health, which encompasses maternal health (both physical and mental), access to and utilization of health care services, perceptions of care, and barriers and facilitators encountered.

Context

We will include literature set in health care or community settings in host countries, including both high-income countries and low- and middle-income countries.

Types of Sources

This scoping review will consider peer-reviewed research using qualitative, quantitative, or mixed methods designs, as well as existing reviews (such as systematic reviews, scoping reviews, and critical literature reviews). Relevant gray literature will also be included to ensure a comprehensive overview of the topic. This includes targeted sources such as the World Health Organization, the IOM, the World Bank, the United Nations High Commissioner for Refugees, the Internal Displacement Monitoring Center, other United Nations agencies, government and nongovernmental organization reports, and other policy documents.

Search Strategy

A scoping review of studies reporting health outcomes and health care among climate-related migrant women, specifically concerning pregnancy and postpartum, will be conducted. A comprehensive literature search will be carried out across four electronic databases—Embase, CINAHL, PsycINFO, and Ovid MEDLINE—to identify relevant articles published in English, French, Portuguese, or Spanish (languages spoken by team members) with no restrictions on publication date. These databases will be selected based on their relevance to the research question and their comprehensive coverage of health, social science, and psychological literature coverage. The search strategy, which will incorporate both keywords and subject headings, will be developed in Ovid MEDLINE and then adapted for use in other databases. The search will be structured around main concepts: climate change, migration, women, and perinatal health care. Within each concept, related keywords and controlled vocabulary will be combined using the Boolean operator OR. The concepts themselves will then be combined using AND to ensure the retrieval of relevant and focused results. To supplement the electronic database search, backward and forward citation tracking of relevant publications will be conducted. Additional searches will also be performed using Semantic Scholar, leveraging tools such as Elicit and Consensus. Finally, targeted searches will be carried out using Google Scholar, where only the first 100 results will be assessed.

Study/Source of Evidence Selection

Following the search, all identified citations will be collated and uploaded into Covidence (Veritas Health Innovation) [36], and duplicates will be removed. Following a pilot test, titles and abstracts will be screened by two independent reviewers for assessment against the inclusion criteria for the review. Two independent reviewers will assess the full text of selected citations in detail against the inclusion criteria. Reasons for the

XSL•FO
RenderX

exclusion of sources of evidence at the full-text review stage that do not meet the inclusion criteria will be recorded and reported in the scoping review. Any disagreements between the reviewers at each stage of the selection process will be resolved through discussion or, if necessary, by consultation with an additional reviewer. The results of the search and the study inclusion process will be reported in full in the final scoping review and presented in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram [37].

Data Extraction

Data will be extracted from papers included in the scoping review by two independent reviewers using a data extraction tool developed by the research team in accordance with JBI guidance [38]. The draft data extraction tool will be piloted on a subset of studies to be included in the review to assess its feasibility. The finalized tool will be reported in the scoping review manuscript. The data extracted will include specific details about the descriptive data (title, year, authors, type of study), population (sample size; key characteristics; type of migrant—international migration, either forced or voluntary; source country; length of time—temporary or permanent; type of climate event), concept (perinatal health and care outcomes), context (country; health care or community settings), study/review design and methodology, and key findings relevant to the review questions. The quality of the included studies will be assessed using the Mixed Methods Appraisal Tool [39].

Data Analysis and Presentation

Data extracted from the included sources will be analyzed using both quantitative and qualitative methods, depending on the type of data. We will first perform a descriptive analysis to summarize characteristics such as publication year, country of origin, study design, population, and type of evidence. This will include frequency counts and tabulations to map the distribution of the literature.

For qualitative data, a thematic analysis will be conducted. Two independent reviewers will code the extracted data using a structured spreadsheet in Microsoft Excel. Initial codes will be developed inductively, then refined iteratively through discussion and consensus. Discrepancies will be resolved by a third reviewer if necessary. Themes will then be systematically mapped to the review's PCC elements to ensure alignment with the research questions, in accordance with the JBI recommendations [38].

A narrative synthesis will accompany the findings to contextualize and interpret the data in relation to the review's objectives. Where appropriate, visual representations such as charts or thematic maps will be used to illustrate the key findings. All methods of data extraction and analysis will be guided by the JBI methodology for scoping reviews [33].

Ethical Considerations

This scoping review does not involve the collection of primary data. Ethics approval was obtained from the *Center intégré universitaire de santé et de services sociaux de la Capitale-Nationale*, through the *Comité d'Éthique de la Recherche Sectoriel* evaluator for Population Health and Primary Care (approval number 2025 - 3314). The project underwent both scientific and ethical review and received favorable evaluations, as well as suitability approval from the managers of the participating facilities and clinical departments.

Results

A preliminary search applying the key terms across all 4 databases yielded 1113 records. After duplicates were removed using the Covidence platform, 741 records were retained for further screening. Next, titles and abstracts will be reviewed, followed by a detailed assessment of full-text articles by two reviewers. A preliminary PRISMA-ScR flow diagram outlining the planned selection process is presented using the PRISMA 2020 flow diagram [37]. The PRISMA diagram can be found in Multimedia Appendix 1. Numbers marked as TBD (to be determined) will be updated once the review is completed. We anticipate finalizing the manuscript in 2026.

Discussion

Anticipated Findings

This scoping review is expected to provide an initial mapping of the evidence describing the perinatal health needs, care experiences, and barriers faced by international climate-related migrant women. Based on preliminary knowledge and the anticipated synthesis of existing literature, we expect to identify recurrent challenges such as limited access to timely and culturally appropriate perinatal care, gaps in service coordination, and broader systemic vulnerabilities linked to both migration status and climate-related displacement. It is also anticipated that the review will highlight substantial gaps in research, including the limited documentation of climate-related migration as a determinant of perinatal health and the fragmented nature of available studies across disciplines.

Comparison to Prior Work

Two previous reviews [31,32] have explored climate-related migration in relation to sexual or reproductive health, and none have focused specifically on perinatal health among international climate-related migrants. By narrowing the scope to this population, our review aims to offer a more granular understanding of how climate-related drivers intersect with migration pathways and perinatal health outcomes. This work is expected to contribute conceptual clarity by identifying inconsistencies in how climate-related migrants are defined and operationalized, while also bringing forward methodological gaps such as the absence of standardized indicators or the underreporting of migration drivers in maternal health research.

Strengths

A key strength of this review is the use of the JBI methodology, which ensures rigor, transparency, and reproducibility throughout each stage of the review process. The comprehensive search strategy—spanning 4 multidisciplinary databases and incorporating gray literature—maximizes the likelihood of capturing diverse forms of evidence across countries and contexts. The multilingual expertise of the review team (English,

XSL•FO
RenderX

French, Portuguese, and Spanish) is an additional strength, increasing the inclusiveness of the review and enhancing its ability to identify relevant literature from multiple regions.

Limitations

Several limitations must be acknowledged. First, there is no universally accepted definition or legal status for climate-related migrants, and many studies do not explicitly identify this group, which may lead to underrepresentation or inconsistent terminology across the included evidence. Second, despite including 4 languages, literature published in other languages may remain inaccessible, potentially limiting the global representativeness of the findings. Third, given the expected diversity in study designs, contexts, and conceptual frameworks, heterogeneity across studies may limit comparability and hinder synthesis. Finally, as a scoping review, the study will not assess the quality of included evidence, and conclusions will be descriptive rather than evaluative. A preliminary quality assessment using the Mixed Methods Appraisal Tool will provide context, but the synthesis will remain descriptive.

Future Directions

The anticipated findings will help identify conceptual, methodological, and empirical gaps that can guide future research. In particular, the review may highlight the need for (1) improved identification and documentation of climate-related migrants in health research; (2) standardized approaches for capturing climate-related drivers of migration in perinatal health studies; (3) longitudinal or mixed methods research examining the long-term perinatal outcomes of climate-related displacement; and (4) studies exploring culturally adapted models of perinatal care that address both social and environmental determinants of health. Additionally, the review may point to the need for coordinated policy responses that integrate climate adaptation strategies with migrant-sensitive maternal health services.

Dissemination Plan

In alignment with *JMIR Research Protocols* guidelines, the findings of this review will be disseminated through publication in an open-access journal; presentations at international conferences on maternal health, migration, and climate change; and knowledge translation activities aimed at policymakers, community organizations, and health care providers. Summaries of results will also be shared through institutional networks and public-facing formats such as infographics, policy briefs, or webinars to maximize accessibility and impact.

This scoping review will provide an evidence-based overview of the perinatal health needs, challenges, and care experiences of international climate-related migrant women. By identifying research gaps and mapping key determinants, it will enhance understanding of the intersecting social and environmental factors affecting this population. The anticipated findings may inform the development of inclusive, culturally responsive, and equitable perinatal health care strategies, as well as guide future research and policy initiatives.

Acknowledgments

We would like to thank Marianne Ruel and Marie-Claude Laferrière, Health Sciences Librarians at Université Laval, for their valuable contributions. Marianne provided methodological support in developing the search strategy and identifying relevant sources, while Marie-Claude offered expert guidance throughout the search process and the development of this study.

Generative artificial intelligence was used to support language refinement and wording adjustments. All scientific content, analysis, and interpretation were entirely developed by the authors.

Funding

This project was funded by the Canadian Institutes of Health Research (CIHR, grant 202409RC3-538715-CCR-CFBA-111141).

Authors' Contributions

Conceptualization and study design: MPG, GK
Funding acquisition: MPG
Methodology: MPG, GK, JP, SO
Search strategy development: MPG, GK
Investigation (planned): GK, JP, SO
Validation: MPG
Writing – original draft: GK
Writing – review & editing: all authors
Supervision: MPG
Final approval of the manuscript: all authors

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

PRISMA flow diagram.
[JPG File, 124 KB - resprot_v15i1e84176_app1.jpg]

Peer Review Report 1
Peer review report by the Canadian Institutes of Health Research.
[PDF File, 1035 KB - resprot_v15i1e84176_app2.pdf]

References

1. IPCC Sixth Assessment Report. Working Group 1: The Physical Science Basis. IPCC. 2021. URL: https://www.ipcc.ch/report/ar6/wg1/ [accessed 2026-02-06]
2. Slow onset events: technical paper. : United Nations Framework Convention on Climate Change; 2012 URL: https://unfccc.int/resource/docs/2012/tp/07.pdf [accessed 2026-02-06] [doi: 10.1177/0973408212475203]
3. Finlayson C. Biogeography and evolution of the genus Homo. Trends Ecol Evol 2005 Aug;20(8):457-463. [doi: 10.1016/j.tree.2005.05.019] [Medline: 16701417]
4. McMichael C, Barnett J, McMichael AJ. An ill wind? Climate change, migration, and health. Environ Health Perspect 2012 May;120(5):646-654. [doi: 10.1289/ehp.1104375] [Medline: 22266739]
5. Foresight: migration and global environmental change: future challenges and opportunities. : Government Office for Science; 2011 URL: https://assets.publishing.service.gov.uk/media/5a74b18840f0b61df4777b6c/11-1116-migration-and-global-environmental-change.pdf [accessed 2026-02-06]
6. Tanner T, Lewis D, Wrathall D, et al. Livelihood resilience in the face of climate change. Nature Clim Change 2015 Jan;5(1):23-26. [doi: 10.1038/nclimate2431]
7. Boano C, Zetter R, Morris T. Environmentally displaced people: understanding the linkages between environmental change, livelihoods and forced migration. : Refugee Studies Centre; 2008 URL: https://www.rsc.ox.ac.uk/files/files-1/pb1-environmentally-displaced-people-2008.pdf [accessed 2026-02-06] [Medline: 19036193]
8. Brown O. Climate change and forced migration: observations, projections and implications. : United Nations Development Programme; 2007 URL: https://www.iisd.org/system/files/publications/climate_forced_migration.pdf [accessed 2026-02-06]
9. Human tide: the real migration crisis. : Christian Aid; 2007 URL: https://documentation.lastradainternational.org/lsidocs/christian_07_human_crisis_0109.pdf [accessed 2026-02-06]
10. Clement V, Rigaud KK, de Sherbinin A. Groundswell part 2: acting on internal climate migration. : World Bank; 2021 URL: https://openknowledge.worldbank.org/entities/publication/2c9150df-52c3-58ed-9075-d78ea56c3267 [accessed 2026-02-02] [doi: 10.1596/36248]
11. Climate change, natural disasters and human displacement: a UNHCR perspective. : United Nations High Commissioner for Refugees; 2009 URL: https://www.unhcr.org/sites/default/files/legacy-pdf/4901e81a4.pdf [accessed 2026-02-06]
12. International migration law no.34 – glossary on migration. : International Organization for Migration; 2019 URL: https://publications.iom.int/books/international-migration-law-ndeg34-glossary-migration [accessed 2026-02-06]
13. McLeman R. Migration and displacement risks due to mean sea-level rise. Bulletin of the Atomic Scientists 2018 May 4;74(3):148-154. [doi: 10.1080/00963402.2018.1461951]
14. World migration report 2022. : International Organization for Migration; 2022 URL: https://publications.iom.int/books/world-migration-report-2022 [accessed 2026-06-02]
15. Global report on internal displacement 2023. : Internal Displacement Monitoring Centre (IDMC); 2023 URL: https://www.internal-displacement.org/global-report/grid2023/ [accessed 2026-06-02]
16. The Paris Agreement. : United Nations; 2015 URL: https://unfccc.int/process-and-meetings/the-paris-agreement [accessed 2026-02-06]
17. Kuehn L, McCormick S. Heat exposure and maternal health in the face of climate change. Int J Environ Res Public Health 2017 Jul 29;14(8):853. [doi: 10.3390/ijerph14080853] [Medline: 28758917]
18. Meherali S, Nisa S, Aynalem YA, et al. Impact of climate change on maternal health outcomes: an evidence gap map review. PLOS Glob Public Health 2024;4(8):e0003540. [doi: 10.1371/journal.pgph.0003540] [Medline: 39159145]
19. Gagnon AJ, Zimbeck M, Zeitlin J. Migration and perinatal health surveillance: an international Delphi survey. Eur J Obstet Gynecol Reprod Biol 2010 Mar;149(1):37-43. [doi: 10.1016/j.ejogrb.2009.12.002] [Medline: 20074849]
20. Kiss V, Pim C, Hemmelgarn BR, Quan H. Building knowledge about health services utilization by refugees. J Immigr Minor Health 2013 Feb;15(1):57-67. [doi: 10.1007/s10903-011-9528-8] [Medline: 21959711]
21. Reitmanova S, Gustafson DL. “They can’t understand it”: maternity health and care needs of immigrant Muslim women in St. John’s, Newfoundland. Matern Child Health J 2008 Jan;12(1):101-111. [doi: 10.1007/s10995-007-0213-4] [Medline: 17592762]
22. Kingston D, Heaman M, Chalmers B, et al. Comparison of maternity experiences of Canadian-born and recent and non-recent immigrant women: findings from the Canadian Maternity Experiences Survey. J Obstet Gynaecol Can 2011 Nov;33(11):1105-1115. [doi: 10.1016/S1701-2163(16)35078-2] [Medline: 22082784]

XSL•FO
RenderX

23. Heaman M, Bayrampour H, Kingston D, et al. Migrant women's utilization of prenatal care: a systematic review. Matern Child Health J 2013 Jul;17(5):816-836. [doi: 10.1007/s10995-012-1058-z] [Medline: 22714797]
24. Finding our way in the age of climate change and migration. International Organization for Migration. 2019. URL: https://www.iom.int/news/finding-our-way-age-climate-change-and-migration [accessed 2026-02-02]
25. Krause SK, Jones RK, Purdin SJ. Programmatic responses to refugees' reproductive health needs. Int Fam Plan Perspect 2000 Dec;26(4):181. [doi: 10.2307/2648256]
26. Climate change and health. World Health Organization. 2018 Sep 7. URL: https://www.who.int/teams/environment-climate-change-and-health/climate-change-and-health [accessed 2026-02-02]
27. Zimmerman C, Kiss L, Hossain M. Migration and health: a framework for 21st century policy-making. PLoS Med 2011 May;8(5):e1001034. [doi: 10.1371/journal.pmed.1001034] [Medline: 21629681]
28. Global trends: forced displacement in 2021. : United Nations High Commissioner for Refugees; 2022 URL: https://www.unhcr.org/media/global-trends-report-2021 [accessed 2026-02-06]
29. Kandasamy T, Cherniak R, Shah R, Yudin MH, Spitzer R. Obstetric risks and outcomes of refugee women at a single centre in Toronto. J Obstet Gynaecol Can 2014 Apr;36(4):296-302. [doi: 10.1016/S1701-2163(15)30604-6] [Medline: 24798666]
30. Lee SA, Corbett GA, McAuliffe FM. Obstetric care for environmental migrants. Ir J Med Sci 2024 Apr;193(2):797-812. [doi: 10.1007/s11845-023-03481-9] [Medline: 37715828]
31. Sundaresan A, Uddin R, Sorensen C. The impacts of climate migration on perinatal health and opportunities to safeguard perinatal well-being. Semin Perinatol 2023 Dec;47(8):151845. [doi: 10.1016/j.semperi.2023.151845] [Medline: 37865558]
32. van Daalen KR, Dada S, Issa R, et al. A scoping review to assess sexual and reproductive health outcomes, challenges and recommendations in the context of climate migration. Front Glob Womens Health 2021;2:757153. [doi: 10.3389/fgwh.2021.757153] [Medline: 34816251]
33. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
34. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
35. Apply PCC. University of South Australia. 2025. URL: https://guides.library.unisa.edu.au/ScopingReviews/ApplyPCC [accessed 2026-02-02]
36. Covidence. 2025. URL: https://www.covidence.org [accessed 2026-02-02]
37. PRISMA flow diagram. PRISMA. 2025. URL: https://www.prisma-statement.org/prisma-2020-flow-diagram [accessed 2026-02-02]
38. Pollock D, Peters MDJ, Khalil H, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth 2023 Mar 1;21(3):520-532. [doi: 10.11124/JBIES-22-00123] [Medline: 36081365]
39. Hong QN, et al. Mixed methods appraisal tool (MMAT), version 2018. : Canadian Institutes of Health Research; 2018 URL: http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf [accessed 2026-02-06]

Abbreviations

IOM: International Organization for Migration
JBI: Joanna Briggs Institute
PCC: population, concept, context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by J Sarvestan; submitted 18.Sep.2025; this is a non–peer-reviewed article;accepted 08.Dec.2025; published 17.Feb.2026.

Please cite as:
Khalaj G, Plaisimond J, Diallo T, Merry L, Ouellet S, Ag Ahmed MA, Maillet L, Amil S, M'bangha M, Corona EA, Gagnon MP
Perinatal Health Care Among Climate Migrant Women: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e84176
URL: https://www.researchprotocols.org/2026/1/e84176
doi:10.2196/84176

© Giti Khalaj, James Plaisimond, Thierno Diallo, Lisa Merry, Steven Ouellet, Mohamed Ali Ag Ahmed, Lara Maillet, Samira Amil, Marielle M'bangha, Erika Adriana Corona, Marie-Pierre Gagnon. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 17.Feb.2026. This is an open-access article distributed under the terms of the Creative

XSL•FO
RenderX

Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Evaluation Models for Undergraduate Nursing Clinical Skills: A Scoping Review Protocol

Kateryna Metersky, BSCN, MN, PhD; Tommy Lin, BScN, MN; Tina Lam, BScN; Areej Al-Hamad, BScN, MScN, PhD (HP), PhD (N); Daria Romaniuk, BN, MN, PhD

Daphne Cockwell School of Nursing, Toronto Metropolitan University, 288 Church St, Toronto, ON, Canada

Corresponding Author:
Kateryna Metersky, BSCN, MN, PhD
Daphne Cockwell School of Nursing, Toronto Metropolitan University, 288 Church St, Toronto, ON, Canada

Abstract

Background: Undergraduate nursing students are expected to perform a high-stakes clinical skills test, which ultimately determines their ability to engage in clinical practice. With intake number of students growing nationally, clinical instructors are modifying these skills tests to be shorter in duration as an attempt to meet scheduled class times, severely decreasing the assessments' accuracy and increasing student stress.

Objective: The objectives of this review are to examine best practices of evaluation structures or models for undergraduate nursing students participating in clinical skills-based assessments as part of their program curriculum to explore alternative, best-practice methods for evaluating key clinical skills. The goal is to use what is known in the literature to help inform the development of a new evidence-based clinical skills evaluation model that is adaptive to student population growth, enhances student learning and faculty teaching experiences, and ensures safe patient care.

Methods: This scoping review will consider studies involving nursing students enrolled in an undergraduate nursing degree program who undergo clinical skills testing as part of their degree fulfillment. Original articles published in English from January 1, 2015, to present day will be included. This review will follow the Joanna Briggs Institute methodology for scoping reviews, as well as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines. The literature search will use the following databases: SCOPUS, CINAHL, Medline, PsycInfo, ProQuest Central, and ProQuest Dissertations and Theses to identify relevant sources. Moreover, nonempirical research, such as editorials, opinion papers, gray literature, and reviews, will be included. Data from professional nursing organizations, including the College of Nurses of Ontario (CNO), the Registered Nurses' Association of Ontario (RNAO), and the Canadian Association of Schools of Nursing (CASN), will be considered. Two independent reviewers will conduct a screening of article titles and abstracts, followed by full-text reviews. Together, they will determine which articles will proceed to the data extraction stage. If discrepancies arise, a consensus discussion will take place between the two reviewers with assistance from a third reviewer.

Results: The project was funded in June 2025 (Figure 1). The scoping review is scheduled to begin in July 2025, with the literature search and study selection processes planned through Fall 2025. Results are expected to be submitted for publication in Winter 2026.

Conclusions: The results of this review will summarize the current breadth of knowledge on clinical skills testing and best-practice methods used for clinical skill evaluation among undergraduate nursing students. Any gaps in the literature will be identified as these can be used to guide future research within this area of nursing education.

Trial Registration: OSF Registries osf.io/k7rv4; https://osf.io/k7rv4

International Registered Report Identifier (IRRID): PRR1-10.2196/79805

(JMIR Res Protoc 2026;15:e79805) doi:10.2196/79805

KEYWORDS

clinical skills assessment; nursing; nursing curricula; nursing student; skills testing

Introduction

Standardized skills testing has been used as a method of evaluation for undergraduate students in nursing programs across Canadian schools to assess and measure student competence in executing essential nursing skills. In such programs, students are prohibited from entering clinical practice or performing such skills in a clinical setting until they have successfully demonstrated competence in performing the skills in the laboratory setting. The assessment typically or most often consists of an individual one-on-one evaluation of a sequential demonstration of a psychomotor skill, such as wound care, Foley catheterization, intramuscular injection as examples, to be carried out within a designated timeframe. Such skills testing

XSL•FO
RenderX

is viewed as a high-stakes assessment, meaning that students who fail typically must undergo remediation until they are deemed successful [1].

This model of testing, in which students would face significant consequences if their performance was not evaluated as successful, is quite stressful for students. Such consequences can include inability to move on to in-person clinical practice or having to remain at a certain level in the program until they are successful. Such consequences may make their educational experience stressful and give rise to increased preparation and test-taking anxiety [1,2]. Such an impact on students could potentially lead to poor performance through increased mistake-making during testing despite the student having the requisite knowledge [3]. Nursing students have often expressed that skill pass-offs invoked negative responses and emotions related to this one-time high-stakes assessment, such as fear and being in a high pressure situation, resulting in decreased confidence when performing the required skills [4].

The supply of Canadian nurses has not kept pace with the country's ever-increasing demand for health care services, with forecasting models documenting "a shortage of 60,000 nurses nationwide by 2022 and more than 117,000 by 2030" [5]. Due to current health human resource shortages experienced in acute care settings nationally, both provincial and territorial governments have put forth a call to expand the health care workforce and consequently, increase the number of seats for nursing students across postsecondary institutions. This continual increase in student enrollment numbers without any increase in physical space, nursing instructors, nursing education resources, or instructional hours has already posed significant challenges for the traditional skills testing model [6]. Class size, time, faculty resistance, and workload concerns remain considerable barriers for nursing educators to effectively teach and evaluate clinical competence [7]. Faculty find it increasingly challenging to adequately and fairly assess their students under such constraints [8]. As a result, instructors have been prompted to come up with creative solutions to ensure all students can get tested within the same amount of class time, despite the significant increase in student numbers. Thus, instructors have started to modify the traditional assessment through either decreasing the time students have to demonstrate the entire skill or asking students to come prepared to perform the entire skill, but assigning only a small component of the skill to demonstrate during the testing.

Altering the skills testing model in this manner may result in an unreliable evaluation of students' competence and capability to perform the requisite skill on a real patient in the clinical setting, decreasing the integrity and validity of this assessment. Presently, these adjustments are seldom aligned with the best practices for effective clinical skills evaluation of students. These modifications are a 'cookie-cutter' approach that can lead to issues such as diminishing the integrity of the evaluation process, heightening student anxiety and instructor fatigue, and necessitating both parties to exceed beyond their scheduled class hours to ensure every student is tested prior to entering the clinical setting at a timely point in the semester. Faculty members conclude that more uniformity is needed in how learning outcomes are interpreted and how students are being assessed, advocating for an evaluation process that is more consistent and reliable with clear expectations for both students and educators [9]. There is a strong need to readdress undergraduate nursing curricula regarding clinical competence in order to prepare nursing graduates for the everchanging profession [10]. Failure to accurately assess students and a lack of appropriate educational support as a means for clinical skills acquisition may constitute a threat to patient safety down the line [11]. Therefore, it is necessary for educators to determine students' learning needs and thoroughly identify skill and knowledge gaps that could hinder the development of clinical competence to adequately meet professional nursing care standards [9-11].

Nursing curricula must be designed to equip graduates with the knowledge and skills necessary to meet present and emerging health care demands. There is a need to strengthen students' competence and confidence by integrating new technologies and innovative teaching approaches into traditional methods [12,13]. It is important for educators to emphasize the learning process as a whole rather than focusing solely on the outcomes by thoughtfully pacing and adjusting student support throughout the nursing program. This foundational pedagogy continues to play an important role in contemporary education [14]. Evidence-based practices within educational approaches can help improve the clinical outcome through increased knowledge, skill acquisition, and competency, as well as improve the beliefs, attitudes and behaviors of undergraduate students [15].

This scoping review aims to explore alternative, best-practice methods for evaluating key clinical skills in nursing students, with the goal of informing the development of a new evidence-based evaluation model. The model could better accommodate student population growth, enhance the overall learning experience for students as well as teaching experience for instructors, and guarantee safe patient care in clinical settings. The review addresses the central question: What is known in the literature about how key nursing skills are evaluated in undergraduate nursing students within Canada? Analyzing relevant scholarly and gray literature will reveal the depth of knowledge within current literature regarding best practices for the assessment of nursing clinical skills in a laboratory environment and generate key themes that will ultimately guide the creation of a new evaluation model.

Methods

Study Design

This scoping review will apply the Joanna Briggs Institute (JBI) methodology due to its iterative structure and comprehensive framework for scoping reviews [16]. The selection of sources, data extraction, and presentation of findings will be guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) for its structured and rigorous review process [17]. The protocol was developed using Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols extension (PRISMA-P) [18].

XSL•FO
RenderX

Inclusion Criteria

The Population, Concept, and Context framework will be used to structure the scope and inclusion criteria of the scoping review [16].

Population

This review focuses on undergraduate nursing students enrolled in an undergraduate nursing degree program.

Concept

The core concept of this scoping review focuses on the skills evaluation model of undergraduate nursing students. The concept of a skills evaluation model is defined as a structured and theory-informed approach used to assess the progressive development of clinical skills among undergraduate nursing students. This is grounded in Benner's Novice to Expert theory, in which skill evaluation is viewed as a developmental process through which learners advance from novice to competent practitioners via guided practice, feedback, and reflection [19]. This concept encompasses the frameworks, methods, and tools employed to evaluate and support the acquisition of nursing competence.

Context

The context includes any undergraduate nursing students enrolled in undergraduate nursing degree programs, inclusive of 3 or 4 years in length, which have clinical skills testing in their program curricula, course evaluation, or as an educational requirement. The scoping review will not focus on accelerated, second-entry, and diploma nursing programs. In addition, registered practical nurse programs will be excluded to ensure the consistency of the scoping review focusing on undergraduate nursing degree programs. This distinction will ensure the findings are pertinent to undergraduate nursing curricula and competencies as the complexity of skills tested might vary between such programs.

Types of Sources

The scoping review will include any type of studies with quantitative, qualitative, or mixed method designs. Additionally, nonempirical research, such as editorials, opinion papers, gray literature, and reviews will be included. Finally, professional nursing organizations, such as the College of Nurses of Ontario (CNO), the Registered Nurses' Association of Ontario (RNAO), and the Canadian Association of Schools of Nursing (CASN), will be accessed for any pertinent documents or information on the topic.

Search Strategy

The scoping review will query the following databases for sources: SCOPUS, CINAHL, Medline, PsycInfo, ProQuest Central, and ProQuest Dissertations and Theses. Moreover, key nursing professional organization websites such as RNAO, CASN, and CNO will also be used. A detailed search syntax for CINAHL has been initiated (see Multimedia Appendix 1). In consultation with a research librarian, the search syntax was developed. This initial search syntax will be modified accordingly to the respective criteria of each identified database and will use standard Boolean operators. This comprehensive search strategy will ensure the breadth of existing literature is captured for this scoping review.

Study Selection

EndNote software will be used to organize and manage all retrieved sources. Duplicates among the compiled sources will be identified and removed [20]. Subsequently, a preliminary screening will be conducted to enhance the source selection process, which will be followed by independent screening of sources by two reviewers according to titles and abstracts to ensure they meet the review's inclusion criteria. Sources considered pertinent to the review's central question will be retrieved and all details will be documented in the JBI's System for the Unified Management, Assessment, and Review of Information (SUMARI). This is to facilitate a systematic method for assessing, analyzing, and synthesizing the results [21]. Two independent reviewers will analyze the full texts of the retrieved sources to verify the eligibility of each study according to the predetermined inclusion criteria. The justification for omitting sources will be recorded in the PRISMA flow diagram [22]. If any conflicts arise between the reviewers regarding a source's eligibility, solutions will be pursued through discussions with the third reviewer, the principal investigator. The complete search and selection procedures will be documented in the final scoping review and illustrated within a PRISMA-ScR flow diagram to guarantee transparency and reproducibility of the review [22].

Data Extraction

All included sources will undergo a data extraction process carried out by two independent reviewers, using a tool specifically developed to gather information relevant to the primary research question of the scoping review. Where applicable, extracted data will capture pertinent information regarding the study's participants, context, and methods, along with key findings crucial to the review inquiry. The data extraction tool may undergo refinement to align more closely with new findings and all changes will be documented accordingly. In instances where disagreements arise between the reviewers, a third reviewer, the principal investigator of this scoping review, will intervene to efficiently reach a consensus and resolve these disputes.

Results Synthesis and Presentation

The extracted data will be organized according to emerging themes, while also examining study contexts, methods, results, and conclusions, particularly focusing on common evaluation models currently being used by instructors to evaluate clinical skills in undergraduate nursing students. A qualitative content analysis will follow to uncover key themes, patterns, and trends within the data to facilitate a comprehensive understanding of the research inquiry [16]. This entails the translation of results from quantitative research into classifications that can be integrated with qualitative narratives, guaranteeing a comprehensive synthesis of data across various types of sources included in the final review. The process will start with the initial coding of data, methodically grouping the data into categories according to pre-established and emerging themes. Subsequently, we will refine these preliminary codes into

XSL•FO
RenderX

broader themes by meticulously identifying and outlining the relationships between them. Our scoping review will assess the prevalence of key concepts, characteristics, and patterns across the included sources. Our objective is to recognize common themes and trends among these elements, offering a comprehensive examination of the data. To uphold rigor and confirm our findings, two researchers will independently conduct the result syntheses, thereby ensuring the strength and validity of our analysis. The concluding summary will integrate the results from various data sources to form a complete overview of the subject that improves the comprehension of the emergent themes. The results from this analysis will be displayed in both table and diagram format to provide a thorough summary of the scope and range of research regarding the methods of clinical skills evaluation used throughout undergraduate nursing education. Furthermore, a narrative summary will be created to incorporate these findings and align the data with the objectives of this review.

Ethical Considerations

This study is a scoping review of the literature and does not involve any human or animal participants. Therefore, research ethics approval was not sought and is not required for this study.

Results

Funding for the project was secured in June 2025. The scoping review is expected to commence in July 2025, with literature searches and the selection of relevant studies continuing throughout Fall 2025 (Figure 1). Findings are anticipated to be ready for publication submission by Winter 2026.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) flow diagram.

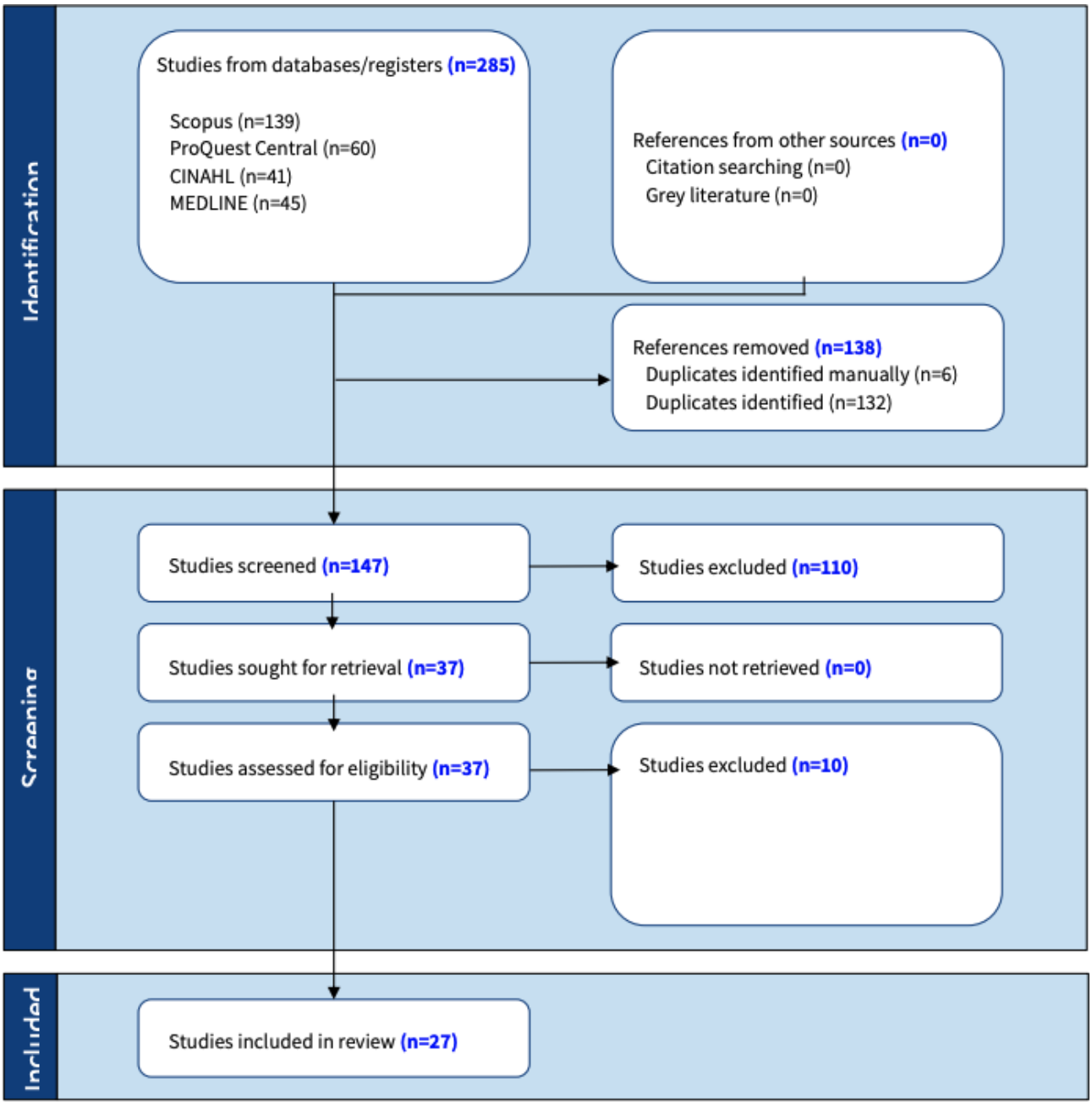

XSL•FO
RenderX

Discussion

Anticipated Findings

This scoping review will provide insights for nursing instructors teaching at postsecondary institutions who are involved in developing, revising, and improving undergraduate nursing curricula, particularly in the area of clinical practice preparation and student readiness evaluation. The results will help identify best practices related to the evaluation of undergraduate nursing students' clinical skills and inform the development of a more inclusive and accurate method of assessment, promoting positive student learning experiences and decreasing academic stress and test-taking anxiety. Furthermore, the findings can highlight gaps in current literature and identify priority areas for research within undergraduate nursing research and education.

Throughout this scoping review, we have recognized two main limitations that may affect the range and depth of our findings. First, due to constraints on translating resources, there is potentially language bias as only English sources will be considered. This limitation may lead to the omission of pertinent studies that have been conducted in other languages, which may provide additional insights or varied perspectives regarding undergraduate nursing education across different geographical regions. However, this scenario is unlikely since the only anticipated omissions would be French-language sources, and given the Canadian context, most literature is expected to be either bilingual or primarily available in English. Second, our review may also be subject to publication bias as studies yielding negative or null results are less likely to be published than those with statistically or clinically significant results.

Conclusion

Psychomotor skills are an essential component to undergraduate nursing education. However, national guidelines do not currently exist to guide nursing programs and educators on the best way to teach and evaluate nursing students to ensure competent and safe practice within the clinical setting [23,24]. This scoping review aims to provide insights that inform future nursing education by identifying gaps and emerging trends in current clinical skills testing models. By synthesizing evidence from the past decade, the review can highlight current practices and inform future innovative clinical skills testing models that effectively accommodate the needs of a growing nursing student population. How students perceive learning psychomotor skills can offer valuable insight for faculty members into how they can create laboratory environments that support effective student-centered learning and encourage active student involvement throughout the teaching process [25,26].

High-stakes evaluation methods require further exploration of the instruments, tools, or approaches used by educators based on validity, reliability, and objectivity. Combinations of assessment methods should also be considered by faculty to minimize instructor bias and the effect of ‘‘outlier’’ evaluations [27,28]. Using carefully designed testing methods with valid and reliable instruments for the evaluation of competency is the bridge to ensuring the delivery of quality health care and patient safety across all disciplines of professional nursing practice.

The findings of this scoping review will serve as a foundation for revamping existing clinical skills assessment frameworks, ensuring they remain practical and scalable. Specifically, the review can inform educators and administrators in designing future assessment methods that improve the student experience, reduce anxiety associated with clinical evaluations, enhance learning outcomes, and better prepare students for real-world clinical demands. Ultimately, the insights gained will contribute to developing a robust clinical skills testing model tailored to contemporary educational challenges.

Funding

This review is funded by the Toronto Metropolitan University Teaching and Learning Grant.

Data Availability

Data will be available upon reasonable request from the corresponding author.

Authors' Contributions

Conceptualization: KM
Funding acquisition: KM, TiL, ToL
Methodology: KM
Project administration: KM
Supervision: KM
Validation: KM
Writing - original draft: KM, TiL, ToL
Writing - review and editing: AA, DR, KM, TiL, ToL

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Search strategy.
[DOCX File, 14 KB - resprot_v15i1e79805_app1.docx]

Peer Review Report 1
Peer review report by Toronto Metropolitan University Teaching and Learning Grant Committee, Centre for Excellence in Learning and Teaching, Toronto Metropolitan University.
[PDF File, 213 KB - resprot_v15i1e79805_app2.pdf]

References

1. Wong CL, Chien WT, Waye MMY, Szeto MWC, Li H. Nursing students' perceived anxiety and heart rate variability in mock skill competency assessment. PLoS ONE 2023;18(10):e0293509. [doi: 10.1371/journal.pone.0293509] [Medline: 37883463]
2. Kaur Khaira M, Raja Gopal RL, Mohamed Saini S, Md Isa Z. Interventional strategies to reduce test anxiety among nursing students: a systematic review. Int J Environ Res Public Health 2023 Jan 10;20(2):1233. [doi: 10.3390/ijerph20021233] [Medline: 36673999]
3. Lavoie-Tremblay M, Sanzone L, Aubé T, Paquet M. Sources of stress and coping strategies among undergraduate nursing students across all years. Can J Nurs Res 2022 Sep;54(3):261-271. [doi: 10.1177/08445621211028076] [Medline: 34192949]
4. Monson NL, Higbee M, Brunger C, Ensign A, Gaul RA, Taylor N. Standardized skill pass-offs and nursing student confidence: a qualitative study. Teaching and Learning in Nursing 2021 Jul;16(3):194-198. [doi: 10.1016/j.teln.2021.04.003]
5. Baumann A, Crea-Arsenio M. The crisis in the nursing labour market: Canadian policy perspectives. Healthcare (Basel) 2023 Jul 6;11(13):1954. [doi: 10.3390/healthcare11131954] [Medline: 37444788]
6. Du L, Zhao L, Xu T, et al. Blended learning vs traditional teaching: the potential of a novel teaching strategy in nursing education - a systematic review and meta-analysis. Nurse Educ Pract 2022 Aug;63:103354. [doi: 10.1016/j.nepr.2022.103354] [Medline: 35580368]
7. Kerns C, Wedgeworth M. Barriers and facilitators experienced by undergraduate nursing faculty teaching clinical judgment: a qualitative study. SAGE Open Nurs 2024;10:23779608241274728. [doi: 10.1177/23779608241274728] [Medline: 39161937]
8. O'Rae A, Peters K, Shajani Z, Burkett J, Laing C. Improving the evaluation of clinical competence in undergraduate students, evidence and technology: An integrative review. J Prof Nurs 2025;56:19-25. [doi: 10.1016/j.profnurs.2024.10.003] [Medline: 39993896]
9. Byermoen KR, Eide T, Egilsdottir H, et al. Nursing students' development of using physical assessment in clinical rotation-a stimulated recall study. BMC Nurs 2022 May 10;21(1):110. [doi: 10.1186/s12912-022-00879-1] [Medline: 35538573]
10. Høegh-Larsen AM, Gonzalez MT, Reierson I, Husebø SIE, Hofoss D, Ravik M. Nursing students' clinical judgment skills in simulation and clinical placement: a comparison of student self-assessment and evaluator assessment. BMC Nurs 2023 Mar 9;22(1):64. [doi: 10.1186/s12912-023-01220-0] [Medline: 36894974]
11. Aldridge MD, Hummel F. Nursing students' perceptions of skills learning. Nurse Educ 2019;44(3):170-174. [doi: 10.1097/NNE.0000000000000569]
12. Morrell S, Ralph J, Giannotti N, Dayus D, Dennison S, Bornais J. Physical assessment skills in nursing curricula: a scoping review protocol. JBI Database System Rev Implement Rep 2019 Jun;17(6):1086-1091. [doi: 10.11124/JBISRIR-2017-003981] [Medline: 31082939]
13. Hernon O, McSharry E, MacLaren I, Carr PJ. The use of educational technology in teaching and assessing clinical psychomotor skills in nursing and midwifery education: a state-of-the-art literature review. J Prof Nurs 2023;45:35-50. [doi: 10.1016/j.profnurs.2023.01.005] [Medline: 36889892]
14. Coombs NM. Educational scaffolding: back to basics for nursing education in the 21st century. Nurse Educ Today 2018 Sep;68:198-200. [doi: 10.1016/j.nedt.2018.06.007] [Medline: 29960084]
15. Patelarou AE, Mechili EA, Ruzafa-Martinez M, et al. Educational interventions for teaching evidence-based practice to undergraduate nursing students: a scoping review. Int J Environ Res Public Health 2020 Aug 31;17(17):6351. [doi: 10.3390/ijerph17176351] [Medline: 32878256]
16. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
17. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
18. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
19. Benner P. From novice to expert. Am J Nurs 1982 Mar;82(3):402-407. [doi: 10.1097/00000446-198282030-00004] [Medline: 6917683]
20. The EndNote Team. EndNote (computer program) version 21. : Clarivate; 2013.

XSL•FO
RenderX

21. Munn Z, Aromataris E, Tufanaru C, et al. The development of software to support multiple systematic review types: the Joanna Briggs Institute System for the Unified Management, Assessment and Review of Information (JBI SUMARI). Int J Evid Based Healthc 2019 Mar;17(1):36-43. [doi: 10.1097/XEB.0000000000000152] [Medline: 30239357]
22. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
23. Reaves C, Martel M, Rose K. Teaching psychomotor skills in undergraduate nursing education: an integrative review. J Nurs Educ 2024 Jul;63(7):421-426. [doi: 10.3928/01484834-20240505-01] [Medline: 38979741]
24. Harvey G, MacLean H, El Hussein M, Zettel S, Benacchio D. Developing formative strategies to support undergraduate nursing student's learning in the lab. J Nurs Educ 2024 Aug;63(8):560-563. [doi: 10.3928/01484834-20240122-02] [Medline: 38302095]
25. Baran L, Özdemir H. Peer education versus traditional education on psychomotor skills and self-efficacy in nursing students: a single-blind randomized controlled trial. BMC Med Educ 2025 Oct 8;25(1):1371. [doi: 10.1186/s12909-025-08038-1] [Medline: 41057851]
26. Aldridge MD. Nursing students' perceptions of learning psychomotor skills: a literature review. Teaching and Learning in Nursing 2017 Jan;12(1):21-27. [doi: 10.1016/j.teln.2016.09.002]
27. Shinnick MA, Woo M. Comparison of simulation assessments: can they identify novice/studentnurses? Clin Simul Nurs 2020 Sep;46:40-49. [doi: 10.1016/j.ecns.2020.04.002]
28. Kardong-Edgren S, Oermann MH, Rizzolo MA, Odom-Maryon T. Establishing inter- and intrarater reliability for high-stakes testing using simulation. Nurs Educ Perspect 2017;38(2):63-68. [doi: 10.1097/01.NEP.0000000000000114] [Medline: 29194298]

Abbreviations

CASN: Canadian Association of Schools of Nursing
CNO: College of Nurses of Ontario
JBI: Joanna Briggs Institute
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols extension
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
RNAO: Registered Nurses' Association of Ontario
SUMARI: System for the Unified Management, Assessment, and Review of Information

Edited by J Sarvestan; submitted 28.Jun.2025; this is a non–peer-reviewed article;accepted 04.Feb.2026; published 13.Mar.2026.

Please cite as:
Metersky K, Lin T, Lam T, Al-Hamad A, Romaniuk D
Evaluation Models for Undergraduate Nursing Clinical Skills: A Scoping Review Protocol
JMIR Res Protoc 2026;15:e79805
URL: https://www.researchprotocols.org/2026/1/e79805
doi:10.2196/79805

© Kateryna Metersky, Tommy Lin, Tina Lam, Areej Al-Hamad, Daria Romaniuk. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Factors Associated With Evolution of the Use of Medical Assistance in Dying: Protocol for a Scoping Review

Isabelle Marcoux[1*], PhD; David Lavoie[2*], PsyD/PhD; Amy Bergeron[3], MSI; Alexandra Beaudin[1], MSc; Simon Lemyre[2], MA; Valérie Bourgeois-Guérin[4], PhD; Gina Bravo[5], PhD; Maude Hébert[6], PhD; Chantale Simard[7], PhD; Deborah Ummel[8], PhD; Catherine Perron[9], PhD; Marie-Eve Bouthillier[2], PhD

[1]Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, 125 University Private, Ottawa, ON, Canada

[2]Department of Family and Emergency Medicine, Faculty of Medicine, Université de Montréal, Montréal, QC, Canada

[3]Health Library, Direction of Libraries, Université de Montréal, Montréal, QC, Canada

[4]Department of Psychology, Université du Québec à Montréal, Montréal, QC, Canada

[5]Department of Community Health Sciences, Faculty of Medicine and Health Sciences, Université de Sherbrooke, Sherbrooke, QC, Canada

[6]Département des sciences infirmières, Université du Québec à Trois-Rivières, Trois-Rivières, QC, Canada

[7]Department of Health Sciences, Nursing Sciences Module, Université du Québec à Chicoutimi, Chicoutimi, QC, Canada

[8]Institute of Psychology, Faculty of Social and Political Sciences, University of Lausanne, Lausanne, Switzerland

[9]École de travail social, Faculté des lettres et sciences humaines, Université de Sherbrooke, Sherbrooke, QC, Canada

[*]these authors contributed equally

Corresponding Author:
Isabelle Marcoux, PhD
Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, 125 University Private, Ottawa, ON, Canada

Abstract

Background: As of June 2025, medical assistance in dying (MAiD) is allowed in over 25 jurisdictions across 12 countries, with varying rates of requests and provision. Hypotheses have been suggested to explain these variations, but they are rarely backed up by empirical evidence. As more jurisdictions consider legalizing MAiD, it is important to better understand what factors may explain the evolution of the use of MAiD worldwide with a systematic approach.

Objective: This scoping review aims to map the available evidence regarding the factors associated with the evolution of the use of MAiD in jurisdictions where it is allowed.

Methods: The scoping review will follow the Joanna Briggs Institute methodology and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. Ten electronic databases (including MEDLINE, Embase, CINAHL Complete, and APA PsycInfo) and additional gray literature sources will be searched from inception to the present. This scoping review will consider multiple types of publications (eg, primary studies, research syntheses, and government reports) and will report factors associated with the use of MAiD for people who have requested or died by MAiD in jurisdictions that have allowed it for at least 5 years. Publications in English, French, Spanish, German, or Dutch will be included. Screening for assessment against the inclusion criteria and data extraction will be carried out independently by pairs of reviewers. Findings will be presented in a narrative format and mapped into tables and graphs to address the review aims.

Results: The database search for scientific publications was completed in March 2025. A total of 8570 publications were identified after removing duplicates. As of October 2025, title and abstract screening is complete, with 216 articles retained for the next stage. The full-text review is underway and scheduled to be completed by December 2025, and results are expected to be submitted for publication in 2026.

Conclusions: MAiD is gaining legal and policy attention worldwide, with wide variation in rates of request and provision over time across jurisdictions. This scoping review will contribute to mapping and synthesizing evidence on the factors that may explain these variations.

Trial Registration: OSF Registries TKNAM; https://osf.io/tknam/overview

International Registered Report Identifier (IRRID): DERR1-10.2196/85963

(JMIR Res Protoc 2026;15:e85963) doi:10.2196/85963

KEYWORDS

medical assistance in dying; MAiD; euthanasia; assisted suicide; legislation; rates; deaths; practices; requests; influencing factors

XSL•FO
RenderX

Introduction

Background

In the past decades, there has been a growing movement in various parts of the world to legalize assisted death practices. Twenty years ago, only a few jurisdictions allowed such practices, including the state of Oregon, Belgium, Colombia, the Netherlands, Luxembourg, and Switzerland. However, since 2015, assisted death practices have been authorized in additional US states and in other countries, such as Canada, New Zealand, Spain, Austria, and Portugal, as well as in most Australian states [1]. Assisted death practice legislation across jurisdictions generally shares core eligibility criteria (eg, minimum age and health conditions) and procedural safeguards (eg, physician or authorized person consultations and reporting processes), but there are also variations [2]. Some regulations of assisted death practices are more general and liberal, whereas others are rather specific and restrictive. In addition, some jurisdictions broadened their eligibility criteria after the law came into force to allow more patients to access the procedure. In Canada, for example, the passing of Bill C-7 in 2021 expanded access to some individuals whose death was not "reasonably foreseeable" [3]. As legislative proposals and court decisions on the issue become more frequent in many legislations, policymakers and researchers are questioning why people request and die by assisted death practices [4].

Different terms are used to define assisted death practices depending on the legislations and regulations. "Euthanasia" is defined as the act of intentionally ending the life of a patient by a qualified health care practitioner through medical means at that patient's explicit request, whereas "physician-assisted suicide" (PAS) refers to the provision or prescription of drugs by a health care practitioner for a patient to end their own life [2]. Some jurisdictions allow only euthanasia, whereas others permit only PAS or both practices. This protocol will use "medical assistance in dying" (MAiD) for both the literature covering euthanasia and PAS and other regulated terms (eg, "medical aid in dying," "termination of life," and "voluntary assisted dying"), but specific terms will be used when there are issues related to either of these practices. Therefore, unless otherwise specified, we will use "MAiD" as the overarching term in this review.

Reported numbers of deaths by MAiD vary greatly over time between jurisdictions where such practices are allowed. In Belgium, where the Act on Euthanasia took effect in 2002, the number of reported deaths by euthanasia increased from 0.2% of all deaths in 2003 [5] to 3.1% in 2023 [6]. In comparison, only 6 years after the legalization of MAiD in Canada, the proportion of all deaths attributed to MAiD rose from 0.7% in the first full year of legalization in 2016-2017 (as this proportion was not available at that time, we made our own calculation based on the total number of MAiD cases reported in the first and second interim updates [7,8] and on mortality data drawn from Statistics Canada [9]) to 4.7% of all deaths in 2023 [10]. Furthermore, within the same country, the rates of MAiD provisions vary significantly across states or provinces. For instance, in 2022, Canadian provinces experienced different variations in MAiD provisions: Saskatchewan had the lowest year-over-year increase (4.0%); Quebec, which has its own law regulating MAiD that came into force 1 year before the federal law [11], saw the highest growth rates (45.5%); whereas Manitoba reported a decline (−9.0%). Regional differences are also notable within a given province: in Quebec, the proportion of deaths by MAiD reported by different health regions in 2023 ranged from less than 4% to over 10.5% [12].

Some assumptions have been formulated to explain the increase in MAiD rates in certain jurisdictions. Boer [13] describes the development of the practice of MAiD in the Netherlands from 1968 onward, noting that the number of deaths by euthanasia tripled between 2007 and 2021, a period characterized by the expansion of eligibility criteria and the implementation of mobile teams to improve access to euthanasia for those without a prior physician-patient relationship. Pullman [14] explores factors that may explain Canada's growing and higher rates of MAiD deaths compared to California's despite their respective laws taking effect in the same year (2016). He suggests that, in Canada, eligibility criteria are more ambiguous, their interpretation is more flexible, and the method of provision (injection rather than prescription of lethal drugs) is more direct. For others, the high rise in MAiD deaths in Canada may be clinician dependent, suggesting that most practices are performed by a few highly active clinicians [15].

What This Scoping Review Will Contribute

To date, most of the published work is based on opinions or demonstrations focused on specific explanations of the different rates without aiming to be comprehensive. Some reviews address parts of the issue—for example, they focus on factors influencing requests [16], decisions [17], or access to MAiD [18]; on one specific factor [19,20]; on barriers to and facilitators of engaging in the practice of MAiD [21]; or on challenges related to implementing MAiD from a health system perspective [22]. One systematic review published in 2013 [23] provides valuable background for our topic, although it focuses on individual characteristics (eg, sex, age, marital status, educational level, and condition leading to PAS or euthanasia) associated with numbers of and variations in assisted death practices in different countries where these practices were legal. Thus, other potential associated factors such as organizational, sociocultural, or legal context were not considered.

These reviews show that there are available data on the subject. However, after a preliminary search of MEDLINE, the Cochrane Database of Systematic Reviews, *JBI Evidence Synthesis*, and PROSPERO, we found no reviews (published or planned) with the specific aims to identify factors associated with the evolution of the use of MAiD (as defined below). To address this gap, the proposed scoping review will provide a comprehensive overview of the available evidence regarding these factors. Better knowledge of the available data regarding these factors could inform decision-makers and health care professionals, as well as jurisdictions contemplating the legalization of these practices.

This proposed scoping review is part of a larger research study funded by the government of Quebec (Canada) [24] to explore the rapid rise in MAiD rates in the province. From April 1, 2023, to March 31, 2024, a total of 7.3% of all deaths in Quebec were

XSL•FO
RenderX

a result of MAiD, the highest rate in jurisdictions where it is legal [25]. In the call for proposals, the government of Quebec also expressed interest in learning from other jurisdictions where MAiD is allowed. Therefore, to provide a comprehensive picture of the factors influencing the evolution in the use of MAiD from an international comparison perspective, the proposed scoping review aims to map the available evidence from jurisdictions where such practices have been allowed for at least 5 years.

Review Questions

Main Question

The main question addressed by this scoping review is as follows: What factors are associated with the evolution in the use of MAiD in jurisdictions where such practices have been allowed for at least 5 years?

Subquestions

The subquestions are as follows: (1) What types of evidence sources are available on factors influencing the evolution in the use of MAiD (eg, primary research, evidence syntheses, and discussion articles)? (2) Who provides the evidence on factors associated with the evolution in the use of MAiD (eg, patients, relatives, MAiD practitioners, governments, or regulatory commissions)? (3) What are the target populations (adults who have requested or died by MAiD) for the available evidence? (4) What factors were identified as influencing the evolution in the use of MAiD?

Findings from this scoping review will also be used to inform a realist review [26], which will explore what are the effects of MAiD legislations and public policies on MAiD practices in different jurisdictions and why they produce such effects, for whom, and in which circumstances.

Methods

Study Design and Review Team

The proposed scoping review will be conducted in accordance with the Joanna Briggs Institute (JBI) methodology for scoping reviews [27] and reported in line with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews; Checklist 1) [28]. This protocol has been registered in Open Science Framework [29]. The interdisciplinary review team consists of researchers in psychology (IM, DL, A Beaudin, DU, and VB-G), nursing (CS and MH), mathematics (GB), and clinical ethics (CP, MEB, and SL), as well as a professional librarian with expertise in information retrieval for knowledge synthesis (A Bergeron). Research team members discussed and revised the drafted protocol, bringing together scientific, clinical, and ethical expertise.

It is worth noting that, prior to this scoping review protocol, we considered conducting a mixed methods systematic review (MMSR). However, in an MMSR, nonempirical evidence (eg, expert opinion–based pieces, government reports, and discussion papers) and literature reviews are excluded. While these sources may not fit the MMSR criteria, they could provide valuable information for our research question. Given the aims of our review, a scoping review offers the most suitable framework [30,31].

Eligibility Criteria

According to the JBI methodology for scoping reviews, the eligibility criteria are based on the population, concept, and context format, as well as the type of sources.

Population

The population of interest is adults who have requested or died by MAiD. However, data concerning this phenomenon may be reported by a variety of populations, such as the patients themselves; relatives; substitute decision-makers, surrogates, or proxies; MAiD providers (physicians, nurse practitioners, or other authorized practitioners); or other health care professionals who were involved in the process of MAiD requests, practices, or deaths (eg, social workers, psychologists, or spiritual care advisors). Studies conducted with key informants, such as members of a regulatory organization for MAiD, representatives of a professional body, or any other type of informant (eg, decision-makers or nonclinical leaders), will also be considered for inclusion. Publications that investigate factors related to the use of MAiD through secondary data (eg, analysis based on death statistics and government reports) will also be included.

Concept

The use of MAiD is conceptualized as MAiD requests, practices, or deaths. MAiD requests are demands made to any health care professional by a patient and that could lead to the assessment of the patient's eligibility. Practices include the assessment of a patient for MAiD or the provision of MAiD to an eligible patient. Assessment refers to a clinician completing an evaluation to ensure that the patient meets all eligibility criteria for MAiD in their jurisdiction. Provision of MAiD is the injection or prescription of medications to intentionally end the life of the patient who meets the eligibility criteria at the person's request. Evolution is conceptualized as the development (change or stability over time) of the use of MAiD.

Factors influencing the evolution in the use of MAiD are defined as a characteristic that can be perceived, observed, measured, or reported. Drawing on the ecological systems model by Bronfenbrenner [32] (from the microsystem to the macrosystem), these factors are identified and categorized into 4 main types: individual characteristics, clinical or organizational factors, sociocultural context, and laws and policies.

Exclusion criteria are publications investigating only passive euthanasia or suicide, focusing solely on opinions or attitudes toward MAiD without an explicit association with the evolution in the use of MAiD, or studying only the factors associated with the use of MAiD (at one point in time) without considering its evolution.

Context

The inclusion criterion of time (selecting jurisdictions where MAiD has been legally allowed for at least 5 years) is important to ensure that sufficient data are available to observe meaningful trends and variations over time. As of December 3, 2024 (date of our database search for preliminary pilot selection;

XSL•FO
RenderX

Multimedia Appendix 1), 16 jurisdictions were considered for inclusion: Australia (Victoria); Belgium; Canada; Colombia; Luxembourg; the Netherlands; Switzerland; and California, Maine, Montana, New Jersey, Hawaii, Colorado, Oregon, Vermont, Washington, and the District of Columbia in the United States. All types of settings are included (eg, hospitals, long-term care facilities, private dwellings, and hospices).

Existing prelegalization data from these jurisdictions will also be included. For example, in the Netherlands, although the law was passed in 2001 and came into force in 2002, the practices had been regulated by the medical authorities since the early 1990s. Relevant data have been published (prior to legalization [33-35]) and are important to capture.

Types of Sources

To provide a thorough understanding of the topic, this review will incorporate all available sources of information on factors influencing the use of MAiD, including both peer-reviewed publications and gray literature. We will consider reports of primary or secondary data sources using any type of design (quantitative, qualitative, or mixed methods), as well as reviews and documents that the JBI describes as "textual evidence" (eg, expert opinion–based pieces, government or institutional policies and/or reports, unpublished [or gray] literature, and discussion papers) [36]. Research protocols, book reviews, media publications, and conference proceedings will be excluded.

Search Strategy

The primary search strategy for MEDLINE was developed iteratively by one author (A Bergeron), a medical librarian, in collaboration with other authors (DL and IM), as recommended in the *JBI Manual for Evidence Synthesis* [27], and reported according to the PRISMA-S (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for reporting literature searches in systematic reviews) guidelines (Checklist 2) [37]. The key concepts for the strategy were identified from an initial set of target articles as well as several test searches: (1) MAiD, (2) requests for or provision of MAiD or statistical data pertaining to cases, and (3) jurisdictions in which MAiD has been allowed for at least 5 years. No restrictions were applied based on publication date. The full search strategies will be reported in the final manuscript according to PRISMA-S guidelines.

Although the evolution in the use of MAiD is central to this scoping review, our preliminary tests revealed that the terminology describing this concept is often heterogeneous and subtle. Furthermore, including this concept in the search strategy sometimes excluded relevant articles. Consequently, it was not retained for the initial search strategy but, rather, serves as an inclusion criterion during the screening process.

The preliminary strategy was revised by the team and additionally reviewed by a second librarian using the PRESS (Peer Review of Electronic Search Strategies) guidelines, as additionally recommended by the JBI [26]. The MEDLINE search strategy was then finalized and has been adapted for other sources using database-appropriate subject headings and platform-specific syntax. In total, 10 electronic databases were searched: MEDLINE (Ovid), Embase (Ovid), CINAHL Complete (EBSCO), APA PsycInfo (Ovid), Web of Science Core Collection, Social Sciences Abstracts (EBSCO), Sociological Abstracts (ProQuest), Dissertations and Theses Closed (ProQuest), and Persée and Érudit (Multimedia Appendix 1). A process is currently being developed to search for gray literature sources for additional documents (eg, government or institutional policies and/or reports and position papers). Subsequently, backward and forward citation tracking will be performed for references in the initial corpus of included articles using Web of Science. References of and articles citing the included articles will be exported to Covidence (Veritas Health Innovation) and screened for additional articles.

Study and Source of Evidence Selection

Following the execution of the full search strategy, references were exported to Covidence, and duplicates were removed using its automatic detection feature. Following a pilot test of 100 records (conducted independently by DL, IM, A Beaudin, and SL), titles and abstracts were screened by at least 2 independent reviewers (DL, A Beaudin, and SL) for assessment against the inclusion criteria, which has now been completed. Publications in English, French, Spanish, German, or Dutch were considered for inclusion. At each stage (past, current, and next) of the selection process, any disagreements that arise between reviewers are resolved through discussion or by including a third party (IM). The full texts of selected citations are currently being assessed in detail against the inclusion criteria by at least 2 independent reviewers. Reasons for the exclusion of sources of evidence at the full-text stage are recorded and will be reported in full, as will the results of the search and the study inclusion process, in the final scoping review and will be presented in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

Data Extraction

Following selection, data will be extracted directly within Covidence by independent reviewers using a data extraction tool developed by the team (Textbox 1). Data will include author; title; source of publication; purpose of study; type of study; date when data were collected; jurisdiction studied; practices studied (eg, MAiD and PAS); settings; and specific details relevant to the review subquestions, such as (1) the type of evidence, (2) the participants providing the information, (3) the target population studied, and (4) factors influencing the evolution in the use of MAiD, which will be divided into 4 main categories (individual characteristics, clinical or organizational factors, sociocultural context, and laws and policies) in accordance with the ecological systems model by Bronfenbrenner [32]. The extraction tool will be pilot-tested on a minimum of 5 articles and then revised. If necessary, authors of the included articles will be contacted to request further information during data extraction. Data extraction, as well as the analysis and presentation of results, will follow the JBI's recommendations [38].

XSL•FO
RenderX

Textbox 1. Preliminary categories for the draft data extraction tool.

- Authors and year
- Title
- Source of publication (eg, scientific journal or government report)
- Purpose of the publication
- Type of evidence (eg, primary research, evidence synthesis, or discussion article)
- Type of study (quantitative, qualitative, or mixed methods)
- Dates of data collection
- Jurisdictions studied
- Performed practices (eg, medical assistance in dying [MAiD], physician-assisted suicide, euthanasia, or voluntary assisted dying)
- Settings (eg, hospitals, long-term care facilities, private dwellings, hospices, or other health care settings)
- Participants who provide the evidence on the factors (eg, patients, relatives, MAiD practitioners, governments, or regulatory commissions)
- Target population (adults who have requested or died by MAiD)
- Factors influencing the evolution of MAiD (individual characteristics, clinical or organizational factors, sociocultural context, and laws and policies)

Data Analysis and Presentation

Findings will be presented in a narrative format and mapped into tables and graphs to address the review aims (eg, a world map of included jurisdictions with numbers of related publications and a chart of included types of evidence and types of performed practices). In particular, reported factors influencing the evolution of MAiD (categorized in accordance with our ecological systems model) will be analyzed and presented graphically while taking into consideration the types of use of MAiD (either requests, practices, or deaths), as well as the types of evidence (eg, primary research, evidence syntheses, and discussion articles). However, as mentioned by Peters et al [39], because of the broad and iterative nature of the scoping review, it is likely that potentially new relevant information will emerge from searches, screening, and extraction. Accordingly, the search, extraction forms, and presentation of data may be modified and expanded during the review process. All deviations from the protocol will be reported in the review.

Results

The project was funded in July 2024. We completed the database search in March 2025 and identified 8570 publications after removing duplicates via Covidence. As of October 2025, the screening of titles and abstracts is complete, resulting in the inclusion of 216 articles for the next stage. The full-text review selection is currently underway and scheduled to be completed by December 2025, followed by the data extraction and analysis phases. The full article is planned for submission to a peer-reviewed journal in 2026.

Discussion

Expected Findings

To the best of our knowledge, this will be the first review to map the available evidence concerning the factors associated with the evolution in the use of MAiD considering all jurisdictions where MAiD has been legally allowed for at least 5 years as of December 2024. This scoping review is also part of a larger research project in Quebec, where MAiD accounts for the highest proportion of deaths among jurisdictions where it is allowed. The scoping review will contribute to the project's overarching objective, which is to explain the rapid increase in MAiD in the province, by providing an overview of associated factors while considering the types and characteristics of the existing evidence. Findings from this scoping review will also inform a realist review on how MAiD laws and policies influence practices, for whom, how, and in what contexts. This will be possible by making comparisons of the included jurisdictions' particularities, which will include specific aspects regarding their laws, their health care systems, and the social acceptability of MAiD, among other aspects. A better knowledge of these factors can also be helpful for health care practitioners, researchers, and policymakers in assessing complex end-of-life care situations and services in their settings and jurisdictions.

Limitations

Our study will face limitations related to its focus and broad concepts. First, the decision to include only jurisdictions where MAiD has been allowed for at least 5 years may exclude recent but potentially relevant experiences from other jurisdictions. However, this criterion should enable us to assess trends in the use of MAiD, which requires a significant time frame. Second, factors associated with the use of MAiD but not related to its evolution are excluded, which may limit our understanding of potentially important information less likely to be reported as evolutionary use factors, such as family decision-making. Third, the predetermined conceptualization of factors related to the evolution in the use of MAiD may limit the breadth and depth of analysis as this approach is guided by the ecological systems model by Bronfenbrenner [32] and purposefully encompasses variables at different ecological systems (eg, individual, organizational, societal, and legislative). Conversely, this approach may strengthen the identification and analysis of

XSL•FO
RenderX

literature gaps based on the types of studied factors, thereby guiding future research.

Conclusions

Since the 90s, and more particularly for the past 10 years, MAiD has been increasingly debated, and related legislation has been adopted in different countries in America, Europe, and Oceania. The frequency of requests and provision of MAiD and the evolution over time vary widely between jurisdictions. The synthesis of available evidence on factors explaining these variations will be greatly helpful to inform policymakers, clinicians, researchers, and all those concerned about assisted death, end-of-life care, and palliative care.

Acknowledgments

The authors would like to thank University of Montreal librarians for their peer review of the initial MEDLINE strategy (Monique Clar and Thien Sa Hoang) and of the Social Sciences Abstracts and Sociological Abstracts search strategies (Caroline Patenaude). The authors also thank the members of the Interdisciplinary Research Consortium on Medical Assistance in Dying, who are not listed as authors of this paper but who have contributed to the larger research project of which this review is part: Michel Dorval, David Lussier, Mathieu Moreau, Ghislaine Rouly, Louise Bernier, Claude Julie Bourque, Bruno Gagnon, Dominique Girard, Michelle Giroux, Mona Gupta, Bertrand Lavoie, Sabrina Lessard, Maude Lévesque, Maryse Soulières, Omid Hosseini, Émilie Allard, Elaine Champagne, Josée Chénard, Marc Dumas, Isabelle Dumont, Nicolas Dupré, Sylvie Fortin, Nathalie Gaucher, Claire Godard-Sebillotte, Diane Guay, Manon Guay, Johanne Hébert, Philippe Laperle, Sylvie Lapierre, Jérôme Leclerc-Loiselle, Marianne Olivier-d'Avignon, Antoine Payot, Ariane Plaisance, Diane Tapp, Bilkis Vissandjee, Carl Bromwich, Andréanne Côté, Thania Crnich-Côté, Dominique Dion, Geneviève Lamothe, Marie-Hélène Marchand, Guy Morissette, Olivia Nguyen, Élodie Petit, Justin Sanders, Jacinthe Dupuis, Patrick Durivage, Sandra Smele, Emilie Richard-Frève, Jean-Thomas Chouinard, Marie-Alexandra Gagné, France Lacharité, Isabelle Martineau, Marta Martisella, Antoine Bédard, Francine Carmichael, Eliana Close, Luc Deliens, James Downar, Jocelyn Downie, Andrea Frolic, Ralf Jox, Thaddeus Pope, and Agnes Van der Heide.

Funding

The planned scoping review is part of a larger research project funded through a partnership between the Quebec Research Fund (Fonds de recherche du Québec) and the Quebec Ministry of Health and Social Services (grant 2025-0AMMR-355830). The project initially received a pilot grant from the Réseau québécois de recherche en soins palliatifs et de fin de vie and another from the Integrated Health and Social Services Centre of Laval. The funders were not involved in the decision to conduct a review related to factors associated with the use of medical assistance in dying and its evolution or to register and publish the review protocol, nor did the funders play a role in designing the study or writing the manuscript.

Disclaimer

The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: proofreading, editing, and translation. The GenAI tools used were Google Gemini 3 and DeepL (DeepL SE). Responsibility for the final manuscript lies entirely with the authors. DeepL was used for French to English translation (words and phrases), whereas Gemini 3 helped improve the phrasing of content drafted directly in English.

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

All authors were actively involved in preparing this protocol through repeated team meetings, including formulating the research questions, determining the eligibility criteria, and proposing keywords for searching the literature. From input provided by the research team, A Bergeron developed the search strategies. DL and IM led the development of the protocol and drafted the manuscript, assisted by A Bergeron. All authors provided feedback on the manuscript and approved its final version.

Members of the Interdisciplinary Research Consortium on Medical Assistance in Dying: Michel Dorval, David Lussier, Mathieu Moreau, Ghislaine Rouly, Louise Bernier, Claude Julie Bourque, Bruno Gagnon, Dominique Girard, Michelle Giroux, Mona Gupta, Bertrand Lavoie, Sabrina Lessard, Maude Lévesque, Maryse Soulières, Omid Hosseini, Émilie Allard, Elaine Champagne, Josée Chénard, Marc Dumas, Isabelle Dumont, Nicolas Dupré, Sylvie Fortin, Nathalie Gaucher, Claire Godard-Sebillotte, Diane Guay, Manon Guay, Johanne Hébert, Philippe Laperle, Sylvie Lapierre, Jérôme Leclerc-Loiselle, Marianne Olivier-d'Avignon, Antoine Payot, Ariane Plaisance, Diane Tapp, Bilkis Vissandjee, Carl Bromwich, Andréanne Côté, Thania Crnich-Côté, Dominique Dion, Geneviève Lamothe, Marie-Hélène Marchand, Guy Morissette, Olivia Nguyen, Élodie Petit, Justin Sanders, Jacinthe Dupuis, Patrick Durivage, Sandra Smele, Emilie Richard-Frève, Jean-Thomas Chouinard, Marie-Alexandra Gagné, France Lacharité, Isabelle Martineau, Marta Martisella, Antoine Bédard, Francine Carmichael, Eliana Close, Luc Deliens, James Downar, Jocelyn Downie, Andrea Frolic, Ralf Jox, Thaddeus Pope, and Agnes Van der Heide.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy for MEDLINE (Ovid).
[DOCX File, 25 KB - resprot_v15i1e85963_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 29 KB - resprot_v15i1e85963_app2.docx]

Checklist 2
PRISMA-S checklist.
[DOCX File, 26 KB - resprot_v15i1e85963_app3.docx]

References

1. Voluntary assisted dying. End of Life Law in Australia. URL: https://end-of-life.qut.edu.au/assisteddying [accessed 2025-06-25]
2. Mroz S, Dierickx S, Deliens L, Cohen J, Chambaere K. Assisted dying around the world: a status quaestionis. Ann Palliat Med 2021 Mar;10(3):3540-3553. [doi: 10.21037/apm-20-637] [Medline: 32921084]
3. Bill C-7: an Act to amend the Criminal Code (medical assistance in dying). Parliament of Canada. 2021. URL: https://www.parl.ca/Content/Bills/432/Government/C-7/C-7_4/C-7_4.PDF [accessed 2025-10-07]
4. Downar J, MacDonald S, Buchman S. What drives requests for MAiD? CMAJ 2023 Oct 16;195(40):E1385-E1387. [doi: 10.1503/cmaj.230259] [Medline: 37844931]
5. Premier rapport aux chambres législatives 22 septembre 2002 - 31 décembre 2003. Commission fédérale de Contrôle et d'Évaluation de l'Euthanasie. URL: https://organesdeconcertation.sante.belgique.be/sites/default/files/documents/federale_controle_en_evaluatiecommissie_euthanasie-fr/14276508_fr.pdf [accessed 2025-10-07]
6. Euthanasie – chiffres de l'année 2023. Commission fédérale de Contrôle et d'Évaluation de l'Euthanasie. URL: https://organesdeconcertation.sante.belgique.be/sites/default/files/documents/cfcee-communiquepresse20240227-chiffreseuthanasie-2023.pdf [accessed 2025-10-07]
7. Interim update on medical assistance in dying in Canada June 17 to December 31, 2016. Government of Canada. URL: https://www.canada.ca/en/health-canada/services/publications/health-system-services/medical-assistance-dying-interim-report-dec-2016.html [accessed 2025-10-07]
8. 2nd interim report on medical assistance in dying in Canada. Government of Canada. 2017. URL: https://www.canada.ca/en/health-canada/services/publications/health-system-services/medical-assistance-dying-interim-report-sep-2017.html [accessed 2025-10-07]
9. Deaths, by month. Statistics Canada. 2026. URL: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1310070801 [accessed 2026-04-04]
10. Fifth annual report on medical assistance in dying in Canada 2023. Health Canada. 2023 URL: https://www.canada.ca/content/dam/hc-sc/documents/services/publications/health-system-services/annual-report-medical-assistance-dying-2023/annual-report-medical-assistance-dying-2023.pdf [accessed 2026-04-04]
11. s-32.0001 - Act respecting end-of-life care. Légis Québec. 2025. URL: https://www.legisquebec.gouv.qc.ca/fr/document/lc/s-32.0001 [accessed 2025-10-07]
12. Fourth annual report on medical assistance in dying in Canada 2022. : Health Canada; 2023 URL: https://www.canada.ca/content/dam/hc-sc/documents/services/medical-assistance-dying/annual-report-2022/annual-report-2022.pdf [accessed 2026-04-04]
13. Boer TA. Assisted dying in the Netherlands: recent legal and religious considerations. Rev Derecho Relig 2022;17:185-198 [FREE Full text]
14. Pullman D. Slowing the slide down the slippery slope of medical assistance in dying: mutual learnings for Canada and the US. Am J Bioeth 2023 Nov;23(11):64-72. [doi: 10.1080/15265161.2023.2201190] [Medline: 37166283]
15. Lyon C, Lemmens T, Kim SYH. Canadian medical assistance in dying: provider concentration, policy capture, and need for reform. Am J Bioeth 2025 May;25(5):6-25. [doi: 10.1080/15265161.2024.2441695] [Medline: 39791998]
16. Castelli Dransart DA, Lapierre S, Erlangsen A, et al. A systematic review of older adults' request for or attitude toward euthanasia or assisted-suicide. Aging Ment Health 2021 Mar;25(3):420-430. [doi: 10.1080/13607863.2019.1697201] [Medline: 31818122]
17. Xu H, Stjernswärd S, Glasdam S, Fu C. Circumstances affecting patients' euthanasia or medically assisted suicide decisions from the perspectives of patients, relatives, and healthcare professionals: a qualitative systematic review. Death Stud 2024;48(4):326-351. [doi: 10.1080/07481187.2023.2228730] [Medline: 37390123]

XSL•FO
RenderX

18. Hewitt J, Wilson M, Bonner A, Bloomer MJ. Factors that influence access to medical assistance in dying services: an integrative review. Health Expect 2024 Oct;27(5):e70058. [doi: 10.1111/hex.70058] [Medline: 39415630]
19. Corcoran E, Bird M, Batchelor R, Ahmed N, Nowland R, Pitman A. The association between social connectedness and euthanasia and assisted suicide and related constructs: systematic review. BMC Public Health 2024 Apr 16;24(1):1057. [doi: 10.1186/s12889-024-18528-4] [Medline: 38627694]
20. Rahimian Z, Rahimian L, Lopez-Castroman J, et al. What medical conditions lead to a request for euthanasia? A rapid scoping review. Health Sci Rep 2024 Mar;7(3):e1978. [doi: 10.1002/hsr2.1978] [Medline: 38515545]
21. Légère K, Doucet S, Luke A, Goudreau A. Barriers and facilitators for engaging in the practice of medical assistance in dying among providers in Canada: a scoping review protocol. JBI Evid Synth 2024 Feb 1;22(2):273-280. [doi: 10.11124/JBIES-22-00278] [Medline: 36734643]
22. Fujioka JK, Mirza RM, Klinger CA, McDonald LP. Medical assistance in dying: implications for health systems from a scoping review of the literature. J Health Serv Res Policy 2019 Jul;24(3):207-216. [doi: 10.1177/1355819619834962] [Medline: 31291767]
23. Steck N, Egger M, Maessen M, Reisch T, Zwahlen M. Euthanasia and assisted suicide in selected European countries and US states: systematic literature review. Med Care 2013 Oct;51(10):938-944. [doi: 10.1097/MLR.0b013e3182a0f427] [Medline: 23929402]
24. Quebec funds a project to better understand the use of medical assistance in dying. CityNews Montreal. 2024. URL: https://montreal.citynews.ca/2024/03/27/quebec-funds-project-medical-assistance-dying/ [accessed 2025-10-07]
25. Rapport annuel d'activités: du 1er avril 2023 au 31 mars 2024. Commission Sur les Soins de Fin de Vie. 2024 URL: https://csfv.gouv.qc.ca/fileadmin/docs/rapports_annuels/csfv_rapport_activites_2023-2024.pdf [accessed 2026-04-04]
26. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist review--a new method of systematic review designed for complex policy interventions. J Health Serv Res Policy 2005 Jul;10 Suppl 1:21-34. [doi: 10.1258/1355819054308530] [Medline: 16053581]
27. Peters MD, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI; 2024. [doi: 10.46658/JBIMES-24-09]
28. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
29. Lavoie D, Marcoux I, Bergeron A, Perron C, Bouthillier ME, et al. Factors associated with medical assistance in dying and its evolution: a scoping review protocol. JMIR Research Protocols. Preprint posted online on Jun 2, 2025. [doi: 10.17605/OSF.IO/TKNAM]
30. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
31. Peters MDJ, Godfrey C, McInerney P, et al. Best practice guidance and reporting items for the development of scoping review protocols. JBI Evid Synth 2022 Apr 1;20(4):953-968. [doi: 10.11124/JBIES-21-00242] [Medline: 35102103]
32. Bronfenbrenner U. The Ecology of Human Development: Experiments by Nature and Design: Harvard University Press; 1981. [doi: 10.2307/j.ctv26071r6]
33. Onwuteaka-Philipsen BD, Brinkman-Stoppelenburg A, Penning C, de Jong-Krul GJF, van Delden JJM, van der Heide A. Trends in end-of-life practices before and after the enactment of the euthanasia law in the Netherlands from 1990 to 2010: a repeated cross-sectional survey. Lancet 2012 Sep 8;380(9845):908-915. [doi: 10.1016/S0140-6736(12)61034-4] [Medline: 22789501]
34. van Alphen JE, Donker GA, Marquet RL. Requests for euthanasia in general practice before and after implementation of the Dutch Euthanasia Act. Br J Gen Pract 2010 Apr;60(573):263-267. [doi: 10.3399/bjgp10X483931] [Medline: 20353671]
35. van der Heide A, van Delden JJM, Onwuteaka-Philipsen BD. End-of-life decisions in the Netherlands over 25 years. N Engl J Med 2017 Aug 3;377(5):492-494. [doi: 10.1056/NEJMc1705630] [Medline: 28767342]
36. Pearson A, Jordan Z, McArthur A, et al. Systematic reviews of textual evidence: narrative, expert opinion or policy. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI; 2024. [doi: 10.46658/JBIMES-24-04]
37. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. Syst Rev 2021 Jan 26;10(1):39. [doi: 10.1186/s13643-020-01542-z] [Medline: 33499930]
38. Pollock D, Peters MDJ, Khalil H, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth 2023 Mar 1;21(3):520-532. [doi: 10.11124/JBIES-22-00123] [Medline: 36081365]
39. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]

Abbreviations

JBI: Joanna Briggs Institute
MAiD: medical assistance in dying

XSL•FO
RenderX

MMSR: mixed methods systematic review
PAS: physician-assisted suicide
PRESS: Peer Review of Electronic Search Strategies
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Reporting Literature Searches in Systematic Reviews
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by A Schwartz; submitted 20.Oct.2025; peer-reviewed by S White; revised version received 13.Mar.2026; accepted 18.Mar.2026; published 21.Apr.2026.

Please cite as:
Marcoux I, Lavoie D, Bergeron A, Beaudin A, Lemyre S, Bourgeois-Guérin V, Bravo G, Hébert M, Simard C, Ummel D, Perron C, Bouthillier ME
Factors Associated With Evolution of the Use of Medical Assistance in Dying: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e85963
URL: https://www.researchprotocols.org/2026/1/e85963
doi:10.2196/85963

© Isabelle Marcoux, David Lavoie, Amy Bergeron, Alexandra Beaudin, Simon Lemyre, Valérie Bourgeois-Guérin, Gina Bravo, Maude Hébert, Chantale Simard, Deborah Ummel, Catherine Perron, Marie-Eve Bouthillier. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Salutogenesis in Mexico and Latin America: Protocol For Scoping Review

Igor Martin Ramos Herrera*, MD, PhD; Rigoberto Antonio Cisneros Garcia*, MGSSS, LE, MPSS

Department of Public Health, University of Guadalajara, Regional Institute for Public Health Research, Sierra Mojada 950, Colonia Independencia, Guadalajara, Mexico

*all authors contributed equally

Corresponding Author:
Rigoberto Antonio Cisneros Garcia, MGSSS, LE, MPSS
Department of Public Health, University of Guadalajara, Regional Institute for Public Health Research, Sierra Mojada 950, Colonia Independencia, Guadalajara, Mexico

Abstract

Background: The salutogenic model, introduced by Aaron Antonovsky, represents a fundamental paradigm shift from the traditional pathogenic orientation—which focuses on the etiology of disease—to a focus on the origins of health (salutogenesis). The core of this model is the Sense of Coherence, a global orientation that allows individuals to mobilize Generalized Resistance Resources to cope with the stressors of daily life. Although salutogenesis has been extensively researched in Europe and North America, its systematic application within diverse and multicultural contexts, marked by the inequity present in Mexico and Latin America, remains fragmented and insufficiently documented. There is a lack of clarity regarding how the model adapts to local realities, specifically in these countries.

Objective: The main objective of this study is to conduct a scoping review to systematically map the extent, range, and nature of research activity on salutogenesis in Mexico and Latin America published between 2010 and 2026. Secondary objectives include identifying the theoretical frameworks employed, describing the psychometric properties of the instruments used to study the Sense of Coherence model in Spanish, English, and Portuguese, as well as determining research gaps to identify future guidelines for health promotion.

Methods: This review will adhere to the methodological framework for scoping reviews developed by Arksey and O'Malley, refined by Levac et al, and the guidelines of the Joanna Briggs Institute. The protocol adheres to the PRISMA criteria for Scoping Reviews (PRISMA-ScR). A trilingual search (ie, English, Spanish, Portuguese) will be conducted in MEDLINE (PubMed), Scopus, Web of Science, PsycINFO, CINAHL, SciELO, LILACS, and Redalyc. Two independent reviewers will screen titles or abstracts and full-text articles, evaluating inter-rater reliability using Cohen Kappa coefficient. Data will be extracted using a piloted form and synthesized through descriptive numerical analysis and qualitative thematic analysis.

Results: This protocol did not involve funding. The search strategy was finalized and preliminary searches were conducted in August 2025. Starting in February 2026, the study will begin the data integration phase. The final analysis is expected to be completed by August 2026, with submission for publication planned for December 2026.

Conclusions: This scoping review will be the first to provide a holistic view of salutogenic research in Latin America. By synthesizing the evidence on how the model is operationalized in this region, the study will provide critical insights for the transition of public health systems from a curative model to one that fosters health assets, supporting the design of culturally congruent interventions.

International Registered Report Identifier (IRRID): PRR1-10.2196/83495

(JMIR Res Protoc 2026;15:e83495) doi:10.2196/83495

KEYWORDS

salutogenesis; Mexico; Latin America; public health; scoping review

Introduction

The Hegemony of the Pathogenic Paradigm

For centuries, biomedical sciences have been dominated by the pathogenic paradigm. This perspective asks, "Why do people get sick?" and focuses its resources on identifying pathogens, risk factors, and treating specific diseases [1]. While this approach has been responsible for monumental achievements in human survival—such as the eradication of smallpox and the management of infectious diseases—it inherently views health as a dichotomous variable (ie, the presence or absence of disease). Under this model, the goal of public health is to return the individual to a state of "normality" or homeostasis, largely

XSL•FO
RenderX

defined by the absence of pathology [2]. However, this "deficit model" fails to explain why despite given the omnipresence of viruses, bacteria, stressors, and socioeconomic challenges, a significant portion of the population remains healthy and thrives.

The Salutogenic Turn

Origins of Health

In the late 1970s, medical sociologist Aaron Antonovsky challenged this orthodoxy by introducing the concept of Salutogenesis (from the Latin *salus*=health, and the Greek *genesis*=origin). Antonovsky posited that chaos and stress are inherent to the human condition, not deviations from the norm. He proposed the "Health Ease/Dis-ease Continuum," suggesting that individuals are not simply "healthy" or "sick," but constantly moving along a spectrum between total health and breakdown [3]. The relevant question then shifts from "how do we treat disease?" to "how can we help a person move toward the health pole on the continuum, regardless of their current physiological state?"

Theoretical Core

The cornerstone of the salutogenic model is the Sense of Coherence (SOC). Antonovsky defined SOC as: "a global orientation that expresses the extent to which one has a pervasive, enduring though dynamic feeling of confidence." This construct consists of three interrelated components:

1. Comprehensibility (Cognitive): The belief that stimuli deriving from internal and external environments are structured, predictable, and explicable.
2. Manageability (Instrumental): The perception that resources are available to meet the demands posed by these stimuli.
3. Meaningfulness (Motivational): The sense that these demands are challenges worthy of emotional investment and engagement.

Research suggests that people with a strong SOC are better able to identify and mobilize Generalized Resistance Resources (GRRs)—such as money, ego strength, social support, and cultural stability—to cope with stressors effectively, thereby maintaining or improving their health.

The Latin American Context

A Unique Epidemiological Landscape

While the salutogenic model has been widely adopted in the Global North, specifically in Scandinavia, the United States, and Israel, its application in Latin America warrants specific investigation. For its part, the Latin American region is characterized by deep structural inequalities, rapid urbanization, and a polarized epidemiological profile where chronic non-communicable diseases (diabetes, hypertension, etc) coexist with infectious diseases and violence [4].

Furthermore, Latin America has a rich history of "Social Medicine" and "Collective Health," which share philosophical roots with salutogenesis but have developed independently. Concepts such as *Buen Vivir* (Sumak Kawsay) in the Andean region emphasize harmony with the community and nature as a prerequisite for health, coinciding with the holistic nature of salutogenesis [5]. However, it is unclear to what extent Europe-derived salutogenic instruments (such as the SOC-29 or SOC-13 questionnaires) are culturally valid in these contexts, or whether they capture the nuances of Latin American resilience.

Rationale for a Scoping Review

Given the potential fragmentation of the literature in this region—scattered across psychology, public health, nursing, and sociology journals, and often published in Spanish or Portuguese—a classic systematic review focusing on intervention effectiveness would be premature. A scoping review is the most appropriate methodology to: (a) map the volume of research activity in the region, (b) clarify concepts and definitions used in the local literature, (c) identify research gaps (eg, lack of qualitative studies, absence of intervention studies), and (d) synthesize the evidence to inform the feasibility of a future full systematic review [6].

To the best of our knowledge, no comprehensive scoping review has specifically addressed the production of salutogenic knowledge in Mexico and Latin America. This protocol outlines the rigorous steps we will take to address this gap.

Methods

Study Design and Registration

This scoping review will be conducted according to the methodological framework originally proposed by Arksey and O'Malley [7] and later enhanced by Levac, Colquhoun, and O'Brien [8]. Additionally, we will follow the guidance provided by the Joanna Briggs Institute (JBI) Manual for Evidence Synthesis [9].

The methodology consists of five key stages: (1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) data charting, and (5) collating, summarizing, and reporting the results.

Stage 1: Identifying the Research Questions

We used the PCC (Population, Concept and Context) mnemonic recommended by the JBI to construct the research questions:

- Population human populations with no restriction on age, gender, or health status (eg, patients, students, health professionals) [10]
- Concept: Salutogenesis, SOC, Generalized Resistance Resources (GRRs)
- Context: Research conducted in Mexico, Central America, South America, and the Spanish/Portuguese-speaking Caribbean

Primary Research Question

What is the current state of scientific knowledge regarding salutogenesis and the SOC in Mexico and Latin America from 2010 to 2026?

Secondary Questions

What theoretical frameworks and definitions of salutogenesis are employed in Latin American research? Which versions of the SOC scale (SOC-13, SOC-29, or adapted versions) are most frequently used and have been validated for specific local

XSL•FO
RenderX

populations? What are the main health outcomes (mental health, chronic disease management, health behaviors) associated with salutogenesis in the region? In what settings (clinical, educational, community, occupational) [11]. Is the model being applied?

Stage 2: Identifying Relevant Studies

A comprehensive search strategy will be implemented to locate both published and gray literature. The search will be limited to documents published between January 1, 2010, and December 31, 2025, to capture the most recent developments and the evolution of the field over the last 15 years.

Information Sources

We will search the following electronic databases through access licenses provided by the University of Guadalajara, Mexico:

International Databases

MEDLINE (via PubMed), Scopus (Elsevier), Web of Science (Core Collection), PsycINFO (APA), and CINAHL (EBSCOhost).

Regional Databases

- SciELO (Scientific Electronic Library Online): covers a vast network of Latin American journals
- LILACS (Latin American and Caribbean Health Sciences Literature): The most comprehensive index of scientific and technical literature in the region
- Redalyc (Network of Scientific Journals of Latin America and the Caribbean, Spain and Portugal): A major open-access database of peer-reviewed journals from Latin America, Spain, and Portugal

Search Strategy

The search strategy will use a combination of controlled vocabulary (MeSH terms for PubMed, DeCS terms for LILACS/SciELO) and free-text keywords. The search will be conducted in English, Spanish, and Portuguese.

- **Keywords (English):** “Salutogenesis,” “Sense of Coherence,”“Sense of Coherence,” “Salutogenic,” “Generalized Resistance Resources,”“Generalized Resistance Resources,” “Latin America,” “Mexico,” “Brazil,” “Argentina,” “Colombia,” “Chile,” “Peru,” “Central America,” “South America.”
- **Keywords (Spanish):** “Salutogénesis,” “Sentido de Coherencia,” “Modelo Salutogénico,” “Recursos Generales de Resistencia,” “América Latina,” “México.”
- **Keywords (Portuguese):** “Salutogênese,” “Senso de Coerência,” “Recursos Gerais de Resistência,” “América Latina.”

A draft of the search strategy for PubMed is provided in Multimedia Appendix 1.

Gray Literature Strategy

To minimize publication bias, we will search for gray literature in:

Google Scholar (reviewing the first 200 results): A broad multidisciplinary search engine that retrieves scholarly literature across formats, useful for capturing a wide range of academic sources

Institutional repositories of major Latin American universities (e.g., UNAM, University of São Paulo, University of Guadalajara): University-based repositories that provide access to theses, dissertations, articles, and other academic outputs produced by institutional researchers and students

ProQuest Dissertations and Theses: A major database of graduate theses and dissertations, offering access to extensive scholarly work from universities worldwide

Stage 3: Study Selection

This process will be managed using Covidence (Veritas Health Innovation) or Rayyan software to ensure transparency and blinding.

Inclusion Criteria

- Study type: Primary empirical studies (quantitative, qualitative, mixed), systematic reviews, meta-analyses, and theoretical articles proposing specific adaptations for the region [12]
- Language: Spanish, English, or Portuguese
- Topic: The study must explicitly address salutogenesis or measure the SOC as a primary variable
- Geography: Conducted in Latin American countries

Exclusion Criteria

- Editorials, commentaries, and opinion pieces without original data or theoretical contribution
- Conference abstracts where the full text is not available
- Studies conducted on Latin American immigrants residing outside the region (eg, in the USA or Spain), as the environmental context differs from the scope of this review

Screening Process

Calibration

Before starting the formal screening, the two reviewers (IMRH and RACG) will independently evaluate a random sample of 50 citations. We will calculate inter-rater reliability using Cohen’s Kappa coefficient. If Kappa<0.75, reviewers will discuss discrepancies and refine the inclusion criteria.

Level 1 Screening (Title and Abstract)

Two reviewers will independently examine all retrieved citations.

Level 2 Screening (Full Text):

The full text of potentially eligible articles will be retrieved and evaluated independently. Reasons for exclusion will be recorded for each article rejected at this stage. Discrepancy Resolution: Any disagreement will be resolved through discussion or by consulting a third senior reviewer if consensus cannot be reached.

Stage 4: Data Charting

A data extraction form will be developed in Microsoft Excel. The form will be iteratively updated; the two reviewers will

XSL•FO
RenderX

pilot test the form on a sample of 5 - 10 studies to ensure it captures all necessary nuances.

The data items to be extracted include:

1. Bibliometric data: Author(s), year, title, journal, country of origin, publication language
2. Methodological data: Study design (eg, cross-sectional, cohort, ethnographic), sample size, sampling method, study duration
3. Population characteristics: Age group (children, adolescents, adults, elderly), gender distribution, socioeconomic status, health status (healthy vs clinical population)

Salutogenic Variables

The methodology framework begins with a clear definition of salutogenesis tailored to the study's context. To quantify the SOC, researchers utilized established instruments such as the SOC-13, SOC-29, SOC-L, or the Family SOC scale. Furthermore, the study details the reliability coefficients (Cronbach α) specifically calculated for the local population to ensure psychometric integrity. Finally, the analysis incorporates additional salutogenic constructs, specifically Generalized Resistance Resources (GRRs).

Main Findings

The main findings included associations between SOC and health outcomes, as well as the main themes identified in qualitative studies. Conclusions included the authors' recommendations for policy or practice.

Stage 5: Collating, Summarizing, and Reporting Results

We will synthesize the data using a hybrid thematic framework synthesis. This approach allows for a systematic link between our research questions and the evidence mapping by combining deductive and inductive coding. The process will follow the framework analysis method proposed by Ritchie and Spencer [13] and the JBI methodological guidance for evidence synthesis [14].

Deductive Coding

An initial framework will be established using categories derived from the study objectives, such as "Theoretical Frameworks" and "SOC Scale Variations." These represent the "expected themes" identified in the pilot search.

Inductive Refinement

Following the reflexive thematic analysis phases of Braun and Clarke [15], reviewers will read the data to identify emerging codes that capture regional nuances, such as the transition from individual to collective resilience or the integration of *Buen Vivir* [16].

Mapping and Synthesis

The final narrative report will integrate numerical descriptions with these themes to provide a holistic view of the salutogenic paradigm in Mexico and Latin America.

Results

Project timeline

The project's timeline began with the conceptualization of the study in January 2025. This was followed by a preparatory phase from August to December 2025, which focused on protocol development and team training. The operational stages for 2026 commenced with database searches and duplicate removal in February and March, leading into the title and abstract screening in April. Looking ahead, the full-text review is projected for May and June, followed by data extraction and synthesis in July. The project is expected to conclude with the reporting phase, involving manuscript writing and submission, scheduled between August and December 2026."

Preliminary Search Results

An initial pilot search in PubMed using the term "Salutogenesis AND Mexico" yielded 109 results, while "Salutogenesis AND Latin America" yielded 1150 results. This confirms the feasibility of the review but highlights the need for a trilingual search strategy to capture the full breadth of the literature. The protocol has been prepared following the PRISMA Extension for Scoping Reviews (PRISMA-ScR) checklist [17] and the preliminary selection process is shown in Figure 1.

XSL•FO
RenderX

Figure 1. PRISMA diagram (PRISMA Extension for Scoping Reviews) [18].

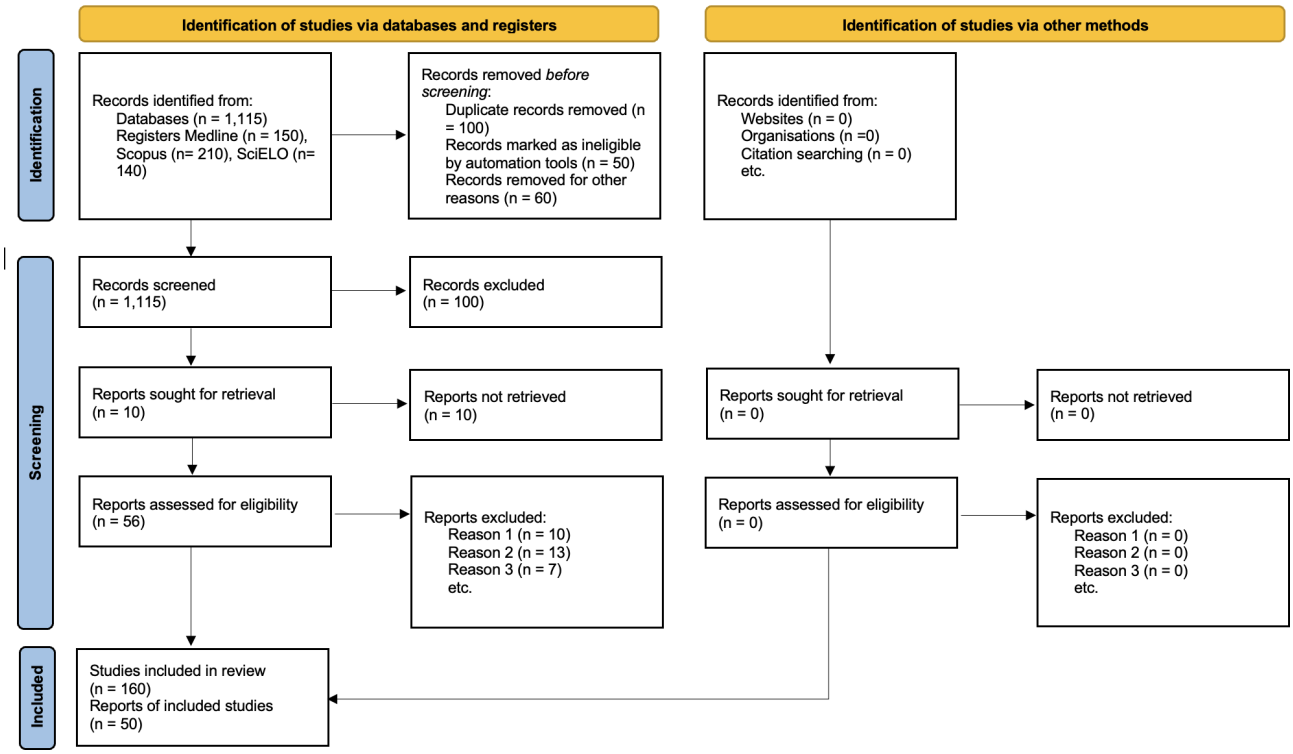

Discussion

Maturation of the Paradigm and the Global-Local Gap

This protocol represents a critical step in the maturation of health promotion research in Latin America [19]. While salutogenesis has been extensively reviewed globally—with a strong predominance of studies in Global North contexts—the translation of this paradigm to Latin American socioeconomic and epidemiological realities remains fragmented. Foundational systematic reviews have mapped the impact of the SOC on specific health outcomes [1]; however, there remains a persistent need for a comprehensive synthesis addressing how the model is operationalized in contexts of high structural inequality. By consolidating literature from regional databases such as LILACS and Redalyc, this study aims to bridge a critical gap, retrieving evidence that is often rendered invisible in Anglophone-centric reviews.

Theoretical Implications: The Tension Between Individualism and Collectivism

We hypothesize that this review will reveal a fundamental epistemological tension between the individualistic genesis of the SOC construct and the inherently collectivistic nature of Latin American cultures. The protocol seeks to clarify whether the use of standard scales (SOC-29, SOC-13) has been uncritical or if there is a transition toward culturally adapted instruments that capture family and community resilience. This analysis is essential for theoretical development, as it allows for testing the universality of Antonovsky's constructs under regional worldviews such as *Buen Vivir* (Sumak Kawsay) [20,21], suggesting that health in the Global South is not merely an individual psychological trait but a socio-ecological process of collective management.

Implications for Practice and Public Policy

From a practical standpoint, the protocol is oriented toward transforming the traditional deficit-based model into an asset-based one. If the review confirms that a robust SOC is associated with better management of chronic conditions (such as diabetes in resource-limited settings), it will strengthen the argument for integrating salutogenic interventions into primary care [22]. Specifically, the findings could support the implementation of "social prescribing," connecting patients with community assets and support networks that bolster their comprehensibility, manageability, and meaningfulness. At a policy level, this study seeks to align with the PAHO Strategy and Plan of Action on Health Promotion 2019 - 2030, providing evidence to shift from a "war on disease" toward governance that generates health through intersectoral interventions [23,24].

Design Strengths and Limitations

The main strength of this protocol lies in its trilingual search strategy and methodological rigor based on JBI and PRISMA-ScR guidelines. By including regional databases, representativeness is ensured, challenging global publication biases. However, it is acknowledged as a limitation that, as a scoping review, a formal assessment of the risk of bias of the included studies will not be conducted; instead, the priority will be an exhaustive mapping of the state of the art in the region.

The results will be disseminated through publication in a high-impact, peer-reviewed indexed journal. Additionally, we plan to create a brief for policymakers and present the findings at the Mexican Public Health Congress to ensure the translation of knowledge into local practice.

XSL•FO
RenderX

Conclusion

The transition toward a health system that truly promotes well-being requires solid theoretical foundations [25]. This protocol establishes the roadmap to assess whether salutogenesis can be that foundation in Latin America [26]. By mapping the existing evidence, it not only identifies academic gaps but also takes a step toward the decolonization of health knowledge, validating how our populations build health amidst adversity [27].

Acknowledgments

Disclosure of Delegation to Generative AI: The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision:

- Translation

The GAI tool used was: GEMINI PRO. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by: Rigoberto Antonio Cisneros Garcia and Igor Martin Ramos Herrera

Additional note: We used GEMINI PRO for the text translation. The translation was also validated by the authors.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Disclosure of Delegation to Generative AI: The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision:
- Translation
The GAI tool used was: GEMINI PRO. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by: Rigoberto Antonio Cisneros Garcia and Igor Martin Ramos Herrera
Additional note: We used GEMINI PRO for the text translation. The translation was also validated by the authors.

Authors' Contributions

IMRH: Conceptualization, Methodology, Supervision, Writing - original draft.
RACG: Investigation, Data curation, Visualization, Writing - review and editing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search Strategy.
[DOCX File, 15 KB - resprot_v15i1e83495_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 275 KB - resprot_v15i1e83495_app2.docx]

References

1. Mittelmark MB, Bauer GF, Vaandrager L, et al. The Handbook of Salutogenesis: Springer Nature Link; 2022. [doi: 10.1007/978-3-030-79515-3]
2. Keyes CLM. The mental health continuum: from languishing to flourishing in life. J Health Soc Behav 2002 Jun;43(2):207-222. [doi: 10.2307/3090197] [Medline: 12096700]
3. Antonovsky A. Health, Stress, and Coping: Jossey-Bass; 1979. URL: https://openlibrary.org/books/OL4433501M/Health_stress_and_coping [accessed 2026-04-03]
4. Pan American Health Organization. Salud En Las Américas+ 2017: Resumen: Panorama Regional y Perfiles de País, 2017th edition: OPS; 2017. URL: https://iris.paho.org/handle/10665.2/34322 [accessed 2026-04-03]
5. Almeida-Filho N. Complexity and multidisciplinarity in collective health: evaluation of concepts and applications [Article in Spanish]. Salud Colect 2006;2(3):241. [doi: 10.18294/sc.2006.61]

XSL•FO
RenderX

6. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
7. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
8. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5(1):69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
9. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Chapter 11: scoping reviews. In: Aromataris E, Muun Z, editors. JBI Manual for Evidence Synthesis: JBI; 2020. URL: https://synthesismanual.jbi.global/ [accessed 2026-04-03]
10. Peruzzo JA, Benjamin CG, Costa MCLD, Silva LDD, Mombelli MA, Lopes AR. Senso de coerência em profissionais de saúde: protocolo de revisão de escopo. Cuad Ed Desar 2024;16(6):e4477. [doi: 10.55905/cuadv16n6-077]
11. Langford R, Bonell CP, Jones HE, et al. The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement. Cochrane Database Syst Rev 2014 Apr 16;2014(4):CD008958. [doi: 10.1002/14651858.CD008958.pub2] [Medline: 24737131]
12. Arroyo-Acevedo HV. The conceptual marriage of salutogenesis and health promotion [Article in Spanish]. Salud Publica Mex 2025;67(5):444-446. [doi: 10.21149/17516]
13. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Analyzing Qualitative Data, Vol. 2010:173. [doi: 10.4324/9780203413081_chapter_9]
14. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Implement 2021 Mar;19(1):3-10. [doi: 10.1097/XEB.0000000000000277] [Medline: 33570328]
15. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
16. Benz C, Bull T, Mittelmark MB, Vaandrager L. Culture in salutogenesis: the scholarship of Aaron Antonovsky. Glob Health Promot 2014 Dec;21(4):16-23. [doi: 10.1177/1757975914528550] [Medline: 24814861]
17. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
19. Lindström B, Eriksson M. The Hitchhiker's Guide to Salutogenesis: Salutogenic Pathways to Health Promotion: Högskolan Vast; 2010. URL: http://hv.diva-portal.org/smash/record.jsf?pid=diva2%3A400361&dswid=-2860 [accessed 2026-04-03]
20. Cubillo-Guevara AP, Hidalgo-Capitán AL, García-Álvarez S. Good living as an alternative to development for Latin America [Article in Spanish]. Iberoamerican Journal of Development Studies 2015;5(2):30-57. [doi: 10.26754/ojs_ried/ijds.184]
21. Mittelmark MB, Sagy S, Eriksson M, et al, editors. The Handbook of Salutogenesis: Cham (CH): Springer; 2017. [doi: 10.1007/978-3-319-04600-6]
22. Morgan A, Ziglio E. Revitalising the evidence base for public health: an assets model. Promot Educ 2007 Jun;14(2_suppl):17-22. [doi: 10.1177/10253823070140020701x]
23. Bauer GF, Pelikan JM, EUPHA (HP), IUHPE Global Working Group of Salutogenesis. 8th Workshop: Advancing Salutogenesis for Health Promotion and Public Health. Eur J Public Health 2017 Nov;27(suppl_3):ckx187.582. [doi: 10.1093/eurpub/ckx187.582]
24. Pan American Health Organization. Strategy and Plan of Action on Health Promotion within the Context of the Sustainable Development Goals 2019-2030: PAHO; 2019. URL: https://iris.paho.org/handle/10665.2/51618 [accessed 2026-04-03]
25. Garzón NE, Heredia LPD. Salutogenesis and cardiovascular health in adults: a scoping review. Texto Contexto - Enferm 2020;29. [doi: 10.1590/1980-265x-tce-2018-0376]
26. Super S, Wagemakers MAE, Picavet HSJ, Verkooijen KT, Koelen MA. Strengthening sense of coherence: opportunities for theory building in health promotion. Health Promot Int 2016 Dec;31(4):869-878. [doi: 10.1093/heapro/dav071] [Medline: 26208654]
27. Rivera de los Santos F, Ramos Valverde P, Moreno Rodríguez C, Hernán García M. Análisis del modelo salutogénico en España: aplicación en salud pública e implicaciones para el modelo de activos en salud. Rev Esp Salud Publica 2011;85(2):129-139. [doi: 10.1590/S1135-57272011000200002]

Abbreviations

GRR: Generalized Resistance Resources
PRISMA-ScR: PRISMA Extension for Scoping Reviews
SOC: Sense of Coherence

XSL•FO
RenderX

Edited by A Schwartz; submitted 08.Sep.2025; peer-reviewed by A Girma; revised version received 13.Mar.2026; accepted 16.Mar.2026; published 16.Apr.2026.

Please cite as:
Ramos Herrera IM, Cisneros Garcia RA
Salutogenesis in Mexico and Latin America: Protocol For Scoping Review
JMIR Res Protoc 2026;15:e83495
URL: https://www.researchprotocols.org/2026/1/e83495
doi:10.2196/83495

© Igor Martin Ramos Herrera, Rigoberto Antonio Cisneros Garcia. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

School-Based Universal Mental Health Prevention Programs for Children Aged 6 to 9 Years in Low- and Middle-Income Countries: Protocol for a Scoping Review

Cherith Langenhoven-Heber, BPsych, MA; Katherine Sorsdahl, PhD; Emily Garman, PhD; Claire van der Westhuizen, MBChB, PhD

Department of Psychiatry and Mental Health, Alan J Flisher Centre for Public Mental Health, University of Cape Town, 46 Sawkins Road, Rondebosch, Cape Town, Western Cape, South Africa

Corresponding Author:
Cherith Langenhoven-Heber, BPsych, MA
Department of Psychiatry and Mental Health, Alan J Flisher Centre for Public Mental Health, University of Cape Town, 46 Sawkins Road, Rondebosch, Cape Town, Western Cape, South Africa

Abstract

Background: While the onset of most mental health conditions occurs in later childhood or adolescence, the early years present an important window for prevention. Worldwide, universal school-based prevention interventions have been shown to be effective in promoting child mental health. Nevertheless, little is known about universal school-based prevention programs targeting children aged 6 to 9 years in low- and middle-income countries (LMICs).

Objective: This scoping review aims to identify promising, potentially feasible, and possibly evidence-based school-based universal prevention mental health programs for children aged 6 to 9 years in LMICs.

Methods: This scoping review is planned according to the Joanna Briggs Institute framework developed by Arksey and O'Malley, with enhancements proposed by Levac et al. Eligible studies will focus on universal prevention school-based mental health interventions with feasibility and/or mental health outcomes that include children aged 6 to 9 years. Additional inclusion criteria are peer-reviewed published and unpublished studies, including gray literature such as theses and dissertations, preprints, conference proceedings and papers, and government reports and publications. There are no limits to the date of publication, but only studies published in English and conducted in LMICs will be included. The search strategy was iteratively developed and finalized, following which major databases, namely, PubMed (MEDLINE), Scopus (Embase), EBSCOhost (PsycInfo, CINAHL, PsycArticles, ERIC, Teacher Reference Center, Academic Search Premier, and SocINDEX), CENTRAL, ProQuest, the Open Science Framework, and the International Clinical Trials Registry Platform, were searched. Titles and abstracts are being screened in addition to manual searches of the reference lists of previous reviews. Experts will be consulted for further study suggestions, and reference mining will be conducted with the included studies. A descriptive and numerical summary will reflect basic program characteristics. A preliminary data extraction chart curated around the research questions will be used for deductive and inductive data analysis. The entire scoping review will be reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines.

Results: In November 2025, the search strategy was finalized. All database searches were completed by early December 2025, yielding 7511 records. After removing duplicates (n=3022), 4489 records underwent title and abstract screening. Of these, 25 records met the criteria for full-text review, which is currently ongoing. Data extraction and analysis should take place between April and May 2026. The scoping review findings will be submitted to a relevant peer-reviewed journal by July 2026 and presented at conferences.

Conclusions: This review will provide a description of feasible and potentially effective school-based universal mental health programs specific to children aged 6 to 9 years living in LMICs. The findings may assist with efficient insight into available programs for implementation or adaptation among the stipulated population in the future.

Trial Registration: Open Science Framework W42CE; https://osf.io/w42ce/overview

International Registered Report Identifier (IRRID): PRR1-10.2196/87105

(JMIR Res Protoc 2026;15:e87105) doi:10.2196/87105

KEYWORDS

mental health; well-being; child mental health; school based; universal prevention; prevention; mental health promotion; children; low- and middle-income countries; LMICs

XSL•FO
RenderX

Introduction

Mental health conditions are highly prevalent in low- and middle-income countries (LMICs), with most conditions emerging early in life. It is estimated that 82% of people with mental health conditions live in LMICs [1]. While most mental health conditions present prior to adulthood [2], worldwide, approximately 34% of people living with mental health conditions have an age of onset younger than 14 years [3]. Furthermore, the global median reported age of onset is 7 to 9 years for attention-deficit/hyperactivity disorder, 7 to 15 years for oppositional defiant disorder, and 9 to 14 years for conduct disorder [4].

In LMICs, although mental health conditions are highly prevalent, up to 90% of those affected receive no care—a disparity known as the "treatment gap" [5,6]. Low political interest, stigma associated with mental health conditions, lack of appropriately trained staff [7], and poor funding for mental health care resulting in insufficient availability and implementation of mental health services [8] have been identified as contributors to the treatment gap in LMICs. This is evident especially for youth in LMICs, who are reported to have a range of 0.05 to 0.12 child and adolescent mental health workers per 100,000 youth [8]. Therefore, the prevalence of mental health conditions, age of onset, and high treatment gap for youth highlighted motivate a focus on mental health prevention interventions for children in LMICs.

The World Health Organization [8], within its Comprehensive Mental Health Action Plan, strongly recommended that interventions should begin in early childhood through universal approaches to prevent poor mental health. Furthermore, worldwide, research evidence has shown that school-based prevention programs are effective in promoting youth mental health, whereas universal school-based programs are especially considered a cost-effective and feasible approach at the population level [7]. More specifically, universal school-based programs have been effective in preventing mental health problems in children as young as 2 years of age [9-12]. Indeed, research indicates that universal social and emotional learning [9,10], resilience [11], and self-regulation [12] school-based programs have significantly improved behaviors, school climate, social skills, emotional skills, and academic abilities and reduced psychological distress, substance use, and school suspensions in youth. Additionally, one meta-analysis conducted by Hayes et al [13] reported that universal prevention school-based programs were effective in preventing depression and anxiety in youth aged 8 to 18 years.

In LMICs, schools offer an opportunity for sustainable mental health promotion and prevention programs among youth [14,15] provided that programs are culturally relevant [16]. One scoping review and one systematic review synthesized evidence primarily on the effectiveness of any universal school-based mental health programs being implemented among children and adolescents in LMICs [16,17]. Bradshaw et al [17] obtained inconclusive results, whereas the scoping review by Harte and Barry [16] supported the efficacy of universal programs among children and adolescents aged 5 to 19 years. Despite a growing evidence base of effective universal school-based programs promoting child mental health, it is still unclear whether such programs are feasible for children aged 6 to 9 years living in LMICs. Indeed, some reviews of the literature have been either too broad in terms of age [13,16,17] or synthesized evidence across mixed prevention approaches (universal, indicated, and selective) [16]. Others have only included randomized controlled trials and/or quasi-randomized controlled trials with specific mental health outcome measures such as anxiety and depression symptoms [13,17,18].

To fill the research gap, we propose to conduct a scoping review to identify and summarize universal school-based mental health programs designed for children aged 6 to 9 years living in LMICs that have demonstrated (1) promising outcomes in feasibility and acceptability or (2) an improvement in well-being and mental health, typically shown through psychosocial measures.

Methods

Overview

The scoping review will be conducted according to the Joanna Briggs Institute (JBI) methodology, developed to guide rigorous knowledge synthesis along with trustworthy and transparent research practice [19]. The JBI scoping review framework, developed by Arksey and O'Malley [20] with enhancements proposed by Levac et al [21], will be used. This framework includes six stages for conducting scoping reviews: (1) identifying the research question, (2) identifying relevant studies, (3) study selection, (4) charting the data, (5) collating findings and summarizing and reporting the results, and (6) stakeholder consultation. As required by the JBI framework, the protocol is reported according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [22] to ensure complete reporting of the scoping review process.

Stage 1: Identifying the Research Question

The scoping review aims to address the following research question: what potentially feasible school-based universal mental health programs are available for children aged 6 to 9 years in LMICs? However, this research question may be further refined if necessary throughout the scoping review process. Secondary research questions are:

1. What components were included in these interventions? Which components were evidence based?
2. Were these interventions developed or adapted for use in their specific context?
3. Were these interventions efficacious, feasible, and acceptable?
4. Which of these interventions reflects a sustainable approach?

Stage 2: Identifying Relevant Studies

Overview

On the basis of the research question, a comprehensive yet feasible search strategy must be developed with clear inclusion and exclusion criteria, including the type of sources searched,

XSL•FO
RenderX

as recommended by the JBI [19,21]. The authors have developed inclusion and exclusion criteria for the scoping review. To retrieve a full range of school-based universal mental health programs for children, all English-language peer-reviewed published and unpublished articles, along with full-text studies from other sources, will be included. These eligibility criteria are delineated and detailed further below.

Eligibility Criteria

There are no limits to the date of publication, but the intervention must have been conducted in LMICs. Due to concerns regarding the dearth of programs available for the stipulated population in LMICs, the authors chose to refrain from applying any limits to the date of publication. All peer-reviewed full-text studies, including nonjournal articles, conference proceedings and papers, and government reports and publications published in English, will be included. The studies must focus on universal prevention school-based mental health interventions and include children aged 6 to 9 years. Further details are provided in Table 1.

Table . Eligibility criteria.

Category	Inclusion criteria	Exclusion criteria
Language	Studies published in English	Studies published in a language other than English
Target population	Studies that include children aged 6 to 9 y at the time of recruitment	Studies that only focus on younger children aged 0 to 5 y or that only include participants aged 10 y and above
Interventions	Studies that investigate universal prevention school-based mental health interventions	Studies that investigate interventions not focused on universal prevention (eg, selective, indicated, or treatment intervention or program studies)
Geographic location	Studies conducted in LMICs[a]	Studies conducted in HICs[b]
Literature sources	Primary studies; secondary studies published in peer-reviewed journals or unpublished; and gray literature, including theses and dissertations and intervention-related literature that appears in conference proceedings and papers and government reports and publications	Gray literature outside of conference proceedings and papers and government reports and publications
Study design	Any study design that includes an intervention	Study designs without an intervention
Outcome measures	Studies with feasibility and/or mental health outcomes, which include the use of psychosocial measures	Studies solely focused on prevalence and/or associations and studies that do not measure any feasibility outcomes or include mental health outcome measures

[a]LMIC: low- and middle-income country.

[b]HIC: high-income country.

Search Strategy

As recommended by the JBI [19], stage 2 also involves developing a comprehensive search strategy. A trial search using 2 databases, namely, EBSCOhost and Scopus, was conducted using preliminary search terms, including MeSH (Medical Subject Headings) terms. This was achieved with the assistance of an institutional librarian from the health science faculty library research department of the University of Cape Town. After analyzing titles, abstracts, and index words of the articles retrieved in the trial search, the key search terms were revised to ensure that the search strategy culminates in articles relevant to this review. Table 2 shows the final search terms.

XSL•FO
RenderX

Table . Search strategy terms.

Keyword	Search terms[a]
Mental health	• “Mental Health” OR “Wellbeing” OR “Emotional Regulation” OR “Resilience” OR “Self-esteem” OR “Coping Skills” OR “Self-regulation” OR “Socio-emotional Characteristics” OR “Socio-emotional Skills” • "Mental Health"[Mesh] OR “Wellbeing” OR "Child Health"[Mesh] OR "Emotional Regulation"[Mesh] OR “Resilience” OR "Resilience, Psychological"[Mesh] OR “Self-esteem” OR "Self-concept"[Mesh] OR "Coping Skills" OR “Self-regulation” OR "Self-control"[Mesh] OR “Socio-emotional Characteristics” OR “Socio-emotional Skills”
Program	• “Program*” OR “Promotion” OR “Life skills” OR “Socio-emotional Learning” OR “Intervention” OR “Psycho-education” • “Program*” OR “promotion” OR “Life Skills” OR “Socio-emotional learning” OR “Intervention” OR "Preventive Health Services"[Mesh] OR "Psychosocial Intervention"[Mesh] OR “Psycho-education”
Children	• “Child*” OR “Primary School*” OR “School Child*” OR “Elementary School*” • "Child"[Mesh] OR “Primary School*” OR “Elementary school*”
School based	• “School-based” OR “School based” OR “School Setting” OR “Curriculum” OR “School” • “School-based” OR “School based” OR “School setting” OR "Curriculum" OR “School” OR “School Mental Health Services"[Mesh] OR "School Health Services"[Mesh]
Universal	• “Universal” OR “Group-based” OR “Group based” OR “Classroom-based” OR “Classroom based” OR “Classroom” • “Universal” OR “Group-based” OR “Group based” OR “Classroom-based” OR “Classroom based” OR “Classroom”

XSL•FO
RenderX

Keyword	Search terms[a]
LMICs[b]	

XSL•FO
RenderX

Keyword	Search terms[a]
	• “Deprived Countries” OR "Deprived Population" OR "Deprived Populations" OR "Developing Countries" OR "Developing Country" OR "Developing Economies" OR "Developing Economy" OR "Developing Nation" OR "Developing Nations" OR "Developing Population" OR "Developing Populations" OR "Developing World" OR "LAMI Countries" OR "LAMI Country" OR "Less Developed Countries" OR "Less Developed Country" OR "Less Developed Economies" OR "Less Developed Nation" OR "Less Developed Nations" OR "Less Developed World" OR "Lesser Developed Countries" OR "Lesser Developed Nations" OR LMIC OR LMICS OR Low GDP OR "Low GNP" OR "Low Gross Domestic" OR "Low Gross National" OR "Low Income Countries" OR "Low Income Country" OR "Low Income Economies" OR "Low Income Economy" OR "Low Income Nations" OR "Low Income Population" OR "Low Income Populations" OR "Lower GDP" OR "lower gross domestic" OR "Lower Income Countries" OR "Lower Income Country" OR "Lower Income Nations" OR "Lower Income Population" OR "Lower Income Populations" OR "Middle Income Countries" OR "Middle Income Country" OR "Middle Income Economies" OR "Middle Income Nation" OR "Middle Income Nations" OR "Middle Income Population" OR "Middle Income Populations" OR "Poor Countries" OR "Poor Country" OR "Poor Economies" OR "Poor Economy" OR "Poor Nation" OR "Poor Nations" OR "Poor Population" OR "Poor Populations" OR "poor world" OR "Poorer Countries" OR "Poorer Economies" OR "Poorer Economy" OR "Poorer Nations" OR "Poorer Population" OR "Poorer Populations" OR "Third World" OR "Transitional Countries" OR "Transitional Country" OR "Transitional Economies" OR "Transitional Economy" OR "Under Developed Countries" OR "Under Developed Country" OR "under developed nations" OR "Under Developed World" OR "Under Served Population" OR "Under Served Populations" OR "Underdeveloped Countries" OR "Underdeveloped Country" OR "underdeveloped economies" OR "underdeveloped nations" OR "underdeveloped population" OR "Underdeveloped World" OR "Underserved Countries" OR "Underserved Nations" OR "Underserved Population" OR "Underserved Populations" OR Afghanistan OR Albania OR Algeria OR “American Samoa” OR Angola OR Armenia OR Azerbaijan OR Bangladesh OR Belarus OR Byelarus OR Belorussia OR Belize OR Benin OR Bhutan OR Bolivia OR Bosnia OR Botswana OR Brazil OR Bulgaria OR Burma OR “Burkina Faso” OR Burundi OR “Cabo Verde” OR “Cape Verde” OR Cambodia OR Cameroon OR “Central African Republic” OR Chad OR China OR Colombia OR Comoros OR Comores OR Comoro OR Congo OR “Costa Rica” OR “Côte d'Ivoire” OR Cuba OR “Democratic People’s Republic of Korea” OR Djibouti OR Dominica OR “Dominican Republic” OR Ecuador OR Egypt OR “El Salvador” OR Eritrea OR Ethiopia OR “Equatorial Guinea” OR Fiji OR Gabon OR Gambia OR Gaza OR “Georgia Republic” OR Georgia OR Ghana OR Grenada OR Grenadines OR Guatemala OR Guinea OR “Guinea Bissau” OR Guyana OR Haiti OR Herzegovina OR Hercegovina OR Honduras OR India OR Indonesia OR Iran OR Iraq OR “Ivory Coast” OR Jamaica OR Jordan OR Kazakhstan OR Kenya OR Kiribati OR Korea OR Kosovo OR Kyrgyz OR Kirghizia OR Kirghiz OR Kyrgyzstan OR “Lao PDR” OR Laos OR Lebanon OR Lesotho OR Liberia OR Libya OR Macedonia OR Madagascar OR Malawi OR Malay OR Malaya OR Malaysia OR Maldives OR Mali OR “Marshall Islands” OR Mauritania OR Mauritius OR Mexico OR Micronesia OR Moldova OR Mongolia OR Montenegro OR Morocco OR Mozambique OR Myanmar OR Namibia OR Nepal OR Nicaragua OR Niger OR Nigeria OR Pakistan OR Palau OR “Papua New Guinea” OR Paraguay OR Peru OR Philippines OR Principe OR Romania OR Rwanda OR Ruanda OR Samoa OR “Sao Tome” OR Senegal OR Serbia OR “Sierra Leone” OR “Solomon Islands” OR Somalia OR “South Africa” OR “South Sudan” OR “Sri Lanka” OR “St Lucia” OR “St Vincent” OR Sudan OR Surinam OR Suriname OR Swaziland OR Syria OR “Syrian Arab Republic” OR Tajikistan OR Tadzhikistan OR Tajikistan OR Tadzhik OR Tanzania OR Thailand OR Timor

XSL•FO
RenderX

Keyword	Search terms[a]
	OR Togo OR Tonga OR Tunisia OR Turkey OR Turkmen OR Turkmenistan OR Tuvalu OR Uganda OR Ukraine OR Uzbek OR Uzbekistan OR Vanuatu OR Venezuela OR Vietnam OR "West Bank" OR Yemen OR Zambia OR Zimbabwe OR "low- and middle-income"

[a]For each keyword, the first bullet point represents the generic search string applied across databases, while the second point includes database-specific MeSH (Medical Subject Headings) terms used exclusively for PubMed searches.

[b]LMIC: low- and middle-income country.

Next, a systematic search incorporating the search terms was performed across all selected databases [19] using various search processes and query strings (Multimedia Appendix 1). PubMed (MEDLINE), Scopus (Embase), EBSCOhost (PsycInfo, CINAHL, PsycArticles, ERIC, Teacher Reference Center, Academic Search Premier, and SocINDEX), and CENTRAL were searched. Additionally, ProQuest, the Open Science Framework, and the International Clinical Trials Registry Platform were searched for gray literature, namely, conference papers, dissertations and theses, and preprints, as mentioned above. The full search was an iterative process as the authors first evaluated whether the full search culminated in sufficient representative articles. The authors decided to exclude the "LMICs" search term during the database search because too few results were found, as was done in a similar review by Bradshaw et al [17]. In addition, colleagues who are experts in the field of school-based mental health programs in LMICs will be consulted for applicable research literature recommendations in case any relevant literature was missed in the search. Moreover, the programs presented in other applicable reviews identified through the database search will be independently assessed by 2 reviewers for inclusion.

Stage 3: Study Selection

This stage involves selecting the relevant articles for the scoping review by screening all the literature identified. Therefore, all the literature identified in stage 2 will be saved on EndNote (version 21; Clarivate Analytics) [23] and loaded onto Rayyan (Qatar Computing Research Institute), a web-based application and mobile app used to facilitate the selection of suitable studies [24]. There are 2 steps to study selection using the study inclusion and exclusion criteria as a guide. First, the study titles and abstracts will be screened independently by 2 reviewers to identify potentially relevant articles for full-text review [24]. Additionally, the review team will meet regularly to discuss any difficulties related to study selection, and this may result in revising the search strategy [21]. At this point, the ascendancy approach, which entails searching the reference list of each source to identify any applicable interventions, will be used with reviews and gray literature from the search results. Relevant studies identified will be added to the title and abstract screening phase for review. Second, the full text of each study will be assessed for eligibility, resulting in the final studies for inclusion in the scoping review [24].

At each step in the selection process, any disagreement on the inclusion or exclusion of literature sources may be resolved through consensus or the involvement of a third party [19]. Once the full-text study selection process is complete, the reviewers will use the ascendancy approach again, with the final included studies, to identify any final study additions to the scoping review. While a scoping review does not require the quality assessment of studies [19], we will apply the JBI critical appraisal tool to each included study. This will highlight the methodological quality of each study, which will be important for the interpretation of the study findings.

Stage 4: Charting the Data

This stage begins with the development of a draft charting table with preliminary domains [21]. In addition to basic study characteristics, data will be extracted according to the study objective and research questions. For example, information will be charted on prior adaptation work and any feasibility testing. A Microsoft Excel spreadsheet will be used to input the extracted data from each study (Multimedia Appendix 2), and the domains will act as headings on the table. The domains currently included are (1) author, (2) publication year, (3) country, (4) setting, (5) target population, (6) sample size, (7) type of intervention, (8) program adaptations for the setting, (9) program components, (10) evidence on the components, (11) method of delivery and facilitator type, (12) duration of program and sessions, (13) outcomes (measures, feasibility, and acceptability), (14) sustainability indicators, and (15) research findings. In addition, a domain for the outcome of the JBI quality appraisal tool for each study will be listed to provide an overview of study rigor. Additional domains may be added after consultation with colleagues in the field of public mental health and during data analysis in the next stage. The table will be piloted independently by 2 reviewers to minimize bias and evaluate whether all the necessary information is extracted. The reviewers will perform an extraction and verification using a 10% sample of the included studies and amend the table if necessary. Once the table is finalized, data from the remaining studies will be extracted and added to the table by the first author.

Stage 5: Collating Findings and Summarizing and Reporting the Results

The results of the study selection process from screening until final selection will be reported [21] according to the PRISMA-ScR guidelines [22], which include a narrative summary of the review selection process and outcomes [19]. Moreover, as required by the JBI framework, it will be accompanied by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram (Figure 1) and tables or charts (Multimedia Appendix 2) used

XSL•FO
RenderX

as visual tools to depict the scoping review process and selected mental health programs [19].

This stage also involves analyzing the data, reporting the results, and discussing the implications of the findings [21]. The first author will compile and present a descriptive and numerical summary of all the study characteristics, including the number of studies, type of study design, year of publication, population, type of intervention, and setting [20]. This stage will also include a thematic summary of the included programs [21] according to program content, outcome measures, method of delivery, duration of the program, type of facilitator, findings, and any other notable theme that may arise from the included studies. The data will be collated and described according to predetermined domains (see Multimedia Appendix 2 for the preliminary charting table). This deductive approach will enable description of the programs according to their components, evidence base, prior adaptations, feasibility, acceptability, and perceived sustainability, thus addressing the research questions. Information regarding other domains identified inductively will also be reported. Although this study is largely descriptive in nature, the collated data on the programs will be assessed for the level of resources and expertise required to deliver them. This initial assessment will inform aspects of the stakeholder consultations described in stage 6.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram (adapted by the first author). MH: mental health.

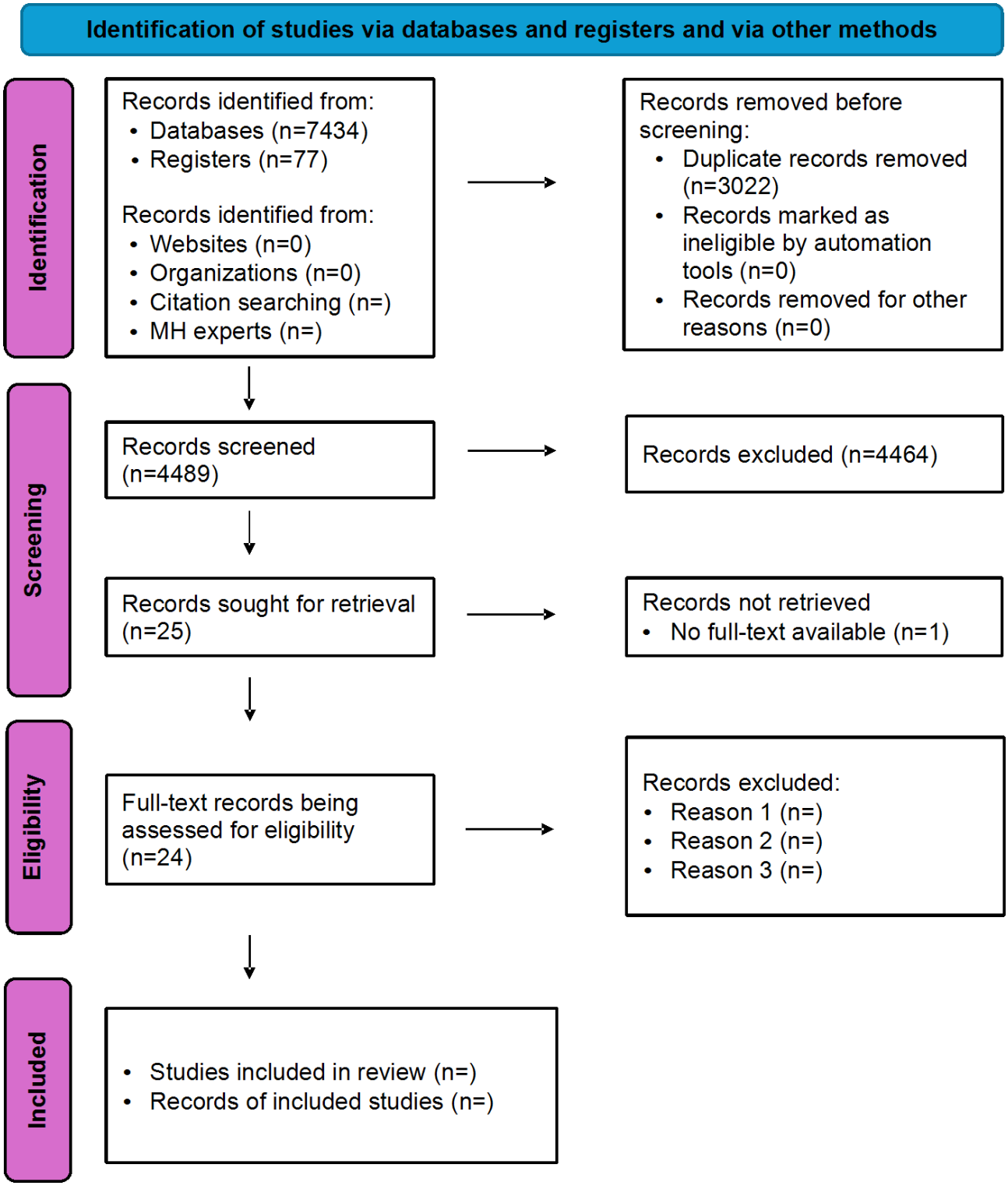

Stage 6: Stakeholder Consultation

According to Levac et al [21], consultation is required throughout the scoping review to establish methodological rigor. Therefore, the entire review process will involve informal consultation with stakeholders such as Centre for Public Mental Health (CPMH) staff, research experts, study supervisors, and a librarian [19]. Thus far, a librarian has assisted with finalizing the search query string and selection of databases; CPMH staff members have provided guidance on databases for gray literature searches and will be involved in regular meetings (1-2 weekly) during the data analysis phase to guide the presentation and interpretation of the data. Experts working in the fields of education or child mental health identified through CPMH staff will be approached to suggest possible programs for inclusion in the review. In addition, they will be contacted to give input on the scoping review findings via email or face-to-face discussions.

XSL•FO
RenderX

Ethical Considerations

As no human participants were involved in this research, no ethics approval is required.

Results

Funding for this study has been ongoing since February 2024, as it is part of a larger study. In November 2025, the search strategy was finalized, and all database searches were completed by December 2, 2025. A pilot search indicated that including the LMIC-related query string greatly limited the search results. Instead, we opted to use this screening criterion at the title and abstract review. In December 2025, we proceeded with the first phase of screening, which includes removing duplicates and the title and abstract review. Currently, the search has yielded 7511 records; 3022 (40.23%) duplicates were removed, and 4489 (59.77%) records were screened at the title and abstract phase, which resulted in 25 (0.33%) studies ready for full-text review. We are in the process of full-text review, and data extraction and analysis will occur between April 2026 and May 2026. We anticipate the completion of the scoping review manuscript for submission by July 2026. Once published, these findings will be presented at a conference to contribute to academic literature and inform future research.

Discussion

From the scoping review, we expect to identify programs that show some level of effectiveness, sustainability, or feasibility and acceptability suitable for children aged 6 to 9 years in LMICs. We know that the value of universal school-based programs is a highly contentious subject in academia currently, especially in LMICs [25,26]. Therefore, it is imperative that we find a program or components within these programs that show promise and may inform future work. Moreover, a detailed review protocol founded on evidence-based guidelines is essential for a well-executed scoping review, resulting in transparent and rigorous research practice [21]. Conducting a formal quality appraisal of each study contributes methodological depth that extends past conventional scoping review practices. This is important as there are various programs that have a distinct focus, such as social and emotional learning, resilience, and self-regulation, and it will be interesting to see the differences in study quality. Moreover, we might find that there is an overlap of components across the different types of programs and that perhaps a combination of these components could yield a more comprehensive program.

It was highlighted by Seekles et al [27] that many reviews have been conducted but lack focus on this specific age group, highlighting a critical gap in research reviews. This is evident by the large availability of adolescent mental health research, yet the age of 6 to 9 years is a highly important developmental stage for children. Two more recent reviews by Hayes et al [13] and Harte and Barry [16] have since included this age range but, unfortunately, focused on the effectiveness of programs, which provides less information on the feasibility and acceptability or sustainability of each program. This review will address this gap.

Apart from the future research contribution of this scoping review, there are additional strengths to its methodology. The application of the JBI framework ensures rigorous, trustworthy, and transparent research practice [21]. Additionally, adhering to the PRISMA-ScR guidelines [22] will ensure comprehensive reporting within the scoping review [28]. In contrast, limiting the geographical location to LMICs may result in the exclusion of potentially feasible school-based universal mental health programs conducted elsewhere. Furthermore, being limited to studies published in English may exclude studies conducted in numerous LMICs where other languages are spoken. The authors are also aware that the lack of limitation on the publication dates could result in the inclusion of outdated programs. However, while conducting the literature review, it was clear that children aged 6 to 9 years in LMICs are often ignored in research, which could already limit the number of programs identified. Additionally, we will be cautious when reporting the findings of outdated interventions. Prior to conducting the scoping review, an interim protocol has been uploaded and registered on the Open Science Framework platform [29], which will be updated iteratively. Furthermore, the findings of the scoping review will be submitted to a relevant academic journal and presented at one or more conferences to contribute to academic literature and inform future research.

Acknowledgments

The authors are grateful to Jamie-Lee Hendricks, a librarian at the University of Cape Town based at the Department of Libraries: Research and Innovation in the Faculty of Health Sciences, who assisted the first author with finalizing the search query string that enabled efficient database searches.

Funding

This research was funded in whole or in part by the African Mental Health Research Initiative 2.0 (Del-22-004). This initiative is being implemented with funding support from the Developing Excellence in Leadership, Training, and Science in Africa program (DEL-22-012), a program of the Science for Africa Foundation being implemented with support from the Wellcome Trust and the UK Foreign, Commonwealth, and Development Office, and is part of the European and Developing Countries Clinical Trials Partnership 2 program supported by the European Union. For purposes of open access, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

XSL•FO
RenderX

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

CL-H, CvdW, and KS led the design and conceptualization of the work, along with formulating the research question. CvdW is the guarantor. Furthermore, all authors (CL-H, CvdW, KS, and EG) were involved in the review design; finalization of the inclusion and exclusion criteria and the search strategy; and, finally, the drafting, review, and editing of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Database search strategies.
[DOCX File, 33 KB - resprot_v15i1e87105_app1.docx]

Multimedia Appendix 2
Data extraction tool.
[DOCX File, 18 KB - resprot_v15i1e87105_app2.docx]

References

1. GBD Compare. Institute for Health Metrics and Evaluation (IHME). 2024. URL: https://vizhub.healthdata.org/gbd-compare/ [accessed 2024-03-15]
2. Lund C, Brooke-Sumner C, Baingana F, et al. Social determinants of mental disorders and the Sustainable Development Goals: a systematic review of reviews. Lancet Psychiatry 2018 Apr;5(4):357-369. [doi: 10.1016/S2215-0366(18)30060-9] [Medline: 29580610]
3. Solmi M, Radua J, Olivola M, et al. Age at onset of mental disorders worldwide: large-scale meta-analysis of 192 epidemiological studies. Mol Psychiatry 2022 Jan;27(1):281-295. [doi: 10.1038/s41380-021-01161-7] [Medline: 34079068]
4. Kessler RC, Angermeyer M, Anthony JC, et al. Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative. World Psychiatry 2007 Oct;6(3):168-176. [Medline: 18188442]
5. Saade S, Parent-Lamarche A, Khalaf T, Makke S, Legg A. What barriers could impede access to mental health services for children and adolescents in Africa? A scoping review. BMC Health Serv Res 2023 Apr 6;23(1):348. [doi: 10.1186/s12913-023-09294-x] [Medline: 37024835]
6. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866. [Medline: 15640922]
7. WHO. World mental health report: transforming mental health for all. : World Health Organisation; 2022 URL: https://www.who.int/publications/i/item/9789240049338 [accessed 2026-03-31]
8. WHO. Mental health atlas 2024. : Worl Health Organization; 2025 URL: https://www.who.int/publications/i/item/9789240114487 [accessed 2026-03-31]
9. Cipriano C, Strambler MJ, Naples LH, et al. The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. Child Dev 2023;94(5):1181-1204. [doi: 10.1111/cdev.13968] [Medline: 37448158]
10. Durlak JA, Weissberg RP, Dymnicki AB, Taylor RD, Schellinger KB. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. Child Dev 2011;82(1):405-432. [doi: 10.1111/j.1467-8624.2010.01564.x] [Medline: 21291449]
11. Dray J, Bowman J, Campbell E, et al. Systematic review of universal resilience-focused interventions targeting child and adolescent mental health in the school setting. J Am Acad Child Adolesc Psychiatry 2017 Oct;56(10):813-824. [doi: 10.1016/j.jaac.2017.07.780] [Medline: 28942803]
12. Pandey A, Hale D, Das S, Goddings AL, Blakemore SJ, Viner RM. Effectiveness of universal self-regulation-based interventions in children and adolescents: a systematic review and meta-analysis. JAMA Pediatr 2018 Jun 1;172(6):566-575. [doi: 10.1001/jamapediatrics.2018.0232] [Medline: 29710097]
13. Hayes D, Deniz E, Nisbet K, et al. Universal, school-based, interventions to improve emotional outcomes in children and young people: a systematic review and meta-analysis. Front Child Adolesc Psychiatry 2025;4:1526840. [doi: 10.3389/frcha.2025.1526840] [Medline: 40529852]
14. Skeen S, Gemmell K, du Toit S, et al. The role of educational institutions in promoting and protecting mental health across childhood, adolescence and youth. In: Tomlinson M, Kleintje S, Lake L, editors. South African Child Gauge 2021/2022: Children's Institute, University of Cape Town; 2022.

XSL•FO
RenderX

15. Fazel M, Patel V, Thomas S, Tol W. Mental health interventions in schools in low-income and middle-income countries. Lancet Psychiatry 2014 Oct;1(5):388-398. [doi: 10.1016/S2215-0366(14)70357-8] [Medline: 26361001]
16. Harte P, Barry MM. A scoping review of the implementation and cultural adaptation of school-based mental health promotion and prevention interventions in low-and middle-income countries. Glob Ment Health (Camb) 2024;11:e55. [doi: 10.1017/gmh.2024.48] [Medline: 38751723]
17. Bradshaw M, Gericke H, Coetzee BJ, Stallard P, Human S, Loades M. Universal school-based mental health programmes in low- and middle-income countries: a systematic review and narrative synthesis. Prev Med 2021 Feb;143:106317. [doi: 10.1016/j.ypmed.2020.106317] [Medline: 33159922]
18. Johnstone KM, Kemps E, Chen J. A meta-analysis of universal school-based prevention programs for anxiety and depression in children. Clin Child Fam Psychol Rev 2018 Dec;21(4):466-481. [doi: 10.1007/s10567-018-0266-5] [Medline: 30105480]
19. Peters MDJ, Godfrey C, McInerney P, et al. Scoping review. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: JBI; 2020.
20. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
22. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
23. EndNote. Clarivate. 2013. URL: https://endnote.com/ [accessed 2026-03-31]
24. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
25. Birrell L, Grummitt L, Smout S, Maulik P, Teesson M, Newton N. Debate: where to next for universal school-based mental health interventions? Child Adolesc Ment Health 2025 Feb;30(1):92-95. [doi: 10.1111/camh.12749] [Medline: 39789719]
26. Coetzee B, Loades M. Debate: Where to next for universal school-based mental health interventions? Addressing the mental health crisis among young people in low- and middle-income countries. Child Adolesc Ment Health 2025 Feb;30(1):105-107. [doi: 10.1111/camh.12752] [Medline: 39644172]
27. Seekles M, Twagira F, Alam A, Obasi A. Promotive and preventive interventions for adolescent mental health in Sub-Saharan Africa: a combined scoping and systematic review. BMJ Public Health 2023 Nov;1(1):e000037. [doi: 10.1136/bmjph-2023-000037] [Medline: 40017854]
28. McGowan J, Straus S, Moher D, et al. Reporting scoping reviews-PRISMA ScR extension. J Clin Epidemiol 2020 Jul;123:177-179. [doi: 10.1016/j.jclinepi.2020.03.016] [Medline: 32229248]
29. Langenhoven-Heber C, van der Westhuizen C, Sorsdahl K, Garman E, Rochat C. Identifying promising, potentially feasible school-based universal mental health programmes for children aged 6 to 9 years in low-income and middle-income countries: a scoping review. OSF. 2026 Jan 16. URL: https://osf.io/yz54t/overview [accessed 2026-03-13]

Abbreviations

CPMH: Centre for Public Mental Health
JBI: Joanna Briggs Institute
LMIC: low- and middle-income country
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by A Schwartz; submitted 14.Nov.2025; peer-reviewed by R Chigevenga; accepted 23.Feb.2026; published 10.Apr.2026.

Please cite as:
Langenhoven-Heber C, Sorsdahl K, Garman E, van der Westhuizen C
School-Based Universal Mental Health Prevention Programs for Children Aged 6 to 9 Years in Low- and Middle-Income Countries: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e87105
URL: https://www.researchprotocols.org/2026/1/e87105
doi:10.2196/87105

© Cherith Langenhoven-Heber, Katherine Sorsdahl, Emily Garman, Claire van der Westhuizen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Apr.2026. This is an open-access article distributed under the terms

XSL•FO
RenderX

of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Methodological Research Practices for the Adaptation of Measuring Instruments in Health Care: Protocol for a Scoping Review

Ninive Pita Gomes de Oliveira, MN; Rosilene Alves Ferreira, MN; Andressa Aline Bernardo Bueno, MN; Flávia Giron Camerini, PhD; Danielle de Mendonça Henrique, PhD; Cristiane Helena Gallasch, PhD; Cintia Silva Fassarella, PhD

Faculdade de Enfermagem, Universidade do Estado do Rio de Janeiro, Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro, Brazil

Corresponding Author:
Ninive Pita Gomes de Oliveira, MN
Faculdade de Enfermagem, Universidade do Estado do Rio de Janeiro, Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro, Brazil

Abstract

Background: The identification of the methodological path followed by researchers in psychometric studies can contribute to maintaining the transparency and rigor necessary for the study.

Objective: This study aims to map the methodological practices of adaptation of health measurement instruments.

Methods: The protocol is based on the method proposed by the Joanna Briggs Institute using the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews). The following databases were consulted: MEDLINE, CINAHL, Web of Science, LILACS, the Cochrane Library, and Embase. The search for data in gray literature sources was carried out in the following databases: WorldCat and the Brazilian Portal of Scientific Publications and Open Access Data (O Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto [OASISBR]). The search strategy was based on the participants, concept, and context framework: participants (researchers in methodological practices), concept (studies of adaptation of measuring instruments), and context (health area). The research will consider studies on methodological practices adopted to conduct adaptation studies of health measurement instruments. The selection of studies, data extraction, and synthesis will be carried out by 2 independent reviewers.

Results: Data collection was carried out in the months of June and October 2025. The results will be presented in a narrative format and with tables or diagrams. The results are being prepared for publication and should be published by March 2026.

Conclusions: This study will identify the current status and challenges faced by researchers in carrying out adaptation studies of health measurement instruments. The findings will contribute to the development of future research that promotes better psychometric evidence and advances science in nursing.

International Registered Report Identifier (IRRID): DERR1-10.2196/86949

(JMIR Res Protoc 2026;15:e86949) doi:10.2196/86949

KEYWORDS

psychometrics; transcultural adaptation studies; health; research instruments; nursing

Introduction

Psychometrics is based on the theory of measurement in sciences to explain the sense that have the answers given by participants to a series of tasks, typically called items, and thus proposes techniques for measuring mental processes. In summary, psychometrics seeks to understand the meaning behind the answers given by the respondents, and from this dynamic, it is possible to develop tools to evaluate psychological aspects [1].

From this, it is possible to understand that studies aimed at the adaptation of measurement instruments are essentially psychometric studies, and these studies concern themselves with the problems related to the challenges of measuring human behavior, through a specific set of mathematical methods and models, in addition to providing essential aspects for conducting research in social and behavioral sciences [2-4].

The adaptation of a test, that is, an instrument, is a complete scientific process, and as a result, it is guided by the principles of the scientific method—the need to offer evidence of the adequacy of this linguistic transformation being the most prominent of all—not only in terms of language but also in terms of other psychometric characteristics [5].

Several categories of professionals have made efforts to apply the measurement instruments in their practices, allowing the

XSL•FO
RenderX

evaluation and measurement of different behaviors and variables. Such behavior has also been seen in health and nursing professionals, who have been developing research aimed at creating or adapting instruments that can improve their professional practices [6-9].

However, it is essential that the conduct of these studies is anchored in a contemporary repertoire of recommendations, which can contribute to resolving the weaknesses inherent in the study and capable of reverberating and impacting positively the reality and designated purpose [10,11]. In this context, it is important to mention that the international scientific community recognizes the standards of the International Test Commission [10] as general guidelines for the process of transcultural adaptation, simultaneously between countries [10]. In 2017, a refinement of the original 1996 proposal was published, in which the guidelines were presented that establish general recommendations and guidelines for conducting studies of this nature and with this purpose [10].

In addition, the periodic updates proposed by the Standards for National and Psychological Test (2014 Edition) [12] are published and disseminated according to the need to conduct psychometric studies aimed at adapting the instruments based on scientific evidence. These updates aim to ensure the required transparency and methodological rigor, as well as to ensure that the tests meet scientifically established quality parameters [12,13].

It should be noted that there is no consensus on how to adapt or create an instrument for use in another cultural context. There are no recipes, manuals, or step-by-step procedures on how to do it. This procedure will depend on the characteristics of the instrument and the context of its application and the population to which it is intended. However, there is a consensus that the process of adaptation goes beyond mere translation. Moreover, validity cannot be guaranteed based on isolated evidence alone, as the validation process requires the accumulation of different relevant evidence to provide a solid scientific basis for interpretations of proposed scores [14,15].

In this sense, the identification of the methodological guidelines that guide these studies is a fundamental strategy for improving the evaluation of quality and safety in the process of adaptation of instruments, allowing the verification of methodological consistency, as well as the alignment between the tools produced for health practice and the theoretical and normative recommendations established by the experts of the method.

Therefore, conducting research through psychometric studies is justified, as it supports scientific advancement in the field, enables the measurement of sensitive and unique variables, and provides interpreted results and analyses that can inform decision-making. In addition, it is estimated that the identification of the practices conducted in these studies may, in the future, be able to guide researchers in the field to improve their practices and expand the contemporary knowledge recommended for studies with this scope.

A preliminary search was conducted on April 4, 2025, to identify other available reviews with a similar objective as that of this study. The survey included the Cochrane database of systematic reviews, the Joanna Briggs Institute's Synthesis of Evidence, and the Open Science Framework repository, and no data were found that would prevent the referral of this research.

The research question of this protocol is as follows: What are the methodological practices of adaptation of measurement instruments in health? Thus, the objective is to map the methodological practices of adaptation of measurement instruments in health care.

Methods

Overview

The proposed protocol will be carried out according to the Joanna Briggs Institute (JBI) methodology for scoping reviews and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [16,17]. The protocol followed the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) and was registered on the Open Science Framework[18].

Eligibility Criteria

The eligibility criteria were developed according to the participants, concept, and context (PCC) framework, and these will be discussed in the following sections [16].

Participants

The participants in this review correspond to the researchers responsible for conducting adaptation studies of measuring instruments. Therefore, scientific productions that describe the methodological decisions, procedures, and steps adopted by researchers in the adaptation process are considered including, where applicable, translation, back-translation, synthesis, evaluation by a committee of experts, and pretest and psychometric analysis.

The adoption of the participants criterion is justified by the fact that, in scoping reviews, as recommended by the JBI [16], participants may correspond to different analytical units, not restricting themselves to individuals but including authors, studies, or processes. In this perspective, the focus on the researchers makes it possible to identify, describe, and analyze systematically the methodological conduct, the technical decisions, and the steps adopted in the processes of adaptation of measuring instruments [17,19].

Concept

For the concept portion of the framework, as a definition, adaptation studies of measurement instruments will be considered, with emphasis on methodological processes used for the transcultural or linguistic adaptation of instruments originally developed in other contexts. The review seeks to map, describe, and systematize the methodological steps adopted in these studies, as well as the theoretical and normative recommendations that guide such processes.

The justification for this criterion lies in the diversity of studies carried out in the field, as some of these studies use obsolete and outdated techniques and procedures that may not produce reliable and adequate evidence to support the adaptation process

XSL•FO
RenderX

of instruments and may introduce benchmarking bias in clinical practice in health [20-22].

Context

In relation to the context of this protocol, the health area will be understood, covering different disciplinary fields, levels of care, and educational and research scenarios. Studies that involve the adaptation of measurement instruments intended for the evaluation of constructions related to human health, clinical practice, public health, health management, and scientific research will be considered, regardless of the target population or level of complexity of care, since inserted in the scope of health sciences.

The justification of the criterion lies in the possibility of intervention for more accurate diagnoses, effective treatments, and patient safety with the use of measuring instruments intended for the health care context. In addition, instruments intended for and applied in the health field allow the quantification of several sensitive variables, and they can assist in the assessment of health status and clinical decision-making [20].

Search Strategy

On the basis of the terms used in the acronym PCC, a mapping of keywords was carried out to identify their descriptors in the controlled health area vocabularies, including Descriptors in Health Sciences, Medical Subject Headings (MeSH), NorthHolland Publishing Company (ELSEVIER), and CINAHL. The terms contained in the titles and abstracts of the relevant articles obtained and the descriptors used will be used to develop a complete search strategy. A librarian with expertise in the health care field developed a structured search strategy with keywords, descriptors, and their synonyms, adapted to each database. There will be a 3-step search strategy.

First, a limited initial search was carried out on June 26, 2025, in the MEDLINE via PubMed and in the CINAHL via EBSCo databases to identify the existence of material that would support this research (Table 1).

Table . Search strategies and preliminary results. Rio de Janeiro, RJ, Brazil, 2025[a].

Database	Search strategy	Results, n (%)
MEDLINE	(“Psychometrics/instrumentation”[Majr] OR “Psychometrics/methods”[Majr] OR “Psychometrics/standards”[Majr]) AND (Validation Studies as Topic[Mesh] OR Adaptation[Title/Abstract] OR Validation[Title/Abstract]) AND (“Health Workforce”[Mesh] OR “Health Personnel”[Mesh] OR “Nursing”[Mesh])	302 (49.3)
CINAHL	(Psychometrics OR “Psychometrics/ST” OR “Psychometrics/MT”) AND MM (“Instrument Adaptation” OR “Instrument Validation” OR “Validation Studies”) AND TX(“Nursing Care” OR “Nursing Diagnosis” OR “Nursing context” OR “health context”)	294 (50.7)

[a]Source: adapted by the authors, 2025.

A second survey using all identified keywords and synonyms was conducted on October 24, 2025, in the following databases: MEDLINE via PubMed, CINAHL, Web of Science, LILACS via BVS, Embase, and the Cochrane Library (Table 2).

XSL•FO
RenderX

Table . Search strategies and preliminary results. Rio de Janeiro, RJ, Brazil, 2025[a].

Database	Search strategy	Results, n (%)
MEDLINE	(“Psychometrics/instrumentation”[Majr] OR “Psychometrics/methods”[Majr] OR “Psychometrics/standards”[Majr]) AND (Validation Studies as Topic[Mesh] OR Adaptation[Title/Abstract] OR Validation[Title/Abstract]) AND (“Health Workforce”[Mesh] OR “Health Personnel”[Mesh] OR “Nursing”[Mesh])	326 (19.3)
CINAHL	(Psychometrics OR “Psychometrics/ST” OR “Psychometrics/MT”) AND MM (“Instrument Adaptation” OR “Instrument Validation” OR “Validation Studies”) AND TX(“Nursing Care” OR “Nursing Diagnosis” OR “Nursing context” OR “health context”)	316 (18.7)
Web of Science	(“Psychometrics” OR “psychometric assessment” OR “psychometric properties”) AND (“Validation Studies” OR “cross-cultural adaptation” OR “Construction and evaluation”) AND (“Health Workforce” OR “Health Personnel” OR “Nursing”)	296 (17.5)
LILACS	((“Psicometria/instrumentação” OR “Psicometria/métodos” OR “Psicometria/normas”) OR (“Propriedades psicométricas” OR “validação psicométrica” OR “Avaliação psicométrica” OR “psychometric properties” OR “psychometric validation” OR “Psicometric evaluation” OR “psychometric assessment” OR “Evaluación psicométrica”)) AND (mh:(“Estudos de Validação como Assunto”) OR “Validation Studies as Topic” OR “Adaptação cultural” OR “Adaptação transcultural” OR “cross-cultural adaptation” OR “Construção e avaliação” OR “Construction and evaluation” OR “Elaboración y evaluación”) AND (mh:(“Mão de Obra em Saúde” OR “Pessoal de Saúde” OR “Enfermagem”) OR “Health Workforce” OR “Fuerza Laboral en Salud” OR “Main-d’oeuvre en santé” OR “Health Personnel” OR “Personal de Salud” OR “Personnel de santé” OR Nursing OR Enfermería OR Soins)	543 (32.0)
Embase	(“psychometry”/exp/mj AND “devices”/de OR (“psychometry”/exp/mj AND “methodology”/de) OR (“psychometry”/exp/mj AND “standard”/de)) AND (“validation study”/exp OR “adaptation”:ti,ab,kw OR “validation”:ti,ab,kw) AND (“health workforce”/exp OR “health care personnel”/exp OR “nursing”/exp)	175 (10.3)
Cochrane Library	(“Psychometrics/instrumentation”[Majr] OR “Psychometrics/methods”[Majr] OR “Psychometrics/standards”[Majr]) AND (Validation Studies as Topic[Mesh] OR Adaptation[Title/Abstract] OR Validation[Title/Abstract]) AND (“Health Workforce”[Mesh] OR “Health Personnel”[Mesh] OR “Nursing”[Mesh])	37 (2.2)

[a]Source: adapted by the authors, 2025.

The research in the gray literature, available in Table 3, was carried out in the same period with the aim of broadening the identification of evidence and reducing publication bias. The WorldCat databases and the Brazilian Portal of Scientific Publications and Data in Open Access (O Portal Brasileiro de Publicações e Dados Científicos em Acesso Aberto [OASISBR]) were consulted, which collect theses, dissertations, technical reports, and other academic documents. The same descriptors and search strategies applied to bibliographic databases were used, with adaptations to the specificities of each platform.

XSL•FO
RenderX

Table . Strategies to search gray literature. Rio de Janeiro, RJ, Brazil, 2025[a].

Database	Search strategy	Results, n (%)
WorldCat	"psychometrics properties" AND (adaptation OR construction) AND health	84 (17.5)
OASISBR	("psychometrics properties" OR "propriedades psicométricas" AND (adaptation OR adaptação) AND Health NOT "revisão sistemática" NOT nutrição NOT desporto NOT (occupational OR ocupação)	397 (82.5)

[a]Source: adapted by the authors, 2025.

In the third step, the reference lists of articles included in the review will be evaluated for possible selection of additional articles. It should be noted that the review will consider studies in any language and without a time restriction, given that the recommendations for conducting research of this scope are part of a historical trajectory rooted in American Psychological Association guidelines [12], which have been disseminated and implemented in practice since 1950, more specifically since 1952.

Types of Sources

Regarding the methodological approach, the review will include cross-cultural, linguistic, or contextual adaptations. It will include studies published in peer-reviewed journals, dissertations, theses, and technical reports detailing methodologies of adaptation and psychometric analysis because these documents provide detailed information on the methodological steps, technical decisions, and procedures adopted by the researchers.

To minimize the risk of double counting of evidence already summarized in reviews, each identified study will be individually checked for its original publication. Whenever a primary study is present in an earlier review, only the original publication will be included in the review. This ensures that each instrument adaptation is considered only once, preserving the integrity and accuracy of the mapping of methodological practices [10,12].

Study Selection

After the search, all identified studies will be grouped and exported to Rayyan [23], and duplicates will be removed. The titles and abstracts will be analyzed by 2 independent reviewers, after training, to assess the studies' eligibility in relation to the inclusion criteria initially defined. A screening pilot test will be conducted independently by both reviewers on an initial total of 25 titles and abstracts. The reviewers will discuss discrepancies and make changes to eligibility criteria and definitions if necessary. This pilot test will continue until at least 75% agreement is reached among the reviewers [16,24].

Complete studies that meet or potentially meet the inclusion criteria will be reviewed. The full text of the selected references will be evaluated according to the inclusion criteria by 2 independent reviewers. Any disagreements between the reviewers at each stage of the selection process will be resolved by consensus or by decision of a third reviewer. Eligible study references will be imported into the Mendeley reference manager. Any reasons for exclusion of studies will be recorded and reported in the scoping review. The research results will be reported in full in the final scoping review and presented in a flowchart.

Data Extraction

Data will be extracted from studies included in the scoping review by 2 reviewers using a data extraction instrument adapted from the model proposed by the JBI [16] (Textbox 1). The data extracted will include specific details about the studies consulted, including information such as title, authors, year of publication, place of study, type of study, research object, research issue, study objectives, study population, sample size, study results, study procedures, study procedures, data analysis, and gaps and limitations of the study.

XSL•FO
RenderX

Textbox 1. Data extraction instrument. Rio de Janeiro, RJ, Brazil, 2025.

Scoping review details

Title

- Methodological research practices for adaptation of measuring instruments in the health field: protocol for a scoping review

Objective

- Map the methodological practices of adaptation of measurement instruments in health.

Inclusion criteria

- Participants: studies that are conducted by researchers on methodological practices, that is, steps conducted by researchers for adaptation studies of measurement instruments
- Concept: adaptation studies of measurement instruments available in the literature
- Context: health studies on the adaptation of measurement instruments to be used and applied in clinical practice

Research issue

- What are the methodological practices of adaptation of measurement instruments in health?

Details extraction from selected studies

- Title
- Authors
- Year of publication
- Place of study
- Type of study
- Object of research
- Research question of the study
- Objective of the study
- Population
- Size of sample
- Results of the authors
- Techniques adopted by the authors
- Procedures adopted by the authors
- Analysis of data
- Gaps or limitations
- Other variables related to the methodological practices adopted by researchers in psychometric studies of the adaptation of measurement instruments (if any)

Source: adapted from the Manual of the Joanna Briggs Institute, by authors.

The draft of the data extraction tool may be modified and revised if necessary, during the data extraction process for each included article, especially if variables related to the methodological practices adopted by researchers in psychometric studies of adaptation of measurement instruments are observed.

Data Analysis and Presentation

The analysis for this scoping review will mainly adopt a content analysis, focusing on data extracted from the literature. As recommended by the JBI, the analysis will be predominantly descriptive and exploratory, without formal critical assessment of the quality of studies. This approach allows the results of the review to contribute directly to the understanding of methodological practices adopted in the adaptation of instruments, offering a comprehensive overview that supports future research and the implementation of good practices in health care [16].

The extracted data will be presented in tables or diagrams format, aligned with the research objective and issue of the scoping review. A descriptive summary will accompany the tabulated results and/or diagrams, providing context on how these data relate to the purpose of the review.

Ethical Considerations

This review does not require approval from an ethics committee because it does not involve human participants.

XSL•FO
RenderX

Dissemination

The results of the study will be disseminated through peer-reviewed publications, conference presentations, and targeted knowledge sharing sessions with relevant stakeholders.

It is believed that this synthesis of knowledge can offer valuable contributions to the academy and practice of nursing, as well as identifying gaps, enabling new paths for applied research, and pointing out transformations in the practice of studies with this scope and that impact on the application of tools that will be used in nursing practice.

Results

Data collection was carried out in the months of June and October 2025. The results will be presented in a descriptive way, through a narrative synthesis, complemented by tables and/or diagrams, according to the nature of the extracted data. At the moment, the results are being organized and analyzed, with submission for publication by March 2026.

Discussion

Anticipated Findings

This scoping review aims to analyze and synthesize the existing literature on methodological practices adopted by researchers in carrying out studies on the adaptation of measuring instruments in the health care field.

To our knowledge, there is a lack of reviews on how the process of adapting measurement instruments occurs, which has resulted in numerous instruments being introduced and used in clinical practice, often with weak, incomplete, or insufficiently satisfactory methodological steps that do not meet the recommended standards for such studies. The use of instruments in clinical practice that do not present robust psychometric evidence may impact decision-making when it is based on the results of the instrument used.

Furthermore, the strength of this review lies in its ability to identify current gaps and possible weaknesses in studies on this topic, and above all, to provide a foundation for the development of future investigations that can deepen knowledge in the area and support the design and execution of measurement instrument adaptation studies based on contemporary methodological recommendations [9,11], directly contributing to the presentation of adequate and reliable psychometric evidence for each study design and phenomenon investigated.

Limitations of the Study

Although the scoping review conducted according to the guidelines of the JBI is appropriate for the broad and systematic mapping of the literature, this methodological design has inherent limitations. The inclusion of studies with different methodological designs, contexts, and analytical approaches may have introduced heterogeneity in the data, limiting the depth of synthesis and comparison between results.

The search strategy, although structured and reproducible according to the JBI guidelines, was restricted to selected databases, which may have resulted in the exclusion of relevant productions not indexed or published in other languages. Finally, considering the exploratory nature of scoping reviews, the results obtained reflect the panorama of the literature available until the search period, not including studies published later, especially in areas of rapid scientific evolution.

Conclusions

We hope that this study can contribute to building a discussion focused on the practice of researchers and the dissemination of results to multiple stakeholders, including nurses, health professionals, health managers, and certainly the populations for whom the measurement instrument is intended.

To maintain the relevance and significance of this context for clinical practice, it is believed that investigations with this type of scope should ensure methodological rigor supported by contemporary recommendations, in addition to ensuring that their decisions are documented so that it is possible to assess the credibility of the process that was followed, as well as assisting in the planning and execution of research of this type. Finally, this study will serve as a basis to equip researchers in the future to make decisions about the methodological framework to be used during their cross-cultural adaptation studies.

Acknowledgments

The authors wish to thank the Faculty of Nursing and the University of the State of Rio de Janeiro for the support and the Research Group Technologies in Health and Nursing in the context of patient safety in the hospital environment (TESPAH) for all contributions and support throughout the construction.

Funding

This work was carried out with the support of Foundation Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ; case: SEI 260003/006194/2024). This study was also supported by CAPES (Coordination of the Improvement of Higher-Level Personnel), Brazil, Financing Code 001.

Data Availability

All data generated or analyzed during this study are included in this published article.

XSL•FO
RenderX

Authors' Contributions

All authors contributed to the conception and planning of the work and to writing and critical review of intellectual content. NPGdO, DdMH, and AABB planned the methodology, including the participants, concept, and context framework criteria and search strategies. NPGdO and CSF were responsible for data acquisition, analysis, and interpretation. CHG and FGC oversaw the project, providing academic guidance and approving the final version of the manuscript. All authors agreed to be responsible for all aspects of the work.

Conflicts of Interest

None declared.

References

1. Pasquali L. Psicometria: Teoria Dos Testes Na Psicologia e Na Educação [Book in Portuguese]: Vozes; 2017.
2. Bado FMR, Rebustini F, Jamieson L, Cortellazzi KL, Mialhe FL. Evaluation of the psychometric properties of the Brazilian version of the Oral Health Literacy Assessment in Spanish and development of a shortened form of the instrument. PLoS ONE 2018;13(11):e0207989. [doi: 10.1371/journal.pone.0207989] [Medline: 30496226]
3. Bandalos DL. Measurement theory and applications for the social sciences. In: American Psychological Association: The Guilford Press; 2018.
4. Dalla Nora CR, Zoboli E, Vieira MM. Validation by experts: importance in translation and adaptation of instruments. Rev Gaúcha Enferm 2017 Sep 18;38(3):e64851. [doi: 10.1590/1983-1447.2017.03.64851]
5. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1986) 2000 Dec;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014]
6. Graeff MDS, Almeida MDA, Porcel-Gálvez AM, Nomura ATG, Lunelli RP, Silva TD. Cross-cultural adaptation and validation of instrument to measure nursing care dependency. Rev Gaucha Enferm 2022;43:e20210135. [doi: 10.1590/1983-1447.2022.20210135.en] [Medline: 36350961]
7. Sampaio H, Araújo NMS, Alves PC, Cabral LA. Adaptation and validity evidence of the Questionnaire for Assessing Educational Podcasts (QAEP) into Brazilian Portuguese: an undisciplined study in health literacy [Article in Portuguese]. Texto livre 2024;17. [doi: 10.1590/1983-3652.2024.47783]
8. Ferreira R, Sucheck N, Sá J, Neto I. Cross-cultural adaptation of the Undergraduate Clinical Education Environment Measure Instrument [Article in Portuguese]. Rev Esp Edu Med 2024;3:603501. [doi: 10.6018/edumed.603501]
9. Oliveira F, Kuznier TP, Souza CC, Chianca TCM. Theoretical and methodological aspects for the cultural adaptation and validation of instruments in nursing [Article in Portuguese]. Texto Contexto Enferm 2018 Jul 23;27(2):e4900016. [doi: 10.1590/0104-070720180004900016]
10. The ITC Guidelines for Translating and Adapting Tests, 2nd edition: International Test Commission; 2017.
11. He J, Vijver FVD. Choosing an adequate design and analysis in cross-cultural personality research. Current Issues in Personality Psychology 2017;5(1):3-10. [doi: 10.5114/cipp.2017.65824]
12. American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME). Standards for Educational and Psychological Testing: American Educational Research Association; 2014.
13. Chan EKH. Standards and guidelines for validation practices: development and evaluation of measurement instruments. In: Validity and Validation in Social, Behavioral, and Health Sciences: Springer; 2014:9-23. [doi: 10.1007/978-3-319-07794-9_2]
14. Pasquali L. Validade dos testes psicológicos: será possível reencontrar o caminho? Psic: Teor e Pesq 2007;23(spe):99-107. [doi: 10.1590/S0102-37722007000500019]
15. Rebustini F. Analysis of the reasons for the regular practice of physical activity in the return of injuries of amateur volleyball athletes. : Revista Terapia Manual; 2018.
16. Khalil H, McInerney P, Munn Z, Tricco AC. Scoping reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI; 2024. [doi: 10.46658/JBIMES-24-09]
17. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
18. Methodological practices for adapting measurement instruments in the healthcare area: scope review protocol. Open Science Framework. URL: https://osf.io/u8hq6/overview [accessed 2025-05-25]
19. Rodrigues RS, Neubert PS. Introdução à Pesquisa Bibliographical [Book in Portuguese]: Editora da UFSC; 2023.
20. Santos Júnior CJD, Costa PDS. Cross-cultural adaptation and validation for Brazilian Portuguese of the Parent Attitudes about the Childhood Vaccine (PACV) questionnaire. Cien Saude Colet 2022 May;27(5):2057-2070. [doi: 10.1590/1413-81232022275.11802021] [Medline: 35544831]
21. Takeda GA, Moraes JMM, Carvalho PD, Alvarenga MDS. Cross-cultural adaptation and evaluation of the psychometric properties of the Brazilian version of the Positive Eating Scale. Cien Saude Colet 2025 Apr;30(4):e13322023. [doi: 10.1590/1413-81232025304.13322023] [Medline: 40298720]

XSL•FO
RenderX

22. Menezes MS, Faro A. Adaptação e Evidências de Validade do Deliberate Self-Harm Inventory - Versão Simpli. Psico-USF 2023;28(1):41-52. [doi: 10.1590/1413-82712023280104]
23. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
24. Moher D, Shamseer L, Clarke M, et al. Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]

Abbreviations

JBI: Joanna Briggs Institute
MeSH: Medical Subject Headings
PCC: participants, concept, and context
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by J Sarvestan; submitted 01.Nov.2025; peer-reviewed by RAD Oliveira; revised version received 03.Feb.2026; accepted 05.Feb.2026; published 27.Mar.2026.

Please cite as:
Gomes de Oliveira NP, Ferreira RA, Bernardo Bueno AA, Camerini FG, Henrique DDM, Gallasch CH, Fassarella CS
Methodological Research Practices for the Adaptation of Measuring Instruments in Health Care: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e86949
URL: https://www.researchprotocols.org/2026/1/e86949
doi:10.2196/86949

© Ninive Pita Gomes de Oliveira, Rosilene Alves Ferreira, Andressa Aline Bernardo Bueno, Flávia Giron Camerini, Danielle de Mendonça Henrique, Cristiane Helena Gallasch, Cintia Silva Fassarella. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 27.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Methods of Analysis in Randomized Noninferiority Trials: Methodological Survey Review Protocol

Daniel Xie[1]; Sameer Parpia[2,3], PhD; Tyler McKechnie[2,4], MD; Phillip Staibano[2,5], MD

[1]Department of Integrated Biomedical Engineering and Health Sciences, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[2]Department of of Health Research Methods, Evidence, and Impact, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[3]Department of Oncology, Faculty of Health Sciences, McMaster University, 711 Concession St G-Wing Room 125, Hamilton, ON, Canada
[4]Department of Surgery, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada
[5]Department of Surgery, Division of Otolaryngology–Head and Neck Surgery, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

Corresponding Author:
Sameer Parpia, PhD
Department of of Health Research Methods, Evidence, and Impact, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada

Abstract

Background: Noninferiority (NI) trial designs that investigate whether an experimental intervention is no worse than the standard of care have been used increasingly in recent years. The robustness of the conclusions depends in part on the analysis population set used. In NI settings, the intention-to-treat (ITT) and per-protocol (PP) analysis sets are most common. The ITT analysis has been considered anticonservative compared with the PP analysis.

Objective: This study aimed to conduct a methodological review assessing the analysis population sets used in contemporary NI trials.

Methods: A comprehensive electronic search strategy will be conducted to identify studies indexed in MEDLINE, Embase, Emcare, and Cochrane CENTRAL. Studies will be included if they are NI trials published in 2024. The primary outcome is the analysis population used for the primary analysis of the trial (ITT, PP, or as-treated). Secondary outcomes include the NI margin, effect estimates, point estimates, and corresponding CIs. Analyses will be performed using descriptive statistics.

Results: The comprehensive search initially identified 1209 studies, of which 403 trials were eligible for data extraction. Data extraction began in January 2025 and is expected to be completed in January 2026.

Conclusions: This methodological survey of NI trials will describe the analysis population used in the primary analysis and assess factors that may be associated with each analysis method.

Trial Registration: PROSPERO CRD420251021089; https://www.crd.york.ac.uk/PROSPERO/view/CRD420251021089

International Registered Report Identifier (IRRID): DERR1-10.2196/76876

(JMIR Res Protoc 2026;15:e76876) doi:10.2196/76876

KEYWORDS

noninferiority; analysis population; analysis; randomized controlled trial; intention-to-treat; per-protocol

Introduction

In biomedical research, randomized controlled trials (RCTs) are generally considered the gold standard for comparing a new treatment with the standard of care (SOC) [1]. Superiority RCTs determine whether a new treatment is more efficacious than the SOC [1]. However, a new treatment may provide similar therapeutic effects while providing additional benefits, such as improved convenience, lower cost, or fewer adverse effects [2]. In such cases, noninferiority (NI) trial designs are used to demonstrate that a new treatment is no worse than the SOC [2] by a prespecified NI margin, which represents the maximum acceptable loss of efficacy of the new treatment relative to the SOC in exchange for the benefits of the new treatment [3].

While NI trials have their benefits, they also present unique challenges, as they require specific trial design, analysis, and interpretation. Researchers continue to debate the optimal application of intention-to-treat (ITT), per-protocol (PP), or as-treated (AT) populations when analyzing NI trials [4]. ITT analysis is preferred in superiority trials because of its conservative nature. Inclusion of patients who are nonadherent to the study protocol tends to attenuate treatment effects toward the null. However, in NI trials, this same characteristic may increase the risk of researchers falsely concluding there is no difference between the treatment and control groups (type I error in the NI setting), because the objective is to demonstrate that the investigational treatment is not worse than the SOC [5].

XSL•FO
RenderX

PP analysis is designed to assess the effect of an intervention among patients who adhere to the study protocol. However, excluding participants after randomization can introduce either positive or negative bias, which may lead to incorrect conclusions [6,7]. Despite this limitation, PP analysis is often used in NI trials, as it is generally considered to be less affected by protocol deviations [8].

Recent reviews of NI trials [9-12] do not adequately capture contemporary trials published after the CONSORT-Outcomes 2022 extension of the original CONSORT 2012 statement, which recommends that authors more clearly define the analysis population used for each outcome and explain how protocol nonadherence is accounted for in the trial protocol [13]. In practice, there is often a lag between the publication of new research guidelines and their consistent application; implementation can take over a year as researchers become aware of and adopt new practices. By analyzing trials published in 2024, we aim to capture trials that have implemented the 2022 CONSORT-Outcomes extension.

Current recommendations suggest using both the ITT and PP analyses, under the assumption that the ITT analysis is anticonservative in the NI setting [13]. The largest previous review focused on trials published in five leading medical journals (*The New England Journal of Medicine*, *The Lancet*, *Journal of the American Medical Association*, *Annals of Internal Medicine*, and *The BMJ*) [7], thereby limiting the generalizability of its findings. Lastly, it remains unclear whether the number of participants excluded from the PP population affects the choice of the primary analysis set.

Therefore, we aim to conduct a methodological review investigating the primary analysis population used in NI trials published in 2024 and whether this choice is associated with the proportion of patients with protocol deviations. Understanding current practices in the contemporary literature will help guide the future design of NI trials.

Methods

Study Design

This study is a methodological survey of contemporary NI trials published across all medical and surgical disciplines. The methodology for this review is reported in accordance with PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) (Checklist 1). The study protocol was registered with the International Prospective Register for Systematic Reviews (PROSPERO; CRD420251021089).

Primary Research Question

What types of analysis population sets (eg, ITT/mITT (modified ITT), PP, ITT and PP, AT, other) are used for the primary analysis in NI trials published in peer-reviewed medical journals between January 1, 2024, and December 31, 2024?

Secondary Research Objective

This study aimed to evaluate the relationship between the primary analysis population set for NI trials and protocol adherence.

Eligibility Criteria

The eligibility criteria are outlined in Textbox 1.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Peer-reviewed noninferiority randomized controlled trials defined as a study self-described as a noninferiority trial in their title, abstract, introduction, or methods of publication.
- Published from January 1, 2024, to December 31, 2024, in any medical or surgical discipline.
- Published in English.

Exclusion criteria

- Studies described as pilot or feasibility randomized trials.
- Secondary analyses of randomized trial data or any other type of study not reporting primary data.
- Conference presentations or abstracts.
- Cluster noninferiority trials.

Since this work is conducted by researchers in Canada and is unfunded, we do not have the resources to invest in software to accurately translate articles published in multiple languages. Therefore, feasibility and resource constraints prevent us from expanding our search to include publications in other languages. We focused on individual randomized trials because cluster randomized trials may have low adherence at both the unit and individual levels, which adds complexity to the definition of ITT and PP analyses in cluster randomized trials.

Information Sources

The following databases were searched from January 1, 2024, to December 31, 2024: MEDLINE, Embase, Emcare, and Cochrane CENTRAL. References of studies meeting the inclusion criteria were manually searched to ensure that all relevant articles were included.

Search Strategy

The search strategy was designed with input from the study investigators and a medical research librarian. The Medical Subject Heading term “non-inferiority” was used for each

XSL•FO
RenderX

database. Search terms related to NI trials such as "noninferiority randomized trial," "noninferiority design," and "noninferiority clinical trial" were also used. A full list of search strategies is available in Multimedia Appendix 1.

Study Selection

Two reviewers will independently evaluate the searched titles and abstracts using Covidence systematic review software (Veritas Health Innovation Ltd). Each reviewer will be adequately trained and familiarized with the inclusion and exclusion criteria to ensure accurate and consistent screening. Discrepancies at the title and abstract and full-text screening stages will be resolved by consensus. If disagreement persists, an additional reviewer will be consulted. Interrater reliability will be measured with the Cohen κ statistic. This statistic is more informative than the proportionate agreement statistic because it adjusts for agreement occurring by chance [14].

Outcome Definitions

The primary outcome for this study will be the analysis population set used in the primary analysis of the NI trials.

Data Management

Two reviewers will independently conduct data extraction into a data collection form designed a priori. Discrepancies will be reviewed in detail by a third reviewer, who will resolve any conflicts. A piloted Microsoft Excel file of the extraction form can be found in Multimedia Appendix 2. The extracted data will include the following:

- Trial characteristics, including author, year of publication, journal of publication, specialty, study period, intervention details, and control details
- Number of patients randomized
- Protocol adherence, including number of patients who crossed over, number of patients who were nonadherent to the assigned intervention, and number of patients lost to follow-up
- Analysis population used in the primary analysis
- Outcome data in each analysis population, including type of primary outcome (continuous, binary, or time of event), direction of outcome (positive or negative), reported point estimate and 95% CI, reported effect estimate and 95% CI, number of primary outcome events (if binary), mean (SD) of the primary outcome (if continuous), and method of statistical analysis for the primary outcome
- Trial conclusion, including inferior, NI, superior, or inconclusive

Statistical Analysis

The analysis population method used in the primary analysis in the trials will be summarized descriptively (ie, the proportion of studies using each method) and corresponding 95% CIs will be estimated using the Wilson score method [15]. Logistic regression will be used to investigate the relationship between the primary analysis population and protocol adherence by regressing the outcome against the proportion of patients in the ITT population who are excluded in the PP analysis, modeled using restricted cubic splines [16]. All statistical analyses will be performed using Stata (version 18; StataCorp LLC).

Risk of Bias Assessment

Risk of bias assessment will not be performed for this methodological review. Risk of bias assessment is not recommended for methodological reviews, because the goal is to understand the methods applied in the studies being assessed [17].

Dissemination

The findings of the survey will be submitted for publication in a peer-reviewed journal aimed at better informing clinical investigators on the current state of NI trial design. The results will be presented at conferences pertaining to clinical trial methodology. Study findings will also be communicated to relevant stakeholders (eg, investigators, biostatisticians).

Ethical Considerations

Ethics approval was not required because this is a methodological survey based entirely on previously published clinical trials.

Results

A total of 2102 studies were captured in the original search, and 893 duplicates were removed. Thus, 1209 studies were available for title and abstract screening. As of December 2025, this review is in the data extraction phase, with 403 studies eligible for extraction. Data extraction started in January 2025, and we expect results to be published in January 2026. Populated data extraction sheets for each included trial will be attached as supplemental files to the final manuscript to promote open science practices.

Discussion

We anticipate that most trials will use the ITT and PP analysis population sets to adhere to the updated CONSORT-Outcomes 2022 extension for NI trials. Based on our initial data extraction, we hypothesize that the ITT will be the most prevalent analysis set despite its tendency to attenuate the treatment effect toward the null, thereby increasing the risk of a type I error. For this reason, we further hypothesize that trial nonadherence will be correlated with the use of the PP analysis set, as the PP set may better reveal true inferiorities that could be masked in an ITT analysis.

Uncertainty regarding which analysis population set should be used for NI trials is ubiquitous in the contemporary literature. Despite the methodological concerns associated with using the ITT population in the NI setting, several recent reviews have found that ITT remains the most prevalent analysis population set. A 2021 review of antibiotic NI trials published before 2019 found that 72.2% (164/227) of studies used both ITT and PP analyses, 18.2% (41/227) of studies used ITT analysis, and 9.7% (22/227) used PP analysis [9], suggesting that when a single analysis population is designated as primary, ITT is the most common. A 2020 review of oncological NI trials published from 2012 to 2018 found that 50% (55/110) of trials reported using only one analysis population, and 87.3% (48/55) of those used the ITT population [10], further reinforcing that ITT is most commonly used when a single primary analysis population is

XSL•FO
RenderX

specified. Most recently, a 2024 review evaluating the methodological quality of NI trials published from 2014 to 2019 also found that most trials used both ITT and PP analyses [11]. It is important to note that all of these reviews examined trials published before 2022. A limitation of these reviews is that they do not report how often each analysis set was designated as the primary, an aspect our study will offer.

In 2022, the CONSORT-Outcomes 2022 extension of the CONSORT 2012 statement was published, recommending that authors further explain the definition of the analysis population used for each outcome and how it relates to nonadherence in the trial protocol [13]. A 2022 paper also suggested that the primary analysis in the NI setting should be the PP or AT populations, given the weaknesses of the ITT approach in NI trials [18]. NI trials published since then may demonstrate more robust reporting of the rationale for the analysis population sets applied, and there may be a shift away from the predominant use of the ITT approach observed in previous reviews.

In addition to assessing the prevalence of analysis population sets in contemporary NI trials, examining whether protocol adherence is associated with certain analysis populations will help researchers understand whether the decision to use the ITT or PP population is influenced by the proportion of patients who do not adhere to the protocol. No previous review has assessed whether an association exists between the primary analysis population set and protocol adherence among trial participants. This information will enable trialists to make more informed decisions regarding analysis population selection in NI trials, contributing to more rigorous and valid synthesis of health evidence. Once this review is completed, the findings will be used to inform a second study comparing the results of NI trials by their designated analysis population set to evaluate how outcomes differ.

A limitation of our study is sampling bias, as we are analyzing only NI trials published in 2024. We focused our review on trials published in 2024 to capture contemporary literature that may have implemented the 2022 CONSORT-Outcomes extension for NI trials. As a result, we will not be able to report longitudinal trends. A second limitation is that restricting our search to NI trials published in English reduces generalizability. Non-English trials may differ systematically in the research methods applied. Moreover, the most impactful papers are typically published in English-language journals with high impact factors [19]. This may bias the dataset because trials published in high-impact journals are more likely to be methodologically robust and to declare NI. For this reason, we designed a broad search strategy to ensure that we captured all NI trials published in peer-reviewed journals.

Despite these limitations, this study will provide valuable insights into the analytic approaches used in contemporary NI trials and the factors that may influence their selection.

Acknowledgments

The authors would like to thank the medical research librarian for their aid and professional guidance in the development of the search strategy.

Funding

No sources of financial support were used in the development of this protocol. The authors have not declared a specific grant for this research from any funding agency in the public, commercial, or not-for-profit sectors.

Disclaimer

No generative artificial intelligence was used in the process of writing and revising this manuscript.

Data Availability

Data sharing is not applicable to this article as no datasets were generated or analyzed during this study.

Authors' Contributions

SP and DX contributed to study design, drafting, and revisions of the manuscript. TM and PS contributed to the study design.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 17 KB - resprot_v15i1e76876_app1.docx]

Multimedia Appendix 2
Extraction template.
[XLSX File, 26 KB - resprot_v15i1e76876_app2.xlsx]

XSL•FO
RenderX

Checklist 1
PRISMA-P checklist.
[PDF File, 121 KB - resprot_v15i1e76876_app3.pdf]

References

1. Braga LH, Farrokhyar F, Dönmez M, et al. Randomized controlled trials - the what, when, how and why. J Pediatr Urol 2025 Apr;21(2):397-404. [doi: 10.1016/j.jpurol.2024.11.021] [Medline: 39701869]
2. Kishore K, Mahajan R. Understanding superiority, noninferiority, and equivalence for clinical trials. Indian Dermatol Online J 2020;11(6):890-894. [doi: 10.4103/idoj.IDOJ_130_20] [Medline: 33344335]
3. Schumi J, Wittes JT. Through the looking glass: understanding non-inferiority. Trials 2011 May 3;12:1-2. [doi: 10.1186/1745-6215-12-106] [Medline: 21539749]
4. Leung JT, Barnes SL, Lo ST, Leung DY. Non-inferiority trials in cardiology: what clinicians need to know. Heart 2020 Jan;106(2):99-104. [doi: 10.1136/heartjnl-2019-315772] [Medline: 31672779]
5. Gupta SK. Non-inferiority clinical trials: practical issues and current regulatory perspective. Indian J Pharmacol 2011 Jul;43(4):371-374. [doi: 10.4103/0253-7613.83103] [Medline: 21844987]
6. Santos-Gallego CG, Requena-Ibanez JA, Badimon J. Per-protocol versus intention-to-treat in clinical trials: the example of GLOBAL-LEADERS trial. J Am Heart Assoc 2022 May 17;11(10):e025561. [doi: 10.1161/JAHA.122.025561] [Medline: 35574954]
7. Mo Y, Lim C, Watson JA, White NJ, Cooper BS. Non-adherence in non-inferiority trials: pitfalls and recommendations. BMJ 2020 Jul 1;370:m2215. [doi: 10.1136/bmj.m2215] [Medline: 32611541]
8. Rehal S, Morris TP, Fielding K, Carpenter JR, Phillips PPJ. Non-inferiority trials: are they inferior? A systematic review of reporting in major medical journals. BMJ Open 2016 Oct 7;6(10):e012594. [doi: 10.1136/bmjopen-2016-012594] [Medline: 27855102]
9. Bai AD, Komorowski AS, Lo CKL, et al. Intention-to-treat analysis may be more conservative than per protocol analysis in antibiotic non-inferiority trials: a systematic review. BMC Med Res Methodol 2021 Apr 19;21(1):75. [doi: 10.1186/s12874-021-01260-7] [Medline: 33874894]
10. Wayant C, Ross A, Vassar M. Methodological quality of oncology noninferiority clinical trials. Crit Rev Oncol Hematol 2020 May;149:102938. [doi: 10.1016/j.critrevonc.2020.102938] [Medline: 32172223]
11. Totton N, Julious S, Walters S, Coates E. A review of UK publicly funded non-inferiority trials: is the design more inferior than it should be? Trials 2024 Dec 4;25(1):809. [doi: 10.1186/s13063-024-08651-3] [Medline: 39627838]
12. Sengul A, Escobar E, Flores JR, et al. Non-Inferiority trials: a systematic review on methodological quality and reporting standards. J Gen Intern Med 2024 Oct;39(13):2522-2530. [doi: 10.1007/s11606-024-08890-9] [Medline: 38954320]
13. Butcher NJ, Monsour A, Mew EJ, et al. Guidelines for reporting outcomes in trial reports: the CONSORT-outcomes 2022 extension. JAMA 2022 Dec 13;328(22):2252-2264. [doi: 10.1001/jama.2022.21022] [Medline: 36511921]
14. Li M, Gao Q, Yu T. Kappa statistic considerations in evaluating inter-rater reliability between two raters: which, when and context matters. BMC Cancer 2023 Aug 25;23(1):799. [doi: 10.1186/s12885-023-11325-z]
15. Dunnigan K. Confidence interval calculation for binomial proportions. 2008 Presented at: InMWSUG Conference.
16. Korn EL, Graubard BI. Appendix C: restricted cubic regression splines. In: Analysis of Health Surveys: Wiley; 1999:345-346.
17. Mbuagbaw L, Lawson DO, Puljak L, Allison DB, Thabane L. A tutorial on methodological studies: the what, when, how and why. BMC Med Res Methodol 2020 Sep 7;20(1):226. [doi: 10.1186/s12874-020-01107-7] [Medline: 32894052]
18. Ranganathan P, Pramesh CS, Aggarwal R. Non-inferiority trials. Perspect Clin Res 2022;13(1):54-57. [doi: 10.4103/picr.picr_245_21] [Medline: 35198430]
19. Strauss P, Manalo E. "Most good papers are published in English": Japanese academics' perspectives on the benefits and drawbacks of writing and publishing in English. J Engl Res Publ Purp 2023 Jul 20;4(1):5-28. [doi: 10.1075/jerpp.22009.str]

Abbreviations

AT: as-treated
ITT: intention-to-treat
NI: noninferiority
PP: per-protocol
RCT: randomized controlled trial
SOC: standard of care

XSL•FO
RenderX

Edited by J Sarvestan; submitted 06.May.2025; peer-reviewed by A Thabane, C Obianyo, H Anandan; revised version received 06.Dec.2025; accepted 08.Dec.2025; published 13.Mar.2026.

Please cite as:
Xie D, Parpia S, McKechnie T, Staibano P
Methods of Analysis in Randomized Noninferiority Trials: Methodological Survey Review Protocol
JMIR Res Protoc 2026;15:e76876
URL: https://www.researchprotocols.org/2026/1/e76876
doi:10.2196/76876

© Daniel Xie, Sameer Parpia, Tyler McKechnie, Phillip Staibano. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Strengthening Undergraduate Medical Education for Inclusive Health Care for People With Down Syndrome and Intellectual and Developmental Disabilities in Medical Schools: Protocol for a Scoping Review

Isabela Amaral de Almeida Bistafa[1], MSc; Marco Antonio Ribeiro Filho[1], MD; Ana Caroline dos Santos Costa[1], MSc; Ana Flávia Pimenta Saad[1]; Rauer Ferreira Franco[1], MSc; Janaína Aparecida de Sales Floriano[1], MSc; Emerson Roberto dos Santos[1], MSc; João Daniel de Souza Menezes[1], MSc; Matheus Querino da Silva[1], MSc; Daniele Nunes Longhi Aleixo[1], MSc; Eliana Fazuoli Chubaci[1], MSc; Andressa Karina Stefani[1], MSc; William Donegá Martinez[1], MSc; Pedro Belchior da Silveira Junior[1], MSc; Ana Mirtes Zecchin[1]; Nathalia Bavaresco Gonçalves Cristóvão[1], MD; Alex Bertolazzo Quitério[1], MSc; Renato Mendonça Ribeiro[2], Prof Dr; Denise Cristina Móz Vaz Oliani[3], Prof Dr; Antônio Hélio Oliani[3], Prof Dr; Maysa Alahmar Bianchin[1], Prof Dr; Luís Cesar Fava Spessoto[4], Prof Dr; Rita de Cássia Helú Mendonça Ribeiro[4], Prof Dr; Júlio César André[1], Prof Dr

[1]Center for Studies and Development of Health Education – CEDES, Avenida Brigadeiro Faria Lima, São Jose do Rio Preto, São Paulo, Brazil

[2]Ribeirão Preto School of Nursing, Av. Dr. Enéas de Carvalho Aguiar, 419, Brazil

[3]University Hospital Center Cova da Beira, University of Beira Interior, Covilhã, Portugal

[4]Faculdade de Medicina de São José do Rio Preto, Sao Jose do Rio Preto, Brazil

Corresponding Author:
João Daniel de Souza Menezes, MSc
Center for Studies and Development of Health Education – CEDES, Avenida Brigadeiro Faria Lima, São Jose do Rio Preto, São Paulo, Brazil

Abstract

Background: Health inequities represent a persistent and multifaceted challenge, particularly pronounced for individuals with intellectual and developmental disabilities (IDDs), including Down syndrome (DS). This population frequently faces systemic barriers to care and is at higher risk of adverse health outcomes. Despite advances, gaps persist in health care professionals' training for caring for this population. Undergraduate medical education constitutes a crucial component in addressing these disparities and promoting inclusive care, especially through practical experiences for future physicians.

Objective: This study aims to map and synthesize evidence regarding undergraduate medical education for the care of individuals with DS and other IDDs in medical schools, identifying gaps and opportunities for enhancement in curricula and educational programs, including pedagogical strategies and experiential learning opportunities.

Methods: A scoping review following the methodological frameworks by Arksey and O'Malley, Levac et al, and the Joanna Briggs Institute will be conducted. The search will be performed in electronic databases and gray literature sources using descriptors in English, Portuguese, and Spanish. Study selection will involve undergraduate medical students as the target population. Two independent reviewers will perform study selection following predefined inclusion and exclusion criteria. Data will be extracted using a standardized form and synthesized narratively, with qualitative (thematic) and descriptive quantitative analyses where appropriate. This protocol has been registered with the Open Science Framework and will adhere to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines.

Results: A comprehensive overview of current undergraduate medical education regarding the care of individuals with DS and other IDDs will be provided, informing the development of more effective and inclusive interventions and yielding insights into existing pedagogical strategies, practical learning opportunities, and medical students' knowledge and attitudes toward this population, providing a foundation for evidence-based recommendations.

Conclusions: This review will address an important gap in the literature, providing insights for the improvement of undergraduate medical education and clinical practice in caring for individuals with DS and other IDDs, contributing to the development of more capable and empathetic professionals. By systematically mapping the current landscape and identifying specific needs, this protocol lays the groundwork for transformative changes in medical curricula worldwide, ultimately enhancing health outcomes for this vulnerable population.

XSL•FO
RenderX

International Registered Report Identifier (IRRID): PRR1-10.2196/80280

(JMIR Res Protoc 2026;15:e80280) doi:10.2196/80280

KEYWORDS

medical education; Down syndrome; intellectual disability; developmental disability; scoping review protocol

Introduction

Health inequities represent a persistent and multifaceted challenge, with studies indicating that they affect a significant portion of patients [1] and are particularly pronounced for individuals with intellectual and developmental disabilities (IDDs), including Down syndrome (DS) [2,3]. This population frequently faces systemic barriers to care and is at higher risk of adverse health outcomes, which is exacerbated by the need for more intensive interactions with the health care system [4]. Despite advances in inclusion policies and disability rights, a significant gap persists in research on health care professionals' self-efficacy in caring for patients with IDDs [5]. This lack of professional preparation can perpetuate inequalities in access and quality of care, contributing to inadequate health outcomes [6]. Furthermore, people with IDDs face systemic barriers such as limited access to preventive health services and unfavorable social determinants [7]. Including the perspective of people with disabilities is essential to understand their unique experiences and challenges; however, a medicalized view still predominates, obscuring the structural and social barriers faced by this population [2]. A critical approach to disability in medical education can help contextualize these experiences and promote transformation in health care systems toward equity and inclusion [8]. The inequities faced by people with IDDs in health care are multifaceted and can be exacerbated by factors such as lack of professional confidence and skills, limited resources, and explicit and implicit biases [9]. These biases may include misconceptions about the quality of life and dignity of people with disabilities, leading to disparities in care [10].

Medical education holds a fundamental responsibility in shaping the perspectives, knowledge, and skills of future physicians to effectively address these disparities. It is within medical schools that basic science, essential clinical skills, and ethical postures toward diverse patient populations are established and developed. However, theoretical knowledge or superficial exposure to specific topics alone is insufficient. The ability to provide truly person-centered and equitable care, particularly for populations that demand individualized and multifaceted approaches, such as those with IDDs, heavily depends on specific training adapted to their unique needs.

Beyond the formal curriculum, training physicians through experiential learning in caring for people with DS and other IDDs is of profound importance and recognized as fundamental [11,12]. Direct exposure and interaction—a crucial "lived experience"—with individuals with these conditions throughout the medical school curriculum is fundamental. This practical experience offers students the following opportunities: developing empathy and understanding by directly interacting with people with IDDs and their families to transcend purely clinical descriptions and understand the daily realities and challenges and, above all, the capabilities and potential of each individual; challenging preconceived notions and biases through direct contact to help deconstruct stereotypes and misconceptions about quality of life, autonomy, and dignity, promoting a more humane and less pathologizing view; building confidence and practical skills via supervised practice in relevant environments to allow students to feel more confident in communicating, conducting adapted clinical assessments, and effectively planning care for this population; and understanding the integral context by experiencing services or community interactions that can reveal the complex interaction among health conditions, social determinants, family and community support, and environmental barriers that impact the health and well-being of people with IDDs.

Therefore, investigating the current state of medical education, and especially the inclusion of opportunities for meaningful practical experience, is of vital importance to identify how well future physicians are being prepared to care for people with IDDs in an inclusive and competent manner. This research is not merely descriptive but a fundamental step toward supporting the need for curricular and pedagogical changes.

Despite growing awareness of these issues, formal education on caring for people with IDDs remains limited in health curricula. A previous systematic review [8] highlighted the scarcity of research in this area, primarily focusing on postgraduate training and the urgent need for interprofessional and competency-based educational programs. Worldwide, various initiatives and frameworks have emerged to address this gap. These programs should adopt a critical perspective on disability and an equity and inclusion framework to guide curriculum development [13]. Innovative educational interventions such as immersive curricula and interprofessional initiatives have shown promising results in reducing prejudice and improving knowledge and skills among health care professionals [3,14]. These approaches can promote transformative learning and improve the quality of care provided to people with IDDs [11]. However, the effective implementation of these interventions requires institutional commitment and cultural change in health schools. Barriers such as lack of time, limited resources, and resistance to change can hinder the integration of these contents into already overburdened curricula [15]. Furthermore, assessing the impact of these initiatives on clinical practice and long-term health outcomes remains a challenge [12]. Despite these challenges, there are inspiring examples of successful educational programs. In the United States, a national consensus on disability health competencies was developed through extensive collaboration [16], establishing a model for training more capable and empathetic health care professionals. Similar initiatives are crucial at an international level to ensure a comprehensive and standardized approach aligned with global best practices.

XSL•FO
RenderX

In undergraduate medical education, there is a recognized scarcity of studies specifically focused on training future physicians to care for people with IDDs, including those with DS. This knowledge and research gap makes it difficult to precisely identify specific challenges and concrete opportunities for improvement in medical school curricula. Therefore, this scoping review differentiates itself from existing systematic reviews by focusing exclusively on undergraduate training and aims to map and synthesize the available evidence on medical education in this area, providing essential insights for the urgent enhancement of training and clinical practice in a global context. A preliminary search conducted in key databases (eg, PubMed and Web of Science) confirmed the absence of existing scoping or systematic reviews that comprehensively address medical education specifically for the care of individuals with DS and other IDDs within the undergraduate medical school context, thus affirming the unique contribution and pressing need for this study.

The specific objectives of this review include (1) identifying the presence of content related to the care of people with DS and other IDDs in undergraduate medical school curricula; (2) assessing the knowledge and attitudes of medical students toward this population; (3) identifying gaps and opportunities for improvement in medical training, including the identification of pedagogical strategies and opportunities for practical experience; and (4) providing evidence-based recommendations for the development and implementation of more comprehensive, effective, and inclusive educational programs.

The results of this review are expected to contribute significantly to the advancement of medical education and clinical practice in caring for people with IDDs in undergraduate medical training, promoting a more inclusive, equitable, compassionate, and evidence-based approach. Additionally, this review has the potential to identify knowledge gaps and guide future research in this area, strengthening the evidence base for policy, curricular, and pedagogical decisions.

Methods

Overview

This paper constitutes the protocol for a scoping review. A review protocol is a pre-established plan that details the methodology to be followed in conducting a systematic or scoping review. Its importance lies in ensuring transparency, reducing the risk of bias, guiding the research team, and allowing for critical evaluation of the method even before data collection begins. This scoping review will follow the methodological frameworks proposed by Arksey and O'Malley [17], Levac et al [18], and the Joanna Briggs Institute (JBI) [19], as described in the following sections.

Protocol Registration

This protocol has been registered with the Open Science Framework (OSF Z5MDB). It will adhere to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines to ensure transparency and high-quality reporting [20,21Checklist 1].

Eligibility Criteria (Inclusion and Exclusion)

Studies will be included in this scoping review based on the population, concept, and context framework.

The population will be undergraduate medical students who participate in educational programs or whose knowledge and attitudes regarding the care of individuals with IDDs are assessed. This specific focus on undergraduate medical students aims to capture foundational training experiences and curricular interventions during the initial stages of medical professional development, directly addressing the core objective of this review.

The concept includes educational programs; curricula; pedagogical interventions; training initiatives; or assessments of knowledge, skills, and attitudes related to health care for people with DS and/or other IDDs. The specific mention of DS is justified due to its prevalence as a genetic cause of intellectual disability often associated with distinct health profiles and societal perceptions, making it a critical exemplar within the broader IDD category. The consistent use of "IDDs" throughout the manuscript aims to harmonize terminology and encompass the wide range of conditions that fall under this umbrella, including autism spectrum disorder, neurodevelopmental disorders, and learning disabilities, ensuring a comprehensive review beyond solely intellectual impairment.

The context is medical schools or medical training environments at the undergraduate level.

Types of evidence will include original research studies (quantitative, qualitative, or mixed methods), reviews (eg, narrative, integrative, or systematic provided that they include primary data relevant to the population, concept, and context framework), experience reports, theses, dissertations, and other relevant gray literature (eg, government reports and conference proceedings). When reviews are included, data extraction will primarily focus on identifying and analyzing the primary studies cited within them. If a review itself provides primary data (eg, a descriptive analysis of curricula across institutions), it will be treated as a primary source for that specific data point, with careful consideration to avoid double counting of evidence.

No restriction will be applied to the year of publication. Studies published in English, Portuguese, and Spanish will be included. This language scope was chosen to balance comprehensive coverage with practical feasibility for translation and analysis by the review team. These languages represent a significant portion of the global scientific output relevant to medical education and disability, allowing for a broad international perspective while managing resources effectively.

Search Strategy

Overview

Searches will be conducted in electronic databases and gray literature sources [22] (Table 1). Descriptors in English, Portuguese, and Spanish will be used combined with Boolean operators. The search strategy will be adapted for each database, as described in Table 1. The search will cover literature published from the inception of each database to the final search date, which is planned for December 31, 2025. Updates to the

XSL•FO
RenderX

search will be performed if the review extends beyond this time frame to ensure that the most current evidence is captured.

The search process will follow the 3-stage strategy recommended by the JBI for scoping reviews [23-28].

Table . Examples of search strategies adapted for each database.

Database	Search strategy
PubMed	(“Medical Education” [MeSH[a]] OR “Medical Training” OR “Medical Teaching” OR “Medical Curriculum” OR “Undergraduate Medical Education” OR “Medical Student Education” OR “Clinical Education” OR “Experiential Learning” OR “Curriculum Development” OR “Competency Based Education” OR “Disability Education” OR “Inclusive Health” OR “Communication Training”) AND (“Down Syndrome” [MeSH] OR “Trisomy 21” OR “Intellectual Disability” [MeSH] OR “Cognitive Impairment” OR “Developmental Disability” OR “Learning Disability” OR “Autism Spectrum Disorder” OR “Neurodevelopmental Disorder” OR “Mental Retardation”)
Scopus	TITLE-ABS-KEY((“Medical Education” OR “Medical Training” OR “Medical Teaching” OR “Medical Curriculum” OR “Undergraduate Medical Education” OR “Medical Student Education” OR “Clinical Education” OR “Experiential Learning” OR “Curriculum Development” OR “Competency Based Education” OR “Disability Education” OR “Inclusive Health” OR “Communication Training”) AND (“Down Syndrome” OR “Trisomy 21” OR “Intellectual Disability” OR “Cognitive Impairment” OR “Developmental Disability” OR “Learning Disability” OR “Autism Spectrum Disorder” OR “Neurodevelopmental Disorder” OR “Mental Retardation”))
Embase	('medical education’/exp OR ’medical training’ OR ’medical teaching’ OR ’medical curriculum’ OR ’undergraduate medical education’ OR ’medical student education’ OR ’clinical education’ OR ’experiential learning’ OR ’curriculum development’ OR ’competency based education’ OR ’disability education’ OR ’inclusive health’ OR ’communication training’) AND ('Down syndrome’/exp OR ’Trisomy 21' OR ’intellectual disability’/exp OR ’cognitive impairment’ OR ’developmental disability’ OR ’learning disability’ OR ’autism spectrum disorder’ OR ’neurodevelopmental disorder’ OR ’mental retardation’)
Cochrane Library	(((“Medical Education” OR “Medical Training” OR “Medical Teaching” OR “Medical Curriculum” OR “Undergraduate Medical Education” OR “Medical Student Education” OR “Clinical Education” OR “Experiential Learning” OR “Curriculum Development” OR “Competency Based Education” OR “Disability Education” OR “Inclusive Health” OR “Communication Training”)) AND ((“Down Syndrome” OR “Trisomy 21” OR “Intellectual Disability” OR “Cognitive Impairment” OR “Developmental Disability” OR “Learning Disability” OR “Autism Spectrum Disorder” OR “Neurodevelopmental Disorder” OR “Mental Retardation”))):ti,ab,kw
Biblioteca Virtual de Saúde	(“Educação Médica” OR “Graduação Médica” OR “Ensino Médico” OR “Currículo Médico” OR “Formação Médica de Graduação” OR “Educação de Estudantes de Medicina” OR “Educação Clínica” OR “Aprendizagem Experiencial” OR “Desenvolvimento Curricular” OR “Educação Baseada em Competências” OR “Educação em Deficiência” OR “Saúde Inclusiva” OR “Treinamento em Comunicação”) AND (“Síndrome de Down” OR “Trissomia 21” OR “Deficiência Intelectual” OR “Deficiência Cognitiva” OR “Deficiência do Desenvolvimento” OR “Dificuldade de Aprendizagem” OR “Transtorno do Espectro Autista” OR “Transtorno do Neurodesenvolvimento” OR “Retardo Mental”)
Web of Science	TS=(“Medical Education” OR “Medical Training” OR “Medical Teaching” OR “Medical Curriculum” OR “Undergraduate Medical Education” OR “Medical Student Education” OR “Clinical Education” OR “Experiential Learning” OR “Curriculum Development” OR “Competency Based Education” OR “Disability Education” OR “Inclusive Health” OR “Communication Training”) AND TS=(“Down Syndrome” OR “Trisomy 21” OR “Intellectual Disability” OR “Cognitive Impairment” OR “Developmental Disability” OR “Learning Disability” OR “Autism Spectrum Disorder” OR “Neurodevelopmental Disorder” OR “Mental Retardation”)

[a]MeSH: Medical Subject Headings.

XSL•FO
RenderX

Step 1: Limited Initial Search

A preliminary search will be conducted in 2 relevant databases (eg, PubMed and Web of Science) to identify keywords and index terms (eg, MeSH [Medical Subject Headings] terms and Emtree terms) that best describe the phenomenon of interest. This initial search will help refine the comprehensive search strategy.

Step 2: Comprehensive Search

On the basis of the keywords and index terms identified in step 1, a comprehensive search strategy will be developed and executed in all selected databases listed in Table 1 (Embase, PubMed, Cochrane Library, Biblioteca Virtual de Saúde, Scopus, and Web of Science), as well as in gray literature sources. The strategy will use a comprehensive network of synonyms and related terms, including controlled vocabulary (eg, MeSH terms in PubMed and Emtree terms in Embase) and free-text terms. Adjacency and proximity operators (eg, NEAR/x and ADJ/x) will be used where appropriate to increase precision. The search terms will focus on core concepts of medical education and the target population, incorporating terms such as “disability education,” “competency,” “inclusive health,” and “communication training” within the educational component. For the population component, terms will be expanded to include various forms of developmental disabilities and historical terminology for maximal retrieval, such as “autism spectrum disorder,” “neurodevelopmental disorder,” “learning difficulties,” and “mental retardation” (used historically), to ensure a broad capture of relevant studies and avoid a narrow focus solely on cognitive decline related to aging. An example of the search strategy adapted for each database is provided in Table 1.

Step 3: Reference List Screening

The reference lists of all studies finally included in the review will be manually screened to identify any additional relevant articles not captured by the electronic searches.

Information Sources

Searches will be conducted in electronic databases and gray literature sources, including institutional repositories (eg, OpenGrey and ProQuest Dissertations and Theses Global), websites of relevant organizations (eg, World Health Organization and American Association on Intellectual and Developmental Disabilities), governmental reports (eg, national health ministries’ guidelines), accreditation body curricula (eg, Association of American Medical Colleges and World Federation for Medical Education), and competency frameworks from professional associations.

Study Selection and Data Extraction

The selection process will be conducted independently by 2 reviewers. Initial deduplication of search results will be performed using automated tools (eg, EndNote and Rayyan) before screening, followed by a manual check for residual duplicates. The process will follow 2 phases:

- Title and abstract screening—all retrieved records will be screened by title and abstract against the predefined inclusion and exclusion criteria using the Covidence software (Veritas Health Innovation). Potentially relevant studies will be moved to the next phase.
- Full-text review—the full texts of studies considered potentially relevant will be retrieved and assessed against the inclusion and exclusion criteria. Any disagreements between the 2 independent reviewers at any screening stage will be resolved through discussion and consensus. If consensus cannot be reached, a third reviewer will be consulted for final arbitration. The study selection process will be presented in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

Data will be extracted by 2 independent reviewers using a standardized form, including information about authorship, publication year, country, objective, design, population, intervention, outcomes, and results (Multimedia Appendix 1). Data extraction will be managed using a customized spreadsheet (eg, Microsoft Excel or Google Sheets). To ensure consistency and reliability of data extraction, the standardized form will be tested on a sample of approximately 5 to 10 included studies by both reviewers before the complete extraction process. Any discrepancies during data extraction will be resolved through discussion and consensus between the 2 reviewers; if necessary, a third reviewer will be consulted to mediate.

Data Synthesis

The data synthesis will be primarily narrative, providing a descriptive summary of the characteristics of the included studies. This will involve grouping studies by key themes, such as educational content, teaching methodologies, and evaluated outcomes. Descriptive quantitative analyses (eg, frequencies and percentages) will be performed to present general trends and distributions of study characteristics (eg, publication year, geographical location, study design, and target population). A thematic analysis approach will be used to identify, analyze, and report patterns and key messages related to educational programs, current gaps, and opportunities for improvement in medical training on the care of people with DS and other IDDs. The insights obtained from the included studies will be presented to map the existing evidence base. No meta-analysis will be conducted as it is outside the scope of a scoping review focused on mapping and synthesizing existing literature.

Expert Consultation

Experts in the field will be consulted to validate the findings and provide additional insights. After the initial data synthesis, a panel of 3 to 5 experts in medical education and/or the care of individuals with DS and other IDDs in medical schools will be consulted. This panel will explicitly include “experts by experience,” meaning individuals with IDDs themselves and their caregivers, recognizing their invaluable insights into lived experiences and care needs. These experts will be selected based on their clinical experience, academic contributions, and advocacy work in the area. The consultation will be conducted through semistructured interviews or an online survey to (1) validate the preliminary findings and their interpretation, (2) identify additional gaps or emerging themes not captured by the literature search, and (3) gather insights on the applicability and implications of the results for medical education and clinical

XSL•FO
RenderX

practice. The experts' contribution will be integrated into the discussion section and used to inform recommendations.

Results

The results of this review will provide a comprehensive overview of current medical training on the care of people with DS and other IDDs, identifying gaps and opportunities for improvement in curricula and educational programs. These findings can inform the development of more effective educational interventions and policies, contributing to the training of more capable and empathetic professionals in the care of this population.

An updated PRISMA-ScR flowchart will be meticulously developed and included in this section to visually and clearly represent the entire study selection process from the initial search results to the final included studies. This graphical representation will enhance the transparency and replicability of our methods, allowing readers to trace the progression of each article through the screening and eligibility phases.

Discussion

Anticipated Findings

The results of this review have the potential to impact medical education and clinical practice by identifying gaps and opportunities for improvement in training on the care of people with DS and other IDDs. However, some limitations can be anticipated, such as the heterogeneity of studies and the scarcity of research in the medical school context. Strategies to mitigate these limitations include adopting robust synthesis methods and consulting with experts to contextualize the findings.

In this section, a detailed and comprehensive discussion of the anticipated findings and their implications will be presented, aligning with the standards for rigorous academic protocols.

On the basis of the proposed methodology, we anticipate that this scoping review will reveal a significant gap in undergraduate medical education specifically tailored to the care of individuals with DS and other IDDs. We expect to identify that, while there is growing awareness of the importance of inclusive care, the implementation of curricular content and practical experiences remains inconsistent and insufficient across many medical institutions. Our findings are predicted to indicate a deficit in medical students' knowledge and confidence regarding the management of health complexities and communication nuances pertinent to this population, as well as the persistence of biases or preconceived notions potentially exacerbated by the absence of meaningful learning experiences. Furthermore, we expect to delineate various pedagogical strategies and experiential learning opportunities that, although extant, may not be widely integrated or systematically evaluated.

Comparison to Prior Work

As highlighted in the Introduction section, previous systematic reviews in this area have predominantly focused on postgraduate training or broader conceptualizations without an in-depth examination of the undergraduate medical curriculum. For instance, Adirim et al [8] conducted a systematic review focused on postgraduate medical training in IDDs. Our scoping review distinguishes itself by its exclusive focus on undergraduate medical education, thereby addressing a critical gap in the literature concerning the foundational training of future physicians. By identifying current educational content, pedagogical approaches, and student perceptions at this formative stage, we can provide a more robust basis for early and effective curricular interventions. Our findings will facilitate a direct comparison with the needs identified at more advanced training levels, pinpointing where fundamental competencies for compassionate and effective care require reinforcement.

Strengths and Limitations of the Protocol

This protocol has several methodological strengths, including adherence to established frameworks (Arksey and O'Malley [17], Levac et al [18], and the JBI) and a comprehensive search strategy across multiple databases and gray literature sources in English, Portuguese, and Spanish. Registration with the OSF and adherence to the PRISMA-ScR guidelines promote transparency and minimize bias. The inclusion of "experts by experience" (individuals with IDDs and their caregivers) in the expert consultation panel represents a strength, ensuring that lived experiences inform the results. Limitations include the potential heterogeneity and scarcity of research focused on undergraduate medical education. This scarcity may constrain quantitative synthesis, increasing reliance on narrative analysis. The search's confinement to selected languages may exclude relevant research. The nature of a scoping review implies that no meta-analysis will be conducted. Dependence on secondary data dictates that the quality of the extracted information is contingent upon its reporting in primary studies.

Future Directions and Implications for Policy and Practice

Insights generated by this review may inform future research and practical initiatives. Future directions include (1) rigorous development and evaluation of innovative educational interventions (eg, immersive curricula and interprofessional programs), (2) investigation of barriers to and facilitators of curricular changes in medical schools, and (3) longitudinal studies assessing the impact of enhanced training on clinical practice and health outcomes for individuals with DS and other IDDs. The results may inform educational policies and guidelines for accreditation bodies aiming to integrate essential competencies for the care of this population into undergraduate curricula. The ultimate goal is to foster an educational ecosystem that produces more prepared, empathetic, and inclusive professionals.

Dissemination Plan

A robust dissemination plan will be implemented. Results will be submitted for publication in leading peer-reviewed journals specializing in medical education and disability health. The findings will be presented at national and international scientific conferences and workshops. Executive summaries and communication materials will be developed for curriculum committees, professional councils, and disability advocacy organizations. Making the protocol and data available on the OSF will ensure public access and transparency. Active

XSL•FO
RenderX

engagement with communities of people with disabilities and their caregivers in dissemination is intended to ensure that the research is informed by the needs derived from their lived experiences.

Conclusions

This scoping review will address an important gap in the literature by analyzing medical training on the care of people with DS and other IDDs in medical school. The results may support improvements in curricula and educational programs, contributing to the training of more capable and empathetic professionals in the care of this population at higher risk of adverse health outcomes.

The contribution of this work is fundamental to reshaping undergraduate medical education, fostering a more inclusive and equitable approach to the care of individuals with disabilities. By providing a detailed map of the literature and identifying critical areas for intervention, this review has the potential to directly influence the training of the next generation of physicians, preparing them to offer person-centered health care. This initiative strengthens the evidence base for curricular and pedagogical decisions and catalyzes the necessary transformation to achieve more inclusive and compassionate health care.

The next steps include conducting searches and selecting studies, extracting and synthesizing data, consulting with experts, and disseminating the results in journals and conferences. This review is expected to be completed within 12 months.

Acknowledgments

No generative artificial intelligence tools were used in the writing, editing, or preparation of this manuscript.

Funding

This research was supported by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Coordination for the Improvement of Higher Education Personnel), Brazil, under grant 88887.902237/2023-00.

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Authors' Contributions

Data curation: RCHMR, EFC, IAAB, MARF, ACSC, AFPS, RFF
Formal analysis: JASF, ERS, JDSM, MQS, DNLA, EFC, AKS
Funding acquisition: MARF, JDSM, MAB, LCFS
Investigation: WDM, PBSJ, AMZ, NBGC, ABQ, RMR, DCMVO, LCFS
Methodology: ACSC, AFPS, RFF, JASF, MQS, DNLA, EFC, AKS
Project administration: IAAB, ERS, AHO, JCA
Resources: EFC, AKS, WDM, RMR
Supervision: MAB, LCFS, RCHMR, ECP, JCA
Visualization: AMZ, NBGC, ABQ, RMR, DCMVO, AHO
Writing—original draft: IAAB, MARF, ACSC, AFPS, ERS, JDSM
Writing—review and editing: RFF, JASF, MQS, DNLA, EFC, AKS, WDM, PBSJ

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data extraction tool.
[DOCX File, 21 KB - resprot_v15i1e80280_app1.docx]

Checklist 1
PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist.
[DOCX File, 110 KB - resprot_v15i1e80280_app2.docx]

References

1. Nong P, Raj M, Creary M, Kardia SLR, Platt JE. Patient-reported experiences of discrimination in the US health care system. JAMA Netw Open 2020 Dec 1;3(12):e2029650. [doi: 10.1001/jamanetworkopen.2020.29650] [Medline: 33320264]
2. Iezzoni LI, Rao SR, Ressalam J, et al. Physicians' perceptions of people with disability and their health care. Health Aff (Millwood) 2021 Feb;40(2):297-306. [doi: 10.1377/hlthaff.2020.01452] [Medline: 33523739]

XSL•FO
RenderX

3. Lagu T, Haywood C, Reimold K, DeJong C, Walker Sterling R, Iezzoni LI. 'I am not the doctor for you': physicians' attitudes about caring for people with disabilities. Health Aff (Millwood) 2022 Oct;41(10):1387-1395. [doi: 10.1377/hlthaff.2022.00475] [Medline: 36190896]
4. Murphy NA, Alvey J, Valentine KJ, Mann K, Wilkes J, Clark EB. Children with medical complexity: the 10-year experience of a single center. Hosp Pediatr 2020 Aug;10(8):702-708. [doi: 10.1542/hpeds.2020-0085] [Medline: 32699000]
5. Roux AM, Rast JE, Shea LL. Family perceptions of health care access for autistic young adults receiving disability services. Pediatrics 2022 Apr 1;149(Suppl 4):e2020049437S. [doi: 10.1542/peds.2020-049437S] [Medline: 35363289]
6. Atenção à saúde da pessoa com deficiência no sistema único de saúde - SUS. Ministério da saúde Departamento de ações programáticas estratégicas; Brasil. 2009. URL: https://www.gov.br/saude/pt-br/search?origem=form&SearchableText=Atenção [accessed 2025-05-29]
7. Agrawal S, Kalocsai C, Capponi P, et al. "It was great to break down the walls between patient and provider": liminality in a co-produced advisory course for psychiatry residents. Adv in Health Sci Educ 2021 May;26(2):385-403. [doi: 10.1007/s10459-020-09991-w]
8. Adirim Z, Sockalingam S, Thakur A. Post-graduate medical training in intellectual and developmental disabilities: a systematic review. Acad Psychiatry 2021 Jun;45(3):371-381. [doi: 10.1007/s40596-020-01378-8] [Medline: 33433827]
9. Agic B, Fruitman H, Maharaj A, et al. Advancing curriculum development and design in health professions education: a health equity and inclusion framework for education programs. J Contin Educ Health Prof 2023;43(4S):S4-S8. [doi: 10.1097/CEH.0000000000000453] [Medline: 35940600]
10. Medical futility and disability bias: part of the bioethics and disability series. National Council on Disability (NCD). 2019. URL: https://www.ncd.gov/assets/uploads/reports/2019/ncd_medical_futility_report_508.pdf [accessed 2025-05-05]
11. Kirshblum S, Murray R, Potpally N, Foye PM, Dyson-Hudson T, DallaPiazza M. An introductory educational session improves medical student knowledge and comfort levels in caring for patients with physical disabilities. Disabil Health J 2020 Jan;13(1):100825. [doi: 10.1016/j.dhjo.2019.100825] [Medline: 31401186]
12. Havercamp SM, Barnhart WR, Robinson AC, Whalen Smith CN. What should we teach about disability? National consensus on disability competencies for health care education. Disabil Health J 2021 Apr;14(2):100989. [doi: 10.1016/j.dhjo.2020.100989] [Medline: 32952097]
13. Chardavoyne PC, Henry AM, Sprow Forté K. Understanding medical students' attitudes towards and experiences with persons with disabilities and disability education. Disabil Health J 2022 Apr;15(2):101267. [doi: 10.1016/j.dhjo.2021.101267] [Medline: 35094955]
14. Bowen CN, Havercamp SM, Karpiak Bowen S, Nye G. A call to action: preparing a disability-competent health care workforce. Disabil Health J 2020 Oct;13(4):100941. [doi: 10.1016/j.dhjo.2020.100941] [Medline: 32467076]
15. Rosenberg S, Carulla LS, Rosen A. Mental health reform in Australia - unfinished business. BJPsych Int 2023 Nov;20(4):99-101. [doi: 10.1192/bji.2023.19] [Medline: 38029430]
16. Hamilton P, Coey-Niebel C, McCaig J, et al. Evaluation of Inter-Professional Education (IPE) with medical, nursing and pharmacy students through a simulated IPL educational intervention. Int J Clin Pract 2021 Nov;75(11):e14725. [doi: 10.1111/ijcp.14725] [Medline: 34382304]
17. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
18. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
19. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
20. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
21. Foster ED, Deardorff A. Open Science Framework (OSF). J Med Libr Assoc 2017;105(2). [doi: 10.5195/jmla.2017.88] [Medline: 5370619]
22. Paez A. Gray literature: an important resource in systematic reviews. J Evid Based Med 2017 Aug;10(3):233-240. [doi: 10.1111/jebm.12266] [Medline: 28857505]
23. Babineau J. Product review: covidence (systematic review software). J Can Health Libr Assoc 2014;35(2):68. [doi: 10.5596/c14-016]
24. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341. [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]
25. Munn Z, Tufanaru C, Aromataris E. JBI's systematic reviews. Am J Nurs 2014;114(7):49-54. [doi: 10.1097/01.NAJ.0000451683.66447.89]
26. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews a product from the ESRC methods programme. Lancaster University. 2006. URL: https://www.researchgate.net/publication/233866356_Guidance_on_the_conduct_of_narrative_synthesis_in_systematic_reviews_A_product_from_the_ESRC_Methods_Programme [accessed 2025-05-25]

XSL•FO
RenderX

27. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
28. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. Introduction to Meta‐Analysis, 2nd edition: Hoboken, USA; 2021. [doi: 10.1002/9781119558378]

Abbreviations

DS: Down syndrome
IDD: intellectual or developmental disability
JBI: Joanna Briggs Institute
MeSH: Medical Subject Headings
OSF: Open Science Framework
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by J Sarvestan; submitted 07.Jul.2025; peer-reviewed by RC Cvejic, R Sheehan; revised version received 28.Nov.2025; accepted 30.Nov.2025; published 16.Mar.2026.

Please cite as:
Amaral de Almeida Bistafa I, Antonio Ribeiro Filho M, dos Santos Costa AC, Pimenta Saad AF, Ferreira Franco R, Aparecida de Sales Floriano J, Roberto dos Santos E, de Souza Menezes JD, Querino da Silva M, Nunes Longhi Aleixo D, Fazuoli Chubaci E, Karina Stefani A, Donegá Martinez W, Belchior da Silveira Junior P, Mirtes Zecchin A, Gonçalves Cristóvão NB, Bertolazzo Quitério A, Mendonça Ribeiro R, Cristina Móz Vaz Oliani D, Hélio Oliani A, Alahmar Bianchin M, Cesar Fava Spessoto L, Helú Mendonça Ribeiro RDC, César André J
Strengthening Undergraduate Medical Education for Inclusive Health Care for People With Down Syndrome and Intellectual and Developmental Disabilities in Medical Schools: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e80280
URL: https://www.researchprotocols.org/2026/1/e80280
doi:10.2196/80280

© Isabela Amaral de Almeida Bistafa, Marco Antonio Ribeiro Filho, Ana Carolina dos Santos Costa, Ana Flávia Pimenta Saad, Rauer Ferreira Franco, Janaína Aparecida de Sales Floriano, Emerson Roberto dos Santos, João Daniel de Souza Menezes, Matheus Querino da Silva, Daniele Nunes Longhi Aleixo, Eliana Fazuoli Chubaci, Andressa Karina Stefani, William Donegá Martinez, Pedro Belchior da Silveira Junior, Ana Mirtes Zecchin, Nathalia Bavaresco Gonçalves Cristóvão, Alex Bertolazzo Quitério, Renato Mendonça Ribeiro, Denise Cristina Móz Vaz Oliani, Antônio Hélio Oliani, Maysa Alahmar Bianchin, Luís Cesar Fava Spessoto, Rita de Cássia Helú Mendonça Ribeiro, Júlio César André. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 16.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Public Engagement Strategies in Digital Health Ethics: Protocol for a Scoping Review

Stella Namuganza[1,2], MSc; Mariacarla Gadebusch Bondio[1*], Dr rer med, Dr phil; Philipp Kellmeyer[3,4*], MD, MPhil, Dr med

[1]Institute for Medical Humanities, University Hospital Bonn, Bonn, Germany

[2]Hertie Institute for AI in Brain Health, University of Tübingen, Tübingen, Germany

[3]Data and Web Science Group, School of Business Informatics and Mathematics, University of Mannheim, B6, 26, Mannheim, Germany

[4]Human-Technology Interaction Lab, Department of Neurosurgery, Faculty of Medicine, University Medical Center Freiburg, Freiburg im Breisgau, Germany

[*]these authors contributed equally

Corresponding Author:
Philipp Kellmeyer, MD, MPhil, Dr med
Data and Web Science Group, School of Business Informatics and Mathematics, University of Mannheim, B6, 26, Mannheim, Germany

Abstract

Background: The ethical, legal, and social issues accompanying the latest advancements in digital health technologies highlight the need to involve the public in their design, development, and deployment to align with societal values and needs. For public engagement to be meaningful, it should be participatory, inclusive, and scalable. However, studies in participatory digital health do not characterize public engagement strategies in terms of scalability, representativeness, and the extent of participation. Moreover, no reviews have examined how ethical debates shape the design and implementation of public engagement strategies in digital health ethics.

Objective: The objectives of the planned study based on this protocol are to (1) identify approaches used to engage the public in digital health ethics; (2) characterize approaches used in digital health ethics in terms of scalability, representativeness, and extent of participation; and (3) explore the role of ethics in the design and implementation of participatory methods.

Methods: The research will be undertaken using the Joanna Briggs Institute scoping review method. To identify relevant literature, the academic databases PubMed, ScienceDirect, IEEE Xplore, and Web of Science will be searched for articles published from January 1, 2015, to November 30, 2025. All retrieved papers will be uploaded to the Rayyan software. Duplicates will be removed, and subsequently, 2 reviewers will independently screen titles and abstracts, followed by full-text screening using a hybrid verification model. Data will be extracted on a spreadsheet, with rows representing individual studies and columns capturing categories of extracted information.

Results: Currently, the search has yielded 1352 articles; at the title and abstract screening phase, 1291 (95.5%) articles have been excluded and 61 (4.5%) have been included. Anticipated results are the number of studies, bibliographic details of the studies (ie, author, publication year, journal, and country), participant characteristics, digital health technology type, participatory methods, and media, as well as the embodiment of participatory, representative, and scalable characteristics in engagement methods and how ethical debates influence the design of public engagement strategies.

Conclusions: This protocol outlines methodology for a scoping review mapping the characteristics and ethics of public engagement methods to improve participatory and ethically responsible innovation in digital health.

International Registered Report Identifier (IRRID): DERR1-10.2196/86280

(JMIR Res Protoc 2026;15:e86280) doi:10.2196/86280

KEYWORDS

digital health; digital health ethics; public engagement; participation; representativeness; scalability; artificial intelligence; AI

Introduction

Background

The convergence of digital health technologies and recent advancements in artificial intelligence (AI) promises substantial improvement in the health care sector. Nonetheless, these innovations are accompanied by ethical, legal, and social concerns [1], which have awakened public skepticism [2] and heightened ethical scrutiny from the public. These dynamics underscore the need to engage the public in the design,

XSL•FO
RenderX

development, and deployment of digital health technologies to ensure that innovations align with societal values and needs.

Public engagement is widely recognized as an essential approach for identifying complex human experiences as well as ethical, legal, and social aspects associated with digital health tools [3-7]. Its goals extend beyond resolution of controversy, public education, democratization of science, widening the representation of voices, enabling responsible innovation, and shaping policy [8]. A range of approaches have been used to engage the public in debates on digital health, including consultation of patient organizations [9], educational campaigns, social media [10], and mainstream media outputs (talks, publications, and podcasts), as well as participatory co-design methods [11] drawing on workshops, think - aloud interviews, focus groups, and surveys to facilitate engagement both in person and online [12].

Two prominent frameworks, ethics by design and embedded ethics, have been proposed to promote the ethical development of digital health tools [13]. They aim to integrate ethical reflection directly into the development and implementation of digital technologies rather than treating ethics as a retrospective or external evaluation. Ethics by design emphasizes the proactive incorporation of ethical values—such as fairness, transparency, and respect for autonomy—into system architectures, design choices, and development processes. Embedded ethics, in contrast, focuses on systematically integrating ethicists or ethical expertise into interdisciplinary research and development teams, enabling continuous ethical reflection throughout the innovation life cycle [14]. Both approaches have been proposed as means to align technological innovation with societal values and are increasingly discussed in the context of digital health and medical AI. These frameworks have been criticized for their top-down orientation, which imposes normative principles externally rather than allowing them to emerge from users' lived experiences. As an alternative, community-led approaches should actively involve end users or potential end users from the earliest stages of design to ensure that local values and social contexts are embedded during the innovation processes [15]. As different communities acknowledge and push for the agenda of public participation, including research governance [16], there is a shift from simply acknowledging the need for public participation to considering how it can be meaningfully achieved [17]. This entails examining the facets that make public participation meaningful, which include but are not limited to inclusivity, representativeness, empowerment, and scalability [15,18]. When these conditions are met, public engagement in the domain of digital health not only fosters ethical and socially responsible innovation but also contributes to improved health outcomes [19,20].

A preliminary search conducted on July 15, 2025, in PubMed, Scopus, and Web of Science retrieved primary research papers on methods used to engage the public in the context of digital health ethics; however, there are no review papers mapping different methods used to engage the public in digital health ethics and characterizing them in terms of scalability, representativeness, and how participatory they are. Moreover, there are no papers reviewing ethical debates to determine how they influence the design of public engagement strategies. However, the growing practice of public engagement in digital health ethics requires critical inquiry into the quality and standards of the methods used to engage the public.

Therefore, this study will provide a comprehensive overview of the approaches used to engage end users and potential end users in discussions on digital health ethics and characterize the methods in terms of scalability, representativeness, and the extent to which they are participatory. Additionally, the study will determine how ethical debates unfold in the design and implementation of participatory methods. This will establish a foundation for enhancing public participation in digital health ethics. This research aims to foster effective public engagement strategies in the field of digital health ethics. This protocol outlines the methodology and reporting plan for a scoping review aimed at addressing these aims.

Operational Definition of Terms

In this study, "digital health ethics" refers to the field of practice concerned with addressing ethical, legal, and social issues in the design, governance, development, and implementation of digital health technologies. This includes ethics-by-design approaches, embedded ethics, value-sensitive design, and other structured efforts to operationalize ethical reflection in digital health innovation. It ensures that these tools benefit patients and society while minimizing harm. Key ethical considerations include privacy and data security, algorithmic bias and fairness, informed consent, patient and professional autonomy, transparency and accountability, and equitable access to digital health solutions.

In this study protocol, "articles" refers to records and publications identified through database searching and screening, while "studies" refers to the empirical research reports included in the scoping review and subjected to data extraction and synthesis. Table 1 provides a full list of operational definitions of terms.

XSL•FO
RenderX

Table . Operational definition of terms.

Term	Definition
Digital health	Encompasses health-related AI[a], mobile health, telemedicine, electronic health records, online health communities, and big data analytics used to enhance health care and improve health outcomes [21,22].
Public engagement methods	Defined as an organized procedure through which certain members of the public (users, affected communities, citizens, and stakeholders) are invited to contribute or exchange with technology developers their views, knowledge, or decisions about AI or algorithmic systems (eg, workshops and interviews) through media that enable the engagement to happen (eg, face-to-face, social media, and digital platforms) [17].
Participation	Defined as the active involvement of end users in the development and implementation of digital health technologies, including transferring decision-making power to these users [23,24].
Representativeness	The degree to which the participants of an engagement activity, such as a workshop, reflect the broader population they are drawn from in terms of experiences and needs [25].
Scalability	Refers to the expansion of an intervention from a small or controlled setting to a broad real-world application without loss of efficacy [26].

[a]AI: artificial intelligence.

Objectives of the Study

The objectives of this study are as follows:

1. To identify approaches (methods and media) used to engage the public in digital health ethics
2. To characterize approaches used in digital health ethics in terms of scalability, representativeness, and extent of participation
3. To explore how ethical debates play a role in the design and implementation of participatory methods

Research Questions

The following research questions (RQs) correspond directly to the study objectives and guide evidence charting and synthesis:

1. What methods are used to engage the public in digital health ethics?
2. What media or communication channels are used to engage the public in digital health ethics?
3. To what extent are the different public engagement methods truly participatory?
4. How scalable and representative are the methods used to engage the public in digital health ethics?
5. How do ethical debates play a role in the design and implementation of participatory methods?

Methods

Overview

The research will be undertaken using the Joanna Briggs Institute (JBI) scoping review method. This framework was first proposed by Arksey and O'Malley [27] and later modified by Levac et al [28] to provide details about what occurs at each stage of the review process. Peters et al [29] further enhanced this approach into the JBI framework to connect eligibility criteria with research objectives. We adopted a scoping review approach because participatory design in digital health is an emerging and conceptually diverse field that benefits from systematic mapping of existing evidence. Therefore, we intend to map the breadth of existing evidence. Given the exploratory nature of our questions and the heterogeneity of the evidence, a scoping review is appropriate to map the literature and systematically identify knowledge gaps. To ensure methodological robustness and transparency, we adopt a systematic approach guided by the *JBI Manual for Evidence Synthesis* for scoping reviews [29]. In addition, the review will be reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [30].

Search Strategy

Title and abstract screening will be conducted using Rayyan (Qatar Computing Research Institute) [31], a web-based software platform designed to facilitate systematic and scoping reviews by enabling independent, blinded screening of records and efficient resolution of reviewer disagreements.

The academic databases PubMed, ScienceDirect, IEEE Xplore, and Web of Science will be searched for studies published from January 1, 2015, to November 30, 2025. The year 2015 marks the onset of three converging developments: (1) the mainstreaming of the Internet of Things and wearables in health care, (2) the first wave of AI-driven digital health applications and policy frameworks (eg, the European Union's Digital Single Market Strategy), and (3) a growing shift in research funding and governance frameworks toward "public and patient involvement" as a quality criterion [16,32,33]. Together, these mark a distinct phase in the evolution of digital health that justifies our temporal scope.

The search terms in Table 2 will be used to search for papers. Reference lists of relevant studies will also be searched for eligible studies. The search strings were adapted to the specific syntax and indexing conventions of each database (eg, MeSH

XSL•FO
RenderX

[Medical Subject Headings] terms in PubMed) rather than applying a single identical string across all platforms.

Table . Search terms.

Database	Search terms
PubMed, ScienceDirect, and Web of Science	• (Digital OR Artificial intelligence) AND participatory ethics • Digital participatory codesign AND Health AND Ethics • (AI ethics OR ethical AI) AND (co-design OR participatory design) AND (inclusion OR empowerment OR representation) • (responsible research and innovation OR RRI) AND (health OR AI in health) AND (participatory OR user involvement) • (deliberative democracy OR deliberative engagement) AND (digital health OR health technology) AND (ethics OR value alignment) • (AI ethics OR artificial intelligence AND ethics) AND (public engagement OR stakeholder participation) AND (inclusion OR justice OR responsible innovation) • (digital health OR e-health) AND (community engagement) AND (ethics OR ethical framework OR value-sensitive design)
IEEE Xplore	• Digital participatory codesign AND Health AND Ethics • “digital health” OR “health technology” OR “Artificial intelligence AND ethics AND participatory” OR “public engagement” AND health • (“responsible research and innovation” OR RRI) AND (health OR “AI in health”) AND (participatory OR “user involvement”)

Eligibility Criteria

The eligibility criteria were developed using the population, concept, and context framework (Textbox 1). The population comprises end users or potential end users of digital health technologies (patients, health care consumers, or the lay public involved in digital health). The concept is the public engagement strategies in the context of digital health ethics (ethics-by-design, embedded ethics, governance consultations, or other structured efforts to address ethical, legal, and social issues). Context is any setting where digital health technologies are designed or implemented (with no geographic limitations). The search is limited to publications from January 1, 2015, to November 30, 2025.

Textbox 1. Eligibility criteria.

Inclusion criteria

- Empirical studies using qualitative, quantitative, or mixed methods with primary data collection from end users or potential end users of digital health technologies
- Studies engaging or collecting data from end users (eg, patients, citizens, and consumers) in the context of digital health
- Studies addressing digital health technologies
- Studies incorporating ethical, public engagement, or participatory design elements
- Articles in English

Exclusion criteria

- Pure ethics papers, conceptual papers, or reviews (systematic, scoping, or narrative reviews) as we aim to map primary evidence; however, for the background discussion, conceptual or normative papers that propose frameworks or arguments (even if not presenting new data) will be included as these inform research question 5
- Nonscholarly sources lacking primary data, such as newspaper articles, blogs, magazine articles, comments, letters, editorials, and review articles
- Studies that collect data only from non–end users of digital health technologies and do not include end users as participants
- Articles that do not include any explicit discussion of ethical considerations, public engagement activities, or participatory design methods in relation to digital health technologies
- Publications outside the time frame of January 1, 2015, to November 30, 2025, reflecting a period of notable growth in relevant literature
- Research protocols or proposals
- Studies not addressing digital health issues
- Animal studies

XSL•FO
RenderX

Conceptual or normative publications are not included in the scoping review's data extraction or synthesis and, therefore, are excluded as evidence sources. Such literature is consulted solely to inform the conceptual framing and interpretation of ethical themes identified within the included empirical studies, particularly in relation to RQ 5. Accordingly, the inclusion criteria focus on studies that explicitly intersect public engagement and digital health ethics within the health domain, whereas conceptual contributions on digital ethics or participation alone are used only for contextual framing and are not included in the data extraction phase.

Selection Process

All retrieved papers from the different databases will be uploaded to Rayyan to ensure a transparent record of all processes and blind screening. Duplicates will be removed, and subsequently, 2 reviewers will screen titles and abstracts. Any disagreements will be resolved through discussions among the 3 reviewers until consensus is reached regarding inclusion or exclusion (interrater disagreements will be documented for transparency). Articles identified as potentially relevant based on title and abstract screening will be retrieved in full text and assessed for eligibility by 1 reviewer, with uncertainties deliberated in team meetings to ensure agreement. Only articles in the English language that the team can reliably understand will be selected. Only studies meeting the inclusion criteria following full-text screening will be retained for the final analysis. A detailed record of the selection process will be reported in the final publication using a PRISMA-ScR flowchart. Reasons for exclusion at the full-text stage will be recorded.

Data Extraction

A structured data extraction spreadsheet will be developed by 1 reviewer initially in consultation with the review team. The data extraction sheet will be piloted across 2 to 3 studies to ensure that all authors agree on the interpretation of the categories and refine it as needed. Two reviewers will independently chart data from the first few studies and compare results to ensure consistency before 1 reviewer completes charting for the remaining studies (with a second reviewer verifying a sample of the extractions for accuracy). The spreadsheet will be organized with rows representing individual studies and columns capturing categories of extracted information, including bibliographic details, study design, target population, and digital health context (eg, technology type). Additional columns will capture conceptual data aligned with the RQs. Specifically, for RQ 1, data on engagement methods will be collected and charted; for RQ 2, information on engagement media will be extracted; for RQ 3, the degree of participation will be mapped against the ladder of participation (ranging from information, consultation, partnership, and co-design to community-led approaches) [18,34]. For RQ 4, any descriptions of the scale of engagement (eg, number of participants and multisite vs single site) and notes on representativeness (eg, demographics of participants compared to the target population) will be extracted and charted to assess scalability and representativeness. If a study does not explicitly address scalability or representativeness, we will record that these were not mentioned. For RQ 5, conceptual and normative publications will be consulted to interpret how ethical frameworks within digital health ethics practice shape public engagement strategies. Any additional data deemed relevant to addressing the RQs but not captured within these predefined categories will also be recorded and analyzed.

Analysis and Presentation of Evidence

Data will be analyzed using descriptive and thematic approaches. A descriptive numerical summary of the included studies (eg, by engagement method, country, or technology type) will be presented in tables and figures to provide an overview of the evidence base (Multimedia Appendix 1).

In the qualitative synthesis, findings will be organized by RQ. Engagement methods and media (RQs 1 - 2) will be summarized and mapped to their level of participation using a modified model of Arnstein's ladder of participation [34], which provides a conceptual continuum from information and consultation to partnership and community control. Issues of scalability and representativeness (RQ 4) will be analyzed by comparing the scope and demographic breadth of participation reported across the studies.

For RQ 5, we will conduct a thematic content analysis [35] of ethical debates and frameworks described in the included literature. The analysis will be partly deductive, guided by key constructs from responsible research and innovation [36] and value-sensitive design [37], and partly inductive to capture emergent ethical themes. These may include privacy, fairness, autonomy, trust, and accountability as recurring normative dimensions of public engagement.

Results will be presented as a combination of quantitative summaries (tables and figures) and narrative thematic synthesis. Consistent with scoping review guidance, we will not conduct a formal risk-of-bias assessment as the purpose is to map all relevant evidence irrespective of methodological quality. However, we will note any apparent limitations of the evidence base in the discussion.

The optional consultation stage will take place after data analysis through discussions with digital health ethics experts and community representatives to validate interpretations and enhance practical relevance.

Any protocol amendments made during the review will be documented in the published record, including the date and rationale. Findings will be interpreted within the broader theoretical discourse on participatory and responsible digital health innovation.

Ethical Considerations

This study analyzes publicly available data and does not involve human participants; therefore, ethics approval was not required.

Results

The timeline for this research is as follows: retrieval and screening of articles will be conducted between July 15, 2025, and November 30, 2025; full-text screening will take place between November 15, 2025, and December 15, 2025; and data extraction and analysis will be conducted in January 2026.

XSL•FO
RenderX

Figure 1 shows a preliminary version of the PRISMA-ScR flowchart.

Figure 1. PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) flowchart. RQ: research question.

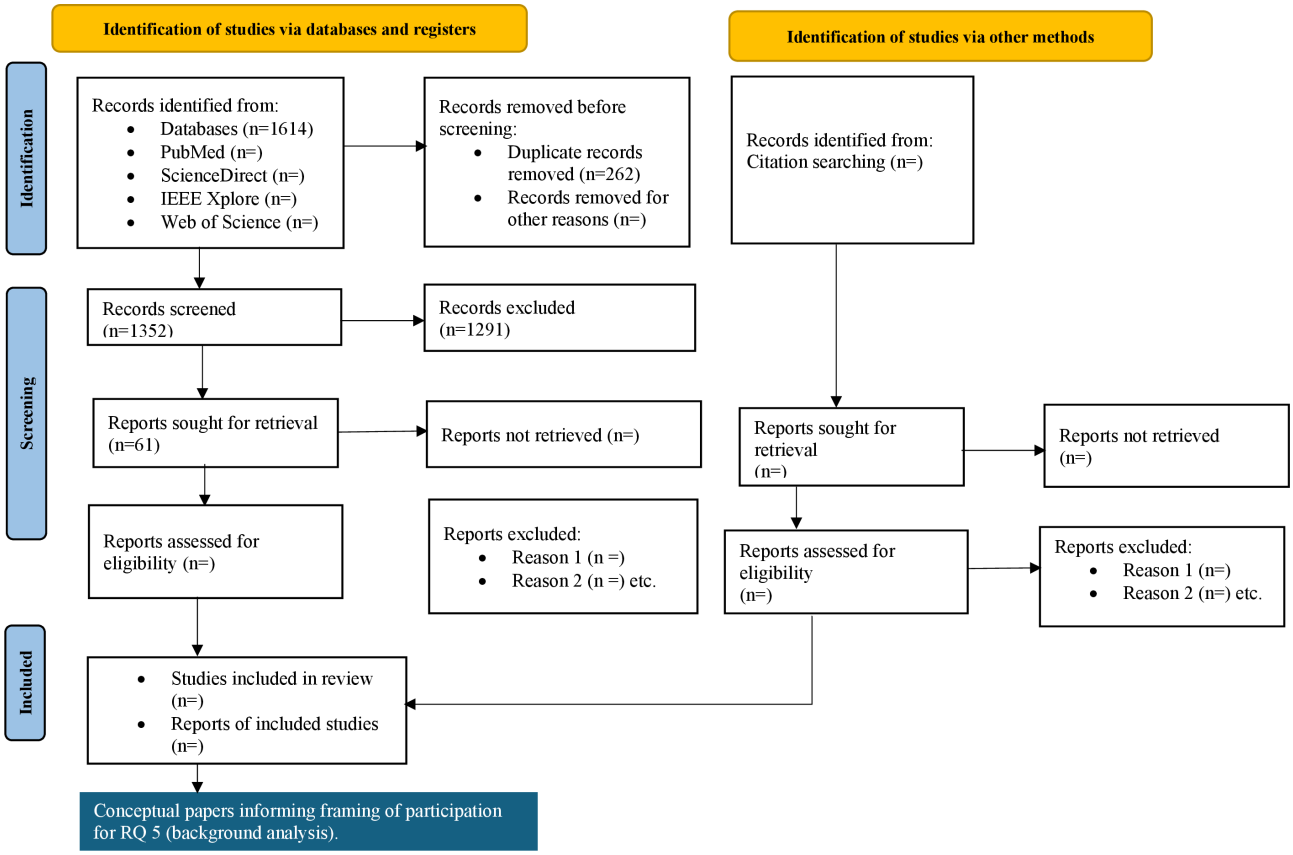

Results are expected to be submitted for publication in the first quarter of 2026. Currently, the search has yielded 1352 articles; at the title and abstract screening phase, 1291 (95.5%) articles have been excluded, and 61 (4.5%) have been included. Funding for this scoping review is blinded. Anticipated results are number of studies, bibliographic details of the studies (ie, author, publication year, journal, and country) of the studies, participant characteristics, funding source, digital health technology type, participatory methods, and media. Additionally, we anticipate results on the embodiment of participatory, representativeness, and scalability characteristics by engagement method and how ethical debates influence the design of public engagement strategies.

Discussion

Anticipated Findings

This scoping review will provide the first comprehensive mapping of public engagement strategies in digital health ethics. By systematically characterizing engagement approaches according to their participatory depth, scalability, and representativeness, the review will identify methodological gaps and highlight examples of effective, community-led practices while comparing them with previous literature. The findings are expected to inform researchers, developers, and policymakers seeking to integrate participatory and ethically grounded approaches into digital health innovation.

Beyond methodological insights, the review will also contribute conceptually to the discourse on digital health ethics by clarifying how ethical debates, such as those on fairness, transparency, and autonomy, shape the design and implementation of engagement strategies. The resulting synthesis will serve as an evidence base for refining participatory frameworks in digital health research and practice.

Limitations

Several limitations should be acknowledged. The availability of English-language abstracts and database indexing constrains coverage. Second, the search will rely primarily on published and indexed literature; while targeted gray literature searches will be conducted, relevant reports or policy documents may remain undiscovered. Third, consistent with scoping review methodology, the review will not include a formal critical appraisal of study quality, which may limit the ability to assess the robustness of individual study findings. Finally, because participatory and ethical terminology varies widely across disciplines, some potentially relevant studies may use different descriptors and, thus, be missed despite a comprehensive search strategy.

Conclusions

This protocol outlines the rationale and methodology for a scoping review that will systematically map public engagement strategies in digital health ethics. By identifying and characterizing participatory approaches across diverse contexts, the review seeks to clarify how ethical considerations are

XSL•FO
RenderX

embedded in engagement strategy design and implementation. The anticipated findings will provide an evidence-informed foundation for improving participatory, ethically responsible innovation in digital health.

Funding

SN was partly funded by a grant (RHUNE, 01GP2310D) of the German Ministry of Research, Technology and Space (BMFTR). Otherwise, no external financial support or grants were received from any public, commercial, or not-for-profit entities for the research, authorship, or publication of this paper.

Data Availability

The data analyzed in this study will consist exclusively of publicly available published literature identified through database searches (eg, PubMed, Scopus, and ScienceDirect). No primary data will be collected. The datasets generated during this study, including the list of included studies, the data charting template, and the data extraction spreadsheet, will be deposited in an open-access repository (eg, the Open Science Framework or the institutional repository of the University of Bonn) upon completion of the review. The persistent identifier (DOI or URL) will be provided in the final published version of the manuscript. Additional materials will be available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data extraction framework.
[DOCX File, 142 KB - resprot_v15i1e86280_app1.docx]

References

1. Zhang J, Zhang ZM. Ethics and governance of trustworthy medical artificial intelligence. BMC Med Inform Decis Mak 2023 Jan 13;23(1):7. [doi: 10.1186/s12911-023-02103-9] [Medline: 36639799]
2. Aicardi C, Fothergill BT, Rainey S, Stahl BC, Harris E. Accompanying technology development in the Human Brain Project: from foresight to ethics management. Futures 2018 Sep;102:114-124. [doi: 10.1016/j.futures.2018.01.005]
3. Weinberger N, Woll S, Kyba CCM, Schulte-Römer N. The value of citizen participation in technology assessment, responsible research and innovation, and sustainable development. Sustainability 2021;13(21):11613. [doi: 10.3390/su132111613]
4. Lampariello R, Ancellin-Panzani S. Mastering stakeholders' engagement to reach national scale, sustainability and wide adoption of digital health initiatives: lessons learnt from Burkina Faso. Fam Med Community Health 2021 Jun;9(3):e000959. [doi: 10.1136/fmch-2021-000959] [Medline: 34144970]
5. Williams CJ, Duff J. The roadblocks to success – identifying challenges in implementing a surgery support e-health solution: a qualitative interview study. Perioper Care Oper Room Manag 2024 Dec;37:100427. [doi: 10.1016/j.pcorm.2024.100427]
6. Tzimourta KD. Human-centered design and development in digital health: approaches, challenges, and emerging trends. Cureus 2025 Jun 13;17(6):e85897. [doi: 10.7759/cureus.85897] [Medline: 40548156]
7. Fiordelli M. Transitioning perspectives in digital health through phenomenology integration. J Med Internet Res 2024 Oct 23;26:e62691. [doi: 10.2196/62691] [Medline: 39442170]
8. Srinivas KR. Why public engagement matters in science. Trends Biotechnol 2017 Apr;35(4):281-283. [doi: 10.1016/j.tibtech.2017.02.004] [Medline: 28268033]
9. Barony Sanchez RH, Bergeron-Drolet LA, Sasseville M, Gagnon MP. Engaging patients and citizens in digital health technology development through the virtual space. Front Med Technol 2022 Nov 25;4:958571. [doi: 10.3389/fmedt.2022.958571] [Medline: 36506474]
10. Teodorowski P, Rodgers SE, Fleming K, Frith L. Use of the hashtag #DataSavesLives on Twitter: exploratory and thematic analysis. J Med Internet Res 2022 Nov 15;24(11):e38232. [doi: 10.2196/38232] [Medline: 36378518]
11. Donia J, Shaw JA. Co-design and ethical artificial intelligence for health: an agenda for critical research and practice. Big Data Soc 2021 Jul;8(2). [doi: 10.1177/20539517211065248]
12. Baines R, Bradwell H, Edwards K, et al. Meaningful patient and public involvement in digital health innovation, implementation and evaluation: a systematic review. Health Expect 2022 Aug;25(4):1232-1245. [doi: 10.1111/hex.13506] [Medline: 35526274]
13. Donia J, Shaw JA. Ethics and values in design: a structured review and theoretical critique. Sci Eng Ethics 2021 Aug 19;27(5):57. [doi: 10.1007/s11948-021-00329-2] [Medline: 34410536]

XSL•FO
RenderX

14. McLennan S, Fiske A, Tigard D, Müller R, Haddadin S, Buyx A. Embedded ethics: a proposal for integrating ethics into the development of medical AI. BMC Med Ethics 2022 Jan 26;23(1):6. [doi: 10.1186/s12910-022-00746-3] [Medline: 35081955]
15. Kellmeyer P. Beyond participation: towards a community-led approach to value alignment of AI in medicine. In: Marcello I, Georg S, editors. Developments in Neuroethics and Bioethics, Volume 7: Brains and Machines; Towards a Unified Ethics of AI and Neuroscience: Elsevier; 2024:249-269. [doi: 10.1016/bs.dnb.2024.02.011]
16. Miller FA, Patton SJ, Dobrow M, Berta W. Public involvement in health research systems: a governance framework. Health Res Policy Syst 2018 Aug 6;16(1):79. [doi: 10.1186/s12961-018-0352-7] [Medline: 30081919]
17. Feffer M, Skirpan M, Lipton Z, Heidari H. From preference elicitation to participatory ML: a critical survey & guidelines for future research. Presented at: AIES '23: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society; Aug 8-10, 2023 URL: https://dl.acm.org/doi/proceedings/10.1145/3600211 [accessed 2026-02-26] [doi: 10.1145/3600211.3604661]
18. Corbett E, Denton R, Erete S. Power and public participation in AI. 2023 Oct 30 Presented at: EAAMO '23: Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization; Oct 30 to Nov 1, 2023 URL: https://dl.acm.org/doi/proceedings/10.1145/3617694 [accessed 2026-02-26] [doi: 10.1145/3617694.3623228]
19. Cho H, Oh O, Greene N, et al. Engagement of older adults in the design, implementation, and evaluation of artificial intelligence systems for aging: a scoping review. J Gerontol A Biol Sci Med Sci 2025 Apr 7;80(5):glaf024. [doi: 10.1093/gerona/glaf024] [Medline: 39909831]
20. Ruh DM. Patient perspectives on digital health. In: Patel D, editor. Digital Health: Telemedicine and Beyond: Academic Press; 2025:481-501. [doi: 10.1016/B978-0-443-23901-4.00035-0]
21. Kickbusch I, Piselli D, Agrawal A, et al. The Lancet and Financial Times Commission on governing health futures 2030: growing up in a digital world. Lancet 2021 Nov 6;398(10312):1727-1776. [doi: 10.1016/S0140-6736(21)01824-9] [Medline: 34706260]
22. Kasoju N, Remya NS, Sasi R, et al. Digital health: trends, opportunities and challenges in medical devices, pharma and bio-technology. CSIT 2023 Apr 11;11(1):11-30. [doi: 10.1007/s40012-023-00380-3]
23. Delgado F, Yang S, Madaio M, Yang Q. The participatory turn in AI design: theoretical foundations and the current state of practice. Presented at: EAAMO '23: Proceedings of the 3rd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization; Oct 30 to Nov 1, 2023. [doi: 10.1145/3617694.3623261]
24. Birhane A, Isaac W, Prabhakaran V, et al. Power to the people? Opportunities and challenges for participatory AI. 2022 Oct 6 Presented at: Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '22); Oct 6-9, 2022 URL: https://dl.acm.org/doi/proceedings/10.1145/3551624 [doi: 10.1145/3551624.3555290]
25. Holetzek T, Desch A, Klingler C, van de Bovenkamp H, Holmberg C. The role of representation in participatory settings of health research in Germany: protocol for a scoping review. Res Involv Engagem 2025 Jul 3;11(1):75. [doi: 10.1186/s40900-025-00736-w] [Medline: 40611341]
26. Zamboni K, Schellenberg J, Hanson C, Betran AP, Dumont A. Assessing scalability of an intervention: why, how and who? Health Policy Plan 2019 Sep 1;34(7):544-552. [doi: 10.1093/heapol/czz068] [Medline: 31365066]
27. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
28. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
29. Peters MD, Godfrey C, McInerney P, et al. Scoping reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis: JBI; 2020. URL: https://jbi-global-wiki.refined.site/space/MANUAL/355862497/10.+Scoping+reviews [accessed 2026-03-09] [doi: 10.46658/JBIMES-24-09]
30. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
31. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
32. Greenhalgh T, Hinton L, Finlay T, et al. Frameworks for supporting patient and public involvement in research: systematic review and co-design pilot. Health Expect 2019 Aug;22(4):785-801. [doi: 10.1111/hex.12888] [Medline: 31012259]
33. Sharma R, Kshetri N. Digital healthcare: historical development, applications, and future research directions. Int J Inf Manage 2020 Aug;53:102105. [doi: 10.1016/j.ijinfomgt.2020.102105]
34. Arnstein SR. A ladder of citizen participation. J Am Inst Plann 1969;35(4):216-224. [doi: 10.1080/01944366908977225]
35. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
36. Stilgoe J, Owen R, Macnaghten P. Developing a framework for responsible innovation. Res Policy 2013 Nov;42(9):1568-1580. [doi: 10.1016/j.respol.2013.05.008]
37. Friedman B, Hendry DG. Value Sensitive Design: Shaping Technology with Moral Imagination 2019. [doi: 10.7551/mitpress/7585.001.0001]

XSL•FO
RenderX

Abbreviations

AI: artificial intelligence
JBI: Joanna Briggs Institute
MeSH: Medical Subject Headings
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RQ: research question

Edited by J Sarvestan; submitted 21.Oct.2025; peer-reviewed by J Shaw, J Grosser; revised version received 09.Jan.2026; accepted 22.Jan.2026; published 19.Mar.2026.

Please cite as:
Namuganza S, Gadebusch Bondio M, Kellmeyer P
Public Engagement Strategies in Digital Health Ethics: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e86280
URL: https://www.researchprotocols.org/2026/1/e86280
doi:10.2196/86280

© Stella Namuganza, Mariacarla Gadebusch Bondio, Philipp Kellmeyer. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 19.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical, Histopathological, and Molecular Prognostic Factors Associated With Survival and Disease Progression in Adult Patients With Primary and Secondary Thyroid Lymphomas: Scoping Review Protocol

Ana M Lemos-Rodriguez[1], MD; Juan Pablo Lenis-Gonzalez[1], MD; Oriana Arias-Valderrama[2,3], MD MSc; Elizabeth Arrieta[1,3], MD; Guillermo Edinson Guzman[3,4], MD; Andres Octavio Garcia[3,4], MD

[1]Departamento Medicina Interna, Unidad de Hemato-oncología, Fundación Valle del Lili, Cra 98 No. 18–49, Santiago de Cali, Colombia

[2]Centro de Investigaciones Clínicas, Fundación Valle del Lili, Cali, Valle del Cauca Department, Colombia

[3]Facultad de Ciencias de la Salud, Universidad Icesi, Santiago de Cali, Colombia

[4]Departamento Medicina Interna, Unidad de endocrinología, Fundación Valle del Lili, Cali, Valle del Cauca Department, Colombia

Corresponding Author:
Ana M Lemos-Rodriguez, MD
Departamento Medicina Interna, Unidad de Hemato-oncología, Fundación Valle del Lili, Cra 98 No. 18–49, Santiago de Cali, Colombia

Abstract

Background: To identify the key prognostic factors in patients with primary and secondary thyroid lymphoma, focusing on clinical, histopathological, biological, and genetic components, as these factors remain unclear.The key prognostic factors in patients with primary and secondary thyroid lymphoma remain unclear, particularly regarding clinical, histopathological, biological, and genetic components.

Objective: This study aimed to analyze and synthesize the available evidence to identify factors that may influence the prognosis of patients with primary and secondary thyroid lymphoma.

Methods: This review will include indexed databases such as Scopus, Ovid-MEDLINE, and EMBASE as the primary search sources. The search strategy was applied during the period from June 1, 2024, to September 1, 2024, encompassing all available literature. Eligibility criteria include observational studies of human patients aged 18 years or older, with a diagnosis of primary and/or secondary thyroid lymphoma. We will include reports of adult cases diagnosed with primary thyroid lymphoma or secondary thyroid lymphoma and exclude studies involving patients with primary or secondary tumors of an origin other than thyroid lymphoma. Additionally, studies reporting patients who, alongside primary or secondary thyroid lymphoma, have concomitant diseases associated with high morbidity or mortality, will be excluded. A total of two independent reviewers will participate in the screening, selection, and data extraction process, all of which will be carried out blindly. Conflicts between reviewers will be resolved through discussion or by involving an additional reviewer. Results will be reported using descriptive statistics, and the information will be stored in RedCap (Electronic Research Data Capture).

Results: The literature search was conducted in August 2024, using the following databases: Scopus, MEDLINE, and EMBASE, to identify observational studies. A total of 1404 results were retrieved. The scoping review is expected to be completed by August 2025.

Conclusions: This scoping review will describe the current understanding of the prognostic factors that influence the course of the disease in patients diagnosed with primary and secondary thyroid lymphoma. The review will highlight characteristics that may identify patients at risk of more aggressive disease behavior and could potentially guide therapeutic decisions. Future studies should assess the incidence of these prognostic factors across different age groups and explore their clinical correlations.

International Registered Report Identifier (IRRID): PRR1-10.2196/73164

(JMIR Res Protoc 2026;15:e73164) doi:10.2196/73164

KEYWORDS

prognostic factors; scoping review methodology; thyroid lymphoma; secondary thyroid lymphoma; primary thyroid lymphoma

XSL•FO
RenderX

Introduction

Thyroid lymphoma may be classified as primary, when neoplastic cells arise de novo within the thyroid gland, or secondary, when the gland is infiltrated by malignant lymphoid cells originating from other sites [1]. Both entities are rare. Primary thyroid lymphoma (PTL) accounts for approximately 2% of extranodal lymphomas and 1% - 5% of all thyroid malignancies [2]. Diagnosis is established by fine-needle aspiration or open biopsy [3]. Accurate subtype identification is essential, as treatment depends on the histologic variant. Diffuse large B-cell lymphoma (DLBCL) is the most frequent subtype, comprising more than half of reported cases [4], followed by mucosa-associated lymphoid tissue (MALT) lymphoma [5].

In addition, PTL is strongly associated with Hashimoto thyroiditis; several studies suggest that up to 80% of patients with PTL have a background of this autoimmune condition [6]. Prognosis varies according to the lymphoma subtype, with reported 5-year survival rates of up to 45% [7].

In contrast, secondary thyroid lymphoma is an even rarer entity. It may be misdiagnosed as other thyroid pathologies, such as primary lymphoma or intrathyroidal metastases. Definitive diagnosis relies on histopathological evaluation [8]. Takashima et al [9] reported older series in which thyroid involvement among patients with systemic lymphoma ranged from 11% to 27%. Among these, aggressive subtypes such as Burkitt lymphoma often present with rapid progression and poor prognosis if treatment is not initiated promptly [7].

Although the prevalence of both entities is low, early recognition of primary and secondary thyroid lymphomas is essential, as their therapeutic approaches differ substantially from those of other thyroid malignancies. However, published data on these conditions remain limited.

Given this knowledge gap, establishing a comprehensive database that compiles information reported in the literature would be valuable for assessing potential prognostic factors and determining optimal treatment strategies—an endeavor complicated by the rarity of these diseases. The aim of this scoping review is to examine the current literature on primary and secondary thyroid lymphomas and to identify key factors associated with improved patient outcomes

Methods

The proposed scoping review will be conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) for scoping reviews.

Review Questions

The research questions for this study are as follows:

- What is the current state of knowledge regarding the clinical course of primary and secondary thyroid lymphoma?
- What is the epidemiology of primary and secondary thyroid lymphoma?
- What are the most frequently observed clinical characteristics in patients with primary and secondary thyroid lymphoma?
- What are the most common histopathological characteristics of primary and secondary thyroid lymphoma?
- What are the most common radiological features of primary and secondary thyroid lymphoma?
- How do these characteristics impact mortality in primary and secondary thyroid lymphoma?
- What clinical factors influence disease progression in primary and secondary thyroid lymphoma?
- What are the existing knowledge gaps and future research directions regarding thyroid lymphoma?

Participants

This review will consider studies of any design reporting adult patients [18 years or older) with a confirmed diagnosis of primary or secondary thyroid lymphoma (Table 1).

Table . Inclusion and exclusion criteria for study selection.

Inclusion criteria	Exclusion criteria
Studies reporting adult patients (≥18 years) with a confirmed diagnosis of primary or secondary thyroid lymphoma	Studies describing neoplastic thyroid pathologies other than primary or secondary thyroid lymphoma
Studies including patients with primary or secondary thyroid lymphoma describing frequency, clinical features, and complications	Studies assessing only pediatric populations
Studies providing information on histopathological features	
Studies providing information on treatment strategies	
Studies providing information on factors influencing disease progression and/or mortality in primary or secondary thyroid lymphoma	
Observational and quasi-experimental studies reporting human data	
Studies published in English or Spanish	

Inclusion Criteria

The following studies will be selected for this review:

- Studies including patients with primary or secondary thyroid lymphoma, describing frequency, clinical features, and complications
- Studies providing information about histopathological features

XSL•FO
RenderX

- Studies providing information about treatment strategies
- Studies providing information about characteristics that condition progression on primary and secondary thyroid lymphoma and mortality

Exclusion Criteria

The exclusion criteria were as follows:

- Studies describing neoplastic thyroid pathologies other than primary or secondary thyroid lymphoma
- We will exclude studies assessing only children, but allow inclusion of studies describing mixed populations of children and adults

Concept

The accurate identification of prognostic factors influencing the course of the disease in patients diagnosed with primary or secondary thyroid lymphoma will enable us to determine the characteristics that may indicate a more aggressive disease behavior and even guide therapeutic decision-making. Considering that no studies have systematically compiled this information to date, addressing this gap is essential.

Context

This review will consider studies on primary or secondary thyroid lymphoma, including population characteristics and histological findings, which may influence treatment selection and survival. Additionally, prognostic factors identified in various databases and search engines will be analyzed.

Types of Sources

This scoping review will include all sources of scientific evidence, including quasi-experimental and observational studies. No restrictions will be placed on the type of quantitative research included. The review will be conducted following the JMIR methodology for scoping reviews and aligned with the PRISMA-ScR checklist.

Search Strategy

The search strategy is intended to locate published scientific studies that meet the inclusion criteria. An initial search was performed by the epidemiologist on August 9, 2024, on Scopus (Elsevier), Embase, and Ovid-Medline; this resulted in 1404 article references. A detailed search strategy is provided in Multimedia Appendix 1.

Keywords and Index Terms

All key terms were adapted for each included database and information source. The search terms will be developed collaboratively by the entire team of reviewers, but a single reviewer will conduct the search itself. Additionally, the reference lists of all included sources of evidence will be screened for relevant studies. Studies published in Spanish and English will be included (Textbox 1).

Textbox 1. Search keywords.

Thyroid lymphoma
Primary thyroid lymphoma
Secondary thyroid lymphoma

Study or Source of Evidence Selection

A total of 1404 article references were compiled and uploaded to Rayyan, with duplicates removed by the platform and reviewed manually by a researcher. Independent researchers will screen the titles and abstracts, and potentially relevant articles will be retrieved in full, with their citation details imported into the JBI System for the Unified Management, Assessment, and Review of Information.

Full-text articles selected for inclusion will be evaluated against the inclusion and exclusion criteria by the same independent researchers. Reasons for exclusion will be recorded and reported in the scoping review.

Any disagreements between researchers at each stage of the selection process will be resolved through discussion or by consulting a third researcher. The search results will be fully reported in the final scoping review and presented in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram.

Data Extraction

Data will be extracted from papers included in the scoping review by 2 independent researchers using a spreadsheet to include the following variables [5]:

Title, authors, year of publication, journal, DOI, origin or country of origin (where the source was published or conducted), authors, study type, aims or purpose, population characteristics (age, sex, comorbidities as history of autoimmune thyroiditis); clinical characteristics (Presence of a painless tumor mass in the anterior cervical region, with rapid growth, with or without cervical lymphadenopathy, usually accompanied by symptoms of aerodigestive obstruction); hormonal status (Thyroid profile); imagenologic characteristics; diagnostic method (fine needle aspiration, open biopsy); primary or secondary thyroid lymphoma, if secondary thyroid lymphoma: site of primary lymphoma; classification according to the World Health Organization (WHO) classification of tumors of hematopoietic and lymphoid tissues; immunohistochemical characteristics; presences of other malignancies; surgery (yes or no); treatment (type of chemotherapy, radiotherapy); and overall survival.

The draft spreadsheet will be modified and revised as needed throughout the data extraction process for each included paper. Any modifications will be documented in the final scoping review. Disagreements between researchers will be resolved through discussion or by consulting a third researcher. If necessary, the authors of the included papers will be contacted to request missing or additional data.

XSL•FO
RenderX

Data Analysis and Presentation

A narrative description of the findings will be discussed in the framework of the review questions. The data will be summarized and presented in tables.

Reference Searches

Citation tracking will be used to identify key articles relevant to the topic of interest. This process will involve reviewing the reference lists of relevant papers, as well as identifying articles that have cited them. Any duplicate papers that have already been examined will be removed. If a paper meets the inclusion criteria, its reference list will be screened to identify additional relevant manuscripts.

Results

The search was conducted on August 9, 2024, using Scopus (Elsevier), Embase, and Ovid, yielding a total of 1404 results. A comprehensive review of the manuscripts is expected to be completed by August 2025. All potential manuscripts will be imported into the reference management software Rayyan QCRI 23. This scoping review aims to identify and describe the clinical, histopathological, and molecular prognostic factors associated with survival and disease progression in adult patients with primary and secondary thyroid lymphomas.

A preliminary PRISMA flow diagram illustrating the identification, screening, eligibility, and inclusion process is provided in Figure 1.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) flowchart.

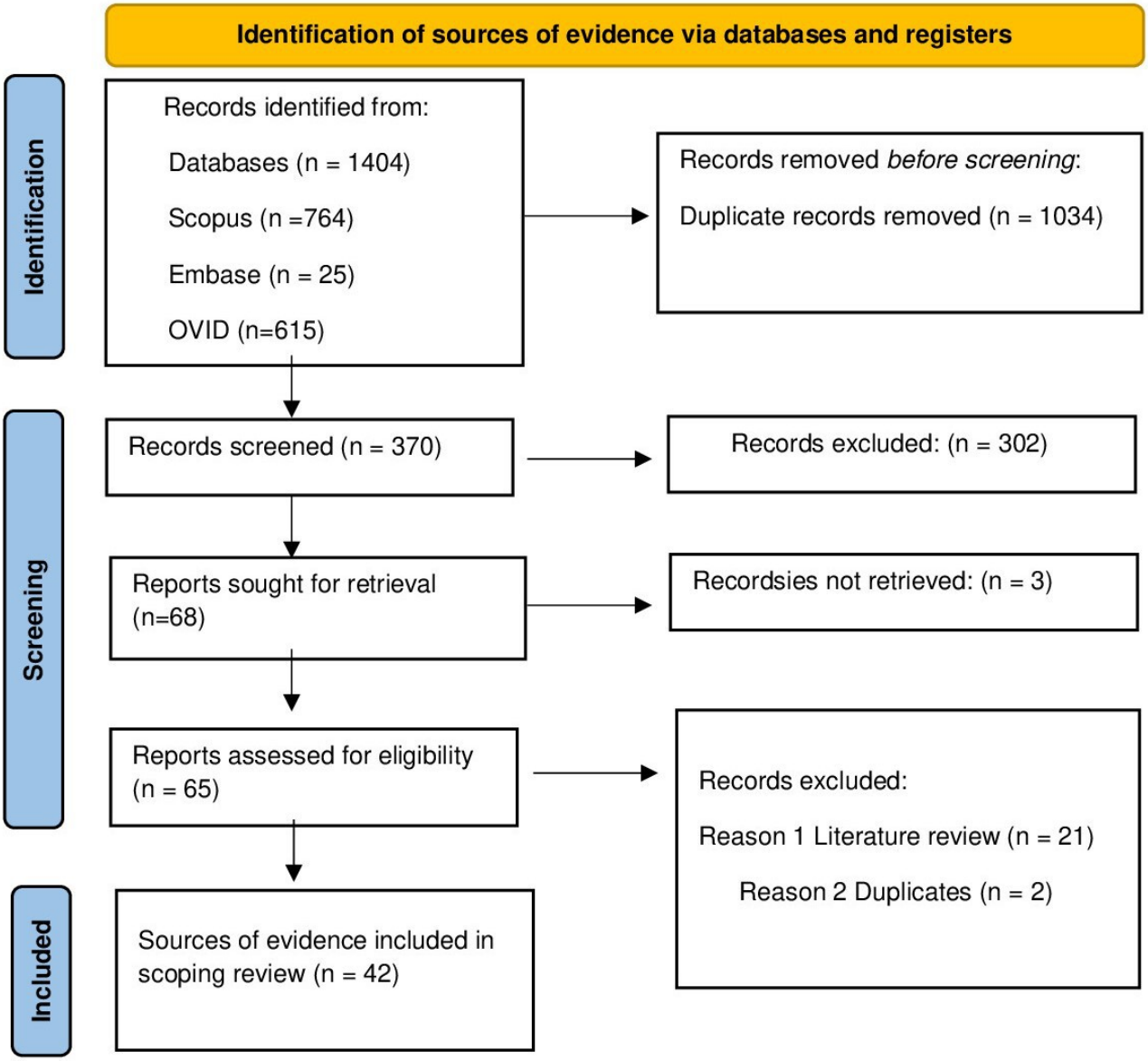

Discussion

Primary and secondary thyroid lymphomas are rare malignant neoplasms that account for only 2% - 5% of all thyroid cancers and approximately 2% of extranodal lymphomas [10,11]. This rarity underscores the need to analyze the various case reports and case series published over time to identify specific characteristics such as histologic subtypes, clinical manifestations, prognosis, and survival, among others [11].

XSL•FO
RenderX

Given their low incidence, the overarching goal is to identify shared features across different patient groups.

With this scoping review, we aim to compile the available observational evidence to determine which clinical factors—such as sex, age, presentation, comorbidities, and disease stage—have been associated with progression and/or survival in patients with thyroid lymphomas [12]. We also aim to identify histopathological characteristics, including histologic subtypes and immunohistochemical markers such as tumor protein p53 (p53) and marker of proliferation Ki-67 (Ki-67), that have been linked to prognosis. This will help confirm that the most frequent histologic subtypes of thyroid lymphoma are DLBCL and MALT, as described in multiple case series [6].

Additionally, we seek to characterize the most relevant and common clinical manifestations—such as dysphagia, hoarseness, and airway obstruction, particularly in rapidly growing masses [13]—to support earlier diagnosis when clinical suspicion arises. However, molecular evidence remains fragmented and heterogeneous, and prognostic factors for secondary thyroid lymphomas are even less consistent and infrequently reported.

Among the main findings, the most frequently evaluated variables across studies include age, the presence of compressive symptoms, a prior diagnosis or documented history of Hashimoto thyroiditis [14], and the different Ann Arbor stages. Certain factors appear to have a more consistent association with survival outcomes[15], including age, tumor size, and the presence of B symptoms [16]. Nonetheless, only a limited number of studies provide comprehensive data on comorbidities at diagnosis [14,17], diagnostic methods [7,17,18], clinical status, and other relevant variables [19].

Regarding the prognostic value of immunomarkers, Ki-67 is perhaps the most extensively studied and most consistently associated with disease aggressiveness. Other markers such as p53, B-cell lymphoma 2 (BCL2), and MYC proto-oncogene (MYC) show variable findings [20], and no adequately standardized protocol exists for their assessment. Molecular factors, including gene rearrangements, are infrequently reported in patients with thyroid lymphoma, making it even more challenging to evaluate somatic mutations or advanced genetic alterations in these tumors [21]. These findings often lack external validation [22], further limiting their incorporation into broader analyses.

This scoping review seeks to highlight the lack of standardization in the characterization of primary and secondary thyroid lymphomas. Examples include the frequent aggregation of primary and secondary cases in previous studies, differences in therapeutic management, and considerable variability in the cut-off values used for Ki-67 [23,24].

Among the strengths of this review, we identify the opportunity to expand the clinical, histologic, and molecular characterization of thyroid lymphomas by including studies with diverse designs and by identifying gaps that had not been clearly documented or systematically addressed in previous literature. Nevertheless, heterogeneity in study designs, sample sizes, and disease stages limits the ability to establish causal relationships. Most of the available evidence is derived from small retrospective series; however, these still contribute valuable insights and facilitate a more structured understanding of the current literature. One of the most significant limitations remains the scarcity of studies focusing specifically on secondary thyroid lymphoma.

These findings underscore the need for prospective multicenter studies that jointly evaluate clinical and molecular factors, standardize immunohistochemical markers and their quantification, and, in the long term, conduct comparative analyses to better establish prognostic differences between primary and secondary thyroid lymphomas.

We hope that disseminating this compiled information and publishing our results in open-access repositories and medical literature will support clinicians and pathologists seeking a deeper prognostic understanding of patients affected by this condition.

This scoping review, therefore, aims to emphasize the importance of clinical and histopathological factors as the predictors with the strongest available evidence in thyroid lymphomas, whereas molecular factors remain fragmented due to their low reporting frequency across studies, making them appear insufficiently explored. Although, as previously mentioned, methodological variability limits the ability to draw solid comparisons, this work constitutes an initial step toward identifying priority areas for standardization and knowledge gaps, providing a clearer path for future research and ultimately improving the prognostic characterization of this rare entity.

Acknowledgments

The authors declare the use of generative AI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision:

- Reformatting

The GAI tool used was: Chat GPT-4,5.

Responsibility for the final manuscript lies entirely with the authors.

GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

XSL•FO
RenderX

Funding

This project was funded by Centro de Investigaciones Clínicas at Fundación Valle del Lili.

Data Availability

The data underlying this scoping review will be made available upon reasonable request to the corresponding author.

Authors' Contributions

JPL and AMLR led the conceptualization of the study. OAV developed the methodology. OAV performed the formal analysis. JPL, AMLR, EA, and OAV conducted the investigation and data acquisition. OAV was responsible for data curation. JPL and AMLR drafted the original manuscript. GEG, AOG, and EA contributed to the critical review and editing of the manuscript. OAV prepared the visualizations. GEG and AOG supervised the project. All authors reviewed and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy thyroid lymphoma.
[DOCX File, 25 KB - resprot_v15i1e73164_app1.docx]

References

1. Surya S, Sathanantham DK, P N S, Doraiswamy S. Secondary thyroid lymphoma due to chronic lymphocytic leukaemia (CLL). Int J Cancer Manag 2023;16(1):e136406. [doi: 10.5812/ijcm-136406]
2. Du Z. Clinicopathologic characteristics, gene mutation profile and prognostic analysis of thyroid diffuse large B-cell lymphoma. Shanghai Jiao Tong Da Xue Xue Ba 2024;44(1):64-71. [doi: 10.3969/j.issn.1674-8115.2024.01.007]
3. Wu SY, Chu CH, Duh QY, Hsieh CB, Yu JC, Shih ML. Management for primary thyroid lymphoma: experience from a single tertiary care centre in Taiwan. Formosan Journal of Surgery 2016 Dec;49(6):201-207. [doi: 10.1016/j.fjs.2016.07.001]
4. Rezkallah E, Hanna RS, Elsaify WM. Thyroid lymphoma: case series review. Am Surg 2023 Nov;89(11):4811-4816. [doi: 10.1177/00031348221111512] [Medline: 35759750]
5. Akcali Z, Sakalli H, Noyan T, Kayaselcuk F, Ozyilkan O. Primary thyroid lymphoma: report of two cases. East Afr Med J 2004 Jul;81(7):378-380. [doi: 10.4314/eamj.v81i7.9195] [Medline: 15490712]
6. Sakhri S, Zemni I, Ayadi MA, Kamoun S, Chargui R, Ben Dhiab T. Primary thyroid lymphoma: a case series. J Med Case Rep 2024 Feb 24;18(1):108. [doi: 10.1186/s13256-024-04434-1] [Medline: 38395921]
7. Vander Poorten V, Goedseels N, Triantafyllou A, et al. Effectiveness of core needle biopsy in the diagnosis of thyroid lymphoma and anaplastic thyroid carcinoma: a systematic review and meta-analysis. Front Endocrinol 2022;13:971249. [doi: 10.3389/fendo.2022.971249]
8. Vallejo Herrera MJ, Vallejo Herrera V. Linfoma secundario de tiroides: una patología engañosa [Secondary thyroid lymphoma: a misleading pathology]. Rev Esp Geriatr Geronto 2025;60:101628. [doi: 10.1016/j.regg.2025.101628]
9. Takashima S, Takayama F, Momose M, Shingu K, Sone S. Secondary malignant lymphoma which simulated primary thyroid cancer. Clin Imaging 2000;24(3):162-165. [doi: 10.1016/s0899-7071(00)00195-9] [Medline: 11150685]
10. Jaafari A, Mehaudens S, Gheysens O, et al. Massive primary thyroid lymphoma on [18F]-FDG-PET/CT: a literature review of a rare case of rapidly progressive goitre. Diagnostics (Basel) 2025 Aug 28;15(17):162. [doi: 10.3390/diagnostics15172180] [Medline: 40941668]
11. Sharma A, Jasim S, Reading CC, et al. Clinical presentation and diagnostic challenges of thyroid lymphoma: a cohort study. Thyroid 2016 Aug;26(8):1061-1067. [doi: 10.1089/thy.2016.0095] [Medline: 27256107]
12. Alalawi Y, Alrashidi TN, Shafie S, Moharram L, Asiri A. Primary thyroid lymphoma: a case series. Cureus 2025 Aug;17(8):e89197. [doi: 10.7759/cureus.89197] [Medline: 40901213]
13. Toyoshima MTK, de Macedo C, de Magalhães IPA, Velasques RD, Hoff AO. Presentation and management of thyroid involvement in systemic and primary thyroid lymphoma: a case series. AIM Clinical Cases 2025 Jun 1;4(6). [doi: 10.7326/aimcc.2024.1289]
14. Hamdy O, Saleh GA, Ayman Selim R, et al. Diagnosis and treatment outcomes of primary lymphoma of the thyroid gland. Discov Onc 2025;16(1):1568. [doi: 10.1007/s12672-025-03218-3]
15. Zhang K, Peng X, Wei T, Li Z, Zhu J, Chen YW. Prognostic nomogram and competing risk analysis of death for primary thyroid lymphoma: a long-term survival study of 1638 patients. Ann Surg Open 2022 Dec;3(4):e226. [doi: 10.1097/AS9.0000000000000226] [Medline: 36590887]
16. Vita O, Dema A, Barna R, et al. Primary thyroid lymphoma: a retrospective-observational study in a single institutional center. Medicina (Kaunas) 2024 Mar 14;60(3):476. [doi: 10.3390/medicina60030476] [Medline: 38541203]

XSL•FO
RenderX

17. Hirokawa M, Suzuki A, Hashimoto Y, et al. Prevalence and diagnostic challenges of thyroid lymphoma: a multi-institutional study in non-Western countries. Endocr J 2020 Nov 28;67(11):1085-1091. [doi: 10.1507/endocrj.EJ20-0202] [Medline: 32641619]
18. Su MW, Beck TN, Knepprath J, Romero-Velez G, Heiden KB, McHenry CR. Primary thyroid lymphoma: a multi-center retrospective review. Am J Surg 2024 Nov;237:115927. [doi: 10.1016/j.amjsurg.2024.115927] [Medline: 39213784]
19. Xiao L, Zhao Z, Zhou L, et al. Clinical and imaging features of primary thyroid MALT lymphoma. Front Oncol 2025;15:1498609. [doi: 10.3389/fonc.2025.1498609] [Medline: 40406269]
20. Tomas-Roca L, Rodriguez M, Alonso-Alonso R, Rodriguez-Pinilla SM, Piris MA. Diffuse large B-cell lymphoma: recognition of markers for targeted therapy. Hemato 2021;2(2):281-304. [doi: 10.3390/hemato2020017]
21. Calafato G, Di Paola FJ, De Leo A, et al. Somatic genetic alterations in the development and progression in thyroid tumors of follicular cells. Eur Thyroid J 2025;14(5). [doi: 10.1530/ETJ-25-0104]
22. Zhu Y, Yang S, He X. Prognostic evaluation models for primary thyroid lymphoma, based on the SEER database and an external validation cohort. J Endocrinol Invest 2022 Apr;45(4):815-824. [doi: 10.1007/s40618-021-01712-3] [Medline: 34865184]
23. Guidobaldi L, Cafiero C, D'Amato G, et al. Primary thyroid lymphoma: how molecular biology and ancillary techniques can help the cytopathologist overcome this diagnostic challenge. J Pers Med 2023 Jul 28;13(8):1203. [doi: 10.3390/jpm13081203] [Medline: 37623454]
24. Harlendea NJ, Harlendo K. Ki-67 as a marker to differentiate Burkitt lymphoma and diffuse large B-cell lymphoma: a literature review. Cureus 2024 Oct;16(10):e72190. [doi: 10.7759/cureus.72190] [Medline: 39583511]

Abbreviations

DLBCL: diffuse large B-cell lymphoma
MALT: mucosa-associated lymphoid tissue
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
PTL: primary thyroid lymphoma

Edited by J Sarvestan; submitted 26.Feb.2025; peer-reviewed by JL Pardal-Refoyo, V Fiorentino; revised version received 29.Dec.2025; accepted 30.Dec.2025; published 20.Feb.2026.

Please cite as:
Lemos-Rodriguez AM, Lenis-Gonzalez JP, Arias-Valderrama O, Arrieta E, Guzman GE, Garcia AO
Clinical, Histopathological, and Molecular Prognostic Factors Associated With Survival and Disease Progression in Adult Patients With Primary and Secondary Thyroid Lymphomas: Scoping Review Protocol
JMIR Res Protoc 2026;15:e73164
URL: https://www.researchprotocols.org/2026/1/e73164
doi:10.2196/73164

© Ana M Lemos-Rodriguez, Juan Pablo Lenis-Gonzalez, Oriana Arias-Valderrama, Elizabeth Arrieta, Guillermo Edinson Guzman, Andres Octavio Garcia. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Affective Health Messages for Adolescents: Protocol for a Scoping Review

Siv Hilde Berg[1], PhD; Raul Ferrer-Conill[2], PhD; Kristine Stifjell[3], MA; Kristin Hjorthaug Urstad[4], PhD; Jo Røislien[1], PhD; Lisbet Fjæran[5], PhD

[1]Department of Quality and Health Technology, Faculty of Health Sciences, University of Stavanger, Innovasjonsparken i Stavanger, Bygg i3, Stavanger, Norway

[2]Department of Media and Social Sciences, Faculty of Social Sciences, University of Stavanger, Stavanger, Norway

[3]Department of Community Medicine, UiT The Arctic University of Norway, Tromsø, Norway

[4]Department of Public Health, Faculty of Health Sciences, University of Stavanger, Stavanger, Norway

[5]Department of Safety, Economics and Planning, Faculty of Science and Technology, University of Stavanger, Stavanger, Norway

Corresponding Author:
Siv Hilde Berg, PhD
Department of Quality and Health Technology, Faculty of Health Sciences, University of Stavanger, Innovasjonsparken i Stavanger, Bygg i3, Stavanger, Norway

Abstract

Background: Affect plays a pivotal role in capturing people's attention with health messages that influence audiences' perceptions, attitudes, and, ultimately, health-related behavior. Adolescents are notably sensitive to affective cues in communication. However, there is a lack of reviews synthesizing the evidence examining the role of affect in health messaging targeted at adolescents.

Objective: The aim of this scoping review is to systematically map and synthesize the empirical studies examining the role of affective health message content targeted at adolescents and the associated communication outcomes.

Methods: We will conduct our review in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist. We will include empirical studies that examine health messages directed at adolescents from formal or informal sources, as well as studies using research-based experimental designs. Only studies focusing on public communication through mediated communication (eg, social media, the internet, and television) will be included. To be eligible, studies must explore message content targeted at adolescents' emotions. We will include studies reporting on communication outcomes in the context of urgent or nonurgent health issues. Only English-language, peer-reviewed empirical studies published in 2014 or later will be included. Data from eligible studies will be extracted into a matrix, summarized in tabular form, and synthesized using a descriptive qualitative analysis.

Results: As of May 2025, we have searched the Embase, Ovid MEDLINE, and PsycInfo databases and screened 29 reviews, which has resulted in 1856 unique hits. Scopus will be added to the search. Screening and full-text assessment will be conducted in 2026.

Conclusions: We anticipate that this scoping review will provide insights into how adolescents behave, form judgments, and make decisions in response to health messages targeted at their emotions. Additionally, the review will conclude by outlining the research gaps and offering recommendations for future studies.

International Registered Report Identifier (IRRID): DERR1-10.2196/78927

(JMIR Res Protoc 2026;15:e78927) doi:10.2196/78927

KEYWORDS

health communication; adolescence; affect; decision-making; health behavior; framing; narratives; emotional appeals

Introduction

The affective-persuasive turn in digital communication is dominated by rapid, emotionally charged messaging across multiple digital platforms [1]. Affect plays an imperative role in designing content that captures people's attention; elicits people's emotions; fosters engagement; and, ultimately, persuades people to change their behavior [1,2]. Persuasion theory proposes that the content of a message shapes its emotional tone, which influences how deeply the audience processes the information [3]. Within this framework, message content refers to what the message communicates and includes, among other things, emotional appeals, framing, narrative elements, and the way in which the arguments are presented [3].

XSL•FO
RenderX

However, the literature lacks a comprehensive overview of how messages targeted at adolescents affect influence health decision-making and behavior within the field of health communication. The role of affect in judgment, decision-making, and behavior [4] is not sufficiently explored in research on public health communication [5]. The gap between theory and empirical application highlights the need for research that connects conceptual models, empirical studies, and real-world health communication practices.

Adolescents are notably receptive to emotionally charged health messages [5]. Adolescence represents a key developmental period marked by emotional, cognitive, and social changes. During this period, extensive remodeling of the brain's dopaminergic system occurs, leading to increased risk-taking and reward-seeking behaviors, particularly in the presence of peers [6]. These behaviors are largely driven by changes in the brain's socioemotional system, which matures earlier than the cognitive control system responsible for self-regulation [6]. This heightened emotional sensitivity is further amplified by the pervasive influence of social media and digital communication in adolescents' daily lives. Digital interactions on social media can intensify affective polarization [7] and significantly impact adolescents' emotional well-being—both positively and negatively [8,9]. Moreover, existing studies rarely take account of the rapidly changing digital environments in which adolescents operate, raising concerns about the relevance of findings predating the social media era.

Increased sensitivity to emotional cues in communication, combined with massive exposure to digital information and the fact that the contingencies for health decision-making are different for adolescents than for their older counterparts and the population at large, highlights the need for better insight into the role of affect in health communication targeted at adolescents. Despite its importance, there is a lack of reviews synthesizing evidence on the role of affect in health message content targeted at adolescents. Existing systematic reviews have primarily focused on messages on specific health issues targeted at adolescents and young adults, such as messaging strategies for preventing and ceasing e-cigarette use [10], mental health awareness campaigns [11], vaccine coverage and timeliness [12], and the promotion of healthy lifestyle behaviors [13]. Systematic reviews have also examined the emotional impact of social media and digital technology [9,14]. Thus, much of the research reviewed is issue specific and lacks integration across health domains.

A generic and broader literature synthesis examining systematically how affective message content—originating from both formal (eg, public health campaigns) and informal (eg, social media) sources—is linked to communication outcomes remains lacking. Such evidence includes a wide variation in methodological designs and outcomes and crosses boundaries among public health, health communication, emotion science, and health decision-making, among others. The lack of a common framework limits comparability across studies and makes it difficult to conduct meta-analyses and assess which strategies are most effective in different contexts. However, the scoping review methodology is particularly useful in this matter to synthesize evidence in a systematic manner and identify knowledge gaps and the scope of a body of literature with different methodologies [15].

This scoping review aims to fill these current knowledge gaps related to lack of literature reviews on the use and outcomes of affective message content in health communication targeted at adolescents by systematically mapping empirical studies. This review is guided by the following research questions: what empirical literature exists regarding health message content targeted at adolescent emotions, and what are the associated communication outcomes?

Methods

This study is a scoping review, supporting a systematic approach to data identification, extraction, and synthesis [16]. We will conduct our review following guidance for systematic scoping reviews in health care [16] and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist (Checklist 1) [17].

Eligibility Criteria

Population, Concept, and Context Framework

The population, concept, and context approach was used to specify our rationale and eligibility criteria [17].

Population

We will include studies in which most of the sample are adolescents (aged 10-24 years) and exclude studies that primarily sample other age groups [18].

Concept

Overview

The conceptual criteria are organized according to the revised integrated change model [19] (Figure 1). The revised integrated change model outlines information factors (personal factors, message factors, channel factors, and source factors), communication outcomes (emotions, awareness, motivation, and action), and preceding factors. The model assumes that the communication outcomes depend on the 2 determinants: information factors and preceding factors.

XSL•FO
RenderX

Figure 1. Revised integrated change model.

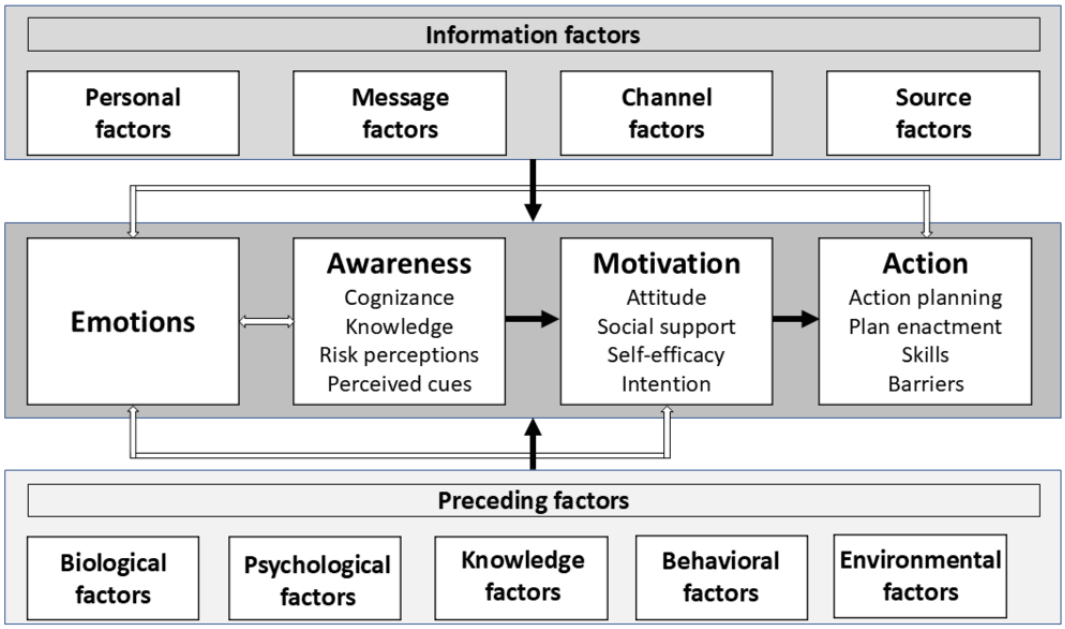

Source Factors

We will include studies of public health messages from formal sources (eg, public health authorities or the World Health Organization or researcher-designed messages in experimental studies), informal sources (eg, influencers, advertisements, and social media), or both. We will exclude studies of personal communication (eg, communication between patients and health care professionals) and interventions that include multiple components beyond the public health messages (eg, group-based patient education interventions).

Channel Factors

We will include studies on diverse modes of communication, which include but are not limited to web-based information, social media, television, newspapers, video, and texts.

Message Factors

In this review, we will include studies that involve affective message content. Affect is defined by Lerner et al [4] as "unspecified feelings; the superordinate umbrella of constructs involving emotion, mood, and emotion-related traits." A precise definition concerning affect in communication does not exist [20]. We understand affective message content in this context as message-level features (message factors) that have been shown to influence the recipients' feelings and emotions as opposed to and/or next to influencing the recipients' thoughts and beliefs [20]. Affective features of message content include emotional appeals (eg, fear, humor, and empathy), framing (eg, gain vs loss framing), narrative elements (eg, storytelling and personal experiences), and the way in which arguments are presented persuasively (eg, emphasizing severe health risks and using moral appeals) [3].

Communication Outcomes

We will include studies reporting on communication outcomes in a broad manner related to emotions (eg, trust and specific emotions), awareness (eg, risk, perception, knowledge, awareness, bias, and heuristics), motivation (eg, attitudes and self-efficacy), and action (eg, behavioral intentions, susceptibility to behave, and self-reported behavior and actions). We will exclude outcomes related to predisposing factors, that is, behavioral factors (eg, lifestyles), psychological factors (eg, personality and message memory), knowledge factors (eg, health literacy and knowledge), biological factors (eg, gender), social and cultural factors (eg, the price of cigarettes and policies), and information factors (the quality of messages and the channels and sources used) [19].

Context

We will include communication outcomes related to both urgent (eg, choosing a cancer treatment) and nonurgent (eg, deciding to quit smoking) health issues or risks. We will include studies in the health context concerning adolescents, which includes communication of health risks or issues and medical decision-making regardless of the geographical location and culture of the receiver. We will exclude judgment and decision-making in nonhealth contexts or regarding nonhealth risks and issues.

Types of Sources of Evidence

English-language peer-reviewed empirical studies will be included irrespective of data collection method. Given the rapidly evolving impact of the digital society and social media in the last decade, we will include only literature published during or after 2014. Commentaries, reviews, opinion pieces, or other papers not reporting primary empirical research will be excluded.

XSL•FO
RenderX

Search Strategy and Information Sources

To identify potentially relevant documents, the following bibliographic databases will be searched: Embase (1974-April 2025), Ovid MEDLINE (1946-April 11, 2025), and PsycInfo (2002-April 2025). We will not undertake reference screening or contact paper authors. The search strategy was drafted by an experienced librarian and SHB, refined, and, finally, validated by the librarian. We combined search terms identified in the literature with MeSH (Medical Subject Headings) terms and proximity operators. The full electronic search strategy is presented in Multimedia Appendix 1.

We will search using the following terms: “dual process theor*,” “dual process model,” “risk as feeling,” “heuristic cues,” “heuristics,” “affect heuristic*,” “appraisal theor*,” “cognitive appraisal*,” “cognitive-experiential self-theory,” “somatic marker,” “self-efficacy,” “intuitive-experiential,” “analytical-rational thinking,” “system 1 thinking,” “system-1 thinking,” “information processing theor*,” “framing*,” “frame,” “narrative,” “empath*,” “message feature*,” “valence,” “message appeal*,” “negative appeal*,” “negative emotion*,” “positive appeal*,” “positive emotion*,” “value approach,” “presentation mode or communication mode*,” “emotions,” “affect,” “fear,” “anger,” “hope,” “shame,” “self-efficacy,” “digital communication,” “health message*,” “social media,” “health communication,” “health promotion,” “health education,” “consumer health information,” “behavior,” “health behavior,” “decision making,” “judgment,” “judgement,” “behaviour*,” “teen*,” “young people,” “young person*,” “youth or adolescen*,” “high school*,” “health*,” and “medical*.”

No filters or limits will be added to the initial literature searches regarding language and peer review. Filters will be added for publication date range (January 2014-present). No searches of gray literature will be conducted due to feasibility constraints and the need for quality assurance as gray literature often lacks peer review. The final search results will be exported into EndNote (Clarivate Analytics), and duplicates will be removed by SHB. All records will thereby be uploaded to Rayyan (software for intelligent systematic review; Qatar Computing Research Institute) for further manual removal of duplicates. In addition, manual searches will be conducted of reviews, which will be screened by 2 reviewers for primary studies fulfilling the eligibility criteria. All identified records will be logged and reported in the PRISMA-ScR flow diagram. The list of screened reviews will be provided as an appendix.

Selection of Sources of Evidence

We will create 3 screening teams (SHB and LF, KHU and JR, and KS and RF-C) to screen titles and abstracts from the databases using the eligibility criteria outlined above and produce a list of articles to be assessed in full text. We will use Rayyan to support the screening and inclusion process of eligible studies. Within each team, members will be blinded to each other’s decisions during the screening process to minimize bias [16]. Dual independent screening will be used at both the title and abstract and full-text stages, with discrepancies resolved through discussion or a third party. The eligibility criteria will be pilot-tested by independently screening 20 articles, and the criteria will be subsequently refined based on feedback from the authors. The authors will divide the final list of articles between them and screen the articles in full text. The final list of included papers will be validated by SHB. Any uncertainties regarding inclusion or exclusion will be discussed within each team, and disagreements will be resolved through discussion with the first author (SHB).

Data Charting Process

Following the guidance for conducting scoping reviews [16], all authors will extract the data from the included articles into a matrix ahead of synthesis. The extracted data will include (1) author names, (2) year of publication, (3) country of origin, (4) aims and research questions, (5) study design, (6) sample characteristics, (7) type of affective feature in the communication content, (8) health communication topic, and (9) key outcomes related to the research objectives. The matrix will be tested by extracting data from 1 paper per reviewer, which will then be validated in collaboration. The extracted data will be condensed and displayed in tables, and the content will be validated by SHB.

Synthesis Method

A narrative synthesis will be applied to the results—an approach to synthesis of findings using words and text to summarize and explain the results [21]. In the context of a scoping review, narrative synthesis addresses a broad range of questions not limited to the effectiveness of specific interventions [21].

Data synthesis will follow the 4-stage narrative synthesis approach proposed by Whittemore and Knafl [22]. The first stage of analysis consists of data reduction, whereby the results relevant to the review question will be condensed and categorized according to 3 predetermined themes and categories (Table 1).

XSL•FO
RenderX

Table . Predetermined themes and categories.

Theme	Predetermined categories
Study characteristics	• Year of publication • Country of origin • Aims and research questions • Study design • Sample characteristics • Health communication topics
Type of affective feature in the communication content	• Use of emotional appeals, framing, and/or narratives targeted at adolescents' emotions • Use of persuasive arguments targeted at adolescents' emotions
Communication outcomes	• Emotions • Awareness • Motivation • Action

Within each category, data-driven inductive coding will ensure that message content targeted at adolescent emotions is integrated into the narrative synthesis. The predetermined categories will be used to preorganize the analysis, while allowing for refinement of categories and themes as new insights emerge during analysis. In the second stage, tables will be created to connect the patterns and relationships among categories. The third stage of analysis, *data comparison*, involves synthesizing the similarities, inconsistencies, and variables in each category. In the fourth stage, the identified themes will be verified by cross-referencing the primary source data.

The analysis will be conducted by SHB and validated by all authors. Disagreements will be resolved through collaborative discussion among the authors. We will not conduct a critical appraisal of individual sources of evidence or assess the risk of bias across studies.

The patterns in the data will be interpreted and contextualized, informed by persuasion theory. Persuasion theory provides a conceptual lens for understanding how message content influences emotional and cognitive processing. It emphasizes message factors (emotional appeals, framing, and narratives) as determinants of persuasion and the interaction between affective responses and information processing depth (eg, heuristic vs systematic processing) [3].

Results

As of May 2025, the search has yielded a total of 2218 hits, of which 784 (35.3%) were from Embase, 942 (42.5%) were from MEDLINE, and 492 (22.2%) were from PsycInfo. In addition, 29 literature reviews were screened for studies sampling adolescents or young people, which resulted in 111 hits. Removing duplicates resulted in 1856 unique hits, as displayed in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram (Figure 2 [23]). An updated search will be conducted in February 2026 to ensure inclusion of recent publications, and Scopus will be added to the database list to strengthen interdisciplinary coverage.

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [23].

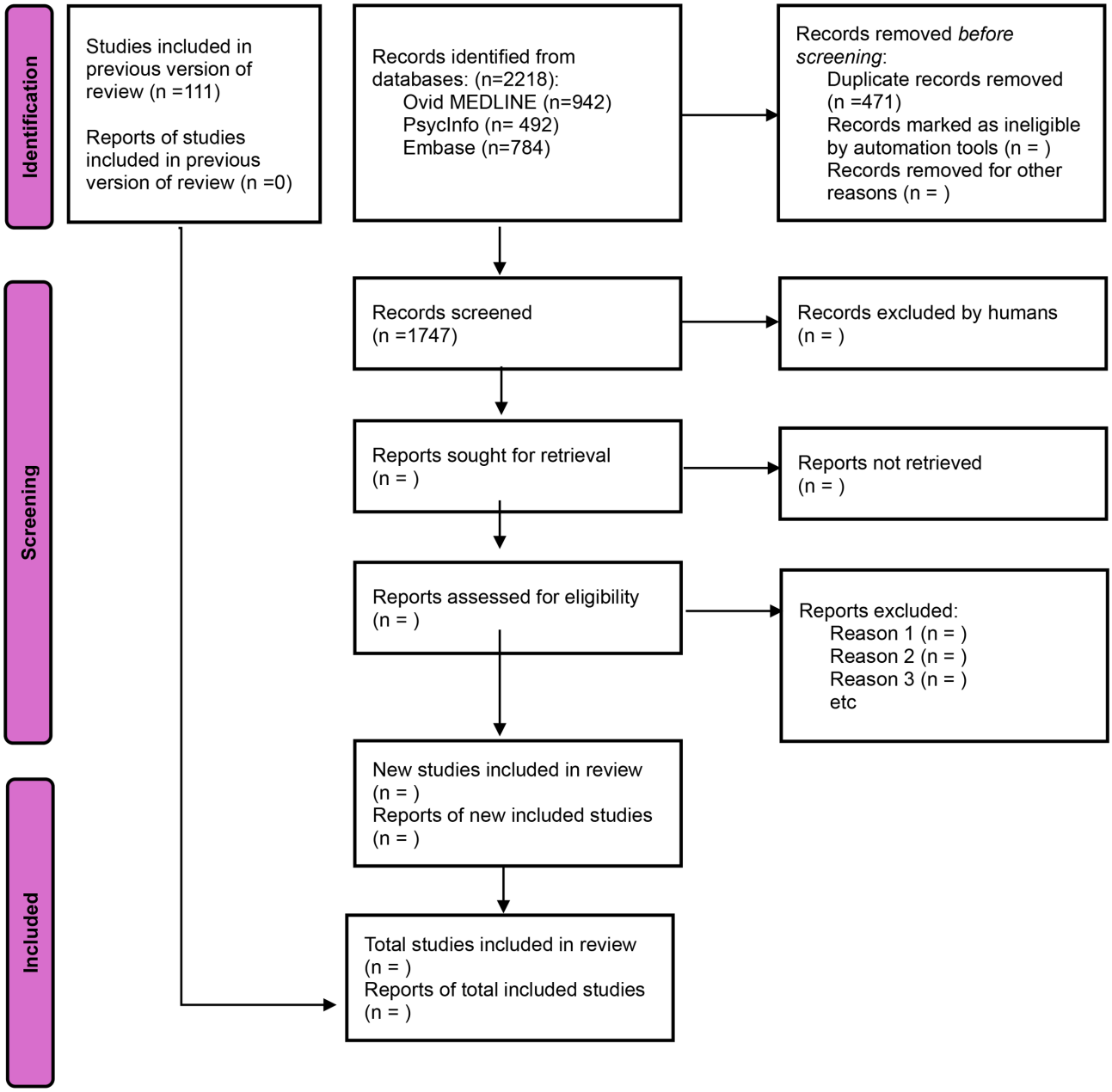

Discussion

This review presents a synthesis of the evidence on the role of affect in health messaging toward adolescents. Through applying systematic scoping review methodology, this study will qualitatively synthesize the outcomes of the identified empirical studies using health message content targeted at adolescents' emotions, the associated communication outcomes, and identified research gaps.

Although this review applies a descriptive synthesis, interpretation of findings will be informed by persuasion theory, which emphasizes how message factors such as emotional appeals, framing, and narratives shape both affective responses and cognitive processing depth [3]. Persuasion theory suggests that affective cues can activate heuristic or peripheral processing routes, influencing attitudes and behavioral intentions even when systematic evaluation is limited. We aim to contextualize how affective message content may lead to positive or negative outcomes among adolescents and discuss the ethical implications of using persuasive messages with this age group, whose heightened emotional sensitivity and developmental characteristics make them particularly responsive to affect-driven communication strategies.

This scoping review will provide insights into how adolescents behave, form judgments, and make decisions in response to affective health message content and highlight the communication outcomes of these studies. We anticipate finding a knowledge gap pertaining to real behavioral outcomes (actions) and that most of the literature measures intention to behave or awareness as the communication outcome.

We will discuss the strengths and limitations of the methods used in the included studies and draw attention to identified ethical challenges of communication targeted at adolescents. The scoping review findings will be discussed in comparison

XSL•FO
RenderX

with previous literature on emotion and decision-making [4], health decision-making [5], health communication targeted at adolescents [10-13], and digital communication [9,14], highlighting the novel insights from this review. To conclude, we will outline policy recommendations regarding considerations that are important when designing health message content targeted at adolescents. Additionally, the review will conclude by outlining the research gaps in, for example, emotion science and communication studies and offering recommendations for future studies and campaigns in selecting theory and communication strategies.

The strength of this scoping review is the systematic application of rigorous research methodology, which strives for transparency and methodological rigor in all phases of the review process. A limitation of this scoping review is the exclusion of gray literature in the search strategy, which may bias findings toward academic and Western contexts in which peer-reviewed publications are more common. A major limitation of this review is the unclear conceptual validity of affect in message content. We address this limitation by using predefined operationalized terms, which may provide transparency regarding the study selection strategy applied and strengthen the theoretical generalization of our findings to other contexts.

Acknowledgments

The authors would like to thank Professor Enny Das at Radboud University for her valuable feedback on the protocol.

Funding

This project has been funded by and supported by the University of Stavanger through a collaborative initiative within the Public Health Communication Research Group.

Authors' Contributions

SHB designed the study and drafted the protocol. All authors contributed to refining the scope of the review, the methodology, and the eligibility criteria and approved the final version of the manuscript. SHB designed the search strategy and conducted the initial search. All authors will contribute to screening and extraction of data. SHB will lead the synthesis of findings, which will be validated by all authors. KS provided subject matter expertise on adolescents' perspectives, with particular focus on digital media and the health communication context. JR and RF-C provided subject matter expertise on digital communication and communication modalities. LF provided subject matter expertise on emotions and risk communication. KHU provided expertise on scoping review methodology and reviewed the search strategy.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 16 KB - resprot_v15i1e78927_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 111 KB - resprot_v15i1e78927_app2.docx]

References

1. Gregg M, Seigworth GJ. The Affect Theory Reader: Duke University Press; 2020.
2. Steinert S, Dennis MJ. Emotions and digital well-being: on social media's emotional affordances. Philos Technol 2022;35:36. [doi: 10.1007/s13347-022-00530-6]
3. Perloff RM. The Dynamics of Persuasion: Communication and Attitudes in the 21st Century: Routledge; 1993. [doi: 10.4324/9781410606884]
4. Lerner JS, Li Y, Valdesolo P, Kassam KS. Emotion and decision making. Annu Rev Psychol 2015 Jan 3;66:799-823. [doi: 10.1146/annurev-psych-010213-115043] [Medline: 25251484]
5. Ferrer RA, Mendes WB. Emotion, health decision making, and health behaviour. Psychol Health 2018 Jan;33(1):1-16. [doi: 10.1080/08870446.2017.1385787] [Medline: 29261038]
6. Steinberg L. A social neuroscience perspective on adolescent risk-taking. In: Walsh A, Beaver KM, editors. Biosocial Theories of Crime: Routledge; 2017:435-463. [doi: 10.4324/9781315096278]
7. Oden A, Porter L. The kids are online: teen social media use, civic engagement, and affective polarization. Soc Media Soc 2023;9(3). [doi: 10.1177/20563051231186364]

XSL•FO
RenderX

8. Weinstein E, Kleiman EM, Franz PJ, et al. Positive and negative uses of social media among adolescents hospitalized for suicidal behavior. J Adolesc 2021 Feb;87:63-73. [doi: 10.1016/j.adolescence.2020.12.003] [Medline: 33493982]
9. Haddock A, Ward N, Yu R, O'Dea N. Positive effects of digital technology use by adolescents: a scoping review of the literature. Int J Environ Res Public Health 2022 Oct 27;19(21):14009. [doi: 10.3390/ijerph192114009] [Medline: 36360887]
10. Wu J, Benjamin EJ, Ross JC, Fetterman JL, Hong T. Health messaging strategies for vaping prevention and cessation among youth and young adults: a systematic review. Health Commun 2025 Mar;40(3):531-549. [doi: 10.1080/10410236.2024.2352284] [Medline: 38742648]
11. Tam MT, Wu JM, Zhang CC, Pawliuk C, Robillard JM. A systematic review of the impacts of media mental health awareness campaigns on young people. Health Promot Pract 2024 Sep;25(5):907-920. [doi: 10.1177/15248399241232646] [Medline: 38468568]
12. Currie GE, McLeod C, Waddington C, Snelling TL. SMS-based interventions for improving child and adolescent vaccine coverage and timeliness: a systematic review. BMC Public Health 2024 Jul 2;24(1):1753. [doi: 10.1186/s12889-024-18900-4] [Medline: 38956527]
13. Loescher LJ, Rains SA, Kramer SS, Akers C, Moussa R. A systematic review of interventions to enhance healthy lifestyle behaviors in adolescents delivered via mobile phone text messaging. Am J Health Promot 2018 May;32(4):865-879. [doi: 10.1177/0890117116675785] [Medline: 27923885]
14. Bottaro R, Faraci P. The use of social networking sites and its impact on adolescents' emotional well-being: a scoping review. Curr Addict Rep 2022;9(4):518-539. [doi: 10.1007/s40429-022-00445-4] [Medline: 36185594]
15. Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
16. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
17. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
18. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence. Lancet Child Adolesc Health 2018 Mar;2(3):223-228. [doi: 10.1016/S2352-4642(18)30022-1] [Medline: 30169257]
19. Lungu DA, Røislien J, Wiig S, et al. The role of recipient characteristics in health video communication outcomes: scoping review. J Med Internet Res 2021 Dec 30;23(12):e30962. [doi: 10.2196/30962] [Medline: 34967758]
20. Ng WJ, See YH, Cheung MW. The Influence of affective and cognitive appeals on persuasion outcomes: a cross-cultural meta-analysis. J Commun 2025 Apr;75(2):101-111. [doi: 10.1093/joc/jqae042]
21. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M. Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A Product from the ESRC Methods Programme: Lancaster University; 2006. URL: https://www.researchgate.net/publication/233866356_Guidance_on_the_conduct_of_narrative_synthesis_in_systematic_reviews_A_product_from_the_ESRC_Methods_Programme [accessed 2026-02-07]
22. Whittemore R, Knafl K. The integrative review: updated methodology. J Adv Nurs 2005 Dec;52(5):546-553. [doi: 10.1111/j.1365-2648.2005.03621.x] [Medline: 16268861]
23. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]

Abbreviations

MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by J Sarvestan; submitted 12.Jun.2025; peer-reviewed by N Hu; accepted 29.Dec.2025; published 02.Mar.2026.

Please cite as:
Berg SH, Ferrer-Conill R, Stifjell K, Hjorthaug Urstad K, Røislien J, Fjæran L
Affective Health Messages for Adolescents: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e78927
URL: https://www.researchprotocols.org/2026/1/e78927
doi:10.2196/78927

XSL•FO
RenderX

© Siv Hilde Berg, Raul Ferrer-Conill, Kristine Stifjell, Kristin Hjorthaug Urstad, Jo Røislien, Lisbet Fjæran. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 2.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding Responsible Development in AI-Based Clinical Prediction Models for Mortality: Protocol for a Scoping Review

Riley Martens[1], BN, RN; Jessalyn K Holodinsky[1,2,3,4], PhD; Jessica Simon[1,5,6], MB ChB, MRCP (UK), FRCPC; Zack Marshall[1,4,7], MSW, PhD

[1]Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, 3280 Hospital Dr NW, Calgary, AB, Canada
[2]Department of Emergency Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[3]Department of Clinical Neuroscience, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[4]O'Brien Institute for Public Health, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[5]Department of Oncology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[6]Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[7]School of Social Work, McGill University, Montréal, QC, Canada

Corresponding Author:
Riley Martens, BN, RN
Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, 3280 Hospital Dr NW, Calgary, AB, Canada

Abstract

Background: Prognostic inequity has been identified as a barrier to accessing end-of-life care for underrepresented groups. Artificial intelligence–based clinical prediction models (AIPMs) for prognostication of mortality have the potential to offer rapid, accessible, and accurate predictions that could streamline care. However, they may also exacerbate preexisting inequities in the health care system rather than address accessibility and quality. This can be caused by erroneous outputs from biased training data, outcomes from out-of-scope operationalization, and inexplicability due to opacity.

Objective: The goal of this study is to synthesize peer-reviewed literature on the creation and application of AIPMs to prognosticate mortality in acute care settings for adult patients, offering new insights into responsible and ethical model development.

Methods: A transdisciplinary, structured search strategy was developed in consultation with librarians from both health sciences and engineering sciences. The academic databases queried were Medline, Embase, IEEE Xplore, ACM Digital Library, Compendex, and Scopus. The search was conducted in spring 2025, and the results were uploaded to Covidence. A team of reviewers will screen in 2 rounds: titles and abstracts, then full texts. Eligibility will be determined by publication in academic journals or as full-length conference proceedings, language, model output, and AI use. Data will be charted using adapted charting tools and then analyzed by descriptive, summary, and qualitative synthesis.

Results: The search was completed on March 25, 2025, with screening starting in May 2025. Results are anticipated for January 2026.

Conclusions: This review will provide a comprehensive summary of AIPMs that predict mortality, highlighting the specific elements included in their development. Informed by the responsible research and innovation (RRI) framework, we will consider interest-holder engagement, interdisciplinary collaboration, and computational and clinical ethics will in the context of the four RRI dimensions: anticipation, reflexivity, inclusion, and responsiveness.

International Registered Report Identifier (IRRID): DERR1-10.2196/80325

(JMIR Res Protoc 2026;15:e80325) doi:10.2196/80325

KEYWORDS

artificial intelligence; AI; responsible development; mortality prediction; responsible research and innovation; RRI; machine ethics; patient engagement; sociotechnical; knowledge synthesis

XSL•FO
RenderX

Introduction

Background

Artificial intelligence–based clinical prediction models (AIPMs) have been developed in a range of medical disciplines to predict the outcomes of different diseases [1]. For example, cardiologists have developed AIPMs to predict mortality in heart failure, and neurologists have trained an AIPM to examine stroke outcomes [2,3]. Another tool, NYUTron, a large language model, showed considerable improvement compared to non-AI CPMs in determining in-hospital mortality for new admissions [4]. AIPMs have demonstrated the ability to predict mortality for patients across various conditions [2-4]. The availability of accurate and timely prognostication could provide tangible benefits to patients, clinicians, and health systems.

AIPMs also have the capacity to exacerbate preexisting inequities in the health care system instead of improving accessibility and equality [5]. This can be caused by erroneous outputs from biased training data, unintentional outcomes from out-of-scope operationalization, and inexplicability due to opacity [5-7]. Training data contain biases from disproportionate representation of some sociodemographic groups [5]. Using AIPMs beyond their intended purposes poses risks because of knowledge limits hardcoded into these tools [7]. Further, low AI literacy can influence patient engagement unfairly and may discourage interest holders from future involvement [6]. Responsible development relies on guiding principles for the design of technology that prioritizes harm reduction and prevention [8]. To safeguard against the exacerbation of inequities, responsible development is an encouraged approach to AIPM creation.

The responsible research and innovation framework (RRI) outlines 4 important components of accountable research [9]: anticipation, reflexivity, inclusion, and responsiveness. Each of these items tied to recommended action points [9]. Anticipation warrants foresight and future-oriented assessment, which can be impacted by limited scopes and siloed disciplines [9,10]. A proposed way to overcome these barriers would be to incorporate a transdisciplinary research approach [10]. Reflexivity encourages critical self-appraisal and evaluation of one's methods and work [9]. Standards and frameworks have been applied to encourage reflexivity, including guidelines regarding development [9]. Inclusion leads to engagement with communities and interest holders who have not always been involved in research [9,11]. This highlights the strategy of public engagement and interest holder consultation [6,11]. Responsiveness, in contrast to the other three aspects, requires extended action to answer questions, beyond information seeking [9]. Responsiveness includes reacting and changing the course of work based on feedback and implications [9]. In relation to AIPM development, impact assessment would be an apt strategy to address responsiveness [1].

This review benefits from the structure of responsible development and respective action items to explore AIPMs. Previous reviews have been conducted of AIPMs, typically highlighting technical performance or a particular pathology [1-3,12]. This review is differentiated by two elements: the responsible development lens and the transdisciplinary approach. This review will examine any AIPMs that use mortality as the primary outcome regardless of machine learning technique or specific pathology. Components of AIPM development will be explored, including interest holder engagement, interdisciplinary collaboration, and ethical frameworks. AIPMs developed for acute care contexts will be the focus, including in emergency rooms, critical care, or other in-patient medical units, as that is representative of the majority of these tools.

The objective is to synthesize peer-reviewed literature on the creation and application of AIPMs to predict mortality in acute care settings with adult patients, offering new insights into responsible and ethical model development.

Review Questions

This review will answer four questions regarding AIPMs predicting mortality. First, how are AI, outputs, and performance reported when describing the development of AIPMs? Second, in the acute care settings the AIPMs are developed and deployed in, what is their intended use, and how are they evaluated? Third, how have responsible research and innovation considerations been integrated into the development of these AIPMs? Do researchers report on ethics, interest-holder engagement, interdisciplinary collaboration, or other elements of responsible development? Fourth, what are the patterns, advances, gaps, evidence, and recommendations (ie, the elements of the PAGER framework) in the literature regarding the responsible development of AIPMs in predicting mortality [13]?

Methods

Overview

This scoping review will follow the extended Arksey and O'Malley [14] framework described by Levac et al [15] through six stages, from identifying the research question to consultation with interest holders. The protocol is formatted according to the JBI template for clear reporting [16]. The protocol will be submitted for publication to support the reproducibility and transparency of the review, pending anticipatory results [17]. Study eligibility will be determined using preestablished criteria. The results will be reported using the PAGER framework to highlight patterns, advances, gaps, evidence, and recommendations [13]. Additionally, this review will follow the PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist to ensure all necessary elements are included [18].

Inclusion Criteria

Formatting the objective into the population, concept, and context elements, as recommended by the JBI guidelines [19], the review objectives are stated in Table 1.

XSL•FO
RenderX

Table . Research objective in the population, concept and context format.

Element	Description
Population	Artificial intelligence–based clinical prediction models that predict mortality
Concept	Examining definitions, concepts, and elements of responsible development
Context	Artificial intelligence–based clinical prediction models used in acute care settings with adult patients

Population

Clinical prediction models that use machine learning or other methods that are considered to be AI and that provide mortality as the primary outcome will be the population for this review. This excludes any models that only use statistical techniques or provide outcomes other than mortality. Novel machine learning applications will be the focus, whereas preexisting clinical tools or external validation methods will not be considered. AI and machine learning in this review will considered to be supervised, nonparametric predictive modeling techniques with flexible, nonlinear relationships between predictors and a mortality outcome. These approaches include decision tree–based and ensemble algorithms, which differ from traditional regression models that impose linearity or distributional assumptions and from unsupervised clustering methods used for risk stratification.

Concept

Responsible development is the concept being examined, as defined by preset criteria, including interest-holder engagement, interdisciplinary collaboration, and computational and clinical ethics [20,21].

Context

AIPMs used in acute clinical contexts for adult inpatients in emergency rooms, in-hospital units, and critical care will be eligible. Any studies conducted in community settings, outpatient care, or pediatric populations will be excluded.

Types of Sources

All study designs will be considered, including experimental, quasi-experimental, qualitative, and mixed methods research. The sources deemed relevant to this review will be peer-reviewed articles published in scientific journals or full-length conference papers. Eligibility will be determined based on the research objectives and outputs reported. No language or date filters will be applied. Systematic reviews and scoping reviews will not be included as full texts but will be identified during the title and abstract screening phase. This is to ensure the quality and robustness of this review by further investigation of reference lists from other relevant reviews.

Any work not published in a scientific journal or in an academic conference proceeding will be excluded, such as gray literature, white papers, or opinion papers. This extends to any editorials or position papers where no development of a machine learning application occurred.

Search Strategy

A transdisciplinary, structured search strategy was developed in consultation with a health sciences librarian and an engineering sciences librarian. Precursory searches in EMBASE and Medline informed the identification of index terms. This was accomplished by using Yale's MeSH Analyzer combined with adapted search filters from the database of the CDA-AMC (Canada's Drug Agency–L'Agence des médicaments du Canada) and the University of Alberta's field-specific search resources [22-25]. A sample search strategy for Medline is shown in Multimedia Appendix 1. Four concepts related to the responsible development of prognostic AIPMs shaped the search strategy based on previously published search filters [22-24]. The first search concept includes any clinical decision support terms to capture the type of AI tool being developed, including clinical prediction models. The second concept in the search strategy relates to artificial intelligence and uses terms related to neural networks and machine learning. The third concept in the search focuses on mortality, including terms such as "death" and "end-of-life." The fourth concept reflects the acute setting in which the AIPMs are used and includes the terms "emergency department," "critical care," and "inpatient units."

The academic databases selected for this review are Medline, Embase, IEEE Xplore, ACM Digital Library, Compendex, and Scopus. Medline and Embase were chosen to capture any AIPMs published in health care journals. IEEE Xplore, ACM, and Compendex were chosen to collect literature from the engineering and computer science fields. Scopus was selected as the interdisciplinary database.

Selection of Studies and Sources of Evidence

After the consolidation and execution of the search strategy across databases, the results will be uploaded to Covidence [26]. Covidence's deduplication tool will be used to remove identical studies. There is potential for a large return of studies given the search strategy. To address this, screening will be done in stages in addition to the use of screening software, Covidence, and a team of multiple screeners. A team of secondary reviewers will begin the first round of screening on the search results. Pilot screening will occur in rounds, with batches of 100 studies, to assess rater agreement, with additional pilot screening to increase agreement as needed to reach an interrater agreement of 75% between the secondary team and primary reviewer. Overall, interrater agreement will be tracked automatically through Covidence, with conflicts being resolved through discussion and consensus. Similar inclusion and exclusion criteria will be applied in the second round of screening, when reviewers progress to full-text review (Multimedia Appendix 2 provides the detailed inclusion and exclusion criteria). For

XSL•FO
RenderX

full-text review, a proportionate agreement of 75% will also be agreed a priori with a pilot screening round of 25 studies. If the agreement rate does not hold after 25 studies, it will be repeated. The exclusion criteria, including gray literature and preprints, were chosen to define and synthesize peer-reviewed research. As the subject matter of AI and machine learning is rapidly evolving, the focus of the review was academic publications and the model development being reported. An updated search in the future will capture any preprints that are peer-reviewed and published.

The exclusion of full texts that are not in English should occur in the title and abstract round of screening, given the language criteria. If there are any non-English texts present in the subsequent screening round, they will be removed according to the eligibility criteria.

Any differences will be resolved through discussion and consensus. Upon completion of two rounds of screening, the studies relevant to the research question will progress to data charting. The review process will be illustrated in a PRISMA flow diagram when reported [18].

Data Charting

Following full-text review, studies will be stratified based on the level of RRI involvement. If the studies report no elements, limited charting will occur, but if the studies report multiple RRI elements, then full data charting will be done. Information from included studies in the scoping review will be charted by two independent reviewers using a data collection tool developed by the team. The data charting form includes adapted criteria from the TRIPOD (Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis)+AI checklist, the EQUALSS (ethnicity and race, qualifications and education, underserved area, age, language and religion, sex, sexual orientation) framework, the artificial intelligence risk management framework, and other guiding principles [27-30]. Data will include specific details about study participants, machine learning techniques, clinical settings, and impact assessments relevant to responsible development.

RRI principles of anticipation, reflexivity, inclusion, and responsiveness have informed the formation of the review and thus structured the data charting and subsequent analysis. Anticipation will be mapped to the level of transdisciplinarity of eligible studies. This will be assessed through any report of multidisciplinary, interdisciplinary, or transdisciplinary collaboration in addition to the academic composition of the study team. Reflexivity will correspond to any standards or frameworks implemented in the work, including best practices, ethical frameworks, and other relevant standards. These will be collated under charting fields and further refined into research, ethical, and clinical practice standards. Further, inclusion correlates to interest-holder involvement, especially patient and family engagement. This will be assessed using the Patient-Oriented Research Level of Engagement Tool (PORLET) [31] to determine level and extent of engagement. This will be followed by responsiveness and data charting relevant to this principle and impact assessment, assessed as the impact a tool has on clinical care, including fairness in patient demographics (determined by EQUALSS) and any reported environmental impacts.

A draft charting form is provided in Multimedia Appendix 3. Pilot data charting by at least two reviewers is planned for the first five eligible studies. For the data charting phase, the calibration of complex topics will involve a charting template with preestablished criteria and tools including TRIPOD+AI, PORLET, and EQUALSS. Tandem charting will occur, where multiple screeners will chart information from eligible studies followed by verification by another screener. If there are differences in the elements charted, adjudication by a third screener, the supervisor of the project, will occur.

The data charting tool will be iteratively revised as necessary during the process of extracting data from each included evidence source. As part of the subsequent analysis, after charting there will be a tally of the elements present that fulfill the RRI principles, including TRIPOD+AI, PORLET, EQUALSS, and the presence of fairness metrics. This tally will be calculated based on the presence of RRI elements, scores from TRIPOD+AI compliance, and PORLET. Any modifications will be detailed in the review.

Data Analysis and Presentation

The RRI principles have informed the elements for data charting, but PAGER provides the structure for the reporting of the charting. The patterns section will detail the most prevalent elements numerically across the four RRI principles. Advancements will showcase the development of the RRI principles chronologically based on date of publication. Gaps will be demonstrated by any anomalies that are present in patterns or advancements. Evidence for practice will be synthesized based on the charted elements that had the highest prevalence of RRI principles and prognostic accuracy. Additionally, any recommendations from authors of eligible studies will be collated. Recommendations for research will be synthesized based on observed gaps and charted data by the study team.

Adhering to the PAGER framework, a patterning chart will be used to organize and visualize advances and gaps in the evidence [13]. Additionally, this will be accompanied by descriptive tabulations and co-occurrence analysis [32]. This will be done to further identify the frequency of each charted element and the overlap between them [32]. Co-occurrence will follow a deductive approach using conceptual codes from the data charting forms. This process will further identify which concepts appear together or appear in isolation, as well as overarching themes.

A narrative summary will follow to highlight the prevalent aspects of the population, concept, and context, which includes machine leaning models, responsible development, and clinical settings [13,19]. An additional component will include an analysis of the chronology of the publications. Through these analyses, the considerations and implications of responsible development of AIPMs predicting mortality will be made clear.

XSL•FO
RenderX

Results

The search for this study was conducted on March 25, 2025, and screening commenced in May 2025 with anticipated reporting by January 2026. The initial search returned 9475 results after deduplication, and the search will be updated in summer 2025 to ensure that newly published literature is assessed in this review.

Discussion

Anticipated Findings

AIPMs have the potential to improve prognostication, but without attention to responsible development, they could also amplify inequities [5]. Certain ethical criteria have been highlighted in the use of AI, including technological literacy, patient engagement, and explainability [6,33]. Reviews or other studies exploring the ethical dimensions of AIPMs typically do so in isolation [6,33]. The exploration of these aspects in combination is a novel facet of this review.

Palliative care is often seen as one of the areas least impacted by AI due to its relational nature and specific focus on the care continuum [34]. New ethical precedents such as surrogate decision-making and goals of care have been established with the introduction of novel technology in end-of-life care, as demonstrated by feeding tubes and ventilators that can prolong life beyond a time that the patient can consent or decline certain treatments [35]. AIPMs may significantly impact palliative care practices by reducing some of the uncertainty in prognostication. With growing interest in large language models and agentic AI, this is an opportune time to codify responsible development [34,36,37].

Palliative care must be patient-centered, as patients and families choose between different care pathways [35]. Information, education, and transparency empower patients to make decisions based on their values and priorities. Opacity and low digital literacy are barriers to patient-centered care in this context, as any AIPM providing prognostic information to a clinician can influence care [6,38]. Opacity means these models are not easily understood, nor are they explainable in their functioning, resulting in a clinician potentially needing more technical knowledge to critically evaluate their output [39]. Additionally, the digital literacy levels of the clinician, patient, and family can hinder decision-making when using or attempting to use information from these tools [6,33]. Clinicians may not be able to explain how the information is derived, and as a result patients and families may not be familiar or comfortable with the technology.

Limitations and Strengths

As machine learning and AI are trending topics, the growth in literature is rapid [37]. Timeliness is one limitation to the relevance of this review's findings, as new guidelines, frameworks, or outlines determining responsible development continue to be developed and introduced [29,40]. Though the inclusion of seven databases should yield strong interdisciplinary results, the noted absence of CINAHL does represent a limitation of this review. We considered excluding PSYCInfo and Web of Science due to their overlaps with other databases, notably Scopus and Medline. We excluded CINAHL to emphasize the technical and biomedical literature, which are more present in other databases. When we obtained nearly 10,000 results, we made the decision to finalize the databases. Limitations related to databases and the timeliness of the review, given the rapidly evolving subject matter, will be addressed in an updated search. Changes in technology could also undermine the salience of the results, as a potential shift from prediction models to large language models or agentic AI could take place. This makes up-to-date knowledge synthesis in this area particularly challenging. The emphasis on ethical elements during development will help inform future work in this area. This is why focusing on concepts in responsible development, including interest holder engagement, interdisciplinary work, and ethics, will make it more likely that this review will remain timely despite changing regulations, frameworks, or technology. In addition, this review will highlight critical concepts related to responsible development across medical specialties for AIPM in acute care. To avoid perpetuating prognostic inequities, exploring the extent of responsible development in AIPM could help identify and reduce concerns related to bias, inexplicability, and consent [9].

Acknowledgments

GPT-5.0 was used for development of the decision tree (Multimedia Appendix 4) and for standardization of the definition of artificial intelligence models.

Funding

RM is a trainee supported through a Collaborative Research and Training Experience (CREATE) grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) in the Responsible Development of AI (grant 554764-2021).

Data Availability

As studies for this review were queried and retrieved from databases, the replication of the search strategy provided should reproduce the relevant studies. Due to the number of studies, a list showing all the studies that met eligibility for data charting will be provided.

XSL•FO
RenderX

Authors' Contributions

RM contributed to the study design, data collection, proposed analyses, and manuscript writing. ZM supported the study design, data collection, manuscript writing, and supervision. JKH and JS contributed to the study design, manuscript revising, and supervision.

Conflicts of Interest

JKH has investments in DESTINE Health Inc. The other authors have no conflicts to declare.

Multimedia Appendix 1
Search strategy.
[DOCX File, 14 KB - resprot_v15i1e80325_app1.docx]

Multimedia Appendix 2
Inclusion and exclusion criteria for screening.
[DOCX File, 15 KB - resprot_v15i1e80325_app2.docx]

Multimedia Appendix 3
Data charting form.
[DOCX File, 15 KB - resprot_v15i1e80325_app3.docx]

Multimedia Appendix 4
Decision tree for artificial intelligence model eligibility.
[DOCX File, 96 KB - resprot_v15i1e80325_app4.docx]

References

1. de Hond AAH, Leeuwenberg AM, Hooft L, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. NPJ Digit Med 2022 Jan 10;5(1):2. [doi: 10.1038/s41746-021-00549-7] [Medline: 35013569]
2. Yang Y, Tang L, Deng Y, et al. The predictive performance of artificial intelligence on the outcome of stroke: a systematic review and meta-analysis. Front Neurosci 2023;17:1256592. [doi: 10.3389/fnins.2023.1256592] [Medline: 37746141]
3. Błaziak M, Urban S, Wietrzyk W, et al. An artificial intelligence approach to guiding the management of heart failure patients using predictive models: a systematic review. Biomedicines 2022 Sep 5;10(9):2188. [doi: 10.3390/biomedicines10092188] [Medline: 36140289]
4. Jiang LY, Liu XC, Nejatian NP, et al. Health system-scale language models are all-purpose prediction engines. Nature New Biol 2023 Jul;619(7969):357-362. [doi: 10.1038/s41586-023-06160-y] [Medline: 37286606]
5. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med 2018 Dec 18;169(12):866-872. [doi: 10.7326/M18-1990] [Medline: 30508424]
6. Adus S, Macklin J, Pinto A. Exploring patient perspectives on how they can and should be engaged in the development of artificial intelligence (AI) applications in health care. BMC Health Serv Res 2023 Oct 26;23(1):1163. [doi: 10.1186/s12913-023-10098-2] [Medline: 37884940]
7. Forsythe D. Studying Those Who Study Us: An Anthropologist in the World of Artificial Intelligence: Stanford University Press; 2001.
8. Stryker C. What is responsible AI? IBM. URL: https://www.ibm.com/think/topics/responsible-ai [accessed 2025-01-04]
9. Stilgoe J, Owen R, Macnaghten P. Developing a framework for responsible innovation. Res Policy 2013 Nov;42(9):1568-1580. [doi: 10.1016/j.respol.2013.05.008]
10. Newman-Griffis D. AI thinking: a framework for rethinking artificial intelligence in practice. R Soc Open Sci 2025 Jan;12(1):241482. [doi: 10.1098/rsos.241482] [Medline: 39780964]
11. Canadian Institutes of Health Research. Strategy for patient-oriented research: patient engagement framework. 2014. URL: https://cihr-irsc.gc.ca/e/documents/spor_framework-en.pdf [accessed 2025-06-18]
12. Xin Y, Li H, Zhou Y, et al. The accuracy of artificial intelligence in predicting COVID-19 patient mortality: a systematic review and meta-analysis. BMC Med Inform Decis Mak 2023 Aug 9;23(1):155. [doi: 10.1186/s12911-023-02256-7] [Medline: 37559062]
13. Bradbury-Jones C, Aveyard H, Herber OR, Isham L, Taylor J, O'Malley L. Scoping reviews: the PAGER framework for improving the quality of reporting. Int J Soc Res Methodol 2022 Jul 4;25(4):457-470. [doi: 10.1080/13645579.2021.1899596]
14. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]

XSL•FO
RenderX

15. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5(1):69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
16. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
17. Hillis C, Bagheri E, Marshall Z. The role of protocol papers, scoping reviews, and systematic reviews in responsible AI research. IEEE Technol Soc Mag 2025 Jun;44(2):72-76. [doi: 10.1109/MTS.2025.3558666]
18. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
19. Pollock D, Peters MDJ, Khalil H, et al. Recommendations for the extraction, analysis, and presentation of results in scoping reviews. JBI Evid Synth 2023 Mar 1;21(3):520-532. [doi: 10.11124/JBIES-22-00123] [Medline: 36081365]
20. Varkey B. Principles of clinical ethics and their application to practice. Med Princ Pract 2021;30(1):17-28. [doi: 10.1159/000509119] [Medline: 32498071]
21. Awad E, Levine S, Anderson M, et al. Computational ethics. Trends Cogn Sci 2022 May;26(5):388-405. [doi: 10.1016/j.tics.2022.02.009] [Medline: 35365430]
22. CEP methods appendix: standard search filters July 2022. Penn Medicine: Center for Evidence-based Practice. URL: https://www.med.upenn.edu/CEP/assets/user-content/documents/cep-methods-appendix-07-2022.pdf [accessed 2024-12-13]
23. Campbell S, Kung J. Google Docs. Filters to retrieve studies related to artificial intelligence. 2024. URL: https://docs.google.com/document/d/1eWyO0jv9_6FYsxyC5LUYwFe9eH_3h83-tPNZ6wmos18/edit?tab=t.0 [accessed 2024-12-13]
24. Glanville J, Lefebvre C, Manson P, Robinson S, Shaw N. Prognosis filters for MEDLINE: performance data. University of York. 2022. URL: https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/home/prognosis-performance-data [accessed 2024-12-13]
25. Hocking R. Yale MeSH Analyzer. J Can Health Libr Assoc 2017 Dec 1;38(3). [doi: 10.29173/jchla/jabsc.v38i3.29336]
26. Covidence systematic review software. Veritas Health Innovation. URL: https://www.covidence.org [accessed 2026-02-12]
27. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 2024 Apr 16;385:e078378. [doi: 10.1136/bmj-2023-078378] [Medline: 38626948]
28. McCann L, Johnson L, Gkiouleka A, Pearce H, Ford J. EQUALSS GUIDE Multiple: a novel framework for equity-focused evidence synthesis. Public Health Pract (Oxf) 2025 Jun;9:100600. [doi: 10.1016/j.puhip.2025.100600] [Medline: 40104249]
29. Badal K, Lee CM, Esserman LJ. Guiding principles for the responsible development of artificial intelligence tools for healthcare. Commun Med (Lond) 2023 Apr 1;3(1):47. [doi: 10.1038/s43856-023-00279-9] [Medline: 37005467]
30. Artificial intelligence risk management framework (AI RMF 1.0). : National Institute of Standards and Technology (US); 2023 Jan URL: https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf [accessed 2026-02-12] [doi: 10.6028/NIST.AI.100-1]
31. Patient-Oriented Research Level of Engagement Tool. Saskatchewan Centre for Patient-Oriented Research. URL: https://static1.squarespace.com/static/5c869fd0e666695abe893b3b/t/61b0f04d878a731b75039cdf/1638985805316/PORLET+2021+12+08.pdf [accessed 2026-02-28]
32. Pirri S, Lorenzoni V, Turchetti G. Scoping review and bibliometric analysis of big data applications for medication adherence: an explorative methodological study to enhance consistency in literature. BMC Health Serv Res 2020 Jul 24;20(1):688. [doi: 10.1186/s12913-020-05544-4] [Medline: 32709237]
33. Salwei ME, Carayon P. A sociotechnical systems framework for the application of artificial intelligence in health care delivery. J Cogn Eng Decis Mak 2022 Dec;16(4):194-206. [doi: 10.1177/15553434221097357] [Medline: 36704421]
34. Stewart J, Lu J, Gahungu N, et al. Western Australian medical students' attitudes towards artificial intelligence in healthcare. PLOS ONE 2023;18(8):e0290642. [doi: 10.1371/journal.pone.0290642] [Medline: 37651380]
35. Youngner SJ, Arnold RM. The Oxford Handbook of Ethics at the End of Life: Oxford University Press; 2016. [doi: 10.1093/oxfordhb/9780199974412.001.0001]
36. Fleming M, Li W, Thompson N. The last mile problem in AI. Brookings. 2024 Aug. URL: https://www.brookings.edu/articles/the-last-mile-problem-in-ai/ [accessed 2026-02-12]
37. Xie Y, Zhai Y, Lu G. Evolution of artificial intelligence in healthcare: a 30-year bibliometric study. Front Med 2025 Jan 15;11:1505692. [doi: 10.3389/fmed.2024.1505692]
38. Bjerring JC, Busch J. Artificial intelligence and patient-centered decision-making. Philos Technol 2021 Jun;34(2):349-371. [doi: 10.1007/s13347-019-00391-6]
39. Price WN. Big data and black-box medical algorithms. Sci Transl Med 2018 Dec 12;10(471):eaao5333. [doi: 10.1126/scitranslmed.aao5333] [Medline: 30541791]
40. Voluntary code of conduct on the responsible development and management of advanced generative AI systems. Innovation, Science and Economic Development Canada. 2024. URL: https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems [accessed 2026-02-12]

Abbreviations

AI: artificial intelligence

XSL•FO
RenderX

AIPM: artificial intelligence–based clinical prediction model
CDA-AMC: Canada's Drug Agency–L'Agence des médicaments du Canada
EQUALSS: ethnicity and race, qualifications and education, underserved area, age, language and religion, sex, sexual orientation
PAGER: patterns, advances, gaps, evidence, and recommendations
PORLET: Patient-Oriented Research Level of Engagement Tool
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
RRI: responsible research and innovation
TRIPOD+AI: Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis + AI

Edited by J Sarvestan; submitted 01.Aug.2025; peer-reviewed by M Elbattah, T Karen; revised version received 14.Jan.2026; accepted 16.Jan.2026; published 05.Mar.2026.

Please cite as:
Martens R, Holodinsky JK, Simon J, Marshall Z
Understanding Responsible Development in AI-Based Clinical Prediction Models for Mortality: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e80325
URL: https://www.researchprotocols.org/2026/1/e80325
doi:10.2196/80325

© Riley Martens, Jessalyn K Holodinsky, Jessica Simon, Zack Marshall. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 5.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Determinants of Street Food Hygiene Practices and the Effectiveness of Interventions for Vendors and Consumers in Low- and Middle-Income Countries: Protocol for a Scoping Review

Abul Kamal[1*], MSc; Tarique Md Nurul Huda[2*], PhD; Sonjida Mesket Simi[3*], MPH; Elizabeth D Thomas[4*], PhD; Farhana Sultana[1*], MPH; Jesmin Sultana[1*], MPH; Rehnuma Haque Sarah[1*], PhD; Soha Ishrak Rafa[1*], MPH; Om Prasad Gautam[5], PhD; Peter John Winch[4*], MPH, MD; Mahbubur Rahman[1,6*], MPH

[1]Environmental Health and WASH, Health Systems and Population Studies Division, International Centre for Diarrhoeal Disease Research, Bangladesh, 68, Shaheed Tajuddin Ahmed Sarani, Mohakhali, Dhaka, Bangladesh

[2]Department of Public Health, College of Applied Medical Sciences, Qassim University, P.O. Box 6666, Qassim, Buraydah, Saudi Arabia

[3]Poverty, Gender, and Inclusion Unit, International Food Policy Research Institute, Dhaka, Bangladesh

[4]Department of International Health, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, United States

[5]WaterAid, Hygiene, London, United Kingdom

[6]Global Health and Migration Unit, Department of Women's and Children's Health, Uppsala University, Sweden

[*]these authors contributed equally

Corresponding Author:
Abul Kamal, MSc
Environmental Health and WASH, Health Systems and Population Studies Division, International Centre for Diarrhoeal Disease Research, Bangladesh, 68, Shaheed Tajuddin Ahmed Sarani, Mohakhali, Dhaka, Bangladesh

Abstract

Background: Street food is readily available food and beverages sold by vendors, frequently situated along streets or other public spaces. Street food is largely popular because it is low cost and readily available. However, unsafe food hygiene behaviors and conditions contribute to a substantial burden of foodborne illness. There is a gap in understanding what factors determine food hygiene behavior among street food vendors and consumers in low-income countries (LICs) and low- and middle-income countries (LMICs).

Objective: This review aims to identify the determinants of safe food hygiene behavior and explore the effectiveness of food hygiene interventions for vendors and consumers in LICs and LMICs.

Methods: We will search the PubMed, Cochrane Library, Web of Science, ProQuest, and Scopus databases for peer-reviewed articles and Google and Google Scholar for gray literature published in the English language from inception to October 31, 2023. We will apply the World Bank 2023 definition of LICs and LMICs to include studies from these countries in this review. Two independent reviewers will search, screen, and analyze the included literature. Data will be analyzed by narrative synthesis for all objectives.

Results: This protocol outlines the methodology for conducting a scoping review. The findings will identify gaps in the existing literature and map the determinants of street food hygiene, as well as the interventions implemented for street food vendors and consumers. The project received funding in October 2022 and institutional review board approval in December 2022. Data analysis is currently underway, and the first results are expected to be submitted for publication in April 2026.

Conclusions: To the best of our knowledge, this will be the first scoping review to investigate the determinants of street food hygiene behaviors and the effectiveness of interventions for vendors and consumers on street food hygiene in LICs and LMICs.

International Registered Report Identifier (IRRID): PRR1-10.2196/68633

(JMIR Res Protoc 2026;15:e68633) doi:10.2196/68633

KEYWORDS

food safety; hazard analysis public health; behavior change; hygiene practices; foodborne disease; street food; low- and middle-income countries; LMICs

XSL•FO
RenderX

Introduction

Background

Ready-to-eat food and beverages sold in public places or in informal settings using temporary or mobile infrastructure are generally referred to as street food [1]. Street-vended food is susceptible to chemical and microbiological contamination due to factors including vending location, raw materials and utensils used, storage and reheating practices, and the personal hygiene of street vendors [2]. Food hygiene is an important component of food safety. It includes behaviors and practices involved in the handling, preparing, storing, and serving of food to minimize the likelihood of individuals falling ill with foodborne illnesses. Street food vendors can play a role in transmitting disease while handling, preparing, serving, and storing food [2-4]. Multiple studies have found an association between contaminated street food and poor hygiene behaviors and practices, including storing food uncovered or at inappropriate temperatures, unhygienic handling of food, using poorly cleaned dishes, using a contaminated water supply, suboptimal reheating of food, and preparing and serving food in unhygienic surroundings and in the absence of handwashing and sanitation facilities [2,4-11]. Contaminated street food has been linked with foodborne disease outbreaks, including hepatitis A and cholera [12-14], and foods sold on the street are frequently contaminated with bacteria including *Salmonella* species, *Shigella* species, and *Escherichia coli*, which can pose serious health concerns to the general public [15,16]. An indicator of fecal contamination in food is *E coli*, which was found in 100% of street food samples in Nakawa, Uganda, and 43.36% of samples in Kathmandu [6,15]. In Coimbatore, Tamil Nadu, India, 70% of food samples from street vendors contained *Salmonella* species [17], and in Bangladesh, *Shigella* species was detected in 12% of food samples collected from both fixed and mobile food vendors [18]. The presence of *E coli*, *Salmonella* species, *and Shigella* species in the food may be linked to poor hand hygiene among food workers and lack of good manufacturing practices [15].

One of the leading causes of morbidity and death worldwide is foodborne illness. In 2010, the World Health Organization estimated that 600 million people experienced foodborne illness, with 420,000 related deaths [14,19]. The disease burden is higher in low- and middle-income countries (LMICs), where street food is popular due to its wide availability and low cost [1,2]. Inadequate sanitation and hygiene facilities and practices, limited surveillance, and a deficit in awareness and food hygiene education contribute to the heightened disease burden in LMICs [20,21]. After the COVID-19 pandemic, there has been a notable increase in informal businesses across many LMICs driven by economic necessity [22-25]. The pandemic significantly disrupted formal employment, prompting many individuals to seek alternative income sources within the informal sector [24,25]. At the same time, the COVID-19 pandemic may have influenced consumer behavior, with some individuals choosing to support street food vendors out of a sense of solidarity and a desire to assist those facing economic hardship [26].

Several food hygiene and safety interventions have been implemented in low-income countries (LICs) and LMICs to target street food vendors and consumers to raise awareness about the importance of street food hygiene to reduce the cases of foodborne illnesses [27-32]. However, the findings from these studies have not yet been synthesized to inform policy recommendations for LICs and LMICs. To inform effective interventions, a comprehensive understanding of the determinants of street food hygiene is needed. This requires looking beyond food hygiene knowledge and taking a closer look at the social, technological, physical, and psychosocial factors that influence the food hygiene behaviors and practices of vendors and consumers, as well as related food contamination [33,34]. Food hygiene promotion programs are likely to be most successful if they address a range of critical behavioral determinants [35]. Several studies have identified determinants of food hygiene behaviors and practices in different settings [2,4,11,36]. For instance, Birgen et al [4] found that vending places, storage facilities, and appropriate clothing are essential factors contributing to *E coli* contamination in raw chicken products vended in Nairobi city [4]. In other studies, knowledge and awareness, along with contextual factors such as settings, were also identified as important behavioral determinants [33,34].

Systematic reviews assessing the effects of education interventions and training on food safety knowledge, attitudes, and behaviors of food handlers have indicated improvements in these domains among restaurant-based vendors as a means to avoid food safety risks [37-40]. Some studies looked at ways to improve food hygiene among street food vendors in LICs and LMICs [36,41]. For instance, a study by Umar et al [36] in Nigeria revealed an enhancement in both knowledge and implementation of food hygiene practices among street food vendors following a training intervention. While several studies have assessed the impact of interventions on food hygiene practices among street food vendors, there is limited evidence on interventions targeting consumers, particularly regarding their behaviors around food handling and consumption. Currently, no systematic reviews explore the effects of food hygiene interventions on behavior and microbial outcomes on street food are registered on PROSPERO (International Prospective Register of Systematic Reviews). There is also no current scoping review on the determinants of street food hygiene in LICs and LMICs, and there are no interventions for street food vendors and consumers in LICs and LMICs to improve hygiene. Given that this area remains relatively unexplored, this review aims to synthesize evidence related to street food hygiene (concept) among vendors (population) in LICs and LMICs (context). The focus will be on the determinants of safe street food hygiene behaviors and the impact of street food hygiene interventions on street food vendors' and consumers' food hygiene behavior and microbiological contamination.

Objectives

The specific objectives of this review are to synthesize existing evidence on (1) the determinants of food hygiene behavior among street food vendors, (2) explore the effects of food hygiene interventions targeting street food vendors on food hygiene behavior and the microbiological contamination of street food in LMICs and LICs, and (3) explore the effects of

XSL•FO
RenderX

food hygiene interventions for consumers on safe street food consumption in LMICs and LICs.

Methods

Protocol Design

We will follow the methodological frameworks proposed by Arksey and O'Malley [42] and Levac et al [43] for conducting a scoping review (Figure 1) [44]. This framework comprises six key phases: (1) specifying the research objectives, (2) finding relevant studies, (3) screening studies in line with predetermined criteria, (4) charting the data, (5) summarizing the findings, and (6) confirming the study results with potential stakeholders involved. Implementing this framework ensures the overall reliability of the review.

Figure 1. Flow diagram of the steps for conducting the scoping review. PCC: population, concept, and context.

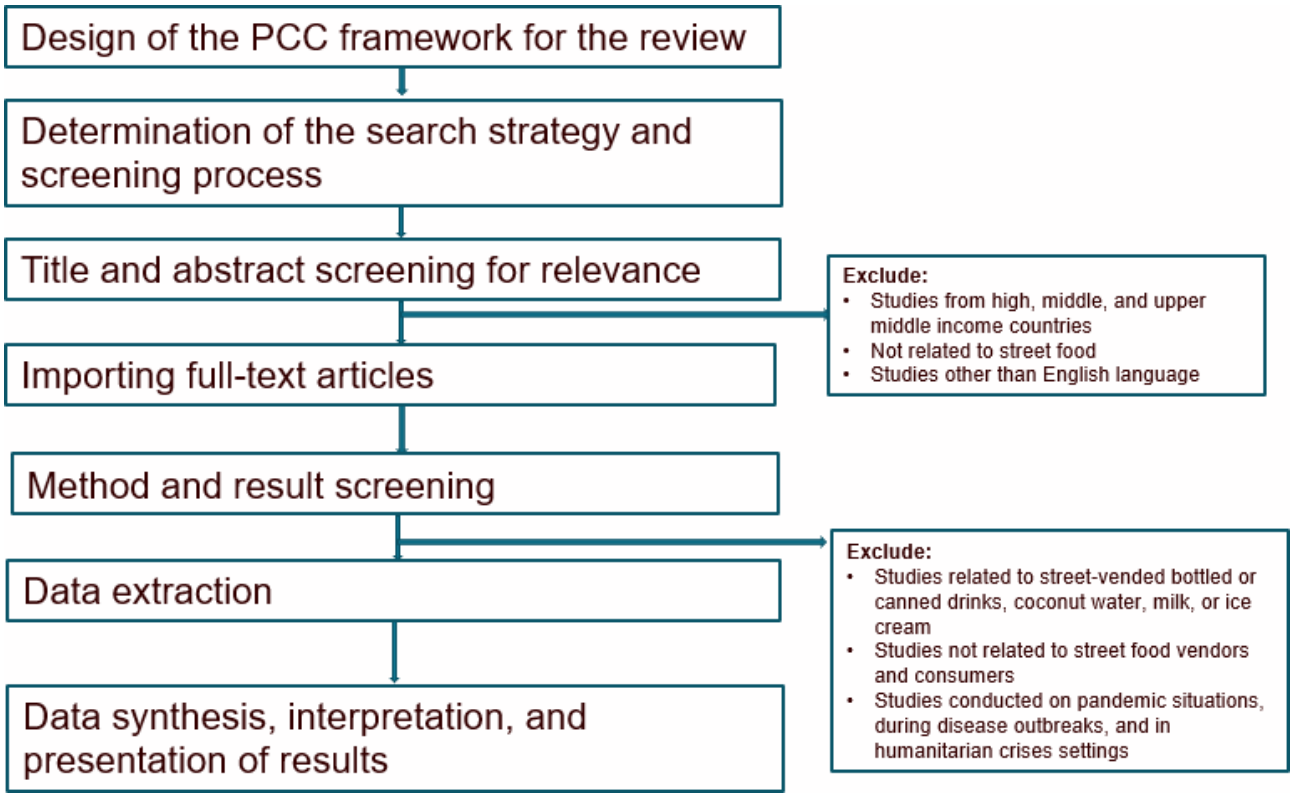

For example, it suggests involving 2 reviewers for the initial screening of titles and abstracts and piloting data extraction to maintain consistent alignment with the research objectives during the extraction process [45]. This framework helps us map available evidence on the determinants of street food hygiene and interventions related to street food hygiene among vendors, focusing on food hygiene behavior and microbiological contamination, as well as interventions targeting consumers in LICs and LMICs. It also enables us to capture diverse behavior change strategies across various informal food vending settings.

Identification of the Research Objectives

We used an iterative process to develop the research objectives for this scoping review, including familiarization with published work on street food and subsequent discussions with the research team.

Study Identification

Inclusion and Exclusion Criteria

Articles published in the English language from inception to October 31, 2023, will be considered for review. Studies will be included based on the PCC (population, concept, and context) framework. Eligible studies must address the PCC framework [46]. We will include all observational and experimental studies for objective 1. We will only include those studies that show an association between determinants and food hygiene behavior (food handling, preparation, serving, and storage) for objective 1. For objectives 2 and 3, we will include only intervention studies (experimental and quasi-experimental). Detailed information on the inclusion criteria for this review can be found in the Population and Context and the Concept sections. Studies related to street-vended bottled or canned drinks, coconut water, milk, and ice cream will be excluded.

Population and Context

The context for this review is low-resource settings, which include both LICs and LMICs as classified by the World Bank in 2023, comprising a total of 82 countries [47]. The population of interest will be street food vendors and consumers from LICs and LMICs. Studies related to street food vendors will be included for objectives 1 and 2, while studies related to street food consumers will be included for objective 3. We define a street food vendor as one who sells food from a temporary or mobile stall in the street or an open public location [48]. We define a street food consumer as an individual who acquires and consumes food or beverages made and sold by vendors in public areas such as streets, marketplaces, or parks, typically for immediate consumption. Articles not related to street food

XSL•FO
RenderX

(eg, home cooking) will be excluded. Regarding context, studies conducted on pandemic situations, during disease outbreaks, and in humanitarian crisis settings will be excluded.

Concept

The core concept for this review is street food hygiene. Defining "street food hygiene" in LICs and LMICs presents particular challenges due to several context-specific factors such as inadequate infrastructure, diverse vendor food handling practices, challenges in country-specific regulations, lack of education, and socioeconomic constraints [49-52]. Street food vendors often face limited access to clean water and sanitation facilities in many LICs and LMICs, which are essential for maintaining hygiene [52]. The complexity is further compounded when there is insufficient government oversight and enforcement of food safety regulations, which allows unhygienic practices to persist unchecked [50]. To guide our review, we adopted the definition of street food hygiene from the World Health Organization and the Food and Agriculture Organization. We characterize street food hygiene as the actions undertaken by food vendors throughout the preparation process (such as handwashing with soap prior to or during food preparation, washing utensils, and maintaining specific cooking temperatures); during food handling (including preventing cross-contamination between cooked and raw foods and maintaining overall kitchen cleanliness); during serving (including handwashing with soap before serving and using clean utensils); and during food storage (ensuring storage at appropriate temperatures and reheating before consumption). These measures also encompass actions taken by consumers during the handling, consumption, and storage of purchased food [53]. The definitions of preparation, handling, serving, and storing are described in Table S4 in Multimedia Appendix 1. The definitions of determinants are described in Table S5 in Multimedia Appendix 1.

Data Sources

The research team will use the search terms to explore a range of electronic databases for peer-reviewed and gray literature articles. Two reviewers will jointly search the PubMed, Cochrane Library, Web of Science, ProQuest, and Scopus databases for peer-reviewed articles. Google and Google Scholar will be searched for gray literature (reports, theses, and dissertations). The reviewers will do a manual search of all included articles' reference lists to find any more pertinent papers for the review and to ensure the thoroughness of the search. We will exclude any type of review articles related to the topic; however, the reference lists within these reviews will be examined to identify relevant articles. Non–peer-reviewed book chapters will also be excluded. Peer-reviewed journal articles will be included, while unpublished theses, dissertations, and reports will be treated as gray literature and will be included. Conference proceedings will be excluded.

Search Strategies

To find keywords for an advanced search, a preliminary search for published scientific articles on the subject of interest will be carried out. Both the PCC and Behavior-Centered Design (BCD) frameworks will be referenced for developing search strategies for all 3 objectives to ensure completeness of possible determinants that influence street food hygiene behaviors [54].

The BCD framework, which draws from multiple disciplines, including psychology and behavioral economics, will be referenced to help capture the behavioral determinants discussed in the literature in the context of street food hygiene [54]. This framework has been widely applied to help identify behavioral determinants in water, sanitation, and hygiene research [33,35]. It includes a broad set of determinants that fall into 4 main categories: environment (climate and society), settings (infrastructure and norms), brain (knowledge and motivation), and body (senses and traits). The range of determinants included in the BCD framework will facilitate the capture of the determinants of food hygiene behavior from the literature.

The search strategy will be adjusted according to the instructions of each database. Truncations will be used to broaden the search where appropriate. Synonyms and Medical Subject Headings terms will be used alongside the "OR" Boolean operator to broaden the search. The "AND" Boolean operator will combine the terms effectively to ensure relevance and prevent an influx of unrelated articles. Search terms will first be piloted in any database, downloading a few relevant articles based on the objectives. If required, we will adjust the search strategies and perform a second search to ensure we retrieve relevant articles. If we successfully identify relevant articles, we will proceed to search another database; if not, we will continue refining the search criteria until we do find the relevant articles.

To address objective 1, we developed search terms for each determinant using the BCD framework and terms related to street food hygiene behaviors. In this review, we defined behavioral determinants as those that influence behaviors positively and negatively within the behavioral context and settings such as physical, social, and biological environment; motives; barriers; social-eco-demo-cultural variability; power relations, etc. Search terms were categorized into 3 groups: terms related to street food or street food vendors (population), terms related to determinants, and terms related to food hygiene behaviors and practices (ie, patterns of behaviors; concept). For objectives 2 and 3, we used terminology associated with behavioral determinants to identify interventions that incorporated behavioral strategies grounded in a theoretical behavior change framework (concept). Additionally, we also used broader intervention-related terms. Textbox 1 presents the selection of search terms used for all 3 objectives. The complete search strategies used for the PubMed database are provided in Tables S1-S3 in Multimedia Appendix 1. A librarian will be invited to review the search strategies. The search strategy will be developed by 2 reviewers with input from a third reviewer. For content validity, search terms will be reviewed by the team members, composed of experts in the field.

XSL•FO
RenderX

Textbox 1. Key search terms to be used to search articles.

Sample search terms for objective 1

- Population-related terms (combined by "OR"): "vendor, street vended, street vendor, mobile food vendor, street market, food truck, food cart, hawker, drink, juice"
- Determinant-related terms (combined by "OR")

o Theory-related terms: "behavior, theory, framework, factor"

o Determinant-related terms: "reward, trade-off, contamination, age, gender, education, skill, capacity"

- Food hygiene behavior– and contamination-related terms (combined by "OR"): "food preparation practice, reheat, temperature, storage, hazard analysis critical control point, *Salmonella, E. coli, Shigella"*

Sample search terms for objectives 2 and 3

- Population-related terms (combined by "OR"): "vendor, street vended, street vendor, mobile food vendor, street market, food truck, food cart, hawker, ready to drink, drink, juice, customer, buyer, consumer"
- General intervention and determinant-related terms (combined by "OR")—intervention-related terms: "intervention, counselling, method, evaluation, video, booklet, demonstration, poster, framework, handwashing facilities, inspection"
- Food hygiene behavior—and contamination-related terms (combined by "OR"): "food preparation practice, reheat, temperature, storage, hazard analysis critical control point, *Salmonella, E. coli, Shigella"*

Article Screening

This review will follow the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) reporting guidelines for the screening process [55]. We will go through four phases in selecting articles: (1) identification, (2) screening, (3) eligibility assessment, and (4) inclusion. The search results will be imported into EndNote (version X9; Clarivate), where duplicate articles will be removed. The updated list will then be imported into Rayyan (Rayyan Systems, Inc) online software, where additional duplicates will be identified and eliminated. To ensure consistency among reviewers, a pilot test will be conducted before the screening. The screening criteria will be refined based on pretesting with approximately 20 titles and abstracts to ensure the capture of relevant studies.

Initially, 2 reviewers will independently assess the titles and abstracts of articles using the specified inclusion and exclusion criteria. In case of any queries or disagreements between the 2 reviewers, a third reviewer will step in to address them. Subsequently, the full text of the selected articles will be reviewed by the 2 reviewers to make the final decision on inclusion. After completing the review, we will use the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram to report the final numbers. The reviewer will review each excluded article again after screening the title and abstract to ensure that no pertinent articles have been mistakenly excluded. Studies for which it is unclear whether the criteria will be met based on the title and abstract will undergo further discussion among the reviewers to reach a decision on inclusion. All the remaining articles at this stage will undergo a comprehensive full-text screening, where the methodology and results sections will be primarily evaluated to determine if the study's approach aligns with the objectives of the scoping review. Any discrepancies in eligibility decisions will be resolved through discussions among the reviewers.

Data Extraction

At least 2 reviewers will work together to extract data from the included articles according to the 3 objectives (eg, determinants of food hygiene, the effect of the intervention for vendors, and the effect of the intervention for consumers). We will also extract data on microbial contaminations including study design; location; sample size; settings of street vendors (eg, urban slum, school, and market); study duration; and intervention characteristics.

For objective 1, we will extract data related to the determinants and hygiene behaviors measured, as well as any reported associations between them. For objectives 2 and 3, we will extract the description of the intervention (eg, health education); outcome indicator (how the intervention was assessed and how microbial contamination was measured); theory, model, and framework used (if applicable); duration of the intervention; type of contact with the intervention (eg, one-to-one counseling); frequency of intervention contact; primary and secondary outcome of the intervention; and outcomes related to microbial contamination. These fields were informed by our preliminary search and refined based on the PCC and BCD frameworks. We will collect and sort key information from the selected studies in a Microsoft Excel spreadsheet following a framework developed and piloted by the team and refined using an iterative approach. Data will be extracted and charted for quantitative and qualitative descriptive analyses. Data from eligible studies will be compiled by objective in the framework in Excel.

Data Analysis and Synthesis

Once the data extraction is finished, for objective 1, we will proceed to encode the determinants identified according to the definitions in the BCD checklist [54]. For objectives 2 and 3, we will use the Behavior Change Technique Taxonomy v1 framework by Michie et al [56] to identify and code behavior change techniques used in the included intervention studies. The Behavior Change Technique Taxonomy v1 is a framework designed to identify and categorize the smallest observable and

XSL•FO
RenderX

reproducible components used in intervention studies. Two reviewers will independently carry out the coding and categorization of the results. To validate the process, a third reviewer will cross-check a random 25% of classifications. In case of any disagreements, the reviewers will discuss their coding decisions and resolve any differing opinions by providing a rationale. To address potential heterogeneity and ensure a realistic scope, we will conduct data extraction, analysis, and quality appraisal separately for each objective. For instance, for objectives 2 and 3, we will present summary tables outlining intervention characteristics, duration, type, and frequency of contact, and intervention outcomes. If a sufficient number of homogeneous intervention types are not identified, a systematic review with meta-analysis may not be feasible. Data analysis will be conducted using a narrative synthesis method in accordance with all 3 objectives [46,57]. Quantitative data from the included studies will be summarized using descriptive statistics where appropriate, while qualitative findings will be analyzed thematically. The outcomes of the interventions for objectives 2 and 3 will be synthesized separately for each objective, enabling relevant comparisons between different interventions. We will map determinants and intervention characteristics, including types and components of interventions and reported outcomes (eg, knowledge, attitude, practices, and microbiological contamination). We will highlight gaps in the existing evidence to inform future research and intervention design.

Quality Appraisal

Gray literature will include unpublished reports, theses, and dissertations. To assess the quality of gray literature, we will use the Authority, Accuracy, Coverage, Objectivity, Date, and Significance checklist [58]. We will conduct a quality appraisal of the included studies. For objective 1, we will use the Joanna Briggs Institute critical appraisal tools according to the included study design [59]. For instance, the Joanna Briggs Institute cross-sectional checklist assesses 8 specific criteria, detailed in Table S6 in Multimedia Appendix 1. Each of the criteria will be rated as “yes” (1 point), “no” (0 points), “unclear” (0.5 points), and “not applicable (N/A).” For objectives 2 and 3, we will use a modified risk of bias (RoB) assessment tool to evaluate the potential biases across different study designs included in this review [60]. This tool was developed by the National Heart, Lung, and Blood Institute to examine controlled intervention and pre-post studies lacking a control group. The adapted RoB assessment tool comprises 14 specific criteria, detailed in Table S7 in Multimedia Appendix 1. Each of the 14 criteria will be assigned a numerical score: “yes”=1; “partially”=0.5 (where relevant); and “no,” “not applicable,” or “not reported”=0. In the context of each study, a score of 1 or “yes” will be considered as good quality, a score of 0.5 or “partially” or “unclear” will be considered as fair quality, and a score of 0 or “no” will be considered as poor quality. A study will be considered “good” when it has minimal bias, making it more likely that results are reliable and valid. A “fair” study acknowledges some bias but not enough to dismiss its findings. In contrast, a “poor” rating indicates a substantial RoB, raising concerns about the credibility of the results. The overall RoB for each study will be determined by adding up the individual scores assigned to each item. For studies under objective 1, the RoB index ranges from 0 to 8. For objectives 2 and 3, the RoB index ranges from 0 to 14. A score of 0 signifies the highest potential bias, whereas scores of 14 for objectives 2 and 3 and 8 for objective 1 represent the lowest potential bias. Two reviewers (AK and SMS) will individually assess the RoB in each study, and disagreements will be resolved with consensus between the 2 reviewers. After completing the quality assessment of the studies included for objective 1, we will present the quality ratings of the included articles alongside each determinant.

Consultation and Dissemination

According to the suggestion of Arksey and O’Malley [42] and Levac et al [43], we will engage hygiene behavior experts as consultants throughout the review process. Consultants will provide input regarding search strategies to integrate food hygiene behavior–related terms to capture relevant studies. They will also guide the formation of data charting tables and data analysis. Through the dissemination program, we will also share the findings of this scoping review with consultants and relevant stakeholders, which will help us gain insight beyond what is reported in the literature. This involvement will be crucial in obtaining insights that may not be captured solely by existing literature. We will invite stakeholders from government, nongovernment, and international organizations who are involved with food hygiene–related programs in Bangladesh.

Ethical Considerations

This study received ethical approval from the Ethical Review Committee of the International Centre for Diarrheal Disease Research, Bangladesh, and the approved research protocol number is PR-22143. Institutional review board Federal wide Assurance number: 00001468.

Results

The project received funding in October 2022 and institutional review board approval in December 2022. We have completed the database search, title and abstract screening, and full-text review. Initially, we identified 19,595 articles, and as of October 2025, 49 (0.3%) articles have been included for synthesis (Figure 2). At present, we are conducting data analysis and quality appraisal and anticipate completing the review by April 2026. This scoping review will map the determinants of street food hygiene and the interventions implemented for street food vendors and consumers in low-resource settings. The findings will provide insights into the key factors influencing street food hygiene in LICs and LMICs and support the design of feasible, sustainable, and scalable interventions targeting both vendors and consumers.

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart highlighting the study selection process for all 3 objectives.

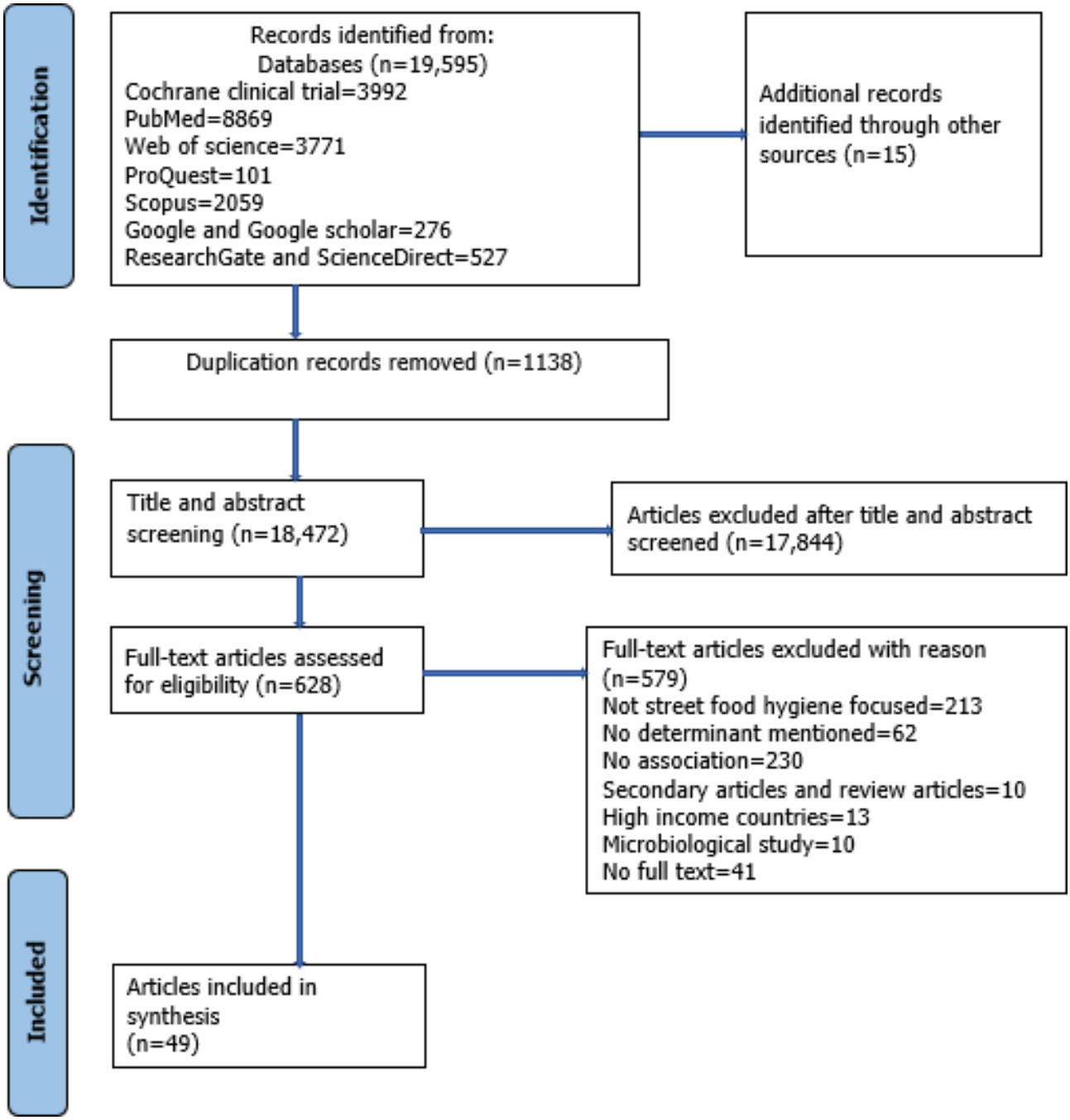

Discussion

The prevalence of foodborne disease is still high worldwide, especially in LICs and LMICs, where street food is popular because it is accessible and reasonably priced [61]. Despite this, synthesized evidence on behavioral determinants and effective hygiene interventions in these settings is lacking. There is an urgent need to inform context-relevant solutions, as repeated outbreaks of foodborne illness are linked to unsafe street food [2]. To help identify and organize behavioral determinants for this review, we took guidance from the BCD framework, which has been frequently used in water, sanitation, and hygiene research. The behavioral determinants outlined in the BCD framework include those related to the body (characteristics and senses); brain (executive, reactive, and motivated); environments (biological, social, and physical); and settings (stage, props, roles, routines, norms, and script) that influence individual behaviors at a given moment [54].

Several interconnected factors, such as vendor practices, environmental constraints, and consumer behavior, may influence street food hygiene [62]. This review methodology tackles a complex problem that needs in-depth investigation by considering the viewpoints of both vendors and consumers and looking at microbiological results in addition to behavior. This review will be the first evidence synthesis that examines street food hygiene determinants and interventions for vendors and consumers in LICs and LMICs [1]. It will also examine the existing evidence on the effectiveness of various interventions

XSL•FO
RenderX

implemented on street food vendors and consumers, facilitating a comparison of these interventions. The understanding of the determinants of street food hygiene in LICs and LMICs identified will inform the development and evaluation of appropriate, targeted street food hygiene interventions and policies. If the evidence on determinants is found suboptimal, the gap identified and disseminated will inform researchers to conduct future studies. It will also underscore the importance of designing appropriate research studies to address any knowledge gaps identified.

Acknowledgments

The International Centre for Diarrhoeal Disease Research, Bangladesh, is grateful to the governments of Bangladesh and Canada for their generous core and unrestricted funding.

Funding

This work was funded by the Reckett Global Hygiene Institute. The ideas expressed are those of the authors and not necessarily those of the Reckett Global Hygiene Institute.

Authors' Contributions

TMNH conceived of the idea and developed the initial research question and study methods. TMNH and AK conceived the idea of the manuscript. TMNH, AK, EDT, PJW, OPG, FS, JS, RHS, and SMS formulated the research question and determined the inclusion and exclusion criteria. TMNH and AK drafted the initial protocol manuscript together with SMS and SIR. AK, SMS, SIR, and TMNH devised the data extraction plan with input from all coauthors. EDT, FS, JS, RHS, OPG, PJW, and MR provided critical feedback on the study design and data extraction plan. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

It provides detailed supplementary materials supporting this scoping review protocol. It includes the complete PubMed search strategies used for objectives 1-3 (Tables S1-S3), operational definitions of food hygiene behaviors and determinants adapted from the behavior-centered design framework (Tables S4-S5), and the Joanna Briggs Institute (JBI) checklist for cross-sectional studies and the risk of bias assessment tool (Tables S6-S7).

[DOCX File, 46 KB - resprot_v15i1e68633_app1.docx]

References

1. Winarno FG, Allain A. Street foods in developing countries: lessons from Asia. Pascal and Francis Bibliographic Databases. URL: https://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=3749119 [accessed 2026-03-04]
2. Rane S. Street vended food in developing world: hazard analyses. Indian J Microbiol 2011 Jan;51(1):100-106. [doi: 10.1007/s12088-011-0154-x] [Medline: 22282636]
3. Vollaard AM, Ali S, van Asten HAGH, et al. Risk factors for transmission of foodborne illness in restaurants and street vendors in Jakarta, Indonesia. Epidemiol Infect 2004 Oct;132(5):863-872. [doi: 10.1017/s0950268804002742] [Medline: 15473149]
4. Birgen BJ, Njue LG, Kaindi DM, Ogutu FO, Owade JO. Determinants of microbial contamination of street-vended chicken products sold in Nairobi County, Kenya. Int J Food Sci 2020 Feb 14;2020:2746492. [doi: 10.1155/2020/2746492] [Medline: 32149074]
5. Khadka S, Adhikari S, Rai T, et al. Bacterial contamination and risk factors associated with street-vended panipuri sold in Bharatpur, Nepal. Int Food Res J 2018;5(3):32-38 [FREE Full text]
6. Mugampoza D, Byarugaba GW, Nyonyintono A, Nakitto P. Occurrence of Escherichia coli and Salmonella spp. in street-vended foods and general hygienic and trading practices in Nakawa Division, Uganda. Am J Food Nutr 2013;3:167-175 [FREE Full text]
7. Derbew G, Sahle S, Endris M. Bacteriological assessment of some street vended foods in Gondar, Ethiopia. Internet J Food Saf 2013;15:33-38 [FREE Full text]
8. Agarwal A, Awasthi V, Dua A, Ganguly S, Garg V, Marwaha SS. Microbiological profile of milk: impact of household practices. Indian J Public Health 2012;56(1):88-94. [doi: 10.4103/0019-557X.96984] [Medline: 22684182]
9. Marwaha P, Pathak S, Alpana S. Bacterial profile of street vended panipuri from different zones of Jabalpur city of MP, India. Int J Res Biosci Agric Technol 2018;VI:132-142 [FREE Full text]
10. Tambekar DH, Kulkarni RV, Shirsat SD, Bhadange DG. Bacteriological quality of street vended food panipuri: a case study of Amravati city (MS) India. Biosci Discov 2011;2(3):350-354 [FREE Full text]

XSL•FO
RenderX

11. Reddi SL. Microbiological quality of street vended fruit juices in Hyderabad, India and their association between food safety knowledge and practices of fruit juice vendors. Int J Curr Microbiol Appl Sci 2015;4(1):970-982 [FREE Full text]
12. Orlando D, Hardianti MS, Setyabudi FM. Mounting an effective response to an outbreak of viral disease involving street food vendors in Indonesia. In: Hoorfar J, editor. Case Studies in Food Safety and Authenticity: Woodhead Publishing; 2012:161-167. [doi: 10.1533/9780857096937.3.161]
13. Pinto M, Fatima A, dos Reis A, et al. An outbreak of suspected food poisoning after eating street food, Dili, Timor-Leste, July 2018. N Terr Dis Control Bull 2018;25(3):17-18 [FREE Full text]
14. Hutin Y, Luby S, Paquet C. A large cholera outbreak in Kano City, Nigeria: the importance of hand washing with soap and the danger of street-vended water. J Water Health 2003 Mar;1(1):45-52. [Medline: 15384272]
15. Kunwor I, Basnet A. Association of multi-drug resistant bacteria with sanitation of street vendors food. Int J Environ Agric Biotechnol 2022;7(6):192-205. [doi: 10.22161/ijeab.76.22]
16. Daniels NA, MacKinnon L, Rowe SM, Bean NH, Griffin PM, Mead PS. Foodborne disease outbreaks in United States schools. Pediatr Infect Dis J 2002 Jul;21(7):623-628. [doi: 10.1097/00006454-200207000-00004] [Medline: 12237592]
17. Narain A, Janani, Narayanaswamy K. Microbiological quality of selected street-vended foods in Coimbatore, India. Afr J Microbiol Res 2015;9(11):757-762. [doi: 10.5897/AJMR2013.6421]
18. Islam S, Nasrin N, Rizwan F, et al. Microbial contamination of street vended foods from a university campus in Bangladesh. Southeast Asian J Trop Med Public Health 2015 May;46(3):480-485. [Medline: 26521522]
19. WHO estimates of the global burden of foodborne diseases: foodborne disease burden epidemiology reference group 2007-2015. : World Health Organization; 2015 URL: https://www.who.int/publications/i/item/9789241565165 [accessed 2026-02-21]
20. Prüss-Ustün A, Bartram J, Clasen T, et al. Burden of disease from inadequate water, sanitation and hygiene in low- and middle-income settings: a retrospective analysis of data from 145 countries. Trop Med Int Health 2014 Aug;19(8):894-905. [doi: 10.1111/tmi.12329] [Medline: 24779548]
21. Makinde OM, Ayeni KI, Sulyok M, Krska R, Adeleke RA, Ezekiel CN. Microbiological safety of ready-to-eat foods in low- and middle-income countries: a comprehensive 10-year (2009 to 2018) review. Compr Rev Food Sci Food Saf 2020 Mar;19(2):703-732. [doi: 10.1111/1541-4337.12533] [Medline: 33325184]
22. Bloem JR, Farris J. The COVID-19 pandemic and food security in low- and middle-income countries: a review. Agric Food Secur 2022;11(1):55. [doi: 10.1186/s40066-022-00391-4] [Medline: 36474782]
23. Arif I, Shourove JH, Sarker T, Rahman MM, Islam GMR. Food purchase and consumption behavior during the COVID-19 pandemic lockdown in Bangladesh: association between sociodemographic composition. BMC Public Health 2024 Oct 24;24(1):2943. [doi: 10.1186/s12889-024-19982-w] [Medline: 39443953]
24. Martínez L, Young G. Street vending, vulnerability and exclusion during the COVID-19 pandemic: the case of Cali, Colombia. Environ Urban 2022 Oct;34(2):372-390. [doi: 10.1177/09562478221113753] [Medline: 36254215]
25. Pitoyo AJ, Aditya B, Amri I. The impacts of COVID-19 pandemic to informal economic sector in Indonesia: theoretical and empirical comparison. Presented at: 1st Geosciences and Environmental Sciences Symposium (ICST 2020); Sep 7-8, 2020. [doi: 10.1051/e3sconf/202020003014]
26. Horodnic IA, Williams CC, ugulea O, Stoian Bobâlcă IC. Exploring the demand-side of the informal economy during the COVID-19 restrictions: lessons from Ia i, Romania. Sustainability 2022;14(12):7228. [doi: 10.3390/su14127228]
27. Wiatrowski M, Czarniecka-Skubina E, Trafiałek J. Consumer eating behavior and opinions about the food safety of street food in Poland. Nutrients 2021 Feb 11;13(2):594. [doi: 10.3390/nu13020594] [Medline: 33670190]
28. Riyanto A, Murwani R, Sulistiyani S, Rahfiludin M. Food safety education using book covers and videos to improve street food safety knowledge, attitude, and practice of elementary school students. Curr Res Nutr Food Sci J 2017 Aug 18;5(2):116-125. [doi: 10.12944/CRNFSJ.5.2.08]
29. Singh AK, Dudeja P, Kaushal N, Mukherji S. Impact of health education intervention on food safety and hygiene of street vendors: a pilot study. Med J Armed Forces India 2016 Jul;72(3):265-269. [doi: 10.1016/j.mjafi.2016.02.008] [Medline: 27546967]
30. Wickrematilake MS, Arnold SM, Karthikeyan P, Jayathilake AD. Improving knowledge of street food vendors through an educational intervention in Kandy district, Sri Lanka. Int J Res Med Sci 2022;10(2):331-335. [doi: 10.18203/2320-6012.ijrms20220276]
31. Daniele G, Mookerjee S, Tommasi D. Informational shocks and street-food safety: a field study in urban India. Institute of Labor Economics. 2020. URL: https://www.econstor.eu/bitstream/10419/227204/1/dp13677.pdf [accessed 2026-02-21]
32. Yu H, Ma J, Liu P, Sirsat SA. Investigating the effects of explanatory-based food safety training: a model of domain knowledge theory perspective. Int J Hosp Manag 2020 Oct;91(8):102688. [doi: 10.1016/j.ijhm.2020.102688]
33. Curtis V, Dreibelbis R, Buxton H, Izang N, Adekunle D, Aunger R. Behaviour settings theory applied to domestic water use in Nigeria: a new conceptual tool for the study of routine behaviour. Soc Sci Med 2019 Aug;235:112398. [doi: 10.1016/j.socscimed.2019.112398] [Medline: 31326766]
34. Tidwell JB, Chipungu J, Chilengi R, Curtis V, Aunger R. Using a theory-driven creative process to design a peri-urban on-site sanitation quality improvement intervention. BMC Public Health 2019 May 14;19(1):565. [doi: 10.1186/s12889-019-6898-7] [Medline: 31088432]

XSL•FO
RenderX

35. White S, Thorseth AH, Dreibelbis R, Curtis V. The determinants of handwashing behaviour in domestic settings: an integrative systematic review. Int J Hyg Environ Health 2020 Jun;227:113512. [doi: 10.1016/j.ijheh.2020.113512] [Medline: 32220763]
36. Umar AA, Mande AT, Umar J. The effect of food hygiene training among street food vendors in Sabon Gari Local Government Area of Kaduna State, Nigeria. Sub Saharan Afr J Med 2018;5(1):20-28. [doi: 10.4103/ssajm.ssajm_30_17]
37. Insfran-Rivarola A, Tlapa D, Limon-Romero J, et al. A systematic review and meta-analysis of the effects of food safety and hygiene training on food handlers. Foods 2020 Aug 25;9(9):1169. [doi: 10.3390/foods9091169] [Medline: 32854221]
38. Young I, Greig J, Wilhelm BJ, Waddell LA. Effectiveness of food handler training and education interventions: a systematic review and meta-analysis. J Food Prot 2019 Oct;82(10):1714-1728. [doi: 10.4315/0362-028X.JFP-19-108] [Medline: 31536416]
39. Young I, Thaivalappil A, Greig J, Meldrum R, Waddell L. Explaining the food safety behaviours of food handlers using theories of behaviour change: a systematic review. Int J Environ Health Res 2018 Jun;28(3):323-340. [doi: 10.1080/09603123.2018.1476846] [Medline: 29852750]
40. Young I, Waddell LA, Wilhelm BJ, Greig J. A systematic review and meta-regression of single group, pre-post studies evaluating food safety education and training interventions for food handlers. Food Res Int 2020 Feb;128:108711. [doi: 10.1016/j.foodres.2019.108711] [Medline: 31955782]
41. Choudhury M, Mahanta LB, Goswami JS, Dutta Mazumder M. Will capacity building training interventions given to street food vendors give us safer food?: A cross-sectional study from India. Food Control 2011 Aug;22(8):1233-1239. [doi: 10.1016/j.foodcont.2011.01.023]
42. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
43. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
44. Daudt HM, van Mossel C, Scott SJ. Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. BMC Med Res Methodol 2013 Mar 23;13:48. [doi: 10.1186/1471-2288-13-48] [Medline: 23522333]
45. Khalil H, Campbell F, Danial K, et al. Advancing the methodology of mapping reviews: a scoping review. Res Synth Methods 2024 May;15(3):384-397. [doi: 10.1002/jrsm.1694] [Medline: 38169156]
46. Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI manual for evidence synthesis. JBI. 2024. URL: https://jbi-global-wiki.refined.site/space/MANUAL [accessed 2026-02-21]
47. World Bank country and lending groups. The World Bank. URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [accessed 2026-02-21]
48. Nizame FA, Alam MU, Masud AA, et al. Hygiene in restaurants and among street food vendors in Bangladesh. Am J Trop Med Hyg 2019 Sep;101(3):566-575. [doi: 10.4269/ajtmh.18-0896] [Medline: 31333161]
49. Kumwenda S. Challenges to hygiene improvement in developing countries. In: Potgieter N, Traore AN, editors. The Relevance of Hygiene to Health in Developing Countries: IntechOpen Publisher; 2019. [doi: 10.5772/intechopen.72024]
50. Mazi IM, Onyeaka H, Akegbe H, et al. Street vended foods in Nigeria: an analysis of the current state of affairs and the way forward. Cogent Food Agric 2023;9(2). [doi: 10.1080/23311932.2023.2266194]
51. Ajayi A, Oluwoye J. Sustainable street vended foods and food safety: a conceptual framework. Int J Food Saf Nutr Public Health 2015;5(3/4):195. [doi: 10.1504/IJFSNPH.2015.070154]
52. Salamandane A, Malfeito-Ferreira M, Brito L. The socioeconomic factors of street food vending in developing countries and its implications for public health: a systematic review. Foods 2023 Oct 14;12(20):3774. [doi: 10.3390/foods12203774] [Medline: 37893667]
53. Biran A, Curtis V, O Gautam P, et al. Background paper on measuring WASH and food hygiene practices: definition of goals to be tackled post 2015 by the joint monitoring programme. London School of Hygiene and Tropical Medicine. 2012. URL: https://www.ircwash.org/resources/background-paper-measuring-wash-and-food-hygiene-practices-definition-goals-be-tackled [accessed 2026-02-21]
54. Aunger R, Curtis V. A Guide to Behaviour Centred Design. : London School of Hygiene and Tropical Medicine URL: https://www.lshtm.ac.uk/sites/default/files/2017-03/BCD%20Guide.pdf [accessed 2026-03-04]
55. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
56. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
57. McKenzie JE, Brennan SE, Ryan RE, Thomson HJ, Johnston RV. Chapter 9: Summarizing study characteristics and preparing for synthesis. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions: The Cochrane Collaboration; 2019:229-240.
58. Tyndall J. AACODS checklist. Flinders University. 2010. URL: https://coilink.org/20.500.12592/kg54zf [accessed 2026-02-21]

XSL•FO
RenderX

59. Moola S, Munn Z, Tufanaru C, et al. Chapter 7: Systematic reviews of etiology and risk. In: Aromatis E, Munn Z, editors. JBI Manual for Evidence Synthesis: JBI; 2020. [doi: 10.46658/JBIRM-17-06]
60. Study quality assessment tools. National Institutes of Health National Heart, Lung, and Blood Institute. 2021. URL: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools [accessed 2026-02-21]
61. Andrade AA, Paiva AD, Machado AB. Microbiology of street food: understanding risks to improve safety. J Appl Microbiol 2023 Aug 1;134(8):lxad167. [doi: 10.1093/jambio/lxad167] [Medline: 37516449]
62. Mwove J, Orwa JD, Njoki JW, Irungu F, Kiruki FG. Factors influencing consumer choices and food safety awareness in street-vended foods: a case study of Meru town, Kenya. J Food Nutr Diet Sci 2024;2(1):94-106. [doi: 10.55976/fnds.22024125794-106]

Abbreviations

BCD: Behavior-Centered Design
LIC: low-income country
LMIC: low- and middle-income country
PCC: population, concept, and context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
RoB: risk of bias

Edited by J Sarvestan; submitted 14.Nov.2024; peer-reviewed by H Kitamura, O Obadimu; revised version received 20.Oct.2025; accepted 24.Oct.2025; published 09.Mar.2026.

Please cite as:
Kamal A, Nurul Huda TM, Simi SM, Thomas ED, Sultana F, Sultana J, Sarah RH, Rafa SI, Gautam OP, Winch PJ, Rahman M
Determinants of Street Food Hygiene Practices and the Effectiveness of Interventions for Vendors and Consumers in Low- and Middle-Income Countries: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e68633
URL: https://www.researchprotocols.org/2026/1/e68633
doi:10.2196/68633

© Abul Kamal, Tarique Md Nurul Huda, Sonjida Mesket Simi, Elizabeth D Thomas, Farhana Sultana, Jesmin Sultana, Rehnuma Haque Sarah, Soha Ishrak Rafa, Om Prasad Gautam, Peter John Winch, Mahbubur Rahman. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 9.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Inequity in Access to and Use of Digital Health Technologies in Routine Heart Failure Care: Protocol for a Scoping Review

Nicklas Vinter[1,2], MD, PhD; Mariam Elmegaard[3], MD; Lars Køber[4], MD, PhD, DMSc; Morten Schou[3], MD, PhD; Søren Paaske Johnsen[1], MD, PhD

[1]Danish Center for Health Services Research, Department of Clinical Medicine, Aalborg University, Selma Lagerløfs Vej 249, Aalborg, Denmark

[2]Department of Cardiology, Regional Hospital Central Jutland, Viborg, Denmark

[3]Department of Cardiology, Herlev and Gentofte University Hospital, Copenhagen, Denmark

[4]Department of Cardiology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Corresponding Author:
Nicklas Vinter, MD, PhD
Danish Center for Health Services Research, Department of Clinical Medicine, Aalborg University, Selma Lagerløfs Vej 249, Aalborg, Denmark

Abstract

Background: Heart failure (HF) is a global health challenge characterized by high mortality, morbidity, and economic burden. The development of digital health technologies offers promising tools for prevention, early detection, and management of HF, potentially improving prognoses and reducing costs. However, these innovations may also widen existing health disparities related to socioeconomic status, geography, and race/ethnicity.

Objective: This scoping review will examine and map existing evidence on socioeconomic, geographic, and racial/ethnic differences in access to and use of digital health technologies for HF care in routine practice.

Methods: The writing of this protocol followed the methodological framework by Arksey and O'Malley, including (1) identifying the research question; (2) identifying relevant studies; (3) selecting studies to be included in the review; (4) charting the data; and (5) collating, summarizing, and reporting the results. Eligible studies must examine digital health technologies in adults (aged ≥18 years) with any type of HF and report on social determinants of health, geography, or race/ethnicity. Observational study designs will be included. Searches will be conducted in Embase, PubMed, Google Scholar, and Scopus. A 2-stage screening process will determine study eligibility, and data will be extracted using a standardized form.

Results: The project is funded. Data collection is expected to begin by the beginning of 2026.

Conclusions: This scoping review will map existing evidence on differences in access to and use of digital health technologies for HF care. The findings are anticipated to highlight patterns and gaps in the literature, informing future research and strategies for equitable implementation.

Trial Registration: Open Science Framework ch2gp; https://osf.io/psyhv/

International Registered Report Identifier (IRRID): PRR1-10.2196/81949

(JMIR Res Protoc 2026;15:e81949) doi:10.2196/81949

KEYWORDS

heart failure; digital health technology; routine practice; inequity; review

Introduction

Heart failure (HF) is a complex clinical syndrome that represents a considerable global health burden, characterized by high mortality, morbidity, and costs [1,2]. The development of digital health technologies for HF is expanding, supporting primary prevention, early detection, and disease management [3,4], and recent reviews highlight the growing role of remote monitoring, wearable technologies, and integrated care models [5-8]. Application of digital health tools in health care services and decision-making has the potential to improve patient care and clinical outcomes and reduce health care expenditures. However, comprehensive investigations are needed to understand the broader implications, including ethical, practical, and systemic concerns, associated with implementation.

Understanding the consequences for health equity of implementing digital health technologies is essential [9]. Health inequities based on socioeconomic status, race/ethnicity, and geography remain a significant challenge in HF care. Inequities can lead to unequal health outcomes, with certain groups experiencing higher rates of hospitalization, poorer quality of life, and increased risk of adverse clinical outcomes [10-14].

XSL•FO
RenderX

While digital health innovations may offer new opportunities for care delivery, uneven access and use across population groups could exacerbate existing inequities. Recent work emphasizes the need for equity frameworks in digital health implementation [15].

Without deliberate efforts to ensure equitable implementation, the digital transformation of health care risks reinforcing or even widening inequities [16,17], yet the extent and nature of evidence available in the literature regarding patients with HF remain unclear. To address this gap, we will conduct a scoping review of digital health technologies used in HF care. The objective of this scoping review is to map and synthesize what is known from existing literature about socioeconomic, geographic, and racial/ethnic differences in access to and use of digital technologies in routine clinical practice among individuals with HF. We hypothesize that individuals from underserved socioeconomic groups, rural or remote areas, or racial/ethnic minority groups have lower access to and use of digital health technologies.

Methods

Overview

The approach follows the methodological framework by Arksey and O'Malley [18] and incorporates refinements proposed by Levac et al [19] and guidance from the Joanna Briggs Institute (JBI) for scoping reviews to enhance methodological rigor [20]. We will adhere to the PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist when reporting the review [21]. This protocol was preregistered at Open Science Framework [22]. Figure 1 illustrates the overall workflow of the scoping review.

Figure 1. Overview of the scoping review workflow. HF: heart failure.

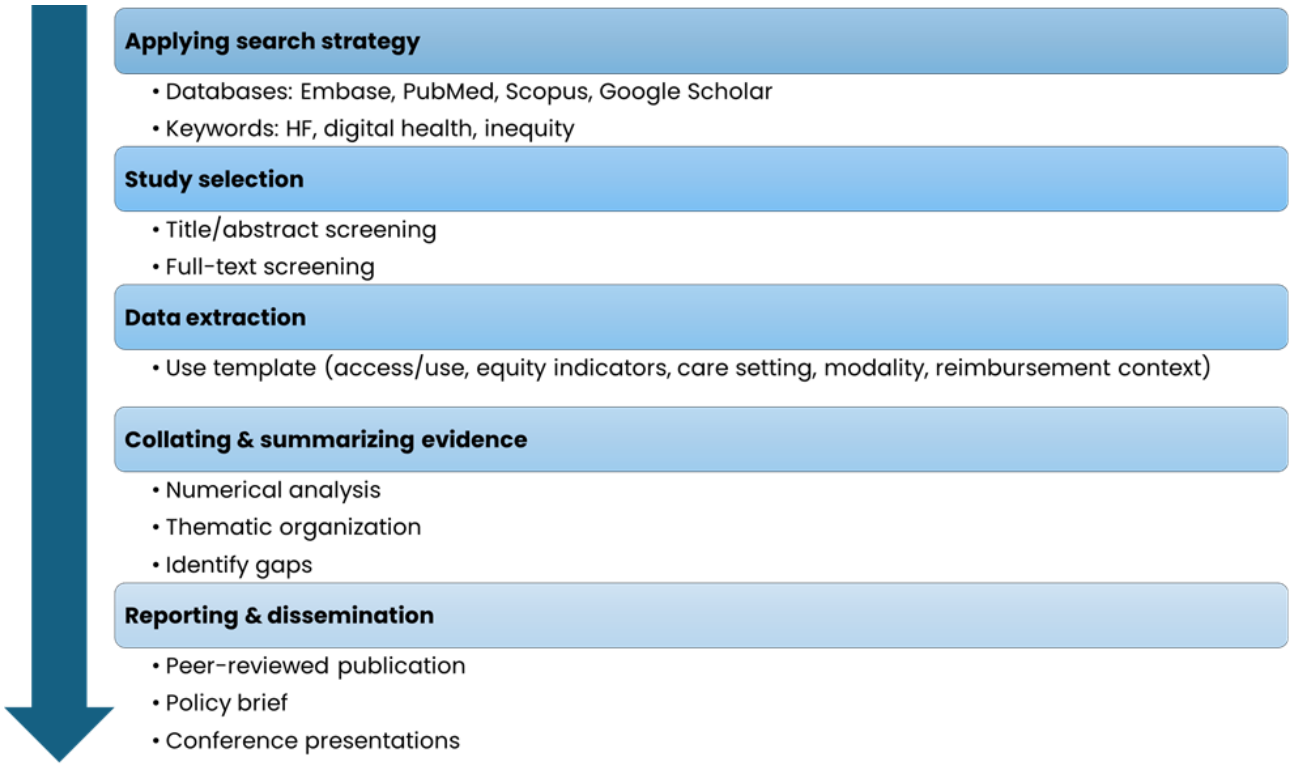

Eligibility Criteria

Textbox 1 summarizes the inclusion and exclusion criteria. We will include studies on individuals aged ≥18 years with any type of HF that report on any digital health technology. We will use the World Health Organization (WHO) definition of "digital health intervention," which states that such an intervention involves the use of digital technology to achieve specific health-related objectives, and that it is integrated into digital health applications and information and communication technology systems [9]. Data on social determinants of health, geography, and race/ethnicity must also be present in the studies. The selection of the social determinants of health was inspired by previous studies on HF (Table 1) [23]. Any study with an observational design is eligible, because this design reflects real-world access and use patterns in routine practice. Randomized controlled trials and qualitative-only studies are excluded as they often occur in controlled settings that may mask real-world inequities. Eligible studies must provide data on access to or use of the digital health technology. We will place no restriction on the calendar year of publication. Finally, we will include only full-text articles published in English in peer-reviewed journals. The restriction is due to resource limitations for translation and to ensure methodological rigor. We acknowledge that this may introduce language and publication bias, which will be discussed as a limitation.

XSL•FO
RenderX

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Adults (≥18 years) with any type of heart failure
- Studies reporting on digital health technologies (World Health Organization definition)
- Observational studies (cross-sectional, cohort, case-control)
- Studies that report social determinants of health, geography, or race/ethnicity
- Full-text articles published in English in peer-reviewed journals
- Any publication year

Exclusion criteria

- Randomized controlled trials, qualitative-only studies
- Conference abstracts, editorials, letters, non–peer-reviewed sources
- Non–English language publications

Table . Domains of social drivers used to examine inequity.

Data domain	Definition
Social determinants of health	• Insurance • Income/wealth • Marital status • Educational attainment • Neighborhood socioeconomic status • Social support • Employment status • Health literacy
Geography	• Urban/rural residency • Distance to hospital • Administrative regions
Race and ethnicity	• Any group

Search Strategy and Information Sources

A professional health care librarian was consulted to build the search strategy (Table 2). Our search strategy combines HF and digital health technology, and then either inequity, social determinants of health, geography, or race/ethnicity. Databases to be searched include Embase, PubMed, and the interdisciplinary sources Google Scholar and Scopus.

XSL•FO
RenderX

Table . Search strategy. Search query: number 1 AND number 2 AND (number 3 OR number 4 OR number 5 OR number 6).

Number	Topic	Database			
		Embase	PubMed	Google Scholar	Scopus
1	Heart failure	"heart failure"/exp OR "heart failure"	"Heart Failure" [Mesh] OR "heart failure"	"heart failure"	"heart failure"
2	Digital health intervention	("digital intervention"/exp OR "digital intervention" OR ("digital" AND ("intervention"/exp OR "intervention"))) OR ("digital health"/exp OR "digital health") OR ("mobile application"/exp) ("information technology device"/exp) OR ("digital technology"/exp) OR ("telehealth"/exp) OR ("mhealth"/exp) OR "mhealth" OR "mobile health"	"Telecommunications" [Mesh] OR "digital intervention" OR "digital health intervention" OR ("digital" AND "intervention") OR "mHealth" OR "mobile health"	"digital intervention" OR "mHealth" OR "mobile health"	"digital intervention" OR ("digital" AND "intervention") OR "mHealth" OR "mobile health"
3	Health inequity	"health disparity"/exp OR "disparity in health" OR "health disparities" OR "health economic disparity" OR "health inequality" OR "health inequities" OR "health inequity" OR "health social disparity" OR "health social economic disparity" OR "health social inequality" OR "health socio-economic disparity" OR "health socioeconomic disparity" OR "health socioeconomic inequity" OR "health status disparities" OR "health status disparity" OR "health status inequality" OR "health status inequity" OR "inequality in health" OR "inequity in health" OR "socioeconomic disparities in health" OR "health disparity"	"Health Inequities" [Mesh] OR "inequity" OR "inequities" OR "disparity" OR "disparities"	"inequity" OR "disparity"	"inequity" OR "inequities" OR "disparity" OR "disparities"
4	Social determinants of health	"social determinants of health"/exp OR "social determinant" OR "social determinants" OR "social determining factor" OR "social factors determining health" OR "social health determinant" OR "social determinants of health"	"Social Determinants of Health" [Mesh] OR "social determinant of health"	"Social Determinants of Health"	"Social Determinants of Health" OR "social determinant of health"

XSL•FO
RenderX

Number	Topic	Database			
		Embase	PubMed	Google Scholar	Scopus
5	Geographic variation	"geography"/exp OR "geographic factor" OR "geographic locations" OR "geography" OR "geography"/exp OR "geographic factor" OR "geographic locations" OR "geography" OR "geographic distribution"/exp OR "distribution, geographic" OR "geographical distribution" OR "geographic distribution" OR "geographic variation"/exp OR "urban area"/exp OR "built-up area" OR "built-up land" OR "urban environment" OR "urban land" OR "urbanized area" OR "urbanized environment" OR "urbanized land" OR "urbanized area" OR "urbanized environment" OR "urbanized land" OR "urban area" OR "rural area"/exp OR "agricultural area" OR "rural environment" OR "rural land" OR "rural area"	"Population" [Mesh] OR "geographic variation" OR "geography"	—[a]	"geography" OR "geographic variation" OR "rural" OR "urban"
6	Race/ ethnicity	"ethnic group"/exp OR "ethnic and racial groups" OR "ethnic and racial minorities" OR "ethnic groups" OR "ethnic minorities" OR "ethnic minority" OR "ethnic origin" OR "ethnic population" OR "ethnic status" OR "ethno-linguistic group" OR "ethnolinguistic group" OR "ethnic group"	"Ethnicity" [Mesh] OR "ethnicity" OR "ethnic" OR "Racial Groups"[Mesh] OR "race"	—	"ethnicity" OR "ethnic" OR "race" OR "racial"

[a]Not applicable.

Study Selection

Two independent reviewers (NV and ME) will use a 2-step screening approach. First, all identified records will be manually screened for eligibility based on titles and abstracts. Second, the full-text reports of the qualifying abstracts will be assessed for eligibility. Finally, the 2 reviewers will screen the list of references of the selected studies. Any inconsistencies will be resolved by formal consensus, which will involve SPJ.

Collecting and Charting Data

Based on the identified literature, NV and ME will independently extract data following the data charging form: (1) authors and year of publication (2); study design and population; (3) aims of the study (4); study location and calendar period; (5) care setting (eg, inpatient, outpatient, or home-based); (6) digital health technology modality; (eg, telemonitoring, mobile app, or wearable); (7) operational definitions of "access" (availability and ability to obtain or use technology) and "use" (actual engagement or use); (8) equity indicators (income, education, race/ethnicity, geography, and insurance status); (9) reimbursement and connectivity context (insurance coverage, internet access, and device ownership); (10) primary outcome measure; (11) main findings, reported according to age, sex, socioeconomic characteristics, geographic characteristics, and race/ethnicity.

XSL•FO
RenderX

Collating, Summarizing, and Reporting the Results

We will systematically organize and present the findings.

Numerical Analysis

Reported measures of association (eg, odds ratios, risk ratios, and hazard ratios) will be tabulated and organized by age, sex, social determinants, geography, and race/ethnicity to examine patterns of inequity. Relevant measures of association may be, for example, high-income vs low-income or urban vs rural residency. We will consider differences in health outcomes as health inequities if they are systematic, avoidable, unnecessary, unfair, and unjust [24]. No meta-analysis or pooled effect estimates will be performed; findings will be summarized narratively.

Thematic Organization

Data from each identified study will be reported by type of digital intervention, which will follow the WHO classification. The WHO classification includes interventions for patients with HF (referred to as clients by the WHO), interventions for health care providers, interventions for health systems or service managers, and interventions for data services [25]. The relevance of the WHO classification system lies in its ability to link interventions with a list of health system challenges, thereby illustrating how technology may address health care needs. Furthermore, it provides easy access to results for researchers planning new randomized controlled trials. Data charting will follow the iterative process recommended by the JBI. We will first pilot the data charting form on a small sample of included studies to ensure all relevant variables (eg, equity indicators and digital modality) are captured. Next, we will refine the form by adding or adjusting fields as new concepts emerge during extraction.

Identification of Research Gaps

We will highlight areas where research is lacking or where findings are inconsistent. Stakeholder consultation will be integrated during interpretation of findings to validate relevance and identify gaps. This will involve engaging clinicians only, who will review preliminary findings and provide input on practical implications.

Methodological Appraisal

To contextualize our interpretation of inequities, we will include a structured appraisal of key methodological features using a checklist adapted from JBI guidance [20]. The appraisal will be descriptive and will not involve scoring or exclusion. The checklist will include the following domains:

1. Study design
2. Sample size adequacy
3. Completeness of equity-related reporting (eg, socioeconomic status, geography, race/ethnicity)
4. Clarity of definitions for access and use

Findings from this appraisal will be summarized narratively to highlight major limitations and inform interpretation.

Results

The project is funded. Data collection is expected to begin by the beginning of 2026. A PRISMA 2020 flow diagram will be included in the final review to illustrate the study selection process. Figure 2 shows a blank version of the flowchart, which will be updated once screening and inclusion are completed.

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flowchart.

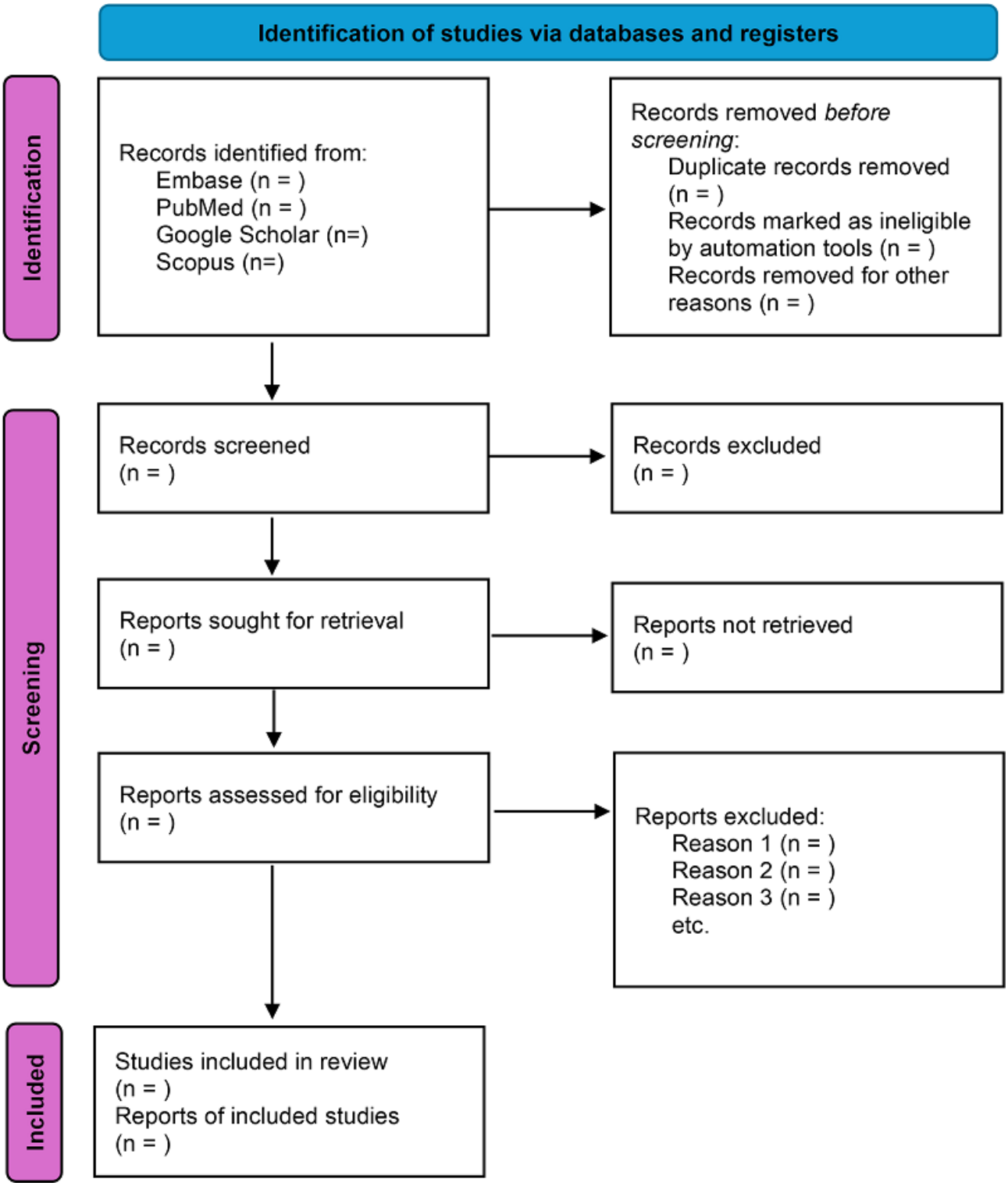

Discussion

Anticipated Main Findings

This scoping review will map existing evidence on socioeconomic, geographic, and racial/ethnic differences in access to and use of digital health technologies for HF care. By synthesizing patterns across observational studies, the review aims to identify where inequities may exist and highlight gaps in the literature. The findings will inform future research and guide strategies for equitable implementation of digital health solutions in routine practice.

Future Directions

The review will identify patterns and gaps in the evidence base, informing the design of future studies and interventions aimed at reducing inequities in digital health adoption. Findings may also support policymakers and health care organizations in developing strategies to ensure equitable access to digital technologies for HF care.

XSL•FO
RenderX

Limitations

A key limitation of this review is the restriction to English-language, peer-reviewed full texts, which may introduce language and publication bias. The restriction is particularly relevant for equity-focused reviews, and future research should consider multilingual searches and inclusion of gray literature to minimize bias.

Dissemination Plan

Results will be disseminated through a peer-reviewed publication, policy briefs, and presentations at conferences and stakeholder meetings, ensuring accessibility to researchers, clinicians, and decision-makers.

Conclusion

The review will synthesize evidence on disparities in digital health adoption for HF care, offering insights to inform policy, practice, and research agendas aimed at reducing inequities. Findings will highlight gaps in current knowledge and support the design of interventions that promote inclusive digital health strategies.

Acknowledgments

We would like to express our appreciation to health care librarian Mette Buje Grundsøe for her expertise and knowledge in developing the search strategies.

Funding

This work was supported by the Novo Nordisk Foundation.

Data Availability

No data are currently associated with this article, and the results of the planned review have not yet been generated. This protocol outlines the intended objectives, methodology, and analytical approach as registered at the time of submission.

Authors' Contributions

ME, LK, and MS developed the hypotheses. NV and SPJ conceptualized and designed the study. NV wrote the first and successive drafts of the protocol. NV collaborated with the health librarian to develop the search strategies. All authors contributed to critical revision of the protocol for important intellectual content.

Conflicts of Interest

NV is supported by a grant from the Danish Cardiovascular Academy (PD2Y-2022002-DCA), which is funded by the Novo Nordisk Foundation (grant NNF20SA0067242) and the Danish Heart Foundation, and has served as an advisory board member and consultant for AstraZeneca. No fees were received personally. LK declares receiving speaker's honoraria from Astra Zeneca, Boehringer, Novartis, and Novo. MS declares receiving lecture fees from Novartis, Bayer, Novo, Astra Zeneca, and Bohringer outside the present work.

References

1. Bragazzi NL, Zhong W, Shu J, et al. Burden of heart failure and underlying causes in 195 countries and territories from 1990 to 2017. Eur J Prev Cardiol 2021 Dec 29;28(15):1682-1690. [doi: 10.1093/eurjpc/zwaa147] [Medline: 33571994]
2. Virani SS, Alonso A, Aparicio HJ, et al. Heart disease and stroke statistics-2021 update: a report from the American Heart Association. Circulation 2021 Feb 23;143(8):e254-e743. [doi: 10.1161/CIR.0000000000000950] [Medline: 33501848]
3. Cowie MR, McBeath KCC, Angermann CE. The digital future of heart failure care. Curr Heart Fail Rep 2022 Jun;19(3):109-113. [doi: 10.1007/s11897-022-00547-0] [Medline: 35441250]
4. Farwati M, Riaz H, Tang WHW. Digital health applications in heart failure: a critical appraisal of literature. Curr Treat Options Cardiovasc Med 2021;23(2):12. [doi: 10.1007/s11936-020-00885-z] [Medline: 33488049]
5. Azizi Z, Broadwin C, Islam S, et al. Digital health interventions for heart failure management in underserved rural areas of the United States: a systematic review of randomized trials. J Am Heart Assoc 2024 Jan 16;13(2):e030956. [doi: 10.1161/JAHA.123.030956] [Medline: 38226517]
6. Noci F, Capodici A, Nuti S, Passino C, Emdin M, Giannoni A. Wearable technologies to predict and prevent and heart failure hospitalizations: a systematic review. Eur Heart J Digit Health 2025 Sep;6(5):868-877. [doi: 10.1093/ehjdh/ztaf079] [Medline: 40984981]
7. Kallas D, Sandhu N, Gandilo C, et al. Use of digital health technology in heart failure and diabetes: a scoping review. J Cardiovasc Transl Res 2023 Jun;16(3):526-540. [doi: 10.1007/s12265-022-10273-6] [Medline: 35639339]
8. Zwack CC, Haghani M, Hollings M, et al. The evolution of digital health technologies in cardiovascular disease research. NPJ Digit Med 2023 Jan 3;6(1):1. [doi: 10.1038/s41746-022-00734-2] [Medline: 36596833]
9. Recommendations on Digital Interventions for Health System Strengthening. World Health Organization. 2019. URL: https://www.who.int/publications/i/item/9789241550505? [accessed 2026-01-09]

XSL•FO
RenderX

10. Lewsey SC, Breathett K. Racial and ethnic disparities in heart failure: current state and future directions. Curr Opin Cardiol 2021 May 1;36(3):320-328. [doi: 10.1097/HCO.0000000000000855] [Medline: 33741769]
11. Lawson CA, Zaccardi F, Squire I, et al. 20-year trends in cause-specific heart failure outcomes by sex, socioeconomic status, and place of diagnosis: a population-based study. Lancet Public Health 2019 Aug;4(8):e406-e420. [doi: 10.1016/S2468-2667(19)30108-2] [Medline: 31376859]
12. Schjødt I, Johnsen SP, Strömberg A, Kristensen NR, Løgstrup BB. Socioeconomic factors and clinical outcomes among patients with heart failure in a universal health care system. JACC Heart Fail 2019 Sep;7(9):746-755. [doi: 10.1016/j.jchf.2019.06.003] [Medline: 31466671]
13. Witte KK, Patel PA, Walker AMN, et al. Socioeconomic deprivation and mode-specific outcomes in patients with chronic heart failure. Heart 2018 Jun;104(12):993-998. [doi: 10.1136/heartjnl-2017-312539] [Medline: 29386325]
14. Vinter N, Fawzy AM, Gent D, et al. Social determinants of health and cardiovascular outcomes in patients with heart failure. Eur J Clin Invest 2022 Nov;52(11):e13843. [doi: 10.1111/eci.13843] [Medline: 35924957]
15. Bitomsky L, Pfitzer EC, Nißen M, Kowatsch T. Advancing health equity and the role of digital health technologies: a scoping review protocol. BMJ Open 2024 Oct 15;14(10):e082336. [doi: 10.1136/bmjopen-2023-082336] [Medline: 39414274]
16. Price-Haywood EG, Arnold C, Harden-Barrios J, Davis T. Stop the divide: facilitators and barriers to uptake of digital health interventions among socially disadvantaged populations. TOJ 2023;23(1):34-42. [doi: 10.31486/toj.22.0101]
17. Cheng C, Beauchamp A, Elsworth GR, Osborne RH. Applying the electronic health literacy lens: systematic review of electronic health interventions targeted at socially disadvantaged groups. J Med Internet Res 2020 Aug 13;22(8):e18476. [doi: 10.2196/18476] [Medline: 32788144]
18. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
19. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5(1):69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
20. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
21. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
22. Digital health technologies in heart failure and health inequity. A scoping review of access and use in routine practice. Open Science Framework. URL: https://osf.io/ch2gp/ [accessed 2026-12-06]
23. Enard KR, Coleman AM, Yakubu RA, Butcher BC, Tao D, Hauptman PJ. Influence of social determinants of health on heart failure outcomes: a systematic review. J Am Heart Assoc 2023 Feb 7;12(3):e026590. [doi: 10.1161/JAHA.122.026590] [Medline: 36695317]
24. Braveman P. What are health disparities and health equity? We need to be clear. Public Health Rep 2014;129 Suppl 2(Suppl 2):5-8. [doi: 10.1177/00333549141291S203] [Medline: 24385658]
25. Melh G, Tamrat T. Classification of Digital Health Interventions v1.0. World Health Organization. 2018. URL: https://www.who.int/publications/i/item/WHO-RHR-18.06? [accessed 2026-01-09]

Abbreviations

HF: heart failure
JBI: Joanna Briggs Institute
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
WHO: World Health Organization

Edited by J Sarvestan; submitted 06.Aug.2025; peer-reviewed by W Wang; revised version received 08.Dec.2025; accepted 09.Dec.2025; published 23.Jan.2026.

Please cite as:
Vinter N, Elmegaard M, Køber L, Schou M, Johnsen SP
Inequity in Access to and Use of Digital Health Technologies in Routine Heart Failure Care: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e81949
URL: https://www.researchprotocols.org/2026/1/e81949
doi:10.2196/81949

XSL•FO
RenderX

© Nicklas Vinter, Mariam Elmegaard, Lars Køber, Morten Schou, Søren Paaske Johnsen. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Role of Shame and Concealment in Patient-Provider Communication Among Individuals With Low Health Literacy: Protocol for a Scoping Review

Tsuyoshi Okuhara[1], MPH, PhD; Hiroko Okada[1], MPH, PhD; Rie Yokota[2], MPH, PhD

[1]Department of Health Communication, School of Public Health, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

[2]Department of Medical Communication, School of Pharmacy and Pharmaceutical Sciences, Hoshi University, Tokyo, Japan

Corresponding Author:
Tsuyoshi Okuhara, MPH, PhD
Department of Health Communication, School of Public Health, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

Abstract

Background: Health literacy, defined as the ability to obtain, understand, evaluate, and use health information, influences health behaviors and outcomes. Low health literacy (LHL) is associated with misunderstandings of treatment instructions, poor adherence, and inadequate preventive behaviors, all of which contribute to health disparities. Although universal precautions such as plain language and the teach-back method are recommended, recent studies indicate that these measures alone cannot fully address the challenges faced by patients with LHL. Previous qualitative studies have examined psychosocial processes through which shame and concealment shape patient-provider communication; however, these findings remain fragmented across settings and disciplines, and no scoping or systematic review has yet synthesized this evidence.

Objective: This review aims to map how patients with LHL experience shame and concealment, how health care providers' behaviors and communication influence these experiences, and how these processes affect patient-provider communication and care. It also seeks to identify strategies to reduce the impact of shame and concealment in clinical practice.

Methods: This review will be conducted and reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. PubMed, MEDLINE, CINAHL, PsycInfo, Web of Science, and Academic Search Complete will be searched using terms related to health literacy, patient-provider communication, and qualitative research. Qualitative and mixed methods studies with qualitative findings will be included, and quantitative-only studies will be excluded. The participants may include patients, health care providers, or both. Data extraction will include specific manifestations of shame (eg, embarrassment, fear of judgment, and self-blame), concealment behaviors (eg, silence, avoidance, and impression management), provider communication behaviors (eg, time pressure, use of jargon, dismissiveness, and validation), and key findings. The accuracy of the data will be verified by multiple reviewers. Data will be synthesized using thematic synthesis, with the findings presented in tables; narrative synthesis; and a conceptual model depicting the interactions among shame, concealment, and provider communication. The findings will describe how shame and concealment are conceptualized, how they shape communication and care, and strategies suggested to reduce their effects. A conceptual diagram will illustrate these dynamics.

Results: This study was funded in April 2025. Database searching is scheduled for October 2025, with study selection and data extraction planned for November 2025. As of September 2025, no data extraction has been completed. Data synthesis is expected to be finalized by December 2025, and results are planned for publication between June and August 2026.

Conclusions: This will be the first scoping review to systematically map the roles of shame and concealment in health care communication among patients with LHL. Synthesizing qualitative evidence will provide insights into relational dynamics, inform professional education and training, and guide organizational- and policy-level strategies to promote equitable and patient-centered health care communication.

International Registered Report Identifier (IRRID): PRR1-10.2196/83451

(JMIR Res Protoc 2026;15:e83451) doi:10.2196/83451

KEYWORDS

health literacy; low health literacy; communication; professional-patient relations; shame; concealment; qualitative research; scoping review

XSL•FO
RenderX

Introduction

Health literacy is defined as the ability to obtain, understand, evaluate, and use health information and has been recognized as a crucial factor influencing health behaviors and outcomes [1-3]. Low health literacy (LHL) has been associated with problems such as misunderstanding medication instructions, poor treatment adherence, and inadequate engagement in preventive behaviors [4-6] and has been identified as a risk factor contributing to health disparities and failures in chronic disease management [7-9]. Recent systematic and scoping reviews continue to demonstrate that health literacy is strongly associated with self-management, medication adherence, quality of life, and clinical outcomes across a wide range of chronic conditions [4-6,10]. To address these challenges, "universal precautions" in health literacy have been recommended, including the use of plain language and the teach-back method, to ensure understanding regardless of patients' literacy levels [11].

However, studies have highlighted the limitations of explaining the difficulties of patients with LHL solely in terms of "deficits in information-processing skills." Particular attention has been paid to the psychosocial mechanisms underlying shame and concealment. Parikh et al [12] reported in their seminal study that 42.6% of patients had inadequate or marginal health literacy, nearly 40% of those acknowledging reading difficulties felt ashamed, 19% never disclosed their difficulties to anyone, and over 60% avoided seeking help. In their subsequent study, 47.8% of patients with the lowest literacy level reported shame compared to only 6.5% at higher levels. Although most patients (83.7%) agreed that documenting literacy difficulties in medical records was useful, those with lower literacy levels were significantly more likely to experience embarrassment, which hindered open communication with providers [13]. These findings have since been interpreted through broader stigma frameworks, which conceptualize shame not merely as an individual emotion but as a socially produced response to anticipated devaluation and power asymmetries in health care interactions [14-16].

Qualitative studies have confirmed that shame and concealment profoundly affect health care communication. Easton et al [17] described how patients with low literacy and fear might be perceived as ignorant, avoid asking questions, and leave consultations without sufficient understanding. Other qualitative studies have also revealed that several patients feel inferior when interacting with those who possess medical expertise, leading them to hesitate to ask questions or express disagreement [18,19]. Recent studies have highlighted the broader social and systemic contexts in which these experiences occur. For example, Stormacq et al [20] observed that socioeconomically disadvantaged adults experience health literacy as an "obstacle course," where financial insecurity heightens their need for information but simultaneously restricts access to physicians, leading many to depend on less reliable sources such as the internet or relatives. In this context, difficulties in understanding or evaluating health information are often accompanied by feelings of shame and a tendency toward concealment, which discourage patients from asking questions or disclosing their struggles [20].

In parallel, a growing body of qualitative literature has examined how health care providers' behaviors and communication styles may unintentionally reinforce these psychosocial dynamics. For example, Shaw et al [21] reported that patients with LHL often felt reluctant to ask questions due to physicians' short consultation times, insufficient information provided, and complex terminology used, which ultimately resulted in patient silence or concealment. Similarly, Jensen et al [22] showed that patients with lower literacy were more likely to feel that their voices were not respected by health care providers and that health care providers did not listen adequately, which contributed to lower satisfaction and inhibited patients' willingness to speak up. Recent studies in primary care, physiotherapy, cardiovascular prevention, and palliative care have further demonstrated that providers often struggle to accurately identify LHL and consistently apply communication strategies tailored to patients' needs despite strong normative support for health literacy–friendly practices [23-26].

Taken together, these findings indicate that shame and concealment are not isolated individual reactions but relational processes shaped by social context, stigma, and the communicative practices of health care professionals. If left unaddressed, these processes reinforce silence, misunderstanding, and inequities in care. Nevertheless, despite a growing body of qualitative literature, existing research remains highly fragmented across clinical domains (eg, primary care, chronic disease management, rehabilitation, and palliative care), populations (eg, older adults, socioeconomically disadvantaged groups, and culturally diverse communities), and conceptual frameworks. Some studies foreground shame without explicitly examining concealment behaviors [17,18], others focus on provider communication without directly theorizing emotional processes [23,26], and still others examine stigma at a general level without anchoring it to health literacy–specific interactions [16].

To date, no scoping or systematic review has synthesized this heterogeneous qualitative evidence to clarify how shame and concealment operate as dynamic psychosocial mechanisms within health care communication for patients with LHL. As a result, it remains unclear how these processes unfold across settings, how they are shaped by provider communication styles, and which strategies are most promising for reducing their harmful effects.

This scoping review will focus exclusively on the findings of qualitative and mixed methods studies. This is because phenomena such as shame and concealment are deeply rooted in patients' subjective experiences and emotions and cannot be adequately captured through quantitative measures alone. Importantly, the rationale for excluding quantitative-only studies is not based on a denial of the value of quantitative emotion research. Rather, this review specifically aims to elucidate how shame and concealment are experienced, negotiated, and managed within patient-provider interactions as social and relational processes. These dynamic meaning-making processes are best accessed through qualitative data such as narratives,

XSL•FO
RenderX

interviews, and observational accounts. Recent conceptual work on epistemic humility and epistemic justice further underscores the importance of examining how patients' voices are validated or silenced in clinical interactions, particularly in contexts marked by informational and power asymmetries [27,28]. By synthesizing qualitative research, this review seeks to comprehensively illustrate how patients experience shame, how concealment behaviors arise, and how these behaviors are shaped by social and cultural factors as well as interactions with health care providers.

Existing studies on this topic remain heterogeneous and fragmented, encompassing diverse populations, including adults, older people, low-income groups, and those from various cultural backgrounds. Therefore, this study aims to map the range of relevant research and organize the conceptual and theoretical frameworks. Therefore, a scoping review approach is more appropriate than a systematic review approach.

The following research questions will guide this review:

- How do patients with LHL experience shame and concealment?
- How do health care providers' behaviors (eg, impression-based judgments and labeling) influence patients' experiences of shame and concealment?
- How do the processes of shame and concealment affect patient-provider communication and quality of care?
- What strategies have been proposed to mitigate the effects of shame and concealment in clinical practice?

Methods

Overview

This scoping review will be conducted following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [29]. A completed PRISMA-ScR checklist will be provided as a supplementary file upon manuscript submission. The methodological framework is guided by the seminal work by Arksey and O'Malley [30] on scoping studies, further refined by the work by Levac et al [31] and updated with methodological guidance for conducting scoping reviews [32]. We plan to initiate the database search in October 2025, complete study selection and data extraction by November 2025, and finalize data synthesis by December 2025. In accordance with the primary objective of this scoping review—to map the range, nature, and conceptualization of qualitative evidence rather than evaluate intervention effectiveness or methodological rigor—we will not conduct a formal critical appraisal of individual studies. This approach is consistent with established scoping review methodology and PRISMA-ScR recommendations [29].

Literature Search

We will search the following databases: PubMed, MEDLINE, CINAHL, PsycInfo, Web of Science, and Academic Search Complete. The search strategy will combine MeSH (Medical Subject Headings) and free-text terms related to health literacy and patient-provider communication. Textbox 1 shows the combination of keywords that will be used in our search. We will not filter our searches based on the year of publication. The reference lists of eligible articles and relevant reviews will also be screened for additional information.

XSL•FO
RenderX

Textbox 1. Search terms.

PubMed

- (doctor*[Title/Abstract] OR physician*[Title/Abstract] OR nurs*[Title/Abstract] OR clinician*[Title/Abstract] OR professional*[Title/Abstract] OR provider*[Title/Abstract] OR expert*[Title/Abstract] OR patient*[Title/Abstract] OR adult*[Title/Abstract] OR “lay person”[Title/Abstract] OR laypeople[Title/Abstract] OR Physicians[MeSH Terms] OR Patients[MeSH Terms] OR Adult[MeSH Terms] OR “Health Personnel”[MeSH Terms]) AND (“health literacy”[Title/Abstract] OR literacy[Title/Abstract] OR “Health Literacy”[MeSH Terms]) AND (qualitative[Title/Abstract] OR interview*[Title/Abstract] OR “focus group”[Title/Abstract] OR “Qualitative Research”[MeSH Terms])

MEDLINE

- XB(“doctor*” OR “physician*” OR nurs*[Title/Abstract] OR “clinician*” OR “professional*” OR “provider*” OR “expert*” OR “patient*” OR “adult*” OR “lay person” OR “laypeople”) AND
- XB(“health literacy” OR literacy) AND XB(“qualitative” OR “interview*” OR “focus group”)

CINAHL

- XB(“doctor*” OR “physician*” OR nurs*[Title/Abstract] OR “clinician*” OR “professional*” OR “provider*” OR “expert*” OR “patient*” OR “adult*” OR “lay person” OR “laypeople”) AND XB(“health literacy” OR literacy) AND XB(“qualitative” OR “interview*” OR "focus group)

PsycInfo

- XB(“doctor*” OR “physician*” OR nurs*[Title/Abstract] OR “clinician*” OR “professional*” OR “provider*” OR “expert*” OR “patient*” OR “adult*” OR “lay person” OR “laypeople”) AND XB(“health literacy” OR literacy) AND XB(“qualitative” OR “interview*” OR “focus group”)

Academic Search Complete

- XB(“doctor*” OR “physician*” OR nurs*[Title/Abstract] OR “clinician*” OR “professional*” OR “provider*” OR “expert*” OR “patient*” OR “adult*” OR “lay person” OR “laypeople”) AND XB(“health literacy” OR literacy) AND XB(“qualitative” OR “interview*” OR “focus group”)

Web of Science

- TS=(“doctor*” OR “physician*” OR nurs*[Title/Abstract] OR “clinician*” OR “professional*” OR “provider*” OR “expert*” OR “patient*” OR “adult*” OR “lay person” OR “laypeople”) AND TS=(“health literacy” OR literacy) AND TS=(“qualitative” OR “interview*” OR “focus group”)

Eligibility Criteria

Textbox 2 shows the inclusion and exclusion criteria for this study.

XSL•FO
RenderX

Textbox 2. Inclusion and exclusion criteria.

Inclusion criteria

- Studies involving health care providers, patients, or both
- Studies focusing on patients with low health literacy and their experiences of shame, concealment, or related psychosocial processes in health care contexts
- Studies examining the role of health care providers' communication or behavior in relation to patient shame and concealment
- Qualitative or mixed methods studies including qualitative findings relevant to shame or concealment
- Studies involving participants of any age, sex, ethnicity, or country
- Peer-reviewed articles and gray literature (eg, theses and conference proceedings) providing sufficient methodological details
- Papers written in English
- Studies published in any year

Exclusion criteria

- Quantitative studies that only measure health literacy levels
- Quantitative-only studies that do not include qualitative data capable of supporting thematic synthesis of shame, concealment, or interactional processes
- Studies exclusively focusing on general stigma unrelated to health literacy
- Unavailable full texts or articles not written in English

Study Selection

We will use the Rayyan software (Qatar Computing Research Institute) [33] to manage and screen the citations. Duplicates will be removed automatically. Two reviewers (TO and HO) will independently screen the titles and abstracts against the inclusion criteria. Disagreements will be resolved through discussion, and if a consensus cannot be reached, a third reviewer (RY) will be consulted. The same process will be applied to the full-text screening. The selection process will be documented using a PRISMA-ScR flow diagram.

Data Extraction

One reviewer (TO) will extract the data using a standardized form developed in Microsoft Excel. The following data will be extracted:

- Bibliographic details (author, year, country, and journal)
- Study design and methodology
- Participant characteristics (sample size, demographics, and clinical setting)
- Specific emotional expressions of shame (eg, embarrassment, humiliation, and fear of negative evaluation), concealment behaviors (eg, silence, pretending to understand, avoidance of questions, and impression management)
- Specific provider communication and behavioral features (eg, consultation time constraints, use of medical jargon, dismissive responses, validation, and teach-back method use)
- Key findings (themes, categories, and representative quotations)
- Implications for clinical practice, education, and policy

To ensure accuracy, the second (HO) and third (RY) reviewers will verify the extracted data for omissions or errors.

Data Synthesis

The initial data synthesis will be conducted by the first reviewer (TO) using thematic synthesis [34]. This approach, endorsed by The Cochrane Collaboration as a rigorous strategy for integrating qualitative evidence [35], involves several stages. In the first stage, TO, who has a background in qualitative methodology and extensive experience in conducting qualitative research, will perform open, line-by-line coding of the findings and quotations presented in the results sections of the included articles. In the second stage, both the first (TO) and second (HO) reviewers will independently cluster the codes that emerged from the initial coding into descriptive themes that will be data driven. Any discrepancies will be resolved through discussion with a third reviewer (RY). In the third stage, TO will generate analytical themes by interpreting and extending the descriptive themes from the previous step. These analytical themes will aim to provide insights beyond the findings of individual studies. The development of these themes will be discussed collaboratively among TO, HO, and RY, and all codes and themes will be created inductively through this iterative synthesis process. In addition, study characteristics (eg, setting, population, country, and clinical context) and intervention-related information will be summarized using descriptive numerical analysis and structured tables. These multiple layers of synthesis will be integrated to construct a conceptual model that visually represents the relational pathways linking health literacy, shame, concealment, provider behaviors, and communication outcomes.

Dissemination

We will present the findings in both a summary table and conceptual diagram and elaborate on them through descriptive and narrative synthesis. A conceptual model will be constructed to depict the interplay between patients' experiences of shame and concealment and the behaviors and communication of health

XSL•FO
RenderX

care providers. Following completion of data synthesis, the findings will be submitted to a peer-reviewed journal and disseminated at international conferences.

Results

This study was funded in April 2025. Database searching is scheduled for October 2025, with study selection and data extraction planned for November 2025. As of September 3, 2025, no data extraction has been completed. Data synthesis is expected to be finalized by December 2025, and results are planned for publication between June and August 2026.

Discussion

Anticipated Findings

This scoping review will be the first to systematically map how shame and concealment are experienced by patients with LHL and how health care providers' behaviors may exacerbate or mitigate these experiences. By synthesizing evidence from qualitative and mixed methods studies with qualitative findings, our review will provide a more comprehensive understanding of this psychosocial phenomenon.

Shame and concealment are not isolated experiences but relational processes shaped by clinical encounters and broader social contexts. Previous studies have shown that patients with LHL often hide their reading or comprehension difficulties out of fear of being judged, which prevents them from accessing the necessary support [18,19]. Such concealment can act as a coping mechanism, protecting patients from embarrassment while simultaneously restricting their ability to obtain information and engage fully in care [13]. Our review will highlight how these dynamics contribute to health inequities, particularly among populations already disadvantaged by socioeconomic or cultural factors.

These findings will have implications for both clinical practice and professional training. Health care providers may unintentionally reinforce shame through hurried communication, the use of jargon, or assumptions about patients' abilities [21]. For example, patients often report that rush consultations and lack of plain language increase their reluctance to ask questions [22]. By mapping the available evidence, this review will identify strategies that can reduce these risks, such as adopting plain language, using the teach-back method, and creating safe spaces for patient questions.

Limitations

This study may have some limitations. First, because we will restrict our inclusion to studies published in English, relevant evidence reported in other languages may be overlooked. This limitation is particularly important given that health literacy and the associated experiences of shame and concealment are strongly influenced by cultural and linguistic contexts. Second, although we will attempt to identify gray literature, some relevant reports, theses, or conference proceedings may not be retrieved because of limited indexing. Third, in accordance with the objectives of this scoping review, we will not quantitatively evaluate the effectiveness of the interventions. Consequently, our findings will not provide pooled estimates or definitive conclusions regarding intervention outcomes. Despite these limitations, the scoping review approach is well suited to our aims as it enables us to capture a broad and heterogeneous body of evidence, clarify conceptual boundaries, and identify important gaps in future research.

Conclusions

To the best of our knowledge, this will be the first scoping review to systematically map how shame and concealment are experienced by patients with LHL in the context of health care communication. By synthesizing evidence from qualitative and mixed methods studies with qualitative findings, this review will provide a comprehensive overview of how these psychosocial processes influence patient-provider interactions.

The insights generated will contribute to a deeper understanding of the relational dynamics that shape health communication and highlight how shame and concealment may reinforce existing inequities in care. These findings can inform the design of targeted interventions, guide professional education and training for health care providers, and support the development of organizational- and policy-level strategies to foster more equitable communication practices.

Ultimately, this review aims to provide a foundation for future empirical research and practical initiatives that seek to reduce the hidden barriers posed by shame and concealment, thereby promoting more inclusive, patient-centered, and effective health care communication.

Funding

This work was supported by a Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (25K13483).

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Authors' Contributions

Conceptualization: TO
Funding acquisition: TO
Methodology: TO
Writing—original draft: TO
Writing—review and editing: TO, HO, and RY

XSL•FO
RenderX

Conflicts of Interest

None declared.

References

1. Nutbeam D, Lloyd JE. Understanding and responding to health literacy as a social determinant of health. Annu Rev Public Health 2021 Apr 1;42:159-173. [doi: 10.1146/annurev-publhealth-090419-102529] [Medline: 33035427]
2. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. Health Promot Int 2000 Sep;15(3):259-267. [doi: 10.1093/heapro/15.3.259]
3. Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80. [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
4. Hyvert S, Yailian AL, Haesebaert J, et al. Association between health literacy and medication adherence in chronic diseases: a recent systematic review. Int J Clin Pharm 2023 Feb;45(1):38-51. [doi: 10.1007/s11096-022-01470-z] [Medline: 36369411]
5. van der Gaag M, Heijmans M, Spoiala C, Rademakers J. The importance of health literacy for self-management: a scoping review of reviews. Chronic Illn 2022 Jun;18(2):234-254. [doi: 10.1177/17423953211035472] [Medline: 34402309]
6. Lee H, La IS. Association between health literacy and self-management among middle-aged women: a systematic review. Patient Educ Couns 2024 Jun;123:108188. [doi: 10.1016/j.pec.2024.108188] [Medline: 38354431]
7. Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy to health outcomes. Am J Health Behav 2007;31 Suppl 1:S19-S26. [doi: 10.5555/ajhb.2007.31.supp.S19] [Medline: 17931132]
8. Svendsen MT, Bak CK, Sørensen K, et al. Associations of health literacy with socioeconomic position, health risk behavior, and health status: a large national population-based survey among Danish adults. BMC Public Health 2020 Apr 28;20(1):565. [doi: 10.1186/s12889-020-08498-8] [Medline: 32345275]
9. Schillinger D. Social determinants, health literacy, and disparities: intersections and controversies. Health Lit Res Pract 2021 Jul;5(3):e234-e243. [doi: 10.3928/24748307-20210712-01] [Medline: 34379549]
10. Marshall N, Butler M, Lambert V, Timon CM, Joyce D, Warters A. Health literacy interventions and health literacy-related outcomes for older adults: a systematic review. BMC Health Serv Res 2025 Feb 26;25(1):319. [doi: 10.1186/s12913-025-12457-7] [Medline: 40011860]
11. Brach C, editor. AHRQ Health Literacy Universal Precautions Toolkit, 3rd edition: Agency for Healthcare Research and Quality; 2024.
12. Parikh NS, Parker RM, Nurss JR, Baker DW, Williams MV. Shame and health literacy: the unspoken connection. Patient Educ Couns 1996 Jan;27(1):33-39. [doi: 10.1016/0738-3991(95)00787-3] [Medline: 8788747]
13. Wolf MS, Williams MV, Parker RM, Parikh NS, Nowlan AW, Baker DW. Patients' shame and attitudes toward discussing the results of literacy screening. J Health Commun 2007 Dec;12(8):721-732. [doi: 10.1080/10810730701672173] [Medline: 18030638]
14. Goffman E. Stigma: Notes on the Management of Spoiled Identity: Prentice-Hall; 1963.
15. Link BG, Phelan JC. Conceptualizing stigma. Annu Rev Sociol 2001;27:363-385. [doi: 10.1146/annurev.soc.27.1.363]
16. Stangl AL, Earnshaw VA, Logie CH, et al. The Health Stigma and Discrimination Framework: a global, crosscutting framework to inform research, intervention development, and policy on health-related stigmas. BMC Med 2019 Feb 15;17(1):31. [doi: 10.1186/s12916-019-1271-3] [Medline: 30764826]
17. Easton P, Entwistle VA, Williams B. How the stigma of low literacy can impair patient-professional spoken interactions and affect health: insights from a qualitative investigation. BMC Health Serv Res 2013 Aug 16;13:319. [doi: 10.1186/1472-6963-13-319] [Medline: 23958036]
18. Pio FH, Nosa V. Health literacy of Samoan mothers and their experiences with health professionals. J Prim Health Care 2020 Mar;12(1):57-63. [doi: 10.1071/HC19026] [Medline: 32223851]
19. Sadeghi S, Brooks D, Goldstein RS. Patients' and providers' perceptions of the impact of health literacy on communication in pulmonary rehabilitation. Chron Respir Dis 2013 May;10(2):65-76. [doi: 10.1177/1479972312471548] [Medline: 23364288]
20. Stormacq C, Oulevey Bachmann A, Van den Broucke S, Bodenmann P. How socioeconomically disadvantaged people access, understand, appraise, and apply health information: a qualitative study exploring health literacy skills. PLoS ONE 2023 Aug 9;18(8):e0288381. [doi: 10.1371/journal.pone.0288381] [Medline: 37556436]
21. Shaw A, Ibrahim S, Reid F, Ussher M, Rowlands G. Patients' perspectives of the doctor-patient relationship and information giving across a range of literacy levels. Patient Educ Couns 2009 Apr;75(1):114-120. [doi: 10.1016/j.pec.2008.09.026] [Medline: 19041210]
22. Jensen JD, King AJ, Guntzviller LM, Davis LS. Patient-provider communication and low-income adults: age, race, literacy, and optimism predict communication satisfaction. Patient Educ Couns 2010 Apr;79(1):30-35. [doi: 10.1016/j.pec.2009.09.041] [Medline: 19879093]
23. van der Scheer-Horst E, Rutten G, Stortenbeker I, et al. Limited health literacy in primary care physiotherapy: does a physiotherapist use techniques to improve communication? Patient Educ Couns 2023 Apr;109:107624. [doi: 10.1016/j.pec.2023.107624] [Medline: 36657334]

XSL•FO
RenderX

24. Roodbeen R, Vreke A, Boland G, et al. Communication and shared decision-making with patients with limited health literacy; helpful strategies, barriers and suggestions for improvement reported by hospital-based palliative care providers. PLoS ONE 2020 Jun 19;15(6):e0234926. [doi: 10.1371/journal.pone.0234926] [Medline: 32559237]
25. Perrin A, Damiolini E, Schott AM, et al. Considering health literacy in cardiovascular disease management: a qualitative study on healthcare professionals' and patients' perspectives. BMC Health Serv Res 2022 Sep 5;22(1):1121. [doi: 10.1186/s12913-022-08455-8] [Medline: 36064395]
26. McKinn S, Chapman N, Sharman JE, et al. How do general practitioners manage patient health literacy differences in cardiovascular disease prevention consultations? An interview study. Patient Educ Couns 2024 Aug;125:108299. [doi: 10.1016/j.pec.2024.108299] [Medline: 38657560]
27. Muyskens K, Ang C, Kerr ET. Relational epistemic humility in the clinical encounter. J Med Ethics 2025 Feb 17:jme-2024-110241. [doi: 10.1136/jme-2024-110241] [Medline: 39922693]
28. Galasiński D, Ziółkowska J, Elwyn G. Epistemic justice is the basis of shared decision making. Patient Educ Couns 2023 Jun;111:107681. [doi: 10.1016/j.pec.2023.107681] [Medline: 36871402]
29. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
30. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
31. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
32. Peters MD, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
33. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
34. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008 Jul 10;8:45. [doi: 10.1186/1471-2288-8-45] [Medline: 18616818]
35. Harden A, Thomas J. Thematic synthesis. Presented at: Proceedings of the ESRC Research Methods Festival; Jun 30 to Jul 3, 2008 URL: https://uel-repository.worktribe.com/output/443812/thematic-synthesis [accessed 2026-01-20]

Abbreviations

LHL: low health literacy
MeSH: Medical Subject Headings
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by A Schwartz; submitted 03.Sep.2025; peer-reviewed by K Culhane-Pera; revised version received 09.Dec.2025; accepted 30.Dec.2025; published 30.Jan.2026.

Please cite as:
Okuhara T, Okada H, Yokota R
Role of Shame and Concealment in Patient-Provider Communication Among Individuals With Low Health Literacy: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e83451
URL: https://www.researchprotocols.org/2026/1/e83451
doi:10.2196/83451

© Tsuyoshi Okuhara, Hiroko Okada, Rie Yokota. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 30.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Safety and Impacts of Physical Activity for Individuals Living With Hypermobility Spectrum Disorders and Hypermobile Ehlers-Danlos Syndrome: Protocol for a Scoping Review

Donald William Golden[1], BHPE; Jonathan L Low[2], PhD; Julia T Daun[2], PhD; C Allyson Jones[3], PT, PhD; Liz Dennett[4], BSc, MLIS; Susan Nicole Culos-Reed[1,2], PhD; Ranita Harpreet Kaur Manocha[2,5], MD, MSc

[1]Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

[2]Faculty of Kinesiology, University of Calgary, Calgary, AB, Canada

[3]Department of Physical Therapy, University of Alberta, Edmonton, AB, Canada

[4]Geoffrey and Robyn Sperber Health Sciences Library, University of Alberta, Edmonton, AB, Canada

[5]Department of Clinical Neurosciences, Cumming School of Medicine, University of Calgary, Room AC144, Special Services Building, Foothills Medical Centre, 1403 – 29 Street NW, Calgary, AB, Canada

Corresponding Author:
Ranita Harpreet Kaur Manocha, MD, MSc
Faculty of Kinesiology, University of Calgary, Calgary, AB, Canada

Abstract

Background: Although physical activity (PA) participation has known health benefits, many individuals with hypermobility spectrum disorders (HSD) and hypermobile Ehlers-Danlos syndrome (hEDS) have difficulties participating in PA. HSD/hEDS affect approximately 1 in 500 people and are more prevalent in females. HSD/hEDS may result in frequent joint injuries, chronic pain, and generalized fatigue. These symptoms, and a fear of new or reinjury, may result in barriers to PA participation. Overall, there is limited research on PA in this population. Previous exercise reviews have not included structured PA such as sports and occupational activities; unstructured PA such as play, household, or leisure activities; or younger ages, including children. Additionally, some females with HSD/hEDS report experiencing more frequent joint injuries and worsening pain and fatigue during times of hormonal transitions, such as puberty, pregnancy, as well as cyclically across the menstrual cycle. Some females also report improvements in symptoms and a reduction in injury frequency following menopause. The impacts of PA during these times of changing hormone levels for females living with HSD/hEDS are uncertain. A clear understanding of what types of structured and unstructured PA are safe and helpful for individuals of all ages with HSD/hEDS, and if PA should be adapted during times of female hormonal changes, is lacking.

Objective: We propose a scoping review protocol to map and synthesize the evidence regarding considerations that may impact structured and unstructured PA participation in HSD/hEDS for individuals of all ages and during times of female hormonal transitions.

Methods: A scoping review will be conducted using Covidence (Veritas Health Innovation Ltd) and Microsoft Excel (Microsoft Corp) to map the evidence regarding the impacts of PA on safety, physical health, and quality of life. These outcomes will be assessed using the World Health Organization International Classification of Functioning, Disability and Health framework.

Results: The database search was performed on August 22, 2024, and updated on September 8, 2025. Data extraction started in September 2025 and is ongoing. The results are expected to be published by August 2026.

Conclusions: This proposed scoping review will aid in defining critical research directions regarding PA in HSD/hEDS, which may help inform guidelines outlining the risks and benefits of structured and unstructured PA. This review will also help define existing evidence for age-specific and hormone-related considerations regarding the impacts of PA in this population. This is particularly important as PA may help ameliorate the physical and mental symptoms associated with HSD/hEDS and may improve quality of life for these individuals across the lifespan.

Trial Registration: OSF Registries osf.io/r6wf8; https://doi.org/10.17605/OSF.IO/GE5AK

International Registered Report Identifier (IRRID): DERR1-10.2196/80394

(JMIR Res Protoc 2026;15:e80394) doi:10.2196/80394

KEYWORDS

Ehlers-Danlos syndrome; exercise; joint instability; safety; International Classification of Functioning, Disability and Health

XSL•FO
RenderX

Introduction

Overview

Hypermobility spectrum disorders (HSDs) and the hypermobile subtype of Ehlers-Danlos syndrome (hEDS) are conditions that predominantly affect females [1]. While the true prevalence of these conditions is unclear and thought to be underestimated [2], the combined prevalence of HSD/hEDS has been reported to be 1 in 500 (Wales, United Kingdom), and recent unpublished data have shown a prevalence of 1 in 227 (Northumberland, United Kingdom) [1,3]. HSD/hEDS may be characterized by recurrent joint injuries, chronic pain, generalized fatigue, and muscular weakness [4]. As a result, many individuals with HSD/hEDS report a fear of new or reinjury due to movement (kinesiophobia), and many tend to be physically inactive [5-7]. Inactivity may further exacerbate HSD/hEDS symptoms [6,7], prompting a vicious cycle of symptoms reinforcing sedentary behavior and vice versa.

Physical activity (PA) is defined as any movement that increases energy expenditure above resting levels [8,9]. PA may be categorized as structured (exercise, sports, and occupational activities) or unstructured (play, household, or leisure activities, and functional daily movements such as walking or climbing stairs) [9]. Exercise is a structured form of PA that is repetitive, structured, and purposeful for improving health or fitness [9]. PA has known health benefits, including improved cardiovascular, bone, and mental health [8]. As the physical and cognitive benefits of PA may improve quality of life in HSD/hEDS [5], a better understanding of optimal guidance and prescription of PA for this population is needed, particularly around the salient risks and tangible benefits of PA.

Females with HSD/hEDS have reported worsening symptoms and more frequent joint injuries during puberty, pregnancy, and at various points in their menstrual cycles [10]. Symptom severity and injury frequency have also been reported to improve after menopause [10]. Changes in symptoms and injury risk during these times may impact an individual's ability to participate in PA. To our knowledge, there has been no mention of female sex hormones, puberty, the menstrual cycle, pregnancy, or menopause in previous reviews of structured exercise in HSD/hEDS [11-17].

The World Health Organization International Classification of Functioning, Disability, and Health (ICF) is an internationally recognized classification system that provides a conceptual framework to determine outcomes relevant to daily function, including physical, cognitive, social, and environmental domains [18]. An ICF model has been developed specifically for Ehlers-Danlos syndromes based on clinical presentations and patient experiences, revealing that HSD/hEDS impacts each domain of the ICF [5,7]. Therefore, the ICF provides a relevant framework to categorize PA-associated outcomes related to quality of life for individuals with HSD/hEDS.

Guidance as to what PAs are safe, appropriate, and beneficial for living with HSD/hEDS may help reduce pain and fear associated with PAs in this population. Prior reviews have focused on exercise interventions as part of physical therapy treatment in HSD/hEDS, including therapeutic exercise and strength training alongside manual therapy or other modalities for rehabilitation [11-17,19]. However, multiple knowledge gaps remain. Prior reviews did not consider children, older adults, effects of female sex hormones, or times of female hormonal changes during the lifespan, unstructured PA, or any structured activities outside of physical therapy and rehabilitation, such as different sports or occupational activities. Therefore, this scoping review will help identify what PAs are appropriate and safe for individuals with HSD/hEDS outside of a physical therapy session, if PA should be adapted during specific times of female hormonal changes, and how PA across the lifespan may be affected. This will help guide clinicians in PA counseling for symptom management and improved quality of life in this patient population [5].

Objectives

This protocol will detail the methodology of the proposed scoping review, which will appraise the evidence regarding the safety and impacts of structured and unstructured PA for individuals of all ages with HSD/hEDS. Categorizing the effects of PA in HSD/hEDS based on ICF domains will provide a comprehensive description of the impacts on quality of life. As female sex hormones may influence symptoms of HSD/hEDS, it is also important to consider hormonal changes across the lifespan when assessing the safety and impacts of PA. Based on the state of the literature, a scoping review is justified [20].

Aims and Research Questions

The aims and research questions are listed in Textbox 1.

XSL•FO
RenderX

Textbox 1. Aims and research questions.

Primary aim: identify all structured and unstructured physical activities (PAs) that have been investigated in hypermobility spectrum disorder (HSD)/hypermobile Ehlers-Danlos syndrome (hEDS). Primary research questions:

- In HSD/hEDS, what types of PAs have been investigated?

Secondary aim: describe the extent of evidence related to the benefits and harms of the identified PAs across several ICF domains within the International Classification of Functioning, Disability, and Health (ICF) "Body Functions and Structures" and "Activities and Participation" categories.

- For each PA identified:
- "Body Function" domain:
- Mental functions: number of studies, benefits, and adverse events (types and frequencies)
- Sensory functions and pain: number of studies, benefits, and adverse events (types and frequencies)
- Functions of the cardiovascular and respiratory systems: number of studies, benefits, adverse events (types and frequencies)
- Functions of the digestive and metabolic systems: number of studies, benefits, and adverse events (types and frequencies)
- Genitourinary and reproductive functions: number of studies, benefits, and adverse events (types and frequencies)
- Neuromusculoskeletal and movement-related functions: number of studies, benefits, and adverse events (types and frequencies)
- Functions of the skin: number of studies and adverse events (types and frequencies)
- "Body Structure" domain:
- Structure of the nervous system: number of studies, benefits, and adverse events (types and frequencies)
- The eye, ear, and related structures: number of studies, benefits, and adverse events (types and frequencies)
- Structure of the cardiovascular and respiratory systems: number of studies, benefits, and adverse events (types and frequencies)
- Structures related to the digestive and metabolic systems: number of studies, benefits, and adverse events (types and frequencies)
- Structures related to movement: number of studies, benefits, and adverse events (types and frequencies)
- Skin structures: number of studies and adverse events (types and frequencies)
- "Activities and Participation" domain
- Mobility: number of studies, benefits, and adverse events (types and frequencies)
- Self-care: number of studies and benefits
- Interpersonal interactions and relationships: number of studies and benefits
- Community, social, and civic life: number of studies and benefits
- "Environmental Factors" domain
- Support and relationships: number of studies and benefits
- Attitudes: number of studies and benefits

Tertiary aim: within the context of the ICF Body Functions: Genitourinary and Reproduction Functions domain, identify and describe the extent of evidence regarding safety, benefits, or harms of identified PAs in HSD/hEDS in relation to the influence of changes in female sex hormones. Tertiary research questions:

- How many studies reported on adverse events during PA or the impact of PA along with:
- Puberty status?
- Menstrual cycle phase, length, and regularity, or cessation of menstruation?
- Hormonal contraceptive use?
- Pregnancy?
- Menopause status?
- What were the findings of these studies?

XSL•FO
RenderX

Methods

Protocol Design

This proposed scoping review will follow the methodology described by Arksey and O'Malley [20], Peters et al [21], and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [22]. Any deviations from this protocol will be included as an update within the OSF (Open Science Framework) registration and a supplementary document in the final review. Ethics approval from an institutional review board was not necessary, as this review will only pursue secondary use of publicly available published data.

Eligibility Criteria

Eligibility criteria will be defined according to the PCC (population, concept, and context) format [22] (Table 1).

Table . Inclusion and exclusion criteria.

	Inclusion criteria	Exclusion criteria
Population		
	• Participants with any of: hEDS[a], HSDs[b], symptomatic generalized joint hypermobility, joint hypermobility syndrome, benign joint hypermobility syndrome, or EDS[c] type III[d]. • Studies that report outcomes of other connective tissue disorders as well as HSD/hEDS will be included; however, extraction will focus on the outcomes of those with HSD/hEDS. • Any age, sex, or gender. • Any PA[e].	• Solely includes genetically defined connective tissue disorders (ie, Marfan syndrome, Loeys-Dietz syndrome, and vascular EDS[c]). • This is necessary as many of these conditions are associated with life- and limb-threatening arterial complications, and as such would have different implications for PA.
Concept		
	• PA safety: adverse events such as falls, heart attack, stroke, aortic dissection, respiratory distress, injury (strain, sprain, subluxation, or dislocation), etc. • Benefits or harms of PA on quality of life (Table 2).	• Multimodal interventions (ie, exercise plus nutritional intervention). • Therapeutic modalities for pain (ie, acupuncture or manipulation). • Psychotherapeutic modalities (ie, exercise counseling and cognitive behavioral therapy).
Context		
	• As the HSD/hEDS nomenclature has undergone reclassification in 1988, 1997, and 2017, no limits will be placed on the timeframe for when studies were published or when data collection occurred. • Studies from any geographical location, country, or institution, and studies including participants from any cultural or social background.	• Non-English language[f] • Nonhuman participants

[a]hEDS: hypermobile Ehlers-Danlos syndrome.

[b]HSD: hypermobility spectrum disorder.

[c]EDS: Ehlers-Danlos syndrome.

[d]Multiple terms have been incorporated to reflect changes in nomenclature over time. A clinical diagnosis for any of these conditions must cite or adhere to established diagnostic criteria in use at the time of diagnosis [4,23-26].

[e]PA: physical activity.

[f]A Multimedia Appendix will be included of potentially relevant studies in languages other than English.

XSL•FO
RenderX

Table . Data items and outcomes according to the ICF[a] domain [5].

Data item	Description
Participant factors	
Diagnosis	• Hypermobile Ehlers-Danlos syndrome, hypermobility spectrum disorders, symptomatic generalized joint hypermobility, joint hypermobility syndrome, benign joint hypermobility syndrome, and Ehlers-Danlos syndrome type III.
Comorbidities or associated conditions	• POTS[b], MCAS[c], fibromyalgia, asthma, etc.
PA[d] details	• Frequency, intensity, and duration. • Structured versus unstructured (Structured PA may include exercise, sports, or occupational activities. Unstructured PA may include play, household or leisure activities, or functional daily movements.).
Comparative intervention	• No PA, nutrition intervention only, etc.
Adverse events	• Occurrence and frequency of falls, strains, sprains, subluxations, dislocations, cardiovascular events, respiratory distress, and syncope.
Impact of PA or inactivity	• Defined as helpful or beneficial if outcomes related to specific ICF domains are improved. • Defined as harmful if outcomes related to ICF domains are diminished.
Hormonal contraceptive use	• If reported, type: intrauterine system (copper or hormone), oral (combined or progestin only, monophasic, biphasic, or triphasic), and other (transdermal, vaginal ring, etc).
Menstrual cycle influences	• If reported: menstrual cycle phase, length, regularity, and amenorrhea.
Menopause status	• If reported: premenopause, perimenopause, or postmenopause.
Puberty status	• If reported: prepubertal, peripubertal, or postpubertal.
Pregnancy status	• If reported: first, second, or third trimester.
Physical outcomes	ICF domain
Vestibular and orthostatic symptoms	• Body function, sensory functions, and pain. • If reported: dizziness, syncope, orthostatic hypertension, etc.
Gastrointestinal and urinary symptoms	• Body function, digestive, metabolic, and endocrine systems. • If reported: gastroesophageal reflux, heartburn, bloating, recurrent abdominal pain, diarrhea, irritable bowel syndrome, etc.
Deconditioning	• Body function, cardiovascular, hematological, immunological, and respiratory systems. • If reported: negative changes in muscle composition, cardiovascular function, and aerobic fitness.
Muscle strength	• Body function, neuromuscular, and movement-related functions.
Muscle endurance	• Body function, neuromuscular, and movement-related functions.
Balance	• Body function, neuromuscular, and movement-related functions.
Proprioception	• Body function, neuromuscular, and movement-related functions.
Functional performance	• Body function, neuromuscular, and movement-related functions.
Exercise capacity	• Body function, neuromuscular, and movement-related functions.
Bone health	• Body structure and structures related to movement.

XSL•FO
RenderX

Data item	Description
Mobility limitations	• Activities and participation, general tasks and demands, mobility, and self-care.
Use of technology or assistive devices to enhance participation	• Environmental factors, products, and technology. • If reported: braces, mobility aids, apps, reminders, etc.
Psychological outcomes	
Fatigue	• Body function and mental functions. • If reported: general, cognitive, or physical fatigue.
Anxiety	• Body function and mental functions.
Depression	• Body function and mental functions.
Sleep quality	• Body function and mental functions.
Stress	• Body function and mental functions.
Kinesiophobia	• Body function and mental functions.
Pain	• Body function, sensory functions, and pain. • If reported: whole body pain, specific joint pain, visceral pain, headache, etc Nociceptive pain and neuropathic pain.
Social outcomes	
Life satisfaction	• Activities and participation, and self-care.
Relationship avoidance or satisfaction	• Activities and participation, interpersonal relations, and relationships.
Work status	• Activities and participation, and major life areas. • If reported: full-time, part-time, sick leave, long-term disability, etc.

[a]ICF: International Classification of Functioning, Disability, and Health.

[b]POTS: postural orthostatic tachycardia syndrome.

[c]MCAS: mast cell activation syndrome.

[d]PA: physical activity.

Information Sources

Databases to be searched include MEDLINE, Embase, CINAHL, SPORT Discus, and Scopus with a search end date of September 5, 2025. Study designs such as randomized controlled trials, nonrandomized controlled trials, prospective and retrospective cohort studies, case-control studies, cross-sectional studies, qualitative studies (phenomenology, ethnography, qualitative descriptions, etc), and case studies will be included. Systematic reviews and meta-analyses that meet eligibility criteria will be excluded but flagged for citation chaining. Study designs that will be excluded from this review include conference abstracts or proceedings, editorials, commentaries, and gray literature. Citation chaining will be used for the studies included in the data extraction phase.

Search Strategy

The search strategy was developed in consultation with DWG, CAJ, and RHKM and performed by a health sciences librarian with experience developing search strategies and performing searches for systematic reviews (LD). A preliminary search was conducted in Ovid MEDLINE to refine the keywords and terms used in the search strategy. This involved running the search with the selected MeSH (Medical Subject Headings) terms and text words in MEDLINE and reviewing the results in consultation with all coauthors. We then performed a preliminary screening of 100 papers to further refine the search terms and keywords. The search strategy will include subject headings (MeSH) and text words related to HSD/hEDS and PA such as Ehlers-Danlos, hypermobility, exercise, gymnastics, stretching, dancing, fitness, PA, sports, tai chi, yoga, qigong, running, swimming, walking, activity tracker, step count, pedometer, daily activities, active lifestyle, housework, manual labor, gardening, and yard work. As the tertiary aim is exploratory, there is no need to add additional keywords related to female hormones, as any relevant papers will be identified using the above strategy. The search will be limited to English, and no date limit will be applied. The search strategy will be adjusted for the specific formatting of each database, and results will be uploaded to the Covidence web-based software platform.

Study Selection

Covidence software will be used to manage the reviews. Three reviewers will independently perform title and abstract screening and full-text review. Reviewer training will involve an

XSL•FO
RenderX

educational session outlining the aims, inclusion and exclusion criteria, and screening procedures for each step of the review. Reviewers will be asked to read a reference instructional manual specific to these procedures. Following this, title-abstract calibration will be performed using the first 100 abstracts. Reviewers will independently review the abstracts, and then conflicts and questions will be discussed during a formal debriefing session to ensure consistency between reviewers and consensus in understanding the eligibility criteria. If needed, the instructional manual will be updated. Full-text calibration will be similarly performed through reviewing the first 2 manuscripts assigned to reviewers B and C with reviewer A, who will be reviewing all manuscripts (Figure 1). Again, if needed, the instructional manual will be updated. The screening protocol is shown in Figure 1.

Studies must meet eligibility criteria based on the agreement of 2 reviewers at the title and abstract screening stage, as well as at the full-text review stage, to be included for data extraction. For any studies where 2 reviewers are uncertain after screening the title and abstract (maybe-maybe), these studies will be included for full text review for further assessment against the eligibility criteria. Disagreements between 2 reviewers (yes-no, yes-maybe, or maybe-no) at any stage will be resolved by discussion between all reviewers. Journal titles and original study authors will not be blinded. Interrater agreement for screening and full text review will be reported as Kappa scores.

Figure 1. Screening and extraction protocol. Figure created in Microsoft PowerPoint (Microsoft Corp). Figure design adapted with permission from Lima et al [27].

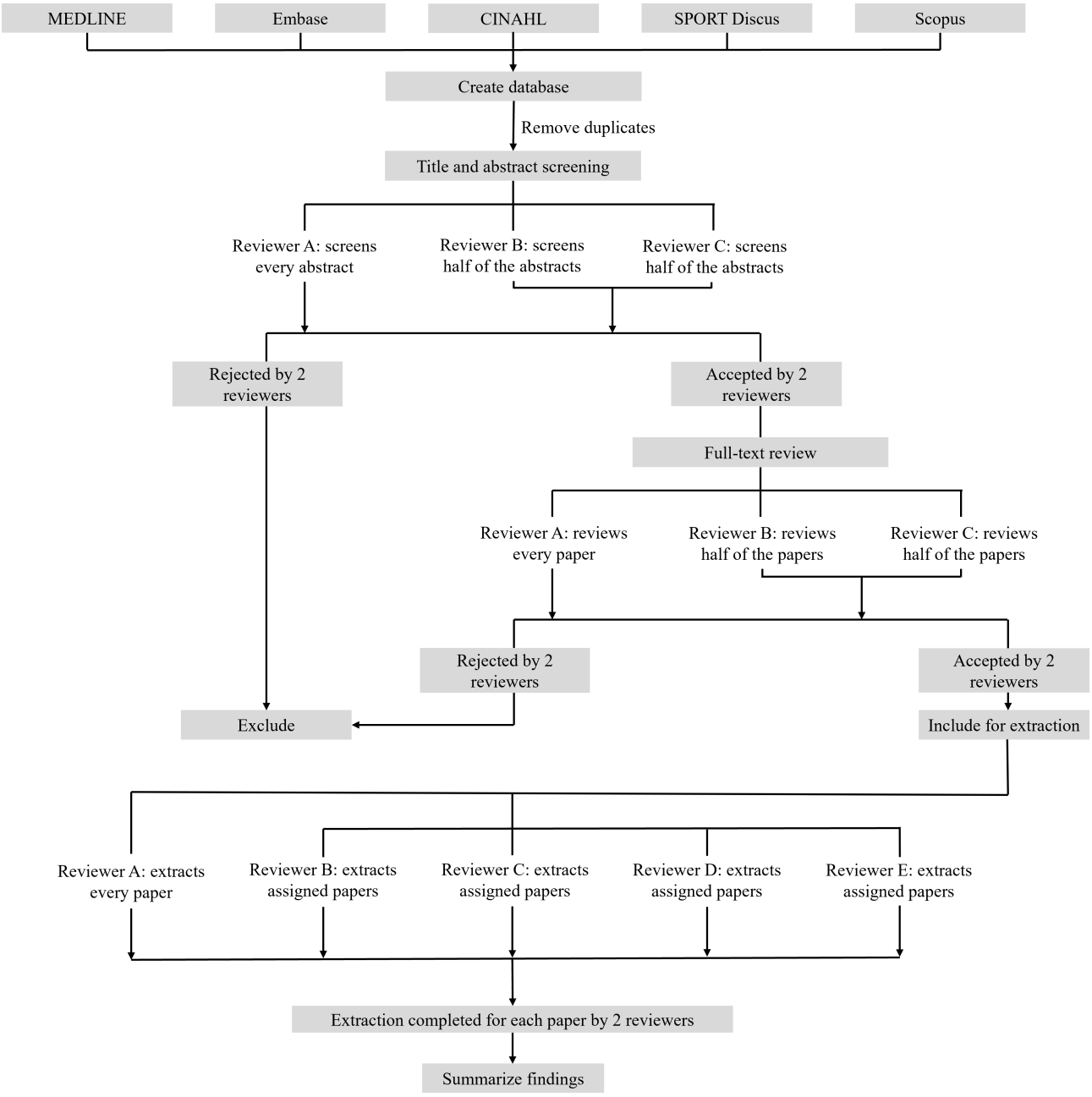

XSL•FO
RenderX

Data Extraction

Data extraction will follow the ICF framework. Data items and outcomes according to ICF domains are listed in Table 2. A data extraction template will be created using this framework. Outcomes will be grouped based on the keywords for data items in Table 2. Data extraction will be performed independently by 5 reviewers. Reviewers onboarded for data extraction (D and E) will be trained as described in the screening protocol. Included research studies will be assigned to each of the 4 reviewers (B, C, D, and E), with the fifth reviewer (A) extracting every study to ensure consistency and extraction by the 2 reviewers per paper (Figure 1). Any discrepancies between reviewers will be resolved by consensus discussion between the 5 reviewers. Similar to screening and review, full-text calibration will be performed through extracting the first 2 research studies assigned to each of the 4 reviewers (B, C, D, and E) with the fifth reviewer (A), who will be extracting all included studies. If needed, the instructional manual will be updated. Operational definitions and methods of measurement for each outcome will be extracted and summarized in a table in the final review. Data management will include secure storage of search strategies, the review protocol, reviewer notes, and extracted data in an online Microsoft OneDrive (Microsoft Corp) shared between all authors. The data extraction sheet will be reported as a Multimedia Appendix in the final scoping review. Authors of the original studies will be contacted if information related to the population, concepts, or outcomes is unclear. Three email attempts will be made (once every 2 wk) to the corresponding author to obtain missing information. If no response is received, the missing data will be reported as "not reported" and will not be included in the data synthesis.

Knowledge Dissemination Plan

Two patient and 4 clinical knowledge users have been involved in the design of this review and the multifaceted integrated knowledge translation plan through formal quarterly meetings. The lived experiences of these knowledge users informed the research questions and specific aims of this review. Specifically, the influence of hormonal contraceptive use on the safety and impacts of PA was included in the research aims based on feedback from these knowledge users. Based on feedback from the knowledge users, the findings of this review will be shared with individuals living with HSD/hEDS, clinicians, and researchers. We will present this at an annual event for patients, researchers, and clinicians and expect attendance to exceed our 2024 event, which was attended by 106 individuals in person and 117 online. We will share the results through a webinar with copresentation with the knowledge users. The knowledge users will also help us craft plain-language summaries of the research findings to be shared on the laboratory websites of the authors. Our research team will generate infographics summarizing the published review to share on social media to reach patients and caregivers. Based on the principal investigator's laboratory social media metrics, we would expect approximately 300 people to access the infographics. The findings will also be shared with clinicians through 2 local education sessions (1 for physicians and 1 for physical therapists) and 1 provincial education session for physical therapists. Clinicians and researchers will be targeted through presentation at 1 international conference, the publication of this protocol, and the final review.

Findings will be summarized in separate tables related to each research aim. In accordance with the guidelines for scoping reviews [28], quality assessment of methodology and risk of bias will not be performed. Analyses will include frequency counts and descriptive statistics, including percentages or proportions. A narrative summary will be performed to articulate the current state of knowledge regarding the safety, benefits, and harms of PA in HSD/hEDS and patient-specific considerations. This approach has been recommended when including both qualitative and quantitative research designs [29].

Results

The search strategy was performed on August 22, 2024, and updated on September 8, 2025. Of 3718 citations identified, 2026 duplicates were removed, 1692 papers were screened, and 57 papers were included for data extraction. Data extraction began in September 2025. The results are expected to be published by August 2026. The number of excluded studies and reasons for exclusion will be reported in the final review using a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart (Figure 2).

XSL•FO
RenderX

Figure 2. Draft PRISMA flowchart for study selection. Figure created in Microsoft PowerPoint. PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

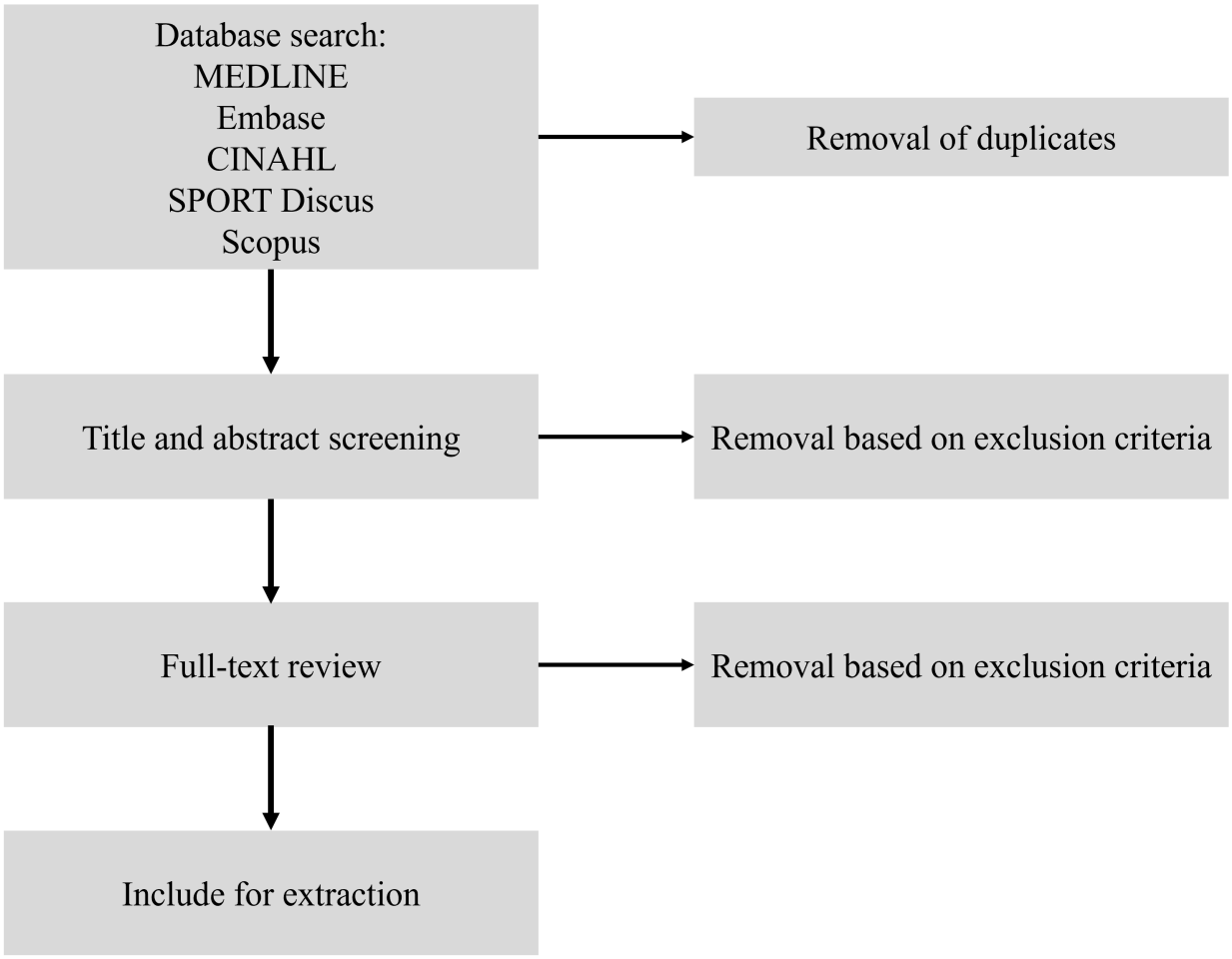

Discussion

While there is some degree of overlap with previous reviews [11-17,19], this proposed scoping review will add to the understanding of the safety and impacts of PA beyond those included in previous reviews. For example, Brittain et al [19] addressed physical therapy-based exercise, manual therapy, and modalities, which would be the focus of the practicing physical therapist to incorporate during clinical sessions with a patient. Reychler et al [14] also focused on therapeutic exercise as part of physical therapy treatment. In contrast, our review will consider structured PA (exercise and sports) and unstructured PAs (play, household, or leisure activities) that occur outside of physical therapy sessions. Thus, this proposed review may be more practically able to be used by clinicians and patients to explore PA interventions outside of the physical therapist's office. As such, this proposed review differs in aim and outcomes compared to previous reviews.

Although we expect to find little data on the role of female sex hormones, given the prior literature the authors are aware of [11-17,19], it is important to search for such information and highlight the knowledge gap, if it exists, based on feedback from patient experiences [10]. As HSD/hEDS results in symptoms that can be severely limiting for physical function, and symptoms have been reported to worsen during times of hormonal fluctuations for some females [10], it is important to highlight what literature exists in this area to direct future research around how times of hormonal fluctuations may influence the safety, benefits, or harms of different PAs. To accommodate the expected scarcity of literature, findings related to the tertiary aim will be reported as count data to map existing evidence and clearly define where literature gaps exist.

To promote generalizability of the findings of this review, all age groups will be included, as well as previous classifications of HSD and hEDS. To minimize overlap in the extraction and synthesis for each aim of this review, the research aims are separated based on the specific data items and outcomes, divided into the domains of the ICF, as shown in Table 2. Thus, overlap between aims will be minimized through specific data extraction items and the presentation of results according to each research aim.

While there are many outcomes for extraction in this review, very few studies identified in preliminary screening included more than 5 outcome variables. As such, most studies are expected to not require in-depth extraction of every listed outcome. To acknowledge the burden on reviewers, biweekly check-in meetings will be implemented to make sure reviewers are not overwhelmed. All reviewers will be trained and undergo

XSL•FO
RenderX

calibration exercises according to the procedure described in the methods of this protocol. To accommodate the number of outcomes during data synthesis, count and frequency data will be presented in concise tables to summarize the findings separately for each research aim.

To help inform future research and clinical practice, it is critical to understand how outcome measures for various ICF domains are being measured and reported in HSD/hEDS. To account for heterogeneity in the definition and measurement of outcomes between studies, variation in outcome definitions and methods will be summarized in the outcome tables for each research aim during data extraction and synthesis. Such heterogeneous outcomes will be grouped based on the keyword (such as fatigue or functional performance), and the various operational definitions and measurements used by each study will be listed.

The ICF model is not without limitations. The individual experience of HSD/hEDS is highly variable [4,5,30,31], and therefore, some nuance could be missed. In addition, there may be domains specific to HSD/hEDS that are not encapsulated within the typical “Body Function” domains. However, we are exploring multiple ICF domains to better capture the psychosocial-cultural impacts of PA.

Appropriate PA may help ameliorate symptoms and improve the quality of life for individuals living with HSD/hEDS. Defining the current evidence and specific knowledge gaps around structured and unstructured PA outside of physical therapy interventions, as well as age-specific and hormone-related considerations regarding the impacts of PA in this population, is critical to informing specific directions for future research and clinical practice recommendations for safe and potentially beneficial PA in HSD/hEDS. This review serves to provide a comprehensive map of the evidence related to PA in this population and will create a call to action for additional research to address the specific identified gaps, an important step toward the development of recommendations for patients and clinicians. As such, this review will inform future research that may have implications for the development of standardized PA guidelines for individuals living with HSD/hEDS, as well as inform clinicians when creating individualized PA guidelines for patients.

Funding

No specific funding was obtained for this project. The funding organizations for DWG's scholarships had no influence on this study's design and will have no input on study conclusions or knowledge dissemination.

Authors' Contributions

Conception and design: DWG, CAJ, and RHKM. Development of search, selection, and data extraction criteria: DWG, CAJ, LD, and RHKM. Literature searches: LD. Paper writing: DWG and RHKM. Critical review of this paper: DWG, JLL, JTD, CAJ, LD, SNC-R, and RHKM. All authors read, provided feedback, and approved the final protocol to be published.

Conflicts of Interest

DWG received financial support for his graduate studies from the following: Hotchkiss Brain Institute Graduate Scholarship in Neuroscience, Alberta SPOR SUPPORT Unit Graduate Studentship, Canadian Institutes of Health Research Canada Graduate Scholarships – Master’s Award, and Alberta Graduate Excellence Scholarship – Research.

References

1. Demmler JC, Atkinson MD, Reinhold EJ, Choy E, Lyons RA, Brophy ST. Diagnosed prevalence of Ehlers-Danlos syndrome and hypermobility spectrum disorder in Wales, UK: a national electronic cohort study and case-control comparison. BMJ Open 2019 Nov 4;9(11):e031365. [doi: 10.1136/bmjopen-2019-031365] [Medline: 31685485]
2. Hakim AJ, Tinkle BT, Francomano CA. Ehlers-Danlos syndromes, hypermobility spectrum disorders, and associated co-morbidities: reports from EDS ECHO. Am J Med Genet C Semin Med Genet 2021 Dec;187(4):413-415. [doi: 10.1002/ajmg.c.31954] [Medline: 34793630]
3. Saravanan V, Reinhold E, Demmler J, Loughran I, Frankel B, The Ehlers-Danlos Support UK. P105 healthcare use by women with joint hypermobility and Ehlers-Danlos syndromes in Northumberland. Rheumatology (Sunnyvale) 2024 Apr 24;63(Supplement_1):keae163.146. [doi: 10.1093/rheumatology/keae163.146]
4. Malfait F, Francomano C, Byers P, et al. The 2017 international classification of the Ehlers-Danlos syndromes. Am J Med Genet C Semin Med Genet 2017 Mar;175(1):8-26. [doi: 10.1002/ajmg.c.31552] [Medline: 28306229]
5. Labanca L, Brizola E, Luciani D, Luppi G, Sangiorgi L, Benedetti MG. Clinical-functional features of individuals with osteogenesis imperfecta and Ehlers-Danlos syndromes: a scoping review of assessment tools and ICF model. Musculoskelet Sci Pract 2023 Apr;64:102742. [doi: 10.1016/j.msksp.2023.102742] [Medline: 36905741]
6. Simmonds JV, Herbland A, Hakim A, et al. Exercise beliefs and behaviours of individuals with joint hypermobility syndrome/Ehlers-Danlos syndrome - hypermobility type. Disabil Rehabil 2019 Feb;41(4):445-455. [doi: 10.1080/09638288.2017.1398278] [Medline: 29125009]
7. Scheper MC, Juul-Kristensen B, Rombaut L, Rameckers EA, Verbunt J, Engelbert RH. Disability in adolescents and adults diagnosed with hypermobility-related disorders: a meta-analysis. Arch Phys Med Rehabil 2016 Dec;97(12):2174-2187. [doi: 10.1016/j.apmr.2016.02.015] [Medline: 26976801]

XSL•FO
RenderX

8. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809. [doi: 10.1503/cmaj.051351] [Medline: 16534088]
9. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Rep 1985;100(2):126-131. [Medline: 3920711]
10. Hugon-Rodin J, Lebègue G, Becourt S, Hamonet C, Gompel A. Gynecologic symptoms and the influence on reproductive life in 386 women with hypermobility type ehlers-danlos syndrome: a cohort study. Orphanet J Rare Dis 2016 Sep 13;11(1):124. [doi: 10.1186/s13023-016-0511-2] [Medline: 27619482]
11. Smith TO, Bacon H, Jerman E, et al. Physiotherapy and occupational therapy interventions for people with benign joint hypermobility syndrome: a systematic review of clinical trials. Disabil Rehabil 2014;36(10):797-803. [doi: 10.3109/09638288.2013.819388] [Medline: 23889528]
12. Castori M, Morlino S, Celletti C, et al. Management of pain and fatigue in the joint hypermobility syndrome (a.k.a. Ehlers-Danlos syndrome, hypermobility type): principles and proposal for a multidisciplinary approach. Am J Med Genet A 2012 Aug;158A(8):2055-2070. [doi: 10.1002/ajmg.a.35483] [Medline: 22786715]
13. Buryk-Iggers S, Mittal N, Mina DS, et al. Exercise and rehabilitation in people with Ehlers-Danlos syndrome: a systematic review. Arch Rehabil Res Clin Transl 2022 Jun;4(2):100189. [doi: 10.1016/j.arrct.2022.100189] [Medline: 35756986]
14. Reychler G, De Backer MM, Piraux E, Poncin W, Caty G. Physical therapy treatment of hypermobile Ehlers-Danlos syndrome: a systematic review. Am J Med Genet A 2021 Oct;185(10):2986-2994. [doi: 10.1002/ajmg.a.62393] [Medline: 34145717]
15. Zabriskie HA. Rationale and feasibility of resistance training in hEDS/HSD: a narrative review. J Funct Morphol Kinesiol 2022 Aug 20;7(3):61. [doi: 10.3390/jfmk7030061] [Medline: 35997377]
16. Palmer S, Bailey S, Barker L, Barney L, Elliott A. The effectiveness of therapeutic exercise for joint hypermobility syndrome: a systematic review. Physiotherapy 2014 Sep;100(3):220-227. [doi: 10.1016/j.physio.2013.09.002] [Medline: 24238699]
17. Corrado B, Ciardi G. Hypermobile Ehlers-Danlos syndrome and rehabilitation: taking stock of evidence based medicine: a systematic review of the literature. J Phys Ther Sci 2018 Jun;30(6):843-847. [doi: 10.1589/jpts.30.847] [Medline: 29950777]
18. Towards a common language for functioning, disability and health: the international classification of functioning, disability and health. : World Health Organization; 2002 URL: https://cdn.who.int/media/docs/default-source/classification/icf/icfbeginnersguide.pdf [accessed 2026-01-16]
19. Brittain MG, Flanagan S, Foreman L, Teran-Wodzinski P. Physical therapy interventions in generalized hypermobility spectrum disorder and hypermobile Ehlers-Danlos syndrome: a scoping review. Disabil Rehabil 2024 May;46(10):1936-1953. [doi: 10.1080/09638288.2023.2216028] [Medline: 37231592]
20. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. Peters MDJ, Godfrey C, McInerney P, et al. Best practice guidance and reporting items for the development of scoping review protocols. JBI Evidence Synthesis 2022;20(4):953-968. [doi: 10.11124/JBIES-21-00242]
22. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
23. Beighton P, de Paepe A, Danks D, et al. International nosology of heritable disorders of connective tissue, Berlin, 1986. Am J Med Genet 1988 Mar;29(3):581-594. [doi: 10.1002/ajmg.1320290316]
24. Beighton P, Paepe AD, Steinmann B, Tsipouras P, Wenstrup RJ. Ehlers‐Danlos syndromes: revised nosology, Villefranche, 1997. Am J Med Genet 1998 Apr 28;77(1):31-37. [doi: 10.1002/(SICI)1096-8628(19980428)77:1<31::AID-AJMG8>3.3.CO;2-P]
25. Kirk JA, Ansell BM, Bywaters EGL. The hypermobility syndrome. Musculoskeletal complaints associated with generalized joint hypermobility. Ann Rheum Dis 1967 Sep;26(5):419-425. [doi: 10.1136/ard.26.5.419] [Medline: 6039590]
26. Grahame R, Bird HA, Child A. The revised (Brighton 1998) criteria for the diagnosis of benign joint hypermobility syndrome (BJHS). J Rheumatol 2000 Jul;27(7):1777-1779. [Medline: 10914867]
27. Lima CA, Alsunaidi S, Lowe S, et al. Exploring the influence of weather variability and climate change on health outcomes in people living with dementia: a scoping review protocol. PLoS ONE 2024;19(6):e0304181. [doi: 10.1371/journal.pone.0304181] [Medline: 38913693]
28. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
29. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. ESRC Methods Programme 2006;1(1):b92. [doi: 10.13140/2.1.1018.4643]
30. Clark NL, Johnson M, Rangan A, Kottam L, Swainston K. The biopsychosocial impact of hypermobility spectrum disorders in adults: a scoping review. Rheumatol Int 2023 Jun;43(6):985-1014. [doi: 10.1007/s00296-023-05298-2] [Medline: 36894757]
31. Fletcher EKS, Fischer AL, Manocha RHK. Depressive symptoms are highly prevalent and associated with fatigue and pain catastrophizing in the hypermobility spectrum disorders and hypermobile Ehlers Danlos syndrome: a cross-sectional study. Rheumatol Int 2025 Apr 28;45(5):1-9. [doi: 10.1007/s00296-025-05869-5] [Medline: 40293579]

XSL•FO
RenderX

Abbreviations

hEDS: hypermobile Ehlers-Danlos syndrome
HSD: hypermobility spectrum disorder
ICF: International Classification of Functioning, Disability, and Health
MeSH: Medical Subject Headings
OSF: Open Science Framework
PA: physical activity
PCC: population, concept, and context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by J Sarvestan; submitted 09.Jul.2025; peer-reviewed by LK Anderson, NL Clark; revised version received 11.Dec.2025; accepted 12.Dec.2025; published 04.Feb.2026.

Please cite as:
Golden DW, Low JL, Daun JT, Jones CA, Dennett L, Culos-Reed SN, Manocha RHK
Safety and Impacts of Physical Activity for Individuals Living With Hypermobility Spectrum Disorders and Hypermobile Ehlers-Danlos Syndrome: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e80394
URL: https://www.researchprotocols.org/2026/1/e80394
doi:10.2196/80394

© Donald William Golden, Jonathan L Low, Julia T Daun, C Allyson Jones, Liz Dennett, Susan Nicole Culos-Reed, Ranita Harpreet Kaur Manocha. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 4.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Mapping the Landscape, Knowledge Gaps, and Areas for Innovation in Brain Health and Dementia Research in Canada: Protocol for a Scoping Review of Reviews

Juanita-Dawne Bacsu[1], PhD; Kiana Mero[2], BA; Megan E O'Connell[3], PhD; Megan Funk[1], MN; Alixe Ménard[4], MSc; Myrna Norman[5]; Sheila Blackstock[1], PhD; Jim Mann[6]; Wendy Hulko[7], PhD; Melba Sheila D'Souza[1], PhD; Sarah Fraser[4], PhD

[1]School of Nursing, Thompson Rivers University, 805 TRU Way, Kamloops, BC, Canada

[2]Department of Psychology, Thompson Rivers University, Kamloops, BC, Canada

[3]Department of Psychology & Health Studies, University of Saskatchewan, Saskatoon, SK, Canada

[4]Interdisciplinary School of Health Sciences, Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada

[5]National dementia advocate and person living with dementia, Maple Ridge, BC, Canada

[6]National dementia advocate and person living with dementia, Surrey, BC, Canada

[7]School of Social Work and Human Service, Thompson Rivers University, Kamloops, BC, Canada

Corresponding Author:
Juanita-Dawne Bacsu, PhD
School of Nursing, Thompson Rivers University, 805 TRU Way, Kamloops, BC, Canada

Abstract

Background: Dementia is one of Canada's most pressing public health challenges, with rates expected to surge in response to the country's aging population. Given the rapidly growing issue of dementia, understanding national research efforts is critical to prioritizing and advancing strategic directions in brain health and dementia research. Recently, the Canadian Institutes of Health Research awarded a 1-year funding grant from the Brain Health and Cognitive Impairment in Aging Research Initiative to map the scope of brain health and dementia research in Canada.

Objective: This scoping review of reviews protocol aims to address this call by outlining the methodology that will be used for mapping the research landscape, documenting the knowledge gaps, and identifying areas of innovation to advance brain health and dementia research in Canada.

Methods: Given the large volume of literature, a scoping review of Canadian-led reviews was selected as the most appropriate method because it would allow for a robust synthesis of nationally relevant research while mapping knowledge gaps and innovation. Our scoping review of reviews will follow the framework by Arksey and O'Malley along with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. The search will focus on peer-reviewed literature reviews published between January 1, 2020, and January 1, 2025, to capture the current state of knowledge since the national dementia strategy's publication in 2019. This search will be conducted using 5 electronic databases: CINAHL, PubMed, PsycInfo, Scopus, and Web of Science. Our data extraction table will include the following categories: author, province, and year; aim; review timeline; method; theme; knowledge gaps; innovations; and findings. The data will be analyzed using a combination of deductive and inductive thematic analysis.

Results: This protocol was registered on June 5, 2025, with the Open Science Framework. This study was funded by the Canadian Institutes of Health Research from November 2024 to November 2025. The anticipated timeline for the publication of the full scoping review of reviews is May 2026. The findings from this review will be shared through targeted knowledge mobilization activities such as presentations at national funding agency meetings, academic conferences, and community workshops.

Conclusions: Our scoping review of reviews will provide a robust synthesis of the brain health and dementia research landscape, helping document critical knowledge gaps and identify areas for innovation. The results of this research will provide critical data to help inform strategic funding initiatives and future research directions. The findings from our scoping review will have implications for research funders, policymakers, community organizations, and researchers that are working to accelerate brain health and dementia research across Canada.

Trial Registration: Open Science Framework A5BSX; https://osf.io/a5bsx/overview

International Registered Report Identifier (IRRID): PRR1-10.2196/79020

(JMIR Res Protoc 2026;15:e79020) doi:10.2196/79020

XSL•FO
RenderX

KEYWORDS

dementia; Alzheimer disease; prevention; treatment; quality of life; Canada

Introduction

Dementia is one of Canada's most pressing public health challenges, with rates expected to surge in response to the country's aging population. It is estimated that 771,939 Canadians live with dementia [1], and approximately 487,000 of them have received a diagnosis [2]. Advancing age is the greatest risk factor for dementia [3], with risk increasing to approximately 25% in adults aged 85 years and older (1 in 4 people) [4]. In 2023, the number of people aged 65 years or older was approximately 7.6 million, nearly one-fifth of Canada's entire population [5]. In Canada, over 414 new cases of dementia are diagnosed every day [1]. By 2030, it is estimated that almost 1 million people will live with Alzheimer disease or another form of dementia in this country [6].

The financial cost of dementia in Canada is substantial. In 2020, Canada's total direct and indirect costs of dementia were an estimated CAD $40.1 billion (US $29.3 billion) per year, with an average cost of CAD $67,200 (US $49,034) per person living with dementia [7]. Moreover, friends and family contribute over 580 million hours of care each year to people living with dementia [1]. Dementia care requires a comprehensive approach, including access to a timely dementia diagnosis, information and resources on treatment options, and support services for people living with dementia and their care partners [8].

Dementia is a leading cause of disability among older adults in Canada and worldwide [9]. By 2050, it is projected that over 1.7 million Canadians will live with dementia [1]. Accordingly, *A Dementia Strategy for Canada: Together We Aspire* emphasizes the importance of dementia research, highlighting the need to make advancements in 3 theme areas: dementia prevention, advancing therapies, and improving quality of life for people with dementia and their care partners [10].

There is extensive research on brain health and dementia in Canada. An initial PubMed search covering the past 5 years (2020 - 2025) using the terms "cognitive impairment," "cognitive decline," "dementia," "dementias," "Alzheimer's Disease," "Alzheimer's," "mild cognitive impairment," "brain health," and "Canadian [Affiliation] OR Canada [Affiliation]" yielded over 10,000 articles. Although numerous Canadian studies exist, there is a paucity of knowledge synthesizing of the brain health and dementia research landscape. Recently, the Canadian Institutes of Health Research awarded a 1-year funding grant from the Brain Health and Cognitive Impairment in Aging Research Initiative to map the scope of Canadian research on brain health and dementia [11].

This scoping review of reviews aims to address this call by mapping the research landscape, documenting knowledge gaps, and identifying areas of innovation to advance brain health and dementia research in Canada. Given the large volume of literature and the 1-year time frame of our funded study, a scoping review of Canadian-led reviews was selected as the most appropriate method because it will allow for a robust synthesis of nationally relevant research while mapping gaps and innovation. To provide a targeted and manageable approach, we will structure our review around Canada's national dementia strategy's 3 theme areas (eg, prevention, treatment, and quality of life) [10] while also ensuring that data are collected to identify research gaps and areas for innovation to inform strategic funding initiatives and future research priorities. The findings from our scoping review will have implications for research funders, policymakers, community organizations, and researchers that are working to elevate brain health and dementia research in Canada.

Methods

Scoping Review Framework

Canadian-led reviews were selected for this scoping review to maximize the national relevance of the findings and their applicability to mapping brain health and dementia research within the Canadian context. Scholars indicate that scoping reviews are highly useful for mapping large bodies of literature and identifying knowledge gaps to determine the breadth of the field [12,13]. Peters et al [13] note that scoping reviews are increasingly being used to inform decision-making and future research directions. This scoping review of reviews protocol was registered with the Open Science Framework (a5bsx) on June 5, 2025.

This scoping review of reviews will adhere to the framework by Arksey and O'Malley [14], comprising six steps: (1) establishing the research question, (2) exploration of relevant studies, (3) selection of studies, (4) extracting and charting the data, (5) synthesizing and reporting key findings, and (6) expert consultation. We will share our findings following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist [12].

Step 1: Establishing the Research Question

Our scoping review will be structured around the dementia strategy of Canada's 3 thematic areas: dementia prevention, advancing treatment and finding a cure, and enhancing the quality of life of people with dementia and their care partners [10]. This scoping review of reviews protocol will consist of primary and secondary research questions.

Primary Research Question

The primary research question is as follows: drawing on the national dementia strategy's 3 research areas (dementia prevention, treatment, and quality of life [10]), what are the research landscape, knowledge gaps, and areas of innovation within brain health and dementia research in Canada?

Secondary Research Questions

The secondary research questions are as follows:

- In the area of dementia prevention, what are the existing knowledge landscape, research gaps, and areas for innovation in brain health and dementia research in Canada?

XSL•FO
RenderX

- In the area of advancing treatments and finding a cure, what are the existing knowledge landscape, research gaps, and areas for innovation in brain health and dementia research in Canada?
- In the area of enhancing quality of life, what are the existing knowledge landscape, research gaps, and areas for innovation in brain health and dementia research in Canada?

Step 2: Exploration of Relevant Studies

Overview

Our search will review 5 electronic databases: CINAHL, PubMed, PsycInfo, Scopus, and Web of Science. In conducting scoping reviews, Peters et al [13] recommend the population, concept, and context (PCC) framework to help support the search strategy by providing meaningful inclusion criteria. Accordingly, our search will be guided by a list of keywords that are documented in our PCC table (Table 1). Moreover, the search string for the PubMed database will be as follows: ((“Literature review*” OR “Systematic Review*” OR “Scoping Review*” OR “Narrative Review*” OR “Review” OR “Rapid Review*” OR “Umbrella Review” OR “Meta-Analysis*” OR “Meta synthesis”) AND (“Cognitive Impairment” OR “Cognitive Decline” OR “Dementia*” OR “Dementias” OR “Alzheimer’s Disease*” OR “Mild Cognitive Impairment*” OR “Alzheimer’s” OR “Brain Health*“) AND (Canada[Affiliation] OR Canadian[Affiliation] OR Canadians[Affiliation] OR “in Canada”[Affiliation])).

The search time frame will focus on peer-reviewed literature reviews published between January 1, 2020, and January 1, 2025. This time frame was selected to focus specifically on the reviews published since the national dementia strategy of Canada was launched [10]. Additionally, we connected with a highly skilled librarian to review and assess our proposed databases and search strategy.

Table . Population, concept, and context criteria.

Element	Description	Keywords
Population	Adults with cognitive impairment or dementia	“Dementia,” “Dementias,” “Alzheimer’s Disease,” “Alzheimer’s,” “Cognitive impairment,” “Mild Cognitive Impairment,” “Brain Health,” “Cognitive Decline”
Concept	Literature review types	“Literature review,” “Systematic Review,” “Scoping Review,” “Narrative Review,” “Rapid Review,” “Umbrella Review,” “Meta-Analysis,” “Meta Synthesis”
Context	Canadian context for national relevance	“Canada [Author Affiliation],” “Canadian [Author Affiliation],” “Canadians [Author Affiliation],” “in Canada [Author Affiliation]”

Inclusion and Exclusion Criteria

Our inclusion criteria will consist of five components: (1) peer-reviewed literature reviews (systematic reviews, scoping reviews, narrative reviews, rapid reviews, umbrella reviews, meta-syntheses, or meta-analyses) that outline the search strategy; (2) a publication date between January 1, 2020, and January 1, 2025; (3) being written in the English or French languages; (4) first author with a Canadian affiliation to ensure national relevance; and (5) a focus on brain health and/or dementia research. The exclusion criteria will be gray literature and empirical research rather than reviews.

It is important to note that we recognize that umbrella reviews may overlap with other review types. To address this issue, umbrella reviews will only be included if they provide a clearly documented search strategy, which will allow us to identify and assess any overlap with other included reviews. Using this strategy will help us support methodological transparency while maintaining consistency across the different types of reviews.

Step 3: Selection of Studies

The titles and abstracts of the extracted reviews will be independently screened by 2 reviewers against the inclusion criteria. The same 2 reviewers will independently conduct the full-text screening of the reviews. Covidence (Veritas Health Innovation) will be used to manage all the literature reviews obtained in our search [15]. To document the review of the literature, we will follow the PRISMA-ScR guidelines [12] and create a flowchart to show the data filtering process. Any discrepancies regarding a review article’s inclusion will be resolved through discussion among the 2 reviewers to achieve consensus. However, any disagreements in which consensus cannot be reached will involve a third reviewer to facilitate discussion and reach consensus through communication among the 3 researchers.

Step 4: Extracting and Charting the Data

Covidence will be used to create a data extraction table that will aid in systematically mapping the data from the literature reviews. The data extraction table will include the following categories: author, province, and year; aim; timeline searched; method; theme (prevention, treatment, or quality of life); knowledge gaps; innovations; and findings (including conclusions). Two reviewers will conduct a pilot test of the data extraction table to ensure the clarity of the data extraction process.

Step 5: Synthesizing and Reporting Key Findings

Overview

The data will be analyzed using a combination of deductive and inductive thematic analysis. Specifically, a deductive approach will be used in which each review will be coded according to

XSL•FO
RenderX

the 3 pre-established themes in Canada's national dementia strategy: dementia prevention, advancing treatments and finding a cure, and enhancing quality of life [10]. Inductive thematic analysis will be performed to identify the underlying themes that exist within the data. For example, this process will be guided by the inductive thematic analysis framework by Braun and Clarke [16] and will include (1) familiarization and immersion with the data (eg, reading and rereading the information in the data extraction table), (2) creating initial codes (eg, developing clear and descriptive code names to capture segments of informative text), (3) establishing themes (eg, creating themes through the use of theme piles by organizing the codes into meaningful groups), (4) reviewing the themes (eg, ensuring that the themes are not redundant, are clear, and capture all the existing data), and (5) naming and defining the themes (eg, ensuring clarity and uniqueness in the theme names and definitions to avoid repetition among the themes). Guided by the framework by Braun and Clarke [16], the full research team will work to develop and refine the themes by assessing whether each theme is clearly understandable or too complex, determining whether more themes are needed, assessing whether the data adequately support each theme, and ensuring that no themes have been missed. The findings from the scoping review of reviews will be reported by adhering to and following the PRISMA-ScR guidelines [12].

Trustworthiness

Guided by the criteria of trustworthiness by Lincoln and Guba [17], this study will include measures of dependability, credibility, and transparency. Dependability will be established by providing a comprehensive audit trail, such as our PCC table with our search terms and identifying clear inclusion criteria for other researchers looking to reproduce our scoping review process. Credibility will be achieved by using a transparent process for selecting studies and documenting each stage of the review using a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart. Confirmability will be supported by having the same 2 researchers independently complete the title and abstract screening and the full-text review. As previously mentioned, a third reviewer will be brought in to handle any discrepancies during the data screening to further reduce issues related to researcher bias. Confirmability will also be supported by engaging in researcher triangulation and having the full research team involved in interpreting the data during the thematic analysis to minimize the potential for researcher bias.

Step 6: Expert Consultation

Our research team includes 2 people living with dementia who are national dementia advocates. They are involved in numerous research projects across Canada and have presented in podcasts, webinars, conferences, and government meetings. In this project, these research partners will be involved in all stages of our research, from refining the scoping review question to collaborating in knowledge mobilization strategies. For example, our research partners with lived experience will be involved in the development and refinement of our themes during the thematic analysis process. Additionally, they will attend our quarterly full-team meetings. Honoraria will be provided to them for their invaluable knowledge, time, and insights shared on the project.

Historically, people with lived experience have not often been included in research teams and have had little involvement in guiding research questions [18]. However, the inclusion of people living with dementia is essential to our work as it contributes to a more robust and critical lens to evaluate the landscape of brain health and dementia research. Brett et al [19] note that incorporating lived experience in reviews can be used to collectively identify knowledge gaps, develop research questions, assess the practical implications of the topic, and recognize relevant areas for future research. Furthermore, our research team consists of multidisciplinary professionals from across Canada with expertise in population health, nursing, social work, sociology, epidemiology, psychology, and family medicine.

Results

This protocol was registered on June 5, 2025, with the Open Science Framework. This study was funded by the Canadian Institutes of Health Research from November 2024 to November 2025. The scoping review's data collection and analysis will be conducted from August 2025 to September 2025. The anticipated timeline for the publication of the full scoping review of reviews is spring May 2026. Our knowledge mobilization strategy will be developed to connect with multiple stakeholders, including funders, policymakers, researchers, and community organizations involved in supporting brain health and dementia research. We will tailor our dissemination approaches to target and align with each stakeholder group. For example, we will share a presentation with the members of the Dementia Research and Innovation Funders Alliance, as well as with policymakers during their annual meeting in Ottawa, Ontario, to inform strategic funding directions and priorities [20]. Additionally, findings will be shared at academic conferences and in journal articles to target researchers. Newsletter articles and community workshops will be used to share findings with organizations at the community level. Our targeted knowledge mobilization strategy will help ensure that our dissemination activities are accessible, relevant, and user-friendly to each of the different audiences.

Discussion

Expected Findings

This scoping review of reviews aims to map the current landscape, knowledge gaps, and areas for innovation to advance brain health and dementia research in Canada. More specifically, this review is anticipated to provide an overview of the current brain health and dementia research ecosystem. By consolidating existing research, this review will reveal gaps in knowledge and illuminate areas for future research. Although our review may highlight policy and practice implications, our study will be conducted to map the current research rather than assess research projects, policies, or practices. By revealing research gaps, our review's findings will offer important insights for policymakers and research funders to help shed light on strategic areas for future research funding. Consequently, the findings from our

XSL•FO
RenderX

study will fuel future investigations and illuminate priority research areas that may require implementation, monitoring, evaluation, and scaling up.

Our protocol adheres to the framework by Arksey and O'Malley [14] and the PRISMA-ScR guidelines [13], ensuring a rigorous, transparent, and systematic methodology. Guided by the criteria of trustworthiness by Lincoln and Guba [17], steps will be taken to ensure dependability, credibility, and transparency throughout the study. Although our research is not without limitations, the findings from our scoping review will be invaluable in helping identify critical research gaps and areas for innovation in the areas of dementia prevention, treatments, and quality of life. The findings from this study will provide essential information for prioritizing strategic funding initiatives and future directions to accelerate brain health and dementia research in Canada.

Limitations

Although this scoping review of reviews will be conducted in a robust manner, it is important to note its limitations. For instance, as we will conduct a review of reviews, it is possible that some relevant studies will be excluded if they are not included in a review. Our selected electronic databases may also present limitations in that some relevant reviews may be overlooked if they are not indexed in the databases. Additionally, our search terms will not include each of the different types of dementia, which may limit the results of our search. Moreover, our search will only include reviews with Canadian authors, which may limit the generalizability of our findings. Accordingly, a future scoping review of reviews to examine the research on brain health and dementia worldwide would be informative. Another limitation is that our review only focuses on peer-reviewed publications from the years 2020 to 2025, which will only provide a snapshot in time. Consequently, future reviews could be augmented by ongoing updates to map the brain health and dementia research landscape longitudinally over time. Future researchers may consider developing an evolving scoping review of reviews to provide periodic updates to ensure that the evidence base remains timely and relevant for informing strategic priorities.

Conclusions

Our scoping review of reviews will provide a robust synthesis of the brain health and dementia research landscape, helping document critical knowledge gaps and identify areas for innovation to advance the field. The results of this research will provide critical data to help inform strategic funding initiatives and future research directions. The findings from our scoping review will have implications for research funders, policymakers, community organizations, and researchers that are working to accelerate brain health and dementia research in Canada.

Funding

This research is supported by the Canadian Institutes of Health Research Institute on Aging (451297) and the Canada Research Chairs Program (451097).

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Authors' Contributions

JDB conceptualized the study and drafted the initial manuscript with support in developing the research questions and methods from MEO, SF, MF, KM, AM, MSD, MN, JM, SB, and WH. All authors contributed to reviewing and approving the manuscript.

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 86 KB - resprot_v15i1e79020_app1.docx]

References

1. Dementia numbers in Canada. Alzheimer Society. 2025. URL: https://alzheimer.ca/en/about-dementia/what-dementia/dementia-numbers-canada [accessed 2025-01-24]
2. Dementia: overview. Government of Canada. 2025. URL: https://www.canada.ca/en/public-health/services/diseases/dementia.html [accessed 2025-11-09]
3. Dementia: risk factors and prevention. Government of Canada. 2025. URL: https://www.canada.ca/en/public-health/services/diseases/dementia/risk-factors-prevention.html [accessed 2025-11-05]
4. Dementia in Canada: summary. Canadian Institute for Health Information. URL: https://www.cihi.ca/en/dementia-in-canada/dementia-in-canada-summary [accessed 2025-06-03]
5. The older people are all right. Statistics Canada. 2024. URL: https://www.statcan.gc.ca/o1/en/plus/7059-older-people-are-all-right [accessed 2025-01-18]

XSL•FO
RenderX

6. Navigating the path forward for dementia in Canada. Alzheimer Society Canada. 2022. URL: https://alzheimer.ca/sites/default/files/documents/Landmark-Study-Report-1-Path_Alzheimer-Society-Canada_0.pdf [accessed 2025-11-08]
7. Dementia in Canada: economic burden 2020 to 2050. Canadian Centre for Economic Analysis. 2023. URL: https://www.cancea.ca/wp-content/uploads/2023/07/CANCEA-Economic-Impact-of-Dementia-in-Canada-2023-01-08.pdf [accessed 2026-01-25]
8. Bacsu J, Mateen FJ, Johnson S, Viger MD, Hackett P. Improving dementia care among family physicians: from stigma to evidence-informed knowledge. Can Geriatr J 2020 Dec;23(4):340-343. [doi: 10.5770/cgj.23.426] [Medline: 33282053]
9. Dementia. World Health Organization. 2025. URL: https://www.who.int/news-room/fact-sheets/detail/dementia [accessed 2025-01-20]
10. Public Health Agency of Canada. Government of Canada. URL: https://www.canada.ca/en/public-health.html [accessed 2025-01-20]
11. Notice of upcoming funding opportunity: summer 2024 ICS planning and dissemination grants: brain health and cognitive impairment in aging research initiative. Canadian Institutes of Health Resarch. 2025. URL: https://cihr-irsc.gc.ca/e/53923.html [accessed 2025-10-10]
12. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
13. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
14. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
15. The world's #1 systematic review tool. Covidence. URL: https://www.covidence.org/ [accessed 2025-08-06]
16. Braun V, Clarke V. Thematic Analysis: A Practical Guide: SAGE Publications; 2021.
17. Lincoln YS, Guba EG. Naturalistic Inquiry: SAGE Publications; 1985.
18. Beames JR, Kikas K, O'Gradey-Lee M, et al. A new normal: integrating lived experience into scientific data syntheses. Front Psychiatry 2021;12:763005. [doi: 10.3389/fpsyt.2021.763005] [Medline: 34777064]
19. Brett J, Staniszewska S, Mockford C, et al. Mapping the impact of patient and public involvement on health and social care research: a systematic review. Health Expect 2014 Oct;17(5):637-650. [doi: 10.1111/j.1369-7625.2012.00795.x] [Medline: 22809132]
20. Dementia research and innovation funders alliance. Canadian Institutes of Health Research. URL: https://cihr-irsc.gc.ca/e/53676.html [accessed 2025-08-05]

Abbreviations

PCC: population, concept, and context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by A Schwartz; submitted 13.Jun.2025; peer-reviewed by N Mirjan; revised version received 05.Jan.2026; accepted 07.Jan.2026; published 13.Feb.2026.

Please cite as:
Bacsu JD, Mero K, O'Connell ME, Funk M, Ménard A, Norman M, Blackstock S, Mann J, Hulko W, D'Souza MS, Fraser S
Mapping the Landscape, Knowledge Gaps, and Areas for Innovation in Brain Health and Dementia Research in Canada: Protocol for a Scoping Review of Reviews
JMIR Res Protoc 2026;15:e79020
URL: https://www.researchprotocols.org/2026/1/e79020
doi:10.2196/79020

© Juanita-Dawne Bacsu, Kiana Mero, Megan E O'Connell, Megan Funk, Alixe Ménard, Myrna Norman, Sheila Blackstock, Jim Mann, Wendy Hulko, Melba Sheila D'Souza, Sarah Fraser. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring the Well-Being, Adaptability, and Sense of Belonging of Undergraduate Nursing Students During the Transition From Simulation to Clinical Practice: Protocol for a Scoping Review

Marjolaine Dionne Merlin, RN, MScN, PhD; Monica McGraw, RN, MScN, PhD; Julie Renaud, RN, MScN; Sandrine Evelyne Roy

School of Nursing, Faculté des sciences de la santé et des services communautaires, Université de Moncton, 51 Avenue Antonine-Maillet, Moncton, NB, Canada

Corresponding Author:
Marjolaine Dionne Merlin, RN, MScN, PhD
School of Nursing, Faculté des sciences de la santé et des services communautaires, Université de Moncton, 51 Avenue Antonine-Maillet, Moncton, NB, Canada

Abstract

Background: The transition from university-based simulation learning to the clinical environment is a pivotal stage in undergraduate nursing education. This period can influence students' psychological well-being, adaptability, and sense of belonging within the clinical setting, which are essential dimensions to professional learning and patient safety. Although simulations aim to prepare students for clinical realities, the extent to which they support students' emotional and social readiness for real practice remains unclear.

Objective: This scoping review aims to map literature on undergraduate nursing students' well-being, adaptability, and sense of belonging during their transition from university-based simulation learning to clinical practice. Secondary objectives include identifying educational interventions and highlighting research gaps for future research.

Methods: The review will follow the JBI methodology for scoping reviews and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) reporting guidelines. A systematic search will be conducted in PubMed, CINAHL, APA PsycINFO, Scopus, and ScienceDirect for studies published between 2015 and 2025 in English or French. A gray literature search will also be performed in Google Scholar. Eligible studies will include empirical (quantitative, qualitative, or mixed method) and review papers exploring nursing students' experiences of transitioning from university-based simulation learning to real clinical settings. Two reviewers will independently screen titles, abstracts, and full texts; extract relevant data; and synthesize findings using thematic analysis.

Results: As of February 2026, data collection is projected to be completed by April 2026, and data analysis is expected to be finalized by June 2026. Results will summarize current definitions, measures, and interventions related to students' well-being, adaptability, and sense of belonging, as well as identify evidence gaps and conceptual trends in literature.

Conclusions: This scoping review will address a critical gap by clarifying how psychosocial dimensions of the transition from university-based simulation learning to clinical practice are conceptualized, measured, and supported in undergraduate nursing education, thereby informing the development of more comprehensive and learner-centered educational strategies.

International Registered Report Identifier (IRRID): DERR1-10.2196/86813

(JMIR Res Protoc 2026;15:e86813) doi:10.2196/86813

KEYWORDS

nursing students; simulation-based learning; clinical education; well-being; adaptability; sense of belonging; transition; scoping review

Introduction

The transition from university-based simulation learning to clinical practice marks a pivotal stage in undergraduate nursing education. Simulation has become a cornerstone of modern pedagogy, providing students with a safe and controlled environment to develop technical skills, clinical reasoning, and decision-making without compromising patient safety [1,2]. Through simulations, standardized patients, and immersive scenarios, educators strive to replicate the complexity of real-world care, equipping nursing students for the challenges of clinical experience. Yet, despite these advantages, moving from simulated settings to authentic clinical environments introduces challenges that go far beyond technical competence. This shift often demands significant psychological, social, and

XSL•FO
RenderX

emotional adjustments, which can profoundly shape students' learning experiences and professional growth [2].

Clinical placements expose students to unpredictable situations and interpersonal dynamics that differ markedly from the controlled conditions of university-based simulation learning. These changes can cause stress, anxiety, and feelings of uncertainty, particularly when students identify a gap between their simulated experiences and real-world expectations [3]. Research suggests that psychological well-being, adaptability, and sense of belonging are pivotal for successful integration into clinical settings, as these dimensions affect confidence, resilience, and ultimately patient safety [3,4]. Students who feel supported and included within the clinical team are more likely to engage actively in learning, demonstrate adaptability, and develop professional identity [5,6]. In contrast, lack of belonging or inadequate coping strategies can lead to disengagement, burnout, and attrition from nursing programs [6,7].

In this review, psychological well-being, adaptability, and sense of belonging are treated as distinct yet interrelated constructs. Psychological well-being is conceptualized as a multidimensional indicator of positive mental health encompassing emotional, psychological, and social dimensions, including life satisfaction, positive affect, and psychological functioning [8,9]. Although resilience is frequently examined alongside well-being in nursing student research, it is considered here as a related protective factor rather than a proxy for well-being. Adaptability is defined as the capacity to adjust cognitive, emotional, and behavioral responses to novel, uncertain, or changing academic and clinical demands [10]. In line with Waldeck et al [11], adaptability is distinguished from resilience in that it reflects ongoing adjustment processes rather than recovery following adversity. Sense of belonging is understood as a relational and context-dependent construct referring to individuals' perceptions of being accepted, valued, and included within academic, clinical, or institutional environments [3,6,12]. While a sense of belonging is recognized as an important determinant of both well-being and academic persistence, it remains analytically distinct from individual adaptive capacities and emotional regulation processes [13].

Although these constructs are sometimes used interchangeably across studies, this review treats them as analytically distinct but potentially overlapping concepts that constitute a psychosocial foundation supporting nursing students' competence development and professional socialization during the transition period.

Despite the recognized importance of these dimensions, evidence on how simulation prepares students for the psychosocial experiences of clinical practice remains insufficient. While simulation can enhance confidence and technical skills [14], its impact on emotional readiness and social integration is less clear. Some studies indicate that structured debriefing and reflective practice within simulation may promote resilience and coping strategies [14], yet others highlight persistent challenges when students encounter authentic clinical pressures [15]. Educational interventions such as orientation programs, mentorship, and peer support have been proposed to bridge this gap, but their effectiveness varies across contexts and remains underexplored in the literature.

Given these complexities, a comprehensive synthesis of existing research is needed to map what is known about nursing students' well-being, adaptability, and sense of belonging during the transition from university-based simulation learning to clinical practice. Scoping reviews are particularly suited for this purpose, as they allow for the examination of diverse study designs, identification of conceptual frameworks, and analysis of interventions aimed at supporting students during this critical phase [16,17]. This study aims to systematically review empirical and theoretical contributions to clarify how key psychosocial dimensions—well-being, adaptability, and sense of belonging—are defined, measured, and supported during the transition from university-based simulation learning to clinical practice. By mapping existing evidence, the review will identify educational strategies that influence these outcomes, highlight gaps in the literature, and propose directions for future research to inform curriculum design and policy development in nursing education. Specifically, the primary objective is to synthesize research on nursing students' experiences of this transition, while secondary objectives include examining interventions that foster psychosocial readiness and outlining areas requiring further investigation. To guide this work, the review will address three questions: (1) What is currently known about well-being, adaptability, and sense of belonging among undergraduate nursing students during the transition from university-based simulation learning to clinical practice? (2) Which interventions or contextual factors support these dimensions during this period? (3) What gaps persist in the literature, and what priorities should shape future research?

In summary, the transition from university-based simulation learning to clinical practice is a multidimensional process that shapes students' professional identity and capacity for safe, effective care. Understanding how well-being, adaptability, and sense of belonging interact during this period is essential for designing interventions that promote successful integration into clinical environments. This scoping review will provide an evidence-informed foundation for enhancing educational practices and supporting nursing students in navigating one of the most challenging stages of their academic journey.

Methods

Overview

This scoping review will follow the JBI methodological framework for scoping reviews and will be reported in alignment with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) checklist (Checklist 1). The review process will include the following stages: (1) defining the research question, (2) identifying relevant studies, (3) selecting eligible evidence, (4) charting the data, and (5) collating, summarizing, and reporting the results.

A comprehensive literature search will be conducted across major bibliographic databases and relevant gray literature sources (eg, dissertations, reports, association websites) to

XSL•FO
RenderX

capture studies published between 2015 and 2025 in English and French. The search will be designed to retrieve studies addressing concepts related to nursing students, university-based simulation learning, clinical practice, well-being, adaptability, and sense of belonging. Reference lists of included publications will also be reviewed to identify additional evidence. Any significant updates to the protocol will be documented and presented in the final scoping review report. An initial exploratory search across relevant databases confirmed that no systematic or scoping reviews currently exist on this specific topic, nor are any in progress.

For this review, the "transition from simulation to clinical practice" is defined as the process through which learners move from university-based simulation learning to real-world clinical practice settings. This transition is not limited to a single point in time or to the first clinical placement but occurs at multiple moments throughout the undergraduate nursing program, as learners progressively engage in clinical experiences following simulation-based learning activities.

Eligibility Criteria

Eligible studies include peer-reviewed empirical research (quantitative, qualitative, or mixed method) and review articles published in English or French between 2015 and 2025, reflecting the rapid evolution of university-based simulation learning in nursing over the past decade. Foundational studies published prior to 2015 informed the background of the review but will not be included in the formal synthesis. Review articles will be used primarily to map and clarify concepts, rather than to contribute quantitative or duplicated findings from primary studies. Data from reviews will support the identification of key constructs, definitions, and conceptual frameworks, while primary studies will remain the main source of empirical data for synthesis.

The papers must focus on undergraduate or prelicensure nursing students enrolled in nursing education programs and engaged in experiential learning activities such as clinical placements, simulation-based learning, or internships. Studies should explore the transition from university-based simulation learning to real clinical settings and may examine interventions such as orientation programs, mentoring, or structured debriefing that support psychological well-being, resilience, adaptability, sense of belonging, and integration into clinical practice. Comparisons across different learning environments (eg, simulation vs real-world practices) are acceptable, as are outcomes related to confidence, satisfaction, perceived competence, and other indicators of clinical learning.

Search Strategy

An academic librarian at the Université de Moncton was consulted by the first author to guide the development of the literature search strategy. The search was structured using the PICO framework, focusing on undergraduate nursing students (Population), transition from university-based simulation learning (Intervention), with or without comparison groups (Comparison), and outcomes related to well-being, adaptability, and sense of belonging (Outcomes). Based on prior knowledge and a preliminary search, a list of relevant keywords and synonyms was developed for each concept. This list was expanded using Medical Subject Headings (MeSH) terms and proximity operators where applicable. The finalized search strategy (Multimedia Appendix 1) was first tested in CINAHL to ensure it captured relevant studies, then adapted for other databases including PubMed, APA PsycINFO, Scopus, and ScienceDirect. Reference lists of included studies will be reviewed, and additional articles may be identified through snowball sampling during the data extraction phase.

Data Management

Following the literature search conducted in collaboration with the academic librarian, all identified references were first imported into the bibliographic management software Zotero [18]. From there, the records were exported and uploaded into Covidence [19] for screening. Duplicate entries were automatically removed by the Covidence system prior to the title and abstract screening process.

Study Selection

All studies will be reviewed to determine whether they meet the inclusion criteria defined in the Covidence software. The screening process will begin with a review of titles and abstracts, as outlined in the standard scoping review methodology. If the abstract is not immediately available but the title appears potentially relevant, the study will be advanced to the next stage for a full-text review. Titles and abstracts will be screened independently by two reviewers, blinded to each other's decisions. Studies will proceed to full-text screening when both reviewers agree, and any conflicts will be resolved by a third reviewer. Full-text screening will also be conducted independently by two blinded reviewers using Covidence, with disagreements resolved by a third blinded reviewer. Studies will be excluded at this stage if they do not meet all inclusion criteria, and reasons for exclusion will be documented and reported in the final review. The study selection process and results will be fully described in the final scoping review and illustrated using a PRISMA-ScR flow diagram. Additionally, a list of excluded studies along with the reasons for exclusion will be provided.

Data Extraction

Data will be extracted from the included studies using a structured tool developed by the review team. The preliminary version of the data extraction tool (Multimedia Appendix 2) will be pilot-tested on two studies by all reviewers before full implementation.

To address conceptual overlap across studies, operational definitions for well-being, adaptability, and sense of belonging will be established a priori and used to guide data extraction. Given the documented overlap among well-being, resilience, adaptability, and coping-related constructs [20], particular attention will be paid to authors' conceptual definitions, theoretical frameworks, and measurement instruments. Constructs such as resilience or flexibility will be coded separately and mapped analytically to the predefined constructs only when explicit conceptual alignment is supported by the authors' definitions or measurement tools. During data extraction, findings will be categorized according to the

XSL•FO
RenderX

construct explicitly targeted by each study, rather than interchangeable terminology.

To ensure reliability, two reviewers will independently extract data from each study. A third member of the research team will verify the extracted data and support discrepancy resolution through discussion. The extraction tool may be refined iteratively, with any modifications documented in the final review.

Data Analysis

Extracted data will be analyzed using a combination of descriptive statistics and narrative synthesis. Descriptive analyses will summarize the number of included studies, publication years, countries of origin, and target populations, which will be presented in tabular format.

The narrative synthesis will address the review objectives and summarize findings related to university-based simulation learning, transition support, well-being, adaptability, and sense of belonging, drawing on both qualitative and quantitative evidence. Overlapping constructs will be examined comparatively and acknowledged explicitly to preserve conceptual nuance, maintain analytical coherence, and accurately reflect the conceptual complexity of the field.

Thematic coding will be applied to categorize types of simulations, educational interventions, and outcomes to support synthesis and interpretation. Barriers and facilitators influencing outcomes will be identified, intervention effectiveness will be examined where applicable, and gaps in the literature will be highlighted. Themes will be developed through an iterative, inductive coding process and organized into higher-order categories through cross-study comparison. Coding will be reviewed collaboratively by the research team to ensure analytic rigor and consistency.

Credibility will be strengthened through duplicate coding and consensus-based discussions. Dependability and confirmability will be supported by a transparent, iterative analytic process with documented analytic decisions, while transferability will be facilitated through detailed reporting of study contexts and characteristics. Discrepancies will be resolved through consensus or consultation with a third reviewer when necessary.

Consistent with scoping review methodology, the methodological quality of included studies will not be formally assessed, as the purpose of this review is to map the breadth of available evidence.

Ethical Considerations

In accordance with institutional and national guidelines, research ethics approval was not sought for this study because it consists of a scoping review of publicly available literature. The study did not involve human participants, identifiable personal information, or primary data collection.

Results

This study was funded by the Faculty of Graduate Studies and Research at the Université de Moncton, with funding received in May 2025. The study was initiated in January 2025, and the final version of the protocol was completed in June 2025. A preliminary search of the published literature (excluding gray literature) was conducted in September 2025, resulting in 1194 articles after duplicate removal.

The formal literature search began in October 2025, with results expected by June 2026. We plan to submit a manuscript to *JMIR Nursing* to share our findings with an academic audience, aiming for submission in the summer of 2026. Additionally, we intend to present the results at conferences to promote broader dissemination. These preliminary results are provisional and will be updated following completion of the final search strategy.

Discussion

Expected Insights

This scoping review is anticipated to provide a structured synthesis of how undergraduate nursing students' psychosocial experiences are conceptualized, measured, and supported during the transition from university-based simulation learning to clinical practice. Preliminary examination of the literature suggests that, although simulation-based learning is well established as a strategy for developing technical competence and clinical reasoning, psychosocial dimensions of transition—in particular, well-being, adaptability, and sense of belonging—remain inconsistently defined and variably operationalized. By systematically mapping definitions, measurement tools, and interventions, the review will clarify how constructs such as well-being, adaptability, and sense of belonging are conceptualized and operationalized across diverse studies. Consistent with prior work, preliminary examination indicates that psychosocial readiness receives less explicit attention than technical competence, often appearing as a secondary or implicit outcome. This gap highlights the need for a structured overview to inform educational strategies and support frameworks.

Implications for Nursing Education

The findings will have direct implications for curriculum design and pedagogical approaches. Identifying interventions, such as structured debriefing, mentorship programs, and orientation sessions, that effectively promote resilience, emotional stability, and social integration will enable educators to implement evidence-based practices. In addition, mapping the range of measurement approaches, including self-report instruments, reflective tools, and emerging digital assessment methods, will clarify how psychosocial outcomes are evaluated and how these tools align with educational objectives and simulation-based pedagogy. Strengthening these dimensions is essential not only for academic success but also for patient safety and professional identity formation. By synthesizing diverse strategies and outcomes, this review will contribute to the development of transition models that integrate both technical and psychosocial aspects of clinical learning.

Research Gaps and Future Directions

The review is expected to reveal significant gaps, including inconsistent definitions of key constructs, variability in measurement approaches, and limited evaluation of intervention

XSL•FO
RenderX

effectiveness. Future research should prioritize longitudinal studies to assess the sustained impact of transition support strategies and explore innovative approaches, such as peer-assisted learning or digital tools, to enhance adaptability and belonging. Additionally, cultural and contextual factors influencing students' experiences should be examined to ensure interventions are applicable across different educational settings.

This protocol addresses a critical yet underexplored dimension of nursing education by systematically mapping the psychosocial factors of well-being, adaptability, and a sense of belonging during the transition from university-based simulation learning to clinical practice. The absence of existing reviews on this topic underscores its originality and the need for evidence-informed strategies to support students in navigating one of the most challenging phases of their academic journey.

Acknowledgments

We would like to thank our research assistants for their valuable contributions to the development of the protocol and the identification of preliminary studies. We also extend our gratitude to the academic librarian, Marc Harper, BA, MLIS, for his expertise and support in designing the search strategy.

Given that English is not the native language of the research team, we used an artificial intelligence tool to support the rephrasing and lexical refinement of certain passages, ensuring clarity and precision in word choice.

Disclosure of delegation to generative artificial intelligence (GAI): The authors declare the use of GAI in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GAI tools under full human supervision: proofreading and editing. The GAI tool used was ChatGPT. Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes. Declaration submitted by MDM.

Funding

The authors gratefully acknowledge the financial support received from the Faculty of Graduate Studies and Research at the Université de Moncton.

Data Availability

All data generated or analyzed during this study will be made available in the final published scoping review paper and its supplementary materials.

Authors' Contributions

MDM is leading the scoping review, and serves as the guarantor of the review. MDM and MM collaboratively developed the study protocol. The search strategy was designed by MDM in consultation with an academic librarian. MDM drafted the initial version of the scoping review protocol. MM, JR and SR contributed by reviewing the manuscript. All authors participated in revising, editing, and approving the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy.
[DOCX File, 23 KB - resprot_v15i1e86813_app1.docx]

Multimedia Appendix 2
Data extraction tool.
[DOCX File, 16 KB - resprot_v15i1e86813_app2.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 111 KB - resprot_v15i1e86813_app3.docx]

References

1. Cant RP, Cooper SJ. The value of simulation-based learning in pre-licensure nurse education: a state-of-the-art review and meta-analysis. Nurse Educ Pract 2017 Nov;27:45-62. [doi: 10.1016/j.nepr.2017.08.012] [Medline: 28843948]
2. Ko YJ, Kim SY. Transition shock experience of nursing students in clinical practice: a phenomenological approach. Healthcare (Basel) 2022 Mar 25;10(4):613. [doi: 10.3390/healthcare10040613] [Medline: 35455791]

XSL•FO
RenderX

3. Levett-Jones T, Lathlean J. Belongingness: a prerequisite for nursing students' clinical learning. Nurse Educ Pract 2008 Mar;8(2):103-111. [doi: 10.1016/j.nepr.2007.04.003] [Medline: 18291327]
4. Wille E, Opheim HMS, Princeton DM, Kisa S, Hjerpaasen KJ. Building resilience and competence in bachelor nursing students: a narrative review based on social cognitive theory. Nurs Rep 2025 Jul 11;15(7):253. [doi: 10.3390/nursrep15070253] [Medline: 40710947]
5. Dionne Merlin M, Lavoie S, Gallagher F. Elements of group dynamics that influence learning in small groups in undergraduate students: a scoping review. Nurse Educ Today 2020 Apr;87:104362. [doi: 10.1016/j.nedt.2020.104362] [Medline: 32065945]
6. Dionne Merlin M, Gallagher F, Lavoie S, Bilodeau C. Éléments qui influencent la dynamique de petits groupes et l'expérience d'apprentissage de stagiaires infirmières [Article in French]. ripes 2023;39(3). [doi: 10.4000/ripes.5201]
7. Sheikoleslami RL, Princeton DM, Mihaila Hansen LI, Kisa S, Goyal AR. Examining factors associated with attrition, strategies for retention among undergraduate nursing students, and identified research gaps: a scoping review. Nurs Rep 2025 May 22;15(6):182. [doi: 10.3390/nursrep15060182] [Medline: 40559473]
8. Chow KM, Tang WKF, Chan WHC, Sit WHJ, Choi KC, Chan S. Resilience and well-being of university nursing students in Hong Kong: a cross-sectional study. BMC Med Educ 2018 Jan 12;18(1):13. [doi: 10.1186/s12909-018-1119-0] [Medline: 29329529]
9. Li ZS, Hasson F. Resilience, stress, and psychological well-being in nursing students: a systematic review. Nurse Educ Today 2020 Jul;90:104440. [doi: 10.1016/j.nedt.2020.104440] [Medline: 32353643]
10. Xu Y, Zhang W, Wang J, Guo Z, Ma W. The effects of clinical learning environment and career adaptability on resilience: a mediating analysis based on a survey of nursing interns. J Adv Nurs 2024 Mar 11. [doi: 10.1111/jan.16144] [Medline: 38468419]
11. Waldeck D, Pancani L, Holliman A, Karekla M, Tyndall I. Adaptability and psychological flexibility: overlapping constructs? J Contextual Behav Sci 2021 Jan;19:72-78. [doi: 10.1016/j.jcbs.2021.01.002]
12. Levett-Jones T, Lathlean J, Higgins I, McMillan M. Staff-student relationships and their impact on nursing students' belongingness and learning. J Adv Nurs 2009 Feb;65(2):316-324. [doi: 10.1111/j.1365-2648.2008.04865.x] [Medline: 19191935]
13. Patel SE, Varghese J, Hamm K. Defining sense of belonging in nursing - an evolutionary concept analysis. J Prof Nurs 2024;54:151-163. [doi: 10.1016/j.profnurs.2024.07.003] [Medline: 39266084]
14. Oliveira Silva G, Oliveira FSE, Coelho ASG, et al. Effect of simulation on stress, anxiety, and self-confidence in nursing students: systematic review with meta-analysis and meta-regression. Int J Nurs Stud 2022 Sep;133:104282. [doi: 10.1016/j.ijnurstu.2022.104282] [Medline: 35679634]
15. Wu CS, Rong JR, Huang MZ. Factors associated with perceived stress of clinical practice among associate degree nursing students in Taiwan. BMC Nurs 2021 Jun 7;20(1):89. [doi: 10.1186/s12912-021-00602-6] [Medline: 34098953]
16. Peters MDJ, Marnie C, Tricco AC, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
17. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
18. Zotero. 2025. URL: https://www.zotero.org/ [accessed 2026-02-09]
19. Covidence.: Veritas Health Innovation; 2025. URL: https://www.covidence.org/ [accessed 2026-02-09]
20. Bowling J, Vercruysse C, Krinner LM, Greene T, Bello-Ogunu F, Webster C. A simultaneous concept analysis of resilience, coping, posttraumatic growth, and thriving. Nurs Forum 2022 Sep;57(5):905-919. [doi: 10.1111/nuf.12754] [Medline: 35643843]

Abbreviations

MeSH: Medical Subject Headings
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews

Edited by A Schwartz; submitted 11.Nov.2025; peer-reviewed by B Charles, G Khalaj; revised version received 26.Jan.2026; accepted 29.Jan.2026; published 13.Feb.2026.

Please cite as:
Dionne Merlin M, McGraw M, Renaud J, Roy SE
Exploring the Well-Being, Adaptability, and Sense of Belonging of Undergraduate Nursing Students During the Transition From Simulation to Clinical Practice: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e86813
URL: https://www.researchprotocols.org/2026/1/e86813
doi:10.2196/86813

XSL•FO
RenderX

© Marjolaine Dionne Merlin, Monica McGraw, Julie Renaud, Sandrine Evelyne Roy. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Shared Decision-Making With a Surrogate for Life-Sustaining Treatment of Critically Ill Patients: Protocol for a Scoping Review

Yoshiyasu Ito[1], MA; Mika Moriyama[2], PhD; Akemi Nasu[3], MA; Keiko Kamitani[4], PhD; Shimpei Hayashi[5], MSc; Satoko Ono[4], MSc; Yasuko Sumida[4], MSc; Keiko Matsumoto[5], PhD; Misae Ito[4], PhD

[1]Faculty of Nursing Science, Tsuruga Nursing University, 2-1, Kizaki 78, Tsuruga City, Fukui, Japan

[2]Faculty of Nursing, Kobe City College of Nursing, Kobe, Japan

[3]Department of Nursing, Sanyo Gakuen University, Okayama, Japan

[4]Graduate School of Medicine, Yamaguchi University, Ube, Japan

[5]Graduate School of Medicine, Kagawa University, Miki, Japan

Corresponding Author:
Yoshiyasu Ito, MA
Faculty of Nursing Science, Tsuruga Nursing University, 2-1, Kizaki 78, Tsuruga City, Fukui, Japan

Abstract

Background: Shared decision-making (SDM) is a collaborative process that integrates patients' values and preferences into health care decisions. In intensive care units, patients who are critically ill often lack the capacity to make decisions, necessitating surrogates to make complex choices regarding life-sustaining treatments (LSTs).

Objective: This scoping review aims to assess the range of research conducted on surrogate SDM for LSTs among patients who are critically ill over the past decade and highlight areas where current research remains limited.

Methods: This scoping review will follow the Joanna Briggs Institute methodology and adhere to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) reporting guidelines. Studies will be included if they examine SDM involving surrogates of adult patients who are critically ill in relation to LST decisions within intensive care unit settings. SDM is defined using 4 criteria: participation of both health care professionals and surrogates, mutual information sharing, consensus building, and agreement on treatment based on the patient's values and preferences. A comprehensive search will be performed across PubMed, CINAHL, PsycInfo, CENTRAL, and Ichushi-Web for English- and Japanese-language studies published between 2016 and 2025. Eligible study designs will include quantitative, qualitative, and mixed methods research. Title and abstract screening, as well as full-text selection, will be conducted independently by 2 reviewers using Rayyan. Data will be extracted on study characteristics, SDM definitions, participant roles, and key findings. Results will be synthesized descriptively and presented in tables and narrative summaries to identify research gaps and inform future investigations.

Results: As of June 13, 2025, the literature search has been completed. A total of 2899 citations were identified through the specified database searches, and 527 (18.2%) duplicates were removed. Title and abstract screening are currently in progress, and full-text review is expected to be completed by September 2025.

Conclusions: This scoping review will systematically map recent evidence on surrogate SDM in the context of LST decisions for patients who are critically ill. By synthesizing diverse studies, it will identify challenges faced by surrogates and summarize existing interventions that aim to improve SDM processes. The findings are expected to inform future interventions and policies and advance patient- and family-centered care in critical care settings.

Trial Registration: OSF Registries 10.17605/OSF.IO/MF2HT; https://osf.io/mf2ht/

International Registered Report Identifier (IRRID): PRR1-10.2196/83284

(JMIR Res Protoc 2026;15:e83284) doi:10.2196/83284

KEYWORDS

intensive care unit; ICU; life-sustaining treatment; protocol; scoping review; shared decision-making; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA

XSL•FO
RenderX

Introduction

Background

In intensive care units (ICUs), patients frequently lack the capacity to make decisions for themselves, and therefore, decisions regarding their health care are delegated to surrogates. To achieve patient-centered care, shared decision-making (SDM) remains essential even when surrogate decision-makers are involved. SDM has been described as a clinician-focused approach that facilitates the provision of family-centered care [1] and is recommended as the primary model for decision-making in the ICU [2]. SDM in the ICU is defined as follows: "SDM is a collaborative process that allows patients—or their surrogates—and clinicians to make healthcare decisions together, taking into account the best available scientific evidence as well as the patient's values, goals, and preferences" [2].

The process of surrogate decision-making in the ICU is frequently reported to be a highly challenging experience [3], with approximately 19% of ICU patient surrogates experiencing moderate to severe long-term decisional regret [4]. Furthermore, such negative experiences with surrogate decision-making have been shown to increase the long-term risk of psychological impairment among surrogates [5,6]. SDM has been shown to improve patient satisfaction and the quality of care [7], and it is also expected to alleviate the negative experiences associated with surrogate decision-making in the ICU, serving as a crucial component in supporting surrogates [1].

SDM is indicated when patients or their surrogates struggle to identify achievable goals or when a patient's clinical trajectory changes substantially as such treatment decisions warrant the use of SDM [8]. Patients in ICUs are critically ill, and situations involving multiple rational options are rare; however, in decisions regarding life-sustaining treatments (LSTs), SDM remains applicable because choices must be made among several possible interventions. Surrogates of patients who are critically ill are often required to make difficult decisions about LSTs, such as whether to withhold, withdraw, or escalate life-sustaining interventions. Decision-making regarding LSTs is highly complex and challenging as it requires balancing the patient's best interests, the surrogate's values, and their understanding of the patient's preferences. Consequently, surrogates frequently experience strong feelings of regret concerning their decisions [4].

In the ICU, where medical decisions often carry significant ethical and emotional implications, SDM serves as a critical framework for aligning medical care with patients' wishes. It is also a key approach to reducing decisional conflict among surrogates and guiding them through the process of making difficult choices [9,10]. More recently, a consensus framework has been developed to guide decision-making processes regarding the continuation or limitation of LSTs in ICU patients, and SDM has been increasingly recommended as an essential component of such decision-making [11].

Building on this growing emphasis on SDM in the ICU, previous systematic reviews have summarized evidence related to SDM in decisions about LSTs in the ICU [12-14]. However, these reviews primarily include studies conducted before a standardized definition of SDM in the ICU was established. For many years, the conceptualization of SDM remained inconsistent [15,16]. In recent years, research on SDM has progressed rapidly [17], and within the ICU context, the clarification of its definition and its wider adoption in both scholarly work and clinical practice did not occur until 2016 [2]. Although several studies examining SDM in decisions about LSTs in the ICU have been published in recent years [18-20], no recent scoping review has systematically assessed the current range of research or identified existing knowledge gaps.

Review Objectives

The aim of this scoping review is to clarify the extent and nature of the evidence on SDM involving surrogates in LST decisions for patients who are critically ill. This review seeks to provide valuable insights into the current state of SDM in critical care, identify existing knowledge gaps, and suggest directions for future research.

Review Questions

The research questions for this scoping review are as follows:

1. What studies have examined SDM involving surrogates in LST decisions for patients who are critically ill, and what are their key findings?
2. What types of SDM-related interventions have been implemented by health care professionals to support surrogate decision-making regarding LSTs for patients who are critically ill?

Methods

Study Design

This review aims to clarify the extent and nature of the evidence on SDM involving surrogates in LST decisions for patients who are critically ill. It does not aim to synthesize evidence for clinical decision-making or generate practice recommendations; rather, its primary objective is to map the existing body of evidence. Scoping reviews are an appropriate methodology for such purposes as they are designed to determine the scope and coverage of research on a given topic and provide a clear understanding of the volume, focus, and characteristics of the available evidence [21]. Therefore, a scoping review approach was considered the most suitable methodology for this study. The proposed scoping review will be conducted in accordance with the Joanna Briggs Institute methodology for scoping reviews [22], and it will be reported following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines (Checklist 1) [23]. This review protocol has been registered on the Open Science Framework [24].

Inclusion Criteria

Population

This scoping review will include studies on SDM involving surrogates who participate in treatment and care decision-making for adult patients aged ≥18 years. Surrogates include not only

XSL•FO
RenderX

family members such as spouses, children, or parents but also individuals legally designated to make health care decisions on the patient's behalf, including legal guardians and those holding durable power of attorney for health care.

Concept

This scoping review will include studies on SDM between health care professionals and surrogates of patients who are critically ill and lack decision-making capacity due to disease progression or functional impairment, including cognitive dysfunction, delirium, decreased level of consciousness, or other neurological impairments.

SDM for patients who are critically ill is defined as a collaborative process that enables patients—or their surrogates—and clinicians to make health care decisions together, taking into account the best available scientific evidence and the patient's values, goals, and preferences [2]. According to Charles et al [25], SDM is characterized by four core features: (1) at least 2 participants—the health care professional and the patient—are involved, (2) both parties share information, (3) both parties engage in building a consensus about the preferred treatment, and (4) an agreement is reached on the treatment to be implemented. Accordingly, in this scoping review, SDM will be defined as meeting all the following criteria: (1) at least 2 participants—the patient's surrogate and the health care professional—are involved; (2) both parties share information; (3) both parties take steps to build a consensus on the preferred treatment based on the patient's values, goals, and preferences; and (4) an agreement is reached on the treatment to be implemented.

Studies involving decision-making support interventions—such as decision coaching, decision aids, and question prompt sheets—will be included if these constitute part of the SDM process and meet the 4 criteria described above. Studies that do not fulfill the SDM definition criteria or that focus on SDM involving patients with decision-making capacity will be excluded.

Context

This scoping review will include studies examining the decision-making process surrounding LSTs for patients who are critically ill. LST does not have a universally accepted definition across countries or clinical guidelines. Comparative analyses indicate that the scope of LST varies substantially by jurisdiction and often includes interventions aimed at maintaining vital organ function or preventing imminent death [26]. International consensus statements similarly define LSTs as treatments such as invasive mechanical ventilation, vasopressor or inotrope support, renal replacement therapy, extracorporeal life support (extracorporeal membrane oxygenation), cardiopulmonary resuscitation, and artificial nutrition or hydration [27-29]. Given this heterogeneity, this scoping review will include studies that explicitly refer to treatment decisions as "life-sustaining treatment," and we will extract and map each study's operational definition of LST.

For the purpose of this review, LST-related decision-making encompasses not only decisions regarding the initiation of such treatments but also decisions to withhold or withdraw them. These decisions most commonly occur in ICUs, where patients are critically ill and often require advanced medical interventions to sustain life. The ICU context is particularly significant because the time available for decision-making is often limited and the decisions made may have profound implications for patient survival. As such, these decisions are highly complex and ethically challenging, involving judgments about whether to continue or discontinue interventions that directly affect life or death. Studies addressing treatments or care other than LSTs, those conducted in settings outside the ICU, or those spanning multiple clinical environments (eg, ICUs and surgical wards) will be excluded.

Types of Sources

This scoping review will consider both experimental and quasi-experimental study designs, including randomized controlled trials, nonrandomized controlled trials, before-and-after studies, and interrupted time-series studies. In addition, analytical observational studies—including prospective and retrospective cohort studies, case-control studies, and analytical cross-sectional studies—will be considered for inclusion. This review will also include descriptive observational study designs, such as case series, individual case reports, and descriptive cross-sectional studies. Studies focusing on qualitative data will also be considered, including but not limited to designs such as phenomenology, grounded theory, ethnography, qualitative description, action research, and feminist research. Although review articles will be excluded from the final analysis, the reference lists of all relevant review articles will be examined to identify additional eligible primary studies. Review articles, protocol papers, opinion papers, letters, editorials, commentaries, and gray literature (eg, conference proceedings, government documents, and institutional websites) will be excluded.

Search Strategy

Although only the literature search has been completed to date, a 3-step search strategy was used. An initial limited search of PubMed was conducted to explore the range of terminology used in the literature and identify key articles relevant to SDM in ICUs. During this preliminary search, the titles and abstracts of retrieved articles were reviewed to extract frequently used text words and indexing terms (MeSH [Medical Subject Headings] terms). These terms were analyzed and incorporated into the development of the final comprehensive search strategy. The text words contained in the titles and abstracts of the relevant articles, as well as the indexing terms used to describe those articles, were used to construct the full search strategy (Multimedia Appendix 1). Specific keywords and indexing terms from each database and information source were integrated into the overall search strategy. The databases searched included PubMed, CINAHL Ultimate (EBSCO), PsycInfo (EBSCO), CENTRAL, and Ichushi-Web (Japan Medical Abstracts Society). Studies published in English or Japanese between 2016 and 2025 were included. The database searches were completed on June 13, 2025.

XSL•FO
RenderX

Study Selection

Following the database searches, all identified citations will be collated and uploaded to the Rayyan software (Qatar Computing Research Institute), and duplicates will be removed. After a pilot screening, titles and abstracts will be reviewed by 2 independent reviewers to assess eligibility according to the inclusion criteria. Potentially relevant sources will be retrieved in full, and their citation details will be imported into Rayyan. Two independent reviewers will then assess the full texts of the selected citations in detail against the inclusion criteria. Reasons for excluding sources that do not meet the inclusion criteria will be documented and reported. Any disagreements between reviewers at any stage of the selection process will be resolved through discussion or, if necessary, consultation with an additional reviewer. The results of the search and study selection process will be presented using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) 2020 flow diagram [30]. The study selection process is scheduled to take place between June 2025 and August 2025.

Data Extraction

Data will be extracted using a standardized data extraction tool developed by the reviewers (the tool will be cited or appended). Extracted data will include the following: source details (eg, citation information, country, and study design), study purpose, definition of LST and SDM, types of health care professionals involved in SDM, context or setting, participants, methodology, and main findings. A draft data extraction form is provided in Multimedia Appendix 2. The extraction tool will be refined and updated as necessary during data extraction from each included source, and all modifications will be documented in the final scoping review. Any disagreements arising during data extraction will be resolved through discussion or consultation with an additional reviewer. Data extraction is scheduled to be conducted between September 2025 and October 2025 following the completion of the study selection process.

Data Analysis and Presentation

In this review, data will be primarily presented in tabular form; however, other formats that better illustrate the results will also be considered. A detailed mapping will be performed by describing each study within relevant domains, including country, study design, purpose, setting, participants, methodology, and key findings. When more than 2 studies address a single domain, the findings will be summarized narratively, and the evidence will be mapped descriptively. Data analysis and presentation are scheduled to be conducted between November 2025 and December 2025 following the completion of data extraction. The results of this scoping review are expected to be published from January 2026 onward.

Results

As of June 13, 2025, the literature search has been completed. A total of 2899 citations were identified through the specified database searches, and 527 (18.2%) duplicates were removed. Title and abstract screening (conducted by YI, MM, AN, KK, SH, SO, YS, KM, and MI) is currently in progress, and full-text review is expected to be completed by September 2025.

Discussion

Anticipated Findings

In ICUs, patients who are critically ill frequently lose the capacity to make health care decisions, and responsibility for these decisions often falls to surrogate decision-makers [2,31]. Although SDM has been endorsed as the preferred framework for aligning medical treatment with patient values, its application in surrogate decision-making for LSTs remains limited [32]. Previous research has underscored the complexity of surrogate decision-making, including decisional conflict, regret, and psychological burden [3,33-35]. However, evidence on how SDM processes specifically address or mitigate these challenges in the ICU remains fragmented. To our knowledge, this scoping review will be the first to systematically map the extent and nature of the evidence on SDM involving surrogates in LST decisions for patients who are critically ill.

This review will clarify how SDM has been conceptualized and operationalized in the ICU setting, identify the types of health care professional interventions that have been implemented, and delineate the roles of these professionals within SDM processes. In ICUs, SDM has been implemented as an interprofessional process that requires collaboration among physicians, nurses, and other health care professionals [36,37]. This review will further contribute to advancing understanding of SDM by clarifying the specific functions of each professional group within these collaborative processes. Moreover, by synthesizing diverse study designs and methodologies, this review will provide a comprehensive overview of the current state of knowledge and identify areas in which empirical evidence remains insufficient. The findings are expected to inform the development of future interventions aimed at enhancing surrogate support and strengthening the integration of SDM into critical care practice.

The results of this scoping review are expected to contribute to both clinical practice and research in several ways. First, by mapping existing interventions and the roles of health care professionals in surrogate SDM, this review will provide valuable insights into multidisciplinary approaches that may help reduce decisional burden. Second, it will help identify which components of SDM are most effective in improving surrogate experiences, such as structured communication tools, decision aids, and family conferences.

A major strength of this review is the use of the Joanna Briggs Institute methodology and adherence to the PRISMA-ScR reporting guidelines [22,23], which ensure a systematic and transparent approach to evidence mapping. The inclusion of a wide range of study designs encompassing both quantitative and qualitative research will facilitate a comprehensive understanding of the topic. Furthermore, by incorporating literature published in both English and Japanese, this review may capture cultural and contextual variations that enrich the interpretation of the findings.

However, several limitations should be acknowledged. Restricting the search to publications from 2016 onward may exclude earlier yet potentially relevant studies on SDM.

XSL•FO
RenderX

Similarly, limiting the review to English- and Japanese-language publications may result in the omission of findings from other cultural contexts where SDM practices differ. Another limitation is that, as a scoping review, this study will not involve an assessment of the methodological quality or risk of bias of the included studies, which may limit the strength of conclusions regarding intervention effectiveness. Despite these limitations, this review will represent an important initial step toward consolidating the existing evidence and identifying key gaps to guide future systematic reviews and interventional research.

Conclusions

Overall, this scoping review will systematically map recent evidence on surrogate SDM in the context of LST decisions for patients who are critically ill. By synthesizing diverse studies, it will identify the challenges faced by surrogates and summarize existing interventions that aim to improve SDM processes. The knowledge generated will not only inform the development of future interventions and policies but also advance patient- and family-centered care in critical care settings.

Acknowledgments

The authors declare the use of generative artificial intelligence (GAI) in the research and writing process. According to the Generative Artificial Intelligence Delegation Taxonomy [38], the following tasks were delegated to GAI tools under full human supervision: translation. The GAI tool used was ChatGPT-5.1 (OpenAI). Responsibility for the final manuscript lies entirely with the authors. GAI tools are not listed as authors and do not bear responsibility for the final outcomes.

Funding

This study was supported by research funds from Tsuruga Nursing University.

Data Availability

The datasets that will be generated or analyzed during this study will be available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: YI, MM, AN, KK, SH, SO, YS, KM, MI
Methodology: YI, MM
Project administration: YI
Resources: YI
Supervision: KM, MI
Writing – original draft: YI
Writing – review & editing: MM, AN, KK, SH, SO, YS, KM, MI

Conflicts of Interest

None declared.

Multimedia Appendix 1
Preliminary search strategy for PubMed.
[DOCX File, 24 KB - resprot_v15i1e83284_app1.docx]

Multimedia Appendix 2
Data extraction instrument.
[DOCX File, 17 KB - resprot_v15i1e83284_app2.docx]

Checklist 1
PRISMA-ScR checklist.
[PDF File, 197 KB - resprot_v15i1e83284_app3.pdf]

References

1. Hwang DY, Oczkowski SJ, Lewis K, et al. Society of critical care medicine guidelines on family-centered care for adult ICUs: 2024. Crit Care Med 2025 Feb 1;53(2):e465-e482. [doi: 10.1097/CCM.0000000000006549] [Medline: 39982184]
2. Kon AA, Davidson JE, Morrison W, et al. Shared decision making in ICUs: an American College of Critical Care Medicine and American Thoracic Society policy statement. Crit Care Med 2016 Jan;44(1):188-201. [doi: 10.1097/CCM.0000000000001396] [Medline: 26509317]

XSL•FO
RenderX

3. Wendler D, Rid A. Systematic review: the effect on surrogates of making treatment decisions for others. Ann Intern Med 2011 Mar 1;154(5):336-346. [doi: 10.7326/0003-4819-154-5-201103010-00008] [Medline: 21357911]
4. Andersen SK, Butler RA, Chang CCH, Arnold R, Angus DC, White DB. Prevalence of long-term decision regret and associated risk factors in a large cohort of ICU surrogate decision makers. Crit Care 2023 Feb 16;27(1):61. [doi: 10.1186/s13054-023-04332-w] [Medline: 36797793]
5. Davidson JE, Jones C, Bienvenu OJ. Family response to critical illness: postintensive care syndrome-family. Crit Care Med 2012 Feb;40(2):618-624. [doi: 10.1097/CCM.0b013e318236ebf9] [Medline: 22080636]
6. Johnson CC, Suchyta MR, Darowski ES, et al. Psychological sequelae in family caregivers of critically III intensive care unit patients. A systematic review. Ann Am Thorac Soc 2019 Jul;16(7):894-909. [doi: 10.1513/AnnalsATS.201808-540SR] [Medline: 30950647]
7. Glaría MG, Fernández MM, Salgado C, Hernández-Leal MJ. A Joanna Briggs Institute framework approach to shared decision making in end-of-life. Health Expect 2024 Oct;27(5):e70041. [doi: 10.1111/hex.70041] [Medline: 39373127]
8. Turnbull AE, Sahetya SK, Needham DM. Aligning critical care interventions with patient goals: a modified Delphi study. Heart Lung 2016;45(6):517-524. [doi: 10.1016/j.hrtlng.2016.07.011] [Medline: 27593494]
9. Cox CE, White DB, Hough CL, et al. Effects of a personalized web-based decision aid for surrogate decision makers of patients with prolonged mechanical ventilation. Ann Intern Med 2019 Mar 5;170(5):285-297. [doi: 10.7326/M18-2335] [Medline: 30690645]
10. Liu X, Humphris G, Luo A, et al. Family-clinician shared decision making in intensive care units: cluster randomized trial in China. Patient Educ Couns 2022 Jun;105(6):1532-1538. [doi: 10.1016/j.pec.2021.10.004] [Medline: 34657779]
11. Kerckhoffs MC, Senekal J, van Dijk D, et al. Framework to support the process of decision-making on life-sustaining treatments in the ICU: results of a Delphi study. Crit Care Med 2020 May;48(5):645-653. [doi: 10.1097/CCM.0000000000004221] [Medline: 32310619]
12. Kryworuchko J, Hill E, Murray MA, Stacey D, Fergusson DA. Interventions for shared decision-making about life support in the intensive care unit: a systematic review. Worldviews Evid Based Nurs 2013 Feb;10(1):3-16. [doi: 10.1111/j.1741-6787.2012.00247.x] [Medline: 22490044]
13. Hajizadeh N, Uhler L, Herman SW, Lester J. Is shared decision making for end-of-life decisions associated with better outcomes as compared to other forms of decision making? A systematic literature review. MDM Policy Pract 2016;1(1):2381468316642237. [doi: 10.1177/2381468316642237] [Medline: 30288399]
14. Kerckhoffs MC, Kant M, van Delden JJ, Hooft L, Kesecioglu J, van Dijk D. Selecting and evaluating decision-making strategies in the intensive care unit: a systematic review. J Crit Care 2019 Jun;51:39-45. [doi: 10.1016/j.jcrc.2019.01.029] [Medline: 30738286]
15. Stiggelbout AM, Pieterse AH, De Haes J. Shared decision making: concepts, evidence, and practice. Patient Educ Couns 2015 Oct;98(10):1172-1179. [doi: 10.1016/j.pec.2015.06.022] [Medline: 26215573]
16. Lee KG, Gagliardi AR, Lofters AK, Pinto AD, Maunder RG. Shared decision-making, the working alliance, and patient-centered care: a simultaneous concept analysis and review of the literature. Patient Educ Couns 2025 Dec;141:109316. [doi: 10.1016/j.pec.2025.109316] [Medline: 40876260]
17. Lu C, Li X, Yang K. Trends in shared decision-making studies from 2009 to 2018: a bibliometric analysis. Front Public Health 2019;7:384. [doi: 10.3389/fpubh.2019.00384] [Medline: 31921749]
18. Choi HR, Rodgers S, Tocher J, Kang SW. Nurse's, physician's and family member's experiences of withholding or withdrawing life-sustaining treatment process in an intensive care unit. J Clin Nurs 2023 Aug;32(15-16):4827-4842. [doi: 10.1111/jocn.16556] [Medline: 36217241]
19. Fleming V, Prasad A, Ge C, et al. Prevalence and predictors of shared decision-making in goals-of-care clinician-family meetings for critically ill neurologic patients: a multi-center mixed-methods study. Crit Care 2023 Oct 21;27(1):403. [doi: 10.1186/s13054-023-04693-2] [Medline: 37865797]
20. Joynt GM, Ling SK, Chang LL, et al. End-of-life practices in Hong Kong intensive care units: results from the Ethicus-2 study. Hong Kong Med J 2024 Aug;30(4):300-309. [doi: 10.12809/hkmj2310944] [Medline: 39143753]
21. Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
22. Peters MD, Godfrey C, McInerney P, et al. Development of scoping review protocol (2020 version). In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: Joanna Briggs Institute; 2020.
23. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
24. Shared decision-making with a surrogate for life-sustaining treatment of critically ill patients: a scoping review protocol. Open Science Framework. URL: https://osf.io/MF2HT/overview [accessed 2025-05-25]
25. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997 Mar;44(5):681-692. [doi: 10.1016/S0277-9536(96)00221-3]

XSL•FO
RenderX

26. Tanaka M, Kodama S, Lee I, Huxtable R, Chung Y. Forgoing life-sustaining treatment - a comparative analysis of regulations in Japan, Korea, Taiwan, and England. BMC Med Ethics 2020 Oct 16;21(1):99. [doi: 10.1186/s12910-020-00535-w] [Medline: 33066771]
27. Truog RD, Campbell ML, Curtis JR, et al. Recommendations for end-of-life care in the intensive care unit: a consensus statement by the American College of Critical Care Medicine. Crit Care Med 2008 Mar;36(3):953-963. [doi: 10.1097/CCM.0B013E3181659096] [Medline: 18431285]
28. Beil M, van Heerden PV, Joynt GM, et al. Limiting life-sustaining treatment for very old ICU patients: cultural challenges and diverse practices. Ann Intensive Care 2023 Oct 27;13(1):107. [doi: 10.1186/s13613-023-01189-8] [Medline: 37884827]
29. Darfelt IS, Neergaard MA, Klepstad P, et al. Consensus statements on end-of-life care in ICU - a Scandinavian multidisciplinary Delphi study. Acta Anaesthesiol Scand 2025 Apr;69(4):e70015. [doi: 10.1111/aas.70015] [Medline: 40066825]
30. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. [doi: 10.1136/bmj.n71]
31. Torke AM, Sachs GA, Helft PR, et al. Scope and outcomes of surrogate decision making among hospitalized older adults. JAMA Intern Med 2014 Mar;174(3):370-377. [doi: 10.1001/jamainternmed.2013.13315] [Medline: 24445375]
32. Lei Y, Zhou Q, Tao Y. Decision aids in the ICU: a scoping review. BMJ Open 2023 Aug 22;13(8):e075239. [doi: 10.1136/bmjopen-2023-075239] [Medline: 37607783]
33. Miller JJ, Morris P, Files DC, Gower E, Young M. Decision conflict and regret among surrogate decision makers in the medical intensive care unit. J Crit Care 2016 Apr;32:79-84. [doi: 10.1016/j.jcrc.2015.11.023] [Medline: 26810482]
34. Sui W, Gong X, Qiao X, et al. Family members' perceptions of surrogate decision-making in the intensive care unit: a systematic review. Int J Nurs Stud 2023 Jan;137:104391. [doi: 10.1016/j.ijnurstu.2022.104391] [Medline: 36442321]
35. Gunnlaugsdóttir T, Jónasdóttir RJ, Björnsdóttir K, Klinke ME. How can family members of patients in the intensive care unit be supported? A systematic review of qualitative reviews, meta-synthesis, and novel recommendations for nursing care. Int J Nurs Stud Adv 2024 Dec;7:100251. [doi: 10.1016/j.ijnsa.2024.100251] [Medline: 39555387]
36. Michalsen A, Long AC, DeKeyser Ganz F, et al. Interprofessional shared decision-making in the ICU: a systematic review and recommendations from an expert panel. Crit Care Med 2019 Sep;47(9):1258-1266. [doi: 10.1097/CCM.0000000000003870] [Medline: 31169620]
37. DeKeyser Ganz F, Engelberg R, Torres N, Curtis JR. Development of a model of interprofessional shared clinical decision making in the ICU. Crit Care Med 2016 Apr;44(4):680-689. [doi: 10.1097/CCM.0000000000001467] [Medline: 26565629]
38. Suchikova Y, Tsybuliak N, Teixeira da Silva JA, Nazarovets S. GAIDeT (Generative AI Delegation Taxonomy): a taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. Account Res 2025 Aug 8;2025:1-27. [doi: 10.1080/08989621.2025.2544331] [Medline: 40781729]

Abbreviations

ICU: intensive care unit
LST: life-sustaining treatment
MeSH: Medical Subject Headings
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
SDM: shared decision-making

Edited by A Mavragani; submitted 31.Aug.2025; peer-reviewed by DY Hwang, J Vick; revised version received 28.Nov.2025; accepted 24.Dec.2025; published 21.Jan.2026.

Please cite as:
Ito Y, Moriyama M, Nasu A, Kamitani K, Hayashi S, Ono S, Sumida Y, Matsumoto K, Ito M
Shared Decision-Making With a Surrogate for Life-Sustaining Treatment of Critically Ill Patients: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e83284
URL: https://www.researchprotocols.org/2026/1/e83284
doi:10.2196/83284

© Yoshiyasu Ito, Mika Moriyama, Akemi Nasu, Keiko Kamitani, Shimpei Hayashi, Satoko Ono, Yasuko Sumida, Keiko Matsumoto, Misae Ito. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 21.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium,

XSL•FO
RenderX

provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Exploring Medical Students' Representations of Future Specialties and Parenthood: Protocol for a Scoping Review

Sylvie Arnoux[1], MSc; Mia Gisselbaek[1,2], MD; Georges Louis Savoldelli[1,2], MEd, MD; Nadia Masood Bajwa[1,3], MHPE, MD, PhD

[1]Unit of Development and Research in Medical Education, Faculty of Medicine, University of Geneva, 1 rue Michel Servet, Geneva, Switzerland

[2]Department of Anaesthesiology, Pharmacology, Intensive Care and Emergency Medicine, Division of Anaesthesiology, University Hospital of Geneva, Geneva, Switzerland

[3]Department of Women, Children, and Adolescents, Division of General Pediatrics, University Hospital of Geneva, Geneva, Switzerland

Corresponding Author:
Sylvie Arnoux, MSc
Unit of Development and Research in Medical Education, Faculty of Medicine, University of Geneva, 1 rue Michel Servet, Geneva, Switzerland

Abstract

Background: Several factors come into consideration when medical students choose their future specialty. Among these factors, the desire to start a family and planning the best timing for pregnancy may interfere with career advancement in certain specialties.

Objective: To the best of our knowledge, this is the first scoping review aimed at understanding medical students' career choice and parental expectations without restriction of the specialty chosen. This protocol describes a scoping review aiming to understand how representations regarding specialties and parenthood influence medical students' career choice. This protocol describes a scoping review aiming to understand how representations regarding specialties and parenthood influence medical students' career choice.

Methods: We will search PubMed, Embase, Web of Science, ERIC, and PsycInfo for literature. Additionally, the reference lists of included articles will be screened for further inclusion. Rayyan and Endnote will be used to organize data screening and extraction. The database selection will allow us to extract and analyze data from various disciplines. This diversity will increase our understanding of medical students' career and personal life decisions. This protocol and the upcoming scoping review have been designed following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines to ensure the quality of the searching process, the data screening, and the data extraction.

Results: This study will conduct a thematic synthesis of how the concepts of representations and perceptions of parenthood are used by medical students in the selected literature, comparing them to theoretical frameworks to clarify their meanings. We also plan to identify key themes related to parenthood and medical specialty choice when planning a career. As of December 2025, we proceeded to data screening. We anticipate publishing our results in the second quarter of 2026.

Conclusions: This scoping review aims to better understand medical students' representations of medical specialties and parenthood, and how these perceptions influence their specialty preferences and career choices. By mapping existing evidence across various disciplines, the review will identify research gaps and provide a foundation for future studies. The findings will offer valuable insights into the challenges of balancing career aspirations and family life, particularly in the context of physician shortages and the growing feminization of the medical profession.

International Registered Report Identifier (IRRID): DERR1-10.2196/78133

(JMIR Res Protoc 2026;15:e78133) doi:10.2196/78133

KEYWORDS

medical student; parenthood; maternity; specialty choice; representations; career choice; scoping review; family planning

Introduction

As medical students complete their studies and start their residency, several factors come into consideration when choosing a career path. For both men and women, initial career intentions may waver as personal life factors are taken into consideration [1-4].

As they complete their initial training, female students often anticipate that their work and their personal life may conflict [1]. When planning their future career, female medical students may take into consideration their desire to become a parent, the best time to have a child, parental obligations, and the tensions created by these responsibilities [5,6]. These reflections lead medical students to delay pregnancy despite the increased risk

XSL•FO
RenderX

of infertility and pregnancy complications [7]. Furthermore, becoming a mother may expose women physicians to discrimination, lack of career opportunities, and limited support from the institution [6,8,9]. Working part-time to take care of children may diminish career opportunities or prolong residency training [10,11]. Social norms, medical culture, and institutional structure and policies are possible sources of this discrimination [8,12,13].

Supervisors and more senior colleagues encountered in the clinical setting are role models for medical students. Role models help students project themselves in a desired specialty [5,14]. Experiences during clerkships and residency allow medical students to synthetize what they see and hear to build their representations [14-16]. However, a lack of role models, especially female ones, can affect how female medical students perceive women's position within the specialty and its environment [5,17,18]. The lack of female mentors can also affect learning and students' motivation [19]. Furthermore, students might encounter more fathers than mothers among their colleagues, especially in specialties where women are underrepresented [20].

Colleagues and supervisors' opinions regarding a specialty and the commitment it demands, their banter, and their comments about their past personal experiences influence students' projections about their future career [5,21-23]. Among those remarks, comments on how difficult it can be to reconcile maternity and career advancement affect the construction of career representations and motivation to pursue a specific specialty [2,17,24]. Such negative attitudes from colleagues contribute to reinforcing negative stereotypes, which in turn generate stereotype threats [25]. Female physicians might embody these negative stereotypes and unconsciously identify themselves with those stereotypes and expose themselves to self-sabotaging [25]. These factors underline the importance of conducting a scoping review on medical student career choices and the influence of parenthood aspirations.

As medical students progress through their training, they face the challenge of envisioning their future as doctors while balancing personal aspirations and life goals. When choosing a specialty, they must take into consideration several factors, such as family planning and the requirements of residency. Starting a career while becoming a parent can be very challenging.

Although there are a number of studies addressing career choice in relation to physician shortages, little is known about the aspirations and representations of medical students regarding their future career as a doctor and their private life. A systematic review has been conducted on experiences and perspectives of women who are already doctors and mothers and focuses on how the family life of trained doctors and their career conflict [6]. Another review has focused on surgery and how the working environment may generate specific challenges, such as long hours working and prolonged periods of standing, and higher rates of pregnancy complications [26]. The use of the terms representations or perception often lacks definition or operationalization in the literature. Drawing on social representation theory will help better define and conceptualize these terms [27-29].

To the best of our knowledge, no comprehensive review has synthetized the intersection of medical students' career perspectives, family planning, and specialty choice. Understanding these interconnected factors is crucial as they influence workforce planning, the well-being of future health care professionals, and the alignment of personal and professional aspirations. The findings may be useful to decrease the attrition of junior doctors when entering practice.

A scoping review allows for a comprehensive exploration of the existing literature, identifying gaps and patterns related to how parental responsibilities and representations influence career trajectories in medicine. Such a review can provide valuable insights into the societal, institutional, and cultural narratives that shape medical students' decisions.

We intend to synthesize existing literature on the representations and aspirations of family planning and career choice of medical students across medical specialties to answer the following research question: How do these representations influence medical students' career choice and desire to become parents?

Methods

Overview

Since we are interested in understanding how medical students project themselves into their career and parenthood and we aim to explore evidence on the topic, we will conduct a scoping review. The review will adhere to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) principles [30,31].

Preliminary searches occurred between August and November 2024 to establish the search strategy. This iterative process helped refine the choice of keywords and relevant databases.

The present protocol may undergo further revision as designing and writing a scoping review is an iterative process. Deviations from the protocol will be explained in the future manuscript to ensure transparency.

Information Sources

The following databases were searched: PubMed, Embase, Web of Science, ERIC, and PsycInfo. Additionally, reference lists of included articles and Google Scholar results have been screened for further inclusion. We focused on peer-reviewed data to map evidence broadly. Excluding gray literature aligns with this exploratory purpose, emphasizing breadth and clarity over exhaustive comprehensiveness. The data collection took place in February and March 2025. The search strategy was elaborated with the help of a librarian specializing in systematic searches. An example of the search strategy is available in Multimedia Appendix 1. The key concepts used to develop the search strategy are outlined in Table 1. The specific thesauri of each database were consulted to identify the most appropriate terminology. A selection of relevant articles was established to test the accuracy of the search strategy.

XSL•FO
RenderX

Table . Key concepts.

Subject		Becoming a parent
Med* student*	Specialt*	Parent*
Med* school student*	Career choice	Motherhood
Medical education	Occupational choice*	Fatherhood
	Occupational aspiration*	Family planning
	Job selection*	Life change*
	Career aspiration*	Having children
	Career goal*	Childbearing
	Vocational aspiration*	Maternity
	Goal orientation*	Pregnancy

Inclusion Criteria

To be included, studies must report on medical students and their career choice, family planning, and representations about having a family while starting a career (Textbox 1). As our goal is to map existing evidence regarding our subject, inclusion criteria do not include geographic area, a specific specialty, or a language. This choice reflects the diversity of existing situations and representations regarding the interaction between parenthood, being a doctor, and medical specialties. As our work is focused on medical students, we aim to explore recent and representative problematics. The review includes publications from 2000 to the present. Studies are excluded if they focus only on medical residents, nursing students, or physicians; on work-life balance only; or on policies.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria

- Focusing on medical students
- Address parenthood aspiration/desire/representation
- Any geographical area
- Any language
- Published from 2000

Exclusion criteria

- Focusing on residents, physicians, or nursing students
- Medical students' parents
- Focus on policies
- Published before 2000

Representations here are defined as values, beliefs, and experiences that interact to help individuals make sense of their experiences [27]. These factors can be shared and built collectively, but individuals interpret them differently. Furthermore, these representations are embedded in a sociocultural and historical context that encompasses various ways of sharing and transmitting representations. Itkonsen et al [29] have previously used social representations to understand the factors related to career choice and planning among university students. Expectations and stereotypes of the role of a doctor and daily life in a specific specialty shape the representation of their career before starting clinical rotations. As medical students are confronted with the reality of the clinical ward and private practice, their experiences enrich their representations. Sharing experiences with their peers also contributes to building their representations. A similar process occurs when students project themselves in their future career and personal lives.

The research team has an academic understanding of French, English, Italian, and Spanish. To ensure best understanding, titles and abstracts of studies published in another language were translated into English using DeepL [32]. The full article will be translated if it is included after the screening process. Both the original and translated version will be used for data extraction. The translated articles will be assigned to researchers based on their comprehension of the original language. Full documentation of the process will be detailed using the PRISMA-ScR chart [31].

A vocabulary list has been established for each key concept and adjusted to the selected database.

XSL•FO
RenderX

Data Screening

The search results were exported in Rayyan [33], a program designed to conduct systematic reviews, identify and remove duplicates, and provide a comprehensive list of all criteria and associated questions. Preliminary research indicated that only a relatively small number of articles were likely to be found from the research equations, enabling all authors (SA, NMB, MG, and GLS) to collaboratively proceed with the title and abstract screening to ensure consistency. The previously defined eligibility criteria were applied to determine whether each article should be included. Regular meetings were held to discuss the inclusion or exclusion of the articles for which no consensus has emerged during the iterative process.

The selected articles were exported to EndNote 21 (Clarivate) [34] and underwent a full-text screening. Any studies recommended for exclusion at this stage were reviewed by SA, NMB, and MG to ensure that exclusion criteria were met. Reference lists of each selected article will be screened to ensure that relevant references are included. Adjustments to the inclusion or exclusion criteria were discussed and agreed upon among the authors throughout the search and screening process.

Data Extraction

Articles meeting the inclusion criteria after the first screening will be listed in an Excel file (Microsoft Corp) providing key elements regarding their content. First, studies will be sorted by characteristics, including country of origin, study design, and data collected, if any. Then, studies will be thematically sorted. Themes of interest involve the conciliation of career choice and desires for parenthood, representations regarding medical specialties and their openness to parenthood, specific challenges regarding either motherhood or fatherhood, or representations about medical specialties and parenthood in general. This list has been established based on the literature cited in the Introduction and findings gathered in interviews during an ongoing project (S Arnoux et al, MSc, unpublished, 2024 data).

We will conduct a critical appraisal of individual sources of evidence to assess their relevance, reliability, validity, and applicability. The Joanna Briggs Institute provides a quality assessment tool specific to each study design type included in the studies [35]. This step will increase the quality of the scoping review and facilitate data collection and analysis.

SA and NMB will perform the data extraction independently. GLS and MG will independently check 20% of the extracted data to ensure that extraction criteria are met. Two meetings will be held at the beginning and the end of the extraction process to define the theme list and discuss the completed chart to reach consensus regarding the extracted data. If necessary, the themes list will be discussed and adjusted during the extraction process to include emerging themes. A final version of the chart will be provided in the scoping review.

Ethical Considerations

Ethics approval will not be necessary as this project does not fall under the Swiss law for human research (LRH) [36]. The scoping review will be published in a scientific journal and its results will be presented at conferences. This scoping review is part of a bigger project focusing on Swiss medical students' career choices. This project aims to understand and describe the factors influencing the career choices of medical students to inform stakeholders and address the workforce shortage.

Results

Representations and perceptions are often used in literature but rarely conceptualized. We intend to conduct a thematic synthesis of the terms used in the selected studies and compare this with the theoretical literature to provide a synthesis of what is understood when speaking of representations. Furthermore, we will describe the main themes associated with parenthood and medical specialty choice.

A total of 32 articles have been identified (Figure 1). As of December 2025, data extraction has started. We expect to publish the results in the second quarter of 2026. This project obtained a grant from the Swiss National Science Foundation in October 2022.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow chart for scoping reviews [37].

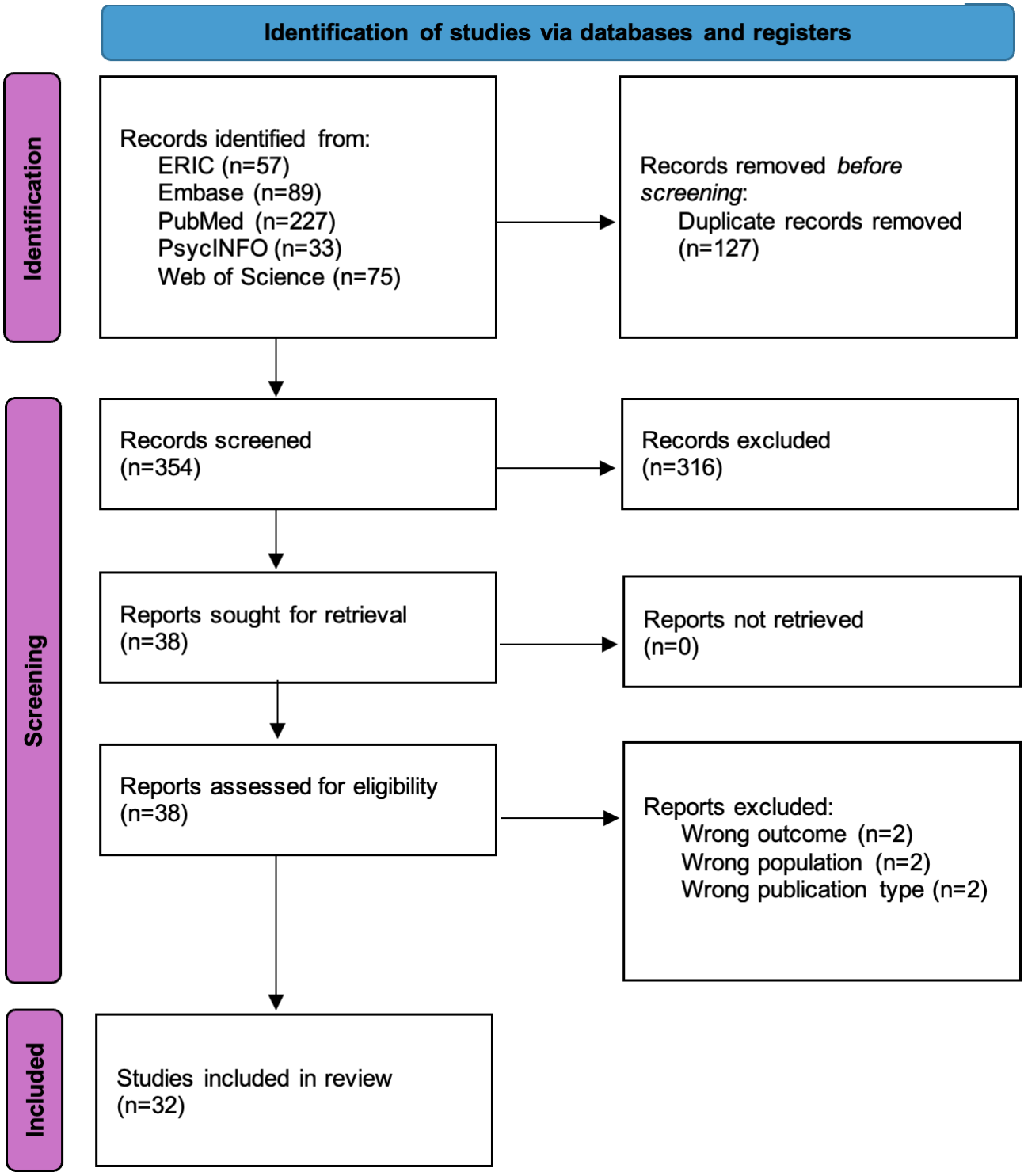

Discussion

This scoping review aims to provide a better understanding of medical students' representations of medical specialties and parenthood. The main objective is to improve our comprehension of the impact of their representations on their specialty preferences and career choices.

Previous studies have focused on specific specialties (such as surgery) and how physicians, residents, or students make decisions regarding their career or their family planning. This scoping review will focus on medical students in various disciplines and the influence of representations on their plans regarding their future career and personal life.

This scoping review will identify and map the existing evidence regarding medical students' representations regarding the interaction between parenthood and medical specialty. Synthesizing the existing evidence will shed light on the gaps and the potential for further research. In the context of physician shortages and feminization of the medical profession, the results of the scoping review will provide valuable insight to address

XSL•FO
RenderX

the challenges faced by physicians to reconcile their career and family planning.

Acknowledgments

The choice of database and the definition of the research equations have been made with the support of the Library of the Faculty of Medicine, University of Geneva.

Funding

This protocol is part of the PhD thesis of SA and is affiliated with the Swiss Medical Career Choice project. This work is supported by the Swiss National Science Foundation for research (grant number 213171). NMB and GLS are the lead investigators on this project. This funder had no role in the design or development of this protocol.

Authors' Contributions

SA: Conceptualization and design, writing and editing
MG: Editing and revision, approval of final manuscript
GLS: Editing and revision, approval of final manuscript
NMB: Editing and revision, approval of final manuscript

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example of search strategy: PubMed.
[DOCX File, 13 KB - resprot_v15i1e78133_app1.docx]

Checklist 1
PRISMA checklist.
[PDF File, 139 KB - resprot_v15i1e78133_app2.pdf]

References

1. Alers M, Verdonk P, Bor H, Hamberg K, Lagro-Janssen A. Gendered career considerations consolidate from the start of medical education. Int J Med Educ 2014 Sep 13;5:178-184. [doi: 10.5116/ijme.5403.2b71] [Medline: 25341228]
2. Bruce AN, Battista A, Plankey MW, Johnson LB, Marshall MB. Perceptions of gender-based discrimination during surgical training and practice. Med Educ Online 2015;20:25923. [doi: 10.3402/meo.v20.25923] [Medline: 25652117]
3. Cronin C, Lucas M, McCarthy A, et al. Are we reaping what we sow? Gender diversity in surgery: a survey of medical students. Postgrad Med J 2019 Mar;95(1121):119-124. [doi: 10.1136/postgradmedj-2018-136136] [Medline: 30975724]
4. Davis G, Allison R. Increasing representation, maintaining hierarchy: an assessment of gender and medical specialization. Social Thought and Research 2013;32:17-45. [doi: 10.17161/STR.1808.12435]
5. Drinkwater J, Tully MP, Dornan T. The effect of gender on medical students' aspirations: a qualitative study. Med Educ 2008 Apr;42(4):420-426. [doi: 10.1111/j.1365-2923.2008.03031.x] [Medline: 18338995]
6. Hoffman R, Mullan J, Nguyen M, Bonney AD. Motherhood and medicine: systematic review of the experiences of mothers who are doctors. Med J Aust 2020 Oct;213(7):329-334. [doi: 10.5694/mja2.50747] [Medline: 32865236]
7. Delbaere I, Pitsillos T, Greek Collaborating Group, Tydén T, Kerckhof L, Iliadis SI. Fertility awareness and parenthood intentions among medical students in three European countries. Eur J Contracept Reprod Health Care 2021 Aug;26(4):312-322. [doi: 10.1080/13625187.2021.1901877] [Medline: 33855916]
8. Halley MC, Rustagi AS, Torres JS, et al. Physician mothers' experience of workplace discrimination: a qualitative analysis. BMJ 2018 Dec 12;363:k4926. [doi: 10.1136/bmj.k4926] [Medline: 30541926]
9. Chun SE, Lee JH, Lee JE, Lee SMK, Leem J, Kim H. Impact of gender on the career development of female traditional Korean medicine doctors: a qualitative study. BMJ Open 2019 Aug 21;9(8):e030390. [doi: 10.1136/bmjopen-2019-030390] [Medline: 31439610]
10. Oberberg S, Enax-Krumova EK, Kruppa C, et al. Career and life planning in the context of the postgraduate medical training - current challenges and opportunities. GMS J Med Educ 2024;41(1):Doc5. [doi: 10.3205/zma001660] [Medline: 38504867]
11. Ziegler S, Zimmermann T, Krause-Solberg L, Scherer M, van den Bussche H. Male and female residents in postgraduate medical education - a gender comparative analysis of differences in career perspectives and their conditions in Germany. GMS J Med Educ 2017;34(5):Doc53. [doi: 10.3205/zma001130] [Medline: 29226221]

XSL•FO
RenderX

12. Acai A, Mahetaji K, Reid SE, Sonnadara RR. The role of gender in the decision to pursue a surgical career: a qualitative, interview-based study. Can Med Educ J 2020 Aug;11(4):e51-e61. [doi: 10.36834/cmej.69292] [Medline: 32821302]
13. Trinh LN, O'Rorke E, Mulcahey MK. Factors Influencing female medical students' decision to pursue surgical specialties: a systematic review. J Surg Educ 2021;78(3):836-849. [doi: 10.1016/j.jsurg.2020.08.050] [Medline: 32933885]
14. Levine RB, Mechaber HF, Reddy ST, Cayea D, Harrison RA. "A good career choice for women": female medical students' mentoring experiences: a multi-institutional qualitative study. Acad Med 2013 Apr;88(4):527-534. [doi: 10.1097/ACM.0b013e31828578bb] [Medline: 23425983]
15. Hill E, Vaughan S. The only girl in the room: how paradigmatic trajectories deter female students from surgical careers. Med Educ 2013 Jun;47(6):547-556. [doi: 10.1111/medu.12134] [Medline: 23662871]
16. Kristoffersson E, Diderichsen S, Verdonk P, Lagro-Janssen T, Hamberg K, Andersson J. To select or be selected - gendered experiences in clinical training affect medical students' specialty preferences. BMC Med Educ 2018 Nov 19;18(1):268. [doi: 10.1186/s12909-018-1361-5] [Medline: 30453953]
17. Cochran A, Hauschild T, Elder WB, Neumayer LA, Brasel KJ, Crandall ML. Perceived gender-based barriers to careers in academic surgery. Am J Surg 2013 Aug;206(2):263-268. [doi: 10.1016/j.amjsurg.2012.07.044] [Medline: 23414631]
18. Nicholson S. "So you row, do you? You don't look like a rower." An account of medical students' experience of sexism. Med Educ 2002 Nov;36(11):1057-1063. [doi: 10.1046/j.1365-2923.2002.01332.x] [Medline: 12406266]
19. Samuriwo R, Patel Y, Webb K, Bullock A. "Man up": medical students' perceptions of gender and learning in clinical practice: a qualitative study. Med Educ 2020 Feb;54(2):150-161. [doi: 10.1111/medu.13959] [Medline: 31746029]
20. Sorouri K, Khan S, Bowden S, Searle S, Carr L, Simpson JS. The glaring gender bias in the operating room: a qualitative study of factors influencing career selection for first-year medical students. J Surg Educ 2021;78(5):1516-1523. [doi: 10.1016/j.jsurg.2021.01.014] [Medline: 33579653]
21. Brown MEL, Hunt GEG, Hughes F, Finn GM. "Too male, too pale, too stale": a qualitative exploration of student experiences of gender bias within medical education. BMJ Open 2020 Aug 13;10(8):e039092. [doi: 10.1136/bmjopen-2020-039092] [Medline: 32792453]
22. Babaria P, Bernheim S, Nunez-Smith M. Gender and the pre-clinical experiences of female medical students: a taxonomy. Med Educ 2011 Mar;45(3):249-260. [doi: 10.1111/j.1365-2923.2010.03856.x] [Medline: 21299600]
23. Hill E, Solomon Y, Dornan T, Stalmeijer R. "You become a man in a man's world": is there discursive space for women in surgery? Med Educ 2015 Dec;49(12):1207-1218. [doi: 10.1111/medu.12818] [Medline: 26611186]
24. P Judge-Golden C, K Dotters-Katz S, Weber JM, Pieper CF, Gray BA. Parenthood and medical training: challenges and experiences of physician moms in the US. Teach Learn Med 2024;36(1):43-52. [doi: 10.1080/10401334.2022.2141750] [Medline: 36370040]
25. Monteiro S, Chan TM, Kahlke R. His opportunity, her burden: a narrative critical review of why women decline academic opportunities. Med Educ 2023 Oct;57(10):958-970. [doi: 10.1111/medu.15141] [Medline: 37312630]
26. Todd AR, Cawthorn TR, Temple-Oberle C. Pregnancy and parenthood remain challenging during surgical residency: a systematic review. Acad Med 2020 Oct;95(10):1607-1615. [doi: 10.1097/ACM.0000000000003351] [Medline: 32271231]
27. Farr RM, Moscovici S. Social Representations: Cambridge University Press; 1984.
28. Jaspars J, Fraser C. Attitudes and social representations. In: Farr RM, Moscovici S, editors. Social Representations: Cambridge University Press; 1984:101-123.
29. Itkonen L, Kosonen T. Exploring social representations of career planning: insights from Finnish university students. Nordic Journal of Transitions, Careers and Guidance 2024 Dec 12;5(1):123-136. [doi: 10.16993/njtcg.83]
30. Tricco AC, Soobiah C, Antony J, et al. A scoping review identifies multiple emerging knowledge synthesis methods, but few studies operationalize the method. J Clin Epidemiol 2016 May;73:19-28. [doi: 10.1016/j.jclinepi.2015.08.030] [Medline: 26891949]
31. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
32. DeepL Translate. 2025. URL: https://www.deepl.com/fr/translator [accessed 2025-12-23]
33. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 5;5(1):210. [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
34. EndNote. 2013. URL: https://www.myendnoteweb.com/EndNoteWeb.html [accessed 2025-12-23]
35. Critical appraisal tools. JBI. URL: https://jbi.global/critical-appraisal-tools [accessed 2025-12-19]
36. Federal act on research involving human beings. Swiss federal authorities. URL: https://www.fedlex.admin.ch/eli/cc/2013/617/en [accessed 2025-12-23]
37. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]

XSL•FO
RenderX

Edited by J Sarvestan; submitted 27.May.2025; peer-reviewed by P Hudelson; revised version received 11.Nov.2025; accepted 13.Nov.2025; published 20.Jan.2026.

Please cite as:
Arnoux S, Gisselbaek M, Savoldelli GL, Bajwa NM
Exploring Medical Students' Representations of Future Specialties and Parenthood: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e78133
URL: https://www.researchprotocols.org/2026/1/e78133
doi:10.2196/78133

© Sylvie Arnoux, Mia Gisselbaek, Georges Louis Savoldelli, Nadia Masood Bajwa. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Understanding the Nurse Champion Concept and the Training Initiatives for Nurse Champions: Protocol for a Scoping Review

Wilmer J Santos[1,2], RN, MScN; Ian D Graham[1,2,3], PhD; Amanda Vandyk[1], RN, PhD; Gillian Harvey[4], RN, PhD; Janet E Squires[1,2], RN, PhD

[1]School of Nursing, Faculty of Health Sciences, University of Ottawa, 200 Lees Avenue, Room 313, Ottawa, ON, Canada

[2]Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON, Canada

[3]School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, ON, Canada

[4]College of Nursing and Health Sciences, Caring Futures Institute, Flinders University, Adelaide, Australia

Corresponding Author:
Wilmer J Santos, RN, MScN
School of Nursing, Faculty of Health Sciences, University of Ottawa, 200 Lees Avenue, Room 313, Ottawa, ON, Canada

Abstract

Background: There is current evidence that a proportion of health care services provided to patients do not align with best evidence. A nurse champion, defined as a nurse who either volunteers or is identified by the management to facilitate or promote the implementation of an innovation (eg, new knowledge or practice), is an important factor for implementation success. The existing literature describes health care champions' attributes, roles, and behaviors, the processes in which they might enable change, and their effectiveness at facilitating implementation. However, a more detailed exploration of the nurse champion concept is needed. Further, despite the prolific use of nurse champions in health care implementation, there is a gap in the literature pertaining to what nurse champion training initiatives exist, what competencies are important to be a nurse champion, and whether current training initiatives are effective in preparing nurse champions. Finally, the extent to which equity, diversity, and inclusion are considered in the nurse champion literature is unknown.

Objective: This study aims to (1) develop a preliminary conceptual understanding of nurse champions, (2) describe the characteristics of existing champion training initiatives in health care that prepare nurse champions and synthesize the competencies that are covered in these champion training initiatives, (3) synthesize the findings of studies that examined the effectiveness of nurse champion training initiatives in preparing nurses to be effective champions, and (4) evaluate the extent to which equity, diversity, and inclusion are considered in studies that define nurse champions and in studies describing or evaluating nurse champion training initiatives attended by nurse champions.

Methods: This series of linked reviews will follow the Joanna Briggs Institute scoping review methodology. We will systematically search 8 electronic databases using a Peer Review of Electronic Search Strategies. We will also search for gray literature (eg, theses and dissertations). We will upload the records from our database searches into Covidence. Two individuals from the research team will perform title and abstract and full-text screening independently and in duplicate using a piloted inclusion and exclusion criteria. Two individuals will perform data extraction and quality appraisal independently and in duplicate. Conflicts will be resolved with consensus. We will perform various forms of content analysis to address our varying research questions and objectives.

Results: As of January 2026, we have completed more than half of our title and abstract screening. We expect to present the results in a scoping review later in 2026.

Conclusions: The results of this research will provide recommendations for current and future nurse champion training initiatives. Although this review is limited to the nursing discipline, the results may be transferable to understanding champions in other contexts.

International Registered Report Identifier (IRRID): PRR1-10.2196/80204

(JMIR Res Protoc 2026;15:e80204) doi:10.2196/80204

KEYWORDS

nurse champions; facilitation; health care; implementation; nursing; equity; diversity; inclusion

XSL•FO
RenderX

Introduction

Background

Despite advancement in health research, many authors identify that a significant proportion of care provided to patients is not in line with the available research evidence [1,2]. For instance, Braithwaite et al [1] described the 60-30-10 phenomenon, indicating that about 60% of care aligns with guideline recommendations, that 30% of care is a waste, because it is duplicated, or of low value, and that at least 10% of patients experience iatrogenic harm or adverse events. A more recent systematic review reported that, on average, 30% of care reported in 174 Canadian studies is not concordant with best evidence or guidelines [2].

Knowledge translation is the dynamic and iterative process that is required to synthesize, disseminate, exchange, ethically apply, evaluate, sustain, or coproduce knowledge using theory-informed and tested processes to provide more effective health services and improve the health care system [3,4]. Many authors emphasize the importance of individuals who enable and lead the implementation of care that is in accordance with the best evidence [5-7]. There has been a growing interest and research on the deployment of champions across different contexts in health care [6-12], such as acute care [8], cancer care [9], long-term care [11], and mental health [10]. The champion concept originates from the management business literature in the 1960s and 1970s; champions were initially deployed by companies to disseminate and promote the use of technological innovations [13,14]. Champions are individuals who often volunteer to work above and beyond their occupational role to promote an innovation (eg, new knowledge, practice, policy, and technology) or enact behaviors to mitigate barriers and reinforce facilitators to implementation [6,8,15]. Champions are often health care workers [8,9], but they can also be individuals in management [16-18] or patients [19,20]. Champions might self-identify or emerge organically, or be appointed formally by management [21,22].

There have been ongoing movements to elevate nursing practice so that it is in accordance with the current research evidence, nurses' experiential knowledge, and patients' preferences [4,23-25]. The importance of deploying nurse champions is emphasized by researchers across multiple sectors [6,8,26] such as acute care [27,28], primary care [29], and home care [30]. As detailed below, despite the plethora of research on champions, further research is required to understand the concept of nurse champions, how nurse champions are trained, and the equity, diversity, and inclusion (EDI) considerations pertaining to the training and deployment of nurse champions.

Conceptualization of Nurse Champions

In previous studies, nurse champions have been described as nurses working in a clinical or managerial role who enact facilitation to increase the uptake of an innovation at their place of work [31-34]. Facilitation is defined as the process that activates implementation by assessing and responding to the characteristics of the innovation, the recipient, and the context, thereby emphasizing existing enablers and addressing barriers to implementation [35]. Although there has not been a concept analysis specifically on nurse champions, previous conceptual studies (published in 2006 [36] and 2017 [15]) about health care champions have focused on clarifying terms that refer to individuals who enable implementation, including champions, change agents, coaches, opinion leaders, knowledge brokers, facilitators, and linking agents [15,36]. Authors have reported that there is confusion between the terms used to describe individuals who facilitate implementation [6,15,36]. The same authors have stated that these different terms or concepts are more similar than different, as their main function is often to facilitate successful implementation, but differ based on their intended purpose, how influence is exerted, their domain of influence [36], and their theoretical origins [15].

Authors have described health care champions based on common personal attributes that they possess or the behaviors they enact [6,15]. In a 2017 scoping review n = 195 studies by Cranley et al [15], the authors presented key health care champion attributes and skills (based on 18 papers), which included (1) expert knowledge of the innovation, (2) persuasiveness, (3) mentorship skills, (4) being visionary, (5) enthusiasm, (6) creativity, (7) driven and passionate about their work, and (8) communication skills. In a 2018 integrative review by Miech et al [6], the authors presented a more comprehensive list of characteristics of health care champions that included all the attributes or skills reported by Cranley et al [15] in addition to the following: (1) negotiation skills, (2) advocacy, (3) communication across organizational boundaries, (4) strong educator and presentation skills, (5) having political acumen, (6) leading teams and recruiting new team members, (7) engaging in team planning and goal setting, and (8) collecting data, tracking progress, and providing feedback. In a 2020 comparative case study (n=78 clinicians and health administrators) by Bonawitz et al [37], health care champions have 6 key attributes: (1) they have influence over other people's opinions and behaviors, (2) they have ownership of the implementation, (3) they are physically present at the point of change, (4) they are persuasive, (5) they have grit, and (6) they enact a participative leadership style.

Authors have formulated conceptual frameworks of champions according to the existing literature [7]. For instance, in a recent conceptual paper published in 2021, Shea [7] presented a conceptual model that describes 7 constructs that pertain to health care champions' performance and impact in implementation activities. According to this conceptual model, the health care champion's performance (defined as activities they perform) is influenced by their level of commitment, experience, and self-efficacy. The health care champions' level of commitment is influenced by their beliefs about the innovation and the organizational support provided to the health care champions. The health care champions' impact (implementation outcomes [eg, changes in perceived acceptability or appropriateness of the innovation]) is related to the health care champions' performance mediated by the level of peer engagement with the health care champion during the phases of implementation [7]. Similarly, Morena et al [38] proposed 2 causal pathway models to illustrate how clinical champion attributes and skills, which function under the mechanism of action of social influence, mediate health care

XSL•FO
RenderX

providers' attitudes, health care providers' self-efficacy, subjective norms, and individuals' perceived abilities to perform the behaviors, which then at the end leads to increased use of the innovation by health care providers [38].

Although the literature discussed above on champions in health care provides some conceptual understanding that may be informative in understanding the nurse champion concept, a fulsome exploration of the concept is needed. According to Rodgers and Knafl [39,40], conceptual clarity relies on understanding the significance, use, and application of a concept over time and within a particular context. Garnering conceptual clarity on nurse champions is important as there may be varied significance, use, and application of the nurse champion concept within the nursing discipline (eg, management, education, and clinical practice), and there may be changes in how nurse champions have been defined through time. Furthermore, conducting a concept analysis on nurse champions using a systematic process may advance the work of Shea [7], as he reported that his conceptual model was not informed by a systematic review. Finally, as stated by Morena et al [38], the causal pathways they constructed may not be applicable to all clinical champions; hence, an in-depth analysis of nurse champions may reveal nuances pertaining to the attributes of nurse champions.

Nurse Champions Training Initiatives and Competencies

Despite the prolific use of champions in nursing and in health care, the evidence pertaining to their effectiveness or ability to facilitate successful implementation is variable [6,8]. In a 2018 integrative review (n=199) on champions in health care, Miech et al [6] concluded that deploying champions is a necessary but not a sufficient condition for implementation success. Champions were stated to be necessary because more than 80% of the studies included in their review cited champions as one of the key factors associated with implementation success, but the extent to which champions influenced implementation success, especially with the presence of other implementation strategies, was not clear [6]. Similarly, a 2022 systematic review conducted by members of this team (n=35) [8] on the effectiveness of health care champions in improving innovation use or outcomes within the context of health care implementation found that the presence of champions was associated with the uptake of practices, programs, and technologies at a system or facility level (in 5 of 7, 71.4% studies). However, the deployment of champions was inconsistently correlated with improvements in conceptual (2/4, 50%) and instrumental (8/17, 47.1%) innovation use by health care providers and in improving outcomes for patients (3/6, 50%) [8]. Conceptual innovation use was defined as increased knowledge or improved attitude toward an innovation [41]. Instrumental innovation use was defined as increased use of the innovation [41]. In both reviews [6,8], the authors reported that champions were often operationalized as merely the lack or presence of a champion during implementation activities in health care. Specifically, most of the studies (25/35, 71.4%) included in the 2022 systematic review conducted by members of our research team [8] operationalized champions simply as a dichotomous item reflecting either the presence or absence of someone labeled as a champion. In the same review, the authors [8] reported 5 subscales that were used in 6 studies to evaluate the existence of champions who have particular attributes, or performed specific roles and behaviors. Similarly, Miech et al [6] reported that in more than 90% of the papers included in their integrative review (n=199 articles), champions were reported simply as the presence or absence of a champion. The lack of robust measures pertaining to champions in health care may be attributed to the murky conceptualization of the concept [6,15,36], which further justifies the concept analysis of nurse champions as discussed above. Furthermore, the lack of detailed descriptions of champions or robust measures used in empirical studies evaluating health care champions is problematic because of the variability of roles and behaviors that they can perform [6,9]; these roles and behaviors require different competencies to perform well.

Competency is defined as the ability of an individual to perform a behavior successfully with the necessary attitudes, judgment, knowledge, and skills [42]. A 2025 rapid systematic review of implementation trials [12] (n=15) reported that there was a lack of description of the training provided to champions in health care, according to 12 of the trials (80%) included in their review. Other authors have echoed the concern regarding the lack of understanding of how effective champions in health care are identified and prepared [6,7,38]. Despite the well-synthesized knowledge pertaining to the roles and behaviors of health care champions (which is assumed to be transferable to nurse champions), there is a gap in knowledge pertaining to the training initiatives that prepare nurse champions and the competencies required by nurse champions to perform behaviors that improve innovation use. Furthermore, there is a need to synthesize the evidence pertaining to the effectiveness of nurse champion training initiatives, as this knowledge can inform future implementation studies on what approaches were helpful in producing effective nurse champions.

EDI Considerations in Training and Deploying Nurse Champions

There was minimal discussion of the EDI considerations pertaining to the role and behaviors of champions in health care in the previous studies discussed above [6-8,12,15,36-38]. In addition, there were minimal or no demographics of the health care champions or the individuals that the champions were intended to influence reported in the studies discussed above [6-8,12,15,36-38]. EDI considerations are important to explore in research pertaining to nurse champions and champions in general because peer or social influence can be more successful at changing behaviors when it is performed by an individual who comes from the same societal group as the people they are trying to influence [43,44]. EDI considerations might also be important to consider when nurse champions or health care champions are deployed to enable the implementation of innovations targeted toward a specific group of people, or when the implementation activity affects equity-deserving populations. For example, in Bonawitz et al's [37] 2020 comparative case study, obstetrician or gynecologist champions were deployed to promote postpartum contraceptive care, which is an innovation that is specific to women of childbearing age. Morena et al [38] also echo the need to examine champions who work

XSL•FO
RenderX

with equity-deserving populations and the EDI considerations in deploying champions.

Leadership styles that have been traditionally associated with gender roles may also influence the deployment of nurse champions. For example, in a 2018 cross-sectional study by Luz et al [21], they hypothesized that men are more likely to be formally appointed by leadership to become nurse champions because men have traditionally been ascribed a leadership style that is more task-oriented, reliant on directing others' behavior, and is more authoritarian [34,45]. On the other hand, Luz et al [21] hypothesized that women are more likely to be informal nurse champions because women have been traditionally attributed a leadership style that is more collaborative and more reliant on their interpersonal relationships with their peers [46]. Luz et al [21] did not find a relationship between the nurse champions' gender and whether they were formally or informally appointed; the authors rationalized that the lack of relationship might be related to the small number of men in their sample. Hence, there is a need to examine the extent to which EDI considerations are pertinent to nurse champions' training, deployment, and behaviors.

Study Objectives

Our study objectives are to (1) develop a preliminary conceptual understanding of nurse champions, (2) describe the characteristics of existing champion training initiatives in health care that prepare nurse champions and synthesize the competencies that are covered in these champion training initiatives, (3) synthesize the findings of studies that examined the effectiveness of nurse champion training initiatives in preparing nurses to be effective champions, and (4) evaluate the extent to which EDI is considered in studies that define nurse champions and in studies describing or evaluating nurse champion training initiatives attended by nurse champions. Although we acknowledge the conclusions made in previous work that there are a lot of similarities between champions and other concepts (eg, facilitators, change agents, and opinion leaders) [6,15,36], focusing our research on individuals who are explicitly referred to as nurse champions will allow for more nuanced understanding of how this concept has been defined and used in different health care sectors and over time, and what training is being provided to nurse champions.

Methods

Study Design

This study is a single scoping review with one prespecified search, screening, eligibility, and data extraction workflow. We will follow the Joanna Briggs Institute (JBI) scoping review approach to conduct this series of link reviews [47], where we will conduct a broad search of the literature, perform varying analyses, and then report the results in 4 papers addressing each of our 4 study objectives. A scoping review is an appropriate methodology because of the broad research objectives. There are five mandatory stages in a scoping review [47,48]: (1) identifying the research question; (2) identifying relevant studies; (3) study selection; (4) data charting; and (5) collating, summarizing, and reporting results. Our review is rooted in a pragmatist paradigm because we are including and gathering knowledge developed using different methodologies and from different sources (ie, published and gray literature) to achieve our research objectives. JBI's approach in evidence implementation and synthesis is also situated within a pragmatist paradigm [49], which demonstrates alignment between our chosen methodology and paradigm. As detailed in a later section, we will follow specific approaches for identifying relevant articles and perform varying data analysis approaches relevant to each review.

Identifying the Research Question

There are 11 research questions for the scoping review: (1) What are the definitions used in the literature to describe nurse champions? (2) What are the disciplinary or theoretical origins of the definitions used to describe nurse champions? (3) What are the defining attributes, antecedents, consequences, references, surrogate terms, and related concepts to nurse champions? (4) What champion training initiatives exist within health care that are attended by nurses? (5) What are the characteristics (eg, target participants and cost) of the nurse champion training initiatives? (6) What competencies are taught in nurse champion training initiatives in health care? (7) What is the evidence that the nurse champion training initiatives in health care increase nurse champions' knowledge and perceived self-efficacy in performing the nurse champion role? (8) What is the evidence that the nurse champion training initiatives in health care improve nurse champions' performance and impact? (9) What are the EDI considerations in studies that discuss the nurse champion concept or nurse champion training initiatives? (10) What equity-related factors, as detailed by the PROGRESS-plus (place of residence; race, ethnicity, culture, and language; occupation; gender and sex; religion; education; socioeconomic status; and social capital–plus) framework [50], are used to describe nurse champions and the study's sample in the included documents? (11) Do any of the studies that evaluate the effectiveness of nurse champion training initiatives integrate EDI considerations? The PROGRESS-plus framework is a widely used framework that outlines categories that potentially influence inequities in health care [50]. There are many research questions, but they pertain to 1 of the 4 research objectives. Specifically, research questions 1 to 3 address objective 1 pertaining to the development of a preliminary conceptual understanding of nurse champions. Similarly, research questions 4-6 address objective 2, which pertains to collating existing nurse champion training initiatives and the competencies that nurse champions gain from their training. Research questions 7 and 8 address objective 3, which pertains to understanding the effectiveness of nurse champion training initiatives. Finally, research questions 9-11 address objective 4 pertaining to examining the EDI considerations related to training and deploying nurse champions. Table 1 details the alignment between the research objectives, research questions, the data extracted, the planned analysis, and how we will publish the results into 4 articles.

XSL•FO
RenderX

Table . Mapping of research objectives, research questions, data extractions, and analysis plan.

Dissemination	Research objective	Research questions	Data extracted	Analysis plan
Article 1	• Develop a preliminary conceptual understanding of nurse champions.	• What are the definitions used in the literature to describe nurse champions? • What are the disciplinary origins of the definitions used to describe nurse champions? • What are the defining attributes, antecedents, consequences, references, surrogate terms, and related concepts to nurse champions?	• Definition used to describe nursing champions. • Citations and disciplinary origins of definitions used to describe nurse champions. • The attributes, antecedents, consequences, references, surrogate terms, and related concepts. • Regarding references (context in which the concept has been used), we will extract the (1) health care sectors, (2) type of innovation for, and (3) the recipients of the champions' facilitation or influence (eg, do champions influence laterally with colleagues or vertically with management). • The date of publication of the included studies.	• Conventional content analysis [51]. We will code and then combine codes into larger themes to synthesize the attributes, antecedents, consequences, surrogate terms, and related concepts that are pertinent to the concept of nurse champions. We will also develop a definition of the term nurse champion from the extracted attributes of the nurse champion concept. To demonstrate the evolution of the nurse champion, we will analyze the attributes, antecedents, consequences over time, disciplinary origin of definitions, the health care sector, type of innovation, and recipient of champions' facilitation or influence.

XSL•FO
RenderX

Dissemination	Research objective	Research questions	Data extracted	Analysis plan
Article 2	• Describe the characteristics of existing champion training initiatives in health care that prepare nurse champions and synthesize the competencies that are covered in these champion training initiatives.	• What champion training initiatives exist within health care that are attended by nurses? • What are the characteristics (eg, target participants, cost) of the nurse champion training initiatives? • What competencies are taught in nurse champion training initiatives in health care?	• Name of champion training initiative. • What sector of health care did the champion training initiative come from? • Is the training initiative specific to implementing a particular innovation? What is the innovation? • Who are the target participants of the champion training initiative (eg, nurse leaders or clinical nurses)? • Cost of training initiative. • Location of training initiative. • Format of training initiative (eg, in person, virtual, self-paced, and hybrid). • Length of training initiative. • The main topics covered by the training initiative. • Did the authors explicitly outline the competencies (attitudes, judgment, knowledge, and skills) that they intended their participants would gain from their training initiative? • The competencies (attitudes, judgment, knowledge, and skills) are emphasized in the champion training initiative. • The behaviors that the champion is expected to be able to perform after gaining the competencies from the training initiative.	• Summative content analysis [51]. We will group and stratify (using frequency counts and descriptive statistics) the characteristics of the nurse champion training initiatives according to the different health care sectors, the type of innovations, and the target participants of the champion training initiatives. • Directed and conventional content analysis [51]: We will organize the data according to AACTT[a] [52] to demonstrate the attitudes, judgment, knowledge, and skills (competencies) that are linked to the expected behaviors (action) that a champion (actor) enacts to promote the implementation of a particular innovation within a context toward knowledge users (target) and at a specific phase or phases of implementation (time).
Article 3	• Synthesize the findings of studies that examined the effectiveness of nurse champion training initiatives in preparing nurses to be effective champions.	• What is the evidence that the training initiatives in health care increase nurse champions' knowledge and perceived self-efficacy in performing the nurse champion role? • What is the evidence that the nurse champion training initiatives in health care improve nurse champions' performance and impact?		

XSL•FO
RenderX

Dissemination	Research objective	Research questions	Data extracted	Analysis plan
			• Study design of the quantitative studies evaluating the effectiveness of champion training initiatives. • Sample size (number of champions attending the training initiative). • Measure of knowledge gained by the champion after completing the training initiative. Measure of the champions' self-efficacy. • Measures evaluating the ability of champions to facilitate implementation (eg, number of nurses who interacted with a nurse champion; satisfaction of nurses with the support provided by the champion). • Measure of successful implementation (eg, innovation use or impact from the use of the innovation) • Reliability and validity of measures. • Statistical tests performed. • Test statistic or effect size. • Statistical significance of results (*P* value).	• Meta-analysis if there are more than 2 experimental or quasi-experimental studies that used the same outcome measures with similar populations and interventions [53]. If the outcome measures, population, and interventions are heterogenous, then we will perform a narrative synthesis of the data based on (1) health care sector (eg, acute care, LTC[b], etc), (2) type of innovation, (3) quantitative research design, (4) outcome measures, (5) significance or lack of significance of results, and (6) quality of studies based on the critical appraisals.
Article 4	• Evaluate the extent to which EDI[c] are considered in studies that define nurse champions and in studies describing or evaluating nurse champion training initiatives attended by nurse champions.	• What are the EDI considerations in studies that discuss the nurse champion concept or nurse champion training initiatives? • What equity-related factors as detailed by the PROGRESS-plus[d] framework (1) are used to describe nurse champions and the study's sample in the included documents? • Do any of the studies that evaluate the effectiveness of nurse champion training initiatives integrate EDI considerations?	• Definition of nurse champions in the context of the implementation of innovations in equity-deserving populations. • The context in which champions facilitate the implementation of innovations for equity-deserving populations. • The name and description of the champion training initiatives that include EDI considerations. • EDI considerations are present in champion training initiatives. • The equity-related factors used to describe champions and the sample. • The positionality of the authors of the documents.	• Conventional and summative content analysis [54]. We will perform conventional content analysis to synthesize themes in the extracted data pertaining to EDI considerations in the included documents that discussed the nurse champion concept or nurse champion training initiatives and nurse champion training initiatives' effectiveness. We will conduct summative content analysis to describe (1) the demographics of the nurse champions and the study sample according to the PROGRESS-plus framework [50]; and (2) the number of included documents that detailed EDI considerations.

XSL•FO
RenderX

[a]AACTT: action, actor, context, target, time.

[b]LTC: long-term care.

[c]EDI: equity, diversity, and inclusion.

[d]PROGRESS-plus: place of residence; race, ethnicity, culture, and language; occupation; gender and sex; religion; education; socioeconomic status; and social capital–plus.

Identifying Relevant Studies

Overview

We performed preliminary searches on Google and Google Scholar to identify papers that are relevant to our study. We developed a search strategy on MEDLINE in consultation with a health science librarian (Victoria Cole) and informed by the key terms we identified from the preliminary online searches. The search strategy includes both Boolean phrases and Medical Subject Headings (MeSH) terms. The key terms in the search strategy include (1) champions and (2) nursing and synonyms (Table 2 for the preliminary search strategy). A second health science librarian evaluated the search strategy using the peer review of electronic search strategy (PRESS) checklist [55] (Multimedia Appendix 1). WJS translated and applied the search strategy to the other databases.

We searched the published literature through 8 health care, education, and management databases: Business Source Complete, CINAHL, Embase, ERIC, MEDLINE, Nursing and Allied Health Premium, PsycINFO, and Scopus. We searched education and management databases to capture articles about nurse champions in nursing education or management. Regarding gray literature, we searched for theses and dissertations from ProQuest and Theses Canada. Other gray literature sources that we will search include (1) documents from the Canadian Agency For Drugs And Technologies In Health Grey Matters, (2) documents from champion training initiative or program websites such as the Registered Nurses' Association of Ontario's Best Practice Champions Network [56] and the Strengthening a Palliative Approach in Long-Term Care Champions' page [57], and (3) government and policy documents through Google search of the keywords champion* and nurs*. We will also contact authors of included documents on an as-needed basis for any other documents detailing champion competencies or the characteristics of their champion training initiatives when more information is required. We will assess the reference lists of included documents, and we will search the list of articles that have cited the included documents using Scopus and Web of Science. We will also evaluate systematic reviews (identified through the online database searches) on champions for any other pertinent articles.

Table . Preliminary search strategy on MEDLINE.

Number	Query	Search results from April 24, 2025, n
1	champion*.tw,kf.	11,896
2	nurs*.tw,kf.	580,298
3	exp Practice Patterns, Nurses'/ or exp Nurse's Role/ or exp Licensed Practical Nurses/ or exp Nurses/ or exp nursing, practical/ or exp faculty, nursing/ or exp nursing staff/	210,885
4	2 or 3	654,444
5	1 and 4	1326

Inclusion Criteria

We will include documents that meet these criteria (1) primary studies (quantitative, qualitative, and mixed-methods) and gray literature; (2) documents that describe individuals who are explicitly referred to as nurse champions; (3) documents written in any language; and (4) documents published in 1970 and onwards as champions were first identified by Schon's study on radical military innovations [14,36] in the 1960s and 1970s, and this cut off was previously used in recent reviews on champions [7,9]; (5) documents that provide a definition or describe the concept nurse champion, describe a nurse champion training initiative and the competencies emphasized by the nurse champion training initiative, evaluate a nurse champion training initiative, or discuss EDI considerations pertaining to deployment or training of nurse champions; and (6) documents that are related to implementation of innovations or evidence-based practices.

Exclusion Criteria

We will exclude documents that (1) are not available in full text; (2) are explicitly only about other knowledge translation roles; (3) are review papers but relevant review papers about champions will be tagged to have their references screened; (4) are conference abstracts but relevant conference abstracts will be tagged and evaluated for relevant articles; (5) are about veterinary nurse champions; (6) are not set in a health care context; and (7) are clinical trial registrations or protocol registration. Any clinical trial registration or protocol registrations that might be relevant to answering the research questions will be tagged, and we will appraise these registrations for any relevant documents.

Study Selection

WJS uploaded the downloaded records from the databases to Covidence, and this system automatically deduplicated the records. Another researcher (Letitia Nadalin-Penno, Megan Greenough, or Priscilla Packiam) and WJS will independently

XSL•FO
RenderX

read and screen the titles and abstracts and then full texts in accordance with the inclusion and exclusion criteria. The 2 reviewers (Letitia Nadalin-Penno, Megan Greenough, or Priscilla Packiam and WJS) will pilot the inclusion and exclusion criteria on 100 titles and abstracts; screening will begin once there is ≥75% agreement (κ higher than 0.75) between the reviewers, which determines substantial agreement between the 2 reviewers (Letitia Nadalin-Penno, Megan Greenough, or Priscilla Packiam and WJS) [58]. The 2 reviewers (Letitia Nadalin-Penno, Megan Greenough, or Priscilla Packiam and WJS) will resolve conflicts through discussion or through consultation of a third senior research member (IDG or JES). We will translate titles and abstracts and full-text documents not written in English using the University of Ottawa's enterprise instance of Microsoft Copilot [54]. We will use our research networks to identify individuals who can confirm the accuracy of the translation of the included non-English documents before we move on to data extraction and quality appraisal for these documents. If we cannot identify a native speaker in our networks, we will hire a professional translator to translate the article into English once we confirm that it meets inclusion.

To address research questions 1-3 related to research objective 1 (to increase conceptual understanding of the nurse champion concept), we will include documents that provide a definition of nurse champions or detail the attributes, antecedents, consequences, related terms, and reference of use of the nurse champion concept and we will place a tag named "concept analysis" for these documents on Covidence [59] during screening. To address research questions 4-6 related to research objective 2 (to collate nurse champion training initiatives and the nurse champion competencies emphasized within the training initiatives), we will include documents that detail training initiatives that are intended to prepare nurse champions. We will place a tag named "competencies" for these documents on Covidence [59]. To address research questions 7 and 8 related to research objective 3 (to synthesize the findings of studies on the effectiveness of the nurse champion training initiatives in increasing nurse champions' self-efficacy or improving their impact), we will prioritize inclusion of randomized controlled trials and quasi-experimental studies that evaluate the effectiveness of champion training initiatives in training nurse champions. However, we will also include other primary nonexperimental quantitative studies that evaluate the correlation between attending a nurse champion training initiative and the nurse champions' knowledge and perceived self-efficacy in performing the nurse champion's role, and the nurse champions' performance and impact. Although these nonexperimental studies do not determine the effectiveness [60], they can highlight existing relationships. We will also include documents that evaluate the effectiveness of training of nurse champions and other knowledge translation roles (eg, physician champions, knowledge broker, or opinion leaders) if details specific to nurse champions can be extracted separately. We will place a tag named "effectiveness" for the documents pertaining to the effectiveness of champion training initiatives. To address research questions 9-11 pertaining to research objective 4 (to evaluate the extent that EDI considerations are present in studies about the nurse champion concept or nurse champion training initiatives), we will place a tag named "EDI" for the documents that detail EDI considerations in relation to either the nurse champion concept or the champion training initiatives.

Data Charting

We will develop a data charting form on Covidence. The data charting form will include (1) study characteristics: year of publication, study's authors, discipline or field of the primary author, study design according to authors' and extractor's interpretation of design, and health care context; (2) characteristics of nurse champions (eg, sex, gender, ethnicity, etc); (3) the innovation being implemented; (4) the intended recipients or users of the innovation; (5) nurse champions' behaviors; (6) definitions used to describe nurse champions and the disciplinary or theoretical origins of the definitions; (7) attributes, antecedents, consequences, references, surrogate, and related concepts to the nurse champion concept; (8) characteristics of champion training initiatives: target participants, cost, length, accessibility (eg, how often are they offered; online or in person), and content of champion training initiatives (eg, main topics covered); (9) competencies emphasized by the champion training initiatives and their definitions; (10) how studies evaluated champions' effectiveness (eg, study design and sample size); (11) number of champions, outcome measures, reliability and validity of measures, statistical tests, test statistic, effect size, and significance of results; (12) EDI considerations in included studies pertaining to the nurse champion concept or the nurse champion training initiatives; and (13) equity-related factors used to describe champions and the study sample in accordance with the PROGRESS-plus framework [50]. Table 1 describes the data planned to be extracted in relation to the research objectives and questions. Data charting will be completed independently and in duplicate by 2 reviewers (Letitia Nadalin-Penno, Megan Greenough, and Priscilla Packiam or WJS). The 2 reviewers (Letitia Nadalin-Penno, Megan Greenough, and Priscilla Packiam or WJS) will pilot the data charting form on 2 - 3 papers. The data charting form may be updated as the study is conducted due to the iterative nature of data charting according to Peters et al [47]. The reviewers will resolve conflicts in data charting through consensus or through consultation with a third senior research member (IDG or JES).

Critical Appraisal

Two research members (Letitia Nadalin-Penno, Megan Greenough, or Priscilla Packiam and WJS) will perform an independent critical appraisal of the included studies. Even though it is not a necessary step in scoping reviews, we will evaluate and report the quality of documents included in the scoping review, as we will be extracting and synthesizing the results reported in these documents. The following JBI critical appraisal tools will be used (1) analytical cross-sectional, (2) case control, (3) case reports, (4) cohort study [61], (5) expert opinion [62], (6) quasi-experimental studies [60], (7) qualitative research [63], (8) narrative [62], (9) policy [62], and (10) randomized controlled trials critical appraisal tools [60] and the Mixed Methods Appraisal Tool for mixed methods studies [64]. The quality of the evidence will be described qualitatively; for instance, the authors will highlight strengths and limitations of

XSL•FO
RenderX

the included articles as per the JBI critical appraisal tools and the Mixed Methods Appraisal Tool.

Stage 5: Collating, Summarizing, and Reporting Results

We will report on the (1) study characteristics (year of publication, study's authors, the discipline or field of the primary author, the study design, and the health care context), (2) the characteristics of nurse champions (eg, sex, gender, ethnicity), (3) the innovation being implemented, and (4) the characteristics of the intended recipients or users of the innovation (profession, sex, gender, etc).

To achieve our 4 research objectives, we sought to conduct 4 subanalyses. The first subanalysis addresses the first research objective and research questions 1-3 pertaining to the development of a preliminary conceptual understanding of nurse champions. We will follow Rodgers' evolutionary concept analysis approach [39] to guide this subanalysis. We will inductively analyze data using conventional content analysis [51] to determine the following regarding nurse champions: (1) attributes of the nurse champion concept (defining characteristics of the concept that makes it identifiable across different contexts and allows for identification of situations in which the concept may be applicable [39]), (2) antecedents (necessary circumstances or characteristics that precede the concept [39]), (3) consequences (circumstances or characteristics that are produced following the existence of the concept), (4) surrogate terms (other words or phrases used to describe the concept other than the ones that are used or results from the analysis [39]), and (5) related terms (associated but distinct words or phrases relating to nurse champions [39]). Conventional content analysis follows an inductive approach: (1) data are coded, (2) codes are then clustered to form subcategories, and (3) these subcategories are then aggregated to create broader categories. The clustering of data into codes and subcategories is based on a subjective interpretation of similarities and differences in the text data [51]. WJS will code the data independently, and another researcher (Letitia Nadalin-Penno, Megan Greenough, and Priscilla Packiam) will check the coding. The coding will also be discussed with senior research members (AV, IDG, JES, or GH) during weekly to biweekly meetings. We will develop a definition of the term nurse champion according to the extracted attributes of the nurse champion concept [39]. We will use simple frequency counts to outline how many documents support a certain code or category. To demonstrate the evolution of the concept nurse champion, we will perform summative content analysis [39] by counting the number of times that each attribute, antecedent, and consequence of the nurse champion concept were mentioned in included documents according to the (1) date of publication, (2) disciplinary origin of definitions or the primary author's disciplines, (3) the theoretical origins (eg, conceptual frameworks, theoretical frameworks, or theories) of the definitions, (4) the health care sector, (5) type of innovation, and (6) the recipient of nurse champions' facilitation or influence.

The second subgroup analysis addresses the second research objective and research questions 4-6 pertaining to describing existing champion training initiatives in health care that prepare nurse champions and the competencies that are emphasized in these training initiatives. We will use summative content analysis [51] to report the characteristics of the nurse champion training initiatives according to different health care sectors, the type of innovations, and the target participants of the champion training initiatives using frequency counts or descriptive statistics. We will use conventional content analysis [51] to code and categorize the champion competencies emphasized in the training initiatives as attitudes, judgment, knowledge, or skills. We will construct definitions for each competency based on the data extracted. A second individual will check the data coding. We will use directed content analysis [51] to organize the data pertaining to nurse champions' behaviors according to action, actor, context, target, time [52], to demonstrate the attitudes, judgment, knowledge, or skills (competencies) that are linked to the expected behaviors (action) that a nurse champion (actor) enacts to promote the implementation of a particular innovation within a context toward knowledge users (target) and at a specific phase or phases of implementation (time). The research team will have weekly to biweekly meetings to discuss the extracted competencies and constructed definitions.

The third subgroup analysis addresses the third research objective and research questions 7 and 8 pertaining to the effectiveness of the champion training initiatives in increasing the champions' knowledge, self-efficacy, performance, and impact toward implementation. We will conduct a meta-analysis if there are more than 2 experimental or quasi-experimental studies that used the same outcome measures with similar populations and interventions [53,60]. The outcomes of interest, in accordance with Shea's conceptual model, intending to guide research on the activities and effects of innovation champions [7], will be (1) change in champions' knowledge and perceived self-efficacy in performing the champion role, and (2) improvement in the champions' performance and impact on the implementation process and outcome. If the outcome measures, populations, or interventions are heterogenous, then we will perform a narrative synthesis [60] of the data based on (1) health care sector (eg, acute care), (2) type of innovation, (3) quantitative research design, (4) outcome measures, (5) significance or lack of significance of results, and (6) quality of studies based on the critical appraisals. The meta-analysis or narrative synthesis will be performed by WJS; senior research members (AV, IDG, JES, or GH) will evaluate the analysis during weekly to biweekly meetings.

To address the fourth research objective and research questions 9-11, we will use conventional content analysis to synthesize the themes pertaining to the EDI considerations reported in the documents that discussed the nurse champion concept or nurse champion training initiatives and nurse champion training initiatives' effectiveness. We will also conduct summative content analysis [51] to describe the demographics of the nurse champions in accordance with equity-related factors detailed by the PROGRESS-plus framework [50] using frequency counts and descriptive statistics. The PROGRESS-plus framework is a widely used framework that describes socially stratifying categories that influence inequities in health care [50]. We will also use summative content analysis to describe the number of

XSL•FO
RenderX

included documents that discussed EDI considerations pertaining to the nurse champion concept or nurse champion training initiatives (using frequency counts or descriptive statistics). WJS will perform the initial coding that will be verified by another research team member (Letitia Nadalin-Penno, Megan Greenough, and Priscilla Packiam). The analysis will also be evaluated by senior research team members (AV, IDG, JES, or GH) during weekly to biweekly meetings.

Review Quality

We will use the PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist [65] and the Evidence-Based Checklist for Improving Scoping Review Quality [66] to ensure review quality and transparent reporting. We will also use the Standards for Reporting Qualitative Research checklist [67] for reporting the subanalyses pertaining to the nurse champion concept and the champion training initiatives, as these subanalyses involve qualitative analyses of textual data. We will use the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) checklist [68] to ensure transparent reporting of the subanalysis pertaining to the effectiveness of the champion training initiatives. We will use the Synthesis Without Meta-analysis in systematic reviews reporting checklist [69] if we conduct narrative synthesis, rather than a meta-analysis of the data extracted from studies reporting the effectiveness of champion training initiatives. We will report these checklists as an additional file in the resulting articles. A preliminary PRISMA-ScR checklist can be found in Checklist 1.

Results

We expect to present the results of the scoping review in 2026. As of January 2026, we have applied our search strategy on the databases we have selected and detailed above, piloted our inclusion and exclusion criteria for title and abstract and full-text screening, and partially completed title and abstract screening. Currently, our online database search has identified 4872 titles and abstracts (Figure 1). We will present the results of this scoping review across 4 peer-reviewed articles, with each article addressing one of our 4 research objectives and the research questions related to each of the objectives. As depicted in Figure 1 (an in-progress PRISMA diagram), we will specify the total number of records screened, the total number included in the scoping review, and the total number of studies included for the subsequent articles derived from the same unified dataset. We will cross-reference the articles included in the 4 articles to ensure accurate reporting of the total number included in the scoping review.

Figure 1. PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) flowchart. EDI: equity, diversity, and inclusion.

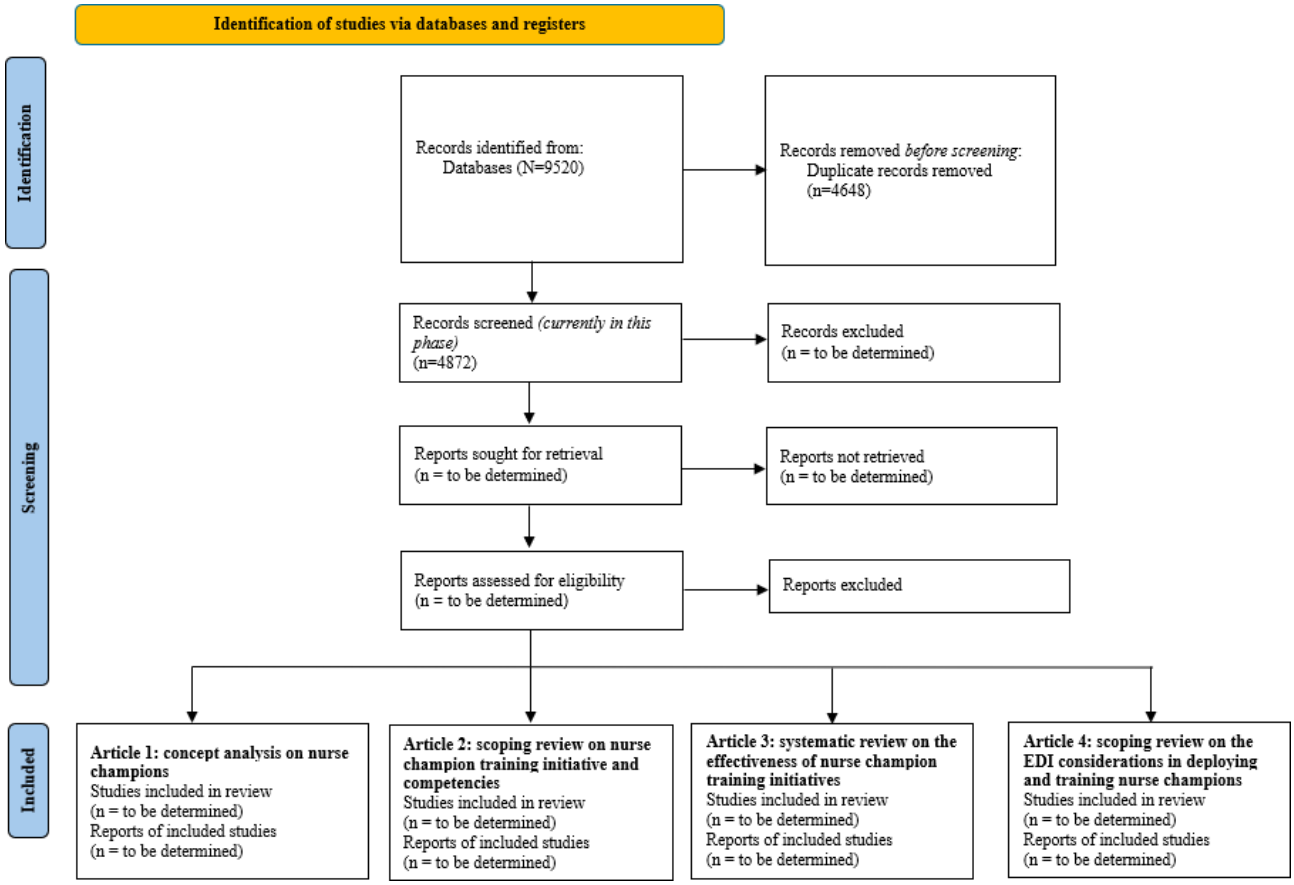

Discussion

Many authors have reported the importance of deploying nurse champions to enable implementation of best practices in health care [6,8,26-30]. The current literature pertaining to the concept of health care champions has advanced knowledge pertaining to the following: (1) differentiated health care champions from other concepts that enable implementation based on their theoretical origin and the mechanisms in which they exert social

XSL•FO
RenderX

influence [15,36], (2) described the plethora of personal attributes, skills, and behaviors that health care champions are reported to perform [6,15,37], and (3) illustrated conceptual models that described the relationships between different constructs that influence how health care champions enable change [7,38]. Since the use and application of concepts can change over time and with different contexts [39,40], there is an opportunity to explore how the nurse champion concept is defined across different health care contexts. Furthermore, there is a lack of knowledge pertaining to the training of nurse champions, the competencies that nurse champions should possess, and whether these training initiatives are effective [8,12]. There is also a lack of understanding of EDI considerations in research about nurse champions and health care champions [38]. This scoping review will (1) advance conceptual understanding of nurse champions, (2) collate existing nurse champion training initiatives and synthesize the competencies emphasized in these training initiatives, (3) synthesize the evidence pertaining to the effectiveness of the training initiatives that prepare nurse champions, and (4) summarize the EDI considerations pertaining to the nurse champion concept or the EDI considerations embedded in nurse champion training initiatives. This research will provide guidance on how to improve existing nurse champion training initiatives, which may result in better-prepared nurse champions and implementation success. Although the focus of this research is limited to nurse champions, it is possible that some of the findings may be transferable to other disciplines in health care.

Acknowledgments

The authors would like to acknowledge Victoria Cole for her guidance in developing the preliminary search strategy. We would also like to thank Letitia Nadalin-Penno, Megan Greenough, and Priscilla Packiam for being the second reviewers for this scoping review.

Funding

WJS is a recipient of the Queen Elizabeth II Excellence Award and the Ottawa Hospital's Sue Robblee Nursing Scholarship. IDG is a Canadian Institutes of Health Research Foundation Grant recipient (FDN 142337).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author on reasonable request.

Authors' Contributions

AV, IDG, GH, JES, WJS contributed to the conceptualization of the study. All the authors contributed to the development of the study design. WJS wrote the original draft of the manuscript. All of the authors critically reviewed and revised the manuscript regarding its content and approved of the version that is to be published and agree to be accountable for all aspects of the work.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRESS - nurse champions scoping review.
[DOCX File, 93 KB - resprot_v15i1e80204_app1.docx]

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 88 KB - resprot_v15i1e80204_app2.docx]

References

1. Braithwaite J, Glasziou P, Westbrook J. The three numbers you need to know about healthcare: the 60-30-10 challenge. BMC Med 2020 May 4;18(1):102. [doi: 10.1186/s12916-020-01563-4] [Medline: 32362273]
2. Squires JE, Cho-Young D, Aloisio LD, et al. Inappropriate use of clinical practices in Canada: a systematic review. CMAJ 2022 Feb 28;194(8):E279-E296. [doi: 10.1503/cmaj.211416] [Medline: 35228321]
3. Graham ID, Logan J, Harrison MB, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
4. Harrison MB, Graham ID. Knowledge Translation in Nursing and Healthcare: A Roadmap to Evidence-Informed Practice: John Wiley & Sons; 2021. [doi: 10.1002/9781119123316]
5. Harvey G, Loftus-Hills A, Rycroft-Malone J, et al. Getting evidence into practice: the role and function of facilitation. J Adv Nurs 2002 Mar;37(6):577-588. [doi: 10.1046/j.1365-2648.2002.02126.x] [Medline: 11879422]

XSL•FO
RenderX

6. Miech EJ, Rattray NA, Flanagan ME, Damschroder L, Schmid AA, Damush TM. Inside help: an integrative review of champions in healthcare-related implementation. SAGE Open Med 2018;6:2050312118773261. [doi: 10.1177/2050312118773261] [Medline: 29796266]
7. Shea CM. A conceptual model to guide research on the activities and effects of innovation champions. Implement Res Pract 2021 Jan 1;2:2633489521990443. [doi: 10.1177/2633489521990443] [Medline: 34541541]
8. Santos WJ, Graham ID, Lalonde M, Demery Varin M, Squires JE. The effectiveness of champions in implementing innovations in health care: a systematic review. Implement Sci Commun 2022 Jul 22;3(1):80. [doi: 10.1186/s43058-022-00315-0] [Medline: 35869516]
9. Astorino Nicola J, Nataliansyah MM, Lopez-Olivo MA, et al. Champions to enhance implementation of clinical and community-based interventions in cancer: a scoping review. Implement Sci Commun 2024 Oct 22;5(1):119. [doi: 10.1186/s43058-024-00662-0] [Medline: 39439009]
10. Wood K, Giannopoulos V, Louie E, et al. The role of clinical champions in facilitating the use of evidence-based practice in drug and alcohol and mental health settings: a systematic review. Implement Res Pract 2020;1:2633489520959072. [doi: 10.1177/2633489520959072] [Medline: 37089122]
11. Hall AM, Flodgren GM, Richmond HL, et al. Champions for improved adherence to guidelines in long-term care homes: a systematic review. Implement Sci Commun 2021 Aug 3;2(1):85. [doi: 10.1186/s43058-021-00185-y] [Medline: 34344486]
12. Jolliffe L, Lannin NA, Larcombe S, Major B, Hoffmann T, Lynch E. Training and education provided to local change champions within implementation trials: a rapid systematic review. Implement Sci 2025 Feb 5;20(1):8. [doi: 10.1186/s13012-025-01416-9] [Medline: 39910625]
13. Chakrabarti AK. The role of champion in product innovation. Calif Manage Rev 1974 Dec;17(2):58-62. [doi: 10.2307/41164561]
14. Schon DA. Champions for radical new inventions. Harv Bus Rev 1963;41:77-86 [FREE Full text]
15. Cranley LA, Cummings GG, Profetto-McGrath J, Toth F, Estabrooks CA. Facilitation roles and characteristics associated with research use by healthcare professionals: a scoping review. BMJ Open 2017 Aug 11;7(8):e014384. [doi: 10.1136/bmjopen-2016-014384] [Medline: 28801388]
16. Kuehl H, Mabry L, Elliot DL, Kuehl KS, Favorite KC. Factors in adoption of a fire department wellness program: champ-and-chief model. J Occup Environ Med 2013 Apr;55(4):424-429. [doi: 10.1097/JOM.0b013e31827dba3f] [Medline: 23302702]
17. Shaw EK, Howard J, West DR, et al. The role of the champion in primary care change efforts: from the State Networks of Colorado Ambulatory Practices and Partners (SNOCAP). J Am Board Fam Med 2012;25(5):676-685. [doi: 10.3122/jabfm.2012.05.110281] [Medline: 22956703]
18. Soo S, Berta W, Baker GR. Role of champions in the implementation of patient safety practice change. Healthc Q 2009;12 Spec No Patient:123-128. [doi: 10.12927/hcq.2009.20979] [Medline: 19667789]
19. Jennings G. Introducing learning disability champions in an acute hospital. Nurs Times 2019;115(4):44-47 [FREE Full text]
20. Lee SJC, Higashi RT, Inrig SJ, et al. County-level outcomes of a rural breast cancer screening outreach strategy: a decentralized hub-and-spoke model (BSPAN2). Transl Behav Med 2017 Jun;7(2):349-357. [doi: 10.1007/s13142-016-0427-3] [Medline: 27402023]
21. Luz S, Shadmi E, Admi H, Peterfreund I, Drach-Zahavy A. Characteristics and behaviours of formal versus informal nurse champions and their relationship to innovation success. J Adv Nurs 2019 Jan;75(1):85-95. [doi: 10.1111/jan.13838] [Medline: 30168170]
22. Maier MA, Brem A. What innovation managers really do: a multiple-case investigation into the informal role profiles of innovation managers. Rev Manag Sci 2018 Oct;12(4):1055-1080. [doi: 10.1007/s11846-017-0238-z]
23. Titler MG, Everett LQ. Translating research into practice. Considerations for critical care investigators. Crit Care Nurs Clin North Am 2001 Dec;13(4):587-604. [doi: 10.1016/S0899-5885(18)30026-1] [Medline: 11778346]
24. Melnyk BM, Fineout-Overholt E, Stillwell SB, Williamson KM. Evidence-based practice: step by step: the seven steps of evidence-based practice. Am J Nurs 2010 Jan;110(1):51-53. [doi: 10.1097/01.NAJ.0000366056.06605.d2] [Medline: 20032669]
25. Aloisio LD, Graham N, Grinspun D, et al. Indicators to measure implementation and sustainability of nursing best practice guidelines: a mixed methods analysis. Heliyon 2023 Sep;9(9):e19983. [doi: 10.1016/j.heliyon.2023.e19983] [Medline: 37809679]
26. Chisholm A, Russolillo A, Carter M, et al. Advancing evidence-based practice through the knowledge translation challenge: nurses' important roles in research, implementation science and practice change. J Adv Nurs 2025 Nov;81(11):7285-7293. [doi: 10.1111/jan.16362] [Medline: 39087775]
27. Campbell J. The effect of nurse champions on compliance with keystone intensive care unit sepsis-screening protocol. Crit Care Nurs Q 2008;31(3):251-269. [doi: 10.1097/01.CNQ.0000325050.91473.0b] [Medline: 18574373]
28. Anand KJS, Eriksson M, Boyle EM, et al. Assessment of continuous pain in newborns admitted to NICUs in 18 European countries. Acta Paediatr 2017 Aug;106(8):1248-1259. [doi: 10.1111/apa.13810] [Medline: 28257153]

XSL•FO
RenderX

29. Bentz CJ, Bayley KB, Bonin KE, et al. Provider feedback to improve 5A's tobacco cessation in primary care: a cluster randomized clinical trial. Nicotine Tob Res 2007 Mar;9(3):341-349. [doi: 10.1080/14622200701188828] [Medline: 17365766]
30. Engel M, van Zuylen L, van der Ark A, van der Heide A. Palliative care nurse champions' views on their role and impact: a qualitative interview study among hospital and home care nurses. BMC Palliat Care 2021 Feb 18;20(1):34. [doi: 10.1186/s12904-021-00726-1] [Medline: 33602177]
31. White CL. Nurse champions: a key role in bridging the gap between research and practice. J Emerg Nurs 2011 Jul;37(4):386-387. [doi: 10.1016/j.jen.2011.04.009] [Medline: 21621832]
32. Siebeck OS, Hoving C. Characteristics of a successful nurse peer champion in the implementation of innovative digital technologies in hospitals: a qualitative study. PEC Innov 2024 Dec 15;5:100339. [doi: 10.1016/j.pecinn.2024.100339] [Medline: 39380690]
33. Sperling D, Shadmi E, Drach-Zahavy A, Luz S. Nurse champions as street-level bureaucrats: factors which facilitate innovation, policy making, and reconstruction. Front Psychol 2022;13:872131. [doi: 10.3389/fpsyg.2022.872131] [Medline: 36081722]
34. Ploeg J, Skelly J, Rowan M, et al. The role of nursing best practice champions in diffusing practice guidelines: a mixed methods study. Worldviews Evid Based Nurs 2010 Dec;7(4):238-251. [doi: 10.1111/j.1741-6787.2010.00202.x] [Medline: 20880009]
35. Harvey G, Kitson A. PARIHS revisited: from heuristic to integrated framework for the successful implementation of knowledge into practice. Implement Sci 2016 Mar 10;11(1):33. [doi: 10.1186/s13012-016-0398-2] [Medline: 27013464]
36. Thompson GN, Estabrooks CA, Degner LF. Clarifying the concepts in knowledge transfer: a literature review. J Adv Nurs 2006 Mar;53(6):691-701. [doi: 10.1111/j.1365-2648.2006.03775.x] [Medline: 16553677]
37. Bonawitz K, Wetmore M, Heisler M, et al. Champions in context: which attributes matter for change efforts in healthcare? Implement Sci 2020 Aug 6;15(1):62. [doi: 10.1186/s13012-020-01024-9] [Medline: 32762726]
38. Morena AL, Gaias LM, Larkin C. Understanding the role of clinical champions and their impact on clinician behavior change: the need for causal pathway mechanisms. Front Health Serv 2022;2:896885. [doi: 10.3389/frhs.2022.896885] [Medline: 36925794]
39. Rodgers BL, Knafl KA. Concept Development in Nursing: Foundations, Techniques, and Applications, 2nd edition: Saunders; 2000.
40. Rodgers BL. Concepts, analysis and the development of nursing knowledge: the evolutionary cycle. J Adv Nurs 1989 Apr;14(4):330-335. [doi: 10.1111/j.1365-2648.1989.tb03420.x] [Medline: 2661622]
41. Straus S, Tetroe J, Bhattacharyya O, Zwarenstein M, Graham I. Chapter 3.5: monitoring knowledge use and evaluating outcomes. In: Knowledge Translation in Health Care: Moving from Evidence to Practice, 2nd edition: Wiley Books; 2013:227-236. [doi: 10.1002/9781118413555]
42. Mrayyan MT, Abunab HY, Abu Khait A, et al. Competency in nursing practice: a concept analysis. BMJ Open 2023 Jun 1;13(6):e067352. [doi: 10.1136/bmjopen-2022-067352] [Medline: 37263688]
43. McKeganey SN. The rise and rise of peer education approaches. Drugs Educ Prev Policy 2000 Jan;7(3):293-310. [doi: 10.1080/09687630050109961]
44. Heidenreich S, Breukers S. Who is telling whose story? The effectiveness of peer-to-peer approaches as inclusive participatory interventions towards sustainability. Sustain Prod Consum 2020 Jan;21:216-227. [doi: 10.1016/j.spc.2019.10.001]
45. Hendy J, Barlow J. The role of the organizational champion in achieving health system change. Soc Sci Med 2012 Feb;74(3):348-355. [doi: 10.1016/j.socscimed.2011.02.009] [Medline: 21444137]
46. Koenig AM, Eagly AH, Mitchell AA, Ristikari T. Are leader stereotypes masculine? A meta-analysis of three research paradigms. Psychol Bull 2011 Jul;137(4):616-642. [doi: 10.1037/a0023557] [Medline: 21639606]
47. Peters MD, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis 2020. [doi: 10.46658/JBIMES-24-09]
48. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
49. The JBI approach in evidence implementation. JBI. URL: https://jbi.global/news/article/jbi-approach-evidence-implementation [accessed 2025-05-29]
50. O'Neill J, Tabish H, Welch V, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. J Clin Epidemiol 2014 Jan;67(1):56-64. [doi: 10.1016/j.jclinepi.2013.08.005] [Medline: 24189091]
51. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
52. Presseau J, McCleary N, Lorencatto F, Patey AM, Grimshaw JM, Francis JJ. Action, actor, context, target, time (AACTT): a framework for specifying behaviour. Implement Sci 2019 Dec 5;14(1):102. [doi: 10.1186/s13012-019-0951-x] [Medline: 31806037]

XSL•FO
RenderX

53. Deeks JJ, Higgins JP, Altman DG, Group CSM. Analysing data and undertaking meta‐analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions 2019:241-284. [doi: 10.1002/9781119536604]
54. Using Microsoft Copilot at uOttawa. University of Ottawa. URL: https://saea-tlss.uottawa.ca/en/teaching-technologies/microsoft-copilot [accessed 2025-12-24]
55. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75(40–6):40-46. [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
56. Best Practice Champions Network. Registered Nurses' Association of Ontario. URL: https://rnao.ca/bpg/champions [accessed 2024-09-11]
57. Champions. SPA LTC. URL: https://spaltc.ca/champions [accessed 2024-09-11]
58. What is inter-rater reliability? Covidence. URL: https://www.covidence.org/blog/what-is-inter-rater-reliability [accessed 2025-11-14]
59. Better systematic review management. Covidence. 2025. URL: https://www.covidence.org [accessed 2025-05-02]
60. Tufanaru C, Munn Z, Aromataris E, Campbell J, Hopp L. Systematic reviews of effectiveness. In: Aromataris E, Munn Z, editors. JBI Reviewer's Manual 2020. [doi: 10.46658/JBIRM-17-03]
61. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, et al. Systematic reviews of etiology and risk. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis JBI 2020. [doi: 10.46658/JBIMES-24-06]
62. McArthur A, Klugarova J, Yan H, Florescu S. Chapter 4: systematic reviews of text and opinion. In: JBI Manual for Evidence Synthesis JBI 2020, Vol. 10:134-174. [doi: 10.46658/JBIRM-17-04]
63. Lockwood C, Munn Z, Porritt K. Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. Int J Evid Based Healthc 2015 Sep;13(3):179-187. [doi: 10.1097/XEB.0000000000000062] [Medline: 26262565]
64. Hong QN, Gonzalez-Reyes A, Pluye P. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the mixed methods appraisal tool (MMAT). J Eval Clin Pract 2018 Jun;24(3):459-467. [doi: 10.1111/jep.12884] [Medline: 29464873]
65. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
66. Cooper S, Cant R, Kelly M, et al. An evidence-based checklist for improving scoping review quality. Clin Nurs Res 2021 Mar;30(3):230-240. [doi: 10.1177/1054773819846024] [Medline: 31088144]
67. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. Acad Med 2014 Sep;89(9):1245-1251. [doi: 10.1097/ACM.0000000000000388] [Medline: 24979285]
68. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021 Mar 29;372:n71. [doi: 10.1136/bmj.n71] [Medline: 33782057]
69. Campbell M, McKenzie JE, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline. BMJ 2020 Jan 16;368:l6890. [doi: 10.1136/bmj.l6890] [Medline: 31948937]

Abbreviations

EDI: equity, diversity, and inclusion
JBI: Joanna Briggs Institute
PRESS: Peer Review of Electronic Search Strategy
PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
PROGRESS-plus: place of residence; race, ethnicity, culture, and language; occupation; gender and sex; religion; education; socioeconomic status; and social capital–plus

Edited by J Sarvestan; submitted 06.Jul.2025; peer-reviewed by H Thongdaeng, M Irfan; revised version received 05.Dec.2025; accepted 08.Dec.2025; published 14.Jan.2026.

Please cite as:
Santos WJ, Graham ID, Vandyk A, Harvey G, Squires JE
Understanding the Nurse Champion Concept and the Training Initiatives for Nurse Champions: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e80204
URL: https://www.researchprotocols.org/2026/1/e80204
doi:10.2196/80204

XSL•FO
RenderX

© Wilmer J Santos, Ian D Graham, Amanda Vandyk, Gillian Harvey, Janet E Squires. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 14.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Efficacy and Safety of Sodium-Glucose Cotransporter-2 Inhibitors in Patients With Refractory Cirrhotic Ascites and Diabetes: Scoping Review Protocol

Tooba Laeeq[1], MD; Ayesha Mehak Kamal[2]; Arbab Burhan Uddin Kasi[3]; Umm E Salma Shabbar Banatwala[2]; Aniqa Baloch[2]; Roberto Sagaribay[4], PhD; Vidhani Goel[4,5], BS; Aditi Singh[1], MD; Kavita Batra[4,6], PhD

[1]Department of Internal Medicine, Kirk Kerkorian School of Medicine at UNLV, University of Nevada, 9850 S Maryland PKWY, Las Vegas, NV, United States

[2]Dow University of Health Sciences, New Labour Colony Nanakwara, Karachi, Pakistan

[3]Mass General Brigham, Boston, MA, United States

[4]Office of Research, Kirk Kerkorian School of Medicine at UNLV, University of Nevada, Las Vegas, NV, United States

[5]School of Public Health, University of Nevada, Las Vegas, NV, United States

[6]Department of Medical Education, Kirk Kerkorian School of Medicine at UNLV, University of Nevada, Las Vegas, NV, United States

Corresponding Author:
Tooba Laeeq, MD
Department of Internal Medicine, Kirk Kerkorian School of Medicine at UNLV, University of Nevada, 9850 S Maryland PKWY, Las Vegas, NV, United States

Abstract

Background: Hepatic cirrhosis is a complex condition leading to multiple complications, including ascites, hepatic encephalopathy, bleeding varices, and eventually liver failure. Patients with diabetes mellitus or insulin resistance are more likely to fail treatment, leading to the worsening of hepatic fibrosis. Sodium-glucose cotransporter-2 (SGLT2) inhibitors are a new class of drugs with the potential for use in cirrhotic ascites. This scoping review will focus on the response of the addition of SGLT2 inhibitors (SGLT2i) on refractory ascites.

Objective: The objective of this scoping review is to understand the extent and type of evidence in relation to the response of the addition of SGLT2i on refractory ascites in patients with diabetes mellitus or insulin resistance.

Methods: A comprehensive literature search will be conducted across five databases using variations of the keywords "Sodium-glucose Cotransporter-2 Inhibitors" and "ascites" to identify original studies published from inception through December 2023. The search will be limited to English-language human studies evaluating the efficacy of SGLT2 inhibitors, excluding studies focused on other drugs or conducted in animal models. This review will include studies involving patients over 21 years of age with cirrhotic ascites and will exclude studies involving noncirrhotic causes. Only studies containing original patient data will be considered, regardless of health care setting, to ensure clinical relevance and consistency in interpretation.

Results: The results of this scoping review will offer a comprehensive overview of the available evidence on the use of SGLT2i in the management of cirrhotic ascites. This review will summarize key study characteristics, patient populations, interventions, and outcomes, and identify existing knowledge gaps and research priorities. Findings will be presented narratively and in tabular format to facilitate interpretation and comparison. The review process began with protocol development and registration in February 2025. A structured literature search and database screening will follow, continuing through June. Title and abstract screening, along with full-text review and eligibility assessment, will be completed by July 2025. Data extraction and charting will take place in July, with synthesis and analysis of findings occurring through August. Manuscript drafting will begin in August, and the final review and submission of the completed scoping review are anticipated by September 2025.

Conclusions: By systematically identifying and summarizing original studies that assess the efficacy of SGLT2i in this unique patient population, the review will provide insight into the therapeutic potential of this drug class beyond glycemic control. It will also highlight the strengths and limitations of the existing literature, identify key knowledge gaps, and suggest directions for future research. Ultimately, this review may support the rationale for further clinical investigation and the development of targeted therapies for a high-risk group of patients with limited treatment options.

Trial Registration: Open Science Framework (OSF) 7eycr; https://osf.io/7eycr

International Registered Report Identifier (IRRID): PRR1-10.2196/77587

(JMIR Res Protoc 2026;15:e77587) doi:10.2196/77587

XSL•FO
RenderX

KEYWORDS

sodium-glucose; citransporter-2; diabetes; cirrhotic ascites; diabetes mellitus

Introduction

Liver cirrhosis is a progressive, chronic disease characterized by diffuse hepatic fibrosis and the formation of regenerative nodules, resulting from various etiologies such as viral hepatitis, alcohol-related liver disease, and nonalcoholic steatohepatitis [1-3]. It is a major global health concern, with complications including ascites, hepatic encephalopathy, variceal bleeding, systemic inflammation, and eventual liver failure [2-4]. Ascites, the most common complication, significantly impacts prognosis; approximately half of affected patients die within three years [4]. Globally, cirrhosis-related deaths increased by 8.1% between 2010 and 2019, reaching nearly 1.5 million annually [5], underscoring the urgent need for more effective management strategies.

The standard approach to ascites management includes sodium restriction and diuretic therapy, with large-volume paracentesis reserved for refractory cases [6]. However, these options often become limited due to diuretic resistance, hyponatremia, or renal impairment. An important comorbidity influencing the course of cirrhosis is type 2 diabetes mellitus (T2DM), which is present in roughly one-third of patients and associated with poorer outcomes, including increased risk of ascites development and mortality [7-9]. More than 80% of individuals with cirrhosis have some degree of glucose intolerance [9]. Recent global estimates indicate that T2DM affects approximately 30% - 40% of individuals with cirrhosis, and its presence is associated with a two-fold increase in mortality risk compared to nondiabetic cirrhotic patients [7-9]. This combination not only accelerates hepatic decompensation but also increases the likelihood of developing refractory ascites, underscoring the need for targeted management strategies. Insulin resistance in these patients accelerates hepatic fibrosis and impairs treatment response [10,11].

Selecting safe and effective antihyperglycemic therapy in the setting of chronic liver disease is challenging due to altered drug metabolism and potential hepatotoxicity [12]. While metformin improves insulin sensitivity, conflicting evidence exists regarding its safety in cirrhosis [12-14]. Other agents, such as sulfonylureas and alpha-glucosidase inhibitors, are generally avoided in decompensated cirrhosis, and insulin therapy, though often used, requires further evidence to support its efficacy [12].

Sodium-glucose cotransporter-2 inhibitors (SGLT2i) are antihyperglycemic agents that block renal glucose reabsorption in the proximal tubule, resulting in osmotic diuresis and mild natriuresis [15-17]. These effects may help improve ascites by reducing sodium retention and modulating the renin-angiotensin-aldosterone system [18,19]. Emerging evidence, including preliminary reports in cirrhotic patients with refractory ascites, suggests that SGLT2i may improve ascites management [18,20]. They may also confer additional hepatic benefits, such as improved liver enzyme profiles in patients with coexisting nonalcoholic fatty liver disease [21,22]. However, safety concerns, such as increased risk of genital infections, urinary tract infections, dehydration, acute kidney injury, and euglycemic ketoacidosis, warrant careful patient selection and monitoring [19,23-27].

Despite growing clinical use of SGLT2i in various cardiometabolic conditions, there remains a lack of synthesized evidence specifically evaluating their role in cirrhotic ascites with concomitant diabetes or insulin resistance. No systematic or scoping reviews on this topic were identified in preliminary searches. This scoping review aims to map the extent and type of available evidence regarding the efficacy and safety of SGLT2i in this high-risk population, identify knowledge gaps, and inform the design of future clinical trials. By systematically mapping the current evidence, this review could provide foundational knowledge to guide updates in clinical practice guidelines for managing refractory ascites in patients with cirrhosis and diabetes. Additionally, the synthesis of safety and efficacy data may inform policy decisions regarding the inclusion of SGLT2i as a therapeutic consideration in specific high-risk subgroups, thereby supporting more evidence-based and targeted treatment strategies.

An additional consideration is the issue of diuretic resistance, which frequently complicates the management of refractory ascites. Emerging evidence from heart failure populations suggests that SGLT2i may enhance diuretic responsiveness by promoting natriuresis and improving renal sodium handling [28]. While this mechanism has not been extensively studied in cirrhotic patients, it represents a promising area for investigation and further underscores the potential clinical utility of SGLT2i in the setting of advanced liver disease with fluid overload.

Objectives

This review aims to examine the impact of SGLT2i on refractory ascites in patients with cirrhosis and coexisting diabetes mellitus or insulin resistance, with a focus on reported safety considerations and adverse events. Additionally, we will assess their effects on clinical outcomes such as glycemic control, fluid status, renal function, and mortality, and explore how their efficacy and safety profile compares to other glucose-lowering agents in this population.

Methods

Protocol Registration

The protocol for this scoping review was officially registered on the Open Science Framework (OSF) Registry to ensure transparency, reproducibility, and adherence to established guidelines for systematic reviews. The registration serves as a public record of the review's objectives, methodology, and planned analysis, allowing for open access to the protocol and facilitating future updates or amendments if needed. The registered protocol is accessible through the following link (10.17605/OSF.IO/7EYCR) [28]. By registering the protocol in this open-access platform, we aim to promote the integrity

XSL•FO
RenderX

of the review process and provide a clear framework for the research that can be scrutinized and reproduced by the scientific community.

Eligibility Criteria

The proposed scoping review will follow the PCC framework (Population, Concept, Context) [29] for structuring review questions, identifying relevant literature, and developing inclusion criteria (Figure 1). The results of the search will be reported and presented using Preferred Reporting Items for Systematic Reviews and Meta-analyses extension for scoping review (PRISMA-ScR) guidelines and flow diagram [30,31]. A scoping review design was selected rather than a systematic review because the literature on SGLT2i use in cirrhotic ascites is limited in volume, varied in methodological approach, and often exploratory. Scoping reviews are particularly suited to mapping the extent and nature of research activity, clarifying key concepts, and identifying evidence gaps in emerging fields [29,30]. This approach enables the inclusion of a broader range of study designs, including case reports and small observational studies, which may provide valuable insights for a condition where large-scale randomized trials are scarce.

Figure 1. Inclusion criteria based on the participants' concept context framework. SGLT2: Sodium-glucose cotransporter-2.

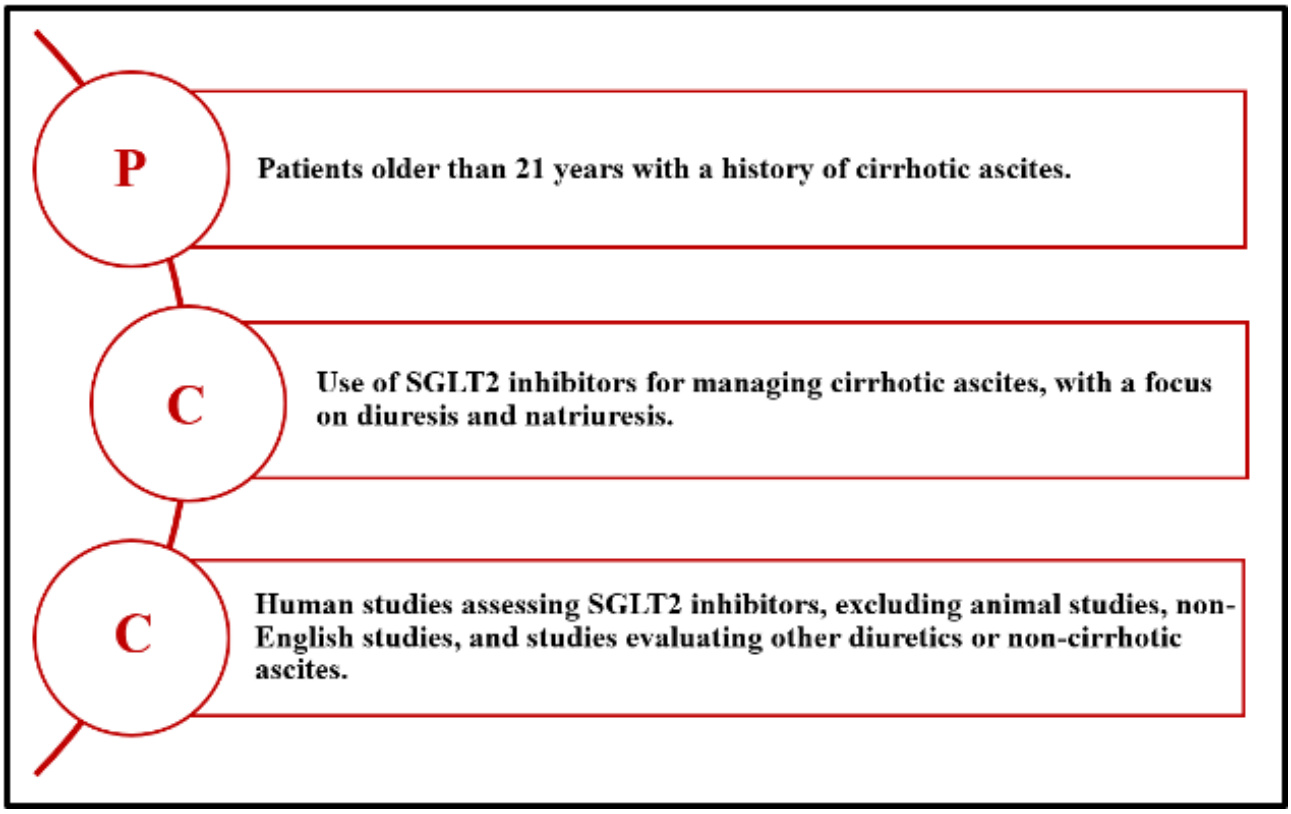

In terms of participants, patients older than 21 years with a history of cirrhotic ascites will be included. Studies involving patients with ascites due to noncirrhotic causes will be excluded, as will be animal studies, to ensure the findings are relevant to human clinical populations. Regarding the concept, studies evaluating the use of SGLT2i in the management of cirrhotic ascites will be included. Only studies that provide patient data will be included to ensure clinically applicable results. Studies that did not report patient data or primarily assessed diuretics other than SGLT2i, will be excluded to maintain the specificity of the review. In terms of context, this review will consider studies conducted in any health care setting that assess the impact of SGLT2i on cirrhotic ascites. Further, only studies published in English will be included due to the unavailability of translation resources within the review team, which may introduce a risk of language bias.

Types of Sources

This scoping review will consider both experimental and quasi-experimental study designs, including randomized controlled trials, nonrandomized controlled trials, before and after studies, and interrupted time-series studies. In addition, analytical observational studies including prospective and retrospective cohort studies, case-control studies, and analytical cross-sectional studies will be considered for inclusion. This review will also consider descriptive observational study designs including case series, individual case reports, and descriptive cross-sectional studies for inclusion. Reviews and meta-analyses will be excluded.

Search Strategy

A comprehensive literature search across five databases (PubMed/Medline, Embase, CINAHL, Cochrane, and Web of Science), using variations of the keywords "Sodium-glucose Cotransporter-2 Inhibitors" and "ascites," will be used to identify original studies published from inception through December 2023. Results will be limited to human studies published in English (see Table 1 for full inclusion and exclusion criteria). The text words contained in the titles and abstracts of relevant articles, and the index terms used to describe the articles, will be used to develop a full search strategy (see Table 2, data extraction framework). The search strategy, including all

XSL•FO
RenderX

identified keywords and index terms, will be adapted for each included database and information source. A draft of the search strategies for each database is provided in Table 3. The reference list of all included sources of evidence will be screened for additional studies.

Table . Inclusion and exclusion criteria for study selection.

Criteria	Inclusion	Exclusion
Population	Patients older than 21 years with a history of cirrhotic ascites	Patients younger than 21 years or those without cirrhotic ascites
Intervention	Use of SGLT2[a] inhibitors for treating cirrhotic ascites	Use of diuretics other than SGLT2 inhibitors
Study type	Original human studies with patient data	Animal studies or studies without patient data
Language	Studies published in English	Non-English studies
Condition	Ascites due to cirrhosis	Ascites from other causes

[a]SGLT2: Sodium-glucose cotransporter-2.

Table . Data extraction framework.

Extraction criteria	Description
Citation	The authors' name, title of paper, year published, and journal name
Type of study	General empirical approach
Duration	Length of time patients participated in the study
Sample size (n)	The number of participants in each study
Patient's age	Mean age of participants in each study
Patient's gender	Gender breakdown (ie, male/female) for each study
Country	The country where the study was conducted
Cause of cirrhosis	ie, PBC[a], alcohol, or nonalcoholic steatohepatitis
Type of SGLT2i[b]	ie, empagliflozin, dapagliflozin, or canagliflozin
Comparator	Comparison group that uses a treatment other than SGLT2i
Primary outcome	Primary outcomes found in each study
Secondary outcome	Secondary outcomes found in each study
Key findings	Summary of overall findings and implications
Comparator intervention(s)	Type of comparator reported (eg, insulin, metformin, placebo, diuretic therapy); outcomes summarized descriptively and narratively. No quantitative synthesis or comparative effectiveness analysis planned
Adverse effects	Summary of adverse effects of SGLT2i (eg, genital infections, UTIs[c], ketoacidosis, hypoglycemia, acute kidney injury, volume depletion); note if severity grading (eg, CTCAE[d] or study-defined)

[a]PBC: Primary biliary cholangitis.

[b]SGLT2i: Sodium-glucose cotransporter-2 inhibitors.

[c]UTI: urinary tract infection.

[d]CTCAE: Common Terminology Criteria for Adverse Events.

XSL•FO
RenderX

Table . Draft search strategies by database.

Database	Search strategy (draft)
PubMed (MEDLINE)	(“Sodium-Glucose Transporter 2 Inhibitors”[Mesh] OR “SGLT2 inhibitors” OR empagliflozin OR dapagliflozin OR canagliflozin OR ertugliflozin) AND (“Ascites”[Mesh] OR ascites OR “fluid retention” OR “refractory ascites”) AND (“Liver Cirrhosis”[Mesh] OR cirrhosis OR “hepatic cirrhosis”)
Embase	(’sodium glucose cotransporter 2 inhibitor’/exp OR ’sglt2 inhibitor’ OR empagliflozin OR dapagliflozin OR canagliflozin OR ertugliflozin) AND ('ascites’/exp OR ascites OR ’fluid retention’ OR ’refractory ascites’) AND ('liver cirrhosis’/exp OR cirrhosis OR ’hepatic cirrhosis’)
Scopus	(TITLE-ABS-KEY (“SGLT2 inhibitors” OR empagliflozin OR dapagliflozin OR canagliflozin OR ertugliflozin) AND TITLE-ABS-KEY (ascites OR “fluid retention” OR “refractory ascites”) AND TITLE-ABS-KEY (cirrhosis OR “hepatic cirrhosis” OR “liver cirrhosis”))
Web of Science	TS= (“SGLT2 inhibitors” OR empagliflozin OR dapagliflozin OR canagliflozin OR ertugliflozin) AND TS= (ascites OR “fluid retention” OR “refractory ascites”) AND TS=(cirrhosis OR “hepatic cirrhosis” OR “liver cirrhosis”)
Cochrane Library	(“SGLT2 inhibitors” OR empagliflozin OR dapagliflozin OR canagliflozin OR ertugliflozin):ti,ab,kw AND (ascites OR “fluid retention” OR “refractory ascites”):ti,ab,kw AND (cirrhosis OR “hepatic cirrhosis” OR “liver cirrhosis”):ti,ab,kw

Study or Source of Evidence Selection

Following the search, all identified citations will be organized and duplicates removed. Following a pilot test, titles and abstracts will then be screened by two independent reviewers for assessment against the inclusion criteria for the review. Potentially relevant sources will be retrieved in full. The full text of selected citations will be assessed in detail against the inclusion criteria by two independent reviewers. Reasons for exclusion of sources of evidence at full text that do not meet the inclusion criteria will be recorded and reported in the scoping review. Any disagreements that arise between the reviewers at each stage of the selection process will be resolved through discussion or with an additional reviewer. The results of the search and the study inclusion process will be reported in full in the final scoping review.

Data Extraction and Synthesis

Data will be extracted from all included studies by two independent reviewers using a predefined data extraction framework (see Table 2). The draft extraction tool will first be pilot-tested on a sample of studies to ensure clarity, consistency, and comprehensiveness. Inter-reviewer reliability will be assessed during the pilot phase (eg, via percent agreement), and discrepancies will be discussed and resolved to refine the tool before full data extraction begins. The finalized extraction form may be iteratively modified during the review process, with all modifications documented. Extracted information will include study characteristics, population details, interventions, outcomes (efficacy and safety) and, where relevant, comparator interventions (eg, insulin, metformin, placebo, diuretics). Comparator data will be summarized descriptively and narratively, without formal statistical pooling. Any disagreements during the extraction process will be resolved through discussion or, if needed, by consultation with a third reviewer.

The data extracted from the identified studies will allow us to map existing research outcomes and highlight gaps in the current literature. Primary findings will be systematically organized into overarching themes and sub-themes, including efficacy, safety, and comparator outcomes. Although formal grading of evidence quality will not be undertaken, results will be organized and presented by study design (eg, randomized controlled trials, nonrandomized trials, cohort studies, case-control studies, case series, and case reports) to provide readers with context regarding the methodological rigor of each source.

Results

The results of this scoping review will provide a comprehensive synthesis of the available evidence regarding the use of SGLT2i in managing cirrhotic ascites. The review will encompass a wide range of studies and will analyze various key elements such as study characteristics, patient demographics, interventions, and clinical outcomes. By focusing on these aspects, the review aims to deliver insights into the effectiveness of SGLT2i in improving refractory ascites in patients with cirrhosis, particularly in those with comorbid diabetes mellitus or insulin resistance.

The review will also highlight the gaps in the current literature, shedding light on areas where further research is needed to optimize treatment strategies and improve patient outcomes. In addition to presenting the findings narratively, a tabular format will be used to systematically summarize the data, facilitating easy comparison across studies and helping identify trends, variations, and patterns in the use of SGLT2i for cirrhotic ascites management.

The anticipated findings will not only contribute to our understanding of the therapeutic potential of SGLT2i in this context but also provide valuable information for clinicians

XSL•FO
RenderX

seeking to make evidence-based decisions in managing patients with cirrhosis and ascites. To ensure a structured and efficient review process, the following timeline has been established (as shown in Table 4). This timeline provides a clear and systematic framework for the execution of the scoping review, ensuring that each phase is completed within the designated time frame. Regular progress checks and adjustments will be made as needed to ensure the review is thorough, efficient, and on schedule.

Table . The suggested timeline of the proposed scoping review.

Milestone	Timeline	Details
Protocol development and registration	February-May 2025	The review protocol was developed and registered on the Open Science Framework (OSF) Registry to ensure transparency and adherence to systematic review guidelines
Structured literature search	June 2025	Comprehensive literature search and database screening across five platforms to identify relevant studies from inception through December 2023
Title and abstract screening	July 2025	Initial screening of identified studies based on title and abstract to determine relevance
Full-text review and eligibility assessment	August 2025	In-depth review of selected studies to assess eligibility according to the inclusion and exclusion criteria
Data extraction and charting	September-November 2025	Systematic extraction of key study data, including patient characteristics, interventions, and outcomes
Synthesis and analysis of findings	December 2025	Analysis of extracted data, synthesis of results, and identification of trends, gaps, and key findings
Manuscript drafting	January-February 2026	Drafting of the manuscript for the scoping review, incorporating the synthesized findings
Final review and submission	March 2026	Final review and submission of the completed scoping review to a peer-reviewed journal

Discussion

Principal Findings

This scoping review seeks to provide a comprehensive evaluation of the available evidence regarding the use of SGLT2i in the management of cirrhotic ascites. As cirrhosis and ascites are associated with significant morbidity and limited therapeutic options, particularly in patients with comorbid diabetes or insulin resistance, SGLT2i offer a potentially valuable adjunctive therapy due to their effects on glucose control, renal function, and fluid balance.

Initial studies suggest that SGLT2i may improve fluid status by promoting natriuresis and reducing renal sodium retention, a key mechanism in the pathogenesis of ascites [18]. Moreover, SGLT2i are known to have beneficial effects on glycemic control and may offer an additional advantage in cirrhotic patients who often struggle with diabetes management [32]. However, while their role in diabetes management in cirrhosis is well documented, their specific impact on refractory ascites remains less clear.

The existing body of literature on this topic is still developing, with limited studies specifically addressing the use of SGLT2i in the context of cirrhotic ascites. Early evidence suggests that while these agents may provide therapeutic benefits in terms of fluid balance and glucose control, safety concerns, such as the risk of dehydration, kidney injury, and ketoacidosis, are prevalent in cirrhotic patients, particularly those with advanced liver disease [33]. This highlights the need for careful patient selection and monitoring in clinical practice.

Moreover, the comparative effectiveness of SGLT2i versus other glucose-lowering therapies in cirrhotic patients with ascites has not been fully explored, necessitating future research to establish their relative safety and efficacy. This review will help identify knowledge gaps and will guide future studies, especially randomized controlled trials, to further evaluate the potential benefits and risks of SGLT2i in this complex patient population.

Another potential clinical implication relates to diuretic resistance, which remains a significant challenge in managing refractory ascites. Although direct evidence in cirrhotic populations is limited, studies in heart failure suggest that SGLT2i may enhance diuretic responsiveness through natriuresis and improved renal sodium handling [34]. This raises the possibility that SGLT2i could serve as an adjunctive therapy for patients with cirrhosis and fluid overload who have inadequate responses to conventional diuretics. Future clinical research should explore this mechanism to determine its translational relevance in cirrhotic populations.

Limitations

Restricting the review to English-language publications may have excluded relevant studies in other languages, introducing potential language bias. The inclusion of diverse study designs,

XSL•FO
RenderX

including case reports and small observational studies, may limit the ability to draw firm conclusions about causality or comparative effectiveness. Additionally, as the evidence base is expected to be heterogeneous in methodology and outcome reporting, formal meta-analysis will not be conducted, and findings will be presented narratively, which may limit quantitative synthesis.

Conclusions

This scoping review will provide a comprehensive overview of the current evidence surrounding the use of SGLT2i in managing cirrhotic ascites, particularly in patients with diabetes or insulin resistance. Although preliminary data suggest that these agents may offer benefits in fluid management and glycemic control, concerns regarding their safety profile in patients with cirrhosis, including potential renal and metabolic risks, warrant further investigation. By identifying key knowledge gaps and highlighting areas requiring additional research, this review aims to contribute to the development of more effective treatment strategies for this challenging and underserved patient population. Further studies, particularly randomized controlled trials, are necessary to confirm the therapeutic role of SGLT2i in cirrhotic ascites and refine treatment protocols.

Acknowledgments

Generative artificial intelligence (AI) tools, including ChatGPT (OpenAI), were not used in the writing, editing, or data generation of this manuscript. All content was created and reviewed by the authors.

Funding

The authors would like to acknowledge Graduate Medical Education (GME) for their Open Access Funds for this paper.

Data Availability

As this is a scoping review protocol, no datasets have been generated or analyzed yet. Upon completion of the review, extracted data will be made available in a publicly accessible repository (eg, Open Science Framework) and/or as supplementary materials accompanying the published manuscript. Any restrictions or limitations regarding data sharing will be clearly indicated at that time.

Authors' Contributions

Conceptualization: KB, TL, AK, AMK, AB, AS, UB
Methodology: TL, KB, RS, VG
Supervision: KB, AS
Writing: Original Draft: TL, AMK, AK, UB, AB, RS, VG, AS, KB
Writing: Review & Editing: KB, RS, VG

Conflicts of Interest

None declared.

Checklist 1
PRISMA-ScR checklist.
[DOCX File, 85 KB - resprot_v15i1e77587_app1.docx]

References

1. Sharma B, John S. Hepatic Cirrhosis: StatPearls Publishing; 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK482419/ [accessed 2025-11-27]
2. Baumgartner K, Cooper J, Smith A, St Louis J. Liver disease: cirrhosis. FP Essent 2021 Dec;511:36-43. [Medline: 34855340]
3. Ginès P, Krag A, Abraldes JG, Solà E, Fabrellas N, Kamath PS. Liver cirrhosis. Lancet 2021 Oct 9;398(10308):1359-1376. [doi: 10.1016/S0140-6736(21)01374-X] [Medline: 34543610]
4. Pedersen JS, Bendtsen F, Møller S. Management of cirrhotic ascites. Ther Adv Chronic Dis 2015 May;6(3):124-137. [doi: 10.1177/2040622315580069] [Medline: 25954497]
5. Liu YB, Chen MK. Epidemiology of liver cirrhosis and associated complications: current knowledge and future directions. World J Gastroenterol 2022 Nov 7;28(41):5910-5930. [doi: 10.3748/wjg.v28.i41.5910] [Medline: 36405106]
6. Wong F. Innovative approaches to the management of ascites in cirrhosis. JHEP Rep 2023 Jul;5(7):100749. [doi: 10.1016/j.jhepr.2023.100749] [Medline: 37250493]
7. Elkrief L, Rautou PE, Sarin S, Valla D, Paradis V, Moreau R. Diabetes mellitus in patients with cirrhosis: clinical implications and management. Liver Int 2016 Jul;36(7):936-948. [doi: 10.1111/liv.13115] [Medline: 26972930]
8. Moreau R, Delègue P, Pessione F, et al. Clinical characteristics and outcome of patients with cirrhosis and refractory ascites. Liver Int 2004 Oct;24(5):457-464. [doi: 10.1111/j.1478-3231.2004.0991.x] [Medline: 15482343]

XSL•FO
RenderX

9. Coman LI, Coman OA, Bădărău IA, Păunescu H, Ciocîrlan M. Association between liver cirrhosis and diabetes mellitus: a review on hepatic outcomes. J Clin Med 2021 Jan 12;10(2):262. [doi: 10.3390/jcm10020262] [Medline: 33445629]
10. Sapra A, Bhandari P. Diabetes: StatPearls; 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK551501/ [accessed 2025-11-27]
11. Garcia-Compean D, Jaquez-Quintana JO, Gonzalez-Gonzalez JA, Maldonado-Garza H. Liver cirrhosis and diabetes: risk factors, pathophysiology, clinical implications and management. World J Gastroenterol 2009 Jan 21;15(3):280-288. [doi: 10.3748/wjg.15.280] [Medline: 19140227]
12. Yen FS, Hsu CC, Wei JCC, Hou MC, Hwu CM. Selection and warning of evidence-based antidiabetic medications for patients with chronic liver disease. Front Med (Lausanne) 2022;9:839456. [doi: 10.3389/fmed.2022.839456] [Medline: 35252271]
13. Yen FS, Huang YH, Hou MC, et al. Metformin use and cirrhotic decompensation in patients with type 2 diabetes and liver cirrhosis. Br J Clin Pharmacol 2022 Jan;88(1):311-322. [doi: 10.1111/bcp.14970] [Medline: 34198358]
14. Kaplan DE, Serper M, John BV, et al. Effects of metformin exposure on survival in a large national cohort of patients with diabetes and cirrhosis. Clin Gastroenterol Hepatol 2021 Oct;19(10):2148-2160. [doi: 10.1016/j.cgh.2020.08.026] [Medline: 32798709]
15. Fonseca-Correa JI, Correa-Rotter R. Sodium-glucose cotransporter 2 inhibitors mechanisms of action: a review. Front Med (Lausanne) 2021;8:777861. [doi: 10.3389/fmed.2021.777861] [Medline: 34988095]
16. Wright EM. SGLT2 inhibitors: physiology and pharmacology. Kidney360 2021 Dec 30;2(12):2027-2037. [doi: 10.34067/KID.0002772021] [Medline: 35419546]
17. Plosker GL. Canagliflozin: a review of its use in patients with type 2 diabetes mellitus. Drugs (Abingdon Engl) 2014 May;74(7):807-824. [doi: 10.1007/s40265-014-0225-5] [Medline: 24831734]
18. Miyamoto Y, Honda A, Yokose S, Nagata M, Miyamoto J. The effects of SGLT2 inhibitors on liver cirrhosis patients with refractory ascites: a literature review. J Clin Med 2023 Mar 14;12(6):2253. [doi: 10.3390/jcm12062253] [Medline: 36983252]
19. Padda IS, Mahtani AU, Parmar M. Sodium-Glucose Transport Protein 2 (SGLT2) Inhibitors: StatPearls Publishing; 2023. URL: https://www.ncbi.nlm.nih.gov/books/NBK576405/ [accessed 2025-11-27]
20. Katsiki N, Mikhailidis DP, Theodorakis MJ. Sodium-glucose cotransporter 2 inhibitors (SGLT2i): their role in cardiometabolic risk management. Curr Pharm Des 2017;23(10):1522-1532. [doi: 10.2174/1381612823666170113152742] [Medline: 28088910]
21. Zhao DM, Li CQ, Sun YM, et al. Sodium glucose cotransporter-2-inhibitor dapagliflozin improves nonalcoholic fatty liver disease by ameliorating dipeptidyl-peptidase-4 protein expression in diabetic mice. Endokrynol Pol 2023;74(2):190-196. [doi: 10.5603/EP.a2023.0018] [Medline: 37039492]
22. Euh W, Lim S, Kim JW. Sodium-glucose cotransporter-2 inhibitors ameliorate liver enzyme abnormalities in Korean patients with type 2 diabetes mellitus and nonalcoholic fatty liver disease. Front Endocrinol (Lausanne) 2021;12:613389. [doi: 10.3389/fendo.2021.613389] [Medline: 34177796]
23. D'Andrea E, Wexler DJ, Kim SC, Paik JM, Alt E, Patorno E. Comparing effectiveness and safety of SGLT2 inhibitors vs DPP-4 inhibitors in patients with type 2 diabetes and varying baseline HbA1c levels. JAMA Intern Med 2023 Mar 1;183(3):242-254. [doi: 10.1001/jamainternmed.2022.6664] [Medline: 36745425]
24. Khouri C, Cracowski JL, Roustit M. SGLT-2 inhibitors and the risk of lower-limb amputation: Is this a class effect? Diabetes Obes Metab 2018 Jun;20(6):1531-1534. [doi: 10.1111/dom.13255] [Medline: 29430814]
25. Chen G, Li X, Cui Q, et al. Acute kidney injury following SGLT2 inhibitors among diabetic patients: a pharmacovigilance study. Int Urol Nephrol 2022 Nov;54(11):2949-2957. [doi: 10.1007/s11255-022-03211-7] [Medline: 35579781]
26. Horii T, Oikawa Y, Kunisada N, Shimada A, Atsuda K. Real-world risk of hypoglycemia-related hospitalization in Japanese patients with type 2 diabetes using SGLT2 inhibitors: a nationwide cohort study. BMJ Open Diab Res Care 2020 Nov;8(2):e001856. [doi: 10.1136/bmjdrc-2020-001856]
27. McGill JB, Subramanian S. Safety of sodium-glucose co-transporter 2 inhibitors. Am J Cardiol 2019 Dec 15;124 Suppl 1:S45-S52. [doi: 10.1016/j.amjcard.2019.10.029] [Medline: 31741440]
28. Role of sodium-glucose cotransporter-2 inhibitors in refractory cirrhotic ascites in diabetics – a scoping review. OSF. URL: https://osf.io/7eycr/overview [accessed 2025-11-27]
29. Peters MDJ, Godfrey CM, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews (2020 version). In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: JBI; 2020. [doi: 10.46658/JBIRM-20-01]
30. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst Rev 2021 Dec;10(1):89. [doi: 10.1186/s13643-021-01626-4]
31. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
32. Siafarikas C, Kapelios CJ, Papatheodoridi M, Vlachogiannakos J, Tentolouris N, Papatheodoridis G. Sodium-glucose linked transporter 2 inhibitors in liver cirrhosis: beyond their antidiabetic use. Liver Int 2024 Apr;44(4):884-893. [doi: 10.1111/liv.15851] [Medline: 38293770]

XSL•FO
RenderX

33. Gao FM, Ali AS, Bellomo R, et al. A systematic review and meta-analysis on the safety and efficacy of sodium-glucose cotransporter 2 inhibitor use in hospitalized patients. Diabetes Care 2024 Dec 1;47(12):2275-2290. [doi: 10.2337/dc24-0946] [Medline: 39602586]
34. Stachteas P, Nasoufidou A, Patoulias D, et al. The role of sodium-glucose co-transporter-2 inhibitors on diuretic resistance in heart failure. Int J Mol Sci 2024 Mar 8;25(6):3122. [doi: 10.3390/ijms25063122] [Medline: 38542096]

Abbreviations

OSF: Open Science Framework
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-analyses-Extension for Scoping Reviews
SGLT2: Sodium-glucose cotransporter-2
SGLT2i: Sodium-glucose cotransporter-2 inhibitors
T2DM: type 2 diabetes mellitus

Edited by J Sarvestan; submitted 16.May.2025; peer-reviewed by M Bhattarai, P Karakasis; revised version received 13.Oct.2025; accepted 15.Oct.2025; published 13.Jan.2026.

Please cite as:
Laeeq T, Kamal AM, Kasi ABU, Shabbar Banatwala UES, Baloch A, Sagaribay R, Goel V, Singh A, Batra K
Efficacy and Safety of Sodium-Glucose Cotransporter-2 Inhibitors in Patients With Refractory Cirrhotic Ascites and Diabetes: Scoping Review Protocol
JMIR Res Protoc 2026;15:e77587
URL: https://www.researchprotocols.org/2026/1/e77587
doi:10.2196/77587

© Tooba Laeeq, Ayesha Mehak Kamal, Arbab Burhan Uddin Kasi, Umm E Salma Shabbar Banatwala, Aniqa Baloch, Roberto Sagaribay, Vidhani Goel, Aditi Singh, Kavita Batra. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Clinical Evidence Profile of Oral Chinese Patent Ethnomedicines: Protocol for a Scoping Review and Evidence Map

Miaomiao Li[1,2], MD; Hui Luo[3], Dr med; Hui Zhao[4], Dr med; Xing Liao[2], Dr med

[1]Medical College of Acu-Moxi and Rehabilitation, Guangzhou University of Chinese Medicine, Guangdong, China

[2]Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, No.16, South Xiao Street, Dongzhimen Nei, Dongcheng District, Beijing, China

[3]China Tibetology Research Center, Beijing, China

[4]China Traditional Chinese Medicine Evidence-based Medical Center, China Academy of Chinese Medical Sciences, 16 South Xiao Street, Dongzhimen Nei, Dongcheng District, Beijing, China, China

Corresponding Author:
Xing Liao, Dr med
Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, No.16, South Xiao Street, Dongzhimen Nei, Dongcheng District, Beijing, China

Abstract

Background: Chinese patent ethnomedicines (CPEs), a form of traditional Chinese patent medicine, originate from the traditional medicines of ethnic minority groups and are widely used in clinical practice. However, existing evidence to support their application remains unclear. Therefore, to address this gap, this comprehensive scoping review will be performed to provide an overview of the available evidence on CPE preparations.

Objective: This review aims to provide the evidence profile for oral CPEs. This study will elucidate the current state of the evidence with respect to these medicines and identify research gaps. The detailed steps for conducting this review are outlined in this protocol. This review will contribute to a better understanding of CPEs.

Methods: This review will include clinical studies of CPEs irrespective of study design. The frameworks described by Arksey and O'Malley, Levac et al, and the Joanna Briggs Institute will be used to guide this scoping review. This review will involve six steps: (1) identifying the research question; (2) collecting information about CPEs from national related drug catalogs; (3) searching MEDLINE (via PubMed), Embase, Web of Science, Cochrane Library, and Chinese databases from inception to February 2025 to identify relevant publications; (4) screening the literature against the eligibility criteria; (5) extracting data using a predefined standardized data extraction form; and (6) summarizing, discussing, analyzing, and reporting the results. We will also present the results via data visualization techniques.

Results: We will synthesize data on CPEs by conducting the scoping review, drawing the evidence maps, identifying the clinical characteristics of CPEs and how they relate to the adverse events, and highlighting the limitations and gaps in the literature. We expect to publish the results in 2027.

Conclusions: The data obtained through this review could inform future research involving CPEs.

Trial Registration: Open Science Framework 10.17605/OSF.IO/E763B; https://osf.io/e763b

International Registered Report Identifier (IRRID): DERR1-10.2196/77741

(JMIR Res Protoc 2026;15:e77741) doi:10.2196/77741

KEYWORDS

Chinese patent medicine; Chinese patent ethnomedicine; complementary and alternative medicine; clinical evidence; scoping review protocol; Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PRISMA

Introduction

Chinese patent ethnomedicines (CPEs), which use indigenous medicines as raw materials, are produced based on ethnomedicine theory and the practical experience of various ethnic minority groups [1]. CPEs are a type of traditional medicine and embody the outstanding wisdom of indigenous people with respect to the use of medicinal plants, animals, and minerals. CPEs contribute significantly to the health care practices of many indigenous communities and are increasingly being used within certain communities in China [2].

China is a multiracial country with 56 nationalities and 56 different cultures, languages, and indigenous medicines due to differences in geographical and climatic conditions [3,4]; 55 of these nationalities are officially recognized as ethnic minority groups across 18 provinces of China [5]. Of China's 55 ethnic

XSL•FO
RenderX

minority groups, 47 possess either an ethnic medicine system or medical knowledge [6]. Among the indigenous medical systems, traditional Chinese medicine (TCM), also known as Han medicine, was developed by the Han people and has held an important and dominant position throughout the entire history of China's medical development. In official documents, TCM includes both Han medicine and ethnomedicine [7]. Ethnomedicine mainly includes Tibetan traditional medicine, Mongolian traditional medicine, Uyghur traditional medicine, Miao traditional medicine, Dai traditional medicine, Yi traditional medicine, Zhuang traditional medicine, and other minority traditional medicines [8].

In recent years, the importance of ethnomedicine has increased in health care practice; in particular, the application of CPEs has become more widespread. The *Chinese Traditional Medicine Statistics*, which was published in 2015 by the National Administration of Traditional Chinese Medicine of China, and the *Investigation and Analysis of Quality Standards of Ethnomedicines in Nine Provinces of China*, which was published in 2015 by the National Medical Products Administration of China, listed 4317 CPEs produced by 161 pharmaceutical enterprises [5]. A total of 93 CPEs were included in the latest *National Reimbursement Drug List* (2021 edition). In the *Pharmacopoeia of the People's Republic of China* (2020 edition) [9], there were 1607 kinds of Chinese patent medicines (CPMs), including 39 CPEs. These CPEs have been used to treat a range of conditions, including cardiovascular and cerebrovascular diseases, hepatobiliary system diseases, and nervous system diseases [6-9]. Furthermore, these CPEs can be administered in multiple forms, such as pills, tablets, capsules, liquids, and injections. Tibetan, Mongolian, and Uyghur traditional medicines are representative of ethnomedicines practiced among ethnic minority groups. For example, Tibetan traditional medicine is beneficial for treating digestive system disorders, rheumatic conditions, and altitude-related illnesses. Popular Tibetan patent medicines, such as the 25-flavor pearl pill (二十五味珍珠丸), are used to treat cardiovascular and cerebrovascular diseases, whereas the *ruyi zhenbao* pill (如意珍宝丸) is known to treat bone and joint diseases [10]. Mongolian patent medicines are known for their significant effects on diabetes, blood diseases, senile conditions, and concussion-induced disorders, among other diseases. For example, *wuwei shaji* powder (五味沙棘散) and *qiwei putao* powder (七味葡萄散) are used to treat chronic bronchitis and asthma in older populations. Uyghur patent medicines offer therapeutic advantages in managing skin diseases, rheumatism, and urogenital and digestive system ailments. For example, *baixuanxiatare* capsules (百癣夏塔热胶囊) are recognized for their ability to address skin diseases, including tinea handi, tinea corporis, tinea pedis, tinea versicolor, psoriasis, atopic dermatitis, herpes zoster, and acne [6].

Traditional CPMs derived from Han medicine have been well developed and widely studied because of the dominance of the Han people and their cultural cultivation and historical heritage; thus, there is an increasing amount of high-quality evidence supporting the efficacy of CPMs [11,12]. However, there is a relative lack of clinical evidence on CPEs, and there have been no systematic explorations of these medicines. To address this gap, this scoping review will identify, describe, and map the evidence profile of the application of CPEs alone or in routine biomedical practice. Previous studies have synthesized information about the statistics and analysis of CPEs listed in the *Pharmacopoeia of the People's Republic of China* [6,9]. Building on these studies, this review aims to identify the CPEs listed in 3 national drug catalogs, systematically search for related clinical studies, and provide a visual analysis of the results to showcase the contemporary research evidence profile of CPEs. The findings of this review will highlight the available clinical research evidence on CPEs and standardized reporting frameworks, thereby informing future research efforts (Textbox 1).

Textbox 1. Strengths and limitations of this study.

- This review will be the first to show the clinical evidence profile of the Chinese patent ethnomedicines (CPEs) commonly used in China, including the current status of clinical research evidence, the relationship between diseases and CPEs, the clinical research collaboration network, and the clinical characteristics of the CPEs related to adverse events.
- This review will be conducted based on the framework proposed by Arksey and O'Malley [13] and Levac et al [14]. Additionally, this review will be conducted in accordance with the Joanna Briggs Institute scoping review guidelines [15]. Furthermore, this review will be reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews).
- The review will only include CPEs listed in the China national authoritative documents and will not include other medicines reported elsewhere.

Methods

Study Design and Registration

The framework described by Arksey and O'Malley [13] and Levac et al [14] and the Joanna Briggs Institute scoping review methodology will be followed and adapted during this scoping review. This review will involve six steps: (1) identifying the research question; (2) collecting information about CPEs; (3) identifying relevant studies; (4) selecting studies; (5) extracting the data; and (6) summarizing, discussing, analyzing, and reporting the results. This scoping review will be reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines [16]. Additionally, this protocol was reported in accordance with the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) guidelines [17]. This review protocol was registered on the Open Science Framework platform.

Stage 1: Identifying the Research Question

This scoping review aims to provide an overview of the available evidence profile of CPEs and identify research gaps by focusing on clinical studies. Suggestions and revisions of

XSL•FO
RenderX

the research question were requested from other related experimental or clinical research experts. The following questions will be addressed in our scoping review following consensus among the research team:

1. What is the current status of clinical research evidence?
2. What is the relationship between diseases and CPEs?
3. What does the clinical research collaboration network look like?
4. What are the clinical characteristics of CPEs related to adverse event (AE) features?

Stage 2: Selecting CPEs

We selected CPEs from the *Pharmacopoeia of the People's Republic of China* (2020 edition), *National Reimbursement Drug List* (2021 edition), and *National Essential Medicines List* (2018 edition). A total of 146 drugs were retrieved. After removing nonoral and repeated medicines, 105 CPEs remained, including Tibetan medicines (n=46, 43.8%), Mongolian medicines (n=35, 33.3%), Tibetan and Mongolian medicines (n=1, 1%), Uyghur medicines (n=18, 17.1%), Miao medicines (n=1, 1%), Yi medicines (n=1, 1%), Dai medicines (n=2, 1.9%), and Jingpo medicines (n=1, 1%). A total of 20 CPEs were included in any 2 of the aforementioned national drug catalogs. In particular, Liuwei Anxiao powder or capsules were included in the 3 national drug catalogs. Textbox 2 and Table S1 in Multimedia Appendix 1 provide basic information on the 105 CPEs included.

XSL•FO
RenderX

Textbox 2. Information sources of the 105 included Chinese patent ethnomedicines [18-20]. Parentheses indicate that the medicine is also available in another form.

Tibetan medicines

- Ershiwuwei songshi pill
- Ershiwuwei shanhu pill
- Ershiwuwei zhenzhu pill
- Shiyiwei nengxiao pill (capsule)
- Shi’erwei yishou powder
- Shisanwei bangga powder
- Shiwuwei chenxiang pill
- 70-flavor zhenzhu pill
- Qiwei tiexie pill
- Bawei chenxiang pill (capsule)
- Bawei chenxiang powder
- Jiuwei shihuihua powder
- Wuwei shexiang pill
- Renqing mangjue
- Renqing changjue
- Jiebai pill
- Cuitang pill
- Duyiwei capsule (pill)
- Changsong bawei chenxiang powder
- Dayuejing pill
- Ershiwei chenxiang pill
- Ershiwei roudoukou pill
- Ershiwuwei datang pill
- Ershiwuwei ercha pill
- Ershiwuwei lvxue pill
- Ershiwuwei zhenzhu pill
- Jiuwei niuhuang pill
- Lishukang capsule
- Liugan pill
- Liuwei nengxiao pill (capsule)
- Nuodikang pill (capsule, granule, or oral solution)
- Pazhu capsule
- Pazhu pill
- Qiwei honghuashusheng powder (pill)
- Ruyi zhenbao pill
- 35-flavor chenxiang pill
- Shanhuqishiwei pill
- Shiwei dida capsule
- Shiweiheibingpian pill
- Shiwei longdanhua granule (capsule)
- Shiwuwei heiyao pill

XSL•FO
RenderX

- Shiwuweilongdanhua pill
- Pomegranate jianwei pill (capsule)
- Pomegranate jianwei powder
- Zhituojiebai pill
- Zuozhu Daxi
- Anshen pill
- Liuwei mingmu pill
- An’erning granule
- Honghua ruyi pill

Mongolian medicines

- Liangxue shiwei powder (pill)
- Sanzi powder
- Sanwei tribulus powder
- Wuwei shaji powder
- Wuwei qingzhuo powder
- Liuwei muxiang powder
- Siwei tumuxiang powder
- Alatan wuwei pill
- Anshen buxin liuwei pill
- Bateri qiwei pill
- Dahuang sanwei pill
- Fengshi ershiwuwei pill
- Hanshuishi ershiyiwei powder
- Honghuaqinggan shisanwei pill
- Huangbaibawei pill
- Jixiang ankun pill
- Naru sanwei pill
- Nuangong qiwei pill (powder)
- Qinggan jiuwei pill
- Qingre bawei pill (powder and capsule)
- Roukou wuwei pill
- Surilao qingfei zhike capsule
- Diaoyuandabu ershiwuwei tang powder
- Wulan shisanweitang powder
- Xiaoji jiebai pill
- Yishen shiqiwei pill
- Xiaoer shikou powder
- Zhachong shisanwei pill
- Zhenbao pill
- Zhenzhu tongluo pill
- Shiliuwei dongqing pill
- Qiwei guangzao pill
- Qiweiputao powder

XSL•FO
RenderX

- Qingxinchenxiang bawei powder or bawei qingxinchenxiang powder
- Bawei tanxiang powder

Tibetan and Mongolian medicines

- Liuwei anxiao powder
- Liuwei anxiao capsule
- Aiweixin oral solution
- Baixuanxiatare pill (capsule)
- Compound gaoziban pill
- Compound muniziqi granule
- Hanchuan zupa granule
- Zukamu granule
- Hugan buzure granule
- Jianxin hemilgaozi ban'an pirated tablets
- Luobufukebiri pill

Uyghur medicines

- Mamulan antidiarrheal capsule
- Meiguihua oral solution
- Niaotongkakenaiqi pill
- Qingre carson granule
- Pomegranate blood-tonifying syrup
- Tongzhisurunjiang capsule (pill)
- Yanxiao dinar syrup
- Yangxindawayimixikemigao
- Yixin badiranjibuya granule
- Yajiaohadun powder

Dai medicines

- Dengzhan shengmai capsule

Jingpo medicines

- Hufengjiu

Miao medicines

- Yindan xinnaotong capsule

Yi medicines

- Lingdancao granule

Stage 3: Identifying Relevant Studies

Search Strategy

We will use a 3-step systematic and comprehensive search strategy [21]. The first step will involve an initial limited search of the MEDLINE database using medicine names, including both brand and generic names, followed by an analysis of the text words contained in the titles and abstracts of the retrieved articles, as well as the index terms used to describe the articles. A second search will then be performed using the identified keywords and index terms across all the included databases. Third, the reference lists of the identified articles and reports will be searched to identify additional eligible studies. Google Scholar will be used to identify additional studies. If necessary, we will contact the authors directly to obtain further information. The search process is expected to be iterative, incorporating additional search terms and sources as the search progresses. The search strategy for the MEDLINE database is provided in Table S2 in Multimedia Appendix 1.

XSL•FO
RenderX

Information Source

We will systematically search PubMed, Embase (Ovid), Cochrane Library, Web of Science, China National Knowledge Infrastructure, Wanfang Data, Chongqing VIP Chinese Scientific Journal Database, and Chinese Biomedical Literature Database from inception to 2025. To ensure the feasibility and accuracy of the retrieval strategy, an information specialist at the China Academy of Chinese Medical Sciences will be consulted. The reference lists of the included articles and related systematic reviews will also be examined to identify additional eligible studies. We will include studies that evaluate eligible interventions and report predefined outcomes of interest, with no restrictions on publication language.

Eligibility Criteria

Concepts

In this scoping review, ethnomedicine is defined as the "traditional medicine" of minority communities whose knowledge and practices have been transmitted through both oral and written traditions and have evolved over the thousands of years of human existence [22,23]. This review will focus on the traditional healing practices of minority communities in China. "Biomedical" refers to modern medicine, also known as "Western medicine" or "allopathic medicine." "Traditional and complementary medicine" is an umbrella term that is often applied to any health care practice outside the biomedical mainstream, especially in industrialized countries, including acupuncture, chiropractic, and herbal medicine [24].

Interventions

The eligible CPEs included in the target studies are listed in the *Pharmacopoeia of the People's Republic of China* (2020 edition), *National Reimbursement Drug List* (2024 edition), or *National Essential Drug List* (2018 edition). We will include CPEs of any type, such as liquids, pills, gels, tablets, and extracts. We will also include studies with treatment groups combined with comparators.

Comparators

The eligible comparators in the target studies include placebo, no intervention, conventional Western medicine treatments, or routine care. Studies comparing different CPEs will be excluded.

Outcomes

The outcomes of the target studies will be categorized as follows: symptom and sign alleviation rate, physicochemical outcomes, quality of life, and other outcomes (TCM syndrome score or effective rate). We will also report safety in terms of the occurrence and types of AEs.

Types of Target Studies

Primary research (which involves the collection of original data for a research project), secondary research (which involves the summary or synthesis of data and literature that have been organized and published by others), and case reports on patient and physician experiences (which involve the opinions and qualitative textual data [eg, interview transcripts or open-ended responses] from patients and physicians) will be included. Other gray literature and some articles that were not published or peer reviewed will also be included if the data are complete. Short or mini reviews, narrative reviews, abstracts from conferences, letters, short communications, and comments will be excluded.

Stage 4: Screening the Studies

All identified references will be collated and uploaded into NoteExpress (version 3.4.0.8878), and duplicate records will be removed. The study selection process will be piloted independently by a reviewer using 20 studies to assess the interpretation of the eligibility criteria and the consistency of use. Following the pilot test, the eligibility criteria will be refined if necessary. During the formal process, the literature will be evaluated against the eligibility criteria; 2 trained reviewers independently and in duplicate will screen the titles and abstracts of the retrieved citations and review the full texts of studies considered potentially eligible. Disagreements will be resolved through discussion or, if necessary, adjudication from clinical experts who will be blinded to the trial results. When clarification of eligibility criteria is needed, we will contact the study authors and exclude studies by authors who do not respond. All the records that do not meet the inclusion criteria will be excluded, and the reasons for exclusion will be provided in an appendix. Multiple articles and reports from the same program will be treated as one with respect to the data extraction process and the presentation of research results. The details of the study selection process are shown in Figure 1.

XSL•FO
RenderX

Figure 1. Flowchart for the selection of studies related to Chinese patent ethnomedicines. CNKI: China National Knowledge Infrastructure; SinoMed: China Biomedical Literature Service System; VIP: VIP Database for Chinese Technical Periodicals.

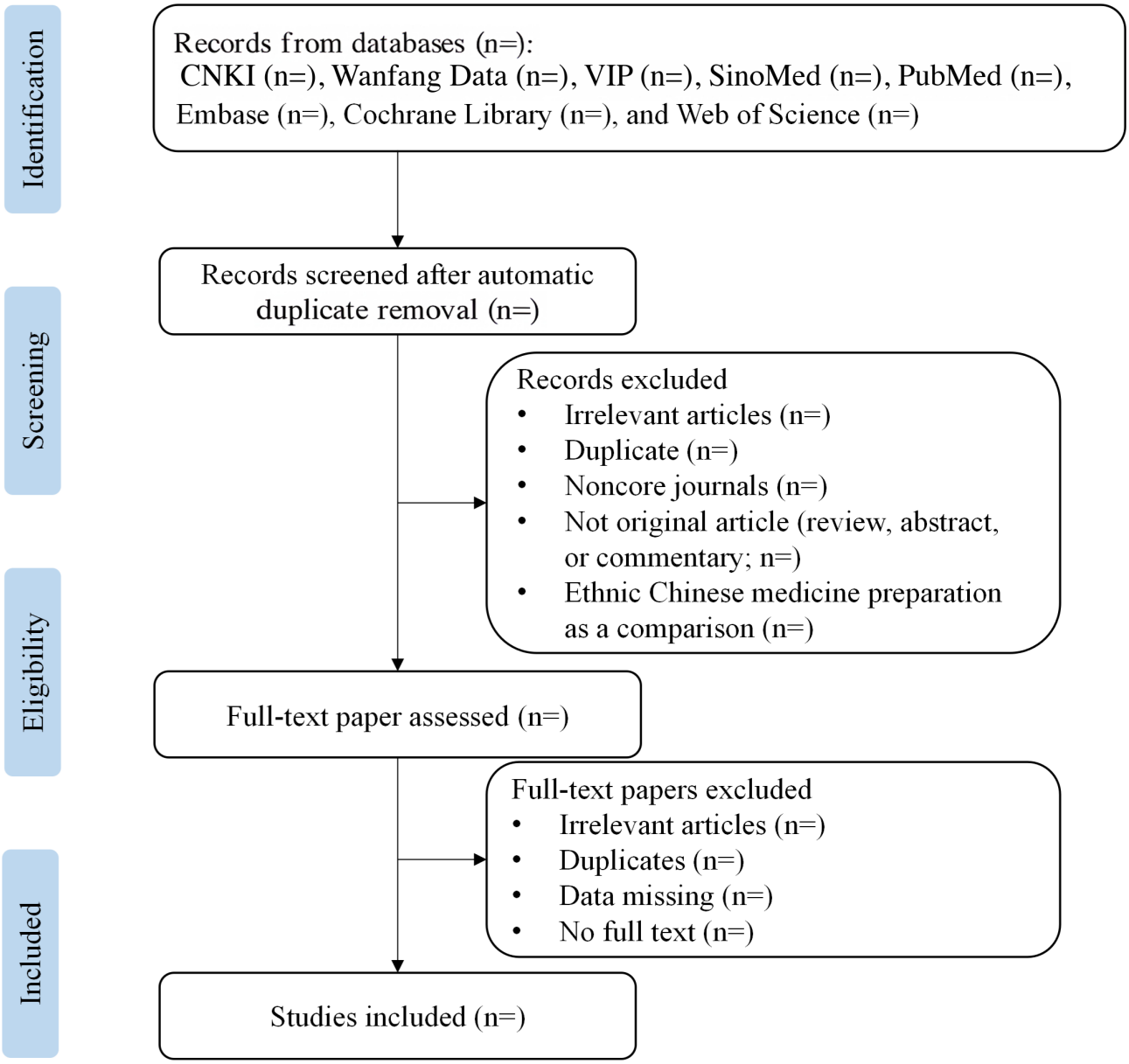

Stage 5: Charting the Data CPE

Data Extraction

The following data about CPEs will be extracted: market prices, formulation type, use, dosage, treatment period, and components. Additionally, treatment symptoms, signs, and diseases will be extracted into a previously developed data extraction table. The Yaozhi website will be used for data extraction of market prices, which will be performed by the first author. The senior investigator will be consulted to address any uncertainties in the data extraction process. Tables S3 and S4 in Multimedia Appendix 1 provide the detailed extraction information of the included CPEs.

Study Data Extraction

On the basis of the pilot-testing form, the final version of the data extraction form will be developed. The following information will be extracted from the included studies: article title, first author, publication year, sample size, participant characteristics (such as sex and age), intervention, comparator, outcomes, and study design. To facilitate the data extraction process, 2 draft charts have been prepared: one to record the characteristics of CPEs and the other to record data relevant to clinical research on CPEs, as shown in Tables S5 and S6 in Multimedia Appendix 1. The completed diagrams will be included in an appendix to the final scoping review report.

Stage 6: Collating, Analyzing, and Reporting the Results

Essential Information Analysis of the Literature

Descriptive and visual analyses (eg, tables or charts, line graphs, bar charts, and bubble charts) of publication trends, CPE distribution in the literature, sample size, intervention measures, and outcomes in clinical studies will be performed. Microsoft Excel 2019 will be used to manage the data, analyze the publication trends, and examine the distribution of the CPEs.

XSL•FO
RenderX

R (R Foundation for Statistical Computing) will be used to analyze outcomes. Cytoscape (Cytoscape Consortium) will be used to analyze and map the relationships between CPEs and diseases identified in the included studies. This will facilitate the analysis of which drugs are frequently used to treat specific disease combinations, thereby revealing potential treatment patterns. Quality assessment is not an integral part of scoping reviews as the aims of these reviews are to clarify concepts, identify gaps in the available evidence, and provide an overview of the literature rather than evaluate the quality of the individual studies [25].

Bibliometric Analysis of the Literature

The selected references will be imported into CiteSpace (version 6.3.R1) to transform the data structure. Subsequently, analyses will be performed, including coauthorship and institutional collaboration networks derived from the references, co-occurrence networks, and keyword citation bursts. The time span will encompass the period from 1994 to 2025, using a 2-year time slice and a network refinement approach that involves a pathfinder network, pruning of sliced networks, and pruning of the merged network.

Clinical Characteristic Identification Related to AEs

This study will apply supervised machine learning methods to identify clinical characteristics associated with AEs reported in randomized controlled trials. This will be an exploratory part of the study based on 5 to 10 published high-quality randomized controlled trials aiming to identify potential patterns rather than produce definitive predictive models. We will first prespecify the AE outcome (occurrence of any serious AE within 30 days of randomization) and harmonize variables across trials using a variable-mapping dictionary. Collected data will include baseline demographics (age and sex), disease severity indicators, treatment assignment, concomitant medications (name, dose, and timing), and trial or center identifiers. Missing data will be assessed and handled using multiple imputation performed separately within training folds to avoid information leakage. Next, through feature engineering, features related to AEs will be selected, and new features will be constructed to capture potential nonlinear relationships. A supervised learning model (random forest or logistic regression) will then be chosen, model performance will be evaluated using 10-fold cross-validation, and hyperparameters will be optimized to improve prediction accuracy. The data will be split into a 70% training set and a 30% test set; the test set will be used to validate the model's generalization ability and analyze feature importance to identify the clinical features that contribute most to the prediction of AEs.

Before analysis, all trial datasets will undergo a standardized data quality control. We will compile a detailed data dictionary documenting variable names, units, coding schemes, and provenance (original case report form field, derived variable definition, and trial identifier). The quality assessment steps will include range and logic checks, detection and resolution of duplicate records, verification of date and temporal consistency (eg, treatment start and stop dates vs AE onset), and cross-checking of AE coding against source documents where available. The variables with excessive missingness (prespecified threshold, eg, >50%) will be considered for exclusion or treated in sensitivity analyses. All data cleaning steps and data transformations will be logged in an audit trail to ensure reproducibility. Regarding data availability, we will respect original trial consent and data sharing agreements: only deidentified datasets approved for secondary analysis will be included.

Results

We will synthesize data on CPEs by conducting the scoping review, drawing the evidence maps, identifying the clinical characteristics related to AEs, and highlighting the limitations and gaps in the literature. CPE selection was completed in 2024. Data collection for clinical studies is expected to begin in December 2025 and finish in December 2026. As of December 2025, data analysis has not yet started, and results are expected to be published in December 2027.

Discussion

Anticipated Findings

Current ethnomedicine studies have focused on identifying classic, well-known prescriptions in ancient indigenous medicine texts and conducting textual research on them through the lens of classic literature rather than clinical research [26,27]. Therefore, research that focuses on analyzing the clinical characteristics of CPEs should be considered. To our knowledge, this scoping review is the first to systematically identify indigenous medicines endorsed by national authoritative drug catalogs and provide an overview of contemporary clinical research evidence. We will analyze the characteristics of previous clinical studies; explore the relationship between indigenous medicines and disorders based on contemporary research; and, ultimately, provide valuable insights and inspiration for future studies concerning CPEs.

The Current Status of CPEs

As an integral aspect of TCM, ethnomedicine continues to be highly valuable in health care and has emerged as a complementary system to contemporary medical practices. Over time, through enduring application and refinement, it has led to the development of a variety of unique CPEs, which have played important roles in the prevention and treatment of chronic conditions, challenging illnesses, and endemic diseases in regions inhabited by ethnic minority groups. Many ethnic minority groups rely on CPEs as an adjuvant or alternative to conventional medical treatments.

However, ethnic medicine is not a mainstream medicine similar to Han Chinese medicine. The minority language, obscure medical theory system, unique processing methods, and wide regional differences have become barriers to the promotion and application of ethnic medicine. In particular, many CPEs contain toxic ingredients, such as *zuotai* (佐太), nux vomica (马钱子), aconite (乌头), aristolochic acid (马兜铃酸), and iron powder (铁粉). Among these, *zuotai* (佐太) is a preparation formed by the special processing of mercury by Tibetans. As an active ingredient, it has been used to treat cardiovascular disorders, liver disorders, and gastrointestinal disorders combined with

XSL•FO
RenderX

Tibetan herbal medicine for 1300 years [28,29]. However, relevant studies have shown that *zuotai* can cause damage to the liver and kidney tissues of rats, especially long-term toxic effects on the spleen [30], which means that a greater number of medicines containing *zuotai* cannot be accepted and applied by a wider range of clinicians outside of minority groups. The same is true for other medicines containing toxic ingredients. Therefore, based on a scoping review of the existing literature, this study will conduct an in-depth analysis of the clinical characteristics of AEs using the collected clinical research evidence base. By identifying key clinical features associated with AEs, this study aims to enhance the safety profile and optimize the clinical treatment outcomes of CPEs.

At present, the exact mechanisms underlying the effects of most CPEs and their disease prevention effects have yet to be fully elucidated. Nonetheless, public enthusiasm for these treatments persists [8]. The natural source of these drugs makes them readily acceptable to the general population [31]. However, few clinical studies have examined the use of CPEs, and the existing evidence supporting their efficacy is inconclusive. Moreover, highly efficient and safe prescriptions have not been thoroughly explored, disseminated, or used. The knowledge regarding the use of ethnomedicine in clinical use and research of CPEs within modern medical practice is also unclear among most medical professionals [32,33]. These factors make it difficult to assess and recognize the potential value of CPEs or explore new avenues to enhance clinical practice. Consequently, there is an imperative need for a comprehensive review and analysis of relevant clinical studies. This effort is vital not only to protect and pass down the knowledge of traditional CPE use but also to provide guidance for ongoing research and development in the field of ethnomedicine.

Implications for Future Research

In the future, we will conduct targeted analyses (eg, systematic reviews or network meta-analyses) to scientifically evaluate the efficacy and safety of CPEs. Additionally, we will perform comparative analyses between similar medicinal products to determine the specific therapeutic advantages of CPEs. To further highlight the characteristics and advantages of CPEs and the role of ethnic Chinese medicine in the health care field, the following suggestions are proposed. The first is to cultivate researchers with the cross-subject background of TCM, ethnic medicine, and Chinese medical literature, among other related fields, to find the intersection of the 2 theories (TCM and Chinese ethnomedicine) from the perspective of historical traceability and interpret and transform the theory of ethnic medicine into the concept understood in Han Chinese medicine to deeply master the theory of ethnic medicine and promote the complementarity of the discipline. The second suggestion is to select the priority disorders that classic CPEs treat, identify the disease stage in which these therapies demonstrate comparative therapeutic advantages, and precisely establish complementary advantages in conjunction with Han Chinese medicine. Third, given that ethnic medicines contain many toxic components, with the patient's informed consent and full consideration of their wishes, pilot studies of advantageous ethnic medicines can be conducted for end-stage patients with critical diseases such as cancer or rare diseases that lack effective therapy in modern medicine. The fourth suggestion is to promote the localization of ethnic medicine production; compare the processing technology of ethnic drugs with that of modern processing methods; carry out the interdisciplinary integration of research methodologies across different medical fields; and, therefore, further understand the value of the processing method of ethnic medicines to increase their acceptability and dissemination.

Strengths and Limitations

Developing a protocol before conducting a scoping review is beneficial for several reasons [17]. First, protocols help predict potential challenges that may arise during the review. Second, protocols prevent arbitrary decisions related to the inclusion criteria and the data extraction process. Third, protocols mitigate the risk of selective reporting. Fourth, protocols minimize the likelihood of study duplication. For this study, we preliminarily identified some indigenous medicines and will retrieve relevant literature, screen the studies, and extract the data.

This protocol describes an exploratory research framework for secondary literature regarding CPEs in which the evidence will be evaluated via comparative analysis rather than quantitative synthesis. Furthermore, this review will only provide a clinical overview of the CPEs listed in Chinese national authoritative documents and will not include other medicines reported elsewhere.

Conclusions

This paper presents a comprehensive scoping review protocol aimed at evaluating the clinical evidence profile of oral CPEs. The review will systematically identify, describe, and map the evidence related to the application of CPEs in clinical practice with a focus on addressing the current lack of clear evidence and identifying research gaps. The findings will contribute to a better understanding of CPEs and provide valuable insights for future research and development in the field of ethnomedicine.

Acknowledgments

The authors used the generative AI tool ChatGPT (OpenAI) to assist with language polishing and initial drafting of the manuscript's background section. All AI-generated content was critically reviewed, substantially revised, and verified by the authors, who take full responsibility for the final work.

XSL•FO
RenderX

Funding

This study was supported by the National Natural Science Foundation of China (No. 82174239) and the Special Project for the Director of the Research Office of the China Center for Evidence-Based Traditional Chinese Medicine Q17 (No. 2020YJSZX-2).

Data Availability

Data sharing is not applicable to this article as no original datasets were generated or analyzed during the conduct of this scoping review protocol.

Authors' Contributions

The idea for this scoping review was put forward by XL. The refinement of the aims and objectives was carried out through consensus between XL and ML. The first draft of the manuscript was written by ML. XL, HL, and HZ provided modification suggestions for this protocol. XL supervised the conduct of the entire study process. HZ is the co-corresponding author and can be reached via email at huizh_519@126.com or phone at 86 135 2148 3754.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Basic information, search strategy, market prices, composition details, and characteristics tables of the 105 traditional Chinese ethnic patent medicines.
[DOCX File, 119 KB - resprot_v15i1e77741_app1.docx]

References

1. Zhu Z, Wang T, Fu D, Gui Y, Wang J, Cui T. Innovative development path of ethnomedicines: an overview of ethnomedicines in China. Front Med 2016 Jun;10(2):166-177. [doi: 10.1007/s11684-016-0448-9] [Medline: 27193007]
2. Li J, Zhang Y. Inheritance and innovation of ethnic medicine under the vision of the Chinese national community. World J Tradit Chin Med 2021;16(15):2211-2216. [doi: 10.3969/j.issn.1673-7202.2021.15.001]
3. Chen R, Duan ZY, Duan XH, Chen QH, Zheng J. Progress in research on gut microbiota in ethnic minorities in China and consideration of intervention strategies based on ethnic medicine: a review. Front Cell Infect Microbiol 2022 Oct 18;12:1027541. [doi: 10.3389/fcimb.2022.1027541] [Medline: 36329820]
4. Wanzala W, Minyoso SI. Ethnomedicines in the 21st century: challenges and opportunities in the contemporary world. J Med Herbs Ethnomed 2024;10:12-36. [doi: 10.25081/jmhe.2024.v10.8832]
5. Li Z, Li C, Zhang X, et al. Policies and problems of modernizing ethnomedicine in China: a focus on the Yi and Dai traditional medicines of Yunnan Province. Evid Based Complement Alternat Med 2020 Aug 14;2020:1023297. [doi: 10.1155/2020/1023297] [Medline: 32855645]
6. Zhao W, Kong L, Wang X, Liu Q, Wang Y, Wang J. Efficacy and safety of Chinese patent medicines for allergic rhinitis based on 2020 Chinese Pharmacopoeia: a protocol for systematic review and meta-analysis of randomized controlled trials. Syst Rev 2025 Mar 24;14(1):69. [doi: 10.1186/s13643-024-02748-1] [Medline: 40128915]
7. Transcript of the press conference on “several opinions on strengthening the work of ethnic minority medicine in the new era”. National Administration of Traditional Chinese Medicine. 2018. URL: http://www.natcm.gov.cn/bangongshi/gongzuodongtai/2018-08-23/7670.html [accessed 2023-07-16]
8. Pan SY, Litscher G, Gao SH, et al. Historical perspective of traditional indigenous medical practices: the current renaissance and conservation of herbal resources. Evid Based Complement Alternat Med 2014;2014:525340. [doi: 10.1155/2014/525340] [Medline: 24872833]
9. Yin XB, Qu CH, Dong XX, Shen MR, Ni J. Preparation regularity of Chinese patent medicine in Chinese Pharmacopoeia (2020 edition, Vol. Ⅰ). Zhongguo Zhong Yao Za Zhi 2022 Aug;47(16):4529-4535. [doi: 10.19540/j.cnki.cjcmm.20220419.601] [Medline: 36046882]
10. Lu S, Han F, Kong XW, Fan X, et al. Analysis of the reasonability of clinical use of folk medicine in the Third Affiliated Hospital of Beijing University of Chinese Medicine. Clin Med J 2017;15(1):21-24. [doi: 10.3969/j.issn.1672-3384.2017.01.006]
11. Liu S, Yao C, Xie J, et al. Effect of an herbal-based injection on 28-day mortality in patients with sepsis: the EXIT-SEP randomized clinical trial. JAMA Intern Med 2023 Jul 1;183(7):647-655. [doi: 10.1001/jamainternmed.2023.0780] [Medline: 37126332]
12. Zhang M, Liu Y, Xu M, et al. Carotid artery plaque intervention with Tongxinluo capsule (CAPITAL): a multicenter randomized double-blind parallel-group placebo-controlled study. Sci Rep 2019 Mar 14;9(1):4545. [doi: 10.1038/s41598-019-41118-z] [Medline: 30872737]

XSL•FO
RenderX

13. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb 23;8(1):19-32. [doi: 10.1080/1364557032000119616]
14. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69. [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
15. Peters MD, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Chapter 11: Scoping reviews (2020 version). In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis: Joanna Briggs Institute; 2020. [doi: 10.46658/JBIMES-20-12]
16. Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 2;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
17. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan 1;4(1):1. [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]
18. National Pharmacopoeia Commission. Pharmacopoeia of the People's Republic of China, 2020th edition: China Medical Science Press; 2020.
19. National Healthcare Security Administration. National Reimbursement Drug List. URL: http://www.nhsa.gov.cn/art/2019/8/20/art_37_1666.html [accessed 2025-12-25]
20. National essential drug directory. National Health Commission of the People's Republic of China. 2018. URL: https://www.nhc.gov.cn/wjw/jbywml/list.shtml [accessed 2018-10-25]
21. Blignault I, Hunter J, Mumford J. Integration of Indigenous healing practices with western biomedicine in Australia, Canada, New Zealand and the United States of America: a scoping review protocol. JBI Database System Rev Implement Rep 2018 Jun;16(6):1354-1360. [doi: 10.11124/JBISRIR-2017-003468] [Medline: 29894403]
22. Chattopadhyay D. Ethnomedicinal phytophores in disease management. Internat J Biomed Pharmaceut Sci 2009;3(Spl1):1-125.
23. Chattopadhyay D, editor. Ethnomedicine: A Source of Complementary Therapeutics: Research Signpost; 2010.
24. WHO traditional medicine strategy: 2014-2023. : World Health Organization; 2013 May 15. URL: https://www.who.int/publications/i/item/9789241506096 [accessed 2025-12-23]
25. Munn Z, Peters MD, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. BMC Med Res Methodol 2018 Nov 19;18(1):143. [doi: 10.1186/s12874-018-0611-x] [Medline: 30453902]
26. Ding K, Cheng L, Zha B, Bao WE, Zhang LJ, Xie YM. Methods and key points of literature collation and mining of classic prescriptions in ethnic medicine. Zhongguo Zhong Yao Za Zhi 2021 Oct;46(19):5123-5129. [doi: 10.19540/j.cnki.cjcmm.20210723.501] [Medline: 34738410]
27. Sun CQ, Li YY, Xie ZN, Xie YM. General screening principles and inheritance paths of classic prescriptions in Tibetan medicine. Zhongguo Zhong Yao Za Zhi 2021 Oct;46(19):5130-5136. [doi: 10.19540/j.cnki.cjcmm.20210702.501] [Medline: 34738411]
28. Kannan N, Nv AK, Balaji S. Chemical, elemental, morphological and toxicological characteristics of traditional Indian Siddha formulation: Kasthuri Karuppu. Adv Tradit Med 2025 Mar;25:221-231. [doi: 10.1007/s13596-024-00790-x]
29. Dou GG, Wei L, Zhou W, Zhang Y, et al. Process development and research status of Tibetan medicine zuotai (mercury preparation). World Sci Technol-Mod Tradit Chin Med 2024;26(1):229-235 [FREE Full text]
30. Qiao Y, Chen H, Guo J, et al. A study on the effects of metacinnabar (β-HgS) on weight and appetite recovery in stressed mice. J Ethnopharmacol 2024 Dec;335:118663. [doi: 10.1016/j.jep.2024.118663] [Medline: 39128797]
31. Gupta AA, Jadhav A, Bhola N, Agrawal P. Insight of ethnomedicines in dentistry: a brief review. Cureus 2022 Aug 18;14(8):e28148. [doi: 10.7759/cureus.28148] [Medline: 36148198]
32. Zhu Z, Fu D, Gui Y, et al. Innovative development path of ethnomedicines: the interpretation of the path. Front Med 2017 Mar;11(1):32-47. [doi: 10.1007/s11684-016-0495-2] [Medline: 28220457]
33. Ping N. Study on the development status and protection countermeasures of naxi ethnic medicine. : China National Knowledge Infrastructure; 2019 URL: https://link.oversea.cnki.net/doi/10.27200/d.cnki.gkmlu.2019.001264 [accessed 2025-05-29]

Abbreviations

AE: adverse event
CPE: Chinese patent ethnomedicine
CPM: Chinese patent medicine
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews
TCM: traditional Chinese medicine

XSL•FO
RenderX

Edited by J Sarvestan; submitted 19.May.2025; peer-reviewed by D Ajayi; revised version received 09.Dec.2025; accepted 10.Dec.2025; published 13.Jan.2026.

Please cite as:
Li M, Luo H, Zhao H, Liao X
Clinical Evidence Profile of Oral Chinese Patent Ethnomedicines: Protocol for a Scoping Review and Evidence Map
JMIR Res Protoc 2026;15:e77741
URL: https://www.researchprotocols.org/2026/1/e77741
doi:10.2196/77741

© Miaomiao Li, Hui Luo, Hui Zhao, Xing Liao. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Mapping Algorithmic Bias in AI-Powered Electrocardiogram Interpretation Across the AI Life Cycle: Protocol for a Scoping Review

Luqman Lawal[1,2], MD, MPH, MBA; Christopher Paton[1,3], BMBS, PhD, MBA; Mike English[1], MBChB; Bruno Holthof[1], MD, PhD, MBA; Tabitha Preston[1]

[1]Nuffield Department of Medicine, University of Oxford, Oxford, England, United Kingdom

[2]Mayo Clinic, Rochester, MN, United States

[3]Liggins Institute and School of Population Health, University of Auckland, Auckland, New Zealand

Corresponding Author:
Luqman Lawal, MD, MPH, MBA
Nuffield Department of Medicine
University of Oxford
Wellington Square
Oxford, England, OX1 2JD
United Kingdom
Phone: 1 6513075288
Email: luqman.lawal@ndm.ox.ac.uk

Abstract

Background: Artificial intelligence (AI)–powered analysis of electrocardiograms (ECGs) is reshaping cardiac diagnostics, offering faster and often more accurate detection of conditions such as arrhythmias and heart failure. However, growing evidence suggests that algorithmic bias, defined as performance disparities across patient subgroups, may undermine diagnostic equity. These biases can emerge at any stage of the AI life cycle, including data collection, model development, evaluation, deployment, and clinical use. If unaddressed, they risk exacerbating health disparities, particularly in underrepresented populations and low-resource settings. Early identification and mitigation of such bias are essential to ensuring diagnostic equity.

Objective: This scoping review protocol outlines a structured approach to mapping the evidence on algorithmic bias in AI-enabled ECG interpretation. Following the population-concept-context framework and PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidance, the planned review will systematically identify and categorize reported sources and types of bias, examine their effects on diagnostic performance across demographic and geographic subgroups, and document mitigation strategies applied throughout the AI life cycle. By synthesizing how bias and fairness considerations are handled in this field, this review aims to clarify existing evidence, highlight key gaps, and inform future efforts toward equitable and clinically trustworthy application of AI in cardiology.

Methods: We will conduct a comprehensive literature search across 5 electronic databases (PubMed, Embase, Cochrane CENTRAL, CINAHL, and IEEE Xplore) and gray literature sources. Eligible studies will include original research (2015-2025) evaluating the performance of AI-based ECG models across different subgroups or reporting on bias mitigation strategies. Two reviewers will independently screen studies, extract data using a standardized form, and resolve disagreements through consensus. This review will follow the PRISMA-ScR reporting framework.

Results: At the time of submission, study identification and screening has been completed. Database searches conducted in August and September 2025 yielded 430 records, with an additional 18 records identified through other sources. After duplicates removal, 398 unique records remained. Title and abstract screening led to the exclusion of 250 records, and 148 articles proceeded to full-text review. Following full-text assessment, 110 articles were evaluated for eligibility, of which 38 studies met the inclusion criteria and were included in the qualitative synthesis. The study selection process is summarized in a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram. Data extraction was conducted between November and December 2025.

Conclusions: This review will be the first to comprehensively map the landscape of algorithmic bias in AI-powered ECG interpretation. By identifying patterns of inequity and evaluating proposed solutions, it will provide actionable insights for developers, clinicians, and policymakers aiming to promote fairness in AI-enabled cardiac care.

XSL•FO
RenderX

International Registered Report Identifier (IRRID): PRR1-10.2196/82486

(JMIR Res Protoc 2026;15:e82486) doi:10.2196/82486

KEYWORDS

algorithmic bias; artificial intelligence; AI; diagnostic fairness; electrocardiography; health equity; machine learning; cardiovascular diagnostics; scoping review

Introduction

Background

Artificial intelligence (AI)–enabled electrocardiogram (ECG; AI-ECG) interpretation has emerged as a transformative tool in cardiovascular diagnostics. These models, particularly those built using deep learning architectures, have demonstrated high accuracy in identifying cardiac conditions such as atrial fibrillation, heart failure, and left ventricular dysfunction [1,2]. Given the global burden of cardiovascular disease, particularly in low-resource settings, AI-ECG tools offer the potential to enhance access to timely diagnosis, especially in areas with limited cardiology expertise [3].

However, as AI adoption accelerates, concerns about algorithmic bias have gained prominence. Algorithmic bias refers to systematic disparities in model performance across subpopulations, often arising from unrepresentative training data, flawed modeling assumptions, or inappropriate deployment environments [4-6]. These biases can lead to unequal diagnostic accuracy and risk misdiagnosis, particularly in underrepresented or structurally marginalized groups. In cardiology, for instance, tools trained predominantly on datasets composed predominantly of White male individuals have underperformed in detecting myocardial infarction in women and individuals of African descent [7,8]. Physical diagnostic tools such as pulse oximeters have similarly demonstrated racial bias in measurement accuracy [9].

Bias may emerge across multiple stages of the AI development pipeline, from data collection and labeling to model training, evaluation and deployment, and clinical integration. Recent frameworks categorize these into 5 key stages: data, model, evaluation, deployment, and postdeployment bias [10]. Addressing each of these stages is critical to ensuring fairness, safety, and clinical utility. High-income countries (HICs) have started to institutionalize fairness evaluations in medical AI development, with regulatory and ethical guidelines now recommending transparency, community engagement, and subgroup performance monitoring [11-13].

Despite this progress, significant gaps persist in understanding algorithmic bias within AI-ECG models, particularly in low- and middle-income countries (LMICs). Most existing studies and validations are based on datasets from North America and Europe, often excluding the very populations that might benefit most from AI-based screening tools [14]. This mismatch, described as “health data poverty” or “digital colonization,” raises concerns that AI tools exported to LMICs may underperform or cause unintended harm [15,16].

To date, there is no comprehensive synthesis of how bias manifests in AI-powered ECG interpretation. Existing reviews focus primarily on the diagnostic accuracy of AI-ECG for specific conditions and seldom report performance stratified by demographic or geographic subgroups [17]. A few studies have started to explore fairness or subgroup performance. However, findings are scattered, and no previous scoping review has mapped the full range of bias-related evidence or mitigation efforts in this field. It remains unclear which populations are most affected, at what stages bias arises, and what strategies have been attempted to reduce disparities.

With the increasing deployment of AI-ECG tools, including in LMICs, there is an urgent need to map the evidence on algorithmic bias in this domain systematically. Doing so will support the equitable development and safe deployment of AI-based ECG technologies across diverse populations and health systems. Early identification and mitigation of such bias are essential to diagnostic equity.

Objectives

This scoping review aims to comprehensively map and characterize the current evidence related to algorithmic bias in AI-powered ECG interpretation across the AI life cycle.

This review is guided by the population-concept-context (PCC) framework, which also informs the eligibility criteria:

- Population—adult patients (≥18 years) undergoing ECG-based cardiac assessment, particularly those from underrepresented groups (eg, women, racial and ethnic minority groups, and LMIC populations)
- Concept—algorithmic bias in AI-powered ECG interpretation, including disparities in diagnostic accuracy, error rates, and subgroup performance as well as documented mitigation strategies
- Context—any health care setting where AI-ECG models are developed, validated, or deployed, including both HICs and LMICs.

By addressing the following research questions, this review aims to provide a comprehensive overview of the field:

1. What types of algorithmic bias are reported in AI-powered ECG interpretation models?
2. At which stages of the AI life cycle (data, model, evaluation, deployment, and postdeployment stage) do these biases occur?
3. How do identified biases affect diagnostic performance and health care equity across different patient groups?
4. What mitigation strategies or fairness interventions have been proposed or tested in the context of AI-ECG?
5. How does AI-ECG performance vary across demographic and geographic subgroups, including sex and gender, race and ethnicity, age, and LMIC versus HIC settings?

XSL•FO
RenderX

This scoping review will inform future research, regulation, and implementation of AI-based cardiac diagnostics to ensure they serve all populations equitably.

Methods

Protocol Design

This study will be conducted as a scoping review, following the framework by Arksey and O'Malley [18], with enhancements by Levac et al [19] and methodological guidance from the Joanna Briggs Institute [20]. This review will be reported following the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist [21].

We will undertake the five standard stages of scoping reviews: (1) identifying the research question; (2) identifying relevant studies; (3) selecting studies; (4) charting the data; and (5) collating, summarizing, and reporting results.

We will skip the optional sixth stage (stakeholder consultation) because of resource and timing constraints.

Eligibility Criteria

Eligibility criteria are defined using the PCC framework recommended for scoping reviews.

Population

We will include studies involving adult human participants (aged ≥18 y) whose ECG data were analyzed using an AI-based tool. Studies must report subgroup performance data (eg, by sex, race and ethnicity, age, or geographic setting), even if they do not explicitly mention "bias." Pediatric populations will be excluded.

Concept

The phenomenon of interest is algorithmic bias in AI-powered ECG interpretation. Eligible studies must evaluate AI models (eg, machine learning and deep learning) used to interpret ECGs for cardiac diagnosis and must report on performance disparities, fairness metrics, or generalizability. Both commercial and research models will be included. Studies that assess model performance across groups, whether or not they use fairness terminology, will be eligible.

Context

We will include studies conducted in any global health care or research setting, including HICs and LMICs, across any level of care. No restrictions will be placed on the study setting or geographic region.

Study Designs

Eligible sources include all empirical studies presenting original data on AI-ECG performance and bias. These include diagnostic accuracy studies, randomized controlled trials, cohort and case-control studies, cross-sectional and retrospective validation studies, and technical validation studies (eg, internal or external validations of AI algorithms).

We will also include preprints, conference abstracts, and dissertations. Excluded sources include narrative and systematic reviews, commentaries, editorials, and purely theoretical or methodological papers without patient-level data.

Comparators

Studies with explicit or implicit comparators (eg, comparing model outputs across subgroups or against clinicians) will be included. No comparator is required if subgroup-specific performance metrics are reported.

Time Frame

Studies published from January 1, 2015, to July 31, 2025, will be eligible. This window reflects the period during which deep learning models became prominent in ECG interpretation [22].

Language

Only studies published in English were included. We acknowledge this as a limitation and will report any potentially relevant non-English studies excluded on this basis.

Information Sources and Search Strategy

Search Strategy Development

A comprehensive search strategy will be developed in collaboration with a medical librarian and peer reviewed using the Peer Review of Electronic Search Strategies (PRESS) checklist [23]. We will use both controlled vocabulary and free-text terms reflecting the PCC elements.

Databases

We will search the following databases including MEDLINE (via PubMed), Embase via Ovid, Cochrane CENTRAL, CINAHL (EBSCO), IEEE Xplore, and Web of Science Core Collection (as supplemental sources).

Gray Literature

Additional sources to be assessed include ProQuest Dissertations and Theses.

To optimize resource use while ensuring comprehensiveness, we will limit gray literature inclusion to a small set of high-yield sources. Specifically, we will search for conference abstracts from 3 major conferences where AI and ECG research is typically presented: American Heart Association, European Society of Cardiology, and American Medical Informatics Association.

We will limit the search to the years 2020 to 2025, using keyword combinations related to "ECG," "artificial intelligence," and "bias" in the title and abstract fields. Only abstracts that report original development or evaluation of AI-ECG tools will be included.

We will not include trial registries (eg, World Health Organization International Clinical Trials Registry Platform) because most entries lack completed results or sufficient methodological detail, and we aim to prioritize sources that offer extractable insights into model bias and evaluation.

Each step in the search and screening process will be documented, with independent screening by 2 reviewers to ensure transparency and reproducibility.

XSL•FO
RenderX

Search Terms

Search strings will include terms such as follows: (“Artificial Intelligence” OR “Machine Learning” OR “Deep Learning”) AND (“ECG” OR “Electrocardiogram”) AND (“Bias” OR “Fairness” OR “Health Equity” OR “Disparities” OR “Subgroup Analysis”).

Search strategies will be customized for each database. Full strategies will be provided in Multimedia Appendix 1. Filters will be applied for publication date (2015-2025) and English language.

To ensure sensitivity to global equity considerations, the primary search strategy will remain unrestricted by income-level terminology, preventing unintentional exclusion of studies that include LMIC data but do not explicitly use LMIC descriptors.

During data extraction, we will record geographic context and setting (LMIC, HIC, mixed, or unspecified) to support structured equity analysis. A supplementary targeted search using LMIC-related terms (eg, “low- and middle-income countries” and “resource-limited settings”) and citation chaining will be conducted to increase the likelihood of capturing LMIC-specific literature.

Study Selection

Screening Procedure

Study selection will occur in 2 stages: title and abstract screening and full-text screening.

Two reviewers will screen independently. Discrepancies will be resolved by consensus or by a third reviewer. A pilot of 100 citations will calibrate interreviewer agreement.

Screening Tool

Covidence (Veritas Health Innovation) will be used to manage the screening process and document inclusion and exclusion decisions.

PRISMA-ScR Flow Diagram

Study selection will be reported using a PRISMA-ScR flowchart, including reasons for exclusion at the full-text stage.

Data Extraction (Charting the Data)

Data Extraction Process

A structured data extraction form will be developed and piloted. Two reviewers will independently extract data from a sample of studies; one reviewer will complete the full extraction, with a second reviewer verifying all entries.

To ensure consistent interpretation of bias across studies, we will use a hierarchical extraction framework with predefined decision rules for each bias domain (sampling, measurement, label, aggregation, evaluation, and deployment). Reviewers will record (1) explicitly reported bias, (2) implicit evidence of bias (eg, subgroup performance differences and nonrepresentative datasets), or (3) insufficient information. Each domain is linked to operational indicators adapted from established AI ethics taxonomies [9], as shown in Multimedia Appendix 1. Two reviewers will independently extract and code bias-related information, with discrepancies resolved through consensus or adjudication by a third reviewer. This structured approach supports transparency and reproducibility in identifying bias manifestation across the AI life cycle.

Data Items

Key variables to be extracted are presented in Textbox 1 [9].

Textbox 1. Data items extracted in this scoping review.

- Study characteristics will include the author, year of publication, country, and journal or source.
- Study design and setting will include the study type and context, including whether the study was conducted in high-income countries or low- and middle-income countries and in primary or tertiary care settings.
- Population characteristics will include sample size, age, sex, race and ethnicity, and reported comorbidities.
- Artificial intelligence model characteristics will include the algorithm type, clinical task, development datasets, and model status (research or commercial).
- Performance and bias findings will include stratified performance metrics (eg, area under the curve, sensitivity, and specificity), subgroup disparities, and statistical significance.
- Bias life cycle stage will be recorded as data, model, evaluation, deployment, or postdeployment stage, classified per the typology by Mehrabi et al [9]. Operational definitions, decision criteria, and exemplar indicators for each bias category will guide reviewer judgment and ensure consistent application of the coding framework.
- Mitigation strategies will include techniques such as resampling, fairness constraints, post hoc calibration, or external validation.
- Author conclusions and implications will be extracted as reported in the included studies.

Data will be stored in Microsoft Excel or equivalent software for analysis. Ethical and contextual domains, including fairness reporting, dataset origin, governance, and community engagement, will be systematically extracted as dedicated fields to evaluate risks of digital colonization and data inequity.

Furthermore, repeated use of canonical ECG datasets may inflate the perceived generalizability. Such overlap will be treated as a data-stage bias signal, specifically related to sampling and aggregation bias. Studies leveraging the same datasets will be clustered to ensure findings are not interpreted as independent replications.

XSL•FO
RenderX

Data Analysis and Synthesis

Synthesis Approach

We will use descriptive and thematic analysis. Results will be synthesized to address the 5 guiding questions of the review as stated in the *Objective* section. Quantitative summaries will describe study counts, subgroup types, and bias stages. Narrative synthesis will highlight bias types and life cycle stages, performance disparities across subgroups, effectiveness of mitigation strategies, and geographic and contextual representation (HIC vs LMIC).

Visualizations

Findings will be presented through evidence tables (eg, bias type by study), charts or maps (eg, study distribution by year or geography), and summary matrices (eg, subgroup performance data).

Gaps and Recommendations

We will identify evidence gaps (eg, lack of LMIC data and underexplored bias stages) and recommend future research priorities. No meta-analysis will be performed due to expected heterogeneity.

Quality Appraisal

Consistent with Joanna Briggs Institute and PRISMA-ScR guidance, we will not conduct a formal quality appraisal. However, we will document evident methodological limitations (eg, small sample sizes or limited subgroup analysis) that may affect interpretation.

In studies reporting AI model development or validation, we will extract selected bias-related indicators aligned with the Prediction model Risk of Bias Assessment Tool–AI framework [24]. Although full risk-of-bias scoring will not be performed, this approach will support structured identification of potential bias-related to participant selection, data quality, model evaluation, and fairness reporting and will help flag limitations affecting subgroup generalizability.

Ethical Considerations

This review involves secondary analysis of publicly available literature and does not include any original research involving human participants. Therefore, it does not require approval from an institutional review board or ethics committee. No individual-level, sensitive, or identifiable data will be accessed or analyzed during this study, consistent with established ethical standards for systematic reviews [23].

Dissemination

The results of this scoping review will be submitted for publication in a peer-reviewed journal focused on digital health, AI in medicine, or cardiovascular care. In addition to traditional academic dissemination through journal articles and conference presentations, findings will also be shared with key stakeholders, such as health care AI developers, clinicians, and policymakers, via webinars, social media, and professional networks. Special emphasis will be placed on sharing the findings in ways that are accessible to stakeholders in LMICs, where bias in AI tools can be particularly impactful.

Results

Timeline

The database search was conducted in August and September 2025. As displayed in Figure 1, study screening and data extraction took place from November 2025 to December 2025. Data analysis and manuscript submission are planned for January and February 2026.

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart.

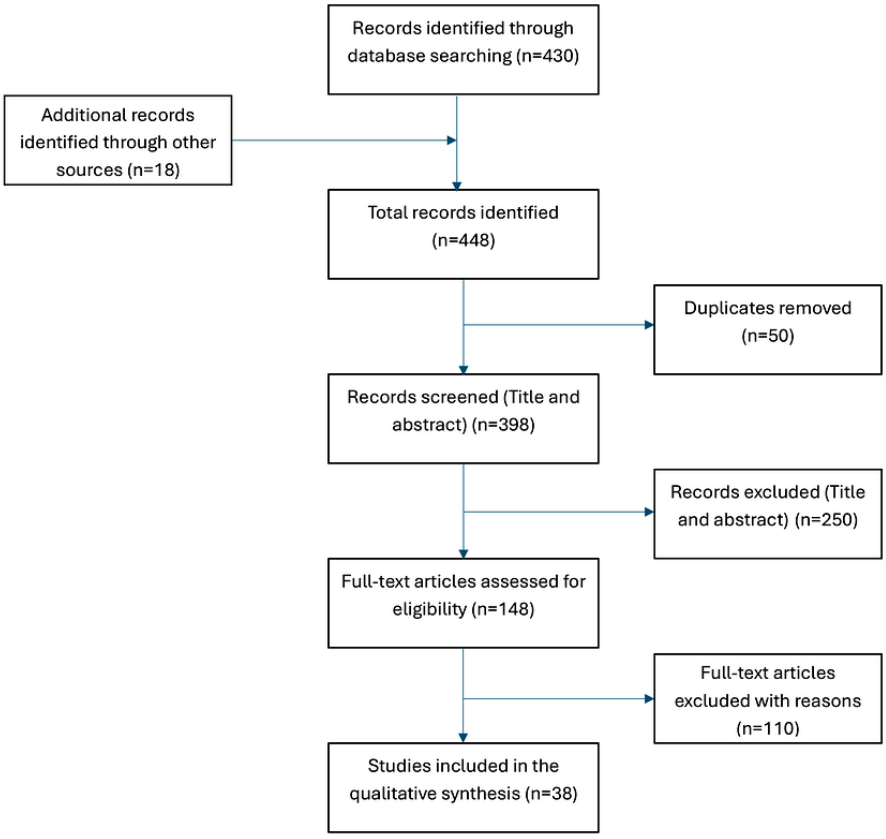

Reporting Standards Compliance

This protocol is reported following the PRISMA-P (Preferred Reporting Items for Systematic Review and Meta-Analyses Protocols) 2015 guidelines.

Progress to Date

The literature search strategy was finalized and executed across the specified databases in August and September 2025. Following duplicate removal, title and abstract screening was completed by two independent reviewers, leading to the identification of 148 records for full-text review. The full-text assessment concluded in November 2025, and 38 studies met the inclusion criteria for the qualitative synthesis. Data extraction was conducted using a standardized charting form between November and December 2025, with all extracted entries verified by a second reviewer. Descriptive and thematic analysis of the collated evidence is currently underway, with completion targeted for February 2026. The study selection process is summarized in a PRISMA flow diagram (Figure 1).

Anticipated Completion

We plan to complete the main scoping review manuscript by the end of February 2026. This will include final synthesis of extracted data, preparation of summary evidence tables and bias classification matrices, and drafting of the Results and Discussion sections. The completed review manuscript will be submitted for peer-reviewed publication shortly thereafter, consistent with PRISMA-ScR guidance and journal requirements.

Discussion

This protocol outlines a robust and comprehensive approach to mapping the evidence on algorithmic bias in AI-powered ECG interpretation. By applying the PCC framework and PRISMA-ScR guidance, this study is designed to identify and categorize systematically reported biases, assess their impact on diagnostic performance, and review mitigation strategies across the AI life cycle.

One of the main anticipated contributions of this review is its focus on both HIC and LMIC contexts, where the implications of algorithmic bias may differ substantially due to variations in data availability, clinical infrastructure, and deployment conditions. In doing so, this review addresses the persistent gap in the global representativeness of AI-ECG research and the underreporting of subgroup performance metrics.

This work is timely given the accelerating integration of AI in cardiovascular diagnostics and the increasing recognition that unchecked bias can exacerbate health inequities. Although previous reviews have examined the accuracy of AI-ECG systems, none have comprehensively mapped evidence on bias manifestations and mitigation efforts. By documenting both the problem space and potential solutions, our findings will provide actionable insights for developers, regulators, and clinicians.

We acknowledge certain limitations inherent to our protocol. Restricting inclusion to English-language publications may omit relevant studies, particularly from non–English-speaking LMICs. The expected heterogeneity in study designs and reporting standards will preclude meta-analysis, limiting the synthesis to

XSL•FO
RenderX

descriptive and narrative methods. In addition, the focus on published and indexed gray literature means that some proprietary or unpublished industry data, where bias assessments may have been conducted, will remain inaccessible.

Despite these constraints, our methodology incorporates multiple strategies to ensure comprehensiveness and transparency, including a librarian-led search strategy, independent screening by 2 reviewers, and standardized data extraction. The planned descriptive and thematic synthesis will enable a nuanced understanding of bias types, affected populations, and life cycle stages as well as the relative maturity of mitigation strategies.

Ultimately, the outputs of this review will serve as an evidence base for equitable AI-ECG development, deployment, and governance, helping to ensure that technological advancements in cardiology benefit all patient populations.

Acknowledgments

The authors gratefully acknowledge the support of the University of Oxford's Nuffield Department of Primary Health Care Sciences and the guidance of the DPhil supervisory team. The authors also thank the librarian they consulted during the development of the search strategy. This work is part of a doctoral research program supported by academic training and collaboration with the Mayo Clinic Platform and other global digital health partners. The authors declare the use of generative artificial intelligence (GenAI) in the research and writing process. According to the Generative AI Delegation Taxonomy, the following tasks were delegated to GenAI tools under full human supervision: proofreading, editing, and reformatting. The GenAI tool used was Grammarly to proofread, edit, and reformat the manuscript. Responsibility for the final manuscript rests entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the outcomes. All scientific content and interpretations are author generated.

Data Availability

The dataset generated from this scoping review will consist of extracted information from published studies. The extracted data will be made available from the corresponding author upon reasonable request after publication, subject to journal policies and any applicable restrictions.

Funding

This is part of a self-funded DPhil research work conducted by the first author at the University of Oxford. No external financial funding or grants were received for this work.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed database search strategies, eligibility criteria, and study selection framework based on the population–concept–context approach, data extraction template and variable definitions, bias classification framework across the artificial intelligence life cycle, operational definitions and coding rules for bias assessment, and summary of the extracted study characteristics and analytical fields.

[DOCX File , 32 KB - resprot_v15i1e82486_app1.docx]

References

1. Attia ZI, Noseworthy PA, Lopez-Jimenez F, Asirvatham SJ, Deshmukh AJ, Gersh BJ, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet 2019 Sep 07;394(10201):861-867. [doi: 10.1016/S0140-6736(19)31721-0] [Medline: 31378392]
2. Kwon JM, Kim KH, Jeon KH, Lee SY, Park J, Oh BH. Artificial intelligence algorithm for predicting cardiac arrest using electrocardiography. Scand J Trauma Resusc Emerg Med 2020 Oct 06;28(1):98 [FREE Full text] [doi: 10.1186/s13049-020-00791-0] [Medline: 33023615]
3. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019 Oct 25;366(6464):447-453 [FREE Full text] [doi: 10.1126/science.aax2342] [Medline: 31649194]
4. Pfohl SR, Foryciarz A, Shah NH. An empirical characterization of fair machine learning for clinical risk prediction. J Biomed Inform 2021 Jan;113:103621 [FREE Full text] [doi: 10.1016/j.jbi.2020.103621] [Medline: 33220494]
5. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med 2018 Dec 18;169(12):866-872 [FREE Full text] [doi: 10.7326/M18-1990] [Medline: 30508424]
6. Authors/Task Force members, Windecker S, Kolh P, Alfonso F, Collet JP, Cremer J, et al. 2014 ESC/EACTS Guidelines on myocardial revascularization: the Task Force on Myocardial Revascularization of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS) developed with the special contribution of the

XSL•FO
RenderX

European Association of Percutaneous Cardiovascular Interventions (EAPCI). Eur Heart J 2014 Oct 01;35(37):2541-2619 [FREE Full text] [doi: 10.1093/eurheartj/ehu278] [Medline: 25173339]
7. Chen IY, Szolovits P, Ghassemi M. Can AI help reduce disparities in general medical and mental health care? AMA J Ethics 2019 Feb 01;21(2):E167-E179 [FREE Full text] [doi: 10.1001/amajethics.2019.167] [Medline: 30794127]
8. Sjoding MW, Dickson RP, Iwashyna TJ, Gay SE, Valley TS. Racial bias in pulse oximetry measurement. N Engl J Med 2020 Dec 17;383(25):2477-2478 [FREE Full text] [doi: 10.1056/NEJMc2029240] [Medline: 33326721]
9. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. ACM Comput Surv 2021 Jul 13;54(6):1-35. [doi: 10.1145/3457607]
10. Panch T, Mattie H, Atun R. Artificial intelligence and algorithmic bias: implications for health systems. J Glob Health 2019 Dec;9(2):010318 [FREE Full text] [doi: 10.7189/jogh.09.020318] [Medline: 31788229]
11. London AJ. Artificial intelligence and black-box medical decisions: accuracy versus explainability. Hastings Cent Rep 2019 Jan;49(1):15-21. [doi: 10.1002/hast.973] [Medline: 30790315]
12. Chen JH, Asch SM. Machine learning and prediction in medicine - beyond the peak of inflated expectations. N Engl J Med 2017 Jun 29;376(26):2507-2509 [FREE Full text] [doi: 10.1056/NEJMp1702071] [Medline: 28657867]
13. Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD) action plan. U.S. Food & Drug Administration. 2021. URL: https://www.fda.gov/media/145022/download [accessed 2025-12-17]
14. Ethics and governance of artificial intelligence for health:WHO guidance. World Health Organization. 2021 Jun 28. URL: https://www.who.int/publications/i/item/9789240029200 [accessed 2025-12-18]
15. Aseidu M, Dieng A, Haykel I, Rostamzadeh N, Pfohl S, Nagpal C, et al. The case for globalizing fairness: a mixed methods study on colonialism, AI, and health in Africa. Preprint on ArXiv 2024 Mar 11:4 [FREE Full text] [doi: 10.48550/arXiv.2403.03357]
16. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019 Oct 25;366(6464):447-453 [FREE Full text] [doi: 10.1126/science.aax2342] [Medline: 31649194]
17. Mittelstadt BD, Floridi L. The ethics of big data: current and foreseeable issues in biomedical contexts. Sci Eng Ethics 2016 Apr;22(2):303-341. [doi: 10.1007/s11948-015-9652-2] [Medline: 26002496]
18. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2007 Feb 23;8(1):19-32. [doi: 10.1080/1364557032000119616]
19. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
20. Peters M, Godfrey CM, McInerney P, Munn Z, Trico A, Khalil H. Scoping reviews. In: Aromataris E, Munn Z, editors. JBI Manual for Evidence Synthesis. Adelaide, Australia: Joanna Briggs Institute; 2020.
21. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
22. Hannun AY, Rajpurkar P, Haghpanahi M, Tison GH, Bourn C, Turakhia MP, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nat Med 2019 Jan;25(1):65-69 [FREE Full text] [doi: 10.1038/s41591-018-0268-3] [Medline: 30617320]
23. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75:40-46 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
24. Moons KM, Damen JA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, et al. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. BMJ 2025 Mar 24;388(1):e082505-e082508 [FREE Full text] [doi: 10.1136/bmj-2024-082505] [Medline: 40127903]

Abbreviations

AI: artificial intelligence
AI-ECG: artificial intelligence–enabled electrocardiogram
ECG: electrocardiogram
HIC: high-income country
LMIC: low- and middle-income country
PRESS: Peer Review of Electronic Search Strategies
PRISMA-P: Preferred Reporting Items for Systematic Review and Meta-Analyses Protocols
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

XSL•FO
RenderX

Edited by J Sarvestan; submitted 15.Aug.2025; peer-reviewed by W Wang; comments to author 04.Nov.2025; revised version received 16.Dec.2025; accepted 17.Dec.2025; published 20.Jan.2026.

Please cite as:
Lawal L, Paton C, English M, Holthof B, Preston T
Mapping Algorithmic Bias in AI-Powered Electrocardiogram Interpretation Across the AI Life Cycle: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e82486
URL: https://www.researchprotocols.org/2026/1/e82486
doi:10.2196/82486
PMID:

©Luqman Lawal, Christopher Paton, Mike English, Bruno Holthof, Tabitha Preston. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 20.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Characteristics of Mobile Health Interventions for Repetitive Negative Thinking: Protocol for a Scoping Review

Judith Martens[1,2], MSc; Thomas Forkmann[1], Prof Dr; Inken Höller[2], Prof Dr

[1]Department of Clinical Psychology and Psychotherapy, University of Duisburg-Essen, Essen, Germany

[2]Department of Clinical Psychology and Psychotherapy, Charlotte Fresenius Hochschule, Düsseldorf, Germany

Corresponding Author:
Judith Martens, MSc
Department of Clinical Psychology and Psychotherapy
University of Duisburg-Essen
Universitätsstr. 2
Essen, 45141
Germany
Phone: 49 201 1835441
Email: judith.martens@stud.uni-due.de

Abstract

Background: Many people are affected by mental disorders. A transdiagnostic symptom and risk factor for most mental disorders is repetitive negative thinking (RNT). Psychotherapy can reduce RNT, but most people in need either do not receive psychotherapy or face long waiting times. In addition, people at risk for developing mental disorders do not receive psychotherapy. Mobile health (mHealth) interventions could overcome treatment barriers and support people at risk.

Objective: This scoping review aims to identify existing mHealth interventions for RNT and to give an overview of their characteristics regarding content, context, and technical features. Another aim is to identify which outcomes and questionnaires are used to measure RNT.

Methods: The scoping review will be conducted according to the JBI (Joanna Briggs Institute) methodology for scoping reviews and will be reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines. English-language, peer-reviewed literature involving mHealth interventions aimed at reducing RNT from 2003 onward will be included. A comprehensive search will be conducted in the following databases: PsycInfo, PubMed, Scopus, and Web of Science. Two independent reviewers will conduct a 2-stage blinded screening process (screening of title or abstract and full text) of the articles according to the inclusion criteria. A data extraction table will be used to extract information on the technology, content, delivery, accessibility, usability, feedback, and outcome measures of the interventions. Data charting, including coding and grouping, will follow an iterative process. The charted data will then be synthesized descriptively.

Results: Data collection began in February 2025 and is now complete, with 21 included articles meeting the inclusion criteria. As of December 2025, data charting is currently underway, and data synthesis will begin shortly. The final scoping review is expected to be submitted to a peer-reviewed journal by April 2026.

Conclusions: The scoping review should lead to a better understanding of the conceptual possibilities and commonalities of current mHealth interventions for the transdiagnostic symptom RNT. It should provide starting points for systematic reviews and for the development of transdiagnostic mHealth interventions.

International Registered Report Identifier (IRRID): PRR1-10.2196/72138

(JMIR Res Protoc 2026;15:e72138) doi:10.2196/72138

KEYWORDS

mobile health; mHealth; digital health; repetitive negative thinking; digital interventions; transdiagnostic; rumination; scoping review

XSL•FO
RenderX

Introduction

Epidemiology

Epidemiological studies show that 1 in 3 persons is affected by a mental disorder in their lifetime, and 1 in 5 persons has experienced a mental disorder within the last 12 months [1,2]. Therefore, the prevalence of mental disorders in the population is alarmingly high, and the burden, societal impacts, and costs are massive [1]. Moreover, mental disorders do not only impair quality of life [3] but also serve as a risk factor for suicide [4] and are related to a greater risk of developing chronic physical illnesses [5]. The detection and especially the treatment of symptoms are therefore indispensable. In particular, symptoms that occur across different disorders should be targeted for treatment.

Repetitive Negative Thinking

A transdiagnostic symptom that occurs in almost all mental disorders is repetitive negative thinking (RNT) [6]. Ehring and Watkins [6] define RNT "as repetitive thinking about one or more negative topics that is experienced as difficult to control." Two common forms of RNT are rumination and worry. Both have so far mostly been recorded in the literature on a disorder-specific basis [6]. For example, rumination was usually measured in the context of depression [7], and worry was usually measured in the context of generalized anxiety disorder [8]. However, as RNT has more similarities than differences between different disorders, it is argued that RNT can be referred to as a transdiagnostic process [6]. Ehring and Watkins [6] found evidence that RNT occurs in up to 13 different disorders, for example, depression, social phobia, and obsessive-compulsive disorder.

Furthermore, RNT can be found not only as a transdiagnostic symptom but also as a risk factor for the development of a mental disorder, which may also be causal for the maintenance of mental disorders. Studies have shown that rumination serves as a transdiagnostic risk factor impacting anxiety, depression, psychosis, insomnia, and impulsive behaviors [9]. In addition, rumination intensifies psychopathological conditions by amplifying and protracting negative affective states, impeding effective problem-solving and instrumental behaviors, and diminishing responsiveness to evolving circumstances [9]. These findings demonstrate the high relevance of RNT as a treatment target in psychotherapy already at an early stage and not only when the mental disorder has already manifested itself.

There are several treatment approaches in psychotherapy offering different interventions that can reduce RNT, for example, cognitive behavioral therapy (CBT) [10]. However, only individuals who are actively engaged in psychotherapy can benefit from these approaches. Studies show that most people with mental disorders do not receive treatment [11-13] or drop out of treatment [14]. Furthermore, the waiting time in Germany for those who eventually receive psychotherapy is, on average, 4.6 months [15]. In addition, individuals, who engage in RNT and may be at risk for developing a mental disorder, may not inherently perceive the imperative for comprehensive therapeutic involvement. This is in line with a study by Andrade et al [16] who found that the greatest barriers to receiving and staying in treatment are low perceived need and attitudinal barriers. For those who recognized a need, one of the most common reasons for not getting treatment was the wish to deal with the problem themselves. Self-stigmatization and label avoidance could also be linked to this desire [16]. Another reason for the lack of treatment are structural barriers, such as finance and availability (particularly in developing countries), especially for severe disorders [16].

Mobile Health

Therefore, the reduction of barriers for receiving fast and low-threshold treatment that addresses RNT is of high importance. One option could be to look for new interventions beyond the classic face-to-face treatments. One approach that has gained prominence in overcoming barriers, particularly in the context of the COVID-19 pandemic, is the use of mobile health (mHealth) interventions [17], first introduced in 2003 [18,19]. mHealth uses mobile and wireless technologies for health, such as smartphones and tablets [20]. These interventions can vary in structure, components, content, duration, and other characteristics. One specific type of mHealth intervention is ecological momentary interventions (EMIs) [21]. EMIs use real-time data from ecological momentary assessment [22]. Through real-time monitoring and data collection in individuals' environments [22], EMIs enable the delivery of treatments within these settings, offering support when needed [21].

An RNT is not only a symptom but also a risk factor for the development of mental disorders [7], and low-threshold treatment is needed for people at risk. Therefore, EMI can offer an important opportunity for the treatment of RNT and the prevention of mental disorders. EMIs provide support when needed, which is critical because RNT fluctuates over the course of time [23]; it has been shown to occur predominantly in the evening in individuals with severe depression [23]. As a treatment option for people with mental disorders, EMI could additionally help to implement techniques learned in psychotherapy in everyday life, such as supporting the practice of techniques as homework, which can be difficult in psychotherapy for various reasons, including patients' characteristics such as low motivation or self-doubts [24]. mHealth interventions can overcome this problem by being accessible anytime and anywhere. This could be beneficial not only for people who are receiving treatment but also for those who are waiting for treatment or have already finished treatment in terms of relapse prevention.

mHealth interventions have gained attention in recent years. Previous reviews about mHealth interventions in the context of mental health have focused on the use and effects of smartphone interventions on mental health in general [25], on specific disorders such as depression [26] or anxiety disorders [27], as well as on the management of suicidal thoughts and behavior [28]. There are also reviews on the use and efficacy of certain interventions for RNT, such as digital mindfulness-based interventions [29].

However, there is a gap in the overview of available mHealth interventions for RNT. As there is a wide range of variations in mHealth interventions (in terms of duration, type of intervention, frequency, content, components, intensity, etc), it

XSL•FO
RenderX

seems useful to first use a broader scoping review to provide an overview of mHealth interventions that address RNT and their specific characteristics. Gaps and potential for improvement of mHealth interventions will be identified and discussed to establish a foundation for the development of subsequent interventions.

Review Questions

The review questions are as follows:

1. What mHealth interventions for RNT are currently available for people who experience RNT?
2. What are the contents, contexts, and technical features of these interventions?
3. What outcomes and questionnaires have been used to measure RNT within the mHealth interventions and as an outcome variable to evaluate the efficacy of the interventions in relation to RNT?

Methods

Overview

To ensure quality and transparency of the scoping review, this scoping review protocol and the final scoping review will be conducted in accordance with the JBI (Joanna Briggs Institute) methodology for scoping reviews [30] and scoping review protocols [31]. The final review will also be reported in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines [32]. The completed PRISMA-ScR checklist can be found in Multimedia Appendix 1. The mHealth evidence reporting and assessment guidelines [20] will be used as the basis for the evaluation criteria for the interventions. The scoping review is preregistered at OSF [33]. Pilot testing of inclusion criteria was conducted after registration but before screening on a small sample of studies to ensure clearly defined boundaries. The inclusion criteria were then further specified before the screening process started. The refined criteria are reported in the Eligibility Criteria section.

Ethical Considerations

As no human participants were involved in this research, no ethics approval is required.

Eligibility Criteria

Eligibility criteria are based on the population-concept-context framework, as recommended for systematic scoping reviews [31] (Table 1).

Table 1. Population-concept-context (PCC) framework.

PCC framework	Definition
Population	People experiencing RNT[a]
Concept	mHealth[b] interventions that aim to reduce RNT
Context	User's everyday life and real-world environment

[a]RNT: repetitive negative thinking.

[b]mHealth: mobile health.

Population

The target population of the mHealth interventions must consist of people who experience RNT. The intervention study must explicitly include participants who exhibit RNT, as verified through appropriate means. This must include the use of specific RNT questionnaires (eg, Perseverative Thinking Questionnaire [34]) or the use of diagnostic or symptom-based questionnaires for disorders in which RNT is a core feature—this includes depression and generalized anxiety disorder [6]. Apart from the presence of RNT, no further restrictions were applied regarding sample type—which means that all kinds of sample types, for example, clinical and nonclinical populations, were included, irrespective of specific diagnoses, as RNT is conceptualized as transdiagnostic [6]. In terms of age, we expanded the inclusion criteria from including all mHealth interventions for RNT for participants above the age of 18 years, as stated in the scoping review protocol, to adolescents. This is because research indicates that RNT already develops in adolescence [35]. Therefore, providing interventions early is essential for the prevention of long-term psychological difficulties. The inclusion age of participants was determined based on other research on RNT in adolescence [35], and therefore, people aged 10 years and older will be included.

Concept

The concept of interest was mHealth interventions that aim to reduce RNT. mHealth was defined such that the interventions had to be delivered via smartphones or tablets or at least had to be accessible with smartphones or tablets, even if they were web based. Computer- or web-based interventions that were only accessible via computers or laptops were not considered. Articles had to aim to address RNT. To clarify this inclusion criterion, the following restrictions were set before the search process started: the aim of the intervention had to be stated as reducing some form of RNT (eg, RNT, rumination, and worry). The intervention had to have been developed to change RNT or at least a theoretical rationale or theoretical model had to be provided as to why this intervention was intended to reduce RNT. Articles or interventions were excluded if the intervention was not explicitly designed for RNT, and there was no theoretical and scientific justification as to why the intervention was intended to help reduce rumination. Articles were also not included if they contained only assessments.

Another concept of interest was the outcomes and assessment tools to assess RNT within these interventions. Therefore, studies had to include an outcome measure that assesses RNT, either as a primary or secondary outcome. Any form of quantitative assessment was eligible regardless of whether the

XSL•FO
RenderX

measurement tool was validated. The study had to report on RNT as part of the intervention outcomes.

Context

There were no restrictions in terms of geographic context or setting. As the idea was to receive the intervention anywhere and anytime, the intervention had to be available to the participants at all times.

Types of Evidence Sources

As the scoping review will focus on intervention design without assessing efficacy, all study designs were included. Only peer-reviewed, full-text articles published in or translated into English were considered. When an intervention appeared in both a review and its primary source, the primary source was prioritized to avoid duplication. Gray literature and other nonpeer-reviewed sources were not included.

Search Strategy

For this review, a search was conducted in the following electronic library databases: PsycInfo, PubMed, Scopus, and Web of Science. The concepts of interests were mHealth interventions and RNT. The search strategy was based on previous literature on the concepts of interest [6,36-38]. As research has shown that worry and depressive rumination belong to a common construct [6], both terms were included. When searching for sources of evidence, articles published between 2003 and shortly before the scoping review was constructed were considered because 2003 is the year when the first paper was published that defined and introduced the term mHealth [18,19]. The full search string for each database can be found in Multimedia Appendix 2.

The reference lists of relevant articles were also searched, and relevant articles were included. If new relevant terms appeared during the search, they were added to the search string and will be reported in the review.

Study Selection

First, all articles identified through the search were downloaded and imported into the reference manager Zotero [39]. After that, all identified articles were uploaded to the online tool Rayyan [40]. Next, duplicates will be removed. The remaining articles will then be screened to determine if they meet the inclusion criteria. First, the titles and abstracts were screened by 2 blinded reviewers according to the inclusion criteria. The blinding was then lifted, and conflicts and uncertainties were discussed, and more reviewers were consulted if no consensus could be reached. Full texts were uploaded to Rayyan, the reviewers were then reblinded, and the full texts of the remaining articles were reviewed according to the inclusion criteria. Conflicts and uncertainties were discussed again, and if no agreement could be reached, more reviewers were again consulted until consensus was reached. The reasons for exclusions will be documented and reported in the review. References for the final selection of articles were uploaded to Zotero [39] and stored in a dedicated folder. To represent the inclusion and exclusion process, the selection procedure will be mapped and illustrated using a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [41].

Data Extraction

First, the data to be extracted are basic study information (eg, author, title, publication year, and location) and sample characteristics (eg, age, gender, and target group). Following the mHealth evidence reporting and assessment guidelines [20], a data extraction table has been designed detailing the data to first evaluate the interventions in case of content, context, technical features, and second, to map which outcomes and assessment tools are used to measure RNT (Table 2). One subsection, called “tailoring,” indicates whether the intervention is an EMI, that is, whether real-time data are used to adapt the interventions. The described table draft will be tested on a subset of studies to evaluate and adjust or expand if necessary. The final table will be documented in the final scoping review. The extracted data from the relevant articles were added verbatim to the data extraction table. The extraction process is being conducted by 1 reviewer (JM), with support from another reviewer (research assistant). In case of disagreements, a third reviewer (IH) is being consulted.

XSL•FO
RenderX

Table 2. Data extraction table for intervention evaluation criteria.

Domain and evaluation criteria	Description with example
Technology platform	
Program used	This item describes the used program, for example, a specific app.
Requirements	This item describes the technical requirements for the program, for example, the operating system such as iOS or Android.
Content	
Therapeutic approach or intervention strategies	This item describes the therapeutic approach of the intervention, for example, cognitive behavioral therapy.
Components	This describes single components of the intervention, for example, psychoeducation, self-monitoring.
Communication	This evaluates how communication with the participant takes place, for example, whether there is human contact or with an artificial intelligence.
Delivery	
Tailoring	This records whether real-time data are collected and if so, how they are used to customize the intervention, for example, by offering certain interventions.
Mode and design of components	This describes the mode and design of the components, that is, the way in which the components are presented, for example, with text messages, videos, or audio.
Timing of components	This describes at what time the components take place, for example, randomly or in the morning.
Frequency of components	This describes how often the components take place, for example, 5 times a day or every hour.
Duration of components	This describes how long the single components take, for example, 10 min for psychoeducation.
Duration of intervention	This describes how long the whole intervention takes, for example, 7 d.
Accessibility	
Costs	This describes whether the intervention costs anything and, if so, how much.
Barriers	This describes whether there are barriers that restrict the use of the intervention, for example, an internet connection.
Language	This describes in which languages the intervention is available or whether it is possibly language free.
Usability	
User involvement in development	This describes whether and how the users were involved in the development of the intervention, for example, through focus groups.
Adherence	This describes the degree to which participants follow and complete the intervention.
Feedback	
User satisfaction	This describes user satisfaction with the intervention, for example, recorded via questionnaires.
Outcomes	
Construct	This describes the construct of interest that is to be reduced, for example, depressive rumination.
Operationalization	This describes the assessment tool that is intended to capture the construct of interest, for example, Ruminative Response Scale.

Data Charting and Synthesis

After data were added verbatim to the data extraction table according to the predefined categories (technology, content, delivery, accessibility, usability, feedback, and outcomes), the data will be charted by coding and grouping using an iterative process. New relevant categories that emerged during analysis will be added, existing categories will be refined, expanded, or merged, and missing information will be marked. Terms and concepts will be standardized, for example, CBT for cognitive behavioral therapy or manual for cognitive behavioral therapy.

The charted data will be synthesized using a descriptive approach with regard to the research questions and presented using text and tables. Highlights and key findings, such as

XSL•FO
RenderX

similarities, differences, or special features regarding the interventions, will be reported.

Results

The study selection process is completed, and 21 articles meeting the inclusion criteria have been identified. The complete inclusion and exclusion process will be illustrated in the final scoping review using a PRISMA flow diagram [41]. Data extraction has been completed, and data charting is underway, with data synthesis starting soon. Data synthesis is expected to be completed by February 2026, and the manuscript is expected to be finalized by April 2026. It will be submitted to an international peer-reviewed journal.

Discussion

The aim of this scoping review is to provide an overview of existing mHealth interventions that address RNT and their specific characteristics. Hypothesized main findings include heterogeneity in the construction of the identified mHealth interventions of RNT. As data extraction and synthesis are still pending, no final results can yet be presented. However, we expect differences in the (1) technology, (2) content, (3) delivery, (4) accessibility, (5) usability, (6) feedback, and (7) outcome measures of the interventions.

Anticipated Findings

Regarding technology, there exist multiple different platforms, which are used to implement mHealth interventions, such as mPath [42] or movisensXS [43]. The content of the interventions might differ due to numerous available approaches for RNT treatment, such as CBT or mindfulness [44]. The delivery of interventions will likely vary, particularly in terms of length and number of delivery as common suggestions in mHealth study designs are lacking, for example, within mHealth for youth [45]. The accessibility might be restricted, especially for older adults [46] and populations in low- and middle-income countries [47]. In terms of usability, it is assumed that adherence might be insufficient [48] and information on adherence and user engagement with the interventions might be lacking [49]. Regarding feedback, assessments of user satisfaction are expected to be heterogeneous, and measures might not have been validated [50]. Finally, outcome measures are expected to vary across studies, reflecting the diversity of instruments available, such as the Perseverative Thinking Questionnaire [34], the Ruminative Response Scale [51], or the Penn State Worry Questionnaire [52].

Another presumed finding is that reporting of information on different aspects of the implemented interventions may be insufficient in many studies. Prior research has, for example, found that safety aspects are insufficiently reported [53] and that information on user engagement is often underreported [49]. Research has also shown that information on the decision-making basis for the design and choice of interventions is lacking and that future studies should empirically justify the design and selection of intervention components [36].

If the results of the final scoping review are as expected, it will be crucial for future research to develop comprehensive study protocols; provide all relevant information; and use existing guidelines and checklists to ensure comprehensibility, transparency, and reproducibility.

Limitations

Although this research is the first to provide an overview of existing mHealth interventions for RNT and thus offers important starting points for further research, some limitations must be considered. As this paper is a scoping review, the efficacy of included interventions is not evaluated and no conclusions of effectiveness of the individual components or contents can be drawn. Given the wide range of design possibilities, it was the goal to first provide an overview of the designs and implementations of mHealth interventions for RNT before evaluating them.

Evaluation should be relevant for future research steps aiming at further advancing the prevention of mental health disorders by treating the risk factor RNT. Future studies might also need to differentiate more closely between interventions for RNT with regard to subgroups of interventions (eg, CBT interventions vs mindfulness).

Conclusions

All in all, mHealth interventions are promising to address gaps in care, provide supplementary support, and enhance access to mental health resources for individuals at risk or individuals who might otherwise face barriers to psychotherapy. The results of the final scoping review will be used as a starting point for new research within our academic institution to develop transdiagnostic mHealth interventions and later investigate their usability and effectiveness.

Acknowledgments

The authors acknowledge support from the Open Access Publication Fund of the University of Duisburg-Essen. The authors declare the use of generative artificial intelligence in the research and writing process. According to the GAIDeT (Generative AI Delegation Taxonomy; 2025), the following tasks were delegated to generative artificial intelligence tools under full human supervision: proofreading and editing, summarizing text, and translation. The generative artificial intelligence tool used was ChatGPT. Responsibility for the final manuscript lies entirely with the authors.

Funding

The publication costs will be covered by the University of Duisburg-Essen, the affiliation of the corresponding author. No further external financial support or grants were received from any public, commercial, or not-for-profit entities for the research or

XSL•FO
RenderX

authorship of this paper. The funder was not involved in the study design, data collection, analysis, and interpretation or writing of the manuscript.

Data Availability

All data used in the final scoping review consist of previously published studies, which will be cited in the manuscript. The data extraction table summarizing study characteristics and key findings will be available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: JM, IH
Data curation: JM
Formal analysis: JM
Methodology: JM, IH
Resources: IH
Supervision: TF, IH
Writing—original draft: JM
Writing—review and editing: TF, JM
All authors approved the final version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PRISMA-ScR checklist.
[DOCX File , 86 KB - resprot_v15i1e72138_app1.docx]

Multimedia Appendix 2
Search strategy.
[DOCX File , 16 KB - resprot_v15i1e72138_app2.docx]

References

1. Kessler RC, Aguilar-Gaxiola S, Alonso J, Chatterji S, Lee S, Ormel J, et al. The global burden of mental disorders: an update from the WHO World Mental Health (WMH) surveys. Epidemiol Psichiatr Soc 2009;18(1):23-33 [FREE Full text] [doi: 10.1017/s1121189x00001421] [Medline: 19378696]
2. Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, et al. The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. Int J Epidemiol 2014 Apr;43(2):476-493 [FREE Full text] [doi: 10.1093/ije/dyu038] [Medline: 24648481]
3. Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, ESEMeD/MHEDEA 2000 Investigators, European Study of the Epidemiology of Mental Disorders (ESEMeD) Project. Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand Suppl 2004 Jun 14;109(420):38-46. [doi: 10.1111/j.1600-0047.2004.00329.x] [Medline: 15128386]
4. Nock MK, Borges G, Bromet EJ, Alonso J, Angermeyer M, Beautrais A, et al. Cross-national prevalence and risk factors for suicidal ideation, plans and attempts. Br J Psychiatry 2008 Feb 02;192(2):98-105 [FREE Full text] [doi: 10.1192/bjp.bp.107.040113] [Medline: 18245022]
5. Scott KM, Lim C, Al-Hamzawi A, Alonso J, Bruffaerts R, Caldas-de-Almeida JM, et al. Association of mental disorders with subsequent chronic physical conditions: world mental health surveys from 17 countries. JAMA Psychiatry 2016 Feb 01;73(2):150-158 [FREE Full text] [doi: 10.1001/jamapsychiatry.2015.2688] [Medline: 26719969]
6. Ehring T, Watkins ER. Repetitive negative thinking as a transdiagnostic process. Int J Cogn Ther 2008 Sep;1(3):192-205. [doi: 10.1521/ijct.2008.1.3.192]
7. Nolen-Hoeksema S. Responses to depression and their effects on the duration of depressive episodes. J Abnorm Psychol 1991 Nov;100(4):569-582. [doi: 10.1037//0021-843x.100.4.569] [Medline: 1757671]
8. Goodwin H, Yiend J, Hirsch CR. Generalized anxiety disorder, worry and attention to threat: a systematic review. Clin Psychol Rev 2017 Jun;54:107-122. [doi: 10.1016/j.cpr.2017.03.006] [Medline: 28448826]
9. Watkins ER, Roberts H. Reflecting on rumination: consequences, causes, mechanisms and treatment of rumination. Behav Res Ther 2020 Apr;127:103573. [doi: 10.1016/j.brat.2020.103573] [Medline: 32087393]

XSL•FO
RenderX

10. Spinhoven P, Klein N, Kennis M, Cramer AO, Siegle G, Cuijpers P, et al. The effects of cognitive-behavior therapy for depression on repetitive negative thinking: a meta-analysis. Behav Res Ther 2018 Jul;106:71-85. [doi: 10.1016/j.brat.2018.04.002] [Medline: 29699700]
11. Kessler RC, Demler O, Frank RG, Olfson M, Pincus HA, Walters EE, et al. Prevalence and treatment of mental disorders, 1990 to 2003. N Engl J Med 2005 Jun 16;352(24):2515-2523 [FREE Full text] [doi: 10.1056/NEJMsa043266] [Medline: 15958807]
12. Thornicroft G. Most people with mental illness are not treated. Lancet 2007 Sep 08;370(9590):807-808. [doi: 10.1016/S0140-6736(07)61392-0] [Medline: 17826153]
13. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. Lancet 2007 Sep 08;370(9590):841-850 [FREE Full text] [doi: 10.1016/S0140-6736(07)61414-7] [Medline: 17826169]
14. Wells JE, Browne MO, Aguilar-Gaxiola S, Al-Hamzawi A, Alonso J, Angermeyer MC, et al. Drop out from out-patient mental healthcare in the World Health Organization's World Mental Health Survey initiative. Br J Psychiatry 2013 Jan 02;202(1):42-49. [doi: 10.1192/bjp.bp.112.113134] [Medline: 23174514]
15. Zepf S, Mengele U, Hartmann S. Zum Stand der ambulanten psychotherapeutischen Versorgung der Erwachsenen in der Bundesrepublik Deutschland. Psychother Psychosom Med Psychol 2003 Mar;53(3-4):152-162. [doi: 10.1055/s-2003-38004]
16. Andrade LH, Alonso J, Mneimneh Z, Wells JE, Al-Hamzawi A, Borges G, et al. Barriers to mental health treatment: results from the WHO World Mental Health surveys. Psychol Med 2014 Apr 09;44(6):1303-1317 [FREE Full text] [doi: 10.1017/S0033291713001943] [Medline: 23931656]
17. John Leon Singh H, Couch D, Yap K. Mobile health apps that help with COVID-19 management: scoping review. JMIR Nurs 2020 Aug 6;3(1):e20596 [FREE Full text] [doi: 10.2196/20596] [Medline: 32897271]
18. Istepanian RS. Mobile health (m-Health) in retrospect: the known unknowns. Int J Environ Res Public Health 2022 Mar 22;19(7):3747 [FREE Full text] [doi: 10.3390/ijerph19073747] [Medline: 35409431]
19. Istepanian RS, Lacal JC. Emerging mobile communication technologies for health: some imperative notes on m-health. In: Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2003 Presented at: IEMBS '03; September 17-21, 2003; Cancun, Mexico p. 1414-1416. [doi: 10.1109/iembs.2003.1279581]
20. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
21. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
22. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008 Apr;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
23. Takano K, Tanno Y. Diurnal variation in rumination. Emotion 2011 Oct;11(5):1046-1058. [doi: 10.1037/a0022757] [Medline: 21517158]
24. Helbig S, Fehm L. Problems with homework in CBT: rare exception or rather frequent? Behav Cogn Psychother 2004 Jun 14;32(3):291-301. [doi: 10.1017/S1352465804001365]
25. Linardon J, Cuijpers P, Carlbring P, Messer M, Fuller-Tyszkiewicz M. The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. World Psychiatry 2019 Oct 09;18(3):325-336 [FREE Full text] [doi: 10.1002/wps.20673] [Medline: 31496095]
26. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
27. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
28. Jiménez-Muñoz L, Peñuelas-Calvo I, Díaz-Oliván I, Gutiérrez-Rojas L, Baca-García E, Porras-Segovia A. Suicide prevention in your pocket: a systematic review of ecological momentary interventions for the management of suicidal thoughts and behaviors. Harv Rev Psychiatry 2022;30(2):85-99. [doi: 10.1097/HRP.0000000000000331] [Medline: 35267251]
29. Vargas-Nieto J, Zambrano D, Montorio I, Sanabria LM. Efficacy of digital mindfulness-based interventions for repetitive negative thinking: a systematic review and bibliometric analysis. Mindfulness 2024 Feb 29;15(3):523-538. [doi: 10.1007/s12671-024-02321-w]
30. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
31. Peters MD, Godfrey C, McInerney P, Khalil H, Larsen P, Marnie C, et al. Best practice guidance and reporting items for the development of scoping review protocols. JBI Evid Synth 2022 Apr 01;20(4):953-968. [doi: 10.11124/JBIES-21-00242] [Medline: 35102103]

XSL•FO
RenderX

32. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
33. Characteristics of mobile health interventions for repetitive negative thinking - a scoping review. Open Science Framework. URL: https://doi.org/10.17605/OSF.IO/N5H7F [accessed 2026-01-16]
34. Ehring T, Zetsche U, Weidacker K, Wahl K, Schönfeld S, Ehlers A. The Perseverative Thinking Questionnaire (PTQ): validation of a content-independent measure of repetitive negative thinking. J Behav Ther Exp Psychiatry 2011 Jun;42(2):225-232 [FREE Full text] [doi: 10.1016/j.jbtep.2010.12.003] [Medline: 21315886]
35. Jose PE, Brown I. When does the gender difference in rumination begin? Gender and age differences in the use of rumination by adolescents. J Youth Adolescence 2007 Jan 23;37(2):180-192. [doi: 10.1007/s10964-006-9166-y]
36. Balaskas A, Schueller SM, Cox AL, Doherty G. Ecological momentary interventions for mental health: a scoping review. PLoS One 2021 Mar 11;16(3):e0248152 [FREE Full text] [doi: 10.1371/journal.pone.0248152] [Medline: 33705457]
37. Devynck F, Rousseau A, Romo L. Does repetitive negative thinking influence alcohol use? A systematic review of the literature. Front Psychol 2019 Jul 03;10:1482 [FREE Full text] [doi: 10.3389/fpsyg.2019.01482] [Medline: 31333536]
38. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. JAMA Pediatr 2017 May 01;171(5):461-469 [FREE Full text] [doi: 10.1001/jamapediatrics.2017.0042] [Medline: 28319239]
39. Zotero: your personal research assistant. Zotero. URL: https://www.zotero.org/ [accessed 2026-01-16]
40. Home page. Rayyan Systems Inc. URL: https://new.rayyan.ai/ [accessed 2025-01-24]
41. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-269 [FREE Full text] [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]
42. Mestdagh M, Verdonck S, Piot M, Niemeijer K, Kilani G, Tuerlinckx F, et al. m-Path: an easy-to-use and highly tailorable platform for ecological momentary assessment and intervention in behavioral research and clinical practice. Front Digit Health 2023 Oct 18;5:1182175 [FREE Full text] [doi: 10.3389/fdgth.2023.1182175] [Medline: 37920867]
43. Fleck L, Wasmus H, Schirmbeck F, de Freitas C, Teixeira R, Sanna L, et al. Digital training for mental health promotion in young people with climate change-related distress: protocol for a feasibility randomized controlled trial. JMIR Res Protoc 2025 Dec 05;14:e77764 [FREE Full text] [doi: 10.2196/77764] [Medline: 41349018]
44. Bell IH, Marx W, Nguyen K, Grace S, Gleeson J, Alvarez-Jimenez M. The effect of psychological treatment on repetitive negative thinking in youth depression and anxiety: a meta-analysis and meta-regression. Psychol Med 2022 Nov 14;53(1):6-16. [doi: 10.1017/s0033291722003373] [Medline: 36373473]
45. Ding X, Wuerth K, Sakakibara B, Schmidt J, Parde N, Holsti L, et al. Understanding mobile health and youth mental health: scoping review. JMIR Mhealth Uhealth 2023 Jun 16;11:e44951 [FREE Full text] [doi: 10.2196/44951] [Medline: 37220197]
46. Pywell J, Vijaykumar S, Dodd A, Coventry L. Barriers to older adults' uptake of mobile-based mental health interventions. Digit Health 2020 Feb 11;6:2055207620905422 [FREE Full text] [doi: 10.1177/2055207620905422] [Medline: 32110429]
47. Poulsen A, Hickie IB, Alam M, Wilson CE, LaMonica HM. Access to mobile health in lower middle-income countries: a review. Health Technol 2025 Mar 01;15(2):333-343. [doi: 10.1007/s12553-025-00948-w]
48. Linardon J, Fuller-Tyszkiewicz M. Attrition and adherence in smartphone-delivered interventions for mental health problems: a systematic and meta-analytic review. J Consult Clin Psychol 2020 Jan;88(1):1-13. [doi: 10.1037/ccp0000459] [Medline: 31697093]
49. Lipschitz JM, Van Boxtel R, Torous J, Firth J, Lebovitz JG, Burdick KE, et al. Digital mental health interventions for depression: scoping review of user engagement. J Med Internet Res 2022 Oct 14;24(10):e39204 [FREE Full text] [doi: 10.2196/39204] [Medline: 36240001]
50. Ng MM, Firth J, Minen M, Torous J. User engagement in mental health apps: a review of measurement, reporting, and validity. Psychiatr Serv 2019 Jul 01;70(7):538-544 [FREE Full text] [doi: 10.1176/appi.ps.201800519] [Medline: 30914003]
51. Roelofs J, Muris P, Huibers M, Peeters F, Arntz A. On the measurement of rumination: a psychometric evaluation of the ruminative response scale and the rumination on sadness scale in undergraduates. J Behav Ther Exp Psychiatry 2006 Dec;37(4):299-313. [doi: 10.1016/j.jbtep.2006.03.002] [Medline: 16678121]
52. Meyer TJ, Miller ML, Metzger RL, Borkovec TD. Development and validation of the Penn State Worry Questionnaire. Behav Res Ther 1990 Jan;28(6):487-495. [doi: 10.1016/0005-7967(90)90135-6] [Medline: 2076086]
53. Papanikolaou PN, Churchill R, Wahlbeck K, Ioannidis JP. Safety reporting in randomized trials of mental health interventions. Am J Psychiatry 2004 Sep 01;161(9):1692-1697. [doi: 10.1176/appi.ajp.161.9.1692] [Medline: 15337661]

Abbreviations

CBT: cognitive behavioral therapy
EMI: ecological momentary intervention
GAIDeT: Generative AI Delegation Taxonomy
JBI: Joanna Briggs Institute

XSL•FO
RenderX

mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic reviews and Meta-Analyses
PRISMA-ScR: Preferred Reporting Item for Systematic Reviews and Meta-Analyses for Scoping Reviews
RNT: repetitive negative thinking

Edited by A Schwartz; submitted 04.Feb.2025; peer-reviewed by ME Heidari; comments to author 07.Nov.2025; revised version received 19.Dec.2025; accepted 05.Jan.2026; published 23.Jan.2026.

Please cite as:
Martens J, Forkmann T, Höller I
Characteristics of Mobile Health Interventions for Repetitive Negative Thinking: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e72138
URL: https://www.researchprotocols.org/2026/1/e72138
doi:10.2196/72138
PMID:41575779

©Judith Martens, Thomas Forkmann, Inken Höller. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Determinants of Equitable Data Governance for African, Caribbean, and Black Communities in Health Research in High-Income Countries: Protocol for a Scoping Review

Josephine Etowa[1], PhD; Shamara Baidoobonso[2], PhD; Doris Kakuru[3], PhD; OmiSoore Dryden[4], PhD; Chinedu Oraka[5], MHA, MD; Egbe Etowa[6], PhD; Peter Farrell[1], BEng, MI; Sandra Mba[1], MPH; Akalewold Gebremeskel[1], PhD; Arone Wondwossen Fantaye[1], MSc

[1]School of Nursing, Faculty of Health Sciences, University of Ottawa, Ottawa, ON, Canada

[2]Chief Public Health Office, Government of Prince Edward Island, Charlottetown, PE, Canada

[3]School of Child and Youth Care, Faculty of Education, University of Victoria, Victoria, BC, Canada

[4]Faculty of Medicine, Dalhousie University, Halifax, NS, Canada

[5]Canadians of African Descent Health Organization (CADHO), Ottawa, ON, Canada

[6]School of Nursing, Toronto Metropolitan University, Toronto, ON, Canada

Corresponding Author:
Josephine Etowa, PhD
School of Nursing
Faculty of Health Sciences
University of Ottawa
200 Lees Avenue
Ottawa, ON, K1N 6N5
Canada
Phone: 1 613 323 1678
Fax: 1 613 562 5443
Email: josephine.etowa@uottawa.ca

Abstract

Background: African, Caribbean, and Black (ACB) communities in high-income countries continue to experience persistent health inequities, driven by systemic anti-Black racism, socioeconomic disadvantage, and exclusion from health decision-making. Historically, data have been extracted from ACB communities without transparency, accountability, or community ownership. These inequitable practices have produced data systems that reinforce harm rather than promote equity. Equitable data governance, which promotes community ownership over data collection, access, and use, is increasingly recognized as a critical but underresearched determinant of health equity.

Objective: This protocol outlines the methodology of a scoping review to identify and synthesize evidence on the determinants of equitable data governance in health research involving ACB communities in high-income countries.

Methods: The review follows the 6-stage Arksey and O'Malley methodological framework, supplemented with updated guidance from the Joanna Briggs Institute. The searches were conducted in the Ovid MEDLINE, Ovid Embase, EBSCO CINAHL, APA PsycInfo, and Scopus databases. Peer-reviewed articles are considered, with no limits placed on study design, publication type, or date. Multiple reviewers will independently extract data by using a standardized form. A 3-phase thematic mapping process, conceptually informed by critical race theory, intersectionality, and community-based participatory research principles, will be conducted to analyze the data, generate themes, and interpret findings.

Results: The final comprehensive database searches were completed on December 17, 2024. The search strategy targeted literature on data management, governance, sharing, security, and ethical principles in relation to ACB populations in high-income countries. A total of 4365 records were screened at the title and abstract level, after deduplication, of which 247 studies were deemed potentially relevant and advanced to full-text screening. Following full-text screening and reference list searching a total of 15 articles were deemed eligible for analysis. The data extraction stage is scheduled to overlap and occur between November 2025 and February 2026. The thematic mapping and stakeholder consultations processes are scheduled between December 2025

XSL•FO
RenderX

and February 2026. The final review and manuscript submission are expected by March 2026, with dissemination activities planned for mid-2026.

Conclusions: This review will synthesize existing information on key pillars, barriers, facilitators, promising data governance policies and practices, and recommendations relevant to ACB communities. The findings may inform the expansion of Ontario's Engagement, Governance, Access, and Protection guidelines and support tailored research and national data governance frameworks. The review is expected to contribute to policy, research, and community-led data initiatives. Dissemination will occur through academic publications, conferences, and community-based knowledge-sharing events. As the review relies solely on publicly available data, ethics approval is not required.

Trial Registration: OSF Registries 10.17605/OSF.IO/Z82AY; https://osf.io/z82ay

International Registered Report Identifier (IRRID): PRR1-10.2196/82403

(JMIR Res Protoc 2026;15:e82403) doi:10.2196/82403

KEYWORDS

African; Caribbean; Black community; data governance; data ownership; scoping review protocol

Introduction

Persistent health inequities in high-income countries (HICs) disproportionately affect African, Caribbean, and Black (ACB) communities, contributing to poorer health outcomes among the population despite positive health behaviors [1,2]. These inequities are notable in the prevalence of chronic conditions such as diabetes, HIV, and most recently COVID-19 [3-6]. The literature indicates that health outcomes are driven by a range of upstream factors, such as historical injustices, socioeconomic disadvantage, and systemic anti-Black racism [7,8]. These inequities stem from historical and structural injustices, including slavery, colonization, and intersecting oppressions, such as sexism, classism, and xenophobia, that mold the lived experiences of ACB communities and perpetuate racial health disparities. Anti-Black racism is rooted in societal structures and contributes to systemic inequalities across sectors such as education, employment, housing, health care, and criminal justice [9]. Its cumulative effects, which include microaggressions and institutional discrimination, have a significant impact on health access, uptake, and outcomes.

Awareness of racial inequities in health has grown, but efforts to address them are often impeded by inequitable data systems, policies, and practices [10,11]. Data shape all aspects of contemporary existence, informing policy development and determining what the world around us looks like [12-14]. The availability, collection, access, and use of health data can either remove or reinforce structural barriers and health inequities in communities. Historically, data have been extracted, manipulated, and deployed to justify and bolster slavery, violence, and anti-Black racism [15-17]. ACB communities have also historically been excluded from decisions about how their health data are collected, protected, interpreted, and used [16,17]. Such inequitable data governance practices persist today, as data continue to be extracted from ACB communities without transparency, accountability, or adequate community engagement [18,19]. In addition, ACB research leaders and ACB-led organizations remain underrepresented in governance structures and decision-making processes that address ACB community concerns [1,20]. It is more common for ACB communities to be engaged primarily as participants or community partners, rather than being positioned as leaders of research within their own communities.

There is growing acknowledgment that ACB communities, ACB-led organizations, and scholars need to be involved in efforts to address health issues that directly affect them [21]. Empowering and involving ACB communities and ACB-led organizations in data systems can nurture trust, strengthen community data ownership, and enhance access to community cultural knowledge and insights that can inform and enrich research [22]. In addition, policies and programs intended to address health issues will be better informed and aligned with the needs and priorities of the communities themselves. The cultivation of community-oriented and equitable data governance is thereby imperative for facilitating equitable data practices that protect against misuse and support self-determined health solutions.

The COVID-19 pandemic and the rise of digital health technologies have intensified attention to the topic of data governance. Increasingly, data governance approaches centering on marginalized populations in health research are recognized as mechanisms for the translation of data into targeted interventions and medical technologies that can be used by both community members and researchers to drive positive health outcomes. The process can bring about transformation, equity, and social justice within health care systems [23]. HICs have demonstrated this growing momentum through the development of inclusive, transparent, and community-led data models and frameworks. For instance, in Canada, the First Nations Information Governance Centre (FNIGC) developed the Ownership, Control, Access, and Possession (OCAP) principles as an expression of self-determination in research [24]. The FNIGC developed strategic partnerships for unique data-gathering initiatives and built data and statistical capacities at national, regional, and community levels for the provision of credible and relevant information on First Nations.

The development of OCAP inspired the creation of the Engagement, Governance, Access, and Protection (EGAP) framework, focused on data management and community governance to address systemic racism and advance health equity in ACB communities [16]. On the basis of the learnings from OCAP, EGAP's guidelines envision ACB communities

XSL•FO
RenderX

gaining control over their collective data through community governance tables, which are decision-making bodies that approve or reject plans for community engagement, data collection, and analysis [25]. Similar calls for community-centered governance have emerged internationally. A recent US study advocating for data governance among Black patients with breast cancer pointed to the urgent need for them to have better access to health data, genetic counseling, education, and inclusion in research [23]. In the United Kingdom, key health care leaders involved in the UK-REACH study acknowledged the importance of minority groups having access to health research data to drive better outcomes [14]. Specifically, they identified a need for "Big Data Ethics by Design" to promote community engagement within health research and in turn drive scientific breakthroughs that can be translated into medical innovation and effective public health intervention [14]. Collectively, these examples illustrate how community-led governance can transform data from a tool of harm into a mechanism for equity and justice.

Despite conceptual arguments and promising recent developments in data governance, there remains a lack of synthesized knowledge on what constitutes equitable data governance for ACB communities in health contexts in HICs. This protocol outlines the methodology for a scoping review that addresses this gap by exploring and synthesizing the evidence on the determinants, components, and promising policies and practices that support equity in data governance for ACB communities in HICs. The review is timely and necessary and will inform future research, policy, and practice to support racial health equity through inclusive, ethical, and community-centered data governance. The evidence from this review can help lay the groundwork for improved health research outcomes and promote health equity in ACB communities.

Methods

Approach

This protocol was registered to the Open Science Framework registry. The proposed scoping review is being conducted according to the 6-stage scoping review methodological framework by Arksey and O'Malley [26], which includes (1) identifying the research questions; (2) identifying relevant studies; (3) selecting the studies; (4) charting the data; (5) collating, summarizing, and reporting the results; and (6) consulting experts and stakeholders. Guidance from the Joanna Briggs Institute (JBI) *Manual for Evidence Synthesis* is being integrated to enhance methodological rigor and address known limitations of the earlier scoping review practices [27-29]. Specifically, the reviewers draw on JBI's guidance for the development and piloting of the data extraction tool (stage 4), the approach to the synthesis and presentation of the results (stage 5), and incorporation of stakeholder consultation as an iterative and reflexive process (stage 6). This combined approach enables us to adhere to foundational scoping review methodology and align with current best practices in evidence synthesis. This protocol is reported in accordance with the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols) guidelines.

Step 1: Identifying the Research Question

In determining the research questions, this scoping review used the population, concept, and context framework to fully examine the scope and gaps within the current literature. Textbox 1 provides key definitions of the population (ACB), concept (data governance), and context (HIC) selections.

Textbox 1. Definitions of key concepts used in this scoping review.

- African, Caribbean, and Black: refers to individuals and communities who identify as Black and trace their heritage to sub-Saharan Africa, the Caribbean, or the African diaspora globally. We use this term inclusively of Black people from diverse cultural, ethnic, and national backgrounds.
- Data governance: refers to data systems, structures, policies, and practices that determine who has influence and control over the agenda, data collection, ownership, interpretation, and application of health data [16,17]
- High-income countries: listed using the 2023 World Bank classification and Organization for Economic Co-operation and Development country membership [30]

Textbox 2 details the primary research question and secondary research questions. For secondary research question 1, *promising* policy and practices are conceptualized as innovative policies and practices that show potential to improve equity in data governance, even if not yet widely implemented, adopted, or evaluated [31,32]. For secondary research question 2, "effective and comprehensive" are being conceptualized as tools, frameworks, or practices that demonstrate evidence of improvement in health equity outcomes, legal and ethical safeguards, community trust and uptake, mechanisms for community engagement, and increased ACB community power over data structures and processes [33]. These components are considered comprehensive when they address multiple levels (eg, policy, organizational, and community) or dimensions of governance, such as ownership, access, protection, use, and accountability.

XSL•FO
RenderX

Textbox 2. Research questions.

Primary research question

- What are the determinants of equitable data governance (concept) for African, Caribbean, and Black (ACB) communities (population) in health contexts in high-income countries (context)?

Secondary research questions

- What are promising policies and practices to address equity in data governance for ACB communities and ACB-led organizations in high-income countries? (secondary research question 1)
- What are the components of effective and comprehensive data governance frameworks to strengthen ACB communities' health in high-income countries? (secondary research question 2)

Step 2: Identifying Relevant Studies

A highly sensitive multidatabase search strategy was developed, using controlled vocabulary and text words to identify peer-reviewed literature that touched on data governance in relation to research on or with diasporic Black populations with African or Caribbean roots in HICs. The following databases were searched by a health sciences librarian (PF): MEDLINE All via Ovid, Embase + Embase Classic via Ovid, APA PsycInfo via Ovid, CINAHL via EBSCOhost, and Scopus. A search strategy was developed in the MEDLINE database and then translated into the other databases, as appropriate (Multimedia Appendix 1). The MEDLINE database search strategy was peer-reviewed using the Peer Review of Electronic Search Strategies tool by another librarian (Marie-Cécile Domecq). All databases were searched from inception to December 17, 2024. The search will be updated before reporting the results. Following the screening of database results, the researchers will conduct forward and backward citation searching on the included articles.

To increase the breadth of the scoping review, source evidence was open to all peer-reviewed publication types, including primary studies, conceptual and opinion papers, secondary studies, reports, etc. Only English and French evidence sources were considered due to feasibility reasons, as they are the primary languages of the researchers. In addition, only evidence sources with a primary focus on ACB populations were considered so that issues specific to this target population are addressed.

Step 3: Study Selection

The research team seeks to identify information sources that examine the factors that contribute to equitable data governance for ACB communities. Sources reporting on effective, comprehensive, or promising practices, policies, tools, and frameworks will also be considered. A summary of the inclusion and exclusion criteria for the primary and secondary research questions is detailed in Tables 1 and 2, respectively. In cases where articles focus on broader racialized or visible minority populations, inclusion will be contingent on the clear identification and description of ACB representation and experiences.

Source selection includes two stages: (1) title and abstract screening and (2) full-text screening. Both stages will use the software program Covidence, which will automatically remove duplicate articles. Four reviewers (AWF and volunteer students) are independently screening titles and abstracts, with disagreements being resolved through discussion with the senior author (JE). Full texts of potentially relevant studies will be independently screened by the same four reviewers. Disagreements between 2 full-text screeners will be resolved by a third screener in Covidence or by discussion with the senior author (JE) if needed by the third screener. The number of excluded articles and the rationale for exclusion will be recorded in the PRISMA-P scoping review flowchart.

XSL•FO
RenderX

Table 1. Eligibility criteria for the primary research question.

	Inclusion criteria	Exclusion criteria
Population	Includes any person or an individual, organizations, community, or stakeholders that are identified as ACB[a] or related terms, and conduct research among the ACB population. Articles will be considered if they report on determinants of equitable data governance in racialized communities, if the ACB community is clearly mentioned and described	Articles that do not include ACB populations, clients, or patients
Concept	The concept of data governance between researchers, agencies, and other data governing bodies with ACB communities or ACB-led organizations. This can encompass data ownership, control, and access to improve health outcomes, as well as research partnerships, collaborations, and knowledge mobilization practices	Articles that are not relevant to data governance
Context	High-income countries listed in the 2023 World Bank classification and Organization for Economic Co-operation and Development country membership [30]	Articles that are not from a high-income country context (eg, articles from a middle- or low-income country)
Source type	Open	Full-text unavailable
Language	English and French	Full-text is only available in languages other than English and French.
Publication date	No date restrictions	—[b]

[a]ACB: African, Caribbean, and Black.

[b]Not applicable.

Table 2. Eligibility criteria for secondary research questions.

	Inclusion criteria	Exclusion criteria
Population	ACB[a] communities and ACB-led organizations	Articles that do not include ACB populations, clients, or patients
Concept	Data governance policies and practices. These may include promising policy and practices that show potential to improve equity in data governance, even if not yet widely implemented, adopted, or evaluated. Effective and comprehensive data governance frameworks between researchers and ACB communities. May include tools, frameworks, or practices that demonstrate evidence of improvement in health equity outcomes, legal and ethical safeguards, community trust and uptake, mechanisms for community engagement, and increased ACB community power over data structures and processes	Articles that are not related to data governance policies and practices
Context	High-income countries listed in the 2023 World Bank classification and Organization for Economic Co-operation and Development country membership [30]	Articles that are not from a high-income country context (eg, articles from a middle- or low-income country)
Source type	Open	Abstracts, full-text unavailable
Language	English and French	Full-text is only available in languages other than English and French
Publication date	No date restrictions	—[b]

[a]ACB: African, Caribbean, and Black.

[b]Not applicable.

Step 4: Charting Data

Data will be extracted from papers included in the scoping review independently by 4 coders (AWF, a senior reviewer, and 2 volunteer students) using a data extraction tool developed by the reviewers. The data will include specific details about the population, concept, context, study methods, and key findings relevant to the scoping review objectives. A draft of the extraction table is provided (Multimedia Appendix 2). This will be modified and revised as necessary during the extraction of the included studies. Modifications will be detailed in the full scoping review report.

To calibrate the table, all 4 coders will independently extract data from 3 evidence sources, which will subsequently be compared. Creating the table is an iterative process, and modifications may be made through the same process as more evidence emerges [34]. Any discrepancies between the coders will be discussed until at least a 75% agreement is achieved. To ensure consistency and catch discrepancies early on, the 4 coders will then independently extract from a subset of the remaining included articles (at least 25%). Two recalibration meetings will be held with the senior author (JE) after extracting approximately a quarter and then half of the articles to review and discuss extractions. Disagreements during subset extraction

XSL•FO
RenderX

or recalibration meetings will be resolved through discussion during these meetings, with oversight by the senior author (JE).

The current template for this table contains (1) first author; (2) date; (3) country of origin; (4) population; (5) organizational resources, infrastructure, knowledge, and skills; (6) culture and climate; and (7) engagement and partnership, study design, sample size, findings, and conclusion. Given that several different types of evidence sources are anticipated, it is not expected that all categories in the general extraction table will be completed for each evidence source; not applicable will be written, where appropriate. However, the extracted data from each evidence source will align with the research question and objectives based on reviewing its entire text. Microsoft Word will be used during the extraction process to record and organize the data. The extracted information will be used to inform the collation and summarization of the findings from the evidence sources.

Step 5: Collating, Summarizing, and Reporting Results

Overview

The data extracted from relevant published literature will be presented in a tabular form that is aligned with the objectives of this scoping review. Data that are presented in tables will reflect the information collected using the data extraction tool. A narrative summary will accompany the tabulated and charted results and will describe how the results relate to the review objectives.

Thematic mapping (TM) [35] will be used to collate and summarize the extracted data, as it is ideal for accommodating the multiple types of evidence sources in this study (Figure 1) [21]. In addition, TM allows for the production of themes while maintaining the breadth of information within each theme through the creation of subthemes (descriptive and analytic themes). This approach is consistent with other methodological approaches from Arksey and O'Malley [26] and Peters et al [27,28] who acknowledge that results should be descriptively mapped rather than synthesized.

Figure 1. Thematic mapping process.

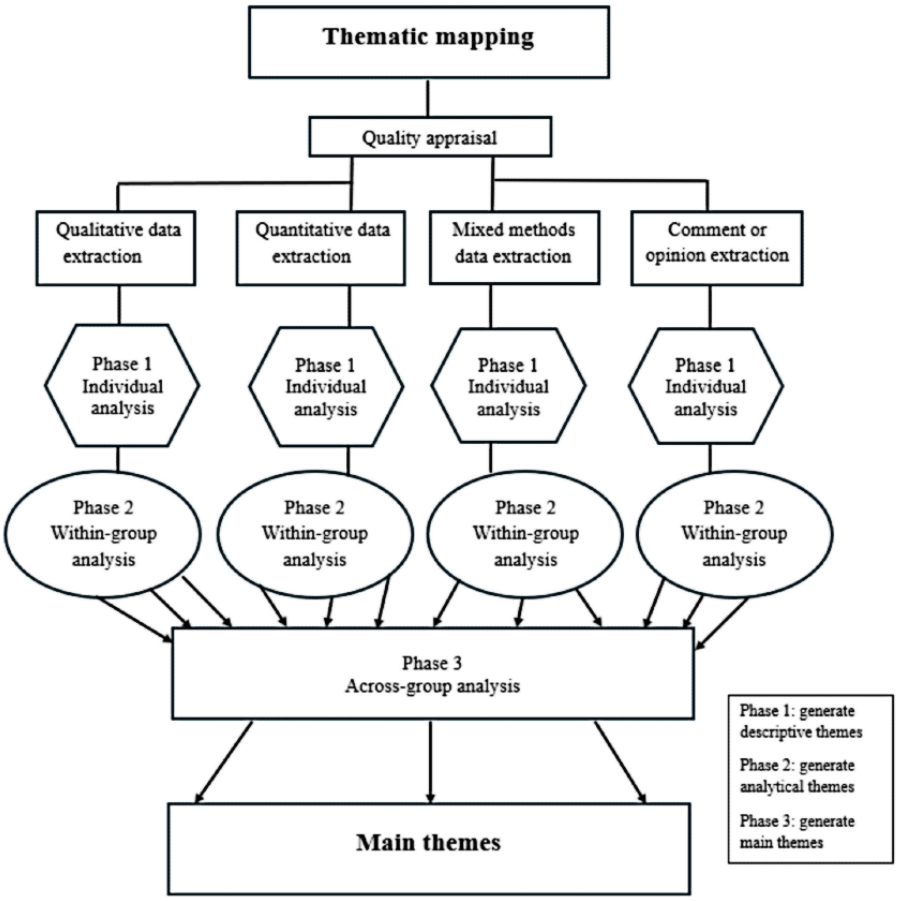

TM Phase 1: Individual Study Analysis

The aim of phase 1 will be to create initial codes and descriptive themes, informed by 3 conceptual underpinnings: critical race theory, intersectionality, and community-based participatory research. The critical race theory and intersectionality lenses will specifically be used as sensitizing concepts to sensitize the coders to the power dynamics, systemic inequities, and overlapping marginalization in data governance structures and processes. Community-based participatory research will also be used as a sensitizing concept to pull out the presence (or lack thereof) and consideration of community voice and priorities in the data. Together, they facilitate consideration of the cause and persistence of systemic racism and discrimination in the structure and processes through which data are collected, managed, analyzed, and used.

XSL•FO
RenderX

TM Phase 2: Within-Group Analysis

In phase 2, analytic themes will be created to gain further insight into the characteristics of the descriptive themes that reflected the content within each type of article grouping (qualitative, quantitative, mixed method, and opinion or conceptual) separately. The descriptive themes from each article will be grouped based on similarities and differences into analytic themes through induction and consistent interpretation with the population, concept, and context research questions, and the theoretical underpinnings, as in phase 1. The iterative process of creating the final analytic themes will be concluded when a consensus is reached between the research team members.

TM Phase 3: Across-Group Analysis

Main themes will be created to give a broad overview of findings based on the similarities and differences across all groups. These main themes will be used to map the evidence on concepts and boundaries, help identify research gaps, and determine existing strategies to dismantle the barriers causing ACB-led organizations to genuinely conduct and lead research. The themes that will be developed will be recorded in the thematic map.

Step 6: Consultation

Scoping review results will be brought to the African, Caribbean, and Black National Expert Working Group (ACB-NEWG) in Canada for collaboration and input. ACB-NEWG is comprised of approximately 20 diverse stakeholder groups, including Black representatives from community-based organizations, Black community–engaged researchers, and Black representatives from mainstream organizations (eg, Public Health Agency of Canada), as well as coapplicants and collaborators. The ACB-NEWG will serve as the Project Advisory Committee for this project. Therefore, the executive group will review the findings and preliminary frameworks and provide feedback on evaluation items that may need to be added, subtracted, or altered. Having a broad Steering Committee, project advisory committee, and a network of potential partners, an effective communication strategy, and clear lines of responsibility will help to assure that the governance structure is sustained over the course of the project.

Ethical Considerations

This manuscript does not report on or involve the use of any animal or human data or tissue.

Results

The comprehensive database searches have been completed across MEDLINE, Embase, CINAHL, APA PsycInfo, and Scopus, with the final searches run on December 17, 2024. A total of 6784 records were imported for screening. After removing 2418 duplicates (2414 via Covidence and 4 manually), 4365 records were screened at the title and abstract level. Preliminary observations from the abstract screening suggest that a diverse range of literature reports on determinants, frameworks, and promising policies and practices to empower and facilitate equitable and community-led data governance in racialized contexts, though the relevance and representation to ACB-led governance appear limited (Figure 2).

XSL•FO
RenderX

Figure 2. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart of the identification of articles via databases and reference lists. ACB: African, Caribbean, and Black; HIC: High income country.

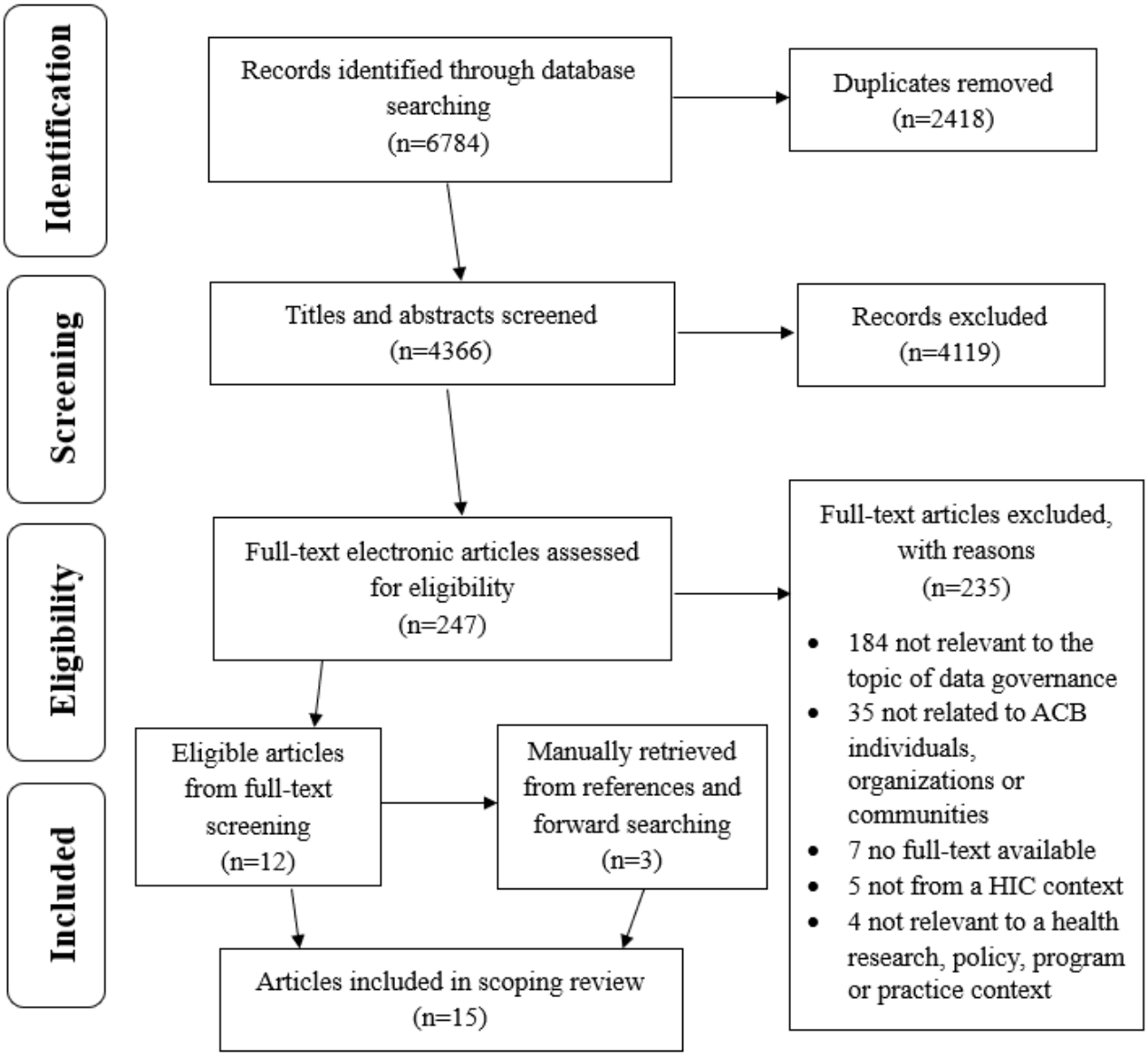

Only 247 records were identified as potentially eligible and have progressed to full-text screening, which assessed each article against the inclusion criteria outlined in Tables 1 and 2. This phase was completed in December 2025, along with manual reference list searching and screening. A total of 12 relevant articles have been retrieved from full-text screening, and 3 additional relevant articles from reference list searching. The data extraction stage has commenced and is scheduled to be completed by mid-February 2026. The TM process and final stakeholder consultations are scheduled to overlap and occur between December 2025 and late-February 2026. The final review and manuscript submission are expected by early March 2026, with dissemination activities planned for mid-2026.

Discussion

Anticipated Findings

This scoping review protocol outlines a rigorous and inclusive methodology for the synthesis of the existing information on equitable data governance in health research involving ACB communities in HICs. The findings will provide a broad perspective from a variety of evidence sources, with a focus on understanding the key factors that contribute to the inclusive, ethical, and participatory approach of data governance for ACB communities and how these contribute to improved outcomes. Moreover, it will also help identify promising practices, policies, tools, and frameworks that contribute to improved health outcomes for ACB communities. Full-text screening has revealed a diverse body of literature, although the representation of ACB-led initiatives remains limited. In multiple articles, ACB communities are represented within broader racialized or visible minority populations. For the purposes of comprehensiveness and relevance, such articles are being included in the review if they clearly identify and describe ACB representation and experiences. This approach allows us to capture insights from multiethnic contexts while maintaining a focused lens on ACB-specific data governance issues.

XSL•FO
RenderX

Frameworks such as OCAP and EGAP have laid the foundational principles for Indigenous and Black data governance in certain pockets of Canada. However, there is limited synthesis of how these principles are operationalized across health research involving ACB communities in HICs. This review will build on existing conceptual work by mapping existing information on determinants, policies, and practices, and identifying opportunities for enhancing data equity and community control over ACB health data. The findings may inform the refinement of EGAP and support the development of a broader and more equity-focused national data governance framework that can be adapted to the Canadian context and beyond, including other HICs where ACB communities face similar data governance, policy, and practice barriers. The review will also offer more practical insights into the policies, practices, and research that are needed to meet the needs of ACB populations, including enhancing their data governance capabilities.

Strengths and Limitations

To limit bias, this review predefines the search question, objectives, methodology, eligibility criteria, search methods, data extraction techniques, summary, and the presentation of anticipated findings. This structured approach enhances transparency and reproducibility. The use of a peer-reviewed, multidatabase search strategy and a structured TM process further enhances methodological rigor and transparency. Another key strength of this review lies in its integration of critical race theory, intersectionality, and community-based participatory research, which provides a theoretically grounded and equity-focused review approach. Finally, stakeholders who are engaged throughout the review will be invited to review preliminary findings and thematic maps to ensure that the interpretations reflect community priorities and lived experiences.

However, multiple practical and operational limitations must be acknowledged. First, the concept of "equitable data governance" is still evolving and may be interpreted differently across studies and contexts. This could lead to difficulties in identifying and categorizing relevant determinants, practices, policies, or recommendations. Second, there may be limited representation of ACB-led data governance initiatives in the peer-reviewed literature, which could restrict the review's ability to capture community-driven models, and existing promising policies and practices. Third, the language restriction to English and French may result in the exclusion of relevant studies published in other languages. Fourth, the inclusion of diverse evidence sources (primary research, reports, opinion, and conceptual papers) that are different in quality, depth, and format may make it difficult to extract consistent data. Despite the limitations, this scoping review will help to examine the evidence available and gauge the current state of the literature on factors influencing data governance for ACB communities and ACB-led organizations.

Future Directions

The findings of this review may inform future empirical research on community-led data governance models, including longitudinal evaluations of their impact on health equity. Additionally, the review may uncover conceptual and methodological gaps, such as the need for standardized definitions, metrics for equity, and frameworks for the evaluation of governance effectiveness across diverse ACB contexts.

Dissemination

The results from the proposed scoping review will be disseminated through peer-reviewed open-access journals and conferences targeting public health stakeholders, vaccination campaigns, and overcoming inequities in health care. In addition to academic dissemination, the reviewers will codevelop and share plain-language summaries of key findings with community partners and stakeholders, including members of the ACB-NEWG. These summaries will be tailored for nonacademic audiences and shared through partner networks, community knowledge-sharing events or workshops (eg, World Café events), and public health forums. Stakeholders engaged during the consultation phase will also advise on other culturally responsive and context-relevant formats for disseminating findings, with the aim of maximizing accessibility, relevance, and potential for policy or practice use within ACB communities.

Acknowledgments

The authors would like to thank all stakeholders, partners, service providers, and study participants who have contributed to this study. They thank Marie-Cécile Domecq (University of Ottawa) for peer review of the MEDLINE search strategy.

Data Availability

This protocol does not report on any datasets. Data generated during the review will be made available upon publication of the final review.

Funding

This project is funded by the Social Sciences and Humanities Research Council (grant 936-2021-00859).

Authors' Contributions

JE (principal investigator) led the design and conceptualization of this review and drafted the protocol with primary support from EE, SB, DK, CO, and AWF. PF developed the search strategy, with review and input from AWF. JE, SM, and PF were involved

XSL•FO
RenderX

in establishing the eligibility criteria and data extraction form. All authors provided feedback on the manuscript and approval for the publication of this protocol manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Search strategy and databases.
[PDF File (Adobe PDF File), 164 KB - resprot_v15i1e82403_app1.pdf]

Multimedia Appendix 2
Extraction instrument.
[PDF File (Adobe PDF File), 106 KB - resprot_v15i1e82403_app2.pdf]

References

1. Dabone C, Mbagwu I, Muray M, Ubangha L, Kohoun B, Etowa E, et al. Global food insecurity and African, Caribbean, and Black (ACB) populations during the COVID-19 pandemic: a rapid review. J Racial Ethn Health Disparities 2022 Apr;9(2):420-435 [FREE Full text] [doi: 10.1007/s40615-021-00973-1] [Medline: 33559109]
2. Etowa J, Demeke J, Abrha G, Worku F, Ajiboye W, Beauchamp S, et al. Social determinants of the disproportionately higher rates of COVID-19 infection among African Caribbean and Black (ACB) population: a systematic review protocol. J Public Health Res 2021 Dec 30;11(2):2274 [FREE Full text] [doi: 10.4081/jphr.2021.2274] [Medline: 34967533]
3. Key health inequalities in Canada: a national portrait. Public Health Agency of Canada. 2018 Aug. URL: https://www.canada.ca/en/public-health/services/publications/science-research-data/key-health-inequalities-canada-national-portrait-executive-summary.html [accessed 2021-03-14]
4. Haddad N, Li JS, Totten S, McGuire M. HIV in Canada-surveillance report, 2017. Can Commun Dis Rep 2018 Dec 06;44(12):348-356 [FREE Full text] [doi: 10.14745/ccdr.v44i12a03] [Medline: 30906230]
5. Choi K, Denice P. Socioeconomic variation in the relationship between neighbourhoods' built environments and the spread of COVID-19in Toronto, Canada. Can Stud Popul 2022;49(3-4):149-181 [FREE Full text] [doi: 10.1007/s42650-022-00070-6] [Medline: 36068823]
6. Subedi R, Greenberg L, Turcotte M. COVID-19 mortality rates in Canada's ethno-cultural neighbourhoods. Statistics Canada. 2020 Oct 28. URL: https://www150.statcan.gc.ca/n1/pub/45-28-0001/2020001/article/00079-eng.htm [accessed 2021-02-13]
7. Churchwell K, Elkind MS, Benjamin RM, Carson AP, Chang EK, Lawrence W, et al. Call to action: structural racism as a fundamental driver of health disparities: a presidential advisory from the American Heart Association. Circulation 2020 Dec 15;142(24):e454-e468 [FREE Full text] [doi: 10.1161/CIR.0000000000000936] [Medline: 33170755]
8. Rakotovao L, Simeoni M, Bennett-AbuAyyash C, Walji T, Abdi S. Addressing anti-Black racism within public health in North America: a scoping review. Int J Equity Health 2024 Jun 28;23(1):128 [FREE Full text] [doi: 10.1186/s12939-024-02124-4] [Medline: 38937746]
9. Dryden O, Nnorom O. Time to dismantle systemic anti-Black racism in medicine in Canada. CMAJ 2021 Jan 11;193(2):E55-E57 [FREE Full text] [doi: 10.1503/cmaj.201579] [Medline: 33431548]
10. Islam M, Valiani AA, Datta R, Chowdhury M, Turin TC. Ethical and equitable digital health research: ensuring self-determination in data governance for racialized communities. Camb Q Healthc Ethics 2024 Apr 03:1-11. [doi: 10.1017/S096318012400015X] [Medline: 38567458]
11. Maddalena V. An exploration of policy options to assist district health authorities in attending to the health needs of African Canadians: a case study. Leadersh Health Serv (Bradf Engl) 2010;23(1):57-74 [FREE Full text] [doi: 10.1108/17511871011013779]
12. Kitchin R. The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences. Thousand Oaks, CA: SAGE Publications; 2014.
13. O'Neil S, Taylor S, Sivasankaran A. Data equity to advance health and health equity in low- and middle-income countries: a scoping review. Digit Health 2021;7:20552076211061922 [FREE Full text] [doi: 10.1177/20552076211061922] [Medline: 34992789]
14. Dove ES, Reed-Berendt R, Pareek M, UK-REACH Study Collaborative Group. "Data makes the story come to life:" understanding the ethical and legal implications of Big Data research involving ethnic minority healthcare workers in the United Kingdom-a qualitative study. BMC Med Ethics 2022 Dec 16;23(1):136 [FREE Full text] [doi: 10.1186/s12910-022-00875-9] [Medline: 36527096]
15. Eisape A, Nogueira A. See change: overcoming anti-Black racism in health systems. Front Public Health 2022;10:895684 [FREE Full text] [doi: 10.3389/fpubh.2022.895684] [Medline: 35784218]

XSL•FO
RenderX

16. Engagement, Governance, Access, and Protection (EGAP): a data governance framework for health data collected from Black communities in Ontario. Alliance for Healthier Communities. 2021. URL: https://www.allianceon.org/research/Engagement-Governance-Access-and-Protection-EGAP-Data-Governance-Framework-Health-Data?language=en [accessed 2021-10-30]
17. Demeke J, Worku F, Sobers M, Roche B, McKenzie K. City of Toronto Black data governance: literature review. Wellesley Institute. 2023. URL: https://www.wellesleyinstitute.com/wp-content/uploads/2024/05/City-of-Toronto-Black-Community-Data-Governance-Literature-Review-Final.pdf [accessed 2025-03-05]
18. Grasser LR, Jovanovic T. Neural impacts of stigma, racism, and discrimination. Biol Psychiatry Cogn Neurosci Neuroimaging 2022 Dec;7(12):1225-1234 [FREE Full text] [doi: 10.1016/j.bpsc.2022.06.012] [Medline: 35811064]
19. Ogunyemi D, Clare C, Astudillo YM, Marseille M, Manu E, Kim S. Microaggressions in the learning environment: a systematic review. J Divers High Educ 2020 Jun;13(2):97-119 [FREE Full text] [doi: 10.1037/dhe0000107]
20. Aden H, Oraka C, Russell K. Mental health of Ottawa's Black community research study. Ottawa Public Health. 2020 Aug. URL: https://www.ottawapublichealth.ca/en/reports-research-and-statistics/resources/Documents/MHOBC_Technical-Report_English.pdf [accessed 2025-03-10]
21. Etowa J, Abrha G, Etowa E, Ghose B. Healthcare during COVID-19 in Canada: need for strengthening providers' capacity for best practices in African, Caribbean and Black community service provision. Am J Public Health Res 2021;9(2):52-56 [FREE Full text] [doi: 10.12691/ajphr-9-2-2]
22. Gahagan J. Women and HIV Prevention in Canada: Implications for Research, Policy, and Practice. Toronto, ON: Canadian Scholars; 2013.
23. Gerido LH, Resnicow K, Stoffel EM, Tomlin T, Cook-Deegan R, Cline M, et al. Big advocacy, little recognition: the hidden work of Black patients in precision medicine. J Community Genet 2023 Dec;14(6):657-665 [FREE Full text] [doi: 10.1007/s12687-023-00673-9] [Medline: 37775604]
24. Ownership, control, access and possession (OCAP™): the path to first nations information governance. First Nations Information Governance Centre. 2014 May. URL: https://fnigc.ca/wp-content/uploads/2020/09/5776c4ee9387f966e6771aa93a04f389_ocap_path_to_fn_information_governance_en_final.pdf [accessed 2021-10-27]
25. Demeke J, Roche B, Orridge C, McKenzie K. DIAGRAM: engagement, governance, access and protection (EGAP) implementation guide. Wellesley Institute. Toronto: Wellesley Institute; 2023 Aug. URL: https://www.wellesleyinstitute.com/wp-content/uploads/2024/07/DIAGRAM-Engagement-Governance-Access-and-Protection-EGAP-Implementation-Guide.pdf [accessed 2024-04-04]
26. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
27. Peters MD, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
28. Peters MD, Godfrey C, McInerney P, Khalil H, Larsen P, Marnie C, et al. Best practice guidance and reporting items for the development of scoping review protocols. JBI Evid Synth 2022 Apr 01;20(4):953-968. [doi: 10.11124/JBIES-21-00242] [Medline: 35102103]
29. Pollock D, Alexander L, Munn Z, Peters MD, Khalil H, Godfrey CM, et al. Moving from consultation to co-creation with knowledge users in scoping reviews: guidance from the JBI Scoping Review Methodology Group. JBI Evid Synth 2022 Apr 01;20(4):969-979. [doi: 10.11124/JBIES-21-00416] [Medline: 35477565]
30. World bank open data. World Bank Group. URL: https://data.worldbank.org [accessed 2024-11-20]
31. Kezar A, Sam C. Strategies for implementing and institutionalizing new policies and practices: understanding the change process. In: Kezar A, editor. Embracing Non-Tenure Track Faculty. New York, NY: Routledge; 2012.
32. Gleason PM. What's the secret ingredient? Searching for policies and practices that make charter schools successful. J Sch Choice 2017 Nov 27;11(4):559-584 [FREE Full text] [doi: 10.1080/15582159.2017.1395620]
33. Hosman CM, Clayton R. Prevention and health promotion on the international scene: the need for a more effective and comprehensive approach. Addict Behav 2000;25(6):943-954. [doi: 10.1016/s0306-4603(00)00127-1] [Medline: 11125781]
34. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
35. Etowa J, Beauchamp S, Fseifes M, Osandatuwa G, Brenneman P, Salam-Alada K, et al. Understanding low vaccine uptake in the context of public health in high-income countries: a scoping review. Vaccines (Basel) 2024 Mar 04;12(3):269 [FREE Full text] [doi: 10.3390/vaccines12030269] [Medline: 38543903]

Abbreviations

ACB: African, Caribbean, and Black
ACB-NEWG: African, Caribbean, and Black National Expert Working Group
EGAP: Engagement, Governance, Access, and Protection
FNIGC: First Nations Information Governance Centre

XSL•FO
RenderX

HIC: high-income country
JBI: Joanna Briggs Institute
OCAP: Ownership, Control, Access, and Possession
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols
TM: thematic mapping

Edited by J Sarvestan; submitted 14.Aug.2025; peer-reviewed by S Mohanadas; comments to author 01.Dec.2025; revised version received 20.Dec.2025; accepted 22.Dec.2025; published 13.Feb.2026.

Please cite as:
Etowa J, Baidoobonso S, Kakuru D, Dryden O, Oraka C, Etowa E, Farrell P, Mba S, Gebremeskel A, Fantaye AW
Determinants of Equitable Data Governance for African, Caribbean, and Black Communities in Health Research in High-Income Countries: Protocol for a Scoping Review
JMIR Res Protoc 2026;15:e82403
URL: https://www.researchprotocols.org/2026/1/e82403
doi:10.2196/82403
PMID:

©Josephine Etowa, Shamara Baidoobonso, Doris Kakuru, OmiSoore Dryden, Chinedu Oraka, Egbe Etowa, Peter Farrell, Sandra Mba, Akalewold Gebremeskel, Arone Wondwossen Fantaye. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 13.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

The Use and Utility of Wearable Devices for Tracking Sleep and Activity in Inpatient Mental Health Settings: Protocol for a Rapid Review

Gillian Strudwick[1]; Hwayeon Danielle Shin[1]; Iman Kassam[1]; Karishini Ramamoorthi[2]; Benjamin Rosen[3]; Michelle A Patriquin[4]; Justin Baker[5]; Nicholas Neufeld[1]

[1]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada

[2]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[3]Sinai Health, Toronto, ON, Canada

[4]John S. Dunn Behavioral Sciences Center, The University of Texas Health Science Center at Houston, Houston, TX, United States

[5]Institute for Technology in Psychiatry, McLean Hospital, Belmont, MA, United States

Corresponding Author:
Gillian Strudwick
Campbell Family Mental Health Research Institute
Centre for Addiction and Mental Health
1001 Queen St W
Toronto, ON, M6J 1H1
Canada
Phone: 1 4165358501
Email: gillian.strudwick@camh.ca

Abstract

Background: Wearable devices offer an opportunity to objectively monitor and capture sleep and activity in psychiatric inpatient settings, where traditional approaches often rely on subjective reporting or staff observation, both of which have inherent flaws. Technologies such as Fitbit, Garmin, Oura Ring, GENEActiv, Empatica, and WHOOP (among others) have been used to passively collect physiological data that may inform care planning and clinical decision-making. Despite growing interest, the extent to which these wearables are feasible, acceptable, and useful in inpatient mental health settings remains unclear. Synthesizing this evidence is essential to inform their potential integration into psychiatric care.

Objective: This study aims to conduct a rapid review to examine the current use, utility, feasibility, and implementation of wearable devices for tracking sleep and/or activity in inpatient mental health settings. To date, no review has examined inpatient settings specifically.

Methods: A rapid review will be conducted using Cochrane Rapid Review methods. Peer-reviewed literature will be searched in four databases: (1) PubMed, (2) Embase, (3) PsycInfo, and (4) CINAHL. Studies will be included if they report on the use of wearable devices in inpatient psychiatric settings and evaluate outcomes related to feasibility, acceptability, clinical utility, or implementation. Data extraction will be conducted using a standardized extraction template, and findings will be synthesized narratively and via descriptive statistics.

Results: Database searches are currently underway. The review is expected to be completed by June 2026. A manuscript detailing the findings of this review will be published thereafter.

Conclusions: This rapid review will provide timely evidence to support the integration of wearable technologies into psychiatric inpatient settings. Findings will inform future research and clinical practice focused on the use of wearables for sleep and activity monitoring, contributing to more responsive and data-informed mental health care.

International Registered Report Identifier (IRRID): PRR1-10.2196/82290

(JMIR Res Protoc 2026;15:e82290) doi:10.2196/82290

KEYWORDS

mental health; wearables; actigraphy; sleep; physical activity; rapid review

XSL•FO
RenderX

Introduction

Accurate monitoring of sleep and physical activity in psychiatric inpatient settings is essential for understanding patients' mental health status and treatment response [1]. Sleep disturbances such as insomnia, hypersomnia, and fragmented sleep are highly prevalent across psychiatric conditions such as depression, bipolar disorder, schizophrenia, and anxiety, serving both as core features of these disorders and potential triggers for relapse [2]. Poor sleep quality in psychiatric inpatients has been linked to worse clinical outcomes, including greater emotion regulation difficulties, elevated depression and anxiety, increased suicidal ideation, and more severe nightmares [3-5]. Similarly, low levels of physical activity are associated with poorer mental health, whereas increased activity can enhance emotional regulation and cognitive function and reduce depressive and anxiety symptoms [6]. As both sleep and activity may signal shifts in mental state, objective and continuous assessment of these domains is important in inpatient settings where patients may lack self-awareness or the capacity to reliably self-report, enabling timely, data-driven clinical interventions.

Traditional methods of assessing or monitoring sleep and activity, such as self-report or staff observation (often by nurses), are often subjective and imprecise [7-9]. These approaches can be limited by recall bias, inconsistency across observers, and challenges in obtaining continuous or real-time data [10]. In settings where clinical acuity is high and patients may be experiencing cognitive or emotional distress, reliable and objective data are required for care such as tailoring care plans, identifying early signs of deterioration, and evaluating treatment effectiveness over time [11].

Wearable technologies offer a promising solution [12] by providing passive, continuous, and generally noninvasive monitoring of sleep and physical activity (among other physiological measures) via a wrist-worn device or ring [13,14]. Devices, including but not limited to Fitbit (Google Inc), Garmin (Garmin Ltd), Oura Ring (Oura Health Ltd), GENEActiv (Activinsights Ltd), Empatica (Empatica Inc), and WHOOP (Whoop, Inc) can generate physiological data that may enhance clinical decision-making and have the potential to reduce clinician documentation burden [15]. In addition to supporting personalized interventions and early detection of changes in mental health status, wearables have the potential to empower patients by increasing engagement with and awareness of their health patterns (eg, sleep length, sleep quality, and step count) [14,16]. However, the utility and implementation of these devices in the context of inpatient psychiatry, where safety concerns, patient acuity, organizational policies, and technology infrastructure differ from outpatient or nonpsychiatric settings, are not well established. Understanding how these technologies are currently being used and whether they are feasible, acceptable, and clinically useful in inpatient environments is an essential step toward integrating them effectively into mental health care [17].

Given the increasing interest in digital health tools in psychiatry and the limited evidence specific to inpatient settings, there is a need to understand how wearables are currently being used and evaluated in these environments. Inpatient psychiatric settings present unique challenges such as restricted environments, high patient acuity, unique privacy concerns, increased vulnerability of patients, and at times limited digital infrastructure [18,19]. Despite these challenges, wearables could offer important benefits: passive data collection, reduced burden on clinical staff, and early detection of symptom exacerbation or treatment effect [12,20]. Synthesizing the current literature on this topic is necessary to identify best practices, feasibility, and gaps in the literature that can guide future implementation and research. A rapid review will therefore be conducted with the following research questions (RQs):

- RQ1: How are wearable devices being used to monitor sleep and/or activity in inpatient psychiatric settings?
- RQ2: What is the reported feasibility, acceptability, and clinical utility of using wearable technologies in inpatient mental health care?

This review will differ from previous reviews of wearable technologies in mental health, which have synthesized studies primarily in outpatient or community settings and have focused on identifying associations between passively collected data and psychiatric symptoms or relapse risk. Previous reviews have also been predominantly conducted for research purposes, while this review will include both research purposes and real-world use. This protocol focuses specifically on inpatient mental health settings, where patient acuity, safety requirements, and clinical workflows introduce distinct considerations for technology use.

Methods

This rapid review will be conducted following the methodological guidance provided by the Cochrane Rapid Reviews Group [21] and the approach described by Tricco et al [22,23]. A rapid review design was selected due to its ability to generate timely insights in a field where the evidence base is still emerging [24]. This rapid review will not include a critical appraisal within its scope, as this is not common practice [25,26] and would extend timelines. The process as described balances efficiency with methodological rigor and will take place over a 6-month timeline.

The objective of this review is to explore the use, utility, feasibility, and implementation of wearable devices for tracking sleep and/or activity in inpatient psychiatric settings. The review will address 3 primary RQs described earlier.

Studies will be eligible for inclusion if they are primary empirical research (quantitative, qualitative, or mixed methods); if they are conducted in inpatient psychiatric or mental health settings; and if they include the use of wearable technologies such as Fitbit, Garmin, Oura Ring, GENEActiv, Empatica, or WHOOP (or other similar devices) to monitor sleep and/or activity. For this review, inpatient psychiatric or mental health settings will include any facility providing mental health care support or treatment to patients staying within that facility for the duration of their treatment. Included studies must report outcomes related to feasibility, acceptability, clinical utility, or implementation. Feasibility refers to the practicality and

XSL•FO
RenderX

achievability of using wearables in a clinical environment, including factors such as cost, workload for staff, and technical and infrastructural requirements. Acceptability assesses the attitudes and willingness of both patients and clinicians to use the technology. The concept of clinical utility focuses on whether the data collected from wearables provide meaningful and actionable insights that can inform patient care (eg, care planning), detect changes, and evaluate treatment effectiveness. Studies will be excluded if they focus solely on outpatient settings or nonpsychiatric mental health inpatient areas. Only published peer-reviewed articles will be considered, and no restrictions will be placed on publication year, although it is likely that articles identified will be primarily in the last 5 to 10 years given the popularity and availability of these wearables during this period.

A comprehensive search strategy was developed (refer to Multimedia Appendix 1 showing the search strategy for PubMed, which will be translated to the other databases) and was peer-reviewed using the Peer Review of Electronic Search Strategies (PRESS) guideline [27]. The literature will be identified from four databases: (1) PubMed, (2) Embase, (3) PsycInfo, and (4) CINAHL. The search uses a combination of subject headings and keywords relating to wearable devices, sleep and activity monitoring, and inpatient psychiatric care.

All references identified from the databases will be imported into Covidence (Veritas Health Innovation Ltd) to manage duplicates and facilitate screening. Titles and abstracts will be screened by one reviewer, with the first 100 citations checked by a second reviewer for consistency. If 90% or more agreement is reached by the 2 reviewers, full-text screening will proceed. If not, the reviewers will meet to discuss discrepancies and carry on to another 100 citations until 90% agreement is reached. There is no predefined limit, and the screening will continue until 90% agreement is reached. Cohen κ will be reported as the measure of agreement. Full-text screening will be conducted independently by 2 reviewers, with disagreements resolved through discussion or consultation with a member of the review team as needed. The selection process will be documented using a PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram [28].

Data extraction will be carried out using a data extraction form, which will be pilot tested initially. One reviewer will extract the data, and a second reviewer will verify their accuracy. Extracted data will include basic study details (authors, title, etc); study characteristics (sample size, etc); type and version of wearable device; purpose of use; duration; outcomes monitored and assessed; and findings related to feasibility, acceptability, and utility. Additionally, ethical, legal, and safety aspects will be systematically extracted and synthesized as outcome domains. For each RQ, data will be characterized to address each review question. Both quantitative and qualitative findings will be brought together in a narrative structure. Descriptive statistics will be used to report discrete data, and all identified data will be accompanied by appropriate visualizations, narrative descriptions, and tables as necessary. For the thematic analysis, thematic findings will be identified using inductive analysis of the data. For the extraction form, pilot testing of the tool will be carried out using pilot studies to ensure consistency of the data extraction process between reviewers, and necessary modifications will be made to ensure alignment among the reviewers.

To maximize resources, our protocol will include 2 reviewers working to code data in tandem. As needed, a senior team member will be consulted to resolve any disagreement between the coders.

To validate findings and enhance interpretation, the research team will engage in a consultation process with key knowledge users. These consultations will include clinicians, researchers, individuals with lived experience of inpatient psychiatric care, and digital health leaders. Consultations will be conducted virtually and will aim to validate the results, identify implications for clinical practice and policy, and help guide future research directions. Feedback obtained will be integrated into the Discussion section of the final manuscript.

No ethics approval is required for this rapid review, as it involves the analysis of secondary data from published sources.

Results

The review will be conducted for a 6-month period and is anticipated to conclude in June 2026. In the first 2 months, the review team will translate the PubMed search strategy to other databases, carry out the database searches, and screen titles and abstracts using the eligibility criteria. In the third and fourth months, full-text articles will undergo review, and if they meet the inclusion criteria, data extraction will occur. In the fifth and sixth months, a narrative synthesis of the findings will be conducted. We will report the results of the searches and selection process via a PRISMA flow diagram. Data will be presented both quantitatively (device types used, common outcomes measured, etc) and qualitatively (themes related to feasibility, acceptability, implementation factors, etc). Results will be structured in alignment with the 3 RQs and key gaps in the literature will be highlighted. Knowledge user consultations will also occur in the final month, during which early synthesized findings will be shared and feedback will be obtained. We anticipate that these consultations will consist of in-person and virtual meetings with clinicians, researchers, and patient advisers who have or who plan to use wearables in their research or clinical care. We will plan to connect with the authors of up to 5 of the studies included in the review, as well as 5 people with lived experience and 5 clinicians from both medicine and nursing. The insights will enhance the contextual relevance of the review and guide the framing of the discussion for the final paper describing the findings. These numbers were selected to both manage feasibility and provide a breadth of opinions and insights. While formal research ethics board approval is not required for this review, we will obtain approval from our quality projects ethics review committee for the consultations that we plan to conduct.

Discussion

Anticipated Findings

This rapid review is expected to synthesize and clarify how wearable devices are currently being used to monitor sleep and

XSL•FO
RenderX

physical activity in inpatient psychiatric settings and to summarize the reported feasibility, acceptability, and clinical utility of these technologies in real-world inpatient contexts. On the basis of preliminary scoping and prior literature in adjacent settings, we anticipate that the findings will show that wearable devices are primarily used for research and observational purposes, with more limited examples of routine clinical integration into inpatient care processes. We further expect that feasibility and acceptability will be reported as generally favorable among many patient populations, particularly when devices are noninvasive, easy to use, and supported by staff; however, we also anticipate that important challenges related to safety policies, data governance, staffing capacity, and integration into clinical workflows may emerge. Regarding clinical utility, it is anticipated that most studies will report potential or perceived value rather than direct evidence of changes in clinical decision-making or patient outcomes. The review may therefore highlight a gap between the promise of wearable-derived data and its actionable use in inpatient psychiatric care.

Previous reviews of wearable technologies in mental health have largely focused on outpatient, community, or ambulatory populations, with an emphasis on associations between passively collected physiological data and psychiatric symptoms, relapse risk, or functional outcomes. These reviews have demonstrated growing interest in consumer-grade wearables and their potential to support mental health research and care, but they have provided limited insight into the unique realities of inpatient psychiatric environments. By focusing on inpatient mental health settings, this review addresses a gap in the literature. Inpatient environments differ from outpatient settings in patient acuity, safety requirements, environmental constraints, and clinical workflows. As a result, findings from outpatient studies may not be directly transferable but still may provide some important and interesting insights.

Findings from this rapid review will be disseminated through publication in a peer-reviewed journal and presentations at relevant national and international conferences. In addition, results will be shared with clinicians, health system leaders, and individuals with lived experience through targeted knowledge translation activities, including presentations, summaries, and stakeholder meetings. Where feasible, accessible summaries will be developed to support broader uptake and inform decision-making related to the use of wearable technologies in inpatient psychiatric care.

Conclusions

This rapid review will generate timely, practical insights into the current use and utility of wearable devices for tracking sleep and activity in inpatient psychiatric settings. By synthesizing evidence across different available wearable technologies and clinical contexts, the review will identify the feasibility, acceptability, and clinical utility of these tools, as well as common barriers and facilitators to implementation. Findings will provide a foundation for future clinical integration and research, particularly in tailoring digital health innovations to meet the needs of complex inpatient populations [29]. Ultimately, this work aims to support the development of safer, more responsive, and data-informed mental health care environments.

Acknowledgments

The generative artificial intelligence tool ChatGPT by OpenAI was used to check spelling and grammar for this paper and to offer copywriting suggestions**.**

Funding

This study received no external funding.

Conflicts of Interest

JB has received consulting fees from Tetricus Lavs inc. MP owns stock in Senseye inc. The authors have no further interests to declare.

Multimedia Appendix 1
PubMed search strategy conducted on August 7, 2025.
[DOCX File , 24 KB - resprot_v15i1e82290_app1.docx]

References

1. Wulff K, Gatti S, Wettstein JG, Foster RG. Sleep and circadian rhythm disruption in psychiatric and neurodegenerative disease. Nat Rev Neurosci 2010 Aug 14;11(8):589-599 [FREE Full text] [doi: 10.1038/nrn2868] [Medline: 20631712]
2. Harvey AG, Murray G, Chandler RA, Soehner A. Sleep disturbance as transdiagnostic: consideration of neurobiological mechanisms. Clin Psychol Rev 2011 Mar;31(2):225-235 [FREE Full text] [doi: 10.1016/j.cpr.2010.04.003] [Medline: 20471738]
3. Blanchard AW, Rufino KA, Nadorff MR, Patriquin MA. Nighttime sleep quality and daytime sleepiness across inpatient psychiatric treatment is associated with clinical outcomes. Sleep Med 2023 Oct;110:235-242 [FREE Full text] [doi: 10.1016/j.sleep.2023.08.011] [Medline: 37647715]

XSL•FO
RenderX

4. Rohr JC, Rufino KA, Alfano CA, Patriquin MA. Sleep disturbance in patients in an inpatient hospital mediates relationship between PTSD and suicidal ideation. J Psychiatr Res 2021 Jan;133:174-180 [FREE Full text] [doi: 10.1016/j.jpsychires.2020.12.039] [Medline: 33348251]
5. Hartwig EM, Rufino KA, Palmer CA, Shepard C, Alfano CA, Schanzer B, et al. Trajectories of self-reported sleep disturbance across inpatient psychiatric treatment predict clinical outcome in comorbid major depressive disorder and generalized anxiety disorder. J Affect Disord 2019 May 15;251:248-255 [FREE Full text] [doi: 10.1016/j.jad.2019.03.069] [Medline: 30953891]
6. Smith PJ, Merwin RM. The role of exercise in management of mental health disorders: an integrative review. Annu Rev Med 2021 Jan 27;72:45-62 [FREE Full text] [doi: 10.1146/annurev-med-060619-022943] [Medline: 33256493]
7. Krahn LE, Lin SC, Wisbey J, Rummans TA, O'Connor MK. Assessing sleep in psychiatric inpatients: nurse and patient reports versus wrist actigraphy. Ann Clin Psychiatry 1997 Dec;9(4):203-210. [doi: 10.1023/a:1022396108587] [Medline: 9511943]
8. Gonzalez R, Tamminga C, Tohen M, Suppes T. Comparison of objective and subjective assessments of sleep time in subjects with bipolar disorder. J Affect Disord 2013 Jul;149(1-3):363-366 [FREE Full text] [doi: 10.1016/j.jad.2013.02.013] [Medline: 23489400]
9. Matousek M, Cervena K, Zavesicka L, Brunovsky M. Subjective and objective evaluation of alertness and sleep quality in depressed patients. BMC Psychiatry 2004 May 26;4:14 [FREE Full text] [doi: 10.1186/1471-244X-4-14] [Medline: 15163350]
10. Ancoli-Israel S, Cole R, Alessi C, Chambers M, Moorcroft W, Pollak CP. The role of actigraphy in the study of sleep and circadian rhythms. Sleep 2003 May 01;26(3):342-392 [FREE Full text] [doi: 10.1093/sleep/26.3.342] [Medline: 12749557]
11. Scott H, Biello SM, Woods HC. Social media use and adolescent sleep patterns: cross-sectional findings from the UK Millennium Cohort Study. BMJ Open 2019 Oct 22;9(9):e031161 [FREE Full text] [doi: 10.1136/bmjopen-2019-031161] [Medline: 31641035]
12. Johnson C, Mathew SJ, Oh H, Rufino KA, Najafi B, Colombo AE, et al. Wearable technology: a promising opportunity to improve inpatient psychiatry safety and outcomes. J Psychiatr Res 2021 Mar;135:104-106 [FREE Full text] [doi: 10.1016/j.jpsychires.2021.01.003] [Medline: 33465608]
13. Piwek L, Ellis DA, Andrews S, Joinson A. The rise of consumer health wearables: promises and barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
14. Hassan L, Milton A, Sawyer C, Casson AJ, Torous J, Davies A, et al. Utility of consumer-grade wearable devices for inferring physical and mental health outcomes in severe mental illness: systematic review. JMIR Ment Health 2025 Jan 07;12:e65143 [FREE Full text] [doi: 10.2196/65143] [Medline: 39773905]
15. Povey J, Sweet M, Nagel T, Mills PP, Stassi CP, Puruntatameri AM, et al. Drafting the Aboriginal and Islander Mental Health Initiative for Youth (AIMhi-Y) app: results of a formative mixed methods study. Internet Interv 2020 Sep;21:100318 [FREE Full text] [doi: 10.1016/j.invent.2020.100318] [Medline: 32477884]
16. Naslund JA, Aschbrenner KA, Marsch LA, Bartels SJ. The future of mental health care: peer-to-peer support and social media. Epidemiol Psychiatr Sci 2016 Apr;25(2):113-122 [FREE Full text] [doi: 10.1017/S2045796015001067] [Medline: 26744309]
17. Robotham D, Satkunanathan S, Doughty L, Wykes T. Do we still have a digital divide in mental health? A five-year survey follow-up. J Med Internet Res 2016 Nov 22;18(11):e309 [FREE Full text] [doi: 10.2196/jmir.6511] [Medline: 27876684]
18. Westheimer JL, Smith RP, Iacobelli P, Oh H, Tezino L, Khan R, et al. The state of (mis) trust: human-centered technology development and implementation in intensive mental health settings. J Affect Disord 2024 Dec 15;367:318-323 [FREE Full text] [doi: 10.1016/j.jad.2024.08.206] [Medline: 39226937]
19. Westheimer JL, Moukaddam N, Lindsay JA, Sabharwal A, Najafi B, Iacobelli PA, et al. Technology implementation for mental health end users: a model to guide digital transformation for inpatient mental health professionals. JMIR Ment Health 2023 Apr 06;10:e40429 [FREE Full text] [doi: 10.2196/40429] [Medline: 37023415]
20. Gazor A, Mathew SJ, Patriquin MA. Q-15 checks and 1:1 observations: exacerbating a problem we are trying to solve? J Affect Disord 2020 Feb 15;263:552-554 [FREE Full text] [doi: 10.1016/j.jad.2019.11.018] [Medline: 31780137]
21. Cochrane Methods Rapid Reviews. URL: https://methods.cochrane.org/rapidreviews/ [accessed 2025-08-11]
22. Tricco AC, Garritty CM, Boulos L, Lockwood C, Wilson M, McGowan J, et al. Rapid review methods more challenging during COVID-19: commentary with a focus on 8 knowledge synthesis steps. J Clin Epidemiol 2020 Oct;126:177-183 [FREE Full text] [doi: 10.1016/j.jclinepi.2020.06.029] [Medline: 32615209]
23. Tricco AC, Langlois EV, Straus SE. Rapid reviews to strengthen health policy and systems: a practical guide. World Health Organization. 2017 Aug 10. URL: https://wkc.who.int/resources/publications/i/item/2017-08-10-rapid-reviews-to-strengthen-health-policy-and-systems-a-practical-guide [accessed 2026-02-25]
24. Garritty C, Nussbaumer-Streit B, Hamel C, Devane D, Cochrane Rapid Reviews Methods Group. Rapid reviews methods series: assessing the appropriateness of conducting a rapid review. BMJ Evid Based Med 2025 Jan 22;30(1):55-60 [FREE Full text] [doi: 10.1136/bmjebm-2023-112722] [Medline: 38485206]
25. Tricco AC, Antony J, Zarin W, Strifler L, Ghassemi M, Ivory J, et al. A scoping review of rapid review methods. BMC Med 2015 Sep 16;13:224 [FREE Full text] [doi: 10.1186/s12916-015-0465-6] [Medline: 26377409]

XSL•FO
RenderX

26. Peters MD, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
27. McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. J Clin Epidemiol 2016 Jul;75:40-46 [FREE Full text] [doi: 10.1016/j.jclinepi.2016.01.021] [Medline: 27005575]
28. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
29. Strudwick G, Kassam I, Torous J, Patenaude S. Building the infrastructure for sustainable digital mental health: it is "prime time" for implementation science. JMIR Ment Health 2025 Jul 21;12:e78791 [FREE Full text] [doi: 10.2196/78791] [Medline: 40690592]

Abbreviations

PRESS: Peer Review of Electronic Search Strategies
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RQ: research question

Edited by J Sarvestan; submitted 12.Aug.2025; peer-reviewed by S Mittal, A Gupta, G Ash; comments to author 01.Dec.2025; revised version received 02.Jan.2026; accepted 05.Jan.2026; published 10.Mar.2026.

Please cite as:
Strudwick G, Shin HD, Kassam I, Ramamoorthi K, Rosen B, Patriquin MA, Baker J, Neufeld N
The Use and Utility of Wearable Devices for Tracking Sleep and Activity in Inpatient Mental Health Settings: Protocol for a Rapid Review
JMIR Res Protoc 2026;15:e82290
URL: https://www.researchprotocols.org/2026/1/e82290
doi:10.2196/82290
PMID:

©Gillian Strudwick, Hwayeon Danielle Shin, Iman Kassam, Karishini Ramamoorthi, Benjamin Rosen, Michelle A Patriquin, Justin Baker, Nicholas Neufeld. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 10.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Managing Metabolic Dysfunction–Associated Steatotic Liver Disease: Protocol for a Scoping Review of Patient Perceptions, Barriers, and Facilitators

Sikyeong Park[1], BSN; Yu Shin Park[2], PhD; Dahye Hong[3], MPH; Bada Kang[2,4], PhD

[1]Yonsei University College of Nursing, Seoul, Republic of Korea

[2]Mo-Im Kim Nursing Research Institute, Yonsei University College of Nursing, Seoul, Republic of Korea

[3]College of Nursing and Brain Korea 21 FOUR Project, Yonsei University, Seoul, Republic of Korea

[4]Yonsei Evidence Based Nursing Centre of Korea: A JBI Affiliated Group, Yonsei University College of Nursing, Seoul, Republic of Korea

Corresponding Author:
Bada Kang, PhD
Mo-Im Kim Nursing Research Institute
Yonsei University College of Nursing
50-1 Yonsei-Ro, Seodaemun-Gu
Seoul, 03722
Republic of Korea
Phone: 82 2 2228 3283
Email: BDKANG@yuhs.ac

Abstract

Background: Metabolic dysfunction–associated steatotic liver disease (MASLD), formerly known as nonalcoholic fatty liver disease (NAFLD), is a major global health concern, affecting over 30% of adults worldwide. Closely associated with metabolic comorbidities such as obesity, type 2 diabetes, and dyslipidemia, MASLD relies heavily on health behavior modification for effective management. However, sustaining healthy behaviors remains challenging, particularly due to the disease's asymptomatic nature in its early stages and low perceived severity among patients. Thus, understanding patient perceptions and identifying barriers and facilitators are essential for developing effective, patient-centered interventions.

Objective: This scoping review protocol aims to systematically map the existing literature on patients' perceptions of MASLD and to identify barriers and facilitators that influence disease management and patient engagement.

Methods: This scoping review will follow the Joanna Briggs Institute scoping review methodology and the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines. A systematic search will be conducted in PubMed, CINAHL, Cochrane Library, and PsycINFO by using MeSH terms and keywords such as "MASLD," "NAFLD," "perception," "awareness," "attitude," "barriers," and "facilitators." Two independent reviewers will screen studies, extract data using a standardized charting form, and resolve discrepancies through discussion or consultation with a third reviewer. Eligible studies will include adults with MASLD/NAFLD and will focus on patient perceptions, health beliefs, behaviors, or barriers and facilitators to management. Data synthesis will be guided by the socioecological model to categorize findings across individual, interpersonal, organizational, community, and public policy levels.

Results: The research protocol was finalized in January 2026. This study received no external funding. A preliminary literature search was conducted by January 30, 2026, and formal screening of titles and abstracts is scheduled to begin in February 2026. Data extraction and synthesis are expected to be completed by April 2026. The final results are anticipated to be submitted for publication in the summer of 2026.

Conclusions: This scoping review will offer a comprehensive overview of patients' perceptions of MASLD, as well as the barriers and facilitators that influence its management. The results will inform the development of patient-centered strategies aimed at improving care delivery and reducing disease impact.

International Registered Report Identifier (IRRID): DERR1-10.2196/81404

(JMIR Res Protoc 2026;15:e81404) doi:10.2196/81404

XSL•FO
RenderX

KEYWORDS

barriers; facilitators; health behaviors modification; metabolic dysfunction–associated steatotic liver disease; nonalcoholic fatty liver disease; patient perception

Introduction

Metabolic dysfunction–associated steatotic liver disease (MASLD) has emerged as a significant global public health concern [1]. Defined by hepatic fat accumulation in the absence of significant alcohol consumption, MASLD is closely associated with metabolic comorbidities such as obesity, type 2 diabetes mellitus, and dyslipidemia [2]. Recent estimates indicate that MASLD affects over 30% of the global adult population, underscoring its high prevalence and clinical significance [3]. While the early stages are typically asymptomatic, MASLD can progress to more severe conditions, including nonalcoholic steatohepatitis, advanced fibrosis, cirrhosis, and hepatocellular carcinoma, underscoring the critical need for early detection and timely intervention [4]. Recently, a multi-society Delphi consensus renamed nonalcoholic fatty liver disease (NAFLD) to MASLD. This change reflects the adoption of affirmative diagnostic criteria that better represent the underlying pathophysiology and aims to remove stigmatizing language associated with the terms “fatty” and “nonalcoholic” [5].

Globally, MASLD imposes a substantial and growing health care and economic impact. This trend is also evident in South Korea, where prevalence estimates range from 10% to 50%, depending on diagnostic criteria and study populations [6]. According to data from the National Health Insurance Service of the Republic of Korea, the incidence of MASLD increased from 1.87% in 2010 to 4.47% in 2022, with projections estimating a prevalence of 43.8% by 2035 [7]. This rising impact extends beyond epidemiology, posing significant health care and economic challenges. The National Health Insurance Service claims analyses reveal that annual medical expenditures for individuals with MASLD are approximately 44% higher than those without the condition (KRW 787,579 [US $531] vs KRW 548,494 [US $327], respectively), largely due to greater pharmaceutical use and health care service utilization [8]. Likewise, data from the Korean Health Insurance Review and Assessment Service demonstrate that the prevalence of MASLD nearly doubled from 15.7 to 34.2 per 1000 population between 2010 and 2021, while the prevalence of metabolic dysfunction–associated steatohepatitis increased nearly 20-fold during the same period [9].

Health behavior modification, including dietary changes, increased physical activity, and weight reduction, is widely recognized as the cornerstone of MASLD management [10-13]. Sustained adherence to these behaviors has been shown to improve hepatic steatosis, reverse early-stage fibrosis, and reduce the risk of progression to advanced liver disease [14,15]. Despite these benefits, maintaining long-term health behavior changes remains a major challenge. Studies consistently report suboptimal adherence rates to dietary and exercise guidelines [16-18].

A growing body of literature has explored factors influencing patient engagement in MASLD management, identifying both barriers and facilitators. However, no comprehensive synthesis has systematically mapped the psychosocial, cultural, and structural factors that influence patient engagement in MASLD care. Existing studies typically focus on specific interventions or isolated populations, with limited attention to the broader sociocultural context shaping patient behaviors [10,19-21]. Moreover, much of the available evidence originates from Western populations, leaving significant gaps in understanding how sociocultural factors, health care systems, and patient beliefs affect disease management in other settings, particularly in Asian contexts such as South Korea [8,22,23]. Addressing these knowledge gaps is essential for informing the development of culturally sensitive, patient-centered interventions and public health strategies aimed at improving long-term outcomes for individuals living with MASLD.

Given these considerations, this scoping review will systematically map the existing literature on (1) how individuals with MASLD perceive the disease, (2) barriers that hinder effective management, and (3) facilitators that support sustained engagement and adherence to health behavior interventions. By synthesizing this evidence, the review seeks to provide a foundational understanding of factors that can inform the development of culturally tailored interventions and guide comprehensive public health responses to the growing impact of MASLD.

Methods

Overview

This scoping review will be conducted in accordance with the Joanna Briggs Institute (JBI) methodological framework [24]. The PRISMA-ScR (Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews) checklist will be used for reporting [25]. A structured approach will be used to systematically identify, evaluate, and map relevant literature on patient perceptions and barriers and facilitators influencing MASLD management.

Eligibility Criteria

To determine the eligibility criteria of the studies to be included in this review, we will apply the *Population*, *Concept*, *Context* framework for scoping review as outlined by JBI. The inclusion and exclusion criteria are specified as follows.

Population

The review will include adults (aged ≥18 years) diagnosed with NAFLD or MASLD based on either clinical confirmation or self-reported diagnosis, as explicitly defined in the included studies.

Concept

This review will include studies that examine how individuals with MASLD perceive the disease, encompassing their

XSL•FO
RenderX

awareness, knowledge, beliefs, and attitudes toward the disease. It will also consider studies that explore barriers and facilitators influencing patient engagement in disease management and adherence to health behavior changes. Specifically, by structuring identified determinants into perceptions, barriers and facilitators within the socioecological model (SEM) framework, this review aims to provide a comprehensive understanding of the psychological, behavioral, and sociocultural factors affecting self-management behaviors.

Context

There will be no restriction on geographic location or health care settings, as the aim is to provide a comprehensive understanding of the literature across diverse cultural and health care contexts.

Types of Studies

This scoping review will consider a wide range of study designs. Eligible studies will include observational designs such as case reports, case-control studies, and cohort studies, as well as experimental designs including quasi-experimental studies and randomized controlled trials. In addition, qualitative observational studies and mixed methods studies that integrate both observational and experimental components will be considered. Scientific conference proceedings that report sufficient methodological detail and outcome data to support extraction and synthesis will be eligible for inclusion. To ensure that the review reflects contemporary evidence from the past decade, only studies published in English from 2016 onward will be included. Editorials, letters, commentaries, and books will be excluded.

Exclusion Criteria

Studies will be excluded if they focus exclusively on pharmacological, genetic, or physiological aspects of MASLD without addressing patient perspectives or behavioral factors. Regarding the study population, studies will be excluded if they do not involve individuals explicitly diagnosed with MASLD or NAFLD, or if they include mixed populations (eg, combined samples of MASLD and other liver diseases) or participants with significant alcohol consumption, unless data specific to the MASLD/NAFLD subgroup can be clearly disaggregated. Furthermore, studies that fail to explicitly describe barriers and facilitators to disease management or patient engagement will also be excluded. Additionally, studies that examine only health care providers' perspectives without incorporating patient-level insights will not be considered.

Data Sources and Search Strategy

The search strategy for this scoping review will follow the JBI 3-step approach. First, a preliminary search of PubMed and CINAHL was conducted to identify relevant keywords and Medical Subject Headings (MeSH) terms related to the research question. Based on this initial search, a comprehensive electronic search will be performed across the following databases: PubMed (MEDLINE), CINAHL, Cochrane Library, and PsycINFO. The search will incorporate both MeSH terms and free-text keywords related to the population and core concepts of interest, including but not limited to "metabolic dysfunction–associated steatotic liver disease (MASLD)" OR "nonalcoholic fatty liver disease (NAFLD)" AND "perception," "awareness," "attitude," "belief," "barriers," "limitations," "facilitators," "enabler," "behavior," and "self-management."

Boolean operators (AND/OR), truncation, and database-specific indexing will be applied to optimize the search. The specific keywords and controlled vocabulary terms (eg, MeSH) used are listed in Multimedia Appendix 1. The complete search strategies for all databases, including PubMed, CINAHL, Cochrane Library, and PsycINFO, are detailed in Multimedia Appendix 2.

Screening and Selection Procedures of Eligible Studies

All records identified through the search will be imported into EndNote 20 (Clarivate Analytics) for reference management, and duplicate entries will be removed. Two independent reviewers will screen the titles and abstracts of retrieved studies for relevance based on the inclusion and exclusion criteria. Full texts of potentially eligible articles will then be assessed independently. Any disagreements during the selection process will be resolved through discussion or, if necessary, consultation with a third reviewer. The study selection process will be documented using a PRISMA-ScR flow diagram.

Data Extraction and Synthesis

Data extraction will use a standardized charting form aligned with JBI guidelines. Two reviewers will independently extract data, including study characteristics (author, year, country, design, objectives, setting, and population details) and key concepts (patient perceptions, barriers, facilitators). Outcomes, key findings, study limitations, and author recommendations will also be documented (Multimedia Appendix 3). A standardized Microsoft Excel spreadsheet will be used to organize the extracted data.

To ensure a structured synthesis, the analysis will be guided by SEM [26]. Patient perceptions, barriers, and facilitators across included studies will be systematically coded and categorized into 5 distinct SEM levels: individual, interpersonal, organizational, community, and public policy (Multimedia Appendix 4). This framework-based approach allows for a stratified analysis of the factors influencing MASLD management, moving beyond a purely narrative summary to identify recurring patterns and evidence across SEM levels.

Ethical Considerations

As this is a review of previously published literature, ethical approval from the institutional review board is not required.

Results

The research protocol was finalized in January 2026. This study received no external funding. A preliminary literature search was conducted by January 30, 2026, and formal screening of titles and abstracts is scheduled to begin in February 2026. Data extraction and synthesis are expected to be completed by April 2026. The final results are anticipated to be submitted for publication in the summer of 2026.

XSL•FO
RenderX

Discussion

This scoping review is expected to provide a foundational understanding of patient perceptions and the barriers and facilitators influencing the implementation of management strategies for MASLD. Although health behavior interventions—such as exercise programs, dietary modifications [10], and digital health strategies [20]—have demonstrated efficacy in controlled research settings, their translation into clinical practice remains limited.

In the broader context of chronic disease management, a substantial body of evidence supports the role of facilitators such as culturally tailored educational interventions [27], digital health technologies [28], and integrated multidisciplinary care models in improving patient engagement, treatment adherence, and clinical outcomes [29]. For instance, in conditions like type 2 diabetes, cardiovascular disease, and obesity, culturally sensitive education has been shown to enhance health literacy, reduce stigma, and promote sustainable behavior change across diverse populations [30-32]. Similarly, digital health platforms—including mobile apps, telemonitoring, and remote coaching—have improved access to care, supported self-management, and facilitated real-time feedback, especially for patients with limited physical access to clinical services [28]. Multidisciplinary care models that incorporate physicians, nurses, dietitians, and behavioral health professionals have also proven effective in delivering coordinated, holistic care for complex chronic conditions [29].

These facilitators are particularly applicable to MASLD, where long-term management heavily depends on patient-driven health behavior changes and sustained engagement with nonpharmacologic interventions [11]. Culturally adapted educational content may enhance disease awareness, encourage health behavior changes, and reduce barriers related to health beliefs or stigma. Digital tools can offer scalable, personalized support for behavior tracking, goal setting, and continuous monitoring outside of clinical visits [20]. Integrated care teams can provide comprehensive and coordinated support that addresses not only hepatic health, but also the psychological, nutritional, and metabolic dimensions of MASLD [33]. Together, these approaches highlight the importance of aligning intervention strategies with the diverse needs, preferences, and contexts of patients to improve the real-world implementation and effectiveness of MASLD management.

Previous studies have identified implementation challenges, including time and resource constraints, limited provider training, and insufficient infrastructure to support sustained patient engagement [34,35]. Although digital health interventions offer promising support for MASLD management, their scalability and adoption are often hindered by system-level barriers such as technical challenges and regulatory constraints. These limitations are further exacerbated by patient-level factors, including limited disease awareness, low intrinsic motivation for behavioral change, and sociocultural or systemic influences, all of which hinder engagement and adherence, thereby diminishing the overall effectiveness of such interventions [33].

These anticipated findings highlight the need for a deeper understanding of multilevel implementation factors influencing MASLD management. This scoping review aims to comprehensively identify and map the psychosocial, cultural, and structural determinants that influence MASLD management. The findings are expected to inform the development of tailored strategies to enhance the uptake, patient engagement, adherence, and long-term effectiveness of interventions across diverse health care settings. By mapping the interplay among patient perceptions, contextual barriers, and facilitators, the study seeks to inform the design of patient-centered approaches that are both feasible and sustainable in real-world practice.

Furthermore, this integrated understanding has implications beyond individual-level behavior change, extending to health system strengthening and policy development. In light of the increasing prevalence of MASLD in South Korea and globally [3,6], the findings from this review may provide empirical evidence to guide national strategies focused on early detection and prevention, the development of provider-facing education and counseling guidelines, and the implementation of public health campaigns aimed at improving health literacy. Collectively, these outcomes will contribute to a multilevel framework for advancing MASLD care and improving long-term patient outcomes.

However, several limitations should be considered. First, the heterogeneity of the included studies across diverse health care systems and sociocultural settings may limit the generalizability of these findings. Second, as a scoping review, this study focuses on mapping existing evidence rather than critically appraising study quality or establishing causality, which limits its direct applicability to clinical guideline development. Finally, this protocol was limited by the absence of patient involvement in the design or search strategy development phases. While this review aims to synthesize patient perspectives reported in the existing literature, the lack of direct consultation with individuals living with MASLD may limit the immediate applicability of the findings. Importantly, such research and subsequent intervention development should prioritize the active engagement of patients to ensure clinical and contextual relevance.

Future research should prioritize context-specific studies across different regions to explore patients' perceptions and cultural influences on disease management. Moreover, identifying and evaluating evidence-based interventions tailored to local sociocultural environments will play a critical role in addressing the global impact of MASLD and improving long-term patient outcomes.

In conclusion, this scoping review protocol outlines a systematic and theory-informed approach to identifying the psychosocial, cultural, and structural determinants that influence MASLD management. By mapping the interplay among patient perceptions, contextual barriers, and facilitators using SEM, the review aims to inform the development of patient-centered strategies that are both feasible and sustainable in real-world clinical practice. Ultimately, the findings are expected to contribute to a comprehensive, multilevel framework for

XSL•FO
RenderX

advancing MASLD care and to help bridge the gap between clinical recommendations and the lived experiences of patients.

Acknowledgments

The authors thank Ms Dami Jeong, Medical Research Librarian at Yonsei University, for her valuable assistance and support. Generative artificial intelligence tools (ChatGPT-4 and Gemini) were used to refine the clarity and grammar of this manuscript. The authors reviewed and revised the output to ensure accuracy.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Authors' Contributions

Concept and design: SP, BK
Manuscript drafting: SP
Critical revisions: SP, YSP, DH, BK
Supervision: BK
All authors interpreted the data and approved the final version for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Specific keywords and controlled vocabulary terms used for this study.
[DOCX File , 25 KB - resprot_v15i1e81404_app1.docx]

Multimedia Appendix 2
Search strategies.
[DOCX File , 30 KB - resprot_v15i1e81404_app2.docx]

Multimedia Appendix 3
Data extraction instrument.
[DOCX File , 26 KB - resprot_v15i1e81404_app3.docx]

Multimedia Appendix 4
Data synthesis instrument.
[DOCX File , 25 KB - resprot_v15i1e81404_app4.docx]

References

1. Younossi ZM, Razavi H, Sherman M, Allen AM, Anstee QM, Cusi K, et al. Addressing the high and rising global burden of metabolic dysfunction-associated steatotic liver disease (MASLD) and metabolic dysfunction-associated steatohepatitis (MASH): From the growing prevalence to payors' perspective. Aliment Pharmacol Ther 2025 May;61(9):1467-1478. [doi: 10.1111/apt.70020] [Medline: 39967239]
2. Cotter TG, Rinella M. Nonalcoholic fatty liver disease 2020: the state of the disease. Gastroenterology 2020 May;158(7):1851-1864. [doi: 10.1053/j.gastro.2020.01.052] [Medline: 32061595]
3. Riazi K, Azhari H, Charette JH, Underwood FE, King JA, Afshar EE, et al. The prevalence and incidence of NAFLD worldwide: a systematic review and meta-analysis. Lancet Gastroenterol Hepatol 2022 Sep;7(9):851-861. [doi: 10.1016/S2468-1253(22)00165-0] [Medline: 35798021]
4. Loomba R, Wong R, Fraysse J, Shreay S, Li S, Harrison S, et al. Nonalcoholic fatty liver disease progression rates to cirrhosis and progression of cirrhosis to decompensation and mortality: a real world analysis of Medicare data. Aliment Pharmacol Ther 2020 Jun;51(11):1149-1159. [doi: 10.1111/apt.15679] [Medline: 32372515]
5. Rinella ME, Lazarus JV, Ratziu V, Francque SM, Sanyal AJ, Kanwal F, NAFLD Nomenclature consensus group. A multisociety Delphi consensus statement on new fatty liver disease nomenclature. J Hepatol 2023 Dec;79(6):1542-1556 [FREE Full text] [doi: 10.1016/j.jhep.2023.06.003] [Medline: 37364790]
6. Choi YJ, Park J, Cho H, Shin MG, Nah E. Metabolic dysfunction-associated steatotic liver disease in the Korean general population: epidemiology, risk factors, and non-invasive screening. Metabolites 2025 Apr 30;15(5):299 [FREE Full text] [doi: 10.3390/metabo15050299] [Medline: 40422876]

XSL•FO
RenderX

7. Kang SY, Kim Y, Park HS. Trends in the prevalence of non-alcoholic fatty liver disease and its future predictions in Korean men, 1998-2035. J Clin Med 2020 Aug 13;9(8):2626 [FREE Full text] [doi: 10.3390/jcm9082626] [Medline: 32823604]
8. Park JW, Yoo J, Lee DH, Chang Y, Jo H, Cho YY, Korean Association for the Study of the Liver. Evolving epidemiology of non-alcoholic fatty liver disease in South Korea: incidence, prevalence, progression, and healthcare implications from 2010 to 2022. Korean J Intern Med 2024 Nov;39(6):931-944 [FREE Full text] [doi: 10.3904/kjim.2024.164] [Medline: 39444335]
9. Lee N, Jeong S, Wang J, Choi Y, Oh H, Cho J, et al. The clinical diagnosis-based nationwide epidemiology of metabolic dysfunction-associated liver disease in Korea. J Clin Med 2023 Dec 12;12(24):7634 [FREE Full text] [doi: 10.3390/jcm12247634] [Medline: 38137703]
10. Mambrini SP, Grillo A, Colosimo S, Zarpellon F, Pozzi G, Furlan D, et al. Diet and physical exercise as key players to tackle MASLD through improvement of insulin resistance and metabolic flexibility. Front Nutr 2024;11:1426551 [FREE Full text] [doi: 10.3389/fnut.2024.1426551] [Medline: 39229589]
11. Chen M, Chen Y, Lin J, Hu R, Liu D, Chen J, et al. Evidence summary of lifestyle interventions in adults with metabolic dysfunction-associated steatotic liver disease. Front Nutr 2024;11:1421386 [FREE Full text] [doi: 10.3389/fnut.2024.1421386] [Medline: 39834455]
12. Rinella ME, Neuschwander-Tetri BA, Siddiqui MS, Abdelmalek MF, Caldwell S, Barb D, et al. AASLD Practice Guidance on the clinical assessment and management of nonalcoholic fatty liver disease. Hepatology 2023 May 01;77(5):1797-1835 [FREE Full text] [doi: 10.1097/HEP.0000000000000323] [Medline: 36727674]
13. Ahmed IA, Mikail MA, Mustafa MR, Ibrahim M, Othman R. Lifestyle interventions for non-alcoholic fatty liver disease. Saudi J Biol Sci 2019 Nov;26(7):1519-1524 [FREE Full text] [doi: 10.1016/j.sjbs.2018.12.016] [Medline: 31762620]
14. Ciardullo S, Mantovani A, Morieri ML, Muraca E, Invernizzi P, Perseghin G. Impact of MASLD and MetALD on clinical outcomes: a meta-analysis of preliminary evidence. Liver Int 2024 Aug;44(8):1762-1767. [doi: 10.1111/liv.15939] [Medline: 38597738]
15. Sohn W, Lee Y, Kim SS, Kim JH, Jin Y, Kim G, Korean Association for the Study of the Liver (KASL). KASL clinical practice guidelines for the management of metabolic dysfunction-associated steatotic liver disease 2025. Clin Mol Hepatol 2025 Feb;31(Suppl):S1-S31 [FREE Full text] [doi: 10.3350/cmh.2025.0045] [Medline: 39967303]
16. Arora C, Malhotra A, Ranjan P, Vikram NK, Dwivedi SN, Singh N, Shalimar. Perceived barriers and facilitators for adherence to lifestyle prescription: perspective of obese patients with non alcoholic fatty liver disease from north India. Diabetes Metab Syndr 2021;15(4):102138. [doi: 10.1016/j.dsx.2021.05.011] [Medline: 34186359]
17. Guan Q, Dong R, Zhang M, Chang D, Zhang R, Wang Y, et al. Factors influencing knowledge-action gap in patients with metabolic dysfunction-associated fatty liver disease: a qualitative study. J Nutr Educ Behav 2025 Apr;57(4):274-284. [doi: 10.1016/j.jneb.2024.12.010] [Medline: 39864004]
18. Stine JG, Soriano C, Schreibman I, Rivas G, Hummer B, Yoo E, et al. Breaking down barriers to physical activity in patients with nonalcoholic fatty liver disease. Dig Dis Sci 2021 Oct;66(10):3604-3611 [FREE Full text] [doi: 10.1007/s10620-020-06673-w] [Medline: 33098023]
19. Bae J, Han E, Lee HW, Park C, Chung CH, Lee DH, Fatty Liver Research Group of the Korean Diabetes Association. Metabolic dysfunction-associated steatotic liver disease in type 2 diabetes mellitus: a review and position statement of the fatty liver research group of the Korean Diabetes Association. Diabetes Metab J 2024 Nov;48(6):1015-1028 [FREE Full text] [doi: 10.4093/dmj.2024.0541] [Medline: 39610131]
20. Kaewdech A, Assawasuwannakit S, Churuangsuk C, Chamroonkul N, Sripongpun P. Effect of smartphone-assisted lifestyle intervention in MASLD patients: a randomized controlled trial. Sci Rep 2024 Jun 17;14(1):13961 [FREE Full text] [doi: 10.1038/s41598-024-64988-4] [Medline: 38886203]
21. Amer J, Abdoh Q, Salous Z, Alsoud EA, AbuBaker S, Salhab A, et al. A cross-sectional study of risk factors associated with sarcopenia in patients with metabolic dysfunction-associated steatotic liver disease. Front Med (Lausanne) 2025;12:1488068 [FREE Full text] [doi: 10.3389/fmed.2025.1488068] [Medline: 40041458]
22. Ng CH, Lim WH, Chin YH, Yong JN, Zeng RW, Chan KE, et al. Living in the non-alcoholic fatty liver disease silent epidemic: a qualitative systematic review of patients' perspectives. Aliment Pharmacol Ther 2022 Aug;56(4):570-579. [doi: 10.1111/apt.17121] [Medline: 35791632]
23. Jang Y, Lee JY, Kim SU, Kim B. A qualitative study of self-management experiences in people with non-alcoholic fatty liver disease. Nurs Open 2021 Nov;8(6):3135-3142 [FREE Full text] [doi: 10.1002/nop2.1025] [Medline: 34405583]
24. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth 2020 Oct;18(10):2119-2126. [doi: 10.11124/JBIES-20-00167] [Medline: 33038124]
25. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473 [FREE Full text] [doi: 10.7326/M18-0850] [Medline: 30178033]
26. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. Health Educ Q 1988;15(4):351-377. [doi: 10.1177/109019818801500401] [Medline: 3068205]

XSL•FO
RenderX

27. Creamer J, Attridge M, Ramsden M, Cannings-John R, Hawthorne K. Culturally appropriate health education for Type 2 diabetes in ethnic minority groups: an updated Cochrane Review of randomized controlled trials. Diabet Med 2016 Feb;33(2):169-183. [doi: 10.1111/dme.12865] [Medline: 26202820]
28. Ma Y, Zhao C, Zhao Y, Lu J, Jiang H, Cao Y, et al. Telemedicine application in patients with chronic disease: a systematic review and meta-analysis. BMC Med Inform Decis Mak 2022 Apr 19;22(1):105 [FREE Full text] [doi: 10.1186/s12911-022-01845-2] [Medline: 35440082]
29. Yeung CHT, Santesso N, Zeraatkar D, Wang A, Pai M, Sholzberg M, et al. Integrated multidisciplinary care for the management of chronic conditions in adults: an overview of reviews and an example of using indirect evidence to inform clinical practice recommendations in the field of rare diseases. Haemophilia 2016 Jul;22 Suppl 3:41-50. [doi: 10.1111/hae.13010] [Medline: 27348400]
30. Rosal MC, Ockene IS, Restrepo A, White MJ, Borg A, Olendzki B, et al. Randomized trial of a literacy-sensitive, culturally tailored diabetes self-management intervention for low-income latinos: latinos en control. Diabetes Care 2011 Apr;34(4):838-844 [FREE Full text] [doi: 10.2337/dc10-1981] [Medline: 21378213]
31. Singh H, Fulton J, Mirzazada S, Saragosa M, Uleryk EM, Nelson MLA. Community-based culturally tailored education programs for Black communities with cardiovascular disease, diabetes, hypertension, and stroke: systematic review findings. J Racial Ethn Health Disparities 2023 Dec;10(6):2986-3006 [FREE Full text] [doi: 10.1007/s40615-022-01474-5] [Medline: 36508135]
32. Kong A, Tussing-Humphreys LM, Odoms-Young AM, Stolley MR, Fitzgibbon ML. Systematic review of behavioural interventions with culturally adapted strategies to improve diet and weight outcomes in African American women. Obes Rev 2014 Oct;15 Suppl 4(0 4):62-92 [FREE Full text] [doi: 10.1111/obr.12203] [Medline: 25196407]
33. Kravchuk S, Bychkov M, Kozyk M, Strubchevska O, Kozyk A. Managing metabolic dysfunction-associated steatotic liver disease (MASLD) in the digital era: overcoming barriers to lifestyle change. Cureus 2025 May;17(5):e84803. [doi: 10.7759/cureus.84803] [Medline: 40568294]
34. Duffy A, Christie GJ, Moreno S. The challenges toward real-world implementation of digital health design approaches: narrative review. JMIR Hum Factors 2022 Sep 09;9(3):e35693 [FREE Full text] [doi: 10.2196/35693] [Medline: 36083628]
35. Silva CS, Godinho C, Encantado J, Rodrigues B, Carraça EV, Teixeira PJ, et al. Implementation determinants of physical activity interventions in primary health care settings using the TICD framework: a systematic review. BMC Health Serv Res 2023 Oct 11;23(1):1082 [FREE Full text] [doi: 10.1186/s12913-023-09881-y] [Medline: 37821932]

Abbreviations

JBI: Joanna Briggs Institute
MASLD: metabolic dysfunction–associated steatotic liver disease
MeSH: Medical Subject Headings
NAFLD: nonalcoholic fatty liver disease
PRISMA-ScR: Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews
SEM: socioecological model

Edited by J Sarvestan; submitted 28.Jul.2025; peer-reviewed by C Sun, C Beresford; comments to author 28.Jan.2026; accepted 03.Mar.2026; published 24.Mar.2026.

Please cite as:
Park S, Park YS, Hong D, Kang B
Managing Metabolic Dysfunction–Associated Steatotic Liver Disease: Protocol for a Scoping Review of Patient Perceptions, Barriers, and Facilitators
JMIR Res Protoc 2026;15:e81404
URL: https://www.researchprotocols.org/2026/1/e81404
doi:10.2196/81404
PMID:41874552

©Sikyeong Park, Yu Shin Park, Dahye Hong, Bada Kang. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 24.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Methodology for Assessing University Members' Health and Well-Being and Their Influencing Factors: Protocol for a Rapid Scoping Review of Survey-Based Research

Varga Zsuzsanna[1*], PhD; László Balkányi[2*], PhD; Anikó Gyulai[3*], PhD; László Lippai[4*], PhD; Andrea Mária Rucska†[5*], PhD; Zsuzsanna Soósné Kiss[6*], PhD; Péter Varsányi[7*], MD; József Vitrai[8*], PhD; Boróka Gács[1*], PhD

[1]Department of Behavioural Sciences, Medical School, University of Pécs, Pécs, Baranya, Hungary

[2]Syreon Research Institute, Budapest, Hungary

[3]Department of Preventive Health Sciences, Faculty of Health Sciences, University of Miskolc, Miskolc, Borsod-Abaúj-Zemplén, Hungary

[4]Institute of Applied Health Sciences and Environmental Education, University of Szeged, Szeged, Hungary

[5]Faculty of Health Sciences, University of Miskolc, Miskolc, Hungary

[6]Department of Obstetrics and Gynecology, Faculty of Health and Sport Sciences, Széchenyi István University, Győr, Győr-Moson-Sopron, Hungary

[7]Society to Renew the Hungarian Public Health, Budapest, Hungary

[8]Department of Psychology and Health Management, Faculty of Sports and Health Sciences, Széchenyi István University, Győr, Győr-Moson-Sopron, Hungary

[*]all authors contributed equally

Corresponding Author:
Varga Zsuzsanna, PhD
Department of Behavioural Sciences
Medical School
University of Pécs
Szigeti str 12
Pécs, Baranya, 7624
Hungary
Phone: 36 307013406
Email: zsuzsanna.varga@aok.pte.hu

Abstract

Background: This scoping review protocol addresses the imperative need for a comprehensive understanding of the health and well-being of university members, aligning with the global recognition of universities as pivotal in promoting holistic well-being. The lack of consensus and diverse definitions surrounding health and well-being in the academic literature necessitate a systematic approach. The scoping review protocol is designed to develop proposals for measures to improve the health and well-being of university members.

Objective: The objective of this scoping review is to systematically map the domains, topics, and methodological characteristics of survey-based studies assessing health and well-being among university members, including students and employees.

Methods: This protocol follows the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines to outline a scoping review to map the existing literature. The review uses the SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, and Research type) tool to define key elements of the research questions. Eligibility criteria include English-language publications reporting on health and well-being surveys of university members published within the past 10 years. Several electronic databases and gray literature repositories were searched.

Results: The scoping review based on this protocol has been completed and published, and this manuscript reports the methodological framework applied during that process. It specifies the primary outputs, including a thematic domain framework, an inventory of survey instruments, and a methodological overview of survey implementation in higher education settings.

Conclusions: By providing a transparent and reproducible methodological description, this protocol supports a comprehensive understanding of health and well-being survey practices in higher education and informs the development of comparable assessment approaches for universities at national and international levels.

Trial Registration: OSF Registries 10.17605/OSF.IO/JMU78; https://osf.io/jmu78/overview

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e83845) doi:10.2196/83845

KEYWORDS

scoping review; health; well-being; university members; survey

Introduction

Recently, it has been increasingly recognized and accepted that, alongside health-supporting public policies and care systems, educating individuals about health-conscious lifestyles is essential [1]. Consistent with this, Pronk et al [2] emphasized that the health and well-being of all people is a shared responsibility. According to Dooris and Doherty [3], universities occupy a prominent, privileged, and unique position that can help foster the health and well-being of their members and ultimately society through their policies and practices. Given their multifaceted roles as educational institutions, workplaces, and community actors, universities are increasingly recognized as strategic settings for promoting health and well-being in a comprehensive and sustainable manner [4]. This recognition has been further reinforced by international frameworks and reviews that conceptualize universities as whole systems capable of influencing health through governance, organizational culture, campus environments, and academic activities [5,6].

Universities committed to health promotion integrate health into their daily operations, business practices, and academic pursuits. This integration enhances a university culture steeped in empathy, well-being, equity, and social justice. International literature on health-promoting universities emphasizes that a "whole university" or systems-based approach moves beyond isolated interventions and instead embeds health considerations into institutional strategy, leadership, and decision-making processes [4,7]. Such initiatives not only improve the health of those within the university but also outside the university and strengthen the social, environmental, economic, organizational, and cultural sustainability of both the university community and the surrounding areas [1]. The operationalization of health and well-being development systems has become a crucial aspect of successful university management, profoundly impacting the future trajectory of these institutions [8]. Managing systems to promote health and well-being is now an essential part of running a successful university and is fundamental to the future of universities [6]. This approach necessitates the gathering and analysis of data concerning the health and well-being of the university community, as well as the factors influencing these areas [9,10]. The lack of comprehensive knowledge regarding the health and well-being of university members (ie, students and employees) [11], coupled with the diversity in definitions of the concepts and the ongoing debate in the literature regarding the proper conceptualization and measurement of health and well-being [12], highlights significant challenges in this area. These challenges underscore the need for robust, context-sensitive frameworks and measurement approaches that can inform evidence-based decision-making within higher education institutions (HEIs) [13,14].

According to the Organisation for Economic Co-operation and Development (OECD), the health of the population in Hungary is below what would be expected given its socioeconomic status, and there is a significant gap in opportunities and capabilities for "healthier choices" in daily life [15]. Lifestyle-related risk factors are responsible for almost 50% of deaths in Hungary. Excessive alcohol consumption is high among both adolescents and adults, as are the proportion of obese adults and the rate of preventable deaths. These indicators are worse than the European Union average [15]. Several research studies have been conducted in Hungary examining university students' health and risk behaviors [16,17] or the methods used to study these behaviors [18,19]. A national survey conducted among Hungarian university students in 2021 (N=7639, 47 universities) examined students' psychological problems, coping strategies, and resources, drawing attention to the unfavorable psychological state of students (10.2% mental illness, more than 40% symptoms of depression, and more than 50% in crisis) [20]. In 2022, a representative Hungarian research (N=1385) was conducted to examine the risk behavior of university students and compare it with that of a nonuniversity population of the same age. Smoking is less common, while alcohol and drug use are much more prevalent among university students than among nonuniversity students in the same age group. Approximately one-quarter of university students have some chronic physical and/or mental illness and also experience problematic internet use [21,22].

Despite all this, health awareness is not treated as a key issue in Hungary, either in people's daily lives or at a systemic level. To change this situation, in mid-2023, several academics representing Hungarian universities initiated the establishment of the Network of Health Promotion Universities in Hungary, following the Okanagan Charter [6]. As a first step, they established the Health and Well-being Survey Working Group to develop and implement a joint survey system in line with the World Health Organization (WHO) guidelines to assess the needs and requirements of university members [23]. The authors of this paper, as members of the working group, planned a scoping review to systematically map previous research in this field and to establish a national survey system to support the development of an integrated health and well-being assessment instrument for universities in the network. Scoping reviews are particularly well suited for this purpose, as they allow for the comprehensive mapping of heterogeneous evidence, including variations in conceptual frameworks, target populations, and methodological approaches [24,25]. The review is undertaken to identify the areas covered by past studies, examine their methodological elements, and evaluate the adequacy of their conditions, contributing to a more holistic understanding of health promotion in higher education. By doing so, the authors intend to provide universities with a solid foundation of knowledge and best practices, enabling them to tailor their health promotion strategies effectively. This endeavor is crucial for ensuring that health promotion in universities is evidence-based, contextually relevant, and capable of making a meaningful impact on the health and well-being of university communities.

XSL•FO
RenderX

This approach is consistent with the broader literature on health promotion in higher education [3,12], which underscores the need for systematic and comprehensive health initiatives in academic institutions. Furthermore, this review is aligned with the WHO Geneva Charter for Well-being [26] and the global framework for health-promoting schools [23] and adapts these principles to the university context. It also aims to facilitate knowledge transfer across different educational levels, in accordance with WHO guidelines [27]. Despite the growing number of surveys, no scoping review has systematically mapped the domains, instruments, and methodological characteristics of university-based health and well-being surveys. A scoping review conducted using the methodological framework described in this protocol has since been completed and published [28].

On the basis of our study objectives, we propose a protocol that is specifically tailored for universities embarking on initiatives to promote health. This scoping review, serving as the initial phase of the planned survey system, aims to examine health-related and well-being–related surveys performed in universities. The goal is to explore which measurement tools are used and which aspects of well-being are assessed.

Methods

Reporting Guidelines

This protocol is reported in accordance with the PRISMA-P (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols) 2015 guidelines. The completed scoping review derived from this protocol followed the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) reporting guidelines [29]. The methodological approach is described below. This protocol documents the methodological framework applied in the completed scoping review [28] and is presented to support transparent reporting and reuse in future scoping reviews in this field.

The protocol was registered in the Open Science Framework (OSF) Registries.

Rationale

As mentioned in the Introduction, our Surveying Health and Well-being Working Group is designing a survey system for assessing the health and well-being of university members, which is intended to be developed based on the information currently available in the literature. Ideally, the information obtained from the survey system may inform the planning and implementation of interventions by Hungarian universities to improve the health and well-being of their members. It is envisaged that the survey system will consist of a questionnaire, a protocol for the implementation of the survey, consultative advice, a common database, analysis procedures, and reporting templates. Universities wishing to use the survey system will agree to adhere to the survey protocol in an agreement with the working group, which is intended to support comparability of results between universities and analysis of national or university trends. For the development of the survey system, this scoping review is intended to collect topics (possibly questions) that have already been used to assess health and well-being, and methodological experience in the design, preparation, implementation, and evaluation of relevant surveys on this topic.

In addition to its applied relevance, this scoping review is intended to support a conceptual synthesis of how health and well-being are operationalized and measured in higher education contexts. By mapping the domains, constructs, and indicators assessed across existing surveys, the review will contribute to the identification of conceptual patterns and gaps in the current measurement landscape. This domain-based mapping allows the findings to be interpreted in relation to international frameworks of health and well-being used in public health and educational research. In this way, the review will provide a conceptual foundation that complements its potential applied and policy-relevant implications.

Objectives

Our scoping review was designed to answer the following two questions: (1) What topics are covered in the articles on university members' health and well-being? and (2) What methodologies and techniques are used in the health and well-being surveys of the university members?

The purpose of this scoping review is to systematically map the domains, instruments, and methodological characteristics of health and well-being surveys conducted in higher education. This mapping provides a structured basis for understanding how health and well-being are currently operationalized and measured in university settings.

Eligibility Criteria

As Table 1 describes, the eligibility criteria were developed in accordance with the general practice of systematic reviews and with the time and human resources available to us. All types of English-language publications published in the past 10 years that meet the search criteria and are available in full text were considered. In this way, the researchers were able to gather as much information as possible on the topics to be considered in the development of the questionnaire and the design of the survey protocol.

According to the literature [30], the SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research Type) tool is appropriate for scoping reviews. This was used to define key elements of the review research question. The sample (the target population) was university members in line with the aims of our project. The phenomena studied were all types of research that aim to measure the health and well-being of university members. Design and evaluation were not restricted; that is, anything could be used, as the topics to be studied could be drawn from any research. For the research type, the review focused on surveys.

XSL•FO
RenderX

Table 1. Eligibility criteria for the review. The criteria were organized following the SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research Type) approach. Eligibility was further informed by the World Health Organization's multidisciplinary framework on well-being [23].

Category	Inclusion criteria	Exclusion criteria
Sample	Adults (aged 18 years and above), university members (students and staff), including mixed populations when general health and well-being were assessed[a]	Non-English studies
Phenomenon of Interest	All types of research that aim to measure the health and well-being (physical, mental, and social) of university members	Studies conducted more than 10 years ago
Design and Evaluation	Not restricted	No full text available in the public domain
Research type	Quantitative questionnaire study	Studies focusing exclusively on disease prevalence or unrelated academic outcomes without an explicit connection to well-being

[a]Studies including university members with and without self-reported health conditions were eligible if the primary focus was the assessment of general health and/or well-being domains at the population level rather than disease-specific outcomes.

Information Sources

The following electronic databases were searched for both published and unpublished (gray) literature on the research topic, covering the period from January 1, 2014, to May 31, 2024: PubMed, Cochrane Library, Scopus, Web of Science, and Google Scholar. These databases were chosen as they are suitable for conducting Boolean searches and for the transparent ranking of results. For gray literature, the search included OSF, PrePrints, SSRN, and Preprints with The Lancet, as well as institutional and organizational repositories containing reports and survey data on health and well-being in higher education (eg, international and national public health and higher education agencies).

Gray literature items were assessed using the AACODS (Authority, Accuracy, Coverage, Objectivity, Date, and Significance) checklist to ensure the relevance and quality of nonpeer-reviewed sources.

Search Strategy

The search strategy was structured in accordance with the Population, Concept, and Context (PCC) framework recommended for scoping reviews. The population comprised university members (including students and employees), the concept focused on health and well-being as assessed through surveys, and the context was HEI. On the basis of this framework, search terms were developed and combined using Boolean operators. Derived from our research questions, the initial concepts and the related term labels for constructing the actual search phrase are as follows:

- HEI-related term labels are *university* and *college*
- Term label for members of HEIs, including students (related term labels: *undergraduate*, *graduate and postgraduate student*, *bachelor*, *master*, *doctorandus*, and *PhD student*) and employees (related term labels: *staff*, *personnel*, *workers*, and *teachers*).
- Health (related term labels: *wellness* and *fitness*)
- Well-being (related term labels: *happiness*, *welfare*, and *quality of life*)
- Survey (related term labels: *assessment*, *investigation*, *analysis*, *exploration*, and *questionnaire*)

The Boolean logical buildup of concepts for searching is as follows: *HEI* AND *member* AND (*health* OR *well-being*) AND *survey*

To build a comprehensive and focused search phrase, we took the following steps. First, to find all relevant synonyms and conceptually related terms, we used the Unified Medical Language System, the largest available health-related metathesaurus, containing well over 4.5 million lexical entities. Second, we developed the search strategy using the PCC framework to systematically define and combine key search terms. Third, on the basis of the result of the step "selection of bibliographic sources," we translated the Boolean expression to the syntax required by the bibliographic source. Fourth, we conducted pilot searches in the selected databases and examined the keywords and index terms (if applicable) of the top 30 most relevant papers. This step was an iteration to refine the search phrase and increase its precision; the results are already in the list of labels above. Fifth, we expanded the search phrase as needed, translated it to the source syntax again, and performed the final searches.

It must be noted that concepts are mentioned above by their labels (descriptive words) themselves—used in the search phrasing as well. The first search phrase we tried is shown in Textbox 1.

Initial overview showed that such a search was not specific for intrainstitutional surveys, resulting in tens of thousands of (mostly nonrelevant) papers. After several pilot searches, the search phrase was finalized to the one in Textbox 2, which was used in all the databases selected.

This broad scope of interest, even with this refined search, resulted in more than several thousand studies to screen. Due to time and resource constraints frequently encountered in practice, the screening process was limited to a subset of the highest-ranked records in each database. Accordingly, this scoping review protocol was conducted as a rapid scoping review. Rather than performing an exhaustive screening of all retrieved records, a structured and transparent sampling approach was applied. In each database, the screening was limited to the highest-ranked records returned by the database-specific relevance algorithms. This approach was

XSL•FO
RenderX

chosen to balance feasibility with the objective of identifying the most relevant survey-based studies and is consistent with pragmatic adaptations described in the methodological literature on rapid evidence syntheses. The limitations associated with this approach, including potential algorithmic bias and reduced reproducibility, were acknowledged and considered in the interpretation of the findings.

Textbox 1. Initial search phrase.

- Population: ("member*" OR "student*" OR "employ*" OR "undergraduat*" OR "graduat*" OR "postgraduat*" OR "bachelor*" OR "master*" OR "doctorand*" OR "staff" OR "personnel" OR "workers" OR "teachers" OR "lecturer*")
- Concept: AND ("health" OR "wellness" OR "fitness") AND ("well-being" OR "happiness" OR "welfare" OR "quality of life")
- Context: AND ("higher education institute*" OR "universit*" OR "colleg*")
- Extra (outcome): AND ("survey" OR "assess*" OR "investigat*" OR "analys*" OR "explor*")

Textbox 2. Final search phrase.

- Population: (student OR staff OR employee OR member OR personnel OR worker OR administrator)
- Concept: AND ("health" OR well-being OR wellbeing)
- Context: ("in university" OR "at university" OR "in college" OR "at college")
- Extra (outcome): AND (survey or questionnaire)

Selection of Sources of Evidence (Screening)

To select the sources, the following steps were performed. First, a deduplication process identified and removed duplicate hits from the search results using EndNote 20 (Clarivate). The screening phase then began with an initial assessment based on the titles and keywords of the identified articles. This was followed by a detailed evaluation of the abstracts. Each article was independently screened by 2 reviewers who excluded studies that were clearly outside the predefined scope of the review. Any discrepancies or disagreements were resolved by consensus and in consultation with a senior researcher. The screening phase was completed within 6 weeks.

Data Charting Process (Extraction)

Citations from each search were downloaded into reference manager software (EndNote), where duplicates were identified and removed. To ensure the degree of agreement among independent observers is appropriate, interrater reliability was calculated using Cohen κ, with a threshold of 0.8, using freely available online or other statistical software (Covidence). Interrater agreement was also assessed and reported using percentage agreement, depending on the screening tool applied by the implementing research team. Also, the 0.8 threshold was intended as a recommended target to guide interrater reliability rather than as a rigid exclusion criterion. The reviewers then evaluated the eligible studies and extracted the information in accordance with predefined criteria listed in the Data Items section. Eligible studies were entered into a study database using a standardized data extraction template developed in Microsoft Excel. The data extraction form was pilot-tested on a predefined subset of studies to calibrate the interpretation of data items, and the form was iteratively refined based on insights gained during this pilot phase. To ensure accuracy, the information extracted by each reviewer was compared and discrepancies were discussed to ensure consistency. Disagreements were resolved through discussion and, when necessary, consultation with an independent third-party senior researcher.

Data Items

Each publication was assessed according to the following criteria.

The following topics were covered by the survey:

- Health status characteristics (physical and/or mental health), well-being characteristics, and health behaviors.
- Influencing factors, including individual or family socioeconomic status, living environment, working or studying environment, social networking, work–study life balance, and other relevant factors.

The following methodological characteristics were used for these surveys:

- Data collection methods, including online, face-to-face, and phone-based approaches
- Data collection tool, including paper-pencil or computer-based instruments
- Purpose of data collection, including screening or health assessment
- Target population, including students and/or employees
- Sampling procedure and sample size
- Survey questions, including the use of validated instruments
- Evaluation methods

Textbox 3 presents how the SPIDER framework was operationalized in this review by mapping each SPIDER element to the corresponding data items extracted.

XSL•FO
RenderX

Textbox 3. Operationalization of the SPIDER (Sample, Phenomenon of Interest, Design, Evaluation, Research Type) framework in data extraction.

Sample

- University members (students and/or employees), as defined in the eligibility criteria

Phenomenon of interest

- Health and well-being topics assessed in surveys, including health status, well-being characteristics, health behaviors, and influencing factors

Design

- Survey-based study designs

Evaluation

- Measurement domains and survey instruments extracted under Data Items

Research type

- Survey-based research

Synthesis of Results

As an initial step, the information extracted from the evaluation of the publications was synthesized in a narrative manner and, where relevant, through descriptive statistics. The findings were summarized according to the topics surveyed and the methodology used. The result of this synthesis process was a thematic domain framework of health and well-being constructs assessed across surveys; an inventory of survey instruments and items; and a methodological and implementation feature map describing survey designs, sampling approaches, and modes of administration used in higher education settings.

Results

The first scoping review to which this protocol was applied has been completed, and the results have been published [28]. Drawing on the empirical and methodological insights of that review, this protocol outlines a structured approach for future scoping reviews in this field. It proposes the development of a thematic domain framework, an inventory of survey instruments, and a methodological overview of survey practices in higher education. A PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram will be used to summarize the identification, screening, and inclusion of studies (Figure 1).

XSL•FO
RenderX

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram summarizing the identification, screening, and inclusion of studies in the completed scoping review.

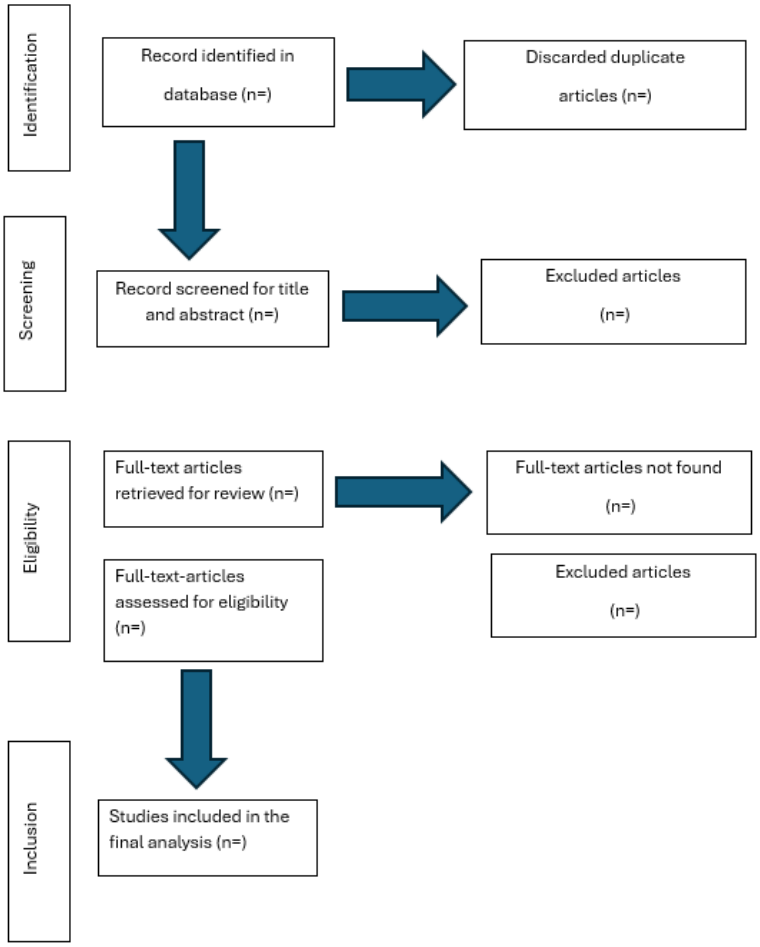

Discussion

Anticipated Findings

This scoping review aimed to bridge gaps in knowledge regarding how health and well-being are assessed through surveys conducted among university members. The protocol follows scoping review methodology and relevant guidelines, ensuring a systematic and comprehensive examination of health and well-being surveys in universities. By clearly defining the scope, eligibility criteria, and data extraction strategy, the protocol supports methodological consistency and reproducibility in the synthesis of health and well-being surveys conducted in universities. As the research addresses a timely and crucial issue, emphasizing the role of universities in promoting health and well-being, the protocol may inform practical applications in the field of university health and well-being assessment. The published scoping review reports the empirical findings of applying this framework [28], while this protocol provides a detailed methodological description intended to support replication and future applications. It is important to note that minor methodological refinements were introduced between the protocol and the published review (eg, SPIDER-based structuring of eligibility criteria, PCC-based structuring of the search strategy, specification of the time frame, explicit framing as a rapid scoping review, inclusion of AACODS criteria for gray literature appraisal, and more detailed procedures for data extraction). These refinements do not alter the underlying methodological approach described in the protocol.

Limitations

A limitation of the protocol is that it restricts its inclusion criteria to publications after 2014 (the past 10 years). Although the 10-year time frame was retained as it reflects a commonly used research practice for defining up-to-date literature in rapidly developing fields, it potentially omits relevant earlier research. Furthermore, focusing solely on English-language publications may introduce a language bias, excluding valuable insights from non-English sources. Restricting inclusion to English-language publications was driven by feasibility and resource constraints, including time limitations and the lack of multilingual screening capacity within the review team. This restriction may result in an overrepresentation of English-speaking and Western European perspectives and an underrepresentation of approaches and survey practices relevant to the Hungarian and Central and Eastern European context. Although artificial intelligence–assisted translation tools could theoretically broaden the scope of included studies, their use entails a risk of translation inaccuracies that require expert linguistic review. In

XSL•FO
RenderX

the absence of sufficient language proficiency and human resources to reliably identify and correct such errors, this approach was not adopted. The decision to limit screening to a subset of records due to resource constraints represents a methodological limitation that may affect the comprehensiveness of the review. These limitations should be considered when interpreting the methodological scope and transferability of the findings. As this review is exploratory in nature and aims to map domains, instruments, and methodological characteristics of survey-based research, the strength of the body of evidence was not formally assessed using tools such as Grading of Recommendations Assessment, Development and Evaluation (GRADE).

Conclusions

This scoping review protocol outlines a systematic approach to mapping health and well-being surveys conducted among university members. By defining clear eligibility criteria, search strategies, and data extraction procedures, the protocol aims to support a transparent and reproducible synthesis of existing survey-based evidence. The findings derived from this protocol are expected to contribute to a clearer understanding of how health and well-being are operationalized and measured in higher education contexts, and to inform future methodological and conceptual developments in this field.

Funding

No funding was received for this study.

Data Availability

All data generated or analyzed during this study were included in the published scoping review article.

Authors' Contributions

All authors participated in the design of the paper. BG, VZ, LB, LL, PV, and JV prepared the first draft. All authors revised the final version of the paper.

Conflicts of Interest

None declared.

References

1. Sedghi F, Mahdizadeh M, Vahedian-Shahroodi M, Gholian-Aval M. Exploring the conceptual framework of a health-promotion faculty from the perspective of members: a qualitative study. BMJ Open 2023 May 11;13(5):e073059. [doi: 10.1136/bmjopen-2023-073059] [Medline: 37169494]
2. Pronk N, Kleinman DV, Goekler SF, Ochiai E, Blakey C, Brewer KH. Promoting health and well-being in healthy people 2030. J Public Health Manag Pract 2021;27(Suppl 6):S242-S248. [doi: 10.1097/PHH.0000000000001254] [Medline: 33278186]
3. Dooris M, Doherty S. Healthy universities--time for action: a qualitative research study exploring the potential for a national programme. Health Promot Int 2010 Mar 18;25(1):94-106. [doi: 10.1093/heapro/daq015] [Medline: 20167825]
4. Dooris M, Powell S, Farrier A. Conceptualizing the 'whole university' approach: an international qualitative study. Health Promot Int 2020 Aug 01;35(4):730-740. [doi: 10.1093/heapro/daz072] [Medline: 31304531]
5. Newton J, Dooris M, Wills J. Healthy universities: an example of a whole-system health-promoting setting. Glob Health Promot 2016 Mar 19;23(1 Suppl):57-65. [doi: 10.1177/1757975915601037] [Medline: 27199018]
6. Okanagan Charter: an international charter for health promoting universities and colleges. NASPA. 2015. URL: https://www.naspa.org/images/uploads/kcs/WHPL_Canon_HP_Okanagan_Charter_12.pdf [accessed 2026-01-14]
7. Sáenz C, Boldú E, Canal J. Health promoting universities: a scoping review. Health Promot J Austr 2025 Apr 06;36(2):e939. [doi: 10.1002/hpja.939] [Medline: 39912192]
8. Gusy B, Lesener T, Wolter C. Time pressure and health-related loss of productivity in university students: the mediating role of exhaustion. Front Public Health 2021 Apr 27;9:653440 [FREE Full text] [doi: 10.3389/fpubh.2021.653440] [Medline: 33987165]
9. Health Promoting Universities: Concept, Experience and Framework for Action. Copenhagen, Denmark: World Health Organization; 1998.
10. Dooris M. Health-promoting higher education. In: Kokko S, Baybutt M, editors. Handbook of Settings-Based Health Promotion. Cham, Switzerland: Springer; 2022:151-165.
11. Hammoudi Halat D, Soltani A, Dalli R, Alsarraj L, Malki A. Understanding and fostering mental health and well-being among university faculty: a narrative review. J Clin Med 2023 Jun 30;12(13):4425 [FREE Full text] [doi: 10.3390/jcm12134425] [Medline: 37445459]
12. Cooke PJ, Melchert TP, Connor K. Measuring well-being: a review of instruments. Couns Psychol 2016 Aug 05;44(5):730-757. [doi: 10.1177/0011000016633507]

XSL•FO
RenderX

13. Suárez-Reyes M, Van den Broucke S. Implementing the Health Promoting University approach in culturally different contexts: a systematic review. Glob Health Promot 2016 Mar;23(1 Suppl):46-56. [doi: 10.1177/1757975915623933] [Medline: 27199017]
14. de Lima MD, Conde JM, Bonito JM. Health promotion in European higher education institutions: an integrative literature review. Public Health Rev 2025 Dec 9;46:1608126 [FREE Full text] [doi: 10.3389/phrs.2025.1608126] [Medline: 41446593]
15. OECD/European Observatory on Health Systems and Policies. Hungary: Country Health Profile 2023. State of Health in the EU. Paris, France: OECD Publishing; 2023.
16. Bencsik A, Csanádi A, Horváth-Csikós G, Juhász T. Az egészségmagatartás jellemzői egyetemista fiatalok körében. Valóság 2018 Oct;61(10):53-67 [FREE Full text]
17. Joó A, Molnár-Tamus V, Kiss J, Kádár A. Tanító szakos hallgatók mentális egészségi állapotának vizsgálata a Debreceni Református Hittudományi Egyetemen. Medárium 2022;XVI(2–3):5-17 [FREE Full text]
18. Elekes Z. Egészségkárosító magatartások és mérési módszerek. Budapesti Corvinus Egyetem. 2011. URL: https://unipub.lib.uni-corvinus.hu/4335/1/646653.pdf [accessed 2026-01-11]
19. Matavovszky D, Nguyen Luu LA, Karner O. Development of an instrument to assess the mental health of university students: validation of the Outcome Questionnaire-45 in a Hungarian sample. Front Psychol 2024 Jan 17;15:1334615 [FREE Full text] [doi: 10.3389/fpsyg.2024.1334615] [Medline: 38298518]
20. Karner O, Kiss M, Szlamka Z, Bernáth S, Kámán P, Lubani A, et al. Magyarországi felsőoktatásban tanuló hallgatók mentális jóllétének felmérése. Felsőoktatási Tanácsadás Egyesület. 2021. URL: https://tinyurl.com/2ncbmkv8 [accessed 2026-01-11]
21. Arnold P, Elekes Z, Felföldi T, Hajdu M, Horváth Á, Kutrovátz K, et al. Egyetemi Hallgatók Alkohol- és Drogfogyasztása, Szabadidő-Eltöltési Szokásai és Jólléte Magyarországon. Budapest, Hungary: Budapesti Corvinus Egyetem; 2024.
22. Bíró É. Az Európai Egészségműveltség Felmérés 2019-2021 magyarországi eredményei. Egészségfejlesztés 2023;64(1):57-59. [doi: 10.24365/ef.9661]
23. Achieving well-being: a global framework for integrating well-being into public health utilizing a health promotion approach. World Health Organization. 2023 Apr 6. URL: https://www.who.int/publications/m/item/wha-76---achieving-well-being--a-global-framework-for-integrating-well-being-into-public-health-utilizing-a-health-promotion-approach [accessed 2026-01-14]
24. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]
25. Peters MD, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Scoping reviews. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, editors. JBI Manual for Evidence Synthesis. Adelaide, Australia: JBI; 2024.
26. The Geneva Charter for well-being. World Health Organization. 2021 Dec 21. URL: https://www.who.int/publications/m/item/the-geneva-charter-for-well-being [accessed 2026-01-14]
27. Making every school a health-promoting school – global standards and indicators. World Health Organization. 2021 Jun 22. URL: https://www.who.int/publications/i/item/9789240025059 [accessed 2026-01-14]
28. Varsányi P, Balkányi L, Gács B, Gyulai A, Lippai L, Rucska A, et al. Health and well-being surveys in higher education: a scoping review. Int J Educ Res Open 2025 Dec;9:100544. [doi: 10.1016/j.ijedro.2025.100544]
29. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018 Oct 02;169(7):467-473. [doi: 10.7326/M18-0850] [Medline: 30178033]
30. Cooke A, Smith D, Booth A. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. Qual Health Res 2012 Oct;22(10):1435-1443. [doi: 10.1177/1049732312452938] [Medline: 22829486]

Abbreviations

AACODS: Authority, Accuracy, Coverage, Objectivity, Date, and Significance
GRADE: Grading of Recommendations Assessment, Development and Evaluation
HEI: higher education institution
OECD: Organisation for Economic Co-operation and Development
OSF: Open Science Framework
PCC: Population, Concept, and Context
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews
SPIDER: Sample, Phenomenon of Interest, Design, Evaluation, Research Type
WHO: World Health Organization

XSL•FO
RenderX

Edited by J Sarvestan; submitted 10.Sep.2025; peer-reviewed by D Perez-Jorge, R Chigevenga, W Costa; comments to author 29.Dec.2025; accepted 26.Feb.2026; published 31.Mar.2026.

Please cite as:
Zsuzsanna V, Balkányi L, Gyulai A, Lippai L, Rucska† AM, Kiss ZS, Varsányi P, Vitrai J, Gács B
Methodology for Assessing University Members' Health and Well-Being and Their Influencing Factors: Protocol for a Rapid Scoping Review of Survey-Based Research
JMIR Res Protoc 2026;15:e83845
URL: https://www.researchprotocols.org/2026/1/e83845
doi:10.2196/83845
PMID:41915417

©Varga Zsuzsanna, László Balkányi, Anikó Gyulai, László Lippai, Andrea Mária Rucska†, Zsuzsanna Soósné Kiss, Péter Varsányi, József Vitrai, Boróka Gács. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 31.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

A Mobile Self-Assessment and Referral Platform for Family Caregivers of Individuals With Alzheimer Disease and Related Dementias: Protocol for a Pilot Randomized Controlled Trial

Francesca Falzarano, PhD; Annabelle Greenfield, BS; Hannah Mason, BS; Katerina Bumbalova, BS; Elizabeth Rojas, MS; Aleksandar Bumbalov, BS

Leonard Davis School of Gerontology, University of Southern California, 3715 McClintock Ave, Los Angeles, CA, United States

Corresponding Author:
Francesca Falzarano, PhD
Leonard Davis School of Gerontology, University of Southern California, 3715 McClintock Ave, Los Angeles, CA, United States

Abstract

Background: Family caregiving for individuals with Alzheimer disease and related dementias (ADRD) is characterized by increasing complexity, intensity, and demand across the disease trajectory. Formal home- and community-based services can provide knowledge, skills, and resources to enhance preparedness and self-efficacy, which may protect against adverse caregiving outcomes; however, awareness and uptake of these services remain low. As caregivers increasingly turn to the internet for information and support in their role, technology offers an opportunity to create a more seamless pipeline between assessment and service referral to match family caregivers with targeted services that meet their specific needs.

Objective: The primary objective of this study is to evaluate the feasibility and acceptability of CarePair—a mobile self-assessment and service referral platform—among ADRD family caregivers. Secondary objectives are to assess the preliminary efficacy of CarePair in reducing stress, depressive symptoms, and anxiety, and enhancing self-efficacy among caregivers randomized to the intervention versus an attention control condition. This study also aims to generate preliminary effect size estimates to inform sample size calculations for a future fully powered randomized controlled trial (RCT).

Methods: This pilot RCT will evaluate the feasibility, acceptability, and preliminary efficacy of CarePair. Eighty ADRD family caregivers will be enrolled and randomized in a 1:1 ratio to the intervention (n=40) or an attention control condition (n=40). Recruitment will be facilitated by the project study site located in an urban metropolitan area of the United States, targeting participants who report residing in and/or being in close proximity to any of the following locations: New York City, Long Island, and Westchester County, New York; Seattle, Washington; and Los Angeles, California. Primary feasibility outcomes include recruitment, retention, and completion rates; website usability; and intervention satisfaction. Exploratory analyses will assess preliminary efficacy on stress, depressive and anxiety symptoms, and self-efficacy.

Results: This trial was funded by the National Institute on Aging in September 2023 and received approval from the institutional review board of the University of Southern California on September 10, 2025. Recruitment began in September 2025 and is scheduled to conclude in May 2026, with data collection scheduled to end in August 2026. As of February 2026, 44 participants have been enrolled and 22 have completed the study.

Conclusions: This pilot trial will offer foundational evidence regarding the feasibility and acceptability of the CarePair intervention. Study findings will determine if “go” criteria are met to warrant the advancement to a larger-scale efficacy trial. Participant insights will also be used to guide intervention refinements and digital platform optimization. By offering a low-burden, caregiver-centered mobile app, CarePair has the potential to facilitate and streamline the timely identification of needs and referral to relevant services for ADRD family caregivers.

Trial Registration: ClinicalTrials.gov NCT06418971; https://clinicaltrials.gov/study/NCT06418971

International Registered Report Identifier (IRRID): DERR1-10.2196/90244

(JMIR Res Protoc 2026;15:e90244) doi:10.2196/90244

KEYWORDS

Alzheimer disease and related dementias; dementia; dementia family caregiving; technology; behavioral intervention

XSL•FO
RenderX

Introduction

Background

Approximately 7.2 million individuals have Alzheimer disease and related dementias (ADRD) in the United States, and 83% of these individuals are reliant on over 12 million unpaid family members and/or friends for basic assistance [1]. While framed as a policy solution to ease strain on the health care system, the lack of viable alternatives for long-term care makes family caregiving unavoidable to meet the needs of the growing ADRD population [2]. Despite their crucial role, the significant physical, psychological, and economic challenges associated with ADRD caregiving significantly increase the risk for a myriad of poor outcomes, including caregiver burden, depression, anxiety, and poor quality of life [3-5]. The unique nature of ADRD requires caregivers to assume increasingly complex and unpredictable responsibilities with disease progression, often learning "on the fly" with little knowledge or preparedness. Such unmet needs also increase the risk for negative care recipient outcomes, including compromised care, poorer well-being, and early long-term care placement [6,7].

Substantial research has examined the caregiving experience to elucidate how family caregivers can be better supported to effectively provide care and to circumvent or delay long-term care placement of the care recipient [7,8]. Findings in this arena have informed numerous intervention efforts aimed at mitigating care-related stress and promoting well-being [9]. For example, multicomponent interventions targeting training, education, and support can improve caregiver outcomes by providing skills to effectively manage care, albeit with small-to-moderate effect sizes [10]. While evidence-based behavioral interventions show promise for improving ADRD caregiver outcomes, the implementation and scalability of such programs in real-world settings remain limited [11]. A major barrier hampering such efforts is the sheer heterogeneity in the family caregiving experience and trajectory, as well as differential outcomes resulting from care-related stress [12]. This variability can limit the effectiveness of interventions designed to alleviate caregiver burden, as a "one-size-fits-all" approach is insufficient for meeting the diverse needs of ADRD family caregivers.

Theoretical Framework

The stress process model [13] provides a theoretical framework to elucidate the dynamic nature of the caregiving experience and the ways by which care-related stress—and appraisals of such stress—differentially influence caregiver outcomes. In addition to background and contextual characteristics, the model is divided into four domains: (1) primary stressors, which are derived directly from the care environment (eg, care recipient's functional and cognitive status); (2) secondary stressors, in which caregiving responsibilities spill over into other life domains (eg, interference with work, family conflicts, and role strains); (3) mediators (eg, support); and (4) outcomes (eg, well-being). The dynamic interplay among these various dimensions of stress differentially shapes caregiver outcomes [14,15]. While prior work underscores the importance of addressing secondary stressors to protect against poor outcomes, they are rarely considered, assessed, or addressed in real-world clinical or community settings [13,16,17].

Further, the model posits that mediating factors, including external (eg, social support) and internal (eg, adaptive coping) factors, can influence the appraisal of care-related stress and subsequently alleviate some of the impact of caregiving stressors. Formal services and supports, such as education, respite, and personal care, can offer knowledge and skill-building opportunities that enhance resilience and adaptive coping. This, in turn, bolsters caregiving self-efficacy, preparedness, and positive caregiving appraisals, thereby promoting quality of care of the care recipient and quality of life among the care dyad [18].

Home- and Community-Based Services

While formal services may enhance caregiver support, actual utilization remains strikingly low in real-world settings. Utilization rates are estimated at 26% and 28% for home care and adult day services, respectively, with home care being the only required home- and community-based service (HCBS) for states participating in Medicaid [19-21]. Findings regarding the uptake of respite (12.7%), skills training (6.3%), and support groups (<5%) further highlight significant underutilization [22]. Barriers contributing to this low prevalence include geographical, occupational, cost, and knowledge-related constraints; cultural incongruence; and uncertainty about services, in addition to information overload by placing the burden on caregivers to seek support [23]. These issues are further compounded by the lack of protocols in most health and long-term care systems for effectively recognizing and integrating caregivers into care delivery processes or for systematically assessing the heterogeneity in caregivers' needs, preferences, and barriers [24].

Our previous work has identified cross-dimensional barriers that impede the identification, access, and actual utilization of formal services and supports [25]. These gaps—described as "nonstarters" to HCBS—exacerbate both primary and secondary stressors, thereby amplifying the likelihood of greater stress and burden. From a structural lens, the siloed, fragmented nature of the service system and the complex bureaucratic structure in which it is embedded contain unclear entry points, which can overall serve as a major deterrent from seeking support, causing those most in need of support to "fall through the cracks" [26]. Despite this pattern, caregivers frequently cite the need for services targeting educational, emotional, and informational support, as well as guidance on navigating the health care system.

Recommendations set forth by The National Academies of Science, Engineering, and Medicine emphasize the need for systematic identification, assessment, and referral to enhance relevant supports to sustain their ability to provide high-quality care [27]. Despite this emphasis, issues related to assessment further impede continuity in the pipeline to HCBS utilization. While there are no shortages of measurement tools to capture caregiver burden, some of which are deemed the "gold standard," from a psychometric perspective, with few exceptions [28,29], these instruments have not been designed to elicit information that can be used to inform appropriate, actionable

XSL•FO
RenderX

referrals [24]. Assessments are further limited by a narrow focus on burden as a single construct and by the time required for clinical administration. In order to achieve its intended purpose in real-world settings, assessments require a holistic understanding of caregiving challenges across a wide range of dimensions such as types of tasks, stress and well-being, and types of training and support needed [22,30]. Effective assessments that account for the multidimensionality of ADRD family caregiving could then be used to facilitate targets for referral to relevant formal services and supports [31].

Technology to Bolster Caregiving Support

While caregivers have expressed a need for personalized guidance, such as formal points of contact, technology is also recognized as a valuable tool with the potential to transform care management, enhancing critical aspects of care such as health monitoring, safety, and communication. As technology becomes increasingly ubiquitous in daily life, it also overcomes barriers to health-related information dissemination to populations previously deemed "hard to reach" [32]. Moreover, technology can support often overlooked aspects of caregiving, including role adaptation, stress reduction, preparedness, and improvement of the overall quality of life and well-being of both caregivers and care recipients [33,34]. As caregivers increasingly turn to the internet for support [25], the potential for digital tools to enhance personalized caregiving support is a promising and cost-effective means to augment support [35,36]. Innovations in technology can facilitate the creation of a more seamless pipeline between caregiver assessment and service referral, thus increasing the potential to bring a "one-stop service" for ADRD caregiving needs to fruition. Further, while the development and implementation of caregiver-focused digital technologies are booming in private industry, research, and clinical settings, the biggest limitation remains the lack of inclusion of family caregivers as co-design partners to ensure that tools are useful, usable, relevant, and aligned with their needs and preferences, which can be considered a major impediment to user adoption and thus scalability. User-centered design approaches are critical for caregiving technology-based interventions to be useful and usable for caregivers, as well as for streamlining pathways to more easily find, access, and utilize formal services.

Among the burgeoning, evidence-based research, tailored support initiatives have emerged; however, many of these initiatives are limited by reliance on an interventionist and a lack of pre-emptive consideration of service use barriers (eg, availability and cost), background characteristics (eg, kinship), and preferences (eg, timing and modality) before referral, which can limit scalability [24,37,38]. Research evaluating the effectiveness of technology-based programs on caregiver outcomes likewise remains limited. Specifically, recent systematic reviews and meta-analyses synthesizing evidence on technology-based behavioral interventions for ADRD caregivers have reported small reductions in depression, anxiety, and caregiver burden, along with modest increases in self-efficacy and quality of life [39-41]. Notably, 1 review found larger effect sizes for mobile-based apps compared to web-based interventions (Hedges g=−0.28) [39]. This pattern is unsurprising, given estimates from the Pew Research Center indicating that 98% of Americans own a smartphone, with high and growing adoption among older adults (90% among those aged 50 - 64 years and 78% among those aged 65 years or older) [42].

At the same time, these reviews highlight important limitations in the existing literature, with most studies categorized as having low-to-moderate quality. Substantial variability in intervention content and delivery, along with inconsistent methodologies, outcome measures, and reported effects, further complicates ascertaining the mechanisms underlying intervention effects, which limits the ability to draw clear conclusions about effectiveness [39,43]. Nonetheless, this limited evidence suggests that multicomponent interventions are most promising, particularly those offering some combination of components, such as training, education, professional support, and relaxation strategies, with interventions that include a personalization component yielding the strongest results [41,43]. Despite the abundance of emerging caregiving-focused digital solutions across both research and industry, few approaches, if any, integrate psychological theory with user-centered design principles to guide both technology and intervention development and delivery in order to combine assessment with referral. Such interdisciplinary approaches are critical for ensuring usability and acceptability and for engaging key stakeholders throughout the development process, all of which are key predictors of adoption.

Our Study

Given the need for systematic identification, assessment, and referral to ensure that caregivers have the skills and preparedness required to sustain high-quality care and considering the gaps in existing research and intervention efforts, the purpose of this study is to examine the feasibility, acceptability, and preliminary efficacy of technology-facilitated caregiver assessment and referral. CarePair is a mobile app consisting of a personalized self-assessment and service referral platform for ADRD family caregivers that applies machine learning algorithms to match caregivers' needs with relevant services and supports while proactively accounting for barriers (eg, cost and modality) and individual preferences (eg, timing and distance). By integrating theory, user-centered design, and principles of psychological science, this study aims to evaluate the potential of this novel, integrative approach to improve well-being and quality of life among a diverse population of ADRD family caregivers.

Aims and Objectives

Our overarching aim is to outline the protocol for a pilot randomized controlled trial (RCT) of the CarePair intervention for ADRD family caregivers. CarePair is designed to enhance awareness and utilization of formal services and supports by directing caregivers to resources that align with their current needs, with the overall goal of promoting caregiver mental health and well-being.

Primary Objective

The primary objective is to evaluate the feasibility and acceptability of the CarePair intervention among ADRD caregivers.

XSL•FO
RenderX

Secondary Objective

The secondary objective is to assess the preliminary efficacy of CarePair in reducing stress, reducing depressive and anxiety symptoms, and enhancing self-efficacy among ADRD family caregivers.

Exploratory Objective

The exploratory objective is to generate preliminary effect size estimates to inform sample size calculations for a future fully powered RCT.

Methods

Study Design

This protocol paper follows the SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) reporting guidelines for pilot protocols. The CarePair SPIRIT checklist is presented in Checklist 1.

This single-site, unblinded pilot RCT will evaluate the feasibility, acceptability, and preliminary efficacy of the CarePair intervention. Eighty ADRD family caregivers will be recruited and randomly assigned to either the intervention group (CarePair) or an attention control group and will participate for a duration of 6 weeks. Intervention participants will receive access to the CarePair app, while control participants will receive access to a digital folder of caregiving articles via the Digify online platform (described below).

Study Setting

Study activities will be conducted virtually and led by researchers at the University of Southern California (USC) in Los Angeles, California. All study activities, including informed consent, intervention delivery, and outcome assessment, will be conducted virtually. Data will be collected and managed using Research Electronic Data Capture (REDCap), a secure, Health Insurance Portability and Accountability Act (HIPAA) compliant electronic data management system hosted at USC [44,45]. Assessments will be administered at baseline and follow-up at weeks 2, 4, and 6 (postintervention). Intervention group participants will also be invited to complete an optional semistructured interview via Zoom (Zoom Communications) to provide feedback about their experiences and provide suggestions for improvement.

Eligibility

Inclusion Criteria

To participate, individuals must (1) be the primary caregiver of a community-dwelling family member or friend with ADRD; (2) provide at least 10 hours of care per week for at least 6 months; (3) be 18 years of age or older; (4) reside in and/or report being in close proximity to any of the following locations: New York City, Long Island, and Westchester County, New York; Seattle, Washington; and Los Angeles, California; (5) have regular internet access; (6) have regular access to a smartphone and proficiency using a smartphone; (7) speak English fluently; (8) not report cognitive impairments (assessed via self-report); and (9) be willing and able to provide informed consent.

Exclusion Criteria

Individuals will be excluded if they (1) are not the primary caregiver of a community-dwelling person with ADRD (eg, bereaved or care recipient resides in memory care); (2) provided care for less than 6 months and/or for fewer than 10 hours per week; (3) are under 18 years of age; (4) do not reside in and/or self-report being near one of the following locations: New York City, Long Island, and Westchester County, New York; Seattle, Washington; and Los Angeles, California; (5) lack regular internet access; (6) lack a smartphone and/or are unable to operate it; (7) are not fluent in English; (8) self-report cognitive impairments; and (9) are unable or unwilling to provide informed consent.

Sample Size

This study proposes to enroll 80 ADRD family caregivers, randomized in equal numbers (n=40 per group) to either the intervention or attention control conditions. Consistent with stages 0-1B of the National Institutes of Health (NIH) stage model for behavioral intervention development [46], the primary goal of this pilot study is to establish the feasibility and acceptability of CarePair. Accordingly, the study is not powered to detect clinically meaningful differences in outcomes. A sample size of 40 participants per arm exceeds the upper limit recommended (n=35) for feasibility trials [43] and allows for adequate precision in estimating key feasibility parameters (recruitment, retention/attrition, and platform usability), which will inform prespecified progression (go/no-go) criteria for advancing to a full-scale RCT. This approach is consistent with methodological guidance for pilot and feasibility studies, which prioritizes the estimation of feasibility parameters over formal hypothesis testing or power-based sample size calculations [47-50]. Preliminary efficacy will be explored as a secondary aim to estimate effect sizes and associated CIs to guide the design of a subsequent, fully powered efficacy trial.

Intervention Description

This study evaluates technology-facilitated ADRD caregiver assessment and referral. CarePair is a mobile app comprised of a self-assessment and service referral platform designed to support ADRD family caregivers by connecting them with formal services and resources tailored to their unique needs and preferences. Upon enrollment, caregivers will complete a brief 12-item survey that evaluates unmet needs across various care-related domains (eg, mental health, legal guidance, and social support). Responses are used to algorithmically match caregivers with content-relevant resources (eg, blogs and educational materials), as well as community services offered both virtually and in-person (eg, support groups, respite care, and financial education) using Zip Code application programming interface functionality to tailor location-based recommendations according to their reported assessment scores and indicated preferences (eg, cost and service timing).

CarePair's design and development followed a user-centered design framework to ensure the intervention reflected the priorities, experiences, and feedback of ADRD caregivers. Prior to the commencement of technology development, a stepwise, mixed-methods approach was utilized to identify caregivers'

XSL•FO
RenderX

most pressing service-related needs. In study 1, conducted from June 2022 to August 2023, 63 participants completed a survey battery of caregiving measures assessing stress and psychosocial well-being to elucidate the most common unmet needs reported by ADRD caregivers. A subgroup of this sample (n=34) participated in one-on-one semistructured interviews exploring their caregiving experiences, including discussion of any family support, the diagnosis process, their current needs, and the utilization of HCBSs, as well as the potential barriers and facilitators of HCBSs. In study 2, data gathered during study 1 activities were used to inform 5 focus group sessions comprised of ADRD family caregivers (n=13) and subject matter experts (SMEs; n=17; eg, care consultants and dementia trainers), which were conducted in April 2023 to gather a holistic, multiperspective understanding of the major barriers and facilitators of service use, as well as potential solutions [25].

Data collected in the preceding fact-finding phases of the study informed the content, design, and functionality of an initial CarePair prototype. Version 1.0 of the CarePair mobile app underwent several rounds of refinement to enhance usability and accessibility based on feedback from the project team. Following internal review, semistructured interviews were conducted (25 participants, 15 caregivers, and 10 SMEs) from August to October 2024 to gather comprehensive feedback related to the app's content, design, usability, and functionality, which informed further refinements to the platform. This stakeholder-informed development and iterative-feedback process helped ensure that the intervention was contextually relevant, acceptable, and usable prior to launching the RCT phase of the project.

Study Conditions

Participants will be randomly assigned to the intervention group or the attention control (Digify) group. Both groups will log into their assigned platform (CarePair or Digify) twice per week and read 1 article per log-in (12 articles total across the intervention) for the 6-week study period. Participants will access their respective platform via their personal smartphone device.

Intervention Group

Participants assigned to the intervention group will have access to the CarePair app. Over the 6-week study period, they will be instructed to log in twice per week to read and engage with 1 article per log-in session (eg, by rating, commenting, saving, or taking notes). Participants will also be encouraged, but not required, to explore and engage with service referrals available through the platform. Backend data will be collected to measure app engagement, including participants' number of log-ins, time spent engaging with materials, specific content accessed, and feature interactions.

The onboarding process will direct caregivers to create a profile for themselves and their care recipients, choose an avatar, and select service preferences (eg, modality, cost, distance based on zip code, and time of the day). Users will also be prompted to complete a 12-item assessment (administered at pre- and postintervention)—the content of which was informed by studies 1 and 2 described above—to measure the primary domains caregivers reported needing assistance with. Each item begins with the stem "In your role as a caregiver, how much support do you need with…" followed by items such as "Understanding your care recipient's brain change?" and "Dealing with your own feelings and worries?" Response options range from 1 (no need) to 4 (high need).

Following the initial onboarding session, users will be provided with an in-app tutorial that introduces the platform's primary features (see Figure 1 for screenshots of the CarePair app interface). The *Home* page provides a personalized feed of resources curated algorithmically based on the user's reported needs and preferences. The *Services* tab includes a catalog of local and virtual services tailored to caregivers' profile preferences and assessment responses gathered during onboarding. The *Library* tab allows participants to browse through the app's vetted ADRD- and caregiving-related resources, thematically organized by domain and relevance. Both the *Services* and *Library* pages allow users to add filters by category (eg, social support, legal, education, and grief).

XSL•FO
RenderX

Figure 1. CarePair app interface.

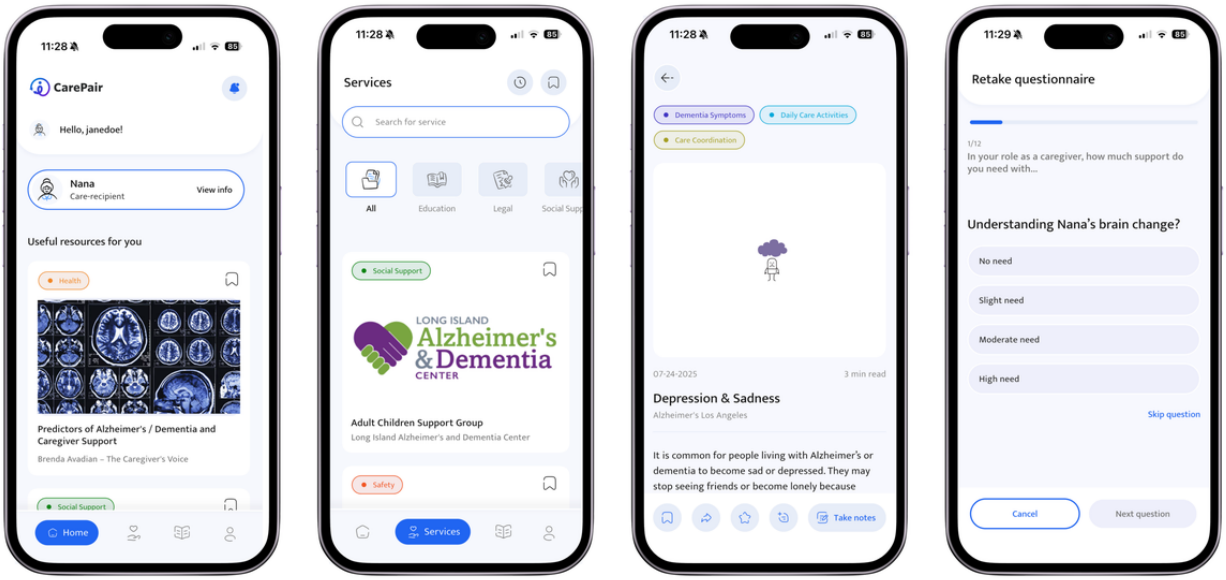

Further, caregivers can interact with resources by commenting, rating, saving, sharing, or note-taking directly in the app. Comments and ratings are visible to other caregiver users. To protect confidentiality, users are instructed during onboarding to create pseudonymous usernames and avoid sharing personally identifiable information. Research staff review account registrations and may modify usernames if they contain personally identifiable information. User-generated content is routinely monitored by research staff. Comments will be removed if they include identifiable information, misinformation, promotional content, or potentially harmful or inappropriate language. These procedures are intended to reduce privacy risks and support a safe user environment. Finally, caregivers can edit their profiles and needs assessment responses, manage settings, review notes and saved resources, and access a directory of national ADRD-focused organizations in the *Account* tab.

All information related to services and resources will be maintained by the study team via a secure database repository. Within the app, caregivers may also flag content that is inaccurate, inaccessible, or outdated for review by the investigative team.

Control Group

Participants in the attention control group will access a digital folder of articles via Digify, a secure document-sharing platform that allows for controlled access of materials and user-engagement tracking on metrics, including log-ins, articles read, and time spent viewing articles. This folder contains most articles available in CarePair, all focused on ADRD and caregiving topics (3 articles with video components were excluded as these formats are not supported by Digify), but they are not organized by relevance or tailored to caregivers' needs. Control participants follow the same engagement schedule as participants in the intervention group (ie, 2 log-ins per week for 6 weeks; read 1 article per session). Control activities are designed to match the intervention condition in terms of time commitment and interface interaction, and to ensure consistent engagement across groups while allowing for the isolation of the impact of CarePair's delivery and personalization features. During onboarding, participants are provided with a deidentified username and password to access the Digify platform, along with step-by-step instructions on the use of the platform.

Recruitment

Sample recruitment will be facilitated by the project's study site using a multipronged outreach strategy targeting ADRD caregivers in the following locations: New York City, Long Island, and Westchester County, New York; Seattle, Washington; and Los Angeles, California. Recruitment strategies will include: (1) posting study flyers in high-traffic areas in the abovementioned cities; (2) study announcements made through community organizations; (3) ResearchMatch, a NIH-funded online registry that connects research volunteers with institutional review board (IRB)–approved studies; (4) digital outreach, including social media advertisements on platforms (eg, Facebook and Instagram); and (5) outreach to organizations provided as resources within the CarePair app. Additionally, individuals who have participated in prior studies and have provided consent for recontact will be approached to gauge interest and eligibility. Snowball sampling through community networks, social media, and currently enrolled participants will also be encouraged. All interested individuals will be directed to complete the study's screening survey to determine eligibility.

Ethical Considerations

This study was approved by USC's IRB on September 10, 2025 (protocol#: UP-25 - 00690). The study IRB approval letter can be found in Multimedia Appendix 1. The trial is registered on ClinicalTrials.gov (NCT06418971). The trial registration record

XSL•FO
RenderX

includes all items from the World Health Organization trial registration dataset (Multimedia Appendix 2), including the primary registry name, trial identifying number, key eligibility criteria, target sample size, primary outcomes, secondary outcomes, funding sources, and sponsor details.

In accordance with the International Council for Harmonisation–Good Clinical Practice E6(R3) guidelines and applicable local regulations, trained IRB-approved research assistants (RAs) will obtain electronic informed consent via REDCap's e-Consent platform after confirming participant eligibility.

During the virtual consent meeting, RAs will review the study purpose, study procedures, and participant rights (eg, the right to withdraw at any time without penalty) to support participant comprehension, and will encourage questions at any time. They will also explain what participation entails, including time commitments, responsibilities, and potential risks. As consent is an ongoing process, participants will be promptly informed of any protocol changes that could affect their willingness to continue. After the consent meeting, participants will be emailed a copy of the informed consent form signed by both themselves and the consenting RA. See Multimedia Appendix 3 for the informed consent form.

All study personnel are trained in human participants' protection, confidentiality, and recruitment procedures. Data collection, sharing, and storage will comply with HIPAA and IRB regulations. Each participant will be assigned a unique study identifier, and all data will be stored in a secure, password-protected REDCap database accessible only to IRB-authorized personnel. To maintain confidentiality, all datasets will remain deidentified throughout the study, including during analysis and dissemination.

Participants will receive US $25 for each completed survey and US $20 for an optional interview, for a maximum total of US $70. All payments will be sent as Amazon digital gift cards.

Data Management

All study data will be stored on a secure, password-protected server with multifactor authentication, accessible only to IRB-authorized study personnel. Supported by the Southern California Clinical and Translational Science Institute (CTSI), REDCap [44,45] will be used for participant tracking, survey data collection, and consent procedures. As a secure, web-based data management platform, REDCap enables the development of customized data management and capture systems with online entry forms, reporting tools, automated export procedures for seamless data download to common statistical packages, and procedures for importing data from external sources. This database system is also utilized for the provision of data to the IRB, NIH, and US Food and Drug Administration and includes robust security features, including role-based access controls, Institutional Lightweight Directory Access Protocol and Shibboleth single sign-on (SSO) authentication, and a complete audit trail of data entry and export activities. REDCap is hosted on CTSI-managed servers, which are backed up daily and support encrypted (SSL-based) connections. The platform is maintained nationally through a multi-institutional consortium led by the Vanderbilt University Clinical and Translational Science Award program.

Randomization Process

After providing consent and completing the baseline survey, participants will be randomized to either the intervention or attention control group using REDCap's randomization module. A stratified permuted block randomized design is utilized to enhance internal validity and minimize potential confounding variables. Randomization tables are generated by an RA, who is not involved in data collection or participant randomization, using a custom Python script.

To ensure demographic balance across study arms while accommodating the target sample size (n=80), randomization is stratified based on caregiver sex and race/ethnicity, resulting in 24 unique strata. Within each stratum, 25 permuted blocks of 4 are created, each containing 2 "intervention" and 2 "control" assignments, randomly shuffled using Python's random shuffle() function to ensure unpredictability. The full table is compiled into a .csv file and uploaded to REDCap's secure randomization module. Once uploaded, the allocation table is locked and inaccessible (ie, no longer visible or modifiable) to study staff. Participants will be randomized in real time via REDCap after stratification variables are entered (sex and race) to preserve allocation concealment and ensure balanced/methodologically sound group assignment. Once randomized, study staff are unable to change participants' group assignments.

Assessments and Outcomes

Prospective participants will complete a brief eligibility screening questionnaire in REDCap, which includes basic demographic items (eg, age and sex) and questions aligned with study inclusion/exclusion criteria (eg, "Do you speak English fluently?"). Study assessments will be administered electronically via REDCap at 4 time points: baseline, brief check-in surveys at the end of weeks 2 and 4, and 6-week follow-up. Intervention participants will also be invited to complete an optional postintervention feedback interview via Zoom or telephone.

A description of study measures, including assessment time points, is shown in Table 1. See Figure 2 for the enrollment and intervention schedule.

XSL•FO
RenderX

Table . Description of survey instruments based on the stress process model [13].

Domain and measure	Time points			
	S[a]	BL[b]	Weeks 2 and 4	FU[c]
Background and context				
Sociodemographic characteristics				
Single-item assessments (age, sex, race/ethnicity, education, geographic location, income, and kinship to care recipient)	✓	—[d]	—	—
Technology proficiency				
Mobile Device Proficiency Questionnaire-16 [51]	—	✓	—	—
Primary stressors: objective				
Functional dependency				
Katz Activities of Daily Living Scale [52,53]	—	✓	—	✓
Lawton-Brody Instrumental Activities of Daily Living Scale [54]	—	✓	—	✓
Caregiver Medical and Nursing Task Inventory [55,56]	—	✓	—	✓
Cognitive impairment				
Revised Memory and Behavior Problems Checklist [57]	—	✓	—	✓
Primary stressors: subjective				
Relational deprivation				
Relational Deprivation Scale [13]	—	✓	—	✓
Secondary stressors: role and intrapsychic strains				
Family conflict				
Family Conflict Scale [13,58]	—	✓	—	✓
Employment conflict				
Job-Caregiving Conflict Scale [13]	—	✓	—	✓
Mood				
Positive and Negative Affect Schedule [59]	—	—	✓	—
Preparedness				
Preparedness for Caregiving Scale [60]	—	✓	—	✓
Cost of care				
Cost of Care Index–Economic Costs Subscale [61]	—	✓	—	✓
Surrogate decision-making				
Advance Care Planning Engagement Survey-9 [62,63]	—	✓	—	✓
Self-care				
Self-Care of Informal Caregivers Inventory [64]	—	✓	—	✓
Role confusion				
Role Captivity Scale	—	✓	—	✓
Predeath grief				

XSL•FO
RenderX

Domain and measure	Time points			
	S[a]	BL[b]	Weeks 2 and 4	FU[c]
Prolonged Grief Disorder-12 Caregiver Version [65,66]	—	✓	—	✓
Caregiver burden				
Zarit Burden Interview-12 [67]	—	✓	—	✓
Caregiving rewards				
Positive aspects of caregiving [68]	—	✓	—	✓
Mediators				
Caregiver needs				
CarePair needs assessment[e]	—	✓	—	✓
Social support				
Lubben Social Network Scale-6 [69]	—	✓	—	✓
Formal support				
Negative service attitudes and experiences [70]	—	✓	—	✓
Satisfaction with the types and quality of services [71]	—	✓	—	✓
Prior use of services, willingness to use services, and types of services used[e]	—	✓	—	✓
Primary outcomes				
Feasibility				
Recruitment: number/proportion screened, eligible, consented/enrolled, declined, and retained [72]	—	✓	—	✓
Usability: System Usability Scale (SUS) [73,74]; user version of the Mobile Application Rating Scale (uMARS) [75]	—	—	—	✓
Intervention adherence: app usage (measured via number of log-ins); feature usage (measured via number of interactions with resources [eg, saves and notes])	—	—	—	✓
Acceptability				
Satisfaction and acceptability survey[e]	—	—	—	✓
Secondary outcomes				
Depression				
Patient Health Questionnaire-9 [76]	—	✓	—	✓
Patient Health Questionnaire-2 [76,77]	—	—	✓	—
Anxiety				
Generalized Anxiety Disorder-7 [78]	—	✓	—	✓
Generalized Anxiety Disorder-2 [78]	—	—	✓	—
Caregiver stress				
Kingston Caregiver Stress Scale [79,80]	—	✓	—	✓
Self-efficacy				
Revised Scale for Caregiving Self-Efficacy [81]	—	✓	—	✓
Quality of life				

XSL•FO
RenderX

Domain and measure	Time points			
	S[a]	BL[b]	Weeks 2 and 4	FU[c]
Satisfaction with Life Scale-5 [82]	—	✓	—	✓
Satisfaction with Life Scale-1 [82,83]	—	—	✓	—

[a]S: screening survey.

[b]BL: baseline survey.

[c]FU: 6-week follow-up survey.

[d]Not applicable.

[e]Developed for this study.

XSL•FO
RenderX

Figure 2. SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) checklist: CarePair trial assessments and timepoints. *Intervention group only.

	Study period				
	Enrollment	Baseline	Postallocation		Follow-up
Timepoint	$-t_1$	t_1	Week 2	Week 4	Week 6
			t_2	t_3	t_4
Enrollment					
Eligibility screening	✓				
Informed consent	✓				
Allocation		✓			
Interventions					
CarePair intervention			←	—	→
Attention control			←	—	→
Assessments					
Screening questions	✓				
Sociodemographic characteristics	✓				
CarePair needs assessment		✓			✓
Mobile Device Proficiency Questionnaire-16		✓			
Katz Activities of Daily Living Scale		✓			✓
Lawton-Brody Instrumental Activities of Daily Living Scale		✓			✓
Caregiver Medical and Nursing Task Inventory		✓			✓
Preparedness For Caregiving Scale		✓			✓
Zarit Burden Interview-12		✓			✓
Kingston Caregiver Stress Scale		✓			✓
Role captivity		✓			✓
Job-caregiving conflict		✓			✓
Family Conflict Scale		✓			✓
Revised Memory and Behavior Problems Checklist		✓			✓
Prolonged Grief Disorder-12		✓			✓
Patient Health Questionnaire-9		✓			✓
Patient Health Questionnaire-2			✓	✓	
Generalized Anxiety Disorder-7		✓			✓
Generalized Anxiety Disorder-2			✓	✓	
Relational deprivation		✓			✓
Positive aspects of caregiving		✓			✓
Satisfaction With Life Scale-5		✓			✓
Satisfaction With Life Scale-1			✓	✓	
Revised Scale for Caregiving Self-Efficacy		✓			✓
Self-Care of Informal Caregivers Inventory		✓			✓
Lubben Social Network Scale		✓			✓
Cost of Care Index		✓			✓
Advance Care Planning Engagement Survey-9		✓			✓
Satisfaction with the types and quality of services		✓			✓
Negative service attitudes and experiences		✓			✓
Positive and Negative Affect Schedule			✓	✓	
Prior formal service use		✓			✓
System Usability Scale*					✓
Mobile Application Rating Scale: User Version*					✓
Satisfaction assessment*					✓
Qualitative feedback interview*					✓

Primary Outcomes

Feasibility

Feasibility will be evaluated across recruitment, retention, and app usability. *Recruitment feasibility* [72] will be assessed by tracking the number and proportion of caregivers who (1) are eligible and consent, (2) are eligible and decline participation, and (3) are screened and found eligible. These metrics will be recorded in REDCap using timestamped research logs and will be used to calculate consent rates (eg, the number of consented

XSL•FO
RenderX

individuals divided by the number of interested individuals). *Retention* will be examined via completion and attrition rates for the full sample and by study arm. We will calculate the number of completed assessments per participant divided by the total number of assessments, with attrition rates calculated as the proportion of noncompleters among enrolled participants. A designated data manager will oversee the collection, monitoring, and reporting of feasibility data.

App feasibility will be assessed through backend usage analytics (eg, number of log-ins and in-app interactions such as saving or reviewing resources). To assess *usability*, participants will complete a 10-item modified version of the System Usability Scale (SUS) [73,74], which measures usefulness, ease of use, and satisfaction at the posttest phase. Sample items include "I thought the CarePair app was easy to use" and "I felt very confident using the CarePair app." Each item is rated on a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]). Total scores are converted to a range from 0 to 100, with higher values indicating greater usability. At follow-up, participants will complete the 20-item user version of the Mobile Application Rating Scale (uMARS) [75], which evaluates app quality across 5 domains: engagement, functionality, esthetics, information, and subjective quality. Items such as "Is the app content (visuals, language, and design) appropriate for the target audience?" are rated on a 5-point Likert scale (1 [inadequate] to 5 [excellent]), with higher scores indicating higher perceived quality. These metrics will inform feasibility across recruitment, engagement, retention, and app usability, as well as guide modifications for a future large-scale trial.

Acceptability

Acceptability will be assessed at the 6-week follow-up using an 8-item survey measuring user satisfaction, usefulness, and ease of use. Participants will rate their agreement with statements such as "I was satisfied with the features in the CarePair application" and "Using the CarePair application was worth my time" on a 5-point Likert scale (1 [strongly disagree] to 5 [strongly agree]). In addition, CarePair intervention participants will complete 4 open-ended questions providing qualitative feedback on their experiences, including which features were engaging, helpful, or in need of improvement. Sample prompts include the following: "Which CarePair features did you find most valuable? Why?"

Secondary Outcomes

Caregivers' Needs

Caregivers' needs will be assessed and compared at baseline and follow-up using a 12-item needs assessment developed for this study. The CarePair Needs Assessment measures caregivers' needs across a variety of caregiving domains, such as social support, legal guidance, and education. Questions, including "In your role as a caregiver, how much support do you need with connecting with others who are in similar care situations?" are rated from 1 (no need) to 4 (high need). Scores range from 12 to 48, with higher scores indicating greater need.

Self-Efficacy

Caregiving self-efficacy will be measured at baseline and follow-up using the 15-item Revised Scale for Caregiving Self-Efficacy [81]. Items assess caregivers' confidence in managing their own lives in addition to care-related responsibilities (eg, "How confident are you that you can ask a friend/family member to stay with [care recipient] for a week when you need time for yourself?"). Responses are rated from 0 (cannot do at all) to 100 (certainly can do). Total scores range from 0 to 100, with higher scores indicating greater perceived self-efficacy (Cronbach α=0.80).

Anxiety

Anxiety symptoms will be assessed using the Generalized Anxiety Disorder-7 (GAD-7) [78]. Participants report how frequently they experienced symptoms (eg, worrying and difficulty relaxing) in the past 2 weeks (eg, "Over the last 2 weeks, how often have you been bothered by... feeling nervous, anxious, or on edge?"). Items are rated on a 4-point scale (0 [not at all] to 3 [nearly every day]). Total scores range from 0 to 21, with higher scores indicating greater anxiety (Cronbach α=0.92).

Depression

Depressive symptoms will be measured using the Patient Health Questionnaire-9 (PHQ-9) [76]. Items capture the severity of symptoms (eg, low interest and fatigue) over the past 2 weeks (eg, "Over the last 2 weeks, how often have you been bothered by... feeling tired or having little energy?"). Responses are rated on a 4-point scale (0 [not at all] to 3 [nearly every day]). Total scores range from 0 to 27, with higher scores indicating greater depressive symptoms (Cronbach α=0.89).

Stress

Caregiver stress will be evaluated using the 10-item Kingston Caregiver Stress Scale (KCSS) [79,80]. Items measure care-related stress (eg, "To what extent do you have concerns regarding the future care needs of [care recipient]?"), with ratings on a 5-point Likert scale (1 [no stress] to 5 [extreme stress]). Total scores range from 20 to 50, with higher scores reflecting greater caregiver stress (Cronbach α=0.85).

Quality of Life

Subjective quality of life will be assessed using the 5-item Satisfaction With Life Scale (SWLS) [82]. Items ask participants to rate agreement with statements, such as "In most ways, my life is close to my ideal," on a 7-point scale (1 [strongly disagree] to 7 [strongly agree]). Scores range from 5 to 35, with 20 representing a neutral midpoint and higher scores indicating greater life satisfaction (Cronbach α=0.87).

Retention and Adherence

A structured, multipronged retention and adherence protocol will be implemented at enrollment to promote engagement, minimize loss to follow-up, and ensure complete data collection. Research staff will be trained to respond sensitively to participant needs, including addressing technical issues, emotional strain, and other barriers to engagement.

XSL•FO
RenderX

To reduce attrition and enhance adherence, participants will receive alternating end-of-week emails from their designated RAs, including check-ins at weeks 1, 3, and 5, and brief check-in surveys at weeks 2 and 4. Engagement will be monitored using CarePair usage analytics (eg, log-in frequency and feature usage), which will guide the content of check-in emails. Participants who have logged in will be contacted to assess their experience, address concerns, and offer optional one-on-one support via Zoom or telephone. Those with little or no engagement will receive reminder emails with encouragement and instructions to participate.

Consistent follow-up will be conducted for participants who do not complete the required study activities (eg, surveys and consent). Weekly reminder emails will be sent in the first 3 weeks of no action, followed by a final phone reminder in the fourth week. Participants who remain unresponsive will be considered lost to follow-up. This structured approach is designed to reinforce expectations, identify participation barriers, and provide personalized support to facilitate continued participation.

Retention will also be supported by incentives. Participants will receive a US $25 Amazon gift card after completing the baseline survey and another US $25 gift card after the 6-week follow-up (after completing all required activities). The fully virtual format further reduces barriers, such as transportation and caregiving responsibilities, making participation more accessible.

Control group activities are designed to match with intervention group activities in terms of interaction frequency and engagement level, minimizing the risk of differential dropout between conditions. All instances of nonadherence or withdrawal, along with reported reasons, will be systematically documented in REDCap to inform feasibility and guide refinements for future trials. While attrition rates of 15% to 30% are common in pilot RCTs and often higher in online interventions [84,85], the study team will monitor retention metrics regularly and, if needed, adjust communication strategies or provide more support to sustain engagement.

Intervention Fidelity

A comprehensive Manual of Operating Procedures has been developed to guide all aspects of training, intervention delivery, and data management. Standardized protocols are in place for participant screening, enrollment, tracking, and communication. All research staff are trained using consistent materials on study procedures, and ongoing team meetings are held to address any issues that may arise with regard to data collection, participant engagement, and CarePair app functionality.

Intervention activities and group assignments will be fixed across participants to ensure consistency. All participants will follow a standardized 6-week study schedule with assessments at predetermined time points. Scripted email templates will be used for all participant communications to maintain uniformity and proactively address potential concerns or barriers to engagement.

To standardize onboarding, intervention participants will receive standardized training materials, including a PDF guide and brief instructional videos on how to use CarePair. The CarePair app's interface, content, and features will be the same across all intervention participants; however, its organization will be tailored by machine learning algorithms to each participant's needs and preferences.

Harms

Given the minimal-risk nature of the study, adverse events (AEs)—defined as any unfavorable or unintended experience associated with participation, whether related to the study or not—are expected to be rare. Potential risks include mild emotional discomfort when responding to survey or interview questions, exposure to ADRD-related content, or minor frustration when navigating digital platforms. There is also a minimal risk of breach of confidentiality. Participants may skip any questions or activities that feel uncomfortable or distressing. Study staff are trained to recognize and respond to participant distress and will provide appropriate referrals, including to the USC Family Caregiver Support Center, for additional support and resources as needed.

All AEs, serious adverse events (SAEs), and unanticipated problems, as well as any protocol deviations, will be documented using designated reporting forms and logs. The principal investigator (PI) will be notified of any AE or SAE within 24 hours, and all required reporting to the USC IRB, designated Safety Officer, and National Institute on Aging (NIA) Program Officer will be conducted in accordance with institutional and federal policies.

Monitoring

Data Monitoring Committee

Given the study's minimal risk, a formal Data Monitoring Committee is not required for this pilot trial. In line with institutional oversight policies, the study team is required to submit an annual continuing review to serve as a postapproval monitoring report to USC's IRB, which includes updates on study progress, protocol adherence, and any unanticipated issues and/or AEs. This reporting ensures ongoing compliance with ethical and regulatory standards throughout the study period.

Trial Monitoring

No interim analyses or stopping guidelines are planned due to the study's short duration and minimal risk for participants. The study will follow a structured data and safety monitoring plan. The PI will oversee day-to-day study conduct, data integrity, and participant safety, and the study team will continuously monitor trial progress and data collection to ensure protocol fidelity and protection of human subjects. A designated Safety Officer—independent of the investigative team and study site—will serve in an advisory capacity to the PI and NIA Program Officer. This individual will evaluate the overall progress of the trial, including participant recruitment, data quality, adherence to study procedures, and participant safety.

Safety monitoring reports will be generated every 3 months beginning at the onset of the trial. These reports will summarize aggregate data on participant accrual, retention, demographic characteristics, protocol deviations, and adverse or unanticipated events. Reports will also include analyses of data completeness, timeliness, and overall study progress. The Safety Officer will

XSL•FO
RenderX

review safety reports quarterly and provide written feedback and recommendations to the PI and NIA Program Officer regarding the continuation, modification, or termination of the study.

Statistical Analysis

Descriptive statistics (eg, frequencies, means, and SDs) will be used to characterize the study sample across arms. The intervention's feasibility and acceptability will be evaluated through recruitment and retention rates, reasons for noncompletion, app usage data (eg, number of log-ins and in-app actions), and participant usability and satisfaction assessments. Independent samples *t* tests for continuous variables and chi-square tests for categorical variables will be used to test for baseline group differences in sociodemographic characteristics and primary and secondary outcome measures (eg, stress and self-efficacy). Relationships among continuous measures will be examined using Pearson correlations. Additional analyses will examine the relationships between covariates (eg, age, education, and kinship), and outcome measures will be examined for inclusion in multivariate models.

To examine preliminary efficacy, within- and between-group differences in secondary outcomes will be assessed. Specifically, we will examine (1) mean values at each time point by group, (2) within-group change from baseline to follow-up, (3) between-group differences in change, and (4) standardized effect sizes (eg, Hedges g) with 95% CIs to guide planning for a future efficacy trial. Missing data will be handled via listwise deletion. A *P* value of .05 will be used to indicate significance in all analyses.

Dissemination Plan

A final report will be prepared for submission to peer-reviewed journals and scientific conferences to support dissemination to stakeholders across the research-practice-community pipeline, as well as to the general public. Participants who have consented to email communication will also receive a lay-language summary of the results (formatted for accessibility). Authorship of all products will be determined by the PI in accordance with journal guidelines and individual contributions. Deidentified participant-level data will be made available upon reasonable request after study completion and publication.

Results

The trial was funded by the NIA in September 2023 and received IRB approval on September 10, 2025. Recruitment began on September 15, 2025, and is expected to conclude in May 2026, with data collection projected to end in August 2026. As of February 2026, 44 participants have been enrolled and 22 have completed the study. The trial flow diagram is presented in Figure 3. Data analysis is planned for August 2026, and the primary results are anticipated to be available in December 2026.

XSL•FO
RenderX

Figure 3. CarePair trial flow diagram. TBD: to be determined.

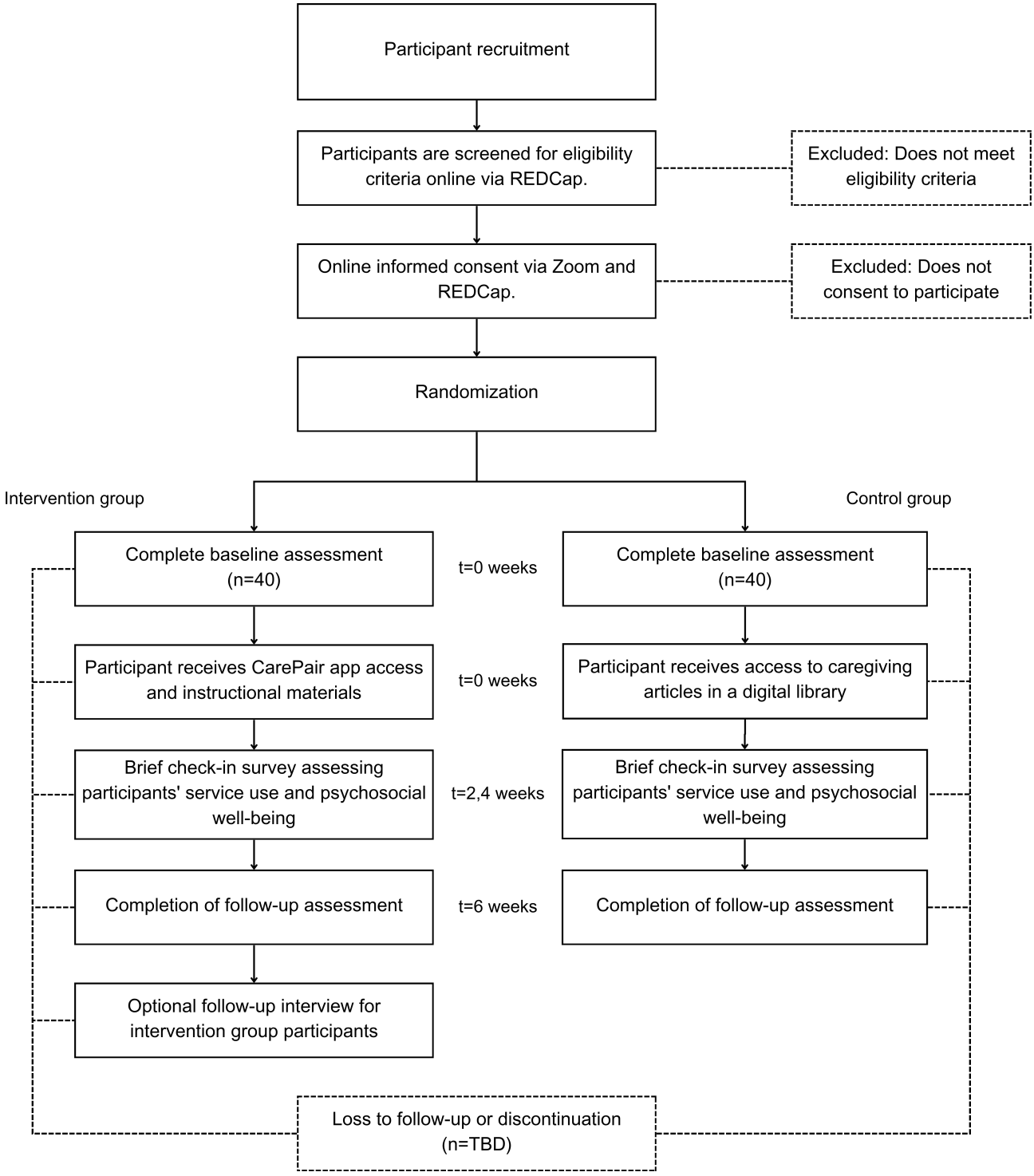

Discussion

Principal Findings

ADRD family caregiving is a chronic, complex role characterized by escalating demands and unpredictability across the disease trajectory and is often assumed with limited knowledge or preparedness. Extensive research has documented the physical and mental health consequences of care-related stress, including increased risks for burden, depression, and anxiety [3-6]. Although the need for systematic identification, assessment, and referral has been emphasized to enhance support, fragmented service systems and navigation barriers contribute to a substantial gap between caregiver needs and the utilization of supportive services [27].

This pilot randomized trial aims to evaluate the feasibility and acceptability of a user-centered, mobile self-assessment and service referral platform for ADRD family caregivers. The current protocol is designed to determine whether CarePair can be feasibly delivered and is perceived as a usable and acceptable digital support tool by family caregivers. We hypothesize that CarePair will (1) demonstrate strong feasibility and acceptability, as reflected by recruitment, retention, adherence,

XSL•FO
RenderX

and usability metrics; (2) provide insights on intervention uptake and engagement to inform iterative refinement and future implementation; and (3) generate preliminary effect sizes to guide the design of a fully powered efficacy trial.

Prior behavioral interventions have demonstrated potential to improve caregiver outcomes; however, their implementation in real-world settings remains limited. Although tailored support initiatives have emerged, many rely on interventionist-led models and do not systematically account for service-use barriers, caregiver characteristics, or individual preferences to inform referrals, thereby limiting personalization and real-world applicability [24,37,38]. Studies on technology-based caregiver interventions have reported small reductions in depression, anxiety, and caregiver burden, along with modest improvements in self-efficacy and quality of life [39-41]. However, the overall strength of the evidence remains mixed, with considerable variability in theoretical grounding, personalization, implementation, and demonstrated effectiveness.

CarePair was developed to address these gaps by integrating self-directed assessments with tailored service referrals that also proactively account for barriers (eg, cost and modality) and individual preferences (eg, timing and distance) to streamline ADRD caregivers' knowledge of and access to relevant support. The design and development of the intervention were guided by the Stress Process Model [13], which conceptualizes caregiving outcomes as arising from the interplay of primary stressors (eg, care demands and functional/cognitive impairment), secondary stressors (eg, role strain and work-family conflict), and mediating factors (eg, social support, coping, and preparedness). A user-centered design framework further informed the platform's development, emphasizing usability, accessibility, and alignment with ADRD caregivers' preferences. Together, these approaches strengthen the intervention's relevance and enhance potential for real-world implementation.

Data on key feasibility indicators, including recruitment, retention, adherence, and participant satisfaction, will determine whether CarePair can be effectively delivered and whether it is perceived as useful and usable among ADRD caregivers. Study findings will guide iterative refinements to CarePair's content, referral logic, and onboarding processes and will directly inform the design of a larger, multisite RCT to evaluate CarePair's efficacy for enhancing the mental health and well-being of ADRD family caregivers.

Limitations

Several limitations warrant consideration. Given that the primary objective of this pilot study is to evaluate feasibility and acceptability, the sample is not statistically powered to examine efficacy for improving caregiver outcomes. Eligibility criteria require participants to have device/internet access, which, in addition to the study's virtual modes of recruitment, may yield a sample with higher digital literacy than the broader caregiving population. Self-selection bias may also impact generalizability, as individuals who choose to participate may be more motivated, more receptive to digital caregiving tools, or differ in caregiving stress levels, which can bias estimates of feasibility and acceptability. Finally, the relatively short intervention period limits the ability to evaluate sustained engagement, longer-term acceptability, and barriers or facilitators influencing CarePair uptake over time.

Conclusion

This study will generate foundational evidence on the feasibility, acceptability, and perceived utility of a mobile self-assessment and referral platform to enhance knowledge of and access to formal services and supports for ADRD family caregivers. Findings will inform the design of a fully powered clinical trial and guide the optimization of the platform. As a scalable, low-cost, and low-burden intervention, CarePair has the potential to bridge gaps in the identification-assessment-referral pipeline, with the overall goal of supporting the mental health and well-being of ADRD family caregivers. Future work should examine the influence of CarePair on both caregiver and care recipient outcomes. Future research should also explore how to integrate digital self-assessment and referral systems into existing clinical workflows within home- and community-based organizations. Dissemination of findings will include sharing results with academic, clinical, and community stakeholders to support continued refinement and implementation of digital support solutions for ADRD family caregivers.

Funding

This research is supported by the National Institute on Aging (NIA) of the National Institutes of Health under award numbers K99AG073509 and R00AG073509 (principal investigator: FF) and by the National Center for Advancing Translational Science (NCATS) of the US National Institutes of Health under UL1TR001855 and UL1TR000130. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funder was not involved in the conceptualization, design, data collection, analysis, decision to publish, or preparation of this manuscript. The sponsor (University of Southern California) had no involvement in the study's conceptualization, design, data collection, analysis, decision to publish, or preparation.

Data Availability

Data sharing is not applicable to this article as no data were analyzed or generated. Upon completion of the project's study aims, data will be available upon reasonable request.

XSL•FO
RenderX

Authors' Contributions

FF is the grant holder and conceived the study. FF and AG initiated and led the study design. FF, AG, and KB were responsible for clinical trial implementation and management. Human subject activities, including consent procedures and data collection, were conducted by KB, ER, and HM, with oversight provided by AG and FF. AG maintained the institutional review board protocols and, together with KB, led recruitment. AB oversaw randomization development and implementation. Participant training materials were developed by KB and AB. REDCap database development was conducted by AG and AB, and both were also responsible for control condition technical development. The manuscript was drafted by FF, AG, and HM. All authors contributed to refining the study protocol and approved the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
CarePair institutional review board approval letter.
[PDF File, 107 KB - resprot_v15i1e90244_app1.pdf]

Multimedia Appendix 2
World Health Organization trial registration dataset.
[DOCX File, 33 KB - resprot_v15i1e90244_app2.docx]

Multimedia Appendix 3
CarePair study informed consent form.
[PDF File, 270 KB - resprot_v15i1e90244_app3.pdf]

Checklist 1
SPIRIT checklist.
[DOCX File, 43 KB - resprot_v15i1e90244_app4.docx]

Peer Review Report 1
Peer Review Report by the NIA-S - Behavior and Social Science of Aging Review Committee, National Institute on Aging Initial Review Group, National Institute on Aging (National Institutes of Health, USA)
[PDF File, 137 KB - resprot_v15i1e90244_app5.pdf]

References

1. 2025 Alzheimer's disease facts and figures. Alzheimers Dement 2025 Apr;21(4):e70235. [doi: 10.1002/alz.70235]
2. Pavolini E, Ranci C. Restructuring the welfare state: reforms in long-term care in Western European countries. J Eur Soc Policy 2008 Aug;18(3):246-259. [doi: 10.1177/0958928708091058]
3. Cooper C, Balamurali TBS, Livingston G. A systematic review of the prevalence and covariates of anxiety in caregivers of people with dementia. Int Psychogeriatr 2007 Apr;19(2):175-195. [doi: 10.1017/S1041610206004297] [Medline: 17005068]
4. Joling KJ, van Hout HPJ, Schellevis FG, et al. Incidence of depression and anxiety in the spouses of patients with dementia: a naturalistic cohort study of recorded morbidity with a 6-year follow-up. Am J Geriatr Psychiatry 2010 Feb;18(2):146-153. [doi: 10.1097/JGP.0b013e3181bf9f0f] [Medline: 20104070]
5. Sallim AB, Sayampanathan AA, Cuttilan A, Ho R. Prevalence of mental health disorders among caregivers of patients with Alzheimer disease. J Am Med Dir Assoc 2015 Dec;16(12):1034-1041. [doi: 10.1016/j.jamda.2015.09.007] [Medline: 26593303]
6. Mittelman MS, Haley WE, Clay OJ, Roth DL. Improving caregiver well-being delays nursing home placement of patients with Alzheimer disease. Neurology (ECronicon) 2006 Nov 14;67(9):1592-1599. [doi: 10.1212/01.wnl.0000242727.81172.91] [Medline: 17101889]
7. Cheng ST, Chan WC, Lam LCW. Are care-recipient outcomes attributable to improved caregiver well-being? A cluster-randomized controlled trial of benefit-finding intervention. Am J Geriatr Psychiatry 2022 Aug;30(8):903-913. [doi: 10.1016/j.jagp.2021.08.009] [Medline: 34563429]
8. Gaugler JE, Reese M, Mittelman MS. Effects of the NYU caregiver intervention-adult child on residential care placement. Gerontologist 2013 Dec;53(6):985-997. [doi: 10.1093/geront/gns193] [Medline: 23339050]
9. Cheng ST, Au A, Losada A, Thompson LW, Gallagher-Thompson D. Psychological interventions for dementia caregivers: what we have achieved, what we have learned. Curr Psychiatry Rep 2019 Jun 6;21(7):59. [doi: 10.1007/s11920-019-1045-9] [Medline: 31172302]

XSL•FO
RenderX

10. Walter E, Pinquart M. How effective are dementia caregiver interventions? An updated comprehensive meta-analysis. Gerontologist 2020 Nov 23;60(8):609-619. [doi: 10.1093/geront/gnz118] [Medline: 33226434]
11. Gitlin LN, Marx K, Stanley IH, Hodgson N. Translating evidence-based dementia caregiving interventions into practice: state-of-the-science and next steps. Gerontologist 2015 Apr;55(2):210-226. [doi: 10.1093/geront/gnu123] [Medline: 26035597]
12. Schulz R, Beach SR, Czaja SJ, Martire LM, Monin JK. Family caregiving for older adults. Annu Rev Psychol 2020 Jan 4;71(1):635-659. [doi: 10.1146/annurev-psych-010419-050754] [Medline: 31905111]
13. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [doi: 10.1093/geront/30.5.583] [Medline: 2276631]
14. Bastawrous M. Caregiver burden--a critical discussion. Int J Nurs Stud 2013 Mar;50(3):431-441. [doi: 10.1016/j.ijnurstu.2012.10.005] [Medline: 23131724]
15. Boumans NPG, Dorant E. The relationships of job and family demands and job and family resources with family caregivers' strain. Scand J Caring Sci 2021 Jun;35(2):567-576. [doi: 10.1111/scs.12873] [Medline: 32400014]
16. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. JAMA 2014 Mar 12;311(10):1052-1060. [doi: 10.1001/jama.2014.304] [Medline: 24618967]
17. Schulz R, Martire LM. Family caregiving of persons with dementia: prevalence, health effects, and support strategies. Am J Geriatr Psychiatry 2004;12(3):240-249. [doi: 10.1097/00019442-200405000-00002] [Medline: 15126224]
18. Laparidou D, Middlemass J, Karran T, Siriwardena AN. Caregivers' interactions with health care services - mediator of stress or added strain? Experiences and perceptions of informal caregivers of people with dementia - a qualitative study. Dementia (London) 2019;18(7-8):2526-2542. [doi: 10.1177/1471301217751226] [Medline: 29385819]
19. Musumeci M, Watts M, Chidambaram P. Key state policy choices about medicaid home and community-based services. KFF. 2020. URL: https://www.kff.org/medicaid/key-state-policy-choices-about-medicaid-home-and-community-based-services/ [accessed 2025-11-03]
20. Sengupta M, Lendon JP, Caffrey C, Melekin A, Singh P. Post-acute and long-term care providers and services users in the United States, 2017-2018. Vital Health Stat 2022 May;3(47):1-93. [Medline: 35604771]
21. Reckrey JM, Morrison RS, Boerner K, et al. Living in the community with dementia: who receives paid care? J Am Geriatr Soc 2020 Jan;68(1):186-191. [doi: 10.1111/jgs.16215] [Medline: 31696511]
22. Riffin C, Van Ness PH, Wolff JL, Fried T. Family and other unpaid caregivers and older adults with and without dementia and disability. J Am Geriatr Soc 2017 Aug;65(8):1821-1828. [doi: 10.1111/jgs.14910] [Medline: 28426910]
23. Waymouth M, Siconolfi D, Friedman EM, Saliba D, Ahluwalia SC, Shih RA. Barriers and facilitators to home- and community-based services access for persons with dementia and their caregivers. J Gerontol B Psychol Sci Soc Sci 2023 May 26;78(6):1085-1097. [doi: 10.1093/geronb/gbad039] [Medline: 36896936]
24. Riffin CA, Wolff JL. Identifying, assessing, and supporting family caregivers in health and long-term care: current progress and future opportunities. In: Gaugler JE, editor. Bridging the Family Care Gap: Academic Press; 2021:341-365. [doi: 10.1016/B978-0-12-813898-4.00011-7]
25. Falzarano F, Greenfield A, Osso F, Bumbalova K, Bloom RF. "Diagnose & adios": multi-perspective insights on formal service use in dementia family caregivers. Gerontologist 2025 May 10;65(6):gnaf081. [doi: 10.1093/geront/gnaf081] [Medline: 39981669]
26. 2015-2017 family support research and training center state coalition report. The Arc. 2019. URL: https://thearc.org/resource/2015-2017-family-support-research-and-training-center-fsrtc-state-coalition-report/ [accessed 2025-11-04]
27. Schulz R, Eden J. Families Caring for an Aging America: National Academies Press; 2016. [doi: 10.17226/23606]
28. Czaja SJ, Gitlin LN, Schulz R, et al. Development of the risk appraisal measure: a brief screen to identify risk areas and guide interventions for dementia caregivers. J Am Geriatr Soc 2009 Jun;57(6):1064-1072. [doi: 10.1111/j.1532-5415.2009.02260.x] [Medline: 19453305]
29. Montgomery R, Kwak J. TCARE: tailored caregiver assessment and referral. Am J Nurs 2008 Sep;108(9 Suppl):54-57. [doi: 10.1097/01.NAJ.0000336418.50193.74] [Medline: 18797229]
30. Gaugler JE, Potter T, Pruinelli L. Partnering with caregivers. Clin Geriatr Med 2014 Aug;30(3):493-515. [doi: 10.1016/j.cger.2014.04.003] [Medline: 25037292]
31. Assessing family caregiver needs: policy and practice considerations. Eldercare Workforce Alliance. URL: https://eldercareworkforce.org/assessing-family-caregiver-needs-policy-and-practice-considerations/ [accessed 2026-03-19]
32. Shu S, Woo BK. Use of technology and social media in dementia care: current and future directions. World J Psychiatry 2021 Apr 19;11(4):109-123. [doi: 10.5498/wjp.v11.i4.109] [Medline: 33889536]
33. Bruinsma J, Peetoom K, Bakker C, et al. Tailoring and evaluating the web-based "Partner in Balance" intervention for family caregivers of persons with young-onset dementia. Internet Interv 2021 Sep;25:100390. [doi: 10.1016/j.invent.2021.100390] [Medline: 33996507]
34. Wang M, He Y, Song J, et al. Digital technology-based dyadic interventions in patients with mild cognitive impairment or dementia and their caregivers: a scoping review. Front Psychiatry 2026;17:1726605. [doi: 10.3389/fpsyt.2026.1726605]

XSL•FO
RenderX

35. Zhou Y, Bai Z, Wan K, Qin T, He R, Xie C. Technology-based interventions on burden of older adults' informal caregivers: a systematic review and meta-analysis of randomized controlled trials. BMC Geriatr 2024 May 4;24(1):398. [doi: 10.1186/s12877-024-05018-w] [Medline: 38704539]
36. Neal DP, Kucera M, van Munster BC, et al. Cost-effectiveness of the FindMyApps eHealth intervention vs. digital care as usual: results from a randomised controlled trial. Aging Ment Health 2024 Nov;28(11):1457-1470. [doi: 10.1080/13607863.2024.2345128] [Medline: 38708873]
37. Bieber A, Nguyen N, Meyer G, Stephan A. Influences on the access to and use of formal community care by people with dementia and their informal caregivers: a scoping review. BMC Health Serv Res 2019 Feb 1;19(1):88. [doi: 10.1186/s12913-018-3825-z] [Medline: 30709345]
38. Navaie M. Accessibility of caregiver education and support programs: reaching hard-to-reach caregivers. In: Toseland R, Haigler D, Monahan D, editors. Education and Support Programs for Caregivers Caregiving: Research, Practice, Policy: Springer; 2011:13-28. [doi: 10.1007/978-1-4419-8031-1_2]
39. Cheng JY, Nurul S, Cheng LJ, He H. Effectiveness of technology‐delivered psychosocial interventions for family caregivers of patients with dementia: a systematic review, meta‐analysis and meta‐regression. Int J Mental Health Nurs 2024 Dec;33(6):1796-1816 [FREE Full text] [doi: 10.1111/inm.13390]
40. Deeken F, Rezo A, Hinz M, Discher R, Rapp MA. Evaluation of technology-based interventions for informal caregivers of patients with dementia-a meta-analysis of randomized controlled trials. Am J Geriatr Psychiatry 2019 Apr;27(4):426-445. [doi: 10.1016/j.jagp.2018.12.003] [Medline: 30642650]
41. Ferrero-Sereno P, Mendoza-Muñoz M, Palomo-López P, et al. A systematic review of the effectiveness of technological interventions for caregivers of people with dementia: effects on quality of life and psychoemotional variables. Front Public Health 2025;13:1579239. [doi: 10.3389/fpubh.2025.1579239] [Medline: 40438051]
42. Mobile fact sheet. Pew Research Center. 2025. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2025-12-17]
43. Pappadà A, Chattat R, Chirico I, Valente M, Ottoboni G. Assistive technologies in dementia care: an updated analysis of the literature. Front Psychol 2021;12:644587. [doi: 10.3389/fpsyg.2021.644587] [Medline: 33841281]
44. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019 Jul;95:103208. [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
45. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381. [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
46. Onken LS, Carroll KM, Shoham V, Cuthbert BN, Riddle M. Reenvisioning clinical science: unifying the discipline to improve the public health. Clin Psychol Sci 2014 Jan 1;2(1):22-34. [doi: 10.1177/2167702613497932] [Medline: 25821658]
47. Teare MD, Dimairo M, Shephard N, Hayman A, Whitehead A, Walters SJ. Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. Trials 2014 Jul 3;15:264. [doi: 10.1186/1745-6215-15-264] [Medline: 24993581]
48. Eldridge SM, Lancaster GA, Campbell MJ, et al. Defining feasibility and pilot studies in preparation for randomised controlled trials: development of a conceptual framework. PLoS ONE 2016;11(3):e0150205. [doi: 10.1371/journal.pone.0150205] [Medline: 26978655]
49. Lancaster GA, Dodd S, Williamson PR. Design and analysis of pilot studies: recommendations for good practice. J Eval Clin Pract 2004 May;10(2):307-312. [doi: 10.1111/j..2002.384.doc.x] [Medline: 15189396]
50. Thabane L, Ma J, Chu R, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010 Jan 6;10(1):1. [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
51. Roque NA, Boot WR. A new tool for assessing mobile device proficiency in older adults: the mobile device proficiency questionnaire. J Appl Gerontol 2018 Feb;37(2):131-156. [doi: 10.1177/0733464816642582] [Medline: 27255686]
52. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. JAMA 1963 Sep 21;185(12):914-919. [doi: 10.1001/jama.1963.03060120024016] [Medline: 14044222]
53. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. Gerontologist 1970;10(1):20-30. [doi: 10.1093/geront/10.1_part_1.20] [Medline: 5420677]
54. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [doi: 10.1093/geront/9.3_Part_1.179] [Medline: 5349366]
55. Wisniewski SR, Belle SH, Coon DW, et al. The resources for enhancing Alzheimer's caregiver health (REACH): project design and baseline characteristics. Psychol Aging 2003 Sep;18(3):375-384. [doi: 10.1037/0882-7974.18.3.375] [Medline: 14518801]
56. Czaja SJ, Lee CC, Perdomo D, et al. Community REACH: an implementation of an evidence-based caregiver program. Gerontologist 2018 Mar 19;58(2):e130-e137. [doi: 10.1093/geront/gny001] [Medline: 29562361]
57. Teri L, Truax P, Logsdon R, Uomoto J, Zarit S, Vitaliano PP. Assessment of behavioral problems in dementia: the revised memory and behavior problems checklist. Psychol Aging 1992;7(4):622-631. [doi: 10.1037/0882-7974.7.4.622]

XSL•FO
RenderX

58. Semple SJ. Conflict in Alzheimer's caregiving families: its dimensions and consequences. Gerontologist 1992 Oct;32(5):648-655. [doi: 10.1093/geront/32.5.648] [Medline: 1427277]
59. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol 1988;54(6):1063-1070. [doi: 10.1037/0022-3514.54.6.1063]
60. Archbold PG, Stewart BJ, Greenlick MR, Harvath T. Mutuality and preparedness as predictors of caregiver role strain. Res Nurs Health 1990 Dec;13(6):375-384. [doi: 10.1002/nur.4770130605] [Medline: 2270302]
61. Kosberg JI, Cairl RE. The Cost of Care Index: a case management tool for screening informal care providers. Gerontologist 1986 Jun;26(3):273-278. [doi: 10.1093/geront/26.3.273] [Medline: 3087822]
62. Sudore RL, Stewart AL, Knight SJ, et al. Development and validation of a questionnaire to detect behavior change in multiple advance care planning behaviors. PLoS ONE 2013;8(9):e72465. [doi: 10.1371/journal.pone.0072465] [Medline: 24039772]
63. Sudore RL, Heyland DK, Barnes DE, et al. Measuring advance care planning: optimizing the advance care planning engagement survey. J Pain Symptom Manage 2017 Apr;53(4):669-681. [doi: 10.1016/j.jpainsymman.2016.10.367] [Medline: 28042072]
64. Dorri N, Riegel B. Development and psychometric evaluation of the self-care of informal caregivers inventory. Int J Nurs Stud Adv 2024 Dec;7:100237. [doi: 10.1016/j.ijnsa.2024.100237] [Medline: 39328836]
65. Kiely DK, Prigerson H, Mitchell SL. Health care proxy grief symptoms before the death of nursing home residents with advanced dementia. Am J Geriatr Psychiatry 2008 Aug;16(8):664-673. [doi: 10.1097/JGP.0b013e3181784143] [Medline: 18669945]
66. Chiambretto P, Moroni L, Guarnerio C, Bertolotti G, Prigerson HG. Prolonged grief and depression in caregivers of patients in vegetative state. Brain Inj 2010;24(4):581-588. [doi: 10.3109/02699051003610490] [Medline: 20235760]
67. Bédard M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden Interview: a new short version and screening version. Gerontologist 2001 Oct;41(5):652-657. [doi: 10.1093/geront/41.5.652] [Medline: 11574710]
68. Tarlow BJ, Wisniewski SR, Belle SH, Rubert M, Ory MG, Gallagher-Thompson D. Positive aspects of caregiving: contributions of the REACH project to the development of new measures for Alzheimer's caregiving. Res Aging 2004;26(4):429-453. [doi: 10.1177/0164027504264493]
69. Lubben J, Blozik E, Gillmann G, et al. Performance of an abbreviated version of the Lubben Social Network Scale among three European community-dwelling older adult populations. Gerontologist 2006 Aug;46(4):503-513. [doi: 10.1093/geront/46.4.503] [Medline: 16921004]
70. Bass DM, Mccarthy C, Eckert S, Bichler J. Differences in service attitudes and experiences among families using three types of support services. American Journal of Alzheimer's Care and Related Disorders & Research 1994 May;9(3):28-38. [doi: 10.1177/153331759400900306]
71. Bass DM, Judge KS, Snow AL, et al. Caregiver outcomes of partners in dementia care: effect of a care coordination program for veterans with dementia and their family members and friends. J Am Geriatr Soc 2013 Aug;61(8):1377-1386. [doi: 10.1111/jgs.12362] [Medline: 23869899]
72. Orsmond GI, Cohn ES. The distinctive features of a feasibility study: objectives and guiding questions. OTJR (Thorofare N J) 2015 Jul;35(3):169-177. [doi: 10.1177/1539449215578649] [Medline: 26594739]
73. Brooke J. SUS: a "quick and dirty" usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester BA, editors. Usability Evaluation in Industry: CRC Press; 1996. URL: https://www.taylorfrancis.com/chapters/edit/10.1201/9781498710411-35/sus-quick-dirty-usability-scale-john-brooke [accessed 2026-03-19]
74. Lewis JR. The system usability scale: past, present, and future. International Journal of Human–Computer Interaction 2018 Jul 3;34(7):577-590. [doi: 10.1080/10447318.2018.1455307]
75. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72. [doi: 10.2196/mhealth.5849] [Medline: 27287964]
76. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613. [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
77. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
78. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
79. Sadak T, Korpak A, Wright JD, et al. Psychometric evaluation of Kingston Caregiver Stress Scale. Clin Gerontol 2017;40(4):268-280. [doi: 10.1080/07317115.2017.1313349] [Medline: 28459351]
80. Kilik LA, Hopkins RW. The relationship between caregiver stress and behavioural changes in dementia. OBM Geriatr 2019;03(2):1-19. [doi: 10.21926/obm.geriatr.1902052]
81. Steffen AM, McKibbin C, Zeiss AM, Gallagher-Thompson D, Bandura A. The revised scale for caregiving self-efficacy: reliability and validity studies. J Gerontol B Psychol Sci Soc Sci 2002 Jan;57(1):74-86. [doi: 10.1093/geronb/57.1.p74] [Medline: 11773226]
82. Diener E, Emmons RA, Larsen RJ, Griffin S. The satisfaction with life scale. J Pers Assess 1985 Feb;49(1):71-75. [doi: 10.1207/s15327752jpa4901_13] [Medline: 16367493]

XSL•FO
RenderX

83. Cheung F, Lucas RE. Assessing the validity of single-item life satisfaction measures: results from three large samples. Qual Life Res 2014 Dec;23(10):2809-2818. [doi: 10.1007/s11136-014-0726-4] [Medline: 24890827]
84. Cooper CL, Whitehead A, Pottrill E, Julious SA, Walters SJ. Are pilot trials useful for predicting randomisation and attrition rates in definitive studies: a review of publicly funded trials. Clin Trials 2018 Apr;15(2):189-196. [doi: 10.1177/1740774517752113] [Medline: 29361833]
85. de Miquel C, Haro JM, van der Feltz-Cornelis CM, et al. Differential attrition and engagement in randomized controlled trials of occupational mental health interventions in person and online: a systematic review and meta-analysis. Scand J Work Environ Health 2024 Dec 1;50(8):588-601. [doi: 10.5271/sjweh.4173] [Medline: 39072699]

Abbreviations

ADRD: Alzheimer disease and related dementias
AE: adverse event
CTSI: Clinical and Translational Science Institute
GAD-7: Generalized Anxiety Disorder-7
HCBS: home- and community-based service
HIPAA: Health Insurance Portability and Accountability Act
IRB: institutional review board
KCSS: Kingston Caregiver Stress Scale
NIA: National Institute on Aging
NIH: National Institutes of Health
PHQ-9: Patient Health Questionnaire-9
PI: principal investigator
RA: research assistant
RCT: randomized controlled trial
REDCap: Research Electronic Data Capture
SAE: serious adverse event
SME: subject matter expert
SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials
SUS: System Usability Scale
SWLS: Satisfaction With Life Scale
uMARS: user version of the Mobile Application Rating Scale
USC: University of Southern California

Edited by J Sarvestan; submitted 23.Dec.2025; this is a non–peer-reviewed article;accepted 03.Mar.2026; published 01.Apr.2026.

Please cite as:
Falzarano F, Greenfield A, Mason H, Bumbalova K, Rojas E, Bumbalov A
A Mobile Self-Assessment and Referral Platform for Family Caregivers of Individuals With Alzheimer Disease and Related Dementias: Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2026;15:e90244
URL: https://www.researchprotocols.org/2026/1/e90244
doi:10.2196/90244

© Francesca Falzarano, Annabelle Greenfield, Hannah Mason, Katerina Bumbalova, Elizabeth Rojas, Aleksandar Bumbalov. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 1.Apr.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

A Need for Strategies to Reduce Alcohol Use After Metabolic and Bariatric Surgery: Technology-Based Intervention and Study Protocol for a Pilot Randomized Controlled Trial

Lisa R Miller-Matero[1,2,3], PhD; Celeste Pappas[2], MPH; Brittany Christopher[2], BA; Roman Grossi[2], BS; Alyssa Vanderziel[2,3], PhD; Nancy P Barnett[4], PhD; Roland S Moore[5], PhD; Sarah Bendit[1], MD; Aaron Hamann[1,6], PsyD; Arthur M Carlin[6], MD; Oliver A Varban[6], MD; Jordan M Braciszewski[2,3], PhD

[1]Behavioral Health, Henry Ford Health System, Detroit, MI, United States

[2]Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, United States

[3]College of Human Medicine, Michigan State University, East Lansing, MI, United States

[4]Department of Behavioral and Social Sciences, Center for Alcohol and Addiction Studies, Brown University, Providence, RI, United States

[5]Pacific Institute For Research and Evaluation, Berkeley, CA, United States

[6]Department of Surgery, Henry Ford Health System, Detroit, MI, United States

Corresponding Author:
Lisa R Miller-Matero, PhD
Behavioral Health
Henry Ford Health System
1 Ford Pl, 5E
Detroit, MI, 48202
United States
Phone: 1 313 214 8240
Email: lmatero1@hfhs.org

Abstract

Background: Individuals who undergo metabolic and bariatric surgery (MBS) are at an increased risk of developing a postoperative alcohol use disorder. Therefore, preventive strategies are needed to mitigate this risk. A technology-based intervention has reduced alcohol use among other populations and could be used after MBS.

Objective: This study aims to describe a technology-based intervention to reduce alcohol use after MBS and to report the study protocol.

Methods: The intervention consists of 2 computerized brief intervention sessions followed by 3 months of daily, tailored, personalized SMS text messages in which the intervention content adapts to an individual's readiness to change. The intervention will be tested in a pilot randomized controlled trial with assessments completed at baseline, 1 month after baseline, after the intervention, and at a 6-month follow-up.

Results: This project was funded by the National Institute on Alcohol Abuse and Alcoholism in September 2020. Enrollment of 60 participants in the trial occurred from November 2023 to August 2024. We anticipate that this intervention will be feasible and acceptable. We hypothesize that the intervention group will have a lower proportion of individuals who report alcohol use at the postintervention and follow-up assessments. We also expect that those in the intervention group will show an increased rating for the importance of avoiding alcohol use. We plan to explore the direction of effects for other outcomes, including hazardous alcohol use, attitudes toward drinking, psychiatric symptoms, and eating and lifestyle behaviors.

Conclusions: A technology-based approach could be a feasible and acceptable method of delivering preventive strategies to individuals after MBS. The long-term goal is to develop an effective, scalable, and sustainable intervention to mitigate the risk of hazardous alcohol use and alcohol use disorder after MBS.

Trial Registration: ClinicalTrials.gov NCT04788316; https://clinicaltrials.gov/study/NCT04788316

International Registered Report Identifier (IRRID): DERR1-10.2196/80068

(JMIR Res Protoc 2026;15:e80068) doi:10.2196/80068

XSL•FO
RenderX

KEYWORDS

metabolic and bariatric surgery; alcohol use; intervention; technology; motivational interviewing

Introduction

Despite the efficacy of metabolic and bariatric surgery (MBS) for long-term weight loss and improvement in medical comorbidities [1-3], patients are at risk of postoperative alcohol use disorder (AUD) [4]. Although the preoperative prevalence of AUD is similar to that in the general population (ie, 6%-8%), the number of individuals with an AUD increases after MBS, and as many as 1 in 5 individuals have an AUD 5 years following MBS [4,5]. During the preoperative evaluation, patients seeking MBS are routinely assessed for a history of alcohol use and AUD and are provided with education about their increased risk of postoperative alcohol misuse and AUD [6]. Consequently, patients are counseled to remain abstinent or minimize postoperative alcohol use. Although patients recall receiving this education, most patients engage in postoperative alcohol use, which places them at risk for an AUD [7]. Thus, there is a need for more novel strategies to mitigate the risk of alcohol misuse and AUD after MBS.

One promising option for reducing the risk of AUD following MBS is delivering intervention content using technology. A systematic review found that various approaches were used to deliver education and intervention content after MBS, including online, web-based, and smartphone delivery [8]. Randomized controlled trials cited in this review suggested that these approaches have the potential to improve outcomes, including weight loss, disordered eating, and physical activity, although the sample sizes were small [8]. However, technology-based tools have not yet been used to reduce alcohol use after MBS. A technology-based intervention could be a reasonable approach to decrease postoperative alcohol use, given the success of technology-based approaches for other outcomes after MBS.

Although there are existing technology-based interventions to reduce alcohol use [9], these are not appropriate for use after MBS in their current form. First, existing interventions were developed and tested for individuals who were engaging in heavy alcohol use [9]. Given the high risk of developing a postoperative AUD, a prevention approach may be best so that we intervene with individuals before escalation to heavy alcohol use. Consequently, strategies to prevent reinitiation of postoperative drinking, or to decrease alcohol use soon after reinitiation, are needed. Intervention content that targets heavy drinking would not be appropriate for those who are not yet drinking or are drinking at lower levels. Second, patients who have undergone MBS may not be motivated to engage in an intensive intervention solely focused on alcohol use if they do not see themselves as having a problem [10]. A technology-based approach may be an appropriate option for delivery of strategies to reduce alcohol use after MBS, as it can offer a lower-intensity intervention, at a low cost, and be tailored to the MBS experience as well as personalized to the individual's goals.

In this paper, we describe the study protocol of a pilot randomized controlled trial of a technology-based approach to reduce postoperative alcohol use, which includes computerized brief intervention (CBI) sessions and SMS text messaging. The goals of the pilot investigation are to study the feasibility, acceptability, and preliminary outcomes.

Methods

Description of the Intervention

The technology-based intervention is composed of 2 sessions of CBI, followed by daily SMS text messaging. Intervention content is grounded in motivational interviewing (MI), which is a patient-centered approach with the goal of eliciting behavior change [11,12]. MI is complementary to the transtheoretical model (TTM)—a theory positing that behavior change occurs in increments along stages—in that the health care provider assesses the patient's stage of readiness and motivation to change, and intervenes accordingly [13]. MI is particularly effective at the initial stages of change within the TTM [14], which is relevant for many individuals who undergo MBS, as they may not be motivated to engage in an alcohol-related intervention if they have not developed a problem with alcohol.

In developing the intervention, we obtained feedback from 20 participants who underwent MBS and were engaging in regular postoperative alcohol use [15]. We then conducted an open trial with 10 participants in which they received the 2 CBI sessions and 1 month of SMS text messaging. On the basis of feedback from the interviews and open trial, we finalized the intervention content and delivery logistics.

CBI Sessions

The CBI was developed using the Computerized Intervention Authoring System [16] and is delivered in a 2-session, web-based format, each lasting 10 to 15 minutes. A 3D, mobile, and emotionally expressive animated narrator reads all the presented material during the sessions. The narration has been found to be respectful, well-understood, and well-liked by other populations [17,18].

CBI session 1 begins with an importance ruler, in which participants are asked, "How important is it for you to avoid alcohol use?" and they respond on a scale of 0 (not at all important) to 10 (very important). The narrator then presents a brief educational component (facts about alcohol use), which incorporates a gender- and race-matched video vignette of a patient who develops an AUD after MBS. Education includes information tailored to this patient population regarding: (1) the new, increased risk for AUD after MBS, (2) reasons for the increased risk (ie, the physiological changes postsurgery and how alcohol is metabolized differently), (3) the health risks of alcohol use (eg, impact on nutrition, weight loss, and physical health), and (4) the "pros" of avoiding alcohol use (eg, impact of alcohol use on MBS, weight loss, mental health, success at work and school, and preventing injury). Participants are again asked about the importance of avoiding alcohol use using the importance ruler. Content is thereafter tailored to each individual's importance rating. Participants who do not think

XSL•FO
RenderX

it is important to avoid alcohol (indicated by scores 0-3) receive intervention content that mirrors phase I of MI (ie, building motivation to abstain or reduce drinking behavior) [12]. Those who are "not sure" (as measured by scores 4-7) receive content on identifying ways to know when it is time to change. For both groups (ie, 0-3 and 4-7), the narrator presents information to evoke change talk (ie, ideas that relate to making behavior changes) [11,12]. Specifically, the narrator asks questions, such as, "If you did to decide to make a change, what would be the best part about having cut down or quit?" "Picture things a year from now. How would you like them to turn out?" and "If you did decide to make a change, what are some reasons you might succeed?" For each of these prompts, participants are presented with myriad response options to choose from, including an open text box, and their responses are reflected by the narrator. A similar approach was successful in increasing confidence for change within a technology-based intervention [19]. Participants who endorse a high level of importance (scores 8-10) are asked about reasons for change, how they might change, and who might help them. All responses that reflect change talk (ie, interest in changing, identification of how they will know it is time to change, and reasons for change) are reinforced with affirmations and reflections. This combination of closed and open questions to elicit change talk, reflections of change talk, and affirmations is among the strongest predictors of further change and commitment talk within MI frameworks [20-23]. The selected change talk from this first CBI session provided by the participant is used in the subsequent SMS text messaging component, which takes place after CBI session 2.

Participants will receive a prompt to complete CBI session 2, 4 days after completing session 1. CBI session 2 begins with an introduction acknowledging that not everyone wants to change or believes that change is necessary. The narrator also states that we want to ensure that participants have the skills they may need in a difficult situation, should they want to make a change. A video vignette using the same gender- and race-matched character from CBI session 1 is presented, illustrating several different coping skills to achieve both weight loss and reductions in alcohol use. The participant then selects tailored coping strategies that they may find useful or want to learn more about. For example, patients can select to learn coping skills that are helpful for both overeating and drinking (ie, relaxation and managing emotional eating and cravings) [24]. The narrator then reviews difficult and risky situations that may arise (eg, "Everyone has the potential to run into difficult or risky situations. Sometimes, they may involve people in your life, and other times could result from feelings you have. Even if you don't think you will soon experience a risky situation, which of these are you most likely to come across?"). The participant can choose from a list of difficult situations they feel are most appropriate for themselves (ie, social pressure, managing difficult emotions, etc). The narrator reflects on the choices selected and acknowledges that it can be helpful to be prepared even if a difficult event is not anticipated. The narrator offers ideas on how to handle the difficult situations and guides the participant through the development of a change plan. After finishing both CBI sessions, participants are asked to provide 1 final importance ruler score, which is used to tailor the SMS text messaging component of the intervention.

SMS Text Messaging

On the basis of feedback from participants in the open trial, daily text messages will be delivered for 3 months. Using results from the CBI, text messaging is personalized. Messages are concise and focused, as both brevity (<160 characters) and immediacy are essential for maximal impact [25]. The content of most of the alcohol-related text messages was adapted from messages used in previous studies [26,27], as well as being theoretically grounded in MI [12] and the TTM [10]. Additional messages were developed that are specific to the MBS experience. Examples of text messages are presented in Textbox 1.

XSL•FO
RenderX

Textbox 1. Examples of SMS text messages. Participants will receive messages tailored to their rating of importance to avoid alcohol on a 0 to 10 scale as follows: low importance (0-3 rating), moderate importance (4-7 rating), and high importance (8-10 rating). All participants will receive personalized goal-related and affirmation text messages.

Text messages based on the rating of the importance of avoiding alcohol use

- Low importance
  - If you decided to be alcohol free, what would be the biggest obstacle?
  - Staying away from alcohol is hard. What's the hardest thing for you?
  - If you decided not to drink, what's the worst that can happen? What's the best that could happen?
- Moderate importance
  - Just considering change is a step in a healthy direction.
  - What goals do you have for the next years? How does alcohol fit into those goals?
  - Like food, alcohol affects chemical pathways in the brain that control reward and pleasure.
- High importance
  - Remind yourself that you can choose to drink or not. Make a promise to yourself to NOT drink and stick with your decision.
  - Being around other people who are drinking is hard! Think about how you are going to handle those situations before they occur.
  - When you get the urge to drink alcohol, distract yourself with something else you enjoy.

Examples of nonalcohol text messages

- One of the most helpful things after surgery is to track your food and liquids. Did you track today?
- Eating slowly allows your brain to register you are full and you may enjoy your food more.
- Parking further away from a store or doing some exercises during a commercial break are great ways to add to your physical activity.
- We all eat things we did not intend to. Getting back on track is a success.

Examples of personalized goal-related text messages

- You mentioned that alcohol could affect your weight loss journey. That could be a powerful motivator!
- You don't want to develop a problem with alcohol later on. That is a good reason to minimize your alcohol use.
- A strategy you find helpful is engaging in positive, fun activities, such as music, art, dancing, or exercise.

Examples of personalized affirmation text messages

- You have not used alcohol since undergoing bariatric surgery. This is something to be proud of!
- You only drank 1 day last week, which was fewer than the 3 days you drank in the week before. Great job!
- You had two weeks in a row where you were alcohol free. It seems you have found helpful ways to avoid drinking!

SMS text messages are tailored to each participant's level of importance to avoid alcohol. Initially, participants will receive messages that correspond with their level of importance rating made at the end of the CBI sessions. Participants will also respond to weekly "poll questions" to assess drinking frequency as well as the importance of avoiding alcohol use. On the basis of their responses, algorithms will determine the tailored messages each participant receives, allowing for fluid change in messages received throughout the 3-month duration of text messages. More specifically, messages will be delivered in 1 of the following 3 ways: those who, at the end of the CBI, indicate a low importance in avoiding alcohol use (ie, scores between 0-3; TTM: precontemplation) will receive message content that helps develop discrepancy between behavior, goals, and values [11,12]. Messages for participants with a moderate level of importance (ie, importance scores between 4-7; TTM: contemplation) will be more motivational in content. Those reporting a high level of importance (ie, scores between 8-10; TTM: preparation and action) will receive messages that include advice and tips. All participants will receive messages that are information-based, and they will also receive content specific to MBS that is nonalcohol focused, based on feedback received in our interviews and open trial.

In addition to the text messages based on their importance rating, participants will also receive personalized text messages twice per week. Selected change talk provided by participants during the CBI will be used to remind participants of the goals they set for themselves. For example, participants are asked to provide reasons for change, people who could help them change, change strategies, events that would precipitate change, and the best parts about potential change. We capture these entries verbatim and provide text messages using the participants' language. Participants also receive personalized affirmation text

XSL•FO
RenderX

messages (ie, positive feedback) when they reduce their drinking or report abstinence, which aligns with MI strategies [11,12].

Study Protocol

Ethical Considerations

This study received funding from the National Institute on Alcohol Abuse and Alcoholism (grant R34 AA027775) and is registered on clinicaltrials.gov (NCT04788316). It is also approved by the health care system's institutional review board (15583). All participants will provide informed consent before engaging in study activities. To protect privacy, participants needed to consent to have their data shared in a data repository. Participants who did not consent to this will not have their data shared. No identifying information of any participant will be shared or published.

Participants and Procedures

We plan to enroll 60 participants who underwent MBS 3 and 18 months before the invitation to participate in the study. This time frame was selected to engage patients who have not yet begun postoperative drinking (most reinitiate between 6 and 12 months postoperatively) [7] and to also engage individuals who began drinking, but are unlikely to have already developed AUD because the risk increases approximately 2 years postoperatively [4]. Patients who underwent MBS at a single institution within this time frame will be sent an email invitation to complete a web-based eligibility survey. Individuals will be eligible to participate if they engaged in alcohol use before MBS, have access to the internet to complete the CBI, and have a cellular phone that can receive SMS text messages. Eligible individuals will be contacted by our study team and provided with a detailed description of the study. Patients will be allowed to ask any questions to the study team and will provide a signature to document their informed consent. Once patients provide informed consent, participants will complete their baseline assessment through a web-based, secure, and Health Insurance Portability and Accountability Act–compliant program (Research Electronic Data Capture). Participants will then be randomized to the intervention group or the treatment-as-usual control group (ie, MBS-related follow-up care) using urn randomization, separating participants by sex and time elapsed since MBS. Participants will complete assessments 1 month after baseline, at the postintervention evaluation (ie, 3 months following baseline), and at the 6-month follow-up through Research Electronic Data Capture. Participants randomized to the intervention group will also complete an exit interview to obtain feedback about the intervention. Participants will be compensated US $50 for each assessment completed.

Measures

Demographics and Background Information

Information collected at baseline will include age, sex and gender, race, weight and BMI, education level, medications, history of psychological symptoms, history of psychological treatment, and history of substance use and substance use treatment.

Alcohol Use

We will measure past 30-day alcohol use, similar to other studies, using adapted items from the Behavioral Risk Factor Surveillance System Questionnaire [28]. Participants will be shown a diagram and description of a standard drink and report the number of days they consumed any alcohol in the past 30 days (0-30). The primary outcome is drinking status (ie, whether alcohol was consumed in the past 30 d). For those who endorse at least 1 drinking day, they will also report on the number of average standard drinks per day, the highest number of drinks in a day, the number of days they consumed 4 or more drinks, and the number of days they perceived that they were drunk. The presence of hazardous alcohol use will be measured with the Alcohol Use Disorder Identification Test-Consumption [29]. This version was chosen because it is shorter than the original Alcohol Use Disorder Identification Test, while still providing a reliable assessment of drinking behavior [29,30]. An Alcohol Use Disorder Identification Test-Consumption score of 3 or greater for women and 4 or greater for men indicates high risk for hazardous drinking.

Other Alcohol-Related Measures

Using the importance ruler, participants will provide their rating of the importance of avoiding alcohol on a scale from 0 (not at all important) to 10 (very important). The Drinking Motives Questionnaire-Revised [31] assesses respondents' reasons for drinking across 4 categories: conformity, coping, enhancement, and social. Each subscale consists of 5 items with responses ranging from 1 (strongly disagree) to 4 (strongly agree).

Psychiatric Symptoms

We will measure depression with the Patient Health Questionnaire-8 depression scale, which is a widely used, validated measure [32]. We will also assess anxiety with the Generalized Anxiety Disorder-7, which is commonly used to assess symptoms of anxiety [33]. Scores greater than or equal to 10 indicate a high likelihood of current depression or anxiety, respectively.

Eating and Lifestyle Behaviors

To explore the potential impact of the intervention on other important MBS-related outcomes, we will include measures related to eating and lifestyle behaviors. We will track the symptoms of food addiction using the Yale Food Addiction Scale 2.0 [34]. We will measure other eating behaviors with the Emotional Eating Scale [35] and the Eating Disorder Examination-Questionnaire [36]. We will also query about the frequency of lifestyle behaviors, including the use of a food diary, exercise, mindful eating, self-weighing, and obtaining 7 to 9 hours of sleep.

Study Feasibility

Feasibility will be assessed in several ways. First, we will calculate eligibility rates as the number who meet eligibility criteria divided by the number screened. Second, enrollment rates will be defined as the number of participants who enroll divided by the number eligible. To improve future recruitment strategies, we will record the reasons for nonenrollment. Finally, retention will be calculated using the rate of follow-up completion, as well as a determination of whether there was

XSL•FO
RenderX

differential completion between the control and intervention groups.

Intervention Adherence

We will examine engagement with the intervention by examining the initiation and completion of CBI sessions and the response rate to the weekly poll questions.

Intervention Acceptability

We will use the Usability and Satisfaction Assessment to capture quantitative data on usability, overall satisfaction, helpfulness, content relevance, and use of intervention skills [37]. Items will query these areas for both the CBI sessions and SMS text messaging to examine acceptability of these components separately (ie, "I liked the 2-session computerized brief intervention," "Is the content in the text messages relevant?" and "How helpful do you think the text messages were?"). The items will be rated on a 5-point scale and could be considered positive (ie, acceptable), neutral, or negative (ie, unacceptable). Exit interviews will provide qualitative feedback on intervention acceptability.

Analytic Plan

We will examine the frequencies of our feasibility, adherence, and acceptability variables. Although the primary focus is on feasibility and acceptability, we also plan to examine the direction of effects on outcomes. Our primary outcome is alcohol use. We will conduct chi-square and logistic regressions for categorical outcomes, and analysis of covariance for continuous variables (controlling for baseline measurements) to examine differences in outcomes between the intervention and control groups at the postintervention evaluation. For the 6-month follow-up, logistic regression will be used for the categorical outcome (eg, any alcohol use at 30 days), and hierarchical linear modeling will be used to examine repeated measurements of alcohol use quantity and frequency, covarying for baseline levels of those variables and demographics. Hierarchical linear modeling is optimal for analyzing longitudinal data, as it allows for control over missing data and comparison of more than 2 waves of data [38,39].

Results

This project was funded by the National Institute on Alcohol Abuse and Alcoholism on September 20, 2020. After completing patient interviews and an open trial to refine the intervention, recruitment for the pilot randomized controlled trial began on November 8, 2023, and the final participant was enrolled on August 14, 2024. A total of 60 patients were enrolled.

Because this is a pilot trial, the primary focus will be on the feasibility of the study protocol, adherence to the intervention, and acceptability of intervention content. We anticipate having an enrollment rate of 50% and at least 75% who complete each assessment. We anticipate that at least 75% of the participants will initiate the intervention (ie, complete at least 1 CBI session). To measure adherence and engagement in the intervention content, we anticipate that at least 75% of the weekly SMS text messages will be responded to and that at least 75% of the participants will complete the intervention (ie, complete both CBI sessions and continue to receive the 3 months of SMS text messaging). We also expect that the majority (>50%) of participants will rate each of the acceptability items as positive and that at least 75% of the participants will believe that other patients would use the intervention after MBS. We have set these benchmarks, as they are deemed reasonable definitions of success for pilot trials [40].

We hypothesize that those in the intervention group will have a lower proportion of individuals engaging in alcohol use at the postintervention and follow-up assessments compared to the control group. We also expect that individuals in the intervention group will report greater importance to avoid alcohol use compared to the control group at the postintervention evaluation and follow-up. Because this is a pilot study and we may not be powered to detect between-group differences, we will also explore changes within groups to determine whether there is potential for an effect (ie, a significant increase in their rating on the importance of avoiding alcohol use among the intervention group). We will explore potential differences between the intervention and control groups for our other outcomes, including proportion of individuals with hazardous drinking, attitudes toward drinking, psychiatric symptoms, and eating and lifestyle behaviors. Although we do not expect statistically significant differences given the small sample size, we hypothesize that those in the intervention group will show a direction toward improvement, and we will examine the effect sizes of the intervention.

Discussion

Anticipated Findings

The goal of this pilot randomized controlled trial is to examine the feasibility and acceptability of a multicomponent technology-based intervention and to examine the direction of intervention effects. We anticipate that this study protocol will be feasible, that participants will engage with and adhere to the components in the technology-based intervention, and that participants will find the intervention content acceptable. We also anticipate that those randomized to the intervention group will have lower alcohol use than those in the control group. Other technology-based interventions for reducing substance use have been feasible, acceptable, and successful at reducing substance use, so this approach will likely translate to individuals who undergo MBS [41]. However, most interventions for substance use are targeted at individuals who have developed substance abuse; therefore, this approach is unique in that the goal is preventive. Moreover, given that the risk of AUD after MBS increases between 1 and 2 years postoperatively, delivering content to the population between 3 and 18 months postoperatively could capture the most vulnerable period for reinitiation and escalation of alcohol use [4].

Strengths and Limitations

This study has several strengths, including the preventive approach to mitigating AUD after MBS; the use of a control group as a comparison to treatment-as-usual; and the collection of data that includes feasibility, acceptability, and preliminary outcomes. However, it is important to note several limitations.

XSL•FO
RenderX

First, this is a pilot study with a small sample size. Consequently, the study may not be appropriately powered to detect changes in alcohol use. Therefore, the goal is to examine the direction of the effects to determine whether a future, fully powered trial is worthwhile. Second, the length of follow-up is short; therefore, the ability to prevent the development of an AUD cannot be assessed. Finally, alcohol use is measured by self-report, and it is possible that patients may minimize their use or report their use inaccurately.

Conclusions and Future Directions

Patients who undergo MBS are at an increased risk of developing a postoperative AUD. Previous research has suggested a need for strategies to reduce postoperative alcohol use to mitigate the risk of an AUD. A technology-based approach could be a feasible and acceptable method of delivering preventive strategies to individuals after MBS. If successful, this project will provide the necessary support to conduct a fully powered randomized controlled trial to examine its efficacy in reducing alcohol use and with longer-term follow-up to determine whether it can reduce the incidence of AUD. Consequently, the next step after this pilot study is to conduct a randomized trial of this technology-based intervention compared to a control group. We plan to determine whether this intervention could prevent the reinitiation of postoperative alcohol use and reduce alcohol use among those who did reinitiate drinking after MBS. We also plan to test mediators and moderators in future work, as well as to examine the factors necessary for the successful implementation with MBS programs. Our long-term goal in this line of research is to develop an effective program that is scalable and sustainable within MBS programs to mitigate the risk of hazardous alcohol use and AUD.

Data Availability

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

Funding

This project was funded by the National Institute on Alcohol Abuse and Alcoholism (grant R34 AA027775).

Authors' Contributions

Conceptualization: LRM-M, JMB, NPB, RSM
Data curation: LRM-M, JMB, CP, BC, RG
Funding acquisition: LRM-M, JMB, NPB, RSM
Investigation: LRM-M, JMB, CP, BC, RG, AV
Methodology: all authors
Project administration: LRM-M, JMB, CP
Writing—original draft: LRM-M
Writing—revising and editing: all authors

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer review report by AA-3 Clinical, Treatment and Health Services Research Review Subcommittee, National Institute on Alcohol Abuse and Alcoholism Initial Review Group (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 153 KB - resprot_v15i1e80068_app1.pdf]

References

1. Buchwald H, Estok R, Fahrbach K, Banel D, Jensen MD, Pories WJ, et al. Weight and type 2 diabetes after bariatric surgery: systematic review and meta-analysis. Am J Med 2009 Mar;122(3):248-56.e5. [doi: 10.1016/j.amjmed.2008.09.041] [Medline: 19272486]
2. Ribaric G, Buchwald JN, McGlennon TW. Diabetes and weight in comparative studies of bariatric surgery vs conventional medical therapy: a systematic review and meta-analysis. Obes Surg 2014 Mar 28;24(3):437-455 [FREE Full text] [doi: 10.1007/s11695-013-1160-3] [Medline: 24374842]
3. Arterburn DE, Telem DA, Kushner RF, Courcoulas AP. Benefits and risks of bariatric surgery in adults: a review. JAMA 2020 Sep 01;324(9):879-887. [doi: 10.1001/jama.2020.12567] [Medline: 32870301]
4. King WC, Chen JY, Courcoulas AP, Dakin GF, Engel SG, Flum DR, et al. Alcohol and other substance use after bariatric surgery: prospective evidence from a U.S. multicenter cohort study. Surg Obes Relat Dis 2017 Aug;13(8):1392-1402 [FREE Full text] [doi: 10.1016/j.soard.2017.03.021] [Medline: 28528115]
5. The National survey on drug use and health. Substance Abuse and Mental Health Services Administration. URL: https://www.samhsa.gov/data/sites/default/files/reports/rpt37924/2020NSDUHNationalSlides072522.pdf [accessed 2025-12-12]

XSL•FO
RenderX

6. Sogg S, Lauretti J, West-Smith L. Recommendations for the presurgical psychosocial evaluation of bariatric surgery patients. Surg Obes Relat Dis 2016 May;12(4):731-749. [doi: 10.1016/j.soard.2016.02.008] [Medline: 27179400]
7. Miller-Matero LR, Coleman JP, LaLonde L, Martens KM, Hamann A, Carlin AM. Patient recall of education about the risks of alcohol use following bariatric surgery. Obes Surg 2019 Aug;29(8):2707-2710. [doi: 10.1007/s11695-019-03940-2] [Medline: 31098786]
8. Messiah SE, Sacher PM, Yudkin J, Ofori A, Qureshi FG, Schneider B, et al. Application and effectiveness of eHealth strategies for metabolic and bariatric surgery patients: a systematic review. Digit Health 2020 Jan 07;6:2055207619898987 [FREE Full text] [doi: 10.1177/2055207619898987] [Medline: 32030193]
9. Kiluk BD, Ray LA, Walthers J, Bernstein M, Tonigan JS, Magill M. Technology-delivered cognitive-behavioral interventions for alcohol use: a meta-analysis. Alcohol Clin Exp Res 2019 Nov 30;43(11):2285-2295 [FREE Full text] [doi: 10.1111/acer.14189] [Medline: 31566787]
10. Prochaska J, Clemente CC. Stages of change in the modification of problem behaviors. Prog Behav Modif 1992;28:183-218. [Medline: 1620663]
11. Rollnick S, Miller WR. What is motivational interviewing? Behav Cogn Psychother 1995;23(4):325-334. [doi: 10.1017/s135246580001643x]
12. Miller WR, Rollnick S. Motivational Interviewing: Helping People Change and Grow. 4th edition. New York, NY: The Guilford press; 2012.
13. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997 Sep 01;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]
14. DiClemente CC, Velasquez MM. Motivational interviewing and the stages of change. In: Miller W, Rollnick S, editors. Motivational Interviewing: Preparing People for Change, 2nd edition. New York, NY: The Guilford Press; Mar 30, 2010:201-216.
15. Braciszewski JM, Hecht LM, Barnett NP, Moore RS, Carlin AM, Haley EN, et al. Preventing alcohol use post-bariatric surgery: patient perspectives on a technology-based approach. Surg Endosc 2023 Nov 05;37(11):8263-8268. [doi: 10.1007/s00464-023-10407-0] [Medline: 37670188]
16. Digital behavioral health interventions made easy. Computerized Intervention Authoring System. URL: https://www.cias.app/ [accessed 2025-05-29]
17. Tzilos GK, Sokol RJ, Ondersma SJ. A randomized phase I trial of a brief computer-delivered intervention for alcohol use during pregnancy. J Womens Health (Larchmt) 2011 Oct;20(10):1517-1524 [FREE Full text] [doi: 10.1089/jwh.2011.2732] [Medline: 21823917]
18. Ondersma SJ, Chase SK, Svikis DS, Schuster CR. Computer-based brief motivational intervention for perinatal drug use. J Subst Abuse Treat 2005 Jun;28(4):305-312 [FREE Full text] [doi: 10.1016/j.jsat.2005.02.004] [Medline: 15925264]
19. Olafsson S, O'Leary T, Bickmore T. Coerced change-talk with conversational agents promotes confidence in behavior change. In: Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare. 2019 Presented at: PervasiveHealth '19; May 20-23, 2019; Trento, Italy p. 31-40 URL: https://dl.acm.org/doi/10.1145/3329189.3329202 [doi: 10.1145/3329189.3329202]
20. Apodaca TR, Jackson KM, Borsari B, Magill M, Longabaugh R, Mastroleo NR, et al. Which individual therapist behaviors elicit client change talk and sustain talk in motivational interviewing? J Subst Abuse Treat 2016 Feb;61:60-65 [FREE Full text] [doi: 10.1016/j.jsat.2015.09.001] [Medline: 26547412]
21. Magill M, Apodaca TR, Borsari B, Gaume J, Hoadley A, Gordon RE, et al. A meta-analysis of motivational interviewing process: technical, relational, and conditional process models of change. J Consult Clin Psychol 2018 Feb;86(2):140-157 [FREE Full text] [doi: 10.1037/ccp0000250] [Medline: 29265832]
22. Jacques-Tiura AJ, Carcone AI, Naar S, Brogan Hartlieb K, Albrecht TL, Barton E. Building motivation in African American caregivers of adolescents with obesity: application of sequential analysis. J Pediatr Psychol 2017 Mar 01;42(2):131-141 [FREE Full text] [doi: 10.1093/jpepsy/jsw044] [Medline: 27246865]
23. Carcone AI, Naar S, Clark J, MacDonell K, Zhang L. Provider behaviors that predict motivational statements in adolescents and young adults with HIV: a study of clinical communication using the Motivational Interviewing framework. AIDS Care 2020 Sep 17;32(9):1069-1077 [FREE Full text] [doi: 10.1080/09540121.2019.1679709] [Medline: 31621396]
24. Miller-Matero LR, Brescacin C, Clark SM, Troncone CL, Tobin ET. Why WAIT? Preliminary evaluation of the weight assistance and intervention techniques (WAIT) group. Psychol Health Med 2019 Oct 07;24(9):1029-1037. [doi: 10.1080/13548506.2019.1587478] [Medline: 30843419]
25. Laursen D. Counseling young cannabis users by text message. J Comput Mediat Commun 2010;15(4):646-665. [doi: 10.1111/j.1083-6101.2009.01510.x]
26. Braciszewski JM, Stout RL, Tzilos GK, Moore RS, Bock BC, Chamberlain P. Testing a dynamic automated substance use intervention model for youths exiting foster care. J Child Adolesc Subst Abuse 2016 Feb 18;25(3):181-187. [doi: 10.1080/1067828x.2014.981771] [Medline: 27081290]
27. Bock BC, Becker BM, Niaura RS, Partridge R, Fava JL, Trask P. Smoking cessation among patients in an emergency chest pain observation unit: outcomes of the Chest Pain Smoking Study (CPSS). Nicotine Tob Res 2008 Oct;10(10):1523-1531. [doi: 10.1080/14622200802326343] [Medline: 18946771]

XSL•FO
RenderX

28. Behavioral Risk Factor Surveillance System survey questionnaire. Center for Disease Control and Prevention. 2021. URL: https://www.cdc.gov/brfss/index.html [accessed 2025-05-29]
29. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol use disorders identification test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
30. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun 24;88(6):791-804. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
31. Cooper ML. Motivations for alcohol use among adolescents: development and validation of a four-factor model. Psychol Assess 1994;6(2):117-128. [doi: 10.1037//1040-3590.6.2.117]
32. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
33. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
34. Gearhardt AN, Corbin WR, Brownell KD. Development of the Yale Food Addiction Scale version 2.0. Psychol Addict Behav 2016 Feb;30(1):113-121. [doi: 10.1037/adb0000136] [Medline: 26866783]
35. Arnow B, Kenardy J, Agras WS. The Emotional Eating Scale: the development of a measure to assess coping with negative affect by eating. Int J Eat Disord 1995 Jul;18(1):79-90. [doi: 10.1002/1098-108x(199507)18:1<79::aid-eat2260180109>3.0.co;2-v] [Medline: 7670446]
36. Luce KH, Crowther JH. The reliability of the Eating Disorder Examination-Self-Report Questionnaire Version (EDE-Q). Int J Eat Disord 1999 Apr;25(3):349-351. [doi: 10.1002/(sici)1098-108x(199904)25:3<349::aid-eat15>3.0.co;2-m] [Medline: 10192002]
37. Braciszewski JM, Tzilos Wernette GK, Moore RS, Tran TB, Bock BC, Stout RL, et al. Developing a tailored substance use intervention for youth exiting foster care. Child Abuse Negl 2018 Mar;77:211-221 [FREE Full text] [doi: 10.1016/j.chiabu.2018.01.013] [Medline: 29367098]
38. Raudenbush SW, Bryk AS. Hierarchical Linear Models: Applications and Data Analysis Methods. Thousand Oaks, CA: Sage Publications; 2002.
39. Singer JD, Willett JB. Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. Oxford, UK: Oxford University Press; 2003.
40. Cooper CL, Whitehead A, Pottrill E, Julious SA, Walters SJ. Are pilot trials useful for predicting randomisation and attrition rates in definitive studies: a review of publicly funded trials. Clin Trials 2018 Apr 23;15(2):189-196 [FREE Full text] [doi: 10.1177/1740774517752113] [Medline: 29361833]
41. Mason M, Ola B, Zaharakis N, Zhang J. Text messaging interventions for adolescent and young adult substance use: a meta-analysis. Prev Sci 2015 Feb 17;16(2):181-188. [doi: 10.1007/s11121-014-0498-7] [Medline: 24930386]

Abbreviations

AUD: alcohol use disorder
CBI: computerized brief intervention
MBS: metabolic and bariatric surgery
MI: motivational interviewing
TTM: transtheoretical model

Edited by J Sarvestan; submitted 03.Jul.2025; this is a non–peer-reviewed article;accepted 05.Nov.2025; published 07.Jan.2026.

Please cite as:
Miller-Matero LR, Pappas C, Christopher B, Grossi R, Vanderziel A, Barnett NP, Moore RS, Bendit S, Hamann A, Carlin AM, Varban OA, Braciszewski JM
A Need for Strategies to Reduce Alcohol Use After Metabolic and Bariatric Surgery: Technology-Based Intervention and Study Protocol for a Pilot Randomized Controlled Trial
JMIR Res Protoc 2026;15:e80068
URL: https://www.researchprotocols.org/2026/1/e80068
doi:10.2196/80068
PMID:

XSL•FO
RenderX

©Lisa R Miller-Matero, Celeste Pappas, Brittany Christopher, Roman Grossi, Alyssa Vanderziel, Nancy P Barnett, Roland S Moore, Sarah Bendit, Aaron Hamann, Arthur M Carlin, Oliver A Varban, Jordan M Braciszewski. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 07.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Interactive Internet-Based Motivational Interviewing Training for HIV Counseling Support Staff to Improve Health Communication in HIV Care Interactions: Protocol for Training Development and a Pilot Randomized Controlled Trial

Iván C Balán[1], PhD; Onna R Brewer[1], PhD; Bryan A Kutner[2], MPH, PhD; Rebecca Giguere[1], MPH

[1]Center for Translational Behavioral Science, Department of Behavioral Science and Social Medicine, College of Medicine, Florida State University, Tallahassee, FL, United States

[2]Division of General Internal Medicine, Department of Psychiatry and Behavioral Sciences, Albert Einstein College of Medicine, Bronx, NY, United States

Corresponding Author:
Onna R Brewer, PhD
Center for Translational Behavioral Science, Department of Behavioral Science and Social Medicine
College of Medicine
Florida State University
2010 Levy Ave Building B
Suite B0266
Tallahassee, FL, 32310
United States
Phone: 1 850 644 2334
Email: onna.brewer@med.fsu.edu

Abstract

Background: HIV counseling support staff (CSS) play a crucial role in HIV care outcomes, providing essential access to HIV test counseling, linkage-to-care support, adherence counseling, peer support, and navigation. Effective training in evidence-based interventions like motivational interviewing (MI) is imperative to maximize the impact of CSS in enhancing HIV care outcomes. MI is a collaborative, goal-oriented communication method aimed at bolstering an individual's motivation and movement toward specific goals by eliciting and exploring personal arguments for change. MI has demonstrated efficacy across various medical and mental health outcomes, including its significant impact on the status-neutral HIV Care Continuum.

Objective: The objective of this study is to develop and pilot an interactive internet-based MI training program for HIV counseling support staff (iMI4HIV), aiming to enhance their communication skills in HIV care interactions.

Methods: The iMI4HIV Study will have 2 phases. In phase 1, Development, we will conduct formative focus groups (FGs) with CSS (n=16) to identify HIV care interactions for demonstration videos and to assess acceptability of gamification components; create MI demonstration videos depicting key interactions from status-neutral HIV Care Continuum; program iMI4HIV (including skills development tasks, gamification components, etc); and pilot a virtual MI (VMI) workshop followed by iMI4HIV (n=8). We will also conduct qualitative interviews with 8 leaders from HIV care agencies to explore their use of existing online MI training programs, acceptability of iMI4HIV, and possible facilitators and impediments to its adoption for training HIV CSS. In phase 2, pilot randomized controlled trial (RCT), we will randomize 30 CSS at a 2:1 ratio (VMI workshop + iMI4HIV vs VMI workshop + waitlist control) to assess the feasibility and acceptability of iMI4HIV and to pilot the RCT processes and assessments for a future efficacy trial. We will also explore the preliminary impact of iMI4HIV on MI skills acquisition.

Results: The CSS focus groups began in August 2024, and programming of the training is near completion. We plan to conduct a pilot with 8 participants by the end of 2025, to obtain feedback and conclude phase 1 of the study. The phase 2 pilot RCT is expected to be launched in February 2026, with all data collection to be completed by October 2026.

Conclusions: Our pilot study aims to demonstrate the feasibility and acceptability of an internet-based MI training program for HIV CSS. If successful, this training program has the potential to enhance the delivery of evidence-based interventions in HIV care settings, ultimately improving patient outcomes and adherence to treatment protocols. Further research and larger trials will be needed to confirm these findings and refine the training approach.

International Registered Report Identifier (IRRID): DERR1-10.2196/82241

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e82241) doi:10.2196/82241

KEYWORDS

HIV; implementation science; internet-based; Motivational Interviewing; training

Introduction

Background

HIV counseling support staff (CSS) provide key services for HIV care outcomes, such as HIV test counseling, linkage to care, adherence counseling, peer support, and care navigation. Interventions delivered by CSS have increased HIV prevention [1,2], testing [3], linkage to care [4,5], treatment engagement [6], and adherence [4-7]. Nevertheless, to maximize patient outcomes, community-based CSS could benefit greatly from training to ensure competent delivery of evidence-based interventions, including motivational interviewing (MI).

MI is "a collaborative, goal-oriented method of communication aimed at strengthening an individual's motivation for and movement toward a specific goal by eliciting and exploring the person's own arguments for change" [8]. MI has proven efficacy in improving substance use [9,10], mental health [11,12], medical [13-17], and other outcomes across the HIV Care Continuum (HCC) in adults and adolescents [3,6,18-36]. MI has also been shown to improve linkage to and retention in care as well as treatment adherence [8-16], including among adolescents [6,26,31]. MI is particularly useful as an intervention because it requires CSS to master one set of skills applicable across the HCC instead of a variety of interventions for each step of the HCC. Furthermore, because it is efficacious even during brief interactions [37], MI can be integrated into all aspects of work typically conducted by CSS in HIV care settings. MI is embedded in HIV care guidelines, including those of the US Department of Health and Human Services and the New York State Department of Health [32,34], and training in MI is often offered through the national network of regional AIDS Education and Training Centers (AETCs).

However, mastering MI is not easy [38-43]. MI entails adhering to standard counseling tools (ie, open-ended questions, affirmations, reflections, and summaries) with an empathic, client-centered, collaborative approach. MI also includes a more difficult, technical component—the systematic elicitation and reinforcement of change talk (statements made by clients that argue for change, eg, "I need to get better at taking my pills") and avoiding evocation of sustain talk (statements made by clients that argue for not changing, eg, "It's hard to remember taking my pills"). Research on MI's mechanisms of action has shown that more change talk and less sustain talk during a session are related to better outcomes [44-46]. Also, counselor behaviors such as advising, confronting, directing, and warning clients are associated with greater sustain talk, whereas counselor behaviors such as affirming strengths or effort, emphasizing client autonomy, and being supportive are associated with increased change talk [47,48]. Thus, to facilitate behavior change, the goal of an MI counselor is to interact with the client in a manner that facilitates change talk and inhibits sustain talk [38,48].

Standard MI training often consists of a 2-day workshop (about 14 hours total) with didactic presentations and many practice and experiential exercises. After this, attendees typically demonstrate an increase in MI relational skills (eg, asking permission before sharing information and more collaborative interactions). Demonstration of MI technical skills, such as using more open-ended questions and reflections, especially complex reflections [49-52] that are key components of MI's mechanisms of action, is less likely after a 2-day workshop. Furthermore, acquired skills tend to diminish over the following 8-12 weeks without postworkshop feedback and coaching [53-55].

Studies have shown that CSS can become competent in MI [6,56], but in real-world settings, the training needed for CSS to attain MI competency is not readily available. This is a key challenge to disseminating and implementing MI [38,42,43]. Thus, we need to develop novel training approaches that are effective at teaching the technical MI skills that are key to MI's efficacy and make these trainings more easily accessible to CSS in community settings. To ensure that MI retains its demonstrated efficacy in improving HCC outcomes, these MI training approaches need to be scalable so that they can both reach CSS working in HIV care settings and effectively increase their MI skills.

The World Health Organization has called for digital training for health care providers [57] and the broad use of online training for other behavioral interventions [58-60]. However, the use of technology in MI training has remained mostly limited to video demonstrations and brief didactic webinars that describe MI and provide examples of MI skills—even among the many MI-based health interventions that have been developed [61-63]. Findings from studies of technology-driven MI training have been encouraging. For example, multisession synchronous online MI training is feasible and effective [63,64], with no difference in counselor skill acquisition between in-person and online workshops [64]. Mitchell et al [65] used Second Life, a virtual world, to train physicians to increase acceptance of colorectal screening using 3 sessions of combined prerecorded MI tutorials, two 90-minute group role-play and feedback sessions with an MI coach, and between session practices, and thereby increased physician's MI global skills related to client-centeredness (there was no assessment of individual technical skills). Other projects have begun to use avatars and simulated patient interactions to build MI skills [66-68]. These allow users to interact with virtual patients who are trained to respond in a certain manner to the words being used by the provider. To date, however, there are no studies that report on the efficacy of these training approaches, whether they are best used for beginner or advanced learners, or whether they need to be combined with other training approaches. Also, there are significant concerns about the costs to develop these patient interactions, especially because each simulation has to be constructed individually using words that are typical in the

XSL•FO
RenderX

subject area the training addresses [68]. Finally, although technology-based MI training has high acceptability among users [63-68] and has shown preliminary evidence of efficacy in building MI skills [64,65], there are challenges to participant engagement and retention [63,68,69], highlighting the need for further research on novel technology-based MI training approaches to improve user engagement.

This study seeks to address a critical need for novel training approaches to build skills among CSS to deliver MI in health care interactions across the status-neutral HCC. Referred to as "iMI4HIV" (internet-based motivational interviewing training program for HIV counseling support staff), this study will offer a new approach to online MI training, specifically for HIV care settings. First, it will offer comprehensive interactive online training (10-12 hours) aimed at helping users learn how to use MI across a range of HCC interactions. Second, it will approach asynchronous online MI training in a completely novel way, using multiple MI video demonstrations combined with gamified exercises to sequentially build skills that are key to MI's mechanisms of action, specifically the relational and technical skills essential for increasing change talk and inhibiting sustain talk. Third, it will integrate a variety of gamification and skills training components to increase engagement and appeal across a range of users [70-84]. Finally, this study includes a detailed framework for developing and evaluating the iMI4HIV strategy that includes HCC stakeholders, MI subject matter experts, and HCC learners in the development and piloting phases of the study, using mixed methods to evaluate the learner experience, engagement, and outcomes for MI learning.

Study Objectives

The overall goal of the study is to develop and pilot an interactive, gamified, online training program in MI for HIV CSS to improve health communication in status-neutral HIV care interactions. This National Institutes of Health (NIH) R34 (Clinical Trial Planning Grant) study will allow us to pilot all the components and build the necessary research infrastructure for a subsequent randomized controlled trial (RCT) that is fully powered to assess the efficacy of iMI4HIV in increasing MI skills among CSS in HIV care interactions. The specific aims of this R34 study are as follows:

- Develop iMI4HIV, an interactive, online, gamified MI training program specifically tailored toward CSS in HIV care settings to improve their MI skills with HIV care clients.
- Conduct a pilot RCT of a virtual motivational interviewing (VMI) workshop + iMI4HIV versus VMI workshop + waitlist control (n=30) to (1) assess the feasibility and acceptability of iMI4HIV as measured by participant completion rate and retrospective acceptability ratings of iMI4HIV, and (2) explore preliminary findings on the effects of iMI4HIV on MI skills acquisition using Standard Patient Interactions (SPIs).
- Explore experiences (including engagement, obstacles, and facilitators) of completing iMI4HIV via in-depth interviews with pilot study participants.

In this paper, we present the proposed procedures for developing iMI4HIV, conducting the pilot RCT, and assessing its acceptability among HIV CSS.

Methods

Study Design

Overview

This study will have 2 phases. In phase 1 (Development), we will conduct formative focus groups (FGs) with CSS (n=16) to identify HIV care interactions for iMI4HIV videos and to assess the acceptability of gamification components; create MI demonstration videos depicting key interactions from status-neutral HCC and program iMI4HIV (including skills development tasks, gamification components, etc); and pilot the VMI workshop and iMI4HIV (n=8). We will also conduct qualitative interviews (QIs) with 8 leaders from HIV care agencies to explore their use of existing online MI training programs, acceptability of iMI4HIV, and possible facilitators and impediments to its adoption for training HIV CSS. In phase 2 (Pilot RCT), we will randomize 30 CSS at a 2:1 ratio (VMI workshop + iMI4HIV vs VMI workshop + waitlist control) to assess the feasibility and acceptability of iMI4HIV and to pilot the RCT processes and assessments for a future efficacy trial. We will also explore preliminary findings on the impact of iMI4HIV on MI skills acquisition. At the outset of the study, we will convene a Community Advisory Board (CAB) composed of 7 members, including individuals involved with HIV-related community health worker (CHW) training programs and HIV CSS with at least 1 year of experience providing services in HIV care settings. The CAB will meet via videoconferencing quarterly throughout the study.

Recruitment of CSS

Recruitment of participants will occur through 2 main recruitment streams. First, we will work with our partners at the Northeast and Caribbean AIDS Education and Training Center (NECA AETC) to help us recruit CSS through their network of collaborating HIV care agencies (a maximum of 2 CSS per clinic). Second, we will recruit through HIV-focused CHW training programs. These 2 recruitment streams will provide us with a group of CSS trainees that will be diverse in terms of HIV counseling experience and training. Eligibility criteria for all CSS study participants will be (1) aged 18 years or older, (2) currently works in a CSS role to improve HIV care outcomes, (3) English-language fluency (since iMI4HIV will only be developed in English during this study), and (4) access to a laptop, desktop, or tablet computer to attend the virtual training and complete iMI4HIV.

Phase 1: Development

MI Training Content Development

We envision approximately 10-12 hours of training content organized by HCC component (Textbox 1). For each component, videos (each 8-18 minutes in duration) will be developed, highlighting different HCC-related interactions. The videos, by themselves, will serve as a resource to demonstrate proficient MI practice in these interactions. Furthermore, we will

XSL•FO
RenderX

incorporate game-like interactive exercises aimed at building key MI skills that are critical to triggering its mechanisms of action, such as differentiating between open and closed questions, simple and complex reflections, and identifying and responding to change talk and sustain talk. In iMI4HIV, we will include some of the most frequently used gamification components (eg, points, feedback, levels, and badges of recognition) [75-78] while eschewing others, such as leaderboards (sharing a user's points with all users to create competition), which may be counterproductive for this type of learning intervention [84]. The gamification components likely to be embedded in iMI4HIV include points to award learners for correct responses, feedback to encourage or prevent specific activities in order to improve performance, levels to stipulate that only users who attain a certain achievement can progress to the next level, and badges as a symbolic representation of a user's achievement or competence.

Textbox 1. Components of an internet-based motivational interviewing training program for HIV counseling support staff.

Video content

- Motivational interviewing (MI) basics
  - MI spirit
  - Four tasks of MI
  - MI micro skills: open-ended questions, affirmations, reflections, and summaries
  - Evoking, recognizing, and reinforcing change talk
  - Responding to sustain talk
  - Planning for change
- HIV testing
  - Overcoming ambivalence to testing
  - Facilitating regular testing
  - Reducing risk behaviors
- Linkage to care
  - Overcoming ambivalence to entry to care after HIV-positive results
  - Linkage to pre-exposure prophylaxis or other prevention services
  - Building motivation for linkage to care to ancillary services (mental health and substance abuse treatment)
- Retention in care
  - Exploration of the patient's perceived benefits of staying in care
  - Collaborative planning to overcome obstacles to retention in care
  - Overcoming ambivalence to entry into care after HIV-positive results
  - Linkage to pre-exposure prophylaxis or other prevention services
  - Building motivation for linkage to care to ancillary services (ie, mental health and substance abuse treatment)
- Treatment adherence and viral suppression
  - Exploring the client's views on adherence
  - Building motivation for adherence
  - Using viral load and CD4 lymphocyte counts to build motivation to adhere

Sample tasks

- Self-reflection exercises
- Brief didactic presentations
- Animation
- Gamification (points for picking open-ended questions or choosing the best reflection; clicking when you hear change talk)
- Coding brief segments for MI fidelity

XSL•FO
RenderX

Development of iMI4HIV will be guided by the User Centered Rapid Application Development process (UCRAD) [85,86]. UCRAD merges the streamlined, iterative Rapid Application Development (RAD) approach [87] with the User Centered Design approach [88,89], which engages intended users throughout the development process. This integrated approach aims to develop successful computer apps with good functionality, simple features, and a usable interface [90]. As such, intended users are given access to prototypes of the program, allowing them to provide feedback before the next iteration. We expect the process to take approximately 12 months in this study to produce the many scenarios and components of iMI4HIV. UCRAD uses a 3-stage process:

Predesign and Interface Prototyping

To accomplish this stage, we will obtain feedback from the CAB regarding the preliminary plans for iMI4HIV, FGs with CSS, and QIs with HIV care agency leaders:

- FG processes. We will conduct 2 real-time, 90-minute synchronous online FGs, each with 8 CSS, using a HIPAA (Health Insurance Portability and Accountability Act)-compliant video conference platform that allows for quantitative polling of participants regarding key attributes of iMI4HIV. The FGs will elicit individual- and group-level perspectives on topics that include (1) challenges faced in helping patients engage in care and obtain or sustain viral suppression, (2) prior MI training and use with patients, (3) challenges to learning and delivering MI, (4) feedback on iMI4HIV and different gamification components (eg, accumulating certain number of points before being able to proceed to next stage; preferences for types of games), and (5) other recommended training content. To obtain feedback on the gamification components, simple gamification components will be integrated into the FG discussion so that participants can provide feedback based on the gamification experience.
- QIs with HIV Care Agency Leadership. We will interview 8 leaders (including program directors who directly oversee CSS) from HIV care agencies associated with the NECA AETC. Given the extensive existing literature on barriers and facilitators to MI implementation, these interviews will focus more specifically on (1) knowledge and use of internet-based MI training, (2) barriers and facilitators to internet-based MI training they are currently using, (3) presentation and feedback on iMI4HIV, (4) recommendations for additional content, and (5) barriers and facilitators to CSS completing iMI4HIV.
- Rapid data analysis. Within 48 hours of completing each FG and QI, we will initiate a rapid data analysis [91,92] using a qualitative matrix [93,94] onto which participant responses will be entered to quickly organize and summarize FG findings. Summarized findings will be shared with the CAB and the external MI training consultants to obtain their feedback.

System Architecture and Coding

We will convene an iMI4HIV development team composed of MI trainers, program developers, and 4 CSS (selected from FG participants for their contributions and comfort in providing feedback). This team will meet monthly online over 6 months. Prior to each meeting, CSS will be asked to use prototype versions of a different iMI4HIV module (including the video, MI learning tasks, and gamification) and provide feedback. During the meeting, we will explore their experience completing the module, including what they liked or disliked, their reactions to the learning tasks and gamification components, and their overall experience of interacting with iMI4HIV.

Deployment

We will recruit 8 CSS from HIV care agencies and CHW training programs to pilot all training components. They will complete a 15-hour (3 days, 5 hours per day) HIV-focused VMI workshop conducted by experienced MI trainers that will consist of an overview of MI and extensive opportunities for joint interactive and experiential exercises conducted as a group and in dyads in different virtual rooms.

After completing the VMI workshop, participants will be sent an online questionnaire to provide feedback on the workshop and will be given access to iMI4HIV, which they will be asked to complete within 4 weeks. After completing iMI4HIV, or 5 weeks after being provided access, they (including those who were assigned to but did not complete iMI4HIV) will be asked to complete (1) an online questionnaire to assess acceptability of iMI4HIV and (2) an audio-recorded qualitative telephone interview to assess (a) overall acceptability of training program (workshop and iMI4HIV), (b) likes and dislikes of the program, (c) experience with gamification components, (d) difficulties encountered, and (e) feedback for the final version of the training program.

Phase 2: Pilot RCT

Overview

The overall aim of phase 2 is to assess the feasibility and acceptability of iMI4HIV and prepare the necessary research infrastructure and processes for a future fully powered RCT to test its efficacy in increasing MI skills among CSS in HIV-related interactions. Thus, we will recruit 30 CSS and randomize them using a 2:1 ratio (VMI workshop + iMI4HIV vs VMI workshop + waitlist control).

Figure 1 presents an overview of the phase 2 RCT procedures, detailing an RCT with a waitlist control arm which will gain access to iMI4HIV after completion of the Follow-Up 1 assessment. This design allows us to achieve the specific aims of the study as well as gain initial insights into the effects of iMI4HIV on MI skills and how these skills may be retained after 3 months. After all participants complete a baseline computer-assisted self-interview (CASI), they will complete SPI-1, then take part in the VMI workshop. After that, they complete SPI-2 and are then randomized into iMI4HIV versus a waitlist control. After 4 weeks, all participants complete Follow-Up 1, which consists of SPI-3, a follow-up CASI, and for only those in the iMI4HIV arm a qualitative interview. This will allow us to assess the feasibility and acceptability of iMI4HIV among those who were randomized into that arm (n=20). At this assessment point, we will also explore the effects of iMI4HIV on MI skills by comparing MI fidelity ratings of the SPI-3 between the intervention and control groups.

XSL•FO
RenderX

After Follow-Up 1, those in the control arm will be given access to iMI4HIV; after 4 weeks, only this group will complete Follow-Up 2, which will consist of SPI-4, a follow-up CASI, and an in-depth interview, which will provide additional acceptability data. Participants originally randomized into iMI4HIV will undergo their Follow-Up 2 after 12 weeks (SPI-4 and CASI, no QI). This will allow us to gain insights into retention of MI skills after 3 months, a key data point for a subsequent efficacy trial of iMI4HIV in which we would want to assess maintenance of MI skills. Community-based organizations rarely offer MI coaching to staff after a training workshop. Thus, we will not offer MI coaching in either study arm to focus on assessing the preliminary effects of iMI4HIV on participants' MI skills.

Figure 1. Phase 2 pilot randomized controlled trial overview. CASI: computer-assisted self-interview; iMI4HIV: internet-based Motivational Interviewing training program for HIV CSS; MI: Motivational Interviewing; QI: qualitative interview.

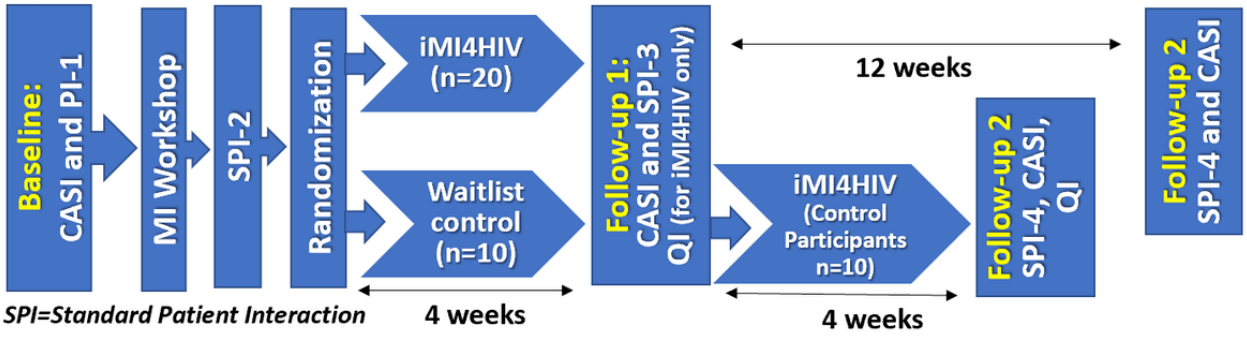

Recruitment of CSS

We will work with HIV care clinics associated with the NECA AETC and CHW training program partners to enroll 30 CSS in the RCT. Interested CSS will be instructed to contact the study research assistant who will screen them for eligibility. Eligibility criteria will be as described above for the phase 1 CSS participants. If these criteria are met, research staff will send them an "enrollment email," which will include a link to an online informed consent form.

Our assessment of acceptability of iMI4HIV will be guided by the Theoretical Framework of Acceptability [95] (TFA; refer to components in Textbox 2), which was derived from a review of the literature on acceptability, and views acceptability as "a multi-faceted construct that reflects the extent to which people delivering or receiving a healthcare intervention consider it to be appropriate, based on anticipated or experienced cognitive and emotional responses to the intervention" [95]. The TFA has guided qualitative and mixed methods research on the acceptability of health care interventions among patients and providers [96-99]. The TFA's 7 domains have informed the development of a generic TFA quantitative assessment with Likert-type items to assess each component of the model [100].

Textbox 2. Components of the Theoretical Framework of Acceptability.

Theoretical Framework of Acceptability construct definitions

- Affective attitude: how the individual feels about the intervention
- Burden: perceived amount of effort that is required to participate in the intervention
- Ethicality: the extent to which the intervention has a good fit with an individual's value system
- Intervention coherence: the extent to which the participant understands the intervention and how it works
- Opportunity costs: the extent to which benefits, profits, or values are given up to engage in the intervention
- Perceived effectiveness: the extent to which the intervention is perceived as likely to achieve its purpose
- Self-efficacy: the participant's confidence that they can perform the behaviors required to participate in the intervention

Baseline Assessments

After enrolling, participants will receive an email with a link to the CASI to be completed within 7 days and will be scheduled for their SPI-Baseline session, which will be audio-recorded and rated for MI fidelity. The CASI and the SPI-Baseline must be completed prior to the MI workshop. The baseline CASI assessment will include the following:

- Demographics. This will capture information on age, gender, race, ethnicity, education, primary language, counseling training, and prior MI training.
- MI Knowledge and Attitudes Test (MIKAT) [101], a 29-item assessment that has shown sensitivity to detect change in MI consistent and inconsistent behavior in pre- and posttraining comparisons [101-104], with acceptable internal consistency [104]. The MIKAT will be adapted to diminish its focus on MI for substance use treatment.
- Perceptions of Evidence-Based Practices (EBPs) will be evaluated using 4 self-report measures:
  1. The Evidence-Based Practice Attitudes Scale-15 [105] is a 15-item assessment of providers' attitudes toward EBPs, with 4 factors: Appeal, Requirements, Openness,

XSL•FO
RenderX

and Divergence. Participants will also complete 2 assessments of their perceptions of how EBPs are viewed by their supervisor and at the agency.

2. The Implementation Climate Scale [106], reliably assesses the extent to which a clinic fosters EBP implementation.
3. The Implementation Leadership Scale [107] assesses leadership support for EBP implementation.
4. TFA quantitative questionnaire [100]. Finally, we will assess the prospective acceptability of iMI4HIV using a tailored version of the generic TFA questionnaire. This questionnaire is composed of a 5-point Likert-type scale for each of the TFA components and 1 general question on acceptability. Findings from phase 1 will also identify other areas of inquiry for the baseline CASI.

- SPIs. We will use scheduling software for study participants to schedule their SPIs. At the specified time, the SPI actor will contact the participant by phone to conduct a 15-minute MI session focused on an HIV care scenario. The session will be audio-recorded and then rated using the Motivational Interviewing Treatment Integrity ratings (MITI 4.2) [108], a valid and reliable measure of MI skills [109]. MI skills will be assessed at baseline, postworkshop, post iMI4HIV, and 3 months later (for participants randomized into iMI4HIV).

MI Virtual Workshop

Synchronous VMI workshops will be conducted through Zoom (Zoom Video Communications, Inc). We expect 2 to 3 workshops with 10-15 CSS in each to facilitate individualized feedback during the workshops. Afterward, participants will receive a brief survey asking for feedback on the workshop.

Randomization

As shown in Figure 1 above, after completing the SPI-2, participants will be randomized into iMI4HIV or waitlist control at a 2:1 ratio. After completing the VMI workshop, iMI4HIV participants will be sent a link to set up their account and access iMI4HIV. They will be asked to complete the online training within 4 weeks. Participants will receive weekly reminders to complete the training within 4 weeks, at which time access to the program will end. Waitlist control individuals will not have access to iMI4HIV during the 4 weeks following the VMI workshop but will get access once their follow-up assessment and SPI-3 are completed.

Follow-up Assessment 1

Four to 5 weeks after randomization, participants in both arms will complete the first follow-up assessment, which forms the basis for comparing the 2 RCT arms and will consist of a CASI and SPI #3. Participants in the iMI4HIV arm will also take part in a QI. Participants in the iMI4HIV arm who did not complete the program will still be invited to complete the follow-up assessment (for which there is an incentive) so we can learn about obstacles to completing the program and obtain their recommendations.

- Follow-up CASI. This assessment will repeat those measures conducted at baseline but tailored to each study arm. For the intervention arm, the follow-up CASI will additionally include iMI4HIV-specific assessments: (1) TFA-based questions to assess acceptability of iMI4HIV, (2) obstacles to completing iMI4HIV, and (3) self-report of any additional MI training (eg, watching YouTube videos and reading MI articles or books) in which participants engaged. Findings from phase 1 will also identify other areas of inquiry for the follow-up CASI.
- SPI-3. After completing the follow-up CASI, participants in both arms will complete the third SPI as previously described, but using a different HIV-related patient scenario.
- QIs with CSS. Participants in the iMI4HIV arm will then have a QI to explore topics including (1) the TFA constructs, (2) iMI4HIV components deemed enjoyable, educational, and helpful, (3) acceptability of different gamification components, (4) facilitators and obstacles to completing iMI4HIV, (5) perceptions of current competency in MI, including strengths and weaknesses, (6) factors that have facilitated or hindered the use of MI in HIV care interactions, and (7) recommendations for further revisions of iMI4HIV. QIs are expected to be approximately 60 minutes in duration, and will be audio-recorded, transcribed, and verified for accuracy before being deidentified and uploaded to NVivo (Lumivero) for coding and analysis.

Follow-up Assessment 2

The second follow-up assessment will vary based on the study arm. Once participants in the control arm have completed their Follow-up Visit 1, they will be given access to iMI4HIV. They will be asked to complete the program within 4 weeks, at which point they will complete Follow-up Visit 2, consisting of (1) follow-up CASI assessment, (2) SPI-4, and (3) a QI. This follow-up CASI will be the same as that completed by participants in the active iMI4HIV arm at their first follow-up assessment. The SPI-4 will use a different scenario from those used in previous SPIs and will be conducted using the same procedures as the baseline SPIs. The QI will use the same as guide for the first follow-up assessment for those in the iMI4HIV arm; however, it will include additional questions exploring the impact of delayed access to iMI4HIV after completing the virtual workshop. Intervention arm participants will complete their Follow-up Visit 2 three months after their Follow-Up Visit 1. The assessment will repeat the CASI assessment from the first follow-up assessment and they will complete SPI-4. This will allow us to explore the degree to which gains in MI skills at SPI-3 are retained after 3 months.

Data Analysis

Qualitative Data Analysis for QIs

Transcripts from the phase 2 QI audio-recordings will be reviewed for accuracy and uploaded into NVivo for management and analysis. Codebook development will begin when 8 transcripts are available. First-level codes will be guided by the TFA components, while second- and third-level codes will emerge from themes identified in the narratives. An initial set of codes will be generated independently by 2 research staff, compared, then synthesized to compile shared coding categories

XSL•FO
RenderX

and subcategories, all with definitions, inclusion and exclusion criteria, and examples. As transcripts become available, coding will continue, refining definitions that will be used to process the remainder of the manuscripts. Coders will discuss discrepancies until they achieve 80% intercoder convergence. As per Patton [110], we will identify indigenous (ie, participant-generated) as well as analyst-constructed typologies. Once data are coded, analysis of coding reports will include categorization, abstraction, comparison, integration, iteration, and refutation of themes.

Integration of Quantitative and Qualitative Data

We will exploit the richness of our mixed methods approach, with quantitative and qualitative findings enriching each other. Quantitative assessments allow us to systematically explore variables of interest across all participants. Qualitative data will add nuance and new insights to quantitative findings. During data analysis, interpretation of quantitative results will be enriched by the summary of codes for specific components of iMI4HIV. In NVivo, we will categorize participants based on their responses in the quantitative assessment and use these categories to compare and contrast themes and concepts that emerge from the qualitative data.

Quantitative Data Analysis

Primary Outcomes

The primary outcomes of the pilot study will be based on the phase 2 RCT pilot, including (1) the percent of study participants that complete iMI4HIV within 4 weeks posttraining workshop (feasibility) and (2) the percentage of participants that respond Acceptable or Completely Acceptable to the question, "How acceptable was iMI4HIV to you?" from the TFA generic acceptability questionnaire (acceptability).

Exploratory Outcomes

We will also explore the following outcomes:

- MI skills acquisition. MITI-4.2 ratings for (1) Technical Component, (2) Relational Component, (3) Percent Complex Reflections, (4) Reflection:Question Ratio, (5) Total MI Adherent, and (6) Total MI Non-Adherent. Although the pilot RCT is not powered to detect differences between the intervention and control groups, we will compare these scores between the 2 groups for the SPI-3 assessment and within groups for each of the SPI assessments.
- Feasibility. Metadata from iMI4HIV, including mean (1) duration to complete specific iMI4HIV tasks, (2) number of times a lesson was completed prior to achieving a passing score, and (3) score for each task.
- Acceptability. Mean ratings on each of the generic TFA items (Affective Attitude, Burden, Intervention Cohesiveness, Self-efficacy, Ethicality, Perceived Effectiveness, and Opportunity Cost).

Data collected via assessments will be imported into IBM SPSS Statistics to calculate descriptive statistics for the sample. Estimation of the feasibility and acceptability outcomes (primary and exploratory) will be accomplished by generating the point estimate with its corresponding 95% CI. We will use a *t* test for continuous outcomes, and Fisher exact test for binary outcomes for the above comparisons. Adjustment for covariates, if necessary, with control baseline values, will be accomplished by multiple linear (continuous outcomes) or logistic (dichotomous outcomes) regression. While this study will not recruit a large enough sample to provide adequate power for assessing the efficacy of iMI4HIV, we will use data from MITI-4 ratings to explore specific changes in MI skills (listed in exploratory outcomes, above). To examine the efficacy of the addition of iMI4HIV on MI skills, the generalized linear model with identity and logit link function will estimate model parameters and their SEs for continuous and dichotomous variables, respectively. The generalized estimating equation method, with a robust SE estimator, will account for the within-subject correlation due to repeated measures from the same participant. Accompanying each of these tests will be point and interval estimates for the parameters of interest, such as group mean differences and odds ratios.

Power Considerations

There is some controversy regarding the use of pilot data to estimate an actual effect size [111]; clinically meaningful effect sizes should be decided based on extrinsic, clinical judgment grounds, and not based on pilot data which are too few, typically, to obtain reliable conclusions. We can use pilot data to rule out unusually large or small true effects through standard 95% CI procedures. We will confirm that extrinsic effect sizes are contained within our CIs from the pilot. We will examine the distribution of each variable and calculate summary statistics by intervention condition. We will estimate key intervention parameters with sample means and proportions together with 2-sided 95% CIs and test the primary null hypotheses at the traditional 2-sided level α=.05 (to simulate the subsequent RCT). For planning the RCT, we will also consider 1-sided 90% confidence limits for mission-critical design parameters such as SDs and reference group end point rates and proportions in the conservative direction. This is because we intend to plan the sample size for the RCT so that power will be "excellent" (at least 80% power) for the clinically relevant effect size to be specified if our estimate of SDs is at the 1-sided 90% confidence limit in the conservative direction. This strategy will make proper allowance for the limited sample size of the pilot study with its consequent uncertainties, and still yield appropriate sample sizes for the future RCT. We will examine the strength of association between our primary outcome and key variables, with the view toward seeing if stratifying on such factors (eg, age and gender) would substantially reduce variability and increase the power of a future RCT. "Mission-critical" parameters include proportions for dichotomous end points; means and SDs for continuous end points; and response profiles for longitudinal studies. Large sample sizes are not required to locate these parameters approximately, but adequately, for planning the subsequent trial, whereas testing the study hypothesis in the pilot will generally not have sufficient statistical power. Estimation of mission-critical design parameters with point and CI estimates will be very important; continuous outcome measures make it feasible to detect promising effects even in small samples. For the continuous measures, means and SDs will be estimated. For approximately

XSL•FO
RenderX

normally distributed variables, an upper 1-sided 90% confidence limit for the population SD, s_U, will be constructed from $s_U = s \times [(n-1)/x^2_{0.10,n-1}]^{1/2}$, where s is the sample SD and $x^2_{0.10,n-1}$ is the upper 10th percentile of the chi-square distribution with $n-1$ degrees of freedom. For approximately log-normally distributed variables, a logarithmic transformation will be applied to achieve approximate symmetry and normality. For the dichotomous variable, exact binomial methods will be used.

Ethical Considerations

The study was reviewed and approved by the Florida State University (FSU) Institutional Review Board (ID: STUDY00004477). This study does not assess health-related biomedical or behavioral outcomes; thus, it does not meet NIH criteria for a clinical trial and will not be registered in ClinicalTrials.gov.

All participants are provided with a written informed consent form to read before participating in the study. The consent form includes an overview of the study, study procedures, risks of participation, procedures to minimize risks of participation, a statement of voluntary participation, and contact information for the principal investigators, along with contact information for the FSU Institutional Review Board.

The results of the study may be published or presented, but no information that can identify individual participants will ever be provided or released in publications or presentations. To protect the privacy and confidentiality of participants, the study staff will keep any records containing identifying information in a locked cabinet accessible only to the research staff. The computer programs used in tablet-based interventions are compliant with US regulations on personal health information, as are the servers on which the data are stored. All data captured on the tablet, although it do not include identifying information, are encrypted in transit to maintain confidentiality. Any other study data will be stored on a secure computer that is protected by passwords. All records will be kept confidential. This is consistent with the guidelines of the US NIH. After removing all identifying information, the remaining deidentified information collected during the study could be used for future research studies or distributed to another investigator for future research studies without additional consent.

All participant incentives will be paid via gift cards. In phase 1, FG participants received US $75; agency leaders received US $100 for the interview; and iMI4HIV pilot participants will receive US $50 plus the opportunity to win a US $250 prize for completing the intervention within the 4-week allotted period. This incentive is provided to motivate the timely completion of the online training program among these participants to inform any final revisions that are necessary before the phase 2 RCT; main study outcomes are based on the phase 2 RCT results, which do not include this raffle incentive. Phase 2 RCT participants will receive up to US $205 for completing questionnaires (3 timepoints), 4 standard patient interactions, and a qualitative interview.

Results

The study launched in February 2024. During the summer of 2024, we conducted 2 FGs (n=11) and QIs (n=8) with HIV Care Agency Leadership to inform the content and interactive components of the iMI4HIV. Preliminary findings from the FGs and QIs were shared with the CAB and expert consultants. In the Fall of 2024, we recorded a series of training videos focused on using MI in key HIV care interactions. Since then, we have been developing the content and layout of the training program. We expect to conduct the phase 1 pilot study in the Fall of 2025 and the pilot RCT during 2026.

Discussion

Overview

To address the need for online MI training geared specifically for the HIV care setting, this mixed methods study seeks to develop iMI4HIV, an interactive, gamified, online MI training program specifically tailored for CSS to improve their MI skills in HIV care communications. The goal is to train CSS to deliver MI competently so they can integrate MI into their health care communications with individuals seeking or receiving HIV-related services (status-neutral). If successful in building MI skills among CSS, iMI4HIV has the potential to shift how MI training is provided to CSS who work in HIV care settings. Furthermore, the program can be housed at FSU for use by AETCs throughout the country to provide additional MI training to their client agencies, used to train HIV treatment counselors in research studies, and provide training opportunities for other community-based agencies working to end the HIV Epidemic.

Strengths

This proposed study has numerous strengths. First, contrary to the currently available simulations for MI training, iMI4HIV will offer an innovative, sequenced approach to comprehensive online MI training with a range of interactive and gamification components used in novel ways to build key MI technical skills. As such, it will also be the first to assess user acceptability, feasibility, and engagement of this new approach to online MI training. Second, at an expected 8-10 hours, iMI4HIV will be the first and most comprehensive interactive online MI training that addresses the full status-neutral HIV care cascade. The only existing interactive HIV-focused MI training is approximately 1 hour in duration and only addresses treatment adherence counseling for nurses. Third, this is the first online MI training study to systematically engage targeted end users throughout the development and testing process to obtain their feedback for revisions. Fourth, based on current scientific literature, this study will be among the first to conduct a rigorous assessment of efficacy (even within an underpowered RCT).

Limitations

Although we anticipate high interest in MI training and iMI4HIV, recruitment and retention challenges can always arise and we consider these the main potential challenges to this study. Additionally, acceptability data (both quantitative and qualitative) are subject to social desirability. To minimize this risk, we will highlight for participants the importance of openly

XSL•FO
RenderX

sharing their experiences and opinions, whether positive or negative.

Conclusions

As designed, this study specifically responds to the Notice of Special Interest: Advancing Health Communication Research on HIV Prevention, Treatment and Cure (NOT-MH-21-105) announced by NIH, which includes an interest in studies that seek to "optimize effective communication and engagement practices between PLHIV and healthcare providers." This R34 begins to address a critical need for novel approaches to train CSS in MI, a counseling or communication approach with proven efficacy in improving health care outcomes across the HIV care cascade. With its focus specifically on MI in HIV care interactions, iMI4HIV gives CSS in HIV care settings an invaluable opportunity to solidify MI skills gained during MI training workshops. Furthermore, as an asynchronous training program, iMI4HIV provides significant flexibility in when it can be completed, increasing accessibility to potential users. If findings on feasibility (≥70% of participants complete the iMI4HIV modules within 4 weeks) and acceptability (≥80% of participants rate iMI4HIV a "4" or "5" on the TFA overall acceptability rating) are favorable, we will submit a proposal for a fully powered RCT to assess the efficacy of iMI4HIV to achieve MI competence. In that study, we may even consider a study arm using an expanded version of iMI4HIV to allow for fully asynchronous MI training, which was cost-prohibitive in this study.

Funding

This research was supported by a grant from the US National Institute of Mental Health (R34 MH133468, Principal Investigator ICB). The authors are solely responsible for the content, which does not necessarily represent the official views of the National Institutes of Health (NIH).

Data Availability

Data to be shared will be deposited with the digital repository of the Inter-university Consortium for Political and Social Research (ICPSR) to ensure that the research community has long-term access to the data. Once data have been deidentified and uploaded to a data repository, it will also be available on request from the investigators.

Authors' Contributions

Conceptualization: ICB, BAK
Methodology: ICB, BAK, RG
Funding acquisition: ICB
Project administration: ICB, RG
Investigation: ICB
Supervision: ICB
Writing – original draft: ICB (study protocol). ORB (manuscript adaptation)
Writing – review & editing: ICB, ORB, BAK, RG

Conflicts of Interest

None declared.

Multimedia Appendix 1
Peer-review report from HIBI - HIV/AIDS Intra- and Inter-personal Determinants and Behavioral Interventions Study Section, Risk, Prevention and Health Behavior Integrated Review Group (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 151 KB - resprot_v15i1e82241_app1.pdf]

References

1. Ambia J, Mandala J. A systematic review of interventions to improve prevention of mother - to - child HIV transmission service delivery and promote retention. Journal of the International AIDS Society 2016 Apr 06;19(1):20309. [doi: 10.7448/ias.19.1.20309]
2. Coleman LM, Ford N. An extensive literature review of the evaluation of HIV prevention programmes. Health Educ Res 1996 Sep;11(3):327-338. [doi: 10.1093/her/11.3.327] [Medline: 10163565]
3. Outlaw AY, Naar-King S, Parsons JT, Green-Jones M, Janisse H, Secord E. Using motivational interviewing in HIV field outreach with young African American men who have sex with men: a randomized clinical trial. Am J Public Health 2010 Apr 01;100 Suppl 1(Suppl 1):S146-S151. [doi: 10.2105/AJPH.2009.166991] [Medline: 20147689]
4. Mwai GW, Mburu G, Torpey K, Frost P, Ford N, Seeley J. Role and outcomes of community health workers in HIV care in sub-Saharan Africa: a systematic review. J Int AIDS Soc 2013 Sep 10;16(1):18586 [FREE Full text] [doi: 10.7448/IAS.16.1.18586] [Medline: 24029015]

XSL•FO
RenderX

5. Global Health Workforce Alliance. Will we achieve universal access to HIV/AIDS services with the health workforce we have? A snapshot from five countries: Literature review. Report of the Technical Working Group on Human Resources for Health and Universal Access. 2010. URL: https://www.who.int/workforcealliance/about/taskforces/access/universalaccesslitrev_en.pdf [accessed 2020-04-23]
6. Naar-King S, Outlaw A, Green-Jones M, Wright K, Parsons JT. Motivational interviewing by peer outreach workers: a pilot randomized clinical trial to retain adolescents and young adults in HIV care. AIDS Care 2009 Jul;21(7):868-873. [doi: 10.1080/09540120802612824] [Medline: 20024744]
7. Kenya S, Chida N, Symes S, Shor-Posner G. Can community health workers improve adherence to highly active antiretroviral therapy in the USA? A review of the literature. HIV Med 2011 Oct;12(9):525-534 [FREE Full text] [doi: 10.1111/j.1468-1293.2011.00921.x] [Medline: 21518221]
8. Miller WR, Rollnick S. Motivational Interviewing: Helping People Change. New York, NY: Guilford Press; 2013.
9. DiClemente CC, Corno CM, Graydon MM, Wiprovnick AE, Knoblach DJ. Motivational interviewing, enhancement, and brief interventions over the last decade: A review of reviews of efficacy and effectiveness. Psychol Addict Behav 2017 Dec;31(8):862-887. [doi: 10.1037/adb0000318] [Medline: 29199843]
10. Lundahl BW, Kunz C, Brownell C, Tollefson D, Burke BL. A Meta-Analysis of Motivational Interviewing: Twenty-Five Years of Empirical Studies. Research on Social Work Practice 2010 Jan 11;20(2):137-160. [doi: 10.1177/1049731509347850]
11. Randall CL, McNeil DW. Motivational Interviewing as an Adjunct to Cognitive Behavior Therapy for Anxiety Disorders: A Critical Review of the Literature. Cogn Behav Pract 2017 Aug;24(3):296-311 [FREE Full text] [doi: 10.1016/j.cbpra.2016.05.003] [Medline: 28871216]
12. Marker I, Norton PJ. The efficacy of incorporating motivational interviewing to cognitive behavior therapy for anxiety disorders: A review and meta-analysis. Clin Psychol Rev 2018 Jun;62:1-10. [doi: 10.1016/j.cpr.2018.04.004] [Medline: 29727863]
13. Miller SJ, Foran-Tuller K, Ledergerber J, Jandorf L. Motivational interviewing to improve health screening uptake: A systematic review. Patient Educ Couns 2017 Feb;100(2):190-198 [FREE Full text] [doi: 10.1016/j.pec.2016.08.027] [Medline: 27599713]
14. Keeley RD, Brody DS, Engel M, Burke BL, Nordstrom K, Moralez E, et al. Motivational interviewing improves depression outcome in primary care: A cluster randomized trial. J Consult Clin Psychol 2016 Nov;84(11):993-1007. [doi: 10.1037/ccp0000124] [Medline: 27599229]
15. Lundahl B, Moleni T, Burke BL, Butters R, Tollefson D, Butler C, et al. Motivational interviewing in medical care settings: a systematic review and meta-analysis of randomized controlled trials. Patient Educ Couns 2013 Nov;93(2):157-168. [doi: 10.1016/j.pec.2013.07.012] [Medline: 24001658]
16. Ekong G, Kavookjian J. Motivational interviewing and outcomes in adults with type 2 diabetes: A systematic review. Patient Educ Couns 2016 Jun;99(6):944-952. [doi: 10.1016/j.pec.2015.11.022] [Medline: 26699083]
17. Palacio A, Garay D, Langer B, Taylor J, Wood BA, Tamariz L. Motivational Interviewing Improves Medication Adherence: a Systematic Review and Meta-analysis. J Gen Intern Med 2016 Aug;31(8):929-940 [FREE Full text] [doi: 10.1007/s11606-016-3685-3] [Medline: 27160414]
18. Hill S, Kavookjian J. Motivational interviewing as a behavioral intervention to increase HAART adherence in patients who are HIV-positive: a systematic review of the literature. AIDS Care 2012;24(5):583-592. [doi: 10.1080/09540121.2011.630354] [Medline: 22292452]
19. Douaihy A, Amico KR. Motivational Interviewing in HIV Care. New York, NY: Oxford University Press; 2020.
20. Dillard PK, Zuniga JA, Holstad MM. An integrative review of the efficacy of motivational interviewing in HIV management. Patient Educ Couns 2017 Apr;100(4):636-646. [doi: 10.1016/j.pec.2016.10.029] [Medline: 27838113]
21. Kiene SM, Bateganya MH, Lule H, Wanyenze RK. The Effect of Motivational Interviewing-Based Counseling During Outpatient Provider Initiated HIV Testing on High-Risk Sexual Behavior in Rural Uganda. AIDS Behav 2016 Sep;20(9):1928-1936 [FREE Full text] [doi: 10.1007/s10461-016-1377-y] [Medline: 27037546]
22. Alemagno SA, Stephens RC, Stephens P, Shaffer-King P, White P. Brief motivational intervention to reduce HIV risk and to increase HIV testing among offenders under community supervision. J Correct Health Care 2009 Jul;15(3):210-221. [doi: 10.1177/1078345809333398] [Medline: 19477803]
23. Bendetson J, Dierst-Davies R, Flynn R, Beymer MR, Wohl AR, Amico KR, et al. Evaluation of a Client-Centered Linkage Intervention for Patients Newly Diagnosed with HIV at an Urban United States LGBT Center: The Linkage to Care Specialist Project. AIDS Patient Care STDS 2017 Jul;31(7):283-289. [doi: 10.1089/apc.2017.0057] [Medline: 28628349]
24. Okeke NL, Ostermann J, Thielman NM. Enhancing linkage and retention in HIV care: a review of interventions for highly resourced and resource-poor settings. Curr HIV/AIDS Rep 2014 Dec;11(4):376-392 [FREE Full text] [doi: 10.1007/s11904-014-0233-9] [Medline: 25323298]
25. Camacho-Gonzalez AF, Gillespie SE, Thomas-Seaton L, Frieson K, Hussen SA, Murray A, et al. The Metropolitan Atlanta community adolescent rapid testing initiative study: closing the gaps in HIV care among youth in Atlanta, Georgia, USA. AIDS 2017 Jul 01;31 Suppl 3(Suppl 3):S267-S275 [FREE Full text] [doi: 10.1097/QAD.0000000000001512] [Medline: 28665885]

XSL•FO
RenderX

26. Naar-King S, Outlaw AY, Sarr M, Parsons JT, Belzer M, Macdonell K, Adolescent Medicine Network for HIV/AIDS Interventions. Motivational Enhancement System for Adherence (MESA): pilot randomized trial of a brief computer-delivered prevention intervention for youth initiating antiretroviral treatment. J Pediatr Psychol 2013 Jul;38(6):638-648 [FREE Full text] [doi: 10.1093/jpepsy/jss132] [Medline: 23359664]
27. Goggin K, Gerkovich MM, Williams KB, Banderas JW, Catley D, Berkley-Patton J, et al. A randomized controlled trial examining the efficacy of motivational counseling with observed therapy for antiretroviral therapy adherence. AIDS Behav 2013 Jul;17(6):1992-2001 [FREE Full text] [doi: 10.1007/s10461-013-0467-3] [Medline: 23568228]
28. DiIorio C, McCarty F, Resnicow K, McDonnell Holstad M, Soet J, Yeager K, et al. Using motivational interviewing to promote adherence to antiretroviral medications: a randomized controlled study. AIDS Care 2008 Mar;20(3):273-283 [FREE Full text] [doi: 10.1080/09540120701593489] [Medline: 18351473]
29. Parsons JT, Golub SA, Rosof E, Holder C. Motivational interviewing and cognitive-behavioral intervention to improve HIV medication adherence among hazardous drinkers: a randomized controlled trial. J Acquir Immune Defic Syndr 2007 Dec 01;46(4):443-450 [FREE Full text] [doi: 10.1097/qai.0b013e318158a461] [Medline: 18077833]
30. Parsons JT, Rosof E, Punzalan JC, Di Maria L. Integration of motivational interviewing and cognitive behavioral therapy to improve HIV medication adherence and reduce substance use among HIV-positive men and women: results of a pilot project. AIDS Patient Care STDS 2005 Jan;19(1):31-39. [doi: 10.1089/apc.2005.19.31] [Medline: 15665633]
31. Naar-King S, Parsons JT, Murphy DA, Chen X, Harris DR, Belzer ME. Improving health outcomes for youth living with the human immunodeficiency virus: a multisite randomized trial of a motivational intervention targeting multiple risk behaviors. Arch Pediatr Adolesc Med 2009 Dec;163(12):1092-1098 [FREE Full text] [doi: 10.1001/archpediatrics.2009.212] [Medline: 19996045]
32. Bartlett J, Cheever L, Johnson M, Paauw D. A Guide to Primary Care of People With HIV/AIDS. Rockville, MD: US Department of Health and Human Services; 2004.
33. Magnus M, Jones K, Phillips G, Binson D, Hightow-Weidman LB, Richards-Clarke C, YMSM of color Special Projects of National Significance Initiative Study Group. Characteristics associated with retention among African American and Latino adolescent HIV-positive men: results from the outreach, care, and prevention to engage HIV-seropositive young MSM of color special project of national significance initiative. J Acquir Immune Defic Syndr 2010 Apr 01;53(4):529-536. [doi: 10.1097/QAI.0b013e3181b56404] [Medline: 19755914]
34. Substance use and dependence among HIV-infected adolescents and young adults. Information for Practice.: New York State Department of Health; 2010. URL: https://ifp.nyu.edu/2010/guidelines-plus/substance-use-and-dependence-among-hiv-infected-adolescents-and-young-adults/ [accessed 2018-12-31]
35. Community health workers. American Public Health Association. URL: https://www.apha.org/apha-communities/member-sections/communityhealth-workers [accessed 2020-04-23]
36. Giblin PT. Effective utilization and evaluation of indigenous health care workers. Public Health Rep 1989;104(4):361-368 [FREE Full text] [Medline: 2502807]
37. Rubak S, Sandbaek A, Lauritzen T, Christensen B. Motivational interviewing: a systematic review and meta-analysis. Br J Gen Pract 2005 Apr;55(513):305-312 [FREE Full text] [Medline: 15826439]
38. Mitcheson L, Bhavsar K, McCambridge J. Randomized trial of training and supervision in motivational interviewing with adolescent drug treatment practitioners. J Subst Abuse Treat 2009 Jul;37(1):73-78. [doi: 10.1016/j.jsat.2008.11.001] [Medline: 19150203]
39. Barwick MA, Bennett LM, Johnson SN, McGowan J, Moore JE. Training health and mental health professionals in motivational interviewing: A systematic review. Children and Youth Services Review 2012 Sep;34(9):1786-1795. [doi: 10.1016/j.childyouth.2012.05.012]
40. Smith JL, Carpenter KM, Amrhein PC, Brooks AC, Levin D, Schreiber EA, et al. Training substance abuse clinicians in motivational interviewing using live supervision via teleconferencing. J Consult Clin Psychol 2012 Jun;80(3):450-464 [FREE Full text] [doi: 10.1037/a0028176] [Medline: 22506795]
41. Darnell D, Dunn C, Atkins D, Ingraham L, Zatzick D. A Randomized Evaluation of Motivational Interviewing Training for Mandated Implementation of Alcohol Screening and Brief Intervention in Trauma Centers. J Subst Abuse Treat 2016 Jan;60:36-44 [FREE Full text] [doi: 10.1016/j.jsat.2015.05.010] [Medline: 26117081]
42. Hallgren KA, Dembe A, Pace BT, Imel ZE, Lee CM, Atkins DC. Variability in motivational interviewing adherence across sessions, providers, sites, and research contexts. J Subst Abuse Treat 2018 Jan;84:30-41 [FREE Full text] [doi: 10.1016/j.jsat.2017.10.011] [Medline: 29195591]
43. Schwalbe CS, Oh HY, Zweben A. Sustaining motivational interviewing: a meta-analysis of training studies. Addiction 2014 Aug;109(8):1287-1294. [doi: 10.1111/add.12558] [Medline: 24661345]
44. Moyers TB, Martin T, Houck JM, Christopher PJ, Tonigan JS. From in-session behaviors to drinking outcomes: a causal chain for motivational interviewing. J Consult Clin Psychol 2009 Dec;77(6):1113-1124 [FREE Full text] [doi: 10.1037/a0017189] [Medline: 19968387]
45. Vader AM, Walters ST, Prabhu GC, Houck JM, Field CA. The language of motivational interviewing and feedback: counselor language, client language, and client drinking outcomes. Psychol Addict Behav 2010 Jun;24(2):190-197 [FREE Full text] [doi: 10.1037/a0018749] [Medline: 20565145]

XSL•FO
RenderX

46. Apodaca TR, Longabaugh R. Mechanisms of change in motivational interviewing: a review and preliminary evaluation of the evidence. Addiction 2009 May;104(5):705-715 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02527.x] [Medline: 19413785]
47. Magill M, Apodaca TR, Barnett NP, Monti PM. The route to change: within-session predictors of change plan completion in a motivational interview. J Subst Abuse Treat 2010 Apr;38(3):299-305 [FREE Full text] [doi: 10.1016/j.jsat.2009.12.001] [Medline: 20149571]
48. Moyers TB, Martin T. Therapist influence on client language during motivational interviewing sessions. J Subst Abuse Treat 2006 Apr;30(3):245-251. [doi: 10.1016/j.jsat.2005.12.003] [Medline: 16616169]
49. Moyers TB, Manuel JK, Wilson PG, Hendrickson SML, Talcott W, Durand P. A Randomized Trial Investigating Training in Motivational Interviewing for Behavioral Health Providers. Behav. Cogn. Psychother 2007 Nov 22;36(2):149-162. [doi: 10.1017/s1352465807004055]
50. Tsai J, Klee A, Shea N, Lawless M, Payne KA, Goggin E, et al. Training peer specialists with mental illness in motivational interviewing: A pilot study. Psychiatr Rehabil J 2017 Dec;40(4):354-360. [doi: 10.1037/prj0000226] [Medline: 27786521]
51. Allicock M, Carr C, Johnson LS, Smith R, Lawrence M, Kaye L, et al. Implementing a one-on-one peer support program for cancer survivors using a motivational interviewing approach: results and lessons learned. J Cancer Educ 2014 Mar;29(1):91-98 [FREE Full text] [doi: 10.1007/s13187-013-0552-3] [Medline: 24078346]
52. Dewing S, Mathews C, Cloete A, Schaay N, Shah M, Simbayi L, et al. From research to practice: lay adherence counsellors' fidelity to an evidence-based intervention for promoting adherence to antiretroviral treatment in the Western cape, South Africa. AIDS Behav 2013 Nov;17(9):2935-2945. [doi: 10.1007/s10461-013-0509-x] [Medline: 23666183]
53. Madson MB, Loignon AC, Lane C. Training in motivational interviewing: a systematic review. J Subst Abuse Treat 2009 Jan;36(1):101-109. [doi: 10.1016/j.jsat.2008.05.005] [Medline: 18657936]
54. de Roten Y, Zimmermann G, Ortega D, Despland J. Meta-analysis of the effects of MI training on clinicians' behavior. J Subst Abuse Treat 2013 Aug;45(2):155-162. [doi: 10.1016/j.jsat.2013.02.006] [Medline: 23537923]
55. Miller WR, Yahne CE, Moyers TB, Martinez J, Pirritano M. A randomized trial of methods to help clinicians learn motivational interviewing. J Consult Clin Psychol 2004 Dec;72(6):1050-1062. [doi: 10.1037/0022-006X.72.6.1050] [Medline: 15612851]
56. Henshaw MM, Borrelli B, Gregorich SE, Heaton B, Tooley E, Santo W, et al. Randomized Trial of Motivational Interviewing to Prevent Early Childhood Caries in Public Housing. JDR Clin Trans Res 2018 Oct;3(4):353-365 [FREE Full text] [doi: 10.1177/2380084418794377] [Medline: 30238060]
57. Report of the consultation meeting on digital health interventions and health workforce capacity building. World Health Organization. 2019. URL: https://www.who.int/hrh/news/2019/meeting_report_Digital-health-interventions-health-workforce-capacity-building.pdf [accessed 2020-05-04]
58. Jackson CB, Quetsch LB, Brabson LA, Herschell AD. Web-Based Training Methods for Behavioral Health Providers: A Systematic Review. Adm Policy Ment Health 2018 Jul;45(4):587-610 [FREE Full text] [doi: 10.1007/s10488-018-0847-0] [Medline: 29352459]
59. Frank HE, Becker-Haimes EM, Kendall PC. Therapist training in evidence-based interventions for mental health: A systematic review of training approaches and outcomes. Clin Psychol (New York) 2020 Sep;27(3):e12330 [FREE Full text] [doi: 10.1111/cpsp.12330] [Medline: 34092941]
60. Calder R, Ainscough T, Kimergård A, Witton J, Dyer KR. Online training for substance misuse workers: a systematic review. Drugs Educ Prev Policy 2017;24(6):430-442. [doi: 10.1080/09687637.2017.1318113]
61. Shingleton RM, Palfai TP. Technology-delivered adaptations of motivational interviewing for health-related behaviors: a systematic review of the current research. Patient Educ Couns 2016;99(1):17-35 [FREE Full text] [doi: 10.1016/j.pec.2015.08.005] [Medline: 26298219]
62. Jiang S, Wu L, Gao X. Beyond face-to-face individual counseling: a systematic review on alternative modes of motivational interviewing in substance abuse treatment and prevention. Addict Behav 2017;73:216-235. [doi: 10.1016/j.addbeh.2017.05.023] [Medline: 28554033]
63. Shafer MS, Rhode R, Chong J. Using distance education to promote the transfer of motivational interviewing skills among behavioral health professionals. J Subst Abuse Treat 2004;26(2):141-148. [doi: 10.1016/s0740-5472(03)00167-3]
64. Mullin DJ, Saver B, Savageau JA, Forsberg L, Forsberg L. Evaluation of online and in-person motivational interviewing training for healthcare providers. Fam Syst Health 2016;34(4):357-366. [doi: 10.1037/fsh0000214] [Medline: 27571542]
65. Mitchell S, Heyden R, Heyden N, Schroy P, Andrew S, Sadikova E, et al. A pilot study of motivational interviewing training in a virtual world. J Med Internet Res 2011;13(3):e77. [doi: 10.2196/jmir.1825] [Medline: 21946183]
66. Radecki L, Goldman R, Baker A, Lindros J, Boucher J. Are pediatricians "Game"? Reducing childhood obesity by training clinicians to use motivational interviewing through role-play simulations with avatars. Games Health J 2013;2(3):174-178. [doi: 10.1089/g4h.2013.0020] [Medline: 24761323]
67. Czart M. Using 3D virtual standardized patients to teach motivational interviewing. J Virtual Worlds Res 2014;7(2). [doi: 10.4101/jvwr.v7i2.6411]

XSL•FO
RenderX

68. Oster C, Schoo A, Litt J, Morello A, Leibbrandt R, Antonello C, et al. Supporting workforce practice change: protocol for a pilot study of a motivational interviewing virtual client software tool for health professionals. BMJ Open 2020;10(2):e033080 [FREE Full text] [doi: 10.1136/bmjopen-2019-033080] [Medline: 32041854]
69. Talley R, Chiang IC, Covell NH, Dixon LB. Comparative initial and sustained engagement in web-based training by behavioral healthcare providers in New York State. J Technol Behav Sci 2018;3(2):41-48 [FREE Full text] [doi: 10.1007/s41347-017-0027-1] [Medline: 29732398]
70. Banna J, Grace Lin M, Stewart M, Fialkowski MK. Interaction matters: strategies to promote engaged learning in an online introductory nutrition course. J Online Learn Teach 2015;11(2):249-261 [FREE Full text] [Medline: 27441032]
71. Bernard RM, Abrami PC, Borokhovski E, Wade CA, Tamim RM, Surkes MA, et al. A meta-analysis of three types of interaction treatments in distance education. Rev Educ Res 2009;79(3):1243-1289. [doi: 10.3102/0034654309333844]
72. Conrad RM, Donaldson JA. Engaging the Online Learner: Activities and Resources for Creative Instruction. San Francisco, CA: John Wiley & Sons; 2011.
73. Lou Y, Bernard RM, Abrami PC. Media and pedagogy in undergraduate distance education: a theory-based meta-analysis of empirical literature. Educ Technol Res Dev 2006;54(2):141-176. [doi: 10.1007/s11423-006-8252-x]
74. Norris DM, Mason J, Lefrere P. Transforming E-Knowledge: A Revolution in the Sharing of Knowledge. Ann Arbor, MI: Society for College and University Planning; 2003.
75. Saleem AN, Noori NM, Ozdamli F. Gamification applications in E-learning: a literature review. Tech Know Learn 2021;27(1):139-159. [doi: 10.1007/s10758-020-09487-x]
76. Majuri J, Koivisto J, Hamari J. Gamification of education and learning: a review of empirical literature. CEUR Workshop Proc 2018;2186:11-19.
77. van Gaalen AEJ, Brouwer J, Schönrock-Adema J, Bouwkamp-Timmer T, Jaarsma ADC, Georgiadis JR. Gamification of health professions education: a systematic review. Adv Health Sci Educ Theory Pract 2021;26(2):683-711 [FREE Full text] [doi: 10.1007/s10459-020-10000-3] [Medline: 33128662]
78. Khaldi A, Bouzidi R, Nader F. Gamification of e-learning in higher education: a systematic literature review. Smart Learn Environ 2023;10(1):10 [FREE Full text] [doi: 10.1186/s40561-023-00227-z] [Medline: 40478066]
79. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining 'gamification'. : ACM; 2011 Presented at: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments; 2021 Sep 28; New York, NY p. 9-15. [doi: 10.1145/2181037.2181040]
80. Hamari J, Koivisto J, Sarsa H. Does gamification work? - a literature review of empirical studies on gamification. : IEEE; 2014 Presented at: 2014 47th Hawaii International Conference on System Sciences; 2014 Mar 10; Washington, DC p. 3025-3034. [doi: 10.1109/hicss.2014.377]
81. Brull S, Finlayson S. Importance of gamification in increasing learning. J Contin Educ Nurs 2016;47(8):372-375. [doi: 10.3928/00220124-20160715-09] [Medline: 27467313]
82. Larson K. Serious games and gamification in the corporate training environment: a literature review. TechTrends 2019;64(2):319-328. [doi: 10.1007/s11528-019-00446-7]
83. Gentry S, Ehrstrom BLE, Gauthier A, Alvarez J, Wortley D, van RJ. Serious gaming and gamification interventions for health professional education. Cochrane Database Syst Rev 2016;6:CD012209. [doi: 10.1002/14651858.cd012209]
84. da Rocha Seixas L, Gomes AS, de Melo Filho IJ. Effectiveness of gamification in the engagement of students. Comput Hum Behav 2016;58:48-63. [doi: 10.1016/j.chb.2015.11.021]
85. Lank E, Withee K, Schile L, Parker T. User centred rapid application development. In: Rapid Integration of Software Engineering Techniques. Berlin, Germany: Springer; 2006:34-49.
86. Kumar S, Sharma T. User-centric rapid application development. Int J Res 2014;1(10):959-964. [doi: 10.1007/11751113_4]
87. Martin J. Rapid Application Development. New York, NY: Macmillan Publishing Co; 1991.
88. Arsand E, Demiris G. User-centered methods for designing patient-centric self-help tools. Inform Health Soc Care 2008;33(3):158-169. [doi: 10.1080/17538150802457562] [Medline: 18850399]
89. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
90. LLC. Simmersion. URL: https://www.simmersion.com/mi [accessed 2025-11-25]
91. Taylor B, Henshall C, Kenyon S, Litchfield I, Greenfield S. Can rapid approaches to qualitative analysis deliver timely, valid findings to clinical leaders? A mixed methods study comparing rapid and thematic analysis. BMJ Open 2018;8(10):e019993 [FREE Full text] [doi: 10.1136/bmjopen-2017-019993] [Medline: 30297341]
92. Gale RC, Wu J, Erhardt T, Bounthavong M, Reardon CM, Damschroder LJ, et al. Comparison of rapid vs in-depth qualitative analytic methods from a process evaluation of academic detailing in the veterans health administration. Implement Sci 2019;14(1):11 [FREE Full text] [doi: 10.1186/s13012-019-0853-y] [Medline: 30709368]
93. Al-Khanbashi M. Using the matrix as a qualitative data display for landscape research and a reflection based on the social constructivist perspective. In: Modern Approaches to the Visualization of Landscapes. Wiesbaden, Germany: Springer VS; 2020:103-118.

XSL•FO
RenderX

94. Miles MB, Huberman AM, Saldaña J. Qualitative data analysis: a methods sourcebook. Thousand Oaks, CA: Sage Publications; 2018.
95. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017;17(1):88 [FREE Full text] [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
96. Murphy AL, Gardner DM. Pharmacists' acceptability of a men's mental health promotion program using the theoretical framework of acceptability. AIMS Public Health 2019;6(2):195-208 [FREE Full text] [doi: 10.3934/publichealth.2019.2.195] [Medline: 31297404]
97. Brook J, Aitken LM, MacLaren JA, Salmon D. An intervention to decrease burnout and increase retention of early career nurses: a mixed methods study of acceptability and feasibility. BMC Nurs 2021;20(1):19 [FREE Full text] [doi: 10.1186/s12912-020-00524-9] [Medline: 33435972]
98. Chakrapani V, Shunmugam M, Rawat S, Baruah D, Nelson R, Newman PA. Acceptability of HIV pre-exposure prophylaxis among transgender women in India: a qualitative investigation. AIDS Patient Care STDS 2020;34(2):92-98 [FREE Full text] [doi: 10.1089/apc.2019.0237] [Medline: 31951490]
99. Msosa TC, Phuka F, Sumari-de Boer M, Mhango O, Twabi HH, Mukoka M, et al. The acceptability of a real-time medication monitoring-based digital adherence tool among young people living with HIV in Malawi. AIDS Behav 2025. [doi: 10.1007/s10461-025-04885-7] [Medline: 40983785]
100. Sekhon M, Cartwright M, Francis JJ. Development of a theory-informed questionnaire to assess the acceptability of healthcare interventions. BMC Health Serv Res 2022;22(1):279 [FREE Full text] [doi: 10.1186/s12913-022-07577-3] [Medline: 35232455]
101. Leffingwell TR. Motivational interviewing knowledge and attitudes test (MIKAT) for evaluation of training outcomes. MINUET 2006;13:10-11. [doi: 10.1002/9781119057840.ch160]
102. Doran N, Hohman M, Koutsenok I. Motivational interviewing training in juvenile corrections: a comparison of outside experts and internal trainers. Legal Criminol Psychol 2011;18(2):262-273. [doi: 10.1111/j.2044-8333.2011.02036.x]
103. Hohman M, Doran N, Koutsenok I. Motivational interviewing training for juvenile correctional staff in California: one year initial outcomes. J Offender Rehabil 2009;48(7):635-648. [doi: 10.1080/10509670903196108]
104. Madson MB, Schumacher JA, Noble JJ, Bonnell MA. Teaching motivational interviewing to undergraduates. Teach Psychol 2013;40(3):242-245. [doi: 10.1177/0098628313487450]
105. Aarons GA. Mental health provider attitudes toward adoption of evidence-based practice: the evidence-based practice attitude scale (EBPAS). Ment Health Serv Res 2004;6(2):61-74. [doi: 10.1023/B:MHSR.0000024351.12294.65]
106. Ehrhart MG, Aarons GA, Farahnak LR. Assessing the organizational context for EBP implementation: the development and validity testing of the implementation climate scale (ICS). Implement Sci 2014;9:157 [FREE Full text] [doi: 10.1186/s13012-014-0157-1] [Medline: 25338781]
107. Aarons GA, Ehrhart MG, Farahnak LR. The implementation leadership scale (ILS): development of a brief measure of unit level implementation leadership. Implement Sci 2014;9(1):45 [FREE Full text] [doi: 10.1186/1748-5908-9-45] [Medline: 24731295]
108. Moyers TB, Manuel JK, Ernst D. Motivational Interviewing Treatment Integrity Coding Manual 4.2.1. Albuquerque, NM: University of New Mexico; 2014.
109. Moyers TB, Rowell LN, Manuel JK, Ernst D, Houck JM. The motivational interviewing treatment integrity code (MITI 4): rationale, preliminary reliability and validity. J Subst Abuse Treat 2016;65:36-42 [FREE Full text] [doi: 10.1016/j.jsat.2016.01.001] [Medline: 26874558]
110. Patton MQ. Qualitative Research & Evaluation Methods. Thousand Oaks, CA: Sage Publications; 2014.
111. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. J Psychiatr Res 2011;45(5):626-629 [FREE Full text] [doi: 10.1016/j.jpsychires.2010.10.008] [Medline: 21035130]

Abbreviations

AETC: AIDS Education and Training Center
CAB: Community Advisory Board
CASI: computer-assisted self-interview
CHW: community health worker
CSS: counseling support staff
EBP: evidence-based practice
FG: focus group
FSU: Florida State University
HCC: HIV Care Continuum
HIPAA: Health Insurance Portability and Accountability Act
iMI4HIV: internet-based motivational interviewing training program for HIV counseling support staff
MI: motivational interviewing

XSL•FO
RenderX

MIKAT: Motivational Interviewing Knowledge and Attitudes Test
MITI: Motivational Interviewing Treatment Integrity
NECA AETC: Northeast and Caribbean AIDS Education and Training Center
NIH: National Institutes of Health
QI: qualitative interview
RAD: Rapid Application Development
RCT: randomized controlled trial
SPI: Standard Patient Interaction
TFA: Theoretical Framework of Acceptability
UCRAD: User Centered Rapid Application Development
VMI: virtual motivational interviewing

Edited by J Sarvestan; submitted 13.Aug.2025; this is a non–peer-reviewed article;accepted 19.Nov.2025; published 07.Jan.2026.

Please cite as:
Balán IC, Brewer OR, Kutner BA, Giguere R
Interactive Internet-Based Motivational Interviewing Training for HIV Counseling Support Staff to Improve Health Communication in HIV Care Interactions: Protocol for Training Development and a Pilot Randomized Controlled Trial
JMIR Res Protoc 2026;15:e82241
URL: https://www.researchprotocols.org/2026/1/e82241
doi:10.2196/82241
PMID:

©Iván C Balán, Onna R Brewer, Bryan A Kutner, Rebecca Giguere. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 07.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Developing a Multimodal Screening Algorithm for Mild Cognitive Impairment and Early Dementia in Home Health Care: Protocol for a Cross-Sectional Case-Control Study Using Speech Analysis, Large Language Models, and Electronic Health Records

Maryam Zolnoori[1], PhD

Columbia University Irving Medical Center, New York, NY, United States

Corresponding Author:
Maryam Zolnoori, PhD
Columbia University Irving Medical Center
560 W 168th St
New York, NY, 10032
United States
Phone: 1 317 515 1950
Email: m.zolnoori@gmail.com

Abstract

Background: Mild cognitive impairment and early dementia (MCI-ED) are frequently unrecognized in routine care, particularly in home health care (HHC), where clinical decisions are made under time constraints and cognitive status may be incompletely documented. Federally mandated HHC assessments, such as the Outcome and Assessment Information Set (OASIS), capture health and functional status but may miss subtle early cognitive changes. Speech, language, and interactional patterns during routine patient-nurse communication, together with information embedded in unstructured clinical notes, may provide complementary signals for earlier identification.

Objective: This protocol describes the development and evaluation of a multimodal screening approach for identifying MCI-ED in HHC by integrating (1) speech and interaction features from routine patient-nurse encounters (verbal communication), (2) large language model–based extraction of MCI-ED–related information from HHC notes and encounter transcripts, and (3) structured variables from OASIS.

Methods: This ongoing cross-sectional case-control study is being conducted in collaboration with VNS Health (formerly Visiting Nurse Service of New York). Eligible participants are adults aged ≥60 years receiving HHC services. Case/control assignment uses a 2-stage process: electronic health record (EHR) prescreening followed by clinician-reviewed cognitive assessment (Montreal Cognitive Assessment and Clinical Dementia Rating) for consented participants without an existing mild cognitive impairment diagnosis. For Aim 1, each participant contributes 3 audio-recorded routine patient-nurse encounters linked to EHR data, including OASIS and free-text clinical notes. Aim 1 extracts acoustic, linguistic, emotional, and interactional features from patient-nurse verbal communication. Aim 2 uses a schema-guided large language model pipeline to extract and normalize MCI-ED–related symptoms, lifestyle risk factors, and communication deficits from HHC notes and encounter transcripts, supported by a human-annotated gold-standard dataset. Aim 3 integrates speech, extracted text variables, and OASIS predictors using supervised machine learning with stratified nested cross-validation; evaluation will include discrimination, calibration, and subgroup performance checks across race, sex, and age.

Results: Between February 2024 and July 2025, a total of 114 HHC patients completed study-administered cognitive assessments and were classified as 55 MCI-ED cases and 59 cognitively normal controls. Audio-recorded patient-nurse encounters had a median duration of 19 (IQR 12-23) minutes and a median of 56 (IQR 31-80) utterances per encounter; nurses contributed more words than patients (median 842, IQR 461-1218 vs median 589, IQR 303-960). In exploratory feasibility analyses, multimodal models integrating speech, interactional features, and structured EHR/OASIS variables outperformed single-source models.

Conclusions: This protocol describes a reproducible multimodal framework for MCI-ED screening in HHC using routinely generated data streams. Initial implementation results support feasibility of data collection and end-to-end processing and suggest potential value of integrating interactional speech features with clinical text and OASIS variables. Final model evaluation, subgroup analyses, and validation will follow the prespecified analytic procedures on the finalized study dataset.

International Registered Report Identifier (IRRID): DERR1-10.2196/82731

XSL•FO
RenderX

(JMIR Res Protoc 2026;15:e82731) doi:10.2196/82731

KEYWORDS

dementia; fairness; home health care; large language models; mild cognitive impairment; multimodal fusion; natural language processing; screening; speech analysis

Introduction

Alzheimer disease and Alzheimer disease–related dementias are among the most pressing global public health challenges. In 2021, an estimated 57 million people were living with dementia worldwide, with over 60% residing in low- and middle-income countries and nearly 10 million new cases occurring each year [1]. Dementia is associated with substantial disability, caregiver burden, and rapidly rising health-system costs [2]. In the United States, an estimated 7.2 million adults aged ≥60 years were living with Alzheimer dementia in 2025, underscoring the scale of need in high-income settings as well [3]. Alongside treatment advances and growing interest in risk reduction, global guidance continues to emphasize the importance of timely detection and prevention-oriented interventions across diverse populations and care contexts [4,5].

Despite this urgency, a large proportion of cognitive impairment, particularly mild cognitive impairment (MCI) and early-stage dementia, remains unrecognized or undocumented in routine care, contributing to delayed diagnosis and missed opportunities to tailor care planning [6,7]. In home-based care settings, documentation gaps can be especially consequential because care teams must make clinical decisions in the context of multimorbidity, limited visit time, and incomplete prior cognitive history [8,9]. Recent evidence from skilled home health care (HHC) shows that dementia is frequently undocumented in home health records, illustrating how care transitions and documentation practices can impede recognition of cognitive impairment [10]. These realities motivate scalable screening approaches that can operate within routine workflows, rather than relying solely on specialist evaluation or resource-intensive testing.

Speech and language have emerged as promising noninvasive, low-burden digital biomarkers for cognitive impairment [11-14]. A recent systematic review and meta-analysis focused on MCI specifically concluded that speech-based biomarkers show meaningful diagnostic use, while also highlighting methodological heterogeneity and the need for validation in diverse settings and populations [15]. At the same time, much of the speech-based Alzheimer disease and Alzheimer disease–related dementias detection literature remains anchored in structured elicitation tasks [16] (eg, cookie-theft picture description) and benchmark corpora (eg, DementiaBank-derived shared tasks such as ADReSSo [17]), which enable comparability but may not capture interactional and pragmatic markers expressed during everyday clinical communication [15,18]. Multilingual and cross-cultural work further indicates that generalization across languages and contexts cannot be assumed; for example, multilingual spontaneous speech studies (eg, Italian and Spanish) demonstrate feasibility outside English-centric benchmarks but also reinforce the importance of ecologically valid sampling and external validation [13,19-21].

These limitations are particularly relevant for MCI and early dementia, where impairments can be subtle, context-dependent, and potentially expressed through conversational dynamics (eg, timing, turn-taking balance, and discourse coherence) rather than only through content produced during structured tasks [16]. This motivates studying routine patient-clinician conversations, where interactional features may provide an additional signal for early-stage cognitive change in real-world contexts [22].

A complementary and underused source of early cognitive signals is unstructured clinical documentation, including HHC nursing notes [10,23]. Recent reviews show that natural language processing (NLP) approaches applied to electronic health record (EHR) notes can identify cognitive impairment with strong median performance across studies, but variability in diagnostic criteria, data sources, and external validation remains a key barrier to translation [24]. In parallel, large language model (LLM) methods are increasingly being evaluated for detecting cognitive decline from clinical notes, including large clinical language model approaches (eg, CD-Tron) and comparative studies of LLMs in real-world clinical text [25-29]. These developments suggest that LLM-enabled extraction can help capture both explicit and implicit mentions of symptoms, risk factors, and functional concerns that are inconsistently represented in structured fields, an especially relevant issue in home-based care workflows.

HHC is therefore a compelling setting for scalable, equity-oriented screening because it provides repeated encounters and routinely generates multiple complementary data streams, including standardized assessments (eg, Outcome and Assessment Information Set [OASIS] in US Medicare–certified home health agencies), narrative nursing notes, and patient-nurse verbal communication. Our prior work in HHC has shown that combining structured assessment data with information extracted from clinical notes can improve risk identification (HomeADScreen [12]). More recently, we demonstrated the potential value of leveraging audio-recorded patient-nurse verbal communication as an additional signal beyond EHR data for early cognitive screening in HHC [16]. However, few studies have jointly leveraged (1) standardized home-care assessments, (2) unstructured home-care clinical notes, and (3) routine patient-clinician conversations within a single integrated screening framework for mild cognitive impairment and early dementia (MCI-ED) in home-based care.

Accordingly, this study describes an ongoing protocol to develop and evaluate a multimodal screening approach for identifying MCI and early dementia in HHC using routinely generated data streams: standardized assessment data (OASIS), HHC nursing notes, and audio-recorded patient-nurse verbal communication. We aim to (1) model speech, language, emotion, and interaction

XSL•FO
RenderX

patterns from patient-nurse conversations using automated speech analysis, (2) apply NLP/LLM methods to identify MCI/early dementia–related symptoms, lifestyle risk factors, and communication deficits from both clinical notes and verbal communication, and (3) integrate these signals with standardized assessment variables to improve screening performance compared with models based on any single data stream.

Methods

Study Setting, Design, and Status

This protocol describes an ongoing cross-sectional case-control protocol in collaboration with VNS Health (formerly Visiting Nurse Service of New York), one of the largest HHC systems in the United States. The study population includes adults aged 60 years and older who receive HHC services from VNS Health. The protocol is designed to develop and evaluate a multimodal screening algorithm for identifying MCI-ED in HHC.

Participant Recruitment, Eligibility, and Group Allocation

Recruitment Focus and Rationale

Aim 1 focuses on modeling speech, language, emotional expression, and interaction patterns during routine patient-nurse encounters (verbal communication) as markers of early cognitive decline in HHC. The primary analytic cohort includes non-Hispanic Black and non-Hispanic White patients receiving HHC services from VNS Health. These groups were selected because they are highly represented in the study setting, enable adequately powered comparisons within a single HHC system, and facilitate evaluation of model performance across racial groups—particularly important given well-documented disparities in dementia diagnosis and care for Black patients.

Recruitment Strategy

Potential participants are identified through EHR-based screening and clinician referral workflows within VNS Health. Prespecified EHR indicators are used to identify likely cases (eg, documented symptoms of cognitive decline) and likely controls (no evidence of impairment). Eligible patients are approached during an active episode of HHC. Recruitment is monitored to achieve representation of both racial groups and, when feasible, balance across key characteristics, age, sex as a biological variable, and education.

Eligibility, Screening Workflow, and Group Allocation

Eligible participants are aged ≥60 years, plan to receive VNS Health services during the study period, have sufficient English proficiency to communicate independently with HHC nurses, have adequate vision/hearing to complete cognitive testing, and can provide written informed consent. Patients are excluded if they (1) are unable to communicate independently with the HHC nurse in English and (2) have speech or language disorders due to neurological conditions other than MCI-ED (eg, Parkinson disease or seizure disorders). Full eligibility criteria are provided in Multimedia Appendix 1.

We use a 2-stage approach to identify patients for case and control groups. First, we identify potential cases using available *ICD-10* (*International Statistical Classification of Diseases, Tenth Revision*) diagnoses in the EHR (*ICD-10* G31.84 for MCI) and identify potential controls as patients without documented cognitive impairment. Second, all consented participants without an existing MCI diagnosis complete cognitive assessments—the Montreal Cognitive Assessment (MoCA) [30,31] and Clinical Dementia Rating (CDR) [31]—in their homes, administered by a trained research assistant who audio-records responses to support final group assignment.

A study clinician with expertise in cognitive impairment detection reviews the recorded cognitive assessments together with relevant clinical context (medical history and nurse assessment information from OASIS) to confirm group assignment. Based on prespecified criteria, participants are classified as MCI-ED cases when findings are consistent with early cognitive impairment (anticipated CDR 0.5-1 and MoCA ~16-25, with consideration of EHR evidence when available) and as cognitively normal controls when findings are within normal limits (CDR 0 and MoCA ≥26), and there is no EHR evidence of cognitive impairment. Participants meeting criteria for moderate to severe impairment (eg, CDR 2-3 or MoCA <16) are excluded because the protocol focuses on MCI-ED.

After group allocation, patients in both the case and control groups are invited to provide additional consent for the next phase of Aim 1, which includes audio-recording routine patient-nurse encounters.

Ethical Considerations

This study was reviewed and approved as human participant research by the Columbia University Irving Medical Center Institutional Review Board (Protocol AAAU3168). The study is conducted in collaboration with VNS Health and complies with all applicable institutional, federal, and regulatory requirements for research involving human participants.

Written informed consent is obtained from all participating patients prior to enrollment. Consent includes permission for the administration of cognitive assessments, including the MoCA and the CDR, audio-recording of patient-nurse encounters, and linkage of audio recordings and assessment data with EHR information. HHC nurses also provide informed consent for participation, including consent for audio-recording of patient-nurse encounters. Participants are informed of the study purpose, procedures, potential risks, and their right to withdraw at any time without affecting their care or employment.

All study data are handled in accordance with Health Insurance Portability and Accountability Act and institutional data protection policies. Audio recordings, transcripts, cognitive assessment data, and clinical text are deidentified prior to analysis, with direct identifiers removed. Data are stored on secure, access-controlled servers at Columbia University and VNS Health with role-based permissions and audit logging. Access to identifiable data is restricted to authorized study personnel only. Deidentified datasets are used for analysis, and results are reported in aggregate to minimize the risk of participant reidentification.

XSL•FO
RenderX

Each participating patient receives a US $50 incentive for completion of cognitive assessments (MoCA and CDR) and an additional US $50 incentive for participation in audio-recording of patient-nurse encounters. HHC nurses also receive a US $50 incentive for participation in audio-recording of patient-nurse encounters. Incentives are provided in accordance with institutional review board–approved procedures and are not contingent on study outcomes.

Data Collection Overview

Number of Audio-Recorded Encounters (Patient-Nurse Conversation)

For each enrolled participant, 3 routine patient-nurse encounters are audio-recorded during the HHC episode of care. Recording multiple encounters provides repeated observations to capture within-person variability in speech, language, and interactional patterns across visits, while minimizing participant and clinician burden. When both patient and nurse provide consent, a trained research assistant attends the visit and operates a Saramonic Blink audio-recording device [32], minimizing burden on clinical staff. This portable device, with dual wireless microphones that attach to clothing, provides clear speech transmission to devices like an iPod and offers dual-channel storage.

Linking Audio-Recorded Encounters to EHR Data and Clinical Notes

Audio-recorded encounters are linked to EHR data extracted from the VNS Health system, including the OASIS [33,34]—a federally mandated HHC assessment capturing patient health status, functional status, and living arrangements—as well as supplemental structured data (eg, medications) and free-text clinical notes. Free-text notes include visit notes, documenting each nurse encounter, and care coordination notes, capturing communications with other clinicians, physicians, and family members.

Preliminary Feasibility and Pilot Studies

Prior to this protocol, we conducted a series of pilot studies to establish the feasibility of audio-recording patient-nurse verbal communication and applying automated speech and machine learning methods in the HHC setting [22]. First, we evaluated several commercially available audio-recording devices in laboratory and real-world HHC settings, assessing usability, transcription quality, and acceptability among HHC nurses and patients. Based on System Usability Scale scores and transcription accuracy measured by word error rate, the Saramonic Blink device [35] demonstrated the best overall performance and was selected for use in the current study. Semistructured interviews with HHC nurses and patients further indicated that audio-recording was acceptable and had minimal perceived impact on routine care delivery.

In a second pilot study, we demonstrated the feasibility of automated speaker type identification in recorded HHC encounters using machine learning models trained on acoustic and lexical features, achieving satisfactory classification performance [36]. In a third pilot study, we developed and validated an end-to-end analytic pipeline for modeling spoken language in cognitive impairment [11] (Figure 1). The pipeline included (1) audio preprocessing for noise reduction; (2) automated speaker type identification to separate patient and clinician speech; (3) extraction of acoustic features capturing phonetic motor planning and voice characteristics (eg, fluency, frequency/spectral measures, intensity, and instability) using OpenSMILE [37] and PRAAT [38]; (4) modeling of emotional expression using the Geneva Minimalistic Acoustic Parameter Set [39] (GeMAPS) complemented by lexicon-based psycholinguistic markers [40] (using Linguistic Inquiry and Word Count [LIWC]); (5) modeling of language organization using transcript-derived lexical and syntactic measures—using Natural Language Toolkit (NLTK) [41]—and contextual language representations (using distilled RoBERTa [42]); and (6) machine learning–based classification with internal validation. We evaluated this pipeline on a benchmark dataset (DementiaBank [43] “Cookie Theft” picture descriptions) and observed strong discrimination between cognitively impaired and cognitively unimpaired participants, supporting the feasibility of extracting informative speech-derived markers and training predictive models. Collectively, these pilot studies informed the design decisions, data collection procedures, and analytic pipelines in this protocol.

XSL•FO
RenderX

Figure 1. Analytic pipeline for modeling spoken language. ADRD: Alzheimer disease and related dementias; LIWC: Linguistic Inquiry and Word Count; NLTK: Natural Language Toolkit.

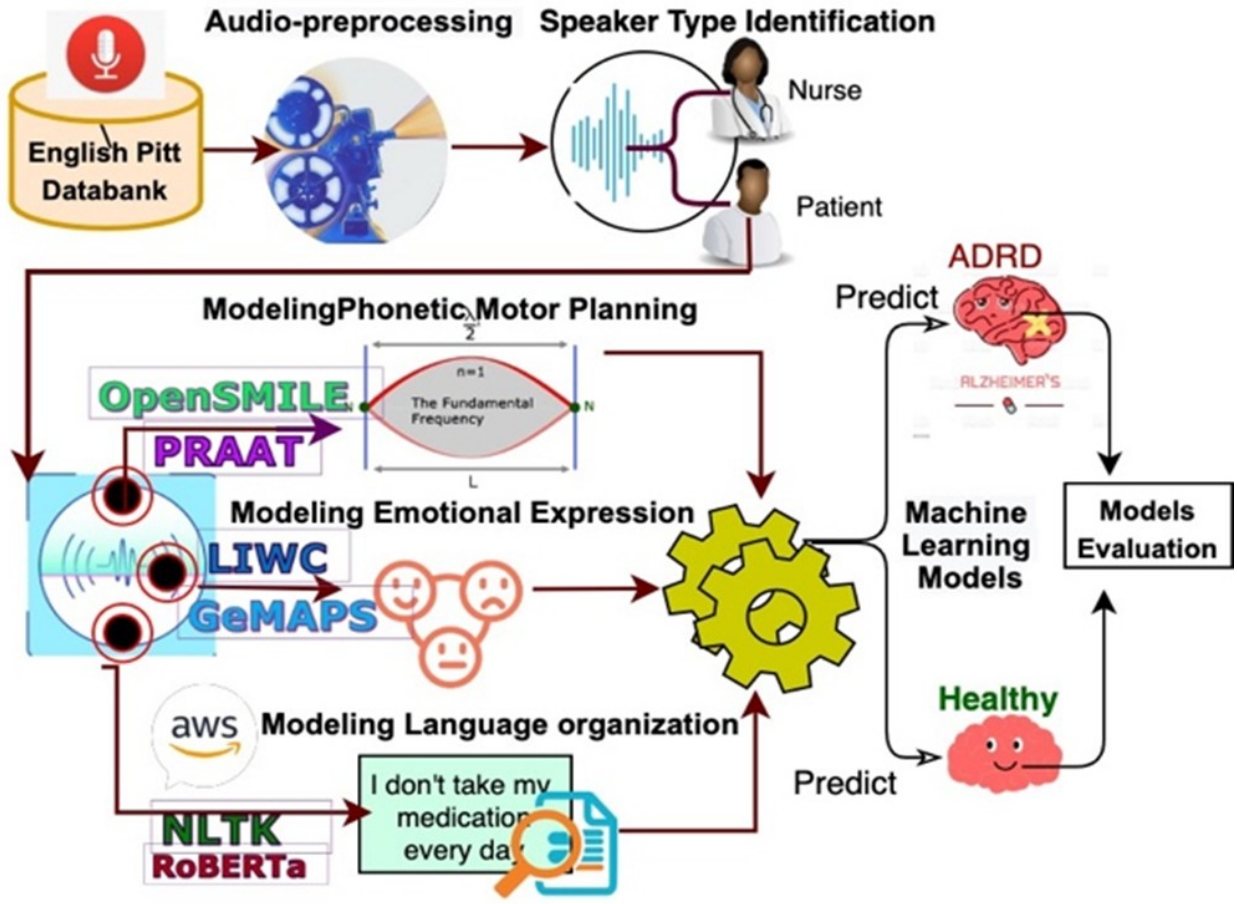

Analytic Method for Aim 1: Model MCI-ED Patients' Verbal Communications With Clinicians Using an Automated Speech Analysis System

Rationale and Overview

Early cognitive decline affects multiple aspects of spoken communication, including speech motor control, language organization, emotional expression, and social interaction. In HHC settings, these changes are expressed during spontaneous patient-nurse verbal communications rather than structured speech production tasks (eg, reading task). Building on our preliminary feasibility and pilot work, Aim 1 focuses on systematically modeling these communication patterns using an automated speech analysis system. The objective is to extract complementary acoustic, linguistic, emotional, and interactional features from naturally occurring patient-nurse communications that may signal MCI-ED. The analytic framework for Aim 1 consists of 5 components, summarized in Table 1, which together capture core dimensions of speech production and interaction relevant to cognitive decline.

XSL•FO
RenderX

Table 1. Modeling mild cognitive impairment and early dementia (MCI-ED) patient-nurse verbal communication in the home health care setting (components 1-4).

Component and domain	Measures
Component 1: modeling phonetic motor planning	
Speech (vocal) fluency	• Articulation: number of phonemes per second without hesitation [44]. • Speech rate: number of phonemes per second with hesitation [44]. • Silent pauses: number of speechless intervals at the beginning of and between words [45]. • Within-word disfluency: within-word silent pauses and sound prolongations [46].
Rhythmic structure of speech	• Syllabic intervals: temporal variability in speech [47]. • Pairwise variability index: durational variability in successive acoustic-phonetic intervals [48]. • Vowel duration: proportion of time of vocalic intervals in a sentence and the standard deviation of inter-vowel intervals [49].
Frequency and spectral domain	• Fundamental frequency: average number of oscillations originating from the vocal folds per second [50]. • Formant frequencies (F1-F4): acoustic resonances of the vocal tract due to changes in the positions of vocal organs [51]. • Spectral center of gravity: amplitude-weighted mean of harmonic peaks averaged over sound duration [52]. • Long-term average spectrum: composite signal representing the spectrum of the glottal source and resonant characteristics of the vocal tract [53]. • Mel-frequency cepstral coefficients: energy variations between frequency bands of a speech signal [54].
Voice instability	• Jitter: cycle-to-cycle period variation of successive glottal cycles [55]. • Shimmer: cycle-to-cycle amplitude variation of successive glottal cycles [55]. • Cepstral peak prominence: measure of periodicity in the speech signal [56].
Voice quality	• Harmonics-to-noise ratio: relative amount of additive noise in the voice signal [50]. • Voice breaks: reduced ability in vocal cord execution resulting in voice breaks [57]. • Acoustic voice quality index: weighted combination of time-frequency and quefrency-domain metrics developed to measure the severity of dysphonia [58].
Voice intensity	• Hammarberg index: articulatory effort computed as the difference between maximum energy in the 0-2 kHz band and the energy in the 2-5 kHz band [59]. • Energy concentration: average spectral frequency [45].
Component 2: modeling the patient's emotional expression	
Frequency parameters	• Pitch: number of vibrations per second produced by the vocal cords [60]. • Jitter [55]. • Center frequency of formants 1-3. • Bandwidth of formants 1-3. Formant frequencies are acoustic resonances of the vocal tract caused by changes in vocal organ positions [51].
Energy/amplitude	• Shimmer. • Loudness: estimate of perceived signal intensity from an auditory spectrum [61]. • Harmonics-to-noise ratio: relative amount of additive noise in the voice signal [50].
Spectral parameters	• Alpha ratio: ratio of summed energy from 50-1000 Hz and 1-5 kHz. • Hammarberg index: measure of articulatory effort [59]. • Spectral slope (0-500 Hz and 500-1500 Hz). • Formant 1-3 relative energy. • Harmonic difference H1-H2: difference between first and second harmonic amplitudes [62]. • Harmonic difference H1-A3: difference between H1 and A3 (energy of the highest harmonic in the third formant range) [62]. • Spectral flux: difference between the spectra of 2 consecutive frames. • Mel-frequency cepstral coefficients: see frequency and spectral domain in component 1.
Component 3: modeling syntactic, semantic, and pragmatic levels of language organization	
Lexical richness	• Moving average type-token ratio: total number of unique words divided by the total number of words for each successive fixed-length window [63]. • Brunet index: variation in word types marked by part-of-speech tagging relative to the total number of words in a sentence [64]. • Honore index: proportion of words used only once relative to the total number of words [65].

XSL•FO
RenderX

Component and domain	Measures
Syntactic level of language organization	• Sentence complexity: score computed using a syntactic parse tree [66]. • Grammatical errors: identified using a parse tree analyzer [67]. • Incomplete (fragment) sentences: identified using an automatic detection algorithm based on syntactic parse trees and part-of-speech tagging [68].
Semantic fluency	• Identification of filled pauses (eg, “um”) in the patient’s spoken language [69-72].
Patient recall ability	• Uncertainty in patient language: computed using the linguistic approximator introduced by Ferson et al [73]. • Memory-related terms: proportion of sentences containing memory-related terms relative to the total number of sentences, computed using the NimbleMiner toolkit [74,75]. • Question ratio: proportion of interrogative sentences relative to the total number of sentences, identified using the NLTK[a] Python package [74].
Component 4: modeling patient-nurse interaction	
Patient turns	• Continuous block of uninterrupted speech by a single patient. • Total number of patient turns indicates frequency of information exchange [76].
Interactivity	• Dialog interactivity: defined as the total number of patient turns divided by the total length of the encounter [76].
Turn density	• Computed using the same parameters specified for lexical richness (component 3).
Turn duration	• Length of time of the patient’s turn; longer durations have been associated with difficulty in turn monitoring in MCI-ED [76].
Relative timing of turns	• Discernible pause rate: proportion of discernible speechless intervals at the start of patient turns relative to total utterances [77]. • Cross-over speaking rate: proportion of patient-nurse utterances with cross-over speaking relative to total utterances during the interaction [77].

[a]NLTK: Natural Language Toolkit.

Feature Specification and Reproducibility

For reproducibility, all speech- and interaction-based parameters extracted in Aim 1 are explicitly specified in Table 1, organized by analytic component (phonetic motor planning, emotional expression, syntactic and semantic language organization, and patient-nurse interaction). Table 1 provides the operational definition and measurement domain for each parameter. Acoustic features are computed using established toolkits (OpenSMILE [78] and PRAAT [38]); linguistic features are derived from automatically transcribed speech using NLTK, LIWC [79], and distilled RoBERTa; and interactional features are computed from speaker-labeled timestamps generated by Amazon Web Services (AWS) Transcribe.

Component 1: Modeling Phonetic Motor Planning

Impairment in phonetic motor planning is a well-documented consequence of neurodegenerative disorders, including MCI-ED, and manifests as reduced articulation precision, altered speech rhythm [44,80,81], and increased disfluency. To characterize these changes, we analyze acoustic parameters across six domains (Table 1, component 1): (1) speech fluency [45,46,82], (2) rhythmic structure [48,49,83], (3) frequency and spectral characteristics [45,84,85], (4) voice instability [45,69,86], (5) voice quality [87-89], and (6) voice intensity [45,90]. These measures quantify temporal and spectral aspects of speech that reflect the patient’s ability to plan and execute vocal motor actions.

Component 2: Modeling Emotional Expression

Alterations in emotional expression often develop alongside cognitive decline and can negatively affect communication quality and interpersonal interaction [91,92]. Emotion is conveyed both through nonverbal vocalization and semantic content [93-96]. To model vocal expression of emotion, we use the GeMAPS [39], which captures affect-related changes in autonomic arousal and vocal musculature via frequency-, energy-, and spectral-domain parameters (Table 1, component 2). To capture the semantic expression of emotion, we extract linguistically encoded emotional indicators using the LIWC dictionary. Emotion-related linguistic markers (eg, sadness and anxiety) have been associated with cognitive dysfunction [97] and adverse health outcomes [98].

Component 3: Modeling Syntactic, Semantic, and Pragmatic Language Organization

Language impairment in MCI-ED is characterized by reduced lexical diversity, simplified syntax [44,97,99], word-finding difficulties [44,100], and impaired memory-related discourse [44,101], which together contribute to reduced coherence. We model language organization using features in four domains (Table 1, component 3): (1) lexical richness [64,65,102], (2) syntactic complexity and grammaticality [66,103], (3) semantic fluency [104], and (4) patient recall ability [73,74]. Linguistic features are derived from automatically transcribed speech and processed using standard NLP tools. In addition, we will use

XSL•FO
RenderX

distilled RoBERTa [42] to generate contextual language representations that capture semantic relationships beyond surface-level lexical features.

Component 4: Modeling Patient-Nurse Interaction

Cognitive impairment also affects social communication, including turn-taking, timing, and responsiveness [105,106]. Patients with MCI-ED may show recurrent interactional patterns, such as longer turns, delayed responses, or reduced interactivity [107]. To capture these phenomena, we will model patient-nurse interaction using easily measurable dialogue features [76,77,108] (Table 1, component 4), including patient turn counts, dialog interactivity, turn density, turn duration, and relative timing of turns. These interactional measures reflect how patients engage with clinicians in real-world care encounters and provide information beyond speech content alone.

Component 5: System Implementation and Feature Extraction

Acoustic parameters for components 1 and 2 will be computed at the utterance level (continuous blocks of uninterrupted patient speech) using OpenSMILE [37] and PRAAT toolkits [38]. Verbal communications are automatically transcribed using AWS Transcribe, after which linguistic features for component 3 are computed at the encounter level using the NLTK toolkit and distilled RoBERTa. Interactional features for component 4 are derived from AWS Transcribe metadata, including speaker labels and time stamps. All extracted features will be aggregated at the patient level for downstream integration with clinical data in Aim 3.

Analytic Method for Aim 2: Extraction of MCI-ED–Related Information From Clinical Text Using LLMs

Rationale and Overview

Many clinical indicators of MCI-ED—including symptoms, lifestyle risk factors, and communication difficulties—are documented in free-text clinical notes or expressed during patient-nurse conversations [109], but are not captured in structured EHR fields [6,110]. Aim 2 uses LLMs to systematically extract this information from HHC clinical notes and transcripts of patient-nurse verbal communication. The goal is to convert unstructured text into standardized, patient-level variables that can be integrated with speech features (Aim 1) and structured assessment data (OASIS) for multimodal screening (Aim 3).

Information Specification and Reproducibility

For reproducibility, all MCI-ED–related information identified in Aim 2 is defined using an information schema summarized in component 1 (Information targets and schema). The schema specifies the target information families (clinical symptoms, lifestyle risk factors, and communication deficits) and associated attributes, and is applied consistently across human annotation and LLM-based identification. MCI-ED–related information is normalized to standard clinical terminologies—Unified Medical Language System (UMLS) [111] concepts, when available—and represented in a structured patient-level format, supporting reproducible integration with Aim 1 features and OASIS variables.

Component 1: Information Targets and Schema

We define a schema for the 3 MCI-ED–related risk factor categories, including clinical symptoms, lifestyle risk factors, and communication deficits. For each identified item, the system records (1) the related terms; (2) a normalized clinical concept identifier when available (UMLS [111] concepts); (3) clinically relevant attributes: assertion (present/absent/possible), temporality (current/historical), and experiencer (patient/caregiver); (4) severity and frequency; and (5) duration. The schema is used consistently across clinical notes and transcripts of patient-nurse communication.

Component 2: Human-Annotated Gold-Standard Dataset

Using the information schema defined in component 1, we create a human-annotated gold-standard dataset to support LLM adaptation and evaluation. The schema specifies the target information categories (clinical symptoms, lifestyle risk factors, and communication deficits) and associated attributes, including assertion, temporality, and experiencer.

Two trained nurse annotators independently annotate a stratified sample of HHC clinical notes and encounter transcripts according to this predefined schema. The annotation sample is stratified by race, sex, and visit type to ensure representation of diverse documentation patterns. Interannotator agreement is assessed using Cohen κ [112] for each information category and attribute, calculated on double-annotated samples prior to adjudication. Discrepancies are resolved through adjudication meetings to produce a finalized gold-standard dataset. The annotated corpus is subsequently partitioned into training, development, and test sets to enable LLM fine-tuning and unbiased performance evaluation.

Component 3: LLM-Based Identification Strategy

We use a hybrid LLM-based strategy that combines prompted extraction and instruction tuning [113], both aligned with the information schema (component 1) and supervised by the human-annotated gold-standard dataset (component 2). First, prompted extraction (baseline): as a baseline approach, we apply structured prompts that explicitly define the 3 information families (clinical symptoms, lifestyle risk factors, and communication deficits) and required attributes (assertion, temporality, and experiencer). Prompts instruct the model to produce schema-compliant JSON outputs and to provide a supporting text span for each identified item to ensure evidence-grounded extraction. This prompted approach provides an interpretable, rapidly adjustable method for early experiments and error analysis. Second, instruction tuning (schema-guided supervised adaptation): to improve reliability on HHC-specific language and documentation patterns, we perform instruction tuning using the training split of the human-annotated gold-standard dataset. Training examples pair the input text (note or transcript segment) with the target output formatted as schema-compliant JSON, including the identified item type, attributes (assertion, temporality, and experiencer), and supporting span. This teaches the model to follow the extraction instructions consistently and to produce outputs that match the

XSL•FO
RenderX

schema across diverse note styles and conversational phrasing. Instruction tuning is implemented using parameter-efficient methods (eg, Low-Rank Adaptation [114]/ Quantized Low-Rank Adaptation [114,115]) to reduce computational burden in the secure environment.

Component 4: Normalization and Patient-Level Aggregation

For each identified item in the text, we normalize the item (mention) to a standardized clinical concept identifier (UMLS, when available) using a controlled vocabulary lookup supplemented by string similarity matching for common variants. When multiple items map to the same concept within a document, we merge them into a single record while preserving the schema attributes (assertion, temporality, experiencer, and—when present—severity and frequency/duration) and retaining the supporting text spans for traceability. We then aggregate document-level outputs to the patient level to produce predictors for Aim 3. Patient-level variables summarize the presence of each normalized concept and its attributes across the patient's available HHC notes and encounter transcripts. The final deliverable is a structured patient-level table of normalized MCI-ED–related symptoms, lifestyle risk factors, and communication deficits for integration with OASIS data and Aim 1 speech and interaction features.

Component 5: Evaluation, Subgroup Checks, and Quality Control

We evaluate performance against the human-annotated gold-standard dataset using (1) span-level precision/recall/F_1 under exact and overlap matching, (2) attribute performance (assertion, temporality, experiencer, and severity/frequency when applicable), and (3) concept normalization accuracy. We report results overall and stratified by race, sex, and age group to monitor for systematic performance differences. We conduct routine error analysis (eg, common false positives from templated note language, negation errors, or transcript artifacts) and use findings to refine prompts, update normalization resources, and adjust fine-tuning settings. Low-confidence outputs are flagged for targeted review during development to guide iteration and reduce systematic errors.

Component 6: Aim 2 Outputs Used in Aim 3

The final Aim 2 deliverable is a structured, patient-level dataset summarizing clinical symptoms, lifestyle risk factors, and communication deficits identified from HHC notes and transcripts, including normalized concept identifiers and clinically meaningful attributes. These variables are used as candidate predictors and complementary signals in the multimodal screening algorithm developed in Aim 3.

Analytic Method for Aim 3: Development of a Multimodal Screening Algorithm for Identifying MCI-ED in HHC

Objective and Overview

The objective of Aim 3 is to develop and evaluate a multimodal screening algorithm for identifying HHC patients with MCI-ED. The algorithm integrates complementary information from three routinely generated data sources: (1) speech and interaction features extracted from patient-nurse communication (Aim 1), (2) MCI-ED–related information identified from clinical notes and transcripts (Aim 2), and (3) structured assessment data from the OASIS.

Data Sources and Candidate Predictors

Input variables include (1) acoustic, linguistic, emotional, and interactional features derived from patient-nurse verbal communication (Aim 1); (2) normalized clinical symptoms, lifestyle risk factors, and communication deficits identified from clinical notes and encounter transcripts (Aim 2); and (3) structured OASIS variables capturing sociodemographic characteristics, diagnoses, medications, functional status, and related clinical information.

Component 1: Data Preprocessing and Feature Preparation

Prior to model development, we assess data quality and address missingness, inconsistency, and integrity issues using a predefined data quality framework [116]. Continuous variables are transformed and scaled as appropriate to ensure comparability across modalities. To reduce dimensionality and mitigate overfitting in the presence of a large number of candidate predictors, we apply Joint Mutual Information Maximization [117] as a feature selection method. Joint Mutual Information Maximization is selected for its suitability in small to moderate sample settings with high-dimensional data, where it balances relevance to the outcome with redundancy among features.

Component 2: Model Development and Multimodal Integration

We develop screening models using supervised discriminative machine learning algorithms that are appropriate for tabular and multimodal clinical data, including logistic regression, support vector machines (SVMs) [118], and ensemble tree-based methods [119-122]. These models are chosen for their interpretability, robustness, and reduced risk of overfitting in clinical datasets. Multimodal integration is performed by combining features from speech, clinical text, and OASIS data within a unified modeling framework. Models are trained to estimate the probability of MCI-ED at the patient level. Temporal aspects of speech-derived features are summarized at the patient level prior to modeling, rather than modeled using complex sequence architectures, to maintain feasibility and stability given sample size considerations.

Component 3: Model Training, Validation, and Fairness Assessment

Model training and hyperparameter tuning are conducted using stratified nested cross-validation, with inner loops for parameter selection and outer loops for performance estimation. Model performance is evaluated using the area under the receiver operating characteristic curve (AUC-ROC) and area under the precision-recall curve, along with calibration measures to assess agreement between predicted risk and observed outcomes. To evaluate equitable performance, we assess model metrics stratified by race, sex, and age group. Fairness-related measures,

XSL•FO
RenderX

including group-wise differences in sensitivity and specificity and calibration across subgroups, are examined. When systematic performance differences are observed, we explore mitigation strategies such as reweighting or threshold adjustment and reassess model performance.

Component 4: Final Model Evaluation and Output

In the final step, the selected model is evaluated on an independent validation dataset to provide an unbiased estimate of performance. The screening algorithm produces a patient-level risk score indicating the likelihood of MCI-ED, which is intended to support clinical awareness and referral for further cognitive evaluation rather than serve as a diagnostic tool. The resulting algorithm and associated feature sets are prepared for downstream evaluation of clinical use and integration into HHC workflows.

Results

Initial Study Population and Clinical Characteristics

Between February 2024 and July 2025, we enrolled 114 HHC patients who met eligibility criteria and completed study-administered cognitive assessments. Following standardized review of the cognitive assessments and prespecified group-allocation procedures, 55 participants were classified as MCI-ED cases and 59 as cognitively normal controls. The cohort had a balanced sex distribution (n=58, 51% female) and was racially and ethnically diverse (n=63, 55.3% Black). Most participants were insured by Medicare (n=75, 66%), and 44% (n=50) lived alone.

In descriptive comparisons, participants classified as cognitively impaired had a higher prevalence of urinary incontinence (21/55, 36.8% vs 13/59, 21.4%), anxiety (32/55, 57.9% vs 23/59, 39.3%), and impaired vision (14/55, 26.3% vs 4/59, 7.1%), as well as greater dependence in activities of daily living (41/55, 73.7% completely dependent vs 38/59, 64.3%). These results are reported to characterize the cohort and should be interpreted descriptively rather than as definitive group differences.

Audio-recorded patient-nurse encounters had a median duration of 19 (IQR 12-23) minutes and a median of 56 (IQR 31-80) utterances per encounter. Across encounters, nurses contributed more words than patients (median 842, IQR 461-1218 vs 589, IQR 303-960), consistent with the structure of routine HHC visits and motivating inclusion of interactional features.

Preliminary Modeling Results

We conducted exploratory modeling analyses to evaluate the feasibility of distinguishing MCI-ED cases from cognitively normal controls using (1) speech-derived measures, (2) clinical text, and (3) structured EHR/OASIS variables, as well as multimodal combinations of these data sources. These analyses are intended to assess feasibility and inform subsequent model refinement and validation, rather than to provide definitive estimates of performance.

Speech-Derived Representations

Acoustic and temporal speech features were encoded using SpeechDETECT, including parameters related to phonetic motor planning (Table 1, component 1). Vocal emotion-related cues were encoded using GeMAPS (Table 1, component 2). Linguistic features included handcrafted measures capturing lexical richness, syntactic complexity, and semantic/fluency markers (eg, repetition and filler words; Table 1, component 3), and psycholinguistic indicators were extracted using LIWC 2015. In addition, we evaluated pretrained transformer language models for transcript-based representations.

Unimodal Performance

When modeling patient speech alone, DistilBERT achieved the strongest performance among evaluated BERT-based models (F_1=69.39; AUC-ROC=69.36). For clinical notes, BioClinicalBERT yielded the best performance among evaluated language models (F_1=64.29; AUC-ROC=69.17). Among traditional classifiers, a linear SVM performed well using patient speech features (F_1=75.0; AUC-ROC=75.94). Models using structured EHR/OASIS variables achieved their best performance with logistic regression (F_1=75.56; AUC-ROC=79.70).

Nurse Speech and Interactional Features

Incorporating nurse speech and interactional measures (Table 1, component 4) resulted in improved discrimination (SVM F_1=85.0; AUC-ROC=86.47), suggesting that patient-nurse interaction captures complementary information beyond patient speech alone.

Multimodal Integration

In multimodal analyses integrating speech features, interactional measures, and structured EHR/OASIS variables, the SVM achieved the highest overall performance (F_1=88.89; AUC-ROC=90.23). Examination of model contributions suggested that reduced lexical diversity, longer patient pauses, increased nurse dominance in conversation, selected psycholinguistic markers, and specific EHR variables (eg, non–insulin-dependent diabetes, pressure ulcers, and living alone) contributed to discrimination.

Overall, these results support the feasibility of extracting and integrating multimodal signals for MCI-ED screening in HHC. Final model evaluation, subgroup performance assessment, and fairness analyses will be conducted using the prespecified validation procedures after completion of recruitment and the finalized analytic dataset.

Discussion

Overview

This study protocol describes a multimodal screening approach for identifying MCI-ED in HHC using routinely generated data streams. The central hypothesis is that spontaneous patient-nurse conversations, combined with structured HHC assessment data (the federally mandated OASIS instrument) and information extracted from free-text clinical documentation, can provide complementary signals for earlier identification of cognitive impairment than any single data stream alone.

The implementation results reported in this manuscript demonstrate the feasibility of an end-to-end workflow in HHC,

XSL•FO
RenderX

including audio capture during routine visits, automated transcription and speaker labeling, extraction of acoustic, linguistic, emotional, and interactional features, and linkage to clinical notes and OASIS variables. Exploratory analyses in the analyzed cohort suggest that incorporating nurse speech and interactional features can improve discrimination beyond patient speech alone, consistent with the premise that conversation structure (eg, timing, pauses, turn-taking balance, and interactivity) contains clinically relevant information. These findings should be interpreted as feasibility and proof-of-concept evidence, rather than definitive estimates of model performance.

A substantial body of prior work has demonstrated that speech markers can differentiate individuals with Alzheimer disease from cognitively unimpaired controls, often using structured or semistructured tasks (eg, picture description, verbal fluency, and reading) collected in controlled environments [13]. While these approaches have been valuable for benchmarking and understanding underlying patterns, they may be less sensitive to the subtle and heterogeneous manifestations of MCI-ED and may not reflect communication behaviors during real-world clinical encounters.

This study extends this literature in 3 important ways. First, it shifts the speech signal from standardized tasks to naturally occurring clinical interactions in HHC, where pragmatic, temporal, and turn-taking patterns can be observed at scale. Second, it models not only patient speech characteristics but also interactional dynamics (including nurse speech), which may reflect clinician adaptation to support patients and/or patient difficulty maintaining conversational flow. Third, it advances a multimodal framework by integrating conversational speech features with (1) structured assessment variables from OASIS and (2) MCI-ED–related information embedded in free-text documentation or spoken conversation but not consistently represented in structured EHR fields. Together, these extensions aim to improve the practical relevance of screening in HHC settings where comprehensive cognitive evaluations may be limited.

Recent advances in minimally invasive approaches to Alzheimer disease detection, including blood-based biomarker testing in symptomatic individuals, reflect increasing clinical emphasis on earlier identification. However, biomarker confirmation alone does not characterize how cognitive decline affects communication during routine care. Changes in speech and language, such as reduced fluency, disrupted discourse organization, and altered vocal control, often emerge early and reflect functional impairment that is not captured by biological measures. The screening approach described here targets this complementary dimension by modeling communication behaviors observed in routine patient-nurse encounters, providing ecologically valid indicators of cognitive change. Integrating speech-based indicators with other clinical information, including biomarker evidence when available, may support a more comprehensive assessment of cognitive decline and its impact on real-world functioning.

Limitations

Several limitations should be considered when interpreting these results. First, findings are based on data from a single HHC organization and a modest sample, which may limit generalizability and yield performance estimates that are sensitive to sampling variability. Second, audio quality, background noise, and automated transcription/speaker-labeling errors can affect the accuracy of extracted acoustic, linguistic, and interactional features. Third, interactional measures may reflect both patient cognitive-linguistic status and clinician communication style or workflow constraints, which can introduce confounding if not explicitly modeled. Fourth, interactional measures may reflect both patient cognitive-linguistic status and clinician communication style or workflow constraints, which can introduce confounding if not explicitly modeled. Fourth, the protocol focuses on English-speaking participants with sufficient hearing/vision to complete cognitive testing; results may not generalize to other language groups or to patients with sensory limitations, who are common in HHC. Finally, cross-sectional classification does not establish whether speech and interaction markers predict future cognitive trajectories, underscoring the need for longitudinal evaluation.

Future Directions

Several avenues for future research could strengthen both the scientific rigor and clinical use of this approach. First, prospective longitudinal studies are needed to move beyond cross-sectional classification and evaluate whether speech and interactional markers can predict cognitive decline trajectories or functional deterioration over time. Such studies would clarify whether these features capture progressive change in addition to baseline differences. Second, external validation across diverse HHC agencies, geographic regions, and care delivery models will be essential to assess model transportability and identify when recalibration is necessary. Third, more granular analysis of conversational dynamics could distinguish clinically meaningful interaction patterns—such as repair sequences, prompting behaviors, and topic maintenance difficulties—from structural features that primarily reflect workflow or documentation practices. Fourth, incorporating clinician feedback through human-in-the-loop development cycles can help identify model failure modes, enhance interpretability of predictions, and establish safe deployment thresholds informed by real-world use cases. Finally, pragmatic clinical use trials are needed to determine whether integrating speech-based screening into HHC workflows improves downstream outcomes, including timeliness of formal cognitive evaluation, care plan modifications, and patient safety. Collectively, these efforts would bridge the gap between technical performance and meaningful improvements in care delivery for older adults at risk of cognitive decline.

Dissemination Plan

Findings will be disseminated through peer-reviewed publications and presentations to clinical and informatics audiences. To support reproducibility while protecting privacy, the study team plans to share (1) detailed feature definitions and extraction procedures, (2) deidentified analytic code and configuration files where permissible, and (3) the annotation schema and evaluation framework for text extraction.

XSL•FO
RenderX

Acknowledgments

I thank Dr James Noble and Margaret McDonald for scientific and clinical consultation that informed components of the study design and for their anticipated roles in study implementation. Their contributions did not include drafting or critically revising this protocol manuscript as authors.

Funding

This study was supported by grant K99/R00AG076808, “Development of a Screening Algorithm for Timely Identification of Patients with Mild Cognitive Impairment and Early Dementia in Home Health Care,” from the National Institute on Aging. The research methodology described in this manuscript derives from K99/R00AG076808, which was previously reviewed by the National Institute on Aging, received a score of 18, and was funded.

Additional support was provided by the Columbia Center for Interdisciplinary Research on Alzheimer’s Disease Disparities (NIH P30 AG059303), a Resource Center for Minority Aging Research that provides mentoring, pilot funding, and interdisciplinary research support to investigators addressing disparities in Alzheimer’s disease and related dementias.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Eligibility criteria.
[DOCX File , 17 KB - resprot_v15i1e82731_app1.docx]

Multimedia Appendix 2
Peer review report from AGCD-1 - Career Development Facilitating The Transition to Independence Study Section, National Institute on Aging (National Institutes of Health, USA).
[PDF File (Adobe PDF File), 127 KB - resprot_v15i1e82731_app2.pdf]

References

1. Dementia. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/dementia?utm_source=chatgpt.com [accessed 2025-12-16]
2. World Alzheimer Report 2025. Alzheimer’s Disease International. URL: https://www.alzint.org/resource/world-alzheimer-report-2025/ [accessed 2025-12-17]
3. 2025 Alzheimer’s disease facts and figures. The Journal of the Alzheimer's Association 2025;21(4):e70235. [doi: 10.1002/alz.70235]
4. Livingston G, Huntley J, Liu KY, Costafreda SG, Selbæk G, Alladi S, et al. Dementia prevention, intervention, and care: 2024 report of the Lancet Standing Commission. Lancet 2024;404(10452):572-628. [doi: 10.1016/S0140-6736(24)01296-0] [Medline: 39096926]
5. Zolnour A, Azadmaleki H, Haghbin Y, Taherinezhad F, Nezhad MJM, Rashidi S, et al. LLMCARE: early detection of cognitive impairment via transformer models enhanced by LLM-generated synthetic data. Front Artif Intell 2025;8:1669896 [FREE Full text] [doi: 10.3389/frai.2025.1669896] [Medline: 41280882]
6. Barrón Y, Ryvicker M, Song J, Zolnoori M, Topaz M. Identifying new ADRD diagnoses in home health care patients using natural language processing of nurses’ notes. Innov Aging. 2023 2023;7(Suppl 1):1060. [doi: 10.1093/geroni/igad104.3406]
7. Song J, Zolnoori M, McDonald MV, Barrón Y, Cato K, Sockolow P, et al. Factors associated with timing of the start-of-care nursing visits in home health care. J Am Med Dir Assoc 2021;22(11):2358-2365 [FREE Full text] [doi: 10.1016/j.jamda.2021.03.005] [Medline: 33844990]
8. Topaz M, Adams V, Wilson P, Woo K, Ryvicker M. Free-Text documentation of dementia symptoms in home healthcare: a natural language processing study. Gerontol Geriatr Med 2020;6:2333721420959861 [FREE Full text] [doi: 10.1177/2333721420959861] [Medline: 33029550]
9. Ryvicker M, Barrón Y, Shah S, Moore SM, Noble JM, Bowles KH, et al. Clinical and demographic profiles of home care patients with Alzheimer's disease and related dementias: implications for information transfer across care settings. J Appl Gerontol 2022;41(2):534-544 [FREE Full text] [doi: 10.1177/0733464821999225] [Medline: 33749369]
10. Burgdorf JG, Amjad H, Barrón Y, Ryvicker M. Undocumented dementia diagnosis during skilled home health care: prevalence and associated factors. J Am Geriatr Soc 2025;73(7):2117-2126. [doi: 10.1111/jgs.19491] [Medline: 40318054]

XSL•FO
RenderX

11. Zolnoori M, Zolnour A, Topaz M. ADscreen: a speech processing-based screening system for automatic identification of patients with Alzheimer's disease and related dementia. Artif Intell Med 2023;143:102624. [doi: 10.1016/j.artmed.2023.102624] [Medline: 37673583]
12. Zolnoori M, Barrón Y, Song J, Noble J, Burgdorf J, Ryvicker M, et al. HomeADScreen: developing Alzheimer's disease and related dementia risk identification model in home healthcare. Int J Med Inform 2023;177:105146. [doi: 10.1016/j.ijmedinf.2023.105146] [Medline: 37454558]
13. Azadmaleki H, Haghbin Y, Rashidi S, Momeni Nezhad MJ, Zolnour A, Zolnoori M. SpeechCARE: dynamic multimodal modeling for cognitive screening in diverse linguistic and speech task contexts. NPJ Digit Med 2025;8(1):677 [FREE Full text] [doi: 10.1038/s41746-025-02026-x] [Medline: 41249382]
14. Azadmaleki H, Zolnour A, Rashidi S, Noble JM, Hirschberg J, Esmaeili E, et al. TransformerCARE: a novel speech analysis pipeline using transformer-based models and audio augmentation techniques for cognitive impairment detection. Int J Med Inform 2026;207:106208. [doi: 10.1016/j.ijmedinf.2025.106208] [Medline: 41353845]
15. Jafari Z, Andrew M, Rockwood K. Diagnostic utility of speech-based biomarkers in mild cognitive impairment: a systematic review and meta-analysis. Age Ageing 2025;54(10):afaf316. [doi: 10.1093/ageing/afaf316] [Medline: 41148189]
16. Zolnoori M, Zolnour A, Vergez S, Sridharan S, Spens I, Topaz M, et al. Beyond electronic health record data: leveraging natural language processing and machine learning to uncover cognitive insights from patient-nurse verbal communications. J Am Med Inform Assoc 2025;32(2):328-340. [doi: 10.1093/jamia/ocae300] [Medline: 39667364]
17. Luz S, Haider F, de LFS, Fromm D, MacWhinney B. Alzheimer's dementia recognition through spontaneous speech: The ADReSS challenge. 2020 Presented at: Proceedings of the Annual Conference of the International Speech Communication Association. INTERSPEECH; 2020 Oct 25-29; Shanghai, China p. 2172-2176. [doi: 10.21437/interspeech.2020-2571]
18. Taherinezhad F, Nezhad M, Karimi S. Speech-based cognitive screening: a systematic evaluation of LLM adaptation strategies. arXiv Preprint posted online on August 18, 2025. [doi: 10.2196/preprints.82608]
19. Zolnour A, Azadmaleki H, Haghbin Y, Taherinezhad F, Nezhad MJM, Rashidi S, et al. LLMCARE: early detection of cognitive impairment via transformer models enhanced by LLM-generated synthetic data. Front Artif Intell 2025;8:1669896 [FREE Full text] [doi: 10.3389/frai.2025.1669896] [Medline: 41280882]
20. Zolnoori M, Azadmaleki H, Haghbin Y. National Institute on Aging PREPARE challenge: early detection of cognitive impairment using speech-the SpeechCARE solution. arXiv Preprint posted online on November 11, 2025. [doi: 10.48550/arXiv.2511.08132]
21. Azadmaleki H, Haghbin Y, Rashidi S, Momeni Nezhad MJ, Naserian M, Esmaeili E, et al. SpeechCARE: harnessing multimodal innovation to transform cognitive impairment detection-insights from the National Institute on Aging Alzheimer's speech challenge. Stud Health Technol Inform 2025;329:1856-1857. [doi: 10.3233/SHTI251249] [Medline: 40776266]
22. Zolnoori M, Vergez S, Kostic Z, Jonnalagadda SR, V McDonald M, Bowles KKH, et al. Audio recording patient-nurse verbal communications in home health care settings: pilot feasibility and usability study. JMIR Hum Factors 2022;9(2):e35325 [FREE Full text] [doi: 10.2196/35325] [Medline: 35544296]
23. Topaz M, Barrón Y, Song J, Onorato N, Sockolow P, Zolnoori M, et al. Risk of rehospitalization or emergency department visit is significantly higher for patients who receive their first home health care nursing visit later than 2 days after hospital discharge. J Am Med Dir Assoc 2022;23(10):1642-1647. [doi: 10.1016/j.jamda.2022.07.001] [Medline: 35931136]
24. Shankar R, Bundele A, Mukhopadhyay A. Natural language processing of electronic health records for early detection of cognitive decline: a systematic review. NPJ Digit Med 2025;8(1):133 [FREE Full text] [doi: 10.1038/s41746-025-01527-z] [Medline: 40025194]
25. Zhang Z, Gupta P, Song J, Zolnoori M, Topaz M. From conversation to standardized terminology: an LLM-RAG approach for automated health problem identification in home healthcare. J Nurs Scholarsh 2025;57(6):1003-1011. [doi: 10.1111/jnu.70039] [Medline: 40785044]
26. Chen YW, Ho W, Vergez SM. Hearing health in home healthcare: leveraging LLMs for Illness scoring and ALMs for vocal biomarker extraction. arXiv Preprint posted online on October 20, 2025.
27. Zhang Z, Momeni Nezhad MJ, Gupta P, Zolnour A, Azadmaleki H, Topaz M, et al. Enhancing AI for citation screening in literature reviews: improving accuracy with ensemble models. Int J Med Inform 2025;203:106035. [doi: 10.1016/j.ijmedinf.2025.106035] [Medline: 40609462]
28. Hosseini S, Momeni Nezhad MJ, Hosseini M, Zolnoori M. Optimizing entity recognition in psychiatric treatment data with large language models. Stud Health Technol Inform 2025;329:784-788. [doi: 10.3233/SHTI250947] [Medline: 40775965]
29. Zhang Z, Nezhad M, Hosseini S, Zolnour A, Zonour Z, Hosseini SM, et al. A scoping review of large language model applications in healthcare. Stud Health Technol Inform 2025;329:1966-1967. [doi: 10.3233/SHTI251302] [Medline: 40776319]
30. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
31. Berg L. Clinical dementia rating. Br J Psychiatry 2018;145(3):339-339. [doi: 10.1192/s0007125000118082]
32. Saramonic wireless microphone. Amazon. URL: https://tinyurl.com/3cwrvzpz [accessed 2025-12-17]

XSL•FO
RenderX

33. Kinatukara S, Rosati RJ, Huang L. Assessment of OASIS reliability and validity using several methodological approaches. Home Health Care Serv Q 2005;24(3):23-38. [doi: 10.1300/J027v24n03_02] [Medline: 16203688]
34. Tullai-McGuinness S, Madigan EA, Fortinsky RH. Validity testing the Outcomes and Assessment Information Set (OASIS). Home Health Care Serv Q 2009;28(1):45-57 [FREE Full text] [doi: 10.1080/01621420802716206] [Medline: 19266370]
35. RX 8: great audio starts with RX. IZOTOPE. URL: https://www.izotope.com/en/products/rx.html [accessed 2026-01-13]
36. Zolnoori M, Vergez S, Sridharan S, Zolnour A, Bowles K, Kostic Z, et al. Is the patient speaking or the nurse? Automatic speaker type identification in patient-nurse audio recordings. J Am Med Inform Assoc 2023;30(10):1673-1683. [doi: 10.1093/jamia/ocad139] [Medline: 37478477]
37. Eyben F, Schuller B. The Munich open-source large-scale multimedia feature extractor. ACM SIGMultimedia Rec 2015;6(4):4-13. [doi: 10.1145/2729095.2729097]
38. Praat Vocal Toolkit. URL: http://www.praatvocaltoolkit.com/ [accessed 2021-04-19]
39. Eyben F, Scherer KR, Schuller BW, Sundberg J, Andre E, Busso C, et al. The geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. IEEE Trans Affective Comput 2016;7(2):190-202. [doi: 10.1109/taffc.2015.2457417]
40. Pennebaker J, Boyd R, Jordan K, Blackburn K. The development and psychometric properties of LIWC2015. University of Texas at Austin. 2015. URL: https://www.liwc.app/ [accessed 2026-01-21]
41. Bird S, Klein E, Loper E. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit. Santa Rosa, CA: O'Reilly Media Inc; 2009.
42. Liu Y, Ott M, Goyal N. Roberta: a robustly optimized bert pretraining approach. arXiv Preprint posted online on July 26, 2019. [doi: 10.48550/arXiv.1907.11692]
43. Lanzi AM, Saylor AK, Fromm D, Liu H, MacWhinney B, Cohen ML. DementiaBank: theoretical rationale, protocol, and illustrative analyses. Am J Speech Lang Pathol 2023;32(2):426-438 [FREE Full text] [doi: 10.1044/2022_AJSLP-22-00281] [Medline: 36791255]
44. Meilán JJG, Martínez-Sánchez F, Martínez-Nicolás I, Llorente TE, Carro J. Changes in the rhythm of speech difference between people with nondegenerative mild cognitive impairment and with preclinical dementia. Behav Neurol 2020;2020:4683573 [FREE Full text] [doi: 10.1155/2020/4683573] [Medline: 32351632]
45. Themistocleous C, Eckerström M, Kokkinakis D. Voice quality and speech fluency distinguish individuals with mild cognitive impairment from healthy controls. PLoS One 2020;15(7):e0236009 [FREE Full text] [doi: 10.1371/journal.pone.0236009] [Medline: 32658934]
46. Huet K, Delvaux V, Piccaluga M, Roland V, Harmegnies B. Inter-syllabic interval as an indicator of fluency in Parkinsonian French speech. 2017 Presented at: 11th International Seminar on Speech Production; 2017 October 16-19; Tianjin, China. [doi: 10.21437/interspeech.2016-349]
47. Martínez-Nicolás I, Llorente TE, Martínez-Sánchez F, Meilán JJG. Ten years of research on automatic voice and speech analysis of people with Alzheimer's disease and mild cognitive impairment: a systematic review article. Front Psychol 2021;12:645 [FREE Full text] [doi: 10.3389/fpsyg.2021.620251] [Medline: 33833713]
48. Calzà L, Gagliardi G, Rossini Favretti R, Tamburini F. Linguistic features and automatic classifiers for identifying mild cognitive impairment and dementia. Comput Speech Lang 2021;65:101113. [doi: 10.1016/j.csl.2020.101113]
49. Dahmani H, Selouani S, O'shaughnessy D, Chetouani M, Doghmane N. Assessment of dysarthric speech through rhythm metrics. J King Saud Univ Comput Inf Sci 2013;25(1):43-49. [doi: 10.1016/j.jksuci.2012.05.005]
50. Boersma P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. 1993 Presented at: Proceedings of the Institute of Phonetic Sciences; 2026 Jan 13; Durham, North Carolina.
51. Viegas F, Viegas D, Guimarães GS, Souza MMGD, Luiz RR, Simões-Zenari M, et al. Comparison of fundamental frequency and formants frequency measurements in two speech tasks. Revista CEFAC 2019;21(6). [doi: 10.1590/1982-0216/201921612819]
52. Kong Y, Mullangi A, Marozeau J, Epstein M. Temporal and spectral cues for musical timbre perception in electric hearing. J Speech Lang Hear Res 2011;54(3):981-994 [FREE Full text] [doi: 10.1044/1092-4388(2010/10-0196)] [Medline: 21060140]
53. Tjaden K, Sussman JE, Liu G, Wilding G. Long-term average spectral (LTAS) measures of dysarthria and their relationship to perceived severity. J Med Speech Lang Pathol 2010;18(4):125-132 [FREE Full text] [Medline: 29422763]
54. On C, Pandiyan P, Yaacob S, Saudi A. Mel-frequency cepstral coefficient analysis in speech recognition. 2006 Presented at: International Conference on Computing & Informatics; 2006 June 6-8; Kuala Lumpur p. 1-5. [doi: 10.1109/icoci.2006.5276486]
55. Teixeira JP, Oliveira C, Lopes C. Vocal acoustic analysis–jitter, shimmer and HNR parameters. Procedia Technology 2013;9:1112-1122. [doi: 10.1016/j.protcy.2013.12.124]
56. Fraile R, Godino-Llorente JI. Cepstral peak prominence: a comprehensive analysis. Biomed Signal Process Control 2014 Nov;14:42-54. [doi: 10.1016/j.bspc.2014.07.001]
57. Simonyan K, Tovar-Moll F, Ostuni J, Hallett M, Kalasinsky VF, Lewin-Smith MR, et al. Focal white matter changes in spasmodic dysphonia: a combined diffusion tensor imaging and neuropathological study. Brain 2008;131(Pt 2):447-459 [FREE Full text] [doi: 10.1093/brain/awm303] [Medline: 18083751]

XSL•FO
RenderX

58. Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. J Commun Disord 2010;43(3):161-174. [doi: 10.1016/j.jcomdis.2009.12.004] [Medline: 20080243]
59. Tamarit L, Goudbeek M, Scherer K. Spectral slope measurements in emotionally expressive speech. 2008 Presented at: Proceedings of Speech Analysis and Processing for Knowledge Discovery; 2008 June 4-6; Aalborg, Denmark p. 169-183.
60. Deutsch D, Henthorn T, Dolson M. Absolute pitch, speech, and tone languageome experiments and a proposed framework. Music Percept 2004;21(3):339-356. [doi: 10.1525/mp.2004.21.3.339]
61. Gramming P, Sundberg J, Ternström S, Leanderson R, Perkins WH. Relationship between changes in voice pitch and loudness. J Voice 1988;2(2):118-126. [doi: 10.1016/s0892-1997(88)80067-5]
62. Narasimhan SV, Vishal K. Spectral measures of hoarseness in persons with hyperfunctional voice disorder. J Voice 2017;31(1):57-61. [doi: 10.1016/j.jvoice.2016.03.005]
63. Fergadiotis G, Wright HH, Green SB. Psychometric evaluation of lexical diversity indices: assessing length effects. J Speech Lang Hear Res 2015;58(3):840-852 [FREE Full text] [doi: 10.1044/2015_JSLHR-L-14-0280] [Medline: 25766139]
64. Sanborn V, Ostrand R, Ciesla J, Gunstad J. Automated assessment of speech production and prediction of MCI in older adults. Appl Neuropsychol Adult 2020;29(5):1-8 [FREE Full text] [doi: 10.1080/23279095.2020.1864733] [Medline: 33377800]
65. Ntracha A, Iakovakis D, Hadjidimitriou S, Charisis VS, Tsolaki M, Hadjileontiadis LJ. Detection of mild cognitive impairment through natural language and touchscreen typing processing. Front Digit Health 2020;2:567158 [FREE Full text] [doi: 10.3389/fdgth.2020.567158] [Medline: 34713039]
66. Cheung H, Kemper S. Competing complexity metrics and adults' production of complex sentences. Appl Psycholinguist 1992;13(1):53e76 [FREE Full text] [doi: 10.1371/journal.pbio.3002968] [Medline: 39836653]
67. Lakretz Y, Dehaene S, King J. What limits our capacity to process nested long-range dependencies in sentence comprehension? Entropy (Basel) 2020;22(4):446 [FREE Full text] [doi: 10.3390/e22040446] [Medline: 33286220]
68. Yeung C, Lee J. Automatic detection of sentence fragments. 2015 Presented at: The 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing; 2015 July 26-31; Beijing, China p. 599-603. [doi: 10.3115/v1/p15-2099]
69. López-de-Ipiña K, Martinez-de-Lizarduy U, Calvo PM, Beitia B, García-Melero J, Fernández E, et al. On the analysis of speech and disfluencies for automatic detection of mild cognitive impairment. Neural Comput Applic 2018;32(20):15761-15769. [doi: 10.1007/s00521-018-3494-1]
70. Tóth L, Gosztolya G, Vincze V. Automatic detection of mild cognitive impairment from spontaneous speech using ASR. 2015 Presented at: Sixteenth Annual Conference of the International Speech Communication Association; 2015 September 6-10; Dresden, Germany. [doi: 10.21437/interspeech.2015-568]
71. Xu Z, Vergez S, Esmaeili E, Zolnour A, Briggs KA, Scroggins JK, et al. Voice for all: evaluating the accuracy and equity of automatic speech recognition systems in transcribing patient communications in home healthcare. Stud Health Technol Inform 2025;329:1904-1906. [doi: 10.3233/SHTI251273] [Medline: 40776290]
72. Zolnoori M, Vergez S, Xu Z, Esmaeili E, Zolnour A, Anne Briggs K, et al. Decoding disparities: evaluating automatic speech recognition system performance in transcribing black and white patient verbal communication with nurses in home healthcare. JAMIA Open 2024;7(4):ooae130. [doi: 10.1093/jamiaopen/ooae130] [Medline: 39659993]
73. Ferson S, O'Rawe J, Antonenko A, Siegrist J, Mickley J, Luhmann CC, et al. Natural language of uncertainty: numeric hedge words. Int J Approx Reason 2015;57:19-39. [doi: 10.1016/j.ijar.2014.11.003]
74. Khodabakhsh A, Yesil F, Guner E, Demiroglu C. Evaluation of linguistic and prosodic features for detection of Alzheimer's disease in Turkish conversational speech. EURASIP J Audio Speech Music Process 2015;2015(1):9. [doi: 10.1186/s13636-015-0052-y]
75. Topaz M. NimbleMiner: a novel multi-lingual text mining application. Stud Health Technol Inform 2019;264:1608-1609. [doi: 10.3233/SHTI190558] [Medline: 31438255]
76. Roter DL, Larson SM, Beach MC, Cooper LA. Interactive and evaluative correlates of dialogue sequence: a simulation study applying the RIAS to turn taking structures. Patient Educ Couns 2008;71(1):26-33 [FREE Full text] [doi: 10.1016/j.pec.2007.10.019] [Medline: 18093788]
77. Drew P, Chatwin J, Collins S. Conversation analysis: a method for research into interactions between patients and health-care professionals. Health Expect 2001;4(1):58-70 [FREE Full text] [doi: 10.1046/j.1369-6513.2001.00125.x] [Medline: 11286600]
78. Eyben F, Wöllmer M, Schuller B. Opensmile: the munich versatile and fast open-source audio feature extracto. 2010 Presented at: Proceedings of the 18th ACM International Conference on Multimedia; 2010 oct 25-29; Firenze, Italy p. 1459-1462. [doi: 10.1145/1873951.1874246]
79. Bahgat M, Wilson S, Magdy W. Classifying online slang terms into LIWC categories. 2022 Presented at: 14th ACM Web Science Conference; 2022 Jun 26-29; Barcelona, Spain p. 422-432. [doi: 10.1145/3501247.3531572]
80. Duffy JR, Josephs KA. The diagnosis and understanding of apraxia of speech: why including neurodegenerative etiologies may be important. J Speech Lang Hear Res 2012;55(5):S1518-S1522 [FREE Full text] [doi: 10.1044/1092-4388(2012/11-0309)] [Medline: 23033445]

XSL•FO
RenderX

81. Ward M, Cecato JF, Aprahamian I, Martinelli JE. Assessment for apraxia in mild cognitive impairment and Alzheimer's dise. Dement Neuropsychol 2015;9(1):71-75 [FREE Full text] [doi: 10.1590/S1980-57642015DN91000011] [Medline: 29213944]
82. Nagumo R, Zhang Y, Ogawa Y, Hosokawa M, Abe K, Ukeda T, et al. Automatic detection of cognitive impairments through acoustic analysis of speech. Curr Alzheimer Res 2020;17(1):60-68. [doi: 10.2174/1567205017666200213094513] [Medline: 32053074]
83. Beltrami D, Gagliardi G, Rossini Favretti R, Ghidoni E, Tamburini F, Calzà L. Speech analysis by natural language processing techniques: a possible tool for very early detection of cognitive decline? Front Aging Neurosci 2018;10:369 [FREE Full text] [doi: 10.3389/fnagi.2018.00369] [Medline: 30483116]
84. Themistocleous C, Eckerström M, Kokkinakis D. Identification of mild cognitive impairment from speech in swedish using deep sequential neural networks. Front Neurol 2018;9:975 [FREE Full text] [doi: 10.3389/fneur.2018.00975] [Medline: 30498472]
85. Bidelman GM, Lowther JE, Tak SH, Alain C. Mild cognitive impairment is characterized by deficient brainstem and cortical representations of speech. J Neurosci 2017;37(13):3610-3620 [FREE Full text] [doi: 10.1523/JNEUROSCI.3700-16.2017] [Medline: 28270574]
86. Yu B, Williamson JR, Mundt JC, Quatieri TF. Speech-based automated cognitive impairment detection from remotely-collected cognitive test audio. IEEE Access 2018;6:40494-40505. [doi: 10.1109/access.2018.2856478]
87. Al-Hameed S, Benaissa M, Christensen H, Mirheidari B, Blackburn D, Reuber M. A new diagnostic approach for the identification of patients with neurodegenerative cognitive complaints. PLoS ONE 2019;14(5):e0217388. [doi: 10.1371/journal.pone.0217388]
88. Mirzaei S, El Yacoubi M, Garcia-Salicetti S, Boudy J, Kahindo C, Cristancho-Lacroix V, et al. Two-stage feature selection of voice parameters for early Alzheimer's disease prediction. IRBM 2018;39(6):430-435. [doi: 10.1016/j.irbm.2018.10.016]
89. König A, Satt A, Sorin A, Hoory R, Toledo-Ronen O, Derreumaux A, et al. Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease. Alzheimers Dement (Amst) 2015;1(1):112-124 [FREE Full text] [doi: 10.1016/j.dadm.2014.11.012] [Medline: 27239498]
90. Ning L, Luo K. Using text and acoustic features to diagnose mild cognitive impairment and Alzheimer’s disease. Research Square Preprint posted online on October 21, 2020. [doi: 10.21203/rs.3.rs-92702/v1]
91. Han KH, Zaytseva Y, Bao Y, Pöppel E, Chung SY, Kim JW, et al. Impairment of vocal expression of negative emotions in patients with Alzheimer's disease. Front Aging Neurosci 2014;6:101 [FREE Full text] [doi: 10.3389/fnagi.2014.00101] [Medline: 24904413]
92. Cadieux NL, Greve KW. Emotion processing in Alzheimer's disease. J Int Neuropsychol Soc 1997;3(5):411-419. [Medline: 9322399]
93. Kamiloğlu RG, Fischer AH, Sauter DA. Good vibrations: a review of vocal expressions of positive emotions. Psychon Bull Rev 2020;27(2):237-265 [FREE Full text] [doi: 10.3758/s13423-019-01701-x] [Medline: 31898261]
94. Weninger F, Eyben F, Schuller BW, Mortillaro M, Scherer KR. On the acoustics of emotion in audio: What speech, music, and sound have in common. Front Psychol 2013;4:292 [FREE Full text] [doi: 10.3389/fpsyg.2013.00292] [Medline: 23750144]
95. Spazzapan EA, de Castro Marino VC, Cardoso VM, Berti LC, Fabron EMG. Acoustic characteristics of voice in different cycles of life: an integrative literature review. Revista CEFAC 2019;21(3). [doi: 10.1590/1982-0216/201921315018]
96. Grabowski K, Rynkiewicz A, Lassalle A, Baron-Cohen S, Schuller B, Cummins N, et al. Emotional expression in psychiatric conditions: new technology for clinicians. Psychiatry Clin Neurosci 2019;73(2):50-62 [FREE Full text] [doi: 10.1111/pcn.12799] [Medline: 30565801]
97. Asgari M, Kaye J, Dodge H. Predicting mild cognitive impairment from spontaneous spoken utterances. Alzheimers Dement (NY) 2017;3(2):219-228 [FREE Full text] [doi: 10.1016/j.trci.2017.01.006] [Medline: 29067328]
98. Shen J, Rudzicz F. Detecting anxiety through reddit. 2017 Presented at: Proceedings of the Fourth Workshop on Computational Linguistics and Clinical Psychology — From Linguistic Signal to Clinical Reality; 2017 August 01; Vancouver, BC p. 58-65.
99. Sung JE, Choi S, Eom B, Yoo JK, Jeong JH. Syntactic complexity as a linguistic marker to differentiate mild cognitive impairment from normal aging. J Speech Lang Hear Res 2020;63(5):1416-1429. [doi: 10.1044/2020_JSLHR-19-00335] [Medline: 32402217]
100. Aramaki E, Shikata S, Miyabe M, Kinoshita A. Vocabulary size in speech may be an early indicator of cognitive impairment. PLoS One 2016;11(5):e0155195 [FREE Full text] [doi: 10.1371/journal.pone.0155195] [Medline: 27176919]
101. Mueller KD, Hermann B, Mecollari J, Turkstra LS. Connected speech and language in mild cognitive impairment and Alzheimer's disease: a review of picture description tasks. J Clin Exp Neuropsychol 2018;40(9):917-939 [FREE Full text] [doi: 10.1080/13803395.2018.1446513] [Medline: 29669461]
102. Ostrand R, Gunstad J. Using automatic assessment of speech production to predict current and future cognitive function in older adults. J Geriatr Psychiatry Neurol 2021;34(5):357-369 [FREE Full text] [doi: 10.1177/0891988720933358] [Medline: 32723128]

XSL•FO
RenderX

103. Davy W, Travis J, Laura W, Dona L, Graciela G. Towards automatic detection of abnormal cognitive decline and dementia through linguistic analysis of writing samples. 2016 Presented at: Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics; 2016 June 12-17; San Diego, California p. 1198-1207. [doi: 10.18653/v1/n16-1143]
104. Gosztolya G, Tóth L, Grósz T. Detecting mild cognitive impairment from spontaneous speech by correlation-based phonetic feature selection. 2016 Presented at: 17th Annual Conference of the International Speech Communication Association; 2016, September 8-12; San Francisco. [doi: 10.21437/interspeech.2016-384]
105. Johnson M, Lin F. Communication difficulty and relevant interventions in mild cognitive impairment: implications for neuroplasticity. Top Geriatr Rehabil 2014;30(1):18-34 [FREE Full text] [doi: 10.1097/TGR.0000000000000001] [Medline: 25356002]
106. Pan CW, Wang X, Ma Q, Sun H, Xu Y, Wang P. Cognitive dysfunction and health-related quality of life among older Chinese. Sci Rep 2015;5:17301 [FREE Full text] [doi: 10.1038/srep17301] [Medline: 26601612]
107. Ilan S, Carmel S. Patient communication pattern scale: psychometric characteristics. Health Expect 2016;19(4):842-853 [FREE Full text] [doi: 10.1111/hex.12381] [Medline: 26172267]
108. Dodge HH, Mattek N, Gregor M, Bowman M, Seelye A, Ybarra O, et al. Social markers of mild cognitive impairment: proportion of word counts in free conversational speech. Curr Alzheimer Res 2015;12(6):513-519 [FREE Full text] [doi: 10.2174/1567205012666150530201917] [Medline: 26027814]
109. Seinen TM, Kors JA, van Mulligen EM, Rijnbeek PR. Using structured codes and free-text notes to measure information complementarity in electronic health records: feasibility and validation study. J Med Internet Res 2025;27:e66910 [FREE Full text] [doi: 10.2196/66910] [Medline: 39946687]
110. Song J, Zolnoori M, Scharp D, Vergez S, McDonald MV, Sridharan S, et al. Do nurses document all discussions of patient problems and nursing interventions in the electronic health record? A pilot study in home healthcare. JAMIA Open 2022;5(2):ooac034 [FREE Full text] [doi: 10.1093/jamiaopen/ooac034] [Medline: 35663115]
111. Bodenreider O. The Unified Medical Language System (UMLS): integrating biomedical terminology. Nucleic Acids Res 2004;32(Database issue):D267-D270 [FREE Full text] [doi: 10.1093/nar/gkh061] [Medline: 14681409]
112. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
113. Zhang S, Dong L, Li X, Zhang S, Sun X, Wang S, et al. Instruction tuning for large language models: a survey. arXiv Preprint posted online on August 21, 2023. [doi: 10.48550/arXiv.2308.10792]
114. Hu EJ, Shen Y, Wallis P, Allen-Zhu Z, Li Y, Wang S, et al. Lora: low-rank adaptation of large language models. arXiv Preprint posted online on June 17, 2021. [doi: 10.48550/arXiv.2106.09685]
115. Li Y, Yu Y, Liang C, He P, Karampatziakis N, Chen W, et al. LoftQ: lora-fine-tuning-aware quantization for large language models. arXiv Preprint posted online on October 12, 2023.
116. Zolnoori M, Williams MD, Leasure WB, Angstman KB, Ngufor C. A systematic framework for analyzing observation data in patient-centered registries: case study for patients with depression. JMIR Res Protoc 2020;9(10):e18366 [FREE Full text] [doi: 10.2196/18366] [Medline: 33118958]
117. Bennasar M, Hicks Y, Setchi R. Feature selection using joint mutual information maximisation. Expert Syst Appl 2015;42(22):8520-8532. [doi: 10.1016/j.eswa.2015.07.007]
118. Murty MN, Raghava R. Kernel-Based SVM. In: Support Vector Machines and Perceptrons. Cham: Springer; 2016:57-67.
119. Zhou ZH, Tang W. Selective ensemble of decision trees. In: Wang G, Liu Q, Yao Y, Skowron A, editors. International Workshop on Rough Sets, Fuzzy Sets, Data Mining, and Granular-Soft Computing. Berlin: Springer; 2003:476-483.
120. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn 2006;63(1):3-42. [doi: 10.1007/s10994-006-6226-1]
121. Chen T, Guestrin C. Xgboost: a scalable tree boosting system. 2016 Presented at: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016 August 13; California, San Francisco, USA p. 785-794.
122. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. Lightgbm: A highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst 2017;30:3146-3154.

Abbreviations

AUC-ROC: area under the receiver operating characteristic curve
AWS: Amazon Web Services
CDR: Clinical Dementia Rating
EHR: electronic health record
GeMAPS: Geneva Minimalistic Acoustic Parameter Set
HHC: home health care
ICD-10: International Statistical Classification of Diseases, Tenth Revision
LIWC: Linguistic Inquiry and Word Count
LLM: large language model

XSL•FO
RenderX

MCI: mild cognitive impairment
MCI-ED: mild cognitive impairment and early dementia
MoCA: Montreal Cognitive Assessment
NLP: natural language processing
NLTK: Natural Language Toolkit
OASIS: Outcome and Assessment Information Set
SVM: support vector machine
UMLS: Unified Medical Language System

Edited by A Schwartz; submitted 20.Aug.2025; this is a non–peer-reviewed article;accepted 31.Dec.2025; published 02.Feb.2026.

Please cite as:
Zolnoori M
Developing a Multimodal Screening Algorithm for Mild Cognitive Impairment and Early Dementia in Home Health Care: Protocol for a Cross-Sectional Case-Control Study Using Speech Analysis, Large Language Models, and Electronic Health Records
JMIR Res Protoc 2026;15:e82731
URL: https://www.researchprotocols.org/2026/1/e82731
doi:10.2196/82731
PMID:

©Maryam Zolnoori. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 02.Feb.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Pilot Test of Diabetes Screening and Referral to Primary Care Within an Alaska Native Dental Clinic: Protocol for a Feasibility Study

Kathryn R Koller[1], PhD, RN; Julie A Beans[2], MPH; Diane K King[3], PhD; Vanessa Y Hiratsuka[2], MPH, PhD; Gretchen M Day[1], MPH; Shiela M Strauss[4], PhD; Barbara J Stillwater[1], PhD, RN

[1]Research Services, Alaska Native Tribal Health Consortium, Anchorage, AK, United States

[2]Research Department, Southcentral Foundation, 4085 Tudor Centre Drive, Research Department, Anchorage, AK, United States

[3]Center for Behavioral Health Research Services, University of Alaska Anchorage, Anchorage, United States

[4]Rory Meyers College of Nursing, New York University, New York, NY, United States

Corresponding Author:
Julie A Beans, MPH
Research Department, Southcentral Foundation, 4085 Tudor Centre Drive, Research Department, Anchorage, AK, United States

Abstract

Background: Recognizing the bidirectional relationship between oral health and diabetes mellitus (DM) and integrating DM screening and referrals to primary care into routine dental hygiene visits has strong potential to expand access to diabetes prevention and early detection services.

Objective: This paper describes methods used in an academic-Tribal partnership to co-design, coimplement, and pilot-test a DM screening and referral process within a Tribal health care setting.

Methods: The project uses implementation science frameworks and applies a 4-step iterative, participatory planning approach. Project implementation steps include (1) conducting key informant interviews with dental and primary care staff and administrators and surveying dental care recipients to assess the context for implementation and receptivity toward the proposed innovation; (2) forming and engaging a multidisciplinary planning team comprising dental, primary care, informatics, and care improvement personnel, care recipients, and researchers; (3) applying survey and interview findings and internal knowledge of clinic procedures and processes to develop a feasible dental DM screening service model to pilot test with iterative data collection; and (4) refining and disseminating the innovation.

Results: Key informant interviews and patient surveys were completed in September 2023. The planning team then developed the pilot process based on these findings in combination with lived experience. Study enrollment for the pilot began in April 2025 and was completed in July 2025. Data analysis is underway.

Conclusions: Offering DM screening during dental appointments aligns dental and primary care services to prevent or delay DM onset. This paradigm shift has the potential to facilitate communication between dental and primary care providers, improve the coordination of care across dental and medical services, capitalize on opportunities to reach individuals in need of DM screening while they are accessing other services, and expand health system capacity to identify individuals at risk and offer them appropriate DM services.

International Registered Report Identifier (IRRID): DERR1-10.2196/75619

(JMIR Res Protoc 2026;15:e75619) doi:10.2196/75619

KEYWORDS

diabetes mellitus type 2; Alaska Native; American Indian; prevention; dental

Introduction

Across the lifespan, oral health both influences and is influenced by general health, and one of the most common health conditions encountered in dental clinics that reflects this relationship is type 2 diabetes mellitus (DM) [1-4]. Elevated blood sugar levels in DM adversely affect oral health, manifesting as oral diseases, including periodontal disease. In the United States, an Indian Health Service brief posted in 2022 reported that while oral health has improved among Alaska Native and American Indian (ANAI) Peoples, this population is more than twice as likely to have periodontal disease as the general US population [5]. DM prevalence among ANAI adults is approximately twice that of US White adults [6]. Clinical interventions to improve access to DM or pre-DM screening through novel venues among this population are underreported in the literature, despite

XSL•FO
RenderX

surveillance data demonstrating the importance of oral health for persons diagnosed with diabetes and the potential for ANAI dental settings to serve as a screening and referral resource [7].

DM is the fifth leading cause of ANAI mortality for both adult men and women, in all age groups [8]. The medical needs of ANAI people with DM are complex because DM in this aggregated population is characterized by early onset and high rates of comorbidities (eg, heart disease, kidney failure, and lower limb amputation) [9]. A disease once rare among Alaska Native people specifically, DM now registers prominently among chronic diseases impacting Alaska Native people; its increased prevalence is attributed to the diet and lifestyle impacts of colonialism [10,11]. Findings from the Alaska Education and Research Towards Health baseline cohort study document the prevalence of modifiable risk factors contributing to DM progression (ie, tobacco use [12], obesity [12,13], large waist circumference [13], gestational diabetes [14], and physical inactivity [15]) among ANAI study participants between 2004 and 2006. A follow-up study conducted with a subset of urban Education and Research Towards Health participants during 2015 to 2017 [16] reported DM incidence of 16.5 per 1000 person-years and a pre-DM incidence of 77.6 per 1000 person-years [17]. DM prevalence (12% at baseline in 2004 - 2006) had increased to 25% and pre-DM prevalence (31% at baseline) had risen to 70% in 10 years, with mean participant age increasing from age 40 to 50 years [17].

The American Diabetes Association recommends DM screening for anyone overweight or obese if they present with additional risk factors [18]. Additional risk factors include (but are not limited to) physical inactivity, first-degree relative with DM, high-risk race or ethnicity (including ANAI people), smoking, hypertension, dyslipidemia, and blood sugar levels in the pre-DM range. For all adults without risk factors, the American Diabetes Association recommends screening beginning at age 35 years. Those with risk factors who do not routinely seek primary care are at the greatest risk.

Any oral infection with associated local and systemic inflammatory responses adversely affects blood glucose levels [4,19]. As DM is intricately interwoven as both a risk and a complicating factor in periodontal disease [1], dental hygiene appointments scheduled more frequently for patients with periodontal disease or those with same-day appointments for urgent dental care provide important opportunities to identify undiagnosed DM through screening in this higher risk group [20]. This study aims to develop, test, and evaluate a DM screening and referral service offered in a Tribal dental clinic.

Methods

Ethics Approvals

The Alaska Area Institutional Review Board (protocol 2017-02-004) reviewed and approved this study. Tribally owned and operated health care organizations, the Alaska Native Tribal Health Consortium (ANTHC) and Southcentral Foundation (SCF) research review committees reviewed and approved this study and this manuscript before journal submission [21]. The study is being conducted by ANTHC and SCF, and data collected are owned by SCF. Data access requests for data collected under this protocol will have to be submitted, reviewed, and approved by both ANTHC and SCF.

Population and Setting

SCF is an Alaska Native–owned, nonprofit health care organization operating under the Tribal authority of Cook Inlet Region, Incorporated. SCF provides dental and primary health care services for more than 70,000 ANAI people [22,23]. ANAI persons with periodontal disease are at increased risk for DM [19,24]. Those without regular primary care visits may not realize their risk for DM in the absence of discomfort or other physical symptoms, as is the case with early DM or pre-DM [19,25-27]. While studies in other population sectors have explored DM screening in dental clinics with varying degrees of success [28-35], none were conducted in a Tribal health system offering primary and dental services with shared electronic health records (EHRs).

SCF and ANTHC have research departments nested within the health care systems. These research departments conduct health research that is driven by and in service of ANAI people by ANAI people. This study's project team comprised ANAI investigators as well as non-ANAI investigators. The study team works to ensure data interpretations include ANAI perspectives and experiences, which is why we also receive community-level review of data interpretation from both Tribal organizations before dissemination.

Preparatory Work

SCF dental providers see more than 25,000 ANAI people from the Southcentral Alaska region annually. These findings have recently been corroborated (Koller KR, Hiratsuka VY, Day G, unpublished data, September 2017) by this research team between November 1, 2015, and April 30, 2017, showed at that time approximately 16% (n=4180) of dental clinic visits were coded as "same-day" (emergent or urgent) visits. Another 6% (n=1493) were coded by dental hygienists for periodontal disease care. Among 6795 unique individuals who did not have diagnosed DM, more than 76% (n=1722) of men and 55% (n=1872) of women seen on a same-day appointment basis had not been screened for DM during that 18-month period, despite the presence of DM risk factors. Specifically, nearly 54% were aged 45 years or older. Among those with scheduled hygiene appointments, 61% of men and 47% of women were unscreened for DM, and 42% were older than 45 years. Among ANAI individuals younger than 45 years, prevalent tobacco use (51% of same-day and 41% of hygiene appointment attendees) and obesity (approximately 40% in both clinic settings) indicate increased DM risk. These data support DM screening in the dental clinic setting for those being seen for same-day urgent care or for periodontal hygiene, with subsequent referral of those identified as at increased DM risk to their primary care provider. As early diagnosis and intervention during pre-DM can prevent DM onset, the service we propose could substantially reduce future DM incidence.

Study Framework

This study uses community-based participatory research principles in the context of Tribal sovereignty [36,37]. Our study

XSL•FO
RenderX

builds on previous research exploring the potential to conduct DM screening and referral during dental appointments, particularly for patients who do not regularly access primary care [20,28,30,35,38,39], and adapts these evidence-informed recommendations for the Tribal health setting.

We adapted the Centers for Disease Control and Prevention's *Replicating Effective Programs* Implementation Science framework, which recommends four steps to guide successful implementation (Figure 1) [40,41]: (1) explore the acceptability and perceived need of the intervention, (2) prepare for preimplementation, (3) pilot-test and refine processes based on monitoring data and feedback, and (4) adapt and disseminate.

Figure 1. Use of dissemination and implementation frameworks to promote external validity. CFIR: Consolidated Framework for Implementation Research; DM: diabetes mellitus; RE-AIM: reach, effectiveness, adoption, implementation, and maintenance.

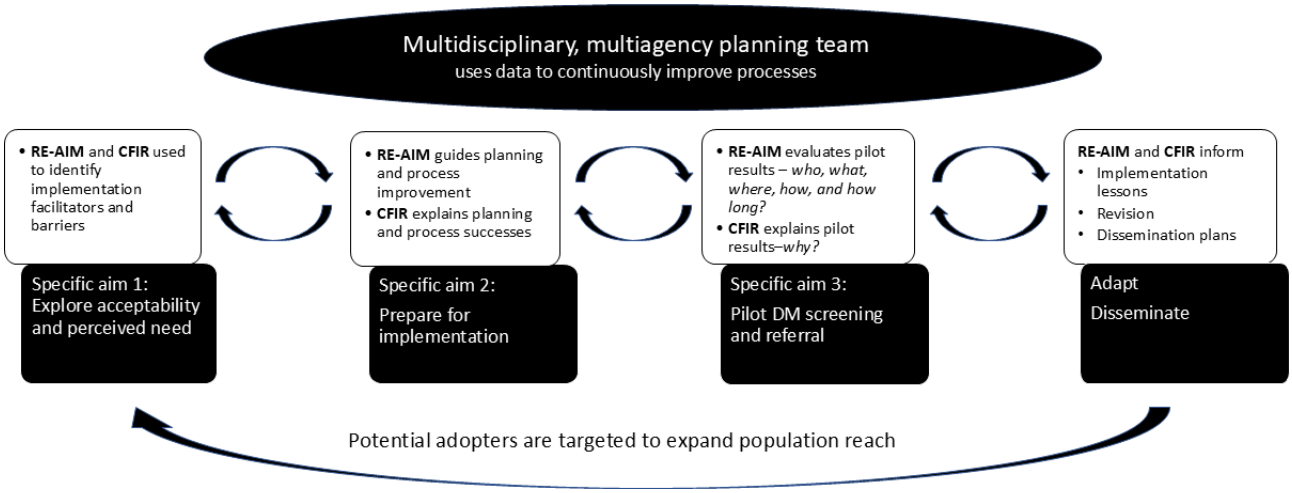

We will use the reach, effectiveness, adoption, implementation, and maintenance (RE-AIM) framework [42] to inform implementation strategies, use relevant process evaluation metrics, and evaluate implementation success [43]. We also included constructs from the Consolidated Framework for Implementation Research (CFIR) [44] to explore factors that explain RE-AIM outcomes (Figure 1) and provide an iterative systematic approach to design, pilot test, refine, and evaluate the screening and referral service [45,46].

Planning Team

Health system user input is critical in designing efficient, effective services that are acceptable to all involved, resulting in its optimal use [47]. Thus, together with an experienced study team serving as practice change facilitators, we will assemble a multidisciplinary, multiagency planning team that includes ANAI individuals who may both receive and provide dental and primary care at SCF, along with SCF internal process improvement staff representing both dental and primary care services. Using dissemination and implementation science frameworks, the planning team will collect formative information from groups involved with the health system to design implementation procedures and pilot test them.

Health system representatives are individuals with authority to make system-level changes (eg, to the EHR, training requirements, and equipment) and staff involved in service delivery (eg, primary care providers, dental providers, and health educators), quality improvement, data services, health information technology, administration, and Tribal leadership. Staff will include both ANAI and non-ANAI individuals. Volunteers from these areas will participate in the study planning team. The team will decide meeting frequency and communication methods. The planning team's meetings will focus on reviewing and providing input on activities across the entire study, including hands-on development of implementation, iterative process data collection and review protocols, pilot test planning, interpretation of preliminary analyses, and review and dissemination of findings.

Study Design and Overview

We will use a within-site, postdesign, phased approach [41] to study the implementation of an evidence-based DM preventive service delivered in this novel dental clinic setting [48]. In the first 2 years, we will use a convergent parallel mixed methods design [49] to analyze health system barriers and facilitators to implementing the DM screening and referral service and codevelop a dental clinic DM screening and referral service with dental and primary care clinic patients and staff (Table 1). In years 3 and 4, we will pilot-test the service, using RE-AIM metrics to evaluate reach, implementation fidelity, and the potential for setting-level maintenance. We will thematically code planning team meeting minutes, surveys, and interview transcripts using CFIR constructs to describe observed determinants that facilitated or were detrimental to reaching patients and implementing the intervention as planned.

XSL•FO
RenderX

Table . Research activities timeline.

	Year 1				Year 2				Year 3				Year 4			
	Q1[a]	Q2[b]	Q3[c]	Q4[d]	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Planning team meetings	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IRB[e] and Tribal approvals	✓	✓														
Dental clinic and primary care staff Interviews data collection and analysis	✓	✓	✓	✓	✓											
Dental patient surveys data collection and analysis		✓	✓	✓	✓											
Develop intervention clinic workflows						✓	✓	✓								
Train dental staff								✓	✓							
Enroll 60 dental patients										✓	✓	✓				

XSL•FO
RenderX

	Year 1				Year 2				Year 3				Year 4			
	Q1[a]	Q2[b]	Q3[c]	Q4[d]	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Electronic health record review at 1, 3, and 6 weeks postenrollment											✓	✓	✓			
CFIR[f] and REAIM[g] evaluation													✓	✓		
Summarize findings and present to Tribal leadership, internal departments, and community members														✓	✓	✓

[a]Q1: quarter 1.

[b]Q2: quarter 2.

[c]Q3: quarter 3.

[d]Q4: quarter 4.

[e]IRB: institutional review board.

[f]CFIR: Consolidated Framework for Implementation Research.

[g]REAIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance.

Dental Clinic and Primary Care Staff Interviews

Staff interviews will be used to understand factors relevant to referral service adoption, implementation, and maintenance. The guide is embellished with questions from the CFIR interview guide [50], which aim to explore the *outer* and *inner setting*, *individual characteristics*, and *process* constructs (Textbox 1). Such constructs include health care professional beliefs that the service is needed; will achieve the desired outcomes; and will fit with existing norms, culture, work processes, and practices.

XSL•FO
RenderX

Textbox 1. Interview guide.

Introduction

Diabetes is a significant health concern and an SCF (Southcentral Foundation) health priority. Our research team is exploring how to pilot diabetes screening for adults in the dental setting. The presence of moderate or severe periodontal disease is a significant risk factor for and a complication of diabetes. Due to this relationship, one method we are exploring is to screen adults for diabetes while they are at the dental clinic using a quick finger-stick blood test. We are also interested in screening adults in the walk-in dental clinic, as some of these individuals do not receive regular dental or primary care and may be at risk for diabetes.

a. Does the SCF dental clinics currently screen for chronic disease? Describe how that happens. Probes: can you describe the protocol in place for referral to primary care? If it does not happen, describe what might need to change to make this feasible. When in a dental visit does chronic disease screening and/or referral occur? What staff are involved? How are other chronic disease screening processes coded or billed? What issues do you foresee with respect to billing or coding diabetes screening in the dental clinic?

b. Does SCF Dental support individuals with diabetes currently? Probe: if yes, can you describe the support provided by SCF Dental for individuals with diabetes? Does SCF Dental support individuals with prediabetes currently? If so, can you describe how SCF Dental supports individuals with prediabetes?

c. Is the relationship between oral health and diabetes recognized by dental professionals? Probe: would diabetes screening and referral during a routine dental clinic visit fill an existing health need? Describe how.

d. Would adding diabetes screening and referral to primary care be feasible for SCF dental clinic to do now? Why or why not? Probe: what competing priorities do you see for implementing diabetes screening and referral in the dental clinic?

e. Whose buy-in will be necessary to successfully implement diabetes screening and referral in SCF dental clinics? Probe: who are current champions? Do you have suggestions on how we can best minimize resistance and/or disruptions to implementing diabetes screening and referral in SCF dental clinics?

f. What changes might be necessary to implement before introducing a new screening protocol (eg, documentation, billing, policies and procedures, and reporting)? Probe: whose work will be most impacted by adding diabetes screening and referral to primary care? How will it be impacted? Who in the dental clinic should or could screen for diabetes?

g. Screening encompasses several steps, including asking the customer-owner (SCF term for patients) about their diabetes status using a questionnaire, reviewing the questionnaire to determine whether to do a diabetes screening blood test, collecting and analyzing the blood sample, and then providing a referral and documenting the results in the patient record. Could or should different people be involved in these steps?* (notes which questions can be skipped if running short on time) Probe: who could or should do what? Can you foresee potential issues getting buy-in or acceptance of these duties? What skills and training will likely be needed? How do you train people now when introducing a new skill or procedure? How are new staff typically trained on other types of screening done in your practice?

h. Given that screening results may identify those currently at risk for diabetes as well as those with probable diabetes, who in your practice would be best positioned to relay the results, provide education, and coordinate next steps? Probe: can you foresee potential issues getting buy-in or acceptance of this role? Describe.*

i. What quality indicators do you currently use to monitor and make improvements to service delivery as well as customer-owner outcomes (eg, chart review, dashboard reports, staff, and customer-owner feedback)? What sorts of data would be helpful to monitor the success of a diabetes screening program in dental clinics?

j. Describe what you see as the main challenges to implementing diabetes screening as a part of routine dental care? Probe: are there system changes needed, such as health record documentation and prompts? Are there logistical barriers, such as staffing, scheduling, and workflow that would need to be addressed? Are there situations where we should not screen for diabetes in the dental clinic?

k. What types of mitigation practices were in place during the COVID-19 pandemic? Probe: are these practices still in place? Will these practices change how diabetes screening is managed in the dental clinic?*

l. Are there any other considerations that should be kept in mind or addressed before, or as part of, the diabetes screening pilot project?

Key informant participants will represent health system staff and be determined by the planning team, including representatives from dental and primary care center administration leadership and clinical providers at all levels. We will conduct at least 15 in-depth interviews to reach data saturation on core constructs. Key informant participants will receive a promotional item following interview completion. Experienced study team members will use ATLAS.ti (version 23.3.2.28729 for Windows) for qualitative analysis. Interviews will be coded using template analysis [51,52], with 2 coders reviewing each transcript to ensure intercoder reliability. Discrepancies will be discussed and resolved by consensus. Guided by CFIR domains and discussion with the planning team, template analysis will be used to identify barriers and facilitators.

Preimplementation Preparation

To plan for implementation, the planning team will engage in audio-recorded reflective planning sessions in which members will discuss the planning process, completed milestones, new or unresolved implementation barriers, and needed refinements. During these sessions, patient surveys and key informant interview data will be reviewed to identify barriers, facilitators, and perceptions (eg, patient and staff beliefs about the relative importance, value, and need for the service, and feasibility). The planning team will also codevelop intervention delivery protocols designed to be integrated, feasible, sustainable, and culturally relevant for ANAI patients. We will use a systems-level design approach to maximize patient reach, effectiveness, and buy-in from clinicians (adoption). The team

XSL•FO
RenderX

will meet monthly via videoconference and quarterly for in-person sessions.

Using ATLAS.ti (version 23.3.2.28729 for Windows), experienced members of the investigative team will code planning team meeting minutes, reflective session notes, and recorded meeting transcripts for RE-AIM domains and CFIR themes. These investigative team members will assess the extent to which reach, effectiveness, staff buy-in and training (adopters), implementation logistics, service quality, and long-term maintenance are addressed and will ensure that mechanisms for periodic refinement of implementation plans are in place. Analysis includes planning team composition characteristics, member participation rates, and other evidence of CFIR readiness for implementation and process determinants. Process determinants include leadership engagement (eg, acceptance and completion of assignments and responsiveness to requests for help), allocation of resources (eg, time for planning and training and funds for infrastructure changes), internal implementation leaders, and communication with stakeholders. Experienced members of the investigative team will provide the qualitative results to the planning team, which will further refine plans for obtaining clinic-level feedback and suggestions and monitor quality throughout the pilot study.

Pilot DM Screening and Referral

Once the planning team finalizes screening and referral protocols, the investigative team will train and certify clinicians or staff in all protocols. The planning team will make a final determination on which dental providers and clinics will be used. Inclusion criteria will be SCF dental patients aged 18 years or older being seen on a walk-in basis or for a prescheduled cleaning by a dental hygienist, ANAI heritage, at least 1 self-reported additional DM risk factor (eg, first-degree relative with DM, smoking, hypertension, dyslipidemia, and older than 35 years), and not currently managed for DM.

Interested individuals will be screened for eligibility by the investigative team in the dental clinic lobby. A study team member will go through the informed consent form with the potential participant and provide time for the individual to ask questions. Participants who provide informed consent will be assigned a unique anonymous study code stored in REDCap (Research Electronic Data Capture (REDCap; Vanderbilt University).. We anticipate demographic, contact, and basic DM risk factor information collected via an iPad survey to be linked in REDCap to each participant's study code. Clinic staff may measure height and weight, perform finger-stick glycated hemoglobin testing using planning team–selected point of care analyzers (eg, A1cNow, Affinion, or DCA Vantage), and initiate the screening and referral service for participants who provide informed consent.

The pilot will enroll 60 participants and will follow those referred to primary care through the dental screening and referral service to monitor process flow and follow-up and to iteratively refine the referral service. Adequate power to assess differences will not be an issue, as we will evaluate the screening and referral implementation rather than measure individual-level health outcomes. Upon signed informed consent, trained and experienced study staff will conduct EHR reviews at 1, 3, and 6 weeks postenrollment to monitor primary care follow-up. Pilot test participants will receeive a US $40 gift card at enrollment.As recommended, members of the investigative team will conduct regular practice facilitation calls with pilot testing clinics to iteratively identify challenges and successes, facilitate problem-solving, update the planning team, and refine protocols [53].

The primary outcome of this study is the assessment of the implementation process, resources required, and management needed for data collected during intervention implementation [54]. We will use descriptive statistics to describe the pilot study sample. We will collect implementation quality data using RE-AIM metrics and CFIR constructs, including EHR documentation of screening rates (reach), training rosters and evidence of leadership and clinician engagement (adoption), documentation of executing screening and referral services as planned (implementation), and planning team completion of a sustainability checklist of evidence-based indicators [47,55]. These indicators support continued postpilot test delivery of the service if demonstrated to be effective, including commitment of available resources (funding, training, and time) for implementation and ongoing operations (maintenance). We will conduct follow-up surveys with providers and pilot participants using REDCap [56] links to evaluate satisfaction with service delivery, suggestions for improvement, and system-level and patient effectiveness outcomes. Follow-up survey respondents will receive a US $15 gift card (pilot participants only) or a promotional item.

Using mixed methods, we will analyze the following RE-AIM data:

- Reach: number, percent, and representativeness of eligible pilot study patients who consent to screening in comparison to the overall population of eligible patients
- Effectiveness: success in engaging patients in testing and follow-up with a primary care provider, measured by the number and percent of referred patients completing a primary care visit where DM risk was addressed, as well as differences in completion rates by key demographic factors (ie, chi-square)
- Adoption: number, percent, and representativeness of invited providers and staff completing training
- Implementation: extent to which screening and referrals were conducted according to service protocols, number and percent of positive screens receiving a referral, number and percent of referred patients who complete a primary care appointment, and use of process data to adapt protocols
- Maintenance: results from a sustainability checklist developed by the planning team and team reflections on persistent barriers and intentions to adopt and disseminate the screening and referral service permanently

In the final step, evaluative data will be provided to the Tribal health care system leadership via the planning team to determine needed adaptations for dissemination beyond the pilot test.

XSL•FO
RenderX

Results

Dental and primary care staff interviews and patient surveys began in March 2023 and were completed in September 2023. The planning team has met a total of 12 times and used data gathered from the interviews and surveys to supplement their lived experience in planning the pilot study process. The pilot study began enrollment in April 2025 and completed it in July 2025. Data analysis is underway.

Discussion

Anticipated Findings

There is strong evidence of the robust bidirectional connection between periodontal disease and DM [57] and the urgent need to reach ANAI adults with and without periodontal disease for early identification and intervention to prevent the onset of undiagnosed or uncontrolled DM [17]. Our proposed methods to study the development and implementation of an evidence-based DM screening and referral service, which expands the role of dental practices to include DM screening and promotes care coordination among dental and primary care providers, are timely and relevant to national goals to address DM [58].

The major strengths of this study are (1) focused attention on dental services likely to have a greater proportion of the ANAI population with increased risk for DM, thereby maximizing screening efforts; and (2) referral to and follow-up with primary care providers for ongoing treatment and monitoring. Results from this study will be applicable to integrated health care systems, Tribal health care beneficiaries, and health systems operating patient-centered medical and/or dental home models.

Limitations

A limitation is that not all individuals at risk will be screened. However, successful implementation may pave the way for extending screening to other services in the future. In addition, our study of a single Tribal health care system will limit the generalizability of our findings, given the heterogeneity of dental and health care delivery across the United States.

Conclusions

The lessons and strategies for implementing a service to reduce the burden of undiagnosed DM and multimorbidity complications will be disseminated throughout SCF dental services as well as among other Tribal health organizations and dental health aide therapist programs. They may then be adapted and applied in Tribal health settings and rural regions of the nation. This project will also set the stage for a large-scale, sufficiently powered hybrid effectiveness [59]—or pragmatic, pretest-posttest, multisetting—implementation trial with long-term follow-up. The goal of such a large trial will be to supply vital information about the most feasible implementation strategies and outcomes when delivering coordinated dental–primary care services to prevent DM.

Acknowledgments

The authors sincerely thank the Alaska Area Institutional Review Board, the Alaska Native Tribal Health Consortium, and Southcentral Foundation's research review bodies for their approval of this study. The authors are also deeply grateful to the Alaska Native volunteers who provided the preliminary data for this proposal; the valuable results would not have been possible without their participation.

Conflicts of Interest

None declared.

References

1. Preshaw PM, Alba AL, Herrera D, et al. Periodontitis and diabetes: a two-way relationship. Diabetologia 2012 Jan;55(1):21-31. [doi: 10.1007/s00125-011-2342-y] [Medline: 22057194]
2. Kudiyirickal MG, Pappachan JM. Diabetes mellitus and oral health. Endocrine 2015 May;49(1):27-34. [doi: 10.1007/s12020-014-0496-3] [Medline: 25487035]
3. Cervino G, Terranova A, Briguglio F, et al. Diabetes: oral health related quality of life and oral alterations. Biomed Res Int 2019;2019:5907195. [doi: 10.1155/2019/5907195] [Medline: 31011577]
4. Borgnakke WS. IDF diabetes atlas: diabetes and oral health - a two-way relationship of clinical importance. Diabetes Res Clin Pract 2019 Nov;157:107839. [doi: 10.1016/j.diabres.2019.107839] [Medline: 31520714]
5. Phipps KR, Ricks TL. The oral health of American Indian and Alaska Native adult dental patients: results of the 2015 IHS oral health survey. Indian Health Service. 2016. URL: https://www.ihs.gov/doh/documents/IHS_Data_Brief_March_2016_Oral_Health%20Survey_35_plus.pdf [accessed 2025-05-25]
6. National Diabetes Statistics Report. Centers for Disease Control and Prevention. 2020. URL: https://www.cdc.gov/diabetes/php/data-research/index.html [accessed 2025-05-25]
7. Schroeder SM, Adamsen C, Besse R. The relationship between diabetes and oral health status, and dental visits among American Indian, Alaska Native, and Native Hawaiian elders. J Am Dent Assoc 2021 Apr;152(4):293-301. [doi: 10.1016/j.adaj.2020.12.008] [Medline: 33775287]
8. Heron M. Deaths: leading causes for 2017. Natl Vital Stat Rep 2019 Jun;68(6):1-77 [FREE Full text] [Medline: 32501203]

XSL•FO
RenderX

9. O'Connell JM, Wilson C, Manson SM, Acton KJ. The costs of treating American Indian adults with diabetes within the Indian health service. Am J Public Health 2012 Feb;102(2):301-308. [doi: 10.2105/AJPH.2011.300332] [Medline: 22390444]
10. Murphy NJ, Schraer CD, Theile MC, et al. Hypertension in Alaska natives: association with overweight, glucose intolerance, diet and mechanized activity. Ethn Health 1997 Nov;2(4):267-275. [doi: 10.1080/13557858.1997.9961835] [Medline: 9526689]
11. Fialkowski MK, Okoror TA, Boushey CJ. The relevancy of community-based methods: using diet within native American and Alaska native adult populations as an example. Clin Transl Sci 2012 Jun;5(3):295-300. [doi: 10.1111/j.1752-8062.2011.00364.x] [Medline: 22686210]
12. Lanier AP, Redwood DG, Kelly JJ. The Alaska Education and Research Towards Health (EARTH) study: cancer risk factors. J Cancer Educ 2012 Apr;27(1 Suppl):S80-S85. [doi: 10.1007/s13187-012-0333-4] [Medline: 22298198]
13. Slattery ML, Ferucci ED, Murtaugh MA, et al. Associations among body mass index, waist circumference, and health indicators in American Indian and Alaska native adults. Am J Health Promot 2010;24(4):246-254. [doi: 10.4278/ajhp.080528-QUAN-72] [Medline: 20232606]
14. Kim SY, England LJ, Shapiro-Mendoza CK, et al. Community and federal collaboration to assess pregnancy outcomes in Alaska native women, 1997-2005. Matern Child Health J 2014 Apr;18(3):634-639. [doi: 10.1007/s10995-013-1287-9] [Medline: 23775248]
15. Redwood DG, Lanier AP, Johnston JM, Asay ED, Slattery ML. Chronic disease risk factors among Alaska Native and American Indian people, Alaska, 2004-2006. Prev Chronic Dis 2010 Jul;7(4):A85 [FREE Full text] [Medline: 20550843]
16. Beans JA, Hiratsuka VY, Shane AL, et al. Follow-up study methods for a longitudinal cohort of Alaska native and American Indian people living within urban south central Alaska: the EARTH study. J Community Health 2019 Oct;44(5):903-911. [doi: 10.1007/s10900-019-00630-z] [Medline: 30798425]
17. Koller KR, Day GE, Hiratsuka VY, et al. Increase in diabetes among urban Alaska native people in the Alaska EARTH follow-up study: a call for prediabetes screening, diagnosis, and referral for intervention. Diabetes Res Clin Pract 2020 Sep;167:108357. [doi: 10.1016/j.diabres.2020.108357] [Medline: 32745696]
18. American Diabetes Association. Standards of medical care in diabetes—2022 abridged for primary care providers. Clin Diabetes 2022 Jan 1;40(1):10-38. [doi: 10.2337/cd22-as01]
19. Lalla E, Papapanou PN. Diabetes mellitus and periodontitis: a tale of two common interrelated diseases. Nat Rev Endocrinol 2011 Jun 28;7(12):738-748. [doi: 10.1038/nrendo.2011.106] [Medline: 21709707]
20. Strauss SM, Russell S, Wheeler A, Norman R, Borrell LN, Rindskopf D. The dental office visit as a potential opportunity for diabetes screening: an analysis using NHANES 2003-2004 data. J Public Health Dent 2010;70(2):156-162. [doi: 10.1111/j.1752-7325.2009.00157.x] [Medline: 20002875]
21. Hiratsuka VY, Beans JA, Robinson RF, Shaw JL, Sylvester I, Dillard DA. Self-determination in health research: an Alaska native example of tribal ownership and research regulation. Int J Environ Res Public Health 2017 Oct 31;14(11):1324. [doi: 10.3390/ijerph14111324] [Medline: 29088111]
22. Gottlieb K. The Nuka system of care: improving health through ownership and relationships. Int J Circumpolar Health 2013;72(1):21118. [doi: 10.3402/ijch.v72i0.21118] [Medline: 23984269]
23. Sherry P. Health care delivery for Alaska natives: a brief overview. Int J Circumpolar Health 2004;63 Suppl 2(sup2):54-62. [doi: 10.3402/ijch.v63i0.17786] [Medline: 15736623]
24. Lalla E, Kunzel C, Burkett S, Cheng B, Lamster IB. Identification of unrecognized diabetes and pre-diabetes in a dental setting. J Dent Res 2011 Jul;90(7):855-860. [doi: 10.1177/0022034511407069] [Medline: 21531917]
25. Elangovan S, Hertzman-Miller R, Karimbux N, Giddon D. A framework for physician-dentist collaboration in diabetes and periodontitis. Clin Diabetes 2014 Oct;32(4):188-192. [doi: 10.2337/diaclin.32.4.188] [Medline: 25646947]
26. Llambés F, Arias-Herrera S, Caffesse R. Relationship between diabetes and periodontal infection. World J Diabetes 2015 Jul 10;6(7):927-935. [doi: 10.4239/wjd.v6.i7.927] [Medline: 26185600]
27. Mawardi HH, Elbadawi LS, Sonis ST. Current understanding of the relationship between periodontal and systemic diseases. Saudi Med J 2015 Feb;36(2):150-158. [doi: 10.15537/smj.2015.2.9424] [Medline: 25719577]
28. Strauss SM, Alfano MC, Shelley D, Fulmer T. Identifying unaddressed systemic health conditions at dental visits: patients who visited dental practices but not general health care providers in 2008. Am J Public Health 2012 Feb;102(2):253-255. [doi: 10.2105/AJPH.2011.300420] [Medline: 22390440]
29. AlGhamdi AST, Bukhari SM, Elias WY, Merdad K, Sonbul H. Dental clinics as potent sources for screening undiagnosed diabetes and prediabetes. Am J Med Sci 2013 Apr;345(4):331-334. [doi: 10.1097/MAJ.0b013e318287c96c] [Medline: 23531966]
30. Strauss SM, Rosedale M, Pesce MA, et al. Point-of-care HbA1c testing with the A1cNow test kit in general practice dental clinics: a pilot study involving its accuracy and practical issues in its use. Point Care 2014 Dec;13(4):142-147. [doi: 10.1097/POC.0000000000000039] [Medline: 25593546]
31. Barasch A, Gilbert GH, Spurlock N, et al. Random plasma glucose values measured in community dental practices: findings from the dental practice-based research network. Clin Oral Investig 2013 Jun;17(5):1383-1388. [doi: 10.1007/s00784-012-0825-y] [Medline: 22903529]

XSL•FO
RenderX

32. Bossart M, Calley KH, Gurenlian JR, Mason B, Ferguson RE, Peterson T. A pilot study of an HbA1c chairside screening protocol for diabetes in patients with chronic periodontitis: the dental hygienist's role. Int J Dent Hyg 2016 May;14(2):98-107. [doi: 10.1111/idh.12140] [Medline: 25809551]
33. Genco RJ, Schifferle RE, Dunford RG, Falkner KL, Hsu WC, Balukjian J. Screening for diabetes mellitus in dental practices: a field trial. J Am Dent Assoc 2014 Jan;145(1):57-64. [doi: 10.14219/jada.2013.7] [Medline: 24379330]
34. Nasseh K, Greenberg B, Vujicic M, Glick M. The effect of chairside chronic disease screenings by oral health professionals on health care costs. Am J Public Health 2014 Apr;104(4):744-750. [doi: 10.2105/AJPH.2013.301644] [Medline: 24524531]
35. Wright D, Muirhead V, Weston-Price S, Fortune F. Type 2 diabetes risk screening in dental practice settings: a pilot study. Br Dent J 2014 Apr;216(7):E15-E1E. [doi: 10.1038/sj.bdj.2014.250] [Medline: 24722119]
36. Smith HS. An Alaska Native leader's views on health research. Int J Circumpolar Health 2013;72:108. [doi: 10.3402/ijch.v72i0.22447]
37. Cochran PAL, Marshall CA, Garcia-Downing C, et al. Indigenous ways of knowing: implications for participatory research and community. Am J Public Health 2008 Jan;98(1):22-27. [doi: 10.2105/AJPH.2006.093641] [Medline: 18048800]
38. Strauss SM, Stefanou LB. Interdental cleaning among persons with diabetes: relationships with individual characteristics. Int J Dent Hyg 2014 May;12(2):127-132. [doi: 10.1111/idh.12037] [Medline: 23790179]
39. Strauss SM, Singh G, Tuthill J, et al. Diabetes-related knowledge and sources of information among periodontal patients: is there a role for dental hygienists? J Dent Hyg 2013 Apr;87(2):82-89. [Medline: 23986141]
40. Kilbourne AM, Neumann MS, Pincus HA, Bauer MS, Stall R. Implementing evidence-based interventions in health care: application of the replicating effective programs framework. Implement Sci 2007 Dec 9;2(1):42. [doi: 10.1186/1748-5908-2-42] [Medline: 18067681]
41. Huynh AK, Hamilton AB, Farmer MM, et al. A pragmatic approach to guide implementation evaluation research: strategy mapping for complex interventions. Front Public Health 2018;6:134. [doi: 10.3389/fpubh.2018.00134] [Medline: 29868542]
42. Glasgow RE, Harden SM, Gaglio B, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64. [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
43. RE-AIM. URL: https://re-aim.org [accessed 2025-05-25]
44. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implementation Sci 2009 Dec;4(1):1-15. [doi: 10.1186/1748-5908-4-50]
45. King DK, Shoup JA, Raebel MA, et al. Planning for implementation success using RE-AIM and CFIR frameworks: a qualitative study. Front Public Health 2020;8:59. [doi: 10.3389/fpubh.2020.00059] [Medline: 32195217]
46. King DK, Gonzalez SJ, Hartje JA, et al. Examining the sustainability potential of a multisite pilot to integrate alcohol screening and brief intervention within three primary care systems. Transl Behav Med 2018 Sep 8;8(5):776-784. [doi: 10.1093/tbm/ibx020] [Medline: 29370421]
47. Hiratsuka VY, Moore L, Dillard DA, et al. Development of a screening and brief intervention process for symptoms of psychological trauma among primary care patients of two American Indian and Alaska native health systems. J Behav Health Serv Res 2017 Apr;44(2):224-241. [doi: 10.1007/s11414-016-9519-6] [Medline: 27328846]
48. Brown CH, Curran G, Palinkas LA, et al. An overview of research and evaluation designs for dissemination and implementation. Annu Rev Public Health 2017 Mar 20;38:1-22. [doi: 10.1146/annurev-publhealth-031816-044215] [Medline: 28384085]
49. Schoonenboom J, Johnson RB. How to construct a mixed methods research design. Kolner Z Soz Sozpsychol 2017;69(Suppl 2):107-131. [doi: 10.1007/s11577-017-0454-1] [Medline: 28989188]
50. Consolidated Framework for Implementation Research (CFIR). URL: https://cfirguide.org/ [accessed 2025-05-25]
51. Crabtree BF, Miller WL. Using codes and code manuals: a template organizing style of interpretation. In: Crabtree BF, Miller WL, editors. Doing Qualitative Research: A Practical Handbook, 2nd edition: SAGE Publications; 1999:163-178.
52. King N. Doing template analysis. In: Symon G, Cassell C, editors. Qualitative Organizational Research: Core Methods and Current Challenges: SAGE Publications; 2012:426-450.
53. Nutting PA, Crabtree BF, Stewart EE, et al. Effect of facilitation on practice outcomes in the National Demonstration Project model of the patient-centered medical home. Ann Fam Med 2010;8 Suppl 1(Suppl 1):S33-S44. [doi: 10.1370/afm.1119] [Medline: 20530393]
54. Kunselman AR. A brief overview of pilot studies and their sample size justification. Fertil Steril 2024 Jun;121(6):899-901. [doi: 10.1016/j.fertnstert.2024.01.040] [Medline: 38331310]
55. Luke DA, Calhoun A, Robichaux CB, Elliott MB, Moreland-Russell S. The program sustainability assessment tool: a new instrument for public health programs. Prev Chronic Dis 2014 Jan 23;11:130184. [doi: 10.5888/pcd11.130184] [Medline: 24456645]
56. About us. Research Electronic Data Capture (REDCap). URL: https://projectredcap.org/about [accessed 2025-05-25]
57. Preshaw PM, Bissett SM. Periodontitis and diabetes. Br Dent J 2019 Oct;227(7):577-584. [doi: 10.1038/s41415-019-0794-5] [Medline: 31605062]

XSL•FO
RenderX

58. Oral health in America: a report of the surgeon general. : U.S. Public Health Service, Department of Health and Human Services; 2000 URL: https://www.nidcr.nih.gov/sites/default/files/2017-10/hck1ocv.%40www.surgeon.fullrpt.pdf [accessed 2025-05-25]
59. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226. [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]

Abbreviations

ANAI: Alaska Native and American Indian
CFIR: Consolidated Framework for Implementation Research
DM: diabetes mellitus
EHR: electronic health record
RE-AIM: reach, effectiveness, adoption, implementation, and maintenance
REDCap: Research Electronic Data Capture
SCF: Southcentral Foundation

Edited by GRML Rosa; submitted 07.Apr.2025; peer-reviewed by AN Ali, C Hardin, J Thompson; revised version received 22.Dec.2025; accepted 20.Jan.2026; published 04.Mar.2026.

Please cite as:
Koller KR, Beans JA, King DK, Hiratsuka VY, Day GM, Strauss SM, Stillwater BJ
Pilot Test of Diabetes Screening and Referral to Primary Care Within an Alaska Native Dental Clinic: Protocol for a Feasibility Study
JMIR Res Protoc 2026;15:e75619
URL: https://www.researchprotocols.org/2026/1/e75619
doi:10.2196/75619

© Kathryn R Koller, Julie A Beans, Diane K King, Vanessa Y Hiratsuka, Gretchen M Day, Shiela M Strauss, Barbara J Stillwater. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 4.Mar.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Protocol

Developing the Link-me+EMPHN Mental Health Model of Care to Improve General Practitioner Capacity for Mental Health Care in Australian Primary Care: Protocol for a Mixed Methods Formative Study

Amy Coe[1], PhD; Jane London[2], BBehavSc, GradDipPsych, MSocSc; Anne-Marie Martin[2], BPharmHons; Aaron van Ree[2], BSW (Hons); Kirsty Lembke[2], BMedSci, MPharm, MPubHealth; Caroline Johnson[1], MBBS, PhD; Bridget Bassilios[3], PhD, DPsych; Catherine Kaylor-Hughes[1], BSc (Hons), DPhil

[1]Department of General Practice and Primary Care, The University of Melbourne, Parkville, Victoria, Australia

[2]Eastern Melbourne Primary Health Network, Box Hill, Victoria, Australia

[3]Melbourne School of Population and Global Health, The University of Melbourne, Carlton, Victoria, Australia

Corresponding Author:
Amy Coe, PhD
Department of General Practice and Primary Care
The University of Melbourne
Corner Grattan Street & Royal Parade
Parkville, Victoria, 3010
Australia
Email: amy.coe@unimelb.edu.au

Abstract

Background: The rising prevalence of mental health conditions continues to place significant pressure on general practitioners (GPs) and general practice. Despite their critical role in managing mental health conditions, GPs and practice staff face substantial barriers in providing effective mental health care, such as time and financial constraints. A new mental health model of care (Link-me+EMPHN) aimed at engaging and building capacity in GPs and practice staff for the provision of mental health services is being designed and implemented in a Primary Health region in Victoria, Australia.

Objective: This protocol describes formative research that aims to support the implementation of the Link-me+EMPHN Mental Health Model of Care in general practice, focusing on identifying current gaps in GP mental health training, assessing barriers and facilitators to implementation, exploring stakeholder perceptions and experiences, and determining needs and priorities for successful integration into routine practice.

Methods: The formative research involves a multimethod approach comprising (1) a desktop audit of currently available mental health training and guidelines for GPs; (2) up to 20 semistructured interviews with GPs and practice nurses; (3) an online survey exploring current mental health care practice of GPs and practice nurses; (4) two co-design workshops with people with lived experience of mental health care help-seeking in general practice and four expert working group workshops with a multidisciplinary primary-care team; and (5) testing of the model of care in a simulated practice setting with GPs, patients, and Care Navigators. The desktop audit, online survey, and interviews will be mapped to the Theoretical Domains Framework to systematically identify gaps in GP knowledge and skills in providing mental health care. Thematic analysis of the interviews will provide context for the gaps found in current mental health care and training so that we can begin to address these. Findings from each of the co-design, expert working group, and simulation sessions will be thematically analyzed and will include key themes, insights, and any practical implications.

Results: The formative research received funding in May 2024, with the formative research components taking place from May 2024 to July 2025. As of June 24, 2025, a total of 5 people participated in co-design, 14 in semistructured interviews, 28 in simulation sessions, 30 completed the online survey, and 4 working groups have been held. The desktop audit yielded 270 results. The findings of the formative research are tentatively planned for publication in 2026.

Conclusions: This formative research will generate a robust understanding of the factors influencing GP and practice engagement in mental health care and will provide practical solutions to improve implementation of the mental health model of care. The

XSL•FO
RenderX

findings will contribute to creating a sustainable and scalable model that improves mental health outcomes for patients and supports GPs.

International Registered Report Identifier (IRRID): DERR1-10.2196/79560

(JMIR Res Protoc 2026;15:e79560) doi:10.2196/79560

KEYWORDS

mental health; decision support; model of care; implementation; primary care; general practice; co-design; formative; Theoretical Domains Framework

Introduction

Background

Mental disorders are a significant global health concern, affecting approximately 14% of the population worldwide within a given year, while lifetime prevalence estimates exceed 30%-50% depending on the study and region [1,2]. In Australia, the burden is also significant, with 45% of people experiencing a mental health condition at some point in their lives and 1 in 5 adults facing a mental disorder each year [3]. Anxiety and affective disorders are the most common, affecting 17% and 8% of Australians each year, respectively [3]. Mental and physical health problems are frequently comorbid, with individuals who experience mental health conditions often having higher rates of chronic physical illnesses such as heart disease, diabetes, and obesity, leading to more complex care needs and worse health outcomes overall [4,5]. Additionally, as neurodevelopmental disorders like attention-deficit/hyperactivity disorder (ADHD) are becoming increasingly recognized in children, adolescents, and adults, there is an increasing need to incorporate ADHD care within the broader scope of mental health care [6,7]. While ADHD is primarily classified as a neurodevelopmental disorder, it is closely linked to mental health, frequently co-occurring with conditions such as anxiety, depression, and substance use disorders, and significantly impacting an individual's mental health and well-being [8,9]. Despite growing awareness of mental ill-health, delivering mental health care remains a challenge, especially in primary care settings. General practitioners (GPs), who are often the first point of professional contact for individuals seeking help with their mental health, play a pivotal role in diagnosing and managing these conditions [10].

In 2021, it was estimated that 38% of Australian general practice consultations involve a mental health component; however, it is thought that this number is greatly underreported [11]. Mental health conditions often require long-term management, complex care coordination, and a nuanced understanding of both psychological and physical health, placing GPs under increasing pressure to balance mental health care with other clinical duties and patient needs [12]. Eighty-three percent of Australian GPs report spending significant unpaid time coordinating care and follow-up for patients with mental health conditions, putting further strain on an already time-poor setting [11]. Over 70% of Australian GPs also agree that mental health consultations take longer than general consultations, creating logistical problems in their practice with an equal number of GPs reporting that mental health work is tiring and emotionally draining [11]. The cost of providing these longer, more complex mental health consultations and coordinating care can also be a significant financial barrier for GPs, especially in rural and remote areas with limited resources [11,13]. GPs also frequently face systemic barriers that limit their ability to provide comprehensive mental health care including difficulties in accessing multidisciplinary teams, such as psychologists and psychiatrists [14], which are crucial for complex mental health cases. This fragmentation between primary care and mental health services can result in delays and gaps in patient care and a lack of continuity in care [15,16]. Even when GPs are well-equipped with theoretical and experiential knowledge, the lack of real-time advice and guidance in managing complex mental health cases, disconnected care pathways, and inadequate infrastructure can prevent them from applying these skills in a holistic manner [15-18]. Together, these factors contribute to reluctance among GPs to take on more mental health patients, unless they have a particular interest in this field [11,19].

Implementing Collaborative Care

The challenges of managing mental health in primary care have led to the development of collaborative care models. Originating in the 1990s with the work of Wayne Katon and colleagues, these models were developed to address the significant gap between the prevalence of mental health conditions in primary care and the capacity of GPs to effectively manage them [20]. The core principles of collaborative care include team-based care delivery, population-based care, measurement-based treatment to target severity, and evidence-based treatment [21,22]. Over the past 30 years, numerous studies have demonstrated the effectiveness of collaborative care in improving outcomes for patients with common mental health conditions such as depression and anxiety [23,24]. The first comprehensive Cochrane review and meta-analysis in 2012 showed that collaborative care was associated with significant improvement in depression and anxiety outcomes compared to usual care [22].

However, despite strong evidence supporting its effectiveness, the implementation of collaborative care in routine practice has faced persistent challenges [25-27]. In the United Kingdom, the Improving Access to Psychological Therapies (IAPT) program (also known as NHS Talking Therapies) has expanded into primary care to make psychological therapies more accessible by offering services in various community settings and through different modalities (face-to-face, telephone, and online) using a collaborative care model [28,29]. The implementation of the IAPT has faced ongoing challenges, including issues in identifying suitable patients, providing holistic care, and ensuring continuity of support [30]. A qualitative study also

XSL•FO
RenderX

found that while a new Health and Wellbeing pathway within IAPT was perceived as having a positive impact on mental health, service- and individual-level barriers need to be addressed to enhance statutory and community support links, manage service users' expectations, and improve accessibility for certain groups [31]. A Norwegian model showed promising results in terms of improving GPs' ability to detect mental disorders in young people and potentially reducing long-term work disability through an adapted collaborative care model based on the successful Hamilton Family Health Team approach from Canada [32]. The model involved colocating psychologists and psychiatrists within general practice settings and demonstrated the feasibility of implementing core elements of collaborative care within a different health care context. However, it also highlighted significant challenges in sustaining such models without dedicated funding for collaborative activities, thus demonstrating the importance of aligning financial incentives with policy goals to support integrated care [32].

Australian Primary Health Networks (PHNs) are agencies that aim to strengthen the primary health care system in Australia. In their local catchment areas, PHNs (previously known as Medicare Locals and Divisions of General Practice) have been commissioning stepped primary mental health care since 2016 and commissioning and/or providing primary mental health care since 2001 [33-35]. Unlike IAPT, which serves the entire UK mental health help-seeking population, primary mental health care commissioned by PHNs targets disadvantaged groups, while universal primary mental health care has been available since 2006 via Australia's *Better Access to Psychiatrists, Psychologists and GPs* under the Medicare Benefits Schedule [36]. PHNs are best placed to support the implementation of collaborative care into clinical practice. Working with PHNs for funding and implementation of collaborative care models offers several advantages: (1) PHNs have a deep understanding of local health needs and resources, allowing for tailored implementation; (2) PHNs can potentially pool resources from various funding streams to support collaborative care initiatives; and (3) PHNs are well-positioned to facilitate cross-sector collaboration, crucial for the success of integrated care models. By working with PHNs, successful collaborative care models may be more easily scaled up across different practices within a region [37,38]. PHNs can also play a key role in monitoring implementation and outcomes, providing valuable data for ongoing improvement and policy development. Additionally, they can support the training and development of primary care staff to enhance their skills in mental health care.

GP Training

Alongside collaborative and stepped care models, emphasis has been placed on the provision of training to improve GP confidence and engagement in mental health care [39-41]. Various types of mental health training programs are available to GPs, ranging from short courses to more comprehensive programs [42]. These have been shown to increase GPs' mental health diagnosis rates, improve their ability to provide counseling and stress management, and manage somatic symptoms, as well as increase GP confidence, motivation, attitudes, and skills in treating patients with mental ill-health [43-45]. In Australia, GPs are not required to complete additional mental health training after completing specialist training; however, a higher level of mental health training correlates with a higher level of billing that a GP may claim for some specific mental health consultations [42,46]. In 2020, the Australian Productivity Commission Inquiry into Mental Health released a report that found "most GPs receive minimal training in mental healthcare" [47]. This finding appears to contradict the fact that 90% of Australian GPs have completed the recommended entry-level mental health skills training required to be eligible for higher billing rebates for mental health treatment plans [11] and suggests that training may be insufficient to fully prepare GPs for the complexities of mental health care in clinical practice. GPs also report feeling inadequately supported to provide effective mental health care [16,41,44]. It has been suggested that this may be due to limited access to, or knowledge of, mental health training and resources, and a lack of infrastructure to support ongoing learning [48]. However, qualitative studies of GP barriers to providing mental health care suggest that there may also be a gap in training related to how to apply theoretical knowledge to real-life challenges when delivering mental health care in practice [12,16].

The Link-me+EMPHN Mental Health Model of Care

To build capacity and support sustainable behavior change among GPs to provide best practice mental health care, the "Link-me+EMPHN Mental Health Model of Care" is being designed by the Eastern Melbourne Primary Health Network (EMPHN) and Primary Care Mental Health Research Team, University of Melbourne, in Victoria, Australia. The Link-me+EMPHN Mental Health Model of Care will consist of three pillars (Figure 1) [49,50].

XSL•FO
RenderX

Figure 1. Expected flow of the 3 pillars of the Link-me+EMPHN Mental Health Model of Care. The Link-me+ tool (pillar 3) has been previously validated. Although patients initially use the tool independently, results are integrated into GP consultations. For detailed information on patient flow and GP-patient interactions when using Link-me+, please refer to Fletcher et al [49,50]. EMPHN: Eastern Melbourne Primary Health Network; GP: general practitioner.

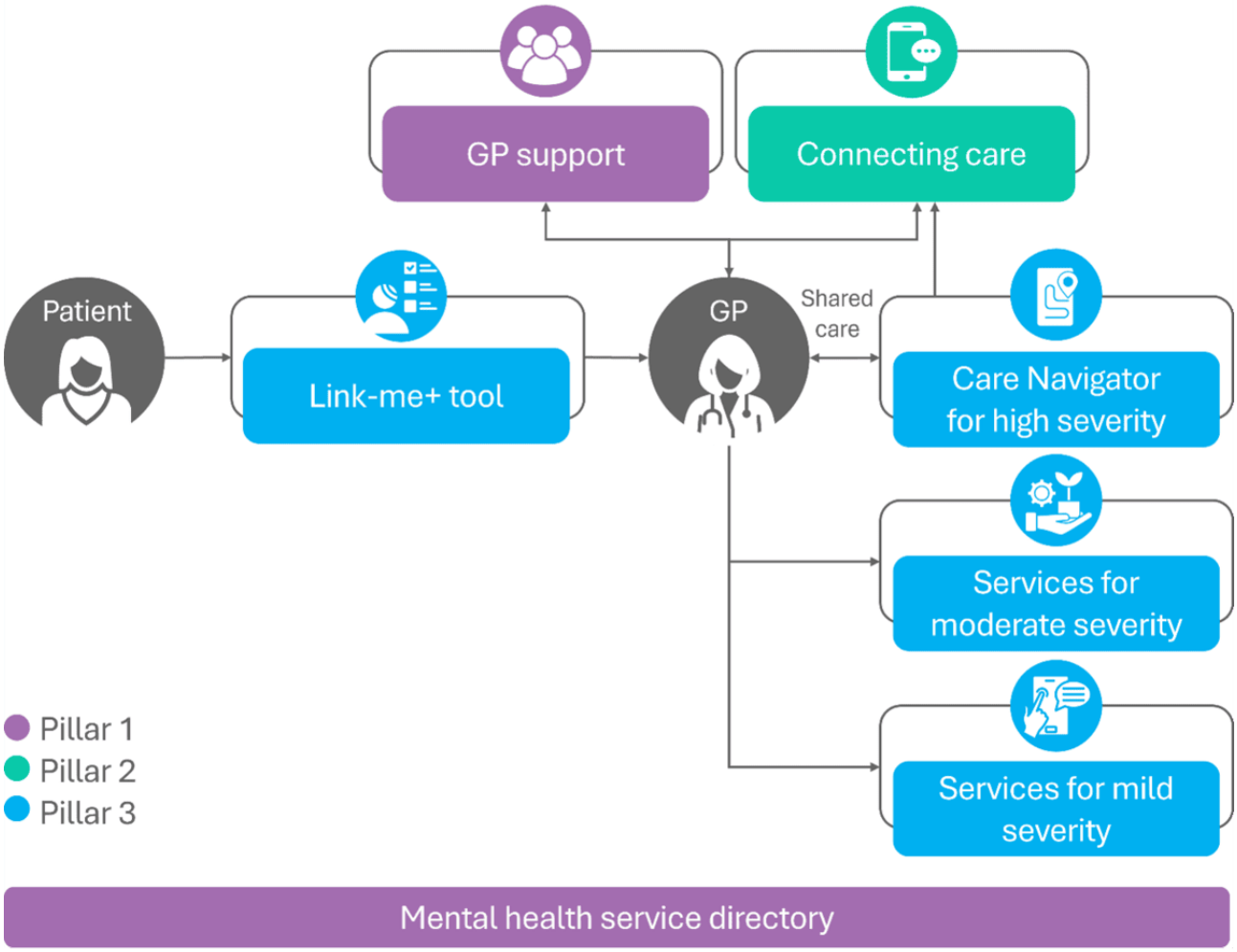

Pillar 1: GP Support

EMPHN will work with subcatchment groups of general practices aiming to fill gaps in clinical practice and enhance skills and confidence in managing mental health. A comprehensive service directory for use across all 3 pillars by all relevant staff and patients will be created. This directory will aim to offer a range of features to enhance service accessibility and coordination.

Pillar 2: SupportConnect

It aims to enhance SupportConnect to support GPs with clinical and practical queries. SupportConnect is a website and telephone service staffed by mental health professionals that helps people to navigate the mental health system and find appropriate services [51]. It assists individuals, families, caregivers, and health professionals with mental health support, suicide prevention, and substance use issues. Current usage of SupportConnect by GPs is low. Between December 2024 and June 2025, only 12% of contacts made to SupportConnect were from GPs. Care Navigators (briefly described in pillar 3) will also use SupportConnect. SupportConnect has been renamed “Medicare Mental Health” as of November 2025.

Pillar 3: Link-me+

The existing Link-me+ model of care will be scaled for implementation into practices to assist GPs with triaging. Link-me+ includes a digitally supported, prognostic tool that uses a stepped care approach to stratify people with mental health issues and then matches treatment to their predicted symptom severity (in 3 months) and level of need [49,50,52]. As part of the Link-me+ model, a Care Navigator will be allocated to patients with severe mental health symptoms. The Care Navigator is a pivotal role in which a trained health professional collaborates with patients and their GPs to develop and implement a holistic, person-centered mental health care plan that is based on patient needs and priorities for recovery. Care navigation is a short-term intervention that aims to identify and link patients with appropriate services [49,50,53]. The Link-me+ prognostic tool has been tested in 2 randomized controlled trials (RCTs) and has been demonstrated to reduce psychological distress in patients with anxiety and depression, especially in those with severe symptoms [49,50]. Additionally, the Link-me+ model of care leads to better outcomes when patients follow the recommended treatment [50]. The prognostic tool was developed from data collected as part of the Diamond 10-year longitudinal study of 789 Victorian (Australia) general

XSL•FO
RenderX

practice patients experiencing depressive symptoms [52]. User-centered design assisted the development of an interface to ensure the usability and acceptability within general practice settings for patients [54]. The Link-me+EMPHN Mental Health Model of Care, while primarily focused on depression and anxiety, can indirectly benefit patients with other mental health conditions due to the high rates of comorbidity [55-57] and the holistic, person-centered approach of the model. Therefore, the expert working group (including representatives from EMPHN and participating GPs) decided to incorporate ADHD as a focus in this project, alongside depression and social anxiety. This decision reflects the increasing recognition of adult ADHD in primary care settings and the need for improved management strategies. This project represents the first implementation of Link-me+ in routine general practice, transitioning it from a research tool to clinical use.

The Link-me+EMPHN Mental Health Model of Care will represent an innovative approach to stepped mental health care in general practice. The addition of pillars 1 and 2 to the existing Link-me+ model of care (pillar 3) represents a significant expansion of the original model, addressing identified needs in primary care mental health management and creating a more integrated, multifaceted approach to care delivery. This evolution from a single-focus intervention to a 3-pillar model may address the complex challenges of implementing collaborative mental health care in primary practice settings. To ensure its effectiveness and sustainability in routine practice, further formative research is necessary. Formative research is essential for developing effective interventions, programs, or strategies as it provides a deep understanding of the target population's needs, behaviors, and context and ensures that the final product is tailored to the end user [58,59]. In this way, formative research supports the creation of evidence-based solutions that are more likely to be accepted and adopted by the target population, ultimately leading to more sustainable and impactful results [59,60]. The Link-me+EMPHN Mental Health Model of Care formative research detailed in this protocol aims to:

1. Identify typical mental health care and knowledge gaps in current GP mental health training and support.
2. Determine the barriers to and enablers of the implementation of the Link-me+EMPHN Mental Health Model of Care in routine practice by GPs and other practice staff.
3. Explore the perceptions, experiences, and other influencing factors for the acceptance and feasibility of the Link-me+EMPHN Mental Health Model of Care in practice.
4. Identify stakeholder needs and priorities for the implementation of the Link-me+EMPHN Mental Health Model of Care in practice.

The findings from this formative research will not only inform the overall implementation strategy for the Link-me+EMPHN Mental Health Model of Care but will also guide the specific design and framing of each element within the 3 pillars. This approach will ensure that each component within the 3 pillars is tailored to address the unique challenges and opportunities identified in the EMPHN region.

Methods

Overview

The formative research will consist of a multipart, mixed methods formative study to develop and implement the Link-me+EMPHN Mental Health Model of Care (Figure 2):

- Part 1: A desktop audit of currently available resources for GPs.
- Part 2: An online survey of GPs, practice nurses, practice managers, and service managers.
- Part 3: Semistructured interviews with GPs.
- Part 4: Co-design sessions with general practice patients with lived experience of mental health care.
- Part 5: GP clinic simulation.

Central to the formative research, the identification of knowledge gaps, implementation factors, and priorities will be mapped to the Theoretical Domains Framework (TDF) [61,62] to help develop tailored supports for the successful implementation of the Link-me+EMPHN Mental Health Model of Care. The TDF is a framework that provides a comprehensive way to analyze behaviors and identify factors influencing change [61,62]. It consists of 14 domains of behavior change, such as knowledge, skills, social/professional role and identity, beliefs about capabilities, and environmental context, allowing for the identification of cognitive, affective, social, and environmental factors that impact decision-making and behavior [61,62]. Additionally, the TDF is underpinned by the COM-B (Capability, Opportunity, Motivation–Behaviour) model, which identifies how these components influence whether a desired behavior will occur [61,63]. The TDF will be used to explore GPs' capability to engage with the model (eg, knowledge and skills to confidently engage in training or use new models of care), external factors that enable or constrain behavior (eg, opportunities available, resources, time, and environmental context), and the motivation to adopt and sustain these changes (eg, weighing pros and cons, beliefs, and emotional regulation). The TDF and COM-B will ensure that the supports developed as part of the 3 pillars are closely aligned with the needs and challenges of GPs and practice staff, to increase the likelihood of successfully integrating the model into routine practice and improving uptake and sustainability.

This protocol adheres to the Comprehensive Criteria for Reporting Qualitative Research (CCQR) guidelines [64]. The CCQR checklist consists of 14 categories and has been used to ensure comprehensive and transparent reporting of both the qualitative and implementation aspects of our study.

XSL•FO
RenderX

Figure 2. The multipart formative research design to inform the 3 pillars of the Link-me+EMPHN Mental Health Model of Care. EMPHN: Eastern Melbourne Primary Health Network; GP: general practitioner.

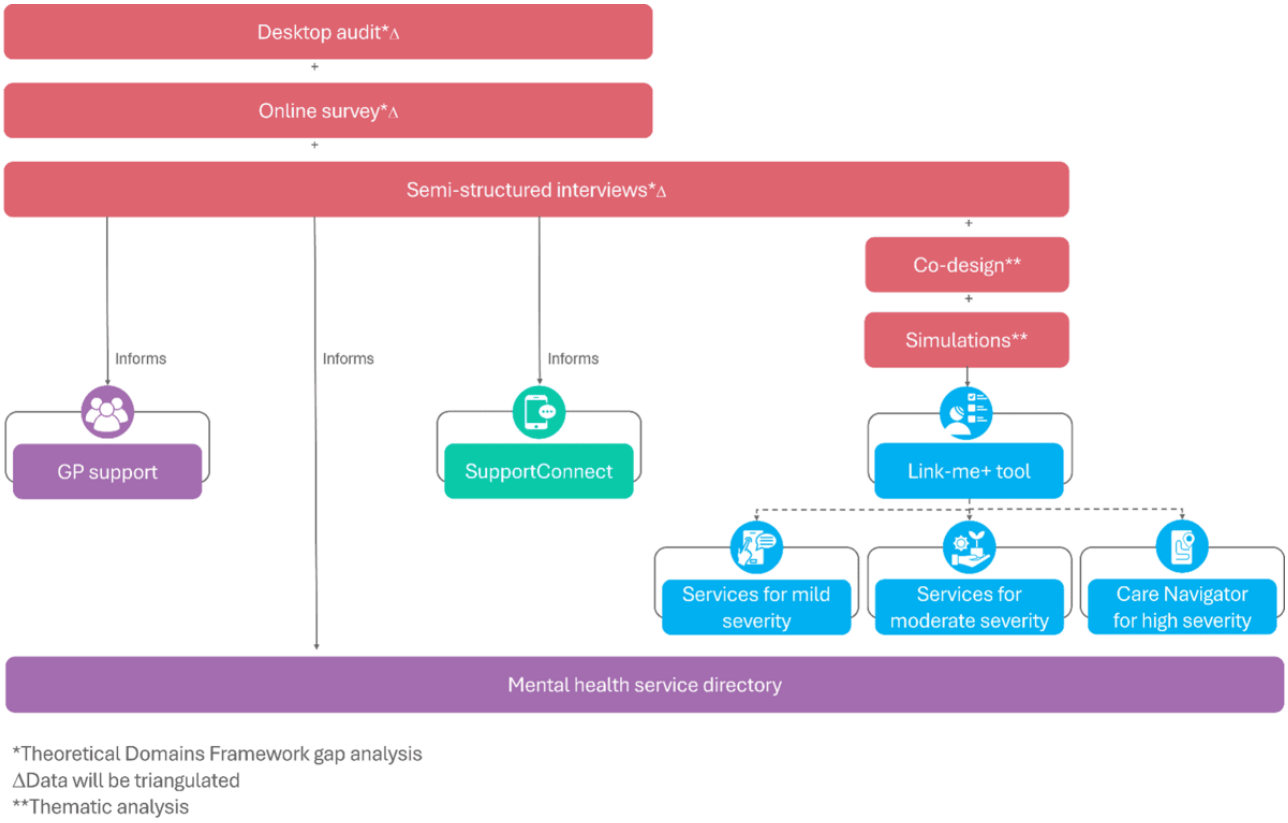

Setting

EMPHN serves a geographically and demographically diverse region in Victoria, Australia. EMPHN encompasses 12 local government areas (also known as municipalities), including the cities of Boroondara, Whitehorse, Knox, Maroondah, Manningham, Monash, Yarra Ranges, and parts of Banyule and Nillumbik. Approximately 1.62 million people live within the EMPHN catchment, where 441 GP clinics operate [65]. EMPHN plays a pivotal role in coordinating and enhancing primary health care services in this region, with a focus on improving health outcomes, particularly for vulnerable and underserved populations. The network is actively involved in implementing initiatives aimed at reducing health disparities, enhancing chronic disease management, and improving mental health services. PHNs support GPs by offering resources, training, and access to data and tools to help improve patient care. They also facilitate collaboration between GPs, specialists, and allied health providers.

Part 1: Desktop Audit

A desktop audit will be conducted to determine the current state of GP training and supports for mental health care. The audit will search for available training materials, curricula, guidelines, and resources used in GP training programs. This will include an examination of both formal educational content provided by medical schools and ongoing professional development offerings. The search will be limited to training available to GPs in the EMPHN catchment (including training that is available nationally).

A systematic search of known GP training providers (Royal Australian College of General Practitioners [RACGP], BlackDog, EMHPrac, Australian College of Rural and Remote Medicine [ACRRM], Inside Out, Mind Gardens, the Victorian and Australian government, Mental Health Professional Online Development [MHPOD] Program) and the EMPHN website will be conducted. A comprehensive Google search will also be conducted. The search for each website will use the key concepts of “mental health,” “mental health skills training,” “mental health training” AND “general practitioners,” and “mental health guidelines.” More specific terms related to mental health conditions will also be used (eg, depression, anxiety, domestic/family violence, substance misuse, and eating disorders). Due to the potentially high volume of available training, searches will continue until 3 pages of search results have passed with no relevant results. Identification of publicly available documents related to mental health quality improvement initiatives in general practices within the EMPHN catchment area will also be searched via Google. This search will specifically target mental health standard operating procedures, Plan-Do-Study-Act (PDSA) cycles, quality improvement plans, and mental health audits. These documents represent formal processes and strategies that practices may use to improve their mental health care delivery. The search will use combinations of keywords such as “mental health,” “quality improvement,” “PDSA,” “audit,” and “standard operating procedure.”

The searches will be conducted from inception to 2024. Only training in English language will be included. Training will need to be accessible to GPs working in the EMPHN catchment

XSL•FO
RenderX

and will be limited to training offered in Australia. All searches were conducted within a 1-week period in December 2024 to ensure consistency in available online content. The online survey (described in part 2) will also elicit information about GP mental health training.

Part 2: Online Survey

A 30-minute online survey will be sent to GPs, practice nurses, practice managers, and mental health service managers working in GP clinics within the EMPHN catchment. A target of over 100 responses will be sought via direct email, noticeboard posts, and advertisement on the EMPHN website. Advertisements will include a link to, or an attachment of, the plain language statement and consent information. Surveys will be completed digitally (via REDCap [Research Electronic Data Capture]) [66,67], although, if required, the surveys can be completed via pen and paper and returned via email, fax, or post. The survey will collect demographic information, including information about length of practice, qualifications, and current level of mental health training. Survey questions (see Multimedia Appendix 1) will be quantitative and will ask about current mental health practices and include a set of questions that relate to the definitions of the 14 domains of the TDF as per Michie et al [61,63]. Specifically, using a 7-point Likert scale ranging from "1=strongly disagree" to "7=strongly agree," participants will be asked to endorse their agreement with statements such as "Addressing mental health issues strengthens the doctor-patient relationship" (Beliefs About Consequences domain) and "I find mental health consultations emotionally draining" (Emotions domain).

The survey was developed through a collaborative process involving our research team's expertise in mental health care, general practice, and implementation science. Definitions of each of the TDF domains as they relate to mental health care were developed in consultation with the research team (see Multimedia Appendix 2). The questions were then formulated to align with the 14 domains of the TDF definitions, ensuring comprehensive coverage of potential barriers and enablers of implementing mental health care in general practice. To capture demographic information and current mental health practices, we adapted items from a survey used in similar research contexts over the past 15 years. This survey was developed from other validated surveys with permissions by the original developers [68-70]. This approach allowed us to build on established measures while tailoring the content to our specific research questions. Although we did not formally validate the entire survey, we conducted an internal review whereby each of us critically assessed each item for clarity, relevance, and alignment with the TDF domains. This process included multiple iterations of refinement based on team feedback.

Part 3: Semistructured Interviews

Semistructured interviews will be conducted with up to 20 GPs working in the EMPHN catchment. Practice nurses will also be invited to be interviewed; however, the priority will be gathering the perspectives of GPs. GPs will be recruited through purposive sampling to ensure a diverse representation of practice sizes, locations, genders, and years of clinical practice. Invitations to participate will (1) appear at the end of the online survey (detailed in part 2), (2) be sent via direct email to GPs and practice nurses, and (3) be posted on online noticeboards and the EMPHN website. The direct invitations and notices will include plain language information and a copy of the consent form. Interested participants will be asked to complete an online expression of interest form or to email the research team directly. The research team will then call or email the participant to arrange the interview at a mutually convenient time. Potential participants who have not previously completed the online survey will be asked to complete the demographic portion of the survey only, prior to attending the interview. Verbal consent will be audio-recorded prior to the commencement of the interview.

Interviews will be conducted via telephone, video call (Zoom), or in person. Each interview will last approximately 60 minutes. During the interview, up to 3 vignettes will be presented to the participant to facilitate discussion around how they usually provide mental health care, potential areas for improvement, and how the Link-me+EMPHN Mental Health Model of Care supports might fit in with their practice. Each vignette will focus on 1 of 3 different mental health diagnoses: depression, ADHD, and anxiety (see Multimedia Appendix 3). The interview questions and prompts will be informed by definitions of the 14 domains of the TDF as per Michie et al [61,63].

Part 4: Co-Designing Link-me+ With Providers and Consumers

To ensure that the implementation and refinement of the Link-me+ service is grounded in the experiences, needs, and preferences of general practice patients with lived experience of mental health issues and mental health care, up to 2 co-design sessions and 4 expert working group workshops will be facilitated.

Each co-design session will involve 5-10 participants. These participants will be recruited from an existing network of people with lived experience of mental health conditions and mental health help-seeking in general practice [71]. Participants will be provided with detailed information about the study and the co-design process prior to their participation. Participants in these co-design sessions have previously provided consent as part of the network members' participation in co-design activities [71]. Each co-design session will be facilitated by a co-design expert who also has a professional background in developing new and evidence-based models of mental health care for the primary care setting. The sessions will be structured to encourage open dialogue and the sharing of personal experiences. They will begin by introducing the purpose of the Link-me+ model of care and the role of participants in shaping its design. Participants will then be guided through a series of structured activities designed to elicit their priorities for how they would like to use Link-me+ when visiting their GP. These activities may include brainstorming sessions, group discussions, and the creation of visual representations (eg, journey mapping or storyboarding) of their ideal interaction with the Link-me+ model of care (ie, the prognostic tool and tailored treatment options). Discussions during the co-design sessions will be audio-recorded and transcribed verbatim.

XSL•FO
RenderX

Four hybrid (in-person and videoconference via Zoom) expert working group sessions will also be facilitated in the style of co-design. The working group consists of 15 people, including GPs, practice managers, nurses, service managers, researchers, and EMPHN staff. Each session will run for 6 hours and aims to determine how all 3 pillars of the Link-me+EMPHN Mental Health Model of Care will be implemented in practice. The sessions will be structured to (1) provide an overview of each pillar and its operationalization (eg, who is involved and where) and (2) facilitate a discussion to identify key priorities, less critical aspects, and potential additions for consideration. Throughout the sessions, participants will use sticky notes on journey boards to independently write down their perspectives. These will be collected by researchers and the EMPHN team. The sessions will also be audio-recorded and transcribed.

Part 5: GP Clinic Simulation

Think-aloud simulation sessions with up to 12 GPs, patients, and Care Navigators from the previous Link-me RCT [50] will be conducted using a prototype of the Link-me+ platform in a GP clinic setting. Some simulations may involve a GP-patient participant pair and/or a GP with a patient actor or a patient with a GP actor. Other simulations will be a participant completing the workflow with a member of the research team. The simulations will take place at the Digital Health Validitron SimLab at Melbourne Connect, University of Melbourne, or via videoconference (Zoom) where necessary. The Validitron is run by a multidisciplinary team of clinicians and research experts in digital health implementation, user design and human factors, and evaluation [72].

GPs working in the EMPHN catchment area will be invited to take part in the simulations via email, post, and online noticeboard advertisements asking them to email the research team to express interest. GPs who complete the online survey (detailed in part 2) will also be invited to take part at the completion of the survey. Patient participants will be recruited from the existing pool of people with lived experience of mental health problems used for the co-design sessions (part 4) [71]. Previous Care Navigators will be invited via email or telephone call. The Validitron team has access to a pool of actors who they will recruit as needed.

Different simulated clinic scenarios, based on insights from parts 1 to 4 of this formative research, will be created to mirror situations GPs, patients, and Care Navigators may face in practice when using the Link-me+EMPHN Mental Health Model of Care supports. Scenarios may include tasks such as completing the Link-me+ prognostic tool, diagnosing and assessing a patient's mental health status, selecting appropriate interventions, or coordinating care with mental health specialists and the Care Navigator. During the simulations, GPs, patients, and Care Navigators will be told to verbalize their thoughts in real time as they navigate a simulated version of the Link-me+ platform. For GPs and Care Navigators, this may include using the Link-me+EMPHN Mental Health Model of Care supports (eg, SupportConnect). All participants will be asked to verbalize their decision-making processes, challenges, and overall experience with the system. Each simulation session will last approximately 60 minutes. For in-person simulation sessions, participants will be observed by the research team through a 2-way mirror at the Validitron, and field notes will be taken. All in-person and videoconference sessions will also be audio-recorded and transcribed.

Data Analysis

Desktop Audit

Relevant data will be extracted from the search results into a purpose-built Microsoft Excel spreadsheet. Qualitative content analysis will be conducted to systematically categorize and interpret the extracted data. This will involve identifying the content of the available GP training and extracting the key elements according to relevance to the 14 TDF domains. For example, training that focuses on increasing understanding of symptom recognition, mental health comorbidities, and available community support services will be mapped to the "Knowledge" domain, while training that reinforces GPs' identity as holistic care providers will be coded under the "Social/Professional Role and Identity" domain. In addition to categorizing the content, quantitative content analysis will be used to count the frequency of particular elements. For instance, how often key areas such as symptom recognition, mental health comorbidities, or community support services are covered in the training programs. This will provide insight into how well-represented different domains are and highlight any gaps in training coverage.

Online Survey

Descriptive statistics for the demographic and TDF items will be calculated using Stata 15. Responses related to the TDF domains will be analyzed by summarizing and categorizing the data under the corresponding domains. This approach will allow for the identification of trends, potential barriers, and gaps within each domain.

Semistructured Interviews

The qualitative analysis will employ a dual approach: content analysis using the TDF framework definitions developed by the research team (Multimedia Appendix 2) and thematic analysis as described by Braun and Clarke [73]. The content analysis will be conducted to identify gaps in the current provision of mental health care. Two researchers will independently familiarize themselves with the data through repeated readings of the transcripts. Relevant sections will be coded according to the TDF domains, followed by a frequency count of each reference to a domain. The coding process will be discussed with the research team at regular intervals. Adjustments to definitions and coding will be made as necessary to ensure that the data are appropriately categorized within the framework.

Thematic analysis [73] will also be conducted on the interview transcripts to explore the COM-B elements surrounding the current provision of mental health care and participant suggestions for improvements. Deidentified transcripts will be read and reread by 2 researchers, and any patterns in the data relating to capability, opportunity, motivation, and other influences on GP and practice nurse behaviors will be noted. Identification of what may support future behavior change in mental health care will also be noted. Initial coding for thematic

XSL•FO
RenderX

analysis will be generated from the notes. Ongoing coding will be iterative and discussed regularly within the research team until final themes are decided upon.

Co-Designing Link-me+ With Providers and Consumers

The data collected as part of the co-design sessions and the expert working groups will be analyzed thematically. In particular, the data will be coded according to challenges, user needs, barriers, and potential solutions to implementation and participants' priorities for the Link-me+ model of care (eg, improving accessibility and user experience with the process of completing the prognostic tool). For both the co-design sessions and expert working group workshops, a summary of the most important findings from each session will be collated and will include key themes, insights, and any actions. The findings will be used to inform the implementation and integration of the Link-me+EMPHN Mental Health Model of Care in practice to ensure it aligns with the needs and expectations of its end users.

GP Clinic Simulation

Qualitative data from the simulation sessions will be analyzed using an inductive thematic analysis approach [73]. The coding will focus on identifying common patterns in both GPs' and patients' thought processes, challenges encountered, and features of the platform that facilitated or hindered their use of the Link-me+ tool and Link-me+EMPHN Mental Health Model of Care supports. Special attention will be given to identifying usability issues, areas of confusion, and potential improvements. The findings from this analysis will be used to refine the Link-me+ platform, ensuring it aligns with the workflows and needs of GPs and Care Navigators as well as the preferences and experience of patients.

Data Triangulation

Data from the desktop audit, semistructured interviews, and online survey will be triangulated to enhance the validity and reliability of the findings. Comparing and contrasting information obtained from these 3 sources will allow for a more comprehensive and nuanced understanding of the factors influencing the use of the Link-me+EMPHN Mental Health Model of Care supports and Link-me+ so that it is optimized for implementation into practice [74]. Triangulation will also enable a systematic assessment of whether current training adequately addresses each behavioral determinant essential for providing mental health care and provide insight into which domains should be focused on to support GPs as part of the Link-me+EMPHN Mental Health Model of Care [74]. Triangulation will involve using the content analysis of each source and looking to identify (1) convergence of findings, where all three sources indicate the same or similar gaps and/or enablers; (2) discrepancies between what is provided in the audit and GPs' expressed needs in the surveys and interviews; and (3) complementarity findings (eg, the desktop audit may show the types and frequency of mental health training available to GPs, while the survey could quantify GPs' participation rates in these trainings. The interviews could then provide in-depth insights into GPs' experiences with the training, including what they found most useful and what challenges they faced in applying the knowledge in practice. Together, these methods provide a more complete picture of the training landscape, from availability to uptake to practical application [74].

Rigor

Rigor for the whole of the formative research has been ensured using the gold standard evaluative criteria for trustworthiness as suggested by Lincoln and Guba [75]: credibility, transferability, dependability, and confirmability. Credibility will be established through data source triangulation, involving participants with diverse backgrounds and experiences across multiple general practices. The relevant multidisciplinary expertise of the research team, including backgrounds in neuroscience, psychology, general practice, pharmacy, mental health services research, and lived experience research, will support confirmability. This diverse perspective will be applied during data analysis, which will be conducted primarily by AC and supported by the authors and a research assistant with expertise in interviewing general practice stakeholders. Confirmability will also be enhanced by providing detailed participant quotes to support the interpretation of data. Transferability will be supported by thoroughly describing the sample in reports and research publications, allowing for potential replication. The sample, which includes GPs, nurses, and practice staff from various general practices, will provide an in-depth analysis of mental health care provision in the Australian general practice context. Dependability will be addressed through documentation of the data gathering and analysis processes, including a detailed protocol, digital notes, meeting minutes, and the publication of study outcomes. The methodology for this study builds upon well-established qualitative research practices and aligns with the theoretical frameworks of the TDF and COM-B model.

Ethical Considerations

Ethical approval has been obtained from the University of Melbourne Human Research Ethics Committee (HREC ID 29143) for all parts of this research. Informed consent will be obtained from all participants before the commencement of all parts of this formative research with the exception of the desktop audit search results. A well-established distress protocol, in use by the research team for over a decade, will be in place to manage any potential participant distress during the study. This protocol outlines clear procedures for researchers to follow if a participant experiences emotional distress, ensuring prompt and appropriate support is provided. All interview, co-design, and clinic simulation data will be deidentified during the transcription process. Participants will be assigned unique identifiers, and any potentially identifying information will be removed from the transcripts. The online survey is anonymous and does not ask for identifying information. All data will be stored securely on password-protected servers accessible only to the research team. Audio recordings will be deleted after transcription and verification. Interview, co-design, and simulation participants will be compensated for their time and contribution to the study with a Aus $100 (approximately US $66.96) gift card. Respondents to the online survey will go into a randomized draw to win 1 of 10 Aus $100 (approximately US $66.96) gift cards.

XSL•FO
RenderX

Results

A systematic search of the RACGP CPD learning, MHPOD, BlackDog, EMHPrac, ACRRM, Inside Out, EMPHN, Mind Gardens, and the Victorian and Australian government websites was conducted using relevant keywords (eg, mental health, mental health guidelines, domestic/family violence, substance misuse, and eating disorders) in November 2024. A Google search was also performed in November 2024. A total of 270 GP training opportunities were found. As only the learning objectives for each training course were readily available, these have been extracted for coding to the 14 domains of the TDF. As of June 24, 2025, a total of 30 people have completed the online survey. The survey remained open until the end of June 2025. The majority of participants were female (n=25, 84%), aged 51-60 years (n=12, 40%), and GPs (n=15, 50%) or practice nurses (n=10, 34%). The average time spent working in general practice was 13.65 years. Fourteen interviews with GPs have been completed with recruitment continuing until June 2025. Thematic analysis of the interview transcripts has commenced. Co-design and GP simulations sessions were completed by July 2025.

Discussion

Anticipated Findings

The Link-me+EMPHN Mental Health Model of Care is an innovative model that aims to build GP capacity and support them to provide better mental health care for their patients, while fostering patient self-awareness and ownership of care. This formative research aims to identify knowledge gaps in current GP mental health training as well as barriers to and enablers of the implementation of the Link-me+EMPHN Mental Health Model of Care in routine practice by GPs and other practice staff. Potential barriers are likely to include systemic and practice-level challenges, and enablers will include potential catalysts for behavior change among GPs and clinic staff.

Implementing models of care in general practice can be challenging due to a range of systemic and practice barriers including time constraints, high patient demand, and the complexity of addressing diverse health needs [76]. Integration of new models of care also requires additional coordination, training, documentation, collaboration with other services, and the need for behavioral shifts in practice which can impact GP and clinic staff buy-in [77]. To ensure the successful implementation and integration of the Link-me+EMPHN Mental Health Model of Care supports, this formative research will investigate catalysts for GP and clinic staff behavior change by collecting and analyzing data from multiple sources and a variety of stakeholders.

The findings from this formative research will inform potential optimizations of the implementation strategy. These could include tailoring the strategy to address specific barriers identified in different practice settings, developing targeted training programs to address skill gaps, adjusting the model to better align with existing practice workflows, and enhancing support systems for practices during the implementation process. While the core evidence-based components of Link-me+ will remain intact, the formative review may lead to adaptations in how the model is presented, integrated, or supplemented within the broader Link-me+EMPHN Mental Health Model of Care. Such adaptations could include expanding resources for managing comorbidities, incorporating additional assessment modules, refining the care navigation process, or developing supplementary materials to address frequently encountered clinical scenarios.

Comparison to Prior Work

While previous projects have often focused on single aspects of implementation, such as GP perspectives or organizational barriers, our formative research integrates views from GPs, nurses, practice staff, and mental health consumers, providing a more holistic understanding of the practice environment. For instance, Whitebird et al [78] conducted mixed methods research to identify successful elements for implementing collaborative care for depression in primary care with a focus primarily on organizational readiness and leadership engagement. Our study expands on this by not only considering organizational factors but also individual practitioner skills, attitudes, and behaviors, as well as patient-level considerations.

The Link-me+EMPHN formative research also shares similarities with Snyder et al [79] as both studies employ mixed methods designs and formative evaluation techniques and recognize the importance of tailoring implementation strategies to local contexts. Both studies also contribute to the evolving field of implementation science in mental health care, demonstrating how formative evaluation methods can be adapted to different health care contexts. However, Snyder et al [79] focused on improving suicide prevention practices within a single community mental health clinic, while Link-me+EMPHN addresses a broader spectrum of mental health issues across a network of primary care practices. Additionally, Snyder et al [79] developed a supervisory toolkit as their primary implementation strategy, but our research will offer a more diverse range of strategies due to the complexity of implementing a digital intervention across multiple primary care sites.

The SMART (Systematic Medical Appraisal, Referral, and Treatment) Mental Health programme provides another relevant comparison. SMART Mental Health heavily emphasized the use of mobile-based applications and an electronic decision support system (EDSS) for community health workers and primary care doctors [80]. Our study, while incorporating digital elements, has a broader focus on GP engagement and capacity building. The formative phase for SMART Mental Health included preliminary feasibility testing of their EDSS, providing quantitative data on screening and diagnosis rates [80]. The Link-me component (pillar 3) of the Link-me+EMPHN Model of Care has already undergone testing in 2 RCTs and additional user-testing. Accordingly, our formative phase is more focused on qualitative insights to inform the subsequent implementation strategies. SMART Mental Health was also implemented in rural India, focusing on low-resource settings, while our study is set in an Australian context with a different health care infrastructure.

XSL•FO
RenderX

The findings of a Danish implementation study highlight the extensive implementation work required by care managers, including solving practical problems of location and logistics, engaging GPs, and tailoring collaboration to meet GP preferences [81]. Although our study focuses on the broader implementation of a mental health model of care, we anticipate similar challenges in integrating Care Navigation and other new processes into existing primary care settings.

Collectively, these studies highlight the importance of context-specific approaches in implementation research, showcasing how similar methodological foundations can be applied to address distinct challenges in different mental health care settings. Additionally, research studies specifically addressing the implementation of comprehensive mental health models of care in primary settings are relatively scarce. Although implementation science in health care has advanced [82], there appears to be a need for more focused formative research on integrating complex mental health interventions in primary care. The current project aims to address this gap, potentially providing a template for future formative research in this area. Specifically, Link-me+EMPHN Mental Health Model of Care is an innovative approach to addressing this gap, strengthened by our comprehensive, theoretically grounded approach to its implementation.

Strengths and Limitations

A key strength of this protocol is its mixed methods approach, combining quantitative surveys with qualitative interviews and desktop audits. This triangulation of data sources will provide rich, contextual insights into the complexities of implementing new mental health care models in general practice. The inclusion of a range of stakeholder perspectives (GPs, practice staff, and mental health consumers) further enhances the comprehensiveness of the study, offering diverse viewpoints on implementation challenges and opportunities. Additionally, the focus on extending and tailoring evidence-based research to optimize adoption, implementation, and successful outcomes in practice addresses a critical gap between research and real-world application. However, potential limitations include the geographical constraint to the EMPHN catchment area, which may limit generalizability to other regions. Additionally, self-selection bias may occur if only practices highly motivated to improve mental health care participate in the study.

Future Directions

Following the formative research and the development of the GP supports, the Link-me+EMPHN Mental Health Model of Care will be implemented in up to 60 GP clinics in the EMPHN catchment area throughout 2025 to 2027. An evaluation (protocol in development) of the implementation will be conducted guided by the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) [83] and the PRISM (Practical, Robust Implementation, and Sustainability Model) [84] frameworks. The evaluation will systematically assess the effectiveness and sustainability of the mental health model of care and ensure that the model is comprehensive and scalable in real-world settings [83,84].

Conclusions

This formative research protocol outlines a comprehensive, mixed methods approach to inform the implementation of the Link-me+EMPHN Mental Health Model of Care in general practice settings. Through desktop audits, interviews, surveys, co-design workshops, and simulated practice testing, we aim to identify key gaps in GP mental health training, barriers to and enablers of implementation, and stakeholder perspectives. These methods are designed to yield practical insights that will directly inform the refinement of the model and its implementation strategy. Although the specific findings are yet to emerge from the data, we anticipate that this research will provide a nuanced understanding of the challenges and opportunities in integrating comprehensive mental health interventions in primary care. Therefore, this study contributes to the broader field of implementation science in mental health.

Acknowledgments

The authors would like to thank the expert working group who continue to lend their knowledge in support of the success of the Link-me+EMPHN Mental Health Model of Care. Thank you to Dr Steven Savvas, Dr Daniela Ramirez-Duran, and Ms Kisanet Tesfu for their assistance in developing the Theoretical Domains Framework survey and interviews. Thank you also to Prof Victoria Palmer for conducting the co-design sessions as part of this formative research. No artificial intelligence tools or technologies were used to prepare this protocol or formative research.

Funding

This formative research is being funded by the Eastern Melbourne Primary Health Network (EMPHN).

Data Availability

The data that will be generated as part of this research will not be publicly available to protect the privacy of research participants who will not be asked to provide consent for the public sharing of their data.

Authors' Contributions

All authors (AC, JL, AMM, AvR, KL, CJ, BB, and CK-H) contributed to the conceptualization and design of the study. AC drafted the manuscript, and all authors (AC, JL, AMM, AvR, KL, CJ, BB, and CK-H) revised all drafts.

XSL•FO
RenderX

Conflicts of Interest

CJ is a researcher involved in the Link-me+EMPHN Mental Health Model of Care. She also serves as an EMPHN Board Director and works as a general practitioner (GP) in the EMPHN catchment; however, she has no influence over the decisions related to the allocation of funding for the current project. A team from EMPHN (authors JL, AMM, KL, and AvR) and CJ have contributed to the development of the overall Link-me+EMPHN Mental Health Model of Care and will be involved in developing the GP supports resulting from the findings of the formative research. They have also been involved in the development of the online survey and interview questions and recruitment of participants for these elements of the formative research. These authors have also contributed to the preparation of this manuscript.

Multimedia Appendix 1
Online survey.
[DOCX File , 55 KB - resprot_v15i1e79560_app1.docx]

Multimedia Appendix 2
Theoretical domains framework.
[DOCX File , 22 KB - resprot_v15i1e79560_app2.docx]

Multimedia Appendix 3
Interview schedule and vignettes.
[DOCX File , 26 KB - resprot_v15i1e79560_app3.docx]

Multimedia Appendix 4
CCQR checklist.
[DOCX File , 21 KB - resprot_v15i1e79560_app4.docx]

References

1. GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Psychiatry 2022;9(2):137-150 [FREE Full text] [doi: 10.1016/S2215-0366(21)00395-3] [Medline: 35026139]
2. Fan Y, Fan A, Yang Z, Fan D. Global burden of mental disorders in 204 countries and territories, 1990-2021: results from the global burden of disease study 2021. BMC Psychiatry 2025;25(1):486 [FREE Full text] [doi: 10.1186/s12888-025-06932-y] [Medline: 40375174]
3. Prevalence and impact of mental illness. Australian Institute of Health and Welfare. 2025 May 20. URL: https://www.aihw.gov.au/mental-health/overview/prevalence-and-impact-of-mental-illness [accessed 2026-01-10]
4. Pizzol D, Trott M, Butler L, Barnett Y, Ford T, Neufeld SA, et al. Relationship between severe mental illness and physical multimorbidity: a meta-analysis and call for action. BMJ Ment Health 2023;26(1):e300870 [FREE Full text] [doi: 10.1136/bmjment-2023-300870] [Medline: 37907331]
5. National Health Survey: mental health and co-existing physical health conditions, Australia, 2014-15. Australian Bureau of Statistics. 2015. URL: https://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/4329.0.00.004Main%20Features22014%20-%2015 [accessed 2026-01-08]
6. Kwon S, Bhurawala H, Munoz A, Kramer J, Poulton A. General practitioners' attitudes and knowledge about attention-deficit hyperactivity disorder (ADHD): insights from a survey. Australas Psychiatry 2024;32(1):18-22 [FREE Full text] [doi: 10.1177/10398562231211156] [Medline: 37936255]
7. French B, Sayal K, Daley D. Barriers and facilitators to understanding of ADHD in primary care: a mixed-method systematic review. Eur Child Adolesc Psychiatry 2019;28(8):1037-1064 [FREE Full text] [doi: 10.1007/s00787-018-1256-3] [Medline: 30552584]
8. Katzman MA, Bilkey TS, Chokka PR, Fallu A, Klassen LJ. Adult ADHD and comorbid disorders: clinical implications of a dimensional approach. BMC Psychiatry 2017;17(1):302 [FREE Full text] [doi: 10.1186/s12888-017-1463-3] [Medline: 28830387]
9. French B, Nalbant G, Wright H, Sayal K, Daley D, Groom MJ, et al. The impacts associated with having ADHD: an umbrella review. Front Psychiatry 2024;15:1343314 [FREE Full text] [doi: 10.3389/fpsyt.2024.1343314] [Medline: 38840946]
10. Britt H, Miller G, Bayram C, Henderson J, Valenti L, Harrison C, et al. A Decade of Australian General Practice Activity 2006-07 to 2015-16. General Practice Series No. 41. Sydney: Sydney University Press; 2016.
11. Delivering mental health care in general practice: implications for practice and policy. General Practice Mental Health Standards Collaboration. 2021 Dec 15. URL: https://gpmhsc.org.au/getattachment/7be109a5-d32f-48c5-bf33-6c16c12e689b/delivering-mental-health-care-in-general-practice_implications-for-practice-and-policy.pdf [accessed 2026-01-08]

XSL•FO
RenderX

12. Parker D, Byng R, Dickens C, McCabe R. "Every structure we're taught goes out the window": general practitioners' experiences of providing help for patients with emotional concerns. Health Soc Care Community 2020;28(1):260-269 [FREE Full text] [doi: 10.1111/hsc.12860] [Medline: 31621140]
13. Woodley M. GPs say systemic Medicare issues impact provision of mental health care. newsGP. 2022 Jul 20. URL: https://www1.racgp.org.au/newsgp/professional/gps-say-systemic-medicare-issues-impact-provision [accessed 2026-01-08]
14. General Practice: Health of the Nation 2023. East Melbourne, Australia: Royal Australian College of General Practitioners; 2023.
15. Durbin A, Durbin J, Hensel JM, Deber R. Barriers and enablers to integrating mental health into primary care: a policy analysis. J Behav Health Serv Res 2016;43(1):127-139. [doi: 10.1007/s11414-013-9359-6] [Medline: 24026235]
16. Fleury MJ, Imboua A, Aubé D, Farand L, Lambert Y. General practitioners' management of mental disorders: a rewarding practice with considerable obstacles. BMC Fam Pract 2012;13(1):19 [FREE Full text] [doi: 10.1186/1471-2296-13-19] [Medline: 22423592]
17. Gask L, Evans M, Kessler D. Personality disorder. BMJ 2013;347:f5276. [doi: 10.1136/bmj.f5276] [Medline: 24022034]
18. Nørøxe KB, Pedersen AF, Bro F, Vedsted P. Mental well-being and job satisfaction among general practitioners: a nationwide cross-sectional survey in Denmark. BMC Fam Pract 2018;19(1):130 [FREE Full text] [doi: 10.1186/s12875-018-0809-3] [Medline: 30055571]
19. Younes N, Gasquet I, Gaudebout P, Chaillet M, Kovess V, Falissard B, et al. General practitioners' opinions on their practice in mental health and their collaboration with mental health professionals. BMC Fam Pract 2005;6(1):18 [FREE Full text] [doi: 10.1186/1471-2296-6-18] [Medline: 15865624]
20. Katon W, Von Korff M, Lin E, Simon G, Walker E, Bush T, et al. Collaborative management to achieve depression treatment guidelines. J Clin Psychiatry 1997;58(1):20-23. [Medline: 9054905]
21. Gunn J, Diggens J, Hegarty K, Blashki G. A systematic review of complex system interventions designed to increase recovery from depression in primary care. BMC Health Serv Res 2006;6(1):88 [FREE Full text] [doi: 10.1186/1472-6963-6-88] [Medline: 16842629]
22. Archer J, Bower P, Gilbody S, Lovell K, Richards D, Gask L, et al. Collaborative care for depression and anxiety problems. Cochrane Database Syst Rev 2012;10(10):CD006525. [doi: 10.1002/14651858.CD006525.pub2] [Medline: 23076925]
23. Whitfield J, Owens S, Bhat A, Felker B, Jewell T, Chwastiak L. Successful ingredients of effective collaborative care programs in low- and middle-income countries: a rapid review. Glob Ment Health (Camb) 2023;10:e11 [FREE Full text] [doi: 10.1017/gmh.2022.60] [Medline: 37854388]
24. Kroenke K, Unutzer J. Closing the false divide: sustainable approaches to integrating mental health services into primary care. J Gen Intern Med 2017;32(4):404-410 [FREE Full text] [doi: 10.1007/s11606-016-3967-9] [Medline: 28243873]
25. Sanchez K. Collaborative care in real-world settings: barriers and opportunities for sustainability. Patient Prefer Adherence 2017;11:71-74. [doi: 10.2147/ppa.s120070]
26. Wolk CB, Last BS, Livesey C, Oquendo MA, Press MJ, Mandell DS, et al. Addressing common challenges in the implementation of collaborative care for mental health: the Penn Integrated Care Program. Ann Fam Med 2021;19(2):148-156 [FREE Full text] [doi: 10.1370/afm.2651] [Medline: 33685876]
27. Overbeck G, Davidsen AS, Kousgaard MB. Enablers and barriers to implementing collaborative care for anxiety and depression: a systematic qualitative review. Implement Sci 2016;11(1):165 [FREE Full text] [doi: 10.1186/s13012-016-0519-y] [Medline: 28031028]
28. Ghosh P. Improving access to psychological therapies for all adults. Psychiatr Bull 2018;33(5):186-188. [doi: 10.1192/pb.bp.108.021972]
29. Clark DM, Layard R, Smithies R, Richards DA, Suckling R, Wright B. Improving access to psychological therapy: initial evaluation of two UK demonstration sites. Behav Res Ther 2009;47(11):910-920 [FREE Full text] [doi: 10.1016/j.brat.2009.07.010] [Medline: 19647230]
30. Martin C, Iqbal Z, Airey ND, Marks L. Improving access to psychological therapies (IAPT) has potential but is not sufficient: how can it better meet the range of primary care mental health needs? Br J Clin Psychol 2022;61(1):157-174. [doi: 10.1111/bjc.12314] [Medline: 34124792]
31. Curtin EL, d'Apice K, Porter A, Widnall E, Franklin M, de Vocht F, et al. Perspectives on an enhanced 'Improving Access to Psychological Therapies' (IAPT) service addressing the wider determinants of mental health: a qualitative study. BMC Health Serv Res 2023;23(1):536 [FREE Full text] [doi: 10.1186/s12913-023-09405-8] [Medline: 37226155]
32. Ruud T, Rugkåsa J, Haavet OR, Dahli MP, Hanssen-Bauer K, Brekke M, et al. A collaborative primary and mental health care model with psychologist and psychiatrist working in GP practices: process evaluation of the implementation, challenges, and sustainability. BMC Health Serv Res 2025;25(1):1178 [FREE Full text] [doi: 10.1186/s12913-025-13408-y] [Medline: 40898247]
33. PHN primary mental health care flexible funding pool implementation guidance: stepped care. Australian Government Department of Health and Ageing. 2016. URL: http://www.health.gov.au [accessed 2026-01-08]
34. Bassilios B, Nicholas A, Reifels L, King K, Fletcher J, Machlin A, et al. Achievements of the Australian access to allied psychological services (ATAPS) program: summarising (almost) a decade of key evaluation data. Int J Ment Health Syst 2016;10(1):61 [FREE Full text] [doi: 10.1186/s13033-016-0092-4] [Medline: 27708698]

XSL•FO
RenderX

35. Operational Guidelines for the Access to Allied Psychological Services Initiative. Canberra: Australian Government Department of Health and Ageing; 2012.
36. Bassilios B, Pirkis J, Fletcher J, Burgess P, Gurrin L, King K, et al. The complementarity of two major Australian primary mental health care initiatives. Aust N Z J Psychiatry 2010;44(11):997-1004. [doi: 10.3109/00048674.2010.495052] [Medline: 21034182]
37. Collaborative care and place-based planning approaches: a scalability assessment for the Regional Health Division of the New South Wales Ministry of Health. Sax Institute. 2023 Dec. URL: https://www.saxinstitute.org.au/wp-content/uploads/collaborative-care-and-place-based-planning-approaches.pdf [accessed 2026-01-08]
38. Rosenberg S, Iorfino F, Hickie I. Redesigning primary mental health care for enhanced access, equity and quality. University of Sydney. 2023 Aug. URL: https://www.sydney.edu.au/content/dam/corporate/documents/brain-and-mind-centre/youth-mental-health/redesigning-primary-mental-health-care.pdf [accessed 2026-01-08]
39. Lam TP, Sun KS, Piterman L, Lam KF, Poon MK, See C, et al. Impact of training for general practitioners on their mental health services: the Hong Kong experience. Aust J Gen Pract 2018;47(8):550-555 [FREE Full text] [doi: 10.31128/ajgp-01-18-4478]
40. Gilbody S, Whitty P, Grimshaw J, Thomas R. Educational and organizational interventions to improve the management of depression in primary care: a systematic review. JAMA 2003;289(23):3145-3151. [doi: 10.1001/jama.289.23.3145] [Medline: 12813120]
41. Papachristopoulos I, Sazakli E, Leotsinidis M. General practitioners' views towards management of common mental health disorders: the critical role of continuing medical education. BMC Prim Care 2023;24(1):63 [FREE Full text] [doi: 10.1186/s12875-023-02017-5] [Medline: 36870977]
42. Mental health training standards 2026-28: a guide for general practitioners. Royal Australian College of General Practitioners. 2025. URL: https://gpmhsc.org.au/guidelinessection/index/86939954-1d7c-4442-9b74-25726081f7f7/mental-health-training-standards-2023-25 [accessed 2025-12-01]
43. Naismith SL, Hickie IB, Scott EM, Davenport TA. Effects of mental health training and clinical audit on general practitioners' management of common mental disorders. Med J Aust 2001;175(S1):S42-S47. [doi: 10.5694/j.1326-5377.2001.tb143789.x] [Medline: 11556436]
44. Richards JC, Ryan P, McCabe MP, Groom G, Hickie IB. Barriers to the effective management of depression in general practice. Aust N Z J Psychiatry 2004;38(10):795-803. [doi: 10.1080/j.1440-1614.2004.01464.x] [Medline: 15369538]
45. England E, Nash V, Hawthorne K. GP training in mental health needs urgent reform. BMJ 2017;356:j1311. [doi: 10.1136/bmj.j1311] [Medline: 28302682]
46. Pirkis J, Ftanou M, Williamson M, Machlin A, Spittal MJ, Bassilios B, et al. Australia's better access initiative: an evaluation. Aust N Z J Psychiatry 2011;45(9):726-739. [doi: 10.3109/00048674.2011.594948] [Medline: 21888609]
47. Mental health (2020). Productivity Commission. 2020 Jun 30. URL: https://assets.pc.gov.au/inquiries/completed/mental-health/report/mental-health-volume1.pdf [accessed 2026-01-08]
48. Browne MO, Lee A, Prabhu R. Self-reported confidence and skills of general practitioners in management of mental health disorders. Aust J Rural Health 2007;15(5):321-326. [doi: 10.1111/j.1440-1584.2007.00914.x] [Medline: 17760916]
49. Fletcher S, Chondros P, Densley K, Murray E, Dowrick C, Coe A, et al. Matching depression management to severity prognosis in primary care: results of the Target-D randomised controlled trial. Br J Gen Pract 2021;71(703):e85-e94 [FREE Full text] [doi: 10.3399/BJGP.2020.0783] [Medline: 33431380]
50. Fletcher S, Spittal MJ, Chondros P, Palmer V, Chatterton M, Densley K, et al. Clinical efficacy of a decision support tool (Link-me) to guide intensity of mental health care in primary practice: a pragmatic stratified randomised controlled trial. Lancet Psychiatry 2021;8(3):202-214 [FREE Full text] [doi: 10.1016/s2215-0366(20)30517-4]
51. SupportConnect. Eastern Melbourne PHN. URL: https://supportconnect.org.au/ [accessed 2026-01-12]
52. Chondros P, Davidson S, Wolfe R, Gilchrist G, Dowrick C, Griffiths F, et al. Development of a prognostic model for predicting depression severity in adult primary patients with depressive symptoms using the diamond longitudinal study. J Affect Disord 2018;227:854-860 [FREE Full text] [doi: 10.1016/j.jad.2017.11.042] [Medline: 29689701]
53. Fletcher S, Chondros P, Palmer VJ, Chatterton ML, Spittal MJ, Mihalopoulos C, et al. Link-me: protocol for a randomised controlled trial of a systematic approach to stepped mental health care in primary care. Contemp Clin Trials 2019;78:63-75. [doi: 10.1016/j.cct.2018.12.014] [Medline: 30593884]
54. Wachtler C, Coe A, Davidson S, Fletcher S, Mendoza A, Sterling L, et al. Development of a mobile clinical prediction tool to estimate future depression severity and guide treatment in primary care: user-centered design. JMIR Mhealth Uhealth 2018;6(4):e95 [FREE Full text] [doi: 10.2196/mhealth.9502] [Medline: 29685864]
55. Fu X, Wu W, Wu Y, Liu X, Liang W, Wu R, et al. Adult ADHD and comorbid anxiety and depressive disorders: a review of etiology and treatment. Front Psychiatry 2025;16:1597559. [doi: 10.3389/fpsyt.2025.1597559] [Medline: 40547117]
56. Cheng S, Chen Z, Wu C, Li X, Shen X, Qiu R, et al. The comorbidity of anxiety and depression symptoms in obsessive-compulsive disorder: a network analysis. Front Psychiatry 2025;16:1567448. [doi: 10.3389/fpsyt.2025.1567448] [Medline: 40386117]

XSL•FO
RenderX

57. Hambleton A, Pepin G, Le A, Maloney D, Touyz S, Maguire S. Psychiatric and medical comorbidities of eating disorders: findings from a rapid review of the literature. J Eat Disord 2022;10(1):132 [FREE Full text] [doi: 10.1186/s40337-022-00654-2] [Medline: 36064606]
58. About the journal: JMIR Formative Research. JMIR Publications. URL: https://formative.jmir.org/about-journal/focus-and-scope [accessed 2025-01-12]
59. Chen HT. Assisting stakeholders as they formulate program rationales. In: Program Evaluation to Help Stakeholders Plan Intervention Programs. Thousand Oaks: SAGE Publications; 2011.
60. Reigeluth C, Frick T. Formative research: a methodology for creating and improving design theories. In: Reigeluth C, editor. Instructional-Design Theories and Models: A New Paradigm of Instructional Theory. New Jersey: Lawrence Erlbaum Associates Publishers; 1999.
61. Michie S. The Behaviour Change Wheel: A Guide to Designing Interventions. Sutton, UK: Silverback Publishing; 2014.
62. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the theoretical domains framework of behaviour change to investigate implementation problems. Implement Sci 2017;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
63. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
64. Sinha P, Paudel B, Mosimann T, Ahmed H, Kovane GP, Moagi M, et al. Comprehensive Criteria for Reporting Qualitative Research (CCQR): reporting guideline for global health qualitative research methods. Int J Environ Res Public Health 2024;21(8):1005 [FREE Full text] [doi: 10.3390/ijerph21081005] [Medline: 39200615]
65. Community snapshot. Eastern Melbourne PHN. URL: https://emphn.org.au/about-us/community-snapshot/ [accessed 2025-01-12]
66. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208 [FREE Full text] [doi: 10.1016/j.jbi.2019.103208] [Medline: 31078660]
67. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research Electronic Data Capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
68. McGarry H, Hegarty K, Johnson C, Gunn J, Blashki G. Managing depression in a changing primary mental healthcare system: comparison of two snapshots of Australian GPs' treatment and referral patterns. Ment Health Fam Med 2009;6(2):75-83 [FREE Full text] [Medline: 22477895]
69. Dunt D, Souther D, Elsworth G, Harris C, Potiriadis M, Young D. Mapping GP integration in the health care system in difference locations in Australia. Aust J Prim Health 2006;12(1):47-55. [doi: 10.1071/py06008]
70. Botega N, Mann A, Blizard R, Wilkinson G. General practitioners and depression—first use of the depression attitude questionnaire. Int J Methods Psychiatr Res 1992;2:16-180. [doi: 10.1037/t08271-000]
71. Palmer VJ, Bibb J, Lewis M, Densley K, Kritharidis R, Dettmann E, et al. A co-design living labs philosophy of practice for end-to-end research design to translation with people with lived-experience of mental ill-health and carer/family and kinship groups. Front Public Health 2023;11:1206620 [FREE Full text] [doi: 10.3389/fpubh.2023.1206620] [Medline: 38115850]
72. The Validitron: making sure digital health actually works for patients and professionals. Centre for Digital Transformation of Health. URL: https://mdhs.unimelb.edu.au/digitalhealth/our-work/validitron [accessed 2026-01-08]
73. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2008;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
74. Creswell J, Plano Clark V. Designing and Conducting Mixed Methods Research. 3rd ed. Thousand Oaks, CA: SAGE Publications; 2018.
75. Lincoln YS, Guba EG. Naturalistic Inquiry. Thousand Oaks, CA: SAGE Publications; 1985.
76. Wood E, Ohlsen S, Ricketts T. What are the barriers and facilitators to implementing collaborative care for depression? A systematic review. J Affect Disord 2017;214:26-43 [FREE Full text] [doi: 10.1016/j.jad.2017.02.028] [Medline: 28266319]
77. Isaacs AN, Mitchell EKL. Mental health integrated care models in primary care and factors that contribute to their effective implementation: a scoping review. Int J Ment Health Syst 2024;18(1):5 [FREE Full text] [doi: 10.1186/s13033-024-00625-x] [Medline: 38331913]
78. Whitebird RR, Solberg LI, Jaeckels NA, Pietruszewski PB, Hadzic S, Unützer J, et al. Effective Implementation of collaborative care for depression: what is needed? Am J Manag Care 2014;20(9):699-707 [FREE Full text] [Medline: 25365745]
79. Snyder SE, Meza J, Liang V, Kennedy J, Wolk CB. Using formative evaluation strategies to support implementation and improvement of suicide prevention practices in a community mental health clinic. Glob Implement Res Appl 2025. [doi: 10.1007/s43477-025-00194-2]
80. Daniel M, Kaur A, Mukherjee A, Bhattacharya A, Tewari A, Sagar R, et al. The systematic medical appraisal, referral and treatment (SMART) mental health programme: formative research informing a cluster randomized controlled trial. SSM Ment Health 2023;3:100223. [doi: 10.1016/j.ssmmh.2023.100223]

XSL•FO
RenderX

81. Overbeck G, Kousgaard MB, Davidsen AS. The work and challenges of care managers in the implementation of collaborative care: a qualitative study. J Psychiatr Ment Health Nurs 2018;25(3):167-175. [doi: 10.1111/jpm.12449] [Medline: 29283474]
82. Moullin JC, Dickson KS, Stadnick NA, Rabin B, Aarons GA. Systematic review of the exploration, preparation, implementation, sustainment (EPIS) framework. Implement Sci 2019;14(1):1 [FREE Full text] [doi: 10.1186/s13012-018-0842-6] [Medline: 30611302]
83. Glasgow RE, Harden SM, Gaglio B, Rabin B, Smith ML, Porter GC, et al. RE-AIM planning and evaluation framework: adapting to new science and practice with a 20-year review. Front Public Health 2019;7:64 [FREE Full text] [doi: 10.3389/fpubh.2019.00064] [Medline: 30984733]
84. Feldstein AC, Glasgow RE. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. Jt Comm J Qual Patient Saf 2008;34(4):228-243. [doi: 10.1016/s1553-7250(08)34030-6] [Medline: 18468362]

Abbreviations

ACRRM: Australian College of Rural and Remote Medicine
ADHD: attention-deficit/hyperactivity disorder
CCQR: Comprehensive Criteria for Reporting Qualitative Research
COM-B: Capability, Opportunity, Motivation–Behaviour
EDSS: electronic decision support system
EMPHN: Eastern Melbourne Primary Health Network
GP: general practitioner
IAPT: Improving Access to Psychological Therapies
MHPOD: Mental Health Professional Online Development
PDSA: Plan-Do-Study-Act
PHN: Primary Health Network
PRISM: Practical, Robust Implementation, and Sustainability Model
RACGP: Royal Australian College of General Practitioners
RCT: randomized controlled trial
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
REDCap: Research Electronic Data Capture
SMART: Systematic Medical Appraisal, Referral, and Treatment
TDF: Theoretical Domains Framework

Edited by A Schwartz; submitted 24.Jun.2025; peer-reviewed by T Thilsing; comments to author 21.Oct.2025; revised version received 09.Dec.2025; accepted 30.Dec.2025; published 23.Jan.2026.

Please cite as:
Coe A, London J, Martin AM, van Ree A, Lembke K, Johnson C, Bassilios B, Kaylor-Hughes C
Developing the Link-me+EMPHN Mental Health Model of Care to Improve General Practitioner Capacity for Mental Health Care in Australian Primary Care: Protocol for a Mixed Methods Formative Study
JMIR Res Protoc 2026;15:e79560
URL: https://www.researchprotocols.org/2026/1/e79560
doi:10.2196/79560
PMID:

©Amy Coe, Jane London, Anne-Marie Martin, Aaron van Ree, Kirsty Lembke, Caroline Johnson, Bridget Bassilios, Catherine Kaylor-Hughes. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 23.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Bridging Neurofeedback and Structural Connectivity in Chronic Pain

Alaeddin Acar[1], MD; Diaa Yahya[2], MD; Eray Tekirdaş[3], MD

[1]Department of Neurosurgery, Kulu State Hospital, No 4, 139518 Street, Dinek, Kulu, Konya, Turkey

[2]Department of Neurosurgery, Bucak State Hospital, Burdur, Turkey

[3]Department of Neurosurgery, Malatya Training and Research Hospital, Malatya, Turkey

Corresponding Author:
Alaeddin Acar, MD
Department of Neurosurgery, Kulu State Hospital, No 4, 139518 Street, Dinek, Kulu, Konya, Turkey

Related Articles:

https://www.researchprotocols.org/2025/1/e78806

Comment in: https://www.researchprotocols.org/2026/1/e89007

(JMIR Res Protoc 2026;15:e87420) doi:10.2196/87420

KEYWORDS

brain-computer interface; brain oscillations; chronic pain; electroencephalography; musculoskeletal; neuromodulation

We read with great interest and appreciation the article by Bialostocki et al [1] published in *JMIR Research Protocols*. The authors present an innovative electroencephalography-based neurofeedback (EEG-NF) protocol targeting abnormal activity in the right insula (RIns) and dorsal anterior cingulate cortex (dACC) [1].

The authors' goal of exploring the functional connectivity between these two key regions is particularly noteworthy. This approach aims to modulate the "salience network," which is involved in processing the emotional-motivational aspects of pain and includes the RIns and dACC as key nodes [2].

At this point, we would like to offer a constructive contribution by approaching the protocol's underlying mechanism from a neurosurgical perspective. This functional circuit involving the RIns and dACC is thought to operate, at least in part, on a structural foundation provided by the cingulum bundle [2].

This connection is important because invasive modulation of this circuit is a long-established neurosurgical procedure. Anterior cingulotomy is used as a "last resort" option for treatment-resistant, intractable chronic pain syndromes, especially when the affective component of pain is dominant [3]. This procedure creates ablative lesions in the cingulum bundle, permanently interrupting the circuit that Bialostocki et al [1] aim to noninvasively "downtrain." It is well documented that anterior cingulotomy reduces the emotional response to pain and the component of suffering, rather than the pain intensity itself [3], which is a close parallel to the known functions of the RIns and dACC.

The EEG-NF protocol presented by Bialostocki et al [1] may therefore be viewed, at least conceptually, as a noninvasive counterpart to this invasive surgical approach. Given this parallel, the inclusion of structural imaging methods, such as diffusion tensor imaging (DTI), could add great value to the protocol. The literature, including a systematic review of DTI in migraine, a chronic pain disorder, has shown that the microstructural integrity of key white matter tracts (eg, reduced fractional anisotropy in thalamic radiations, corpus callosum, cingulum, and association fibers) is frequently altered [4].

This raises two important questions for future studies:

1. Could the pretreatment structural integrity of the cingulum bundle, as measured by DTI, be used as a biomarker to predict a patient's response to this EEG-NF therapy?
2. Can a successful neurofeedback intervention lead to measurable neuroplastic changes in the cingulum bundle, as measured by DTI? Indeed, a strong relationship between functional and structural connectivity has been reported [5].

In conclusion, we believe that adding structural DTI analysis to the innovative protocol by Bialostocki et al [1] may deepen the understanding of the targeted mechanisms, help bridge the gap between noninvasive neuromodulation and invasive neurosurgery, and aid in optimizing patient selection for larger clinical trials in the future.

XSL•FO
RenderX

Acknowledgments

Google Gemini was used for language editing.

Conflicts of Interest

None declared.

References

1. Bialostocki LS, Adhia DB, Mudiyanselage DR, et al. Neurofeedback training for managing neuropathic pain-like features in chronic musculoskeletal pain: protocol for an open-label pilot feasibility clinical trial. JMIR Res Protoc 2025 Nov 4;14(1):e78806. [doi: 10.2196/78806] [Medline: 41187327]
2. Seeley WW, Menon V, Schatzberg AF, et al. Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci 2007 Feb 28;27(9):2349-2356. [doi: 10.1523/JNEUROSCI.5587-06.2007] [Medline: 17329432]
3. Sharim J, Pouratian N. Anterior cingulotomy for the treatment of chronic intractable pain: a systematic review. Pain Physician 2016;19(8):537-550. [Medline: 27906933]
4. Rahimi R, Dolatshahi M, Abbasi-Feijani F, et al. Microstructural white matter alterations associated with migraine headaches: a systematic review of diffusion tensor imaging studies. Brain Imaging Behav 2022 Oct;16(5):2375-2401. [doi: 10.1007/s11682-022-00690-1] [Medline: 35710680]
5. Greicius MD, Supekar K, Menon V, Dougherty RF. Resting-state functional connectivity reflects structural connectivity in the default mode network. Cereb Cortex 2009 Jan;19(1):72-78. [doi: 10.1093/cercor/bhn059] [Medline: 18403396]

Abbreviations

dACC: dorsal anterior cingulate cortex
DTI: diffusion tensor imaging
EEG-NF: electroencephalography neurofeedback
RIns: right insula

Edited by A Schwartz; submitted 08.Nov.2025; this is a non–peer-reviewed article;accepted 08.Dec.2025; published 28.Jan.2026.

Please cite as:
Acar A, Yahya D, Tekirdaş E
Bridging Neurofeedback and Structural Connectivity in Chronic Pain
JMIR Res Protoc 2026;15:e87420
URL: https://www.researchprotocols.org/2026/1/e87420
doi:10.2196/87420

© Alaeddin Acar, Diaa Yahya, Eray Tekirdaş. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 28.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Authors' Reply: Bridging Neurofeedback and Structural Connectivity in Chronic Pain

Luke Spencer Bialostocki[1,2]; Divya Bharatkumar Adhia[2], PhD; Damith Rathnayake Mudiyanselage[1], MBBS; Mark Llewellyn Smith[3], LCSW; Yusuf Ozgur Cakmak[1], MD, PhD; Dirk De Ridder[2], MD, PhD; Ramakrishnan Mani[4], PhD; Jerin Mathew[1], PhD

[1]Department of Anatomy, Faculty of Biomedical Sciences, University of Otago, 270 Great King Street, Dunedin, Otago, New Zealand

[2]Department of Surgical Sciences, Faculty of Medicine, University of Otago, Dunedin, Otago, New Zealand

[3]Neurofeedback Services of New York, New York, NY, United States

[4]Centre for Health, Activity, and Rehabilitation Research, School of Physiotherapy, University of Otago, Dunedin, Otago, New Zealand

Corresponding Author:
Jerin Mathew, PhD
Department of Anatomy, Faculty of Biomedical Sciences, University of Otago, 270 Great King Street, Dunedin, Otago, New Zealand

Related Articles:

https://www.researchprotocols.org/2025/1/e78806

https://www.researchprotocols.org/2026/1/e87420

(JMIR Res Protoc 2026;15:e89007) doi:10.2196/89007

KEYWORDS

brain-computer interface; brain oscillations; chronic pain; electroencephalography; musculoskeletal; neuromodulation

We appreciate the opportunity to respond to the letter [1] written in response to our recent publication "Neurofeedback Training for Managing Neuropathic Pain–Like Features in Chronic Musculoskeletal Pain: Protocol for an Open-Label Pilot Feasibility Clinical Trial" [2]. The paper outlines a protocol for a novel electroencephalography-based neurofeedback approach, which aims to simultaneously downregulate infraslow activity in the right insula and the dorsal anterior cingulate cortex.

We thank the authors for their interest in our publication and for suggesting the incorporation of diffusion tensor imaging (DTI) to assess neuroplastic alterations in the cingulum bundle. We agree that DTI could provide valuable complementary information on the structural substrates underlying functional changes induced by electroencephalography-based neurofeedback. While we recognize the merit of this proposal, particularly the possibility that structural differences could influence training outcomes, we are unfortunately unable to incorporate DTI into the present study. Access to DTI is not available within the scope or resources of this project, as medical imaging technologies such as DTI are costly to implement and limited to specialized centers with the necessary equipment and trained personnel. In addition, our study has now completed its recruitment and data collection phase. For these reasons, integrating DTI into this particular study is not feasible.

Structural imaging techniques such as DTI may indeed be incorporated into future studies to further elucidate the neurophysiological changes associated with endogenous neuromodulation. These suggestions underscore a promising direction for integrating neuroanatomical perspectives with neuromodulatory approaches, offering a more complete understanding of the anatomical processes involved. Integrating DTI also represents a promising direction for future research aiming to link functional neuromodulation with white matter pathways [3]. However, it is also important to note that structural and functional connectivity often show weak correlations [4].

Conflicts of Interest

MLS is the owner of Neurofeedback Services of New York.

References

1. Acar A, Yahya D, Tekirdaş E. Bridging neurofeedback and structural connectivity in chronic pain. JMIR Res Protoc 2026;15:e87420. [doi: 10.2196/87420]

XSL•FO
RenderX

2. Bialostocki LS, Adhia DB, Mudiyanselage DR, et al. Neurofeedback training for managing neuropathic pain–like features in chronic musculoskeletal pain: protocol for an open-label pilot feasibility clinical trial. JMIR Res Protoc 2025 Nov 4;14:e78806. [doi: 10.2196/78806] [Medline: 41187327]
3. Silva NA, Barrios-Martinez J, Yeh FC, et al. Diffusion and functional MRI in surgical neuromodulation. Neurotherapeutics 2024 Apr;21(3):e00364. [doi: 10.1016/j.neurot.2024.e00364] [Medline: 38669936]
4. Huang H, Ding M. Linking functional connectivity and structural connectivity quantitatively: a comparison of methods. Brain Connect 2016 Mar;6(2):99-108. [doi: 10.1089/brain.2015.0382] [Medline: 26598788]

Abbreviations

DTI: diffusion tensor imaging

Edited by A Schwartz; submitted 04.Dec.2025; this is a non–peer-reviewed article;accepted 08.Dec.2025; published 28.Jan.2026.

Please cite as:
Bialostocki LS, Adhia DB, Mudiyanselage DR, Smith ML, Cakmak YO, De Ridder D, Mani R, Mathew J
Authors' Reply: Bridging Neurofeedback and Structural Connectivity in Chronic Pain
JMIR Res Protoc 2026;15:e89007
URL: https://www.researchprotocols.org/2026/1/e89007
doi:10.2196/89007

© Luke Spencer Bialostocki, Divya Bharatkumar Adhia, Damith Rathnayake Mudiyanselage, Mark Llewellyn Smith, Yusuf Ozgur Cakmak, Dirk De Ridder, Ramakrishnan Mani, Jerin Mathew. Originally published in JMIR Research Protocols (https://www.researchprotocols.org), 28.Jan.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on https://www.researchprotocols.org, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX